JMIR Human Factors

Impact Factor (2023): 2.6
Volume 12 (2025) ISSN 2292-9495 Editor in Chief: Andre Kushniruk, BA, MSc, PhD, FACMI

Contents

Original Papers

Patient-Centered Televisit for Chronic Obstructive Pulmonary Disease Discharge Transitions: User-Centered Design Study (e77953)
Joanna Abraham, Alicia Meng, Nicolas Caravelli, Leah Traeger, May Nguyen, Vineet Arora, Valerie Press. 23

A Messaging App Empowering Lifestyle Modification in Chronic Kidney Disease (LINE Official Account “Kidney Lifestyle”): Platform Development and Usability Study (e73935)
Chun-Yi Ho, Deborah Siregar, Miaofen Yen, Junne-Ming Sung, Ming-Cheng Wang, Wei-Hung Lin. 53

Impact of Digital Phenotypes and Question-Asking on Emotional Disorders in Adolescents: 4-Week Field Study (e66536)
Minseo Cho, Doeun Park, Myounglee Choo, Doug Han, Jinwoo Kim. 80

Applying Human-Centered Design to Develop Smartphone-Based Intervention Messages to Help Young Adults Quit Using E-Cigarettes and Cigarettes: A Remote User Testing Study (e76503)
Thi Tran, Christine Tran, Pamela Ling, Lucy Popova, Nhung Nguyen. 97

User Requirements and Conceptual Design for an Electronic Data Platform for Interhospital Transfer Between Acute Care Hospitals: User-Centered Design Study (e67884)
Pamela Garabedian, Jazzarae Kain, Srinivas Emani, Stephanie Singleton, Ronen Rozenblum, Lipika Samal, Stephanie Mueller. 109

Growth Mindset Intervention's Impact on Positive Response to eHealth for Older Adults With Chronic Disease: Randomized Controlled Trial (e65519)
Meijuan Cao, Xiaojuan Xu, Yaling Zeng, Binyu Zhao, Chunqi Xie, Hailu Wu, Jianlin Lou. 121

Digital Informed Consent/Assent in Clinical Trials Among Pregnant Women, Minors, and Adults: Multicountry Cross-Sectional Evaluation of Comprehension and Satisfaction (e65569)
Jaime Fons-Martinez, Carlos Murciano-Gamborino, Cristina Ferrer-Albero, Carlos Vergara-Hernandez, Javier Diez-Domingo. 135

Improving Access to and Delivery of Maternal Health Care Services to Prevent Postpartum Hemorrhage in Selected States in Nigeria: Human-Centered Design Study (e58577)
Bosun Tijani, Uchenna Igbokwe, Temi Filani, Adefemi Adewemimo, Lola Ameyan, Martins Iyekekpolor, Steven Karera, Olatunji Oluyide, Emmanuela Ezike, Temidayo Akinreni, Obruche Ogefere, Victor Adetimilehin, Valentine Amasiatu, Chukwunonso Nwaokorie, Naanma Kangkum, Olufunke Fasawe, Eric Aigbogun Jr. 151

A Just-in-Time Adaptive Intervention (Shift) to Manage Problem Anger After Trauma: Co-Design and Development Study (e62960)
Olivia Metcalf, David Forbes, Lauren Henry, Tianchen Qian, Tracey Varker, Melissa Brotman, Sean Cowlishaw, Karen Lamb, Meaghan O'Donnell. 163

XSL•FO
RenderX

A Health Professional Mentorship Platform to Improve Equitable Access to Abortion: Development, Usability, and Content Evaluation (e63364)
Abdul-Fatawu Abdulai, Cam Duong, Eleni Stroulia, Efrat Czerniak, Rachel Chiu, Aashay Mehta, Ken Koike, Wendy Norman. 175

Developing an App for Real-Time Daily Life Observations in a Nursing Home Setting: Qualitative User-Centered Co-Design Approach (e57911)
Coen Hacking, Bram de Boer, Hilde Verbeek, Jan Hamers, Sil Aarts. 188

Oncology Clinicians' Perspectives of a Remote Patient Monitoring Program: Multi-Modal Case Study Approach (e60585)
Ann Mazzella-Ebstein, Robert Daly, Jennie Huang, Camila Bernal, Clare Wilhelm, Katherine Panageas, Jessie Holland, Rori Salvaggio, Jill Ackerman, Jennifer Cracchiolo, Gilad Kuperman, Jun Mao, Aaron Begue, Margaret Barton-Burke. 197

Evaluating the Safety and Usability of an Over-the-Counter Medical Device for Adults With Mild to Moderate Hearing Loss: Formative and Summative Usability Testing (e65142)
Megan Salwei, Shilo Anders, Carrie Reale, Jason Slagle, Todd Ricketts, Matthew Weinger. 211

Predictive Factors and the Predictive Scoring System for Falls in Acute Care Inpatients: Retrospective Cohort Study (e58073)
Chihiro Saito, Eiji Nakatani, Hatoko Sasaki, Naoko E Katsuki, Masaki Tago, Kiyoshi Harada. 228

An Explainable AI Application (AF'fective) to Support Monitoring of Patients With Atrial Fibrillation After Catheter Ablation: Qualitative Focus Group, Design Session, and Interview Study (e65923)
Wan She, Panote Siriaraya, Hibiki Iwakoshi, Noriaki Kuwahara, Keitaro Senoo. 244

Preadolescent Children Using Real-Time Heart Rate During Moderate to Vigorous Physical Activity: A Feasibility Study (e58715)
Lincoln Lu, Danielle Jake-Schoffman, Hannah Lavoie, Maedeh Agharazidermani, Kristy Boyer. 260

User-Centered Design of an Electronic Dashboard for Monitoring Facility-Level Basic Emergency Obstetric Care Readiness in Amhara, Ethiopia: Mixed Methods Study (e64131)
Kylie Dougherty, Yihenew Tesfaye, Heran Biza, Mulusew Belew, Natalie Benda, Abebe Gebremariam Gobezayehu, John Cranmer, Suzanne Bakken. 276

Co-Designing a Web-Based and Tablet App to Evaluate Clinical Outcomes of Early Psychosis Service Users in a Learning Health Care Network: User-Centered Design Workshop and Pilot Study (e65889)
Kathleen Burch, Valerie Tryon, Katherine Pierce, Laura Tully, Sabrina Ereshefsky, Mark Savill, Leigh Smith, Adam Wilcox, Christopher Hakusui, Viviana Padilla, Amanda McNamara, Merissa Kado-Walton, Andrew Padovani, Chelyah Miller, Madison Miles, Nitasha Sharma, Khanh Nguyen, Yi Zhang, Tara Niendam. 291

Exploring Mental Health Content Moderation and Well-Being Tools on Social Media Platforms: Walkthrough Analysis (e69817)
Zoë Haime, Lucy Biddle. 313

Examining Patient Characteristics Associated With Digital Outpatient Care for Type 1 Diabetes (DigiDiaS): Cross-Sectional Study (e73708)
Ingeborg Spildo, Heidi Holmen, Annesofie Lunde Jensen, Milada Hagen, Tone Singstad, Jacob Winther, Eirik Årsand, Astrid Torbjørnsen. . .
3 3 0

Gamified Physical-Digital Smoking Cessation Intervention for Young Adults: Mixed Methods Development and Usability Study (e72749)
Natalia Bartłomiejczyk, Alessio De Santo, Adrian Holzer. 350

XSL•FO
RenderX

Expanding the Design Space for Fall Prevention in Acute Orthopedic Hospital Care: Human-Centered Design Study (e73110)
Maria Ehn, Åsa Revenäs, Helena Tobiasson. . . . 370

Redesigning Telemedicine for Traditional Chinese Medicine: Service Design Approach to Digital Transformation (e76752)
Arisara Jiamsanguanwong, Romanee Luo, Ratchapoom Kaingam, Oran Kittithreerapronchai, Waratta Authayarat. . . . 648

An AI-Based Clinical Decision Support System for Antibiotic Therapy in Sepsis (KINBIOTICS): Use Case Analysis (e66699)
Juliane Düvel, David Lampe, Maren Kirchner, Svenja Elkenkamp, Philipp Cimiano, Christoph Düsing, Hannah Marchi, Sophie Schmiegel, Christiane Fuchs, Simon Claßen, Kirsten-Laura Meier, Rainer Borgstedt, Sebastian Rehberg, Wolfgang Greiner. . . . 664

The Impact of Human-Robot Collaboration Levels on Postural Stability During Working Tasks Performed While Standing: Experimental Study (e64892)
Daniele Bibbo, Giovanni Corvini, Maurizio Schmid, Simone Ranaldi, Silvia Conforto. . . . 675

Feasibility of a 12-Week, Therapist-Independent, Smartphone-Based Biofeedback Treatment for Episodic Migraine in Adults: Single-Center, Open-Label, 1-Armed Trial (e59622)
Amalie Poole, Ingunn Winnberg, Melanie Simpson, Anker Stubberud, Kjersti Vetvik, Marte-Helene Bjørk, Lise Øie, Petter Holmboe, Alexander Olsen, Erling Tronvik, Tore Wergeland. . . . 689

Usability Evaluation of an Electrically Powered Orthopedic Exerciser: Focus Group Interview and Satisfaction Survey Study (e60607)
Seojin Hong, Hyun Choi, Hyosun Kweon. . . . 704

Factors Shaping Public Perceptions of a Range of Robotic Technologies in Surgery: Cross-Sectional Web-Based Survey (e64224)
Sarek Shen, Deborah Xie, Andy Ding, Lisa Zhang, Francis Creighton. . . . 714

Evaluating the Usability of an HIV Prevention Artificial Intelligence Chatbot in Malaysia: National Observational Study (e70034)
Zhao Ni, Sunyoung Oh, Rumana Saifi, Iskandar Azwa, Frederick Altice. . . . 724

User Experience of a Bespoke Videoconferencing System for Web-Based Family Visitation for Patients in an Intensive Care Unit: 1-Year Cross-Sectional Survey of Nursing Staff (e54560)
Aoife Murray, Irial Conroy, Frank Kirrane, Leonie Cullen, Hemendra Worlikar, Derek O'Keeffe. . . . 733

Feasibility of Long-Term Physical Activity Measurement With a Wearable Activity Tracker in Patients With Axial Spondyloarthritis: 1-Year Longitudinal Observational Study (e68645)
Emil Thomassen, Anne Tveter, Inger Berg, Eirik Kristianslund, Andrew Reiner, Sarah Hakim, Laure Gossec, Gary J Macfarlane, Annette de Thurah, Nina Østerås. . . . 744

Building and Beta-Testing Be Well Buddy Chatbot, a Secure, Credible and Trustworthy AI Chatbot That Will Not Misinform, Hallucinate or Stigmatize Substance Use Disorder: Development and Usability Study (e69144)
Adam Salyers, Sheana Bull, Joshva Silvasstar, Kevin Howell, Tara Wright, Farnoush Banaei-Kashani. . . . 755

The mChoice App, an mHealth Tool for the Monitoring of Preexposure Prophylaxis Adherence and Sexual Behaviors in Young Men Who Have Sex With Men: Usability Evaluation (e59780)
Fabiana Dos Santos, Maeve Brin, Mary Tanner, Carla Galindo, Rebecca Schnall. . . . 763

Use of Wearable Transdermal Alcohol Sensors for Monitoring Alcohol Consumption After Detoxification With Contingency Management: Pilot Randomized Feasibility Trial (e64664)
Eileen Brobbin, Colin Drummond, Stephen Parkin, Paolo Deluca. . . . 772

XSL•FO
RenderX

Requirement Analysis for Data-Driven Electroencephalography Seizure Monitoring Software to Enhance Quality and Decision Making in Digital Care Pathways for Epilepsy: A Feasibility Study from the Perspectives of Health Care Professionals (e59558)
Pantea Keikhosrokiani, Johanna Annunen, Jonna Komulainen-Ebrahim, Jukka Kortelainen, Mika Kallio, Päivi Vieira, Minna Isomursu, Johanna Uusimaa. . . . 792

Integrating a Large Language Model Into a Socially Assistive Robot in a Hospital Geriatric Unit: Two-Wave Comparative Study on Performance, Engagement, and User Perceptions (e81936)
Lauriane Blavette, Sébastien Dacunha, Xavier Alameda-Pineda, Jeanne Cattoni, Anne-Sophie Rigaud, Maribel Pino. . . . 817

Enhancing Safety in Clinical Practice: Cross-Sectional, Within-Subjects, Simulated-Use Compliance Study Using BD PosiFlush SafeScrub (e78967)
Kyle De Boer, John Simpson, Purnima Sharma, Trevor Steele. . . . 839

Optimizing SloMo, a Digitally Supported Therapy Targeting Paranoia, for Implementation: Inclusive, Human-Centered Design Study (e75377)
Thomas Gant, Kathryn Taylor, Thomas Ward, Philippa Garety, Amy Hardy. . . . 850

Patient-Derived Design Principles for Technology-Enabled Healing at Home Following Hospital Discharge: Mixed Methods Study (e72913)
Lindsey Philpot, Abhinav Singla, Sagar Dugani, Rachel Canning, Christina Smith, Meredith DeZutter, Priya Ramar, Jennifer Hovell, Jon Ebbert.
8 6 7

Patient Experiences With a Mobile Self-Care Solution for Low-Complex Orthopedic Injuries: Mixed Methods Study (e53074)
Jelle Spierings, Gijs Willinge, Marike Kokke, Sjoerd Repping, Wendela de Lange, Thijs Geerdink, Ruben van Veen, Detlef van der Velde, J Goslings, Bas Twigt, Collaboration Group. . . . 880

Effect of SMS Ward Round Notifications on Inpatient Experience in Acute Medical Settings: Retrospective Cohort Study (e57470)
Jongchan Lee, Soyeon Ahn, Jung Ohn, Eun Kim, Yejee Lim, Hye Kim, Hee-Sun Park, Jae Cho, Sun-wook Kim, Jiwon Ryu, Jihye Kim, Hak Jang, Nak-Hyun Kim. . . . 893

Empowering Community Health Workers With Scripted Medicine: Design Science Research Study (e57545)
Dario Staehelin, Damaris Schmid, Felix Gerber, Mateusz Dolata, Gerhard Schwabe. . . . 902

Knowledge Sharing Maturity Model for Medical Imaging Departments: Development Study (e54484)
Maryam Almashmoum, James Cunningham, John Ainsworth. . . . 922

Device Failures and Adverse Events Associated With Rhinolaryngoscopes: Analysis of the Manufacturer and User Facility Device Experience (MAUDE) Database (e67036)
Shao-Hsuan Chang, Daishi Chen, Chi-Sheng Chen, Dong Zhou, Lung-Kun Yeh. . . . 964

Completion Rates for Ecological Momentary Assessments of Food Intake During Pregnancy and Post Partum: Descriptive Study (e67081)
Sarah Sanders, Serwaa Omowale, Andrea Casas, Alexis Kiyanda, Abigail Kosbie, Yu-Hsuan Lai, Meredith Wallace, Stephen Rathbun, Tiffany Gary-Webb, Esa Davis, Lora Burke, Dara Méndez. . . . 973

Text Messaging Between Patients With Inflammatory Rheumatic Diseases and Pharmacists to Solve Drug-Related Problems: Prospective Feasibility Study (e66514)
Lex Haegens, Charlotte Bekker, Marcel Flendrie, Bart van den Bemt, Victor Huiskes. . . . 984

XSL•FO
RenderX

Intention to Use Automated Diagnosis and Clinical Risk Perceptions Among First Contact Clinicians in Resource-Poor Settings: Questionnaire-Based Study Focusing on Acute Burns (e56300)
Constance Boissin, Lisa Blom, Zara Taha, Lee Wallis, Nikki Allorto, Lucie Laflamme. . . . 998

Mobile-Based Cognitive Behavioral Therapy for Health Care Workers' Mental Health in Ecuador: Quasi-Experimental Study (e58943)
Sandra Muñoz-Ortega, Rubén Alvarado Muñoz, Daniela Santamaria Guayaquil, Jade Pluas-Borja, Marco Faytong-Haro. . . . 1010

Finnish Registry-Based Protocol for Screening and Management of Fatigue and Cognitive Problems in Multiple Sclerosis: Observational Study (e67990)
Päivi Hämäläinen, Elina Lämsä, Matias Viitala, Hanna Kuusisto, Marja Niiranen, Sari Avikainen, Juha Puustinen, Mervi Ryytty, Juhani Ruutiainen, Merja Soilu-Hänninen. . . . 1026

Health Care Providers' Experiences and Perceptions With Telehealth Tools in a Hospital-at-Home Program: Mixed Methods Study (e56860)
Shi Low, Stephanie Ko, Ian Ang. . . . 1036

Assessing the Impact on Electronic Health Record Burden After Five Years of Physician Engagement in a Canadian Mental Health Organization: Mixed-Methods Study (e65656)
Tania Tajirian, Brian Lo, Gillian Strudwick, Adam Tasca, Emily Kendell, Brittany Poynter, Sanjeev Kumar, Po-Yen Chang, Candice Kung, Debbie Schachter, Gwyneth Zai, Michael Kiang, Tamara Hoppe, Sara Ling, Uzma Haider, Kavini Rabel, Noelle Coombe, Damian Jankowicz, Sanjeev Sockalingam. . . . 1052

Swedish Version of the System Usability Scale: Translation, Adaption, and Psychometric Evaluation (e64210)
Helene Åvik Persson, Charlotte Castor, Nilla Andersson, Mia Hylén. . . . 1068

Evaluating the Development, Reliability, and Validation of the Tele-Primary Care Oral Health Clinical Information System Questionnaire: Cross-Sectional Questionnaire Study (e53630)
Rosnah Sutan, Shahida Ismail, Roszita Ibrahim. . . . 1078

Creation of Text Vignettes Based on Patient-Reported Data to Facilitate a Better Understanding of the Patient Perspective: Design Study (e58077)
Sofia Kelly, Emilie Wegener, Lars Kayser. . . . 1093

Development of a Web-Based Intervention for Middle Managers to Enhance Resilience at the Individual, Team, and Organizational Levels in Health Care Systems: Multiphase Study (e67263)
Eva Gil-Hernández, Irene Carrillo, Jimmy Martin-Delgado, Daniel García-Torres, José Mira. . . . 1109

A Web-Based Lifestyle-Related Course for People Living With Multiple Sclerosis: Quantitative Evaluation of Course Completion, Satisfaction, and Lifestyle Changes Among Participants Enrolled in a Randomized Controlled Trial (e59363)
Maggie Yu, Sandra Neate, Steve Simpson-Yap, Rebekah Davenport, William Bevens, George Jelinek, Jeanette Reece. . . . 1127

Patients' Perceptions of Using a Digital Previsit Tool in Outpatient Settings (Part 2): Qualitative Study (e73477)
Petra Pohl, Emma Kjörk. . . . 1149

Pursuit of Digital Innovation in Psychiatric Data Handling Practices in Ireland: Comprehensive Case Study (e64919)
Rana Zeeshan, John Bogue, Amna Gill, Mamoona Asghar. . . . 1164

Usability and Quality of the JoyPop App: Prospective Evaluation Study (e65472)
Ishaq Malik, Teagan Neufeld, Aislin Mushquash. . . . 1177

XSL•FO
RenderX

Usability and Usefulness of Occupational Health Care Patient Portals: Patient-Based Cross-Sectional Study (e73197)
Sari Nissinen, Pauliina Toivio, Erja Sormunen. 1192

A Comprehensive Profiling System Integrating Myers-Briggs Type Indicator (MBTI) and Dominance, Influence, Steadiness, and Conscientiousness (DISC) for Personalized Health Training: Correlational Analysis and Usability Evaluation (e73397)
Donghyun Kim, Dong Lee, Mi Hwang. 1207

Personalized Digital Care Pathways Enable Enhanced Patient Management as Perceived by Health Care Professionals: Mixed-Methods Study (e68581)
David Rodrigues, Clara Jasmins, Ricardo Ladeiras-Lopes, Luis Patrao, Eduardo Rodrigues. 1219

ChatGPT-Delivered Physical Activity Intervention for Children With Autism Spectrum Disorder: Pre-Post Feasibility Study (e71119)
U ur Aydemir. 1228

Using a Mobile Health App (ColonClean) to Enhance the Effectiveness of Bowel Preparation: Development and Usability Study (e58479)
Hui-Yu Chen, Ming-Hsiang Tu, Miao-Yen Chen. 1238

NutriDiary, a Smartphone-Based Dietary Record App: Description and Usability Evaluation (e62776)
Linda Klasen, Stefanie Koch, Maike Benz, Johanna Conrad, Ute Alexy, Konrad Blaszkiewicz, Ionut Andone, Ute Nöthlings. 1250

Think-Aloud Testing of a Companion App for Colonoscopy Examinations: Usability Study (e67043)
Christine Jacob, Roman Müller, Sonja Schüler, Alix Rey, Guillaume Rey, Berj Armenian, Alain Vonlaufen, Michael Drepper, Marius Zimmerli. .
1 2 6 2

Comparison of a Novel Machine Learning–Based Clinical Query Platform With Traditional Guideline Searches for Hospital Emergencies: Prospective Pilot Study of User Experience and Time Efficiency (e52358)
Hamza Ejaz, Hon Tsui, Mehul Patel, Luis Ulloa Paredes, Ellen Knights, Shah Aftab, Christian Subbe. 1278

Usability and User Experience Testing of a Co-Designed Electronic Patient-Reported Outcomes App (“MyPal for Adults”) for Palliative Cancer Care: Mixed Methods Study (e57342)
Panos Bonotis, Pantelis Angelidis, Pantelis Natsiavas. 1287

Satisfaction and Usability of a Commercially Available Medication Adherence App (Medisafe) Among Medically Underserved Patients With Chronic Illnesses: Survey Study (e63653)
Christa Hartch, Mary Dietrich, B Lancaster, Shelagh Mulvaney, Deonni Stolldorf. 1303

Evaluating the Usability, Acceptability, User Experience, and Design of an Interactive Responsive Platform to Improve Perinatal Nurses’ Stigmatizing Attitudes Toward Substance Use in Pregnancy: Mixed Methods Study (e67685)
Michael Rubyan, Yana Gouseinov, Mikayla Morgan, Deborah Rubyan, Divya Jahagirdar, David Choberka, Carol Boyd, Clayton Shuman. . . .
1 3 1 8

Effects of a Mobile Storytelling App (Huiyou) on Social Participation Among People With Mild Cognitive Impairment: Pilot Randomized Controlled Trial (e70177)
Di Zhu, Abdullah Al Mahmud, Wei Liu. 1336

XSL•FO
RenderX

Understanding the Behavioral Determinants of First Responder App Adoption by Integrating Perspectives From the Unified Theory of Acceptance and Use of Technology and Health Belief Model: Cross-Sectional Survey (e69934)
Cas von Winckelmann, Robyn Vanherle, Lara Schreurs, Olivier Hoogmartens, Heidi Salaets, Jan De Spiegeleer, Marc Sabbe, Kathleen Beullens. . . . 1356

Participants' Perspectives on the iCareBreast Mobile-Based Perioperative Care Program for Women Undergoing Breast Cancer Surgery: Qualitative Process Evaluation (e71686)
Yan Pang, Honggu He, Minna Pikkarainen, Swee-Ho Lim. . . . 1375

Patient's Perceptions of a Centralized Virtual Ward for Remote Patient Monitoring in Primary Care: Qualitative Study (e78780)
Alex Jaranka, Gunnar Nilsson, Terese Stenfors, Maria Hägglund, Panagiotis Papachristou, Marina Taloyan. . . . 1387

Family Caregivers' Perspectives on the Potential of Drone-Based Medication Delivery in Palliative Home Care: Qualitative Focus Group Study (e80320)
Franziska Fink, Sabrina Otto, Martin Grünthal, Anne Lehmann, Patrick Jahn. . . . 1400

Capturing Movement Behaviors in Latinas: Feasibility, Validity, and Acceptability Study of an Ecological Momentary Assessment Protocol (e75855)
Jaclyn Maher, Peyton Greco, Eugenia Fernandez, Brynn Hudgins, Sandra Echeverria. . . . 1413

Perceptions of User-Generated Content as a Source of Health Messages in Smoking Cessation Mobile Interventions: Focus Group Study (e76804)
Michael Wakeman, Lydia Tesfaye, Tim Gregory, Erin Leahy, Gunnar Baskin, Greg Gruse, Brandon Kendrick, Sherine El-Toukhy. . . . 1427

Improving Mental Health Referral Systems in Rural Australia: Co-Design Study With Health Professionals and Consumers (e73460)
Kate Bartel, Asini Siriwardene, Paul Worley, Kim Pearce, Niranjan Bidargaddi, Bronwin Patrickson, Simon Moody, Sharon Wingard, Martin Jones, Brian McKenny, Darryl Cameron, Sharon Lawn, Amy Mendham. . . . 1439

Exploratory Research With a Health Consumer Group on Social Robot Use Among Older Adults: Qualitative Study (e70462)
James Sadler, Aila Khan, Omar Mubin, Michael Lwin. . . . 1452

Testing the Acceptability and Feasibility of a Gender-Informed Smoking Cessation mHealth App for Women: Mixed Methods Approach (e71683)
Osnat Melamed, Kamna Mehra, Allison Gayapersad, Roshni Panda, Nadia Minian, Laurie Zawertailo, Leslie Buckley, Marta Maslej, Lorraine Greaves, Andreea Brabete, Jonathan Rose, Matt Ratto, Peter Selby. . . . 1461

Design Guidelines for a Game-Based Physical Rehabilitation System: Focus Group Study (e67336)
Ahmad S Elaklouk, Ratna Ramli, Samah M Alakklouk, Norsafinar Rahim. . . . 1475

Clinician Needs and Requirements for a Decision Aid Navigator: Qualitative Study (e69756)
Brad Morse, Carrie Reale, An Nguyen, Erin Latella, Hannah Bauguess, Shilo Anders, Pamela Roberts, Spencer SooHoo, Robert El-Kareh, Andrey Soares, Lisa Schilling. . . . 1493

Symptom Management Preference and Persona Development for Mobile Health Design Targeting Chinese Older Adult Patients With Breast Cancer: Descriptive Qualitative Study (e71448)
Danyu Li, Yanlin Zhu, Changrong Yuan, Qingmei Huang, Yang Yang, Zhaohui Geng, Xiaotong Yuan, Fulei Wu. . . . 1504

Exploring Clinician Experiences With a Digital Platform Supporting Orthopedic Care That Integrates Patient-Generated Health Data: Qualitative Study of Early Users (e65216)
Dmitri Katz, Daniel Gooch, Linda Price, Irum Rauf, Oliver Pearce, Blaine Price. . . . 1515

XSL•FO
RenderX

An Informatics-Based, Payer-Led, Low-Intensity Multichannel Educational Campaign Designed to Decrease Postdischarge Utilization for Medicare Advantage Members: Retrospective Evaluation (e63841)
Danica Fernandes, Elise Kokonas, Jai Bansal, Ken Hayashima, Brian Hurley, Annabel Ryu, Snehal Mhatre, Mohammed Ghori, Kelly Craig, Amanda Zaleski, Lily Vogel, Alena Baquet-Simpson, Daniel Reif. 1529

Exploring Youth Perspectives on Digital Mental Health Platforms: Qualitative Descriptive Study (e69907)
Sarah Daniel, Lauren Volcko, Emilie Bassi, Julia Hews-Girard, Katherine Bright, Marianne Barker, Lia Norman, Karina Pintson, Geneca Henry, Sumaya Soufi, Chukwudumbiri Omorotionmwan, Melanie Fersovitch, Leanne Stamp, Karen Moskovic, David Johnson, Gina Dimitropoulos. .
1 5 4 4

Preuse Acceptance of a Family-Centered, Need-Based, and Interprofessional Perinatal Care Mobile Health Intervention: Exploratory Study (e66658)
Kristina Killinger, Verena Seyfried, Katharina Brusniak, Markus Wallwiener, Michael Abou-Dakn, Dorothea Scholle, Stephanie Wallwiener. . .
1 5 5 8

Exploring Clinicians' and Patients' Acceptance and Utilization of a Digital Solution to Support Individualized Care in Diabetes Specialist Outpatient Care (DigiDiaS): Qualitative Study (e70301)
Maria Mollestad, Annesofie Jensen, Heidi Holmen, Tone Singstad, Eirik Årsand, Jacob Winther, Astrid Torbjørnsen. 1570

Adapting a Text Messaging Intervention to Improve Diabetes Medication Adherence in a Spanish-Speaking Population: Qualitative Study (e66668)
Jacqueline Seiglie, Seth Tobolsky, Ricaurte Crespo Trevino, Lluvia Cordova, Sara Cromer, A Caballero, Margarita Alegria, J Miranda, Deborah Wexler, Lindsay Mayberry. 1585

Perception and Evaluation of a Knowledge Transfer Concept in a Digital Health Application for Patients With Heart Failure: Mixed Methods Study (e56798)
Madeleine Flaucher, Sabrina Berzins, Katharina Jaeger, Michael Nissen, Jana Rolny, Patricia Trißler, Sebastian Eckl, Bjoern Eskofier, Heike Leutheuser. 1601

The Utilization of Electronic Consultations (eConsults) to Address Emerging Questions Related to Long COVID-19 in Ontario, Canada: Mixed Methods Analysis (e58582)
Jatinderpreet Singh, Michael Quon, Danica Goulet, Erin Keely, Clare Liddy. 1614

Understanding Appropriation of Digital Self-Monitoring Tools in Mental Health Care: Qualitative Analysis (e60096)
Lena de Thurah, Glenn Kiekens, Jeroen Weermeijer, Lotte Uyttebroek, Martien Wampers, Rafaël Bonnier, Inez Myin-Germeys. 1622

Development and Testing of the Kids Hurt App, a Web-Based, Pain Self-Report App for First Nations Youths: Mixed Methods Study (e48370)
Karlee Francis, Julie Francis, Margot Latimer, Hayley Gould, Shante Blackmore, Emily MacLeod. 1634

Understanding Experiences of Telehealth in Palliative Care: Photo Interview Study (e53913)
Mahima Kalla, Teresa O'Brien, Olivia Metcalf, Rashina Hoda, Xiao Chen, Andy Li, Catriona Parker, Michael Franco, Sam Georgy, Kit Huckvale, Christopher Bain, Peter Poon. 1647

The Effects of Digital eHealth Versus Onsite 2-Day Group-Based Education in 255 Patients With Irritable Bowel Syndrome: Cohort Study (e43618)
Birgitte Berentsen, Camilla Thuen, Eline Hillestad, Elisabeth Steinsvik, Trygve Hausken, Jan Hatlebakk. 1667

Unveiling Sociocultural Barriers to Breast Cancer Awareness Among the South Asian Population: Case Study of Bangladesh and West Bengal, India (e53969)
Fahmida Hamid, Tania Roy. 1681

XSL•FO
RenderX

The Significance of a Cerebrovascular Accident Outcome Prediction Model for Patients, Family Members, and Health Care Professionals: Qualitative Evaluation Study (e56521)
Corinne Allaart, Sanne van Houwelingen, Pieter Hilkens, Aart van Halteren, Douwe Biesma, Lea Dijksman, Paul van der Nat. 1697

Exploring the Users' Perspective of the Nationwide Self-Exclusion Service for Gambling Disorder, "Spelpaus": Qualitative Interview Study (e66045)
Johanna Tjernberg, Sara Helgesson, Anders Håkansson, Helena Hansson. 1712

Barriers and Facilitators to User Engagement and Moderation for Web-Based Peer Support Among Young People: Qualitative Study Using the Behavior Change Wheel Framework (e64097)
Ananya Ananya, Janina Tuuli, Rachel Perowne, Leslie Gutman. 1724

The Safety of Digital Mental Health Interventions: Findings and Recommendations From a Qualitative Study Exploring Users' Experiences, Concerns, and Suggestions (e62974)
Rayan Taher, Daniel Stahl, Sukhi Shergill, Jenny Yiend. 1743

Provider Perspectives on the Use of Mental Health Apps, and the BritePath App in Particular, With Adolescents at Risk for Suicidal Behavior: Qualitative Study (e64867)
Frances Lynch, Julie Cavese, Lucy Fulton, Nancy Vuckovic, David Brent. 1755

Benefits and Barriers to mHealth in Hypertension Care: Qualitative Study With German Health Care Professionals (e52544)
Susann May, Felix Muehlensiepen, Eileen Wengemuth, Frances Seifert, Martin Heinze, Dunja Bruch, Sebastian Spethmann. 1789

Perceived Risks, Mitigation Strategies, and Modifiability of Telehealth in Rural and Remote Emergency Departments: Qualitative Exploration Study (e58851)
Christina Tsou, Justin Yeung, Melanie Goode, Josephine Mcdonnell, Aled Williams, Stephen Andrew, Jenny Tetlow, Andrew Jamieson, Delia Hendrie, Christopher Reid, Sandra Thompson. 1803

Experiences of Wheelchair Users With Spinal Cord Injury With Self-Tracking and Commercial Self-Tracking Technology ("In Our World, Calories Are Very Important"): Qualitative Interview Study (e65207)
Vasiliki Mylonopoulou, Katerina Cerna, Alexandra Weilenmann, Mattias Rost, Tobias Holmlund. 1820

Blended Mobile-Based Interventions With Integrated Virtual Reality Exposure Therapy for Anxiety Disorders: Thematic Analysis of Patient Perspectives (e60957)
Jari Planert, Anne-Sophie Hildebrand, Alla Machulska, Kati Roesmann, Marie Neubert, Sebastian Pilgramm, Juliane Pilgramm, Tim Klucken.
1 8 3 8

Using Gesture and Speech to Control Surgical Lighting Systems: Mixed Methods Study (e70628)
Nima Zargham, Anke Reinschluessel, Andre Mühlenbrock, Thomas Muender, Timur Cetin, Verena Uslar, Dirk Weyhe, Rainer Malaka, Tanja Döring. 1855

Experiences With a Multicomponent Digital Behavioral Pain Management Intervention for Adults With Sickle Cell Disease: Qualitative Analysis of the CaRISMA Trial (e73719)
Olubusola Oluwole, Lakeya McGill, Kadeem Gayle, Julia O'Brien, Kaleab Abebe, Charles Jonassaint, Megan Hamm. 1876

A Sociotechnical Approach to Bring-Your-Own-Device Security in Hospitals: Development and Pilot Testing of a Maturity Model Using Mixed Methods Action Research (e71912)
Tafheem Wani, Antonette Mendoza, Kathleen Gray. 1890

Diverse Attitudes and Experiences With Technology Use During the COVID-19 Pandemic Among Asian American and Pacific Islander Adults (the COMPASS Study): Survey Study (e64999)
Linda Park, Serena Chi, Myka Lay, Nicole Phan, Janice Tsoh, Oanh Meyer, Bora Nam, Van Park. 1915

XSL•FO
RenderX

Barriers and Enablers to Young People's Posting, Responding, and Reading Behaviors on Mental Health Forums Using the Behavior Change Wheel: Qualitative Study (e71549)
Zhen Lim, Si Lim, Siqi Lu, Leslie Gutman. . . . 1932

Cycling-Based Telerehabilitation: Acceptability and Feasibility Study (e71099)
Sara Arlati, Vera Colombo, Marta Mondellini, Roberta Nossa, Chiara Grasso, Mauro Rossini, Emilia Biffi, Alessia Fumagalli, Eleonora Diella, Eleonora Guanziroli, Nicole Sanna, Emilia Ambrosini, Marco Sacco, Franco Molteni. . . . 1948

Immersive Virtual Reality eHealth Intervention to Reduce Anxiety and Depression in Pregnant Women: Randomized Controlled Trial (e71708)
Marta Jimenez-Barragan, Amparo Del Pino Gutierrez, Gloria Sauch Valmaña, Olga Monistrol, Carme Monge Marcet, Mar Pallarols Badia, Ignasi Garrido, Anna Carmona Ruiz, Oriol Porta Roda, Cristina Esquinas, Gemma Falguera Puig. . . . 1967

Smartphone Apps and Wearables for Health Parameters in Young Adulthood: Cross-Sectional Study (e64629)
Gaia Leuzzi, Mirko Job, Aldo Scafoglieri, Marco Testa. . . . 1988

Developing a Data Trust Model (Not Only) for Sleep Research: Conceptual Study and Quantitative Survey (e66513)
Raphael Dressle, Dieter Riemann, Nicole Thoma, Christina Erler, Rodger Burmeister, Bianka Jogwitz, Katharina Domschke, Kai Spiegelhalder, Joachim Boldt, Svenja Wiertz, Bernd Feige. . . . 2004

Digital Health Literacy in Adults With Low Reading and Writing Skills Living in Germany: Mixed Methods Study (e65345)
Saskia Muellmann, Rebekka Wiersing, Hajo Zeeb, Tilman Brand. . . . 2020

The Chinese Version of the DigiHealthCom (Digital Health Competence) Instrument for Assessing Digital Health Competence of Health Care Professionals: Translation, Adaptation, and Validation Study (e65373)
Lu Gao, Meilian Chen, Jingxin Wei, Jinni Wang, Xiaoyan Liao. . . . 2038

Application of Clinical Department–Specific AI-Assisted Coding Using Taiwan Diagnosis-Related Groups: Retrospective Validation Study (e59961)
An-Tai Lu, Chong-Sin Liou, Chia-Hsin Lai, Bo-Tsz Shian, Ming-Ta Li, Chih-Yen Sun, Hao-Yun Kao, Hong-Jie Dai, Ming-Ju Tsai. . . . 2051

Enhancing Early Language Disorder Detection in Preschools: Evaluation and Future Directions for the Gades Platform (e60424)
María Dolón-Poza, Ana-Marta Gabaldón-Pérez, Santiago Berrezueta-Guzman, David López Gracia, María-Luisa Martín-Ruiz, Iván Pau De La Cruz. . . . 2065

Developing a Novel Mobile App to Support HIV Testing and Pre-Exposure Prophylaxis Uptake Among Men Who Have Sex With Men: Formative and Technical Pilot Study (e62830)
Albert Liu, Thiago Torres, Cat-Dancing Alleyne, Janie Vinson, Kelly Bojan, Pedro Serrano, Temitope Oyedele, Amayvis Garcia, Elizabeth Enriquez-Bruce, Patricia Emmanuel, Jeb Jones, Kathryn Muessig, Susan Buchbinder, Patrick Sullivan, Lisa Hightow-Weidman, Hyman Scott.
2 0 8 5

Designing a Visual Analytics Tool to Support Data Analysis Tasks of Digital Mental Health Interventions: Case Study (e64967)
Gyuwon Jung, Heejeong Lim, Kyungsik Han, Hyungsook Kim, Uichin Lee. . . . 2104

Acceptability and Usability of a Socially Assistive Robot Integrated With a Large Language Model for Enhanced Human-Robot Interaction in a Geriatric Care Institution: Mixed Methods Evaluation (e76496)
Lauriane Blavette, Sébastien Dacunha, Xavier Alameda-Pineda, Daniel Hernández García, Sharon Gannot, Florian Gras, Nancie Gunson, Séverin Lemaignan, Michal Polic, Pinchas Tandeitnik, Francesco Tonini, Anne-Sophie Rigaud, Maribel Pino. . . . 2125

XSL•FO
RenderX

Providers' Perspective on the Feasibility of Digital Self-Management of Blood Pressure in Refugees: Mixed Methods Study (e66176)
Carol Gonzalez, Marcos Real, Nargis Ahmadi, Raghad Aljenabi, Lana Bridi, Nour Makarem, Job Godino, Tala Al-Rousan. 2141

Factors Determining Acceptance of Internet of Things in Medical Education: Mixed Methods Study (e58377)
Khadija Alhumaid, Kevin Ayoubi, Maha Khalifa, Said Salloum. 2154

The Impact of Trust and the Role of the Opt-Out Mechanism in Willingness to Share Health Data via Electronic Health Records in Germany: Telephone Survey Study (e65718)
Felix Wilke. 2174

Collecting at-Home Biometric Measures for Longitudinal Research From the i3C: Feasibility and Acceptability Study (e71103)
Marta Russell, Erin Cain, Lydia Bazzano, Ileana De Anda, Jessica Woo, Elaine Urbina. 2187

Assessing Postoperative Pain in Patients Who Underwent Total Knee Arthroplasty Using an Automated Self-Logging Patient-Reported Outcome Measure Collection Device: Retrospective Cohort Study (e65271)
Prabjit Ajrawat, Blaine Price, Daniel Gooch, Rudolf Serban, Ruqaiya Al-Habsi, Oliver Pearce. 2197

Factors Influencing Health Workers' Acceptance of Guideline-Based Clinical Decision Support Systems for Preventive Services in Thailand: Questionnaire-Based Study (e57314)
Tullaya Sitasuwan, Prapat Suriyaphol, Saranath Lawpoolsri, Ngamphol Soonthornworasiri, Wirichada Pan-ngum. 2206

Evaluating the Usability, Technical Performance, and Accuracy of Artificial Intelligence Scribes for Primary Care: Competitive Analysis (e71434)
Emily Ha, Isabelle Choon-Kon-Yune, LaShawn Murray, Siying Luan, Enid Montague, Onil Bhattacharyya, Payal Agarwal. 2220

Clinician Perspectives of a Magnetic Resonance Imaging–Based 3D Volumetric Analysis Tool for Neurofibromatosis Type 2–Related Schwannomatosis: Qualitative Pilot Study (e71728)
Shelby Desroches, Alice Huang, Rithvik Ghankot, Steven Tommasini, Daniel Wiznia, Frank Buono. 2233

Leveraging Dual Usability Methods to Evaluate Clinical Decision Support Among Patients With Traumatic Brain Injury: Mixed Methods Study (e60268)
Rubina Rizvi, Sameen Faisal, Mark Sussman, Patricia Mendlick, Sam Brown, Elizabeth Lindemann, Sean Switzer, Genevieve Melton, Christopher Tignanelli. 2244

Digital Capability, Open-Source Use, and Interoperability Standards Within the National Health Service in England: Survey of Health Care Trusts (e66398)
Matthew Bennion, Ross Spencer, Roger Moore, Richard Kenyon. 2257

Health-Related Internet Use Among Outpatients Undergoing Cancer Treatment During the COVID-19 Pandemic: Cross-Sectional Survey Study (e72614)
Nicole Erickson, Daniel Nasseh, Sarah Seynstahl, Nicole Jost, Rachel Wuerstlein, Jozefina Casuscelli, Sebastian Theurich, Volker Heinemann, Theres Fey. 2274

User-Centered Delivery of AI-Powered Health Care Technologies in Clinical Settings: Mixed Methods Case Study (e76241)
Meredith Schreier, Randall Brandt, Hien Brown, T Saensuksopa, Christine Silva, Laura Vardoulakis. 2286

Collecting Real-World Data via an In-Home Smart Medication Dispenser: Longitudinal Observational Study of Survey Panel Persistency, Response Rates, and Psychometric Properties (e60438)
Benjamin Ogorek, Thomas Rhoads, Erica Smith. 2303

Comparisons of Physicians', Nurses', and Social Welfare Professionals' Experiences With Participation in Information System Development: Cross-Sectional Survey Study (e51495)
Susanna Martikainen, Johanna Viitanen, Samuel Salovaara, Ulla-Mari Kinnunen, Tinja Lääveri. 2319

XSL•FO
RenderX

Patient-Generated Collections for Organizing Electronic Health Record Data to Elevate Personal Meaning, Improve Actionability, and Support Patient–Health Care Provider Communication: Think-Aloud Evaluation Study (e50331)
Drashko Nakikj, David Kreda, Karan Luthria, Nils Gehlenborg. . . . 2334

Mental Health Providers' Challenges and Solutions in Prescribing Over Telemedicine: Content Analysis of Semistructured Interviews (e65419)
Julia Ivanova, Mollie Cummins, Hiral Soni, Triton Ong, Brian Bunnell, Esteban López, Brandon Welch. . . . 2353

Health Care Professionals' Perspectives on Using eHealth Tools in Advanced Home Care: Qualitative Interview Study (e60582)
Eric Rivas, Ulf Lesley, Nadia Davoody. . . . 2365

Need Analysis of Clinician-Oriented Integrated Precision Oncology Decision Support Tools: Qualitative Descriptive Study (e67476)
Sheng Zheng, Yi Feng, Jieran Long, Xiang Li, Meng Zhang, Hui Chen, Xue Du, Huilong Duan, Shuqin Jia, Nan Wu, Xudong Lu. . . . 2379

Personal Health Record Software for Neuroendocrine Tumors: Patient-Centered Design Approach (e68788)
Juan Hourcade, Michael O'Rorke, Elizabeth Chrischilles, Nicholas Rudzianski, Brian Gryzlak, Kerry Peterman, Chris Ortman, Sam Wolstencroft, Margaret Bean, Elyse Gellerman, Josh Mailman, Maryann Wahmann. . . . 2398

Using Social Media Platforms to Raise Health Awareness and Increase Health Education in Pakistan: Structural Equation Modeling Analysis and Questionnaire Study (e65745)
Malik Munir, Nabil Ahmed. . . . 2409

Patients' and Physicians' Experience With and Acceptability of a Telemedicine Cabin: Mixed Methods Study (e55430)
Caroline de França, Paola Segalla, Felipe Reis, José Pereira, Alexandre de Mattos, Roberta Ferron, Cleyton de Oliveira, Jéssica Borges, Lilian Hoffmann, Edmundo Caboclo. . . . 2423

Gender-Inclusive Language in Public-Facing Labor and Delivery Web Pages in the New York Tristate Area: Cross-Sectional Study (e53057)
Sarah Isaac, Mark Dawes, Emily Howell, Antonia Oladipo. . . . 2438

Leveraging Smart Telemedicine Technology to Enhance Nursing Care Satisfaction and Revolutionize COVID-19 Care: Prospective Cohort Study (e53456)
You-Lung Chang, Chi-Ying Lin, Jiun Hsu, Sui-Ling Liao, Chun-Ti Yu, Hung-Chueh Peng, Chung-Yu Chen, Matthew Ma, Juey-Jen Hwang. . .
2 4 4 4

Examining Individuals' Use of the Internet for Health Care Activities Over Time: Results from the US National Health Interview Survey (e58362)
Zachary Junkins, Nusrath Zahan, David Neyens. . . . 2454

Development and Systematic Evaluation of a Progressive Web Application for Women With Cardiac Pain: Usability Study (e57583)
Monica Parry, Tony Huang, Hance Clarke, Ann Bjørnnes, Paula Harvey, Laura Parente, Colleen Norris, Louise Pilote, Jennifer Price, Jennifer Stinson, Arland O'Hara, Madusha Fernando, Judy Watt-Watson, Nicole Nickerson, Vincenza Spiteri DeBonis, Donna Hart, Christine Faubert.
2 4 6 8

Co-Designing, Developing, and Testing a Mental Health Platform for Young People Using a Participatory Design Methodology in Colombia: Mixed Methods Study (e66558)
Laura Ospina-Pinillos, Débora Shambo-Rodríguez, Mónica Sánchez-Nítola, Alexandra Morales, Laura Gallego-Sanchez, María Riaño-Fonseca, Andrea Bello-Tocancipá, Alvaro Navarro-Mancilla, Jaime Pavlich-Mariscal, Alexandra Pomares-Quimbaya, Carlos Gómez-Restrepo, Ian Hickie, Jo-An Occhipinti. . . . 2485

XSL•FO
RenderX

Use of Online Tools for Mental Health Among Racially and Ethnically Diverse College Students: Mixed Methods Study (e60628)
Sarah Hamza, Yesenia Aguilar Silvan, Lauren Ng. . . . 2503

Investigating How Clinicians Form Trust in an AI-Based Mental Health Model: Qualitative Case Study (e79658)
Anthony Kelly, Niharika Bhardwaj, Trine Holmberg Sainte-Marie, Pepijn Van de Ven, Ruth Melia, John Williams, Kim Mathiasen, Amalie Nielsen.
2 5 2 2

TeleAllergy: Potential of Telemedicine in Management of Patients With Allergies (e75483)
Hanna Lindemann, Emil Hammer, Luca Bonifacio, Christian Greis, Karin Hartmann. . . . 2545

Exploring Young Adults' Attitudes Toward AI-Driven mHealth Apps: Qualitative Study (e76075)
Ali Aboueldahab, Gabriele Damaschi, Marco D'Addario, Patrizia Steca. . . . 2553

Digital Health Intervention for Patient Monitoring in Immune-Mediated Inflammatory Diseases: Cocreation and Feasibility Study of the IMIDoc Platform (e58095)
Diego Benavent, Jose Iniesta-Chamorro, Marta Novella-Navarro, Miguel Pérez-Martínez, Nuria Martínez-Sánchez, Mónica Kaffati, Manuel Juárez-García, Marina Molinari-Pérez, Andrea González-Torbay, Mariana Gutiérrez, Natalia López-Juanes, Victoria Navarro-Compán, Irene Monjo-Henry, Germán Rodríguez-Rosales, Javier Bachiller, Enrique Calvo-Aranda, Xabier Michelena, Laura Berbel-Arcobé, Alejandro Balsa, IMIDOC Research Team, Enrique Gómez, Chamaida Plasencia-Rodríguez. . . . 2566

Understanding Dermatologists' Acceptance of Digital Health Interventions: Cross-Sectional Survey and Cluster Analysis (e59757)
Patrick Reinders, Matthias Augustin, Marina Otten. . . . 2579

Device Functionalities and Technology Acceptance for Innovations in Neonatal Ventilation and Enhanced, Immediate Newborn Care: International, Multicenter, Web-Based Survey Study (e64701)
Anna-Sophie Käferböck, Meggy Hayotte, Daniel Sieber, Martin Pillei, Martin Wald. . . . 2595

Pharmaceutical Analysis of Inpatient Prescriptions: Systematic Observation of Hospital Pharmacists' Practices in the Early User-Centered Design Phase (e65959)
Jesse Butruille, Natalina Cirnat, Mariem Alaoui, Jérôme Saracco, Etienne Cousein, Noémie Chaniaud. . . . 2605

Web-Based Delivery of an Effective Church-Based Intervention Program to Promote Cancer Screening (Community-based Health litEracy-focused intervention for breast and cervical Cancer Control) Among Korean Immigrant Women in the United States: Randomized Controlled Trial (e66092)
Hae-Ra Han, Yoon-Jae Lee, Deborah Min, Joyline Chepkorir, DaSol Hwang, Steve Chae. . . . 2622

Following Up Patients With Chronic Pain Using a Mobile App With a Support Center: Unicenter Prospective Study (e60160)
Marta Gómez-González, Nicolas Cordero Tous, Javier De la Cruz Sabido, Carlos Sánchez Corral, Beatriz Lechuga Carrasco, Marta López-Vicente, Gonzalo Olivares Granados. . . . 2633

Exploring Older Adults' Needs for a Healthy Life and eHealth: Qualitative Interview Study (e50329)
Paula Valkonen, Sari Kujala, Kaisa Savolainen, Riina-Riitta Helminen. . . . 2644

Exploring the Impact of Digital Peer Support Services on Meeting Unmet Needs Within an Employee Assistance Program: Retrospective Cohort Study (e68221)
Harpreet Nagra, Robert Mines, Zara Dana. . . . 2662

Identifying Strategies for Home Management of Ostomy Care: Content Analysis of YouTube (e66634)
Marketa Haughey, David Neyens, Casey Hopkins, Christofer Gonzaga, Melinda Harman. . . . 2677

Development of a Patient-Centered Symptom-Reporting Application in Pharmacy Settings Using a Hierarchical Patient-Friendly Symptom List: Developmental and Usability Study (e71439)
Seiya Watanabe, Hayato Kizaki, Satoko Hori. . . . 2692

XSL•FO
RenderX

Illustrating User Needs for eHealth With Experience Map: Interview Study With Chronic Kidney Disease Patients (e48221)
Paula Valkonen, Sini Hölsä, Johanna Viitanen, Sini Leinonen, Nina Karisalmi, Virpi Rauta. 2703

Best Practice Guide for Reducing Barriers to Video Call–Based Telehealth: Modified Delphi Study Among Health Care Professionals (e64079)
Lena Rettinger, Lea Aichinger, Veronika Ertelt-Bach, Andreas Huber, Susanne Javorszky, Lukas Maul, Peter Putz, Sevan Sargis, Franz Werner, Klaus Widhalm, Sebastian Kuhn. 2720

Evaluating the Needs and Characteristics of Individuals of Low Socioeconomic Status Using Digital Health Technology to Address Health-Related Social Needs: Mixed Methods Study With Patients and Care Providers (e69545)
Kayla Haydon, Cynthia LeRouge, Polina Durneva, Maria Diaz Campo, David Brown. 2736

Collaborative Design and Development of a Patient-Centered Digital Health App for Supportive Cancer Care: Participatory Study (e73829)
Sonia Difrancesco, Matthia Bauert, Claude Lehmann, Steven Häsler, Yi Zhang, Sven Hirsch, Philipp Ackermann, Kurt Stockinger, Monika Reif, Sunjoy Mathieu, Anna Götz, Andreas Wicki, Michael Krauthammer, Claudia Witt. 2758

Key Factors Shaping Successful Implementation of the Internet of Things (IoT) in Health Care: Qualitative Study (e71546)
Klas Palm, Carl Kronlid, Marie Elf, Anders Brantnell. 2777

Optimizing Participant Engagement in Cyberhealth Co-Design: Course-of-Action Framework Analysis (e70772)
Melanie Tremblay, Christine Hamel, Anabelle Viau-Guay, Dominique Giroux. 2791

Clinician Perceptions of a Novel Multicomponent Digital Care Assistant and Support Program for People After Stroke or Transient Ischemic Attack (CAPS) for the Secondary Prevention of Stroke: Qualitative Study (e72873)
Liam Allan, Jane Li, Tara Purvis, David Silvera-Tawil, Jan Cameron, Marlien Varnfield, Vanessa Smallbon, Julia Bomke, Natasha Lannin, Dominique Cadilhac. 2806

Designing for Patient-Centered Care and Equity in Virtual Hospital-at-Home Models: Quality Improvement Initiative Using Experience-Based Co-Design (e79679)
Mahabhir Kandola, Emma Wong, Robert Paquin, Kamal Arora, Harroop Sharda, Roman Deol, Mary Jung, Maria Montenegro, Megan MacPherson. 2822

Exploring User Experience and the Therapeutic Relationship of Short-Term Avatar-Based Psychotherapy: Qualitative Pilot Study (e66158)
Byeul Jang, Chisung Yuh, Hyeri Lee, Yu-Bin Shin, Heon-Jeong Lee, Eun Kang, Jeongyun Heo, Chul-Hyun Cho. 2832

The Research Agenda for Perinatal Innovation and Digital Health Project: Human-Centered Approach to Multipartner Research Agenda Codevelopment (e60825)
Haneen Amhaz, Sally Chen, Amanee Elchehimi, Kylin Han, Jade Gil, Lu Yao, Marianne Vidler, Kathryn Berry-Einarson, Kathryn Dewar, May Tuason, Nicole Prestley, Quynh Doan, Tibor van Rooij, Tina Costa, Gina Ogilvie, Beth Payne. 2840

Building a Decentralized Biobanking App for Research Transparency and Patient Engagement: Participatory Design Study (e59485)
Ananya Dewan, M Eifler, Amelia Hood, William Sanchez, Marielle Gross. 2854

Impact of Bottom-Up Cocreation of Nursing Technological Innovations: Explorative Interview Study Among Hospital Nurses and Managers (e60543)
Saskia van Steenis, Onno Helder, Helianthe Kort, Thijs van Houwelingen. 2871

XSL•FO
RenderX

Turning Patients' Open-Ended Narratives of Chronic Pain Into Quantitative Measures: Natural Language Processing Study (e80269)
Raquel Norel, Jennifer Gewandter, Zhengwu Zhang, Anika Tahsin, Chadi Abdallah, John Markman, Zhiyao Duan, Guillermo Cecchi, Paul Geha. . . . 2884

Adolescent Voices in Action—Co-Designing Digital Sexual and Reproductive Health Knowledge Translation Interventions: Community-Based Participatory Action Project (e71148)
Salima Meherali, Amyna Rehmani, Mariam Ahmad. . . . 2898

Experiences of Using Online Peer Forums Among People With Postpartum Psychosis: Interpretative Phenomenological Study (e80717)
Katherine Williams, Fiona Lobban, Elizabeth Chamberlain. . . . 2912

Requirements and Value Elicitation for a High-Fidelity Pelvic Floor Simulator for Physiotherapists: Mixed Methods Study (e72119)
Yael Zekaria, Antonia Tzemanaki, Jonathan Rossiter. . . . 2929

Identifying Design Requirements for an Interactive Physiotherapy Dashboard With Decision Support for Clinical Movement Analysis of Musicians With Musculoskeletal Problems: Qualitative User Research Study (e65029)
Eduard Wolf, Karsten Morisse, Sven Meister. . . . 2944

Capturing Requirements for a Data Annotation Tool for Intensive Care: Experimental User-Centered Design Study (e56880)
Marceli Wac, Raul Santos-Rodriguez, Chris McWilliams, Christopher Bourdeaux. . . . 2962

Internet-Based Mental Health Intervention for Depressive Symptoms in Young Adults: Cost-Effectiveness Analysis (e68167)
Joyce You, Scotty Luk, Dilys Chow, Arthur Mak, Winnie Mak. . . . 2975

Use of the ME-BYO Index, a Mobile Health App, During an Online Strength Training Program in Adults: Fidelity, Feasibility, and Acceptability Study (e63123)
Yoshinobu Saito, Naoki Kikuchi, Kaname Watanabe, Sho Nakamura, Hiroto Narimatsu. . . . 2987

Using Crowdsourcing Internet of Things Technology to Reduce Caregiver Worry in Dementia-Related Lost Episodes: Longitudinal Observational Study (e73670)
Bel Wong, Tobi Cheng, Nicole Fung, Zhongming Lin, Ki-Kit Lai, Florence Ho, S-H Chan, Timothy Kwok. . . . 3000

Mobile Health App for Adults with Persisting Postconcussion Symptoms: Development and Usability Study (e75323)
Gøril Storvig, Anker Stubberud, Johanne Rauwenhoff, Liv Rønhovde, Martijn Smits, Simen Saksvik, Toril Skandsen, Erling Tronvik, Alexander Olsen. . . . 3018

Usability Evaluation of Digital Health Applications for Older People With Depressive Disorders: Prospective Observational Study in a Mixed Methods Design (e66271)
Magdalini Chatsatrian, Katharina Kunde, Jennifer Bosompem, Jan Dieris-Hirche, Nina Timmesfeld, Rainer Wirth, Georg Juckel, Magdalena Pape, Anna Mai, Chantal Giehl, Bianca Ueberberg, Horst Vollmar, Ina Otte, Theresa Busse. . . . 3032

Mobile Health Adoption in High-Risk Pregnancies Using Cluster Analysis of Biopsychosocial Outcomes: Observational Longitudinal Cohort Study (e67680)
Fernanda Schier de Fraga, Mayara Narita, Monique Schreiner, Flavio Belli, Jaqueline Leonel Celestino, Karolayne Braz Pereira, Gabriella Soecki, Vitória Bevervanso, Rogério de Fraga. . . . 3046

Designing Chatbots to Treat Depression in Youth: Qualitative Study (e66632)
Florian Kuhlmeier, Luise Bauch, Ulrich Gnewuch, Stefan Lüttke. . . . 3063

XSL•FO
RenderX

Understanding Barriers to Home Safety Assessment Adoption in Older Adults: Qualitative Human-Centered Design Study (e66854)
Jonathan Lee, Devika Patel, Meghana Gadgil, Simone Langness, Christiana von Hippel, Amanda Sammann. . . . 3080

Self-Determined Health App Evaluation Questionnaire Development: Mixed Methods Study (e63739)
Angelika Rzepka, Kurt Edegger, Stefan Welte, Diotima Bertel, Anja Mandl, Günter Schreier. . . . 3090

User-Centered Prototype Design of a Health Care Robot for Treating Type 2 Diabetes in the Community Pharmacy: Development and Usability Study (e48226)
Ching-Ju Chiu, Lin-Chun Hua, Jung-Hsien Chiang, Chieh-Ying Chou. . . . 3102

Telemedicine Booths for Screening Cardiovascular Risk Factors: Prospective Multicenter Study (e57032)
Mélanie Decambron, Christine Tchikladze Merand. . . . 3114

Education and Symptom Reporting in an mHealth App for Patients With Cancer: Mixed Methods Development and Validation Study (e60169)
Carolina Muñoz Olivar, Miguel Pineiro, Juan Gómez Quintero, Carlos Avendaño-Vásquez, Pablo Ormeño-Arriagada, Silvia Palma Rivadeneira, Carla Taramasco Toro. . . . 3127

Mobile Therapeutic Attention for Treatment-Resistant Schizophrenia (m-RESIST) Solution for Improving Clinical and Functional Outcomes in Treatment-Resistant Schizophrenia: Prospective, Multicenter Efficacy Study (e67659)
Jussi Seppälä, Eva Grasa, Anna Alonso-Solis, Alexandra Roldan-Bejarano, Marianne Haapea, Matti Isohanni, Jouko Miettunen, Johanna Caro Mendivelso, Cari Almazán, Katya Rubinstein, Asaf Caspi, Zolt Unoka, Kinga Farkas, Elisenda Reixach, Jesus Berdun, Judith Usall, Susana Ochoa, Iluminada Corripio, Erika Jääskeläinen, m-Resist Group. . . . 3140

Factors Related to Pain and Disability Outcomes After an Internet-Delivered or Physiotherapist-Led Exercise Program for Individuals With Chronic Whiplash Symptoms: Secondary Analysis of a Randomized Controlled Study (e67991)
Gunnel Peterson, Stefan Ljunggren, Anneli Peolsson. . . . 3156

A User-Centered Design Approach for a Screening App for People With Cognitive Impairment (digiDEM-SCREEN): Development and Usability Study (e65022)
Michael Zeiler, Nikolas Dietzel, Fabian Haug, Julian Haug, Klaus Kammerer, Rüdiger Pryss, Peter Heuschmann, Elmar Graessel, Peter Kolominsky-Rabas, Hans-Ulrich Prokosch. . . . 3174

Effectiveness of Using a Digital Wearable Plantar Pressure Device to Detect Muscle Fatigue: Within-Subject, Repeated Measures Experimental Design (e65578)
Fu-Yu Chen, Tzu-Yao Lin, Yi-Cheng Huang, Evina Widianawati. . . . 3187

Usability Testing of a Bystander Bullying Intervention for Rural Middle Schools: Mixed Methods Study (e67962)
Aida Midgett, Diana Doumas, Claudia Peralta, Matt Peck, Blaine Reilly, Mary Buller. . . . 3197

Evaluation of Satisfaction With a Secure, Connected Mobile App for Women in Assisted Reproductive Technology Programs: Prospective Observational Study (e63570)
Pauline Plouvier, Romaric Marcilly, Geoffroy Robin, Chaymae Benamar, Camille Robin, Virginie Simon, Anne Piau, Isabelle Cambay, Jessica Schiro, Christine Decanter. . . . 3211

Digital Cognitive Behavioral Therapy–Based Treatment for Insomnia, Nightmares, and Posttraumatic Stress Disorder Symptoms in Survivors of Wildfires: Pilot Randomized Feasibility Trial (e65228)
Fadia Isaac, Britt Klein, Huy Nguyen, Shaun Watson, Gerard Kennedy. . . . 3219

XSL•FO
RenderX

Design Preferences, Routines, and Well-Being of Older Adults Using Voice-Guided Digital Mindfulness: Qualitative Interview Study (e67533)
Lucy McCarren, Sanna Kuoppamäki. . . . 3250

A Digital Mental Health Intervention for Paranoia (the STOP App): Qualitative Study on User Acceptability (e70181)
Laura Eid, Alexandra Kenny, Rayan Taher, Pamela Jacobsen, Che Hsu, Jenny Yiend. . . . 3266

Use and User Experience of a Preconception Lifestyle App for Couples Undergoing in Vitro Fertilization: Mixed Methods Study (e65815)
Tessy Boedt, Sharon Lie Fong, Eline Dancet, Merijn Mestdagh, Joke Verbeke, David Geerts, Carl Spiessens, Christophe Matthys. . . . 3282

Story Retelling and Verbal Working Memory in Young Adults With a History of COVID-19: Cross-Sectional Study (e71303)
Hyunsoo Yoo. . . . 3297

Evaluating the Acceptance and Usability of an Independent, Noncommercial Search Engine for Medical Information: Cross-Sectional Questionnaire Study and User Behavior Tracking Analysis (e56941)
Lisa Specht, Raphael Scheible, Martin Boeker, Erik Farin-Glattacker, Nikolas Kampel, Marina Schmölz, Andrea Schöpf-Lazzarino, Stefan Schulz, Christian Schlett, Fabian Thomczyk, Sebastian Voigt-Radloff, Constanze Wegner, Katharina Wollmann, Andy Maun. . . . 3304

Telehealth Acceptance and Perceived Barriers Among Health Professionals: Pre-Post Evaluation of a Web-Based Telehealth Course (e74107)
Lena Rettinger, Lukas Maul, Peter Putz, Veronika Ertelt-Bach, Andreas Huber, Susanne Javorszky, Elisabeth Kupka-Klepsch, Sevan Sargis, Franz Werner, Klaus Widhalm, Rosmarie Joseph, Stefanie Doci, Sebastian Kuhn. . . . 3323

Multiple Large Language Models' Performance on the Chinese Medical Licensing Examination: Quantitative Comparative Study (e77978)
Yanyu Diao, Mengyuan Wu, Jingwen Xu, Yifeng Pan. . . . 3341

Practitioner Perspectives on the Uses of Generative AI Chatbots in Mental Health Care: Mixed Methods Study (e71065)
Jessie Goldie, Simon Dennis, Lyndsey Hipgrave, Amanda Coleman. . . . 3354

User Intent to Use DeepSeek for Health Care Purposes and Their Trust in the Large Language Model: Multinational Survey Study (e72867)
Avishek Choudhury, Yeganeh Shahsavar, Hamid Shamszare. . . . 3367

Prioritizing Trust in Podiatrists' Preference for AI in Supportive Roles Over Diagnostic Roles in Health Care: Qualitative Interview and Focus Group Study (e59010)
Mohammed Tahtali, Chris Snijders, Corné Dirne, Pascale Le Blanc. . . . 3382

Acceptability of a Conversational Agent–Led Digital Program for Anxiety: Mixed Methods Study of User Perspectives (e76377)
Pearla Papiernik, Sylwia Dzula, Marta Zimanyi, Edward Millgate, Malika Bouazzaoui, Jessica Buttimer, Graham Warren, Elisa Cooper, Ana Catarino, Shaun Mehew, Emily Marshall, Valentin Tablan, Andrew Blackwell, Clare Palmer. . . . 3401

Bridging Policy and Practice in Telemedicine Follow-Up Identification: Multicenter Mixed Methods Study in Beijing (e75964)
Yanli Lyu, Xi Li, Huixian Li, Mingyu Gu, Xuedong Xu, Yipei Wang, Changxiao Jin. . . . 3424

Dynamic Indicators of Adherence and Retention in Adults Using a Digital Mental Health App: Longitudinal Observational Analysis From the Brighten Study (e69464)
Dylan Hamitouche, Youcef Barkat, Deven Parekh, Eva Hammer, David Benrimoh. . . . 3442

XSL•FO
RenderX

Use and Acceptance of Innovative Digital Health Solutions Among Patients and Professionals: Survey Study (e60779)
Fritz Seidl, Florian Hinterwimmer, Ferdinand Vogt, Günther Edenharter, Karl Braun, Rüdiger von Eisenhart-Rothe, AG Digitalisierung der DGOU, Peter Biberthaler, Dominik Pförringer. . . . 3456

Public Attitudes and Predictors of Public Awareness of Personal Digital Health Data Sharing for Research: Cross-Sectional Study in Japan (e64192)
Yasue Fukuda, Koji Fukuda. . . . 3465

Patient Satisfaction With a Comprehensive Remote Care Program for Chronic Condition Management for Adults Under Medical Care: Observational Study (e69586)
Michael Merrill, Susan Zbikowski, Garrett Jenkins. . . . 3483

Predicting Engagement With Conversational Agents in Mental Health Therapy by Examining the Role of Epistemic Trust, Personality, and Fear of Intimacy: Cross-Sectional Web-Based Survey Study (e70698)
Fanny Guglielmucci, Daniela Di Basilio. . . . 3494

Weighing Costs and Benefits of Delay and the Acceptance of Two Decision Support Tools in Mental Health Care: Scoping Study Using Quantitative and Qualitative Data (e71678)
Sarah McKenna, Min Chong, Adam Poulsen, Ashlee Turner, Carla Gorban, Jacob Crouse, William Capon, Mathew Varidel, Melissa Aji, Elizabeth Scott, Ian Hickie, Frank Iorfino. . . . 3510

Understanding the Thoughts and Preferences for Technologies Designed to Detect Feelings of Loneliness: Interview Study Among Older Adults (e73694)
Jessica Rees, John Ratcliffe, Wei Liu, Yi Zhou, Sebastien Ourselin, Michela Antonelli, Ashay Patel, Yu Shi, Jingqi Liu, Anthea Tinker, Faith Matcham. . . . 3527

Perspectives of Older Adults on Assistive Technology: Qualitative Study (e74214)
Mirou Jaana, Maude Lévesque Ryan, Haitham Tamim, Edward Riachy, Guy Paré. . . . 3539

Factors Associated With the Intention to Use mHealth Among Thai Middle-Aged Adults and Older Adults: Cross-Sectional Study (e63607)
Nida Buawangpong, Wachiranun Sirikul, Penprapa Siviroj. . . . 3554

Assessing the Relationship Between the Type of Internet Use and Internet Addiction in Early and Middle Adolescents: Cross-Sectional Study From Qatar (e62955)
Khansa Chemnad, Maryam Aziz, Sanaa Al- Harahsheh, Azza Abdelmoneium, Ahmed Baghdady, Diana Alsayed Hassan, Raian Ali. . . . 3566

Employee Preference and Use of Employee Mental Health Programs: Mixed Methods Study (e65750)
Benedict Sevov, Robin Huettemann, Maximillian Zinner, Sven Meister, Leonard Fehring. . . . 3583

mHealth Use, Preferences, Barriers, and eHealth Literacy Among Patients With Inflammatory Bowel Disease: Survey Study (e64471)
Christopher Kretzschmar, Johannes Knitza, Robert Pietschner, Raja Atreya, Markus Neurath, Till Orlemann. . . . 3603

Exploring Perspectives of Health Care Professionals on AI in Palliative Care: Qualitative Interview Study (e79514)
Osamah Ahmad, Stephen Mason, Sarah Stanley, Amara Nwosu. . . . 3615

Organizational Readiness for Change in the Era of Smart Hospital Wards: Qualitative Study of Health Care Workers' Insights (e81932)
Hui Lim, Jennifer Sumner, Abigail Ang, Camille Keck, Emily Chew, Alexander Yip. . . . 3624

XSL•FO
RenderX

Association of Virtual Nurses' Workflow and Cognitive Fatigue During Inpatient Encounters: Cross-Sectional Study (e67111)
Saif Khairat, Jennifer Morelli, Wan-Ting Liao, Julia Aucoin, Barbara Edson, Cheryl Jones. . . . 3635

The Effects of Presenting AI Uncertainty Information on Pharmacists' Trust in Automated Pill Recognition Technology: Exploratory Mixed Subjects Study (e60273)
Jin Kim, Vincent Marshall, Brigid Rowell, Qiyuan Chen, Yifan Zheng, John Lee, Raed Kontar, Corey Lester, Xi Yang. . . . 3643

Longitudinal and Combined Smartwatch and Ecological Momentary Assessment in Racially Diverse Older Adults: Feasibility, Adherence, and Acceptability Study (e69952)
Sophia Holmqvist, Marina Kaplan, Riya Chaturvedi, Haochang Shou, Tania Giovannetti. . . . 3661

Impact of a Free Influenza Vaccination Policy on Older Adults in Zhejiang, China: Cross-Sectional Survey of Vaccination Willingness and Determinants (e73940)
Yusui Zhao, Jinhang Xu, Xuehai Zhang, Yue Xu, Xiaotong Yan, Shaofan Ren, Lixin Wang. . . . 3678

Needs and Expectations for the myNewWay Blended Digital and Face-to-Face Psychotherapy Model of Care for Depression and Anxiety (Part 2): Participatory Design Study including Mental Health Professionals (e68789)
Rosemaree Miller, Kathleen O'Moore, Katarina Kikas, Julie-Anne Matheson, Alexis Whitton, Peter Baldwin, Sophie Li, Melissa Black, Laura Kampel, Nicole Cockayne, Fiona Tuttlebee, Caitlin Fraser, Victoria Carr, Kathleen Varghese, Jill Newby. . . . 3688

Effects of Choice Set Sizes and Moderations of Anxiety and State Emotions on Mental Health Self-Care Uptake, Engagement, and User Experience: Experimental Study (e71165)
Siu Yeung, Florence Leung, Jason Lee, Gabriel Cheung, Ching Li, Winnie Mak. . . . 3709

Designing Digital Mental Health Support for Paramedics Exposed to Trauma: Qualitative Study of Lived Experiences and Design Preferences (e76158)
Nicola Cogan, Spence Whittaker, Ashleigh Craig, Lucy Milligan, Robyn McCluskey, Tara Burns, Alison Kirk, Susan Rasmussen, William Hodgson. . . . 3731

Theory-Based Mobile App Intervention to Promote Healthy Salt Intake Among Adults: Randomized Controlled Trial (e54174)
Milena Perin, Maria Gallani, Eduardo Nilson, Titilayo Agbadje, Marilia Cornélio. . . . 3745

Setting Goals and Accepting Challenges for Behavior Change—Analysis of Participants' Interactions With a Digital Multiple Health Behavior Intervention: Mixed Methods Study (e66208)
Katarina Åsberg, Marie Löf, Marcus Bendtsen. . . . 3760

A Mobile App Designed to Promote Shared Decision-Making in the Treatment of Psychotic Disorders: Feasibility and Acceptability Study (e68813)
Mari Skoge, Sofie Aminoff, Elizabeth Barrett, Gina Bryhni, Kristine Kling, Kari Kværner, Ingrid Melle, Erlend Mork, Carmen Simonsen, Linn Støme, Josina Vink, Tor Værnes, Kristin Romm. . . . 3775

Using Personalized Intervention Criteria in a Mobile Just-in-Time Adaptive Intervention for Increasing Physical Activity in University Students: Pilot Study (e66750)
Mai Ikegaya, Jerome Foo, Taiga Murata, Kenta Oshima, Jinhyuk Kim. . . . 3794

A Novel Just-in-Time Intervention for Promoting Safer Drinking Among College Students: App Testing Across 2 Independent Pre-Post Trials (e69873)
Philip Chow, Jessica Smith, Ravjot Saini, Christina Frederick, Connie Clark, Maxwell Ritterband, Jennifer Halbert, Kathryn Cheney, Katharine Daniel, Karen Ingersoll. . . . 3804

Comparison of an AI Chatbot With a Nurse Hotline in Reducing Anxiety and Depression Levels in the General Population: Pilot Randomized Controlled Trial (e65785)
Chen Chen, Kok Lam, Ka Yip, Hung So, Terry Lum, Ian Wong, Jason Yam, Celine Chui, Patrick Ip. . . . 3813

XSL•FO
RenderX

Development of Digital Strategies for Reducing Sedentary Behavior in a Hybrid Office Environment: Modified Delphi Study (e59405)
Iris Parés-Salomón, Cristina Vaqué-Crusellas, Alan Coffey, Bette Loef, Karin Proper, Anna Señé-Mir, Anna Puig-Ribera, Kieran Dowd, Judit Bort-Roig. . . . 3823

Innovative Digital Cognitive Behavioral Treatment for Insomnia Disorder in Adults (dCBT-i): Framework Development (e70193)
Emilie Stern, Severine Brune, Stéphane Mouchabac, Anthony Dubroc, Charles de la Personne, Pierre Geoffroy. . . . 3841

Needs and Expectations for the myNewWay Blended Digital and Face-to-Face Psychotherapy Model of Care for Depression and Anxiety (Part 1): Participatory Design Study including People with Lived and Living Experience (e69499)
Katarina Kikas, Kathleen O'Moore, Rosemaree Miller, Julie-Anne Matheson, Sophie Li, Kathleen Varghese, Peter Baldwin, Nicole Cockayne, Alexis Whitton, Jill Newby. . . . 3851

Associations Between Sleep Duration and Cognitive Function Among Older Adults: Cross-Sectional Study (e72886)
Guolin Guo, Yanling Jiang, Jianteng Dong, Xu Zhao, Xiaofan Lai, Xiumei Wang, Hongguo Rong, Jian Li. . . . 3874

Efficacy and Compliance of a Working Memory Multitasking Task Mobile Intervention for Children With Attention-Deficit/Hyperactivity Disorder: Single-Arm, Pre-Post Pilot Study (e70479)
Minyoung Jung, Jimin Woo, Sung Kim, You Lim, Younglae Kim, Dongwon Kang, Jung-sang Min, Bung-Nyun Kim. . . . 3886

Understanding the Differences Between Online and Offline Mental Health Help Seekers: Cross-Sectional Comparative Study (e69305)
Mohamed Adwi, Bahaa Mahmoud, Noha Amer, Roa Alamrawy, Ismail Sadek, Mohamed Elsheikh. . . . 3900

Reviews

Representation of Rural Older Adults in AI for Health Research: Systematic Literature Review (e70057)
Kristina Shiroma, Jacqueline Miller. . . . 389

User and Provider Experiences With Health Education Chatbots: Qualitative Systematic Review (e60205)
yung-Eun Choi, Sebastian Fitzek. . . . 399

User Experience of Older Adults With an Age-Friendly Transportation Planning E-Tool: Scoping Review (e63273)
Sara Bahrampoor Givi, Mireille Gagnon-Roy, Hélène Pigot, Véronique Provencher. . . . 413

Effectiveness of Participatory Ergonomic Interventions on Work-Related Musculoskeletal Disorders, Sick Absenteeism, and Work Performance Among Nurses: Systematic Review (e68522)
Guganesan Krishnanmoorthy, Sanjay Rampal, Sumitra Karuthan, Faiz Baharudin, Rama Krishna. . . . 431

Personalized Interactive Music Systems for Physical Activity and Exercise: Exploratory Systematic Review and Meta-Analysis (e70372)
Andrew Danso, Tiia Kekäläinen, Friederike Koehler, Keegan Knittle, Patti Nijhuis, Iballa Burunat, Pedro Neto, Anastasios Mavrolampados, William Randall, Niels Hansen, Alessandro Ansani, Timo Rantalainen, Vinoo Alluri, Martin Hartmann, Rebecca Schaefer, Johanna Ihalainen, Rebekah Rousi, Kat Agres, Jennifer MacRitchie, Petri Toiviainen, Suvi Saarikallio, Sebastien Chastin, Geoff Luck. . . . 455

Driving and Restraining Forces in the Implementation of Information Systems in the Public Sector: Scoping Review (e71575)
Arja Lemmettylä, Ulla-Mari Kinnunen. . . . 481

XSL•FO
RenderX

Tailoring for Health Literacy in the Design and Development of eHealth Interventions: Systematic Review (e76172)
Jan Hovingh, Cynthia Elderson-van Duin, Derek Kuipers, Yanda van Rood, Geke Ludden, Denise Hanssen, Judith Rosmalen. 493

Framing the Human-Centered Artificial Intelligence Concepts and Methods: Scoping Review (e67350)
Roberta Bevilacqua, Tania Bailoni, Elvira Maranesi, Giulio Amabili, Federico Barbarossa, Marta Ponzano, Michele Virgolesi, Teresa Rea, Maddalena Illario, Enrico Piras, Matteo Lenge, Elisa Barbi, Garifallia Sakellariou. 511

On the Necessity of Multidisciplinarity in the Development of at-Home Health Monitoring Platforms for Older Adults: Systematic Review (e59458)
Chris Lochhead, Robert Fisher. 524

Challenges to Rehabilitation Services in Sub-Saharan Africa From a User, Health System, and Service Provider Perspective: Scoping Review (e58841)
Callixte Cyuzuzo, Marie Dukuzimana, Clement Muhire, Mathew Sheldon Ames, Emmanuel Ngwakongnwi. 538

AI Chatbots for Psychological Health for Health Professionals: Scoping Review (e67682)
Gumhee Baek, Chiyoung Cha, Jin-Hui Han. 545

Optimizing Clinical Decision Support System Functionality by Leveraging Specific Human-Computer Interaction Elements: Insights From a Systematic Review (e69333)
Ali Azadi, Francisco García-Peñalvo. 564

Advancing Regional and Remote Health Care With Virtual Hospital Implementation: Rapid Review (e64582)
Artika Kumari, Tafheem Wani, Michael Liem, James Boyd, Urooj Khan. 588

Toward Human-Centered Artificial Intelligence for Users' Digital Well-Being: Systematic Review, Synthesis, and Future Directions (e69533)
Youngsoo Shin. 609

Endometriosis Support and Development of Digital Technology–Based Interventions: Systematic Review (e71859)
Tivizio Pavic, Kévin Nadarajah, Alain Somat, Geneviève Cabagno, Florence Terrade. 629

Legitimacy as Social Infrastructure: A Critical Interpretive Synthesis of the Literature on Legitimacy in Health and Technology (e48955)
Sydney Howe, Carin Uyl-de Groot, Rik Wehrens. 1767

Viewpoints

Auxiliary Teaching and Student Evaluation Methods Based on Facial Expression Recognition in Medical Education (e72838)
Xueling Zhu, Roben Juanatas. 939

Human-Centered Design and Digital Transformation of Mental Health Services (e66040)
William Fleming, Adam Coutts, Diane Pochard, Daksha Trivedi, Kristy Sanderson. 949

XSL•FO
RenderX

Original paper

Screening Workers for Occupational Exposure to Respirable Crystalline Silica: Development and Usability of an Electronic Data Capture Tool (e64111)
Fiona Hore-Lacy, Christina Dimitriadis, Ryan Hoy, Javier Jimenez-Martin, Malcolm Sim, Jane Fisher, Deborah Glass, Karen Walker-Bone. . . 2 2 9 5

Corrigenda and Addendas

Correction: Patient Experiences With a Mobile Self-Care Solution for Low-Complex Orthopedic Injuries: Mixed Methods Study (e75812)
Jelle Spierings, Gijs Willinge, Marike Kokke, Sjoerd Repping, Wendela de Lange, Thijs Geerdink, Ruben van Veen, Detlef van der Velde, J Goslings, Bas Twigt, Collaboration Group. 2926

Correction: Requirement Analysis for Data-Driven Electroencephalography Seizure Monitoring Software to Enhance Quality and Decision Making in Digital Care Pathways for Epilepsy: A Feasibility Study from the Perspectives of Health Care Professionals (e79484)
Pantea Keikhosrokiani, Johanna Annunen, Jonna Komulainen-Ebrahim, Jukka Kortelainen, Mika Kallio, Päivi Vieira, Minna Isomursu, Johanna Uusimaa. 2927

Correction: Identifying Contextual Factors and Strategies for Practice Facilitation in Primary Care Quality Improvement Using an Informatics-Driven Model: Framework Development and Mixed Methods Case Study (e84390)
. 3352

Letter to the Editor

Enhancing Methodological Rigor in Mobile Health Care Research (e68243)
Shuhan Tang. 3295

Research Letter

Older Adults' Experiences With an Online Survey (e65684)
Kristie Weir, Yehya Maitah, Sarah Vordenberg. 3419

XSL•FO
RenderX

Patient-Centered Televisit for Chronic Obstructive Pulmonary Disease Discharge Transitions: User-Centered Design Study

Joanna Abraham[1], PhD; Alicia Meng[1], MA; Nicolas Caravelli[2], MA; Leah Traeger[2], MHA; May Nguyen[2], MA; Vineet Arora[2], MD; Valerie G Press[2], MD

[1]Department of Anesthesiology, School of Medicine, Washington University in St. Louis, 660 S Euclid Ave, St. Louis, MO, United States

[2]Department of Medicine, University of Chicago, Chicago, IL, United States

Corresponding Author:
Joanna Abraham, PhD
Department of Anesthesiology, School of Medicine, Washington University in St. Louis, 660 S Euclid Ave, St. Louis, MO, United States

Abstract

Background: Chronic obstructive pulmonary disease (COPD) affects approximately 16 million Americans and often results in avoidable readmissions due, in part, to medication errors and lack of education. Telehealth interventions can support medication reconciliation and inhaler education following hospital discharge for patients with COPD.

Objective: This study aimed to design and prototype TELE-TOC (Telehealth Education: Leveraging Electronic Transitions of Care), a post-discharge, in-home, televisit intervention, and to map its workflow to ensure integration into the routine discharge care transition process for patients with COPD.

Methods: A user-centered design approach across 3 phases was followed to develop and prototype TELE-TOC. Participants included adult patients hospitalized for COPD exacerbations, their caregivers, clinicians involved in COPD care, and organizational leaders. Data collection methods included semi-structured interviews, system usability scale surveys, and cognitive walkthroughs of the TELE-TOC prototype to assess participants' perceptions on usability and feasibility of TELE-TOC implementation as part of routine COPD discharge care transitions. Qualitative data were analyzed using inductive thematic analysis and an inductive-deductive approach guided by the Agency for Healthcare Research and Quality–endorsed Care Transitions Framework. Quantitative data were summarized using basic descriptive statistics.

Results: Participants included 18 patients, 18 clinicians, 8 organizational leaders, and 2 caregivers. Phase 1 identified 3 interdependent stages of COPD hospital-to-home discharge: inpatient pre-discharge, at-home post-discharge, and outpatient clinic visit post-discharge. Key facilitators of discharge care transitions included the hospital's "meds-to-beds" program and high patient health literacy, while barriers to discharge included poor timing of education and conflicting patient priorities. Phase 2 delineated the core televisit components (eg, dedicated clinician, medication reconciliation, inhaler use, and self-management education) and flexible components (eg, reminder system and session frequency). Potential implementation enablers included multiple techniques for clinicians to access and support patient education and backup communication strategies in the event of technical issues. Potential implementation barriers included insufficient patient technology access and limited technology and health literacy, as well as limited clinician bandwidth for thorough COPD education and medication reconciliation. Phase 3 TELE-TOC prototype walkthroughs demonstrated a positive patient experience (average system usability scale score of 97.5/100), attributed to the benefits of videoconferencing technology for hands-on teaching and the use of the virtual teach-back method. Identified barriers included varying levels of patient technology literacy, insufficient inhaler education, limited patient understanding of medication lists, and clinician uncertainty around TELE-TOC documentation. Suggestions for mitigating these barriers included patient training for TELE-TOC sessions, amendments to pharmacists' "visit note," and enhanced patient preparation for medication reconciliation.

Conclusions: Using a co-design approach, we integrated multiple perspectives to develop and optimize TELE-TOC, a patient-centered televisit intervention aimed at supporting discharge care transitions to improve continuity of care and outcomes for patients with COPD. Future research will evaluate the impact of TELE-TOC on readmissions from acute exacerbations.

(JMIR Hum Factors 2025;12:e77953) doi:10.2196/77953

KEYWORDS

virtual; remote; chronic obstructive pulmonary disease; readmission; telehealth; telemedicine; participatory design; education

XSL•FO
RenderX

Introduction

Chronic obstructive pulmonary disease (COPD) affects over 16 million Americans [1]. A leading cause of hospital readmissions [2] and mortality [3], COPD is projected to cost American health care systems over US $60 billion by 2029 [4]. To reduce costs related to preventable readmissions after an index admission for COPD [5], the Centers for Medicare and Medicaid Services included COPD as part of the Hospital Readmissions Reduction Program (HRRP) financial penalty to reduce hospital readmissions and their associated costs [6,7].

Several factors contribute to readmissions, including inadequate patient self-management of COPD and suboptimal care following a hospital discharge. These factors are often attributed to poor implementation of COPD care guidelines among clinicians during discharge [8], fragmented and inadequate discharge education on strategies to improve skills for COPD self-management at home [9,10], lack of follow-up with patients [11], and insufficient understanding about proper use of medications, including inhaler techniques [12-14]. To address these gaps in medication and symptom self-management, care transition interventions have been explored by many, including our team [15,16]. For instance, pharmacist-led medication reconciliation programs at discharge have been shown to reduce medication discrepancies and 30-day readmission rates for older patients with COPD [17]. Similarly, inhaler technique education using a "Teach-To-Goal" (TTG) approach, consisting of rounds of assessment and education prior to discharge, was shown to reduce all-cause readmissions after initial admissions for COPD compared to simply providing verbal instruction [13,16].

While promising, both medication reconciliation and TTG at discharge can be resource-intensive and challenging to implement at large scale within existing hospital workflows [16,17]. Moreover, although in-hospital pre-discharge interventions may improve patient self-management and potentially reduce readmissions, medication errors still occur in more than half of patients after hospital discharge [18]. Inpatient-only care is likely insufficient for many patients, and researchers recommend that interventions (eg, medication reconciliation and education) extend post-discharge for sustaining medication and COPD self-management skills [19].

To address sustainability and effectiveness concerns, our team helped to develop a pharmacist-led, at-home telehealth program, which has improved inhaler technique and quality of life among patients with COPD, while reducing symptoms [20,21]. Other research has also shown that post-discharge televisit medication reconciliation reduces medication errors across populations [22,23]. The virtual approach for medication review and reconciliation and teaching inhaler technique could enhance COPD education, medication management, and care quality without increasing costs. Informed by published evidence [20,21,24], we developed an end user–informed, comprehensive, post-discharge televisit service termed TELE-TOC (Telehealth Education: Leveraging Electronic Transitions of Care) to facilitate inhaler education and reinforcement of COPD self-management skills and provide opportunities for medication reconciliation and counseling by clinical pharmacists. Here we report on the development and prototyping of the TELE-TOC intervention and identify the components and implementation process necessary to fit TELE-TOC into our current COPD HRRP and DTOC (Discharge Transitions of Care) workflow.

Methods

Study Setting

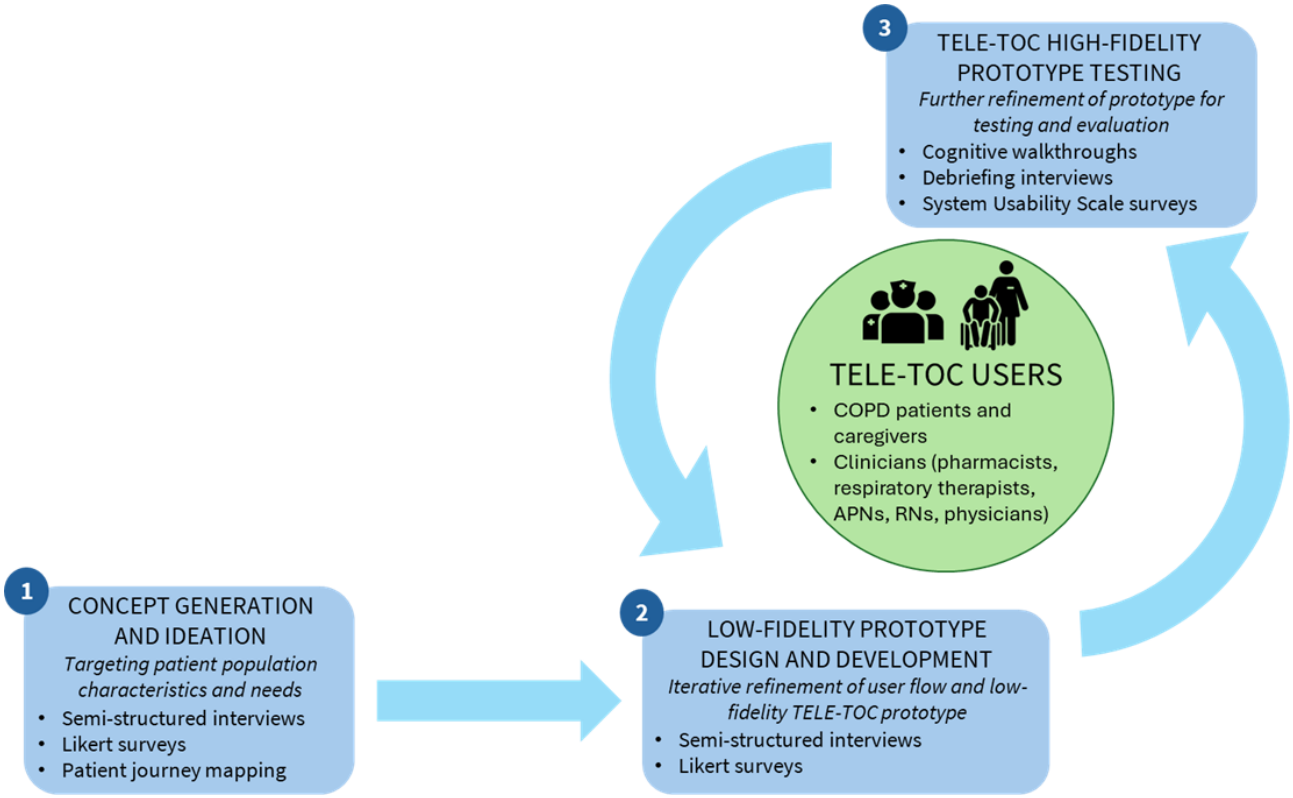
The study was conducted at the University of Chicago Medicine (UCM). We are developing the TELE-TOC intervention for integration into our existing COPD HRRP [25].

Study Approach

We followed a user-centered design (UCD) approach to ensure TELE-TOC met the needs, preferences, and behaviors of end users (eg, patients, clinicians, and organizational leaders [OLs]) while considering usability and accessibility (eg, older age, limited health, and technology literacy) [26,27]. Multiple end users were actively engaged in the development of TELE-TOC across 3 iterative phases (Figure 1). Phase 1 mapped the current hospital-home-outpatient care transition workflow for patients with COPD, with emphasis on perceived workflow facilitators and barriers from multiple perspectives (ie, patient, caregiver, clinician, and OLs). Data-informed strategies were then developed to address care transition barriers and inform the TELE-TOC role within the current workflow. Phase 2 identified design ideas and implementation considerations using a low-fidelity TELE-TOC prototype with emphasis on potential facilitators and barriers to adoption and implementation. Phase 3 focused on evaluating the usability and feasibility of a high-fidelity TELE-TOC prototype.

XSL•FO
RenderX

Figure 1. User-centered design phases for Telehealth Education: Leveraging Electronic Transitions Of Care development. APN: advanced practice nurse; COPD: chronic obstructive pulmonary disease; RN: registered nurse; TELE-TOC: Telehealth Education: Leveraging Electronic Transitions of Care.



Participants

Phases 1 and 2: (1) adult patients admitted for COPD exacerbations enrolled in our COPD HRRP, (2) clinicians (pharmacists, advanced practice nurses [APNs], hospitalists, primary care physicians, pulmonologists, and respiratory therapists), and (3) OLs (patient safety and quality improvement hospital leaders). Phase 1 also included patient caregivers. Phase 3: TELE-TOC end users (adult patients and inpatient clinical pharmacists). While we planned to include caregivers in all phases, they were ultimately only included in Phase 1 due to challenges with further caregiver participation (eg, patient discomfort with caregiver involvement, lack of caregiver availability, passive caregiver, and lack of caregiver altogether).

Data Collection

We used semi-structured interviews to gather participant perspectives on current discharge care transitions. Interview guides were iteratively developed and refined with input from pilot interviews with 2 patients and caregivers from the COPD Foundation, a clinician involved in caring for patients with COPD, and an OL.

Phase 1 mapped current DTOC workflows, ascertained facilitators and barriers, and delineated potential for televisit interventions. Phase 2 assessed attitudes toward the proposed televisit intervention, a low-fidelity prototype of PowerPoint slides with process maps and instructions for inhaler education and identified potential facilitators and barriers to implementation. We asked participants to rate their experiences with medication reconciliation and COPD medication education using a Likert scale from 1 to 5, where 1 indicated “not prepared” or “not good” and 5 indicated “prepared” or “very good.” We requested patient feedback on preferred televisit format (phone, text, or video), comfort levels with video technology, experiences with technical difficulties, and confidence in internet usability. Phase 3 comprised televisit walkthrough usability sessions with 4 patients post-discharge from UCM and 2 clinical pharmacists using the TELE-TOC prototype, as previous research demonstrated that only 3 - 5 sessions were required to identify most usability issues [28]. High-fidelity TELE-TOC prototypes consisted of interactive web-based sessions integrated into MyChart (Epic Systems). Additionally, pharmacists piloted a TELE-TOC–specific templated “visit note” in the electronic health record (EHR) to document medication reconciliation and inhaler education (Figure 2). Participants verbalized their actions during the TELE-TOC process. After medication review and inhaler education, patients rated TELE-TOC usability on the system usability scale (SUS) [29] and completed debriefing interviews to assess user satisfaction and suggestions for improvement. Walkthroughs were conducted and recorded via a HIPAA (Health Insurance Portability and Accountability Act)–compliant Zoom platform (Zoom Video Communications Inc) (refer to guides in Multimedia Appendix 1).

XSL•FO
RenderX

Figure 2. Screenshot of electronic health record–embedded “visit note.” COPD: chronic obstructive pulmonary disease.

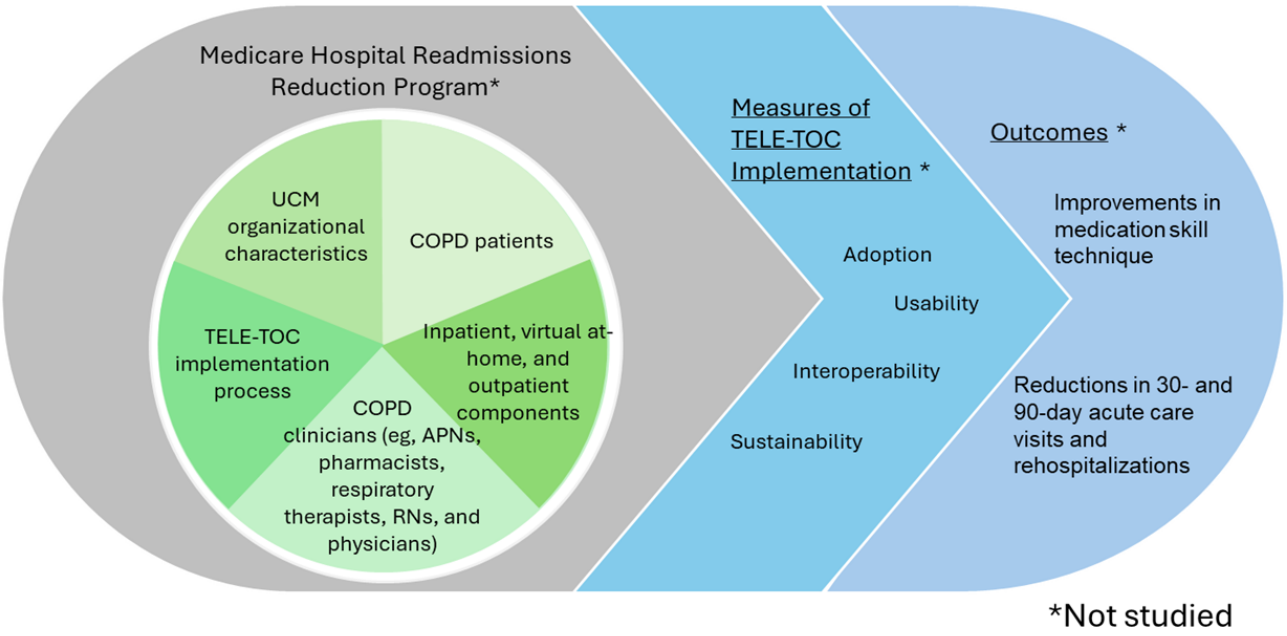
XSL•FO
RenderX

Data Coding and Analysis

We followed an inductive (data-driven) thematic analysis approach for Phase 1 [30]. Two researchers (JA and EZ) independently read the data transcripts multiple times and openly coded the data with keywords representing the data, then identified relationships among codes, highlighting areas of similarity and overlap to form subthemes; repeated patterns among established subthemes identified overarching themes. Additionally, interview responses informed the development of a process map for care transitions of patients with COPD.

In Phase 2, we used an inductive-deductive thematic analysis approach to integrate insights on current discharge care transition processes with elements of an evidence-based AHRQ-endorsed care transition framework and implementation science guidelines, informing the design and implementation of our intervention [31]. This framework offers relevant guidelines for addressing challenges in discharge care transition processes and identifies contextual determinants that can inform intervention design and implementation (Figure 3). We used 5 of the 8 AHRQ-endorsed Care Transition Framework domains: intervention characteristics, organizational characteristics, characteristics and roles of clinicians, characteristics and roles of patients, and the process of implementation (Table 1).

Figure 3. Domains of the Agency for Healthcare Research and Quality–endorsed care transitions framework. APN: advanced practice nurse; COPD: chronic obstructive pulmonary disease; RN: registered nurse; TELE-TOC: Telehealth Education: Leveraging Electronic Transitions of Care; UCM: University of Chicago Medicine.



XSL•FO
RenderX

Table . Agency for Healthcare Research and Quality–endorsed care transitions framework for the Telehealth Education: Leveraging Electronic Transitions of Care context.

Domain	Definition	Domain contextualized for TELE-TOC[a]
Intervention characteristics	Characteristics and features of the intervention that impact implementation success	TELE-TOC characteristics (TELE-TOC visit scheduling, set-up of the reminders, TELE-TOC visit content, and TELE-TOC documentation)
Organizational characteristics	Characteristics of organizations involved in the intervention, including structural characteristics, networks and communications, culture, climate, and readiness that influence implementation	UCM[b] characteristics (insurance coverage, funding and sustainability, and ownership of TELE-TOC)
Characteristics and roles of clinicians	Attributes (eg, type of clinician, affiliations, and duties) of individuals who are engaged in the provision of care or treatment	Clinicians (medical assistant, patient service representative, pharmacist, APNs[c], RNs[d], etc) and their roles in TELE-TOC
Characteristics and roles of patients	Attributes (eg, type of patients) of individuals who are the recipients of care or treatment in the given intervention setting	Adult, low-acuity patients with COPD[e] enrolled in HRRP[f]
Process of implementation	Processes including planning, engaging, and reflecting to achieve the intended use of the intervention at both individual and organizational levels	Implementation strategies for TELE-TOC

[a]TELE-TOC: Telehealth Education: Leveraging Electronic Transitions of Care.

[b]UCM: University of Chicago Medicine.

[c]APN: advanced practice nurse.

[d]RN: registered nurse.

[e]COPD: chronic obstructive pulmonary disease.

[f]HRRP: Hospital Readmissions Reduction Program.

First, we (JA and LP) followed an inductive approach consistent with our Phase 1 analysis to generate open codes about the current discharge care transition process and contextual determinants for intervention design and implementation. We then applied the constructs underlying the AHRQ-endorsed Care Transitions Framework (CTF) [32] as a structured coding template to guide deductive thematic analysis. This step helped us to align and integrate our open codes with the constructs within each of the 5 domains in the CTF as outlined in Table 1 [33]. Following this, we then formed overarching themes, or patterns across the data that related to constructs within each care transition domain, highlighting the design and implementation needs for TELE-TOC.

In Phase 3, we (JA and AM) analyzed the cognitive walkthrough transcripts using an inductive thematic analysis approach. Usability, TELE-TOC ease-of-use, and televisit effectiveness were coded and categorized into themes and subthemes (eg, barriers and facilitators). In addition, debriefing interviews were thematically analyzed to identify TELE-TOC user experiences and suggestions for improvement.

Phases 1 and 2 surveys were summarized using basic descriptive statistics (mean). Phase 3 SUS surveys had 10 questions rated from 1 to 5. In accordance with recommendations and standard practice for analysis [29,34], we subtracted 1 from each value for odd-numbered questions sharing positive statements, and we subtracted each value from 5 for even-numbered questions sharing negative statements. After summing the new values across questions, we multiplied this final value by 2.5 and then found the average score across participants.

To ensure validity and reliability, we followed best methodological practices, including member checking and team debriefing [35]. Two researchers independently coded all data, and any inconsistencies were resolved through team discussions to reach a 100% consensus.

Ethical Considerations

This study was approved by the University of Chicago Institutional Review Board (IRB21-0947). All participants were informed of the study purpose, procedures, risks, and rights. Patients and caregivers provided written consent, and clinicians and OLs provided oral consent after reviewing the consent form with study staff. Trained personnel conducted all activities using standardized protocols, with strict confidentiality safeguards to protect participant privacy. Paper records were stored in a locked office, and electronic data were encrypted, password-protected, and housed on secure institutional servers. Hospital data were collected and stored in REDCap (Research Electronic Data Capture; Vanderbilt University), and interview data were maintained on a secure, HIPAA-compliant shared drive. Virtual sessions used the HIPAA-compliant Zoom platform with full security features enabled. Study identification numbers were used in place of personal identifiers for all analytic files. Any identifiable contact information used for scheduling or compensating participants was stored separately from study data and will be destroyed, along with audio recordings, at the conclusion of the study.

Participants in the semi-structured interviews were compensated for their time. Patients and caregivers received a US $50 gift card for each study phase they completed, and clinicians and

XSL•FO
RenderX

organizational leaders received a US $15 gift card for each phase in which they participated.

Results

Study Overview

Forty-five participants were enrolled across Phases 1, 2, and 3 (Table 2). Findings are presented across the following themes: Phase 1: current DTOC workflow for patients with COPD enrolled in HRRP; Phases 2 and 3: considerations for TELE-TOC design and implementation.

Table . Participant numbers and characteristics by phase.

Participants	Phase 1	Phase 2	Phase 3
Clinicians	12	4	2
Pharmacist	2	2	2
Respiratory therapist	3	2	0
Advanced practice nurse	3	0	0
Nurse	1	0	0
Physician	3	0	0
Organizational leaders	5	3	0
Caregivers	2	0	0
Patients	8	6	4
Age (years), mean (SD)	63.3 (9.7)	65 (11.3)	56.3 (11.0)
Race			
Black	8	6	4
White	0	0	0
Other	0	0	0
Non-Hispanic	8	5	4
Sex			
Female	4	2	3
Male	4	4	1
Other	0	0	0

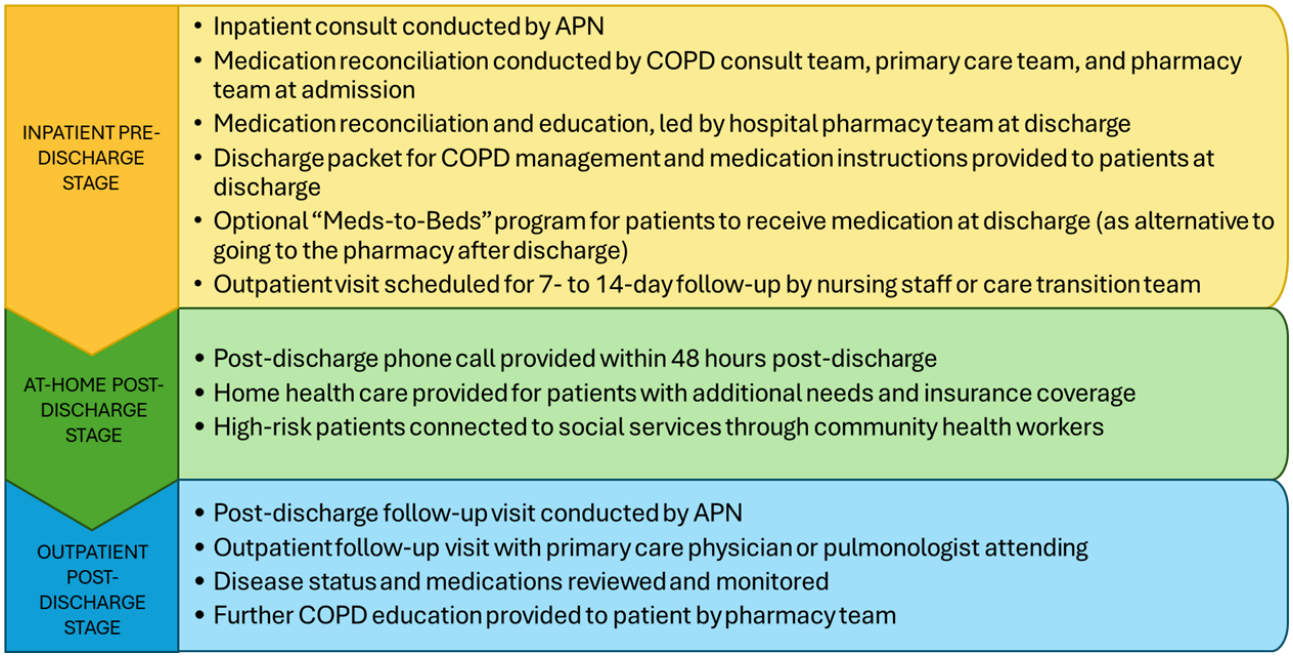
The mean (SD) patient age across phases ranged from 56.3 (11.0) to 65 (11.3) years (Table 2). Most patients were identified as Black and non-Hispanic, with no representation from White or other racial groups. Sex distribution was evenly split, with 9 female and 9 male patients overall.

Phase 1: Current DTOC Workflow for Patients With COPD

The current inpatient-outpatient discharge transition of care workflow for acute exacerbations of COPD consists of 3 interdependent stages: (1) inpatient pre-discharge (hospital admission to discharge to home), (2) at-home post-discharge (home to outpatient follow-up), and (3) outpatient visit - (Figure 4).

XSL•FO
RenderX

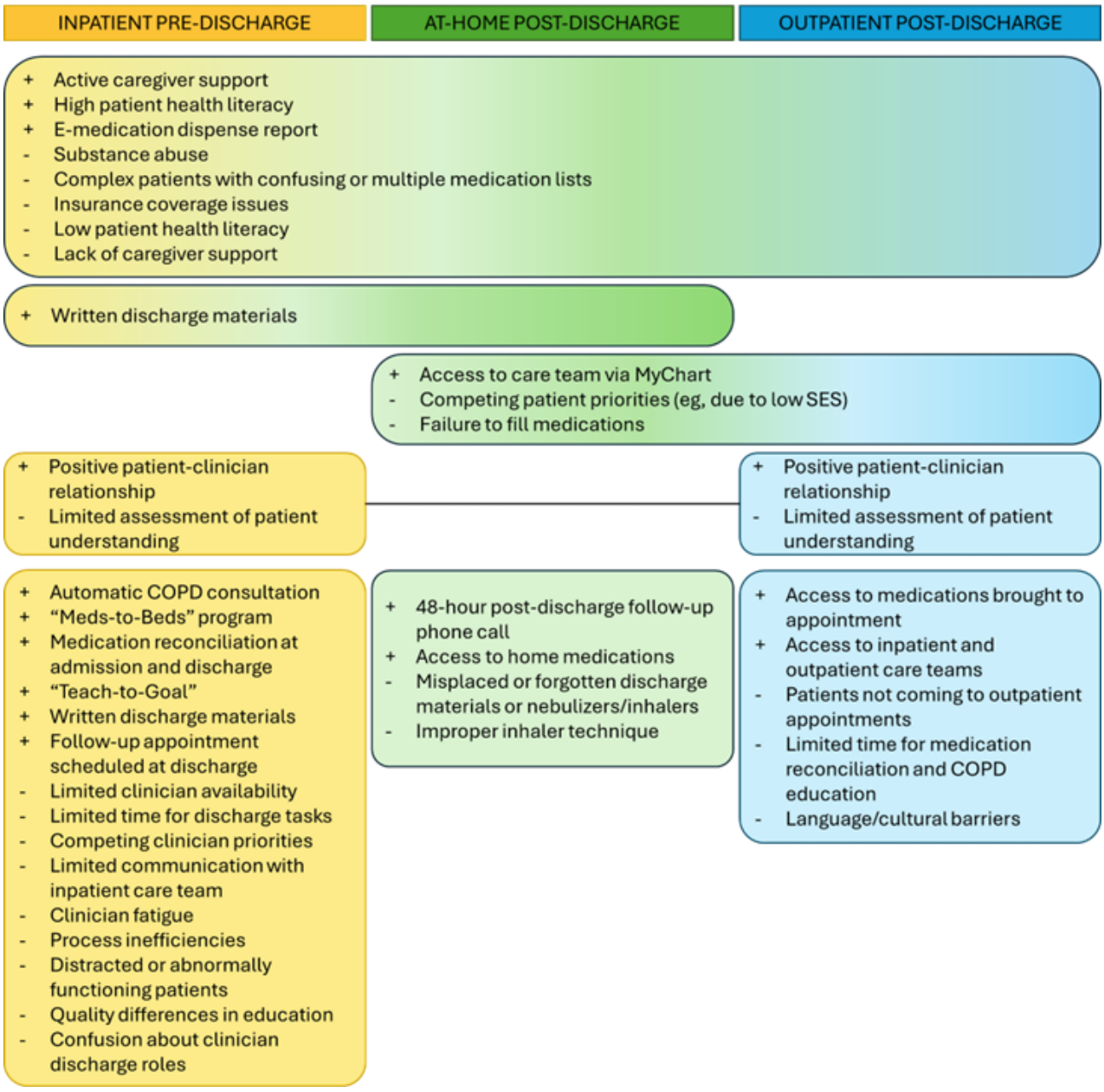
Figure 4. The interdependent stages of the Discharge Transitions of Care workflow. APN: advanced practice nurse; COPD: chronic obstructive pulmonary disease.



All participants emphasized that each stage relied on the successful execution of activities in the preceding stage, given the interdependence in the workflow. We identified additional key facilitators and barriers to the current DTOC workflow across these stages (Table 3)—with potential impacts on safety and care continuity outcomes (Figure 5).

XSL•FO
RenderX

Figure 5. Facilitators and barriers to standard inpatient to outpatient Discharge Transitions of Care workflow. COPD: chronic obstructive pulmonary disease; SES: socioeconomic status.



XSL•FO
RenderX

Table . Phase 1 topics with representative quotes. (+) Observed facilitators to current workflow; (–) observed barriers to current workflow.

Stage and topic	Quote
Inpatient pre-discharge	
(+) Medication reconciliation conducted with inpatient team	• Q1. "[The nurses] always keep contact. Because he doesn't have—he's not aware of a lot of things, and we always keep everything on my phone, his charts and everything. He always tells them, 'Ask my wife. I don't know anything.'" [CG1[a]] • Q2. "We try...talking to caregivers and involving caregivers on that discharge med reconciliation process." [C5[b]] • Q3. "I think that one of the tools that pharmacists sometimes use for medication reconciliation...It's a button called the dispense report in Epic, which can show you what patients have picked up at local pharmacies." [C11]
(+) Efficient meds-to-beds program	• Q4. "Meds-to-beds was the big thing, just because pharmacies are super short right now to make sure that they got their inhalers before they are able to go home, or a bronchodilator or steroids, anything really crucial for a COPDer." [C10]
(+) Teach-to-goal COPD[c] education	• Q5. "Our pharmacists will make sure everything's clinically relevant...[The prescription] goes over to a technician who enters it, fills it, gets it ready. Before we deliver it to the patient, my pharmacist will do a med education over the phone with the patient. So, they call the patient, do education. Once the patient confirms they understand what they're taking, what they're receiving, the technician will take it and deliver it to the room. And then the patient is discharged home." [OL3[d]]
(+) Teach-back of inhaler education	• Q6. "Really, the first step is just to have them try to explain it to me...what they're doing. And not only what they're doing, we try to get them to verbalize why they think they're doing it and connect those dots. So that's the primary mechanism." [C2]
(+) Building rapport with patients	• Q7. "Part of our needs assessment when they come in on admission is asking them about their history of smoking, how many packs, and for how many years. Sometimes they won't be as honest with you as others, that's why you try to establish rapport with them first so they kind of open up with you a little bit better." [C10]
(+) High patient health literacy	• Q8. "What I like a lot is when...the patient's...very well-educated, and they take care of their own health...I'll go over the medications. They're like, 'Yup, I know what this is for. I take it at this time.' And they verbally understand, and they can kind of repeat back to me, what's going on. And they already know that their scripts are ready at home as well." [C10]
(–) Discharge delays	• Q9."On the day that I'm supposed to be released, it was either you're waiting for the doctors to do the final paperwork, and the nurses having to bring it together to even the pharmacy gathering up the medicines that you'll be taking home with you to carry you until you get it going with your primary care and your pharmacist. So those were things that you just simply have to wait. And if one couldn't do something without—the pharmacist [can't] do their thing until the doctor does theirs. And the nurse can't do their thing until the doctor does their thing." [P4[e]]
(–) Insurance coverage issues	• Q10. "I would say 50% of drugs that we receive require prior authorization. If it's a prior authorization and we can't get it approved in time or we can't get it switched to a different therapy, we tell the patient to go home, and then we'll work on the prior authorization." [OL3]

XSL•FO
RenderX

Stage and topic	Quote
(–) Absent or incomplete medication reconciliation	• Q11. “The information contained (in the expense report) is limited because sometimes people will pay cash rather than use insurance or may obtain their inhalers through different sample programs or free drug programs if they haven’t been affordable to them. So, I think the incompleteness of records that people are going into those scenarios with can be a barrier and inaccuracy in existing records, and that temptation to just confirm what’s already there can be a barrier.” [C11]
(–) Low patient health literacy	• Q12. “Unfortunately, the great majority of our patients don’t know their meds. And so, trying to speak to them about what is accurate and what might not be accurate on their medication list is a challenge.” [C2] • Q13. “I think it really depends on the patient’s health literacy and level of understanding and...we sometimes are able to get pharmacy teaching, but we don’t always get all of those patients. So, I think they’re not always as equipped as we would like them to be [for discharge].” [C5]
(–) Poor timing of COPD education	• Q14. “When a patient is getting discharged, all they think about is going home. So, we counsel them at bedside, and we give them their meds, and we teach them how to use an inhaler or whatever medication they’re receiving. But I think all they’re thinking about is going home.” [OL2]
(–) Loss of patient follow-up	• Q15. "But here, once a patient is discharged, the inpatient team is 100% done with that patient. There is no follow-up.” [C2]
(–) Varied quality of education	• Q16. “For patients that are started on biologic therapies that are hopefully helping them with their asthma/COPD, I think there’s sometimes confusion about whether or not they’re supposed to continue first-line therapy.” [C11]
(–) Limited knowledge of patient medications	• Q17. “I think a lot of it is done by pharmacy or nursing. The physician teams are not sitting and going through medications with patients for the most part.” [OL1] • Q18. “I’d say mostly just because it’s not often done by a pharmacist, and it’s done by another person who works in the institution...They don’t look to see what [the patient has] been taking, the dosage and frequencies are wrong. Routes are often wrong as well, so maybe due to lack of knowledge of some of the medications and how they’re supposed to be taken.” [C4]
(–) Time barriers and competing priorities	• Q19. “The other times when med reconciliations kind of fall apart are if they get admitted in very severe illness in the MICU, they’re focused on stabilizing the patient rather than on the med reconciliation, and so sometimes...it’s not prioritized as much as it could be...” [C5]
At-home post-discharge	

XSL•FO
RenderX

Stage and topic	Quote
(+) Active post-discharge communication	• Q20. "We do have a nurse that is part of the team that does make a 48-hour phone call. And so that's another mechanism where—if the patient actually answers the phone and is willing to engage in that conversation, that's another opportunity for the patient to voice like, 'Oh, I didn't get this inhaler.' I think she maybe reaches the patient 50 to 60 percent of the time." [C2] • Q21. "And then we also definitely found that there had to be an education component before the patient leaves the hospital, of course...I think because I think that one round of education might not be enough. And again, things might seem okay after you meet with a patient or a family, and they might feel okay about it, but then they have questions later or they forgot something." [OL4]
(+) Use of MyChart portal to facilitate communication	• Q22. "You see all them doctors I got? [sharing MyChart screen] I use [MyChart] for everything. Doctor, pharmacy, everything." [P3]
(+) Caregiver support	• Q23. "I'm thinking about one patient in particular with COPD that—and this is where the family caregivers come into play, when I need to talk to her, it's actually through her daughter." [C12]
(+) High patient health literacy	• Q24. "...the papers given explaining about different things concerning my treatment and the issues that I have, they were pretty thorough." [P4]
(–) Low patient health literacy	• Q25. "Someone who has a low literacy, you can explain it to them, but they still wouldn't get it." [C8] • Q26. "We do want our patients to be involved as much as they can be with their plan of care...But it is a lot for them at times...Sometimes we'll have charts of different inhalers, because they'll be trying to explain to us, 'Oh, I lost this one.'" [C4]
(–) Cost-related nonadherence to COPD management	• Q27. "Unfortunately, again, it's this demographic where there is such a high amount of poor health literacy, competing priorities with patients, housing instability, food insecurity, all of those things. So taking all [medications] of this in is just not on the top of their priority list." [C1] • Q28. "I think in terms of barriers...they may have different social situations or socioeconomic factors that kind of contribute to them being able to even take care of themselves when it comes to their health." [C5]
(–) Comorbidities impacting patient health	• Q29. "Quite a few of our [patients]...with COPD are still heavy smokers, and a lot of it's like, 'Yeah, I still smoke,' even though they're on oxygen. And so that's something else you also have to educate them on, and they're like, 'Well, I've been doing this for years. I'm okay.'" [C9]
Outpatient post-discharge	

XSL•FO
RenderX

Stage and topic	Quote
(+) Medication review in clinic	• Q30. “[In my] follow-up visit with primary care...[We go over] full medication review, how I’m feeling, [ask about] medicine I’m taking and everything.” [P7] • Q31. “But sometimes people aren’t really sure which medications they are taking. I mean, the names are different, and they get changed around from time to time. And so having someone actually at home verifying, ‘Oh, these are the inhalers that they actually have at home...’” I try to have them bring in their medications.” [C3]
(+) Caregiver support	• Q32. ”Also encourage them, if there is a family member or someone else that is sort of helping them with meds and kind of to keep things straight, to bring that patient—bring that person, rather, to the appointment“ [C2]
(–) Missed appointments due to travel barriers	• Q33. “...it’s often not super easy for them to leave the house. A lot of them require oxygen. They’re not extremely mobile. And so actually physically getting to the pharmacy or having someone pick their medications up can sometimes be difficult.” [C4]
(–) Missed appointments due to costs	• Q34. “They discuss[ed] price, like how would a medicine be obtained or any co-pays, any substitutions, or generic types. Because I was switched over to this Eliquis and more could have been explained about that.” [P4]
(–) Incomplete information/ communication with outpatient team	• Q35. “Patient comes for follow-up and generally, when you’re a primary care doctor or any doctor, I would imagine, you’re sort of triaging the most important to the least important. And education and med rec tend to fall further on the bottom than direct patient care and acuity of the patient.” [OL2] • Q36. “Every ambulatory clinic has a different system of how they reconcile or who reconciles or if they reconcile. So, for patients who see a lot of providers, they may or may not have a very up-to-date list.” [C2] • Q37. “I think the other thing, too, is, depending on the patient’s environment...there might be a language barrier as well.” [OL3]

[a]CG: caregiver.

[b]C: clinician.

[c]COPD: chronic obstructive pulmonary disease.

[d]OL: organizational leader.

[e]P: patient.

Inpatient Pre-Discharge Stage

Facilitators

Several facilitators enabled successful inpatient management and preparation for discharge. First, the COPD APN consultation was a mandatory step for enrolling in the COPD HRRP. In addition to clinical recommendations for the primary inpatient team, APNs provided written discharge materials to patients in the form of a COPD Action Plan. When possible, the pharmacy team provided medication reconciliation, which helped clinicians and patients to identify and address issues concerning missing or misused medications. At times, caregivers facilitated this process by playing an active role in medication reconciliation, aiding patients in recalling their home medications (Table 3, Q1 and Q2). Clinicians further noted that the Epic EHR medication dispense report aided in the medication reconciliation process (Table 3, Q3).

At discharge, eligible patients could receive additional medication reconciliation facilitated by the “meds-to-beds” program that delivered prescription medications directly to the bedside before discharge (Table 3, Q4). Finally, the pharmacy team provided inhaler instructions, using the TTG method to assess and encourage patient understanding (Table 3, Q5) [14]. Clinicians and OLs especially emphasized that using TTG at discharge significantly improved patients’ understanding of inhaler usage and follow-up care instructions (Table 3, Q6). Therefore, 67% of patients who responded in Phase 1 (4/6) felt prepared or extremely prepared to manage their COPD following inpatient discharge, with a mean (SD) survey score of 4.5 (1.03). Additionally, 83% of Phase 1 patients (5/6) agreed or strongly agreed that their inhaler education was helpful, with a mean (SD) survey score of 4.5 (0.84).

Clinicians and OLs further noted that an established positive clinician-patient relationship facilitated effective COPD education by creating rapport and improving patient comfort in speaking honestly about their needs (Table 3, Q7).

High patient health literacy also significantly aided the discharge process: both clinicians and patients felt that patients with a

XSL•FO
RenderX

clearer understanding of their health were better able to retain and use COPD self-management instructions. They also felt that patients' health literacy significantly improved medication reporting (Table 3, Q8). Finally, patients benefited from scheduling an outpatient follow-up appointment at discharge, which promoted care continuity and coordination of the next care steps with our program.

Barriers

Several barriers prevented a streamlined workflow and successful COPD education. Only 22% of clinicians in Phase 1 (2/9) felt satisfied with pre-discharge medication reconciliation, rating current practices a mean (SD) of 3 (0.83).

Frustrations were also reported around discharge delays due to inpatient processes that depended on multiple clinicians performing patient care tasks in a timely manner (Table 3, Q9). Another concern was inconsistent insurance coverage for "meds-to-beds." Clinicians reported other issues with insurance, stating that time-intensive processes (eg, prior authorization and communication with pharmacies) delayed medication delivery (Table 3, Q10). While medication reconciliation was highly recommended at admission and at discharge, limited pharmacy availability during some off-hour admissions resulted in missed consultations. In cases where pharmacists were available, other complications led to a lack of clear or accurate information on patient medication, hindering medication reconciliation (eg, complex medication lists, multiple patient records, undocumented and uninsured pharmacy purchases) (Table 3, Q11). Furthermore, clinicians and OLs observed that patients with poor health literacy struggled to correctly report their home medications, often relying on caregivers to understand their medication lists and provide accurate information (Table 3, Q12). Many clinicians noted that fast-paced education at discharge coupled with poor health literacy often resulted in patients failing to retain instructions (Table 3, Q13).

A few participants noted that educating patients immediately before discharge was not the optimal time, as patients with COPD were disoriented or eager to go home, failing to engage with the discharge COPD education and medication review, and returning home confused about self-management (Table 3, Q14). Furthermore, some patients were unable to contact their inpatient team following discharge when they had questions or were unsure about self-management (Table 3, Q15). Depending on the clinician's role (eg, respiratory therapist vs nurse vs pharmacist) and experience, the quality of patient education varied widely.

Additionally, with several inpatient team members visiting the patient and multiple lines of COPD therapy, confusion sometimes arose about instructions (Table 3, Q16). A few clinicians reported that with 6‐12 patients to manage simultaneously, challenges such as fatigue and limited understanding of roles and responsibilities in the discharge process prevented complete and successful education (Table 3, Q17 and Q18). Due to excessive clinician workloads, many stated that they frequently prioritized direct care over medication reconciliation and COPD education. As a result, only 17% of Phase 1 clinicians (2/12) felt that patients were prepared to manage their COPD following discharge (Table 3, Q19).

At-Home Post-Discharge Stage

Facilitators

Despite clinician dissatisfaction, 86% of Phase 1 patients (6/7) felt that they were prepared to take their medications following discharge from the hospital, with a mean (SD) score of 4.5 (0.76). Several aspects of the DTOC workflow enabled successful COPD self-management. Patients received a follow-up 48-hour phone call from a COPD registered nurse to check in on health and address any concerns or questions (eg, obtaining medications and post-discharge visit scheduling) (Table 3, Q20). Most patients appreciated the clarification on COPD self-management that the phone call offered (Table 3, Q21). Furthermore, patients enrolled in the MyChart patient portal had easy access to communication with both inpatient and outpatient teams during transitions (Table 3, Q22). Some patients received further support from caregivers who helped with medication compliance and proper inhaler use (Table 3, Q23). Patients with higher health literacy took a proactive role in their COPD self-management and remarked that the discharge materials given to them were useful and effective (Table 3, Q24).

Barriers

Conversely, patients with lower health literacy or patients of lower socioeconomic status are more likely to throw their discharge instructions away or neglect to read them, resulting in improper COPD self-management. In a few cases, patients lost their nebulizers or inhalers (Table 3, Q25 and Q26). Lower socioeconomic status patients also reported competing priorities (eg, housing and food insecurity) that prevented them from focusing on COPD self-management. In addition, some could not afford the cost of their medications and therefore did not fill their prescriptions (Table 3, Q27 and Q28). Finally, clinicians reported poor patient self-management due to individual factors, including substance abuse and mental health issues (Table 3, Q29).

Outpatient Post-Discharge Stage

Facilitators

At outpatient post-discharge follow-up, clinicians continued to monitor patient condition and COPD self-management. Patients were scheduled with COPD APNs for a post-discharge visit within 1‐2 weeks, where ongoing COPD care needs were evaluated and attended to; patients also received additional inhaler training from the pharmacy team. All Phase 2 patients (6/6) felt prepared or extremely prepared to manage their COPD following their outpatient follow-up, with a mean (SD) score of 4.5 (0.55). Like the inpatient setting, successful outpatient follow-ups were supported by an established relationship between the patient and the outpatient team.

Some patients were able to discuss medications more clearly by physically bringing the medications to their appointment or having their caregivers attend. This, along with access to the medication dispense report, helped clinicians to clearly

XSL•FO
RenderX

understand the medications and dosages that patients were using (Table 3, Q30 and Q31).

Additional caregiver responsibilities included coordinating transportation to outpatient clinics and providing accurate descriptions of medication and inhaler use (Table 3, Q32).

Barriers

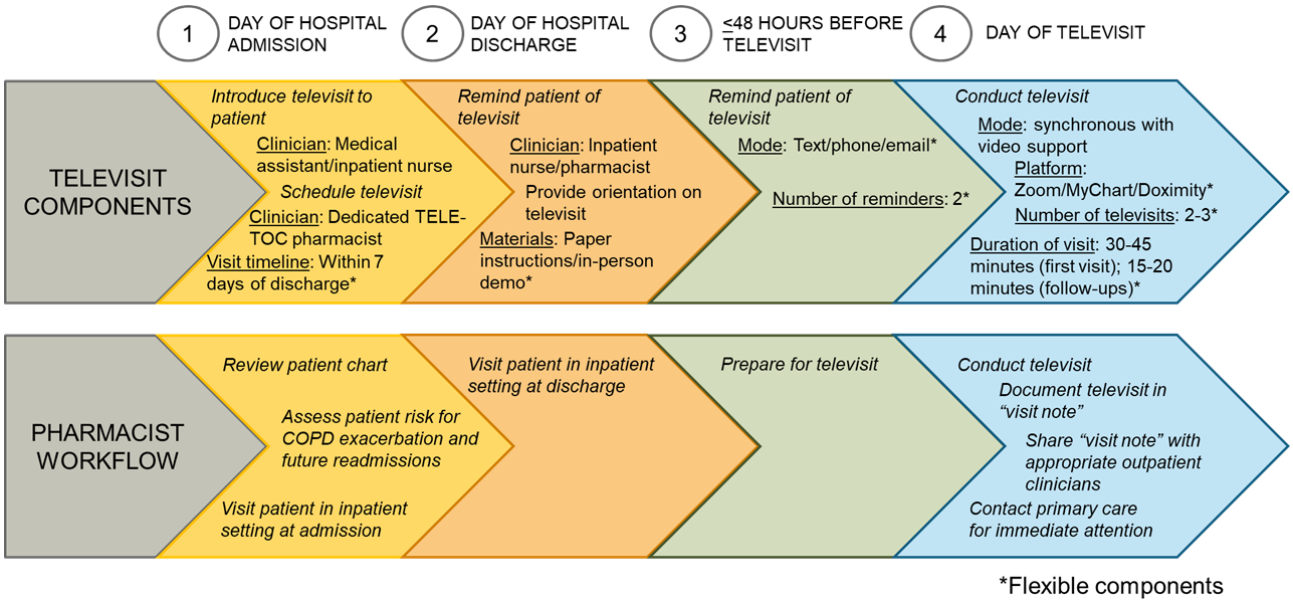
Clinicians reported that many patients missed their follow-up appointments for several reasons. Some could not physically travel to their outpatient follow-ups without assistance (Table 3, Q33). Furthermore, issues with insurance and medication costs continued into the outpatient stage, with patients sacrificing their follow-up appointments to pay for other necessities. Patients also reported going without medications due to expensive copays or if they were uninsured, leading to medication adherence discrepancies (Table 3, Q34).

Even though many patients spoke with their outpatient clinicians, the short appointment duration often resulted in incomplete medication reconciliation and limited assessment of patient literacy about medication use and COPD self-management. These issues, in addition to occasional cultural or language barriers, prevented clear communication and understanding between patient and clinician and sometimes resulted in misunderstandings about at-home care and medication use (Table 3, Q35-Q37).

Phase 2: Potential Role and Impact of Televisit Intervention on DTOC Workflow

Following our assessment of the DTOC workflow, we designed the televisit service with an initial, post-discharge, at-home televisit, followed by additional follow-up televisits before or after outpatient follow-up to maintain continuity of care. We developed a low-fidelity, televisit prototype for virtual pharmacist-led follow-up to provide inhaler training and medication reconciliation. Based on participant feedback, we categorized intervention design requirements as core (fundamental) or flexible (aspects tailored to patient preference) (Figure 6).

Figure 6. Telehealth Education: Leveraging Electronic Transitions of Care–integrated inpatient to outpatient Discharge Transitions of Care workflow. COPD: chronic obstructive pulmonary disease; TELE-TOC: Telehealth Education: Leveraging Electronic Transitions of Care.



Participant-proposed core components included a dedicated care transition clinician role, medication reconciliation, and coordination with outpatient care teams, and COPD education to reinforce inhaler techniques and skills with teach-back. Flexible components, adapted to patient needs and preferences, would include format, structure, reminders, and platform options. Participants felt that multiple options for these components would allow for flexibility and customization based on patients' preferences and needs. During the televisits, clinicians would use a checklist to ensure thorough inhaler education and teach-back. Additional patient COPD resources would be provided, including an inhaler use handout with visual aids and step-by-step instructions.

Most participants stated that televisits could address several barriers and augment the current DTOC workflow. They specifically mentioned potential improvements in patients' accessibility to virtual care, which could increase outpatient follow-up adherence. Participants' insights informed the iterative design of the televisit service, along with details on workflow and components according to patient and clinician preferences. We highlight the potential facilitators and barriers to televisit implementation (informed by the AHRQ-endorsed care transition framework) that may impact its buy-in and adoption (Table 4).

XSL•FO
RenderX

Table . Phase 2 topics with representative quotes. (+) Potential televisit facilitators; (–) potential televisit barriers.

CTF[a] domain and topic	Quote
Intervention characteristics	
(+) Use of televisit reminders	• Q1. "Maybe there's an initial text message that you're set for your appointment, this is your day, maybe 24 hours prior and then an hour maybe prior to their visit...Because I think that will remind a patient...like, 'Hey, it's time for you to get ready for your tele visit. Please sign on.' And maybe there's the link, and they go right to that link, and it takes them right to the visit." [OL6][b] • Q2. "A text would be best [to remind me]." [P9][c]
(+) Convenience of televisit	• Q3. "When it be cold and they don't have me coming out, that's when I do it on the phone with the doctors...Yeah. And raining. I ain't going out in the rain." [P11] • Q4. "We've actually seen an increase of [patients] attending visit(s) when it was telehealth because they didn't have to worry about getting a ride or commuting here." [OL3]
(+) Flexible televisit timing and frequency	• Q5. "[Wait to schedule the visit for] at least about a week...And that way, they...try to get familiarized with their house again because...coming home from the hospital is an experience." [P14] • Q6. "I think most of the time, patients are still overwhelmed when leaving the hospital...I think the sooner, the better to ensure that they start off on the right foot as far as utilizing their medications appropriately to prevent readmission." [C16][d] • Q7. "I think two to three visits would be reasonable..." [C15]
(+) Convenient documentation and education via televisit	• Q8. "One of the cool things about telehealth is you can actually have someone show you. So as long as they were doing it through a cell phone or a laptop or was mobile, I would say, 'Okay. Show me where you're storing your medications.' But through having patients show me their meds, I had a checklist of...is it being stored at the correct temperature, right? Is it in their fridge or not? Is it being stored mindfully with regard to if it's something that shouldn't be exposed to light?" [C11] • Q9. "I think the first visit is going to be primarily teaching, me showing them, 'This is how it's done, essentially.' And then second and third visit would be more assessing, 'Now that you've learned how to do this properly, how are you actually doing it?' And then additionally, outside of even inhaler technique, "Are you remembering to take it every day? How can we do those adherence sort of things, like setting reminders on your phones, engaging family members to help you, things like that?" [C13] • Q10. "In the first half, making sure there's no contraindications or medication, like disease and drug interactions. I always document that it's a virtual visit. And I also document if an interpreter or translator is needed. In the second half, I talk about medications: 'Are you still taking amlodipine? What are you taking for blood pressure...What medications are you taking?'" [C14]
(+) Dedicated role to support patients post-discharge	• Q11. "Just need the support (for a patient), either over the phone or in person to help navigate." [P9]
(–) Patient lack of access to technology	• Q12. "You want to make sure people have a camera, obviously, to use the video visits...So, getting a webcam. Making sure you either have headphones or a speaker or some way that you're going to physically communicate because a lot of the webcams may have speakers but may not necessarily have a microphone to pick up sound." [C15] • Q13. "I mean, there's plenty of barriers. Do they have phone service? Do they have cell service? Do they have internet access? Are they able to use Zoom or other video-conferencing software or MyChart software to access the visit?" [C5]
Characteristics of clinicians	

XSL•FO
RenderX

CTF[a] domain and topic	Quote
(+) Clear roles and responsibilities for clinicians	• Q14. “I think you also need a PSR to schedule them...It’s a big burden to worry about constantly scheduling the patients. What we do is we provide the schedulers with a template of the times where we are available, and then the schedulers place patients on that schedule.” [C14] • Q15. “I think the person that probably has the least time (to conduct a tele-visit) is probably the MD...It’s hard to say. My mind drifts towards RN maybe having more time.” [C13]
(–) Lack of clinician bandwidth	• Q16. “One big barrier for myself has kind of been a theme all along...doing the notes was a barrier for me. These notes...were relatively in-depth in that they were kind of like progress notes...But I think that was a major barrier for me in that if the average visit was 45 minutes and I had 60 minutes allotted to it, that meant I had to write my note in 15 minutes, which wasn’t always possible. And then sometimes staying late taking work home.” [C13]
(–) Inability to conduct in-person physical assessments	• Q17. “Some patients feel like it’s not a good visit, like it’s not a real visit because we didn’t touch them or really get to see them or do full vitals.” [C1]
Characteristics of patients	
(+) High patient health literacy	• Q18. “[I want to learn about] my health, medications, the do’s and the don’ts.” [P9] • Q19. “I do think you’re right. I think that there are some patients who you could probably do this in 10, 15 minutes, and it just goes by really quickly, and other patients that are going to take 45.” [C3]
(–) Low patient health literacy	• Q20. “Oftentimes, our patients have a gap in health literacy which impedes some things, which this would be a nice bridge to. But there would be also the concern that their technical literacy might also prevent them from doing a visit...” [OL4]
(–) Low patient technology literacy	• Q21. “What is their, I guess you can say, digital literacy? Do they know how to use it? Do they know how to get on to that?...When I think about the population that are most likely COPD, it’s usually an older population. I don’t know if they understand how to use that technology.” [OL3]
Process of implementation	

XSL•FO
RenderX

CTF[a] domain and topic	Quote
(+) Provision of training strategies for patients	• Q22. "So, I mean, if we're doing these visits through like a telehealth type visit...maybe it's techniques of how to be able to engage with a patient through a Zoom type of link and different techniques to make sure that you can assess the patient is able to do the proper techniques if they're using an inhaler or something." [OL6]
(+) Back-up strategies for implementation	• Q23. "Usually, we'll just ask the patient if we can call them on their cell phone. So that way, at least if there's anything major that we can still conduct most of the visit and still address anything that's immediate." [C16]
(+) Use of hands-on teaching strategies	• Q24. "I ask them to show me basically how do you do it, without actually doing it. And if they don't, then what I usually try and do is, if there's a video, I'll show them that, because I think that's really nice and typically, they can explain it better than I can. And then I'll supplement that with, "Okay. What questions do you have? And kind of show me what they showed you," and kind of do hands-on teaching from there." [C13]
(+) Incorporation of buy-in strategies	• Q25. "Yeah. And then you make sure that their inner (support) circle...knows [about COPD management]." [P14]
(–) Health care hierarchy clashes	• Q26. "I think doing the telemedicine at the skilled nursing facility is...really difficult. And the reason...is because they basically have somebody there giving them the medication...And theoretically, it's a trained professional who should be watching them do it. So I think that's difficult. Are you kind of stepping on someone's toes by doing it there? I don't know." [C13]
Organizational characteristics	
(–) Insufficient funding	• Q27. "I think that you're going to want to think about how...if the program is successful, how are you going to sustain it? A lot of these things that—it's great. They get grant funded. They get developed. And then once the grant funding runs out, then sustainability becomes an issue." [OL1]
(–) Lack of TELE-TOC[e] ownership	• Q28. "One of the things we had to figure out was who's going to own this process? Was it going to be nursing? Was it going to be my team? Was it going to be ambulatory? And so, coming up with that solution, I think, is one of the initial steps." [OL5]

[a]CTF: Care Transitions Framework.

[b]OL: organizational leader.

[c]P: patient.

[d]C: clinician.

[e]TELE-TOC: Telehealth Education: Leveraging Electronic Transitions of Care.

Intervention Characteristics

Facilitators

All participants felt that televisits could be supported with 2 - 3 text or phone call reminders to increase attendance and effective usage. While most were flexible with reminder timing, some specifically requested a reminder the day before or on the day of the televisit. Patients who preferred text reminders also felt that helpful messages should include tele-visit details (eg, clinician information, timing, and connection instructions) (Table 4, Q1 and Q2). Several patients believed that clarification for questions and concerns was more easily addressed via televisit between discharge and outpatient follow-up, allowing easy access to a pharmacist in the comfort of their homes (Table 4, Q3 and Q4). All patients and most clinicians and OLs believed that 2 - 3 televisits would be optimal (Table 4, Q7). They also felt that patients needed to settle in at home, observe treatment effectiveness, and gather any questions or concerns they have regarding medication use or discharge instructions, so televisits should be scheduled at least a week past discharge (Table 4, Q5 and Q6).

A few clinicians highlighted that using video televisits allowed patients to more easily show pharmacists their medications and inhaler routines (Table 4, Q8 and Q9). Most clinicians and several OLs also felt that the televisit service could be facilitated with the Epic EHR visit note template to document patient information, medication regimen, adverse events, quality of life, and patient counseling. In the event of an emergency, this

XSL•FO
RenderX

template could be forwarded to other clinicians. Clinicians stated that documentation could reduce information omission and maintain consistency (Table 4, Q10). Additionally, clinicians suggested that visits should be led by a nurse or pharmacist, as they were knowledgeable about their patients and often had more availability. Some patients also believed that televisits could be more effective with technical support (Table 4, Q11).

Barriers

A few clinicians and OLs highlighted that a fraction of patients might not have internet access, limiting their ability to connect to the televisit. Furthermore, issues with internet connectivity or video quality could prevent clinicians from seeing patients and medications clearly (Table 4, Q12 and Q13).

Clinician Characteristics

Facilitators

To avoid confusion about responsibilities and promote successful televisit implementation, most clinicians felt that a medical assistant, patient service representative, or nurse should schedule televisits. Many also recommended that clear roles and responsibilities be outlined to ensure better coordination and improvement of patient outcomes (eg, medical assistant or patient service representative for scheduling, nurse or pharmacist for televisit) (Table 4, Q14 and Q15).

Barriers

Several clinicians mentioned that time and workload could significantly limit their ability to conduct and document televisits. They worried that extra time after the visit would be required to accurately document the session, increasing the workload and potentially affecting documentation quality (Table 4, Q16). Others also stated that they might struggle with assessing physical symptoms through virtual televisits, impacting patient satisfaction (Table 4, Q17).

Patient Characteristics

Facilitators

Several clinicians felt that patients with high health literacy would be the most proficient users of the televisit service, as they would be more likely to engage during COPD education (Table 4, Q18 and Q19).

Barriers

Some clinicians and OLs worried that poor patient health literacy or technology literacy could negatively impact patient engagement during the visit, limiting the effectiveness of televisit outcomes (Table 4, Q20 and Q21). A total of 50% of Phase 2 patients (3/6) stated they had never used video technology platforms or conducted televisits, and several were not comfortable with FaceTime (Apple, Inc), WhatsApp (WhatsApp LLC), or other video chat technologies, scoring their comfort a mean (SD) score of 2 (2.00). Furthermore, some clinicians predicted that patients might have difficulty using apps such as Zoom videoconferencing software or the MyChart patient portal. Only 1 patient respondent was 100% comfortable using the internet, and many stated that they would need help, with a mean (SD) score of 2.3 (1.63).

Process of Implementation

Facilitators

To successfully implement the televisit, participants all suggested training and education for both clinicians and patients on the visit procedure (Table 4, Q22). A few clinicians and OLs also suggested conducting televisits via phone call if technical issues occurred (Table 4, Q23). Several clinicians believed that using demonstrational videos or other visuals would help to engage patients during televisits, increasing their understanding of COPD management and improving patient outcomes (Table 4, Q24). To encourage buy-in, a few clinicians and patients suggested involving caregivers in the televisit. As caregivers played an active role in patients' lives, participants speculated that they could encourage treatment adherence and aid the medication reconciliation process (Table 4, Q25).

Barriers

Several clinicians mentioned that some patients might receive care at skilled nursing facilities, which could introduce patient management conflict between tele-visit clinicians and in-facility clinicians (Table 4, Q26).

Organizational Characteristics

Facilitators

Our participants did not report any organizational facilitators within our data.

Barriers

All clinicians and some OLs strongly felt that secured funding for and clear, defined ownership of the televisit service were essential for long-term sustainability and consistent clinician accountability (Table 4, Q27). A small number of clinicians also highlighted their lack of dedicated space to conduct televisits, potentially making both the clinician and patient uncomfortable and affecting the quality of communication (Table 4, Q28).

Phase 3: Participant Perceptions of TELE-TOC Prototype

Following 5 walkthroughs with our high-fidelity TELE-TOC prototype, patients and pharmacists discussed their perceptions of TELE-TOC and identified facilitators and barriers to TELE-TOC overall, medication reconciliation, and inhaler education. Furthermore, suggestions for improvement were made (Table 5).

XSL•FO
RenderX

Table . Phase 3 topics with representative quotes. (+) Observed facilitators to Telehealth Education: Leveraging Electronic Transitions of Care; (–) observed barriers to Telehealth Education: Leveraging Electronic Transitions of Care.

Theme and topic	Quote
Participant perceptions of TELE-TOC[a]	
Appropriate content and duration of TELE-TOC sessions	• Q1. "I still have some questions too. And with the pharmacists too, we could sit down and we could talk a little bit more. I figured we didn't talk long enough." [P16][b] • Q2. "[The tele-visit] gave you all the—it gave me all the information I needed to help me, like I said, because I have to help me, not just let somebody else help. So that's what I like about this." [P17]
Preference for hands-on engagement	• Q3. "Just it was helpful to have someone actually really explain it rather than just hand you a piece of paper and your meds and say, 'Here you go.'" [P16] • Q4. "[The inhaler instruction] was easy because I can actually see the way he was showing me how to do it because I'm a visualist. I need to be able to see you do it so I know, 'Okay. This is how I need to do it.' And that helped me." [P17]
Improved confidence in COPD self-management	• Q5. [Upon being asked by the clinician whether the tele-visit made the patient more confident in COPD self-management] "Yes, because I know as long as I do what I'm supposed to do, my chances of going to the emergency room are decreased. I prefer to be able to—like you said, I'm in the comfort of my own home...I just want to have everything that I need right here with me and it'll be accessible to me so I can take care of me better." [P17] • Q6. "For me, it's just like a sparkle that somebody even cared enough to talk later...You came in there—you came in there and talked to me, and I feel like we've been knowing each other for—do you understand what I'm saying?" [P18]
Experiences with medication reconciliation component of TELE-TOC	
(+) Straightforward medication reconciliation with use of video technology	• Q7. [Upon patient confusion about two medications with the same cover but different names] "No, no, no, no, no. We got the Symbicort. I told you I had two of them. It was only one...Can you see that?" [P16]
(+) Easy access to patient health information via MyChart	• Q8. [Upon clinician asking if the patient was taking Lasix] "I was taking some Lasix. And I was just told to leave that alone too. So I'm not taking Lasix anymore...They cut that out." [P16] • Q9. [Upon clinician asking if the patient received azithromycin recently] "No, I don't even know what that is." [P16]
(+) Improved understanding of medication through in-depth patient education	• Q10. "When you give somebody or tell somebody to take stuff, tell them what it's for and make sure they understand what it's for. Just don't [say] like, 'Here, take this here three times a day and call me next week and let me know how you feel.'" [P16] • Q11. [Upon patient asking about what a spacer does in relationship to inhalers] "Oh, that's not for your inhalers. That helps strengthen your lungs. So that helps kind of control—so when you push this button, the medication comes out really quickly. This allows it to just kind of—allows us to take a much easier and deep breath with the medicine, and it allows it to get fully into our lungs." [C17][d]
(–) Lack of patient preparation for medication reconciliation	• Q12. "It would have been better if I had a copy of that [medication] list to go over it with you. [When asked if the patient could find their medications to show the clinician] I can't find them right now. That's what I was looking around for, but I was getting short of breath just looking for them." [P16]
Experiences with inhaler education component of TELE-TOC	

XSL•FO
RenderX

Theme and topic	Quote
(+) Enhanced inhaler education through visual cues from video technology	• Q13. “So what I’m going to have you do is you’re just going to act. As we go through this, you’re just going to act like you’re taking it. I’m not asking you to take an extra dose. But as we go through it, I want you to show me what you would normally do. So from the beginning, and don’t open the lid all the way, that’ll waste a dose, but we’ll act like the lid is open.” [C18]
(+) Improved understanding of inhaler education through teach-back method	• Q14. “I’m going to go ahead and show you really quickly. I’m just going to show you the steps and how I would use the inhaler and then have you show me one more time on how you go ahead and use it, okay?...[Upon explaining the steps and demonstrating inhaler use] That’s all I would do...Can you just go ahead and show me one more time how you would do it?” [C17]
(–) Lack of patient understanding despite education	• Q15. “[My confidence in my inhaler use is] still up in the air because I don’t know what triggers it. I don’t have a full understanding of my concentrators, of how they’re working and what to look for or when do I need to change the hoses on that and stuff. So there is still a lot more to be done with this topic.” [P16]
TELE-TOC use and implementation	

XSL•FO
RenderX

Theme and topic	Quote
(+) Smooth TELE-TOC process	• Q16. “Yeah. It was not something that I had to think hard about. It was easy for me to just go with the flow. I didn’t have to struggle trying to understand, “Okay. What are they going to ask?” I just went with it, and everything flowed right through naturally for me.” [P17] • Q17. “In terms of the last one, I think, honestly, overall, the flow was a little bit better. Obviously, this patient didn’t have their medications, but that is real world. But I think overall, say we were able to do the full visit top to bottom...” [C17]
(+) Convenient use of video technology	• Q18. [Interviewer asks whether the order of the tele-visit was easy to follow] “Yeah. Because I [inaudible] I see it.” [P15] • Q19. “[The tele-visit was] Great. I don’t have to come out. [laughter] I kind of struggle with [leaving the house]. And then I got anxiety. So, the fact is, when I can’t breathe, I panic, and it’s bad.” [P15] • Q20. “I felt good because I’m comfortable here. I didn’t have to come out into the cold or sitting in an office where I’d be uncomfortable. I’m comfortable sitting here. I’ve got my oxygen on too, so. But it felt good to be able to be at home and do something like this.” [P17]
(+) Helpful reminders for sessions	• Q21. “I appreciated [the reminder]. Yeah, it was very helpful because sometime I will forget some stuff. Then this morning they sent me one at 7:47 to say, ‘It’s time to log in.’ So I was able to log on in.” [P17]
(–) Patient desire for more in-depth education	• Q22. “I’m concerned about it all. And I need someone to be there to go over all of it. What should I do? If there’s something should happen, what should I do? I don’t want to be—I don’t want to be the one doctoring on myself.” [P16]
(–) Technical issues with Zoom	• Q23. “I can’t hear them. Can’t hear anybody.” [P16] • Q24. “My phone is not muted. My phone is not—hold on. Let me see. Microphone, iPhone microphone. We have your audio. I don’t know why you can’t hear me.” [C18]
(–) Low patient technology literacy	• Q25. [When asked about difficulty of logging into MyChart for tele-visit] “Usually, it’s really hard for me to get in, and I just give up. But it really wasn’t that bad. I guess I needed the hands-on to show me how to do it.” [P15]
(–) Confusing “visit note” for pharmacists	• Q26. “[It was confusing inputting information into the “visit note”]. Specifically under the section, quote, ‘script for pharmacist.‘ It just doesn’t feel like this necessarily belongs in a ‘visit note.‘ I’ve never seen an outpatient note say, ’This is your script that we want the doctor to say or the pharmacist to say, and then these are the responses.‘ So I just find it difficult to know what’s supposed to be included there. I’ll include his answers, but his answers are also included two or three other times in the note. For example, the question, ‘What medication do you take at home?‘ And then further down, it has a list of his medications. So I don’t really know. Do I type every medication there? It doesn’t make sense.” [C18]
Suggestions for improving content and implementation	

XSL•FO
RenderX

Theme and topic	Quote
Ask patients to prepare medications prior to sessions	• Q27. “Right. And then I would like to have— like to know to have all my medicines with me that I’m taking so that we can go over them. Don’t wait till I get on the meeting and say, ‘Dude, where’s your medicine? You got this. You got that.‘ You have all your medicines with you before we start this meeting. And that way, you got my undivided attention.” [P16]
Provide detailed education on COPD	• Q28. “[I’d like to know more about...] Maybe the causes of COPD. And how it could be treated. [Clinicians] usually just give you a paper and tell you to read aftercare, but I think if you talk about it more, people will understand it more.” [P15]
Provide clearer inhaler instructions at discharge	• Q29. “The [first day] I made it to a visit anywhere and somebody gave me an inhaler, we should have went over what we went over today. That should have been from day one...Anything that’s being used to help you breathe, it should be an in-depth— get a video of something that you should watch. All right? I mean, okay, a video is something. It’s a start.” [P16]
Provide a tutorial for TELE-TOC video chat	• Q30. “It probably would take maybe some training for them, or I would say if they were going to participate, don’t you know how you came in the room with me before we—you probably want to just ask them if you can see their phone and make sure the Zoom is set up on there.” [P18]
Amend “visit note” for streamlined sessions	• Q31. “Probably, just the template of the note I would reorganize a little bit because it kind of goes from med rec to COPD med rec and then—or COPD meds, and then it goes back to med rec again, which definitely threw me off. And there’s a lot of noise in the note, so it’s kind of hard to—while you’re trying to look at a patient, look at a note, but then also conduct something, it’s a lot to go back and forth. So I think probably rearranging the note a little more will be pretty beneficial by kind of just grouping med rec and then all the COPD stuff itself.” [C17]
Conduct TELE-TOC in a group setting	• Q32. “I would like to do it as a group. I’d like to have more than one, besides me, I want to have another patient there with me also. So we can compare our notes together. Me and that patient can talk. And that way, you’ll get a better understanding [of] how we feel...” [P16]
Target patients with medication needs	• Q33. “I think we were able to still go through the entire visit, identify some things that the patient can do better from their inhaler standpoint and then also how we can go ahead and assist the patient in terms of their medicines, not necessarily with adherence because we can’t dive too much into that, but in terms of medicines they need. Medicines that they’re out of and they need refills and stuff like that and also duplicate therapies.” [C17]

[a]TELE-TOC: Telehealth Education: Leveraging Electronic Transitions of Care.

[b]P: patient.

[c]COPD: chronic obstructive pulmonary disease.

[d]C: clinician.

Participant Perceptions of TELE-TOC

Patients and pharmacists described an overall positive experience with TELE-TOC. Patient SUS scores were reported as an average of 97.5/100, indicating that patients found TELE-TOC to be user-friendly and easy to use. Furthermore, when asked to rate the overall quality of TELE-TOC on a Likert scale from 1 (very poor) to 5 (very good), on average, patients rated TELE-TOC 5/5, while pharmacists rated TELE-TOC 3.8/5.

Most participants felt that TELE-TOC was not difficult to engage with and even suggested lengthening the sessions to create more time for questions and patient-specific concerns (Table 5, Q1 and Q2). When asked about TELE-TOC content, patients responded that they did not feel that any content needed to be removed; on the contrary, they requested the addition of more in-depth COPD information. Pharmacists and patients agreed that TELE-TOC was engaging and helpful, especially the inhaler education, which patients described as more hands-on than their discharge education and in-person follow-up visits

XSL•FO
RenderX

(Table 5, Q3 and Q4). Patients stressed that TELE-TOC's follow-up helped them to feel cared for and improved their confidence in medication management and inhaler use (Table 5, Q5 and Q6). They also felt that TELE-TOC motivated them to learn more and improve their COPD management.

Experiences With Medication Reconciliation Component of TELE-TOC

Facilitators

Regarding medication reconciliation, a pharmacist noted that it was easier to confirm which medications each patient was taking based on what they presented on video (Table 5, Q7). Furthermore, use of MyChart allowed pharmacists to easily access medication lists to double-check medication names, dosages, and frequency of use (Table 5, Q8 and Q9). Through TELE-TOC, pharmacists could identify prescription discrepancies while helping patients to better understand the purpose of each medication (Table 5, Q10 and Q11).

Barriers

Challenges with medication reconciliation included lack of patient preparation and low patient health literacy. Pharmacists noted that patients were not always adequately prepared to review their medication lists and may need reminders to have their medications ready for review in future sessions (Table 5, Q12). Some patients also simply read their medication labels but were unable to fully understand their medication management, even with help from the pharmacists.

Experiences With Inhaler Education Component of TELE-TOC

Facilitators

During inhaler education, video technology enabled pharmacists to see how patients used their inhalers and instruct them accordingly (Table 5, Q13). Most patients agreed that video technology made instructions easier to follow through visual feedback. They also enjoyed the teach-back method used by pharmacists, noting that it felt more engaging and helped them process and understand clinician instructions (Table 5, Q14).

Barriers

While most patients felt satisfied with their inhaler education from the TELE-TOC pharmacist, others reported that the lack of further in-depth education left them with an incomplete understanding of how to use an inhaler (Table 5, Q15). They felt that even with the education provided during the session, they were still not fully confident in their inhaler use.

TELE-TOC Use and Implementation

Facilitators

Overall, patients noted that TELE-TOC's flow was easy to follow and led to a more streamlined process of medication reconciliation and inhaler education (Table 5, Q16). Both pharmacist interventionists agreed that both components of TELE-TOC were easy to complete smoothly, and TELE-TOC did not require too much extra work to conduct (Table 5, Q17). Using video technology also allowed patients to speak to their pharmacists from the comfort of their home. Unanimously, patients agreed that the telemedicine nature of the intervention allowed for improved access to their clinicians and increased the likelihood of attending the virtual session, in comparison to a similar in-person follow-up (Table 5, Q18-Q20). Additional facilitators to TELE-TOC implementation included patient reminders via MyChart to ensure patient attendance (Table 5, Q21), as well as technology literacy among pharmacists to correctly set up visits. Both pharmacists noted that they did not have issues with navigating TELE-TOC or MyChart.

Barriers

Patients and pharmacists listed different barriers to TELE-TOC implementation. Several patients believed that TELE-TOC was helpful for those who were unfamiliar with their medications or inhalers, but others voiced the desire to learn more about their COPD and management beyond the general information provided alongside medication and inhaler education. For example, some patients wanted to better understand COPD epidemiology, along with signs and symptoms to watch out for at home; lack of further in-depth discussion about their condition and experiences left them feeling mildly dissatisfied (Table 5, Q22).

Technology literacy and technical issues were observed as impediments to conducting TELE-TOC. Over multiple sessions, technical issues within Epic EHR and over Zoom made TELE-TOC difficult to conduct (Table 5, Q23 and Q24). Further, while patients stated that they did not have too much difficulty or concern with navigating the technology required to conduct TELE-TOC, issues such as forgotten passwords, confusion about MyChart login information, and misunderstandings about Zoom settings were observed during walkthroughs. One patient even noted that without technology assistance from the clinician or a caregiver, they would likely have given up on TELE-TOC (Table 5, Q25). These issues were quickly resolved with the help of TELE-TOC pharmacists, who were trained to provide several solutions to technical barriers (eg, switching to telephone calls, moving from MyChart to Doximity, and providing clear instructions in the event of patient problems with log-in or calling).

Both pharmacists also provided feedback on the TELE-TOC workflow and tools, primarily voicing initial dissatisfaction with the "visit note." They initially felt that the "visit note" was not structured in a way that allowed for streamlined conversation and were unsure of the needed depth for the notes (Table 5, Q26). As both pharmacists conducted more walkthroughs, they verbalized that the structure and clarity of the "visit note" greatly improved over time.

Suggestions for Improving TELE-TOC Content, Function, and Implementation

Following the discussion of the barriers to their TELE-TOC experiences, participants provided a few suggestions to mitigate these challenges. To combat a lack of patient preparation for and to streamline medication reconciliation, it was suggested that pharmacists ask patients to preemptively prepare their medications prior to TELE-TOC (Table 5, Q27). Furthermore, more detailed education and anticipatory instructions on inhaler use and general COPD management were suggested as a method

XSL•FO
RenderX

to support patients who still felt unsure in how to care for their conditions (Table 5, Q28 and Q29).

Participants also felt that providing clear instructions and a tutorial on how to use Zoom and MyChart technology would be helpful in assisting patients with lower technology literacy (Table 5, Q30). Furthermore, amendments to the "visit note" were suggested: sequence the note in alignment with the structure of each session and provide clearer instructions for pharmacists to complete the document (Table 5, Q31). Other suggestions included establishing an additional group format of TELE-TOC to provide inhaler and COPD education for multiple patients at once (Table 5, Q28), as well as using TELE-TOC as an additional source of support for patients with medication needs, such as checking on refills or checking for duplicate therapies (Table 5, Q32).

Discussion

Principal Results

Our 3-phase, UCD approach tailored and optimized an evidence-based televisit medication management and education intervention for patients with COPD transitioning from hospital to home. We identified facilitators, barriers, and interdependencies across the current DTOC workflow for patients with COPD, seeking to amend the workflow and improve patient outcomes. Key facilitators across settings included the "meds-to-beds" program and caregiver support, while challenges to effective care included inopportune education timing, inaccurate medication records, and limited outpatient follow-ups. These findings suggested the need for more comprehensive post-discharge patient support, which indicates the utility of the TELE-TOC intervention.

While the study was open to patients of all races and ethnicities, all enrolled patients identified as Black and non-Hispanic. This aligns with trends observed in UCM's COPD HRRP, in which a large proportion of eligible patients are Black. As a result, our study sample reflects the demographic makeup of UCM's patient population. Several factors, including lower utility of health care services, health insurance coverage, and behavioral differences (eg, smoking habits), have increased COPD's impact on Black patients [36,37]. Hence, Black patients hospitalized with COPD exhibit greater symptom severity and have significantly higher odds of 30-day readmission in comparison to non-Hispanic White patients [38-40]. Despite this disproportionate effect, very few studies examining the implementation and effectiveness of COPD DTOC interventions have explicitly included Black populations [41].

Comparisons With Prior Work

Our study corroborates existing literature, which emphasizes clinician COPD training and positive patient-clinician relationships as key facilitators to successful post-discharge care, alongside key barriers such as resource constraints, limited patient understanding of the education provided, and inappropriate allocation of tasks among clinical staff [17,42,43]. However, we identified additional barriers not widely discussed in prior research, such as patients using cash online pharmacies, resulting in undocumented medication use. This underscores the importance of accurate medication reconciliation, which is often limited in hospital settings.

We then explored how to address these barriers using pharmacy-led, virtual, in-home televisits for medication reconciliation and inhaler education. Our data suggested that televisits would be well-received and engaging due to their convenience and the addition of a dedicated pharmacist during the critical care transition period. Prior literature also highlighted the significant advantages of telehealth in medication management and virtual health education. For example, a previous study found that patients benefited more from telehealth (vs usual care) medication reconciliation, showed higher engagement and buy-in, and maintained better adherence to prescribed medications [44]. Further, multiple reviews [24,45,46] noted that telehealth-based interventions have improved patients' post-discharge healthcare outcomes and quality of life.

Based on these insights, we refined the TELE-TOC components and workflow. Notably, TELE-TOC offers a customizable telehealth experience that tailors televisit format (eg, visit mode, dose, frequency, and structure) to meet patient preferences and clinical needs, enhancing engagement and ensuring more personalized care while retaining core patient discharge components. Both patients and clinicians found TELE-TOC effective and efficient for medication reconciliation and comprehensive inhaler education. Despite some patients suspecting they would struggle to log in and navigate to TELE-TOC on their own, SUS scores remained high due to support and clear instructions from clinicians. Furthermore, due to a 766% increase in telemedicine encounters following the COVID-19 pandemic [47,48], many patients and providers report improvements in their televisit competency, which could explain the high success rate of TELE-TOC among our participants. Apart from a few technical difficulties, both patients and providers noted positive experiences with TELE-TOC usability and effectiveness [49]; other studies have reported similar feasibility and usability of eHealth interventions for patients with chronic conditions [50]. Other key success factors included remote monitoring and feedback from clinicians, visual aids via videoconferencing tools, and access to medication lists at home. Use of previously successful implementation strategies, such as patient-personalized televisit format, inhaler education teach-back, and TTG methods, was crucial to the success of our TELE-TOC prototype [13,16,51]. Several implementation barriers were addressed, including clinician availability for COPD education and patient forgetfulness about medications [52-58]. However, structural barriers, such as patient financial burden related to payment structures, remain unaddressed [52,59-61].

To evaluate the clinical impact of TELE-TOC, we are currently conducting a randomized controlled trial evaluating the effectiveness and implementation potential of TELE-TOC compared to usual care at UCM (NCT06461403), which is expected to be completed in 2026 [62]. Our primary effectiveness outcome assesses inhaler technique at 30 days after discharge, based on a validated 12-step checklist. Secondary outcomes include acute care revisits, rehospitalizations, and mortality rates at 30, 90, and 180 days

XSL•FO
RenderX

post-discharge, in addition to medication errors, changes in COPD symptoms based on the COPD Assessment Test and the Modified Medical Research Council dyspnea scale, and inhaler technique. Our primary implementation outcome is intervention reach. If successful, TELE-TOC could serve as a scalable model for improving COPD discharge care transitions and reducing avoidable acute exacerbations of COPD readmissions.

Limitations

We acknowledge our study limitations. First, this is a single-site study at an urban academic medical center, which may limit generalizability. In particular, the racial homogeneity of our study sample may limit the generalizability of our findings to the broader US population. Nevertheless, this study attempts to address the research gap in developing COPD discharge care transition interventions to meet the needs of Black patients. Second, we acknowledge that within our small sample for Phase 3, some patients mentioned familiarity and experience with using phones and computers to perform tasks, including video calls. This experience with technology may strongly contribute to their usability ratings, and we acknowledge that our SUS findings may reflect limited sample diversity. However, as previously mentioned, televisit proficiency increased significantly during the COVID-19 pandemic, and we anticipate that within our current randomized controlled trial, with a larger, more diverse sample, we may see similar TELE-TOC proficiency among our target population. Third, the feasibility of televisit interventions may not be applicable across health care organizations, depending on institutional resources. To best address this, we will develop an evidence-based roadmap for pragmatic implementation and dissemination of TELE-TOC across diverse rural, urban, and suburban US hospitals.

Conclusions

We adopted a UCD approach to codevelop and optimize a pragmatic and patient-centered COPD televisit intervention to support hospital-to-home care transitions. The goals of the proposed TELE-TOC intervention were to reduce unnecessary readmissions and improve continuity of care, provide impactful medication management and education to improve outcomes such as COPD symptom control, and ascertain implementation outcomes, including reach, sustainability, and scalability of TELE-TOC.

Acknowledgments

The authors wish to thank all participants of the Telehealth Education: Leveraging Electronic Transitions of Care (TELE-TOC) study and our TELE-TOC pharmacist interventionists who participated in Phase 3 of the study. We would also like to thank Ms. Kelley Foyil for proofreading the article. The support for this work comes from the Agency for Healthcare Research and Quality (R01AS027804).

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

Authors' Contributions

Conceptualization: JA, VA, VGP. Methodology: JA, VGP. Data curation/data collection: JA, NC, LT, MN. Formal analysis: JA, AM. Investigation: JA, NC, LT, MN. Writing – original draft: JA. Writing – review & editing: JA, AM, NC, LT, MN, VA, VGP.

Conflicts of Interest

VGP reports receiving grant funding from the National Institutes of Health (K24HL163408 and R01HL146644) and the Agency for Healthcare Research and Quality (R01AS027804). VGP also reports receiving consulting fees for her work with Humana. All other authors have no conflicts to report.

Multimedia Appendix 1
Telehealth Education: Leveraging Electronic Transitions of Care interview guides.
[DOCX File, 66 KB - humanfactors_v12i1e77953_app1.docx]

References

1. COPD chronic disease indicators. US Centers for Disease Control and Prevention. 2024 Jun 25. URL: https://www.cdc.gov/cdi/indicator-definitions/chronic-obstructive-pulmonary-disease.html [accessed 2025-11-06]
2. Weiss AJ, Jiang JH. Overview of clinical conditions with frequent and costly hospital readmissions by payer, 2018, in HCUP statistical brief #278. : Agency for Healthcare Research and Quality; 2021 URL: https://hcup-us.ahrq.gov/reports/statbriefs/sb278-Conditions-Frequent-Readmissions-By-Payer-2018.jsp [accessed 2025-11-06]
3. Liu Y, Carlson SA, Watson KB, Xu F, Greenlund KJ. Trends in the prevalence of chronic obstructive pulmonary disease among adults aged ≥18 years - United States, 2011-2021. MMWR Morb Mortal Wkly Rep 2023 Nov 17;72(46):1250-1256. [doi: 10.15585/mmwr.mm7246a1] [Medline: 37971940]

XSL•FO
RenderX

4. Mannino DM, Roberts MH, Mapel DW, et al. National and local direct medical cost burden of COPD in the United States from 2016 to 2019 and projections through 2029. Chest 2024 May;165(5):1093-1106. [doi: 10.1016/j.chest.2023.11.040] [Medline: 38042365]
5. Jencks SF, Williams MV, Coleman EA. Rehospitalizations among patients in the Medicare fee-for-service program. N Engl J Med 2009 Apr 2;360(14):1418-1428. [doi: 10.1056/NEJMsa0803563] [Medline: 19339721]
6. Press VG, Myers LC, Feemster LC. Preventing COPD readmissions under the hospital readmissions reduction program: how far have we come? Chest 2021 Mar;159(3):996-1006. [doi: 10.1016/j.chest.2020.10.008] [Medline: 33065106]
7. Defining the PCMH. US Department of Health and Human Services and Agency for Healthcare Research and Quality. 2025. URL: https://www.ahrq.gov/ncepcr/research/care-coordination/pcmh/define.html [accessed 2025-11-06]
8. Han MK, Martinez CH, Au DH, et al. Meeting the challenge of COPD care delivery in the USA: a multiprovider perspective. Lancet Respir Med 2016 Jun;4(6):473-526. [doi: 10.1016/S2213-2600(16)00094-1]
9. Doos L, Bradley E, Rushton CA, Satchithananda D, Davies SJ, Kadam UT. Heart failure and chronic obstructive pulmonary disease multimorbidity at hospital discharge transition: a study of patient and carer experience. Health Expect 2015 Dec;18(6):2401-2412. [doi: 10.1111/hex.12208] [Medline: 24831061]
10. Lindenauer PK, Pekow P, Gao S, Crawford AS, Gutierrez B, Benjamin EM. Quality of care for patients hospitalized for acute exacerbations of chronic obstructive pulmonary disease. Ann Intern Med 2006 Jun 20;144(12):894-903. [doi: 10.7326/0003-4819-144-12-200606200-00006] [Medline: 16785478]
11. Health Quality Ontario. Effect of early follow-up after hospital discharge on outcomes in patients with heart failure or chronic obstructive pulmonary disease: a systematic review. Ont Health Technol Assess Ser 2017;17(8):1-37. [Medline: 28638496]
12. Paasche-Orlow MK, Riekert KA, Bilderback A, et al. Tailored education may reduce health literacy disparities in asthma self-management. Am J Respir Crit Care Med 2005 Oct 15;172(8):980-986. [doi: 10.1164/rccm.200409-1291OC] [Medline: 16081544]
13. Press VG, Arora VM, Shah LM, et al. Teaching the use of respiratory inhalers to hospitalized patients with asthma or COPD: a randomized trial. J Gen Intern Med 2012 Oct;27(10):1317-1325. [doi: 10.1007/s11606-012-2090-9] [Medline: 22592354]
14. Press VG, Arora VM, Shah LM, et al. Misuse of respiratory inhalers in hospitalized patients with asthma or COPD. J Gen Intern Med 2011 Jun;26(6):635-642. [doi: 10.1007/s11606-010-1624-2] [Medline: 21249463]
15. Press VG, Arora VM, Kelly CA, Carey KA, White SR, Wan W. Effectiveness of virtual vs in-person inhaler education for hospitalized patients with obstructive lung disease: a randomized clinical trial. JAMA Netw Open 2020 Jan 3;3(1):e1918205. [doi: 10.1001/jamanetworkopen.2019.18205] [Medline: 31899529]
16. Press VG, Arora VM, Trela KC, et al. Effectiveness of interventions to teach metered-dose and diskus inhaler techniques. A randomized trial. Ann Am Thorac Soc 2016 Jun;13(6):816-824. [doi: 10.1513/AnnalsATS.201509-603OC] [Medline: 26998961]
17. Eisenhower C. Impact of pharmacist-conducted medication reconciliation at discharge on readmissions of elderly patients with COPD. Ann Pharmacother 2014 Feb;48(2):203-208. [doi: 10.1177/1060028013512277] [Medline: 24259651]
18. Mixon AS, Myers AP, Leak CL, et al. Characteristics associated with postdischarge medication errors. Mayo Clin Proc 2014 Aug;89(8):1042-1051. [doi: 10.1016/j.mayocp.2014.04.023] [Medline: 24998906]
19. Press VG, Au DH, Bourbeau J, et al. Reducing chronic obstructive pulmonary disease hospital readmissions. An official American Thoracic Society workshop report. Ann Am Thorac Soc 2019 Feb;16(2):161-170. [doi: 10.1513/AnnalsATS.201811-755WS] [Medline: 30707066]
20. Locke ER, Thomas RM, Woo DM, et al. Using video telehealth to facilitate inhaler training in rural patients with obstructive lung disease. Telemed J E Health 2019 Mar;25(3):230-236. [doi: 10.1089/tmj.2017.0330] [Medline: 30016216]
21. Thomas RM, Locke ER, Woo DM, et al. Inhaler training delivered by internet-based home videoconferencing improves technique and quality of life. Respir Care 2017 Nov;62(11):1412-1422. [doi: 10.4187/respcare.05445] [Medline: 28720676]
22. Emadi F, Dabliz R, Moles R, et al. Medication-focused telehealth interventions to reduce the hospital readmission rate: a systematic review. J Pharm Policy Pract 2025;18(1):2457411. [doi: 10.1080/20523211.2025.2457411] [Medline: 39917474]
23. Thai T, Plotke M, Downing G, Olmsted E, Cook B, Jafri FN. Telehealth pharmacist approach to comprehensive medication management in post-discharge high-risk patients: a quality improvement initiative. Telemed J E Health 2024 Apr;30(4):976-986. [doi: 10.1089/tmj.2023.0239] [Medline: 37976134]
24. Akula M, Nguyen M, Abraham J, et al. Determining if COPD self-management televisit-based interventions are evaluated among and equitably effective across diverse patient populations to reduce acute care use: a scoping review. Chest 2024 Dec;166(6):1371-1393. [doi: 10.1016/j.chest.2024.06.3799] [Medline: 39002815]
25. Waltman A, Konetzka RT, Chia S, et al. Effectiveness of a bundled payments for care improvement program for chronic obstructive pulmonary disease. J Gen Intern Med 2023 Sep;38(12):2662-2670. [doi: 10.1007/s11606-023-08249-6] [Medline: 37340256]
26. Abras C, Maloney-Krichmar D, Preece J. User-centered design. In: Berkshire Encyclopedia of Human-Computer Interaction: Sage Publications: Thousand Oaks; 2004:445-456.

XSL•FO
RenderX

27. Mao JY, Vredenburg K, Smith PW, Carey T. The state of user-centered design practice. Commun ACM 2005 Mar;48(3):105-109. [doi: 10.1145/1047671.1047677]
28. Turner CW, Lewis JR, Nielsen J. Determining usability test sample size. In: International Encyclopedia of Ergonomics and Human Factors 2006:3084-3088.
29. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the System Usability Scale. Int J Hum Comput Interact 2008 Jul 29;24(6):574-594. [doi: 10.1080/10447310802205776]
30. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
31. Fereday J, Muir-Cochrane E. Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. Int J Qual Methods 2006 Mar;5(1):80-92. [doi: 10.1177/160940690600500107]
32. Rojas Smith L, et al. AHRQ methods for effective health care. In: Contextual Frameworks for Research on the Implementation of Complex System Interventions: Agency for Healthcare Research and Quality (US): Rockville (MD); 2014. [Medline: 24783308]
33. Proudfoot K. Inductive/deductive hybrid thematic analysis in mixed methods research. J Mix Methods Res 2023 Jul;17(3):308-326. [doi: 10.1177/15586898221126816]
34. Bangor A, Kortum PT, Miller JT. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009;4(3):114-123 [FREE Full text]
35. Cypress BS. Rigor or reliability and validity in qualitative research: perspectives, strategies, reconceptualization, and recommendations. Dimens Crit Care Nurs 2017;36(4):253-263. [doi: 10.1097/DCC.0000000000000253] [Medline: 28570380]
36. Dransfield MT, Bailey WC. COPD: racial disparities in susceptibility, treatment, and outcomes. Clin Chest Med 2006 Sep;27(3):463-471. [doi: 10.1016/j.ccm.2006.04.005] [Medline: 16880056]
37. Shaya FT, Maneval MS, Gbarayor CM, et al. Burden of COPD, asthma, and concomitant COPD and asthma among adults: racial disparities in a medicaid population. Chest 2009 Aug;136(2):405-411. [doi: 10.1378/chest.08-2304] [Medline: 19318663]
38. Eisner MD, Blanc PD, Omachi TA, et al. Socioeconomic status, race and COPD health outcomes. J Epidemiol Community Health 2011 Jan;65(1):26-34. [doi: 10.1136/jech.2009.089722] [Medline: 19854747]
39. Goto T, Faridi MK, Gibo K, Camargo CA Jr, Hasegawa K. Sex and racial/ethnic differences in the reason for 30-day readmission after COPD hospitalization. Respir Med 2017 Oct;131:6-10. [doi: 10.1016/j.rmed.2017.07.056] [Medline: 28947044]
40. Nastars DR, Rojas JD, Ottenbacher KJ, Graham JE. Race/ethnicity and 30-day readmission rates in Medicare beneficiaries with COPD. Respir Care 2019 Aug;64(8):931-936. [doi: 10.4187/respcare.06475] [Medline: 30914490]
41. Fan VS, Gaziano JM, Lew R, et al. A comprehensive care management program to prevent chronic obstructive pulmonary disease hospitalizations: a randomized, controlled trial. Ann Intern Med 2012 May 15;156(10):673-683. [doi: 10.7326/0003-4819-156-10-201205150-00003] [Medline: 22586006]
42. Griffiths S, Stephen G, Kiran T, Okrainec K. “She knows me best”: a qualitative study of patient and caregiver views on the role of the primary care physician follow-up post-hospital discharge in individuals admitted with chronic obstructive pulmonary disease or congestive heart failure. BMC Fam Pract 2021 Sep 7;22(1):176. [doi: 10.1186/s12875-021-01524-7] [Medline: 34488652]
43. Miravitlles M, Bhutani M, Hurst JR, et al. Implementing an evidence-based COPD hospital discharge protocol: a narrative review and expert recommendations. Adv Ther 2023 Oct;40(10):4236-4263. [doi: 10.1007/s12325-023-02609-8] [Medline: 37537515]
44. Noel K, Messina C, Hou W, Schoenfeld E, Kelly G. Tele-transitions of care (TTOC): a 12-month, randomized controlled trial evaluating the use of telehealth to achieve triple aim objectives. BMC Fam Pract 2020 Feb 7;21(1):27. [doi: 10.1186/s12875-020-1094-5] [Medline: 32033535]
45. Rezende LC, Ribeiro EG, Parreiras LC, et al. Telehealth and telemedicine in the management of adult patients after hospitalization for COPD exacerbation: a scoping review. J Bras Pneumol 2023;49(3):e20220067. [doi: 10.36416/1806-3756/e20220067] [Medline: 37132694]
46. Rush KL, Hatt L, Janke R, Burton L, Ferrier M, Tetrault M. The efficacy of telehealth delivered educational approaches for patients with chronic diseases: a systematic review. Patient Educ Couns 2018 Aug;101(8):1310-1321. [doi: 10.1016/j.pec.2018.02.006] [Medline: 29486994]
47. Le TV, Galperin H, Traube D. The impact of digital competence on telehealth utilization. Health Policy Technol 2023 Mar;12(1):100724. [doi: 10.1016/j.hlpt.2023.100724]
48. Shaver J. The state of telehealth before and after the COVID-19 pandemic. Prim Care 2022 Dec;49(4):517-530. [doi: 10.1016/j.pop.2022.04.002] [Medline: 36357058]
49. Sinha Gregory N, Shukla AP, Noel JJ, et al. The feasibility, acceptability, and usability of telehealth visits. Front Med 2023;10:1198096. [doi: 10.3389/fmed.2023.1198096]
50. Fomo M, Borga LG, Abel T, et al. Empowering capabilities of people with chronic conditions supported by digital health technologies: scoping review. J Med Internet Res 2025 Jun 27;27:e68458. [doi: 10.2196/68458] [Medline: 40577044]

XSL•FO
RenderX

51. Dantic DE. A critical review of the effectiveness of 'teach-back' technique in teaching COPD patients self-management using respiratory inhalers. Health Educ J 2014 Jan;73(1):41-50. [doi: 10.1177/0017896912469575]
52. O'Toole J, Krishnan M, Riekert K, Eakin MN. Understanding barriers to and strategies for medication adherence in COPD: a qualitative study. BMC Pulm Med 2022 Mar 19;22(1):98. [doi: 10.1186/s12890-022-01892-5] [Medline: 35305609]
53. Ramachandran HJ, Oh JL, Cheong YK, et al. Barriers and facilitators to the adoption of digital health interventions for COPD management: a scoping review. Heart Lung 2023;59:117-127. [doi: 10.1016/j.hrtlng.2023.02.004] [Medline: 36801546]
54. Lennox L, Green S, Howe C, Musgrave H, Bell D, Elkin S. Identifying the challenges and facilitators of implementing a COPD care bundle. BMJ Open Respir Res 2014;1(1):e000035. [doi: 10.1136/bmjresp-2014-000035] [Medline: 25478183]
55. Siu DCH, Gafni-Lachter L. Addressing barriers to chronic obstructive pulmonary disease (COPD) care: three innovative evidence-based approaches: a review. Int J Chron Obstruct Pulmon Dis 2024;19:331-341. [doi: 10.2147/COPD.S426050] [Medline: 38317666]
56. Smith BJ, Dalziel K, McElroy HJ, et al. Barriers to success for an evidence-based guideline for chronic obstructive pulmonary disease. Chron Respir Dis 2005;2(3):121-131. [doi: 10.1191/1479972305cd075oa] [Medline: 16281435]
57. Bivolaru S, Constantin A, Vlase CM, Gutu C. COPD patients' behaviour when involved in the choice of inhaler device. Healthcare (Basel) 2023 May 30;11(11):1606. [doi: 10.3390/healthcare11111606] [Medline: 37297746]
58. Davis JR, Wu B, Kern DM, et al. Impact of nonadherence to inhaled corticosteroid/LABA therapy on COPD exacerbation rates and healthcare costs in a commercially insured US population. Am Health Drug Benefits 2017 Apr;10(2):92-102. [Medline: 28626506]
59. Baldwin-Medsker A, Holland J, Rodriguez ES. Access to Care: using eHealth to limit location-based barriers for patients with cancer. Clin J Oncol Nurs 2020 Jun 1;24(3):16-23. [doi: 10.1188/20.CJON.S1.16-23] [Medline: 32441695]
60. Galvin E, Desselle S, Gavin B, McNicholas F, Cullinan S, Hayden J. Training service users in the use of telehealth: scoping review. J Med Internet Res 2024 Jul 31;26:e57586. [doi: 10.2196/57586] [Medline: 39083789]
61. Lai AM, Kaufman DR, Starren J, Shea S. Evaluation of a remote training approach for teaching seniors to use a telehealth system. Int J Med Inform 2009 Nov;78(11):732-744. [doi: 10.1016/j.ijmedinf.2009.06.005] [Medline: 19620023]
62. Ramadurai D, Lee CT, Traeger L, et al. Telehealth Education Leveraging Electronic Transitions Of Care for COPD Patients (TELE-TOC): a study protocol for a type II hybrid effectiveness-implementation randomised, pragmatic clinical trial of a pharmacist-led intervention. BMJ Open 2025 Nov 4;15(11):e105521. [doi: 10.1136/bmjopen-2025-105521] [Medline: 41193196]

Abbreviations

AHRQ: Agency for Healthcare Research and Quality
APN: advanced practice nurse
COPD: chronic obstructive pulmonary disease
CTF: Care Transitions Framework
DTOC: Discharge Transition of Care
EHR: electronic health record
HIPAA: Health Insurance Portability and Accountability Act
HRRP: Hospital Readmissions Reduction Program
OL: organizational leader
SUS: system usability scale
TELE-TOC: telehealth education: leveraging electronic transitions of care
TTG: Teach-to-Goal
UCD: user-centered design
UCM: University of Chicago Medicine

Edited by A Kushniruk; submitted 22.05.25; peer-reviewed by M Shawn McFarland, U Elechi, V Ajibade; revised version received 07.07.25; accepted 14.10.25; published 05.12.25.

Please cite as:
Abraham J, Meng A, Caravelli N, Traeger L, Nguyen M, Arora V, Press VG
Patient-Centered Televisit for Chronic Obstructive Pulmonary Disease Discharge Transitions: User-Centered Design Study
JMIR Hum Factors 2025;12:e77953
URL: https://humanfactors.jmir.org/2025/1/e77953
doi:10.2196/77953

XSL•FO
RenderX

© Joanna Abraham, Alicia Meng, Nicolas Caravelli, Leah Traeger, May Nguyen, Vineet Arora, Valerie G Press. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 5.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Messaging App Empowering Lifestyle Modification in Chronic Kidney Disease (LINE Official Account "Kidney Lifestyle"): Platform Development and Usability Study

Chun-Yi Ho[1], MEd; Deborah Siregar[1], MA; Miaofen Yen[1], PhD, RN; Junne-Ming Sung[2], MD, Prof Dr Med; Ming-Cheng Wang[2], MD, Prof Dr Med; Wei-Hung Lin[3], MD, Prof Dr Med, PhD

[1]Department of Nursing, College of Medicine, National Cheng Kung University, No. 1, Daxue Rd, East Dist, Tainan City, Taiwan

[2]Division of Nephrology, Department of Internal Medicine, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, Tainan City, Taiwan

[3]Division of General Internal Medicine, Department of Internal Medicine, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, Tainan City, Taiwan

Corresponding Author:
Miaofen Yen, PhD, RN
Department of Nursing, College of Medicine, National Cheng Kung University, No. 1, Daxue Rd, East Dist, Tainan City, Taiwan

Abstract

Background: Regular lifestyle modification is crucial for chronic kidney disease (CKD) management; yet, older patients often struggle to sustain behavior change and rely on support from their significant others such as family caregivers or partners. In such cases, both members of the dyad require accessible, jointly usable tools to maintain healthier behaviors over time. Given the ubiquity of instant messaging platforms, a digital intervention delivered via such a platform offers strong potential to empower CKD dyads in active lifestyle modification.

Objective: Guided by the Digital Dyadic Empowerment Framework, this study aimed to develop, optimize, and test the usability of a digital platform named "Kidney Lifestyle," using the LINE Official Account (OA) and an integrated extended app to facilitate collaborative lifestyle modification among CKD dyads.

Methods: We adopted a three-phase Agile-based development cycle: (1) iterative development and trial use, (2) heuristic evaluation, and (3) usability testing. In phase 1, the platform prototype was codeveloped with health care professionals and trialed by CKD dyads who provided feedback on interface clarity, ease of use, acceptance, intention to continue usage, and overall satisfaction. In phase 2, multidisciplinary experts conducted heuristic evaluations, rating compliance with Nielsen's 10 usability principles and suggesting improvements. In phase 3, experienced CKD dyads from phase 1 performed 6 representative tasks using the platform. Task success rates, completion times, and operational errors were recorded, and usability perceptions were assessed using the After-Scenario Questionnaire (1 - 7) and the System Usability Scale (0 - 100).

Results: In phase 1, 10 CKD dyads (19 individuals) reported high acceptance (mean overall satisfaction 4.1/5), valuing real-time interaction, convenient health data monitoring, and educational resources. In phase 2, 5 experts found high usability compliance (89% - 93%) but noted navigation complexity and the need for more interactive feedback. In phase 3, usability testing with 5 dyads showed high task success (60% - 100%) and short completion times (1 - 5 minutes). Extended app tasks used for structured self-monitoring achieved higher satisfaction, reflecting simpler navigation than tasks within the LINE OA (mean After-Scenario Questionnaire 5.64 vs 3.87). Navigation difficulties within LINE OA were likely due to multilayered menus and limited customization. The average System Usability Scale was 67.5, indicating marginally acceptable usability.

Conclusions: The LINE-based digital dyadic empowerment platform "Kidney Lifestyle" demonstrated promising usability and engagement. It has clinical potential to improve CKD control by extending health education, enabling continuous self-monitoring, and allowing clinicians to track patients' daily living conditions. To enhance effectiveness, future work should include a larger-scale feasibility trial while pursuing ongoing platform optimization, specifically by simplifying navigation pathways, adding a return option, and improving interactive feedback. The platform is now publicly accessible via LINE ID search, as provided in phase 1 results.

(JMIR Hum Factors 2025;12:e73935) doi:10.2196/73935

KEYWORDS

chronic kidney disease; dyadic empowerment; lifestyle modification; digital health; instant messaging; usability testing; LINE app

XSL•FO
RenderX

Introduction

Background

Chronic kidney disease (CKD) poses a serious burden to global health, affecting approximately 9.1% of the population or around 700 million people worldwide [1]. In Taiwan, the increasing prevalence of end-stage kidney disease has escalated health care costs, imposing substantial strain on the National Health Insurance system [2]. Effective CKD management can delay disease progression to end-stage kidney disease, reduce the risk of complications, and enhance patients' quality of life [3,4]. However, significant barriers remain, including insufficient continuous monitoring, limited patient education, and restricted access to health care providers [5,6].

CKD primarily affects older adults, especially those older than 60 years with lower educational levels, underscoring the need for tailored interventions for this group [7-10]. Digital health interventions (DHIs) offer unique opportunities to maintain patient engagement and provide tailored health education, particularly for older adults to improve autonomy and overall well-being [11]. With the widespread use of digital and mobile technologies, DHIs have become a key part of chronic care, supporting various health needs effectively [12]. Previous studies have shown that DHIs can improve self-management among patients with CKD [13]. This potential has led to the ongoing development of diverse tools—ranging from smartphone apps to web-based platforms—designed to support CKD and dialysis self-care [14-17].

Building on these advances, instant messaging platforms have emerged as promising tools for ubiquitous CKD management support. LINE, in particular, has gained significant popularity in Asia due to its user-friendly interface and broad demographic reach, effectively engaging users of all ages [18]. In Taiwan, LINE dominates the instant messaging market with more than 90% market share, making it a highly suitable platform for health communication [19]. The LINE app has already demonstrated potential in facilitating health education and self-management among middle-aged patients with CKD [20,21]. However, a LINE-based platform promoting lifestyle modification for older (older than 60 years) patients with CKD and their caregivers remains scarce. The extensive user base of LINE presents a unique opportunity to deliver cost-effective DHIs supporting multifaceted CKD management for this older population through a LINE Official Account (LINE OA).

Previous Work

In our previous work, we developed the Digital Dyadic Empowerment Framework (DDEF) as a conceptual and practical guide for designing DHIs that empower both patients with CKD and their significant others (SOs) (eg, spouses and children) to sustain lifestyle modification. The framework identifies 5 core domains of dyadic empowerment, including *perceived control*, *understanding*, *adaptation*, *dyadic efficacy*, and *dyadic support*. It further proposes 5 digital strategies: *digital education*, *digital communication*, *digital monitoring*, *digital feedback*, and *digital analysis*, through which DHIs can be designed to strengthen these dyadic empowerment domains. Together, these domains and strategies specify *what* to empower in the dyad and *how* to operationalize dyadic empowerment through concrete digital features, providing a structured pathway for designing and evaluating DHIs that support collaborative CKD self-management.

Objectives

This study aimed to develop, evaluate, and optimize the usability of the LINE-based digital platform "Kidney Lifestyle" for CKD dyads. The platform was directly developed under the guidance of the DDEF, applying its 5 digital strategies to strengthen dyadic empowerment and promote sustainable lifestyle modification among CKD dyads. This research specifically investigates the platform's usability, user acceptance, and potential clinical application in CKD management.

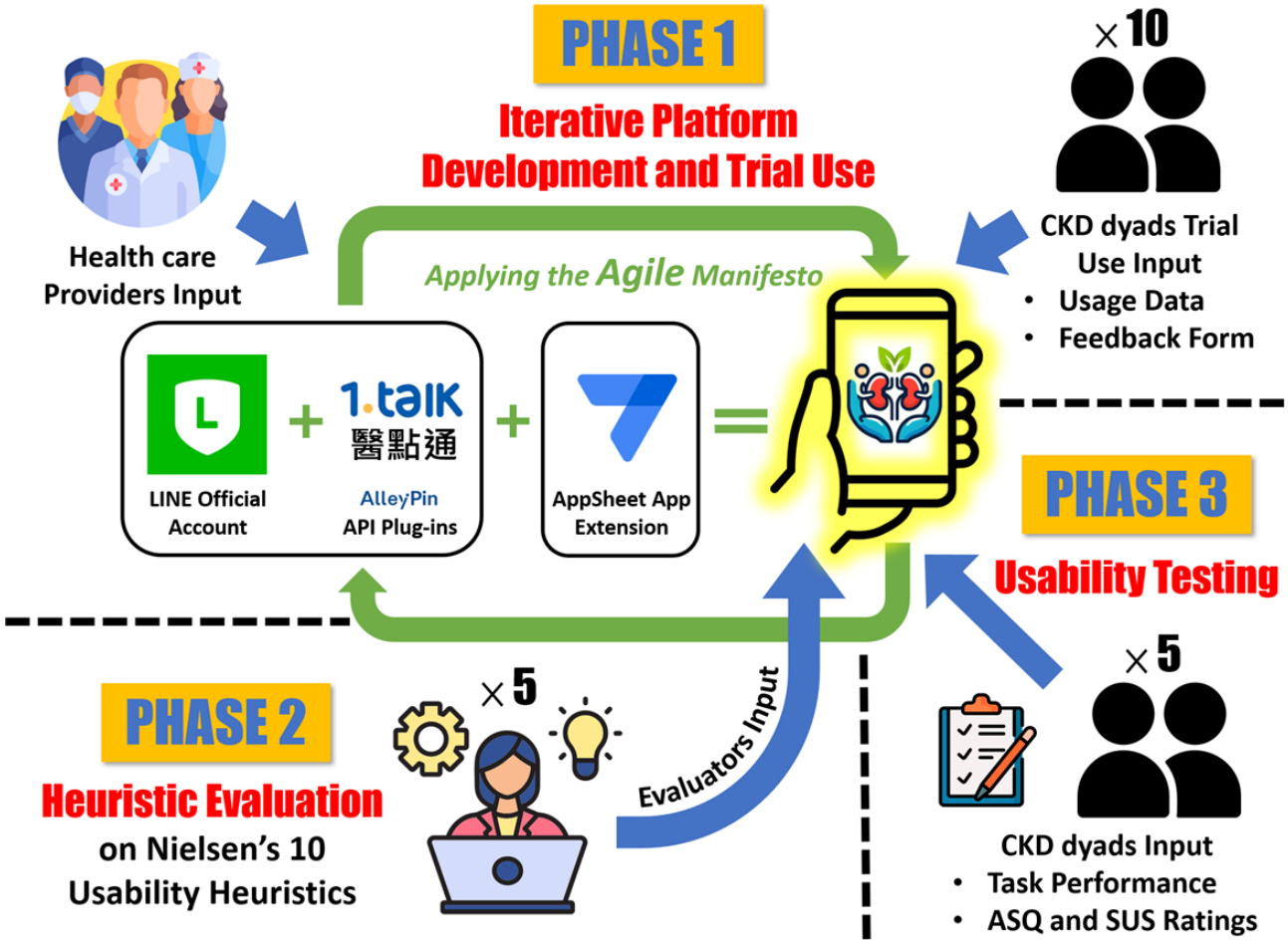
Methods

Study Design

Following the guidelines outlined by Kushniruk and Patel [22], a 3-phase iterative system development cycle was used to develop and test our LINE-based digital dyadic empowerment platform. These phases included (1) iterative development and trial use, (2) heuristic evaluation, and (3) usability testing (Figure 1). An Agile project management methodology was applied throughout the development cycle to manage initial uncertainties and facilitate continuous stakeholder interaction [23]. During each phase, participants identified usability issues, which guided subsequent refinements of the digital platform. The details of each phase are described in the following sections.

XSL•FO
RenderX

Figure 1. A visualization of the 3-phase iterative system development cycle. Blue arrows denote the input provided by the participants from each phase. ASQ: After-Scenario Questionnaire; CKD: chronic kidney disease; SUS: System Usability Scale.



Phase 1: Iterative Development and Trial Use

Phase 1 focused on developing the platform, especially the functionalities and content. Summarizing retrospectively, the platform development process involved the following four iterative steps: (1) platform establishment, (2) technical partner collaboration, (3) extended platform development, and (4) all platform integration and intervention design.

Step 1: Platform Establishment

A regular LINE OA was first set up to explore its basic functionalities, applying two functions aligned with 3 DDEF digital strategies: (1) chat: "Digital Feedback" and "Digital Communication," and (2) multimedia and multipage messages: "Digital Education." For example, CKD-related information (eg, symptoms, staging, and risk factors) could be delivered through autoresponse messages triggered by specific keywords. However, such keyword-based use was inconvenient and insufficient for promoting lifestyle modification among CKD dyads, as prior medical applications of LINE OAs mainly focused on clinic operations (eg, appointment scheduling and reminders) rather than ongoing lifestyle support [24]. These limitations highlighted the need for integration with third-party application programming interface (API) plug-ins to enable more advanced functionalities.

Step 2: Technical Partner Collaboration

AlleyPin Interactive Ltd Co is a LINE-certified Gold-level technical partner in the health care industry, providing API technical resources to meet diverse needs. Their cloud-based doctor-patient interaction management system, "DoctorConnect 1.Talk," when integrated with the team's LINE OA, can provide more advanced and user-friendly functionalities, including (1) service menu setup: it allows up to 8 buttons with multilayered structures for resource integration, beyond the 6-button limit of LINE OA; (2) enhanced multimedia and multipage messages: multiple buttons can be set within a single message, enabling users to access tailored health information via menu clicks without manual text input; and (3) case tracking system: it supports member management, scheduled "Push" messages, and interaction through AlleyPin's independent backend.

The service menu content and layout were designed not only by following the DDEF but also in consultation with a health education nurse from a university-affiliated hospital and 2 nephrologists, ensuring the platform met clinical needs. Through the integration of LINE OA and AlleyPin Interactive's API plug-in, our platform now effectively adopted 3 digital strategies—"Digital Education," "Digital Feedback," and "Digital Communication"—to support CKD dyads. To further apply the other 2 digital strategies—"Digital Monitoring" and "Digital Analysis," an extended app was developed and integrated with the LINE OA in the next step.

XSL•FO
RenderX

Step 3: Extended App (AppSheet) Development

After exploring various options and evaluating alternative platforms, the team discovered that Google's AppSheet platform could effectively compensate for the current platform's shortcomings and integrate with LINE OA. AppSheet is a platform for developing mobile and web apps that do not require programming code and can integrate with cloud spreadsheets or databases such as Google Drive and Office 365 as data sources [25]. The team immediately began developing the app prototype and conducted iterative testing.

Step 4: Platform Integration and Intervention Design

As the digital platform was not selected during the development of DDEF, the team, after the platform selection and development process outlined in steps 1-3, reviewed the correspondence between the DDEF and the functions provided by this integrated digital platform. Based on this, interventions were designed to promote lifestyle modification for CKD dyads.

Platform Trial Use

After establishing our digital platform prototype, the team invited 10 early-stage CKD (stages 1-3) dyads receiving regular treatment at a university hospital to use the platform in their daily lives. Besides assisting dyads in installing the digital platform on their smartphones, a user manual (Multimedia Appendix 1) was provided for each dyad to learn the operations at home. At the time of enrollment, operational instructions for various functions were provided by a research assistant (CYH) according to the manual.

A user feedback form (Multimedia Appendix 2) was provided to be filled out by the dyads at least 1 month apart during their next clinic visit. For each of the 6 major functions, the form assessed the following five elements: (1) interface clarity, (2) ease of operation, (3) acceptability: like or not, (4) intention to continue usage, and (5) satisfaction rating from 1 to 5 (Table 1). A brief interview was conducted with each dyad to obtain further feedback. The platform was iteratively optimized based on both system usage data and user feedback.

XSL•FO
RenderX

Table . Content of user feedback form for phase 1 dyadic trial use[a].

Element (item number) and question		Type
Interface clarity (N-1)		
	Are there any unclear elements on the screen?	Yes/No
	If yes, please specify.	Short answer
Ease of operation (N-2)		
	Is this function easy to use?	Yes/No
	If not, please explain why.	Short answer
Acceptability (N-3)		
	Do you like this function?	Yes/No
	Tell us what you like about this function?	Short answer
	If you dislike this function, how can we improve?	Short answer
Intention to continue usage (N-4)		
	Will you keep using this function?	Yes/No
	Why or why not?	Short answer
Satisfaction (N-5)		
	Please rate your satisfaction with this function on a scale from 1 to 5.	5-Point Scale
Previous experience (7-1)		
	Have you ever used any other apps or websites for managing chronic kidney disease?	Yes/No
	If yes, what did you like or dislike about them?	Short answer
General platform satisfaction (7-2)		
	What is your overall satisfaction rating with our platform?	5-Point Scale
	Will you continue to use our platform for chronic kidney disease management?	Yes/No

[a]Items N-1 to N-5 are inquired repeatedly for each of the 6 functions of the platform, resulting in a total of 30 questions. Items 7 - 1 and 7 - 2 are general questions asked once, adding 2 more questions. Therefore, the total number of questions in this form is 32.

Phase 2: Heuristic Evaluation

To further examine whether our digital platform is user-friendly in its design, a heuristic evaluation involving 5 evaluators with backgrounds in health care or information engineering was conducted. Evaluators tried out the digital platform and assessed how well each of the 6 major functions in the interface design complies with Nielsen's 10 usability principles (10 usability heuristics) [26].

The team referred to Moran and Gordon's [27] suggested heuristic evaluation steps and designed an evaluation form based on their published *Heuristic Evaluation Handbook*. Considering that the digital platform's access is not limited by time or space and acknowledging the need for evaluators to independently conduct the evaluation, the team opted for evaluators to fill out Google Forms for remote participation.

The form begins by inviting evaluators to join our digital platform on LINE and set up the initial settings. Evaluators were then asked to use the platform for 3 days. Specifically, they explored the buttons' functionalities and information on the LINE OA service menu and the extended app, similar to the real-use scenarios in phase 1 dyadic trial use. Before starting the evaluation, evaluators are invited to visit relevant websites to familiarize themselves with the content of Nielsen's 10 usability heuristics [28]. The explanations and the assessment questions for each principle are shown in Table 2.

XSL•FO
RenderX

Table . Explanation of 10 usability heuristics and assessment questions[a].

Heuristic	Explanation	Assessment questions
Principle 1: Visibility of System Status	The design should always keep users informed about what is going on through appropriate feedback within a reasonable amount of time.	• Does the design clearly communicate its state? • Is feedback presented quickly after user actions?
Principle 2: Match Between System and the Real World	The design should speak the users' language. Use words, phrases, and concepts familiar to the user, rather than internal jargon. Follow real-world conventions, making information appear in a natural and logical order.	• Will the user be familiar with the terminology used in the design? • Do the design's controls follow real-world conventions?
Principle 3: User Control and Freedom	Users often perform actions by mistake. They need a clearly marked "emergency exit" to leave the unwanted action without having to go through an extended process.	• Does the design allow users to go back a step in the process? • Are exit links easily discoverable? • Can users easily cancel an action? • Is *Undo* and *Redo* supported?
Principle 4: Consistency and Standards	Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform and industry conventions.	• Does the design follow industry conventions? • Are visual treatments used consistently throughout the design?
Principle 5: Error Prevention	Good error messages are important, but the best designs carefully prevent problems from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.	• Does the design prevent slips by using helpful constraints? • Does the design warn users before they perform risky actions?
Principle 6: Recognition Rather Than Recall	Minimize the user's memory load by making elements, actions, and options visible. The user should not have to remember information from one part of the interface to another. Information required to use the design (eg, field labels[b] or menu items) should be visible or easily retrievable when needed.	• Does the design keep important information visible, so that users do not have to memorize it? • Does the design offer help in context?
Principle 7: Flexibility and Efficiency of Use	Shortcuts—hidden from novice users—may speed up the interaction for the expert user such that the design can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.	• Does the design provide accelerators such as keyboard shortcuts and touch gestures? • Is content and functionality personalized or customized for individual users?
Principle 8: Esthetic and Minimalist Design	Interfaces should not contain information that is irrelevant or rarely needed. Every extra unit of information in an interface competes with the relevant units of information and diminishes their relative visibility.	• Is the visual design and content focused on the essentials? • Have all distracting, unnecessary elements been removed?
Principle 9: Help Users Recognize, Diagnose, and Recover From Errors	Error messages should be expressed in plain language (no error codes), precisely indicate the problem, and constructively suggest a solution.	• Does the design use traditional error message visuals, such as bold, red text? • Does the design offer a solution that solves the error immediately?
Principle 10: Help and Documentation	It is best if the system does not need any additional explanation. However, it may be necessary to provide documentation to help users understand how to complete their tasks.	• Is help documentation easy to search? • Is help provided in context right at the moment when the user requires it?

[a]Adapted from the "*Heuristic Evaluation Workbook*," by K. Moran & K. Gordon, 2023, Nielsen Norman Group.

[b]Field labels are commonly used in forms to provide context for information input.

Consistent with Donald et al [29], the evaluators assessed the degree to which each of the 6 major functions from our digital platform complies with the 10 usability principles. They gave ratings of "−1 (does not comply)," "0 (partially complies)," or "1 (complies)." If an item scores "does not comply" or "partially complies" with a principle, the evaluator will be asked to note specific issues in not adhering to that principle and provide suggestions for fixing these problems. If the evaluator finds that

XSL•FO
RenderX

the assessed principle does not apply to a specific function, they can select "N/A (not applicable)."

Phase 3: Usability Testing

Usability testing involves observing how users interact with the platform to fulfill their needs; understanding their behaviors, preferences, operational strategies, and paths; and collecting feedback to identify existing interface problems and possibilities for improvement.

We designed a task for each of the 6 major functions of our digital platform, which a realistic user would typically use in their daily lives. Five dyads who participated in phase 1, with at least 1 month of experience using the platform, were invited to perform the designed tasks. As the DDEF emphasizes helping relationships, tasks were performed jointly by the dyads, with each assisting the other throughout the process. Researchers provided instructions on task execution and responded to participant queries but refrained from providing guidance to avoid influencing their behavior. Task design and testing followed usability testing guidelines from Moran [30] and Nielsen [31]. Task success, time spent, and the number of operational errors during execution were recorded for each task.

Upon completion of each task, whether successful or not, participants filled out a postscenario questionnaire (After-Scenario Questionnaire [ASQ]) consisting of three questions to assess the perceived (1) ease of the task, (2) time required, and (3) support received during execution (help from the manual) on a scale from 1 to 7 [32]. A higher average ASQ score indicates greater user satisfaction and perceived usability. After completing all tasks, participants also filled out the System Usability Scale (SUS), a 10-item questionnaire designed to evaluate the overall usability and learnability of the digital platform [33]. A SUS score ranges from 0 to 100, with 68 indicating marginally acceptable perceived usability by the user and suggesting the existence of areas for improvement [34]. See Multimedia Appendix 3 for the full usability testing instrument. The task scenario and instruction for each of the 6 functions are shown in Table 3.

Table . Usability testing task scenarios and instructions for tested functions.

Task scenario	Instruction	Tested function
Task 1: Here are the values of your blood pressure and blood test results from today's visit: systolic blood pressure 125, diastolic blood pressure 78, heart rate 65, and creatinine (CREA) 1.4 mg/dL.	Please attempt to log the above values into our digital platform.	Record Values
Task 2: You want to know which stage of chronic kidney disease you are currently in. However, you found that your recent blood test report does not provide your "estimated glomerular filtration rate (eGFR)" value. You need to calculate your eGFR using "Creatinine (CREA)" (1.4 mg/dL), combined with your "gender" and "age."	Please try to explore our digital platform and calculate your eGFR.	Health Education Information
Task 3: Here is the date and time of your next appointment: December 28, 2023, afternoon clinic; suggested check-in time 14:10.	Please update the above information on our digital platform and notify the researchers through the platform.	Reminder Settings
Task 4: You want to learn about the symptoms of "hypoglycemia."	Please try to explore our digital platform and find the 7 possible symptoms of hypoglycemia.	Inquiry & Consultation
Task 5: Here is the content of your most recent meal: Half a bowl of white rice, 2 servings of bok choy, and 1 serving of chicken thigh.	Please record the above dietary information on our digital platform via photo or text note.	Kidney Health Mission
Task 6: Your partner (family member) has been trying hard to adjust their lifestyle recently. You want to know some ways to support their change.	Please try to explore our digital platform and write out support strategies corresponding to the following specific behaviors.	Kidney Support Teammate

Data Analysis

Data were compiled and analyzed using Google Sheets. Categorical variables (eg, user responses in feedback forms) were summarized as frequencies and percentages. Continuous variables, such as ASQ and SUS scores, were reported using means and ranges. Task success rates, completion times, and error frequencies were recorded and descriptively summarized for each usability task. No inferential statistical analyses were conducted, as the study primarily aimed to assess usability outcomes through descriptive measures.

Ethical Considerations

This study received approval from the institutional review board of National Cheng Kung University Hospital, Tainan, Taiwan (institutional review board no. B-ER-110-110). All participants were fully informed about the study's purpose, procedures, their rights, and data protection measures. Participation was voluntary, and participants could withdraw at any time without consequence. Paper source data were collected and stored in locked cabinets on site at the institution. Deidentified research data were stored on the institution's cloud servers and on password-protected personal computers, with access restricted to authorized study personnel only. All data will be retained

XSL•FO
RenderX

until 2029 and then securely destroyed. Each participant received a signed copy of the informed consent form. Participants received a gift voucher of NT $200 (approximately US $6) as compensation each time they completed a study task.

Results

Phase 1: Iterative Development and Trial Use

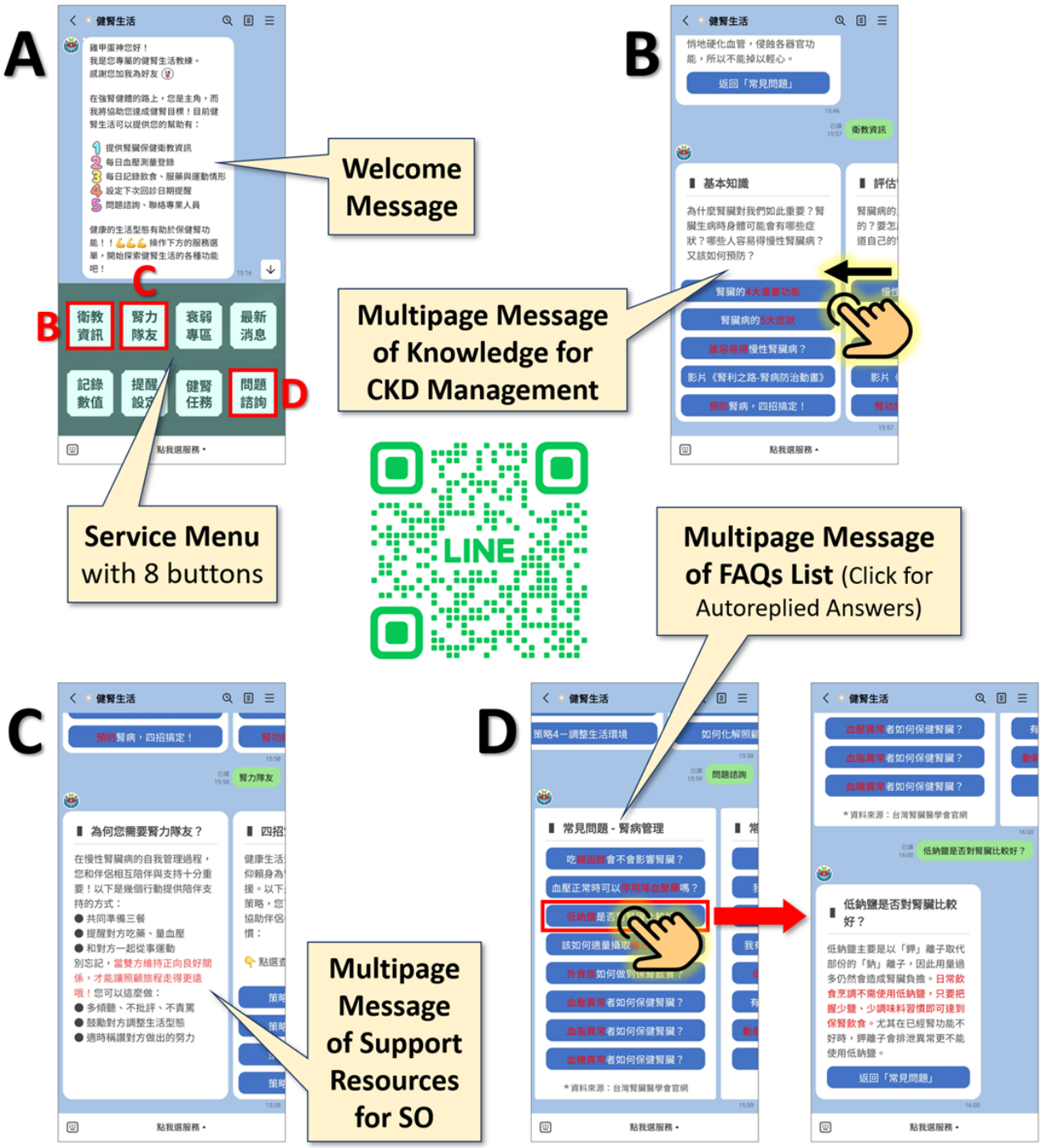
After collaborating with health care providers and technical partners for 3 months to develop our digital platform for intervention, we officially launched the prototype on July 25, 2023. Our LINE OA, named "Kidney Lifestyle," can be accessed by searching for LINE ID: @509kgajt. In addition, the extended app, named "Kidney Lifestyle App," is currently available through a dedicated link.

LINE OA "Kidney Lifestyle" With API Plug-Ins

Our LINE OA acts as the main hub for real-time interaction with users via text or voice chat. Moreover, by clicking buttons from the service menu (Figure 2A), users can access various resources for CKD management, retrieved or adapted from the official websites of *Taiwan Society of Nephrology* and *Taiwan Kidney Foundation*. The eight buttons featured on the service menu were introduced: (1) Health Education Information: providing basic kidney disease knowledge, evaluating kidney function, dietary and exercise guidelines, and so on (Figure 2B). (2) Kidney Support Teammate: multipage messages providing support resources for caregivers (SOs) (Figure 2C). (3) Frailty Section: providing health education information about frailty prevention and intervention. (4) Latest News: linking to the LINE OA's post area, forwarding information about relevant hospital events. (5) Record Values: users can log self-measured physiological values, and the system records and generates reports for feedback. (6) Reminder Settings: users can log their next appointment dates, and reminders are scheduled through the AlleyPin Interactive backend. (7) Kidney Health Mission: encouraging daily completion of "Record Values" and "Kidney Health Diary" (recording daily diet, medication, and physical activity). (8) Inquiry & Consultation: providing answers to frequently asked questions for CKD management; a consultation channel is provided for further queries (Figure 2D).

XSL•FO
RenderX

Figure 2. Interfaces of the LINE Official Account "Kidney Lifestyle" with a demonstration of user making inquiry about CKD management via automatic reply messages. The QR code provides access to "Kidney Lifestyle." (A) Basic Interface, (B) Health Education Information, (C) Kidney Support Teammate, and (D) Inquiry & Consultation. CKD: chronic kidney disease; FAQs: frequently asked questions; SO: significant other.



The Extended App "Kidney Lifestyle App" (AppSheet)

Our extended app mainly interconnects with the functions of "Record Values," "Reminder Settings," and "Kidney Health Mission" provided by our LINE OA with API plug-ins. Users can directly access the corresponding pages of the extended app by clicking on LINE OA service menu buttons, with the specific paths as follows: (1) Clicking "Record Values" on LINE OA service menu jumps to the "New Entry" page of the extended app, where users can log blood pressure, heart rate, and weight, and optionally enter blood sugar or blood test report data (Figure 3A). (2) Clicking "Reminder Settings" → "Go to Reminder Setting Page" on LINE OA service menu jumps to the extended app to view the next appointment date and time, which users can update manually (Figure 3B). (3) Clicking "Kidney Health Mission" on LINE OA service menu allows users to perform 2 "Daily Tasks," jump to the extended app to log physiological values or write a kidney health diary (recording daily diet, medication, and physical activity) (Figure 3C1). Alternatively, they can click "Challenge Missions" → "View Title Challenge Progress" on the extended app to confirm task execution status (Figure 3C2).

XSL•FO
RenderX

Figure 3. Interfaces of the extended app "Kidney Lifestyle App" and their connections with the LINE Official Account "Kidney Lifestyle" service menu. (A) The entry form of "Recording Values" and the chart viewing for self-measured physiological values, (B)"Reminder Settings" interface, and (C)"Kidney Health Mission," directing to (C1) Kidney Health Diary, records of daily diet, medication, and physical activity, and to (C2) Usage Report, presenting daily tasks execution status and progress.



In addition to the functionality accessible via LINE, the extended app interface also has a bottom-placed menu with five buttons (Figure 3), each briefly described as follows: (1) My Entries: viewing the "Usage Report" (the page redirected after clicking "View Title Challenge Progress" on LINE), "Physiological Value View," and "Blood Test Record" to check entered values. (2) Chart View: visualizing self-entered physiological values in graphical form. (3) New Entry: same as clicking "Record Values" on LINE OA. (4) Kidney Health Diary: adding, viewing, or editing written kidney health diaries. (5) Reminder Settings: same as clicking "Go to Reminder Setting Page" on LINE OA.

The data entered by users on various pages of the extended app were synchronized with their device IDs and sent back to the corresponding Google Sheets worksheets. Through these worksheets, the team can view all user-entered data and identify them based on the device ID and a custom user identification code. Subsequently, security filters were set up, distinguishing by device ID, enabling individual users on the front end (ie, extended app interfaces) to access only their entered data.

With the completion of the extended app's development and integration with the LINE OA, it can now fully use the "Digital Monitoring" strategy to allow both the research team and the users to track physiological values, daily diet, medication, and physical activity. For the application of the "Digital Analysis" strategy, we calculate user metrics such as accumulated entry counts, entry activity rates (EARs), and consecutive entry days. These serve as reference indicators for users' lifestyle modification progress, enabling personalized interventions through the LINE OA.

Interventions to Promote Platform Use

Referring to the DDEF, interventions delivered through our integrated digital platform were designed and finalized through the following three processes:

XSL•FO
RenderX

1. Defined 4 specific lifestyle modification behaviors: physiological values monitoring, healthy diet, adequate physical activity, and proper medication.
2. Designed 2 daily tasks on the digital platform to execute these four specific behaviors:
    - Record Physiological Values: measure physiological values and log them through "Record Values" (multiple times allowed).
    - Write Kidney Health Diary: record daily diet (photographs or textual notes); check whether medication was taken correctly and, if not, select a reason; and check whether physical activity for more than 30 minutes was done and record the number of steps taken.

After consulting health care providers and observing practical scenarios, we confirmed the benefits of regular execution of the 2 daily tasks for both patients and health care providers. Recording physiological values generates easily readable reports, facilitating quick understanding of trends in physiological value changes. Writing a kidney health diary allows health care providers to confirm whether the patient's meals align with the principles of a CKD diet, and it serves as a reminder for the patient to engage in physical activity and take medication correctly. Inappropriate medication check items reference the Chronic Kidney Disease Self-Care Scale [35] and the Elderly Medication Adherence Assessment [36], consolidated into 4 items: "missed medication," "not taking medication on time," "self-adjusted medication or dosage changes," and "using nonprescribed medication."

1. Designed strategies to promote the execution of daily tasks based on the principles of behavior change and processes of change listed in the DDEF, including:
    - Gamification Mode for Behavior Shaping: Different "Kidney Health Titles" are granted upon completing 2 daily tasks consecutively for 3 (*novice*), 7 (*intermediate*), 14 (*veteran*), 30 (*master*), and 60 (*grandmaster*) days. Users can view title challenge progress (in percentage) through the "Usage Report" page on the extended app.

To implement the gamification strategy, the team needed to calculate the number of consecutive days a user completed daily tasks from today backward. With the assistance of generative artificial intelligence, ChatGPT, for syntax writing, we used Google Apps Script to execute custom functions to solve this issue. Furthermore, the team further simulated data using RStudio (Posit Software, PBC) and attempted to interpret it using Item Response Theory to link the 2 indicators, "consecutive days" and "entry activity rate." The calculation method for the EAR is as follows:

$$\text{Entry activity rate} = \frac{\text{Days performing daily tasks}}{\text{Total observation days}} \times 100\%$$

The EAR can be viewed as an "ability" indicator for users performing daily tasks. A higher EAR indicates a more regular execution of daily tasks. The team simulated different levels of EARs (from 0% to 100%, at intervals of 5%) for 180 observation days, with or without task execution, a total of 10,000 times. The frequency of achieving various Kidney Health Titles was calculated among these 10,000 simulations (Table 4).

XSL•FO
RenderX

Table . Occurrences of achieving different titles within 180 days based on various entry activity rates (simulation counts=10,000)[a,b].

EAR[c], %	Novice achieving counts	Intermediate achieving counts	Veteran achieving counts	Master achieving counts	Grandmaster achieving counts
0	0	0	0	0	0
5	211	0	0	0	0
10	1507	0	0	0	0
15	4040	1	0	0	0
20	6839	13	0	0	0
25	8873	64	0	0	0
30	9723	248	0	0	0
35	9954	698	0	0	0
40	9996	1590	2	0	0
45	10,000	2999	13	0	0
50	10,000	5003	48	0	0
55	10,000	7191	192	0	0
60	10,000	8852	502	0	0
65	10,000	9712	1275	0	0
70	10,000	9964	3001	17	0
75	10,000	9999	5629	82	0
80	10,000	10,000	8179	355	0
85	10,000	10,000	9746	1699	12
90	10,000	10,000	9994	5394	250
95	10,000	10,000	10,000	9541	3119
100	10,000	10,000	10,000	10,000	10,000

[a]EAR= (Number of days executing daily tasks/Total observation days) ×100%.

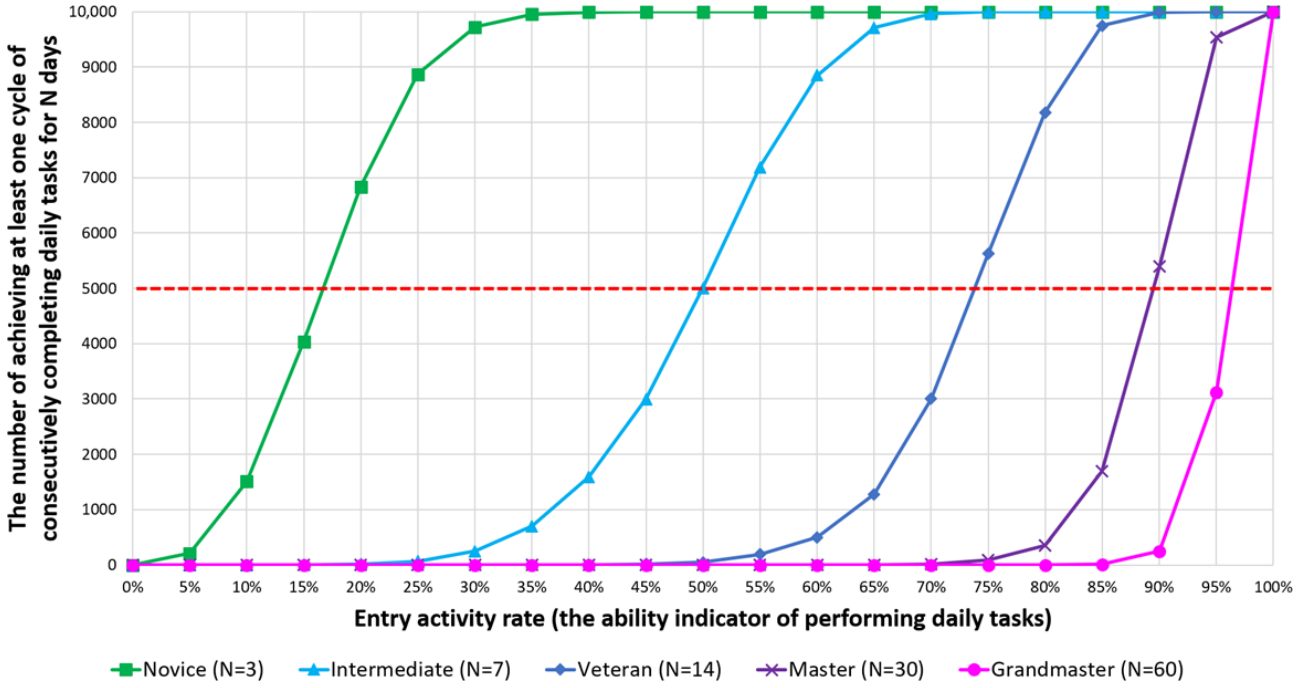
[b]Continuous execution days for each title: Kidney Health Novice—3 days; Kidney Health Intermediate—7 days; Kidney Health Veteran—14 days; Kidney Health Master—30 days; and Kidney Health Grandmaster—60 days.

[c]EAR: entry activity rate.

The simulation results provided item characteristic curve for each Kidney Health Title (Figure 4). These curves suggested that higher ability to perform daily tasks (indicated by a higher EAR) leads to a higher chance of achieving specific titles (indicated by more achievement instances). Using 5000 times of achievement as a baseline, its corresponding EAR on the item characteristic curve represents the difficulty level in achieving that specific Kidney Health Title. These results provide adequate rationality for the gamification strategy: assuming users in different stages of change have different EARs, this strategy sets appropriate challenges accordingly, aligning with our aim of providing corresponding interventions based on CKD dyads' different motivational states.

XSL•FO
RenderX

Figure 4. Item characteristic curves (ICCs) of achieving various kidney health titles within 180 days (simulation counts=10,000). According to Item Response Theory, with an achievement rate of 50% (ie, achieving 5000 times out of 10,000), the entry activity rate (EAR) corresponding to the ICC indicates the difficulty level to achieve that title. This simulation demonstrated the relationship between consecutively completing daily tasks for N days and its corresponding ability level, indicated by EAR. EAR = (Number of days executing daily tasks/Total observation days) × 100%. Continuous execution days for each title: Kidney Health Novice—3 days; Kidney Health Intermediate—7 days; Kidney Health Veteran—14 days; Kidney Health Master—30 days; and Kidney Health Grandmaster—60 days.



- Reinforcement management: Upon completing 2 daily tasks, users receive encouragement messages through the LINE OA; when obtaining a Kidney Health Title, the team sends encouragement messages and title certificates via LINE and informs their SOs. It also informs the patient's physician to provide face-to-face encouragement during the next visit (providing social reinforcement).
- Nudging through helping relationships: If a patient cannot complete tasks independently, the SO assists in recording. For patients in the early stages of change (eg, EAR below 50%) or those who have not performed tasks for more than 2 days, the system schedules LINE reminders via AlleyPin Interactive's case tracking system or notifies SOs to assist in reminders. References are drawn from the resources in "Kidney Support Teammate," providing assistance based on the patient's lifestyle context.

Platform Trial Use to Optimize the Platform

After the initial prototype of our digital platform went live, 10 CKD dyads visiting nephrology clinics regularly at a university hospital were invited to participate in the trial use from August to October 2023. One of the SOs (no. 003S) was invited by the patient but did not decide to participate in the end, resulting in a total of 19 participants. All dyads completed the user feedback form and returned it during the next visit. Table 5 shows the basic demographics of the participants. The following sections separately analyzed (1) the LINE message activity, (2) usage of the extended app, and (3) dyadic feedback content.

XSL•FO
RenderX

Table . Basic information of participants in phase 1 platform trial use (n=19).

ID[a]	Sex	Age (years)	Education	Relations	Stage of change	Primary user[b]
001P and 001S				Spouse	Maintenance	P
001P	Male	67	College			
001S	Female	59	Junior college			
002P and 002S				Spouse	Maintenance	P
002P	Male	63	College			
002S	Female	62	Senior high			
003P and 003S				N/A	Preparation	P
003P	Male	64	College			
003S[c]	N/A[d]	N/A	N/A			
004P and 004S				Mother-son	Maintenance	S
004P	Female	81	Elementary			
004S	Male	53	College			
005P and 005S				Father-son	Maintenance	S
005P	Male	82	Illiterate			
005S	Male	56	Junior college			
006P and 006S				Mother-daughter	Preparation	P
006P	Female	37	College			
006S	Female	73	Junior college			
007P and 007S				Spouse	Maintenance	P
007P	Male	64	Junior high			
007S	Female	59	Elementary			
008P and 008S				Spouse	Contemplation	P
008P	Female	67	Junior college			
008S	Male	68	Junior high			
009P and 009S				Spouse	Action	P
009P	Male	52	College			
009S	Female	51	College			
010P and 010S				Mother-daughter	Maintenance	P
010P	Female	70	Senior high			
010S	Female	42	Junior college			

[a]ID marked with P denotes patients; S denotes significant others.

[b]Primary user refers to the one using the app for daily tasks, primarily patients.

[c]003S was invited by 003P to participate in the study but eventually chose not to.

[d]N/A: not applicable.

LINE Message Activity

Based on data from the LINE OA backend, a total of 849 messages were sent during the initial 28 days (October 30 to November 26, 2023) after participant recruitment completion (as shown in Table 6), averaging approximately 30 messages per day. This indicates continuous interaction and usage.

XSL•FO
RenderX

Table . Number of messages sent via LINE Official Account and application programming interface module in the initial 28 days after participant recruitment completion[a].

Date	LINE Official Account			API[b] module		Daily total
	Auto Response	Welcome	Manual Chat	Push[c]	Reply[d]	
10/30	1	5	2	12	55	75
10/31	0	0	0	4	20	24
11/1	0	1	0	15	22	38
11/2	0	3	6	17	46	72
11/3	0	0	7	13	17	37
11/4	0	0	4	11	6	21
11/5	0	1	0	9	11	21
11/6	0	0	0	12	22	34
11/7	0	0	1	11	2	14
11/8	0	1	1	18	11	31
11/9	0	2	2	15	13	32
11/10	0	0	2	11	13	26
11/11	0	0	0	9	14	23
11/12	0	0	1	11	17	29
11/13	0	0	1	9	8	18
11/14	0	2	3	18	16	39
11/15	0	1	1	15	9	26
11/16	0	1	3	15	21	40
11/17	0	0	3	11	19	33
11/18	0	0	2	11	6	19
11/19	0	0	1	9	3	13
11/20	0	0	2	15	12	29
11/21	0	1	3	11	12	27
11/22	0	1	3	18	45	67
11/23	0	0	1	10	5	16
11/24	0	1	1	11	2	15
11/25	0	0	2	9	5	16
11/26	0	0	1	12	1	14
Total	1	20	53	342	433	849

[a]Data obtained from LINE Official Account Manager.

[b]API: application programming interface.

[c]Push messages sent to remind users of follow-up visits and daily tasks execution.

[d]Messages returned when clicking service menu layers or push message buttons.

Notably, the proportion of messages sent through API plug-in modules is extremely high: there were 342 Push messages, accounting for approximately 40% (342/849) of all messages, used for reminders on returning for checkups or daily task reminders and regular automatic care. There were also 433 Reply messages, constituting around 51% (433/849) of all messages, automatically sent by the system when users clicked buttons in the service menu or Push message prompts. In comparison, manually sent messages amounted to only 53, which is around 6% (53/849) of the total, primarily aimed at assisting users with operational issues or scheduling research-related appointments. These data demonstrate that our LINE OA with API plug-in modules shows a high degree of automation and convenience, allowing remote and real-time intervention, significantly reducing manpower costs, and reserving resources for handling more complex and specific user queries.

XSL•FO
RenderX

Usage of the Extended App

Upon observing the data on daily task execution by the primary users among the 10 dyads on the extended app (as shown in Table 7), since the start of enrollment on August 29, 2023, until 2:30 PM on April 29, 2024, a total of 2402 sets of physiological values have been logged, and 1079 kidney health diaries have been uploaded. Notably, it was found that the average EAR for logging physiological values (63.70%) was higher than that for writing kidney health diaries (50.74%). This difference might be due to the ease of recording physiological values and the immediate feedback obtained through trend chart viewing.

Table . Phase 1 primary users' usage of extended app for daily tasks (n=10)[a].

ID	Enrollment date (YY/M/D)	Days	Physiological values		Kidney health diary		Highest kidney health title[d]
			Counts	Activity rate, %[b]	Counts	Activity rate, %[c]	
001P	23/8/29	245	530	95.92	232	94.69	Master
002P	23/9/18	225	438	100.00	209	92.89	Master
003P	23/9/21	222	164	62.61	133	59.91	Intermediate
004S	23/10/2	211	2	0.95	0	0.00	None
005S	23/10/3	210	579	89.52	168	80.00	Master
006P	23/10/17	196	100	51.02	7	3.57	Novice
007P	23/10/19	194	125	63.92	25	12.89	Intermediate
008P	23/10/23	190	0	0.00	0	0.00	None
009P	23/10/23	190	307	90.53	161	84.74	Master
010P	23/10/30	183	157	82.51	144	78.69	Master
Total (Average)		206.6	2402	63.70	1079	50.74	N/A[e]

[a]This table is updated until April 29, 2024, at 2:30 PM.

[b]Activity rate for physiological values = (Days logged/Days enrolled) × 100%.

[c]Activity rate for kidney health diary = (Days written/Days enrolled) × 100%.

[d]Continuous execution days for each Kidney Health Title: novice—3 days; intermediate—7 days; veteran—14 days; and master—30 days.

[e]N/A: not applicable.

Eight users had an activity rate higher than 50% for logging physiological values, and among them, 6 also had an activity rate higher than 50% for writing kidney health diaries, indicating these users have a moderate to high level of acceptance and capability in performing daily tasks on the extended app. Regarding the challenges for Kidney Health Titles: 8 out of 10 users achieved consecutive logging of physiological values and writing kidney health diaries for at least 3 days (granted the "Kidney Novice" title) during the observation period; among the previously mentioned highly active 6 users, they achieved challenges of logging physiological values and writing kidney health diaries consecutively for 7 days (*intermediate*) or even 30 days (*master*). This demonstrates that higher activity levels correlate with a higher chance of long-term adoption of a healthy lifestyle, aligning well with the conclusions drawn from the simulation study.

Dyadic Feedback Content

Table 8 summarized the average ratings for each of the 6 major functions of our integrated platform, as provided by the 10 dyads in their user feedback forms. Generally, more than 80% of the dyads (more than 8) perceived each function to have a clear interface and to be easy to use. Similarly, most of the dyads (more than 80%) found each function acceptable and expressed their intention to continue usage. The average satisfaction rating for each function ranged from 3.5 to 4.3, with an overall rating of 4.1 for the platform as a whole.

XSL•FO
RenderX

Table . Phase 1 dyadic trial use feedback form rating summary (N=10)[a].

Functions of the digital platform	Assessed elements				
	#N-1: Interface clarity	#N-2: Ease of operation	#N-3: Acceptability	#N-4, 7‐2: Intention to continue usage	#N-5, 7‐2: Satisfaction rating (from 1 to 5)
Health Education Information	90%	90%	90%	90%	3.9
Kidney Support Teammate	90%	100%	90%	90%	3.5
Record Values	90%	80%	90%	80%	4.3
Reminder Settings	100%	90%	90%	90%	4.1
Kidney Health Mission	80%	90%	80%	80%	3.8
Inquiry & Consultation	100%	90%	80%	90%	3.7
Average	91.67%	90%	86.67%	86.67%	3.9
Overall platform evaluation	N/A[b]	N/A	N/A	90%	4.1

[a]The percentages in this table reflect the proportion of participant dyads who answered "Yes" to the corresponding Yes/No questions in the user feedback form.

[b]N/A: not applicable.

Qualitative feedback on each function, gathered through user feedback forms and brief interviews, is organized according to its positive and negative aspects:

1. Health Education Information (LINE): Dyads generally appreciated this function for its simplicity, clarity, and informative nature, noting that "it provided access to various, accurate CKD-related information at any time" (no. 005). However, one dyad (no. 010) suggested that "the font size should be enlarged, and the navigation of this function needs simplification to avoid getting lost."
2. Kidney Support Teammate (LINE): The majority of dyads found this function to be beneficial, appreciating its provision of "mutual support, encouragement, reminders, and detailed knowledge" (nos. 001 and 002). However, 2 dyads expressed concerns regarding its usage, with one noting "not using this function" (no. 005) and the other mentioning that "the colors are too vibrant, making it difficult to grasp key points, and navigating within the function could lead to getting lost" (no. 007).
3. Record Values (LINE+ App extension): Overall, dyads expressed satisfaction with this function, noting its advantages in "replacing traditional paper records, convenience, ease of use, and the ability to track data" (nos. 001, 002, and 009). Specifically, 1 dyad (no. 005) appreciated the support for chart view upon logging in data, facilitating easy tracking. However, 1 dyad (no. 010) mentioned dissatisfaction with the small font size and confusion regarding certain actions due to unclear wording.
4. Reminder Settings (LINE+ App extension): Dyads expressed overall satisfaction with this function, particularly highlighting its capability to provide advance notifications for upcoming clinical appointments. This function was deemed convenient and provided peace of mind by several dyads (no. 001, 005, and 009), eliminating the need to remember appointment dates. No negative feedback was received about this function.
5. Kidney Health Mission (LINE+ App extension): Overall, dyads who appreciated this function found its operation to be straightforward. Some dyads specifically noted its convenience for recording daily diet by uploading a photograph (no. 005) and its provision of daily data for tracking (no. 002). Conversely, 1 dyad found this function "troublesome, causing psychological pressure," and despite its good intent, felt "it lacked feedback from healthcare providers" (no. 001). Another dyad expressed inconvenience due to the small font size and difficulties logging in data (no. 010).
6. Inquiry & Consultation (LINE): Dyads provided positive feedback regarding the helpfulness of this function, particularly emphasizing its ability to provide answers to various relevant questions (nos. 002, 005, and 009). One dyad noted that "for new CKD patients, this function could serve as a substitute for in-person health education; however, for experienced patients, this function may seem unnecessary as they are already familiar with the answers" (no. 001). Another dyad commented on the potential for getting lost while navigating through the content of this function (no. 007).

Regarding the overall evaluation of the digital platform, neither dyad had previously used other apps or websites for managing CKD. Almost all of the dyads (90%) expressed willingness to continue using it to aid in managing their CKD. Upon further in-depth interviews with researchers about dyadic feedback, all agreed that the digital platform's functions were comprehensive, with no suggestions for additional functions. The first patient (no. 001P) mentioned that using the platform provided a "routine" after retirement. His SO (no. 001S) also mentioned the lack of diverse social activities after retirement for the patient

XSL•FO
RenderX

and expressed interest in participating in hospital support groups to share their experiences with other CKD dyads.

Phase 2: Heuristic Evaluation

By early November 2023, the team collected responses from 5 evaluators who participated in the heuristic evaluation. The evaluators' basic information is shown in Table 9.

Table . Basic information of phase 2 heuristic evaluation participants (N=5).

No.	Sex	Age (years)	Education	Background	Current occupation
1	Female	55	PhD	Nursing	Professor
2	Female	42	PhD	Health care science	Associate professor
3	Female	45	PhD	Nursing	Assistant professor
4	Female	54	PhD	Nursing	Assistant professor
5	Male	28	Master's degree	Electrical engineering	Engineer

General Compliance Rating

Their scores regarding the compliance of various functions with usability principles were tallied, averaged, and summarized in Table 10. On average, the compliance scores for each principle ranged between 0 (*partially complies*) and 1 (*complies*). The functions' average compliance with the 10 usability principles ranged from 89% to 93%. This suggests that the evaluators generally perceived that the digital platform's various functions mostly complied with the usability principles.

Table . Phase 2 heuristic evaluation compliance rating summary (N=5)[a].

Degree of compliance with each usability principle	Functions of the digital platform					
	Health Education Information, %	Kidney Support Teammate, %	Record Values, %	Reminder Settings, %	Kidney Health Mission, %	Inquiry & Consultation, %
Principle 1	60	80	100	80	80	100
Principle 2	100	100	100	100	80	100
Principle 3	100	80	80	60	100	75
Principle 4	100	100	100	100	100	100
Principle 5	100	100	100	100	100	100
Principle 6	80	80	100	100	100	75
Principle 7	80	80	100	60	100	80
Principle 8	100	100	100	100	100	100
Principle 9	100	100	33.33	100	33.33	100
Principle 10	100	100	100	100	100	100
Overall result	92	92	91.33	90	89.33	93

[a]Compliance percentages represent the sum of evaluator ratings (−1, 0, or 1) for each function's compliance with each usability principle, divided by the number of valid responses (ie, excluding "N/A [not applicable]" selections), and then multiplied by 100%.

Recommendations for Improvement

The breakdown of issues raised by evaluators about each function's incompliance with certain principles and their recommendations is as follows:

1. Health Education Information

- Principle 1 (60% compliance): Evaluator 1 found that "clicking into the five stages of CKD and lab test reports didn't lead to corresponding content." Evaluator 5 suggested "transforming health education information into daily articles and actively pushing them to users with thematic context."
- Principle 6 (80% compliance): Evaluator 2 felt that the design here might cause information loss or be easily forgotten after closing the linked web page and suggested implementing a memory function for highly clicked queries.
- Principle 7 (80% compliance): Evaluator 2 believed that this function lacked personal flexibility and suggested establishing personalized zones based on click rates to access or push commonly used information.

2. Kidney Support Teammate

- Principle 1 (80% compliance): Evaluator 5 felt that while the information was comprehensive, the incentive for users to actively read this information was slightly weak and suggested recommending correct concepts based on user records.
- Principle 3 (80% compliance): Evaluator 2 noticed that once clicked, the buttons for some information, returning to the previous message is not supported, requiring scrolling

XSL•FO
RenderX

or clicking on the service menu again, recommending immediate display of original options after dialogue box content appears.
- Principle 6 (80% compliance): Same as evaluator 2's feedback in (1) "principle 6 (80% compliance)."
- Principle 7 (80% compliance): Same as evaluator 2's feedback in (1) "principle 7 (80% compliance)."

3. Record Values

- Principle 3 (80% compliance): Evaluator 1 noted that after entering values, the entire app needed closing to return to the main screen and suggested adjusting this.
- Principle 9 (33% compliance): Evaluator 5 noticed that there were no notifications for abnormal physiological values and suggested providing automatic or manual messages based on professional or objective facts for abnormal values.

4. Reminder Settings

- Principle 1 (80% compliance): No specific issues or suggestions were mentioned by the evaluator.
- Principle 3 (60% compliance): Evaluator 5 discovered that alterations in the reminder function required notification to researchers and suggested allowing users to freely set daily or appointment reminders, similar to setting alarms on a phone.
- Principle 7 (60% compliance): Same as evaluator 5's feedback in (4) "principle 3 (60% compliance)."

5. Kidney Health Mission

- Principle 1 (80% compliance): Evaluator 5 felt that the current daily diet recording was based on self-assessment without professional guidance and suggested having professionals review information and provide guidance.
- Principle 2 (80% compliance): Same as evaluator 5's feedback in (5) "principle 1 (80% compliance)."
- Principle 9 (33% compliance): Same as evaluator 5's feedback in (3) "principle 9 (33% compliance)."

6. Inquiry & Consultation

- Principle 3 (75% compliance): Same as Evaluator 2's feedback in (2) "principle 3 (80% compliance)."
- Principle 6 (75% compliance): Same as evaluator 2's feedback in (1) "principle 6 (80% compliance)."
- Principle 7 (80% compliance): Same as evaluator 2's feedback in (1) "principle 7 (80% compliance)."

Overall, based on the evaluation scores and feedback from evaluators on each function of the digital platform concerning various principles, this phase successfully acquired specific recommendations for improving the platform's usability. These suggestions can generally be categorized into three aspects: (1) enhancing user operation convenience, such as adding a return option in Health Education Information, Kidney Support Teammate, and Inquiry & Consultation, reducing repetitive operation hassles; (2) pushing more relevant and engaging information, for instance, pushing corresponding educational content based on user records; and (3) enhancing interactivity and feedback of functions, such as providing personalized memory functions for highly clicked content, warning after inputting abnormal physiological values, or providing feedback from professionals. Results of the heuristic evaluation have guided the modification of the digital platform for the next iteration to enhance user experience.

Phase 3: Usability Testing

Following the design of tasks, the selection of participants for phase 3 involved convenience sampling by inviting 5 dyads who participated in phase 1 trial use during their next clinic visits. Eventually, dyad numbers 001, 002, 005, 007, and 009 agreed to take part in the testing. The results of their average performance on our designed tasks and their average ASQ ratings are summarized in Table 11.

Table . Phase 3 usability testing task performance and rating summary (N=5).

	Functions of the digital platform					
	Health Education Information	Kidney Support Teammate	Record Values	Reminder Settings	Kidney Health Mission	Inquiry & Consultation
Task success[a]	60%	80%	100%	100%	100%	80%
Task time (mm:ss)	04:23	05:16	01:06	01:52	01:04	02:55
Average error times (range)	1 (0 - 2)	1.4 (1-2)	0.2 (0 - 1)	0.8 (0 - 1)	0	1.2 (0 - 2)
ASQ: ease of use[b]	3.6	4	6.2	5.4	6	4.2
ASQ: time required	3.4	4.2	6.4	5.2	6.2	4.2
ASQ: help received	3.8	3.8	6	4.2	5.2	3.6
Average ASQ score	3.6	4	6.2	4.9	5.8	4

[a]Task success = (Number of dyads successfully completed the task/N) × 100%.

[b]ASQ: After-Scenario Questionnaire. The score of ASQ ranges from 1 to 7, with higher score indicating higher level of satisfaction.

XSL•FO
RenderX

Task Performance and ASQ

This section describes the designed task scenario for each function and summarizes participants' general task performance as well as their ASQ ratings:

1. Health Education Information: The task scenario reflected the need for patients with CKD to calculate their estimated glomerular filtration rate (eGFR) independently. To accomplish this task, users can navigate through the Health Education Information section in our LINE OA to find a link to a web page with an eGFR calculator. Results indicated that 3 out of 5 dyads completed this task successfully (success rate 60%), with an average completion time of approximately 4.5 minutes. Operational errors were observed with numbers 002, 005, and 009, who navigated to the wrong sections, causing them additional time to locate the eGFR calculator. Another error occurred when number 007 could not find the box to enter the value of creatinine (CREA), as the enlarged font size setting obscured a clear presentation of the calculator elements on the smartphone. The average ASQ score for this task is 3.6, which is the lowest score among all tasks, indicating that dyads were the least satisfied with the ease and the time required for this task. They also perceived minimal help from the user manual.

2. Kidney Support Teammate: The task scenario involved an SO finding effective ways to support a patient who is attempting lifestyle modification. The Kidney Support Teammate in our LINE OA provides such information for SOs to implement supportive strategies. The task success rate is 80%, with only 1 dyad failing to complete this task. The average completion time exceeded 5 minutes, which is the longest among all tasks. All dyads took time to navigate different sections before they located the target information for this task. Dyads numbers 005 and 009 experienced uncertainty as they browsed through the correct section but did not recognize the target information. The average ASQ score is 4, indicating a moderate but relatively low level of user satisfaction compared with the other tasks.

3. Record Values: The task scenario reflects the need for patients with CKD to monitor their own physiological data. Completing this task requires the user to click "Record Values" on the service menu of our LINE OA or click "New Entry" on the extended app and input the numbers. The task success rate is 100%, with an average completion time of only 1 minute. Dyad number 007 encountered initial difficulty when trying to input the CREA value. The average ASQ score is 6.2, which is the highest among all tasks. These results suggested that this function is user-friendly and intuitive, leading to high user satisfaction.

4. Reminder Settings: The task scenario arises when patients with CKD wish to set a reminder for their upcoming clinical appointment. Dyads will need to input the date and time of their next appointment using the extended app and notify the research team via LINE OA to set an automatic message reminder. The task success rate is 100%, with an average completion time of around 2 minutes. Dyads numbers 002 and 007 seemed unfamiliar with this function as they clicked "New Entry" on the extended app, which redirected them to the wrong page. The average ASQ score is 4.9, indicating a moderate satisfaction level with this function.

5. Kidney Health Mission: The task scenario revolves around the essential practice of tracking daily diet, a key aspect of managing CKD. Users can access this task by clicking "Kidney Health Mission" on the service menu of our LINE OA or selecting "Kidney Health Diary" on the extended app to upload a photograph or text note regarding their diet. The task success rate is 100%, with an average completion time of just 1 minute. All dyads completed the task without encountering any operational errors, indicating that the design of this function is sufficiently intuitive. The average ASQ score is 5.8, demonstrating a relatively high level of satisfaction compared with the other tasks.

6. Inquiry & Consultation: This task scenario arises when patients with CKD encounter health-related inquiries at home without access to a health care professional. Users can explore the frequently asked question section in our LINE OA to seek answers to their questions. For instance, this task involves users wanting to learn about the symptoms of hypoglycemia (low blood sugar) using our platform. The task success rate is 80%, with dyad number 002 failing to complete the task. On average, it takes around 3 minutes to complete the task, as all dyads navigate through different sections of our LINE OA to find the desired information. Errors occurred when all dyads, except number 001, attempted to locate the answer in the Health Education Information section. The average ASQ score is 4, indicating a moderate but relatively low level of user satisfaction, highlighting areas for usability improvement.

Overall Platform Usability

The SUS scores provided by the 5 dyads range from 57.5 to 87.5, with an average score of 67.5, just below the industry standard threshold of 68. This indicates that users found the platform to be marginally acceptable in terms of usability, indicating areas for improvement. In combination with the performance and ASQ ratings of the 6 tasks, it was observed that the 3 tasks involving operations on the extended app (ie, Record Values, Reminder Settings, and Kidney Health Mission) all achieved a 100% success rate, with shorter completion times, fewer operational errors, and higher satisfaction (mean average ASQ score: 5.64 vs 3.87 for the other 3 tasks). Conversely, tasks with lower performance and ASQ scores (ie, Health Education Information, Kidney Support Teammate, and Inquiry & Consultation) all involved finding information within LINE OA, suggesting that underlying significant usability issues in this area may have contributed to the low overall platform SUS score.

Discussion

Principal Results

In this study, we demonstrated that through an instant messaging app, our digital platform aimed at empowering lifestyle modification for CKD dyads holds promise for its cost-effectiveness, high user engagement, and potential clinical practicality. Key insights emerged from the 3-phase development cycle are as follows.

First, the application of Agile methods facilitated a flexible and efficient process in developing our digital platform. Notably,

XSL•FO
RenderX

the platform design and deployment were accomplished within a remarkably short time frame of 3 months, with the ability to accommodate modifications, typically within a week. Such adaptability of our platform ensured responsiveness to user needs throughout the development cycle.

Second, guided by the DDEF, this study introduced 2 quantifiable daily self-management tasks—logging physiological values and maintaining a kidney health diary—to address the multifaceted nature of CKD management. These tasks, coupled with behavior change techniques, aimed to promote sustained user engagement. The EAR of these tasks emerged as a simple, innovative, and promising indicator of users' lifestyle modification abilities, which could facilitate more precise tailored interventions based on different user activity levels.

Third, comprehensive evaluation across 3 phases revealed high user acceptance and satisfaction with our digital platform. However, usability challenges, particularly concerning information layout within the LINE OA, were identified. The absence of CKD dyad involvement during the initial prototype development hindered the alignment of the content of LINE OA service menu with varying user mental models, potentially causing users to struggle with navigation [37]. This finding underscores the importance of considering user expectations early in the development cycle to ensure optimal usability.

Comparison With Prior Work

Previous DHIs for CKD management have adopted various approaches, including a mobile app [15], a web-based platform [17], a chatbot [21], and the combination of a mobile app and wearable devices with instant messaging support [20]. These interventions have demonstrated potential in supporting CKD health education and self-management. However, challenges were faced such as limited evidence of sustained use—many reported low [15] or unclear usage rates [17,20,21]—and participant samples skewed toward younger, tech-savvy patients (mean age around 50 years), due to the necessity of internet access and smartphone proficiency. Moreover, prior systems typically focused on a single aspect of care (education or tracking) and rarely involved patients' family members, potentially excluding older patients who require caregiver support. In contrast, our LINE-based digital dyadic empowerment platform addresses these gaps through the following strengths mentioned in the subsections.

Diverse Participant Backgrounds

Our study encompassed participants with diverse backgrounds, ranging from health care providers (including nephrologists and nurses) within the institution involved in phase 1 system development, to patients with CKD and SOs engaged in both phase 1 platform trial use and phase 3 usability testing. Health care and electrical engineering experts from external institutions were involved in phase 2 heuristic evaluation. Of these participants, 41.7% (10/24) were male, spanning ages from 28 to 82 years, and educational attainment ranging from elementary school to doctoral degrees. Such diversity maximized the input from all potential stakeholders who could use or benefit from our digital platform. We ensured that various perspectives were considered throughout the development cycle, thereby optimizing platform usability while accommodating user differences.

Notably, the design of including SOs with patients with CKD in our study was pivotal. It not only enriched the diversity of perspectives but also prevented the exclusion of patients who might face challenges participating independently, such as those from older age groups or with lower educational levels, potentially constituting digitally disadvantaged groups. This inclusive approach, exemplified by dyads numbers 004 and 005 (Table 5), served to mitigate selection bias and promote health equity at the digital resource level.

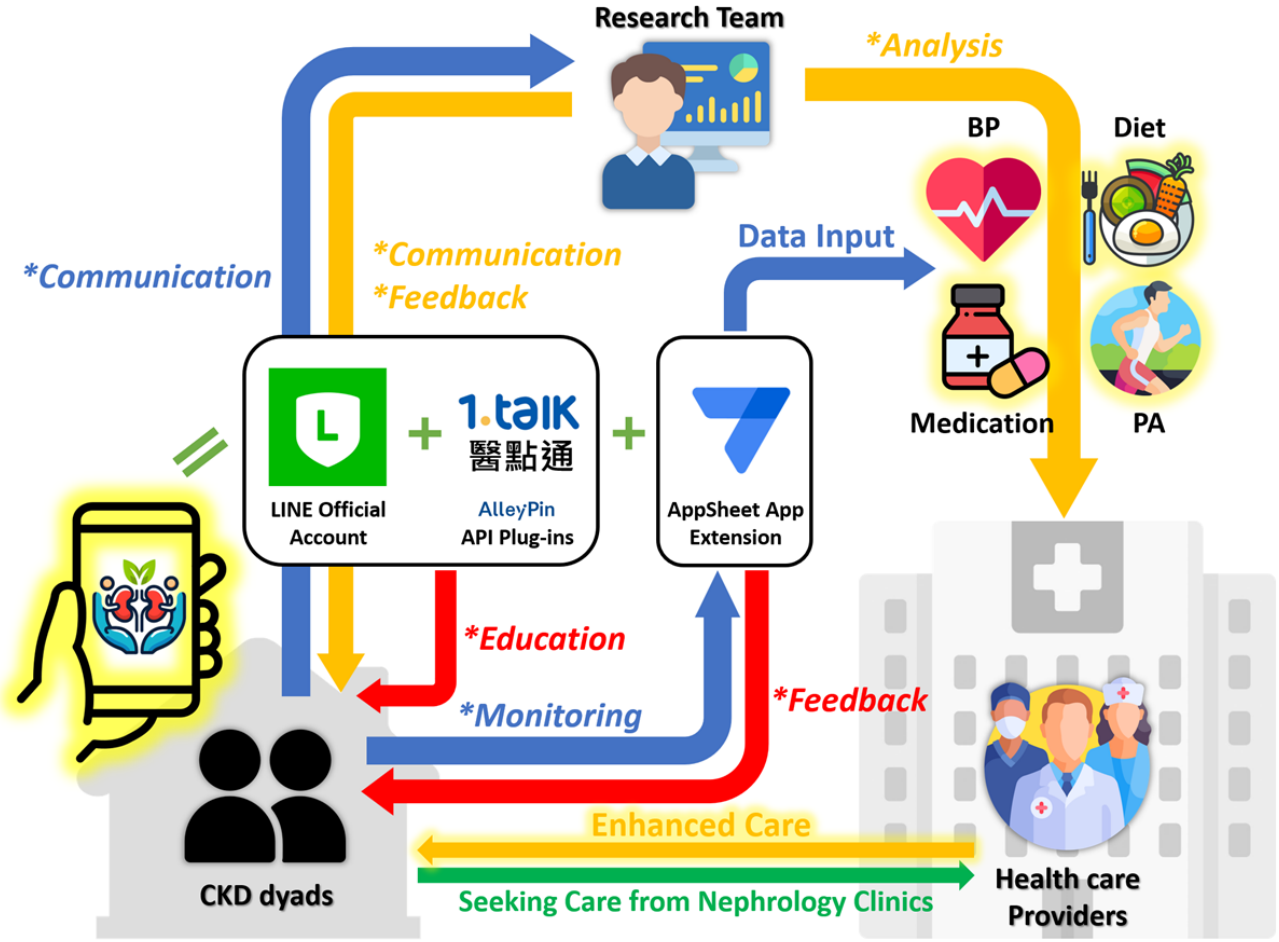
A Scalable and Practical Digital Platform

In a 2023 internet survey conducted in Taiwan, it was found that 98.51% of internet users use instant messaging tools, with LINE holding a dominant market share of 91.61% [38]. Our digital platform operates within Taiwan's most prevalent instant messaging app, significantly improving its accessibility and ubiquity. This integration closely aligns CKD self-management with the most frequent daily use of digital technology. Data extracted from the backend of the LINE OA further revealed consistent user engagement with our digital platform, with more than 90% of interactions being automatic Push or Reply messages (Table 6). This underscores the high automation and low-cost model of our digital platform, demonstrating its potential for scalable application in medical settings.

Figure 5 further illustrates the interactive relationship between CKD dyads, the research team, and health care providers, elucidating how our digital platform can enhance CKD care in medical settings. The green arrow at the bottom represents patients with CKD and their SOs seeking routine care at the hospital. Blue arrows depict CKD dyads using LINE OA, either through encouragement or voluntary interaction, to stay connected with the research team (*communication* strategy); or using the extended app to log home-based health records, including blood pressure, diet, medication, and physical activity (*monitoring* strategy). Red arrows signify the fully automated services offered by our digital platform, such as providing health education information through LINE OA Push or Reply messages (*education* strategy); or accessing physiological trend charts and usage reports via the extended app (*feedback* strategy). Orange arrows indicate the interventions that the research team can deliver through the digital platform, such as maintaining contact with CKD dyads (*communication* strategy); or answering CKD dyads' queries and sending encouraging messages (*feedback* strategy) via LINE OA. Finally, the research team can further analyze and aggregate home-based self-monitoring records logged by CKD dyads through the extended app (*analysis* strategy) and provide reports to health care providers. This additional information will furnish health care providers with supplementary insights during CKD dyads' visits, aiding in treatment recommendations to enhance the quality of CKD care.

XSL•FO
RenderX

Figure 5. The schematic diagram of the LINE Official Account "Kidney Lifestyle" and the extended app in enhancing CKD care. The arrows marked with an asterisk (*) represent the practical application of the 5 digital strategies mentioned in the Digital Dyadic Empowerment Framework. Blue arrows represent the active engagement from CKD dyads. Orange arrows denote service flows involving human support, while red arrows denote service flows with full automation. BP: blood pressure; CKD: chronic kidney disease; PA: physical activity.



Empirical Evidence for Sustained User Engagement

This study provides empirical evidence of high user acceptance and satisfaction with our digital platform, as indicated by self-reported satisfaction ratings and intentions to continue usage during phase 1 trial use, as well as by the ASQ scores obtained from phase 3 usability testing. While self-reported ratings could potentially be inflated by demand characteristics [39,40], the robustness of these findings is further supported by users' consistent execution of the 2 daily self-management tasks. As depicted in Table 7, more than half of the 10 trial users maintained an EAR from 50% to 100% for the 2 daily tasks for more than 6 months. Notably, 5 users had a record of executing the 2 daily tasks consecutively for 30 days or more, demonstrating their ability to sustain regular usage. These findings imply their persistent commitment to using our digital platform for managing CKD, even after they completed the study. Such ongoing engagement serves as compelling evidence of our digital platform's potential as a highly acceptable digital health solution, empowering CKD dyads in the process of lifestyle modification.

Limitations

Potential Selection Bias Induced by Digital Divides

Despite the wide availability of our digital platform on LINE OA and the inclusive dyadic design, this study might still have suffered from selection bias due to digital divides between those with access to our platform and those without [41]. During phase 1 participant recruitment, we encountered instances where some CKD dyads, particularly spouses, reported lacking access to smartphones or the internet. Conversely, some dyads had smartphones with internet access but lacked diverse digital skills or did not have LINE installed on their phones. All these factors could impede the implementation of our study and raise concerns over health inequity. Although the extent of these digital divides was not thoroughly investigated within the scope of this study, their existence warrants careful consideration in our future work to prevent selection bias favoring participants with higher digital proficiency.

For participants facing difficulties in using our platform, personalized assistance will be provided through interactive learning sessions and the user manual, reducing the likelihood of their attrition from the study. Furthermore, future research could investigate the prevalence and characteristics of digitally disadvantaged groups within the CKD population. Assessing the severity and underlying causes of digital divides may help

XSL•FO
RenderX

mitigate their impact on the implementation of our DHI. Alternatively, programs empowering lifestyle modification tailored to CKD dyads who are not internet or smartphone users could be developed.

Usability Issues With LINE OA Navigation

Despite continuous optimization, the SUS scores obtained from phase 3 usability testing indicate that the usability of our digital platform remains marginally acceptable. Particularly, significant usability issues were identified within the information layout of the service menu on LINE OA. During phase 1 trial use, dyads reported difficulties navigating sections including Health Education Information, Kidney Support Teammate, and Inquiry & Consultation, often getting lost. Similarly, task performance and user satisfaction related to these sections were relatively low in phase 3 usability testing. However, tracking individual user interactions with specific sections in LINE OA is challenging due to its backend limitations, as only aggregate user activity data are available. This impedes the establishment of personalized sections, as suggested by the heuristic evaluators in phase 2.

Analysis of phase 3 error records further revealed 2 major sources of confusion (Multimedia Appendix 4). First, users were often unsure whether to seek answers in Health Education Information or Inquiry & Consultation, as both sections were perceived to serve a similar purpose of obtaining CKD-related knowledge. This mismatch between users' mental models and the platform's menu structure frequently led them to attempt tasks in the wrong section, causing navigation failures. Second, the design of LINE OA as an instant messaging interface posed inherent challenges: the service menu occupies nearly one-third of the small mobile screen, while new messages push earlier information upward, forcing users to scroll back or reselect menu buttons to revisit prior content. When combined with multilayered menu depth, these constraints increased cognitive load and disorientation, further compounding the likelihood of users getting lost.

To overcome these challenges, several measures have been introduced and others are planned. In response to phase 2 feedback, we incorporated a "return to previous level" option within the Kidney Support Teammate and Inquiry & Consultation sections (Figure 2D), providing a route back when disoriented. Future platform refinements could also consolidate Health Education Information and Inquiry & Consultation into a single section, for instance, by embedding Inquiry & Consultation under Health Education Information to reduce functional overlap and user confusion. Moreover, involving CKD dyads in early stages of prototype development remains essential. Interviews on information needs and common questions could align menu structures more closely with users' mental models. Finally, harnessing the instant messaging nature of LINE OA, the inclusion of a chatbot could provide more efficient and context-sensitive access to disease-related information through automated dialogue interactions, thereby enhancing overall usability [21].

Concerns for Low EARs

Despite our efforts to simplify the design of our digital platform, reducing hierarchies and operational paths, along with manual support and relevant skill teachings, some users, as shown in Table 7, have an activity rate below 50% for logging physiological values (2 users) or writing kidney diaries (4 users). We identified 3 potential reasons for these low EARs, indicating the need for more comprehensive interventions as mentioned in the following sections.

Dyads in the Preaction Stages of Change

Observations from Tables 5 and 7 revealed a positive relationship between dyads' self-reported readiness for lifestyle modification (ie, stages of change) and their execution of daily tasks. The 5 most active dyads, with EARs exceeding 75% for both tasks and having achieved the *Master* kidney health title, were all in the late stages of change (*action* or *maintenance*), while of the other 5 dyads, 3 were in the preaction stages (*contemplation* or *preparation*). Despite daily task reminders, there was no observed increase in activity rates among lower-activity users. Future interventions may focus on leveraging helping relationships with SOs and providing proactive caring phone calls to enhance their motivation for change.

Dyads Perceive No Need for Use

Among the previously mentioned 5 dyads with lower activity rates, 2 self-reported being in the late stage of change (*maintenance*). These dyads might have already engaged in regular lifestyle modifications but perceive minimal help from our digital platform and hence feel no necessity in using it. To address this, future implementation could focus on highlighting the platform's strengths and collaborating with health care providers to promote the platform's benefits to patients and their caregivers.

Dyads Face Usage Obstacles and Perceive High Execution Difficulty

Another possible cause of low EARs is shortcomings in the digital platform's design and user manual, leading dyads to perceive high cognitive load and execution difficulty. To address this, helping relationships from SOs and providing assistance are crucial. Regular caring and reminder messages, platform design optimization, and personalized teaching during return visits can help address usage obstacles and enhance usability.

Another promising future direction is to assess the feasibility of organizing support groups for CKD dyads, based on our dyadic feedback in phase 1. Support groups may facilitate the process of social liberation, exposing dyads to a supportive social environment for behavior change [42]. Dyads in the preaction stages might find positive support from the group, enhancing their motivation for change. Support groups can also promote community participation through connections and information exchange and allow their efforts to gain recognition and reinforcement from the group. For dyads perceiving minimal help from the digital platform or finding it too challenging to use, they can observe and learn from other dyads in the support groups, enhancing their confidence and skills in usage.

XSL•FO
RenderX

Lack of Feedback From Health Care Providers

In phase 1 dyad trial use and phase 2 heuristic evaluation, participants highlighted the need for health care provider feedback on user self-monitoring records. However, our digital platform currently lacks feedback mechanisms beyond system-generated physiological trend charts and platform usage reports from the extended app (Figures 3A and 3C). Specifically, feedback from health care providers on dietary, medication, and exercise content in the kidney health diary is absent. As observed in Figure 5, health care providers currently do not interact with CKD dyads through our digital platform and are thus unable to provide feedback on the content of the kidney health diary. This gap may contribute to lower activity rates for the kidney health diary compared with physiological values (Table 7). Workflow integration challenges are common in the development and application of new digital health tools, as they might potentially burden health care providers with additional tasks or even necessitate practice reorganization [15,43].

Recognizing the temporary difficulty in overcoming this challenge, an alternative approach involves compiling CKD dyads' self-monitoring records into paper-based reports for health care providers. These reports are provided to health care providers on CKD dyads' follow-up visit days for reference. During consultations, health care providers can then access the extended app with CKD dyads to review the kidney health diary content as needed. While conducting trials to assess the effect of our digital platform on CKD management is imperative, our future work should also tackle workflow integration issues through close collaboration with health care system personnel to promote platform adoption.

Conclusions

Guided by DDEF and using Agile methods, this study successfully developed, evaluated, and optimized an instant messaging–based digital platform tailored for CKD care across a 3-phase development cycle. The resulting product, the LINE OA "Kidney Lifestyle" and its associated extended app, aims to empower lifestyle modification for CKD dyads. Our investigation uncovered robustly high user acceptance and satisfaction with the platform, evidenced by both self-report ratings and ongoing engagement beyond the study period, lasting over 6 months. The platform's wide availability, high automation, and cost-effective operational model underscore its potential for scalable application in practical health care settings. While our digital solution shows considerable promise, a feasibility study is the logical next step to evaluate implementation factors and potential impacts on CKD management outcomes. Meanwhile, our future works should also bridge digital divides, enhance usability, foster support groups, and integrate our platform into current health care workflow to maximize its effectiveness and scalability. In sum, this study contributes significantly to the growing field of digital health for CKD management by showcasing the potential of an innovative platform integrated with the LINE instant messaging app.

Acknowledgments

The authors would like to thank all participants who took part in this study. The authors are especially grateful to the *Taiwan Kidney Foundation* for authorizing the use of their official educational resources within the platform developed in this study. The authors also extend their sincere thanks to Ms Yu-Chen Kuo, health education nurse at the Health Education Center of National Cheng Kung University Hospital, and to their technical development partner, AlleyPin Interactive Co, LTD, for their valuable suggestions and support throughout the platform development process. The authors additionally thank 2 research assistants, Jung-Jung Liang and Wei-Tung Kao, for their assistance in participant enrollment and data processing. During intervention development, OpenAI's ChatGPT was used to assist in the creation of Google Apps Script functions for backend automation, such as calculating daily task patterns and simulating gamification logic. It was also used to improve language clarity and structure in the manuscript. All outputs were reviewed and edited by the authors, who assume full responsibility for the final content.

Funding

This study was supported by the National Health Research Institutes, Taiwan, through the Integrated Research Grants in Health and Medical Sciences (Innovative Research Grant, grant no. NHRI-EX112-11106PI). The funding agency had no role in the design and conduct of the study; collection, management, analysis, and interpretation of data; or preparation, review, and approval of the manuscript.

Data Availability

The complete dataset generated during this study—including raw item-level feedback from phase 1 dyads, heuristic evaluation scores and qualitative comments from phase 2 experts, and individual-level task performance metrics and questionnaire responses from phase 3 usability testing—is available in Multimedia Appendix 4. Multimedia Appendix 4 also contains summary tables that correspond to the results reported in Tables 8, 10, and 11 of the manuscript.

Authors' Contributions

MY conceived the study idea, developed the initial study design, secured research funding, and supervised the overall implementation of the study. CYH optimized the study design based on practical considerations, led the development and integration of the digital platform, and was responsible for participant recruitment, data collection, analysis, interpretation, and

XSL•FO
RenderX

manuscript preparation. DS contributed to the initial drafting of the "Introduction" section. JMS, MCW, and WHL provided suggestions on platform design and assisted with participant referral. All authors critically reviewed and revised the manuscript for intellectual content and approved the final version for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
User manual.
[PDF File, 2359 KB - humanfactors_v12i1e73935_app1.pdf]

Multimedia Appendix 2
Phase 1 user feedback form.
[PDF File, 232 KB - humanfactors_v12i1e73935_app2.pdf]

Multimedia Appendix 3
Phase 3 usability testing instrument.
[PDF File, 268 KB - humanfactors_v12i1e73935_app3.pdf]

Multimedia Appendix 4
Complete dataset and summary tables.
[XLSX File, 39 KB - humanfactors_v12i1e73935_app4.xlsx]

References

1. Bikbov B, Purcell CA, Levey AS. Global, regional, and national burden of chronic kidney disease, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2020 Feb 29;395(10225):709-733. [doi: 10.1016/S0140-6736(20)30045-3] [Medline: 32061315]
2. Lee CC, Hsu CC, Lin MH, Sung JM, Kuo TH. Healthcare utilization and expenditure among individuals with end-stage kidney disease in Taiwan. J Formos Med Assoc 2022 Feb;121 (Suppl 1):S47-S55. [doi: 10.1016/j.jfma.2021.12.010] [Medline: 34980549]
3. Megari K. Quality of life in chronic disease patients. Health Psychol Res 2013 Sep 24;1(3):e27. [doi: 10.4081/hpr.2013.e27] [Medline: 26973912]
4. Neale EP, Rosario VD, Probst Y, Beck E, Tran TB, Lambert K. Lifestyle interventions, kidney disease progression, and quality of life: a systematic review and meta-analysis. Kidney Med 2023 Jun;5(6):100643. [doi: 10.1016/j.xkme.2023.100643] [Medline: 37235039]
5. Chen J, Fowler KJ, Grams ME. Knowledge is power: patient education as a tool for patient activation. Am J Kidney Dis 2020 Aug;76(2):163-165. [doi: 10.1053/j.ajkd.2020.03.012] [Medline: 32560974]
6. Hirsch JS, Danna SC, Desai N, et al. Optimizing care delivery in patients with chronic kidney disease in the United States: proceedings of a multidisciplinary roundtable discussion and literature review. J Clin Med 2024 Feb 20;13(5):1206. [doi: 10.3390/jcm13051206] [Medline: 38592013]
7. Wen CP, Cheng TYD, Tsai MK, et al. All-cause mortality attributable to chronic kidney disease: a prospective cohort study based on 462 293 adults in Taiwan. Lancet 2008 Jun 28;371(9631):2173-2182. [doi: 10.1016/S0140-6736(08)60952-6] [Medline: 18586172]
8. Chao S, Yen M, Lin TC, Sung JM, Wang MC, Hung SY. Psychometric Properties of the Kidney Disease Quality of Life-36 Questionnaire (KDQOL-36. West J Nurs Res 2016 Aug;38(8):1067-1082. [doi: 10.1177/0193945916640765] [Medline: 27026384]
9. Chen TK, Knicely DH, Grams ME. Chronic kidney disease diagnosis and management: a review. JAMA 2019 Oct 1;322(13):1294-1304. [doi: 10.1001/jama.2019.14745] [Medline: 31573641]
10. Krishnamurthy S, Ks K, Dovgan E, et al. Machine learning prediction models for chronic kidney disease using national health insurance claim data in Taiwan. Healthcare (Basel) 2021 May 7;9(5):546. [doi: 10.3390/healthcare9050546] [Medline: 34067129]
11. Daniels K, Bonnechère B. Harnessing digital health interventions to bridge the gap in prevention for older adults. Front Public Health 2023;11:1281923. [doi: 10.3389/fpubh.2023.1281923] [Medline: 38259780]
12. Classification of digital health interventions v10: a shared language to describe the uses of digital technology for health. World Health Organization. 2018. URL: https://www.who.int/publications/i/item/WHO-RHR-18.06 [accessed 2025-11-12]
13. Jeddi FR, Nabovati E, Amirazodi S. Features and effects of information technology-based interventions to improve self-management in chronic kidney disease patients: a systematic review of the literature. J Med Syst 2017 Sep 18;41(11):170. [doi: 10.1007/s10916-017-0820-6] [Medline: 28920174]

XSL•FO
RenderX

14. Ren Q, Lian M, Liu Y, Thomas-Hawkins C, Zhu L, Shen Q. Effects of a transtheoretical model-based WeChat health education programme on self-management among haemodialysis patients: a longitudinal experimental intervention study. J Adv Nurs 2019 Dec;75(12):3554-3565. [doi: 10.1111/jan.14182] [Medline: 31441525]
15. Hsiao PN, Hsu P, Hsiao SM, Chen TH, Lee CL. Project to increase the usage rate of iCKD application mobile healthcare on chronic kidney disease patients. Hu Li Za Zhi 2019 Feb;66(1):84-92. [doi: 10.6224/JN.201902_66(1).10] [Medline: 30648248]
16. Yang Y, Chen H, Qazi H, Morita PP. Intervention and evaluation of mobile health technologies in management of patients undergoing chronic dialysis: scoping review. JMIR Mhealth Uhealth 2020 Apr 3;8(4):e15549. [doi: 10.2196/15549] [Medline: 32242823]
17. Donald M, Beanlands H, Straus S, et al. An eHealth self-management intervention for adults with chronic kidney disease, My Kidneys My Health: a mixed-methods study. CMAJ Open 2022;10(3):E746-E754. [doi: 10.9778/cmajo.20210332] [Medline: 35973709]
18. LINE Business Guide (Summary) v11. LY Corporation. 2024. URL: https://www.linebiz.com/sites/default/files/media/jp-en/download/EN_LINE%20Business%20Guide_202304-09_summary.pdf [accessed 2025-04-15]
19. 2022 Taiwan Internet Report. Taiwan Network Information Center. 2022. URL: https://report.twnic.tw/2022 [accessed 2025-04-15]
20. Li WY, Chiu FC, Zeng JK, et al. Mobile health app with social media to support self-management for patients with chronic kidney disease: prospective randomized controlled study. J Med Internet Res 2020 Dec 15;22(12):e19452. [doi: 10.2196/19452] [Medline: 33320101]
21. Chen NJ, Huang CM, Fan CC, et al. User evaluation of a chat-based instant messaging support health education program for patients with chronic kidney disease: preliminary findings of a formative study. JMIR Form Res 2023 Sep 19;7:e45484. [doi: 10.2196/45484] [Medline: 37725429]
22. Kushniruk AW, Patel VL. Cognitive and usability engineering methods for the evaluation of clinical information systems. J Biomed Inform 2004 Feb;37(1):56-76. [doi: 10.1016/j.jbi.2004.01.003] [Medline: 15016386]
23. Serrador P, Pinto JK. Does Agile work? A quantitative analysis of agile project success. Int J Proj Manag 2015 Jul;33(5):1040-1051. [doi: 10.1016/j.ijproman.2015.01.006]
24. 用LINE官方帳號做到約診提醒-醫療業 [Using LINE Official Account for appointment reminders—medical industry]. LY Corporation. URL: https://tw.linebiz.com/smb/industry-application/medical/ [accessed 2025-04-15]
25. AppSheet: the fastest way to build apps and automate work. Google. URL: https://about.appsheet.com/home/ [accessed 2025-04-15]
26. Nielsen J. Heuristic evaluation. In: Nielsen J, Mack RL, editors. Usability Inspection Methods: John Wiley & Sons; 1994:25-62.
27. Moran K, Gordon K. How to conduct a heuristic evaluation. Nielsen Norman Group. 2023 Jun 25. URL: https://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation/ [accessed 2025-04-15]
28. Nielsen J. 10 usability heuristics for user interface design. Nielsen Norman Group. 1994 Apr 24. URL: https://www.nngroup.com/articles/ten-usability-heuristics/ [accessed 2025-04-15]
29. Donald M, Beanlands H, Straus SE, et al. A web-based self-management support prototype for adults with chronic kidney disease (My Kidneys My Health): co-design and usability testing. JMIR Form Res 2021 Feb 9;5(2):e22220. [doi: 10.2196/22220] [Medline: 33560245]
30. Moran K. Usability testing 101. Nielsen Norman Group. 2019 Dec 1. URL: https://www.nngroup.com/articles/usability-testing-101/ [accessed 2025-04-15]
31. Nielsen J. Why you only need to test with 5 users. Nielsen Norman Group. 2000 Mar 18. URL: https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/ [accessed 2025-04-15]
32. Lewis JR. IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. Int J Hum Comput Interact 1995 Jan;7(1):57-78. [doi: 10.1080/10447319509526110]
33. Brooke J. SUS—a quick and dirty usability scale. In: Jordan PW, editor. Usability Evaluation in Industry: Taylor & Francis; 1996:189-194.
34. Sauro J. Measuring usability with the System Usability Scale (SUS). Measuring U. 2011 Feb 3. URL: https://measuringu.com/sus/ [accessed 2024-04-11]
35. Wang SL, Kung LF, Chen TH, Hsiao SM, Hsiao PN, Chiou CJ. Construction and validation of a chronic kidney disease self-care scale. Hu Li Za Zhi 2016 Aug;63(4):90-99. [doi: 10.6224/JN.63.4.90] [Medline: 27492299]
36. Lin CY, Chang JH, Shih CD. Assessment of medication adherence in elderly residents in Pingtung City. J Taiwan Pharm 2019;35(1) [FREE Full text]
37. Norman D. The psychopathology of everyday things. In: The Design of Everyday Things: Revised and Expanded: Basic Books; 2013:1-34.
38. Taiwan internet report. Taiwan Network Information Center. 2023. URL: https://report.twnic.tw/2023 [accessed 2025-04-15]
39. Demand characteristics. Interaction Design Foundation. 2015 Jul 5. URL: https://www.interaction-design.org/literature/book/the-glossary-of-human-computer-interaction/demand-characteristics [accessed 2025-04-15]

XSL•FO
RenderX

40. Orne MT, Whitehouse WG. Demand characteristics. In: Kazdin AE, editor. Encyclopedia of Psychology: American Psychological Association and Oxford University Press; 2000:469-470. [doi: 10.1037/10517-173]
41. Ragnedda M, Muschert GW, editors. The Digital Divide: The Internet and Social Inequality in International Perspective, 1st edition: Routledge; 2013. [doi: 10.4324/9780203069769]
42. Banbury A, Nancarrow S, Dart J, Gray L, Parkinson L. Telehealth interventions delivering home-based support group videoconferencing: systematic review. J Med Internet Res 2018 Feb 2;20(2):e25. [doi: 10.2196/jmir.8090] [Medline: 29396387]
43. Steele Gray C, Gill A, Khan AI, Hans PK, Kuluski K, Cott C. The electronic patient reported outcome tool: testing usability and feasibility of a mobile app and portal to support care for patients with complex chronic disease and disability in primary care settings. JMIR Mhealth Uhealth 2016 Jun 2;4(2):e58. [doi: 10.2196/mhealth.5331] [Medline: 27256035]

Abbreviations

API: application programming interface
ASQ: After-Scenario Questionnaire
CKD: chronic kidney disease
CREA: creatinine
DDEF: Digital Dyadic Empowerment Framework
DHI: digital health intervention
EAR: entry activity rate
eGFR: estimated glomerular filtration rate
OA: LINE Official Account
SO: significant other
SUS: System Usability Scale

Edited by E Borycki; submitted 08.05.25; peer-reviewed by M Gasmi, R Wiraguna; accepted 21.10.25; published 26.11.25.

Please cite as:
Ho CY, Siregar D, Yen M, Sung JM, Wang MC, Lin WH
A Messaging App Empowering Lifestyle Modification in Chronic Kidney Disease (LINE Official Account "Kidney Lifestyle"): Platform Development and Usability Study
JMIR Hum Factors 2025;12:e73935
URL: https://humanfactors.jmir.org/2025/1/e73935
doi:10.2196/73935

© Chun-Yi Ho, Deborah Siregar, Miaofen Yen, Junne-Ming Sung, Ming-Cheng Wang, Wei-Hung Lin. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 26.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Impact of Digital Phenotypes and Question-Asking on Emotional Disorders in Adolescents: 4-Week Field Study

Minseo Cho[1], MEng; Doeun Park[2], PhD; Myounglee Choo[3], MEng; Doug Hyun Han[4], MD, PhD; Jinwoo Kim[1,3], PhD

[1]Human Computer Interaction Lab, School of Business, Yonsei University, 50, Yonsei-ro, Seodaemun-gu, Building 212, Seoul, Republic of Korea

[2]School of Management of Technology, Human Computer Interaction Lab, Yonsei University, Seoul, Republic of Korea

[3]HAII Corporation, Seoul, Republic of Korea

[4]Department of Psychiatry, College of Medicine, Chung-Ang University, Seoul, Republic of Korea

Corresponding Author:
Jinwoo Kim, PhD
Human Computer Interaction Lab, School of Business, Yonsei University, 50, Yonsei-ro, Seodaemun-gu, Building 212, Seoul, Republic of Korea

Abstract

Background: Adolescence is the period with the highest incidence of mental disorders, with approximately one-third, half, and two-thirds of cases emerging by ages 14, 18, and 25 years, respectively. Proactive interventions are essential, and digital phenotyping has emerged as a promising approach for timely detection and management. However, passive digital phenotyping is limited to sensor-detectable behaviors, while active phenotyping is often confined to clinical scales, missing the opportunity to capture users' subjective perspectives and emotional nuances. Furthermore, the potential therapeutic effect of the data collection process itself on emotional disorder management remains underexplored.

Objective: This study developed and tested a mobile app that collects passive and active digital phenotypes related to adolescents' emotions and daily behaviors. The study aimed to assess the app's impact on managing emotional disorders through self-monitoring and to identify daily lifestyle indicators that can predict and track the development of such disorders.

Methods: A 4-week parallel, nonequivalent control group design was employed. The intervention group installed a digital phenotype collection tool on their mobile devices for 28 days. Passive data (location, sleep, and screen time) were continuously recorded. Active data were collected through ecological momentary assessments delivered randomly up to 8 times daily, prompting participants to report their current mood and levels of depression, anxiety, and stress. The control group received no intervention. Both groups were assessed at time points on emotional disorders, self-efficacy, and time management. Postintervention interviews were conducted with the intervention group.

Results: Thirty-six Korean adolescents participated (19 control, 17 intervention). The intervention group showed significant reductions in depression (P=.04, d=0.42) and stress (P=.03, d=0.46) and improvements in self-efficacy (P=.002, d=0.50) and time management abilities (P<.001, d=0.39), with small to large effect sizes. No significant change was observed in anxiety levels (P=.11). Correlational analysis revealed weak but significant links between passive digital phenotypes and daily emotional states.

Conclusions: Integrating active and passive digital phenotypes through a mobile collection tool can help manage emotional disorders in adolescents. Use of the tool was associated with moderate reductions in depression and stress, as well as improvements in self-efficacy and time management, while anxiety levels remained unchanged, possibly due to adolescents' differing perceptions of anxiety. Passive digital phenotypes such as location variability and phone usage showed modest correlations with daily emotional states, supporting their potential as ecological markers. These findings suggest that digital phenotype collection not only aids in monitoring but may also have therapeutic benefits by promoting self-reflection on mood and behavior. High adherence rates further support the practicality and acceptability of this approach for long-term emotional disorder management in adolescents.

(JMIR Hum Factors 2025;12:e66536) doi:10.2196/66536

KEYWORDS

adolescent; emotional disorder; digital phenotype; digital health; field study; EMA; self-monitoring; ecological momentary assessment

XSL•FO
RenderX

Introduction

Background

Alarming statistics reveal that nearly half of all mental illnesses begin before the age of 18 years, placing adolescents at the highest risk of developing emotional disorders [1,2]. Depression and anxiety are the most diagnosed conditions in adolescents, affecting nearly every aspect of their lives. These conditions can disrupt school, strain family relationships [3], and fuel risky behaviors such as smoking, drinking, and substance abuse. Most disturbingly, depression is a leading risk factor for suicide among adolescents [4]. Adolescence is a critical period of physiological, psychological, and social development; as such, mental health during this phase is intricately linked to well-being in adulthood [5]. This underscores the urgent need for early diagnosis and timely intervention in adolescent emotional disorders.

Digital phenotypes, the real-time collection of physiological and behavioral data through computer-based tools, have gained attention as a means to prevent and address mental disorders [4]. By capturing continuous, comprehensive data often unavailable through traditional methods, digital phenotypes offer new opportunities for the early detection and treatment of emotional disorders. These phenotypes can be categorized as active and passive. Active digital phenotypes require conscious user input, such as responses to self-reported surveys. Passive digital phenotypes collect data unobtrusively through smartphone sensors like accelerometers and Global Positioning System (GPS) as individuals go about their daily lives [6]. Research shows that digital phenotypes can effectively detect the severity of social anxiety, predict depression recurrence [6], identify suicidal ideation, and even prevent symptom deterioration with real-time monitoring [7].

However, opinions vary on whether passive digital phenotypes are sufficient to understand and explain emotional disorders. Context is crucial in shaping an individual's emotional perceptions and reactions, yet passive digital phenotypes often fail to capture this nuanced information. While behaviors such as movement or physical activity can be easily inferred from GPS and accelerometer data, more complex activities or emotional states that require inferences or are inaccessible to sensors remain unknown through passive sensing alone.

To gain a deeper understanding of emotional disorders, active digital phenotypes must be used to complement passive methods [8]. However, existing research primarily relies on clinical scales that may not fully account for the contextual factors relevant to emotional disorders [6,9-13]. Moreover, much of the literature treats digital phenotypes merely as data collection outcomes [14,15]. In contrast, active digital phenotyping functions as a form of self-monitoring, empowering individuals to gain insights into their emotional responses and develop targeted strategies to improve their mental health [16]. Therefore, the process of collecting active digital phenotypes may itself contribute to addressing emotional disorders.

In this context, our study is among the first to integrate both passive and active data collection in a single tool and to deploy it in the real world among adolescents. This dual-mode approach enables the collection of context-rich, ecologically valid data that neither passive nor active methods alone can provide. By analyzing the relationship between these digital phenotypes and daily assessed emotional disorder levels, we aim to determine how passive and active digital phenotypes function as complementary, parallel indicators of adolescent emotional health. Furthermore, this study investigates how the process of data collection can promote emotional disorder levels, self-efficacy, and time management skills.

Related Works

Digital Phenotypes and Emotional Disorders

Passive digital phenotypes are data collected about people's daily lives using digital devices [6]. This definition contains 4 further elements [9]. First, it involves continuous data collection and analysis rather than collecting data at specific time points. Second, the individual is the unit of analysis. Third, data are collected in natural, uncontrolled environments to capture real-world behavior. This point distinguishes passive from active digital phenotypes, such as ecological momentary assessments (EMAs), which will be discussed later. Fourth, passive digital phenotypes can be collected by devices that individuals already own, such as smartphones, without additional instruments [9].

A variety of physiological and behavioral passive digital phenotypes can be used to identify abnormal behavioral patterns associated with mental disorders [6,10]. People with mental disorders exhibit symptoms that affect mobility, life satisfaction, and social interactions, some of which can be quantified through digital phenotypes [11,12]. Symptoms of depression include negative affect, hopelessness, devaluation of life, and inertia, while anxiety is characterized by physiological arousal such as sweating, trembling, increased heartbeat, and avoidance tendencies [13]. Tracking these symptoms with passive digital phenotypes can allow for the early diagnosis and timely treatment of certain diseases, thereby more effectively managing health.

Adolescence is a particularly vulnerable developmental stage, often characterized by exposure to various stressors and their debilitating mental health effects [17]. Stress is accompanied by physical, psychological, or social complaints or dysfunctions resulting from individuals feeling unable to meet requirements or expectations placed on them. Stress during adolescence is associated with negative outcomes such as decreased well-being, impaired mental health, anxiety, and depression [17-20]. These challenges often co-occur, and symptoms of one disorder may predict the concurrent or future development of other mental health disorders. Therefore, addressing stress and emotional disorders is crucial for promoting long-term mental health.

Digital phenotyping could be particularly effective in addressing adolescent mental health challenges as adolescents often lack self-awareness or face communication barriers, particularly due to externalizing psychiatric disorders [21]. Given that approximately 95% of adolescents own or have access to smartphones, most by age 11 years, these technologies could be especially effective at tracking and intervening in mental health issues [22]. Mobile sensing and digital phenotypes can

XSL•FO
RenderX

support health behaviors by enhancing individuals' responsibility for their health and have been shown to increase health literacy [23]. These technologies also promote stronger relationships between health care providers and adolescents by facilitating meaningful conversations based on collected data, enabling timely treatment and encouraging help-seeking behaviors [24].

Ecological Momentary Assessment

An EMA [25] is used to build a comprehensive understanding of individuals [26,27]. It collects real-time self-reports of individuals' behaviors, thoughts, and emotions in their daily lives [28]. EMAs examine the present moment and so are conducted repeatedly over an extended period. With advancements in technology, EMAs are often implemented through digital platforms, such as smartphones and wearable devices [25].

Emotions are not fixed or stable; they are highly responsive to immediate circumstances and daily experiences [29]. Traditional data collection methods that rely on retrospective reports may struggle to capture these dynamics due to several biases, such as recency and novelty effects, which may compromise their validity [30]. EMAs overcome this limitation by gathering data at the moment when these experiences occur. This real-time approach allows for a detailed examination of how emotions evolve throughout the day [31].

To effectively manage emotional disorders, it is necessary to understand the conditions that trigger them [32]. Clinical scales and interviews with health care providers rarely address the context of the symptoms. In contrast, EMAs solicit a wide range of information from individuals, such as their location, time, and interpersonal influences, which can identify specific contexts in which emotional disorders are triggered [25,33]. Such detailed information can be used to better manage these conditions.

Unlike passive digital phenotypes, which rely on data collected by sensors, EMAs can directly ask individuals about their subjective experiences, which allows them to capture contextual data that sensors may not be able to [33]. Active and passive digital phenotypes complement each other, and neither is inherently superior. By integrating EMAs with passive digital phenotypes, researchers can achieve a more comprehensive understanding of emotional disorders [34]. This study collected active digital phenotypes through EMAs and passive digital phenotypes to develop detailed insights into the contextual factors influencing emotional disorders in adolescents.

Self-Monitoring and Self-Efficacy

Self-monitoring is the process in which individuals observe and record their own behaviors, emotions, and thoughts and adjust themselves accordingly [35]. Through this process of data collection and reflection, individuals can enhance their self-awareness and choose to engage in behaviors to achieve specific goals. Emotional self-monitoring helps identify factors that influence mood, both positively and negatively, allowing individuals to develop strategies to improve their mental health [36]. Self-efficacy is an individual's belief in their ability to plan and execute actions to achieve desired outcomes [37] and is closely linked to mental health. By boosting self-efficacy, self-monitoring can enhance mental health, increasing an individual's belief that they can autonomously manage and overcome emotional challenges [38,39].

Previous studies have identified several benefits of self-monitoring. Regularly measuring symptoms of emotional disorders can help predict their recurrence and evaluate the effectiveness of the ongoing treatments [40]. It provides opportunities to reflect on the context surrounding symptom occurrences, aiding in understanding the causes. Moreover, self-monitoring helps identify factors that either alleviate or worsen emotional disorders, providing insights into their patterns and impacts [41]. Finally, self-monitoring is done independently without external interventions, so it encourages individuals to actively participate in managing their mental health [36,42,43], which can enhance an individual's agency, positively influencing their self-efficacy.

Combined with these benefits, self-monitoring has been shown to improve mental health symptoms [44]. Studies on adolescents in the early stages of depression have demonstrated that self-monitoring gives them insights into their health, which is the first step in a stepped-care approach [33,45,46]. Furthermore, self-monitoring extends symptom-free periods in patients with depressive disorders [47] and helps them focus on behaviors that can help them control their emotional disorders [48]. Even for individuals not clinically diagnosed with emotional disorders, self-monitoring has been shown to have positive effects, such as motivating them to pursue health management [49].

To effectively define the role of digital phenotypes in managing emotional disorder, it is crucial to understand them as self-monitoring. However, the role of digital phenotypes as a tool for self-monitoring has not been investigated yet. This study explored whether digital phenotypes could help individuals engage in self-monitoring of emotional disorders. Accordingly, a service was developed for this study that integrates the collection of digital phenotypes, which have been validated for their use in the early diagnosis and treatment of adolescent mood disorders. The study evaluated the effects of this service on adolescents' depression, anxiety, stress, and self-efficacy levels and time management abilities.

Methods

Study Design

Participants were not randomly assigned due to constraints related to the study tool's operational environment as the app was only compatible with Android smartphones. Given the limited number of adolescents using Android devices who were available to participate, it was not feasible to recruit a sufficiently large sample for both groups exclusively from Android users. The study used a 4-week parallel, nonequivalent control group design.

The intervention group consisted of Android users who installed and used the app for 28 days. The control group, which was not required to use the app, had no restrictions on smartphone type. Both groups were assessed on key variables at baseline (day 0), midpoint (day 14), and postintervention (day 28). Following

XSL•FO
RenderX

the intervention, semistructured interviews were conducted with the intervention group to explore their experiences using the digital phenotype collection tool.

Interviews were conducted via the Zoom online conferencing tool within 1 week after the intervention period ended. With participants' consent, all interviews were audio-recorded, transcribed verbatim, and then deleted to maintain confidentiality. Topics covered the perceived impact of app usage on personal life and mood and suggestions for improvement. Data were analyzed using thematic analysis, with 2 researchers independently coding to identify recurring themes and patterns.

Recruitment

This study included individuals who: (1) were South Korean residents in the age group of 12 - 18 years, (2) were enrolled in middle or high school, (3) spoke Korean as their native language and had no difficulty communicating in it, and (4) understood the study purpose and voluntarily agreed to participate. Participants were recruited to the intervention group if they owned an Android smartphone and had no issues using it to ensure the efficient operation of research tools.

Participants were recruited through multiple channels, including digital flyers posted on social media platforms (eg, Instagram, KakaoTalk channels), online youth community forums, and outreach via collaborating clinicians. Recruitment materials briefly explained the study purpose and provided a link to an online preassessment survey. Those who completed the survey received the download link to the study tool via a text message.

This recruitment strategy may have introduced selection bias as participants were more likely to be adolescents who actively engage with digital platforms and possess higher digital literacy. Additionally, the requirement for Android smartphone ownership in the intervention group may have introduced systematic differences between groups, particularly in terms of socioeconomic status and smartphone usage patterns.

Sample Size Calculation

The sample size was calculated using a 2-sided significance level of .05, a statistical power of 80%, 2 groups, and 3 repeated measures. Based on these parameters, a sample size of 26 participants per group was estimated to be sufficient to detect an effect size of .33 in the reduction of depression levels. This effect size was selected based on findings from a systematic review of digital mental health interventions targeting young people, which reported small to moderate effect sizes for similar outcomes. To account for a potential 5% dropout rate, we increased the target sample size to 56 participants in total.

Intervention

The study tool was an app that collects digital phenotypes validated as predictive indicators for emotional disorders. The app was developed based on exploratory research into adolescents' intentions and behaviors in voluntarily recording emotions and daily activities [50]. To develop the tool, we first observed how adolescents recorded target active digital phenotypes regarding emotions and lifestyle behaviors within a digital environment and inquired about the tools they used, including their advantages and disadvantages.

These observations were then translated into system design goals, which included designing interfaces that resemble digital planners, streamlining data input processes with shortcut interactions, providing visual reports on collected data to reward users' efforts, and writing friendly EMA prompts. By aligning the system's features with users' natural behaviors, the app was designed to enhance the quality of collected active digital phenotypes.

Emotional recording followed an EMA methodology where users responded to notifications delivered randomly throughout the day (Figure 1). These notifications were sent upon waking, before school, during lunchtime, after school, before private tutoring, after private tutoring, before bedtime, and if the phone was used for more than 30 minutes after midnight. Users were restricted to recording their emotions only when prompted by these notifications.

The recording of lifestyle behaviors was facilitated through a screen displaying a table that divided the 24-hour day into 30-minute intervals, resulting in 48 cells (Figure 2). By clicking on each cell, users could access a pop-up screen with sleep, exercise, leisure, and study categories, each of which included specific activities, like napping or meeting friends. The recorded data were then processed into reports, allowing users to gain detailed insights into the structure of their daily lives and their levels of emotional disorders.

XSL•FO
RenderX

Figure 1. Process of mood recording with corresponding system interface screens. EMA: ecological momentary assessment.

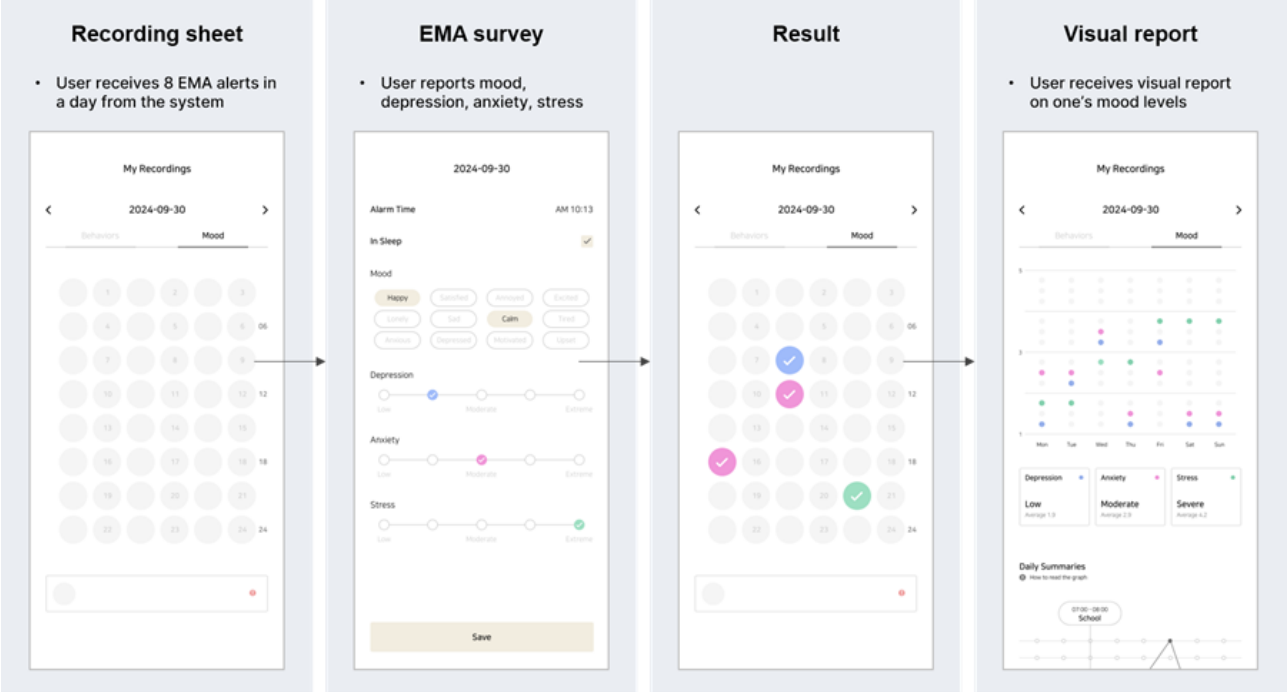
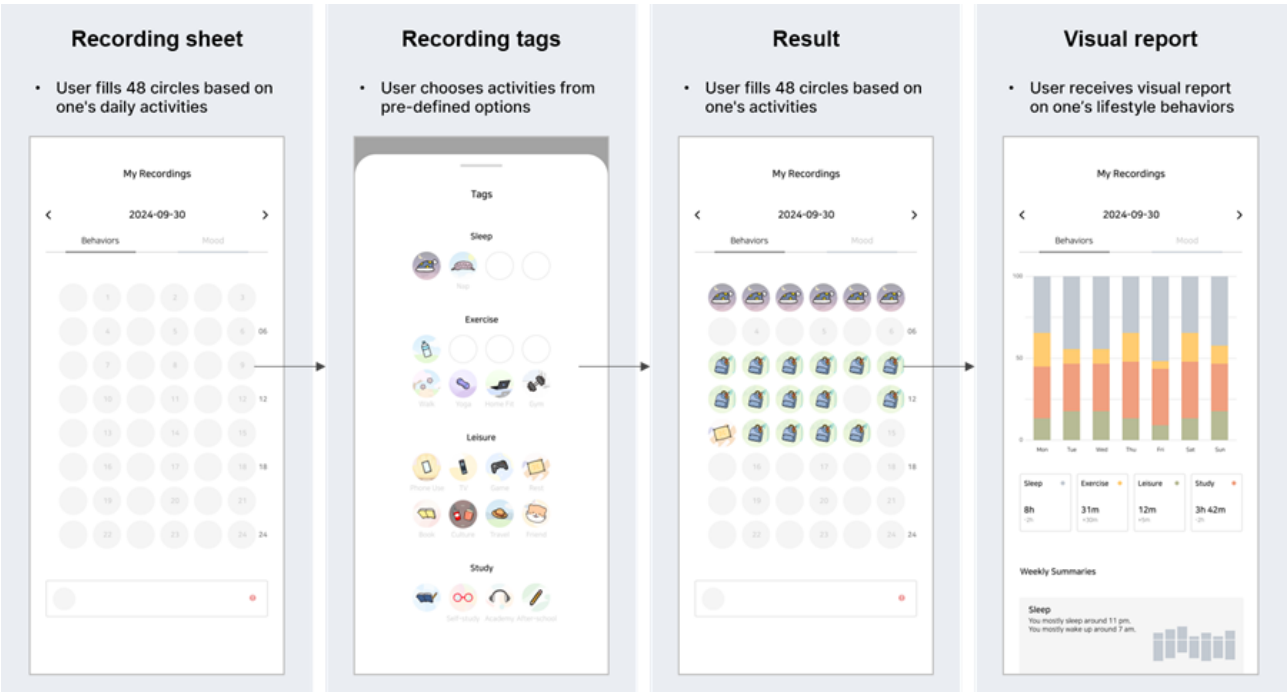


Figure 2. Process of lifestyle behavior recording with corresponding system interface screens.



Measures

Demographic Information

Participants provided demographic information, namely gender, age, level of education, and address. The level of effort typically invested in mood and life management was assessed using a 5-point Likert scale, ranging from 1 (“very low”) to 5 (“very high”).

Korean Version of the Depression Anxiety Stress Scale

The Korean version of the Depression Anxiety Stress Scale, adapted by Lee (2002) from the version shortened by Henry of the original developed by Lovibond, was used to measure levels of depression, anxiety, and stress [51-53]. This shortened scale consists of 21 items, with 7 items assessing each symptom. Each item is rated on a 4-point Likert scale from 0 (“does not apply at all”) to 3 (“applies most of the time”), with higher scores indicating greater severity of symptoms.

Self-Efficacy Scale

Self-efficacy was measured using a scale developed by Sherer and translated by Kim [54,55]. The scale includes 22 items divided into 3 subfactors: confidence, self-regulatory efficacy, and preference for task difficulty. Each item is rated on a 5-point

XSL•FO
RenderX

Likert scale from 1 (“strongly disagree”) to 5 (“strongly agree”), with higher scores indicating higher self-efficacy. The internal consistency coefficient, Cronbach α, for this scale is 0.88.

Time Management Ability Scale

Time management ability was evaluated with a scale developed by Jeong and Jang, which includes 30 items categorized into 3 subfactors: planning ability, execution ability, and evaluation ability [56]. Participants rated each item on a 5-point Likert scale from 1 (“not at all”) to 5 (“always”), with higher scores indicating better management ability. The internal consistency coefficients, Cronbach α, for these subfactors are 0.91, 0.86, and 0.89, respectively.

mHealth App Usability Questionnaire

To evaluate the usability of the app developed for this study, we used a self-reported scale that assesses the usefulness and usability of mobile health apps [57]. The scale consists of 21 items rated on a 7-point scale from 1 (“strongly disagree”) to 7 (“strongly agree”). The questionnaire is divided into 3 subfactors: ease of use and satisfaction, assessed by 8 items, system information arrangement, assessed by 6 items, and usefulness, assessed by 7 items.

Passive Digital Phenotypes

The location digital phenotypes, collected using the devices’ GPS functions, included total distance traveled, total travel time, location variance, entropy, and normalized entropy. These data were preprocessed by estimating travel speed through time derivatives. Using a threshold of 1 km/h, the data were classified as either stationary or moving [57,58]. The stationary data were analyzed with the K-means clustering algorithm to identify the locations where participants spent most of their time.

The sleep digital phenotypes included total amount of sleep, total daytime sleep, and total nighttime sleep. Considering the daily routine and schedule of adolescents, daytime was defined as 7 AM to 11 PM and nighttime as from 11 PM to 7 AM. Sleep states were estimated using data collected from cameras, microphones, accelerometers, and smartphone log data, considering factors such as ambient light, noise, and smartphone screen time [59].

Digital phenotypes calculated from screentime data contained in device logs included total phone usage duration, total daytime phone usage, total nighttime phone usage, total phone lock duration, and the frequency of phone unlocks [60]. We handled outliers using the interquartile range method, where values outside 1.5 times the IQR were considered outliers and excluded. We did not attempt to impute missing values as the data were passive and imputation could introduce artifacts.

Active Digital Phenotypes

Mood records were collected using EMAs. System notifications were delivered randomly throughout the day, except during school, private tutoring, and sleep. Upon receiving a notification, participants recorded their current mood by selecting the mood-related adjectives (eg, happy, sad, relaxed) and rated their levels of depression, anxiety, and stress on a scale from 1 (“very low”) to 5 (“very high”).

Participants recorded their daily activities using predefined options. Key lifestyle areas include sleep, exercise, leisure, and study-related activities. Each 24-hour day was divided into 30-minute intervals and participants were asked to select the corresponding activity for each time slot (eg, studying, walking). The recorded lifestyle behaviors were organized by type, start time, and end time.

Statistical Analysis

Analyses were conducted using SPSS 26.0. Descriptive statistics were used to describe and compare participants’ baseline characteristics. Univariate analyses were performed to compare baseline values using the *t* test for independent samples. A 2-way ANOVA model followed by Scheffé post hoc tests was used to test the effect of app usage on each variable. Data were analyzed using 2×3 repeated measures ANOVAs with 1 between-subject factor (intervention group and control group) and 1 within-subject factor (preintervention, midintervention, and postintervention). The statistical significance level was defined as α=.05.

Ethical Considerations

This study was reviewed and approved by the Yonsei University Institutional Review Board (IRB no. 7001988-202406-HR-2045-05). All participants provided written informed consent after receiving a full explanation of the study purpose and procedures from the research team. Participant confidentiality was maintained throughout the study, with each individual assigned a numerical code according to their order of participation. Upon completion of the study, participants received compensation of 100,000 KRW (US $72.00).

Results

Descriptive Statistics

Of the 21 participants in the intervention group, 17 completed the study, resulting in an 81% retention rate. Dropouts occurred due to technical issues (n=2), failure to complete the preintervention survey (n=1), and personal reasons (n=1). In the control group, all 19 participants completed the study, resulting in a 100% retention rate. The intervention group consisted of 3 male and 14 female participants, with a mean age of 16.70 (SD 0.98) years. The control group included 2 male and 17 female participants, with a mean age of 16.63 (SD 1.02) years. There were no statistically significant differences in gender distribution or age between the two groups.

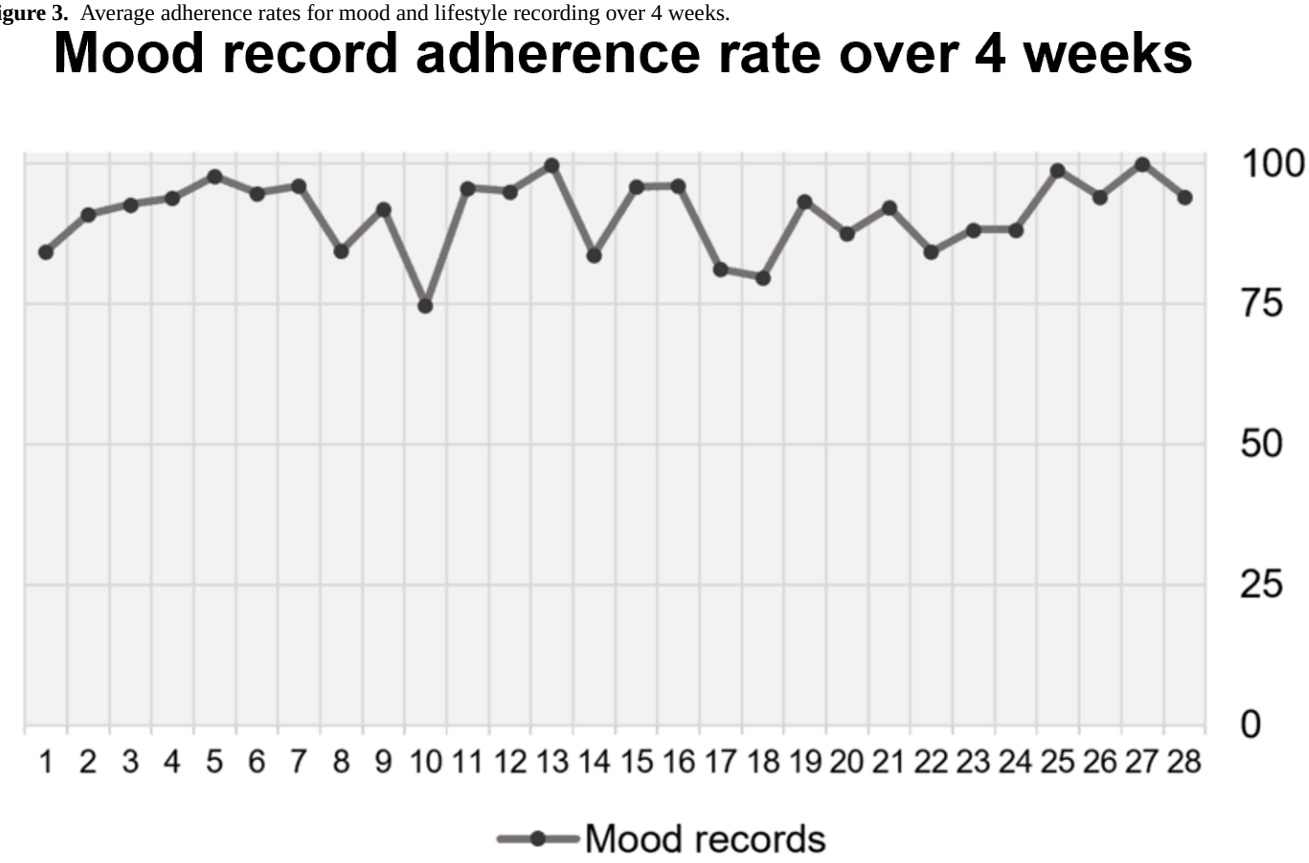
At baseline, there were no statistically significant differences between the groups in anxiety, self-efficacy, or time management abilities. However, the intervention group showed significantly higher levels of depression and stress. For depression, the control group had a mean score of 10.21 (SD 2.20), while the intervention group had a mean score of 18.12 (SD 3.60); this difference was statistically significant (t_{34}=–2.25, P=.03). For stress, the control group’s mean score was 12.00 (SD 3.10) and the intervention group’s score was 19.76 (SD 3.80). A Mann–Whitney U test indicated a significant difference between the groups (U=234.5, P=.02).

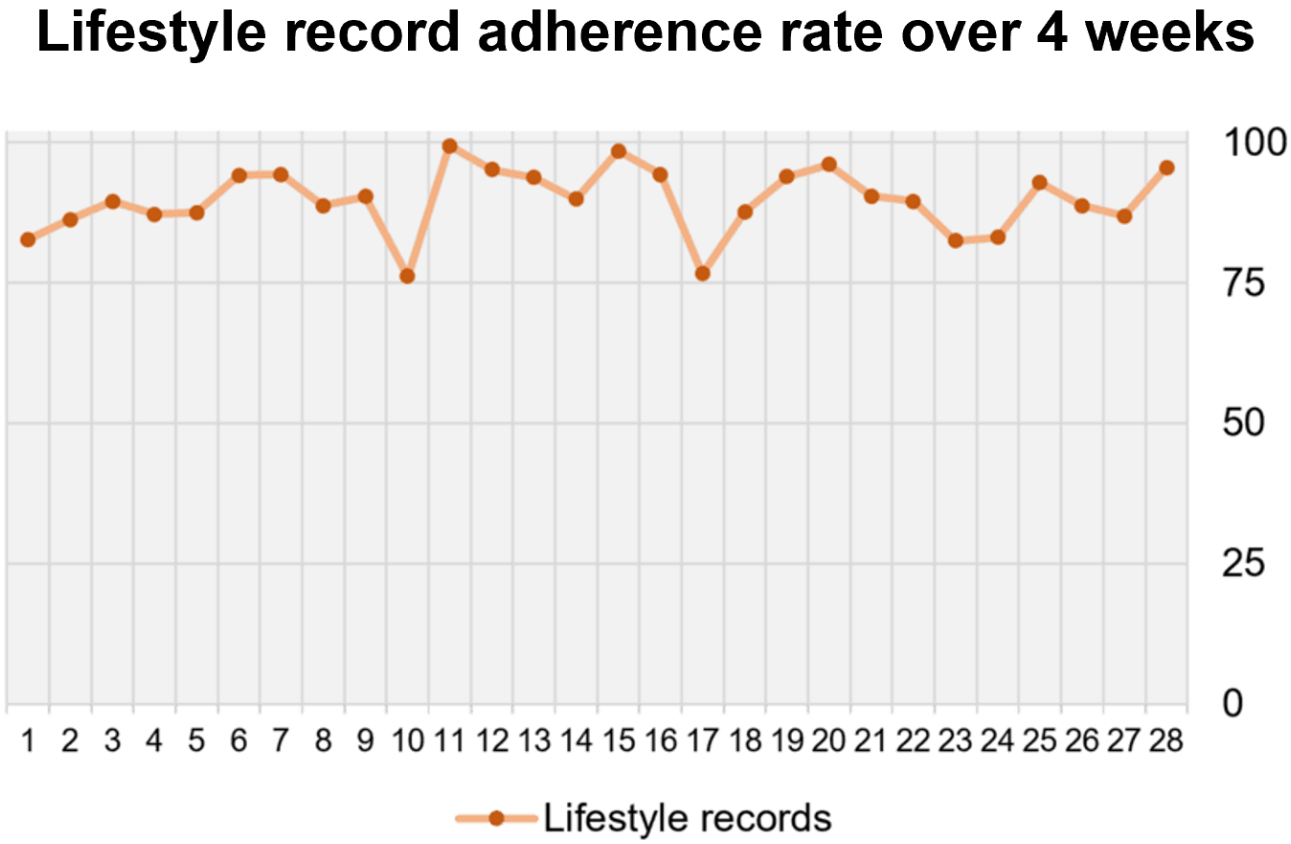
XSL•FO
RenderX

Adherence to lifestyle records was defined as the percentage of records made from the 48 requests. Adherence to mood records was defined as the percentage of responses to the 28 daily EMA alerts sent. Lifestyle record adherence had an average of 89.68% and a range of 71.5% - 99.6% (Multimedia Appendix 1). Mood record adherence had an average of 90.65% and a range of 79.3% - 99.6%. The average adherence rates per participant indicate that most participants complied with the system's requests for collecting digital phenotypes (Figure 3).

XSL•FO
RenderX

Figure 3. Average adherence rates for mood and lifestyle recording over 4 weeks.





XSL•FO
RenderX

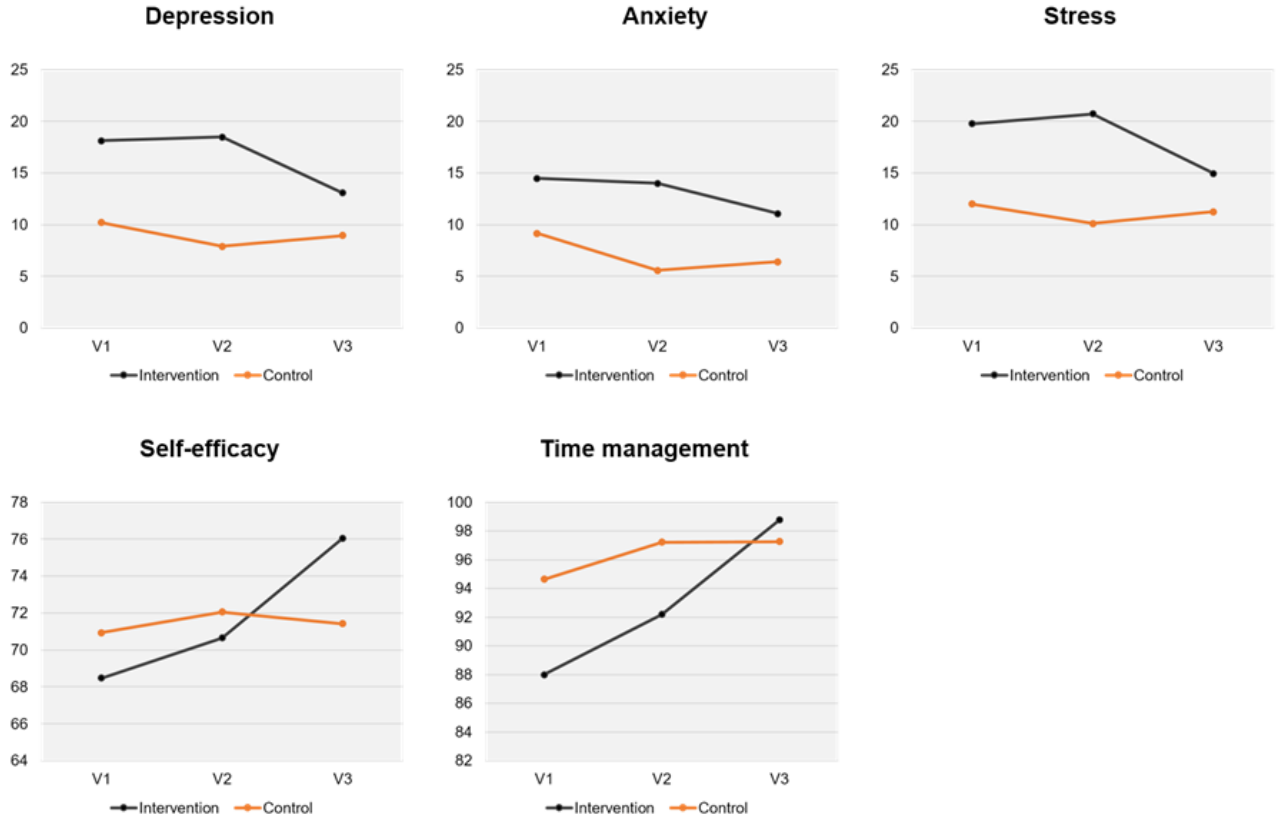
Intervention group satisfaction was assessed using the mHealth App Usability Questionnaire at the end of the intervention. The survey consisted of 21 items rated on a 7-point scale, evaluating the tool's ease of use, user satisfaction, system information arrangement, and usefulness. Participants' responses showed a high level of satisfaction, with an average score of 6.31 of 7. The median score was 6.29, indicating that half of the participants rated the tool at or above this value.

Scores ranged from a minimum of 5.81 to a maximum of 6.86, demonstrating relatively low variability in user evaluations. The SD was 0.31, further confirming consistent positive feedback across the participant group. These findings suggest that the digital phenotype collection tool was generally well-received by adolescents, both in terms of usability and usefulness, which is also reflected in the high adherence rate throughout the intervention period [57].

Changes in Repeated Measures

A repeated measures ANOVA revealed a significant main effect of time ($F_{2, 68}$=3.62, P=.04), indicating that depression levels changed significantly across the 3 time points, regardless of group. A repeated measures ANOVA revealed a statistically significant interaction effect between time and group for depression levels ($F_{2, 68}$=3.65, P=.04), indicating that the intervention group experienced a different trajectory of change compared with the control group (Figure 4). For anxiety, although a significant main effect of time was observed ($F_{2, 68}$=5.64, P=.009), the interaction effect was not significant ($F_{2, 68}$=2.35, P=.11), suggesting similar changes across both groups.

Figure 4. Repeated measures ANOVA results comparing the intervention and control groups.



In contrast, stress levels showed no significant main effect of time ($F_{2, 68}$=2.56, P=.07), but there was a significant interaction between time and group ($F_{2, 68}$=3.55, P=.03). Time management ability significantly improved over time ($F_{2, 68}$=11.54, P<.001), with a significant interaction effect ($F_{2, 68}$=4.78, P=.01), indicating a greater improvement in the intervention group. Similarly, self-efficacy scores showed both a significant main effect of time ($F_{2, 68}$=6.99, P=.002) and a significant interaction effect ($F_{2, 68}$=5.27, P=.007). These results suggest that the intervention had a positive and differential effect on depression, stress, time management, and self-efficacy when compared with the control group.

Passive Digital Phenotypes and EMA

We analyzed the relationship between passive digital phenotypes and emotional disorder levels over time. While previous studies often clustered subjects based on their baseline emotional disorder scores, this study used ecologically assessed emotional disorder levels collected over a 4-week period [61,62]. This approach aimed at the same timeframe of the data in the analysis and to capture the dynamic nature of an individual's emotional disorders, which fluctuate within a day.

Collected passive digital phenotypes were categorized by participants and averaged daily, resulting in 28 sets of passive digital phenotypes per participant. Outliers and missing values were excluded from the analysis. The results of the analysis are

XSL•FO
RenderX

presented in Tables 1-3, summarizing Pearson correlation coefficients with a significance level of α=.05.

Table . The correlation analysis of passive digital phenotypes and depression ecological momentary assessment.

	Depression	*P* value
Location		
Total travel distance	−0.0932[a]	.04
Total travel time	−0.0806	.08
Location variance	0.1041[a]	.02
Location entropy	−0.1186[b]	.01
Normalized location entropy	−0.1089[a]	.02
Sleep		
Total sleep duration	−0.0053	.91
Daytime sleep duration	0.0988[a]	.03
Nighttime sleep duration	−0.0429	.35
Phone log		
Total phone usage time	−0.1123[a]	.01
Total phone screen lock time	0.0980[a]	.03
Daytime phone usage time	−0.0511	.27
Nighttime phone usage time	−0.1421[b]	.002
Phone screens unlock frequency	−0.1511[c]	<.001

[a]*P*<.05.
[b]*P*<.01.
[c]*P*<.001.

XSL•FO
RenderX

Table . The correlation analysis of passive digital phenotypes and anxiety ecological momentary assessment.

	Anxiety	*P* value
Location		
Total travel distance	−0.0782	.09
Total travel time	−0.0735	.11
Location variance	0.1376[b]	.003
Location entropy	−0.1234[b]	.007
Normalized location entropy	−0.1296[b]	.005
Sleep		
Total sleep duration	0.0452	.33
Daytime sleep duration	0.1342[b]	.003
Nighttime sleep duration	−0.0057	.90
Phone log		
Total phone usage time	−0.9078[c]	<.001
Total phone screen lock time	−0.1159[a]	.01
Daytime phone usage time	−0.0401	.38
Nighttime phone usage time	−0.1329[b]	.004
Phone screens unlock frequency	−0.1159[a]	.01

[a]*P*<.05.
[b]*P*<.01.
[c]*P*<.001.

Table . The correlation analysis of passive digital phenotypes and stress ecological momentary assessment.

	Stress	*P* value
Location		
Total travel distance	−0.0156	.73
Total travel time	0.0118	.80
Location variance	0.0955[a]	.04
Location entropy	−0.0566	.22
Normalized location entropy	−0.0896	.05
Sleep		
Total sleep duration	0.0281	.54
Daytime sleep duration	−0.1262[b]	.006
Nighttime sleep duration	−0.0197	.67
Phone log		
Total phone usage time	−0.0881	.06
Total phone screen lock time	0.0694	.13
Daytime phone usage time	−0.0461	.32
Nighttime phone usage time	−0.1074[a]	.02
Phone screens unlock frequency	−0.1798[c]	<.001

[a]*P*<.05.
[b]*P*<.01.
[c]*P*<.001.

XSL•FO
RenderX

For location digital phenotypes, there were no statistically significant correlations between total travel time and emotional disorders. Total travel distance was negatively correlated with depression (r=−0.0932, P=.04). Statistically significant positive correlations emerged between location variance and depression (r=0.1041, P=.02), anxiety (r=0.1376, P=.003), and stress (r=0.0955, P=.04). Location entropy was negatively correlated with depression (r=−0.1186, P=.01) and anxiety (r=−0.1234, P=.007), while normalized location entropy was statistically significantly negatively correlated with depression (r=−0.1186, P=.01) and anxiety (r=−0.1089, P=.02).

Regarding sleep digital phenotypes, there were no statistically significant correlations between total sleep duration, nighttime total sleep duration, and the level of emotional disorders. However, daytime total sleep duration showed statistically significant positive correlations with depression (r=0.0988, P=.03), anxiety (r=0.1342, P=.003), and stress (r=0.1262, P=.006). These results suggest that as daily levels of emotional disorders increase, daytime sleep duration or nap duration also tends to increase.

The relationships between digital phenotypes of phone use and daily depression, anxiety, and stress levels are presented in Tables 1-3. Statistically significant correlations were found between all variables except for daytime phone usage. Total phone usage exhibited a statistically significant negative correlation with depression (r=−0.1123, P=.01) and anxiety levels (r=−0.9078, P<.001). Phone screens unlock frequency showed statistically significant negative correlations with depression (r=−0.1151, P<.001), anxiety (r=−0.1159, P=.01), and stress levels (r=−0.1798, P<.001).

Total phone screen lock time showed a statistically significant positive correlation with depression (r=0.0980, P=.03) and negative correlation with anxiety (r=−0.1159, P=.01), whereas nighttime phone usage time exhibited statistically significant negative correlations with depression (r=−0.1421, P=.002), anxiety (r=−0.1329, P=.004), and stress levels (r=−0.1074, P=.02). These findings suggest that as daily levels of emotional disorders increase, nighttime phone usage decreases. Overall phone usage tends to decrease on days with higher depression and anxiety levels, while total phone screen lock time increases on days with higher anxiety levels.

Discussion

Main Results

Overview

This study developed and deployed a tool that captures both passive and active digital phenotypes in real-world settings to build a comprehensive dataset for understanding emotional disorders in adolescents. The tool collects behavioral data commonly used in prior studies of passive digital phenotypes, such as location, sleep patterns, and phone usage, and extends data collection into domains that are not easily captured by sensors, including study and rest routines, self-reported mood, and subjective emotional distress.

Unlike studies that focused solely on passive sensing or retrospective self-reports, this study positioned self-monitoring as both a data collection strategy and an intervention mechanism. As participants recorded their emotions, they became more aware of their levels of depression, anxiety, and stress and examined the underlying causes of these feelings. This process caused participants to integrate their reflections on both emotions and daily life, allowing them to construct causal relationships between daily experiences and emotional states based on the gathered digital phenotypes. This aligns with prior findings that self-monitoring can facilitate emotional awareness and behavior change, but this study extends that work by integrating real-time active digital phenotypes with passively collected context.

Furthermore, the correlation analysis between passive and active digital phenotypes revealed patterns consistent with earlier findings. For instance, higher late-night phone use and irregular sleep patterns were associated with elevated stress and depression levels. These observations suggest that integrated digital phenotype systems can surface meaningful, co-occurring patterns between lifestyle behaviors and emotional states among adolescents. The strength of this study is further supported by the high adherence rate across both active and passive data streams. Overall, the findings demonstrate how combining passive sensing with active self-reporting can generate a comprehensive understanding of emotional well-being and support light-touch interventions through self-reflection.

User-Centric Design for Digital Phenotypes

As demonstrated in this study, having individuals consistently record their mood and lifestyle behaviors can aid in managing emotional disorders by leveraging the benefits of self-monitoring. While developing our study tool, we applied exploratory research to enhance both the quantity and quality of active phenotype data. A pilot study with adolescents helped us understand how their daily routines and digital habits could inform a design aligned with their lived experiences.

Based on these insights, we implemented a timetable-based interface that reflected the structure of a typical adolescent day, allowing participants to intuitively locate and record emotions in the context of their daily schedules. We designed context-aware EMA nudges that prompted emotion recordings at relevant points in the day, increasing compliance without causing notification fatigue. The visualization of the collected data prioritized the types of insights adolescents reported to value most, making the reflection more meaningful.

The impact of these design features is reflected in the tool's high usability and strong adherence. Our findings suggest that in designing tools to collect active digital phenotypes from adolescents, it is essential to find an appropriate balance between the burden of self-reporting and the perceived value of data reflection. A well-designed interface that minimizes friction and maximizes personal relevance can significantly improve both engagement and the quality of self-reported data.

Digital Phenotypes and Personalized EMAs

The potential of digital phenotype collection as an intervention also emerged using EMA notifications. In follow-up interviews,

XSL•FO
RenderX

participants reported that when EMA alerts were tailored to their daily life context rather than generic reminders, they experienced a stronger sense of social support and were more motivated to engage with the tool. The effect was particularly notable when prompts included words of encouragement and comfort, such as cheering them on to attend school or private tutoring sessions. Additionally, the tool issued EMAs during extended periods of late-night device usage. Participants interpreted these messages as expressions of care, which encouraged them to reflect on their behavior and, in many cases, reduce nighttime screen time.

In this study, EMA notifications were primarily time-based and followed a fixed template designed by the researchers. However, by integrating both passive and active digital phenotypes, EMA prompts could be further personalized to match each user's real-time context. For example, if an adolescent logs study time that exceeds recommended limits, the tool could deliver a supportive message acknowledging their effort while gently encouraging balance. Likewise, if an increase in physical activity is detected, a message might inquire about their experience, reinforcing awareness and positive habits. Such personalization could be significantly enhanced by generative artificial intelligence, which would allow EMA content to evolve dynamically based on ongoing user interactions and behavioral patterns.

This personalized, context-aware approach can extend to facilitating more proactive interventions. By analyzing trends in digital phenotype data when heightened distress is reported, we can infer individual patterns related to emotional disorders. For instance, this study found a positive correlation between increased nap duration and elevated emotional disorder levels. If the system detects more frequent or longer naps, it could adjust EMA notification such as, "You've been napping more than usual; have you been feeling more down lately?" However, it is important that these prompts are framed with care. Messages should invite introspection rather than make assumptions and always be mindful of user privacy, especially when dealing with sensitive behavioral data.

Our findings extend the current discourse on digital phenotypes in 2 significant ways. First, EMA can be more than simply assessments by indirectly helping to alleviate emotional disorders. EMAs involve frequent interactions between the system and the user, with each session shaping the user experience. In this study, we particularly designed empathetic survey prompts, which helped build rapport with and provide emotional support to participants. Second, this study showed how passive digital phenotypes, which are commonly used to determine the timing of EMA prompts, could be used. Participants viewed the tool's apparent interest in them as a major benefit, which suggests that passive digital phenotypes could be used more broadly to optimize timing and personalize other aspects of EMAs based on users' behavioral and psychological characteristics.

Limitations

This study has several limitations. First, although it included both an intervention group that used a tool for 4 weeks and a control group that did not use it, participants were not randomly assigned to the groups and there was a gap in recruitment timing. The tool was exclusively designed for Android phones, which limited our ability to randomly assign participants to groups. While we attempted to ensure homogeneity by considering factors such as age, gender, residential area, and grade, there may have been unexamined characteristics that could have influenced the study results.

Second, active digital phenotypes were collected via self-reporting, which was only possible when participants were not otherwise engaged. To accurately assess participants' daily routines, the study needed to be conducted during school terms but not around major school events, such as national examinations, which could affect levels of depression, anxiety, and stress. Thus, there was a 1-month difference in the start times between the intervention and control groups, which might have affected the results. Third, there were statistically significant differences in depression and stress level baseline scores between the two groups. The intervention group reported higher levels of both depression and stress initially. This discrepancy indicates that the intervention group had greater potential for improvement, so the positive impact of the digital phenotype collection service should be interpreted with caution.

Fourth, there was an uneven distribution of gender and grade among participants. Female participants were more represented than male participants in both groups, and there were more high school students than middle school students. These imbalances may limit the generalizability of the findings to male and middle school students. However, emotional disorders are generally more frequent and severe among female adolescents, so this distribution may reflect real-life conditions. Additionally, as the study was conducted exclusively with Korean adolescents, cultural factors unique to Korea may have influenced participants' behaviors and responses. This cultural specificity could limit the applicability of the findings to adolescents in other cultural contexts. Future studies should aim for a more balanced representation by gender and grade and include more diverse cultural settings to improve the generalizability of the results.

Fifth, this study did not specifically target adolescents diagnosed with emotional disorders. Although some participants had high emotional disorder scores, these scores did not necessarily equate to clinical diagnoses. Furthermore, the EMAs assessed depression, anxiety, and stress levels using single items for ease of response, which may not fully represent the participants' clinical status. Future research should focus on participants with clinically diagnosed emotional disorders to enhance the generalizability and clinical relevance of the study findings.

Conclusions

This study explored how digital phenotypes can be used to manage emotional disorders among adolescents. To this end, we developed a digital phenotype collection tool that integrates active digital phenotypes into EMAs for mood disorders and related daily contexts. This study included a 4-week field test to investigate the potential diagnostic and therapeutic roles of active digital phenotypes to assess how participating in digital phenotype collection affects the emotional disorders of depression, anxiety, and stress; self-efficacy; and time

XSL•FO
RenderX

management abilities. We also examined the relationships between passive digital phenotypes and adolescents' daily emotional states.

The results showed that tool use was negatively correlated with depression and stress levels and positively correlated with self-efficacy and time management abilities. However, it was not correlated with anxiety levels and interviews with participants provided insights as to why. Adolescents understood depression in a way that matched clinical definitions, such as lethargy and sadness. However, adolescents tend to interpret anxiety in a more situational manner, associating it with specific stressors like upcoming examinations or incomplete homework, rather than recognizing it as a clinical condition. This discrepancy might explain why no significant improvement in anxiety levels was observed when assessed using clinical scales.

Furthermore, we analyzed the relationship between passive digital phenotypes and their relationships with adolescents' levels of depression, anxiety, and stress via EMAs. This correlation analysis revealed that location variance, location entropy, daytime sleep duration, phone usage time, and phone screens unlock frequency were weak but statistically significantly correlated with daily levels of depression, anxiety, and stress, which is like trends observed in other studies [4,63]. Unlike past research that typically compared baseline scores, our study used long-term, ecologically assessed data, showing that EMAs can be used to establish predictive correlations with passive digital phenotypes.

These findings suggest that digital phenotype collection services can be used to understand and manage adolescent emotional disorders. Despite the recognized potential of digital phenotypes for this purpose, few studies have collected active digital phenotypes in the context of emotional disorders, and none have focused on their effectiveness in managing these disorders. This study provides preliminary evidence that the use of a digital phenotype collection tool, particularly through the process of reflecting on both mood and lifestyle-related behaviors, may contribute to improvements in depression, self-efficacy, and time management among adolescents. These findings suggest that the integration of passive and active digital phenotypes could play a supportive role in better understanding and potentially addressing emotional challenges in this population. Additionally, the high level of adherence maintained throughout the study, despite the extensive user participation required, indicates that the methods used to collect digital phenotypes were successful.

Acknowledgments

This research was supported by a grant of the 2024 Bio-industry Technology Development Project, funded by the Ministry of Trade, Industry and Energy, Republic of Korea (grant RS-2024-00431485).

Conflicts of Interest

M Cho, DP, M Choo are employees of the company (HAII Co, Ltd), founded by JK. The study was conducted using software designed by the company, and conflicts of interest may arise due to the affiliation with the company.

Multimedia Appendix 1
Average adherence rates for mood and lifestyle recording per participant.
[PNG File, 104 KB - humanfactors_v12i1e66536_app1.png]

References

1. Gulliver A, Griffiths KM, Christensen H. Perceived barriers and facilitators to mental health help-seeking in young people: a systematic review. BMC Psychiatry 2010 Dec 30;10(1):113. [doi: 10.1186/1471-244X-10-113] [Medline: 21192795]
2. Solmi M, Radua J, Olivola M, et al. Age at onset of mental disorders worldwide: large-scale meta-analysis of 192 epidemiological studies. Mol Psychiatry 2022 Jan;27(1):281-295. [doi: 10.1038/s41380-021-01161-7] [Medline: 34079068]
3. Moon KY, Choi WJ, Son S, Jung S, Lee S, Shin D. Correlation with psychosocial, workplace coping skills, pain, and degree of disability of office workers with non-specific chronic neck pain. Phys Ther Rehabil Sci 2022 Dec 30;11(4):565-570. [doi: 10.14474/ptrs.2022.11.4.565]
4. Bufano P, Laurino M, Said S, Tognetti A, Menicucci D. Digital phenotyping for monitoring mental disorders: systematic review. J Med Internet Res 2023 Dec 13;25:e46778. [doi: 10.2196/46778] [Medline: 38090800]
5. Felitti VJ, Anda RF, Nordenberg D, et al. Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults. The Adverse Childhood Experiences (ACE) Study. Am J Prev Med 1998 May;14(4):245-258. [doi: 10.1016/s0749-3797(98)00017-8] [Medline: 9635069]
6. Onnela JP, Rauch SL. Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health. Neuropsychopharmacology 2016 Jun;41(7):1691-1696. [doi: 10.1038/npp.2016.7] [Medline: 26818126]
7. Kleiman EM, Turner BJ, Fedor S, et al. Digital phenotyping of suicidal thoughts. Depress Anxiety 2018 Jul;35(7):601-608. [doi: 10.1002/da.22730] [Medline: 29637663]
8. de Vries LP, Baselmans BML, Bartels M. Smartphone-based ecological momentary assessment of well-being: a systematic review and recommendations for future studies. J Happiness Stud 2021;22(5):2361-2408. [doi: 10.1007/s10902-020-00324-7] [Medline: 34720691]

XSL•FO
RenderX

9. Onnela JP. Opportunities and challenges in the collection and analysis of digital phenotyping data. Neuropsychopharmacology 2021 Jan;46(1):45-54. [doi: 10.1038/s41386-020-0771-3] [Medline: 32679583]
10. Torous J, Kiang MV, Lorme J, Onnela JP. New tools for new research in psychiatry: a scalable and customizable platform to empower data driven smartphone research. JMIR Ment Health 2016;3(2):e16. [doi: 10.2196/mental.5165]
11. Lardier DT, Lee CYS, Rodas JM, Garcia-Reid P, Reid RJ. The effect of perceived college-related stress on depression, life satisfaction, and school satisfaction: the coping strategies of Hispanic college students from a Hispanic serving institution. Educ Urban Soc 2020 Nov;52(8):1204-1222. [doi: 10.1177/0013124519896845]
12. Rezaei A, Mousanezhad Jeddi E. Relationship between wisdom, perceived control of internal states, perceived stress, social intelligence, information processing styles and life satisfaction among college students. Curr Psychol 2020 Jun;39(3):927-933. [doi: 10.1007/s12144-018-9804-z]
13. Szabo M, Lovibond PF. Development and Psychometric Properties of the DASS-Youth (DASS-Y): an extension of the Depression Anxiety Stress Scales (DASS) to adolescents and children. Front Psychol 2022;13:766890. [doi: 10.3389/fpsyg.2022.766890] [Medline: 35496218]
14. Birk RH, Samuel G. Digital phenotyping for mental health: reviewing the challenges of using data to monitor and predict mental health problems. Curr Psychiatry Rep 2022 Oct;24(10):523-528. [doi: 10.1007/s11920-022-01358-9] [Medline: 36001220]
15. H. Birk R, Samuel G. Can digital data diagnose mental health problems? A sociological exploration of 'digital phenotyping'. Sociology Health & Illness 2020 Nov;42(8):1873-1887. [doi: 10.1111/1467-9566.13175]
16. Bakker D, Rickard N. Engagement in mobile phone app for self-monitoring of emotional wellbeing predicts changes in mental health: MoodPrism. J Affect Disord 2018 Feb;227:432-442. [doi: 10.1016/j.jad.2017.11.016] [Medline: 29154165]
17. Byrne DG, Davenport SC, Mazanov J. Profiles of adolescent stress: the development of the Adolescent Stress Questionnaire (ASQ). J Adolesc 2007 Jun;30(3):393-416. [doi: 10.1016/j.adolescence.2006.04.004] [Medline: 16750846]
18. Troy AS, Mauss IB. Resilience in the Face of Stress: Emotion Regulation as a Protective Factor: Cambridge University Press eBooks; 2011:30-44. [doi: 10.1017/CBO9780511994791.004]
19. Bouma EMC, Ormel J, Verhulst FC, Oldehinkel AJ. Stressful life events and depressive problems in early adolescent boys and girls: the influence of parental depression, temperament and family environment. J Affect Disord 2008 Jan;105(1-3):185-193. [doi: 10.1016/j.jad.2007.05.007] [Medline: 17574686]
20. Waaktaar T, Borge AIH, Fundingsrud HP, Christie HJ, Torgersen S. The role of stressful life events in the development of depressive symptoms in adolescence--a longitudinal community study. J Adolesc 2004 Apr;27(2):153-163. [doi: 10.1016/j.adolescence.2003.09.006] [Medline: 15023515]
21. Nisenson M, Lin V, Gansner M. Digital phenotyping in child and adolescent psychiatry: a perspective. Harv Rev Psychiatry 2021;29(6):401-408. [doi: 10.1097/HRP.0000000000000310] [Medline: 34313626]
22. Richter A, Adkins V, Selkie E. Youth perspectives on the recommended age of mobile phone adoption: survey study. JMIR Pediatr Parent 2022 Oct 31;5(4):e40704. [doi: 10.2196/40704] [Medline: 36315219]
23. Riiser K, Helseth S, Haraldstad K, Torbjørnsen A, Richardsen KR. Adolescents' health literacy, health protective measures, and health-related quality of life during the Covid-19 pandemic. PLoS ONE 2020;15(8):e0238161. [doi: 10.1371/journal.pone.0238161] [Medline: 32857806]
24. Sequeira L, Perrotta S, LaGrassa J, et al. Mobile and wearable technology for monitoring depressive symptoms in children and adolescents: a scoping review. J Affect Disord 2020 Mar 15;265:314-324. [doi: 10.1016/j.jad.2019.11.156] [Medline: 32090755]
25. Stone AA, Shiffman S. Ecological momentary assessment (EMA) in behavioral medicine. Ann Behav Med 1994 Sep 1;16(3):199-202. [doi: 10.1093/abm/16.3.199]
26. Fahrenberg J, Myrtek M, Pawlik K, Perrez M. Ambulatory assessment - monitoring behavior in daily life settings. Eur J Psychol Assess 2007 Jan;23(4):206-213. [doi: 10.1027/1015-5759.23.4.206]
27. Piasecki TM, Hufford MR, Solhan M, Trull TJ. Assessing clients in their natural environments with electronic diaries: rationale, benefits, limitations, and barriers. Psychol Assess 2007 Mar;19(1):25-43. [doi: 10.1037/1040-3590.19.1.25] [Medline: 17371121]
28. Trull TJ, Ebner-Priemer UW. Using experience sampling methods/ecological momentary assessment (ESM/EMA) in clinical assessment and clinical research: introduction to the special section. Psychol Assess 2009 Dec;21(4):457-462. [doi: 10.1037/a0017653] [Medline: 19947780]
29. Ebner-Priemer UW, Eid M, Kleindienst N, Stabenow S, Trull TJ. Analytic strategies for understanding affective (in)stability and other dynamic processes in psychopathology. J Abnorm Psychol 2009 Feb;118(1):195-202. [doi: 10.1037/a0014868] [Medline: 19222325]
30. Solhan MB, Trull TJ, Jahng S, Wood PK. Clinical assessment of affective instability: comparing EMA indices, questionnaire reports, and retrospective recall. Psychol Assess 2009 Sep;21(3):425-436. [doi: 10.1037/a0016869] [Medline: 19719353]
31. Wilhelm FH, Roth WT. Taking the laboratory to the skies: ambulatory assessment of self-report, autonomic, and respiratory responses in flying phobia. Psychophysiology 1998 Sep;35(5):596-606. [doi: 10.1017/s0048577298970196] [Medline: 9715103]

XSL•FO
RenderX

32. Stiglmayr CE, Ebner-Priemer UW, Bretz J, et al. Dissociative symptoms are positively related to stress in borderline personality disorder. Acta Psychiatr Scand 2008 Feb;117(2):139-147. [doi: 10.1111/j.1600-0447.2007.01126.x] [Medline: 18028248]
33. Armey MF, Schatten HT, Haradhvala N, Miller IW. Ecological momentary assessment (EMA) of depression-related phenomena. Curr Opin Psychol 2015 Aug 1;4:21-25. [doi: 10.1016/j.copsyc.2015.01.002] [Medline: 25664334]
34. Oudin A, Maatoug R, Bourla A, et al. Digital phenotyping: data-driven psychiatry to redefine mental health. J Med Internet Res 2023 Oct 4;25:e44502. [doi: 10.2196/44502] [Medline: 37792430]
35. Rapp A, Cena F. Personal informatics for everyday life: how users without prior self-tracking experience engage with personal data. Int J Hum Comput Stud 2016 Oct;94:1-17. [doi: 10.1016/j.ijhcs.2016.05.006]
36. Sarzotti F. Self-monitoring of emotions and mood using a tangible approach. Computers 2018;7(1):7. [doi: 10.3390/computers7010007]
37. Bandura A. Self‐efficacy. In: The Corsini Encyclopedia of Psychology: John Wiley & Sons; 2010:1-3. [doi: 10.1002/9780470479216]
38. Fan L, Cui F. Mindfulness, self-efficacy, and self-regulation as predictors of psychological well-being in EFL learners. Front Psychol 2024;15:1332002. [doi: 10.3389/fpsyg.2024.1332002] [Medline: 38601825]
39. Clarke J, Proudfoot J, Birch MR, et al. Effects of mental health self-efficacy on outcomes of a mobile phone and web intervention for mild-to-moderate depression, anxiety and stress: secondary analysis of a randomised controlled trial. BMC Psychiatry 2014 Sep 26;14(1):272. [doi: 10.1186/s12888-014-0272-1] [Medline: 25252853]
40. Hoffman AJ. Enhancing self-efficacy for optimized patient outcomes through the theory of symptom self-management. Cancer Nurs 2013;36(1):E16-E26. [doi: 10.1097/NCC.0b013e31824a730a] [Medline: 22495550]
41. Beiwinkel T, Hey S, Bock O, Rössler W. Supportive mental health self-monitoring among smartphone users with psychological distress: protocol for a fully mobile randomized controlled trial. Front Public Health 2017;5:249. [doi: 10.3389/fpubh.2017.00249] [Medline: 28983477]
42. Burns JR, Rapee RM. Screening for mental health risk in high schools: the development of the Youth RADAR. Psychol Assess 2016;28(10):1220-1231. [doi: 10.1037/pas0000237]
43. Jorm AF, Wright A, Morgan AJ. Beliefs about appropriate first aid for young people with mental disorders: findings from an Australian national survey of youth and parents. Early Interv Psychiatry 2007 Feb;1(1):61-70. [doi: 10.1111/j.1751-7893.2007.00012.x] [Medline: 21352109]
44. Kauer SD, Reid SC, Crooke AHD, et al. Self-monitoring using mobile phones in the early stages of adolescent depression: randomized controlled trial. J Med Internet Res 2012 Jun 25;14(3):e67. [doi: 10.2196/jmir.1858] [Medline: 22732135]
45. Wichers M, Simons CJP, Kramer IMA, et al. Momentary assessment technology as a tool to help patients with depression help themselves. Acta Psychiatr Scand 2011 Oct;124(4):262-272. [doi: 10.1111/j.1600-0447.2011.01749.x] [Medline: 21838742]
46. Suveg C, Hoffman B, Zeman JL, Thomassin K. Common and specific emotion-related predictors of anxious and depressive symptoms in youth. Child Psychiatry Hum Dev 2009 Jun;40(2):223-239. [doi: 10.1007/s10578-008-0121-x]
47. Kordy H, Wolf M, Aulich K, et al. Internet-delivered disease management for recurrent depression: a multicenter randomized controlled trial. Psychother Psychosom 2016;85(2):91-98. [doi: 10.1159/000441951] [Medline: 26808817]
48. Torous J, Friedman R, Keshavan M. Smartphone ownership and interest in mobile applications to monitor symptoms of mental health conditions. JMIR Mhealth Uhealth 2014 Jan 21;2(1):e2. [doi: 10.2196/mhealth.2994] [Medline: 25098314]
49. Harrington R, Loffredo DA. Insight, rumination, and self-reflection as predictors of well-being. J Psychol 2010 Nov 30;145(1):39-57. [doi: 10.1080/00223980.2010.528072]
50. Cho M, Park D, Choo M, Kim J, Han DH. Development and initial evaluation of a digital phenotype collection system for adolescents: proof-of-concept study. JMIR Preprints. Preprint posted online on Apr 17, 2024. [doi: 10.2196/preprints.59623]
51. Lovibond PF, Lovibond SH. The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. Behav Res Ther 1995 Mar;33(3):335-343. [doi: 10.1016/0005-7967(94)00075-u] [Medline: 7726811]
52. Henry JD, Crawford JR. The short-form version of the Depression Anxiety Stress Scales (DASS-21): construct validity and normative data in a large non-clinical sample. Br J Clin Psychol 2005 Jun;44(Pt 2):227-239. [doi: 10.1348/014466505X29657] [Medline: 16004657]
53. Lee EH, Moon SH, Cho MS, et al. The 21-Item and 12-Item versions of the depression anxiety stress scales: psychometric evaluation in a Korean population. Asian Nurs Res (Korean Soc Nurs Sci) 2019 Feb;13(1):30-37. [doi: 10.1016/j.anr.2018.11.006] [Medline: 30503903]
54. Sherer M, Adams CH. Construct validation of the Self-Efficacy Scale. Psychol Rep 1983 Dec;53(3):899-902. [doi: 10.2466/pr0.1983.53.3.899]
55. Lee YS, Park SH, Kim JK. Nursing students' self-leadership, self-efficacy, interpersonal relation, college life satisfaction. The Journal of the Korea Contents Association 2014 Jun 28;14(6):229-240. [doi: 10.5392/JKCA.2014.14.06.229]
56. Kim SM, Yoo MH. Comparison of time management behavior, life satisfaction between gifted middle school students and general students, and relationship between the time management behaviors and life satisfaction. Journal of Gifted/Talented Education 2013 Jun 30;23(3):315-333. [doi: 10.9722/JGTE.2013.23.3.315]

XSL•FO
RenderX

57. Zhou L, Bao J, Setiawan IMA, Saptono A, Parmanto B. The mHealth App Usability Questionnaire (MAUQ): development and validation study. JMIR Mhealth Uhealth 2019;7(4):e11500. [doi: 10.2196/11500]
58. Saeb S, Lattie EG, Schueller SM, Kording KP, Mohr DC. The relationship between mobile phone location sensor data and depressive symptom severity. PeerJ 2016;4:e2537. [doi: 10.7717/peerj.2537] [Medline: 28344895]
59. Saeb S, Lattie EG, Kording KP, Mohr DC. Mobile phone detection of semantic location and its relationship to depression and anxiety. JMIR Mhealth Uhealth 2017 Aug 10;5(8):e112. [doi: 10.2196/mhealth.7297] [Medline: 28798010]
60. De Angel V, Lewis S, White K, et al. Digital health tools for the passive monitoring of depression: a systematic review of methods. NPJ Digit Med 2022 Jan 11;5(1):3. [doi: 10.1038/s41746-021-00548-8] [Medline: 35017634]
61. Price GD, Heinz MV, Song SH, Nemesure MD, Jacobson NC. Using digital phenotyping to capture depression symptom variability: detecting naturalistic variability in depression symptoms across one year using passively collected wearable movement and sleep data. Transl Psychiatry 2023 Dec 9;13(1):381. [doi: 10.1038/s41398-023-02669-y] [Medline: 38071317]
62. Jacobson NC, Summers B, Wilhelm S. Digital biomarkers of social anxiety severity: digital phenotyping using passive smartphone sensors. J Med Internet Res 2024;22(5):e16875. [doi: 10.2196/16875]
63. Sheikh M, Qassem M, Kyriacou PA. Wearable, environmental, and smartphone-based passive sensing for mental health monitoring. Front Digit Health 2021;3:662811. [doi: 10.3389/fdgth.2021.662811] [Medline: 34713137]

Abbreviations

EMA: ecological momentary assessment

Edited by A Kushniruk; submitted 22.09.24; peer-reviewed by R Nooripour; revised version received 30.05.25; accepted 03.06.25; published 07.10.25.

Please cite as:
Cho M, Park D, Choo M, Han DH, Kim J
Impact of Digital Phenotypes and Question-Asking on Emotional Disorders in Adolescents: 4-Week Field Study
JMIR Hum Factors 2025;12:e66536
URL: https://humanfactors.jmir.org/2025/1/e66536
doi:10.2196/66536

© Minseo Cho, Doeun Park, Myounglee Choo, Doug Hyun Han, Jinwoo Kim. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 7.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Applying Human-Centered Design to Develop Smartphone-Based Intervention Messages to Help Young Adults Quit Using E-Cigarettes and Cigarettes: A Remote User Testing Study

Thi Phuong Thao Tran[1,2], PhD; Christine Tran[3,4], BSc; Pamela M Ling[4,5], PhD; Lucy Popova[1], PhD; Nhung Nguyen[3,4,5], PhD

[1]School of Public Health, Georgia State University, Atlanta, GA, United States

[2]College of Health Sciences, VinUniversity, Hanoi, Vietnam

[3]HEARTY lab, University of California San Francisco, San Francisco, CA, United States

[4]Division of General Internal Medicine, Department of Medicine, University of California San Francisco, San Francisco, CA, United States

[5]UCSF Center for Tobacco Control Research and Education and Division of General Internal Medicine, 530 Parnassus Ave, San Francisco, CA, United States

Corresponding Author:
Nhung Nguyen, PhD
HEARTY lab, University of California San Francisco, San Francisco, CA, United States

Abstract

Background: Despite the popularity of concurrent use of electronic cigarettes (e-cigarettes) and cigarettes (dual tobacco use) among young adults, few interventions address the cessation of both tobacco products. The application of a human-centered design (HCD) approach in the development of such interventions remains limited.

Objective: This study used an HCD approach to develop smartphone-based intervention messages for dual tobacco cessation for young adults.

Methods: Intervention messages were developed based on theories, cessation guidelines, existing messages, and our previous formative study. Three rounds of message testing were conducted asynchronously via an online platform with 35 young adults (18 - 29 years old) who currently used both e-cigarettes and cigarettes and were motivated to quit either smoking or vaping in the next 6 months. In each round, a new sample of 10 - 12 participants evaluated the messages individually. For the quantitative assessment, participants viewed and rated each message on a scale from 1 ("very low degree") to 5 ("very high degree") across 4 components: Comprehension ("This message is easy to understand"), Usefulness ("This message is useful for encouraging tobacco cessation"), Tone ("The language is clear and non-judgmental"), and Design ("The design is appealing"). For the qualitative assessment, participants used a platform-enabled feature to place markers on specific parts of messages they liked, disliked, or found confusing and then provided brief explanations for their feedback. Initial messages were assessed during the first 2 rounds of testing, and those with low mean scores were revised based on participants' feedback and retested in the third round.

Results: We found significant improvements in message ratings after refinement. The overall mean score increased from 3.6 (SD 0.4) to 4.6 (SD 0.2) (P<.001), using paired t tests. Specifically, the mean score of "Comprehension" improved from 4.0 (SD 0.5) to 4.9 (SD 0.2) (P<.001), the mean score of "Usefulness" increased from 3.0 (SD 0.6) to 4.4 (SD 0.4) (P<.001), the mean score of "Tone" increased from 3.8 (SD 0.6) to 4.8 (SD 0.2) (P<.001), and the mean score of "Design" increased from 3.4 (SD 0.48) to 4.4 (SD 0.3) (P<.001). The qualitative assessments highlighted design elements related to message liking, such as clear layout, minimalistic imagery, italicized quotes, and highlighted keywords. Conversely, design features related to message dislike included color shades, lengthy text, and confusing wording.

Conclusions: This study demonstrated the use of HCD in developing smartphone-based intervention messages to support dual tobacco cessation among young adults. Integrating remote message testing improved the feasibility of rapid prototyping while enhancing the relevance and appeal of message content and design. Future interventions targeting emerging health behaviors among young adults may benefit from incorporating a remote testing method to efficiently gather user feedback and refine intervention messages in a timely manner.

(JMIR Hum Factors 2025;12:e76503) doi:10.2196/76503

KEYWORDS

tobacco cessation; vaping; smoking; dual tobacco use; human-centered design; mHealth; smartphone intervention; young adults; mobile health

XSL•FO
RenderX

Introduction

Concurrent use of electronic cigarettes (e-cigarettes) and cigarettes in the past 30 days (dual tobacco use) is prevalent among US young adults aged 18 to 29 years. In 2021, among young adults who smoked cigarettes, 15% - 30% also used e-cigarettes; among young adults who used e-cigarettes, 30% - 67% also smoked cigarettes [1,2]. This dual-use behavior is associated with a higher risk of various health conditions, such as stroke, metabolic dysfunction, asthma, chronic obstructive pulmonary disease, and oral disease [3], and has been linked to an increased risk of nicotine dependence, subsequent substance use [4], and more difficulty with cessation [5]. Most young adults who use tobacco express a strong desire to quit and want to stop using all tobacco and nicotine products [6,7]. Yet, cessation interventions addressing both e-cigarettes and cigarettes remain limited [8].

Smartphone-based apps could offer cost-effective access to evidence-based cessation support for young adults [9], as over 95% of them own smartphones [10]. Of several available smartphone-based cessation apps, almost all were either for smoking cessation or vaping cessation [11], offered limited features for vaping cessation [12], and lacked rigorous development and testing [13-15]. Furthermore, existing apps typically adopt a "one-size-fits-all" approach and often fail to incorporate meaningful input from end users early in the development process [16]. This limited evidence on effective strategies addressing dual tobacco use among young adults highlights the need for developing new intervention messages addressing the use of both e-cigarettes and cigarettes among young adults [8,17].

A human-centered design (HCD) approach can be used to develop a smartphone app intervention that meets a variety of individual cessation needs [18]. HCD is defined as "an approach to systems design and development that aims to make interactive systems more usable by focusing on the use of the system and applying human factors/ergonomics and usability knowledge and techniques" [19]. Its systematic approach engages with and prioritizes the needs and preferences of end users in the development of an intervention, ensuring that interventions are likely to be used effectively in a real-world context [18]. We used the IDEAS framework (including 4 stages: Integrate, DEsign, Assess, and Share, described in the next paragraph), which integrates behavioral theory, design thinking, and intervention evaluation and dissemination into systematic guidance for the development of effective interventions [20]. This framework has guided the development of previous substance use interventions, including tobacco smoking cessation [21-23]. There are 10 phases (Empathize, Specify, Ground, Ideate, Prototype, Gather, Build, Pilot, Evaluate, and Share) across the 4 stages. Although HCD is increasingly used in mobile health interventions, there is limited published research on how the HCD approach is applied to developing interventions for young adult dual tobacco cessation [24].

By applying the HCD approach, the first 3 phases in the "*Integrate*" stage (empathizing with the target users, specifying targets for future interventions, and grounding findings in behavioral theories) were conducted in our previous formative research using in-depth interviews with young adults to elicit needs and preferences for dual tobacco cessation interventions [6]. This formative work led to our "*Design*" stage, in which we aim to ideate intervention strategies, prototype intervention messages, gather user feedback, and build an intervention prototype that includes smartphone-based intervention messages for dual tobacco cessation among young adults. The developed intervention prototype will then be used in the next stage, "*Access*," where we plan to pilot and evaluate its efficacy in a randomized controlled trial (RCT). Focusing on the "*Design*" stage, this study aims to describe our process for ideating, prototyping, and refining intervention message mockups through iterative feedback from young adults. Our remote HCD approach may serve as a methodological model for message development and remote testing in future digital health interventions.

Methods

Design

This study is the second phase of a project aiming to develop a smartphone-based intervention targeting dual tobacco cessation among young adults. In the first phase, we conducted a needs assessment study to understand young adults' desires and preferences for cessation of both e-cigarette vaping and cigarette smoking [6]. In the second phase (ie, this study), we developed intervention messages for use in the smartphone-based intervention, which will be evaluated in an RCT in the next phase.

Participants

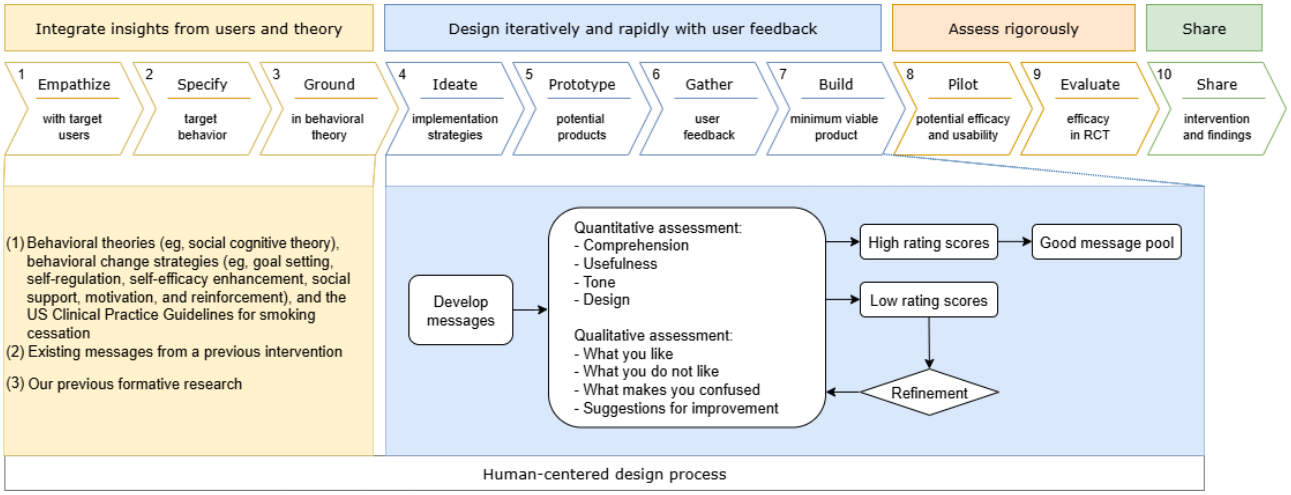
Participants were recruited through targeted advertisements on Instagram linking to the study website. Eligible participants were young adult men and women (aged 18 to 29 years) who: (1) speak English, (2) own a smartphone, (3) report use of both cigarettes and e-cigarettes at least weekly during the past month [6], and (4) intend to quit either smoking or vaping in the next 6 months. Overall, 35 participants were enrolled in message testing.

Message Development

The HCD approach in the development of tobacco cessation messages is displayed in Figure 1. The message development was guided by behavioral theories (eg, Social Cognitive Theory), behavioral change strategies (eg, goal setting, self-regulation, self-efficacy enhancement, social support, motivation, and reinforcement) [25], and the US Clinical Practice Guidelines for smoking cessation [26]. The message content was also based on existing messages in a previous intervention that effectively increased smoking abstinence among young adults [26]. We further refined message content to specifically target young adult dual tobacco users' needs and preferences determined previously in our formative research, including providing cessation support to quit one prioritized product while preventing switching between 2 products, and emphasizing unique facilitators and barriers to quitting each product (eg, unpleasant smell of cigarettes facilitating smoking cessation and accessibility of e-cigarettes hindering vaping cessation) [6].

XSL•FO
RenderX

Figure 1. Human-centered design approach in the development of tobacco cessation messages for young adults who use both electronic cigarettes and cigarettes. RCT: randomized controlled trial.



To fit with smartphone interfaces, we then designed 90 mockups of the messages. We added design elements (eg, color and style of text, graphics, and pictures) to the text messages to enhance visual aesthetics and user engagement.

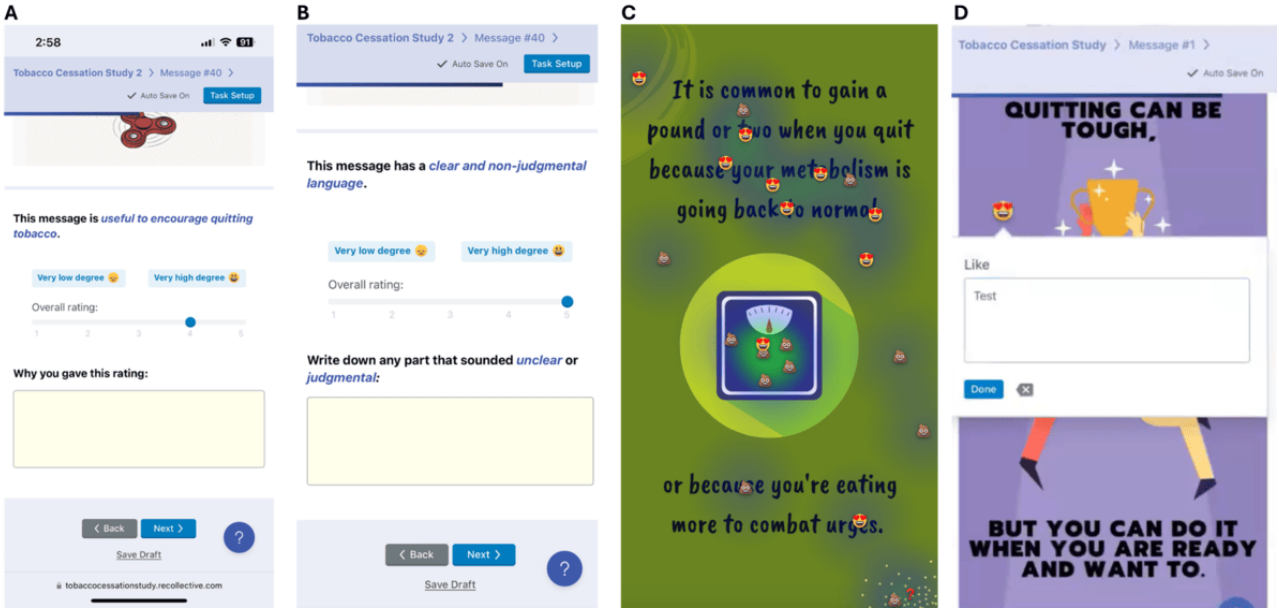
Message Testing and Refinement

The testing was conducted remotely via Recollective (Recollective Inc), a platform for conducting user experience research. Each participant logged into the platform via phone or computer to complete the message rating tasks at a time convenient for them. This flexible and self-paced approach with asynchronous features could enhance accessibility to the testing process. We conducted 3 rounds of testing, with the first 2 rounds assessing the developed messages (prerefinement) and the third round retesting the refined messages (postrefinement). To enhance the representativeness of the target users and gather diverse insights, we included a new sample of 10 - 12 young adults in each round. This approach also minimized participant fatigue and disengagement through repeated rounds of testing. We provided participants with a video with detailed instructions to facilitate their message assessment. Participants evaluated the messages individually, with no access to others' assessments. Each message rating task took 3 - 5 minutes to complete. To avoid overburdening participants and allow sufficient time to complete all message ratings, we asked participants to evaluate 45 messages within 2 weeks in each round of testing (a total of 2 testing rounds). Our team monitored participants' progress daily and sent reminders to ensure the message evaluation was completed within the predefined timeframe.

Messages were evaluated using both quantitative and qualitative assessments. For the quantitative assessment, participants viewed and rated each message on a scale from 1="very low degree" to 5="very high degree" across 4 components, with 1 question for each component: Comprehension (ie, "This message is easy to understand"), Usefulness (ie, "This message is useful for encouraging tobacco cessation"), Tone (ie, "The language is clear and non-judgmental"), and Design (ie, "The design is appealing") [27,28], together with optional text box provided for participants to explain their ratings (Figure 2A and 2B). For the qualitative assessment, participants could place an unlimited number of markers on specific parts of the messages they liked, disliked, or found confusing, without seeing others' markers (Figure 2C). For each marker placed, participants were asked to type in a pop-up text box to briefly explain their specific reasons for liking, disliking, or confusion (Figure 2D).

XSL•FO
RenderX

Figure 2. Screenshots illustrating the testing process: (A and B) examples of component rating, (C) markers indicating whether participants liked, disliked, or found messages confusing, and (D) a text box for providing brief explanations of their specific reasons. The icons on the original message indicate the specific parts that participants liked ( ), disliked ( ), or found confusing ( ).



All 90 developed messages were revised based on participants' feedback. Due to time and resource constraints for rapid testing, we retested only 35 out of 90 messages with low rating mean scores (eg, a mean score of less than 4 out of the maximum score of 5) in the subsequent round, while the remaining messages with high rating scores were also revised but not retested [24].

We encountered a small number of open-ended comments that were ambiguous or lacked sufficient context for accurate interpretation. Our team carefully reviewed these comments and collaboratively discussed potential meanings. When a comment could not be reasonably interpreted through consensus discussion, it was excluded from the message revision to avoid potential misrepresentation. These exclusions were minimal and ensured that only responses with sufficient clarity informed the message refinement.

Data Analysis

For the qualitative analysis of participants' comments, we summarized their feedback on design features, confusing content, notes on usefulness and tone, and suggestions for message improvement.

To evaluate the overall improvement of the 35 retested messages, we compared the mean scores across the 4 components (Comprehension, Usefulness, Tone, and Design) and the overall score before and after refinement. Messages that initially had high scores and did not require refinement were excluded from the analysis. Paired *t* tests were conducted to examine whether the postrefinement mean scores were significantly higher than the prerefinement mean scores. The quantitative data analysis was performed using STATA software version 18 (StataCorp LLC).

Ethical Considerations

The study was approved by the Institutional Review Board at the University of California, San Francisco (no. 352216). All participants were provided with the study information and completed electronic informed consent on the study website. An identification number was assigned to each participant to maintain anonymity, and all data were deidentified. Participants received a US $100 e-gift card for their participation.

Results

Participants' Characteristics

The demographic characteristics of study participants are shown in Table 1. The study sample had a mean age of 22 years (SD 3). In total, 20 out of 35 participants (57%) were female and 26 participants (74%) were not Hispanic, Latino, or of Spanish origin. Regarding race, the largest group was White (n=20, 57%), followed by multiracial individuals (n=6, 17%), Asian (n=4, 11%), and Black or African American (n=3, 9%). In addition, 27 out of 35 participants (77%) were currently enrolled in college.

XSL•FO
RenderX

Table . Demographic characteristics of study participants (N=35).

Characteristics	Value
Age (years), n (%)	
Min-max	18 - 29
Mean (SD)	22 (3)
Gender, n (%)	
Female	20 (57)
Male	15 (43)
Ethnicity, n (%)	
Hispanic, Latino, or Spanish	9 (26)
Not Hispanic, Latino, or Spanish	26 (74)
Race, n (%)	
White	20 (57)
Asian (Chinese, Vietnamese, Asian Indian)	4 (11)
Black or African American	3 (9)
American Indian or Alaska Native	1 (3)
Middle Eastern	1 (3)
Multiracial	6 (17)
Education attainment, n (%)	
Currently in college	27 (77)
Currently in graduate school	2 (6)
Currently in professional or technical school	1 (3)
Not attending school	5 (14)

Results of Message Development

To address dual tobacco cessation, we developed messages that encourage quitting all tobacco and nicotine products by using terms “any tobacco” or “nicotine.” Another strategy was to develop messages that address shared barriers (eg, craving) and facilitators (eg, social support and rewarding) for quitting both e-cigarettes and cigarettes, as identified in our previous formative research. In addition, our previous study indicated that users usually did not attempt to quit both tobacco products simultaneously but instead prioritized quitting one before the other. Therefore, in addition to promoting dual tobacco cessation, we also developed single-product cessation messages. This tailored approach allows message delivery to align with an individual’s preferred quitting sequence, which can be implemented in the next phase. For instance, if a user prioritized quitting smoking, they would receive more messages focused on smoking cessation while still being exposed to messages encouraging the cessation of all tobacco and nicotine products in our future RCT. Moreover, we also developed messages to prevent switching between e-cigarettes and cigarettes. Multimedia Appendix 1 provides examples of these messages before refinement, including those promoting dual tobacco cessation, addressing shared barriers and facilitators to quitting both products (eg, craving management, social support, and rewarding), targeting single-product cessation, and discouraging switching between e-cigarettes and cigarettes.

Results of Message Testing and Refinement

Overview

After each round of testing, we revised the messages based on participants’ feedback. One key challenge in this refinement process was balancing the diverse needs and preferences of young adults. Participants often provided multiple, sometimes conflicting, suggestions for improving the messages, making it difficult to determine the best approach to fully address the diverse feedback. To address this issue, our team balanced their needs with practical feasibility and resource restraints. In some cases, we created multiple versions of messages tailored to different subgroups. Multimedia Appendix 2 shows an example of two revised versions of a message promoting exercise to manage nicotine cravings, tailored separately for males and females.

Quantitative Results

To meet the rapid timeline of iterative testing cycles, we prioritized refining the 35 messages with low ratings and retested them in the subsequent round. Multimedia Appendix 3 presents average scores for the 35 retested messages before and after refinement, covering 4 components. All 4 components showed significant improvement after refinement: Comprehension increased from 4.0 (SD 0.5) to 4.9 (SD 0.2), Usefulness rose from 3 (SD 0.6) to 4.4 (SD 0.4), Tone increased from 3.8 (SD 0.6) to 4.8 (SD 0.2), and Design increased from 3.4 (SD 0.48)

XSL•FO
RenderX

to 4.4 (SD 0.3). Overall, the mean score increased significantly after refinement, rising from 3.6 (SD 0.4) to 4.6 (SD 0.2). All paired *t* tests showed statistically significant results with $P<.001$.

Qualitative Results

Overall, based on the participants' comments during the message testing, the design features that participants liked the most included a clear layout, minimalistic imagery, italicized quotes, and highlighted keywords. The design features of message disliking included color shades, lengthy text, and confusing words. Given the large number of refined messages, below we provide 2 examples illustrating how specific messages were revised based on participants' feedback.

Figure 3A provides an example of a message encouraging self-reward after quitting. While many participants liked illustrations and colors, some found the original messages confusing, with one stating, "*a bit confused as to why the bottom half of the left tooth looks like cheese.*" To address this, we revised the graphic by removing the yellow streaks and replacing them with smiley teeth illustrations to maintain consistency across messages. In addition, we removed the word "*dentist,*" as many participants expressed their dislike of this word. Some participants also found the message somewhat unclear and "*seems kind of out of left field, not extremely relevant,*" and suggested "*change the text to make it more convincing.*" In response, we revised the bottom text to be more direct and affirmative, aligning with a common theme of the message.

XSL•FO
RenderX

Figure 3. Two examples of messages before and after refinement based on participants' feedback.

XSL•FO
RenderX

Figure 3B provides an example of a message promoting gratitude and emotional well-being as supportive strategies for tobacco cessation. Some participants expressed that the message seemed unrelated, stating that "*not see how this relates to vaping*," "*not specific to tobacco*," "*not sure if they are trying to convince people to give up tobacco*," and "*not sure what this message is trying to address*." To strengthen its relevance to tobacco cessation, we added an element acknowledging that "*quitting tobacco is hard*." In addition, given that messages telling people to be grateful can come off wrong to some people, we refined it as a coping strategy for quitting. The bottom part was refined to be more specific and actionable, encouraging participants to identify 3 things they are grateful for.

Discussion

Principal Results

This study showcases an HCD approach to develop messages for a smartphone-based intervention to support young adults' dual tobacco cessation. As an early stage of intervention development, we built a low-cost and low-fidelity prototype of the intervention (mockups of intervention messages) that could be developed rapidly to gather feedback from target users early and often, allowing exploration of alternative solutions before investing significant resources in any one particular or potentially suboptimal solution. We found that the HCD approach significantly improved the comprehension, usefulness, tone, and design of the intervention messages, which can ultimately increase the intervention's real-world impact and end-user engagement. As a methodological example of remote message testing, our study demonstrated the feasibility of engaging young adults in remote user experience testing without the need for face-to-face interactions, which can speed up the intervention development process while saving research costs. Another notable contribution of this study is to develop intervention messages specifically for dual tobacco cessation among young adults—a critical gap in the literature.

In this study, we applied the HCD principles to translate evidence-based tobacco cessation treatment into a real-world solution. Aligned with a traditional approach, we initially developed core message content based on a comprehensive evidence base, combining best practices for tobacco treatment, behavioral theories, behavior change techniques, previous interventions, and our formative needs assessment. Traditional message development methods are typically content-focused and expert-driven, which could overlook the lived experiences, preferences, and contextual needs of the end users. Rooted in empathy-driven focus, our HCD approach addresses these issues by prioritizing the preferences of end users and integrating design strategies that enhance the engagement and usability of the core content. Particularly, our participants' iterative feedback revealed that messages could be made more appealing by incorporating a clear layout, minimalistic imagery, italicized quotes, and highlighted keywords while avoiding lengthy text and confusing wording. This is especially important for smartphone apps, as well-designed message interfaces can significantly influence how young adults perceive messages' clarity, appeal, and the overall effectiveness in supporting tobacco cessation.

Comparison with Prior Work

Unlike previous smartphone app interventions that primarily focused on single-product smoking cessation (eg, cigarette use) and were often grounded in single behavioral change theory, our study used a more robust approach to intervention design. We integrated multiple behavioral theories or techniques and clinical guidance with users' needs and preferences. This user-centered design process ensured that the messages were not only scientifically evidence-based but also relevant and engaging in a real-world context. Previous studies used an HCD approach to design mHealth apps for smoking cessation among pregnant women [29], people with serious mental illness [30], and people with HIV [29-33]. Consistent with our findings, user-centered design and usability testing in these studies helped align the interventions with users' needs for smoking cessation. Moreover, this approach also enhances the ability to create multiple tailored message versions (eg, for males and females), enhancing the personalization of interventions, which is an important direction highlighted in previous mHealth research [34].

A key contribution of this study is the use of remote quantitative and qualitative assessments from target users, offering greater flexibility and accessibility than traditional in-person feedback collection. This approach enabled participants to provide detailed input not only on content but also on specific design elements, such as colors, symbols, and other visual features. The entire process of message testing was conducted remotely, with the self-guided video instruction provided beforehand, which allowed participants to complete the tasks independently. Unlike prior studies that relied heavily on researcher-facilitated interviews or focus groups, our method reduced the need for real-time involvement from the research team, minimizing the time and resource burden on both researchers and participants. This low-cost, asynchronous approach also enhances access to user testing and may facilitate more open and honest feedback as participants provide responses privately, potentially reducing social desirability bias often present in interviews and the "groupthink" effect commonly observed in focus groups [35]. Overall, our remote testing method presents a promising and scalable methodology for future message development efforts targeting emerging health behaviors among young adults.

Future Research

The final message set will be implemented in the next phase, "*Assess*," in which we will deliver the intervention via smartphone app and evaluate its efficacy and effectiveness on dual tobacco cessation among young adults. Past research highlighted the need to investigate important mediators of smartphone interventions' success, such as psychological empowerment (the intervention helping one find ways to think and behave differently), hedonic well-being (the intervention making one happier), and inspiration (the intervention motivating one to be a person they want to be) [36]. These perceptions are influenced by the content of the intervention and are, in turn, predictive of behavior, such that psychological empowerment and hedonic well-being are positively, but

XSL•FO
RenderX

inspiration is negatively, associated with quitting and reducing smoking [36]. Thus, in future studies, it would be worthwhile to not only assess the impact of the intervention on tobacco cessation, but also examine these and other potential mechanisms of impact.

Limitations

First, the participants in our study may not fully represent the broader young adult population, and recruiting through social media may have restricted our reach to more diverse groups. This may reflect the perspectives and experiences of White younger individuals who are more digitally connected, limiting the breadth of feedback for the intervention messages and the generalizability of our findings to the broader target population. Therefore, future studies could consider alternative recruitment strategies to ensure a more representative sample. Second, some open-ended comments from participants were unclear, which posed challenges for interpretation. Finally, since our repeated testing involved different participant samples, significant differences in quantitative assessments may reflect not only improvements to the messages but also variations in sample characteristics. While engaging different groups allowed us to capture feedback from more young adults, this approach may introduce the possibility of confounding due to between-sample differences. To mitigate this limitation, future studies may consider a longitudinal design in which the same participants assess messages over time, allowing for greater control over individual-level variability and clearer attribution of changes to message revisions.

Conclusions

In conclusion, this study offers valuable insights into using a remote HCD approach to develop smartphone-based intervention messages for young adults aiming to quit e-cigarettes and cigarettes. The use of self-instructional videos and asynchronous feedback collection enhanced the accessibility and feasibility of remote message testing. This approach improved message comprehension, usefulness, tone, and design. These findings can inform future message development and remote user testing efforts for tobacco cessation. Moreover, our HCD process is adaptable to other behavioral topics, populations, delivery formats (eg, video and audio), and platforms (eg, social media). Future research can build on this work to optimize engagement and effectiveness in digital health interventions targeting health behavior changes among young adults.

Acknowledgments

We are grateful to all study participants who helped to advance the science of tobacco cessation for young people. We would also like to thank Dr Johannes Thrul for inspiring us to conduct remote user experience research, Dr Kimberly Koester for guiding the use of the Recollective platform, and Sara Perez for helping with the initial study setup. Preliminary findings were presented as a poster at the Society of Behavioral Medicine Annual Meeting 2024. This research was supported by the National Institutes of Health (NIH)/National Institute on Drug Abuse (K01DA056693), the California Tobacco-Related Disease Research Program (T32KT5071), and the NIH/National Center for Advancing Translational Sciences through UCSF-CTSI (UL1 TR001872). NN is also supported by the IPA HSR & D from the US Department of Veterans Affairs Health Services and the Hellman Fellowship Program. The content is solely the authors' responsibility and does not necessarily represent the official views of the funding agencies.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

All authors have been involved in analyzing the data, interpreting the results, and writing the manuscript. TPTT conducted the analysis, drafted the initial manuscript, and contributed to all subsequent drafts of the manuscript. CT assisted with the data collection and analysis and reviewed and revised the manuscript. LP assisted with the interpretation of the result and reviewed and revised the manuscript. PML assisted with the interpretation of the result and reviewed and revised the manuscript. NN obtained funding, conceptualized the study, planned and supervised the data collection and analysis, reviewed and revised the manuscript, and contributed to all subsequent drafts of the manuscript. All authors have read and approved the final manuscript for submission.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Examples of messages before refinement in this study.
[DOCX File, 1021 KB - humanfactors_v12i1e76503_app1.docx]

Multimedia Appendix 2
Examples of messages tailored to males and females.

XSL•FO
RenderX

[PNG File, 152 KB - humanfactors_v12i1e76503_app2.png]

Multimedia Appendix 3
Average scores for the 35 retested messages before and after refinement, covering 4 components: Comprehension, Usefulness, Tone, and Design.
[PNG File, 39 KB - humanfactors_v12i1e76503_app3.png]

References

1. Erhabor J, Boakye E, Obisesan O, et al. E-cigarette use among US adults in the 2021 behavioral risk factor surveillance system survey. JAMA Netw Open 2023 Nov 1;6(11):e2340859. [doi: 10.1001/jamanetworkopen.2023.40859] [Medline: 37921768]
2. Mayer M, Reyes-Guzman C, Grana R, Choi K, Freedman ND. Demographic characteristics, cigarette smoking, and e-cigarette use among US adults. JAMA Netw Open 2020 Oct 1;3(10):e2020694. [doi: 10.1001/jamanetworkopen.2020.20694] [Medline: 33048127]
3. Glantz SA, Nguyen N, Oliveira da Silva AL. Population-based disease odds for E-cigarettes and dual use versus cigarettes. NEJM Evid 2024 Mar;3(3):EVIDoa2300229. [doi: 10.1056/EVIDoa2300229] [Medline: 38411454]
4. Zhang BY, Bannon OS, Tzu-Hsuan Chen D, Filippidis FT. Dual and poly-nicotine and tobacco use among adolescents in the United States from 2011 to 2022. Addict Behav 2024 May;152:107970. [doi: 10.1016/j.addbeh.2024.107970] [Medline: 38277994]
5. Nguyen N, Koester KA, Kim M, Watkins SL, Ling PM. “I’m both smoking and vaping”: a longitudinal qualitative study of US young adults who tried to quit smoking cigarettes by using electronic cigarettes. Tob Control 2024 Sep 25;33(5):596-602. [doi: 10.1136/tc-2022-057804] [Medline: 37072166]
6. Nguyen N, Koester KA, Tran C, Ling PM. Desires and needs for quitting both e-cigarettes and cigarettes among young adults: formative qualitative study informing the development of a smartphone intervention for dual tobacco cessation. JMIR Form Res 2024 Oct 22;8:e63156. [doi: 10.2196/63156] [Medline: 39437386]
7. Zhang L, Gentzke A, Trivers KF, VanFrank B. Tobacco cessation behaviors among U.S. middle and high school students, 2020. J Adolesc Health 2022 Jan;70(1):147-154. [doi: 10.1016/j.jadohealth.2021.07.011] [Medline: 34930566]
8. Kundu A, Kouzoukas E, Zawertailo L, et al. Scoping review of guidance on cessation interventions for electronic cigarettes and dual electronic and combustible cigarettes use. CMAJ Open 2023;11(2):E336-E344. [doi: 10.9778/cmajo.20210325] [Medline: 37072138]
9. Vilardaga R, Casellas-Pujol E, McClernon JF, Garrison KA. Mobile applications for the treatment of tobacco use and dependence. Curr Addict Rep 2019 Jun;6(2):86-97. [doi: 10.1007/s40429-019-00248-0] [Medline: 32010548]
10. Demographics of mobile device ownership and adoption in the United States. Pew Research Center. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2025-08-23]
11. Micalizzi L, Mattingly DT, Hart JL, Jensen JK, Mahabee-Gittens EM, Garrison KA. Smartphone apps targeting youth tobacco use prevention and cessation: an assessment of credibility and quality. Curr Addict Rep 2023 Dec;10(4):649-663. [doi: 10.1007/s40429-023-00524-0] [Medline: 38680515]
12. Sanchez S, Kundu A, Limanto E, Selby P, Baskerville NB, Chaiton M. Smartphone apps for vaping cessation: quality assessment and content analysis. JMIR Mhealth Uhealth 2022 Mar 28;10(3):e31309. [doi: 10.2196/31309] [Medline: 35343904]
13. McKay F, Chan L, Cerio R, et al. Assessing the quality and behavior change potential of vaping cessation apps: systematic search and assessment. JMIR Mhealth Uhealth 2024 Mar 15;12:e55177. [doi: 10.2196/55177] [Medline: 38532616]
14. Whittaker R, McRobbie H, Bullen C, Rodgers A, Gu Y, Dobson R. Mobile phone text messaging and app-based interventions for smoking cessation. Cochrane Database Syst Rev 2019 Oct 22;10(10):CD006611. [doi: 10.1002/14651858.CD006611.pub5] [Medline: 31638271]
15. Guo YQ, Chen Y, Dabbs AD, Wu Y. The effectiveness of smartphone app-based interventions for assisting smoking cessation: systematic review and meta-analysis. J Med Internet Res 2023 Apr 20;25:e43242. [doi: 10.2196/43242] [Medline: 37079352]
16. An Q, Kelley MM, Hanners A, Yen PY. Sustainable development for mobile health apps using the human-centered design process. JMIR Form Res 2023 Aug 25;7:e45694. [doi: 10.2196/45694] [Medline: 37624639]
17. Berg CJ, Romm KF, Patterson B, Wysota C, Abroms LC. Appeal of novel cessation intervention approaches among young-adult users of traditional and alternative tobacco products. Tob Use Insights 2021;14:1179173X211041123. [doi: 10.1177/1179173X211041123] [Medline: 34658625]
18. Göttgens I, Oertelt-Prigione S. The application of human-centered design approaches in health research and innovation: a narrative review of current practices. JMIR Mhealth Uhealth 2021 Dec 6;9(12):e28102. [doi: 10.2196/28102] [Medline: 34874893]
19. Ergonomics of human system interaction-part 210: human-centred design for interactive systems. : International Standardization Organization (ISO); 2009.

XSL•FO
RenderX

20. Mummah SA, Robinson TN, King AC, Gardner CD, Sutton S. IDEAS (Integrate, Design, Assess, and Share): a framework and toolkit of strategies for the development of more effective digital interventions to change health behavior. J Med Internet Res 2016 Dec 16;18(12):e317. [doi: 10.2196/jmir.5927] [Medline: 27986647]
21. Beck AK, Kelly PJ, Deane FP, et al. Developing a mHealth routine outcome monitoring and feedback app ("SMART Track") to support self-management of addictive behaviours. Front Psychiatry 2021;12:677637. [doi: 10.3389/fpsyt.2021.677637] [Medline: 34220583]
22. Dick S, Vasiliou VS, Davoren MP, et al. A digital substance-use harm reduction intervention for students in higher education (MyUSE): protocol for project development. JMIR Res Protoc 2020 Aug 27;9(8):e17829. [doi: 10.2196/17829] [Medline: 32851980]
23. Sinicrope PS, Koller KR, Prochaska JJ, et al. Social media intervention to promote smoking treatment utilization and cessation among Alaska Native people who smoke: protocol for the Connecting Alaska Native People to Quit Smoking (CAN Quit) Pilot Study. JMIR Res Protoc 2019 Nov 22;8(11):e15155. [doi: 10.2196/15155] [Medline: 31755867]
24. Harte R, Glynn L, Rodríguez-Molinero A, et al. A human-centered design methodology to enhance the usability, human factors, and user experience of connected health systems: a three-phase methodology. JMIR Hum Factors 2017 Mar 16;4(1):e8. [doi: 10.2196/humanfactors.5443] [Medline: 28302594]
25. Michie S, Richardson M, Johnston M, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
26. Ramo DE, Thrul J, Delucchi KL, et al. A randomized controlled evaluation of the tobacco status project, a Facebook intervention for young adults. Addiction 2018 May 24;doi. [doi: 10.1111/add.14245] [Medline: 29797621]
27. Künzler F, Mishra V, Kramer JN, Kotz D, Fleisch E, Kowatsch T. Exploring the state-of-receptivity for mHealth interventions. Proc ACM Interact Mob Wearable Ubiquitous Technol 2019 Dec;3(4):140. [doi: 10.1145/3369805] [Medline: 36159209]
28. Redfern J, Thiagalingam A, Jan S, et al. Development of a set of mobile phone text messages designed for prevention of recurrent cardiovascular events. Eur J Prev Cardiol 2014 Apr;21(4):492-499. [doi: 10.1177/2047487312449416] [Medline: 22605787]
29. Derksen ME, Jaspers MW, van Strijp S, Fransen MP. Mobile health for smoking cessation among disadvantaged young women during and after pregnancy: user-centered design and usability study. JMIR Form Res 2021 Aug 4;5(8):e24112. [doi: 10.2196/24112] [Medline: 34346895]
30. Vilardaga R, Rizo J, Zeng E, et al. User-centered design of learn to quit, a smoking cessation smartphone app for people with serious mental illness. JMIR Serious Games 2018 Jan 16;6(1):e2. [doi: 10.2196/games.8881] [Medline: 29339346]
31. Hartzler AL, BlueSpruce J, Catz SL, McClure JB. Prioritizing the mHealth design space: a mixed-methods analysis of smokers' perspectives. JMIR Mhealth Uhealth 2016 Aug 5;4(3):e95. [doi: 10.2196/mhealth.5742] [Medline: 27496593]
32. McClure JB, Heffner J, Hohl S, Klasnja P, Catz SL. Design considerations for mHealth programs targeting smokers not yet ready to quit: results of a sequential mixed-methods study. JMIR Mhealth Uhealth 2017 Mar 10;5(3):e31. [doi: 10.2196/mhealth.6845] [Medline: 28283465]
33. Brin M, Trujillo P, Huang MC, et al. Development and evaluation of visualizations of smoking data for integration into the Sense2Quit app for tobacco cessation. J Am Med Inform Assoc 2024 Jan 18;31(2):354-362. [doi: 10.1093/jamia/ocad162] [Medline: 37632226]
34. Tudor-Sfetea C, Rabee R, Najim M, et al. Evaluation of two mobile health apps in the context of smoking cessation: qualitative study of Cognitive Behavioral Therapy (CBT) versus non-CBT-based digital solutions. JMIR Mhealth Uhealth 2018 Apr 18;6(4):e98. [doi: 10.2196/mhealth.9405] [Medline: 29669708]
35. MacDougall C, Baum F. The devil's advocate: a strategy to avoid groupthink and stimulate discussion in focus groups. Qual Health Res 1997 Nov;7(4):532-541. [doi: 10.1177/104973239700700407]
36. Lin Y, Tudor-Sfetea C, Siddiqui S, Sherwani Y, Ahmed M, Eisingerich AB. Effective behavioral changes through a digital mHealth app: exploring the impact of hedonic well-being, psychological empowerment and inspiration. JMIR Mhealth Uhealth 2018 Jun 15;6(6):e10024. [doi: 10.2196/10024] [Medline: 29907557]

Abbreviations

e-cigarette: electronic cigarette
HCD: human-centered design
RCT: randomized controlled trial

XSL•FO
RenderX

Edited by A Parush; submitted 24.04.25; peer-reviewed by A Eisingerich, WJ He; revised version received 18.07.25; accepted 20.07.25; published 18.09.25.

Please cite as:
Tran TPT, Tran C, Ling PM, Popova L, Nguyen N
Applying Human-Centered Design to Develop Smartphone-Based Intervention Messages to Help Young Adults Quit Using E-Cigarettes and Cigarettes: A Remote User Testing Study
JMIR Hum Factors 2025;12:e76503
URL: https://humanfactors.jmir.org/2025/1/e76503
doi:10.2196/76503

© Thi Phuong Thao Tran, Christine Tran, Pamela M Ling, Lucy Popova, Nhung Nguyen. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 18.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

User Requirements and Conceptual Design for an Electronic Data Platform for Interhospital Transfer Between Acute Care Hospitals: User-Centered Design Study

Pamela Garabedian[1], MS; Jazzarae Kain[2], MPH; Srinivas Emani[2,3], MA, PhD; Stephanie Singleton[2], BA; Ronen Rozenblum[2,3], MPH, PhD; Lipika Samal[2,3], MPH, MD; Stephanie Mueller[2,3], MPH, MD

[1]Research Information Science & Computing, Mass General Brigham, 399 Revolution Drive, Somerville, MA, United States

[2]Division of General Internal Medicine, Brigham and Women's Hospital, Boston, MA, United States

[3]Harvard Medical School, Boston, MA, United States

Corresponding Author:
Pamela Garabedian, MS
Research Information Science & Computing, Mass General Brigham, 399 Revolution Drive, Somerville, MA, United States

Abstract

Background: The transfer of patients between hospitals, that is, interhospital transfer (IHT), introduces discontinuity of care, including gaps in health information transfer, which may worsen patient outcomes.

Objective: This is the first phase of a 5-year research study. Our goals are (1) to understand the gaps in health information exchange (HIE) and the clinician experience in accessing and using the electronic health record (EHR) during IHT and (2) to identify clinician user requirements for the development of an internal EHR solution for IHT.

Methods: We used prior work on HIE during IHT, coupled with a user-centered design (UCD) process to engage in discussions with clinical users and gather input on EHR workflow during IHT patient admission and planning. A total of 8 UCD sessions were held between February and July 2023, involving 18 clinicians who interact with the EHR during IHT, including 3 medicine residents, 10 advanced practice providers (APPs), and 5 direct care attendings—all responsible for caring for IHT patients at Brigham and Women's Hospital Cardiology, Medicine, Oncology, and intensive care unit services. Discussions highlighted facilitators and barriers and suggested improvements for data access and availability at the time of transfer. UCD sessions were recorded, analyzed, and coded by 2 independent reviewers to identify common themes driving suboptimal HIE. User requirements were derived from the sessions with users and iteratively refined throughout the process.

Results: Qualitative analysis revealed that a significant number of frontline clinicians experience suboptimal availability of clinical information in the EHR at the time of IHT, including gaps in communication, incomplete data, and inefficient access to clinical data. User requirements emerged from these themes and primarily focused on information prioritization, data accessibility, and workflow and efficiency.

Conclusions: Notable levels of missing information and inefficient access to clinical data were reported by end users caring for IHT patients at the time of transfer. Conducting user research to understand the current process of IHT, involving users in conceptual design and information architecture, and generating prototypes for feedback from users can aid in designing a solution that meets user needs. The results of these early UCD activities will be used to develop and implement a data platform to support clinicians during IHT.

(JMIR Hum Factors 2025;12:e67884) doi:10.2196/67884

KEYWORDS

interhospital transfer; health information exchange; user-centered design; electronic health record; electronic data; acute care; conceptual design; health information; user requirement; qualitative; data access; clinical data; design; user design; decision-making; decision support

Introduction

The transfer of patients between acute care hospitals (interhospital transfer [IHT]) occurs regularly, with one study of nationally representative Medicare data reporting over 100,000 patients undergoing IHT each year and a particularly high frequency of IHT among patients with multiple chronic conditions [1-5]. During IHT, health information exchange (HIE), which can be defined as "appropriate, timely sharing of vital patient information," is an essential part of the transfer process for appropriate care to be provided to the transferred patient [6]. Importantly, effective HIE encompasses both timely

XSL•FO
RenderX

access to data (ie, data exchange), as well as an organized presentation of clinical information (ie, data visualization) so that the receiving or accepting clinician can safely and efficiently provide care to transferred patients without unnecessary cognitive load [7].

Though the concept of HIE is easy to understand, achieving effective HIE has proven difficult due to several factors: (1) accepting clinicians who care for IHT patients after arrival are working within a busy hospital environment. Thus, any solution aimed at improving HIE during IHT must fit seamlessly into the workflow of these busy clinicians, while also not creating delays in patient transfer which could impact patient outcomes; (2) existing interoperability of different electronic health records (EHRs) between different hospital systems is often insufficient, resulting in lack of complete or adequate data exchange [8-13]. In some cases, this is due to active information blocking by developers of certified HIEs or others, but in many cases, it is due to insufficient adoption of interoperability standards [14,15]. Therefore, effective HIE must include adequate interoperability; and (3) even when HIE occurs effectively, there is often a lack of organized presentation of the data; up to 34% of clinicians report delays in their care of transferred patients due to disorganization of medical records requiring extra time to review [16,17]. Thus, improving HIE during IHT must also address the user interface.

There is a critical need for an integrated, interoperable HIE that addresses these 3 factors, along with other domains such as ease of use and usability of the information. At the same time, there is limited research on user requirements to support the development of such HIE.

This is the first phase of a 5-year study that aims to design, implement, and rigorously evaluate the implementation of an HIE platform to improve data access during IHT [18]. This paper reports on the utilization of user-centered design principles to better understand the clinician experience of HIE in the IHT process and identify key user requirements for the design of an effective HIE platform to support IHT. User requirements reflect the user needs of a system and can be used broadly to design solutions to address them.

Methods

Design, Setting, Participants, and Recruitment

This study was conducted at Brigham and Women's Hospital (BWH), Boston, Massachusetts, an 826-bed quaternary care hospital that cares for a high volume of patients who are transferred from other acute care facilities for specialized care. BWH uses Epic (Epic, MyChart, Care Everywhere; Epic Systems Corporation) as their enterprise EHR, which allows the transfer of health information between institutions using Epic through an HIE feature called Care Everywhere [19].

We recruited APPs, direct care attendings, and residents who regularly care for patients who undergo IHT to the general medical, oncology, cardiology, and intensive care unit services at BWH to participate in the design sessions. A general recruitment letter with information describing the goals of the study, human research committee approval, and study team contact information was emailed to all eligible clinicians requesting volunteers to participate in design sessions, with recurring emails sent up to 3 times until target recruitment was reached, ensuring appropriate representation among each participant role. Interested participants could sign up via a link included in the email and were contacted by the study team to coordinate available timing for the design session.

Ethical Considerations

The study was approved by the Mass General Brigham Institutional Review Board for Human Research (2022P001284). Participants were informed that their participation was voluntary, would not affect their employment, and that all collected data would remain confidential, with results presented only in the aggregate. Participants gave verbal consent to participate in the design sessions. At the completion of the session, participants were given a US $50 gift card for their time.

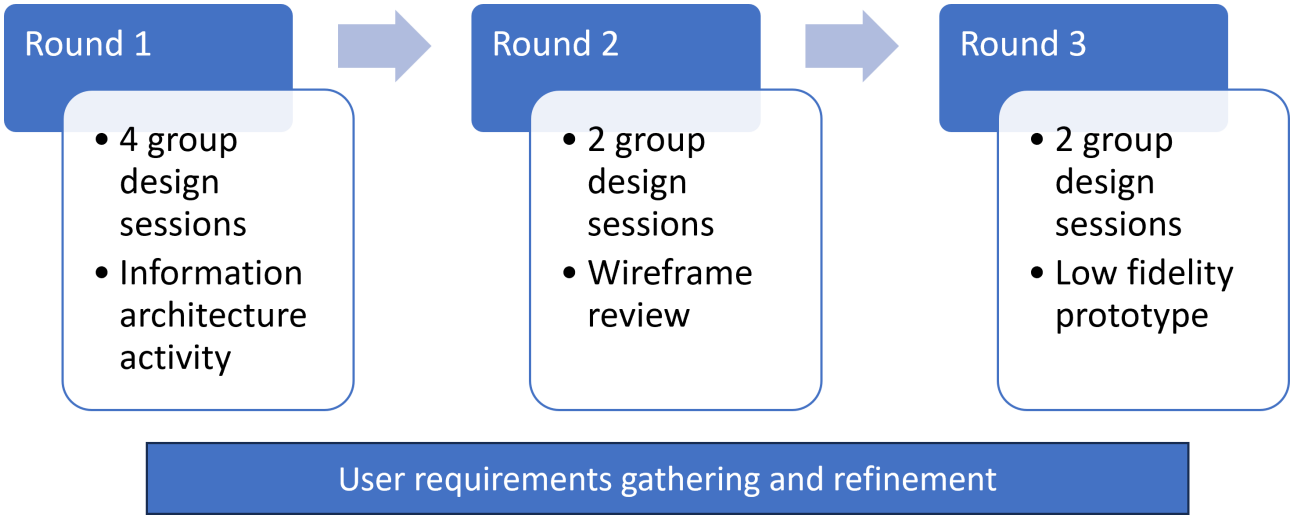
Design Sessions

All sessions were conducted virtually over Teams (Microsoft Corp). Sessions were led by a user experience specialist (PG) and assisted by a research team member with content expertise (SM).

Following a user-centered design process (Figure 1), we conducted the first round of 4 group design sessions to understand existing processes around HIE during IHT, using a semistructured group discussion guide (Multimedia Appendix 1). The interview guide was developed by the user experience specialist (PG) and qualitative research expert (RR) with input from the content expert (SM) and additional research team members and piloted prior to participant sessions. The discussion guide covered topics including prompts for sharing the current state of workflow during IHT and interaction with the EHR, with a particular focus on accessing clinical data from hospitals in different systems and with different EHRs, as well as a discussion of barriers and facilitators with the current systems of data access during IHT. During this part of the discussion, we included current screenshots of Epic using redacted real-life patient transfer examples to understand their data-gathering activities more explicitly. Design sessions also gathered participant input on essential clinical data elements available during IHT and included an interactive information architecture activity using previously identified key data elements during IHT (Multimedia Appendix 2), where participants were prompted to group, add, and organize data elements according to their preferences for optimal data review when caring for a transferred patient [20].

XSL•FO
RenderX

Figure 1. User research process.



Following this first round of design sessions, we performed rapid analyses to generate a preliminary list of user requirements to be refined and discussed during later rounds of design sessions. We drafted an initial wireframe based on these user requirements to be used during our next round of design sessions. The second round of 2 design sessions focused on iteratively refining the draft wireframe and preliminary list of user requirements, with directed input from participants on what they liked and disliked, what was missing, and how they would use, reorganize, group, and manipulate the data. We again performed rapid analysis to develop the draft wireframes into low-fidelity prototypes of an IHT data platform within Epic, developed by Mass General Brigham Digital, which we then presented to end users in the third round of 2 design sessions. In this last stage, we gathered iterative end user feedback to optimize data elements and presentation prior to further development and testing of the IHT data platform within the EHR.

Analysis

We used a modified framework and rapid analysis approach to analyze the focus group data starting after the first round of design sessions and continuing with each additional session. Sessions were transcribed using the automated feature of Teams, reviewed for accuracy, and then reviewed independently by 2 study team members (JK and PG) to develop a preliminary set of codes. The codebook was refined during multiple meetings among the 2 coders and a third research team member to provide content expertise (SM) until a final codebook was developed. Data were input into MAXQDA qualitative analysis software (VERBI GmbH), and the final analytical framework was applied to participant comments. To further analyze the data, a matrix was created to identify larger themes and subthemes among the data. Data from the information architecture activity was organized using Excel (Microsoft Corp) to identify how often data elements were sorted into any given category.

User requirements emerged from the participant comments, including the information architecture activity. The user requirements reflect the user's needs, expectations for the system, and what is required to support the user's goals and tasks. Although we identified solutions to these requirements within our Epic environment, the user needs and expectations are applicable regardless of the platform or technology. The requirements were iterated on throughout the design sessions using a rapid analysis approach and summarized as refinements to the user requirements.

Results

Participant Characteristics

Of the 34 clinicians who expressed interest in participating in one of our design sessions during this period of the study, 18 clinicians were selected. We conducted 8 design sessions between February and November 2023, involving 3 medicine residents, 10 APPs, and 5 direct care attendings. Most participants were female (n=14, 78%), general internal medicine clinicians (n=13, 72%), and APPs (n=10, 63%; Table 1).

XSL•FO
RenderX

Table . Group design session participant demographics.

Characteristics[a]	Participants, n (%)
Age (years)	
<30	3 (17)
30 - 35	5 (28)
36 - 44	7 (39)
45 - 54	3 (17)
Sex	
Female	14 (78)
Male	4 (22)
Race	
Asian	3 (17)
Black or African American	1(6)
White	14 (78)
Ethnicity	
Hispanic	1 (6)
Non-Hispanic	17 (94)
Role	
Advanced practice provider	10 (64)
Direct care attending	5 (23)
Resident	3 (14)
Department	
Cardiology	2 (11)
General or internal medicine	13 (72)
Oncology	2 (11)
Intensive care unit	1 (6)
Years of experience in role	
0 - 4	10 (56)
5 - 9	5 (28)
10 - 15	1 (6)
Over 15	2 (11)
Frequency of IHT[b] patient admission	
Around once per month	6 (33)
Once every 2 - 4 weeks	7 (39)
At least once per week	5 (28)
Time conducting direct care as a responding clinician (%)	
0 - 25	3 (17)
26 - 50	3 (17)
51 - 75	3 (17)
>76	9 (50)

[a]All characteristics were self-reported, collected via recruitment survey prior to participation.

[b]IHT: interhospital transfer.

XSL•FO
RenderX

Design Sessions

Summary

The analytic process resulted in multiple codes which were categorized into three main themes and subthemes that reflect the essential aspects of effective HIE during IHT from the perspective of frontline clinicians: (1) communication, (2) data availability and completeness, and (3) efficiency of access to data.

Theme 1: Communication

Overview

Clinicians tended to frame their responses about data in a way that was often tied to a desire for complementary communication to accompany the discrete patient data available to them within the EHR. Participants expressed that, in current IHT processes, communication is often lacking, and felt as though adding information on the timing of patient arrival or a "warm handoff" (verbal communication to help synthesize the clinical information) would add valuable contextual information to improve their ability to care for IHT patients on arrival. Communication was categorized into 2 subthemes: written and verbal.

Written Communication

In addition to specific data elements (as described below), participants also identified a requirement for additional communication related to aspects of the transfer process that would be essential in guiding their workflow. These included transfer process characteristics such as status or timeline of the transfer (eg, final decision to accept and anticipated time of arrival), any other consultants or care team members that were involved in the transfer process (eg, transferring clinician contact information, accepting physician, and consultants involved prior to transfer), and clinical aspects of care that required follow up (eg, pending tests or reports).

Verbal Communication

One of the challenges mentioned by accepting clinicians was a lack of verbal communication or "warm handoff." Participants relayed that with select, complicated patient transfers, the opportunity to have a "warm handoff" was helpful to synthesize the clinical data available to them within the EHR, to help them optimize care for IHT patients. As such, participants identified that the availability of accurate contact information for the transferring physician, as well as the accepting physician, would be beneficial, such that they could then directly contact these individuals to help clarify information as needed. One participant shared, "Patient information can be lost or forgotten due to unclear communication ... if communication for pass off is not addressed, patients can be harmed."

Notably, there were some overlaps between these 2 subthemes. For example, in many circumstances, participants expressed a need for clearer communication without distinction as to whether they prefer that this occurs via written or verbal communication, particularly related to the contextual aspects of the transfer (eg, timing of arrival).

Theme 2: Data Availability and Completeness

Overview

Participants additionally shared their challenges in being able to access complete data about a patient from the transferring hospital. Data completeness was defined as access to essential clinical patient information to effectively care for the transferred patient on arrival, and was categorized into 6 subthemes: brief summary, imaging, laboratory data, notes, medications, and contact information.

Brief Summary

Commonly mentioned as one of the most important data elements that was commonly unavailable to frontline clinicians was a "brief summary" of the clinical scenario. Clinicians described the need for access to a "brief summary" or "snapshot" of the patient, including a "one-liner" on the patient's active medical issues and pertinent past medical history, the reason for transfer and the patient's immediate needs, and the patient's acuity level or illness severity. Participants relayed the need for this summary for all transferred patients and endorsed that this information was especially critical when patients were unable to provide history on their own. Importantly, participants also indicated the need for real-time updates to this summary if the clinical status of the patient were to evolve while awaiting transfer.

Radiology Imaging

In addition, one of the most challenging data completeness issues expressed by the participants was the availability of radiology images (eg, the computed tomography scan or chest x-ray images) rather than solely the interpretive reports of the imaging studies completed by the radiologist at the transferring hospital. Participants relayed that EHR system incompatibility, specifically the part of the EHR system that manages radiology images, contributes to the lack of access.

Laboratory Data

Participants requested access to specific laboratory data depending on the clinical condition of the patient (eg, cardiac enzymes for patients with cardiac symptoms), as well as information about laboratory tests that were pending at the time of transfer (eg, sent by the transferring hospital but not yet resulted), which was often unavailable at the time of transfer. One participant stated, "…we want to be able to have previous data for what's happening, is patient getting better or worse."

Clinical Notes

Participants also relayed that incomplete availability of clinical notes felt to be important in caring for transferred patients, including notes related to prior admissions (eg, discharge summaries), prior office visits, and notes related to the patient's current admission at the transferring hospital (eg, their initial history and physical, or "H & P"). One participant stated their need for "the most recent pertinent data from the hospitalization, if they do not get most of the care from our hospital, we want a bigger time frame of information ... is their health problem new? We need back data or a link out to prior data. That would be helpful."

XSL•FO
RenderX

Medications

Participants relayed challenges with the availability of an accurate medication list at the time of transfer. Importantly, they expressed the desire for access to several different aspects of the patient's medication list, including an accurate list of medications the patient was taking prior to admission to the transferring hospital to provide historical context, an accurate list of medications the patient was taking at the time of transfer, and particularly a desire for access to the medication administration record, which is a documentation of the timing of medication administration. This latter list was felt essential to be able to make efficient decisions about medication orders when the patient arrives at the hospital.

Contact Information

Participants expressed that they infrequently have available to them contact names and numbers they need throughout the transfer process, including the contact information for the transferring clinician (ie, to be able to contact with any questions), as well as the contact information for the patient's specialists, which in many cases can help provide significant insight into the case.

Theme 3: Efficiency of Access to Data

Overview

In addition to challenges with the availability of data from transferring hospitals, participants noted that, even when available, they experienced inefficient access to it. Efficiency of access to data was categorized into 5 subthemes: dependence on manual processes, EHR organization and presentation, clinical notes, laboratory data and vital signs, and awareness of existing EHR capabilities.

Dependence on Manual Processes

Participants explained that in cases when the patient is transferred from a hospital outside the health care system, many processes are manual, such as searching through paper or electronically faxed records that are often PDF scans of paper. They also relayed that occasionally, the clinical data for the patient is on a disc or USB that requires manual uploading and sorting. This dependence on manual processes is felt to be extremely cumbersome and an inefficient way to access patient data.

EHR Organization and Presentation

Participants explained that for patients transferring from a hospital within the health care system or from a hospital with the same EHR, the "Care Everywhere" feature of Epic provides access to data but processes of distilling that information and finding relevant data points to the current patient situation is challenging and inefficient. Participants described it as "disorganized, difficult to navigate" and requiring "lots of digging." In addition, participants describe organizational issues that impact their workflow. For example, one participant shared this about imaging and imaging reports: "Putting images and encounters together would be helpful. They are usually separate from the report."

Clinical Notes

In addition to the incomplete availability of select clinical notes (described above), relevant notes such as the discharge summary and history and physical from the transferring hospital are not always easy to find. In addition, participants shared that, when available, they use the filtering feature for notes and different specialty tabs, such as the cardiology tab, to find relevant notes more quickly.

Laboratory Data and Vital Signs

In addition to incomplete access to laboratory data (described above), participants also described frequent inefficiencies in the ability to view historical laboratory data and vital signs to be able to effectively evaluate for trends, particularly if the patient had data in disparate hospital settings or EHRs. Participants expressed a desire for direct access to the most recent laboratory data and vital signs, but with easily accessible historical data organized in a coherent chronological way, even when these data are coming from disparate sources. A participant shared, "I think it's great to have it [lab data] in one spot. I think not everybody knows how to grab things from Care Everywhere..."

Awareness of Existing EHR Capabilities

Most participants noted a lack of familiarity with or use of the "access center note"—a document reflecting a brief summary of the most recent transfer hospital information entered by access center nurses at the time the patient was accepted for transfer until the time of patient arrival. Additionally, most participants relayed not knowing how to locate the "transfer module," also used by the access center nurses during the acceptance process, which includes a call log and running timeline of transfer hospital information and communication. "Two reasons why people cannot see the transfer module is either because they do not have access to see it or they don't know what it is (that the information exists) and don't bother to look for the information in notes."

Information Architecture: Organization of Data Elements

In parallel to the above analyses, we analyzed participant feedback from the information architecture activity conducted as part of the first 4 design sessions to identify the preferred organization of essential data elements, naming for the groupings, and additional data elements to include. Participants grouped patient demographic data and transferred hospital contact information together. Participants organized data elements such as the reason for transfer, needs of transferring hospital, and a patient "one-liner" into a "brief summary" or "snapshot" section. Most participants grouped clinical patient data together, including vital signs, laboratory data, microbiology data, radiology data, and other clinical data. Narrative data, like clinical notes and summaries, were grouped together, and medications often got placed in a category of their own.

User Requirements

Throughout the design sessions, we generated and refined a list of user requirements. The user requirements represent how the

XSL•FO
RenderX

system needs to support the work of the end user and were identified from the themes and subthemes outlined above, as well as the results of the information architecture activity. In total, we generated a list of 17 user requirements categorized into information prioritization, data accessibility, and workflow and efficiency:

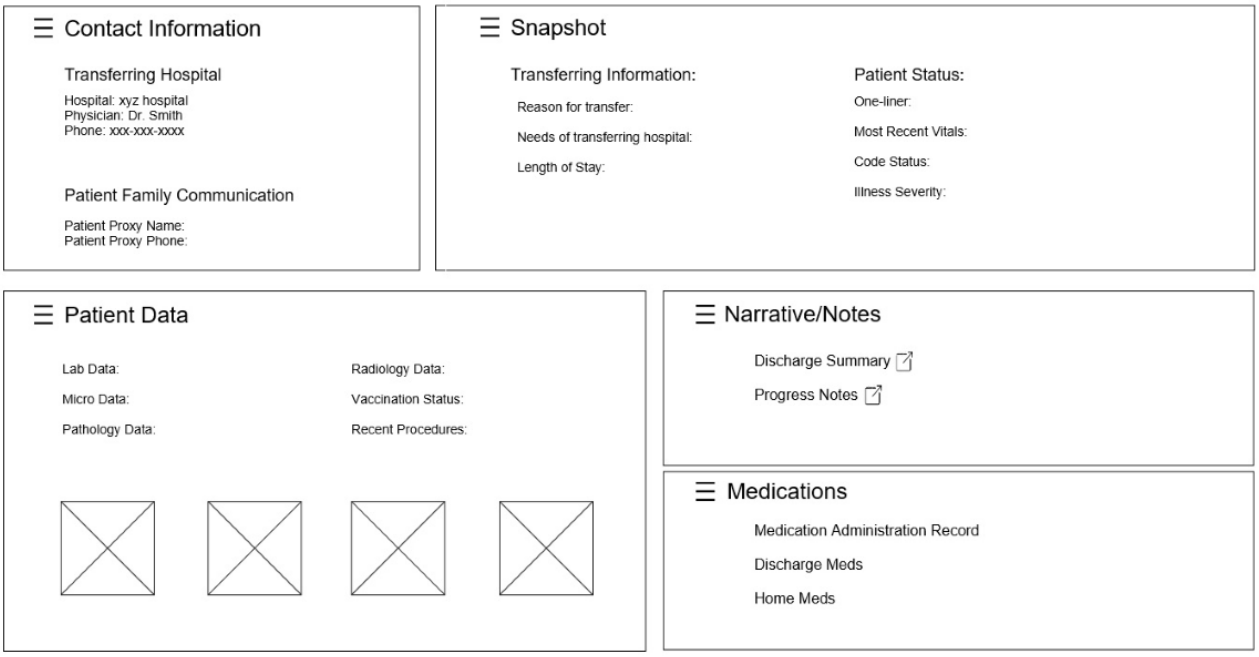
Information prioritization includes requirements that reflect the user preferences for the order in which they view data (ie, user needs to see select patient data in line-of-sight vs via click-to-expand, eg, second or third priority).

Data accessibility requirements cover the users' needs for relevant data elements and additional clinical information that is not currently readily accessible to them. For example, users describe a desire to access the various data elements within "medications," including access to the patient's home medications (prior to hospitalization at the transferring hospital), medications at the time of transfer, and the medication administration record, which includes details of the timing of most recent medication administration.

Workflow and efficiency include requirements that address the user needs for solutions that work within their current workflow, minimize steps, and provide easier ways to accomplish their task. (ie, user wants flexible displays of data to be able to view the most critical and choose to access more detail).

For each identified user requirement within these 3 categories, we brainstormed potential solutions to address the requirement and iteratively refined each requirement and proposed solution via additional design sessions. Wireframes were developed based on these user requirements and used during subsequent group sessions to elicit feedback and validate requirements (Figure 2). The wireframes included sections for demographic data and transferring hospital contact information; a collated patient "snapshot" or "brief summary" including a "one-liner" and reason for transfer; and sections of clinical data grouped into classical categories including vital signs, laboratory data, microbiology data, radiology data, other clinical data, clinical notes, and medications. Participants from later design sessions shared more specific feedback on how they prefer to view and organize certain data based on refinement from earlier sessions. More details on our user requirements and proposed solutions can be found in Table 2.

Figure 2. Wireframe of the electronic interhospital data platform.



XSL•FO
RenderX

Table . User requirements and proposed design features.

User requirement	Proposed solutions
Information prioritization	
System places contact information as high priority	• Suggested data elements: number for transferring floor, name or contact of transferring physician, name of accepting hospital and other clinicians involved, contact information for health care proxy. • Contact information section in the top left corner of the summary page.
System places critical clinical information as high priority	• Suggested data elements: a brief summary of the patient's clinical status, reason for transfer, anticipated patient needs in the first 24 hours, length of stay at the transferring hospital, code status, and level of illness severity. • "Snapshot" section in the top right corner of the summary page.
System places clinical patient data as mixed high and low priority	• Suggested data elements: vitals data, laboratory data, radiology data, microbiology data, pathology data, vaccination status including COVID-19, and recent procedures. • Data section in the left center of the summary page. • Combination of preview and controls that expand or link out.
System places narrative of clinical course at transferring hospital as mixed higher and lower priority	• Suggested data elements: progress notes, discharge or transfer summery, prior discharge summaries, consult notes, history and physical note. • Notes section on the right side of the summary page, below the snapshot. • Combination of preview and controls that expand or link out.
Data accessibility	
System supports user in efficiently accessing contact information for the patients care team	• Provide contact information within line of sight.
System allows easy access to home medications, medications at time of transfer, and medication administration record	• Active medication list shown before home medication list. • Medications listed in alphabetical order. • Access to medication administration record via a link or collapsible content (to reduce overcrowding on the summary page).
System includes a way to view trends in vitals, laboratory data, and historical out of range lab data	• Display last 3 values or days. • Search function for specific laboratory data. • Present an option to view laboratory data in tabular format.
System supports access to current and prior admission notes, progress notes, recent PCP[a] notes, specialist consult notes, and discharge summary	• "Notes" displays patient clinical notes in reverse chronological order. • Links to admission notes, progress notes, recent PCP notes, specialist consult notes, and discharge summary.
System allows user to access all prior patient records	• Display links to all available outside hospital records.
System allows user to access pending data at time of transfer	• Links to pending data included within pertinent sections (eg, laboratory data and microbiology data).
Workflow and efficiency	
System supports user in accessing transferred patient data in one location and in one logical timeline to avoid searching through outside records and faxes	• Generate a Patient Summary Page with all outside records pulled into structured sections. • List data in reverse chronological order. • Provide links to specific data points or outside records for additional information.
System provides flexible displays of data so users can view most critical and choose to access more detail	• Provide progressive disclosure for lengthy content. • Combination of preview and controls that expand or link out
System provides user with access to the status of a transfer as it evolves during the transfer process	• Provide access to the patient timeline documentation as documented by the accepting hospital (generated at time of acceptance for transfer).

XSL•FO
RenderX

User requirement	Proposed solutions
System provides user with access to radiology images (not just reports)	• Provide access to the radiology image and report within same location. • Organize this within the "radiology data" section.
System allows user to choose their timeline for select data presentation	• Provide link out to additional historical data.
System allows user to view Patient Summary Page while viewing other windows with detailed information to reduce back and forth between tabs	• Design Patient Summary Page as an embedded tab within the electronic health record. • Option to open other parts of the patient chart on the side panel while viewing the Patient Summary Page. • Option to "pin" page to workbench for easier access.
System allows user to pull data from the Patient Summary Page into their current patient documentation	• Provide functionality to pull data from the Patient Summary Page to patient documentation.

[a]PCP: primary care provider.

Discussion

Principal Findings

We identified 3 major themes that are important to address in the design and development of an IHT summary: communication, data completeness, and efficiency of access to data. The user requirements generated from the design sessions were centered around these themes and focused on prioritizing critical information, easily accessing and viewing data from outside sources in one place, and organizing the data within the existing workflow to maximize efficiency. This study highlights the need for one location within the EHR where clinicians can access all the relevant and available data from a transferring hospital as it relates to the care of the transferred patient.

Most prior studies of HIE have focused on the technical requirements for the exchange of data, such as addressing the fragmentation caused by a lack of standards and the governance of EHR vendors [21,22]. Within this framework, several studies have targeted select aspects of HIE; for example, several studies targeted the transfer of medications between systems. One such study relied on qualitative semistructured interviews of key stakeholders to assess challenges and opportunities for more effective information exchange to achieve more accurate medication lists. This study identified similar challenges to our findings, including incomplete patient records and gaps in the availability of digital data, along with barriers related to lack of interoperability and poor IT and change management [23]. Another study reported the use of a design science research methodology, including assessing stakeholder requirements, to develop an interoperable open standards solution for the transfer of medication information between acute and postacute care facilities using 2 different EHRs [24]. Although the transfer of medication information is an essential component of HIE, it does not encompass the entirety of clinical information that should be included with effective HIE. A third study, which also used qualitative semistructured interviews of key stakeholders to address HIE more broadly, identified 3 themes related to user requirements for interoperability of HIE between systems: technological (eg, interoperability challenges), organizational (facilitators of interoperability), and environmental (policy domain) [21]. Other studies have identified similar themes, such as technological and organizational barriers to interoperability [22,25-27]. This study contributes to this body of literature, importantly by highlighting user needs beyond interoperability, particularly regarding the presentation and usability of the data once integrated into the EHR, and the ability to act upon the data within clinical workflow. Our rigorous methodology, using tools such as prototyping and participatory design, ensured that we adequately represented the clinician's perspective and defined the user requirements [28]. This distinction is essential to ensuring that HIE is effective in allowing data to be easily accessible and understandable by the end user.

While addressing the technical aspects of data availability is critical, our findings suggest that improving clinician experiences during IHT requires more than back-end technical solutions alone. This study found that access to critical data remains a challenge for providers, particularly when it comes to imaging, medication records, and laboratory values. In addition, this study highlights the burden on the clinician when the data is available but difficult to find or digest. Many providers are experiencing significant effort required to "click around" and find the data they need, which has been shown in other studies to contribute to clinician burnout [29]. Even within existing HIE tools that allow access to external data (eg, CareEverywhere within Epic), this study highlights that clinicians still struggle to navigate the information and quickly retrieve essential details needed to care for the patient in front of them. Although EHR vendors are continually improving and working to address these issues, our findings show that there are still significant challenges in distributing these solutions to end users [30]. This may lead to delays in patient care and increased effort for clinicians, negatively impacting the transfer process. This is supported by prior work that demonstrates a high frequency of suboptimal HIE, with up to 35% of patient transfers missing essential clinical data on transfer [16,17,31]. Suboptimal HIE during IHT has also been associated with patient harm, such as therapeutic errors and delays in care, and it may contribute to greater mortality observed among IHT patients with select diagnoses, in addition to other outcomes such as greater resource utilization [4,9,17,32-35].

One significant subtheme that we identified within the data availability and completeness themes was the mention of the "brief summary" or "snapshot" of the patient transfer. This data

XSL•FO
RenderX

request was unique in that it is an unstructured piece of information, often evolving and dynamic; therefore, unlike more discrete pieces of data such as lab values or medications, there are challenges in identifying the source and responsibility for this information and making it available to clinicians in an accurate and useful presentation. With advances in existing technologies such as artificial intelligence and natural language processing [36], our findings highlight the potential to leverage these newer modalities to assist in this identified need to synthesize clinical information and present a "brief summary" to the clinician. For example, natural language processing algorithms can be used to analyze unstructured text and parse the data to create a structured patient summary. In addition, artificial intelligence algorithms can pull data from disparate sources and consolidate it into a unified summary for a full picture of a patient's relevant history.

Limitations

Our study has several limitations. We gathered requirements from clinicians working at one health care institution using Epic, therefore, the results may not be generalizable to other settings and EHRs. While the user requirements can be applied broadly, the information exchange may vary depending on the health system, so the successful application of these findings may vary depending on the environment. Indeed, any solution that meets these identified user requirements must also comply with data privacy and security regulations, which may restrict the extent of successful information exchange. Future research will include additional sites and EHR systems to understand whether there are additional needs across different health care environments and geographical regions. Furthermore, we restricted the qualitative sessions to direct care attending physicians, residents, and APPs on the medicine, oncology, cardiology, and intensive care units. Thus, the providers who participated may not be representative of all providers. However, we wanted to focus on those who had been most relevant to this study in order to get more detailed feedback about our potential intervention. In exploring solutions to address the user requirements, constraints of the Epic system and compatibility issues surfaced as limitations in addressing the user needs fully. Additional iterative refinement and usability testing will be essential during the development of an HIE platform to ensure its usability and relevance in improving the IHT process, as technology advances and health care culture and workflows evolve.

Conclusions

Optimal HIE has the potential to improve the safety of IHT and to meet the workflow needs of the clinicians caring for transferred patients. Current EHR interoperability and disorganized, insufficient information make these conditions difficult and can cause delays in patient planning and treatment. We have identified key themes within existing HIE and end user requirements for effective HIE, incorporating both data needs and interface design of key data elements. The next steps include using these key findings to develop an organized IHT HIE platform within the EHR, capable of gathering the most pertinent transfer patient information quickly and securely, with the goal of both bettering patient outcomes and reducing provider burden.

Acknowledgments

We would like to acknowledge the Mass General Brigham Digital eCare Development Team for their contribution to the development of the prototypes. We would also like to acknowledge the clinicians at Brigham and Women's Hospital in the Cardiology, Medicine, Oncology, and Intensive Care services for their support and participation in our design process.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

RR reports holding equity in Hospitech Respiration, which makes Airway Management Solutions. RR also receives equity from TRI-O, which makes a medical device for diabetic foot ulcers; AEYE Health, which provides an automated artificial intelligence–based diagnostic screening solution for retinal imaging; RxE2, which integrates the practice of pharmacy into clinical trials; and OtheReality, which provides a virtual reality technology to boost empathy in health care, all of which is unrelated to this work. RR is also receiving research funding from Telem, Calosense Health, Breath of Health, and BriefCam, all of which are unrelated to this work.

Multimedia Appendix 1
User research session semistructured discussion guide.
[DOCX File, 26 KB - humanfactors_v12i1e67884_app1.docx]

Multimedia Appendix 2
Key data elements for interhospital transfer.
[DOCX File, 16 KB - humanfactors_v12i1e67884_app2.docx]

References

XSL•FO
RenderX

1. Mueller SK, Zheng J, Orav EJ, Schnipper JL. Rates, predictors and variability of interhospital transfers: a national evaluation. J Hosp Med 2017 Jun;12(6):435-442. [doi: 10.12788/jhm.2747] [Medline: 28574533]
2. Iwashyna TJ, Kahn JM, Hayward RA, Nallamothu BK. Interhospital transfers among Medicare beneficiaries admitted for acute myocardial infarction at nonrevascularization hospitals. Circ Cardiovasc Qual Outcomes 2010 Sep;3(5):468-475. [doi: 10.1161/CIRCOUTCOMES.110.957993] [Medline: 20682917]
3. Roe MT, Chen AY, Delong ER, et al. Patterns of transfer for patients with non-ST-segment elevation acute coronary syndrome from community to tertiary care hospitals. Am Heart J 2008 Jul;156(1):185-192. [doi: 10.1016/j.ahj.2008.01.033] [Medline: 18585515]
4. Sokol-Hessner L, White AA, Davis KF, Herzig SJ, Hohmann SF. Interhospital transfer patients discharged by academic hospitalists and general internists: characteristics and outcomes. J Hosp Med 2016 Apr;11(4):245-250. [doi: 10.1002/jhm.2515] [Medline: 26588825]
5. Iwashyna TJ. The incomplete infrastructure for interhospital patient transfer. Crit Care Med 2012 Aug;40(8):2470-2478. [doi: 10.1097/CCM.0b013e318254516f] [Medline: 22809914]
6. What is HIE? ASTP. 2020. URL: https://www.healthit.gov/topic/health-it-and-health-information-exchange-basics/what-hie [accessed 2024-10-01]
7. Harry EM, Shin GH, Neville BA, et al. Using cognitive load theory to improve posthospitalization follow-up visits. Appl Clin Inform 2019 Aug;10(4):610-614. [doi: 10.1055/s-0039-1694748] [Medline: 31434160]
8. Britton MC, Hodshon B, Chaudhry SI. Implementing a warm handoff between hospital and skilled nursing facility clinicians. J Patient Saf 2019 Sep;15(3):198-204. [doi: 10.1097/PTS.0000000000000529] [Medline: 30095538]
9. Knudson J. Streamlining patient transfers. ACP Hospitalist. 2014. URL: https://acphospitalist.acponline.org/archives/2014/10/transfer.htm [accessed 2025-05-26]
10. Rome BN, Schnipper JL, Maviglia SM, Mueller SK. Effect of shared electronic health records on duplicate imaging after hospital transfer. J Gen Intern Med 2020 May;35(5):1617-1619. [doi: 10.1007/s11606-019-05355-2] [Medline: 31650403]
11. Usher M, Rhodes M, Baum K, Joseph A, Melton-Meaux G, Weinert C. Harnessing electronic health record inter-operability to improve inter-hospital transfer outcomes: a pilot study. Presented at: Society of General Internal Medicine Annual Meeting; Apr 11-14, 2018; Denver, CO.
12. Everson J, Adler-Milstein J. Gaps in health information exchange between hospitals that treat many shared patients. J Am Med Inform Assoc 2018 Sep 1;25(9):1114-1121. [doi: 10.1093/jamia/ocy089] [Medline: 30010887]
13. Thompson MP, Graetz I. Hospital adoption of interoperability functions. Healthcare (Amst) 2019 Sep;7(3):100347. [doi: 10.1016/j.hjdsi.2018.12.001] [Medline: 30595558]
14. Information blocking. ASTP. 2024. URL: https://www.healthit.gov/topic/information-blocking [accessed 2024-10-01]
15. Samal L, Dykes PC, Greenberg JO, et al. Care coordination gaps due to lack of interoperability in the United States: a qualitative study and literature review. BMC Health Serv Res 2016 Apr 22;16:143. [doi: 10.1186/s12913-016-1373-y] [Medline: 27106509]
16. Mueller SK, Shannon E, Dalal A, Schnipper JL, Dykes P. Patient and physician experience with interhospital transfer: a qualitative study. J Patient Saf 2021 Dec 1;17(8):e752-e757. [doi: 10.1097/PTS.0000000000000501] [Medline: 29901654]
17. Mueller SK, Muralidhar R, Schnipper JL. Sub-optimal communication during inter-hospital transfer (IHT). Presented at: Society of Hospital Medicine Converge Conference; May 3-7, 2021.
18. Mueller SK, Garabedian P, Goralnick E, Bates DW, Samal L. Advancing health information during interhospital transfer: an interrupted time series. J Hosp Med 2023 Dec;18(12):1063-1071. [doi: 10.1002/jhm.13221] [Medline: 37846028]
19. Care everywhere 101: Epic Systems care everywhere training companion. Epic. 2024. URL: https://galaxy.epic.com/?#Browse/page=8400!35!240!3729927 [accessed 2024-10-01]
20. Mueller S, Murray M, Goralnick E, et al. Implementation of a standardised accept note to improve communication during inter-hospital transfer: a prospective cohort study. BMJ Open Qual 2023 Oct;12(4):e002518. [doi: 10.1136/bmjoq-2023-002518] [Medline: 37899076]
21. Walker DM, Tarver WL, Jonnalagadda P, Ranbom L, Ford EW, Rahurkar S. Perspectives on challenges and opportunities for interoperability: findings from key informant interviews with stakeholders in Ohio. JMIR Med Inform 2023 Feb 24;11:e43848. [doi: 10.2196/43848] [Medline: 36826979]
22. Vest JR, Gamm LD. Health information exchange: persistent challenges and new strategies. J Am Med Inform Assoc 2010;17(3):288-294. [doi: 10.1136/jamia.2010.003673] [Medline: 20442146]
23. Tolley C, Seymour H, Watson N, Nazar H, Heed J, Belshaw D. Barriers and opportunities for the use of digital tools in medicines optimization across the interfaces of care: stakeholder interviews in the United Kingdom. JMIR Med Inform 2023 Mar 10;11:e42458. [doi: 10.2196/42458] [Medline: 36897631]
24. Gottumukkala M. Design, development, and evaluation of an automated solution for electronic information exchange between acute and long-term postacute care facilities: design science research. JMIR Form Res 2023 Feb 17;7:e43758. [doi: 10.2196/43758] [Medline: 36800213]
25. Li Z, Merrell MA, Eberth JM, Wu D, Hung P. Successes and barriers of health information exchange participation across hospitals in South Carolina from 2014 to 2020: longitudinal observational study. JMIR Med Inform 2023 Sep 28;11:e40959. [doi: 10.2196/40959] [Medline: 37768730]

XSL•FO
RenderX

26. Sarkar IN. Transforming health data to actionable information: recent progress and future opportunities in health information exchange. Yearb Med Inform 2022 Aug;31(1):203-214. [doi: 10.1055/s-0042-1742519] [Medline: 36463879]
27. Dobrow MJ, Bytautas JP, Tharmalingam S, Hagens S. Interoperable electronic health records and health information exchanges: systematic review. JMIR Med Inform 2019 Jun 6;7(2):e12607. [doi: 10.2196/12607] [Medline: 31172961]
28. Arnowitz J, Arent M, Berger N. Effective Prototyping for Software Makers 2007: Elsevier, Inc; 2007.
29. Johnson KB, Neuss MJ, Detmer DE. Electronic health records and clinician burnout: a story of three eras. J Am Med Inform Assoc 2021 Apr 23;28(5):967-973. [doi: 10.1093/jamia/ocaa274] [Medline: 33367815]
30. Epic interoperability creates one virtual system worldwide. Epic Systems Corporation. 2018. URL: https://www.epic.com/epic/post/epic-interoperability-one-virtual-system-worldwide [accessed 2024-10-01]
31. Mueller SK, Schnipper JL. Physician perspectives on interhospital transfers. J Patient Saf 2019 Jun;15(2):86-89. [doi: 10.1097/PTS.0000000000000312] [Medline: 27811599]
32. Mueller S. Triaging inter-hospital transfers. : WebM&M Cases & Commentaries; 2019.
33. Mueller S, Zheng J, Orav EJ, Schnipper JL. Inter-hospital transfer and patient outcomes: a retrospective cohort study. BMJ Qual Saf 2019 Nov;28(11):e1. [doi: 10.1136/bmjqs-2018-008087] [Medline: 30257883]
34. Mueller SK, Fiskio J, Schnipper J. Interhospital transfer: transfer processes and patient outcomes. J Hosp Med 2019 Aug 1;14(8):486-491. [doi: 10.12788/jhm.3192] [Medline: 30986189]
35. Lawrence RH, Tomolo AM, Garlisi AP, Aron DC. Conceptualizing handover strategies at change of shift in the emergency department: a grounded theory study. BMC Health Serv Res 2008 Dec 16;8:256. [doi: 10.1186/1472-6963-8-256] [Medline: 19087251]
36. Alowais SA, Alghamdi SS, Alsuhebany N, et al. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. BMC Med Educ 2023 Sep 22;23(1):689. [doi: 10.1186/s12909-023-04698-z] [Medline: 37740191]

Abbreviations

APP: advanced practice provider
BWH: Brigham and Women's Hospital
EHR: electronic health record
HIE: health information exchange
IHT: interhospital transfer

Edited by A Parush; submitted 23.10.24; peer-reviewed by J Walsh, T Akindahunsi; revised version received 17.04.25; accepted 21.04.25; published 30.05.25.

Please cite as:
Garabedian P, Kain J, Emani S, Singleton S, Rozenblum R, Samal L, Mueller S
User Requirements and Conceptual Design for an Electronic Data Platform for Interhospital Transfer Between Acute Care Hospitals: User-Centered Design Study
JMIR Hum Factors 2025;12:e67884
URL: https://humanfactors.jmir.org/2025/1/e67884
doi:10.2196/67884

© Pamela Garabedian, Jazzarae Kain, Srinivas Emani, Stephanie Singleton, Ronen Rozenblum, Lipika Samal, Stephanie Mueller. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 30.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Growth Mindset Intervention's Impact on Positive Response to eHealth for Older Adults With Chronic Disease: Randomized Controlled Trial

Meijuan Cao[1,2], PhD; Xiaojuan Xu[1], BSc; Yaling Zeng[1], BSc; Binyu Zhao[2], MN; Chunqi Xie[2], MN; Hailu Wu[1], BSc; Jianlin Lou[1], PhD

[1]School of Medicine and Nursing, Huzhou University, 759 Erhuan Rd, Huzhou, China

[2]School of Nursing, Hangzhou Normal University, Hangzhou, China

Corresponding Author:
Meijuan Cao, PhD
School of Medicine and Nursing, Huzhou University, 759 Erhuan Rd, Huzhou, China

Abstract

Background: Although eHealth has shown promise in managing chronic diseases, there remains a substantial digital divide among older adults. The concept of a growth mindset, based on psychological theory, offers a new direction and potential breakthrough for addressing this dilemma.

Objective: This study aims to develop and explore the feasibility and efficacy of a growth mindset intervention for older adults with chronic diseases and their positive response to eHealth.

Methods: A randomized controlled trial was conducted at the internal medicine departments of a hospital in Hangzhou, Zhejiang Province, China, from September 2021 to October 2022. A total of 77 older patients with chronic disease initially participated in the study. The mean age of the participants was 67.16 (SD 7.04) years, with 42.86% (33/77) being women and 57.14% (44/77) being men. The experimental group received an eHealth program intervention plus a growth mindset intervention over 12 weeks, with weekly sessions for the first 6 weeks and biweekly follow-up phone calls for the next 6 weeks. Each session lasted at least 25 - 45 minutes. Data were collected using a personal information form, the Implicit Theories of Intelligence Scale-6 (ITIS-6), and a questionnaire on knowledge, willingness, confidence, and practice of smart medicine (KWCP-SM). Measurements were taken at the beginning of the study (T0), immediately after the 6 weeks of training provided to the experimental group (T1), and after the 12 weeks of training for the intervention (T2). Data were analyzed using repeated-measures analysis of variance and analysis of covariance.

Results: The final sample comprised 74 participants, of which 36 were in the experimental group and 38 in the control group. After 12 weeks of intervention, the level of growth mindset was significantly higher in the intervention group (P<.05) and significant group × time interaction was observed (Wald=11.57; P<.05) between the two groups. KWCP-SM scores increased in both groups (P<.05), with more significant changes in the intervention group.

Conclusions: This study demonstrated the effectiveness of the intervention program in improving the growth mindset level of older adults with chronic diseases and bridging the “digital divide” among them. Future studies should refine this intervention, considering the characteristics and needs of this population, to create fault-tolerant and lifelong growth environments that enhance growth mindset in older adults.

Trial Registration: ClinicalTrials.gov NCT06550817; https://clinicaltrials.gov/study/NCT06550817

(JMIR Hum Factors 2025;12:e65519) doi:10.2196/65519

KEYWORDS

e-health; older adults; chronic diseases; growth mindset intervention; positive response; randomized controlled trial

Introduction

Chronic disease is one of the major causes of morbidity and mortality worldwide, which occurs predominantly among older adults [1]. According to the Chinese Bureau of Statistics, the population aged 60 years and older reached 260 million in 2020, while 75% were diagnosed with at least one chronic disease [2,3]. As a health-vulnerable group, older adults with chronic disease require long-term health care service and self-management to adapt to physiological changes, posing a tremendous burden on patients, caregivers, and the health care system [4-6], and this is where eHealth can contribute.

eHealth is characterized by information technology, automation, and intelligence, and enables patients, especially those with

XSL•FO
RenderX

chronic disease, to monitor their health parameters automatically, receive quality clinical care remotely, and achieve health management goals through interactions with eHealth systems, medical personnel, and health care institutions [7,8]. It is regarded as an efficient approach to solve the uneven distribution of limited medical resources and reduce the burden of health care expenses [7-9]. By 2020 , more than 5000 hospitals in China provided online medical services, and over 900 internet hospitals were established, covering national prefecture-level cities [10]. Studies have shown that eHealth such as eHealth apps, monitoring devices [11], wearable devices [12], and rehabilitation systems [13], can improve the outcomes of chronic disease significantly. However, patients with chronic diseases in old age, especially those with age-related stereotypes and fixed mindsets, possess a low degree of willingness and confidence in using eHealth services. Vailati Riboni and Pagnini [14] found that negative stereotypes can trigger stereotype threat, causing older adults to experience anxiety and inferiority when encountering new technologies, thereby impairing their operational abilities. They frequently struggle to comprehend overwhelming and boring information offered across eHealth services [15,16], and may be reluctant to try eHealth service due to previous unsuccessful experiences [17,18]. Consequently, the online health care resource use rate among older adults is merely 15.2% [19]. Previous studies have demonstrated that an increased growth mindset is a potential key strategy for improving positive responses to eHealth because older adults with a growth mindset can improve age-related stereotypes, increase their willingness to adopt healthy behaviors and resilience in the face of difficulties, and tend to use eHealth more [20-23].

The growth mindset theory, proposed by Dweck and Leggett [24], theorizes that individuals possess one of two types of mindsets—fixed mindset and growth mindset—regarding the malleability of abilities, intelligence, and other attributes. Individuals with fixed mindsets view the attributes as unchangeable and uncontrollable internal parameters, and often withdraw or abandon efforts when faced with difficulties. Conversely, growth mindset individuals believe the attributes are malleable and controllable, and they always view failure as an opportunity to learn and improve [25]. In addition, the literature suggests that a mindset intervention is particularly effective in changing behavior by improving participants' persistence in goal-oriented behavior [26]. Therefore, the eHealth mindset for an intervention program appeared to be a promising approach.

Many studies have shown that interventions fostering a growth mindset are beneficial to behavior change, such as improving performance in students [27-29]. mitigating weight gain resulting from dieting setbacks [30], controlling healthy eating and exercise in obese youth [31], and smoking cessation [23]. In the geriatric population, existing studies have shown that a growth mindset is significantly associated with intention to engage in health behavior [32], cognitive function [33], and successful aging [34]. Experimental studies have demonstrated its benefits in improving age-related stereotypes [35] and memory performance [36] as well. Dang and Liu [37] have provided evidence that individuals who believe human mental capacity is incremental report more positive feelings about robots and show higher willingness to interact with them.

In summary, previous studies have investigated the growth mindset in different populations and have demonstrated potential relationships between growth mindset and positive responses to eHealth. Studies have also been conducted on the growth mindset levels of older adults, the results have demonstrated that this aspect needs to be improved [20,38], and that growth mindset may be a key target in bridging the "digital divide" among older adults. However, most of these studies were cross-sectional and descriptive, making it difficult to provide direct guidance for practice. Therefore, this study was conducted to develop and investigate the effectiveness of a growth mindset intervention based on growth mindset theory. It aimed to improve the levels of growth mindset and knowledge, willingness, belief, and practice of smart medicine among older patients with chronic disease in the context of global trends in aging and chronic diseases.

Methods

Design

This study used a quasi-randomized controlled trial design conducted over 12 weeks. Before conducting the study, the researchers numbered the wards of internal medicine departments (excluding intensive care units) according to the order of bed number. The wards were then randomly divided into intervention wards and control wards using a random number generator. Patients were recruited to the control group (eHealth only) in the control ward and the intervention group (growth mindset intervention + eHealth) in the intervention unit. Participants in the control group could undergo the training after the study concluded. The study was registered on ClinicalTrials.gov (NCT06550817).

Setting and Participants

Participants (N=77) were enrolled from a hospital internal medicine department in Hangzhou, Zhejiang Province in China between September 2021 and October 2022. The inclusion criteria were: (1) participants aged 60 years and older, (2) diagnosed with ≥1 chronic disease, (3) owned at least one smartphone, (4) ability to communicate clearly in writing or orally, and (5) willingness to participate in this study. The exclusion criteria were: (1) inability to use WeChat (Tencent Holdings Limited) or phone to communicate, (2) diagnosed as having a mental illness or cognitive impairment, (3) acute severe disease or loss of self-care ability, and (4) concurrent participation in other similar training.

This study used a quasi-experimental design, and the sample size was calculated based on comparing the means of the two samples [39]. An a priori power analysis showed that 34 participants per group (total N=68), accounting for anticipated attrition, would achieve 90% power at a 5% significance level to detect the expected mean difference. Finally, 77 participants started the intervention, and 74 completed the 12-week intervention period, with an attendance rate of 96.10% (74/77).

XSL•FO
RenderX

Ethical Considerations

The ethics committee of the School of Nursing at Hangzhou Normal University has approved this study (2022015). Informed consent was obtained from all participants involved in this study. Participants were informed of the study's purpose, methods, and process, as well as their right to withdraw at any time. Only the relevant data of the participants were used for this study and all collected data were pseudonymized to protect participants' confidentiality. As compensation for study participation, participants received a ¥50 (US $6.95) supermarket voucher and a vegetable package upon study completion.

Blinding and Randomization

The single-blind principle was strictly applied in this study. Neither the study participants nor the data collectors had access to the grouping and the content of the intervention, thus mitigating subjective bias. Patients were recruited into the control group in the control ward and into the intervention group in the intervention ward. Random grouping by ward avoided cross-contamination when patients in the same ward were divided into different groups. Participants were debriefed at the end of the study regarding the full purpose and differences between groups in the study.

Intervention

The eHealth training program was delivered to participants by nursing students with a master's degree, who have received training in growth mindset knowledge, and training related to the construction and implementation of nursing training studies. Psychological counselors, professors in geriatric nursing, and clinical experts in geriatric nursing served as consultants. The purpose, duration, and main contents of the training were introduced to the patients by the faculty before the training. Patients were asked about their learning experience and whether they were confused about the content of the lectures after the training. The faculty used the video training platform to track patients' learning, and for those who did not complete the training in time, telephone communication was implemented to encourage them to improve their learning progress. This ensured that the patients mastered the content of the training and assistance was offered in solving problems.

Growth Mindset Intervention

Based on the previous literature review, this study found that previous intervention [27], deeply aligned with the growth mindset theory. The program is mature and verified by large samples in multiple regions, and hence, based on the framework of the intervention, the following 3 strategies were used for creating the intervention content: recognize growth mindset, explore growth mindset deeply, and apply growth mindset. The purpose of "recognize growth mindset" was to convince patients that they could still change their intelligence through learning, strengthening the belief that their brain can be exercised. The purpose of "explore growth mindset deeply" was to increase patients' understanding of growth mindset and promote their willingness to cultivate a growth mindset. The purpose of "apply growth mindset" was for patients to learn to apply a growth mindset and connect it to real life. In addition, some studies have noted that there are adaptation issues in the implementation of the growth mindset [40]. The physical and mental functions of older adults show an irreversible decline trend, thus the cultivation and adaptation of the growth mindset in older adults are more difficult. Therefore, this study takes the characteristics of older adults into account and adds the theme "growth mindset in life" under the guidance of ife span theory by Baltes et al [41], to further cultivate the habit of using growth mindset in practice in older patients with chronic diseases. In this manner, the above 4 strategies were identified for the study.

Research shows that the introduction of brain science, story rewriting, skill training, and other methods can enhance the subjects' identification and internalization of the growth mindset [42,43], so the research group formulated the intervention program based on this and the actual situation of older patients with chronic diseases. Table 1 shows the specific content. The intervention duration is 12 weeks. Themes 1 - 6: once a week, online intervention is implemented and it lasts about 20 minutes each time, while face-to-face intervention lasts about 30 minutes each time, and they are carried out for a total of six weeks; theme 7: face-to-face or telephone intervention is implemented once every two weeks for about 15 minutes each time, and it is carried out for a total of six weeks.

XSL•FO
RenderX

Table . Implementation form of the growth mindset intervention program.

Intervention strategy	Intervention theme	Intervention content	Intervention form
Recognize growth mindset	Is intelligence innate?	• Introduced the structure and function of the brain and neurons. • Explained the plasticity of the brain. • Emphasized the importance of learning and challenging new things. • Strengthened the belief of "exercising the brain" and "the more the brain is used, the more flexible it is."	Video lessons
	Can older adults become smart?	• Introduce the changes in brain structure and function in older adults. • List relevant scientific research to encourage patients to believe that the brain of older adults is still plastic. • Encourage patients to talk about the relationship between "the plasticity of older adults' brain" and their own intelligent medical.	Video lessons and face-to-face interaction
Explore growth mindset deeply	The 2 different mindsets	• Introduce the concept of growth mindset. • Share different mindsets in the process of learning eHealth, and guide patients to understand different behaviors, and results of the 2 mindsets. • Emphasize the benefits of the growth mindset for older adults. • Encourage patients to rewrite stories from the perspective of the growth mindset.	Video lessons, telelecture, and story rewriting
	Facing difficulties and challenges	• Encourage patients to share their experiences of overcoming difficulties and guide them to discover the importance of a growth mindset. • Make a case of the common difficulties and challenges of older adults with chronic diseases in the use of eHealth. • Ask patients to imagine how to solve the problem through efforts, strategies, seeking help, and other ways if they are the parties.	Experience sharing and case studies

XSL•FO
RenderX

Intervention strategy	Intervention theme	Intervention content	Intervention form
Apply growth mindset	The 4 steps of growth mindset	• Introduce Dweck's 4 step strategy to change the mindset: accept, observe, name, and educate. • Encourage patients to describe the experience of the emergence of a fixed mindset during eHealth use and discuss how to implement the 4 steps of the growth mindset.	Video lessons and telelecture
	Growing mindset habit cultivation plan	• Review the interventions that have been implemented. • Encourage patients to talk about the challenges related to eHealth, and guide them to think about how to use a growth mindset to face the challenges. • Complete the "growth mindset habit cue card."	Face-to-face interaction
Growth mindset in life	Growth mindset in life	• Encourage patients with good application of the growth mindset, and help patients with difficulties to solve problems.	Telephone follow-up

eHealth Program

The eHealth program intervention aimed to increase patients' knowledge of eHealth and help them master common operational skills. The intervention forms included face-to-face instruction, video teaching, training manual, and telephone follow-up. Video teaching included "Smart Phone Basics, Online Registration, Mobile Phone Payment and Report Query," "Introduction and Operation of Online Health Consultation," "Introduction and Operation of Online Prescriptions and Residents' Electronic Health Archives," "Introduction and Operation of Hospital Online Disease Management," "Methods to Obtain and Identify Reliable Health Knowledge," and "Introduction to Wearable Mobile Medical Devices." The training manual was a video supporting content, which was used with the videos. Finally, telephone follow-up was conducted to understand the learning situation and difficulties of the participants. The program lasted for 12 weeks. In the first 6 weeks, face-to-face eHealth teaching in the internal medicine department and telephone follow-up were conducted once a week, for about 10 minutes each time. In the last 6 weeks, only telephone follow-up was conducted.

Measures

Implicit Theories of Intelligence Scale-6 (ITIS-6)

The Implicit Theories of Intelligence Scale-6 (ITIS-6) [44] was used to assess participants' level of growth mindset. Participants completed the scale at baseline, 6, and 12 weeks of intervention. The scale contains 3 growth mindset items and 3 fixed mindset items. The Likert-6 scoring method was used, and a reverse scoring method was adopted for the fixed mindset dimension, with a higher score on the scale signifying a higher level of growth mindset. In this study, the Cronbach α coefficient for ITIS-6 was 0.95, indicating good internal consistency.

Questionnaire of Knowledge, Willingness, Confidence, and Practice of Smart Medicine (KWCP-SM)

According to the relevant policy documents on the promotion and implementation of eHealth, this study invited 5 experts with 10 years or more of work experience (including 2 clinical nurses, 1 community nurse, and 2 community doctors) to conduct face-to-face interviews, and developed the questionnaire of knowledge, willingness, confidence, and practice of smart medicine. The final questionnaire had the 4 dimensions of knowledge, willingness, confidence, practice, and 9 items including online registration, and mobile phone payment. There were 36 items in total. The questionnaire adopted the Likert-5 scoring method. The higher score indicated a higher level of knowledge, willingness, confidence, and practice of smart medicine. The Cronbach α coefficient for the scale was 0.97, indicating good internal consistency.

Statistical Analysis

Baseline demographic and medical characteristics within groups were compared with an independent-sample *t* test (for quantitative variables), $\chi 2$ test (qualitative variables), or nonparametric rank-sum test (qualitative variables following grade data). The Mann-Whitney *U* test was applied to assess the difference in growth mindset and KWCP-SM between the two groups because of their abnormal distributions. Within-group comparisons were performed by the Wilcoxon signed ranks test. To evaluate the between-group differences over time, a generalized estimation equation (GEE) was performed. Statistical analysis was carried out with SPSS 25.0 version (IBM). $P<.05$ was considered statistically significant.

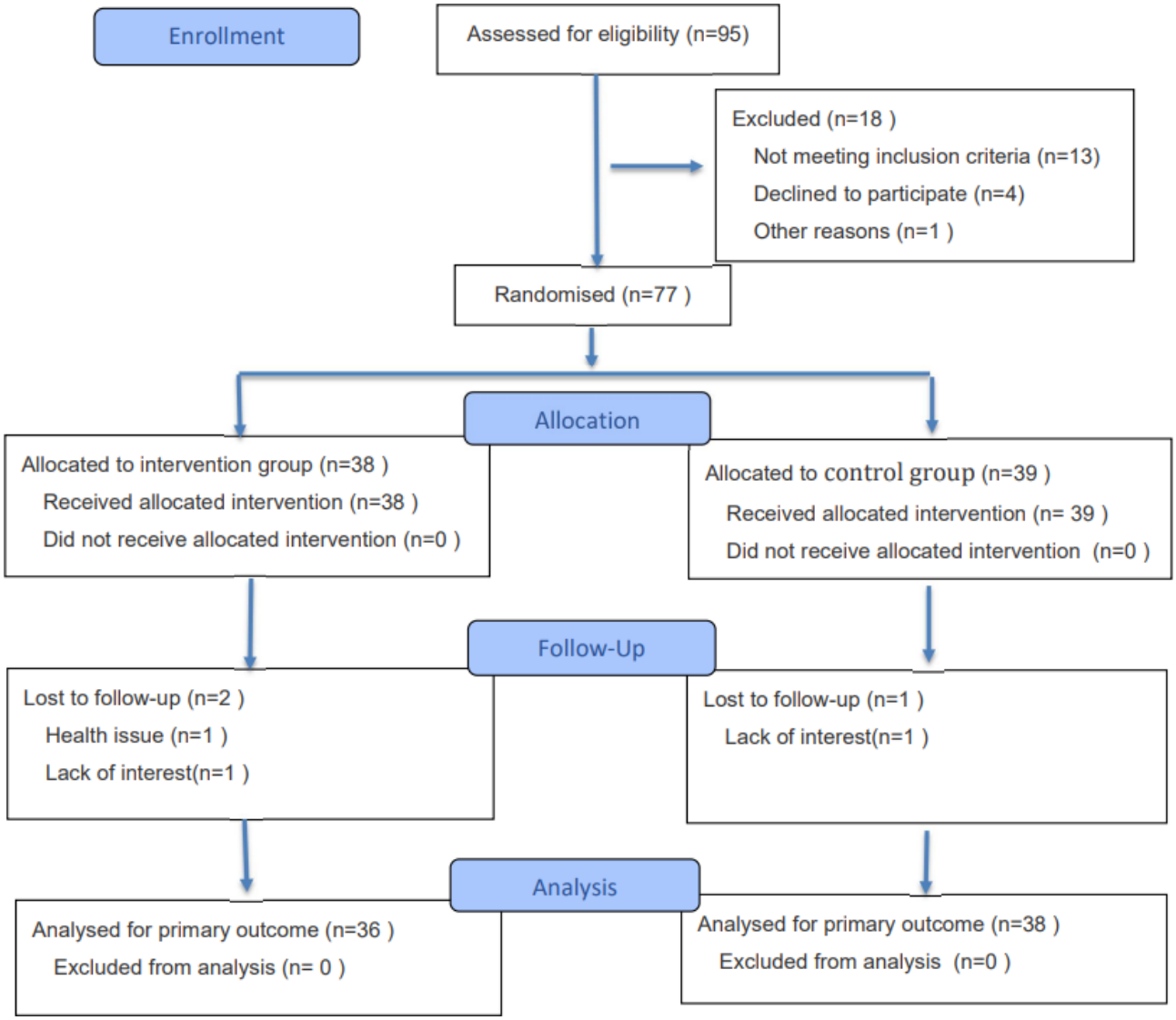
XSL•FO
RenderX

Results

Participants

Out of the 77 participants who participated in the intervention, 74 completed the 12-week intervention period, resulting in an attendance rate of 96.10% (74/77), with 36 in the intervention group and 38 in the control group. Among the 3 participants in the two groups who withdrew from the study, 2 participants in the intervention group reported health issues and lost interest, respectively, and 1 participant in the control group lost interest. Figure 1 shows the CONSORT (Consolidated Standards of Reporting Trials) diagram of the participant flow and the EHEALTH checklist is provided (see Checklist 1). Table 2 shows the baseline characteristics of participants. No significant differences were observed between the intervention and control groups regarding demographic and medical characteristics, growth mindset, and KWCP-SM of participants.

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flow chart of participation in the eHealth training program.



XSL•FO
RenderX

Table . Baseline demographics, growth mindset, and knowledge, willingness, confidence, and practice of smart medicine (KWCP-SM) of participants.

Characteristic	Intervention (n=36)	Control (n=38)	Z, *t* test, or χ^2(df)	*P* value
Demographics				
Age (years), mean (SD)	67.47 (7.44)	67.03 (6.64)	0.27[a](72)	.53
Sex, n (%)			2.60[b](1)	.16
Male	18 (50)	26 (68.42)		
Female	18 (50)	12 (31.58)		
Religious belief			2.81[b](1)	.11
Have	12 (33.33)	20 (52.63)		
Do not have	24 (66.67)	18 (47.37)		
Educational level			−0.63[c]	.53
Primary school and below	15 (41.67)	14 (36.84)		
Middle school	14 (38.89)	14 (36.84)		
High school	5 (13.89)	7 (18.42)		
Junior college and above	2 (5.56)	3 (7.89)		
Marital status			0.35[b](1)	.74
Married	32 (88.89)	32 (84.21)		
Others (including divorced, etc.)	4 (11.11)	6 (15.79)		
Place of residence			0.45[b](1)	.64
Urban	17 (47.22)	15 (39.47)		
Rural	19 (52.78)	23 (60.53)		
Occupation (previous main occupation if retired)			3.10[b](5)	.71
Public institution personnel	4 (11.11)	6 (15.79)		
Enterprise employee	10 (27.78)	13 (34.21)		
Worker	9 (25)	5 (13.16)		
Farmer	9 (25)	9 (23.68)		
Self-employed	3 (8.33)	2 (5.26)		
Others (including freelance, etc.)	1 (2.78)	3 (7.89)		
Monthly income (RMB)			−0.60[c]	.55
<1000（<US $139.10)	5 (13.89)	4 (10.53)		
1000 - 2999（US $139.10 - US $417.14)	7 (19.44)	7 (18.42)		
3000 - 4999（US $417.28 - US $695.33)	13 (36.11)	13 (34.21)		
>5000（>US $695.47)	11 (30.56)	14 (36.84)		
Living places			2.94[b](4)	.57
Living with a spouse and children	13 (36.11)	11 (28.95)		
Living with a spouse	20 (55.56)	19 (50)		
Living with children	1 (2.78)	4 (10.53)		

XSL•FO
RenderX

Characteristic	Intervention (n=36)	Control (n=38)	Z, *t* test, or χ^2(df)	*P* value
Living in institutions (including nursing home, etc)	1 (2.78)	1 (2.63)		
Living alone	1 (2.78)	3 (7.89)		
Health insurance			2.05[b](3)	.56
Urban resident basic medical insurance	17(47.22)	22(57.89)		
New rural cooperative medical scheme	16(44.44)	11(28.95)		
Commercial medical insurance	1(2.78)	2(5.26)		
No medical insurance	2(5.56)	3(7.89)		
The number of comorbidities			−0.31[c]	.76
1	12(33.33)	9(23.68)		
2	13(36.11)	19(50)		
≥3	11(30.56)	10(26.32)		
Self-rated health status (0 - 10), mean (SD)	6.32 (1.72)	5.91 (2.05)	−0.54[a](72)	.60
ITIS[d] score, mean (SD)	18.50 (8.65)	17.53 (9.09)	−0.47[a](72)	.81
KWCP-SM[e] score, mean (SD)	80.06 (22.25)	79.58 (28.05)	−0.32[f]	.75
Knowledge score, mean (SD)	23.75 (7.24)	23.47 (8.21)	−0.52[f]	.60
Willingness score, mean (SD)	23.19(7.63)	23.29 (9.46)	−0.06[f]	.95
Confidence score, mean (SD)	20.56 (8.47)	20.74 (9.68)	−0.19[f]	.85
Practice score, mean (SD)	12.56 (4.77)	12.08 (4.72)	−0.23[f]	.82

[a]Denotes the independent-samples *t* test (t).

[b]Denotes the χ^2 test (χ^2).

[c]Denotes the Wilcoxon or Mann-Whitney Z test (Z).

[d]ITIS: implicit theories of intelligence scale.

[e]KWCP-SM: knowledge, willingness, confidence, and practice of smart medicine.

[f]Denotes the Mann-Whitney *U* test (Z).

Growth Mindset

In terms of growth mindset level, there was no significant difference between the two groups at baseline (*P*>.05), but there was a significant difference between the two groups after intervention (*P*<.05). The intervention group demonstrated a significant improvement in group changes in growth mindset at 12 weeks (9.89% change from baseline; *z* score=−5.13, *P*<.05), while no statistically significant change was observed in the control group (2.68% change from baseline; *z* score=−5.13, *P*>.05). The GEE analysis revealed the growth mindset changes over time in the intervention and the control groups. A significant group × time interaction was observed at 12 weeks of intervention (Wald=11.57; *P*<.05), indicating that there was a significant improvement in the intervention group than the control group.

KWCP-SM

Table 3 shows the mean, SE, percentage change of outcomes, and the results of the Wilcoxon signed ranks test. Table 4 shows the GEE results of outcome valuables over time between the two groups. Regarding knowledge of smart medicine, at 12 weeks both the intervention group (29.35% change from baseline; *z*=−5.20; *P*<.001) and control group (27.06% change from baseline; *z*=−4.85; *P*<.001) had statistically significant within-group changes, but the difference between groups (Wald=1.14; *P*=.57) was not significant.

For willingness to use smart medicine, at 12 weeks the intervention group (19.53% change from baseline; *z*

XSL•FO
RenderX

score=−3.25, *P*=.001) showed a marked improvement compared with baseline score, while the control group (1.12% change from baseline; *z* score=−0.67; *P*=.50) showed a slight decrease. The GEE analysis revealed a significant interactive effect between time and groups (Wald=10.63; *P*=.005).

For confidence in smart medicine, both groups experienced improvement during the 12-week intervention, while this improvement was larger in the intervention group (30.25% change from baseline; *z* score=−3.75; *P*<.001) than the control group (5.45% change from baseline; *z* score=−1.77; *P*=.08). In addition, a significant difference in group × time interactive effect was observed (Wald=25.09; *P*<.001).

For the practice of smart medicine, both the intervention group (6.61% change from baseline; *z* score=−1.69; *P*=.09) and control group (4.14% change from baseline; *z* score=−1.50; *P*=.13) showed a slight improvement at 12 weeks compared with baseline. However, no significant between-group difference was observed (Wald=1.51; *P*=.47).

Table . Comparison of pre- and postintervention mean of outcomes (growth mindset; knowledge, willingness, confidence, and practice of smart medicine [KWCP-SM]).

Variables	Groups	T0[a], mean (SD)	T1[a], mean (SD)	T2[a], mean (SD)	Relative to baseline, %	*z* score	*P* value[b]
Growth mindset, ITIS[c]	IG[d]	18.50 (8.65)	21.89 (7.99)	20.33 (8.53)	9.89	−5.13	<.001
	CG[e]	17.53 (9.09)	17.79 (10.03)	18.00 (9.16)	2.68	−5.13	>.001
Knowledge, KWCP-SM[f]	IG	23.75 (7.24)	31.67 (5.13)	30.72 (6.09)	29.35	−5.20	<.001
	CG	23.47 (8.21)	31.61 (6.82)	29.82 (7.45)	27.06	−4.85	<.001
Willingness, KWCP-SM	IG	23.19 (7.63)	29.69 (7.52)	27.72 (8.16)	19.53	−3.25	.001
	CG	23.29 (9.46)	24.92 (8.55)	23.55 (7.13)	1.12	−0.67	.50
Confidence, KWCP-SM	IG	20.56 (8.47)	28.64 (8.23)	26.78 (9.13)	30.25	−3.75	<.001
	CG	20.74 (9.68)	21.87 (9.45)	21.87 (9.45)	5.45	−1.77	.08
Practice, KWCP-SM	IG	12.56 (4.77)	14.56 (5.57)	13.39 (5.18)	6.61	−1.69	.09
	CG	12.08 (4.72)	13.21 (5.31)	12.58 (4.79)	4.14	−1.50	.13

[a]Time points. T0: baseline; T1: 6-week; T2: 12-week.

[b]Within-group changes were examined with Wilcoxon signed ranks test.

[c]ITIS: Implicit theories of intelligence scale.

[d]IG: intervention group.

[e]CG: control group.

[f]KWCP-SM: Questionnaire of knowledge, willingness, confidence, and practice of smart medicine.

Table . Effects of intervention on outcome variables between two groups: Growth mindset and knowledge, willingness, confidence, and practice of smart medicine (KWCP-SM).

Outcomes	Group		Time		(Group × time)[a]	
	Wald	*P* value	Wald	*P* value	Wald	*P* value
Growth mindset	1.63	.20	15.84	<.001	11.57	.003
Knowledge	0.08	.78	204.28	<.001	1.14	.57
Willingness	3.26	.07	26.53	<.001	10.63	.005
Confidence	3.89	.049	40.95	<.001	25.09	<.001
Practice	0.68	.41	19.47	<.001	1.51	.47

[a]Results of growth mindset and knowledge, willingness, confidence, and practice of smart medicine (KWCP-SM) outcome analysis using the generalized estimation equation (GEE).

Discussion

Principal Findings

This study constructed a growth mindset intervention program for older patients with chronic diseases and evaluated its effects on their growth mindset and positive response to eHealth. After 12 weeks of intervention, the findings showed a significant improvement in the total ITIS-6 and KWCP-SM scores of patients in the intervention group compared with the control group, which suggests that the intervention was effective. We tailored our growth mindset program to address the mindset and micro-psychology of older adult patients with chronic diseases. This approach offers a new direction in bridging the digital divide targeting older adults.

XSL•FO
RenderX

Comparison With Previous Work

While the intervention program of this study effectively improves the growth mindset of older adults with chronic diseases,research on the effectiveness of growth mindset interventions in increasing the level of growth mindset in different populations is not entirely consistent. For example, Plaks and Chasteen [36] conducted a growth mindset psychological experiment with 88 older adults and found significant increases in the growth mindset in the experimental group. Orvidas et al [31] conducted an online growth mindset intervention with 48 obese children and adolescents and showed that the experimental group had better levels of growth mindset than the control group and significantly improved perceptions of health behaviors and self-efficacy. However, Sridharan et al [23] conducted an online growth mindset intervention with 398 smokers for 24 days and found that there was no significant difference in the level of growth mindset between the intervention group and the control group after the intervention, which may be due to the excessive amount of growth mindset intervention content provided to the participants. In this study, the intervention group showed better improvement in growth mindset than the control group, which could be attributed to the following reasons: First, we consulted previous growth mindset interventions that worked and adopted the framework of the intervention program previously published [27] as the main theme. At the same time, we combined the characteristics of psychological dependence of older adults to construct the themes, a total of 4 strategies along with the themes were included. The “Recognizing growth mindset” strategy used video lessons and face-to-face interaction to let the participants believe the plasticity of the brain of their own; the “Exploring growth mindset deeply” strategy used video lessons, telelecture, and story rewriting to promote deeply understanding and internalization of the growth mindset; the “Applying growth mindset” strategy used daily practice to promote the integration of growth mindset into the patient’s real life and learn to apply it; the “growth mindset in life” strategy used phone follow-up to help older adults with difficulties to solve problem and promote habit formation. Second, considering that effective training and dependence behavior are important measures for individual growth in old age [45], we helped patients identify the timing of a fixed mindset and adjust and cultivate a growth mindset through four steps: acceptance, observation, naming, and education. Furthermore, by teaching older adults about brain plasticity and the experience of growth mindset activities, this study, to a certain extent, broke the stereotype of “old age is frailty,” and instilled the hope of improvement in older adults. In addition, patients overcame difficulties through the growth mindset intervention and successfully mastered the experience of eHealth treatment, enabling them to realize that the growth of older adults is controllable, achievable, and recognized by society. This hopeful and emotional experience enhanced their belief that their ability can still be developed, improving their level of growth mindset.

After the 12-week eHealth program intervention, the total score of KWCP-SM, which measured the level of willingness and confidence in smart medicine, improved significantly in both the intervention and control groups compared to the pre-intervention period. This suggests that the traditional teaching model can improve the ability of older adults to use eHealth to some extent, and the result is consistent with the study by Zhao et al [46]. There are some possible explanations for this. First, existing studies have found that older adults experience major negative psychological emotions such as age-related stereotype, frustrated self-esteem, and dwindling motivation, which hinder the integration of eHealth in their lives [47,48]. Growth mindset intervention can effectively improve age-related stereotypes by instilling the belief in the plasticity of older adults’ intelligence and ability [49,50]. It also reduces patients’ excessive attention to learning outcomes, regards failure as a process of learning and improvement, protects their self-esteem and secondary fear of science and technology in the event of failure to a certain extent, and stimulates their internal motivation [51-53]. Second, the theory of growth mindset emphasizes that efforts, strategies, and timely help-seeking are important paths for learning and growth [27]. This study not only teaches patients to learn knowledge and skills, emphasizing that the key premise is hard work, but also empowers patients with eHealth learning strategies and help-seeking methods. Furthermore, our intervention met the practical needs of older adults who urgently need eHealth health care but are unable to take full advantage of it because they have not kept pace with technological developments. As a result, the willingness and interest of older people to learn are relatively high and the outcomes of this eHealth program are more positive.

Strengths and Limitations

The growth mindset intervention in this trial was developed in a theoretical and evidence-based manner and contains a tailored education program taking the characteristics of the older adults with chronic diseases into account. In addition, the training program organically integrates local slang into the training, increases the patients’ trust in the plasticity of older adults’ brains and their belief in lifelong learning, and sets up regular and multiple nurse-patient communication sessions, which facilitate real-time interaction and feedback, thus meeting the patients’ personalized needs.

Several limitations of this study should be noted here. First, the study was carried out at only one comprehensive hospital in Hangzhou. More multicenter and large sample studies should be carried out to further verify the scheme’s effectiveness. Second, the between-group difference in actual eHealth behavior change was not statistically significant. This may due to time constraints, only the effect of the growth mindset intervention for 12 weeks was confirmed. In the future, research should continue to track participants for 12 weeks or more to identify the effectiveness of the intervention program over the long term. Third, this study took the lead in building a growth mindset intervention program for older patients with chronic diseases and preliminarily verified the effectiveness. However, during the intervention, individual differences were found among older adults which may partly explain the nonsignificant finding, so future studies need to constantly adjust and improve the intervention plan according to the characteristics and needs of the participants.

XSL•FO
RenderX

Future studies should also implement the double-blind method and increase the objective evaluation indicators related to eHealth (such as frequency of visiting doctors, compliance, clinical nursing outcomes, etc.) to verify the efficacy of this program. Finally, future studies can build a platform for elderly growth mindset intervention with emerging technologies to achieve large-scale and accurate intervention, to improve the level of growth mindset in older adults to a greater extent.

Conclusions

The growth mindset intervention program for older adults with chronic diseases constructed in this study effectively improves the growth mindset of older adults with chronic diseases, enhances their mastery of eHealth knowledge, and boosts their confidence and willingness to use eHealth tools. It is expected to help patients meet their own diagnostic, treatment, disease control, and other practical needs through eHealth solutions. A growth mindset shows high potential in promoting eHealth use among older adults. This study is referable due to its creation of training content based on the growth mindset theory and investigation of the effect of the training on a range of variables. Further research is needed to improve the growth mindset intervention, in particular by considering older patients' needs and interests, improving its adoption, and developing fault-tolerant and lifelong growth environments that enhance growth mindset among older adults with chronic diseases. The growth mindset intervention program in this study could also be applied to enhance growth mindset and eHealth behaviors among older adults with chronic diseases in other countries by integrating localized case examples, community specific customs, and beliefs.

Acknowledgments

The authors would like to acknowledge the support given by School of Medicine and Nursing Huzhou University and School of Nursing Hangzhou Normal University.

Authors' Contributions

MC was responsible for conceptualization, writing, preparing the original draft, and reviewing and editing this paper. XX, YZ, BZ, and CX were responsible for conceptualization, reviewing, and editing this paper. XX, YZ, CZ, HW, and JL were responsible for methodology and data curation of this paper. All authors have read and agreed to the published version of the paper.

Conflicts of Interest

None declared.

Checklist 1
CONSORT-EHEALTH checklist V 1.6.1.
[PDF File, 367 KB - humanfactors_v12i1e65519_app1.pdf]

References

1. Marín-Cascales E, Rubio-Arias J, Alcaraz PE. Effects of two different neuromuscular training protocols on regional bone mass in postmenopausal women: a randomized controlled trial. Front Physiol 2019;10:846. [doi: 10.3389/fphys.2019.00846] [Medline: 31354513]
2. Zhongguo tongji nianjian [2020 China statistical year book]. National Bureau of Statistics of China. 2020. URL: https://www.stats.gov.cn/sj/ndsj/2020/indexch.htm [accessed 2025-06-17]
3. Guojia weisheng jiankang weiyuanhui erlingyijiunian shiyiyue shiyiri zhuanti xinwen fabuhui wenzi shilu [transcript of the NHSC's november 11, 2019 special press conference]. 2019 Nov 11. URL: http://www.nhc.gov.cn/xcs/s7847/201911/b01a5ca22bef4570ab44605c12940f97.shtml [accessed 2023-08-11]
4. Rashidi K, Jalali A, Salari N, Abbasi P. Validation of an instrument for assessing elder care needs in Iran. Clin Interv Aging 2020;15(275-83):275-283. [doi: 10.2147/CIA.S243142] [Medline: 32161451]
5. Falck L, Zoller M, Rosemann T, Martínez-González NA, Chmiel C. Toward standardized monitoring of patients with chronic diseases in primary care using electronic medical records: development of a tool by adapted delphi procedure. JMIR Med Inform 2020 Mar 25;8(3):e14483. [doi: 10.2196/14483] [Medline: 32209535]
6. Towle RM, Tan CG, Saptu K, et al. What do caregivers value and is there agreement in perception of met needs between nurses and caregivers? Singapore Med J 2019 Nov;60(11):575-582. [doi: 10.11622/smedj.2019040] [Medline: 31044258]
7. Zhou C, Hu J, Chen N. Remote care assistance in emergency department based on smart medical. J Healthc Eng 2021;2021(9971960):9971960. [doi: 10.1155/2021/9971960] [Medline: 34249300]
8. He J, Fu JY, Chen L, et al. Multicentre, prospective registry study of amyotrophic lateral sclerosis in mainland China (CHALSR): study protocol. BMJ Open 2020 Dec;10(12):e042603. [doi: 10.1136/bmjopen-2020-042603]
9. Chen Y, Zhang L, Wei M. How does smart healthcare service affect resident health in the digital age? empirical evidence from 105 cities of China. Front Public Health 2021;9(833687):833687. [doi: 10.3389/fpubh.2021.833687] [Medline: 35127633]

XSL•FO
RenderX

10. China Internet Network Information Center. Di sishiqici zhongguo hulianwangluo fazhanzhuangkuang tongjibaogao [the 47th statistical report on the development of the internet in china]. 2021. URL: https://www.cac.gov.cn/2021-02/03/c_1613923423079314.htm [accessed 2021-03-02]
11. Huang F, Wu X, Xie Y, et al. An automated structured education intervention based on a smartphone app in Chinese patients with type 1 diabetes: a protocol for a single-blinded randomized controlled trial. Trials 2020 Nov 23;21(1):944. [doi: 10.1186/s13063-020-04835-9] [Medline: 33225982]
12. Ren T. The management effect of mobile medical chronic disease management system based on wearable devices on middle-aged and elderly hypertension and diabetes patients in community. : Chongqing Medical University; 2019 URL: https://kns.cnki.net/kcms2/article/[illegible] [accessed 2025-06-26]
13. Miao S, Shen C, Feng X, Zhu Q, Shorfuzzaman M, Lv Z. Upper limb rehabilitation system for stroke survivors based on multi-modal sensors and machine learning. IEEE Access 2021;9:30283-30291. [doi: 10.1109/ACCESS.2021.3055960]
14. Vailati Riboni F, Pagnini F. Age-based stereotype threat: a scoping review of stereotype priming techniques and their effects on the aging process. Aging Ment Health 2022 Jul;26(7):1444-1450. [doi: 10.1080/13607863.2021.1932741] [Medline: 34060405]
15. Pywell J, Vijaykumar S, Dodd A, Coventry L. Barriers to older adults' uptake of mobile-based mental health interventions. Digit Health 2020;6:2055207620905422. [doi: 10.1177/2055207620905422] [Medline: 32110429]
16. Zibrik L, Khan S, Bangar N, Stacy E, Novak Lauscher H, Ho K. Patient and community centered eHealth: Exploring eHealth barriers and facilitators for chronic disease self-management within British Columbia's immigrant Chinese and Punjabi seniors. Health Policy Technol 2015 Dec;4(4):348-356. [doi: 10.1016/j.hlpt.2015.08.002]
17. Paige SR, Miller MD, Krieger JL, Stellefson M, Cheong J. Electronic health literacy across the lifespan: measurement invariance study. J Med Internet Res 2018 Jul 9;20(7):e10434. [doi: 10.2196/10434] [Medline: 29986848]
18. Kuang Y. Barriers to digital integration of the elderly group: influencing factors, user portraits and policy responses. J Huazhong Univ Sci Technolog Med Sci 2022:46-53. [doi: 10.19648/j.cnki.jhustss1980.2022.01.06]
19. Zhang S, Hao X, Ma C, Liu Q, Zheng Y. Analysis on the elderly's recognition and willingness to use of the "Internet+medical care. Health Economics Research 2021:36-39. [doi: 10.14055/j.cnki.33-1056/f.2021.11.010]
20. Hwang YK, Lee CS. Relationship between body image, growth mindset, grit, and successful aging in Korean elderly: moderated mediation effect of hope. Medico-legal Update 2020;20(1). [doi: 10.37506/mlu.v20i1.707]
21. Heslin PA, Burnette JL, Ryu NG. Does a growth mindset enable successful aging? Work, Aging and Retirement 2021 Mar 25;7(2):79-89. [doi: 10.1093/workar/waaa029]
22. Hiu SKW, Rabinovich A. Lay theories and ageist attitudes at work: essentialist beliefs about cognitive aging, fixed mindsets, and cyclical perception of time are linked to ageist attitudes toward older workers. J Applied Social Pyschol 2021 Mar;51(3):190-207 [FREE Full text] [doi: 10.1111/jasp.12725]
23. Sridharan V, Shoda Y, Heffner J, Bricker J. A pilot randomized controlled trial of a web-based growth mindset intervention to enhance the effectiveness of a smartphone app for smoking cessation. JMIR Mhealth Uhealth 2019 Jul 9;7(7):e14602. [doi: 10.2196/14602] [Medline: 31290404]
24. Dweck CS, Leggett EL. A social-cognitive approach to motivation and personality. Psychol Rev 1988;95(2):256-273. [doi: 10.1037/0033-295X.95.2.256]
25. Lewis LS, Williams CA, Dawson SD. Growth mindset training and effective learning strategies in community college registered nursing students. Teaching and Learning in Nursing 2020 Apr;15(2):123-127. [doi: 10.1016/j.teln.2020.01.006]
26. Burnette JL, O'Boyle EH, VanEpps EM, Pollack JM, Finkel EJ. Mind-sets matter: a meta-analytic review of implicit theories and self-regulation. Psychol Bull 2013 May;139(3):655-701. [doi: 10.1037/a0029531] [Medline: 22866678]
27. Yeager DS, Hanselman P, Walton GM, et al. A national experiment reveals where a growth mindset improves achievement. Nature New Biol 2019 Sep;573(7774):364-369. [doi: 10.1038/s41586-019-1466-y] [Medline: 31391586]
28. Burnette JL, Pollack JM, Forsyth RB, et al. A growth mindset intervention: enhancing students' entrepreneurial self-efficacy and career development. Entrepreneurship Theory and Practice 2020 Sep;44(5):878-908. [doi: 10.1177/1042258719864293]
29. Bettinger E, Ludvigsen S, Rege M, Solli IF, Yeager D. Increasing perseverance in math: evidence from a field experiment in Norway. J Econ Behav Organ 2018 Feb;146:1-15. [doi: 10.1016/j.jebo.2017.11.032]
30. Burnette JL, Finkel EJ. Buffering against weight gain following dieting setbacks: an implicit theory intervention. J Exp Soc Psychol 2012 May;48(3):721-725. [doi: 10.1016/j.jesp.2011.12.020]
31. Orvidas K, Burnette JL, Schleider JL, Skelton JA, Moses M, Dunsmore JC. Healthy body, healthy mind: a mindset intervention for obese youth. J Genet Psychol 2020;181(6):443-457. [doi: 10.1080/00221325.2020.1796573] [Medline: 32734811]
32. Marquet M, Plaks JE, Balasubramaniam L, Brunet S, Chasteen AL. Older adults' lay theories of abilities predict their willingness to engage in preventive health behaviors. Arch Gerontol Geriatr 2021;93:104291. [doi: 10.1016/j.archger.2020.104291B3] [Medline: 33166728]
33. Wu R, Rebok GW, Lin FV. A novel theoretical life course framework for triggering cognitive development across the lifespan. Hum Dev 2016;59(6):342-365. [doi: 10.1159/000458720]

XSL•FO
RenderX

34. Hwang YK, Lee CS. Relationship between body image, growth mindset, grit, and successful aging in Korean elderly: moderated mediation effect of hope. Medico Legal Update 2020;20(1):2196-2202. [doi: 10.37506/mlu.v20i1.707]
35. Weiss D. On the inevitability of aging: essentialist beliefs moderate the impact of negative age stereotypes on older adults' memory performance and physiological reactivity. J Gerontol B Psychol Sci Soc Sci 2018 Aug 14;73(6):925-933. [doi: 10.1093/geronb/gbw087] [Medline: 27440918]
36. Plaks JE, Chasteen AL. Entity versus incremental theories predict older adults' memory performance. Psychol Aging 2013 Dec;28(4):948-957. [doi: 10.1037/a0034348B1] [Medline: 24128076]
37. Dang J, Liu L. A growth mindset about human minds promotes positive responses to intelligent technology. Cognition 2022 Mar;220:104985. [doi: 10.1016/j.cognition.2021.104985] [Medline: 34920301]
38. Liu K, Chen JM. The mindset, lifelong learning, and gerotranscendence of elderly women in Taiwan: an exploratory study. In: Liu K, J-m C, editors. Womens Stud Int Forum 2022 Jan;90:102552. [doi: 10.1016/j.wsif.2021.102552]
39. Cheng WLS. The effects of mutual goal-setting practice in older adults with chronic illness. Geriatr Nurs (Lond) 2018 Mar;39(2):143-150. [doi: 10.1016/j.gerinurse.2017.07.007]
40. Babinski LM, Murray DW, Wilson WA, Kuhn CM, Malone PS. Impact of a neuroscience-based health education course on high school students' health knowledge, beliefs, and behaviors. J Adolesc Health 2018 Oct;63(4):489-496. [doi: 10.1016/j.jadohealth.2018.05.016] [Medline: 30286902]
41. Baltes PB, Lindenberger U, Staudinger UM. Life span theory in developmental psychology. In: Handbook of Child Psychology: Wiley; 1998:1029-1143 URL: https://pure.mpg.de/rest/items/item_2103192/component/file_2103191/content [accessed 2025-06-27]
42. Brez C, Hampton EM, Behrendt L, Brown L, Powers J. Failure to replicate: testing a growth mindset intervention for college student success. Basic Appl Soc Psych 2020 Nov 1;42(6):460-468. [doi: 10.1080/01973533.2020.1806845]
43. Lee J, Lee HJ, Song J, Bong M. Enhancing children's math motivation with a joint intervention on mindset and gender stereotypes. Learn Instr 2021 Jun;73:101416. [doi: 10.1016/j.learninstruc.2020.101416]
44. Blackwell LS, Trzesniewski KH, Dweck CS. Implicit theories of intelligence predict achievement across an adolescent transition: a longitudinal study and an intervention. Child Dev 2007;78(1):246-263. [doi: 10.1111/j.1467-8624.2007.00995.x] [Medline: 17328703]
45. Baltes PB, Lindenberger U, Staudinger UM. Life Span Theory in Developmental Psychology 2007. [doi: 10.1002/9780470147658.chpsy0111]
46. Zhao X, Wang L, Ge C, et al. Smartphone application training program improves smartphone usage competency and quality of life among the elderly in an elder university in China: a randomized controlled trial. Int J Med Inform 2020 Jan;133(104010):104010. [doi: 10.1016/j.ijmedinf.2019.104010] [Medline: 31675540]
47. Ivan L, Cutler SJ. Ageism and Technology: The Role of Internalized Stereotypes: National University of Political Studies and Public Administration; 2021:127-139. [doi: 10.3138/utq.90.2.05]
48. Reinwand DA, Crutzen R, Zank S. Online activities among elder informal caregivers: results from a cross-sectional study. Digit Health 2018;4(2055207618779715):2055207618779715. [doi: 10.1177/2055207618779715] [Medline: 31463072]
49. Mrazek AJ, Ihm ED, Molden DC, Mrazek MD, Zedelius CM, Schooler JW. Expanding minds: growth mindsets of self-regulation and the influences on effort and perseverance. J Exp Soc Psychol 2018 Nov;79:164-180. [doi: 10.1016/j.jesp.2018.07.003]
50. Pagnini F, Cavalera C, Volpato E, et al. Ageing as a mindset: a study protocol to rejuvenate older adults with a counterclockwise psychological intervention. BMJ Open 2019 Jul 9;9(7):e030411. [doi: 10.1136/bmjopen-2019-030411] [Medline: 31289097]
51. Wolcott MD, McLaughlin JE, Hann A, et al. A review to characterise and map the growth mindset theory in health professions education. Med Educ 2021 Apr;55(4):430-440. [doi: 10.1111/medu.14381] [Medline: 32955728]
52. Liu WC. Implicit theories of intelligence and achievement goals: a look at students' intrinsic motivation and achievement in mathematics. Front Psychol 2021;12(593715):593715. [doi: 10.3389/fpsyg.2021.593715] [Medline: 33692718]
53. Passalent L, Cyr A, Jurisica I, Mathur S, Inman RD, Haroon N. Motivators, barriers, and opportunity for e-health to encourage physical activity in axial spondyloarthritis: a qualitative descriptive study. Arthritis Care Res (Hoboken) 2022 Jan;74(1):50-58. [doi: 10.1002/acr.24788] [Medline: 34928533]

Abbreviations

GEE: generalized estimation equation
ITIS-6: implicit theories of intelligence scale-6
KWCP-SM: knowledge, willingness, confidence, and practice of smart medicine

XSL•FO
RenderX

Edited by A Kushniruk; submitted 18.08.24; peer-reviewed by F Pagnini; revised version received 28.04.25; accepted 05.05.25; published 03.07.25.

Please cite as:
Cao M, Xu X, Zeng Y, Zhao B, Xie C, Wu H, Lou J
Growth Mindset Intervention's Impact on Positive Response to eHealth for Older Adults With Chronic Disease: Randomized Controlled Trial
JMIR Hum Factors 2025;12:e65519
URL: https://humanfactors.jmir.org/2025/1/e65519
doi:10.2196/65519

© Meijuan Cao, Xiaojuan Xu,Yaling Zeng, Binyu Zhao, Chunqi Xie, Hailu Wu, Jianlin Lou. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 3.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Digital Informed Consent/Assent in Clinical Trials Among Pregnant Women, Minors, and Adults: Multicountry Cross-Sectional Evaluation of Comprehension and Satisfaction

Jaime Fons-Martinez[1,2], PhD; Carlos Murciano-Gamborino[1,2], BA; Cristina Ferrer-Albero[3], PhD; Carlos Vergara-Hernandez[4], MSc; Javier Diez-Domingo[1,2], PhD

[1]Vaccines Research Department, FISABIO (Foundation for the Promotion of Health and Biomedical Research of Valencia Region), Avenida Cataluña, 21, Valencia, Spain

[2]CIBER de Epidemiología y Salud Pública, Instituto de Salud Carlos III, Madrid, Spain

[3]Universidad Católica de Valencia San Vicente Mártir. Facultad de Medicina y Ciencias de la Salud, Valencia, Spain

[4]Statistics Service, Platforms and Services, FISABIO (Foundation for the Promotion of Health and Biomedical Research of Valencia Region), Valencia, Spain

Corresponding Author:
Jaime Fons-Martinez, PhD
Vaccines Research Department, FISABIO (Foundation for the Promotion of Health and Biomedical Research of Valencia Region), Avenida Cataluña, 21, Valencia, Spain

Abstract

Background: Informed consent (IC) is a cornerstone of ethical clinical research, yet comprehension gaps persist. The i-CONSENT guidelines aim to improve IC materials by enhancing clarity, accessibility, and tailoring them to the needs of diverse populations. This study evaluates the effectiveness of electronic IC (eIC) materials developed under these guidelines for 3 target populations—minors, pregnant women, and adults—across Spain, the United Kingdom, and Romania.

Objective: The primary aim of this study is to assess participants' comprehension of and satisfaction with eIC materials tailored to their specific needs. The secondary objectives are to identify demographic predictors of comprehension, evaluate the cross-cultural applicability of materials, and explore format preferences.

Methods: A cross-sectional study was conducted with 1757 participants (620 minors, 312 pregnant women, and 825 adults), who reviewed eIC materials through a digital platform offering layered web content, narrative videos, printable documents, and infographics. Materials were co-designed using participatory methods, including design thinking sessions with minors and pregnant women, and online surveys with adults. Comprehension was assessed using an adapted version of the Quality of the Informed Consent questionnaire. Objective comprehension (part A) was categorized as low (<70%), moderate (70% - 80%), adequate (80% - 90%), or high (≥90%). Subjective comprehension (part B) was measured using a 5-point Likert scale. Satisfaction was evaluated through Likert scales and usability questions, with scores ≥80% considered acceptable. Multivariable regression models were applied to identify predictors of comprehension.

Results: Objective comprehension exceeded 80% across all groups: minors (mean 83.3, SD 13.5), pregnant women (mean 82.2, SD 11.0), and adults (mean 84.8, SD 10.8). Women/girls outperformed men/boys (β=+.16 to +.36). Generation X adults scored higher than millennials (β=+.26, *P*<.001), while prior trial participation was associated with lower comprehension scores (β=−.47 to −1.77). Among minors, compared with participants from Spain with no previous clinical trial experience, comprehension was significantly lower in Spain (*P*=.03), Romania (*P*<.001), and the United Kingdom (*P*<.001). Format preferences varied: 382 out of 620 (61.6%) minors and 152 out of 312 (48.7%) pregnant women preferred videos, whereas 452 out of 825 (54.8%) adults favored text (*P*<.001). Satisfaction rates surpassed 90% in all groups (minors, 604/620, 97.4%; pregnant women, 303/312, 97.1%; and adults, 804/825, 97.5%), with 777 out of 825 (94.2%) adults also indicating that the materials facilitated understanding. While translated materials maintained high efficacy across countries, comprehension scores in Romania were lower among participants with lower educational levels (β=−1.05, *P*=.001). Materials cocreated in Spain were effective across countries but yielded higher comprehension within the original target population.

Conclusions: eIC materials developed following the i-CONSENT guidelines achieved high levels of comprehension and satisfaction across diverse populations, demonstrating scalability for multinational trials. Cocreation and multimodal design effectively addressed participant preferences; however, cultural adaptation remained crucial for optimizing outcomes. The negative impact of prior trial participation highlights the need for tailored engagement strategies for returning participants. Future research should explore regional disparities, evaluate interventions for overconfident returning participants, and validate these tools across broader cultural contexts.

XSL•FO
RenderX

(JMIR Hum Factors 2025;12:e65569) doi:10.2196/65569

KEYWORDS

informed consent; eConsent; clinical trials; comprehension; pregnant women; adults; participant satisfaction; demographic factors; eHealth; electronic consent materials; pregnant; women; minor; youth; adolescent; teenager; older adult; electronic informed consent; digital informed consent; statistical analysis

Introduction

Background

Informed consent (IC) is a key element in ensuring the autonomy of potential participants when deciding whether to partake in a study. To be effective, IC must meet 5 key criteria: voluntariness, capacity, disclosure, understanding, and decision-making. However, numerous studies have highlighted persistent gaps in participants' comprehension of the information provided during the IC process [1].

The i-CONSENT project was initiated to address these challenges by improving IC materials to make them more comprehensible, accessible, and tailored to the specific needs of diverse populations. As part of this initiative, the project developed the "Guidelines for Tailoring the Informed Consent Process in Clinical Studies" [2] (i-CONSENT guidelines). Among the recommendations included in the guidelines, involving potential participants in the preparation of the IC and its associated materials has been identified as a key factor [3].

To evaluate the impact of these guidelines, 3 mock studies were conducted, each focusing on a distinct population: minors, pregnant women, and adults. The IC materials were presented in electronic formats (electronic IC [eIC]), incorporating innovative features such as layered web content, narrative videos, printable documents, and customized infographics. These formats allowed participants to engage with the information in ways that suited their preferences and cognitive styles.

This study aims to assess participants' comprehension of and satisfaction with eIC materials tailored to their specific needs. Additionally, it seeks to explore cross-cultural applicability by implementing these materials across 3 countries—Spain, the United Kingdom, and Romania—and examining demographic predictors of comprehension. The findings contribute to advancing inclusive and participant-centered IC processes in clinical research while promoting informed decision-making.

Objective

The main objective of this study is to estimate the proportion of potential participants who understand the 3 different eIC materials prepared following the i-CONSENT guidelines. Additional objectives are to (1) analyze differences in comprehension across countries (Spain, the United Kingdom, and Romania) and evaluate whether materials designed for one region apply to other regions or languages; (2) assess comprehension within specific domains of IC information; and (3) evaluate participant satisfaction with the eIC materials. Additional objectives for adults include exploring differences in comprehension by gender and age, while for minors, the objective is to explore differences in comprehension by gender.

Methods

Elaboration of e-Consent Materials

Overview

Electronic information materials have been prepared for 3 mock studies on vaccine clinical trials, each addressed to a different population group (see Table 1).

XSL•FO
RenderX

Table . Description of the 3 case studies.

Variable	Minors	Pregnant women	Adults
Target population			
Age (years)	12-13	≥18	Millennials: 18 - 38; Generation X: 39 - 53
Genders included	Both male and female	N/A[a]	Both male and female
Sample size, n	620	312	825
Case scenarios	Human papillomavirus vaccination in minors, considering gender differences.	Respiratory syncytial virus vaccination in pregnant women, considering both their needs and those of their babies.	Meningococcal vaccination in adults, considering gender and generational differences.
How was the target population involved?	One design thinking session with children and parents, 1 session with children alone, and piloting of the contents of the information sheet and the survey.	Two design thinking sessions with pregnant women and piloting of materials.	An online-based self-administered survey used for piloting.
Materials prepared	eConsent[b], including information in 3 formats: a website page (layered approach), a video (storytelling format), and a document (improved format).	eConsent included information in 3 formats: a website page (layered approach), 1 infographic (about the study procedures), a video (question-and-answer format), and a document (improved format).	eConsent included information in 3 formats: a website page (layered approach), 5 infographics (covering general information, procedures, benefits and risks, legal aspects, and data protection), and a document (improved format).

[a]N/A: not applicable.

[b]eConsent: electronic consent.

The materials have been prepared following the i-CONSENT guidelines, taking into account potential participants' preferences and needs [4]. Different formats were included, and participants could choose among these formats or combine them. Table 1 provides information about the materials prepared.

Development Process

The materials were originally prepared in Spanish through a cocreation process involving representatives from the target population. A multidisciplinary team comprising clinical trial physicians, epidemiologists, a sociologist, a journalist, and a nurse collaborated on the design. This approach ensured that the materials were scientifically accurate while addressing the cognitive and cultural needs of participants. The cocreation methodology included participatory sessions with minors and pregnant women in Spain to ensure that the materials were relevant and engaging for these groups. For adults, online surveys provided insights into preferences and usability. This iterative process allowed for refinement based on user feedback before finalizing the materials.

To facilitate cross-cultural implementation, the materials were professionally translated into English and Romanian by native speakers of the target languages. The translation process adhered to a rigorous rubric that prioritized fidelity to meaning, contextual appropriateness, and adaptation to local customs and linguistic conventions. Each translation was independently reviewed by another professional translator to ensure quality and consistency.

Formats Offered

Participants accessed the eIC materials via a dedicated website, where they could choose from multiple formats based on their preferences or combine them as needed. The formats offered (depending on the specific case study) included the following:

- Layered web content: A modular approach that allowed participants to access additional details or definitions by clicking on specific terms.
- Narrative videos: Tailored video formats for minors and pregnant women (a question-and-answer style for pregnant women and narrative storytelling for minors).
- Printable documents: Text-based materials with integrated images, designed for participants who prefer physical copies.
- Infographics: Simplified visual representations of complex topics such as legal aspects or participant rights.

The formats were not mutually exclusive, and participants could engage with any combination of formats according to their preferences.

Comprehension Assessment Tool

Adaptation of the Quality of the Informed Consent Questionnaire for Target Populations

To evaluate comprehension, 3 tailored adaptations of the Quality of the Informed Consent questionnaire (QuIC) [5] were used, 1 for each mock study. The adaptations were based on the original version of the QuiC developed by Joffe et al [5] and subsequent modifications by Paris et al [6,7]. The adjustments accounted for the nature of the study (vaccine clinical trials in healthy populations), the characteristics of the target groups (minors, pregnant women, and adults), and European regulations on data protection and clinical trials.

XSL•FO
RenderX

Development and Validation

Each adaptation process was conducted independently for the respective population. For minors and pregnant women, cocreation sessions were held with representatives from these groups to ensure that the questionnaires were appropriate and comprehensible. Feedback from these sessions informed iterative revisions, resulting in the final versions of the tools. The adapted questionnaires were originally developed in Spanish and subsequently translated into English and Romanian by professional translators. These translations followed a rigorous rubric that prioritized fidelity to meaning, cultural appropriateness, and linguistic clarity.

Structure of the Questionnaires

The surveys, as their predecessors, consist of 2 parts to assess understanding: part A measures objective understanding, and part B measures subjective understanding.

In the surveys for adults and pregnant women, part A includes 22 questions with 3 response options ("no," "don't know," and "yes"). The scoring system follows that of the original version of the QuIC [5]: 0 points for an incorrect answer, 50 points for a "don't know" response, 100 points for a correct answer. The questions in part A are grouped into 7 domains: "Nature/purposes of research," "Risks and benefits," "Alternative procedures," "Legal, participant protection," "Compensation," "Contacts, information," and "Participants' rights." Part B consists of 13 questions scored using a 5-point Likert scale. Scores are assigned as follows: 0 points for a rating of 1, 25 points for 2, 50 points for 3, 75 points for 4, and 100 points for 5.

The survey to measure children's understanding was designed in a cocreation session with children and resulted in the Assent Comprehension Questionnaire for vaccine studies (abbreviated *CCAsIn*, from its Spanish title *Cuestionario de Comprensión del Asentimiento Informado*) [4]. In this survey, part A includes 14 questions with 2 response options ("agree" and "disagree"); correct answers score 100 points, and incorrect answers score 0 points. Part B consists of 10 questions, scored using the same values as in the adult version. The minors' survey includes 3 domains: "Nature/purposes of research," "Risks and benefits," and "Participants' rights."

The global score of each part was calculated as the average score of all questions in that part. Domain-specific scores were also calculated for part A by averaging the scores within each domain.

Additionally, 2 sections were added: 1 with sociodemographic questions and 1 with questions addressing acceptance, preferences, and satisfaction with the eIC materials. A multiple-choice question was included to capture participants' preferences regarding information formats.

The final versions of the questionnaires underwent pilot testing with representatives from each target population to ensure validity and reliability. Feedback from participants informed minor adjustments before implementation in the main study.

Parts A and B of each survey are included in Multimedia Appendix 1, along with the frequency of responses for each question. Multimedia Appendix 1 also presents the questions grouped by domain.

Ethical Considerations

Ethics Review and Approvals

This study was conducted in accordance with the Declaration of Helsinki and received ethical approval from the Research Ethics Committee of the Foundation for the Promotion of Health and Biomedical Research of the Valencia Region (FISABIO; approval number 20200109/09).

Informed Consent

All participants provided IC before participation (see Multimedia Appendix 1). For minors, parental/guardian consent was obtained alongside assent from the children themselves. Online participants provided consent/assent electronically via a digital form integrated into the survey platform. Participants were explicitly informed that their involvement was voluntary and not related to any actual clinical trial.

Privacy and Confidentiality

All data collected were anonymized and securely stored on password-protected servers. Personal identifiers were removed before analysis, and participants were assured that their responses would remain confidential. Access to the raw data was restricted to the research team.

Compensation

Participants received compensation based on their recruitment group:

- Minors in Spain: Schools where the surveys were conducted received an Amazon voucher worth €200 (US $235) as an institutional incentive. For minors outside Spain, no direct compensation was provided due to logistical constraints; instead, recruitment was facilitated by the market research company (GfK) that visited schools to oversee participation.
- Adults and pregnant women: Participants recruited through the market research company's panel database received points equivalent to €4 to €5 (US $5 to US $6), which could be redeemed for gift cards or other rewards through the panel's system.

Recruitment

Rationale for Selecting Target Populations

The study targeted 3 distinct populations: minors, pregnant women, and adults. These groups were selected due to their specific ethical and regulatory considerations in vaccine clinical trials. Minors were included because their participation requires both assent and parental authorization, presenting ethical challenges due to their classification as a vulnerable population. Pregnant women were chosen to examine ethical safeguards and risk perceptions that may influence their willingness to participate, particularly in the context of vaccines administered during pregnancy to protect the unborn child. Adults were included as they represent the primary target population for vaccine trials, allowing for an assessment of IC comprehension and decision-making within a generally healthy population, as well as enabling intergenerational comparisons.

XSL•FO
RenderX

Surveys were conducted in 3 countries —Spain, the United Kingdom, and Romania—through the market research company with established household panel databases (30,000 panelists in Spain, 700,000 in Romania, and 2 million in the United Kingdom). The selection of these countries followed a cluster sampling approach based on their geographical distribution across Northern, Southern, and Eastern Europe. Factors such as cultural and linguistic diversity, country size, the number of registered clinical trials, and affiliation with the consortium were considered in the selection process.

Recruitment of Minors

For minors, recruitment was initially intended to be conducted face-to-face in schools. However, due to COVID-19 restrictions, this approach was adapted to online surveys in the United Kingdom and Romania. In Spain, surveys were partially conducted at state-subsidized schools under controlled conditions and complemented by online surveys completed at home. To ensure comparability across countries, recruitment targeted children aged 12 - 13 attending state-subsidized schools or equivalent educational settings.

A recruitment questionnaire was sent to randomly selected panelists with children aged 12 or 13 years. Parents were asked whether their children attended subsidized schools; if they responded affirmatively, they were invited to participate. Once parental IC was obtained, children were asked to provide their assent to participate. Parents also signed a certification confirming that survey responses were provided by their children rather than themselves. Materials and surveys were then sent online to participants who agreed to participate.

In Spain, 187 out of 312 students completed the survey in information technology classrooms under the supervision of study representatives, while the remainder participated from home. The data obtained across all 3 countries were fully comparable, as the sample profiles of children were matched across regions and the surveys that were conducted using similar procedures.

Recruitment of Adults and Pregnant Women

Adults and pregnant women were recruited using the market research company's household panel database for online studies. Invitations were sent to randomized samples of panelists who met the age and sex criteria. Sampling was controlled by area, age group, and socioeconomic level to ensure representativeness.

To identify pregnant women, the recruitment questionnaire included the question, "Are you currently pregnant?" All women who replied in the affirmative were included in the study group for pregnant women.

In Spain, due to the smaller panel size for pregnant women and the larger sample requirements, part of the recruitment was conducted face-to-face near health centers in Valencia and Madrid. This was carried out by professional recruiters using convenience sampling criteria. Pregnant women approached by recruiters were asked whether they wished to participate after being presented with the study details. Interested participants provided their email addresses via a dedicated form for follow-up communication. Subsequently, documentation and participation links were sent electronically in the same manner as for other panelists.

Instructions Provided to Participants

Participants received materials via email along with a standardized script explaining the study's purpose and procedures. This script included (1) an introduction to FISABIO's research within the framework of the European i-CONSENT project; (2) a statement clarifying that participation was voluntary and unrelated to any actual clinical trial; (3) an overview of the study structure (reviewing IC materials followed by a survey); and (4) estimated times for viewing the materials (35-40 minutes) and completing the survey (a minimum of 12 minutes).

Participants were instructed to review all materials before completing the survey. They accessed the materials via a private link sent to their email or provided during recruitment. The website allowed participants to open the materials in a browser tab while completing the survey in another, enabling them to refer back to the information as needed during data collection.

Fieldwork Period

All fieldwork for all study groups and countries was conducted between September and October 2020.

Survey Platform and Data Management

The surveys were administered using Confirmit Horizons v24, a multimodal platform employed by GfK for data collection via computer-assisted telephone interviewing, computer-assisted personal interviewing, or computer-assisted web interviewing methodologies. Data were securely transmitted from the platform to the research team, ensuring anonymization and compliance with data protection standards.

Adaptations Due to the COVID-19 Pandemic

A pilot test was conducted in Spain in July 2020 to evaluate the feasibility of face-to-face recruitment. The results (530 people approached, 467 uninterested, 5 completed surveys on-site, and 58 completed surveys from home via email) highlighted significant challenges with in-person recruitment and demonstrated the practicality of online data collection methods. Consequently, all surveys were conducted online, except for minors in Spain, where a hybrid approach was used (see the CHERRIES [Checklist for Reporting Results of Internet E-Surveys] checklist in Multimedia Appendix 1).

Statistical Analysis

Summary of Descriptive Statistics

Descriptive statistics were calculated for the study database, stratified by country, age, and gender (except for pregnant women, where gender was not applicable). Categorical variables were summarized using frequencies and percentages, while quantitative variables were described using means and SDs.

Comprehension Score

Comprehension of the study materials was assessed using the total score from part A of the adapted QuIC surveys, as well as by individual domains of IC. Scores were categorized as follows, based on the maximum score: below 70% indicated low

XSL•FO
RenderX

comprehension; between 70% and 79% indicated moderate comprehension; between 80% and 89% indicated adequate comprehension; and 90% or higher indicated high comprehension.

These cut-off points were established based on previous studies [8-10].

Regression Modeling

To evaluate differences in comprehension scores (part A) by country, age (adults), or gender (adults and minors), a linear regression model was constructed. Given the exploratory nature of the objectives, the model was adjusted for age (adults only), country, gender, previous participation in a clinical trial, and educational level (up to primary vs tertiary education, adults only), as these factors were anticipated to be strong predictors of comprehension. Interactions between variables were tested, and the Akaike information criterion (AIC) was used to select the optimal model by penalizing for model complexity. As the mean score did not follow a normal distribution, an ordered quantile normalizing transformation was applied to the data before regression analysis [11].

The analysis was carried out using the statistical software R, version 4.0.3 (R Foundation).

Data Exclusion Criteria

Participants whose response pattern indicated that more than 80% of items in the part A questionnaire were answered as "don't know" were excluded from the analyses (n=6).

Results

Study Sample and Participant Characteristics

After exclusions, a total of 1757 participants were included in the analysis: 825 adults, 312 pregnant women, and 620 minors. Key sociodemographic characteristics are summarized in Tables 2-4.

Table . Characteristics of the minors (n=620).

Characteristics	Spain (n=312), n (%)	Romania (n=125), n (%)	United Kingdom (n=183), n (%)	All (n=620), n (%)
Age				
12 years old	149 (47.8)	72 (57.6)	103 (56.3)	324 (52.3)
13 years old	163 (52.2)	53 (42.4)	80 (43.7)	296 (47.7)
Gender				
Male	167 (53.5)	47 (37.6)	88 (48.1)	302 (48.7)
Female	145 (46.5)	78 (62.4)	95 (51.9)	318 (51.3)
Previous participation in a real clinical trial				
No	293 (93.9)	118 (94.4)	176 (96.2)	587 (94.7)
Yes	19 (6.1)	7 (5.60)	7 (3.8)	33 (5.3)

Table . Characteristics of the pregnant women (n=312).

Characteristics	Spain (n=163), n (%)	Romania (n=89), n (%)	United Kingdom (n=60), n (%)	All (n=312), n (%)
Age groups				
<29 years old	41 (25.2)	30 (33.7)	28 (46.7)	99 (31.7)
29 - 38 years old	100 (61.3)	47 (52.8)	29 (48.3)	176 (56.4)
39 - 45 years old	18 (11.0)	12 (13.5)	3 (5.0)	33 (10.6)
>45 years old	4 (2.5)	0 (0)	0 (0)	4 (1.3)
Previous participation in a real clinical trial				
No	142 (87.1)	76 (85.4)	45 (75.0)	263 (84.3)
Yes	21 (12.9)	13 (14.6)	15 (25.0)	49 (15.7)
Educational level				
None completed/primary school/high school	52 (31.9)	17 (19.1)	10 (16.7)	79 (25.3)
University or higher	111 (68.1)	72 (80.9)	50 (83.3)	233 (74.7)

XSL•FO
RenderX

Table . Characteristics of the adults (n=825)

Characteristics	Spain (n=420), n (%)	Romania (n=245), n (%)	United Kingdom (n=160), n (%)	All (n=825), n (%)
Gender				
Male	219 (52.1)	122 (49.8)	79 (49.4)	420 (50.9)
Female	201 (47.9)	123 (50.2)	81 (50.6)	405 (49.1)
Age group				
<29 years old	105 (25.0)	57 (23.3)	37 (23.1)	199 (24.1)
29 - 38 years old	112 (26.7)	63 (25.7)	43 (26.9)	218 (26.4)
39 - 45 years old	99 (23.6)	65 (26.5)	40 (25.0)	204 (24.7)
>45 years old	104 (24.8)	60 (24.5)	40 (25.0)	204 (24.7)
Previous participation in a real clinical trial				
No	401 (95.5)	230 (93.9)	134 (83.8)	765 (92.7)
Yes	19 (4.5)	15 (6.1)	26 (16.3)	60 (7.3)
Educational level				
None completed/primary school/high school	179 (42.6)	23 (9.4)	46 (28.8)	248 (30.1)
University or higher	241 (57.4)	222 (90.6)	114 (71.3)	577 (69.9)

Objective Comprehension (Part A)

Comprehension Outcomes by Population Group

The mean objective comprehension scores (adapted QuIC part A) were high across all groups: minors, mean 83.3 (SD 13.5); pregnant women, mean 82.2 (SD 11.0); and adults, mean 84.8 (SD 10.8). All domains achieved mean scores above 70%, with no domains falling below 80% in any of the target populations (Table 5). The distribution of comprehension levels is presented in Table 6.

Table . Mean scores and SDs of each domain by study case (adapted QuIC[a] part A).

Domain	Minors, mean (SD)	Pregnant women, mean (SD)	Adults, mean (SD)
Nature/purpose of research	84.3 (16.2)	85.8 (17.3)	85.7 (15.3)
Risks and benefits	85.4 (13.2)	70.5 (21.4)	82.8 (19.3)
Alternative procedures	N/A[b]	90.2 (24.9)	77.0 (23.4)
Legal, participant protection	N/A	78.9 (19.9)	87.3 (15.6)
Compensation	N/A	84.0 (20.3)	82.7 (21.3)
Contacts, information	N/A	80.8 (16.4)	79.2 (20.7)
Participants' rights	88.8 (23.2)[c]	90.0 (14.7)	92.3 (13.0)
Final score for QuIC part A	83.3 (13.5)	82.2 (11.0)	84.8 (10.8)

[a]QuIC: Quality of the Informed Consent questionnaire.

[b]N/A: not applicable.

[c]In the minors' survey "Contacts, information" and "Participants' rights" were grouped into a single domain, as there was only 1 question on "Contacts, information" (A12), which made it very weak for analysis.

Table . Distribution of participants in each level of comprehension by case study (N=1757).

Degree of comprehension (part A)	Minors (n=620), n (%)	Pregnant (n=312), n (%)	Adults (n=825), n (%)
High (≥90%)	259 (41.8)	91 (29.2)	322 (39.0)
Adequate (≥80% and <90%)	109 (17.6)	100 (32.1)	287 (34.8)
Moderate (≥70% and <80%)	153 (24.7)	80 (25.6)	117 (14.2)
Low (<70%)	99 (16.0)	41 (13.1)	99 (12.0)

XSL•FO
RenderX

Factors That Have an Influence on the Understanding

Regression analyses identified several predictors of comprehension: (1) in minors, lower scores were associated with being male, aged 13, non-Spanish, and having prior clinical trial participation (Table 7); (2) among pregnant women, lower scores were linked to previous trial participation and lower education levels, particularly in Romania (Table 8); and (3) for adults, higher scores were observed among women, members of Generation X, Spanish participants, those with higher education, and those without prior trial participation (Table 9).

Table . Association between adapted QuIC[a] part A and covariates: minors.

Variables	Univariable model, β coefficient (95% CI); *P* value	Multivariable model, β coefficient (95% CI); *P* value
Gender		
Boy	N/A[b]	N/A
Girl	0.32 (0.18 to 0.47); <.001	0.36 (0.22 to 0.50); <.001
Age		
12 years old	N/A	N/A
13 years old	−0.17 (−0.32 to −0.02); .02	−0.20 (−0.34 to −0.06); .006
Which country do you live in?/Previous participation in a clinical trial		
Spain/No	N/A	N/A
Spain/Yes	−0.52 (−0.95 to −0.10); .01	−0.46 (−0.87 to −0.04); .03
Romania/No	−0.18 (−0.38 to 0.01); .06	−0.26 (−0.45 to −0.07); .009
Romania/Yes	−1.71 (−2.40 to −1.03); <.001	−1.77 (−2.43 to −1.10); <.001
United Kingdom/No	−0.25 (−0.42 to −0.08); .004	−0.29 (−0.45 to −0.12); .001
United Kingdom/Yes	−1.37 (−2.05 to −0.69); <.001	−1.34 (−2.01 to −0.67); <.001

[a]QuIC: Quality of the Informed Consent questionnaire.

[b]N/A: not applicable.

Table . Association between adapted QuIC[a] part A and covariates: pregnant women.

Variables	Univariable model, β coefficient (95% CI); *P* value	Multivariable model, β coefficient (95% CI); *P* value
Generation		
Millennials	N/A[b]	N/A
Generation X	0.25 (−0.09 to 0.59); .15	0.21 (−0.12 to 0.54); .22
Previous participation in a clinical trial		
No	N/A	N/A
Yes	−0.71 (−1.00 to −0.41); <.001	−0.68 (−0.97 to −0.39); <.001
Which country do you live in? (Educational level)		
Spain (none completed/primary school/high school)	N/A	N/A
Spain (university or higher education)	0.06 (−0.26 to 0.38); .72	0.03 (−0.28 to 0.34); .87
United Kingdom (none completed/primary school/high school)	−0.16 (−0.69 to 0.38); .56	−0.26 (−0.78 to 0.26); .33
United Kingdom (university or higher education)	−0.03 (−0.38 to 0.32); .86	−0.02 (−0.36 to 0.32); .91
Romania (none completed/primary school/high school)	−1.15 (−1.81 to −0.49); .001	−1.05 (−1.69 to −0.41); .001
Romania (university or higher education)	0.12 (−0.26 to 0.50); .53	0.20 (−0.17 to 0.56); .30

[a]QuIC: Quality of the Informed Consent questionnaire.

[b]N/A: not applicable.

XSL•FO
RenderX

Table . Association between adapted QuIC[a] part A and covariates: adults.

Factors	Univariable model, β coefficient (95% CI); *P* value	Multivariable model, β coefficient (95% CI); *P* value
Gender		
Male	N/A[b]	N/A
Female	0.18 (0.04 to 0.31); .01	0.16 (0.03 to 0.29); .02
Generation		
Millennials	N/A	N/A
Generation X	0.25 (0.12 to 0.38); <.001	0.26 (0.12 to 0.39); <.001
Which country do you live in?		
Spain	N/A	N/A
United Kingdom	−0.19 (−0.34 to −0.04); .02	−0.27 (−0.43 to −0.11); .001
Romania	−0.27 (−0.45 to −0.09); .003	−0.25 (−0.43 to −0.07); .005
Educational level		
None completed/primary school/high school	N/A	N/A
University or higher education	0.15 (0 to 0.30); .04	0.22 (0.07 to 0.38); .004
Previous participation in a clinical trial		
No	N/A	N/A
Yes	−0.58 (−0.84 to −0.33); <.001	−0.47 (−0.73 to −0.21); <.001

[a]QuIC: Quality of the Informed Consent questionnaire.

[b]N/A: not applicable.

Materials cocreated in Spain were effective across all countries, although comprehension was generally higher in the original target population. Lower educational levels in Romania were associated with reduced comprehension.

Subjective Comprehension (Part B)

Comprehension Outcomes by Population Group

This was measured using a 5-point Likert scale ranging from 1 (“I did not understand anything”) to 5 (“I understood everything”), and a final score was calculated. Subjective comprehension was high across all groups, with mean scores exceeding 85 out of 100 (Table 10).

XSL•FO
RenderX

Table . Mean scores and SDs by study case (adapted QuIC[a] part B).

Study case	Minors, mean (SD)	Pregnant women, mean (SD)	Adults, mean (SD)
B1	4.31 (0.81)	4.56 (0.72)	4.60 (0.72)
B2	4.68 (0.66)	4.73 (0.62)	4.81 (0.53)
B3	4.38 (0.87)	4.50 (0.82)	4.50 (0.83)
B4	4.42 (0.79)	4.57 (0.70)	4.50 (0.79)
B5	4.49 (0.79)	4.44 (0.90)	4.42 (0.85)
B6	4.16 (0.98)	4.35 (0.81)	4.54 (0.78)
B7	4.39 (0.78)	4.33 (0.86)	4.43 (0.83)
B8	4.07 (1.07)	4.29 (0.91)	4.47 (0.84)
B9	4.54 (0.72)	4.58 (0.67)	4.61 (0.70)
B10	4.66 (0.66)	4.51 (0.72)	4.46 (0.86)
B11	N/A[b]	4.51 (0.81)	4.57 (0.75)
B12	N/A	4.77 (0.54)	4.76 (0.58)
B13	N/A	4.56 (0.63)	4.56 (0.65)
Total score (QuIC part B)	85.2 (14.1)	87.9 (13.7)	88.9 (13.2)

[a]QuIC: Quality of the Informed Consent questionnaire.

[b]N/A: not applicable.

Preferences and Satisfaction

Minors and pregnant women preferred video, while adults favored text (Table 11). Satisfaction was high, with over 90% of participants in all groups (minors, 604/620, 97.4%; pregnant women, 303/312, 97.1%; and adults, 804/825, 97.5%) reporting that the materials were easy to understand and helpful (Table 12).

Table . Preferred information formats by target population.

How would you prefer to be given the information if you were to participate in a clinical trial? (You may choose various answers)	Minors (n=620), n (%)	Pregnant women (n=312), n (%)	Adults (n=825), n (%)
Written text on paper			
No	438 (70.6)	166 (53.2)	373 (45.2)
Yes	182 (29.4)	146 (46.8)	452 (54.8)
Web			
No	385 (62.1)	210 (67.3)	397 (48.1)
Yes	235 (37.9)	102 (32.7)	428 (51.9)
Video			
No	238 (38.4)	160 (51.3)	586 (71.0)
Yes	382 (61.6)	152 (48.7)	239 (29.0)
To be told by someone in charge of the trial			
No	488 (78.7)	213 (68.3)	457 (55.4)
Yes	132 (21.3)	99 (31.7)	368 (44.6)

XSL•FO
RenderX

Table . Satisfaction and perceived difficulty of information by the target population.

Question	Minors (n=620), n (%)	Pregnant (n=312), n (%)	Adults (n=825), n (%)
What is your overall satisfaction with the information you have read/seen?			
Not at all satisfied/not happy at all[a]	16 (2.6)	9 (2.9)	21 (2.5)
Satisfied/happy[a]	358 (57.7)	188 (60.3)	466 (56.5)
Very satisfied/very happy[a]	246 (39.7)	115 (36.9)	338 (41.0)
Has the information you have read/seen helped you to understand the clinical trial?			
No	16 (2.6)	5 (1.6)	8 (1.0)
Somewhat/a bit[a]	98 (15.8)	17 (5.4)	40 (4.8)
Yes	506 (81.6)	290 (92.9)	777 (94.2)
What is your impression of the information you have read/seen in terms of understanding?			
Very difficult	9 (1.5)	6 (1.9)	7 (0.8)
Difficult	44 (7.1)	23 (7.4)	64 (7.8)
Easy	410 (66.1)	201 (64.4)	526 (63.8)
Very easy	157 (25.3)	82 (26.3)	228 (27.6)
Have you felt the need to ask questions about the clinical trial to the medical professional who appeared in the video/informational materials?			
No	N/A[b]	189 (60.6)	604 (73.2)
Yes	N/A	123 (39.4)	221 (26.8)

[a]Wording used in the minor's survey.

[b]N/A: not applicable.

Discussion

Principal Findings

This exploratory, cross-sectional study provides preliminary evidence that eIC materials developed according to the i-CONSENT guidelines may be associated with high levels of comprehension and satisfaction among minors, pregnant women, and adults in 3 European countries. Objective comprehension scores (part A) exceeded 80% in all groups, and subjective comprehension (part B) was similarly high. However, a discrepancy was observed between subjective and objective understanding, consistent with previous research using QuIC-based tools [6,12-16]. This suggests that self-perceived understanding may not always reflect actual comprehension, underscoring the importance of ongoing assessment and targeted improvements in consent materials.

Minors exhibited greater variability in comprehension scores, with a higher proportion of both high (≥90%) and low (<70%) results compared with other groups. This finding highlights the challenges in tailoring materials for younger populations and suggests that additional support may be needed to facilitate their understanding. Certain domains, such as randomization and placebo concepts, remained difficult to understand for all groups, consistent with prior studies [1,14,17].

Notably, some survey questions yielded poor comprehension rates, particularly those requiring a "disagree" response or containing negative phrasing, consistent with findings from the original QuIC questionnaire [5]. Negatively worded items are known to increase cognitive load and reduce response accuracy [18]. Although these questions aimed to assess understanding of critical concepts such as randomization and placebo, their structure may have contributed to the observed difficulties. In the case of minors, the questionnaire was cocreated with children to ensure suitability [4], making it difficult to determine whether the phrasing of the questions or the complexity of the concepts was the primary factor influencing comprehension.

Formats designed using a layered approach allowed participants to engage with content according to their needs and preferences, which may have contributed to the overall high satisfaction and comprehension rates. Format preferences varied by population: minors and pregnant women tended to prefer video, while adults favored written text. These findings are consistent with previous research suggesting that offering multiple formats can accommodate diverse learning styles and improve engagement [19,20]. The separation of information delivery from discussion with the investigator was reinforced, aligning with literature indicating that combining both methods—especially when accompanied by extended discussion—can further enhance understanding [21].

Satisfaction with the materials was high across all groups, with 1604 out of 1757 (91.29%) participants rating the materials as easy or very easy to understand, and fewer than 3% (46/1757, 2.62%) expressing dissatisfaction. These results suggest that the cocreation process, which actively involved representatives from each target population, contributed to the development of user-friendly materials tailored to diverse needs [3,4,19]. The integration of multimedia tools, including layered content presentation and interactive features, likely supported engagement and comprehension [19], although the impact of

XSL•FO
RenderX

these features on actual decision-making remains to be further explored. It is also important to consider the potential limitations of digital resources, particularly for populations with limited access to technology or low digital literacy, as these factors could influence both comprehension and satisfaction outcomes [22,23].

Comparison With Previous Work

The results of this exploratory study are consistent with previous research indicating that structured, multimodal information materials can enhance participant comprehension and satisfaction in the IC process. Compared with the original QuIC study by Joffe et al [5], which reported a mean objective score of 77.8, our study achieved higher mean scores across all target populations (ranging from 82 to 84.8), while subjective comprehension scores remained similar (mean 85.2 - 88.9). This suggests that the application of the i-CONSENT guidelines, particularly the use of layered and cocreated materials, may contribute to improved understanding without inflating self-perceived comprehension.

The layered approach and the inclusion of multiple formats (text, video, infographics) align with recommendations from the i-CONSENT guidelines and recent systematic reviews, which highlight the importance of tailoring information to participant needs and preferences to address health literacy barriers and support informed decision-making [3,4,19]. Our findings reinforce the value of cocreation, as actively involving representatives from the target population in the development of materials has been shown to increase relevance, accessibility, and user satisfaction [4]. This is consistent with the i-CONSENT project's emphasis on participatory design and the RAND/UCLA (University of California, Los Angeles) expert consensus, which identified cocreation and a layered approach as key strategies for improving IC [3].

Demographic factors such as gender, age, and education were significant predictors of comprehension, consistent with previous literature. Women consistently achieved higher comprehension scores than men, a trend also observed by Raich et al [24] and Tam et al [1]. Age-related differences were apparent: older adults generally outperformed younger adults, in line with findings by Klima et al [25] and Tam et al [1]. By contrast, among minors, older participants scored lower than their younger peers, suggesting that factors such as motivation or engagement may influence comprehension in this subgroup and warrant further investigation [26].

A notable and unexpected finding was the negative association between prior clinical trial participation and comprehension. While some previous studies have suggested that experience in clinical trials enhances understanding [22,25], our results indicate that returning participants may overlook critical information due to perceived familiarity, potentially leading to overconfidence and knowledge gaps. Given the relatively small proportion of prior participants in our sample (adults, 60/825, 7.3%; pregnant women, 49/312, 15.7%; and minors, 33/620, 5.3%), this trend should be interpreted with caution. Future research should explore whether repeated exposure to trial materials influences learning patterns or whether targeted interventions can address overconfidence in returning participants.

Socioeconomic and cultural factors also played a role in comprehension outcomes. Higher education attainment was associated with better understanding, as reported in previous studies [1,16,22,27], yet participants with lower education still achieved reasonable scores, suggesting that the design of the materials may have mitigated some barriers related to health literacy. Comprehension remained high (>80%) in translated materials, consistent with Addissie et al [22], indicating that cocreated materials developed in one country can be effectively utilized in others. However, the observed differences in comprehension between Spain and other countries suggest that further adaptation to local linguistic and cultural contexts may be beneficial to optimize understanding across all populations.

Format preferences also influenced comprehension. The preference for video formats among minors and pregnant women aligns with Gesualdo et al [19], while adults' preference for text-based materials is consistent with health literacy principles described by Lorenzen et al [28]. These findings underscore the importance of offering multiple formats to accommodate diverse learning styles and ensure accessibility for all participants.

Strengths and Limitations

Overview

This section presents the main strengths that support the validity and applicability of the findings, as well as the limitations that should be taken into account to interpret the results and guide future research.

Strengths

One of the key strengths of this study is its multinational design, which evaluated eIC materials across 3 distinct cultural contexts—Spain, the United Kingdom, and Romania. This approach enhances the external validity of the findings and demonstrates the potential applicability of the i-CONSENT guidelines in diverse settings. The large sample size (1757 participants) and the inclusion of 3 target populations—minors, pregnant women, and adults—allowed for robust subgroup analyses and increased the generalizability of the results.

The cocreation process, which actively involved representatives from each target population, is another notable strength. This participatory methodology ensured that the materials were tailored to real-world needs and preferences, likely contributing to the high satisfaction and comprehension rates observed [3]. The use of multiple content formats (layered web content, videos, printable documents, and infographics) provided participants with flexibility in how they accessed information, accommodating different learning styles and potentially reducing barriers related to health literacy [19].

The study also benefited from the use of adapted versions of the QuIC questionnaire, which enabled a comprehensive assessment of both objective and subjective understanding. This dual assessment helped identify specific comprehension challenges, allowing for targeted recommendations for improvement [5]. For minors, the survey was codeveloped with

XSL•FO
RenderX

children, further enhancing its relevance and appropriateness for this group [4].

Limitations

Several limitations should be considered when interpreting these findings. First, the study used a cross-sectional design without a control group, which limits the ability to attribute observed outcomes directly to the i-CONSENT guidelines.

Second, the materials tested represent only 3 specific examples of eIC products developed according to the i-CONSENT guidelines. While these materials were tailored to the needs of minors, pregnant women, and adults, they do not encompass the full range of possible designs or clinical contexts. Additionally, participant engagement with specific formats was not monitored, so the impact of each format on comprehension could not be directly assessed.

Third, recruitment was conducted primarily online due to COVID-19 restrictions, which may have excluded individuals with limited digital access or lower digital literacy. Although printable documents were offered as an alternative format to mitigate this limitation, the sample may not fully represent populations with the greatest barriers to digital participation.

Fourth, while professional translators prepared the English and Romanian versions of the materials, end-user validation of translations was not performed. Although the translations were reviewed by native speakers within the consortium, the lack of direct feedback from target users may have affected comprehension, particularly in non-Spanish populations. Prior studies have highlighted that lower education levels can negatively impact comprehension of IC materials [22], and this effect has been particularly noted among Romanian participants [23].

Finally, the proportion of participants with prior clinical trial experience was relatively small, which may limit the interpretation of findings related to this subgroup. The exploratory nature of the study and the specific sample characteristics further limit the generalizability of the results.

Future Directions

Building on the exploratory findings of this study, several avenues for future research are warranted to further advance the development and implementation of IC materials tailored to diverse populations.

First, additional research should investigate the underlying causes of the polarized comprehension scores among minors and the negative association between prior clinical trial participation and understanding. Qualitative studies or mixed-method approaches could help elucidate whether factors such as motivation, engagement, or overconfidence contribute to these trends. Interventions designed to address these factors, such as interactive modules, adaptive content, or targeted reminders, should be evaluated for their effectiveness in improving comprehension, particularly among returning participants.

Second, the impact of multimedia formats on comprehension and satisfaction merits further exploration. While video content was preferred by minors and pregnant women, its overall effectiveness in improving comprehension remains unclear [19]. Future studies could apply experimental designs, user engagement analytics, or eye-tracking methodologies to assess how participants interact with various formats and which features most effectively support comprehension and retention. Additionally, the role of specific multimedia elements, such as animations or interactive elements, should be systematically evaluated.

Third, the generalizability of cocreated materials across broader cultural and linguistic contexts requires further validation. Although this study demonstrated that materials cocreated in one country can be applied in others, local adaptation and end-user validation may be necessary to optimize comprehension, especially in populations with lower educational attainment or limited health literacy. Comparative studies across additional countries and languages, as well as research on the impact of different translation and adaptation strategies, would provide valuable insights.

Finally, longitudinal research is needed to assess whether improved comprehension at the time of consent translates into better retention of information and more informed decision-making over the course of clinical trial participation. Such studies could also explore the long-term effects of digital consent tools on participant engagement, satisfaction, and trust in research.

Conclusions

This exploratory study provides preliminary evidence supporting the effectiveness of eIC materials developed according to the i-CONSENT guidelines in enhancing comprehension and satisfaction among diverse populations. Across 3 countries—Spain, the United Kingdom, and Romania—and 3 target groups—minors, pregnant women, and adults—objective comprehension scores consistently exceeded 80%, while subjective comprehension and satisfaction rates surpassed 90%. These findings highlight the potential of cocreated, multimodal consent materials to address diverse participant needs and preferences.

Demographic factors such as gender, age, and educational attainment significantly influenced comprehension outcomes. Women consistently outperformed men, and older adults achieved higher scores than younger ones. However, the negative association between prior clinical trial participation and comprehension underscores the need for targeted strategies to engage returning participants and address potential overconfidence or reduced attention to consent materials. These findings call for further investigation into how prior experience shapes participant behavior during the IC process.

The study also demonstrated that materials cocreated in one cultural context can be effectively applied in others, provided that cultural and linguistic adaptations are carefully implemented. However, differences in comprehension scores between countries suggest that additional localization efforts may further optimize outcomes. Format preferences varied across populations, reinforcing the importance of offering

XSL•FO
RenderX

multiple formats to accommodate different learning styles and align with cognitive needs.

While these results are promising, they should be interpreted with caution due to the study's cross-sectional design and exploratory nature. The absence of a control group limits causal inference, and the findings may not be generalizable beyond the specific populations and settings studied. Future research should employ longitudinal designs to assess comprehension retention over time and explore whether improved understanding translates into more informed decision-making throughout clinical trial participation.

These results highlight the potential of user-centered design principles, such as cocreation, multimodal delivery, and cultural adaptation, to improve IC processes in clinical research. The present findings provide a foundation for further development and validation of digital consent tools across broader populations and contexts. Continued efforts to refine these approaches will be essential to ensuring that IC processes remain inclusive, effective, and responsive to participant needs.

Acknowledgments

We thank the participants in the study, Gfk for conducting the fieldwork, and the entire i-CONSENT Consortium for their support and contribution throughout the project. This research was funded by Horizon 2020 SwafS-17-2016 GA n.741856 and conducted in the framework of the European project Improving the Guidelines for Informed Consent, including vulnerable populations, under a gender perspective (i-CONSENT).

This research was funded by Horizon 2020 SwafS-17-2016 GA n.741856 and conducted in the framework of the European project Improving the Guidelines for Informed Consent, including vulnerable populations, under a gender perspective (i-CONSENT).

Disclaimer

During the preparation of this work, the authors used ChatGPT (OpenAI, Inc) and Perplexity (Perplexity AI, Inc) to enhance the writing quality and improve the expression of the text in English. After using these tools, the author(s) reviewed and edited the content as needed and take full responsibility for the content of the publication.

This publication reflects only the authors' view. The Research Executive Agency (REA) is not responsible for any use that may be made of the information that it contains.

Data Availability

The datasets used or analyzed during this study are available online [29].

Authors' Contributions

Conceptualization: JFM, JDD
Data curation: CVH, JFM
Formal analysis: JFM, CMG, CVH
Funding acquisition: JDD
Investigation: JFM
Methodology: JFM, CFA, JDD
Supervision: JFM
Visualization: JFM, CMG
Writing–original draft: JFM
Writing–review & editing: JFM, CMG, CFA, CVH, JDD
All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study questionnaires and additional documentation.
[DOCX File, 54 KB - humanfactors_v12i1e65569_app1.docx]

Checklist 1
CHERRIES (Checklist for Reporting Results of Internet E-Surveys).
[DOCX File, 25 KB - humanfactors_v12i1e65569_app2.docx]

References

XSL•FO
RenderX

1. Tam NT, Huy NT, Thoa LTB, et al. Participants' understanding of informed consent in clinical trials over three decades: systematic review and meta-analysis. Bull World Health Organ 2015 Mar 1;93(3):186-98H. [doi: 10.2471/BLT.14.141390] [Medline: 25883410]
2. Fons-Martinez J, Ferrer-Albero C, Diez-Domingo J. Keys to improving the informed CONSENT process in research: Highlights of the i-CONSENT project. Health Expect 2022 Aug;25(4):1183-1185. [doi: 10.1111/hex.13427] [Medline: 35894807]
3. Fons-Martinez J, Ferrer-Albero C, Diez-Domingo J. Assessment of the appropriateness of the i-CONSENT guidelines recommendations for improving understanding of the informed CONSENT process in clinical studies. BMC Med Ethics 2021 Oct 13;22(1):138. [doi: 10.1186/s12910-021-00708-1] [Medline: 34645425]
4. Fons-Martinez J, Ferrer-Albero C, Diez-Domingo J. Co-creation of information materials within the assent process: from theory to practice. Health Expect 2023 Feb;26(1):429-439. [doi: 10.1111/hex.13675] [Medline: 36416386]
5. Joffe S, Cook EF, Cleary PD, Clark JW, Weeks JC. Quality of informed consent: a new measure of understanding among research subjects. J Natl Cancer Inst 2001 Jan 17;93(2):139-147. [doi: 10.1093/jnci/93.2.139] [Medline: 11208884]
6. Paris A, Cornu C, Auquier P, et al. French adaptation and preliminary validation of a questionnaire to evaluate understanding of informed consent documents in phase I biomedical research. Fundamemntal Clinical Pharma 2006 Feb;20(1):97-104. [doi: 10.1111/j.1472-8206.2005.00391.x]
7. Kerlan-Paris A. Qualité de l'information écrite aux personnes se prêtant à une recherche biomédicale en France. Etudes QuIPs. HAL Thèses. 2008. URL: https://theses.hal.science/tel-00333771 [accessed 2024-08-14]
8. Tattersall MHN, Jefford M, Martin A, et al. Parallel multicentre randomised trial of a clinical trial question prompt list in patients considering participation in phase 3 cancer treatment trials. BMJ Open 2017 Mar 1;7(3):e012666. [doi: 10.1136/bmjopen-2016-012666] [Medline: 28249847]
9. Bergenmar M, Molin C, Wilking N, Brandberg Y. Knowledge and understanding among cancer patients consenting to participate in clinical trials. Eur J Cancer 2008 Nov;44(17):2627-2633. [doi: 10.1016/j.ejca.2008.08.013] [Medline: 18818068]
10. Ferrer Albero C. Diseño de un documento de asentimiento informado en formato comic para los menores que participan en un ensayo clínico. Dialnet. 2017. URL: https://dialnet.unirioja.es/servlet/tesis?codigo=190201 [accessed 2024-10-08]
11. Peterson RA, Cavanaugh JE. Ordered quantile normalization: a semiparametric transformation built for the cross-validation era. J Appl Stat 2020;47(13-15):2312-2327. [doi: 10.1080/02664763.2019.1630372] [Medline: 35707424]
12. Hoyos M, Villamañán-Bueno E, Uzquiano E, Gómez-Salcedo P, Río-Durango M, Frías-Iniesta J. Informed consent process in clinical trials: development of a patient-reported questionnaire. Farm Hosp 2020;44(6):254-271. [doi: 10.7399/fh.11430]
13. Atal S, Dunne F. Assessment of the understanding of informed consent including participants' experiences, and generation of a supplemental consent decision aid for gestational diabetes mellitus (GDM) research. HRB Open Res 2018;1:12. [doi: 10.12688/hrbopenres.12811.1] [Medline: 32002505]
14. Shiono YN, Zheng YF, Kikuya M, et al. Participants' understanding of a randomized controlled trial (RCT) through informed consent procedures in the RCT for breast cancer screening, J-START. Trials 2014 Sep 25;15(1):375. [doi: 10.1186/1745-6215-15-375] [Medline: 25257667]
15. Joffe S, Cook EF, Cleary PD, Clark JW, Weeks JC. Quality of informed consent in cancer clinical trials: a cross-sectional survey. The Lancet 2001 Nov;358(9295):1772-1777. [doi: 10.1016/S0140-6736(01)06805-2]
16. Paris A, Deygas B, Cornu C, et al. Improved informed consent documents for biomedical research do not increase patients' understanding but reduce enrolment: a study in real settings. Br J Clin Pharmacol 2015 Nov;80(5):1010-1020. [doi: 10.1111/bcp.12716] [Medline: 26147763]
17. Wu C, Wang N, Wang Q. Participants' understanding of informed consent in clinical trials: a systematic review and updated meta-analysis. PLoS One 2024;19(1):e0295784. [doi: 10.1371/journal.pone.0295784]
18. Marsh HW. Positive and negative global self-esteem: a substantively meaningful distinction or artifactors? J Pers Soc Psychol 1996 Apr;70(4):810-819. [doi: 10.1037//0022-3514.70.4.810] [Medline: 8636900]
19. Gesualdo F, Daverio M, Palazzani L, et al. Digital tools in the informed consent process: a systematic review. BMC Med Ethics 2021 Feb 27;22(1):18. [doi: 10.1186/s12910-021-00585-8] [Medline: 33639926]
20. Chapman N, McWhirter R, Armstrong MK, et al. Self-directed multimedia process for delivering participant informed consent. BMJ Open 2020 Jul 26;10(7):e036977. [doi: 10.1136/bmjopen-2020-036977] [Medline: 32713850]
21. Hereu P, Pérez E, Fuentes I, Vidal X, Suñé P, Arnau JM. Consent in clinical trials: what do patients know? Contemp Clin Trials 2010 Sep;31(5):443-446. [doi: 10.1016/j.cct.2010.05.004] [Medline: 20462521]
22. Addissie A, Abay S, Feleke Y, Newport M, Farsides B, Davey G. Cluster randomized trial assessing the effects of rapid ethical assessment on informed consent comprehension in a low-resource setting. BMC Med Ethics 2016 Jul 12;17(1):40. [doi: 10.1186/s12910-016-0127-z] [Medline: 27406063]
23. Alexa-Stratulat T, Neagu M, Neagu AI, Alexa ID, Ioan BG. Consent for participating in clinical trials - is it really informed? Dev World Bioeth 2018 Sep;18(3):299-306. [doi: 10.1111/dewb.12199] [Medline: 29933502]
24. Raich PC, Plomer KD, Coyne CA. Literacy, comprehension, and informed consent in clinical research. Cancer Invest 2001;19(4):437-445. [doi: 10.1081/cnv-100103137] [Medline: 11405182]

XSL•FO
RenderX

25. Klima J, Fitzgerald-Butt SM, Kelleher KJ, et al. Understanding of informed consent by parents of children enrolled in a genetic biobank. Genet Med 2014 Feb;16(2):141-148. [doi: 10.1038/gim.2013.86] [Medline: 23807615]
26. Bedard K, Dhuey E. The persistence of early childhood maturity: international evidence of long-run age effects. Q J Econ 2006 Nov 1;121(4):1437-1472. [doi: 10.1093/qje/121.4.1437]
27. Flory J, Emanuel E. Interventions to improve research participants' understanding in informed consent for research: a systematic review. JAMA 2004 Oct 6;292(13):1593-1601. [doi: 10.1001/jama.292.13.1593] [Medline: 15467062]
28. Lorenzen B, Melby CE, Earles B. Using principles of health literacy to enhance the informed consent process. AORN J 2008 Jul;88(1):23-29. [doi: 10.1016/j.aorn.2008.03.001] [Medline: 18603032]
29. Fons-Martinez J. Dataset: assessing understanding and satisfaction with electronic consent materials tailored to 3 target populations (pregnant women, minors and adult population). Zenodo. URL: https://zenodo.org/records/15144013 [accessed 2025-07-08]

Abbreviations

CCAsIn: Cuestionario de Comprensión del Asentimiento Informado
CHERRIES: Checklist for Reporting Results of Internet E-Surveys
eIC: electronic informed consent
FISABIO: Foundation for the Promotion of Health and Biomedical Research of the Valencia Region
IC: informed consent
PIAAC: Program for the International Assessment of Adult Competencies
PISA: Programme for International Student Assessment
QuIC: Quality of the Informed Consent questionnaire
UCLA: University of California, Los Angeles

Edited by A Kushniruk; submitted 19.08.24; peer-reviewed by A Azizi, H Vanaken; revised version received 23.04.25; accepted 12.05.25; published 15.08.25.

Please cite as:
Fons-Martinez J, Murciano-Gamborino C, Ferrer-Albero C, Vergara-Hernandez C, Diez-Domingo J
Digital Informed Consent/Assent in Clinical Trials Among Pregnant Women, Minors, and Adults: Multicountry Cross-Sectional Evaluation of Comprehension and Satisfaction
JMIR Hum Factors 2025;12:e65569
URL: https://humanfactors.jmir.org/2025/1/e65569
doi:10.2196/65569

© Jaime Fons-Martinez, Carlos Murciano-Gamborino, Cristina Ferrer-Albero, Carlos Vergara-Hernandez, Javier Diez-Domingo. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 15.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Improving Access to and Delivery of Maternal Health Care Services to Prevent Postpartum Hemorrhage in Selected States in Nigeria: Human-Centered Design Study

Bosun Tijani[1*], PhD; Uchenna Igbokwe[2*], MSc; Temi Filani[1*], MSc; Adefemi Adewemimo[1], MBA; Lola Ameyan[3*], MBA; Martins Iyekekpolor[1], MSc; Steven Karera[1*], MSc; Olatunji Oluyide[1*], MSc; Emmanuela Ezike[1], MSc; Temidayo Akinreni[1], MSc; Obruche Ogefere[2*], BPharm; Victor Adetimilehin[2*], MSc; Valentine Amasiatu[2,4*], PhD; Chukwunonso Nwaokorie[2], MBBS; Naanma Kangkum[2], BPharm; Olufunke Fasawe[3], MSc; Eric Aigbogun Jr[2,5*], PhD

[1]Co-creation Hub Nigeria, Lagos, Nigeria

[2]Solina Centre for International Development and Research, 8 Libreville Crescent, Wuse, Abuja, Nigeria

[3]Clinton Health Access Initiative, Abuja, Nigeria

[4]Department of Anatomy, Alex Ekwueme Federal University Ndufu-Alike, Abakiliki, Nigeria

[5]College of Medicine, Enugu State University of Science and Technology, Enugu, Nigeria

[*]these authors contributed equally

Corresponding Author:
Eric Aigbogun Jr, PhD
Solina Centre for International Development and Research, 8 Libreville Crescent, Wuse, Abuja, Nigeria

Abstract

Background: A significant cause of postpartum hemorrhage (PPH) is access to and delivery of maternal health care services. Several multisectoral strategies have been deployed to address the challenges with little success, thereby necessitating the use of human-centered design (HCD) to enhance health care delivery, particularly in PPH management.

Objective: This study aims to develop facility-level solutions for optimizing uterotonic supply chain systems and health service delivery in PPH management through an HCD approach in selected Nigerian states.

Methods: The research used a four-phase HCD methodology: (1) co-research, (2) co-design, (3) co-refinement, and (4) implementation. However, this paper focused on the first 3 phases. In the co-research phase, 203 interviews were conducted, involving 80 pregnant women and nursing mothers, 97 health care workers, and 26 key stakeholders. Additionally, 33 sites were observed across a 3-level continuum of care. Interviews and focus group discussions revealed insights into the distribution of health workers and observed PPH cases, alongside knowledge and administration of uterotonics. Data analysis was carried out using three key steps: (1) identifying key themes from the collected data, (2) developing insight statements that encapsulate these themes, and (3) translating each insight statement into actionable design opportunities.

Results: About 150 ideas were produced and translated into 12 solution prototypes in the co-design phase. Progressive refinement following feedback from 140 stakeholders led to the selection of three final solutions: (1) implementing a referral linkage system to improve the transportation of pregnant women to nearby health facilities, (2) increasing demand for antenatal care services among pregnant women and their families, and (3) delivering a comprehensive uterotonic logistics management program for streamlined uterotonic storage and management.

Conclusions: This approach aligns with global health trends advocating for HCD integration in health care programming and aims to empower local champions to drive sustainable improvements in maternal health outcomes. Judicious implementation of the developed prototypes across the states can strengthen clinical care and potentially reduce maternal health service delivery gaps.

(JMIR Hum Factors 2025;12:e58577) doi:10.2196/58577

KEYWORDS

human-centered design; postpartum hemorrhage; maternal health; health care services; maternal; health care; service; postpartum; facility; uterotonic; supply chain; Nigeria; interview; focus group; pregnant; female; health care workers; stakeholder; participatory; prevention; treatment; hospital setting; community

XSL•FO
RenderX

Introduction

Maternal mortality occurs due to complications that arise during and after pregnancy and childbirth. The major complications responsible for most maternal deaths include postpartum hemorrhage (PPH), postpartum infections, pre-eclampsia and eclampsia, complications during delivery, and unsafe abortion [1].

Globally, PPH is most pronounced in low-income countries [2], inflicting suffering on women and their families and also causing a strain on local and national health systems. Generally, PPH varies among regions, largely due to demographic and socioeconomic factors such as age, race, and social status, with low-income countries having a 12% higher occurrence [3]. One of the major drivers is limited access to timely, safe, and quality maternal care, which results in inequalities in maternal health services globally. Following a systematic review, 24.5% of maternal deaths in sub-Saharan Africa were attributed to PPH [4]. Over the past two decades, substantial efforts have been made to improve maternal care in sub-Saharan Africa, resulting in a 40% reduction in maternal mortality [5]. However, there are still significant challenges, such as substandard care, poor management skills, lack of knowledge, and delays in transferring women to the next level of care [6].

In Nigeria, PPH is the leading cause of maternal mortality, responsible for at least 21% of maternal deaths [7]. The country also bears a significant burden of maternal deaths, with 28.5% (82,000) of global maternal deaths occurring there annually [8]. Roughly 31% of female deaths in Nigeria are attributed to maternal causes, and in 2018, the country recorded a maternal mortality ratio of 512 maternal deaths per 100,000 live births [9]. Some factors contributing to this sad reality include poor access to essential medicines, including uterotonics for PPH management, and inadequate health care providers [10]. To address this threat, timely dissemination of information, primary preventive measures, comprehensive antenatal care (ANC), and strong political commitment are essential [3]. The key lies in proactive prevention, underscoring the need to analyze PPH comprehensively to develop effective mitigation and treatment strategies. Although several interventions have been carried out to prevent PPH, such as assessing coverage, acceptability, and feasibility of a comprehensive program [11-14], these interventions were health facility-centered and overlooked the broader ecosystem of care. Therefore, it was imperative to introduce the human-centered design (HCD), a collaborative method to enhance clinical care and bridge gaps in maternal health services delivery. This method involves not only health care providers but also other actors integral to providing and receiving care, including policy makers and patients [15,16].

The HCD approach has been applied successfully in global health programming, offering a human-centered approach to entrenched issues [15]. One example is the Support Sisters Intervention in the United States, a community-based peer-support intervention to help Medicaid-insured pregnant women access services and appointments [16]. In light of the global successes of using HCDs, its application to address service delivery complexities associated with maternal health care in Nigeria became imperative.

Hence, this study aimed to describe the use of HCD as an approach to develop facility-level solutions for optimizing uterotonic supply chain systems and health service delivery in managing PPH in selected Nigerian states.

Methods

Study Design and Setting

This study was designed as participatory research using the HCD through co-research, co-design, and co-refinement phases involving desk reviews, interviews, brainstorming, and rapid prototyping. The aim was to identify and develop facility-level solutions to strengthen the appropriate use of uterotonics for PPH prevention and treatment, reduce gaps in maternal health services, and drive improvements in uterotonic use for PPH prevention and treatment in Kano, Lagos, and Niger states in Nigeria through the development and integration of solutions to achieve a safe pre- and post-childbirth delivery experience for women in Nigeria.

This study was guided by the principle that the quality of PPH care spans the entire lifecycle of the uterotonic administration, beginning from production to proper usage within the hospital setting. It recognizes that people play a pivotal role at every stage of this process and emphasizes the importance of understanding the system to facilitate necessary improvements for better outcomes.

HCD Approach

The data collection and synthesis involved incorporating the principles of HCD. This is a user-inclusive approach that has been adopted to address and proffer customized solutions to complex challenges that impact the public health space, as well as suit the beneficiaries of the solutions [16]. It entails a mindset that starts with the people for whom the intervention is designed and ends with new solutions that are tailor-made to suit their needs. This approach disrupts traditional methods where researchers, health care providers, and administrators design new care models based solely on studies and expert opinions, instead integrating health care users' perspectives across all stages of the process [17-19].

The HCD process used in this study included co-research, co-design, and co-refinement phases (Table 1). Each of these phases enabled the discovery and an understanding of the problem and allowed us to co-create and test the developed solutions using iterative processes (Figure 1).

XSL•FO
RenderX

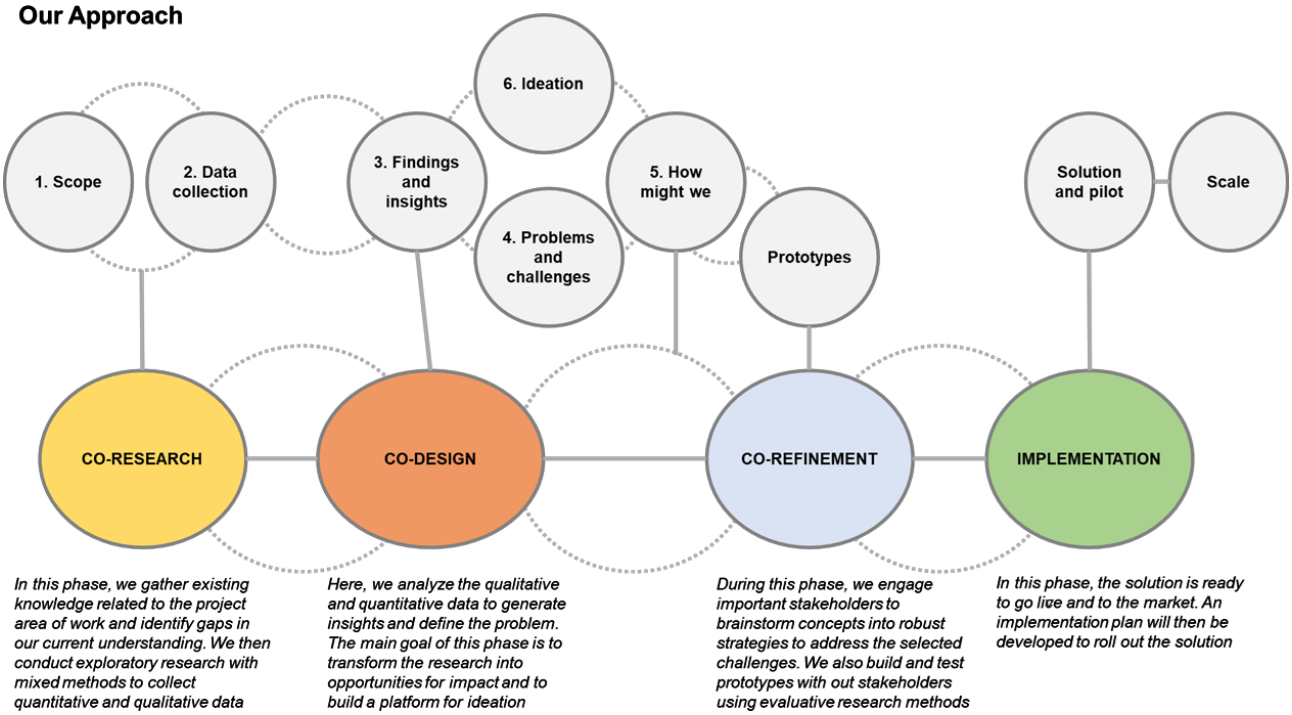
Table . The study's HCD[a] phases, approaches, and rationale.

HCD phases	Participatory approach	Rationale or comment
Co-research	Desk reviews, interviews, and site observations	Understand the PPH[b] landscape in the states, frame research questions, design the data collection tools, and collect data
Co-design	Brainstorming and alignment	Promote creativity, idea generation, and collaborative problem-solving
Co-refinement	Rapid prototyping and streamlining (redoing, discarding, and refining)	Generate tangible results iteratively through early feedback and revision of solutions
Co-implementation	Not part of the study	Recommendations were made to the states on the need to implement the developed prototypes

[a]HCD: human-centered design.

[b]PPH: postpartum hemorrhage.

Figure 1. The study's human-centered design framework.



Co-Research Phase: Understanding the User Needs, Goals and Contexts, and Tool Design (Phase 1)

The co-research phase was conducted across the 3 program states—Kano, Lagos, and Niger—from May 24 to October 24, 2022. It kick-started the HCD process and involved identifying the knowledge gaps around maternal mortality due to PPH through desk reviews to (1) understand the current status of PPH in Kano, Niger, and Lagos states; (2) frame research questions; and (3) design the data collection. The insights gleaned from this phase were subsequently translated into opportunities for impact.

Co-Design Phase: Designing, Piloting, and Testing (Phase 2)

The co-design phase was conducted across the 3 program states from November 7, 2022, to January 10, 2023. The co-design phase aimed at transforming insights that were identified during the co-research phase into new and more structured ideas via the brainstorming process, which encouraged the champions to think expansively.

Co-Refinement Phase: Refining the Prototypes (Phase 3)

The co-refinement phase took about 1 month to achieve between January 2 and February 23, 2023. During the co-refinement phase, each of the context-based solutions was validated through rapid prototyping.

Data Collection

In the co-research phase, 58 champions were trained in data collection across these 3 states. Training on proper interviewing techniques and mock sessions were held across the states. This enabled the government stakeholders from the program State Ministries of Health and the State Primary Health Care Boards, assigned as HCD champions, to review existing knowledge, identify gaps, and define the research methodology needed to deepen understanding of the problem.

XSL•FO
RenderX

The methodology for data collection was selected based on the timeframe, understanding of the local context, and resources available. The following exploratory methods were selected: (1) site visit observation (live-birth observation); (2) interviews with key informants (government officials); and (3) one-on-one interviews with pregnant women, health care providers (ie, nurses, midwives, doctors, and other health professionals), traditional birth attendants (TBAs), or health facility representatives (ie, facility in-charges and administrators).

The data collection tools were developed based on identified gaps from the landscape review and to answer the research question on these core areas: facility profile, knowledge and information on the use of uterotonics, the types and challenges regarding uterotonic use, facility equipment, and uterotonic availability.

Three unique tools were developed to solicit responses from the participants: (1) an observatory tool for the site visit (health facilities); (2) key informant interview guides; and (3) survey tools for health care providers, TBAs, and pregnant women. These tools differed for each group, ensuring the questions were tailored to their various contexts.

A 3-day workshop was organized in each of the 3 program states to gather the data collected from the field activities. This workshop focused on identifying the knowledge gaps gathered, having a clear and robust understanding of the challenges to be addressed, as well as generating insights to be translated into areas for opportunities.

Data Analysis: Synthesis and Intervention Development

The data collected (site visit observation, interviews with key government informants, and one-on-one interviews with pregnant women, health care providers, TBAs, or health facility representatives) were analyzed using a descriptive analysis method to constructively summarize data points.

Data analysis followed three steps in accordance with the HCD methodology: (1) identifying key themes from the collected data, (2) developing “insight statements” that encapsulate these themes, and (3) translating each insight statement into actionable “design opportunities” [16,20]. The synthesis and actionability of the findings from the co-research phase were carried out in the co-design and co-refinement phases.

The co-design phase featured a 2-day workshop in the program states where the HCD champions collectively brainstormed over 60 unique ideas from individual concepts and bundled them into 12 distinct robust prototypes of substance that were best suited to address the challenge. In the co-refinement phase, the solutions (prototypes) were tested with the intended users, and their feedback was integrated. This involved redoing, discarding, and refining the solutions based on a rich understanding of the context, and many potential points of failure that have been uncovered and resolved. Following the testing and feedback activity, the most viable solution was selected, one for each state, and a plan of action was created for implementation on a larger scale. These co-design and co-refinement processes were iterative to ensure the most appropriate models were achieved.

Ethical Considerations

The study was carried out between May 24, 2022, and February 23, 2023, and was conducted in strict compliance with the ethical standards outlined in the Helsinki Declaration, ensuring the well-being of the participants. These standards encompassed obtaining informed consent, guaranteeing confidentiality, and safeguarding participants’ rights. Prior to commencement, the research received its ethical approval from the National Health Research Ethics Committee of Nigeria (NHREC) under the approval number NHREC/01/01/2007-15/03/2022 on March 15, 2022.

Results

Participant Groups

In total, 203 interviews were conducted among 97 health care providers, TBAs, and CHIPS (Community Health Influencers, Promoters and Services) Programme agents; 80 pregnant women or nursing mothers; and 26 government officials, with observations conducted at 33 sites across a 3-level continuum of care. Table 2 presents the participant distribution across the states.

The findings were classified based on the 3 HCD phases used in this research. Following the analysis of the data collected in the field during the co-research phase, the strong findings were subsequently categorized under common themes according to the different participant groups as described in Table 2. The findings were further synthesized into insights that were translated into opportunity areas for impact in the co-research phase.

Table . Distribution of study participants interviewed during the co-research phase.

State	Site observation (n=33), n	Doctors, nurses or midwives, TBAs[a], and CHIPS[b] agents interviewed (n=97), n	Pregnant women or nursing mothers interviewed (n=80), n	Government officials interviewed (n=26), n
Kano State	12	33	28	8
Niger State	11	35	26	10
Lagos State	10	29	26	8

[a]TBAs: traditional birth attendants.

[b]CHIPS: Community Health Influencers, Promoters and Services Programme.

XSL•FO
RenderX

Co-Research: Understanding the User Needs, Goals and Contexts, and Tool Design

All of the activities in the co-research phase ensured a robust and clear understanding of the challenge to be addressed. The key insights from the co-creation phase across the 3 program states included the challenges faced by pregnant women in rural areas concerning affordability, accessibility, and awareness of health care services; health care workers' strong commitment to maternal health despite issues with service efficiency and working conditions; observations indicating inadequacies in equipment and space management within health care facilities; and key informants expressing concerns about the current health care system's ability to deliver adequate maternal care, emphasizing the need for ongoing improvements in service delivery and coordination.

Co-Design: Designing, Piloting, and Testing

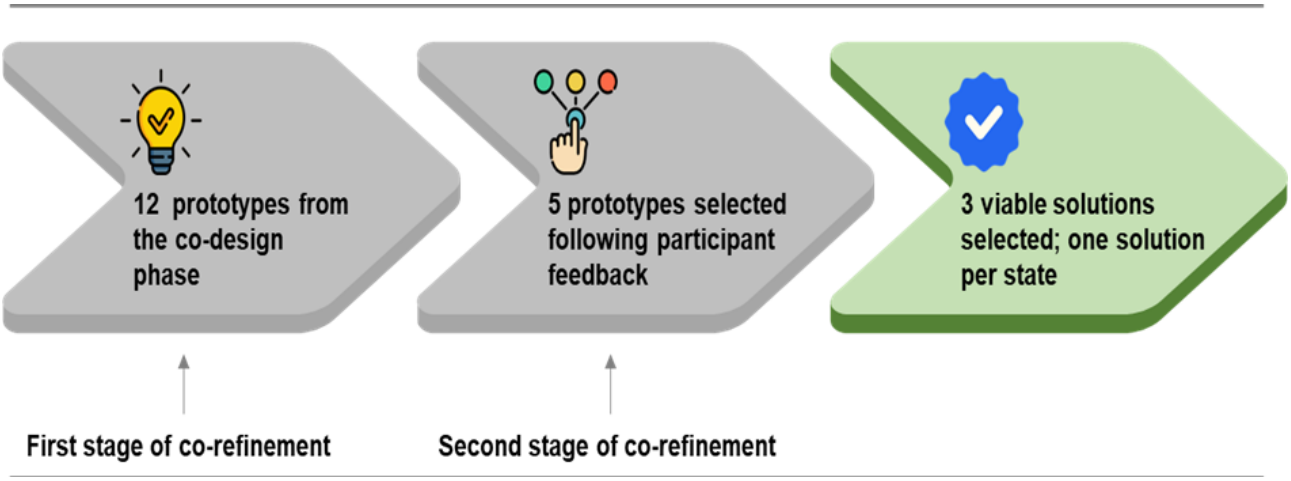
The workshop focused on finding solutions to the challenges from the findings from the co-research phase. These solutions aimed to provide improvements in clinical care and appropriate use of uterotonics for PPH management and treatment.

During the co-design session, ideas and concepts were synthesized, reviewed, and narrowed down to a set of concepts that could be developed for testing. The main components of this phase were framing the design challenge, generating "how might we" questions to proffer solutions to the identified challenge area, brainstorming ideas, creating appropriate storyboards for promising solutions, and conducting rapid prototyping sessions. In this phase, over 100 different ideas were submitted by champions across the various states to resolve the identified challenge statement.

Co-Refinement: Refining the Prototypes

The co-refinement phase comprised 2 stages to contextualize the prototypes developed from the co-creation phase (Figure 2). In the first stage, 12 initial prototypes were carefully evaluated for their suitability to address the state-specific issues and narrowed down to 5 prototypes with the most promising qualities. These 5 prototypes underwent further evaluation in the second co-refinement stage to produce 3 prototypes, one for each state, that best addressed the specific issues related to optimizing uterotonic supply chain systems and health service delivery in managing PPH.

Figure 2. Stages of the co-refinement phase.



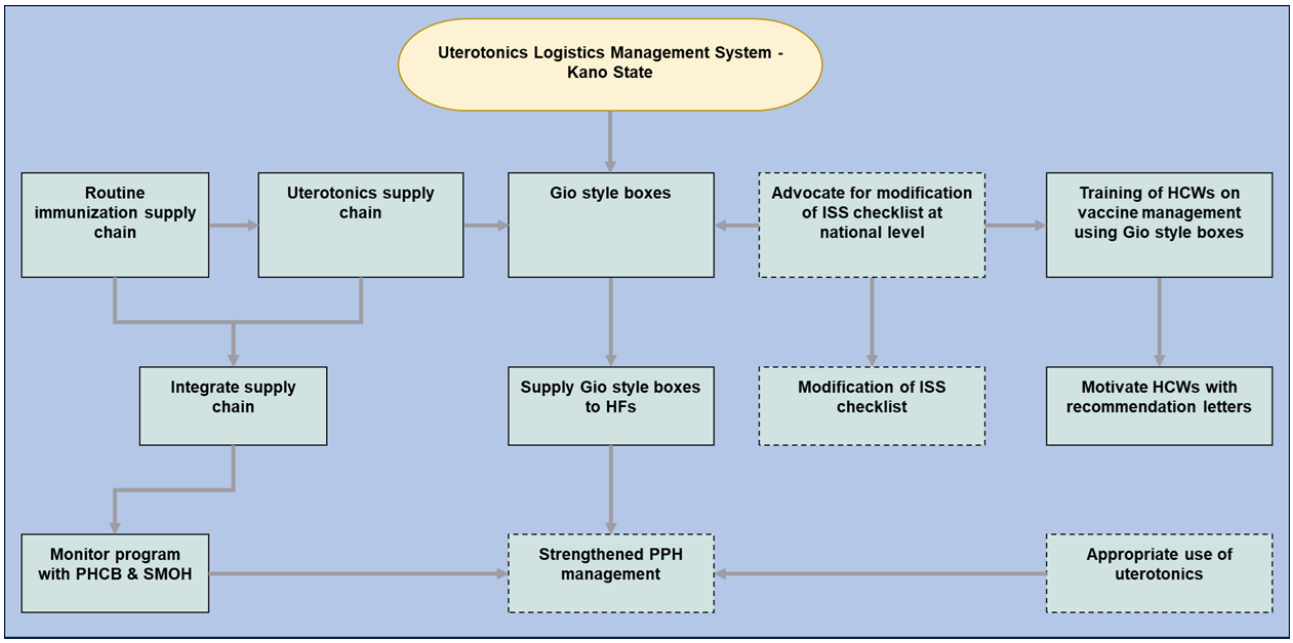
Prototype Profile 1

Challenge Statement 1: How Might We Increase the Storage of Uterotonic Drugs in the Health Facilities in Kano State?

The Uterotonics Logistics Management Program is an integrative program that aims to include uterotonics in the general supply chain and storage for routine vaccinations within the state (Figure 3). This program aims to create an advocacy strategy by creating an alliance within the State Ministry of Health to include uterotonics into the integrative supply list for routine immunization.

XSL•FO
RenderX

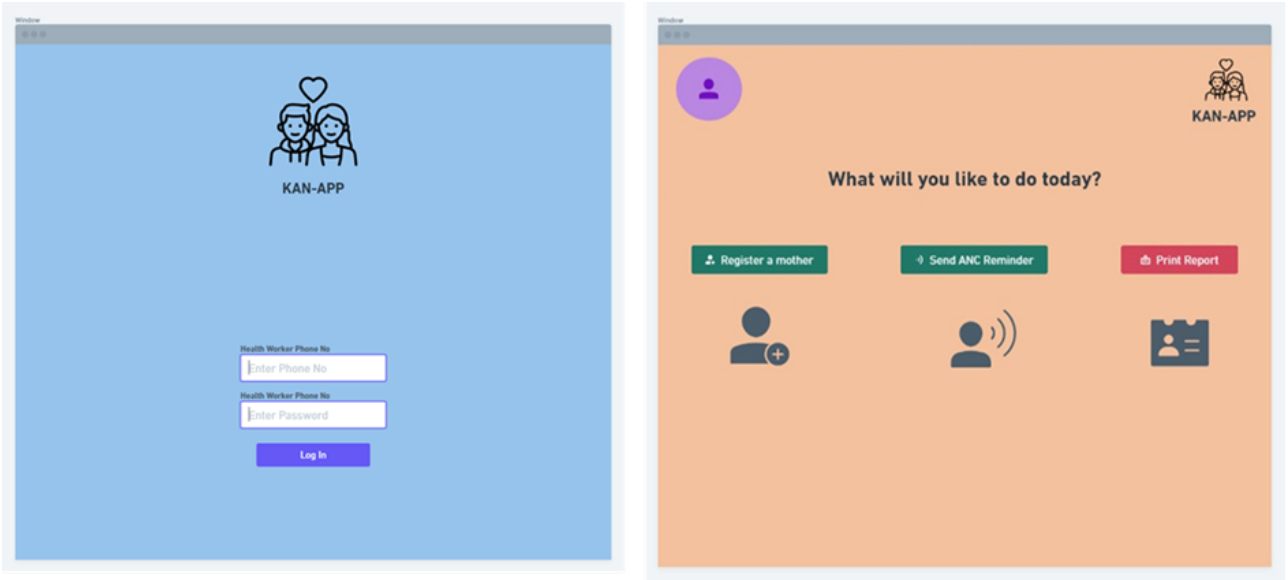
Figure 3. Flowchart of the Uterotonics Logistics Management System. HCW: health care worker; HF: health facility; ISS: integrated supportive supervision; PHCB: Primary Health Care Board; PPH: postpartum hemorrhage; SMOH: State Ministry of Health.



Challenge Statement 2: How Might We Improve Care-Seeking Behavior for Handling PPH Cases Among Women of Childbearing Age in Kano State?

The need to increase the level of awareness about PPH in Kano State is high, particularly due to women's preference for home delivery (Figure 4), a decision that is highly influenced by their spouses. Hence, women of childbearing age may tend to demand health services like ANC following reinforcement from their husbands. The Kan-App program has three core elements for improving care-seeking behavior among women of childbearing age in Kano: (1) health care workers will be able to send reminders to husbands about the ANC visit schedules of their wives; for unmarried women or single mothers, the app will send reminders to their close relatives; (2) distribution of maternity blankets designed with key information on PPH and why women should attend ANC sessions, highlighting the role of the community in identifying the danger signs of pregnancy and the safety of facility-based delivery; and (3) radio programs to educate women on the dangerous signs of pregnancy and delivery.

Figure 4. Kan-App for the increase in the demand for antenatal care services among pregnant women and their families, particularly their spouses.



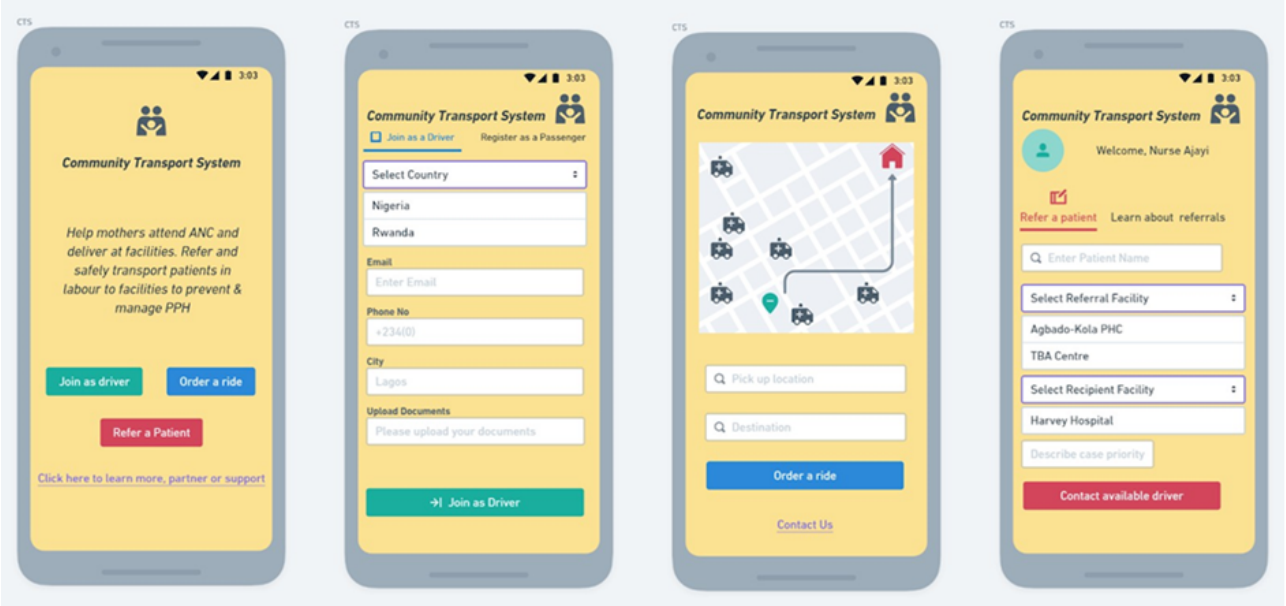
XSL•FO
RenderX

Prototype Profile 2

Challenge Statement 1: How Might We Improve the Transportation and Referral Linkage for the Primary Health Care Facilities?

The community-based transport system is a simplified intervention designed to leverage the already existing network of Keke Napep drivers, Yellow Cab drivers, or the National Union of Road Transport Workers (NURTW), as well as the local security agencies in Lagos State (Figure 5). The State Primary Health Care Board in conjunction with the State Ministry of Health will develop a memorandum of understanding with the NURTW, Keke Napep, and the security agencies. A database of drivers will be developed for the 20 local government areas in Lagos State, and drivers will be informed of the MoU (memorandum of understanding) and the program, allowing them to volunteer their services and provide their contact details. Each primary health care (PHC) facility would be furnished with a list of drivers and emergency contacts, which will be shared with pregnant women as well as TBAs nearby. TBAs can use the drivers to refer women to PHC facilities. The key elements of the program include (1) pre-established communication between the health facilities, drivers, and security agencies; (2) a database of the drivers and unique communication lines; (3) unique terrain and level of care specific to hard-to-reach areas and gridlock areas (Keke Napep and Yellow Cabs); and (4) selection (volunteer), remuneration, and recognition of the drivers.

Figure 5. Community transport system app for improving the transportation and referral linkage for the primary health care facilities.



Prototype Profile 3

Challenge Statement 1: How Might We Increase the Storage of Uterotonic Drugs in the Health Facilities in Niger State?

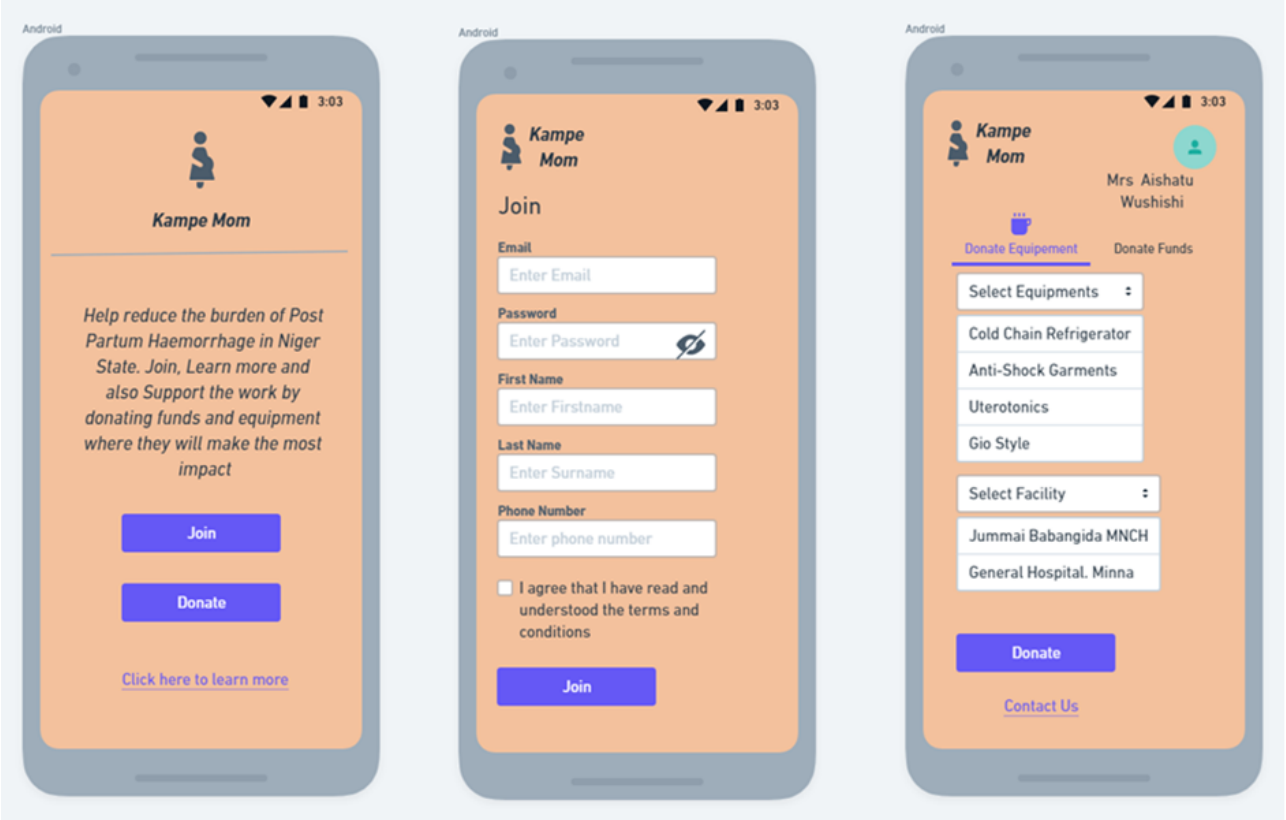
The Niger State Primary Health Care Development Agency (NSPHCDA) Mobile and Electronic Application Platform (Kampe Mom) is an open-source web-based application that visualizes the equipment needs of PHC facilities within the state. This platform will be open to government actors (ministries, departments, and agencies), philanthropists, and organizations who are willing to contribute toward purchasing equipment for PHC facilities based on identified needs (Figure 6). The equipment will be displayed based on the order of priority and needs of the PHC facility at the time. In addition, there would be key events such as charity walks and fundraising events to create awareness of PPH and safe motherhood and generate funds that will feed into the purchase of this equipment for the PHC facilities.

The elements of the solution include the following:

- Registration of 1400 PHC facilities in Niger State on the platform. These PHC facilities will be disaggregated by location, accessibility, patient in-flow, etc.
- Equipment needed in all 25 local government areas will be captured on the platform and open to philanthropists, key decision-makers, legislators, community-based organizations, civil society organizations, women groups, etc.
- Charity walk can be placed on the platform for fundraising.
- Create awareness once a year on May 22, which is the world's safe motherhood day. Progress updates will be provided on programs about maternal and child health.
- The target audience will be pregnant women, fathers, families, policy makers, community and religious leaders and gatekeepers, TBAs/CHIPS agents, other community volunteers, community-based organizations, civil society organizations, etc.

XSL•FO
RenderX

Figure 6. Kampe Mom app for increasing the storage of uterotonic drugs in the health facilities in Niger State.



Discussion

HCD Process and Outcomes

While much of the global effort has focused on a health system–strengthening approach to reducing pregnancy-related deaths in low- and middle-income countries [21,22], less energy has been focused on solutions that take into consideration the major challenges mothers go through during pregnancy, so as to develop a client-centered solution that effectively addresses their needs. This research was aimed at proffering solutions to develop context-specific solutions to enhance service delivery for postpartum care in Lagos, Kano, and Niger states using HCD.

In the co-discovery phase, the study was able to uncover critical challenges faced by both pregnant women and health care workers in rural settings, with a focus on maternal health care access and service delivery. Pregnant women primarily struggle with affordability, accessibility, and awareness. Many of these women are housewives in rural areas, with limited financial resources and logistical difficulties in accessing health care facilities [23]. These findings are consistent with existing literature, which highlights that rural women often face significant socioeconomic barriers, including poverty, transportation challenges, and limited access to health care services [24-26].

Health care workers, on the other hand, are very committed to improving maternal health but face challenges resulting from gaps in care levels between primary and secondary facilities. The strain is particularly evident in facilities with high client volumes and limited human resources [27,28], underscoring the need for better support systems for health care providers in these settings. There were also issues of inadequate health care infrastructure, such as insufficient storage systems and overcrowded delivery wards, which could compromise the quality of care provided during critical moments like prolonged labor [28,29]. There is also the issue of the service delivery system, which requires continuous strengthening to meet the needs of these vulnerable populations. Addressing these issues will require a multifaceted approach, which should include a better system for managing the human resource problems, improved health communication strategy, demand generation, and enhanced coordination of care to ensure that both health care workers and pregnant women receive the support they need to achieve better health outcomes [11].

In this research, the co-design phase was aimed at synthesizing relevant questions to create potential solutions for the identified problems affecting the appropriate use of uterotonics and PPH care, particularly in resource-limited settings. This process highlighted the importance of involving the intended users in the HCD co-design process to ensure that the developed solutions are tailored to the specific contexts and challenges identified [30,31]. For instance, the HCD co-design phase was used to develop health solutions to enhance the comfort, health outcomes, and overall satisfaction and experience of patients [16,32].

More than 100 questions and ideas gathered from the initial co-research phase were ideated in brainstorming sessions to identify the 12 most feasible prototypes. These solutions were narrowed down by selecting the questions that most exhaustively discussed the challenges with service delivery for PPH care, suited the contexts in each state, and considered the most

XSL•FO
RenderX

frequently occurring challenges and issues. This was done to ensure the final prioritized questions would generate solutions that would incite the most impact in each state. The brainstorming approach to co-design was used to further foster creative thinking and ideation, as well as generate comprehensive, user-centric solutions [16,33]. The 12 prototypes were further refined to select the most fitting solutions for the 3 states.

The implementation of brainstorming sessions in the co-design phase aided the co-refinement of the 12 prototypes. The collaborative relationships built during this session fostered the iteration and feedback process, ensuring that the prototypes selected aligned with the needs of the users and were actionable and sustainable [33,34]. The 12 prototypes were narrowed down to 3 final solutions, which were selected based on their operationalizability (how well they worked), feasibility (how possible it is to implement), and specificity (how they matched each state's unique challenges).

Challenges Faced and Lessons Learned

Stakeholder Engagement

The project team had to put mechanisms in place to consistently check in and follow up with stakeholders on aligned timelines to ensure the timely execution of activities. Moreover, additional meeting time was allocated with stakeholders for the review of data and findings to generate insights for further validation.

Technical Execution

Priority and preference were given to hybrid engagements or blended delivery to ensure that the entire process was worthwhile and all planned activities were completed in record time. Due to the emergence of the COVID-19 pandemic, the development of an engagement protocol for physical sessions had to be developed in accordance with globally accepted guidelines.

Teamwork

It became apparent that the sharing of lessons learned was instrumental to team bonding and fostered synergy within the team. The project leveraged the expertise and strengths within the consortium in cases where such was necessary.

Program Management

During the course of the project, appropriate and real-time communication channels and structures had to be put in place for the effective passage of information across partners and stakeholders.

Schedule Conflict

Some of the champions were involved in other programs that were running simultaneously as the HCD co-research workshop. This necessitated the modification of the workshop agendas to suit the needs and time demands of the participants. Going forward, the scheduling process should be improved to enable more champions to attend the workshop.

Conclusion

The study has successfully applied HCD to develop unique solutions capable of strengthening clinical care, reducing gaps in the delivery of maternal health services, and driving improvements in the use of uterotonics for PPH prevention and treatment using statements generated from the co-research and co-design phases. The study objective of developing facility-level solutions for optimizing uterotonic supply chain systems and health service delivery in PPH management was achieved through the creation of solutions that were able to increase the demand for ANC services among pregnant women and their families, particularly their spouses, improve the storage and supply of uterotonics, and generate equitable distribution and improved capacity of health care workers to handle PPH. Although the designed solutions do not address all the issues raised, their implementation can strengthen an evidence-based approach to maternal and child health services and, in turn, improve PPH management. This implementation phase will involve deploying the uptake building of the solutions and documentation outcomes, building a comprehensive picture of the entire process, and identifying potential points for intervention.

Study Limitations

First, this study was conducted in only 3 states within Nigeria, indicating that the developed solutions may not be applicable in other regions with different cultural, socioeconomic, and health care contexts. Second, it focused on proffering solutions to improve uterotonic supply chain systems and health service delivery in PPH management, overlooking other critical aspects like cultural and religious beliefs, as well as the quality of care. Third, the short timeframe of the study and the lack of implementation mean that the long-term outcomes or effectiveness of the solutions were not captured.

Acknowledgments

The authors are grateful to all the respondents who participated in this study. We also wish to appreciate the Kano, Lagos, and Niger States' Ministries of Health, and their respective Primary Health Care Boards for their collaboration and assistance throughout the study lifecycle. This publication was supported by funding from MSD, through its MSD for Mothers initiative. MSD had no role in the design, collection, analysis, and interpretation of data, in the writing of the manuscript, or in the decision to submit the manuscript for publication. The content of this publication is solely the responsibility of the authors and does not represent the official views of MSD. MSD for Mothers is an initiative of Merck & Co, Inc, Rahway, NJ, USA.

XSL•FO
RenderX

Data Availability

The datasets generated and analyzed for this study are available in the Harvard Dataverse repository [35]. The data are publicly available and can be reused under the terms specified in the repository and the publishing journal's data use policy.

Authors' Contributions

Conceptualization: BT, UI, TF, EA Jr, CN
Data curation: LA, MI, O Oluyide, V Adetimilehin, V Amasiatu, EA Jr, NK
Formal analysis: V Amasiatu, UI, EA Jr
Funding acquisition: BT, UI, EA Jr, CN
Investigation: AA, LA, MI, EE, SK, O Oluyide, TA, O Ogefere, NK
Methodology: UI, O Oluyide, TA, O Ogefere, V Amasiatu, EA Jr, NK
Project administration: UI, O Oluyide, V Adetimilehin, V Amasiatu, EA Jr, CN
Resources: BT, UI, TF, CN
Software: UI, V Amasiatu, TA, EA Jr
Supervision: BT, UI, TF, AA, SK, O Ogefere, EA Jr, NK
Validation: AA, LA, MI, EE, SK, O Oluyide, V Adetimilehin, NK
Visualization: O Ogefere, V Adetimilehin, V Amasiatu, EA Jr
Writing – original draft: UI, V Adetimilehin, V Amasiatu, EA Jr
Writing – review & editing: BT, TF, AA, LA, MI, EE, SK, O Ogefere, TA, CN, NK

Conflicts of Interest

None declared.

References

1. Maternal mortality. World Health Organization. 2023. URL: https://www.who.int/news-room/fact-sheets/detail/maternal-mortality [accessed 2024-03-14]
2. Amanuel T, Dache A, Dona A. Postpartum hemorrhage and its associated factors among women who gave birth at Yirgalem General Hospital, Sidama Regional State, Ethiopia. Health Serv Res Manag Epidemiol 2021;8. [doi: 10.1177/23333928211062777] [Medline: 34869791]
3. Abdullahi HM, Aliyu LD, Yusuf M, Miko MA. Obstetric hemorrhage: effective methods for addressing the menace in sub-Saharan Africa. J Perinat Med 2022 Nov 25;50(9):1157-1162. [doi: 10.1515/jpm-2022-0054] [Medline: 35850787]
4. Say L, Chou D, Gemmill A, et al. Global causes of maternal death: a WHO systematic analysis. Lancet Glob Health 2014 Jun;2(6):e323-e333. [doi: 10.1016/S2214-109X(14)70227-X] [Medline: 25103301]
5. Onambele L, Ortega-Leon W, Guillen-Aguinaga S, et al. Maternal mortality in Africa: regional trends (2000-2017). Int J Environ Res Public Health 2022 Oct 12;19(20):13146. [doi: 10.3390/ijerph192013146] [Medline: 36293727]
6. Braddick L, Tuckey V, Abbas Z, et al. A mixed-methods study of barriers and facilitators to the implementation of postpartum hemorrhage guidelines in Uganda. Int J Gynaecol Obstet 2016 Jan;132(1):89-93. [doi: 10.1016/j.ijgo.2015.06.047] [Medline: 26475077]
7. Tasneem S, Nnaji˙ A, Artac M. Causes of maternal mortality in Nigeria: a systematic review. Int J Health Manag Tourism 2019 Dec 12;4(3):200-210. [doi: 10.31201/ijhmt.633942]
8. Trends in maternal mortality 2000 to 2020: estimates by WHO, UNICEF, UNFPA, World Bank Group and UNDESA/Population Division. World Health Organization. 2023 Feb 23. URL: https://www.who.int/publications/i/item/9789240068759 [accessed 2024-03-17]
9. Nigeria demographic and health survey 2018. The Demographic and Health Surveys (DHS) Program. 2019. URL: https://dhsprogram.com/pubs/pdf/FR359/FR359.pdf [accessed 2024-03-14]
10. Sageer R, Kongnyuy E, Adebimpe WO, Omosehin O, Ogunsola EA, Sanni B. Causes and contributory factors of maternal mortality: evidence from maternal and perinatal death surveillance and response in Ogun state, southwest Nigeria. BMC Pregnancy Childbirth 2019 Feb 11;19(1):63. [doi: 10.1186/s12884-019-2202-1] [Medline: 30744576]
11. Akter S, Forbes G, Miller S, et al. Detection and management of postpartum haemorrhage: qualitative evidence on healthcare providers' knowledge and practices in Kenya, Nigeria, and South Africa. Front Glob Womens Health 2022;3:1020163. [doi: 10.3389/fgwh.2022.1020163] [Medline: 36467287]
12. Henriquez D, Bloemenkamp KWM, van der Bom JG. Management of postpartum hemorrhage: how to improve maternal outcomes? J Thromb Haemost 2018 Jun 8;16(8):1523-1534. [doi: 10.1111/jth.14200] [Medline: 29883040]
13. Althabe F, Therrien MNS, Pingray V, et al. Postpartum hemorrhage care bundles to improve adherence to guidelines: a WHO technical consultation. Int J Gynaecol Obstet 2020 Mar;148(3):290-299. [doi: 10.1002/ijgo.13028] [Medline: 31709527]
14. Dao B, Ngabo F, Zoungrana J, et al. Increasing access to prevention of postpartum hemorrhage interventions for births in health facilities and at home in four districts of Rwanda. Afr J Reprod Health 2015 Dec;19(4):58-67. [Medline: 27337854]

XSL•FO
RenderX

15. LaFond A, Cherney M. A theory of change for guiding the integration of human-centered design into global health programming. Glob Health Sci Pract 2021 Nov 29;9(Suppl 2):S209-S216. [doi: 10.9745/GHSP-D-21-00334] [Medline: 34845044]
16. Nijagal MA, Patel D, Lyles C, et al. Using human centered design to identify opportunities for reducing inequities in perinatal care. BMC Health Serv Res 2021;21(1):714. [doi: 10.1186/s12913-021-06609-8] [Medline: 34284758]
17. What's the difference between human-centered design and design thinking? IDEO Design Thinking. 2018. URL: https://designthinking.ideo.com/faq/whats-the-difference-between-human-centered-design-and-design-thinking [accessed 2024-03-17]
18. Holeman I, Kane D. Human-centered design for global health equity. Inf Technol Dev 2019 Sep 29;26(3):477-505. [doi: 10.1080/02681102.2019.1667289] [Medline: 32982007]
19. Human-centered design and child, early, and forced marriage programming. Breakthrough ACTION and RESEARCH. 2021. URL: https://breakthroughactionandresearch.org/wp-content/uploads/2021/10/CEFM-Technical-Brief-HCD.pdf [accessed 2024-03-14]
20. The Field Guide to Human-Centered Design: IDEO; 2015. URL: https://design-kit-production.s3.us-west-1.amazonaws.com/Field_Guides/Field+Guide+to+Human-Centered+Design_IDEOorg_English.pdf [accessed 2025-05-07]
21. Gülmezoglu AM, Lawrie TA, Hezelgrave N, et al. Interventions to reduce maternal and newborn morbidity and mortality. In: Black R, Laxminarayan R, Temmerman M, Walker N, editors. Disease Control Priorities, Volume 2: Reproductive, Maternal, Newborn, and Child Health, 3rd edition: World Bank; 2016. [doi: 10.1596/978-1-4648-0348-2]
22. Jones AJ, Federspiel JJ, Eke AC. Preventing postpartum hemorrhage with combined therapy rather than oxytocin alone. Am J Obstet Gynecol MFM 2023 Feb;5(2S):100731. [doi: 10.1016/j.ajogmf.2022.100731] [Medline: 36028160]
23. Calvello EJ, Skog AP, Tenner AG, Wallis LA. Applying the lessons of maternal mortality reduction to global emergency health. Bull World Health Organ 2015 Jun 1;93(6):417-423. [doi: 10.2471/BLT.14.146571] [Medline: 26240463]
24. Habib SS, Jamal WZ, Zaidi SMA, et al. Barriers to access of healthcare services for rural women-applying gender lens on TB in a rural district of Sindh, Pakistan. Int J Environ Res Public Health 2021 Sep 26;18(19):10102. [doi: 10.3390/ijerph181910102] [Medline: 34639403]
25. Dahab R, Sakellariou D. Barriers to accessing maternal care in low income countries in Africa: a systematic review. Int J Environ Res Public Health 2020 Jun 16;17(12):4292. [doi: 10.3390/ijerph17124292] [Medline: 32560132]
26. Barriers to improving rural maternal health. Rural Health Information Hub. 2021 May 17. URL: https://www.ruralhealthinfo.org/toolkits/maternal-health/1/barriers [accessed 2025-01-24]
27. Ramavhoya TI, Maputle MS, Lebese RT, Makhado L. Midwives' challenges in the management of postpartum haemorrhage at rural PHC facilities of Limpopo province, South Africa: an explorative study. Afr H Sci 2021;21(1):311-319. [doi: 10.4314/ahs.v21i1.40]
28. Braddick L, Tuckey V, Abbas Z, et al. A mixed‐methods study of barriers and facilitators to the implementation of postpartum hemorrhage guidelines in Uganda. Intl J Gynecology Obste 2016 Jan;132(1):89-93. [doi: 10.1016/j.ijgo.2015.06.047]
29. Maternal and obstetric care challenges in rural America: policy brief and recommendations to the secretary. : National Advisory Committee on Rural Health and Human Services; 2020 May URL: https://www.hrsa.gov/sites/default/files/hrsa/advisory-committees/rural/2020-maternal-obstetric-care-challenges.pdf [accessed 2025-01-24]
30. Erstad E, Hopcraft R, Vineetha Harish A, Tam K. A human-centred design approach for the development and conducting of maritime cyber resilience training. WMU J Marit Affairs 2023 Jun;22(2):241-266. [doi: 10.1007/s13437-023-00304-7]
31. Leveraging human-centered design for family planning: lessons and considerations. Breakthrough ACTION and RESEARCH. 2020 Oct. URL: https://breakthroughactionandresearch.org/wp-content/uploads/2021/02/Leveraging-HCD-for-FP.pdf [accessed 2025-01-24]
32. Robert G, Cornwell J, Locock L, Purushotham A, Sturmey G, Gager M. Patients and staff as codesigners of healthcare services. BMJ 2015 Feb 10;350:g7714. [doi: 10.1136/bmj.g7714] [Medline: 25670179]
33. Bazzano AN, Martin J, Hicks E, Faughnan M, Murphy L. Human-centred design in global health: a scoping review of applications and contexts. PLoS ONE 2017;12(11):e0186744. [doi: 10.1371/journal.pone.0186744] [Medline: 29091935]
34. Bird M, McGillion M, Chambers EM, et al. A generative co-design framework for healthcare innovation: development and application of an end-user engagement framework. Res Involv Engagem 2021 Mar 1;7(1):12. [doi: 10.1186/s40900-021-00252-7] [Medline: 33648588]
35. Barriers to utilization of maternal health services for pre-eclampsia and eclampsia management in Kano, Nigeria: a mixed-method study. Harvard Dataverse. URL: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/RXLHK1 [accessed 2025-02-14]

Abbreviations

ANC: antenatal care
CHIPS: Community Health Influencers, Promoters and Services Programme
HCD: human-centered design

XSL•FO
RenderX

MDA: Ministries Departments and Agencies
MoU: memorandum of understanding
NHREC: National Health Research Ethics Committee of Nigeria
NSPHCDA: Niger State Primary Health Care Development Agency
NURTW: National Union of Road Transport Workers
PHC: primary health care
PPH: postpartum hemorrhage
SPHCDA: State Primary Health Care Development Agency
TBA: traditional birth attendant

Edited by A Kushniruk; submitted 19.03.24; peer-reviewed by G Luntsi, J Kaswija; revised version received 17.08.24; accepted 01.12.24; published 07.05.25.

Please cite as:
Tijani B, Igbokwe U, Filani T, Adewemimo A, Ameyan L, Iyekekpolor M, Karera S, Oluyide O, Ezike E, Akinreni T, Ogefere O, Adetimilehin V, Amasiatu V, Nwaokorie C, Kangkum N, Fasawe O, Aigbogun Jr E
Improving Access to and Delivery of Maternal Health Care Services to Prevent Postpartum Hemorrhage in Selected States in Nigeria: Human-Centered Design Study
JMIR Hum Factors 2025;12:e58577
URL: https://humanfactors.jmir.org/2025/1/e58577
doi:10.2196/58577

© Bosun Tijani, Uchenna Igbokwe, Temi Filani, Adefemi Adewemimo, Lola Ameyan, Martins Iyekekpolor, Steven Karera, Olatunji Oluyide, Emmanuela Ezike, Temidayo Akinreni, Obruche Ogefere, Victor Adetimilehin, Valentine Amasiatu, Chukwunonso Nwaokorie, Naanma Kangkum, Olufunke Fasawe, Eric Aigbogun Jr. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 7.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Just-in-Time Adaptive Intervention (Shift) to Manage Problem Anger After Trauma: Co-Design and Development Study

Olivia Metcalf[1,2], PhD; David Forbes[2], PhD; Lauren M Henry[3], PhD; Tianchen Qian[4], PhD; Tracey Varker[2], PhD; Melissa A Brotman[3], PhD; Sean Cowlishaw[5], PhD; Karen E Lamb[6], PhD; Meaghan L O'Donnell[2], PhD

[1]Centre for Digital Transformation of Health, University of Melbourne, 700 Swanston St, Carlton, Australia

[2]Phoenix Australia - Centre for Posttraumatic Mental Health, University of Melbourne, Carlton, Australia

[3]Neuroscience and Novel Therapeutics Unit, Emotion and Development Branch, National Institute of Mental Health Intramural Research Program, Bethesda, MD, United States

[4]Department of Statistics, University of California, Irvine, Irvine, CA, United States

[5]Monash School of Psychological Sciences, Turner Institute for Brain and Mental Health, Monash University, Clayton, Australia

[6]MISCH (Methods and Implementation Support for Clinical Health) Research Hub, Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne, Carlton, Australia

Corresponding Author:
Olivia Metcalf, PhD
Centre for Digital Transformation of Health, University of Melbourne, 700 Swanston St, Carlton, Australia

Abstract

Background: Problem anger is common after experiencing trauma and is under-recognized relative to other posttraumatic mental health issues. Previous research has shown that digital mental health tools have significant potential to support individuals with problem anger after trauma.

Objective: The objective of this study was to describe the co-design and development of a just-in-time adaptive intervention (JITAI) targeting problem anger in individuals who have experienced trauma.

Methods: We used a participatory design process following the double-diamond framework. Phase 1 involved one-on-one qualitative interviews with trauma-exposed individuals with problem anger (n=10). Using an inductive approach (interpretative phenomenological analysis), we thematically coded interview data to create design principles for this population and generate potential content for the intervention. Phase 2 involved academic and clinical experts in trauma and experts in digital health reviewing the Phase 1 results and an evidence-based cognitive behavioral approach to treating anger. We then created intervention content and prototypes, which we then took to workshops with all participants for feedback, using group discussions and ratings of desirability and feasibility.

Results: From Phase 1, core considerations for a JITAI included look and feel preferences, self-led and personalized support and content, and different support needed for each anger stage. A JITAI was developed with the following components: (1) personalized schedules and content onboarding; (2) psychoeducation about problem anger; (3) crisis support; (4) mood monitoring via anger check-ins; (5) self-led and personalized circuit breakers; (6) cognitive-behavioral based skills; (7) and a digital Coach embedded in the app. Some suggested features, such as social networking and sharing data with loved ones, were not pursued due to feasibility reasons relating to participant safety or technical costs.

Conclusions: The resulting JITAI, termed “Shift,” is the first digital mental health tool designed with end users to manage anger after trauma.

(JMIR Hum Factors 2025;12:e62960) doi:10.2196/62960

KEYWORDS

co-design; just-in-time adaptive intervention; mental health; anger; trauma; design; tool; digital tool; post-traumatic; mental health; participatory design; development; user; support; feedback; mood; monitoring; manage

Introduction

Problem anger, defined as anger occurring at an intensity, frequency, and duration that negatively impacts an individual and their relationships, is one of the most common posttraumatic mental health issues, and is highly prevalent in predominately male military and first responder populations [1,2], as well as women who have experienced trauma [3]. While cognitive behavioral therapy (CBT)-based treatments can be effective, there are difficulties engaging individuals with problem anger in treatment due to impacts of hostility, mistrust, and interpersonal conflicts on therapeutic relationships [4]. Mental

XSL•FO
RenderX

health practitioners commonly report feeling more drained, more therapeutic ruptures, and providing shorter treatments in response to antagonistic client presentations, resulting in worse outcomes [5].

Advancements in digital mental health tools may provide nontraditional treatment options for problem anger [6]. As an emotion, anger is unique because it manifests as primarily externally-focused, in that an external agent often activates anger, such as how other people have behaved in a hurtful, disrespectful, or otherwise anger-triggering way [7]. Accordingly, interventions that increase internal focus are likely to support greater self-awareness to offset focus on anger-provoking events. Early efforts in this area used digital technologies such as smartphones and wearables to record physiological data, monitor symptoms, and support practicing of CBT-based skills as an adjunct to traditional CBT, and while effective in reducing anger symptoms, did not have additive effects beyond standard psychological treatment [8]. Recent research using wearables and regular mood monitoring delivered via smartphone have shown promise [9]. Despite the potential, there are many challenges to the development of innovative digital mental health tools for anger. First, the overwhelming majority of research into digital mental health tools has focused on depression and anxiety, and the foundation for building anger-focused digital solutions is limited and nonspecific [10]. Second, digital mental health tools more broadly face significant challenges around uptake and usage [11]. Modest engagement with digital mental health apps is due to a variety of reasons including that may include a lack of user involvement in the design, as well as lack of tailoring to suit the times an individual most needs support [12]. As such, digital mental health tools that are designed alongside users and support real-world moments of high need may be beneficial.

JITAI are 1 type of new digital health approach that may overcome challenges associated with low engagement by providing the right information at the right time [13]. JITAIs deliver intervention content that are adapted to an individual's context, such as the level and type of support matched to an individual's current emotional state [14]. Efforts to develop such interventions in physical activity [15], depression [16], substance use [17], and adolescent mental health [18] have shown promise in improving affect, cognition, and behavior. A JITAI includes 6 key elements, namely a distal outcome, proximal outcomes, decision points, intervention options, tailoring variables, and decision rules. While research has established that digital mental health tools for problem anger are feasible, much remains unknown about how to design a JITAI for this population, and more empirical work including data-driven approaches to designing JITAI is needed [14,17]. The aim of this study was to understand the tailoring variables and co-design the intervention options of a JITAI with individuals who have experienced trauma and problem anger, alongside experts in digital and mental health, by understanding:

1. What is the preferred language, visual features, and content of a digital mental health tool to help individuals manage their problem anger?
2. When do individuals need support to manage their anger differently?
3. Taking the findings from above, and combining them with evidence-based CBT-approaches, what is acceptable and feasible content for a JITAI for problem anger?

Methods

Overview

This study had 2 phases, following the double diamond framework (ie, discover, define, develop, and deliver) [19] which is user-centered and well validated in digital health tool research [20]. Phase 1 involved qualitative interviews with individuals with problem anger after trauma (termed "experts by experience") to understand their experiences with digital mental health tools and treatment for anger, their "pain points" regarding day-to-day management of their anger, and views about essential content, features, and functionality components of smartphone-delivered interventions for problem anger. Phase 2 involved co-design workshops with a mixture of experts by experience, and experts by profession, which included digital health experts, trauma researchers, and clinical psychologists to converge on core components of a prototype JITAI, using CBT for anger after trauma principles [21]. Both phases were conducted in line with trauma-informed research principles outlined in the recent Trauma and Resilience Informed Research Principles and Practice (TRIRPP) Framework [22], which aims to improve participation in research by individuals who have experienced trauma and to avoid retraumatization of participants. Specifically, our study involved actively minimizing retraumatization by structuring interview guides so that participants do not have to discuss their trauma, framing the research process as relational, addressing power imbalances by having twice the number of experts by experience than experts by profession, and making available women-identifying groups only for the group sessions, recognizing many women experience trauma as a result of gendered violence. We also made available individual co-design workshops for participants who did not want to engage in group work for any reason.

Phase 1

Participants

A total of 10 experts by experience (5 out of 10 were women; aged 19 - 49 years old) with a history of trauma exposure and problem anger were invited to participate via the researchers' database, which contained a list of individuals who had participated in previous trauma and anger–related research at the University of Melbourne. Participants needed to meet the criteria for problem anger and have reported experiencing a traumatic event previously, be residing in Australia, have access to Zoom, be over the age of 18 years, and be fluent in English. All participants had sought help for their mental health previously, and 4 had received a diagnosis of posttraumatic stress disorder. One individual identified as living with a disability. In addition, 8 participants had previously tried apps to support their mental health. All participants verbally consented to participate in the qualitative interviews. Data were collected virtually via Zoom (Zoom Communications) using audio-video recordings across August-September 2023 by author OM using a semistructured interview guide and interviews lasted

XSL•FO
RenderX

approximately 45 minutes. Interview guides were developed by the research team after reviewing the literature around problem anger experiences, and digital mental health tools. Interview guides are in Multimedia Appendix 1.

Analyses

Qualitative interview data were transcribed verbatim from audio-video recordings. Data were analyzed using NVivo (Lumivero) and using an inductive approach (interpretive phenomenological analysis) by 2 independent coders who were part of the research team, to allow identification of themes about the experiences of problem anger treatments, and views and preferences about content, features, and functionality of the JITAI [23]. Each coder read the transcripts, coded the insights, and clustered the codes into themes. Coders discussed consensus at the insights and themes stage to reach convergence. No formal inter-rater reliability was conducted.

Phase 2

Participants

All participants from Phase 1 were invited to attend Phase 2 and agreed to participate. In addition, 6 experts by profession (ie, clinicians and researchers in problem anger and digital health) were invited to participate. This number was determined to balance the power dynamics between expert types and to ensure experts by experience were not outnumbered in workshops. Experts by profession were identified by the research team based on their expertise in the field of anger, trauma, or digital health and invited directly. Data were collected virtually across October 2023. Focus group sizes varied between 4 - 6 individuals per group and lasted approximately 90 minutes and were conducted via Zoom. Phase 2 were not audio-recorded.

Analyses

The themes from Phase 1 were reviewed by the research team alongside the content from an evidence-based CBT manual developed by one of the authors (DF) for managing anger after trauma [21] to distill into a range of possible content features that could help users manage their anger. These content features, as well as 3 different aesthetics styles developed by the digital health experts after reviewing the Phase 1 data, were then taken to Phase 2 workshops in the form of low fidelity prototypes for discussion about content to keep, alter, or remove [24], with the goal of converging on a final prototype. The app title, and the stages of anger was also the focus of the workshops, in terms of converging on a consensus for how to represent them. Finally, the workshop participants charted components of the possible prototype on a 2 by 2 feasibility (low to high) and desirability (low to high) matrix. Workshop discussions were captured as detailed written notes by the facilitator [OM] and another member of the research team who was observing, and comments typed into the Zoom chat function from participants within each of the workshops relating to suggestions to keep, alter, or remove aspects of content, and the charting results of feasibility and acceptability.

Ethical Considerations

This study was approved by the University of Melbourne Human Research Ethics Committee (Approval number: 2023-26108-39829-4). Written informed consent was obtained for the study. Experts by experience were reimbursed for their time in the form of a voucher ($96.06; AUD $1 = US $1.56) for Phase 1 and 2, while experts by profession were not.

Results

Phase 1

The results from the individual interviews revealed several themes, listed in Table 1, relating to preferences for a JITAI for anger. First, participants described wanting an app that was free, with no ads, and easy to use. Second, several themes relating to design principles were identified. These included the “Look & Feel,” which detailed specific preferences for the language and imagery used in the app. Experts by experience felt that many mental health apps overused the word “calm” and imagery with strong meditation and wellness themes, as well as being too directive, using terms such as “should” and “need” to direct users. As a result, they felt that many mental health apps were condescending and irritating, and worsened anger when used for self-management. The theme “Strengths-based” described preferences of experts by experience for an app that felt supportive and addressed the guilt and shame they felt about their anger. Experts by experience reflected wanting digital mental health tools that feel like a companion: “a trusted mate.” Under the theme of “Personalized,” there was divergence regarding the types of content individuals wanted, such as military and first-responder-specific information explaining the role of occupation in anger, the links between anger and substance use, as well as menstrual health. Strongly linked was the theme “Self-led” which encompassed accounts of who support should come from and how information should be delivered, particularly during heightened states of anger. While the voice of professional experts and lived experience was valued, overwhelmingly, experts by experience conveyed that the person they would listen to was themselves. Conversely, some experts by experience conveyed how complex trauma left them uncertain about trusting their own voice, and thus they would want information delivered to them by an avatar they could create.

XSL•FO
RenderX

Table . Design principles and desired intervention components from qualitative interviews with ten participants aged 19 - 49 years old with a history of trauma and problem anger.

Themes	Example quote
Design principles	
Free and no ads	__[a]
Easy to use	*The most annoying and triggering thing about using an app is when it doesn't do what you want it to do, or you're trying to do something, and then you['re] hit with roadblocks*. [Expert by Experience 1]
Look & Feel	*It would have to be written in a way that it comes from someone that isn't condescending; that doesn't make you angry reading it*. [Expert by Experience 1]
	It tells you to take three deep breaths when you log in on that [mental health app]. I'm like, I don't want to take three deep breaths. I can't. When I breathe, it makes me angry. [Expert by Experience 2]
Strengths-based	*I'm not going to use anything that is just focused on my anger. Like who would want that, just some negative app going [beep] 'you are mad again'. [laughs] praise me for stuff too*. [Expert by Experience 2]
Personalized	*One option could be you could tick in the app, 'these are the things that really affect me and are related to my anger.' And for you, that might be anxiety, it might not be pain or sleep. And then you could also perhaps get support specifically around that. So you have a really personalized kind of list*. [Expert by Experience 1]
Self-led	*Don't tell me what to do straight up*. [Expert by Experience 2]
	I don't want to listen to some expert, no offence. I don't read your textbooks, I don't like studies, theses. I didn't read them. [Expert by Experience 7]
Preferred features of an anger-focused app	
Anger is experienced in distinct stages with different support needed for each stage	*'1 to 10' for my anger- that doesn't mean anything to me. It's out of control, or it's not*. [Expert by Experience 3]
	I want [an app] that … would recognize over time –'okay, he needs an aid, so I sort of guess he's feeling these sorts of things'. And then it would give me something structured more for that for that level for me, rather than just a generic one. [Expert by Experience 4]
Targeting cognition and behavioral aspects of anger	*Rumination is something I've had to work really hard on, because it's caused me a lot of problems recently…I like to say to myself, 'it's done. It's finished, you're safe, it's past.' And I almost take it out of myself and then put it down and say I'm leaving it there. And it's not going any further. And then when I catch myself ruminating about it if I get home, I say no, that's outside of I haven't brought that in I'm not gonna let that come into my calm area of home*. [Expert by Experience 3]
Leveraging physiological data	*I've used a wearable and with the blood pressure and seeing the heartrate, that's very intriguing*. [Expert by Experience 6]
Complimenting current mental health treatments	*If I trusted [a psychologist] and felt like things were progressing, [an app] would be something that I would seek out her opinion on*. [Expert by Experience 7]
Regular check ins	*The process of a check-in and reflecting your anger [is] good, because I never really do that. Like, I don't think anyone does. I don't think anyone really sits there and reflects on how they're feeling or anything like that*. [Expert by Experience 5]
Add-ons	
Social features with other users	*An app that could connect us all, so we don't feel like we're all alone, that we still got an identity, something that although we're doing the app by ourselves, but gives us a feeling that we're not by ourselves*. [Expert by Experience 7]
Sharing data with loved ones	*I would like the app to send something to my husband saying I've had a shocking day, because then that way, he'll be prepared before he comes home*. [Expert by Experience 3]

XSL•FO
RenderX

Themes	Example quote
Parenting support	*This is hard for me to admit, but my anger with my kids is such a huge trigger [pause interview due to distress]. Yeah, so anyway, things that help me be a better mum.* [Expert by Experience 8]

[a] —: not available.

In addition to these design features, several other themes emerged in relation to a JITAI for anger. In identifying opportunities to intervene, participants reported that their anger is experienced in distinct stages, highlighting that for them, anger is either increasing, or completely out of control, with different support needed for each stage. Participants also wanted support in the form of ”targeting cognition and behavioral aspects of anger” such as rumination and supporting interactions with others when angry, ”leveraging physiological data” to detect their mood, an app that was “complimenting current mental health treatments,” and “regular check ins” of their anger.

Phase 2

After reviewing Phase 1 results, the trauma experts selected content from an anger-focused CBT-based manual [21] that related to psychoeducation, managing rumination, circuit breakers, cognitive reframing, managing physiological arousal, and assertive communication. This potential content and low-fidelity prototype design options were taken to co-design workshops. The results from the co-design workshop included convergence on the way anger stages should be represented, convergence on the look and feel of the app, convergence on the CBT-based content, and discussions around the feasibility of add-ons. While experts by profession discussed anger in terms of Likert scales (ie, 1 to 10), the experts by experience spoke about how “anger was experienced in stages,” with some individuals likening this to a runaway train. The workshops converged on the following stages of anger, shown in Figure 1: (1) “no anger,” (2) “frustrated” (ie, anger is still within an individual’s control, and while heightened, the individual can engage in cognitive, behavioral, or physiological activities), (3) “out of control” (ie, the individual and/or others are not safe), and (iv) “after a crash” (ie, anger at others has dissipated, and feelings of shame, regret, guilt, anger at self, sadness, and loneliness are high). In terms of content, all the recommended CBT skills were considered feasible. While the add-ons such as social networking and sharing data with loved ones was highly desirable by experts by experience, they were rated low in feasibility by the experts by profession, due to technical challenges, cost, and the lack of information on how to protect users’ safety, so were not included. Parenting support was considered as out of scope for the first iteration of the app.

XSL•FO
RenderX

Figure 1. The tailoring variables of the “Shift” just-in-time adaptive intervention in the form of anger stages.

XSL•FO
RenderX

The results of both phases led to the development of a JITAI named "Shift" (Figure 2 and Textbox 1), combines mood monitoring (ie, 4 daily check-ins on anger stages, with ability to check-in outside of these 4) with tailored CBT-based support based on the stage of anger. When users check-in as "frustrated," 6 CBT and arousal management skills, under the categories of "Body, Mind, and Actions" in (Textbox 1) are made available. When users check-in as "out of control," a personalized circuit breaker preset by the user at the download of the app sends a safe, supportive message, video, photo, image, or song to circuit break the situation. And when users check-in as "after a crash," warm and emphatic audio messages from their "Shift" Coach are sent reassuring them that set-backs are normal in recovery and reminding them to do some self-care to get back on track. In addition to these JITAI components are psychoeducation content available on demand.

XSL•FO
RenderX

Figure 2. Home screen and psychoeducation section of “Shift.”

XSL•FO
RenderX

Textbox 1. Content structure of the just-in-time adaptive intervention "Shift."

Onboarding:

- Welcome and explanation that "Shift" was created by people with lived experience.
- Set up personalized circuit breakers for when anger is out of control.
- Set-up notification schedule for check-ins.
- Switch-off for unwanted information, such as alcohol-related information, menstrual health, and military and first responder.
- Select from one of the 2 coaches, who is a digital voiceover that provides the audio for the app, the welcome message, and the messages provided when an individual reports losing control of their anger.

Psychoeducation:

- Anger after trauma.
- Healthy versus unhealthy anger.
- Baseline arousal and how to reduce stress levels using progressive muscle relaxation.
- How anger is linked to military or first responder roles.
- How menstrual health and anger might be connected.
- How diet, exercise, sleep, connection, and substance use is connected to overall well-being.
- Rumination and how to stop it.
- Assertive communication versus unhealthy communication styles.

Mood monitoring:

- Four daily check-ins on anger stage (ie, no anger, escalating, out of control, and after a crash; Figure 1).

Intervention components:

- When users check-in as having no anger, they are prompted to either practice progressive muscle relaxation, or engage with the psychoeducation components of "Shift."
- When users check-in that their anger is escalating, they are prompted to practice a cognitive-behavioral-therapy–based skill that focuses on the physiological aspect of anger (termed, "body"), the cognitive aspect of anger (termed "mind"), or their behavior when angry (termed "action"). The "body" intervention components are isometric exercises, a targeted form of tension release designed specifically for discharging anger from the body, and cyclic sighing, recognized as the simplest and most effective way to increase relaxation. The "mind" intervention components target stopping anger rumination, and cognitive reframing for anger-provoking situations. The "action" intervention components provide assertive communication skills to use.
- When users check-in that their anger is out of control, they receive one of their circuit breakers established at the onboarding of the app.
- When users check-in that they have had an episode of anger, and behaved in a way they regret, (termed "after a crash"), a warm supportive message and suggestions for self-care are provided by the app coach.

Crisis:

- At any time, users can go into the app and receive tailored helpline information to support a crisis state.

Discussion

Principal Findings

The aim of this study was to co-design and develop a JITAI for individuals who have experienced trauma and problem anger. We engaged experts by experience and experts by profession to establish design principles and content, as well as opportunities to intervene, resulting in the JITAI "Shift." Building on previous work that has used mood monitoring for problem anger [2,9], "Shift" is a JITAI that leverages regular check-ins on anger state as the tailoring variables with bite-sized, digitally delivered CBT-based skills to manage lower levels of anger, and a personalized, self-led circuit breaker for when anger is out of control. "Shift" is now developed because of these findings and a micro-randomized trial is being conducted.

From a design perspective, using individual interviews before group workshops was valuable for managing trauma, sensitive personal issues, and the feelings of shame that can accompany problem anger in experts by experience, as it allowed individuals to generate information about the intervention privately, but then also benefit from sharing ideas in a group format. Other methods within the workshop to improve the participant experience include not using digital tools to collect data, which can be frustrating or challenging to learn to navigate, and not getting participants to actively criticize or provide feedback on each other's ideas. Although there is potential for more work in this area, these findings speak to the importance of research and development of digital mental health tools being

XSL•FO
RenderX

co-designed with potential end users. Emerging research is showing that co-design is an important step to ensuring that digital mental health tools are accessible, appropriate, and effective [24]. Without end-user involvement, interventions could miss the mark in terms of developing effective and acceptable digital health programs. Very little research has used co-design in problem anger, or in the design of JITAI. Our findings are consistent with a study that co-designed a JITAI for children to limit sun exposure, in that both opportune times to intervene, and the content and design of the intervention were uncovered as a result of co-design [25].

Several learnings from this study have significant implications for design of digital mental health tools for understudied populations and problem anger more broadly. First, our results showed that individuals with problem anger often try digital mental health tools focused on symptoms of anxiety and depression and can find these anger-provoking. Second, while previous research has established the value of lived experience stories in digital mental health tools for trauma-exposed populations [26], our study found that individuals with problem anger preferred a self-led approach rather than hearing others stories. Third, our results shows that individuals with problem anger have significant unmet needs in the aftermath of a high anger episode. Finally, our results showed that digital mental health tools must balance a focus on changing pathological thoughts, feelings, and behavior with providing warmth, empathy, validation, and support, all of which are typically provided via the therapeutic relationship in psychological treatment and are at risk of being lost in digital approaches. Important next steps are evaluating the components of "Shift" using methods such as a micro-randomized trial, followed by a randomized controlled trial to validate efficacy.

Limitations

While the study had diversity in terms of gender, some representation of disability, and a diversity of trauma experiences, this study is limited by a sample that contained English-speaking individuals only. In addition, the study only collected qualitative data and work is needed to pilot the intervention to determine acceptability and feasibility of the intervention to individuals who have experienced trauma and have problem anger. Finally, EMA-driven JITAIs can become burdensome for participants, and innovations in sensor-based detection of mood states will likely advance "Shift" [27]. As yet, preliminary work in detecting anger in this population from wearables has found promising results that remain technically challenging to deploy, and sensor-based JITAIs in this space warrant further investigation [28].

Acknowledgments

This study was funded by the Australian National Health and Medical Research Council. This research was supported in part by the Intramural Research Program of the NIMH (National Institute of Mental Health; No.: ZIA-MH002969). The funders were not involved in any aspect of this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interviews.
[DOCX File, 23 KB - humanfactors_v12i1e62960_app1.docx]

References

1. Adler AB, LeardMann CA, Roenfeldt KA, Jacobson IG, Forbes D, Millennium Cohort Study Team. Magnitude of problematic anger and its predictors in the Millennium Cohort. BMC Public Health 2020 Jul 27;20(1):1168. [doi: 10.1186/s12889-020-09206-2] [Medline: 32718306]
2. Varker T, Cowlishaw S, Baur J, et al. Problem anger in veterans and military personnel: prevalence, predictors, and associated harms of suicide and violence. J Psychiatr Res 2022 Jul;151:57-64. [doi: 10.1016/j.jpsychires.2022.04.004] [Medline: 35453092]
3. Cowlishaw S, Metcalf O, Varker T, et al. Anger dimensions and mental health following a disaster: distribution and implications after a major bushfire. J Trauma Stress 2021 Feb;34(1):46-55. [doi: 10.1002/jts.22616] [Medline: 33136348]
4. DiGiuseppe R, Tafrate R, Eckhardt C. Critical issues in the treatment of anger. Cogn Behav Pract 1994 Jun;1(1):111-132. [doi: 10.1016/S1077-7229(05)80089-6]
5. Hyatt CS, Sleep CE, Hemmy Asamsama O, Reger MA. Surveying veterans affairs mental health care providers on experiences working with veteran patients with antagonistic clinical presentations. Psychol Serv 2024 May;21(2):379-387. [doi: 10.1037/ser0000782] [Medline: 37358554]
6. Forbes D, Adler AB. Advancing anger research. In: Adler AB, Forbes D, editors. Anger at Work: Prevention, Intervention, and Treatment in High-Risk Occupations: American Psychological Association; 2021:305-315. [doi: 10.1037/0000244-011]
7. Berkowitz L, Harmon-Jones E. Toward an understanding of the determinants of anger. Emotion 2004 Jun;4(2):107-130. [doi: 10.1037/1528-3542.4.2.107] [Medline: 15222847]

XSL•FO
RenderX

8. Mackintosh MA, Niehaus J, Taft CT, Marx BP, Grubbs K, Morland LA. Using a mobile application in the treatment of dysregulated anger among veterans. Mil Med 2017 Nov;182(11):e1941-e1949. [doi: 10.7205/MILMED-D-17-00063] [Medline: 29087863]
9. Metcalf O, Pham L, Lamb KE. A mixed-methods investigation of A digital mental health tool to manage posttrauma anger. J Trauma Stress 2025 Apr;38(2):296-304. [doi: 10.2196/preprints.58032]
10. Miralles I, Granell C, Díaz-Sanahuja L, et al. Smartphone apps for the treatment of mental disorders: systematic review. JMIR Mhealth Uhealth 2020 Apr 2;8(4):e14897. [doi: 10.2196/14897] [Medline: 32238332]
11. Kerst A, Zielasek J, Gaebel W. Smartphone applications for depression: a systematic literature review and a survey of health care professionals' attitudes towards their use in clinical practice. Eur Arch Psychiatry Clin Neurosci 2020 Mar;270(2):139-152. [doi: 10.1007/s00406-018-0974-3] [Medline: 30607530]
12. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. Evid Based Ment Health 2018 Aug;21(3):116-119. [doi: 10.1136/eb-2018-102891] [Medline: 29871870]
13. Wang L, Miller LC. Just-in-the-moment adaptive interventions (JITAI): a meta-analytical review. Health Commun 2020 Nov;35(12):1531-1544. [doi: 10.1080/10410236.2019.1652388] [Medline: 31488002]
14. Nahum-Shani I, Smith SN, Spring BJ, et al. Just-in-time adaptive interventions (JITAIs) in mobile health: key components and design principles for ongoing health behavior support. Ann Behav Med 2018 May 18;52(6):446-462. [doi: 10.1007/s12160-016-9830-8] [Medline: 27663578]
15. Hardeman W, Houghton J, Lane K, Jones A, Naughton F. A systematic review of just-in-time adaptive interventions (JITAIs) to promote physical activity. Int J Behav Nutr Phys Act 2019 Apr 3;16(1):31. [doi: 10.1186/s12966-019-0792-7] [Medline: 30943983]
16. Teepe GW, Da Fonseca A, Kleim B, et al. Just-in-time adaptive mechanisms of popular mobile apps for individuals with depression: systematic app search and literature review. J Med Internet Res 2021 Sep 28;23(9):e29412. [doi: 10.2196/29412] [Medline: 34309569]
17. Perski O, Hébert ET, Naughton F, Hekler EB, Brown J, Businelle MS. Technology-mediated just-in-time adaptive interventions (JITAIs) to reduce harmful substance use: a systematic review. Addiction 2022 May;117(5):1220-1241. [doi: 10.1111/add.15687] [Medline: 34514668]
18. Bell IH, Arnold C, Gilbertson T, et al. Pilot randomised controlled trial of mello: a personalised, transdiagnostic smartphone intervention targeting repetitive negative thinking in young people with depression and anxiety. SSRN. . [doi: 10.2139/ssrn.4356608]
19. Design Council UK. 2019. URL: https://www.designcouncil.org.uk/our-resources/framework-for-innovation/ [accessed 2025-06-02]
20. Bucher A, Chaudhry BM, Davis JW, et al. How to design equitable digital health tools: a narrative review of design tactics, case studies, and opportunities. PLOS Digit Health 2024 Aug;3(8):e0000591. [doi: 10.1371/journal.pdig.0000591] [Medline: 39172776]
21. Cash R, Varker T, McHugh T, et al. Effectiveness of an anger intervention for military members with PTSD: a clinical case series. Mil Med 2018 Sep 1;183(9-10):e286-e290. [doi: 10.1093/milmed/usx115] [Medline: 29584890]
22. Edelman NL. Trauma and resilience informed research principles and practice: a framework to improve the inclusion and experience of disadvantaged populations in health and social care research. J Health Serv Res Policy 2023 Jan;28(1):66-75. [doi: 10.1177/13558196221124740] [Medline: 36326603]
23. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
24. Malloy J, Partridge SR, Kemper JA, Braakhuis A, Roy R. Co-design of digital health interventions with young people: a scoping review. Digit Health 2023;9:20552076231219117. [doi: 10.1177/20552076231219117] [Medline: 38107986]
25. Huh J, Lee KJ, Roldan W, et al. Making of mobile SunSmart: co-designing a just-in-time sun protection intervention for children and parents. Int J Behav Med 2021 Dec;28(6):768-778. [doi: 10.1007/s12529-021-09987-9] [Medline: 33846955]
26. Metcalf O, Gibson K, Fredrickson J, Finlayson-Short L, Varker T, O'Donnell M. Design, development and randomised controlled trial protocol of a smartphone-delivered version of "SOLAR" for emergency service workers to manage stress and trauma. BMJ Open 2023 Feb 10;13(2):e062710. [doi: 10.1136/bmjopen-2022-062710] [Medline: 36764725]
27. Perski O, Kale D, Leppin C, et al. Supervised machine learning to predict smoking lapses from Ecological Momentary Assessments and sensor data: Implications for just-in-time adaptive intervention development. PLOS Digit Health 2024 Aug;3(8):e0000594. [doi: 10.1371/journal.pdig.0000594] [Medline: 39178183]
28. Metcalf O, Lamb KE, Forbes D, et al. Predicting high anger intensity using ecological momentary assessment and wearable-derived physiological data in a trauma-affected sample. Eur J Psychotraumatol 2025 Dec;16(1):2472485. [doi: 10.1080/20008066.2025.2472485] [Medline: 40135377]

Abbreviations

CBT: cognitive behavioral therapy

XSL•FO
RenderX

JITAI: just-in-time adaptive intervention
NIMH: National Institute of Mental Health
TRIRP: Trauma and Resilience Informed Research Principles and Practice

Edited by A Kushniruk; submitted 05.06.24; peer-reviewed by C Leppin, M Gasmi, TCC Hsu; revised version received 12.02.25; accepted 14.02.25; published 22.05.25.

Please cite as:
Metcalf O, Forbes D, Henry LM, Qian T, Varker T, Brotman MA, Cowlishaw S, Lamb KE, O'Donnell ML
A Just-in-Time Adaptive Intervention (Shift) to Manage Problem Anger After Trauma: Co-Design and Development Study
JMIR Hum Factors 2025;12:e62960
URL: https://humanfactors.jmir.org/2025/1/e62960
doi:10.2196/62960

© Olivia Metcalf, David Forbes, Lauren M Henry, Tianchen Qian, Tracey Varker, Melissa A Brotman, Sean Cowlishaw, Karen E Lamb, Meaghan L O'Donnell. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 22.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Health Professional Mentorship Platform to Improve Equitable Access to Abortion: Development, Usability, and Content Evaluation

Abdul-Fatawu Abdulai[1], PhD; Cam Duong[1], PhD; Eleni Stroulia[2], PhD; Efrat Czerniak[1], PhD; Rachel Chiu[3], BSc; Aashay Mehta[1], MA; Ken Koike[4], MPH; Wendy V Norman[5], FCFP, DTM&H, MD, MHSc

[1]School of Nursing, University of British Columbia, Vancouver, BC, Canada

[2]Computing Science, Faculty of Science, University of Alberta, Edmonton, AB, Canada

[3]Cognitive Sciences, University of British Columbia, Vancouver, BC, Canada

[4]The Society of Obstetricians and Gynaecologists of Canada, Ottawa, ON, Canada

[5]Department of Family Practice, Faculty of Medicine, University of British Columbia, Vancouver, ON, Canada

Corresponding Author:
Abdul-Fatawu Abdulai, PhD
School of Nursing, University of British Columbia, Vancouver, BC, Canada

Abstract

Background: Access to safe abortion care is a reproductive right for all individuals across Canada. Underserved populations are overrepresented among those with unintended pregnancies and particularly those seeking abortion. Yet, few resources exist to help health care and allied helping professionals provide culturally competent and gender-affirming abortion care to such a population group.

Objective: This project aimed to redesign and adapt an existing subscription-based medication abortion mentorship platform into a culturally appropriate and gender-affirming open-access website of curated health professional resources to promote equitable, accessible, high-quality abortion care, particularly for underserved populations.

Methods: We drew on a user-centered design framework to redesign the web platform in 5 iterative phases. Health care and allied helping professionals were engaged in each stage of the development process including the initial design of the platform, curation of the resources, review of the content, and evaluation of the wireframes and the end product.

Results: This project resulted in an open-access bilingual (English and French) web-based platform containing comprehensive information and resources on abortion care for health care providers (physicians, nurse practitioners, and pharmacists) and allied helping professionals (midwives, medical officers, community workers, and social workers). The website incorporated information on clinical, logistical, and administrative guidance, including culturally competent and gender-affirming toolkits that could equip health care professionals with the requisite knowledge to provide abortion care for underserved populations.

Conclusions: This platform contains resources that can increase the competencies of health care professionals to initiate and sustain culturally and contextually appropriate abortion care for underserved groups while clarifying myths and misconceptions that often militate against initiating abortion. Our resource also has the potential to support equitable access to high-quality abortion care, particularly for those among underserved populations who may have the greatest unmet need for abortion services yet face the greatest barriers to accessing care.

(JMIR Hum Factors 2025;12:e63364) doi:10.2196/63364

KEYWORDS

medication abortion; mifepristone; web-based platform; user-centered design; underserved populations; abortion; equitable; accessibility; open-access website; gender-affirming; user-centered; Canada; unwanted pregnancy; framework

Introduction

Access to abortion is a fundamental human right and a critical component of sexual and reproductive health care [1]. Each year, approximately 56 million abortions occur across the globe—translating to 35 abortions per 1000 women. In Canada, nearly half (~40%) of pregnancies are unplanned, and one-third of women and pregnancy-capable people will have at least 1 abortion in their lifetime [2,3]. Medication abortion using mifepristone is the international gold standard for first-trimester medication abortions and currently accounts for over 50% of all abortions in most European countries and the United States. Canada's unique regulations allow nonphysician prescribers to prescribe and dispense the medication and allow people without

XSL•FO
RenderX

access to ultrasound to access medication abortion services. Despite the loosening of restrictions in Canada, timely access to safe abortion services can be difficult for historically, persistently, or systemically marginalized (HPSM) populations including racialized groups, migrants, Indigenous people, people with disabilities, homeless and underhoused people, sex workers, two-spirit, lesbian, gay, bisexual, transgender, queer, intersex, and gender-diverse (2SLGBTQI+) people, youths, and people living in rural and remote areas [4,5]. For instance, Indigenous people in Canada experience more abortion access barriers than non-Indigenous Canadian people [6]. In addition, approximately 18% of people in Canada traveled more than 100 km to access abortion services [7]. While most clinicians in Canada indicated providing abortion services to underserved populations, the majority of them did not have the appropriate cultural training to provide care to underserved populations. For instance, a recent study suggests that most clinicians in Canada who provide abortion services to underserved populations (91.2%) reported not receiving any form of training for providing care to such populations [8]. A lack of training or mentorship in culturally safe and gender-affirming care would further exacerbate barriers to accessing abortion services that already bedevil underserved populations including Indigenous people [9,10].

To provide culturally safe and gender-affirming abortion services, there is an urgent need to develop, evaluate, and implement training and mentorship resources that will equip health care professionals in providing equitable abortion services to underserved populations in Canada. In this paper, we reported on the processes we adopted in redesigning our already-established web-based community of practice platform into an open-access website. We also highlighted on the considerations for developing a gender-affirming, culturally appropriate, and inclusive web platform to support health professionals in providing abortion to equity-deserving populations. With the internet increasingly becoming a major source of information on abortion [11], an open-access web platform could not only help in disseminating evidence-based abortion resources to all health care professionals but also help in dispelling the myths and misconceptions that often fuel abortion-related stigmas. Furthermore, an open-access web platform could assist in countering abortion-related misinformation or disinformation that is often targeted at health professionals by antichoice groups.

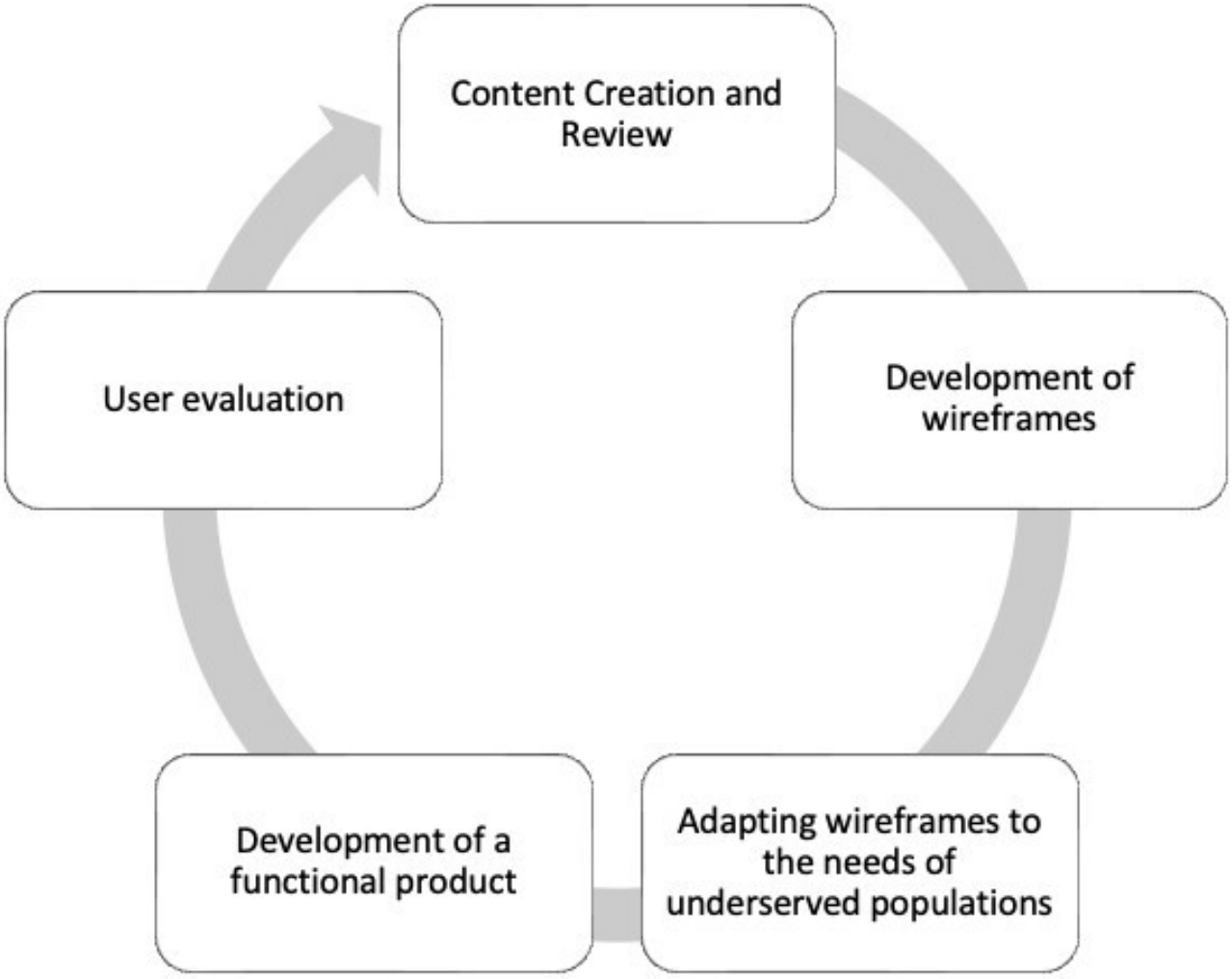
Methods

Study Design

We used a mixed methods user-centered design approach in developing the content and structure of the website [12]. In compliance with the user-centered design process, we engaged with the relevant stakeholders in Canada (health professional associations and abortion advocacy organizations) in 5 iterative design phases. The five phases include (1) content creation and review, (2) development of wireframes, (3) adaptation of wireframes to end-user needs, (4) development of a functional product, and (5) user evaluation. We adopted a user-centered design process to emphasize the clinician-centeredness of the website while revealing the knowledge gaps that need to be addressed to create an inclusive web platform. Figure 1 shows a schematic presentation of the website's development process.

XSL•FO
RenderX

Figure 1. A schematic depiction of the website development process.



Ethical Considerations

This project was approved by the University of British Columbia Behavioral Research Ethics Board (approval H22-03342). All participants who took part in the interviews were provided with CAD $150 (US $103.34) honoraria.

Phase 1: Content Creation and Review

Overview

The content creation phase comprised 2 steps to ensure that the content is up-to-date and reliable. These include (1) the content curation process, in which we searched for resources that are relevant to medication abortion, and (2) the content review process.

Content Curation

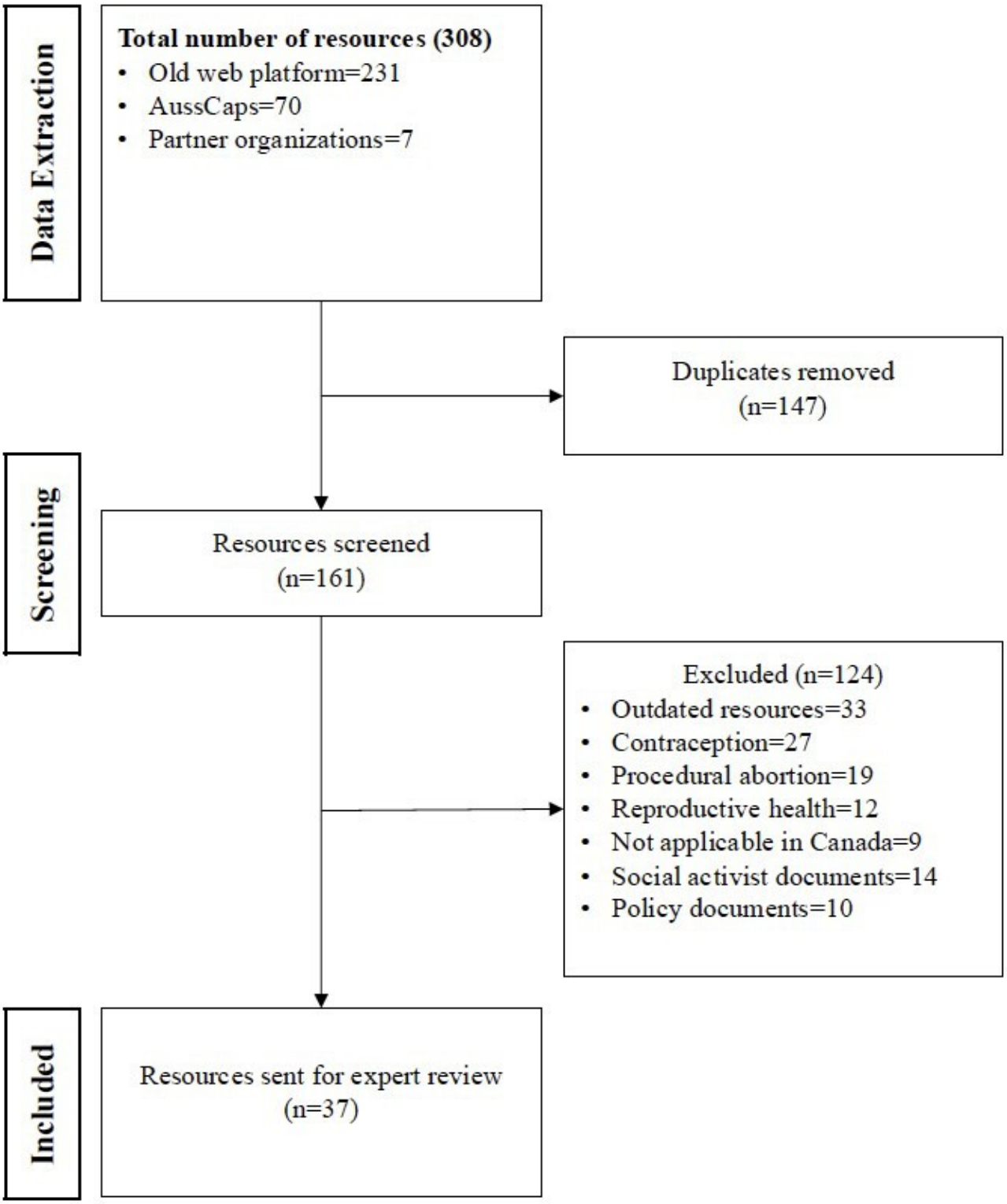
We actively engaged with health professional associations, health regulatory agencies, and sexual and reproductive health advocacy organizations across Canada to curate relevant content for the website. To ensure high-quality, appropriate, evidence-based materials specific to Canadian health care professionals, we focused our content curation efforts on sources within Canada as well as trusted and reliable international evidence on medication abortion. The main sources for curating our content include the national health professional organizations (Society of Obstetricians and Gynaecologists of Canada, Canadian Pharmacy Association, Canadian Association of Midwives, and Canadian Association of Schools of Nursing) and advocacy groups (Action Canada for Sexual Health and Rights and National Abortion Federation) as well as our already-established web-based community of practice platform [13]. We examined relevant resources from a partner website in Australia (the Australian Contraception and Abortion Primary Care Practitioner Support Network) [14]. We limited our focus to medication abortion to enhance relevance and usability for clinicians looking to adopt, improve, and sustain the practice of this newly available service in Canada. After identifying the data sources and establishing a working relationship with stakeholders, we undertook focused data extraction from May 2023 to November 2023. The data extraction process yielded a total of 308 documents detailing resources on medication abortion. These were compiled into a Microsoft Excel sheet, and 147 duplicates were removed. Following this, we conducted an initial screening of the resources by examining the titles, abstracts, headings, and subheadings to determine their relevance to medication abortion. To be eligible for inclusion, each resource must be focused on medication abortion services including clinical (guidelines, protocols, checklists, and factsheets) and nonclinical services (financial, legal, and social). Since our objective was to improve equitable access to abortion among underserved populations, medication abortion resources

XSL•FO
RenderX

focusing on Indigenous populations, racialized groups, and gender minorities were included. Resources were excluded if they were dated more than 20 years ago, not relevant to the Canadian context, beyond the scope of medication abortion, or not targeted to health care providers (ie, physicians, pharmacists, social workers, and midwives). Most papers were not relevant to the Canadian context because of the constant regulatory changes regarding the prescription, dispensing, and use of medication abortion in Canada since 2017 [15-17]. We also excluded resources that were exclusively focused on procedural abortion, contraception, or general reproductive health or illnesses. Procedural abortion resources were excluded because procedural abortion demands specialized training, and we considered a web platform as an inappropriate training source for those seeking to provide this skilled task. After screening the resources against the inclusion and exclusion criteria, 124 resources were excluded, leaving us with 37 resources for a detailed expert review. Figure 2 shows the data extraction and review process.

XSL•FO
RenderX

Figure 2. Resource extraction and screening process.



Content Review

Even though the resources were extracted from credible sources, medication abortion regulations and professional practice guidelines in Canada have evolved over time, and with these changes, some of the resources might have become outdated. To ensure our website contains up-to-date resources and is in line with current practices, we engaged 10 health care and allied helping professionals providing abortion services to review the 37 resources for accuracy, currency, and relevance for providing medication abortion, particularly for equity-deserving populations. These health care professionals include physicians, nurse practitioners, pharmacists, and midwives. We asked each health care professional to review each document and respond to the following questions: (1) How do you think this resource will be relevant or useful for you to initiate or strengthen your medication abortion services and why? (2) Please indicate specific parts that you feel are most useful for your practice. (3) Please indicate specific parts that you think are out of date, not useful, or contradict your current practice.

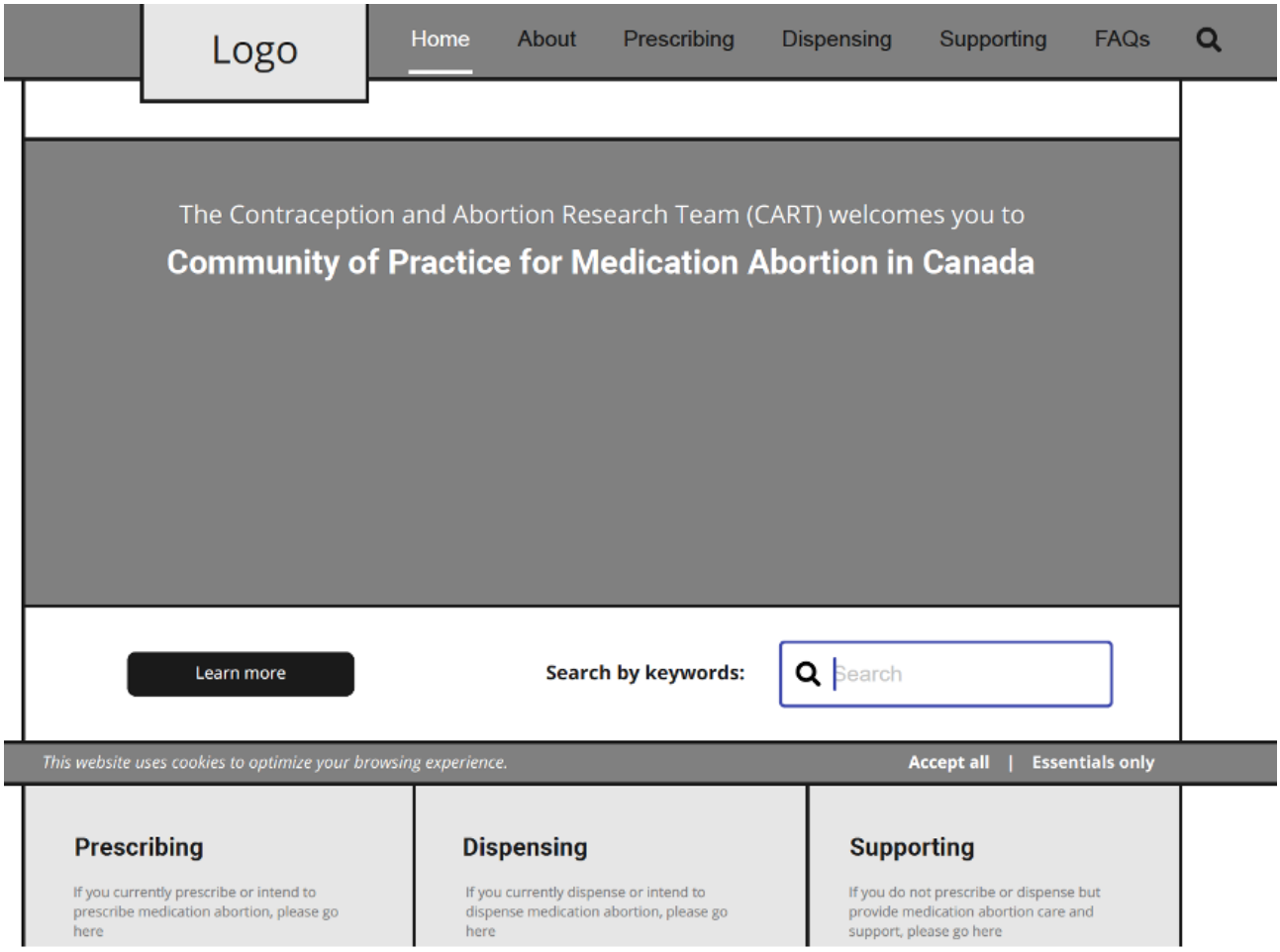
XSL•FO
RenderX

All reviewers provided their feedback within 1 month, and all feedback was compiled into a single document. Following the review, we then engaged 2 experts in abortion care (a family physician and an obstetrician and gynecologist) to discuss and reconcile any differences or contradictory comments by the reviewers. The 2 expert reviewers resolved any differences and wrote a master document containing up-to-date resources on current medication abortion practices. The master file was organized into three main sections including (1) prescriber resources (how medication abortion works, preabortion medication evaluation, postabortion assessment, web-based and hybrid care, physician billing information, regulations, and inclusivity toolkit), (2) dispenser resources (coverage information, patient communication guidelines, checklist for dispensing abortion, and pre- and postabortion assessment), and (3) supporting resources (patient counseling before and after abortion and social worker support resources). This master document was used to inform the content of the website.

Phase 2: Development of Wireframes

Using a web-based design tool known as Figma (Figma Inc), we created wireframes of the website. A wireframe is a visual guide that shows a skeletal framework of a web interface [18]. The wireframe was designed to follow the format of the resources on the master file. The framework depicted the page layout and arrangement of content on the website including interface elements, navigational features, the range of functions to include, and the relative prioritization of content. This wireframe was reviewed by 3 experts in abortion care, and revisions were made regarding space allocation and content prioritization for each of the resource sections outlined in the master file. Following this review, the skeletal framework was then populated with the content from the master file. Figure 3 shows the wireframe of the website.

Figure 3. A screenshot of the wireframe.



Phase 3: Adapting the Wireframes to the Needs of Underserved Populations

After populating the skeletal framework with the revised content, we then conducted a focus group discussion with 14 health care professionals who currently provide, or intend to provide, medication abortion in Canada. The essence of the focus group was to understand how the wireframe can be adapted to equip health care professionals to provide inclusive, gender-affirming, and culturally safe abortions to underserved populations in Canada. We also wanted to identify potential areas for improvement before moving on to developing a functioning website.

Before the start of each focus group, we distributed a visual layout of the website containing relevant content to all participants. The participants were asked to review the visual layouts ahead of the web-based meetings. During the focus group discussions, participants were asked about (1) how the

XSL•FO
RenderX

information on the wireframe could be adapted to support them in providing evidence-based, culturally safe, and accessible abortion services to HPSM populations and (2) any additional features, functionalities, and considerations for ensuring confidentiality, appropriateness, culturally affirmativeness, and supportive resources for improving access and quality of abortion seeking experience for HPSM populations.

The focus groups were led by author AFA, with assistance from RC. Each session lasted approximately 1 hour, and the data were transcribed verbatim. The data from this focus group were analyzed thematically, and the findings (reported under the Results section) were used to modify the Figma designs before moving on to developing a functional product. The findings from the focus groups helped to minimize the design flaws and reduce the possibilities for major changes once a product was eventually developed.

Phase 4: Development of a Functional Website

After making changes to the Figma design, we engaged our software engineers to translate the Figma design into a functioning website. To further explain the Figma design to the development team, we wrote a Systems Requirements Specification Document to further explain the Figma design to the development team. This document included a breakdown of the Figma designs into various components and specifying each design feature as contained in the guide [19]. Figure 4 shows the interface of the functional website. To make sure that the functional product reflects the Figma design and meets the knowledge needs of the end-user clinicians, the development team sought ongoing feedback from our team as well as from the health care professionals who reviewed the resources. We held biweekly consultation meetings with the software development team starting from August 2023 until we had a functional website in February 2024. In compliance with the user-centered design process, we evaluated the functional website with end-user clinicians.

Figure 4. The home page of the functional website.



Phase 5: User Evaluation

We conducted a user evaluation of the website using a think-aloud observation method with 26 health care professionals across Canada. A think-aloud observation is a user evaluation method, in which participants verbalize their thought processes as they interact with a system [20]. The purpose of the user evaluation was to determine the relevance and usefulness of the website in providing abortion services to equity-deserving and HSPM populations, understand the ease of use of the website, detect any design flaws, and make final changes before the website is launched. During the think-aloud observation, we provided participants with 5 tasks related to information retrieval on our website. As participants navigated the website, they verbalized their thought processes by commenting on aspects of the website that they found confusing, difficult to navigate, or irrelevant to their purpose. The think-aloud data were analyzed using simple content analysis. The findings (reported under the Results section) were used to revise the website.

XSL•FO
RenderX

Results

Overview

In this section, we focus on the findings from the focus group discussions of the wireframes as well as the user evaluation of the functional website. Overall, the health professionals who took part in the focus group discussions and the end-user evaluation acknowledged that our web platform was an important educational resource that would equip them with the necessary knowledge to provide abortion care. They also provided recommendations on how this web-based platform could be adapted to support them in providing inclusive abortion care to underserved populations.

Findings From the Focus Group Discussion

Overview

The participants expressed their satisfaction with the layout and the content displayed on the wireframes. They indicated that the various clinical guidelines, factsheets, and checklists on medication abortion would help demystify the myths and misconceptions associated with providing abortion. The usefulness of the resources in demystifying abortion provision was a very important feature considering participants' self-described limited understanding of medication abortion. While acknowledging the importance and potential usefulness of the wireframes, participants noted some limitations on the interfaces of the wireframes and suggested ways in which they could be adapted to improve the capacity of health care professionals to provide abortion services to equity-deserving and HPSM populations. These recommendations fell into 4 thematic areas.

Inclusive and Culturally Sensitive Content

Participants expressed the need for the website to contain inclusive and culturally sensitive language that has the least chance of causing harm to underserved populations including Indigenous, trans, and nonbinary people who seek abortion. To enable health professionals to provide inclusive abortion services to trans people, the participants suggested that the website should contain gender-neutral terms as well as handouts and factsheets on abortion and postabortion care in multiple languages. A physician who provides abortion stated:

> *A lot of my patients can speak English but there are instances [where] I get patients that cannot speak English properly. In such cases, it will be good to have some resources in their language that I can give out to them.*

Some of them also suggested having a print-out function to provide materials for patients who do not speak English. The participants also expressed the need for low bandwidth connectivity to enable health care providers in rural areas with poor internet access to access resources.

Indigenous and Gender-Diverse Resources

The participants also called for the inclusion of Indigenous and 2SLGBTQI+ resources on the website because they indicated that one of the groups facing the most barriers to accessing health services includes Indigenous groups and the 2SLGBTQI+ populations. This could be seen in one of the statements given by one of the participants:

> *it isn't just about understanding the clinical guidelines and factsheets. But we also need to understand the other aspects of Indigenous cultural safety when providing abortion to Indigenous patients like the relationship between abortion care and legacies of forced apprehension/the millennial scoop, information on the knowledge and risk of reproductive coercion (either for or against abortion) etc.*

These calls were particularly made by participants who provide services to Indigenous communities. Other participants indicated a need for a learning hub or short courses on Indigenous cultural safety, gender-affirming, and trauma-informed abortion care—arguing that these additional knowledge sources will help them to provide services to abortion seekers who experienced historical and intergenerational traumas.

Providing Resources for Noninsured Clients

The participants also indicated how the website could be a "one-stop shop" of information if they included information on how to support noninsured clients. This recommendation was made by a midwife who revealed the challenges of providing abortion services for noninsured clients such as undocumented migrants:

> *More than half of my clients were unhoused, or street-involved or living in extreme poverty, or they were refugees, recently landed immigrants, or people of color and trans communities. You know there is a significant intersection between such populations and those living in poverty so there needs to be some resources around how to make sure that these populations are also getting access to abortion services.*

Protecting the Privacy of Website Users and Abortion Seekers

The participants acknowledged the stigmatizing nature of abortion and the need to protect people's privacy. Thus, they suggested including a built-in exit button that would enable users to immediately exit the website, particularly in public places. They also indicated the need for the website to contain geolocation features that can map out nearby abortion providers and pharmacies outside of people's communities. One participant stated:

> *I provide services in a local community, and some don't feel comfortable seeking abortion services from me because I may probably know them. When this happens, it's difficult to refer because they won't even come to you. So, it will be good to have a feature on this platform where patients themselves can search for nearby abortion providers that they can approach with comfort.*

According to the participants, this will make visible nearby abortion providers that could serve as alternative sources for patients who may be reluctant to seek abortion services within

XSL•FO
RenderX

their (particularly very small) communities due to privacy concerns.

Findings From the User Evaluation

Overview

Iterative feedback from the focus group informed changes to the website, as participants serially identified gaps and problems to be addressed before it was launched. Some health care professionals indicated that the functional website felt too medical and was insufficiently focused on the social aspects of abortion. They perceived our banner image on the home page as overrepresenting physicians and not sufficiently acknowledging the role of midwives and supporting professionals. They also made recommendations on how the website could be adapted for navigability and to touch on the stories of underserved populations. The recommendations from the user evaluation fell into 3 thematic areas.

Portraying a Positive Image of Abortion via Visual Effects

Some health care professionals commented that the initial wireframes lacked a human touch and thus suggested adding more vibrant and bright colors that reflect the identities of diverse groups of health professionals and the diversity of patients. Participants felt that such adaptations could portray abortion as something that is considered normal for everyone to seek, normal for every health professional to provide or support, and to ensure the abortion-seeking experience is positive.

> *maybe some colors will be nice because I feel that when I am looking into abortion care, I just want to take this stigma away. So, having colors that are dark for me is like you are doing something bad, right? So maybe having some lighter colors and highlighting on socio-cultural aspects of abortion would be helpful. People need to feel that abortion is normal after going through this website.*

Connecting With Patients

Other health professionals explained that although the website targets health care professionals, it is ultimately about caring for people. They thus suggested that it would be great to have images of health care providers taking care of patients from diverse gender and racial backgrounds or better still include images of patients having children to remove the myth that people seeking abortion are irresponsible adults without children or families. One participant stated:

> *This website is about people getting care. So, we need to see how so many people who have pregnancy termination also have children already. So, it's not just like the unwed single woman who wants to terminate her pregnancy.*

The participants also expressed the need for the website to help health care professionals understand the patient populations most likely to seek abortion services from them. This way, the website would not only provide clinical information but also help providers to connect with the patients' stories.

Making Compromises Between Being Comprehensive and Practical

Many participants acknowledged the comprehensiveness of the website and were confident that they would be able to find relevant information should the need arise. They also acknowledged that being comprehensive might make the website a less suitable quick reference guide, particularly during clinician-patient encounters. This concern could be a challenge for high-volume providers and providers in rural and remote areas. One midwife stated:

> *the information that I'm seeing is actually really good. It's really good information, and I'm intrigued to read more about what's in here. But [it's] too much. It would be really good for like education and training, but not when you have a patient in front of you.*

While this theme was not specifically related to providing abortion to underserved populations, it highlighted the busy schedule of clinicians vis-à-vis looking for information in a dense and cluttered web platform.

Discussion

Principal Findings

In this study, we report on the cocreation of an open-access web-based abortion platform and highlight considerations for promoting access to abortion for underserved populations. This study followed an intersectionality study we conducted with end-user clinicians before embarking on the development of the website [21]. Our study incorporated rigorous engagement of end-user clinicians, and our user evaluation data demonstrated an accessible, acceptable website that could better equip health and allied helping professionals to provide equitable, culturally safe, and gender-affirming resources to equity-deserving and HSPM populations. It is important to note that many health professionals in Canada, including nurse practitioners and midwives, had not received any formal training on abortion as part of their education, leading to limited knowledge and general uncertainties in providing abortion [22]. The limited knowledge may lead to some hesitation in providing abortion, particularly for patients from underserved populations who present with complex health challenges. We believe that this mentorship platform will equip health professionals with the necessary knowledge to provide abortion to patients who face intersecting barriers to access services while also enhancing timely decision-making for health professionals in rural and remote areas. For a time-sensitive procedure such as abortion, expeditious decision-making is crucial to high-quality care for women and pregnancy-capable people who request abortion services.

In addition to information dissemination and helping with decision-making, an open-access abortion web platform may also serve to counter abortion-related misinformation and disinformation that are often targeted at health care professionals. This is important at a time when the proliferation of disinformation, together with the stigma and belief-based refusal to provide legal and appropriate abortion care, creates interpersonal tensions that make it difficult for some health

XSL•FO
RenderX

professionals to meet the reproductive needs of their patients [23,24]. Health professionals may further be disadvantaged when there are frequent changes in medication abortion guidelines and regulations [17]. These changes, together with the need for quality improvement, highlight the importance of such a web-based platform for updating the competencies of health care professionals who provide or intend to provide abortion. This is more relevant, as there is an increasing call for the timely integration of research evidence in routine clinical practice [25].

Our findings suggest that adopting a user-centered design in this project resulted in a platform that was largely considered by potential end-user clinicians as inclusive. While engaging stakeholders from diverse clinicians and organizations was quite challenging, we acknowledged that the process was essential to fulfill the principles of user-centered design. Our team believes that the active engagement of clinicians, health professional organizations, and abortion advocacy groups in cocreating this web-based platform enhanced the appropriateness of the content while facilitating clinician recruitment, dissemination, and subsequent uptake of the resource among end-user health care professionals.

It is also important to note that few digital health projects are designed from gender equity and Indigenous rights perspective [26]. In the context of abortion-related websites, intersectionality is particularly important, as it helps illuminate ways in which the website could be tailored to the needs of the population subgroups facing the most barriers in access to abortion (racialized groups, migrants, people with disabilities, homeless and underhoused people, sex workers, 2SLGBTQI+ people, Indigenous people, and youths) [21,27]. Indeed, available evidence suggests that such minority groups are overrepresented among those with unintended pregnancies and particularly among those seeking abortion [28-30]. Therefore, engaging with a group of health professionals who serve diverse populations in developing and evaluating our web platform made the website unique, as it cedes all decisions on the content and structure in the hands of the end-user clinicians who may have first-hand experience in providing services to underserved populations. This approach to developing our website supports the vision of “design justice” where the community owns the design artifacts [31]. By cocreating with health care professionals, we were able to generate website content that illuminates ways health care professionals can provide abortion services to underserved populations who present with complex, historical, and intersecting health challenges. This approach is more likely to lead to a seamless translation of the web platform into practice since the clinicians and partner organizations were more or less coowners of the end product.

Creating web platforms that are inclusive of equity-deserving and HSPM populations requires more than just technically sophisticated algorithms but a deep understanding of users’ needs and specific considerations that promote inequities in access to health care [32]. The inclusive design recommendations from the participants in this study were aimed at equipping health care professionals with the necessary tools to address inequities in access to abortion care. The integration of low-bandwidth connectivity for instance would improve accessibility of the resources to health professionals in rural and remote areas with low internet connectivity. Improved web access could help improve abortion access in rural and remote areas where medication abortion might be the only, or most accessible, option [33]. Furthermore, the recommendation for including geolocation services could help address abortion-related stigma and provide better privacy for abortion seekers in communities with conservative values around abortion. While the argument for geolocation features was seen as a way of helping patients easily locate abortion providers and pharmacists, some were concerned that geolocation services could be used by antichoice groups to target abortion providers. Additional security measures may be needed if geolocation features are to be implemented on abortion-related web platforms.

Limitations

Our study was not without limitations. Even though our objective was to design a web platform that supports health care professionals to provide equitable access to abortion for underserved populations, we did not rigorously engage abortion seekers the same way we engaged with clinicians. A future improvement would be to engage the patients’ voices in resource development to better reflect the peculiar needs of underserved populations. Our ultimate goal for this website is to empower health care professionals to provide equitable and inclusive abortion to underserved populations. However, we cannot ascertain if we have achieved that objective at the time of writing this paper. Future evaluation reports from users would tell us if the resources provided in this platform are indeed useful in providing abortion services to underserved populations. Furthermore, the participants made several recommendations in the focus group and the user evaluation that were not implemented mainly due to time and resource constraints. While we acknowledge this as a limitation, it is not uncommon to see such limitations in digital health projects, as it is practically impossible to implement all findings from user evaluation [34,35]. Finally, even though we planned to recruit health professionals who identify as or have expertise in providing abortion services to underserved populations (ie, racialized people, migrants, people with disabilities, homeless and underhoused people, sex workers, 2SLGBTQI+ people, and youths as well as those living in or providing abortion services in rural or remote areas), it was difficult to find health professionals who identify as or has expertise in all these diverse populations. Therefore, some phases were completed without adequate input from health professionals with expertise or experience in providing abortion services to underserved populations.

Conclusions

Drawing on a user-centered design approach, we cocreated a gender-affirming and culturally appropriate open-access mentorship platform with and for health care professionals to support the equitable provision of abortion in Canada. We believe that a platform of this nature would increase medication abortion awareness among health care professionals and destigmatize abortion care while clarifying the myths and misconceptions that often create tensions and foment general

XSL•FO
RenderX

uncertainties in initiating and sustaining abortion care. With underserved populations more likely to have unintended pregnancies than the general population [36,37], a website that specifically seeks to improve the competencies of health care professionals in providing abortion services to this population is important and timely. We believe that this platform has the potential to enhance the knowledge and expertise of health care professionals, particularly those in rural and remote areas on how to initiate and sustain abortion care for HPSM populations in Canada who face the most barriers in access to health services. This platform is also expected to provide the necessary resources needed to support health care professionals in providing evidence-based and culturally safe medication abortion for health care professionals in rural and remote areas.

Acknowledgments

The funding for this project was received from the Health Canada Sexual and Reproductive Health Fund under the Health Care Policy and Strategies Program. The authors would also like to acknowledge the Health Professional Association (ie, Society of Obstetricians and Gynaecologists of Canada, Canadian Pharmacy Association, Canadian Association of Midwives, and Canadian Association of Schools of Nursing), advocacy groups (Action Canada for Sexual Health and Rights and National Abortion Federation), and health care professional participants who provided diligent reviews of the resources.

Conflicts of Interest

None declared.

References

1. Sexual and reproductive health and rights. Office of the High Commissioner for Human Rights. URL: https://www.ohchr.org/en/women/sexual-and-reproductive-health-and-rights [accessed 2024-06-05]
2. Black AY, Guilbert E, Hassan F, et al. The cost of unintended pregnancies in Canada: estimating direct cost, role of imperfect adherence, and the potential impact of increased use of long-acting reversible contraceptives. J Obstet Gynaecol Can 2015 Dec;37(12):1086-1097. [doi: 10.1016/s1701-2163(16)30074-3] [Medline: 26637081]
3. Dunn S, Cook R. Medical abortion in Canada: behind the times. CMAJ 2014 Jan 7;186(1):13-14. [doi: 10.1503/cmaj.131320] [Medline: 24277708]
4. Schummers L, Norman WV. Abortion services in Canada: access and safety. CMAJ 2019 May 13;191(19):E517-E518. [doi: 10.1503/cmaj.190477] [Medline: 32392474]
5. Monchalin R, Jubinville D, Pérez Piñán AV, et al. “I would love for there not to be so many hoops … ”: recommendations to improve abortion service access and experiences made by Indigenous women and 2SLGTBQIA+ people in Canada. Sex Reprod Health Matters 2023 Dec;31(1):2247667. [doi: 10.1080/26410397.2023.2247667] [Medline: 37799036]
6. Monchalin R, Pérez Piñán AV, Wells M, et al. A qualitative study exploring access barriers to abortion services among Indigenous Peoples in Canada. Contraception 2023 Aug;124:110056. [doi: 10.1016/j.contraception.2023.110056] [Medline: 37116606]
7. Sethna C, Doull M. Spatial disparities and travel to freestanding abortion clinics in Canada. Womens Stud Int Forum 2013 May;38:52-62. [doi: 10.1016/j.wsif.2013.02.001]
8. Ennis M, Renner R, Olure B, Begun S, Norman WV, Munro S. Provision of care to diverse populations: results from the 2019 Canadian Abortion Provider Survey. BMJ Sex Reprod Health 2024 Oct 15;50(4):308-309. [doi: 10.1136/bmjsrh-2023-202175] [Medline: 38729768]
9. Decolonizing reproductive justice—indigenous voices. Abortion Rights Coalition of Canada. URL: https://www.arcc-cdac.ca/activist-workshops/decolonizing-rj-indigenous-voices/ [accessed 2024-06-06]
10. Monchalin R. Novel coronavirus, access to abortion services, and bridging Western and Indigenous knowledges in a postpandemic world. Womens Health Issues 2021;31(1):5-8. [doi: 10.1016/j.whi.2020.10.004] [Medline: 33371932]
11. Georgsson S, Carlsson T. Readability of web-based sources about induced abortion: a cross-sectional study. BMC Med Inform Decis Mak 2020 Jun 5;20(1):102. [doi: 10.1186/s12911-020-01132-y] [Medline: 32503524]
12. Anderson NS, Norman DA, Draper SW. User centered system design: new perspectives on human-computer interaction. Am J Psychol 1988 ;101(1):148. [doi: 10.2307/1422802]
13. CAPS CPCA Forum. URL: https://caps-cpca.ubc.ca/index.php?title=Main_Page [accessed 2023-09-07]
14. AusCAPPS | Medcast. URL: https://medcast.com.au/communities/auscapps [accessed 2024-06-06]
15. Munro S, Wahl K, Soon JA, et al. Pharmacist dispensing of the abortion pill in Canada: diffusion of innovation meets integrated knowledge translation. Implement Sci 2021 Aug 3;16(1):76. [doi: 10.1186/s13012-021-01144-w] [Medline: 34344393]
16. Stirling-Cameron E, Carson A, Abdulai AF, et al. Nurse practitioner medication abortion providers in Canada: results from a national survey. BMJ Sex Reprod Health 2024 Dec 30. [doi: 10.1136/bmjsrh-2024-202379] [Medline: 39414352]

XSL•FO
RenderX

17. Munro S, Guilbert E, Wagner MS, et al. Perspectives among Canadian physicians on factors influencing implementation of mifepristone medical abortion: a national qualitative study. Ann Fam Med 2020 Sep;18(5):413-421. [doi: 10.1370/afm.2562] [Medline: 32928757]
18. What is wireframing? The Interaction Design Foundation. URL: https://www.interaction-design.org/literature/topics/wireframe?srsltid=AfmBOoo2DUHZGcYHDMRnK52HW6gksg6tgPoUvH6MLDMpQFe3BS_-q1xy [accessed 2023-09-13]
19. How to write system requirement (SRS) documents. Visure Solutions. URL: https://visuresolutions.com/requirements-management-traceability-guide/how-write-system-requirement-documents/ [accessed 2024-01-24]
20. Boren MT, Ramey J. Thinking aloud: reconciling theory and practice. IEEE Trans Profess Commun 2000;43(3):261-278. [doi: 10.1109/47.867942]
21. Abdulai AF, Czerniak E, Duong C, et al. Intersectionality in developing a virtual community of practice platform on abortion. Stud Health Technol Inform 2024 Jul 24;315:392-397. [doi: 10.3233/SHTI240176] [Medline: 39049289]
22. Paynter M, Norman WV, Martin-Misener R. Nurses are key members of the abortion care team: why aren't schools of nursing teaching abortion care? CJCND 2019 Dec 10;1(2):17-29. [doi: 10.25071/2291-5796.30]
23. Pagoto SL, Palmer L, Horwitz-Willis N. The next infodemic: abortion misinformation. J Med Internet Res 2023 May 4;25:e42582. [doi: 10.2196/42582] [Medline: 37140975]
24. Ennis M, Renner RM, Olure B, et al. Experience of stigma and harassment among respondents to the 2019 Canadian abortion provider survey. Contraception 2023 Aug;124:110083. [doi: 10.1016/j.contraception.2023.110083] [Medline: 37263373]
25. Johnston HB, Footman K, Ali MM, et al. Measuring availability of and facility readiness to deliver comprehensive abortion care: experiences and lessons learnt from integrating abortion into WHO's health facility assessments. BMJ Glob Health 2024 Aug 9;8(Suppl 4):e015097. [doi: 10.1136/bmjgh-2024-015097] [Medline: 39122445]
26. Figueroa CA, Luo T, Aguilera A, Lyles CR. The need for feminist intersectionality in digital health. Lancet Digit Health 2021 Aug;3(8):e526-e533. [doi: 10.1016/S2589-7500(21)00118-7] [Medline: 34325855]
27. Price K. It's not just about abortion: incorporating intersectionality in research about women of color and reproduction. Womens Health Issues 2011;21(3 Suppl):S55-S57. [doi: 10.1016/j.whi.2011.02.003] [Medline: 21530842]
28. Mark A, Foster AM, Aguilar G, et al. Access for all: research at the National Abortion Federation's 47th Annual Meeting. Contraception 2024 Jun;134:110402. [doi: 10.1016/j.contraception.2024.110402]
29. Foster AM, Darney BG, Nguyen BT, Russo J, Arora KS, Westhoff C. Advancing diversity, equity, inclusion, and anti-racism in contraception: a statement of commitment from the Editorial Board. Contraception 2023 Jul;123:110066. [doi: 10.1016/j.contraception.2023.110066] [Medline: 37207728]
30. Hulme J, Dunn S, Guilbert E, Soon J, Norman W. Barriers and facilitators to family planning access in Canada. Healthc Policy 2015 Feb;10(3):48-63. [Medline: 25947033]
31. Udoewa V. An introduction to radical participatory design: decolonising participatory design processes. Des Sci 2022;8:e31. [doi: 10.1017/dsj.2022.24]
32. Inclusive technology: design for underserved and marginalized communities. The Department of Computer Science at the University of Chicago. URL: https://www.classes.cs.uchicago.edu/archive/2020/winter/20370-1/ [accessed 2024-04-27]
33. Sandoval S, Chen G, Rafie S, Rabin B, Mody S, Averbach S. A qualitative analysis of pharmacists' attitudes towards provision of medication abortion. BMC Health Serv Res 2023 May 30;23(1):557. [doi: 10.1186/s12913-023-09543-z] [Medline: 37254171]
34. Chapter 10: MoSCoW prioritisation. Agile Business Consortium Limited. URL: https://www.agilebusiness.org/dsdm-project-framework/moscow-prioririsation.html [accessed 2024-01-06]
35. Lehtola L, Kauppinen M, Kujala S. Requirements prioritization challenges in practice. In: Bomarius F, Iida H, editors. Product Focused Software Process Improvement: Springer; 2004:497-508. [doi: 10.1007/978-3-540-24659-6_36]
36. Reynolds CA, Beccia A, Charlton BM. Multiple marginalisation and unintended pregnancy among racial/ethnic and sexual minority college women. Paediatr Perinat Epidemiol 2021 Jul;35(4):493-500. [doi: 10.1111/ppe.12744] [Medline: 33345309]
37. Oulman E, Kim THM, Yunis K, Tamim H. Prevalence and predictors of unintended pregnancy among women: an analysis of the Canadian Maternity Experiences Survey. BMC Pregnancy Childbirth 2015 Oct 13;15(1):260. [doi: 10.1186/s12884-015-0663-4] [Medline: 26462914]

Abbreviations

2SLGBTQI+: two-spirit, lesbian, gay, bisexual, transgender, queer, intersex, and gender-diverse
HPSM: historically, persistently, or systemically marginalized

XSL•FO
RenderX

Edited by A Kushniruk; submitted 19.06.24; peer-reviewed by L Patchen, WR Dullius; revised version received 31.12.24; accepted 05.01.25; published 19.02.25.

Please cite as:
Abdulai AF, Duong C, Stroulia E, Czerniak E, Chiu R, Mehta A, Koike K, Norman WV
A Health Professional Mentorship Platform to Improve Equitable Access to Abortion: Development, Usability, and Content Evaluation
JMIR Hum Factors 2025;12:e63364
URL: https://humanfactors.jmir.org/2025/1/e63364
doi:10.2196/63364

© Abdul-Fatawu Abdulai, Cam Duong, Eleni Stroulia, Efrat Czerniak, Rachel Chiu, Aashay Mehta, Ken Koike, Wendy V Norman. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 19.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Developing an App for Real-Time Daily Life Observations in a Nursing Home Setting: Qualitative User-Centered Co-Design Approach

Coen Hacking, MSci; Bram de Boer, PhD; Hilde Verbeek, PhD; Jan Hamers, PhD; Sil Aarts, PhD

Department of Health Services Research, Maastricht University, Duboisdomein 30, Maastricht, the Netherlands

Corresponding Author:
Bram de Boer, PhD
Department of Health Services Research, Maastricht University, Duboisdomein 30, Maastricht, the Netherlands

Abstract

Background: Assessing the daily lives of older adults, including their activities, social interactions, and well-being is essential, particularly in nursing homes, as it gains insights into their quality of life. Methods such as the Microsoft Excel-based Maastricht Electronic Daily Life Observation (MEDLO) tool are time-consuming and require extensive manual input, making them difficult to use.

Objective: This study aimed to develop an app-based version of the MEDLO using a user-centered design (UCD) and co-design approach to enhance efficiency and usability. We looked to actively involve researchers and care professionals who have used the MEDLO before, throughout the development process.

Methods: Participants included a diverse group of researchers and care professionals experienced in using the MEDLO tool. The UCD approach involved multiple iterative phases including semistructured interviews, user research sessions, and application development. Data were analyzed using a qualitative (thematic) approach of UCD and user research sessions. The app, which was preferred to the traditional Excel-based MEDLO, underwent multiple iterations. This method primed the continuous iterative development of the app, aimed for a minimum viable product (MVP).

Results: This study included 14 participants, primarily female, from diverse professional backgrounds. Their feedback highlighted the need for efficiency improvements in tool preparation and data management. Key improvements included automated data handling, an intuitive tablet interface, and functionalities such as randomization and offline data syncing.

Conclusions: The iterative development process led to an app that aligns with end-user needs, indicating potential for improved usability. Early and continuous user involvement was key in enhancing the application's usability, demonstrating the importance of user feedback in the development process.

(JMIR Hum Factors 2025;12:e57911) doi:10.2196/57911

KEYWORDS

co-design; user-centered design; app development; nursing home; user-centered; design; efficiency; usability; tablet; mobile phone

Introduction

Assessing the daily lives of older adults, particularly in the context of nursing homes, is essential for gaining insights into quality of life [1,2]. It allows health care providers to detect and address issues related to their physical health (eg, mobility limitations, pain), mental well-being (eg, mood disturbances, cognitive decline), and social interactions (eg, isolation, engagement in activities) in a timely manner [1,3,4]. This information can then be used to personalize care plans, thereby potentially improving health outcomes and enhancing the overall quality of life of older adults [5,6].

Ecological momentary assessments (EMAs) have been developed to facilitate this process by providing real-time, on-site evaluations of an individual's well-being [5,7,8]. EMA involves collecting data on individuals' behaviors and experiences in their natural environments, which can then be used to identify patterns and inform interventions [7]. However, existing EMA tools are mainly used in research settings, and their implementation in clinical practice has been challenging due to the time-consuming nature of data collection and the complexity of the tools [9,10]. The Maastricht Electronic Daily Life Observation tool (MEDLO-tool) is designed to assess the daily lives of nursing home residents using EMA methodologies [1,4]. The tool captures real-time information across several key dimensions of daily life, including activity levels (eg, participation in communal or solitary activities), agitation, mood states, and interactions with staff, fellow residents, and visitors [1,4,11]. By systematically observing and recording these

XSL•FO
RenderX

aspects, the MEDLO-tool provides a nuanced view of residents' experiences, which can inform personalized care strategies [4]. However, the MEDLO-tool relies on Excel templates that require significant manual input and are not user-friendly for care professionals [1,4]. The complexity of the Excel-based system poses significant challenges, including data entry errors, time inefficiency, and barriers to widespread adoption in clinical settings [12,13]. This complexity can be particularly problematic in nursing homes, where staff may have limited time and technological proficiency [13].

Developing a mobile application for the MEDLO-tool could address these challenges by automating data collection processes, reducing manual input, and providing an intuitive interface for users [14,15]. An app could streamline observations, enable real-time data analysis, and facilitate immediate feedback to care providers, thereby enhancing the tool's use in clinical practice. Furthermore, considering the unique needs of residents with dementia, it is important that such an app is designed to account for cognitive impairments and communication difficulties [16,17]. The development of such an app requires a thorough understanding of the needs and preferences of the end-users. A user-centered design (UCD) approach is vital for developing software that truly meets users' needs [14,18]. UCD emphasizes involving users throughout the design process to ensure that the product aligns with their requirements and preferences [19]. This involves iterative cycles of user need assessment, design, prototyping, and testing [20].

Co-design, a component of UCD, involves direct collaboration between users and designers, requiring the active participation of users in generating ideas, making decisions, and solving problems [21,22]. This not only ensures the software's functionality but also its usability [21].

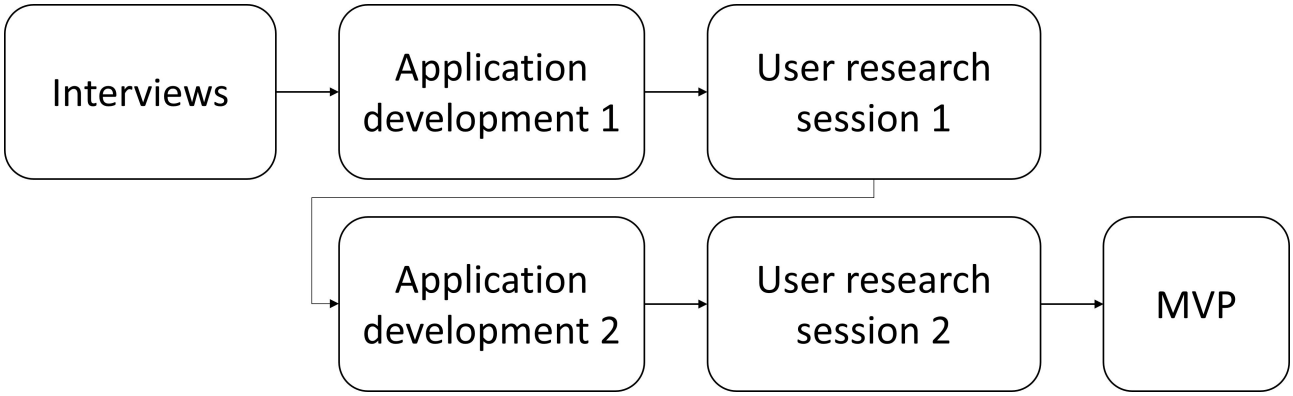
Potential barriers to implementing such technology include technological literacy among staff, user engagement, and ensuring data privacy and security [13,23]. By actively involving care professionals and stakeholders in the design process, we aim to create a tool that is both functional and user-friendly, facilitating its adoption in clinical practice and ultimately improving the quality of care for residents with dementia. Therefore, this study is aimed at developing an app for measuring the daily life of residents with dementia living in nursing homes, using a user-centered and co-design approach.

Methods

Study Design

This methodological study aimed to develop a mobile application version of the MEDLO-tool for use in long-term care facilities, using a UCD approach [24-26]. The study unfolded in multiple iterative phases, including semistructured interviews, user research sessions, application development, and comprehensive data analysis, as illustrated in Figure 1. The project culminated in the achievement of a minimum viable product (MVP).

Figure 1. The iterative process of app development. MVP: minimum viable product.



Participants

Participants were selected through purposive sampling to include individuals with experience using the MEDLO-tool. The sample consisted of researchers and health care professionals affiliated with Maastricht University and local nursing homes. Inclusion criteria were previous use of the MEDLO-tool and willingness to participate in the study. A total of 14 participants were recruited, comprising research assistants, PhD candidates, postdoctoral researchers, senior researchers, occupational therapists, research coordinators, and associate professors.

Data Collection

Data collection involved semistructured interviews and user research sessions conducted between January and June 2023.

Interviews

Semistructured interviews were guided by a comprehensive list of topics derived from available literature and consultation with the researchers who developed the original MEDLO-tool [1,4]. The interview guide can be found in Multimedia Appendix 1. The interviews aimed to gather insights into participants' experiences with the Excel-based MEDLO-tool, usability issues, functionality requirements, data analysis techniques, and general impressions [27]. Each interview lasted approximately 45 to 60 minutes and was audio-recorded with participants' consent. The audio was transcribed verbatim for analysis. The interview guide is provided in the supplementary materials.

The interview transcriptions were subjected to a 6-phase thematic analysis process, aligning with the methodologies supported by previous research [28-30]. The inductive analysis was conducted by one researcher. A second researcher reviewed

XSL•FO
RenderX

the coding to enhance trustworthiness. This comprehensive process involved familiarization with the data, generating initial codes, identifying themes, reviewing themes, defining and naming themes, and ultimately, reporting the results. The resulting report summarized the key findings, interpretations, and implications, to provide valuable insights for the next stages of the development of the app.

User Research Sessions

A total of 2 user research sessions were conducted monthly to involve participants in the iterative development process. The session guide can be found in Multimedia Appendix 2. Each session lasted approximately 2 hours and was held at Maastricht University. Participants interacted with app prototypes using tablets and smartphones provided by the research team. The sessions included guided tasks (eg, completing an observation using the app), open-ended discussions, and interactive feedback activities facilitated by researchers using prototypes and interactive demos. The structure of these sessions was informed by previous co-design methodologies [1,3,31].

Data Analysis

Qualitative Analysis of Interviews

Interview recordings were transcribed verbatim and analyzed using thematic analysis following Braun and Clarke's 6-phase approach [28]. A total of 2 researchers independently coded the transcripts to enhance reliability. Initial codes were generated and organized into potential themes. Discrepancies in coding were discussed and resolved through consensus meetings. Themes were reviewed and refined to ensure they accurately represented the data. The final themes captured key insights into the usability and functionality of the MEDLO-tool, informing the app development. The overall process of thematic analysis was guided by earlier works in this field [29,30]. The session guide is provided in Multimedia Appendix 1.

User Research Sessions

Each user research session included guided tasks (eg, completing an observation using the app), open-ended discussions, and interactive feedback activities facilitated by researchers using mock-ups and interactive demos. Feedback from user research sessions was documented through field notes and audio recordings. An inductive content analysis was conducted by one researcher, who coded feedback and categorized issues into 4 themes: layout, functionality, errors, and appearance. A second researcher reviewed the coding to enhance trustworthiness. This categorization allowed the development team to prioritize user concerns based on frequency and severity, ensuring a user-centric approach to app refinement.

Application Development

The application was developed using ".NET MAUI" (ie, a framework designed by Microsoft) for cross-platform compatibility on iOS and Android devices. The server-side (ie, where the data are sent to) used "ASP.NET Core" (ie, a framework designed by Microsoft) for a maintainable and performant application. The development process was iterative, with weekly meetings between developers and researchers to integrate user feedback from the interviews and user research sessions.

Usability testing involved participants completing specific tasks using the app prototypes while researchers observed and noted any issues. Testing sessions were conducted in a controlled environment to simulate actual usage conditions. Feedback from these tests informed further refinements and addressed potential adoption or feasibility issues.

Ethical Considerations

This study did not require formal ethics approval under the Medical Research Involving Human Subjects Act (WMO). In the Netherlands, research that collects only anonymized or nonsensitive feedback—without requiring participants to undergo procedures that affect their physical or psychological integrity—is not considered WMO research and is therefore exempt from ethics review. The guidelines provided by the Central Committee on Research Involving Human Subjects clarify that retrospective or noninvasive studies using aggregated or anonymized data do not fall under the scope of the WMO [32]. The Maastricht University Medical Center+ Research Code also confirms that studies conducted with minimal risk, such as those analyzing collective feedback, do not require formal ethical review. All participant feedback was anonymized and is presented in aggregate form to ensure confidentiality [33].

Results

Sample

The study involved 14 participants with an average age of 36 (SD 10; median 39; range 24 - 57) years, and 93% (13/14) of them were female. The participants represented a diverse range of professions, including research assistants (n=3), PhD candidates (n=4), postdoctoral researchers (n=2), a senior researcher (n=1), an occupational therapist (n=1), a research coordinator (n=1), an associate professor (n=1), and an individual who was currently unemployed. Most participants (9/14, 64%) were affiliated with Maastricht University, while others worked at health care providers (4/14, 29%) or were unemployed (1/14, 7%). The average duration of employment at their respective institutions was 6 (SD 5; median 6; range 0.5 - 16) years.

All 14 participants had previous experience with the MEDLO-tool; 12 (86%) participants had used it directly in their research, while 2 (14%) participants had used it as inspiration for developing other tools different from MEDLO. All participants had also been in contact with the developers of the MEDLO-tool. Table 1 presents the demographic information of the participants in detail.

XSL•FO
RenderX

Table . Demographic information of the participants.

Characteristic		Values
Age (years), mean (SD)		
	24 - 57	36 (7.7)
Gender, n (%)		
	Female	13 (93)
	Male	1 (7)
Time at institution, mean (SD)		
	6 months to 16 years	6 (3.3)
Use of the tool, n (%)		
	Used in research	12 (86)
	Inspiration for another tool	2 (14)
Contact with developers, n (%)		
	All participants	14 (100)
Participants' affiliations, n (%)		
	University	9 (64)
	Health care provider	4 (29)
	Unemployed	1 (7)

Interviews

The participants described the process of using the MEDLO-tool in research projects as comprising 7 distinct phases: (1) acquiring informed consent, (2) preparing the Excel sheets before the observation, (3) familiarizing themselves with the faces of the residents who would be observed, (4) conducting the observations, (5) fixing issues with the data (such as missing values or inconsistencies), (6) analyzing the data, and (7) communicating the data back to the care organizations.

Participants emphasized that the preparation of the tool and the subsequent data cleaning were particularly time-consuming and would significantly benefit from automation. One participant stated, "Every time before we do a new observation, we have to look at the participants, randomize them several times, and input all of this into the Excel sheet." Another participant highlighted the challenges faced after data collection, "After the observation is done, I have to go through the whole document to check all the fields that I didn't have time to fill out or that I didn't know how to rate." These comments reflect the manual and labor-intensive nature of the existing process, which participants found cumbersome.

Participants reported that during observations, they used separate pieces of paper to write down descriptions of each resident's appearance because the observation periods were too short to input this information into the Excel sheet in real time. They suggested that having an additional field to securely write this information directly on the tablet would make the process more accessible. One participant explained, "We usually arrive at least half an hour earlier, so we can ask the nurses which participant corresponds to which name, and write down what they look like. We do that to find them back more easily during the observation." Another participant elaborated on the logistical challenges, "The data entry process on the tablet is intuitive, but we also have to keep track of the patient names and descriptions on paper, we carry the manual, and we use a timer on our phone to ensure that we don't spend too long on any resident. This is a lot to carry around." Regarding data reporting, participants indicated that the data reported back to the nursing homes primarily consisted of quantitative measures, which could potentially be automated through a dashboard (eg, Microsoft PowerBI). Participants expressed that automating this process would reduce the time between data collection and reporting, which currently could be several months. One participant noted, "When we want to report back the numbers to the care organization, we make a PowerPoint presentation that includes aggregated numbers. This usually takes place several months after the observations." Another participant suggested: "I think some of the numbers could be reported back automatically."

Overall, participants preferred using a tablet compared to using pen and paper, citing the intuitive data entry process. However, they also mentioned that a tool compatible with smartphones would be beneficial, especially for care professionals who might find tablets less convenient. One participant remarked, "I think the tool works well, but I'm not sure if it would be usable like this for nurses."

Initial User Research Session

Based on the information gathered during the interviews and the suggested points for improvement, an initial version of the application was created and provided to the participants for the first user research session. This first draft of the application consisted of an dashboard web application and a mobile app, both of which included the core components but had limited functionality. Multimedia Appendix 3 shows the app and dashboard as they were presented to the participants. Multimedia

XSL•FO
RenderX

Appendix 4 provides a detailed overview the feedback received during the initial user research session.

During the initial user research session, participants expressed largely positive sentiments towards the concept of the application and appreciated the efforts made in its development. One participant commented, "I know we have a lot of complaints about the app, but we really appreciate the work [the development team] has done already." However, participants also identified several areas for improvement. They expressed concerns regarding certain features of the app, such as the timer functionality, which lacked a reset option. Participants noted that in their current practice, they used interval timers on their phones to manage observation times for each resident, and the app's timer did not adequately support this need. One participant suggested, "What could be nice is some kind of interval timer. That's what we currently use on our phones." In addition, participants missed the comprehensive overview provided by the Excel-based MEDLO-tool, as the prototype app had been designed primarily for phone screens and did not offer the same level of data visibility. One participant observed, "Normally, you could easily scroll back in the overview." Furthermore, there were some unclear aspects of the app, such as certain functionalities being available only through the dashboard and not within the app itself. Participants found labels on buttons to be unclear, which led to confusion during navigation. They recommended improving the screen layout for better usability and enabling easy switching between participants by a list.

Participants raised concerns about data storage and security, even though these aspects had been addressed in the app's design. They questioned how the data was stored and whether it was encrypted. It was explained to them that the data was stored on the device (ie, phone or tablet) in encrypted form and could be synced to a secure server. Regarding randomization, participants asked how it would be conducted within the app and where they could access this feature. It was clarified that randomization was performed automatically upon creating a new observation group. A participant inquired, "How do I randomize the participants in the app?" To which the response was, "That is done automatically when an observation group is added."

Participants inquired about features that had not yet been implemented but were planned for future updates. For example, they asked if the app would work offline, which was an important consideration for use in environments with limited internet connectivity. They also asked whether the data could be exported to Excel for further analysis. These features were already part of the application's development roadmap but were not yet available in the user interface at the time of the session. One participant expressed concern, "I don't know if it's an issue when I lose connection, or whether I lose my data." Another participant discussed data export capabilities, "These are the data you see when you download the Excel file. With Excel, it looks a bit different, but what you see here are all the column names."

Second User Research Session

Following the initial user research session, the app was updated to address the feedback received. Adjustments were made to the timer functionality, screen layout, language translation, and several other features as per the participants' suggestions. A second version of the application was then provided to the participants for the subsequent user research session. In Multimedia Appendix 5, the feedback received during the second user research session is detailed.

In the second session, participants expressed fewer concerns regarding the application, indicating that many of their initial issues had been resolved. However, they still provided valuable feedback on functionality and appearance. One of the main concerns was that the app was not yet fully translated into Dutch; while significant progress had been made, some parts remained in English. Participants emphasized the importance of complete translation before the app could be effectively used by nurses and other care professionals. One participant remarked, "Most of the app has been translated, but some of the texts are still only in English." Another participant added, "This all needs to be translated before we can hand this to nurses."

Participants also suggested that the criteria for showing and hiding fields during observations should be further refined based on relevance, to streamline the data entry process and reduce cognitive load. They discussed various small changes in the appearance of the app to enhance usability and visual appeal. For instance, they recommended that the "star" button used to mark important observations be made a different color to stand out more prominently. One participant suggested, "Could you make the star button orange to make it stand out a bit more?" In addition, participants found the criteria for marking an observation as "complete" to be somewhat confusing. They proposed that once an activity has been set for an observation, it should be automatically marked as complete to provide clearer feedback to the user. One participant explained, "I would prefer it if the observation could be marked as 'complete' when the activity has been set." Participants also recommended offering flexibility in the number of observation rounds that could be added, as different research protocols or care routines might require varying numbers of observations. They suggested making the process more efficient by creating a priority list or extending the default time block for certain observations that typically take longer. Furthermore, participants emphasized the importance of ensuring that the app's manual could be accessed from within the app itself. They stressed that the manual should be clear, comprehensive, and align with users' practical experiences to facilitate ease of use, especially for new users.

By the conclusion of the second user research session, participants expressed optimism about the app's potential to improve their workflow and reduce the time spent on manual data entry and processing. They acknowledged the responsiveness of the development team to their feedback and looked forward to future iterations of the app that would incorporate their latest suggestions.

Discussion

Principal Findings

This study successfully developed an app-based version of the MEDLO-tool aimed at assessing the daily life of residents with

XSL•FO
RenderX

dementia in nursing homes. By using a UCD approach, we incorporated feedback from participants throughout the development process, resulting in a minimum viable product that aligns with the needs and preferences of end-users.

The iterative development process revealed that larger issues, such as bugs and significant feature requests, were identified and addressed in the initial phases. Subsequent user research sessions showed a decrease in the number and severity of issues, indicating progressive refinement of the app. Participants expressed that the app improved upon the Excel-based MEDLO-tool by streamlining data collection and reducing manual input, thereby potentially enhancing usability.

Balancing end-user feedback with established best practices and user interface guidelines was crucial during development [34-37]. While participants provided valuable suggestions, not all could be implemented due to conflicting requests and constraints related to usability standards. For instance, some users desired a more complex timer feature, whereas others preferred simplicity. Decisions were made to benefit the majority and adhere to platform guidelines to ensure familiarity and ease of use [34,35].

The effectiveness of the UCD approach in this study aligns with previous research demonstrating its benefits in software development [38,39]. A key factor in our success was the development team's familiarity with the context of long-term care for older adults. This domain knowledge facilitated insightful conversations between developers and participants, reducing misunderstandings and accelerating the development process. Previous studies have highlighted that cognitive similarity and shared jargon can enhance team performance [40,41].

Limitations

Despite these strengths, the study has limitations that warrant consideration. The participant sample consisted primarily of researchers affiliated with the University, which may limit the generalizability of the findings to care professionals who are the intended end-users in clinical settings. While these researchers had extensive experience with the MEDLO-tool, involving a broader range of stakeholders, including nurses and other care staff, could provide additional insights into usability and practical implementation. Furthermore, the app was not tested in a real-life nursing home environment, which could have revealed context-specific challenges and opportunities.

Future Work

Future research should focus on conducting pilot studies to evaluate the feasibility and effectiveness of the MEDLO app in real-world nursing home settings. Involving care professionals in these studies would help assess the app's usability, accessibility, and impact on daily workflows. In addition, gathering feedback from residents and their families could provide a more comprehensive understanding of the app's influence on care quality.

Further development could explore integrating advanced technologies, such as artificial intelligence and machine learning, to enhance data analysis and provide predictive insights [42,43]. For example, natural language processing could be used to automate the interpretation of observational data, aiding care providers in identifying patterns and potential areas for intervention. However, implementing such technologies would require careful consideration of ethical implications, data privacy, and user acceptance.

Future research may also investigate the usability in practice, by letting nurses use the app in their daily work. This could provide valuable insights into the app's integration into existing workflows and its impact on the quality of care provided to residents. In addition, longitudinal studies could assess the long-term effects of using the app on care outcomes and staff satisfaction. This information could inform further refinements and improvements to the MEDLO app, ensuring its relevance in long-term care.

Conclusion

This study successfully demonstrates the viability and demand for an app-based MEDLO-tool for assessing residents with dementia in nursing homes. By using a UCD approach, this study addressed the limitations of the existing Excel-based system by offering a more efficient and user-friendly alternative. This study shows that involving users early on in the process and keeping them involved can have a positive effect on the usability of an application. The MEDLO app shows that a UCD approach can provide real benefits in the development of a digital tools in nursing homes.

Data Availability

The datasets generated or analyzed during this study are available from the Living Lab in Aging and Long-term Care (ouderenzorg@maastrichtuniversity.nl) on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The interview guide used to the conduct the initial interviews.
[DOCX File, 15 KB - humanfactors_v12i1e57911_app1.docx]

Multimedia Appendix 2

XSL•FO
RenderX

The interview guide used for the user research sessions.
[DOCX File, 14 KB - humanfactors_v12i1e57911_app2.docx]

Multimedia Appendix 3
A screenshot of what the dashboard of the application looks like.
[PNG File, 165 KB - humanfactors_v12i1e57911_app3.png]

Multimedia Appendix 4
Feedback from the initial user research session.
[DOCX File, 16 KB - humanfactors_v12i1e57911_app4.docx]

Multimedia Appendix 5
Feedback from the second user research session.
[DOCX File, 15 KB - humanfactors_v12i1e57911_app5.docx]

References

1. de Boer B, Beerens HC, Zwakhalen SMG, Tan FES, Hamers JPH, Verbeek H. Daily lives of residents with dementia in nursing homes: development of the Maastricht electronic daily life observation tool. Int Psychogeriatr 2016 Aug;28(8):1333-1343. [doi: 10.1017/S1041610216000478]
2. Logsdon RG, Gibbons LE, McCurry SM, Teri L. Assessing quality of life in older adults with cognitive impairment. Psychosom Med 2002 May;64(3):510-519. [doi: 10.1097/00006842-200205000-00016]
3. Dunton GF. Ecological momentary assessment in physical activity research. Exerc Sport Sci Rev 2017 Jan;45(1):48-54. [doi: 10.1249/JES.0000000000000092] [Medline: 27741022]
4. Beerens HC, de Boer B, Zwakhalen SMG, et al. The association between aspects of daily life and quality of life of people with dementia living in long-term care facilities: a momentary assessment study. Int Psychogeriatr 2016 Aug;28(8):1323-1331. [doi: 10.1017/S1041610216000466]
5. Bajenaru L, Marinescu IA, Dobre C, Prada GI, Constantinou CS. Towards the development of a personalized healthcare solution for elderly: from user needs to system specifications. Presented at: 2020 12th International Conference on Electronics, Computers and Artificial Intelligence (ECAI); Jun 25-27, 2020; Bucharest, Romania. [doi: 10.1109/ECAI50035.2020.9223254]
6. Stewart AL, King AC. Evaluating the efficacy of physical activity for influencing quality-of-life outcomes in older adults. Ann Behav Med 1991 Sep 1;13(3):108-116. [doi: 10.1093/abm/13.3.108]
7. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008 Apr 1;4(1):1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415]
8. Moore RC, Depp CA, Wetherell JL, Lenze EJ. Ecological momentary assessment versus standard assessment instruments for measuring mindfulness, depressed mood, and anxiety among older adults. J Psychiatr Res 2016 Apr;75:116-123. [doi: 10.1016/j.jpsychires.2016.01.011] [Medline: 26851494]
9. Or CKL, Karsh BT. A systematic review of patient acceptance of consumer health information technology. J Am Med Inform Assoc 2009;16(4):550-560. [doi: 10.1197/jamia.M2888] [Medline: 19390112]
10. Gagnon MP, Desmartis M, Labrecque M, et al. Systematic review of factors influencing the adoption of information and communication technologies by healthcare professionals. J Med Syst 2012 Feb;36(1):241-277. [doi: 10.1007/s10916-010-9473-4] [Medline: 20703721]
11. Beerens HC, Zwakhalen SMG, Verbeek H, et al. The relation between mood, activity, and interaction in long-term dementia care. Aging & Mental Health 2018 Jan 2;22(1):26-32. [doi: 10.1080/13607863.2016.1227766]
12. Lee H, Chapiro J, Schernthaner R, et al. How I do it: a practical database management system to assist clinical research teams with data collection, organization, and reporting. Acad Radiol 2015 Apr;22(4):527-533. [doi: 10.1016/j.acra.2014.12.002] [Medline: 25641319]
13. Kruse CS, Kristof C, Jones B, Mitchell E, Martinez A. Barriers to electronic health record adoption: a systematic literature review. J Med Syst 2016 Dec;40(12):252. [doi: 10.1007/s10916-016-0628-9] [Medline: 27714560]
14. Adipat B, Zhang D. Interface design for mobile applications. Presented at: 11th Americas Conference on Information Systems, AMCIS 2005; Aug 11-14, 2005; Omaha, NE, United States.
15. Martínez-Pérez B, de la Torre-Díez I, López-Coronado M. Mobile health applications for the most prevalent conditions by the World Health Organization: review and analysis. J Med Internet Res 2013 Jun 14;15(6):e120. [doi: 10.2196/jmir.2600] [Medline: 23770578]
16. dos Santos Picanco LC, Ozela PF, de Fatima de Brito Brito M, et al. Alzheimer's disease: a review from the pathophysiology to diagnosis, new perspectives for pharmacological treatment. CMC 2018 Sep 4;25(26):3141-3159. [doi: 10.2174/0929867323666161213101126]

XSL•FO
RenderX

17. Bernstein Sideman A, Harrison KL, Garrett SB, Naasan G, Group D, Ritchie CS. Practices, challenges, and opportunities when addressing the palliative care needs of people living with dementia: Specialty memory care provider perspectives. Alzheimers Dement (N Y) 2021;7(1):e12144. [doi: 10.1002/trc2.12144] [Medline: 33969177]
18. Wang W, Grundy J, Khalajzadeh H, Madugalla A, Obie HO. Designing adaptive user interfaces for mhealth applications targeting chronic disease: a user-centric approach. arXiv. Preprint posted online on May 14, 2024. [doi: 10.48550/arXiv.2405.08302]
19. Norman DA. User Centered System Design: New Perspectives on Human-Computer Interaction: CRC Press; 1986.
20. Gould JD, Lewis C. Designing for usability: key principles and what designers think. Commun ACM 1985 Mar;28(3):300-311. [doi: 10.1145/3166.3170]
21. Tan L, Szebeko D. Co-designing for dementia: the Alzheimer 100 project. AMJ 2008 Nov 1:185-198. [doi: 10.4066/AMJ.2009.97]
22. Slattery P, Saeri AK, Bragge P. Research co-design in health: a rapid overview of reviews. Health Res Policy Syst 2020 Feb 11;18(1):17. [doi: 10.1186/s12961-020-0528-9] [Medline: 32046728]
23. Martı´nez-Pérez B, Torre-Dı´ez I, López-Coronado M. Privacy and security in mobile health apps: a review and recommendations. J Med Syst Springer 2015 Jan;39(1). [doi: 10.1007/s10916-014-0181-3]
24. McCusker D, Guzik K. User-Centered Design for Middleware: Elsevier; 2007.
25. Couture B, Lilley E, Chang F, et al. Applying user-centered design methods to the development of an mHealth application for use in the hospital setting by patients and care partners. Appl Clin Inform 2018 Apr;9(2):302-312. [doi: 10.1055/s-0038-1645888] [Medline: 29742756]
26. Abras C, Maloney-Krichmar D, Preece J. Bainbridge, W Encyclopedia of Human-Computer Interaction Thousand: Sage Publications; 2004, Vol. 37:445-456.
27. Wilson C. Interview Techniques for UX Practitioners: A User-Centered Design Method: Elsevier; 2013.
28. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
29. Maguire M, Delahunt B. Doing a thematic analysis: a practical, step-by-step guide for learning and teaching scholars. AISHE-J 2017;9(3). [doi: 10.62707/aishej.v9i3.335]
30. Strauss A, Corbin J. Basics of Qualitative Research Techniques: Sage Publications; 1998.
31. Fox S, Brown LJE, Antrobus S, et al. Co-design of a smartphone app for people living with dementia by applying agile, iterative co-design principles: development and usability study. JMIR Mhealth Uhealth 2022 Jan 14;10(1):e24483. [doi: 10.2196/24483] [Medline: 35029539]
32. Your research: is it subject to the WMO or not? Central Committee on Research Involving Human Subjects. URL: https://english.ccmo.nl/investigators/legal-framework-for-medical-scientific-research/your-research-is-it-subject-to-the-wmo-or-not [accessed 2025-02-13]
33. Research code version 2024. Maastricht UMC+. 2024. URL: https://www.mumc.nl/sites/default/files/2024-02/Research_Code_v2024.pdf [accessed 2025-02-13]
34. Visnapuu K. Consistency, efficiency, and scalability in design systems: a comparative analysis of material design, Apple's human interface guidelines, and IBM design systems. : Uppsala University; 2023.
35. Gülenman T. Designing better mobile apps: an experimental evaluation of apple's and google's design guidelines: how analysing the human interface guidelines for ios and material design for android better our understanding of the usability challenges app users face and what we can do to overcome key issues [Thesis]. : Jönköping University School of Engineering; 2022 URL: https://www.diva-portal.org/smash/get/diva2:1677778/FULLTEXT01.pdf [accessed 2025-01-31]
36. Eberts RE. User Interface Design: Prentice-Hall, Inc; 1994.
37. Peter Dey P, Raj Sinha B, Amin M, Badkoobehi H. Best practices for improving user interface design. IJSEA 2019 Sep 30;10(5):71-83 [FREE Full text] [doi: 10.5121/ijsea.2019.10505]
38. Mao JY, Vredenburg K, Smith PW, Carey T. The state of user-centered design practice. Commun ACM 2005 Mar;48(3):105-109. [doi: 10.1145/1047671.1047677]
39. Van Elzakker C, Delikostidis I. User-centered design of mobile geo-applications. In: Handbook of Research on Mobile Software Engineering: Design, Implementation, and Emergent Applications: IGI Global; 2013:772-787. [doi: 10.4018/978-1-4666-4301-7.ch037]
40. Kang HR, Yang HD, Rowley C. Factors in team effectiveness: cognitive and demographic similarities of software development team members. Hum Rel 2006 Dec;59(12):1681-1710. [doi: 10.1177/0018726706072891]
41. He J, Butler BS, King WR. Team cognition: development and evolution in software project teams. J Manag Inf Syst 2007 Oct;24(2):261-292. [doi: 10.2753/MIS0742-1222240210]
42. Ashraf A, Imam A. ChatGPT's use case for software engineers. IET Conf Proc 2023 Jul 26;2023(5):487-492. [doi: 10.1049/icp.2023.1537]
43. Roziere B, Gehring J, Gloeckle F, Sootla S, Gat I, Tan XE, et al. Code Llama: open foundation models for code. arXiv. Preprint posted online on Aug 24, 2023. [doi: 10.48550/arXiv.2308.12950]

XSL•FO
RenderX

Abbreviations

EMA: ecological momentary assessment
MEDLO: Maastricht Electronic Daily Life Observation
MEDLO-tool: Maastricht Electronic Daily Life Observation tool
MVP: minimum viable product
UCD: user-centered design
WMO: Medical Research Involving Human Subjects Act

Edited by A Kushniruk, Y Quintana; submitted 29.02.24; peer-reviewed by H Wörtche, ZVD Heijden; revised version received 22.11.24; accepted 25.11.24; published 27.02.25.

Please cite as:
Hacking C, de Boer B, Verbeek H, Hamers J, Aarts S
Developing an App for Real-Time Daily Life Observations in a Nursing Home Setting: Qualitative User-Centered Co-Design Approach
JMIR Hum Factors 2025;12:e57911
URL: https://humanfactors.jmir.org/2025/1/e57911
doi:10.2196/57911

© Coen Hacking, Bram de Boer, Hilde Verbeek, Jan Hamers, Sil Aarts. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 27.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Oncology Clinicians' Perspectives of a Remote Patient Monitoring Program: Multi-Modal Case Study Approach

Ann Marie Mazzella-Ebstein[1], RN, PhD; Robert Daly[2], MBA, MD; Jennie Huang[3], BA; Camila Bernal[4], MPH; Clare Wilhelm[5], PhD; Katherine S Panageas[6], Dr PH; Jessie Holland[7], RN, MSN, NE-BC; Rori Salvaggio[7], RN, MS, NEA-BC; Jill Ackerman[4], MBA; Jennifer Cracchiolo[5], MD; Gilad Kuperman[8], MD, PhD; Jun Mao[9], MSCE, MD; Aaron Begue[10], RN, MS, CNP; Margaret Barton-Burke[1], RN, PhD

[1]Nursing Research, Department of Nursing, Memorial Sloan Kettering Cancer Center, New York, NY, United States

[10]Hospital Administration, Advanced Practice Providers, Memorial Sloan Kettering Cancer Center, New York, NY, United States

[2]Thoracic Services, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, United States

[3]Memorial Sloan Kettering Cancer Center, New York, NY, United States

[4]Hospital Administration, Memorial Sloan Kettering Cancer Center, New York, NY, United States

[5]Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY, United States

[6]Epidemiology - Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY, United States

[7]Nursing Ambulatory Services, Department of Nursing, Memorial Sloan Kettering Cancer Center, New York, NY, United States

[8]Digital Informatics & Technology Solutions (DigITs), Memorial Sloan Kettering Cancer Center, New York, NY, United States

[9]Integrative Medicine Services, Memorial Sloan Kettering Cancer Center, New York, NY, United States

Corresponding Author:
Ann Marie Mazzella-Ebstein, RN, PhD
Nursing Research, Department of Nursing, Memorial Sloan Kettering Cancer Center, New York, NY, United States

Abstract

Background: Remote patient monitoring (RPM) aims to improve patient access to care and communication with clinical providers. Overall, understanding the usability of RPM applications and their influence on clinical care workflows is limited from the perspectives of clinician end users at a cancer center in the Northeastern United States.

Objective: This study aims to explore the usability and functionality of RPM and elicit the perceptions and experiences of oncology clinicians using RPM for oncology patients after hospital discharge.

Methods: The sample included 30 of 98 clinicians (31% response rate) managing at least 5 patients in the RPM program and responding to the mHealth usability between March 2021 and October 2021. Overall, clinicians responded positively to the survey. Item responses with the highest proportion of disagreement were explored further. A nested sample of 5 clinicians who responded to the study survey (30% response rate) participated in interview sessions conducted from November 2021 to February 2022, averaging 60 minutes each.

Results: Survey responses highlighted that RPM was easy to use and learn and verified symptom alerts during follow-up phone calls. Areas to improve identified practice changes from reporting RPM alerts through digital portals and its influence on clinicians' workload burden. Interview sessions revealed 3 main themes: clinician understanding and usability constraints, patient constraints, and suggestions for improving the program. Subthemes for each theme were explored, characterizing technical and functional limitations that could be addressed to enhance efficiency, workflow, and user experience.

Conclusions: Clinicians support the value of RPM for improving symptom management and engaging with providers. Improvements to address RPM challenges include functional changes to enhance the program's utility, such as input from patients about temporal changes in their symptoms and technical resources for home monitoring devices.

(JMIR Hum Factors 2025;12:e60585) doi:10.2196/60585

KEYWORDS

cancer; oncology; clinician end users; remote patient monitoring; digital health; implementation science; patient monitoring; patient access; care; communication; usability; functionality; survey; interview; efficiency; workflow; user; clinician support

XSL•FO
RenderX

Introduction

Improving the patient's access to care and enhancing their quality of life by preventing readmission is the primary goal of posthospital care delivery, yet traditional oncology care models often lack communication and monitoring after discharge [1-5]. Consequently, re-admissions for oncology patients within 30 days of discharge for symptoms that could be mitigated with proactive, remote patient monitoring (RPM) is an opportunity for advancing oncology patient care [6-8]. Our prior research at Memorial Sloan Kettering Cancer Center (MSK) highlights the potential of RPM to facilitate transitions of care through optimized patient-provider communication and proactive engagement through digital technology [7,8]. The RPM program was designed to reduce unplanned care visits following discharge from the hospital by proactively monitoring patient status via daily questionnaires. Electronic patient-reported outcome (ePRO) assessments generated alerts for patients enrolled in the RPM that were sent to the patient's primary oncology care clinician team. Details regarding the early development and integration of the organizational RPM project, as well as the patients recruited have been previously published [3,9-11]. However, limited research exists describing the perspectives and experiences of the clinician end users regarding how digital interventions are integrated into clinical care and the influence of technological health care models on patient care workflows.

To address this gap, our team developed this study to gain a greater understanding of the influence of RPM on clinical workflows in daily practice as part of a larger institutional initiative supporting center-wide symptom management during the transition of care from the inpatient to the outpatient setting [3,9-11]. The purpose of this study is to explore the usability of the RPM program and the perceptions and experiences of oncology clinicians using RPM to care for patients in practice after hospital discharge. The broader aim was to identify the impact, perceived usefulness ease of use, user control of the RPM, and barriers and facilitators experienced during the initial RPM implementation. The Consolidated Framework for Implementation Research [12] guided this study to query oncology clinicians, the key contributors to this management strategy, about their experiences, satisfaction, usefulness, and value of RPM.

Methods

Overview

The study's design used a multi-modal case study approach, using quantitative (survey) and qualitative (interviews and focus groups) to characterize the phenomenon and context of integrating RPM into the practice setting [13-18]. This approach encourages multiple sources of evidence to understand the clinician's perspectives about the use and function of RPM in real-world clinical practice and its influence on patient care workflows.

Initially, survey methods assessed clinicians' responses to the overall usability, functionality, and value of RPM. Second, clinicians participated in semistructured interviews and focus group sessions to elicit feedback and explanations about items of concern found in the survey responses [13,14,17]. The interview sessions aimed to delve deeper into the survey concerns and examine the use of the RPM application from a clinical perspective.

Setting and Clinician Sample

This study focused on clinician end users, specifically nurses involved in clinical office practices that integrated the initial organization's RPM initiative [3,9-11] from October 15, 2018, to July 10, 2019, at MSK in the Northeastern United States. The clinicians in this study were the first group of end users involved in caring for oncology patients using RPM. The RPM program was rolled out sequentially, beginning with the oncology services that care for patients with the highest symptom burden (Table 1). Clinicians from physician office practice settings, who had experience in caring for at least 5 patients enrolled in RPM, were invited to participate in the usability survey. This criterion ensured that the clinicians had ample experience with using the RPM program. Additionally, a nested sample of clinicians who completed the survey were recruited to take part in semistructured interviews to provide further information about the challenges of the RPM identified from the survey [19-22].

XSL•FO
RenderX

Table . Description of practice clinicians by percent of the sample. Clinicians: registered nurse in practice settings; responses (n=30; response rate 31%).

Oncology specialty disease management service	m-Adapted health questionnaire disseminated, n (%)[a]
Thoracic (THR)	23 (24)
Gastro-intestinal (GI)	30 (31)
Breast (BR)	15 (15)
Sarcoma (SAR)	6 (6)
Genito-urinary (GU)	6 (6)
General medicine oncology (GMO)	5 (5)
Head & neck (HN)	4 (4)
Melanoma (MEL)	2 (2)
Myeloma (MYL)	2 (2)
Bone marrow transplant (BMT)	1 (1)
Lymphoma (LYMP)	1 (1)
Leukemia (LEU)	1 (1)
Covering clinicians/no specific service	2 (2)
Total surveys sent	98 (100)

[a]Percentage of the sample.

Data Collection

Survey

Eligible clinicians in this study who met the study criteria were identified using the RPMs dashboard. Clinicians were recruited using the organization's email with a link to a web-based consent and m-Adapted Health Usability Questionnaire from March 2021 to October 2021. To maintain the anonymity of the clinicians, broad demographic data were collected to quantify clinicians' years of experience in their current organizational roles and use RPMs within electronic health record systems.

The m-Adapted Health Usability Questionnaire [23,24], is a 29-item self-report survey, that assessed an overall understanding of the functionality of the RPM from the perspectives of the clinician end users including the quality of work life, perceived usefulness, and ease of use, and user control of RPM in a 7-point Likert Scale (1 [strongly disagree] to 7 [strongly agree]). The questionnaire takes about 13 minutes to complete in REDCap (research electronic data capture; Vanderbilt University) [25]. The Cronbach α for all scale values was >0.7 with scores ranging from 0.85 to 0.92. Permission to use the questionnaire was not required [24].

Interview Sessions and Guide

Clinicians who completed the study's survey were also recruited through the organization's email to take part in 60-minute interviews or focus group sessions. The interview guide was developed by the research team based on understanding areas where the usability of RPM was challenging for the clinicians. Items from the usability survey where the percentage of responders who disagree, disagree, and strongly disagree exceeded 44%, informed the interview and focus group guide. Topics in the guide were broad enough to elicit narrative data describing the background and contextual characteristics related to the clinician's experience with RPM [26,27]. The interview guide was organized into four topics: (1) General understanding and information about the RPM; (2) Experiences, barriers, and challenges with using the RPM; (3) The Influence of the RPM on the current workflow; and (4) Suggestions for improvement (Table 2). Interviews and focus groups were conducted via the web by the PI (AMME) and the qualitative methods specialist (MBB) from November 2021 to February 2022. Verbal consent was obtained from clinicians to record sessions.

XSL•FO
RenderX

Table . Interview guide.

Topic	Main statements	Probe questions	Aligns survey
Topic 1: provide general information about the topic of study and ask questions related to the specific service and the specific patient needs.	• Where do you work? • Can you tell me about your work with RPM[a]? • What can you tell us/each other about RPM?	• Can you briefly describe the patient group in your office practices (patient symptom burden, any factors in this group that need to be addressed through RPM)? • Can you describe your workflow and the technology systems used in your daily work activities? • Can you describe the workflow, and the systems used when providing care? • How does the RPM fit into the care workflow?	Q[b] 3,4
Topic 2: issues and barriers that are specific to using the RPM.	• What are the strengths/barriers to using RPM from the nursing perspective? • What are the strengths/barriers to using RPM from the patient's perspective? • Do you think RPM is valuable: to nurses the patients?	• What functionality is found in the current systems and applications used during patient care activities to make caring for patients easy? • What functionality in the RPM is not working? Describe? • Do you think the patients like using the RPM? • Does the RPM provide value for the patients? Care? • Do you think patient education about the RPM helps?	Q 7,9,10,13
Topic 3: practice workflow; getting to the issue about the RPM, frequency of using the toolkit, and the influence of workflow.	• How does the RPM toolkit work with your workflow? • How do you use the RPM apps to provide care? • How do you use the RPM toolkit using the digital apps?	• Describe how you use the RPM (and patients with remote monitoring devised, ie, pulse oximetry) • How do you use the information in caring for your patients? • Can you describe the communication among practitioners using the RPM?	Q 18 - 20
Topic 4: functionality, if modified, would be meaningful and helpful for the RPM, improved workflow, user satisfaction, and improved patient outcomes.	• What are your thoughts about how RPM functions? • How could it be improved?	• If you could fix anything about the RPM, what would it be? • To what extent do you think the modifications described could influence patient outcomes that patient care? • What current functionality has the greatest influence on workflow?	Q 23 - 25

[a]RPM: remote patient monitoring.

[b]Q: question.

Data Analysis

Survey

Survey response data were extracted from REDCap [25] to a deidentified Excel spreadsheet before analysis. To ensure anonymity, demographic information was only provided for clinicians who were sent the surveys and not collected for clinicians who responded. Data responses were initially scored in REDCap [25] for the frequency and proportion of agreement or disagreement with survey items consistent with the 7-point response scale. The scale responses were then collapsed into 3 groups, disagreement (strongly disagree, disagree), neutral (neither agree nor disagree), and agreement (agree and strongly agree). The Strengthening the Reporting of Observational Studies in Epidemiology guidelines were used to report the study's quantitative findings [28].

Interviews Sessions

Data from the interviews and focus groups followed the Consolidated Criteria for Reporting Qualitative Research

XSL•FO
RenderX

guidelines and were used to report the study's findings [29]. The rigor and validity of the data were supported by clarifying participants' statements during interview sessions. The transcripts were audio recorded, transcribed verbatim, and independently reviewed by 2 reviewers using thematic content analysis [30]. Themes and subthemes were constructed based on the verbal responses from the clinicians. Team consensus was determined by agreeing upon the best representation of the data [24,26,27,30].

Ethical Considerations

This research is part of an ongoing organizational program of RPM initiatives. This research obtained ethical approval from the MSK Institutional Review Board (X20-086) as exempt research and follows the ethical principles and guidelines of the Belmont Report. All responders to surveys completed web-based informed consent to participate. Participants in focus groups and interviews provided informed verbal consent.

Results

Overview

A total of 35 clinicians participated in this study from both survey and interview sessions. Of the 98 clinicians who had experiences with at least 5 patients enrolled in RPM and received the study's survey, 30 responded (31% response rate). A nested sample of 5 clinicians (17% response rate) from this original group of 30 clinicians who completed the survey, also agreed to take part in 4 sessions (3 interviews; 1 focus group of 2 clinicians). Demographic data were not collected with the survey to maintain the anonymity of the clinician responders. However, the majority of clinicians who responded to the survey and interview sessions represented the Gastrointestinal, Thoracic, and Breast oncology services known to have the highest symptom burden. Of the clinicians who took part in the interview sessions, 3 clinicians had 5 or more years of organizational work experience, while 2 clinicians had 2 years or less of experience.

Survey Responses

Overall, the clinicians reported that the RPM was easy to use and learn and that symptoms communicated through the RPM program were confirmed during follow-up phone calls. Suggested areas for improvement included communication and practice changes related to symptom management and the clinicians' workload burden. Findings highlighting survey items with the highest frequency and proportion of survey responses are presented (Table 3).

XSL•FO
RenderX

Table . Frequency distribution of m-Adapted health questionnaire responses (n=30)[a].

Statements about remote patient monitoring (RPM)	Total responses, n	Agree/somewhat agree/strongly agree, n (%)[b]	Neither agree nor disagree, n (%)	Disagree/somewhat disagree/strongly disagree, n (%)	Unknown, n
Ease of use and functionality as easy to learn					
1. RPM is easy to use	27	18 (67)	2 (7)	7 (26)	3
2. RPM is easy for me to learn	27	18 (67)	7 (15)	5 (18)	3
3. Liked the digital interface of the patient data received through portal secure message alerts	26	8 (31)	4 (15)	14 (54)	4
4. I liked the digital interface of the patient data received through the Splunk/summary dashboard	26	7 (27)	12 (46)	7 (27)	4
5. Information in the Splunk/summary dashboard was well-organized	25	12 (48)	9 (36)	4 (16)	5
Integrating the RPM into workflows					
6. RPM has usable functions and capabilities	27	14 (52)	6 (22)	7 (26)	3
7. RPM has been appropriate for me to care for patients	27	11 (44)	1 (4)	14 (52)	3
8. Easy to integrate into my current clinical workflow	27	11 (41)	4 (15)	12 (44)	3
Acceptable for practice					
9. An acceptable way to coordinate health care services	26	9 (35)	5 (19)	12 (46)	4
10. Improved communication between my colleague's office practice teams for patient symptoms	26	10 (38)	1 (4)	15 (58)	4
11. Prompts me to refer patients to a specialist for symptom management	25	9 (36)	5 (20)	11 (44)	5
12. RPM is useful for my health care practice	26	12 (46)	4 (15)	10 (39)	4
13. Improved my ability to deliver health care services	27	7 (26)	4 (15)	16 (59)	3
14. Helped me manage my patient's symptoms effectively	27	11 (41)	5 (18)	11 (41)	3

XSL•FO
RenderX

Statements about remote patient monitoring (RPM)		Total responses, n	Agree/somewhat agree/strongly agree, n (%)[b]	Neither agree nor disagree, n (%)	Disagree/somewhat disagree/strongly disagree, n (%)	Unknown, n
RPM convenience						
	15. Is convenient for me to communicate with patients	26	10 (39)	4 (15)	12 (46)	4
	16. Had many more opportunities to interact with patients	27	11 (41)	4 (15)	12 (44)	3
	17. Felt comfortable communicating with my patients about symptoms using portal secure messaging	26	12 (46)	1 (4)	13 (50)	4
Devices and symptom management						
	18. Highlighted the high-risk symptoms (not pulse oximetry) provided the correct corresponding severity level for the patient-reported symptoms	27	10 (37)	5 (19)	12 (44)	3
	19. Highlighted symptoms related to pulse oximetry appropriately	27	10 (37)	9 (33)	8 (30)	3
	20. Patients were appropriately identified for pulse oximeter monitoring.	27	12 (44)	7 (26)	8 (30)	3
	21. Pulse oximeter monitoring enabled me to more effectively manage my patient's symptoms	26	10 (38)	8 (31)	8 (31)	4
	22. High-risk symptom alerts were confirmed upon communication with the patient by telephone	27	20 (74)	2 (7)	5 (19)	3
Satisfaction, value, and recommendations						
	23. Using RPM has improved my job satisfaction	27	3 (11)	9 (33)	15 (56)	3
	24. Using RPM decreased my workload	26	2 (11)	2 (8)	21 (81)	4
	25. Adds value to how I can care for my patients	27	8 (30)	7 (26)	12 (44)	3
	26. Patients reported the value of participating in RPM	27	8 (30)	7 (26)	12 (44)	3

XSL•FO
RenderX

Statements about remote patient monitoring (RPM)		Total responses, n	Agree/somewhat agree/strongly agree, n (%)[b]	Neither agree nor disagree, n (%)	Disagree/somewhat disagree/strongly disagree, n (%)	Unknown, n
	27. Overall, I am satisfied with the using RPM	27	9 (33)	4 (15)	14 (52)	3
	28. I would use RPM again to monitor the symptoms of patients	27	10 (37)	6 (22)	11 (41)	3
	29. I would recommend the RPM to my colleagues	26	8 (31)	5 (19)	13 (50)	4

[a]Number of participants who answered the survey.

[b]Percentage of the group responses.

Usability

The RPM was easy for clinicians to use but was influenced by a shift in basic assumptions with patients now instructed to report symptoms through portal messages rather than calling the medical offices (14/26, 54%).

Integrated Workflows

The proportion of responses was fairly divided between agreement and disagreement. While 14 (52%) clinicians agreed that the RPM had usable functions and capabilities, 14 (52%) clinicians disagreed that the RPM was suitable for their patient care needs.

Acceptability in Practice

A total of 16 (59%) clinicians reported a disagreement with the notion that RPM improved their ability to care for their patients and, 12 (46%) clinicians reported that the RPM was useful in their practice.

Convenience in Patient Care

Despite similar proportions of agreement and disagreement overall, a higher proportion of clinicians (13/26, 50%) disagreed with feeling comfortable communicating with patients about symptoms through portal messages.

Symptom Management

Clinicians reported that devices were appropriately used for 12 (44%) patients and that high-risk symptoms were confirmed during follow-up phone calls to 20 (74%) patients. The responses were fairly split between agreement (10/27, 37%) and disagreement (12/27, 44%) on how well the devices provided severity levels. Additionally, 10 (38%) patients agreed that RPM helped them manage their patients.

Satisfaction and Value

A total of 21 (81%) clinicians reported that the RPM did not decrease their workload and 15 (56%) clinicians reported it did not improve job satisfaction. In addition, 14 (52%) of the clinicians were not satisfied with RPM, and 13 (50%) clinicians would not recommend the RPM to colleagues. The proportion of clinicians who neither agreed nor disagreed was similar to those who reported disagreement and was further explored in the interview sessions.

Interview Sessions

The interviews and focus group sessions consistently provided similar information about the RPM thus achieving thematic saturation [30]. All clinicians reported a limited understanding of the RPM during its implementation and suggested that ongoing educational modules and supporting technical support would enhance the RPM program for the clinician end users. Three major themes emerged from the interviews: clinician understanding and usability constraints, patient constraints, and suggestions for improvement. Subthemes were further explored (Table 4).

Table . Themes and subthemes.

Theme	Subthemes
Theme 1: Clinician understanding and usability constraints	• Clinician uncertainty • Repetitive alerts • Alignment with clinician workflows • Program value
Theme 2: Patient constraints	• Appropriate patient enrollment • Timing of enrollment • Communication during and after enrollment
Theme 3: Suggestions for improvement	• Clinical champions needed • Program modifications • Information technology support for end users

XSL•FO
RenderX

Theme 1: Clinician Understanding and Usability Constraints

Overview

Theme 1 encompasses the information given to the clinicians during their RPM orientation and how they applied the application to their current workflow. Three subthemes included clinician uncertainty, repetitive alerts, alignment with clinician workflows, and program value.

Subtheme: Clinician Uncertainty

The RPM was rolled out in stages starting with oncology services known to have patients with high symptom burdens. The clinicians involved in the initial rollout reported more knowledge and understanding about the RPM compared with clinicians who were involved later in the rollout. Clinicians joining later in the rollout reported having little training about the RPM. However, they reported that the increased frequency of using the RPM helped them navigate the program for addressing symptom management alerts and functionality.

> *I have only worked with it (RPM) for a couple of patients in the outpatient setting, but I get a notification (through) the portal … to notify us that the patient has enrolled in the RPM program…During this timeframe patients (complete) a survey every day about their symptoms and how they feel.* [R32]

Subtheme: Repetitive Alerts

All clinicians reported concerns about repetitive alerts from daily patient surveys. Each day that a patient reported a symptom and generated an alert [1,3,10], clinicians called the patient to verify it and confirm if the symptom was worse, better, or the same as the previous survey. Although they could see past patient response trends in the system, this required substantial effort and did improve the patient's care for symptoms. When alerts were consistently reported, the clinicians used their judgment to decide when to call patients for verification.

> *There are a couple of categories that we feel were a little monotonous. ie, was (the pain/symptom) worse? If they are at stage four lung cancer, (the patient) is on treatment and they have fatigue, every day… (the patient cannot modify) their answers, (for example) moderate fatigue; I am having trouble doing my activities of daily living, so they are clicking that every day.* [R11]

Subtheme: Aligning With Clinician Workflow

Clinicians reported that before implementing the RPM, patients were instructed to call the physician's medical offices to report symptoms. Although none of the clinicians reported significant changes to their clinical workflows, they all raised concerns about the shift from patients calling the medical oncology office for unrelieved symptoms through a digital portal. This change made clinicians apprehensive about potentially missing symptoms reported in patient portals.

> *Before RPM, patients were not supposed to report symptoms through the portal, but a lot of patients, ended up just doing that because they were home.* [R11]

Subtheme: Program Value

Clinicians reported that the RPM was of immense value to them and their patients and supported its inclusion in the organization's future care delivery. Patients liked that their office practice clinician proactively contacted and interacted with them after they were discharged from the hospital.

> *I do not think they (the patients) mind and love to be followed up closely, I mean my patients would love for me, a call them every day. I think patients prefer (clinicians) calling them over calling the office.* [R41]

Theme 2: Patient Constraints

Overview

The patient constraint themes involved the transition of care from the inpatient setting to RPM following discharge. Three subthemes emerged: the appropriate patient enrollment, the timing of enrollment, and communication during and after enrollment.

Subtheme: Appropriate Patient Enrollment

Nurses from the discharging inpatient unit were responsible for educating patients about using RPM and program enrollment. However, office practice clinicians stated that not all patients enrolled by the discharge team were appropriate due to a lack of technical proficiency, ability, and understanding of the purpose of the RPM. Clinicians consistently reported that they should be included in RPM enrollment decisions for their patients.

> *There is a disconnect, (between enrolling the patient and educating them about the program) and sending symptoms via the portal messaging using the patient portal. The policy is that we are given two business days to answer portal messages…clearly, we cannot do that when it is a separate message. We need to speak to that person (By phone).* [R21]

Subtheme: Timing of Enrollment

Office practice clinicians reported that the time of discharge is overwhelming for patients and inpatient unit teams have limited time to prepare the patients to leave the hospital. During this study, pandemic-related social distancing restrictions prevented caregivers and family members from taking part in the discharge process, further complicated RPM enrollment.

> *The second thing that I wanted to suggest for patients wanting to quickly leave the hospital and be discharged and suddenly just get bombarded with all this information, is to discuss the RPM program the day (before) discharge.* [R32]

Subtheme: Communication During and After Enrollment

Clinicians reported that a patient's age was not a factor in RPM enrollment, as older adults familiar with technology could engage digitally with providers. However, patients with language barriers or limited access to electronic devices needed assistance from caregivers or family members similar to findings in a prior

XSL•FO
RenderX

study [8]. The RPM assessments were intended to be completed during weekday work hours, from 9 AM to 5 PM Monday through Friday. Clinicians described receiving the surveys late in the day, or after hours added stress for the nursing team to address alerts.

> *We have a lot of foreign-speaking patients ---, you know it is not them who are filling out the survey ... and it may be one of their relatives or their daughter. But the daughter may not live with them ... so how do you know what is going on?* [R11]

Theme 3: Suggestions for Improvement

Overview

The Clinicians supported using RPM to sustain patient care posthospital discharge. From their perspective, 3 subthemes emerged, the need for clinical champions, technology support for end users (clinicians and patients), and specific RPM modifications.

Subtheme: Clinical Champions Needed

Clinicians expressed the need for continuing education and updates on RPM improvements, infrequent use required relearning the system for each new patient. While there were few Clinical Champions during the initial rollout, additional experts would have helped them navigate the program and supported novice clinicians. They also suggested developing informative slides or videos addressing specific program issues.

> *I think nursing (clinician end-user) … needs more education on the portal because it (the portal) is such a big part of our job. More training is needed for using (connecting) pulse ox to an iPhone (or other devices ie, Android).* [R11]

Subtheme: Program Modifications

Clinicians reported that patient enrollment before hospitalization or early in their hospital admission would improve the transition from inpatient to outpatient care. They also requested allowing patients to change or clarify their symptom responses within the assessment. Frequent patient alerts were seen as potentially problematic [31] and suggested that enabling patients the option to modify their assessment responses would provide more accurate symptom information to the clinical team.

> *There should be a way for them (patients) to say (respond in the survey) no changes or something… so that we don't get the same exact thing (response) that we spoke to the patient yesterday---about because it's not realistic, that I call the patient every day to talk to them about the same time (symptom).* [R42]

Subtheme: Information Technology Support for End Users

Clinicians raised concerns about the lack of integrated IT systems they use for delivering patient care. Additionally, clinicians expressed the need for greater IT support for both patients and staff. Clinicians conveyed concern about the difficulty in resolving connectivity issues and requested dedicated assistance from IT teams.

> *The patients are calling in a panic and you are trying to walk them through over the phone how to set up the device (RPM on patient's device ie, phone, tablet). If it was not set up correctly on discharge, or the device is just not working. (These issues) add more stress to the patient, but it also adds stress to you (the clinician) This becomes the added work.* [R11]

Discussion

Principal Results

This study presented the perceptions and experiences of oncology clinicians when caring for their patients using RPM after hospital discharge. Initially, clinician survey responses highlighted concerns about their understanding, perceptions, and challenges when using RPM as well as its influence on their clinical workflows. These topics were further explored in interview sessions and revealed 3 themes: clinician understanding and usability constraints, patient constraints, and suggestions for improvement. Most clinicians found the RPM easy to use and learn, allowing them to confirm the patient's reported symptoms during follow-up phone calls. They supported RPM's value for both clinicians and patients as a care delivery method in oncology practice, aligning with findings from other organizational studies using ePROs [11,31-36], as well as the challenges and limitations of addressing symptom alerts after the patient transitions from an inpatient to the home setting [7,35,36].

Challenges and Opportunities From the Clinical Setting

This research further emphasized multiple challenges faced by clinicians during the initial RPM roll-out. A major concern was the practice change in clinical workflows whereby patients reported symptoms through electronic portals rather than contacting the medical offices by phone. Despite this change, clinicians felt their comfort level would improve with more experience, and as RPMs aligned with their workflows [7]. An opportunity to improve clinician confidence includes continuing educational tutorials from the initial RPM orientation and more technical and clinical support for both patients and clinicians. These efforts would facilitate aligning clinician workflows with clinical practice, thus leading to reduced stress and improved job satisfaction when caring for oncology patients using RPM [7].

Another challenge was patient constraints and their lack of understanding of RPM, which clinicians suggested might contribute to their uncertainty in digital symptom management reporting. Some patients completed surveys late in the clinic hours making it difficult to address symptoms. In many instances, clinicians reported that family members completed the surveys instead of the patients which raised concerns about the integrity of the symptom alerts. Further research could enhance the understanding of logistical limitations involving patients and caregivers with completing postdischarge symptom assessments at home. A similar study recommended that concerns revealed by patients and caregivers postdischarge could be included in future clinical outcomes [37]. Additionally, the timeframe for enrolling patients in RPM was another

XSL•FO
RenderX

constraint identified by clinicians. Although older patients were comfortable using the RPM [38,39], enrollment at discharge was considered suboptimal. They suggested educating patients in a relaxed environment or office practices before hospitalizations would benefit patients and improve the enrollment process.

A major opportunity for RPM improvement involves addressing repetitive alerts from the daily ePRO assessments, also identified in other ePRO studies [23]. Clinicians verify all patient-reported symptom alerts. When prior symptoms are stable, these alerts could be averted to reduce their workflow burden [40,41]. To enhance effective communication with the patients, clinicians recommended adding context about symptom changes relative to the prior day's ePRO assessment to improve the integrity of alerts.

The perspectives of all clinicians involved in this study expressed a need for additional resources to improve the effectiveness of RPM and their ability to clinically support it. The RPM was initially implemented using a core team of designated clinicians for addressing patient alerts. Studies of RPM cited the inclusion of core teams for this purpose which may be a preferred approach for clinicians in other studies [42]. However, clinicians in this research preferred a hybrid RPM program with primary care clinicians in the clinical practices addressing alerts on weekdays and centralized after-hours and on weekends. Clinicians proposed that leaders provide consistent status updates and education about RPM functionality through tutorial videos in an accessible location. Additional technological support could improve data collection and relieve clinicians of the burden of providing technical support for patients at home [41,42], which is not the best use of their skills [43].

Strengths and Limitations

The consistent reports from survey responses and interviewed clinicians strengthen the study's findings and provide pragmatic recommendations for workflow redesign and enhancing access to care for oncology patients postdischarge [1,2,7,31], As a result of this study, selected modifications were implemented into practice to improve the program's experience for both clinicians and patients. These results provide important foundational work for future pragmatic trials and implementation science to enhance the usability and value of RPM in oncology practice.

The sample was limited to only clinicians who had experiences with at least 5 patients enrolled in the RPM, but typical for case studies [13-18]. They expressed strong positive or negative experiences with the RPM which may have created a selection bias. The study's design was multi-modal which could have included both interviewer and report biases. Clinicians were also employees of the organization and therefore specific demographics were not collected to ensure their anonymity. This study was conducted at one comprehensive cancer center, limiting the generalizability of results. However, the perspectives of clinician end users from oncology services with known symptom burdens can apply to similar RPM implementation initiatives in other organizations worldwide.

Conclusion

The National Cancer Institute calls for studies about cancer-related interventions, which also include program effectiveness from the perspectives of both patients and clinician end users [44]. This study contributes to the National Cancer Institute's initiatives, demonstrating that RPM is a valuable method for communicating with clinical providers and managing patient symptoms during transitions of care from inpatient to the home setting.

Acknowledgments

The authors want to thank the clinicians who participated in this study for their insight and dedication to oncology practice.

Authors' Contributions

AMME drafted the nursing perspective aspect of the proposal and conducted the interviews, transcribed and analyzed the interview data, and drafted the original manuscript. MBB conducted interviews with AMME, reviewed the transcript data, and edited the manuscript. RD is the principal investigator of the organizational protocol involving remote patient monitoring and reviewed and edited the manuscript. J Huang, CB, and JA are involved in providing protocol support for the daily activities of this study. CW was involved with drafting the manuscript and editing the study proposal. KSP, GK, and JM were involved with the organizational protocol and its development, and manuscript review. J Holland, RS, JC, and AB are involved with the organizational project development and review of the manuscript.

Conflicts of Interest

RD reports grant support from the National Institutes of Health and the Emerson Collective, participation on a data safety monitoring board or advisory board with Varian Medical Systems, and stock or stock options in Roche. The remaining authors do not have conflicts of interest to declare.

References

1. Basch E, Deal AM, Kris MG, et al. Symptom monitoring with patient-reported outcomes during routine cancer treatment: a randomized controlled trial. J Clin Oncol 2016 Feb 20;34(6):557-565. [doi: 10.1200/JCO.2015.63.0830] [Medline: 26644527]

XSL•FO
RenderX

2. Basch E, Deal AM, Dueck AC, et al. Overall survival results of a trial assessing patient-reported outcomes for symptom monitoring during routine cancer treatment. JAMA 2017 Jul 11;318(2):197. [doi: 10.1001/jama.2017.7156]
3. Daly B, Kuperman G, Zervoudakis A, et al. InSight care pilot program: redefining seeing a patient. JCO Oncol Pract 2020 Oct;16(10):e1050-e1059. [doi: 10.1200/OP.20.00214]
4. Denis F, Basch E, Septans AL, et al. Two-year survival comparing web-based symptom monitoring vs routine surveillance following treatment for lung cancer. JAMA 2019 Jan 22;321(3):306-307. [doi: 10.1001/jama.2018.18085] [Medline: 30667494]
5. Mooney KH, Beck SL, Wong B, et al. Automated home monitoring and management of patient-reported symptoms during chemotherapy: results of the symptom care at home RCT. Cancer Med 2017 Mar;6(3):537-546. [doi: 10.1002/cam4.1002] [Medline: 28135050]
6. Innovation: oncology care first model: informal request for information. Centers for Medicare and Medicaid Services. 2021. URL: https://innovation.cms.gov/files/x/ocf-informalrfi.pdf [accessed 2024-12-20]
7. Hassett MJ, Cronin C, Tsou TC, et al. eSyM: an electronic health record-integrated patient-reported outcomes-based cancer symptom management program used by six diverse health systems. JCO Clin Cancer Inform 2022 Jan;6:e2100137. [doi: 10.1200/CCI.21.00137] [Medline: 34985914]
8. Doolin JW, Berry JL, Forbath NS, et al. Implementing electronic patient-reported outcomes for patients with new oral chemotherapy prescriptions at an academic site and a community site. JCO Clin Cancer Inform 2021 Dec;5(5):631-640. [doi: 10.1200/CCI.20.00191]
9. Daly RM, Cracchiolo JR, Huang J, et al. Remote symptom monitoring after hospital discharge. JCO 2022 Jun 1;40(16_suppl):1517-1517. [doi: 10.1200/JCO.2022.40.16_suppl.1517]
10. Simon BA, Assel MJ, Tin AL, et al. Association between electronic patient symptom reporting with alerts and potentially avoidable urgent care visits after ambulatory cancer surgery. JAMA Surg 2021 Aug 1;156(8):740-746. [doi: 10.1001/jamasurg.2021.1798] [Medline: 34076691]
11. Daly B, Cracchiolo J, Holland J, et al. Digitally enabled transitional care management in oncology. JCO Oncol Pract 2024 May;20(5):657-665. [doi: 10.1200/OP.23.00565] [Medline: 38382002]
12. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009 Aug 7;4(1):50. [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
13. Yin RK. Case study evaluations: a decade of progress? In: Evaluation Models: Viewpoints on Educational and Human Services Evaluation: Dordrecht: Springer Netherlands; 2000:185-193. [doi: 10.1007/0-306-47559-6_11]
14. Yin RK. Case Study Research and Applications: Sage; 2018.
15. Hoon C. Meta-synthesis of qualitative case studies: an approach to theory building. Org Res Methods 2013 Oct;16(4):522-556. [doi: 10.1177/1094428113484969]
16. Hollweck T. Robert K. Yin. (2014). Case Study Research Design and Methods (5th ed.). Can J Program Eval 2016;30(1):282. [doi: 10.3138/cjpe.30.1.108]
17. Moss S. Case studies - the YIN approach. Charles Darwin University. 2020. URL: https://www.cdu.edu.au/files/2020-07/Introduction%20to%20case%20studies%20-%20the%20Yin%20approach.docx [accessed 2024-04-30]
18. Baškarada S. Qualitative case study guidelines. TQR 2014;19(40):1-25. [doi: 10.46743/2160-3715/2014.1008]
19. Sargeant J. Qualitative research part II: participants, analysis, and quality assurance. J Grad Med Educ 2012 Mar;4(1):1-3. [doi: 10.4300/JGME-D-11-00307.1] [Medline: 23451297]
20. Boddy CR. Sample size for qualitative research. QMR 2016 Sep 12;19(4):426-432. [doi: 10.1108/QMR-06-2016-0053]
21. Ayton D. Choose your qualitative study qualitative research–a practical guide for health and social care researchers and practitioners. Open Educational Resources Collective. 2023. URL: https://oercollective.caul.edu.au/qualitative-research/chapter/__unknown__-s2/ [accessed 2024-04-17]
22. Pirani SA. Navigating the complexity of sample size determination for robust and reliable results. IJMRR 2024;3(2):73-86. [doi: 10.56815/IJMRR.V3I2.2024/73-86]
23. Schnall R, Cho H, Liu J. Health information technology usability evaluation scale (Health-ITUES) for usability assessment of mobile health technology: validation study. JMIR Mhealth Uhealth 2018 Jan 5;6(1):e4. [doi: 10.2196/mhealth.8851] [Medline: 29305343]
24. Zhou L, Bao J, Setiawan IMA, Saptono A, Parmanto B. The mHealth app usability questionnaire (MAUQ): development and validation study. JMIR Mhealth Uhealth 2019 Apr 11;7(4):e11500. [doi: 10.2196/11500] [Medline: 30973342]
25. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381. [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
26. Roller MR. The in-depth interview methods: 12 articles on design and implementation. Roller Research. URL: http://rollerresearch.com/MRR%20WORKING%20PAPERS/IDI%20Text%20April%202020.pdf [accessed 2024-12-20]
27. Creswell JW, Creswell JD. Research Design: Qualitative, Quantitative, and Mixed-Methods Approaches: Sage; 2014.

XSL•FO
RenderX

28. Vandenbroucke JP, von Elm E, Altman DG, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. Epidemiology (Sunnyvale) 2007 Nov;18(6):805-835. [doi: 10.1097/EDE.0b013e3181577511]
29. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
30. Cernasev A, Axon DR. Research and scholarly methods: thematic analysis. J Am Coll Clin Pharm 2023 Jul;6(7):751-755. [doi: 10.1002/jac5.1817]
31. Basch E, Stover AM, Schrag D, et al. Clinical utility and user perceptions of a digital system for electronic patient-reported symptom monitoring during routine cancer care: findings from the PRO-TECT trial. JCO Clin Cancer Inform 2020 Oct;4(4):947-957. [doi: 10.1200/CCI.20.00081] [Medline: 33112661]
32. Daly B, Lauria TS, Holland JC, et al. Oncology patients' perspectives on remote patient monitoring for COVID-19. JCO Oncol Pract 2021 Sep;17(9):e1278-e1285. [doi: 10.1200/OP.21.00269] [Medline: 34085536]
33. Daly B, Nicholas K, Flynn J, et al. Analysis of a remote monitoring program for symptoms among adults with cancer receiving antineoplastic therapy. JAMA Netw Open 2022 Mar 1;5(3):e221078. [doi: 10.1001/jamanetworkopen.2022.1078] [Medline: 35244701]
34. Sandhu S, King Z, Wong M, et al. Implementation of electronic patient-reported outcomes in routine cancer care at an academic center: identifying opportunities and challenges. JCO Oncol Pract 2020 Nov;16(11):e1255-e1263. [doi: 10.1200/OP.20.00357]
35. Keith RE, Crosson JC, O'Malley AS, Cromp D, Taylor EF. Using the consolidated framework for implementation research (CFIR) to produce actionable findings: a rapid-cycle evaluation approach to improving implementation. Implement Sci 2017 Feb 10;12(1):15. [doi: 10.1186/s13012-017-0550-7] [Medline: 28187747]
36. Lawrence K, Singh N, Jonassen Z, et al. Operational implementation of remote patient monitoring within a large ambulatory health system: multimethod qualitative case study. JMIR Hum Factors 2023;10:e45166. [doi: 10.2196/45166PMID]
37. Dowzicky PM, Shah AA, Barg FK, Eriksen WT, McHugh MD, Kelz RR. An assessment of patient, caregiver, and clinician perspectives on the post-discharge phase of care. Ann Surg 2021;273(4):719-724. [doi: 10.1097/SLA.0000000000003479]
38. Melstrom LG, Zhou X, Kaiser A, et al. Feasibility of perioperative remote monitoring of patient-generated health data in complex surgical oncology. J Surg Oncol 2023 Jan;127(1):192-202. [doi: 10.1002/jso.27106] [Medline: 36169200]
39. Beaney T, Clarke J. Home oxygen monitoring and therapy: learning from the pandemic. Curr Opin Crit Care 2023 Feb 1;29(1):34-39. [doi: 10.1097/MCC.0000000000001010] [Medline: 36539965]
40. Handley NR, Schuchter LM, Bekelman JE. Best practices for reducing unplanned acute care for patients with cancer. J Oncol Pract 2018 May;14(5):306-313. [doi: 10.1200/JOP.17.00081] [Medline: 29664697]
41. Penedo FJ, Oswald LB, Kronenfeld JP, Garcia SF, Cella D, Yanez B. The increasing value of eHealth in the delivery of patient-centred cancer care. Lancet Oncol 2020 May;21(5):e240-e251. [doi: 10.1016/S1470-2045(20)30021-8] [Medline: 32359500]
42. Coffey JD, Christopherson LA, Glasgow AE, et al. Implementation of a multisite, interdisciplinary remote patient monitoring program for ambulatory management of patients with COVID-19. NPJ Digit Med 2021 Aug 13;4(1):123. [doi: 10.1038/s41746-021-00490-9] [Medline: 34389787]
43. Offodile AC II, Seitz AJ, Peterson SK. Digital health navigation: an enabling infrastructure for optimizing and integrating virtual care into oncology practice. JCO Clin Cancer Inform 2021 Dec;5(5):1151-1154. [doi: 10.1200/CCI.21.00111]
44. Division of cancer control and population sciences. National Cancer Institute. URL: https://cancercontrol.cancer.gov/ [accessed 2024-05-22]

Abbreviations

ePRO: electronic patient-reported outcome
MSK: Memorial Sloan Kettering Cancer Center
REDCap: research electronic data capture
RPM: remote patient monitoring

Edited by A Kushniruk, E Borycki; submitted 03.06.24; peer-reviewed by J Sharp, L Zheng; accepted 23.09.24; published 24.01.25.

Please cite as:
Mazzella-Ebstein AM, Daly R, Huang J, Bernal C, Wilhelm C, Panageas KS, Holland J, Salvaggio R, Ackerman J, Cracchiolo J, Kuperman G, Mao J, Begue A, Barton-Burke M
Oncology Clinicians' Perspectives of a Remote Patient Monitoring Program: Multi-Modal Case Study Approach
JMIR Hum Factors 2025;12:e60585
URL: https://humanfactors.jmir.org/2025/1/e60585
doi:10.2196/60585

XSL•FO
RenderX

© Ann Marie Mazzella-Ebstein, Robert Daly, Jennie Huang, Camila Bernal, Clare Wilhelm, Katherine S Panageas, Jessie Holland, Rori Salvaggio, Jill Ackerman, Jennifer Cracchiolo, Gilad Kuperman, Jun Mao, Aaron Begue, Margaret Barton-Burke. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 24.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluating the Safety and Usability of an Over-the-Counter Medical Device for Adults With Mild to Moderate Hearing Loss: Formative and Summative Usability Testing

Megan Elizabeth Salwei[1], PhD; Shilo Anders[1], PhD; Carrie Reale[1], MSN, RN, NI-BC; Jason M Slagle[1], PhD; Todd Ricketts[2], PhD; Matthew B Weinger[1], MD, MS

[1]Center for Research and Innovation in Systems Safety, Department of Anesthesiology, Vanderbilt University Medical Center, 2525 West End Avenue, Suite 800, Nashville, TN, United States

[2]Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, Nashville, TN, United States

Corresponding Author:
Megan Elizabeth Salwei, PhD
Center for Research and Innovation in Systems Safety, Department of Anesthesiology, Vanderbilt University Medical Center, 2525 West End Avenue, Suite 800, Nashville, TN, United States

Abstract

Background: Only 15% of the nearly 30 million Americans with hearing loss use hearing aids, partly due to high cost, stigma, and limited access to professional hearing care. Hearing impairment in adults can lead to social isolation and depression and is associated with an increased risk of falls. Given the persistent barriers to hearing aid use, the Food and Drug Administration issued a final rule to allow over-the-counter hearing aids to be sold directly to adult consumers with perceived mild to moderate hearing loss at pharmacies, stores, and online retailers without seeing a physician or licensed hearing health care professional.

Objective: We evaluated the safety and usability of an over-the-counter hearing aid prior to Food and Drug Administration approval and market release.

Methods: We first conducted a formative usability test of the device and associated app with 5 intended users to identify outstanding safety and usability issues (testing round 1). Following design modifications, we performed a summative usability test with 15 intended users of the device (testing round 2). We concurrently conducted a test with 21 nonintended users (ie, users with contraindications to use) to ascertain if consumers could determine when they should not use the device, based on the packaging, instructions, and labeling (testing round 3). Participants were asked to complete 2 - 5 tasks, as if they were using the hearing aid in real life. After each task, participants rated the task difficulty. At the end of each session, participants completed a 10-question knowledge assessment and the System Usability Scale and then participated in debriefing interviews to gather qualitative feedback. All sessions were video recorded and analyzed to identify use errors and design improvement opportunities.

Results: Usability issues were identified in all 3 usability testing rounds. There were minimal safety-related issues with the device. Round 1 testing led to several design modifications which then increased task success in round 2 testing. Participants had the most difficulty with the task of pairing the hearing aids to the cell phone. Participants also had difficulty distinguishing the right and left earbuds. Nonintended users did not always understand device contraindications (eg, tinnitus and severe hearing loss). Overall, test findings informed 9 actionable design modifications (eg, clarifying pairing steps and increasing font size) that improved device usability and safety.

Conclusions: This study evaluated the usability and safety of an over-the-counter hearing aid for adults with mild to moderate hearing loss. Human factors engineering methods identified opportunities to improve the safety and usability of this direct-to-consumer medical device for individuals with perceived mild-moderate hearing loss.

(JMIR Hum Factors 2025;12:e65142) doi:10.2196/65142

KEYWORDS

usability; human factors; patient safety; over-the-counter hearing aids; direct-to-consumer hearing aids; medical device; hearing loss; adult; hearing impairment; hearing aid use; hearing care; formative usability test; safety; mobile phone

Introduction

Approximately 1 in 4 people older than 12 years of age in the United States have hearing loss in 1 or both ears [1], and 30 million American adults could benefit from hearing aid use [2]. Hearing loss is associated with an increased risk for falls [3] and may lead to cognitive decline [4] and dementia [5]. Yet, only 1 in 3 older adults [6] and 15% of all adults with hearing loss use hearing aids [7]. Some barriers preventing adoption

XSL•FO
RenderX

include lack of access, stigma, and financial challenges with traditional hearing aids costing between US $1000 and US $6000 per ear [8]. Aiming to mitigate these barriers, the Food and Drug Administration (FDA) issued a final rule in 2022 [9] permitting the sale of hearing aid devices without a prescription or medical exam. Yet, there are few studies of the usability and safety design challenges of over-the-counter medical devices intended to be sold widely in consumer markets.

Over-the-counter or direct-to-consumer medical devices do not require a prescription and can be sold at retail outlets including pharmacies, big-box stores, and online. Examples include self-monitoring blood glucose test systems and at-home pregnancy tests. Typically used outside of health care settings, these devices must have (1) low potential for misuse; (2) benefits that outweigh their safety risks; and (3) sufficient labeling so that lay-users can self-diagnose their condition, self-determine that the device is appropriate for their condition, and understand how to use the device without clinician assistance or instruction [10]. To minimize use errors and their associated risks, the FDA requires premarket review of the safety and effectiveness for most direct-to-consumer medical devices. This review process includes rigorous human factors usability testing with representative users to validate that the warnings, cautions, and contraindications are transparent, assess if users can complete critical tasks of the device, and identify areas of foreseeable misuse (including use of the device by unintended users) [11].

Numerous studies have evaluated the safety and usability of clinician-facing medical devices (eg, smart intravenous infusion pumps) [12-15], as well as patient-facing medical devices (eg, glucometers, ventilators, and epinephrine injectors) [16]. Typically, clinicians still mediate the use of these patient-facing medical devices either by providing a prescription to use the device or instructions on how to use the device appropriately [17]. There remains limited research on the safety and usability of direct-to-consumer medical devices, especially those used by underserved populations (eg, people with hearing impairment) [18-20].

Compared to clinician-facing medical devices, patient-facing devices present a challenge to designers as the potential environment of use and characteristics of end users is much more diverse [16]. The usability of direct-to-consumer devices is especially important given the lack of clinical supervision to support appropriate use. For instance, over-the-counter hearing aids will be on the shelf at consumer electronics stores next to other, nonmedical devices (eg, wireless headphones) and consumers may come across the device without previously identifying a need for a hearing aid. While the risk is minimal, use of hearing aids by those with normal hearing could result in noise-induced hearing loss. In this context, the packaging of the device and communication of its intended use (and contraindications to use such as an ear infection, tinnitus, or severe hearing loss) is especially important. In this study, we evaluated the safety and usability of an over-the-counter hearing aid among adults in the United States.

Methods

Ethical Considerations

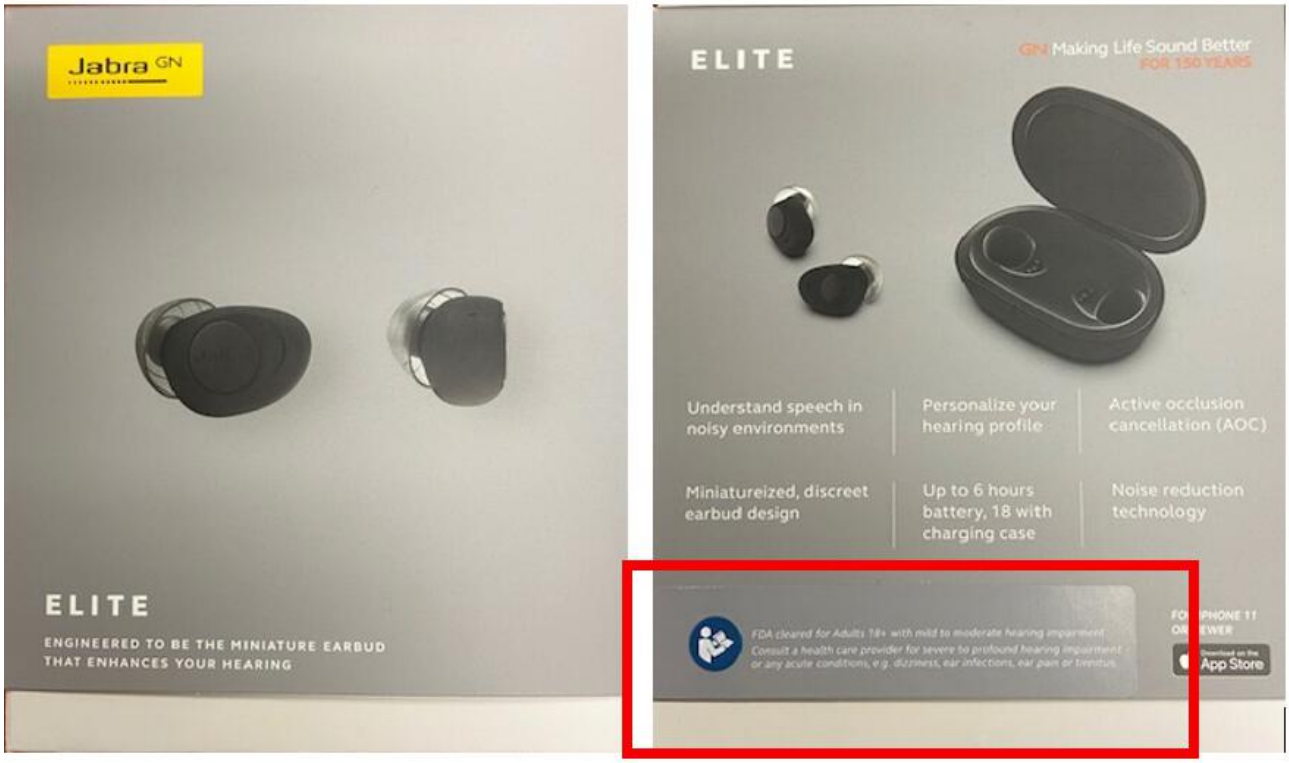
We conducted usability testing of an over-the-counter hearing aid to evaluate the safety and usability of the device prior to market release. This study was approved by the Vanderbilt University Medical Center's institutional review board (201894). At the start of each session, we reviewed the consent document with participants and answered any questions they had. All participants provided written consent to be in the study. Data have been deidentified.

Hearing Aid

We evaluated the Jabra Enhance Plus (formerly called the "Jabra Elite" during the study), an over-the-counter hearing aid that is intended to enhance hearing but can also be used for phone calls and listening to music. Users can purchase these hearing aids at a suggested retail price of US $799 per pair. The hearing aids are physically similar to wireless earbuds with adjustable ear gels. Users can pair them to a smartphone and personalize their own hearing aid profile using a free smartphone app (ie, Jabra Enhance app). The miniaturized in-the-ear style earbuds (see Figure 1) come in a small charging case and can be used up to 12 hours on a single charge.

XSL•FO
RenderX

Figure 1. Front and back of the device box.



Participants and Setting

We conducted usability testing with production equivalent prototype devices (ie, packaging, hearing aids, charging case, and user manual) including an iPhone with the Jabra Enhance app installed. We only used iPhones in the usability testing since the Jabra Enhance app was only available for iPhones at the time of testing. We recruited two participant populations for the study, (1) intended users of the device and (2) nonintended users (Textbox 1). We identified potential participants meeting our study criteria through medical record chart review, university newsletters, and word of mouth. Eligible participants conducted a prescreening questionnaire, via phone call or email, to verify they met the inclusion criteria for the study before scheduling. The testing occurred from December 2020 to August 2021.

Textbox 1. Eligibility criteria for intended and nonintended users.

Intended users (adults aged 18+) with:

- Mild to moderate hearing impairment

Nonintended users (adults aged 18+) with:

- Normal hearing,
- Severe to profound hearing impairment, and/or
- Acute conditions including tinnitus, dizziness, ear infections, and ear pain

Intended Users

To start, we conducted usability testing with 5 intended users of the device. Based on the safety and usability issues identified in the testing round (round 1), we made recommendations for modifications to the device packaging and app screens. We then conducted round 2 of usability testing of the revised product with 15 intended users. In each usability testing session, participants were asked to complete 5 tasks (see Textbox 2) within 90 minutes. Participants received a US $50 gift card at the end of the session.

XSL•FO
RenderX

Textbox 2. Task descriptions.

Tasks for intended users (rounds 1 and 2)

- Read outer label and indicate if use is appropriate
- Set up the hearing aids and app including
- Find and open hearing aid app
- Register product and accept terms and conditions
- Pair the hearing aids
- Finalize the set up of the hearing aids
- Fit the hearing aids into ear
- Recharge the hearing aids
- Clean the hearing aids

Tasks for nonintended users (round 3)

- Read outer label and indicate if use is appropriate
- Read the inner box contents (note that the inner box contents were available to the participants, but they did not receive explicit instructions on how to read them) and indicate if use is appropriate

Nonintended Users

We conducted a third round of usability testing with 21 nonintended users of the devices to evaluate if consumers could ascertain that they should not use the device based on the external and internal labeling, user manual, and app screens. We recruited participants who were contraindicated to use the device including adults with normal hearing, adults with severe to profound hearing loss, and adults with tinnitus or severe dizziness. Round 3 participants were scheduled for a 30-minute test session, in which they were asked to complete 2 tasks (see Textbox 2). Participants received a US $20 gift card at the end of the session.

Data Collection

We conducted functional simulated use studies with end users completing real-world tasks. One human factors engineer facilitated the sessions. In line with real-world use of the device, participants received no training prior to testing. We read task instructions out loud while displaying the same instructions on a computer monitor with closed captioning. Participants were directed to complete 5 (rounds 1 and 2) or 2 tasks (round 3, see Textbox 2). After each task, participants rated how easy or difficult the task was to complete on a scale from 1 (very difficult) to 5 (very easy). After all tasks were completed, we asked intended and nonintended users 10 questions to assess their knowledge of the device and its contraindications for use (see Multimedia Appendix 1). Participants then completed the System Usability Scale (SUS), a 10-item validated survey for assessing the usability of interactive systems [21] (see Multimedia Appendix 2). To conclude the session, we conducted debrief interviews to gather additional qualitative feedback about the safety and usability of the device. The interview guide can be found in Multimedia Appendix 3. All sessions were video recorded and uploaded to Morae (TechSmith), a software suite for usability testing that supports video annotation (eg, of use errors) and logging of tasks.

Data Analysis

One researcher coded the videos in Morae, annotating areas of difficulty, confusion, safety-related issues, and errors. We analyzed the participant's path (ie, steps) through the attempted completion of each task. Any deviation that occurred in the attempt to complete the task (eg, participant clicks on an incorrect menu item or participant struggled to get the user manual out of the box) was coded as a "use error." For task 2, we specifically noted any instances where participants' workflow deviated from the expected workflow (see steps in Textbox 2). We analyzed the time spent on each task and if the tasks were (1) completed, (2) completed with difficulty, or (3) not completed (task failure). We then calculated the task success rate. A task was only scored as a success if the user was able to complete the task requirements on their own without facilitator assistance. We calculated participants' rating of difficulty for each task and the SUS scores. One researcher listened to the audio recording of each debrief interview and created a comprehensive list of feedback (eg, longer battery life) organized into two categories (1) what participants liked about the device and (2) areas for improvement. A team of human factors researchers reviewed all identified errors and participant difficulties completing tasks to develop design recommendations to improve the device, app, and packaging prior to market release. We mapped each design recommendation to specific human factors design principles, using the list of design heuristics described in Barton et al [22], which combined the commonly used heuristics for medical device design [23], interactive systems [24], and medical documents [25].

Results

Overview

Table 1 shows the participant demographics for each round of usability testing. Table 2 depicts the task success rates, number

XSL•FO
RenderX

of use errors, task durations, and difficulty ratings for each task and round of usability testing.

While we identified minimal safety concerns with the device, we identified numerous usability issues, which led to several design modifications. We outline the key issues identified from the usability testing in the following sections.

Table . Participant demographics for each usability testing round.

	Round 1 (intended users, n=5)	Round 2 (intended users, n=15)	Round 3 (nonintended users, n=21)
Hearing loss, n (%)			
None or normal hearing	0 (0)	0 (0)	18 (86)
Mild	2 (40)	7 (47)	0 (0)
Mild to moderate	1 (20)	2 (13)	0 (0)
Moderate	2 (40)	6 (40)	0 (0)
Severe	0 (0)	0 (0)	3 (14)
Tinnitus (ringing in ears), n (%)	0 (0)	0 (0)	5 (24)
Average age (years), n (range)	65 (57 - 73)	66 (56 - 79)	65 (52 - 77)
Women, n (%)	2 (40)	8 (53)	13 (62)
Education, n (%)			
High school or General Educational Development	0 (0)	5 (33)	5 (24)
Associate's or trade school	1 (20)	2 (13)	4 (19)
4-year college	2 (40)	3 (20)	7 (33)
Graduate or professional school	2 (40)	5 (33)	5 (24)

Table . Usability testing performance.

Measure and round	Task 1	Task 2	Task 3	Task 4	Task 5
Task success rate, n (%)					
Round 1	3 (60)	2 (40)	2 (40)	5 (100)	3 (60)
Round 2	12 (80)	10 (67)	7 (47)	15 (100)	10 (67)
Round 3	10 (48)	10 (48)	—[a]	—	—
Average task duration (minutes), mean (SD) [b]					
Round 1	0.72 (0.3)	9.33 (0.5)	1.37 (0.6)	0.50 (0.2)	1.08 (1.1)
Round 2	0.83 (0.5)	11.24 (4.0)	2.21 (0.9)	0.67 (0.6)	0.95 (0.8)
Round 3	—	—	—	—	—
Average difficulty rating, mean (SD) [c]					
Round 1	4.6 (0.5)	3.4 (1.2)	4.2 (0.8)	5 (0)	5 (0)
Round 2	4.87 (0.4)	3.27 (1.4)	3.73 (1.2)	4.53 (0.8)	4.67 (0.6)
Round 3	4.71 (0.6)	4.43 (0.9)	—	—	—
Number of use errors per participant, n					
Round 1	0	4	0	0	0
Round 2	0	4.9	0.5	0.6	0
Round 3	0	0.7	—	—	—

[a]Not applicable.

[b]Only includes times for participants who successfully completed the task.

[c]Scale from 1 to 5 (1=very difficult, 5=very easy).

XSL•FO
RenderX

Package Labeling

Overall, we found minimal safety issues with the set up and use of the device across all 3 rounds of testing. The one issue that presented potential, albeit minimal, concerns to safety was related to the accessibility of intended use and contraindications to use of the device. We found that the font size and color on the outside of the box made the intended use of the device hard to see and read (right side of Figure 1). Few participants saw or read this information on the box, and several participants indicated that they would not know the device was a hearing aid based on the outside of the box, stating "to know they are hearing aids, off the bat, I would have walked right by them."

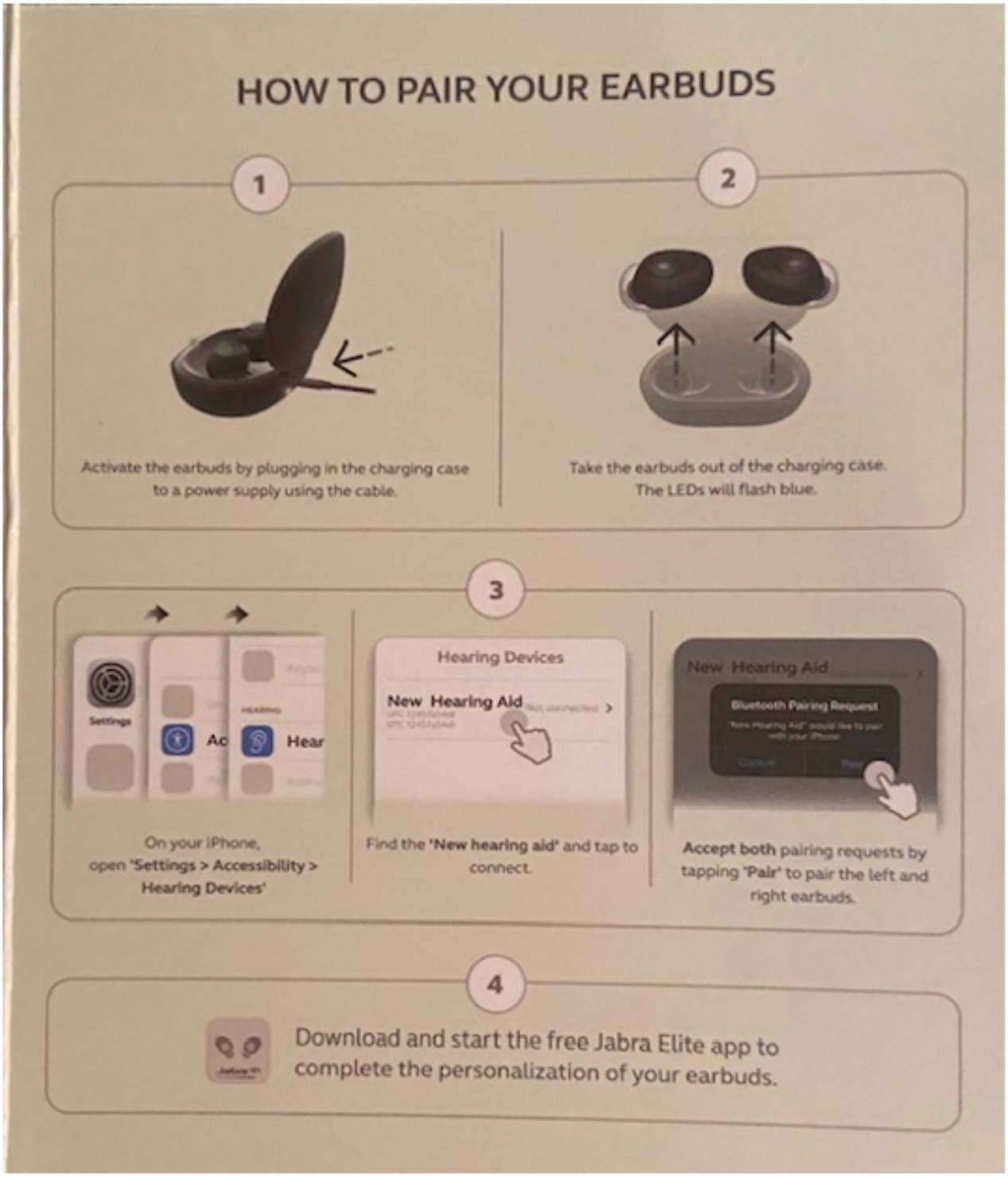
A total of 3 participants in usability testing rounds 1 and 2 failed task 1 (read outer label) since they could not determine if it was appropriate for them to use the product based on the outside of the box. We identified the same problem with nonintended users of the device, with over half of the participants in round 3 stating that they did not know that the device was a hearing aid from the labeling on the outside of the box. All 3 of the participants with severe hearing loss and 3/4 of participants with tinnitus said (incorrectly) that the device would be appropriate for them to use based on the labeling on the outside of the box.

Device Pairing

The primary usability challenge with the device was setting up and pairing the hearing aids to the iPhone (task 2). In round 1, a total of 3 of 5 participants failed this task, and a fourth participant had difficulty completing it. A common challenge was that participants did not know that they needed to return to the app to complete the hearing aid set up after pairing the hearing aids to the iPhone. Rather, participants thought they were done setting up the device once they followed the Quick Guide pairing instructions (see Figure 2). The only participant to complete the task easily followed the workflow: open the Jabra Enhance app → pair the hearing aids using the iPhone → return to the app to complete the set up. As a result, we recommended design changes to the Quick Guide prior to round 2. As shown in Figure 3, the Quick Guide was revised to inform users to first go to the Jabra Enhance app before pairing the hearing aids to the phone. We further emphasized that users should return to the Jabra Enhance app after pairing to finish set up.

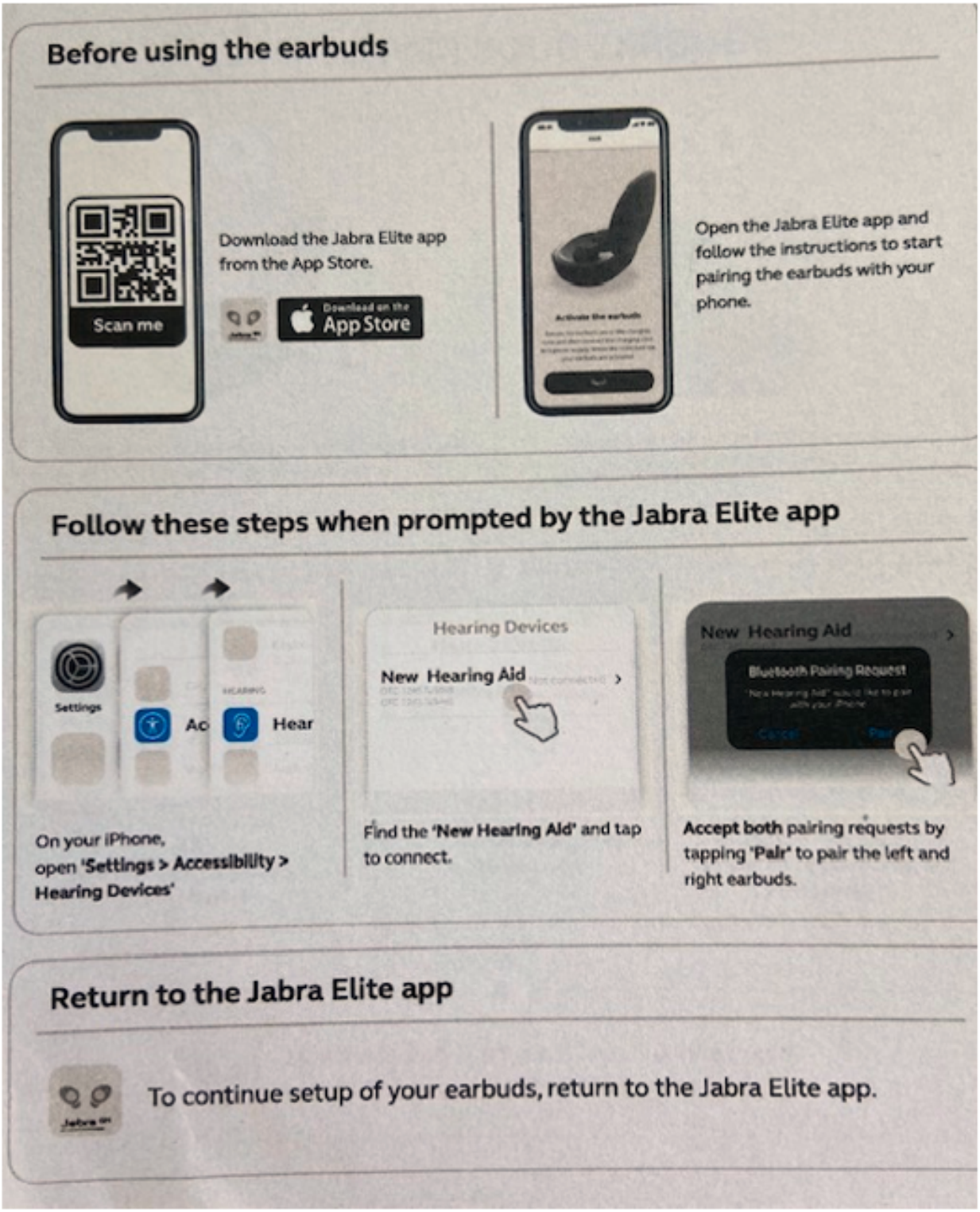
XSL•FO
RenderX

Figure 2. Quick guide instructions (December 2020 version—round 1 testing).



XSL•FO
RenderX

Figure 3. Quick guide (February 2021 version—round 2 testing).



Following these design changes, task 2 success rate increased from 40% (2 out of 5 in round 1) to 67% (10 out of 15 in round 2). Yet, there were still several use errors during pairing and setting up the earbuds in round 2. For instance, 2 (13%) participants did not understand that they needed to go to an app to set up the earbuds. Several participants had trouble reading the Quick Guide instructions as the text was too small (see Figure 3). A total of 3 participants did not know what to do after pairing the earbuds to the phone. Finally, it was unclear to participants which section of the iPhone settings to use to pair the hearing aids. Some participants had difficulty finding the "Accessibility" section in the iPhone's system settings, and some participants then went to the "Bluetooth" rather than the "Hearing Devices" section to pair the hearing aids.

Distinguishing the Right and Left Earbud

Participants had difficulty determining which earbud was for the right versus the left ear. Several participants stated that they

XSL•FO
RenderX

saw a "R" and "L" on the charging case, but not on the earbuds themselves. For instance, one participant stated, "so, there's not a right and a left?" while another participant said, "I don't see a left and right, they don't indicate that right?" Some participants were eventually able to identify the "R" and "L" markings on the earbuds stating,

> *Ah, there's an R on the end, that's really hard to see. I'd make that bigger...yeah that's really hard to see.*

The inability to distinguish the right from the left earbud resulted in several use errors, including difficulty placing the earbuds in the ears and difficulty placing the earbuds back in the charging case.

User Manual

The user manual contained information about the intended use and about most of the hazards of device use. The location of the user manual at the bottom of the box led to several use errors. Several participants were not able to find the manual. Finding the user manual was challenging since the text "User manual" was partially covered by the cardboard insert (see Figure 4). One participant said,

> *it's kind of hidden though...that's a little deceiving maybe, because I would have looked at this first.*

Another participant asked, "Where are the directions?" In total, 5 (33%) participants, all of whom identified the user manual, struggled to get it out of the box. Several more participants never took the manual out of the box; it is possible that these participants could not find the manual but did not verbalize this. Additionally, in round 1 testing, we found that there was not an entry in the table of contents to indicate on which page the "intended use" and "health and safety information" was within the user manual. This was corrected prior to round 2.

XSL•FO
RenderX

Figure 4. User manual cutout in box.



Product Design

Overall, participants were very positive about the device. They liked the size and design of the charging case (eg, small and easy to put in a pocket or purse), and they liked that the device was rechargeable, and thus did not require batteries. Participants liked the sleek and modern looking design stating, "it feels high-end" and said that the hearing aids were easy to place in their ears. They appreciated the different sized ear gels and the magnets that assured that the earbuds clicked into the correct charging position in the case. Participants were excited that they could purchase the device "off-the-shelf" and stated it would be helpful to use when they were in crowded rooms. Despite observed difficulties with setting up and pairing the hearing aids to the iPhone (task 2), participants reported that the app was easy to use. Finally, participants liked that they could adjust the hearing aids' settings on their phone.

XSL•FO
RenderX

Participants recommended several areas for improvement. Most participants desired a longer battery life (which was 6 h at the time of testing) so that they could wear the hearing aids all day without needing to recharge them. They also recommended that technical terms be better explained in the manual and on the box. One participant recommended that the hearing aids come in more discreet colors (ie, skin colored) instead of black. Participants recommended some modifications to the front of the box saying,

> *it should have more on the front about like if it's better for a noisier environment or something, because I'd be more likely to use it if it said, "will improve your hearing in a crowded environment" because I have a very hard time hearing when there is background noise.*

Some participants also found the contrast of white and red text on a gray background hard to read.

Knowledge Assessment

Table 3 details the scores for each of the 10 knowledge assessment questions across all 3 rounds of testing. Participants who used the user manual performed better overall compared to those who did not access or refer to the user manual. The user manual appeared to help the most on question 9—whether someone who had a recent episode of dizziness should use the earbuds.

Table . Knowledge assessment (percent correct answers).

#	Question topic	Round 1 (n=5), n (%)[a]	Round 2 (n=15), n (%)[b]	Round 3 (n=21), n (%)[c]
1	Tinnitus	1 (20)	13 (87)	13 (62)
2	Sinus infection	0 (0)	15 (100)	15 (71)
3	Severe hearing loss	4 (80)	8 (53)	5 (24)
4	Rapid hearing loss	1 (20)	11 (73)	11 (52)
5	Preference for earbuds	1 (20)	12 (80)	15 (71)
6	Laryngitis	3 (60)	4 (27)	7 (33)
7	Sore in ear	5 (100)	13 (87)	17 (81)
8	Ear wax	3 (60)	12 (80)	11 (52)
9	Dizziness	4 (80)	9 (60)	11 (52)
10	Ear infection	5 (100)	15 (100)	21 (100)

[a]Average percent correct: 50%.

[b]Average percent correct: 75%.

[c]Average percent correct: 60%.

Design Recommendations

Based on the identified usability issues, we recommended 9 design changes to the earbud system and its packaging (Table 4). We rated these recommendations as "high,""medium," or "low" priority. A total of 4 of the proposed changes (items 1, 2, 3, and 7 in Table 4) were implemented prior to subsequent testing rounds and all but 2 (items 4 and 5) were implemented prior to initial product release. Each design recommendation corresponds to a human factor's design principle (descriptions of the design principles can be found in Barton et al [22]).

XSL•FO
RenderX

Table . Design recommendations derived from the results of 3 rounds of usability testing.

#	Priority	Addressed before round 2 testing?	Addressed before product release?	Design recommendation	Corresponding human factors principle
1	High	Yes	Yes	• Increase the font size of the text in the Quick Guide	Readability (font and capitalization)
2	High	Yes	Yes	• Revise the Quick Set-up Guide and user manual to clearly show the following steps: • Download Jabra Enhance app from app store • Open the Jabra Enhance app (show a picture of the app) • Register and set up the hearing aids following the steps on the app • Pair the hearing aids with the iPhone (show pictures to show the steps of going to settings > accessibility > hearing devices), select "New Hearing Aid" device, accept both pairing requests • Return to the Jabra Enhance app (show picture of returning to app)	Organization (order)
3	High	Yes	Yes	• Add a section in the user manual's table of contents for "Intended use and warnings"	Organization (navigation tools)
4	High	No	No	• Add a screen in the app that lists the contraindications of use (eg, severe hearing loss, tinnitus, and dizziness). This could be a screen after the registration in the app that participants acknowledge they understand reasons they should not use the device.	Content (emphasis)

XSL•FO
RenderX

#	Priority	Addressed before round 2 testing?	Addressed before product release?	Design recommendation	Corresponding human factors principle
5	Medium	No	No	• Increase the font size of text about contraindications and intended use on the outside of the box.	Readability (font and capitalization)
6	Medium	No	Yes	• More clearly indicate on the box that the device is to support those with hearing loss. It may be helpful to add the below text from the user manual, page 1, somewhere onto the outside of the box: "This product may help you, if you: • Strain to follow conversations in both quiet and noisier environments • Miss important information during conversations • Have trouble hearing at a distance • Have trouble understanding the television or telephone calls"	Content (clarity of content)
7	Low	Yes	Yes	• Revise the app name and other aspects in the app and user materials to have a consistent term for the Jabra earbuds, for example, "Jabra Elite," "Jabra earbuds," and "Jabra hearing aid."	Comprehensibility (terminology)
8	Low	No	Yes	• Make the cut-out in the white box bigger so that the entire "User Manual" text can be seen (Figure 4), and add text directing users that the user manual is under the white box. Could say "User Manual" with an arrow pointing below the box.	Readability (layout and position)
9	Low	No	Yes		

XSL•FO
RenderX

#	Priority	Addressed before round 2 testing?	Addressed before product release?	Design recommendation	Corresponding human factors principle
				• Add a note on the "manual pairing" page of the user manual that participants should refer to the "first time use" section if it is their first time setting up the device: "if this is the first-time use, please refer to page X."	Organization (navigational tools)

Device Packaging

We recommended that the warnings and intended use information on the outside of the box be better emphasized (eg, larger font size and more color contrast) to ensure appropriate use. Due to participants' difficulty finding the user manual and getting it out of the bottom of the box, we recommended placing the user manual on the top so the user sees and touches it before accessing the product. This could improve safety as the user manual contains most of the warnings and intended use information about the device.

Pairing the Device

To better support the pairing of the device to the smartphone, we recommended that the Quick Guide be revised. Despite modifications after round 1 testing, participants still could not read or follow the Quick Guide instructions due to the small text size. We recommended that the font size in the Quick Guide be significantly increased given the older population intended to use this device. We also recommended modifying the Quick Guide to clearly outline the pairing steps (eg, go to "Settings" then "Accessibility" then "Hearing devices"). As participants did not know to return to the app to complete set up after pairing the earbuds to the iPhone, we recommended that the steps in the Quick Guide be more clearly sequentially marked (eg, 1, 2, 3,... or a, b, c,...).

Discussion

Principal Findings

In this study, we evaluated the safety and usability of an over-the-counter hearing aid for people with mild to moderate hearing loss. We conducted usability testing with 20 intended and 21 nonintended users of the device. We identified numerous usability issues hindering the optimal use of the hearing aid, and generated design recommendations to mitigate use errors and improve the overall design of the device.

This work expands our understanding of the safety and usability of novel, over-the-counter hearing devices [18,26]. Overall, we found minimal safety-related issues with the device. The primary safety risk is that the diagnosis and treatment of a serious condition, such as sudden deafness due to infection, may be delayed due to the lack of medical screening prior to the device use [27]. Consistent with this, the main challenge we found was related to the lack of understanding by nonintended users, including those with tinnitus and severe hearing loss, that they should not use the device. This issue underscores the importance of ensuring that device contraindications and safety warnings are clearly displayed and legible, preferably on the outside packaging. Further, redundant messaging should be incorporated; we recommended a verification dialog in the phone app during device set up to ensure that users understood device use contraindications.

Usability testing identified numerous usability issues with the device, which largely related to violations of human factors design principles. Some examples of violations of the principle, readability (font and capitalization), were the small font size of the contraindications on the outside of the box, the Quick Guide, and the markings of "L" and "R" on the hearing aids. We also identified issues with comprehensibility (terminology), as the device was referred to by different terms (eg, "Elite" and "Rogue") throughout the packaging and app. Based on our design recommendations, there were significant improvements in the usability of the device between the testing rounds. Human factors design principles [23,24], including those developed specifically for patient-facing medical documents [22,25], should be incorporated into the design of patient-facing medical devices. In hearing care, specific design guidelines for older adults (eg, use 12-14 pt font sizes, write in third-sixth grade level) [28] should be followed to maximize the ease of use of these devices given the substantial diversity in patients' physical and cognitive attributes and abilities, experience, and background.

We observed appreciable tension between the goals of designing for safety and designing for ease of use, aesthetics, and marketability. For instance, from a safety perspective, the device's contraindications for use (eg, tinnitus) should have been displayed in a large font such that they covered much of the back of the box. However, this would not be aesthetically appealing and there were concerns that it would deter customers from purchasing the device. Similarly, from a safety perspective, we recommended adding a hard stop in the app setup to ensure that the user understood the intended use of the device. However, this would have decreased the ease of pairing the earbuds to the phone. With over-the-counter medical devices, both product safety and aesthetic appeal are important [16] and

XSL•FO
RenderX

it can be challenging to achieve both concurrently resulting in trade-off decisions. For devices with greater safety risks, safety should be prioritized.

We found that participants were excited about an over-the-counter hearing aid and appreciated this product's design attributes. Similar to prior studies [18,26], participants liked the earbuds' discreet size and form factor, which made hearing aid use less obvious, especially given the growing general use of earbuds for listening to music or talking on the phone. Participants also liked that the device did not require batteries and that it had a small carrying case that could easily fit in a pocket or purse. Our study identified excitement about the device from the intended user population, which may result in increased access to and use of hearing aids.

We uncovered challenges in conducting usability testing with nonintended users of the device. One goal of the FDA device approval process is to identify areas of foreseeable misuse including use of the device by unintended users [11]. This is a unique issue with over-the-counter devices, as there is no clinical mediation to ensure appropriate use. During testing, we found that only 50% of nonintended users were able to determine that they should not use the hearing device. While the design of the device and packaging likely contributed to this, another challenge is that participants may have assumed they should use the device since they were recruited to participate in the study. In essence, participants may have been "tricked" into thinking they should use the device due to the nature of usability testing. Future research should explore additional methods or approaches for eliciting information from nonintended users of over-the-counter medical devices.

Another limitation of this study is that the findings are based on users' interaction with the device over only a short period of time; it is possible that as users learn to use the device, ease of use would increase or alternatively, new usability issues may emerge over prolonged use that we did not observe. We only used iPhones for testing and some of the participants were Android users. Despite diversity in participants' age, hearing ability, comfort with technology, and education level, our sample may not have included the full range of end user cognitive abilities, experiences, and backgrounds. In the real environment of use, older adults may rely on adult children or other family members to assist with device pairing and set up; our study highlights the usability issues encountered by older adults setting up the device on their own, without assistance. We also did not test actual use of the hearing aid to enhance participant hearing so we cannot draw conclusions on the function of the hearing aid technology. Future studies should evaluate the use of the hearing aid in the real environment over time to determine if it supports users' needs. There is increasing evidence, however, demonstrating efficacy of over-the-counter hearing aids in comparison to prescriptive devices [29,30].

Conclusions

This study describes the safety and usability of an over-the-counter hearing aid for adults with mild to moderate hearing loss. Based on the usability testing, we proposed human factors design recommendations to enhance the usability and safety of the device. For instance, the intended user group and contraindications for use (eg, tinnitus) should be clearly displayed on the outside of the box, as well as throughout the set up materials. Clear step-by-step guidelines should be available to support user set up and pairing of the device. We identified challenges to the design and testing of direct-to-consumer devices. As the population ages and technologies continue to pervade every aspect of our lives, the direct use of medical devices by laypersons will continue to expand. This study lays the foundation for future studies on best practices for the user interface design of direct-to-consumer medical devices.

Acknowledgments

This research was supported in part by GN Group (Ballerup, Denmark). Some of MES's effort in preparing this study was supported by the National Library of Medicine Institutional Training Program in Biomedical Informatics and Data Science through the National Institutes of Health (NIH; grant T15LM007450-19) and the Agency for Healthcare Research and Quality (AHRQ; grants K12HS026395 and K01HS029042). The content is solely the responsibility of the authors and does not necessarily represent the official views of the GN Group, NIH, or AHRQ.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Knowledge assessment.
[DOCX File, 17 KB - humanfactors_v12i1e65142_app1.docx]

Multimedia Appendix 2
System Usability Scale.
[DOCX File, 20 KB - humanfactors_v12i1e65142_app2.docx]

Multimedia Appendix 3

XSL•FO
RenderX

Interview questions.
[DOCX File, 19 KB - humanfactors_v12i1e65142_app3.docx]

References

1. Goman AM, Lin FR. Prevalence of hearing loss by severity in the United States. Am J Public Health 2016 Oct;106(10):1820-1822. [doi: 10.2105/AJPH.2016.303299] [Medline: 27552261]
2. Lin FR, Thorpe R, Gordon-Salant S, Ferrucci L. Hearing loss prevalence and risk factors among older adults in the United States. J Gerontol A Biol Sci Med Sci 2011 May;66(5):582-590. [doi: 10.1093/gerona/glr002] [Medline: 21357188]
3. Lin FR, Ferrucci L. Hearing loss and falls among older adults in the United States. Arch Intern Med 2012 Feb 27;172(4):369-371. [doi: 10.1001/archinternmed.2011.728] [Medline: 22371929]
4. Lin FR, Yaffe K, Xia J, et al. Hearing loss and cognitive decline in older adults. JAMA Intern Med 2013 Feb 25;173(4):293-299. [doi: 10.1001/jamainternmed.2013.1868] [Medline: 23337978]
5. Thomson RS, Auduong P, Miller AT, Gurgel RK. Hearing loss as a risk factor for dementia: a systematic review. Laryngoscope Investig Otolaryngol 2017 Apr;2(2):69-79. [doi: 10.1002/lio2.65] [Medline: 28894825]
6. Reed NS, Garcia-Morales EE, Myers C, et al. Prevalence of hearing loss and hearing aid use among US Medicare beneficiaries aged 71 years and older. JAMA Netw Open 2023 Jul 3;6(7):e2326320. [doi: 10.1001/jamanetworkopen.2023.26320] [Medline: 37505496]
7. Chien W, Lin FR. Prevalence of hearing aid use among older adults in the United States. Arch Intern Med 2012 Feb 13;172(3):292-293. [doi: 10.1001/archinternmed.2011.1408] [Medline: 22332170]
8. Hearing Health Care for Adults: Priorities for Improving Access and Affordability: National Academies Press; 2016.
9. Food and Drug Administration, Department of Health and Human Services (HHS). Medical devices; ear, nose, and throat devices; establishing over-the-counter hearing aids. Fed Regist 2022 Aug 17;87.
10. Over-the-counter (OTC) medical devices: considerations for device manufacturers. U.S. Food & Drug Administration. URL: https://www.fda.gov/medical-devices/products-and-medical-procedures/over-counter-otc-medical-devices-considerations-device-manufacturers [accessed 2023-08-10]
11. Applying human factors and usability engineering to medical devices. : U.S. Department of Health and Human Services Office of Device Evaluation; 2016.
12. Weinger MB, Wiklund ME, Gardner-Bonneau DJ. Handbook of Human Factors in Medical Device Design: CRC Press; 2010.
13. Anders S, Miller A, Joseph P, et al. Blood product positive patient identification: comparative simulation-based usability test of two commercial products. Transfusion 2011 Nov;51(11):2311-2318. [doi: 10.1111/j.1537-2995.2011.03185.x] [Medline: 21599676]
14. Harris B, Weinger MB. An insidious failure of an oxygen analyzer. Anesth Analg 2006 May;102(5):1468-1472. [doi: 10.1213/01.ane.0000205814.47841.16]
15. Weinger MB. Anesthesia equipment and human error. J Clin Monit Comput 1999 Jul;15(5):319-323. [doi: 10.1023/a:1009919207870] [Medline: 12568138]
16. Gardner-Bonneau DJ. Home health care. In: Weinger M, Wiklund M, Gardner-Bonneau DJ, editors. Handbook of Human Factors in Medical Device Design: FL: CRC Press; 2011:747-770.
17. Kingsley PA, Backinger CL, Brady MW. Medical-device user instructions: the patient's need, the nurse's role. Home Healthc Nurse 1995;13(1):31-37. [Medline: 7860339]
18. Maidment DW, Ali YHK, Ferguson MA. Applying the COM-B model to assess the usability of smartphone-connected listening devices in adults with hearing loss. J Am Acad Audiol 2019 May;30(5):417-430. [doi: 10.3766/jaaa.18061] [Medline: 31044693]
19. Story MF, Luce AC, Rempel DM. Accessibility and usability of diabetes management devices for users with vision disabilities. HFES Ann Conf Proc 2009 Oct 19;53(11):699-703. [doi: 10.1518/107118109X12524442636706]
20. Benda NC, Creber RMM. Underserved populations: integrating social determinants of health into the study of patient work. In: Valdez RS, Boca R, Holden RJ, editors. The Patient Factors: Applications of Patient Ergonimcs: CRC Press; 2021, Vol. 1:163-186. [doi: 10.1201/9780429293009-13]
21. Brooke J. SUS: a "quick and dirty" usability scale. In: Jordan PW, Werdmeester BA, McClelland AL, editors. Usability Evaluation in Industry: Taylor & Francis; 1996:189-194.
22. Barton HJ, Salwei ME, Rutkowski RA, et al. Evaluating the usability of an emergency department after visit summary: staged heuristic evaluation. JMIR Hum Factors 2023 Mar 9;10(1):e43729. [doi: 10.2196/43729] [Medline: 36892941]
23. Zhang J, Johnson TR, Patel VL, Paige DL, Kubose T. Using usability heuristics to evaluate patient safety of medical devices. J Biomed Inform 2003;36(1-2):23-30. [doi: 10.1016/s1532-0464(03)00060-1] [Medline: 14552844]
24. Scapin DL, Bastien JMC. Ergonomic criteria for evaluating the ergonomic quality of interactive systems. Behav Inf Technol 1997 Jan;16(4-5):220-231. [doi: 10.1080/0144929971198806]
25. Tremoulet P, Krishnan R, Karavite D, et al. A heuristic evaluation to assess use of after visit summaries for supporting continuity of care. Appl Clin Inform 2018 Jul;9(3):714-724. [doi: 10.1055/s-0038-1668093] [Medline: 30208496]

XSL•FO
RenderX

26. Gomez R, Habib A, Maidment DW, Ferguson MA. Smartphone-connected hearing aids enable and empower self-management of hearing loss: a qualitative interview study underpinned by the behavior change wheel. Ear Hear 2022;43(3):921-932. [doi: 10.1097/AUD.0000000000001143] [Medline: 34653030]
27. Manchaiah V, Taylor B, Dockens AL, et al. Applications of direct-to-consumer hearing devices for adults with hearing loss: a review. Clin Interv Aging 2017;12:859-871. [doi: 10.2147/CIA.S135390] [Medline: 28553093]
28. Caposecco A, Hickson L, Meyer C. Assembly and insertion of a self-fitting hearing aid: design of effective instruction materials. Trends Amplif 2011 Dec;15(4):184-195. [doi: 10.1177/1084713811430837] [Medline: 22389434]
29. Humes LE, Rogers SE, Quigley TM, Main AK, Kinney DL, Herring C. The effects of service-delivery model and purchase price on hearing-aid outcomes in older adults: a randomized double-blind placebo-controlled clinical trial. Am J Audiol 2017 Mar 1;26(1):53-79. [doi: 10.1044/2017_AJA-16-0111] [Medline: 28252160]
30. De Sousa KC, Manchaiah V, Moore DR, Graham MA, Swanepoel DW. Effectiveness of an over-the-counter self-fitting hearing aid compared with an audiologist-fitted hearing aid: a randomized clinical trial. JAMA Otolaryngol Head Neck Surg 2023 Jun 1;149(6):522-530. [doi: 10.1001/jamaoto.2023.0376] [Medline: 37052929]

Abbreviations

FDA: Food and Drug Administration
SUS: System Usability Scale

Edited by K Adapa; submitted 06.08.24; peer-reviewed by D Urbanski, M Nickbakht; revised version received 30.10.24; accepted 01.11.24; published 20.01.25.

Please cite as:
Salwei ME, Anders S, Reale C, Slagle JM, Ricketts T, Weinger MB
Evaluating the Safety and Usability of an Over-the-Counter Medical Device for Adults With Mild to Moderate Hearing Loss: Formative and Summative Usability Testing
JMIR Hum Factors 2025;12:e65142
URL: https://humanfactors.jmir.org/2025/1/e65142
doi:10.2196/65142

© Megan Elizabeth Salwei, Shilo Anders, Carrie Reale, Jason M Slagle, Todd Ricketts, Matthew B Weinger. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 20.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Predictive Factors and the Predictive Scoring System for Falls in Acute Care Inpatients: Retrospective Cohort Study

Chihiro Saito[1,2], MPH; Eiji Nakatani[2,3,4], BSP, MS, PhD; Hatoko Sasaki[2], MPH, DrPH; Naoko E Katsuki[5], MD, PhD; Masaki Tago[5], MD, PhD; Kiyoshi Harada[6], MD, PhD

[1]Department of Nursing, Shizuoka General Hospital, Shizuoka, Japan

[2]Graduate School of Public Health, Shizuoka Graduate University of Public Health, 4-27-2, Kita-ando, Aoi-ku, Shizuoka, Japan

[3]Research Support Center, Shizuoka General Hospital, Shizuoka, Japan

[4]Department of Biostatistics and Health Data Science, Graduate School of Medical Science, Nagoya City University, Nagoya, Japan

[5]Department of General Medicine, Saga University Hospital, Saga, Japan

[6]Department of Medical Safety, Shizuoka General Hospital, Shizuoka, Japan

Corresponding Author:
Eiji Nakatani, BSP, MS, PhD
Graduate School of Public Health, Shizuoka Graduate University of Public Health, 4-27-2, Kita-ando, Aoi-ku, Shizuoka, Japan

Abstract

Background: Falls in hospitalized patients are a serious problem, resulting in physical injury, secondary complications, impaired activities of daily living, prolonged hospital stays, and increased medical costs. Establishing a fall prediction scoring system to identify patients most likely to fall can help prevent falls among hospitalized patients.

Objectives: This study aimed to identify predictive factors of falls in acute care hospital patients, develop a scoring system, and evaluate its validity.

Methods: This single-center, retrospective cohort study involved patients aged 20 years or older admitted to Shizuoka General Hospital between April 2019 and September 2020. Demographic data, candidate predictors at admission, and fall occurrence reports were collected from medical records. The outcome was the time from admission to a fall requiring medical resources. Two-thirds of cases were randomly selected as the training set for analysis, and univariable and multivariable Cox regression analyses were used to identify factors affecting fall risk. We scored the fall risk based on the estimated hazard ratios (HRs) and constructed a fall prediction scoring system. The remaining one-third of cases was used as the test set to evaluate the predictive performance of the new scoring system.

Results: A total of 13,725 individuals were included. During the study period, 2.4% (326/13,725) of patients experienced a fall. In the training dataset (n=9150), Cox regression analysis identified sex (male: HR 1.60, 95% CI 1.21 - 2.13), age (65 to <80 years: HR 2.26, 95% CI 1.48 - 3.44; ≥80 years: HR 2.50, 95% CI 1.60 - 3.92 vs 20-<65 years), BMI (18.5 to <25 kg/m²: HR 1.36, 95% CI 0.94 - 1.97; <18.5 kg/m²: HR 1.57, 95% CI 1.01 - 2.44 vs ≥25 kg/m²), independence degree of daily living for older adults with disabilities (bedriddenness rank A: HR 1.81, 95% CI 1.26 - 2.60; rank B: HR 2.03, 95% CI 1.31 - 3.14; rank C: HR 1.23, 95% CI 0.83 - 1.83 vs rank J), department (internal medicine: HR 1.23, 95% CI 0.92 - 1.64; emergency department: HR 1.81, 95% CI 1.26 - 2.60 vs department of surgery), and history of falls within 1 year (yes: HR 1.66, 95% CI 1.21 - 2.27) as predictors of falls. Using these factors, we developed a fall prediction scoring system categorizing patients into 3 risk groups: low risk (0-4 points), intermediate risk (5-9 points), and high risk (10-15 points). The c-index indicating predictive performance in the test set (n=4575) was 0.733 (95% CI 0.684 - 0.782).

Conclusions: We developed a new fall prediction scoring system for patients admitted to acute care hospitals by identifying predictors of falls in Japan. This system may be useful for preventive interventions in patient populations with a high likelihood of falling in acute care settings.

(JMIR Hum Factors 2025;12:e58073) doi:10.2196/58073

KEYWORDS

falls; inpatient falls; acute care hospital; predictive factor; risk factors

Introduction

In 2022, the proportion of the population aged 65 years or older in Japan reached a record high of 29% [1]. A 2017 survey by the Ministry of Health, Labour and Welfare of Japan revealed that patients aged 75 years or older accounted for 41.5% of all patients admitted to acute care hospitals [2]. Given the trend in population aging, it is projected that the number of older adult

XSL•FO
RenderX

patients with a high risk of falling will further escalate in the future. Falls are not limited to older adults, and falls in hospitalized patients can lead to severe physical injuries, secondary complications, a marked decline in activities of daily living (ADL), and even death in extreme cases [3,4]. Therefore, fall prevention has become an important issue to protect patients' lives and quality of life [5].

Interventions for fall prevention must be strategically targeted to populations with a high risk of falling during hospitalization. Furthermore, previous studies have emphasized the importance of patient exercise therapy [6,7] and education for both patients and health care providers in fall prevention [8-12]. Educating patients about the risks of falls and strategies to mitigate these risks is crucial in reducing the incidence of falls in hospitalized patients. To effectively conduct patient education, it is imperative to construct a fall prediction model for the accurate identification of these high-risk patients. Currently, fall prevention measures in hospitals include fall prediction models using information from electronic health record (EHR) systems [13-22], as well as predictive models that analyze patient information from EHRs and nursing records using artificial intelligence [15,23]. Here, we present several fall prediction models that can be used with EHRs [17-20]. The STRATIFY scale [17] uses a history of falls, visual impairment, mental status, frequency of elimination, and ability to transfer and move as factors in a prediction model. The Morse Fall Scale [18] includes 6 items related to a history of falls, comorbidities, use of walking aids, intravenous fluids, ability to walk and move, and mental status. The Medication Fall Risk Score and Evaluation Tool [19] assesses the medication-related fall risk. This tool considers a patient's use of medications as predictors, classified according to the associated risk levels. Tago et al [20] reported 8 predictors of falls in people with disabilities in Japan: age, sex, emergency hospitalization, admission to neurosurgery, use of sleeping pills, history of falls, independence in eating, and level of independence in daily living.

Although fall prediction tools are widely used in Japanese hospitals, existing models vary in predictors and are often difficult to apply due to differences in facility and patient characteristics. Many hospitals also rely on tools that lack a strong evidence base [24,25] and are not well integrated with EHR systems. In our clinical environment, we found it challenging to implement existing models due to their reliance on numerous variables that are either difficult to extract from the EHR or unavailable. Therefore, we aimed to develop a streamlined fall risk assessment tool that considers facility-specific factors and can be seamlessly integrated with EHR systems, enabling real-time insights and more efficient fall prevention strategies.

The purpose of this study was to identify predictive factors for falls in patients admitted to an acute care hospital as well as to develop a scoring system using these factors and evaluate its validity.

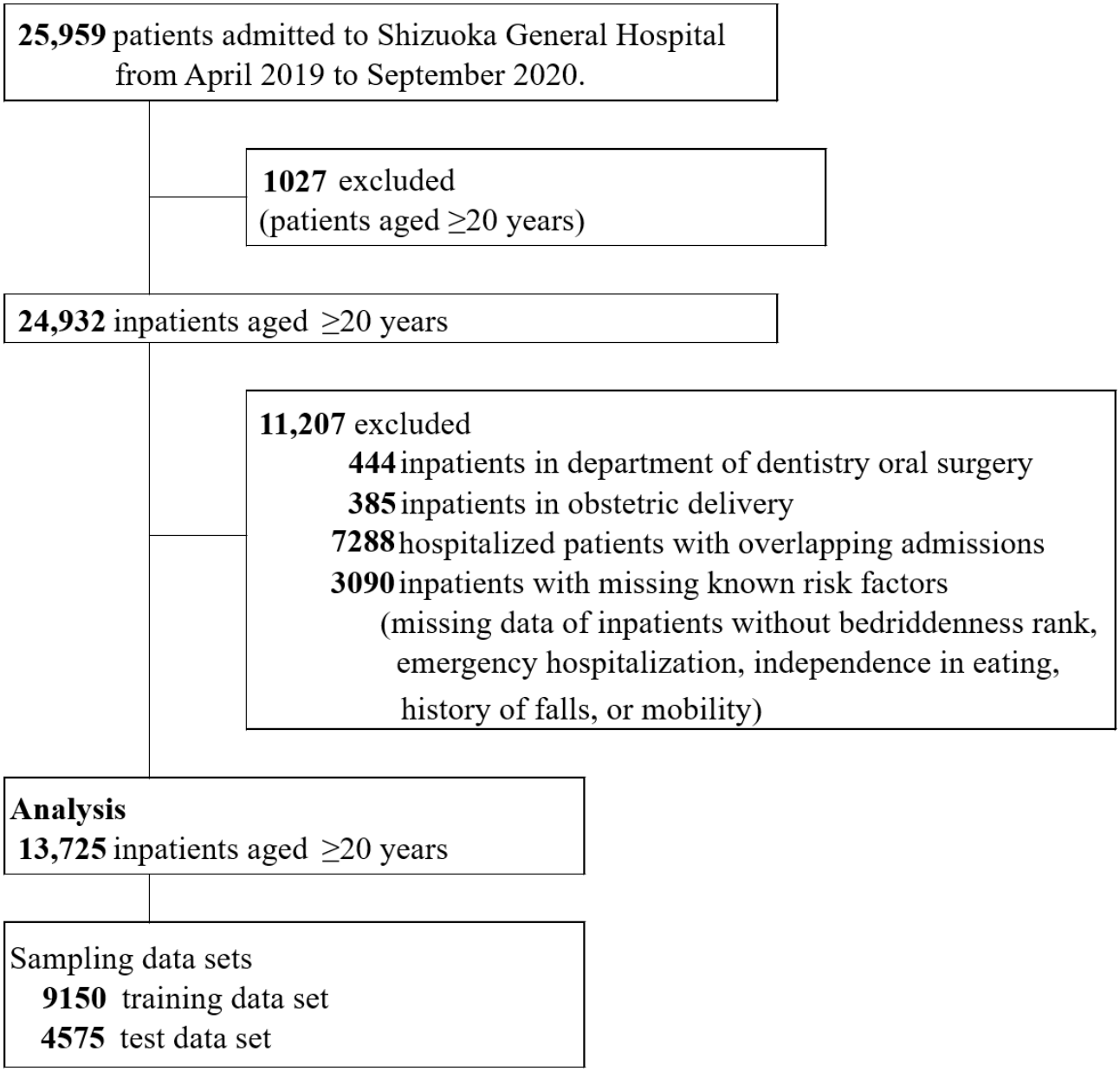
Methods

Participants and Study Design

We conducted a retrospective cohort study of patients aged ≥20 years admitted to Shizuoka General Hospital between April 2019 and September 2020. Inpatients excluded from the study included those not covered by the Diagnosis Procedure Combination system, such as dental and oral surgery patients and obstetrics and gynecology inpatients during pregnancy, childbirth, and postpartum. In addition, inpatients lacking data on known risk factors such as the degree of independence in daily living for the older adults with disabilities, bedriddenness rank (BR), emergency admission, dietary independence, mobility, and history of falls in the past year were also excluded, as illustrated in Figure 1.

XSL•FO
RenderX

Figure 1. Flowchart illustrating the study population.



Variables at Hospitalization as Candidate Predictive Factors

All data used in this study were extracted from patients' medical records as of February 3, 2022. Preadmission medical history variables included dementia [5,6], Parkinson disease [7,26], stroke [6], visual impairment (with or with no diagnosis of glaucoma or cataracts) [4,5], history of falls [5,16,20,27], and use of sleep medications [20,28-30]. The following variables at admission were also collected: age [6,16,20,27], sex [20,27,28], BMI [31-35], date of admission [20], disease name at admission, department [20], mode of admission, ambulance transport, consciousness disorders [16,23,28], requirement for nursing care, good sleep condition, use of sleeping medication [20,28,30], status of medication management, BR [20,36], Cognitive Function Scores [5,6], ADL at admission (eating, transferring, dressing, toilet transfer or use, bathing, level walking, stair use, changing clothes, defecation management, and urinary management) [17,18,20], fall assessment end points at admission (history of falls or falls within 1 year, inability to stand without holding on [28], impaired judgment and comprehension, toilet assistance, and use of portable toilet), and presence of physical restraint screening at admission.

The BR [20,36] is an official assessment tool in Japan's long-term care insurance system [37]. The BR is an assessment of the degree to which a person's daily life is restricted; this degree is mainly assessed in terms of mobility in daily life, such as whether the person is independent, in a wheelchair, or bed bound. The Ministry of Health, Labour and Welfare ranks the degree of BR based on evaluations by nurses and other health care professionals according to the daily care they provide during hospitalization, as well as on reports from family members. The procedure for assessing BR and its reliability have been reported [20,36].

Falls During Hospitalization as an Outcome

The primary end point was the time from the date of admission to a fall at incident level [38] 2 or higher (hereafter referred to

XSL•FO
RenderX

as “fall”), which requires medical resources. For patients who died during hospitalization and those who did not have a fall, the date of death or date of discharge, respectively, was used as the censoring date. The classification of incident levels is shown in Table S1 in Multimedia Appendix 1.

Statistical Analysis

Demographic data and potential predictors at the time of hospital admission were summarized as follows. Continuous variables are described using mean (SD) or median (range), considering the distribution type. Categorical variables are summarized as frequency (%). For comparisons between the groups with and with no fall occurrence, *t* tests were used for continuous variables and chi-squared tests for categorical variables. The Kaplan-Meier method was applied to estimate the fall rate.

We explored predictors of falls and constructed a predictive model using two-thirds of the total cases, randomly selected as the training group, with the remaining one-third serving as the validation group for the scoring system. In the training group, predictive factors were identified using the Cox proportional hazards model, and we calculated hazard ratios (HRs), 95% confidence intervals (CIs), and *P* values. Predictive candidates that were significant ($P<.05$) in the comparison of backgrounds among patients with and with no falls, along with known predictors, were included in a multivariable model. Factors with *P* value of <.2 in this model were identified as predictors of inhospital falls. Independence between explanatory variables was confirmed using an absolute value of Spearman rank correlation coefficient of >0.4. Among 2 correlated variables, 1 was chosen based on ease of collection or clinical significance; this variable was then included in the multivariable model. According to the identified predictive factors, a score was created for each HR, and these scores were summed. We used a method called “conditional inference tree analysis” to categorize patients into 3 groups based on their risk of falling. This approach works by first dividing the data into 2 groups based on their overall scores. Then, a statistical test is performed to see whether these 2 groups are significantly different, and the variable that shows the strongest difference (the one with the lowest *P* value) is used to split the groups. This process is repeated within each subgroup until no further meaningful divisions can be made or the smallest group size allowed is reached. The predictive performance of this score and the fall risk groups in the validation group were evaluated using the c-index.

The significance level of the 2-tailed test was set at .05. Missing values were not imputed in the analyses. All analyses were performed using R (version 4.1.1; The R Foundation for Statistical Computing), EZR (version 1.54; Saitama Medical Center, Jichi Medical University) [39], and IBM SPSS (version 28; IBM Corp).

Ethical Considerations

This study conformed to the Ethical Principles for Medical Research Involving Human Subjects issued by the Ministry of Health, Labour and Welfare and the Ministry of Education, Culture, Sports, Science and Technology of Japan. Following these guidelines, the Shizuoka General Hospital research ethical committee determined that individual patient informed consent was not required because we analyzed existing information in this study, and patients were given the right to refuse participation via disclosure. After obtaining committee approval (SGHIRB #2020075; January 15, 2022) and publishing the disclosure document on Shizuoka General Hospital’s website, the information for each individual was anonymized, and the analysis was conducted.

Results

Patient Background and Falls at Incident Level 2 or Higher on Admission

From April 1, 2019, to September 30, 2020, a total of 24,932 inpatients aged 20 years or older were admitted to Shizuoka General Hospital. We excluded 3.3% (829/24,932) of patients not covered by the Diagnosis Procedure Combination, 29.2% (7288/24,932) with duplicate admissions, and 12.4% (3090/24,932) with missing known risk factors of falls. Consequently, 55% (13,725/24,932) of patients were included in the analysis (Figure 1).

During the observation period, defined as the length of hospital stay (median [range]: 13 [1-271] days), 3.6% (489/13,725) of patients experienced falls across all incident levels, of which 2.4% (326/13,725) falls were classified as incident level 2 or higher. For this study, we used fall data meticulously managed by the hospital’s medical safety department, ensuring accuracy and reliability. The details are shown in Table S2 in Multimedia Appendix 1. The median age (range) of the patients included in the analysis was 66 (20-103) years, with 52.1% (7150/13,725) male patients, and the median BMI (range) was 22.8 (9.6 - 58.1) kg/m^2. Table S3 in Multimedia Appendix 1 shows the results of the comparison between the backgrounds of patients in the training and test datasets. In a univariable analysis of the training dataset (n=9150; two-thirds of the study population), we compared patient backgrounds according to the presence of falls (Table 1).

XSL•FO
RenderX

Table . Patient background on admission[a].

Variable		With no fall (n=8934)	With fall (n=216)	*P* value
Age				<.001
	Age (years), mean (SD)	65.9 (17)	75.0 (11.5)	
Age				<.001
	20 to <65 years, n (%)	3347 (37.5)	28 (13)	
	65 to <80 years, n (%)	3673 (41.1)	105 (48.6)	
	≥80 years, n (%)	1914 (21.4)	83 (38.4)	
Sex, n (%)				<.001
	Female	4287 (48)	75 (34.7)	
	Male	4647 (52)	141 (65.3)	
BMI				<.001
	BMI, mean (SD)	22.9 (4.2)	21.1 (3.9)	
BMI				<.001
	<18.5 kg/m^2, n (%)	1142 (12.8)	48 (22.2)	
	18.5 to <25 kg/m^2, n (%)	5363 (60)	132 (61.1)	
	≥25 kg/m^2, n (%)	2403 (26.9)	36 (16.7)	
	Missing, n (%)	26 (0.3)	0 (0)	
Dementia, n (%)				.927
	No	8697 (97.3)	211 (97.7)	
	Yes	237 (2.7)	5 (2.3)	
Parkinson disease, n (%)				.592
	No	8861 (99.2)	213 (98.6)	
	Yes	73 (0.8)	3 (1.4)	
Stroke, n (%)				.003
	No	8121 (90.9)	183 (84.7)	
	Yes	813 (9.1)	33 (15.3)	
Visual impairment, n (%)				.778
	No	7991 (89.4)	195 (90.3)	
	Yes	943 (10.6)	21 (9.7)	
Cognitive function score, n (%)				<.001
	No	7855 (87.9)	179 (82.9)	
	Yes	710 (7.9)	37 (17.1)	
	Missing	369 (4.1)	0 (0)	
Ambulance transport, n (%)				<.001
	No	7432 (83.2)	156 (72.2)	
	Yes	1502 (16.8)	60 (27.8)	
Emergency admission, n (%)				<.001
	Scheduled hospitalization	5528 (61.9)	74 (34.3)	
	Emergency hospitalization	3406 (38.1)	142 (65.7)	
Department, n (%)				.007
	Internal medicine	4496 (50.3)	118 (54.6)	
	Department of surgery	4025 (45.1)	80 (37)	

XSL•FO
RenderX

Variable		With no fall (n=8934)	With fall (n=216)	*P* value
	Emergency department	413 (4.6)	18 (8.3)	
Consciousness disorders, n (%)				<.001
	No	8022 (89.8)	167 (77.3)	
	Yes	912 (10.2)	49 (22.7)	
Bedriddenness rank, n (%)				<.001
	Rank J	5500 (61.6)	65 (30.1)	
	Rank A	1568 (17.6)	58 (26.9)	
	Rank B	526 (5.9)	34 (15.7)	
	Rank C	1340 (15)	59 (27.3)	
Eating, n (%)				<.001
	Independent	7067 (79.1)	107 (49.5)	
	Requiring assistance	1867 (20.9)	109 (50.5)	
Transferring, n (%)				<.001
	Independent	6579 (73.6)	89 (41.2)	
	Requiring assistance	2355 (26.4)	127 (58.8)	
Dressing, n (%)				<.001
	Independent	7065 (79.1)	112 (51.9)	
	Requiring assistance	1853 (20.7)	104 (48.1)	
	Missing	16 (0.2)	0 (0)	
Toilet transfer or use, n (%)				<.001
	Independent	6820 (76.3)	94 (43.5)	
	Requiring assistance	2099 (23.5)	121 (56)	
	Missing	15 (0.2)	1 (0.5)	
Bathing, n (%)				<.001
	Independent	6673 (74.7)	95 (44)	
	Requiring assistance	2042 (22.9)	114 (52.8)	
	Missing	219 (2.5)	7 (3.2)	
Level walking, n (%)				<.001
	Independent	6707 (75.1)	93 (43.1)	
	Requiring assistance	2107 (23.6)	122 (56.5)	
	Missing	120 (1.3)	1 (0.5)	
Stair use, n (%)				<.001
	Independent	6541 (73.2)	86 (39.8)	
	Requiring assistance	1986 (22.2)	111 (51.4)	
	Missing	407 (4.6)	19 (8.8)	
Changing clothes, n (%)				<.001
	Independent	6799 (76.1)	97 (44.9)	
	Requiring assistance	2116 (23.7)	119 (55.1)	
	Missing	19 (0.2)	0 (0)	
Defecation management, n (%)				<.001
	Independent	7254 (81.2)	127 (58.8)	
	Requiring assistance	1618 (18.1)	87 (40.3)	
	Missing	62 (0.7)	2 (0.9)	

XSL•FO
RenderX

Variable		With no fall (n=8934)	With fall (n=216)	*P* value
Urination management, n (%)				<.001
	Independent	7243 (81.1)	126 (58.3)	
	Requiring assistance	1625 (18.2)	88 (40.7)	
	Missing	66 (0.7)	2 (0.9)	
History of falls within 1 year, n (%)				<.001
	No	7830 (87.6)	153 (70.8)	
	Yes	1104 (12.4)	63 (29.2)	
Inability to stand without holding, n (%)				<.001
	No	5318 (59.5)	53 (24.5)	
	Yes	3486 (39)	159 (73.6)	
	Missing	130 (1.5)	4 (1.9)	
Impaired judgment and comprehension, n (%)				<.001
	No	7592 (85)	157 (72.7)	
	Yes	1116 (12.5)	53 (24.5)	
	Missing	226 (2.5)	6 (2.8)	
Toileting assistance, n (%)				<.001
	No	6694 (74.9)	106 (49.1)	
	Yes	1795 (20.1)	99 (45.8)	
	Missing	445 (5)	11 (5.1)	

[a]Between-group comparisons were made using *t* tests and χ^2 tests for continuous and categorical variables, respectively. The *P* value was calculated using the Wald test. Bedriddenness rank J=independent or autonomous, rank A=housebound, rank B=chair, and rank C=bedridden.

Predictors of Falls

In the training dataset, univariable Cox regression analysis compared patient backgrounds based on the presence or absence of falls and identified factors affecting the time from admission to the date of fall (Table 2, left side). Correlations between explanatory variables were checked using Spearman correlation coefficients, and independent explanatory variables were entered into a multivariable regression model. The variables were narrowed down by applying a high absolute value of the Spearman correlation coefficient (>0.4) (Table S4 in Multimedia Appendix 1). Among the 21 variables that were significant in the univariable analysis, 5 variables (age, sex, BMI, department, and history of falls within 1 year) were not correlated with each other. Sixteen variables (including 12 ADL-related variables, emergency admission, and consciousness disorders) had a correlation coefficient of 0.4 or higher. From the correlated variables, BR was ultimately chosen. This decision was influenced by the fact that 12 other ADL variables showed correlation, and BR could potentially explain physical severity, including factors such as emergency transport and impaired consciousness. The variables were chosen based on ease of collection or clinical significance.

XSL•FO
RenderX

Table . Univariable and multivariable Cox regression analysis results for fall rates in the training dataset[a].

Variable (reference) and category	Training dataset (n=9150)					
	Univariable model			Multivariable model		
	HR[b]	95% CI	*P* value	HR	95% CI	*P* value
Age						
1 year	1.03	1.02 - 1.04	<.001	—[c]	—	—
Age (20 to <65 years)						
65 to <80 years	2.71	1.79 - 4.12	<.001	2.26	1.48 - 3.44	<.001
≥80 years	3.53	2.29 - 5.43	<.001	2.50	1.60 - 3.92	—
Sex (female)						
Male	1.60	1.21 - 2.12	.001	1.60	1.21 - 2.13	.001
BMI						
1 kg/m^2	0.92	0.89 - 0.96	<.001	—	—	—
BMI (≥25 kg/m ^{2})						
18.5 to <25 kg/m^2	1.57	1.09 - 2.27	.017	1.36	0.94 - 1.97	.127
≤18.5 kg/m^2	2.01	1.30 - 3.10	.002	1.57	1.01 - 2.44	—
Ambulance transport (no)						
Yes	1.77	1.33 - 2.36	<.001	—	—	—
Emergency admission (scheduled hospitalization)						
Emergency hospitalization	1.22	0.90 - 1.64	.203	—	—	—
Stroke (no)						
Yes	1.18	0.81 - 1.72	.376	—	—	—
Bedriddenness rank (rank J)						
Rank A	2.27	1.59 - 3.23	<.001	1.81	1.26 - 2.60	.001
Rank B	2.93	1.93 - 4.45	<.001	2.03	1.31 - 3.14	—
Rank C	1.94	1.35 - 2.77	<.001	1.23	0.83 - 1.83	—
Cognitive function score (no)						
Yes	1.40	0.98 - 2.01	.063	—	—	—
Department (department of surgery)						
Internal medicine	1.36	1.02 - 1.81	.034	1.23	0.92 - 1.64	.119
Emergency department	2.16	1.29 - 3.60	.003	1.81	1.26 - 2.60	—
Consciousness disorders (no)						
Yes	1.47	1.06 - 2.03	.020	—	—	—
Eating (independent)						
Requiring assistance	2.05	1.56 - 2.69	<.001	—	—	—
Transferring (independent)						
Requiring assistance	2.12	1.60 - 2.80	<.001	—	—	—
Dressing (independent)						

XSL•FO
RenderX

Variable (reference) and category	Training dataset (n=9150)					
	Univariable model			Multivariable model		
	HR[b]	95% CI	*P* value	HR	95% CI	*P* value
Requiring assistance	1.88	1.43 - 2.48	<.001	—	—	—
Toilet transfer or use (independent)						
Requiring assistance	2.22	1.68 - 2.93	<.001	—	—	—
Bathing (independent)						
Requiring assistance	2.10	1.59 - 2.78	<.001	—	—	—
Level walking (independent)						
Requiring assistance	2.18	1.65 - 2.88	<.001	—	—	—
Stair use (independent)						
Requiring assistance	2.17	1.63 - 2.90	<.001	—	—	—
Changing clothes (independent)						
Requiring assistance	2.10	1.60 - 2.77	<.001	—	—	—
Defecation management (independent)						
Requiring assistance	1.63	1.23 - 2.15	.001	—	—	—
Urination management (independent)						
Requiring assistance	1.64	1.24 - 2.17	.001	—	—	—
History of falls within 1 year (no)						
Yes	2.01	1.50 - 2.70	<.001	1.66	1.21 - 2.27	.002
Inability to stand without holding (no)						
Yes	2.51	1.83 - 3.45	<.001	—	—	—
Impaired judgment and comprehension (no)						
Yes	1.40	1.03 - 1.92	.035	—	—	—
Toileting assistance (no)						
Yes	2.10	1.59 - 2.77	<.001	—	—	—

[a]Between-group comparisons were made using *t* tests and χ^2 tests for continuous and categorical variables, respectively. The *P* value was calculated using the Wald test. Bedriddenness rank J= independent or autonomous, rank A=housebound, rank B=chair, and rank C= bedridden.

[b]HR: hazard ratio.

[c]Not applicable.

In multivariable analysis, age (65 to <80 years: HR 2.26, 95% CI 1.48 - 3.44; ≥80 years: HR 2.50, 95% CI 1.60 - 3.92 vs 20 to <65 years), sex (male: HR 1.60, 95% CI 1.21 - 2.13), BMI (18.5 to <25 kg/m²: HR 1.36, 95% CI 0.94 - 1.97; <18.5 kg/m²: HR 1.57, 95% CI 1.01 - 2.44 vs ≥25 kg/m²), BR (rank A: HR 1.81, 95% CI 1.26 - 2.60; rank B: HR 2.03, 95% CI 1.31 - 3.14; rank C: HR 1.23, 95% CI 0.83 - 1.83 vs rank J), emergency department (internal medicine: HR 1.23, 95% CI 0.92 - 1.64; emergency department: HR 1.81, 95% CI 1.26 - 2.60 vs department of surgery), and history of falls within 1 year (yes: HR 1.66, 95% CI 1.21 - 2.27) are shown as predictors of inhospital falls (Table 2, right side).

Construction of a Fall Scoring System

Based on the results of multivariable analysis using the training set, we weighted the scores based on each HR (Table 2, right side) and formed 3 fall-risk groups (low risk: 0-4 points, moderate risk: 5-9 points, and high risk: 10-15 points) using conditional inference tree analysis (Figure S1 in Multimedia Appendix 1). The new fall prediction scoring system built on this basis is shown in Figure 2. As a predictor of the performance of these 3 classifications of fall prediction, the c-index in the validation set (n=4561) was 0.733 (95% CI 0.684 - 0.782). The cumulative fall incidences in training and test datasets are

XSL•FO
RenderX

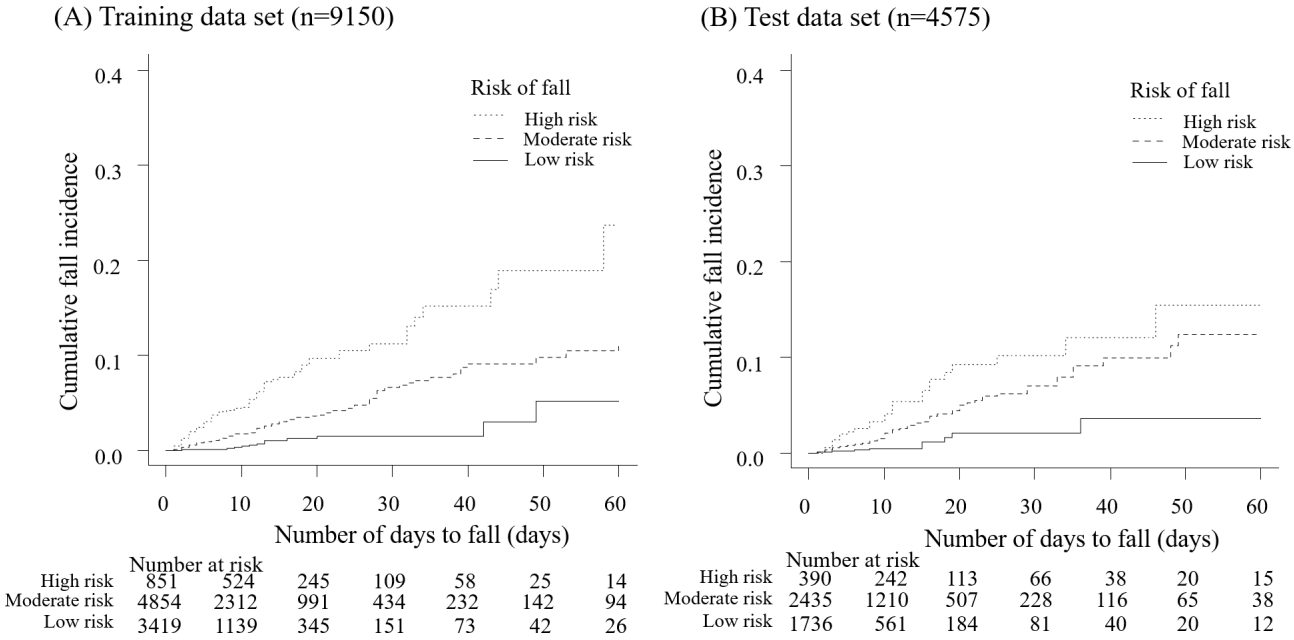
shown, with Kaplan-Meier curves presented for the training dataset (Figure 3A) and the test dataset (Figure 3B).

Figure 2. Fall prediction scoring system to be implemented at the time of admission.

Predictive score for falls at admission		
Sex		
☐ Female		0 points
☐ Male		2 points
Age		
☐ 20 to < 65 years		0 points
☐ 65 to < 80 years		3 points
☐ ≥ to 80 years		4 points
BMI		
☐ ≤ 18.5 kg/m^2		2 points
☐ 18.5 to < 25 kg/m^2		1 points
☐ ≥ to 25 kg/m^2		0 points
Bedriddenness rank		
☐ Rank J		0 points
☐ Rank A		2 points
☐ Rank B		3 points
☐ Rank C		1 points
Department		
☐ Internal medicine		1 points
☐ Department of surgery		0 points
☐ Emergency department		2 points
History of falls with in 1 years		
☐ Absence		0 points
☐ Presence		2 points
Sum of the points		____points
Risk of fall	0-4	Low risk
	5-9	Moderate risk
	10-15	High risk

XSL•FO
RenderX

Figure 3. Fall risk classification-specific cumulative fall incidence. Cumulative fall incidences classified by fall risk scoring are shown for the training dataset (**A**) and test dataset (**B**).



Discussion

Principal Results

This study was a retrospective cohort investigation conducted at an institution specialized in acute inpatient care, aimed at identifying the risk factors for falls using the time from admission to fall as the outcome variable. Fall risk factors included age, sex, BMI, BR, emergency department, and history of falls within 1 year. Specifically, the study found that older patients (aged 80 years and older) had a higher risk of falls, with men being more at risk than women. Patients with a BMI of <18.5 and those admitted through the emergency department had an increased risk. In addition, those with a history of falls within the past year were particularly vulnerable. We constructed a new predictive scoring system for falls by weighting scores based on each HR according to the results of multivariable analysis and using statistical methods to classify fall risk groups into 3 categories (low risk: 0-4 points, moderate risk: 5-9 points, and high risk: 10-15 points). The present fall prediction scoring system could facilitate preventive interventions for high-risk patients, potentially reducing the likelihood of falls among the most susceptible patient populations and improving patient safety and care in the hospital environment.

Comparison With Prior Work

In terms of age-specific fall incidence, it was evident that a higher proportion of falls occurred among people aged 65 years and older. This result is in alignment with previous reports [6,16,20,22,27,40] identifying advanced age as a predictive factor for falls. Consistent with past studies [5,16,20,22,40], a history of falls was determined to be a predictive factor.

Previous reports on predictors of falls have shown that sex can be a predictor for both men and women. In this study, being male was identified as a risk factor. In addition, men have been found to experience multiple falls more frequently [27]. Past reports indicating that being male increases the risk of falls [20,28] have focused on hospitalized patients whereas those suggesting an increased risk for women [41,42] have focused on community-dwelling individuals. These differences may be owing to the different characteristics of the study populations, that is, relatively healthy community residents and patients in health care facilities. This may be related to the fact that hospitalized patients tend to have reduced physical activity, which increases the risk of falls. However, the relationship between sex differences and fall risk factors in the hospitalized population remains unclear, and further research is needed to elucidate these aspects.

A systematic review targeting community-dwelling older adults showed that a low BMI (<17 kg/m^2) is associated with a greater risk of falls, when using 23.5 kg/m^2 as the baseline [34]. In addition, some reports indicate that both high (25 - 35 kg/m^2 or above) and low (below 18.5 kg/m^2) BMI values are associated with increased fall risks [31-33,35]. These findings suggest that extreme BMI values influence the risk of falling. However, in our study, we identified low BMI (below 18.5 kg/m^2) and normal BMI (18.5 - 25 kg/m^2) as risk factors for falls, using high BMI (25 kg/m^2 or above) as the reference category. This divergence in results might be attributed to the fact that in our Japanese study population, few patients met the international high BMI criterion (>30 kg/m^2). In the study, because only 5.7% (6/741) of people in this training category fell, BMI >30 kg/m^2 was not able to be included as the category for a categorical variable BMI.

The BR degree, which is strongly correlated with ADL, was identified as a risk factor for inhospital falls. This finding aligns with several fall prediction models [17,18,20], demonstrating the significant role of a decline in ADL, as it is known that exercise therapy is effective in fall prevention [6,7]. This underscores that impaired ADL is a crucial factor in determining the outcomes of patients who experience falls. In addition,

XSL•FO
RenderX

because the overall assessment of ADL is predominantly based on mobility, a substantial correlation [43] between ADL and BR has been observed. In our study, of the 15 variables that showed a correlation with BR, 12 were related to components of ADL.

Being admitted to the emergency department (as an inpatient department) was identified as a risk factor for inhospital falls. The emergency department inpatient population comprises patients who have been transported to the emergency department or who otherwise came voluntarily to the emergency department. This population typically has more severe illness, which may explain why emergency department admission is a risk factor for inhospital falls. Moreover, a history of falls within the past 1 year was identified as a risk factor for inhospital falls. This result was similar to previous reports that identified a fall history as a risk factor for subsequent falls [17,18,20].

Falls are internationally recognized as a serious health issue [3,30], and various efforts to prevent falls are undertaken worldwide. Particularly in Japan, where rapid aging is prevalent, fall prevention has become an increasingly critical issue. Education for patients and staff is considered a fundamental approach to addressing this problem [8,10-12,44], and the introduction and enhancement of educational programs in Japanese hospitals are desirable [10]. To achieve this, it is essential to identify high-priority patients among inpatients and implement fall prevention measures as a high-risk approach. Multifactorial interventions to comprehensively assess and address multiple fall risk factors have proven effective [10]. However, some tools for evaluating fall risk have been criticized for their time-consuming nature and limited effectiveness [45-47], necessitating judicious selection and effective use. Recent research supports these multifactorial interventions and highlights the importance of tailored educational programs and effective risk assessment tools. Guidelines from the United Kingdom's National Institute for Health and Care Excellence offer specific approaches for fall prevention [48], which could serve as valuable references for fall prevention strategies in Japan. In the approach to fall prevention in Japanese hospitals, international insights and guidelines should be considered while also tailoring unique approaches according to facility characteristics, health care delivery systems, and patient backgrounds.

Statistical Validation and Clinical Application of Risk Categorization

In our research, we used the conditional inference tree method for statistical analysis to categorize patients into low-, middle-, and high-risk groups, as detailed in Table S4 in Multimedia Appendix 1. This methodological choice ensured that our risk stratifications were based on solid data analysis, avoiding arbitrary determinations. In addition, this scoring approach permits adjustment of cutoff values based on each health care facility's resources and circumstances, enhancing its applicability in diverse clinical settings. Our aim is not to change the patient's fall risk level based on resources but to adjust the intensity of preventive interventions according to available resources. The fall risk classification is based solely on clinical characteristics, while resource availability guides the prioritization and distribution of these interventions.

For patients identified as high risk for falls according to our predictive model, it is important to recognize that identifying these individuals is only the first step; providing effective interventions is a separate and critical challenge. Based on prior research demonstrating their effectiveness, we recommend several interventions tailored to implement fall prevention measures as a high priority. Here, we outline interventions such as increased monitoring, personalized environmental adjustments, nonslip footwear, assistive devices such as walkers or canes, one-on-one support, and immediate assistance. In addition, patient education and rehabilitation are crucial components. Educating both patients and health care providers about fall risks and preventive strategies, combined with physical therapy to enhance strength and balance, can significantly reduce the risk of falls [6-12]. However, it may be necessary to adjust these cutoffs based on sensitivity and specificity considerations to enhance the accuracy of patient risk identification, aligning more closely with practical prediction practices in health care.

The fall prediction tool we developed integrates seamlessly with a hospital's EHR system. Upon a patient's admission, it automatically retrieves EHR data and calculates a fall risk score based on various predictors. This integration allows all health care providers, including emergency and trauma physicians, to easily access the patient's fall risk assessment. Emergency or trauma physicians can use this tool to quickly identify patients at high risk for falls and implement appropriate interventions. This proactive approach can significantly reduce the incidence of falls and enhance patient safety in fast-paced and high-stress environments such as emergency departments and trauma centers.

Limitations

This study has several limitations. First, owing to the identification of inhospital fall predictors based on the characteristics of the facility and patient population in an acute care hospital, generalizing the findings to other all hospitals may be challenging. While the generalizability of our model is not fully guaranteed due to the lack of external validation, hospitals with similar facilities and patient characteristics might find the identified predictors and scoring system developed in this study applicable to their setting. Second, our study did not examine psychiatric symptoms, including delirium, as a potential predictive factor [49,50]. Because previous studies have shown that psychiatric symptoms can be a risk factor for falls [17,18,28], it can be necessary to reconstruct the prediction model by including additional predictors of falls, including psychiatric symptoms such as delirium, in future studies. Third, we did not examine the association between medications and falls. Previous studies have shown that taking sleeping pills [19,20,28-30,51] and the number of medications [42,52-54] are risk factors for falls. At the time of our study, limitations in our EHR system made it difficult to collect accurate medication data. This prevented us from including medication type and number as predictors. Since then, our EHR system has been upgraded, and we can now reliably obtain medication data. We

XSL•FO
RenderX

plan to incorporate these variables in future studies to enhance the accuracy of our fall prediction model. Fourth, we could not compare the results of our model with those of existing fall prediction models. However, we believe that our newly developed fall prediction model is highly useful in that it is easy to apply in many Japanese hospitals with acute care settings, similar to that in this study.

To further enhance the practical application of our fall prediction scoring system, we plan to integrate it into our hospital's EHRs to gather real-world evidence. This will allow us to evaluate its usability, accuracy, and impact on reducing inhospital falls. Future research will include multisite longitudinal studies to validate the tool's effectiveness across different health care settings.

Conclusions

We successfully identified predictors of falls within a patient population admitted to an acute care hospital and developed a novel prediction model in Japan. This model could serve as an effective tool to guide preventive interventions for both individual hospitalized patients and high-risk populations in many hospitals with acute care settings.

Acknowledgments

We thank Motoe Yumioka, Satomi Unno, Kazuyo Okushio, Sumiko Sano, Hideko Mochizuki, and Kazue Sano for their guidance in conducting this study. We thank Ms Etsuko Nagahashi and Mr Hiroki Ueda of the Information Support Office of Shizuoka General Hospital for assistance with data acquisition. We thank Katherine Thieltges and Analisa Avila, MPH, ELS, from Edanz for editing a draft of the manuscript. This study was supported by a grant-in-aid for Medical Research Promotion from the Shizuoka Prefectural Hospital Organization and a research instruction budget from the Shizuoka Graduate University of Public Health.

Data Availability

The datasets generated and analyzed during this study are not publicly available since the authors failed to obtain informed consent regarding the disclosure of raw patient data at the start of the study, which means the data in the manuscript, supporting information files, or public repositories cannot be freely accessed. However, the datasets are available from the Clinical Trial Management Office at Shizuoka General Hospital, located at 4-27-1 Kita-ando, Aoi-ku, Shizuoka, Japan 420-8527 (chiken-sougou@shizuoka-pho.jp), on reasonable request.

Authors' Contributions

CS and EN contributed to study concept and design and data analysis; CS and KH contributed to data acquisition; CS, EN, HS, NEK, MT, and KH did data interpretation; and CS contributed to first draft of the manuscript. All authors read drafts of the manuscript and approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables and figure.
[DOCX File, 151 KB - humanfactors_v12i1e58073_app1.docx]

References

1. Ministry of Health, Labour and Welfare. White paper on aging society [Article in Japanese]. Cabinet Office. 2023. URL: https://www8.cao.go.jp/kourei/whitepaper/w-2023/zenbun/pdf/1s1s_01.pdf [accessed 2024-12-13]
2. Status of inpatients by category of basic charge for hospitalization in general ward [Article in Japanese]. Ministry of Health, Labour and Welfare. 2017. URL: https://www.mhlw.go.jp/file/05-Shingikai-12404000-Hokenkyoku-Iryouka/0000155341.pdf [accessed 2024-12-13]
3. Dykes PC, Curtin-Bowen M, Lipsitz S, et al. Cost of inpatient falls and cost-benefit analysis of implementation of an evidence-based fall prevention program. JAMA Health Forum 2023 Jan 6;4(1):e225125. [doi: 10.1001/jamahealthforum.2022.5125] [Medline: 36662505]
4. Aihara H, Tago M, Oishi T, Katsuki NE, Yamashita SI. Visual impairment, partially dependent ADL and extremely old age could be predictors for severe fall injuries in acute care settings. Int J Gerontol 2018 Sep;12(3):175-179. [doi: 10.1016/j.ijge.2018.02.014]
5. Panel on Prevention of Falls in Older Persons, American Geriatrics Society and British Geriatrics Society. Summary of the updated American Geriatrics Society/British Geriatrics Society clinical practice guideline for prevention of falls in older persons. J Am Geriatr Soc 2011 Jan;59(1):148-157. [doi: 10.1111/j.1532-5415.2010.03234.x]
6. Cuevas-Trisan R. Balance problems and fall risks in the elderly. Clin Geriatr Med 2019 May;35(2):173-183. [doi: 10.1016/j.cger.2019.01.008] [Medline: 30929881]

XSL•FO
RenderX

7. Canning CG, Sherrington C, Lord SR, et al. Exercise for falls prevention in Parkinson disease: a randomized controlled trial. Neurology (ECronicon) 2015 Jan 20;84(3):304-312. [doi: 10.1212/WNL.0000000000001155] [Medline: 25552576]
8. Hill AM, Francis-Coad J, Haines TP, et al. "My independent streak may get in the way": how older adults respond to falls prevention education in hospital. BMJ Open 2016 Jul 26;6(7):e012363. [doi: 10.1136/bmjopen-2016-012363] [Medline: 27466244]
9. Heng H, Kiegaldie D, Slade SC, et al. Healthcare professional perspectives on barriers and enablers to falls prevention education: a qualitative study. PLoS One 2022;17(4):e0266797. [doi: 10.1371/journal.pone.0266797] [Medline: 35476840]
10. Morris ME, Webster K, Jones C, et al. Interventions to reduce falls in hospitals: a systematic review and meta-analysis. Age Ageing 2022 May 1;51(5):afac077. [doi: 10.1093/ageing/afac077] [Medline: 35524748]
11. Shaw L, Kiegaldie D, Morris ME. Educating health professionals to implement evidence-based falls screening in hospitals. Nurse Educ Today 2021 Jun;101:104874. [doi: 10.1016/j.nedt.2021.104874] [Medline: 33774527]
12. Heng H, Jazayeri D, Shaw L, Kiegaldie D, Hill AM, Morris ME. Hospital falls prevention with patient education: a scoping review. BMC Geriatr 2020 Apr 15;20(1):140. [doi: 10.1186/s12877-020-01515-w] [Medline: 32293298]
13. Kawazoe Y, Shimamoto K, Shibata D, Shinohara E, Kawaguchi H, Yamamoto T. Impact of a clinical text-based fall prediction model on preventing extended hospital stays for elderly inpatients: model development and performance evaluation. JMIR Med Inform 2022 Jul 27;10(7):e37913. [doi: 10.2196/37913] [Medline: 35896017]
14. Cho I, Boo EH, Chung E, Bates DW, Dykes P. Novel approach to inpatient fall risk prediction and its cross-site validation using time-variant data. J Med Internet Res 2019 Feb 19;21(2):e11505. [doi: 10.2196/11505] [Medline: 30777849]
15. Cho I, Jin IS, Park H, Dykes PC. Clinical impact of an analytic tool for predicting the fall risk in inpatients: controlled interrupted time series. JMIR Med Inform 2021 Nov 25;9(11):e26456. [doi: 10.2196/26456] [Medline: 34626168]
16. Jung H, Yoo S, Kim S, et al. Patient-level fall risk prediction using the observational medical outcomes partnership's common data model: pilot feasibility study. JMIR Med Inform 2022 Mar 11;10(3):e35104. [doi: 10.2196/35104] [Medline: 35275076]
17. Oliver D, Britton M, Seed P, Martin FC, Hopper AH. Development and evaluation of evidence based risk assessment tool (STRATIFY) to predict which elderly inpatients will fall: case-control and cohort studies. BMJ 1997 Oct 25;315(7115):1049-1053. [doi: 10.1136/bmj.315.7115.1049] [Medline: 9366729]
18. Morse JM, Morse RM, Tylko SJ. Development of a scale to identify the fall-prone patient. Can J Aging 1989;8(4):366-377. [doi: 10.1017/S0714980800008576]
19. Beasley B, Patatanian E. Development and implementation of a pharmacy fall prevention program. Hosp Pharm 2009 Dec;44(12):1095-1102. [doi: 10.1310/hpj4412-1095]
20. Tago M, Katsuki NE, Oda Y, Nakatani E, Sugioka T, Yamashita SI. New predictive models for falls among inpatients using public ADL scale in Japan: a retrospective observational study of 7,858 patients in acute care setting. PLoS One 2020;15(7):e0236130. [doi: 10.1371/journal.pone.0236130] [Medline: 32673366]
21. Hirata R, Katsuki NE, Yaita S, et al. Validation of the Saga Fall Injury Risk Model. Int J Med Sci 2024;21(8):1378-1384. [doi: 10.7150/ijms.92837] [Medline: 38903917]
22. Yaita S, Tago M, Katsuki NE, et al. A simple and accurate model for predicting fall injuries in hospitalized patients: insights from a retrospective observational study in Japan. Med Sci Monit 2023 Aug 14;29:e941252. [doi: 10.12659/MSM.941252] [Medline: 37574766]
23. Nakatani H, Nakao M, Uchiyama H, Toyoshiba H, Ochiai C. Predicting inpatient falls using natural language processing of nursing records obtained from Japanese electronic medical records: case-control study. JMIR Med Inform 2020 Apr 22;8(4):e16970. [doi: 10.2196/16970] [Medline: 32319959]
24. Soyano A, Suzuki M, Harada A, et al. JSFP Research Report〉 fall risk assessment tool usage and issues: results of a fall risk assessment tool usage and issues: results of a. Jpn J Fall Prev 2018;5:41-49. [doi: 10.11335/tentouyobou.5.1_41]
25. Higaonna M. The predictive validity of a modified Japanese Nursing Association fall risk assessment tool: a retrospective cohort study. Int J Nurs Stud 2015 Sep;52(9):1484-1494. [doi: 10.1016/j.ijnurstu.2015.05.009] [Medline: 26138481]
26. Okuma Y. Freezing of gait and falls in Parkinson's disease. J Parkinsons Dis 2014;4(2):255-260. [doi: 10.3233/JPD-130282] [Medline: 24577502]
27. Hodgson G, Pace A, Carfagnini Q, et al. Risky business: factors that increase risk of falls among older adult in-patients. Gerontol Geriatr Med 2023;9:23337214231189930. [doi: 10.1177/23337214231189930] [Medline: 37533770]
28. Hendrich AL, Bender PS, Nyhuis A. Validation of the Hendrich II Fall Risk Model: a large concurrent case/control study of hospitalized patients. Appl Nurs Res 2003 Feb;16(1):9-21. [doi: 10.1053/apnr.2003.YAPNR2] [Medline: 12624858]
29. Obayashi K, Araki T, Nakamura K, et al. Risk of falling and hypnotic drugs: retrospective study of inpatients. Drugs R D 2013 Jun;13(2):159-164. [doi: 10.1007/s40268-013-0019-3] [Medline: 23760758]
30. Noh HM, Song HJ, Park YS, Han J, Roh YK. Fall predictors beyond fall risk assessment tool items for acute hospitalized older adults: a matched case-control study. Sci Rep 2021 Jan 15;11(1):1503. [doi: 10.1038/s41598-021-81034-9] [Medline: 33452370]
31. Nakamura K, Kitamura K, Watanabe Y, et al. Body mass index and risk of recurrent falls in community-dwelling Japanese aged 40-74 years: the Murakami cohort study. Geriatr Gerontol Int 2021 Jun;21(6):498-505. [doi: 10.1111/ggi.14167] [Medline: 33881216]

XSL•FO
RenderX

32. Ylitalo KR, Karvonen-Gutierrez CA. Body mass index, falls, and injurious falls among U.S. adults: findings from the 2014 Behavioral Risk Factor Surveillance System. Prev Med 2016 Oct;91:217-223. [doi: 10.1016/j.ypmed.2016.08.044] [Medline: 27575319]
33. Yi SW, Kim YM, Won YJ, Kim SK, Kim SH. Association between body mass index and the risk of falls: a nationwide population-based study. Osteoporos Int 2021 Jun;32(6):1071-1078. [doi: 10.1007/s00198-020-05725-1] [Medline: 33411009]
34. Trevisan C, Crippa A, Ek S, et al. Nutritional status, body mass index, and the risk of falls in community-dwelling older adults: a systematic review and meta-analysis. J Am Med Dir Assoc 2019 May;20(5):569-582. [doi: 10.1016/j.jamda.2018.10.027] [Medline: 30554987]
35. Ogliari G, Ryg J, Andersen-Ranberg K, Scheel-Hincke LL, Masud T. Association between body mass index and falls in community-dwelling men and women: a prospective, multinational study in the Survey of Health, Ageing and Retirement in Europe (SHARE). Eur Geriatr Med 2021 Aug;12(4):837-849. [doi: 10.1007/s41999-021-00485-5] [Medline: 33817765]
36. Tago M, Katsuki NE, Yaita S, et al. High inter-rater reliability of Japanese bedriddenness ranks and cognitive function scores: a hospital-based prospective observational study. BMC Geriatr 2021 Mar 9;21(1):168. [doi: 10.1186/s12877-021-02108-x] [Medline: 33750305]
37. Ozawa MN, Nakayama S. Long-term care insurance in Japan. J Aging Soc Policy 2005;17(3):61-84. [doi: 10.1300/J031v17n03_04] [Medline: 16219594]
38. Council for Medical Safety Management of National University Hospitals influence classification [Article in Japanese]. Japan National University Hospital Alliance on Patient Safety. URL: https://square.umin.ac.jp/anzenhc/information/pdf/23_business_text2022.pdf [accessed 2024-12-13]
39. Kanda Y. Investigation of the freely available easy-to-use software "EZR" for medical statistics. Bone Marrow Transplant 2013 Mar;48(3):452-458. [doi: 10.1038/bmt.2012.244] [Medline: 23208313]
40. Hirata R, Tago M, Katsuki NE, et al. History of falls and bedriddenness ranks are useful predictive factors for in-hospital falls: a single-center retrospective observational study using the Saga Fall Risk Model. Int J Gen Med 2022;15:8121-8131. [doi: 10.2147/IJGM.S385168] [Medline: 36389017]
41. Masumoto T, Yamada Y, Yamada M, et al. Fall risk factors and sex differences among community-dwelling elderly individuals in Japan. A Kameoka study. Nihon Koshu Eisei Zasshi 2015;62(8):390-401. [doi: 10.11236/jph.62.8_390] [Medline: 26511610]
42. Bally ELS, Ye L, van Grieken A, et al. Factors associated with falls among hospitalized and community-dwelling older adults: the APPCARE study. Front Public Health 2023;11:1180914. [doi: 10.3389/fpubh.2023.1180914] [Medline: 37457268]
43. Tsuboi A, Murakami T, Kurumadani H, et al. Relationship between criteria for evaluating the degree of independence of disabled elderly persons in performing activities of daily living. J Jpn Occup Ther Assoc 2002;21:255-262. [doi: 10.11477/mf.6003900624]
44. Haines TP, Hill AM, Hill KD, et al. Patient education to prevent falls among older hospital inpatients: a randomized controlled trial. Arch Intern Med 2011 Mar 28;171(6):516-524. [doi: 10.1001/archinternmed.2010.444] [Medline: 21098343]
45. Montero-Odasso M, van der Velde N, Martin FC, et al. World guidelines for falls prevention and management for older adults: a global initiative. Age Ageing 2022 Sep 2;51(9):afac205. [doi: 10.1093/ageing/afac205] [Medline: 36178003]
46. Jellett J, Williams C, Clayton D, Plummer V, Haines T. Falls risk score removal does not impact inpatient falls: a stepped-wedge, cluster-randomised trial. J Clin Nurs 2020 Dec;29(23-24):4505-4513. [doi: 10.1111/jocn.15471] [Medline: 32945020]
47. Morris ME, Haines T, Hill AM, et al. Divesting from a scored hospital Fall Risk Assessment Tool (FRAT): a cluster randomized non-inferiority trial. J Am Geriatr Soc 2021 Sep;69(9):2598-2604. [doi: 10.1111/jgs.17125] [Medline: 33834490]
48. Swift CG, Iliffe S. Assessment and prevention of falls in older people—concise guidance. Clin Med (Northfield) 2014 Dec;14(6):658-662. [doi: 10.7861/clinmedicine.14-6-658]
49. Wong CL, Holroyd-Leduc J, Simel DL, Straus SE. Does this patient have delirium?: value of bedside instruments. JAMA 2010 Aug 18;304(7):779-786. [doi: 10.1001/jama.2010.1182] [Medline: 20716741]
50. Inouye SK, Westendorp RGJ, Saczynski JS. Delirium in elderly people. Lancet 2014 Mar;383(9920):911-922. [doi: 10.1016/S0140-6736(13)60688-1]
51. Herzig SJ, Rothberg MB, Moss CR, et al. Risk of in-hospital falls among medications commonly used for insomnia in hospitalized patients. Sleep 2021 Sep 13;44(9):zsab064. [doi: 10.1093/sleep/zsab064] [Medline: 33710329]
52. Kojima T, Akishita M, Nakamura T, et al. Association of polypharmacy with fall risk among geriatric outpatients. Geriatr Gerontol Int 2011 Oct;11(4):438-444. [doi: 10.1111/j.1447-0594.2011.00703.x] [Medline: 21545384]
53. Zaninotto P, Huang YT, Di Gessa G, Abell J, Lassale C, Steptoe A. Polypharmacy is a risk factor for hospital admission due to a fall: evidence from the English Longitudinal Study of Ageing. BMC Public Health 2020 Nov 26;20(1):1804. [doi: 10.1186/s12889-020-09920-x] [Medline: 33243195]
54. Dhalwani NN, Fahami R, Sathanapally H, Seidu S, Davies MJ, Khunti K. Association between polypharmacy and falls in older adults: a longitudinal study from England. BMJ Open 2017 Oct 16;7(10):e016358. [doi: 10.1136/bmjopen-2017-016358] [Medline: 29042378]

XSL•FO
RenderX

Abbreviations

ADL: activities of daily living
BR: bedriddenness rank
EHR: electronic health record
HR: hazard ratio

Edited by P Santana-Mancilla; submitted 13.04.24; peer-reviewed by S Haag, C Udeh, L Lehtonen; revised version received 24.09.24; accepted 30.09.24; published 13.01.25.

Please cite as:
Saito C, Nakatani E, Sasaki H, E Katsuki N, Tago M, Harada K
Predictive Factors and the Predictive Scoring System for Falls in Acute Care Inpatients: Retrospective Cohort Study
JMIR Hum Factors 2025;12:e58073
URL: https://humanfactors.jmir.org/2025/1/e58073
doi:10.2196/58073

© Chihiro Saito, Eiji Nakatani, Hatoko Sasaki, Naoko E Katsuki, Masaki Tago, Kiyoshi Harada. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 13.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

An Explainable AI Application (AF'fective) to Support Monitoring of Patients With Atrial Fibrillation After Catheter Ablation: Qualitative Focus Group, Design Session, and Interview Study

Wan Jou She[1], PhD; Panote Siriaraya[1], PhD; Hibiki Iwakoshi[2,3], MD, PhD; Noriaki Kuwahara[1,4], PhD; Keitaro Senoo[2,3], MD, PhD

[1]Faculty of Information and Human Sciences, Kyoto Institute of Technology, Kyoto, Japan

[2]Department of Cardiovascular Medicine, Graduate School of Medical Science, Kyoto Prefectural University of Medicine, Kyoto, Japan

[3]Department of Cardiac Arrhythmia Research and Innovation, Graduate School of Medical Science, Kyoto Prefectural University of Medicine, Kyoto, Japan

[4]Department of Advanced Fibro-Science, Kyoto Institute of Technology, Kyoto, Japan

Corresponding Author:
Keitaro Senoo, MD, PhD
Department of Cardiac Arrhythmia Research and Innovation
Graduate School of Medical Science
Kyoto Prefectural University of Medicine
465 Kajii-cho Hirokoji, Kawaramachi-dori
Kamigyo-ku Kyoto
Kyoto, 602-8566
Japan
Phone: 81 75 251 5511
Fax: 81 75 251 5514
Email: k-senoo@koto.kpu-m.ac.jp

Abstract

Background: The opaque nature of artificial intelligence (AI) algorithms has led to distrust in medical contexts, particularly in the treatment and monitoring of atrial fibrillation. Although previous studies in explainable AI have demonstrated potential to address this issue, they often focus solely on electrocardiography graphs and lack real-world field insights.

Objective: We addressed this gap by incorporating standardized clinical interpretation of electrocardiography graphs into the system and collaborating with cardiologists to co-design and evaluate this approach using real-world patient cases and data.

Methods: We conducted a 3-stage iterative design process with 23 cardiologists to co-design, evaluate, and pilot an explainable AI application. In the first stage, we identified 4 physician personas and 7 explainability strategies, which were reviewed in the second stage. A total of 4 strategies were deemed highly effective and feasible for pilot deployment. On the basis of these strategies, we developed a progressive web application and tested it with cardiologists in the third stage.

Results: The final progressive web application prototype received above-average user experience evaluations and effectively motivated physicians to adopt it owing to its ease of use, reliable information, and explainable functionality. In addition, we gathered in-depth field insights from cardiologists who used the system in clinical contexts.

Conclusions: Our study identified effective explainability strategies, emphasized the importance of curating actionable features and setting accurate expectations, and suggested that many of these insights could apply to other disease care contexts, paving the way for future real-world clinical evaluations.

(JMIR Hum Factors 2025;12:e65923) doi:10.2196/65923

KEYWORDS

atrial fibrillation; explainable artificial intelligence; explainable AI; user-centered design; prevention; postablation monitoring

XSL•FO
RenderX

Introduction

Background

While the role of artificial intelligence (AI) in health care is growing, concerns about its reliability and transparency remain largely unaddressed [1]. The opaque nature of AI algorithms challenges trust, particularly in cardiology, where the lack of interpretability in deep learning models limits their adoption and utility [2,3]. This has led to increased interest in developing explainable AI models that offer accurate predictions along with clear explanations of the rationale behind the predictive results [4].

Our study explored the advantages and challenges of using AI models in cardiovascular treatment, focusing on designing an explainable AI system to aid posttreatment care for patients with atrial fibrillation (AF) after catheter ablation. There were >3 million new cases of AF, a common arrhythmic disorder [5], in 2017, with numbers expected to rise due to aging [6]. Although often undetected, AF is a significant risk factor for severe conditions such as stroke or heart failure [7] and can greatly impact patients' quality of life [5,8]. While catheter ablation is effective, AF can recur after ablation, requiring careful monitoring of risk factors [9].

Several machine learning models, from screening algorithms [10,11] to those predicting AF recurrence after ablation [12], have been developed, but integrating them into health care remains challenging. Physicians often require detailed justifications due to prediction uncertainties. While explainable AI algorithms [5,13] could assist, a more holistic approach is needed in posttreatment scenarios to effectively integrate additional information, which remains poorly understood in real-world care contexts (discussed in the next section).

To better integrate a machine learning model into the ongoing decision-making ecosystem and identify unmet clinical needs beyond performance optimization, we used a user-centered, iterative design approach to develop AF'fective, an explainable AI system for posttreatment monitoring of patients with AF. By conducting semistructured interviews, design sessions, and focus groups with 23 cardiologists, we aimed to identify key principles for the effective design and use of explainable AI in real-life cardiac treatment and management. In this section, we first review the development of explainable AI in relevant fields in the second subsection and summarize our first 2 study stages in the third subsection. We then detail and discuss our final-stage findings and conclude with reflections on the study's implications and the generalizability of our insights.

High Accuracy but (Still) Rejected: Machine Learning Models in Cardiology

Cardiology is one field of medicine that has seen extensive use of AI to aid in medical practice. Various machine learning models have been developed and used to assist in a wide range of tasks, from helping predict the prognosis or readmission after heart failure [14,15] to detecting various cardiovascular diseases from medical images [16]. In the area of arrhythmia, previous studies have focused particularly on two key areas: (1) prevention, which involves assisting in the early detection of various arrhythmic conditions; and (2) monitoring, which involves supporting the management of such conditions after treatment [17]. In the preventive area, researchers have shown how machine learning models could be developed and used to detect irregular conditions using electrocardiography (ECG) signals [18] or even through photoplethysmography sensors from commercially available smartwatches [19,20]. In most cases, such models were able to diagnose these irregular conditions with a high degree of accuracy (97%) [20]. However, the potential for a high rate of false positives when implemented through commercial smartwatches has triggered concerns among researchers, and as a result, they have cautioned against using such systems for population-wide screening and emphasized maintaining the 12-lead ECG as the gold standard [21]. Given the prevalence of AF, several of the developed models in this domain have chosen to focus specifically on this condition, with various deep learning ECG models being developed to help screen people for AF, stratify patients based on their risk level, and even predict the chances of the condition occurring in the future (see the study by Sehrawat et al [21] for a comprehensive review).

In the monitoring area, to support posttreatment care, previous studies have developed models capable of predicting patient mortality or echocardiographic response after procedures such as cardiac resynchronization therapy [22]. Other studies have even shown how the risk of recurrence for conditions such as AF could be predicted using patient demographics and 3D computed tomography images of the left atrium [23]. Similarly, several studies have demonstrated how the recurrence of AF or 30-day hospital readmission after catheter ablation can be predicted using various deep learning and non–deep learning strategies (eg, convolutional neural network [CNN] and Extreme Gradient Boosting) [12,13,24]. In other cases, deep learning algorithms have also been developed to detect an AF episode 4.5 minutes before onset, thus enabling prompt interventions to prevent their occurrence [25].

Overall, while the findings of these studies all appear to demonstrate significant potential for implementing predictive models to support AF treatment, we have encountered substantial pushback from clinical partners when attempting to implement such a model in clinical practice. Despite the promising performance of the model in the study by Nishimura et al [26] (area under the curve of 0.72 with 83% sensitivity and 58% specificity), there was still strong resistance from physicians regarding trusting the predicted outcomes. Such hesitation and resistance to adopting machine learning models among clinical staff have also been frequently identified in previous studies [1,27]. Given the delicate and often critical nature of treatment in this field, it is not surprising that the lack of interpretability undermines trust in these models, preventing clinicians from using them in medical treatment (see the study by Petch et al [2]). Moreover, the opaque process raises concerns regarding how mistakes could be rectified, not to mention that it may cause potentially disruptive emotions when discussing such results with patients. Therefore, at present, the high accuracy, accessibility, and (possibly) low cost of machine learning models often do not lead to their adoption in medical practice.

XSL•FO
RenderX

Explainable AI in AF

Recently, explainable AI has gained attention as a potential approach to address these issues of transparency and interpretability [28-30]. In the context of AF treatment, ECG data are generally used as a main data source for screening irregular rhythm and predicting the risk of AF recurrence, and as such, explainable AI techniques such as Shapley Additive Explanations (SHAP) have been used with the ECG data to identify features contributing to the predictive outcomes [31]. In previous studies, researchers have proposed an explainable deep learning model that not only detects AF but also describes the reasons behind their decisions and visualizes key regions within the ECG signal identified as important predictors using techniques such as Gradient-Weighted Class Activation Mapping (Grad-CAM) or a deep visual attention [29,32,33]. Building upon earlier models, researchers such as Raza et al [34] incorporated federated learning techniques alongside explainable AI to preserve patient ECG data privacy. In more recent studies, explainable models have even been created that are capable of identifying patients at high risk after catheter ablation and highlighting the features (eg, type of AF, age, and left atrial diameter) that the model used to decide on the risk level for each patient [30]. The use of "explainable" models in this manner has been argued to have the added benefit of allowing clinicians to better understand the relationship between each contributing factor and the predicted outcome of an individual case, helping clinicians justify their decisions and treatments more effectively [35].

While explainable AI techniques are believed to enhance the interpretability of algorithms, incorporating them to truly improve patient outcomes in real-life medical contexts remains challenging [36]. Understanding of the strengths and limitations of these approaches in different use cases and how human and AI-based diagnoses can complement each other is still lacking [36]. Deploying these models effectively requires a thorough examination of the patient journey and adapting the explainable strategy accordingly. Recent studies [37,38] have begun to address these issues by examining the use of explainable AI in more realistic settings with various stakeholders to ensure safe adoption in clinical practice.

Overall, although the integration of explainable AI into medical contexts holds great promise, it requires a concerted effort to address technical, ethical, and practical challenges. Insights from near-live and real-life case studies can help develop best practices and trustworthy methods for implementing such systems in clinical settings, thereby enabling users and professionals to fully benefit from AI technology.

Early Study Stages and Findings in a Nutshell

In this section, we summarize and report the in-depth insights from the 2 early stages of interviews (partially reported as a poster in the study by She et al [39]) and group design sessions to conceptualize and pilot-test our design strategies and prototype. We chose to include some key findings from these early stages to enhance readers' understanding of our design decisions in developing the final prototype for evaluation.

Stage 1

Overview

In the first stage, we conducted semistructured interviews with physicians to gain insights into their experiences and best practices for posttreatment care of patients with AF. Physicians were asked to share their positive and negative experiences, communication strategies, and effective procedures for monitoring posttreatment patients with AF. To sensitize and prepare participants, we provided a 1-week workbook with reflection exercises sent 1 to 2 weeks before the interview. Each interview lasted approximately 45 minutes and was fully recorded with the consent of the physicians.

Participants

We invited 8 physicians from different specialties, such as emergency medicine, internal medicine, and cardiology, to join us for interviews at Kyoto Prefectural University of Medicine in Japan. We made sure to include a mix of experience levels and backgrounds in cardiology to obtain insights that could be more broadly generalized to other medical contexts.

Measurements and Study Apparatus

A week before the interview, each physician received a toolkit to reflect on their clinical practice. The toolkit included a workbook with 7 daily assignments on their experiences and patient interactions as well as stickers featuring emotional words and contextual images to help express their thoughts and feelings.

Study Process

We asked physicians to reflect on their process for diagnosing patients and explaining treatment plans, especially in difficult situations such as delivering bad news. They were also asked to share strategies for handling uncooperative patients or those who rejected treatment. Physicians were encouraged to use the 2 sticker sets to respond to workbook questions or use them as interview resources. During the interview, physicians were guided through the workbook to discuss their specific approaches, experiences, motivations, stressors, and methods for negotiating treatment plans.

Analysis

In total, 2 researchers from the human-computer interaction (HCI) field and a cardiologist analyzed the interview data, with all recordings fully transcribed. The HCI researchers developed design strategies through thematic analysis, whereas the cardiologist reviewed and selected strategies for further evaluation in stage 2.

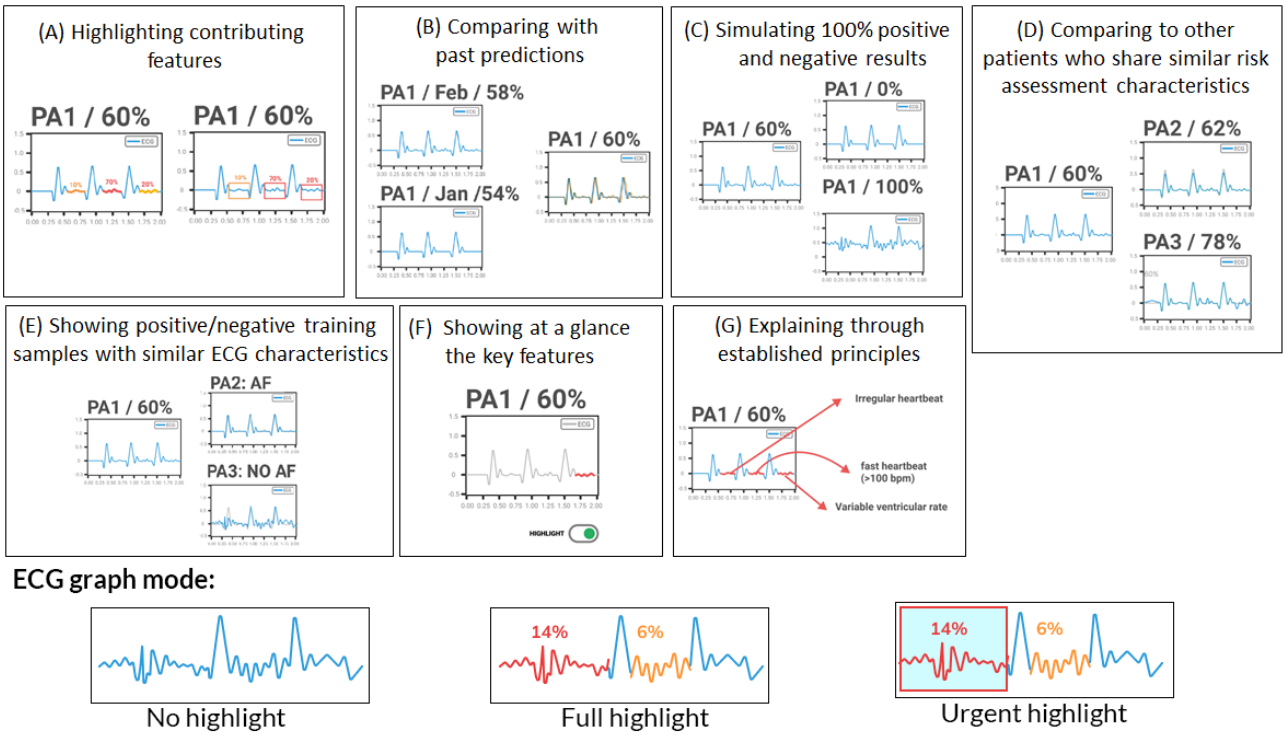
Results

Our initial interviews with 8 physicians produced 2 key outputs: 4 physician personas (*teamworker*, *minimum effort professional*, *the veteran*, and *self-actualizer*) and 7 design strategies for explainable prediction. Each persona included pseudodemographics and progress bars depicting their personalities across 4 parameters: *passionate/realistic*, *rookie/veteran*, *work-oriented/life-oriented*, and *individualism/collectivism*.

XSL•FO
RenderX

The research team discovered that participants had common concerns about self-efficacy in patient treatment, feeling validated when treatments succeeded but disempowered when seen as incompetent. This insight led to the development of 7 design strategies for explainable prediction. Participants emphasized the need to justify decisions by clearly explaining issues such as ECG results to patients. We reviewed and adapted explainable AI strategies from previous studies [40] for cardiology. The seven strategies were (1) highlighting key ECG features (strategy A), (2) comparing with historical predictions (strategy B), (3) simulating extreme ECG outcomes (strategy C), (4) comparing with similar patients (strategy D), (5) showing similar training examples (strategy E), (6) providing an at-a-glance mode (strategy F), and (7) explaining through established ECG principles (strategy G; Figure 1 [39]).

Figure 1. The 7 explainable artificial intelligence strategies (A-G) that were examined in the initial 2 stages of the study, expanded from the previous publication [39]. AF: atrial fibrillation; bpm: beats per minute; ECG: electrocardiography; PA1: patient 1; PA2: patient 2; PA3: patient 3.



Stage 2

Overview

In the second stage, we evaluated the feasibility of the 7 design strategies from the first stage, focusing on those that could be implemented using explainable AI. Physicians first described their usual process for assessing recurrence risk and discussed the potential role of predictive models and explainable AI. We then introduced the explainable AI strategies and gathered detailed feedback on which would most effectively clarify predicted outcomes for other physicians and patients.

Study Process and Findings to Guide the Prototype Development

We conducted 2 group design sessions with 8 cardiologists to critically review and assess our design strategies. Initially, 5 cardiologists were asked to evaluate the advantages and disadvantages of each strategy, describe scenarios in which they would be useful, and reflect on their concerns regarding implementing such explainable features.

From the 7 proposed strategies, we selected 4 (strategies A, B, D, and G) for further development into a web-based prototype using a high-fidelity explainable CNN model with real patient data. In total, 3 cardiologists reviewed the prototype using the concurrent think-aloud approach [41] and provided feedback. They also discussed concerns about using AI during posttreatment care, specifically on whether the explainable features enhanced their trust in the AI system and how the system might add or reduce value in their practice.

Participants

A total of 5 cardiologists with experience treating AF were recruited from Kyoto Prefectural University of Medicine to assess the explainable AI strategies. In addition, 3 cardiologists from the same institution were brought on to evaluate a web-based prototype.

Results and Study Apparatus Development

The 4 effective design strategies identified were highlighting contributing features (strategy A), comparing with past predictors (strategy B), comparing with other patients (strategy D), and explaining through established principles (strategy G). Participants' choices were guided by 3 principles: discernibility (highlighting relevant noteworthy parts), comparability (intra- and interpersonal comparisons), and evidence-based approaches (derived from diagnostic examples or clinical literature).

XSL•FO
RenderX

Strategies to avoid included non–evidence-based methods and those showing only partial graphs.

In general, physicians were highly receptive to the AF'fective prototype, particularly its explainable features, which they found useful for evaluating treatment effectiveness in posttreatment monitoring. On the basis of their feedback, we advanced to web-based prototyping.

The web-based prototype was developed using Figma (Figma, Inc), a popular tool among user experience (UX) designers for creating high-fidelity and web-based prototypes at minimal cost during the development process. Our prototype comprised 2 main screens: a patient overview screen and an individual health record dashboard (to avoid repetition of similar images, we share the screenshots from the final system in this paper instead; Figures 2 and 3).

The overview screen displays patients' demographic details and predicted recurrence risks, allowing cardiologists to sort by various demographic and risk factors. The health record dashboard displays (comparable and evidence-based) health information gathered from a patient's previous visits. Our predictive model will automatically extract and highlight keywords or risk factors based on the patient's visits and health data. In addition, the model differentiates the importance of the risk factors and offers 3 tiers of ECG result presentation based on the level of urgency for physicians: no highlight, full highlight, and critical highlight only.

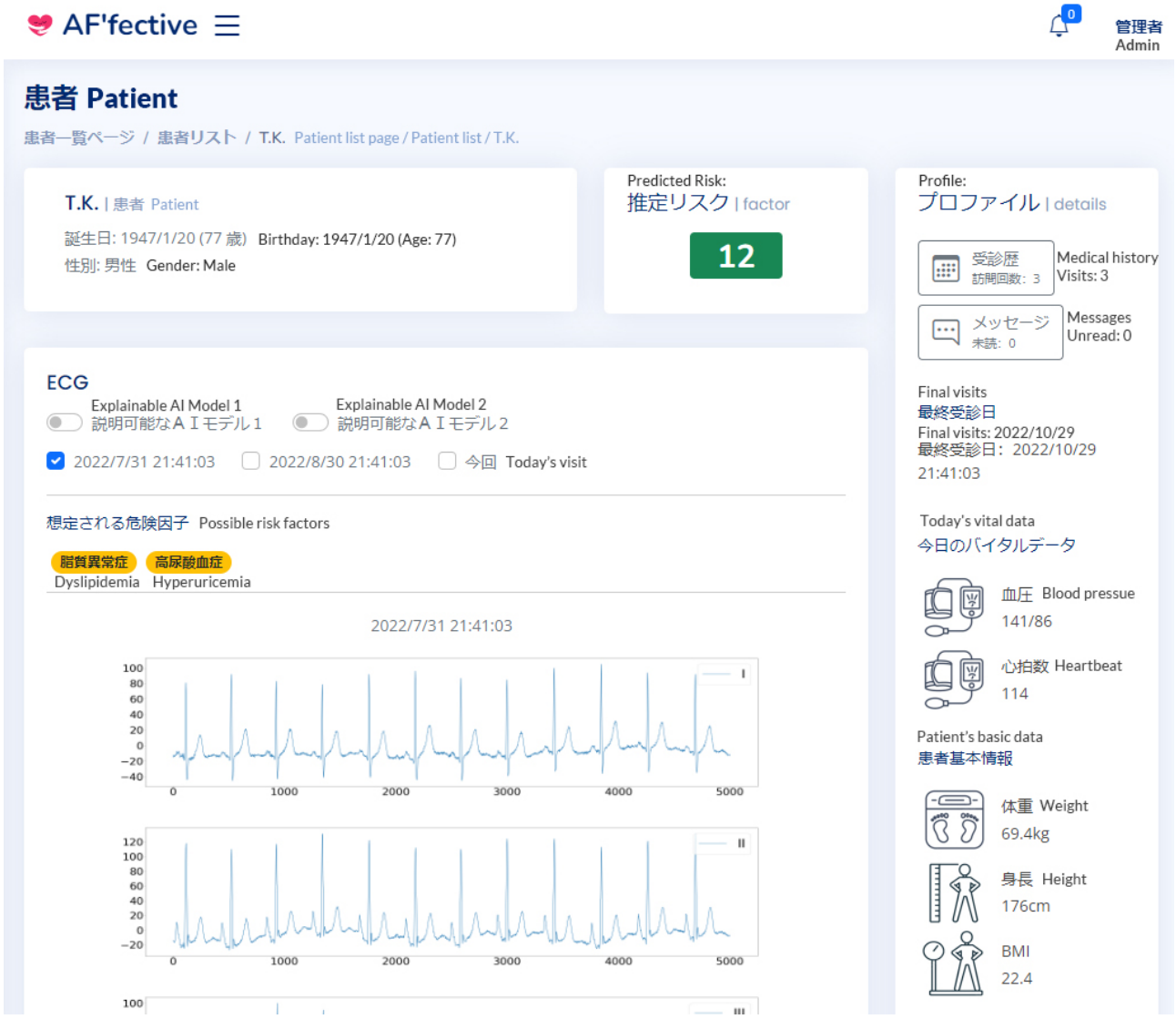
In critical highlight mode, only urgent health information is shown, whereas the full highlight mode displays all details, including less urgent AI highlights. Validated risk factors are listed as diagnostic keywords in the right panel, helping physicians quickly identify the main contributors to the predictive outcome.

Figure 2. The patient overview screen.



XSL•FO
RenderX

Figure 3. Screenshot of the patient page, including extra risk factors. ECG: electrocardiography.



Methods

Overview

Our research followed a 3-phase approach collaborating with stakeholders familiar with AF management. In the first phase, we conducted semistructured interviews with 8 cardiologists to explore potential explainable AI strategies for postablation care. From these insights, we developed 7 design concepts, which were then critically assessed and refined by another 8 cardiologists in the second phase. This process led to the selection of 4 explainable strategies. In this section (the final phase of the study), we built a functional prototype incorporating real patient data and invited a third group of 7 cardiologists to evaluate its effectiveness and feasibility in clinical settings.

Ethical Considerations

This study received ethics approval from the relevant institutional review boards (IRBs) in Japan, including the Kyoto Institute of Technology IRB (2021-08) and the Kyoto Prefectural University of Medicine IRB (ERB-C-2014). Informed consent was obtained from all participating physicians, who volunteered to take part in the study without receiving any compensation. To ensure the privacy and confidentiality of patient health data, all personally identifiable information, including patients' names, was removed and replaced with unique codes. This anonymization process ensured that no data could be traced back to individual patients.

Stage 3

The web-based prototype from the previous stage was developed into a fully functional progressive web application (PWA) prototype, which was then tested and evaluated by cardiologists.

Patient Data

A total of 11 postablation patients consented to provide data for the prototype. Protected health information, such as names and birthdays, was masked or replaced due to IRB requests, but other health data (eg, ECG, body measurements, blood pressure, and visit history) were used to closely resemble the original patient health data for a realistic prototype.

Participants

In total, 7 cardiologists involved in the treatment of AF were recruited from the Kyoto Prefectural University of Medicine in

XSL•FO
RenderX

Japan. All of them had experience diagnosing patients with AF and performing catheter ablation.

Measurements and Study Apparatus

In total, 3 widely adapted UX-related questionnaires were implemented in our study: the User Experience Questionnaire (UEQ), technology acceptance model (TAM), and Mobile App Rating Scale (MARS). The rationales for including these questionnaires are explained in the following sections.

UEQ Inclusion Rationale

The UEQ, developed by Laugwitz et al [42] in 2008, consists of 26 items across 6 dimensions: attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty. The scale ranges from −3 (*most negative*) to +3 (*most positive*). Scores closer to 0 are considered "neutral." A data analysis tool is provided for interpreting the results, with benchmarks in 5 tiers: excellent, good, above average, below average, and bad [43]. Our goal during prototyping was to achieve above-average ratings. The English version was translated into Japanese.

TAM Inclusion Rationale

Previous research suggests that a user's intention to use is the primary predictor of actual system use [44]. To ensure that our PWA was both useful and usable in a medical context, we implemented the TAM to assess user acceptance and intention to use our explainable AI system. Originally developed by Davis et al [45], the TAM is a key model in new technology development. We adapted it for our PWA evaluation focusing on 3 dimensions: efficiency (perceived usefulness), ease of use (perceived ease of use), and intention to use (behavioral intention to use). For the TAM, a 7-point Likert scale is used, ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). A score of 4 represents neutrality.

MARS Inclusion Rationale

Despite its name, the MARS is widely used to assess the quality of health applications across various types of digital platforms. It evaluates 4 dimensions: engagement, functionality, esthetics, and information quality [46,47]. We selected the MARS to assess the information quality of our PWA using the Japanese translation by Yamamoto et al [48]. The MARS uses a 5-point Likert scale ranging from 1 (*inadequate*) to 5 (*excellent*). Scores of >3 are considered above average.

AF'fective: The Web-Based PWA

A fully web-based PWA was developed using the Ionic framework (Drifty) for the front end and Python (Python Software Foundation) for the back end. Access was strictly controlled through an authentication system to ensure data security. The PWA included 2 main pages that were developed based on our explainable strategies: a patient overview screen (corresponding to strategy D) and an individual health record dashboard (corresponding to strategies A, B, and G). One noteworthy change was made to our predictive model, which also influenced our explainable features. On the basis of physicians' feedback, we omitted direct highlights on the raw ECG graph in stage 2 (using Grad-CAM; see the demonstration in Figure 1 [39]) and, instead, used the 7-item features derived from the standardized clinical interpretation of ECG graphs (corresponding to strategies A and G in particular; see the demonstration in Figures 4 and 5). In particular, we replaced the explainable AI model used in previous stages with a SHAP model [49], which was constructed to highlight the extent to which different key clinical features—commonly used as established interpretation principles in AF prediction for ECG data (eg, maximum P wave duration and augmented vector right [aVR]/first precordial lead [V1])—contributed to the decision-making process for strategy G. Therefore, the predictive model was updated from a CNN model (image based) used in the previous 2 stages to a Cox regression model (feature based) [26] to accommodate the introduction of 7 standard features for prediction (and, later on, SHAP for explaining). While physicians were not shown the Cox regression model's performance to avoid biasing them, it achieved strong evaluations, with an area under the curve of 0.72, sensitivity of 83%, and specificity of 58%. The model was trained on data from 502 patients, with an average follow-up of 6.2 (SD 3.5) years. A total of 13.1% (66/502) of the patients developed new-onset AF [26].

These standardized clinical explainable features are included in the following list. These features were derived from the interpretation standards that cardiologists commonly apply when reading an ECG graph and assessing a patient's risk level. As indicated in our previous studies, while directly highlighting sections of the ECG graph that contributed to the predictive model's output can explain the model's prediction, it does not aid physicians in interpreting the graph. To enhance the interpretability and physician-friendly explanation, we asked cardiologists to standardize their interpretation of the ECG graph into higher-level features that are more straightforward and easier to understand for their fellow cardiologists. Our previous work [26] provides details on training a predictive model based on these new features:

1. Max P wave duration >125 ms: whether the maximum P wave duration is >125 ms
2. aVR/V1<1: whether the amplitude ratio of the P wave is <1
3. Amplitude V1≥10: whether the P wave amplitude in V1 is >0.1 mV
4. Amplitude aVR<4=0: whether the P wave amplitude in the lead aVR is <0.04 mV
5. PAC on ECG=1.0: whether there was one or more supraventricular ectopies during recording (removed in the latest model for simplicity from the work by Nishimura et al [26])
6. RV5SV1≥2.2=0: whether the amplitude (height) of the R wave in the fifth precordial lead plus the amplitude (depth) of the S wave in V1 is >2.2 mV
7. PR≥185=0: whether the interval between the P and R waves is >185 ms

The individual health record dashboard (as shown in Figure 3) also allowed physicians to check past visit ECG graphs and explainable features (corresponding to strategy B). The physicians can obtain an overview of the basic patient data in the top left corner and then proceed to the raw ECG graph. In the main graph section, they can choose between 2 types of explainable models (see Figures 4 and 5 for the corresponding

XSL•FO
RenderX

graphs) and add data from the patient's history for comparison. The right panel provides additional patient correspondence data and physical measurements such as weight and blood pressure. The standardized features made it easier to compare patients when physicians navigated the patient overview board (corresponding to strategy G).

We also developed 2 interfaces to summarize the explainable features. The first interface (Figure 4) was developed using SHAP, a popular tool for explaining the outcomes of machine learning models [49]. It features a single bar chart that shows the combined influence of risk and protective factors. The second interface uses a pie chart to show the relative importance of the 7 items and their risk or protective contributions to the predictive outcome (Figure 5). Cardiologists were advised to access it using a computer, although the site is also available on mobile devices, to simulate their normal use context.

Figure 4. Screenshot of the explainable items and how they were displayed in the system—explainable model 1. ECG: electrocardiography.

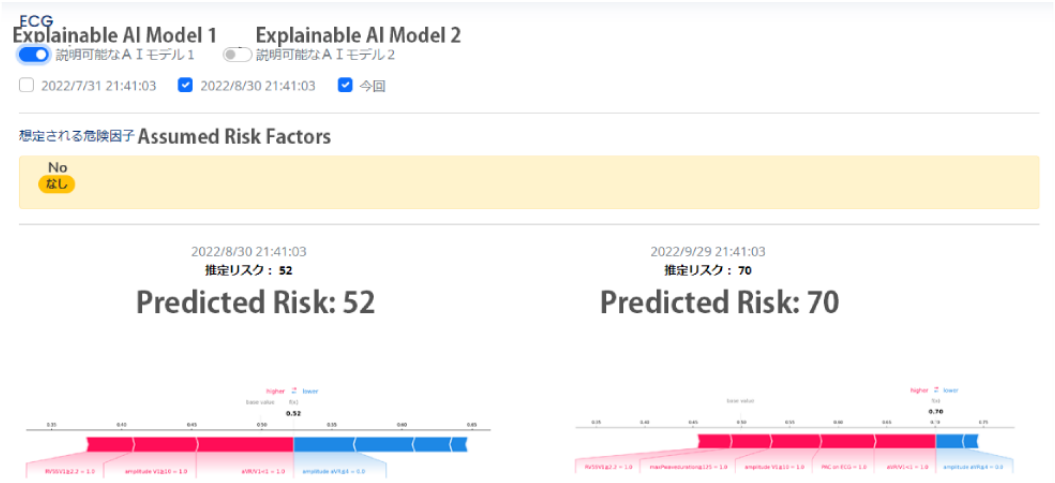
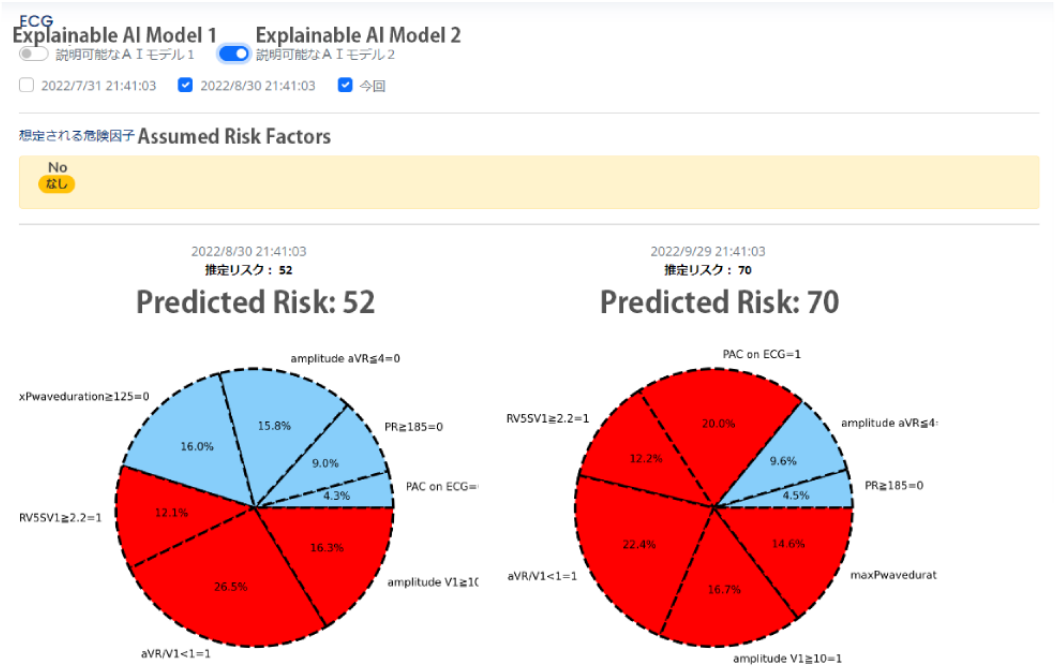


Figure 5. Screenshot of the explainable items and how they were displayed in the system—explainable model 2. aVR: Augmented Vector Right; ECG: electrocardiography. PAC: Premature Atrial Contraction; PR: PR Interval (the onset of the P wave to the start of the QRS complex).



Study Process

Each cardiologist participated in a 1-on-2 interview with an HCI researcher and a cardiologist. Before the interview began, the HCI researcher set the scene for the cardiologist using the following statement: "You have 11 patients that you are currently caring for. You can see a list of them in the AF'fective system, which indicates their risks levels for recurrence." Participants were first briefed on the system and then asked to test it by imagining the following four scenarios: (1) viewing a patient's status, (2) checking a patient's ECG graph, (3) viewing explainable model 1, and (4) viewing explainable model 2. We then asked the cardiologists to complete the study measurements (UEQ, TAM, and MARS, as described in a previous section)

XSL•FO
RenderX

and began the semistructured interviews. Each interview lasted 45 minutes to an hour. All the interviews were conducted in Japanese and fully recorded, including screen recordings of the participants testing the AF'fective PWA.

Analysis

For the survey data, we calculated the mean and SD for each item in all surveys except for the UEQ, which has its own analysis tool in Microsoft Excel (Microsoft Corp) format. We compared our mean scores for the TAM and MARS against the theoretical average scores to determine whether physicians experienced any usability issues related to the interface design, functionality, information quality, or accessibility.

For the interview data, we reviewed and analyzed the transcripts following the thematic analysis framework proposed by Clarke et al [50]. The analysis was carried out by 3 researchers, 2 from the HCI field and 1 with a cardiology background. As the 2 HCI researchers were more familiar with the analysis method selected and had a long history of conducting thematic analysis, the initial formation of the themes was first conducted by them. The HCI researchers first read through all the transcripts to familiarize themselves with the content and highlighted interesting quotes. The quotes were then summarized and independently interpreted as "topics of interest" by the HCI researchers. The researchers then grouped the topics into several potential "themes" and involved the cardiologist to approve and edit the themes. The cardiologist researcher then joined the evaluation to come up with the final findings that were deemed meaningful in both the HCI and medical fields.

Results

Overview

In the final stage, we recruited 7 male cardiologists (aged 31 to 55 years) with 8 to 30 years of cardiology experience. After the interviews, they completed the surveys, and we calculated the mean scores to identify potential usability or design issues. The results suggested no significant usability problems, and the prototype showed strong potential to motivate physicians' intention to use it due to its ease of use, reliable information, and explainable functionality. The detailed survey results are reported in the following sections.

UEQ Results: Positive Perspicuity

The results from the UEQ showed that perspicuity received a positive evaluation (mean 1.21, SD 0.96), whereas attractiveness (mean 0.56, SD 0.71), efficiency (mean 0.625, SD 0.27), dependability (mean 0.33, SD 0.27), stimulation (mean 0.5, SD 0.3), and novelty (mean 0.58, SD 0.27) received neutral evaluations. Perspicuity stood out as the most positively perceived aspect of the UX, suggesting that ease of use and clarity are strong points of the AF'fective system. The neutral evaluations for the other aspects (attractiveness, efficiency, dependability, stimulation, and novelty) indicated that physicians had average or mixed perceptions of these factors.

TAM Results: Above-Average Technical Acceptance

In addition, the TAM measures showed that perceived usefulness (mean 4.22, SD 0.31), perceived ease of use (mean 5.13, SD 0.77), and behavioral intention to use (mean 4.77, SD 0.68) scores were all above the average (mean 3.5). The behavioral intention to use score indicated a strong intention to use the AF'fective system. These results suggest that the identified factors motivated users to engage with the system.

MARS Results: Perceived High Functionality and Information Reliability

Finally, the results of the MARS showed above-average scores for functionality (mean 4.25, SD 0.42), esthetics (mean 3.61, SD 0.68), information reliability (mean 3.38, SD 0.54), and perceived impact on practice (mean 3.33, SD 0.76). These results indicated that users perceived high functionality and reliable information from the AF'fective system, whereas they rated its esthetics and perceived impact on practice as somewhat mediocre. We were particularly interested in the physicians' high ratings for information reliability as these suggested that they felt confident in understanding and agreeing with the predictive model's decisions. Notably, this rating was achieved without physicians knowing the predictive model's performance. This serves as a strong indicator that insights derived from explainable models play a crucial role in physicians' adaptation to our technology. This was deemed acceptable for a system in its prototype phase.

Discussion

Principal Findings

Overview

Aligned with our positive survey results, cardiologists expressed strong acceptance of the explainable features and willingness to use the system during the in-depth interviews. They also praised the system's enhanced efficiency, comprehensibility, comparability, consistency, and trustworthiness:

> *I think it is useful to be able to express in a general, universal, numerical way what we used to say in the past, like, "This person is likely to have a recurrence," by using a risk score. I think it is useful.* [P04; male]

> *In clinical practice, we honestly don't have that much time to look at ECGs and scores in detail, so I think AI is a way to make it easier for us to get the numbers out.* [P07; male]

> *I think it is easy to pick up if you can tell from a quick glance at the list that a person in red is at higher risk of recurrence, or that a person in green has a low probability of recurrence.* [P06; male]

While most cardiologists in our study viewed the real data–backed explainable AI system positively, we identified potential concerns and (possibly) false expectations for its real-life implementation. During the interviews, cardiologists were asked to envision using the system in their daily routines, leading to a realistic discussion after 2 rounds of iterative design. This probing step was crucial for understanding the feasibility and challenges of implementing the system in clinical settings. Our findings are discussed in the following sections through 4 key themes.

XSL•FO
RenderX

Explaining the Symptoms Does Not Equal to Identifying the Causes

Perhaps due to our intentional mimicry of cardiologists' methods for identifying irregularities in ECG graphs, participants showed high acceptance of the predicted results after viewing our 7-item AF clinical explainable features. In fact, P04 gave a very positive endorsement of why the explainable items were necessary:

> *I am of course curious as to why the risk was higher, so the AI [predictive] model would be rather meaningless without a graph [with explainable AI items] that can explain it, or perhaps it would make me wonder why the model [makes such a decision].* [P04; male]

According to P04, the predictive model's results may fail to convince them without an explanation of why the model makes a particular decision. P04's reaction suggests that physicians find explanations based on the new clinical explainable features both agreeable and trustworthy. However, a concern emerged regarding the limitations of these features. A key request was, "Then what should I do?"—cardiologists wanted to understand the features influencing the risk score and actionable strategies to address them. This aligns with previous insights showing a preference for clear indicators and actionable steps. In one case, P05 even suggested using the system to nudge patients toward behavior change:

> *We are at a stage where only these numbers [of explainable items] are available. 40 or something like this is the only thing that comes out, so what should I do [with each of the items]?* [P04; male]
>
> *I feel that the numbers alone don't mean much, or don't seem to lead to a change in [patient's] behavior.* [P05; male]

It should be noted that physicians treated each explainable feature as a "treatable item," which was not possible with an ECG graph–based explainable model. This improvement suggests that converting graphical highlights into standardized feature explanations can foster the expectation that addressing each item will reduce the risk score. In our current phase, it is still too early to determine how to further link our features with treatment approaches; however, this does indicate an interesting and operationalizable future implementation for explainable models.

Hence, in our current system, while these features influence the predictive outcome, clinical causal insights should be interpreted with caution. Our 7-item explainable features are based on standard ECG wave interpretation, identifying AF symptoms but not directly indicating underlying causes such as excessive drinking or sleep apnea. As noted in the document by Shapley [51], estimating causal effects from explainable items can be misleading. Cardiologists should be cautioned about this limitation and guided to set appropriate expectations for decision-making.

A key lesson from our interviews was the need for better curation of explainable features in the model, especially in clinical contexts. While various approaches provide transparency in AI decision-making [31,39,52,53], it is crucial that the chosen features are relevant, operationalizable, and actionable for medical professionals. This focus ensures that AI predictions are both understandable and practically applicable, enhancing postablation patient care.

Window for Interpretation: When 0% and 100% Recurrence Risk Does Not Exist

While a predictive model can theoretically provide risk scores from 0% to 100%, in real life, extreme scores such as 0% or 100% are unattainable. Participants noted that the possibility of such extreme scores diminishes the perceived realism and reliability of the model's predictions:

> *Theoretically speaking, it has to be 0, so even if it's low, it's still bad. You can't say 30 is a good thing. There is still a chance for a recurrence...So there are cases where it comes out as 0? (Cardiologist interviewer: That would be considered extreme score. Well, some might be zero, but I think the smallest is probably 3 or 4 [in the natural clinical contexts].)* [P03; male]
>
> *I'm not quite sure what [the risk score] means. There are a lot of questions about what 12 means and what we should think about it. It's not 0 [even for a normal person], So are you saying that AF can happen to anyone? (Interviewer: I don't think we can lower that number any further.)* [P05; male]

In a similar manner, participants also questioned whether a predictive score of 100% was at all meaningful in clinical practice:

> *Are you saying that even if all of [the items] were full, it would not be 100?(Interviewer: Even if [all of the items were] full, I think it will probably be around 80-92, which is the highest risk score [in the natural clinical contexts].)* [P01; male]

It should be noted that it is uncommon for a predictive model to yield a full 100% score in clinical contexts, particularly in postsurgery contexts. The contributing factors would generally contain complex interactions and dependencies.

Cardiologists emphasized that recurrence risk scores should be viewed as reference points, not absolute indicators. Establishing a threshold, such as 50, is essential for guiding actions, but borderline scores (eg, 49 or 51) still warrant patient warnings or follow-ups. Similar "borderline windows" were noted in a grief-related study [52] where psychologists suggested taking the score as a reference for concerning cases. Although the width of the borderline window can also be highly subjective for each individual cardiologist and possibly for patients, this approach is common in clinical practice. Future studies should develop nuanced protocols for borderline cases to prevent misdiagnoses and patient stress, potentially standardizing risk level interpretation, as suggested by P05:

> *I think it would be better to indicate the risk of developing the disease on a 5-point scale, such as low, medium, and high, because I don't think there is such a difference between 40 and 41.* [P05; male]

XSL•FO
RenderX

In treatment decision-making, another notable implication from our findings was the patient's presumptive belief that the risk score should be 0%, particularly after undergoing ablation. It was apparently conflicting with what the predictive model would reveal:

> *However, if you think about it, after AF ablation, the number of people who think that AF recurrence is quite small. For example, more than 70% to 80% of the patients think it is zero.* [P03; male]

Even cardiologists familiar with the postablation risk might expect that addressing the explainable items (see the Methods section for more discussion on these items) could eliminate the recurrence risk:

> *Considering the importance of risk scoring...I am very interested in the explanatory factors or items involved, and which of them are [positively correlated to the risk score]. If the purpose of this system is to follow up the patient, it would be even better if we could get such risk factors and know what kind of intervention is being done here.* [P01; male]

It struck the research team that popular explainable AI approaches could lead to misinterpretations and unrealistic expectations, such as believing that a 0% recurrence risk is easily attainable. Our study highlights the need to set accurate expectations about what predictive results and explainable features truly imply in a clinical context and how much they can be influenced through intervention.

Offer a Holistic Risk Overview by Integrating Other Risk Factors and Temporal Data

In our study, we adopted the flow of cardiologists' diagnostic process by reviewing ECG graphs and assessing the presence of known risk factors one by one. Therefore, other known risk factors, such as high blood pressure, smoking, and drinking, were included as additional "risk factors" alongside the predicted risk score derived from the patient's ECG data. While we believed that this approach resembled a cardiologist's normal decision-making procedure and would increase their acceptance of the model's predicted results, our interviewees expressed concerns about the potential for the predictive model to yield biased risk estimates:

> *There are many known risks that are commonly associated with AF, such as high blood pressure, obesity, lack of exercise, etc... Inputting it would further make this risk [score], or perhaps, the evaluation, more correct.* [P03; male]

> *So we are predicting the risk score purely based on ECG...If you are going to make progress, I think it would be good to include obesity information, bmi, waist circumference, and so on.* [P06; male]

Furthermore, some of our participants also indicated that they would prefer a clear display and association with the medications or ongoing treatments that the patients are receiving. For instance, P02 requested us to explore whether the explainable model could reflect the effects of the patient's medications, and P01 indicated that it would be better to see which interventions were done to lead to the risk score:

> *If there was information on arrhythmia medications, it would be more useful. I think the doctor might be able to think about it more holistically, such as, "This ECG variation [highlighted by the explainable model] is caused by the medicine the patient is taking."* [P02; male]

> *If the purpose of this system is to follow up the patient, it would be even better if we could get such risk scores and know what kind of intervention is being done here.* [P01; male]

As the cardiologists pointed out, although our model provided an overview of the other risk factors, it would only be meaningful if they could establish a connection between the variation in risk scores and the specific risk factors. Our findings indicated that, for explainable elements to be truly useful in real-life practice, they need to relate to more holistic and treatment-related risk factors or medications. Otherwise, the contribution of such a predictive, even explainable, model to preventing disease recurrence would be quite limited.

In the end, perhaps the most important takeaway from the interviews was the need to provide a holistic overview and explanation of intrapatient temporal data and risk scores. Echoing the insights from previous interviews, the design strategy of comparing interpersonal ECGs (strategy G) was heavily criticized by the physicians as individual risk factors can vary significantly. However, comparing a patient's current and historical risk factors could provide valuable insights into the evolution of these factors and potential trends in recurrence risk. For instance, participants P04 and P06 described how they would like to plan their interventions by tracking variations in risk scores and identifying the underlying causal factors over different time points:

> *For a patient with a high estimated risk this time, but lower estimated risk in the previous times, why did it become so high? What factors contributed to it? [I would expect] this AI model to estimate and explain.* [P04; male]

> *I think it would be better to look at it...from a longitudinal point of view...to see what kind of intervention I should use, what's going wrong, and what the reason is for the increase [of the risk score].* [P06; male]

We were intrigued by the design opportunities revealed in our findings. As cardiologists reminded us, a significant and crucial aspect of their profession involves deciding on appropriate interventions. Although explainable AI models have shown promise in diagnosing and identifying patients at risk, they have yet to offer actionable insights for intervention planning. We believe that explainable AI tools should integrate comprehensive patient data, including temporal variations in risk factors, medications, and treatments, to enhance the model's relevance and utility in real-world settings. In addition, the insights generated should be clearly linked to potential intervention strategies that could influence risk scores. Without such integration, the potential of explainable AI tools to contribute meaningfully to patient care, particularly in preventing disease recurrence, may remain limited.

XSL•FO
RenderX

A Sense of Control and Learning From the Model

The findings of our interviews revealed that cardiologists valued the sense of control and the learning opportunities provided by these models, aligned with the findings in the previous rounds that suggested a self-actualizing and constant learning side of cardiologists.

One of the key insights that emerged from our study is the potential of explainable AI models to help cardiologists prioritize patients at higher risk. As P04 addressed, there was a potential for these models to become an integral part of daily practice:

> *If it becomes a habit, I will probably look at people with high estimated risk and wonder why they were at higher risk.* [P04; male]

Moreover, P02 also offered a similar thought regarding triaging patients based on the risk scores:

> *So the follow-up should be more thorough, and those who are more likely to be affected [by the recurrence risk] should be followed up in a shorter period of time.* [P02; male]

By discriminating the patients that need urgent attention, physicians gain a sense of control over who to allocate more time and treatment resources to. Cardiologists indicated that such a potential new routine would enhance their work efficiency and allow them to spare cognitive capacity for more critical treatment planning. Interestingly, our interview findings suggest that cardiologists may appreciate useful "hints" or "interpretations" from explainable models to help them narrow down potential problems:

> *I understand that the ECG is used as the reference, but I think it would be easier to understand if there was an explanation of what is of interest in this ECG, even though those who understand may [already] understand it.* [P01; male]

The power dynamic between cardiologists and the AI model is worth exploring. To what extent should an AI model's interpretations be considered, and to what extent should cardiologists accept its suggestions? These boundaries are critical to explore to ensure that the final AI models can be effectively integrated into the patient care routines.

Further enhancement of control over the prevention procedure might also occur as cardiologists begin to learn more from AI-augmented interpretation and actionability. Interestingly, the explainability of AI models appears to increase cardiologists' desire to validate their treatment approaches:

> *So, at the time of the ablation, if I burn the tricuspid valve, you know [the score will go down], or if I burn the superior vena cava, you know [the score will go down]...you can learn more and more from time to time.* [P03; male]
>
> *I would take another ECG before prescribing too much medication, and see if [the risk score] goes up or down on the next ECG...For those in yellow...I will adjust my [treatment plan] according to the risk level of the next ECG.* [P06; male]

P06 went further and emphasized that, while AI models are valuable for providing quantitative assessments, their true worth lies in delivering actionable insights that can guide clinical decision-making. This sentiment reflects a broader interest among cardiologists in understanding how AI models arrive at their conclusions, particularly in relation to identifying intervention points:

> *I wonder if you are highlighting these risk factors to tell us these are the areas for intervention. I wonder if there is a reason for this risk factor, such as if we treat these factors, it could lower the risk.* [P06; male]

In general, our findings indicate that cardiologists, particularly self-actualizers, are not only open to integrating AI suggestions into their decision-making processes but also eager to enhance their knowledge through interaction with the AI model provided that these suggestions are accompanied by clear rationales and practical recommendations.

Implications of Creating a Software as a Medical Device

In the development of AI algorithms for clinical applications, it is crucial to not only meet the user interface requirements but also ensure compliance with regulatory standards. Software as a medical device must adhere to international standards such as International Organization for Standardization (ISO) 13485 (medical device quality management systems), ISO 14971 (medical device risk management), and ISO 62304 (software life cycle processes). In addition, compliance with Food and Drug Administration approval and the European Medical Device Regulation is essential from the perspectives of patient safety, effectiveness, and legal requirements.

While the AF'fective study provides valuable foundational knowledge on user interface, it is still considered in the early stages of technology readiness level (1-3). However, with advancements in AI technology, the insights from this research could contribute to the revision of existing standards, such as the International Electrotechnical Commission 62366-1:2015 (usability engineering for medical devices). Specifically, the consideration of explainability in the AI result display methods proposed in this research—such as visually clarifying the rationale behind suggestions and intuitively conveying risk and uncertainty in interface design—will play a crucial role in helping health care professionals accurately understand AI-generated proposals and integrate them into clinical decision-making. Moreover, our study highlighted that health care professionals do not need to fully comprehend the technical background or mechanisms of AI. Instead, it is crucial to create a state in which users can reasonably understand "why this result was reached," which is effective in building trust in AI. Such efforts extend beyond display design and have the potential to be incorporated into new standards that facilitate smooth communication between AI and health care professionals.

Conclusions

This paper reports a 3-stage study involving 23 physicians to co-design, assess, and pilot an explainable AI system for AF postablation monitoring called AF'fective. Our co-design approach effectively identified 4 feasible explainable strategies.

XSL•FO
RenderX

Furthermore, while we actively used explainable AI to identify the features contributing to the predictive results, our explainable strategies enabled the design of a more holistic patient monitoring system that incorporates contextual factors such as patient history and established principles to support physicians' decision-making rather than merely explaining the features indicated by ECG graphs. In the third stage, the system prompted cardiologists to envision its use in their routines, leading to highly contextualized feedback. This study highlighted the need to curate actionable explainable features and set correct expectations for interpreting predictive scores. In addition, cardiologists were interested in understanding the AI's reasoning and identifying strategies to address recurrence risk factors.

Interestingly, although our focus was on AF recurrence prevention, many insights appear generalizable to other disease care contexts, such as temporal risk score monitoring and setting realistic expectations for scoring, particularly as extreme scores (0% and 100%) are practically unattainable. Further investigation is needed to determine whether these insights can be universally applied to enhance the feasibility of implementing explainable AI in clinical settings.

A future study will build on these findings and further evaluate the potential challenges and advantages when implementing the AF'fective system in near-live or live clinical care routines. We believe that AI possesses the potential to revolutionize medical practices and we can only realize it through putting the technology in actual use case scenarios.

Acknowledgments

This work was supported by a research grant from the Japanese Circulation Society (2021).

Conflicts of Interest

None declared.

References

1. von Eschenbach WJ. Transparency and the black box problem: why we do not trust AI. Philos Technol 2021 Sep 01;34(4):1607-1622. [doi: 10.1007/S13347-021-00477-0]
2. Petch J, Di S, Nelson W. Opening the black box: the promise and limitations of explainable machine learning in cardiology. Can J Cardiol 2022 Feb;38(2):204-213 [FREE Full text] [doi: 10.1016/j.cjca.2021.09.004] [Medline: 34534619]
3. Bizopoulos P, Koutsouris D. Deep learning in cardiology. IEEE Rev Biomed Eng 2019;12:168-193. [doi: 10.1109/RBME.2018.2885714] [Medline: 30530339]
4. Nakamura T, Sasano T. Artificial intelligence and cardiology: current status and perspective. J Cardiol 2022 Mar;79(3):326-333 [FREE Full text] [doi: 10.1016/j.jjcc.2021.11.017] [Medline: 34895982]
5. Nattel S. New ideas about atrial fibrillation 50 years on. Nature 2002 Jan 10;415(6868):219-226. [doi: 10.1038/415219a] [Medline: 11805846]
6. Lippi G, Sanchis-Gomar F, Cervellin G. Global epidemiology of atrial fibrillation: an increasing epidemic and public health challenge. Int J Stroke 2021 Feb;16(2):217-221. [doi: 10.1177/1747493019897870] [Medline: 31955707]
7. Odutayo A, Wong CX, Hsiao AJ, Hopewell S, Altman DG, Emdin CA. Atrial fibrillation and risks of cardiovascular disease, renal disease, and death: systematic review and meta-analysis. BMJ 2016 Sep 06;354:i4482 [FREE Full text] [doi: 10.1136/bmj.i4482] [Medline: 27599725]
8. Lip GY, Tse HF. Management of atrial fibrillation. Lancet 2007 Aug 18;370(9587):604-618. [doi: 10.1016/S0140-6736(07)61300-2] [Medline: 17707756]
9. Dretzke J, Chuchu N, Agarwal R, Herd C, Chua W, Fabritz L, et al. Predicting recurrent atrial fibrillation after catheter ablation: a systematic review of prognostic models. Europace 2020 May 01;22(5):748-760 [FREE Full text] [doi: 10.1093/europace/euaa041] [Medline: 32227238]
10. Shashikumar SP, Shah AJ, Li Q, Clifford GD, Nemati S. A deep learning approach to monitoring and detecting atrial fibrillation using wearable technology. In: Proceedings of the 2017 IEEE EMBS International Conference on Biomedical & Health Informatics. 2017 Presented at: BHI '17; February 16-19, 2017; Orlando, FL p. 141-144 URL: https://ieeexplore.ieee.org/document/7897225 [doi: 10.1109/bhi.2017.7897225]
11. Andersen RS, Peimankar A, Puthusserypady S. A deep learning approach for real-time detection of atrial fibrillation. Expert Syst Appl 2019 Jan;115:465-473. [doi: 10.1016/j.eswa.2018.08.011]
12. Zhou X, Nakamura K, Sahara N, Takagi T, Toyoda Y, Enomoto Y, et al. Deep learning-based recurrence prediction of atrial fibrillation after catheter ablation. Circ J 2022 Jan 25;86(2):299-308 [FREE Full text] [doi: 10.1253/circj.CJ-21-0622] [Medline: 34629373]
13. Hung M, Lauren E, Hon E, Xu J, Ruiz-Negrón B, Rosales M, et al. Using machine learning to predict 30-day hospital readmissions in patients with atrial fibrillation undergoing catheter ablation. J Pers Med 2020 Aug 09;10(3):82 [FREE Full text] [doi: 10.3390/jpm10030082] [Medline: 32784873]

XSL•FO
RenderX

14. Przewlocka-Kosmala M, Marwick TH, Dabrowski A, Kosmala W. Contribution of cardiovascular reserve to prognostic categories of heart failure with preserved ejection fraction: a classification based on machine learning. J Am Soc Echocardiogr 2019 May;32(5):604-15.e6. [doi: 10.1016/j.echo.2018.12.002] [Medline: 30718020]
15. Mortazavi BJ, Downing NS, Bucholz EM, Dharmarajan K, Manhapra A, Li SX, et al. Analysis of machine learning techniques for heart failure readmissions. Circ Cardiovasc Qual Outcomes 2016 Nov;9(6):629-640. [doi: 10.1161/circoutcomes.116.003039]
16. Isensee F, Jaeger PF, Full PM, Wolf I, Engelhardt S, Maier-Hein KH. Automatic cardiac disease assessment on cine-MRI via time-series segmentation and domain specific features. In: Proceedings of the 8th International Workshop, STACOM 2017, Held in Conjunction with MICCAI 2017 on Statistical Atlases and Computational Models of the Heart, ACDC and MMWHS Challenges. 2017 Presented at: STACOM '17; September 10-14, 2017; Montreal, QC p. 120-129 URL: https://link.springer.com/chapter/10.1007/978-3-319-75541-0_13 [doi: 10.1007/978-3-319-75541-0_13]
17. Nagarajan VD, Lee SL, Robertus JL, Nienaber CA, Trayanova NA, Ernst S. Artificial intelligence in the diagnosis and management of arrhythmias. Eur Heart J 2021 Oct 07;42(38):3904-3916 [FREE Full text] [doi: 10.1093/eurheartj/ehab544] [Medline: 34392353]
18. Ebrahimi Z, Loni M, Daneshtalab M, Gharehbaghi A. A review on deep learning methods for ECG arrhythmia classification. Expert Syst Appl 2020 Sep;7:100033. [doi: 10.1016/j.eswax.2020.100033]
19. Tison GH, Sanchez JM, Ballinger B, Singh A, Olgin JE, Pletcher MJ, et al. Passive detection of atrial fibrillation using a commercially available smartwatch. JAMA Cardiol 2018 May 01;3(5):409-416 [FREE Full text] [doi: 10.1001/jamacardio.2018.0136] [Medline: 29562087]
20. Nazarian S, Lam K, Darzi A, Ashrafian H. Diagnostic accuracy of smartwatches for the detection of cardiac arrhythmia: systematic review and meta-analysis. J Med Internet Res 2021 Aug 27;23(8):e28974 [FREE Full text] [doi: 10.2196/28974] [Medline: 34448706]
21. Sehrawat O, Kashou AH, Noseworthy PA. Artificial intelligence and atrial fibrillation. J Cardiovasc Electrophysiol 2022 Aug 15;33(8):1932-1943 [FREE Full text] [doi: 10.1111/jce.15440] [Medline: 35258136]
22. Feeny AK, Chung MK, Madabhushi A, Attia ZI, Cikes M, Firouznia M, et al. Artificial intelligence and machine learning in arrhythmias and cardiac electrophysiology. Circ Arrhythm Electrophysiol 2020 Aug;13(8):e007952 [FREE Full text] [doi: 10.1161/CIRCEP.119.007952] [Medline: 32628863]
23. Kim JY, Kim Y, Oh GH, Kim SH, Choi Y, Hwang Y, et al. A deep learning model to predict recurrence of atrial fibrillation after pulmonary vein isolation. Int J Arrhythm 2020 Nov 16;21(1):95. [doi: 10.1186/S42444-020-00027-3]
24. Budzianowski J, Kaczmarek-Majer K, Rzeźniczak J, Słomczyński M, Wichrowski F, Hiczkiewicz D, et al. Machine learning model for predicting late recurrence of atrial fibrillation after catheter ablation. Sci Rep 2023 Sep 14;13(1):15213 [FREE Full text] [doi: 10.1038/s41598-023-42542-y] [Medline: 37709859]
25. Li Z, Wheelock KM, Lathkar-Pradhan S, Oral H, Clauw DJ, Gunaratne P, et al. Predicting atrial fibrillation episodes with rapid ventricular rates associated with low levels of activity. BMC Med Inform Decis Mak 2021 Dec 28;21(1):364 [FREE Full text] [doi: 10.1186/s12911-021-01723-3] [Medline: 34963444]
26. Nishimura T, Senoo K, Makino M, Munakata J, Tomura N, Shimoo S, et al. Prediction model for the new onset of atrial fibrillation combining features of 24-hour Holter electrocardiogram with 12-lead electrocardiogram. Int J Cardiol Heart Vasc 2023 Aug;47:101245 [FREE Full text] [doi: 10.1016/j.ijcha.2023.101245] [Medline: 37521520]
27. Hallowell N, Badger S, Sauerbrei A, Nellåker C, Kerasidou A. "I don't think people are ready to trust these algorithms at face value": trust and the use of machine learning algorithms in the diagnosis of rare disease. BMC Med Ethics 2022 Nov 16;23(1):112 [FREE Full text] [doi: 10.1186/s12910-022-00842-4] [Medline: 36384545]
28. K S, V S, EA G, KP S. Explainable artificial intelligence for heart rate variability in ECG signal. Healthc Technol Lett 2020 Dec 09;7(6):146-154. [doi: 10.1049/htl.2020.0033]
29. Jo YY, Cho Y, Lee SY, Kwon J, Kim KH, Jeon KH, et al. Explainable artificial intelligence to detect atrial fibrillation using electrocardiogram. Int J Cardiol 2021 Apr 01;328:104-110. [doi: 10.1016/j.ijcard.2020.11.053] [Medline: 33271204]
30. Ma Y, Zhang D, Xu J, Pang H, Hu M, Li J, et al. Explainable machine learning model reveals its decision-making process in identifying patients with paroxysmal atrial fibrillation at high risk for recurrence after catheter ablation. BMC Cardiovasc Disord 2023 Feb 17;23(1):91 [FREE Full text] [doi: 10.1186/s12872-023-03087-0] [Medline: 36803424]
31. Anand A, Kadian T, Shetty MK, Gupta A. Explainable AI decision model for ECG data of cardiac disorders. Biomed Signal Process Control 2022 May;75:103584. [doi: 10.1016/j.bspc.2022.103584]
32. Mousavi S, Afghah F, Acharya UR. HAN-ECG: an interpretable atrial fibrillation detection model using hierarchical attention networks. Comput Biol Med 2020 Dec;127:104057 [FREE Full text] [doi: 10.1016/j.compbiomed.2020.104057] [Medline: 33126126]
33. Mousavi S, Afghah F, Razi A, Acharya UR. ECGNET: learning where to attend for detection of atrial fibrillation with deep visual attention. IEEE EMBS Int Conf Biomed Health Inform 2019 May;2019:854 [FREE Full text] [doi: 10.1109/BHI.2019.8834637] [Medline: 33083788]
34. Raza A, Tran KP, Koehl L, Li S. Designing ECG monitoring healthcare system with federated transfer learning and explainable AI. Knowl Based Syst 2022 Jan;236:107763. [doi: 10.1016/j.knosys.2021.107763]

XSL•FO
RenderX

35. Wesołowski S, Lemmon G, Hernandez EJ, Henrie A, Miller TA, Weyhrauch D, et al. An explainable artificial intelligence approach for predicting cardiovascular outcomes using electronic health records. PLOS Digit Health 2022 Jan 18;1(1):e0000004 [FREE Full text] [doi: 10.1371/journal.pdig.0000004] [Medline: 35373216]
36. Matthiesen S, Diederichsen SZ, Hansen MK, Villumsen C, Lassen MC, Jacobsen PK, et al. Clinician preimplementation perspectives of a decision-support tool for the prediction of cardiac arrhythmia based on machine learning: near-live feasibility and qualitative study. JMIR Hum Factors 2021 Nov 26;8(4):e26964 [FREE Full text] [doi: 10.2196/26964] [Medline: 34842528]
37. Gerlings J, Jensen MS, Shollo A. Explainable AI, but explainable to whom? An exploratory case study of xAI in healthcare. In: Lim CP, Chen YW, Vidya A, Mahorkar C, Jain LC, editors. Handbook of Artificial Intelligence in Healthcare. Volume 2. Cham, Switzerland: Springer; 2022:169-198.
38. Payrovnaziri SN, Chen Z, Rengifo-Moreno P, Miller T, Bian J, Chen J, et al. Explainable artificial intelligence models using real-world electronic health record data: a systematic scoping review. J Am Med Inform Assoc 2020 Jul 01;27(7):1173-1185 [FREE Full text] [doi: 10.1093/jamia/ocaa053] [Medline: 32417928]
39. She WJ, Senoo K, Iwakoshi H, Kuwahara N, Siriaraya P. AF'fective design: supporting atrial fibrillation post-treatment with explainable AI. In: Proceedings of the 27th International Conference on Intelligent User Interfaces. 2022 Presented at: IUI '22; March 22-25, 2022; Helsinki, Finland p. 22-25 URL: https://dl.acm.org/doi/10.1145/3490100.3516455 [doi: 10.1145/3490100.3516455]
40. Xie Y, Chen M, Kao D, Gao G, Chen XA. CheXplain: enabling physicians to explore and understand data-driven, AI-enabled medical imaging analysis. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. 2020 Presented at: CHI '20; April 25-30, 2020; Honolulu, HI p. 1-13 URL: https://dl.acm.org/doi/10.1145/3313831.3376807 [doi: 10.1145/3313831.3376807]
41. Alhadreti O, Mayhew P. Rethinking thinking aloud: a comparison of three think-aloud protocols. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI '18; April 21-26, 2018; Montreal, QC p. 1-12 URL: https://dl.acm.org/doi/10.1145/3173574.3173618 [doi: 10.1145/3173574.3173618]
42. Laugwitz B, Held T, Schrepp M. Construction and evaluation of a user experience questionnaire. In: Proceedings of the 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society. 2008 Presented at: USAB '08; November 20-21, 2008; Graz, Austria p. 63-76 URL: https://link.springer.com/chapter/10.1007/978-3-540-89350-9_6 [doi: 10.1007/978-3-540-89350-9_6]
43. Schrepp M, Hinderks A, Thomaschewski J. Construction of a benchmark for the user experience questionnaire (UEQ). Int J Interact Multimed Artif Intell 2017;4(4):40. [doi: 10.9781/ijimai.2017.445]
44. Davis FD, Venkatesh V. A critical assessment of potential measurement biases in the technology acceptance model: three experiments. Int J Hum Comput Stud 1996 Jul;45(1):19-45. [doi: 10.1006/ijhc.1996.0040]
45. Davis F, Bagozzi R, Warshaw P. User acceptance of computer technology: a comparison of two theoretical models. Manag Sci 1989;35(8):1003 [FREE Full text]
46. Terhorst Y, Philippi P, Sander LB, Schultchen D, Paganini S, Bardus M, et al. Validation of the Mobile Application Rating Scale (MARS). PLoS One 2020 Nov 2;15(11):e0241480 [FREE Full text] [doi: 10.1371/journal.pone.0241480] [Medline: 33137123]
47. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
48. Yamamoto K, Ito M, Sakata M, Koizumi S, Hashisako M, Sato M, et al. Japanese version of the Mobile App Rating Scale (MARS): development and validation. JMIR Mhealth Uhealth 2022 Apr 14;10(4):e33725 [FREE Full text] [doi: 10.2196/33725] [Medline: 35197241]
49. Lundberg S, Lee SI. A unified approach to interpreting model predictions. In: Proceedings of the 31st International Conference on Neural Information Processing Systems. 2017 Presented at: NIPS '17; December 4-9, 2017; Long Beach, CA p. 4768-4777 URL: https://dl.acm.org/doi/10.5555/3295222.3295230
50. Clarke V, Braun V, Hayfield N. Thematic analysis. In: Smith JA, editor. Qualitative Psychology: A Practical Guide to Research Methods. Thousand Oaks, CA: Sage Publications; 2015:23-51.
51. Lundberg S. Be careful when interpreting predictive models in search of causal insights. Medium. URL: https://towardsdatascience.com/be-careful-when-interpreting-predictive-models-in-search-of-causal-insights-e68626e664b6 [accessed 2024-04-29]
52. She WJ, Ang CS, Neimeyer RA, Burke LA, Zhang Y, Jatowt A, et al. Investigation of a web-based explainable AI screening for prolonged grief disorder. IEEE Access 2022;10:41164-41185. [doi: 10.1109/access.2022.3163311]
53. Taniguchi H, Takata T, Takechi M, Furukawa A, Iwasawa J, Kawamura A, et al. Explainable artificial intelligence model for diagnosis of atrial fibrillation using Holter electrocardiogram waveforms. Int Heart J 2021;62(3):534-539 [FREE Full text] [doi: 10.1536/ihj.21-094] [Medline: 34053998]

XSL•FO
RenderX

Abbreviations

AF: atrial fibrillation
AI: artificial intelligence
aVR: augmented vector right
CNN: convolutional neural network
ECG: electrocardiography
Grad-CAM: Gradient-Weighted Class Activation Mapping
HCI: human-computer interaction
IRB: institutional review board
ISO: International Organization for Standardization
MARS: Mobile App Rating Scale
PWA: progressive web application
SHAP: Shapley Additive Explanations
TAM: technology acceptance model
UEQ: User Experience Questionnaire
UX: user experience
V1: first precordial lead

Edited by A Kushniruk; submitted 29.08.24; peer-reviewed by D Almonfrey, HJ Noordmans; comments to author 14.10.24; revised version received 27.11.24; accepted 05.01.25; published 13.02.25.

Please cite as:
She WJ, Siriaraya P, Iwakoshi H, Kuwahara N, Senoo K
An Explainable AI Application (AF'fective) to Support Monitoring of Patients With Atrial Fibrillation After Catheter Ablation: Qualitative Focus Group, Design Session, and Interview Study
JMIR Hum Factors 2025;12:e65923
URL: https://humanfactors.jmir.org/2025/1/e65923
doi:10.2196/65923
PMID:

©Wan Jou She, Panote Siriaraya, Hibiki Iwakoshi, Noriaki Kuwahara, Keitaro Senoo. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 13.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Preadolescent Children Using Real-Time Heart Rate During Moderate to Vigorous Physical Activity: A Feasibility Study

Lincoln Lu[1], PhD; Danielle E Jake-Schoffman[2], PhD; Hannah A Lavoie[2], MS; Maedeh Agharazidermani[1], PhD; Kristy Elizabeth Boyer[1], PhD

[1]LearnDialogue Lab, Computer and Information Science and Engineering, University of Florida, Gainesville, FL, United States

[2]Exhale Lab, Health and Human Performance, University of Florida, Gainesville, FL, United States

Corresponding Author:
Lincoln Lu, PhD
LearnDialogue Lab
Computer and Information Science and Engineering
University of Florida
1889 Museum Road
Gainesville, FL, 32611
United States
Phone: 1 352 392 1133
Email: lincolnlu@ufl.edu

Abstract

Background: Given the global burden of insufficient physical activity (PA) in children, effective behavioral interventions are needed to increase PA levels. Novel technologies can help expand the reach and accessibility of these programs. Despite the potential to use heart rate (HR) to target moderate- to vigorous-intensity PA (MVPA), most HR research to date has focused on the accuracy of HR devices or used HR for PA surveillance rather than as an intervention tool. Furthermore, most commercial HR sensors are designed for adults, and their suitability for children is unknown. Further research about the feasibility and usability of commercial HR devices is required to understand how children may use HR during PA.

Objective: This study aimed to explore the use of a chest-worn HR sensor paired with a real-time HR display as an intervention tool among preadolescent children and the usability of a custom-designed app (Connexx) for viewing real-time HR.

Methods: We developed Connexx, an HR information display app with an HR analytics portal to view HR tracking. Children were recruited via flyers distributed at local public schools, word of mouth, and social media posts. Eligible participants were children aged 9 to 12 years who did not have any medical contraindications to MVPA. Participants took part in a single in-person study session where they monitored their own HR using a commercial HR sensor, learned about HR, and engaged in a series of PAs while using the Connexx app to view their real-time HR. We took field note observations about participant interactions with the HR devices. Participants engaged in a semistructured interview about their experience using Connexx and HR during PA and completed the System Usability Scale (SUS) about the Connexx app. Study sessions were audio and video recorded and transcribed verbatim.

Results: A total of 11 participants (n=6, 55% male; n=9, 82%, non-Hispanic White) with an average age of 10.4 (SD 1.0) years were recruited for the study. Data from observations, interviews, and SUS indicated that preadolescent children can use real-time HR information during MVPA. Observational and interview data indicated that the participants were able to understand their HR after a basic lesson and demonstrated the ability to make use of their HR information during PA. Interview and SUS responses demonstrated that the Connexx app was highly usable, despite some accessibility challenges (eg, small display font). Feedback about usability issues has been incorporated into a redesign of the Connexx app, including larger, color-coded fonts for HR information.

Conclusions: The results of this study indicate that preadolescent children understood their HR data and were able to use it in real time during PA. The findings suggest that future interventions targeting MVPA in this population should test strategies to use HR and HR monitoring as direct program targets.

(JMIR Hum Factors 2025;12:e58715) doi:10.2196/58715

XSL•FO
RenderX

KEYWORDS

smartphone app; physical activity; heart rate; wearable sensors; youth; commercial wearable device; Garmin; mobile phone

Introduction

Background

Physical activity (PA) levels are insufficient across most of the world's population [1], constituting a persistent problem in global health. PA engagement in the United States is reflective of global trends: only 25% of US children meet the minimum daily recommended amount of moderate- to vigorous- PA (MVPA) [2]. PA behaviors and insufficient PA developed during childhood often worsen during adolescence if not addressed through intervention [3]. Insufficient PA is also a risk factor for obesity, a condition related to various comorbidities [4] that can persist into adulthood [5,6]. Alarmingly, children who develop obesity at a young age are more likely to have obesity into adulthood [7]. Therefore, addressing obesogenic factors (eg, PA) during childhood has the potential to positively impact long-term overall health [8,9].

Effective behavioral interventions are essential to address insufficient PA, particularly before the significant decline that often occurs during adolescence, around the age of 13 years [10]. Novel technologies can help expand the reach and accessibility of PA promotion efforts. To date, most research leveraging commercial PA devices has focused on adult populations [11], although some studies focusing on children have examined step counts as a PA outcome [12]. Step count measures basic movement, but this metric insufficiently captures the intensity of movement and, therefore, may be inadequate for measuring or promoting MVPA.

In contrast to step count, heart rate (HR) holds great promise for measuring MVPA. Moreover, it is possible that novel technologies allowing children to view their HR in real time could foster PA awareness and engagement. However, this great potential is understudied; a recent systematic review found only two studies that examined the use of existing commercial sports-style HR sensors for use in preadolescent children [13]. The studies identified by the systematic review focused on whether children would wear an HR sensor on their chest paired with a wrist-worn HR display over a long time period [14] and the accuracy of commercial HR sensors on a preadolescent child's wrist against a medical-grade control [15]. Another recent study involved the use of chest-worn HR sensors paired with a wrist-worn HR display to monitor PA in adolescents and young adults. However, the study did not provide participants with education on HR or guidance on effectively using HR to achieve MVPA goals [16].

These studies demonstrate the accuracy of commercial HR sensors and displays for preadolescent children. The adoption of devices capable of displaying real-time HR information is growing in this age group, matching trends across the demographic spectrum. Recent statistics indicate wide adoption of smartphones: 97% of Americans own a smartphone [17], and ownership is equitable across racial and socioeconomic groups [18]. Statistics from a survey conducted in July 2019 reported that approximately 21% of the US population owned a smartwatch [19]. Smart device adoption is growing among preadolescent children in the United States; by their teenage years, 95% of children either have their own smartphone or have access to one through their family [20]. Despite the growing ubiquity of these devices, research has only focused on the technical ability of smart devices to surveille user HR [16]. Research should now be conducted to explore the use of real-time HR as an intervention tool to target specific MVPA goals.

Objectives

Pairing chest-worn HR sensors with existing smart devices in the form of smartphones and smartwatches is a novel opportunity that has not been sufficiently explored. However, providing real-time HR information without exploring preadolescent children's ability to understand and use HR in real time to effectively target MVPA would only replicate existing studies with new hardware [16]. Therefore, we provided participants with information about HR and how to target different HR zones during PA. This study primarily aimed to explore whether preadolescent children can use real-time HR information to target specific MVPA goals. A chest-worn HR sensor was paired to a smart device (eg, smartphone or smartwatch) that displayed real-time HR information. As the HR sensors used in this study were originally designed for adult use, a secondary objective of this study was to explore whether preadolescent children could feasibly use a chest-worn HR sensor paired with a smart device.

Another secondary objective of this study was to explore the usability of a custom app (Connexx) developed by the research team. As part of a broader program of research, the team is exploring the impact of collaboration on PA, and one facet of this research explores collaborative pairs in a family dyad consisting of a parent or guardian and child. Dyads collaborate by viewing each other's real-time HR information during PA. There are currently no commercial apps capable of transmitting and displaying real-time HR information between two users. We developed the Connexx app with this functionality. However, a key aim of this study was to explore the usability of Connexx for preadolescent children, who represent half of the intended pair in parent or guardian and child collaborative dyads. Thus, this study explored the usability of the Connexx app in solo mode to establish its feasibility in use with preadolescent children.

Methods

Study Setting

This feasibility study was conducted at the University of Florida, Gainesville, Florida. Data collection occurred between October 2023 and February 2024.

Ethical Considerations

The study was approved by the University of Florida Institutional Review Board (IRB202301463) and complies with the ethical guidelines mandated therein. Informed consent was

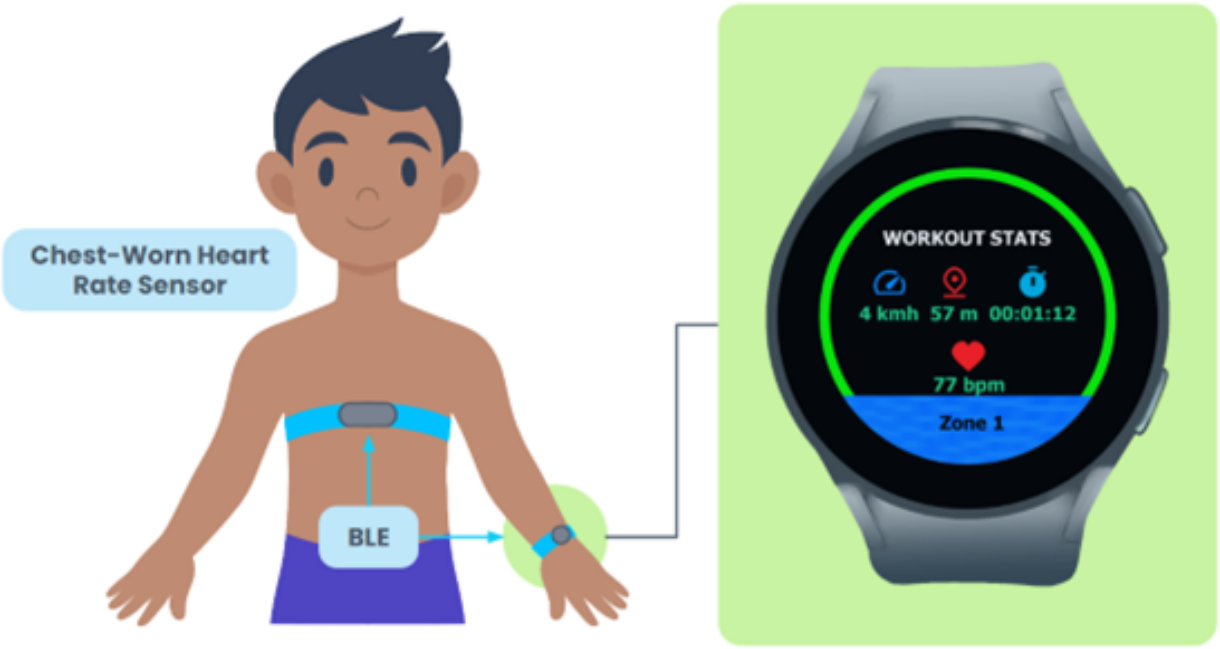
XSL•FO
RenderX

obtained from a parent or guardian electronically, and written assent was obtained from the participants themselves at the start of each study session. Audio and video recordings from study sessions were stored on secured servers mandated as part of the University of Florida Institutional Review Board parameters. A US $25 Amazon e-gift card was provided to the participant's parent or guardian at the conclusion of the study session.

Study Equipment

Several pieces of commercial technology were used to measure and display HR information (Figure 1). HR was measured using a sport-style chest-worn HR sensor (Garmin HRM-Pro Plus), and HR information was transmitted via Bluetooth Low Energy and displayed on either a smartphone or smartwatch. The smartphone used was a Google Pixel 4A, selected due to its low cost and the ability of Android phones to sideload apps, allowing us to install the Connexx app on the smartphone. The smartwatch used was a Samsung Galaxy Watch 4, selected for its Android-based operating system, which allows the Connexx app to be modified from the smartphone version to be compatible with the Galaxy Watch 4.

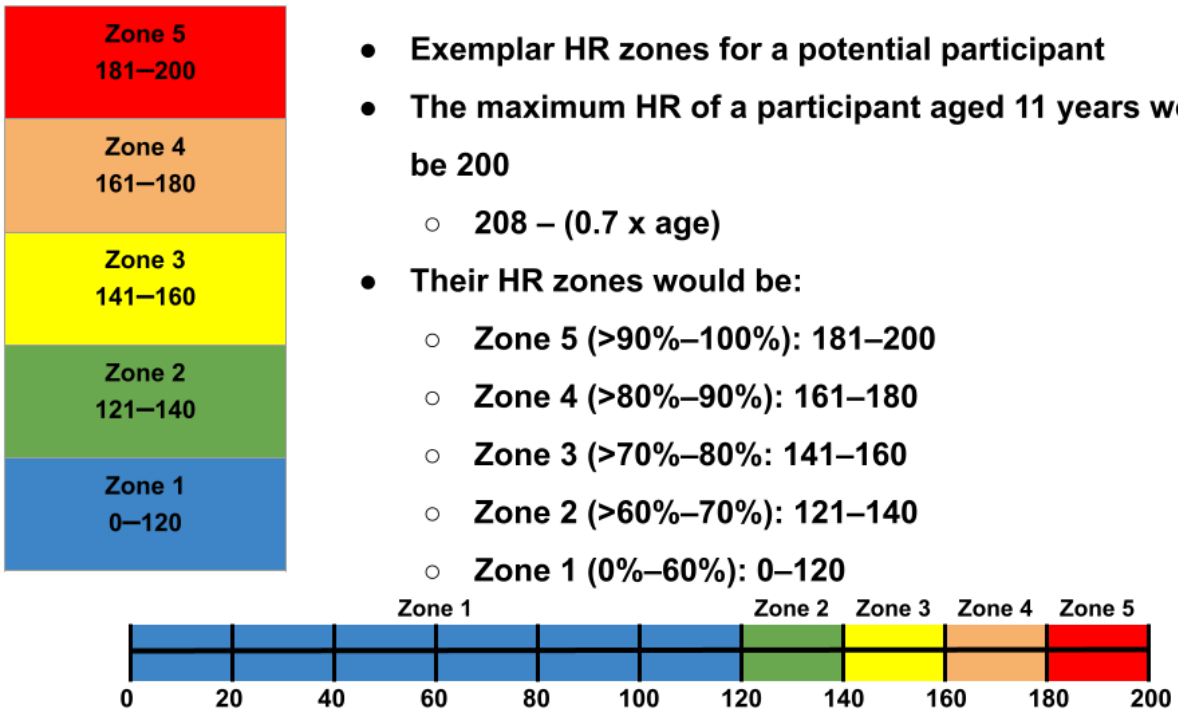
Figure 1. Smartwatch with solo workout screen connected to Bluetooth heart rate sensor. bpm: beats per minute; BLE: Bluetooth Low Energy.



This study used HR rather than accelerometry readings to measure MVPA. HR was chosen because it can be used to display effort in real time across various types of activities, which can contribute to a more accurate measure of the amount of time spent in MVPA [16]. We developed the Connexx app to display real-time HR information between two users. The Connexx app transmits real-time HR information between two devices via a traditional Bluetooth connection. However, Connexx is also capable of operating in solo mode, displaying the real-time HR of a single user, as we did in this study (Figure 1). Information on the user's current speed, distance covered during the workout, and workout duration was displayed along the top row of the Connexx solo workout screen. Speed and distance fields can be toggled between metric and imperial units by tapping on the display. The middle row of the workout screen displays the user's real-time HR. Connexx is able to display real-time HR information as either beats per minute (BPM) or as a percentage of maximum HR; users can toggle between these two display options by tapping on the HR display. A color-coded HR zone indicator is displayed below the HR field. This color coding matched the colors on HR exemplar cards used to explain HR information to participants (Figure 2). In this study, maximum HR was calculated using an established formula that is accurate for preadolescent children: 208 – (0.7 × age) [21]. The Connexx app automatically calculates HR zones using a straightforward 5-zone division based on the percentage of maximum HR, selected because there was no clear consensus or definitive definition of specific HR zones for preadolescent children [22-24] (zone 1: 0%-60%, zone 2: >60%-70%, zone 3: >70%-80%, zone 4: >80%-90%, and zone 5: >90%-100%). Furthermore, the Connexx app has an analytics portal; after the app records a workout, the information is sent to this portal, capable of displaying data from the workout on a graph.

XSL•FO
RenderX

Figure 2. Exemplar heart rate (HR) zone explanation card for a participant aged 11 years.



Participants, Eligibility, Recruitment, and Screening

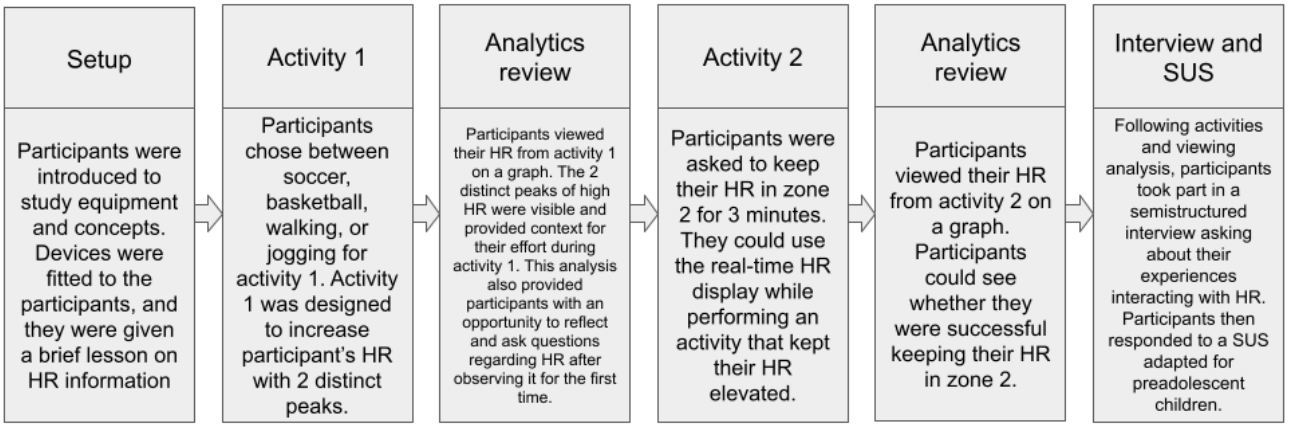
This study aimed to enroll a sample of 12 diverse preadolescent children to complete the full protocol. This sample size was chosen based on best practices in usability testing, showing that a sample of 10 to 12 participants is able to identify most major user issues [25]. Children were eligible to participate if they were aged 9 to 12 years and in good health with no physical conditions, injuries, or other conditions precluding them from engaging in MVPA for up to 30 minutes (assessed via a modified Physical Activity Readiness questionnaire for everyone [26]). We sought an equal number of male and female participants and equal distribution between Black, Hispanic, and non-Hispanic White participants (the three largest racial or ethnic groups in the local area).

Recruitment targeted parents of potentially eligible children using an electronic flyer distributed via local public elementary and middle schools, social media posts, and word of mouth. The flyer provided basic information regarding the study, eligibility criteria, compensation (US $25 Amazon e-gift card), and a link or QR code to a Qualtrics screening questionnaire. Parents or guardians completed the eligibility questionnaire, which included demographic information about their child. If the child was eligible, the parent or guardian was contacted via email to sign an electronic consent form through REDCap (Research Electronic Data Capture; Vanderbilt University). Once consent was obtained, a study team member (LL) scheduled the child for a study session.

Study Session

Study sessions were conducted at an indoor multiuse facility at the University of Florida. The space measures approximately 1021 square meters (27.7 meters by 36.9 meters) and can be used for a variety of uses, including basketball and volleyball; it has a maximum capacity of 1140 people. The study session began with obtaining assent for participation from the preadolescent child and followed the format outlined in Figure 3.

Figure 3. Study session activities flowchart. HR: heart rate; SUS: System Usability Scale.



XSL•FO
RenderX

Device Setup

Following assent, participants were fitted with Rode wireless lapel microphones. A researcher (LL) explained the study equipment and demonstrated the functionality of the chest-worn HR sensor. Parents aided participants in putting on the HR sensor. The session was recorded with a video camera. Next, participants chose between a smartphone and a smartwatch to view their real-time HR information. The Connexx app was available to all participants on either a smartphone or a smartwatch. However, all participants chose to use the smartwatch. The participants were guided through the process of entering their study information, such as their research identification number, age, and maximum HR, into the Connexx app. A workout was then initiated on the app, allowing participants to view their HR in real time, with more HR information provided to them.

HR Information

HR and HR zones were explained to the participants using a real-time HR display in Connexx (Figure 1) and an exemplar HR information card (Figure 2) based on their age (Table 1). Additional language used to explain HR and HR zones is provided in Multimedia Appendix 1. The goal of the brief explanation was to frame the upcoming activities in a way that participants could understand in relation to their HR zones. The researcher encouraged the participants to ask questions during the HR explanation.

Table 1. Example of the 5-zone heart rate (HR) categorization and description for participants aged 11 years.

HR zones	BPM[a]	Zone effort category	Example activities associated with the zone	Description of the zone
Zone 5	181 to 200	Maximum	Very high effort activities, which are often repeated (eg, sprinting in games such as tag multiple times in a row)	Extremely difficult to hold a conversation and most people must wait until after this effort to speak
Zone 4	161 to 180	Vigorous	Vigorous activities, such as playing a sport (eg, soccer or basketball), strenuous hiking, running, or climbing flights of stairs continually at a brisk pace	Difficult to hold conversation and often can only squeeze in short 1- or 2-word responses around breathing
Zone 3	141 to 160	Endurance	Shorter vigorous activities or long endurance activities (eg, climbing 2 to 3 flights of stairs or walking or hiking for an hour)	Slight difficulty to hold a normal conversation and speech is negotiated around breathing
Zone 2	121 to 140	Low endurance	Watching an exciting program and slow walk for short distances (eg, walking to the kitchen and walking around the block at a slow pace)	Can hold normal conversations without thinking about breathing
Zone 1	≤120	Rest	Sleeping, sitting and reading a book, or watching television	Can hold normal conversations without thinking about breathing

[a]BPM: beats per minute.

Activities

Next, we outlined the structure of the study to participants (Figure 3). The study consisted of 2 activity intervals; activity 1 lasted for 10 minutes and activity 2 lasted for 3 minutes, each followed by a brief rest period during which participants could view their HR information via an analytics portal on a laptop.

Activity 1

This study primarily aimed to explore whether preadolescent children can use real-time HR information to target specific MVPA goals. Activity 1 was designed to provide participants with an opportunity to observe their HR as it elevated through several different zones aligning with MVPA, providing practical, experiential use of HR information for participants. For activity 1, participants chose from playing basketball or soccer, or walking or jogging around the gym space. A researcher moderated participant effort during activity 1 by speeding up or slowing down the activity to take participants through a range of HR values. During this activity, the researcher asked participants to report their real-time HR at regular intervals. Querying for HR regularly was intended to help participants notice how their HR responded to various activity levels. Activity 1 was designed to increase participants' HR through multiple HR zones twice. Participants viewed their HR through the Connexx app while engaged in PA. After completing activity 1, the researcher guided participants through ending the workout on the Connexx app, and then, participants viewed a graphical representation of their HR in a Connexx analytics portal. We guided participants through the HR observation process, drawing attention to how HR can vary at rest, how HR responds to effort levels during PA, and how the body feels at various HRs.

Analytics Reflection

Once a workout ended on the Connexx app, workout data were uploaded to the analytics portal. The analytics portal displayed the participant's color-coded HR zones in the background of the graph, with HR in BPM on the vertical axis and duration on the horizontal axis (Figure 4). The researcher pointed to the connection between the HR graph and PA level: the more intense the PA level, the steeper the HR graph line became. Furthermore, the researcher pointed to how variable HR was: when at rest, HR did not stay at a consistent BPM; rather, it varied up and down slightly. In addition to the HR timeline graph, the analytics portal also presents an HR zone breakdown, which is a bar graph showing the total cumulative time the participants' HR spent in each zone (Figure 5). Participants were encouraged to ask questions throughout the process of viewing their HR graph.

XSL•FO
RenderX

Figure 4. Sample heart rate graph following activity 1. bpm: beats per minute.

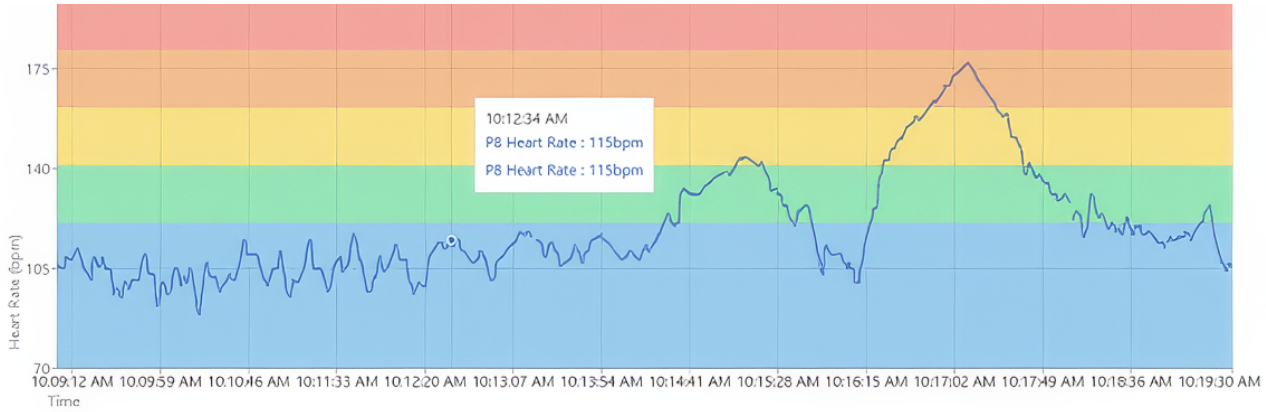
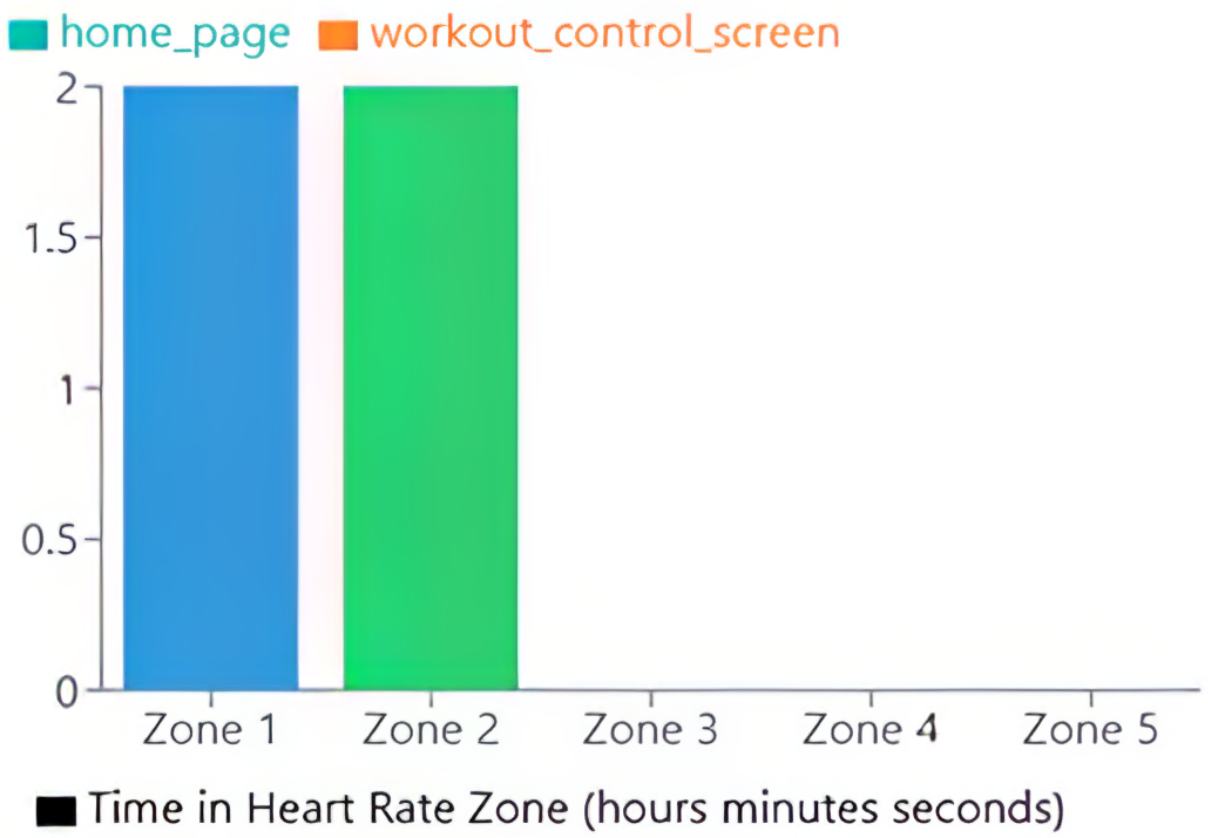


Figure 5. Sample heart rate zone cumulative time.



Activity 2

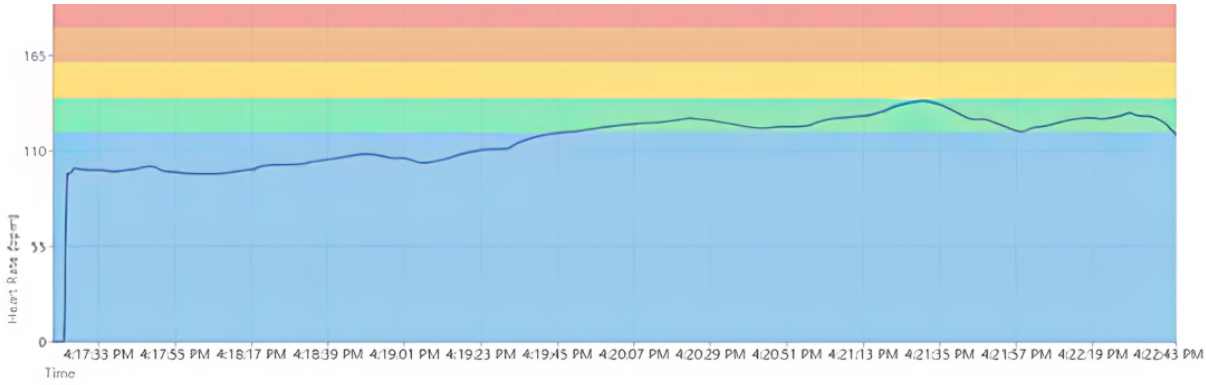
Activity 2 was designed to determine whether participants could make use of the tools provided to them so far (such as the information they have learned about HR, the HR sensor they were wearing, and the Connexx display, which provided a real-time display of their HR in BPM and HR zones) to maintain their HR in a targeted zone. Participants were asked to maintain their HR in zone 2 while engaged in the PA of their choice (basketball, soccer, walking, or jogging). Zone 2 was chosen because it was accessible to all participants, regardless of their fitness level, and they had to use the tools provided to maintain zone 2 HR, as they could easily exceed zone 2 HR by exercising too intensely. Participants could use the Connexx display to monitor their own HR and HR zone in real time to maintain zone 2 HR. The only prompts we provided were for timing (eg, when each minute elapsed). After 3 minutes, the participant was invited to save and end their workout on the Connexx app and view their HR graph for the second interval.

Analytics Review

The analytics portal displayed whether participants were successful in maintaining their HR in zone 2 (Figure 6). The HR zone breakdown also indicated the total time participants spent in each HR zone.

XSL•FO
RenderX

Figure 6. Sample heart rate graphical displays following activity 2. bpm: beats per minute.



Interview and System Usability Scale

After completing the 2 activities, participants engaged in a semistructured interview about their experiences. The interview explored participants' preference for smartwatch or smartphone, the comfort of commercial chest-worn HR sensors and smart devices, the participant's real-time actions to adjust their HR to meet a target goal, what HR zones mean to them, their preference for HR zones, whether they had previous experience with HR, and their enjoyment and potential future use of the Connexx app (refer to Multimedia Appendix 2 for the full interview script). During the interview, relevant objects and visual aids were presented to the participants as needed (eg, smartwatches and HR zone cards).

Finally, after the interview, participants filled out a modified version of the System Usability Scale (SUS) specifically modified and validated for use with preadolescent children [27]. This 10-item scale is used to calculate a usability score between 0 and 100. The commonly accepted threshold for a usable technology is a score of 68 [28].

Observational Data Collection Measures

We conducted observational data collection during each study session. In the event of recording capture device failure, a physical copy of the start and end times of each activity and a log of the participant's HR report and its corresponding value during the study session provide an overview of the participant's HR. Fortunately, this backup was not necessary during analysis.

The researcher who primarily interacted with the participants recorded field notes after each session. Notes focused on participant interaction with the instruments that might not be captured through the interview or SUS questionnaire, such as whether additional adjustments to the chest-worn HR sensor were required and how quickly participants adapted to the input methods for the Connexx app.

Analysis

Demographic data were summarized with descriptive statistics. Observational data were collated with the session transcripts to provide more depth and nuance to the thematic analysis. Additional observations that were not captured in the transcripts or SUS results are included as separate findings in the Observational Findings section.

Interviews were recorded and automatically transcribed via Microsoft OneDrive; we reviewed and revised autogenerated transcripts. Three researchers (LL, MA, and HAL) with qualitative research experience performed an inductive thematic analysis [29]. They followed the 6 steps of thematic analysis: data familiarization, initial code generation, searching for themes, theme review, defining and naming themes, and reporting [30,31]. While this study was designed with several objectives, we remained open to accepting any informative findings in the data. LL, MA, and HAL collaborated to familiarize themselves with the data and generate codes using a sample of the first 2 participant transcripts. The researchers worked together to search and identify themes, creating a codebook. The team then coded transcripts independently, meeting after completing the third and fourth participant transcripts to refine the codebook. Then, they coded the rest of the participant transcripts independently. Intercoder reliability was calculated using Krippendorff α [32] on a sample of 4 independently coded transcripts across 3 coders using the SPSS software (IBM Corp) extension developed by Hayes and Krippendorff [33]. Krippendorff α reliability estimate for the 3 coders was 0.907 (95% CI 0.8348-0.9621), corresponding to a high degree of agreement among them.

Results

Participant Demographics

A total of 11 preadolescent children with an average age of 10.5 (SD 1.0) years participated in this study, of whom 6 (55%) were male and 9 (82%) were non-Hispanic White. While the team sought to recruit a sample with equal numbers of male and female children and an even distribution of races and ethnicities, this diversity was not achieved in the final sample.

Observational Findings

Several observations demonstrated that preadolescent children can use real-time HR information to target specific MVPA goals. Participants shared a strong interest in their real-time HR information; some participants were quick to explore how their HR would respond to changes in PA. Participants looked at their HR information multiple times after initially starting a workout in the Connexx app and observed how controlling their breathing or fidgeting in their chairs may impact their HR. Their ability to use real-time HR information was further demonstrated by the actions of participants in activity 2. For activity 2,

XSL•FO
RenderX

participants were asked to maintain their HR in zone 2 for 3 minutes of PA. A total of 7 participants successfully used the Connexx display to hold their HR in zone 2. Of the remaining 4 participants, 1 (25%) initially ran at a high intensity that raised her HR above zone 2; however, throughout the rest of the 3-minute activity, she ran progressively slower to bring her HR back into zone 2. The 3 (75%) other participants took the opportunity to exert themselves as hard as possible in an attempt to achieve the highest HR reading instead of holding their HR in zone 2. These participants knew how to use the Connexx display to observe their HR in real time, as they looked at their smartwatch throughout activity 2. However, these 3 participants had a strong desire to see how high their HR could go. The parent of 1 of these 3 participants noted that she was highly competitive and always wanted to outcompete her older brother (who was also a participant in this study). One of the participants who successfully held his HR in zone 2 during the second activity demonstrated a similar desire to achieve a high HR value when he asked to be allowed to run around the gym space after completing the study so that he could see how fast his HR could read in the Connexx app.

Regarding hardware suitability for the children, all 11 participants chose to use a smartwatch to observe their HR in real time, and the watch straps successfully fit across a range of wrist sizes. However, there were some challenges with the chest-strap HR sensor. Two (18%) of the 11 participants needed the strap on the HR sensor to be shortened beyond its normal range to fit securely around their bodies. In both cases, the strap was shortened by <5 cm (by folding the excess fabric over and securing it with a safety pin) from the lower limit of the original strap range.

Furthermore, participants were able to quickly learn how to use the Connexx app. All participants were able to report their real-time HR and HR zone when prompted. Several participants (7/11, 64%) were also able to navigate the Connexx app themselves after the initial setup. They could swipe through the different screens in the app and start and end workouts themselves, and 2 (18%) out of 11 participants were able to navigate the text input interface with minimal guidance from the researcher. However, participants identified a substantial issue with the Connexx app in its current state. During the activities, participants noted that the font size of their HR display was too small. This was of particular concern during activity 1, as participants were moving vigorously at points to raise their HR through multiple HR zones. Participants had to adopt several different strategies to be able to see their HR values, including bringing the watch close to their face or slowing down their movement. The small font size of the Connexx app contributed to interruptions in the MVPA participants were engaged in.

Interview Findings

The interview provided rich data that generated themes related to using commercial devices to observe real-time HR to target specific MVPA goals. Three major themes emerged: one focused on real-time use of HR, one related to HR knowledge, and another on device and app use. Specific subthemes ranged from reports of real-time HR to preferences for the construction of HR zones. Themes, subthemes, definitions, and sample quotes are provided in Table 2. Major subthemes are described in the following sections in detail.

XSL•FO
RenderX

Table 2. Interview themes and sample quotes.

Theme and subtheme	Subtheme definition	Sample quotes
Real-time sensemaking of HR[a]		
Reporting HR and HR zones	Declarations of participant's real-time HR and HR zone	"Oh, I'm at Zone 4!... (which is) 160 (BPM)." [Participant 2]
Real-time HR use	How participants make use of real-time HR during activities	"I just looked at the watch a couple of times throughout the exercise and I saw Zone 2 and I'm like, OK, I'm, I was and I, you know, thought that I was doing good. So, I just kept it at that pace." [Participant 9]
HR and correlation to body sensations	How participant's body feels at different HR and HR zones	"I feel like I have to breathe more, but it mostly feels similar." [Participant 1]
HR-related knowledge		
Confusion regarding HR	Confusion about HR or HR zones	"I was just walking up and going over there to play soccer and it (HR) was (already) at 117." [Participant 4]
HR as an achievement	Desire for a high HR or HR zone displayed	"... with five, it's, I just like feel kind of accomplished." [Participant 8]
Previous knowledge of HR	Previous experience learning about HR	"In elementary school in PE class... count... how many beats we have in six seconds? And then he would tell us to, like, multiply it by 10 and that's beats per minute" [Participant 4]
HR zone preferences	Preference for HR zones, 3 larger or 5 smaller zones	"I personally like breaking stuff down to like smaller pieces and understanding it better." [Participant 9]
Device and app		
Device user experience	Comments on function and design of the device	"The smartwatch, you could just turn it over to look..." [Participant 7]
Device comfort or discomfort	How the devices felt for participants	"Wasn't too tight, too heavy... it feels comfortable." [Participant 3]
App usability	Like or dislike the app and willingness to use (or not use) the app in the future	"Probably just for like for fun, just to see (my HR)." [Participant 7]

[a]HR: heart rate.

Using Real-Time HR Information

During the interview, participants were asked about their strategies for adjusting their efforts to meet specific MVPA goals, including holding their HR in zone 2 for 3 minutes in activity 2. Participants who successfully held their HR in zone 2 discussed how they adjusted their effort based on the HR display (coded as "Real-time HR use" in Table 2):

> *I tried to keep the same speed when it was in Zone 2, because if I got faster, it [referring to his HR] will go higher, and if I went slower, it will go to Zone 1.* [Participant 3]

Often, this involved reducing their effort to not exceed the upper limit of zone 2 (coded as "HR and correlation to body sensations" in Table 2):

> *I tried using less power for [basketball] shots and like, try and walk or running across slower to conserve [energy].* [Participant 1]

The participants who were successful in keeping their HR steady in zone 2 used their smartwatch to adjust their effort much more often:

> *I just kept looking at the watch.* [Participant 2]

Some participants (2/11, 18%) noticed that their HR would change even when they did not think their actions should correspond to the change (coded as "Confusion regarding HR" in Table 2):

> *I was just walking up and going over there to play soccer and it [HR] was [already] at 117.* [Participant 4]

However, noting this unexpected result demonstrated that the participant was actively making connections between their PA and the displayed HR.

Not all participants were able to successfully keep their HR in zone 2 during activity 2. These participants acknowledged that they knew the goal was to maintain their HR in zone 2, but they wanted to achieve the highest HR possible. When asked what their HR was, participant 1 stated that "it just hit Zone 4, it is 161... I can go way faster than that!" (coded as "HR as achievement" in Table 2), and participant 8 expressed interest in his HR throughout the course of the study session. After finishing activity 1, he told the researcher that he "would like to see what it (his HR) looks like on a graph." Furthermore, after activity 2, where he was not successful in keeping his HR in zone 2, he told the researcher that he did "want to see how long I was in [zone] 5 for."

XSL•FO
RenderX

HR-Related Knowledge

In addition to findings on the specific research objectives, several themes emerged during the inductive thematic analysis process.

Previous Experience With HR

Four (36%) of the 11 participants had learned about HR in school before taking part in the study (coded as "Prior knowledge of HR" in Table 2); however, none of the participants had a complete understanding of HR. Learning about HR in school was summed up by a participant as follows:

> *[We would] do jumping jacks or jog in place for like 30 seconds, 20 seconds, and then... count our heart, how many beats we have in six seconds, and then he would tell us to, like multiply it by 10, and that's beats per minute.* [Participant 4]

None of the participants had used a chest-worn HR sensor before the study.

HR Zones

Participants were asked to describe what HR zones meant in their own words, and overall, they were able to provide a basic description of each zone. A participant explained the zones in a manner similar to how we had explained them before activity 1, as follows:

> *Zone 1, just calm, sitting down and not doing anything. Zone 2 I might be walking around, and I don't know. Zone 3 I would start to feel really hard to breathe and a little faster [breathing]. Zone 4, I might be breathing kind of heavier and having to slow down a little bit.* [Participant 4]

Other participants described the zones in relation to their own experiences during the study:

> *[Zone] 4 was like, I got to 4 when I was like shooting the ball and running.* [Participant 7]

Participants were asked whether they had a preference for the 5 HR zones used in this study or a more simplified 3 HR zones (coded as "HR zone preferences" in Table 2). Three (27%) of the 11 participants preferred the simplicity of having fewer zones:

> *It would be easier to look at if there was just easy, medium, hard. It would be easier to keep it in the easy-medium zone [referring to Zone 2].* [Participant 4]

The other participants (8/11, 73%) preferred 5 HR zones. Reasons for preferring 5 zones differed between participants; some preferred the greater granularity offered by more zones, and others preferred more HR zones as it provided a sense of achievement. A participant explained the meaning of having more HR zones as follows:

> *It would narrow it down and be easier just to keep it at that zone.* [Participant 6]

Another participant explained his preference for more HR zones as follows:

> *Oh, it's just like an achievement. It kind of feels like an achievement.* [Participant 7]
>
> *I just want to know if I'm in Zone 5, because, like, it's exciting.* [Participant 7]

Feasibility of Chest-Worn HR Sensors Paired to Smart Devices

Device Preference

To explore device preferences, participants were asked why they chose to use a smartwatch over a smartphone to observe their HR during PA (which all participants did). While only 2 (18%) of the 11 participants had consistently used a smartwatch before the study (one had their own, and another used their father's), all participants commented on the greater convenience they could foresee with a smartwatch, particularly while engaging in PA (coded as "Device user experience" in Table 2):

> *The smartwatch, you could just turn it over to look. But for a smartphone, you have to grab it, then you have to [mimes the action of pulling a phone out of their pocket to look at it]. If you're like working out, and you want to know [your HR] you could just look.* [Participant 7]

Device Comfort

When asked about the comfort or discomfort of the devices, all participants stated that the smartwatch felt comfortable (coded as "Device comfort or discomfort" in Table 2):

> *I'm used to wearing watches, so it's pretty normal for me to put one [smartwatch] on.* [Participant 1]

While there was some initial discomfort with the chest-worn HR sensor, it soon became unnoticeable:

> *When I first put it on, it kind of feels weird. But I got used to it when I used it for quite a while.* [Participant 3]

Only one participant reported continued discomfort and a desire to take off the HR sensor before the conclusion of the interview:

> *It was hot and kind of squeezing, so you get uncomfortable feeling.* [Participant 7]

Furthermore, one participant noted that using a softer fabric for the strap would increase the comfort of the instrument:

> *If like the material is different, so you didn't feel it as much. Like if it's like a smoother material or something maybe not as thick.* [Participant 6]

Usability of the Connexx App

Overall, participants were enthusiastic about the Connexx app and the ability to observe their own HR data (coded as "App usability" in Table 2). Participant 7 represented most participants when he declared that using Connexx "was kind of fun!" However, the small font size on the Connexx display made it challenging to view real-time HR while engaged in PA. For instance, participant 1 stated that what he "found frustrating was when we were running that when I wanted to check my HR, I couldn't see it straight because of it shaking." Many

XSL•FO
RenderX

participants (9/11, 82%) expressed interest in using the Connexx app in the future. The final response of participant 11 to a series of questions regarding the Connexx app and the ability to view her own HR was that she could "use it for playing games" with her friends. Similarly, participant 7 noted that he would like to use the Connexx app again in the future, "probably just for like for fun."

SUS Findings

The Connexx app can be considered highly usable in its current iteration; participants rated the Connexx app with an average score of 77.1 (SD 13.7), well above the threshold score of 68 [28]. However, feedback from this study will contribute to its future refinement to further increase usability.

Discussion

Principal Findings

While significant research has demonstrated the accuracy of commercially available HR devices for children [15], there has been insufficient research on how real-time HR information can be used to target MVPA goals. In this study, preadolescent children were provided with a basic lesson about their HR and HR zones, and then, they engaged in 2 PA sessions. Participants demonstrated an understanding of the HR display during PA by effectively using it to regulate their efforts. Therefore, real-time HR information can be used by children as a tool to meet PA guidelines more effectively; children can monitor their own HR information to ensure that they are engaged in MVPA. In addition, this study provides insight into the feasibility of HR sensors and smart devices, smart device preferences, the usability of a custom HR app, preadolescent children's previous HR experiences, and HR zone preferences.

Preadolescent Children Using Real-Time HR Information to Target Specific MVPA Goals

The findings of this study are unique in research on HR device use. Previous research often focused exclusively on accuracy [34], overlooked how users interact with commercial fitness trackers [35], or did not include HR information [36] in the design. In fact, preadolescent children in one study complained that the only real-time information they could see were step count and distance [36]. Participants in a study that paired chest-worn HR sensors with a wrist-worn display regulated their effort based on running speed; these participants did not use their available real-time HR information to adjust their effort [16]. Participants in this study demonstrated their ability to use real-time HR and HR zone information during PA by successfully keeping their HR in zone 2. We observed participants adjusting their effort while exercising (during activity 2) to keep their HR in zone 2. Participants confirmed in the interview that they adjusted their movement speed based on their real-time HR display, demonstrating an ability to use their real-time HR information to target specific MVPA goals during the study session.

Feasibility of Chest-Worn HR Sensors Paired to Smart Devices

Results of this study suggest that chest-worn HR sensors are feasible for use by preadolescent children. While some participants (2/11, 18%) mentioned slight discomfort when first putting on the chest-worn HR sensor, this did not seem to be an ongoing issue and was not brought up during activity 1 or 2. This study provides a more comprehensive investigation into the comfort and usability of these devices compared to previous research. One study that examined the use of a chest-strap HR sensor in conjunction with a wrist-worn device for preadolescent children focused on the accuracy of measurement rather than how preadolescent children felt about using these devices [14]. Another study focused on the accuracy of commercial devices for HR measurements, in this case, wrist-worn devices, without consideration for the comfort of the device [15]. Participants in this study were undergoing surgery, which may not be an appropriate situation to consider the comfort of an HR monitoring device. Furthermore, previous research that focused on the comfort of devices [16] examined an older population, whose more developed bodies may be better suited for devices typically designed for adults. However, the findings of this study indicate that chest-worn HR sensors and smartwatches were generally comfortable for preadolescent children.

Preference for Smartwatch

An interesting finding of this study on the feasibility of smart devices is the appeal of a smartwatch over a smartphone. Researchers did not anticipate all participants to choose a smartwatch over a smartphone for observing their HR information during the study. This is an issue that was not discussed in previous literature [37,38]. Upon reflection, this preference aligns with current commercial trends. Most preadolescent children (9/11, 82%) who participated in this study did not have previous experience with smartwatches or wristwatches in general, which may have contributed to a novelty effect. However, during the interview, they identified the convenience of using a wrist-worn device to view their HR in real time. Several participants (6/11, 54%) commented on perceived ease of use, where they would only have to move their wrists up to be able to view their HR rather than pulling a phone out of a pocket. Other participants (5/11, 45%) favored the security of a worn device, noting that they do not have to worry about dropping a watch compared to a phone. Furthermore, some participants (2/11, 18%) noted that watches may be more comfortable during PA, as a phone in their pocket would move undesirably.

Usability

Research on the design of mobile apps and the use of commercial wearable sensors has focused on the design process; technical limitations, such as loading speed; and efforts to ensure hardware capabilities and accuracy [39,40]. We developed the Connexx app following best practices based on industry standards and previous research [40,41]. Participants rated the Connexx app with a SUS score of 77, which is above the threshold score of 68 [28], indicating that the app has above-average usability. In addition, the Connexx app received a higher score than other digital health apps, more in line with

XSL•FO
RenderX

other PA apps [42]. The usability of the Connexx app was demonstrated by participants throughout the study.

For most participants (9/11, 82%), this was their first time using a smartwatch, and for all participants, this was their first time observing their real-time HR. We helped participants with the initial setup, such as inputting their age and other information in the Connexx app. However, many participants (7/11, 64%) quickly figured out how to operate the Connexx app themselves, and they could input text, navigate the menus, and begin and end workouts on their own without our intervention. This ability not only demonstrated the Connexx app's usability but also this population's familiarity with modern technology and its use, which was a concern for developing health-related apps for preadolescent children [43].

Despite the high SUS score, there were issues with the Connexx app. The Connexx app was initially developed for smartphones and then adapted to smartwatches, creating unanticipated usability challenges. The main PA screen for Connexx featured live tracking of several data fields, including HR, activity time, and speed. While this was easy to read on a smartphone display, on a smaller smartwatch display, the presence of multiple data fields resulted in each field being too small to be easily legible during PA. Thus, participants used multiple strategies to stabilize the smartwatch to be able to read what their HR was. The inclusion of a large color-coded portion at the bottom of the smartwatch display was helpful in informing participants of their HR zone at a glance. Despite difficulty reading portions of the app display, participants found it easy and fun to use. Furthermore, some participants (2/11, 18%) indicated that they not only wanted to use the app again in the future but also would like to share it with their friends.

Other HR Findings

Additional findings from this study indicated that an introductory lesson on HR helped provide participants with a basic understanding of their HR and demonstrated the feasibility of using HR during MVPA. Some participants (2/11, 18%) viewed high HR as an achievement, and the preferences for HR zones are linked to HR achievements for some participants.

Many participants (7/11, 64%) did not have previous experience with HR. An introductory lesson on HR and HR zones was provided for participants at the beginning of the study session, using a simple description of HR zones, including color-coded visual aids (Figure 2) and multiple descriptors that demonstrated how the participants' bodies might respond to various HR zones (Table 1). Participants subsequently demonstrated an understanding of their HR information while engaged in PA, responding to our queries about their HR with reports of both their HR and the corresponding HR zones. Furthermore, participants recalled the introductory lesson on HR well enough after engaging in PA to describe HR zones during the interview; some repeated the descriptors from the lesson, while others applied the descriptors to their own experiences. These findings suggest that participants were able to quickly grasp the HR concepts taught during the session and that using HR to target MVPA with preadolescent children is feasible.

This understanding of HR is even demonstrated by participants who were unable to keep their HR in zone 2 during activity 2. Participants who did not keep their HR in zone 2 demonstrated their understanding of HR by fully exerting themselves in an attempt to achieve a high HR reading. These participants were more interested in hitting the highest HR and HR zone possible; they viewed higher HR and HR zone as an accomplishment. While participants used the technology in a way not originally intended by the study's design, it still showed their understanding of HR and its application during PA.

Finally, many participants (8/11, 73%) preferred having a more granular HR zone division, favoring the 5-zone model over the 3-zone option. Some participants (2/11, 18%) equated zone 5 with a higher sense of accomplishment, mirroring the drive to achieve the highest HR possible. The drive for higher values is reflected in commercially available products for PA tracking. Platforms such as Strava [44] reward users by keeping track of their personal records. Furthermore, when a user's best performance is faster than any other user for a particular segment, the fastest user is rewarded with a crown icon and the title "King of the Mountain" or "Queen of the Mountain." The desire for high achievement of these participants is likely a reason why gamification and leaderboards are more effective in PA promotion than health benefits [45].

Limitations

Despite these findings, this study is not without its limitations. First, the participants had never encountered an app such as Connexx before; therefore, the novelty factor of a new measurement offered by the Connexx app may have offset potential shortcomings arising from font size or other use limitations [46]. As participants continue to use the Connexx app, the SUS score may decrease, as the novelty of displaying real-time HR information wears off. However, as participants indicated having little previous experience or education with HR, this highlights an educational opportunity missing from school curricula, as the participants were interested in learning about their HR.

In addition, the small sample size of this feasibility study limits the generalizability of the findings. While there were almost equal numbers of male and female participants, there was little racial or ethnic diversity (9/11, 82% non-Hispanic White). However, the sample size was still sufficient to demonstrate the usability of these HR monitoring technologies with preadolescent children [25]. Future research is needed to explore the technologies with a larger, more diverse sample.

Because this was a feasibility study to explore whether adolescent children could understand and use their HR information in real time as opposed to measuring HR change in response to PA, we did not control for other confounding variables (eg, the child's fitness level, diet, and resting HR). Rather, as a first step, this study simply examined whether HR use was possible. Moreover, due to the study design, this study did not take advantage of the full scope of capabilities developed as part of the Connexx app, including displaying a paired partner's HR information in real time. Future studies will further explore these capabilities of the Connexx app.

XSL•FO
RenderX

Conclusions

This study aimed to investigate the feasibility and usability of multiple technologies to measure and display HR information for preadolescent children. It examined whether preadolescent children could effectively use these devices to view their HR and make practical use of the information. The results indicate that preadolescent children understood and used their HR to regulate their MVPA and enjoyed using Connexx for real-time HR observation. Participant responses demonstrate the ability to use and the desire for access to real-time HR information during PA. Future research should build on these findings regarding the use of real-time HR as an intervention tool to target specific MVPA goals. This could involve testing strategies that use HR and real-time monitoring as direct targets in program designs. Preadolescent children's interest in using real-time HR should be leveraged by researchers, health professionals, and educators to promote MVPA in an engaging way for this population.

Acknowledgments

The authors would like to thank Sunny Dhama for assisting with software and investigation, Lidong Lin for assisting with the investigation, and Wesly Menard and Mollie Brewer for assisting with the conceptualization and methodology of the study. The funding was provided by the University of Florida College of Health and Human Performance and the University of Florida College of Engineering.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Explanation of heart rate zones.
[DOCX File , 14 KB - humanfactors_v12i1e58715_app1.docx]

Multimedia Appendix 2
Semistructured interview script.
[DOCX File , 10 KB - humanfactors_v12i1e58715_app2.docx]

References

1. Katzmarzyk PT. Expanding our understanding of the global impact of physical inactivity. Lancet Glob Health 2023 Jan;11(1):e2-e3. [doi: 10.1016/s2214-109x(22)00482-x]
2. Katzmarzyk PT, Denstel KD, Beals K, Bolling C, Wright C, Crouter SE, et al. Results from the United States of America's 2016 report card on physical activity for children and youth. J Phys Act Health 2016 Nov;13(11 Suppl 2):S307-S313 [FREE Full text] [doi: 10.1123/jpah.2016-0321] [Medline: 27848726]
3. Guthold R, Stevens GA, Riley LM, Bull FC. Global trends in insufficient physical activity among adolescents: a pooled analysis of 298 population-based surveys with 1·6 million participants. Lancet Child Adolesc Health 2020 Jan;4(1):23-35. [doi: 10.1016/s2352-4642(19)30323-2]
4. Finer N. Medical consequences of obesity. Medicine 2015 Feb;43(2):88-93. [doi: 10.1016/j.mpmed.2014.11.003]
5. Griffiths LJ, Parsons TJ, Hill AJ. Self-esteem and quality of life in obese children and adolescents: a systematic review. Int J Pediatr Obes 2010 Aug;5(4):282-304. [doi: 10.3109/17477160903473697] [Medline: 20210677]
6. Sahoo K, Sahoo K, Choudhury AK, Sofi NY, Kumar R, Bhadoria AS. Childhood obesity: causes and consequences. J Family Med Prim Care 2015;4(2):187-192 [FREE Full text] [doi: 10.4103/2249-4863.154628] [Medline: 25949965]
7. Obesity. World Health Organization. URL: https://www.who.int/health-topics/obesity#tab=tab_1 [accessed 2025-01-27]
8. Wyszyńska J, Ring-Dimitriou S, Thivel D, Weghuber D, Hadjipanayis A, Grossman Z, et al. Physical activity in the prevention of childhood obesity: the position of the European Childhood Obesity Group and the European Academy of Pediatrics. Front Pediatr 2020 Nov 5;8:535705 [FREE Full text] [doi: 10.3389/fped.2020.535705] [Medline: 33224905]
9. Dimitri P, Joshi K, Jones N, Moving Medicine for Children Working Group. Moving more: physical activity and its positive effects on long term conditions in children and young people. Arch Dis Child 2020 Nov 20;105(11):1035-1040. [doi: 10.1136/archdischild-2019-318017] [Medline: 32198161]
10. Davison KK, Werder JL, Trost SG, Baker BL, Birch LL. Why are early maturing girls less active? Links between pubertal development, psychological well-being, and physical activity among girls at ages 11 and 13. Soc Sci Med 2007 Jun;64(12):2391-2404 [FREE Full text] [doi: 10.1016/j.socscimed.2007.02.033] [Medline: 17451855]
11. Dooley EE, Golaszewski NM, Bartholomew JB. Estimating accuracy at exercise intensities: a comparative study of self-monitoring heart rate and physical activity wearable devices. JMIR Mhealth Uhealth 2017 Mar 16;5(3):e34 [FREE Full text] [doi: 10.2196/mhealth.7043] [Medline: 28302596]

XSL•FO
RenderX

12. St Fleur RG, St George SM, Leite R, Kobayashi M, Agosto Y, Jake-Schoffman DE. Use of Fitbit devices in physical activity intervention studies across the life course: narrative review. JMIR Mhealth Uhealth 2021 May 28;9(5):e23411 [FREE Full text] [doi: 10.2196/23411] [Medline: 34047705]
13. Fuller D, Colwell E, Low J, Orychock K, Tobin MA, Simango B, et al. Reliability and validity of commercially available wearable devices for measuring steps, energy expenditure, and heart rate: systematic review. JMIR Mhealth Uhealth 2020 Sep 08;8(9):e18694 [FREE Full text] [doi: 10.2196/18694] [Medline: 32897239]
14. Brazendale K, Decker L, Hunt ET, Perry MW, Brazendale AB, Weaver RG, et al. Validity and wearability of consumer-based fitness trackers in free-living children. Int J Exerc Sci 2019;12(5):471-482 [FREE Full text] [doi: 10.70252/BBTU4889] [Medline: 30899354]
15. Pelizzo G, Guddo A, Puglisi A, De Silvestri A, Comparato C, Valenza M, et al. Accuracy of a wrist-worn heart rate sensing device during elective pediatric surgical procedures. Children (Basel) 2018 Mar 08;5(3):38 [FREE Full text] [doi: 10.3390/children5030038] [Medline: 29518020]
16. Sun J, Liu Y. Using smart bracelets to assess heart rate among students during physical education lessons: feasibility, reliability, and validity study. JMIR Mhealth Uhealth 2020 Aug 05;8(8):e17699 [FREE Full text] [doi: 10.2196/17699] [Medline: 32663136]
17. Sidoti O, Dawson W, Gelles-Watnick R, Faverio M, Atske S, Radde K, et al. Mobile fact sheet. Pew Research Center. 2024 Nov 13. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2025-01-22]
18. Broffman L, Harrison S, Zhao M, Goldman A, Patnaik I, Zhou M. The relationship between broadband speeds, device type, demographic characteristics, and care-seeking via telehealth. Telemed J E Health 2023 Mar 01;29(3):425-431. [doi: 10.1089/tmj.2022.0058] [Medline: 35867048]
19. Vogels EA. About one-in-five Americans use a smart watch or fitness tracker. Pew Research Center. 2020 Jan 9. URL: https://www.pewresearch.org/short-reads/2020/01/09/about-one-in-five-americans-use-a-smart-watch-or-fitness-tracker/ [accessed 2025-01-27]
20. Richter A, Adkins V, Selkie E. Youth perspectives on the recommended age of mobile phone adoption: survey study. JMIR Pediatr Parent 2022 Oct 31;5(4):e40704 [FREE Full text] [doi: 10.2196/40704] [Medline: 36315219]
21. Tanaka H, Monahan KD, Seals DR. Age-predicted maximal heart rate revisited. J Am Coll Cardiol 2001 Jan;37(1):153-156 [FREE Full text] [doi: 10.1016/s0735-1097(00)01054-8] [Medline: 11153730]
22. Goh DY, Mun MW, Lee WL, Teoh OH, Rajgor DD. A randomised clinical trial to evaluate the safety, fit, comfort of a novel N95 mask in children. Sci Rep 2019 Dec 12;9(1):18952 [FREE Full text] [doi: 10.1038/s41598-019-55451-w] [Medline: 31831801]
23. Lind RR, Beck MM, Wikman J, Malarski K, Krustrup P, Lundbye-Jensen J, et al. Acute high-intensity football games can improve children's inhibitory control and neurophysiological measures of attention. Scand J Med Sci Sports 2019 Oct 24;29(10):1546-1562 [FREE Full text] [doi: 10.1111/sms.13485] [Medline: 31125468]
24. Miklós M, Komáromy D, Futó J, Balázs J. Acute physical activity, executive function, and attention performance in children with attention-deficit hyperactivity disorder and typically developing children: an experimental study. Int J Environ Res Public Health 2020 Jun 07;17(11):4071 [FREE Full text] [doi: 10.3390/ijerph17114071] [Medline: 32517384]
25. Alroobaea R, Mayhew PJ. How many participants are really enough for usability studies? In: Proceedings of the Science and Information Conference. 2014 Presented at: SAI 2014; August 27-29, 2014; London, UK. [doi: 10.1109/sai.2014.6918171]
26. Warburton DE, Jamnik VK, Bredin SS, Gledhill N. The physical activity readiness questionnaire for everyone (PAR-Q+) and electronic physical activity readiness medical examination (ePARmed-X+). Health Fit J Can 2011;4(2):3-17. [doi: 10.14288/hfjc.v4i2.103]
27. Putnam C, Puthenmadom M, Cuerdo MA, Wang W, Paul N. Adoption of the system usability scale for user testing with children. In: Proceedings of the Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems. 2020 Presented at: CHI EA '20; April 25-30, 2020; Honolulu, HI. [doi: 10.1145/3334480.3382840]
28. Lewis JR, Sauro J. Item benchmarks for the system usability scale. J Usability Stud 2018;13(3):158-167.
29. Sandborg J, Henriksson P, Larsen E, Lindqvist AK, Rutberg S, Söderström E, et al. Participants' engagement and satisfaction with a smartphone app intended to support healthy weight gain, diet, and physical activity during pregnancy: qualitative study within the HealthyMoms trial. JMIR Mhealth Uhealth 2021 Mar 05;9(3):e26159 [FREE Full text] [doi: 10.2196/26159] [Medline: 33666554]
30. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
31. D'Addario M, Baretta D, Zanatta F, Greco A, Steca P. Engagement features in physical activity smartphone apps: focus group study with sedentary people. JMIR Mhealth Uhealth 2020 Nov 16;8(11):e20460 [FREE Full text] [doi: 10.2196/20460] [Medline: 33196450]
32. Krippendorff K. Content Analysis: An Introduction to Its Methodology. Thousand Oaks, CA: SAGE Publications; 2019.
33. Hayes AF, Krippendorff K. Answering the call for a standard reliability measure for coding data. Commun Methods Meas 2007 Apr;1(1):77-89. [doi: 10.1080/19312450709336664]

XSL•FO
RenderX

34. Martinko A, Karuc J, Jurić P, Podnar H, Sorić M. Accuracy and precision of consumer-grade wearable activity monitors for assessing time spent in sedentary behavior in children and adolescents: systematic review. JMIR Mhealth Uhealth 2022 Aug 09;10(8):e37547 [FREE Full text] [doi: 10.2196/37547] [Medline: 35943763]
35. Chevance G, Golaszewski NM, Tipton E, Hekler EB, Buman M, Welk GJ, et al. Accuracy and precision of energy expenditure, heart rate, and steps measured by combined-sensing Fitbits against reference measures: systematic review and meta-analysis. JMIR Mhealth Uhealth 2022 Apr 13;10(4):e35626 [FREE Full text] [doi: 10.2196/35626] [Medline: 35416777]
36. Mackintosh KA, Chappel SE, Salmon J, Timperio A, Ball K, Brown H, et al. Parental perspectives of a wearable activity tracker for children younger than 13 years: acceptability and usability study. JMIR Mhealth Uhealth 2019 Nov 04;7(11):e13858 [FREE Full text] [doi: 10.2196/13858] [Medline: 31682585]
37. Böhm B, Karwiese SD, Böhm H, Oberhoffer R. Effects of mobile health including wearable activity trackers to increase physical activity outcomes among healthy children and adolescents: systematic review. JMIR Mhealth Uhealth 2019 Apr 30;7(4):e8298 [FREE Full text] [doi: 10.2196/mhealth.8298] [Medline: 31038460]
38. Asgari Mehrabadi M, Azimi I, Sarhaddi F, Axelin A, Niela-Vilén H, Myllyntausta S, et al. Sleep tracking of a commercially available smart ring and smartwatch against medical-grade actigraphy in everyday settings: instrument validation study. JMIR Mhealth Uhealth 2020 Nov 02;8(10):e20465 [FREE Full text] [doi: 10.2196/20465] [Medline: 33038869]
39. Düking P, Fuss FK, Holmberg HC, Sperlich B. Recommendations for assessment of the reliability, sensitivity, and validity of data provided by wearable sensors designed for monitoring physical activity. JMIR Mhealth Uhealth 2018 Apr 30;6(4):e102 [FREE Full text] [doi: 10.2196/mhealth.9341] [Medline: 29712629]
40. Newton A, Bagnell A, Rosychuk R, Duguay J, Wozney L, Huguet A, et al. A mobile phone-based app for use during cognitive behavioral therapy for adolescents with anxiety (MindClimb): user-centered design and usability study. JMIR Mhealth Uhealth 2020 Dec 08;8(12):e18439 [FREE Full text] [doi: 10.2196/18439] [Medline: 33289671]
41. Wu H, Li G. Innovation and improvement of visual communication design of mobile app based on social network interaction interface design. Multimed Tools Appl 2019 Apr 13;79:1-16. [doi: 10.1007/s11042-019-7523-6]
42. Hyzy M, Bond R, Mulvenna M, Bai L, Dix A, Leigh S, et al. System usability scale benchmarking for digital health apps: meta-analysis. JMIR Mhealth Uhealth 2022 Aug 18;10(8):e37290 [FREE Full text] [doi: 10.2196/37290] [Medline: 35980732]
43. Schwarz A, Winkens LH, de Vet E, Ossendrijver D, Bouwsema K, Simons M. Design features associated with engagement in mobile health physical activity interventions among youth: systematic review of qualitative and quantitative studies. JMIR Mhealth Uhealth 2023 Mar 06;11:e40898 [FREE Full text] [doi: 10.2196/40898] [Medline: 36877551]
44. Community-powered motivation. STRAVA. URL: https://www.strava.com/ [accessed 2024-01-22]
45. Millonig A, Wunsch M, Stibe A, Seer S, Dai C, Schechtner K, et al. Gamification and social dynamics behind corporate cycling campaigns. Transp Res Proc 2016;19:33-39. [doi: 10.1016/j.trpro.2016.12.065]
46. Rutten I, Bogaert LV, Geerts D. From initial encounter with mid-air haptic feedback to repeated use: the role of the novelty effect in user experience. IEEE Trans Haptics 2021 Jul 1;14(3):591-602. [doi: 10.1109/toh.2020.3043658]

Abbreviations

BPM: beats per minute
HR: heart rate
MVPA: moderate- to vigorous- physical activity
PA: physical activity
REDCap: Research Electronic Data Capture
SUS: System Usability Scale

Edited by A Kushniruk; submitted 22.03.24; peer-reviewed by H Namba, M Fraser; comments to author 01.12.24; revised version received 14.12.24; accepted 23.12.24; published 06.03.25.

Please cite as:
Lu L, Jake-Schoffman DE, Lavoie HA, Agharazidermani M, Boyer KE
Preadolescent Children Using Real-Time Heart Rate During Moderate to Vigorous Physical Activity: A Feasibility Study
JMIR Hum Factors 2025;12:e58715
URL: https://humanfactors.jmir.org/2025/1/e58715
doi:10.2196/58715
PMID:40053729

©Lincoln Lu, Danielle E Jake-Schoffman, Hannah A Lavoie, Maedeh Agharazidermani, Kristy Elizabeth Boyer. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 06.03.2025. This is an open-access article distributed under

XSL•FO
RenderX

the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

User-Centered Design of an Electronic Dashboard for Monitoring Facility-Level Basic Emergency Obstetric Care Readiness in Amhara, Ethiopia: Mixed Methods Study

Kylie Dougherty[1], RN, PhD; Yihenew Tesfaye[2], PhD; Heran Biza[3], BS, MS; Mulusew Belew[4], MSC; Natalie Benda[5], PhD; Abebe Gebremariam Gobezayehu[3,4], MD; John Cranmer[2,3,4,6,7], DNP, MPH, MSN, ANP, CPH, EBP-C; Suzanne Bakken[5], PhD

[1]Department of Medical Social Sciences, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

[2]College of Health Sciences, Bahir Dar University, Bahir Dar, Ethiopia

[3]Nell Hodgson Woodruff School of Nursing, Emory University, Atlanta, GA, United States

[4]Emory-Ethiopia Partnership, Bahir Dar, Ethiopia

[5]School of Nursing, Columbia University, New York, NY, United States

[6]Woodruff Health Sciences Center, Emory University, Atlanta, GA, United States

[7]Center for the Study of Human Health, Emory University, Atlanta, GA, United States

Corresponding Author:
Kylie Dougherty, RN, PhD
Department of Medical Social Sciences
Feinberg School of Medicine
Northwestern University
633 N St Clair St
Chicago, IL, 60611
United States
Phone: 1 (312) 503 1725
Email: kylie.dougherty@northwestern.edu

Abstract

Background: Maternal mortality remains a persistent public health concern in sub-Saharan African countries such as Ethiopia. Health information technology solutions are a flexible and low-cost method for improving health outcomes with proven benefits in low- to middle-income countries' health systems.

Objective: This study aimed to develop and assess the usability of an electronic dashboard to monitor facility-level readiness to manage basic emergency obstetric care (BEmOC) in Amhara, Ethiopia.

Methods: The study used three methods to iteratively refine the dashboard: (1) user-centered design sessions with individuals who interact with the BEmOC supply chain, (2) review and feedback from domain and information visualization subject matter experts (SMEs) to refine the dashboard, and (3) usability heuristic evaluation with human-computer interaction (HCI) SMEs.

Results: User-centered design sessions resulted in a preliminary version of the dashboard informed by end-user preferences and perceptions, with recommendations focusing on aesthetic design, filtering and sorting, and matching with the real world. An example of an end-user recommendation included increasing font sizes on the dashboard and using a red, yellow, and green color-coding scheme. Next, domain and visualization SMEs continued the dashboard's iterative refinement, focusing on aesthetic design and navigation, by confirming design choices incorporated from the user-centered design sessions and recommending changes to enhance user experience moving through the dashboard, such as adding more filtering options. HCI SMEs rated the dashboard as highly usable (0.82 on a scale of 0-4, with 0 being no usability concern and 4 being a catastrophic usability concern). The principle with the highest usability severity scores was a match between the system and the real world with a score of 1.4. The HCI SMEs also rated the information visualization aspects of the dashboard favorably with 2 usability principles, spatial organization and information coding, scoring 0.

Conclusions: Dashboards are a novel method for promoting and tracking facility capacity to manage BEmOC. By including targeted end users and SMEs in the design process, the team was able to tailor the dashboard to meet user needs, fit it into the existing government health systems, and ensure that the dashboard follows design best practices. Collectively, the novel, customized

XSL•FO
RenderX

BEmOC dashboard can be used to track and improve facility-level readiness in Amhara, Ethiopia, and similar global BEmOC facilities.

(JMIR Hum Factors 2025;12:e64131) doi:10.2196/64131

KEYWORDS

health information technology; design and evaluation; Ethiopia; usability; nursing informatics; user-centered design; basic emergency obstetric care; obstetric; nurse; user-centered; design; maternal mortality; maternal; develop; sub-Saharan Africa; Africa; dashboard; tracking; emergency care

Introduction

Maternal mortality is a critical public health issue, particularly in low-resource settings like sub-Saharan Africa, which accounts for over three-fourths of global maternal deaths [1]. Despite progress in reducing adverse maternal outcomes in countries like Ethiopia, high maternal mortality ratios persist, largely due to gaps and stockouts of essential supplies for managing obstetric emergencies [2,3]. Inadequate supplies for basic emergency obstetric care (BEmOC) can lead to delayed or suboptimal care. A study on facility-level readiness to manage BEmOC in Amhara, Ethiopia, using the Clinical Cascades reported a 63.3% mean overall BEmOC readiness for managing the 6 most common emergencies, which underscores the need for improving facility-level BEmOC readiness in this region [4].

A promising approach to improve facility-level BEmOC readiness is the development and implementation of health information technology (HIT) to monitor critical supply availability. HIT involves the processing, storage, and exchange of health information in an electronic environment, and has been shown to improve health outcomes and strengthen health systems in low-resource settings [5,6]. Electronic dashboards are a form of HIT that can be used to track critical information, provide alerts, assess performance indicators, develop reports, and customize data views, which makes them a useful tool for monitoring inventory data and facility-level readiness [7,8]. For example, dashboards were used in Ethiopia to monitor the supply inventory of malaria medications and prevent stockouts from occurring, and a simple dashboard was merged into the existing Integrated Pharmaceutical Logistics System (IPLS) in the Amhara and Oromia regions of Ethiopia, resulting in improved facility-level readiness to treat tuberculosis by ensuring health care facilities had adequate supplies [7,8]. Beyond Ethiopia, HIT has been used to anticipate medical stockouts and improve maternal health outcomes in low-resource settings [9,10]. This evidence highlights the potential for HIT to improve facility-level BEmOC readiness in Amhara, ultimately contributing to improved maternal care quality and reduced maternal mortality ratios.

Previous studies, including those conducted in low-resource settings and Ethiopia, have demonstrated the benefits of HIT and dashboards in improving facility-level readiness [5-8]. Furthermore, the growing utilization of technology and cellular services across Africa further supports the feasibility of implementing successful electronic solutions, driven by the increasing adoption of HIT [11,12]. Finally, Ethiopia has recognized the value of HIT in strengthening health care systems and has been implementing electronic medical records in its health care facilities since 2013 [13].

However, while dashboards can be a powerful tool for improving health outcomes, their success depends on ensuring high usability, enabling end users to effectively accomplish their tasks. Usability is defined as “the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use” [14]. Despite the potential of HIT, many implementations have failed due to insufficient consideration of end users’ needs and the environment in which the technology is applied [15,16]. Similarly, while dashboards show promise for improving facility-level BEmOC readiness, their success relies on development that incorporates input from targeted end users, accounts for the context of use, and adheres to usability principles. This study aims to develop and assess the usability of an electronic dashboard for monitoring facility-level BEmOC readiness.

Methods

Overview

To create the initial version of the dashboard, the team incorporated findings from previous qualitative interviews with individuals who worked in Ethiopia’s obstetric emergency supply chain, used information and designs from the preexisting dashboards in Ethiopia’s IPLS, and followed best practices found in the current literature for dashboard development [17-20]. The initial dashboard and its components were developed in Microsoft PowerPoint (refer to Multimedia Appendix 1).

To ensure adequate fit between the dashboard, the required tasks, the current system, the environment, and intended users, the research team conducted participatory user-centered design sessions, performed informal evaluations with domain and information visualization subject matter experts (SMEs), and completed a heuristic usability evaluation with human-computer interaction (HCI) SMEs.

Ethics Considerations

The research team obtained ethics approval from Columbia University’s institutional review board (IRB) (IRB-AAAU2006), Emory University’s IRB (MOD005-STUDY00005335), and the Amhara Public Health Institute (NoH/RfftTlDlo7l44). Before collecting data, the research team provided participants with an information sheet and gave them time to ask questions about the study. Once all questions had been answered, the team obtained verbal informed consent from all participants before

XSL•FO
RenderX

collecting data. All ethics review boards waived the requirement for written documentation of informed consent because the study procedures met the criteria for minimal risk to participants. All data was de-identified and participants data was linked to a random identification number. Only KD had access to the list of participant names linked to the ID numbers. None of the targeted end-users, domain experts, or subject matter experts received compensation for their participation in the study. Human-computer interaction experts participating in the heuristic evaluation received a $50 amazon gift card for their participation.

Sample Selection

User-Centered Participatory Design Sample Selection

Consistent with existing literature in the field of participatory design, the team estimated that a sample size of 15 would be sufficient to reach design saturation, a state where the design sessions reveal no major revisions to the designs reviewed [21,22]. The research team purposively recruited individuals with firsthand experience working within the BEmOC supply chain in Amhara, Ethiopia, at both the regional and local or facility levels. Participants were required to be at least 18 years old. Study participants were Ethiopian citizens and full-time employees in one aspect of Amhara's BEmOC supply chain, such as hospital supply managers, regional hub employees, and pharmacists.

Informal Expert Review Sample Selection

The domain SME was a research team member with over 6 years of experience conducting research in Amhara, Ethiopia; evaluating the supply chain; and measuring facility-level readiness to manage BEmOC. The information visualization SMEs were individuals from Columbia University School of Nursing's Visualization Design Studio, where individuals with varying levels of expertise can learn and gain advice on their current projects. The discussions at this studio were led by 2 PhD-level faculty, a nurse scientist with expertise in visualization design and evaluation, and a human-factors engineer. Key contributions also came from other faculty members at the university who have training in the areas of nursing research and biomedical informatics. All information visualization SMEs have published in the areas of information visualization, HCI, and user experience.

Heuristic Evaluation Sample Selection

The study sample included 5 HCI SMEs who were recruited through the team's professional networks. Previous research has found that 5-8 SMEs can identify over 80% of HIT usability violations [23,24]. HCI SMEs were eligible for inclusion if they had conducted research or published in the field of user interfaces or information visualizations and had not previously evaluated the dashboard.

Procedures

User-Centered Design Procedures

Staff from Emory-Ethiopia Partnership and Amhara Regional Health Bureau identified BEmOC supply chain stakeholders and reached out to determine if they were interested in participating in this study. A member of the research team contacted interested individuals via phone call, email, and WhatsApp (Meta) messaging to explain the study in detail, answer any questions, and set a time to meet in person with the participant to obtain verbal consent and perform the research activities. We used multiple communication channels to reach potential participants; in particular, WhatsApp has been found to be a useful tool for recruiting participants, specifically in global settings [25,26]. Design sessions were conducted in English and Amharic, and data collection tools were available in both languages since Amharic is the national language and Ethiopia's health care professionals receive clinical training in English [27,28].

Design sessions lasted approximately 60 minutes and occurred in small groups of 2-4 people within the health care facility or health bureau offices. Participants were given printed copies of the initial dashboard and asked to verbalize their thought processes as they explored the visualizations and information in the preliminary dashboard [29-32]. Design sessions were audio recorded and transcribed into English for analysis by a team member who is bilingual in Amharic and English. The individual conducting the analysis is a native Amharic speaker with over a decade of experience translating Amharic to English for health care research. Given their extensive expertise and the time constraints of the study, the team opted not to perform back translations. The dashboard development followed an iterative process, consisting of 4 design sessions. After each session, the research team incorporated recommended changes into the dashboard, which were then presented in the subsequent sessions [30].

Informal Expert Review Procedures

The team presented the Microsoft PowerPoint dashboard to a domain expert and a group of information visualization SMEs. Team member (KD) used Zoom (Zoom Video Communications) screen-sharing capabilities to walk the domain expert through the dashboard and their various features. The domain expert then reviewed the dashboard and provided recommendations for improvement.

KD presented the same dashboard to information visualization SMEs at 2 sessions of the Visualization Design Studio [33]. KD walked the SMEs through the dashboard, answered questions about their intended use, and elicited feedback and recommendations. Collectively, both rounds of feedback were used to optimize and further develop a medium-fidelity version of the dashboard for use in the heuristic usability evaluation.

Heuristic Evaluation Procedures

The team used Figma to develop a medium-fidelity prototype, which allowed the simulation of the dashboard's workflow [34]. Using Zoom Pro, a Health and Insurance Portability Accountability Act (HIPAA)–compliant videoconference platform, and its screen-sharing technology, KD asked the HCI SMEs to complete several tasks using the dashboard and encouraged them to explore other components of the dashboard (Multimedia Appendix 2). During their exploration, participants completed a heuristic evaluation checklist to identify and rank the severity of the usability violations and had the opportunity

XSL•FO
RenderX

to explain the violations they encountered [23,24,32]. They were also asked to describe what they were thinking, feeling, and seeing while completing the various tasks. The sessions were video recorded, and KD took field notes. At the end of the sessions, the HCI SMEs and KD discussed potential solutions for the identified usability violations. Participants were asked to complete tasks during the heuristic usability evaluation mentioned in Textbox 1.

Textbox 1. List of tasks participants were asked to complete during the heuristic usability evaluation.

Heuristic usability evaluation tasks

- Filter the main screen to only view primary hospitals.
- Determine which health care facilities are not ready to manage retained placenta.
- Determine if [hospital name] is ready to manage hypertensive emergencies. If not, identify which items are missing.
- Determine if [hospital name] is ready to manage maternal sepsis or infection. If not, identify which items are missing.
- Determine what emergencies [hospital name] is ready to manage.
- For prolonged labor at [hospital name] filter the supplies table to see the items that are at the emergency order point.
- For prolonged labor at [hospital name] filter the supplies to see the items that are overstocked.
- Determine how many health care facilities are at risk for not being able to manage retained placentas.

The Heuristic Evaluation Checklist used in this study was guided by Nielsen's 10 usability heuristics (ie, best practices; Table 1) [24,35] and includes a 5-item scale where experts rate each usability heuristic on a scale of 0 (not a problem) to 4 (catastrophic violation; Table 2) [23,32,36]. The research team modified the original checklist to fit the needs of a medium-fidelity dashboard (Multimedia Appendix 2). For example, questions for heuristics related to error prevention, help users with errors, and help and documentation were excluded since those tasks could not be simulated with the medium-fidelity prototype. In addition, 3 usability heuristics specific to information visualizations in the dashboards were included: spatial organization, information coding, and orientation [19].

Table 1. Usability heuristics and their definitions used in the heuristic usability evaluation.

Usability heuristic	Type	Definition
Visibility of system status	HIT[a]	The system should always keep users informed about what is going on through appropriate feedback in a reasonable time.
System and real-world match	HIT	The system should speak the user's language, with words, phrases, and concepts familiar to the user. The system should follow real-world conventions and ensure the dashboard fits within the existing workflow and technology system.
User control and freedom	HIT	Users should be free to select and sequence tasks and make their own decisions regarding the cost of exiting current work. Users should have clearly marked "emergency exit" to leave the unwanted state.
Consistency and standards	HIT	Users should not have to wonder whether different words, situations, or actions mean the same thing. Systems should maintain interface design choices in similar contexts and differ in different contexts.
Recognition rather than recall	HIT	The user should not have to remember information from one part of the dialogue to another. Objects, actions, and options should be easily visible, and instructions should be visible or easily retrievable whenever appropriate.
Flexibility and efficiency of use	HIT	The system should offer users several options for finding content. Users should be able to customize their interface and achieve their goals in an efficient manner and have the capacity to adapt to users' needs.
Aesthetic and minimalist design	HIT	The main dashboard should not contain information that is irrelevant or rarely needed. The system should present the largest amount of data with the least amount of ink.
Spatial organization	Information visualization	The overall layout of a visual representation should make it easy for the user to locate an information element in the display.
Information coding	Information visualization	The symbols and numbers used in the visualization should aid perception. The numeracy and graph literacy of the visualization should match the intended users' ability.
Orientation and help	Information visualization	The system should provide support for the user and help to orient them in their visualization.

[a]HIT: health information technology.

XSL•FO
RenderX

Table 2. Severity ratings for usability violations in the heuristic evaluation.

Score	Usability problem	Explanation
0	Not a usability problem at all	No problem identified
1	Cosmetic problem	Need not be fixed unless extra time is available on the project
2	Minor problem	Fixing this should be a low priority
3	Major usability	Important to fix, so should be given high priority
4	Usability catastrophe	Imperative to fix this before the product can be released

Data Analysis

User-Centered Design Analysis

Analysis was concurrent with data collection with modifications of the dashboard occurring after the initial 2 design sessions and following the third and fourth sessions. Using the translated transcripts of audio-recorded sessions and field notes, KD extracted key preferences, identified poorly performing graphics and dashboard components, and improved the design based on participant feedback. The team maintained detailed notes on design decisions and the rationale for those decisions alongside the notes and transcripts, which functioned as an audit trail.

Heuristic Usability Evaluation Analysis

The responses on the Heuristic Evaluation Checklist were transferred into Microsoft Excel for analysis [19,20]. The research team used unique participant numbers to link the responses to a confidential participant list. The data were analyzed with descriptive statistics (frequencies and means) [37]. Both session field notes and qualitative responses reported in the free response comments section of the checklist were used to understand usability and enhance the dashboard's design. During the evaluations, KD documented key areas of concern that HCI SMEs verbalized during the sessions. Combing those notes with the free-response comments, KD used the frequency that those concerns were mentioned to understand their importance. The usability concerns and potential solutions reported from the free-response comments were also compiled into a single list of action items, with duplicates removed, and KD used this list to guide the refinement of the next version of the dashboard.

Results

User-Centered Design Sessions

Sample Characteristics

In total, 6 individuals, who work at the regional health bureau or regional hubs for supply distribution, participated in the first 2 design sessions and reviewed the initial version of the dashboard. Following the refinement of the dashboard by KD, 5 individuals who work at health care facilities in the region participated in the third and fourth design sessions and reviewed the updated version of the dashboard. Recruitment and data collection ended prematurely at 11 participants due to civil unrest and safety concerns in the region [38].

Evolution of the Dashboard Design

During design sessions 1 and 2, participants reviewed the initial version of the prototype, and during design sessions 3 and 4, they reviewed a revised version that incorporated feedback from study participants in design sessions 1 and 2. During design sessions 1 and 2, participants viewed the initial design of the dashboard and were able to easily grasp the idea the dashboard was trying to portray related to facility readiness to manage BEmOC. The users felt the dashboard, specifically, the regional view, helped them to understand which facilities would be ready to handle the emergencies. One regional respondent stated,

> *This image [dashboard] demonstrates how they [individual health care facilities] would be able to handle each health issue considering the level of the supplies that each hospital has for these health problems. In general, we can conclude from this image [regional dashboard] that some hospitals are in good condition for these health issues and others are not in good condition for these health issues.*

Generally, participants in the first two sessions focused on (1) individual graphical component preferences such as which type of charts they preferred to see the data presented and (2) how they would want the data to be filtered and sorted. The third and fourth sessions with the refined dashboard focused on (1) confirming the terminology in the dashboard, and that it matched terms used in the supply chain, and (2) affirming design changes that were incorporated from the earlier sessions. Respondents in the third and fourth sessions also provided feedback on which job roles should have access to the dashboard. A participant during the fourth design session stated,

> *This dashboard [emergency-specific view] is essential for the dispensary unit as well. The store and supply officer bears most of the workload or duty in this workplace, but if the pharmacists and hospital managers have access to the data, it would be very good.*

The changes that were incorporated during the design sessions, as well as the session number, iterations of the dashboard, and job level that provided the recommendations, can be found in Multimedia Appendix 3. In the following paragraphs, more details are provided on aesthetic design, filtering and sorting, and matching with real-world preferences.

Aesthetic Design

For colors used throughout the dashboard, participants wanted to use multiple colors that were consistent with those currently used in the existing dashboard. Furthermore, participants in the

XSL•FO
RenderX

first 2 design sessions identified pieces of information that they thought should be removed from the dashboard because they did not assist the end users with making decisions related to supply ordering or maintaining facility-level readiness. However, to ensure users could find information that was not included in the dashboard, the research team added a button that would allow users to view the comprehensive drug list, which is already embedded in the electronic components of the existing IPLS. Facility-level participants, during design sessions 3 and 4, explained that not all health posts will manage certain obstetric emergencies, such as sepsis, so they should not receive readiness classifications for those emergencies. In response to this, the research team grayed out the classifications for those facilities and their corresponding emergencies.

Another aesthetic preference in the early design sessions included requests to include keys in the dashboard that define the colors and terms. Participants in design sessions 3 and 4 agreed that the keys assisted in their comprehension of the dashboard. Participants in design sessions 3 and 4 affirmed the design choices of the previous sessions and provided only small recommendations for the overall look of the dashboard, such as increasing the font size. Related to aesthetic decisions, design saturation was met with the smaller sample size.

Filtering and Sorting

Participants in the early sessions wanted to ensure there were multiple ways to filter the data so they could view pieces of information that were most important to them. Preferred filtering and sorting options included viewing inventory alphabetically within preestablished categories from the Ethiopian Pharmaceutical Supply System (eg, medical equipment, pharmaceuticals, and supplies), filtering supplies based on quantity available at the facility, and filtering facility by health care facility tier (eg, primary, general, referral, and center) [17]. Participants during the third and fourth design sessions endorsed the filtering options identified in the earlier sessions.

The participants from design sessions 3 and 4 also wanted employees at individual facilities to be able to view inventory data at other facilities besides their own. They felt having access to these data could support transferring excess supplies from one facility to another when neighboring facilities faced stockouts or were at an emergency order point. Thus, the research team added the ability for facility-level users to select different facility readiness data, similar to the regional-level users. Since there was very little new information identified during the second 2 design sessions, the research team determined that the study achieved design saturation for filtering preferences even with the smaller respondent sample.

Match With the Real-World Preferences

A feature that was a very important point of discussion in all the design sessions was using appropriate terminology that aligned with the participants' current workflow and job expectations. Participants wanted to ensure that all the terms used within the dashboard aligned with terms from the IPLS. For example, supply names and categories as well as terms for quantities available were updated (eg, overstock changed to excess, and normal changed to functional for the category medical supplies).

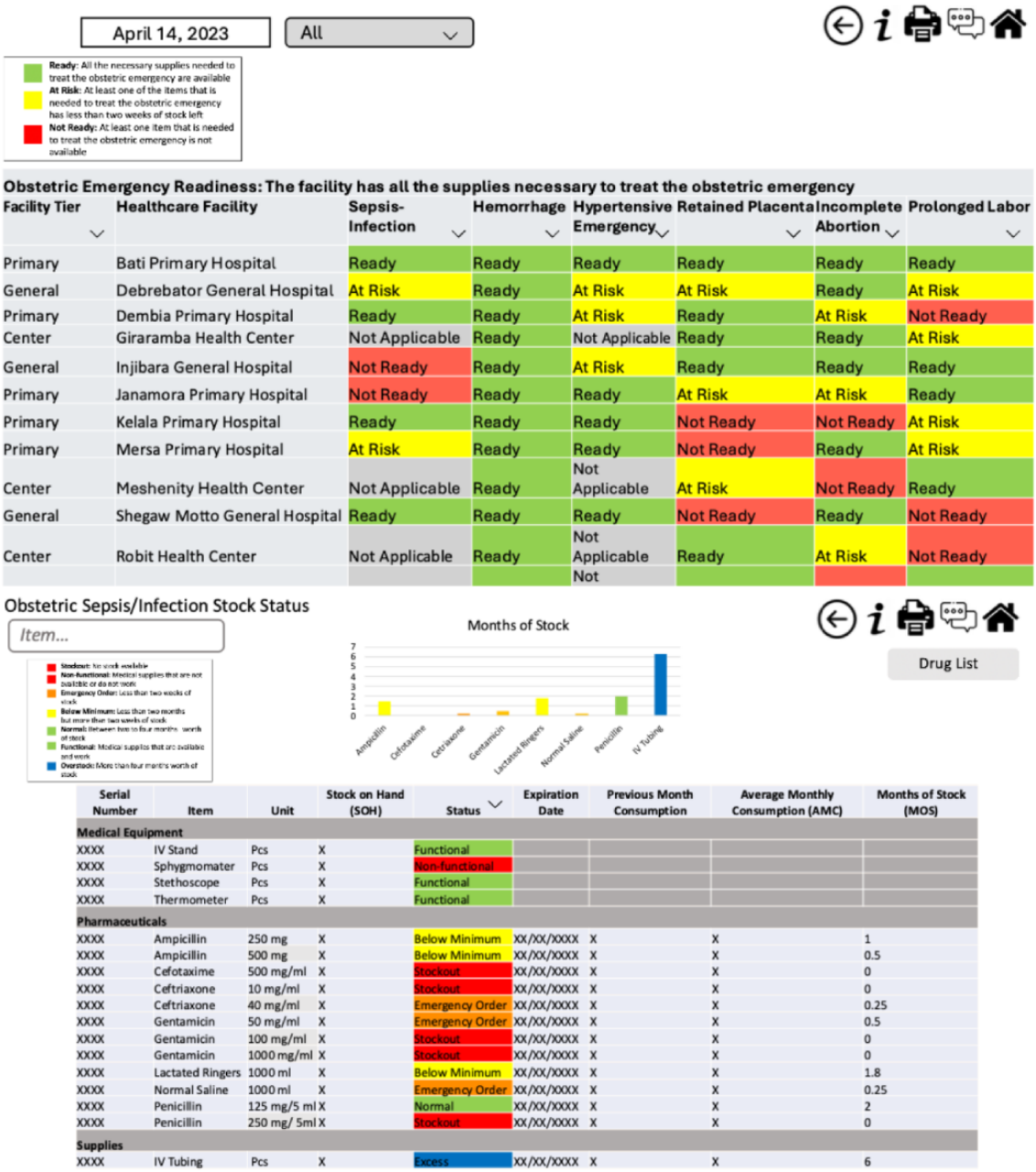
One important change to note is that during the third design session, participants recommended changing the terminology for BEmOC readiness from ready, at risk, and not ready to normal, emergency order, and stockout so they would be the same as the terms used with the emergency-specific dashboard view. Based on this recommendation, the research team created 3 different readiness terminology keys. The 3 options included the original terms; the newly recommended terms; and a third option of yes, no, and emergency order. Since the team could not confirm their preferred choice for readiness terminology due to the civil conflict in the region, the domain expert selected the final terms—yes, at risk, and no. Since the team was unable to obtain user preferences related to this final change, the study did not achieve design saturation in the "match with the real-world" domain.

Final Design

The final design comprised 2 different dashboard views for monitoring BEmOC readiness—one for the regional level and another for the facility level. The regional-level dashboard view provides a summary of readiness for the 6 obstetric emergencies for all health care facilities in the region. The facility-level dashboard view focuses on one facility and one obstetric emergency at a time. Overall study participants viewed the dashboard positively, with one respondent from design session 4 reporting "The hard copies [of the dashboard] you showed us today are beautiful and wonderful." Figure 1 shows the design of the dashboard following the integration of the findings from the third and fourth user-centered design sessions, and Multimedia Appendix 3 lists all the changes that were incorporated into the dashboard following the user-centered design sessions.

XSL•FO
RenderX

Figure 1. Regional basic emergency obstetric care (BEmOC) readiness dashboard and emergency-specific, facility-level dashboard following user-centered design sessions. Hospital names are real, but all readiness data are fictitious.



Informal Expert Review

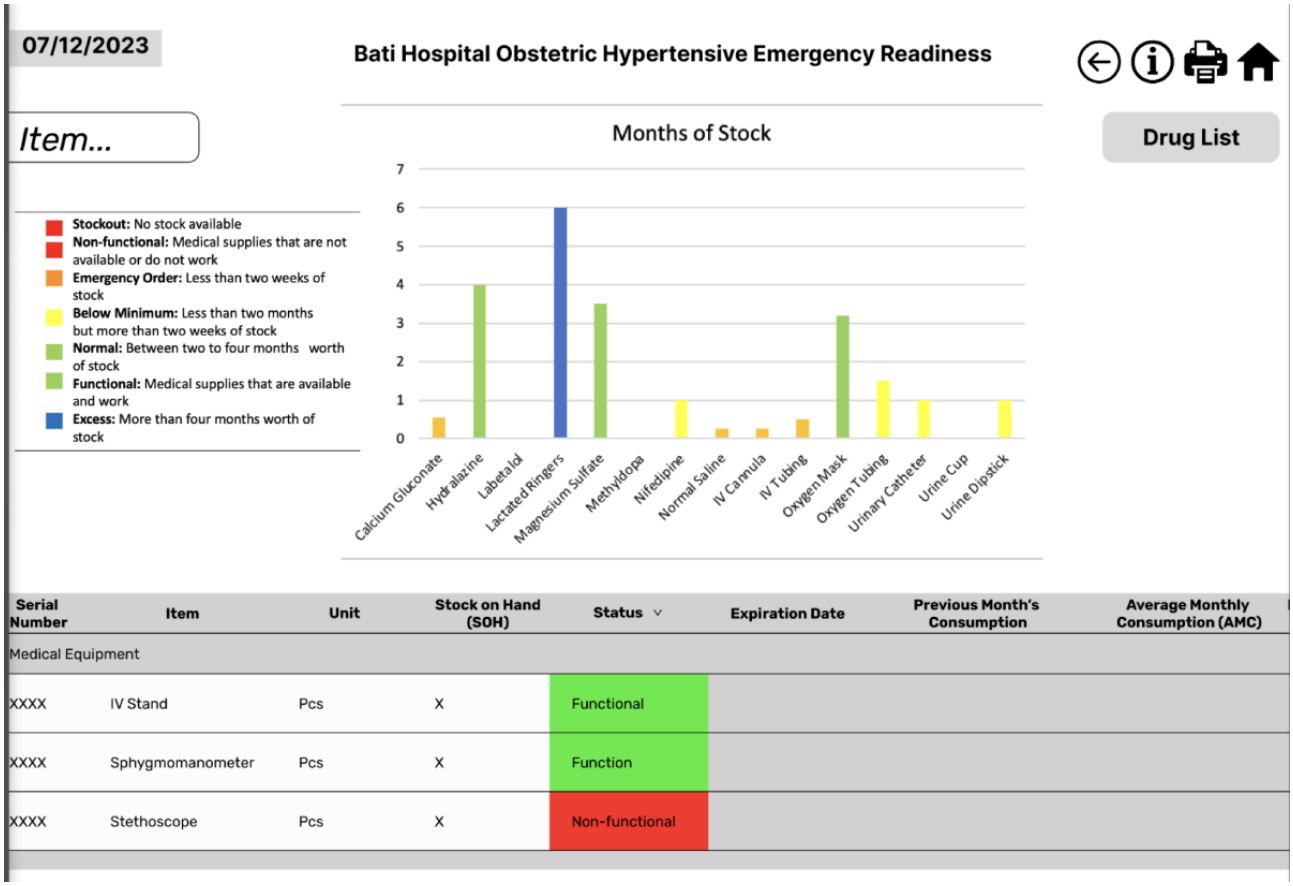
The domain expert's largest contribution to the refinement of the dashboard was his input on which terms to use to define BEmOC readiness, specifically for the regional view dashboard. To make the final decision on readiness classification terms that were discussed in the final user-centered design session, KD presented the domain expert with 3 different options created following the fourth design session. After reviewing the options, the domain expert recommended KD use yes, emergency order, and no. The rationale for this decision was the yes and no choice was the simplest option, and he believed the end users would be able to readily act based on these classification terms. The domain expert also affirmed KD's decisions to group data at the regional and facility levels and felt that no critical information was missing from the dashboard.

During the review by information visualization SMEs, most recommendations aesthetics related to aesthetics and navigation. In terms of aesthetic design, the recommendations mainly focused on where to place different pieces of information or graphics to best grab the users' attention or prevent overcrowding on a screen, for example, moving several columns of data to the right in the emergency-specific dashboard view. This would allow users to scroll horizontally to see that information, but it would not be visible without scrolling so there would be less information crowding the page. To improve navigation between the various dashboard views, the SMEs during the design studio recommended adding additional filtering options, such as for emergency readiness categories. They also recommended adding an affordance feature to the dashboard so that users would be able to see which items on the screen are clickable or not. In addition, the experts

XSL•FO
RenderX

recommended adding several additional titles to different dashboard views to enhance clarity. These changes were incorporated when the dashboard was updated from PowerPoint to Figma (Figure 2).

Figure 2. Regional view dashboard following informal expert review. All inventory data in this dashboard are fictitious.



Heuristic Usability Evaluation With Human-Computer Interaction Experts

The heuristic evaluation took approximately 30 minutes per session. The years of experience for HCI SMEs ranged from 4 to 14 years, with a mean of 9. All HCI SMEs were female.

Overall, severity scores for heuristic evaluation ranged from 0 to 3 for the 10 usability principles. The mean usability score for all ten heuristics combined was 0.82 (SD 0.98), with a mode of zero. The principle with the highest severity rating was a match between the system and the real world, with a score of 1.4, which falls between a minor and cosmetic usability concern. For this principle, 4 HCI SMEs identified inappropriate color choices for both potential users with color blindness and using colors not congruent with common color-coding expectations as a common issue. Additional problems included the dashboard using concepts and phrases that were unfamiliar to the evaluators, and 3 evaluators felt the section headings were not ordered in the most logical fashion. HCI SMEs identified issues with usability principles visibility of system status, user control and freedom, consistency and standards, and recognition rather than recall (mean 1.2 for each, SD 0.84, 1.3, 0.84, and 1.3, respectively), which also falls between a cosmetic and minor usability concern. Concerns related to these principles included 2 of the HCI SMEs identifying the lack of clear exits for every dashboard and feeling that the current design hindered their ability to efficiently navigate the system. Table 3 shows the mean value, range, and mode for the 10 heuristic principles.

XSL•FO
RenderX

Table 3. Identified heuristic usability problems and their severity scores.

Heuristic category	Problems	Score, mean (SD)	Score, range	Score, mode
Visibility of system status	• Title font is too small • Color choices may be inappropriate for color-blind users • Vertical bar charts make reading supply labels more difficult • Need hyperlinks to improve navigation • Menu-naming terminology is not consistent with the user's task domain	1.2 (0.84)	2	1 and 2
Match between system and real-world	• Color choices may be inappropriate for color-blind users • Section headings and subheadings are not ordered in the most logical way • Not all words, concepts, and phrases were familiar to the human-computer interaction subject matter experts • Some of the colors used in the dashboard do not correspond to common expectations about color codes	1.4 (0.89)	2	1
User control and freedom	• Need hyperlinks to improve navigation • Not a clear exit on each dashboard screen • Not all screens are accessible across the system • Users could not easily move forward and backward between fields	1.2 (1.3)	3	0
Consistency and standards	• Titles on the regional view dashboard do not update when the user filters the data • Abbreviations not clearly explained • Some colors are too similar to distinguish from other categories • Not enough or inconsistent visual cues to identify active screens	1.2 (0.84)	2	1 and 2
Recognition rather than recall	• White space is not optimized within the emergency-specific dashboard view • Prompts, cues, and messages are not placed where the eye is likely to be looking on the screen	1.2 (1.3)	3	0
Flexibility and efficiency of use	• Need hyperlinks to improve navigation	0.6 (0.89)	2	0
Aesthetic and minimalist design	• Not all field labels are brief, familiar, or descriptive • Large objects, bold fonts, and simple areas have not been used to distinguish sections • There is not enough white space between color representation • Too much text is present in the keys which makes the screens look busy	0.8 (0.84)	2	0 and 1
Spatial organization	• Font size is too small throughout the dashboard • Information does not follow a logical flow	0 (0.0)	0	0
Information coding	• No problems identified	0 (0.0)	0	0
Orientation	• Measurement units are not displayed clearly • Users cannot control the level of detail they see in a representation	0.6 (1.3)	3	0

The experts provided several recommendations for how to remedy the problems identified during the heuristic evaluation. They believed adding hyperlink functions would improve users' ability to navigate between the dashboard views, thus increasing usability scores for visibility of system status, user control and freedom, and flexibility and efficiency of use. Aesthetically, the experts believed increasing the font size of titles and data within the dashboard, as well as moving different components of the dashboard around to maximize white space would improve scores for spatial organization, aesthetic and minimalist design, recognition rather than recall, and visibility of system status.

HCI SMEs also made recommendations that disagreed with user preferences identified during the design sessions. These differences included the experts preferring horizontal bar charts compared to vertical bar charts. An expert explained the rationale for this preference stating a horizontal bar chart allows for the labels in the table to be read horizontally, which can be easier for users. However, during the user-centered design sessions, the participants had a resounding preference for the vertical bar charts because they were more familiar with that format since some of their preexisting dashboards included vertical bar charts in the IPLS. Given the importance of the new dashboard to mimic the look of the existing ones, as well as the

XSL•FO
RenderX

fact that the users explicitly preferred the vertical presentation, KD chose to retain the vertical bar charts. The color of the tables and graphs was also a point of disagreement between experts and users. The experts did not think the red, yellow, and green color choices were appropriate because these colors can be difficult to distinguish if individuals are color blind [39]. However, the targeted end users of the dashboard preferred those colors because the colors reminded them of traffic lights, and the users could make assumptions related to the significance of the colors based on what those colors mean by a traffic light. The research team chose to keep the traffic light color scheme.

Beyond these recommendations, the reviewers cited several positive features of the dashboard. This includes things such as the belief that the icons used throughout the dashboard were clear and easy to associate with their function. In addition, the experts found the labels and keys to be clear and assisted with comprehension of the data. One respondent reported, "Your labels are quite clear. It's [the dashboard] very accessible and you have nice keys right here." The heuristic violation severity scores for information visualization-specific heuristics were low with scores of zero for spatial organization and information coding, and 0.6 for orientation. Table 4 summarizes the changes incorporated into the dashboard following the completion of the heuristic usability evaluation. The final dashboard is displayed in Figure 3.

Table 4. Summary of changes incorporated into the dashboard following the heuristic evaluation

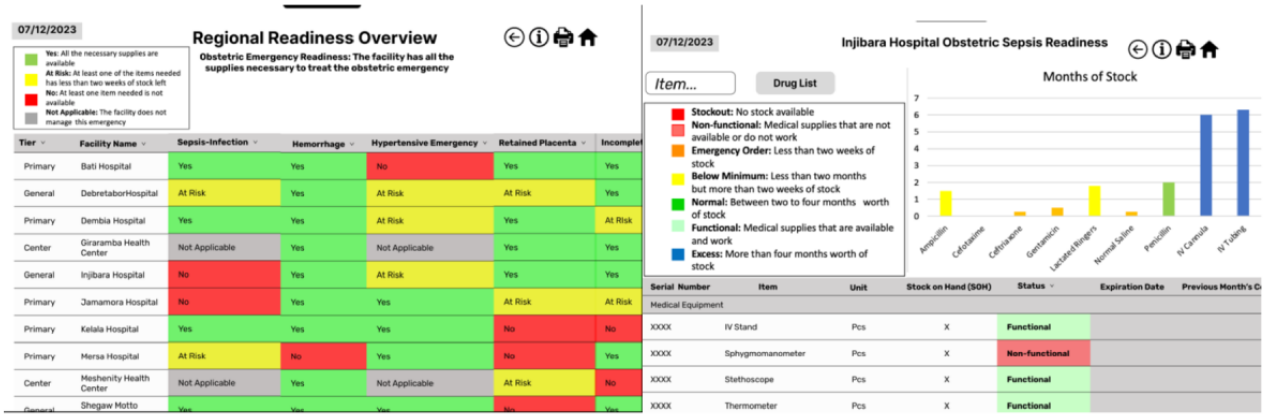
Heuristic category	Changes following heuristic evaluation
Visibility of system status	• Increased font size throughout the dashboard • Added hyperlinks to the dashboard
Match between system and real-world	• Updated section headings to reflect filtering capabilities • Updated the terminology defining BEmOCa readiness on the regional-level dashboard
User control and freedom	• Added hyperlinks to the dashboard • Ensured exit buttons on every screen are activated • Activated more functions within the dashboard to make it a higher-fidelity prototype
Consistency and standards	• Updated section headings to reflect filtering capabilities • Defined all abbreviations used in the dashboard • Changed the colors for functional and nonfunctional on the emergency-specific screen so that they are more easily distinguishable from the other colors used on the screen • Added additional visual cues to identify active screens
Recognition rather than recall	• Moved graphics at the top of the dashboard to optimize white space • Added additional visual cues to identify active screens
Flexibility and efficiency of use	• Added hyperlinks to the dashboard
Aesthetic and minimalist design	• Shortened information in the keys to make them one line of text • Shortened data within the tables to make it one line of text • Used bold fonts for all titles • Moved graphics at the top of the dashboard to optimize white space
Spatial organization	• Increased font size throughout the dashboard • Improved the navigation between dashboard views to make them a higher-fidelity prototype
Information coding	• N/Ab
Orientation	• Increased font size for measurement units • Activated more functions within the dashboard to make it a higher-fidelity prototype

[a]BEmOC: basic emergency obstetric care.

[b]N/A: not applicable.

XSL•FO
RenderX

Figure 3. Final regional view and emergency-specific view of the dashboard following the heuristic usability evaluation. Hospital names are real, but all readiness data are fictitious.



Discussion

Principal Findings

The dashboard underwent iterative refinement based on user input, domain expert feedback, information visualization evaluations, and heuristic usability assessments. Regional participants emphasized preferences for vertical bar charts, traffic-light color coding, and keys to define readiness levels, while facility-level users focused on aligning terminology with existing workflows and tailoring readiness classifications. The information domain expert provided critical input on readiness terminology, and information visualization experts suggested improvements in layout, navigation, and aesthetics, such as optimizing white space and increasing font sizes. Collectively, the central tendency of heuristic severity scores, combined with a mode of zero, indicated high overall usability of the dashboard. However, the HCI experts did highlight minor usability concerns, including navigation and accessibility improvements, which were addressed in the final design. While some expert recommendations, such as horizontal bar charts and alternative color schemes, differed from user preferences, the team prioritized end-user input to ensure practical usability. The final dashboard is intuitive, visually clear, and aligned with real-world needs, achieving high usability and user satisfaction.

This study highlights the importance of including participants from all targeted end-user groups. By including employees from both regional and facility levels, the team gathered diverse, relevant perspectives, ensuring the dashboard met the real-world needs of users across different supply chain roles and experiences. Feedback from both end users and experts enhanced the overall usability and functionality of the dashboard. This aligns with existing usability literature, which emphasizes the value of involving both end users and experts to identify distinct concerns and improve HIT design [40]. Participants' spontaneous interpretations of the dashboard underscore their understanding of the content, and obtaining these perceptions and insights can assist in preventing usability errors from occurring later in the piloting or implementation phase of HIT development. These refinements will likely maximize the ease of implementation and promote consistent uptake and utilization by future users [41,42].

There were 2 areas of disagreement between HCI SMEs' recommendations and end-user preferences: bar chart orientation and color schemes. The research team chose to keep the vertical bar charts for their familiarity with the end users and alignment with existing dashboards. In addition, lower color blindness prevalence among Black African and African American populations (4.2% compared to 6.6% and 1.4% compared to 5.6% respectively) compared to White populations (6.6% and 5.6%) [43,44] supported the decision to maintain the preferred color scheme, as most targeted end users were unlikely to face difficulties distinguishing the colors.

This study reinforces the notion that obtaining expert opinions and performing usability testing can be a low-cost, efficient method for exploring technology needs [45]. Performing the evaluations with the domain expert and the information visualization SMEs took less than an hour for both groups and provided critical information on how to improve upon the design of the dashboard. Performing usability evaluations can identify problems and usability concerns that may not have been noticed before HIT implementation, avoiding costly and time-consuming corrections [24].

Finally, this study contributes to the growing body of literature on the use of dashboards to support BEmOC in sub-Saharan Africa. Similar to findings by Banke-Thomas et al [46] and Wang et al [47], stakeholders expressed enthusiasm for dashboards and prioritized features like reporting facility characteristics on the dashboards. However, there remains a gap in the literature regarding the specific design preferences of end users. This study addresses that gap by highlighting key preferences, such as the use of stoplight color coding and IPLS-specific terminology, which enhance usability and are tailored specifically for dashboards intended for Ethiopia and BEmOC contexts.

Limitations

During data collection, there was unanticipated civil unrest in the Amhara region, which prevented the research team from recruiting participants outside of the regional capital of Bahir Dar and from completing participant recruitment from multiple facilities outside the regional capital. Despite this unforeseen barrier, design saturation was reached for 2 out of the 3 categories during the design sessions. There was also a lack of

XSL•FO
RenderX

gender diversity among our participants. Only one user-centered design participant from Ethiopia was female. However, this gender imbalance mirrored the existing gender imbalance in the existing health system for these roles. However, future research related to this dashboard and facility-level BEmOC readiness would benefit from an emphasis on opinions and participation from female participants with relevant expertise. Conversely, all HCI SMEs were female. While the incorporation of a female perspective is critical since the sample for user-centered design sessions was predominately male, the effect of having completely female HCI SMEs is unknown. Future research should incorporate technical and gender-balanced perspectives during both user and HCI SMEs' review of this dashboard.

Furthermore, the dashboard design was informed by responses from participants in qualitative interviews. However, due to time constraints, the initial prototype was developed by the research team rather than being cocreated with participants. While this approach allowed the team to efficiently move forward with development, it may have limited the incorporation of deeper, iterative feedback from participants during the initial design phase. This limitation could have influenced the alignment of the prototype with user needs and preferences, which may require further refinement in future iterations.

To prospectively limit social desirability bias, the research team created a low-stress, comfortable environment for data collection and emphasized our desire to obtain personal responses to the dashboard while reiterating that all opinions would be blinded so no one could identify specific feedback from specific participants. The team chose to have an Amharic-speaking Ethiopian team member conduct the user-centered design sessions rather than the principal investigator, who is from the United States and does not speak Amharic. The team took steps to ensure confidentiality and used indirect questioning techniques in an attempt to garner the participants' true opinions [48]. Furthermore, the dashboard reviewed in the heuristic usability evaluation was medium fidelity. This means the dashboard was not able to complete all required tasks and certain usability heuristics, such as *help and documentation* and *error prevention*, were excluded from this evaluation, since their tasks were not applicable at this stage in the design process. Future research will need to explore all usability heuristics before implementation to ensure no new concerns arise once the dashboard transitions from a medium- to high-fidelity model.

Conclusion

In conclusion, this study demonstrates the value of user-centered design and usability evaluations in developing HIT for low-resource settings. By integrating diverse perspectives from regional and facility-level participants, as well as domain and usability experts, the dashboard was refined to meet real-world needs and achieve high usability. This iterative approach not only addressed usability concerns but also incorporated end-user preferences, ensuring alignment with existing workflows. The study fills a gap in the literature by identifying design preferences tailored to Ethiopia's health care context, offering a model for future HIT development aimed at improving BEmOC readiness and maternal health outcomes. These findings emphasize the importance of engaging end users and experts early in the design process to create functional, user-friendly systems that support effective implementation and long-term adoption.

Acknowledgments

We would like to thank Emory-Ethiopia Partnership's Addis Ababa and Bahir Dar offices for their support and guidance in this research.

This work was supported by the National Institute of Nursing Research (F31NR020569), a Fulbright US Student research award to Ethiopia, and a research grant from Sigma Theta Tau Alpha Zeta chapter.

Authors' Contributions

YA performed data collection with support from MB and LA. HB supported the project administration and maintained ethics approvals throughout the study. KD analyzed the data and created the dashboard with assistance from SB and JNC. KD drafted the manuscript, and all other authors assisted in editing and refining the final manuscript. JNC conceptualized the nested study design and obtained parent funding for the Saving Little Lives study which this research was nested within. AG provided contextual expertise on Amhara and Ethiopian clinical care.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The graph options and preliminary dashboard prototypes used in the user-centered design sessions.
[DOCX File , 206 KB - humanfactors_v12i1e64131_app1.docx]

Multimedia Appendix 2
Heuristic usability evaluation survey.
[DOCX File , 28 KB - humanfactors_v12i1e64131_app2.docx]

XSL•FO
RenderX

Multimedia Appendix 3
Changes incorporated into the dashboard following the user-centered design sessions.
[DOCX File , 22 KB - humanfactors_v12i1e64131_app3.docx]

References

1. Trends in maternal mortality 2000 to 2017: estimates by WHO, UNICEF, UNFPA, World Bank Group and the United Nations Population Division. World Health Organization. 2021. URL: https://www.unfpa.org/featured-publication/trends-maternal-mortality-2000-2017 [accessed 2025-02-20]
2. Emory-Ethiopia partnership. Emory University's Nell Hodgson Woodruff School of Nursing. URL: https://www.nursing.emory.edu/overview/emory-ethiopia-partnership [accessed 2021-07-20]
3. Tessema GA, Laurence CO, Melaku YA, Misganaw A, Woldie SA, Hiruye A, et al. Trends and causes of maternal mortality in Ethiopia during 1990-2013: findings from the Global Burden of Diseases Study 2013. BMC Public Health 2017;17(1):160 [FREE Full text] [doi: 10.1186/s12889-017-4071-8] [Medline: 28152987]
4. Dougherty K, Gebremariam Gobezayehu A, Lijalem M, Alamineh Endalamaw L, Biza H, Cranmer JN. Comparison of obstetric emergency clinical readiness: a cross-sectional analysis of hospitals in Amhara, Ethiopia. PLoS One 2023;18(8):e0289496 [FREE Full text] [doi: 10.1371/journal.pone.0289496] [Medline: 37535678]
5. Brenner SK, Kaushal R, Grinspan Z, Joyce C, Kim I, Allard RJ, et al. Effects of health information technology on patient outcomes: a systematic review. J Am Med Inform Assoc 2016;23(5):1016-1036 [FREE Full text] [doi: 10.1093/jamia/ocv138] [Medline: 26568607]
6. Ngwa W, Olver I, Schmeler KM. The use of health-related technology to reduce the gap between developed and undeveloped regions around the globe. Am Soc Clin Oncol Educ Book 2020;40:1-10 [FREE Full text] [doi: 10.1200/EDBK_288613] [Medline: 32223667]
7. Fact sheet Ethiopia. USAID. 2020. URL: https://www.usaid.gov/sites/default/files/2022-05/Ethiopia-Fact-Sheet_PMI_July-2020.pdf [accessed 2025-02-20]
8. Rabiei R, Almasi S. Requirements and challenges of hospital dashboards: a systematic literature review. BMC Med Inform Decis Mak 2022;22(1):287 [FREE Full text] [doi: 10.1186/s12911-022-02037-8] [Medline: 36348339]
9. Jbaily A, Feldhaus I, Bigelow B, Kamareddine L, Tolla MT, Bouvier M, et al. Toward health system strengthening in low- and middle-income countries: insights from mathematical modeling of drug supply chains. BMC Health Serv Res 2020;20(1):776 [FREE Full text] [doi: 10.1186/s12913-020-05549-z] [Medline: 32838778]
10. Till S, Mkhize M, Farao J, Shandu LD, Muthelo L, Coleman TL, et al. Digital health technologies for maternal and child health in Africa and other low- and middle-income countries: cross-disciplinary scoping review with stakeholder consultation. J Med Internet Res 2023;25:e42161 [FREE Full text] [doi: 10.2196/42161] [Medline: 37027199]
11. Bukachi F, Pakenham-Walsh N. Information technology for health in developing countries. Chest 2007;132(5):1624-1630. [doi: 10.1378/chest.07-1760] [Medline: 17998362]
12. Lewis T, Synowiec C, Lagomarsino G, Schweitzer J. E-health in low- and middle-income countries: findings from the center for health market innovations. Bull World Health Organ 2012;90(5):332-340 [FREE Full text] [doi: 10.2471/BLT.11.099820] [Medline: 22589566]
13. Bekele TA, Gezie LD, Willems H, Metzger J, Abere B, Seyoum B, et al. Barriers and facilitators of the electronic medical record adoption among healthcare providers in Addis Ababa, Ethiopia. Digit Health 2024;10:20552076241301946 [FREE Full text] [doi: 10.1177/20552076241301946] [Medline: 39659403]
14. Usability of consumer products and products for public use-part 2:summative test method. International Organization for Standardization. URL: https://www.iso.org/obp/ui/#iso:std:iso:ts:20282:-2:ed-2:v1:en [accessed 2021-10-09]
15. Abugabah A, Alfarraj A. Issues to consider in designing health care information systems: a user-centred design approach. Electronic Journal of Health Informatics 2015;9.
16. Gnanlet A, Choi M, Davoudpour S. Impediments to the implementation of healthcare information technology : a systematic literature review. Journal of Supply Chain and Operations Management 2019;17(1).
17. Standard operating procedures (SOP) manual for the Integrated Pharmaceuticals Logistics System in health facilities of Ethiopia. International Institute for Primary Health Care - Ethiopia. 2020. URL: http://repository.iphce.org/handle/123456789/2613 [accessed 2025-02-20]
18. Dougherty K, Gebremariam A, Biza H, Belew M, Benda N, Tesfaye Y, et al. Obstetric emergency supply chain dynamics and information flow among obstetric emergency supply chain employees: key informant interview study. JMIR Form Res 2024;8:e59690 [FREE Full text] [doi: 10.2196/59690] [Medline: 39235860]
19. Dowding D, Merrill JA. The development of heuristics for evaluation of dashboard visualizations. Appl Clin Inform 2018;9(3):511-518 [FREE Full text] [doi: 10.1055/s-0038-1666842] [Medline: 29998455]
20. Nielsen J. 10 Usability heuristics for user interface design. Nielsen Norman Group. 1994 Apr 24. URL: https://www.nngroup.com/articles/ten-usability-heuristics/ [accessed 2025-02-20]
21. Faulkner L. Beyond the five-user assumption: benefits of increased sample sizes in usability testing. Behav Res Methods Instrum Comput 2003;35(3):379-383. [doi: 10.3758/bf03195514] [Medline: 14587545]

XSL•FO
RenderX

22. Guest G, Bunce A, Johnson L. How many interviews are enough? Field Methods 2006;18(1):59-82. [doi: 10.1177/1525822X05279903]
23. Allen M, Currie LM, Bakken S, Patel VL, Cimino JJ. Heuristic evaluation of paper-based Web pages: a simplified inspection usability methodology. J Biomed Inform 2006;39(4):412-423 [FREE Full text] [doi: 10.1016/j.jbi.2005.10.004] [Medline: 16321575]
24. Nielsen J. Finding usability problems through heuristic evaluation. 1992 Presented at: CHI '92: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; May 3-7, 1992; Monterey, CA p. 373-380 URL: https://doi.org/10.1145/142750.142834 [doi: 10.1145/142750.142834]
25. Darko EM, Kleib M, Olson J. Social media use for research participant recruitment: integrative literature review. J Med Internet Res 2022;24(8):e38015 [FREE Full text] [doi: 10.2196/38015] [Medline: 35925655]
26. Manji K, Hanefeld J, Vearey J, Walls H, de Gruchy T. Using WhatsApp messenger for health systems research: a scoping review of available literature. Health Policy Plan 2021;36(5):594-605 [FREE Full text] [doi: 10.1093/heapol/czab024] [Medline: 33860314]
27. Medicine. Addis Ababa University. 2021. URL: http://www.aau.edu.et/medicine/ [accessed 2025-02-20]
28. Nursing. Addis Ababa University. 2021. URL: http://www.aau.edu.et/nursing/ [accessed 2025-02-20]
29. Lai TY, Larson EL, Rockoff ML, Bakken S. User acceptance of HIV TIDES--tailored interventions for management of depressive symptoms in persons living with HIV/AIDS. J Am Med Inform Assoc 2008;15(2):217-226 [FREE Full text] [doi: 10.1197/jamia.M2481] [Medline: 18096915]
30. Lucero RJ, Sheehan B, Yen PY, Velez O, Nobile-Hernandez DL, Tiase VL, et al. Developing self-management tools with vulnerable populations for use in personal health information management systems. NI 2012 (2012) 2012;2012:248 [FREE Full text] [Medline: 24199096]
31. Vélez O, Okyere PB, Kanter AS, Bakken S. A usability study of a mobile health application for rural Ghanaian midwives. J Midwifery Womens Health 2014;59(2):184-191 [FREE Full text] [doi: 10.1111/jmwh.12071] [Medline: 24400748]
32. Yen PY, Bakken S. A comparison of usability evaluation methods: heuristic evaluation versus end-user think-aloud protocol - an example from a web-based communication tool for nurse scheduling. AMIA Annu Symp Proc 2009;2009:714-718 [FREE Full text] [Medline: 20351946]
33. Arcia A, Benda N. Visualization Design Studio. URL: https://vis-studio.mailchimpsites.com/ [accessed 2023-08-08]
34. Figma. 2016. URL: https://www.figma.com/ [accessed 2025-02-20]
35. Nielsen J, Molich R. Heuristic evaluation of user interfaces. 1990 Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; April 1-5, 1990; Seattle, WA p. 249-256 URL: https://doi.org/10.1145/97243.97281 [doi: 10.1145/97243.97281]
36. Lai TY, Bakken S. Heuristic evaluation of HIV-TIDES - tailored interventions for management of depressive symptoms in HIV-infected individuals. AMIA Annu Symp Proc 2006;2006:996 [FREE Full text] [Medline: 17238615]
37. Choi J, Bakken S. Web-based education for low-literate parents in neonatal intensive care unit: development of a website and heuristic evaluation and usability testing. Int J Med Inform 2010;79(8):565-575 [FREE Full text] [doi: 10.1016/j.ijmedinf.2010.05.001] [Medline: 20617546]
38. Alene GD, Ayenew GM, Bantie GM, Zeru T, Yismaw G. Impact of ware on maternal health services in the war affected areas of the Amhara region, ethiopia: disruption of antenatla care, institutional delivery and postnatal services. Ethiopian Journal of Translational Sciences 2023;1(1).
39. Types of color vision deficiency. National Eye Institute. URL: https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/color-blindness/types-color-vision-deficiency#:~:text=The%20most%20common%20type%20of,of%20green%20look%20more%20red [accessed 2025-02-20]
40. Følstad A. Users' design feedback in usability evaluation: a literature review. Hum Cent Comput Inf Sci 2017;7(1). [doi: 10.1186/s13673-017-0100-y]
41. Ergonomics of human-system interaction part 210: human-centred design for interactive systems. International Organization for Standardization. 2010. URL: https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/05/20/52075.html [accessed 2023-08-01]
42. Lau MK, Bounthavong M, Kay CL, Harvey MA, Christopher MLD. Clinical dashboard development and use for academic detailing in the U.S. department of veterans affairs. J Am Pharm Assoc (2003) 2019;59(2S):S96-S103.e3. [doi: 10.1016/j.japh.2018.12.006] [Medline: 30713078]
43. Mashige KP, van Staden DB. Prevalence of congenital colour vision deficiency among Black school children in Durban, South Africa. BMC Res Notes 2019;12(1):324 [FREE Full text] [doi: 10.1186/s13104-019-4374-1] [Medline: 31182158]
44. Color blindness most common in White boys. Review of Optometry. URL: https://tinyurl.com/2ztwwm6w [accessed 2025-02-20]
45. Farzandipour M, Nabovati E, Sadeqi Jabali M. Comparison of usability evaluation methods for a health information system: heuristic evaluation versus cognitive walkthrough method. BMC Med Inform Decis Mak 2022;22(1):157 [FREE Full text] [doi: 10.1186/s12911-022-01905-7] [Medline: 35717183]

XSL•FO
RenderX

46. Banke-Thomas A, Abejirinde IO, Ogunyemi O, Gwacham-Anisiobi U. Innovative dashboard for optimising emergency obstetric care geographical accessibility in Nigeria: qualitative study with technocrats. Health Policy and Technology 2023;12(2):100756 [FREE Full text] [doi: 10.1016/j.hlpt.2023.100756]
47. Wang J, Wong KLM, Olubodun T, Gwacham-Anisiobi U, Ogunyemi O, Afolabi BB, et al. Developing policy-ready digital dashboards of geospatial access to emergency obstetric care: a survey of policymakers and researchers in sub-Saharan Africa. Health Technol 2023;14(1):69-80. [doi: 10.1007/s12553-023-00793-9]
48. Cobo B, Castillo E, López-Torrecillas F, Rueda MDM. Indirect questioning methods for sensitive survey questions: modelling criminal behaviours among a prison population. PLoS One 2021;16(1):e0245550 [FREE Full text] [doi: 10.1371/journal.pone.0245550] [Medline: 33481882]

Abbreviations

BEmOC: basic emergency obstetric care
HCI: human-computer interaction
HIPAA: Health and Insurance Portability Accountability Act
HIT: health information technology
IPLS: Integrated Pharmaceutical Logistics System
IRB: institutional review board
SME: subject matter expert

Edited by A Kushniruk; submitted 16.07.24; peer-reviewed by M Gasmi , A Banke-Thomas; comments to author 20.12.24; revised version received 21.01.25; accepted 10.02.25; published 03.04.25.

Please cite as:
Dougherty K, Tesfaye Y, Biza H, Belew M, Benda N, Gebremariam Gobezayehu A, Cranmer J, Bakken S
User-Centered Design of an Electronic Dashboard for Monitoring Facility-Level Basic Emergency Obstetric Care Readiness in Amhara, Ethiopia: Mixed Methods Study
JMIR Hum Factors 2025;12:e64131
URL: https://humanfactors.jmir.org/2025/1/e64131
doi:10.2196/64131
PMID:

©Kylie Dougherty, Yihenew Tesfaye, Heran Biza, Mulusew Belew, Natalie Benda, Abebe Gebremariam Gobezayehu, John Cranmer, Suzanne Bakken. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 03.04.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Co-Designing a Web-Based and Tablet App to Evaluate Clinical Outcomes of Early Psychosis Service Users in a Learning Health Care Network: User-Centered Design Workshop and Pilot Study

Kathleen E Burch[1], BA; Valerie L Tryon[1], PhD; Katherine M Pierce[1], PhD; Laura M Tully[1], PhD; Sabrina Ereshefsky[1], PhD; Mark Savill[1,2], PhD; Leigh Smith[3], PhD; Adam B Wilcox[4], PhD; Christopher Komei Hakusui[1], BA; Viviana E Padilla[1], BA; Amanda P McNamara[5], MPH; Merissa Kado-Walton[5], MPH; Andrew J Padovani[6], PhD; Chelyah Miller[1], BA; Madison J Miles[1], BA; Nitasha Sharma[1], BA; Khanh Linh H Nguyen[1], MS; Yi Zhang[6], PhD; Tara A Niendam[1], PhD

[1]Department of Psychiatry & Behavioral Sciences, University of California-Davis, Sacramento, CA, United States

[2]Department of Psychiatry and Behavioral Sciences, Weill Institute for Neuroscience, University of California-San Francisco, San Francisco, CA, United States

[3]Department of Psychology, University of California-Davis, Davis, CA, United States

[4]Institute for Informatics, Washington University in St. Louis, St. Louis, MO, United States

[5]Herbert Wertheim School of Public Health and Human Longevity Science, University of California San Diego, San Diego, CA, United States

[6]Center for Healthcare Policy and Research, University of California-Davis, Sacramento, CA, United States

Corresponding Author:
Kathleen E Burch, BA
Department of Psychiatry & Behavioral Sciences
University of California-Davis
4701 X St
Sacramento, CA, 95817
United States
Phone: 1 916 699 5193
Email: knye@ucdavis.edu

Abstract

Background: The Early Psychosis Intervention Network of California project, a learning health care network of California early psychosis intervention (EPI) programs, prioritized incorporation of community partner feedback while designing its eHealth app, Beehive. Though eHealth apps can support learning health care network data collection aims, low user acceptance or adoption can pose barriers to successful implementation. Adopting user-centered design (UCD) approaches, such as incorporation of user feedback, prototyping, iterative design, and continuous evaluation, can mitigate these potential barriers.

Objective: We aimed to use UCD during development of a data collection and data visualization web-based and tablet app, Beehive, to promote engagement with Beehive as part of standard EPI care across a diverse user-base.

Methods: Our UCD approach included incorporation of user feedback, prototyping, iterative design, and continuous evaluation. This started with user journey mapping to create storyboards, which were then presented in UCD workshops with service users, their support persons, and EPI providers. We incorporated feedback from these workshops into the alpha version of Beehive, which was also presented in a UCD workshop. Feedback was again incorporated into the beta version of Beehive. We provided Beehive training to 4 EPI programs who then piloted Beehive's beta version. During piloting, service users, their support persons, and EPI program providers completed Beehive surveys at enrollment and every 6 months after treatment initiation. To examine preliminary user acceptance and adoption during the piloting phase, we assessed rates of participant enrollment and survey completion, with a particular focus on completion of a prioritized survey: the Modified Colorado Symptom Index.

Results: UCD workshop feedback resulted in the creation of new workflows and interface changes in Beehive to improve the user experience. During piloting, 48 service users, 42 support persons, and 72 EPI program providers enrolled in Beehive. Data were available for 88% (n=42) of service users, including self-reported data for 79% (n=38), collateral-reported data for 42% (n=20), and clinician-entered data for 17% (n=8). The Modified Colorado Symptom Index was completed by 54% (n=26) of

XSL•FO
RenderX

service users (total score: mean 24.16, SD 16.81). In addition, 35 service users had a support person who could complete the Modified Colorado Symptom Index, and 56% (n=19) of support persons completed it (mean 26.71, SD 14.43).

Conclusions: Implementing UCD principles while developing the Beehive app resulted in early workflow changes and produced an app that was acceptable and feasible for collection of self-reported clinical outcomes data from service users. Additional support is needed to increase collateral-reported and clinician-entered data.

(JMIR Hum Factors 2025;12:e65889) doi:10.2196/65889

KEYWORDS

eHealth; user-centered design; learning health system; psychosis; early psychosis; user-driven development; web-based; data visualization; surveys and questionnaires; measurement-based care

Introduction

Background

Research estimates that the lifetime prevalence of psychotic disorder diagnoses is approximately 1.5%, and the prevalence of psychotic symptoms is between 4.2% and 17.5% [1]. California, the most populous and second most diverse state in the United States [2], had a population of 39.11 million in 2023, according to the California Census Bureau. This indicates that 586,650 to 6.8 million Californians may experience psychosis symptoms in their lifetime. In response to this, many California counties have developed specialty early psychosis intervention (EPI) services, which vary widely in their implementation approach [3]. The Early Psychosis Intervention Network of California (EPI-CAL) [4] was developed to support the provision of quality EPI services and to create an infrastructure to conduct standardized measurement of the impact of early psychosis care delivery. To support this goal, the EPI-CAL team, in collaboration with several California counties, developed a learning health care network (LHCN) consisting of EPI programs across the state. The EPI-CAL LHCN later joined the national Early Psychosis Intervention Network [5], which allowed additional California EPI programs to participate. Members of the LHCN agreed to gather standardized information and outcomes from their clinics as part of measurement-based care (MBC). Collecting this information is critical to support quality early psychosis care provision within clinical programs as well as enhance statewide learning and development. For example, a narrative review of an MBC approach in behavioral health clinics found such benefits as significantly improving clinical outcomes, improving symptoms more quickly, and decreasing treatment costs [6].

To this end, the EPI-CAL team chose to design and implement a web-based and tablet app called Beehive. We chose the name Beehive to reflect the envisioned purpose for the app: to help LHCN programs learn and grow together for the betterment of the collective, just as bees work together to build a hive for the benefit of the colony. Beehive is a robust, stand-alone eHealth app for use by service users, their support persons, and EPI program providers. In this text, we use the term *service user* to refer to the individual with a psychosis diagnosis who is receiving mental health care from an early psychosis program. We use the term *support person* to refer to any person that the service user has chosen to involve in their care. This is typically the individual's parent but might be another family member, a friend, a partner, or some other close relative. Beehive's purpose is to promote MBC in EPI programs by standardizing data collection across a network of programs focusing on community partner priorities; supporting key components of care such as assessment, safety monitoring, and ongoing care delivery; supporting program-level management of care; and aggregating data across a large network to support evaluation and research at state and even national levels [4].

We chose an eHealth approach to implementing MBC due to its appropriateness for the EPI setting and its potential benefits. Despite the perceived challenges related to experiences of suspicion or paranoia, individuals experiencing mental health difficulties, such as schizophrenia and bipolar disorder, have widely found use of eHealth both feasible and acceptable [7-10]. Furthermore, the use of eHealth can support the advancement of MBC. For example, eHealth apps have been previously demonstrated to promote symptom and outcomes monitoring in both early psychosis care [11,12] and LHCNs [13]. Conducting MBC with eHealth enhances its benefits as it allows for data collection to be standardized across programs and instantly available. For example, MBC may promote collaboration across a care team [14-16], which is relevant for EPI programs for which the evidence-based treatment, coordinated specialty care, is inherently team based [17]. Use of eHealth to collect data in this setting allows data collected by one team member or entered by a service user to be instantly available for all team members. eHealth also enhances the benefits of MBC through data aggregation, which enables evaluation of program performance [16,18,19] or promotion of evidence-based treatments [16,20].

Though there are many potential benefits, there are also numerous barriers to implementation of both new eHealth technology and MBC. According to a systematic literature analysis, the top factors posing a barrier to eHealth app implementation include lack of digital health literacy, lack of devices, financing issues, service-user cognition, and security [21]. These barriers contribute to low adoption and user acceptance, which limit the success of implementation [22,23]. Barriers to implementation of MBC include training burden, concern that negative feedback causes harm to service users, and the time required for survey completion [24-26].

To pursue the benefits of using eHealth to implement MBC and mitigate the potential barriers, we developed Beehive with user-centered design (UCD) principles. UCD prioritizes the needs and expectations of the end user [27,28]. UCD approaches include dedicated design activities, active involvement of users

XSL•FO
RenderX

in the design process, incorporation of their feedback, prototyping, and continuous evaluation [29], which can address low user acceptance and low adoption [30]. Beehive's iterative development began with a collaborative process with service users, support persons, and EPI program providers to identify and prioritize which outcome measures should be collected in the app [31]. We also explored how service users, support persons, and EPI program providers wanted to be informed about data-sharing options in Beehive and built Beehive's end user license agreement (EULA) workflow to incorporate user perspectives [32]. With the survey content and EULA workflow finalized, we moved onto the development of the user-facing parts of Beehive.

Objectives

In this study, we aimed to (1) use UCD principles to create a co-designed web-based and tablet app, called Beehive, to support MBC in EPI programs, and (2) assess Beehive's initial feasibility in clinical settings by piloting it in 4 EPI programs.

Methods

Design

To promote Beehive engagement across multiple types of users and across multiple domains of engagement, we integrated UCD principles of incorporation of user feedback, prototyping, iterative design, and continuous evaluation. Service users, support persons, and EPI program providers had multiple opportunities to provide feedback, which was incorporated throughout Beehive development. Figure 1 shows the study design from conceptualization through data collection.

Figure 1. Study design for Beehive development.



The development process began with conceptualization of user journeys. User journey mapping envisions how specific types of users, such as a service user or a service provider, will interact with an app from access point through all required activities [33]. User journey mapping allowed us to identify which storyboards we should develop to present in UCD workshops and which user-types we needed to recruit for those workshops. Storyboards are a tool to visualize app workflows and the user interface [34]. We developed storyboards to present as prototypes in UCD workshops so that feedback could modify the app design before time was invested in coding the alpha version of the app. The alpha version of the app included core workflows and was both evaluated internally and presented in another UCD workshop to gather more feedback before coding the beta version of the app. The beta version of the app incorporated remaining feedback from storyboard workshops, new feedback from the alpha stage, and added the remaining core functionality that was not in-scope for the alpha version (eg, reports). The beta version of Beehive was piloted by 3 EPI programs over 6 months to further refine the app before launching it across all EPI-CAL programs. We used pilot data to assess initial use and uptake of Beehive's beta version.

This UCD approach allowed the EPI-CAL team to receive and incorporate feedback during conception, design, and testing phases of eHealth app development, and include multiple perspectives to facilitate user engagement in eHealth. Notably, UCD has also been demonstrated to increase eHealth adoption and user acceptance in research and clinical settings [30,35].

Participant Recruitment

UCD Workshops

For UCD workshops of the Beehive storyboard, we recruited participants from the following three EPI community partner

XSL•FO
RenderX

groups: (1) EPI service users, (2) their support persons, and (3) EPI providers. Eligible participants were (1) actively or formerly affiliated with an EPI-CAL EPI program, (2) English speaking, and (3) able to provide written informed consent and assent (minors or conserved adults). EPI program providers were recruited through contact with the team lead of the 12 active EPI-CAL EPI programs. Service users and support person participants were invited either through EPI program provider referral or by the research team directly contacting individuals who had previously consented to be contacted for future research opportunities.

One EPI-CAL clinic agreed to participate in a workshop for the alpha version of the app to support the refinement of Beehive before piloting.

Piloting

In total, 4 EPI-CAL clinics agreed to participate in 6 months of Beehive beta testing before Beehive's full launch across the entire EPI-CAL LHCN. During this period, programs integrated the Beehive app into standard clinical care. Pilot sites registered service users who were active in their program at the time of launching Beehive and new service users who started after the launch. Each participating program has different acceptance criteria for service users, and this has been described in a separate protocol paper for the EPI-CAL study (NCT04007510) [4]. Service users and support persons were excluded from piloting if they did not speak English because Beehive beta version was only available in English. At their first point of contact with Beehive, service users and primary support persons completed the Beehive EULA and were asked if they gave permission for their clinical data collected in Beehive to be used for research purposes [32]. We trained clinics to involve the legal guardians of service users aged <18 during EULA completion, and these service users were required to have a primary support person registered in Beehive. Individuals could update their data-sharing permissions at any time.

Methods

User Journey Mapping

The EPI-CAL research team worked collaboratively with the contracted app developer in the user journey mapping and storyboard design phases for the Beehive tablet and web apps. Three primary user groups with distinct roles were identified: (1) service users or support persons, (2) direct-service providers, and (3) program administrators. Beehive user journeys were developed for each group.

For all user groups, user journeys were designed to guide individuals smoothly through Beehive onboarding and account creation to the EULA explainer video detailing the types of information Beehive collects, who can access their data, and how to select their preferred permissions for who can view their data [32]. Users then choose their data-sharing permissions.

Beehive then presents service users and support persons with a series of one-time and longitudinal surveys to measure clinical outcomes at 6-month intervals. Specific survey items associated with risk, such as suicidal or homicidal ideation, send real-time alerts to EPI program providers if they are endorsed by service users. Beehive creates visualizations of this survey data. While service users and support persons cannot independently access survey visualizations, we considered this is as part of their user journey because it should be shown to them by EPI program staff as part of regular care.

EPI program provider user journeys were designed to facilitate easy management of service user records and smooth navigation to service user data. A dashboard presents the most important information to users, such as outstanding survey alerts or other action items. A client list presents all registered service user records in list format with the most relevant information displayed. EPI program provider users can click into service user records to view survey results, survey visualizations, and complete provider-entered surveys. Providers can display survey results and survey visualizations as part of ongoing care with service users and their support persons to facilitate understanding and coordinate treatment priorities.

Administrator user journeys were designed to promote clinic-level management tasks, such as Beehive implementation and quality assurance. Administrator dashboards present aggregate-level information of survey data and allow for the ability to compare clinic averages to the average of the EPI-CAL LHCN. Administrator dashboards also present summaries of Beehive activity across the clinic. Administrator users can also download reports, including survey data reports, to use their clinic data for quality assurance or reporting requirements.

UCD Workshops

Next, the development team created dynamic storyboards of the above user journeys. These storyboards were presented to community partners from early psychosis clinics in 90-minute UCD workshops. Figures 2 and 3 are images from the storyboards presented to community partners.

In the storyboard workshops, we presented major features of the app and asked for feedback on the app's look and feel, as well as functionality as it related to existing clinical workflow, and ease of use and acceptability for service users, their support persons, and EPI program providers. Because we determined 3 user-types during user journey mapping, we held 3 different types of storyboard workshops tailored to the journeys of these users: service users or support persons, direct-service providers, and program administrators. If EPI providers had roles in both direct-service provision and program administration, they could attend both groups. UCD workshops were conducted through videoconferencing to comply with the COVID-19 social distancing restrictions. Each session was audio recorded and included 2 facilitators (KEB, LMT, TAN, or SE) and a notetaker (VLT). There were no other individuals present other than researchers and participants. The notetaker took detailed notes, as close to verbatim as possible. Audio recordings were used as a reference to fix any portions of the notes which were not clear.

After storyboard workshops, we integrated feedback to make design changes to the alpha version of the app. During a 90-minute alpha testing workshop, we solicited feedback on the alpha app, with a special emphasis on how compatible it was with the existing clinical workflows. We created test

XSL•FO
RenderX

accounts for each participant and had them complete various core workflows in the app, such as registering a service user, completing surveys, and reviewing data visualizations. This workshop included one facilitator (KEB) and one notetaker (LS). After the conclusion of all workshops, we continued to integrate feedback to make design changes in the beta version of the app.

Figure 2. Survey item in storyboard.

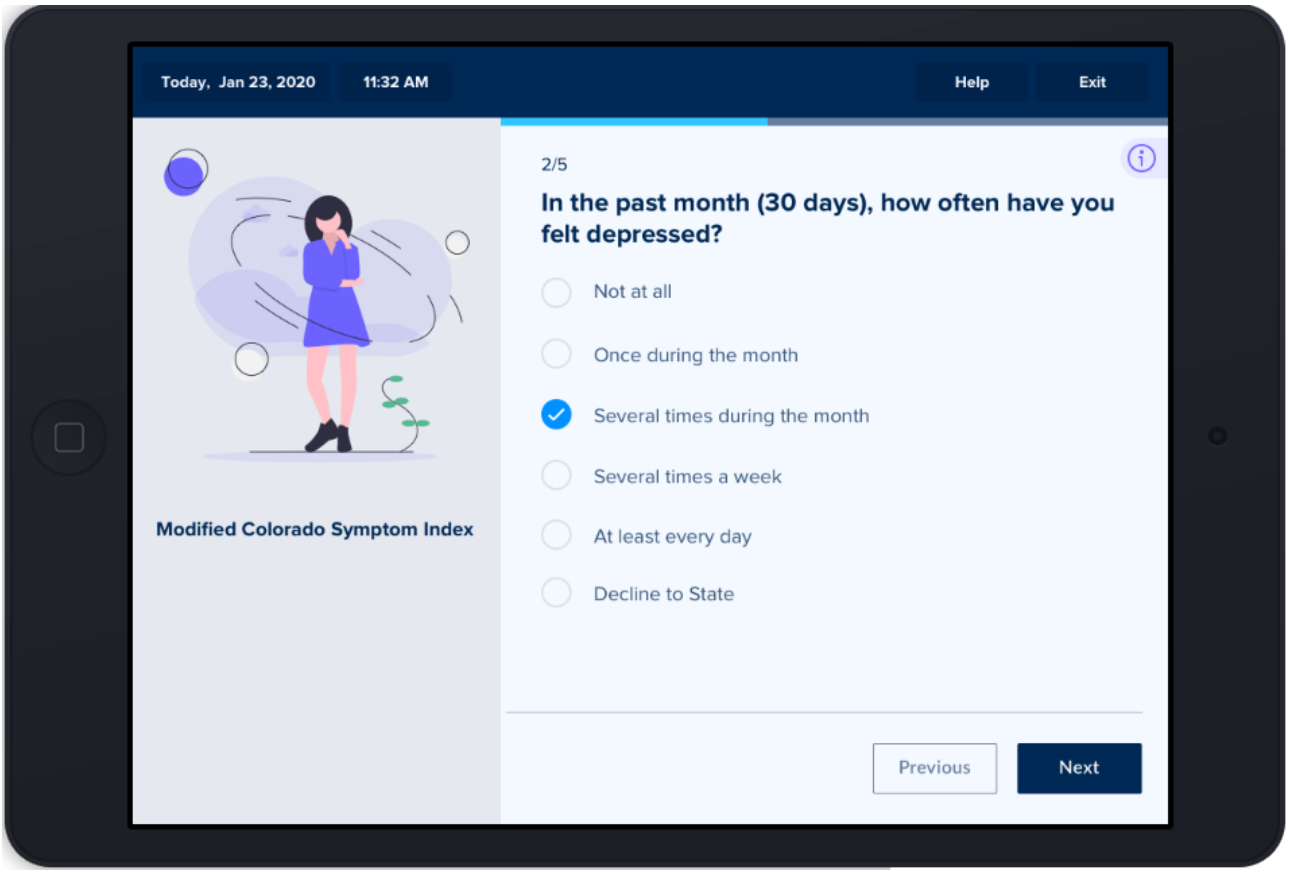
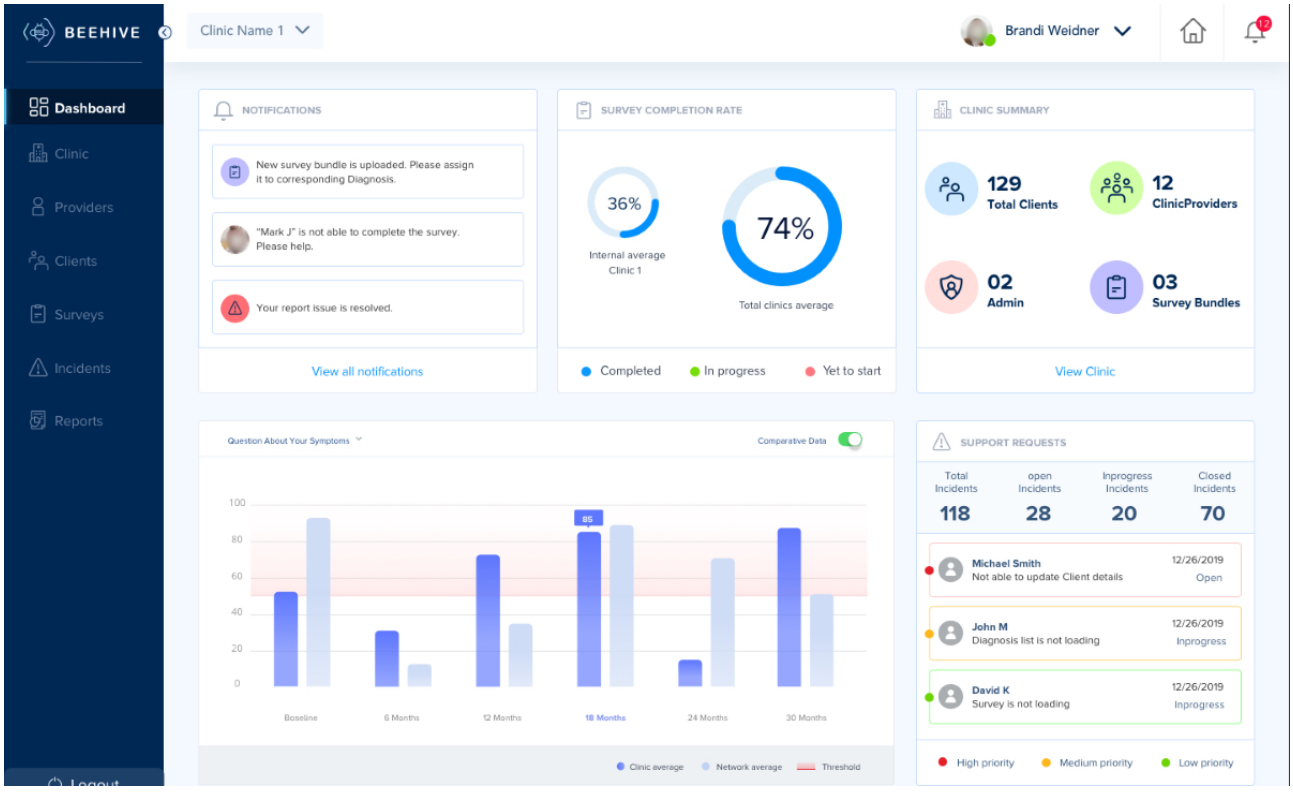


Figure 3. Clinical administrator dashboard in storyboard.



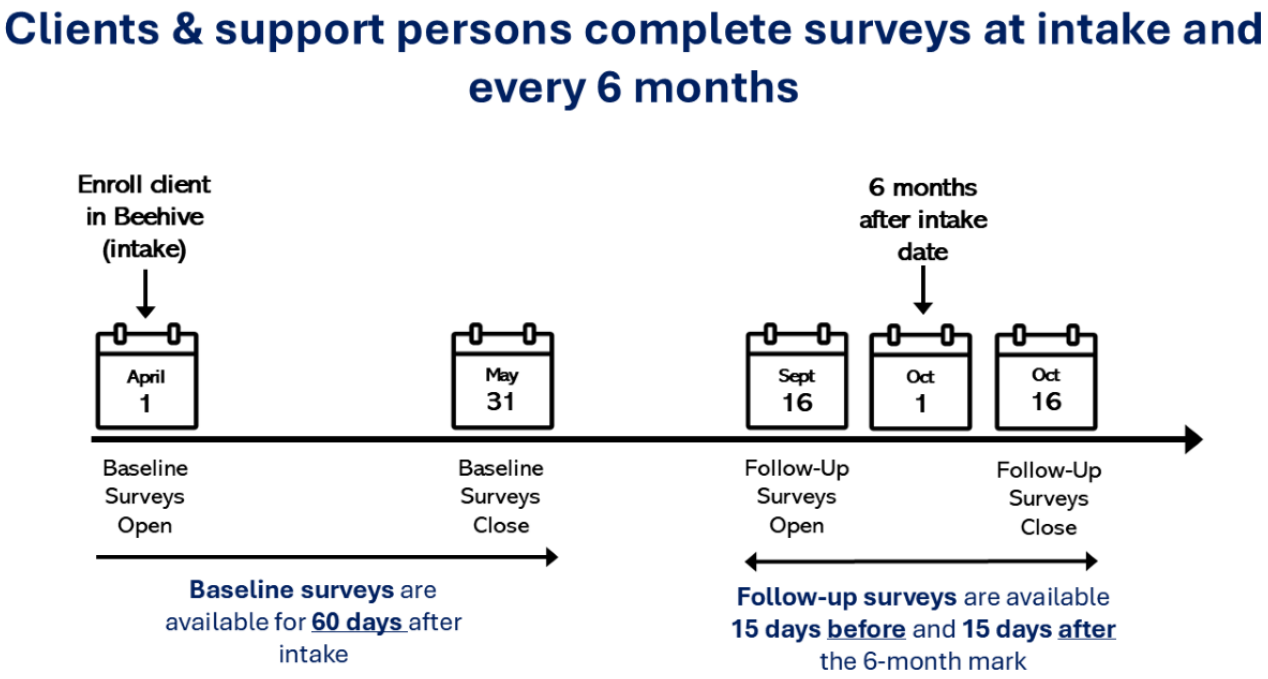
XSL•FO
RenderX

Piloting

Finally, piloting of the beta version of Beehive was conducted over a 6-month period with EPI-CAL LHCN programs that were identified as pilot sites. We provided training to each pilot site and assigned them a point person from our research team to provide regular support, troubleshoot implementation challenges, and escalate app bugs and implementation barriers. We trained all program staff, regardless of clinical role. The training series showed users how to complete key Beehive workflows, introduced the EPI-CAL core assessment battery, and included activities on how to interpret data visualizations. We also met with key staff at each program to support them to devise a plan to integrate Beehive into their existing clinical workflows, such as how to integrate service user registration during the clinical intake process. The full description of this training and support is described in a separate paper [4]. During this training and support process, we received informal feedback from program staff about implementation successes and challenges.

During piloting, programs integrated the Beehive app into standard clinical care for all service users. Service users and their support persons were registered in Beehive web app by staff at their EPI program. Program staff entered the service users' treatment start date which was considered the start of the service users' baseline window in Beehive. At enrollment, baseline, and every 6 months thereafter, surveys were assigned to service users, their support person, and their provider. One-time surveys assessing lifetime experiences were available for all respondents at enrollment. Service users had 3 enrollment surveys, support persons had 1, and clinicians had 1. Longitudinal surveys were assessed at treatment baseline and every 6 months thereafter. Service users had 17 longitudinal surveys, support persons had 6 surveys, and clinicians had 9 surveys. The duration of baseline survey window was 60 days; the duration of follow-up survey windows was 30 days (15 days before and after the target completion date). Figure 4 shows a Beehive training slide with a visualization of survey windows.

Figure 4. Beehive training slide showing survey windows during piloting.



Programs were instructed to enroll all service users, regardless of how long they had been affiliated with the EPI program. Therefore, some individuals may have been enrolled in Beehive after their baseline window had closed, and their first time point during piloting may have been a 6-month time point or a 12-month time point. Service users and support persons could access and complete surveys in-person at the clinic on the tablet app or they could complete them remotely via "web link." The "web link" was a unique link that was texted or emailed to them weekly during survey windows if surveys were not fully completed. Surveys accessed via web link could be completed on any personal device that had access to the internet and a web browser. Service providers completed surveys and could review data on the web app. Baseline surveys were intended to support the clinical intake process, including initial assessment and collaborative treatment planning. Follow-up surveys were intended to support ongoing assessment, adjustments to treatment planning, and monitoring of treatment goals.

The EPI-CAL core assessment battery, including how it was created and all included measures, is described in a separate paper [4]. Briefly, Beehive survey content includes both the Early Psychosis Intervention Network core assessment battery [36] and additional measures based on EPI community partner feedback determined in earlier qualitative work for the EPI-CAL study [4,31]. A table of measures is included in the Multimedia Appendix 1 [4]. One survey of particular interest is the Modified Colorado Symptom Index (MCSI) [37], a measure central to the aims of the broader EPI-CAL study and which we asked sites to prioritize [4]. The MCSI is a 14-item, self-report scale

XSL•FO
RenderX

which measures the frequency of psychiatric symptoms, including symptoms of mood, psychosis, cognition, forgetfulness, and risk to self and others. Respondents indicate frequency of symptoms over the past 30 days on a 0 to 4 scale of "not at all" to "at least every day." Total scores range between 0 and 56, with higher scores indicating higher frequency and number of psychiatric symptoms.

Data Analysis

In UCD workshops, we asked highly structured questions to solicit feedback on the storyboard and alpha version of Beehive. Subsequently, we organized our data categories relevant to the workflows and features we were evaluating. We then organized comments by whether they were supportive of the existing features or critical and asked for change so that we could focus on what features to move forward and what features to change as we created the alpha and beta versions of Beehive.

To investigate the initial feasibility of Beehive in EPI-CAL clinics, we reviewed descriptive statistics of pilot participants, including registration, enrollment, participant characteristics, and survey completion. Engagement with surveys and survey completion were examined in three ways as follows: (1) determine the proportion of service users for whom any data were entered, regardless of respondent type, (2) determine the rate of survey completion across all available surveys, and (3) evaluate whether participants completed all, partial, or no surveys across survey time points during piloting phase. Partial survey completion indicates that the respondent completed at least 1 survey, but did not complete all of their surveys in the specified time point. We also evaluated completion of MCSI because it is a measure that we asked programs to prioritize.

Ethical Considerations

The institutional review board of the University of California, Davis, approved the study (1403828-21, California Collaborative Network to Promote Data-Driven Care and Improve Outcomes in Early Psychosis). In addition, several of the counties and universities with a program participating in EPI-CAL required a separate review and approval of the project by their institutional review board. All study participants provided written informed consent and assent (as appropriate). Participants received US $30 compensation for each workshop they participated in. Participants in the piloting phase were not compensated because integration of Beehive was part of routine care in the EPI program. Audio recordings UCD workshops include voice print identifiers and are stored in compliance with University of California, Davis HIPAA (Health Insurance Portability and Accountability Act) policies and procedures. Data collected during piloting for research includes limited identifiers including zip code, dates of service, and month and year of birth. Only trained research staff with a need-to-access have access to identifiable data.

Results

UCD Workshops

We conducted 14 storyboard workshops with 77 total participants between April 3, 2020, and August 28, 2020. In total, 4 workshops were with service users (n=8, 10%) and their support persons (n=9, 12%). In addition, 10 workshops were with EPI program providers (n=60, 78%), including 6 for service providers, 3 for administrators, and 1 for both service providers and administrators. Demographics for workshop participants are provided in Table 1.

We completed an interim analysis of storyboard workshop data in May 2020. We completed the final analysis of workshop data in August 2020, after all groups were completed. After each analysis, we discussed and synthesized the feedback for the developers to support app development. We attempted to balance the needs of all types of participants. However, if there were needs or feedback in direct contrast with one another, we prioritized service-user feedback due to our value of centering service-user feedback in this app. This feedback and the action taken to address it are summarized in Table 2.

We conducted 1 alpha workshop in October 2020 with 4 EPI program provider participants. Feedback from this workshop was analyzed in October 2020. During this workshop, participants identified a few bugs in the app, but their feedback primarily focused on ideas for integrating Beehive into clinical workflows. For example, they suggested that Beehive training should include best practices for how providers can review the data, engage with the data, and make the most out of Beehive. They also shared concerns about using technology in telehealth settings. For example, switching to telehealth in response to the COVID-19 pandemic had been very difficult for some families, and they predicted those same families would find using Beehive challenging if the clinic could not meet with them in-person to teach them how to use it. They were less concerned about service users and support people using Beehive on a tablet in the clinic where they could provide in-person support. Finally, participants brought up the importance of shifting the culture of clinics to view data collection as an important part of treatment, not just an extra task where information is being extracted from service users. For example, surveys should be directly related to service-user recovery goals. Participants discussed how visualizations could be used to demonstrate the clinical utility of gathering these data. For example, 1 clinician said they would want to use the graphs to point out the way a service user is improving or doing better and that they would want to highlight their strengths. Another participant cautioned that some service users may not want to look at data visualizations and that this should be an optional part of their care. Figures 5 and 6 illustrate design changes present in the beta version of app after the conclusion of all workshops.

XSL•FO
RenderX

Table 1. Demographics of user-centered design workshops[a].

	Service users (n=8)	Support persons (n=9)	EPI[b] program providers (n=60)
Clinic type, n (%)			
Medi-Cal	7 (88)	6 (67)	30 (50)
Private insurance	<5 (<63)	<5 (<55)	30 (50)
Age (y), mean (range)	22.50 (16-33)	41.50 (14-60)	36.25 (26-50)
Sex at birth, n (%)			
Female	<5 (<63)	8 (89)	43 (72)
Male	6 (75)	<5 (<55)	17 (28)
Gender, n (%)			
Female	<5 (<63)	8 (89)	43 (72)
Male	6 (75)	<5 (<55)	16 (27)
Nonbinary	<5 (<63)	—[c]	—
Missing	—	—	<5 (<8)
Race, n (%)			
African or African American or Black	<5 (<63)	—	5 (8)
Asian	—	—	6 (10)
White or Caucasian	<5 (<63)	6 (67)	33 (55)
Other	<5 (<63)	<5 (<55)	10 (17)
More than one	<5 (<63)	<5 (<55)	<5 (<8)
Missing	—	<5 (<55)	<5 (<8)
Ethnicity, n (%)			
Latinx	<5 (<63)	5 (56)	26 (43)
Not Latinx	5 (63)	<5 (<55)	33 (55)
Missing	—	—	<5 (<8)
Sexual orientation, n (%)			
Bisexual or gay or lesbian	<5 (<63)	—	5 (8)
Heterosexual	5 (63)	9 (100)	54 (90)
Other	<5 (<63)	—	<5 (<8)

[a]Cells with fewer than 5 individuals are masked to protect the identity of participants.

[b]EPI: early psychosis intervention.

[c]Not available.

XSL•FO
RenderX

Table 2. Implementation of feedback from user-centered design workshops.

Problem or need identified in workshop	Solution implemented in alpha version
The color scheme and layout seemed "overly clinical"	Brought in more color into the palette and added icons for visual information
Some important aspects of the user-interface were too subtle, such as the survey progress bar or the urgent clinical issues widget	Changed color and design to make it more prominent
Service-user and support person registration were only available as self-registration and could not be completed by EPI program providers	Added this workflow to the web app so that EPI program providers may complete it in advance of service users engaging with surveys
Clinic-level data for service-user demographics was not visualized	Added clinic-level visualizations for race, ethnicity, sex, gender identity, and other demographic metrics
Not all service users or support persons wanted to see score thresholds or comparative data on clinical measures, but some did	Added a toggle to individual-level visualizations so that users can turn on the threshold information or comparative data if they want to see it, or turn it off if they do not want to see it
Service users might have differed on which individual-level survey visualization they wanted to see by default on their data view page	Added feature that allows users to set which measure displays by default for each service user
Some language used in the app needed to be clarified for users to understand what data was being collected or how certain features worked	Changed "homeless" to "without a permanent address" when assessing housing status; changed "help" to "Ask for help" to make it clearer that selecting button will alert the EPI program provider; changed "Diagnosis" to "Primary diagnosis"
Different programs used different words to refer to service users, and individuals might have varied on their preference for what word to use regardless of what their program tended to use	Wherever possible, implemented dynamic text so the service user's preferred name shows throughout the app, rather than any specific word to denote "service user"
Program staff wanted to see overall progress on completion of all surveys at any given time point	Added in visual indicator to show survey completion across multiple surveys (not just while completing one individual survey)
When visualizing a survey, they wanted to have more than just the global score visualized. Also, they wanted a visualization that showed responses to individual items	Added in a visualization that shows individual items as well as the global score
Service users and support persons might not have preferred the official names for measures and might have preferred a more simplified title	Added the ability to enter a display name for surveys (eg, "Family Impact" instead of "Burden Assessment Scale")
Early psychosis intervention program providers needed a way to see both the official measure name as well as the display name	Added a hover modal on survey titles to show the display name for the survey
The provision of clinic services might have been fully remote for the foreseeable future, and the current design of Beehive only allowed service users and support persons to complete surveys on a tablet in-person at the clinic	Design a web link solution which allows service users and support persons to answer surveys remotely. A link to complete their surveys can be emailed or texted to them

XSL•FO
RenderX

Figure 5. Survey item in beta.

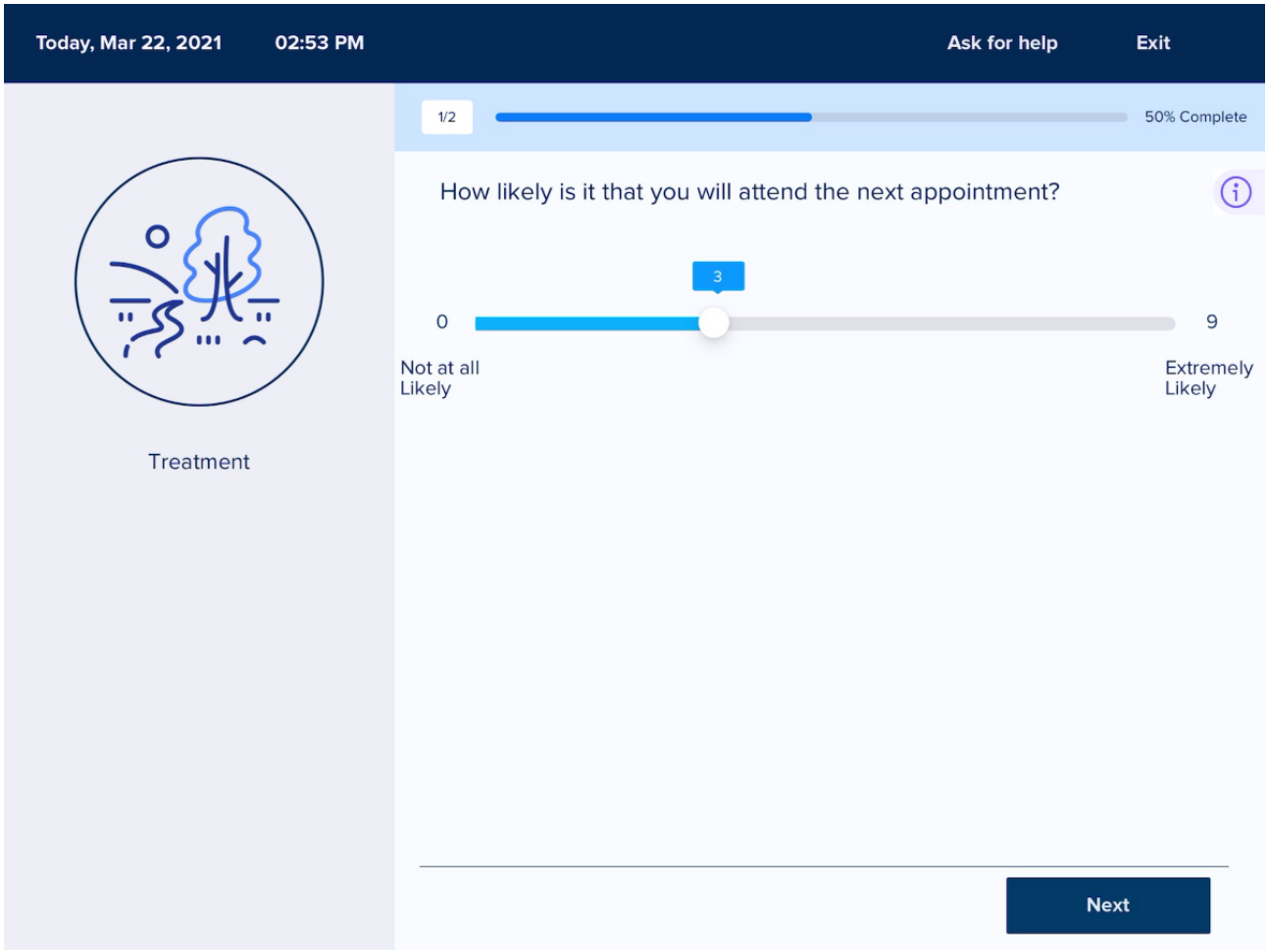
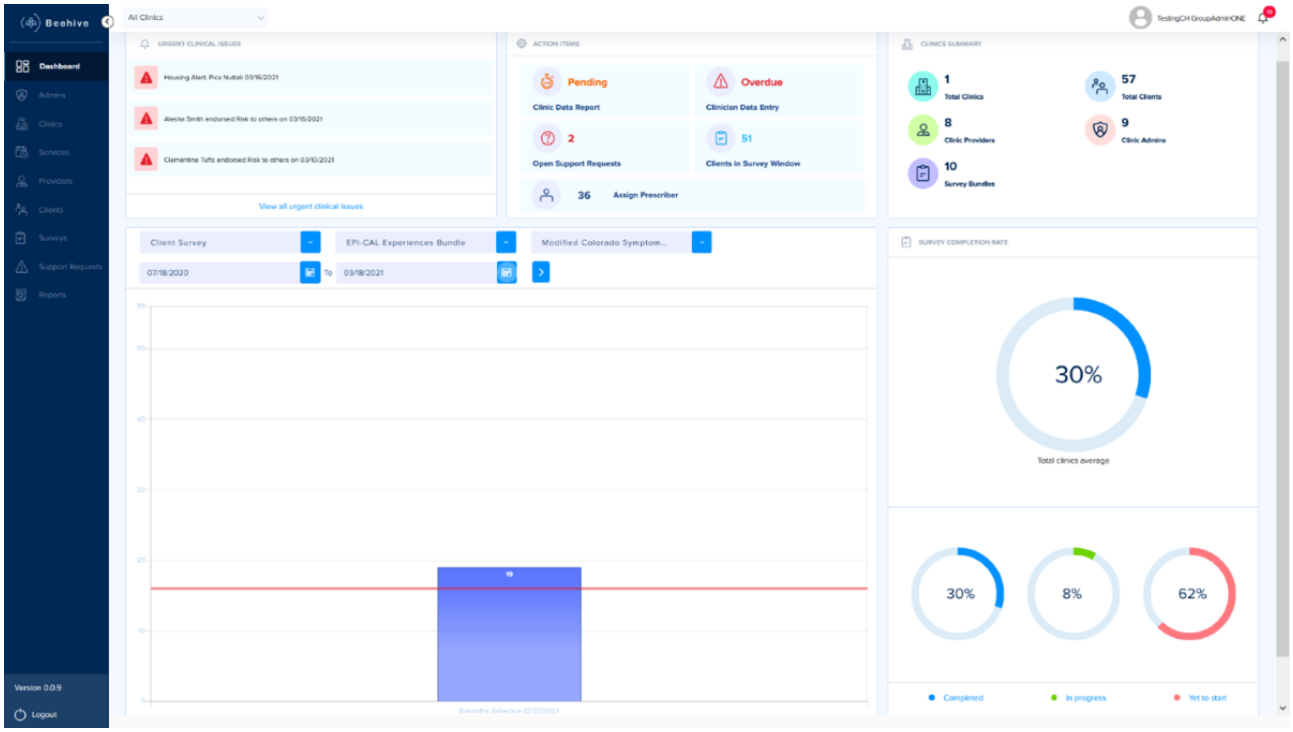


Figure 6. Clinical administrator dashboard in beta.



XSL•FO
RenderX

Piloting

We conducted piloting of Beehive beta app between March 2021 and September 2021. Our training and ongoing support of pilot sites allowed us to gather informal feedback about both the training and the Beehive app. We used this feedback to make adjustments in real time, when possible, or to plan for future changes to Beehive.

We made real-time changes to the training approach in response to program needs based on our observations and their feedback. At the time of training pilot sites in early 2021, these EPI programs were navigating constant uncertainty related to the COVID-19 pandemic, including influx of service users, uncertainty about work location, reduced workforce, etc. In response to this environment, we found it necessary to ask sites to focus on small implementation steps even though we trained them on all available workflows. For example, we asked pilot sites to initially focus on engaging service users and their support persons to complete enrollment and complete surveys. When that was mastered, we asked them to focus on engaging new service users and support persons with Beehive during their clinical intake process. Finally, toward the end of the piloting period, we encouraged them to focus on entering EPI program provider-entered data. Even if pilot sites registered existing service users, they were asked to set the service user's survey baseline date to align with their start in the EPI program. Therefore, some service users may have never been assigned surveys during their baseline survey window.

During the piloting phase, 93 service users, 78 support persons, and 86 service providers were registered across 4 clinics. Of the 93 service users who were registered into Beehive by their program, 59 (63%) completed the Beehive EULA during piloting, including 48 (51%) individuals who gave permission to use their data for research. Of the 78 support persons registered to a service user in Beehive, 52 (66%) completed the Beehive EULA, including 42 (54%) individuals who gave permission to use their data for research. Of these 42 support persons, 5 were excluded from analysis because the service user they were registered with did not give permission to use data in research, and we prioritize the data-sharing decision of the service user regarding use of collateral data for research purposes.

While most users entered at least one survey window during the piloting phase, 3 were discharged from their program before longitudinal surveys were available. Of the 86 service providers registered, 78 (91%) completed the Beehive EULA, including 72 (84%) individuals who gave permission to use their data for research. This information is presented in Figure 7.

Participant demographics and EPI program provider professional background are provided for individuals who agreed to share their data for research purposes in Tables 3 and 4. Of note, age is missing for some EPI program providers and support persons. This data was collected during registration, and this field was not included in the first release of the beta version of the app. Therefore, some users were not able to complete this field during registration and did not return later to update it, yielding missing data for 15 (36%) support persons and 37 (51%) EPI program providers.

First, to examine engagement during piloting, we assessed the number of service users that had entered any data that could be used in care. During piloting, respondents entered survey data for 85% (n=41) of service users. This includes self-reported data for 75% (n=36), collateral-reported data for 42% (n=20), and EPI program provider–entered data for 17% (n=8).

Second, we evaluated survey completion across the total amount of available surveys. A total of 1517 surveys were assigned across all respondent types during piloting and 35.4% (n=537) of those surveys were completed across all time points. We also evaluated survey completion by respondent type. Across all time points, service users completed 396 (49.4%) of 802 assigned surveys. Support persons completed 113 (47.5%) of 238 assigned surveys. EPI program providers completed 28 (5.9%) of 477 assigned surveys.

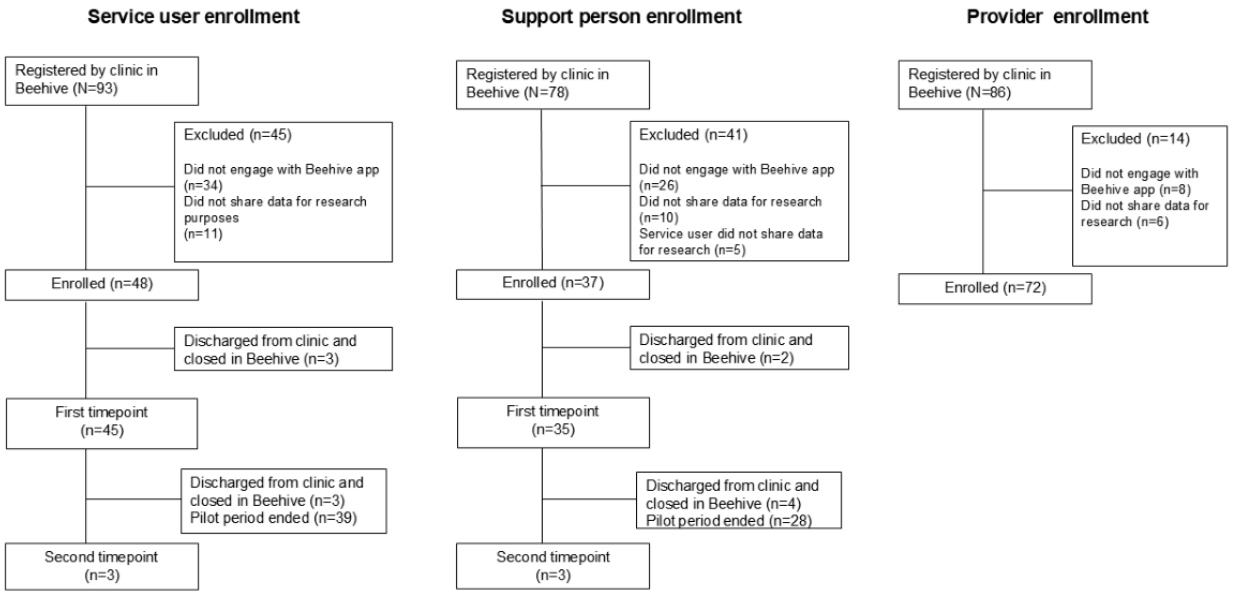
Finally, we evaluated how many respondents completed all, partial, or no surveys at each time point. Because participants could be enrolled at any point in treatment, the first time point for a service user may not have been their "baseline" appointment. These data are presented in Table 5.

The MCSI was completed by 54% (26) of service users, and 7 were excluded for responses of "prefer not to say" (Total Score: mean 19.58, SD 16.81). Additionally, 35 service users had a support person who could complete the Modified Colorado Symptom Index, and 56% (n=19) of support persons completed it. From these 19, 5 were excluded for responses of "prefer not to say" (mean 26.71, SD 14.43).

Through piloting, we also gathered informal feedback about workflows that could be improved in Beehive and that we would address with future change orders to the app after the testing of the beta version of Beehive. For example, we received feedback that the survey windows, initially chosen to mirror the data collection windows of clinical trials, were far too narrow and restrictive for data collection in community mental health programs. Program staff users also indicated that they wanted a better way of seeing a summary of what surveys service users and support persons completed. Service user and support person feedback was relayed to our team via program staff. For example, service users and support persons wanted to customize the day and time they received the web link via SMS text messaging and email.

XSL•FO
RenderX

Figure 7. Registration and enrollment during Beehive piloting.



XSL•FO
RenderX

Table 3. Demographics of Beehive pilot participants[a].

	Service users (n=48)	Support persons (n=42)	EPI[b] program providers (n=72)
Age (y), mean (range)	18.88 (12-31)	44.11 (31-61)[c]	33.89 (22-50)[d]
Sex at birth, n (%)			
Female	24 (50)	21 (70)	30 (42)
Male	24 (50)	6 (20)	5 (7)
Prefer not to say	__[e]	<5 (<12)	—
Missing	—	—	37 (51)
Gender, n (%)			
Female	18 (38)	21 (70)	30 (42)
Male	22 (46)	6 (20)	5 (7)
Nonbinary	<5 (<10)	—	—
Questioning or unsure of gender identity	<5 (<10)	—	—
Prefer not to say	<5 (<10)	<5 (<12)	37 (51)
Race, n (%)			
African or African American or Black	15 (31)	7 (17)	5 (7)
American Indian or Alaskan native	<5 (<10)	—	—
Asian	<5 (<10)	<5 (<12)	8 (11)
Hispanic or Latinx only	10 (21)	<5 (<12)	19 (26)
White or Caucasian	10 (21)	10 (24)	33 (46)
More than one race	8 (17)	<5 (<12)	5 (7)
Other	—	—	<5 (<7)
Prefer not to say	<5 (<10)	<5 (<12)	—
Unsure	—	<5 (<12)	—
Missing	—	12 (29)	<5 (<7)
Ethnicity, n (%)			
No—I do not identify as Hispanic or Latinx	27 (56)	17 (40)	41 (57)
Yes—I identify as Hispanic or Latinx	14 (29)	5 (12)	28 (39)
Unsure or do not know	5 (10)	5 (12)	—
Prefer not to say	<5 (<10)	<5 (<12)	<5 (<7)
Missing	—	12 (29)	<5 (<7)
Service user diagnosis, n (%)			
Clinical high risk	—	—	—
Attenuated psychosis symptoms	6 (13)	—	—
First episode psychosis	—	—	—
Substance induced psychotic disorder with onset during intoxication	<5 (<10)	—	—
Mood disorders with psychotic features	9 (19)	—	—
Schizoaffective disorder (bipolar or depressive type combined)	10 (21)	—	—
Schizophrenia	5 (10)	—	—
Other specified schizophrenia spectrum disorder	<5 (<10)	—	—
Unspecified psychosis	<5 (<10)	—	—
Other first episode psychosis	7 (15)	—	—

XSL•FO
RenderX

	Service users (n=48)	Support persons (n=42)	EPI[b] program providers (n=72)
Clinical high risk or first episode psychosis status not confirmed	—	—	—
Anxiety disorders	<5 (<10)	—	—
Number of support persons registered in Beehive, n (%)			
None	14 (29)	—	—
1	29 (60)	—	—
2	5 (10)	—	—
Relationship of support persons with service user, n (%)			
Parent (adoptive)	—	5 (12)	—
Parent (biological)	—	34 (81)	—
Stepparent	—	<5 (<12)	—
Spouse or partner	—	<5 (<12)	—
Sibling	—	<5 (<12)	—

[a]Cells with less than 5 individuals are masked to protect the identity of participants.

[b]EPI: early psychosis intervention.

[c]Data missing for 15 individuals.

[d]Data missing for 37 individuals.

[e]Not available.

XSL•FO
RenderX

Table 4. Professional background of early psychosis intervention program providers registered during Beehive pilot (N=72)[a].

Background	Values, n (%)
Education level	
HS[b] diploma or GED[c]	7 (10)
Associate's degree	<5 (<7)
BA[d] or BS[e]	16 (22)
MA[f] or MS[g]	18 (25)
MFT[h]	7 (10)
MSW[i]	<5 (<7)
PsyD[j]	5 (7)
PhD[k]	6 (8)
MD[l]	9 (13)
Professional role	
Administrative support staff	<5 (<7)
Case manager or recovery coach	<5 (<7)
Clinic coordinator	6 (8)
Clinical supervisor or team lead	<5 (<7)
Clinician or therapist	30 (42)
Family advocate	<5 (<7)
Peer support specialist	<5 (<7)
Prescriber or psychiatrist or Other medical personnel	9 (13)
Program director	<5 (<7)
Research staff	<5 (<7)
Supported education and employment specialist	<5 (<7)
Other	7 (10)
Licensure status	
Unlicensed	49 (68)
Licensed	23 (32)
Years licensed (n=23)	
≤1	8 (38)
1 to 6	7 (33)
≥7	6 (29)
Number of languages in which services are provided	
1	47 (65)
2	18 (25)
Missing	7 (10)
Languages for service provision[m]	
English	60 (92)
Spanish	18 (25)
Arabic	<5 (<7)
Hmong	<5 (<7)
Tagalog	<5 (<7)

XSL•FO
RenderX

Background	Values, n (%)
Other	<5 (<7)

[a]Cells with less than 5 individuals are masked to protect identity of participants.
[b]HS: high school.
[c]GED: general educational development.
[d]BA: Bachelor of Arts.
[e]BS: Bachelor of Science.
[f]MA: Master of Arts.
[g]MS: Master of Science.
[h]MFT: Master of Marriage and Family Therapy.
[i]MSW: Master of Social Work.
[j]PsyD: Doctor of Psychology.
[k]PhD: Doctor of Philosophy.
[l]MD: Doctor of Medicine.
[m]Respondents could select more than one response, so percentages will be greater than 100%.

Table 5. Survey completion by respondent type[a].

	Service users, n (%)	Support persons, n (%)	EPI[b] program providers, n (%)
Survey completion at enrollment[c]			
All	30 (63)	17 (50)	10 (21)
Partial	5 (10)	__[d]	—
None	13 (27)	17 (50)	38 (79)
Survey completion at first time point[e]			
All	18 (40)	17 (55)	—
Partial	8 (18)	3 (10)	4 (9)
None	19 (42)	13 (42)	41 (91)
Survey completion at second time point[f]			
All	1 (33)	—	—
Partial	—	—	—
None	2 (67)	2 (100)	3 (100)

[a]Partial survey completion indicates that respondents completed at least one survey, but did not complete all assigned surveys.
[b]EPI: early psychosis intervention.
[c]Total respondents for service users, n=48; support persons, n=34; and service providers, n=48.
[d]Not available.
[e]Total respondents for service users, n=45; support persons, n=31; and service providers, n=45.
[f]Total respondents for service users, n=3; support persons, n=2; and service providers, n=3.

Discussion

Principal Findings

This study describes the EPI-CAL program's design and acceptability testing approach for a custom web-based and tablet app, Beehive, to support systematic data collection, care delivery, program evaluation, and research across a statewide network of EPI programs. Our goal was to develop an app that was clinically useful for, usable by, and acceptable to diverse EPI programs across the state of California.

To ensure the app best matched the needs of the EPI participants, we adopted a UCD approach to develop Beehive. Previous research in the mental health digital space supports that active involvement from the app's intended users during the development phase can improve the appropriateness of the end product for the users of interest [38]. Initial feedback across the 3 development phases was primarily collected in workshops (storyboard and alpha version) and during pilot implementation (beta version). In storyboard and alpha workshops, we presented prototypes to demonstrate major features of the app and asked for feedback on the app's "look and feel," compatibility with existing clinical workflow, and ease of use and acceptability for service users, their support persons, and EPI program providers. Consistent with other studies who have included end users during the design phase of their eHealth apps [38-41],

XSL•FO
RenderX

feedback from these workshops resulted in immediate changes to the alpha and beta apps that would not have otherwise been made.

During piloting, we continued to collect user feedback around Beehive features, as well as assess acceptability of the app by examining preliminary enrollment and survey completion. Our enrollment and survey completion rates are consistent with the acceptability of other mental health apps developed using a UCD approach [42,43], although there is wide variability depending on the implementation approach.

During the design and testing phase, we observed that different types of community partners expressed different, and at times conflicting, needs. For example, we asked participants about their preferences for seeing score thresholds or comparative data as part of the visualization for their clinical measures. Some service users said that, in times of relapse or increasing intensity of symptoms, additional information on the visualization would be demoralizing. In contrast, many participants could imagine scenarios where that information would be useful as a form of psychoeducation to normalize service-user experiences or understand the relative severity of symptoms. To address these diverse needs and promote engagement with Beehive we added a toggle to individual-level visualizations so that users can turn the threshold information or comparative data on or off. A flexible design approach that is tailored to an individual's needs has been shown to be more efficacious in a mobile health app setting [43,44]. Therefore, design changes incorporated flexibility where possible to enable our team to meet the various needs of individuals while maintaining consistent implementation to meet evaluation and research goals.

Similarly, user feedback informed our training approach. For example, some EPI program provider users expressed that they would use the graphs in the app in clinical care with service users to highlight strengths and progress. In contrast, another EPI program provider participant cautioned that some service users may not want to look at data visualizations and that this should be an optional part of their care. Thus, our training approach highlights how visualizations may be used in direct care but are not prescriptive. Feedback from workshop participants also highlighted the importance of shifting clinic culture to view data collection as a key part of care provision. Our team considers EPI program providers to be integral in promoting engagement for service users and their support persons as it is the EPI program providers who communicate why Beehive is being used in care. To begin addressing potential barriers of buy-in and engagement for all users, we designed our trainings to include the context of why Beehive and MBC were being implemented in their programs, including a presentation on the potential value of Beehive (designed and delivered by author LS) [45].

During piloting, we observed barriers to integration. For example, despite our designing the first training such that the programs could start registering service users immediately afterward, the programs failed to do so. When asked, programs informed us that they were nervous to receive questions about Beehive that they did not know how to answer. This resulted in our team creating materials to provide more structure for programs as they introduced Beehive to service users, such as an introduction script, Beehive infographics, and other handouts. Once programs started enrolling existing service users, many found it difficult to transition enrolling new service users, given that they already had numerous documentation requirements during their clinical intake process. In response, we added a "workflow meeting" to our training series where we asked program leadership and key staff involved in intakes to walk us through their existing procedures so that we could help brainstorm where the required Beehive workflow steps could be implemented and who from their program would be responsible for each step. We additionally observed that clinician-entered data were hard for sites to prioritize. For example, there was a lack of clarity within teams about who was responsible for entering these data and what training was required. Therefore, we chose to add another "workflow meeting" between key program staff and our team to help programs identify who was responsible for which surveys, who needed training, and how programs could monitor survey completion. We added these workflow meetings to our formal training protocol and made the support materials available to all sites who joined after the piloting phase.

Furthermore, our team worked with programs beyond the piloting phase to ensure that we continued to incorporate individual feedback and offer continuous support, which is key to successful adoption and can improve engagement [46]. After the piloting of the beta version of Beehive concluded, we continued to make development changes to meet users' needs, such as further design changes to the admin dashboard, widening survey completion windows, adjusting and eventually allowing customization of the frequency and timing of web link notifications, allowing the EULA to be completed before survey baseline date, simplifying registration fields, adding a survey status page, adding additional survey visualizations, adding a workflow for providers to enter data collected outside of Beehive, and prioritizing the order of additional languages in the app based on active need in the participating clinics. This iterative approach in response to user feedback is consistent with the development process of other eHealth apps [47].

This study highlights how critical it is for programs using a continuous improvement approach, such as UCD, to budget appropriately for ongoing development needs and staff time for ongoing support. As long as an app is in use and collecting data from real users, there should be a plan for ongoing project management and app development to address feedback from users, improve engagement and useability, and respond to changing needs. Implementing UCD from the outset allowed our team to be aware of and address user concerns before investing valuable time and resources in initial development and implementation. Focusing on workflow during the storyboard and alpha phases of app and continuing this collaborative relationship throughout the implementation phase resulted in an app that represents the interests and needs of users.

During the piloting, we observed that survey completion rates varied among different types of users. This variance may be partially explained by our training approach during piloting (see the Results section). As we continue to collect data after the piloting phase, we can evaluate if this trend continues beyond

XSL•FO
RenderX

the initial onboarding period throughout multiple years of data collection. These varied results may also reflect the challenges of implementing MBC, with or without an eHealth app, such as the training burden and limited time to conduct new duties associated with eHealth implementation [24-26,31]. When implementing outcomes data collection in these settings, it may be critical to gather only the minimum required data from EPI program providers (eg, diagnosis) and rely on service-user self-report measures whenever possible. Future analyses will examine the relationship between characteristics of EPI program providers (eg, degree and years licensed) and completion rates of clinician-entered data. Future work, including barriers and facilitators interviews with users after they gain more experience using Beehive, will be used to prioritize the needs and perspectives of our users in the ongoing development of Beehive and to better understand the reasons why users do or do not engage with the app [48].

Limitations

The COVID-19 pandemic introduced multiple challenges for our study, which may have reduced the breadth and diversity of participation in various phases of the project. We offered workshops over remote teleconferencing instead of in-person, which may have excluded individuals who are less comfortable with using technology. This may have disproportionately impacted on the recruitment of service users and support persons for participation in workgroups, as our participant numbers were lower than previous studies where we were conducting in-person research [31]. To reduce bias that may have resulted from this imbalance, we chose to prioritize the feedback of service users and support persons if there was conflicting feedback between these participants and EPI program provider participants. In addition, the beta and full versions of Beehive have been introduced to all service users in participating programs, regardless of their comfort with technology, and this has allowed us to incorporate informal feedback from these individuals as we have continued to make improvements to Beehive.

Much of our data in UCD workshops were gathered at the start of the COVID-19 pandemic, before anyone had experienced the long-term shift in daily practices brought on by the increased use of telehealth and remote working. We sought feedback on an app that was intended for in-person use and received feedback based on participants' experience of using in-person services. This highlights the importance of planning multiple opportunities for soliciting and incorporating feedback from sites so that apps can be responsive to changing environments.

Our workshops and piloting were only conducted in English. To serve the diverse population of California, Beehive needs to be both translated and adapted, a process known as localization [49], into at least 15 languages. Since Beehive's launch, we have localized into 7 additional languages. We continue to solicit feedback from users, including those whose primary language is not English, to inform the ongoing development of Beehive, and we will continue localize this app into additional languages.

While we used our prior knowledge of mental health development in the development of Beehive [11,12], we did not use a structured analysis approach for the feedback obtained during workshops due to time constraints imposed by project deliverables. To reduce the impact of subjective biases, the researchers who conducted each group debriefed afterward to review the notes, and recordings were referenced if notes were unclear or vague. In addition, all decisions about how to incorporate feedback from these notes into app development were made collaboratively by authors KEB, LMT, TAN, and VLT. Future work in this area will benefit from organized approaches to data collection and formal qualitative analysis [50,51].

Conclusions

Working with community partners to co-design an eHealth app for use in community EPI programs helped us to anticipate and resolve barriers earlier in the app development and implementation pipeline. On the basis of our observation and the data, there appeared to be high levels of engagement with Beehive. This resulted in feedback and continued design improvements which allowed our team to be better poised to launch Beehive across the EPI-CAL LHCN. Variance in survey completion rates among respondent types suggests that support persons and EPI program providers especially may need additional support.

Acknowledgments

The authors would like to thank all service users, support persons, and EPI program providers who participated in this study, Binda Mangat of Quorum Technologies, and its team of professionals for supporting the development of Beehive, and the Early Psychosis Intervention Network of California learning health care network.

This research made was supported by One Mind, the National Institute of Mental Health under award R01MH120555-01, and the following California counties: Los Angeles, Napa, Orange, San Diego, Solano, Stanislaus, Lake, Nevada, Mono, Colusa, and Sonoma. The content is solely the responsibility of the authors and does not necessarily represent the official views of any sponsors of the research.

Data Availability

The datasets generated and analyzed during this study are available from the corresponding author on reasonable request.

XSL•FO
RenderX

Authors' Contributions

KEB conceptualized and designed the study, created the focus group guides, conceptualized and designed storyboards, recruited participants, collected data, trained pilot sites, audited and cleaned data, analyzed data, drafted the manuscript, and reviewed and revised the manuscript for important intellectual content. VLT conceptualized and designed the study, created the focus group guides, conceptualized and designed storyboards, collected data, trained pilot sites, drafted the manuscript, and reviewed and revised the manuscript for important intellectual content. KMP audited and cleaned data, drafted the manuscript, and reviewed and revised the manuscript for important intellectual content. LMT conceptualized and designed the study, created the focus group guides, conceptualized and designed storyboards, collected data, trained pilot sites, and reviewed and revised the manuscript for important intellectual content. SE conceptualized and designed the study, created the focus group guides, collected data, trained pilot sites, and reviewed and revised the manuscript for important intellectual content. MS, ABW, and AJP conceptualized and designed the study and reviewed the manuscript for important intellectual content. LS collected data, trained pilot sites, and reviewed and revised the manuscript for important intellectual content. CKH trained pilot sites, audited and cleaned data, and reviewed and revised the manuscript for important intellectual content. VEP, APM, and MK-W trained pilot sites and reviewed and revised the manuscript for important intellectual content. CM, MJM, NS, KLHN, and YZ cleaned data and reviewed and revised the manuscript for important intellectual content. TAN obtained funding, conceptualized and designed the study, created the focus group guides, conceptualized and designed storyboards, collected data, reviewed and revised the manuscript for important intellectual content.

Conflicts of Interest

KEB consulted with ChatOwl Inc after data collection and before submission. LMT was employed by ChatOwl Inc, a digital mental health company after data collection and before submission. LMT was a founder and owner of shares in Safari Health, Inc. during project implementation, data collection, and write up (no longer true at time of writing) and is an employee and shareholder at Kooth LLC at time of publication. TAN is a cofounder and shareholder in Safari Health, Inc.

Multimedia Appendix 1
The Early Psychosis Intervention Network of California outcomes collected in Beehive.
[DOCX File , 46 KB - humanfactors_v12i1e65889_app1.docx]

References

1. van Os J, Hanssen M, Bijl RV, Vollebergh W. Prevalence of psychotic disorder and community level of psychotic symptoms: an urban-rural comparison. Arch Gen Psychiatry 2001 Jul;58(7):663-668. [doi: 10.1001/archpsyc.58.7.663] [Medline: 11448373]
2. Menchaca A, Pratt B, Jensen E, Jones N. Examining the racial and ethnic diversity of adults and children. United States Census Bureau. 2023 May 22. URL: https://tinyurl.com/3wbpwc8t [accessed 2024-08-16]
3. Niendam TA, Sardo A, Savill M, Patel P, Xing G, Loewy RL, et al. The rise of early psychosis care in California: an overview of community and university-based services. Psychiatr Serv 2019 Jun 01;70(6):480-487. [doi: 10.1176/appi.ps.201800394] [Medline: 30890048]
4. Tryon VL, Nye KE, Savill M, Loewy R, Miles MJ, Tully LM, et al. The California collaborative network to promote data driven care and improve outcomes in early psychosis (EPI-CAL) project: rationale, background, design and methodology. BMC Psychiatry 2024 Nov 14;24(1):800 [FREE Full text] [doi: 10.1186/s12888-024-06245-6] [Medline: 39543502]
5. Heinssen RK, Azrin ST. A national learning health experiment in early psychosis research and care. Psychiatr Serv 2022 Sep 01;73(9):962-964 [FREE Full text] [doi: 10.1176/appi.ps.20220153] [Medline: 35895842]
6. Lewis CC, Boyd M, Puspitasari A, Navarro E, Howard J, Kassab H, et al. Implementing measurement-based care in behavioral health: a review. JAMA Psychiatry 2019 Mar 01;76(3):324-335 [FREE Full text] [doi: 10.1001/jamapsychiatry.2018.3329] [Medline: 30566197]
7. Firth J, Cotter J, Torous J, Bucci S, Firth JA, Yung AR. Mobile phone ownership and endorsement of "mHealth" among people with psychosis: a meta-analysis of cross-sectional studies. Schizophr Bull 2016 Mar;42(2):448-455 [FREE Full text] [doi: 10.1093/schbul/sbv132] [Medline: 26400871]
8. Naslund JA, Marsch LA, McHugo GJ, Bartels SJ. Emerging mHealth and eHealth interventions for serious mental illness: a review of the literature. J Ment Health 2015;24(5):321-332 [FREE Full text] [doi: 10.3109/09638237.2015.1019054] [Medline: 26017625]
9. Meyer N, Kerz M, Folarin A, Joyce DW, Jackson R, Karr C, et al. Capturing rest-activity profiles in schizophrenia using wearable and mobile technologies: development, implementation, feasibility, and acceptability of a remote monitoring platform. JMIR Mhealth Uhealth 2018 Oct 30;6(10):e188 [FREE Full text] [doi: 10.2196/mhealth.8292] [Medline: 30377146]
10. Torous J, Friedman R, Keshavan M. Smartphone ownership and interest in mobile applications to monitor symptoms of mental health conditions. JMIR Mhealth Uhealth 2014 Jan 21;2(1):e2 [FREE Full text] [doi: 10.2196/mhealth.2994] [Medline: 25098314]

XSL•FO
RenderX

11. Kumar D, Tully LM, Iosif AM, Zakskorn LN, Nye KE, Zia A, et al. A mobile health platform for clinical monitoring in early psychosis: implementation in community-based outpatient early psychosis care. JMIR Ment Health 2018 Feb 27;5(1):e15 [FREE Full text] [doi: 10.2196/mental.8551] [Medline: 29487044]
12. Niendam TA, Tully LM, Iosif AM, Kumar D, Nye KE, Denton JC, et al. Enhancing early psychosis treatment using smartphone technology: a longitudinal feasibility and validity study. J Psychiatr Res 2018 Jan;96:239-246. [doi: 10.1016/j.jpsychires.2017.10.017] [Medline: 29126059]
13. Institute of Medicine, Committee on the Learning Health Care System in America, McGinnis JM, Stuckhardt L, Smith M, Saunders R. Best Care at Lower Cost: The Path to Continuously Learning Health Care in America. Washington, DC: National Academies Press; 2013.
14. Unützer J, Katon W, Callahan CM, Williams JWJ, Hunkeler E, Harpole L, et al. Collaborative care management of late-life depression in the primary care setting: a randomized controlled trial. JAMA 2002 Dec 11;288(22):2836-2845. [doi: 10.1001/jama.288.22.2836] [Medline: 12472325]
15. Katon WJ, Lin EH, Von Korff M, Ciechanowski P, Ludman EJ, Young B, et al. Collaborative care for patients with depression and chronic illnesses. N Engl J Med 2010 Dec 30;363(27):2611-2620 [FREE Full text] [doi: 10.1056/NEJMoa1003955] [Medline: 21190455]
16. Scott K, Lewis CC. Using measurement-based care to enhance any treatment. Cogn Behav Pract 2015 Feb;22(1):49-59 [FREE Full text] [doi: 10.1016/j.cbpra.2014.01.010] [Medline: 27330267]
17. Heinssen RK, Goldstein AB, Azrin ST. Evidence-based treatments for first episode psychosis: components of coordinated specialty care. National Institutes Mental Health. 2014 Apr 14. URL: https://www.nimh.nih.gov/sites/default/files/documents/health/topics/schizophrenia/raise/evidence-based-treatments-for-first-episode-psychosis.pdf [accessed 2025-03-31]
18. Garland AF, Kruse M, Aarons GA. Clinicians and outcome measurement: what's the use? J Behav Health Serv Res 2003;30(4):393-405. [doi: 10.1007/BF02287427] [Medline: 14593663]
19. Bickman L. A measurement feedback system (MFS) is necessary to improve mental health outcomes. J Am Acad Child Adolesc Psychiatry 2008 Oct;47(10):1114-1119 [FREE Full text] [doi: 10.1097/CHI.0b013e3181825af8] [Medline: 20566188]
20. Trivedi MH, Daly EJ. Measurement-based care for refractory depression: a clinical decision support model for clinical research and practice. Drug Alcohol Depend 2007 May;88 Suppl 2(Suppl 2):S61-S71 [FREE Full text] [doi: 10.1016/j.drugalcdep.2007.01.007] [Medline: 17320312]
21. Schreiweis B, Pobiruchin M, Strotbaum V, Suleder J, Wiesner M, Bergh B. Barriers and facilitators to the implementation of eHealth services: systematic literature analysis. J Med Internet Res 2019 Nov 22;21(11):e14197 [FREE Full text] [doi: 10.2196/14197] [Medline: 31755869]
22. Treisman GJ, Jayaram G, Margolis RL, Pearlson GD, Schmidt CW, Mihelish GL, et al. Perspectives on the use of eHealth in the management of patients with schizophrenia. J Nerv Ment Dis 2016 Aug;204(8):620-629 [FREE Full text] [doi: 10.1097/NMD.0000000000000471] [Medline: 26828911]
23. Ossebaard HC, van Gemert-Pijnen L. eHealth and quality in health care: implementation time. Int J Qual Health Care 2016 Jun;28(3):415-419. [doi: 10.1093/intqhc/mzw032] [Medline: 27029590]
24. Hatfield DR, Ogles BM. Why some clinicians use outcome measures and others do not. Adm Policy Ment Health 2007 May;34(3):283-291. [doi: 10.1007/s10488-006-0110-y] [Medline: 17211715]
25. Zimmerman M, McGlinchey JB. Why don't psychiatrists use scales to measure outcome when treating depressed patients? J Clin Psychiatry 2008 Dec;69(12):1916-1919. [doi: 10.4088/jcp.v69n1209] [Medline: 19192467]
26. Jensen-Doss A, Haimes EM, Smith AM, Lyon AR, Lewis CC, Stanick CF, et al. Monitoring treatment progress and providing feedback is viewed favorably but rarely used in practice. Adm Policy Ment Health 2018 Jan;45(1):48-61 [FREE Full text] [doi: 10.1007/s10488-016-0763-0] [Medline: 27631610]
27. Gould JD, Lewis C. Designing for usability: key principles and what designers think. Commun ACM 1985;28(3):300-311. [doi: 10.1145/3166.3170]
28. Norman DA, Draper SW. User Centered System Design: New Perspectives on Human-computer Interaction. Mahwah, NJ: Lawrence Erlbaum Associates; 1986.
29. Gulliksen J, Göransson B, Boivie I, Blomkvist S, Persson J, Cajander Å. Key principles for user-centred systems design. Behav Inform Technol 2003 Nov;22(6):397-409. [doi: 10.1080/01449290310001624329]
30. van Gemert-Pijnen JE, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011 Dec 05;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
31. Savill M, Banks LM, Tryon VL, Ereshefsky S, Nye KE, Botello RM, et al. Exploring data collection priorities of community partners in early psychosis care. Psychiatr Serv 2024 Sep 01;75(9):854-862. [doi: 10.1176/appi.ps.20230455] [Medline: 38595117]
32. Tully LM, Nye KE, Ereshefsky S, Tryon VL, Hakusui CK, Savill M, et al. Incorporating community partner perspectives on eHealth technology data sharing practices for the California early psychosis intervention network: qualitative focus group study with a user-centered design approach. JMIR Hum Factors 2023 Nov 14;10:e44194 [FREE Full text] [doi: 10.2196/44194] [Medline: 37962921]

XSL•FO
RenderX

33. Endmann A, Keßner D. User journey mapping – a method in user experience design. i-com 2016;15(1):105-110. [doi: 10.1515/icom-2016-0010]
34. Newman MW, Landay JA. Sitemaps, storyboards, and specifications: a sketch of Web site design practice. In: Proceedings of the 3rd Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques. 2000 Presented at: DIS '00; August 17-19, 2000; New York, NY. [doi: 10.1145/347642.347758]
35. Ludden GD, van Rompay TJ, Kelders SM, van Gemert-Pijnen JE. How to increase reach and adherence of web-based interventions: a design research viewpoint. J Med Internet Res 2015 Jul 10;17(7):e172 [FREE Full text] [doi: 10.2196/jmir.4201] [Medline: 26163456]
36. Core Assessment Battery (CAB). Early Psychosis Intervention Network. URL: https://nationalepinet.org/core-assessment-battery-cab/ [accessed 2024-08-16]
37. Boothroyd RA, Chen HJ. The psychometric properties of the Colorado Symptom Index. Adm Policy Ment Health 2008 Sep;35(5):370-378. [doi: 10.1007/s10488-008-0179-6] [Medline: 18561020]
38. de Beurs D, van Bruinessen I, Noordman J, Friele R, van Dulmen S. Active involvement of end users when developing web-based mental health interventions. Front Psychiatry 2017;8:72 [FREE Full text] [doi: 10.3389/fpsyt.2017.00072] [Medline: 28515699]
39. van Bruinessen IR, van Weel-Baumgarten EM, Snippe HW, Gouw H, Zijlstra JM, van Dulmen S. Active patient participation in the development of an online intervention. JMIR Res Protoc 2014 Nov 06;3(4):e59 [FREE Full text] [doi: 10.2196/resprot.3695] [Medline: 25379679]
40. van Hierden Y, Dietrich T, Rundle-Thiele S. Designing an eHealth well-being program: a participatory design approach. Int J Environ Res Public Health 2021 Jul 06;18(14):7250 [FREE Full text] [doi: 10.3390/ijerph18147250] [Medline: 34299700]
41. Honary M, Fisher NR, McNaney R, Lobban F. A web-based intervention for relatives of people experiencing psychosis or bipolar disorder: design study using a user-centered approach. JMIR Ment Health 2018 Dec 07;5(4):e11473 [FREE Full text] [doi: 10.2196/11473] [Medline: 30530457]
42. Mak C, Whittingham K, Cunnington R, Boyd RN. Efficacy of mindfulness-based interventions for attention and executive function in children and adolescents—a systematic review. Mindfulness 2017 Jul 26;9(1):59-78. [doi: 10.1007/s12671-017-0770-6]
43. Kajitani K, Higashijima I, Kaneko K, Matsushita T, Fukumori H, Kim D. Short-term effect of a smartphone application on the mental health of university students: a pilot study using a user-centered design self-monitoring application for mental health. PLoS One 2020;15(9):e0239592 [FREE Full text] [doi: 10.1371/journal.pone.0239592] [Medline: 32976515]
44. Bakker D, Kazantzis N, Rickwood D, Rickard N. Mental health smartphone apps: review and evidence-based recommendations for future developments. JMIR Ment Health 2016 Mar 01;3(1):e7 [FREE Full text] [doi: 10.2196/mental.4984] [Medline: 26932350]
45. The value of beehive. YouTube. URL: https://www.youtube.com/watch?v=-7MDPRLU1BM&feature=youtu.be [accessed 2024-08-29]
46. Glomann L, Hager V, Lukas CA, Berking M. Patient-centered design of an e-mental health app. In: Joint Proceedings of the AHFE 2018 International Conference on Human Factors in Artificial Intelligence and Social Computing, Software and Systems Engineering, The Human Side of Service Engineering and Human Factors in Energy. 2018 Presented at: AHFE 2018; July 21-25, 2018; Orlando, FL. [doi: 10.1007/978-3-319-94229-2_25]
47. Kip H, Keizer J, da Silva MC, Beerlage-de Jong N, Köhle N, Kelders SM. Methods for human-centered eHealth development: narrative scoping review. J Med Internet Res 2022 Jan 27;24(1):e31858 [FREE Full text] [doi: 10.2196/31858] [Medline: 35084359]
48. Ereshefsky S, Gemignani R, Savill M, Sanford KC, Banks LM, Tryon VL, et al. A mixed-methods study exploring the benefits, drawbacks, and utilization of data in care: findings from the EPI-CAL early psychosis learning health care network. Schizophr Res 2025 Feb;276:157-166 [FREE Full text] [doi: 10.1016/j.schres.2025.01.018] [Medline: 39892249]
49. Schäler R. Localization and translation. In: Handbook of Translation Studies Online. Amsterdam, The Netherlands: John Benjamins Publishing Company; 2010:209-214.
50. Molina-Recio G, Molina-Luque R, Jiménez-García AM, Ventura-Puertos PE, Hernández-Reyes A, Romero-Saldaña M. Proposal for the user-centered design approach for health apps based on successful experiences: integrative review. JMIR Mhealth Uhealth 2020 Apr 22;8(4):e14376 [FREE Full text] [doi: 10.2196/14376] [Medline: 32319965]
51. Ozkaynak M, Sircar CM, Frye O, Valdez RS. A systematic review of design workshops for health information technologies. Informatics 2021 May 14;8(2):34. [doi: 10.3390/informatics8020034]

Abbreviations

EPI: early psychosis intervention
EPI-CAL: Early Psychosis Intervention Network of California
EULA: end user license agreement
HIPAA: Health Insurance Portability and Accountability Act

XSL•FO
RenderX

LHCN: learning health care network
MBC: measurement-based care
MCSI: Modified Colorado Symptom Index
UCD: user-centered design

Edited by C Jacob; submitted 04.11.24; peer-reviewed by J D'Arcey; comments to author 06.12.24; revised version received 19.12.24; accepted 12.03.25; published 09.04.25.

Please cite as:
Burch KE, Tryon VL, Pierce KM, Tully LM, Ereshefsky S, Savill M, Smith L, Wilcox AB, Hakusui CK, Padilla VE, McNamara AP, Kado-Walton M, Padovani AJ, Miller C, Miles MJ, Sharma N, Nguyen KLH, Zhang Y, Niendam TA
Co-Designing a Web-Based and Tablet App to Evaluate Clinical Outcomes of Early Psychosis Service Users in a Learning Health Care Network: User-Centered Design Workshop and Pilot Study
JMIR Hum Factors 2025;12:e65889
URL: https://humanfactors.jmir.org/2025/1/e65889
doi:10.2196/65889
PMID:

©Kathleen E Burch, Valerie L Tryon, Katherine M Pierce, Laura M Tully, Sabrina Ereshefsky, Mark Savill, Leigh Smith, Adam B Wilcox, Christopher Komei Hakusui, Viviana E Padilla, Amanda P McNamara, Merissa Kado-Walton, Andrew J Padovani, Chelyah Miller, Madison J Miles, Nitasha Sharma, Khanh Linh H Nguyen, Yi Zhang, Tara A Niendam. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 09.04.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Exploring Mental Health Content Moderation and Well-Being Tools on Social Media Platforms: Walkthrough Analysis

Zoë Haime[1,2], BSc, MSc, PGDip, PhD; Lucy Biddle[1,2,3], BA, MA, PhD

[1]Population Health Sciences, University of Bristol, Bristol, United Kingdom

[2]NIHR Bristol Biomedical Research Centre, Bristol, United Kingdom

[3]NIHR Applied Research Collaboration West, Bristol, United Kingdom

Corresponding Author:
Zoë Haime, BSc, MSc, PGDip, PhD
Population Health Sciences
University of Bristol
Bristol Medical School, University of Bristol
Canynge Hall 39 Whatley Road Bristol
Bristol, BS8 2PS
United Kingdom
Phone: 44 0117123456
Email: zoe.haime@bristol.ac.uk

Abstract

Background: Social networking site (SNS) users may experience mental health difficulties themselves or engage with mental health–related content on these platforms. While SNSs use moderation systems and user tools to limit harmful content availability, concerns persist regarding the implementation and effectiveness of these methods.

Objective: This study aimed to use an ethnographic walkthrough method to critically evaluate 4 SNSs—Instagram, TikTok, Tumblr, and Tellmi.

Methods: Walkthrough methods were used to identify and analyze mental health content moderation and safety and well-being resources of SNS platforms. We completed systematic checklists for each of the SNS platforms and then used thematic analysis to interpret the data.

Results: Findings highlighted both successes and challenges in balancing user safety and content moderation across platforms. While varied mental health resources were available on platforms, several issues emerged, including redundancy of information, broken links, and a lack of non–US-centric resources. In addition, despite the presence of several self-moderation tool options, there was insufficient evidence of user education and testing around these features, potentially limiting their effectiveness. Platforms also faced difficulties addressing harmful mental health content due to unclear language around what was allowed or disallowed. This was especially evident in the management of mental health–related terminology, where the emergence of "algospeak," where users adopt alternative codewords or phrases to avoid having content removed or banned by moderation systems, highlighted how users easily bypass platform censorship. Furthermore, platforms did not detail support for reporters or reportees of mental health–related content, leaving users susceptible.

Conclusions: Our study resulted in the production of preliminary recommendations for platforms regarding potential mental health content moderation and well-being procedures and tools. We also emphasized the need for more inclusive user-centered design, feedback, and research to improve SNS safety and moderation features.

(JMIR Hum Factors 2025;12:e69817) doi:10.2196/69817

KEYWORDS

online moderation; mental health; user experience; social networking; online well-being

XSL•FO
RenderX

Introduction

Background

Social networking sites (SNSs) can serve multiple purposes for individuals with mental health difficulties [1]. These include facilitating self-disclosure of mental health disorders, acting as discursive spaces, and offering exposure to both harmful and helpful mental health–related content [2,3]. However, several research studies have identified relationships between SNS use and development or exacerbation of mental health symptoms, including depression, anxiety, psychosis, eating disorder, self-harm, and suicide behaviors [4,5].

With approximately 1 in 8 people globally living with a mental health condition [6] and considering the multifaceted role SNSs can have in shaping mental health experiences, platforms have a significant responsibility to support and guide users in managing their well-being. SNSs currently use a variety of moderation and safety techniques to regulate their online communities, including blocking and removing content, banning users, outlining community guidelines, and providing mental health resources, through automatic and manual processes [7].

Despite this, concerns remain about the effectiveness of SNS safety practices [8]. For instance, users engaging with self-harm and suicide material online have reported actively avoiding online help during crises, including dismissing pop-ups and support links [9]. Furthermore, SNS users engaging with self-harm and suicide content online have criticized platform inconsistencies regarding content removal and the communication of decisions around this, reporting that these moderation actions can lead to feelings of stigmatization and harm [8,10]. In addition, studies have found that due to strict censorship of mental health–related terms, platform users may turn to "algospeak," where they adopt alternative codewords or phrases to avoid having content removed or banned by moderation systems [11], or to circumnavigate search filters [12,13]. Another study, which explored how people use platforms and technology to support their mental health [14] found that users wanted the provision of mental health support on SNSs to be more nuanced and trustworthy, with clear, evidence-based guidance. Concerns have also been raised more generally about users' ability to read and understand SNS policies and guidance. One study identified significant variation between community guidelines across platforms, highlighting how this can negatively impact user safety and trust [15]. Another study stated that users' contract requirements with SNSs whereby they state they have read and agreed to the terms of service (ToS) is "the biggest lie on the internet" [16].

These findings highlight possible inefficiencies in current moderation and safety mechanisms used by SNSs and significant potential risks, particularly for vulnerable users, including exposure to harmful content, inadequate crisis responses, and possible disengagement from regulated spaces in favor of less-moderated ones. The full extent of such moderation and safety challenges across SNSs remains unclear, highlighting the need for more in-depth explorations of user experiences. The introduction of the Online Safety Act (OSA) [17] in the United Kingdom, which aims to tighten safety regulations and standards for digital platforms, further emphasizes this need for comprehensive data on current SNS practices and their effectiveness, to fully assess and address issues.

The walkthrough method is a valuable tool for understanding user experiences on these online platforms [18,19]. It is a structured ethnographic approach examining application and platform features, settings, and interactions by "touring" the platform interactively. The method focuses on understanding how platform design and resources potentially encourage and impact user behavior. For instance, in their paper on a walkthrough of the dating platform Tinder, Gillett [20] identified several user safety issues, including the problematic placement and design of the reporting icon. The author proposed that these design choices could significantly influence user behavior, leading to potential unintended interactions with harmful content and adverse effects on user well-being. This study demonstrated how the walkthrough method can be used to reveal critical aspects of user experience that might otherwise be overlooked.

This Study

Using the walkthrough method, this study aimed to systematically identify how SNSs manage mental health content and user well-being through moderation, safety features, resources, and policies and to explore the potential impact of these platform attributes on users. By identifying platform successes and areas for improvement, the study aimed to contribute to the creation of safer and more supportive digital environments.

Methods

Research Design

This study used an ethnographic walkthrough method to systematically and critically assess the approaches of various SNSs in moderating mental health content and promoting user well-being. Researchers actively participated in the platforms as observers to evaluate their features, guidance, and resources.

Ethical Considerations

This study adhered to ethical guidelines for internet-mediated research [21]. Researchers ensured that their participation did not interfere with the functioning of the platforms, mislead other users, or go against platform regulations. Confidentiality was maintained by avoiding the recording of personally identifiable information, such as user-generated names or specific images. No interventions were made in response to harmful content observed, ensuring that the user experience and platform responses remained unaltered. No financial compensation was provided to researchers as they undertook the research as part of their routine professional role. There were well-being procedures in place for researchers, and they were able to withdraw from conducting the walkthrough checklists if they wished.

Ethics approval was received from the Faculty of Health Sciences Research Ethics Committee at the University of Bristol (approval number 17972).

XSL•FO
RenderX

Positionality Statement

As researchers in the field of mental health and digital technology, we acknowledge that our personal and professional experiences shape our perspectives and approach to this study. Researchers ZH and LB are both White women from the United Kingdom. ZH identifies as a digital native, reflecting a comfort with digital technology and early adoption of various online platforms throughout her life. In contrast, LB is a digital immigrant.

Data Collection

We selected 4 SNS platforms for evaluation. TikTok and Instagram were selected due to their status as two of the most widely used social media platforms in the United Kingdom, particularly among adolescents [22,23]. Tumblr was chosen for its established reputation as a significant SNS platform, offering a contrasting perspective to the newer platforms. In addition, although Tumblr has less representation in the current research literature and discussion around online safety [23], media reports indicate a recent shift in their user base toward the Gen Z population [24,25]. Tellmi was chosen as a smaller platform specializing in mental health support and discussion among young people, which offered an opportunity to evaluate a platform specifically designed with the mental health of users in mind, providing a useful comparison to SNSs that prioritize broader aspects of social interaction. For this study, Facebook was excluded from our SNS evaluation despite it being the most accessed platform in the United Kingdom. This was because of its lower popularity among younger UK users [23] and its likely similarities with Instagram, given their shared ownership by Meta.

To conduct our walkthroughs, we used 2 checklists (Multimedia Appendices 1 and 2) to systematically evaluate each platform's approach to moderating mental health content and supporting user well-being. Checklists were developed from previous research findings [8,26] and industry guidelines for suicide and self-harm content [27]. To conduct the walkthrough checklist, ZH used a new iOS device and LB used a web browser. Table 1 depicts the platform areas that the checklist focused on, and Table 2 shows the stages of use that were observed and documented through field notes and screenshots.

Table 1. Checklist focus areas.

Checklist focus areas	Description
Design, language, and features	Assessment of the design, language, and features of platform content related to mental health, safety, and well-being.
Policies and processes	Evaluation of the accessibility, availability, and relevance of information and policies concerning age limits, mental health rules, and safety guidelines.
Resources	Analysis of the accessibility, availability, and relevance of mental health resources, including those tailored for specific populations such as young people and parents or guardians.
Moderation processes	Examination of moderation processes, including self-moderation tools.

Table 2. Documented platform stages of use.

Stages of use	Details
Sign-up	For sign-up purposes, researchers entered a date of birth, making the account user <13 years, 15 years, and 19 years old, respectively. No account was created to comply with platform policies. New accounts were created by the researcher using their own date of birth for checklist completion. Tellmi also required users to enter a UK postcode. Researchers entered their own UK postcode. No other demographic data (eg, gender, ethnicity, and sexual orientation) was provided by researchers.
Standard platform use	Observations during typical use of the platform environment. Researchers only observed content and did not engage (eg, like, share, or comment) with content during this phase. However, in the case of Tumblr, some interactions (eg, following hashtags) were required during the sign-up process to proceed, as detailed in the Results section.
Searching for mental health terms	Searches were conducted for mental health–related terms, including algospeak terms related to eating disorders, self-harm, and suicide. Algospeak terms were informed by the Digital Dialogues Young Person's Group [28]. Algospeak terms are not listed to avoid unintentionally increasing their visibility. The full list is available to researchers on request to the authors.
Reporting mental health content	Reporting processes were observed, but no content or user reports were submitted to the platform.
Exploration of moderation, safety, and well-being information	Information provided by platforms about moderation, safety, and well-being was observed.
Exploration of self-moderation tools	Self-moderation tools available on platforms were observed and used.

Data Analysis

Data were analyzed using a thematic analysis approach [29,30]. Initially, ZH reviewed the field notes, observations, and screenshots, applying a deductive approach to generating codes based on the checklist format. These codes were then further refined through an iterative process. The codes were then organized into groups, based on similarities in the issues, ideas, or patterns they addressed. These grouped codes were then

XSL•FO
RenderX

defined as themes that represented meaningful constructs of the coded data. Data for each theme can be viewed in a comparative table, by platform (Multimedia Appendix 3).

Results

Overview

Data were analyzed to understand the current provisions by platforms regarding mental health–related content moderation and user well-being. The themes we defined are summarized in Table 3.

Table 3. Theme and theme descriptions.

Theme	Theme description
Observations of safety features	Includes how users are introduced to safety on the platform, particularly through explicit safety instructions during the user journey, and moderation interventions, such as during searches for mental health–related terms
Platform moderation and safety guidance and systems	Examines how platforms describe the mechanisms for content moderation, including automatic filters, content warnings, and moderation policies that impact what users see and how content is managed
User "self-moderation" tools and reporting	Explores functionality, accessibility, and availability of self-moderation tools for users
Platform support and resources	Explores the information and resources provided by platforms related to mental health and well-being
Platform interface and communication	Explores the style, format, and presentation of mental health–related information and tools to users via the platform

Observation of Safety Features

Sign-Up Processes

Across platforms, sign-up procedures generally lacked emphasis on key safety information. Instagram, TikTok, and Tumblr gave subtle displays of their ToS and privacy policies, without proactively engaging users with them. In addition, at sign-up, no platform provided direct links to their community guidelines (Moderation Guidelines on Tellmi), which detail important mental health content moderation information. Tellmi provided more safety-related prompts during the sign-up process, ensuring users agreed to data-sharing practices and were aware of the platform's limitations in handling mental health crises. In addition, they directed users to the ToS, making it clear that this was a key part of their agreement with the platform.

Age requirements varied, Instagram and TikTok set the minimum at 13 years, Tellmi set the minimum at 11 years, and Tumblr set the minimum at 16 years. During sign-up, TikTok and Instagram restricted users from changing their birthdate if they had originally attempted to sign up below the minimum age, while Tumblr and Tellmi allowed users to enter a different date and continue, exposing a lack of age verification procedures. Notably, during sign-up, TikTok offered additional videos to users aged 13 to 17 years on data protection and default settings, reflecting an attempt to mitigate risks for younger users.

As part of the sign-up process, Tumblr required users to follow other accounts before accessing content, by selecting from the platform's recommendation system for popular hashtags and accounts. We selected recommended hashtags that we believed could be problematic, which exposed us to potentially harmful disordered eating content early on. This highlighted insufficient moderation within Tumblr's algorithm-driven suggestions. In contrast, Instagram also directed users to engage with popular content, but all suggestions were verified, and we did not identify any potentially problematic accounts. Users also had the option to skip this step. TikTok and Tellmi did not prompt users to follow any specific content tags or accounts during sign-up.

Content Exposure

Initial platform use provided varying experiences in terms of content exposure. Tellmi, Instagram, and TikTok took deliberate steps to provide users with safe, moderated content. Tellmi presented users with a wide range of nonharmful posts on mental health–related topics, while Instagram and TikTok's algorithms effectively prioritized popular content such as humor, sports, and celebrities, avoiding explicitly harmful material. In contrast, due to Tumblr's requirements to follow accounts and the potentially problematic ones we chose, we were immediately exposed to harmful mental health–related content in our feed including images of graphic self-harm and severely underweight female-presenting figures, as well as further recommendations to follow other seemingly harmful accounts. In addition, other than during the sign-up of users aged 13 to 17 years on TikTok, where some self-moderation tools were mentioned in videos, none of the platforms offered explicit guidance on how to find or implement self-moderation or safety tools during initial use.

Mental Health–Related Content Searches

When searching for mental health–related terms, using default sensitive or mature content levels, the platforms differed in their handling of material. Tellmi allowed search results for discussions on topics such as suicide, self-harm, and eating disorders, but the content was moderated and nonharmful, focusing on recovery and support. In comparison, the same suicide and self-harm term searches returned no results on Instagram, TikTok, or Tumblr, although all 3 offered relevant support. Instagram provided details for multiple external support organizations, and their own general support resource on self-care tips. TikTok gave details for Samaritans, as well as their own support resource on self-harm and suicide, and Tumblr

XSL•FO
RenderX

provided details for mainly US-based organizations, as well as their internal "counseling and prevention" page. Neither Instagram nor TikTok is linked to their more developed internal self-harm and suicide help pages, available in their platform safety centers. For searches on the terms "eating disorders," "anorexia," and "bulimia," Tumblr showed a similar help message with details for several more US organizations and no results. TikTok showed no results but directed users to its internal "eating disorders" resource and contact details for Beat An Instagram search for "eating disorders" gave the same resource page as for "suicide" and "self-harm" searches on the platform. For the terms "anorexia" and "bulimia," Instagram gave these resources again but also provided the option for users to continue their search. Results included recovery and support posts, images of self-harm scars, content promoting disordered eating, related potentially harmful hashtags, access to user biographies that stated goal weights (below healthy adult weights), and the phrase "block don't report" (also found in Tumblr results). On Instagram, sensitive content warnings also appeared on some posts but were inconsistently applied.

In searches for algospeak, all platforms returned results for at least 1 algospeak term related to self-harm, suicide, or eating disorders. On Instagram, TikTok, and Tumblr, algospeak terms led to more graphic and harmful content, including identification of further algospeak hashtags. Tellmi again returned moderated content, meaning the discussions presented were not harmful.

Platform Moderation and Safety Guidance and Systems

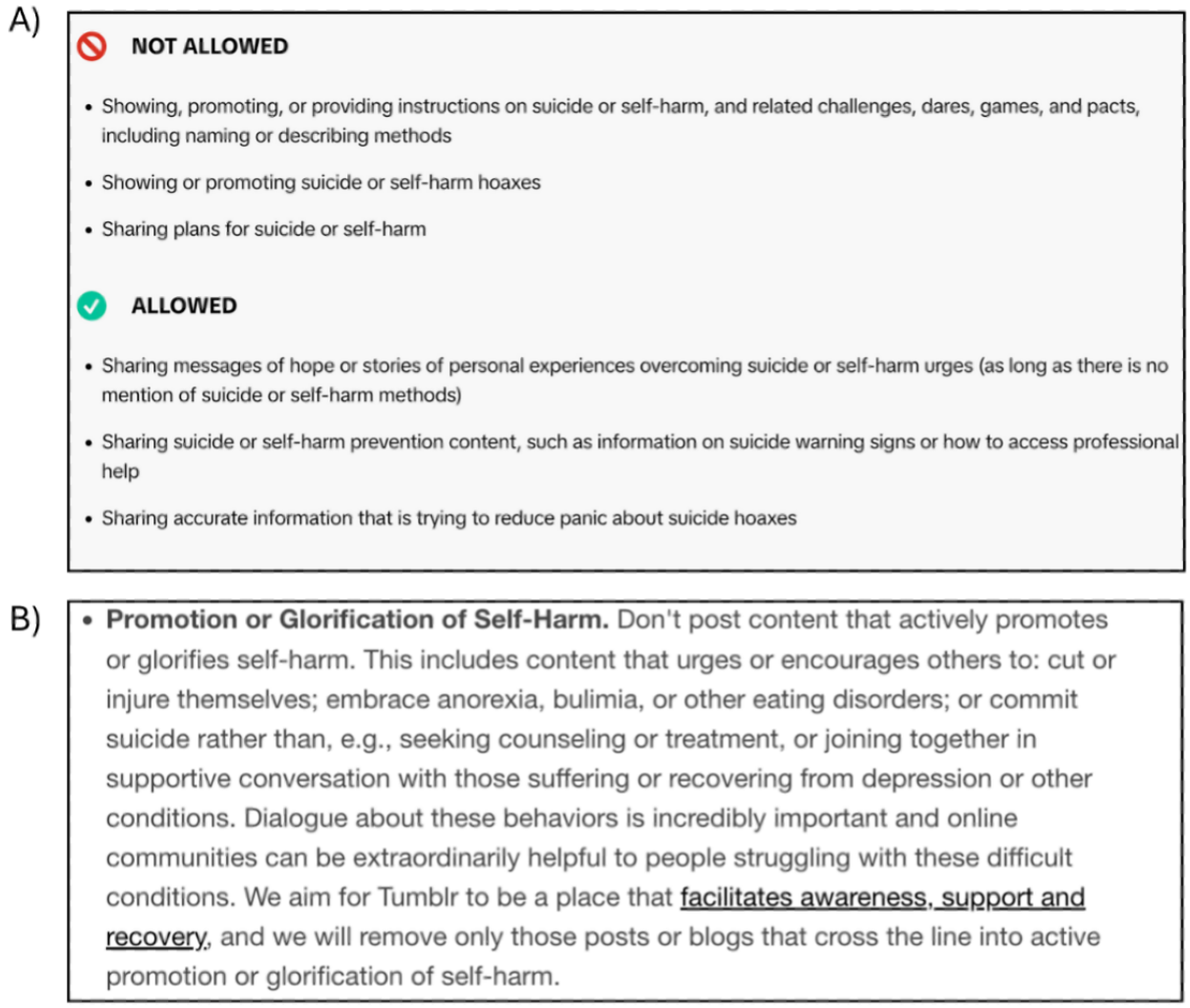
Moderation Transparency

All platforms provided information on the moderation of mental health content, primarily through their community guidelines or equivalent. The level of detail and transparency varied, with Instagram and TikTok offering extensive information on mental health content moderation and user well-being across various internally dedicated "centers," "guides," and "resources." However, this spread of information often led to repetition, required users to navigate multiple links, and sometimes failed to provide essential resources for help and support such as details of emergency or external support services. Tumblr and Tellmi provided limited internal spaces for users to access detailed information when encountering specific mental health–related content or seeking support for their own mental health needs. Users on Tellmi could access relevant areas of the directory, and on Tumblr, their "counselling and prevention resources" page. However, in both cases, resources were external to the platform.

TikTok and Instagram provided the most concise and specific information about what mental health–related content was and was not allowed on their platform, including clear examples. Instagram further outlined content they may label as sensitive, such as "older instances of self-harm such as healed cuts" or "content that depicts ribs, collar bones, thigh gaps, hips, concave stomach or protruding spine or scapula in a recovery context." In contrast, Tumblr and Tellmi guidance offered less clarity regarding disallowed content, often merging eating disorder, self-harm, or suicide content into a single category and using more abstract language, such as Tumblr prohibiting content that "embrace[s] anorexia" (Figure 1).

XSL•FO
RenderX

Figure 1. (A) Screenshot of TikTok community guidelines for allowed and disallowed suicide and self-harm content. (B) Screenshot of Tumblr user guidelines for promotion or glorification of self-harm.



Content Moderation Systems and Enforcement

TikTok, Tumblr, and Instagram described the use of automated moderation systems alongside human moderators, although roles were sometimes unclear. TikTok and Instagram further provided quarterly updates through their transparency centers, offering insights into the volume of mental health–related content removed both automatically and through user reporting. Tellmi detailed their unique approach of using entirely human moderation, emphasizing how this allowed for more nuanced decision-making and a personal approach to safety. However, we noted Tellmi moderators went offline from 11 PM to 8:30 AM (UK time), meaning no posting, and we also noted that during periods of heavy user engagement, posting would also be restricted to manage moderation load.

All platforms indicated that they could ban users or remove their content if it did not meet their required standards. However, the mechanisms and systems for enforcing these rules varied significantly. For instance, TikTok used a “strike” system to manage violations, but the details were somewhat vague, as the specifics of how strikes were issued and accumulated were not clearly communicated. Instagram stated that users “may” receive warnings before content was removed; however, the language used was noncommittal, leading to ambiguity about the enforcement process. TikTok and Instagram did allow users to track their account status, which indicated violations to the user and allowed for efficient use of the appeals process. Alternatively, Tumblr stated that they would inform users only after removing content or terminating accounts, except in the case of sexually explicit content, where users were given an opportunity to review, amend, or remove content. Tellmi’s approach was more explicit, stating that posts or replies that did not meet their moderation standards would be rejected and users would receive a brief explanation for the rejection to which they could directly respond to a human moderator.

Age-Based Protections

TikTok and, more recently, Instagram have implemented age-related defaults aimed at protecting young users, particularly those aged between 13 and 17 years. These focus on limiting potentially sensitive content exposure, managing screen time, and restricting interactions with unknown users. Both platforms enforce age-related settings with distinct approaches. On TikTok, younger teenagers (aged 13-15 years) face restrictions that cannot be altered. These include preventing others from downloading their videos, blocking all direct messaging, keeping

XSL•FO
RenderX

accounts private, and disabling features such as video stitching, duetting, and comments (stitching refers to combining an existing video with one you are creating; duetting refers to you posting your video side-by-side with an existing video on TikTok). Instagram similarly defaults those aged 13 to 15 years to private accounts and enforces restrictions including time limits, sleep mode, and disabling interaction features. Unlike TikTok, Instagram allows teenagers to request changes to these settings, but parental approval is required for modifications. Teenagers aged 16 to 17 years on both platforms have slightly more flexibility, such as enabling public accounts, but on Instagram, parental supervision is still necessary for approving changes. Alternatively, Tellmi adopted an age-banding approach to ensure users <20 years primarily interact with peers within 2 years of their age. Despite this, Tellmi informs users that responses to discussions may introduce wider age gaps, potentially affecting their intended interactions. In addition, Tumblr limits potentially mature content to audiences aged >18 years but does not implement any other age-based functions.

User "Self-Moderation" Tools and Reporting

User Control and Access to Tools

Self-moderation tools across platforms were often presented as options in user profile menus but with limited context regarding their potential effects. Even where their functionality was described in detail on platform help or safety pages, it was rare that well-being benefits were reflected to the user. Instagram and TikTok provided 21 and 23 self-moderation tools, respectively, offering time management, filtering, and privacy controls. Generally, these tools were in similar places on the platforms, making them intuitive to locate for those familiar with these types of online environments, but occasionally, they were in unexpected spaces or were less straightforward to implement. Some tools seemed to have overlapping functions, which may inhibit users' understanding of their functionality. Time management tools offered by Instagram and TikTok included prompts to take screen breaks, have quiet time, or go to sleep. These were easily accessible in time management controls of the respective platforms, where users could also view a visual tracker of their engagement. Both platforms offered preset time options for these tools, with TikTok's settings being slightly more customizable than Instagram's. However, the imposed limits on these presets suggested that neither platform was actively promoting extended or frequent disengagement. Self-moderation tools on the other platforms were much more limited. Tumblr offered 12 tools, primarily focused on privacy with some options for content filtering, and while Tellmi's role as a mental health support platform may make disengagement tools less intuitive, they provided 3 relevant self-moderation tools, with only 1 for hiding content related to certain mental health categories, limiting user control over their experiences.

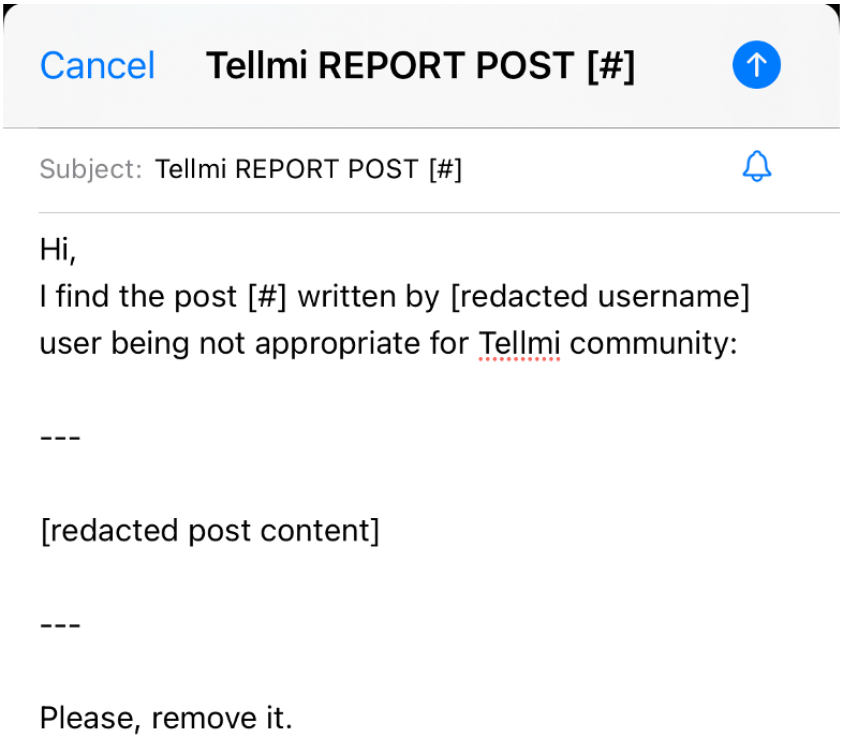
All platforms also implemented some form of sensitivity content rating, mostly to prevent access to what they termed mature content. TikTok, Tumblr, and Instagram automatically restricted such content for users aged <18 years, although definitions of mature or sensitive content for TikTok and Tumblr were vague. Tellmi offered greater transparency by clearly indicating that higher sensitivity would limit access to potentially triggering content. However, the specific content covered by each sensitivity level remained unclear. On Instagram and Tellmi, where sensitivity ratings could be adjusted by users, more restrictive settings limited search results for mental health and algospeak terms, suggesting that mental health content was included in these sensitivity categories for both platforms.

Reporting Systems

All platforms also shared similarities in encouraging users to report harmful content throughout their health, safety, and policy pages, often referring to user responsibility to report others' content that violates platform terms. Instagram and TikTok allowed users to report specific categories of mental health content such as suicide, self-harm, and disordered eating. Tumblr only provided a mental health content option for reporting suicide and self-harm, potentially hindering users aiming to report disordered eating content. However, Tumblr's reporting process was more detailed, requiring users to indicate if the person posting content was actively threatening suicide and to explain their concerns. In contrast, TikTok included reminders of platform rules regarding mental health–related content before users submitted their report, reinforcing guidelines and giving users a clear understanding of the platform's expectations. Tellmi offered a more manual process, where users reported any content by sending an email to the platform's support team via an automated template that could be personalized before submission (Figure 2). This method contrasted with the more automated systems of the other platforms.

XSL•FO
RenderX

Figure 2. Screenshot of an automated Tellmi template report email.



The platforms also differed in how they handled report follow-ups. Instagram and TikTok allowed users to check the status of their reports and appeal any decisions made, creating a straightforward process for feedback. Notably, Instagram also indicated that user reports may influence the content algorithm for their feed:

> *We may use your report to make content similar to what you reported appear lower in your Feed or in the recommendations we make for you, such as Suggested posts and Explore.* [Instagram's "Staying Safe" section in the Help Centre]

The other platforms lacked tracking features for their reporting systems, although Tellmi did offer email correspondence regarding decisions made on reports, allowing users to engage directly with a human moderator. None of the platforms provided information about whether they offered follow-up with support or resources for users who reported mental health–related content or for those whose content was reported, and as we did not report mental health content within this study, we cannot confirm their processes.

Platform Support and Resources

Platform resources for mental health support varied significantly. Instagram and TikTok provided a range of internal resources on self-harm, suicide, and disordered eating, including pages with practical advice, educational content, and emergency procedures that were written in consultation with credible external organizations. Uniquely, TikTok also published a page focused on creator burnout, and Tellmi allowed users to complete and track scores in mental health quizzes. In addition, Instagram hosted a variety of resources around mental health that had been developed and published by external services and charities, which users were able to download. Similarly, Tellmi directed users to a wealth of resources hosted externally. Both Instagram and TikTok provided a somewhat overwhelmingly wide range of guidance and policy pages directed at young people and their parents or guardians, again providing information on well-being support and practical online advice. Tumblr only provided 1 brief page for users that aimed to help with mental health needs, giving specific advice on what to do in response to immediate harm. On this page, Tumblr also gave contact details for a couple of UK-based mental health organizations. Tellmi and Instagram offered a broader range of relevant organizations, including those focused on mental health, online harms, and intersectional factors such as sexuality and religion. Notably, TikTok, Instagram, and Tumblr often featured US-centric resources across help and safety pages, which sometimes excluded other countries.

Platform Interface and Communication

Platforms generally communicated with their users using clear and professional language, suitable for their audiences. However, Tumblr's casual style veered into potentially patronizing territory:

> *You have to be the Minimum Age to use Tumblr. We're serious: it's a hard rule. "But I'm, like, almost old enough!" you plead. Nope, sorry. If you're not old enough, don't use Tumblr. Ask your parents for a PlayStation 4, or try books.* [Tumblr's ToS]

All platforms occasionally relied on nondefinite language regarding processes, which could potentially confuse users and create opportunities for inconsistent practices. When conveying information, Instagram and TikTok were reasonably text heavy, although attempts were made to use more mixed media in youth-centered areas of the platform. Tellmi used images and short descriptions, likely in an attempt to keep younger users

XSL•FO
RenderX

engaged. Tumblr was more effective at using a mix of media—such as illustrations, GIFs, and videos—throughout their information and guidance pages aimed at all audiences.

Tellmi organized its resources and help pages in a single directory, and Tumblr, across a few pages, resulting in a concise and user-friendly experience. In contrast, Instagram and TikTok disseminated information about mental health policies, guidance, and resources across various centers, with Instagram often redirecting users to additional Meta sites as well. Although this approach offered multiple valuable resources and ways to discover them, it complicated user navigation and increased information burden and repetition.

Within the app interface, Tumblr, TikTok, and Instagram provided a trigram "hamburger"  icon for access to most tools, pages, and policies, whereas Tellmi did not use familiar icons along a navigation bar. On the browser, access to all relevant documentation was available in the footer of the platform home page. In addition, several platforms functioned differently depending on device, occasionally restricting direct access to self-moderation options on browsers. Finally, Instagram's use of hidden links—accessible only through specific paths—and TikTok and Instagram's occasional reliance on blog posts for self-moderation tool descriptions limited accessibility for users.

Discussion

Principal Findings

This study extended our understanding of how a selection of SNSs manage user safety and well-being. Using walkthrough analysis, we observed both accomplishments by platforms and challenges they face in balancing content management, moderation, and user empowerment, leading to the development of preliminary recommendations for the industry (Textbox 1). Findings raised some critical concerns regarding the effectiveness of existing practices and the need for further research.

XSL•FO
RenderX

Textbox 1. Preliminary recommendations for the industry.

User onboarding and education

- Use the first-time new user's engagement with the platform to guide them to self-moderation tools and inform them of default settings (including those related to age restrictions) and moderation procedures.
- During sign-up, consider introducing prompts to encourage users to read the terms of service, privacy policies, and community guidelines. Alternatively, give users access to brief explanations of relevant key information from policies.
- Consider giving users the option to not follow accounts during their initial interactions on the platform.
- Consider implementing transition periods when users move into the 16- to 17-year and then >18-year age bands on age-restricted platforms. During this period, provide tailored education information on self-moderation tools and the benefits of using them.

Reporting content or users

- Offer clear guidance on reporting any potentially harmful mental health–related content, ensuring the process is simple and user-friendly, with any information about outreach and resources sent to reported users. Consider adding a system for users to track reports.
- Offer immediate follow-up resources to users who report mental health–related content.
- Consider a warning system for users whose content or accounts are reported for harmful mental health–related content, giving them the opportunity to review and amend their post. When content is removed or accounts are banned, ensure transparency around decisions and consider sending mental health support resources.
- Consider allowing users access to information about the effectiveness of platform moderation systems through reports on content removal and user bans for posting mental health–related content.
- Consider shifting away from language that expects users to report content and instead highlight the platform's commitment to proactive moderation and safety measures.

Support and resources

- Provide evidence-based internal support resources around mental health so users can access information in a familiar and trusted space. Consolidate and streamline these resources to avoid overwhelming vulnerable users.
- Ensure that resources and support are regionally and culturally relevant across platform pages. Platforms should strive to include organizations that address intersectional factors, promoting inclusivity and better serving diverse user needs.
- Consider specifying the evidence base for default safety mechanisms (including age restrictions).

Self-moderation tools

- Ensure users have access to a variety of self-moderation tools, offering them flexibility in managing and controlling their platform use. Clearly explain the purpose and functionality of each tool, ensuring they are distinct, practical, and easy to use.
- Ensure that users have access to clear explanations about the benefits of self-moderation tools for their well-being.
- Consider adding clear step-by-step written instructions and visual aids (eg, GIFs, screenshots, and short videos) when instructing users on the implementation of self-moderation and well-being tools.
- Consider the introduction of more default user options, such as automatically clearing search history or readjusting content level settings after a session to reduce the burden on users, giving them opt-out options.

Platform algorithms and content allowance

- Be explicit about what mental health–related content is allowed and restricted, using examples and unambiguous language.
- Where platforms use recommendation systems to influence content users engage with, prioritize giving users the ability to "refresh" their platform algorithm.
- Strive to enhance the robustness of automatic moderation systems, aiming for improved detection of potentially harmful mental health content (including severe, graphic, competitive, and instructive content). Recognize the complexity of this task and consider an approach that integrates human oversight.
- Consider alternatives to censoring common mental health terms, such as increasing the number of steps needed for users to access content, using overlays, and "pushing" recovery or positive mental health related content via platform algorithms.
- Consider implementing rate limits on accounts for repeated searches of potentially harmful content or algospeak terms. This could involve introducing novel help messaging or alerts after engagement attempts, which change depending on the frequency or severity of search terms.

Early Intervention

Previous studies have shown that application onboarding periods are critical in various contexts, including assisting new users' engagement with features [31,32], and helping patients with mental health disorders manage expectations and build confidence and proficiency with technology [33,34]. However, findings from this study showed that platforms were limited in

XSL•FO
RenderX

their education of users during initial engagements, particularly lacking clear direction to "community guidelines" and "self-moderation" tools. Such a passive approach may contribute to users feeling less informed about platform features and potentially susceptible [35].

Despite this, there was some evidence of platforms using nudges—subtle prompts designed to positively influence user behavior [36]—during this period. This included TikTok's data protection and safety videos for younger teenagers, Instagram and Tumblr's prompts for users to follow accounts, personalizing recommendations, and Tellmi's information regarding its limitations in terms of crisis support. These nudges served as proactive ways of engaging users with platform mechanisms, leveraging early interactions to educate or apply settings that could benefit users in the future [36], though there were some limitations, TikTok's instructional videos could be overwhelming, especially given that online users perceive learning about privacy settings as burdensome [37]. In addition, targeting younger teenagers excluded other users who may need the guidance. Furthermore, TikTok and Instagram allowed users to skip their processes, which, although encouraging autonomy, could result in users being less informed [36]. Tumblr did not offer the same skipping option, but findings indicated that the recommendation system offered options to engage with potentially harmful content, exposing broader moderation issues [38]. This suggests that while platforms have some capacity to engage users through nudges, they may be missing a valuable opportunity for early intervention around safety features and content moderation for their respective user bases. In addition, recommendation system inadequacies need to be considered significant in the safety of new or susceptible users.

Navigating the Platform

Instagram and TikTok offered comprehensive internal mental health–related support, including resources on suicide, self-harm, and disordered eating, as well as areas for younger or parent and guardian audiences. Integrating evidence-based resources internally and matching platform branding may be an effective approach for capturing users' attention through familiarity and perceived reliability [39,40]. However, the volume of mental health–related resources often led to repetition, resulting in the redundancy of information and potentially user fatigue [41]. Given evidence that individuals with severe mental illness struggle to understand complex content or navigation, platforms should aim to avoid overwhelming users with excessive psychoeducation, safety, and well-being content [42,43]. Furthermore, the presence of broken links and spelling mistakes in some of these sections emphasized poor maintenance, highlighting potential concerns around the platforms' commitment to providing high-quality, up-to-date resources [44].

Instagram also gave access to several mental health–related resources produced by other organizations, and Tellmi provided similar links to external content in various media forms. This approach allowed platforms to offer a wider range of resources, including those addressing intersectional factors in mental health, without needing to produce the content themselves, thereby improving efficiency. However, directing users away from internal platform resources may have drawbacks, including the need for monitoring external updates and possible inconsistencies in resource branding that could confuse users [44].

Following searches for harmful content, all platforms provided contact information for mental health organizations and emergency services. However, these details were often absent in other mental health–related platform spaces, missing opportunities for preventive intervention [45]. In addition, except for Tellmi, which is for young people in the United Kingdom, platforms overrepresented US resources. This reflects a broader issue of US-centric content bias on global platforms that may marginalize non–US users by limiting accessibility to relevant information [46].

Design decisions on platforms, particularly those impacting usability and accessibility to features, were also considered. Tumblr, TikTok, and Instagram used the "hamburger" icon for access to settings, policies, and resources, although some features were directly accessible during platform use. While the hamburger icon is widely recognized by experienced users, it may appear hidden or unclear to those without prior online knowledge [47]. Tellmi used a ring-bound notepad symbol for their "directory" resources page, which may be less recognizable. Research shows that users with mental health disorders often find abstract icons challenging, highlighting potential usability barriers associated with these icon designs [48].

Encouraging User Agency

Study findings indicated that while major platforms have implemented some "self-moderation" strategies to comply with guidelines in the OSA [17], significant gaps remain in how they promote user empowerment, particularly for those engaging with mental health–related content or experiencing mental health difficulties. For instance, TikTok and Instagram introduced screen break tools to encourage user control, but these tools allowed users to return after short periods or to be easily overridden, potentially undermining their effectiveness in favor of user retention. This reinforces criticisms that commercial interests may outweigh ethical considerations in social networking design [49].

In contrast, Tumblr's limited self-moderation tools and weak content moderation increased users' potential exposure to harmful content, while Tellmi, although also offering fewer self-moderation tools, compensated with robust content moderation offering better avoidance from unsafe material. However, the supportive nature of interactions on Tellmi could create a sense of responsibility among users, potentially resulting in excessive engagement that current "self-moderation" tools do not address [50]. In addition, the absence of Tellmi moderators overnight could result in user migration to other platforms, leading to riskier engagements elsewhere [51]. Alternatively, TikTok's "refresh your for you feed" feature was a novel self-moderation tool that empowered users and allowed for continual platform engagement, highlighting benefits for both users and SNSs. However, as with other self-moderation tools, its effectiveness relied on users recognizing the need for online behavior change and having the capacity to act, which

XSL•FO
RenderX

may be challenging for individuals experiencing poorer mental health or with lower metacognition skills [26].

This highlights a critical limitation found across platforms—the lack of proactive intervention to raise awareness about "self-moderation" tools. These findings align with the survey by Ofcom video-sharing platform [52], which found that only 42% of UK-based video-sharing platform users were aware of platform safety mechanisms. Along with minimal prompts directing users to tools, our findings revealed insufficient education on their functionality and usefulness. There was also a significant lack of information on tool adoption or any evidence that they underwent user testing, raising questions about their efficacy. This is especially relevant considering a recent research by Bright et al [35] indicating users' dissatisfaction with online safety technology. These findings suggest platforms may be superficially complying with OSA requirements rather than genuinely committing to empowering users. They also indicate a broader failure across the industry to address the importance of digital literacy surrounding "self-moderation" tools.

Platforms in this study also attempted to improve user safety by implementing default feature restrictions. However, default sensitivity or mature content controls often lacked clarity on whether they addressed mental health–related content or the types of content covered under each level, potentially confusing users [8]. On TikTok and Tumblr, the strictest restrictions were largely applied until users reached the age of 18 years, suggesting platforms may assume age as a qualifier of user efficacy. This overlooks the risks of removing protective measures for users who have not developed the skills or awareness needed to manage their safety independently [53]. Alternatively, Instagram, via its "teen accounts" approach allowed teenaged users to modify defaults with parental approval, encouraging a more open dialogue between users and their families. However, this approach shifts responsibility from platforms to parents. These measures also fail to address prevalent issues of age falsification among young users [54] and parents permitting their children to use social media before reaching the minimum age [52,55]. In contrast, Tellmi stood out by allowing all users to adjust their sensitivity content controls, prioritizing user agency over imposed controls [53].

Despite a lack of proactive user empowerment strategies, the changes made by platforms to improve user safety have likely been driven by the introduction of recent regulations such as the Digital Services Act in Europe [56] and the OSA in the United Kingdom [17]. These frameworks have specifically prompted platforms to implement changes that focus on the safety of young users and the moderation of dangerous or harmful content. However, the effectiveness of regulation implementation remains unclear [57,58]. For instance, an initial case study of Instagram's "teen accounts" revealed that young people using the feature remain exposed to inappropriate content [59].

In addition, while platforms adapt to meet the technical requirements of regulations, they likely remain unable to overcome long-standing limitations in moderating harmful content without further support or guidance from legislators. Current regulations also overlook variability between types of SNS platforms, including differences in resource availability and community make-up. This lack of insight risks the introduction of regulation that may inadvertently suppress or disempower online users. Country-specific regulations further fail to account for workarounds by users, such as the use of virtual private networks, allowing them to bypass geographical restrictions and access platforms from countries where regulations do not exist. This suggests that although platforms may implement changes in response to legislation, ongoing challenges in user safety and content moderation could remain unaddressed due to gaps in regulatory focus and scope.

Our findings also showed that platforms often implied it was the user's responsibility to identify and report harmful content, which may be problematic for susceptible users [8] and those unaware of reporting mechanisms [60]. In addition, Instagram's claim that reported content would feature less for the reporter could discourage the behavior, as users who engage with mental health–related content may fear limiting their own access to resources or information [26]. Observations also revealed that transparency reports on harmful content removal, when available, were difficult for users to locate and interpret. Improving their visibility could alleviate user pressure by showing that much disallowed mental health–related content is automatically flagged, but the continued reliance on user reporting for content that evades moderation systems may undermine confidence.

Platforms did make efforts to streamline reporting, with Instagram and TikTok offering feedback mechanisms and Tellmi providing direct communication with human moderators. However, it remains unclear whether the reporter or reportee receives any support after harmful content is reported, and the absence of feedback or mental health resources for reporters may discourage action [8]. Furthermore, if reported users are not offered appropriate resources, platforms may inadvertently punish individuals for sharing mental health struggles, reinforcing stigma [61]. This highlights the need for greater transparency around platform reporting procedures and clearer communication about the resources and support available to those whose content is reported.

Posting Harmful Mental Health–Related Content

All platforms gave some indication of the types of mental health content that were prohibited. However, while TikTok and Instagram included in-depth examples of allowed and disallowed content, other platforms lacked similar specificity. Furthermore, guidelines often overlooked mental health–related terminology, and there was minimal mention of rules around mental health topics unrelated to suicide, self-harm, or disordered eating. This lack of clarity could lead to subjective interpretations, resulting in self-censorship or unintentional posting of harmful material. Furthermore, only Instagram addressed the allowance of content depicting self-harm scars, a significant oversight by other platforms given that users have previously criticized moderation practices regarding this issue [26]. In addition, consequences for posting harmful content were frequently vague, with platforms using ambiguous language to describe repercussions. While this may allow for nuanced moderation, it risks confusing

XSL•FO
RenderX

users and deterring them from posting around their mental health, potentially perpetuating feelings of isolation and stigma [62].

Availability of Harmful Mental Health–Related Content

Harmful mental health content was not immediately visible on TikTok, Instagram, or Tellmi. However, on Tumblr, it was available following recommendation choices, posing higher risks to susceptible users [63]. After searching common mental health–related terms, all platforms other than Tellmi provided help resources, as expected for Tellmi as a platform focused on mental health discussion. In some instances, sensitive content overlays were used on harmful images on Instagram, which may empower users by giving them the opportunity to make informed decisions about engagement, encouraging metacognitive skills [64]. Although, if similar messages are presented repeatedly, it may result in desensitization [65] and could increase curiosity among susceptible users [26]. Furthermore, we found the application of overlays on Instagram was inconsistent, which may reduce their effectiveness.

Notably, one algospeak term resulted in content on Tellmi. Although this content was not harmful, its inclusion by human moderators indicated an acceptance of the language. This presents a broader issue in mental health content moderation where censoring common mental health terms can normalize algospeak terms as substitutes, effectively reverting platforms to their original state concerning mental health language moderation [11]. Similarly, mental health–related algospeak terms bypassed moderation on TikTok, Tumblr, and Instagram and resulted in harmful content. This highlights the danger that algospeak terms can more readily expose users to risky content, and their use instead of common mental health terms may exclude users from finding and sharing genuine information and support [12,13].

On Instagram and Tumblr, some user biographies featured the phrase “block don’t report.” This showed an online culture where users actively attempted to circumnavigate content moderation [66]. This direction to others may indicate that those engaging with mental health content believe users are responsible for their own safety through “self-moderation,” rather than relying on the platform to provide protection. It may also reflect a community norm, which could result in newer members feeling pressured to conform and avoid reporting. At the same time, the statement highlights the existence of a space where users feel empowered to express themselves by sharing content. This highlights an ongoing tension for platforms between allowing users to feel empowered through self-expression and ensuring their safety.

Future Research

Our work found no evidence that platforms evaluated the effectiveness of their self-moderation tools and guidance relevant to user mental health and well-being. This included assessing the usability, accessibility, and function of mental health–related resources and tools as well as the acceptability and practicality of moderation procedures. Notably, there was a lack of evidence of user input and testing, potentially raising concerns about the current efficacy and value of existing measures among users. If such evaluations are being conducted internally, platforms should be encouraged to share these insights publicly to enhance transparency, adding to user confidence and supporting learning across the industry. Future research evaluating these features remains essential to ensure that platforms meet users’ needs and do not promote ineffective or even risky features or guidance.

While this study did expose gaps in the current availability and accessibility of mental health and well-being resources and functions across a selection of SNS platforms, there remains a spectrum of platforms further research should aim to investigate. This includes safety and well-being resources, features, and guidance available on social media, dating, communication, e-commerce, pornography, mental health, and gaming platforms. In addition, research is needed into newly developed SNSs such as those including artificial intelligence chatbot integration plus smaller online communities that are less likely to undergo media or governmental attention.

This study resulted in a set of preliminary recommendations for industry, which give potential ways for platforms to address mental health–related content and moderation concerns. However, these recommendations need further validation and refinement, with stakeholder input, to ensure their applicability and effectiveness.

Limitations

Due to the rapidly evolving nature of online platforms, we acknowledge that the data collected may become outdated. This research aims to capture a moment in time, and it is possible and hopeful that platforms have since updated their practices to address some of the issues highlighted and align with our recommendations. Nevertheless, we believe these observations remain valuable in understanding user experiences. In addition, the recommendations provided are likely to be relevant for similar emerging or popular platforms.

We also understand the limitations of using the walkthrough method in this study. This approach relies on subjective interpretations of researchers’ observations, which may have led to biases. Although the checklist approach allowed the researchers to be systematic in their investigation of the platforms, it did not represent the real-world experience of a user, limiting the ecological validity. To mitigate these limitations, we included both a digital native and a digital immigrant researcher in an attempt to understand different interpretations and perspectives of platforms. However, as a next step, research should engage directly with users to reflect on their experiences of online mental health–related content moderation and safety.

Tellmi was included in this study as a smaller UK-based SNS, characterized by its user-generated content and social interaction around mental health. We recognize that Tellmi is not equivalent to Instagram, TikTok, or Tumblr in terms of target user populations, content diversity, and platform functionalities. However, given its direct focus on mental health and younger users, we considered Tellmi valuable as a comparator to platforms not specifically related to mental health in evaluating

XSL•FO
RenderX

their management of user features, guidance, and resources. Through our analysis, we identified notable practice differences and approaches across all platforms that helped to refine our recommendations.

Conclusions

These findings indicate that while platforms use some successful strategies, they are not effectively addressing the mental health moderation and well-being needs of all users. While larger platforms that undergo governmental scrutiny may introduce changes, these can seem superficial, appearing to prioritize user retention and engagement over meaningful safety intervention. Smaller platforms which face less scrutiny could also present risks due to weaker automated moderation systems resulting in access to harmful content or challenges in managing high user traffic, causing potential engagements with alternative less-moderated platforms. Using human moderation before material is posted may be the most effective approach to avoiding harmful content availability but is likely unfeasible for larger, global platforms due to scalability issues. It also raises concerns about limiting online freedoms, which could have negative consequences for users, particularly for those intentionally sharing harmful content. Platforms should focus on empowering users by educating them about safety mechanisms early, giving them agency over their decisions, and providing clear information regarding content moderation. In addition, while platforms have a responsibility toward the safety of young users, they should also consider other susceptible users, such as those with mental health disorders, during their moderation policy decisions. This study also highlights the need for improvements in the scope and global relevance of government regulations to better address user needs on platforms. Further research is essential to understand the effectiveness of existing interventions used by platforms, including self-moderation tools. User involvement is also necessary for meaningful improvement.

Data Availability

The datasets generated and analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

LB has previously collaborated with Tellmi on research to evaluate and enhance their platform.

Multimedia Appendix 1
Walkthrough checklist template.
[DOCX File , 21 KB - humanfactors_v12i1e69817_app1.docx]

Multimedia Appendix 2
Secondary walkthrough checklist template.
[DOCX File , 16 KB - humanfactors_v12i1e69817_app2.docx]

Multimedia Appendix 3
Comparison table (trigger warning: description of graphic and harmful content).
[XLSX File (Microsoft Excel File), 52 KB - humanfactors_v12i1e69817_app3.xlsx]

References

1. Bayer JB, Tri u P, Ellison NB. Social media elements, ecologies, and effects. Annu Rev Psychol 2020 Jan 04;71(1):471-497. [doi: 10.1146/annurev-psych-010419-050944] [Medline: 31518525]
2. Naslund JA, Bondre A, Torous J, Aschbrenner KA. Social media and mental health: benefits, risks, and opportunities for research and practice. J Technol Behav Sci 2020 Sep 20;5(3):245-257 [FREE Full text] [doi: 10.1007/s41347-020-00134-x] [Medline: 33415185]
3. Bucci S, Schwannauer M, Berry N. The digital revolution and its impact on mental health care. Psychol Psychother 2019 Jun 28;92(2):277-297. [doi: 10.1111/papt.12222] [Medline: 30924316]
4. Stiglic N, Viner RM. Effects of screentime on the health and well-being of children and adolescents: a systematic review of reviews. BMJ Open 2019 Jan 03;9(1):e023191 [FREE Full text] [doi: 10.1136/bmjopen-2018-023191] [Medline: 30606703]
5. Mittal VA, Tessner KD, Walker EF. Elevated social Internet use and schizotypal personality disorder in adolescents. Schizophr Res 2007 Aug;94(1-3):50-57 [FREE Full text] [doi: 10.1016/j.schres.2007.04.009] [Medline: 17532188]
6. World mental health report: transforming mental health for all. World Health Organization. 2022. URL: https://www.who.int/publications/i/item/9789240049338 [accessed 2024-10-17]
7. Jhaver S, Birman I, Gilbert E, Bruckman A. Human-machine collaboration for content regulation: the case of reddit automoderator. ACM Trans Comput Hum Interact 2019 Jul 19;26(5):1-35 [FREE Full text] [doi: 10.1145/3338243]

XSL•FO
RenderX

8. Haime Z, Kennedy L, Grace L, Biddle L. Experiences of moderation, moderators, and moderating by online users who engage with self-harm and suicide content. medRxiv Preprint posted online February 17, 2024 [FREE Full text] [doi: 10.1101/2024.02.15.24302878]
9. Biddle L, Derges J, Goldsmith C, Donovan JL, Gunnell D. Using the internet for suicide-related purposes: contrasting findings from young people in the community and self-harm patients admitted to hospital. PLoS One 2018 May 24;13(5):e0197712 [FREE Full text] [doi: 10.1371/journal.pone.0197712] [Medline: 29795637]
10. Smith H, Cipolli W. The Instagram/Facebook ban on graphic self‐harm imagery: a sentiment analysis and topic modeling approach. Policy Internet 2021 Oct 29;14(1):170-185 [FREE Full text] [doi: 10.1002/poi3.272]
11. Lorenz T. Internet 'algospeak' is changing our language in real time, from 'nip nops' to 'le dollar bean'. The Washington Post. URL: https://www.washingtonpost.com/technology/2022/04/08/algospeak-tiktok-le-dollar-bean/ [accessed 2024-05-29]
12. Gerrard Y. Beyond the hashtag: circumventing content moderation on social media. New Media Soc 2018 May 28;20(12):4492-4511 [FREE Full text] [doi: 10.1177/1461444818776611]
13. Chancellor S, Pater JA, Clear T, Gilbert E, De Choudhury M. #thyghgapp: Instagram content moderation and lexical variation in pro-eating disorder communities. In: Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing. 2016 Presented at: CSCW '16; February 27-March 2, 2016; Francisco, CA p. 1201-1213 URL: https://dl.acm.org/doi/10.1145/2818048.2819963 [doi: 10.1145/2818048.2819963]
14. Stawarz K, Preist C, Coyle D. Use of smartphone apps, social media, and web-based resources to support mental health and well-being: online survey. JMIR Ment Health 2019 Jul 12;6(7):e12546 [FREE Full text] [doi: 10.2196/12546] [Medline: 31301126]
15. Schaffner B, Bhagoji AN, Cheng S, Mei S, Shen JL, Wang G, et al. "Community guidelines make this the best party on the internet": an in-depth study of online platforms' content moderation policies. In: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems. 2024 Presented at: CHI '24; May 11-16, 2024; Honolulu, HI p. 1-16 URL: https://dl.acm.org/doi/10.1145/3613904.3642333 [doi: 10.1145/3613904.3642333]
16. Obar JA, Oeldorf-Hirsch A. The biggest lie on the Internet: ignoring the privacy policies and terms of service policies of social networking services. Inf Commun Soc 2018 Jul 03;23(1):128-147 [FREE Full text] [doi: 10.1080/1369118X.2018.1486870]
17. Online safety act. Government of UK. 2023. URL: https://www.legislation.gov.uk/ukpga/2023/50/contents [accessed 2025-05-19]
18. Troeger J, Bock A. The sociotechnical walkthrough – a methodological approach for platform studies. Stud Commun Sci 2022 Jun 29;22(1):43-52. [doi: 10.24434/j.scoms.2022.01.3064]
19. Light B, Burgess J, Duguay S. The walkthrough method: an approach to the study of apps. New Media Soc 2016 Nov 11;20(3):881-900 [FREE Full text] [doi: 10.1177/1461444816675438]
20. Gillett R. "This is not a nice safe space": investigating women's safety work on Tinder. Fem Media Stud 2021 Jun 29;23(1):199-215. [doi: 10.1080/14680777.2021.1948884]
21. Ethics guidelines for internet-mediated research. British Psychological Society. 2021. URL: https://doi.org/10.53841/bpsrep.2021.rep155 [accessed 2025-05-29]
22. Most popular social media platforms in the United Kingdom (UK) as of the third quarter 2023, by usage reach. Statista. 2023. URL: https://www.statista.com/statistics/284506/united-kingdom-social-network-penetration/ [accessed 2025-05-29]
23. Online Nation 2024 Report. Ofcom. URL: https://www.ofcom.org.uk/siteassets/resources/documents/research-and-data/online-research/online-nation/2024/online-nation-2024-report.pdf?v=386238 [accessed 2025-05-22]
24. Hoover A. Gen Z flocks to Tumblr, the one social media millennials didn't ruin. Business Insider. URL: https://www.businessinsider.com/gen-z-flocking-tumblr-millennials-musk-zuckerberg-safe-space-2025-4 [accessed 2025-05-29]
25. Upton-Clark E. The Tumblr revival is real—and Gen Z is leading the charge. The Fast Company. 2025 Apr 3. URL: https://www.fastcompany.com/91310237/the-tumblr-revival-is-real-and-gen-z-is-leading-the-charge [accessed 2025-05-22]
26. Haime Z, Kennedy L, Grace L, Cohen R, Derges J, Biddle L. The journey of engaging with web-based self-harm and suicide content: longitudinal qualitative study. JMIR Infodemiology 2024 Mar 28;4:e47699 [FREE Full text] [doi: 10.2196/47699] [Medline: 38546718]
27. Managing self-harm and suicide content online: guidelines for sites and platforms hosting user-generated content. Samaritans. 2020. URL: https://media.samaritans.org/documents/Online_Harms_guidelines_FINAL_1.pdf [accessed 2025-05-29]
28. How should we talk about mental health and digital technology with young people? Digital Dialogues. URL: https://www.digitaldialogues.co.uk/ [accessed 2025-05-29]
29. Campbell KA, Orr E, Durepos P, Nguyen L, Li L. Reflexive thematic analysis for applied qualitative health research. Qual Rep 2021;26(6):2011-2028 [FREE Full text] [doi: 10.46743/2160-3715/2021.5010]
30. Braun V, Clarke V. Thematic analysis. In: Cooper H, Camic PM, Long DL, Panter AT, Rindskopf D, Sher KJ, editors. APA handbook of research methods in psychology, Vol 2: Research designs: Quantitative, qualitative, neuropsychological, and biological. Washington, DC: American Psychological Association; 2012:57-71.
31. Wernerson N, Söderblom Carlsson E. Increasing user engagement and feature discoverability through user onboarding in business-to-business applications. Lund University. 2019. URL: https://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=8988351&fileOId=8988353 [accessed 2025-05-29]

XSL•FO
RenderX

32. Kaipainen K, Välkkynen P, Kilkku N. Applicability of acceptance and commitment therapy-based mobile app in depression nursing. Transl Behav Med 2017 Jun;7(2):242-253 [FREE Full text] [doi: 10.1007/s13142-016-0451-3] [Medline: 27896798]
33. Cardoso MC. The onboarding effect: leveraging user engagement and retention in crowdsourcing platforms. In: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems. 2017 Presented at: CHI EA '17; May 6-11, 2017; Denver, CO p. 263-267 URL: https://doi.org/10.1145/3027063.3027128 [doi: 10.1145/3027063.3027128]
34. Robotham D, Satkunanathan S, Doughty L, Wykes T. Do we still have a digital divide in mental health? A five-year survey follow-up. J Med Internet Res 2016 Nov 22;18(11):e309 [FREE Full text] [doi: 10.2196/jmir.6511] [Medline: 27876684]
35. Bright J, Enock FE, Johansson P, Margetts HZ, Stevens F. Understanding engagement with platform safety technology for reducing exposure to online harms. arXiv Preprint posted online January 3, 2024 [FREE Full text]
36. Auf H, Dagman J, Renström S, Chaplin J. Gamification and nudging techniques for improving user engagement in mental health and well-being apps. Proc Des Soc 2021 Jul 27;1:1647-1656. [doi: 10.1017/pds.2021.426]
37. Smullen D, Feng Y, Zhang SA, Sadeh N. The best of both worlds: mitigating trade-offs between accuracy and user burden in capturing mobile app privacy preferences. Proc Priv Enh Technol 2020:195-215 [FREE Full text] [doi: 10.2478/popets-2020-0011]
38. Russell I. Debate: more, not less social media content moderation? How to better protect youth mental health online. Child Adolesc Ment Health 2024 Sep;29(3):319-321 [FREE Full text] [doi: 10.1111/camh.12717] [Medline: 38738680]
39. Kim J, Min JE, Le LH. Impacts of brand familiarity and brand responses on perceived brand credibility, similarity, and blog recommendation intention: a study of corporate blogs. J Fash Mark Manag 2021 Jun 22;26(2):328-343. [doi: 10.1108/jfmm-09-2020-0189]
40. Connolly B. Engagement and trust. In: Connolly B, editor. Digital Trust: Social Media Strategies to Increase Trust and Engage Customers. New York, NY: Bloomsbury Business; Feb 18, 2020.
41. Matthes J, Karsay K, Schmuck D, Stevic A. “Too much to handle”: impact of mobile social networking sites on information overload, depressive symptoms, and well-being. Comput Hum Behav 2020 Apr;105:106217. [doi: 10.1016/j.chb.2019.106217]
42. Cohen R, Rifkin-Zybutz R, Moran P, Biddle L. Web-based support services to help prevent suicide in young people and students: a mixed-methods, user-informed review of characteristics and effective elements. Health Soc Care Community 2022 Nov 06;30(6):2404-2413 [FREE Full text] [doi: 10.1111/hsc.13819] [Medline: 35514202]
43. Rotondi AJ, Spring MR, Hanusa BH, Eack SM, Haas GL. Designing eHealth applications to reduce cognitive effort for persons with severe mental illness: page complexity, navigation simplicity, and comprehensibility. JMIR Hum Factors 2017 Jan 05;4(1):e1 [FREE Full text] [doi: 10.2196/humanfactors.6221] [Medline: 28057610]
44. Kington RS, Arnesen S, Chou WS, Curry SJ, Lazer D, Villarruel AM. Identifying credible sources of health information in social media: principles and attributes. NAM Perspect 2021 Jul 16;2021:29 [FREE Full text] [doi: 10.31478/202107a] [Medline: 34611600]
45. Kirtley OJ, O'Connor RC. Suicide prevention is everyone's business: challenges and opportunities for Google. Soc Sci Med 2020 Oct;262:112691 [FREE Full text] [doi: 10.1016/j.socscimed.2019.112691] [Medline: 32247626]
46. Celeste E. From content moderation to content governance. In: Celeste E, Palladino N, Redeker D, Yilma K, editors. The Content Governance Dilemma: Digital Constitutionalism, Social Media and the Search for a Global Standard. Cham, Switzerland: Palgrave Macmillan; 2023.
47. Khadka N. An usability and universal design investigation into hamburger menus. OsloMet-storbyuniversitetet. URL: https://oda.oslomet.no/oda-xmlui/bitstream/handle/11250/2777325/khadka-mauu2021.pdf?sequence=1&isAllowed=y [accessed 2025-05-29]
48. Ferron JC, Brunette MF, Geiger P, Marsch LA, Adachi-Mejia AM, Bartels SJ. Mobile phone apps for smoking cessation: quality and usability among smokers with psychosis. JMIR Hum Factors 2017 Mar 03;4(1):e7 [FREE Full text] [doi: 10.2196/humanfactors.5933] [Medline: 28258047]
49. Geeling S, Brown I. Towards ethical social media practice: a grounded theory for analyzing social media platform ethics. SSRN Journal Preprint posted online October 18, 2021 [FREE Full text] [doi: 10.2139/ssrn.3942451]
50. Cliffe B, Linton MJ, Haime Z, Biddle L. Young adults' experiences of using a young person's mental health peer support app: a qualitative interview study. PLOS Digit Health 2024 Jul 31;3(7):e0000556 [FREE Full text] [doi: 10.1371/journal.pdig.0000556] [Medline: 39083458]
51. Ribeiro MH, Jhaver S, Zannettou S, Blackburn J, Stringhini G, De Cristofaro E, et al. Do platform migrations compromise content moderation? Evidence from r/the_Donald and r/Incels. Proc ACM Hum Comput Interact 2021 Oct 18;5(CSCW2):1-24 [FREE Full text] [doi: 10.1145/3476057]
52. Online nation: 2023 report. Ofcom. 2023. URL: https://www.ofcom.org.uk/siteassets/resources/documents/research-and-data/online-research/online-nation/2023/online-nation-2023-report.pdf?v=368355 [accessed 2025-05-29]
53. Livingstone S, Byrne J. Parenting in the digital age - the challenges of parental responsibility in comparative perspective. In: Mascheroni G, Ponte C, Jorge A, editors. Digital Parenting: The Challenges for Families in the Digital Age. Gothenburg, Sweden: Nordicom; Feb 15, 2019:90-91.
54. Lichy J, McLeay F, Burdfield C, Matthias O. Understanding pre‐teen consumers social media engagement. Int J Consum Stud 2022 May 23;47(1):202-215. [doi: 10.1111/ijcs.12821]

XSL•FO
RenderX

55. Boyd D, Hargittai E. Connected and concerned: variation in parents' online safety concerns. Policy Internet 2013;5(3):245-269. [doi: 10.1002/1944-2866.POI332]
56. Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) (Text with EEA relevance. European Union. URL: https://eur-lex.europa.eu/eli/reg/2022/2065/oj/eng [accessed 2025-05-22]
57. Nash V, Felton L. Treating the symptoms or the disease? Analysing the UK online safety act's approach to digital regulation. Policy Internet 2024 Jul 19;16(4):818-832 [FREE Full text] [doi: 10.1002/poi3.404]
58. Neudert LM. Regulatory capacity capture: the United Kingdom's online safety regime. Internet Policy Rev 2023;12(4):e009173. [doi: 10.14763/2023.4.1730]
59. Is Instagram now safe for teens? Instagram teen accounts case study. 5rights Foundation. 2025. URL: https://5rightsfoundation.com/wp-content/uploads/2025/04/TeenAccounts_Case-Study_Final.pdf [accessed 2025-05-29]
60. Record RA, Straub K, Stump N. #Selfharm on #Instagram: examining user awareness and use of Instagram's self-harm reporting tool. Health Commun 2020 Jun 08;35(7):894-901. [doi: 10.1080/10410236.2019.1598738] [Medline: 30961389]
61. Suzor NP, West SM, Quodling A, York JC. What do we mean when we talk about transparency? Toward meaningful transparency in commercial content moderation. Int J Commun 2019;13(1):60 [FREE Full text]
62. Zhang CC, Zaleski G, Kailley JN, Teng KA, English M, Riminchan A, et al. Debate: social media content moderation may do more harm than good for youth mental health. Child Adolesc Ment Health 2024 Feb 13;29(1):104-106. [doi: 10.1111/camh.12689] [Medline: 38088464]
63. Cavazos-Rehg PA, Krauss MJ, Sowles SJ, Connolly S, Rosas C, Bharadwaj M, et al. An analysis of depression, self-harm, and suicidal ideation content on Tumblr. Crisis 2017 Jan;38(1):44-52 [FREE Full text] [doi: 10.1027/0227-5910/a000409] [Medline: 27445014]
64. Burr C, Morley J. Empowerment or engagement? Digital health technologies for mental healthcare. In: Burr C, Milano S, editors. The 2019 Yearbook of the Digital Ethics Lab. Digital Ethics Lab Yearbook. Cham, Switzerland: Springer; 2020.
65. Daryanani M. Desensitizing the user: a study of the efficacy of warning messages. University of Oxford. 2011. URL: https://www.europoli.org/MSc/Dissertation.pdf [accessed 2025-05-29]
66. Savic M, McCosker A. Circumventing social media restrictions: preteens' use of Musical.ly. ICA Preconference. 2018. URL: https://www.academia.edu/38593383/Circumventing_social_media_restrictions_Preteens_use_of_Musical_ly [accessed 2025-05-29]

Abbreviations

OSA: Online Safety Act
SNS: social networking site
ToS: terms of service

Edited by A Kushniruk; submitted 09.12.24; peer-reviewed by J Walsh, B Cahill; comments to author 20.04.25; revised version received 29.04.25; accepted 12.05.25; published 29.05.25.

Please cite as:
Haime Z, Biddle L
Exploring Mental Health Content Moderation and Well-Being Tools on Social Media Platforms: Walkthrough Analysis
JMIR Hum Factors 2025;12:e69817
URL: https://humanfactors.jmir.org/2025/1/e69817
doi:10.2196/69817
PMID:

©Zoë Haime, Lucy Biddle. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 29.05.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Examining Patient Characteristics Associated With Digital Outpatient Care for Type 1 Diabetes (DigiDiaS): Cross-Sectional Study

Ingeborg Spildo[1], MSc; Heidi Holmen[1], RN, MSc, PhD; Annesofie Lunde Jensen[2,3], RN, MSc, PhD; Milada Hagen[1], MSc, PhD; Tone Singstad[4], RN, MSc; Jacob Andreas Winther[4], MD, PhD; Eirik Årsand[5], PhD; Astrid Torbjørnsen[1], RN, MSc, PhD

[1]OsloMet – Oslo Metropolitan University, Oslo, Norway

[2]Aarhus University, Aarhus, Denmark

[3]Steno Diabetes Centers, Aarhus, Denmark

[4]Akershus University Hospital, Lørenskog, Norway

[5]UiT The Arctic University of Norway, Tromsø, Norway

Corresponding Author:
Ingeborg Spildo, MSc
OsloMet – Oslo Metropolitan University
Pilestredet 32
Oslo, 0166
Norway
Phone: 47 67236106
Email: ingspi@oslomet.no

Abstract

Background: People with type 1 diabetes require ongoing self-management and frequent follow-up care. Digital care models might offer flexible solutions and increased sustainability, and while the clinical opportunities of these care models have been explored, the characteristics of patients inclined to opt for digital care remain unclear.

Objective: This study aimed to investigate which selected sociodemographic and disease-related patient characteristics are associated with opting for digital outpatient care among patients with type 1 diabetes.

Methods: This cross-sectional study was conducted at the endocrinology department of Akershus University Hospital in Norway, as part of a larger longitudinal study. Adult patients with type 1 diabetes were eligible to participate and were recruited consecutively. Patients could choose either a novel, mobile health–based, digital, tailored, outpatient care model (DigiDiaS care) or continuation of usual care. DigiDiaS care is delivered via an app comprising a message service; preconsultation questionnaires; options for physical, video, or telephone consultations; an information page; and an e-learning course. Sociodemographic and clinical data were collected from the participants' medical records and the national diabetes registry. Self-reported data comprised self-management measured using the Patient Activation Measure 13 (PAM-13), diabetes distress (20-item Problem Areas in Diabetes; PAID-20), well-being (World Health Organization-Five Well-Being Index; WHO-5), and health literacy (12-item short version of the European Health Literacy Survey Questionnaire; HLS-19 Q12) questionnaires. We explored group differences and conducted logistic regression to identify factors associated with opting for DigiDiaS care versus usual care.

Results: A total of 237 patients consented to participate in the study, with 185 (78.1%) opting for DigiDiaS care and 52 (21.9%) opting for usual care. The DigiDiaS care group had a statistically significantly shorter duration of diabetes (median 19, range 0-51 years) compared with the usual care group (median 29, range 3-58 years; P<.001); higher proportions of users of insulin pumps than insulin pens for insulin delivery (DigiDiaS care: 74/185, 40%; usual care: 12/185, 23.1%; P=.007), and a lower median well-being care score (DigiDiaS care: median score 60, range 4-96; usual care: median score 68, range 16-100; P=.04). The DigiDiaS care and usual care groups did not differ in sociodemographic variables, presence of late complications from diabetes or comorbidities, self-management, diabetes distress, or health literacy.

Conclusions: This study reveals that most patients with type 1 diabetes choose digital outpatient care when it is offered. Our study suggests that those opting for DigiDiaS care are already familiar with using diabetes-related technology, have a shorter diabetes duration, and have lower well-being. It is essential to understand the characteristics of patients who opt for usual care

XSL•FO
RenderX

to ensure high-quality health care services. Further studies should also focus on how the implementation of digital solutions in outpatient care can affect how and if patients use them.

International Registered Report Identifier (IRRID): RR2-10.2196/52766

(JMIR Hum Factors 2025;12:e73708) doi:10.2196/73708

KEYWORDS

diabetes mellitus; type 1; telemedicine; remote consultation; telenursing; distance counseling; delivery of health care; patient-reported outcome measures

Introduction

Background

Living with type 1 diabetes can be challenging and is associated with health complications such as poor mental well-being and poor quality of life, and it requires lifelong self-management and support to prevent these complications [1-3]. A life with type 1 diabetes can affect patients' ability to self-manage, increasing their risk of diabetes complications [4,5]. Often, diabetes care and self-management education and support are provided through multidisciplinary teams [6-8]. However, the traditional calendar-based approach in outpatient care is often rigid, giving patients little opportunity to influence their care. Consequently, the calendar-based approach threatens health care service resource sustainability and presents several challenges; it is not given that outpatient care is appropriate or necessary at the time of the consultation [7,9,10]. Conversely, a patient-centered approach in outpatient care, where patients themselves influence the timing and nature of treatment based on patients preferences and needs, has been suggested as a solution to these problems [11]. Furthermore, individualization through patient-centered care can strengthen the patient's autonomy and participation in their own treatment.

The increasing focus on individualizing care aligns with an increased recognition of how digital solutions in outpatient care can provide closer interaction with health care services [8]. Such digital outpatient care may include various forms of synchronous and asynchronous remote patient monitoring; telephone and video consultations; messaging services; connection to wireless tools, such as glucose sensors and insulin pumps; or other forms of telecommunication tools [12]. Allowing and facilitating a more active patient role, patient-reported outcomes (PROs) can be integrated into digital tools for follow-up care [8]. PROs offer direct feedback from patients themselves about their health, which, among other things, provides a broader perspective on clinical goals and serves as a conversational tool [13,14]. While PROs in clinical care have gained increasing attention in recent years, a scoping review found the use of PRO as part of routine treatment in diabetes outpatient settings to be scarce [15].

Although interest in and recognition of the opportunities for digital outpatient care have been acknowledged by international panels of health care professionals in recent years [8], few studies have investigated PROs in digital outpatient care [15]. A pioneering study from Denmark used the PRO measure DiabetesFlex before scheduled consultations delivered in a digital solution to allow patients to voice their need for consultations or cancel if not deemed necessary [16]. This resulted in reduced face-to-face visits and improved well-being and diabetes distress while providing safe diabetes management. In addition, DiabetesFlex altered the interactions between clinicians and patients as it shifted toward the patients' perspectives on living with diabetes [17]. It increased patient involvement, strengthened their sense of responsibility, and tailored diabetes care to meet individual needs [16,17].

Despite previous research emphasizing the positive contribution of digital outpatient care solutions in diabetes outpatient care, there remains a lack of research to establish an evidence base. For instance, not everyone wanted to continue with DiabetesFlex after the end of the study period, despite their positive experiences [17]. Furthermore, there is a lack of knowledge on the characteristics of patients choosing digital or flexible diabetes follow-ups. Characteristics such as sociodemographics, clinical measures, and disease-specific measures can reveal whether users and nonusers of digital outpatient care differ and whether they possess specific traits or factors. Exploring such characteristics can provide insights into the actions and choices of patients with type 1 diabetes in their outpatient care [18-20].

Objective

This study aims to investigate which selected sociodemographic and disease-related patient characteristics are associated with opting for digital outpatient care among patients with type 1 diabetes.

Methods

Study Design

This study used a cross-sectional design to investigate characteristics of patients with type 1 diabetes opting for digital outpatient care compared with those opting for usual care. The data used in this study is collected as part of a larger multimethod prospective observational study assessing the integration of digital outpatient care procedures for patients with type 1 diabetes and their clinicians [21]. In this paper, we explored the baseline characteristics of these participants, and the recruitment of participants and data collection took place from October 2022 to October 2023. This study is reported according to the Strengthening the Reporting of Observational studies in Epidemiology (STROBE) reporting guideline [22].

Setting

The study was carried out at the Endocrinological Outpatient Clinic at Akershus University Hospital. In November 2021, the clinic developed and introduced a voluntary digital mobile health (mHealth) supplement outpatient care program specifically tailored to patients with type 1 diabetes called DigiDiaS.

XSL•FO
RenderX

Participants

Eligibility Criteria

The inclusion and exclusion criteria for participants are listed in Textbox 1. Both users and nonusers of DigiDiaS care were eligible for participation.

Textbox 1. Eligibility criteria.

Inclusion criteria

- Aged ≥18 years
- Having type 1 diabetes and receiving treatment and care at the endocrinological outpatient clinic at Akershus University Hospital
- Ability to read and understand Norwegian

Exclusion criteria

- Having type 2 diabetes or gestational diabetes
- Individuals with impaired cognitive abilities who are unable to complete the self-reported data independently or understand the nature of the study

Recruitment

Participants were recruited through consecutive sampling. There was no randomization, and the participants could choose whether they opted for DigiDiaS care or to continue with usual care. Participants were introduced to the study and invited to participate during a routine consultation with a diabetes specialist nurse or an endocrinologist. If they were not already using the digital outpatient care DigiDiaS, they were provided with details on both the study and the digital outpatient care solution. For patients interested in DigiDiaS care and new to this digital outpatient care solution, their diabetes specialist nurse provided a brief introduction to the app and offered support to download it, if needed. The patients received brief written and oral information about the study and were asked if a researcher (IS) could contact them. Those consenting to further information received a phone call before they received a link to a digital consent form. If a patient expressed a desire to participate in the study over the phone but did not consent, a reminder was sent via text message. The written informed consent form was also made available in paper format, if needed.

Diabetes Outpatient Care

Usual Care

At the endocrinology outpatient clinic in which this study was conducted, usual care for type 1 diabetes follows a routine, calendar-based system in accordance with Norwegian national diabetes guidelines [6]. In Norway, all patients with type 1 diabetes are offered follow-up care in specialist outpatient care, provided by interdisciplinary health care teams, including endocrinologists, diabetes-specific nurses, and dieticians. This care includes an extended annual check-up and a routine midyear check-up. Patients with recent diagnoses of diabetes, comorbidities, late complications, or poorly controlled diabetes may receive more frequent consultations. Under this system, appointments are approximately prescheduled, giving patients limited flexibility to influence the timing and nature of their own outpatient care. The usual care group in this study continued with this routine—calendar-based system of receiving a prescheduled time for consultation with limited flexibility to influence their follow-up care.

DigiDiaS Care

DigiDiaS care serves as a flexible digital supplement designed to support the patients' interactions with clinicians during outpatient care, allowing the individual patient's needs to guide contact with clinicians [21]. DigiDiaS care is delivered through a free app that patients can download onto their private smartphones from the App Store or Google Play, following an invitation from a health care professional. The main components in DigiDiaS care provided by Dignio Connected Care are simple and intuitive to use, comprising the patient app MyDignio and the clinician's software DignioPrevent delivered through a web browser via a direct link from the medical record system. Dignio Connected Care meets the necessary safety requirements, including CE certification [23].

The digital outpatient care solution includes several features in the MyDignio app for patients. First, a messaging service enables patients to contact the health care service outside of planned outpatient care, if necessary, with asynchronous messages. Second, before scheduled follow-up consultations, patients receive a PRO-based questionnaire, which includes questions about well-being, diabetes-specific outcomes, and the need for a consultation and questions for patients to prepare topics they want to address in the scheduled outpatient care. The PRO-based questionnaire is based on the Danish DiabetesFlex [16]. It was adapted to Norwegian and further refined in collaboration with user representatives, clinicians, and managers to suit the digital platform used at the clinic [13]. The third feature of DigiDiaS care is that patients can choose between remote video consultation, telephone consultation, or in-person attendance for routine outpatient care consultations. Finally, clinicians can provide patients with e-learning courses and create a tailored information page for each patient, if needed. Patients receive notifications about new activities from the app through push notifications or text messages.

For clinicians, digital outpatient care is conducted through software accessible from web app DignioPrevent, providing clinicians with features for flexible patient care [21]. Through this system, clinicians manage the asynchronous chat, send out the PRO-based questionnaire before consultations, and edit the information page. The answers on the PRO-based questionnaires

XSL•FO
RenderX

have defined thresholds dependent on the patients' responses, using a traffic light model, with green, yellow, and red alerts and an additional orange alert if the questionnaire has not been completed. Green questionnaires indicate no issues and are not visible in DignioPrevent; however, all yellow and red questionnaires from the past are available for comparison. Clinicians also provide patients with information and e-learning courses, both through premade templates and an open-text field. It is also possible to contact the patient directly through video consultation with DignioPrevent. However, for scheduled appointments, a separate system called Whereby is used for video consultations as it communicates with the calendar within the medical record system that DignioPrevent does not.

Data Collection

Data were collected through 3 primary sources: self-reported questionnaires, electronic medical records, and the national quality registry for diabetes, the Norwegian Organization for Quality Improvement of Laboratory Examinations (NOKLUS). Self-reported data and the digital consent form were collected using Nettskjema [24], a secure platform that stores data in the Services for Sensitive Data (TSD) platform for confidential handling. To deliver the digital consent form and self-reported data, the patients received a secure link via SMS text message or email, requiring them to log in with their national ID for secure access. Patients who preferred the consent form and self-reported data on paper received this and submitted these papers back to the clinic which was stored in a locked safety box. Data from medical records and the national diabetes quality registry (NOKLUS) were manually extracted into Microsoft Excel files, stored, and managed in TSD. To ensure the accuracy of the data extraction, a coresearcher cross-checked all extracted data against the original medical records and NOKLUS registry, as described in the protocol [15].

In addition, to compare our sample with the overall hospital sample comprising all patients with type 1 diabetes and the national population of people with type 1 diabetes, we received aggregated data directly from NOKLUS on the variables age, gender, education, years since diabetes duration, insulin delivery method, and glycated hemoglobin (HbA_{1c}), all provided in median and range. Comparing our sample with the national population is essential to identify and assess biases, ensuring that our findings are representative and applicable to the broader hospital sample and national population.

To explore the role of patients' travel time from home to the outpatient clinic, an estimate of travel time was calculated based on the patients' postal code. Google Maps was used to determine the travel time by car to the main premises of the hospital. Akershus University Hospital is a local hospital with a defined coverage area. Participants living outside this area were typically students or recent movers who had not switched to their local hospital or individuals who, for other reasons, did not receive outpatient care at their local hospital. Public transport is only available for parts of the coverage area and was therefore not used in the calculation.

Measures

Sociodemographic and Clinical Variables

Sociodemographic variables, gathered from the medical records, included age, gender, education, employment, ethnicity, comorbidities, cohabitation, single resident status, travel time to the hospital, and use of other health-related applications (Table 1). Clinical variables comprised diabetes duration, HbA_{1c}, insulin delivery method, blood glucose monitoring method, time-in-range, diabetic ketoacidosis, hypoglycemia, late complications from diabetes, BMI, cholesterol, blood pressure, and comorbidities (Table 2). While time-in-range can be provided in various ranges when aiming to present the ambulatory glucose profiles of the patients, in this study, we present time-in-range for the last 14 days as this is most frequently used in this study's clinical setting. All clinical variables are from the last measurement before consent to participate in the study, except for HbA_{1c}, which is the last measurement inclusion in the digital solution for the DigiDiaS care group, and the last measurement before consent to the study for the usual care group. Self-reported data encompassed self-management, diabetes distress, well-being, and health literacy (Table 3).

XSL•FO
RenderX

Table 1. Sociodemographic variables.

Variables	Categories or data type	Merged into categories for analysis	Data source
Age	Years	__[a]	Medical records
Gender	Woman or man	—	Medical records
Education	Not completed primary school (10 y), primary upper secondary school (13 y), vocational school (13 y), college or university (≤4 y), college or university (>4 y), unknown	“13 years or less” and “over 14 years”	Medical records
Employment status	Full- or part-time employment, being a student, job seeking, receiving disability benefits, and retirement (based on the Norwegian retirement age of 67 y)	—	Self-reported
Ethnicity	European, Asian, African, or unknown	European, Asian, or African, with a separate category for missing data	Medical records
Cohabitation	Living in a partnership or not living in a partnership	Yes or no	Medical records
Single resident	Living alone or not living alone	Yes or no	Medical records
Travel time to hospital	Estimate based on postal code using Google Maps to identify travel time by car, in min	>30 min travel time, <30 min, or living outside of coverage area	Medical records and Google Maps
Use of other health-related apps	One question with the response alternatives “daily” “weekly,” “monthly,” “not using this now but having used it before.” or “not using other health-related apps”	Daily or weekly use, monthly use or used before, never or previous use	Self-reported

[a]Not applicable.

XSL•FO
RenderX

Table 2. Clinical variables.

Variables	Categories or data type	Merged into categories for analysis	Data source
Diabetes duration	• The year from onset to when participants consented to participate in this study	—[a]	NOKLUS[b] registry
HbA_{1c}[c]<75 mmol/mol	• HbA1c (mmol/mol) on a continuous scale • The proportion of participants with HbA1c above 75 mmol/L	—	NOKLUS registry
Insulin delivery	• Pump or pen	—	NOKLUS registry
Blood glucose monitoring	• CGMd or glucometer	—	NOKLUS registry
Time-in-range, for those using CGM	• Percentage of time-in-range (3.9-10 mmol/L) last 14 d	—	NOKLUS registry
Diabetic ketoacidosis	• Never • Once • Several times or unknown	Recoded into none, 1≥, missing	NOKLUS registry
Hypoglycemia	• Never • Once • Several times or unknown	Recoded into none, ≥1, missing	NOKLUS registry
Late complications from diabetes	• Albuminuria • Dialysis treatment • Kidney transplantation • Retinopathy • Neuropathy • Stroke • Arterial vascular surgery • Amputation • Diabetic foot ulcers	Recoded into none, 1, ≥2, and missing	NOKLUS registry
BMI (kg/m^2)	• Weight and height	BMI=weight (kg)/[height (m)]^2	NOKLUS registry
Cholesterol	• Low-density lipoprotein	—	NOKLUS registry
Blood pressure	• Systolic blood pressure and • Diastolic blood pressure in mm Hg	—	NOKLUS registry
Comorbidities	• Charlson Comorbidity Index, which includes a range of conditions such as myocardial infarction, congestive heart failure, peripheral vascular disease, and others	None, 1, ≥2, and missing	Self-reported

[a]Not applicable.

[b]NOKLUS: Norwegian Organization for Quality Improvement of Laboratory Examinations.

[c]HbA_{1c}: glycated hemoglobin.

[d]CGM: continuous glucose monitoring.

XSL•FO
RenderX

Table 3. Measures for self-management, diabetes distress, well-being, and health literacy.

Measures	Items, scale, and domains	Interpretation

Self-management (PAM-13[a])	• 13 items, rated from 1 “disagree strongly,” to 4, “strongly agree,” with an additional “not applicable”
	• 4 domains: knowledge, beliefs, confidence, and skills for managing one’s health.
		• The total PAM-13 score is calculated using a formula that weighs the various questions and that forms a scale from 0-100. Higher scores indicate higher activation [25]. A 3.2 change on a 1-100 scale is considered small but clinically relevant [26]
		• There are 4 levels:
		  • Level 1 (score ≤47), disengaged and overwhelmed
		  • Level 2 (score 47.1-55.1), becoming aware but still struggling
		  • Level 3 (score 55.2-67), taking action
		  • Level 4 (score≥67.1), maintaining behavior and pushing further
		• Levels 1 and 2, 3 and 4 were merged into categories for analysis

Diabetes distress (PAID[b])	• 20 items, ranked on a 5-point Likert scale from 0, “not a problem” to 4, “serious problem”	• Sum of all 20 items and multiplied by 1.25, giving a total score ranging from 0-100 [26] • A higher score reflects greater emotional distress. A score of ≥40 indicates severe emotional distress.
Well-being (WHO-5[c])	• The 5 items are measured in a 6-point Likert scale from 5 "all the time" to 0 "at no time".	• Sum of 5 items multiplied by 4 to create a scale from 0-100 [27] • A higher score indicates higher well-being • Score <50 may be an identification of mild to severe depressive affect [28]

Health literacy (HLS19-Q12[d])	• The 12 items are measured in a 4-point Likert scale from 1, “very hard” to 4, “very easy” with added “I don’t know”	• Sum of the 12 items is calculated. Higher scores reflect higher health literacy [29]
		• The scale is divided into 4 levels with expected characteristics:
		  • Below level 1; lack key knowledge and skills about health and may have challenges in understanding and applying health information
		  • Level 1; where you are expected to be able to access, understand, and apply health information relevant to staying healthy
		  • Level 2; access, appraise, understand, and apply health information and advice relevant to enhancing physical and mental health
		  • Level 3; access, appraise, understand, and apply health information and advice relevant to making informed health care choices by critically evaluating health claims and judiciously comparing treatments
		  • The items can also be distributed into the 3 different domains they measure—health care, disease prevention, and health promotion, and the individual domain can be scored with a minimum of points 0 maximum of 16 points per domain

[a]PAM-13: Patient Activation Measure 13.

[b]PAID: Problem Areas in Diabetes.

[c]WHO-5: World Health Organization-Five Well-Being Index.

[d]HLS19-Q12: 12-item short version of the European Health Literacy Survey Questionnaire.

Self-Management

Self-management and the ability to manage one’s health was assessed with the “Patient activation measure” short version (Patient Activation Measure 13 [PAM-13]) [30] (Table 3). Hibbard et al [30,31] developed and validated the PAM-13 questionnaire as a generic questionnaire for people with and without chronic conditions. The PAM-13 questionnaire has been validated in Norwegian and has previously been used to evaluate patient education in various diagnoses, including type 2 diabetes [25,32,33]. It has been used in multiple contexts across several countries, with findings suggesting that Norwegians generally exhibit higher activation than other populations [33]. Furthermore, the PAM is also suitable for the evaluation of digital interventions [34]. The reliability of the PAM-13 scale in this study was confirmed with a Cronbach α of 0.894, indicating high internal consistency among the items.

Diabetes Distress

Diabetes distress was assessed using the disease-specific questionnaire 20-item Problem Area in Diabetes (PAID-20) [26] (Table 3). PAID-20 has been previously translated and validated in Norwegian by Graue et al [35] and has been used as a tool to prepare for consultations for patients with type 1 diabetes in a previous study [36]. The reliability of the PAID-20 scale in this study was confirmed with a Cronbach α of 0.939, indicating excellent internal consistency among the items.

Well-Being

Psychological well-being and quality of life were measured with the generic tool World Health Organization–Five Well-Being Index (WHO-5) [37] (Table 3). WHO-5 has been translated to Norwegian and is validated for the type 1 diabetes population [27,38,39]. A score <50 may indicate that the patient is feeling unwell and should be examined further to investigate mild to moderate depression [40]. The reliability of the WHO-5

XSL•FO
RenderX

in this study was confirmed with a Cronbach α of 0.826, indicating high internal consistency among the items.

Health Literacy

Health literacy and the ability to make informed health choices were measured using the 12-item short version of the European Health Literacy Survey Questionnaire (HLS19-Q12) [41]. HLS_{19}-Q12 is translated into many languages and has been used in 19 European countries, including Norway [42] (Table 3). The HSL_{19}-Q12 has been validated and used in both a large general population and in patients with type 2 diabetes in Norway [41,43,44]. The reliability of the HLS_{19}-Q12 scale in this study was confirmed with a Cronbach α of 0.871, indicating high internal consistency among the items.

Ethical Considerations

As this research project involves human subjects, the project has been reported to the Norwegian Agency for Shared Services in Education and Research (SIKT; reference number: 456954) and approved by the local data protection officer at Akershus University Hospital (reference: 2022_125). The project was presented to the Regional Committees for Medical and Health Research Ethics (REK; reference number: 407539), and they deemed it to be outside their scope and defined as a quality improvement project.

All participants received oral and written information before providing informed consent to participate in this study and filling out the self-reported data. Digital consent forms, medical records data, and self-reported data were stored and managed in TSD. Participants provided their social security numbers for in the consent form but were assigned ID numbers for deidentification. The code key was stored in a separate file within TSD. Data received from NOKLUS for comparison with the overall hospital sample and the national population were aggregated and thus anonymous. The study participants were not compensated for participating in this study.

Statistical Analysis

Categorical variables are described with counts and percentages; continuous variables are presented with median values and ranges. To compare DigiDiaS care and usual care, we used nonparametric chi-square tests for pairs of categorical variables and Mann-Whitney *U* tests for pairs of continuous data. Missing data were not imputed, and the counts and proportions of missing values for each variable are presented in tables. Responses such as “don’t know” or “not applicable” in self-reported questionnaires were treated as missing values. To assess internal consistency of the scales used in the study, Cronbach α was calculated for the validated questionnaires [45].

To assess the strength of the possible association between the selected patient- and disease-related variables and the outcome, univariate logistic regression was conducted. Variables that were statistically significant in univariate analyses were entered into a multiple regression model, and the final model was fitted using the backward conditional selection process. The results are expressed as odds ratios (ORs) with 95% CIs. All tests were 2-sided, and *P* values <.05 were considered statistically significant.

In addition, to compare our sample with the overall hospital sample comprising all patients with type 1 diabetes and the national population of people with type 1 diabetes, we calculated CIs for the median values in our study sample to compare with the median and min-max values of the hospital sample and national population.

All data were processed using Excel (Microsoft Office Professional Plus 2016) and SPSS (version 29.0.0.0).

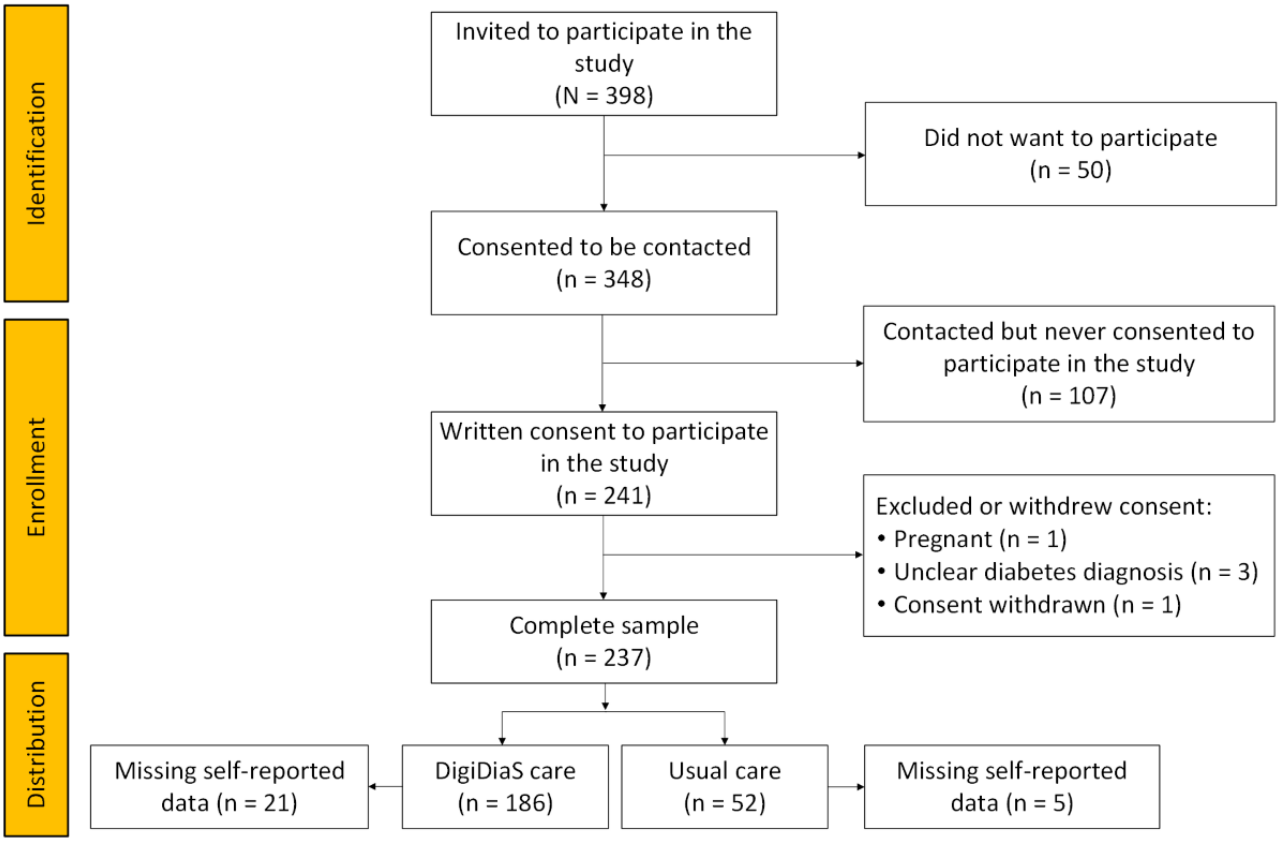
Results

Overview

Of the 398 patients invited to participate in the study, 241 consented to participate (Figure 1). Of the 237 patients included, 185 patients opted for DigiDiaS care, and 52 patients opted to continue with usual care. In the DigiDiaS care groups, 21 self-reported data were missing and 5 from the usual care group.

XSL•FO
RenderX

Figure 1. Flowchart of the participants using DigiDiaS care and usual care.



Demographics

In the complete sample, the included participants ranged in age from 19 to 81 years, with a median age of 49 years (Table 4). The median duration of diabetes was 21 (range 0-58; Table 5) years. Overall, one-third of the participants had university or college education at the bachelor level or higher (81/237, 34%), and the majority were either currently employed or students (139/237, 58.6%; Table 4). Just under half of the participants had an estimated travel time to the hospital of >30 minutes (98/237, 41.2%). Over half of all participants used other health-related apps daily (137/237, 57.8%).

XSL•FO
RenderX

Table 4. Demographic variables.

Demographic variables	Total (N=237)	DigiDiaS care (n=185)	Usual care (n=52)	*P* value
Age (y), median (range)	49 (19-81)	48 (19-79)	59 (23-81)	<.001
Gender, n (%)				
Women	108 (45.8)	91 (49.2)	17 (32.7)	.04
Education, n (%)				.41
Education up to upper secondary school or vocational school (13 y)	87 (36.7)	70 (37.8)	17 (32.7)	
University or college education up to 4 y	49 (20.7)	41 (22.2)	8 (15.4)	
University or college education >4 y	32 (13.5)	23 (12.4)	9 (17.3)	
Missing	69 (29.1)	51 (27.6)	18 (34.6)	
Employment, n (%)				.01
Yes	139 (58.6)	114 (61.6)	25 (48)	
No	44 (18.5)	28 (15.1)	16 (30.7)	
Retired[a]	28 (11.8)	19 (10.2)	8 (15.3)	
Missing	26 (10.9)	21 (11.3)	5 (9.6)	
Ethnic origin, n (%)				—[b]
European	216 (91.1)	166 (89.7)	50 (96.2)	
Asian or African	8 (3.3)	8 (4.3)	0 (0)	
Missing	12 (5.5)	11 (5.9)	2 (3.8)	
Cohabitation, n (%)				.45
Yes	148 (62.4)	117 (63.2)	31 (59.6)	
Missing	22 (9.3)	18 (9.7)	4 (7.7)	
Single resident, n (%)				
Yes	42 (17.6)	29 (15.7)	13 (25)	.12
Travel time to hospital by car, median (range)	27.5 (3-73)	27.5 (3-75)	26.5 (3-60)	.75
Travel time >30 min, n (%)	98 (41.2)	77 (41.4)	21 (40.4)	.87
Living outside of coverage area[c], n (%)	28 (11.8)	22 (11.8)	6 (11.5)	—
Use of other health apps, n (%)				.27
Daily or weekly use	137 (57.8)	111 (67.6)	26 (55.3)	
Monthly use	50 (21.1)	39 (23.7)	11 (23.4)	
Never or previous use	24 (11.3)	14 (8.5)	10 (21.2)	

[a]Not included in analysis due to filling the test assumption.

[b]Not applicable; cannot be computed due to small numbers.

[c]Do not have a postal code in the area where Akershus University Hospital is the local hospital.

XSL•FO
RenderX

Table 5. Diabetes treatment, complications, and clinical variables.

	Total (N=237)	DigiDiaS care (n=185)	Usual care (n=52)	*P* value
Diabetes duration (y), median (range)	21 (0-58)	19 (0-51)	29 (3-58)	<.001
HbA_{1c} mmol/mol), median (range)	58 (35-117)	59 (35-117)	52.5 (41-90)	.004
Missing, n (%)	2 (0.8)	1 (0.5)	1 (1.9)	—[a]
HbA_{1c} >75 mmol/mol, n (%)	36 (15.2)	32 (17.3)	4 (7.7)	—
Insulin delivery (pump), n (%)	86 (36.3)	74 (40)	13 (23.1)	.02
Missing	1 (0.4)	1 (0.5)	0 (0)	—
Blood glucose monitoring (continuous glucose monitoring), n (%)	225 (94.9)	178 (96.2)	47 (90.4)	.05
Missing	1 (0.4)	1 (0.5)	0 (0)	—
Time-in-range in percentage[b], median (range)	62 (0-99)	61 (0-98)	68 (7-99)	.21
Diabetic ketoacidosis, n (%)				
≥1[c]	39 (16.5)	37 (20)	2 (3.8)	—
Missing	41 (17.3)	34 (18.4)	7 (13.5)	—
Hypoglycemia requiring assistance, n (%)				
≥1[c]	82 (35)	65 (35.1)	18 (34.6)	.91
Missing	33 (13.9)	26 (14)	7 (13.5)	—
Late complications from diabetes, n (%)				.76
None[d]	136 (57.4)	107 (57.8)	29 (55.8)	
1[d]	63 (26.5)	51 (27.4)	12 (23.1)	
≥2[d]	38 (16)	27 (15.5)	11 (21.1)	
Missing	18 (7.6)	13 (7)	5 (9.6)	
BMI (kg/m^2), median (range)	26.5 (16.9-43.9)	26.7 (16.9-43.9)	26.4 (19.4-42.7)	.71
Missing, n (%)	4 (1.7)	3 (1.6)	1 (1.9)	—
Low-density lipoprotein cholesterol (mmol/L), median (range)	2.4 (0.9-5.6)	2.4 (0.9-5.6)	2.1 (0.9-3.4)	.71
Missing, n (%)	2 (0.8)	2 (1.1)	0 (0)	—
Blood pressure, median (range)				
Systolic blood pressure (mm Hg)	132.5 (94-178)	132 (94-178)	132 (107-166)	.28
Diastolic blood pressure (mm Hg)	79 (55-106)	80 (55-103)	77 (63-106)	.12
Missing, n (%)	18 (7.6)	13 (7)	5 (9.6)	—
Comorbidities, n (%)				.62
None	97 (40.9)	78 (42.2)	19 (36.5)	
1	77 (32.5)	59 (31.8)	18 (34.6)	
≥2	37 (19.7)	27 (14.5)	10 (19.2)	
Missing	26 (11)	21 (11.4)	5 (9.6)	—

[a] The Norwegian retirement age is 67 years.

[b]For continuous glucose monitoring users only.

[c]One or more times.

[d]Merged into categories for analysis.

XSL•FO
RenderX

Clinical Variables

The median HbA_{1c} for the complete sample was 58 mmol/mol (range 35-117 mmol/mol; Table 5). Furthermore, 15.5% (36/237) had an HbA_{1c} >75 mmol/mol, considered poorly regulated. Over one-third of the participants used an insulin pump (87/237, 36.7%), while the remaining participants used an insulin pen (150/237, 63.2%) for insulin delivery. Over half of the participants had no late complications of diabetes (136/237, 57.4%), and nearly half reported no comorbidities (97/237, 40.8%).

Self-Management, Diabetes Distress, Well-Being, and Health Literacy

The median self-management score of 72.5 (range 9-100) corresponds with a high self-management level (Table 6). This high median score corresponds with the highest PAM level: "Maintaining behaviors and pushing forward." Regarding the distribution across the 4 levels, half of the participants were at level 4 (116/237, 54.9%), and one-third were at level 3 for "taking action" (68/237, 32.2%).

Table 6. Self-reported outcomes for self-management, problem areas in diabetes, well-being, and health literacy.

	Total (n=211)	DigiDiaS care (n=164)	Usual care (n=47)	*P* value
Self-management (PAM-13[a]) score, median (range)	72.5 (9-100)	72.5 (9-100)	75 (27-100)	.33
Level 1: disengaged and overwhelmed, n (%)	14 (5.9)	11 (5.9)	3 (5.8)	
Level 2: becoming aware, but still struggling, n (%)	13 (5.5)	8 (4.3)	5 (9.6)	
Level 3: taking action, n (%)	68 (28.6)	59 (31.7)	9 (17.3)	
Level 4: maintaining behaviors and pushing forward, n (%)	116 (48.7)	86 (46.2)	30 (57.7)	.33[e]
Diabetes distress (PAID [b]) score, median (range)	23.75 (0-71.25)	23.75 (0-71.25)	21.25 (0-50)	.25
>40, n (%)	42 (17.6)	35 (18.8)	7 (13.5)	.33
Well-being (WHO-5[c]) score, median (range)	60 (4-100)	60 (4-96)	68 (16-100)	.01
<50, n (%)	46 (21.8)	40 (21.5)	6 (11.5)	.09
Health literacy (HLS19-Q12[d]) score, median (range)	35 (13-48)	35 (13-48)	34 (19-45)	.85
Below level 1, n (%)	27 (12.8)	23 (14)	4 (8.5)	
Level 1: access, understand, and apply, n (%)	52 (24.6)	38 (23.2)	14 (29.8)	
Level 2: appraise, understand, and apply, n (%)	79 (37.4)	62 (37.8)	17 (36.2)	
Level 3: making informed choices by critically evaluating, n (%)	53 (25.1)	41 (25)	12 (25.5)	.67[e]
Health literacy (HLS19-Q12) domains, median (range)				.63
Health care	12 (2-16)	12 (2-16)	12 (5-15)	
Disease prevention	11 (3-16)	11 (3-16)	11 (4-15)	
Health promotion	12 (4-16)	12 (4-16)	12 (5-16)	

[a]PAM-13: Patient Activation Measure; the score ranges from 0 to 100, with higher scores representing higher activation. Level 1: ≤47; level 2: 47.1 to 55.1; level 3: 55.2 to 67; level 4: ≥67.1.

[b]PAID: 20-item Problem Area in Diabetes; the score ranges from 0 to 100, with a higher score reflecting greater emotional distress. A score of ≥40 indicates severe emotional distress.

[c]WHO-5: World Health Organization-Five Well-Being Index; the score ranges from 0 to 100 with a higher score indicating higher well-being. A score <50 may be an identification of mild to severe depressive affects.

[d]HLS19-Q12: 12-item short version of the European Health Literacy Survey Questionnaire. Higher scores reflect higher health literacy. Below level 1: <27; level 1: ≥27; level 2: ≥33; and level 3: ≥39 (with associated characteristics). Domains can have a minimum score of 0 and a maximum score of 16.

[e]The *P* value represents the differences across all levels combined, not between individual levels. The level of diabetes distress was low, with a median score of 23 (range 0-71.25) of a possible maximum score of 100. However, almost 1 in 5 (42/237, 17.6%) of the participants reported scores >40, indicating diabetes distress.

Regarding well-being, the participants had a median score of 60 (range 4-100). However, 21.8% (46/237) of the participants had a WHO-5 score <50, indicating symptoms of mild to severe depression.

Overall, the participants had a fairly good health literacy score with a median health literacy score of 35 (range 13-48), characterized by the ability to appraise, understand, and apply health information (level 2).

Study Sample, Hospital Sample, and National Population

When comparing the study participants with the hospital sample of patients with type 1 diabetes at Akershus University Hospital, which includes all individuals with type 1 diabetes followed up

XSL•FO
RenderX

at Akershus University Hospital, and the national population of patients with type 1 diabetes, which includes all individuals with type 1 diabetes in Norway, the 3 samples were comparable regarding gender, education, insulin administration, HbA_{1c}, age, and diabetes duration (Table 7). However, regarding HbA_{1c} ≥75 mmol/mol, the proportion of such individuals was significantly lower in the national population compared with the study sample, indicating that our sample comprised individuals with less well-regulated diabetes.

Table 7. Comparison of study sample, hospital sample, and national population on sociodemographic and clinical variables.

Variables	Study sample (N=237)	Hospital sample (n=1643)	National population in Norway (n=22,408)
Age (y; 95% CI 46-53), median (range)	49 (19-81)	46 (18-98)	47 (18-98)
Missing, n (%)	0 (0)	0 (0)	0 (0)
Gender, n (%)			
Women (95% CI 39.1-52.1)	108 (45.8)	698 (42.5)	9877 (44.1)
Men (95% CI 47.4-60.4)	128 (54.8)	945 (57.5)	12,531 (55.9)
Other or missing	0 (0)	0 (0)	0 (0)
Education, n (%)			
Education up to upper secondary school or vocational school (13 y; 95% CI 30.5-43.1)	87 (36.7)	625 (38)	8054 (35.9)
University or college education up to 4 y (95% CI 15.7-26.3)	49 (20.7)	268 (16.3)	4089 (18.2)
University or college education >4 y (95% CI 9.4-18.5)	32 (13.5)	168 (10.2)	2487 (11.1)
Missing	69 (29.1)	582 (35.4)	7778 (34.7)
Years with diabetes (95% CI 18-23), median (range)	21 (0-58)	18 (0-78)	20 (0-78)
Missing, n (%)	0 (0)	9 (0.5)	107 (0.5)
HbA_{1c}[a]			
HbA_{1c} (≥75 mmol/mol; 95% CI 10.8-20.4), n (%)	36 (15.2)	190 (12.3)	2076 (9.5)
HbA_{1c} (mmol/mol; 95% CI 55-59), median (range)	58 (35-117)	58 (21-159)	56 (21-159)
Missing, n (%)	2 (0.8)	93 (5.6)	667 (3)
Insulin delivery, n (%)			
Pump (95% CI 30.1-42.7)	86 (36.3)	520 (31.6)	8911 (39.8)
Pen (95% CI 57.2-69.8)	151 (63.7)	1103 (67.1)	12,960 (57.8)
Missing	1 (0.4)	20 (1.2)	537 (2.4)

[a]HbA_{1c}: glycated hemoglobin.

Comparing Patient Characteristics in DigiDias Care Versus Usual Care

Demographic Variables

Most participants opted for DigiDiaS care (185/237, 78.1%), while the remaining participants (52/237, 21.9%) opted to continue with usual care. Among those choosing DigiDiaS care, half had been included in DigiDiaS care more than 90 days before consenting to participate in the study (92/185, 49.7%), thereby having had access to the digital outpatient care solution from before.

The participants in the DigiDiaS care group were significantly younger, with a median age of 47 (range 19-79) years, compared with those choosing usual care (median age 59, range 23-81 years; *P*<.001; Table 4). DigiDiaS care patients have had diabetes for a median of 10 years shorter than the usual care group (*P*<.001; Table 5). A statistically significant linear relationship was found between age and duration of diabetes in the complete sample (*P*<.001).

There was no statistically significant difference in the education level between DigiDiaS care and usual care (*P*=.42). However, there was a statistically significant higher proportion of individuals in the usual care group (16/52, 31%) who were outside the labor market or received disability benefits of working age than with DigiDiaS care (28/185, 15.1%; *P*=.01). A higher proportion lived alone and were single residents in usual care (13/52, 25%) than in DigiDiaS care (29/185, 15.7%). However, the difference was not statistically significant (*P*=.12). Further demographic variables are summarized in Table 4.

Concerning the use of other health-related mobile apps, there were more individuals in the usual care group who had never or only previously used other health-related apps (10/52, 21%)

XSL•FO
RenderX

compared with the participants in the DigiDiaS care group (14/185, 8.5%); however, the difference was not statistically significant (*P*=.27; Table 4).

Clinical Variables

In the DigiDiaS care group, the participants had a median HbA_{1c} significantly higher at 59 (range 35-117) mmol/mol compared with 52 (range 41-90) mmol/mol in the usual care group (*P*=.004; Table 5). In terms of insulin delivery, there was a significantly higher use of an insulin pump in the DigiDiaS care group (74/185, 40%) than in the usual care group (23/52, 23%; *P*=.02).

There were no statistically significant differences in late complications from diabetes (*P*=.76). Concerning comorbidities, there were no differences between the DigiDiaS care group and usual care (*P*=.63). Further details on between-group differences in diabetes clinical variables are presented in Table 5.

Self-Management, Diabetes Distress, Well-Being, and Health Literacy Across the Two Groups

There were no differences between the DigiDiaS participants compared with usual care in self-management (*P*=.33), diabetes distress (*P*=.25), or the health literacy scale (*P*=.85; Table 6).

Well-Being

The usual care group (median 68, range 16-100) had a statistically significant higher well-being score than the DigiDiaS care group (*P*=.01) indicating better well-being, with the usual care group having a higher well-being score than DigiDiaS care (median 60, range 4-96; Table 6).

Furthermore, we assessed the correlation between well-being scores and HbA_{1c} levels using Spearman correlation, finding no statistically significant relationship in the study sample, within the DigiDiaS care group (*P*=.37), or the usual care group (*P*=.33).

If the median value or percentage for the hospital selection and national population falls within the CI of the study sample, there is no significant difference. However, if the median value or percentage falls outside the CI, the groups may be significantly different.

Group Comparison

Univariate Analyses

In the univariate logistic regression analyses, we identified statistically significant differences between those who chose DigiDiaS care and usual care regarding the following variables: gender, employment, diabetes duration, HbA_{1c}, insulin delivery method, and well-being, as seen in Table 8. However, age was not associated with the choice of DigiDiaS or usual care.

Table 8. Logistic regression univariate and multivariate of the significant outcomes from Tables 2-4.

Variables	Unadjusted odds ratio (95% CI)	*P* value	Adjusted odds ratio (95% CI)	*P* value
Age (y)	0.99 (0.98-1.01)	.94	—[a]	—
Gender: man or woman (ref)	1.93 (1.01-3.70)	.04	1.10 (0.52-2.36)	.97
Employment: yes (ref) and no	2.60 (1.22-5.52)	.01	2.1 (0.94-5.07)	.68
Diabetes duration (y)	0.96 (0.93-0.98)	<.001	0.95 (0.93-0.97)	<.001
HbA_{1c}	1.03 (1.00-1.06)	.01	1.02 (0.99-1.05)	.07
Insulin delivery: pen and pump (ref)	0.32 (1.09-4.52)	.02	3.16 (1.37-7.32)	.007
Well-being (WHO-5[b] score)	0.97 (0.95-0.99)	.01	0.97 (0.95-0.99)	.03

[a]Not included in the multivariate analysis because it was not statistically significant in the univariate analysis.

[b]WHO-5: World Health Organization-Five Well-Being Index.

Multivariate Analysis

In multiple analyses with variables that were statistically significant in univariate analyses (Table 8), we found that a longer duration of diabetes was associated with statistically significant lower odds of opting for DigiDiaS care (OR 0.95, 95% CI 0.93-0.97). Similarly, a higher well-being score was associated with lower odds of choosing DigiDiaS care (OR 0.97, 95% CI 0.95-099). Participants using an insulin pump for insulin delivery were found to have 3 times higher odds of choosing DigiDiaS care (OR 3.16, 95% CI 1.37-7.32) in comparison to insulin pump users. Furthermore, the insulin delivery method was the variable with the strongest association with choosing DigiDiaS care versus usual care.

The multivariate logistic regression is divided into 2 rounds to avoid overfitting of the model, marked with the hard line in between—round 1 consists gender, employment in working age, and diabetes duration and round 2 consists HbA_{1c}, insulin delivery, and well-being.

Discussion

Principal Findings

This cross-sectional study was designed to explore the characteristics of patients with type 1 diabetes who were given a choice between a novel, flexible digital mHealth supplement for outpatient care, referred to as DigiDiaS care, or continuing with their usual care.

XSL•FO
RenderX

Our main findings reveal, as expected, that the vast majority of the 238 participants included in the study opted for DigiDiaS care rather than usual care. When comparing the two groups, the DigiDiaS care group had a shorter duration of diabetes, more use of insulin pumps than insulin pens for insulin delivery, and a lower well-being score compared with the usual care group. Those who opted for DigiDiaS care did not differ from those who continued with their usual care in terms of sociodemographic variables, presence of late complications or comorbidities from diabetes, self-management, diabetes distress, or health literacy. Furthermore, our study sample did not differ from the outpatient sample or the national population on relevant variables.

Comparison With Prior Evidence

Not surprisingly, most participants opted for DigiDiaS care, aligning well with their profile of being both younger and already users of technology for insulin administration. Using an insulin pump requires certain technical skills, which are also necessary to use the DigiDiaS care app. Using an insulin pump instead of an insulin pen was the strongest characteristic associated with the choice of DigiDiaS care over usual care. This aligns well with previous research findings on how familiarity with diabetes-related technology might drive the adoption of other digital health tools [46,47]. These findings indicate that patients already accustomed to using advanced diabetes management technologies are more likely to embrace new digital health solutions, such as DigiDiaS care. In general, patients in the usual care group had longer diabetes duration compared with those in the DigiDiaS care group, and they may have established a routine for their outpatient care, making them less inclined to try out something new. Thus, when introducing digital care solutions in diabetes management, those already using diabetes technology equipment may be more prone to exploring the possibilities within new ways of receiving outpatient care.

In our sample, those opting for DigiDiaS care had lower well-being scores than those in the usual care group. Living with a chronic condition requiring ongoing self-management, such as type 1 diabetes, can affect well-being, even though well-being can vary throughout life and with age [4]. This variation in well-being is also observed in the Norwegian general population, where well-being appears to increase with age from the early 40s to the mid-70s [48]. This might correspond with how our participants opting for DigiDiaS care had a median age of 48 years—that is, 11 years younger than those opting for usual care and reporting a higher well-being score. Compared with the Norwegian population of adults with type 1 diabetes, both groups in our sample had a median well-being score higher than the national average [49]. Thus, despite our DigiDiaS group presenting a lower well-being score compared with those continuing with usual care, all participants had a fair level of well-being without symptoms of depression interfering with their abilities to self-manage.

Although most of the recruited participants opted for DigiDiaS care, a small proportion of the participants chose to continue with usual care, thus declining the opportunity to explore digital follow-up. In our study, we did not investigate actual use but rather the willingness to use DigiDiaS care. To the best of our knowledge, DigiDiaS care is the first outpatient care model comprising specific components, such as a message service, PRO-based questionnaires, an information page, and video consultations for diabetes care [15]. The closest alternative to our knowledge is DiabetesFlex, and our study is a response to the call by Jensen et al [17] after interviewing patients with type 1 diabetes who received flexible PRO-based telehealth as part of their diabetes follow-up in Denmark, questioning whether this type of solution is acceptable for all patients. Our findings show that not all patients opt for digital outpatient care. A previous scoping review on reasons for nonuse of digital PRO identified factors, such as ability to use the PRO, engagement, emotional distress, and technical barriers [50]. Our identified differences in the use of insulin delivery equipment may represent differences in the technical skills required to use an insulin pump. Furthermore, the scoping review highlighted how participants with few symptoms or those in good health stopped using digital PRO solutions more frequently [50]. This aligns well with the differences in well-being among our groups, where those with higher well-being chose usual care.

In contrast to previous research highlighting the importance of socioeconomic background when offering digital health services, differences in such characteristics were not observed in our sample. Previous research highlights that health literacy and socioeconomically disadvantaged backgrounds may be less likely to be considered for digital PRO-based outpatient care [51]. Our findings show that there are no statistical differences in health literacy or socioeconomic status between the DigiDiaS care and usual care groups, suggesting that there is no reason not to consider such patients to use these solutions. However, this may be contextual, as the Norwegian population in general has a high socioeconomic status. Yet, further research can provide a deeper understanding of the specific characteristics that influence patients' decisions to choose digital outpatient care.

Although we examined various characteristics from medical records and self-reported questionnaires, we did not investigate all potential influencing factors. For instance, while we assessed general health literacy to understand the ability to find, understand, evaluate, and apply health information, digital health literacy could provide deeper insights into patients' ability to effectively use such solutions [52]. In addition, efforts should be made to explore whether the use of digital outpatient care solutions such as DigiDiaS care varies with factors, such as geography, travel time, and the availability of health care services. The introduction of new technology in health care services is a complex process [53], and patient acceptance of new solutions is influenced by how health care professionals adopt and support these technologies [54]. Further understanding and investigation of these dynamics may reveal factors for successful implementation in practice.

Strengths and Limitations

This study presents strengths and limitations that are important to consider when interpreting the findings. First, this study explores a fully implemented digital outpatient care service [21], hindering the use of established gold standard research

XSL•FO
RenderX

designs. However, our nonrandomized study design was essential in capturing the participants' preferences, allowing them to choose either DigiDiaS care or continuing with usual care. This approach enabled us to investigate characteristics based on the participants' choice of follow-up, which would not have been possible with other study designs, such as randomization or a stepped wedge design. The resulting skewed distribution was, therefore, expected and reflects these treatment choices. To compensate for the skewed distribution, nonparametric tests were used; nevertheless, the results should be interpreted with caution due to the uneven distribution.

While a none-randomized design with consecutive sampling may introduce some bias, it provides valuable insights into real-world decision-making and patient preferences that are crucial for understanding the adoption and effectiveness of the DigiDiaS care model [21], especially considering the needs and choices of the minority group who continued with usual care. Acknowledging that some participants do not wish to use this type of outpatient care is important, and the reasons for this should be further investigated. Qualitative interviews may reveal factors that we may not have detected in this study.

Second, this study does not examine the usability of individual components in the patient app or the specific solution of the app, but rather its clinical implications. Our investigation does not exhaustively cover all factors that may influence the choice of digital outpatient care given the complexity of such decisions. However, by collecting a broad range of data from medical records and self-reported data from the PRO questionnaire, we have gathered a comprehensive dataset. Our cross-sectional study, incorporating both generic and disease-specific well-validated measures, allowed us to add depth to the medical records data [55]. The richness of PRO data lies in the patient perspective it provides, reflecting their experiences, symptoms, and the impact of the disease on their daily lives [56].

Third, the digital outpatient care solution was introduced to half of the participants in the DigiDiaS care group before they were included in this study. The extent to which this was used is unknown, as its use is mainly related to the patients' need to contact the clinic outside of fixed appointments, preparation for consultation or information, and e-learning courses received in connection with consultations. The time before we started recruiting for this study was also an introductory period in the use of DigiDiaS care for the clinic. We do not see it as a limitation that participants had access to the app before the start of the study, as we investigated opting for use of DigiDiaS care or not in this paper. However, as previously addressed, experiences from the Danish DiabetesFlex indicated that not everyone wanted to continue with DiabetesFlex after the study period [17], and even though our participants were opting for DigiDiaS care, actual use should be further investigated.

Conclusions

This study reveals that patients with type 1 diabetes are opting to use a digital outpatient care solution when it is offered. Most of participants opted for DigiDiaS care, the novel mHealth solution, while a smaller proportion opted to continue with usual care. Those opting for DigiDiaS care are characterized by being already familiar with using diabetes-related technology, having a shorter diabetes duration and lower well-being than the usual care group.

Importantly, our findings indicate that factors such as demographic and clinical variables, as well as self-management, health literacy, and diabetes distress did not differ significantly between DigiDiaS care and usual care. Even though only a small proportion continued with usual care, the characteristics of this group should be further explored to understand how these individuals can also be included in digital outpatient care solutions. Further studies should also focus on how the implementation of digital solutions in outpatient care can affect how and if patients use them.

Acknowledgments

The authors would like to express their heartfelt thanks to all the participants in the study, the diabetes nurses, and endocrinologists at Akershus University Hospital for their assistance in recruitment. They are also grateful to the NOKLUS Diabetes Registry for its contribution to the data. Finally, they would like to thank their colleague Maria Aadland Mollestad for patience in controlling the data.

The paper does not contain text generated by artificial intelligence (AI). However, AI was used to assist with proofreading, alongside a professional proofreader.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

IS and AT conceptualized the research questions. TS and JAW facilitated recruitment of participants. IS recruited participants and collected data. IS, AT, and MH managed the data analysis. IS drafted the manuscript with support from HH and AT. ALJ, MH, TS, JAW, and EÅ revised the manuscript. All authors contributed to the manuscript and approved the final version.

Conflicts of Interest

None declared.

XSL•FO
RenderX

References

1. Speight J, Holmes-Truscott E, Hendrieckx C, Skovlund S, Cooke D. Assessing the impact of diabetes on quality of life: what have the past 25 years taught us? Diabet Med 2020 Mar;37(3):483-492. [doi: 10.1111/dme.14196] [Medline: 31797443]
2. Young-Hyman D, de Groot M, Hill-Briggs F, Gonzalez JS, Hood K, Peyrot M. Psychosocial care for people with diabetes: a position statement of the American diabetes association. Diabetes Care 2016 Dec;39(12):2126-2140 [FREE Full text] [doi: 10.2337/dc16-2053] [Medline: 27879358]
3. Umpierrez G, Korytkowski M. Diabetic emergencies - ketoacidosis, hyperglycaemic hyperosmolar state and hypoglycaemia. Nat Rev Endocrinol 2016 Apr;12(4):222-232. [doi: 10.1038/nrendo.2016.15] [Medline: 26893262]
4. van Duinkerken E, Snoek FJ, de Wit M. The cognitive and psychological effects of living with type 1 diabetes: a narrative review. Diabet Med 2020 Apr;37(4):555-563 [FREE Full text] [doi: 10.1111/dme.14216] [Medline: 31850538]
5. Weissberg-Benchell J, Shapiro JB, Bryant FB, Hood KK. Supporting Teen Problem-Solving (STEPS) 3 year outcomes: preventing diabetes-specific emotional distress and depressive symptoms in adolescents with type 1 diabetes. J Consult Clin Psychol 2020 Nov;88(11):1019-1031 [FREE Full text] [doi: 10.1037/ccp0000608] [Medline: 33136423]
6. Diabetes: nasjonal faglig retningslinje. Helsedirektoratet.: Helsedirektoratet; 2016 Sep 14. URL: https://www.helsedirektoratet.no/retningslinjer/diabetes [accessed 2025-01-05]
7. American Diabetes Association Professional Practice Committee. 1. Improving care and promoting health in populations: standards of care in diabetes-2024. Diabetes Care 2024 Jan 01;47(Suppl 1):S11-S19. [doi: 10.2337/dc24-S001] [Medline: 38078573]
8. Phillip M, Bergenstal RM, Close KL, Danne T, Garg SK, Heinemann L, et al. The digital/virtual diabetes clinic: the future is now-recommendations from an international panel on diabetes digital technologies introduction. Diabetes Technol Ther 2021 Feb;23(2):146-154 [FREE Full text] [doi: 10.1089/dia.2020.0375] [Medline: 32905711]
9. Bour C, Ahne A, Aguayo G, Fischer A, Marcic D, Kayser P, et al. Global diabetes burden: analysis of regional differences to improve diabetes care. BMJ Open Diabetes Res Care 2022 Oct;10(5):e003040 [FREE Full text] [doi: 10.1136/bmjdrc-2022-003040] [Medline: 36307139]
10. Institute of Medicine, Committee on Quality of Health Care in America. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: National Academies Press; 2001.
11. Powell PW, Corathers SD, Raymond J, Streisand R. New approaches to providing individualized diabetes care in the 21st century. Curr Diabetes Rev 2015;11(4):222-230 [FREE Full text] [doi: 10.2174/1573399811666150421110316] [Medline: 25901504]
12. American Diabetes Association Professional Practice Committee. 7. Diabetes technology: standards of care in diabetes-2024. Diabetes Care 2024 Jan 01;47(Suppl 1):S126-S144. [doi: 10.2337/dc24-S007] [Medline: 38078575]
13. Holmen H, Singstad T, Ribu L, Jensen AL, Weldingh NM, Torbjørnsen A. Adapting a patient-reported outcome measure to digital outpatient specialist health care services for type 1 diabetes: user involvement study. JMIR Hum Factors 2022 Nov 15;9(4):e38678 [FREE Full text] [doi: 10.2196/38678] [Medline: 36378513]
14. Greenhalgh J, Gooding K, Gibbons E, Dalkin S, Wright J, Valderas J, et al. How do patient reported outcome measures (PROMs) support clinician-patient communication and patient care? A realist synthesis. J Patient Rep Outcomes 2018 Sep 15;2:42 [FREE Full text] [doi: 10.1186/s41687-018-0061-6] [Medline: 30294712]
15. Torbjørnsen A, Jensen AL, Singstad T, Weldingh NM, Holmen H. Patient-reported outcome measures in diabetes outpatient care: a scoping review. BMJ Open Diabetes Res Care 2023 Nov;11(6):e003628 [FREE Full text] [doi: 10.1136/bmjdrc-2023-003628] [Medline: 37963648]
16. Laurberg T, Schougaard LM, Hjollund NH, Lomborg KE, Hansen TK, Jensen AL. Randomized controlled study to evaluate the impact of flexible patient-controlled visits in people with type 1 diabetes: the DiabetesFlex trial. Diabet Med 2022 May;39(5):e14791 [FREE Full text] [doi: 10.1111/dme.14791] [Medline: 35028992]
17. Jensen AL, Schougaard LM, Laurberg T, Hansen TK, Lomborg K. Flexible patient-reported outcome-based telehealth follow-up for type 1 diabetes: a qualitative study. Scand J Caring Sci 2023 Sep;37(3):662-676. [doi: 10.1111/scs.13154] [Medline: 36775917]
18. Liu X, Haagsma J, Sijbrands E, Buijks H, Boogaard L, Mackenbach JP, et al. Anxiety and depression in diabetes care: longitudinal associations with health-related quality of life. Sci Rep 2020 May 20;10(1):8307 [FREE Full text] [doi: 10.1038/s41598-020-57647-x] [Medline: 32433470]
19. Sánchez-Urbano C, Pino MJ, Herruzo C. Personality prototypes in people with type 1 diabetes and their relationship with adherence. Int J Environ Res Public Health 2021 Apr 30;18(9):4818 [FREE Full text] [doi: 10.3390/ijerph18094818] [Medline: 33946437]
20. Sińska BI, Kucharska A, Panczyk M, Matejko B, Traczyk I, Harton A, et al. Identification of sociodemographic and clinical factors influencing the feeling of stigmatization in people with type 1 diabetes. Healthcare (Basel) 2023 Aug 02;11(15):2185 [FREE Full text] [doi: 10.3390/healthcare11152185] [Medline: 37570425]
21. Torbjørnsen A, Spildo I, Mollestad MA, Jensen AL, Singstad T, Weldingh NM, et al. Investigating digital patient-reported outcome measures in patient-centered diabetes specialist outpatient care (DigiDiaS): protocol for a multimethod prospective observational study. JMIR Res Protoc 2024 Mar 05;13:e52766 [FREE Full text] [doi: 10.2196/52766] [Medline: 38441955]

XSL•FO
RenderX

22. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol 2008 Apr;61(4):344-349. [doi: 10.1016/j.jclinepi.2007.11.008] [Medline: 18313558]
23. Instructions for use - MyDignio. Dignio. URL: https://hjelp.dignio.com/en/page/instructions-for-use-mydignio [accessed 2024-12-15]
24. Short introduction to Nettskjema. University of Oslo. URL: https://www.uio.no/english/services/it/adm-services/nettskjema/about-nettskjema.html [accessed 2024-11-10]
25. Steinsbekk A. Måling av effekt av pasientopplæring. J Norwegian Med Assoc 2008 Oct 23;128(20):2316-2324 [FREE Full text]
26. Polonsky WH, Anderson BJ, Lohrer PA, Welch G, Jacobson AM, Aponte JE, et al. Assessment of diabetes-related distress. Diabetes Care 1995 Jun;18(6):754-760. [doi: 10.2337/diacare.18.6.754] [Medline: 7555499]
27. de Wit M, Pouwer F, Gemke RJ, Delemarre-van de Waal HA, Snoek FJ. Validation of the WHO-5 Well-Being Index in adolescents with type 1 diabetes. Diabetes Care 2007 Aug;30(8):2003-2006. [doi: 10.2337/dc07-0447] [Medline: 17475940]
28. Newnham EA, Hooke GR, Page AC. Monitoring treatment response and outcomes using the World Health Organization's Wellbeing Index in psychiatric care. J Affect Disord 2010 Apr;122(1-2):133-138. [doi: 10.1016/j.jad.2009.06.005] [Medline: 19592116]
29. Guttersrud Ø, Le C, Pettersen KS, Helseth S, Finbråten HS. Towards a progression of health literacy skills: establishing the HLS-Q12 cutoff scores. Res Square 2019:1-20 [FREE Full text] [doi: 10.21203/rs.2.13456/v2]
30. Hibbard JH, Mahoney ER, Stockard J, Tusler M. Development and testing of a short form of the patient activation measure. Health Serv Res 2005 Dec;40(6 Pt 1):1918-1930 [FREE Full text] [doi: 10.1111/j.1475-6773.2005.00438.x] [Medline: 16336556]
31. Hibbard JH, Stockard J, Mahoney ER, Tusler M. Development of the Patient Activation Measure (PAM): conceptualizing and measuring activation in patients and consumers. Health Serv Res 2004 Aug;39(4 Pt 1):1005-1026 [FREE Full text] [doi: 10.1111/j.1475-6773.2004.00269.x] [Medline: 15230939]
32. Steinsbekk A. [Patient activation measure]. Tidsskr Nor Laegeforen 2008 Oct 23;128(20):2316-2318 [FREE Full text] [Medline: 19096487]
33. Rademakers J, Maindal HT, Steinsbekk A, Gensichen J, Brenk-Franz K, Hendriks M. Patient activation in Europe: an international comparison of psychometric properties and patients' scores on the short form Patient Activation Measure (PAM-13). BMC Health Serv Res 2016 Oct 12;16(1):570 [FREE Full text] [doi: 10.1186/s12913-016-1828-1] [Medline: 27729079]
34. Abdelraheem O, Salama M, Chun S. Impact of digital interventions and online health communities in patient activation: systematic review and meta-analysis. Int J Med Inform 2024 Aug;188:105481 [FREE Full text] [doi: 10.1016/j.ijmedinf.2024.105481] [Medline: 38776718]
35. Graue M, Haugstvedt A, Wentzel-Larsen T, Iversen MM, Karlsen B, Rokne B. Diabetes-related emotional distress in adults: reliability and validity of the Norwegian versions of the Problem Areas in Diabetes Scale (PAID) and the Diabetes Distress Scale (DDS). Int J Nurs Stud 2012 Feb;49(2):174-182. [doi: 10.1016/j.ijnurstu.2011.08.007] [Medline: 21924422]
36. Hernar I, Graue M, Richards D, Strandberg RB, Nilsen RM, Tell GS, et al. Electronic capturing of patient-reported outcome measures on a touchscreen computer in clinical diabetes practice (the DiaPROM trial): a feasibility study. Pilot Feasibility Stud 2019 Feb 20;5:29 [FREE Full text] [doi: 10.1186/s40814-019-0419-4] [Medline: 30820340]
37. Wellbeing measures in primary health care/the DepCare Project: report on a WHO meeting: Stockholm, Sweden, 12–13 February 1998. World Health Organization Regional Office for Europe. 1998 Feb 13. URL: https://iris.who.int/handle/10665/349766 [accessed 2025-06-27]
38. The World Health Organisation- Five Well-Being Index (WHO-5). Child Outcomes Research Consortium. URL: https://www.corc.uk.net/outcome-experience-measures/the-world-health-organisation-five-well-being-index-who-5/ [accessed 2024-11-12]
39. Kaiser S, Kyrrestad H. Måleegenskaper ved den norske versjonen av Verdens helseorganisasjon Well Being Index (WHO-5). PsykTestBarn 2019;1:4 [FREE Full text]
40. Christensen KS, Haugen W, Sirpal MK, Haavet OR. Diagnosis of depressed young people--criterion validity of WHO-5 and HSCL-6 in Denmark and Norway. Fam Pract 2015 Jun;32(3):359-363. [doi: 10.1093/fampra/cmv011] [Medline: 25800246]
41. Finbråten HS, Wilde-Larsson B, Nordström G, Pettersen KS, Trollvik A, Guttersrud Ø. Establishing the HLS-Q12 short version of the European Health Literacy Survey Questionnaire: latent trait analyses applying Rasch modelling and confirmatory factor analysis. BMC Health Serv Res 2018 Jun 28;18(1):506 [FREE Full text] [doi: 10.1186/s12913-018-3275-7] [Medline: 29954382]
42. Sørensen K, Pelikan JM, Röthlin F, Ganahl K, Slonska Z, Doyle G, et al. Health literacy in Europe: comparative results of the European health literacy survey (HLS-EU). Eur J Public Health 2015 Dec;25(6):1053-1058 [FREE Full text] [doi: 10.1093/eurpub/ckv043] [Medline: 25843827]
43. Finbråten HS, Pettersen KS, Wilde-Larsson B, Nordström G, Trollvik A, Guttersrud Ø. Validating the European Health Literacy Survey Questionnaire in people with type 2 diabetes: latent trait analyses applying multidimensional Rasch

XSL•FO
RenderX

modelling and confirmatory factor analysis. J Adv Nurs 2017 Nov;73(11):2730-2744. [doi: 10.1111/jan.13342] [Medline: 28543754]
44. Befolkningens helsekompetanse. Helsedirekotratet. 2021 Jun 22. URL: https://www.helsedirektoratet.no/rapporter/befolkningens-helsekompetanse [accessed 2022-11-18]
45. Pallant J. SPSS Survival Manual: A Step By Step Guide to Data Analysis Using SPSS Program Sixth Edition. London, UK: McGraw-Hill Education; 2016.
46. Toschi E, Adam A, Atakov-Castillo A, Slyne C, Laffel L, Munshi M. Use of telemedicine in adults with type 1 diabetes: do age and use of diabetes-related technology matter? Telemed J E Health 2023 Sep;29(9):1374-1382. [doi: 10.1089/tmj.2022.0397] [Medline: 36695656]
47. Alaslawi H, Berrou I, Al Hamid A, Alhuwail D, Aslanpour Z. Diabetes self-management apps: systematic review of adoption determinants and future research agenda. JMIR Diabetes 2022 Jul 28;7(3):e28153 [FREE Full text] [doi: 10.2196/28153] [Medline: 35900826]
48. Hansen T, Blekesaune M. The age and well-being "paradox": a longitudinal and multidimensional reconsideration. Eur J Ageing 2022 Dec 23;19(4):1277-1286 [FREE Full text] [doi: 10.1007/s10433-022-00709-y] [Medline: 36506681]
49. Løvaas KF, Madsen TV, Cooper J, Sandberg S, Ernes T, Ueland G. Data fra diabetespoliklinikker diabetes type 1 Årsrapport for 2023. NOKLUS. 2024 Jun. URL: https://www.noklus.no/media/ia5ot02z/type-1_-%C3%A5rsrapport-norsk-diabetesregister-for-voksne-2023.pdf [accessed 2025-06-27]
50. Nielsen AS, Kidholm K, Kayser L. Patients' reasons for non-use of digital patient-reported outcome concepts: a scoping review. Health Informatics J 2020 Dec;26(4):2811-2833 [FREE Full text] [doi: 10.1177/1460458220942649] [Medline: 32731773]
51. Aiyegbusi OL, Hughes SE, Peipert JD, Schougaard LM, Wilson R, Calvert MJ. Reducing the pressures of outpatient care: the potential role of patient-reported outcomes. J R Soc Med 2023 Feb;116(2):44-64 [FREE Full text] [doi: 10.1177/01410768231152222] [Medline: 36758615]
52. Campanozzi LL, Gibelli F, Bailo P, Nittari G, Sirignano A, Ricci G. The role of digital literacy in achieving health equity in the third millennium society: a literature review. Front Public Health 2023 Feb 20;11:1109323 [FREE Full text] [doi: 10.3389/fpubh.2023.1109323] [Medline: 36891330]
53. Kho J, Gillespie N, Martin-Khan M. A systematic scoping review of change management practices used for telemedicine service implementations. BMC Health Serv Res 2020 Sep 01;20(1):815 [FREE Full text] [doi: 10.1186/s12913-020-05657-w] [Medline: 32873295]
54. Longhini J, Rossettini G, Palese A. Digital health competencies among health care professionals: systematic review. J Med Internet Res 2022 Aug 18;24(8):e36414 [FREE Full text] [doi: 10.2196/36414] [Medline: 35980735]
55. Terwee CB, Elders PJ, Blom MT, Beulens JW, Rolandsson O, Rogge AA, et al. Patient-reported outcomes for people with diabetes: what and how to measure? A narrative review. Diabetologia 2023 Aug;66(8):1357-1377 [FREE Full text] [doi: 10.1007/s00125-023-05926-3] [Medline: 37222772]
56. Aiyegbusi OL, Roydhouse J, Rivera SC, Kamudoni P, Schache P, Wilson R, et al. Key considerations to reduce or address respondent burden in patient-reported outcome (PRO) data collection. Nat Commun 2022 Oct 12;13(1):6026 [FREE Full text] [doi: 10.1038/s41467-022-33826-4] [Medline: 36224187]

Abbreviations

HbA1c: glycated hemoglobin
mHealth: mobile health
OR: odds ratio
PAID-20: 20-item Problem Area in Diabetes
PAM-13: Patient Activation Measure 13
PRO: patient-reported outcome
STROBE: Strengthening the Reporting of Observational studies in Epidemiology
HLS19-Q12: 12-item short version of the European Health Literacy Survey Questionnaire
WHO-5: World Health Organization-Five Well-Being Index

XSL•FO
RenderX

Edited by A Kushniruk; submitted 10.03.25; peer-reviewed by I Milani, S do Vale; comments to author 13.05.25; revised version received 30.05.25; accepted 10.06.25; published 22.08.25.

Please cite as:
Spildo I, Holmen H, Lunde Jensen A, Hagen M, Singstad T, Winther JA, Årsand E, Torbjørnsen A
Examining Patient Characteristics Associated With Digital Outpatient Care for Type 1 Diabetes (DigiDiaS): Cross-Sectional Study
JMIR Hum Factors 2025;12:e73708
URL: https://humanfactors.jmir.org/2025/1/e73708
doi:10.2196/73708
PMID:

©Ingeborg Spildo, Heidi Holmen, Annesofie Lunde Jensen, Milada Hagen, Tone Singstad, Jacob Andreas Winther, Eirik Årsand, Astrid Torbjørnsen. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 22.08.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Gamified Physical-Digital Smoking Cessation Intervention for Young Adults: Mixed Methods Development and Usability Study

Natalia Bartłomiejczyk[1*], MSc; Alessio De Santo[2*], PhD; Adrian Holzer[1], PhD

[1]University of Neuchâtel, Neuchâtel, Switzerland

[2]HEG Arc, HES-SO // University of Applied Sciences and Arts Western Switzerland, Neuchâtel, Switzerland

[*]these authors contributed equally

Corresponding Author:
Alessio De Santo, PhD
HEG Arc
HES-SO // University of Applied Sciences and Arts Western Switzerland
Espace de l'Europe 21
Neuchâtel, 2000
Switzerland
Phone: 41 32 930 23 65
Email: alessio.desanto@he-arc.ch

Abstract

Background: Smoking remains a leading cause of death worldwide, with young adults particularly at risk due to the lack of targeted cessation initiatives. While mobile apps show promise in supporting smoking cessation, they primarily target smokers already motivated enough to install them, highlighting the need for interventions that reach those who are not yet ready to take that step.

Objective: This paper focuses on designing and evaluating Smokwit, a digital smoking cessation intervention aimed at young adults during the act of smoking. Smokwit seeks to investigate the early stages of smoking cessation (precontemplation and contemplation) that are important yet rarely investigated.

Methods: The paper is based on the design science research methodology where a digital intervention—Smokwit—was designed and evaluated in the wild using a mixed method approach combining quantitative results of a quasi-experiment with qualitative insights from users and experts. More specifically, Smokwit is a novel gamified ambient intervention that integrates a connected ashtray with a mobile app. The ashtray aims to trigger processes of change, in particular consciousness raising and social liberation (as part of the transtheoretical model of change) by provoking curiosity, self-reflection, and ad-hoc peer discussions among smokers. The linked mobile app is designed to reinforce this goal by providing smoking cessation self-help material and coaching possibilities. We evaluated the effectiveness of this intervention through a 3-month field study designed as a quasi-experiment with a treatment and control group (n=46). A qualitative analysis with users (n=10) and smoking cessation experts (n=7) provides insights into the type of interactions that happened within and outside the system as well as practical implications for smoking cessation organizations.

Results: The qualitative findings revealed that the intervention promoted smokers' self-reflection, peer discussions, and mobile app interactions. Furthermore, the quantitative analysis uncovered a possible trend toward increased readiness to quit among smokers in the treatment group compared to the control group; however, this did not reach conventional levels of statistical significance (b=1.33; z=1.91; P=.06).

Conclusions: Smokwit provides encouraging insights into how to design a bottom-up digital intervention that targets young adults at an opportune moment to support them on their smoking cessation journey.

(JMIR Hum Factors 2025;12:e72749) doi:10.2196/72749

KEYWORDS

smoking cessation; ambient intelligence; gamification; user study; behavioral research

XSL•FO
RenderX

Introduction

Background

Despite many years of prevention, tobacco addiction is still the cause of approximately 8 million deaths annually, making it the foremost preventable cause of death globally [1]. Efforts to pass comprehensive laws and regulations can be hindered in countries where the tobacco industry provides economic benefits and thus wields significant lobbying power. For example, in Switzerland, where the industry has some of its major headquarters, this lobbying power is particularly strong [2]. Specifically, the tobacco industry in Switzerland contributes approximately US $7 billion in annual revenue, representing about 1% of the country's gross domestic product, and provides employment for around 11,500 people (0.2% of the labor force) [3]. However, it can be argued that these economic benefits are offset by the health-related expenses of US $11 billion yearly [4], including the 10,000 smoking-related deaths annually [5]; thus, passing top-down regulation remains highly challenging. As a result, in such countries, it is crucial to devise effective bottom-up prevention measures to support smoking cessation.

There are effective evidence-based interventions that could be helpful for young smokers; these include brief advice, behavioral support, pharmacotherapy, and abstinence evaluation [6]. Furthermore, digital technologies, such as mobile apps, have been identified as a potentially effective, lightweight, and scalable complementary support for smoking cessation [7-9]. These technologies could represent a particular interest, potentially lowering current funding, time, and human resource barriers for health promotion [10]. For instance, there is a wide range of specialized mobile apps that could act as catalysts to provide this support [9,11]. However, despite their potential benefits, there remains a challenge with user engagement, especially when the apps are not well-designed or integrated into clinical or daily workflows [12]. Indeed, even minor frictions prevent smokers from engaging in effective behavior change interventions [13]. For instance, it is challenging to get a smoker to seek digital self-help information, sign up for peer support, or even register for face-to-face counseling [14]. As such, one of the main challenges in smoking cessation interventions is effectively getting smokers on board [15,16]. Most existing digital solutions require individuals to first be aware of the app's existence, and then take the additional steps of downloading and installing it; there are barriers that can deter engagement, especially for those not actively seeking help. Conversely, our approach initiates onboarding directly at smoking hotspots through physical devices, offering immediate and contextual engagement. This paper aims to address this gap by exploring an alternative, low-threshold entry point into cessation support.

A particularly important demographic to target prevention measures are young adults (ages 18 to 35 years). Most smokers start early, with 87% beginning before the age of 21 years [17], and young adults often represent the largest smoking demographic [17]. This is particularly problematic because the earlier one starts smoking regularly, the faster addiction sets in [18], thus making it harder to quit. Despite approximately half of young people wishing to quit [19], only an estimated 5% manage to do so during their youth [20]. Furthermore, smoking cessation initiatives that are specifically aimed at young adults are lacking [21]. As such, we seek to address the following research question: How can a digital intervention be designed to increase young adult engagement with evidence-based smoking cessation interventions?

Contributions

This research makes 3 main contributions to the literature on digital interventions for smoking cessation. First, it introduces Smokwit—a novel digital system specifically designed to reach smokers at physical smoking hotspots—immediately before, during, or after the act of smoking. Unlike conventional tools that rely on user motivation or active digital engagement, Smokwit reduces entry barriers by engaging users in situ and at opportune moments [22,23]. It leverages connected ashtrays as ambient objects [24], facilitating a seamless transition from the physical environment to digital smoking cessation resources. Furthermore, Smokwit is conceptualized as a motivational affordance [25], a system intentionally designed to trigger behavioral outcomes by incorporating gamification mechanisms such as competition and rewards [26], with the goal of increasing engagement with evidence-based interventions. Second, this paper presents a real-world validation of the Smokwit system through a 3-month quasi-experimental field study (n=46) conducted across 2 university campuses. This study combines quantitative data (surveys and system logs) with qualitative insights from interviews with participants (n=10) and experts (n=7), thus providing a comprehensive understanding of system engagement and user experience in a natural context. Third, the paper contributes empirical insights into the early stages of smoking cessation, specifically the precontemplation and contemplation stages as defined by the transtheoretical model (TTM) of health behavior change [27]. These stages are critical yet underexplored in digital health research [28]; our findings shed light on how low-friction, context-aware interventions can effectively engage smokers who are not yet actively seeking help.

Smoking Cessation Process

Giving up smoking is more of a process rather than a single decision or action [29]. A useful theoretical lens for analyzing this process is the TTM [27], which conceptualizes behavior change as a progression through discrete stages. Unlike continuum models, which assume interventions can be applied at any point, stage models such as the TTM emphasize that different strategies are appropriate at different stages in the change process [29-31]. In this context, the stages toward smoking cessation can be seen as a smoker's readiness to quit (RTQ). TTM hypothesizes that this change occurs in 6 distinct steps, also called stages of change:

- Stage 1 – precontemplation: no intention of taking action within the next 6 months.
- Stage 2 – contemplation: intention of taking action within the next 6 months.
- Stage 3 – preparation: intention of taking action within the next 30 days; thus, some behavioral steps have already been taken.

XSL•FO
RenderX

- Stage 4 – action: changed overt behavior for less than 6 months.
- Stage 5 – maintenance: changed overt behavior for more than 6 months.
- Stage 6 – termination: no temptation to relapse and 100% confidence.

In this paper, we focus on the early stages—precontemplation and contemplation—that are especially underaddressed in digital health interventions. These are stages where smokers are not yet ready to take direct action; however, they may be more receptive to change through targeted support. According to TTM, progress between these stages is mediated by cognitive and affective processes such as raising consciousness, meaning increasing awareness through information and feedback, and social liberation, thus reinforcing social norms that support healthy behavior [31].

Our intervention, Smokwit, is deliberately designed to activate these processes in contextually appropriate ways. By placing connected ashtrays at smoking hotspots, Smokwit raises awareness of quitting resources (consciousness raising) and introduces quitting as a socially supported behavior (social liberation). The use of ambient technology lowers friction for engagement, helping reach users who are not yet actively seeking help. Moreover, Smokwit incorporates gamification elements, such as rewards and peer comparison, to further reinforce social motivation and self-reflection [25,26]. In line with TTM, these mechanisms aim to help users move from passive awareness (stage 1) to the contemplation of change (stage 2).

In addition to these motivational design elements, Smokwit supports 2 evidence-based engagement strategies. First, peer interaction is integrated into the system to support social comparison and group influence, which have been shown to strengthen motivation and behavior change in stopping smoking [6,32,33]. This leads to our first premise:

- Premise 1: Engagement with peer groups helps smokers increase their RTQ.

Second, Smokwit offers access to curated self-help material, which can also support users in the early stages of change. Prior studies on smoking cessation apps show that such content can improve RTQ, especially among less motivated users [9,34-37]. This confirms our second premise:

- Premise 2: Engagement with self-help material helps smokers increase their RTQ.

By aligning Smokwit's design with key TTM processes and stages, we aim to offer a context-sensitive intervention that addresses a critical gap in digital smoking cessation efforts.

Gamifying Smoking Cessation Interventions

There is a growing number of technological solutions, such as forums and mobile apps, that give smokers access to self-help material and leverage peer support by facilitating multilateral interactions between smokers on dedicated social media channels [6,9,38,39]. However, it is unclear how many smokers actually engage with such resources [9]. A particular challenge is to capture smokers who are still in the precontemplation phase and, as such, are not interested in quitting at present and to engage for the first time.

One approach to increase engagement with digital or mobile content is to conceptualize interventions as motivational affordances, such as gamification or digital nudging [40]. These are artifacts designed to foster specific behaviors [25]. An immediate behavioral outcome in our case would be increased engagement with digital self-help resources or peer group reflection. A distal outcome would be to move smokers from the precontemplation to the contemplation stage by increasing their RTQ.

The gamification approach [26], where game elements such as points, leaderboards, or badges are integrated into nongame systems, has already been extensively used in mobile apps targeting smoking cessation [9]. Typically, around 65% of the top 100 smoking cessation apps include some kind of reward system, such as points or digital badges and trophies awarded when milestones are reached (eg, number of days without smoking) [9]. However, these systems primarily use these mechanisms to encourage existing users to stay on the platform and are not designed to engage new users. In this paper, we want to address this gap by focusing on bringing new users to the system.

Opportune Moment Prompting

One approach to bringing users to the system can be inspired by digital nudging—a motivational affordance approach that modifies a user's choice to make certain behaviors more attractive [41]. The idea with certain types of digital nudges is to prompt users at opportune times to transition from one context to another [23,42]. An example of such prompting is simply to display push notifications on the phone, which indicate to users that something is happening, for example, a message has been received, a location reached, or a time limit exceeded. The opportune timing of prompting is an important aspect [22]. These moments can be life events, where stopping smoking could be linked to a health-related event such as a doctor's visit, pregnancy, or hospitalization [43]. However, there could also be moments close to more mundane behavior. For instance, road safety, reminding drivers to use their seat belts shortly before they drive away, significantly increases seat belt compliance in comparison to reminding drivers 5 minutes before they drive away [44]. In the smoking context, a recent study found promising results for face-to-face interventions in areas where smokers gather to smoke (smoking hotspots) [45]. In this paper, we build on these insights to investigate how a digital artifact could prompt users at smoking hotspots without the need for in-person interventions.

Ambient Transition Device

A promising approach to transition from a physical context, smoking, to a digital context, accessing digital content, is the use of ambient devices. An ambient device is a regular object from physical space augmented through a digital artifact, which can enhance awareness of digital activities [46]. In a knowledge-sharing system, a connected aquarium was designed to visualize digital contributions to motivate users to log in and make their own contributions [24]. Similarly, in the

XSL•FO
RenderX

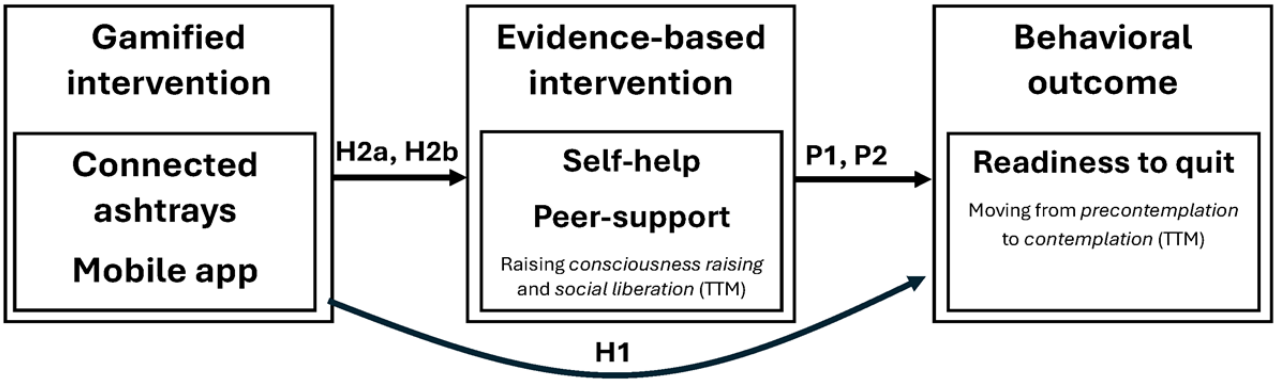
energy-saving context, artistic screens connected to smart meter data have been shown to increase awareness of energy consumption and encourage energy-saving behaviors [47]. In the medical domain, researchers provided patients with asthma with an "Asthma Buddy" teddy bear [48]. In this system, simple feedback is provided through the teddy bear, and more elaborate data are displayed on a mobile app. We build on these previous findings to hypothesize that a gamified ambient device at a hotspot could help reduce the friction for smokers to engage more with digital self-help and peer support. This increased engagement will lead to progress in the process of stopping smoking such as RTQ. Figure 1 illustrates our conceptual model. The gamified intervention, implemented through connected ashtrays in public smoking areas, acts as a playful and tangible nudge toward 2 core evidence-based strategies, self-help and peer support. We state our main hypothesis as follows:

- Hypothesis 1: Gamifying an ambient device at smoking hotspots increases RTQ.

We argue that this increase can happen by increasing engagement with self-help material and reflection among peers. Formally stated:

- Hypothesis 2a: Gamifying an ambient device at smoking hotspots increases engagement with self-help material.
- Hypothesis 2b: Gamifying an ambient device at smoking hotspots increases engagement with peers.

Figure 1. Overview of the full conceptual model illustrating our hypotheses (H1, H2a, and H2b) and premises (P1 and P2). TTM: transtheoretical model.



Methods

Overview

This research uses a user-centered design approach [49], which incorporates elements of design thinking [50] and design science research methodology [51,52]. The artifact's design was guided by practice and theory, with psychological and behavioral outcomes used to measure its effectiveness. This project involved an 18-month collaborative effort with a multidisciplinary team of human-computer interaction and information systems researchers, embedded systems engineers, and partners from smoking cessation organizations. Co-design workshops were conducted with these partners to understand the issue and explore the design space before creating initial prototypes. Then the design was evaluated in the wild in a field study. The study takes a mixed method approach [53], combining elements from a quantitative quasi-experimental design setup with qualitative interview analyses. We adopted a multimodal data collection approach [54], with survey data, system logs, and interview transcripts.

Ethical Considerations

The regional ethics committee of the Canton of Vaud, Switzerland (CER-VD) reviewed the study and determined that it fell outside the scope of the Swiss Human Research Act (Req-2022-01567); thus, no formal ethics approval was required. At the beginning of the survey, participants were informed about the purpose of the study, the voluntary nature of participation, and their right to withdraw at any time without consequences. Informed consent was implied by clicking "Next" to begin the survey. Data were pseudonymized, stored securely on access-restricted servers, and reported only in aggregate. No direct financial compensation was provided; participants completing both surveys were entered into a prize draw, while interviewees received no compensation.

Smokwit System Design

Overview

To help young smokers in the precontemplation phase to increase their RTQ, Smokwit leverages two evidence-based behavioral support approaches directly at the smoking hotspot: (1) self-help material and (2) peer group support. The first approach, self-help material such as digital videos, can aid smokers without needing outside assistance, and it is deemed useful if it is tailored to individual smokers [55]. In the second approach, which may be more effective [56,57], peers provide mutual support and encouragement, either through a digital forum [11] or during in-person group sessions.

At the center of the Smokwit system are connected voting ashtrays, which act as ambient transition devices, providing a link between the physical smoking hotspot and a mobile app, which provides digital support (Figure 2). The interaction with the Smokwit system can be described through 6 phases, representing a typical user flow (Figure 3). The journey begins when a smoker arrives at a hotspot and notices the connected ashtray during the "smoking break starts" phase. In the "connected ashtray engagement" phase, the smoker interacts with the ashtray by reading reflective questions and scanning a QR code to access additional digital resources, linking the

XSL•FO
RenderX

physical and digital components of the system. The ashtray presents a question (Table 1) along with 2 answer options, each corresponding to a separate ashtray compartment. Smokers cast their vote by placing their cigarette butt in the slot, which reflects their answer. The questions were divided into 2 sets. The first set aimed to capture individual smokers' perceptions, while the second set presented aggregated feedback from the broader smoking community on each campus. At the hotspot, smokers could also engage in peer discussions fostered by the ashtray's presence, creating opportunities for shared reflection and mutual support. Scanning the QR code transitions the smoker to the "mobile app engagement" phase, where they can access self-help materials, coaching resources, maps, and forums. The journey may continue into the "postsmoking engagement" phase, where smokers further interact with the mobile app. They can create an account to participate in forums, analyze personal statistics, or contact support organizations.

Figure 2. The Smokwit system design concept includes the connected voting ashtray with a link redirecting users to a mobile app with digital support materials.

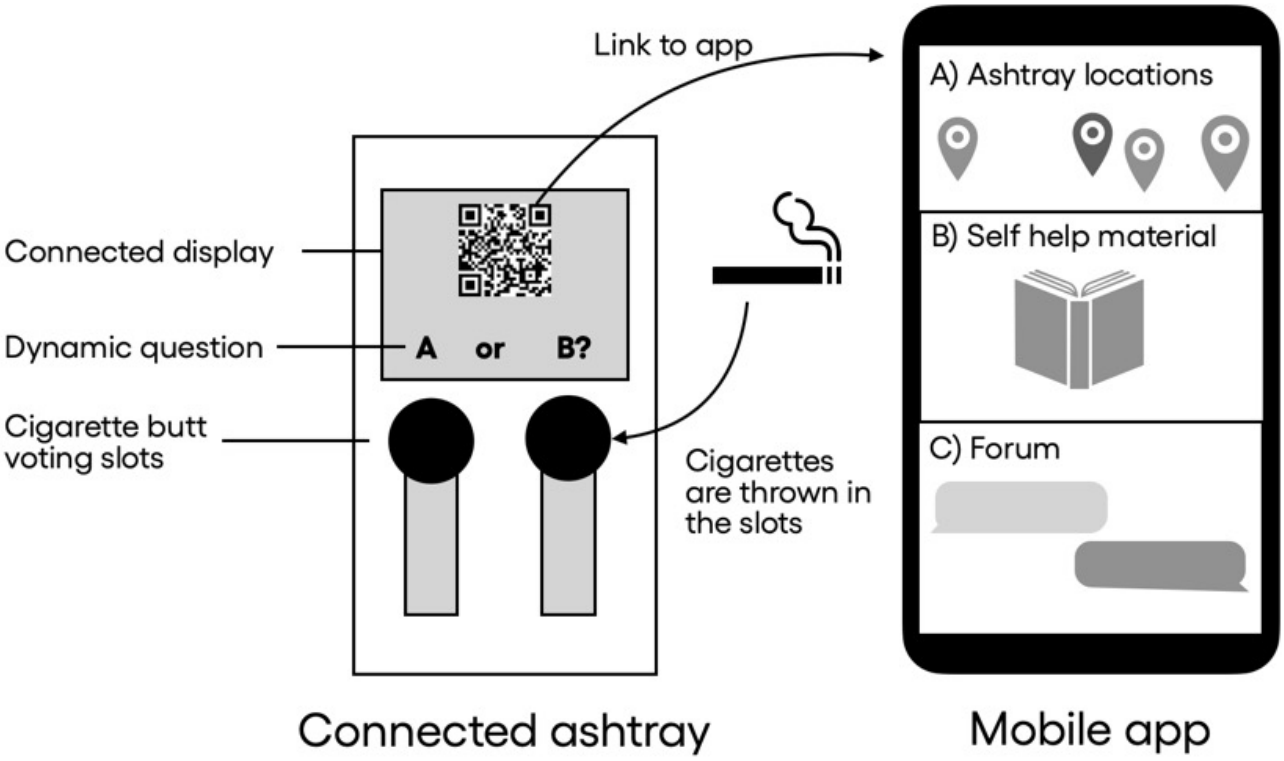
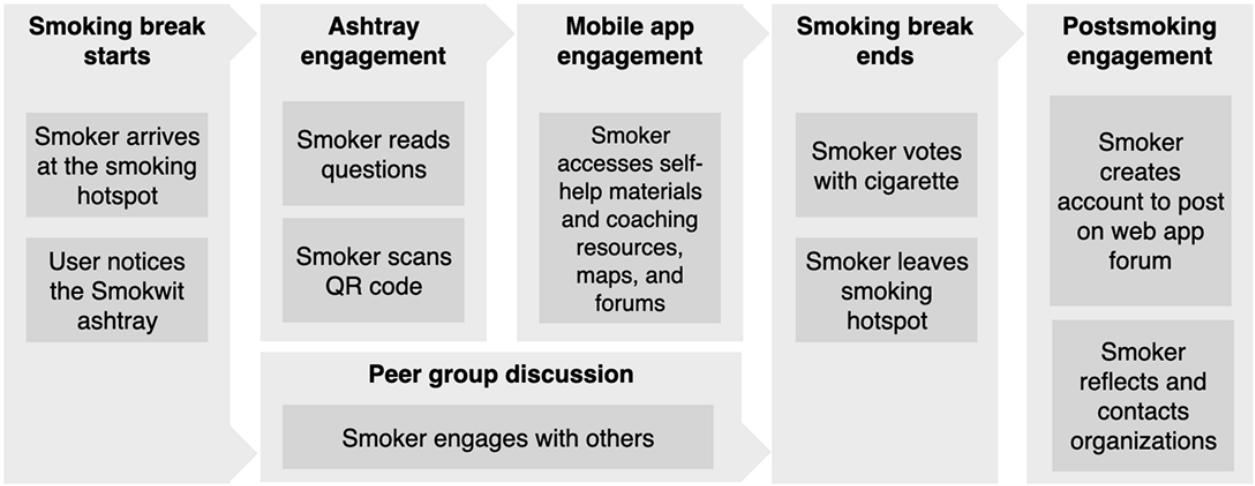


Figure 3. Smokwit smoker journey represented by 6 phases: smoking break starts, ashtray engagement, mobile app engagement, peer group discussion, smoking break ends, and postsmoking engagement.



XSL•FO
RenderX

Table 1. Poll questions displayed on Smokwit ashtrays and the number of cigarette butts collected for each response from 1 hotspot.

Poll question examples	Set	Option 1	Butts for option 1, n	Option 2	Butts for option 2, n
I smoked this cigarette to give myself a boost	1	Yes	51	No	82
I found this cigarette ...	1	Great	67	Not great	30
This cigarette will help me focus	1	I think so	32	No need!	22
On Campus 1, what % of cigarettes are smoked to get a boost?	2	<50%	20	≥50%	41
What % of Campus 1 smokers actually enjoy cigarettes?	2	<50%	5	≥50%	16
On Campus 1, what % of cigarettes are smoked to help focus?	2	<50%	10	≥50%	18

One of the main system design priorities for the user experience was lowering friction. We focused on the solution's simplicity to ensure that anyone, regardless of prior experience or technical proficiency, can become a potential user. To engage users with self-help materials and group discussions, the system was designed following Zhang's [25] 5 key principles for creating motivational affordances, which are design elements intended to promote specific user behaviors. These design principles (DPs) are (DP1) allowing autonomy and self-control, (DP2) supporting competence and achievement, (DP3) fostering social connections, (DP4) facilitating leadership and followership, and (DP5) triggering emotions. Moreover, the system design incorporates the Fogg Behavior Model [58], ensuring that all 3 essential elements as in motivation, ability, and behavior triggers are present to support behavior change in cigarette addiction effectively.

The Smokwit ashtray's simple design eliminates the need for advanced abilities or complex instructions, ensuring that the system is intuitive and accessible to all users. Smokers can cast their "vote" by placing their cigarette butt in the compartment, which corresponds to their answer (DP1). This low physical effort supports the ability component [58] by minimizing barriers to engagement while simultaneously strengthening motivation by fostering a sense of autonomy and ease of use. Users are free to choose whether to participate and which ashtray they want to interact with. The system is designed to support competency and achievement (DP2) through a check-in mechanism where scanning a new ashtray's QR code unlocks it in the mobile app and grants access to its statistics and forum. The app displays both discovered and locked ashtrays on a campus map and presents poll results, giving users a clear sense of progress. This feature is designed to motivate users to go to different places and engage in reflection both in person and within the app's forum (DP3). This incorporates facilitator triggers, which reduce friction with spark triggers such as gamified rewards and visible progress to prompt participation in discussions [58]. When joining discussions, users create a username and can upload an avatar; however, their true identity is not required. Conversations are anchored in the ashtray's shared question and poll results, offering opportunities for asynchronous peer reflection and social connections. The system also provides experience points (XPs) to users who participate in forum discussions to support the leadership or followership, social comparison, aspect (DP4). These points are displayed under the usernames in the forum, making active users more visible to others and potentially encouraging further engagement through subtle social referencing. This feature incorporates motivational elements such as social acceptance [58]. Finally, the system supports curiosity (DP5) with a bright yellow ashtray that draws attention, a rotating set of reflective questions on its display, and prompts to discover other ashtrays on campus. These features initiate a pleasure motivator, a spark trigger, which both arouse interest and motivate users to engage further [58].

Connected Ashtrays

The Smokwit connected ashtrays (Figure 4) are based on modified commercial voting ashtrays. A voting ashtray consists of 2 compartments, each labeled with a distinct choice in the form of a printed question or poll. In the Smokwit version of the ashtrays, the paper has been replaced with a connected e-paper screen, and sensors were installed on the inside of the ashtray to allow the counting of cigarette butts in both ashtray compartments. An ESP32 microcontroller unit is used to pilot the e-paper screen, sensors, and Wi-Fi connectivity to (1) update questions remotely and (2) send counted votes to the mobile app. To make the ashtrays energy autonomous, they include a 10,000 mAh battery and a 6 W solar panel.

XSL•FO
RenderX

Figure 4. Smokwit connected ashtray at the smoking hotspot.

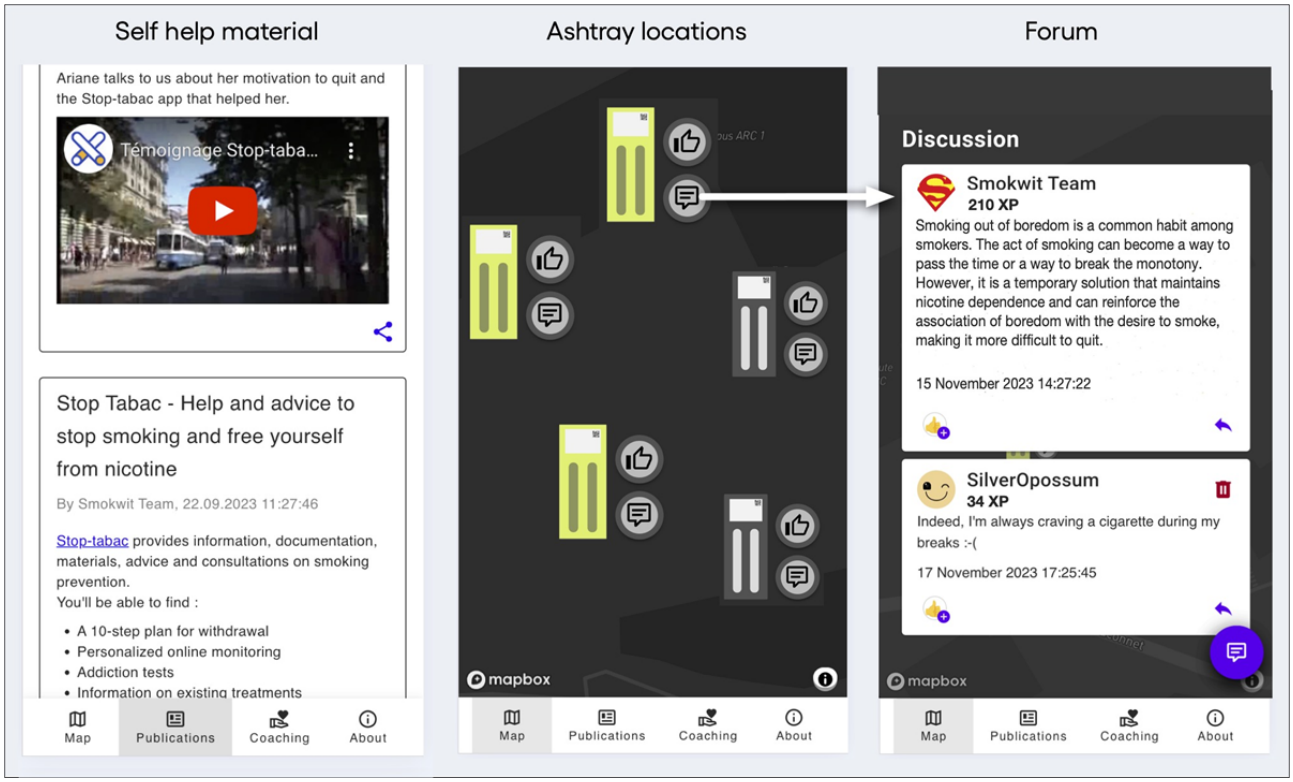


Mobile App

The Smokwit mobile app (Figure 5) was developed as a responsive web app. The mobile app provides self-help material with tobacco prevention content as well as a coaching section, where users can contact regional smoking cessation experts. The main screen of the mobile app is the campus 1 map, on which the Smokwit ashtrays have been geolocated. Each ashtray has a scannable QR code. By scanning the QR code, the user unlocks the ashtray on the mobile app and earns XPs. The physical ashtray can be used independently; however, the app enhances the experience by providing additional insights. When scanning an ashtray, it unlocks it on the map by becoming yellow. By unlocking the ashtray on the platform, the user directly arrives at the ashtray page and has access to the digital discussion forum associated with the ashtray's current poll question. Indeed, users can access this information without registration. Open access promotes inclusivity and user engagement, making it easier for individuals to join in forum discussions and explore resources at their own pace. The user is invited to comment on the poll question by leaving a message or using emojis in the dedicated community section. Each action is rewarded with XPs. Posting and reacting to messages require an account, which can be created on the go by creating only a username and password. A username is automatically generated and suggested to the user to minimize friction and maintain anonymity. When the ashtray is unlocked, the user can also consult information regarding current and past polls as well as additional statistics such as the daily and hourly inflow of cigarette butts. There is an ashtray page for each Smokwit ashtray.

XSL•FO
RenderX

Figure 5. Smokwit mobile app with views of self-help material, a map showing connected ashtrays, and an example of the digital discussion associated with a poll question.



Field Study

A field study was designed in the wild on 2 university campuses to assess the impact of the Smokwit intervention on enhancing smokers' progress in their quitting process through their RTQ (hypothesis 1) by fostering self-reflection (hypothesis 2a) and peer support (hypothesis 2b).

Measures

To better understand the participants, in addition to age and sex, we also inquire about the age the smokers started, if they were smoking daily, and their number of previous attempts to stop smoking as a preliminary measure of their desire to stop [59,60].

To assess the user experience and understand self-reflection and peer support processes, we log activity on the mobile app as well as the number of votes from the ashtrays. Furthermore, to obtain more in-depth insights, we conducted semistructured interviews with smokers exposed to Smokwit and experts from smoking cessation organizations.

The selected metric to measure progress on the quitting smoking process is the RTQ ladder [59] (Table 2). This provides a graphic representation of the continuum of change, ranking smokers' RTQ smoking from 1 to 10. Level 1 for smokers having no interest in quitting at all to level 10 for smokers who have become permanent nonsmokers and who have little concern that they will relapse. Levels 2 to 5 are precontemplators, according to the TTM. Level 6 is equivalent to the contemplation stage. Levels 7 and 8 indicate that the smoker is, respectively, in the preparation and action stages, whereas level 9 indicates maintenance.

XSL•FO
RenderX

Table 2. Readiness to quit ladder with stages of change, their level equivalents, and descriptions.

Stage of change	Level	Description
Termination	10	I have quit smoking, and I will never smoke again.
Maintenance	9	I have quit smoking, but I still worry about slipping back, so I need to keep working on living smoke-free.
Action	8	I still smoke, but I have begun to change, for example, cutting back the number of cigarettes I smoke. I am ready to set a quit date.
Preparation	7	I definitely plan to quit smoking in the next 30 days.
Contemplation	6	I definitely plan to quit smoking in the next 6 months.
Precontemplation	5	I often think about quitting smoking, but I have no plans to quit.
Precontemplation	4	I sometimes think about quitting smoking, but I have no plans to quit.
Precontemplation	3	I rarely think about quitting smoking, and I have no plans to quit.
Precontemplation	2	I never think about quitting smoking, and I have no plans to quit.
None	1	I enjoy smoking and have decided not to quit smoking for my lifetime. I have no interest in quitting.

Evaluation Setup

Overview

We conducted the field study on 2 university campuses following a quasi-experimental design with a treatment (campus 1) and a control (campus 2) condition. We performed the study in a campus setting because the student population would, by design, consist of young adults, who are our target demographic. Furthermore, in such contexts, students tend to smoke at the same time, meaning during lecture breaks, which is prone to group discussion. It is a quasi-experiment, as the participants were the natural members of each geographically distant campus of the same university and thus were not randomly assigned. Table 3 gives an overview of the field study.

Table 3. Experimental phases with study groups and participant amount.

	Phase 1 (June to September 2023): presurvey, n (%)	Phase 2 (September to December 2023): intervention	Phase 3 (January to March 2024): postsurvey, n (attrition %)	Phase 4 (June to November 2024): postinterviews, n (%)
Treatment	47 (65)	Prevention day + 4 Smokwit ashtrays	27 (43)	10 (21)
Control	25 (35)	Prevention day	19 (24)	—[a]
Experts	—	—	—	7 (—)
Total	72 (100)	—	46 (36)	10 (—)

[a]Not applicable.

Phase 1: Recruitment Pretest

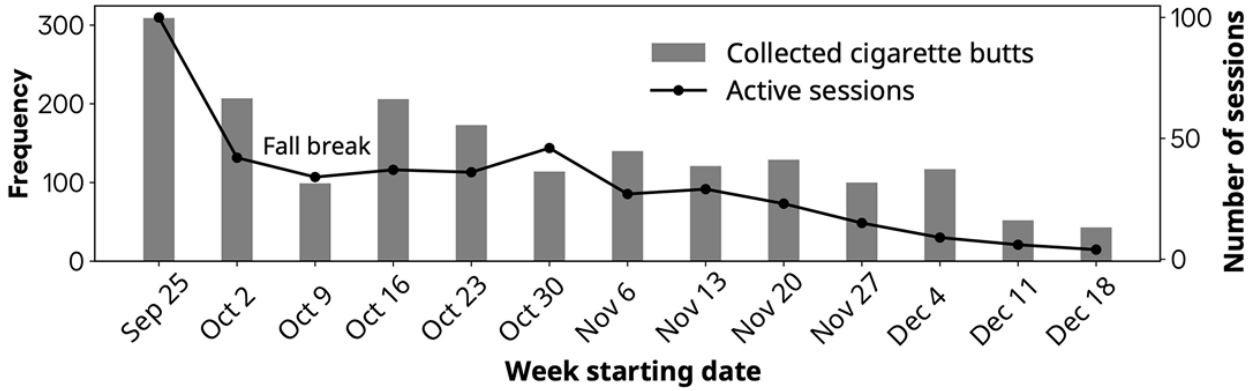
Before the experiment began, we recruited participants by asking smokers in various smoking locations on the 2 campuses for their email addresses. A total of 119 participants registered for the study, and 72 fully completed the presurveys (47 for the treatment and 25 for the control conditions). Two researchers developed the survey and pretested it in collaboration with regional smoking cessation experts (Cipret Neuchâtel). It included 9 items on demographics, smoking habits, and the literature-validated RTQ scale, which participants completed in approximately 6.50 (SD 7.08) minutes.

Phase 2: Intervention

The intervention phase lasted 3 months. It began with a smoking prevention day on both campuses, led by 2 experts from a local smoking prevention organization (Cipret Neuchâtel). The event featured a free information stand aimed at encouraging participants in both conditions to consider quitting and providing contact information. This setup was designed to reduce the risk of social desirability bias by giving participants on both campuses a sense of being part of an experiment. During the entire intervention phase, 4 Smokwit ashtrays were deployed only on campus 1 (the treatment condition). Each Smokwit ashtray showed a different poll question, and the polls were updated every 3 days. In total, 70 poll items were created. The first set of 35 poll questions (ie, set 1 in Table 1) aimed to gather community data (eg, “I smoked this cigarette to give myself a boost,” yes or no). The second set of 35 poll questions (ie, set 2 in Table 1) aimed to raise awareness about consumption and encourage contemplation on quitting smoking (eg, “On Campus 1, what % of cigarettes are smoked to get a boost?,” <50% or ≥50%). Every butt tossed into a Smokwit ashtray was detected by sensors and considered for the poll question statistics. Figure 6 illustrates the number of cigarette butts tossed weekly alongside the number of sessions on the mobile app. These data indicate that smokers were more active in the early weeks of September. Reduced smoking activity happened when the weather gets colder, and this has been observed previously [61].

XSL•FO
RenderX

Figure 6. Number of cigarette butts collected by ashtrays and number of active sessions on the mobile app.



Phase 3: Postsurvey

The postsurvey included the same questions on smoking habits and RTQ as presurvey. A survey was sent to all participants after the experiment in mid-January 2024, with a follow-up email in late February. The postsurvey was fully completed by 46 valid participants who also completed the presurvey (27 in the treatment condition and 19 in the control condition). This took on average 5.30 (SD 2.06) minutes to complete.

Phase 4: Postinterviews

We conducted 2 rounds of follow-up interviews with participants to understand user experience and long-term impressions. The first round was conducted in June 2024 with 5 participants, and the second one in November 2024 with 5 additional participants. We invited participants who had completed both pre- and postsurvey and expressed interest in a follow-up interview. Through these interviews, we sought to understand motivations, barriers, and decision-making processes, providing deeper insight into their behavior and perceptions of the system. Furthermore, in November 2024, we interviewed 7 smoking cessation experts to gather their professional feedback.

Results

Overview

Hereafter, we present the quantitative results of the quasi-experiment. We conducted all analyses using Python scripts, which are available alongside the datasets in Multimedia Appendix 1. The demographics and smoking habits are presented in Figure 7 and summarized in Tables 4 and 5. They show that the participants belong to the demographic that we planned to address, namely, mainly young adults in their 20s, who are in their precontemplation phase when it comes to stopping smoking.

Figure 7. Presurvey results for demographics, smoking habits, and readiness to quit (treatment group: n=46 and control group: n=27).

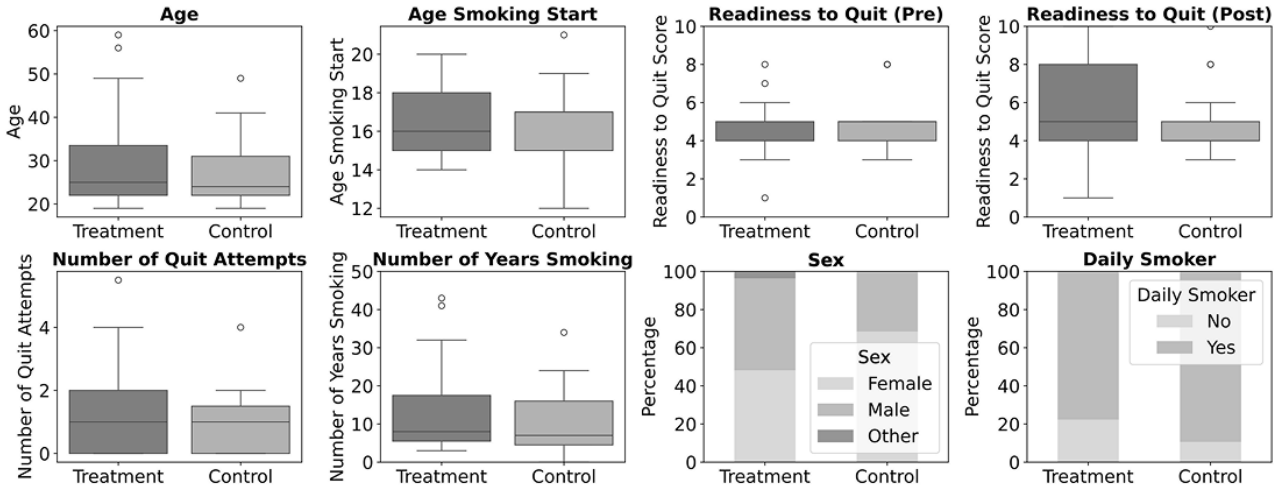


Table 4. Summary statistics for treatment and control groups (n=46) collected in presurvey from participants who completed both pre- and poststudy surveys.

Variable	Treatment (n=27), n (%)	Control (n=19), n (%)
Female	13 (48)	13 (68)
Other	1 (4)	0 (0)
Daily smoker	21 (78)	17 (89)

XSL•FO
RenderX

Table 5. Descriptive statistics for age, age smoking start, number of quit attempts, readiness to quit (RTQ) in presurvey, and RTQ in postsurvey for treatment and control groups.

Variable	Treatment group			Control group		
	Mean (SD)	Range	25%, 50%, 75%	Mean (SD)	Range	25%, 50%, 75%
Age (years)	30.00 (11.73)	19-59	22, 25, 33	27.37 (8.72)	19-49	22, 24, 31
Age smoking start (years)	16.59 (1.74)	14-20	15, 16, 18	16.42 (2.06)	12-21	15, 17, 17
Number of quit attempts	1.52 (1.47)	0-5.5	0, 1, 2	0.89 (1.10)	0-4	0, 1, 1.5
RTQ (pre)	4.37 (1.36)	1-8	4, 4, 5	4.95 (1.51)	3-8	4, 5, 5
RTQ (post)	5.67 (2.62)	1-10	4, 5, 8	5.00 (1.83)	3-10	4, 4, 5

In terms of interaction with the Smokwit system, the QR codes on the ashtrays were scanned 245 times. Based on session IDs, we could account for 74 people with an average interaction duration of 2 minutes and 11 seconds (SD 70 seconds). The ashtrays' webpages were viewed 2409 times, with a mean of 50.9 (SD 35.0) views per user and an average length of engagement of 59 (SD 36.1) seconds. The page containing self-help material was viewed 44 times, with a mean of 3.67 (SD 1.57) views per user and an average length of engagement of 20 (SD 7.54) seconds. This use is in line with what could be expected with the number of participants in the treatment group (77 people in the beginning and 46 at the end).

Attrition Rate

The overall attrition rate of the study was 36% (n=46), with 24% (n=19) in the control group and 43% (n=27) in the treatment group. This attrition rate is rather high and common in smoking cessation studies over several months, where attrition rates have been found to be as high as 77% [62] and even 90% [37]. Nevertheless, it is important to understand that this attrition might have created an imbalance in the sample and potentially introduced bias. We assessed if there was a statistical difference between participants who provided full answers to both surveys (n=46) and those who only answered the presurvey (n=72) across all measures taken in the presurvey (sex, age, age smoking start, group, number of quit attempts, and RTQ). The sample was balanced, although 2 motivation-related variables showed marginally significant outcomes. Smokers with more quitting attempts (t_{25}=1.81; *P*=.08) and higher RTQ scores (t_{25}=1.79; *P*=.08) may have left the study more often than those with lower RTQ scores and fewer attempts to quit. The results suggest that, even if such imbalances exist, the primary target population of this study, namely, individuals in the early stages of their smoking cessation process, has been retained in the sample. Nevertheless, future studies with larger samples should replicate our findings to ensure that attrition bias did not modify the results.

Ambient Gamification and RTQ (Hypothesis 1)

To test hypothesis 1, we analyzed whether the Smokwit intervention increased participants' RTQ compared to a control group. A linear mixed effects model was used with restricted maximum likelihood estimation [63]. The model included time (pre or post), treatment (Smokwit vs control), and their interaction, along with relevant covariates: sex, age, number of quit attempts, age smoking start, and smokes daily. A random intercept was included for each participant to account for individual differences in baseline RTQ.

The results (Table 6) show a possible interaction effect (Figure 8) between time and treatment: when controlling for all other variables, participants in the treatment group showed a potential greater increase in RTQ scores compared to the control group after the 3-month experiment (b=1.33; *z*=1.91; *P*=.06). Even though the result is close to statistical significance, it does not reach it. This finding hints at possible support for hypothesis 1. The effect size for this interaction, measured using Cohen f^2, was 0.04, representing a small effect size according to conventional benchmarks [64]. While modest, such effects are typical and meaningful in behavioral science and digital health research [65]. To evaluate model fit, we computed the marginal and conditional R^2 values. The marginal R^2 (variance explained by fixed effects) was 0.27, and the conditional R^2 (fixed + random effects) was 0.38, indicating that both predictors and individual variability contributed meaningfully to the variance in RTQ. The results indicate no significant association between RTQ and sex or the age of smoking start.

Additionally, the results show that the number of past quit attempts is positively associated with RTQ (b=0.60; *z*=3.81; *P*<.001), and daily smokers reported lower RTQ scores than nondaily smokers (b=–1.07; *z*=–1.99; *P*=.047), meaning that the higher the number of previous quit attempts, the higher the smoker's RTQ, and the more a smoker smokes, the lower their RTQ. Furthermore, the results also show a negative marginally significant trend (b=–0.04; *z*=–1.83; *P*=.07), indicating that younger smokers may have a higher RTQ than the older ones.

Note that the results also indicated a potential imbalance in the baseline sample when controlling for the other variables, because the main effect of treatment was significant (b=–1.09; *z*=–1.97; *P*=.049) when pooling pre- and postresults together. This means that the treatment group may have had a lower baseline RTQ than the control group. Figure 8 visualizes these results.

To further investigate the difference between pre- and postlevels of RTQ in both conditions, we conducted a complementary paired sample for each group. In the treatment group, RTQ scores increased significantly from pre- to postintervention (pre: mean 4.37, SD 1.36; post: mean 5.67, SD 2.62; t_{26}=2.72; *P*=.01), with a medium effect size (Cohen *d*=0.52). Conversely, the control group showed no significant change (pre: mean 4.95, SD 1.51; post: mean 5.00, SD 1.83; t_{18}=0.11; *P*=.91; *d*=0.03).

XSL•FO
RenderX

A post hoc power analysis based on the observed effect size (d=0.52) and sample size (n=27) indicated a power of 0.74 to detect the change in RTQ, which is close to the conventional threshold of 0.80 [64]. This suggests that while the study may be slightly underpowered, the observed effect is meaningful and warrants further investigation with larger samples.

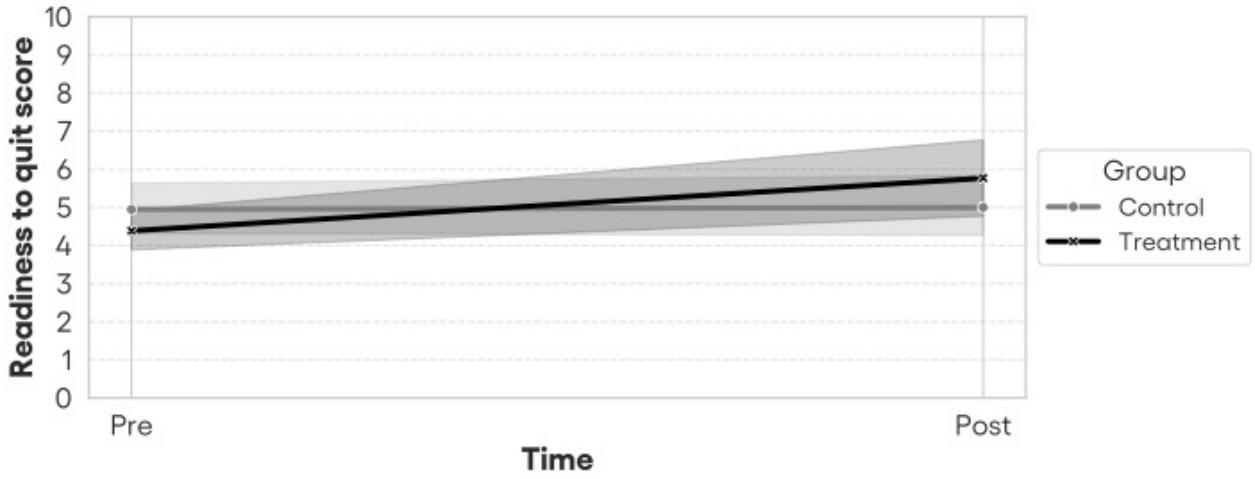
Taken together, these results cautiously support hypothesis 1: ambient gamification through the Smokwit device may help increase RTQ among young adults; however, future studies with larger samples are needed to confirm the robustness of this effect.

Table 6. Results of the mixed linear model regression on readiness to quit scores.

Effect	Coefficient b (SE)	*z* score	*P* value	95% CI
Intercept	7.69 (2.03)	3.78	<.001	3.71 to 11.66
Time (post)	0.05 (0.53)	0.10	.92	–0.99 to 1.09
Treatment	–1.09 (0.55)	–1.97	.049	–2.16 to –0.01
Treatment over time (interaction)	1.33 (0.70)	1.91	.06	–0.03 to 2.70
Number of quit attempts	0.60 (0.16)	3.81	<.001	0.29 to 0.91
Age of smoking start	–0.09 (0.11)	–0.81	.42	–0.30 to 0.13
Smokes daily	–1.07 (0.54)	–1.99	.047	–2.12 to –0.02
Sex (male)	0.68 (0.42)	1.63	.10	–0.14 to 1.49
Age	–0.04 (0.02)	–1.83	.07	–0.08 to 0.00
Group variance	0.45 (0.34)	—[a]	—	—

[a]Not applicable.

Figure 8. Visual representation of the interaction effect between time and group on readiness to quit scores with 95% CIs.



Qualitative Analysis (Hypothesis 2a and Hypothesis 2b)

Overview

To gain additional insights into the user experience with Smokwit, we conducted interviews with 10 participants and 7 smoking cessation experts. We conducted semistructured interviews with 10 participants (Participant 1-Participant 10) from the treatment group between 6 and 9 months after the end of the study. These interviews focused on four key themes: (1) interactions with the ashtrays, (2) self-reflection, (3) peer support, and (4) user feedback. The discussions explored participants' memories of the study, including their interactions with the ashtrays, the system, and other smokers, as well as their smoking habits and suggestions for improving the intervention. Additionally, we presented our project at the Swiss Association for Tobacco Control (AT Switzerland) Conference 2024, where we collected feedback from 7 smoking cessation experts (Expert 1-Expert 7) on our system.

We followed a qualitative data analysis via a general inductive approach [66]. Interviews were audio-recorded, transcribed uniformly, and reviewed multiple times to identify recurring patterns and derive significant themes.

Ashtrays Interactions

The interviews revealed that participants found the Smokwit ashtrays' questions engaging, particularly at the beginning of the intervention. Participant 1 remarked, "I liked these questions, not all of them, but often, they were funny. They made me laugh," highlighting the positive reception to the reflective prompts. Participant 3 also described the questions as highly

XSL•FO
RenderX

engaging and innovative, noting how they encouraged smokers to reflect on their habit. Similarly, Participant 4 appreciated the weekly changing questions, which sparked curiosity and led to meaningful reflections on smoking habits. The gamified and humorous design of the questions was central to sustaining participants' interest and stimulating self-reflection. Participant 5 highlighted the ashtrays' visibility, noting their "super visible yellow" color, which made them accessible and noticeable on campus. Participant 9 and Participant 8 also remarked on the ashtrays' appealing design, describing it as "hard to miss" and "innovative and engaging, especially during the initial weeks of use." The ashtrays managed to attract attention beyond smokers, as Participant 3 observed, "We were all going to see the question, even non-smokers were going to see." However, this broad appeal diminished over time as the novelty wore off, as Participant 3 noted, "At first, it was attractive, but afterward, I think they didn't change the questions every day." Practical challenges, such as crowding during peak hours, also reduced engagement, as Participant 10 explained, "The voting process sometimes felt crowded, which could have discouraged participation." Experts empathized with the nonintrusive nature of the ashtrays, noting their passive presence in smokers' environments. Expert 2 remarked, "I like the passive side of the object. It is present in the smoker environment, but it is not intrusive," which allowed participants to engage without feeling pressured. Expert 1 highlighted the ashtrays' potential for workplace deployment, suggesting that companies could use them to gather data and justify targeted interventions. Expert 6 further noted, "Smoking prevention centers would be able to use these ashtrays as an extra tool prior to traditional in-person meetings and hold discussions on various subjects depending on the information gathered." Expert 7 recommended testing the ashtrays in public spaces such as bus stops to evaluate their effectiveness in diverse contexts, "It could be interesting to test it on bus stops also." These suggestions highlight the flexibility of intervention design for alternative environments. Despite challenges such as diminished novelty and logistical issues, the intervention successfully sparked curiosity and reflection among participants. The ashtrays' design and passive approach made them a versatile tool, with potential apps extending beyond individual behavior changes to organizational and community contexts.

Self-Reflection

The interviews revealed that the Smokwit system fostered self-reflection, disrupting habitual smoking behavior. The questions posed by the ashtrays prompted smokers to think more critically about their habits. As Participant 1 remarked, "it (ashtray) makes you think, yes, think about what you're doing," while Participant 2 emphasized, "It makes smoking more conscious." These prompts heightened awareness and occasionally sparked discussions about the broader implications of smoking, as observed by Participant 10. Participant 4 highlighted the intervention's dual impact, stating, "It's true that these are not necessarily reflections that we usually ask ourselves." By encouraging such reflections, the ashtray helped participants reconsider their behaviors and the broader implications of their smoking habits. Participant 3 acknowledged, "It (ashtray) told me even more that I should quit smoking," suggesting that while the ashtrays did not directly reduce smoking, they reinforced the motivation to stop. This indicates that the intervention planted seeds of change, providing a foundation for future behavior modification. The gamified aspect of the ashtrays, such as combining playful and thought-provoking questions, further stimulated self-reflection. As Participant 10 remarked, "I liked the fact that it was a bit playful. And then, it made us think, so they were surely good questions." However, Participant 6 noted that without frequent updates, the questions became less relevant over time. Participant 2, who initially felt uncomfortable with the ashtrays' focus on her smoking habit, engaged indirectly through discussions with colleagues who had interacted directly with the system. Even by indirect engagement, her smoking became more conscious, demonstrating the ashtrays' broader impact on mindfulness. Although participants reported limited immediate changes in their smoking behavior, the intervention reinforced awareness and the desire to quit. Experts emphasized the ashtrays' ability to disrupt smoking rituals, prompting deeper reflection. Expert 3 explained, "Smoking is seen by smokers as a ritual. Intervening during the ritual is absolutely fantastic and has, in my opinion, great potential. It also breaks the routine." External stressors, such as life challenges and academic pressure, influenced participants' engagement with reflective prompts. Participant 3 mentioned going through a particularly difficult period, which may have impacted both smoking habits and engagement with the ashtrays: "We've always got excuses, so you had put them (ashtrays) in July, you had put them in whenever. It would all be the same. I had a lot of problems at that time, it would not change anything," highlighting the impact of personal difficulties on the intervention's effectiveness. Participant 7 similarly emphasized, "Quitting is harder than reflecting, it's the execution that counts." Additionally, Participant 5 pointed out that academic stress often led students who did not normally smoke to resume during high-pressure periods, such as examinations. These insights suggest that while immediate behavior change was limited, the intervention laid the foundation for future efforts by increasing awareness and prompting critical self-reflection. By encouraging smokers to think more critically about their habits, even amid external stressors, the Smokwit system shows promise as a long-term catalyst for change. These findings provide some qualified support for hypothesis 2a.

Colocated Peer Support on Smoking Hotspots

Our results show that the Smokwit system fostered peer support and facilitated meaningful discussions among smokers. The intervention encouraged conversations about smoking cessation, helping smokers build a support network. Participant 1 shared, "We talked a little about our experiences. Of our attempts to quit," illustrating how the ashtrays encouraged smokers to share their personal experiences and reflections. The gamified elements of the ashtrays, such as the changing questions and the visible, playful design, served as conversation starters, breaking down barriers and making it easier for smokers to engage with one another. According to Participant 10, the ashtrays "sparked meaningful debates among colleagues about smoking," while Participant 9 noted, "The ashtrays encouraged casual debates among smokers, sometimes involving

XSL•FO
RenderX

non-smokers too." These interactions fostered a sense of community, as observed by Participant 4: "And then the interactions we can have with the smokers around." Experts emphasized the strength of leveraging smoker communities for behavior change. Expert 3 explained, "A smoker community is felt by the smokers as an entity. Acting in the community itself is a real strength of the project. Smokers in the community will talk with each other and are influenced by each other." Expert 6 suggested that the ashtrays could be deployed in workplace settings to facilitate discussions before formal interventions, stating, "Smoking prevention center could use these ashtrays to promote dialogue around smoking behaviors and cessation efforts prior to in-person meetings." Expert 7 proposed that using ashtrays over an extended period before interventions would help target those most willing to engage in cessation programs. The ashtrays promoted social connections and collective reflection, creating a supportive environment, which encouraged participants to reconsider their smoking habits. Participant 4 mentioned that the ashtrays often led to small debates, reinforcing the collective effort to rethink smoking behaviors. Similarly, Participant 3 observed that both smokers and nonsmokers were drawn to the ashtrays, leading to some level of interaction and conversation, although this later declined as the novelty wore off.

By facilitating conversations and fostering social connections, the Smokwit system showed potential in influencing attitudes concerning smoking within the campus community. These findings suggest that the intervention not only supported individuals but also laid the groundwork for a more collaborative approach to addressing smoking habits, providing some qualified support for hypothesis 2b.

User Feedback

Participants offered valuable suggestions for improving the Smokwit intervention. Participant 1 and Participant 10 proposed extending the playful voting system using cigarette butts, which Participant 10 described as "a simple yet engaging mechanism." Participant 9 and Participant 5 proposed expanding the forum functionalities by creating a dedicated space for smokers to log their reflections digitally, emphasizing that "a platform for shared experiences could boost long-term involvement." Participant 9 suggested enhancing engagement by offering rewards following interactions with the ashtray, such as displaying live statistics not only within the mobile app but also directly on the ashtray screen. To keep the content relevant, Participant 3 highlighted the importance of updating questions more frequently and incorporating humor. Technological challenges with the app were sometimes a concern. Participant 6 and Participant 1 experienced difficulties scanning QR codes due to outdated smartphones, while Participant 7 and Participant 8 preferred direct interactions with the ashtray over using a QR code. Participant 7 noted, "Taking out the phone and scanning the QR code, even if it's only two minutes, can be a hindrance," and Participant 8 added, "Interacting directly with the physical object would be more immediate and engaging." Experts recommended tailoring ashtray questions to different stages of behavioral change and integrating self-management tools. Expert 1 remarked, "It could be a good idea to have a similar kind of artifact for sugar and alcohol but also other forms of tobacco/nicotine consumption." Expert 5 highlighted the value of deploying ashtrays in diverse settings, such as universities, to compare results across demographic groups, noting, "I would like to have this ashtray at the university where I work to compare the results between different types of cultural backgrounds." Both participants and experts offered constructive feedback for improvement, offering a clear path for refining the Smokwit intervention. By addressing these technological opportunities and incorporating feedback, future iterations of the Smokwit intervention can become more engaging and impactful.

Discussion

Principal Findings

The results provide tentative support for the fact that including a gamified ambient object—a connected voting ashtray—at a smoking hotspot visited by young adults can complement existing evidence-based interventions and lead to increased RTQ among this demographic. Our findings align with previous research, which shows that simple interventions can potentially yield positive results [6,34]. Our controlled quasi-experiment results indicate a potential tendency for participants in the treatment condition to progress from the precontemplation to the contemplation stage, which could potentially lead to an increase in RTQ compared to the control condition (hypothesis 1). This tentative result may represent an early indication of the intervention's potential to motivate change in young smokers, though its significance needs investigation in larger samples. While previous research has shown promise for in-person interventions at the smoking hotspots to leverage the act of smoking as an opportune moment to incentivize change [45], our research extends these findings by showing that ambient objects could be used as transitional devices. Although this study extended over 3 months and included follow-up interviews at 6 and 9 months, future research should aim to replicate and extend these findings in studies with longer durations and larger sample sizes. This would help clarify the role of novelty effects, the influence of community dynamics, and participants' ongoing progress in the smoking cessation journey. Building on our work, future studies could also use path analysis to test the underlying conceptual model and examine both direct and indirect effects of the intervention components.

Characteristics of the Target Population

The Smokwit intervention focused on young adults, aged 18 to 35 years, an important and understudied demographic. Indeed, smoking initiation often occurs before 21 years [17], and young adults tend to remain in the precontemplation stage longer due to lower perceived risks [45]. By embedding gamified ashtrays in habitual smoking environments, the intervention facilitated reflection and dialogue without requiring formal participation. The gamified intervention was tailored to young adults with gamification and fun poll questions, which were appreciated by participants. Although the study targeted young adults, participants older than 35 years of age were not excluded, reflecting the diversity of real-world smoking environments.

In our qualitative analysis, stress emerged as a key influence on smoking behaviors, with participants citing academic

XSL•FO
RenderX

pressure and personal challenges as triggers. This is aligned with previous findings, which list smoking urges and stress as smoking triggers [67,68]. Participants often indicated that life challenges, such as academic or personal stress, played a role in their smoking patterns and their engagement with the intervention. Future work could aim to better understand the role of stress in the outcome of such interventions and how they could be designed to better cater to this characteristic.

Ambient Interaction

Interaction analyses and interviews shed light on the mechanisms, which led to this outcome. First, the connected ashtrays, with their thought-provoking questions and gamified features, led participants to engage with the digital material by scanning the QR codes. The qualitative data suggest that the engagement with the ambient object played a critical role in raising awareness about smoking behaviors and encouraging self-reflection (hypothesis 2a). This aligns with previous research in other contexts, which showed the use of an ambient object to transition from physical to digital environments [24]. This also aligns with literature suggesting that behavior change apps should be designed to give users maximum control over their level of engagement and disengagement [69]. In this study, the Smokwit ashtray acted as a passive entry point into the digital intervention. Participants did not need to sign up or commit to a program to begin reflecting on their behavior. Instead, smokers encountered provocative, low-effort prompts in a familiar space, which encouraged casual interaction. This dynamic encouraged ongoing self-assessment, even among those not actively seeking to quit, and appeared to contribute to increased RTQ over time. This approach differs from typical behavior change apps, which rely on continuous digital engagement [70]. In Smokwit, a minimal physical interaction, voting with a cigarette butt, initiated a reflective process and offered a seamless path into digital support. Future research could investigate how to deepen these interactions directly at the smoking hotspot. For instance, integrating chatbots into the ashtray's interface could increase playfulness and interaction while reducing friction to access self-help material [59]. Moreover, they could contribute to increasing emotional attachment to the system and potentially motivation to change behavior [71]. One participant mentioned that even though she felt the intervention was positive in the end, it could come across as moralizing. This could be problematic, as it might create reactance and backfire [72]. Future work should clarify how widespread this feeling is and how design strategies can mitigate unintended negative effects.

Sense of Community

Our findings align with previous literature, supporting that engagement in digital smoking cessation communities is positively associated with the process of change [8]. While face-to-face peer groups have long been recognized as effective for fostering behavior change [32], this study extends that understanding by demonstrating the unique value of spontaneously formed, physically colocated peer groups. These groups emerge naturally at smoking hotspots, therefore eliminating the need for individuals to actively register or join formal programs. By delivering the intervention directly to smokers in their habitual environments, this approach reduces many of the typical barriers associated with participating in structured peer-support programs and leverages a gamified, spontaneous design to encourage engagement. Smokers may find it easier to make changes in their own lives if they are encouraged to share their experiences in digital communities dedicated to quitting smoking [8], but also in real-life, directly at the smoking hotspot. The ashtrays not only facilitated individual self-reflection but also promoted social interactions among smokers (hypothesis 2b). The gamified questions functioned as conversation starters, enabling smokers to discuss their experiences and challenges in real time, directly at the smoking hotspot. This sense of community fostered through the intervention may have reinforced participants' RTQ by providing social support, an essential component in the process of behavior change [6]. Initiatives such as Weight Watchers have shown that "the magic of the meeting" has long been seen as an effective means of enhancing behavior [32]. Future work should explore how these spontaneous peer groups could be sustained over longer periods and investigate the mechanisms underlying their formation and effectiveness. Additionally, future research could examine the interplay between the various components of this intervention, including the app, forum, information, and ambient devices, as well as the influence of the implementation context. The context in which the intervention is deployed may mediate the sense of community it fosters, which is critical for its success [73].

Implications for Practice

Smokwit facilitated discussions among campus smokers about the intervention and, to some extent, quitting smoking. The physical-digital integration, combining ashtrays at smoking hotspots with a mobile app, engages young smokers in both physical and digital environments. This approach increases accessibility and encourages spontaneous peer support by enabling interactions at familiar locations while providing digital resources and community forums without time and space limits. Moreover, it also improved the RTQ.

The findings from this study, informed by participants' experiences and smoking cessation experts' feedback, suggest several practical strategies for designing and implementing smoking cessation interventions. First, embedding interventions in habitual environments, such as smoking hotspots, fosters spontaneous peer groups. This approach encourages immediate social support and organic discussions about smoking behaviors. Second, tailoring intervention content to align with different stages of behavioral change enhances relevance and effectiveness, particularly when paired with gamified elements, which sustain engagement. To address varying user preferences, digital tools such as web platforms for sharing reflections can complement physical interventions and encourage long-term involvement. Experts highlighted the ashtrays' passive and accessible design, which reduces barriers to participation and makes them suitable for diverse settings, including workplaces, bus stops, and public spaces. The ashtrays' ability to collect data while engaging participants underscores their potential for broader applications in other contexts. These insights emphasize the importance of integrating interventions into users' daily routines and combining physical and digital elements to

XSL•FO
RenderX

maximize accessibility and impact. Future research could further explore how such an intervention could leverage context particularities to increase engagement. For instance, research could investigate if social mechanisms such as competition between groups or reciprocity could be used to foster higher peer support or a sense of community around a digital intervention. Such refinements could improve RTQ and deepen the intervention's long-term impact.

Limitations

This research is not without limitations. First, as with other experiments performed in the wild over several months, sample decline and potential differences between conditions (campus 1 and campus 2) might be beyond the researchers' control. To address these sample differences, we controlled demographic characteristics such as age and sex in the analysis. However, other factors may still have influenced the outcome. Second, the intervention included several aspects of gamification, such as initial prompts (the poll questions), a treasure hunt feature (finding other ashtrays), and status (the XPs), making it challenging to isolate the effects of individual components. Third, while the ashtrays initially attracted significant attention, engagement declined over time. The reason for this cannot be clearly identified because it could be due to the waning of the novelty effect or lower smoking activity during cooler months to name 2 potential reasons. Future studies could aim to better understand optimal timing and durations for such interventions. Finally, external factors such as personal challenges and academic stress further moderate the intervention's effectiveness on smoking behavior. The postintervention interviews demonstrated that the Smokwit ashtrays positively affect smokers' awareness and social interactions regarding smoking. Although the immediate impact on smoking behavior was limited, the intervention raised consciousness concerning smoking habits and sparked discussions about quitting, providing a foundation for future enhancements, which could further amplify the intervention's effectiveness and reach. It is important to note that the higher education context of the study may limit generalizability, because participants were more likely to adopt technology than the general population. Finally, the RTQ scale relied on self-reported data, which may be subject to social desirability bias. Additionally, the postsurvey timing at the beginning of the year, from January 17 to March 5, 2024, coincided with the "New Year's Resolution" effect, which may have inflated RTQ scores. While these overestimations likely affected both conditions similarly, it cannot be excluded that one condition was more susceptible to these biases than the other.

Conclusions

In this study, we investigated how a gamified ambient intervention called Smokwit deployed directly at a smoking hotspot could help smokers progress at the early stage of their smoking cessation process. This is a crucial public health issue, as this demographic represents the largest tobacco consumer base. Our approach was designed to complement existing evidence-based interventions and act as an entry point toward self-help material and peer support. We designed gamified voting ashtrays, which would engage smokers at the smoking hotspot and lead them to discuss with each other and access digital material, leading them to increase their RTQ. The qualitative results of our field experiment showed that the intervention did indeed foster digital interaction and discussions among peers. Furthermore, our quantitative analysis provided preliminary evidence that smokers in the intervention condition may have experienced a greater increase in RTQ, more than those in the control group, although the statistical strength of this observation remains limited. Overall, these findings are encouraging and indicate that well-designed digital interventions can potentially act as important complements to existing interventions, driving significant behavior change and promoting healthier lifestyles within the campus community. Despite encouraging trends, the study's limited power necessitates cautious interpretation of the results. These preliminary findings should be confirmed in future trials with larger and more balanced samples.

Acknowledgments

The Smokwit project was partially funded by the Swiss Tobacco Control Fund.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Dataset and analysis scripts (Jupyter Notebook) for transparency and reproducibility.
[ZIP File (Zip Archive), 659 KB - humanfactors_v12i1e72749_app1.zip]

References

1. Tobacco. World Health Organization. 2023. URL: https://www.who.int/news-room/fact-sheets/detail/tobacco [accessed 2025-06-01]
2. Global Tobacco Industry Interference Index. Global Center for Good Governance in Tobacco Control (GGTC). 2024. URL: https://globaltobaccoindex.org/ [accessed 2024-06-21]

XSL•FO
RenderX

3. Project SUN: a study of the illicit cigarette market in the EU, Norway and Switzerland: 2017 results. KPMG. 2018. URL: https://assets.kpmg.com/content/dam/kpmgsites/uk/pdf/2018/07/project_sun_executive_summary_2018.pdf [accessed 2024-06-21]
4. Tabagisme. Unisanté. 2016. URL: https://www.unisante.ch/fr/promotion-prevention/tabagisme/ [accessed 2024-06-21]
5. Tobacco-attributable mortality. MonAM—Swiss Monitoring System of Addiction and Noncommunicable Diseases. 2025. URL: https://ind.obsan.admin.ch/en/indicator/monam/tobacco-attributable-mortality [accessed 2024-06-21]
6. Verbiest M, Brakema E, van der Kleij R, Sheals K, Allistone G, Williams S, et al. National guidelines for smoking cessation in primary care: a literature review and evidence analysis. NPJ Prim Care Respir Med 2017 Jan 20;27(1):2 [FREE Full text] [doi: 10.1038/s41533-016-0004-8] [Medline: 28108747]
7. Hartmann-Boyce J, Stead LF, Cahill K, Lancaster T. Efficacy of interventions to combat tobacco addiction: Cochrane update of 2013 reviews. Addiction 2014 Sep;109(9):1414-1425. [doi: 10.1111/add.12633] [Medline: 24995905]
8. De Santo A, Moro A, Kocher B, Holzer A. From digital community engagement to smoking cessation: insights from the Reddit r/stopsmoking thread. 2021 Presented at: Proceedings of the 54th Hawaii International Conference on System Sciences (HICSS); January 4, 2021; Grand Wailea, HI, United States. [doi: 10.24251/hicss.2021.479]
9. De Santo A, Holzer A. Assessing digital support for smoking cessation. In: Lecture Notes in Business Information Processing. Cham: Springer; 2019. [doi: 10.1007/978-3-030-44322-1_25]
10. Duncan LR, Pearson ES, Maddison R. Smoking prevention in children and adolescents: a systematic review of individualized interventions. Patient Educ Couns 2018 Mar;101(3):375-388. [doi: 10.1016/j.pec.2017.09.011] [Medline: 28987451]
11. De Santo A, Moro A, Kocher B, Holzer A. Helping each other quit online: understanding user engagement and real-life outcomes of the r/StopSmoking digital smoking cessation community. Trans Soc Comput 2023 Apr 08;6(1-2):1-30. [doi: 10.1145/3564745]
12. Portz J, Moore S, Bull S. Evolutionary trends in the adoption, adaptation, and abandonment of mobile health technologies: viewpoint based on 25 years of research. J Med Internet Res 2024 Sep 27;26:e62790 [FREE Full text] [doi: 10.2196/62790] [Medline: 39331463]
13. Privitera GJ, Zuraikat FM. Proximity of foods in a competitive food environment influences consumption of a low calorie and a high calorie food. Appetite 2014 May;76:175-179. [doi: 10.1016/j.appet.2014.02.004] [Medline: 24560689]
14. Paay J, Kjeldskov J, Brinthaparan U, Lichon L, Rasmussen S, Srikandaraja N. Quitty: using technology to persuade smokers to quit. 2014 Presented at: Proceedings of the 8th Nordic Conference on Human-Computer Interaction (NordCHI '14); October 26-30, 2014; Helsinki, Finland. [doi: 10.1145/2639189.2639195]
15. Bold KW, Garrison KA, DeLucia A, Horvath M, Nguyen M, Camacho E, et al. Smartphone apps for smoking cessation: systematic framework for app review and analysis. J Med Internet Res 2023 Jul 13;25:e45183 [FREE Full text] [doi: 10.2196/45183] [Medline: 37440305]
16. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010 Feb 17;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
17. Tobacco consumption (age: 15+). MonAM—Swiss Monitoring System of Addiction and Concommunicable Diseases. 2025. URL: https://ind.obsan.admin.ch/en/indicator/monam/tobacco-consumption-age-15 [accessed 2024-06-21]
18. Tabac. Addiction Suisse. 2020. URL: https://shop.addictionsuisse.ch/fr/tabac-nicotine/82-159-focus-tabac.pdf [accessed 2024-06-21]
19. Kuendig H, Notari L, Gmel G. Le tabagisme chez les 15 à 25 ans en 2014: analyse des données 2014 du monitorage suisse des addictions. Addiction Suisse. 2015. URL: https://www.suchtmonitoring.ch/docs/library/kuendig_dmnxxct80wxl.pdf [accessed 2024-06-21]
20. Keller R, Krebs H, Hornung R. Enquête sur la consommation de tabac en Suisse (monitoring tabac). Universität Zürich. 2003. URL: https://www.addictionsuisse.ch/wp-content/uploads/2023/01/enquete-sur-la-consommation-de-tabac-en-suisse-monitoring-tabac-resume-du-rapport-sur-les-jeunes.pdf [accessed 2024-06-21]
21. Arrêter avec succès. Stop-Tabac. 2022. URL: https://stop-tabac.ch/arreter-avec-succes [accessed 2024-06-21]
22. Rogers T, Frey E. Changing behavior beyond the here and now. In: Gideon K, W W, editors. The Wiley Blackwell Handbook of Judgment and Decision Making. Chichester, West Sussex: Wiley-Blackwell; 2015:723-748.
23. Purohit AK, Holzer A. Functional digital nudges: identifying optimal timing for effective behavior change. 2019 Presented at: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems; May 4-9, 2019; Glasgow, Scotland. [doi: 10.1145/3290607.3312876]
24. Holzer A, Kocher B, Bendahan S, Vonèche Cardia I, Mazuze J, Gillet D. Gamifying knowledge sharing in humanitarian organisations: a design science journey. Eur J Inf Syst 2020;29(2):153-171. [doi: 10.1080/0960085x.2020.1718009]
25. Zhang P. Technical opinion: Motivational affordances: reasons for ICT design and use. Commun ACM 2008;51(11):145-147. [doi: 10.1145/1400214.1400244]
26. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness. 2011 Presented at: Proceedings of the 15th International Academic MindTrek Conference (MindTrek '11); September 28-30, 2011; Tampere, Finland. [doi: 10.1145/2181037.2181040]

XSL•FO
RenderX

27. Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: toward an integrative model of change. J Consult Clin Psychol 1983 Jun;51(3):390-395. [doi: 10.1037//0022-006x.51.3.390] [Medline: 6863699]
28. Lin H, Wang Y, Xing Y, Han Y, Zhang C, Luo T, et al. A personalized mobile cessation intervention to promote smokers from the preparation stage to the action stage: double-blind randomized controlled trial. J Med Internet Res 2023 Apr 26;25:e41911 [FREE Full text] [doi: 10.2196/41911] [Medline: 37099360]
29. Prochaska JO, Goldstein MG. Process of smoking cessation. Implications for clinicians. Clin Chest Med 1991 Dec;12(4):727-735. [Medline: 1747990]
30. Schwarzer R, Luszczynska A. How to overcome health-compromising behaviors: the health action process approach. Eur Psychol 2008 Jan;13(2):141-151. [doi: 10.1027/1016-9040.13.2.141]
31. Prochaska J, Redding C, Evers K. The transtheoretical model and stages of change. In: Glanz K, Rimer BK, Viswanath K, editors. Health Behavior: Theory, Research, and Practice. 5th Edition. San Francisco, CA: Jossey-Bass; 2015:125-148.
32. Goetz T. The Decision Tree: Taking Control of Your Health in the New Era of Personalized Medicine. Emmaus, PA: Rodale; 2010.
33. Faranda M, Roberts LD. Social comparisons on Facebook and offline: the relationship to depressive symptoms. Pers Individ Dif 2019;141:13-17. [doi: 10.1016/j.paid.2018.12.012]
34. Roberts N, Kerr S, Smith S. Behavioral interventions associated with smoking cessation in the treatment of tobacco use. Health Serv Insights 2013;6:79-85. [doi: 10.4137/hsi.s11092]
35. Etter JF, Khazaal Y. The Stop-tabac smartphone application for smoking cessation: a randomized controlled trial. Addiction 2022 May;117(5):1406-1415 [FREE Full text] [doi: 10.1111/add.15738] [Medline: 34738687]
36. Rajani NB, Mastellos N, Filippidis FT. Self-efficacy and motivation to quit of smokers seeking to quit: quantitative assessment of smoking cessation mobile apps. JMIR Mhealth Uhealth 2021 Apr 30;9(4):e25030 [FREE Full text] [doi: 10.2196/25030] [Medline: 33929336]
37. Bustamante Perez LA, Romo L, Zerhouni O. Feasibility and engagement of a mobile app preparation program (Kwit) for smoking cessation in an ecological context: quantitative study. JMIR Mhealth Uhealth 2024 Oct 02;12:e51025 [FREE Full text] [doi: 10.2196/51025] [Medline: 39357053]
38. Lin Y, Tudor-Sfetea C, Siddiqui S, Sherwani Y, Ahmed M, Eisingerich AB. Effective behavioral changes through a digital mHealth app: exploring the impact of hedonic well-being, psychological empowerment and inspiration. JMIR Mhealth Uhealth 2018 Jun 15;6(6):e10024 [FREE Full text] [doi: 10.2196/10024] [Medline: 29907557]
39. Christakis NA. Health care in a web. BMJ 2008 Jun 28;336(7659):1468 [FREE Full text] [doi: 10.1136/bmj.a452] [Medline: 18583678]
40. Bergram K, Djokovic M, Bezençon V, Holzer A. The digital landscape of nudging: a systematic literature review of empirical research on digital nudges. 2022 Presented at: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22); April 29-May 5, 2022; New Orleans, LA, United States. [doi: 10.1145/3491102.3517638]
41. Thaler RH, Sunstein CR. Nudge: Improving Decisions About Health, Wealth, and Happiness. New Haven, CT: Yale University Press; 2008.
42. Caraban A, Karapanos E, Gonçalves D, Campos P. A review of technology-mediated nudging in human-computer interaction. 2019 Presented at: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19); May 4-9, 2019; Glasgow, Scotland. [doi: 10.1145/3290605.3300733]
43. McBride CM, Emmons KM, Lipkus IM. Understanding the potential of teachable moments: the case of smoking cessation. Health Educ Res 2003 Apr;18(2):156-170. [doi: 10.1093/her/18.2.156] [Medline: 12729175]
44. Austin J, Sigurdsson SO, Rubin YS. An examination of the effects of delayed versus immediate prompts on safety belt use. Environ Behav 2006;38(1):140-149. [doi: 10.1177/0013916505276744]
45. Cheung YTD, Lam TH, Li WHC, Wang MP, Chan SSC. Feasibility, efficacy, and cost analysis of promoting smoking cessation at outdoor smoking "hotspots": a pre-post study. Nicotine Tob Res 2018 Nov 15;20(12):1519-1524. [doi: 10.1093/ntr/ntx147] [Medline: 28655173]
46. Pousman Z, Stasko J. A taxonomy of ambient information systems: four patterns of design. 2006 Presented at: Proceedings of the Working Conference on Advanced Visual Interfaces (AVI 2006); May 23-26, 2006; Venezia, Italy. [doi: 10.1145/1133265.1133277]
47. Rodgers J, Bartram L. Exploring ambient and artistic visualization for residential energy use feedback. IEEE Trans Vis Comput Graph 2011 Dec;17(12):2489-2497. [doi: 10.1109/TVCG.2011.196] [Medline: 22034370]
48. Gisvold A, Aarseth E. Using gamification and tangible user interfaces to treat asthmatic children. Norwegian University of Science and Technology. 2014. URL: https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/253710 [accessed 2024-09-21]
49. Abras C, Maloney-Krichmar D, Preece J. User-centered design. In: Bainbridge W, editor. Encyclopedia of Human-Computer Interaction. Thousand Oaks, CA: Sage Publications; 2004:445-456.
50. Bason C, Austin RD. The right way to lead design thinking. Harv Bus Rev 2019;97(2):82-91 [FREE Full text]
51. Peffers K, Tuunanen T, Rothenberger MA, Chatterjee S. A design science research methodology for information systems research. J Manag Inf Syst 2014 Dec 08;24(3):45-77. [doi: 10.2753/MIS0742-1222240302]
52. Dolak F, Uebernickel F, Brenner W. Design thinking and design science research. University of St. Gallen. 2013. URL: https://www.alexandria.unisg.ch/handle/20.500.14171/89211 [accessed 2024-09-21]

XSL•FO
RenderX

53. Van TK, Bennis A, Craenmehr S, Holwerda R, Jacobs M, Niels R. Design patterns for mixed-method research in HCI. 2014 Presented at: Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational (NordiCHI '14); October 26-30, 2014; Helsinki, Finland. [doi: 10.1145/2639189.2639220]
54. Ochoa X, Worsley M. Augmenting learning analytics with multimodal sensory data. J Learn Anal 2016;3(2):213-219. [doi: 10.18608/jla.2016.32.10]
55. Abroms LC, Lee Westmaas J, Bontemps-Jones J, Ramani R, Mellerson J. A content analysis of popular smartphone apps for smoking cessation. Am J Prev Med 2013 Dec;45(6):732-736 [FREE Full text] [doi: 10.1016/j.amepre.2013.07.008] [Medline: 24237915]
56. Reno C, Poole ES. It matters if my friends stop smoking: social support for behavior change in social media. 2016 Presented at: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16); May 7-12, 2016; San Jose, CA, United States. [doi: 10.1145/2858036.2858203]
57. Stead LF, Carroll AJ, Lancaster T. Group behaviour therapy programmes for smoking cessation. Cochrane Database Syst Rev 2017 Mar 31;3(3):CD001007 [FREE Full text] [doi: 10.1002/14651858.CD001007.pub3] [Medline: 28361497]
58. Fogg BJ. A behavior model for persuasive design. : ACM; 2009 Presented at: Proceedings of the 4th International Conference on Persuasive Technology (Persuasive '09); April 26-29, 2009; Claremont, CA, United States. [doi: 10.1145/1541948.1541999]
59. Abrams DB, Niaura R. The Tobacco Dependence Treatment Handbook: A Guide to Best Practices. New York, NY: Guilford Press; 2003.
60. Farkas AJ, Pierce JP, Zhu SH, Rosbrook B, Gilpin EA, Berry C, et al. Addiction versus stages of change models in predicting smoking cessation. Addiction 1996 Sep;91(9):1271-1280; discussion 1281. [Medline: 8854358]
61. Momperousse D, Delnevo CD, Lewis MJ. Exploring the seasonality of cigarette-smoking behaviour. Tob Control 2007 Feb;16(1):69-70 [FREE Full text] [doi: 10.1136/tc.2006.018135] [Medline: 17297078]
62. Belita E, Sidani S. Attrition in smoking cessation intervention studies: a systematic review. Can J Nurs Res 2015 Dec;47(4):21-40. [doi: 10.1177/084456211504700402] [Medline: 29509476]
63. West BT, Welch KB, Galecki AT. Linear Mixed Models: A Practical Guide Using Statistical Software. 3rd Edition. New York, NY: Chapman and Hall/CRC; 2022.
64. Cohen J. Statistical Power Analysis for The Behavioral Sciences. New York, NY: Routledge; 1988.
65. Selya AS, Rose JS, Dierker LC, Hedeker D, Mermelstein RJ. A practical guide to calculating Cohen's f(2), a measure of local effect size, from PROC MIXED. Front Psychol 2012;3:111 [FREE Full text] [doi: 10.3389/fpsyg.2012.00111] [Medline: 22529829]
66. Thomas DR. A general inductive approach for analyzing qualitative evaluation data. Am J Eval 2006;27(2):237-246. [doi: 10.1177/1098214005283748]
67. Businelle MS, Ma P, Kendzor DE, Frank SG, Vidrine DJ, Wetter DW. An ecological momentary intervention for smoking cessation: evaluation of feasibility and effectiveness. J Med Internet Res 2016 Dec 12;18(12):e321 [FREE Full text] [doi: 10.2196/jmir.6058] [Medline: 27956375]
68. Brewer JA, Elwafi HM, Davis JH. Craving to quit: psychological models and neurobiological mechanisms of mindfulness training as treatment for addictions. Psychol Addict Behav 2013 Jun;27(2):366-379 [FREE Full text] [doi: 10.1037/a0028490] [Medline: 22642859]
69. Smith W, Ploderer B, Wadley G, Webber S, Borland R. Trajectories of engagement and disengagement with a story-based smoking cessation app. 2017 Presented at: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17); May 6-11, 2017; Denver, CO, United States. [doi: 10.1145/3025453.3026054]
70. Zhu Y, Long Y, Wang H, Lee KP, Zhang L, Wang SJ. Digital behavior change intervention designs for habit formation: systematic review. J Med Internet Res 2024 May 24;26:e54375 [FREE Full text] [doi: 10.2196/54375] [Medline: 38787601]
71. Berney M, Ouaazki A, Macko V, Kocher B, Holzer A. Care-based eco-feedback augmented with generative AI: fostering pro-environmental behavior through emotional attachment. 2024 Presented at: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24); May 11-16, 2024; Honolulu, HI, United States. [doi: 10.1145/3613904.3642296]
72. Brehm SS, Brehm JW. Psychological Reactance: A Theory of Freedom and Control. New York, NY: Academic Press; 2013.
73. Hamari J, Koivisto J, Sarsa H. Does gamification work? A literature review of empirical studies on gamification. 2024 Presented at: Proceedings of the 47th Hawaii International Conference on System Sciences (HICSS); January 6-9, 2014; Waikoloa, HI, United States p. 3025-3034. [doi: 10.1109/hicss.2014.377]

Abbreviations

DP: design principle
RTQ: readiness to quit
TTM: transtheoretical model
XP: experience point

XSL•FO
RenderX

Edited by A Kushniruk; submitted 17.02.25; peer-reviewed by A Eisingerich, J King; comments to author 04.04.25; revised version received 30.06.25; accepted 15.07.25; published 19.09.25.

Please cite as:
Bartłomiejczyk N, De Santo A, Holzer A
Gamified Physical-Digital Smoking Cessation Intervention for Young Adults: Mixed Methods Development and Usability Study
JMIR Hum Factors 2025;12:e72749
URL: https://humanfactors.jmir.org/2025/1/e72749
doi:10.2196/72749
PMID:40971186

©Natalia Bartłomiejczyk, Alessio De Santo, Adrian Holzer. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 19.09.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Expanding the Design Space for Fall Prevention in Acute Orthopedic Hospital Care: Human-Centered Design Study

Maria Ehn[1], MSc, PhD; Åsa Revenäs[2,3,4], PhD; Helena Tobiasson[5], PhD

[1]School of Innovation, Design and Technology, Mälardalen University, Västerås, Sweden

[2]Ortopedic Clinic Västerås, Västerås Hospital, Region Västmanland, Västerås, Sweden

[3]Centre for Clinical Research, Uppsala University, Västerås, Sweden

[4]School of Health Care and Social Welfare, Mälardalen University, Västerås, Sweden

[5]School of Innovation, Design and Technology, Mälardalen University, Eskilstuna, Sweden

Corresponding Author:
Maria Ehn, MSc, PhD
School of Innovation, Design and Technology
Mälardalen University
Box 883
Västerås, S-721 23
Sweden
Phone: 46 21 107093
Email: maria.ehn@mdu.se

Abstract

Background: In-hospital fall prevention is a complex phenomenon most efficiently addressed via a wide range of multifactorial interventions. Technology may contribute, but research in this field has so far mainly focused on detecting falls. As a result, new knowledge from a system perspective is needed regarding when and how new technologies may support fall prevention among patients who have been hospitalized.

Objective: This study aimed to explore and describe clinical practices in an acute orthopedic hospital ward for fall prevention from a system perspective; determine the needs and possibilities related to support for clinical practices for fall prevention; and test whether a framework for studying interactions between people, activities, contexts, and technologies can be used to support observations of complex phenomena such as clinical fall prevention.

Methods: This qualitative study followed the principles of human-centered design while combining focused ethnography with a workshop. Eight health care professionals representing different staff categories in an acute hospital ward of an orthopedic clinic participated in on-site interviews or were observed in their clinical practice. Data from these events were subjected to qualitative content analysis to describe the clinical practices for fall prevention observed in terms of people, activities, context, and tools. In a workshop, a larger group of clinic personnel provided their views on fall prevention, described the activities and tools they observed to prevent falls, and discussed needs for further support.

Results: This study determined that health personnel considered fall prevention in all their interactions with patients, which included a wide range of activities for fall prevention wherein staff categories played complementary roles. These staff-patient meetings were goal oriented, responsive, and patient centered. The staff often served as key “tools” in assessment, communication, and coaching, while digital tools (mainly computer-based software programs) were used for information retrieval, documentation, and communication. The personnel worked to prevent patient falls both during hospitalization and after discharge. They believed that the long-term perspective was much more difficult to address in their clinical practice, and they expressed a need for more homelike environments in the hospital.

Conclusions: The view on technology-based in-hospital fall prevention can be broadened not only to mainly include monitoring and alarm systems, information systems in general, or computer-based information in particular systems but also to support activities performed by health personnel that engage patients in fall prevention. For example, tools such as these can be implemented in training involving daily activities and mobility within safe yet more homelike clinical contexts.

(JMIR Hum Factors 2025;12:e73110) doi:10.2196/73110

XSL•FO
RenderX

KEYWORDS

falls; prevention; hospital; technology; interaction

Introduction

Background

Falls are the most common adverse event reported in hospitals [1,2]. In-hospital falls may be effectively prevented by multifactorial interventions, but this may apply mostly to a subacute setting [3], and the interventions' content can vary widely [4]. Home hazard minimization has been proven to reduce the number of falls among older patients recently discharged from hospitals, particularly if they had a history of recent falls [5]. However, interventions effective in preventing falls among older adults in general are not automatically transferable to this group [5]. So far, research on technology-based fall prevention [6] has mainly focused on systems for older adults living in the community [7] and resulted in few implementations in real-world situations. A systematic review of health technologies, which are defined as "...the application of organised knowledge and skills in the form of devices, medicines, vaccines, procedures and systems developed to solve a health problem and improve the quality of lives" [8], identified that highly heterogenous multicomponent or multifactorial interventions were most common in hospital settings, followed by technical devices to detect falls [9]. A few examples of fall detection technologies have been tested with patients [10-13], but most of them were emerging technologies that had not yet been tested on patients beforehand [9]. Nevertheless, an artificial intelligence–based motion alarm that notifies health personnel of fall risk situations has recently proved promising for in-hospital fall prevention [14]. To enable new complementary fall prevention technologies to evolve, a system perspective on fall prevention is needed.

Prior Work

Previous studies on in-hospital fall prevention from a system perspective include the literature review by Taylor and Hignett [15], which analyzed risk factors and interventions to prevent in-hospital falls based on human factors and ergonomics. This review also presented a model describing the complex interplay between organizations, people, and environments related to in-hospital fall prevention [15]. The importance of human factors and ergonomics in the design of health care environments to support well-being and high performance has also been advocated by Hignett [16] and Hignett and Masud [17]. Through ethnographic observations and interviews of older persons on the subject of orthopedics in hospital wards, Mcvey et al [18] studied interactions between hospital staff and patients who have been hospitalized with cognitive impairments to engage patients in fall prevention. Mcvey et al [18] determined that patients who engaged with staff regarding fall prevention were able to express their needs and, to a certain extent, collaborate with the staff on fall prevention. The positive impact of patient and care team partnerships regarding fall prevention was also observed by Dykes et al [19] in their evaluation of the tailored prevention of inpatient falls tool. This tool, whose usability and patient centeredness have been improved in an iterative user-centered design process [20], clearly illustrates the importance of human factors in fall prevention. Similarly, Bianco et al [21] determined that insufficient dialogue with older adults about their personal goals results in the development of stigmatizing solutions for fall prevention that will not be accepted and implemented by the older users. Older adults' experiences and perceptions of digital technologies for home-based fall prevention have also been investigated in a field study by Ogonowski et al [22]. This study concluded that the self-monitoring of results and self-assessment of fall risks made the system valuable and motivating for the users and that the successful integration of the system into daily life could increase the users' quality of life, technical confidence, and awareness of measures for fall prevention [22].

The system perspective is also essential in human-centered design (HCD), which supports systems design based on human needs and functions from a holistic point of view [23]. HCD has been widely used in health research and innovations [24], such as in systems designed to prevent falls by assessing a person's balance [25] and fall risk [26]. An essential part of HCD is developing activities to understand and specify the context of use, defined as "the combination of users, goals and tasks, resources and environment" [23]. Therefore, holistic studies of humans in their natural settings based on ethnographic principles are common in HCD [27]. Because ethnography is useful for producing detailed descriptions of phenomena, the method has also been suggested as suitable for health care research [28,29] and health care improvements [30]. Focused ethnography is a type of ethnography that studies specific aspects of field work such as interactions and activities; compared to classical ethnography, it typically involves shorter field visits, more extensive data and analysis work, and a greater focus on communicative activities [31]. This approach has been used to study older adults' and their children's perceptions of fall risks [32]. Workshops are another method commonly used in HCD for creating a common arena, enabling participants to discuss issues from different perspectives and areas of competence to develop a common experience within a complex phenomenon [33]. Workshops are widely used in health care innovation [34] and have been used by McKenzie et al [35] to support interprofessional learning to reduce the number of falls experienced by older adults. McKenzie et al [35] reported that their workshop participants believed the workshops provided them with the confidence to implement the fall-preventive practices presented and discussed in the workshops.

Purpose of This Study

The three objectives of this study were as follows: (1) to explore and describe the clinical practices for fall prevention in an acute orthopedic hospital ward from a system perspective; (2) to determine the needs of patients and the possibilities of further supporting clinical practices for fall prevention; and (3) to test whether a framework for studying interactions between people, activities, contexts, and technologies (PACT) can be used to support observations of complex phenomena such as clinical fall prevention.

XSL•FO
RenderX

Methods

The methods used in this study have been reported in accordance with the COREQ (Consolidated Criteria for Reporting Qualitative Studies) guidelines [36].

Study Design

Because this study aimed to provide a detailed description of clinical practices for fall prevention, it had a qualitative research design based on HCD principles [23], signifying that it combined holistic studies of humans in their natural settings via focused ethnography [31] and a workshop [34].

Theoretical Frameworks

PACT Framework

The PACT framework supports the analysis of people who use innovative technologies or tools, the activities that these people are willing to undertake, the contexts in which these activities take place, and the key features of meaningful innovative technologies [37]. PACT analyses have been used to transfer design knowledge from non–health care domains to health care to increase the usefulness of health care information systems [38].

Activities to Prevent Falls

Due to the well-documented need for effective strategies to prevent falls among inpatients during their high-risk hospital-to-home transition [39-41], this study included activities to prevent falls during the patients' hospital stay and the hospitals' efforts to prepare inpatients to take an active role in fall prevention after their discharge. On the basis of the clinical practices and guidelines recommended for Swedish caregivers and health care personnel [42,43], evidence-based methods to prevent in-hospital falls [44], and activities aiming to increase the patients' possibilities to implement fall prevention actions at home [45] after their hospital discharge, this study has defined "activities to prevent falls" as the following:

- Fall risk assessments (related to fall histories, hospital personnel's overall assessment, clinical instruments, and aspects related to the fall prevention activities mentioned subsequently)
- Fall prevention activities (related to medication, cognitive impairment, nutrition, care environment, physical exercise and mobilization, mobility aids, bathroom visits, and personal hygiene)
- Documentation regarding fall risks and fall prevention activities (from previous care providers regarding patients considered to be at risk, risk activities, risk moments, and any fall risk mentioned in a patient's health record)
- Informational activities (regarding the fall risk of patients considered to be at risk as well as handover)
- Activities supporting patients' physical activity and home modifications to prevent falls after hospital discharge

Ethical Considerations

On the basis of the Swedish Act Concerning the Ethical Review of Research Involving Humans, the research in this study (ethnography and workshop) did not require ethical vetting. An application for the ethical review of the ethnographic work was submitted to the Swedish Ethical Review Authority. The written decision from the authority (Dnr 2022-05412-01) confirmed that this research did not require ethical vetting, and the authority stated that it did not see any ethical difficulties related to the research. All the research performed in this study was conducted in accordance with good research practices. The persons interested in participating in the clinical observations and on-site interviews received oral and written information about the study, and informed written consent was collected from all participants before their participation. The workshop participants received oral and written information through a Microsoft PowerPoint presentation before completing the questionnaire and provided written consent on the questionnaire to participate in the research.

Research Setting

This study was performed in an orthopedic clinic at a hospital located in a medium-sized Swedish city with approximately 150,000 inhabitants. The clinic performed pre and postoperative rehabilitation and outpatient and inpatient care, including elective and acute surgical operations. The ethnography was conducted at the clinic's acute hospital ward, and the user workshop was held during a seminar open to staff from the entire clinic. The study was performed by ME, a researcher in medical and health technology with expertise in user-centered development. ÅR, a researcher in physiotherapy also working at the orthopedic clinic, and HT, a human-computer interaction researcher, participated in peer debriefing. Both had clinical experience as physiotherapists. All 3 researchers are female.

Participants and Recruitment

Participants in Clinical Observations and On-Site Interviews

The 8 participants came from the approximately 100 health care workers in the acute ward (Table 1). They had been recruited according to principles of purposive sampling, combining criterion-based (personnel active in fall prevention within a specific clinical setting) and stratified purposeful (personnel representing a variety of staff categories) sampling strategies [46]. The participants represented the following staff categories: assistant nurse, clinical pharmacist, nurse, occupational therapist, physician, and physiotherapist. One of the nurses was a discharge coordinator. The study aimed to recruit at least 1 person per staff category and cover both morning and afternoon shifts.

XSL•FO
RenderX

Table 1. Data collection events and participants.

Events specified by consecutive number	Participants	Time
Clinical observations		
1	Physiotherapist and occupational therapist	Morning
2	Assistant nurse	Afternoon
3	Nurse	Afternoon
4	Nurse	Morning
5	Physician	Morning
On-site interviews		
6	Clinical pharmacist	Morning
7	Discharge coordinator	Morning
Workshop		
8	Representatives of several health care occupations (refer to the Workshop Participants subsection)	Lunch

To recruit participants, ME presented the study at staff meetings and on the clinic's intranet. Thereafter, she booked appointments with 5 interested staff members by email and found 3 interested staff members while visiting the ward to see if any health personnel would volunteer to be observed.

Workshop Participants

The participants were clinical staff members who attended a research seminar at the orthopedic clinic and filled in a questionnaire. The participants represented the following health occupations: assistant nurse, health counselor, medical secretary, nurse, occupational therapist, physician, and physical therapist.

Data Collection and Curation

Overview

An overview of the data collection events and their respective participants is presented in Table 1. The methods used in each event are thoroughly described subsequently. The clinical observations and on-site interviews were performed from February to May 2023, and the workshop was held in November 2023.

On-Site Observations and Interviews

On-site observations of health personnel were performed on 5 different occasions (Table 1). Each time, a researcher followed 1 or 2 participants for 2 to 4 hours to observe them during routine patient care. Apart from the participants' health care occupation, no personal information was collected from any of the persons observed. Instead, general types of attributes relevant to the persons' involvement in fall prevention activities were observed along with any attributes the health care personnel discussed and considered in their work with patients. The observations were documented via written notes in a structured protocol that recorded the activities observed to prevent falls, any other persons involved, and the context in which the activities took place (Multimedia Appendix 1). At the end of each observation, the researcher asked the participant a few questions regarding whether something unusual had happened or the person had done anything more or less than usual during the shift; whether the person had used any specific templates or forms to provide or receive a report before, during, or after the shift; and whether the person would like to correct or add something to the researchers' observational notes by following a brief interview guide (Multimedia Appendix 1). Of the 8 health care personnel, 2 (25 %) who worked in an office participated in unstructured on-site interviews for 60 to 90 minutes instead of the observations (Table 1). These interviews focused on the same aspects as the observations. All the interviews were documented through written notes. Audio recordings were avoided to help participants feel more at ease. Before the first observation, the researcher examined an empty patient room and used the observation protocol to document what they saw via photos and written notes. After each observation (the same or following day), the researcher wrote down their reflections, ideas, and questions in notes, which were later transferred to the observation protocol. Thereafter, the researcher prepared a text integrating the observational data and their own reflections (based on the protocol), the interview data, and any additional information that they recalled during the writing process.

Workshop

Data were collected during a research seminar for the entire orthopedic clinic's staff. The seminar included presentations of 3 projects, of which 1 was the ethnographic work from this study The presentation of our ethnographic work provided a brief background on in-hospital fall prevention and current research regarding supporting technologies; the study's approach, aim, expected results, and methodology; how the researcher defined the "activities used to prevent falls" in the study; the potential benefits of staff participating in the study; and the clinical observations. Next, the researcher presented the preliminary results regarding the observed activities that the respective health professions implemented to prevent falls, as well as the tools used during these activities. Finally, the researcher introduced the workshop approach, which included group discussions on the 3 questions related to the staff's views on the definition of "activities to prevent falls" and how these activities contribute to preventing falls both in the short and

XSL•FO
RenderX

long term; whether adjustments were needed in the presentation of in-hospital activities to prevent falls as well as the tools supporting the activities; and what may be needed to further develop any aspect of the activities, context, or tools to further improve the work to prevent falls (Multimedia Appendix 2). Participants were also invited to contribute their own responses to the research by providing written consent and submitting the completed forms to the researcher.

Data Analysis

On-Site Observations and Interviews

A qualitative content analysis was performed based on principles described by Elo and Kyngas [47] using Microsoft Excel as the software tool for the categorizations. All descriptive texts integrating the observational and interview data, as well as the researcher's own reflections for each individual observation and on-site interview, were first analyzed separately in an unconstrained categorization matrix based on PACT [37]. In this adaptation, the category "tools" was used in place of "technology," as originally defined in the PACT model. ME, who was familiar with the material from data collection and curation, read the running text and coded phrases based on their content as people, activities, contexts, or tools. Data categorized as people and context were further categorized into health personnel; patients; external people; and organizational, social, and physical environment, respectively. Data from the people and context subcategories, as well as the tools and activities categories, were analyzed following principles of inductive content analysis [47]. Thereafter, the main and subcharacteristics of the matrix elements related to people, contexts, and tools from all observations and interviews, as well as from the activities of all health care occupations, were reviewed and integrated by ME, who subsequently wrote new descriptive texts for all characteristics in the individual matrix elements. ÅR and HT, who had not been involved in data collection or curation, participated in peer debriefing, during which they gave their views on the descriptive texts, coding, and categorizations, as well as the overall analysis process [32,48]. The activities from each health care personnel category were visualized, and information on the type of tool used in specific activities was added to the visualizations. In some cases, the titles and structures of the categorization's main and subgroups were modified during the writing process.

Workshop

Free-text answers from the groups on the open-ended questionnaire were summarized in a table. The answers were read and discussed by all the authors.

Strategies Used in the Study to Achieve Trustworthiness

Overview

The strategies this study applied to achieve trustworthiness are described subsequently and based on the parallel criteria for trustworthiness in qualitative research developed by Lincoln and Guba [49], who recommend applying strategies to qualitative research that better achieve their 4 criteria for trustworthiness (credibility, transferability, dependability, and confirmability). Application of the 4 criteria by Lincoln and Guba [49] in qualitative research has been reported by other researchers, for example, by Enworo [50] and Forero et al [51].

Credibility

This study applied the following strategies to achieve credibility: *prolonged and varied engagement with the setting*, by ME visiting the clinical environment on several occasions over several months during different activities; *in-method triangulation*, by integrating observations, interviews, and the workshop; *data source triangulation*, by involving personnel representing different staff categories within different settings at the orthopedic clinic; *peer review*, with ÅR and HT reviewing data, analysis process, and the result; and *member check*, which was performed through the brief interviews at the end of the observations and the workshop, during which the study findings were presented and discussed with a broader group of health personnel.

Dependability

To achieve dependability, this study applied the following 2 strategies: *rich description of study methods*, as reflected in the methods description guided by the COREQ checklist [36], and *member check* (refer to the description mentioned earlier under the Credibility subsection).

Confirmability

To achieve confirmability, this study applied 2 strategies: *reflexivity*, via ME's self-reflections after each observation and on-site interview, the researchers' common reflections in the peer review meetings, and the workshop participants' reflections on the study results presented; and *triangulation* (refer to the description mentioned earlier under the Credibility subsection).

Transferability

To achieve transferability, this study applied the 2 strategies: *purposeful sampling*, by combining criterion-based (personal active in fall prevention in a specific clinical setting) and stratified purposeful (personal representing a variety of staff categories) sampling strategies, and *thick description*, in which the integration of data collected from several sources (observations, interviews, reflections, and the workshop) on several different occasions contributed to a rich and detailed description of in-hospital fall prevention.

Results

Clinical Observations and On-Site Interviews

People

The people involved in activities to prevent falls were health personnel, patients, and external persons. Main and subcategories are presented in Multimedia Appendix 3. The clinical practices implemented by individual members of the *health personnel* depended on which health occupation they were licensed in, their amount of clinical experience, whether they held specific positions in management or coordination, and their individual competencies. Moreover, the health personnel customized fall prevention activities for the individual *patients*, who were all adults of varied age. The individualization

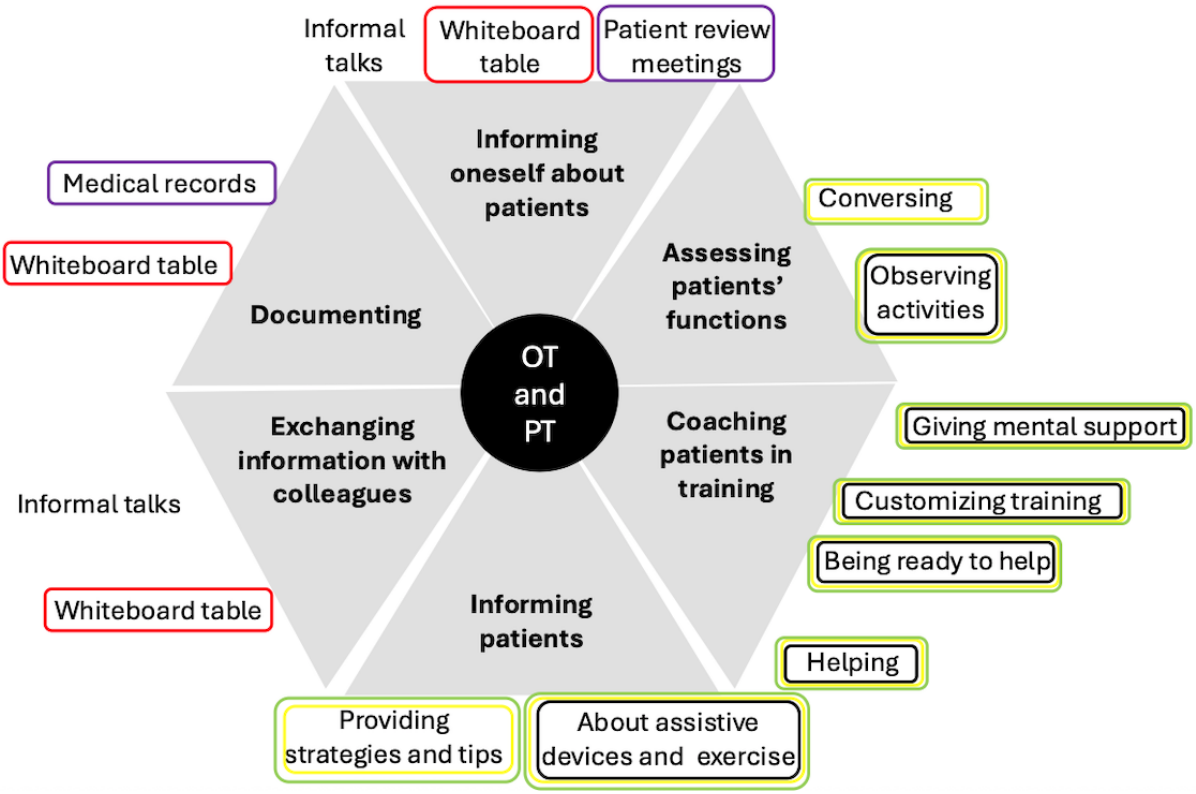
XSL•FO
RenderX

was based on the patient's overall health status, which could include both the reason that the patient had sought acute orthopedic care and any other potential additional health problems. The patient's personal factors, such as communication, life situation, self-awareness, and preferences, also needed to be taken into account. In some cases, the patients and health personnel also interacted with *external persons* in fall prevention. This could include health personnel from other clinical units in patient transportation, medical consultations, and referrals to primary care, as well as municipal staff in planning continued care along with the patients' family members.

Activities and Tools

Fall prevention activities and supporting tools are presented by staff category in Figures 1-6 and in the descriptive texts subsequently.

Figure 1. The clinical practices for fall prevention performed by physical and occupational therapists, with the type of activity (main categories) presented in bold text in gray fields. Specific activities (associated subcategories) are presented in normal font text outside their respective gray field. The supporting tools are represented by circles around the activities. The color of the circle indicates the type of tool used: purple represents computer-based systems, red represents physical tools and systems, green represents the staff, black represents mechanical aids and assistive devices, and yellow represents physical barriers and mechanical tools. OT: occupational therapist; PT: physical therapist.



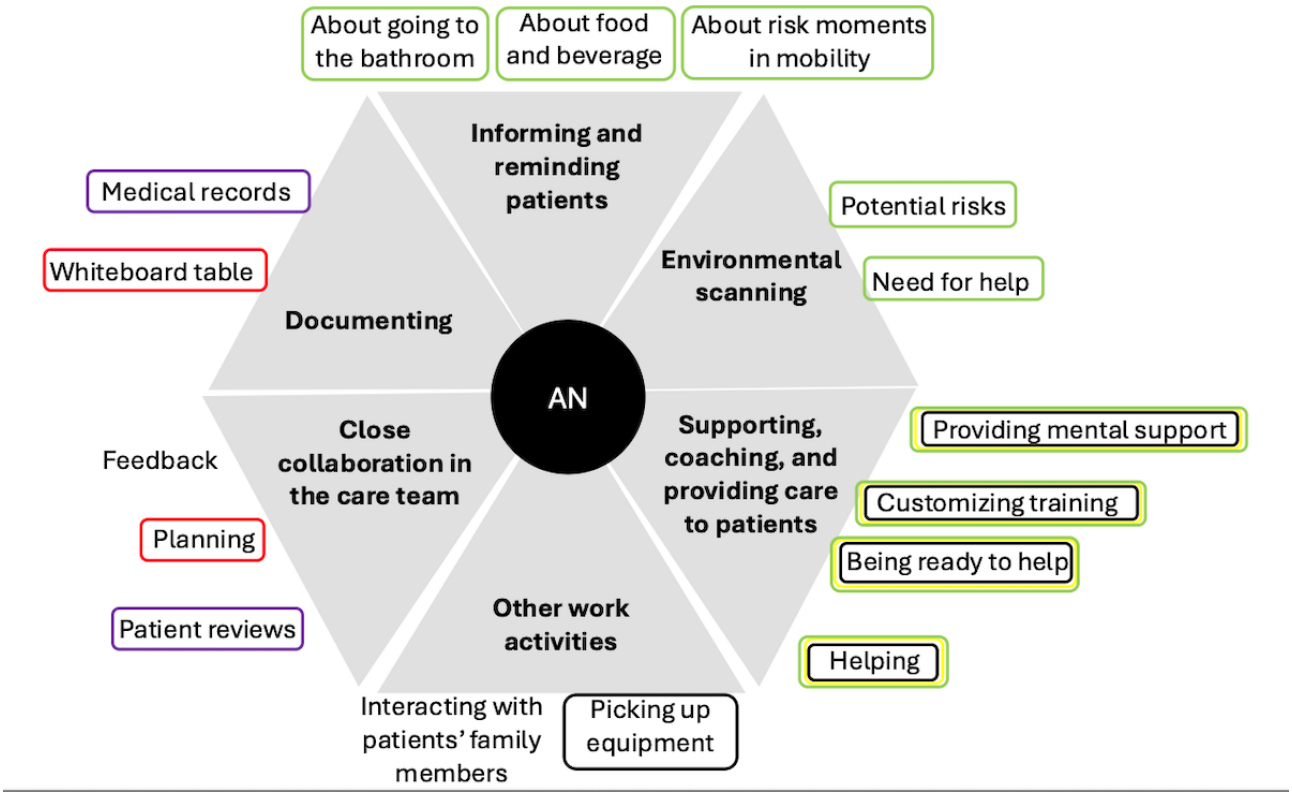
The physical and occupational therapists mainly supported patients in fall prevention during mobility training and activities in daily life (ADLs; Figure 1).

The therapists prioritized meeting patients who were new or close to being discharged. Their meetings focused on evaluating and supporting the patients as they underwent training in tasks essential for ADL. To initiate this, the therapists asked the patient to perform basic tasks, such as sitting up in bed. As the patient accomplished tasks, the therapists gradually introduced the patient to more challenging tasks, such as rising, standing, walking with mobility aids, and performing ADLs. Although the therapists encouraged the patients to try these tasks independently, they were constantly ready to support them immediately if needed. The therapists assessed the patient's functions and well-being by observing how the patient seemed to feel (eg, if they were affected by pain) and by listening to what the patient said (eg, regarding well-being, pain, mobility, their living situation, and assistive devices). They provided the patient with coaching and mental support, calming and encouraging them, based on the patient's individual needs. As they customized the training to the patient's current abilities, they prevented the patient from experiencing "failures." Tasks that seemed too challenging in training on 1 occasion could be postponed until an up-coming meeting or be supported by assistant nurses later the same day. The therapists gave advice related to safe mobility habits, ADL, and body positioning; provided practical help to patients; and eliminated safety risks in the physical environment near the patient.

Assistant nurses provided close care and support to patients, such as when they got dressed, visited the bathroom, and engaged in mobility training (Figure 2).

XSL•FO
RenderX

Figure 2. The clinical practices for fall prevention performed by assistant nurses, with the type of activity (main categories) presented by bold text in gray fields. Specific activities (associated subcategories) are presented in normal font text outside the respective gray field. The supporting tools are represented by circles around the activities. The color of the circle indicates the type of tool used: purple represents computer-based systems, red represents physical tools and systems, green represents the staff, black represents mechanical aids and assistive devices, and yellow represents physical barriers and mechanical tools. AN: assistant nurse.

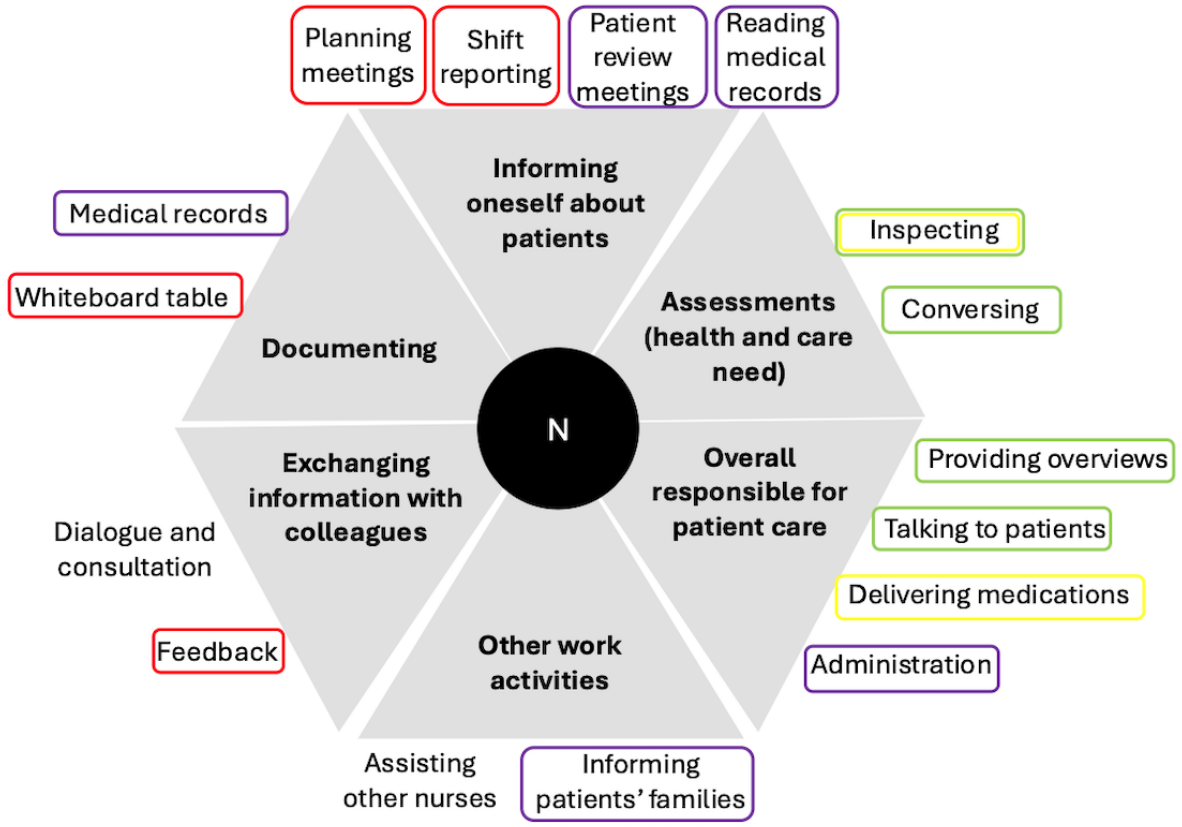


The assistant nurses moved along the corridor, scanning for patients in need of support, such as helping them go to the bathroom, and performing scheduled visits to see patients and to take food orders. They assisted patients with food orders by conversing, checking their documentation, and ensuring access to beverages. They also scanned for potential safety risks in the patient rooms and modified their physical environment by removing tripping hazards; helping patients to vary body position when needed; and reminding the patients about food, beverages, and bathroom visits, etc. They also supported patients in mobility training by reminding them, for example, of how to safely position their feet or lock their walker's wheels. While interacting with the patients, the assistant nurses were attentive, responsive, encouraging, calm, and clear and always one step ahead of the patient while focusing on safety. The assistant nurses collaborated closely with nurses, assisted colleagues, and interacted with the patients' family members at the ward.

Nurses were responsible for the care of a small group of patients, and there were approximately 4 to 6 patients per nurse. Their work included planning follow-up care, which in many cases involved prevention (Figure 3).

XSL•FO
RenderX

Figure 3. The clinical practices for fall prevention performed by nurses, with type of activity (main categories) presented by bold text in gray fields. Specific activities (associated subcategories) are presented in normal font text outside the respective gray field. The supporting tools are represented by circles around the activities. The color of the circle indicates the type of tool used: purple represents computer-based systems, red represents physical tools and systems, green represents the staff, black represents mechanical aids and assistive devices, and yellow represents physical barriers and mechanical tools. N: nurse.

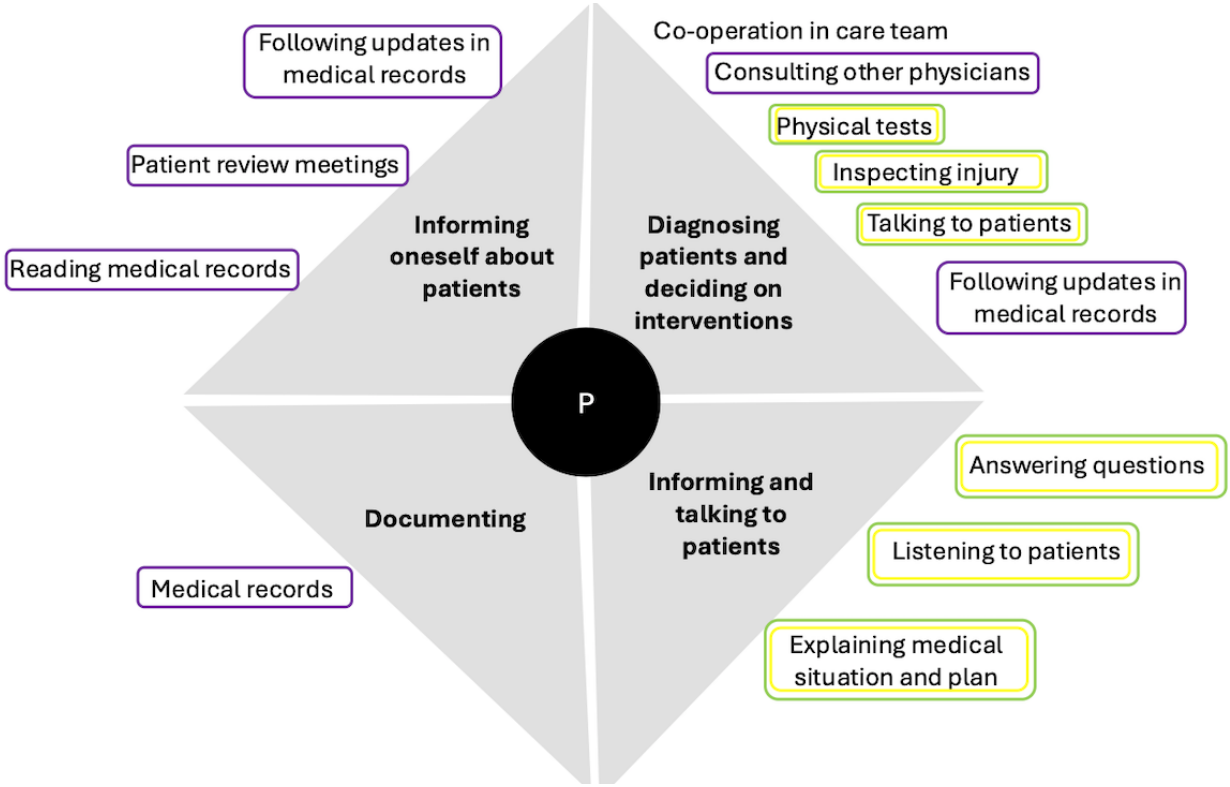


The nurses consistently kept themselves well informed of their patients' history and status through the patients' medical records, their own written notes from previous shifts, conversations with staff members from the outgoing and current shifts, and talking to patients during care activities and patient review meetings. Moreover, they continuously informed the care team's staff and interacted with the care team, physicians, and care coordinators through discussions, consultations, and information exchange. Together with the care team, they planned their activities and documented them on a whiteboard table and in their own written notes. Their interactions with assistant nurses, who carried out most of the point-of-care activities, were crucial and included continuous 2-way reporting performed orally and on the whiteboard table. The nurses documented all care activities provided in the patients' medical records, including notes from patient review meetings when necessary. They performed structured patient rounds to introduce themselves, check the patients' health status, and deliver medications and blood transfusions to patients. They also administered and followed up on medical and care activities, assisted colleagues, and talked to the patients' family members.

Physicians were responsible for the patients' medical care, including providing medication that could be closely related to fall prevention (Figure 4).

XSL•FO
RenderX

Figure 4. The clinical practices for fall prevention performed by physicians, with the type of activity (main categories) presented by bold text in gray fields. Specific activities (associated subcategories) are presented in normal font text outside the respective gray field. The supporting tools are represented by circles around the activities. The color of the circle indicates the type of tool used: purple represents computer-based systems, red represents physical tools and systems, green represents the staff, black represents mechanical aids and assistive devices, and yellow represents physical barriers and mechanical tools. P: physician.

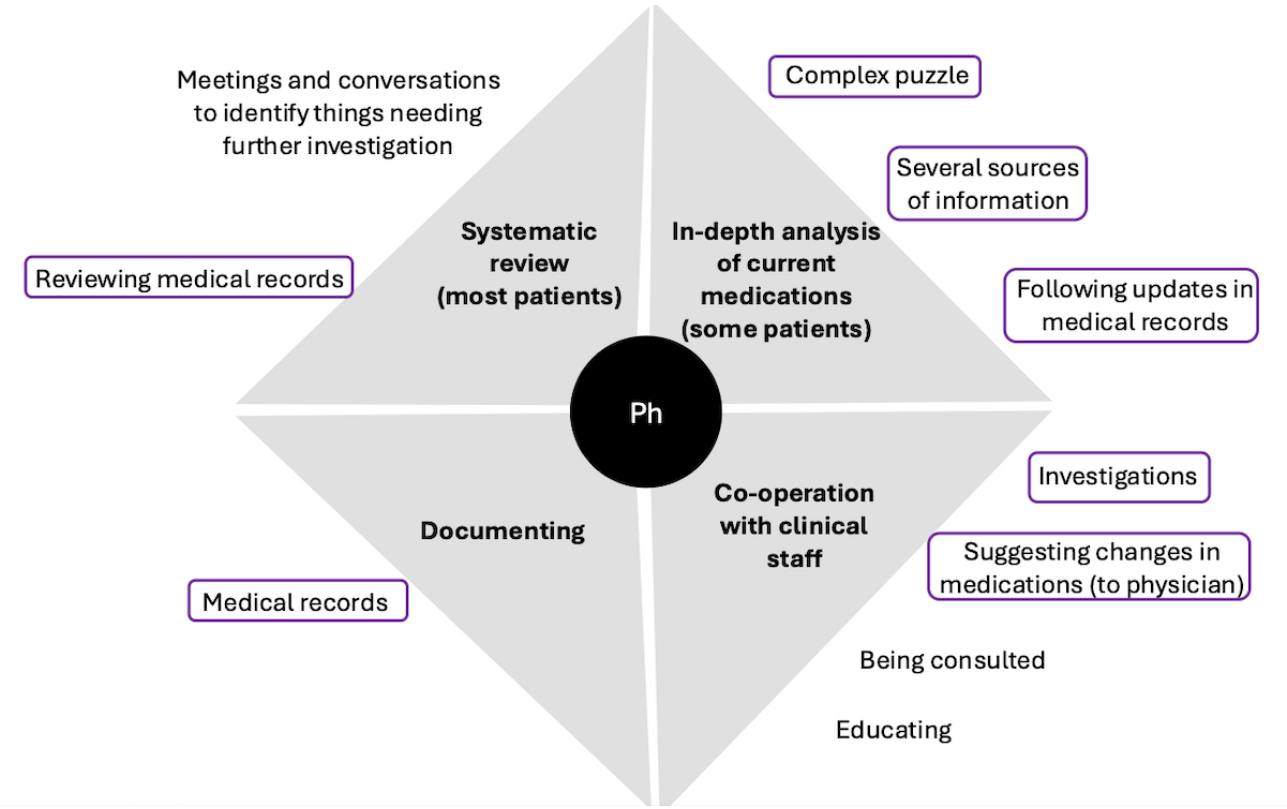


The ward physician was available in the ward during the daytime. They had support from a more senior colleague who, after having attended patient review meetings with patient care teams and participated in the clinical rounds at the ward, was accessible over phone. The ward physician consistently stayed well informed by reading medical records, leading the patient review meetings, and interacting with all the patients during their clinical rounds. In the patient review meetings, the care team informed the physicians about events from the past shift. In some cases, the ward physician reported on consultations with specialists in related domains. On the basis of reports, discussions, and consultations, the senior physician decided on the medical interventions, which the ward physician documented in the medical records. Thereafter, the physicians met all the patients in their rooms during their daily clinical rounds. Here, they assessed the patients' state of well-being and cognition based on conversations with them and their muscular and neural functions based on physical examinations, which included observations of the patients' responses to specific stimuli, visual inspections of any bandages, and subcutaneous bleeding. The specialist physician informed the patients of the results of their examinations, their medications, and any plans for ongoing evaluations and interventions. The physicians demonstrated where injuries were located by pointing at their own body. Before their discharge, patients received a written summary of the examinations and interventions they underwent, the results, and the plan for continued care. The physicians listened to the patients' experiences, answered questions, and asked them if they felt comfortable with the plans. In their conversations with patients, the physicians appeared clear and calm, and the patients showed them confidence and respect.

The clinical pharmacist supported the physicians and other health personnel by answering questions related to pharmacology, including those related to fall risks (Figure 5).

XSL•FO
RenderX

Figure 5. The clinical practices for fall prevention performed by the clinical pharmacist, with type of activities (main categories) presented by bold text in gray fields. Specific activities (associated subcategories) are presented in normal font text outside the respective gray field. The supporting tools are represented by circles around the activities. The color of the circle indicates the type of tool used: purple represents computer-based systems. Ph: clinical pharmacist.

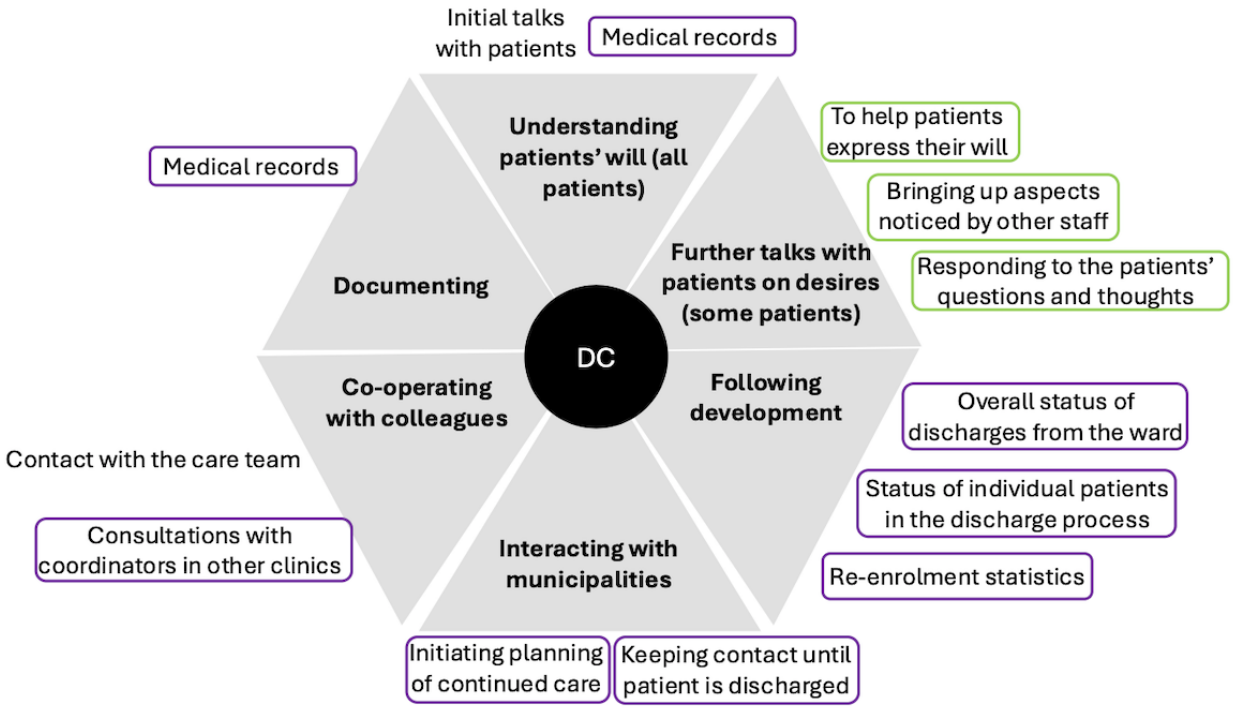


The pharmacist systematically reviewed the drugs prescribed for most patients to identify any need to change them based on the information from the medical records, a national system for prescribed medications, patient review meetings with physicians and the patient care team, and consultations with the discharge coordinators. They screened for potential drug-related medical issues by checking the patients' medication lists and physiological parameters in the medical records. They also made thorough reviews of drug prescriptions for patients at risk of having drug-related health problems, such as patients with polypharmacy issues. These reviews were complex puzzles balancing combinations of risks and benefits. They focused on medications with documented issues affecting fall risks, either by themselves or in combination with other drugs, particularly among older patients. This may include prescribed combinations of medications that may cause drowsiness and fatigue, drug prescriptions deviating from general recommendations or those lacking evaluation, prescribed doses of antihypertensive drugs high enough to cause dizziness, and diuretics prescribed to be taken at certain times that could increase the risk of nocturnal bathroom visits. The analysis was done by manually comparing complex information from separate systems on the computer screen. The pharmacist documented suggested changes in drug prescriptions in the patients' medical records for the physicians to consider. The physicians then decided whether and at what point a change should be implemented. Changes more suitable to implement after the patients' hospital discharge were referred to physicians in primary care.

The discharge coordinators supported the patients in their discharge and safe transition home by addressing fall risks and considering the patients' desires (Figure 6).

XSL•FO
RenderX

Figure 6. The clinical practices for fall prevention performed by discharge coordinators, with type of activities (main categories) presented by bold text in gray fields. Specific activities (associated subcategories) are presented in normal font text outside the respective gray field. The supporting tools are represented by circles around the activities. The color of the circle indicates the type of tool used: purple represents computer-based systems and green represents the staff. DC: discharge coordinator.



The discharge coordinators gathered information on the patients' desires by reading their medical records and talking to them. Some patients were able to clearly describe their living and occupational situations and express their requirements for discharge in a single conversation. Others needed more time to process what type of help they wanted or needed after their discharge. Therefore, the coordinator talked to them again to guide them in their decision-making process by asking them concrete questions about their daily life (eg, home environment, ADL capacity, etc) and by describing the possibilities of home-based help for patients interested in such support. Because the coordinators had a professional role different from that of the care team, they could spend more time talking to the patients about things they brought up or about needs that other staff members had noticed. The coordinators cooperated closely with the care team staff, listened to their assessments, and reminded them to carefully document the patients' state of well-being and activities in their health records. They also cooperated with municipal staff to plan for the continued care of patients who had given their consent and patients who needed such care but were unable to give their consent. The coordinators conveyed the patients' desires via telephone and a computerized system to the municipal care administrators, who decided on which activities the patient were entitled to. The coordinators also communicated with the patients' family members, who sometimes contacted the clinic to emphasize their relative's needs and were then reminded to talk to their relative about his or her desires. Although the coordinators tracked 30-day readmission statistics, their involvement ended once the patient was discharged.

The tools used in the wide range of activities described earlier to prevent falls (Figures 1-6) can be categorized into the 5 main categories:

- Computer-based systems, for example, medical records software (purple in figures)
- Physical tools and systems, for example, the care teams' whiteboard table (red)
- The staff, for example, their sight and hearing while assessing patients (green)
- Mechanical aids and assistive devices, for example, walkers for mobility (black)
- Physical barriers and mechanical tools, for example, sliding mats when lifting (yellow)

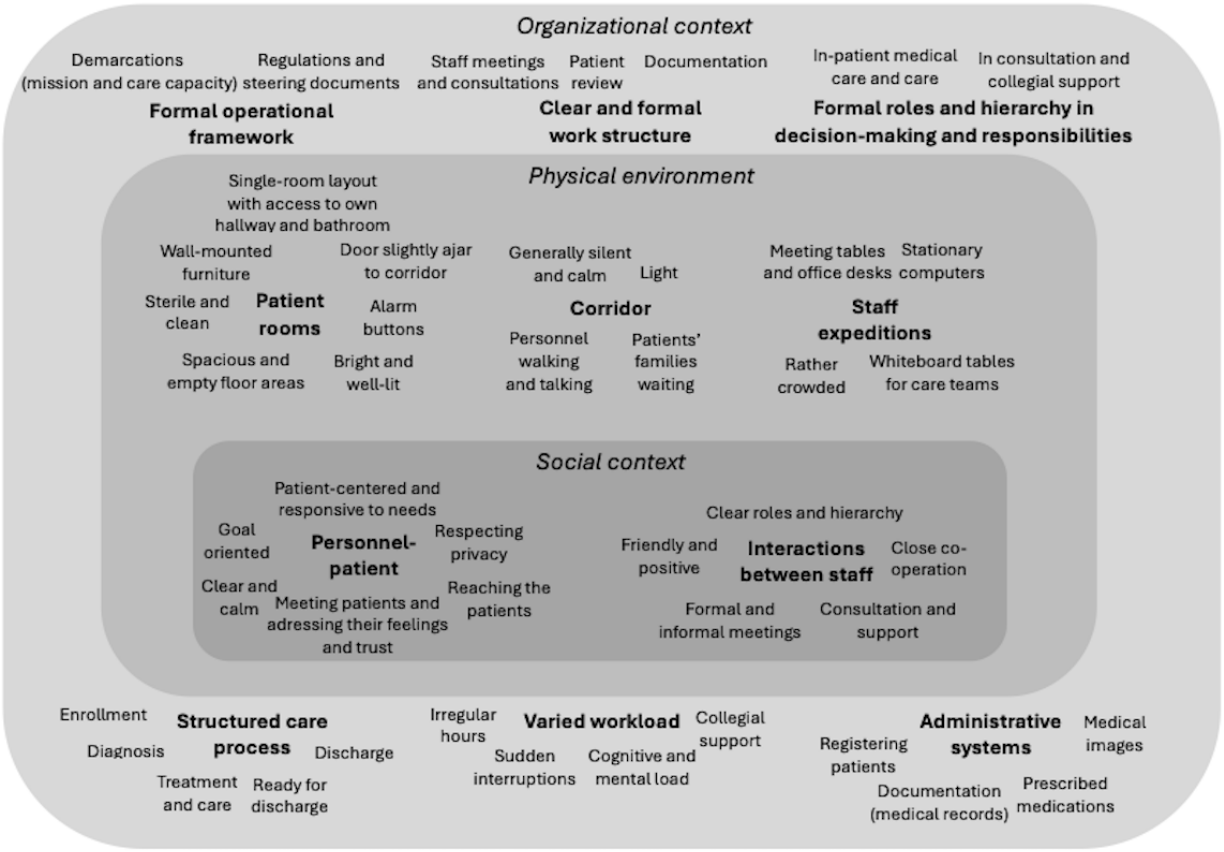
These are all summarized in Multimedia Appendix 4.

Context

Within the PACT framework, fall prevention is described in terms of the physical environment as well as the social and organizational contexts. The main and subcategories are presented in Multimedia Appendix 5 and Figure 7.

XSL•FO
RenderX

Figure 7. A description of the context, in which the main categories (in bold) and associated subcategories (nonbold) are summarized for the organizational context (outer circle area), physical environment (middle circle area), and social context (inner circle area).



The organizational context of fall prevention was highly structured in the work itself and the patient care process, and it was formal in the operational framework and the health personnel's roles and responsibilities. Other essential elements within the organizational context included a varied workload, which could be increased by mental and cognitive loads, sudden interruptions, and irregular work hours. The workload could also be reduced through collegial support and several administrative computer-based systems that the personnel used regularly.

The physical environment for fall prevention included the location where the health personnel met patients, which was primarily the patients' rooms; the location where the patient care teams met, which was primarily the office where the physicians did administrative work and held teaching meetings with the care team staff; and the office where nurses, licensed practical nurses, as well as physiotherapists and occupational therapists did administrative work, planning, and reporting. Both offices had meeting tables and office desks with desktop computers, but the physicians' office had a wall-mounted screen to project the physician's computer display during the teaching meetings. The care team's office had a large whiteboard on which the status and planned activities for all patients were documented. The location dedicated to the care teams (which included the patient rooms, day room, and health personnel team offices) was connected to long corridors, which were brightly lit, mostly silent, and calm. In the corridor, the health personnel could be seen walking on their way to see a patient or a colleague, occasionally pausing to have a quick word with a colleague. Occasionally, a patient walked along the corridor, and sometimes, the patients' families waited in the corridor outside a patient's room. Patient rooms were on 1 side of the corridor, with most of the doors opened slightly. Most patient rooms were single rooms with a small hallway leading into a spacious and bright bedroom with double windows and large open spaces. The walls and ceiling were painted white. The floor was light gray and free from raised sills and thresholds. The patients' rooms were well lit by several light sources on the ceiling and on the wall over a height-adjustable bed, which could be raised or lowered via a panel on 1 side. The bed also had height-adjustable safety rails on the sides. A wired patient alarm control unit was placed beside the bed, mostly near the patient on a mobile bedside table whose tabletop could be placed over the bed. The bedrooms also contained a basin along with disinfectants, a chair, and a large television, all of which were wall mounted. All the patient rooms had their own spacious bathrooms with a shower area, and they were equipped with several supportive grips placed at strategic places for safely sitting down and rising from the toilet seat and moving to and standing in the shower. Alarm buttons were placed near the toilet seat, door entrance, etc. A movable laundry basket was the only item set on the floor.

The social context for fall prevention included settings for the health personnel's interactions with patients, colleagues, and outside persons. The health personnel interacted with patients in different constellations, mostly when 1 or 2 members of the

XSL•FO
RenderX

health personnel visited the patient's room. The patient-staff interactions were patient centered, goal oriented, and empathetic. In sensitive situations, the personnel closed the door to the corridor to protect the patient's privacy. To reach out to the patients, the personnel needed to be perceptive, flexible, and prepared to handle difficult topics brought up by patients in a professional way. The personnel also had to act in a calm, clear, direct, and pedagogic manner to enable the patients to understand and assimilate information, as well as to instill trust and confidence. The patients listened to the personnel and occasionally expressed their feelings toward them, including trust and confidence. Interactions between health personnel took place between colleagues in the patient care team or between the patient care team and staff in specific positions. Here, their roles and responsibilities were noticeably clear. The health personnel consulted each other in both formal and informal meetings within a friendly and positive atmosphere. On the basis of their distinct roles, the health personnel could also complement each other when in contact with a patient. The personnel's interactions with external persons were also patient centered. The personnel stated that they sometimes needed to remind the patients' family members that care planning must be based on the patient's own preferences. Moreover, they stated that they had to respect and confide in the division of responsibilities between the clinic and municipality.

The Staff's View on Fall Prevention, Descriptions of Observed Activities and Tools, and the Need to Further Support Fall Prevention

A larger number of staff members from the orthopedic clinic provided their views on the analysis of observed activities to prevent falls and the tools used to support these activities, which were presented in the Activities and Tools subsection earlier. The staff provided feedback in groups representing specific health staff categories, which included assistant nurse, health counselor, medical secretary, nurse, occupational therapist, physician, and physical therapist (Table 2). As can be seen in Table 2, the staff supported the broad definition of activities to prevent falls and said that they considered fall prevention in all interactions with patients. They also stated that long-term fall prevention was more difficult due to short hospital visits and their lack of influence over the home environment. They agreed on the presentation of activities to prevent falls and the tools used. The staff described the need to further support fall prevention, such as by using more practical "real-life tools" and "making the ward's physical environment more similar to the home environment."

Table 2. Health personnel's views on fall prevention, descriptions of observed activities and tools used to prevent falls, and identified needs for further support.

Group	View on the definition of "activities to prevent falls" as activities contributing to preventing falls both in the short and long term	View on whether adjustments are needed in the presentation of in-hospital activities to prevent falls and the tools supporting these activities	View on the need to further develop any part of the activities, context, or tools to further improve the work to prevent falls
Group 1 (occupational therapist, physiotherapist, and physician)	"It is possible to prevent falls in the short term at the hospital but more difficult to prevent falls in the long term because hospital stays are short and we cannot control the home environment."	"No."	"Make the hospital environment more like the home environment by incorporating features such as doorsteps, carpets, narrow spaces, and doors without automatic opening mechanisms."
Group 2 (assistant nurse and nurse)	"We agree."	"No."	"No."
Group 3 (medical secretary and nurse)	"Good, important work"	"Feels like you got it all."	Continue to develop more practical, real-life tools and access to the latest news
Group 4 (health counselor, physiotherapist, and occupational therapist)	"We consider fall prevention in all interactions with patients at the ward and automatically assess fall risks in everything we do together with the patients."	"No."	—[a]

[a]Not available.

Discussion

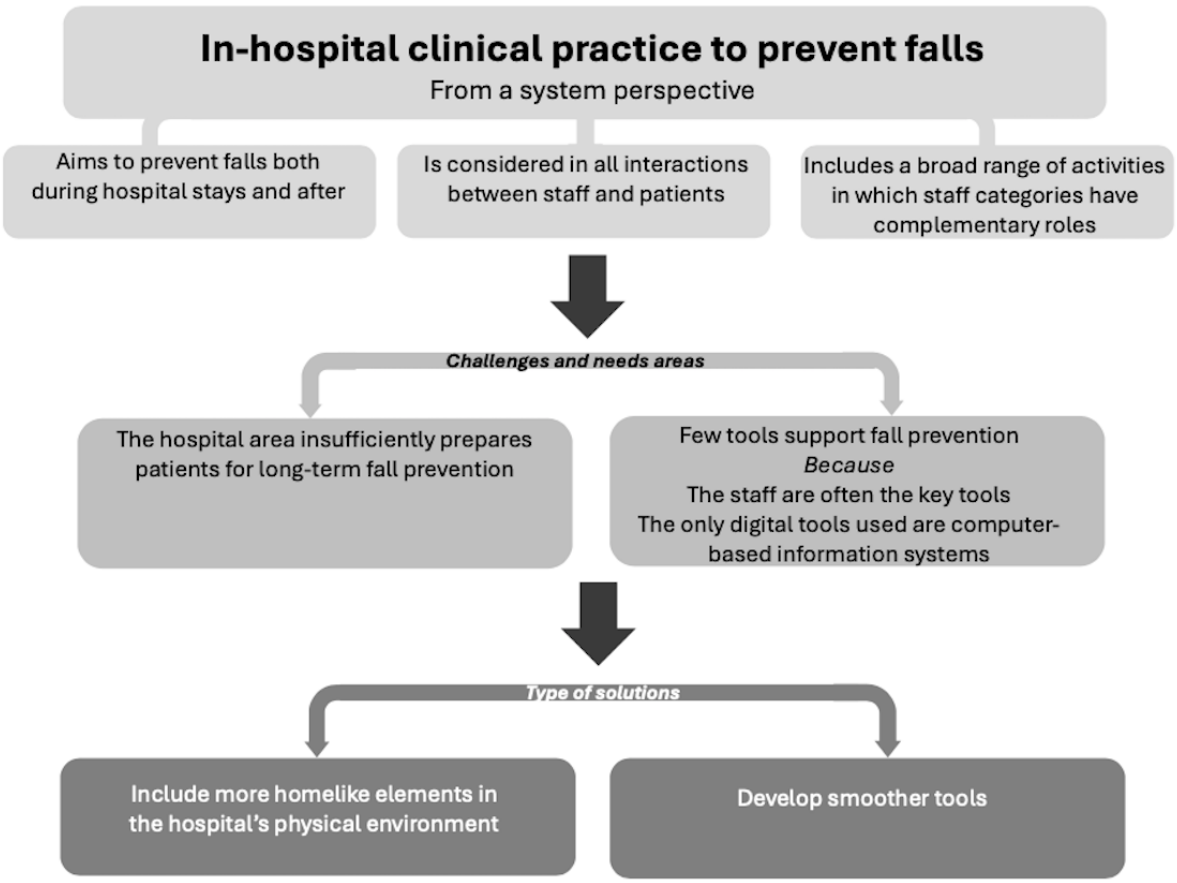
Principal Findings

The principal findings of this study (summarized in Figure 8) showed that the health staff had a holistic view of fall prevention, which was recognized in the clinical observations of the people, activities, context, and tools in a clinical ward. This view was also expressed verbally by the personnel in a workshop. As a result, the study determined that the PACT framework can be useful to support the observation of complex phenomena such as clinical fall prevention. The health personnel stated that they automatically assessed fall risks and considered fall prevention in all their interactions and activities with patients to prevent falls both in the hospital setting and after hospital discharge. It was clearly seen in the observations that the health personnel strived to make the hospital stay safe for the patients and to prepare the patients for a safe transition home (or to a care home). This was observed in the wide range of activities in which the respective health staff categories had complementary roles. For example, physiotherapists and occupational therapists focused on mobility training and ADLs

XSL•FO
RenderX

in their assessments and training with patients, while assistant nurses screened the patient environments for risk factors, supported patients in mobility training, and cooperated with nurses to provide all the patients with effective and safe care. The pharmacist and the physicians reviewed medications prescribed to each patient with the aim of minimizing fall risks, while the discharge coordinators interacted with patients to plan a safe home transition. In the workshop, the health personnel mentioned that the short hospital stays and their own lack of control over the patient's home environment made it more difficult for them to address long-term fall prevention in their clinical practice; therefore, they saw a need for making the hospital environment more homelike by incorporating elements such as doorsteps, carpets, narrow spaces, and manually operated doors. Observations in the wards confirmed that the physical environment was clean, spacious, and free from homelike physical elements that could both increase fall risk as well as enable ADL and mobility training in situations representative of the home setting. The responsive, goal-oriented, and patient-centered interactions between the patients and personnel, which seemed very important for engaging the patients in fall prevention, were also observed in the ward. The observations revealed that although the personnel used different tools in their work and possessed a broad range of skills and functions, they often served as key "tools" themselves in many of their fall prevention interactions by assessing and communicating with and coaching the patients, all while using digital tools (mainly computer-based software) for information retrieval, documentation, and communication. The personnel expressed that they were interested in more practical, real-life tools for fall prevention.

Figure 8. A summary of the main findings of the study, including key aspects of the in-hospital clinical practice to prevent falls studied from a system perspective based on the people, activities, contexts, and technologies framework, along with areas identified for further support and the types of solutions required by staff.



Limitations

This study has several limitations. The method used (ie, collecting clinical practice data relative to fall prevention via observations, interviews, and the workshop without collecting any personal data from the participants except for the health personnel's staff category) limited the results because they encompassed only observable actions, verbal expressions, and contexts related to the people, activities, and tools involved in fall prevention. Consequently, no conclusions were drawn in this study regarding the staff's personal experiences; feelings; mental models; or the psychological, social, and attitudinal attributes related to fall prevention. Moreover, the study explored the health staff's perspectives on fall prevention, and no data were collected regarding the patients' perspective other than what the patients expressed and did in their interactions with the staff.

XSL•FO
RenderX

Although the study applied several strategies (described in the Methods section) to achieve trustworthiness, the study also had methodological limitations. For example, the sampling, performed via criterion-based (personnel involved in fall prevention within a specific clinical setting) and stratified purposeful (personnel representing a variety of staff categories) sampling strategies [46], may have introduced biases into the study and thereby limited the generalizability of the result. It is possible that subgroups within staff categories (other than those described in the Results section) were not represented in the sample due to the researcher's unawareness. For example, it cannot be ruled out that a health staff member's interest in participating in a research study on fall prevention practices may have been affected by their interest in fall prevention. The relatively low number of participants in the study may have contributed further to this risk of sampling bias.

The competence and background of the observing researcher may also have limited the context-specific in-depth understanding of the observations. Although this researcher was knowledgeable in fall prevention due to participating in multidisciplinary research projects, they had no previous training or experience in health care work. Finally, because this study was performed in 1 specific clinical setting, these results may not be generalized to any other in-hospital setting.

Comparison With Prior Work

The holistic view on in-hospital fall prevention practices expressed by the health staff and observed in the clinical context of this study opens up new possibilities for using fall prevention technologies in hospital settings, such as to support staff activities that prepare patients for long-term fall prevention. This view is well in line with previous research confirming the positive impact of patient and care team partnerships to prevent falls and fall injuries and how this can be supported by human-centered tools [19,20]. The level of importance of contextual support to create more realistic "homelike" conditions during mobility and ADL training, which was pointed out by the staff in this study, is different and thus complements the previously studied technology-based in-hospital fall prevention tools and approaches, which aimed to detect early signs of acutely elevated fall risks and falls, such as those described by Cooper et al [9]. This includes chair sensor alarms [10], video- and image-based monitoring [11,12], wearable devices [13], and artificial intelligence–based motion alarms, such as those mentioned by Pagels [14]. The staff's request to make the clinical environment more "homelike" can be seen to reflect more holistic models of fall prevention that involve complex interplays between people, activities, processes, organization, environment, and technology, as seen in the models proposed by Choi et al [52] and Taylor and Hignett [15]. Although the review by Hignett and Masud [17] on the associations between inpatient falls and environmental modifications identified a lack of supporting scientific evidence for most recommendations on how to manage environmental hazards in hospital wards, which included flooring, lighting, and bed alarms, it is likely that making the ward's physical context more "homelike" could introduce new environmental hazards related to falls. As a result, environmental modifications in care settings must be implemented with precaution. Keeping patient rooms intact and introducing homelike elements in other locations outside the rooms is one strategy. Temporarily anchoring mobile, homelike modules in patient rooms offers a more flexible approach, allowing the hospital environment to be tailored to each patient's training needs. These modules could also incorporate embedded digital systems to support and enhance training. One idea would be to develop modules that could be used both in clinical settings and in-home settings, thus allowing patients to continue the training that they were introduced to in the hospital during their rehabilitation. This would allow them to gradually achieve a more active life after their hospital discharge. Information gained from previous research on how to create meaningful and feasible interactions between older adults and home-based digital systems for fall prevention training, such as that from Ogonowski et al [22], for example, can be of value in the design of these modules. It is also essential to build on the knowledge of how to engage in a creative dialogue with patients regarding their personal goals because incorporating these goals into nonstigmatizing designs of solutions to prevent falls is essential, as described by Bianco et al [21]. The dialogue between staff and patients is also essential when the patients are preparing themselves for the challenges posed by their home settings after hospital discharge [53]. For example, opportunities to experience homelike situations under the supervision of the clinical staff can contribute to strengthening the patients' experience with their postinjury bodies and thereby contribute to building self-confidence and self-efficacy in gross motor skills.

Conclusions

The following conclusions can be drawn from this study:

- Expanding the current view of technology-based in-hospital fall prevention to both include computer-based information- and communication systems and digital fall-protection tools aligns well with clinical practice. These technologies will provide support to the health personnel's activities to engage patients in fall prevention.
- Health professionals see a need for further support in their activities so that they can prepare patients who have been hospitalized for the challenges they will face after their hospital discharge. They also suggested that the clinical environment should contain more home-based elements. These could provide the patients with opportunities to experience homelike situations under the supervision of clinical staff and may contribute to strengthening the patients' experience with their postinjury bodies.
- The PACT model can serve as a useful framework for supporting observations, and workshops can be fruitful for enabling health personnel to discuss clinical practices for fall prevention and express their need for further support.

XSL•FO
RenderX

Acknowledgments

The authors would like to thank all study participants and all other personnel at the orthopedic clinic who contributed to the execution of this study. This study was funded by Vinnova (grant number: 2022-01193). Finalization of the manuscript and language review were funded by the multidisciplinary research initiative on human-centered technology at Mälardalen University.

Authors' Contributions

ME (the main researcher) designed the study, collected and analyzed the data, and was one of the main authors of the manuscript. ÅR and HT participated in peer debriefing (including reviewing) during the data analysis, critical review, and revision stages of the study. All the authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Observation protocol and post-observation interview guide.
[DOCX File , 23 KB - humanfactors_v12i1e73110_app1.docx]

Multimedia Appendix 2
Workshop paper material.
[DOCX File , 438 KB - humanfactors_v12i1e73110_app2.docx]

Multimedia Appendix 3
Main categories and subcategories describing the people involved in fall prevention.
[DOCX File , 18 KB - humanfactors_v12i1e73110_app3.docx]

Multimedia Appendix 4
Type of tools used in activities to prevent falls.
[DOCX File , 18 KB - humanfactors_v12i1e73110_app4.docx]

Multimedia Appendix 5
Main categories and subcategories describing the context for fall prevention.
[DOCX File , 20 KB - humanfactors_v12i1e73110_app5.docx]

References

1. Tan KM, Austin B, Shaughnassy M, Higgins C, McDonald M, Mulkerrin EC, et al. Falls in an acute hospital and their relationship to restraint use. Ir J Med Sci 2005 Jul;174(3):28-31. [doi: 10.1007/bf03169144]
2. Weintraub D, Spurlock M. Change in the rate of restraint use and falls on a psychogeriatric inpatient unit: impact of the health care financing administration's new restraint and seclusion standards for hospitals. J Geriatr Psychiatry Neurol 2002;15(2):91-94. [doi: 10.1177/089198870201500207] [Medline: 12083599]
3. Cameron ID, Dyer SM, Panagoda CE, Murray GR, Hill KD, Cumming RG, et al. Interventions for preventing falls in older people in care facilities and hospitals. Cochrane Database Syst Rev 2018 Sep 07;9(9):CD005465 [FREE Full text] [doi: 10.1002/14651858.CD005465.pub4] [Medline: 30191554]
4. Oliver D, Healey F, Haines TP. Preventing falls and fall-related injuries in hospitals. Clin Geriatr Med 2010 Nov;26(4):645-692. [doi: 10.1016/j.cger.2010.06.005] [Medline: 20934615]
5. Naseri C, Haines TP, Etherton-Beer C, McPhail S, Morris ME, Flicker L, et al. Reducing falls in older adults recently discharged from hospital: a systematic review and meta-analysis. Age Ageing 2018 Jul 01;47(4):512-519 [FREE Full text] [doi: 10.1093/ageing/afy043] [Medline: 29584895]
6. Ren L, Peng Y. Research of Fall Detection and Fall Prevention Technologies: A Systematic Review. IEEE Access 2019;7:77702-77722. [doi: 10.1109/ACCESS.2019.2922708]
7. Hamm J, Money AG, Atwal A, Paraskevopoulos I. Fall prevention intervention technologies: a conceptual framework and survey of the state of the art. J Biomed Inform 2016 Feb;59:319-345 [FREE Full text] [doi: 10.1016/j.jbi.2015.12.013] [Medline: 26773345]
8. Health technologies. World Health Organization. 2023 Nov 24. URL: https://www.who.int/europe/news-room/fact-sheets/item/health-technologies [accessed 2025-07-19]

XSL•FO
RenderX

9. Cooper K, Pavlova A, Greig L, Swinton P, Kirkpatrick P, Mitchelhill F, et al. Health technologies for the prevention and detection of falls in adult hospital inpatients: a scoping review. JBI Evid Synth 2021 Oct;19(10):2478-2658. [doi: 10.11124/JBIES-20-00114] [Medline: 34149020]
10. Sahota O, Drummond A, Kendrick D, Grainge MJ, Vass C, Sach T, et al. REFINE (REducing Falls in In-patieNt Elderly) using bed and bedside chair pressure sensors linked to radio-pagers in acute hospital care: a randomised controlled trial. Age Ageing 2014 Mar;43(2):247-253 [FREE Full text] [doi: 10.1093/ageing/aft155] [Medline: 24141253]
11. Coahran M, Hillier LM, Van Bussel L, Black E, Churchyard R, Gutmanis I, et al. Automated fall detection technology in inpatient geriatric psychiatry: nurses' perceptions and lessons learned. Can J Aging 2018 Jul 03;37(3):245-260. [doi: 10.1017/s0714980818000181]
12. Cournan M, Fusco-Gessick B, Wright L. Improving patient safety through video monitoring. Rehabil Nurs J 2018;43(2):111-115. [doi: 10.1002/rnj.308]
13. Clarke HD, Timm VL, Goldberg BR, Hattrup SJ. Preoperative patient education reduces in-hospital falls after total knee arthroplasty. Clin Orthop Relat Res 2012 Jan;470(1):244-249 [FREE Full text] [doi: 10.1007/s11999-011-1951-6] [Medline: 21691907]
14. Pagels S. AI förebyggde fallolyckor på sjukhus. Dagens Medicin. 2023 Sep 21. URL: https://www.dagensmedicin.se/specialistomraden/medicinsk-teknik/ai-forebyggde-fallolyckor-pa-sjukhus/ [accessed 2025-08-12]
15. Taylor E, Hignett S. The SCOPE of hospital falls: a systematic mixed studies review. HERD 2016 Jul 30;9(4):86-109. [doi: 10.1177/1937586716645918] [Medline: 27240563]
16. Hignett S. Why design starts with people. The Health Foundation. 2013 May. URL: https://www.lboro.ac.uk/media/media/schoolanddepartments/design-and-creative-arts/downloads/Hignett-Why-design-starts-with-people-2013.pdf [accessed 2025-08-12]
17. Hignett S, Masud T. A review of environmental hazards associated with in-patient falls. Ergonomics 2006 Feb 20;49(5-6):605-616. [doi: 10.1080/00140130600568949] [Medline: 16717012]
18. Mcvey L, Alvarado N, Zaman H, Healey F, Todd C, Issa B, et al. Interactions that support older inpatients with cognitive impairments to engage with falls prevention in hospitals: an ethnographic study. J Clin Nurs 2024 May 19;33(5):1884-1895. [doi: 10.1111/jocn.17006] [Medline: 38240045]
19. Dykes PC, Burns Z, Adelman J, Benneyan J, Bogaisky M, Carter E, et al. Evaluation of a patient-centered fall-prevention tool kit to reduce falls and injuries: a nonrandomized controlled trial. JAMA Netw Open 2020 Nov 02;3(11):e2025889 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.25889] [Medline: 33201236]
20. Katsulis Z, Ergai A, Leung WY, Schenkel L, Rai A, Adelman J, et al. Iterative user centered design for development of a patient-centered fall prevention toolkit. Appl Ergon 2016 Sep;56:117-126. [doi: 10.1016/j.apergo.2016.03.011] [Medline: 27184319]
21. Bianco ML, Pedell S, Renda G, Kapoor A. HCI methods for empowering discussion on person-centered fall prevention with older adults. In: Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction. 2015 Presented at: OzCHI '15; December 7-10, 2015; Parkville, Australia. [doi: 10.1145/2838739.2838767]
22. Ogonowski C, Aal K, Vaziri D, Rekowski TV, Randall D, Schreiber D, et al. ICT-based fall prevention system for older adults: qualitative results from a long-term field study. ACM Trans Comput Hum Interact 2016 Oct 10;23(5):1-33. [doi: 10.1145/2967102]
23. ISO 9241-210:2019 ergonomics of human-system interaction part 210: human-centred design for interactive systems. International Organization for Standardization. 2019. URL: https://www.iso.org/standard/77520.html [accessed 2025-08-12]
24. Göttgens I, Oertelt-Prigione S. The application of human-centered design approaches in health research and innovation: a narrative review of current practices. JMIR Mhealth Uhealth 2021 Dec 06;9(12):e28102 [FREE Full text] [doi: 10.2196/28102] [Medline: 34874893]
25. Lan Hing Ting K, Dessinger G, Voilmy D. Examining usage to ensure utility: co-design of a tool for fall prevention. IRBM 2020 Oct;41(5):286-293. [doi: 10.1016/j.irbm.2020.03.001]
26. Harte R, Quinlan LR, Glynn L, Rodríguez-Molinero A, Baker PM, Scharf T, et al. Human-centered design study: enhancing the usability of a mobile phone app in an integrated falls risk detection system for use by older adult users. JMIR Mhealth Uhealth 2017 May 30;5(5):e71 [FREE Full text] [doi: 10.2196/mhealth.7046] [Medline: 28559227]
27. Blomberg J, Burrel M. The ethnographic approach to design. In: Sears A, Jacko JA, editors. The Human-Computer Interaction Handbook. Boca Raton, FL: CRC Press; 2007.
28. Savage J. Ethnography and health care. BMJ 2000 Dec 02;321(7273):1400-1402 [FREE Full text] [doi: 10.1136/bmj.321.7273.1400] [Medline: 11099288]
29. Goodson L, Vassar M. An overview of ethnography in healthcare and medical education research. J Educ Eval Health Prof 2011;8:4 [FREE Full text] [doi: 10.3352/jeehp.2011.8.4] [Medline: 21637319]
30. Black GB, van Os S, Machen S, Fulop NJ. Ethnographic research as an evolving method for supporting healthcare improvement skills: a scoping review. BMC Med Res Methodol 2021 Dec 05;21(1):274 [FREE Full text] [doi: 10.1186/s12874-021-01466-9] [Medline: 34865630]
31. Knoblauch H. Focused ethnography. Forum Qual Soc Res 2005;6(3) [FREE Full text] [doi: 10.17169/fqs-6.3.20]

XSL•FO
RenderX

32. Kilian C, Salmoni A, Ward-Griffin C, Kloseck M. Perceiving falls within a family context: a focused ethnographic approach. Can J Aging 2008;27(4):331-345. [doi: 10.3138/cja.27.4.331] [Medline: 19416795]
33. Chamorro-Koc M, Swann L, Haskell N, Dwyer J, Wainwright L, Hosking J. Co-designing design thinking workshops: from observations to quality improvement insights for healthcare innovation. In: Miller E, Winter A, Chari S, editors. How Designers Are Transforming Healthcare. Singapore, Singapore: Springer; 2024.
34. Ozkaynak M, Sircar CM, Frye O, Valdez RS. A systematic review of design workshops for health information technologies. Informatics 2021 May 14;8(2):34. [doi: 10.3390/informatics8020034]
35. McKenzie G, Lasater K, Delander GE, Neal MB, Morgove M, Eckstrom E. Falls prevention education: interprofessional training to enhance collaborative practice. Gerontol Geriatr Educ 2017 Feb 17;38(2):232-243. [doi: 10.1080/02701960.2015.1127809] [Medline: 26886245]
36. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
37. Benyon D. Spaces of Interaction, Places for Experience. Cham, Switzerland: Springer; 2014.
38. Morana S, Dehling T, Reuter-Oppermann M, Sunyaev A. User assistance for health care information systems. In: Proceedings of the 2017 SIG-Health Pre-ICIS Workshop. 2017 Presented at: SIG-Health Pre-ICIS Workshop; December 10, 2017; Seoul, South Korea.
39. Hill AM, Hoffmann T, McPhail S, Beer C, Hill KD, Oliver D, et al. Evaluation of the sustained effect of inpatient falls prevention education and predictors of falls after hospital discharge--follow-up to a randomized controlled trial. J Gerontol A Biol Sci Med Sci 2011 Sep 09;66(9):1001-1012. [doi: 10.1093/gerona/glr085] [Medline: 21743091]
40. Sherrington C, Lord SR, Vogler CM, Close JCT, Howard K, Dean CM, et al. A post-hospital home exercise program improved mobility but increased falls in older people: a randomised controlled trial. PLoS One 2014;9(9):e104412 [FREE Full text] [doi: 10.1371/journal.pone.0104412] [Medline: 25180702]
41. Hill AM, McPhail SM, Haines TP, Morris ME, Etherton-Beer C, Shorr R, et al. Falls after hospital discharge: a randomized clinical trial of individualized multimodal falls prevention education. J Gerontol A Biol Sci Med Sci 2019 Aug 16;74(9):1511-1517 [FREE Full text] [doi: 10.1093/gerona/glz026] [Medline: 30721940]
42. Halvarsson A. Åtgärder på sjukhus. Vårdhandboken. URL: https://www.vardhandboken.se/vard-och-behandling/basal-och-preventiv-omvardnad/fallprevention/atgarder-pa-sjukhus/ [accessed 2023-11-30]
43. Halvarsson A. Teamarbete och ansvar. Vårdhandboken. URL: https://www.vardhandboken.se/vard-och-behandling/basal-och-preventiv-omvardnad/fallprevention/teamarbete-och-ansvar/ [accessed 2023-11-30]
44. Spoelstra SL, Given BA, Given CW. Fall prevention in hospitals: an integrative review. Clin Nurs Res 2012 Feb;21(1):92-112. [doi: 10.1177/1054773811418106] [Medline: 21862700]
45. Enhörning T, Svensson M, Nystrand C, Fastbom J. Fallprevention – en kostnadseffektiv åtgärd? Hälsoekonomiska beräkningar av fallpreventiva åtgärder för äldre. Socialstyrelsen. 2022. URL: https://www.socialstyrelsen.se/contentassets/a14087b2443142d5b91b21454daf22e8/2022-5-7923.pdf [accessed 2025-08-12]
46. Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. Adm Policy Ment Health 2015 Sep;42(5):533-544 [FREE Full text] [doi: 10.1007/s10488-013-0528-y] [Medline: 24193818]
47. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
48. Barber JP, Walczak KK. Conscience and critic: peer debriefing strategies in grounded theory research. In: Proceedings of the American Educational Research Association. 2009 Presented at: AERA 2009; April 13-17, 2009; San Diego, CA.
49. Lincoln YS, Guba EG. But is it rigorous? Trustworthiness and authenticity in naturalistic evaluation. New Dir Prog Eval 2004 Nov 04;1986(30):73-84. [doi: 10.1002/ev.1427]
50. Enworo OC. Application of Guba and Lincoln's parallel criteria to assess trustworthiness of qualitative research on indigenous social protection systems. Qual Res J 2023 Mar 07;23(4):372-384. [doi: 10.1108/QRJ-08-2022-0116]
51. Forero R, Nahidi S, De Costa J, Mohsin M, Fitzgerald G, Gibson N, et al. Application of four-dimension criteria to assess rigour of qualitative research in emergency medicine. BMC Health Serv Res 2018 Feb 17;18(1):120. [doi: 10.1186/s12913-018-2915-2] [Medline: 29454350]
52. Choi YS, Lawler E, Boenecke CA, Ponatoski ER, Zimring CM. Developing a multi-systemic fall prevention model, incorporating the physical environment, the care process and technology: a systematic review. J Adv Nurs 2011 Dec;67(12):2501-2524. [doi: 10.1111/j.1365-2648.2011.05672.x] [Medline: 21545639]
53. Damar HT, Cirito?lu B, Aslan G, Öztürk M. Investigating the relationship between fall prevention behaviors and health literacy in hospitalized older patients. Geriatrik Bilimler Dergisi 2023;6(1):1-12 [FREE Full text] [doi: 10.47141/geriatrik.1218729]

Abbreviations

ADL: activity in daily life

XSL•FO
RenderX

COREQ: Consolidated Criteria for Reporting Qualitative Research
HCD: human-centered design
PACT: people, activities, contexts, and technologies

Edited by C Jacob; submitted 25.02.25; peer-reviewed by S Hignett, A Garcia; comments to author 09.05.25; revised version received 02.07.25; accepted 05.07.25; published 02.10.25.

Please cite as:
Ehn M, Revenäs Å, Tobiasson H
Expanding the Design Space for Fall Prevention in Acute Orthopedic Hospital Care: Human-Centered Design Study
JMIR Hum Factors 2025;12:e73110
URL: https://humanfactors.jmir.org/2025/1/e73110
doi:10.2196/73110
PMID:

©Maria Ehn, Åsa Revenäs, Helena Tobiasson. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 02.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Representation of Rural Older Adults in AI for Health Research: Systematic Literature Review

Kristina Shiroma, PhD; Jacqueline Miller, MLIS

School of Information Studies, Louisiana State University, 223T Peabody Hall, Baton Rouge, LA, United States

Corresponding Author:
Kristina Shiroma, PhD
School of Information Studies, Louisiana State University, 223T Peabody Hall, Baton Rouge, LA, United States

Abstract

Background: Older adults in rural communities face unique challenges, including geographic isolation, a shortage of health care providers, and limited access to specialized services. Artificial intelligence (AI) has emerged as a promising solution for improving health care access and delivery. However, concerns persist about equitable access and representation in these innovations, especially for marginalized populations where technological literacy and infrastructure may present additional barriers to effective use.

Objective: This systematic literature review aims to: (1) identify existing literature on AI for health research focused on rural older adults, and (2) critically evaluate the identified literature to highlight gaps and inform the development of inclusive AI for health research and design practices.

Methods: Between January 2024 and March 2025, we followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 Protocol to conduct a systematic search of health and computer science literature. We searched 7 databases (PubMed, CINAHL Plus with Full Text, PsycINFO, Web of Science, IEEE Explore, ACM Digital Library, and Scopus) to identify relevant research papers published between January 2013 and December 2023. We used predetermined search terms and built-in result limiters, including English language, human subjects, full text, and aged 65+. Publications were excluded if they were not empirical or did not include a focus on older adults, rural populations, or AI. The resulting data were reviewed, coded, and analyzed using thematic analysis.

Results: A total of 23 papers comprised the final sample. The results showed that the representation of rural older adults in AI for health literature is limited. We identified three salient themes: (1) Numbers over Narratives: The Quantitative Focus in AI for Health Research on Older Adults, (2) Efficacy over Impact: Prevalence Clinical Outcomes in AI for Health Research, and (3) Deepening Disparities: Representation of Rurality Missing in AI for Health Research. These themes underscore the need for a more nuanced understanding of how AI for health research can be tailored to the specific needs of rural older populations.

Conclusions: Our systematic analysis identified a robust body of research on AI for older adults. However, a critical gap emerged with a dearth of studies explicitly focusing on older adults in rural communities. This lack of representation raises concerns about the generalizability of findings and the potential for exacerbating existing health care disparities in rural areas. Future research should: (1) prioritize targeted recruitment strategies for rural older adult participants to ensure better representation in AI for health research; (2) develop community-based AI policies, practices, and products that reflect the specific needs and contexts of rural populations; and (3) explore solutions that address the limited representation of rural communities, ensuring that AI interventions are equitable, accessible, and beneficial for all.

(JMIR Hum Factors 2025;12:e70057) doi:10.2196/70057

KEYWORDS

artificial intelligence; aging; older adult; rural; health care research; Preferred Reporting Items for Systematic Reviews and Meta-Analyses; PRISMA; systematic literature review

Introduction

Background

The global aging population presents significant opportunities to expand and extend health care, and artificial intelligence (AI) offers promising avenues to address these issues. As the population gets older, successful AI integration into health care systems necessitates a deeper understanding of the specific needs and perspectives of older adults. The increasing prevalence of chronic health conditions among older adults intensifies the demand for health care services while also offering new opportunities for engagement through digital health tools [1]. AI holds promise for meeting these needs, from early disease detection [2] to personalized care management [3] and even robotic assistance for independent living [4]. However,

XSL•FO
RenderX

the integration of AI into health care is not without significant challenges. To be effective, AI-powered health care tools must be user-friendly, accessible to individuals with varying levels of technological literacy, and free from algorithmic biases that could disproportionately affect older populations [5].

Rurality and Older Adults' Health

The demographic shift toward an aging population is particularly pronounced in rural areas, where access to health care is often limited. Over the past decade, 106 rural hospitals in the United States have closed [6,7], and another 600 are at risk of following suit [8], leaving rural communities increasingly vulnerable to limited access to essential health care services. These hospitals have traditionally provided vital care to large populations of older adults, but as closures continue, there presents a need to rethink traditional in-person health care models. Approximately 20% of the US population, or 60 million people, live in rural areas, and they face disproportionately worse health outcomes compared to urban populations [9]. Rural older adults, in particular, experience higher rates of comorbidities, such as obesity, cardiovascular disease, and stroke, as well as a growing number of emergency department visits [9,10]. This has contributed to widening life expectancy gaps between rural and urban populations [11]. In addition, rural communities face unique barriers that make it harder for older adults to access the specialized care required to address these complex health challenges [10]. Although online care models and other AI for health solutions are beginning to demonstrate promise in improving access to specialized health care, many rural areas still lack the necessary infrastructure to implement these solutions [12]. Ensuring AI health care solutions address these challenges will be crucial in providing equitable care for rural older adults.

AI for Health Research

We use the phrase "AI for health research" to refer to the investigation and development of AI technologies to improve various aspects of health care. By leveraging machine learning, natural language processing, and other AI techniques, researchers aim to create systems that can process vast amounts of health data to enhance clinical decision-making as well as to deliver highly personalized and sometimes remote care. However, despite its potential, AI for health research faces several significant challenges [13]. One of the most pressing issues is the quality and representativeness of the data used to train AI systems [14]. Large datasets are essential for AI models; however, many of these datasets may be incomplete, biased, or not sufficiently diverse, raising concerns about the accuracy and fairness of AI-generated decisions [2]. Furthermore, the sensitive nature of health data raises ethical concerns, particularly around privacy, security, and the risk of misuse. Ensuring robust data protection mechanisms is crucial to maintaining public trust in AI-driven health care solutions [15].

Another major challenge lies in the "black box" nature of many AI algorithms, which makes it difficult for health care providers and patients to understand how decisions are made [16]. This lack of transparency undermines accountability and raises the potential for unrecognized biases and amplified health care disparities [17,18]. Research continues to show that top-down or prescriptive design processes in the development of technology, including AI, often fail to include the perspectives of older adults, overlooking their unique experiences within their communities [19]. However, when leveraging the knowledge of this group through community-based research practices, technology design can be more effectively tailored to address issues, such as usability, literacy, privacy, and data management [19]. In addition, there is a significant gap in understanding the perspectives of older adults and other marginalized populations, such as those in rural areas. Without incorporating their diverse experiences and needs, AI tools risk failing to address the specific challenges faced by these groups, potentially exacerbating existing health inequities [20]. Addressing these challenges is essential to ensuring that AI for health research leads to equitable, effective, and ethically sound health care solutions.

Study Aims

Building on the insights from the existing literature, it is clear that gaps remain in understanding how AI for health research can effectively address and incorporate the health care needs of rural older adults. This systematic literature review aims to: (1) identify existing literature on AI for health research specifically focused on older adults residing in rural communities, (2) critically evaluate the identified literature to highlight gaps, and (3) inform the development of inclusive AI for health research and design practices.

Methods

Study Design

We followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 protocol for systematic literature reviews [21]. The PRISMA 2020 checklist is provided as Checklist 1. A protocol was not registered for this systematic literature review. A systematic review methodology was selected to enable a focused, critical appraisal of how rural older adults are represented in empirical AI for health research. Given the equity implications of AI in health care, a systematic review was appropriate to assess not only the presence but also the depth and nature of rural older adult representation using predefined inclusion criteria, structured data extraction, and formal critical appraisal tools (eg, Joanna Briggs Institute [JBI]).

We performed 4 rounds of systematic selection and review for the 7 selected databases (PubMed, CINAHL Plus with Full Text, PsycINFO, Web of Science, IEEE Explore, ACM Digital Library, and Scopus). In Round 1, we searched the databases using predetermined keywords. In Round 2, we screened titles and abstracts using predetermined inclusion and exclusion criteria. In Round 3, we further screened the full text of the selected papers using the same predetermined inclusion and exclusion criteria. Finally, in Round 4 we reviewed and coded each study in the final sample. A detailed description of our search strategies can be found in Multimedia Appendix 1. Once the articles were coded, the researchers met to analyze the data using thematic analysis [22]. The main themes identified centered on the representation of older adults in AI for health research.

XSL•FO
RenderX

In addition, we assessed the levels of evidence reported in the 23 papers [23-45] using the appraisal tools developed by the JBI to evaluate study quality [46]. To calculate assessment scores, 1 point was assigned to each question answered as "Yes," while responses of "No" or "Unclear" received 0 points. The total score was then divided by the number of questions and multiplied by 100 to yield a percentage. Based on the resulting score, studies were categorized as very poor (0% - 30%), poor (31% - 50%), fair (51% - 70%), good (71% - 90%), or excellent (>90%). Both authors reviewed and independently completed the checklist for each study; disagreements were resolved through discussion.

Ethical Considerations

This systematic literature review was deemed exempt by the Louisiana State University Institutional Review Board.

Results

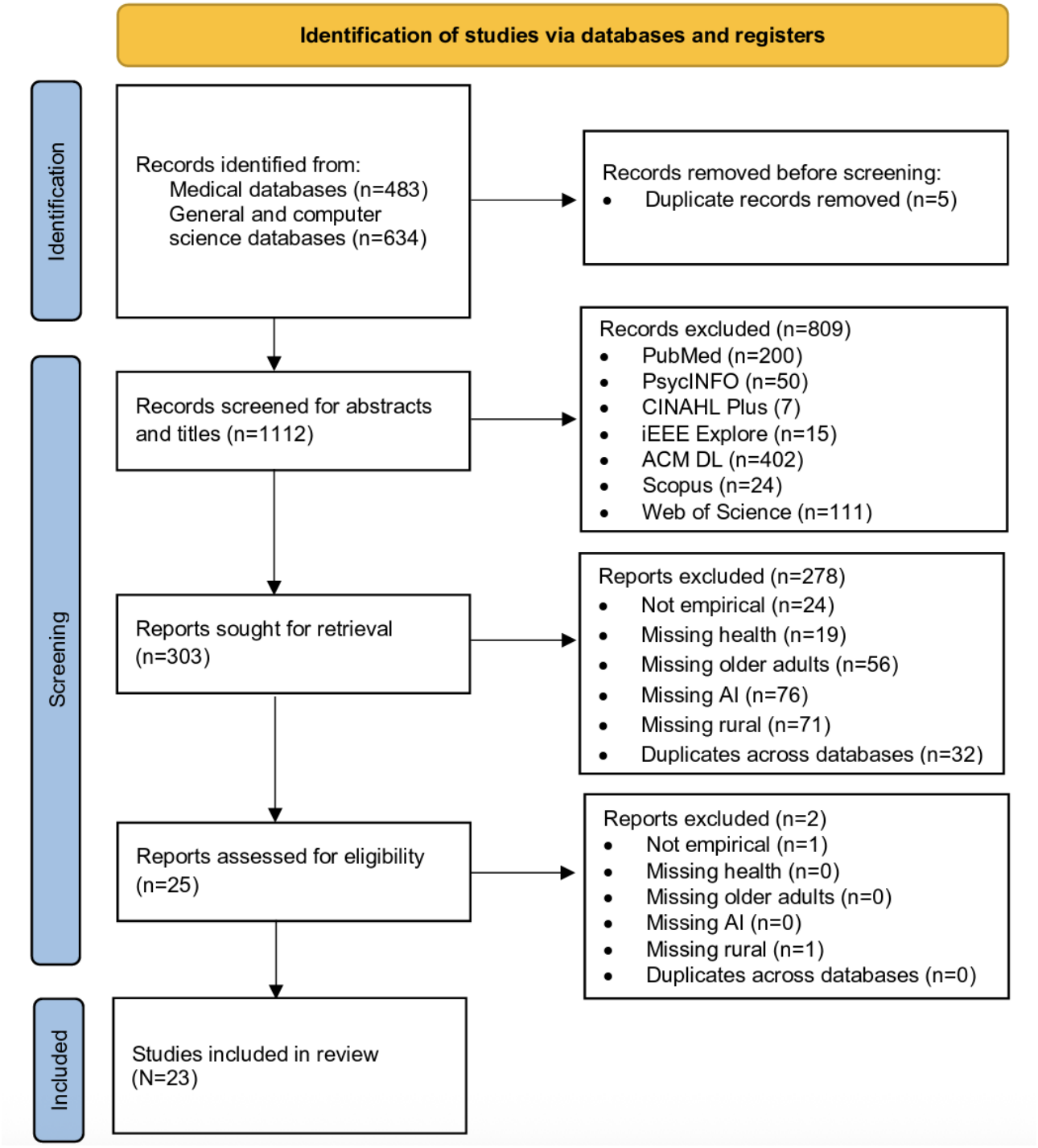
Search and Screening Outcomes

Round 1, keyword search, yielded 1112 publications across all databases, with 479 publications from medical databases and 633 from general and computing databases. After initial screening for the predetermined exclusion criteria, 303 remained for title and abstract screening in Round 2. During Round 2, a total of 278 publications were removed based on the predetermined exclusion criteria, leaving 25 publications for screening in Round 3. During Round 3, full-text screening, 2 publications were excluded.

This round resulted in a final sample comprised of 23 publications [23-45]. Among these, 8 publications [23-30] selected from the medical databases (PubMed, CINAHL Plus with Full Text, and PsycINFO) and 15 publications [31-45] selected from the general and computer science databases (Web of Science, Scopus, ACM DL, and IEEE Explore). The final PRISMA search strategy with results is illustrated in Figure 1 below.

XSL•FO
RenderX

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram. AI: artificial intelligence.



Quality Assessment Outcomes

To evaluate the methodological quality of the included studies, we used the JBI Critical Appraisal Checklists. Each study was independently reviewed by 2 researchers, with discrepancies resolved through discussion to ensure consistency and minimize potential bias. Studies that met the majority of JBI criteria were retained for inclusion in the final synthesis, ensuring a high-quality evidence base. Results of the quality appraisal are provided in Multimedia Appendix 2. Of the included studies, 10/23 (43%) [23,25-27,36-39,43,45] were rated as excellent, 10/23 (43%) [24,28,29,31-34,40,41,44] as good, and 3/23 (13%) [30,35,42] as fair in overall methodological quality.

Descriptive Statistics

A summary of the characteristics of all 23 studies [23-45] is presented in Multimedia Appendices 3 and 4. Of the 23 [23-45] publications, the years of publication ranged from 2014 to 2023; the majority of the research has been published since 2021 (16, 70%) [23-27,29,33-38,42-45]. Studies were conducted in Asia (13, 57%) [23-27,29,33-35,39-41,45], Europe (2, 9%) [30,43], North America (7, 30%) [28,31,32,36,38,42,44], and Africa (1, 4%) [37]. Fields of study represented in this sample included health sciences (15, 65%) [23,26,30,32,33,35-40,42-45], psychology (3, 13%) [24,25,31], neuroscience (2, 9%) [29,34], gerontology (2, 9%) [27,41], and rehabilitation (1, 4%) [28].

XSL•FO
RenderX

Quantitative methods were almost exclusively used (19, 83%) [23-27,29,30,34-45]. The quantitative studies focused on aspects of AI in health research, such as prediction [24-26,42], screening [23,37], investigating [27,43], estimating [29], evaluating [30,38-41,44], and design and development [34-36,45]. The qualitative studies (4, 17%) [28,31-33] focused on patient experiences of robotic stroke therapy [28], the demand for smart health care services for older adults [33], and participatory design of socially assistive robots for mental health [31,32].

Sample sizes across the included studies ranged from as few as 10 to over 1.85 million older adult participants. Definitions of what constituted an "older adult" varied considerably. Participant ages spanned from 50 to 105 years, reflecting a lack of consensus on age-based eligibility criteria. Two studies did not specify age-based criteria but noted that participants were recruited from older care facilities, implying an assumed classification based on setting rather than age. It is important to note that there is no universally accepted standard for defining "older adult." The World Health Organization generally defines older adults as individuals aged 60 years and older, particularly in global and low- to middle-income contexts [47]. In contrast, the United States Census Bureau designates age 65 years as the threshold for older adulthood [48]. These variations highlight how classifications of "older adult" may differ by region, cultural context, or institutional policy.

The representation of rurality similarly varied across the studies. Only 7/23 (30%) [23,26,28,30,35,41,44] of the publications included in this review focused entirely on a rural population. The context of rurality in these studies included rural villages in Yangxi County, China [23]; villages and towns in rural districts of Beijing [35]; home-based care in rural China [41]; rural areas in the state of Georgia, United States [28]; small rural hospitals in the United States [44]; rural regions of South Korea [26]; and a rural health care center in Sweden [30]. However, rurality was often (14/23, 57%) [24,25,29,31-34,36-38,40,42,43,45] treated as a background variable used for comparison against urban settings rather than as a distinct sociocultural or structural context worthy of focused analysis. Furthermore, nearly half of the studies (11/23, 49%) [23-25,27,33-35,39-41,45] examining AI for health in rural areas were conducted in China, suggesting a potential geographic bias or overrepresentation that may limit the broader applicability of findings and obscure rural-specific insights from other regions.

The types of AI used across the included studies were primarily described as machine learning (n=7, 30%), which supported a range of predictive tasks, including identifying depression risk [24], forecasting cognitive decline [25], modeling active aging [27], predicting suicidal ideation in rural older adults [26], and developing risk scores for stroke and coronary heart disease [34,36,45]. Deep learning (n=1, 4%) was used for image-based screening, specifically for detecting retinal lesions in rural older adult populations [23]. Robotics (n=2, 9%) was applied in studies involving telerehabilitation for rural veterans [28] and remote consultation through robot-assisted echocardiography [30]. One additional study [31] described the ongoing participatory design of socially assistive robots with older adults, though it had not yet reported implementation outcomes. Notably, no studies examined the use of generative AI (eg, large language models). Moreover, few studies included qualitative data capturing older adults' own perceptions, experiences, or understanding of AI in their health care, limiting insight into how this population interprets, evaluates, or engages with AI technologies in practice.

AI was primarily examined as a strategy to extend health care services to remote and underserved areas. The largest proportion of studies (10/23, 43%) [23-26,31,34,36,37,42,45] focused on AI as a screening and predictive tool in rural health care settings, supporting early identification of retinal lesions [23], depression [24,31], cognitive decline [25], suicidal ideation [26], stroke risk [34], early hospitalizations due to ambulatory-care sensitive conditions [36], diabetes [37], novel disease association [42], and cardiovascular disease [45]. Fewer studies explored other applications of AI, including its use in robotic telerehabilitation, particularly for rural veterans with limited access to in-person care [28]; its role in supporting long-term public planning and social support through epidemiological modeling in aging populations [29]; and its utility in facilitating remote specialist consultation, such as through robot-assisted echocardiography that reduced time to care [30]. These findings highlight a dominant focus on AI's predictive capabilities in rural contexts, while pointing to emerging but underexplored applications in rehabilitation, planning, and remote access.

Discussion

Principal Findings

Across the literature, we identified three salient themes: (1) Numbers over Narrative: The Quantitative Focus in AI for Health Research on Older Adults, (2) Efficacy over Impact: Prevalence of Clinical Outcomes in AI for Health Research, and (3) Deepening Disparities: Rural Older Adults Missing in AI for Health Research. We examine these themes in the following sections.

Theme 1: Numbers Over Narratives: The Quantitative Focus in AI for Health Research on Older Adults

Across this systematic literature review, we identified that AI for health research predominantly emphasizes quantitative findings, focusing on measurable results, such as diagnostic accuracy, predictive modeling, and algorithmic efficacy. This trend reflects broader scientific priorities placed on numerical data and technical performance, often at the expense of deeper, qualitative insights into the lived experiences of older adults.

The studies reviewed generally included older adults as subjects of analysis, rather than as active participants in the research process. This suggests, and in many cases perpetuates, a research approach that focuses on older adults, rather than with them. By emphasizing AI's reach and capabilities in health research, such as screening [23,37], prediction [24-26,40,42], identification [27,38,43], and feasibility [30], the literature largely overlooks the personal narratives and lived experiences that are crucial for designing patient-centered, effective AI solutions. For example, Cui et al [23] designed their quantitative study to test the efficacy of AI as a screening tool for rural older adults who are particularly more susceptible to retinal lesions.

XSL•FO
RenderX

This application of AI, like most of the studies in this review, is provider-focused, as older adults were not asked to contribute their perspectives on the use of AI within their health care.

This focus on quantitative findings, while valuable in establishing initial data and baseline efficacy, remains a limitation in fully understanding the potential and impact of AI for health technologies [5]. The minimal presence of qualitative research undermines the broader relevance and applicability of AI solutions for marginalized populations, particularly rural older adults. Among the few studies that incorporated qualitative components, many offered limited depth or treated qualitative insights as secondary. For example, Lee et al [33] conducted a qualitative study with older adults in China and found that 81.6% of participants had never independently purchased or leased smart health care products, highlighting low levels of engagement and understanding of these tools, despite their increasing availability. Studies focused on clinical outcomes and algorithmic performance often fail to address the real-world implications of these technologies, such as usability, trust, and accessibility, which present critical factors in adoption and long-term success, especially in rural communities.

Most qualitative studies explore the experiences, preferences, and barriers faced by older adults [19]. This approach not only enriches the findings of quantitative studies but also ensures that AI for health tools is designed in ways that truly meet the needs of older adults, especially those in underrepresented communities. Without this qualitative dimension, the development of AI technologies will likely continue to miss crucial elements that could optimize their utility in everyday health care settings, particularly in rural and underserved populations. The current literature gap underscores the need for a more balanced research approach that integrates both quantitative and qualitative methods, ensuring that emerging technologies are both scientifically rigorous and human-centered.

Theme 2: Efficacy Over Impact: Prevalence of Clinical Outcomes in AI for Health Research

Our systematic literature review highlighted the prevalence of clinical outcomes in AI for health research within the context of rural older adult populations. A majority of studies in this sample concentrated on the efficacy of AI interventions designed to evaluate existing models or outcomes (6/23, 26%) [28,30,37-39,44], predicting health outcomes (4/23, 17%) [24-26,42], investigating determinants of health (3/23, 13%) [27,33,38], screening medical data (2/23, 9%) [23,45], estimating disease prevalence (1/23, 4%) [29], and analyzing associated factors (1/23, 4%) [41]. While the investigations of these outcomes are pivotal and promising, our review suggests that a more inclusive, participatory approach may ensure that AI tools are better aligned with the challenges and needs of rural older adults. In contrast to the clinical-centric focus, other technology sectors have increasingly adopted community-based participatory research models to engage users in the design and development process [20]. However, such approaches have yet to be widely applied in AI for health research, as seen in this systematic literature review.

Our findings suggest that while AI-driven tools demonstrate potential in improving diagnostic precision and personalizing care for older adults [2-4], the understanding of their real-world application in rural areas remains underexplored. Rural health care systems are complex, with challenges, such as limited broadband access, shortages of health care professionals, and transportation barriers, all of which may impede the successful implementation and sustainability of AI solutions [6-8]. In addition to those challenges, AI solutions also present unique challenges, including black box algorithms [16], privacy issues [19], and limited AI literacy [19,20]. None of the studies focused on any of these user-centered issues. This gap offers rich research opportunities for technology-driven, health-related fields, such as information science and computer science, to address these issues and contribute to a user-centered approach in understanding the development and implementation of AI for health research.

Furthermore, without a qualitative, community-centered approach to AI for health research, which includes understanding the barriers and experiences of rural older adults, the focus may likely remain on clinical outcomes rather than the broader community health outcomes that truly affect these populations [49].

Our literature review revealed a notable absence of initiatives aimed at improving digital and health literacy for rural older adults, highlighting a significant gap in current research. Beyond clinical outcomes, there are valuable opportunities to support the health of rural older adults through efforts focused on enhancing digital literacy as well as education on AI's impact on health care [19,20]. However, none of the studies within this review investigated older adults' digital literacy or AI literacy specifically focused on health care tools.

Theme 3: Deepening Disparities: Representation of Rurality Missing in AI for Health Research

The findings from this systematic review of the literature highlight a glaring gap in the representation in AI for health research of the perspectives of older adults from rural areas. During the initial screening, we found that using the term "rural" in combination with other key terms frequently yielded no results. As a result, we manually identified references to rural populations within the titles and abstracts. This process underscored not only the lack of focus on rural communities but also the challenges of finding research that includes rural older adults in a meaningful way. AI tools for health care and health decision support are developed using large datasets that, if not inclusive, will continue to marginalize communities like rural older adults [2].

Our review also reveals a geographic imbalance in the literature, with a substantial proportion of studies (11/23, 49%) [23-25,27,33-35,39-41,45] conducted in China. This concentration may suggest a geographic bias in the global development of AI for health research, raising concerns about the applicability and generalizability of findings to rural populations in other countries with different health care systems, cultural contexts, and infrastructure. Moreover, most of the studies we reviewed were heavily urban-centric, emphasizing data from metropolitan populations while overlooking the

XSL•FO
RenderX

distinct challenges and health contexts of rural communities. As a result, AI models developed in these contexts may fail to reflect the lived experiences or health care needs of rural older adults, further widening existing disparities [9-11].

Rurality was rarely a central focus in the studies reviewed. In most cases, it was treated as a background demographic variable rather than an analytical lens. Even in studies that addressed rural-related issues, such as health care access or social isolation, the quantitative data often lacked input from rural older adults themselves. This absence of direct perspectives limits the field's ability to design AI applications that are contextually relevant and responsive to diverse rural settings [13]. Beyond merely including rural populations as a demographic factor, it is essential to deeply understand the lived experiences of rural older adults within AI for health research. In many of the studies reviewed, rurality was treated as a simple comparative statistic, with little exploration of how rural-specific factors might impact research findings or health care outcomes. For example, in many of the studies (15/23, 65%) [24,25,27,29,31-34,36-38,40,42,43,45], rural populations were only mentioned in descriptive statistics, without any further analysis of how rurality influenced the research outcomes. This limited approach prevents a comprehensive understanding of the needs and challenges of rural older adults, ultimately undermining the development of AI solutions and algorithms that are equitable and effective for all populations [17].

Limitations and Future Directions

Like all systematic literature reviews, this study has limitations. Our search was constrained by the language spoken by the research team, limiting inclusion to studies published in English and potentially omitting relevant research in other languages. In addition, the set of keywords used in the search was not exhaustive, and future reviews using alternative or broader search terms may yield different findings. Despite these limitations, this review serves as an initial step in advancing a research agenda focused on the inclusion of older adults, particularly those from underrepresented, rural populations, in AI for health research. Our primary aim was to identify opportunities where AI can help address unmet health needs and enhance health care delivery for these communities. The findings from this review provide a foundation for the development and implementation of responsible, inclusive AI-driven health care solutions.

Conclusions

These findings suggest that AI for health care research is currently focused on (1) Numbers over Narrative and (2) Efficacy Over Impact, while (3) Deepening Disparities. While older adults are often described as lagging or in triple jeopardy when it comes to technology in health care, this literature review suggests that it is AI for health research that is actually lagging when it comes to understanding the specific community health care needs of this population. In addition, our findings identify a growing need for increased attention to qualitative research to investigate rural older adults' perspectives on the role of AI in the development of emerging health care solutions. In addition, there is a need for concerted efforts to examine the impacts of AI for health care research. This includes investigating the transparency of AI tools as well as older adults' access to education on the implications of AI within their health care.

Acknowledgments

We would like to acknowledge that preliminary findings of this systematic literature review were originally presented as a conference presentation at the 2024 Annual Scientific Meeting of the Gerontological Society of America. The authors attest that there was no use of generative artificial intelligence (AI) technology in the generation of text, figures, or other informational content of this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Detailed search strategy.
[PDF File, 71 KB - humanfactors_v12i1e70057_app1.pdf]

Multimedia Appendix 2
Quality assessment for the final sample.
[PDF File, 154 KB - humanfactors_v12i1e70057_app2.pdf]

Multimedia Appendix 3
Summary of 23 studies in the final sample.
[PDF File, 133 KB - humanfactors_v12i1e70057_app3.pdf]

Multimedia Appendix 4
Descriptive results (N=23).

XSL•FO
RenderX

[PNG File, 126 KB - humanfactors_v12i1e70057_app4.png]

Checklist 1
Preferred Reporting Items for Systematic Reviews and Meta-Analyses checklist
[DOCX File, 29 KB - humanfactors_v12i1e70057_app5.docx]

References

1. Buttorff C, Ruder T, Bauman M. Multiple chronic conditions in the united states. RAND Corporation. 2017. URL: https://www.rand.org/pubs/tools/TL221.html [accessed 2025-09-12]
2. Beam AL, Kohane IS. Big data and machine learning in health care. JAMA 2018 Apr 3;319(13):1317-1318. [doi: 10.1001/jama.2017.18391] [Medline: 29532063]
3. Miotto R, Wang F, Wang S, Jiang X, Dudley JT. Deep learning for healthcare: review, opportunities and challenges. Brief Bioinform 2018 Nov 27;19(6):1236-1246. [doi: 10.1093/bib/bbx044] [Medline: 28481991]
4. Malik P, Pathania M, Rathaur V. Overview of artificial intelligence in medicine. J Family Med Prim Care 2019;8(7):2328. [doi: 10.4103/jfmpc.jfmpc_440_19]
5. Stypinska J. AI ageism: a critical roadmap for studying age discrimination and exclusion in digitalized societies. AI Soc 2023;38(2):665-677. [doi: 10.1007/s00146-022-01553-5] [Medline: 36212226]
6. 196 Rural Hospital Closures and Conversions since January 2005. The Cecil G. Sheps Center for Health Services Research. 2023. URL: https://www.shepscenter.unc.edu/programs-projects/rural-health/rural-hospital-closures/ [accessed 2024-12-18]
7. Center for Healthcare Quality & Payment Reform. The impact of the pandemic on rural hospitals. 2023 URL: https://ruralhospitals.chqpr.org/downloads/Pandemic_Impact_on_Rural_Hospitals.pdf [accessed 2024-10-25]
8. McCarthy S, Moore D, Smedley WA, et al. Impact of rural hospital closures on health-care access. J Surg Res 2021 Feb;258:170-178. [doi: 10.1016/j.jss.2020.08.055] [Medline: 33011448]
9. Harrington RA, Califf RM, Balamurugan A, et al. Call to action: rural health: a presidential advisory from the American Heart Association and American Stroke Association. Circulation 2020 Mar 10;141(10):e615-e644. [doi: 10.1161/CIR.0000000000000753] [Medline: 32078375]
10. Krause KE, Kokorelias KM, Sinha SK. A systematic review and qualitative analysis of geriatric models of care for rural and remote populations. Rural Remote Health 2022 Aug;22(3):7486. [doi: 10.22605/RRH7486] [Medline: 35975281]
11. Abrams LR, Myrskylä M, Mehta NK. The growing rural-urban divide in US life expectancy: contribution of cardiovascular disease and other major causes of death. Int J Epidemiol 2022 Jan 6;50(6):1970-1978. [doi: 10.1093/ije/dyab158] [Medline: 34999859]
12. Cowie P, Townsend L, Salemink K. Smart rural futures: will rural areas be left behind in the 4th industrial revolution? J Rural Stud 2020 Oct;79:169-176. [doi: 10.1016/j.jrurstud.2020.08.042] [Medline: 32863561]
13. Sharma S, Rawal R, Shah D. Addressing the challenges of AI-based telemedicine: best practices and lessons learned. J Educ Health Promot 2023;12:338. [doi: 10.4103/jehp.jehp_402_23] [Medline: 38023098]
14. Chen Y, Clayton EW, Novak LL, Anders S, Malin B. Human-centered design to address biases in artificial intelligence. J Med Internet Res 2023 Mar 24;25:e43251. [doi: 10.2196/43251] [Medline: 36961506]
15. Gerlings J, Jensen MS, Shollo A. Explainable AI, but explainable to whom? an exploratory case study of xai in healthcare. In: Lim CP, Chen YW, Vaidya A, Mahorkar C, Jain LC, editors. Handbook of Artificial Intelligence in Healthcare: Springer International Publishing; 2022, Vol. 212:169-198. [doi: 10.1007/978-3-030-83620-7_7]
16. Wadden JJ. Defining the undefinable: the black box problem in healthcare artificial intelligence. J Med Ethics 2021 Jul 21:medethics-2021-107529. [doi: 10.1136/medethics-2021-107529] [Medline: 34290113]
17. Hassija V, Chamola V, Mahapatra A, et al. Interpreting black-box models: a review on explainable artificial intelligence. Cogn Comput 2024 Jan;16(1):45-74. [doi: 10.1007/s12559-023-10179-8]
18. Durán JM, Jongsma KR. Who is afraid of black box algorithms? On the epistemological and ethical basis of trust in medical AI. J Med Ethics 2021 Mar 18:medethics-2020-106820. [doi: 10.1136/medethics-2020-106820] [Medline: 33737318]
19. Quinn TP, Jacobs S, Senadeera M, Le V, Coghlan S. The three ghosts of medical AI: can the black-box present deliver? Artif Intell Med 2022 Feb;124:102158. [doi: 10.1016/j.artmed.2021.102158] [Medline: 34511267]
20. Wang S, Bolling K, Mao W, et al. Technology to support aging in place: older adults' perspectives. Healthcare (Basel) 2019 Apr 10;7(2):60. [doi: 10.3390/healthcare7020060] [Medline: 30974780]
21. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
22. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
23. Cui T, Lin D, Yu S, et al. Deep learning performance of ultra-widefield fundus imaging for screening retinal lesions in rural locales. JAMA Ophthalmol 2023 Nov 1;141(11):1045-1051. [doi: 10.1001/jamaophthalmol.2023.4650] [Medline: 37856107]

XSL•FO
RenderX

24. Su D, Zhang X, He K, Chen Y. Use of machine learning approach to predict depression in the elderly in China: a longitudinal study. J Affect Disord 2021 Mar 1;282:289-298. [doi: 10.1016/j.jad.2020.12.160] [Medline: 33418381]
25. Wu Y, Jia M, Xiang C, Lin S, Jiang Z, Fang Y. Predicting the long-term cognitive trajectories using machine learning approaches: a Chinese nationwide longitudinal database. Psychiatry Res 2022 Apr;310:114434. [doi: 10.1016/j.psychres.2022.114434] [Medline: 35172247]
26. Kim J, Gwak D, Kim S, Gang M. Identifying the suicidal ideation risk group among older adults in rural areas: developing a predictive model using machine learning methods. J Adv Nurs 2023 Feb;79(2):641-651. [doi: 10.1111/jan.15549] [Medline: 36534434]
27. Yu J, Huang W, Kahana E. Investigating factors of active aging among Chinese older adults: a machine learning approach. Gerontologist 2022 Mar 28;62(3):332-341. [doi: 10.1093/geront/gnab058] [Medline: 33942091]
28. Cherry CO, Chumbler NR, Richards K, et al. Expanding stroke telerehabilitation services to rural veterans: a qualitative study on patient experiences using the robotic stroke therapy delivery and monitoring system program. Disabil Rehabil Assist Technol 2017 Jan;12(1):21-27. [doi: 10.3109/17483107.2015.1061613] [Medline: 26135221]
29. Jin H, Crimmins E, Langa KM, Dey AB, Lee J. Estimating the prevalence of dementia in India using a semi-supervised machine learning approach. Neuroepidemiology 2023;57(1):43-50. [doi: 10.1159/000528904] [Medline: 36617419]
30. Boman K, Olofsson M, Berggren P, Sengupta PP, Narula J. Robot-assisted remote echocardiographic examination and teleconsultation: a randomized comparison of time to diagnosis with standard of care referral approach. JACC: Cardiovascular Imaging 2014 Aug;7(8):799-803. [doi: 10.1016/j.jcmg.2014.05.006] [Medline: 25124011]
31. Lee HR, Šabanović S, Chang WL, et al. Steps toward participatory design of social robots: mutual learning with older adults with depression. 2017 Presented at: Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction; Mar 6-9, 2017; Vienna, Austria p. 244-253. [doi: 10.1145/2909824.3020237]
32. Randall N, Šabanović S, Chang W. Engaging older adults with depression as co-designers of assistive in-home robots. 2018 May 21 Presented at: PervasiveHealth '18: 12th EAI International Conference on Pervasive Computing Technologies for Healthcar; May 21-24, 2018; New York, NY p. 304-309. [doi: 10.1145/3240925.3240946]
33. Lee CH, Zhang Z, Zhao X. A survey of smart healthcare for the elderly based on user requirements and supply accessibility. Presented at: 5th International Conference on Crowd Science and Engineering; Oct 16-18, 2021; Jinan, China p. 108-112. [doi: 10.1145/3503181.3503217]
34. Ding Z, Zhang L, Niu M, et al. Stroke prevention in rural residents: development of a simplified risk assessment tool with artificial intelligence. Neurol Sci 2023 May;44(5):1687-1694. [doi: 10.1007/s10072-023-06610-5] [Medline: 36653543]
35. Zhu T, Wu H, Chen C, Zhu H. Research on the rural community elderly care service management systems based on the internet of things. Presented at: Proceedings of the 2022 6th International Conference on Electronic Information Technology and Computer Engineering; Oct 21-23, 2022; Xiamen China p. 157-164. [doi: 10.1145/3573428.3573456]
36. Yi SE, Harish V, Gutierrez J, et al. Predicting hospitalisations related to ambulatory care sensitive conditions with machine learning for population health planning: derivation and validation cohort study. BMJ Open 2022 Apr 1;12(4):e051403. [doi: 10.1136/bmjopen-2021-051403] [Medline: 35365510]
37. Mathenge W, Whitestone N, Nkurikiye J, et al. Impact of artificial intelligence assessment of diabetic retinopathy on referral service uptake in a low-resource setting: the RAIDERS randomized trial. Ophthalmol Sci 2022 Dec;2(4):100168. [doi: 10.1016/j.xops.2022.100168] [Medline: 36531575]
38. Parry MW, Markowitz JS, Nordberg CM, Patel A, Bronson WH, DelSole EM. Patient perspectives on artificial intelligence in healthcare decision making: a multi-center comparative study. Indian J Orthop 2023 May;57(5):653-665. [doi: 10.1007/s43465-023-00845-2] [Medline: 37122674]
39. Zou FW, Tang YF, Liu CY, Ma JA, Hu CH. Concordance study between IBM Watson for oncology and real clinical practice for cervical cancer patients in China: a retrospective analysis. Front Genet 2020;11:200. [doi: 10.3389/fgene.2020.00200] [Medline: 32265980]
40. Tang H, Zhao J, Yang X. Explore machine learning for analysis and prediction of lung cancer related risk factors. Presented at: Proceedings of the 2018 2nd International Conference on Computer Science and Artificial Intelligence; Dec 8-10, 2018; Shenzhen, China p. 41-45. [doi: 10.1145/3297156.3297274]
41. Liu F, Wang W, Zhou Y, Huang F. Research on the impact of education level for the elderly over 60 years old on the elderly care mode and the corresponding industry. Presented at: 2020 3rd International Conference on Education Technology Management; Dec 17-19, 2020; London, United Kingdom p. 155-161. [doi: 10.1145/3446590.3446615]
42. Manemann SM, St Sauver JL, Liu H, et al. Longitudinal cohorts for harnessing the electronic health record for disease prediction in a US population. BMJ Open 2021 Jun 8;11(6):e044353. [doi: 10.1136/bmjopen-2020-044353] [Medline: 34103314]
43. Miró Catalina Q, Femenia J, Fuster-Casanovas A, et al. Knowledge and perception of the use of AI and its implementation in the field of radiology: cross-sectional study. J Med Internet Res 2023 Oct 13;25:e50728. [doi: 10.2196/50728] [Medline: 37831495]
44. Jones KJ, Haynatzki G, Sabalka L. Evaluation of automated video monitoring to decrease the risk of unattended bed exits in small rural hospitals. J Patient Saf 2021 Dec 1;17(8):e716-e726. [doi: 10.1097/PTS.0000000000000789] [Medline: 33009181]

XSL•FO
RenderX

45. Zhang L, Niu M, Zhang H, et al. Nonlaboratory-based risk assessment model for coronary heart disease screening: Model development and validation. Int J Med Inform 2022 Mar 18;162:104746. [doi: 10.1016/j.ijmedinf.2022.104746] [Medline: 35325662]
46. Critical appraisal tools. JBI. 2017. URL: https://jbi.global/critical-appraisal-tools [accessed 2025-04-10]
47. Active ageing: a policy framework. : World Health Organization; 2002 URL: https://extranet.who.int/agefriendlyworld/wp-content/uploads/2014/06/WHO-Active-Ageing-Framework.pdf [accessed 2025-05-29]
48. The older population. : U.S. Census Bureau; 2020 URL: https://www2.census.gov/library/publications/decennial/2020/census-briefs/c2020br-07.pdf [accessed 2025-03-29]
49. Chu CH, Nyrup R, Leslie K, et al. Digital ageism: challenges and opportunities in artificial intelligence for older adults. Gerontologist 2022 Aug 12;62(7):947-955. [doi: 10.1093/geront/gnab167] [Medline: 35048111]

Abbreviations

AI: artificial intelligence
JBI: Joanna Briggs Institute
PRISMA : Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by A Kushniruk, P Santana-Mancilla; submitted 20.12.24; peer-reviewed by MP Gagnon, MR Ortiz; revised version received 01.08.25; accepted 01.08.25; published 15.09.25.

Please cite as:
Shiroma K, Miller J
Representation of Rural Older Adults in AI for Health Research: Systematic Literature Review
JMIR Hum Factors 2025;12:e70057
URL: https://humanfactors.jmir.org/2025/1/e70057
doi:10.2196/70057

© Kristina Shiroma, Jacqueline Miller. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 15.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

User and Provider Experiences With Health Education Chatbots: Qualitative Systematic Review

K yung-Eun (Anna) Choi[1,2,3*], Prof Dr; Sebastian Fitzek[1*], PhD

[1]Health Services Research Group, Medical Images Analysis and Artificial Intelligence, Danube Private University, Steiner Landstraße 124, Krems an der Donau, Austria

[2]Center for Health Services Research, Brandenburg Medical University, Neuruppin, Germany

[3]Evidence-Based Practice in Brandenburg—A JBI Affiliated Group, Hochstraße 29, Brandenburg an der Havel, Germany

[*]all authors contributed equally

Corresponding Author:
Sebastian Fitzek, PhD
Health Services Research Group, Medical Images Analysis and Artificial Intelligence, Danube Private University, Steiner Landstraße 124, Krems an der Donau, Austria

Abstract

Background: Chatbots, as dialog-based platforms, have the potential to transform health education and behavior-change interventions. Despite the growing use of chatbots, qualitative insights into user and provider experiences remain underexplored, particularly with respect to experiences and perceptions, adoption factors, and the role of theoretical frameworks in design.

Objective: This systematic review of qualitative evidence aims to address three key research questions (RQs): (RQ1) user and provider experiences; (RQ2) facilitators and barriers to adoption; and (RQ3) role of theoretical frameworks.

Methods: We systematically searched PubMed, the Cochrane Library, and ScienceDirect from January 1, 2018, to October 1, 2023, for English- or German-language, peer-reviewed qualitative or mixed methods studies. Studies were included if they examined users' or providers' experiences with chatbots in health education or behavior-change contexts. Two reviewers independently screened titles, abstracts, and full texts (Cohen κ=0.82). We used the Joanna Briggs Institute Critical Appraisal Checklist for quality assessment and conducted a reflexive thematic analysis following Braun and Clarke's framework.

Results: Among the 1754 records identified, 27 studies from 10 countries met the inclusion criteria, encompassing 241 qualitative-only participants and 10,802 mixed method participants (657 contributing qualitative data). For RQ1, users emphasized empathy and emotional connection. For RQ2, accessibility and ease of use emerged as facilitators, whereas trust deficits, technical glitches, and cultural misalignment were key barriers. For RQ3, the integration of behavior-change theories emerged as underutilized despite their potential to increase motivation.

Conclusions: Chatbots demonstrate strong potential for health education and behavior-change interventions but must address privacy and trust issues, embed robust theoretical underpinnings, and overcome adoption barriers to fully realize their impact. Future directions should include evaluations of cultural adaptability and rigorous ethical considerations in chatbot design.

Trial Registration: OSF Registries osf.io/4px23; https://osf.io/4px23

(JMIR Hum Factors 2025;12:e60205) doi:10.2196/60205

KEYWORDS

chatbot; health education; behavior change; user experience; privacy concerns; personalization; qualitative research.

Introduction

Chatbots are software applications designed for 2-way dialogue that simulate human conversation through text or speech and are increasingly integrated into health care systems to deliver health education and behavior-change interventions [1,2]. These applications provide tailored guidance and real-time support by leveraging the widespread availability of smartphones and internet connectivity, thereby delivering standalone health interventions [3]. Users receive timely information, reminders, and motivational messages that may improve health outcomes [4,5].

Despite their growing use, previous systematic reviews and meta-analyses have focused predominantly on quantitative outcomes, such as symptom reduction and behavioral compliance [1-3,6], leaving significant gaps in our understanding of the qualitative dimensions of chatbot interactions [7]. These gaps, namely, the limited qualitative insights into user experiences, the underutilization of theoretical frameworks in chatbot design, and an insufficient understanding of adoption

XSL•FO
RenderX

factors, form the foundation of this study and directly inform our 3 research questions (RQs).

Established theoretical frameworks, such as the health belief model (HBM), the technology acceptance model (TAM), and social cognitive theory, offer valuable constructs for understanding and motivating behavior change [8-11]; however, these models are rarely integrated into current chatbot designs. This underutilization may limit chatbots' capacity to enhance user engagement and sustain behavior change.

This review aims to synthesize qualitative insights into the following:

1. What are the experiences and perceptions of users and providers regarding chatbots in health education and behavior change?
2. What are the key facilitators and barriers affecting the adoption of healthcare chatbots?
3. How do theoretical frameworks guide the design and implementation of these chatbots?

Methods

Overview

This systematic review aligns with the meta-aggregation principles outlined by Lockwood et al [12] (adapted for qualitative evidence synthesis), which we have applied to the domain of health education and behavior change interventions delivered via conversation-based digital tools, including chatbots. This approach was chosen to explore in-depth user and provider experiences through a specific focus on qualitative data, directly addressing the RQs on perceptions, theoretical frameworks, and adoption facilitators and barriers. We adhered to the guidelines of the Joanna Briggs Institute (JBI) [13] for critical appraisal and reported our methods in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart and standard systematic review protocols. The review was registered with the Open Science Framework to ensure methodological transparency.

Search Strategy

We searched PubMed, the Cochrane Library, and ScienceDirect from January 1, 2018, to October 1, 2023, for English- or German-language, peer-reviewed qualitative or mixed methods studies. A medical librarian assisted in developing the search strategy to ensure comprehensive coverage. Our search syntax included terms such as "chatbot," "conversational agent," "digital health interventions," "qualitative," and "mixed methods." These terms were applied to titles, abstracts, and keywords to identify relevant studies examining conversation-based digital tools for health education or behavior change. The detailed search syntax is presented in Multimedia Appendix 1.

Study Selection

Two reviewers (SF and KEC) independently screened all titles and abstracts (Cohen κ=0.82), retrieving full texts when either deemed an article potentially relevant. Discrepancies were resolved by discussion, obviating the need for a third reviewer. Mixed methods studies were included if they provided relevant qualitative data, which were prioritized over quantitative findings during synthesis to ensure a comprehensive understanding of user and provider experiences. The final inclusion criteria limited articles to those presenting qualitative data on conversation-based digital tools in health education or behavior change, with participants or providers offering experiential insights. References were tracked and managed via EndNote (Clarivate) and manual cross-checking. An overview of the complete selection process is shown in Multimedia Appendix 2, along with the PRISMA flow diagram.

Inclusion and Exclusion Criteria

Studies were selected on the basis of the detailed inclusion and exclusion criteria outlined in Table 1. In brief, we included qualitative or mixed methods studies published in English or German between 2018 and 2023 that investigated user or provider experiences with conversation-based digital tools designed for health education or behavior change. These tools encompass various designations, including "chatbots," "conversational agents," and "digital health assistants," all of which rely fundamentally on dialog-based engagement. This inclusive approach enables a comprehensive examination of automated conversational systems within health intervention contexts. The time frame (2018 - 2023) was selected to capture recent developments in chatbot technology, while the language restriction was based on the research team's proficiency. Studies were excluded if they lacked conversational functionality, did not provide qualitative data, or focused on domains unrelated to health education and behavior change. These criteria were systematically applied to ensure consistent study selection and reduce potential bias.

XSL•FO
RenderX

Table . Inclusion and exclusion criteria.

Criterion	Inclusion	Exclusion
Study type	Primary qualitative studies, mixed methods studies with qualitative components, and peer-reviewed articles	Purely quantitative studies, reviews, editorials, opinion pieces, conference abstracts, and non–peer-reviewed articles
Population	Patients and health consumers, health care professionals, adults (aged 18 years and older), adolescents (aged 12 - 19 years) with specific health conditions	Studies focused exclusively on children and studies without direct user and provider perspectives
Intervention	Chatbots and conversational agents for health education, chatbots for behavior change, artificial intelligence–driven health dialog systems	Mobile apps without conversational features, static health information systems, and noninteractive digital tools
Outcomes	User experiences, provider experiences, perceptions of chatbot use, qualitative feedback on usability, and implementation insights	Only quantitative outcomes, technical performance metrics, and cost-effectiveness analyses
Language	English and German	All other languages
Publication period	Published between 2018 and 2023	Studies published before 2018 and after 2023
Context	Health care settings, health education contexts, and behavior change interventions	Non–health care settings, commercial customer service, and general technology evaluation
Study design quality	Clear methodological description, appropriate data collection methods, and rigorous analysis procedures	Poor methodological quality, insufficient description of methods, and lack of ethical considerations

Data Extraction and Quality Appraisal

A standardized form was used to extract data on study design, participant demographics, chatbot features, outcomes, and key qualitative findings, ensuring consistency across studies. Two reviewers initially extracted data from 5 articles to establish a consistent approach before dividing the remaining articles; ongoing discussions resolved any ambiguities. The JBI Critical Appraisal Checklist for Qualitative Research was used to assess the methodological rigor and risk of bias (Multimedia Appendix 3). Rather than excluding studies failing to meet certain criteria, our appraisal informed the weighting of findings during synthesis.

Data Analysis

We conducted reflexive thematic analysis [14], beginning with immersive reading and coding of text fragments and progressing through 6 phases: (1) familiarization, (2) initial coding, (3) theme development, (4) review, (5) definition and naming, and (6) reporting. Subsequent iterative rounds of coding and discussion refined the analysis, and the themes were validated against the context of each article. To ensure rigor, investigator triangulation was used, with both reviewers achieving high interrater reliability (Cohen κ=0.82 for the screening process), along with regular analytic meetings and reflective journaling to mitigate researcher bias. The identified themes were organized in alignment with our 3 RQs, and representative participant quotes were integrated to illustrate key findings and enhance the credibility of the analysis.

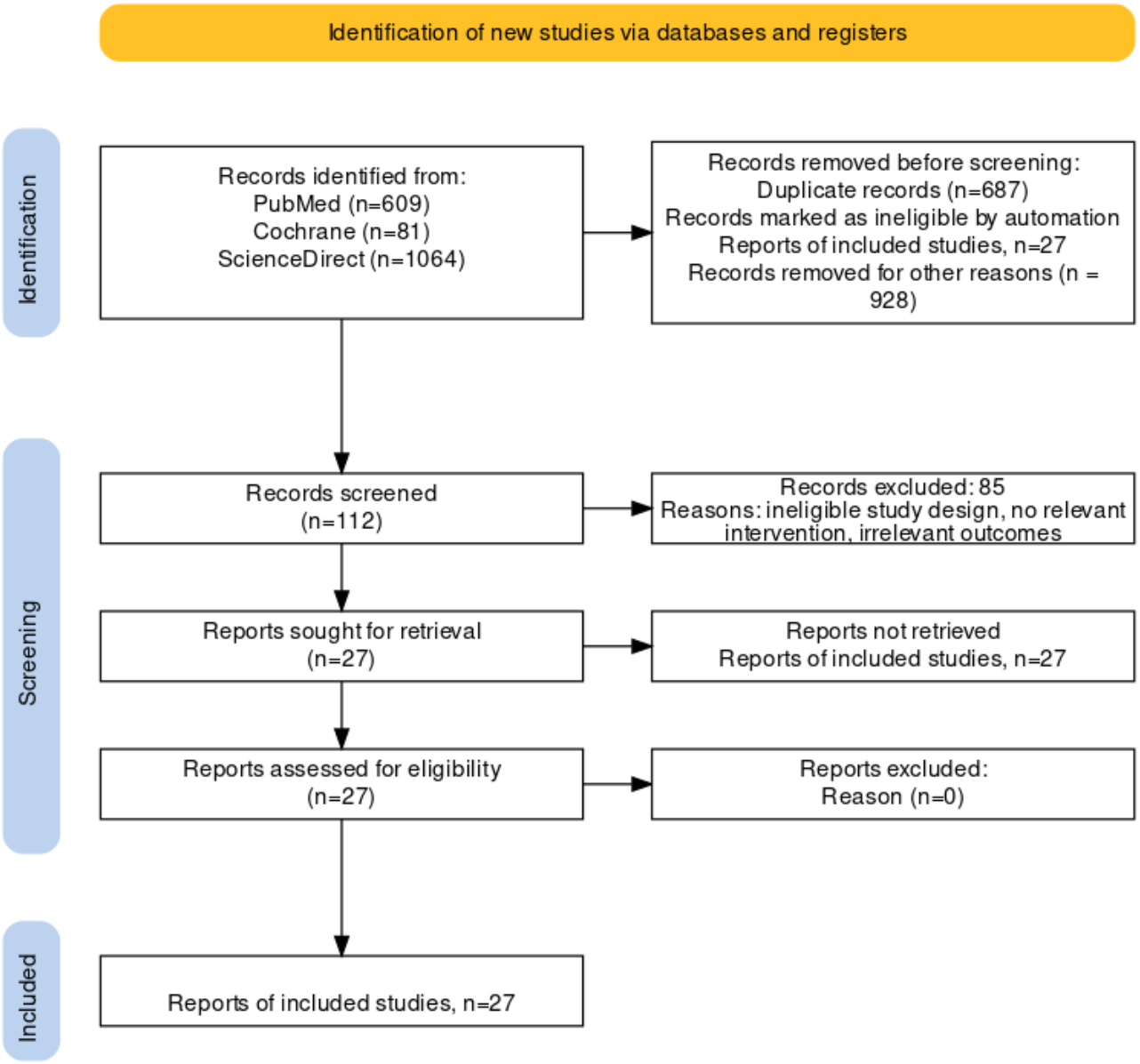
Results

Overview of Included Studies

From the 1754 records identified, 27 articles from 10 countries satisfied our criteria, encompassing 169 participants in exclusively qualitative studies and 10,802 in mixed methods studies, of whom 657 contributed qualitative data (Figure 1). The studies varied in focus, encompassing chatbot design, technological innovations, behavior change, mental health, and user experience research. Participant ages ranged from adolescents to older adults, highlighting the broad adaptability of chatbots across diverse health contexts. The participants included patients, health care providers, and general users, with studies primarily using qualitative interviews or mixed methods designs. For a detailed summary of the study characteristics, see Table 2.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of the study selection process.



XSL•FO
RenderX

Table . Summary of study characteristics.

First author	Title	Year	Country	Aim of the study	Study design	Participants
Baptista et al [15]	Acceptability of an embodied conversational agent for type 2 diabetes self-management education and support	2020	Australia and New Zealand	To evaluate the acceptability of an embodied conversational agent, named Laura, used to deliver diabetes self-management education and support in the My Diabetes Coach (MDC) app.	Randomized controlled trial and mixed methods	93 participants in the intervention arm
Barnett et al [16]	Enacting "more-than-human" care: clients' and counsellors' views on the multiple affordances of chatbots in alcohol and other drug counseling	2021	Australia	The paper explores the experiences and perceptions of clients and counsellors of chatbots used in web-based alcohol and other drug counseling. It focuses on how these technologies afford (provide) or constrain online care and client interactions.	Mixed	20 clients (10 men and 10 women) and 8 counsellors (5 men and 3 women)
Beaudry et al [17]	Getting ready for adult healthcare: designing a chatbot to coach adolescents with special health needs through the transitions of care	2019	United States	To engage adolescents with chronic medical conditions using a chatbot and text messaging platform to promote skill attainment in self-care and ease the transition from pediatric to adult-focused care	Pilot study	13 adolescents with chronic conditions
Biro et al [18]	The effects of a health care chatbot's complexity and persona on user trust, perceived usability, and effectiveness: mixed methods study	2023	Clemson University, Clemson, United States	To examine how various design elements impact the efficacy of a health care chatbot intended for educational purposes.	Mixed methods	71 university students aged 18 - 26 years (28 males, 43 females; 69% Caucasian, 11.3% African American, 19.7% Asian)
Boggiss et al [19]	Improving the well-being of adolescents with type 1 diabetes during the COVID-19 pandemic: qualitative study exploring acceptability and clinical usability of a self-compassion chatbot	2023	Aotearoa, New Zealand	To assess the effectiveness and acceptance of a self-compassion chatbot (COMPASS) designed for adolescents with type 1 diabetes.	Qualitative	19 adolescents with type 1 diabetes and 11 diabetes health care professionals
Chang et al [20]	Why would you use medical chatbots? Interview and survey	2022	Taiwan	To understand the factors influencing individuals' attitudes and intentions to use medical chatbots.	Two-stage mixed-method approach	20 interview participants and 205 survey respondents

XSL•FO
RenderX

First author	Title	Year	Country	Aim of the study	Study design	Participants
Chen et al [5]	Developing a heart transplantation self-management support mobile health app in Taiwan: qualitative study	2020	Taiwan	To investigate the information needs of post–heart transplantation patients and develop a preliminary framework for a mobile health app to support their self-management.	Qualitative	17 post–heart transplantation patients and 4 health professionals
Galvão Gomes da Silva et al [21]	Experiences of a motivational interview delivered by a robot: qualitative study	2018	United Kingdom	To explore the experiences of individuals who have engaged in a motivational interview delivered by a social robot.	Qualitative	20 participants from the School of Psychology's pool of research volunteers
Griffin et al [22]	A chatbot for hypertension self-management support: user-centered design, development, and usability testing	2023	United States	To incorporate users' feedback into Medicagent through usability testing, leveraging the Information-Motivation-Behavioral skills model and the model of medication self-management	Mixed	10 adults with hypertension
Griffin et al [23]	Information needs and perceptions of chatbots for hypertension medication self-management: a mixed methods study	2021	Chapel Hill, North Carolina, United States	To understand information needs and perceptions toward using a chatbot to support hypertension medication self-management	Convergent mixed methods design	15 adults with hypertension
Han et al [24]	Preliminary evaluation of a conversational agent to support self-management of individuals living with posttraumatic stress disorder: interview study with clinical experts	2023	United States	To conduct a preliminary evaluation of the PTSDialogue, focusing on its usability and acceptance from the viewpoint of clinical experts.	Qualitative	10 clinical experts with experience in posttraumatic stress disorder care
Hurmuz et al [25]	Evaluation of a digital coaching system eHealth intervention: a mixed methods observational cohort study in the Netherlands	2022	The Netherlands	To assess the use, user experience, and potential health effects of a conversational agent-based eHealth platform among older adults.	Observational cohort study	51 older adults, 70.6% female, and average age of 65 years
Kornfield et al [26]	A text messaging intervention to support the mental health of young adults: user engagement and feedback from a field trial of an intervention prototype	2023	United States	To investigate young adults' engagement and experiences with a prototype of an interactive text messaging program for managing mental health concerns.	User-centered design	48 individuals (66.7% female and 33.3% male)

XSL•FO
RenderX

First author	Title	Year	Country	Aim of the study	Study design	Participants
Lin et al [27]	Exploring pictorial health education tools for long-term home care: a qualitative perspective	2020	Taiwan	To explore a theoretical framework for developing pictorial health education tools to enhance communication between medical and nonmedical staff in home care.	Grounded theory	6 designers, 5 medical staff (hospital director, supervisor, nurses), and 8 groups of home caregivers for intubated patients
Ly et al [28]	A fully automated conversational agent for promoting mental well-being: a pilot RCT using mixed methods	2017	Sweden	To map and synthesize qualitative evidence on the use of chatbots in health education and behavioral change settings.	Mixed methods	9 individuals, women (n=4) and men (n=5), mean age of 28.8 years
Mash et al [29]	Evaluating the implementation of the GREAT4 Diabetes WhatsApp chatbot to educate people with type 2 diabetes during the COVID-19 pandemic: convergent mixed methods study	2022	South Africa	To assess the implementation of the GREAT4Diabetes chatbot, focusing on adoption, appropriateness, acceptability, and other implementation outcomes.	Mixed methods	8158 people connected with the chatbot
Nadarzynski et al [30]	Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: a mixed-methods study	2019	United Kingdom	To assess the acceptability of artificial intelligence–driven health chatbots and uncover challenges and driving factors impacting their use.	Mixed methods	29 university students for interviews and 215 individuals for survey
Papadopoulos et al [7]	Socially assistive robots in health and social care: acceptance and cultural factors. Results from an exploratory international internet-based survey	2022	United Kingdom	To explore registered nurses' and midwives' views on socially assistive robots and the impact of cultural dimensions on their acceptance.	Exploratory, cross-sectional, and descriptive study	1341 participants from 19 countries
Roman et al [31]	"Hey assistant, how can I become a donor?" The case of a conversational agent designed to engage people in blood donation	2020	Brazil	To develop and assess a conversational agent aimed at engaging people in blood donation.	User experience assessment study	50 participants (16 men and 34 women)
Schmidlen et al [32]	Patient assessment of chatbots for the scalable delivery of genetic counseling	2019	United States	To gather feedback on chatbots as new communication tools for facilitating genetic counseling.	Qualitative	62 participants in 6 focus groups

XSL•FO
RenderX

First author	Title	Year	Country	Aim of the study	Study design	Participants
Scholten et al [33]	An empirical study of a pedagogical agent as an adjunct to an eHealth self-management intervention	2019	Netherlands	To investigate the support level technology, specifically a pedagogical agent, can provide to users of a self-guided positive psychology psycho-education.	Between-subjects experiment	230 psychology students
Siglen et al [34]	Ask Rosa – the making of a digital genetic conversation tool, a chatbot, about hereditary breast and ovarian cancer	2022	Norway	To design and develop a pilot version of the Rosa chatbot as a reliable source of information on hereditary breast and ovarian cancer.	Participatory methodology	58 participants including patient representatives, IT engineers, and medical staff
Svendsen et al [35]	One size does not fit all: participants' experiences of the selfBACK app to support self management of low back pain	2022	Denmark and Norway	To investigate patients' experiences with the selfBACK app for self-management of low back pain.	Qualitative interview study	26 participants from the selfBACK trial
Swendeman et al [36]	Feasibility and acceptability of mobile phone self-monitoring and automated feedback to enhance telephone coaching for people with risky substance use	2021	United States	To evaluate the feasibility and acceptability of mobile-phone delivered self-monitoring and feedback integrated into a health coaching intervention for risky drug use.	Mixed methods	20 participants with risky substance use
Ter Stal et al [37]	An embodied conversational agent in an eHealth self-management intervention for chronic obstructive pulmonary disease and chronic heart failure	2021	The Netherlands	To investigate users' perceptions of an embodied conversational agent's design in a real-life setting.	Mixed methods design	11 participants with chronic obstructive pulmonary disease and congestive heart failure
Escobar-Viera et al [38]	A chatbot-delivered intervention for optimizing social media use and reducing perceived isolation among rural-living LGBTQ+ youth	2023	United States	To evaluate REALbot, a chatbot designed to deliver an educational program to rural living LGBTQ+ youth to reduce perceived isolation.	Exploratory pilot study	20 adolescents aged 14 - 20 years old, identifying as LGBTQ+
Wang et al [39]	Revealing the complexity of users' intention to adopt healthcare chatbots: a mixed-method analysis of antecedent condition configurations	2023	China	To understand users' attitudes and experiences with online health care chatbots.	Mixed methods	347 survey respondents and 12 participants in semistructured interviews

RQ1: User and Provider Experiences

RQ1 explores the general experiences and perceptions of users and providers regarding chatbots in health education and behavior change, which is distinct from the specific adoption factors addressed in RQ2.

XSL•FO
RenderX

Empathy and Emotional Connection

The participants valued human-like interactions in chatbots, particularly for emotional support. Ly et al [28] noted that participants perceived chatbots as living characters capable of forming relationships, emphasizing their potential to mimic human guidance and empathy. However, providers expressed skepticism about chatbots' empathetic capabilities. Nadarzynski et al [30] raised significant concerns, stating that chatbots are not capable of empathy, notably in recognizing users' emotional states and tailoring responses accordingly, potentially compromising user engagement.

Trust and Privacy Concerns

Privacy and data security are significant concerns affecting trust in chatbots. A participant in the study by Nadarzynski et al [30] stated, "Some things are confidential and you wouldn't just type it on the internet. You would want the confidentiality of [a doctor]," highlighting concerns over the chatbot's ability to maintain the confidentiality expected in traditional medical interactions. These concerns highlight the tension between accessibility and data protection.

RQ2: Facilitators and Barriers to Adoption

RQ2 investigates specific factors influencing the adoption of healthcare chatbots, distinguishing them from the broader experiences covered in RQ1.

Accessibility and Ease of Use

Round-the-clock availability and intuitive design were key facilitators. Boggiss et al [19] explicitly stated that chatbots offer unique advantages, including "24-hour availability, accessibility, remote delivery, scalability, and real-time personalized responses."

Personalization

Personalization enhanced adoption. Chang et al [20] highlighted user desire for personalized health information, rather than generic content, to enhance relevance and usefulness. Boggiss et al [19] further supported this, noting that nearly all users wanted to customize chatbot interactions.

Trust Deficits

Trust issues, particularly those related to privacy, hinder adoption. Barnett et al [16] expressed concerns about confidentiality, noting uncertainty regarding who has access to chatbot interactions and thus preferring traditional doctor–patient confidentiality.

Technical Glitches

Technical issues such as app freezes and connectivity problems were barriers. Mash et al [29] explicitly described technical problems such as repeated messages and system overloads affecting message dissemination to users, significantly disrupting user experiences. Svendsen et al [35] reported that faulty synchronization and login issues impact chatbot use.

Cultural Misalignment

Chatbots' inability to interpret cultural nuances posed obstacles. Papadopoulos et al [7] highlighted the chatbot's difficulty in understanding cultural contexts, stating, "It may interfere with the patient, the patient may not be able to explain the problem to the robot because the robot does not know how people grow in cultures in places."

RQ3: Role of Theoretical Frameworks

RQ3 examines how theoretical frameworks of health behavior change and how technology acceptance guides chatbot design and implementation.

Underutilization of Behavior Change Models

While some studies have integrated frameworks such as the HBM, TAM, or social cognitive theory, many have not, missing opportunities to increase user motivation. Griffin et al [22] explicitly used the IDEAS (Integrate, Design, Assess, and Share) framework and information-motivation-behavioral skills model to inform chatbot design for hypertension management, demonstrating the positive impact of theory-driven approaches. Conversely, Nadarzynski et al [30] identified the absence of a clear theoretical foundation as problematic, noting risks such as incorrect self-diagnosis and inadequate care from theoretically ungrounded chatbots.

Integration With Theoretical Constructs

Studies incorporating theoretical constructs reported improved engagement. Griffin et al [22] positively illustrated the integration of behavioral theories, notably the IDEAS framework and the information-motivation-behavioral skills model, effectively informing users and motivating behavior change.

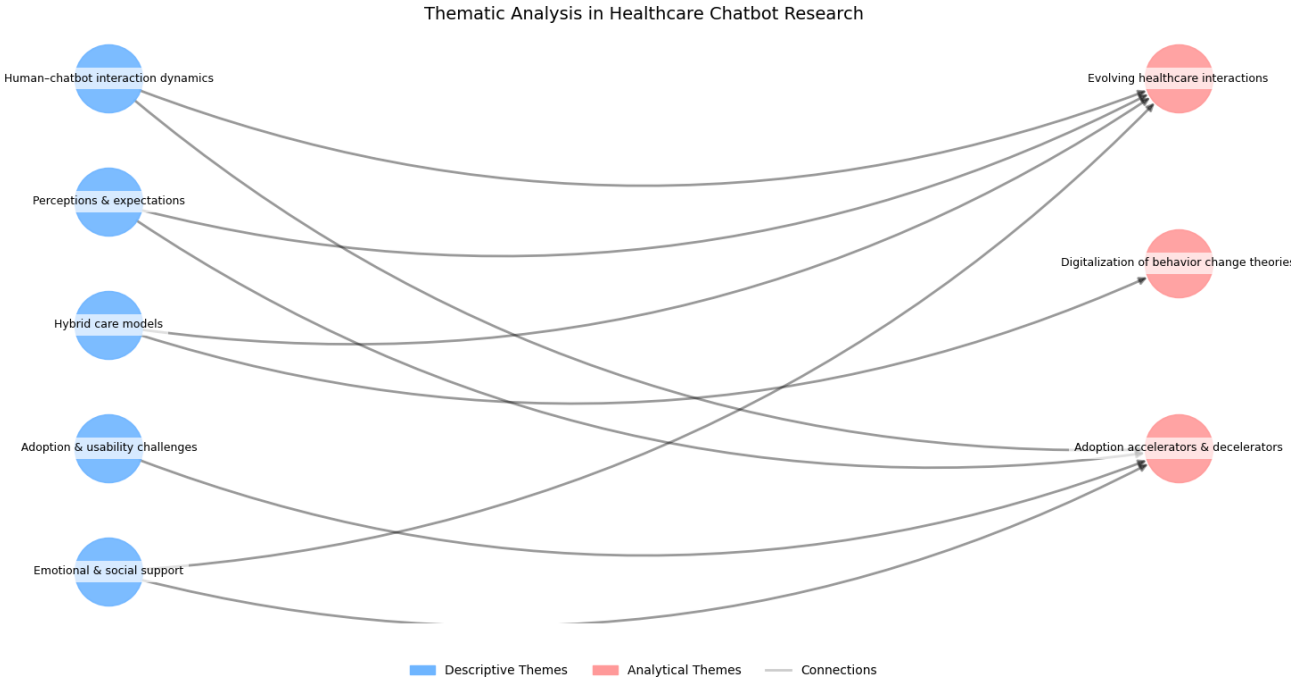
Additional Analytical Themes

This section synthesizes insights from RQ1-RQ3 to illustrate how themes of empathy, trust, and theoretical grounding converge in the experiences of users and providers. Chatbots were seen as evolving interfaces capable of triaging minor questions before escalation, as highlighted by participants in Nadarzynski et al [30], who described chatbots effectively directing users either to immediate care or reassurance. They were also viewed as expansions of social networks, offering psychosocial support for isolated populations. Barnett et al [16] emphasized that chatbots might attract users needing extreme privacy and confidentiality, potentially engaging more individuals in treatment. By transparently integrating artificial intelligence (AI)–driven personalization with behavior-change theories, these digital tools not only strengthened emotional connections but also addressed key adoption barriers, creating a unified and trustworthy health-education experience.

Figure 2 provides a visual representation of the descriptive and analytical themes identified in our review, illustrating how user experiences (eg, empathy, trust) intersect with theoretical frameworks and adoption factors.

XSL•FO
RenderX

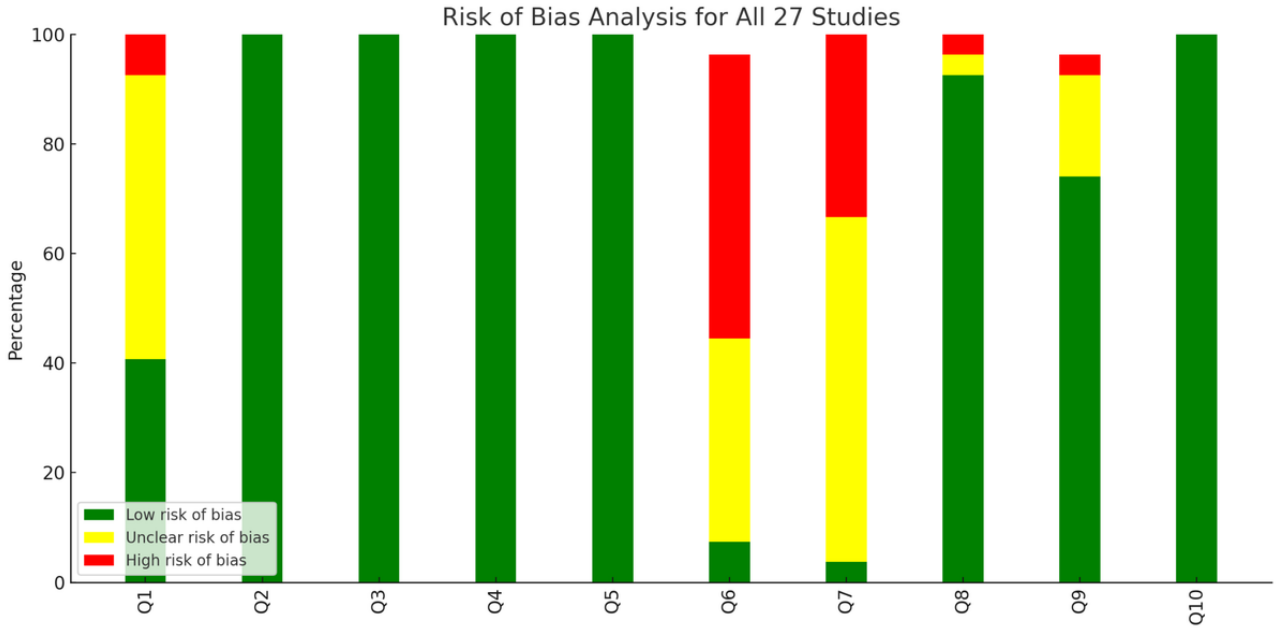
Figure 2. Thematic interconnections in health care chatbot research.



Methodological Quality and Risk of Bias

The methodological rigor of the 27 studies varied. Many aligned studies aim with qualitative methods, but some lack transparency in reporting researcher biases or ethical considerations. Using the JBI Critical Appraisal Checklist, 13 studies were at low risk of bias, 12 were at unclear risk, and 2 were at high risk, primarily due to limited ethical reporting or reflexivity. Figure 3 provides a summary of the risk of bias assessment for the included studies.

Figure 3. Summary of the risk of bias assessment for the 27 included studies.



Discussion

Principal Findings

Our synthesis reveals robust enthusiasm for chatbots as tools for health education and behavior change. In addressing RQ1 (User and Provider Experiences), we found that users value empathy and emotional support, yet they express concerns about trust and privacy. For RQ2 (Facilitators and Barriers to Adoption), 24/7 accessibility, ease of use, and personalization emerged as key facilitators, whereas trust deficits, technical glitches, and cultural misalignment were identified as major barriers. With respect to RQ3 (Role of Theoretical Frameworks), the underutilization of models such as the HBM or TAM in chatbot design has emerged as a critical gap, limiting the potential for enhanced motivational and behavioral outcomes.

Earlier meta-analyses, such as those by Laranjo et al [2], reported that chatbots can improve health outcomes, but our

XSL•FO
RenderX

qualitative review complements this by emphasizing user-centered dimensions such as trust, empathy, and cultural fit—factors often overlooked in quantitative studies. For example, Nadarzynski et al [30] identified privacy as a barrier to adoption, a concern supported by our findings and detailed in the results section. Our review also identified specific barriers, such as technical errors [19], privacy gaps [16,30], and cultural mismatches [7], which previous studies have not extensively addressed. This qualitative depth highlights that chatbot success depends not only on technical functionality but also on resonating with users personally and culturally—an insight crucial for developers and policymakers aiming to foster trust and adoption across diverse populations.

Our findings offer guidance for developers to embed behavior change theories and culturally adaptive content into chatbot design. Developers should prioritize incorporating prompts that address perceived barriers (eg, tailored reminders to overcome reluctance) and foster personalized feedback loops on the basis of theoretical constructs such as the HBM or information-motivation-behavioral skills model [22]. Developers should also embed culturally relevant content and empathetic responses to build trust and engagement. This systematic integration of theories is essential for optimizing chatbot effectiveness, addressing a gap in current designs.

Healthcare providers can integrate chatbots as triage tools to complement face-to-face care and reduce routine workloads. Providers should advocate for chatbots that support, rather than replace, human interaction. Training healthcare staff in how to use chatbot systems effectively can improve both patient outcomes and provider acceptance. Our qualitative insights suggest that providers are more likely to support chatbot integration when they perceive these tools as enhancing, not threatening, their professional roles.

Policymakers should promote transparent privacy standards and fund theory-based chatbot initiatives. Regulating data security and ensuring transparent privacy policies are essential to fostering user trust. Policymakers should support the integration of theoretical frameworks into digital health tools to ensure that they are evidence-based and effective. Funding initiatives promoting culturally sensitive chatbot designs could further support adoption. The emphasis on cultural relevance and ethical considerations in our findings points to a broader need for policies that prioritize user trust and inclusivity in digital health innovations.

Strengths and Limitations

A primary strength is the focus on qualitative evidence, offering granular insights into the emotional, cultural, and trust-based dimensions of chatbot use. The broad timescale (2018 - 2023) captures recent innovations, including AI-based chatbots. However, the 2018 - 2023 window may exclude older studies or those in languages beyond English/German, limiting the diversity of perspectives. Additionally, the rapid evolution of AI-driven chatbots, particularly large language models (LLMs), means that our results may not fully capture their latest affordances. The inclusion of broader conversational technologies [e.g., Biro et al [18], ; Chang et al [20], ; Svendsen et al [35],] to offer holistic insights may reduce the specificity of findings for strictly defined chatbots. Selection bias may have influenced the findings, as many studies recruited tech-savvy participants, potentially overrepresenting positive experiences. The heterogeneity of study designs and limited focus on specific populations (eg, older adults) also constrain generalizability.

Future Directions

Future research should examine how user perceptions (trust, empathy) evolve over time (RQ1), assess cultural influences on adoption barriers/facilitators (RQ2), and pilot systematic integration of theories into chatbot workflows (RQ3). For example, longitudinal studies should explore adherence and user engagement across time, whereas cross-cultural research could illuminate how regional dialects and health beliefs affect adoption. Methodologically, future work should strengthen transparency through preregistration and the use of qualitative reporting frameworks (eg, Consolidated Criteria for Reporting Qualitative Research), enhance ethical safeguards, adopt inclusive recruitment strategies, ensure human support during chatbot interactions, and establish clear data usage and privacy protocols. Collaborative efforts among developers, clinicians, behavioral scientists, and ethicists are essential for refining chatbot designs, ensuring data security, and embedding effective behavior

Conclusion

This systematic review highlights robust interest in chatbots for health education and behavior change while revealing significant challenges related to privacy, cultural alignment, and theoretical underpinnings. By addressing RQ1, RQ2, and RQ3, we identified key facilitators of and barriers to chatbot adoption and underscored the importance of embedding behavior change theories into chatbot design. Although chatbots extend the reach of healthcare by providing empathetic, accessible support, their full potential remains constrained without systematic integration of behavior change models and rigorous data safeguards. Developers should prioritize user trust, cultural relevance, and theoretical grounding to enhance chatbot adoption and effectiveness. Addressing these challenges through interprofessional collaboration, ethical oversight, and ongoing research will be crucial for realizing chatbots' transformative potential in global health systems. This review's qualitative insights provide a foundation for designing innovative, user-centered chatbots that address diverse populations' complex needs in digital health.

Acknowledgments

We thank Danube Private University for institutional support. We also appreciate the medical librarians who assisted with our search strategy.

XSL•FO
RenderX

This study was conducted with the internal resources of Danube Private University. The funder played no role in the design, interpretation, or publication decisions.

Data Availability

All data (search queries, extracted study details, quality appraisals) are available in the Multimedia Appendices or upon request from the corresponding author. No additional restrictions apply.

Authors' Contributions

Conceptualization: AC Methodology: SF, AC Data Curation: SF Formal Analysis: SF, AC Writing—Original Draft: SF Writing—Review & Editing: SF, AC Supervision: AC Both authors approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search Strategy.
[DOCX File, 20 KB - humanfactors_v12i1e60205_app1.docx]

Multimedia Appendix 2
Comprehensive overview of selected studies.
[PDF File, 173 KB - humanfactors_v12i1e60205_app2.pdf]

Multimedia Appendix 3
Risk of bias assessment.
[PDF File, 155 KB - humanfactors_v12i1e60205_app3.pdf]

Checklist 1
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 checklist.
[PDF File, 232 KB - humanfactors_v12i1e60205_app4.pdf]

References

1. Montenegro JLZ, da Costa CA, da Rosa Righi R. Survey of conversational agents in health. Expert Syst Appl 2019 Sep;129:56-67. [doi: 10.1016/j.eswa.2019.03.054]
2. Laranjo L, Dunn AG, Tong HL, et al. Conversational agents in healthcare: a systematic review. J Am Med Inform Assoc 2018 Sep 1;25(9):1248-1258. [doi: 10.1093/jamia/ocy072] [Medline: 30010941]
3. Aggarwal A, Tam CC, Wu D, Li X, Qiao S. Artificial intelligence-based chatbots for promoting health behavioral changes: systematic review. J Med Internet Res 2023 Feb 24;25(1):e40789. [doi: 10.2196/40789] [Medline: 36826990]
4. Bickmore TW, Pfeifer LM, Paasche-Orlow MK. Using computer agents to explain medical documents to patients with low health literacy. Patient Educ Couns 2009 Jun;75(3):315-320. [doi: 10.1016/j.pec.2009.02.007] [Medline: 19297116]
5. Chen YW, Wei J, Chen HL, Cheng CH, Hou IC. Developing a heart transplantation self-management support mobile health app in Taiwan: qualitative study. JMIR Mhealth Uhealth 2020 Aug 19;8(8):e18999. [doi: 10.2196/18999] [Medline: 32812883]
6. Milne-Ives M, de Cock C, Lim E, et al. The effectiveness of artificial intelligence conversational agents in health care: systematic review. J Med Internet Res 2020 Oct 22;22(10):e20346. [doi: 10.2196/20346] [Medline: 33090118]
7. Papadopoulos I, Wright S, Koulouglioti C, et al. Socially assistive robots in health and social care: acceptance and cultural factors. Results from an exploratory international online survey. Jpn J Nurs Sci 2023 Apr;20(2):e12523. [doi: 10.1111/jjns.12523] [Medline: 36732396]
8. Rosenstock IM. The health belief model and preventive health behavior. Health Educ Monogr 1974 Dec;2(4):354-386. [doi: 10.1177/109019817400200405]
9. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
10. Glanz K, Bishop DB. The role of behavioral science theory in development and implementation of public health interventions. Annu Rev Public Health 2010;31(1):399-418. [doi: 10.1146/annurev.publhealth.012809.103604] [Medline: 20070207]
11. Prochaska JO, DiClemente CC. Stages of change in the modification of problem behaviors. Prog Behav Modif 1992;28(183–218):183-218. [Medline: 1620663]

XSL•FO
RenderX

12. Lockwood C, Munn Z, Porritt K. Qualitative research synthesis: methodological guidance for systematic reviewers utilizing meta-aggregation. Int J Evid Based Healthc 2015 Sep;13(3):179-187. [doi: 10.1097/XEB.0000000000000062] [Medline: 26262565]
13. JBI critical appraisal tools. JBI. URL: https://jbi.global/critical-appraisal-tools [accessed 2024-01-16]
14. Braun V, Clarke V. Toward good practice in thematic analysis: avoiding common problems and be(com)ing a knowing researcher. Int J Transgend Health 2023;24(1):1-6. [doi: 10.1080/26895269.2022.2129597] [Medline: 36713144]
15. Baptista S, Wadley G, Bird D, Oldenburg B, Speight J, My Diabetes Coach Research Group. Acceptability of an embodied conversational agent for type 2 diabetes self-management education and support via a smartphone app: mixed methods study. JMIR Mhealth Uhealth 2020 Jul 22;8(7):e17038. [doi: 10.2196/17038] [Medline: 32706734]
16. Barnett A, Savic M, Pienaar K, et al. Enacting "more-than-human" care: clients' and counsellors' views on the multiple affordances of chatbots in alcohol and other drug counselling. Int J Drug Policy 2021 Aug;94:102910. [doi: 10.1016/j.drugpo.2020.102910] [Medline: 33059955]
17. Beaudry J, Consigli A, Clark C, Robinson KJ. Getting ready for adult healthcare: designing a chatbot to coach adolescents with special health needs through the transitions of care. J Pediatr Nurs 2019;49:85-91. [doi: 10.1016/j.pedn.2019.09.004] [Medline: 31644960]
18. Biro J, Linder C, Neyens D. The effects of health care chatbot's complexity and persona on user trust, perceived usability, and effectiveness: mixed methods study. JMIR Hum Factors 2023 Feb 1;10:e41017. [doi: 10.2196/41017] [Medline: 36724004]
19. Boggiss A, Consedine N, Hopkins S, et al. Improving the well-being of adolescents with type 1 diabetes during the COVID-19 pandemic: qualitative study exploring acceptability and clinical usability of a self-compassion chatbot. JMIR Diabetes 2023 May 5;8:e40641. [doi: 10.2196/40641] [Medline: 36939680]
20. Chang IC, Shih YS, Kuo KM. Why would you use medical chatbots? Interview and survey. Int J Med Inform 2022 Sep;165:104827. [doi: 10.1016/j.ijmedinf.2022.104827] [Medline: 35797921]
21. Galvão Gomes da Silva J, Kavanagh DJ, Belpaeme T, Taylor L, Beeson K, Andrade J. Experiences of a motivational interview delivered by a robot: qualitative study. J Med Internet Res 2018 May 3;20(5):e116. [doi: 10.2196/jmir.7737] [Medline: 29724701]
22. Griffin AC, Khairat S, Bailey SC, Chung AE. A chatbot for hypertension self-management support: user-centered design, development, and usability testing. JAMIA Open 2023 Oct;6(3):ooad073. [doi: 10.1093/jamiaopen/ooad073] [Medline: 37693367]
23. Griffin AC, Xing Z, Mikles SP, et al. Information needs and perceptions of chatbots for hypertension medication self-management: a mixed methods study. JAMIA Open 2021 Apr;4(2):ooab021. [doi: 10.1093/jamiaopen/ooab021] [Medline: 33898936]
24. Han HJ, Mendu S, Jaworski BK, Owen JE, Abdullah S. Preliminary evaluation of a conversational agent to support self-management of individuals living with posttraumatic stress disorder: interview study with clinical experts. JMIR Form Res 2023 May 29;7:e45894. [doi: 10.2196/45894] [Medline: 37247220]
25. Hurmuz MZM, Jansen-Kosterink SM, Beinema T, Fischer K, Op den Akker H, Hermens HJ. Evaluation of a virtual coaching system eHealth intervention: a mixed methods observational cohort study in the Netherlands. Internet Interv 2022 Mar;27:100501. [doi: 10.1016/j.invent.2022.100501] [Medline: 35198411]
26. Kornfield R, Stamatis CA, Bhattacharjee A, et al. A text messaging intervention to support the mental health of young adults: user engagement and feedback from a field trial of an intervention prototype. Internet Interv 2023 Dec;34:100667. [doi: 10.1016/j.invent.2023.100667] [Medline: 37746639]
27. Lin FS, Shi HC, Fang KT. Exploring pictorial health education tools for long-term home care: a qualitative perspective. Healthcare (Basel) 2020 Jul 9;8(3):205. [doi: 10.3390/healthcare8030205] [Medline: 32659998]
28. Ly KH, Ly AM, Andersson G. A fully automated conversational agent for promoting mental well-being: a pilot RCT using mixed methods. Internet Interv 2017 Dec;10:39-46. [doi: 10.1016/j.invent.2017.10.002] [Medline: 30135751]
29. Mash R, Schouw D, Fischer AE. Evaluating the implementation of the GREAT4Diabetes WhatsApp chatbot to educate people with type 2 diabetes during the COVID-19 pandemic: convergent mixed methods study. JMIR Diabetes 2022 Jun 24;7(2):e37882. [doi: 10.2196/37882] [Medline: 35537057]
30. Nadarzynski T, Miles O, Cowie A, Ridge D. Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: a mixed-methods study. Digit Health 2019;5:2055207619871808. [doi: 10.1177/2055207619871808] [Medline: 31467682]
31. Roman MK, Bellei EA, Biduski D, Pasqualotti A, De Araujo C, De Marchi ACB. "Hey assistant, how can I become a donor?" The case of a conversational agent designed to engage people in blood donation. J Biomed Inform 2020 Jul;107:103461. [doi: 10.1016/j.jbi.2020.103461] [Medline: 32504669]
32. Schmidlen T, Schwartz M, DiLoreto K, Kirchner HL, Sturm AC. Patient assessment of chatbots for the scalable delivery of genetic counseling. J Genet Couns 2019 Dec;28(6):1166-1177. [doi: 10.1002/jgc4.1169] [Medline: 31549758]
33. Scholten MR, Kelders SM, Van Gemert-Pijnen J. An empirical study of a pedagogical agent as an adjunct to an eHealth self-management intervention: what modalities does it need to successfully support and motivate users? Front Psychol 2019;10:1063. [doi: 10.3389/fpsyg.2019.01063] [Medline: 31143152]

XSL•FO
RenderX

34. Siglen E, Vetti HH, Lunde ABF, et al. Ask Rosa - the making of a digital genetic conversation tool, a chatbot, about hereditary breast and ovarian cancer. Patient Educ Couns 2022 Jun;105(6):1488-1494. [doi: 10.1016/j.pec.2021.09.027] [Medline: 34649750]
35. Svendsen MJ, Nicholl BI, Mair FS, Wood K, Rasmussen CDN, Stochkendahl MJ. One size does not fit all: participants' experiences of the selfBACK app to support self-management of low back pain-a qualitative interview study. Chiropr Man Therap 2022 Oct 3;30(1):41. [doi: 10.1186/s12998-022-00452-2] [Medline: 36192724]
36. Swendeman D, Sumstine S, Aguilar E, Gorbach PM, Comulada WS, Gelberg L. Feasibility and acceptability of mobile phone self-monitoring and automated feedback to enhance telephone coaching for people with risky substance use: the QUIT-mobile pilot study. J Addict Med 2021 Apr 1;15(2):120-129. [doi: 10.1097/ADM.0000000000000707] [Medline: 32732684]
37. Ter Stal S, Sloots J, Ramlal A, Op den Akker H, Lenferink A, Tabak M. An embodied conversational agent in an eHealth self-management intervention for chronic obstructive pulmonary disease and chronic heart failure: exploratory study in a real-life setting. JMIR Hum Factors 2021 Nov 4;8(4):e24110. [doi: 10.2196/24110] [Medline: 34734824]
38. Escobar-Viera CG, Porta G, Coulter RWS, Martina J, Goldbach J, Rollman BL. A chatbot-delivered intervention for optimizing social media use and reducing perceived isolation among rural-living LGBTQ+ youth: development, acceptability, usability, satisfaction, and utility. Internet Interv 2023 Dec;34:100668. [doi: 10.1016/j.invent.2023.100668] [Medline: 37746640]
39. Wang X, Luo R, Liu Y, Chen P, Tao Y, He Y. Revealing the complexity of users' intention to adopt healthcare chatbots: a mixed-method analysis of antecedent condition configurations. Inf Process Manag 2023 Sep;60(5):103444. [doi: 10.1016/j.ipm.2023.103444]

Abbreviations

AI: artificial intelligence
HBM: health belief model
IDEAS: Integrate, Design, Assess, and Share
JBI: Joanna Briggs Institute
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RQ: research question
TAM: technology acceptance model

Edited by A Kushniruk; submitted 04.05.24; peer-reviewed by A Luke MacNeill, A Arbabisarjou, D Swendeman, D Neyens, E Siglen, IC Hou; revised version received 20.04.25; accepted 20.04.25; published 13.06.25.

Please cite as:
Choi K E(, Fitzek S
User and Provider Experiences With Health Education Chatbots: Qualitative Systematic Review
JMIR Hum Factors 2025;12:e60205
URL: https://humanfactors.jmir.org/2025/1/e60205
doi:10.2196/60205

© K yung-Eun (Anna) Choi, Sebastian Fitzek. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 13.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

User Experience of Older Adults With an Age-Friendly Transportation Planning E-Tool: Scoping Review

Sara Bahrampoor Givi[1,2], PhD; Mireille Gagnon-Roy[1,3], OTR, PhD; Hélène Pigot[1,4], CEng, PhD; Véronique Provencher[1,5], OTR, PhD

[1]Research Center on Aging, Centre intégré universitaire de santé et de services sociaux de l'Estrie - Centre hospitalier universitaire de Sherbrooke, 1036 Rue Belvédère S, Sherbrooke, QC, Canada

[2]Faculty of Letters and Human Sciences, Université de Sherbrooke, Sherbrooke, QC, Canada

[3]Faculty of Medicine and Health Sciences, Université de Sherbrooke, Sherbrooke, QC, Canada

[4]Department of Computer Science, Université de Sherbrooke, Sherbrooke, QC, Canada

[5]School of Rehabilitation, Université de Sherbrooke, Sherbrooke, QC, Canada

Corresponding Author:
Sara Bahrampoor Givi, PhD
Research Center on Aging, Centre intégré universitaire de santé et de services sociaux de l'Estrie - Centre hospitalier universitaire de Sherbrooke, 1036 Rue Belvédère S, Sherbrooke, QC, Canada

Abstract

Background: Aging is associated with various challenges, especially concerning mobility. Transportation planning e-tools are currently available to provide older adults with real-time travel information and help them choose trip options. However, many older adults find them challenging to use, as they are not tailored to their specific needs, such as lack of accessibility to the different means of transportation. It is necessary to identify knowledge gaps about design based on older adults' experience using transportation planning e-tools.

Objective: This study aims to identify knowledge gaps regarding the user experience design and its evaluation for older adults using transportation planning e-tools.

Methods: A scoping review of the scientific literature was conducted, based on Arksey and O'Malley's guidelines. The search covers sources published in English from January 2002 to October 2022 through 7 scientific databases, including MEDLINE, AgeLine, CINAHL, SCOPUS, ProQuest, IEEE Explore, and TRID (Transportation Research International Documentation), and was updated in October 2023. Data selection and extraction were performed by the first author and were co-validated by 2 co-authors. The identified sources were analyzed based on source characteristics (authors, year of publication, title of the article, source of article, country, and context of the studies), the purpose of the studies (objectives and study orientation), and the mapping of the methodology and evaluation of user experience (design approach, setting, type of data collection and analysis, type of usability evaluation, and sample size). Both descriptive-analytical methods and thematic analysis were used to analyze the categorized data.

Results: Overall, 1905 sources were identified through databases, and 40 sources were selected for full-text analysis. Data analysis revealed in recent years that there has been significant growth in e-tools designed for older adults, but only 2 studies were related to the field of transportation. In total, 12 studies aimed to evaluate user experience, and 22 studies focused on a user-centered design approach. Most of these studies (n=31) were carried out in a laboratory setting, using summative usability evaluation (user-based testing). The System Usability Scale (SUS) was the most prevalent tool (15 studies) to measure user efficiency and satisfaction. However, no studies have been found that specifically aim to improve the mobility experience of older adults using an age-friendly transportation planning e-tool in a real-life context.

Conclusions: There is a lack of studies assessing older adults' experience when using transportation planning e-tools in real-life situations. To bridge this gap, a participatory approach is necessary to better consider the needs of older adults in a real-life context and create age-friendly design guidelines for the development of transportation planning e-tools. This will not only enhance their user experience in trip planning but also promote their social engagement by improving their mobility.

(JMIR Hum Factors 2025;12:e63273) doi:10.2196/63273

KEYWORDS

older adults; transportation; user experience; E-tool; review; design

XSL•FO
RenderX

Introduction

Background

The population of individuals aged 65 years and above is projected to increase from 900 million in 2015 to 2 billion by 2050 [1]. Recent data reported that 37.8% of Canadians over the age of 65 years live with a disability, which increases mobility issues in their daily travel [2]. Indeed, some age-related changes may impede the use of transportation for many older adults to participate in meaningful activities. Decreased vision [3-5] and hearing loss [6,7] make it difficult to see other road users (especially at night), recognize traffic signs (reading text and distinguishing colors) while driving, and hear high-frequency sounds when crossing the street. In addition, decreased flexibility, range of motion, and grip strength can limit driving control and impair driving safety [7], which makes it difficult for them to get around. Therefore, it is necessary to support the ability of older adults to know and use transportation options that are tailored to their needs so that they continue to be able to leave their homes, which is an essential characteristic of social participation [8].

Transportation-planning e-tools are currently available to provide older adults with information about transportation and trip options. These tools can take the form of websites or applications designed to assist users in planning, personalizing, and managing their travel. However, many older adults find them challenging to use. With increased aging, reduced dexterity and lower vision make it difficult to locate the necessary information on the screen or click on the appropriate button, which can impede their experience when planning travels in the community [9]. Specifically, visual decline also impacts the contrast sensitivity in using travel planning e-tools (eg, recognizing the elements in maps or e-tools). The decline of fine motor skills [4] may interfere with the use of electronic devices that allow movement planning. Moreover, lower executive abilities can lead to difficulty in planning travel and understanding public transportation options [10]. Besides, some older adults experience difficulty in travel planning due to their lack of digital literacy [11], which may raise challenges, such as access to the up-to-date flow of information for travel planning [12]. The interface of the developed tool should thus consider these age-related changes and the variability in numerical literacy [13] in planning trips. When it comes to designing an interface, user experience is essential [14].

User experience is defined as an "iterative set of decisions leading to a successful outcome with an interactive tool, as well as a productive and satisfying process while arriving at this outcome" [15]. According to Hassenzahl and Tractinsky [16], it is created by the interaction of the user's internal state, the characteristics of the designed system, and context. It should be noted that 2 other concepts should be introduced in relation to user experience: user interface and usability. User interface is "the part of technology that people interact with. The interaction between a computer and a user is a two-way interaction" [17]. Therefore, user interface design refers to "the iterative set of decisions leading to a successful implementation of an interactive tool" [15]. Usability refers to the fact that "when a product or service is truly usable, the user can do what she or he wants to do, the way she or he expects to be able to do it, without hindrance, hesitation or questions" [18]. Hassan and Galal-Edeen [19] argue that usability is a subset of user experience, and none alone can determine the degree of conformity of a product or service to the expectations and needs of its users. The usability criteria include usefulness, efficiency, effectiveness, satisfaction, learnability, accessibility, memorability, and error tolerance. Considering these definitions in the framework of this study, the user experience refers to the interaction of older adults with the aesthetic and functional aspects of the transportation planning e-tool to plan a trip that satisfactorily meets their needs and preferences.

A co-design design is a participatory approach that can be relevant to older adults to address their challenges in planning a trip and improve the user experience when using e-tools. Participatory design refers to "a process of investigating, understanding, reflecting upon, establishing, developing, and supporting mutual learning between multiple participants in collective reflection-in-action" [20]. Understanding the interactions and collaboration of older adults in co-design approaches such as participatory design [20] is an important aspect of user experience design [21] as older adults' needs and physical condition can affect their participation in the design process. Participatory design could help evaluate user experience. As there is a lack of research on user experience design for older adults [22], it is essential to support older adults' mobility through travel options that are tailored to their needs and preferences by identifying knowledge gaps regarding the user experience design of older adults when using a transportation planning e-tool.

Context and Objective

The aim of this scoping review is to identify knowledge gaps regarding the user experience design and its evaluation for older adults using transportation planning e-tools.

Methods

Study Design

This scoping review was based on Arksey and O'Malley's [23] framework with the following stages: (1) identifying the research question; (2) identifying the relevant sources; (3) selecting the studies; (4) charting the data; and (5) collecting, summarizing, and reporting results.

Identifying the Research Question

The question was formulated using a broad approach to ensure sufficient breadth during the literature search. The question is: What is known about the user experience design, co-design, and its evaluation for older adults using transportation-planning e-tools to move around in the community? This scoping review followed the PCC framework: (P) older adults, (C) user experience design, co-design, and usability evaluation, and (C) transportation-planning e-tools.

Identifying the Relevant Sources

This stage consisted of searching scientific literature published in English from January 2002 to October 2022, with an updated

XSL•FO
RenderX

search conducted on October 3, 2023. The search was performed across 7 databases: MEDLINE (PubMed, NCBI), AgeLine (EBSCOhost), CINAHL Plus with Full Text (EBSCOhost), SCOPUS (Elsevier), ProQuest Dissertations and Theses (ProQuest), IEEE Xplore (IEEE Xplore Digital Library), and TRID (Transportation Research Board).

Databases and keywords were selected and confirmed during a multi-step process. First, the keywords and databases were selected by two of the authors (SBG and VP). Then, an experienced librarian (FL) reviewed and refined them, ensuring that the selected terms and databases were relevant to the research question. As part of the search strategy confirmation, FL provided feedback aligned with the PRESS 2015 Guidelines [24], offering expert guidance on refining keywords, optimizing Boolean operators, and using controlled vocabulary where available. These adjustments helped ensure both the accuracy and comprehensiveness of the search. Some decisions were made to retrieve the largest number of relevant sources. During the pretesting of the research strategy, incorporating transportation-related keywords alongside other search terms significantly reduced the number of relevant studies. To address this issue, the librarian recommended removing these keywords and adding the TRID (Transportation Research International Documentation) and IEEE Xplore databases. This adjustment ensured a more comprehensive search and improved the scope of the review. Finally, the keywords and databases were confirmed by a senior researcher (HP), an experienced professor in the field of user experience. The full search strategies, including controlled vocabulary and keywords, are detailed in Table 1. To ensure comprehensive coverage, we first identified controlled vocabulary terms from each database's thesaurus (eg, MeSH in PubMed, CINAHL Headings, and IEEE Xplore's indexing terms). Based on these terms, we then created a structured list of keywords for each concept to capture both indexed and natural variations of relevant terms.

Table . Concepts and keywords.

Concept	Controlled Vocabulary (by Database)	Keywords
User experience	User-computer interface (MeSH—MEDLINE, PubMed) Usability testing (CINAHL headings) Human-computer interaction (IEEE Thesaurus)	"User Interface" OR "User interface design" OR "UI" OR "UI design" OR "User experience" OR "User experience design" OR "UX" OR "UX design" OR "Usability"
Aging	Aged (MeSH—MEDLINE, PubMed) Older adults (CINAHL headings) Older adults (TRB thesaurus)	"Older adult" OR "Old people" OR "Elderly" OR "Senior" OR "Aging" OR "Normal aging" OR "Frailty"
E-tool	Mobile applications (MeSH—MEDLINE, PubMed) Software applications (IEEE thesaurus) Web applications (CINAHL headings)	"Website" OR "Application" OR "E-tool" OR "Web application" OR "App" OR "Web"
Co-design	User-centered design (MeSH—MEDLINE, PubMed, CINAHL, SCOPUS) Participatory design (IEEE thesaurus)	"Co-design" OR "Participatory Design" OR "Co-operative Design" OR "User-Centered Design" OR "Co-creation"

The search results were exported into Zotero (Corporation for Digital Scholarship) for reference management. Duplicate records were initially removed using Zotero's automated deduplication feature, followed by a manual review by the first author (SBG) to identify and eliminate any remaining duplicates.

Selecting the Studies

The inclusion and exclusion criteria were established before the study commenced and applied during the screening process (SBG, VP, and HP). Sources were included if they were (1) focused on developing or evaluating e-tools, and (2) discussed e-tool usability or user experience. Sources were excluded if they did not apply to the target population (older adults). To evaluate a larger range of studies, articles whose target group was 50 years or older were considered. Some studies define 'older adults' as individuals aged 65 years and above, while others include individuals aged 50 years and older, especially in the context of research related to aging and technology use [25]. Considering individuals aged 50 years and older as "older adults" helped to retrieve a broader set of articles, particularly those focusing on promoting the use of e-tools for the aging population. This approach ensured that relevant studies addressing the needs of both younger and older segments of the aging population were included, capturing a more comprehensive view of user experience with transportation-planning e-tools. The first author (SBG) screened all identified sources by title and abstract, followed by a full-text review to assess their eligibility. To mitigate the risk of bias, 20% of the study selection process was independently reviewed and confirmed by co-authors (VP and HP).

Charting the Data

A data charting form was developed by SBG and validated by VP, using a Microsoft Excel software grid that included source characteristics (authors, year of publication, title of the article, source of article, country, and context of the studies), the purpose of the studies (objectives and study orientation), and mapping of the methodology and evaluation of user experience (design approach, setting, type of data collection and analysis, type of usability evaluation, and sample size).

Collecting, Summarizing, and Reporting Results

Both descriptive-analytical method and thematic analysis [26] were used to analyze all categories of data (characteristics of

XSL•FO
RenderX

sources, purpose of the studies, and the reflection of the methodology and evaluation). Results were discussed, analyzed, and combined to identify the knowledge gap. Table 2 shows the list of selected articles.

XSL•FO
RenderX

Table . List of the selected articles.

Title	Year	Aim	Context	Design approach	Study setting	Type of methodology
A New Tourist Audio Guide Service for Elderly People Integrated in the Mobile Phone: Preliminary Results [27]	2010	Evaluating user experience	Tourism	—[a]	Laboratory	Qualitative and quantitative
A User-Centered Interface Design for a Pill Dispenser [28]	2020	Evaluating user experience	Health	User-centered design	Laboratory	Qualitative and quantitative
A User-centred Design Approach for Mobile-Government Systems for the Elderly [29]	2018	Evaluating user experience	Service provider	User-centered design	Laboratory	Qualitative and quantitative
Adoption and feasibility of a communication app to enhance social connectedness amongst frail institutionalized oldest old: an embedded case study [30]	2018	Evaluating user experience	Communication	Participatory design	Real-world environment	Qualitative and quantitative
B2C Websites' Usability for Chinese Senior Citizens [31]	2014	Evaluating user experience	Shopping	User-centered design	Laboratory	Quantitative
Co-Creation with Older Adults to Improve User-Experience of a Smartphone Self-Test Application to Assess Balance Function [14]	2020	Developing or redesigning an e-tool	Health	Participatory design	Laboratory	Qualitative and quantitative
Creating a digital memory notebook application for individuals with mild cognitive impairment to support everyday functioning [25]	2019	Evaluating user experience	Daily planner	—	Laboratory	Qualitative and quantitative
Design and Development of an eHealth Service for Collaborative Self-Management among Older Adults with Chronic Diseases: A Theory-Driven User-Centered Approach [32]	2021	Developing or redesigning an e-tool	Health	ISO 9241 - 210 standard for human-centered design	Real-world environment	Qualitative

XSL•FO
RenderX

Title	Year	Aim	Context	Design approach	Study setting	Type of methodology
Design and Usability Evaluation of Mobile Voice-Added Food Reporting for Elderly People: Randomized Controlled Trial [33]	2020	Evaluating user experience	Nutrition	User-centered design	Laboratory	Qualitative and quantitative
Design and development of a mobile app to support the care of the elderly [34]	2021	Evaluating user experience and developing or redesigning an e-tool	Health	User-centered design	Laboratory	Quantitative
Development and Field Testing of a Long-Term Care Decision Aid Website for Older Adults: Engaging Patients and Caregivers in User-Centered Design [35]	2020	Evaluating user experience and developing or redesigning an e-tool	Health	User-centered design co-design	Laboratory	Qualitative and quantitative
Evaluation of an App to Support Healthy Living by Older Adults [36]	2017	Developing or redesigning an e-tool	Nutrition	User-centered design	Laboratory	Qualitative and quantitative
Home Monitoring System for Comprehensive Geriatric Assessment in Patient's Dwelling: System Design and UX Evaluation [37]	2021	Evaluating User Experience & Developing or redesigning an e-tool	Health	User-centered design, participatory design	Laboratory	Qualitative and quantitative
Human Factors Analysis, Design, and Evaluation of Engage, a Consumer Health IT Application for Geriatric Heart Failure Self-Care [38]	2017	Developing or redesigning an e-tool	Health	User-centered design	Laboratory	Qualitative and quantitative
Human-Centered Design Study: Enhancing the Usability of a Mobile Phone App in an Integrated Falls Risk Detection System for Use by Older Adult Users [39]	2017	Evaluating user experience and developing or redesigning an e-tool	Health	Human-centered design	Laboratory	Qualitative and quantitative
Improving Messenger Accessibility for Elderly Users using User Centered Design (UCD) Methods (Study Case: WhatsApp) [40]	2021	Evaluating user experience and developing or redesigning an e-tool	Communication	User-centered design	Laboratory	Qualitative and quantitative

XSL•FO
RenderX

Title	Year	Aim	Context	Design approach	Study setting	Type of methodology
Older Adults Can Successfully Monitor Symptoms Using an Inclusively Designed Mobile Application [41]	2020	Developing or redesigning an e-tool	Health	Inclusive design	Real-world environment	Qualitative and quantitative
Measure It Super Simple (MISS) activity tracker: (re)design of a user-friendly interface and evaluation of experiences in daily life [42]	2022	Evaluating user experience and developing or redesigning an e-tool	Health	User-centered design	Real-world environment	Qualitative and quantitative
Mobile App Prototype in Older Adults for Post fracture Acute Pain Management: User-Centered Design Approach [43]	2022	Evaluating user experience and developing or redesigning an e-tool	Health	Human-centered design	Laboratory	Qualitative
Participatory Design: Apps from The Older Adults to The Older Adults [44]	2021	Developing or redesigning an e-tool	Health	Participatory design	Laboratory	Qualitative
Promoting social and collaborative language learning among older adults in the digital era: Development and evaluation of a smartphone app prototype using a design-thinking approach [45]	2022	Evaluating user experience and eveloping or redesigning an e-tool	Learning	Specific type of user-centered design approach known as design thinking (DT)	Laboratory	Qualitative and quantitative
Requirements Elicitation to Develop Mobile Application for Elderly [46]	2017	Evaluating user experience	Not specified	User-centered design	Laboratory	Qualitative and quantitative
Sahayak: An Application for Social and Physical Well-Being for the Elderly [47]	2020	Evaluating user experience and developing or redesigning an e-tool	Social and physical well-being	User-centered design	Laboratory	Qualitative
Tailoring digital apps to support active ageing in a low income community [48]	2020	Developing or redesigning an e-tool	Physical activity	Co-design	Laboratory	Qualitative
The Developments and Iterations of a Mobile Technology-Based Fall Risk Health Application [49]	2022	Evaluating user experience and developing or redesigning an e-tool	Health	User-centered design	Real-world environment	Qualitative and quantitative

XSL•FO
RenderX

Title	Year	Aim	Context	Design approach	Study setting	Type of methodology
Untold Stories in User-Centered Design of Mobile Health: Practical Challenges and Strategies Learned From the Design and Evaluation of an App for Older Adults With Heart Failure [50]	2020	Evaluating user experience and developing or redesigning an e-tool	Health	User-centered design	Laboratory	Qualitative and quantitative
Usability and feasibility of consumer-facing technology to reduce unsafe medication use by older adults [51]	2020	Evaluating user experience	Health	User-centered design	Laboratory	Qualitative and quantitative
User Centered Design Approach for Elderly People in using Website [52]	2012	Developing or redesigning an e-tool	Transportation	User-centered design	Laboratory	Qualitative
User Interface for Social Networking Application for the Elderly [53]	2013	Evaluating user experience and developing or redesigning an e-tool	Social network	User-centered design	Real-world environment	Qualitative and quantitative
Usability testing of tablet-based cognitive behavioral intervention application to improve a simple walking activity for older adults with arthritis fatigue [54]	2021	Evaluating user experience and developing or redesigning an e-tool	Health	User-centered design	Real-world environment	Qualitative
The TV-WEB project—combining internet and television—lessons learnt from the user experience studies [55]	2017	Evaluating user experience	Entertainment	—	Laboratory	Qualitative and quantitative
Usability Study on Railway Self-Service Terminal Interface for the Elderly [56]	2018	Evaluating user experience	Transportation	—	Real-world environment	Qualitative and quantitative
Redesigning websites for older adults: a case study [57]	2014	Evaluating user experience and developing or redesigning an e-tool and documenting or extracting usability principles	Tourism	—	Laboratory	Qualitative and quantitative

XSL•FO
RenderX

Title	Year	Aim	Context	Design approach	Study setting	Type of methodology
Evaluating websites for older adults: adherence to 'senior-friendly' guidelines and end-user performance [58]	2008	Evaluating user experience	Not specified	—	Laboratory	Qualitative and quantitative
Evaluating 3D Printed VR Controller Prototypes to Increase VR Accessibility for Older Adults [59]	2022	Evaluating user experience	Not specified	—	Laboratory	Qualitative and quantitative
The Co-Design/Co-Development and Evaluation of an Online Frailty Check Application for Older Adults: Participatory Action Research with Older Adults [60]	2023	Evaluating user experience and developing or redesigning an e-tool	Health	—	Laboratory	Qualitative and quantitative
Applicability of the User Experience Methodology: Communication and Employment Web Portal for Older Adults [61]	2023	Evaluating user experience and developing or redesigning an e-tool	Communication and employment	—	Laboratory	Qualitative and quantitative
Integrated health system to assess and manage frailty in community dwelling: Co-design and usability evaluation [62]	2023	Evaluating user experience and developing or redesigning an e-tool	Health	Co-creation human-centered design	Real-world environment	Qualitative and quantitative
User-centered design (UCD) of time-critical weather alert application [63]	2023	Evaluating user experience and developing or redesigning an e-tool	Weather	User-centered design	Laboratory	Qualitative and quantitative
Web-Based Cognitive Behavioral Therapy for Depression Among Homebound Older Adults: Development and Usability Study [64]	2023	Evaluating user experience and developing or redesigning an e-tool	Health	Co-design	Laboratory	Qualitative and quantitative

[a]not available.

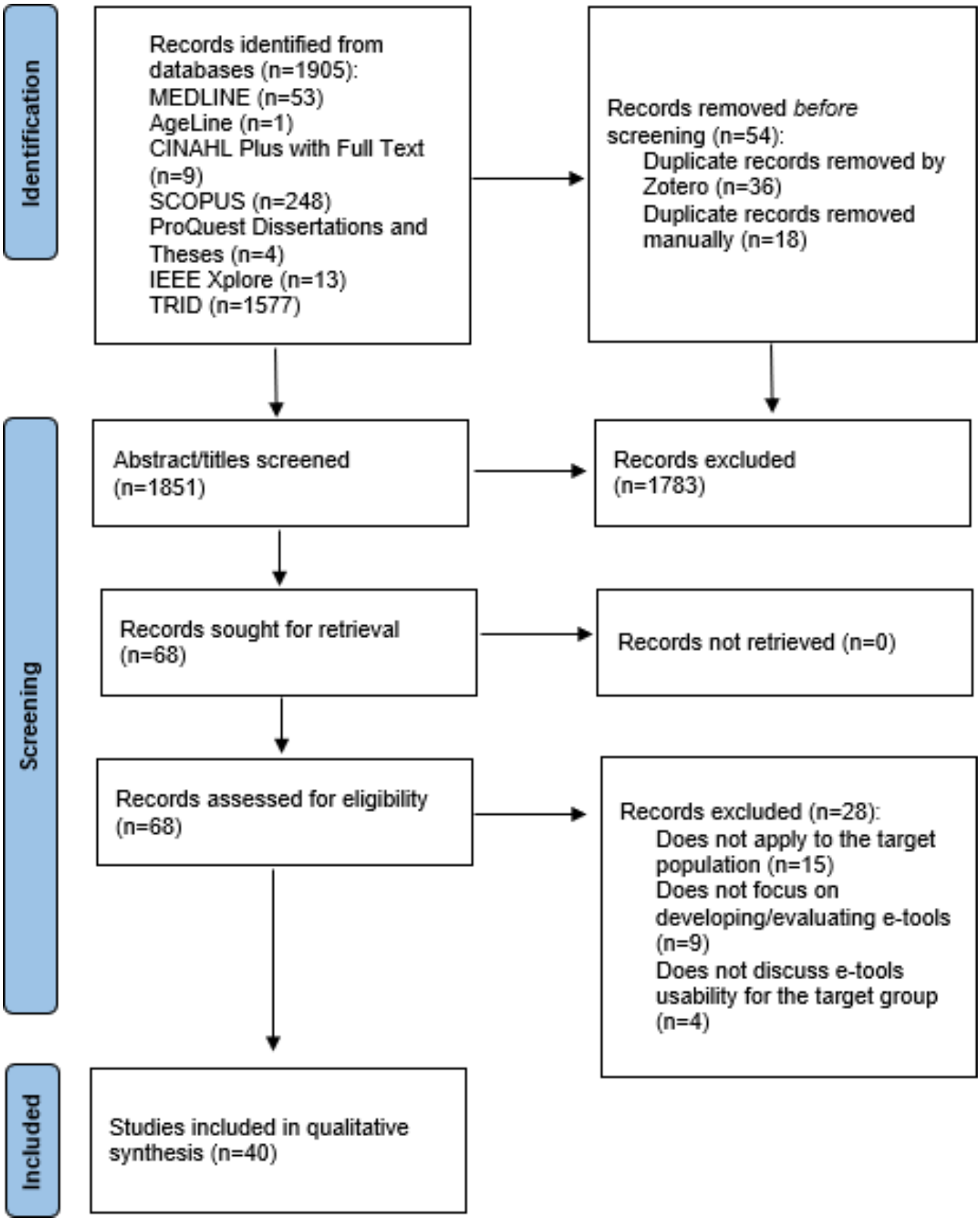
Results

Screening Results

Overall, 1905 studies were identified through the databases. The record after duplicate elimination was 1854. In total, 68 sources were retained after screening for title and abstract. Finally, since 28 records did not meet the mentioned inclusion and exclusion criteria, 40 sources were included in this review (Figure 1; Checklist 1).

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram detailing the process of literature search, screening, and selection.



Characteristics of Sources

Out of a total of 40 studies, 4 studies [27,52,57,58] were published during the first decade of the scope (2002 - 2012) (10%), while 36 studies [14,25,28-51,53-56,59-64] were published after 2012 (90%). These studies were conducted across different continents, with 14 (35%) in Europe [14,27-29,32,36,37,39,42,45,53,55,57,62], 10 (25%) in Asia [31,33,38,40,44,46,47,52,56,60], 12 (30%) in North America [25,30,35,43,49-51,54,58,59,63,64], and 4 (10%) in South America [34,41,48,61]. The majority of the sources (29 studies) were peer-reviewed articles, representing 72.5% of the total [14,25,30,32,33,35-45,48-51,54,55,57,58,60-64], while the

XSL•FO
RenderX

remaining 11 studies were conference articles (27.5%) [27-29,31,34,46,47,52,53,56,59].

Most contexts of the identified studies were related to health [14,28,32,34,35,37-39,41-44,49-51,54,60,62,64]. The other identified study contexts included tourism [27,57], shopping [31], daily planning [25], learning [45], social and physical well-being [47], physical activity [48], service provider [29], communication [30,40,61], nutrition [33,36], social network [53], entertainment [55], and weather [63]. In total, 3 articles did not specify the context of the study [46,58,59]. In total, 2 studies aimed to develop an e-tool and evaluate user experience in transportation [52,56].

The Purpose of the Selected Studies

The purpose of the selected articles was categorized into three groups: (1) documenting or extracting usability design principles specified to older adults, (2) evaluating user experience, and (3) developing or redesigning an e-tool for older adults. Out of the 40 articles, only 1 (2.5%) followed all 3 objectives [57]. In total, 12 (30%) articles [25,27,29-31,33,46,51,55,56,58,59] focused on evaluating user experience, while 19 (47.5%) articles [28,34,35,37,39,40,42,43,45,47,49,50,53,54,60-64] concentrated on both evaluating user experience and developing or redesigning an e-tool. The remaining 8 (20%) articles [14,32,36,38,41,44,48,52] only focused on developing or redesigning an e-tool. Evaluating user experience had the highest rate among all the studies, representing 80% of the total (32 studies) [25,27-31,33-35,37,39,40,42,43,45-47,49-51,53-64].

Empirical Methods of Usability Evaluation

Design Approach

Out of the total 40 articles, 22 (55%) of them were based on user-centered design [28,29,31-34,36,38-40,42,43,45-47,49-54,63], while 5 (12.5%) articles followed a co-design or participatory design approach [14,30,44,48,64]. In total, 3 (7.5%) articles used both user-centered design and co-design approaches [35,37,62], and 1 (2.5%) article followed an inclusive design approach [41]. There were 9 (22.5%) articles that did not specify any design approach [25,27,55-61].

Settings

Regarding experimental setting, 31 (77.5%) articles [14,25,27-29,31,33-40,43-48,50-52,55,57-61,63,64] were conducted in a laboratory setting, while the remaining 9 (22.5%) articles [30,32,41,42,49,53,54,56,62] took place in a real-world environment.

Type of Data Collection and Analysis

Most of the articles (31 studies) used both qualitative and quantitative methods for data collection and analysis [14,25,27-30,33,35-42,45,46,49-51,53,55-64], which represented 77.5% of studies. In total, 7 (17.5%) articles used the qualitative method [32,43,44,47,48,52,54], and just 2 (5%) articles [28,34] used only the quantitative method.

Usability Evaluation

Usability evaluation is usually divided into 2 general categories: (1) formative usability (usability inspection and expert-based): evaluating systems by experienced professionals using predetermined principles to identify usability issues [65], such as heuristic evaluation [66] and cognitive walkthrough [67]; and (2) summative usability (usability testing and user-based): evaluating the system by observing and recording the objective performance and subjective opinion of target users while interacting with a product to identify usability issues [68]. In usability tests, data are collected through diverse methods of objective performance and subjective opinions. A great number of studies (36/40) used summative usability evaluation (user-based), representing 90% of the studies [14,25,27-47,49-51,53,56-64]. Among the 40 articles, 10 articles (25%) used both summative and formative usability evaluation [29,39,44-46,50,57,58,61,62]. Only 1 (2.5%) article used formative usability evaluation [54], and 3 (7.5%) articles did not specify that they evaluated usability [48,52,55].

Data were collected through various methods during usability testing among older adults. Objective performance information was gathered from 19 articles, with task performance metrics being the most commonly used method [14,25,29,33,37,39,45,46,51,56-58,63], followed by observation 7 [14,30,40,42,43,47,64] and screen recording 3 [14,45,64] as less frequent methods. Subjective opinions (older adults' experiences and their design preferences) were predominantly collected via questionnaires [14,25,27,29-31,33-39,41,42,45,50,51,53,57-60,62-64], with additional methods including think-aloud protocol [14,35,37,38,40,44,45,50,51,54], interviews [30,32,35,36,38,40,42,47,49-51,53,56,57], focus groups [27,28,59,60], and open-ended questions [33,36,37,43,58,64]. A subset of studies also used formative usability evaluation, some including heuristic evaluations [29,39,44-46,50,57,58,61], and cognitive walkthrough a few articles [54,62].

The questionnaire was the most frequently used data collection method. Among them, the System Usability Scale (SUS), measuring efficacy and satisfaction with usability when users perform specific tasks [69], was the most prevalent one. Table 3 demonstrates a structured summary of the usability evaluation methods and their frequency.

XSL•FO
RenderX

Table . Usability testing data collection methods.

Usability evaluation type	Data collection method	n (%)
Objective performance	Task performance metrics	13 (32.5)
	Observation	7 (17.5)
	Screen recording	3 (7.5)
Subjective opinions	Questionnaire • System Usability Scale (SUS) (15 articles) • After-Scenario Questionnaire (ASQ) (2 articles) • Technology Acceptance Model (TAM) (3 articles) • Post-Study System Usability Questionnaire (PSSUQ) (1 article) • User Interface Satisfaction (QUIS) (2 articles) • NASA Task Load Index (NASA-TLX) (2 articles) • Health Information Technology Usability Evaluation Scale (Health-ITUES) (1 article) • Self-Assessment Manikin (SAM) (1 article) • Self-designed questionnaires (7 articles) • Usability Metric for User Experience (UMUX) (1 article) • Single Ease Question (SEQ) (1 article) • User Experience Questionnaire (UEQ) (1 article)	26 (65)
	Think-aloud	10 (25)
	Interview Semi-structured (5 articles)	14 (35)
	Focus group	4 (10)
	Open-ended questions	6 (15)

Sample Size

The number of participants varied, based on the purpose of the evaluation (ranging from 3 to 168). In the studies with a participatory or co-design approach (8 studies), the average sample size was 10, ranging from 3 to 22.

Synthesis of Results

Most of the studies focused on developing eHealth, and only 2 articles [52,56] were related to transportation. The first one [52] was about the usability evaluation of an existing travel planning website via a user-centered design approach and did not concentrate on developing the website, while the other one [56] concentrated on evaluating the user experience of existing rail terminal interfaces for older people without using any design approach. The first one was conducted in a laboratory setting with 3 older adults via various qualitative methods, while the second one was conducted in a real-life setting with 5 older adults and 5 adults via a combination of both qualitative and quantitative methods. Although 1 of these 2 articles was performed in a real-life setting, this study aimed to evaluate older adults' experience using existing services and did not concentrate on developing a transportation planning age-friendly e-tool.

Discussion

Principal Findings and Comparison With Previous Works

This scoping review aimed to identify knowledge gaps regarding the user experience design of older adults in using transportation planning e-tools. Our study reveals that e-tools designed for older adults have grown dramatically in recent years in health-related fields, while only 2 studies were related to transportation. Despite the large number of e-tools that have been developed to support transportation planning [70], documenting the user experience is rarely considered. This lack of focus may explain why it remains challenging for many older adults to use these e-tools effectively while planning their travels and finding their way from one place to another. For example, real-time travel information, as a crucial aspect of travel planning [71], is expected to significantly impact the user experience of older adults, by reducing the anxiety of getting lost [72], protect them from the weather traffic congestion [73,74] and enhance the use of modes of transportation that are tailored to their needs and preferences [75]. Statistics Canada [76] reported that only 54% of older adults have a smartphone for personal use, many of whom rarely use its various functions, which creates a gap between the intended goal of transportation planning e-tools and their current use. The lack of studies aiming

XSL•FO
RenderX

to document older adults' experience in using transportation planning e-tools might thus affect their accessibility in terms of content and functionality. In this context, Subasi et al [77] argue that developing a system that is not only universally accessible but that also meets the specific expectations of older adults will optimize their usage of the digital system.

This review also revealed the lack of studies concentrating on developing transportation planning e-tools using participatory design. This result might be explained by the challenges related to this design approach, such as substantial time, resources, and institutional commitment [78]. In participatory design, designers try to learn the realities of users' situations while the users strive to articulate their desired aims and learn the appropriate technological means to acquire them [20]. Although it seems that participatory design is an approach to the user-centered design [79], it is actually a shift in attitude from designing "for" users (user-centered design) to designing "by" users (participatory design) [80]. To address the challenges of participatory design, it would be good to understand the optimal methods and timing for involving users in the study process. Providing concrete deliverables and regular updates [81] could motivate participants by demonstrating the tangible results of their contributions to the development of transportation planning e-tools. Additionally, establishing a welcoming and enjoyable atmosphere at the beginning of co-designed sessions, such as icebreaking [9], can help participants feel comfortable and build trust with the researchers. Simplifying some consultation sessions into questionnaires or virtual sessions could lessen the burden of in-person meetings for both older adults and the research team, particularly during winter. This strategy would also facilitate frequent and balanced engagement with older adults while reducing expenses [82].

Although the aim of a considerable number of studies was to evaluate the experience of older adults using existing or newly developed e-tools, only 1 [57] considered not only evaluating user experience and redesigning an e-tool but also documenting usability design principles. It should be noted that developing usability design principles and guidelines for age-friendly e-tools is crucial due to variations in older adults' design characteristics [83]. Some user-centered methods, such as persona [84], might help to consider the needs and preferences of a wide range of older adults in developing age-friendly e-tools. More specifically, persona is used to identify user expectations through the definition of personality characteristics and specific personal details [84]. According to Cooper [85], the aim of persona is to have a comprehensive and global view of the target population to know for whom the product is not intended, and its approach is concentrating on "specific or canonical users" [84]. As none of the reviewed studies mentioned the use of persona, it would be interesting for future studies to explore how this method could help classify older adults with, for example, varying levels of digital literacy, diverse mobility profiles, and different travel planning habits. This could help identify variations in older adults' design characteristics and aid in the development of usability design principles.

Based on our results, combining formative (expert-based) and summative (user-based) usability evaluation using a mixed-method and providing the intersectoral participation of different disciplines might help to perform a comprehensive evaluation of the transportation planning e-tools for older adults. Indeed, most of the sources reviewed (77.5%) used both qualitative and quantitative methods for data collection and analysis, which helps to gain a detailed understanding of the needs and preferences of older adults. As argued by Creswell and Plano Clark [86], using a mixed method could provide a deeper understanding of older adults' expectations in travel planning via transportation planning e-tools that would help to further improve knowledge in the production of guidelines and—accordingly—develop useful e-tools. Furthermore, most of the articles conducted in standardized situations used summative usability evaluation (user-based testing), highlighting the urgent need to capture the experience of older adults, more specifically in real-life settings, to help design products as close to actual usage as possible [87]. However, conducting a summative usability evaluation in a real-life context is challenging, due to its resource-intensive and time-consuming nature, which makes it costly. To address these challenges effectively, it would be good to adopt a participatory design approach involving a minimum of 5 participants. These participants should encompass diverse mobility profiles, digital knowledge, and travel planning habits, representing the targeted end-users to address 80% of usability problems, as proposed by Nielsen [65].

Limitations

This study identified knowledge gaps regarding the design user experience of older adults when using transportation planning e-tools. However, it has certain limitations. Although the data selection and extraction processes were confirmed by co-authors, study screening was performed by the first author. It is possible that some articles were not retrieved due to the selected keywords and lack of consideration for gray literature. Moreover, we did not use the methodology suggested by Levac et al [88], that includes a consultation step. This step would be practical in transferring development knowledge for e-tools for a broad spectrum of older adults through different disciplines such as gerontology, web developing, and design. Finally, although we did not systematically account for all plural forms, alternative spellings, or flexible word order, the use of controlled vocabulary (eg, MeSH terms) and keyword variations helped maximize the retrieval of relevant literature, while the librarian's review ensured the rigor of the search strategy and adherence to best practices in systematic reviews.

Conclusions

Supporting the mobility of older adults should be a priority in developing a transportation-related e-tool. According to this review, there are no studies aimed at specifically improving the mobility experience of older adults using an age-friendly transportation planning e-tool in a real-life context. Therefore, it is necessary to fill this gap by using a participatory approach (co-design) to better understand the needs of older adults in a real context to create age-friendly design principles and guidelines in the development of a transportation planning e-tool. This will not only improve their usability experience when planning a trip using a transportation planning e-tool but also strengthen their social participation.

XSL•FO
RenderX

Acknowledgments

We thank the Fonds de recherche du Québec–Santé for its financial support. We also thank Francis Lacasse, librarian, for his invaluable assistance in reviewing and refining our search strategy, ensuring its accuracy and comprehensiveness.

Conflicts of Interest

None declared.

Checklist 1
PRISMA-ScR Checklist.
[PDF File, 112 KB - humanfactors_v12i1e63273_app1.pdf]

References

1. Ageing and health. WHO. 2018 Feb 5. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2021-01-15]
2. Morris SP, Fawcett G, Brisebois L, Hughes J. A demographic, employment and income profile of canadians with disabilities aged 15 years and over, 2017. Statistics Canada (Statistque Canada). 2018. URL: https://www150.statcan.gc.ca/n1/pub/89-654-x/89-654-x2018002-eng.htm [accessed 2021-11-30]
3. Andersen GJ. Aging and vision: changes in function and performance from optics to perception. Wiley Interdiscip Rev Cogn Sci 2012 May;3(3):403-410. [doi: 10.1002/wcs.1167] [Medline: 22919436]
4. Robnett RH, Chop WC. Gerontology for the Health Care Professional: Third Jones & Bartlett Publishers; 2013.
5. Whiteside MM, Wallhagen MI, Pettengill E. Sensory impairment in older adults: part 2: vision loss. Am J Nurs 2006 Nov;106(11):52-61. [doi: 10.1097/00000446-200611000-00019] [Medline: 17068436]
6. Weinstein BE, Ventry IM. Hearing impairment and social isolation in the elderly. J Speech Lang Hear Res 1982 Dec;25(4):593-599. [doi: 10.1044/jshr.2504.593]
7. Young KL, Koppel S, Charlton JL. Toward best practice in human machine interface design for older drivers: a review of current design guidelines. Accid Anal Prev 2017 Sep;106:460-467. [doi: 10.1016/j.aap.2016.06.010] [Medline: 27374247]
8. Levasseur M, Richard L, Gauvin L, Raymond E. Inventory and analysis of definitions of social participation found in the aging literature: proposed taxonomy of social activities. Soc Sci Med 2010 Dec;71(12):2141-2149. [doi: 10.1016/j.socscimed.2010.09.041] [Medline: 21044812]
9. Dixon J, Crooks H, Henry K. Breaking the ice: supporting collaboration and the development of community online. CJLT / RCAT 2006;32(2). [doi: 10.21432/T25S3M]
10. Murman DL. The impact of age on cognition. Semin Hear 2015 Aug;36(3):111-121. [doi: 10.1055/s-0035-1555115] [Medline: 27516712]
11. Van De Watering M. The impact of computer technology on the elderly. 2005 Jun. (2008) p. 12 URL: https://www.marekvandewatering.com/texts/HCI_Essay_Marek_van_de_Watering.pdf [accessed 2025-06-26]
12. Zhou J, Rau PLP, Gao Q, Qin J, Liao Q. Requirements of transport information service and route guidance service for older adults. 2009 Presented at: 17th World Congress on Ergonomics (IEA 2009 URL: https://www.researchgate.net/publication/230710849_Requirements_of_transport_information_service_and_route_guidance_service_for_older_adults [accessed 2025-07-07]
13. Dodd C, Athauda R, Adam M. Designing user interfaces for the elderly: a systematic literature review. 2017 Presented at: ACIS 2017 Proceedings URL: https://aisel.aisnet.org/acis2017/61/ [accessed 2025-07-07]
14. Mansson L, Wiklund M, Öhberg F, Danielsson K, Sandlund M. Co-creation with older adults to improve user-experience of a smartphone self-test application to assess balance function. Int J Environ Res Public Health 2020 May 26;17(11):3768. [doi: 10.3390/ijerph17113768] [Medline: 32466484]
15. Roth R. User interface and user experience (UI/UX) design. GIS&T BoK 2017;2017(Q2) [FREE Full text] [doi: 10.22224/gistbok/2017.2.5]
16. Hassenzahl M, Tractinsky N. User experience - a research agenda. Behav Inf Technol 2006 Mar;25(2):91-97. [doi: 10.1080/01449290500330331]
17. Weinschenk S, Barker DT. Designing Effective Speech Interfaces: John Wiley & Sons, Inc; 2000.
18. Rubin J, Chisnell D. Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests: Second. John Wiley & Sons; 2008.
19. Hassan HM, Galal-Edeen GH. From usability to user experience. Presented at: 2017 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS); Okinawa p. 216-222. [doi: 10.1109/ICIIBMS.2017.8279761]
20. Robertson T, Simonsen J. Participatory design: an introduction. In: Routledge International Handbook of Participatory Design, 1st edition: Routledge, Vol. 2012:1-17. [doi: 10.4324/9780203108543]

XSL•FO
RenderX

21. Sakaguchi-Tang DK, Cunningham JL, Roldan W, Yip J, Kientz JA. Co-design with older adults: examining and reflecting on collaboration with aging communities. Proc ACM Hum-Comput Interact 2021 Oct 13;5(CSCW2):1-28. [doi: 10.1145/3479506]
22. Yang M, Huang H. Research on interaction design of intelligent mobile phone for the elderly based on the user experience. : Springer Presented at: Human Aspects of IT for the Aged Population Design for Aging: First International Conference, ITAP 2015, Held as Part of HCI International 2015; Aug 2-7, 2015. [doi: 10.1007/978-3-319-20892-3_51]
23. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
24. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS peer review of electronic search strategies: 2015 guideline statement. J Clin Epidemiol 2016 Jul;75:40-46. [doi: 10.1016/j.jclinepi.2016.01.021] [Medline: 27005575]
25. Raghunath N, Dahmen J, Brown K, Cook D, Schmitter-Edgecombe M. Creating a digital memory notebook application for individuals with mild cognitive impairment to support everyday functioning. Disabil Rehabil Assist Technol 2020 May;15(4):421-431. [doi: 10.1080/17483107.2019.1587017] [Medline: 30907223]
26. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
27. Díaz U, García Á, de Felipe A. A new tourist audio guide service for elderly people integrated in the mobile phone. 2010 Jun 23 Presented at: PETRA '10; Samos Greece URL: https://dl.acm.org/doi/proceedings/10.1145/1839294 [doi: 10.1145/1839294.1839315]
28. Sahlab N, Wiebelitz L, Schmid P, et al. A user-centered interface design for a pill dispenser. Presented at: 2020 IEEE 9th Global Conference on Consumer Electronics (GCCE); Kobe, Japan p. 761-764. [doi: 10.1109/GCCE50665.2020.9291946]
29. Ko A, Molnar T, Matyus B. A user-centred design approach for mobile- government systems for the elderly. Presented at: 2018 12th International Conference on Software, Knowledge, Information Management & Applications (SKIMA); Phnom Penh, Cambodia. [doi: 10.1109/SKIMA.2018.8631531]
30. Neves BB, Franz RL, Munteanu C, Baecker R. Adoption and feasibility of a communication app to enhance social connectedness amongst frail institutionalized oldest old: an embedded case study. Information, Communication & Society 2018 Nov 2;21(11):1681-1699. [doi: 10.1080/1369118X.2017.1348534]
31. Kang L, Dong H. B2C websites' usability for chinese senior citizens. : Springer; 2014 Presented at: Human-Computer Interaction Applications and Services: 16th International Conference, HCI International 2014; Jun 22-27, 2014. [doi: 10.1007/978-3-319-07227-2_2]
32. Ekstedt M, Kirsebom M, Lindqvist G, et al. Design and development of an eHealth service for collaborative self-management among older adults with chronic diseases: a theory-driven user-centered approach. Int J Environ Res Public Health 2021 Dec 30;19(1):391. [doi: 10.3390/ijerph19010391] [Medline: 35010652]
33. Liu YC, Chen CH, Lin YS, et al. Design and usability evaluation of mobile voice-added food reporting for elderly people: randomized controlled trial. JMIR Mhealth Uhealth 2020 Sep 28;8(9):e20317. [doi: 10.2196/20317] [Medline: 32985999]
34. Rodriguez-Duenas WR, Aguia-Rojas K, Valencia-Daza V. Design and development of a mobile app to support the care of the elderly. Presented at: 2021 IEEE 2nd International Congress of Biomedical Engineering and Bioengineering (CI-IB&BI); Bogota DC, Colombia p. 1-4. [doi: 10.1109/CI-IBBI54220.2021.9626120]
35. Hoffman AS, Bateman DR, Ganoe C, et al. Development and field testing of a long-term care decision aid website for older adults: engaging patients and caregivers in user-centered design. Gerontologist 2020 Jul 15;60(5):935-946. [doi: 10.1093/geront/gnz141] [Medline: 31773140]
36. Haslinda Z, Sani A, Petrie H. Evaluation of an app to support healthy living by older adults. 2017 Presented at: Electronic Visualisation and the Arts (EVA 2017. [doi: 10.14236/ewic/HCI2017.42]
37. Villalba-Mora E, Ferre X, Pérez-Rodríguez R, et al. Home monitoring system for comprehensive geriatric assessment in patient's dwelling: system design and UX evaluation. Front Digit Health 2021;3:659940. [doi: 10.3389/fdgth.2021.659940] [Medline: 34713133]
38. Srinivas P, Cornet V, Holden R. Human factors analysis, design, and evaluation of Engage, a consumer health IT application for geriatric heart failure self-care. Int J Hum Comput Interact 2017;33(4):298-312. [doi: 10.1080/10447318.2016.1265784] [Medline: 30429638]
39. Harte R, Quinlan LR, Glynn L, et al. Human-centered design study: enhancing the usability of a mobile phone app in an integrated falls risk detection system for use by older adult users. JMIR Mhealth Uhealth 2017 May 30;5(5):e71. [doi: 10.2196/mhealth.7046] [Medline: 28559227]
40. Zahida W, Effendy V, Hadikusuma A. Improving messenger accessibility for elderly users using user centered design (UCD) methods (study case: whatsapp). Presented at: 2021 International Conference on Software Engineering & Computer Systems and 4th International Conference on Computational Science and Information Management (ICSECS-ICOCSIM); Pekan, Malaysia p. 273-278. [doi: 10.1109/ICSECS52883.2021.00056]
41. Reading Turchioe M, Grossman LV, Baik D, et al. Older adults can successfully monitor symptoms using an inclusively designed mobile application. J Am Geriatr Soc 2020 Jun;68(6):1313-1318. [doi: 10.1111/jgs.16403] [Medline: 32157679]

XSL•FO
RenderX

42. Ummels D, Braun S, Stevens A, Beekman E, Beurskens A. Measure it super simple (MISS) activity tracker: (re)design of a user-friendly interface and evaluation of experiences in daily life. Disabil Rehabil Assist Technol 2022 Oct;17(7):767-777. [doi: 10.1080/17483107.2020.1815089] [Medline: 32970493]
43. Tran-Nguyen K, Berger C, Bennett R, Wall M, Morin SN, Rajabiyazdi F. Mobile app prototype in older adults for postfracture acute pain management: user-centered design approach. JMIR Aging 2022 Oct 17;5(4):e37772. [doi: 10.2196/37772] [Medline: 36251348]
44. Abdullah Sani ZH, Mohd Nizam D, Baharum A. Participatory design: apps from the older adults to the older adults. JST 2021;29(4) [FREE Full text] [doi: 10.47836/pjst.29.4.35]
45. Puebla C, Fievet T, García J, Tsopanidi M, Clahsen H. Promoting social and collaborative language learning among older adults in the digital era: development and evaluation of a smartphone app prototype using a design-thinking approach. Gerontechnology 2022 Jan 1;21(1):1-12 [FREE Full text] [doi: 10.4017/gt.2022.21.1.483.06]
46. Jakkaew P, Hongthong T. Requirements elicitation to develop mobile application for elderly. Presented at: 2017 International Conference on Digital Arts, Media and Technology (ICDAMT); Chiang Mai, Thailand p. 464-467. [doi: 10.1109/ICDAMT.2017.7905013]
47. Bhayana R, Agrawal K, Aggarwal M, Devgon R, Kar R. Sahayak: an application for social and physical well-being for the elderly. 2020 Nov 5 Presented at: IndiaHCI 2020 p. 124-129. [doi: 10.1145/3429290.3429300]
48. Castro PC, Romano LB, Frohlich D, et al. Tailoring digital apps to support active ageing in a low income community. PLoS One 2020;15(12):e0242192. [doi: 10.1371/journal.pone.0242192] [Medline: 33301455]
49. Hsieh KL, Frechette ML, Fanning J, Chen L, Griffin A, Sosnoff JJ. The developments and iterations of a mobile technology-based fall risk health application. Front Digit Health 2022;4:828686. [doi: 10.3389/fdgth.2022.828686] [Medline: 35574255]
50. Cornet VP, Toscos T, Bolchini D, et al. Untold stories in user-centered design of mobile health: practical challenges and strategies learned from the design and evaluation of an app for older adults with heart failure. JMIR Mhealth Uhealth 2020 Jul 21;8(7):e17703. [doi: 10.2196/17703] [Medline: 32706745]
51. Holden RJ, Campbell NL, Abebe E, et al. Usability and feasibility of consumer-facing technology to reduce unsafe medication use by older adults. Res Social Adm Pharm 2020 Jan;16(1):54-61. [doi: 10.1016/j.sapharm.2019.02.011] [Medline: 30853507]
52. Balakrishnan S, Salim SSB, Hong JL. User centered design approach for elderly people in using website. 2012 Presented at: 2012 international conference on advanced computer science applications and technologies (ACSAT); Kuala Lumpur, Malaysia p. 382-387. [doi: 10.1109/ACSAT.2012.11]
53. Kivimäki T, Kölndorfer P, Vainio AM, et al. User interface for social networking application for the elderly. 2013 May 29 Presented at: PETRA '13; Rhodes Greece p. 1-8 URL: https://dl.acm.org/doi/proceedings/10.1145/2504335 [doi: 10.1145/2504335.2504358]
54. Choi J, Cody JL, Fiske S. Usability testing of tablet-based cognitive behavioral intervention application to improve a simple walking activity for older adults with arthritis fatigue. Geriatr Nurs (Lond) 2021 Mar;42(2):473-478. [doi: 10.1016/j.gerinurse.2021.02.014]
55. Guna J, Stojmenova E, Kos A, Pogačnik M. The TV-WEB project - combining internet and television – lessons learnt from the user experience studies. Multimed Tools Appl 2017 Oct;76(20):20377-20408. [doi: 10.1007/s11042-016-3243-3]
56. Luo Z, Wang Y, Wang P. Usability study on railway self-service terminal interface for the elderly. Presented at: 2018 11th International Symposium on Computational Intelligence and Design (ISCID); Hangzhou, China p. 157-160. [doi: 10.1109/ISCID.2018.00043]
57. Patsoule E, Koutsabasis P. Redesigning websites for older adults: a case study. Behav Inf Technol 2014 Jun 3;33(6):561-573. [doi: 10.1080/0144929X.2013.810777]
58. Hart TA, Chaparro BS, Halcomb CG. Evaluating websites for older adults: adherence to 'senior-friendly' guidelines and end-user performance. Behav Inf Technol 2008 May;27(3):191-199. [doi: 10.1080/01449290600802031]
59. Gluck A, Solini H, Maiti K, Brinkley J. Evaluating 3D printed VR controller prototypes to increase VR accessibility for older adults. Presented at: 2022 IEEE 3rd International Conference on Human-Machine Systems (ICHMS); Orlando, FL, USA p. 1-6. [doi: 10.1109/ICHMS56717.2022.9980720]
60. Son BK, Miura T, Yabu KI, et al. The co-design/co-development and evaluation of an online frailty check application for older adults: participatory action research with older adults. Int J Environ Res Public Health 2023 Jun 10;20(12):6101. [doi: 10.3390/ijerph20126101] [Medline: 37372688]
61. Ferrer-Mavárez MDLÁ, Aguirre-Villalobos ER, Valecillos-Pereira JB. Applicability of the User Experience Methodology: Communication and Employment Web Portal for Older Adults. MaC 2023;11(3):63-76. [doi: 10.17645/mac.v11i3.6775]
62. Moral C, Pérez-Rodríguez R, Villalba-Mora E, Barrio-Cortes J, Ferre X, Rodríguez-Mañas L. Integrated health system to assess and manage frailty in community dwelling: Co-design and usability evaluation. Digit Health 2023;9:20552076231181229. [doi: 10.1177/20552076231181229] [Medline: 37361432]
63. Ali AM, Khamaj A, Kang Z, Moosa M, Alam MM. User-centered design (UCD) of time-critical weather alert application. IJACSA 2023;14(1):797-808. [doi: 10.14569/IJACSA.2023.0140188]
64. Xiang X, Kayser J, Ash S, et al. Web-based cognitive behavioral therapy for depression among homebound older adults: development and usability study. JMIR Aging 2023 Sep 19;6:e47691. [doi: 10.2196/47691] [Medline: 37725423]

XSL•FO
RenderX

65. Nielsen J. Usability Engineering: Morgan Kaufmann; 1994. [doi: 10.1016/B978-0-08-052029-2.50009-7]
66. Nielsen J, Molich R. Heuristic evaluation of user interfaces. 1990 Presented at: the SIGCHI conference; Seattle, Washington, United States p. 249-256 URL: http://portal.acm.org/citation.cfm?doid=97243 [doi: 10.1145/97243.97281]
67. Blackmon MH, Polson PG, Kitajima M, Lewis C. Cognitive walkthrough for the web. 2002 Apr 20 Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems p. 463-470 URL: https://dl.acm.org/doi/proceedings/10.1145/503376 [doi: 10.1145/503376.503459]
68. Wang Q, Liu J, Zhou L, et al. Usability evaluation of mHealth apps for elderly individuals: a scoping review. BMC Med Inform Decis Mak 2022;22(1):317. [doi: 10.1186/s12911-022-02064-5]
69. Johnson R, Shaw J, Berding J, Gather M, Rebstock M. European national government approaches to older people's transport system needs. Transp Policy (Oxf) 2017 Oct;59:17-27. [doi: 10.1016/j.tranpol.2017.06.005]
70. Tahir S, Abdulrazak B, Baillargeon D, Girard C, Provencher V. Adapting transportation planning e-tools to older adults' needs: scoping review. JMIR Aging 2023 May 16;6:e41938. [doi: 10.2196/41938] [Medline: 37191985]
71. Hounsell NB, Shrestha BP, McDonald M, Wong A. Open data and the needs of older people for public transport information. Transportation Research Procedia 2016;14:4334-4343. [doi: 10.1016/j.trpro.2016.05.355]
72. Fiedler M. Older people and public transport: challenges and changes of an ageing society. European Metropolitan Transport Authorities. 2007. URL: https://www.emta.com/wp-content/uploads/2023/08/Final_Report_Older_People_protec.pdf [accessed 2023-06-02]
73. Vechione M, Marrufo C, Vargas-Acosta RA, et al. Smart mobility for seniors: challenges and solutions in el paso, TX, and new york, NY. Presented at: 2018 IEEE International Smart Cities Conference (ISC2); Kansas City, MO, USA p. 1-8. [doi: 10.1109/ISC2.2018.8656975]
74. Vargas-Acosta RA, Becerra DL, Gurbuz O, Villanueva-Rosales N, Nunez-Mchiri GG, Cheu RL. Smart mobility for seniors through the urban connector. Presented at: 2019 IEEE International Smart Cities Conference (ISC2); Casablanca, Morocco p. 241-246. [doi: 10.1109/ISC246665.2019.9071732]
75. Stein M, Meurer J, Boden A, Wulf V. Mobility in later life: appropriation of an integrated transportation platform. : Association for Computing Machinery; 2017 May 2 Presented at: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems p. 5716-5729. [doi: 10.1145/3025453.3025672]
76. Statistics Canada. So long landline, hello smartphone. Statistics Canada. 2023 May 11. URL: https://www.statcan.gc.ca/o1/en/plus/3582-so-long-landline-hello-smartphone [accessed 2024-05-24]
77. Subasi Ö, Leitner M, Hoeller N, Geven A, Tscheligi M. Designing accessible experiences for older users: user requirement analysis for a railway ticketing portal. Univ Access Inf Soc 2011 Nov;10(4):391-402. [doi: 10.1007/s10209-011-0223-2]
78. Spinuzzi C. The methodology of participatory design. Tech Commun 2005;52(2):163-174 [FREE Full text]
79. Carroll JM. Encountering others: reciprocal openings in participatory design and user-centered design. Human–Computer Interaction 1996 Sep;11(3):285-290. [doi: 10.1207/s15327051hci1103_5]
80. Sanders EBN. From user-centered to participatory design approaches. In: Design and the Social Sciences, 1st edition: CRC Press; 2002:18-25. [doi: 10.1201/9780203301302-8]
81. Mind. Mind's participation and co-design practice framework. : Mind Australia; 2021 URL: https://www.mindaustralia.org.au/sites/default/files/2023-06/Participation_and_codesign_practice_framework.pdf [accessed 2024-05-30]
82. Kennedy A, Cosgrave C, Macdonald J, Gunn K, Dietrich T, Brumby S. Translating co-design from face-to-face to online: an Australian primary producer project conducted during COVID-19. Int J Environ Res Public Health 2021 Apr 14;18(8):4147. [doi: 10.3390/ijerph18084147] [Medline: 33919920]
83. Petrovčič A, Taipale S, Rogelj A, Dolničar V. Design of mobile phones for older adults: an empirical analysis of design guidelines and checklists for feature phones and smartphones. Int J Hum–Comput Interact 2018 Mar 4;34(3):251-264. [doi: 10.1080/10447318.2017.1345142]
84. Brangier E, Bornet C. Persona: a method to produce representations focused on consumers' needs. In: Human Factors and Ergonomics in Consumer Product Design, 1st edition: CRC Press; 2011:37-61. [doi: 10.1201/b10950-5]
85. Cooper A. Why High-Tech Products Drive Us Crazy and How to Restore the Sanity, 2nd edition: Pearson Higher Education; 2004. URL: https://z-lib.id/book/31979274-the-inmates-are-running-the-asylum-2 [accessed 2025-07-07]
86. Creswell JW, Clark VLP. Designing and Conducting Mixed Methods Research, 3rd edition: Sage publications; 2017.
87. Coorevits L, Georges A, Schuurman D. A framework for field testing in living lab innovation projects. TIM Review 2018 Dec 20;8(12):40-50 [FREE Full text] [doi: 10.22215/timreview/1204]
88. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5:69. [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]

Abbreviations

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SUS: System Usability Scale

XSL•FO
RenderX

Edited by A Kushniruk; submitted 17.06.24; peer-reviewed by B Skidmore, D Patil; revised version received 04.04.25; accepted 17.04.25; published 28.07.25.

Please cite as:
Bahrampoor Givi S, Gagnon-Roy M, Pigot H, Provencher V
User Experience of Older Adults With an Age-Friendly Transportation Planning E-Tool: Scoping Review
JMIR Hum Factors 2025;12:e63273
URL: https://humanfactors.jmir.org/2025/1/e63273
doi:10.2196/63273

© Sara Bahrampoor Givi, Mireille Gagnon-Roy, Hélène Pigot, Véronique Provencher. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 28.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effectiveness of Participatory Ergonomic Interventions on Work-Related Musculoskeletal Disorders, Sick Absenteeism, and Work Performance Among Nurses: Systematic Review

Guganesan Krishnanmoorthy[1], MPH; Sanjay Rampal[2], PhD; Sumitra Ropini Karuthan[1], MPH; Faiz Baharudin[3,4], DrPH; Rama Krishna[1,5], DrPH

[1]Department of Social and Preventive Medicine, Faculty of Medicine, Universiti Malaya, Wilayah Persekutuan, Kuala Lumpur, Malaysia

[2]Centre for Epidemiology and Evidence-based Practice, Department of Social and Preventive Medicine, Faculty of Medicine, Universiti Malaya, Kuala Lumpur, Malaysia

[3]Occupational Safety, Health and Environment Centre, Universiti Malaya, Kuala Lumpur, Malaysia

[4]Occupational Safety, Health and Environment Unit, Universiti Malaya Medical Centre, Kuala Lumpur, Malaysia

[5]Occupational Health Clinic, Universiti Malaya Medical Centre, Kuala Lumpur, Malaysia

Corresponding Author:
Rama Krishna, DrPH
Department of Social and Preventive Medicine, Faculty of Medicine, Universiti Malaya, Wilayah Persekutuan, Kuala Lumpur, Malaysia

Abstract

Background: Nurses face a higher risk of developing work-related musculoskeletal disorders (WMSDs) due to their primary roles in patient care. Participatory ergonomics (PE), an approach that integrates large-scale interventions performed at organizational and systems levels with small-scale interventions, is widely considered a promising approach to mitigate health problems at the workplace. However, its effectiveness in addressing WMSDs and secondary outcomes such as sickness absence and work performance among nurses is not fully understood.

Objective: This systematic review assessed the effectiveness of PE interventions in preventing WMSDs and mitigating two related outcomes, sickness absence and work performance, among nurses.

Methods: A literature search was performed in four electronic databases, PubMed, ScienceDirect, Scopus, and PsycNet, guided by the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) guidelines to retrieve relevant papers published between 2017 and 2023. Papers fulfilling the eligibility criteria were analyzed and subjected to quality appraisal.

Results: Overall, 19 papers were included in the final analysis. Various categories of ergonomic interventions were identified, with the predominant being exercise and physical activities, health promotional activities and training, educational programs, and patient handling devices. Multicomponent interventions, especially those involving physical activities and exercise, demonstrated stronger effects in reducing the risk of WMSDs at 6 months (OR 1.64, 95% CI 1.12 - 4.54) and 12 months postintervention (OR 2.70, 95% CI 1.52 - 4.51) compared with single interventions. However, most ergonomic interventions had no statistically significant effect (P>.05) on sickness absence and work performance. More than half (n=13) of the studies demonstrated moderate to high risk of bias, reflecting the need for better quality interventions.

Conclusions: Multicomponent interventions, particularly those involving physical activities and exercise, are more effective in reducing the risk of WMSDs among nurses compared with individual interventions. However, their long-term effects in addressing WMSDs, sick absenteeism, and work performance are still unclear. These gaps could be addressed by integrating organizational factors and prevention policies into existing ergonomic interventions, thereby offering opportunities to improve psychological health, job satisfaction, and work dynamics.

Trial Registration: PROSPERO CRD42023403981; https://www.crd.york.ac.uk/PROSPERO/view/CRD42023403981

(JMIR Hum Factors 2025;12:e68522) doi:10.2196/68522

KEYWORDS

musculoskeletal disorders; nurses; sick absenteeism; work performance; ergonomics

XSL•FO
RenderX

Introduction

Background

Nurses are highly exposed to risks of work-related disorders and injuries, culminating in burnout and low quality of care [1]. Such injuries may involve all body organs, especially the musculoskeletal system [2]. Musculoskeletal disorders (MSDs) refer to health problems affecting the locomotor apparatus, such as the skeleton, muscles, tendons, cartilage, ligaments, and nerves [3]. MSDs comprise several degenerative and inflammatory conditions affecting the joints, tendons, ligaments, muscles, peripheral nerves, and supporting blood vessels, leading to soreness and physical discomfort [3].

Among health care workers, nurses have a higher risk of developing work-related musculoskeletal disorders (WMSDs) [4]. Nurses play pivotal roles within the health care system, which can be physically challenging as they frequently help patients to mobilize, transfer between positions, and execute other daily life activities [5]. Multiple factors have been found to increase the risk of WMSDs among nurses, including physical load and work posture [6], psychosocial factors such as the presence of psychosomatic symptoms and personality, work tasks, and work organizational factors [5]. Thus, it is not surprising that nursing staff stationed in the operating room recorded a higher prevalence of occupational injuries compared with intensive care nurses, nonspecialized nurses, and radiologic technicians [7]. Events including static body posture, prolonged standing, using tools such as retractors during surgery, and manual tasks such as pushing, pulling, and lifting patients and surgery sets could elicit discomfort and WMSDs, particularly among nurses working in operating units [8]. Recent studies have also highlighted factors such as low educational level, increasing age, high BMI, and lifestyle as predictors of WMSDs among nurses [4].

Given the aforementioned points, several authors have advocated and explored effective workplace interventions to address the risk of WMSDs among nurses. Occupation-specific interventions and rehabilitation options have the potential to reduce functional disabilities among nursing staff that may result from these injuries [8,9]. Ergonomic interventions, particularly participatory ergonomics (PE), are a promising approach to reducing WMSDs. PE integrates large-scale interventions that are performed at organizational and systems levels with small-scale interventions where workers are opportune to use their knowledge in addressing work-related ergonomic issues [8]. It entails the active involvement of workers in controlling and planning a significant amount of their tasks or work activities while introducing ergonomic knowledge, changes, and procedures in order to enhance working conditions, productivity, safety, quality, and comfort [9].

Positive outcomes such as lower musculoskeletal symptoms and occupational injuries, improved psychological health, as well as a decrease in workers' compensation claims, productivity loss, and work absenteeism, were reported following PE interventions among various high-risk groups [9,10]. Despite ergonomic interventions being considered essential to reduce the increasing trend of WMSDs, there is no agreement or consensus on the most appropriate intervention, as the component and approach may vary with the workplace setting [[11]].

Importance of Ergonomics in Nursing

The risk of WMSDs can be predicted based on ergonomic, individual, and psychosocial factors [8]. Ergonomics scientifically discusses how a job can be aligned with a person's physical and psychological characteristics without causing any harm to the person's well-being and efficiency [11]. The onset and development of WMSDs can be prevented by acquiring more knowledge and applying ergonomics [12]. Nurses are motivated to become more efficient, experience higher job satisfaction, and reduce job-related stress, occupational-related diseases, accidents, and absenteeism when they work in appropriate ergonomic conditions [13].

An earlier study revealed several advantages and enhanced occupational health among community nurses following the implementation of a multifaceted ergonomic program [14]. Moazzami et al [15] reported that alterations in nurses' body movements were facilitated by ergonomic educational intervention, particularly from the contemplation and preparation phases to the action phase for adopting correct body postures in the operating room. Likewise, a significant reduction in musculoskeletal pain in the shoulder, neck, and knee regions was recorded among nurses in general wards after implementing an ergonomic educational intervention [16]. The use of ergonomic posture training among nursing assistants also led to a significant reduction in work-related low back pain 6 months after the intervention but did not affect disability scores [15]. However, some studies depicted limited success of ergonomic interventions on measured outcomes such as WMSDs [16,17], work productivity, and absenteeism [18], which may be linked to differences in the design of interventions, samples, data collection tools, settings, and research design. Thus, there is still limited information on the appropriate and effective ergonomic interventions for preventing and addressing WMSDs and their related consequences among nurses.

Rationale for Conducting This Review

Although more research is required to develop conclusive evidence on the effectiveness of ergonomics interventions in reducing WMSDs among nurses, pertinent information is yet to be gleaned from the several available studies in the literature. In the last decade, numerous studies have demonstrated the importance of diverse programs such as participatory, educational, single, and multidisciplinary ergonomics interventions among nursing professionals [12,19,20]. Nevertheless, findings from these studies are yet to be systematically reviewed. Most of the systematic reviews focused on office workers or workplace settings [21] and dental and health care professionals [22]. However, Sun et al [23] in a recent systematic review and meta-analyses, compared the efficacy of nondrug interventions on nonspecific low back pain among nurses and found that the most effective intervention was a combination of low back exercises and health education, followed by single low back exercises and yoga.

XSL•FO
RenderX

Previous literature reviews have investigated the effectiveness of interventions among nurses and in nurses' homes [14], including nursing assistants' risk of handling obese and overweight patients and interventions to reduce low back pain [24]. However, none of the past reviews looked into how ergonomic interventions ameliorate WMSDs and their consequences, particularly among nurses. The present review differs from prior attempts by systematically analyzing the various ergonomic interventions adopted in preventing WMSDs and other closely related outcomes, including sickness absenteeism and work performance or productivity in nurses. It seeks to provide evidence-based data and scientific support for developing an ergonomic intervention to prevent WMSDs among nursing professionals.

The objective of this systematic review is to assess the effectiveness of ergonomic interventions in preventing WMSDs and mitigating two related outcomes, sickness absence and work performance, among nurses.

Methods

Overview

This systematic review was conducted according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement. The PRISMA statement was developed to assist researchers in preparing a complete, accurate, and transparent review that is beneficial to users [25]. The PRISMA 2020 statement includes new reporting guidelines, particularly for systematic reviews of studies that appraise the effects of health interventions. The PRISMA statement was considered suitable for the present systematic review, given its provision for a clear definition of research questions, identification of the exclusion and inclusion criteria, and assessment of relevant and accessible scientific papers within a specific timeframe. In addition, this review has been registered in International Prospective Register of Systematic Reviews (PROSPERO) with the registration number CRD42023403981.

Formulation of Research Questions

Formulation of a concise and clear research question is an important phase to guide the authors in many aspects of the systematic review. PICO, a research question development tool (RQDT), was used in the formulation of the research question for this review. PICO is an acronym in which P is population, I is intervention, C is comparison, and Ois outcome, and the 3 main concepts of PICO are population or problem, interest, and context. The authors of this review made an effort to incorporate all the aspects in synthesizing the main research question. Considering the population of interest as nurses, ergonomic intervention as the intervention, other nonergonomic interventions as the comparison group, and WMSDs, work performance, and sickness absenteeism as the outcomes, the research questions for this systematic review are as follows:

1. What are the commonly implemented participatory ergonomic interventions to prevent WMSDs among nurses?
2. Are ergonomic interventions effective in preventing WMSDs, sickness absenteeism and enhancing work performance among nurses?

Systematic Literature Search

Identification

The identification process entailed the determination of suitable keywords to search for sufficient information. Aligning with the research question, a few main keywords were recognized, namely MSDs, WMSDs, musculoskeletal pain, and musculoskeletal symptoms. In terms of the study population, the selected keywords encompass nurse, nursing, nurse assistants, and health care professionalsor workers. While the former is more specific for the present study context, the term "health care workers" was considered since such studies may include diverse professionals including those from the nursing field. Moreover, by taking the term "health care" into account in the identified keywords, health care professions that entail heavy-lifting jobs (ie, important risk factors for WMSDs) can be covered in the literature search.

Search Strategy

Search functions, including truncation and Boolean operators (OR and AND), were used in combining the selected keywords, leading to the formation of search strings. Search strings relating to study participants or study population, intervention, and study outcomes were synthesized as follows:

1. Study participants: "Nurse" OR "Nursing" OR "Nurse assistant" OR "Healthcare professionals" OR "Healthcare workers."
2. Intervention: "Ergonomic" OR "Ergonomic Intervention" OR "Intervention" OR "Participatory Ergonomics"
3. Outcomes: "Work-related musculoskeletal disorders" OR "Musculoskeletal disorder" OR "Musculoskeletal Pain" OR "Musculoskeletal symptom" OR "Sick Absence" OR "Sick leave" OR "Sick Absenteeism" OR "Work Performance" OR "Work Productivity."

All the search strings outlined above were combined using the Boolean operator "AND," leading to the final search string used for each database. However, since each database is unique and programmed to operate differently, the search strings were modified accordingly to suit each database and to maximize the number of hits. For instance, ScienceDirect only accepts a maximum of 8 search terms, while PubMed requires all keywords to be placed in parentheses.

A detailed literature search was then carried out independently by the researcher and a research collaborator in four electronic databases: PubMed, ScienceDirect, Scopus, and PsycNet. These databases were selected as they offer an extensive collection of relevant journals, provisions for customizing an advanced search, and storage of retrieved papers. The literature search was first conducted in March 2023 and updated in March 2024 to gather original research papers written in English and published between 2017 and 2023. The search string used in all 4 databases, the number of papers retrieved, and the date of acquisition are presented in Table 1.

XSL•FO
RenderX

Table . Search terms and the total number of papers retrieved from each database.

Search engines	Search string and search terms		Number of hits (N=5867)	Acquisition date
PubMed	Main search terms using document title, abstract, and keywords	(“Nurse” OR “Nursing” OR “Nurse assistant” OR “Healthcare professionals” OR “Healthcare workers”) AND (“Ergonomic” OR “Ergonomic Intervention” OR “Intervention” OR “Participatory Ergonomics”) AND (“work-related musculoskeletal disorders” OR “Musculoskeletal disorder” OR “Musculoskeletal Pain” OR “Musculoskeletal symptom” OR “Sick Absence” OR “Sick leave” OR “Sick Absenteeism” OR “Work Performance” OR “Work Productivity”)	525	March 4, 2024
Scopus	Main search terms using document title, abstract, and keywords	(“Nurse” OR “Nursing” OR “Nurse assistant” OR “Healthcare professionals” OR “Healthcare workers”) OR (“Ergonomic” OR “Ergonomic Intervention” OR “Intervention” OR “Participatory Ergonomics”) AND (“work-related musculoskeletal disorders” OR “Musculoskeletal disorder” OR “Musculoskeletal Pain” OR “Musculoskeletal symptom” OR “Sick Absence” OR “Sick leave” OR “Sick Absenteeism” OR “Work Performance” OR “Work Productivity”)	1009	March 5, 2024
ScienceDirect	Main search terms using document title, abstract, and keywords	(“Nurse” OR “Healthcare professional”) AND (“Ergonomic” AND “Intervention” OR “Participatory Ergonomics”) AND (“Musculoskeletal” OR “Sick” OR “Work Performance”)	1617	March 7, 2024
PsycNet	Main search terms using document title, abstract, and keywords	(“Nurse” OR “Nursing” OR “Nurse assistant” OR “Healthcare professionals” OR “Healthcare workers”) AND (“Ergonomic” OR “Ergonomic Intervention” OR “Intervention” OR “Participatory Ergonomics”) AND (“work-related musculoskeletal disorders” OR “Musculoskeletal disorder” OR “Musculoskeletal Pain” OR “Musculoskeletal symptom” OR “Sick Absence” OR “Sick leave” OR “Sick Absenteeism” OR “Work Performance” OR “Work Productivity”)	2716	March 8, 2024

XSL•FO
RenderX

Eligibility Criteria and Screening

Upon completing the literature search, the retrieved papers were screened according to the eligibility criteria. The first screening process was performed by using the filter feature provided in all 4 search engines. In terms of type of publication, only original papers were selected for further review, whereas chapters in books, books, and conference proceedings were removed. As for review papers, only those relevant to our review objective were assessed for vital information and to identify research papers that might have been missed during the literature search.

In terms of study duration, only papers in the last 7 years were considered for full-text reading. Based on preliminary analysis, the proposed timeline is sufficient to yield papers to facilitate a representative and comprehensive review of the research topic. Finally, only papers written and published in English were included in this study. This decision was taken to prevent any challenges during the data extraction since the researcher is only proficient in English. Issues arising during the process were resolved by consensus between all the review authors.

Data Extraction and Analysis

All the final papers were subjected to qualitative analysis. An Excel (Microsoft) spreadsheet form was generated before reviewing the papers to document basic data from the studies, such as the authors' names, titles, journals, and publication years. Data were extracted from each paper by scanning through the abstracts, methods, and results before proceeding to the full text. These processes assisted the researcher in recording the year of publication, study location, study design, specific ergonomic intervention used, instrument used, and main results or emerging themes.

First, available data were analyzed thematically to generate themes and subthemes. The potential themes considered were as follows: (1) study characteristics, (2) quality assessment, (3) common types of participatory ergonomic interventions, (4) multicomponents versus single components, and (5) effects of participatory ergonomic interventions on WMSDs, sickness absence, and work performance. Contrast analyses, specifically pairwise comparisons, were performed to ascertain the effect size report in each study involving intervention and control groups. For studies reporting WMSDs as categorical outcomes, the odds ratio (OR, 95% CI) was computed to measure the strength and direction of the association. Meanwhile, the mean difference (95% CI) was computed for studies reporting the outcome variables as continuous outcomes.

Quality Appraisal

Quality appraisal of each paper was performed using the widely accepted quality assessment tools for intervention studies, the *Cochrane Handbook for Systematic Reviews of Interventions* and the Risk of Bias in Non-randomized Studies-of Interventions (ROBINS-I) tool [26,27]. The Cochrane tool was used for experimental studies reporting randomized interventions.

A total of 5 domains were considered for explicit reporting, comprising the sequence generation, allocation concealment, blinding, and completeness of data. For the first domain, the use of a randomized sampling method was assessed either by a random number table, tossing a coin, or computerized-generated numbers. Regarding the blinding of recruited patients, studies were considered adequate if both the intervention providers and outcome assessors were blinded (ie, minimal risk of compromised blinding) and inadequate if no blinding or partial blinding was performed. For outcome measures, trials were recorded as adequate if there were no missing data or if they stated the reasons for missing data that were unlikely to be associated with the true outcome. In addition, the studies were evaluated for selective reporting. Studies without sufficient information regarding the outcome reporting process were considered unclear. For all the aforementioned criteria, studies with insufficient information about each criterion were considered unclear. Other potential sources of bias were assessed, such as demographic characteristics between the treatment groups, by considering the baseline information of the participants.

Using the ROBINS-I tool, similar procedures described above were applied for the nonrandomized studies. The quality assessment tool focused on the randomization method, blinding of drug administration and assessment of treatment outcomes, data analysis techniques, sample size, and steps taken in addressing missing data and loss to follow-up. Studies reporting at least 70% (8/12) of the items in the quality assessment tool were considered of higher methodological quality or low risk of bias [27].

Ethical Considerations

Systematic review is part of a module development.Module development has obtained study ethics approval from the Medical Research Ethical Committee of Universiti Malaya Medical Centre (MREC ID No:20221019-11632) and National Medical Research Register (NMRR ID-23-01804-F44).

Results

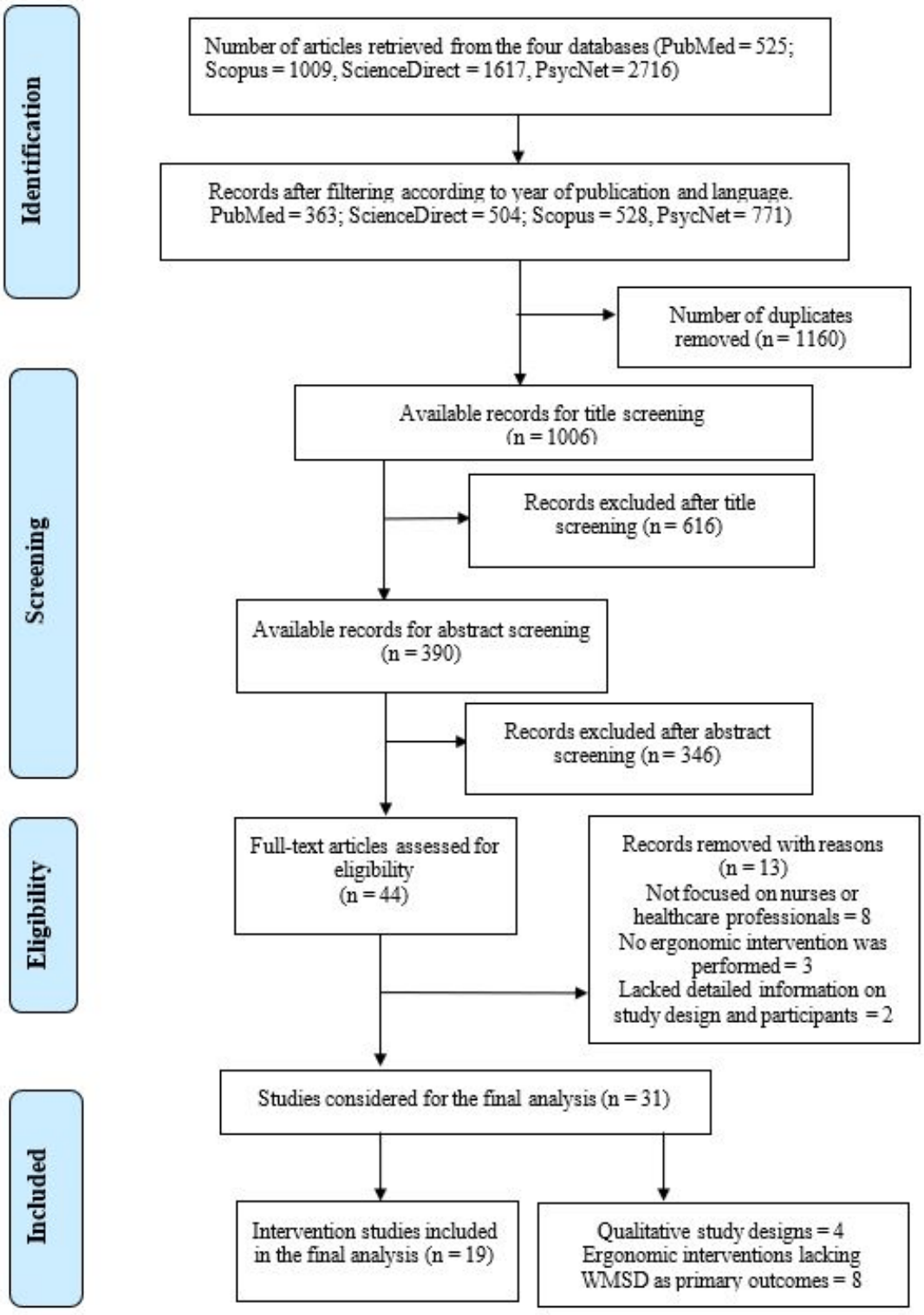
Search Results

The initial search process before filtering and screening resulted in a total of 5867 papers, comprising 525 from PubMed, 1009 from Scopus, 1617 from ScienceDirect, and 2716 from PsycNet (Figure 1). Upon considering the years of publication (2017 - 2023) and language (ie, English), the remaining papers were reduced to 2166 (ie, PubMed=363, ScienceDirect=504, Scopus=528, and PsycNet=771).

Thereafter, 1160 papers were identified as duplicates and removed accordingly. To further reduce the number of records, the remaining results from each database were sorted and subjected to title screening. Out of the remaining papers (n=1006), those with irrelevant titles (n=616) and abstracts (n=346) were removed, leading to a total of 44 papers for full-text reading. Around 25 studies were excluded after reading the full texts because of irrelevant study populations (ie, not involving nurses or health care professionals, n=8), not reporting any ergonomic intervention (n=3), unclear study design (n=2), qualitative study designs (n=4), and ergonomic interventions with WMSDs as primary outcomes (n=8).

XSL•FO
RenderX

Figure 1. Literature search process using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines. WMSD: work-related musculoskeletal disorder.



Descriptive Findings

Overall, 19 papers were experimental studies involving ergonomic interventions among nurses and other related health care professionals. These papers reported at least one of the following outcomes: WMSDs symptoms or pain, sickness absence, and work performance. Hence, all 19 papers were included in the final analysis and systematically reviewed. Table S1 in Multimedia Appendix 1 depicts the studies and participants' characteristics of types of ergonomic interventions and a summary of the main findings.

Ergonomic Interventions Reported in the Papers

As shown in Table 2, 7 categories of ergonomic interventions were identified in the 19 studies as follows: (1) exercise and physical activities, (2) educational programs, (3) health

XSL•FO
RenderX

promotional activities and training, (4) patient handling devices or equipment (acquired ergonomics equipment), (5 programmes) workstation and administrative interventions (organizational management), (6) counseling, and (7) others (case management and transcutaneous electrical nerve stimulation). The predominant ergonomic intervention was exercise and physical activities (11/19, 58%), followed by health promotional activities and training (8/19, 42%), educational programs (7/19, 36%), patient handling devices (4/19, 21%), workstation and administrative interventions (n=3), counseling (n=3), and case management (n=2). Furthermore, 3 studies reported specific exercises for back pain, 2 for neuromuscular pain, and 1 for neck pain. Around 10 studies (53%) each entailed either only multicomponent or single intervention, whereas 1 study compared both types of interventions.

Table . Types of participatory ergonomic interventions applied in the 19 studies.

Reference	Educational ergonomics	Counseling	Patient handling device or ergonomics equipment	Health promotion activities	Workstation adjustment and administrative	Exercise and physical activities	Others (social media/website, case management, nerve stimulation, and online apps)
Abdollahi et al [8]	P[a]	A[b]	A	A	A	A	A
Jalalvandi et al [16]	A	A	P	A	A	P	A
Soler-Font et al [19]	A	A	A	P	A	P	P
Yang et al [28]	P	A	A	P	A	A	A
Rasmussen et al [10]	A	A	A	P	A	P	A
Nguyen et al [29]	A	A	A	P	A	P	A
Pereira et al [30]	A	A	A	P	P	P	A
Chaiprateep et al [31]	A	A	A	A	A	P	A
Suni et al [32]	A	P	A	A	A	P	A
Higuchi et al [33]	P	P	A	A	A	A	A
Jakobsen et al [34]	P	A	A	A	A	A	A
Al-Quasi et al [35]	A	A	P	A	A	A	A
Sezgin and Esin [36]	P	A	A	A	A	A	A
Beyan et al [17]	A	P	P	P	P	A	A
Wulff Risør et al [37]	A	A	P	A	A	A	A
Imai et al [38]	P	A	A	A	A	P	A
Marshall et al [39]	P	A	P	A	P	A	A
Taulaniemi et al [40]	A	A	A	A	A	P	A
Hosseini et al [41]	A	A	A	P	A	A	A

[a]P: present.

[b]A: absent.

[c]NA: not available.

XSL•FO
RenderX

Work-Related Musculoskeletal Disorders and Pain

Single Interventions

Among the 10 papers involving a single ergonomic intervention, significant reductions in either the prevalence, frequency, or pain intensity of WMSDs were reported in 6 studies [8,31,34,36,40,41]. The educational programs delivered by experts on ergonomic principles and risk factors of WMSDs led to a significant decrease (P<.05) in the risk of WMSDs, particularly in the ankle, wrist, lower back, and neck, pain intensity and Rapid Upper Limb Assessment ergonomic risk scores among nursing personnel with similar low effect sizes in both studies (0.30 vs 0.40) [8,36]. Likewise, ergonomic exercise (ie, neuromuscular and back exercise) performed in 3 studies was associated with significantly lower scores for low back pain (P=.05, effect size=0.45) [40] and overall pain scores (effect size=0.30 - 0.52) compared with either a control group or other nonergonomic single intervention such as transcutaneous electrical nerve stimulation [16,31].

In addition, 2 studies involving single ergonomic interventions, particularly online applications, were identified in this review [8,33]. Both studies recorded significantly lower low back pain scores in the intervention group compared to the control, with moderate to low effect size (ie, 0.30 - 0.50). Higuchi et al [33] focused on individual online therapy for low back pain, whereas Hosseini et al [41] used a Nursing Stretch Break app and reflected a significant reduction in WMSDs symptoms and pain intensity in nurses' body parts except the elbows and knees.

In contrast, 3 of the remaining studies found no significant effects of single ergonomic interventions such as educational programs [34], individual-based interventions [17], and patient handling devices [37] in preventing or managing WMSDs among nursing personnel. In Jakobsen et al [34] and Risør et al [37], despite improvement in the usage of patient-assisted devices posteducational program, no significant change was observed in the risk of low back pain and other WMSDs at 12-month follow-up. Similarly, individual-based ergonomic programs, including training, engineering, and administrative interventions, did not affect WMSDs scores at 18 months follow-up [17].

Multicomponent Interventions

Among the studies involving multicomponent ergonomic interventions, despite several combinations of interventions being reported, the results were analyzed based on whether they included exercise or physical activities or health promotional activities.

Around, 3 studies entailed a combination of health promotion activities and case management [19], workstation adjustment, and training without performing any exercise or physical activities [18,28]. Results from the studies revealed a significant reduction in the risk of WMSDs with an OR ranging from 1.64 (95% CI 1.12 - 4.54) at 6 months postintervention to 2.70 (95% CI 1.52 - 4.51) at 12 months postintervention [19,28]. In contrast, Rasmussen et al [10] found no significant effect of an intervention comprising three training workshops centred on focusing on workstation adjustment and health promotion in addressing WMSDs.

As for the 3 multicomponent interventional studies involving physical exercise, the pain intensity arising from WMSDs such as the lower back and shoulder or upper arm, neck, and hand or wrists body regions reduced significantly either 6 months or 12 months postintervention [28,32,38]. The effect size reported in these studies ranged from moderate to high (0.54 - 0.80), reflecting the greater efficacy of multicomponent PE comprising general or specific body exercise, for the management of WMSDs in nurses. Notably, the strongest effect size was observed in studies whereby exercise or physical activities were combined with either health promotion activities or educational or counseling programs such as pain neuroscience education and back care counseling [32,38]. Finally, the study by Marshall et al [39] found a significantly moderate reduction (OR=0.40) in subsequent acute cases of WMSD following multifactorial interventions comprising equipment intervention, educational ergonomics, and no-lift policies (administrative interventions).

Work Performance and Related Outcomes

Only 4 of the 19 papers considered work performance or presenteeism as musculoskeletal-related work performance, with 3 comprising multicomponent [19,30,38] and 1 including only a single intervention [40]. Both Imai et al [38] and Pereira et al [30] performed a dual-component intervention (ie, exercise and Pain Neuroscience Education (PNE) and workstation ergonomics and neck-specific exercise), leading to significant improvements in presenteeism (moderate to strong effect size=0.52 - 0.78) at 12 months follow-up. Meanwhile, using the same follow-up period, Soler-Font et al [19] found no significant difference in work functioning and performance in nurses exposed to a combination of PE, health promotion activities, and case management relative to the control group.

For the single-component studies, Taulaniemi et al [40] revealed a significant reduction in work-interfering pain (effect size=0.40; P=.046) in heavy nursing duties. Suni et al [32] also depicted that operating room nurses subjected to only back care counseling or neuromuscular exercise had a higher risk (OR 0.46) of work-related fear of pain compared with those administered a combination of both ergonomic interventions.

Sick Absenteeism

Sick absenteeism was reported in 6 studies included in this review, with 3 reporting multicomponent [10,19,30], 2 involving only single ergonomic interventions [17,32,37], and 1 study comparing both types [32].

For the multicomponent interventions, 2 studies found a significant reduction in musculoskeletal-related sickness absence in the intervention relative to comparison groups at 20 weeks [10] and 12 months postintervention [30]. Specifically, Rasmussen et al [10] entailed 20 weeks of PE comprising training workshops and health promotion activities, whereas Pereira et al [30] compared nurses subjected to workstation ergonomics and neck-specific exercises to those exposed to health promotion activities or information. Both studies found a moderate to strong effect size (−0.48 days per month, 95% CI −0.8 to −0.10 vs −0.72 days per month, 95% CI −0.52 to −0.94). In contrast, no significant difference was detected in sickness absence between nurses who administered PE, health

XSL•FO
RenderX

promotion activities, and case management compared with the control at 12 months postintervention [19].

The 2 studies involving single interventions, such as individual-based training, stretching exercises, motivation meetings, and patient handling devices or equipment, found no significant effect on sickness absence among nurses at 12 months [37] and 18 months of foled a singlelow-up [17]. These studies acknowledged the low likelihood of single-component ergonomic interventions in addressing musculoskeletal-related sickness absence, thus advocating for multicomponent programs, particularly administrative measures. This was further evidenced in the study comparing 4 different intervention arms, whereby only the combined arm (ie, neuromuscular exercise and back care counseling) demonstrated a significantly lower mean number of sickness absence days (effect size=0.55, *P*=.025) relative to the exercise-only groups at 12-month follow-up.

Quality Assessment Results

Out of the 11 RCTs evaluated using the Cochrane tool, 6 studies demonstrated a low risk of bias [10,19,28,30,32,34], 3 recorded a moderate risk of bias [16,33,40], and 2 papers had a high risk of bias [39,42], as shown in Table 3. All 6 studies with a low risk of bias showed high methodological quality in terms of randomization methods, concealment of treatment allocation, no significant difference between treatment groups at baseline, intervention delivery and adherence, as well as accounting for missing data and adjustment for possible confounders. Only one of the 6 studies performed a single-blinding procedure for the outcome assessment [32]. Meanwhile, the main issues observed in studies with moderate to high risk of bias were unclear or lack of information on the participants' baseline characteristics (n=4), a strong potential for confounding of the intervention's effect (n=4), lack of appropriate analysis method for confounder adjustment (n=5), no information on how missing data were addressed (n=3), incoherent onset of intervention and follow-up in most participants (n=3), unclear definition of intervention groups (n=2), poor adherence to the intervention regimen (n=2), and indications of selective reporting (n=2).

XSL•FO
RenderX

Table . Risk of bias assessment of the experimental studies using the Cochrane tool (n=11).

Items	Jalalvandi et al [16]	Soler-Font et al [19]	Yang et al [28]	Rasmussen et al [10]	Pereira et al [30]	Suni et al [32]	Higuchi et al [33]	Jakobsen et al [34]	Imai et al [38]	Marshall et al [39]	Taulaniemi et al [40]
Bias due to confounding											
Is there potential for confounding of the effect of the intervention in this study?	N[a]	N	N	N	N	N	N	N	Y[b]	Y	N
Was the analysis based on splitting participants' follow-up time according to the intervention received?	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y
Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome?	NA[c]	N	N	N	NA	NA	NA	NA	N	Y	NA
Baseline confounding and time-varying confounding											
Did the authors use an appropriate analysis method that controlled for all the important confounding domains?	Y	Y	Y	NA	NA	Y	Y	Y	Y	N	Y

XSL•FO
RenderX

Items	Jalalvandi et al [16]	Soler-Font et al [19]	Yang et al [28]	Rasmussen et al [10]	Pereira et al [30]	Suni et al [32]	Higuchi et al [33]	Jakobsen et al [34]	Imai et al [38]	Marshall et al [39]	Taulaniemi et al [40]
Did the authors control for any postintervention variables that could have been affected by the intervention?	N	Y	N	Y	Y	Y	N	Y	Y	N	N
Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding?	Y	Y	Y	NA	NA	Y	Y	Y	Y	Y	Y
Bias in selection of participants for the study											
Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of the intervention?	N	N	N	NA	N	N	N	Y	N	N	N
Do the start of follow-up and the start of intervention coincide for most participants?	Y	Y	Y	Y	Y	Y	Y	Y	N	N	Y

XSL•FO
RenderX

Items	Jalalvandi et al [16]	Soler-Font et al [19]	Yang et al [28]	Rasmussen et al [10]	Pereira et al [30]	Suni et al [32]	Higuchi et al [33]	Jakobsen et al [34]	Imai et al [38]	Marshall et al [39]	Taulaniemi et al [40]
Bias in classification of interventions											
Were intervention groups clearly defined?	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y
Was the information used to define intervention groups recorded at the start of the intervention?	Y	Y	N	Y	N	Y	Y	Y	Y	N	Y
Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome?	N	N	N	N	N	N	N	N	N	Y	N
Bias due to deviations from intended interventions											
Were there deviations from the intended intervention beyond what would be expected in usual practice?	NA	N	N	N	N	N	N	N	N	NA	NA
Effect of starting and adhering to intervention											

XSL•FO
RenderX

Items	Jalalvandi et al [16]	Soler-Font et al [19]	Yang et al [28]	Rasmussen et al [10]	Pereira et al [30]	Suni et al [32]	Higuchi et al [33]	Jakobsen et al [34]	Imai et al [38]	Marshall et al [39]	Taulaniemi et al [40]
Were important cointerventions balanced across intervention groups?	N	Y	Y	Y	Y	Y	Y	Y	Y	N	N
Was the intervention implemented successfully for most participants?	NA	Y	Y	Y	Y	Y	Y	Y	Y	Y	NA
Did study participants adhere to the assigned intervention regimen?	NA	NA	NA	Y	NA	NA	Y	NA	Y	Y	NA
Bias due to missing data											
Were outcome data available for all, or nearly all, participants?	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
Were participants excluded due to missing data on intervention status?	Y	N	N	N	N	N	Y	NA	NA	Y	Y
Were participants excluded due to missing data on other variables needed for the analysis?	N	N	N	N	N	N	Y	NA	N	Y	N
Bias in measurement of outcomes (blinding)											

XSL•FO
RenderX

Items	Jalalvandi et al [16]	Soler-Font et al [19]	Yang et al [28]	Rasmussen et al [10]	Pereira et al [30]	Suni et al [32]	Higuchi et al [33]	Jakobsen et al [34]	Imai et al [38]	Marshall et al [39]	Taulaniemi et al [40]
Could the outcome measure have been influenced by knowledge of the intervention received?	N	N	N	N	N	N	N	N	Y	N	N
Were outcome assessors aware of the intervention received by study participants?	N	N	Y	N	Y	Y	Y	Y	N	Y	N
Were the methods of outcome assessment comparable across intervention groups?	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y
Were any systematic errors in measurement of the outcome related to intervention received?	N	N	N	N	N	N	N	N	N	N	N
Is the reported effect estimate likely to be selected based on the result?											
Multiple outcome measurements within the outcome domain?	N	N	N	Y	N	N	N	N	N	N	N

XSL•FO
RenderX

Items	Jalalvandi et al [16]	Soler-Font et al [19]	Yang et al [28]	Rasmussen et al [10]	Pereira et al [30]	Suni et al [32]	Higuchi et al [33]	Jakobsen et al [34]	Imai et al [38]	Marshall et al [39]	Taulaniemi et al [40]
Multiple analyses of the intervention-outcome relationship?	N	N	N	N	N	N	N	N	N	N	N
Different subgroups?	N	N	N	Y	N	N	N	N	N	N	N
Overall risk of bias grade	Moderate	Low	Low	Low	Low	Low	Moderate	Low	High	High	Moderate

[a]N: no.

[b]Y: yes.

[c]NA: not available.

Table 4 depicts the quality appraisal for the nonrandomized or quasi-experimental studies (n=8), with 6 [8,17,31,35-37], and 2 papers identified as having high risk and moderate risk of bias, respectively [17,29]. The main methodological weaknesses observed in these studies were unclear or lack of information on the participants' baseline characteristics (n=3), the potential for confounding of the intervention's effect (n=3), limited information on missing data analysis (n=5), unclear definition of intervention groups (n=4), poor adherence to the intervention regimen (n=3), low sample size for acceptable statistical power (n=3), lack of appropriate analysis method for confounder adjustment (n=6), and disparity in the onset of intervention, and follow-up in most participants (n=4). In addition, 4 of the studies did not report the measures taken to ensure consistency in intervention delivery, and only 2 reported measures of intervention adherence. None of these studies performed a blinding procedure, as the active treatment group was usually recognisable.

XSL•FO
RenderX

Table . Risk of bias assessment of the experimental studies using the Risk of Bias in Non-randomized Studies - of Interventions tool (n=8).

Items	Chaiprateep et al [31]	Abdollahi et al [8]	Nguyen et al [29]	Al-Qaisi et al [35]	Sezgin and Esin [36]	Beyan et al [17]	Risør et al [37]	Hosseini et al [41]
Bias due to confounding								
Is there potential for confounding of the effect of intervention in this study?	Y[a]	Y	N[b]	Y	Y	N	Y	N
Was the analysis based on splitting participants' follow up time according to intervention received?	Y	N	Y	Y	N	Y	N	Y
Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome?	N	Y	NA[c]	Y	Y	NA	Y	NA
Baseline confounding and time-varying confounding								
Did the authors use an appropriate analysis method that controlled for all the important confounding domains?	Y	N	Y	Y	N	Y	N	Y
Did the authors control for any post-intervention variables that could have been affected by the intervention?	Y	N	Y	N	N	Y	N	N
Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding?	Y	Y	Y	N	Y	Y	Y	Y

XSL•FO
RenderX

Items	Chaiprateep et al [31]	Abdollahi et al [8]	Nguyen et al [29]	Al-Qaisi et al [35]	Sezgin and Esin [36]	Beyan et al [17]	Risør et al [37]	Hosseini et al [41]
Bias in selection of participants into the study								
Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention?	N	N	Y	N	N	N	N	N
Do start of follow-up and start of intervention coincide for most participants?	N	N	Y	N	N	Y	N	Y
Bias in classification of interventions								
Were intervention groups clearly defined?	Y	N	Y	Y	N	Y	N	Y
Was the information used to define intervention groups recorded at the start of the intervention?	N	N	Y	Y	N	Y	N	Y
Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome?	Y	Y	N	N	Y	N	Y	N
Bias due to deviations from intended interventions								
Were there deviations from the intended intervention beyond what would be expected in usual practice?	NA	NA	NA	N	NA	NA	NA	NA

XSL•FO
RenderX

Items	Chaiprateep et al [31]	Abdollahi et al [8]	Nguyen et al [29]	Al-Qaisi et al [35]	Sezgin and Esin [36]	Beyan et al [17]	Risør et al [37]	Hosseini et al [41]
Effect of starting and adhering to intervention								
Were important co-interventions balanced across intervention groups?	NA	NA	N	Y	N	N	N	N
Was the intervention implemented successfully for most participants?	Y	NA	NA	Y	Y	NA	Y	Y
Did study participants adhere to the assigned intervention regimen?	NA	NA	Y	Y	Y	Y	Y	NA
Bias due to missing data								
Were outcome data available for all, or nearly all, participants?	Y	Y	Y	Y	Y	Y	Y	Y
Were participants excluded due to missing data on intervention status?	NA	N	Y	NA	Y	Y	Y	Y
Were participants excluded due to missing data on other variables needed for the analysis?	N	N	Y	N	Y	Y	NA	NA
Bias in measurement of outcomes (blinding)								
Could the outcome measure have been influenced by knowledge of the intervention received?	Y	Y	Y	Y	Y	Y	Y	N

XSL•FO
RenderX

Items	Chaiprateep et al [31]	Abdollahi et al [8]	Nguyen et al [29]	Al-Qaisi et al [35]	Sezgin and Esin [36]	Beyan et al [17]	Risør et al [37]	Hosseini et al [41]
Were outcome assessors aware of the intervention received by study participants?	N	Y	N	N	Y	N	Y	N
Were the methods of outcome assessment comparable across intervention groups?	N	N	Y	Y	N	Y	N	Y
Were any systematic errors in measurement of the outcome related to intervention received?	N	N	N	N	N	N	N	N
Is the reported effect estimate likely to be selected on the basis of the result?								
Multiple outcome measurements within the outcome domain?	N	N	N	N	N	N	N	N
Multiple analyses of the intervention-outcome relationship?	N	N	N	N	N	N	N	N
Different subgroups?	N	N	N	N	N	N	N	N
Overall risk of bias	High	High	Moderate	High	High	Moderate	High	High

[a]Y: yes.

[b]N: no.

[c]NA: not available.

Discussion

Principal Findings

Despite the heterogeneity of the 19 studies in this review in terms of study design, interventions used, sample size, measured outcomes, and research instruments, we were able to identify the participatory ergonomic interventions designed specifically for nursing professionals. In addition, the research designs and results revealed the methodological quality of the studies and evidence of the interventions in ameliorating WMSDs among this high-risk group of health care workers.

Descriptive findings from the systematic review revealed that the predominant PE interventions, specifically for preventing and managing WMSDs among nurses, were physical activities or exercise, health promotional activities, educational programs, and patient handling devices. These findings are consistent with previous systematic and integrative reviews focusing on interventions to address WMSDs and their secondary outcomes among health care professionals [43,44]. Physical activities and specific exercises are frequently considered in PE interventions either as individual or multifaceted approaches, given their efficacy in reducing musculoskeletal pain and associated symptoms [44]. Educational programs and health promotional activities are also integrated with most multicomponent

XSL•FO
RenderX

interventions to enhance nurses' knowledge, attitude, and self-efficacy toward lifestyle changes and behaviors to mitigate risk factors of WMSDs [43,45].

Single or individual interventions were applied in 10 studies ranging from physical activity to educational programs and the use of patient handling devices. Nevertheless, contradictory findings were gleaned from the studies, with few reflecting a low to moderate reduction in WMSD risk and ergonomic risk scores following educational programs [8,36]. Ergonomic exercise (ie, neuromuscular and back exercise) conducted in 3 studies also led to significantly lower scores for low back pain [40] and overall body pain scores [16,31]. These findings corroborate reports of individual exercise and physical activity proposed by Jakobsen et al [34] for the generality of health care professionals. Several prior studies have equally emphasized the positive effect of physical exercise on MSDs [20,29]. Moreover, physical exercise helps in preventing recurring and new episodes of WMSD pain [46], improving endurance, strength, and neuromuscular control in rehabilitation patients [47]. In contrast, 3 studies found no significant preventive effects of single interventions such as educational programs, individual-based interventions, and patient handling devices in ameliorating WMSDs among nurses and health care professionals [17,34,37]. These studies reflect the need for a multifaceted approach to addressing musculoskeletal pain, especially when designing interventions that include educational ergonomics. In other words, rather than using educational programs in isolation, they should be combined with other interventional approaches such as exercise.

As observed in this review, several studies entailed multifactorial interventions—a combination of procedures considering numerous factors and events acting on multiple levels [48]. This approach aligns with the multifaceted causes of WMSDs [19], which require integrated and comprehensive management strategieseducational program rather than isolated approches [42]. Comparisons between multicomponent and single interventions depicted that the former was more effective in reducing the risk of WMSDs among nurses relative to single type of ergonomic interventions. These results were evidenced in several studies involving a combination of health promotion activities and case management [19], workstation adjustment and training [10,28], educational ergonomics and patient handling equipment [39], as well as those involving physical activities and exercise [29,32,38]. In addition, moderate to high effect sizes were reported in studies where multicomponent interventions included at least one or more physical activities and exercise, whereas effects in those without exercise were generally low to moderate. These findings are consistent with a recent review conducted among health care professionals, in which the strongest reduction in WMSD injuries and pain stemmed from combined manual or mechanical patient lifting and handling training [44]. This could be explained by the fact that exercise, patient handling devices, and ergonomic training offer opportunities to mitigate physical risk factors and insufficient handling knowledge, which are the main contributing events to injuries among nurses and health care professionals [49,50]. More importantly, these studies highlight the significance of multifaceted PE aimed at comprehending the appropriate handling of patients or using the principles of biomechanics in executing such activities [30] and guidelines implementation [18]. These findings align with the Occupational Safety and Health Administration recommendations for increased availability of assistive facilities and devices, including both sophisticated and basic devices, that are crucial for safe patient handling [40].

Only a few reviewed studies investigated work performance and sickness absence, which may be because these events are secondary outcomes of WMSDs. This finding is not surprising, given that a systematic review of RCTs aimed at reducing sickness absence among health care workers identified only 7 relevant papers [45]. Results from the present review highlight the need to further explore how nurses' work performance and sickness absence can be improved by ergonomic interventions. Most studies found no significant effects of multicomponent and individual interventions on these secondary outcomes at different follow-up points [19,28,32]. For instance, complaints of low back pain and its associated cost persisted among nurses for several weeks and months after performing neuromuscular exercise and back care counseling [32].

These studies raise the question of whether some multifaceted PE interventions, such as ergonomic education, health promotion activities and case management, neuromuscular exercise, and back care counseling, are either productive or counterproductive in improving secondary outcomes primarily linked to WMSDs among nurses. On the contrary, reduced work-day absences and increased safe patient handling were observed in health care professionals subjected to task-specific and multifactorial interventions such as mechanical and practical in-ward activities [50]. These promising results from studies conducted among health care workers emphasized changing the work dynamics, increasing job satisfaction, productivity, and the complicity between workers [32]. Moreover, psychological stress constitutes approximately 28% and 84% of upper back injuries in health care professionals [51].

Although qualitative studies were not included in this review, results from the studies highlighted the importance of considering psychological health problems, ergonomic knowledge, and preventive behaviors that could shape the risk of WMSDs, sickness absence, and work performance [11,18]. Organizational, physical, psychosocial, and personal factors were found to play pivotal roles in the development of musculoskeletal problems and their related consequences, thereby supporting the use of multifaceted interventions. Hence, as recommended by Carneiro et al [52], the organizational aspects of the working environment have to be considered in order to address WMSDs and their direct consequences on nurses' health and productivity.

Multiple approaches are required to effect changes in practices and encourage a safety culture, such as ongoing training, skills, workflow processes, supervision, and communication between health care professionals about the risks [43,44]. This diverse and multifaceted approach is also crucial considering that long-term adherence to interventions is necessary for sustainable positive impacts on WMSDs, especially their secondary outcomes. For instance, although nurses enduring from WMSDs

XSL•FO
RenderX

can experience significant relief following a short-term ergonomic intervention, a prolonged duration is required for the effects on sick leaves and performance [45]. Hence, the long-term beneficial effects of adhering to these multicomponent interventions need to be analyzed and relentless efforts to effect the necessary changes are equally important [41].

Quality assessment results underscored various areas to be considered when interpreting the systematic review findings. For instance, moderate and high risk of bias was observed in 5 and 8 studies, respectively. Thus, only 6 papers were of high quality. Methodological issues identified in the studies mainly entailed limited information on participants' baseline characteristics, potential confounders that may affect the intervention's effect, not addressing missing data, incoherent onset of intervention and follow-up, and poor adherence to the intervention regimen. These findings highlight the need for well-designed and high-quality RCTs.

The strength of the systematic review stems from being among the first attempts to gather current evidence on ergonomic interventions for nursing professionals and their effects on WMSDs, sickness absence, and work performance. In addition, a comprehensive and systematic approach was used, thereby assisting the researcher in retrieving relevant papers from selected databases. Quality appraisal was also performed to identify the methodological gaps that need to be bridged in future research. Nevertheless, the limitations are well acknowledged. Given that only papers published between 2015 and 2023 were considered in the review, papers published before these dates might contain important findings that are relevant to the research topic. The literature search and screening processes were performed only by the principal investigator, reflecting the possibility of bias and unknown intraobserver reliability. Moreover, only papers written in English in 4 databases were considered, thus limiting relevant studies written in other languages that might have assisted in answering the review questions.

Conclusions

This systematic review synthesized evidence published between 2017 and 2024 on ergonomic interventions to reduce WMSDs and sickness absence and improve work performance among nurses. The most common interventions entailed physical activities or exercise, educational programs, health promotional activities, and training on the usage of patient handling devices or equipment, which were usually applied either as a single or multicomponent approach. Overall, the findings suggest that multifaceted interventions are more effective in reducing the risk of WMSDs and the associated pain in nursing professionals compared with individual interventions. The strongest effect size was observed in multicomponent interventions comprising at least one or more physical activities and exercise. Although these positive results are expected to lead to improved work functioning, work performance, and reduced sick leaves, available results are convincing to support the use of either single or combined ergonomic programs in achieving these secondary outcomes.

Apart from conducting high-quality RCTs, addressing the identified research gaps requires the integration of organizational factors and prevention policies to enhance psychological health, job satisfaction, and work dynamics into multicomponent ergonomic interventions. Since most interventions are designed for short-term periods, efforts to ensure their long-term effects are crucial for sustainable positive impacts on WMSDs, sickness absence, and work performance. For instance, introducing curricula content on the management of WMSDs and improving work performance, sickness absence and psychological health in nursing education may assist nurses in adopting the measures discussed in this review in clinical practice. This review may guide health care institutions on the most appropriate ergonomic interventions to prevent or reduce WMSDs among nurses, given their significance for well-being and economic implications in terms of absenteeism, work performance, and care for clients enduring mobility problems.

Acknowledgments

This research is funded by the Universiti Malaya Research Excellence Grant (UMREG) 2023 (project number UMREG023-2023)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Descriptive characteristics of the studies included in this systematic review and analysis (n = 19).
[DOCX File, 28 KB - humanfactors_v12i1e68522_app1.docx]

Checklist
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.
[DOCX File, 275 KB - humanfactors_v12i1e68522_app2.docx]

References

1. Sousa AD, Baixinho CL, Presado MH, Henriques MA. The effect of interventions on preventing musculoskeletal injuries related to nurses work: systematic review. J Pers Med 2023 Jan 20;13(2):185. [doi: 10.3390/jpm13020185] [Medline: 36836419]

XSL•FO
RenderX

2. Summers K, Bevan S. Musculoskeletal disorders, workforce health and productivity in the united states. : Lancaster University; 2015 URL: https://www.lancaster.ac.uk/media/lancaster-university/content-assets/documents/lums/work-foundation/AccessibilitycheckedMusculoskeletaldisorders.pdf [accessed 2024-03-20]
3. El-Tallawy SN, Nalamasu R, Salem GI, et al. Management of musculoskeletal pain: an update with emphasis on chronic musculoskeletal pain. Pain Ther 2021 Jun;10(1):181-209. [doi: 10.1007/s40122-021-00235-2] [Medline: 33575952]
4. Krishnan KS, Raju G, Shawkataly O. Prevalence of work-related musculoskeletal disorders: psychological and physical risk factors. Int J Environ Res Public Health 2021 Sep 4;18(17):9361. [doi: 10.3390/ijerph18179361] [Medline: 34501950]
5. Jradi H, Alanazi H, Mohammad Y. Psychosocial and occupational factors associated with low back pain among nurses in Saudi Arabia. J Occup Health 2020 Jan;62(1):e12126. [doi: 10.1002/1348-9585.12126] [Medline: 32515887]
6. Ciezar-Andersen SD, Hayden KA, King-Shier KM. A systematic review of yoga interventions for helping health professionals and students. Complement Ther Med 2021 May;58(102704-10):102704. [doi: 10.1016/j.ctim.2021.102704] [Medline: 33652090]
7. Bos E, Krol B, van der Star L, Groothoff J. Risk factors and musculoskeletal complaints in non-specialized nurses, IC nurses, operation room nurses, and X-ray technologists. Int Arch Occup Environ Health 2007 Jan;80(3):198-206. [doi: 10.1007/s00420-006-0121-8] [Medline: 16799823]
8. Abdollahi T, Pedram Razi S, Pahlevan D, et al. Effect of an ergonomics educational program on musculoskeletal disorders in nursing staff working in the operating room: a quasi-randomized controlled clinical trial. Int J Environ Res Public Health 2020 Oct 8;17(19):7333. [doi: 10.3390/ijerph17197333] [Medline: 33049927]
9. Harcombe H, McBride D, Derrett S, Gray A. Prevalence and impact of musculoskeletal disorders in New Zealand nurses, postal workers and office workers. Aust N Z J Public Health 2009 Oct;33(5):437-441. [doi: 10.1111/j.1753-6405.2009.00425.x] [Medline: 19811479]
10. Rasmussen CDN, Hendriksen PR, Svendsen MJ, et al. Improving work for the body - a participatory ergonomic intervention aiming at reducing physical exertion and musculoskeletal pain among childcare workers (the TOY-project): study protocol for a wait-list cluster-randomized controlled trial. Trials 2018 Jul 31;19(1):411. [doi: 10.1186/s13063-018-2788-z] [Medline: 30064464]
11. Van Eerd D, D'Elia T, Ferron EM, et al. Implementation of participatory organizational change in long term care to improve safety. J Safety Res 2021 Sep;78(9-18):9-18. [doi: 10.1016/j.jsr.2021.05.002] [Medline: 34399935]
12. Smeland AH, Twycross A, Lundeberg S, et al. Educational intervention to strengthen pediatric postoperative pain management: a cluster randomized trial. Pain Manag Nurs 2022 Aug;23(4):430-442. [doi: 10.1016/j.pmn.2021.09.007] [Medline: 34836822]
13. Shojaei S, Tavafian SS, Jamshidi AR, Wagner J. A multidisciplinary workplace intervention for chronic low back pain among nursing assistants in Iran. Asian Spine J 2017 Jun;11(3):419-426. [doi: 10.4184/asj.2017.11.3.419] [Medline: 28670410]
14. Szeto GPY, Wong TKT, Law RKY, et al. The impact of a multifaceted ergonomic intervention program on promoting occupational health in community nurses. Appl Ergon 2013 May;44(3):414-422. [doi: 10.1016/j.apergo.2012.10.004] [Medline: 23153515]
15. Moazzami Z, Dehdari T, Taghdisi MH, Soltanian A. Effect of an ergonomics-based educational intervention based on transtheoretical model in adopting correct body posture among operating room nurses. Glob J Health Sci 2015 Nov 3;8(7):26-34. [doi: 10.5539/gjhs.v8n7p26] [Medline: 26925897]
16. Jalalvandi F, Ghasemi R, Mirzaei M. Effects of back exercises versus transcutaneous electric nerve stimulation on relief of pain and disability in operating room nurses with chronic non-specific LBP: a randomized clinical trial. BMC Musculoskelet Disord 2022 Mar 26;23(1):291. [doi: 10.1186/s12891-022-05227-7] [Medline: 35337314]
17. Coskun Beyan A, Dilek B, Demiral Y. The effects of multifaceted ergonomic interventions on musculoskeletal complaints in intensive care units. Int J Environ Res Public Health 2020 May 25;17(10):3719. [doi: 10.3390/ijerph17103719] [Medline: 32466147]
18. Horppu R, Väänänen A, Kausto J. Evaluation of a guidelines implementation intervention to reduce work disability and sick leaves related to chronic musculoskeletal pain: a theory-informed qualitative study in occupational health care. BMC Musculoskelet Disord 2022 Mar 22;23(1):272. [doi: 10.1186/s12891-022-05234-8] [Medline: 35317760]
19. Soler-Font M, Ramada JM, van Zon SKR, et al. Multifaceted intervention for the prevention and management of musculoskeletal pain in nursing staff: results of a cluster randomized controlled trial. PLoS ONE 2019;14(11):e0225198. [doi: 10.1371/journal.pone.0225198] [Medline: 31738798]
20. Serra C, Soler-Font M, García AM, et al. Prevention and management of musculoskeletal pain in nursing staff by a multifaceted intervention in the workplace: design of a cluster randomized controlled trial with effectiveness, process and economic evaluation (INTEVAL_Spain). BMC Public Health 2019 Mar 28;19(1):348. [doi: 10.1186/s12889-019-6683-7] [Medline: 30922285]
21. Sundstrup E, Seeberg KGV, Bengtsen E. A systematic review of workplace interventions to rehabilitate musculoskeletal disorders among employees with physical demanding work. J Occup Rehabil 2020 Dec;30(4):588-612. [doi: 10.1007/s10926-020-09879-x] [Medline: 32219688]

XSL•FO
RenderX

22. Lietz J, Ulusoy N, Nienhaus A. Prevention of musculoskeletal diseases and pain among dental professionals through ergonomic interventions: a systematic literature review. Int J Environ Res Public Health 2020 May 16;17(10):3482. [doi: 10.3390/ijerph17103482] [Medline: 32429439]
23. Sun W, Zhang H, Lv C, et al. Comparative efficacy of 12 non-drug interventions on non-specific chronic low back pain in nurses: a systematic review and network meta-analysis. BMR 2021;34(4):499-510. [doi: 10.3233/BMR-200147]
24. Vaismoradi M, Tella S, A Logan P, et al. Nurses' adherence to patient safety principles: a systematic review. Int J Environ Res Public Health 2020 Mar 19;17(6):2028. [doi: 10.3390/ijerph17062028] [Medline: 32204403]
25. Page MJ, Moher D, Bossuyt PM, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ 2021 Mar 29;372(160):n160. [doi: 10.1136/bmj.n160] [Medline: 33781993]
26. Higgins JPT, Altman DG, Gøtzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011 Oct 18;343:d5928. [doi: 10.1136/bmj.d5928] [Medline: 22008217]
27. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016 Oct 12;355:i4919. [doi: 10.1136/bmj.i4919] [Medline: 27733354]
28. Yang S, Li L, Wang L. Effectiveness of a multidimensional intervention program in improving occupational musculoskeletal disorders among intensive care unit nurses: a cluster-controlled trial with follow-up at 3 and 6 months. BMC Nurs 2021 Dec;20(1):46. [doi: 10.1186/s12912-021-00561-y]
29. Nguyen TT, Nguyen TH, Hoang DL, et al. Effectiveness of interventions to prevent musculoskeletal disorders among district hospital nurses in Vietnam. Biomed Res Int 2022;2022:1539063. [doi: 10.1155/2022/1539063] [Medline: 35309175]
30. Pereira M, Comans T, Sjøgaard G, et al. The impact of workplace ergonomics and neck-specific exercise versus ergonomics and health promotion interventions on office worker productivity: a cluster-randomized trial. Scand J Work Environ Health 2019 Jan 1;45(1):42-52. [doi: 10.5271/sjweh.3760] [Medline: 30132008]
31. Kumthornthip W. Effectiveness of back exercise and education for lower back pain prevention among nurses at a Tertiary Hospital in Bangkok, Thailand. Smj 2020 Apr 27;72(2):109-116. [doi: 10.33192/Smj.2020.15]
32. Suni JH, Kolu P, Tokola K, et al. Effectiveness and cost-effectiveness of neuromuscular exercise and back care counseling in female healthcare workers with recurrent non-specific low back pain: a blinded four-arm randomized controlled trial. BMC Public Health 2018 Dec 17;18(1):1376. [doi: 10.1186/s12889-018-6293-9] [Medline: 30558592]
33. Higuchi S, Funatsu K, Nawata K, et al. Effect of online physical therapy on workplace accident-related outcomes in nursing care worker: study protocol of a multicentre randomised controlled trial. BMJ Open 2022 Oct 31;12(10):e061804. [doi: 10.1136/bmjopen-2022-061804] [Medline: 36316077]
34. Jakobsen MD, Aust B, Kines P, et al. Participatory organizational intervention for improved use of assistive devices in patient transfer: a single-blinded cluster randomized controlled trial. Scand J Work Environ Health 2019 Mar 1;45(2):146-157. [doi: 10.5271/sjweh.3769] [Medline: 30821335]
35. Al-Qaisi SK, El Tannir A, Younan LA. An ergonomic assessment of using laterally-tilting operating room tables and friction reducing devices for patient lateral transfers. Appl Ergon 2020 Sep;87:103122. [doi: 10.1016/j.apergo.2020.103122] [Medline: 32501251]
36. Sezgin D, Esin MN. Effects of a PRECEDE-PROCEED model based ergonomic risk management programme to reduce musculoskeletal symptoms of ICU nurses. Intensive Crit Care Nurs 2018 Aug;47:89-97. [doi: 10.1016/j.iccn.2018.02.007] [Medline: 29609839]
37. Risør BW, Casper SD, Andersen LL, Sørensen J. A multi-component patient-handling intervention improves attitudes and behaviors for safe patient handling and reduces aggression experienced by nursing staff: a controlled before-after study. Appl Ergon 2017 Apr;60:74-82. [doi: 10.1016/j.apergo.2016.10.011] [Medline: 28166902]
38. Imai R, Konishi T, Mibu A, et al. Effect of pain neuroscience education and exercise on presenteeism and pain intensity in health care workers: a randomized controlled trial. J Occup Health 2021 Jan;63(1):e12277. [doi: 10.1002/1348-9585.12277] [Medline: 34587662]
39. Marshall L, Villeneuve J, Grenier S. Effectiveness of a multifactorial ergonomic intervention and exercise conditioning kinesiology program for subsequent work related musculoskeletal disorder prevention. Work 2018;61(1):81-89. [doi: 10.3233/WOR-182782] [Medline: 30223415]
40. Taulaniemi A, Kankaanpää M, Tokola K, et al. Neuromuscular exercise reduces low back pain intensity and improves physical functioning in nursing duties among female healthcare workers; secondary analysis of a randomised controlled trial. BMC Musculoskelet Disord 2019 Jul 13;20(1):328. [doi: 10.1186/s12891-019-2678-x] [Medline: 31301741]
41. Hosseini E, Sharifian R, Bashiri A, Daneshmandi H. Effect of a developed nursing stretch break application on work-related musculoskeletal complications and fatigue among nurses: an interventional study. Pain Res Manag 2022;2022:7870177. [doi: 10.1155/2022/7870177] [Medline: 36561913]
42. Macdonald W, Oakman J. Requirements for more effective prevention of work-related musculoskeletal disorders. BMC Musculoskelet Disord 2015 Oct 14;16:293. [doi: 10.1186/s12891-015-0750-8] [Medline: 26466897]
43. Gideon Asuquo E, Tighe SM, Bradshaw C. Interventions to reduce work-related musculoskeletal disorders among healthcare staff in nursing homes; An integrative literature review. Int J Nurs Stud Adv 2021 Nov;3:100033. [doi: 10.1016/j.ijnsa.2021.100033] [Medline: 38746711]

XSL•FO
RenderX

44. Albanesi B, Piredda M, Bravi M, et al. Interventions to prevent and reduce work-related musculoskeletal injuries and pain among healthcare professionals. A comprehensive systematic review of the literature. J Safety Res 2022 Sep;82:124-143. [doi: 10.1016/j.jsr.2022.05.004] [Medline: 36031239]
45. Simmons L, Jones AW, Siriwardena N, Bridle C. Interventions to reduce sickness absence among healthcare workers: a systematic review. IJES 2019 Aug 5;8(2):147-162. [doi: 10.1108/IJES-05-2018-0028]
46. Smith BE, Hendrick P, Bateman M, et al. Musculoskeletal pain and exercise-challenging existing paradigms and introducing new. Br J Sports Med 2019 Jul;53(14):907-912. [doi: 10.1136/bjsports-2017-098983] [Medline: 29925503]
47. Voorn EL, Koopman FS, Nollet F, Brehm MA. Individualized aerobic exercise in neuromuscular diseases: a pilot study on the feasibility and preliminary effectiveness to improve physical fitness. Phys Ther 2021 Mar 3;101(3):pzaa213. [doi: 10.1093/ptj/pzaa213] [Medline: 33332538]
48. Oakman J, Macdonald W, Kinsman N. Barriers to more effective prevention of work-related musculoskeletal and mental health disorders. Appl Ergon 2019 Feb;75:184-192. [doi: 10.1016/j.apergo.2018.10.007] [Medline: 30509526]
49. Garzillo EM, Monaco MGL, Corvino AR. Healthcare workers and manual patient handling: a pilot study for interdisciplinary training. Int J Environ Res Public Health 2020 Jul 10;17(14):4971. [doi: 10.3390/ijerph17144971] [Medline: 32664243]
50. Hegewald J, Berge W, Heinrich P, et al. Do technical aids for patient handling prevent musculoskeletal complaints in health care workers?—A systematic review of intervention studies. Int J Environ Res Public Health 2018 Mar 9;15(3):476. [doi: 10.3390/ijerph15030476] [Medline: 29522440]
51. Van Hoof W, O'Sullivan K, O'Keeffe M, et al. The efficacy of interventions for low back pain in nurses: a systematic review. Int J Nurs Stud 2018 Jan;77:222-231. [doi: 10.1016/j.ijnurstu.2017.10.015] [Medline: 29121556]
52. Carneiro P, Braga AC, Barroso M. Work-related musculoskeletal disorders in home care nurses: study of the main risk factors. Int J Ind Ergon 2017 Sep;61:22-28. [doi: 10.1016/j.ergon.2017.05.002]

Abbreviations

MSD: musculoskeletal disorder
OR: odds ratio
PE: participatory ergonomics
PICO: Population, Intervention, Comparison, and Outcome
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO: International Prospective Register of Systematic Reviews
WMSD: work-related musculoskeletal disorder

Edited by J Waterson; submitted 07.11.24; peer-reviewed by A Varghese, MS Sohrabi; revised version received 18.03.25; accepted 15.04.25; published 18.06.25.

Please cite as:
Krishnanmoorthy G, Rampal S, Karuthan SR, Baharudin F, Krishna R
Effectiveness of Participatory Ergonomic Interventions on Work-Related Musculoskeletal Disorders, Sick Absenteeism, and Work Performance Among Nurses: Systematic Review
JMIR Hum Factors 2025;12:e68522
URL: https://humanfactors.jmir.org/2025/1/e68522
doi:10.2196/68522

© Guganesan Krishnanmoorthy, Sanjay Rampal, Sumitra Ropini Karuthan, Faiz Baharudin, Rama Krishna. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 18.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Personalized Interactive Music Systems for Physical Activity and Exercise: Exploratory Systematic Review and Meta-Analysis

Andrew Danso[1], PhD; Tiia Kekäläinen[2,3], PhD; Friederike Koehler[1], PhD; Keegan Knittle[4], PhD; Patti Nijhuis[1], PhD; Iballa Burunat[1], PhD; Pedro Neto[1], PhD; Anastasios Mavrolampados[1], PhD; William M Randall[1], PhD; Niels Chr Hansen[1,5,6], PhD; Alessandro Ansani[1], PhD; Timo Rantalainen[4], PhD; Vinoo Alluri[7], PhD; Martin Hartmann[1], PhD; Rebecca S Schaefer[8,9,10], PhD; Johanna K Ihalainen[4,11], PhD; Rebekah Rousi[12], PhD; Kat R Agres[13], PhD; Jennifer MacRitchie[14], PhD; Petri Toiviainen[1], PhD; Suvi Saarikallio[1], PhD; Sebastien Chastin[15,16], PhD; Geoff Luck[1], PhD

[1]Department of Music, Arts and Culture Studies, Centre of Excellence in Music, Mind, Body and Brain, University of Jyväskylä, Seminaarinkatu 15, Jyväskylän yliopisto, Jyväskylä, Finland

[10]Academy of Creative and Performing Arts, Faculty of Humanities, Leiden University, Leiden, The Netherlands

[11]Finnish Institute of High Performance Sport KIHU, Jyväskylä, Finland

[12]School of Marketing and Communication, Communication Studies, University of Vaasa, Vaasa, Finland

[13]Centre for Music and Health, Yong Siew Toh Conservatory of Music, National University of Singapore, Singapore, Singapore

[14]Department of Music & Healthy Lifespan Institute, University of Sheffield, Sheffield, United Kingdom

[15]School of Health and Life Sciences, Glasgow Caledonian University, Glasgow, United Kingdom

[16]Department of Movement and Sports Sciences, Ghent University, Ghent, Belgium

[2]Laurea University of Applied Sciences, Vantaa, Finland

[3]Gerontology Research Center and Faculty of Sport and Health Sciences, University of Jyväskylä, Jyväskylä, Finland

[4]Faculty of Sport and Health Sciences, University of Jyväskylä, Jyväskylä, Finland

[5]Royal Academy of Music Aarhus, Aarhus, Denmark

[6]Cognitive Musicology and Performance Science Lab, Department of Communication and Psychology, Aalborg University, Aalborg, Denmark

[7]Cognitive Science Lab, International Institute of Information Technology, Hyderabad, India

[8]Health, Medical, and Neuropsychology Unit, Institute of Psychology, Faculty of Social and Behavioural Sciences, Leiden University, Leiden, The Netherlands

[9]Leiden Institute for Brain and Cognition, Leiden University, Leiden, The Netherlands

Corresponding Author:
Andrew Danso, PhD
Department of Music, Arts and Culture Studies, Centre of Excellence in Music, Mind, Body and Brain, University of Jyväskylä, Seminaarinkatu 15, Jyväskylän yliopisto, Jyväskylä, Finland

Abstract

Background: Personalized Interactive Music Systems (PIMSs) are emerging as promising devices for enhancing physical activity and exercise outcomes. By leveraging real-time data and adaptive technologies, PIMSs align musical features, such as tempo and genre, with users' physical activity patterns, including frequency and intensity, enhancing their overall experience.

Objective: This exploratory systematic review and meta-analysis evaluates the effectiveness of PIMSs across physical, psychophysical, and affective domains.

Methods: Searches across 9 databases identified 18 eligible studies, of which 6 (comprising 17 intervention arms) contained sufficient data for meta-analysis. Random-effects meta-analyses and meta-regression were performed to assess outcomes for physical activity levels, physical exertion, ratings of perceived exertion, and affective valence.

Results: Results showed significant improvements in physical activity levels (g=0.49, CI 0.07 to 0.91, P=.02, k=4) and affective valence (g=1.65, CI 0.35 to 2.96, P=.01, k=4), with faster music tempo identified as a significant moderator (P=.03). No significant effects were observed for ratings of perceived exertion (g=0.72, CI −0.13 to 1.58, P=.10, k=3) or physical exertion (g=0.78, CI −0.55 to 2.11, P=.25, k=5).

Conclusions: Substantial heterogeneity and limited study quality indicate the need for more robust, randomized controlled trials to establish the efficacy of PIMSs in diverse populations.

Trial Registration: PROSPERO (International Prospective Register of Systematic Review) CRD42023465941; https://www.crd.york.ac.uk/PROSPERO/view/CRD42023465941

XSL•FO
RenderX

(JMIR Hum Factors 2025;12:e70372) doi:10.2196/70372

KEYWORDS

music intervention; health promotion; exercise; affect; systematic review; meta-analysis; mobile phone

Introduction

Background

Regular physical activity and exercise are fundamental to maintaining and enhancing overall health and well-being. Despite their recognized role in preventing and managing noncommunicable diseases such as cardiovascular diseases, cancer, and diabetes, engagement in regular physical activity and exercise remains below the suboptimal level [1]. This deficiency undermines the potential for mental health benefits of physical exercise and its contributions to quality of life [2]. The World Health Organization defines physical activity broadly, encompassing all forms of bodily movement generated by skeletal muscles that require energy expenditure, including activities such as walking, sports, and dance [1]. In contrast, exercise has been defined as "physical activity that is planned, structured, repetitive, and purposive, aiming to improve or maintain one or more components of physical fitness" [3]. However, the broad spectrum of activities categorized as physical activity and exercise often presents challenges in promoting consistent engagement and uptake, including individual-level barriers such as motivation and time constraints [4]. Efforts to increase engagement in physical activity and exercise have faced significant challenges, frequently yielding inconsistent outcomes, as exemplified by interventions such as pedometer-based programs, which demonstrate variable effectiveness depending on factors including participant motivation and engagement [4,5].

Role of Music in Enhancing Physical Activity and Exercise

Music's rhythmic properties have been shown to influence perceptions, ergonomics, and physiological markers associated with physical activity and exercise [6-10]. Available evidence suggests that auditory-motor coupling facilitates predictive synchronization in physical activity and exercise settings, which can reduce perceived exertion and enhance endurance [11,12]. Additionally, when music aligns with individual preferences, such as through self-selection, it may further increase motivation, improve affective states, induce distraction, and lower perceived effort during physical activities and exercise [7,11].

The integration of music into exercise contexts can be further understood through theoretical frameworks such as the Affective-Reflective Theory (ART) and Dual-Mode Theory. ART emphasizes the importance of momentary affective responses—such as pleasure or displeasure—in shaping future exercise behaviors [13,14]. These responses, encapsulated in the construct of "affective valence," reflect the intrinsic pleasantness or unpleasantness of emotional states that fluctuate in response to internal and external stimuli. Conversely, Dual-Mode Theory posits that music's impact on affective responses is most pronounced at moderate exercise intensities, within a zone of response variability. This zone refers to the range of exercise intensity where affective responses—such as feelings of pleasure or displeasure—are particularly sensitive to individual differences (eg, fitness level and psychological state) and contextual factors (eg, music, environment, and social setting). In this range, attentional focus and physiological cues mediate affective experiences [15]. While both theories acknowledge the importance of affective responses in exercise, Dual-Mode Theory provides a more nuanced perspective by emphasizing intensity-dependent variability and its interaction with individual and contextual factors.

Extending the principles of ART and Dual-Mode Theory [16], the framework highlights how music's intrinsic properties—such as tempo, rhythm, and harmony—interact with personal and situational moderators, including exercise intensity and individual preferences, to influence affective and behavioral outcomes in exercise. Music operates through 3 primary mechanisms: regulating affective states, dissociating attention from exertional discomfort, and facilitating temporal prediction and rhythmic synchronization. These mechanisms are most effective within the zone of response variability, where affective valence dynamically influences exercise engagement [15,17]. Empirical studies consistently demonstrate that personalized music enhances energy efficiency, reduces perceived exertion, and improves adherence by fostering positive affect [17]. Such evidence positions personalized music systems as a key tool for optimizing both the immediate and long-term benefits of exercise.

Personalized Interactive Music Systems in Physical Activity and Exercise

Recently, advances in personalized music technologies have led to the development of Personalized Interactive Music Systems (PIMSs), which leverage software, sensors, and computer algorithms to deliver a dynamic, tailored music experience during physical activity and exercise [18,19]. These systems integrate with smartphones and wearable devices to monitor user movements and adjust musical features, such as tempo, style, and timbre, in real time to align with exercise routines, enhancing engagement and adherence to activity [20,21].

PIMSs have been designed for diverse contexts, targeting both intrinsic factors (eg, motivation and attentional focus) and extrinsic factors (eg, training guidance). For example, a PIMS, the moBeat system, used real-time interactive music and biophysical feedback to enhance cycling performance by increasing intrinsic motivation and maintaining pace and intensity [12]. Similarly, PIMS interventions for older adults have demonstrated benefits for physical endurance and engagement relative to conventional exercise conditions [22]. As mobile interventions incorporating personalization have been shown to be more effective at enhancing physical activity than nonpersonalized approaches [23], PIMSs hold promise for

XSL•FO
RenderX

improving physical activity adherence, reducing the ratings of perceived exertion (RPE), and fostering positive affective states during exercise by dynamically tailoring music to individual physiological, affective, and contextual needs [7,12,22].

Due to the relatively recent advancements of PIMSs, there is yet limited empirical evidence on their effectiveness across physical activity and exercise-related domains. Such information is essential for informing implementation, replication, and comparative evaluation of interventions aimed at promoting adherence to physical activity and exercise [24]. While systematic reviews and meta-analyses have explored the general effects of music on physical activity and exercise-related outcomes [6,7,25], these reviews predominantly focused on traditional music listening interventions and did not systematically evaluate the impact of personalized and interactive music systems. By specifically examining PIMSs, this review and meta-analysis contribute to understanding how tailored, interactive music interventions influence physical, psychophysical, and affective dimensions of physical activity and exercise engagement, thereby addressing a critical gap in the existing literature.

Therefore, this study combines a systematic review and exploratory meta-analysis to evaluate the effectiveness of PIMSs on physical activity and exercise-related outcomes. Specifically, this study synthesizes findings on physical activity levels, psychophysical measures (eg, RPE and physical exertion), and affective outcomes (eg, affective valence and mood states).

Our main research question is: How effective are PIMSs across physical, psychophysical, and affective outcomes during physical activity and exercise? This analysis intends to provide early insights into the specificity of PIMSs' effects and identify gaps in the literature that warrant further investigation.

Methods

Study Design

This systematic review and meta-analysis were designed based on the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) protocol [26]. The full search strategy can be found in the review registration document (CRD42023465941).

Eligibility Criteria

We included (1) studies investigating the effect of PIMSs on physical activity or exercise, including their effects on motivation, exercise intensity, adherence, or related outcomes, (2) studies including participants from diverse populations (eg, sufficiently active and not sufficiently active individuals), and (3) papers in the English language, published from January 2010 to May 2024 in peer-reviewed journals or as published proceedings (conference papers were considered due to the limited number of peer-reviewed studies).

We excluded (1) studies from nonpeer-reviewed sources, books, dissertations, and theses; (2) papers written in languages other than English; and (3) studies that were not directly related to the effect of PIMSs on physical exercise or physical activity.

Information Sources

We searched the following databases: (1) Web of Science, (2) SPORTDiscus, (3) Medline, (4) Embase, (5) ACM Digital Library databases, (6) Springer, (7) Google Scholar, (8) IEEE Xplore, and (9) Scopus. The database search was supplemented by a backward snowball search, whereby the reference list of all papers was scanned for potential sources. The snowball search continued until no new sources could be identified. The initial interrater agreement for the identification of relevant sources was k=0.83, indicating a strong level of agreement among the 2 individuals performing 2 independent snowball searches (AD and TK). Full search strings for all databases used in this review are provided in Section S3 in Multimedia Appendix 1.

Search Strategy

A literature search was performed using terminology related to the effects of PIMSs on physical activity and exercise, ("Personali*ed Interactive Music System*" OR "Music Recommendation Algorithm" OR "Music Recommendation System*" OR "Streaming" OR "MP3" OR "Digital Music") AND ("Physical Activity" OR "Exercise" OR "Recovery" OR "Recuperation" OR "Sedentary Behav*" OR "Physical Inactivity").

Selection Process and Data Collection Process

The citations of all retrieved papers were imported into Zotero (Digital Scholar), where duplicates were systematically identified and removed. Subsequently, 2 authors (AD and TK) independently screened the titles and abstracts of the studies using ASReview (Utrecht University) [27] and Rayyan [28]. Papers that could not be definitively excluded based on the title or abstract underwent full-text retrieval for further evaluation. The full-text papers were then independently assessed for inclusion by the same 2 authors (AD and TK). Disagreements at any stage were resolved through discussion, with a third author consulted to achieve consensus when necessary.

Data Extraction

The studies' information was extracted to a spreadsheet, including study characteristics, such as the type of PIMS, the study design, PIMS measurement, and the target behavior of the PIMS (Target Physical Activity or Exercise; Table 1).

XSL•FO
RenderX

Table . Characteristics of included studies.

Reference	Country	Age (years)	Sample size (N)	Population	Type of PIMS[a]	Study design	PIMS measurement	Target behavior or target physical activity or exercise	Physical activity results	*G* (95% CI) of PIMS on outcomes of interest
[29]	Canada	• 47.3 - 79.2[b]	• 34	Patients with cardiovascular disease	Personalized music audio-playlists	Randomized experimental design	Triaxial accelerometer	Adherence	Improved PA[c] volumes (P<.001)	• g=0.51 (−0.47 to 1.49) for physical activity level (RAS[d]) • g=−0.06 (−1.04, 0.92) for physical activity level (no RAS)
[30]	Spain	• N/A[e]	• N/A	N/A	Personalized music recommendation system	Proof of concept	Sensors[f]	Motivation or performance enhancement	N/A	• N/A
[31][g]	Germany	• N/A	• 1	Older adult participant	Music feedback for rehabilitation	Proof of concept	Accelerometer	Rehabilitation	N/A	• N/A
[32][g]	Taiwan	• 21.56 (SD 1.04)[h]	• 10 female and 26 male	Participants from the National Yang Ming Chiao Tung University	Exercise system for middle-distance running	Experimental	Smartphone's built-in triaxial accelerometer	Adapting music selection to the user's pace during walking	N/A	• g=−0.73 (−1.40 to −0.06) for physical exertion • g=1.63 (0.88 to 2.39) for RPE[i] • g=2.17 (1.34 to 2.99) for affective valence
[33]	Taiwan	• N/A	• N/A	N/A	Music assisted run trainer	Proof of concept	Triaxial accelerometer	Physiological, perceptual, and affective responses	N/A	• N/A

XSL•FO
RenderX

Reference	Country	Age (years)	Sample size (N)	Population	Type of PIMS[a]	Study design	PIMS measurement	Target behavior or target physical activity or exercise	Physical activity results	*G* (95% CI) of PIMS on outcomes of interest
[34][g]	Singapore	• N/A	• 60	Students	A music recommendation system	Within-subjects crossover design	Music recommendation ratings	Motivation	N/A	• N/A
[35]	N/A	• N/A	• 45	Nonathletes, nonbody builders, nonmusicians	JYMMiN sensor attached to fitness devices to provide musical feedback	Experimental design	Movement sensor[j]	Workout	N/A	• *g*=0.96 (0.35 to 1.56) for affective valence
[36][g]	N/A	• N/A	• 27	N/A	Runner's Jukebox: music tempo matching the user's pace during exercise	User testing design	Smartphone app to recognize user pace or adjust music tempo	Walking or running pace monitor[k]	N/A	• *g*=2.76 (1.63 to 3.89) for physical exertion (fixed BPM[l]) • *g*=2.29 (1.24 to 3.34) for physical exertion (pace-matched) • *g*=−0.74 (−1.61 to 0.13) for physical exertion (random)
[37][g]	Denmark	• N/A	• N/A	N/A	Music clips with dynamic BPM ranging from 110 - 170	Proof of concept	Sensors[m]	Cycling	N/A	• N/A
[38]	Switzerland	• 18 - 45	• 7 females and 8 males	Cyclists	SoundBike: musical sonification to improve spontaneous synchronization of cyclists	Experimental	Sensors	Cycling	Enhanced cyclist synchronization to external music	• N/A
[39]	Finland	• N/A	• 2						N/A	• N/A

XSL•FO
RenderX

Reference	Country	Age (years)	Sample size (N)	Population	Type of PIMS[a]	Study design	PIMS measurement	Target behavior or target physical activity or exercise	Physical activity results	*G* (95% CI) of PIMS on outcomes of interest
				Older adult participants	Processing accelerometry data to create musical sonifications of physical activity	Proof of concept	Sonification of PA data[n]	Awareness of PA		
[40][g]	Belgium	• N/A	• 33	Participants from public event	DSaT[o] algorithm for music selection and real-time adaptation	Pilot study	Triaxial accelerometer	Synchronization to the beat of music	The majority (19/33, 58%) synchronized their steps with music	• N/A
[41]	Belgium, Czech Republic	• 21.9 (SD 12.9)[h] • 20.2 (SD 0.8)[h] • 21.2 (SD 1.7)[h] • 23 (SD 3)[h]	• 82 male and 68 female • 56 male and 44 female • 12 female • 6 female and 4 male	N/A	Synchronize music with the participant's movements	Case study	Recordings of footfalls and music alignment strategies[p]	Synchronization to the beat of music	Improved entrainment	• N/A
[42][g]	N/A	• N/A	• N/A	N/A	Context-aware recommender system	Mixed methods design [q]	Automatic learning algorithm	Motivate users to complete PA	N/A	• N/A
[22]	Germany	• 70.6 (SD 3.9)[h]	• 11 females and 5 males	Nonphysically active	JYMMiN: sensor attached to fitness devices to provide musical feedback	Within-subjects design	Movement sensor[j]	Strength-endurance exercises	N/A	• *g*=0.73 (0.00 to 1.46) for physical activity level • *g*=0.20 (−0.27 to 0.67) for RPE • *g*=0.09 (−0.47 to 0.64) for affective valence
[43][g]			• 24							• N/A

XSL•FO
RenderX

Reference	Country	Age (years)	Sample size (N)	Population	Type of PIMS[a]	Study design	PIMS measurement	Target behavior or target physical activity or exercise	Physical activity results	*G* (95% CI) of PIMS on outcomes of interest
	Netherlands	• 18 - 25		Office workers	Smart cushion providing musical feedback	Within-subjects design	Movement sensor pad	Posture changes	No effect on breaking sedentary behavior	
[44]	Norway	• N/A	• 3-6[r]	Seniors with early-stage Alzheimer disease	Interactive music system	Qualitative research design	Sensor pad	Stimulate or motivate PA	N/A	• N/A
[12]	Netherlands	• 23 - 51	• 26	Philips employees	moBeat: interactive music system	Within-subject experiment	Cadence sensor, heart rate	Motivation	N/A	• *g*=0.54 (−0.22 to 1.30) for physical activity level • *g*=0.54 (−0.22 to 1.30) for physical exertion • *g*=0.43 (−0.32 to 1.19) for RPE • *g*=3.79 (2.52 to 5.06) for affective valence

[a]PIMS: Personalized Interactive Music System.

[b]Lowest lower bound: 47.3 years (from the second subgroup). Highest upper bound: 79.2 years (from the first subgroup). The estimated entire age range for all 3 groups combined would be from approximately 47.3 to 79.2 years.

[c]PA: physical activity.

[d]RAS: rhythmic auditory stimulation.

[e]N/A: not applicable.

[f]Galvanic skin response, oxygen saturation sensor, and pulse sensor.

[g]Conference papers.

[h]Mean (SD).

[i]RPE: ratings of perceived exertion.

[j]JYMMiN: the movement of the sensor-equipped fitness device is mapped to musical parameters, creating an acoustic feedback signal.

[k]SWPM: swings per minute.

[l]BPM: beats per minute.

[m]Monitor cycling pace and heart rate, influencing audio feedback (soundscape sounds) in real-time.

[n]Accelerometry data.

XSL•FO
RenderX

[o]DSaT: Dynamic Song and Tempo.

[p]The methodology involved recording footfalls and various music alignment strategies to synchronize music with participants' walking or running movements.

[q]Includes elements of a proof of concept design and an experimental design.

[r]Exact numbers are not specified, but a mention of a group size of 3 to 6 participants.

Preregistration Deviations

Where available, quantitative data suitable for meta-analysis were extracted. This was done for the preregistered outcome of physical activity level, as well as for affective valence, RPE, and physical exertion, which were not preregistered as outcomes. The decision to extract data on these additional outcomes was taken because of the close relationships between these variables and physical activity and exercise participation, their prevalence as outcomes in the included studies, and the limited number of studies reporting data on physical activity and exercise behavior. In cases where effect sizes could not be readily calculated based on the published papers, their authors (n=2) were contacted at least twice for additional data, resulting in the provision of calculations for 5 additional effect sizes.

Operationalization of Terms

This review operationalizes 4 key terms central to physical activity and exercise research. Physical activity level is defined by the quantified volume (eg, daily activity counts and weekly minutes), intensity (eg, metabolic equivalent of task [MET] and percent oxygen uptake reserve), and compliance (eg, adherence to heart rate zones or regimens) [45,46] of physical activity. Affective valence refers to the pleasure-displeasure dimension of emotional responses during or after physical activity, assessed using self-report scales such as the Feeling Scale (FS) [47] and the "good versus bad mood" subscale of the Multidimensional Mood Questionnaire (MDMQ) [48]. These measures capture subjective ratings of positivity or negativity without incorporating arousal [14,35,49]. Physical exertion encompasses physiological (eg, heart rate), biomechanical (eg, stride length), and perceptual demands, providing a comprehensive assessment of effort [50]. These constructs serve as the primary outcomes of interest in this review. The constructs are summarized in Table 2 and described further in Section S1 in Multimedia Appendix 1.

XSL•FO
RenderX

Table . Operationalization of terms.

Term	Definition	Operational metrics	References
Physical activity level	Encompasses the volume, intensity, and compliance with physical activity recommendations or exercise regimens.	• Volume: total activity counts per day via accelerometer, mean weekly minutes. Intensity: absolute intensity using metabolic equivalent of tasks, relative intensity as oxygen consumption reserve percentage. • Compliance: adherence to recommendations or regimens via changes in volume, device usage, or adherence to heart rate zones.	[45,46]
Physical exertion	Effort exerted to perform physical activity, involving physiological, biomechanical, and perceptual demands.	• Physiological: heart rate as an indicator of cardiovascular response. • Biomechanical: stride length and pace for activities such as running and walking. • Perceptual: integration of physiological and biomechanical cues to assess overall effort.	[50]
RPE[a]	Subjective numerical value reflecting perceived effort during physical activity, integrating sensory cues, and physiological sensations.	• Scale: Borg RPE scale for aerobic activities (cycling and running). • Category-Ratio Scale: Borg Category-Ratio 10 Scale to measure perceived exertion or other sensations. • Responses: local sensations (muscles, skin, and joints) and central factors (cardiopulmonary system).	[7,51,52]

Affective valence	The subjective feeling of pleasure or displeasure experienced during or after physical activity. It is independent of perceived exertion and reflects emotional responses to exercise, influenced by individual, contextual, and social factors.	Affective valence is measured using self-report scales, such as:
		• Scale:
		  • Feeling Scale: bipolar scale from +5 (very good) to −5 (very bad).
		  • Positive and negative affect schedule: assesses positive and negative emotions.
		  • Multidimensional Mood Questionnaire (MDMQ): evaluates mood during exercise using subscales for “good versus bad mood,” “calmness versus agitation,” and “alertness versus tiredness.” Only the “good versus bad mood” subscale aligns with the pleasure-displeasure dimension of affect.
		• Context: Measurement occurs before, during or immediately after exercise.
			[14,47-49]

[a]RPE: ratings of perceived exertion.

XSL•FO
RenderX

Study Risk of Bias Assessment

The quality of the studies was assessed by 2 authors (AD and TK) using the JBI's (Joanna Briggs Institute) critical appraisal checklist, including tools for quasi-experimental appraisal, qualitative research appraisal, and the revised checklist for randomized controlled trials [53].

Data Synthesis and Analysis Methods

We conducted a narrative synthesis, categorizing studies into two groups based on design: (1) experimental studies, including randomized, quasi-experimental, pilot, and within-subject designs, and (2) proof-of-concept and user-testing studies. This classification enabled the identification of trends within and across these categories.

For experimental studies, we examined outcomes related to physical activity levels, physical exertion, RPE, and affective valence. Proof-of-concept and user-testing studies were analyzed for their focus on PIMS design features and effectiveness, including synchronization, user engagement, and personalization.

Our synthesis followed the methodological framework of [54], facilitating systematic comparisons across study groups. Trends and variations in PIMSs' outcomes were interpreted through subgroup analyses, accounting for methodological rigor and study design. We also considered sample characteristics, including demographic variability (eg, age, fitness level, and population type) and sample size heterogeneity (ranging from n=10 to n=150). Limitations arising from study heterogeneity were explicitly addressed to provide transparency regarding factors affecting generalizability.

Hedges *g* effect sizes and SEs were calculated using the tool by Wilson [55]. Meta-analytic models were conducted in *R* (version 4.5.1) using the *metafor* package [56], applying a random-effects model with the DerSimonian-Laird estimator for physical activity level, physical exertion, RPE, and affective valence. These outcomes were selected based on the preregistration criterion: "meta-analyses will be performed when at least three studies provide data sufficient for effect size calculation." For inclusion in the meta-analysis, physical activity outcomes analyzed included behaviors such as walking, running, weight training, cycling, housework, and gardening, while studies focusing on nonphysical activity outcomes (eg, subjective feasibility of PIMSs) were excluded. Six studies (comprising 17 intervention arms) met this criterion, while outcomes with insufficient data were excluded.

Heterogeneity was assessed using the I^2 statistic (relative proportion of variability attributable to heterogeneity), τ^2 statistic (absolute variance), and Cochran *Q* statistic (a formal test of homogeneity). To address the dispersion of effects across studies, the prediction interval was calculated, as it provides insights into the range of effects expected in future comparable studies, beyond the mean effect size [57]. A sensitivity analysis was conducted to evaluate publication bias by examining the relationship between SEs and effect size estimates. Following the studies by Sterne and Egger [58] and Sterne and Harbord [59], funnel plots were produced to assess asymmetry, while forest plots were used to summarize the data. Data and syntax files for these analyses are available in Multimedia Appendix 2.

An exploratory meta-regression analysis was conducted to investigate potential moderators contributing to variability in the effectiveness of PIMSs on physical activity level, affect, RPE, and physical exertion. Candidate moderators were selected based on their theoretical relevance to physical activity and exercise research: study size, participant age, exercise intensity, and music tempo. Music tempo was categorized into tempo ranges to standardize data across studies with differing methodologies, reflecting its established influence on motivational and psychophysical responses [60]. Exercise intensity was classified using MET guidelines to enhance comparability [61]. Participant age and study size were included to address population-level and methodological variability, respectively. Due to the small number of studies included in the meta-analyses, the meta-regression encompassed all outcomes of interest, with a focus on generating hypotheses for future research.

Specifically, within the meta-regression analysis, music tempo was categorized into 3 distinct groups based on beats per minute (BPM): slow (60 - 90 BPM), coded as 1; medium (91 - 130 BPM), coded as 2; and fast (131+ BPM), coded as 3. When studies reported variable tempos, the average BPM or the dominant tempo range was used for classification. Exercise intensity was categorized using a 3-level scale aligned with MET guidelines [61]: low (<3 METs), coded as 1; moderate (3 - 6.9 METs), coded as 2; and high (≥7 METs), coded as 3. For studies that did not explicitly report METs, intensity was inferred from descriptions of the exercise type or target heart rate zones. Participant age was handled as follows: for studies reporting mean age directly, the provided value was used. In studies reporting age ranges, the midpoint of the range was used as an estimate. If group-specific mean ages were available, a weighted average was calculated based on group sizes to derive an overall mean age for the study.

Results

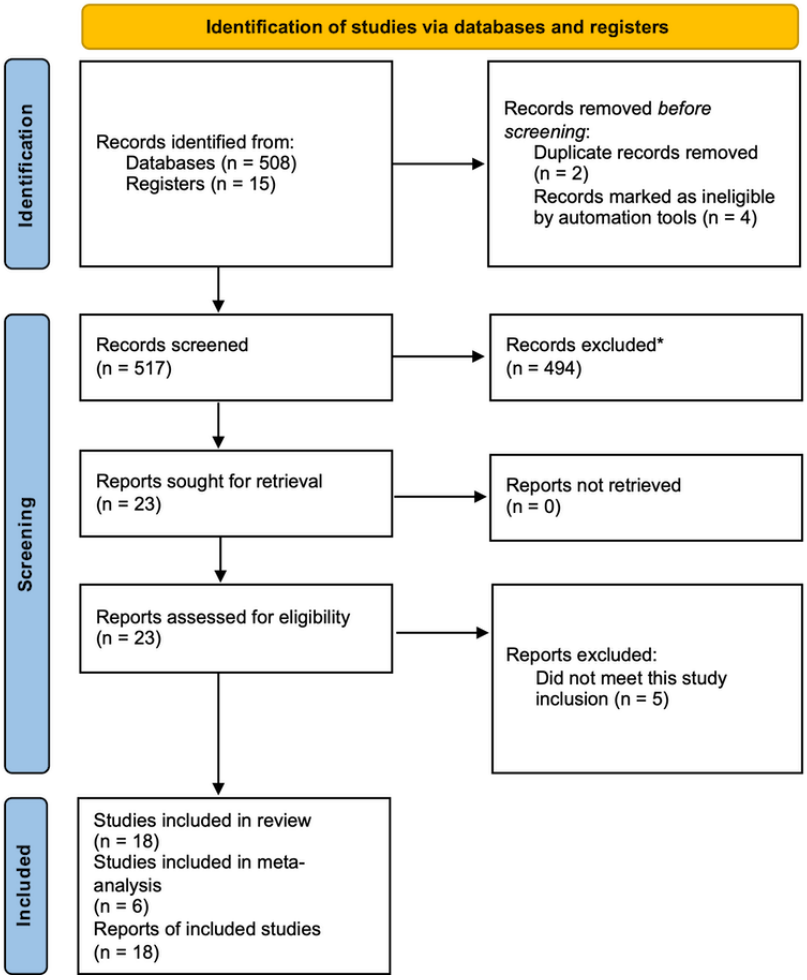
Study Selection

All records were excluded by ASReview [27,28].

A total of 523 papers were identified through the initial strategic search using the specified keywords. During the screening process, 3 papers were excluded as duplicates, while 4 additional papers were identified as ineligible based on the inclusion criteria. After screening titles and abstracts, 494 papers were excluded for not meeting the inclusion criteria. Subsequently, 23 full-text papers were assessed for eligibility. Of these, 5 papers were excluded [21,62-65] because they did not evaluate the desired effect or outcome. In total, 18 papers were eligible to be included in this review study (Figure 1).

XSL•FO
RenderX

Figure 1. PRISMA information flow describing the screening process. PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses.



Study Characteristics

The study characteristics (Table 1) encompass a diverse range of studies conducted across various countries, including Canada, Spain, Germany, Taiwan, Singapore, Denmark, Finland, Belgium, Switzerland, the Czech Republic, the Netherlands, Norway, and locations not specified. These studies, conducted between 2010 and 2024, provide a broad age range among participants, with some studies focusing on specific groups such as older adults, patients with cardiovascular disease, students, nonathletes, and office workers. The PIMSs used in these studies vary in their design and objectives, ranging from personalized music audio playlists [30,34,42] to interactive music systems linked to fitness devices [22,35]. These systems are used in different settings and for various purposes, ranging from synchronizing movement during physical activity and exercise to enhancing the experience of physical activity and exercise.

Reported Outcome Measures

A variety of outcome measures were reported across studies to explore the effects of PIMSs on physical activity and exercise-related behaviors. The outcome measures included assessments of physical activity levels, such as accelerometer-based metrics and adherence to specific heart rate zones, as well as psychological and perceptual outcomes such as mood (measured through tools such as the MDMQ and FS) and intrinsic motivation (measured via the Intrinsic Motivation Inventory, IMI). The RPE was frequently captured using the Borg Category-Ratio 10 Scale [51,52]. Table 3 presents this information. Further information on these outcome measurements can be found in Section S2 in Multimedia Appendix 1.

Studies used diverse technologies and protocols to assess PIMSs' effects on physical activity and exercise behaviors. Reported technologies included accelerometers, heart rate monitors, and systems such as JYMMiN, which integrate real-time musical feedback with gym equipment. Analytical methods, such as ANOVA and multivariate analysis of variance, were used to evaluate outcomes, with specific systems adapting music based on cadence, heart rate, and intensity. Table 4 presents a detailed overview of these technologies and protocols.

XSL•FO
RenderX

Table . Outcome measures as reported in the studies.

Outcome measure and measurement method		Study
Physical activity level		
	Mean weekly minutes of physical activity measured using a triaxial accelerometer.	[29]
	Duration of exercise until exhaustion, timed with a stopwatch.	[22]
	Compliance with exercise regime by monitoring adherence to target heart rate zones during cycling.	[12]
Affective valence		
	Feeling Scale based on Russell circumplex model of affect.	[32]
	Multidimensional Mood Questionnaire, evaluating “good versus bad” mood dimensions during acute physical exercise.	[22,35]
	Interest or enjoyment subscale of the Intrinsic Motivation Inventory for intrinsic motivation.	[12,66,67]
Ratings of perceived exertion	Borg Category-Ratio 10 Scale, with ratings of perceived exertion collected at specific time intervals during exercise.	[12,22,32]
Physical exertion		
	Heart rate measured using a Polar Verity Sense (Polar Electro) device based on photoplethysmography.	[32]
	Pace measured via swings per minute using smartphone accelerometer data.	[36]
	Heart rate measured using a Polar T61 (Polar Electro) heart rate belt.	[12]

XSL•FO
RenderX

Table . Technologies and analysis protocols of PIMSs[a].

Type of PIMSs	PIMS description	Data analysis protocol	Reference
Accelerometer	Music synchronization with step cadence	N/A[b]	[36]
Heart rate monitor	Music tempo adjustments based on physiological data	ANOVA and MANOVA[c]	[32]
JYMMiN	Music feedback system	ANOVA, MANOVA, and Wilcoxon Signed-Ranks Test	[22,35]
Magnet and heart rate sensors	Magnet sensors detect the ratings of perceived exertion, paired with heart rate, to optimize cycling rhythms	N/A	[37]
moBeat	Music feedback system	ANOVA	[12]
Musical sonification systems	Converts movement data into sound to enhance engagement and differentiate activity patterns	One proportion z-test	[39]
Musical sonification (custom pedals)	Pedals with load sensors and a microcontroller adjust musical feedback	ANOVA, Friedman test, and pairwise comparisons	[38]
Personal activity monitor	N/A	Generalized linear modeling	[29]
Three-axis accelerometer (smartphone)	Adjusts music tempo to synchronize with the user's pace using swings per minute	N/A	[36]
Triaxial accelerometer	Uses accelerometer and heart rate to adjust music tempo for maintaining the target heart rate during cardio training	N/A	[33]

[a]PIMS: Personalized Interactive Music System.

[b]N/A: not applicable.

[c]MANOVA: multivariate analysis of variance.

Risk of Bias in Studies

Following the assessment of the study quality using the JBI critical appraisal checklist tools, the nine criteria were adapted to the 5 risk-of-bias domains found in the R package for risk-of-bias assessments (robvis) in the study by McGuinness and Higgins [68]. This assessment tool tests the risk of bias resulting from the randomization process (domain [D1]), deviations from intended intervention (D2), missing outcome data (D3), measurement of the outcome (D4), and selection of the reported result (D5). Each domain is assessed with a judgment scale indicating a high risk of bias (red cross), some concerns (yellow circle), low risk of bias (green plus), and no information (blue question mark; cf Figure 2).

We included all 18 studies in the review regardless of their overall risk of bias rating (see Figure 2, column "overall"). The overall risk of bias rating for each study was assigned conservatively, reflecting the highest risk level present across any of the 5 domains (D1-D5). For example, if 1 domain was judged to have a high risk of bias, the overall rating for that study was classified as high risk. Of the 18 studies, 1 randomized experimental design study [29] was rated for low risk of bias. Seven studies received a moderate (some concerns) rating of risk of bias, and 10 were rated for a high risk of bias.

XSL•FO
RenderX

Figure 2. Evaluation of risk of bias in the included studies, categorized across 5 domains from D1 to D5 (cf [68]). An overall bias risk assessment for each study is also provided, conservatively summarizing the findings across all 5 domains [12,22,29-44]. D: domain.

XSL•FO
RenderX

PIMSs Used in Experimental Studies

PIMSs were explored in experimental studies for their influence on physical, psychophysical, and affective exercise–related outcomes. Several studies focused on synchronization and auditory-motor coupling. Moens et al [41] examined beat synchronization using the D-Jogger adaptive music player. They found that initiating music in phase synchrony significantly enhanced consistent sensorimotor patterns, while strategies relying on tempo adjustments alone were less effective. Maes et al [38] provided detailed analyses of synchronization strength using SoundBike (Ghent University), where musical sonification significantly increased pedal cadence synchronization with external music. Similarly, Jun et al [36] found significant increases in step frequency (swings per minute) when music tempo aligned with user pace, enhancing consistency and efficiency of the activity.

Rehfeld et al [22] and Fritz et al [35] reported on the JYMMiN system's role in improving mood and exercise duration. Notably, Fritz et al [35] observed mood enhancements in younger adults, while [22] noted prolonged exercise durations in older adult participants despite no significant mood changes. This may be potentially due to age-related differences in energy pacing. Rosseland [44] explored a tempo-responsive system for Alzheimer patients, observing improved synchronization and engagement. Additionally, van der Vlist et al [12] reported the moBeat system maintained exercise compliance while enhancing intrinsic motivation and attentional dissociation from discomfort. Sample sizes varied (N=10 to N=150), with participants aged 18 - 79 years across diverse demographics. Detailed descriptions of PIMSs used in these studies can be found in Table S1 in Multimedia Appendix 3.

PIMSs Used in Proof of Concept and User Testing Studies

Proof-of-concept and user-testing studies used PIMSs to adapt music or audio feedback based on real-time physical activity and exercise-related data (eg, heart rate, oxygen saturation, and galvanic skin response), with a focus on music recommendation systems and synchronization features. Álvarez et al [30] tested DJ-Running (University of Zaragoza), which integrates environmental (GPS) and galvanic skin response data to provide personalized music recommendations using algorithms such as artificial neural networks. Ospina-Bohórquez et al [42] developed a context-aware recommender system using smartphone sensors to adjust music based on exercise intensity, providing evidence for preliminary efficacy in low-concentration activities (eg, low-to-moderate intensity activities that require minimal concentration, such as walking).

Two synchronization-based systems were included: [33] music-assisted run trainer, which adjusts music tempo to heart rate or step frequency; and [31] music feedback exercise system, which synchronizes music with movement intensity through advanced audio processing. For example, as exercise intensity increases, additional layers of musical elements such as rhythm guitar, bass, or drums are progressively added to the audio track. Mendoza et al [39] introduced musical sonification, converting movement data into music for users to identify different physical activity patterns. Maculewicz and Serafin [37] examined ecological soundscapes to influence cycling behavior. Soundscapes were, for example, dynamically altered based on users' cycling speed and heart rate.

Moens et al [40] reported optimal movement entrainment at ~120 BPM using D-Jogger but noted disruptions during song transitions. The reinforcement learning–based system by Fang et al [34] found improved user satisfaction and fewer track rejections, while Rosseland [44] found tempo-responsive music systems beneficial for older adults with Alzheimer disease. Details on these systems can be found in Table S2 in Multimedia Appendix 3.

Meta-Analyses

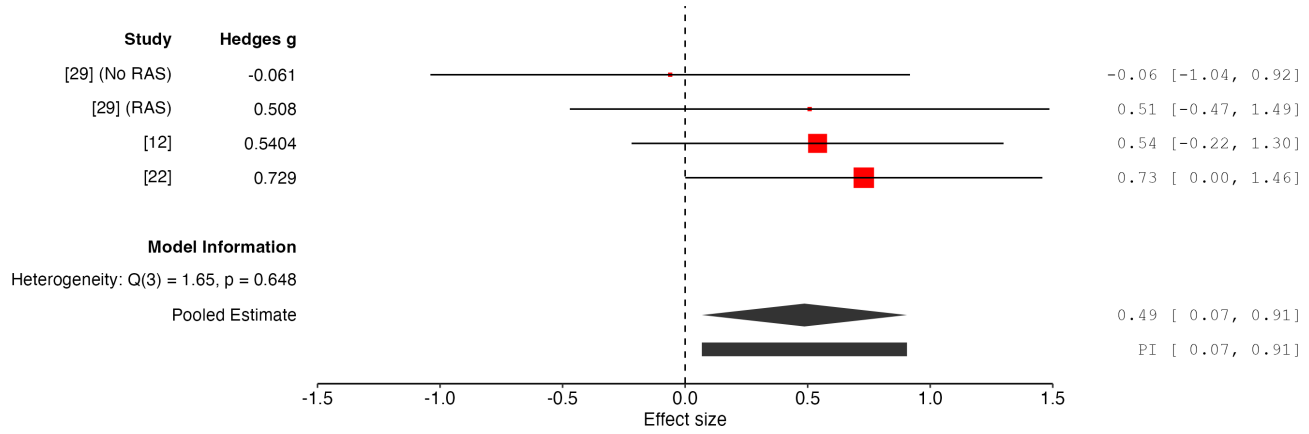
A single overall meta-analysis of the studies was not achievable due to heterogeneity across datasets and outcomes [69]. Instead, the outcomes were reported separately based on their focus. The reported outcomes distinguished between (1) physical activity levels, (2) physical exertion, (3) RPE, and (4) affective valence.

Results for Physical Activity Level

The overall effect size is 0.49 with a 95% CI of 0.07 to 0.91, and a *P* value of .02 (k=4, n=76). This indicates that the results are statistically significant, supporting the effectiveness of PIMSs in improving outcomes relating to physical activity level (Figure 3). The random-effects model indicates low heterogeneity (Q=1.65, P=.65, I^2=0%, τ^2=0.00) between the studies, suggesting it to be negligible. The calculated 95% prediction interval for the true effect size is 0.07 to 0.91, indicating that while the average effect is positive, the range of potential true effects across future studies could include larger positive outcomes.

XSL•FO
RenderX

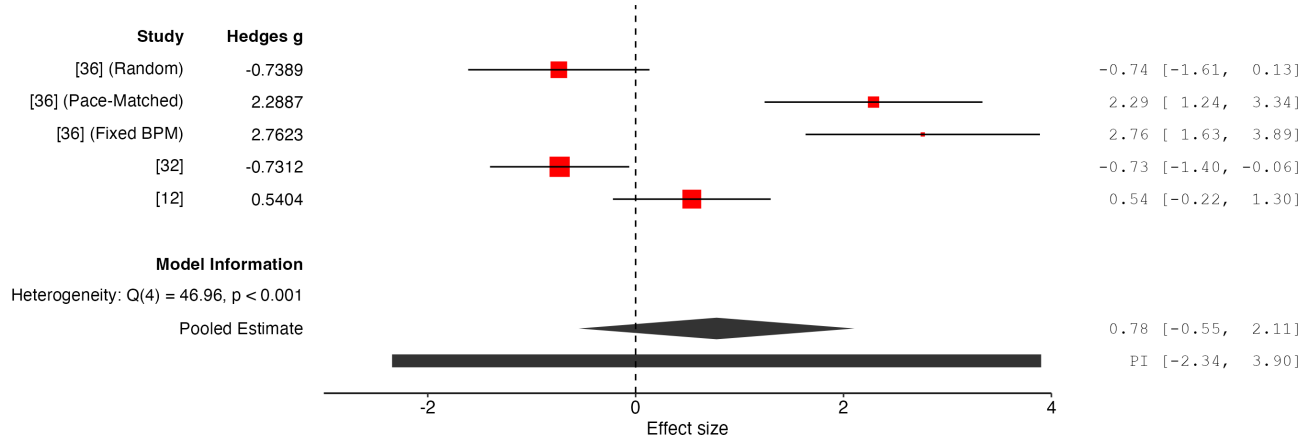
Figure 3. Forest plot of effect sizes for physical activity level outcomes associated with PIMSs [12,22,29]. PIMS: Personalized Interactive Music System; RAS: rhythmic auditory stimulation.



Results for Physical Exertion

The overall effect size is 0.78 with a 95% CI of −0.55 to 2.11, and a *P* value of .25 (*k*=5, n=142), indicating that the results are not statistically significant and do not support the effectiveness of PIMSs in improving physical exertion outcomes (Figure 4). The random-effects model indicates high heterogeneity (Q=46.96, P≤.001, I^2=91%, τ^2=2.08) between the studies. The calculated 95% prediction interval for the true effect size is −2.34 to 3.90, indicating the potential for considerable variation in the effects of PIMSs on physical exertion across future studies.

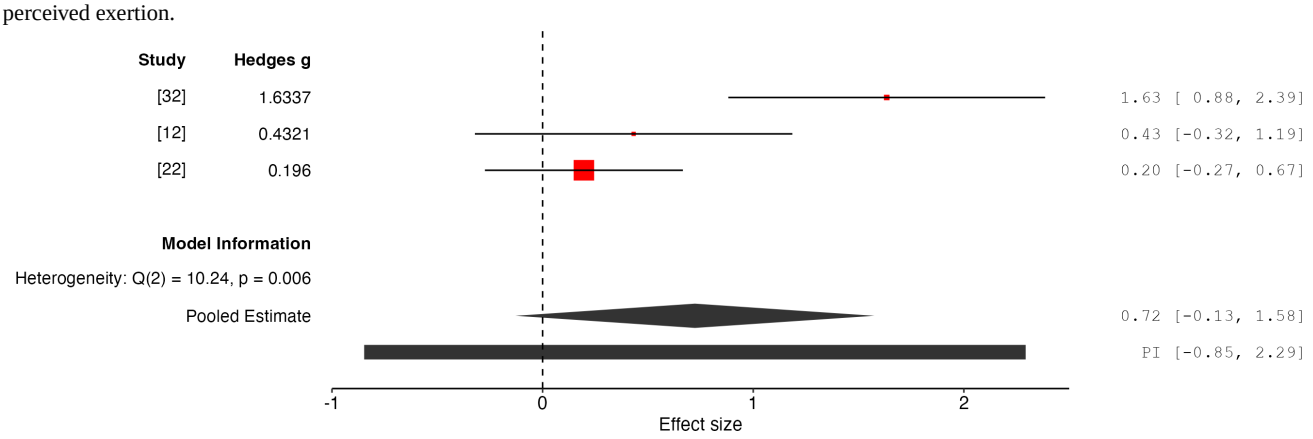
Figure 4. Forest plot of effect sizes for physical exertion outcomes associated with PIMSs [12,32,36]. BPM: beats per minute; PIMS: Personalized Interactive Music System.



Results for RPE

The overall effect size is 0.72 with a 95% CI of −0.13 to 1.58, and a *P* value of .10 (*k*=3, n=77), indicating that the results are not statistically significant and do not conclusively support the effectiveness of PIMSs in improving RPE outcomes (Figure 5). The random-effects model indicates substantial heterogeneity (Q=10.24, P=.01, I^2=80%, τ^2=0.45) between the studies. The calculated 95% prediction interval for the true effect size is −0.85 to 2.29, reflecting the significant variability in potential outcomes across future studies.

Figure 5. Forest plot of effect sizes for RPE outcomes associated with PIMSs [12,22,32]. PIMS: Personalized Interactive Music System; RPE: ratings of perceived exertion.

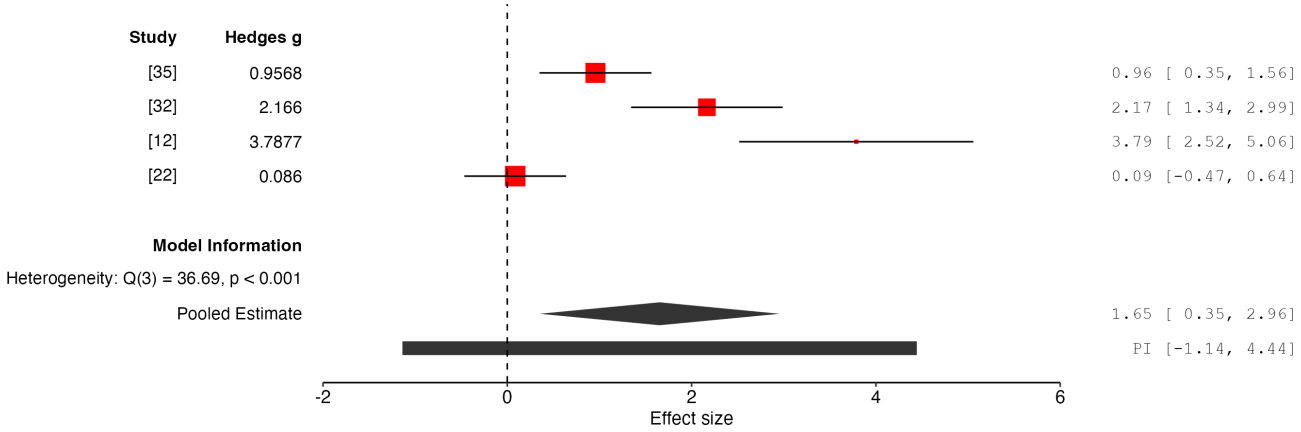


XSL•FO
RenderX

Results for Affective Valence

The overall effect size is 1.65 with a 95% CI of 0.35 to 2.96, and a *P* value of .01 (*k*=4, n=122), indicating that the results are statistically significant and thus consistent with the effectiveness of PIMSs in improving affective valence outcomes (Figure 6). The random-effects model indicates substantial heterogeneity (Q=36.69, P<.001, I^2=92%, τ^2=1.59) between the studies. The calculated 95% prediction interval for the true effect size is −1.14 to 4.44, highlighting significant variability in potential outcomes across future studies.

Figure 6. Forest plot of effect sizes for affective valence outcomes associated with PIMSs [12,22,32,35]. PIMS: Personalized Interactive Music System.



Meta-Regression Analysis

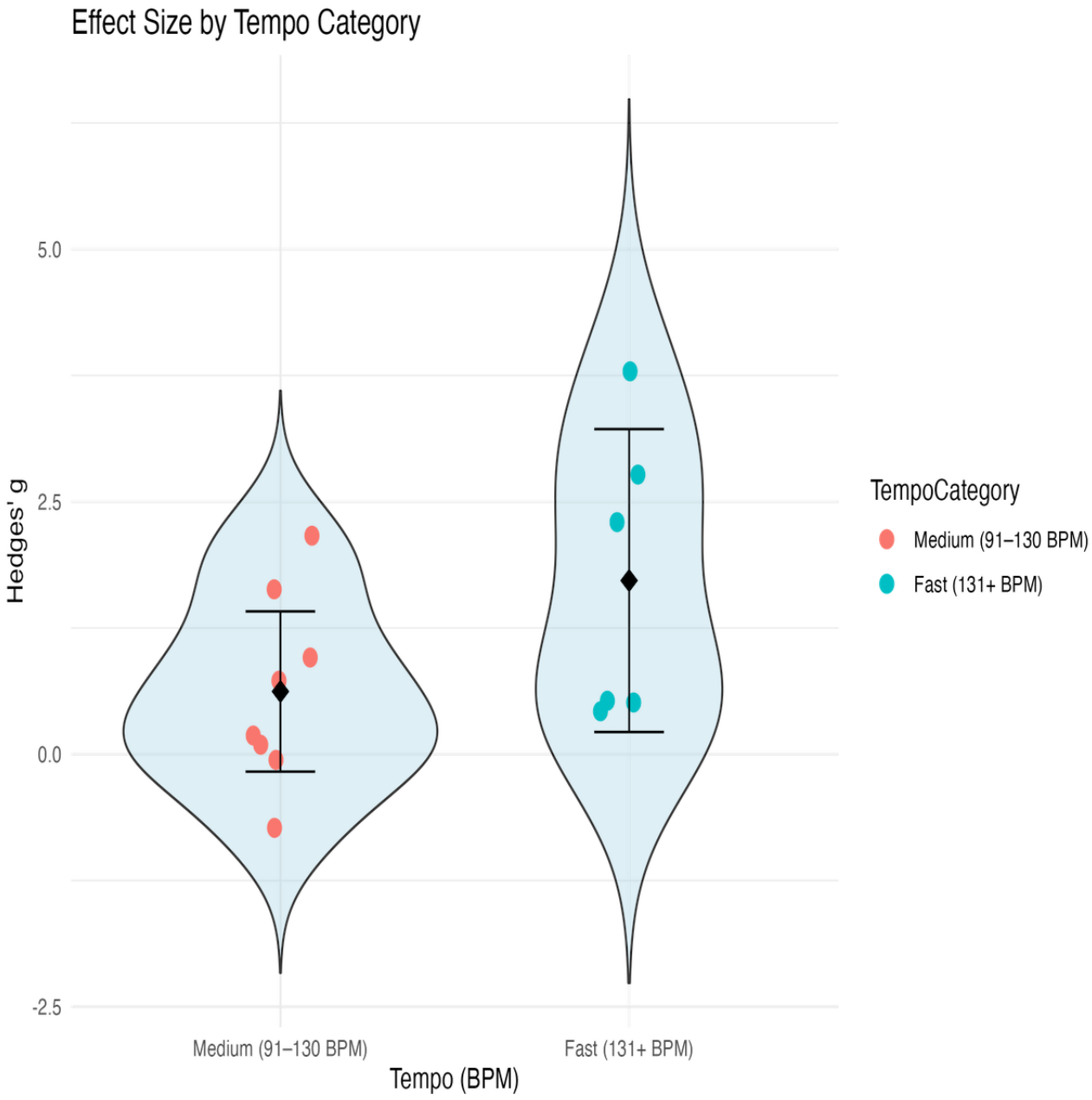
Heterogeneity was identified in the meta-analyses, prompting the use of meta-regression analysis to explore potential moderators of effect sizes. Music tempi showed a statistically significant positive association with effect sizes (β=.62, SE=0.29, *P*=.031), suggesting that faster tempi may have a significant effect across the outcomes of interest. None of the other predictors, including participant age, exercise intensity, or sample size, demonstrated a significant effect on effect sizes (Table 5 and Figure 7). The overall meta-regression model was not statistically significant, $Q_M(4)$=7.03, P=.135, and a substantial portion of heterogeneity remained unexplained, $Q_E(11)$=76.78, P<.001, I^2=85.67%, τ^2=0.92. This indicates that other, unexplored factors likely contribute to the variability in outcomes. Given the inclusion of all outcomes of interest in this analysis, the potential for residual variability and unaccounted-for heterogeneity is high.

Table . A summary of the meta-regression analysis.

Predictor	Estimate	SE	z	*P* value	95% CI
Intercept	1.194	2.965	0.402	.69	−4.619 to 7.006
Age	−0.035	0.040	−0.878	.38	−0.114 to 0.043
Music tempo	0.617	0.286	2.155	.031	0.056 to 1.178
Exercise intensity	0.212	1.262	0.168	.87	−2.262 to 2.686
Sample size	−0.025	0.062	−0.402	.69	−0.147 to 0.097

XSL•FO
RenderX

Figure 7. Distribution of study-level effect sizes across music tempo categories. Violin plots illustrate the density and spread of Hedges *g* values for medium (91 - 130 BPM) and fast (131+ BPM) tempo groups. Dots represent individual study estimates; diamonds and error bars indicate group means and 95% CIs, respectively. BPM: beats per minute.

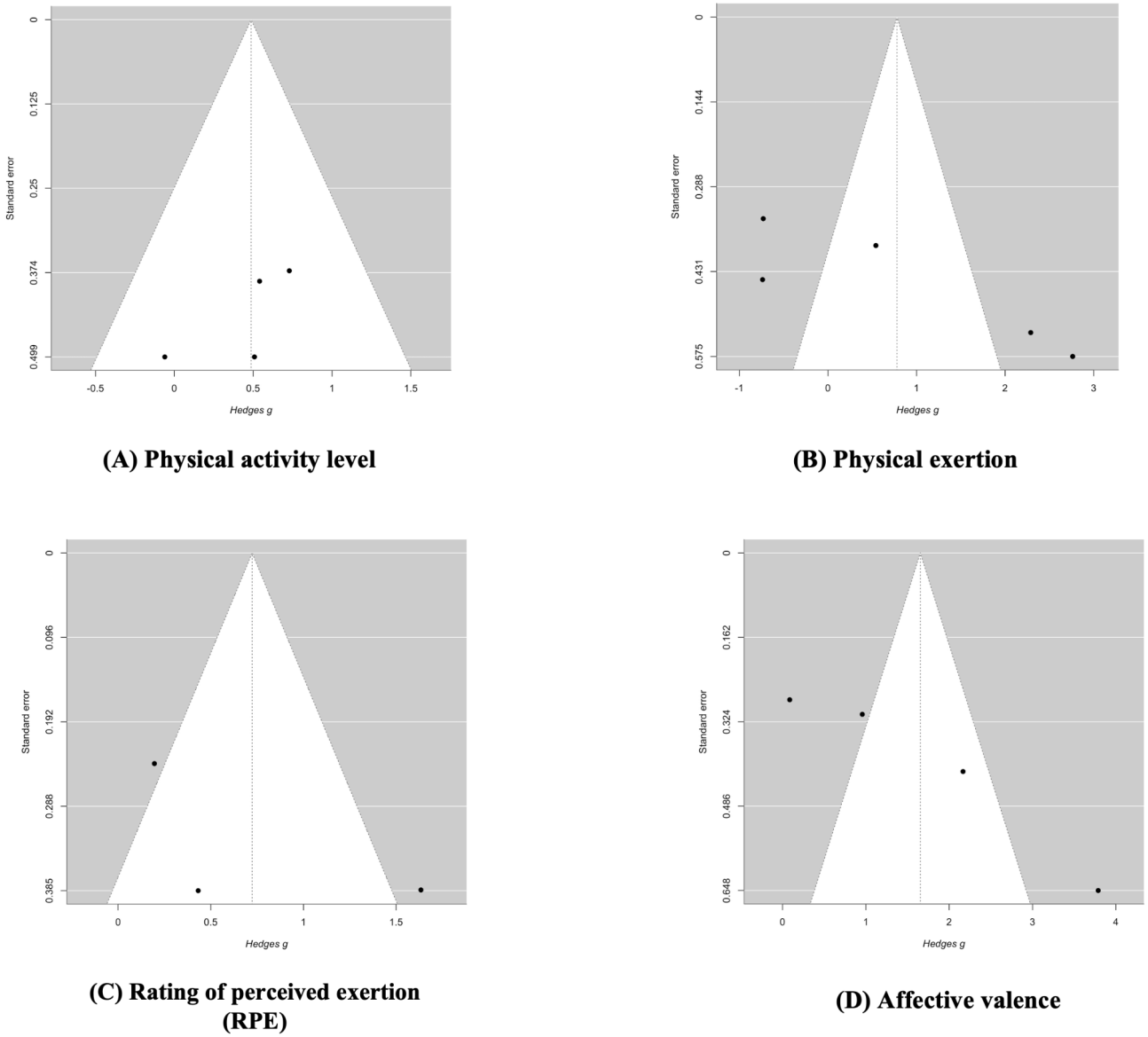


Publication Bias Analysis (Egger Test)

Egger test [70] indicated nonsignificant asymmetry for physical activity level (z=−0.968, P=.333), significant asymmetry for physical exertion (z=2.927, P=.003), nonsignificant asymmetry for RPE (z=0.832, P=.405), and significant asymmetry for affective valence (z=4.961, P<.001; Figure 8). Due to potential publication bias, the summary effect sizes for physical exertion and affective valence outcomes may thus be slightly inflated.

XSL•FO
RenderX

Figure 8. Funnel plots for (A) physical activity level, (B) physical exertion, (C) RPE, and (D) affective valence. RPE: ratings of perceived exertion.



Discussion

Principal Findings

This review aimed to systematically evaluate the effectiveness of PIMSs across physical activity levels, physiological outcomes (eg, heart rate and step frequency), psychophysical outcomes (eg, the RPE), and affective valence in relation to physical activity and exercise behaviors. A central focus was an exploratory meta-analysis of PIMSs across these outcome domains.

The exploratory meta-analysis revealed that PIMSs demonstrate favorable effects on physical activity levels and affective valence, with effect size estimates surpassing those of general music listening [6]. However, the certainty of evidence is limited by methodological inconsistencies, a moderate to high risk of bias, and the limited number of published studies eligible for meta-analyses. Importantly, no significant effects were observed for RPE or measured physical exertion. This reflects variability in the psychophysical outcomes associated with interventions using PIMSs.

When examining the findings of individual studies separately, they offer preliminary evidence that PIMSs may improve physical, psychophysical, and affective outcomes related to physical activity and exercise. For example [22], observed longer exercise durations during sessions using JYMMiN compared to routines with passive music listening, without significant increases in perceived exertion. Similarly, Alter et al [29] reported increased weekly physical activity volumes among cardiovascular disease patients using personalized rhythmic auditory stimulation–enhanced playlists. Additionally, Ren et al [43] provided qualitative evidence suggesting that PIMSs may prompt physical activity, such as reducing sitting time in office settings.

However, interpreting these findings is challenging due to methodological limitations and variability in population characteristics. Some studies focus on clinical populations, such as cardiovascular disease patients [29], while others target healthy younger adults [33] or older adult participants [22]. Several studies lack demographic details entirely, further complicating the assessment of population-specific efficacy.

XSL•FO
RenderX

Sample sizes also vary widely, from single participants [31] to larger groups (N=36) [42].

Heterogeneity in PIMSs' Outcomes: Methodological Influences

The wide prediction intervals observed across outcome domains reflect the substantial heterogeneity in PIMSs' effects. For example, prediction intervals for physical activity levels and affective valence highlight significant variability in potential effect sizes. This suggests that while PIMSs may provide positive average effects, individual study outcomes could range from substantial benefits to negligible or even negative impacts. Similarly, the prediction intervals for RPE and physical exertion emphasize uncertainty surrounding these psychophysical outcomes, pointing to inconsistencies in measurement and intervention design.

Specifically, variations in study methodologies and control group conditions contribute significantly to this heterogeneity. Some studies used passive music or other auditory stimuli as controls, while others used no-music conditions. This negatively affects comparability. Well-powered randomized designs, such as that by Alter et al [29] produced robust findings, whereas smaller studies, such as that by Rehfeld et al [22] yielded nonsignificant results, pointing to the influence of study design and statistical power. Additionally, short intervention durations and small sample sizes [37,39] constrain the generalizability of findings. The absence of standardized metrics and protocols across studies further hinders the ability to synthesize outcomes and develop systematic guidelines for PIMS interventions. To alleviate this, future research should adopt standardized protocols and outcome measures. This could be achieved via a music selection and delivery protocol, ensuring uniformity through a predefined library of music tracks categorized by tempo and intensity, delivered via standardized systems (eg, wireless headphones at consistent volumes). Validated tools such as the Borg RPE and the FS for measuring affective valence, administered at fixed intervals, may enhance comparability.

Feasibility of PIMSs on Physical Activity Levels and Affective Outcomes

Despite methodological inconsistencies, our findings suggest that PIMSs may have a positive influence on physical activity levels. Studies in this cluster were rated as having low [29] to moderate [12,22] risk of bias, with both the studies by Alter et al [29] and Rehfeld et al [22] focusing on older adult populations. Positive effects include increased exercise duration (~66 seconds) [22] and overall weekly physical activity (~105.4 additional minutes per week on average) [29]. However, van der Vlist et al [12] found no significant impact of PIMSs on physical activity levels. The low heterogeneity in this cluster indicates consistent findings despite variations in study design and participant populations. This is promising and calls for further investigation.

Our results align with that of Clark et al [6], who noted that music listening, when combined with physical activity, enhances exercise outcomes in older adults. Both Alter et al [29] and van der Vlist et al [12] used synchronization strategies—rhythmic auditory stimulation and auditory-motor coupling, respectively—consistent with frameworks by Bood et al [8] and Clark et al [71] that link synchronized music to improved physical activity and exercise performance. However, the exploratory nature of the meta-analysis and the small number of studies limit the potential generalizability of these findings. Further research with diverse populations and robust methodologies is required to confirm whether PIMSs are effective adjuncts for increasing physical activity levels.

For affective valence, the large effect size estimate suggests PIMSs contribute to elevated affective experiences during physical activity and exercise [12,22,32,35]. However, this finding is strongly influenced by van der Vlist et al [12], whose notably high effect size estimate substantially raised the overall meta-analytic effect size estimate. In contrast, smaller effects observed in other studies [22,35] reduced the precision and generalizability of the overall meta-analytic finding. The differences in these outcomes likely reflect variations in music selection methods: researcher-selected music in the study by van der Vlist et al [12] prompted synchronization and enjoyment ("fun and enjoyment" ratings via IMI), while self-selected music [32] and device-generated feedback [22,35] influenced affective outcomes in distinct ways. In the study by van der Vlist [12], researcher-selected music facilitated synchronization, while Chen et al [32] used self-selected music based on participants' individual preferences. Rehfeld et al [22] and Fritz et al [35] used device-generated musical feedback, where participants' movements influenced the music. These differences suggest that PIMSs may enhance affective valence outcomes during physical activity and exercise through both self-selected and researcher-selected music, with evidence of positive effects for music tailored to individual preferences (aligning with prior research by Terry et al [7] and Khalfa et al [11]) as well as for standardized, researcher-selected stimuli.

Curiously, van der Vlist et al [12] reported no significant benefits for RPE, despite using auditory-motor coupling strategies. This discrepancy may find alignment with the Dual-Mode Theory, as even though music can enhance automatic synchronization and facilitate improved physical performance, it does not always mitigate RPE if reflective processes (eg, cognitive appraisal of effort) are less engaged [13]. The substantial heterogeneity within the affective valence cluster, driven by variability in musical strategies, participant demographics, and inconsistent measurement tools (eg, MDMQ, IMI, and FS), further supports ART's assertion that individual and contextual factors critically shape affective outcomes during exercise.

All studies in the affective valence cluster were deemed to have a moderate risk of bias. Furthermore, the reliance on measurement scales without strong theoretical grounding, as noted in the study by van der Vlist et al [12], suggests the need for alignment with validated frameworks such as ART. For instance, the FS used by Chen et al [32] directly measures the pleasure-displeasure dimensions central to ART, aligning with validated frameworks in physical activity and exercise contexts [72]. The FS provides a theoretically robust and context-specific assessment of affective responses, capturing the transient emotional states during exercise that ART posits are critical for

XSL•FO
RenderX

shaping future behavioral intentions. These findings tentatively indicate that these PIMSs leverage momentary affective responses to improve exercise experiences [6,7,17]. In sum, findings across the physical activity and affective valence meta-analytic clusters suggest PIMSs may support affect augmentation during physical activity, highlighting their potential to enhance both physical activity levels and affective outcomes [5,17].

PIMSs' Tempo Adjustments and Synchronization in Physical Activity and Exercise Outcomes

The identification of faster music tempi as a statistically significant moderator in the meta-regression aligns with evidence supporting the role of synchronization strength and auditory-motor coupling in enhancing exercise outcomes [8,60]. For instance, faster tempi provide consistent rhythmic cues that facilitate the alignment of motor actions with auditory stimuli. This can optimize auditory-motor coupling [8-10], which, in turn, enables predictive synchronization to reduce RPE [7]. For example, Chen et al [32] reported that real-time tempo adjustments based on heart rate significantly reduced RPE and improved affective responses. This indicates that synchronized music facilitated participants' dissociation from internal sensory signals and promoted enjoyment during exercise [7].

Limitations and Future Directions

This review presents the first systematic exploration of PIMSs exclusively within physical, psychophysical, and affective domains of physical activity and exercise. While it provides valuable insights, several limitations must be acknowledged. A significant proportion of the included studies (14 of 18) primarily assessed the feasibility of PIMSs, with few investigating direct outcomes related to physical activity or exercise. Many experimental studies were limited by short durations, small sample sizes, and insufficiently rigorous methodologies. Similarly, proof-of-concept and user-testing studies largely focused on system feasibility rather than assessing objective psychophysiological outcomes. Consequently, the high risk of bias in 10 studies underscores the overall low quality of evidence. Additionally, the small number of eligible studies precluded sensitivity analyses, which further emphasizes the preliminary nature of this review's findings.

Few studies identified physical activity as a primary outcome, often relegating it to secondary importance. Objective assessments of physical activity—such as measures of frequency, intensity, and duration—were notably absent, making it difficult to draw robust conclusions or compare results across studies. Standardizing methods for quantifying physical activity would enhance future research by enabling more meaningful cross-study comparisons.

Furthermore, the methodology used in this study was limited by substantial heterogeneity across studies. This prevents a unified meta-analysis and necessitates the reporting of separate outcomes. Variability in study designs, participant demographics, and measurement tools contributed to unexplained heterogeneity, while the small number of studies precluded sensitivity analyses. These factors, combined with the exploratory nature of the meta-analysis, point to the need for standardized methodologies and rigorous reporting in future research. Additional limitations include the possibility of publication and retrieval bias, as only English-language studies from selected databases were included. Furthermore, although screening and data extraction were independently conducted by 2 reviewers, the use of automated tools and subjective judgment may have introduced bias.

To address these limitations, future research should prioritize larger, randomized controlled trials with diverse populations and longer intervention periods. Longitudinal studies are particularly needed to evaluate the sustained impact of PIMSs on physical activity and exercise. Additionally, investigating the mechanisms underlying individual variability in PIMSs' responses could optimize these systems for different populations and exercise contexts. This highlights the need for more rigorous research to validate these effects and refine PIMSs' interventions, particularly through the development of dynamic systems that can adapt tempo in real time to suit diverse user needs and exercise contexts [41,71].

Emerging trends in PIMSs, such as music recommender systems examined by Álvarez et al [30], Fang et al [34], and Ospina-Bohórquez et al [42], highlight the potential for integration with streaming services such as Spotify (Spotify AB). These systems demonstrated promising user feedback [34] and feasibility, suggesting they could serve as a foundation for future hypothesis-driven studies. Incorporating feedback from wearable and smartphone devices offers another avenue for development, allowing PIMSs to adapt based on physical activity and exercise metrics as well as music preferences. Finally, many PIMSs are relatively low-cost interventions (eg, the devices in the study by Alter et al [29] cost approximately US $75 per patient) and could have significant cost-effectiveness implications as part of broader health policy strategies to enhance physical activity and exercise participation at the population level [5].

Conclusions

This systematic review provides exploratory evidence that PIMSs may positively impact physical activity levels and affective valence in physical activity and exercise contexts. The meta-analysis revealed moderate effect sizes for physical activity levels and significant but heterogeneously distributed effects for affective valence. However, outcomes for RPE and physical exertion were inconclusive due to high heterogeneity and limited study quality.

The findings are constrained by methodological limitations, including high risk of bias, small sample sizes, short study durations, and inconsistent measures across studies. Furthermore, the lack of theoretical frameworks for informing PIMSs' designs and the absence of standardization in quantifying physical activity outcomes limit the generalizability of these findings. PIMSs remain considerably underexplored, and further research is essential.

Overall, PIMSs provide promising potential for enhancing physical activity levels and elevated affective valence, offering engaging physical activity and exercise opportunities for the

XSL•FO
RenderX

public at large. With advancements in adaptive systems capable of real-time tempo adjustments, PIMSs may emerge as effective adjuncts for physical activity and exercise, pending rigorous validation in diverse populations.

Acknowledgments

This project has received funding from the Research Council of Finland (346210).

Authors' Contributions

AD, TK, GL, and SC conceptualized this study. AD, TK, and GL handled the methodology. AD, TK, IB, P Neto, AM, and WMR investigated. AD, TK, P Nijhuis, IB, P Neto, AM, and WMR curated the data. AD and KK worked on the formal analysis. AD, FK, P Nijhuis, RR, KRA, JM, and NCH wrote the original draft. AD, TK, FK, KK, P Nijhuis, IB, P Neto, AM, WMR, NCH, AA, TR, VA, MH, RSS, JKI, RR, KRA, JM, PT, SS, SC, and GL reviewed and edited the writing.

Conflicts of Interest

None declared.

Checklist 1
PRISMA checklist. PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses.
[DOCX File, 23 KB - humanfactors_v12i1e70372_app1.docx]

Multimedia Appendix 1
Descriptions and outcomes of the PIMSs used across studies in sections. PIMS: Personalized Interactive Music System.
[DOCX File, 29 KB - humanfactors_v12i1e70372_app2.docx]

Multimedia Appendix 2
Data and syntax for meta-analysis.
[DOCX File, 14 KB - humanfactors_v12i1e70372_app3.docx]

Multimedia Appendix 3
Descriptions and outcomes of the PIMSs used across studies in tables. PIMS: Personalized Interactive Music System.
[DOCX File, 33 KB - humanfactors_v12i1e70372_app4.docx]

References

1. Bull FC, Al-Ansari SS, Biddle S, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. Br J Sports Med 2020 Dec;54(24):1451-1462. [doi: 10.1136/bjsports-2020-102955]
2. Posadzki P, Pieper D, Bajpai R, et al. Exercise/physical activity and health outcomes: an overview of Cochrane systematic reviews. BMC Public Health 2020 Nov 16;20(1):1724. [doi: 10.1186/s12889-020-09855-3] [Medline: 33198717]
3. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep 1985;100(2):126-131. [Medline: 3920711]
4. Chase JAD. Systematic review of physical activity intervention studies after cardiac rehabilitation. J Cardiovasc Nurs 2011;26(5):351-358. [doi: 10.1097/JCN.0b013e3182049f00] [Medline: 21263342]
5. Moy ML, Teylan M, Weston NA, Gagnon DR, Garshick E. Daily step count predicts acute exacerbations in a US cohort with COPD. PLoS ONE 2013;8(4):e60400. [doi: 10.1371/journal.pone.0060400] [Medline: 23593211]
6. Clark IN, Taylor NF, Peiris CL. Music listening interventions for physical activity: a systematic review and meta-analysis of randomised controlled trials. Disabil Rehabil 2024 Jan;46(1):13-20. [doi: 10.1080/09638288.2022.2155715] [Medline: 36523132]
7. Terry PC, Karageorghis CI, Curran ML, Martin OV, Parsons-Smith RL. Effects of music in exercise and sport: a meta-analytic review. Psychol Bull 2020 Feb;146(2):91-117. [doi: 10.1037/bul0000216] [Medline: 31804098]
8. Bood RJ, Nijssen M, van der Kamp J, Roerdink M. The power of auditory-motor synchronization in sports: enhancing running performance by coupling cadence with the right beats. PLoS ONE 2013;8(8):e70758. [doi: 10.1371/journal.pone.0070758] [Medline: 23951000]
9. Karageorghis CI, Priest DL. Music in the exercise domain: a review and synthesis (part I). Int Rev Sport Exerc Psychol 2012 Mar;5(1):44-66. [doi: 10.1080/1750984X.2011.631026] [Medline: 22577472]
10. Edworthy J, Waring H. The effects of music tempo and loudness level on treadmill exercise. Ergonomics 2006 Dec 15;49(15):1597-1610. [doi: 10.1080/00140130600899104] [Medline: 17090506]

XSL•FO
RenderX

11. Khalfa S, Roy M, Rainville P, Dalla Bella S, Peretz I. Role of tempo entrainment in psychophysiological differentiation of happy and sad music? Int J Psychophysiol 2008 Apr;68(1):17-26. [doi: 10.1016/j.ijpsycho.2007.12.001] [Medline: 18234381]
12. van der Vlist B, Bartneck C, Mäueler S. moBeat: using interactive music to guide and motivate users during aerobic exercising. Appl Psychophysiol Biofeedback 2011 Jun;36(2):135-145. [doi: 10.1007/s10484-011-9149-y] [Medline: 21380562]
13. Ekkekakis P. Pleasure and displeasure from the body: perspectives from exercise. Cogn Emot 2003 Mar;17(2):213-239. [doi: 10.1080/02699930302292] [Medline: 29715726]
14. Russell JA. Core affect and the psychological construction of emotion. Psychol Rev 2003 Jan;110(1):145-172. [doi: 10.1037/0033-295x.110.1.145] [Medline: 12529060]
15. Brand R, Ekkekakis P. Affective–Reflective Theory of physical inactivity and exercise. Ger J Exerc Sport Res 2018 Mar;48(1):48-58. [doi: 10.1007/s12662-017-0477-9]
16. Karageorghis CI. The scientific application of music in exercise and sport: towards a new theoretical model. In: Sport and Exercise Psychology: New York, NY, US: Routledge/Taylor & Francis Group; 2016:276-322.
17. Karageorghis CI. Music‐Related interventions in the exercise domain. In: Handbook of Sport Psychology: Wiley; 2020:929-949. [doi: 10.1002/9781119568124]
18. Agres KR, Schaefer RS, Volk A, et al. Music, computing, and health: a roadmap for the current and future roles of music technology for health care and well-being. Music Sci 2021 Jan 1;4. [doi: 10.1177/2059204321997709]
19. Danso A. The use of technology in music-based interventions for health and education [Dissertation]. : Jyväskylä: University of Jyväskylä; 2023.
20. Marcus BH, Owen N, Forsyth LH, Cavill NA, Fridinger F. Physical activity interventions using mass media, print media, and information technology. Am J Prev Med 1998 Nov;15(4):362-378. [doi: 10.1016/s0749-3797(98)00079-8] [Medline: 9838978]
21. Wijnalda G, Pauws S, Vignoli F, Stuckenschmidt H. A personalized music system for motivation in sport performance. IEEE Pervasive Comput 2005 Jul;4(3):26-32. [doi: 10.1109/MPRV.2005.47]
22. Rehfeld K, Fritz TH, Prinz A, Schneider L, Villringer A, Witte K. Musical feedback system JYMMiN leads to enhanced physical endurance in the elderly-a feasibility study. Front Sports Act Living 2022;4:915926. [doi: 10.3389/fspor.2022.915926] [Medline: 36032261]
23. Laranjo L, Ding D, Heleno B, et al. Do smartphone applications and activity trackers increase physical activity in adults? Systematic review, meta-analysis and metaregression. Br J Sports Med 2021 Apr;55(8):422-432. [doi: 10.1136/bjsports-2020-102892] [Medline: 33355160]
24. Michie S, Richardson M, Johnston M, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
25. Ziv G, Lidor R. Music, exercise performance, and adherence in clinical populations and in the elderly: a review. J Clin Sport Psychol 2011 Mar;5(1):1-23. [doi: 10.1123/jcsp.5.1.1]
26. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
27. van de Schoot R, de Bruin J, Schram R, et al. An open source machine learning framework for efficient and transparent systematic reviews. Nat Mach Intell 2021 Feb;3(2):125-133. [doi: 10.1038/s42256-020-00287-7]
28. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 5;5(1):210. [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
29. Alter DA, O'Sullivan M, Oh PI, et al. Synchronized personalized music audio-playlists to improve adherence to physical activity among patients participating in a structured exercise program: a proof-of-principle feasibility study. Sports Med Open 2015;1(1):23. [doi: 10.1186/s40798-015-0017-9] [Medline: 26284164]
30. Álvarez P, Zarazaga-Soria FJ, Baldassarri S. Mobile music recommendations for runners based on location and emotions: the DJ-Running system. Pervasive Mob Comput 2020 Sep;67:101242. [doi: 10.1016/j.pmcj.2020.101242]
31. Carôt A, Fritz TH, Englert K. Towards a system supporting music feedback exercise in physical tele-rehabilitation. Presented at: 2023 4th International Symposium on the Internet of Sounds; Oct 26-27, 2023; Pisa, Italy p. 1-7. [doi: 10.1109/ieeeconf59510.2023.10335285]
32. Chen Y, Chen CC, Tang LC, Chieng WH. Mid-distance running exercise assistance system via iot and exercise-oriented music. Presented at: 2023 IEEE World AI IoT Congress (AIIoT); Jun 7-10, 2023; Seattle, WA, USA p. 264-272. [doi: 10.1109/AIIoT58121.2023.10174324]
33. Chen LY, Tung YJ, Jang JSR. MART-music assisted run trainer. Presented at: 15th International Society for Music Information Retrieval Conference; Oct 27-31, 2014; Taipei, Taiwan.
34. Fang J, Grunberg D, Lui S, Wang Y. Development of a music recommendation system for motivating exercise. 2018 Presented at: 2017 International Conference on Orange Technologies (ICOT); Dec 8-10, 2018; Singapore p. 83-86. [doi: 10.1109/ICOT.2017.8336094]
35. Fritz TH, Halfpaap J, Grahl S, Kirkland A, Villringer A. Musical feedback during exercise machine workout enhances mood. Front Psychol 2013;4(DEC):921. [doi: 10.3389/fpsyg.2013.00921] [Medline: 24368905]

XSL•FO
RenderX

36. Jun S, Rew J, Hwang E. Runner's jukebox: a music player for running using pace recognition and music mixing. Presented at: MMEDIA 2015: The Seventh International Conferences on Advances in Multimedia; Apr 19-24, 2015; Barcelona, Spain.
37. Maculewicz J, Serafin S. A stationary bike in augmented audio reality. 2015 Oct Presented at: REHAB '15; Oct 1-2, 2015; Lisbon Portugal p. 164-167. [doi: 10.1145/2838944.2838984]
38. Maes PJ, Lorenzoni V, Six J. The SoundBike: musical sonification strategies to enhance cyclists' spontaneous synchronization to external music. J Multimodal User Interfaces 2019 Sep;13(3):155-166. [doi: 10.1007/s12193-018-0279-x]
39. Mendoza JI, Danso A, Luck G, Rantalainen T, Palmberg L, Chastin S. Musification of accelerometry data towards raising awareness of physical activity. Presented at: Proceedings of the Conference on Sonification of Health and Environmental Data (SoniHED 2022); Oct 27-28, 2022; Stockholm, Sweden. [doi: 10.5281/ZENODO.7243875]
40. Moens B, van Noorden L, Leman M. D-jogger: syncing music with walking. 2010 Presented at: 7th Sound and music computing Conference; Jul 21-24, 2010; Barcelona p. 451-456 URL: https://biblio.ugent.be/publication/1070528 [accessed 2025-08-06]
41. Moens B, Muller C, van Noorden L, et al. Encouraging spontaneous synchronisation with D-Jogger, an adaptive music player that aligns movement and music. PLoS ONE 2014 Dec;9(12):e114234. [doi: 10.1371/journal.pone.0114234]
42. Ospina-Bohórquez A, Gil-González AB, Moreno-García MN, Luis-Reboredo A. Context-Aware Music Recommender System Based on Automatic Detection of the User's Physical Activity: Springer; 2021:142-151. [doi: 10.1007/978-3-030-53036-5_15]
43. Ren X, Lu Y, Visser V, Le PDH, van den Burg R. Interaction matters? Exploring interactive music as a reminder to break sedentary office time. 2017 Presented at: 34th International Symposium on Automation and Robotics in Construction; Jun 27-30, 2017; Taipei, Taiwan p. 1099-1106. [doi: 10.22260/ISARC2017/0151]
44. Rosseland RB. Design and evaluation of an interactive music system for exercise and physical activity with Alzheimer's patients. SoundEffects 2016;6(1):4-22. [doi: 10.7146/se.v6i1.24910]
45. Bassett DR, Troiano RP, McClain JJ, Wolff DL. Accelerometer-based physical activity: total volume per day and standardized measures. Med Sci Sports Exerc 2015 Apr;47(4):833-838. [doi: 10.1249/MSS.0000000000000468] [Medline: 25102292]
46. Kujala UM, Pietilä J, Myllymäki T, et al. Physical activity: absolute intensity versus relative-to-fitness-level volumes. Med Sci Sports Exerc 2017 Mar;49(3):474-481. [doi: 10.1249/MSS.0000000000001134] [Medline: 27875497]
47. Hardy CJ, Rejeski WJ. Not what, but how one feels: the measurement of affect during exercise. J Sport Exerc Psychol 1989 Sep;11(3):304-317. [doi: 10.1123/jsep.11.3.304]
48. Steyer R, Schwenkmezger P, Notz P, Eid M. Development of the Multidimensional Mood State Questionnaire (MDBF) Primary Data: Primärdatensatz. ZPID Leibniz Institute for Psychology; 2003. [doi: 10.5160/PSYCHDATA.SRRF91EN15]
49. Ekkekakis P, Petruzzello SJ. Analysis of the affect measurement conundrum in exercise psychology. Psychol Sport Exerc 2000 Oct;1(2):71-88. [doi: 10.1016/S1469-0292(00)00010-8]
50. Nystoriak MA, Bhatnagar A. Cardiovascular effects and benefits of exercise. Front Cardiovasc Med 2018;5:135. [doi: 10.3389/fcvm.2018.00135] [Medline: 30324108]
51. Borg G. Perceived exertion as an indicator of somatic stress. Scand J Rehabil Med 1970;2(2):92-98. [doi: 10.2340/1650197719702239298]
52. Borg GA. Psychophysical bases of perceived exertion. Med Sci Sports Exerc 1982;14(5):377-381. [Medline: 7154893]
53. Grammatopoulos T, Hunter JWS, Munn Z, Stone JC, Barker TH. Reporting quality and risk of bias in JBI systematic reviews evaluating the effectiveness of interventions: a methodological review protocol. JBI Evidence Synthesis 2023 Jan;21(3):584-591. [doi: 10.11124/JBIES-22-00317]
54. Campbell M, McKenzie JE, Sowden A, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. BMJ 2020 Jan 16;368:l6890. [doi: 10.1136/bmj.l6890] [Medline: 31948937]
55. Wilson D. Practical meta-analysis effect size calculator. Campbell Collaboration. 2023. URL: https://campbellcollaboration.org/research-resources/effect-size-calculator.html [accessed 2025-07-30]
56. Viechtbauer W. Conducting meta-analyses in R with the metafor package. J Stat Soft 2010;36(3):1-48. [doi: 10.18637/jss.v036.i03]
57. Borenstein M. How to understand and report heterogeneity in a meta-analysis: the difference between I-squared and prediction intervals. Integr Med Res 2023 Dec;12(4):101014. [doi: 10.1016/j.imr.2023.101014] [Medline: 38938910]
58. Sterne JAC, Egger M. Funnel plots for detecting bias in meta-analysis. J Clin Epidemiol 2001 Oct;54(10):1046-1055. [doi: 10.1016/S0895-4356(01)00377-8] [Medline: 11576817]
59. Sterne JAC, Harbord RM. Funnel plots in meta-analysis. Stata J 2004 Jun;4(2):127-141. [doi: 10.1177/1536867X0400400204]
60. Elliott D. Music during exercise: does tempo influence psychophysical responses? : Philica; 2007.
61. Jetté M, Sidney K, Blümchen G. Metabolic equivalents (METS) in exercise testing, exercise prescription, and evaluation of functional capacity. Clin Cardiol 1990 Aug;13(8):555-565. [doi: 10.1002/clc.4960130809] [Medline: 2204507]
62. Elliott GT, Tomlinson B. Presented at: CHI EA '06: CHI '06 Extended Abstracts on Human Factors in Computing Systems; Apr 22-27, 2006; Montréal Québec Canada p. 736-741. [doi: 10.1145/1125451.1125599]

XSL•FO
RenderX

63. Gallo D, Shreepriya S, Willamowski J. RunAhead: exploring head scanning based navigation for runners. Presented at: CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems; Apr 25-30, 2020; Honolulu HI USA p. 1-13. [doi: 10.1145/3313831.3376828]
64. Antón AM, Rodríguez BR. Runtastic PRO app: an excellent all-rounder for logging fitness. Br J Sports Med 2016 Jun;50(11):705-706. [doi: 10.1136/bjsports-2015-094940] [Medline: 26294442]
65. Nymoen K, Song S, Hafting Y, Torresen J. Funky sole music: gait recognition and adaptive mapping. Presented at: Proceedings of the International Conference on New Interfaces for Musical Expression; Jun 30 to Jul 3, 2014; Goldsmiths, University of London, UK. [doi: 10.5281/ZENODO.1178895]
66. Ryan RM, Mims V, Koestner R. Relation of reward contingency and interpersonal context to intrinsic motivation: a review and test using cognitive evaluation theory. J Pers Soc Psychol 1983 Oct;45(4):736-750. [doi: 10.1037/0022-3514.45.4.736]
67. McAuley E, Duncan T, Tammen VV. Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: a confirmatory factor analysis. Res Q Exerc Sport 1989 Mar;60(1):48-58. [doi: 10.1080/02701367.1989.10607413] [Medline: 2489825]
68. McGuinness LA, Higgins JPT. Risk-of-bias VISualization (robvis): an R package and Shiny web app for visualizing risk-of-bias assessments. Res Synth Methods 2021 Jan;12(1):55-61. [doi: 10.1002/jrsm.1411] [Medline: 32336025]
69. Cumpston M, Li T, Page MJ, et al. Updated guidance for trusted systematic reviews: a new edition of the Cochrane Handbook for Systematic Reviews of Interventions. Cochrane Database Syst Rev 2019 Oct 3;10(10):ED000142. [doi: 10.1002/14651858.ED000142] [Medline: 31643080]
70. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997 Sep 13;315(7109):629-634. [doi: 10.1136/bmj.315.7109.629] [Medline: 9310563]
71. Clark IN, Baker FA, Taylor NF. The modulating effects of music listening on health-related exercise and physical activity in adults: a systematic review and narrative synthesis. Nord J Music Ther 2016 Jan 2;25(1):76-104. [doi: 10.1080/08098131.2015.1008558]
72. Bastos V, Rodrigues F, Davis P, Teixeira DS. Assessing affective valence and activation in resistance training with the feeling scale and the felt arousal scale: a systematic review. PLOS ONE 2023;18(11):e0294529. [doi: 10.1371/journal.pone.0294529] [Medline: 37972201]

Abbreviations

ART: Affective-Reflective Theory
BPM: beats per minute
D: domain
FS: Feeling Scale
IMI: Intrinsic Motivation Inventory
JBI: Joanna Briggs Institute
MDMQ: Multidimensional Mood Questionnaire
MET: metabolic equivalent of task
PIMS: Personalized Interactive Music System
RPE: ratings of perceived exertion

Edited by A Kushniruk; submitted 20.12.24; peer-reviewed by A Carot, B Lai, C br, Y Zhang; revised version received 28.05.25; accepted 26.06.25; published 08.09.25.

Please cite as:
Danso A, Kekäläinen T, Koehler F, Knittle K, Nijhuis P, Burunat I, Neto P, Mavrolampados A, Randall WM, Hansen NC, Ansani A, Rantalainen T, Alluri V, Hartmann M, Schaefer RS, Ihalainen JK, Rousi R, Agres KR, MacRitchie J, Toiviainen P, Saarikallio S, Chastin S, Luck G
Personalized Interactive Music Systems for Physical Activity and Exercise: Exploratory Systematic Review and Meta-Analysis
JMIR Hum Factors 2025;12:e70372
URL: https://humanfactors.jmir.org/2025/1/e70372
doi:10.2196/70372

© Andrew Danso, Tiia Kekäläinen, Friederike Koehler, Keegan Knittle, Patti Nijhuis, Iballa Burunat, Pedro Neto, Anastasios Mavrolampados, William M Randall, Niels Chr Hansen, Alessandro Ansani, Timo Rantalainen, Vinoo Alluri, Martin Hartmann, Rebecca S Schaefer, Johanna K. Ihalainen, Rebekah Rousi, Kat R Agres, Jennifer MacRitchie, Petri Toiviainen, Suvi Saarikallio, Sebastien Chastin, Geoff Luck. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 8.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License

XSL•FO
RenderX

(https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Driving and Restraining Forces in the Implementation of Information Systems in the Public Sector: Scoping Review

Arja Lemmettylä[1*], MSocSc, MHSc; Ulla-Mari Kinnunen[1,2*], PhD

[1]Department of Health and Social Management, Faculty of Social Sciences and Business Studies, University of Eastern Finland, Yliopistonranta 1, Kuopio, Finland

[2]Research Centre for Nursing Science and Social and Health Management, Kuopio University Hospital, Wellbeing Services County of North Savo, Kuopio, Finland

[*]all authors contributed equally

Corresponding Author:
Arja Lemmettylä, MSocSc, MHSc
Department of Health and Social Management, Faculty of Social Sciences and Business Studies, University of Eastern Finland, Yliopistonranta 1, Kuopio, Finland

Abstract

Background: Public sector organizations increasingly adopt information systems (ISs) to improve economic efficiency, service quality and overall adaptability. These projects represent substantial financial investments, making their success critical for organizational performance and societal impact.

Objective: This scoping review aimed to identify the driving and restraining forces influencing IS implementation in public sector organizations and to explore strategies that support successful change processes.

Methods: A total of 25 peer-reviewed articles were analyzed using Lewin's change theory to categorize and interpret driving and restraining forces. In addition, the narrative emerging from previous research on IS implementation was examined to explore how previous research portrays the success of IS implementation processes.

Results: The findings highlight that IS implementation is influenced by 6 domains: organizational practices and challenges, technological factors and barriers, management practices and issues, change project factors and challenges, end user factors and concerns, as well as institutional factors and barriers. Key driving forces include leadership support, stakeholder involvement and system usability, while restraining forces encompass user resistance, technical challenges, and organizational silos.

Conclusions: Despite the challenges, IS implementation offers significant opportunities for improving public sector operations and societal outcomes. Addressing restraining forces and leveraging driving forces is essential for achieving sustainable digital transformation. This study provides actionable insights for future IS implementation in the public sector.

(JMIR Hum Factors 2025;12:e71575) doi:10.2196/71575

KEYWORDS

information systems; implementation; public sector; digital transformation

Introduction

The implementation of information systems (ISs) is one of the greatest transformation challenges that organizations face. Digitalization of services and operations is fundamentally transforming public sector organizations. It is seen as a key component of administrative reforms, and as a solution to high public expenditure, inefficiency, increased user involvement, and the need for greater transparency in public administration [1,2]. In addition, the digital transformation observed in other sectors has led citizens to expect real time, high-value services from the public sector [3]. In this study, public sector organizations are defined as government-funded and operated entities, such as health and social care institutions.

In health care, ISs play a critical role in operational processes, directly influencing health service outcomes. The term IS refer to a set of digital technologies and processes designed to collect, process, store, and disseminate information within organizations, including tools like electronic health records (EHRs) and other information technology (IT) solutions to support organizational operations. The digitalization of health care has potential to enhance care effectiveness, improve cost efficiency, and enable new service delivery methods [4,5].

Despite the advanced quality of health ISs, they also represent significant financial investments [6]. Furthermore, IS development is filled with challenges [7-10]. There are numerous ways to cause large information technology projects to fall short of expectations [11-13]. The failure rates have been persistently high [14] with estimates suggesting that up to

XSL•FO
RenderX

60% - 70% of system projects in health care fail [15,16]. Despite the best intentions in software development, delays, failures, and complete abandonment remain possible outcomes [8,17]. The success of IS implementation refers to their ability to meet organizational needs, achieve user acceptance, and deliver intended outcomes, such as operational efficiency and financial sustainability.

Successful adoption of new technology is supported by considering factors related to the organization, people, work, and technology [18,19]. IS-related changes require effective leadership [20-22] and capability to modify practices both at the individual and organizational levels to achieve success [18,23,24]. However, the importance of context and process is often overlooked [25,26]. Implementation in this context refers to the process of planning, deploying, and integrating ISs into organizational workflows, encompassing technical aspects as well as organizational factors such as stakeholder engagement and change management.

Given the clear necessity for the IS development and increasing financial constraints faced by public sector organizations, successful implementation is essential. The complexity of these projects highlights the need for deeper understanding of the forces that influence their success. Previous research has approached IS implementations from multiple perspectives, including their impact on organizations, services and operations [15,27], individual responses to technology adoption and resistance [16,28-30], and project management and governance [11,31]. However, these studies have often remained fragmented, addressing isolated aspects rather than forming a cohesive understanding of IS implementation processes in public sector organizations. A comprehensive examination from the public sector's viewpoint provides a novel contribution to the existing literature. The purpose of the study is to improve the understanding of the factors influencing IS implementation in public sector organizations and to provide insights into preparing for a successful change process.

The analysis was conducted using Lewin's change model and its force fields [32], which categorizes factors as either driving or restraining forces [33-35]. According to model, change is influenced by driving forces which push toward change and support it, and restraining forces which seek to maintain the status quo and resist change [36]. With Lewin's model, it is possible to analyze and identify forces that either propel the change forward or create obstacles that can halt the desired change altogether [34]. Using Lewin's model, this study reveals new insights into the forces influencing IS implementations, building on previous research. Both the driving and restraining forces are identified and analyzed, offering a deeper understanding of the dynamics during a system adoption in public sector organizations.

Research questions are: (1) What are the driving and restraining forces in the implementation of ISs within public sector organizations? (2) What kind of narrative emerges about the success of the change process?

Methods

Research questions were addressed through a scoping review. This method allows to summarize the breadth and depth of the research field on the topic [37]. It helps to identify relevant literature, clarify concepts, and map research activity within the subject area [38]. Through a scoping review, existing knowledge can be identified, evaluated, interpreted, and combined [39,40].

Population, concept, context method [41] was used for topic definition and search strategy formulation. The databases selected were Scopus, Web of Science, and PubMed due to the relevance of the topic and the coverage.

The search queries and results are detailed in Table 1. The database searches were subject to general inclusion criteria described in Textbox 1.

XSL•FO
RenderX

Table . Research queries and results.

Database	Research queries	Results (n=509), n
Scopus	(TITLE-ABS-KEY (“information system*” OR “health information system*” OR “computer system*“) AND TITLE-ABS-KEY (implementation OR deployment) AND TITLE-ABS-KEY (transformation OR change OR “organizational change”) AND TITLE-ABS-KEY (“public sector” OR “health care organization*” OR “social care organization*"))	22
Web of Science	(((ALL=(“information system*” OR “health information system*” OR “computer system*“)) AND ALL=(implementation OR deployment)) AND ALL=(transformation OR change OR “organizational change”)) AND ALL=(“public sector” OR “health care organization*” OR “social care organization*")	23
PubMed	(((((((“information system”[Title/Abstract] OR “health information system”[Title/Abstract] OR “computer system”[Title/Abstract]) AND “implementation”[Title/Abstract]) OR “deployment”[Title/Abstract]) AND “transformation”[Title/Abstract]) OR “change”[Title/Abstract] OR “organizational change”[Title/Abstract]) AND “public sector”[Title/Abstract]) OR “health care organization”[Title/Abstract] OR “social care organization”[Title/Abstract])	464

Textbox 1. Inclusion criteria.

- Available electronically
- Peer-reviewed
- Accessible full text
- Published in English
- Published between 2018 - 2023
- Relevant to the research questions

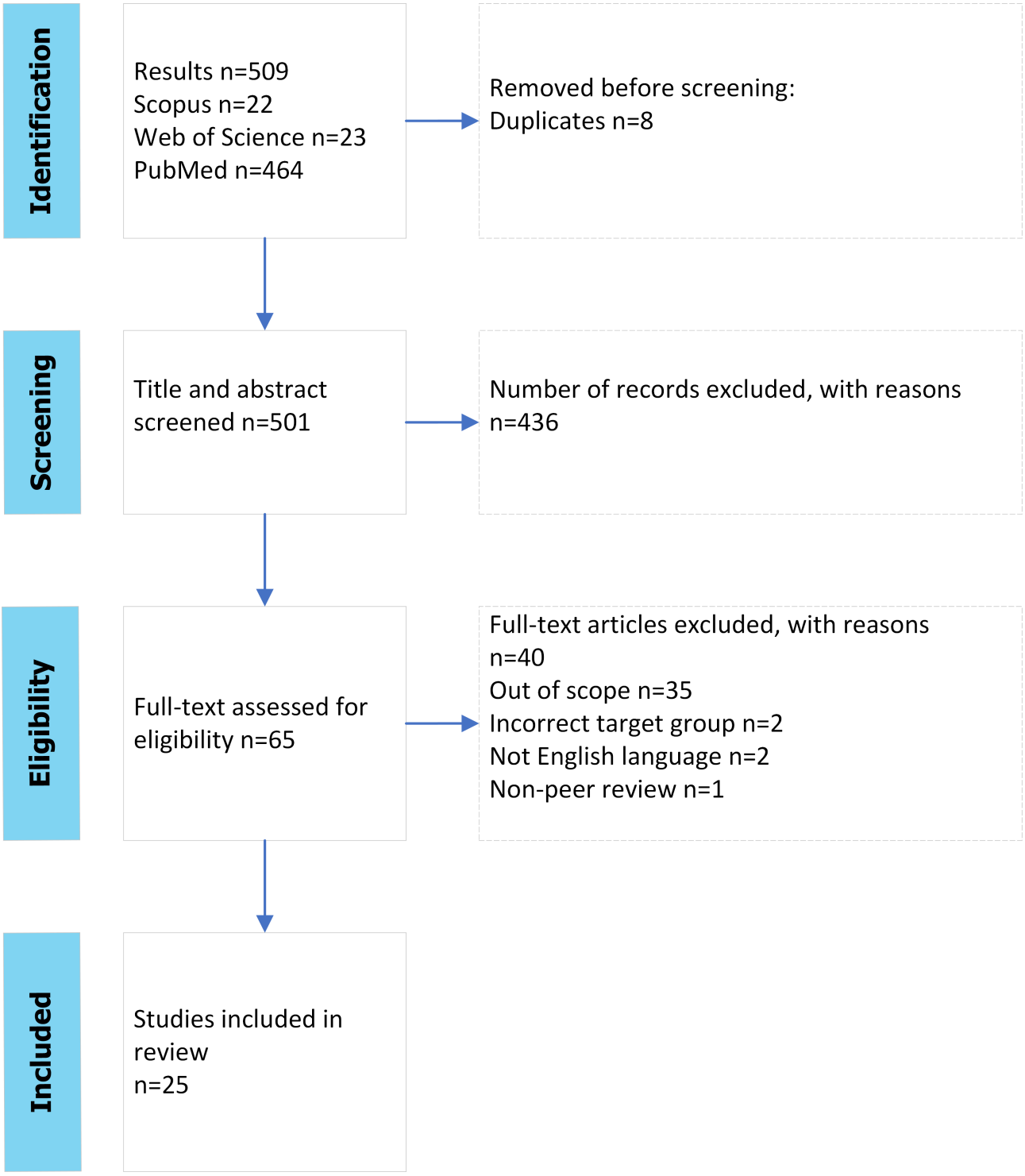
Results

Overview

The data handling process is illustrated in the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) flow diagram (Figure 1) [42] and the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) checklist is present as Checklist 1. The final study included 25 primary research articles, and the data are summarized in Multimedia Appendix 1. Narratives about IS implementations presented in research articles are summarized in Multimedia Appendix 2. The data analysis was conducted by using force fields of the Lewin’s change model as it provides a framework to assess the success of IS reform in relation to planning, implementation, and user acceptance [34,35].

XSL•FO
RenderX

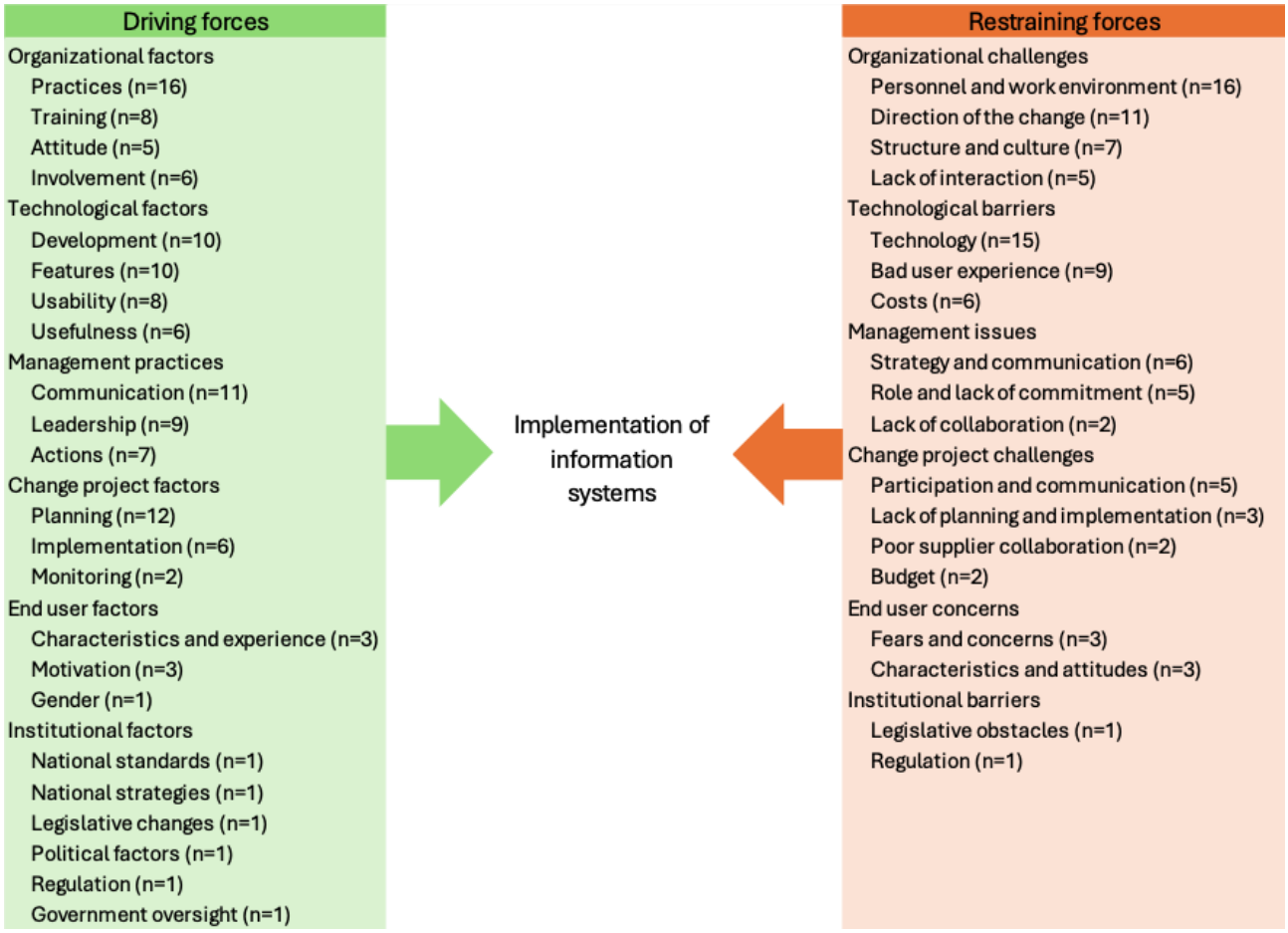
Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram [42].



The articles were published between 2018 and 2024. The publications originated from a total of 15 different countries and 5 different continents. Most of the studies were qualitative (n=18). In addition, there was 1 mixed methods study, 3 quantitative studies, and 3 literature reviews. Public sector domains were health care (n=18), government agencies (n=3), taxation (n=2), and social welfare (n=2). Forces that drive and restrain the implementation of ISs in public sector organizations were grouped into broader categories (Figure 2). Subgroups were created under the categories, highlighting the most frequently mentioned factors identified from the data (Figure 2). While numerous studies addressed avoidable restraining factors, the emphasis in the data was on the driving factors.

XSL•FO
RenderX

Figure 2. Driving and restraining forces affecting on implementation of information systems.



Driving Forces

The implementation of ISs in public sector organizations is influenced by a range of driving forces. However, the majority individual forces fall under the categories of organizational and technological factors.

Organizational Factors

Existing practices within the organization can significantly promote the adoption of new systems. Key factors that enhance the personnel's readiness for change include openness [43], encouragement [44], providing safe space for expressing concerns [44], and proactively addressing issues [45]. If the organization learns through negotiation and interaction [465], it is more capable of implementing changes. Centralized coordination and standardization functions [46], implementation practices [47], and decision-making capability [48] within the organization are also important. Change is facilitated if the organization is built around integrated processes that evolve with changes [49] and if these processes are monitored and adjusted when necessary [50]. IT governance reform [48] should be considered as market-driven solution development [46] emerging as a driving factor. The use of a same brand [51] can promote the implementation of ISs in large entities. In extreme cases, personnel can be mandated to use the system [52].

Training of the personnel and end users was highlighted in 8 studies [44,51-57]. Training should occur before the implementation of the ISs, but postimplementation training and guidance should also be emphasized. In addition to training, participation was noted in 6 studies in the dataset [43,52,55,58-60]. All stakeholders, including frontline professionals [58], should be involved from the very beginning.

Attitudinal factors can manifest within the organization as a whole and among its members. In health care, there is a noticeable sense of moral and ethical obligation among professionals to promote change [59], particularly among those in leadership roles. A positive attitude toward innovation [43] and determined efforts [51] supports the change. The organization's ability to create space for creativity [43], agility, and communication along with confidence in the success of the change [59] eases the process.

Technological Factors

Key aspects of development include good software designing [61] and its feasibility [46], implementation strategy [61], testing and validation practices [55,61], and the sensitivity of software developers to the client's need [61]. Innovative solutions are more likely to emerge through local development [46], but it is crucial that solutions also pass pre-established quality gates [53].

ISs possess various features that either drive or restrain their implementation. The most significant driving factor is solution that support workflows [50,55,57,62]. Other factors include the system's observability [54], interoperability [55], and compatibility [54]. The more the system is used, the more likely

XSL•FO
RenderX

the usage will increase [56]. The IS must also fit within the organizational culture [53]. When the IS is perceived to provide more relative benefits, it further drives its adoption [54].

Ease of use was identified as a driving factor in 4 studies [50,56,58,62]. Usability of the IS was discussed in 2 studies [47,55]. It is essential that needed information must be accessible [62] and the system is available for user testing [54]. Perceived usefulness [55,56,58] and even forecasts of positive impacts [56] are considered driving factors. It is crucial that the IS has the potential to enhance aspects such as the quality of care [45] and it supports professionals in performing their tasks [63].

Management Practices

In the context of implementing ISs in public sector organization, management involves leadership activities, communication skills, continuous assessment, and collaboration. In terms of communication, the involvement of individuals in leadership positions emerged as significant factor [43,59]. At a more detailed level, it is essential for management to convey expectations [50] and objectives [60] related to the change. Ensuring that the organization's vision is clearly understood in the context of the change is important [64]. Expectations are placed on leaders' communication, as listening to concerns and doubts [55,60], thereby fostering greater understanding [59]. Managers must be able to address questions regarding changes [59] and be capable of negotiating potential tensions [65]. They should also possess the capability to motivate users [58].

Overall, leadership and change management were highlighted in 3 studies [44,52,55]. Those in leadership positions must ensure the success of the change process [60] and have confidence in the success of the change [65]. The alignment of differing perspectives [64] and the involvement of leaders [48] are also seen as driving factors. Notably, Fennelly et al [55] observed that successful implementation is more likely when the change is not managed directly by top leadership as otherwise the data emphasized the importance of active leadership engagement.

When leading the change, management should focus on preventing the consolidation of the conflicting views [64] and addressing any ambiguities or issues [60] to promote the acceptance of innovations. Managing the change process [58] and monitoring potential performance deficiencies [49] are key factors in implementation process. Management must be able to develop and communicate context-specific goals [58] for the change. Although it was previously noted that change should not be directly led by top management, a collaborative IT group led by management can promote the IS change [49].

Change Project Factors

At the beginning, it is essential to secure resources [55] and ensure clear roles and responsibilities [48,55]. A project team must be established [52] and the supporting structures for the change processes addressed [66]. Further driving factors include stakeholder involvement from the perspective of the change project [66] and the early engagement of various stakeholders [45]. Maintaining strong and trustworthy relationships with software vendors and consulting firms is critical [55]. In early stages of the change project, attention should be paid to the implementation strategy [55], as well as to the practices for adoption that will arise in later phases of the process [43].

During the implementation phase, the importance of the project management team and governance [55] becomes important. A well-planned implementation process [55] supports the change management of the project [50]. Establishing a command center for the adoption phase [48], involving operational-level leadership roles [55], facilitates the actual system rollout. The change project must have clear and measurable goals from a monitoring perspective [55]. Strong communication practices are essential to promoting the desired practical implementation [46].

End User Factors

Regarding user characteristics, a high level of education [56], technological proficiency [62], and overall competence [53] are identified as factors that drives the adoption of the IS. In addition, the motivation [44], commitment [44], and efforts [65] of end users are critical. In 1 study, being female was also identified as a driving factor [56].

Institutional Factors

Institutional factors refer to government guidance and external elements that influence the organizational environment. They are essential factors from the perspective of public sector organizations. National strategies [51] and standards [55] guide the development of the social and health care sectors. Legislative changes [51] and regulations [67] can support or accelerate system reforms within organizations. Government oversight [52] and political factors [67] are also driving forces.

Restraining Forces

Most individual factors restraining the implementation of ISs fell under the categories of organizational challenges and technological barriers. In total, 4 studies [43,47,51,62] did not address restraining factors.

Organizational Challenges

Concerns related to personnel and the work environment include various issues as the loss of professional identity, patient safety, reduced visibility of work and workforce retention, which negatively impact change processes related to ISs [44]. There are concerns about the possibility of professional expertise being replaced by technology or technology being designed for management rather than professionals [63]. There may be difficulties in harmonizing different interpretations of the technology within the organization [64]. Professionals are worried about potentially having less time with patients [44], compromising core values [63] and experiencing general fear [60]. The reluctance of personnel to adopt new systems [52] and resistance to change [54] are barriers that may be related to a lack of incentive systems [49]. Poor job satisfaction [44] and experienced stress among personnel [60] can restrain the change process and may trigger a wave of resignations due to the reform or stress it brings [66].

Factors restraining change also include phenomena related to the direction of the organization concerning the transformation. There may be active resistance to change within the organization, with a lack of belief in improvements as well as

XSL•FO
RenderX

concerns about successful implementation [59]. In addition, there may be covert resistance [67] to change. If there is a sense of being an outsider in the organization concerning the development and planning of ISs [44], the reform is likely to face opposition. The organization's inadequate learning capacity [49], especially the lack of administrative learning [46], restrain the change process. Factors restraining the direction of the reform within the organization may involve viewing technology as a limiting factor [63] or simply a lack of information [57]. External factors include a heavy reliance on partners or their lack of commitment to common changes, both of which can be obstructive factors [54].

As its worst, organizational culture can inherently restrain change [47]. Passive resistance and indifference embedded in its structures may surface during cycles of change [59]. The lack of a shared vision [49] and collaboration [53] are factors that inhibit change as well as a lack of trust between organizations [49]. Insufficient communication and tensions within the organization negatively impact not only the organization itself but also the implementation of change and ISs [64]. Hidden power dynamics [67], inappropriate behavior [60], and problems in social interactions [48] are restraining factors.

Technological Barriers

The complexity of the technology being implemented was identified as a restraining factor in 4 studies [46,50,53,54]. Another technological barrier may arise when software developers do not operate in the same environment where the software and tools are used [61]. If ISs are developed with a tool-centric rather than a user-centric approach [55], issues such as system underutilization [52], may occur. Technical problems [57], system downtime [44], and slow performance [44] can delay the stabilization of change. Technology can impose limitations [54] and cause workflow issues [48], which may lead to the adoption of undesirable practices [63]. Delays in transferring data from the old system to the new one [52] and the complete absence of data [44] do not support the wanted change.

In terms of user experience, the inflexibility of the system was highlighted in 2 studies [50,53]. Concerns related to the reform of ISs include issues regarding data protection, information security, and the doctor-patient relationship [55]. The potential narrowing of professional discretion [63] may be a concern for professionals. The lack of clinical efficiency and safety [65] affects the user experience. Negative previous experiences of professionals [54] influence their attitudes toward changes.

Cost factors related to ISs were identified in 2 studies [48,53]. In an ever-changing operational environment, costs can easily escalate [53]. The extensive need for customization [53] and the requirement for unique system adjustments [48] increase costs when using so called off-the-self-software. If the IS demands the use of specialized resources [53], this raises associated costs. High maintenance costs and inadequate support from the system provider to the organization are restraining factors [53].

Management Issues

Leadership must not leave strategic objectives vaguely defined [53]. A lack of communication regarding the vision [66], insufficient communication of the purpose [66], and inadequate presentation of goals [58] create significant obstacles to change. The lack of information [66] is one of the greatest challenges in transformations. Change initiatives may be deemed unsuccessful due to their scale [53]. If the upcoming implementation project is not adequately communicated to the personnel [59], the consequences may restrain the change.

The leader's role is to be visible and committed to the change [66], as lack of responsibility [52] is unfavorable from the perspective of reform. It is also common for personnel to not be informed at all due to uncertainty [59], leading to feelings of abandonment among employees [66]. Challenges arise especially when management's expectations for change conflict the values of the employees [66], such as with productivity expectations. In terms of collaboration, failure can occur in estimating the size of the project [53], often resulting in the project being larger than initially anticipated. Another identified barrier is the lack of close cooperation between clinical leadership and IT professionals [65].

Change Project Challenges

Inadequate communication about the project's tasks [48] provides an undesirable start to a change project. If the participation efforts of various stakeholders are not considered [59], this can increase resistance to change. Involving end users appropriately in the selection of the system [53] drives the implementation. Regarding the implementation process, the lack of transparency and improper execution [59] were identified as barriers. Poor planning [53], delays [53], and the inability to manage a centralized timeline [48] are challenges related to project planning and execution. Budget constraints and overruns [53] are restraining factors. As for supplier collaboration, inadequate screening of suppliers [53] and a nontransparent and nonopen relationship with the supplier can hinder the process.

End User Concerns

End users experience concerns and fears about the impact of artificial intelligence, data privacy, and the replacement of humans by technology [45]. Among user characteristics, advanced age was identified as a restraining factor [56]. In terms of attitudes, user frustration [44] and general resistance to acceptance [57] are also barriers.

Institutional Barriers

As with driving factors, institutional barriers have been identified in only few studies, but they are included due to their significance. Legislative obstacles refer to factors such as data protection and security laws [46]. Regulatory factors [58] can restrain or delay IS reforms, especially in the public sector.

A Narrative of the Change Process in IS Implementations

Based on the data, the procurement of ISs is one of the largest decisions a public sector organization can make [48]. These projects are complex [57] and often fail to meet their objectives [58,62,64]. The IS project represents a significant financial

XSL•FO
RenderX

investment [48,53]. If the system proves impractical and offers limited value, it is considered a poor investment [63]. Although health care has seen rapid technological growth over the past 50 years, it has lagged behind other industries in its use of technology [54,58] and integrating technology into health care remains challenging [50].

Large health care organizations frequently face difficulties in adopting high-quality, modern technological solutions [43]. Although the benefits of systems such as EHRs are widely recognized, their potential is not always fully realized, often due to the implementation process [55]. Furthermore, the introduction of new technology has been shown to increase stress among professionals, potentially worsen existing issues such as personnel shortages and heavy workloads [44].

Changes often lead to uncertainty [66] and anxiety about how they will affect employees' work [59]. The implementation of an IS requires time and resources from organization's perspective [53], and it can disrupt routines [45]. In public sector, digital projects can easily become stalled, resulting delays of several years [46]. Leaders should have necessary expertise to carry out these projects [49].

It appears that the implementation of ISs remains an unresolved challenge for organizations [43,58], but innovations also offer an opportunity to transform organizational practices [67]. Over the years, technological advancements have facilitated the organization and provision of services [62], and they can radically transform how business is conducted [44,52]. For instance, the new generation of EHR systems can offer various forms of support related to patient care and operational planning [62].

ISs play a crucial role in meeting the demands for reform in public sector organizations and the use of technology has increased within public services [63]. Technology offers significant advantages in reducing errors, improving communication and enhancing patient and customer satisfaction [56]. The use of health technology in hospital setting has resulted in several benefits, such as improved services and reduced medication errors, thus promoting patient safety [61]. However, new customer and patient safety errors caused by technology continue to be a significant issue [61]. Nonetheless, electronic health care has a clear and growing impact on health care delivery worldwide and contributes to the efficiency of health care systems [51]. Digitally operated public services can make citizens' lives easier, more satisfying, and safer [49].

Discussion

Principal Findings

The purpose of the study was to enhance the understanding of factors influencing the implementation of ISs in public sector organizations and how to prepare for such change. The main findings indicate that the implementation of ISs in public sector organizations is influenced by a range of driving and restraining forces. These forces can be categorized into six domains, which vary depending on whether the force is driving or restraining: (1) organizational practices and challenges, (2) technological factors and barriers, (3) management practices and issues, (4) change project factors and challenges, (5) end user factors and concerns as well as (6) institutional factors and barriers. The results align with previous research underscoring the persistent challenges in IS implementation within public sector, such as high failure rates, significant financial losses, and diminished trust in technology [6,11-16,43,64]. However, they also highlight the transformative potential of digital technologies, including improved efficiency, reduced errors, and enhanced user satisfaction [4,5,51,56,62].

Key driving forces (leadership, stakeholder involvement, and usability) and restraining forces (user resistance, lack of managerial expertise, and technical challenges) were identified as major factors influencing IS implementation public sector. These findings align with Lewin's change theory, which emphasizes that successful IS implementation depends on overcoming restraining forces and reinforcing driving forces to facilitate a sustainable transformation [34]. The Lewin's change model provides framework for understanding and managing the dynamics of the change as it helps to identify and analyze both driving and restraining forces to enable targeted actions to support the wanted outcome. Force field analysis helps organizations to plan their transformation journey [34,35].

Leadership support was identified as a critical driver, reducing resistance, and aligning organizational goals [43,44,49]. Leadership activities, such as addressing concerns [55,60] and fostering a shared vision for change [59] were pivotal in building trust and guiding the implementation process [64]. These findings align with earlier research highlighting the central role of active and engaged leadership in complex organizational transformations [11,18,20-22,31].

Training and stakeholder involvement were equally important, with 8 studies emphasizing pre- and postimplementation training as critical for building user competence and confidence [44,51-57]. Stakeholder participation, particularly from frontline professionals, facilitated a sense of ownership and reduced resistance, as noted in 6 studies [43,52,55,58-60].

On the other hand, user resistance [16,28-30], end users' key restraining force, was often rooted in fears of increased administrative burdens and reduced professional autonomy [44,60]. Technical challenges, such as bad user experience [50,53,54,65], system downtime [44] and slow performance [50] further interfered with the process. Organizational silos and a lack of a shared vision were also highlighted as significant barriers to successful implementation [49,64].

Previous research emphasizes the need to understand that even the best IS alone cannot meet the expectations placed on it and it is crucial to change practices at individual and organizational level [18,21,23]. Results of this study indicates that the most significant feature of an IS is its alignment with organizational workflows [50,55,57,62], which means that the objectives set for the IS project must consider the processes and the context of the operating environment [25,26].

A narrative analysis of the introduction sections of the studies included in this scoping review revealed how previous research has described IS implementations in public sector organizations. The findings indicate that IS implementation remains a complex

XSL•FO
RenderX

and unresolved challenge [43,57,58]. These projects require significant financial investment [48,53], yet they often fail to meet their objectives [58,62,64]. One major barrier is the disruption IS implementation causes to organizational routines, increasing stress among personnel and worsening workforce shortages [44,45]. Uncertainty about changes and a lack of leadership further hinder the progress [49]. Despite these challenges, the positive narrative from analyzed data emphasizes the benefits of the technology advancement and digitalization such as improved service organization [44,52,62] and increased efficiency [51] as well as reduced errors and enhanced satisfaction [49,55,61]. To conclude, studies acknowledge the significant challenges and risks involved in IS implementation, but they also emphasize the transformative potential in improving operations and service quality. This approach suggests that IS implementation is not inherently problematic, but rather that its success depends on how effectively the implementation process is managed and aligned with organizational needs.

Limitations

This scoping review aimed to provide a broad overview and summarize the breadth and depth of the known forces influencing IS implementation in public sector organizations and to examine the narrative presented in previous search. Existing knowledge was identified, evaluated, interpreted, and synthesized. The objective of this study was not to assess the quality of the evidence, although only peer-reviewed scientific papers were included in the review. While specific search queries and inclusion criteria were applied, it is possible that some relevant studies were not captured in the process.

Conclusions

The decision to implement an IS is a significant and complex process, often encountering challenges in achieving its objectives. These challenges relate to high costs, technical difficulties and the complexity of implementation. However, the adoption of technology also brings positive impacts. Successfully implemented ISs improve services, reduce errors, and increase customer and patient satisfaction. Despite the challenges, a well-managed implementation process can bring significant benefits for customers, organizations, and society.

This study contributes to a growing body of knowledge of public sector technology implementation, offering valuable insights for future projects, which should be examined and planned comprehensively from various identified perspectives. Previous research has primarily focused on the failure of IS implementations, highlighting the need for further studies on the factors influencing successful implementation in the public sector organizations.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Data of the review.
[PDF File, 161 KB - humanfactors_v12i1e71575_app1.pdf]

Multimedia Appendix 2
Narrative about information system implementations.
[PDF File, 62 KB - humanfactors_v12i1e71575_app2.pdf]

Checklist 1
PRISMA-ScR (Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews) checklist.
[PDF File, 97 KB - humanfactors_v12i1e71575_app3.pdf]

References

1. Plesner U, Justesen L, Glerup C. The transformation of work in digitized public sector organizations. J Organ Change Manag 2018 Sep 12;31(5):1176-1190. [doi: 10.1108/JOCM-06-2017-0257]
2. Kitsios F, Kamariotou M, Mavromatis A. Drivers and outcomes of digital transformation: the case of public sector services. Information 2023;14(1):43. [doi: 10.3390/info14010043]
3. Mergel I, Edelmann N, Haug N. Results from expert interviews. Gov Inf Q 2019;36(4). [doi: 10.1016/j.giq.2019.06.002]
4. Senbekov M, Saliev T, Bukeyeva Z, et al. The recent progress and applications of digital technologies in healthcare: a review. Int J Telemed Appl 2020;2020:8830200. [doi: 10.1155/2020/8830200] [Medline: 33343657]
5. Valkeapää E, Peltonen LM. Terveydenhuollon digitaalisten tietojärjestelmien vaikutus hoitohenkilöstöön ja potilastuloksiin: Kartoittava kirjallisuuskatsaus hoitotyön näkökulmasta [Article in Finnish]. FinJeHeW 2022;14(2):150-165. [doi: 10.23996/fjhw.110621]
6. Neittaanmäki P, Lehto M, Ruohonen T, Kaaselainen K, Karla T. Suomen Terveysdata Ja Sen Hyödyntäminen: Loppuraportti [Book in Finnish]: Yliopistopaino Jyväskylä; 2019, Vol. 4.
7. Dybå T, Dingsøyr T. Empirical studies of agile software development: a systematic review. Inf Softw Technol 2008 Aug;50(9-10):833-859. [doi: 10.1016/j.infsof.2008.01.006]

XSL•FO
RenderX

8. Chow T, Cao DB. A survey study of critical success factors in agile software projects. J Syst Softw 2008 Jun;81(6):961-971. [doi: 10.1016/j.jss.2007.08.020]
9. Goulielmos M. Systems development approach: transcending methodology. Inf Syst J 2004 Oct;14(4):363-386. [doi: 10.1111/j.1365-2575.2004.00175.x]
10. Marwaha JS, Landman AB, Brat GA, Dunn T, Gordon WJ. Deploying digital health tools within large, complex health systems: key considerations for adoption and implementation. NPJ Digit Med 2022 Jan 27;5(1):13. [doi: 10.1038/s41746-022-00557-1] [Medline: 35087160]
11. Money WH, Kaisler SH, Cohen SJ. Understanding failed software projects through forensic analysis. Journal of Computer Information Systems 2022 Sep 3;62(5):940-953. [doi: 10.1080/08874417.2021.1950076]
12. Jones C. Social and technical reasons for software project failures. J Defense Softw Eng 2006;19(6):4-9 [FREE Full text]
13. Sanchez OP, Terlizzi MA, de Moraes HDO. Cost and time project management success factors for information systems development projects. Int J Proj Manag 2017 Nov;35(8):1608-1626. [doi: 10.1016/j.ijproman.2017.09.007]
14. Dendere R, Janda M, Sullivan C. Are we doing it right? We need to evaluate the current approaches for implementation of digital health systems. Aust Health Rev 2021 Dec;45(6):778-781. [doi: 10.1071/AH20289] [Medline: 34488938]
15. Ammenwerth E, Iller C, Mahler C. IT-adoption and the interaction of task, technology and individuals: a fit framework and a case study. BMC Med Inform Decis Mak 2006 Jan 9;6(3):3. [doi: 10.1186/1472-6947-6-3] [Medline: 16401336]
16. Ekholm S, Kinnunen UM. Tietojärjestelmän käyttöönottoa tukevat teoreettiset mallit terveydenhuollossa [Article in Finnish]. FinJeHeW 2016;8(2-3):63-73 [FREE Full text]
17. Coccia M. New perspectives in innovation failure analysis: a taxonomy of general errors and strategic management for reducing risks. Technol Soc 2023 Nov;75:102384. [doi: 10.1016/j.techsoc.2023.102384]
18. Gill R. Change management--or change leadership? J Change Manag 2002 Dec;3(4):307-318. [doi: 10.1080/714023845]
19. Weber E, Büttgen M, Bartsch S. How to take employees on the digital transformation journey: an experimental study on complementary leadership behaviors in managing organizational change. J Bus Res 2022 Apr;143:225-238. [doi: 10.1016/j.jbusres.2022.01.036]
20. Agreli H, Huising R, Peduzzi M. Role reconfiguration: what ethnographic studies tell us about the implications of technological change for work and collaboration in healthcare. leader 2021 Jun;5(2):134-141. [doi: 10.1136/leader-2020-000224]
21. Valta M. Sähköisen potilastietojärjestelmän sosiotekninen käyttöönotto. Seitsemän vuoden seurantatutkimus odotuksista omaksumiseen [Dissertation in Finnish]. : University of Eastern Finland.
22. Gooch P, Roudsari A. Computerization of workflows, guidelines, and care pathways: a review of implementation challenges for process-oriented health information systems. J Am Med Inform Assoc 2011;18(6):738-748. [doi: 10.1136/amiajnl-2010-000033] [Medline: 21724740]
23. Bartoli A, Hermel P. Managing change and innovation in IT implementation process. J Manuf Technol Manag 2004 Jul 1;15(5):416-425. [doi: 10.1108/17410380410540417]
24. Konopik J, Jahn C, Schuster T, Hoßbach N, Pflaum A. Mastering the digital transformation through organizational capabilities: a conceptual framework. Digital Business 2022;2(2):100019. [doi: 10.1016/j.digbus.2021.100019]
25. Ke W, Wei KK. Organizational culture and leadership in ERP implementation. Decis Support Syst 2008 May;45(2):208-218. [doi: 10.1016/j.dss.2007.02.002]
26. Lang A, Mertes A. Governance of large innovation projects: the implementation of the electronic health card in germany. In: Innovation Policy and Governance in High-Tech Industries: Springer; 2012. [doi: 10.1007/978-3-642-12563-8_11]
27. Tortorella GL, Fogliatto FS, Tlapa Mendoza D, Pepper M, Capurro D. Digital transformation of health services: a value stream-oriented approach. Int J Prod Res 2023 Mar 19;61(6):1814-1828. [doi: 10.1080/00207543.2022.2048115]
28. Doolin B. Power and resistance in the implementation of a medical management information system. Inf Syst J 2004 Oct;14(4):343-362. [doi: 10.1111/j.1365-2575.2004.00176.x]
29. Lapointe L, Rivard S. A multilevel model of resistance to information technology implementation. MIS Q 2005;29(3):461. [doi: 10.2307/25148692]
30. Kim JS, Chung GH. Implementing innovations within organizations: a systematic review and research agenda. Innovation (Abingdon) 2017 Jul 3;19(3):372-399. [doi: 10.1080/14479338.2017.1335943] [Medline: 28129321]
31. Ahimbisibwe A, Cavana R, Daellenbach U. A contingency fit model of critical success factors for software development projects. A comparison of agile and traditional plan-based methodologies. J Enterp Inf Manag 2015;28(1):7-33. [doi: 10.1108/JEIM-08-2013-0060]
32. Lewin K. Field Theory in Social Science: Selected Theoretical Papers: Harper & Brothers Publishers; 1951.
33. Burnes B. The origins of Lewin's three-step model of change. J Appl Behav Sci 2020 Mar;56(1):32-59. [doi: 10.1177/0021886319892685]
34. Bozak MG. Using Lewin's force field analysis in implementing a nursing information system. Comput Inform Nurs 2003;21(2):80-85. [doi: 10.1097/00024665-200303000-00008] [Medline: 12802948]
35. Trzeciak M. Factors of success in the change management process of IT programs. J Organ Change Manag 2024 Feb 6;37(1):58-74. [doi: 10.1108/JOCM-04-2023-0110]
36. Kaminski J. Theory applied to informatics – Lewin´s change theory. Can J Nurs Inform 2011;6(1).

XSL•FO
RenderX

37. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5(69):69. [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
38. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol 2018 Nov 19;18(1):143. [doi: 10.1186/s12874-018-0611-x] [Medline: 30453902]
39. Vilkka H. Kirjallisuuskatsaus Metodina, Opinnäytetyön Osana Ja Tekstilajina: Printon; 2023.
40. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
41. Aromataris E, Lockwood C, Porrit K, Pilla B, Jordan Z, editors. JBI Manual for Evidence Synthesis: JBI; 2024. [doi: 10.46658/JBIMES-24-01]
42. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372(71):n71. [doi: 10.1136/bmj.n71]
43. Esdar M, Hübner U, Thye J, Babitsch B, Liebe JD. The effect of innovation capabilities of health care organizations on the quality of health information technology: model development with cross-sectional data. JMIR Med Inform 2021 Mar 15;9(3):e23306. [doi: 10.2196/23306] [Medline: 33720029]
44. Jedwab RM, Manias E, Hutchinson AM, Dobroff N, Redley B. Nurses' experiences after implementation of an organization-wide electronic medical record: qualitative descriptive study. JMIR Nurs 2022 Jul 26;5(1):e39596. [doi: 10.2196/39596] [Medline: 35881417]
45. Nitiéma P. Artificial intelligence in medicine: text mining of health care workers' opinions. J Med Internet Res 2023 Jan 27;25:e41138. [doi: 10.2196/41138] [Medline: 36584303]
46. Di Giulio M, Vecchi G. Multilevel policy implementation and the where of learning: the case of the information system for school buildings in Italy. Policy Sci 2019 Mar;52(1):119-135. [doi: 10.1007/s11077-018-9326-4]
47. Hörhammer I, Kujala S, Hilama P, Heponiemi T. Building primary health care personnel's support for a patient portal while alleviating eHealth-related stress: survey study. J Med Internet Res 2021 Sep 22;23(9):e28976. [doi: 10.2196/28976] [Medline: 34550087]
48. Kiepek W, Sengstack PP. An evaluation of system end-user support during implementation of an electronic health record using the model for improvement framework. Appl Clin Inform 2019 Oct;10(5):964-971. [doi: 10.1055/s-0039-3402450] [Medline: 31853937]
49. Pittaway JJ, Montazemi AR. Know-how to lead digital transformation: the case of local governments. Gov Inf Q 2020 Oct;37(4):101474. [doi: 10.1016/j.giq.2020.101474]
50. Tossaint-Schoenmakers R, Versluis A, Chavannes N, Talboom-Kamp E, Kasteleyn M. The challenge of integrating eHealth into health care: systematic literature review of the Donabedian model of structure, process, and outcome. J Med Internet Res 2021 May 10;23(5):e27180. [doi: 10.2196/27180] [Medline: 33970123]
51. Haverinen J, Keränen N, Tuovinen T, Ruotanen R, Reponen J. National development and regional differences in eHealth maturity in Finnish public health care: survey study. JMIR Med Inform 2022 Aug 12;10(8):e35612. [doi: 10.2196/35612] [Medline: 35969462]
52. Agyei-Ababio N, Ansong E, Assa-Agyei K. Digitalization of revenue mobilization in an emerging economy: the new institutional theory perspective. IJISPM 2023;11(2):5-22. [doi: 10.12821/ijispm110201]
53. Shimange TP, Pillay K. A South African institution perspective of a framework for enterprise resource planning systems. SA Journal of Information Management 2023;25(1). [doi: 10.4102/sajim.v25i1.1578]
54. Han L, Liu J, Evans R, Song Y, Ma J. Factors influencing the adoption of health information standards in health care organizations: a systematic review based on best fit framework synthesis. JMIR Med Inform 2020 May 15;8(5):e17334. [doi: 10.2196/17334] [Medline: 32347800]
55. Fennelly O, Cunningham C, Grogan L, et al. Successfully implementing a national electronic health record: a rapid umbrella review. Int J Med Inform 2020 Dec;144:104281. [doi: 10.1016/j.ijmedinf.2020.104281] [Medline: 33017724]
56. Al-Otaibi J, Tolma E, Alali W, Alhuwail D, Aljunid SM. The factors contributing to physicians' current use of and satisfaction with electronic health records in Kuwait's public health care: cross-sectional questionnaire study. JMIR Med Inform 2022 Oct 7;10(10):e36313. [doi: 10.2196/36313] [Medline: 36206039]
57. Saiyed S, Joshi R, Khattab S, Dhillon S. The rapid implementation of an innovative virtual diabetes boot camp program: case study. JMIR Diabetes 2022 Jan 14;7(1):e32369. [doi: 10.2196/32369] [Medline: 35029529]
58. De Benedictis A, Lettieri E, Gastaldi L, Masella C, Urgu A, Tartaglini D. Electronic Medical Records implementation in hospital: an empirical investigation of individual and organizational determinants. PLOS ONE 2020;15(6):e0234108. [doi: 10.1371/journal.pone.0234108] [Medline: 32497058]
59. Stefánsdóttir NT, Nilsen P, Lindstroem MB, et al. Implementing a new emergency department: a qualitative study of health professionals' change responses and perceptions. BMC Health Serv Res 2022 Apr 5;22(1):447. [doi: 10.1186/s12913-022-07805-w] [Medline: 35382815]
60. Raza MA, Imran M, Rosak-Szyrocka J, Vasa L, Hadi NU. Organizational change and workplace incivility: mediated by stress, moderated by emotional exhaustion. Int J Environ Res Public Health 2023 Jan 21;20(3):2008. [doi: 10.3390/ijerph20032008] [Medline: 36767374]

XSL•FO
RenderX

61. Ndabu T, Mulgund P, Sharman R, Singh R. Perceptual gaps between clinicians and technologists on health information technology-related errors in hospitals: observational study. JMIR Hum Factors 2021 Feb 5;8(1):e21884. [doi: 10.2196/21884] [Medline: 33544089]
62. Kayser L, Karnoe A, Duminski E, et al. Health professionals' eHealth literacy and system experience before and 3 months after the implementation of an electronic health record system: longitudinal study. JMIR Hum Factors 2022 Apr 29;9(2):e29780. [doi: 10.2196/29780] [Medline: 35486414]
63. Sarwar A, Harris M. Children's services in the age of information technology: what matters most to frontline professionals. J Soc Work (Lond) 2019 Nov;19(6):699-718. [doi: 10.1177/1468017318788194]
64. Bunduchi R, Tursunbayeva A, Pagliari C. Coping with institutional complexity: intersecting logics and dissonant visions in a nation-wide healthcare IT implementation project. Inf Technol People 2020;33(1):311-339. [doi: 10.1108/ITP-08-2018-0373]
65. Bernardi R, Exworthy M. Clinical managers' identity at the crossroad of multiple institutional logics in it innovation: the case study of a health care organization in England. Information Systems Journal 2020 May;30(3):566-595. [doi: 10.1111/isj.12267]
66. Morrison VJ, Jensen AL. Between a rock and a hard place: Nurse managers' experiences of large-scale organizational change in the public health service. J Adv Nurs 2022 Oct;78(10):3385-3397. [doi: 10.1111/jan.15413] [Medline: 35975292]
67. Kemal AA, Shah MH. Digital innovation in social cash organizations – the effects of the institutional interactions for transforming organizational practices. Inf Technol People 2024 Jul 5;37(5):2092-2126. [doi: 10.1108/ITP-02-2023-0176]

Abbreviations

EHR: electronic health record
IS: information system
IT: information technology
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews

Edited by A Kushniruk; submitted 22.01.25; peer-reviewed by A Mertes, NS Aziz; revised version received 08.04.25; accepted 05.05.25; published 11.06.25.

Please cite as:
Lemmettylä A, Kinnunen UM
Driving and Restraining Forces in the Implementation of Information Systems in the Public Sector: Scoping Review
JMIR Hum Factors 2025;12:e71575
URL: https://humanfactors.jmir.org/2025/1/e71575
doi:10.2196/71575

© Arja Lemmettylä, Ulla-Mari Kinnunen. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 11.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Tailoring for Health Literacy in the Design and Development of eHealth Interventions: Systematic Review

Jan Wessel Hovingh[1,2,3], MA; Cynthia Elderson-van Duin[1], MSc; Derek A Kuipers[3], PhD; Yanda van Rood[4], PhD; Geke D S Ludden[2], PhD; Denise J C Hanssen[1], PhD; Judith G M Rosmalen[1,5], PhD

[1]Department of Psychiatry, University Medical Center Groningen, University of Groningen, Hanzeplein 1, Groningen, The Netherlands

[2]Department of Interaction Design, Department of Design, Production and Management, University of Twente, Enschede, The Netherlands

[3]Future Design Center, Faculty of Tech and Design, NHL Stenden University of Applied Sciences, Groningen, The Netherlands

[4]Department of Psychiatry, Leiden University Medical Center, Leiden University, Leiden, The Netherlands

[5]Department of Internal Medicine, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

Corresponding Author:
Jan Wessel Hovingh, MA
Department of Psychiatry, University Medical Center Groningen, University of Groningen, Hanzeplein 1, Groningen, The Netherlands

Abstract

Background: Tailoring is an important strategy to improve uptake and efficacy of medical information and guidance provided through eHealth interventions. Given the rapid expansion of eHealth, understanding the design rationale of such tailored interventions is vital for further development of and research into eHealth interventions aimed at improving health and healthy behavior.

Objective: This systematic review examines the use of health literacy concepts through tailoring strategies in digital health interventions (eHealth) aimed at improving health and how these elements inform the overall design rationale.

Methods: A systematic search of PubMed, PsycINFO, Web of Science, and ACM databases yielded 31 eligible randomized trials that focused on adult health improvement through eHealth interventions. Eligible studies compared tailored versus nontailored eHealth interventions for adults, excluding non-English papers and those addressing solely readability or targeting populations with accessibility barriers. Data extraction focused on study characteristics, health literacy components, tailoring methods, and design rationales, with study quality evaluated using Quality Assessment for Diverse Studies (QuADS) by independent reviewers.

Results: Most interventions applied both cognitive and social health literacy concepts and predominantly used content matching as a tailoring strategy. Of all studies using content matching, most used one or more supporting theories as well as end-user data to inform the content matching. While choices for individual intervention components were mostly explicated, detailed descriptions of the design process were scarce, with only a few studies articulating an underlying narrative that integrated the most important chosen components.

Conclusions: While tailored eHealth interventions demonstrate promise in enhancing health literacy and the trial design of the interventions overall was of good quality, inconsistent documentation of design rationales impedes replicability and broader application of the used eHealth concepts. This calls for more detailed reporting on the design choices of the intervention in efficacy studies, so that reported outcomes can be easier connected to choices made in the design of the eHealth intervention.

Trial Registration: PROSPERO 225731; https://www.crd.york.ac.uk/PROSPERO/view/CRD42023225731

(JMIR Hum Factors 2025;12:e76172) doi:10.2196/76172

KEYWORDS

eHealth; tailoring; health literacy; design; design rationale; development; PRISMA

Introduction

In recent years, the design, development, and implementation of digital health interventions, or eHealth interventions, has transformed the landscape of health care and health care delivery [1]. Through these interventions, a wide range of technologies, including mobile apps, portal websites, serious games, and wearables, have been conceived, tested, and implemented in various health care settings. These technologies have shown great promise to enhance accessibility [2] and usability of various health care concepts and services [3]. However, despite the potential benefits of eHealth interventions in improving health outcomes, their impact is not equally distributed. Many end users encounter difficulties in comprehending and using the information provided by these interventions, leading to

XSL•FO
RenderX

underperformance compared with their anticipated potential [4].

One of the most important mechanisms associated with health outcomes is health literacy [5]. Health literacy has been defined as people's knowledge, motivation, and competencies to access, understand, appraise, and apply health information in order to make judgments and take decisions in everyday life concerning health care, disease prevention, and health promotion [6]. Initially, health literacy was defined as a mainly cognitive skill focused on information processing, including understanding, reading, and knowledge of health information. Recent perspectives have adopted a more comprehensive approach to the concept by including the significance of social skills and executive capacities, such as goal-setting, taking action, confidence, motivation, beliefs, and role expectations [7,8]. Most likely, health literacy in its broadest sense impacts the use and effectiveness of eHealth interventions.

One strategy to increase the impact of eHealth interventions is to tailor these interventions toward the specific needs and capacity of the end-user, ensuring that desired knowledge and behavior cues are transferred in the most suitable manner. Tailoring refers to the optimization of communication and functionality based on assessable characteristics unique to that individual. Choosing tailoring tactics, guided by theoretical and empirical evidence, can lead to better understanding and uptake of information and subsequently increased efficacy of the intervention [9].

Regarding health literacy, especially the strategy of content matching, seems important. Named "the crux" of tailoring [10], content matching attempts to direct messages to individuals' status on key theoretical determinants (such as knowledge, outcome expectations, normative beliefs, efficacy, or skills) of the behavior of interest [9]. Ideally, these assessable determinants are established by an evaluative framework, supported by the use of specific supporting theories such as the Transtheoretical Model or the Social Cognitive Theory [10].

Indeed, tailoring for health literacy could follow 2 main conceptual approaches. The first one involves optimizing communication and functionality, focusing on the cognitive capacity of an end-user: in short, an end-user needs to understand and become susceptible to the provided health-related information to act accordingly [9]. The second involves focusing on the social capacity of an end-user. The end-user is advised or encouraged to show certain health-related behaviors [11]. Of course, these 2 approaches are not mutually exclusive and may both be present in an eHealth intervention at the same time, ideally both supporting the same targeted behavior.

The question is how these tailoring strategies, especially content matching, can help to translate the cognitive and social health literacy concepts into a usable and assessable eHealth intervention that meets the desired outcomes. This can be done by establishing the design rationale of the intervention: a representation of the reasoning behind the design of that intervention [12]. It explains choices such as type of technology, content, and usability from an underlying story of the concept that ties together the different elements of the solution into one coherent argument for the solution as a whole [13].

Unfortunately, there is a lack of an overview of how health literacy concepts and how tailoring strategies are used in eHealth interventions, including how health literacy and tailoring are combined. Several reviews have shown that tailored interventions are important for increasing health literacy itself [4,14,15]. However, they did not clarify how social and cognitive health literacy specifically contributed to the tailoring process. Moreover, they did not describe in what way the forthcoming tailoring strategies added to the design rationale of the intervention [4].

Therefore, in this study, we aim:

1. To provide a structured overview of the use of cognitive and social health literacy concepts in tailored eHealth interventions that aim to improve health.
2. To provide a structured overview of tailoring strategies used in eHealth interventions that aim to improve health.
3. To explore how these cognitive and social health literacy concepts and tailoring strategies contributed to the design rationale of eHealth interventions.

Methods

Overview

This systematic review was executed and reported according to PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) criteria for systematic reviews [16]. The review was registered with PROSPERO (International Prospective Register of Systematic Reviews; 225731). The review has been slightly altered after registration in PROSPERO in terms of focus and scope. While the original goal also included studying eHealth intervention efficacy and health literacy-related outcomes, the selected studies varied too much in terms of setup, execution, dosage, and population to draw useful overarching conclusions. Instead, we decided to focus on the design rationales of the eHealth interventions.

Search Strategy

Studies were identified by searching PubMed (MEDLINE), PsycINFO, Web of Science, and Association for Computing Machinery (ACM) for publications before June 2021. The search strategy was based on previous health literacy reviews [4,17]. For all databases, search strings were compiled using 3 main concepts: eHealth, Health Literacy, and Tailoring. For the string "eHealth", variations of the term eHealth as well as alternative descriptions for digital health interventions (ie, mobile health [mHealth], internet intervention, and online intervention) were used. For the search string "Health Literacy", both cognitive (knowledge, competence, and health literacy) and social (patient activation and motivation) health-related terms were used, based on the search strategy from the European Health Literacy Survey (HLS-EU) consortium [4]. For the string "Tailoring", the search strategy of an earlier review was adapted [17], which included terms such as "tailoring", "personalization." The 3 resulting search strings were combined using the Boolean operator "AND". The queries were then adjusted and optimized for any

XSL•FO
RenderX

of the databases. The complete strategy per database can be found in Table S1 in Multimedia Appendix 1.

Eligibility Criteria

When selecting relevant literature, we included (1) trials with at least one control group; (2) eHealth studies aimed at improving health, such as studies on patient education, decision support, risk assessment, health behavior change, treatment, or self-management of physical and psychological illnesses; (3) studies focused on adults; (4) studies that measured health literacy (cognitive, social); (5) studies that compared a tailored eHealth intervention with a nontailored intervention. We excluded all studies that (1) were not written in English; (2) targeted populations lacking basic accessibility requirements (functionally illiterate, deaf, or blind); (3) focused on cultural tailoring; (4) were only focusing on wording and readability of written content; (5) had active human involvement in the intervention, other than the patient himself; (6) only compared differences in medium type (eg. printed media compared with online media).

Study Selection

After extraction from the databases, all identified studies were uploaded to Covidence Systematic Review software (Veritas Health Innovation Ltd) for screening and removal of duplicates. First, titles and abstracts of uploaded studies were independently screened by 2 reviewers for the eligibility criteria (JH and CE). Any disagreements on the title and abstract screening were negotiated by accessing the full text to determine its appropriateness. In the second stage, full-text papers were further screened for the eligibility criteria. Any disagreements in all previous steps were extensively discussed and resolved by the reviewers (JH, CE) or, when they were unable to agree, by a third reviewer (JR, DK).

Data Extraction

All data were extracted independently by 2 reviewers (JH, CE for tailoring questions; JH, DK for design rationale questions, Quality Assessment for Diverse Studies [QuADS] by JH and CE).

General Characteristics

We extracted the first author, year of publication, country of origin, titles, and targeted illness or targeted behavior.

Health Literacy Concepts

When extracting information on health literacy, we focused on descriptions of both cognitive as well as social approaches to health literacy concepts in the intervention and extracted the reported details on these concepts.

Tailoring Strategies

Tailoring strategy types were categorized into Personalization, Feedback, and Content Matching, plus their corresponding subsets [10]. We also extracted information on the theoretical framework for content matching in the form of reported supporting theories and whether personal data of end-users was used for content matching.

Design Rationale

To establish the design rationale of the intervention, we looked at the intervention itself, its intended goals, its components, the design process, and the reasoning behind the choices made in that process.

We therefore categorized (1) the medium (website, app, email, or message service), (2) the hardware (desktop, mobile app, wearable, or tablet), (3) the components (eLearning, Game or Game mechanics, Diary logging, or Questionnaires). We further extracted (4) hypotheses concerning envisioned or desired workings of the intervention connected to its intended goals; (5) information on the design process (if any such process was reported and whether a design method was used); and (6) the reasoning related to the choices for the health literacy, tailoring, and design characteristics of the intervention. We used a classification of 3 types of reasoning: pragmatic (based on ease of accessibility or availability), theoretical (based on relevant scientific evidence in other contexts) and contextual (based on evidence from the end user's context) and (7) whether an underlying story was provided which ties together the different elements of the solution.

Quality and Bias Assessment

We only included randomized controlled trials, and while this strategy was initially intended to study efficacy, we did choose to include the study quality review. Quality and bias assessment was based on the criteria of QuADS [18] by 2 independent reviewers using Covidence. Any disagreements were resolved by discussion between the reviewers (JH and CE) or, when they were unable to agree, by a third reviewer (JR).

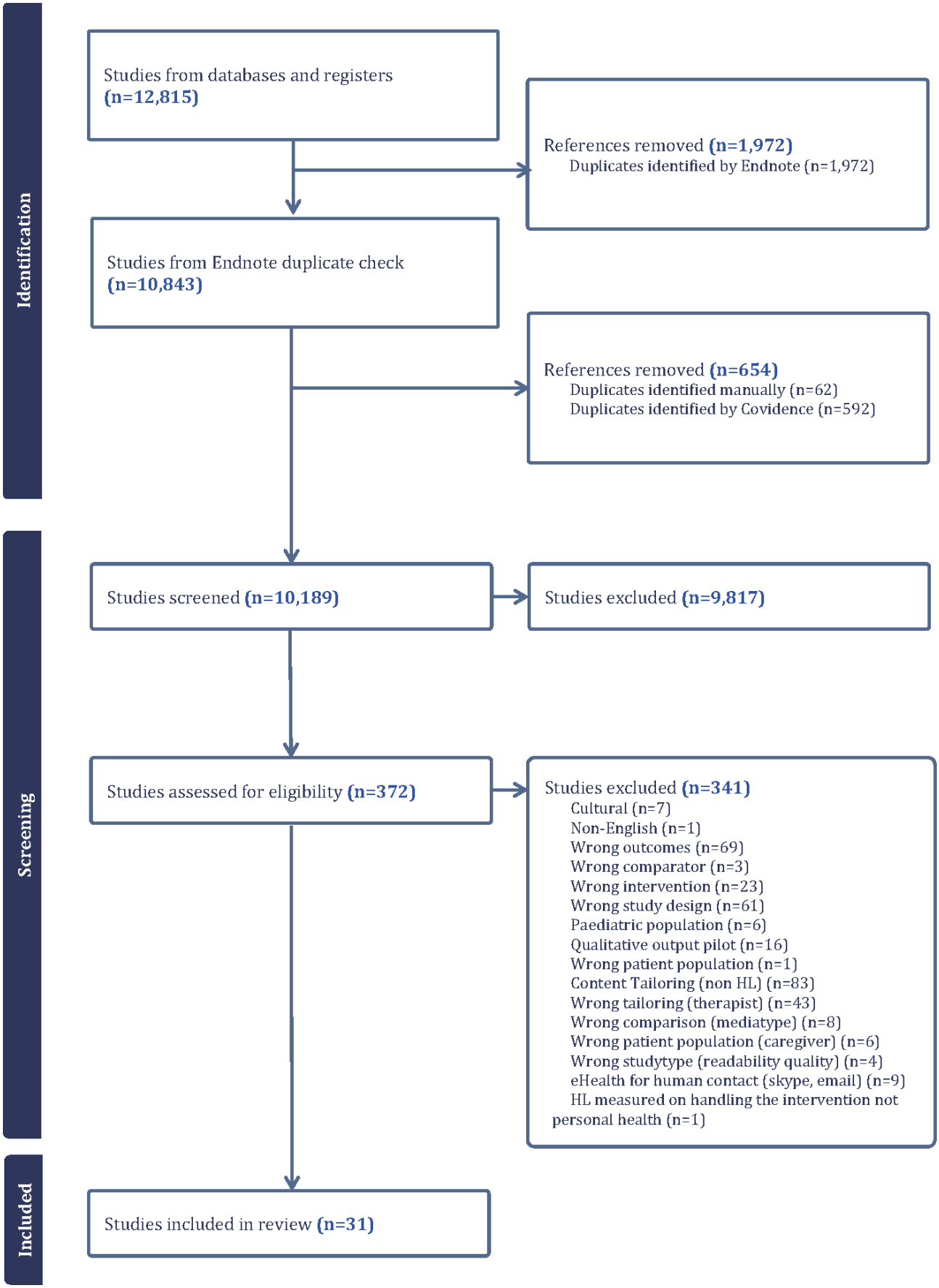
Results

Search Results

The search yielded 12,815 results (Pubmed: 3308; Psychinfo 1916; Web of Science 3188, ACM 4403). EndNote was used to discard duplicates for all databases, after which 10,843 references were imported into Covidence. Using Covidence, a further 592 duplicates were removed before screening. Eventually, the screening resulted in 31 studies that were eligible for data extraction and analysis (Figure 1).

XSL•FO
RenderX

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow chart.



XSL•FO
RenderX

General Characteristics

Of the 31 selected papers [19-49], 16 studies were from the United States [20,22,23,26-30,33,36,37,41-45], 4 from the Netherlands [25,31,39,40], 3 from Canada [21,24,32], 3 from Australia [19,47,49], and 1 each from New Zealand [38], Germany [34], Switzerland [34], and Hong Kong [48]; 1 study used an international sample [35].

Only 2 papers were published before 2010 [23,37]. Between 2010 and 2015, 11 papers were published [19,26,28,30,36,38,39,41,45-47], with 4 papers in 2015 alone [36,38,46,47]. Between 2016 and 2020, 17 papers were published [20-22,24,25,27,29,31-35,40,42-44,49], with 9 in 2020 alone [22,24,27,29,31,32,40,42,43]. For 2021, 1 paper was included [48], but this may be partly due to the extraction date.

Most of the interventions addressed a single clinical or behavioral topic, 1 addressed 2 topics [46], and 1 addressed 4 topics simultaneously [44].

Of the screened interventions, 52% (16) were targeted at specific illnesses, complaints, or symptoms, such as cancer (5) [19,27,33,45,47], diabetes (4) [22,34,43,46], cardiovascular diseases (4) [29,31,37,38], back pain (3) [28,36,46], menopausal symptoms (1) [23], and migraine (1) [26].

The remaining interventions (15) were focused on behavioral, preventative, or informative health-related activities such as food intake (3) [20,24,44], smoking (3) [25,44,49], physical activity (2) [40,44], STIs and prevention (2) [39,48], weight management (1) [44], or other preventative care (4) [30,35,41,48]. Two interventions informed end users on specific medical procedures, such as medicine adherence (1) [21], breast reconstruction (1) [42], and one intervention focused on coping strategies for domestic violence [32] (Table S1 in Multimedia Appendix 2).

Health Literacy Concepts

A total of 18 studies [19,23-33,36,39,42,43,45,48] described interventions that used both cognitive and social aspects of Health Literacy concepts. 6 studies [20,35,41,44,46,47] only contain interventions that take a cognitive approach to Health Literacy, while 7 studies [21,22,34,37,38,40,49] only aim to convey one or more social components of health literacy. Some of the health literacy concepts are used to explicitly convey knowledge or awareness on a certain health-related topic with very little or no behavioral cues [19,20], while others use a more practical approach and focus on the social component of health literacy by helping end users develop new behavioral patterns [21,22]. Interestingly, while all included interventions make use of health literacy related concepts, only a very few studies explicitly mention health literacy or a derivative such as food literacy [50], as a concept used to inform the design of the intervention (Table 1).

XSL•FO
RenderX

Table . Tailoring strategies and health literacy concepts.

Author	Tailoring strategy types	Supporting theory or theories	Personal data used for content matching	HL[a] cognitive	HL[a] social
Barnabei et al [23]	Content Matching, Feedback (Descriptive)	None reported	Yes	Explanation of Hormone therapy	Customized questions to ask the provider about HT[c]
Bomfim et al [24]	Content matching, personalization (identification, raising expectation, contextualization), feedback (descriptive, comparative, evaluative)	Self-Affirmation Theory, Self-Determination Theory, and Food Literacy Theory	Yes	Develop the knowledge and awareness of nutrition and food	"Make informed food decisions when grocery shopping."
Bommele et al [25]	Content matching, personalization (identification) feedback (descriptive, evaluative, comparative)	Self-Affirmation Theory and Self-Efficacy Theory	Yes	Emphasized the cons of smoking and the pros of quitting	Increase receptivity to antismoking information
Boyd et al [29]	Content matching, personalization (contextualization)	None reported	Yes	Patient narratives about the importance of DAPT	Helping the patient plan and overcome the common barriers to medication adherence
Bromberg et al [26]	Content matching, feedback (evaluative, comparative), personalization (identification, contextualization)	Cognitive Behavior Therapy	Yes	Migraine-specific knowledge, medication safety	Migraine self-management skills, emotional coping, communication skills
Burgermaster et al [20]	Content matching, feedback (descriptive, evaluative, comparative)	Health Literacy, Nutrition Literacy, and Social Cognitive Theory	No	Vicarious learning about caloric contents of food	None reported
Carter-Harris et al [27]	Content matching, personalization (contextualization)	Conceptual Model on Lung Cancer Screening Participation: Health Belief Model Precaution Adoption Process Model	Yes	Lung health knowledge, knowledge of risk factors for lung cancer, knowledge of option of lung cancer screening; knowledge of risks and benefits of lung cancer screening	Tailored summary that individuals can use to guide a discussion with their clinician
Chiauzzi et al [28]	Content matching, personalization (contextualization), feedback	Cognitive Behavior Therapy	Yes	Motivational enhancement	Collaborative decision making with health professionals 2) CBT to improve self-efficacy, manage thoughts and mood, set clinical goals, work on problem-solving life situations, and prevent pain relapses; 3. wellness activities to enhance good sleep, nutrition, stress management, and exercise practices.

XSL•FO
RenderX

Author	Tailoring strategy types	Supporting theory or theories	Personal data used for content matching	HL[a] cognitive	HL[a] social
Côté et al [21]	Content matching, feedback (comparative), personalization (identification)	Social Learning Theory, Behavior Change techniques	No	None reported	Reinforcing self-management skills required for medication intake. The sessions aimed to help users incorporate the therapeutic regimen in their daily routine, cope with medication side effects, handle situations or circumstances that could interfere with medication intake, interact with health care professionals, and mobilize social support. The learning objectives included strengthening various capacities such as self-motivation and self-monitoring
Dingle and Carter [49]	Content matching, feedback (comparative)	Elaborated Intrusion Theory of Desire for Smoking	No	None reported	Psychoeducation about smoking and its triggers, particularly emotional states and cravings Learning how to use music to mediate negative affective triggers that may lead to smoking behavior
Drieling et al [30]	Content matching, personalization (contextualization)	SMCR communication model, Social Cognitive Theory, Transtheoretical Model	Yes	Education on risk factors, knowledge, and attitudes concerning osteoporosis	Tutorials for healthy behavior concerning osteoporosis
Engelen et al [31]	Content matching, personalization, feedback	I-change Model	Yes	Knowledge, awareness, Intention, Attitude, Self-Efficacy CVD, Coping CVD, Boundaries in daily life CVD, Lifestyle CVD, Healthy nutrition CVD, Physical Activity CVD, Interaction with Health Professionals	Habits, skills, Norm, Intention, Attitude, Self-Efficacy CVD: Coping CVD: Boundaries in daily life CVD: Lifestyle CVD: Healthy nutrition CVD: PA CVD: Interaction with Health Professionals
Flight et al [19]	Content matching, personalization (identification, contextualization)	Preventive Health Model, Precaution Adoption Process Model	Yes	Educational content with “generalized risk information (ie,>50 years, certain bowel conditions, and family history”	Messages tailored to an individual user’s decision stage for screening and responses to PHM, self-efficacy, and fecal aversion variables
Ford-Gilboe et al [32]	Content matching, personalization (identification, raising expectation, contextualization), feedback (comparative, evaluative)	Danger Assessment Calendar	Yes	Increase their awareness of safety risks	Reflecting on (.) plans for their relationships and priorities, a personalized detailed action plan of strategies and resources for addressing their safety and health concerns
Fowler et al [33]	Content matching, personalization (contextualization), feedback (descriptive, comparative, evaluative)	None reported	Yes	Informing participants of the extent to which behavior change could reduce their risk.	“It targets self-efficacy by suggesting specific strategies for changing behavior.”

XSL•FO
RenderX

Author	Tailoring strategy types	Supporting theory or theories	Personal data used for content matching	HL[a] cognitive	HL[a] social
Gimbel et al [22]	Content matching, personalization (contextualization), feedback (descriptive, evaluative)	Patient Activation Measure Summary of Diabetes Self-Care Activities	Yes	None reported	Behavioral messaging concerning nutrition, home monitoring, physical activity, blood pressure, foot care, medications, smoking, glucose control, and general behavioral reinforcement.
Höchsmann et al [34]	Content matching, feedback (descriptive, evaluative)	Self-Determination Theory	Yes	None reported	"In the game, regular PA is rewarded with water or building materials that are needed to restore the garden and proceed in the storyline. In-game workouts consist of 130 variations of strength, endurance, balance, and flexibility exercises whose execution, as well as daily PA, is tracked via the phone's sensors (camera, accelerometer, and gyroscope)."
Hopkin et al [35]	Content matching, feedback (evaluative)	Subjective Expected Utility Theory	Yes	Knowledge of benefit and harm outcome of statins according to their own preferences.	None reported
Irvine et al [36]	Content matching, personalization (identification, contextualization), feedback (descriptive, evaluative)	Social Cognitive Theory, Theory of Planned Behavior	Yes	General aspects of pain and pain management	Pain self-care activities in 4 categories (rest and relief, mindfulness, general fitness, and back pain-specific stretching and strength exercises), cognitive and behavioral strategies to manage and prevent pain (eg, controlling fear of pain, mindfulness and relaxation, use of heat and ice, over-the-counter medications, benefits of staying active), and instructional videos on specific strength and stretching exercises tailored by job type (sitter, stander, driver, lifter)
Kukafka et al [37]	Content matching, feedback	Self-Efficacy Change Theory	Yes	None reported	"Seeking help in responding to AMI symptoms specific to each dimension of self-efficacy scale"
Maddison et al [38]	Feedback (descriptive), personalization (contextualization)	None reported	No	None reported	Regular exercise prescription, provision of behavior change strategies

XSL•FO
RenderX

Author	Tailoring strategy types	Supporting theory or theories	Personal data used for content matching	HL[a] cognitive	HL[a] social
Mevissen et al [39]	Content matching, personalization (identification, contextualization), feedback (evaluative)	AIDS Risk Reduction Model, Extended Parallel Process Model, Motivational Interviewing	Yes	Addressing perceived risk of STI[b] infections within established relationships; Targeting normative beliefs associated with condom use and STI testing in established relationships	self-efficacy and skills related to condom use and STI testing
Middelweerd et al [40]	Content matching, personalization (contextualization), feedback (descriptive, comparative, evaluative)	Health Behavior Theory	Yes	None reported	Coaching on sports participation, taking the stairs, or active transport.
Milan and White [41]	Content matching	Transtheoretical Model	Yes	“Five educational modules, each corresponding to the five stages of readiness for meeting folic acid multivitamin recommendations.” - “Each module consisted of 4 Web pages, designed to be read over the course of 4 weeks in a 5- to 10-minute block of time”	None reported
Politi et al [42]	Feedback (Evaluative)	None reported	No	breast reconstruction knowledge	Asking the right questions during a consultation
Sittig et al [43]	Content matching, personalization (contextualization), feedback (evaluative)	Social Cognitive Theory, Fogg Behavior Model, Persuasive Technology	No	Knowledge on type 2 diabetes	Self-efficacy, Self-care for type 2 diabetes, goal setting
Valle et al [44]	Content matching, personalization (identification, raising expectation), feedback (comparative)	Social Cognitive Theory, Theory of Planned Behavior	Yes	Feedback messages on health risk assessment	None reported
Vernon et al [45]	Content matching, personalization (identification), feedback (evaluative)	Transtheoretical Model	Yes	Knowledge about CRCS, CRC risk perception, Awareness of CRC as a serious problem	Self-efficacy, Goal Setting, Getting CRCS
Weymann et al [46]	Content matching, personalization	Summary of Diabetes Self-Care Activities Measure, Avoidance Endurance Model	Yes	Presenting and testing information on T2D and CLBP, self-management education	None reported
Wilson et al [47]	Content matching	Preventive Health Model	Yes	Information sheet; tailored messages; consumer information booklet	None reported
Wong et al [48]	Content matching, feedback (evaluative)	Health Literacy, Health Belief Model, Harvard Cancer Risk Index, Continuum of Conflict and Control Theory	Yes	Disseminate knowledge about STIs, cervical cancer, and condom use.	Take action regarding condom use, communication and negotiation about condom use, and sexual coercion in daily life

[a]HL: health literacy

[b]STI: Sexually Transmitted Infection

[c]Hormone Therapy

XSL•FO
RenderX

Tailoring Strategies

Content Matching (29 studies, [19-37,39-41,43-49]) was found in almost all studies, followed by Feedback (24 studies, [20-26,28,31-40,42-45,48,49]) and Personalization (21 studies, [19,21,22,24-33,36,38-40,43-46]); this includes 15 studies using all 3 tailoring strategies [21,22,24-26,28,31-33,36,39,40,43-45]. In one study, we also could discern all subsets [24]. 13 studies used 2 strategies [19,20,23,27,29,30,34,35,37,38,46,48,49] and 3 used only one [41,42,47]. Content Matching was often guided by personal data (25 studies, [23-37,39-41,44-48]) and by theoretical frameworks (28 studies [19-22,24-28,30-32,34-37,39-41,43-49]); only 2 studies [20,48] explicitly mention Health Literacy as a guiding principle for tailoring, while 2 name Food Literacy as a supporting theory [20,24]. Of all interventions, 13 used one supporting theory, 7 used 2 theories, 5 used 3 theories, and one intervention used 4 supporting theories. 5 interventions did not report any use of supporting theories. Of the interventions that did not use any supporting theories, 3 relied on personal data from end users for content matching. Apart from 7 theories that were reported twice, the most prevalent supporting theories were Social Cognitive Theory (5) and the Transtheoretical Model (3) (Table 2).

XSL•FO
RenderX

Table . Design rationale and process.

Author	Rationale for HL[a] concepts	Rationale for tailoring strategy	Rationale for technology or modules	Solid overarching reasoning for intervention	Design process	Development or design method
Barnabei et al [23]	None reported	Theoretical	Theoretical, Contextual	None reported	None reported	None reported
Bomfim et al [24]	Theoretical, Contextual	Theoretical, Contextual	Theoretical, Contextual	Reported[b]	External	None reported
Bommele et al [25]	Theoretical	Contextual	Theoretical	None reported	None reported	Intervention Mapping
Boyd et al [29]	Contextual, Theoretical	Contextual	Contextual, Theoretical	None reported		
Bromberg et al [26]	Theoretical	Theoretical	Theoretical	None reported	None reported	The self-management education model by Lorig and colleagues at Stanford University
Burgermaster et al [20]	Theoretical	Theoretical	Theoretical	Reported[d]	None reported	None reported
Carter-Harris et al [27]	Contextual	Contextual	Pragmatic	None reported	External	USPSTF[c] Lung Cancer Screening Guidelines, International Patient Decision Aid Standards
Chiauzzi et al [28]	Theoretical	Contextual	Theoretical, Contextual	None reported	External	None reported
Côté et al [21]	Theoretical	None	Theoretical, Pragmatic	Reported[e]	External	None reported
Dingle and Carter [49]	Theoretical	Theoretical	Theoretical, Pragmatic	None reported	None reported	None reported
Drieling et al [30]	Theoretical, Contextual	Theoretical	Theoretical, Pragmatic	None reported	None reported	None reported
Engelen et al [31]	Theoretical, Contextual	Contextual	Theoretical, Contextual	None reported	External	Intervention Mapping
Flight et al [19]	Theoretical	Theoretical	Theoretical	None reported	None reported	None reported
Ford-Gilboe et al [32]	Theoretical	Theoretical	Pragmatical, Theoretical	None reported	None reported	None reported
Fowler et al [33]	Theoretical	Theoretical	Theoretical, Pragmatic	None reported	External	None reported
Gimbel et al [22]	Theoretical, Contextual	Theoretical	Theoretical, Contextual	None reported	Reported	“User-Centered Design”
Höchsmann et al [34]	Contextual, Theoretical	Contextual, Theoretical	Contextual, Theoretical	Reported[f]	External	None reported
Hopkin et al [35]	Theoretical	Theoretical	Theoretical	None reported	None reported	None reported
Irvine et al [36]	Theoretical	Theoretical, Contextual	Pragmatic	None reported	None reported	None reported
Kukafka et al [37]	Theoretical	Theoretical	Pragmatic	None reported	External	None reported
Maddison et al [38]	Theoretical	Pragmatic	Pragmatic, Theoretical	None reported	None reported	mHealth Development and Evaluation Framework
Mevissen et al [39]	Contextual, Theoretical	Theoretical, Contextual	Theoretical	None reported	None reported	Intervention Mapping
Middelweerd et al [40]	Theoretical	Theoretical	Pragmatic, Theoretical	None reported	External	“A systematic and stepwise approach”

XSL•FO
RenderX

Author	Rationale for HL[a] concepts	Rationale for tailoring strategy	Rationale for technology or modules	Solid overarching reasoning for intervention	Design process	Development or design method
Milan and White [41]	Theoretical	Theoretical	None reported	None reported	None reported	None reported
Politi et al [42]	Theoretical, Contextual	Theoretical	None reported	None reported	Reported	International Patient Decision Aids Standards
Sittig et al [43]	Theoretical, Contextual	Theoretical, Contextual	Pragmatic, Theoretical	None reported	Reported	Interactive Health Communication Applications framework, semi structured focus group question model
Valle et al [44]	Theoretical	Theoretical, Contextual	Theoretical	None reported	None reported	None reported
Vernon et al [45]	Theoretical	Theoretical	Pragmatic	None reported	None reported	Intervention Mapping
Weymann et al [46]	Theoretical	Theoretical	Theoretical	None reported	Reported	None reported
Wilson et al [47]	Theoretical	Theoretical	Theoretical	None reported	None reported	None reported
Wong et al [48]	Theoretical, Contextual	Theoretical	Contextual, Theoretical, Pragmatic	None reported	None reported	None reported

[a]HL: health literacy

[b]"Our gameful app, Pirate Bri's Grocery Adventure (PBGA), incorporates gameful design elements, such as challenges, personalization, and meaningful choices to motivate young adults to develop food literacy, increase awareness, and improve choices at the grocery store."

[c]United States Preventive Services Task Force

[d]"Taken together, these theoretical foundations point to the potential for casual observational learning to occur when people are intrinsically motivated to participate in learning activities to improve nutrition literacy on a social computing platform."

[e]"This program taught participants to identify which emotions were triggers for their smoking, and how music can be used as a substitute for the emotion regulating effect of smoking"

[f]"The game uses the self-determination theory as the theoretical framework. The self-determination theory is a widely researched theory of motivation that addresses both intrinsic and extrinsic motives for acting and has shown its utility in explaining processes that underpin exercise behavior as well as motivation to play video games. The goal of the game is to restore a decayed garden by planting trees and flowers. In doing so, the player attracts animals that used to live in the garden to come back and help with the restoration process. At the same time, the Schweinehund, the game's adversary, is kept in check. In German, "innerer Schweinehund" (inner swine hound) refers to the weak or lazy part of one's nature, often regarding PA, that has to be overcome to get one's self going. The garden setting was deliberately chosen, as its restoration stands metaphorically for the restoration of the player's body through regular PA. In addition, it has been shown that gardening is among the target group's preferred forms of PA and that gardening-themed games are quite popular and comprehensible to a wide range of players because of their straightforward mechanics."

Design Rationale

Most interventions used a website (24, [19-21,23,24,27,28,30-33,35-39,41,42,44-49]), while 7 interventions were described as an app or native application [22,24,25,29,34,40,43], 28 studies describe interventions using eLearning [19-21,23-34,36-39,41-49], 25 used a form of automated feedback [19-25,27,28,30-35,37-40,42-46,48] while 24 incorporated one or more questionnaires [19,20,23,25-37,39,41-46,48]. 11 interventions used diaries or logging [22,24,26,29,31,34,36,38,40,43,49], 5 interventions made use of a virtual agent [21,25,36,39,44], and 4 interventions used serious games or game mechanics [24,34,40,43] (Table S1 in Multimedia Appendix 3).

A clear trend in all studies is the use of eLearning. 28 studies report the use of this component to convey health-related information [19-21,23-34,36-39,41-49]. Out of the 3 studies that did not use eLearning [22,35,40], 2 did not use cognitive health literacy concepts either [22,40]. These studies relied heavily on providing automated feedback on end-user reporting and behavioral data, thus aiming to reinforce desired health-related behavior [22,40].

All studies (31) explicitly described envisioned or intended workings of the intervention by connecting the functional description of the intervention to its intended goals. Only 4 studies provided specifics concerning the design process of the intervention in the publication [22,42,43,46] while 10 studies referred to other publications for details on the design process [21,24,27-29,31,33,34,37,40]. The majority, however, did not report any details on the design process of the intervention (17). 12 studies mentioned a design method [22,25-27,29,31,38,40,42,43,45], of which 4 [25,31,39,45] referred to Intervention Mapping [51] and 2 pointed towards the International Patient Decision Aids Standards [52] as a framework for designing the intervention. Other examples of methods named were the mHealth Development and Evaluation Framework [53] and "User-Centered Design."

XSL•FO
RenderX

In 4 studies, we could pinpoint an "underlying story" that explicitly tied the elements of the solution together [20,21,24,34]. These "stories," however, varied significantly in detail and clarity. For example, Höchsmann et al [34] extensively explain how and why the choice of medium and components contribute to the transfer of health literacy concepts to the target group by using the metaphor used in the intervention as the underlying story—in this case the maintenance of a garden as a metaphor for living healthy—that ties the elements of the intervention together. Burgermaster et al [19-37,39-41,43-49], on the other hand, is much more succinct: the authors tie the intervention's goal, the medium, and the health literacy content together in one sentence from a theoretical starting point.

When it comes to reasoning, a clear trend is noticeable. The vast majority of reasoning types that were used was theoretical: reasoning based on earlier research in other contexts than the research or intervention's end user context. All but one intervention used theoretical reasoning for one or more components of the interventions. 16 of the 31 interventions use contextual reasoning to support choices for one or more of the components of the intervention [22-25,27-31,34,36,39,42-44,48] (Table 2).

Quality and Bias Assessment

The results of the QuADS assessments can be found in Table S1 in Multimedia Appendix 4. Of the included studies in this systematic review, 26 studies (79%) were of high quality (scoring above 66%), and 7 (21%) were of moderate quality (scoring between 33% and 66%). No low-quality studies were found. The lowest score was 51%. Main areas for quality improvements were: (1) strengths and limitations critically discussed; and (2) appropriate sampling to address the research aim.

Discussion

Principal Findings

The goal of this review was threefold: to provide an overview of the health literacy concepts and tailoring strategies used in eHealth interventions that aim to improve health and to determine how these elements contributed to the design choices of the interventions. In all selected studies, health literacy concepts are used to inform the design of the interventions. The types of concepts and their application vary significantly, as does the choice of tailoring strategies in these interventions. Content matching is present in almost all studies, but a clear trend in applied design strategies or patterns for tailoring strategies to inform the design of the interventions cannot be distinguished.

Concerning the overview of the health literacy concepts, in all extracted papers, we can identify the use of health literacy concepts as guiding principles to aid the development of the inner workings and content of the interventions. Most eHealth interventions used both cognitive and social health literacy concepts. However, there is a broad variety in how these concepts are described, used, and translated into design choices for the intervention and how these concepts are tailored towards the end users. Considering the differences in contexts and targeted illnesses and behaviors, this is not entirely surprising, but there are no clear patterns to be found in approaching health literacy on the whole.

As for the tailoring strategies, content matching can be found in all but 2 studies [19-37,39-41,43-48] while personalization and feedback are used significantly less. Personalization and feedback have been described as mechanisms aimed at "message processing" and "self-referencing," hinting at "creating preconditions for message processing" [9], while content tailoring is described as "the crux" of tailoring. This may explain why personalization and feedback are less common and why only 15 studies show the use of all tailoring strategies [21,22,24-26,28,31-33,36,39,40,43-45].

Only 2 studies [19,44] explicitly refer to the reporting recommendations for tailored interventions [10], in which these strategies are proposed. The way in which content matching is applied varies.

A minority of the studies reported a design process and even fewer reported a methodology for the design process of the intervention. Only 4 studies described a design process in the paper itself [22,42,43,46], although in most studies, rationales for choices for the main components of the interventions were provided. Interestingly, 2 of these interventions involved serious games [24,34]. This may be partly explained by the fact that serious games in the context of health literacy are a relatively novel approach to eHealth interventions, provoking more extensive reporting on how the product was conceived. Also, it is quite common that serious games are designed around a metaphor, which increases the chances of an overarching story that ties together the elements of the solution [54].

The main type of rationale provided for the choice of health literacy concepts, tailoring strategy, and technology was theoretical. Although a qualitative assessment on the rationales is difficult to make, choices for technology and modules would require some contextual reasoning: reasoning based on needs, affordances, and preferences of an end user's context [55]. In half of the studied interventions in this review, such reasoning was not reported.

As all extracted studies aimed to determine the intervention efficacy, the outcomes and conclusions of these studies are focused on answering the question whether the intervention works as intended and expected. However, for insights into why these interventions work and how the combined workings of their components contribute to efficacy, it is important to know what specific reasoning lies behind the choices for concepts, tailoring strategy, and choice of technology. Reporting on these aspects is thus important for the development and improvement of new and existing interventions.

However, we did not find any indication of formal methods such as Toulmin [56] or IBIS (Issue-Based Information Systems) [57], or use of the guidelines for reporting on tailored interventions [10]. While some of the studies were conducted before the latter guidelines were made, more recent papers also did not completely adhere to the provided recommendations. Moreover, some of these recommendations can be interpreted

XSL•FO
RenderX

as a call for a summarized design rationale of the intervention, rather than just a description of the study design. On the other hand, the recommendations do not move beyond describing how supporting theories translate into tailored messages. From an eHealth perspective, a reasoned choice of technology and technological components would certainly add to these recommendations.

eHealth interventions bring together both evidence-based health practices and human-computer interaction, and thus 2 highly different paradigms for design, development, and evaluation. For example, in 4 studies, intervention mapping was named as a design process. While this method does offer a well-documented evidence-based approach, it was not specifically developed for creating digital interventions and may not offer the most suitable procedural approach when designing digital interventions.

Literature calls for a deeper understanding of and appreciation for the differences in development methods to make more effective eHealth interventions and provide more useful documentation on its research [58]. Health intervention development focuses on choosing proven "mechanisms of action" or "active ingredients" [59]. Human-computer interaction, on the other hand, has a deeply rooted tradition in ethnological research and iterative design cycles where anticipated end-users are actively participating in the complete design process [55]. We have found little evidence of these types of design practices in the reporting.

Strengths and Limitations of This Review

This review is the first review on this topic. A strength of our approach is that it was multidisciplinary, involving researchers and databases from the fields of health and technology. Incorporating ACM as an extra source of data widened the scope of this review significantly. It also revealed that multiarmed randomized trials are far less common in this database, and eHealth evaluations tend to be more qualitative and iterative in smaller samples. In addition, we followed the PRISMA guidelines for systematic studies.

When interpreting the results of the review, the following limitations need to be considered. Because we initially planned to report on study health outcomes in relation to tailoring, we only included interventions that were used in a randomized trial. This sometimes resulted in the exclusion of interesting studies, such as research protocols, (speculative) design studies, and single-armed efficacy tests or pilot studies. However, limiting the included studies to multiarmed trials also entailed a certain level of product maturity, compared with more speculative or single-arm pilot studies. This is probably one reason why the quality of the included studies was moderate to high.

In addition, we initially aimed to study the comparative effectiveness of different tailoring strategies to optimize intervention outcomes; however, the degree of variation among the selected studies made this infeasible. A comparative effectiveness study is recommended for future research.

Another limitation is that we were dependent on the quality and focus of reporting. If a design rationale is not described, this does not necessarily mean that it was not present or part of the development process. The constraints of this review limited us from studying other publications linked to the intervention, for instance, study protocols or design studies where more information on the intervention's design rationale could have been present. In any case, our study suggests that design rationales for digital interventions are often not reported or even summarized in efficacy studies, implying that there are limited possibilities to be informed or inspired by approaches and choices of other researchers published in the scientific literature.

The quality of included studies in this review was moderate to high. For future studies, quality can be improved by improved samples selection and by a more critical discussion of strengths and limitations. Of course, QuADS focuses on the study design in general and whether useful conclusions can be drawn from the outcomes, not so much on whether the intervention itself was well designed or if the choices made in the design process are justifiable. As discussed, there may be opportunities to explicate standards on reporting that include (summaries of) design rationales for eHealth interventions.

Conclusion

In conclusion, we can see a broad variety of approaches to the use of health literacy concepts and tailoring strategies in tailored eHealth interventions. We identified rationales for choices of components in all studies, and in a small minority, we could discern overarching design rationales. Although we have been able to discern health literacy concepts, tailoring strategies, and design rationales in the interventions, only a few studies provide sufficient insight into the inner workings and overarching rationale behind the intervention. A significant number of studies refers to other sources for details on the design and design process of the intervention, and while this strategy of reporting on the origins of an intervention certainly is preferable over no reporting at all, it does provide an extra barrier in linking intervention design choices to observed outcomes. Indeed, existing frameworks already provide opportunities to consistently report on the design of tailored health interventions and could be used to report on the use of cognitive and social health literacy concepts as well.

In the field of human-computer interaction, extensive reporting on the design process, as well as describing an overarching rationale for a product, is far more common. Literature has already called for a stricter reporting standard on tailored interventions [10], and since these recommendations are not specifically aimed at digital interventions, there is room to iterate on these recommendations with digital health in mind. Indeed, there are clear opportunities to explore novel reporting frameworks that incorporate the key design rationales of eHealth interventions, in addition to their efficacy. If such frameworks are used regularly, it will be feasible to study comparative effectiveness of eHealth interventions with design rationales informed by specific tailoring strategies. Future studies should therefore prioritize comprehensive documentation on the design of the intervention to facilitate replicability and understanding of intervention efficacy.

XSL•FO
RenderX

Authors' Contributions

Conceptualization: JH (lead), CE (equal), JR (supporting), GL (supporting), YR (supporting) Data curation: JH (lead), CE (equal), JR (supporting), DH (supporting), DK (supporting), GL (supporting) Formal analysis: JH (lead), JR (supporting), DH (supporting), DK (supporting) Methodology: JH (lead), CE (equal), JR (supporting), DH (supporting), DK (supporting), GL (supporting), YR (supporting) Project administration: JH (lead), CE (equal), JR (supporting) Resources: JR, DH, GL, DK, YR Supervision: JR, DH, GL, DK Validation: JR, DH, GL, DK Writing – original draft: JH (lead), CE (supporting), JR (supporting), DH (supporting), DK (supporting), GL (supporting) Writing – review & editing: JH (lead), JR (supporting), DH (supporting), DK (supporting), GL (supporting)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Table S1: Search strategy per database.
[DOCX File, 15 KB - humanfactors_v12i1e76172_app1.docx]

Multimedia Appendix 2
Table S1: general characteristics.
[DOCX File, 26 KB - humanfactors_v12i1e76172_app2.docx]

Multimedia Appendix 3
Table S1: intervention descriptions and components.
[DOCX File, 26 KB - humanfactors_v12i1e76172_app3.docx]

Multimedia Appendix 4
Table S1: QuADS scoring results.
[DOCX File, 17 KB - humanfactors_v12i1e76172_app4.docx]

Checklist 1
Preferred Reporting Items for Systematic Reviews and Meta-Analyses 2020 checklist as requested.
[PDF File, 156 KB - humanfactors_v12i1e76172_app5.pdf]

References

1. Marques ICP, Ferreira JJM. Digital transformation in the area of health: systematic review of 45 years of evolution. Health Technol 2020 May;10(3):575-586. [doi: 10.1007/s12553-019-00402-8]
2. Bennett CB, Ruggero CJ, Sever AC, Yanouri L. eHealth to redress psychotherapy access barriers both new and old: A review of reviews and meta-analyses. J Psychother Integr 2020;30(2):188-207. [doi: 10.1037/int0000217]
3. Kraaijkamp JJM, van Dam van Isselt EF, Persoon A, Versluis A, Chavannes NH, Achterberg WP. eHealth in geriatric rehabilitation: systematic review of effectiveness, feasibility, and usability. J Med Internet Res 2021 Aug 19;23(8):e24015. [doi: 10.2196/24015] [Medline: 34420918]
4. Consortium H-E. Comparative report of health literacy in eight EU member states. : The European Health Literacy Survey (HLS-EU); 2012 URL: http://cpme.dyndns.org:591/adopted/2015/Comparative_report_on_health_literacy_in_eight_EU_member_states.pdf [accessed 2025-08-26]
5. Berkman ND, Sheridan SL, Donahue KE, et al. Health literacy interventions and outcomes: an updated systematic review. Evid Rep Technol Assess (Full Rep) 2011 Mar(199):1-941. [Medline: 23126607]
6. Sørensen K, Van den Broucke S, Fullam J, et al. Health literacy and public health: a systematic review and integration of definitions and models. BMC Public Health 2012 Jan 25;12:80. [doi: 10.1186/1471-2458-12-80] [Medline: 22276600]
7. Hibbard J. Patient activation and health literacy: what's the difference? How do each contribute to health outcomes. Stud Health Technol Inform 2017;240:251-262. [Medline: 28972522]
8. Rademakers J, Heijmans M. Beyond reading and understanding: health literacy as the capacity to Act. Int J Environ Res Public Health 2018 Aug 7;15(8):1676. [doi: 10.3390/ijerph15081676] [Medline: 30087254]
9. Hawkins RP, Kreuter M, Resnicow K, Fishbein M, Dijkstra A. Understanding tailoring in communicating about health. Health Educ Res 2008 Jun;23(3):454-466. [doi: 10.1093/her/cyn004] [Medline: 18349033]
10. Harrington NG, Noar SM. Reporting standards for studies of tailored interventions. Health Educ Res 2012 Apr;27(2):331-342. [doi: 10.1093/her/cyr108] [Medline: 22156230]

XSL•FO
RenderX

11. Dijkstra A, De Vries H, Roijackers J. Computerized tailored feedback to change cognitive determinants of smoking: a Dutch field experiment. Health Educ Res 1998 Jun;13(2):197-206. [doi: 10.1093/her/13.2.197] [Medline: 10181018]
12. Shum SB, Hammond N. Argumentation-based design rationale: what use at what cost? Int J Hum Comput Stud 1994 Apr;40(4):603-652. [doi: 10.1006/ijhc.1994.1029]
13. Knudsen LS, Haase LM, Goncalves MG. Design rationale in conceptual design: a longitudinal study of professional design teams' practice. 2020 May Presented at: Proc Des Soc: Des Conf p. 1315-1324. [doi: 10.1017/dsd.2020.11]
14. Jacobs RJ, Lou JQ, Ownby RL, Caballero J. A systematic review of eHealth interventions to improve health literacy. Health Informatics J 2016 Jun;22(2):81-98. [doi: 10.1177/1460458214534092] [Medline: 24916567]
15. Aida A, Svensson T, Svensson AK, Chung UI, Yamauchi T. eHealth delivery of educational content using selected visual methods to improve health literacy on lifestyle-related diseases: literature review. JMIR Mhealth Uhealth 2020 Dec 9;8(12):e18316. [doi: 10.2196/18316] [Medline: 33295296]
16. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015 Jan 1;4(1):1. [doi: 10.1186/2046-4053-4-1] [Medline: 25554246]
17. Ryan K, Dockray S, Linehan C. A systematic review of tailored eHealth interventions for weight loss. Digit Health 2019;5:2055207619826685. [doi: 10.1177/2055207619826685] [Medline: 30783535]
18. Harrison R, Jones B, Gardner P, Lawton R. Quality assessment with diverse studies (QuADS): an appraisal tool for methodological and reporting quality in systematic reviews of mixed- or multi-method studies. BMC Health Serv Res 2021 Feb 15;21(1):144. [doi: 10.1186/s12913-021-06122-y] [Medline: 33588842]
19. Flight IH, Wilson CJ, Zajac IT, Hart E, McGillivray JA. Decision support and the effectiveness of web-based delivery and information tailoring for bowel cancer screening: an exploratory study. JMIR Res Protoc 2012 Sep 26;1(2):e12. [doi: 10.2196/resprot.2135] [Medline: 23611950]
20. Burgermaster M, Gajos KZ, Davidson P, Mamykina L. The role of explanations in casual observational learning about nutrition. 2017 May 2 Presented at: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems p. 4097-4145. [doi: 10.1145/3025453.3025874]
21. Côté J, Fortin MC, Auger P, et al. Web-based tailored intervention to support optimal medication adherence among kidney transplant recipients: pilot parallel-group randomized controlled trial. JMIR Form Res 2018 Jul 19;2(2):e14. [doi: 10.2196/formative.9707] [Medline: 30684400]
22. Gimbel RW, Rennert LM, Crawford P, et al. Enhancing patient activation and self-management activities in patients with type 2 diabetes using the US department of defense mobile health care environment: feasibility study. J Med Internet Res 2020 May 26;22(5):e17968. [doi: 10.2196/17968] [Medline: 32329438]
23. Barnabei VM, O'Connor JJ, Nimphius NM, Vierkant RA, Eaker ED. The effects of a web-based tool on patient-provider communication and satisfaction with hormone therapy: a randomized evaluation. J Womens Health (Larchmt) 2008;17(1):147-158. [doi: 10.1089/jwh.2007.0369] [Medline: 18240991]
24. Bomfim MCC, Kirkpatrick SI, Nacke LE, Wallace JR. Food literacy while shopping: motivating informed food purchasing behaviour with a situated gameful app. 2020 Apr 21 Presented at: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems p. 1-13. [doi: 10.1145/3313831.3376801]
25. Bommelé J, Schoenmakers TM, Kleinjan M, Peters GJY, Dijkstra A, van de Mheen D. Targeting hardcore smokers: The effects of an online tailored intervention, based on motivational interviewing techniques. Br J Health Psychol 2017 Sep;22(3):644-660 [FREE Full text] [doi: 10.1111/bjhp.12256] [Medline: 28703470]
26. Bromberg J, Wood ME, Black RA, Surette DA, Zacharoff KL, Chiauzzi EJ. A randomized trial of a web-based intervention to improve migraine self-management and coping. Headache 2012 Feb;52(2):244-261 [FREE Full text] [doi: 10.1111/j.1526-4610.2011.02031.x] [Medline: 22413151]
27. Carter-Harris L, Comer RS, Slaven Ii JE, et al. Computer-tailored decision support tool for lung cancer screening: community-based pilot randomized controlled trial. J Med Internet Res 2020 Nov 3;22(11):e17050. [doi: 10.2196/17050] [Medline: 33141096]
28. Chiauzzi E, Pujol LA, Wood M, et al. painACTION-back pain: a self-management website for people with chronic back pain. Pain Med 2010 Jul;11(7):1044-1058. [doi: 10.1111/j.1526-4637.2010.00879.x] [Medline: 20545873]
29. Boyd AD, Ndukwe CI, Dileep A, et al. Elderly medication adherence intervention using the my interventional drug-euting stent educational app: multisite randomized feasibility trial. JMIR Mhealth Uhealth 2020 Jun 24;8(6):e15900. [doi: 10.2196/15900] [Medline: 32579120]
30. Drieling RL, Ma J, Thiyagarajan S, Stafford RS. An internet-based osteoporotic fracture risk program: effect on knowledge, attitudes, and behaviors. J Womens Health (Larchmt) 2011 Dec;20(12):1895-1907. [doi: 10.1089/jwh.2010.2515] [Medline: 21970565]
31. Engelen MM, van Dulmen S, Puijk-Hekman S, et al. Evaluation of a web-based self-management program for patients with cardiovascular disease: explorative randomized controlled trial. J Med Internet Res 2020 Jul 24;22(7):e17422. [doi: 10.2196/17422] [Medline: 32706708]
32. Ford-Gilboe M, Varcoe C, Scott-Storey K, et al. Longitudinal impacts of an online safety and health intervention for women experiencing intimate partner violence: randomized controlled trial. BMC Public Health 2020 Feb 26;20(1):260. [doi: 10.1186/s12889-020-8152-8] [Medline: 32098633]

XSL•FO
RenderX

33. Fowler SL, Klein WMP, Ball L, McGuire J, Colditz GA, Waters EA. Using an internet-based breast cancer risk assessment tool to improve social-cognitive precursors of physical activity. Med Decis Making 2017 Aug;37(6):657-669. [doi: 10.1177/0272989X17699835] [Medline: 28363033]
34. Höchsmann C, Infanger D, Klenk C, Königstein K, Walz SP, Schmidt-Trucksäss A. Effectiveness of a behavior change technique-based smartphone game to improve intrinsic motivation and physical activity adherence in patients with type 2 diabetes: randomized controlled trial. JMIR Serious Games 2019 Feb 13;7(1):e11444. [doi: 10.2196/11444] [Medline: 30758293]
35. Hopkin G, Au A, Collier VJ, Yudkin JS, Basu S, Naci H. Combining multiple treatment comparisons with personalized patient preferences: a randomized trial of an interactive platform for statin treatment selection. Med Decis Making 2019 Apr;39(3):264-277. [doi: 10.1177/0272989X19835239] [Medline: 30873906]
36. Irvine AB, Russell H, Manocchia M, et al. Mobile-Web app to self-manage low back pain: randomized controlled trial. J Med Internet Res 2015 Jan 2;17(1):e1. [doi: 10.2196/jmir.3130] [Medline: 25565416]
37. Kukafka R, Lussier YA, Eng P, Patel VL, Cimino JJ. Web-based tailoring and its effect on self-efficacy: results from the MI-HEART randomized controlled trial. Proc AMIA Symp 2002:410-414. [Medline: 12463857]
38. Maddison R, Pfaeffli L, Whittaker R, et al. A mobile phone intervention increases physical activity in people with cardiovascular disease: results from the HEART randomized controlled trial. Eur J Prev Cardiol 2015 Jun;22(6):701-709. [doi: 10.1177/2047487314535076] [Medline: 24817694]
39. Mevissen FEF, Ruiter RAC, Meertens RM, Zimbile F, Schaalma HP. Justify your love: testing an online STI-risk communication intervention designed to promote condom use and STI-testing. Psychol Health 2011 Feb;26(2):205-221. [doi: 10.1080/08870446.2011.531575] [Medline: 21318930]
40. Middelweerd A, Mollee J, Klein MM, Manzoor A, Brug J, Te Velde SJ. The Use and Effects of an App-Based Physical Activity Intervention “Active2Gether” in Young Adults: Quasi-Experimental Trial. JMIR Form Res 2020 Jan 21;4(1):e12538. [doi: 10.2196/12538] [Medline: 31961330]
41. Milan JE, White AA. Impact of a stage-tailored, web-based intervention on folic acid-containing multivitamin use by college women. Am J Health Promot 2010;24(6):388-395. [doi: 10.4278/ajhp.071231143] [Medline: 20594096]
42. Politi MC, Lee CN, Philpott-Streiff SE, et al. A randomized controlled trial evaluating the BREASTChoice tool for personalized decision support about breast reconstruction after mastectomy. Ann Surg 2020 Feb;271(2):230-237. [doi: 10.1097/SLA.0000000000003444] [Medline: 31305282]
43. Sittig S, Wang J, Iyengar S, Myneni S, Franklin A. Incorporating behavioral trigger messages into a mobile health app for chronic disease management: randomized clinical feasibility trial in diabetes. JMIR Mhealth Uhealth 2020 Mar 16;8(3):e15927. [doi: 10.2196/15927] [Medline: 32175908]
44. Valle CG, Queen TL, Martin BA, Ribisl KM, Mayer DK, Tate DF. Optimizing tailored communications for health risk assessment: a randomized factorial experiment of the effects of expectancy priming, autonomy support, and exemplification. J Med Internet Res 2018 Mar 1;20(3):e63. [doi: 10.2196/jmir.7613] [Medline: 29496652]
45. Vernon SW, Bartholomew LK, McQueen A, et al. A randomized controlled trial of a tailored interactive computer-delivered intervention to promote colorectal cancer screening: sometimes more is just the same. Ann Behav Med 2011 Jun;41(3):284-299. [doi: 10.1007/s12160-010-9258-5] [Medline: 21271365]
46. Weymann N, Dirmaier J, von Wolff A, Kriston L, Härter M. Effectiveness of a web-based tailored interactive health communication application for patients with type 2 diabetes or chronic low back pain: randomized controlled trial. J Med Internet Res 2015 Mar 3;17(3):e53. [doi: 10.2196/jmir.3904] [Medline: 25736340]
47. Wilson CJ, Flight IH, Turnbull D, et al. A randomised controlled trial of personalised decision support delivered via the internet for bowel cancer screening with a faecal occult blood test: the effects of tailoring of messages according to social cognitive variables on participation. BMC Med Inform Decis Mak 2015 Apr 9;15:25. [doi: 10.1186/s12911-015-0147-5] [Medline: 25886492]
48. Wong JYH, Zhang W, Wu Y, et al. An interactive web-based sexual health literacy program for safe sex practice for female Chinese University students: multicenter randomized controlled trial. J Med Internet Res 2021 Mar 12;23(3):e22564. [doi: 10.2196/22564] [Medline: 33709941]
49. Dingle GA, Carter NA. Smoke into sound: a pilot randomised controlled trial of a music cravings management program for chronic smokers attempting to quit. Music Sci 2017 Jun;21(2):151-177. [doi: 10.1177/1029864916682822]
50. Cullen T RD, MS, Hatch J RD, MHSc, Martin W RN, PhD, Higgins JW PhD, Sheppard R RN, BScN. Food literacy: definition and framework for action. Can J Diet Pract Res 2015 Sep;76(3):140-145. [doi: 10.3148/cjdpr-2015-010]
51. Bartholomew LK, Parcel GS, Kok G. Intervention mapping: a process for developing theory- and evidence-based health education programs. Health Educ Behav 1998 Oct;25(5):545-563. [doi: 10.1177/109019819802500502] [Medline: 9768376]
52. Stacey D, Volk RJ, IPDAS Evidence Update Leads (Hilary Bekker, Karina Dahl Steffensen, Tammy C. Hoffmann, Kirsten McCaffery, Rachel Thompson, Richard Thomson, Lyndal Trevena, Trudy van der Weijden, and Holly Witteman). The International Patient Decision Aid Standards (IPDAS) collaboration: evidence update 2.0. Med Decis Making 2021 Oct;41(7):729-733. [doi: 10.1177/0272989X211035681] [Medline: 34416841]
53. Sadegh SS, Khakshour Saadat P, Sepehri MM, Assadi V. A framework for m-health service development and success evaluation. Int J Med Inform 2018 Apr;112:123-130. [doi: 10.1016/j.ijmedinf.2018.01.003] [Medline: 29500009]

XSL•FO
RenderX

54. Allegra M, Bongiovanni A, Città G, Cusimano A, Dal Grande V. In: Gentile M, editor. The Role of Metaphor in Serious Games Design: The BubbleMumble Case Study 2021: Springer International Publishing. [doi: 10.1007/978-3-030-92182-8_19]
55. Blandford A, Furniss D, Makri S. Qualitative HCI Research: Going Behind the Scenes, 1st edition: Springer International Publishing; 2016. [doi: 10.1007/978-3-031-02217-3]
56. Toulmin S. The Uses of Argument Rev Ed: Cambridge University Press; 2003:247 URL: https://johnnywalters.weebly.com/uploads/1/3/3/5/13358288/toulmin-the-uses-of-argument_1.pdf [accessed 2025-08-26]
57. Kunz W. Issues as Elements of Information Systems: Universitat Stuttgart Institut fur Grundlagen der Planung iA; 1970. URL: http://magrawal.myweb.usf.edu/phd/articles/ibis_wp_70.pdf [accessed 2025-08-26]
58. Metz G, Peters GJY, Crutzen R. Acyclic behavior change diagrams: a tool to report and analyze interventions. Health Psychol Behav Med 2022;10(1):1216-1228. [doi: 10.1080/21642850.2022.2149930] [Medline: 36518605]
59. Blandford A, Gibbs J, Newhouse N, Perski O, Singh A, Murray E. Seven lessons for interdisciplinary research on interactive digital health interventions. Digit Health 2018;4:2055207618770325. [doi: 10.1177/2055207618770325] [Medline: 29942629]

Abbreviations

ACM: Association for Computing Machinery
HLS-EU: European Health Literacy Survey
IBIS: Issue Based Information Systems
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO: International Prospective Register of Systematic Reviews
QuADS: Quality Assessment for Diverse Studies

Edited by P Santana-Mancilla; submitted 17.04.25; peer-reviewed by C Obianyo, MA Coman, Y Zhang; revised version received 09.06.25; accepted 10.06.25; published 02.09.25.

Please cite as:
Hovingh JW, Elderson-van Duin C, Kuipers DA, van Rood Y, Ludden GDS, Hanssen DJC, Rosmalen JGM
Tailoring for Health Literacy in the Design and Development of eHealth Interventions: Systematic Review
JMIR Hum Factors 2025;12:e76172
URL: https://humanfactors.jmir.org/2025/1/e76172
doi:10.2196/76172

© Jan Wessel Hovingh, Cynthia Elderson-van Duin, Derek A Kuipers, Yanda van Rood, Geke D S Ludden, Denise J C Hanssen, Judith G M Rosmalen. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 2.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Framing the Human-Centered Artificial Intelligence Concepts and Methods: Scoping Review

Roberta Bevilacqua[1], MScPH; Tania Bailoni[2], MSc; Elvira Maranesi[1], PhD; Giulio Amabili[1], MEng; Federico Barbarossa[1], MEng; Marta Ponzano[3], PhD; Michele Virgolesi[4], PhD; Teresa Rea[4], PhD; Maddalena Illario[4,5], PhD; Enrico Maria Piras[6], PhD; Matteo Lenge[7], PhD; Elisa Barbi[7], BSc; Garifallia Sakellariou[8,9], PhD

[1]Scientific Direction, IRCCS INRCA, Via Santa Margherita 5, Ancona, Italy

[2]Intelligent Digital Agents (IDA) Research Group, Fondazione Bruno Kessler (FBK), Trento, Italy

[3]Department of Health Sciences, Section of Biostatistics, University of Genoa, Genoa, Italy

[4]Department of Public Health, University of Naples "Federico II", Naples, Italy

[5]Department of Endocrinology, Diabetes, Andrology and Nutrition, Federico II University Hospital, Naples, Italy

[6]Digital Health & Wellbeing, Fondazione Bruno Kessler, Trento, Italy

[7]Neuroscience and Human Genetics Department, Meyer Children's Hospital IRCCS, Florence, Italy

[8]Department of Internal Medicine and Therapeutics, University of Pavia, Pavia, Italy

[9]General Medicine, Istituti Clinici Scientifici Maugeri SpA SB, IRCCS, Pavia, Italy

Corresponding Author:
Elvira Maranesi, PhD
Scientific Direction, IRCCS INRCA, Via Santa Margherita 5, Ancona, Italy

Abstract

Background: With the rapid expansion of artificial intelligence (AI) applications, researchers have begun focusing on the concept of human-centered artificial intelligence (HCAI). This field is dedicated to designing AI systems that augment and improve human abilities, rather than substituting them.

Objective: The objective of the paper was to review the information on design principles, techniques, applications, methods, and outcomes adopted in the field of HCAI, in order to provide some insights on the discipline, in relation with the broader concepts of human-centered and user-centered design.

Methods: Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR) checklist guidelines, we conducted a scoping review in PubMed, ScienceDirect, and IEEE Xplore, including all study types, excluding narrative reviews and editorials.

Results: Out of the 1035 studies retrieved, 14 studies conducted between 2018 and 2023 met the inclusion criteria. The main fields of application were the health sector and AI applications. Human-centered design methodologies were adopted in 3 studies, personas in 2 studies, while the remaining methodologies were adopted in individual studies.

Conclusions: HCAI emphasizes designing AI systems that prioritize human needs, satisfaction, and trustworthiness, but current principles and guidelines are often vague and difficult to implement. The review highlights the importance of involving users early in the development process to enhance trust, especially in fields like health care, but notes that there is a lack of standardized HCAI methodologies and limited practical applications adhering to these principles.

(JMIR Hum Factors 2025;12:e67350) doi:10.2196/67350

KEYWORDS

human-centered artificial intelligence; HCAI; usability; acceptability; design principles; UCD; human-centered design; AI principles; ethics; AI applications; machine learning; ML; algorithm; analytics; predictive model; predictive analytics; deep learning; early warning; early detection; review; user-centered design

Introduction

User-centered design (UCD) is an iterative methodology that places the user at the center of the design of innovative solutions, allowing the information gathered since its early stages to define the product features and end user experience. This approach is typically enabled by interdisciplinary teams and different methodologies that work together synergistically to optimize the user experience of systems, products, and processes.

The interaction of users with innovation prototypes not only accelerates the identification of usability issues, highlighting improvement opportunities, but also strengthens the capacity

XSL•FO
RenderX

of researchers to define cost and benefits evaluation methods that are propaedeutic to identify the return of investments also in terms of economic benefit. The requirements and recommendations for human-centered design principles have been formalized in ISO 9241‐210, which details the role of UCD and its benefits for human-centered design applied to interactive technologies.

The use of appropriate UCD methods, especially when tailored to specific stakeholders' contexts, can reduce the risk of a given product not meeting stakeholder requirements or being rejected by users.

Similar to UCD, human-centered design (HCD) is an approach that prioritizes human needs, capabilities, and behavior. Therefore, HCD aims to address problems by putting people with their human perspective at the center of the processes, involving them in all stages of problem solving, from observation to brainstorming, conceptualizing, solution development, and final implementation [1]. It is believed that such an approach can improve the usability of an innovation, increasing product acceptance and user satisfaction, proving effective in all situations where solutions are needed that are not only useful, but also involve the emotional sphere of users in some way [2-4].

Recently, due to the increasing growth of artificial intelligence (AI) applications, researchers have shifted their attention to the construct called human-centered artificial intelligence (HCAI), a discipline aimed at creating AI systems to amplify and enhance human capabilities, rather than replace them [5]. As HCD, HCAI places people at the center by seeking to improve, experience after experience, their lives [6]. On the basis of the design phase of any technological product, there are a number of useful techniques that can be adopted to obtain quick feedback from the end users, even in the absence of a fully functioning solution, to provide creative design alternatives to fit the users' preferences and needs. If properly trained, HCAI could offer useful solutions tailored to the peculiar characteristics of the final users, for example, considering new approaches coming from the original combination of many variables.

Despite the undeniable richness of HCAI, there is the need to map the main concepts and constructs behind it, as well as to understand the methods and design principles that underpin it, in order to ensure that these metrics are adequately included in the experimental methodology and understand their role in supporting the UCD and HCD approach. The objective of this paper is to review the information on design principles, techniques, applications, methods, and outcomes adopted in the field of HCAI, in order to provide some insights on the discipline, in relation to the broader concepts of HCD and UCD.

Methods

The scoping review was carried out according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR; see Checklist 1) [7]. Before the start of the project, a protocol was developed to guide the review process, shared among the authors and registered on the Open Science Framework (OSF) platform. The topic of interest for the review was any application of HCAI for UCD, in any population of adults. The outcomes of interest were descriptions of the HCAI approach, description of the HCAI design, outcomes of the application of HCAI, user involvement in HCAI, description of use cases, methods for understanding and the user's mental model, description of user needs, and new working practices. Any other design methodology could serve as a comparator. All English-language studies, with the exception of narrative and editorial reviews, could be included. Clinical questions, exploring the applications of HCAI for design, were translated into epidemiological terms using the Patient, Intervention, Comparator, Outcome, and Study Type (PICOS) methodology (see Textbox 1).

Based on the identified PICOS, search strategies were created and applied to the electronic PubMed, ScienceDirect and IEEE Xplore databases from January 1, 2018, to October 27, 2023. Searches were performed by a single reviewer. The complete search strategies are presented in Multimedia Appendix 1.

The records retrieved from the search were imported into the review management platform Rayyan, where duplicate entries were removed. A total of 4 investigators (RB, EM, TB, and ML) working in pairs performed screening, selection, and data extraction. Specifically, 2 blinded reviewers assessed the title and abstracts to define eligibility for full-text assessment, and, at the subsequent step, assessed full-texts for inclusion. Disagreement was resolved by consensus. Data extraction from the included studies was performed by using a standardized extraction form, including general information on the paper, the field of application, design methods, the adopted model, design principles and properties, technology, and type of AI; extraction was performed by a reviewer and checked by a second one. Results were summarized in summary of findings tables. The methodological quality of the included studies was assessed by a single author (GS) with the SANRA (Scale for the Assessment of Narrative Review Articles) or the Critical Appraisal Skills Programme (CASP) checklist for qualitative studies, depending on the study design.

XSL•FO
RenderX

Textbox 1. PICOS (Patient, Intervention, Comparator, Outcome, Study Type) driving the search strategies and the inclusion criteria.

Patient:

- Any type of patient.

Intervention:

- Human-centered artificial intelligence (HCAI) for design.

Comparator:

- Any other technique adopted for design.

Outcome:

- Description of HCAI approaches.
- Description of HCAI design.
- Outcome of HCAI application.
- User involvement in HCAI.
- Description of study cases.
- Description of methods for understanding.
- Description of user's mental model.
- User's needs and new working practices.

Study type:

- All study types excluding narrative reviews and editorials.

Results

The search in the electronic databases retrieved a total of 1035 articles, of which 8 were duplicates. Finally, the titles and abstracts of 1027 studies were screened. Based on the inclusion criteria, 27 studies were eligible for full-text review; of these, 14 studies fulfilled the criteria and were included in the scoping review. Studies excluded during the full-text analysis are listed in Multimedia Appendix 2, where the title of the paper and the reason for exclusion are indicated. Figure 1 shows the flowchart of the selection process.

Conference abstracts represented a relevant proportion of the retrieved evidence (7/15 studies), no study allowed comparison between techniques. The main fields of application were the health sector and AI applications (see Table 1).

The method adopted for design was described in 8/14 studies. In particular, HCD methodologies were adopted in 3 studies, personas in 2 studies, while the remaining methodologies were adopted in individual studies.

In 10/14 studies, a model was adopted, while in 3 studies this was not applied. Information was not available for a single study.

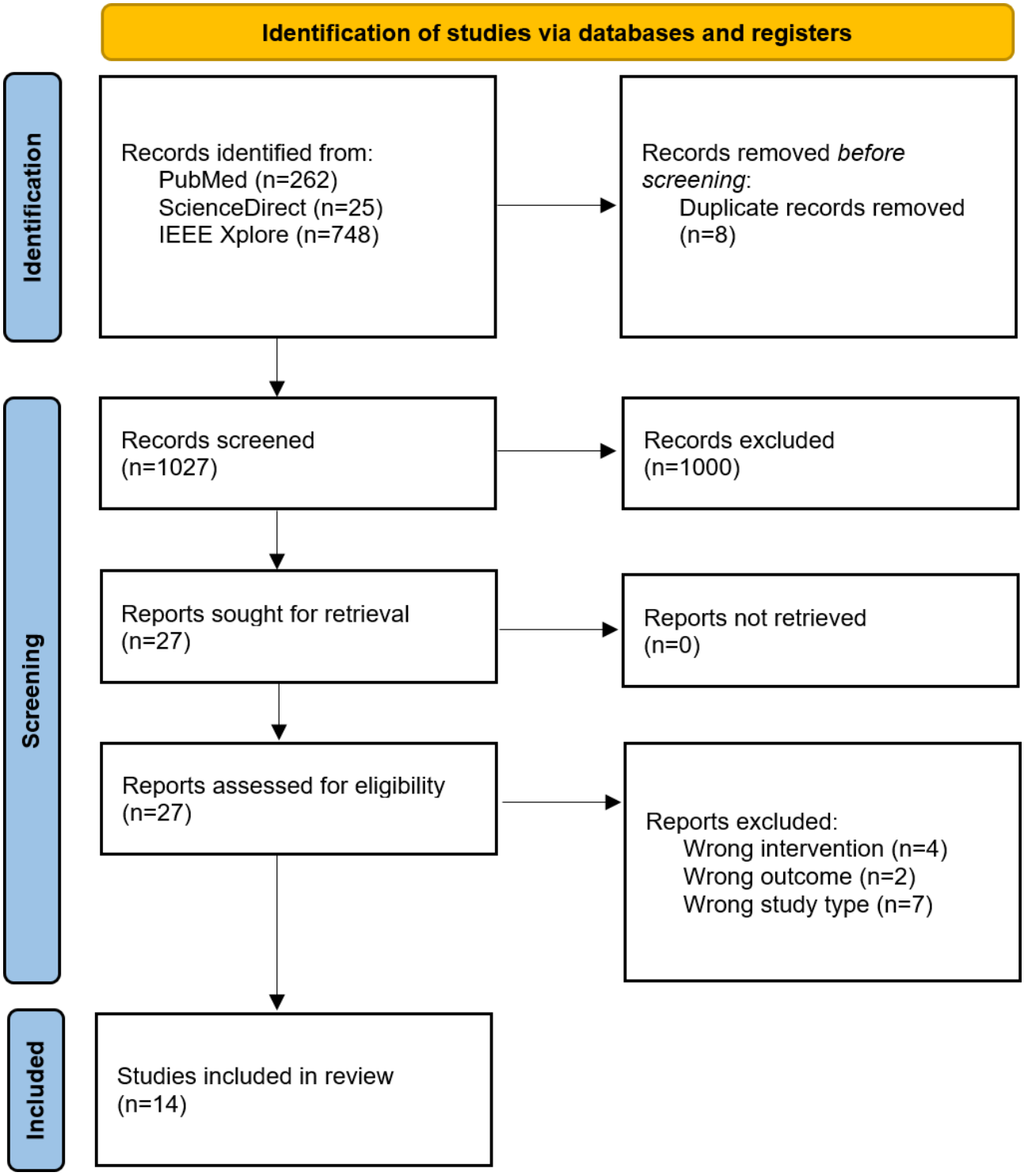
Design principles and properties were not described in 2 studies, in the remaining studies trustworthiness and explainability were the most frequently adopted principles (see Table 2).

The technology and type of AI were not described in 3 studies, while the remaining studies adopted very heterogeneous technologies.

Table 3 shows the characteristics of each study in terms of scope, presence of a model, design principles used, and technology.

XSL•FO
RenderX

Figure 1. Flowchart showing the selection process.



XSL•FO
RenderX

Table . Different applications of human-centered artificial intelligence (HCAI) for design.

Field of application	Studies, n
Health	6
AI[a] application	5
Privacy	1
General	1
Education	1
Security	1
Safety	1

[a]AI: artificial intelligence.

Table . Design principles and properties.

Design principles and properties	Studies, n
Trustworthiness	6
Explainability	4
Transparency	3
User needs	3
Privacy	3
Fairness	3
Safety	2
Acceptability	2
Reliability	1
Understandability	1
Protection	1
Model confidence	1
Failure	1
Errors	1
Developer needs	1
Data needs	1
Model needs	1
Interpretability	1
Accountability	1
Decision making authority	1
Emotion	1

XSL•FO
RenderX

Table . Descriptive analysis of the included studies.

Study	Type of publication	Field of application	UCD[a] or design methods	Yes or no (model)	Yes or no (design principles or properties)	Technology or type of AI[b]
He et al, 2022 [8]	Article	Safety, security, and health	—[c]	Yes (acceptance model of TRAS[d])	Yes (trustworthiness)	SAR[e]
Tyagi et al, 2023 [9]	Conference abstract	Education	Value-sensitive design	Yes (enhanced HCAI[f] framework)	No	AIED[g] application (intelligent tutoring system; exploratory learning environments, dialogue-based tutoring system, automatic writing assessment)
Holzinger et al, 2022 [10]	Conference abstract	Health	Personas for AI (explainability)	Yes (for the design and development of the personas AI)	Yes (emotion, decision-making authority and explainability as well as ethical issues are added in HAII[h])	—
Usmani et al, 2023 [11]	Conference abstract	General	—	Yes (user empowerment, ethical consideration, and human AI collaboration)	Yes (fairness, transparency, accountability, and privacy protection)	—
Fagbola and Thakur, 2019 [12]	Conference abstract	Development of AI systems	—	No	Yes (interpretability, explainability, fairness, transparency, and safety)	Toolkit for developing AI system (FairML, Aequitas, FairTest, IBM AIFairness 360, Mash, Concept Activation Vectors, LIME, DeepLIFT, and Themis)
Beltrão et al, 2022 [13]	Conference abstract	AI-mind	Yes (HCD[i] methods: personas, scenarios, and journey maps)	N o	No	AI-based decision-support system for clinical decision
Correia et al, 2021 [14]	Conference abstract	Human-AI integration	—	Yes (acceptance model and use of technology [UTAUT[j] and TAM[k]])	Yes (trustworthiness and acceptability)	—
Elahi et al, 2021 [15]	Article	Privacy and health	HCD methods	Yes (design shared responsibility privacy model)	Yes (privacy protection)	AAL system (ambient assisted living)
Ahmad et al, 2023 [16]	Article	Development of AI system using HCD.	Human-centered methods	Yes (human-centered AI-based software [RE4HCAI])	Yes (requirements include user needs, model needs, data needs, explainability, and trust, as well as error handling and failure mitigation)	Model applied in a system of VR (virtual reality)
Bingley et al, 2023 [6]	Article	HCAI	Qualitative survey for HCAI	—	Yes (AI developers and user's needs [functionality, social impact, understandability, ethic, privacy, security])	—

XSL•FO
RenderX

Study	Type of publication	Field of application	UCD[a] or design methods	Yes or no (model)	Yes or no (design principles or properties)	Technology or type of AI[b]
Ventura et al, 2023 [17]	Article	Health	—	N O	Yes (user needs, acceptance of technology, and perceived truthfulness)	MAIA[l] technology
Soliman et al, 2023 [18]	Article	Health	Yes (stakeholder input, low-fidelity sketches and high-fidelity prototype for usability test)	Yes (model performance and explainability)	Yes (model confidence, trustworthiness, and explainability)	Clinical decision support application to predict heart failure patient risk of readmission
Chen et al, 2023 [19]	Article	Health care	—	Yes (for mitigating biases in AI life cycle)	Yes (Ethical, privacy protection, fairness, understandable, and transparent)	—
Kim et al, 2022 [20]	Article	Health (home care for older adults)	Yes (focus group interviews and scenarios)	Yes (human-AI collaboration and user satisfaction)	Yes (reliable, safe, and trustworthy)	DORI - older adult guided and caregiver-monitored robot

[a]UCD: user-centered design.

[b]AI: artificial intelligence.

[c]Not available.

[d]TRAS: trustworthy robots and autonomous system.

[e]SAR: socially assistive robot.

[f]HCAI: human-centered artificial intelligence.

[g]AIED: artificial intelligence in education.

[h]HAII: human–artificial intelligence interaction.

[i]HCD: human-centered design.

[j]UTAUT: Unified Theory of Acceptance and Use of Technology.

[k]TAM: technology acceptance model.

[l]MAIA: multifunctional adaptive and interactive artificial intelligent system.

The study by He et al [8] explores the critical role of trustworthiness in the acceptance of AI systems within the domains of safety, security, and health. By employing the Trustworthiness and Reliability Assessment System (TRAS) model, the research demonstrates that user acceptance is significantly influenced by the perceived trustworthiness and reliability. Socially assistive robots are used as a case study, illustrating how trust can be enhanced in AI systems within these critical fields.

Tyagi et al [9] investigates the application of AI in education through value-sensitive design methods. The study presents an enhanced HCAI framework and explores various AI applications such as intelligent tutoring systems and dialogue-based tutoring systems. The findings underscore the importance of aligning AI systems with educational values and addressing user needs to improve acceptance.

Kim et al [20], Chen et al [19], Soliman et al [18], Ventura et al [17], and Holzinger et al [10] focus on health care applications with distinct approaches. Kim et al [20] examines home care for older adults, highlighting human-AI collaboration and user satisfaction as critical design principles. Chen et al [19] address biases in the AI lifecycle within health care, emphasizing ethical principles, privacy protection, fairness, and transparency. Soliman et al [18] investigates clinical decision support systems for predicting heart failure patient readmission risks, emphasizing model performance, explainability, and trustworthiness. Ventura et al [17] explores health applications, emphasizing user needs and technology acceptance, particularly the perceived trustworthiness of AI systems. Holzinger et al [10] uses personas in the health sector to enhance AI explainability, integrating emotional and ethical considerations to improve user understanding and trust in AI technologies. These studies collectively contribute to advancing AI applications in health care by addressing various challenges and enhancing user confidence and acceptance.

Usmani et al [11] addresses general AI applications with an emphasis on user empowerment and ethical considerations. The research highlights the importance of human-AI collaboration and identifies design principles such as fairness, transparency, accountability, and privacy protection. These principles are crucial for developing trustworthy and ethical AI systems.

Fagbola and Thakur [12] and Ahmad et al [16] discuss the development of AI systems from complementary perspectives. Fagbola's study emphasizes interpretability, explainability, fairness, transparency, and safety as crucial design principles, introducing tools like FairML and IBM AIFairness 360. Ahmad's research focuses on human-centered methods, presenting RE4HCAI, a software that addresses user needs,

XSL•FO
RenderX

model requirements, data considerations, explainability, and trust. These approaches collectively aim to foster the development of fair, transparent, and user-centric AI systems, enhancing both usability and trust in AI applications, including virtual reality environments.

Beltrão et al [13] explores the use of human-centered design methods such as personas, scenarios, and journey maps in developing AI-based decision-support systems for clinical settings. The study highlights the importance of aligning AI systems with clinical user needs to improve decision-making processes.

Correia et al [14] delves into human-AI integration, focusing on acceptance models and technology use. The study uses models like Unified Theory of Acceptance and Use of Technology (UTAUT) and technology acceptance model to evaluate user acceptance, emphasizing trustworthiness and acceptability. These insights are vital for designing AI systems that users are more likely to adopt.

Elahi et al [15] examines privacy and health applications, using human-centered design methods. The study introduces a shared responsibility privacy model, underscoring privacy protection as a key design principle. This research highlights the importance of safeguarding user privacy in health-related AI systems.

Bingley et al [6] conducts a qualitative survey on HCAI. The study emphasizes the needs of AI developers and users, addressing aspects such as functionality, social impact, understandability, ethics, privacy, and security. These insights are crucial for developing AI systems that meet diverse stakeholder requirements.

The methodological quality of the included studies was overall adequate, with an unclear quality of the reporting of the relationship between the researcher and participants and ethical aspects in qualitative research, while the aspect of the literature research was not explicitly addressed in most studies with a narrative design (see Figures 2 and 3).

Figure 2. Methodological quality of the included studies presenting qualitative research, assessed with the Critical Appraisal Skill Programme (CASP) checklist for qualitative studies. Correspondence between the color and the items in the checklist: green=yes; red=no; yellow=can't tell [6,10,13-18,20].

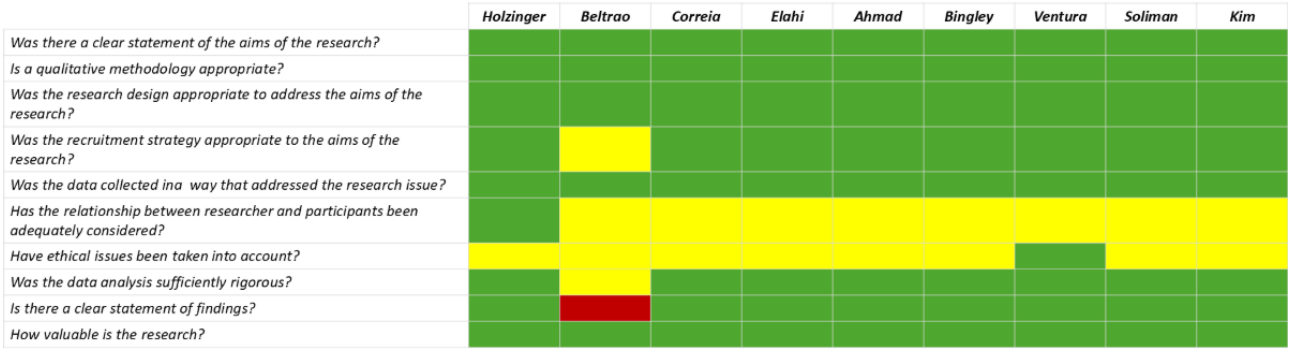


Figure 3. Methodological quality of the included studies presenting narrative results, assessed with the Scale for the Assessment of Narrative Review Articles (SANRA) checklist. Correspondence between the color and the items in the checklist: green=2 points; red=0 points; yellow=1 point [8,9,11,12,19].



Discussion

HCAI and Its Challenges in Practice

HCAI focuses on human experiences, satisfaction, and the needs, with the goal of "amplifying, enhancing, and improving human performance in ways that make systems reliable, safe, and trustworthy." This is pivotal to "support human self-efficacy, encourage creativity, clarify responsibility, and facilitate social participation" [21].

The HCAI seeks to shift the focus in AI development, from technologies to people. However, it is unclear whether existing HCAI principles and practices adequately achieve this goal.

The results of our review indicate that although many studies address the issue of trust and acceptability in HCAI systems, few adopt systematic strategies to mitigate demographic biases and biases in health data. However, in the absence of a standardized methodology to control bias, there is a need to develop more structured guidelines to ensure trustworthiness in AI systems applied to health. In order to formalize these developments, several guidelines have been proposed by governments, organizations, and researchers to translate the ideals of HCAI into practice [22]. For example, the European Union lists 7 key requirements that AI systems should meet to be trustworthy, including transparency, accountability, and promotion of social and environmental well-being [23].

XSL•FO
RenderX

Unfortunately, despite this proliferation of guidelines, IAE ideals have proven difficult to implement [21,24,25]. At this point, Shneiderman [26] argued that although ethical guidelines are a step in the right direction, they are often too vague to be useful to software engineers. Similarly, Mittelstadt [27] criticized AI ethics for having vague principles and lofty value statements that lack the detail and precision needed to make specific recommendations for improving practice.

In line with the literature in the field, our review highlighted the need for a deeper analysis of the design principles promoted by HCAI to understand how they can be put into practice from the outset and how they differ from those proposed by HCD in terms of their impact on end consumers.

In general, it can be said that there is a need to raise the awareness of researchers and developers on the special issues offered by HCAI: the review showed that there is still a limited number of applications for AI design and solutions already available developed in accordance with HCAI principles, mainly in the field of social assistive robotics. This can partly be explained by the innovative nature of the proposed concept, but also by the partial overlap with the more general definition of human-centered design and user-centered design, highlighting the need for a clear definition and standardization of the framework.

From the analysis of recent literature [28], the explainability represents one feature deeply investigated by developers and researchers in the field of AI applications. In the health care sector, in particular, it is possible to notice a relevant focus of attention given to this principle especially from the stakeholders' point of view. In fact, they are in charge of suggesting the use of AI systems to the final users and/or rely on AI decision support systems for complex diagnosis, for example. This evidence highlights the crucial need for transparency and understandability as the basis of explainability of AI, in particular in sensitive contexts as the healthcare sectors, in which erroneous or correct decisions have a significant impact on the well-being of older patients. At the same time, the concept of explainability overcame the mere role of technical prerequisite and became the essential driver to favor the systems' trustworthiness.

On the end users side, in fact, it seems that the value that mostly drives and supports the uptake of AI solutions is represented by the trustworthiness, as also supported by our findings.

Our review, in fact, showed that there is the need to focus more on the concept of trustworthiness when designing a system. Even if all the articles reviewed are in accordance with observing that a user needs to trust an AI-based system in order to use it, however, they differ in the approaches to reach this outcome.

This concept is in fact multifaceted and not easily defined or achieved, as dependent from intrinsic and extrinsic characteristics of the users and of the AI technology. Several theories on technology acceptance have underlined the role of trust especially in the field of social robotics, giving AI applications a primary role in the study of the topic of trust [29,30].

Embodiment, social presence, capacity to interact following social norms and proximity, and advanced dialog features are in fact key ingredients of social assistive robots, which stimulate reciprocity and trust in the relationship between human and robot [29,30].

In addition to explainability and trustworthiness, the main concepts that are taken into consideration in the papers reviewed are transparency, privacy, and safety, which contribute to making the system more trustworthy and acceptable for the user.

The way to better achieve these aspects is still an open challenge, but this review demonstrates the importance of involving real users and stakeholders from the early conceptualization and development stage, through the entire lifecycle and testing phases of a system. This becomes a crucial aspect, especially in the development of applications to support the decision-making by health professionals in clinical fields.

Furthermore, by incorporating user feedback into the design process, users can have a sense of ownership and control over AI systems, potentially promoting trust, acceptance and better adherence, even if in most cases they have had no previous experience with AI solutions.

In addition to the need for specific HCAI methods and techniques for the design and development of new AI systems, the articles examined either did not describe the methodologies used in the design process or used the same methods as those proposed by the user-centered design approach, such as focus groups, personification, scenarios, interviews, and others, as they are independent of the type of technology while maintaining the focus on users as a pillar.

Role of AI in Health Care

AI holds significant potential across various health care domains. AI-driven tools, particularly those based on machine learning algorithms, enhance early disease detection through advanced medical imaging analysis, such as radiography, CT, and MRI. Soliman et al [18] demonstrated the efficacy of AI in clinical decision support systems for predicting hospital readmission risk in heart failure patients, emphasizing model performance and explainability. Similarly, Chen et al [19] highlighted the use of AI to mitigate biases in healthcare applications, with a focus on ethical principles, privacy protection, and transparency. In addition, AI contributes to personalized medicine by analyzing large datasets to predict treatment responses and optimize drug dosages. Beltrão et al [13] explored AI-based clinical decision support systems, emphasizing human-centered methodologies, such as personas and journey maps [13]. Soliman et al [18] further demonstrated AI's role in improving clinical prediction models for readmission risk.

AI also enhances patient care through virtual agents and remote monitoring systems. Kim et al [20] investigated the use of assistive robots (DORI) for home care in older adults, highlighting human–AI collaboration and patient satisfaction. Elahi et al [15] introduced a shared-privacy model for ambient assisted living applications, emphasizing privacy protection as a key design principle. Furthermore, AI-powered medical imaging applications automate routine tasks and reduce

XSL•FO
RenderX

radiologists' cognitive load. Chen et al [19] underscored AI's role in reducing biases in image analysis. Holzinger et al [10] applied the concept of personas to improve explainability in AI systems for health care, integrating emotional and ethical dimensions to foster user trust. Despite these advancements, several challenges remain overlooked. A critical issue is the "black box" problem, referring to the opacity of AI decision-making processes, which prevents users from understanding how outcomes are generated. Although explainability and transparency have been discussed [10-12], the specific concept of the "black box"—which highlights the difficulty in interpreting model behavior—has not been directly addressed. Usmani et al [11] stressed the importance of fairness, transparency, accountability, and privacy protection in developing ethical and reliable AI systems. Fagbola et al [12] discussed tools such as FairML and IBM AI Fairness 360 to ensure algorithm interpretability and transparency. In addition, integrating AI into clinical workflows poses a challenge. Ventura et al [17] explored co-design processes with poststroke patients and caregivers, demonstrating that user involvement enhances AI solution acceptance and facilitates integration into care pathways. Correia et al [14] emphasized the importance of technological acceptance, using models like the UTAUT and technology acceptance model to assess AI's reliability and acceptability in health care environments.

Ethical Concerns in AI for Health Care

In addition to the technical and practical challenges, the integration of AI in health care also raises critical ethical concerns, particularly in sensitive contexts. AI decision-making in health care can have significant implications for vulnerable populations, such as those with serious or chronic conditions, the elderly, or minors. These decisions need to be made with the utmost care to avoid potential biases and discrimination. To address these concerns, AI systems must be designed with ethical safeguards that ensure fairness, transparency, and accountability, especially in critical decision-making processes.

Furthermore, AI's role in processing sensitive patient data underscores the importance of safeguarding privacy and ensuring data security. As AI systems handle increasingly personal and confidential information, adherence to privacy laws, such as the General Data Protection Regulation, becomes essential. Health care providers must ensure that AI technologies respect patient confidentiality while maintaining trust in the system.

Another ethical challenge involves the issue of informed consent. As AI becomes more integrated into clinical decision-making, patients must be fully informed about how AI systems are used in their care. This includes understanding the role of AI in diagnostics, treatment planning, and decision support, as well as the extent to which AI influences clinical outcomes. Clear and transparent consent protocols should be developed to ensure that patients have a clear understanding of their rights and the implications of AI technologies on their care.

Limitations and Future Directions for HCAI

Although this review provides important insights into HCAI, several limitations must be recognized. First, the number of studies focusing on HCAI remains limited and many of these overlap with human-computer interaction. This overlap hinders the differentiation between the two fields and underlines the need for further research to establish the differences between the two fields. A review of this work will be necessary in the coming years to follow the evolution of HCAI and assess the progress that will be made in establishing unique methodologies and contributions.

Another limitation is the need for a more systematic approach to define the design principles underlying HCAI. While co-design and participatory design are widely used, there is no clear guidance on how to integrate the specific principles of HCAI within these processes. Future studies should aim to develop frameworks that not only focus on technical aspects, but also address ethical, social, and contextual considerations.

Trustworthiness, explainability, and reliability remain the most critical issues for the adoption of AI-driven systems, particularly in health care. The ability to provide clear and complete explanations of AI-generated results is crucial to promote user trust. Future research should analyze practitioners' perceptions of explainability, exploring what factors enhance or hinder trust in AI systems.

Furthermore, these emerging tools represent an opportunity to improve the participatory design process by capturing both user needs and expectations. Research is needed to understand how these tools can complement traditional HCD methods and contribute to a more user-centered AI development.

Conclusions

HCAI emphasizes designing AI systems that prioritize human needs, satisfaction, and trustworthiness, but current principles and guidelines are often vague and difficult to implement. The review highlights the importance of involving users early in the development process to enhance trust, especially in fields like health care, but notes that there is a lack of standardized HCAI methodologies and limited practical applications adhering to these principles.

From our perspective, the future of HCAI will depend on its ability to establish itself as a structured and distinct research field. To achieve this, the development of standardized methodologies and principles tailored specifically to HCAI is essential. Rather than relying solely on existing human-centered approaches, the field should work toward codifying new design principles that account for the unique challenges posed by AI systems.

We anticipate that interdisciplinary collaboration will play a crucial role in advancing HCAI. Insights from psychology, cognitive science, and human factors engineering should be incorporated to refine HCAI methodologies, ensuring that AI systems align with human values and cognitive capabilities. Furthermore, integrating AI-based tools for user need assessment could significantly enhance the participatory design process, providing more precise and scalable insights into user expectations.

While AI in health care holds significant potential, there are critical gaps that require further investigation. Our review

XSL•FO
RenderX

highlights the need for a deeper understanding of how biases in training datasets could impact equitable health care outcomes, and how improving transparency in AI decision-making is crucial for building trust. Privacy and data security concerns also need ongoing attention, particularly as AI systems increasingly handle sensitive patient information.

To bridge these gaps, future research should prioritize the integration of AI into clinical workflows, addressing potential challenges to adoption and ensuring that AI solutions are practical, ethical, and aligned with the needs of both health care providers and patients. Moving forward, interdisciplinary approaches that combine technical advancements with human-centered design will be essential for creating AI systems that are not only efficient but also ethically sound and trustworthy

Acknowledgments

This research was supported by the National Plan for Complementary Investments to the PNRR Hub Life Science-Digital Health (PNC-E3-2022-23683267; DHEAL COM). This publication reflects only the authors' view and the Italian Ministry of Health is not responsible for any use that may be made of the information it contains.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategies.
[DOCX File, 12 KB - humanfactors_v12i1e67350_app1.docx]

Multimedia Appendix 2
Studies excluded during the full text analysis indicating exclusionion motivation.
[DOCX File, 14 KB - humanfactors_v12i1e67350_app2.docx]

Checklist 1
Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (Prisma-ScR) checklist.
[DOCX File, 84 KB - humanfactors_v12i1e67350_app3.docx]

References

1. Melles M, Albayrak A, Goossens R. Innovating health care: key characteristics of human-centered design. Int J Qual Health Care 2021 Jan 12;33(Supplement_1):37-44. [doi: 10.1093/intqhc/mzaa127] [Medline: 33068104]
2. Norman D. Emotion & design: attractive things work better. Interactions 2002;9. [doi: 10.1145/543434.543435]
3. Alaniz T, Biazzo S. Emotional design: the development of a process to envision emotion-centric new product ideas. Procedia Comput Sci 2019;158:474-484. [doi: 10.1016/j.procs.2019.09.078]
4. Hanington B. Chapter 6 - design and emotional experience. In: Jeon M, editor. Emotions and Affect in Human Factors and Human-Computer Interaction: Academic Press; 2017:165-183. [doi: 10.1016/B978-0-12-801851-4.00006-9]
5. Schmager S, Pappas I, Vassilakopoulou P. Defining human-centered AI: a comprehensive review of HCAI literature. Presented at: Proceedings of MCIS 2023; Sep 6-9, 2023 p. 13 URL: https://aisel.aisnet.org/mcis2023/13 [accessed 2025-05-16]
6. Bingley WJ, Curtis C, Lockey S, et al. Where is the human in human-centered AI? Insights from developer priorities and user experiences. Comput Human Behav 2023 Apr;141:107617. [doi: 10.1016/j.chb.2022.107617]
7. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
8. He H, Gray J, Cangelosi A, Meng Q, McGinnity TM, Mehnen J. The challenges and opportunities of human-centered AI for trustworthy robots and autonomous systems. IEEE Trans Cogn Dev Syst 2022 Dec;14(4):1398-1412. [doi: 10.1109/TCDS.2021.3132282]
9. Tyagi PK, Singh AK, Tyagi P, Kumar Chanchal H, Gautam P, Nibhoria S. Human-centred artificial intelligence in education, present and future opportunities. Presented at: Proceedings of the 2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE); May 12-13, 2023; Greater Noida, India. [doi: 10.1109/ICACITE57410.2023.10182545]
10. Holzinger A, Kargl M, Kipperer B, Regitnig P, Plass M, Muller H. Personas for artificial intelligence (AI) an open source toolbox. IEEE Access 2022;10:23732-23747. [doi: 10.1109/ACCESS.2022.3154776]
11. Usmani UA, Happonen A, Watada J. Human-centered artificial intelligence: designing for user empowerment and ethical considerations. Presented at: Proceedings of the 2023 5th International Congress on Human-Computer Interaction,

XSL•FO
RenderX

Optimization and Robotic Applications (HORA); Jun 8-10, 2023; Istanbul, Turkiye. [doi: 10.1109/HORA58378.2023.10156761]
12. Fagbola TM, Thakur SC. Towards the development of artificial intelligence-based systems: human-centered functional requirements and open problems. Presented at: Proceedings of the 2019 4th International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS); Nov 21-24, 2019; Shanghai, China. [doi: 10.1109/ICIIBMS46890.2019.8991505]
13. Beltrão G, Paramonova I, Sousa S. User interface design for ai-based clinical decision-support system: preliminary study. Presented at: Proceedings of the 2022 17th Iberian Conference on Information Systems and Technologies (CISTI); Jun 22-25, 2022; Madrid, Spain. [doi: 10.23919/CISTI54924.2022.9820378]
14. Correia A, Fonseca B, Paredes H, Chaves R, Schneider D, Jameel S. Determinants and predictors of intentionality and perceived reliability in human-AI interaction as a means for innovative scientific discovery. Presented at: Proceedings of the 2021 IEEE International Conference on Big Data (Big Data); Dec 15-18, 2021; Orlando, FL, USA p. 3681-3684. [doi: 10.1109/BigData52589.2021.9671358]
15. Elahi H, Castiglione A, Wang G, Geman O. A human-centered artificial intelligence approach for privacy protection of elderly app users in smart cities. Neurocomputing 2021 Jul;444:189-202. [doi: 10.1016/j.neucom.2020.06.149]
16. Ahmad K, Abdelrazek M, Arora C, Agrahari Baniya A, Bano M, Grundy J. Requirements engineering framework for human-centered artificial intelligence software systems. Appl Soft Comput 2023 Aug;143:110455. [doi: 10.1016/j.asoc.2023.110455]
17. Ventura S, Ottoboni G, Lullini G, et al. Co-designing an interactive artificial intelligent system with post-stroke patients and caregivers to augment the lost abilities and improve their quality of life: a human-centric approach. Front Public Health 2023;11:1227748. [doi: 10.3389/fpubh.2023.1227748] [Medline: 37808976]
18. Soliman A, Nair M, Petersson M, et al. Interdisciplinary human-centered AI for hospital readmission prediction of heart failure patients. Stud Health Technol Inform 2023 May 18;302:556-560. [doi: 10.3233/SHTI230204] [Medline: 37203747]
19. Chen Y, Clayton EW, Novak LL, Anders S, Malin B. Human-centered design to address biases in artificial intelligence. J Med Internet Res 2023 Mar 24;25:e43251. [doi: 10.2196/43251] [Medline: 36961506]
20. Kim JW, Choi YL, Jeong SH, Han J. A care robot with ethical sensing system for older adults at home. Sensors (Basel) 2022 Oct 3;22(19):7515. [doi: 10.3390/s22197515] [Medline: 36236614]
21. Shneiderman B. Bridging the gap between ethics and practice: guidelines for reliable, safe, and trustworthy human-centered AI systems. ACM Trans Interact Intell Syst 2020;10(4):26-31. [doi: 10.1145/3419764]
22. Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. Nat Mach Intell 2019;1(9):389-399. [doi: 10.1038/s42256-019-0088-2]
23. Ethics guidelines for trustworthy AI. Futurium - European Commission. 2019. URL: https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines/1 [accessed 2025-05-16]
24. Oppermann L, Boden A, Hofmann B, Prinz W, Decker S. Beyond HCI and CSCW: challenges and useful practices towards a human-centred vision of AI and IA. Proceed Halfway to the Future Sympos 2019. [doi: 10.1145/3363384.3363481]
25. Winfield AFT, Jirotka M. Ethical governance is essential to building trust in robotics and artificial intelligence systems. Phil Trans R Soc A 2018 Nov 28;376(2133):20180085. [doi: 10.1098/rsta.2018.0085]
26. Shneiderman B. Human-centred AI. Iss Sci Technol 2021;37(2):56-61 [FREE Full text]
27. Mittelstadt B. Principles alone cannot guarantee ethical AI. Nat Mach Intell 2019;1(11):501-507. [doi: 10.1038/s42256-019-0114-4]
28. Beger J. The crucial role of explainability in healthcare AI. Eur J Radiol 2024 Jul;176:111507. [doi: 10.1016/j.ejrad.2024.111507] [Medline: 38761444]
29. Heerink M, Kröse B, Evers V, Wielinga B. The influence of social presence on acceptance of a companion robot by older people. J Phys Agents 2008;2(2):33-40. [doi: 10.14198/JoPha.2008.2.2.05]
30. Bevilacqua R, Di Rosa M, Riccardi GR, et al. Design and development of a scale for evaluating the acceptance of social robotics for older people: the robot era inventory. Front Neurorobot 2022;16:883106. [doi: 10.3389/fnbot.2022.883106] [Medline: 35874107]

Abbreviations

AI: artificial intelligence
CASP: Critical Appraisal Skills Programme
HCAI: human-centered artificial intelligence
HCD: human-centered design
OSF: Open Science Framework
PICOS: Patient, Intervention, Comparator, Outcome, and Study type
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews
SANRA: Scale for the Assessment of Narrative Review Articles
TRAS: Trustworthiness and Reliability Assessment System

XSL•FO
RenderX

UCD: user-centered design
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by A Choudhury; submitted 09.10.24; peer-reviewed by A Jafarizadeh, D Li; revised version received 06.03.25; accepted 17.03.25; published 28.05.25.

Please cite as:
Bevilacqua R, Bailoni T, Maranesi E, Amabili G, Barbarossa F, Ponzano M, Virgolesi M, Rea T, Illario M, Piras EM, Lenge M, Barbi E, Sakellariou G
Framing the Human-Centered Artificial Intelligence Concepts and Methods: Scoping Review
JMIR Hum Factors 2025;12:e67350
URL: https://humanfactors.jmir.org/2025/1/e67350
doi:10.2196/67350

© Roberta Bevilacqua, Tania Bailoni, Elvira Maranesi, Giulio Amabili, Federico Barbarossa, Marta Ponzano, Michele Virgolesi, Teresa Rea, Maddalena Illario, Enrico Maria Piras, Matteo Lenge, Elisa Barbi, Garifallia Sakellariou. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 28.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

On the Necessity of Multidisciplinarity in the Development of at-Home Health Monitoring Platforms for Older Adults: Systematic Review

Chris Lochhead, BSCE, MSc; Robert B Fisher, BSc, MSc, PhD

Advanced Care Research Centre, University of Edinburgh, Old College, South Bridge, Edinburgh, United Kingdom

Corresponding Author:
Chris Lochhead, BSCE, MSc
Advanced Care Research Centre, University of Edinburgh, Old College, South Bridge, Edinburgh, United Kingdom

Abstract

Background: The growth of aging populations globally has increased the demand for new models of care. At-home, computerized health care monitoring is a growing paradigm, which explores the possibility of reducing workloads, lowering the demand for resource-intensive secondary care, and providing more precise and personalized care. Despite the potential societal benefit of autonomous monitoring systems when implemented properly, uptake in health care institutions is slow, and a great volume of research across disciplines encounters similar common barriers to real-world implementation.

Objective: The goal of this systematic review was to construct an evaluation framework that can assess research in terms of how well it addresses already identified barriers to application and then use that framework to analyze the literature across disciplines and identify trends between multidisciplinarity and the likelihood of research being developed robustly.

Methods: This paper introduces a scoring framework for evaluating how well individual pieces of research address key development considerations using 10 identified common barriers to uptake found during meta-review from different disciplines across the domain of health care monitoring. A scoping review is then conducted using this framework to identify the impact that multidisciplinarity involvement has on the effective development of new monitoring technologies. Specifically, we use this framework to measure the relationship between the use of multidisciplinarity in research and the likelihood that a piece of research will be developed in a way that gives it genuine practical application.

Results: We show that viewpoints of multidisciplinarity; namely across computer science and medicine alongside public and patient involvement (PPI) have a significant positive impact in addressing commonly encountered barriers to application research and development according to the evaluation criteria. Using our evaluation metric, multidisciplinary teams score on average 54.3% compared with 35% for teams made up of medical experts and social scientists, and 2.68 for technical-based teams, encompassing computer science and engineering. Also identified is the significant effect that involving either caregivers or end users in the research in a co-design or PPI-based capacity has on the evaluation score (29.3% without any input and between 48.3% and 36.7% for end user or caregiver input respectively, on average).

Conclusions: This review recommends that, to limit the volume of novel research arbitrarily re-encountering the same issues in the limitations of their work and hence improve the efficiency and effectiveness of research, multidisciplinarity should be promoted as a priority to accelerate the rate of advancement in this field and encourage the development of more technology in this domain that can be of tangible societal benefit.

(JMIR Hum Factors 2025;12:e59458) doi:10.2196/59458

KEYWORDS

multi-disciplinarity; gait assessment; machine learning; at-home health monitoring; older adults; elderly; artificial intelligence; AI; gait; development; health monitoring; monitoring; systematic review; monitoring system; barriers; caregiver; efficiency; effectiveness

Introduction

Background

With the general increase in global life expectancy and the ever-increasing ratio of retirement-age to working-age people, research in the scope of health and social care for older people is greatly increasing. As people live longer and the aging population grows, governments and international organizations are recognizing the need for a paradigm shift in the way we care for our older populations [1]. Current methods of care typically include a social care system to look after the most vulnerable and frail older adults, but as the size of this population increases at a faster rate than the working population, these institutional

XSL•FO
RenderX

models of care at the country level become progressively less practical and more costly. In response to this growing crisis, one solution researchers have sought to apply is the use of novel health monitoring technologies to tackle the problem of labor shortages in health care systems and to improve productivity and efficiency in care [2-11]. This can be in direct terms, for example monitoring systems that assist care professionals carrying out their duties within care settings [12] and also indirectly with at-home monitoring devices designed to reduce the occurrence of injuries by, for example, assessing gait [13,14] and predicting fall events [3,15], with the goal of reducing the potential burden on secondary care and ultimately reducing the burden on the care-home sector and keeping people at home, healthy, and independent for longer [9,16,17].

Research Hypothesis and Goals

The goals of this work can be summarized by the following 2 research questions:

1. Can we construct an evaluation framework for at-home health monitoring research and justify that a positive score in said framework broadly correlates with a higher likelihood of the research being effective in real-world use?
2. Can we identify the trends if they exist, between the method of research (single-discipline, multidiscipline, and the use of PPI or co-design) and the increase or decrease in effectiveness of said research?

In the context of this paper, “effective research” can be defined as research that is more efficient by being less prone to making oversights already identified in the existing literature across disciplines. We assert that research that addresses the already existing issues in the literature (and thereby scoring highly on our proposed metric) leads to more effective research and ultimately leads to products and platforms more likely to be usable in real-life applications (refer to the Evaluation Framework Justification in the Methods section for evidence of this).

To answer the research questions, this work is split into 2 phases, a scoping phase and evaluation phase. In the scoping phase, an overview is provided of the various systematic reviews across disciplines in the field of at-home health care monitoring for older adults. In doing this, we synthesize an evaluation framework based on the consensus and concatenation of these studies to measure the degree to which individual pieces of research encounter similar common problems. By defining this framework, observations can be made about the degree to which these problems are acknowledged by different types of research teams operating with or without multidisciplinarity, which types of issues are likely to be identified by which types of research, and how stakeholder involvement can improve the likelihood of developing an effective product. To justify this evaluation criteria and to address the first research question, examples of existing technologies in commercial use are positively evaluated by this approach to assert the positive relationship between a high evaluation score and the generation of more effective research, considerate to existing identified issues with real-world use (refer to the Methods section). To our knowledge, there is currently no evaluation methodology designed to measure research application effectiveness, based on multidisciplinary metrics.

In the evaluation phase. a review of individual works in the application of computerized at-home health-monitoring systems is then presented, with these works evaluated against the metric developed in the scoping phase, and with the results analyzed to address the second research question (refer to the Results section). For this evaluation phase, a total of 350 papers were found from the IEEE Explore, PubMed, and ArXiv databases, of which 60 papers were ultimately used, with the inclusion constraints being that the papers had to be individual pieces of original research relating either to the development of technology associated with the various applications of at-home older adult health monitoring or a review into the effectiveness of existing at-home health monitoring applications. This is including, but not limited to, development and deployment of disease diagnosis and progression analysis, fall detection and prevention, lifestyle monitoring, vital-sign monitoring, and smart-home systems. Candidate papers must present methods that have the prospect of or are already actively being tested in an at-home environment in whole or in part.

For this evaluation phase, a total of 350 papers were found from the IEEE Explore, PubMed, and ArXiv databases, of which 60 papers were ultimately used, with the inclusion constraints being that the papers had to be individual pieces of original research relating either to the development of technology associated with the various applications of at-home older adult health monitoring or a review into the effectiveness of existing at-home health monitoring applications. This is including but not limited to the development and deployment of disease diagnosis and progression analysis, fall detection and prevention, lifestyle monitoring, vital-sign monitoring, and smart-home systems. Candidate papers must present methods that have the prospect of or are already actively being tested in an at-home environment in whole or in part.

The Methods section contains the scoping phase and provides an overview of the literature in health monitoring technologies, constructing from them a concise evaluation framework in the form of a 10-point checklist criteria for research application effectiveness, based on consensus drawn from these works. The Results section presents the evaluation phase, reviewing individual pieces of research across disciplines using this evaluation framework. The interaction between multidisciplinarity and stakeholder involvement through PPI and co-design and the effectiveness of research are also analyzed here. Finally, the Discussion section summarizes the findings of this research and makes recommendations for how future health monitoring research should be conducted in order to improve the effectiveness of research in this field.

Research Contributions

The contributions of this paper can be summarized as follows:
:

1. A review of the recent literature of at-home health care monitoring. This is achieved by providing a 2-stage review of the literature, with the first (scoping phase) being a meta-review of reviews in the relevant literature, and the

XSL•FO
RenderX

second (evaluation phase) being a review of individual works across multiple disciplines in the area of at-home health monitoring.

2. The introduction of a comprehensive evaluation framework for assessing the likely real-world effectiveness of health monitoring application research, based on 10 key issues consistently identified from analysis of 15 literature reviews in the field across disciplines.
3. The identification and demonstration of various links between the manner of research and the effectiveness of research. We find that multidisciplinarity has a consistently positive effect in this respect, reflected by higher scores on the evaluation framework. We also find a convincing increase in scores when PPI and co-design methods are used.

The goal of this research is to concretely illustrate the necessity of multidisciplinarity for the success of at-home medical technology development and deployment and provide a framework for future research to reference when assessing the usefulness of their own applications.

Methods

Scoping Phase: Overview of at-Home Health-Monitoring in Older Adults

To identify the most common barriers to uptake in this field, an overview of 12 systematic and other review papers from 2014 onwards was conducted to identify and collate the common issues within some of the common subdisciplines of at-home health care monitoring: encompassing any autonomous monitoring methods used in a domestic context such as camera-based applications, wearable sensors, remote sensors, and ensemble smart-home systems. Table 1 provides a concise summary of the main reviews investigated, and the presence of each of the commonly identified barriers among them. Broken down in Table 1 are the makeup of the teams involved in the research, multidisciplinary, technical, or application-based. In the context of this work, “Technical” concerns research conducted by engineering or computer science teams, with “Application” making up all other types of research team, but predominantly teams in the fields of medicine and social science. In total, 10 core barriers are identified across these reviews and synthesized to be inclusive of all the barriers identified in all the systematic reviews evaluated in this section. An exact breakdown of the presence or absence of the 10 points in each review is available in the supplementary materials (refer Table S1 in Multimedia Appendix 1).

XSL•FO
RenderX

Table . Overview of systematic reviews and their evaluation score according to the evaluation metric.

Paper title and reference	Description	Research team	Score (%)
Are Active and Assisted Living applications addressing the main acceptance concerns of their beneficiaries? [18]	Overview of opinions of older adults regarding concerns with ambient assisted living technologies.	Multidisciplinary	50
A critical review of smart residential environments for older adults with a focus on pleasurable experience [5]	Review of older adults in focus groups for a variety of smart home applications.	Multidisciplinary	30
Smart homes and home health monitoring technologies for older adults: A systematic review [6]	Investigation of the abilities of various smart-home technologies, alongside feasibility and technical limitations.	Application	70
Health Monitoring Using Smart Home Technologies: Scoping Review [8]	Review of smart home environments, specifically on existing study design limitations alongside technical limitations.	Application	50
Older persons have ambivalent feelings about the use of monitoring technologies [2]	Series of 5 focus groups of older adults attempting to build consensus on the concerns of smart home implementations.	Application	60
Older adults' perceptions of technologies aimed at falls prevention [3]	Systematic review of focus group-based studies to build consensus on the main factors hindering at home uptake.	Application	40
Unobtrusive sensing and wearable devices for health informatics [11]	Overview of 4 main sensor-based monitoring technologies including relative benefits and drawbacks.	Technical	40
Wearable sensors for remote health monitoring [7]	Overview of specifically wearable health monitoring technologies, mostly centered on technical benefits and limitations regarding data.	Technical	40
Unobtrusive health monitoring in private spaces: The smart home [10]	Overview of smart home. Applications with a focus on perceived "unobtrusive" applications.	Technical	20
Detection and assessment of Parkinson's disease based on gait analysis: A survey [14]	Overview of gait assessment monitoring for age-related disease detection using a variety of classical and ML[a]-powered monitoring technology.	Technical	0
Remote patient monitoring using AI[b] [9]	Overview of classical and deep learning-based AI applications in primarily at-home patient monitoring.	Technical	20
Factors Determining the Success and Failure of eHealth Interventions [16]	Overview of smart home. Applications with a focus on discovering the key factors behind the success or failure of med-tech applications.	Multidisciplinary	70

[a]ML: machine learning.

[b]AI: artificial intelligence.

"Research Team" denotes whether the teams conducting the research were technical or medical-based. See further in this section for a breakdown of the 10 common identified points and refer to the supplementary material for a specific breakdown of which barriers are present in each. The score value was computed by aggregating which of the 10 features were addressed in each review. Beyond these key reviews from which the evaluation framework was constructed, there are numerous other reviews and surveys whose findings are inclusive of the framework outlined at the end of this section (refer to Table S2 in Multimedia Appendix 1 for a full breakdown of the 13 reviews). Only the core 12 are included here for brevity, with these 12 being selected as they encompass all of the core at-home health care methodologies, include examples of all 10

XSL•FO
RenderX

key points that make up the framework, and represent a roughly equal variety of research teams, namely multidisciplinary teams (n=3), computer science (n=3) engineering (n=2), medical (n=2), and social science (n=2).

Table . Scoring of 2 current at-home monitoring projects being developed with the intention of widespread use. The factors are numbered corresponding to the attribute list given in the attribute list above. Y or N indicates "yes" or "no" as to the presence of the factors.

Project Name	Description	Factors (Y or N)									
		1	2	3	4	5	6	7	8	9	10
SPHERE[a] [19-21]	Smart-home system for behavior monitoring, under development for several years and the feature of multiple research studies.	Y[b]	Y	Y	Y	Y	Y	Y	Y	N[c]	Y
HALLEY[d] [22]	Internet of things-based Smart-home development project currently in commercial development.	Y	Y	Y	N	Y	Y	Y	N	N	Y

[a] SPHERE is a multi-year smart-home development project based in England which has been involved both in commercial smart-home production as well as research.

[b] Y: yes.

[c] N: no.

[d] HALLEYASSIST is a company based in Australia involved in developing multiple healthcare monitoring devices for a smart home context.

The study by Ghorayeb et al [23] presents a number of insights in their review of the literature regarding older adults using smart-homes, such as identifying solutions to the problems in most research applications. For example, they assert that gradual introduction of smart-home technology combined with the ability to “pause” it at will to provide “emotional release” is highly desired in end users for a more pleasant experience. They also find that more tech-literate people tend to have less concerns with the technology due to an improved understanding of the data being collected and the privacy risks involved, if any. One other concern identified was transmission of data and the insecurities associated with this, leaning to a user-preference that data should be handled manually instead of through remote connection. These findings are all further reinforced by the findings in [24], which similarly conduct interview sessions with 20 older adults regarding which aspects of health-monitoring systems concern them. These findings inform points 4 and 5 in the evaluation metric, defined in Methods section.

Regarding data access, the study by Robinson et al [25] finds that older adults could in some cases want control of data access to withhold the data from certain parties, namely their family and friends, for fear of burdening them as well as the parallel desire some express that they do not want to be micromanaged by their loved ones or health care providers, informing points 6 and 8 of the framework. Cost was also an identified challenge (informing point 3), for example focus group attendees in [2] were concerned by the cost of long-term use of monitoring devices, with the implication being that they would have to buy them outright.

The study by Hawley-Hague et al [3] discovers that when surveyed, there is a broad perception that older people view new advancements in smart-home related technology as good but “unnecessary for them,” as they don’t perceive themselves as being “unhealthy enough” to merit using what they see as a drastic action toward greater care. Mann et al [26] found in their study of 661 older people that a slim majority (56.3%) perceived smart-home technology as not being of use to them specifically.

XSL•FO
RenderX

Ghorayeb et al [23] identify a series of factors through their investigation with older participants which are crucial to account for when considering the integration of new technology into real life. These factors range from societal stigma to technical reliability, covering points 1, 2, 3, and 6 in the evaluation framework. Research in studies by Boström et al and by Boström et al [2] and Liu et al [6] also point to the issues surrounding the implication of autonomous monitoring being a decline in human contact, addressed by point 7. In the systematic review conducted by Ghorayeb et al [23], observations were made regarding the progression of smart-home technologies. Regarding patient acceptability, they note that less than half of the papers reviewed take into consideration the acceptability of their technology where privacy is concerned: with the focus of the paper instead being explicitly about the functions of the novel technology. Some steps taken toward improving privacy in certain papers included encrypting collected data [27] and locking data access behind authorization [28].

Regarding health care professionals, their main concerns with smart-home technology and health care monitoring concern the feasibility of use. Unlike patient acceptability, acceptability by health care professionals is underresearched, where many papers will either focus on the technology being developed and not address it in application or they will focus only on the opinions of the end users and how they will use and accept the technology [29]. This finding is echoed in the lack of surveys found in the scoping review in the Results section, where the focus is on health care professionals rather than patients. Of the 60 papers in the scoping review, only 7 made any explicit use of external medical caregivers in their research development.

Evaluation Framework Criteria

Across the review papers both in Table 1 and the other reviews referenced in this section, the following list of 10 factors were collated to encompass all the common considerations identified when designing and implementing novel monitoring technology for use with older adults in a domestic environment. They are segmented by category of concern (technical, application, or multidisciplinary), where the technical are purely engineering or computer-science implementation concerns and application are concerns involving human interaction with the technology and any adjacent concerns important to medical and social science-based disciplines. The framework is as follows (Textbox 1).

Textbox 1. The evaluation framework

- Usability-technical: concerning the use of the technology and how feasible it is to be used by caregivers, health care professionals, and end users.
- Accessibility-technical: this concerns the ease of use by laypeople of the computer technology and the barriers for entry in terms of effectively using the technology.
- Reliability-technical: covers issues relating to the long-term viability of the application, for example, is it expensive, does it require upkeep, charging or maintenance.
- Control-multidisciplinary: concerns the level of access both patients and health care staff should have to the application and the data it records.
- Privacy-multidisciplinary: constitutes issues relating to the intrusiveness of the data being collected, and the manner in which it is collected.
- Stigma-application: anything relating to the societal pressures and negative connotations people may feel by using monitoring technologies for the purposes of personal health.
- Lack of human response-application: concerns the issue of the perception that increased autonomous monitoring would result in a decrease in person-to-person interaction as a result.
- Burden to others-application: regarding the perception of older adults that additional at-home monitoring is representative of applying additional pressure on their caregivers.
- Lack of perceived need-application: concerns a commonly identified phenomenon in the literature that people have a tendency to think a technology is useful but unnecessary for them personally.
- Affordability-application: this concerns the cost both on a personal and institutional level to implement solution applications in real life.

Evaluation Framework Justification

To demonstrate the descriptive power this evaluation system has to the likelihood of success in application, Table 2 provides an overview of papers concerning 2 technologies being developed for commercial use (Sphere [21] and HalleyAssist [22]), and the degree to which the prior research and development of these systems adhere to the criteria in the evaluation metric, with scores being calculated based on all existing literature regarding each technology, as opposed to individual articles. Both have demonstrated a commitment to addressing barriers relating to both technical and human elements, addressing 90% and 70% of the evaluation metrics respectively (the average score across the scoping review in the Results section is 36.2% for comparison). They also both were developed by a multidisciplinary research team from computer science and signal processing to nursing, geriatric medicine and social care, with both systems now being in varying stages of commercialization. The relationship between the absence of certain metrics such as multidisciplinarity and the lack of consideration for common barriers to uptake is further concretely illustrated during the analysis at the beginning of the Results section. While these are only 2 data points and cannot be said to constitute a trend on their own, it can be asserted that they provide a strong positive indication that successful applications are likely to score highly on the introduced metric.

XSL•FO
RenderX

Results

Evaluation Phase: Scoping Review and Parameters

Figure 1 illustrates the selection process for papers considered for the scoping review. From the 3 databases, the following search term was used:

((healthcare monitoring) AND (older adult)) AND ((home) OR (at-home) OR (domestic)) AND ((obtrusive) OR (unobtrusive) OR (intrusive)) AND ((machine learning) OR (AI) OR (artificial intelligence))

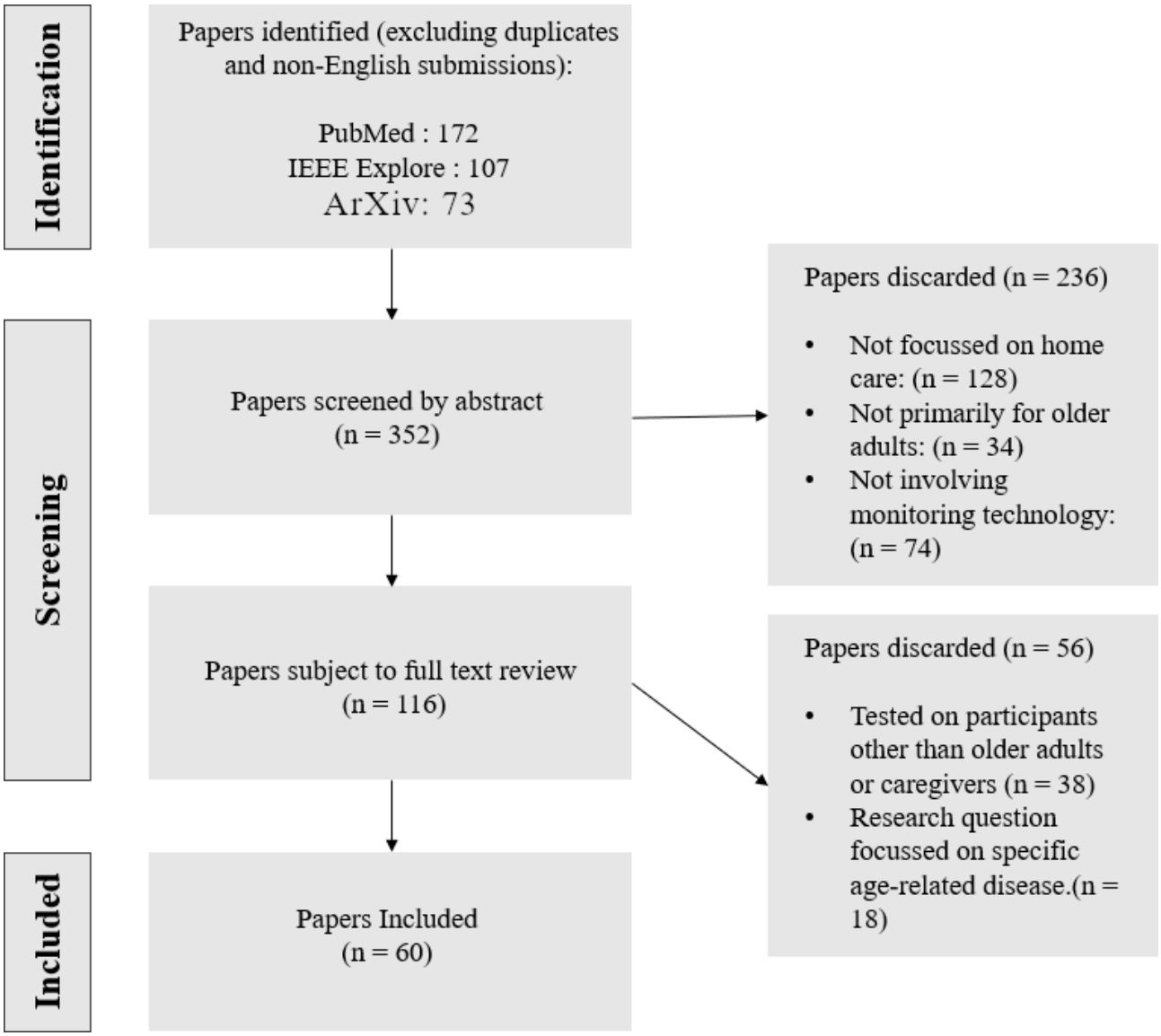
with the exception of the ArXiv database, where the

((obtrusive) OR (unobtrusive) OR (intrusive))

section was omitted for a larger range of initial texts. Language was restricted to English and the date was restricted to 2014 or later, with the search commencing on January 7-24, 2024. Excluded were papers that did not in whole or in part reference older adult monitoring as the purpose of their application or review, and papers where the application was specifically designed for care homes or hospitals. The initial search yielded a total of 352 papers. Paper title and abstract analysis excluded 236 papers, which resulted in 116 remaining, with the final count being 60 papers after a full-paper review resulting in 56 more deletions (Figure 1).

The following section is broken down by technical and application-based research, following the definitions for both established in the beginning of the Methods section. Papers defined as multidisciplinary are those in which the research team consist of at least 1 person from each category.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram: screening process illustration for the scoping review portion of this research.



Technical Research

The majority of research in this space relies either on traditional or neural network–based machine learning methods to make their systems effective and largely automatic. Researchers in [30] construct a machine learning model that can differentiate between regular, prefrail, and frail gait within a population of 50 older adults at an accuracy of 88.5% based on gait data

XSL•FO
RenderX

collected from wearable accelerometer signals. They rely mostly on “traditional” methods of machine learning: that is, machine learning methods that don’t use deep neural networks.

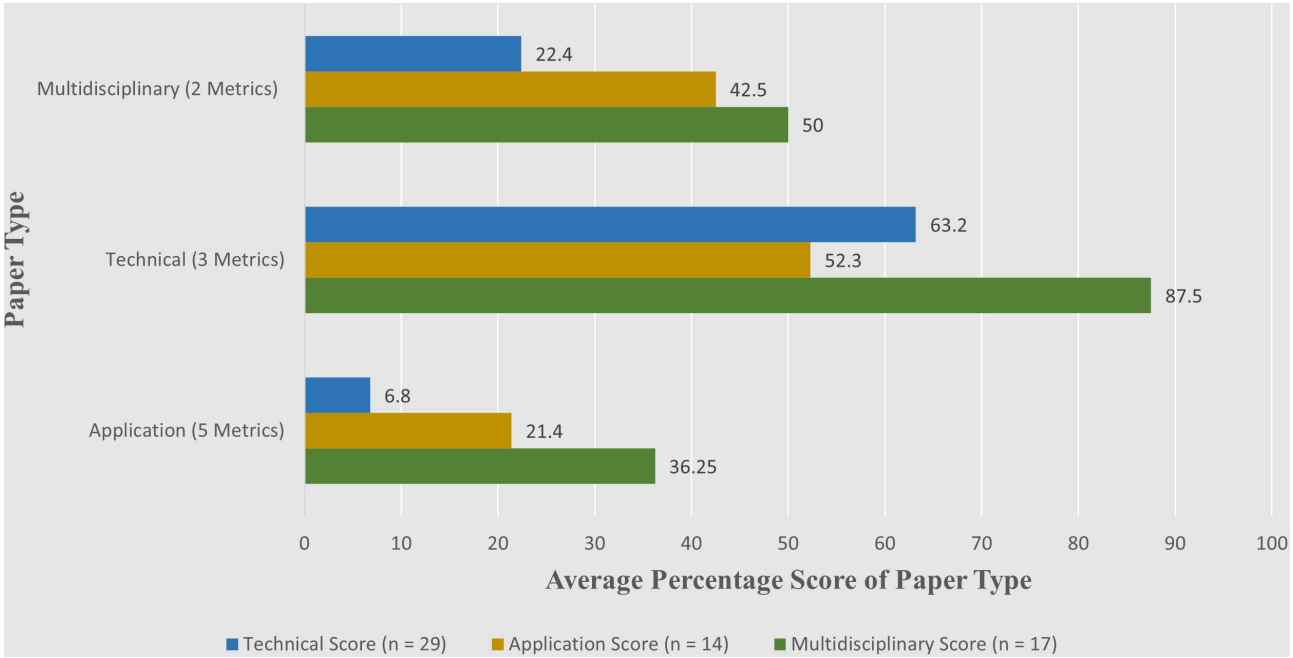
With vision-based applications, on the other hand, deep neural networks, especially convolutional networks designed specifically for processing image data, are frequently used. Researchers in the study by Lin et al [31] use a novel 3D convolutional neural network (CNN) to process gait silhouette images from the CASIA-B dataset [32] to achieve a person-identification accuracy of 97.6%. While the question of person identification through gait can be fairly trivially solved with modern CNN frameworks, the question of gait analysis, or assessing health requires a more nuanced approach and is still an active research area.

Researchers in [33] use a Spatiotemporal Graph Convolutional Network (ST-GCN) [34] framework to assess Parkinson disease severity as classified under the Unified Parkinson’s Disease Rating Scale [35]. The ST-GCN is an extension of the standard CNN model that works specifically on skeletal graph data as opposed to raw images, and works by processing input data with sequential attention to both the spatial and temporal dimensions of the input. As gait assessment is a far more complicated problem than simple person-identification or even traditional action recognition, they achieved 53% F_1-score using a dataset of 53 individuals with Parkinson disease.

Recently, in the study by Yin et al [36], the Spatio-temporal joint adjacency GCN (STJA-GCN) was developed that uses 3 input streams for joints, bones and velocities, a novel joint attention module to emphasize spatial data and a simplified skeletal graph input architecture to achieve state-of-the-art results (93.17% and 92.08%) on the domain of recognizing and classifying different types of abnormal gaits compared to other ST-GCN–based methodologies. This was tested on a pair of synthetic gait datasets introduced by the researchers, totalling 22 participants and 8600 instances of gait data collected from a total of 7 sensors across the 2 datasets.

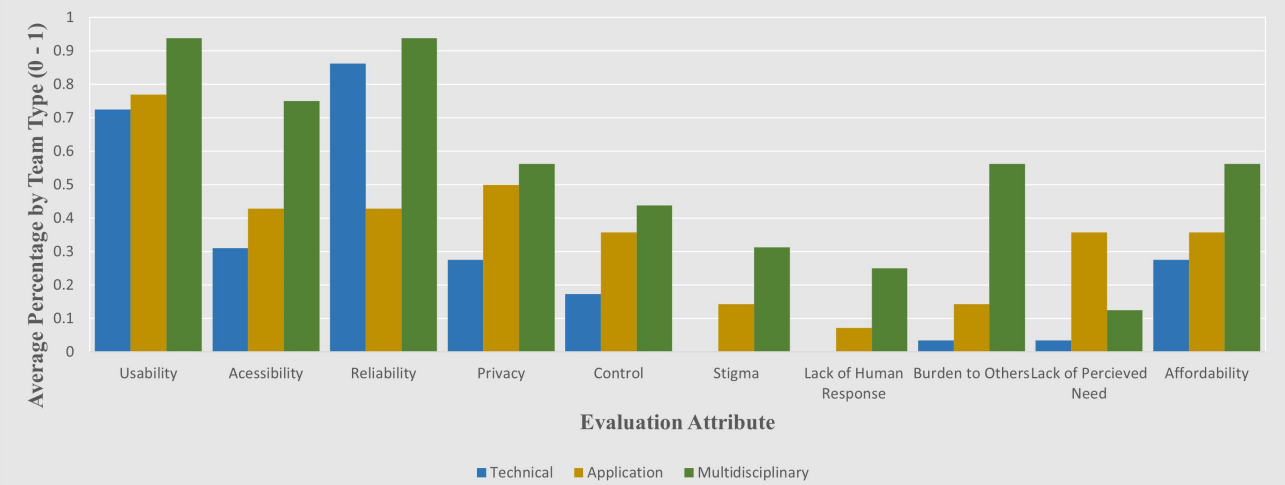
Demonstrably, machine learning methods are extremely capable in multiple health care monitoring applications, necessitating the inclusion of highly technical fields in this type of research. The issue with the papers in this field, however, is that they score poorly on the proposed evaluation framework due to an overt focus on technical innovation at the expense of applicational concerns as well as scoring consistently lower than papers in the applicational and multidisciplinary categories on all but explicitly technical metrics (Figure 2). This analysis indicates a narrow focus almost purely on the effectiveness of the technology at the expense of an omission of application considerations, such as cost or perceived need by end users and medical stakeholders (refer to Figure 3 for a per-category breakdown). End users in this context refers both to patients and health care professionals who would be involved with the use of the technology (such as general practitioners, nurses, physiotherapists, occupational therapy and care workers). Stakeholders, on the other hand, refers to family members, friends or dependents of the patient using the technology. To take a more concrete example of this narrowed focus, research by Brunzini et al [13] is explicitly geared toward the development of a decision-support algorithm for health care professionals, complete with bespoke visual aids. However, there is no evidence of consideration of the requirements of health care professionals who would prospectively use such a system (in this case namely GPs, geriatricians, and nurses) being used in the design of the system itself. Similar issues were found in several studies where technical teams sought to develop assistive tools for caregivers or patients without direct involvement of prospective caregivers or end users, ranging from tools designed for doctors and nurses to carers, physiotherapists, and occupational therapists [37-41].

Figure 2. Illustration of evaluation performance per-evaluation category. Shown here is the impact that the use of multidisciplinarity has on the likelihood of each of the 10 metrics being taken into consideration.



XSL•FO
RenderX

Figure 3. Breakdown of the average presence of each evaluation attribute by research team type.



Application and Multidisciplinary Research

Of the surveyed technical papers (n=29), they scored an average 6.8% on the application-centered metrics in the proposed evaluation framework (Figure 2). In contrast, the application papers (n=14) mainly authored by health care researchers scored 21.4% on application-based metrics. While this indicates that application focused papers typically address more concerns with real-life deployment of research solutions, there is also a notable trend that this focus on application comes at a slightly increased tendency to neglect technical aspects of research (with both technical and multidisciplinary papers outperforming application-based papers on technical metrics). A clear trend can be seen in multidisciplinary papers consistently outperforming technical or application-based papers, even on the metrics specific to their own discipline (Figure 3).

As indicated in the second research question in the Introduction, we speculate that there exists a relationship between the inclusion of different stakeholders in PPI and co-design with the effectiveness of the resultant applications and research. As indicated by Table 1, this type of inclusion is usually implemented using focus groups. In this section, we highlight the typically moderate-to-high scores where this stakeholder input is prioritised, whilst also outlining other areas of shortcomings that follow from lack of multidisciplinarity in other, mainly technical domains.

In the study by Mireles et al [42], a total of 75 health care professionals (mostly nurses, n=47) are interviewed as a means of investigating the effectiveness of 3 different eHealth technologies for at-home monitoring of older adults with chronic conditions. Scoring 40%, this research focuses on the implications of e-health solutions as they affect end users, namely stakeholders, with less consideration for the technical implications of their recommendations. Research in the study by Bjornsdottir and Ceci [43] is similar, conducting qualitative interviews with both health care professionals and also patients. By including the patients and getting their insight, this research scores higher (60%) but suffers on the same technical metrics, highlighting the shortcomings in the systems they are evaluating without explicit consideration of the technical implications of their conclusions. In [44], 5 focus groups were conducted with both older adults and caregivers; namely family and care staff, to help synthesize a consensus on the design of future at-home sensor-based systems, with a focus on acceptability, respect for privacy and how to best provide control of care to the end users themselves. Scoring 50%, the authors neglect aspects such as system reliability at the expense of greater privacy, and other application-driven barriers such as societal stigma and the threat of greater isolation resulting in the deployment of remote monitoring technology. Expanding on this focus-group centered research [45], conduct a series of interviews with both formal and informal caregivers to identify the key parameters which need to be addressed to achieve effective at-home lifestyle monitoring systems. Scoring 60%, their strengths come from the inclusion of both medical and end user input, however they similarly lack extensive consideration for how the parameters being set by stakeholders would affect the performance of the monitoring systems themselves. Itoh et al [46] investigated issues of application from a medical perspective but score only 40%. Not only does their investigation consist only of testing an existing technology without the intention of developing said technology further, they also concern their research only with application concerns as far as the caregivers and medical stakeholders are concerned, with no involvement being afforded to the end users and patients. As a result, issues around stigma and even privacy are completely unaddressed.

As can be seen in Table 3, there is a clear lack of involvement in research and application design from the medical community and caregivers. Overall, 13.7% and 64.2% of single-discipline papers involve end user participation versus a mere 14.2% and 0% exhibiting caregiver or medical participation in the development of the research (for application and technical based papers. respectively). While the likelihood of involving both types of stakeholders (caregivers and end users) is more than double when multidisciplinary teams are involved (14.2% vs 37.5%), this is not conclusive due to the relatively small number of papers in each category when divided by the inclusion of different stakeholder types (end user and caregiver). Likewise, there is a definite trend across all research team types that exhibit lower evaluation scores when neither caregivers nor end users participate in the design or evaluation of research. However, discerning the differing impact between the 2 stakeholder types is difficult due to the small positive sample size (with only 2%

XSL•FO
RenderX

separating end user only and caregiver only scores for both research team types, refer to Table 4). The trend in the research, especially in survey-style research seems to be to focus on end users rather than caregivers or other medical stakeholders with only 7 of 60 papers involving medical stakeholders and 22 of 60 involving end user input.

Table . The presence or absence of caregivers and end users in the co-design of research across the review, split by research team type.

Team type	Caregivers (%)	End users (%)	Total (%)
Multidisciplinary[a]	37.5	56.25	*37.5*
Application[a]	38.4	64.2	14.2
Technical[a]	0	13.7	0

[a] Values are not mutually exclusive and some papers have both caregivers and end users involved. As a result, values may not necessarily add up to 100%.

Table . The average evaluation score by research team type when caregivers and end users are included or not in the research.

Team type	Neither (score %)	Caregivers (score %)	End users (score %)	Caregiver and End users (score %)
Multidisciplinary	35.7	70	68	70
Application	27.5	40	42	43
Technical	24.8	0	35	0

The strongest scores using the evaluation framework come from those where interdisciplinary teams are used (average score of 5.43 versus 3.5 and 2.68 for application-based and technical teams respectively), with a distinct trend shown in Table 4 indicating the presence of caregiver and end user involvement in a co-design or PPI capacity leads to a stronger evaluation score.

In general, as shown in Figure 3, different categories of research team tend to neglect certain aspects identified in the evaluation framework. Technical papers struggle with “Stigma” and “Human Response,” and application-based papers, mostly led by medical teams often omit factors such as “Accessibility“ and “Burden to Others” from their consideration. Both tend to struggle particularly with addressing societal stigma, possibly due to the varying levels and types of stigma for at-home monitoring from researcher’s home countries, or potentially due to considerations for more general issues not directly related to their specific research being deemed as outside of the scope of their work.

A common theme across all of these works is the tendency to make assumptions on the necessity of their application in the eyes of end users (Figure 3), as the lack of perceived need is seldom addressed regardless of the make-up of the research team. Technical issues are generally well addressed such as usability and reliability, however broader human issues such as societal stigma and feelings of end users becoming a burden, or the technology itself being a burden to use by caregivers are likewise rarely addressed. Across the entire evaluation phase, there is a consistent trend of greater representation of every attribute when multidisciplinary teams are involved in the research, with specifically technical teams being especially susceptible to a lack of focus on application-based concerns.

Discussion

Principal Findings

The research presented here concludes that a root problem underlying the issues affecting real-life uptake of research applications is the lack of multidisciplinarity in the research. We find that the developed evaluation metric is effective for quantifying the likelihood an application will transition into further development and effective deployment in real-world use. We also identified a series of common themes across the literature, namely the consistent underrepresentation of end users and other stakeholders in research, a distinct lack of multidisciplinarity in technical applications and the consistent underappreciation of societal factors such as stigma and human contact when developing at-home medical technologies.

Using a novel evaluation metric based on the 10 key barriers to uptake identified in the literature, a concrete trend can be observed that that multidisciplinarity between engineers, computer scientists, medical experts, and social scientists, alongside co-design or PPI-based inclusion of caregivers (such as family, friends, or dependants) and end users (health care professionals or patients) is highly beneficial to the effective development of technology that has direct practical application and tangible social benefit. The evaluation framework itself was also justified by the demonstration of a trend of extremely high evaluation scores on research projects that had reached the point of successful practical application, namely HalleyAssist and SPHERE. Using this evaluation metric, we find an increase of between 19.3% - 27.5% for multidisciplinary teams versus single discipline, and an increase of between 7.4% - 19% when research includes the use of caregivers or end users in a PPI or co-design capacity.

XSL•FO
RenderX

Limitations and Future Work

The main limitation of this study is that the results can only demonstrate the lack of multidisciplinarity is a root issue, not necessarily the root issue. For example, many extenuating factors not included in the research itself can explain why effective technology was researched but not developed, such as personal material conditions for the researchers, leading them to be uninterested or unable to develop effective research further. One other limitation is in the methodology of the framework itself. Ways to improve the descriptive power and nuance of the framework could be to introduce more criteria, break the criteria down into further subdisciplines or introduce criteria weighting so that criteria deemed more “vital” or less common are weighted higher when calculating the score. Had more resources been available, a broader review in both phases could have been conducted, using a greater number of papers from a greater number of databases, an especially important factor given the number of different disciplines relevant to this type of research.

In future, a larger study could be conducted on the literature to further verify the effectiveness of the evaluation framework, including the application of the framework to more smart-home technologies in commercial development. The framework itself could also benefit from being imbued with greater complexity for example, additional factors or a bespoke weighting of them to make certain factors more important to the success of a technology in application, and thus scoring a higher score. To our knowledge there is no comparable work in quantitatively assessing the effects of multidisciplinarity on research in at-home health monitoring, so in general, more research is needed to saturate the domain and allow for trends to be more concretely identified. As HalleyAssist and SPHERE are leading examples in the area of in-home monitoring of ageing singles, it would be good to follow up their development and experiences in the future, to see how well the assessment undertaken here was predictive of their successes.

Acknowledgments

This research was funded by the Legal and General Group (research grant to establish the independent Advanced Care Research Centre at University of Edinburgh).

Authors' Contributions

CL and RBF were both involved in the conception of the paper, with CL carrying out the literature reviews and being the primary drafter of the manuscript. RBF was also responsible for supervisory support and redrafting and revising the manuscript. Both authors have given final approval for the manuscript presented for publishing.

Conflicts of Interest

This research was funded by the Legal & General Group (research grant to establish the independent Advanced Care Research Centre at University of Edinburgh). The funder had no role in the conduct of the study, interpretation or the decision to submit for publication. The views expressed are those of the authors and not necessarily those of Legal & General. To the authors knowledge there are no conflicts of interest during the undertaking of this research.

Multimedia Appendix 1
Supplementary materials referred to in the manuscript. There is also a live hyperlink.
[XLSX File, 51 KB - humanfactors_v12i1e59458_app1.xlsx]

Checklist 1
PRISMA checklist.
[PDF File, 182 KB - humanfactors_v12i1e59458_app2.pdf]

References

1. United Nations. World Population Ageing 2019 Highlights: United Nations; 2019. [doi: 10.18356/9df3caed-en]
2. Boström M, Kjellström S, Björklund A. Older persons have ambivalent feelings about the use of monitoring technologies. Info Technol Disab 2013;25(2):117-125. [doi: 10.3233/TAD-130376]
3. Hawley-Hague H, Boulton E, Hall A, Pfeiffer K, Todd C. Older adults’ perceptions of technologies aimed at falls prevention, detection or monitoring: A systematic review. Int J Med Inform 2014 Jun;83(6):416-426. [doi: 10.1016/j.ijmedinf.2014.03.002] [Medline: 24798946]
4. Kaur B, Kumar S, Kaushik BK. Novel wearable optical sensors for vital health monitoring systems-a review. Biosensors (Basel) 2023 Jan 23;13(2):181. [doi: 10.3390/bios13020181] [Medline: 36831947]
5. Lee LN, Kim MJ. A critical review of smart residential environments for older adults with a focus on pleasurable experience. Front Psychol 2019;10:3080. [doi: 10.3389/fpsyg.2019.03080] [Medline: 32038424]

XSL•FO
RenderX

6. Liu L, Stroulia E, Nikolaidis I, Miguel-Cruz A, Rios Rincon A. Smart homes and home health monitoring technologies for older adults: A systematic review. Int J Med Inform 2016 Jul;91:44-59. [doi: 10.1016/j.ijmedinf.2016.04.007] [Medline: 27185508]
7. Majumder S, Mondal T, Deen MJ. Wearable sensors for remote health monitoring. Sensors (Basel) 2017 Jan 12;17(1):130. [doi: 10.3390/s17010130] [Medline: 28085085]
8. Morita PP, Sahu KS, Oetomo A. Health monitoring using smart home technologies: scoping review. JMIR Mhealth Uhealth 2023 Apr 13;11:e37347. [doi: 10.2196/37347] [Medline: 37052984]
9. Shaik T, Tao X, Higgins N, et al. Remote patient monitoring using artificial intelligence: Current state, applications, and challenges. WIREs Data Min Knowl 2023 Mar;13(2). [doi: 10.1002/widm.1485]
10. Wang J, Spicher N, Warnecke JM, Haghi M, Schwartze J, Deserno TM. Unobtrusive health monitoring in private spaces: the smart home. Sensors (Basel) 2021 Jan 28;21(3):864. [doi: 10.3390/s21030864] [Medline: 33525460]
11. Zheng YL, Ding XR, Poon CCY, et al. Unobtrusive sensing and wearable devices for health informatics. IEEE Trans Biomed Eng 2014 May;61(5):1538-1554. [doi: 10.1109/TBME.2014.2309951] [Medline: 24759283]
12. Tao X, Yun Z. Fall prediction based on biomechanics equilibrium using Kinect. Int J Distrib Sens Netw 2017 Apr;13(4):155014771770325. [doi: 10.1177/1550147717703257]
13. Brunzini A, Caragiuli M, Massera C, Mandolini M. Healthy ageing: a decision-support algorithm for the patient-specific assignment of ICT devices and services. Sensors (Basel) 2023 Feb 7;23(4):1836. [doi: 10.3390/s23041836] [Medline: 36850433]
14. Guo Y, Yang J, Liu Y, Chen X, Yang GZ. Detection and assessment of Parkinson's disease based on gait analysis: a survey. Front Aging Neurosci 2022;14:916971. [doi: 10.3389/fnagi.2022.916971] [Medline: 35992585]
15. Kulurkar P, Dixit CK, Bharathi VC, Monikavishnuvarthini A, Dhakne A, Preethi P. AI based elderly fall prediction system using wearable sensors: a smart home-care technology with IOT. Measurement: Sensors 2023 Feb;25:100614. [doi: 10.1016/j.measen.2022.100614]
16. Granja C, Janssen W, Johansen MA. Factors determining the success and failure of eHealth interventions: systematic review of the literature. J Med Internet Res 2018 May 1;20(5):e10235. [doi: 10.2196/10235] [Medline: 29716883]
17. Higuma M, Sanjo N, Mitoma H, Yoneyama M, Yokota T. Whole-day gait monitoring in patients with Alzheimer's disease: a relationship between attention and gait cycle. J Alzheimers Dis Rep 2017 Apr 28;1(1):1-8. [doi: 10.3233/ADR-170001] [Medline: 30480224]
18. Colantonio S, Jovanovic M, Zdravevski E, et al. Are active and assisted living applications addressing the main acceptance concerns of their beneficiaries? preliminary insights from a scoping review. Presented at: Proceedings of the 15th International Conference on PErvasive Technologies Related to Assistive Environments; Jun 29 to Jul 1, 2022; Corfu Greece. [doi: 10.1145/3529190.3534753]
19. Burrows A, Gooberman-Hill R, Craddock I, Coyle D. Sensors for healthcare: would you want them in your home. arXiv. Preprint posted online on 2014arXiv:1408.2476.
20. Morgan C, Craddock I, Tonkin EL, et al. Protocol for PD SENSORS: Parkinson's disease symptom evaluation in a naturalistic setting producing outcome measuRes using SPHERE technology. an observational feasibility study of multi-modal multi-sensor technology to measure symptoms and activities of daily living in Parkinson's disease. BMJ Open 2020 Nov 30;10(11):e041303. [doi: 10.1136/bmjopen-2020-041303] [Medline: 33257491]
21. Zhu N, Diethe T, Camplani M, et al. Bridging e-Health and the internet of things: the SPHERE project. IEEE Intell Syst 2015;30(4):39-46. [doi: 10.1109/MIS.2015.57]
22. Forkan ARM, Branch P, Jayaraman PP, Ferretto A. HalleyAssist: a personalised internet of things technology to assist the elderly in daily living. 2019 Jan 8 Presented at: Hawaii International Conference on System Sciences. [doi: 10.24251/HICSS.2019.507]
23. Ghorayeb A, Comber R, Gooberman-Hill R. Older adults' perspectives of smart home technology: are we developing the technology that older people want? Int J Hum Comput Stud 2021 Mar;147:102571. [doi: 10.1016/j.ijhcs.2020.102571]
24. McNeill A, Briggs P, Pywell J, Coventry L. Functional privacy concerns of older adults about pervasive health-monitoring systems. Presented at: Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments; 2017; Island of Rhodes Greece. [doi: 10.1145/3056540.3056559]
25. Robinson L, Brittain K, Lindsay S, Jackson D, Olivier P. Keeping In Touch Everyday (KITE) project: developing assistive technologies with people with dementia and their carers to promote independence. Int Psychogeriatr 2009 Jun;21(3):494-502. [doi: 10.1017/S1041610209008448] [Medline: 19193255]
26. Mann WC, Belchior P, Tomita MR, Kemp BJ. Older adults perception and use of PDAs, home automation system, and home health monitoring system. Top Geriatr Rehabil 2007 Jan;23(1):35-46. [doi: 10.1097/00013614-200701000-00006]
27. Austin J, Dodge HH, Riley T, Jacobs PG, Thielke S, Kaye J. A smart-home system to unobtrusively and continuously assess loneliness in older adults. IEEE J Transl Eng Health Med 2016;4:2800311. [doi: 10.1109/JTEHM.2016.2579638] [Medline: 27574577]
28. Moretti G, Marsland S, Basu D, Sen Gupta G, The MUSE Group. Towards a monitoring smart home for the elderly: one experience in retrofitting a sensor network into an existing home. J Ambient Intell Smart Environ 2013;5(6):639-656. [doi: 10.3233/AIS-130233]

XSL•FO
RenderX

29. Mohammed KI, Zaidan AA, Zaidan BB, et al. Real-time remote-health monitoring systems: a review on patients prioritisation for multiple-chronic diseases, taxonomy analysis, concerns and solution procedure. J Med Syst 2019 Jun 11;43(7):223. [doi: 10.1007/s10916-019-1362-x] [Medline: 31187288]
30. Abbas M, Le Bouquin Jeannès R. Acceleration-based gait analysis for frailty assessment in older adults. Pattern Recognit Lett 2022 Sep;161:45-51. [doi: 10.1016/j.patrec.2022.07.006]
31. Lin B, Zhang S, Bao F. Gait recognition with multiple-temporal-scale 3D convolutional neural network. Presented at: Proceedings of the 28th ACM International Conference on Multimedia; 2020; Seattle WA USA. [doi: 10.1145/3394171.3413861]
32. Yu S, Tan D, Tan T. A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition. Presented at: 18th International Conference on Pattern Recognition (ICPR'06); 2006; Hong Kong, China. [doi: 10.1109/ICPR.2006.67]
33. Sabo A, Mehdizadeh S, Iaboni A, Taati B. Estimating Parkinsonism severity in natural gait videos of older adults with dementia. IEEE J Biomed Health Inform 2022 May;26(5):2288-2298. [doi: 10.1109/JBHI.2022.3144917] [Medline: 35077373]
34. Yan S, Xiong Y, Lin D. Spatial temporal graph convolutional networks for skeleton-Based action recognition. AAAI 2018;32(1). [doi: 10.1609/aaai.v32i1.12328]
35. Goetz CG. Unified Parkinson's disease rating scale (UPDRS) and the movement-disorder society sponsored-unified Parkinson's disease rating scale (MDS-UPDRS). Encyclo Mov Disord 2010:307-309. [doi: 10.1016/B978-0-12-374105-9.00091-5]
36. Yin Z, Jiang Y, Zheng J, Yu H. STJA-GCN: a multi-branch spatial–temporal joint attention graph convolutional network for abnormal gait recognition. Appl Sci (Basel) 2023;13(7):4205. [doi: 10.3390/app13074205]
37. Du Y, Lim Y, Tan Y. A novel human activity recognition and prediction in smart home based on interaction. Sensors (Basel) 2019 Oct 15;19(20):4474. [doi: 10.3390/s19204474] [Medline: 31619005]
38. Islam M, Kabir M, Mridha MF, Alfarhood S, Safran M, Che D. Deep Learning-Based IoT System for Remote Monitoring and Early Detection of Health Issues in Real-Time. Sensors (Basel) 2023;23(11):5204. [doi: 10.3390/s23115204]
39. Megalingam RK, Pocklassery G, Jayakrishnan V, Mourya G, Thulasi AA. Smartphone based continuous monitoring system for home-bound elders and patients. Presented at: 2014 International Conference on Communications and Signal Processing (ICCSP); 2014; Melmaruvathur, India. [doi: 10.1109/ICCSP.2014.6950039]
40. Rachakonda L, Marchand DT. Fall-sense: an enhanced sensor system to predict and detect elderly falls using iomt. Presented at: 2022 IEEE Computer Society Annual Symposium on VLSI (ISVLSI); Jul 4-6, 2022; Nicosia, Cyprus. [doi: 10.1109/ISVLSI54635.2022.00100]
41. Rajaei H. An iot based smart monitoring system detecting patient falls. Presented at: 2022 Annual Modeling and Simulation Conference (ANNSIM); Jul 18-20, 2022; San Diego, CA, USA. [doi: 10.23919/ANNSIM55834.2022.9859346]
42. Mireles C, Sanchez M, Cruz-Ortiz D, Salgado I, Chairez I. Home-care nursing controlled mobile robot with vital signal monitoring. Med Biol Eng Comput 2023 Feb;61(2):399-420. [doi: 10.1007/s11517-022-02712-y]
43. Bjornsdottir K, Ceci C. Home care nursing practice for older persons with heart failure living at home. J Clin Nurs 2023 Jul;32(13-14):4103-4115. [doi: 10.1111/jocn.16575] [Medline: 36369989]
44. Robinson EL, Park G, Lane K, Skubic M, Rantz M. Technology for healthy independent living: creating a tailored in-home sensor system for older adults and family caregivers. J Gerontol Nurs 2020 Jul 1;46(7):35-40. [doi: 10.3928/00989134-20200605-06] [Medline: 32597999]
45. Braspenning AM, Cranen K, Snaphaan L, Wouters EJM. A multiple stakeholder perspective on the drivers and barriers for the implementation of lifestyle monitoring using infrared sensors to record movements for vulnerable older adults living alone at home: a qualitative study. Int J Environ Res Public Health 2022 Jan 5;19(1):570. [doi: 10.3390/ijerph19010570] [Medline: 35010829]
46. Itoh S, Tan HP, Kudo K, Ogata Y. Comparison of the mental burden on nursing care providers with and without mat-type sleep state sensors at a nursing home in Tokyo, Japan: quasi-experimental study. JMIR Aging 2022 Mar 23;5(1):e19641. [doi: 10.2196/19641] [Medline: 35319474]

Abbreviations

CNN: convolutional neural network
PPI: public and participant involvement
ST-GCN: Spatiotemporal Graph Neural Network
STJA-GCN: Spatio-temporal Joint Adjacency Graph Neural Network

XSL•FO
RenderX

Edited by A Kushniruk; submitted 12.04.24; peer-reviewed by M Jordan-Marsh, TA El-Hafeez; revised version received 14.11.24; accepted 24.11.24; published 27.02.25.

Please cite as:
Lochhead C, Fisher RB
On the Necessity of Multidisciplinarity in the Development of at-Home Health Monitoring Platforms for Older Adults: Systematic Review
JMIR Hum Factors 2025;12:e59458
URL: https://humanfactors.jmir.org/2025/1/e59458
doi:10.2196/59458

© Chris Lochhead, Robert B Fisher. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 27.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Challenges to Rehabilitation Services in Sub-Saharan Africa From a User, Health System, and Service Provider Perspective: Scoping Review

Callixte Cyuzuzo[1*], MSc, MPH; Marie Josee Dukuzimana[2*], MSc; Clement Muhire[3*], MSc; Mathew Sheldon Ames[4*], PhD; Emmanuel Ngwakongnwi[2*], PhD

[1]Center for Population Health, Institute of Global Health Equity Research, University of Global Health Equity, Kigali, Rwanda

[2]Institute of Global Health Equity Research, University of Global Health Equity, KG 7 Ave, 5th Floor, PO Box 6955, Kigali, Rwanda

[3]Campus Operations and Community Engagement, University of Global Health Equity, Burera, Rwanda

[4]Educational Development and Quality Center, University of Global Health Equity, Kigali, Rwanda

[*]all authors contributed equally

Corresponding Author:
Marie Josee Dukuzimana, MSc
Institute of Global Health Equity Research, University of Global Health Equity, KG 7 Ave, 5th Floor, PO Box 6955, Kigali, Rwanda

Abstract

Background: Rehabilitation aims to restore and optimize the functioning of impaired systems for people with disabilities. It is an integral part of universal health coverage, and access to it is a human right.

Objective: We aimed to identify the key challenges to rehabilitation services in Sub-Saharan Africa from a user, health system, and service provider perspective.

Methods: This scoping review was conducted in accordance with the 5-stage framework proposed by Arksey and O'Malley. A comprehensive electronic search was run to identify published articles on rehabilitation services in Sub-Saharan Africa. Of the 131 articles retrieved, 83 articles were assessed for eligibility and 15 papers that met the inclusion criteria were considered.

Results: The results show that people with disabilities in Sub-Saharan Africa face multifactorial challenges to access rehabilitation services. Poor access to rehabilitation services is associated with less attention given to rehabilitation by governments, which leads to less funding, negative cultural and social beliefs, fewer rehabilitation centers, poorly equipped rehabilitation units, failure of health systems, lack of training to rehabilitation practitioners, and logistical and financial constraints. This review also reveals that digital rehabilitation reduces costs and improves access to services in hard-to-reach geographical areas. However, digital rehabilitation faces challenges as well, including connectivity issues, inaccessibility to technology, a lack of technical knowledge, a lack of privacy, and ethical concerns.

Conclusions: People with disabilities face multifactorial challenges to access rehabilitation services in Sub-Saharan Africa. It is therefore critical to address these challenges to optimize patients' health outcomes and offer better rehabilitation services.

(JMIR Hum Factors 2025;12:e58841) doi:10.2196/58841

KEYWORDS

challenges; users; health system; service providers; Sub-Saharan Africa; scoping review; rehabilitation service

Introduction

Rehabilitation refers to a set of therapeutic approaches that are structured to address physical, developmental, emotional, and mental challenges and improve patients' health outcomes and quality of life [1]. Rehabilitation is an integral part of universal health coverage and contributes greatly to the achievement of the United Nations' Sustainable Development Goal 3, which aims to ensure healthy lives and promote well-being for people of all ages [2]. Access to rehabilitation for people with disabilities is a human right, as stated in Article 26 of the United Nations Convention on the Rights of Persons with Disabilities [3]. Rehabilitation services can be either delivered through traditional face-to-face methods or, using remote approaches, facilitated by existing technologies [4,5]. Digital rehabilitation has reduced costs and improved access to rehabilitation services in hard-to-reach geographical areas [5].

Demographic changes that lead to chronic health conditions and accidents have gradually contributed to the increased need for rehabilitation services [6]. It is estimated that 2.4 billion people have a health condition that may benefit from rehabilitation services; however, these services are still

XSL•FO
RenderX

inaccessible due to a shortage of rehabilitation practitioners [1]. A World Health Organization (WHO) report disclosed that the ratio of skilled rehabilitation practitioners to patients is 10 to 1,000,000 in low- and middle-income countries. This leads to a persistent scarcity of health-related human resources and reduced access to rehabilitation services in low-resource regions [7].

Sub-Saharan Africa (SSA) has a shortage of health-related human resources, with the region accounting for only 3.5% of the world's health workers, despite it having a high disease burden that includes several kinds of disabilities [1]. In this low-resource region, there is a significant difference in the availability of and access to rehabilitation services, and this impedes individuals with disabilities from achieving their desired health outcomes [8,9]. The WHO reported that in SSA, 50% of people do not get the rehabilitation services they need, and a high proportion of people with disabilities with unmet needs are in low- and middle-income countries [10].

Demand for rehabilitation services in SSA is well established given the prevalence of unmet rehabilitation needs, the rising cases of noncommunicable diseases, and the significant incidence of road traffic accidents that result in disabilities. Additionally, lack of access to rehabilitation services for people with chronic conditions can lead to a need for assistance with activities of daily living and long-term hospital stays [11]. It is important to explore the rehabilitation challenges faced by users, health systems, and service providers. This scoping review aims to enhance our understanding of the complex demands imposed by these factors in SSA and examine the challenges associated with rehabilitation services from those 3 domains.

Methods

Study Design

This review follows the framework of conducting scoping reviews as proposed by Arksey and O'Malley [12]. This framework consists of 5 distinct stages: identifying the research question; identifying relevant studies; selecting studies; charting the data; and collating, summarizing, and reporting the results.

Identifying the Research Question

The research topic and its associated objectives helped to determine the scope of the review, appropriate literature, search strategy, and inclusion and exclusion criteria. To attain the main aim of this review, the scope of rehabilitation services in SSA was explored and key challenges that users, service providers, and health systems face on the availability of and access to rehabilitation services were identified.

Identifying Relevant Studies

Search Strategy

A comprehensive electronic search was conducted to retrieve published data on the availability of and access to rehabilitation services in SSA. Research studies published in reputable journals and databases such as PubMed, National Center for Biotechnology Information, Scopus, PLOS, BioMed Central, Taylor and Francis, Science Direct, Frontiers, Springer Nature, and Web of Science were chosen, as they yielded the most topic-relevant articles. A search strategy using the keywords "rehabilitation services," "availability," "access," and "Sub-Saharan Africa" was used. To increase the number of accessed literatures, "Sub-Saharan Africa" was replaced by the name of the country, and "rehabilitation services" was replaced by specific types of rehabilitation services such as physiotherapy, occupational therapy, speech and language therapy, and prosthetics services. To maximize the literature coverage and to provide additional evidence, reference lists of primary studies and review articles were screened to identify additional relevant literature.

Inclusion and Exclusion Criteria

Inclusion criteria were developed to ensure that all relevant publications focusing on availability of and access to rehabilitation services in SSA were included. Works written in English and published in the above-mentioned reputable journals and databases between 2018 and 2023 were considered. Articles not written in English, published before 2018, and that did not report on rehabilitation services in SSA countries were excluded.

Study Selection

One team member led the search and screened the titles and abstracts of all retrieved papers based on the inclusion and exclusion criteria. Articles retained for review were confirmed by the research team after discussion. Duplicates were removed and the articles that satisfied the inclusion criteria were selected for full-text review.

Charting the Data

With the help of Microsoft Excel, data on availability of and access to rehabilitation services were extracted from the full texts. Challenges faced by users, service providers, and health systems were recorded. Information such as the title of the paper, names of authors, year of publication, study setting, study design, objectives of the paper, inclusion criteria, and study findings were extracted and recorded in the Excel table. Reference lists of included papers were also screened to optimize the search. The data extraction document is attached as a supplementary file (Multimedia Appendix 1).

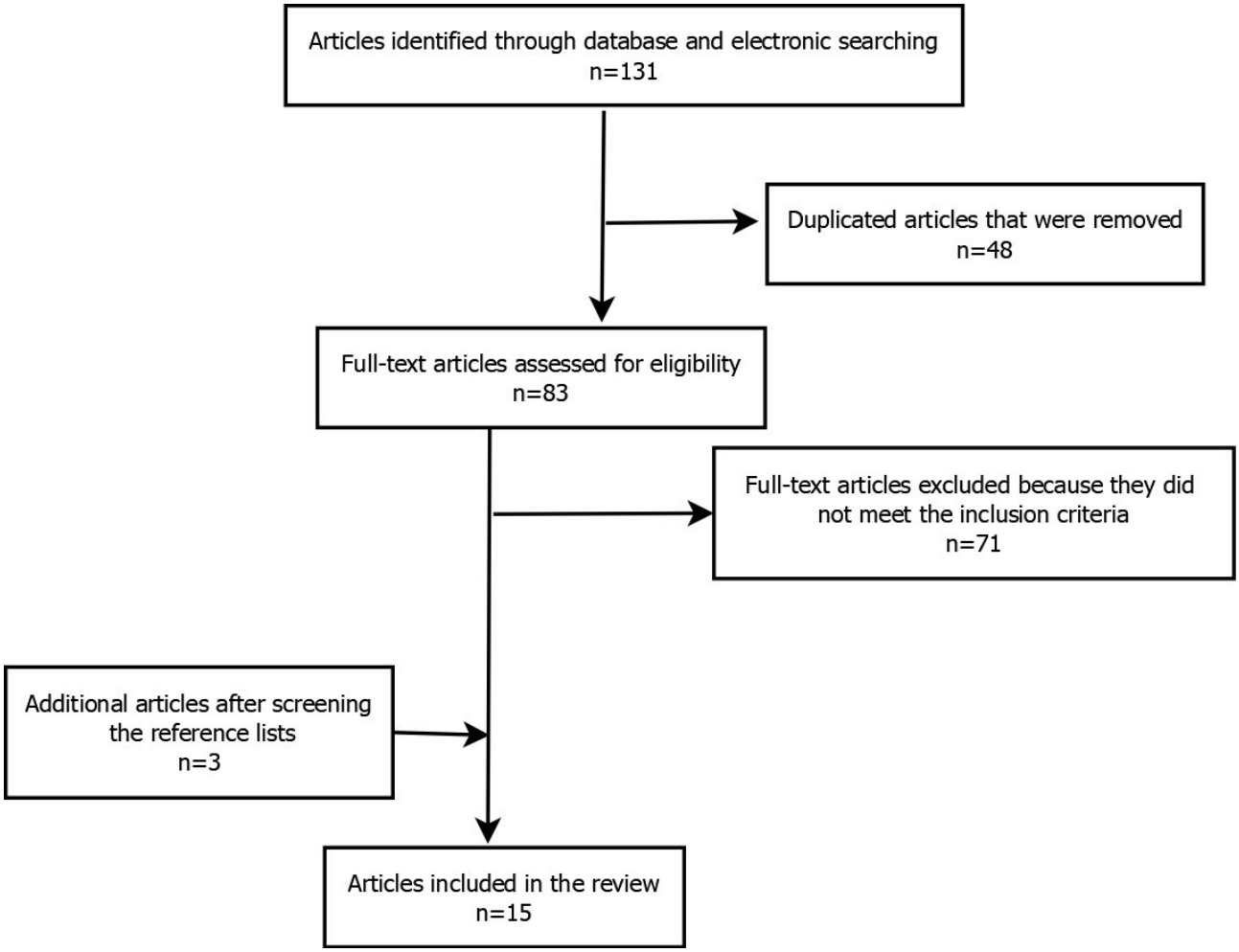
Results

Overview

The search yielded 131 research and review articles overall. Forty-eight articles were duplicates and 71 articles did not satisfy the inclusion criteria. The full texts of the remaining 12 articles were screened, and 3 more articles were identified using the reference lists, resulting in a total of 15 articles eligible for analysis. Six of the 15 (40%) articles were original studies while 9 (60%) were scoping or systematic reviews covering other aspects of rehabilitation. Figure 1 illustrates how the articles were selected. All 15 considered articles indicated the existence of rehabilitation challenges; of these, 6 papers discussed the challenges faced by users, 5 discussed those faced by health systems, 3 discussed those faced by service providers, and 1 review article reported challenges for all.

XSL•FO
RenderX

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart illustrating the article selection process.



Challenges to Rehabilitation Services in SSA

A synthesis of the literature revealed 2 categories of challenges to rehabilitation services from user, service provider, and health system perspectives (Table 1).

XSL•FO
RenderX

Table . Challenges to rehabilitation services reported in studies involving users, providers, and health systems.

	Category	
	Accessibility	Availability
Users	Overall challenges: • Unsuitable environment and buildings for people with disabilities • Inaccessibility to technology • Stigma, negative attitudes, and discrimination from therapists and other hospital staff • Long queues and waiting time • Financial constraints to purchasing medicines • Mindset that nothing can be done to help them • Lack of psychological support • Nonsupportive family members • Lack of trust in service providers • Pain during the rehabilitation process • Religious beliefs and socially accepted norms and traditions that claim that disability results from witchcraft or divine origins • Language barriers Digital rehabilitation–specific challenges: • Lack of secure platforms and privacy • Lack of technical knowledge or digital literacy • Connectivity issues • Inconsistency in power supply	• Do not know where to find the services • Long distance between users' homes and health facilities • Lack of reliable transport and logistical affordability • Distant health units
Service providers	Overall challenges: • Insufficient time for consultation • Inadequate follow-up • Inappropriate treatment procedures • Lack of patience • Ethical challenges Digital rehabilitation–specific challenges: • Lack of technical knowledge or digital literacy • Connectivity issues	• Lack of communication skills • Lack of professional training and necessary skills • Lack of maintenance for the assistive tools • Failure to schedule online appointments
Health systems	Overall challenges: • Consideration of rehabilitation as a less important health care strategy that is not well integrated into services • Fragmented health services • Lack of community awareness on rehabilitation services • Failure of the health system Digital rehabilitation–specific challenges: • Financial constraints	• Irregular referral to well-equipped health facilities • Lack of medicines and services • Shortage of health-related human resources • Low attention from governments and underfunding • Limited rehabilitation centers and poorly equipped rehabilitation units

Discussion

Principal Findings

This review revealed that people with disabilities in SSA have limited access to rehabilitation services. The limited access can be explained partly by user, service provider, and health system factors [8,13]. Studies included in this review [5,14] found that rehabilitation was given less priority by governments and not recognized as a crucial health care service. Consequently, the sector faced underfunding; experienced a deficiency in health-related human resources, particularly in therapists; encountered a scarcity of rehabilitation centers; and witnessed inadequately equipped rehabilitation units within health facilities. This contributed to the compromised access to high-quality rehabilitation services [5,14,15].

Another study reported that limited access to rehabilitation services in SSA is attributed to the financial inability to afford treatment, lack of insurance covering service expenses, long waiting times at the health facilities, and lack of drugs at health facilities [10]. Other studies have also found that some patients lack psychological support, are unaware of where to find

XSL•FO
RenderX

services, believe that nothing could be done to help them, and are not involved in making decisions on their health [8,16].

This review identified stigma as an impediment to accessing rehabilitation services. A study conducted in Rwanda [17] found that stigma around disability, coupled with negative attitudes and discrimination from therapists and other hospital staff, negatively impacted the uptake of rehabilitation services. Likewise, the role of family in accessing rehabilitation services is well documented. Patients with family members who do not perceive the need for rehabilitation services and who do not trust service providers may be demotivated to seek rehabilitation services [13,17,18]. Additionally, compromised access to health information, irregularities in referral to health facilities that are well equipped, inadequate policies and standards that govern the services, lack of adequate follow-up, insufficient time for consultations, and pain during the rehabilitation process are all hindrances to accessing rehabilitation services [9,15].

This review found that people with disabilities in SSA lack reliable, affordable, and accessible transportation means to reach health facilities. Furthermore, long distances to services and health units, unsuitable environments, and a lack of wheelchair-accessible buildings prevent users from accessing rehabilitation services. A review conducted in Brazil documented patients' determinants (residential location, economic resources, and social characteristics) and characteristics of services (cost, location, and status of the available facilities) as factors that limit access to rehabilitation services [8].

The results published by Bright et al [18] revealed that compromised health systems, lack of professional training and necessary skills for therapists, inappropriate treatment procedures, and lack of maintenance for the assistive tools reduce access to high-quality rehabilitation services. Health care providers who have not received enough training may lack the skills required to provide successful rehabilitation services. This lack of competence implies that misdiagnoses or outdated and inappropriate techniques could be used, causing harm to patients.

The study conducted by Jones et al [5] on telerehabilitation reported how digital rehabilitation services have addressed some logistical challenges; however, they come with additional challenges, including the absence of secure platforms and privacy, inaccessibility to technology, and lack of technical knowledge for both users and therapists [8]. Other challenges identified include a failure to schedule online appointments, connectivity issues, ethical challenges, and failure of the health systems. Patients and health care providers may struggle to use digital tools for rehabilitation, limiting the effectiveness of remote health services. Additionally, the inability to schedule and failure to maintain digital appointments may result in missed sessions, decreased patient participation, and care delays. Health systems face the challenge of poor policies that lead to a lack of financial capacity to afford digital technologies in health service delivery [19]. Digital technologies such as online medical records, wearable devices, and telemedicine can improve patient care and accessibility to health services, particularly in hard-to-reach areas. When financial limitations prevent the use of these technologies, patients may not receive the necessary high-quality care on time.

This review discerned that religious beliefs, cultural norms, and traditions that attribute disabilities to witchcraft or divinity serve as deterrents for individuals with disabilities seeking rehabilitation services at health facilities. A study conducted in Rwanda found that people with mental disorders believe that traditional and faith healers are more effective at treating mental problems than hospital specialists [20]. Such cultural beliefs dissuade people with disabilities from rehabilitation services, casting doubt on their quality and efficacy. Uncertainties regarding the precise origins of their conditions, which are often intertwined with mystical beliefs, add to this dissuasion [21].

Considering the rehabilitation challenges identified in this review, there is a need to bridge the gaps of infrastructure and social and cultural awareness regarding disability and rehabilitation services. This should be done by decentralizing services, providing continuous professional training to therapists, and promoting regular community awareness about rehabilitation services to inform service seekers and break the stigma around it. While this review has enumerated the challenges to rehabilitation services in SSA, it is worth noting that a majority of articles retained for consideration were scoping and systematic reviews that examined other aspects of rehabilitation. User and health system challenges were mostly highlighted in comparison to service provider challenges, which were limited in content.

Conclusion

This review reveals that rehabilitation services in SSA face multifactorial challenges that negatively impact timely access and quality of rehabilitation services for people with disabilities. As a consequence, those who require rehabilitation services experience longer periods of decreased mobility and functioning, inferior quality of life, and lower socioeconomic well-being. Future studies should examine the application of digital technologies to improve rehabilitation services' accessibility, especially in remote settings.

Authors' Contributions

All authors contributed to all phases of data collection and manuscript preparation.

Conflicts of Interest

None declared.

Multimedia Appendix 1

XSL•FO
RenderX

Data extraction table.
[DOCX File, 19 KB - humanfactors_v12i1e58841_app1.docx]

Checklist 1
Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) checklist.
[PDF File, 110 KB - humanfactors_v12i1e58841_app2.pdf]

References

1. Mapulanga M, Dlungwane T. Mapping evidence of community health workers delivering physical rehabilitation services in sub-Saharan Africa: a scoping review protocol. BMJ Open 2022 Mar 29;12(3):e053736. [doi: 10.1136/bmjopen-2021-053736] [Medline: 35351705]
2. Goal 3: ensure healthy lives and promote well-being for all at all ages. United Nations. URL: https://www.un.org/sustainabledevelopment/health/ [accessed 2025-02-10]
3. Convention on the rights of persons with disabilities and optional protocol. United Nations. URL: https://www.un.org/disabilities/documents/convention/convoptprot-e.pdf [accessed 2025-02-10]
4. Décary S, De Groote W, Arienti C, et al. Scoping review of rehabilitation care models for post COVID-19 condition. Bull World Health Organ 2022 Nov 1;100(11):676-688. [doi: 10.2471/BLT.22.288105] [Medline: 36324552]
5. Jones L, Lee M, Castle CL, Heinze N, Gomes RSM. Scoping review of remote rehabilitation (telerehabilitation) services to support people with vision impairment. BMJ Open 2022 Aug 1;12(8):e059985. [doi: 10.1136/bmjopen-2021-059985] [Medline: 35914903]
6. Jesus TS, Landry MD, Hoenig H. Global need for physical rehabilitation: systematic analysis from the Global Burden of Disease Study 2017. Int J Environ Res Public Health 2019 Mar 19;16(6):980. [doi: 10.3390/ijerph16060980] [Medline: 30893793]
7. Magnusson L, Kebbie I, Jerwanska V. Access to health and rehabilitation services for persons with disabilities in Sierra Leone - focus group discussions with stakeholders. BMC Health Serv Res 2022 Aug 5;22(1):1003. [doi: 10.1186/s12913-022-08366-8] [Medline: 35932077]
8. Clemente KAP, da Silva SV, Vieira GI, et al. Barriers to the access of people with disabilities to health services: a scoping review. Rev Saude Publica 2022;56:64. [doi: 10.11606/s1518-8787.2022056003893] [Medline: 35792776]
9. Naidoo U, Ennion L. Barriers and facilitators to utilisation of rehabilitation services amongst persons with lower-limb amputations in a rural community in South Africa. Prosthet Orthot Int 2019 Feb;43(1):95-103. [doi: 10.1177/0309364618789457] [Medline: 30044179]
10. Tannor AY, Nelson MES, Steere HK, Quao BO, Haig AJ. Building PRM in sub-Saharan Africa. Front Rehabil Sci 2022;3:910841. [doi: 10.3389/fresc.2022.910841] [Medline: 36189021]
11. Treger I, Kosto A, Vadas D, Friedman A, Lutsky L, Kalichman L. Crafting the future of community-based medical rehabilitation: exploring optimal models for non-inpatient rehabilitation services through a narrative review. Int J Environ Res Public Health 2024 Oct 8;21(10):1332. [doi: 10.3390/ijerph21101332] [Medline: 39457305]
12. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
13. Padmavathi J, Gandhi S, Siva Kumar T. Systematic review on end-users' perception of facilitators and barriers in accessing tele-rehabilitation services. J Psychosoc Rehabil Ment Health 2023 Mar 27;10(3):1-12. [doi: 10.1007/s40737-023-00333-8] [Medline: 37360916]
14. Conradie T, Berner K, Louw Q. Rehabilitation workforce descriptors: a scoping review. BMC Health Serv Res 2022 Sep 17;22(1):1169. [doi: 10.1186/s12913-022-08531-z] [Medline: 36115976]
15. Hussein El Kout NAR, Pilusa S, Masuku KD. A review of the framework and strategy for disability and rehabilitation services in South Africa. Afr J Disabil 2022;11:893. [doi: 10.4102/ajod.v11i0.893] [Medline: 36567920]
16. Louw Q, editor. Collaborative Capacity Development to Complement Stroke Rehabilitation in Africa: AOSIS; 2021. [doi: 10.4102/aosis.2020.BK85]
17. Kumurenzi A, Richardson J, Thabane L, Kagwiza J, Musabyemariya I, Bosch J. Provision and use of physical rehabilitation services for adults with disabilities in Rwanda: a descriptive study. Afr J Disabil 2022;11:1004. [doi: 10.4102/ajod.v11i0.1004] [Medline: 36092478]
18. Bright T, Wallace S, Kuper H. A systematic review of access to rehabilitation for people with disabilities in low- and middle-income countries. Int J Environ Res Public Health 2018 Oct 2;15(10):2165. [doi: 10.3390/ijerph15102165] [Medline: 30279358]
19. Owhor GA, Adeniyi AA, Shomuyiwa D. Overview of digital health in Sub-Saharan Africa: challenges and recommendations. IOSR J Nurs Health Sci 2023;12(1):19-21. [doi: 10.9790/1959-1201041921]
20. Muhorakeye O, Biracyaza E. Exploring barriers to mental health services utilization at Kabutare District Hospital of Rwanda: perspectives from patients. Front Psychol 2021;12:638377. [doi: 10.3389/fpsyg.2021.638377] [Medline: 33828506]

XSL•FO
RenderX

21. Grandpierre V, Milloy V, Sikora L, Fitzpatrick E, Thomas R, Potter B. Barriers and facilitators to cultural competence in rehabilitation services: a scoping review. BMC Health Serv Res 2018 Jan 15;18(1):23. [doi: 10.1186/s12913-017-2811-1] [Medline: 29334968]

Abbreviations

SSA: Sub-Saharan Africa
WHO: World Health Organization

Edited by A Kushniruk, E Borycki; submitted 26.03.24; peer-reviewed by D Franzsen, S Levine; revised version received 04.12.24; accepted 08.12.24; published 28.02.25.

Please cite as:
Cyuzuzo C, Dukuzimana MJ, Muhire C, Sheldon Ames M, Ngwakongnwi E
Challenges to Rehabilitation Services in Sub-Saharan Africa From a User, Health System, and Service Provider Perspective: Scoping Review
JMIR Hum Factors 2025;12:e58841
URL: https://humanfactors.jmir.org/2025/1/e58841
doi:10.2196/58841

© Callixte Cyuzuzo, Marie Josee Dukuzimana, Clement Muhire, Mathew Sheldon Ames, Emmanuel Ngwakongnwi. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 28.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

AI Chatbots for Psychological Health for Health Professionals: Scoping Review

Gumhee Baek[1,2*], MS, RN; Chiyoung Cha[1,3*], PhD, RN; Jin-Hui Han[1*], MS, RN

[1]College of Nursing, Ewha Womans University, 52 Ewhayeodae-gil, Daehyun-dong, Seodaemun-gu, Seoul, Republic of Korea

[2]Graduate Program in System Health Science and Engineering, Ewha Womans University, Seoul, Republic of Korea

[3]College of Nursing, Ewha Research Institute of Nursing Science, Ewha Womans University, Seoul, Republic of Korea

[*]all authors contributed equally

Corresponding Author:
Jin-Hui Han, MS, RN
College of Nursing, Ewha Womans University, 52 Ewhayeodae-gil, Daehyun-dong, Seodaemun-gu, Seoul, Republic of Korea

Abstract

Background: Health professionals face significant psychological burdens including burnout, anxiety, and depression. These can negatively impact their well-being and patient care. Traditional psychological health interventions often encounter limitations such as a lack of accessibility and privacy. Artificial intelligence (AI) chatbots are being explored as potential solutions to these challenges, offering available and immediate support. Therefore, it is necessary to systematically evaluate the characteristics and effectiveness of AI chatbots designed specifically for health professionals.

Objective: This scoping review aims to evaluate the existing literature on the use of AI chatbots for psychological health support among health professionals.

Methods: Following Arksey and O'Malley's framework, a comprehensive literature search was conducted across eight databases, covering studies published before 2024, including backward and forward citation tracking and manual searching from the included studies. Studies were screened for relevance based on inclusion and exclusion criteria, among 2465 studies retrieved, 10 studies met the criteria for review.

Results: Among the 10 studies, six chatbots were delivered via mobile platforms, and four via web-based platforms, all enabling one-on-one interactions. Natural language processing algorithms were used in six studies and cognitive behavioral therapy techniques were applied to psychological health in four studies. Usability was evaluated in six studies through participant feedback and engagement metrics. Improvements in anxiety, depression, and burnout were observed in four studies, although one reported an increase in depressive symptoms.

Conclusions: AI chatbots show potential tools to support the psychological health of health professionals by offering personalized and accessible interventions. Nonetheless, further research is required to establish standardized protocols and validate the effectiveness of these interventions. Future studies should focus on refining chatbot designs and assessing their impact on diverse health professionals.

(JMIR Hum Factors 2025;12:e67682) doi:10.2196/67682

KEYWORDS

artificial intelligence; AI chatbot; psychological health; health professionals; burnout; scoping review

Introduction

Health professionals are on the front line of patient care, enduring enormous responsibilities and constant pressure. Their work often occurs in an environment in which they must make life-and-death decisions, making them particularly vulnerable to psychological burnout [1]. According to recent studies, more than 50% of health professionals worldwide experience symptoms of burnout, such as fatigue, cynicism, and decreased effectiveness, leading to low job satisfaction [2], high turnover [3], and low-quality patient care [4]. This phenomenon not only causes an individual crisis but also has broader implications for health care systems [2,4]. Since the COVID-19 pandemic, psychological health issues such as anxiety, depression, and burnout among health professionals have become more prevalent because of long working hours, high-stress working environments, and exposure to traumatic events [5]. These issues directly affect health professionals' well-being and work performance and can ultimately negatively impact the quality of health care [6]. Therefore, supporting the psychological health of health professionals and preventing burnout are urgent challenges.

While traditional face-to-face counseling has been used to prevent psychological illness among health professionals, this

XSL•FO
RenderX

method presents various limitations including temporal and spatial constraints, limited availability of mental health professionals, accessibility barriers, and anonymity concerns [7-9]. Health professionals face significant challenges in accessing psychological care owing to demanding work schedules and high work intensity [10]. Additionally, they may feel psychologically burdened when seeking help in situations in which anonymity is not guaranteed [8]. To address these challenges, artificial intelligence (AI) chatbots have emerged as a potential solution, offering a new way to deliver psychological health interventions that can compensate for the limitations of traditional methods [7,8]. AI chatbots have the potential to be accessible anytime and anywhere while maintaining user anonymity, which may provide an alternative approach to psychological health support [7].

Recently, rapid advances in AI and natural language processing (NLP) have opened new possibilities for psychological health interventions, with AI chatbots gaining particular attention in the field of psychological health. Studies have demonstrated the potential of AI chatbots—AI-driven conversational agents—in providing mental health support [11] while emphasizing the importance of their safe and explainable implementation [12]. AI chatbots analyze users' conversation histories and data to provide advice and support tailored to their specific needs. This can help health professionals better manage psychological health issues, such as burnout, and enable timely interventions tailored to individual situations [13]. Previous research has shown that AI chatbots can improve coping skills through natural conversations with users [7], enhance the user experience, facilitate effective intervention delivery [13,14], and provide interventions tailored to specific needs [13,14]. This conversational approach enhances the effectiveness of interventions by fostering positive user engagement, as people are more likely to empathize with and respond to tools that feel engaging and relatable [6,7,14].

Despite the potential of AI chatbots for enhancing psychological health, there remains a lack of research on their effectiveness for health professionals. Much of the existing research has focused on using AI chatbots in patient care, whereas relatively little research exists on the use of AI chatbots to support health professionals. This research gap highlights the need to systematically investigate the characteristics and effectiveness of AI chatbots that target health professionals. Therefore, this study aimed to fill this gap and analyze the characteristics and effectiveness of AI chatbot interventions targeting health professionals to contribute to future research and practice.

Methods

Study Design

We conducted a scoping review to identify the available evidence regarding AI chatbot interventions for health professionals' psychological health. A scoping review maps the relevant literature in the field and is useful for identifying gaps in knowledge [15]. This review was conducted in accordance with the scoping review framework of Arksey and O'Malley. The study adopted the following five steps: (1) identifying the research question; (2) identifying relevant studies; (3) study selection; (4) charting the data; and (5) collating, summarizing, and reporting the results. The five stages provided a structured approach to assist in understanding the current evidence by examining the literature. There was no rating of quality performed on the included studies, as scoping studies do not require a quality evaluation [16]. Therefore, a scoping review is suitable for providing a broader perspective on the current research and mapping the relevant literature on AI chatbot interventions for health professionals' psychological health.

Identifying the Research Question and Relevant Studies

The research question was, "What is the status of the current body of knowledge regarding AI chatbots for the psychological health of health professionals?" According to Arksey and O'Malley, the search strategy involves identifying relevant studies, including published and unpublished works, with an emphasis on being as comprehensive as possible [15]. A literature search was performed between May and August 2024 from the following eight databases: seven global search engines in English (Cochrane Library, CINAHL, Embase, ProQuest Dissertations & Theses Global, PsyclNFO, PubMed, and Web of Science) and one Korean search engine in Korean (Research Information Sharing Service). Various combinations of search terms (Textbox 1) were used to either broaden or narrow the search, depending on the results in a specific database. The reference lists of the retrieved studies were searched manually. Finally, gray literature, relevant letters, and editorials were searched.

Textbox 1. Query box

medical team OR medical staff* OR healthcare worker* OR healthcare provider* OR health personnel* OR health professional* OR doctor* OR surgeon* OR physician* OR nurse* OR midwife* OR Licensed Practical Nurse* OR licensed vocational nurse* OR physician assistant OR nurse practitioner OR clinical nurse specialist* OR nurse clinician* OR registered nurse* OR resident* OR therapist* OR pharmacist* OR nurse assistant* OR nurse aid* AND burnout OR burn-out* OR depress* OR depression OR stress* OR exhaust* OR depersonalization OR personal accomplishment OR anxiety* AND chatbot* OR artificial intelligence OR conversational agent* OR dialogue system* OR machine learning OR natural language process*

Study Selection

Overview

Studies published before May 2024 were selected for this review. The studies retrieved from the literature search were reviewed by three researchers using the following inclusion and exclusion criteria.

Inclusion Criteria

(1) The purposes of the studies included the development or implementation of AI chatbot interventions. (2) Participants of the studies included licensed health professionals such as

XSL•FO
RenderX

physicians or surgeons, dentists, medical doctors, registered nurses, licensed practice or vocational nurses, pharmacists, and allied health professionals such as respiratory and physical therapists. Studies that included nonhealth professionals in addition to health professionals as part of their participants were included if the main purpose of the AI chatbot was to alleviate the psychological distress of health professionals. (3) The intention of the AI chatbots was to improve the psychological health of health professionals, focusing on outcomes such as stress, burnout, depression, and anxiety. (4) The studies were written in English or Korean.

Exclusion Criteria

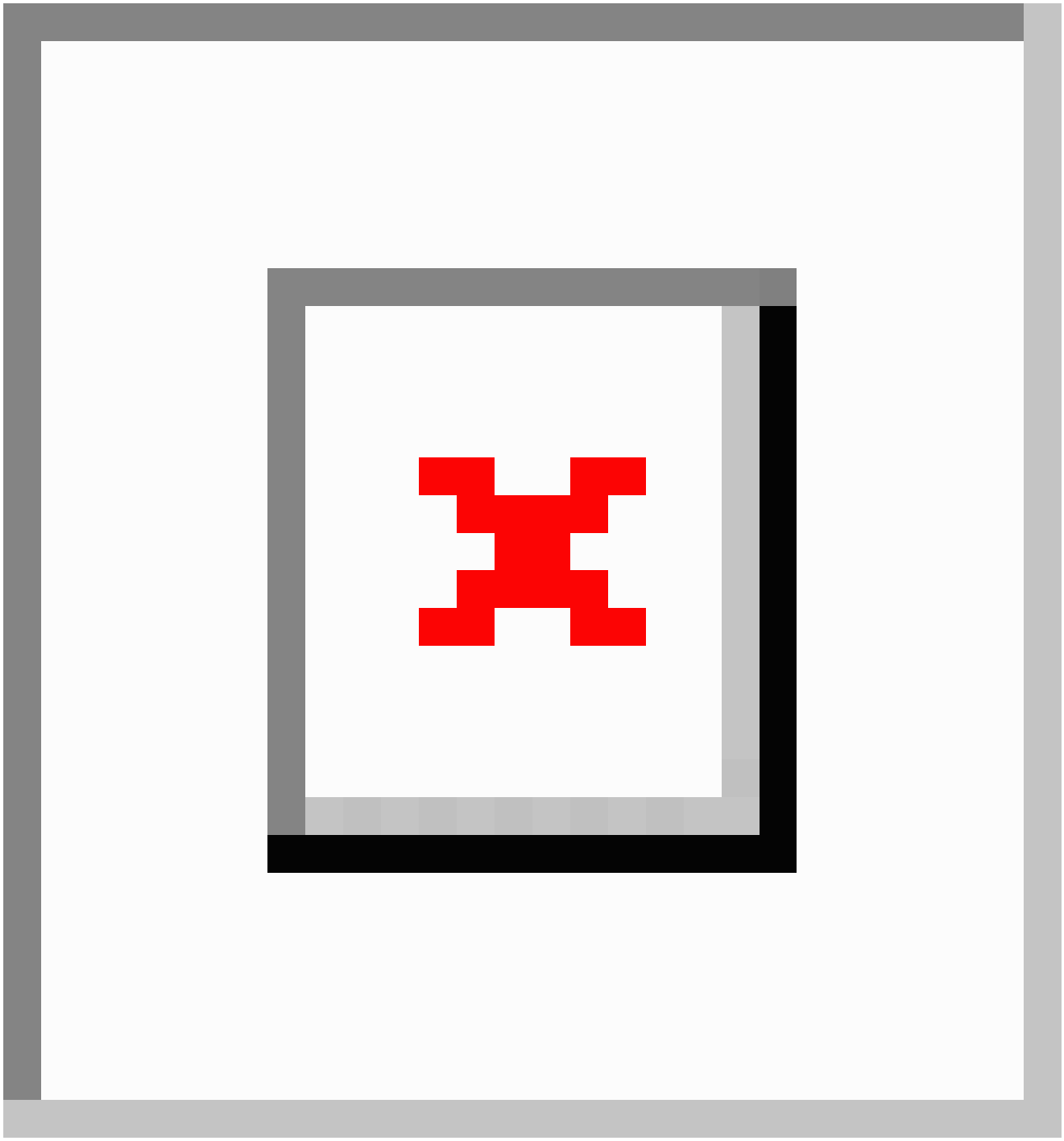
Studies in which we could not retrieve the full text were excluded.

A total of 2465 studies were retrieved from the initial database search (Figure 1). After removing duplicates, 1766 studies remained. The titles and abstracts were reviewed, and 1750 studies were excluded, leaving 16 studies for a full-text review. Six studies completely met the inclusion criteria after full-text review. Then, by reviewing the reference lists and manually searching, five additional studies were obtained, of which four studies met the criteria for the scoping review. Finally, 10 studies were included in this scoping review.

Throughout the review process, two researchers (GB and JH) took the lead in data extraction, and all researchers reviewed the content. Any uncertainties and disagreements were resolved through weekly meetings. The literature review software Covidence (Veritas Health Innovation Ltd), a study-managing tool for collaborating researchers [17], was used to manage the retrieved studies and to screen for redundant studies.

XSL•FO
RenderX

Figure 1. Flowchart. AI: artificial intelligence.



Charting the Data

The data are presented in tables, summarized, and synthesized qualitatively according to the purpose and research questions of the review.

Collating, Summarizing, and Reporting the Results

The final stage of a scoping review involves collating, summarizing, and reporting the results [15]. Data extraction was performed by two researchers (GB and JH) and checked for accuracy by all the researchers. Uncertainties and disagreements were resolved through discussions. The data were extracted using a common data extraction format.

Results

General Characteristics

Table 1 presents the characteristics of the analyzed studies. Ten studies published between 2018 and 2024 were retrieved. The studies were published worldwide: in Canada [7,18], Japan [19], Malawi [20], Netherlands [21], Singapore [13,22], Spain [23], and the United States [24,25]. Four studies examined diverse health professionals, including doctors, nurses, nurse assistants, clinical officers, pharmacists, lab technicians, and other allied health professionals [20,21,23,26]. Two studies specifically targeted doctors and nursing staff [13,27] such as nurse practitioners, registered nurses, licensed practical nurses, and nursing assistants, while one study exclusively covered chatbots

XSL•FO
RenderX

for nurses [19]. Three studies identified their participants as health professionals, without specifying the specific types [7,18,22]. Of the ten studies, eight studies focused solely on health professionals [13,19-23,26,27]. Two studies included both health professionals and their families [7,18].

The study designs were diverse. Of the 10 studies, three were development studies [18,23,27], one was a design verification study [26], one was a pilot study [22], one was a randomized controlled trial [20], one was a mixed-method feasibility study [21], one was an intervention study [19], one was a cross-sectional study protocol study [7], and one was a qualitative study [13]. Three studies evaluated the development and feasibility of AI chatbots for psychological health support [23,26,27].

XSL•FO
RenderX

Table . Study characteristics.

Author (year)/ country	Study design	Participants	Source of participants	Sample size (M:F[a] ratio)	Age (years), mean (SD)	Publication type
Anmella et al [27](2023)/Spain	Development study	Primary care and health care professionals; medical doctors (GPs[b] and specialists from diverse disciplines), nurses, and nurse assistants	• GP referrals to PCMHSP[c] • Hospital Clínic de Barcelona health professionals	• Simulation study: 17 (24:76) • Feasibility and effectiveness study: 34 (24:76)	• Simulation study: 36.5 (9.7) • Feasibility and effectiveness study: 35.3 (10.12)	Journal study
Chaudhry and Islam [26](2021)/United States	Design verification study	Health professionals in their workplaces during the COVID-19 pandemic; RNs[d], NPs[e], LPN[f], CNAs[g], PAs[h], RT[i], PTs[j], students or intern nurses	Local medical services departments and health professionals' mailing lists	22 (9.1:90.9)	34.98 (8.55)	Journal study
Chang et al [22] (2024)/Singapore	Pilot exercise (service evaluation study)	Health professionals (types unspecified)	National tertiary health care cluster	527 (N/A)[k]	N/A	Journal study
Jackson-Triche et al [23] (2023)/United States	Development and usability study	Health professionals including faculty, staff, and trainees across various disciplines within the health care system	UCSF[l]	Over 26,000 (N/A)	N/A	Journal study
Kleinau et al [20] (2024)/Malawi	Randomized controlled trial	Health professionals (doctors, nurses, clinical officers, medical assistants, pharmacists, lab technicians, physiotherapy technicians)	Public and private health care facilities within Blantyre and Lilongwe districts	1584 (N/A) Completed: 511 (E=296, C=215)	N/A	Journal study
Kroon [21](2018)/Netherlands	Mixed method feasibility study (one group pre-posttest design)	Health professionals (radiation expert, social worker, nurse in training, training officer, pedagogue, health professional disability care HRD[m] adviser	Email using personal and iThrive connections.	12 (8.3:91.7) (interview participant=9)	Half of the participants were aged 18 to 25 years	Master's thesis
Matsumoto et al [19](2021)/Japan	Intervention study	Nurses	Two hospitals in Tokyo	70 (N/A)	N/A	Journal study
Noble et al [7](2022)/Canada	Cross-sectional study protocol	Health professionals and their families	The Canadian provinces of Alberta and Nova Scotia	2200 (N/A)	N/A	Journal study
Yoon et al [13] (2021)/Singapore	Qualitative study	Frontline health professionals (doctors and nurses)	Singapore general hospital and community care facilities (large scale institutional isolation units)	42 (24:76)	29.6 (3.9)	Journal study
Zamani [18](2022)/Canada	Development study	Health professionals and their families	The Canadian provinces of Alberta and Nova Scotia	N/A	N/A	Master's thesis

XSL•FO
RenderX

[a]M:F: male:female.
[b]GP: general practitioner.
[c]PCMHSP: Primary Care Mental Health Support Program.
[d]RN: registered nurse.
[e]NP: nurse practitioner.
[f]LPN: licensed practical nurse.
[g]CNA: certified nurse assistant.
[h]PA: physician assistant.
[i]RT: respiratory therapist.
[j]PT: physical therapist.
[k]N/A: not available.
[l]UCSF: University of California, San Francisco
[m]HRD: human resource development.

AI Chatbot Characteristics

Table 2 shows the characteristics of the AI chatbots developed or implemented in the retrieved studies. The types of AI chatbot algorithms were diverse; NLP-based AI was the most commonly used algorithm, appearing in six studies [7,13,18,19,21,23]; rule-based AI, an algorithmic method that does not learn from data or create rules on its own, was used in two studies [20,22]; and learning-based AI that learns from data was adopted in two studies [26,27]. AI chatbots were delivered through mobile apps in six studies [13,19,20,22,26,27], and web-based platforms via web links in four studies [7,18,21,23]. The interaction type in all studies was 1:1 interaction: Among them, one study included 1:1 interaction alongside group chat functionality [26], and another study included 1:1 interaction with optional human coach support [22]. In seven studies, the chatbot actively initiated the conversation [7,13,18-20,23,27], while in three studies, the conversation was passively initiated [21,22,26].

Participant completion of the AI chatbot interventions was reported in four studies [20,21,23,27]; among these, three studies provided the percentage of participants who completed the intervention: 14.3% [21], 67.7% [20], and 73% [27] while one study reported the employee access rate of the intervention as 10.88% [27]. Of the 10 studies, six evaluated the AI chatbot's usability [13,20,22,23,26,27]. Four studies relied on participant feedback: 54.6% rated it as useful [26], the average satisfaction score was 4.07 out of five [22], more than 80% reported it as helpful [23], and 91% of the users found the chatbot easy to use [20]. One study used user engagement indicators to identify high subjective usability (acceptability, usability, and satisfaction) but low objective engagement (completion, adherence, compliance, and engagement) [27]. Another study assessed chatbot performance metrics, reporting an intent detection accuracy of 99.1% and an entity extraction accuracy of 95.4% for the Mira Chatbot [18]. Chatbot usability was not evaluated in the remaining four studies [7,13,19,21].

XSL•FO
RenderX

Table . AI[a] chatbot characteristics.

Author (year)/ Country	Chatbot name	Purposes of chatbot	AI chatbot type	AI algorithm	Mode of delivery	Interaction type	Initiates conversation	Completion related	Chatbot usability evaluation
Anmella et al [27] (2023)/Spain	Vickybot	To screen anxiety, depression, and burnout with interventions	Digital decision support platform (PRESTO[b])	• Machine-learning severity prediction model	Mobile app	1:1 interaction	Active	27% completed second self-assessment	• High subjective UEI[c] (acceptability, usability, and satisfaction). • Low objective UEI (completion, adherence, compliance, and engagement)
Chaudhry and Islam [26] (2021)/United States	iChatBot (AIO) PeerChat (AI X)	To provide stress management support during COVID-19	Intelligent chatbot	• iChatBot: machine learning model • PeerChat: immediate communication with expert peers	Mobile app	iChatBot: 1:1 PeerChat: group chat	Passive (on-demand queries)	Not reported	• 54.6% participants rated the app as either useful or very useful • 59.1% participants were willing to use it
Chang et al [22] (2024)/Singapore	Wysa	To evaluate Wysa for health professionals' support	Conversational chatbot	• Rule-based AI model	Mobile app	1:1 interaction with optional human coach support	Passive	Not reported	• High engagement rate with positive feedback (91-92%) • Mean feedback score of 4.07 (SD 0.95) out of 5

XSL•FO
RenderX

Author (year)/ Country	Chatbot name	Purposes of chatbot	AI chatbot type	AI algorithm	Mode of delivery	Interaction type	Initiates conversation	Completion related	Chatbot usability evaluation
Jackson-Triche et al [23]/United States	UCSF[d] Cope	To provide behavioral health support via chatbot	An algorithm-based, automated, web-based AI conversational tool	• Natural language model	Web-based interface	1:1 interaction	Active	10.88% of employees accessed the chatbot	• >80%of attendees reported the experience as helpful
Kleinau et al [20] (2024)/Malawi	Viki (Vitalk's virtual mental health care assistant)	To assess chatbot impact on mental health in Malawi	Automated conversational agent	• Rule-based AI model	Mobile app	1:1 interaction	Active	Of 1584 participants, 511 completed	• 91% found the app easy to use; 87% found the content relevant; 92% benefited from the app, and 83% felt more resilient • Common barriers: lack of time (45%), app duration (30%), and usability issues (13%)
Kroon [21] (2018)/Netherlands	iThrive	To help health professionals manage inner critics and enhance self-compassion	Preprogrammed conversations and interaction	• Natural language model	Web-based interface	1:1 interaction	Passive (users initiate interaction with the chatbot)	12 of 14 participants included in the final analysis	• N/A[e]
Matsumoto et al [19] (2021)/Japan	CB[f] app (a chatbot app)	To support mental health self-care for busy professionals, especially nurses	Digital decision support platform	• Natural language model	Mobile app, VR[g]	1:1 interaction	Active	Not reported	• N/A
Noble et al [7](2022)/Canada	MIRA[h]		Hybrid NLP[i] and decision tree AI chatbot	• Natural language model	Web-based platform	1:1 interaction	Active	Not reported	• N/A

XSL•FO
RenderX

Author (year)/ Country	Chatbot name	Purposes of chatbot	AI chatbot type	AI algorithm	Mode of delivery	Interaction type	Initiates conversation	Completion related	Chatbot usability evaluation
		To provide tailored mental health information and services for health professionals and their families							
Yoon et al [13](2021)/Singapore	mHealth[j] app	To support frontline health professionals's psychosocial well-being during COVID-19	N/A	• Natural language model	Mobile app	1:1 interaction	Active	Not reported	• N/A
Zamani [18] (2022)/Canada	MIRA	To offer strategic mental health resources for health professionals and their families	Hybrid NLP and decision tree AI chatbot	• Natural language model	Web-based platform	1:1 interaction	Active	Not reported	• Intent detection: 99.1% accuracy in identifying user needs • Entity extraction: 95.4% accuracy in recognizing key terms for resource recommendations.

[a]AI: artificial intelligence.

[b]PRESTO: primary care digital support tool in mental health.

[c]UEI: user engagement indicator.

[d]UCSF cope: University of California, San Francisco faculty, staff, and trainee coping and resiliency program.

[e]Not available.

[f]CB: cognitive behavioral.

[g]VR: virtual reality

[h]MIRA: mental health intelligent information resource assistant.

[i]NLP: natural language processing.

[j]mHealth: mobile health.

AI Chatbot Intervention Characteristics and Outcomes

Table 3 lists the contents and outcomes of the AI chatbot interventions. The primary component is the screening function for mental health conditions such as anxiety, depression, and stress, which was included in seven studies [19-23,26,27]. Four studies included cognitive behavioral therapy techniques [13,20,22,27]. Additionally, four studies implemented stress management techniques, such as breathing exercises and meditation [13,19,20,22]. Three studies incorporated problem-solving strategies [7,18,22] and psychoeducation [19,22,27]. Two studies used behavioral activation [20,22], cognitive restructuring [21,22], and gratitude training [20,22]. One study featured storytelling [21]. In addition, there was a

XSL•FO
RenderX

function to support portal sites connected to community resources in three studies [7,18,23].

XSL•FO
RenderX

Table . AI[a] chatbot intervention characteristics and outcomes.

Author (year)/ country	Chatbot contents	Additional features	Intervention doses	Main outcomes
Anmella et al [27](2023)/Spain	• Screening and monitoring for self-assessment tools • Personalized psychological modules: CBT[b], mindfulness, DBT[c], and ACT[d] • Psychoeducation and coping strategies for stress management • Chatbot-guided navigation	• Emergency alert for suicidal thoughts • Weekly objective reminders • Audio recording option for reflections after self-assessments	• 1 month (self-assessments conducted at baseline and biweekly)	• Anxiety (t_8=1.000; P=.34) and Depressive (t_8=0.40; P=.70) symptoms were not significant differences • Work-related burnout was moderately reduced (z=–2.07, P=.04, r=0.32)
Chaudhry and Islam [26] (2021)/United States	• iChatBot—provides real-time answers (self-monitoring and stress-tracking information related to patient care and workplace stress management) • PeerChat—enables peer-to-peer support	• AssignedTasks: records stress levels and task completion times • MyNotifications: Sends reminders for health activities like sleep, exercise, and meals • SmartMonitor: connects with devices to track vitals (eg, blood pressure, oxygen, pulse, and temperature) • InfectionCheck: Manual COVID-19 symptom checker via chatbot	• N/A[e]	• Usefulness and willingness: 54.6% found the app useful; 59.1% willing to use it. • Six feature feedback: MyNotifications: useful reminders; PeerChat: positive for peer support; AssignedTasks: good for tracking, but time-consuming; InfectionCheck: useful, potential redundancy; StressMonitor: very helpful for coping; iChatBot: concerns about emergency responsiveness
Chang et al [22](2024)/Singapore	• Mental health assessments: anxiety—GAD-7[f], depression—PHQ[g]-2 • Intervention modules: mindfulness, sleep meditation, guided visualization, thought recording, behavioral activation, psychoeducation, breathing exercises, cognitive restructuring, acceptance, grounding, social support, problem-solving, habit building, gratitude training	• Option to connect with a human coach • Weekly objective reminders and notifications • Gratitude challenge to boost engagement • Audio recording for self-reflection	• Varied among users (on average, users completed 10.9 sessions over 3.80 weeks)	• 495 completed at least one full session; 422 completed two or more sessions • The interventions most used were for sleep and anxiety, with a strong repeat-use rate • 46.2% reported symptoms of anxiety • 15.2% reported symptoms of depression
Jackson-Triche et al [23] (2023)/United States	• Mental health assessment screening: suicide or self-harm thoughts, substance use concerns, emotional distress, social withdrawal and work-related struggles, resource navigation • Web-based self-care resources • Behavioral health services: emergency support	• Telehealth assessments for crisis evaluation • Self-management website offering curated resources • In-person navigator for personalized support • Support groups with regular peer meetings	• Varied based on user interaction	• UCSF[h] cope reached all units • 10.88% (3785/34,790) of employees accessed the technology • 39.7% (708/1783) of employees with psychological distress requested in-person services • Positive feedback on all program elements. • Over 80% of attendees found the experience helpful

XSL•FO
RenderX

Author (year)/ country	Chatbot contents	Additional features	Intervention doses	Main outcomes
Kleinau et al [20](2024)/Malawi	• Mental health assessments: anxiety (GAD-7), depression (PHQ-9), burnout (OLBI[i]), loneliness (UCLA Loneliness Scale)[j], resilience (RS-14)[k] • Careline programs (thematic mental health conversation tracks) • Intervention: CBT techniques, positive psychology strategies, behavioral activation, gratitude exercises, breathing, relaxation, and meditation exercises,	• Mood tracking with emojis for human-like interaction • Gamification for engagement • Emergency support with local psychologist contact info for high-risk users	• 8 weeks -First 4 weeks: daily interactions • Second 4 weeks: interactions every other day	• GAD-7 showed a decrease (–0.44), but the result was not statistically significant. • OLBI showed a decrease (–0.58), with borderline significance. • PHQ-9 showed a significant decrease. • UCLA loneliness scale: no significant • RS-14 showed a significant increase.
Kroon [21](2018)/Netherlands	• Awareness of inner critics: storytelling, encouraging reflection • Identification of inner critics: assessment tool • Cognitive restructuring techniques • Self-compassion development	• Humor: adds humor for engaging interactions • Tokens: rewards for task completion • Compliments: encourages users with motivating feedback	• 2 weeks	• ProQOL[l] and SCS[m] no significant. • Quantitative results: no significant changes in self-compassion or compassion fatigue. • Qualitative results: Positive feedback on the iThrive intervention, with increased awareness of inner critics and changes in thinking patterns. • Chatbot interactions reportedly reduced stress and improved self-compassion.
Matsumoto et al [19](2021)/Japan	• Digital-SAT[n] method: self-guided stress management, techniques to convert stress-related physical discomfort into positive sensations, image-based therapeutic exercises • Mental health assessments: self-esteem, anxiety, emotional support, depression, problem-solving abilities • Stress management: guided stress relief sessions, emotional stabilization exercises, visualization techniques	• Automated daily messages with reminders and psychoeducational content • Gamification elements such as progress tracking • Immersive VR[o] experiences	• 4 weeks	• Self-esteem and family support significantly improved. • Self-repression showed a trend of improvement but was not significant. • SDS scores and counseling needs increased in the VR + CB[p] group. • Maximal blood pressure decreased significantly in the VR group and showed a decreasing trend in the VR + CB group. • Minimal blood pressure significantly decreased in the VR group.
Noble et al [7](2022)/Canada	• Mental health education and information: problem-solving • Personalized resource navigation • Self-help strategies and coping skills		• N/A	• The findings will be reported in follow-up research

XSL•FO
RenderX

Author (year)/ country	Chatbot contents	Additional features	Intervention doses	Main outcomes
		• Emergency contacts for urgent mental health support • Hybrid AI approach: free-text input and predefined options • Crisis support referrals for suicidal thoughts		
Yoon et al [13](2021)/Singapore	• Personalized goal setting: tailored goals for lifestyle improvements (eg, sleep, exercise), feedback on progress to encourage behavioral changes, reminders to support adherence to personal goals • Educational resources: mindfulness exercises and short wellness studies, stress management techniques	• In-app counseling and peer support • Reminders and notifications • Gamification elements: point-based rewards • Engaging features to promote continued app use • Mood tracking and progress monitoring	• N/A	• Mood-tracking reminders were helpful, but workers struggled with self-awareness. • Goal-setting and resources were valued; frequent notifications were distracting. • A built-in counselor chat was preferred for accessibility and privacy. • Gamification was not well-received. • Peer support was needed, but app-based interactions raised concerns. • Motivation, usability, and rewards were crucial for app use.
Zamani [18](2022)/Canada	• Information on mental health topics, including substance use and coping strategies: problem-solving • Delivers personalized recommendations • Facilitates access to local mental health services and programs through a resource portal	• Emergency contact option for users experiencing distress • Stores conversation logs • Feedback system	• N/A	• Intent detection: 99.1% accuracy in identifying user needs. • Entity extraction: 95.4% accuracy in recognizing key terms for resource recommendations.

[a]AI: artificial intelligence.

[b]CBT: cognitive behavior therapy.

[c]DBT: dialectical behavioral therapy.

[d]ACT: acceptance and commitment therapy.

[e]N/A: not available.

[f]GAD-7: general anxiety disorder-7.

[g]PHQ: patient health questionnaire.

[h]UCSF cope: University of California, San Francisco faculty, staff, and trainee coping and resiliency program

[i]OLBI: Oldenburg burnout inventory.

[j]UCLA: University of California, Los Angeles loneliness scale.

[k]RS-14: 14-item resilience scale.

[l]ProQOL: Professional Quality of Life Scale.

[m]SCS: Self-Compassion Scale

[n]SAT: structured association technique.

[o]VR: virtual reality.

[p]VR-CB: virtual reality-chatbot.

XSL•FO
RenderX

Across the studies, a variety of additional features were incorporated. Seven studies included AI chatbot functionalities for crisis management and support [7,13,18,20,22,27]. Specific features encompassed emergency contact information for crises [7,18,20,27], peer support groups and in-app forums for social support [13], expert support through in-app counseling, and options for connecting with human coaches [13,22,23]. Six studies incorporated reminders to encourage chatbot use [13,18,19,22,26,27]. Gamification elements, such as point-based rewards to encourage participation, were featured in four studies [13,19-21]. Peer support functionalities were available in three studies [13,23,26]. Two studies included mood-tracking features for health and lifestyle management [13,20], while another two studies provided an audio recording option [22,27]. Additionally, one study integrated devices for tracking vital signs [26] and provided an immersive virtual reality experience [19]. A total of 4 out of 10 studies explicitly stated the duration of chatbot use or intervention dose ([19]: 4 wk; [20]: 8 wk; [21]: 2 wk; [23]: 1 mo) while it was not specified in six studies. The dose varied among users in two studies [22,23].

The research in four studies evaluated psychological health outcomes [19-21,27], whereas, in six studies, the focus was primarily on the development and description of user perceptions of the AI chatbot without providing specific outcome measures [7,13,18,22,23,26]. The most commonly measured psychological health outcomes were anxiety and depression, which were examined in three studies [19,20,27]. Burnout was investigated in two studies [20,27]. Compassion fatigue [21], loneliness [20], resilience [20], and self-compassion [21] were also evaluated. Anxiety was measured in three studies: one reported a significant reduction of scores [20] while the other two reported no significant differences in anxiety scores [19,27]. Burnout was assessed in two studies, and each reported a significant decrease in its scores [20,27]. Depressive symptoms were measured in two studies: one reported a significant reduction of scores [20] while the other did not [27]. Other psychological variables were measured in one study each; of those, the variables that showed significant change in scores were health counseling needs [19] and resilience [20].

Discussion

Principal Results

The studies retrieved for this scoping review were published during the past seven years, between 2018 and 2024, and reflect the rapidly evolving trend of AI chatbot technology for supporting psychological health. The global distribution of studies across countries indicates a widespread interest in using AI for psychological health support and health professionals' psychological health.

Our review revealed that chatbots were used by a broad spectrum of health professionals, highlighting their diverse applications. This diversity underscores the potential for universally designed chatbots to effectively address psychological health needs across different professional settings. However, prior research has shown that nurses, who have direct and continuous patient contact, experience greater mental health distress compared to other health professionals [24]. Therefore, future studies should consider incorporating role-specific options into chatbot designs to create tailored interventions for diverse professional groups.

The diverse research methods used in the retrieved studies signify the nascent stage of research on AI chatbots for psychological health support, emphasizing technology refinement and user experience comprehension. Such a variety of research approaches is crucial for validating the efficacy and safety of AI chatbot technology, enabling iterative improvements and a deeper understanding of how health professionals interact with AI chatbots [25]. However, it is noteworthy that only 4 out of 10 studies measured psychological health outcomes. This limitation underscores the significant gap between technological developments and clinical validation. As the field matures, it becomes imperative to quantitatively assess the impact of AI chatbots on psychological health indicators. Further studies measuring outcome variables to demonstrate the clinical effectiveness of AI chatbots are needed.

The studies reviewed here demonstrate the variety of AI algorithms used in chatbot design, including NLP-based-, rule-, and learning-based approaches. NLP-based AI chatbots which provide personalized responses by analyzing user inputs [28] are the most commonly used AI algorithms. Evidence has shown that an NLP-based AI chatbot, which simulates natural conversations, is successful at delivering tailored interventions [29]. Rule-based AI algorithms, adopted in 2 studies out of the 10 retrieved studies, are effective in providing fixed and predetermined responses; however, they lack the adaptability required to provide more dynamic and engaging psychological interventions [30]. Previous research has shown that NLP technology, which interprets psychological health cues from user-generated texts, significantly improves the effectiveness of interventions [31,32], a finding consistent with our review [7,13,18,19,21,23]. Moreover, the development of deep learning techniques has led to the widespread use of general-purpose pretrained language models in various NLP applications [33]. This suggests that NLP-based algorithms are particularly suitable for delivering personalized psychological support in real-time interactions. Machine learning and deep learning algorithms show promise in enhancing AI chatbots' effectiveness by continuously learning through user interactions. These algorithms can provide accurate and personalized interventions over time [34]. Future studies should investigate how these advanced algorithms can be used to offer more precise psychological support tailored to individual user experiences.

User engagement is a critical factor in the success of mental health interventions; however, several studies did not report intervention completion rates [20,21,23,27]. These issues might become more pronounced in real-world settings, where factors such as interface complexity, the chatbot's ability to provide meaningful and personalized feedback, and technical issues impact the overall user satisfaction and perceived effectiveness of the intervention [35]. This underscores the need for further exploration of how chatbot interactions can be designed to maintain user engagement, possibly through improved user experiences and personalization techniques [36,37]. From a different perspective, unlike traditional interventions, which are typically designed to be completed from start to finish, chatbot

XSL•FO
RenderX

interactions often do not follow this linear structure. Users may terminate interactions once their immediate concerns are addressed, reflecting the unique flexibility of chatbots [38]. For instance, if a chatbot provides critical information for crisis management and the user responds affirmatively, the user is likely to exit the chat after their need is met. This immediate resolution highlights the efficiency of chatbot interventions in addressing user concerns promptly, even if the interaction appears incomplete. Recognizing this distinction is essential for accurately interpreting dropout rates and evaluating chatbot effectiveness.

The AI chatbots reviewed here comprised various therapeutic approaches. The incorporation of established psychological treatment techniques such as cognitive behavioral therapy, mindfulness, and dialectical behavior therapy into AI chatbots represents a novel approach to delivering psychological health interventions. This approach is considered novel because it uses AI to deliver evidence-based psychological techniques 24/7, serve multiple users simultaneously, tailor interventions based on individual responses, and combine various therapeutic modalities on a single platform.

This multidisciplinary therapeutic approach has the potential to address health professionals' diverse psychological needs [31]. AI chatbot interventions are characterized by a digital and responsive design, which allows for integrating practical tools such as stress management techniques, mood tracking, and personalized goal setting [39]. These features align with the immediate and personalized support often required by health professionals. However, while personalization is a key advantage of AI chatbots, a need also exists for some level of standardization to ensure the interventions' quality and effectiveness [40]. This standardization should focus on establishing evidence-based core components and assessment methods while still allowing for personalized delivery [31]. Future research should focus on identifying the most effective components of these chatbots and tailoring them to the specific psychological health needs of health professionals while maintaining a balance between standardization and personalization.

In this study, a temporary increase in depression scores was observed in one study [20], whereas previous research has reported AI chatbots' positive effects in reducing anxiety, depression, and work-related burnout among health professionals [41]. These contrasting findings may reflect the challenges faced by such professionals, who frequently experience high levels of stress and emotional demands [42]. This suggests the complexity of psychological interventions and highlights the importance of careful monitoring to determine the appropriate duration and intensity of chatbot use.

Some studies in our review demonstrated the potential of AI chatbots to act as first-line responders in identifying and managing mental health crises. Prior research has also shown the potential benefits of AI chatbots not only in improving mental and emotional well-being but also in promoting behavioral changes [43]. While these features may only benefit a subset of users, their impact on preventing escalation to severe mental health conditions or emergencies is invaluable. Future research should prioritize the integration and evaluation of these features to enhance the safety and efficacy of chatbot interventions.

Although AI chatbots address the unique time constraints and stress among health professionals, balancing automated responses with genuine therapeutic interactions remains a significant challenge. Personalization technologies that provide tailored features to AI chatbots have emerged as key factors in their effectiveness, with studies reporting increased user engagement and satisfaction through customized responses and immediate support [29]. Consequently, the development of tailored intervention strategies that consider the individual needs and preferences of health professionals is essential to maximize AI chatbots' effectiveness. These strategies should integrate ongoing user feedback and dynamically adapt interventions to maintain long-term user engagement and effectiveness. Future research should focus on larger sample sizes, extended follow-up periods, and diverse health care professional groups of health professionals. Such comprehensive studies will play a crucial role in developing evidence-based guidelines for implementing AI chatbots in psychological health and in advancing the sophistication of personalization algorithms.

Limitations

This scoping review has several limitations. First, the limited number of studies adopting diverse study designs and using various AI algorithms complicated making direct comparisons among the retrieved studies. Second, the findings' generalizability is restricted because many studies were published in specific regions, particularly high-income countries. Additionally, the short duration of many interventions and limited follow-up periods make it difficult to evaluate the long-term effectiveness and sustainability of AI chatbots. Future research should address these limitations by conducting more rigorous trials, standardizing the outcome measures, and exploring the long-term applications of AI chatbots across diverse populations and settings.

Conclusions

This scoping review explored the current state of AI chatbots aimed at supporting the psychological health of health professionals. Although the reviewed studies demonstrated AI chatbots' potential to reduce stress, anxiety, depression, and burnout, research in this area remains in its early stages. The diversity of study design, AI algorithms, therapeutic approaches, and outcome measures highlights this field's innovative but fragmented nature.

Despite promising results, particularly with NLP-based chatbots, a significant need exists for more rigorous and standardized studies to fully evaluate their clinical efficacy. Challenges such as usability issues and limited generalizability of the findings must be addressed to enhance the long-term application and effectiveness of AI chatbots in real-world settings. Future research should focus on refining chatbot designs, expanding the research to diverse populations, and conducting long-term studies to clarify the role of AI chatbots in supporting the psychological health of health professionals. Thus, AI chatbots could become a valuable solution for addressing the

XSL•FO
RenderX

psychological health challenges faced by health professionals worldwide.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2024-00351194).

Authors' Contributions

CC conceptualized the study, acquired funding, and oversaw project administration. GB and JH were responsible for data curation and methodology development. CC, GB, and JH conducted the formal analysis. GB created the visualizations. CC, GB, and JH collaboratively wrote the original draft, and all authors contributed to reviewing and editing the manuscript. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Checklist 1
PRISMA-ScR checklist. The checklist has been completed in accordance with the journal's guidelines and includes all the items required for a systematic scoping review.
[PDF File, 98 KB - humanfactors_v12i1e67682_app1.pdf]

References

1. Moss M, Good VS, Gozal D, Kleinpell R, Sessler CN. An official Critical Care Societies Collaborative statement: burnout syndrome in critical care health care professionals: a call for action. Am J Crit Care 2016 Jul;25(4):368-376. [doi: 10.4037/ajcc2016133] [Medline: 27369038]
2. Aronsson G, Theorell T, Grape T, et al. A systematic review including meta-analysis of work environment and burnout symptoms. BMC Public Health 2017 Mar 16;17(1):264. [doi: 10.1186/s12889-017-4153-7] [Medline: 28302088]
3. Maslach C, Schaufeli WB, Leiter MP. Job burnout. Annu Rev Psychol 2001;52(1):397-422. [doi: 10.1146/annurev.psych.52.1.397] [Medline: 11148311]
4. Bakker AB, Costa PL. Chronic job burnout and daily functioning: a theoretical analysis. Burn Res 2014 Dec;1(3):112-119. [doi: 10.1016/j.burn.2014.04.003]
5. Smiley RA, Lauer P, Bienemy C, et al. The 2017 National Nursing Workforce survey. J Nurs Regul 2018 Oct;9(3):S1-S88. [doi: 10.1016/S2155-8256(18)30131-5]
6. Gamble A. Artificial intelligence and mobile apps for mental healthcare: a social informatics perspective. AJIM 2020 Jun 2;72(4):509-523. [doi: 10.1108/AJIM-11-2019-0316]
7. Noble JM, Zamani A, Gharaat M, et al. Developing, implementing, and evaluating an artificial intelligence-guided mental health resource navigation chatbot for health care workers and their families during and following the COVID-19 pandemic: protocol for a cross-sectional study. JMIR Res Protoc 2022 Jul 25;11(7):e33717. [doi: 10.2196/33717] [Medline: 35877158]
8. Fortuna KL, Naslund JA, LaCroix JM, et al. Digital peer support mental health interventions for people with a lived experience of a serious mental illness: systematic review. JMIR Ment Health 2020 Apr 3;7(4):e16460. [doi: 10.2196/16460] [Medline: 32243256]
9. Jacob C, Sanchez-Vazquez A, Ivory C. Social, organizational, and technological factors impacting clinicians' adoption of mobile health tools: systematic literature review. JMIR Mhealth Uhealth 2020 Feb 20;8(2):e15935. [doi: 10.2196/15935] [Medline: 32130167]
10. Grover S, Dua D, Shouan A, Nehra R, Avasthi A. Perceived stress and barriers to seeking help from mental health professionals among trainee doctors at a tertiary care centre in North India. Asian J Psychiatr 2019 Jan;39:143-149. [doi: 10.1016/j.ajp.2018.12.020] [Medline: 30639973]
11. Chaudhry BM, Debi HR. User perceptions and experiences of an AI-driven conversational agent for mental health support. Mhealth 2024;10:22. [doi: 10.21037/mhealth-23-55] [Medline: 39114462]
12. Sarkar S, Gaur M, Chen LK, Garg M, Srivastava B. A review of the explainability and safety of conversational agents for mental health to identify avenues for improvement. Front Artif Intell 2023;6:1229805. [doi: 10.3389/frai.2023.1229805] [Medline: 37899961]
13. Yoon S, Goh H, Nadarajan GD, et al. Perceptions of mobile health apps and features to support psychosocial well-being among frontline health care workers involved in the COVID-19 pandemic response: qualitative study. J Med Internet Res 2021 May 31;23(5):e26282. [doi: 10.2196/26282] [Medline: 33979296]
14. Omarov B, Zhumanov Z, Gumar A, Kuntunova L. Artificial intelligence enabled mobile chatbot psychologist using AIML and cognitive behavioral therapy. IJACSA 2023;14(6):137-146. [doi: 10.14569/IJACSA.2023.0140616]

XSL•FO
RenderX

15. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
16. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5(69):69. [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
17. Babineau J. Product review: Covidence (systematic review software). J Can Health Libr Assoc 2014;35(2):68. [doi: 10.5596/c14-016]
18. Zamani A. Developing a mental health virtual assistance (chatbot) for healthcare workers and their families. : University of Alberta; 2022. [doi: 10.7939/r3-dwgt-jz32]
19. Matsumoto A, Kamita T, Tawaratsumida Y, et al. Combined use of virtual reality and a chatbot reduces emotional stress more than using them separately. J Univers Comput Sci 2021;27(12):1371-1389. [doi: 10.3897/jucs.77237]
20. Kleinau E, Lamba T, Jaskiewicz W, et al. Effectiveness of a chatbot in improving the mental wellbeing of health workers in Malawi during the COVID-19 pandemic: a randomized, controlled trial. PLoS ONE 2024;19(5):e0303370. [doi: 10.1371/journal.pone.0303370] [Medline: 38805444]
21. Kroon YHE. The appreciation of the ithrive chatbot and its impact on self-compassion and compassion fatigue in healthcare workers: a mixed method feasibility study. : University of Twente; 2018 URL: https://purl.utwente.nl/essays/75151
22. Chang CL, Sinha C, Roy M, Wong JCM. AI-led mental health support (Wysa) for health care workers during COVID-19: service evaluation. JMIR Form Res 2024 Apr 19;8:e51858. [doi: 10.2196/51858] [Medline: 38640476]
23. Jackson-Triche M, Vetal D, Turner EM, Dahiya P, Mangurian C. Meeting the behavioral health needs of health care workers during COVID-19 by leveraging chatbot technology: development and usability study. J Med Internet Res 2023 Jun 8;25:e40635. [doi: 10.2196/40635] [Medline: 37146178]
24. Lai J, Ma S, Wang Y, et al. Factors associated with mental health outcomes among health care workers exposed to coronavirus disease 2019. JAMA Netw Open 2020 Mar 2;3(3):e203976. [doi: 10.1001/jamanetworkopen.2020.3976] [Medline: 32202646]
25. Laranjo L, Dunn AG, Tong HL, et al. Conversational agents in healthcare: a systematic review. J Am Med Inform Assoc 2018 Sep 1;25(9):1248-1258. [doi: 10.1093/jamia/ocy072] [Medline: 30010941]
26. Chaudhry B, Islam A. Design validation of a workplace stress management mobile app for healthcare workers during COVID-19 and beyond. : Springer; 2021 Presented at: Proceedings of EAI MobiQuitous 2021. [doi: 10.1007/978-3-030-94822-1_17]
27. Anmella G, Sanabra M, Primé-Tous M, et al. Vickybot, a chatbot for anxiety-depressive symptoms and work-related burnout in primary care and health care professionals: development, feasibility, and potential effectiveness studies. J Med Internet Res 2023 Apr 3;25:e43293. [doi: 10.2196/43293] [Medline: 36719325]
28. Lim SM, Shiau CWC, Cheng LJ, Lau Y. Chatbot-delivered psychotherapy for adults with depressive and anxiety symptoms: a systematic review and meta-regression. Behav Ther 2022 Mar;53(2):334-347. [doi: 10.1016/j.beth.2021.09.007] [Medline: 35227408]
29. Inkster B, Sarda S, Subramanian V. An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: real-world data evaluation mixed-methods study. JMIR Mhealth Uhealth 2018 Nov 23;6(11):e12106. [doi: 10.2196/12106] [Medline: 30470676]
30. Denecke K, Vaaheesan S, Arulnathan A. A mental health chatbot for regulating emotions (SERMO)—concept and usability test. IEEE Trans Emerg Topics Comput 2020;9(3):1170-1182. [doi: 10.1109/TETC.2020.2974478]
31. Vaidyam AN, Wisniewski H, Halamka JD, Kashavan MS, Torous JB. Chatbots and conversational agents in mental health: a review of the psychiatric landscape. Can J Psychiatry 2019 Jul;64(7):456-464. [doi: 10.1177/0706743719828977] [Medline: 30897957]
32. Thunström AO, Carlsen HK, Ali L, Larson T, Hellström A, Steingrimsson S. Usability comparison among healthy participants of an anthropomorphic digital human and a text-based chatbot as a responder to questions on mental health: randomized controlled trial. JMIR Hum Factors 2024 Apr 29;11:e54581. [doi: 10.2196/54581] [Medline: 38683664]
33. Ouali Y, Hudelot C, Tami M. An overview of deep semi-supervised learning. arXiv. Preprint posted online on Jun 9, 2020 URL: https://arxiv.org/abs/2006.05278
34. Chakraborty C, Bhattacharya M, Pal S, Lee SS. From machine learning to deep learning: advances of the recent data-driven paradigm shift in medicine and healthcare. Curr Res Biotechnol 2024;7:100164. [doi: 10.1016/j.crbiot.2023.100164]
35. Agnihotri A, Bhattacharya S. Chatbots' effectiveness in service recovery. Int J Inf Manage 2024 Jun;76:102679. [doi: 10.1016/j.ijinfomgt.2023.102679]
36. Xu Y, Niu N, Zhao Z. Dissecting the mixed effects of human-customer service chatbot interaction on customer satisfaction: an explanation from temporal and conversational cues. J Retail Consum Serv 2023 Sep;74:103417. [doi: 10.1016/j.jretconser.2023.103417]
37. Abd-Alrazaq AA, Alajlani M, Ali N, Denecke K, Bewick BM, Househ M. Perceptions and opinions of patients about mental health chatbots: scoping review. J Med Internet Res 2021 Jan 13;23(1):e17828. [doi: 10.2196/17828] [Medline: 33439133]
38. Ashfaq M, Yun J, Yu S, Loureiro SMC. I, Chatbot: modeling the determinants of users' satisfaction and continuance intention of AI-powered service agents. Telematics and Informatics 2020 Nov;54:101473. [doi: 10.1016/j.tele.2020.101473]
39. Balcombe L. AI chatbots in digital mental health. Informatics 2023;10(4):82. [doi: 10.3390/informatics10040082]

XSL•FO
RenderX

40. Aggarwal A, Tam CC, Wu D, Li X, Qiao S. Artificial intelligence-based chatbots for promoting health behavioral changes: systematic review. J Med Internet Res 2023 Feb 24;25:e40789. [doi: 10.2196/40789] [Medline: 36826990]
41. Abd-Alrazaq AA, Rababeh A, Alajlani M, Bewick BM, Househ M. Effectiveness and safety of using chatbots to improve mental health: systematic review and meta-analysis. J Med Internet Res 2020 Jul 13;22(7):e16021. [doi: 10.2196/16021] [Medline: 32673216]
42. Tamminga SJ, Emal LM, Boschman JS, et al. Individual-level interventions for reducing occupational stress in healthcare workers. Cochrane Database Syst Rev 2023 May 12;5(5):CD002892. [doi: 10.1002/14651858.CD002892.pub6] [Medline: 37169364]
43. Bragazzi NL, Crapanzano A, Converti M, Zerbetto R, Khamisy-Farah R. The impact of generative conversational artificial intelligence on the lesbian, gay, bisexual, transgender, and queer community: scoping review. J Med Internet Res 2023 Dec 6;25:e52091. [doi: 10.2196/52091] [Medline: 37864350]

Abbreviations

AI: artificial intelligence
NLP: natural language processing

Edited by A Kushniruk; submitted 17.10.24; peer-reviewed by C Pankhurst, EM Turner; revised version received 02.02.25; accepted 14.02.25; published 19.03.25.

Please cite as:
Baek G, Cha C, Han JH
AI Chatbots for Psychological Health for Health Professionals: Scoping Review
JMIR Hum Factors 2025;12:e67682
URL: https://humanfactors.jmir.org/2025/1/e67682
doi:10.2196/67682

© Gumhee Baek, Chiyoung Cha, Jin-Hui Han. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 19.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Optimizing Clinical Decision Support System Functionality by Leveraging Specific Human-Computer Interaction Elements: Insights From a Systematic Review

Ali Azadi[1], MSc; Francisco José García-Peñalvo[1], PhD

GRIAL Research Group, Computer Science Department, Universidad de Salamanca, Salamanca, Spain

Corresponding Author:
Ali Azadi, MSc
GRIAL Research Group
Computer Science Department
Universidad de Salamanca
Paseo de Canalejas, 169
Salamanca, 37008
Spain
Phone: 34 691797246
Email: ali.azadi@usal.es

Abstract

Background: Clinical decision support systems (CDSSs) play a pivotal role in health care by enhancing clinical decision-making processes. These systems represent a significant advancement in medical information systems. However, optimizing their effectiveness requires accounting for various human-computer interaction (HCI) elements that influence their functionality and user acceptance.

Objective: This study aimed to identify and categorize key HCI elements that impact CDSS performance to enhance system usability, adaptability, and decision-making accuracy.

Methods: We conducted a systematic literature review, identifying 923 studies from the databases PubMed, Scopus, and Web of Science. Papers were screened and selected based on predefined inclusion criteria. A rigorous quality assessment process was applied to ensure the relevance and reliability of the included studies. Ultimately, of the 923 papers identified, 43 (4.7%) that specifically addressed HCI elements applicable to CDSS environments were included in the final analysis. Data extraction and synthesis were performed to answer the research questions regarding HCI elements.

Results: A total of 12 distinct HCI elements were identified, each with the potential to influence CDSS functionality. These elements align with the International Organization for Standardization (ISO) 9241-11 framework, which defines usability in terms of effectiveness, efficiency, and satisfaction. “User satisfaction,” “flexibility,” and “individuality” enhance satisfaction by improving system adaptability and user acceptance. “Visibility,” “explainability,” and “user control” strengthen effectiveness by supporting decision-making and error prevention. “Ease of use” improves efficiency by streamlining interactions and reducing cognitive load. Some elements influence effectiveness and efficiency, such as “data entry,” which ensures structured inputs for decision accuracy while optimizing workflows. Likewise, “alerts” provide timely information for effective decision-making and, simultaneously, are designed to avoid overwhelming users and maintain system efficiency. “Simplification” and “mental effort” also optimize workflows and reduce complexity. Furthermore, “interface” impacts effectiveness and efficiency by supporting accurate decision-making and streamlining user interaction. This categorization, aligned with ISO 9241-11, underscores the context and task dependency of usability, highlighting that HCI elements must be adapted to different user needs and environments for effective clinical decision-making.

Conclusions: This study addresses a critical gap in CDSS research by offering a comprehensive framework of HCI elements tailored to the CDSS environment. Incorporating these elements into system design can improve user satisfaction, reduce data errors, and enhance the accuracy of medical decisions. The findings lay the groundwork for future research, offering practical guidelines for developing more reliable and efficient CDSS systems in medical informatics fields.

(JMIR Hum Factors 2025;12:e69333) doi:10.2196/69333

XSL•FO
RenderX

KEYWORDS

human-computer interaction; clinical decision support system; usability; user-centered design; artificial intelligence

Introduction

Background

Clinical decision support systems (CDSSs) provide clinicians with computer-generated clinical knowledge and patients with relevant health information presented at the right time to support optimal patient care [1]. Poorly designed user interfaces (UIs) may impede structured data entry, potentially compromising data quality and resulting in incomplete datasets [2]. Such shortcomings can negatively impact the effectiveness of integrated medical information systems when used within a CDSS framework [3]. CDSSs, as a prominent method for enhancing clinical decisions effectively, rely on patient-entered data in electronic medical records (EMRs). Data scientists face challenges in analyzing data obtained from EMRs due to the prevalence of errors, including typos and incorrect entries. In other words, incomplete medical advice originates from incomplete and unstructured prepared data for CDSSs [4]. Physicians often express these errors as stemming from an inadequate interaction with the system [5]. Human-computer interaction (HCI), based on the solid principles of ergonomics, cognitive science, and psychology, is pivotal in creating practical and beneficial technology for all users [6]. Indeed, HCI science advocates for physicians through technological solutions and enhances their comprehension of the significance of their daily processes and the subsequent use of data stored in electronic health records (EHRs) [7].

Although the HCI concept typically encompasses how computer systems are designed for ease of use, efficiency, and effectiveness [8], HCI in the medical realm implies the investigation and design of computer systems that expedite productive and intuitive interaction between physicians and technology to boost user experience (UX) and achieve the desired result [9].

Neglecting HCI elements in medical systems can lead to inadequate graphical UIs, potentially hindering physician adoption and efficacy within the digital health care landscape [10]. Conversely, some studies have demonstrated that incorporating HCI elements improves user acceptance and enhances the effective use of CDSS functionalities [11]. Despite several studies addressing HCI elements and their components, the lack of studies on HCI elements that can affect CDSS performance is quite noticeable. In this regard, related concepts, such as usability, UX, or user-centered design (UCD), all of which contribute to physician satisfaction, need to be investigated. HCI assists in identifying system requirements, including text style, fonts, layout, graphics, and color. Meanwhile, usability focuses on efficiency, effectiveness, utility, ease of learning, and ease of assessment. Combining these perspectives ensures a comprehensive approach to designing a CDSS that meets technical and user-centric demands [12].

The necessity of distinguishing HCI elements specifically for CDSSs, in comparison with decision support systems (DSSs), is derived from the unique demands of the health care environment [13]. CDSSs must seamlessly integrate complex medical knowledge, ensure adherence to stringent clinical guidelines, and accurately handle sensitive patient data, all while minimizing the cognitive load on health care professionals [14]. These systems support clinical decisions that directly impact patient outcomes. This is why elements of usability and UX must be considered. Unlike generic DSSs, CDSSs must navigate the intricate nature of medical decision-making**.** Hence, a tailored approach to HCI elements is needed to address these specialized requirements effectively [15]. In addition, EMRs must be designed using comprehensive HCI elements to serve as a gateway to the data required by the CDSS. Poorly designed HCI in EMRs can lead to improper data entry or retrieval [16]. These errors can have serious consequences, potentially impacting patient safety and even leading to loss of life. Conversely, well-designed HCI can facilitate accurate data handling, improving patient care and significantly enhancing outcomes [17].

Due to the absence of targeted studies that focus on leveraging HCI elements within the CDSS environment, conducting this research is crucial. By systematically identifying and evaluating practical HCI elements specific to CDSSs, this study provides a comprehensive guide to enhance the performance and usability of these systems. This focused investigation offers a unique framework that health care professionals and system developers can use to implement more effective and user-friendly CDSS solutions. The scientific contribution of this study lies in providing a structured methodology to improve CDSS performance, ensuring better integration of HCI elements tailored to the health care context.

Systematic Literature Review

Conducting a systematic literature review (SLR) [18] was the most effective method for this research because it allowed for a rigorous and comprehensive collection of existing studies in the field. An SLR provides a transparent and replicable process for identifying, evaluating, and synthesizing relevant literature. Using the SLR approach, we endeavored to extract and categorize HCI elements that are especially applicable within the CDSS environment, ultimately paving the way for improved CDSS performance.

On the basis of the concerns highlighted in this study, we formulated the following research questions (RQs):

- RQ1: Which HCI elements in medical information systems can affect CDSS functionality?
- RQ2: How does incorporating the identified elements improve CDSS performance?

By seeking a response to the first RQ, we aimed to identify and extract HCI components that have the potential to affect CDSS functionality, as pointed out in the reviewed papers. The second RQ focuses on investigating how the identified HCI elements, when applied within medical environments and information systems, contribute to enhancing CDSS performance.

XSL•FO
RenderX

The remainder of this work is organized as follows. The Methods section describes the methodology and the review steps. The Review Planning subsection details the SLR planning phase, while the Review Process subsection presents the review and data extraction steps. The Results section presents the findings from the analyzed studies to answer the RQs. Finally, the Discussion section presents the principal findings, with some clues and future research lines outlined in the Conclusions subsection.

Methods

Overview

This study followed the SLR methodology outlined by Kitchenham [19] and Kitchenham and Charters [20] to ensure a thorough and unbiased synthesis of existing research. The SLR was complemented by a systematic literature mapping following the method proposed by Kitchenham et al [21]. This segment outlines the protocol used to perform the SLR, providing the necessary information to support the subsequent findings. According to the guidelines provided by Kitchenham [19] and Kitchenham and Charters [20], the SLR process consists of 3 principal phases: planning, conducting, and reporting the study. We adhered strictly to the defined SLR protocols and methodological nuances articulated in the study by Pati and Lorusso [22] to ensure transparency in both the research process and results.

Before planning this SLR, a preliminary search was conducted to ascertain that no recent SLR had been undertaken to identify HCI elements within CDSSs. This preparatory appraisal involved searching various electronic databases, including Scopus, Web of Science, and PubMed, using terms related to the methodology (eg, *SLR* or *systematic literature review*) and the review focus (eg, *HCI*, *usability*, *user experience*, *elements in CDSS*, and equivalent concepts). The results of this preliminary search confirmed the absence of an SLR on the specified theme, thereby justifying the implementation of this review to fill the identified research gap.

Review Planning

The review planning process involves identifying and defining various dimensions to provide the foundation for executing the review, such as formulating the RQs, detailing the protocol followed, and any other relevant information to ensure the traceability of the review process. We outline these dimensions in this subsection.

Mapping Questions

This study's main concerns and RQs are outlined in the Introduction section. To further explore the context and breadth of our SLR, we formulated 4 mapping questions (MQs) to guide our analysis:

- MQ1 (geographic distribution of studies): Where have studies related to HCI elements in CDSSs been conducted worldwide?
- MQ2 (year-wise distribution of studies): What is the distribution of studies by identified HCI element over the years?
- MQ3 (publication venues): What is the distribution of studies between journal papers and conference papers?
- MQ4 (focus distribution): What percentage of the identified studies discuss each HCI element (eg, alerts, interface, system design defects, and mental effort)?

It is important to note that MQ4 (focus distribution) does not have a dedicated subsection. Instead, it has been addressed throughout the identification and discussion of each HCI element. By addressing these MQs, we aim to provide a comprehensive overview of the current research landscape, highlighting key trends and patterns that inform the influence of HCI elements on CDSS performance and user interaction.

Population, Intervention, Comparison, Outcome, and Context Framework

Regarding the questions posed, the population, intervention, comparison, outcome, and context method proposed by Petticrew and Roberts [23] (outlined in the following subsections) and which provides a structured framework for research reviews was used to define the review scope.

Population

We considered all studies that involved software solutions, regardless of their implementation or design. Although the review focuses on the properties and attributes of EMRs, the target studies addressed HCI elements within CDSSs.

Intervention

The interventions of interest included those that suggested or introduced some elements addressing HCI issues in EMRs or eHealth systems, excluding stand-alone solutions such as mobile health apps.

Comparison

No comparison between interventions was planned.

Outcome

The fundamental outcome of this review is to identify key HCI elements that can enhance the functionality of CDSSs.

Context

We considered all studies relevant to EMR and eHealth systems. The review includes all papers describing the successful implementations and designs of EMRs in medical environments, such as hospitals and medical centers worldwide.

Inclusion and Exclusion Criteria

After defining the review's scope, we established inclusion and exclusion criteria (Textbox 1) to choose pertinent literature for addressing the RQs.

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- The paper proposes a pragmatic and implementable solution (eg, method, technique, model, tool, or framework).
- The proposed solution is applied to software, applications, platforms, services, infrastructures, or systems.
- The study focuses on electronic medical records, electronic health records, or eHealth.
- The proposed solution supports or addresses the tailoring of attributes and criteria to improve medical decision support capabilities.
- The paper is written in English.
- The paper was published between 2000 and July 2024 (when the search queries were executed).
- The full paper is available.
- The tailoring capabilities relate to human-computer interaction, usability, user interface, user experience, or user-centered design.
- The article is published in a peer-reviewed journal, book, or conference (including only conferences ranked B or higher in the CORE Conference Ranking list or classified in the top 2 quartiles in the Scimago Journal Rank index).

Exclusion criteria

- The paper does not propose a pragmatic and implementable solution (eg, method, technique, model, tool, or framework).
- The proposed solution is not applied to software, applications, platforms, services, infrastructures, or systems.
- The study is not focused on electronic medical records, electronic health records, or eHealth.
- The proposed solution does not support or address the tailoring of attributes and criteria to improve medical decision support capabilities.
- The paper is not written in English.
- The paper was published before 2000 or after July 2024 (when the search queries were executed).
- The full paper is not available.
- The tailoring capabilities do not relate to human-computer interaction, usability, user interface, user experience, or user-centered design.
- The article is not published in a peer-reviewed journal, book, or conference (including only conferences ranked B or higher in the CORE Conference Ranking list or classified in the top 2 quartiles in the Scimago Journal Rank index).

Search Strategy

It is imperative to identify the most significant databases regarding the search domain in which the queries will be executed to obtain relevant outcomes from the investigation. For this study, 3 electronic databases were selected: Scopus, Web of Science, and PubMed. These databases were chosen based on the following criteria:

- They serve as reference databases in the research domain.
- They are highly relevant to the study context.
- They support the use of search strings and Boolean operators to augment the results of the retrieval process.

On the basis of the research context used to construct the search query, the following terms were included:

- Usability—this concept refers to efficiency, effectiveness, and physician satisfaction [24,25]. It emphasizes maximizing system facilitators to enhance usability [26,27].
- All terms related to HCI—HCI involves interaction and communication between users and computer systems, encompassing the exchange of information, symbols, and actions to facilitate seamless interaction between humans and computers [28]. In this regard, *HCI elements* are defined as the specific elements or attributes that influence user interaction and experience within the determined framework [29]. *HCI elements* and *HCI factors* in medical settings highlight the importance of UCD, ensuring that medical staff are integrally involved in developing and refining medical systems for enhanced performance and satisfaction [30,31]. The terms *HCI factors* and *HCI elements* are sometimes used interchangeably in HCI literature. *HCI factors* often refer to broader influences on usability, while *HCI elements* represent specific design components.
- User-centered design (UCD)—UCD emphasizes the essentiality of considering end users' viewpoints in shaping or evolving a product or software, predominantly based on end-user feedback [32].
- User interface (UI)—in medical information systems, the UI serves as the direct interface between the system and physicians. It can be categorized as static, appearing the same for all users; or dynamic, adapting to varying circumstances based on user interactions with the system [33].
- CDSS—CDSSs encompass a broad range of tools and interventions, both computerized and noncomputerized, to aid clinicians in their complex decision-making processes [34,35]. They are being integrated into EMRs and computerized clinical workflows worldwide, empowering health care providers to access timely and pertinent information, ultimately leading to better patient outcomes through informed decision-making during clinical care [14].
- Artificial intelligence (AI)—in medicine, AI refers to the ability of medical systems to respond to environmental factors and support decision-making, achieving results

XSL•FO
RenderX

comparable to human medical professionals [36]. CDSSs that use AI function as digital dynamic knowledge systems that leverage patient data to create tailored recommendations for clinicians [37]. In other words, CDSSs are meant to assist rather than perform clinical decision-making [38].

In line with the research goals, we only investigated HCI elements that directly or indirectly impact systems designed for medical decision-making. Hence, HCI elements not relevant to this context were excluded from consideration.

Quality Assessment

Overview

Although inclusion and exclusion criteria help identify relevant works for a study, they do not address the quality of the papers retrieved regarding their ability to address the RQs. Therefore, we developed a separate set of quality assessment criteria. We adopted a scoring system in which each coauthor independently evaluated the studies using the quality questionnaire proposed by Kitchenham and Brereton [39]. The scoring was as follows: a "yes" response earned 1 point, a "partially" response earned 0.5 points, and a "no" response earned 0 points. The maximum score a study could receive was 7 points. The quality questionnaire and its corresponding assessment phases are detailed in the following subsections.

Design Clarity and Relevance

Are the research aims and objectives clearly stated and aligned with the study's focus on HCI elements in CDSSs?

Data Collection and Relevance

Does the paper provide comprehensive and relevant data pertinent to HCI elements in CDSSs, including metrics for evaluation?

Empirical Measurement and Methodology

Are the HCI elements empirically measured with well-defined metrics? Is the methodology for evaluating these elements clearly described?

Analysis and Documentation

Are the research results documented with granularity, including participant information, observational units, and analysis methods?

Conclusion Validity and RQs

Do the study's conclusions adequately answer the RQs? Are the implications for HCI and CDSSs clearly articulated?

HCI Evaluation Method

Is the HCI evaluation method used in the study justified and described in sufficient detail?

Relationship Between HCI Elements and CDSS Outcomes

Does the study adequately illuminate the relationship between the applied HCI elements and the CDSS outputs or results and their consequences?

Quality Assessment and Interrater Reliability

The maximum score a study could receive was 7 points. To advance to the next phase of the review, studies were required to score at least 5 out of 7 points. To validate the quality assessment results and demonstrate process consistency, we computed a measure of interrater reliability using Krippendorff α [40]. The Krippendorff α value was 75.33%, indicating that the data were interpreted similarly and acceptably among the coauthors. When discrepancies arose, we engaged in discussion sessions to reach a consensus.

Query Strings

Overview

The search strings for each considered source were generated using relevant search terms derived from the population, intervention, comparison, outcome, and context methodology outcomes [41], connected by Boolean "and" or "or" operators. The canonical search equation provides a standardized template that ensures consistency across different databases. This canonical search equation is then adapted to fit the specific syntax of each database. The canonical search equation is as follows: ("human-computer interaction" OR "HCI" OR "usability" OR "user experience" OR "user interface" OR "user-centered") AND ("electronic medical records" OR "EMR" OR "electronic health records" OR "EHR") AND ("clinical decision support systems" OR "CDSS" OR "decision support system*" OR "DSS" OR "artificial intelligence" OR "AI").

Using the canonical search equation, we formulated the following queries tailored to each of the 3 selected databases to retrieve all relevant studies, including those addressing the implications of physicians' interactions with systems. Each database required distinct syntax, as outlined in Textbox 2.

XSL•FO
RenderX

Textbox 2. Search syntax for each database.

Scopus

(TITLE-ABS-KEY (“human-computer interaction” OR “HCI” OR “usability” OR “user experience” OR “user interface” OR “user-centered”) AND TITLE-ABS-KEY (“electronic medical records” OR “EMR” OR “electronic health records” OR “EHR”) AND TITLE-ABS-KEY (“clinical decision support systems” OR “CDSS” OR “*decision support system*” OR “DSS” OR “artificial intelligence” OR “AI”))

Web of Science

(TS= (“human-computer interaction” OR “HCI” OR “usability” OR “user experience” OR “user interface” OR “user-centered”) AND TS= (“electronic medical records” OR “EMR” OR “electronic health records” OR “EHR”) AND TS= (“clinical decision support systems” OR “CDSS” OR “decision support system*” OR “DSS” OR “artificial intelligence” OR “AI”))

PubMed

((“human-computer interaction” OR “HCI” OR “usability” OR “user experience” OR “user interface” OR “user-centered”) AND (“electronic medical records” OR “EMR” OR “electronic health records” OR “EHR”) AND (“clinical decision support systems” OR “CDSS” OR “decision support system*” OR “DSS” OR “artificial intelligence” OR “AI”))

Limiting Retrieved Results

Each database required distinct syntax, as outlined in the query strings presented above. In subsequent stages of this review, the retrieved results were limited to those that had the potential to address the predefined RQs.

Review Process

The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [42] were followed for reporting this SLR. Refer to Multimedia Appendix 1 for the PRISMA checklist.

The data-gathering process was segmented into several stages, during which various inspections were performed.

Once the search was completed on July 25, 2024, the paper selection process was conducted in the following stages:

1. The raw data, consisting of records retrieved from the 3 databases, were compiled into a Zenodo repository (Multimedia Appendix 2). This dataset, which contains data from 923 papers, is organized into separate sheets for each database: Scopus (n=506, 54.8%), Web of Science (n=188, 20.4%), and PubMed (n=229, 24.8%).
2. After arranging the records, duplicate papers were eliminated. Of the 923 papers, 326 (35.3%) were removed, leaving 597 (64.7%) for the next stage.
3. At the first step of the screening, reading the titles, abstracts, and keywords and applying the inclusion and exclusion criteria resulted in 444 (74.4%) of the 597 papers being excluded, leaving 153 (25.6%) for the next phase.
4. Of these 153 papers, the full text of 5 (3.3%) was not accessible; hence, we read in detail and further scrutinized 148 (96.7%) papers. After an exhaustive examination of these 148 papers, we removed 105 (70.9%) for the following reasons: 49 (46.7%) involved HCI elements in medical environments, but these elements were unrelated to CDSSs; 38 (36.2%) did not explicitly focus on HCI or usability elements, although they discussed CDSS functionality; and 18 (17.1%) did not clearly explain the relationship between the identified HCI elements and CDSS functionality.
5. Ultimately, of the initially identified 923 papers, 43 (4.7%) were selected for the final analysis and review (Multimedia Appendix 3 [7,43-84]).

Results

Overview

This section provides a detailed analysis of the findings derived from our SLR on HCI elements that influence CDSSs. By thoroughly examining the 43 articles included for analysis, we aimed to answer the formulated RQs and MQs, thereby delineating the boundaries of this research area.

Our RQs focused on identifying the HCI elements that influence medical data management and CDSS functionality (RQ1) and understanding how these identified HCI elements can impact CDSS performance (RQ2). These questions were essential for understanding the relationship between HCI elements and the efficiency, accuracy, and usability of CDSSs.

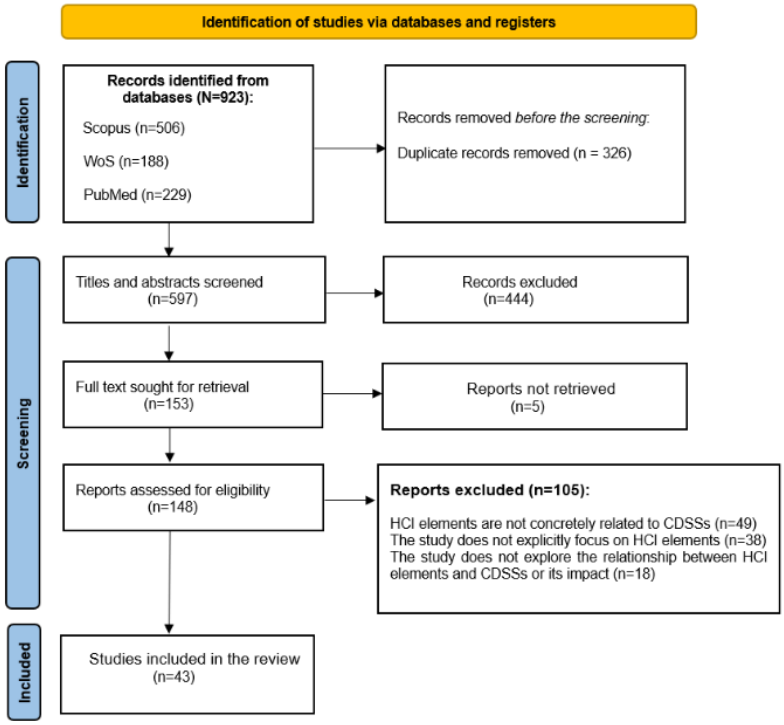
In addition to these RQs, we formulated several MQs to comprehensively explore the scope and context of the studies. These included understanding the geographic distribution of research efforts in this area, analyzing the temporal distribution of the included studies, and identifying their publication venues. Furthermore, we examined the topic distribution in this SLR to determine the percentage of attention each identified HCI element received across the reviewed studies.

This analysis provides insights into the HCI elements investigated in the studies and how they influence CDSS functionality. This comprehensive overview offers valuable information on how HCI elements can be optimized to enhance the overall performance and effectiveness of CDSSs, ultimately contributing to improved patient care and decision-making within medical settings.

The PRISMA flow diagram [42] illustrates the steps taken to extract the required data (Figure 1).

XSL•FO
RenderX

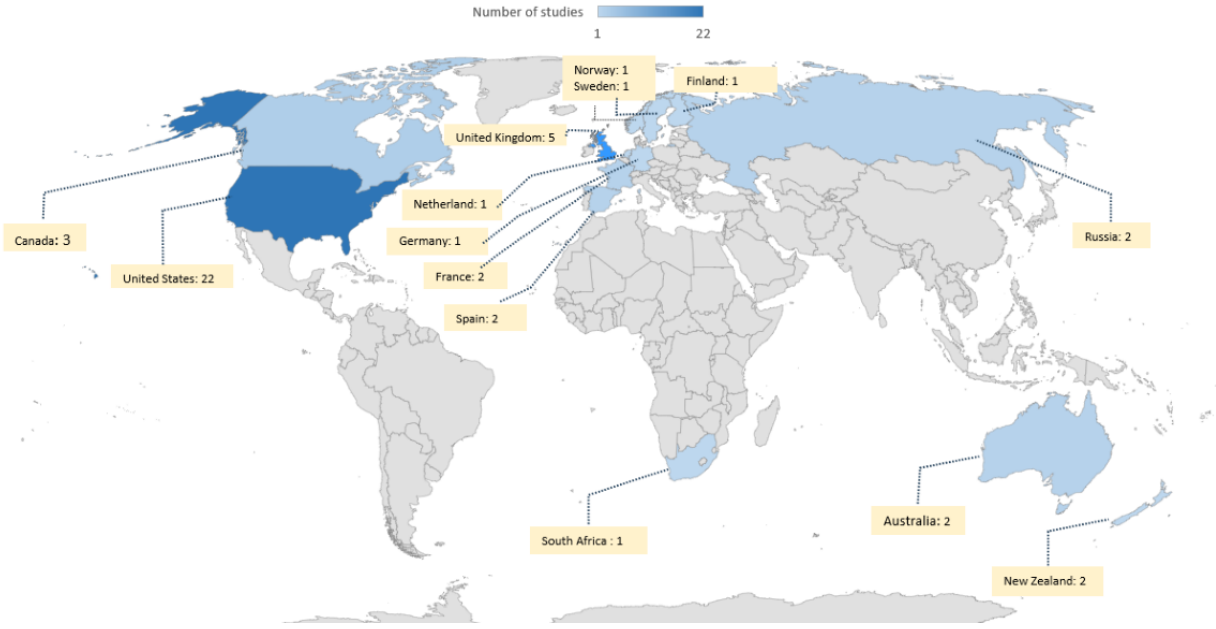
Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram showing the number of studies identified, screened, assessed for eligibility, and included in the final analysis. CDSS: clinical decision support system; HCI: human-computer interaction; WoS: Web of Science.



Geographic Distribution of Studies

Our investigation to address MQ1 was carried out by considering the affiliations of all authors. Consequently, a study with authors from Spain, the United Kingdom, and Sweden was attributed to all 3 countries. In total, 43 studies were conducted across 14 countries. As depicted in Figure 2, the geographic distribution reveals a significant concentration of research in the United States and Europe, reflecting global interest in enhancing CDSSs through improved HCI elements. The United States contributed the most publications, with 51% (22/43) of the relevant studies, followed by the United Kingdom with 12% (5/43); Canada with 7% (3/43); and Russia, Spain, France, and Australia with 5% (2/43). The remaining countries—Sweden, South Africa, Norway, New Zealand, Finland, and Germany—each accounted for 2% (1/43) of the studies.

Figure 2. Geographic distribution of the included studies.



Despite the rapid technological advancements and growing health care needs in Asia, our analysis revealed a notable lack of studies relevant to HCI elements in the CDSS realm from Asian countries. Addressing this discrepancy is crucial for developing cultural and contextual usability solutions to enhance CDSS performance.

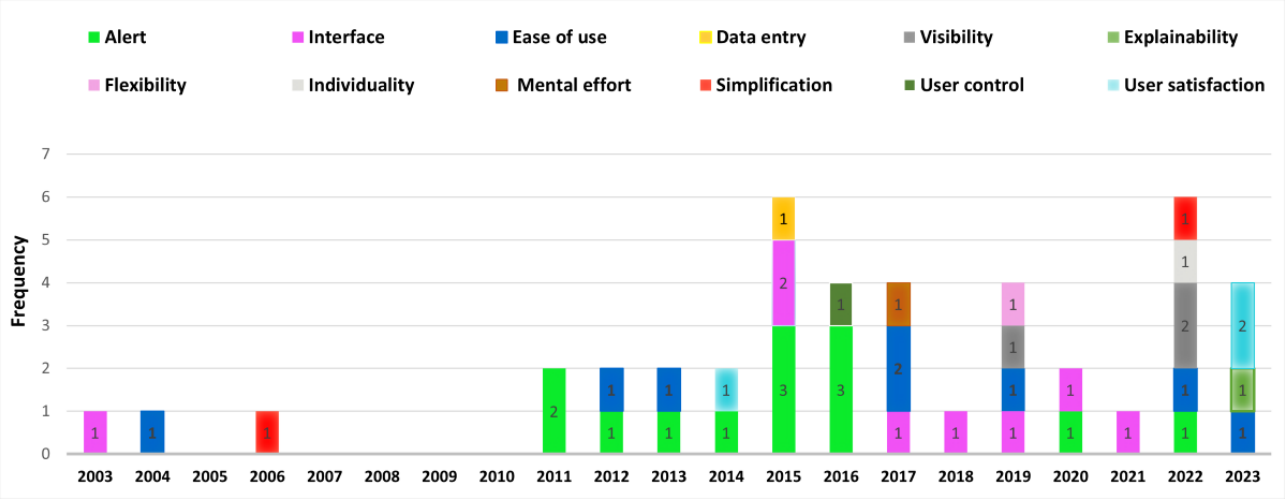
XSL•FO
RenderX

Our exploration indicates that the emphasis on distinct HCI elements varies by country based on specific demands, policies, and priorities. In the United States, which has the highest research activity in this field, studies have predominantly investigated HCI elements related to CDSS alerts (9/22, 41%), UI (3/22, 14%), and ease of use (2/22, 9%). Although these core HCI elements are imperative, some specific HCI demands may be considered depending on cultural differences and local needs [85].

Year-Wise Distribution of Studies by Identified HCI Elements

Through the review process, we identified 12 HCI elements. Each included paper primarily focuses on 1 main element, as shown in Figure 3. While some papers discuss multiple HCI elements, the figure presents the dominant HCI element addressed in each paper. The bar graph illustrates which HCI elements related to CDSS received attention each year. Although the gradual formation and evolution of CDSS technology began in 1990 [14], HCI elements have gained significant attention in medical settings since the 1990s, with increasing emphasis on usability, UX, and interface design in health care technology [86]. This study depicts the conjunction of 2 concepts—HCI and CDSS functionality—that have emerged since 2003 and have seen significant advances, especially since 2015. Figure 3 highlights this growth, which is tightly related to MQ2. As shown in Figure 3, an investigation gap existed in this realm between 2006 and 2011, and most studies on this subject were carried out in 2015 and 2022. Since 2011, there has been a growing interest in exploring the relationship between HCI elements and CDSS functionality.

Figure 3. Scattering leverage of human-computer interaction elements in clinical decision support systems over the years.



This trend reflects the increasing comprehension of the importance of UCD in health care technology, aiming to improve the usability and effectiveness of CDSSs in supporting clinical decision-making processes [87].

Publication Venues: Distribution by Study Type (Journal, Conference, or Book Chapter)

Nearly all of the included papers (42/43, 98%) were published in academic journals. This indicates a strong preference for disseminating research findings related to HCI elements in CDSSs through journal publications, which are often peer reviewed and highly regarded in the academic community. Moreover, 1 (2%) of the 43 papers was a book chapter, which provides comprehensive overviews and in-depth discussions on specific topics, contributing valuable insights to the field.

This distribution, as the explicit answer to MQ3, highlights the credibility and rigor of the selected studies, ensuring that the insights and conclusions drawn from this review are thoroughly examined and well substantiated.

Identified HCI Elements That Influence CDSS Functionality and Performance

Overview

As discussed in the previous subsections, regarding our explicit question about HCI or user perspective elements with the potential to impact CDSS functionality, we identified 12 elements after a full-text study of 43 selected papers. These elements were categorized based on their inherent nature, application context, and significance to the UX, particularly from the perspective of medical staff. As illustrated in Table 1, these elements address different HCI concerns. Although HCI elements such as “interface” inherently encompass subelements such as “visibility,” “user control,” and “data entry,” these subelements have been presented independently to highlight their unique roles and distinct contributions to CDSS functionality and usability; for example, “data entry” focuses on minimizing input errors through standardized workflows, whereas “visibility” prioritizes intuitive data presentation—distinct concerns that have been highlighted in studies on clinical efficiency and user trust [60,88].

XSL•FO
RenderX

Table 1. Extracted human-computer interaction (HCI) elements and insights.

HCI elements and topics and addressed concerns	References
Alerts	
Usability flaws	[43,44]
Alert fatigue	[7,45-47]
Alert design recommendations	[48-53]
Barriers (too many alerts)	[54]
Ease of use	
Intention to use	[55]
Understandable	[56]
Customizability	[57]
User-friendly features	[58]
Jargon avoidance	[59]
Consistency of operations	[60]
Goal-directed design	[61]
Facilitators (automating data fetching)	[62]
Interface	
Acceptability	[63]
Information recall issues	[64]
Data presentation	[65]
Visualization	[66]
Interface features	[67]
Physical layout	[68]
Facilitators and barriers to presenting diagnostic information	[69]
Heuristic semantic tags (to improve interface design)	[70]
Visibility	
Error visibility	[71]
Patient status visibility	[72]
Importance-based highlighting	[73]
Simplification	
Removing extraneous information and procedures	[74]
Resolving the conflict between system simplicity and usefulness	[75]
User satisfaction	
Users' expectations and learnability	[76]
Effectiveness, efficiency, and accessibility	[77]
Barriers (mismatch between meaningful-use criteria and physicians' expectations)	[78]
Explainability	
Transparent and informative recommendations	[79]
Flexibility	
Ability to change order and revise data entry	[80]
User control	
Controllability for error prevention	[81]
Data entry	

XSL•FO
RenderX

HCI elements and topics and addressed concerns	References
Accurate and structured data entry	[82]
Individuality	
Individualized and user-centered approach	[83]
Mental effort	
Cognitive load	[84]

This categorization of HCI elements aligns with the International Organization for Standardization (ISO) 9241-11 framework [89], which defines usability as a combination of effectiveness, efficiency, and satisfaction in achieving specific goals within a particular context. We found some studies (3/43, 7%) that directly evaluated CDSSs to assess "user satisfaction," focusing on the outcomes of user-system interaction. These studies align with the satisfaction component of the ISO 9241-11 standard, emphasizing users' subjective experience and acceptance of the system. In comparison, "visibility" supports the effectiveness dimension of this standard by ensuring that users remain informed about the system's status in a timely manner, facilitating accurate decision-making. Conversely, "ease of use" corresponds directly to the efficiency dimension of ISO 9241-11 because it minimizes the cognitive and physical effort required to interact with the system, enabling users to achieve their goals with minimal waste of resources or time. Moreover, "explainability" enhances effectiveness by improving users' understanding of system outputs, while "alerts" contribute to both effectiveness and efficiency, ensuring that critical information is promptly communicated without overwhelming users. Another HCI element, "user control," also strengthens effectiveness by aiding error prevention and recovery, allowing users to rectify mistakes and maintain workflow reliability. As the "interface" element impacts decision accuracy and interaction speed, it is strongly related to effectiveness and efficiency. "Flexibility" and "individuality" primarily support satisfaction because they enable system adaptability to different user needs and preferences, fostering trust and engagement. "simplification" and "mental effort" impact both efficiency and effectiveness because they reduce cognitive load, eliminate unnecessary complexity, and optimize workflow processes. Another HCI element, "data entry," contributes to accurate information for decision-making (effectiveness) and streamlines the input process to save time and reduce effort (efficiency).

The rationale for this categorization is clarified in Table 1, which outlines the main concern addressed by each study and explains how it relates to a specific HCI element. The subsections that follow provide a detailed explanation of each HCI element listed in the table.

Alerts

Overview

Alerts have been identified in this study as a major HCI element. Alerts in CDSSs are essential for enhancing patient safety and clinical efficiency. They provide immediate, actionable information to health care providers. They can warn about potential drug interactions, highlight critical patient allergies, or flag abnormal laboratory results [48]; for instance, an alert might notify a physician about a dangerously high potassium level in a patient, prompting immediate intervention. These alerts help prevent errors, ensure timely responses, and support adherence to clinical guidelines, ultimately improving patient outcomes and streamlining clinical workflows [90].

To explicitly investigate alert-related concerns within CDSSs, it is crucial to differentiate between alerts and reminders because conflating these elements can hinder effective system design. In the context of HCI within CDSSs, alerts and reminders should be discussed separately, given their distinct functions [91]. Alerts are immediate notifications requiring urgent attention, supporting real-time decision-making with minimal cognitive load. Reminders are persistent prompts designed to ensure that critical tasks are not overlooked, thereby improving adherence to clinical guidelines. They are defined as a communication or message to ensure that physicians remember critical tips [92]. Reminders can cause cognitive overload and frustration if they are too frequent or irrelevant—in such cases, they are categorized as barriers. Conversely, when reminders assist physicians, they are considered facilitators [93]. This distinction ensures a more effective optimization of HCI elements in CDSSs, balancing support and usability.

The findings from this SLR demonstrate that 31% (13/43) of the selected papers addressed alert-related topics, while Figure 3 shows that this element was mostly studied between 2011 and 2016 (11/13, 85%), although some studies (2/13, 5%) have been conducted in recent years (2020 and 2022). Research on alerts has focused on 3 main areas. Table 1 presents the studies that have addressed these categories, which are detailed in the following subsections.

Alert Fatigue

Alert fatigue poses a significant challenge for CDSS instruments because users may begin to neglect or override activated alerts due to their frequent occurrence [94]. In 2016, Gong and Kang [47] presented five solutions to address alert fatigue: (1) augmenting alert specificity, (2) tiering alerts based on severity, (3) applying human factors principles (such as format, content, legibility, and color), (4) customizing alerts based on patient attributes, and (5) providing tailored alerts for medical practitioners.

The timing of alerts is crucial because they should be presented at moments conducive to informed medical decision-making. It is important to strike a balance: alerts should be intrusive enough to capture attention, but they should not distract or divert users from the primary care pathway [45]. Maintaining this balance prevents alert fatigue, sometimes referred to as habitual override.

XSL•FO
RenderX

Usability Flaws

Usability flaws relevant to CDSSs that diminish alert performance have been identified as follows [43]:

- The distinguishing visualizations do not illustrate the variety of severity levels in the alerts.
- The information presented within alerts is dense and lacks brevity.
- A low signal-to-noise ratio, reflecting a high proportion of erroneous alerts compared to correct ones, often results from nonupdated or incorrect data.
- Alert content issues stem from missing information relevant to alert goals, data interpretation, or necessary practical recommendations.
- There is a scarcity of transparency regarding the reasons that have led to the triggering of the alert.
- The clinicians are not directed at fixing the problem identified by the alert.
- Adaptability issues point out the insufficient adjustability of the system to accommodate all types of users.
- Workload issues arise when too many tasks need to be performed to correct errors or obtain the required information.

In medical information systems, 2 main types of alert flaws are commonly observed: general and specific. General alert flaws occur when alerts lack specificity or fail to distinguish between different severity levels [44]. Specific alert flaws arise when alerts do not adequately address individual patients' or health care providers' unique contexts or needs, leading to ineffective or irrelevant notifications [95].

Alert Design Recommendations

Alert design recommendations have been addressed across the following aspects to enhance alert effectiveness (Textbox 3). Enhancements in alert functionality within CDSSs can optimize clinical workflows, ensure timely access to relevant information, mitigate alert fatigue, foster user trust, and ultimately contribute to more informed and effective clinical decision-making processes [68].

Textbox 3. Alert design recommendations.

- The visual presentation of alerts can be improved by using different colors, bullet points, and clear textual guidelines [48].
- Alerts should be triggered at appropriate points within the clinical workflow and should align with real-world practices [50].
- False alarms that can occur when the alert's logic is flawed or not based on updated information must be avoided. In some cases, rigid sensitivity calibrations can result in false alarms [49].
- The presentation of alerts should be determined based on their level of importance; for instance, higher-priority alerts may be presented in a way that interrupts the current workflow [50].
- Habituation pertains to repeated exposure to insignificant alerts. This underscores the importance of reducing the occurrence of false alarms, along with adopting an alarm philosophy aimed at minimizing alert overrides [49].
- Consistent terminology emphasizes the use of standardized and predefined words and expressions. This enhances the ability to locate specific words or data on the screen through visual screening and promotes uniformity in the generated data irrespective of geographic location [96].
- Mental models refer to an individual's interpretation of a specific matter that influences their reaction when an alert is triggered; for instance, the color red typically evokes an immediate association with "STOP" [49].
- The content of the alert—the reason for its activation and potential medical consequences—should be kept concise [97], with additional details accessible through related data links [98].
- Font style and size can be used to convey the importance and prioritization of the alert [51,99].
- Alert visibility is a critical HCI component that should be considered in the design of alerts within the CDSS environment [52]. The alert's dimensions as displayed on the screen (target size), luminosity, background contrast, and typography attributes should be carefully considered to quickly capture the user's attention [100,101].
- The degree of workflow interruption must be proportional to the severity of the medical issue that triggered the alert [51]. For high-priority alerts, it may be necessary to interrupt the workflow to prompt immediate user intervention [52].
- To ensure the proximity of decision-making components, decision support tools must be incorporated within the alert; for example, the alert should include hyperlinks to medical reference websites [53].

Excessive and Unnecessary Alerts

Particularly in specific patient populations or under uncommon circumstances, excessive and unnecessary alerts represent a significant barrier to the effective use of alerts within CDSSs. This obstacle can be resolved by grouping similar alerts into a single notification and tailoring alert thresholds and delivery methods based on patient condition and clinical context [54].

Ease of Use

Ease of use was identified as a principal user concern in 19% (8/43) of the selected papers. This element is foreseeable and represents a standard expectation among users, particularly physicians. In this respect, systems should be designed so that users easily understand how to operate the system. This is referred to as perceived ease of use. A positive perceived ease of use is associated with an increased intention to use (IU) [55].

The concept of consistency will augment a system's ease of use. It can be achieved in 2 distinct areas: jargon-free language and consistent operative patterns. Consistent terminology and operation across all system components minimize user confusion and cognitive load, enabling users to quickly and effectively achieve their goals [59].

XSL•FO
RenderX

In essence, a system that is easy to use enhances learnability because users can understand and remember operational procedures more quickly [102].

In a CDSS environment, ease of use can significantly boost system functionality by ensuring that health care professionals can quickly and without complexity access and apply clinical guidelines and decision aids [58]. In this regard, a key facilitator is automated data fetching, which streamlines the retrieval of relevant information, minimizing complexity and allowing users to focus on patient care [62].

Moreover, goal-directed design is a user-centered approach that prioritizes understanding user needs and goals throughout the design process. By engaging end users throughout the design process, goal-directed design ensures that the system effectively supports their tasks and contexts. This alignment directly contributes to ease of use by streamlining interactions and enhancing UX [61].

Interface

Interface, another important HCI element**,** was addressed in 19% (8/43) of the selected papers, especially between 2015 and 2021. While interface design inherently encompasses various HCI elements, the analyzed papers offer unique insights that extend beyond the scope of other discussed elements. These papers provide specific recommendations for interface improvements that can significantly enhance user interaction by optimizing system design and fostering intuitive use.

The interface of a CDSS directly influences acceptability by determining how intuitively users can access and act on clinical information. When the interface aligns with users' workflows and preferences, it increases their willingness to adopt and use the system effectively [63]. Despite recent advances in computerized technologies, the defective design of graphical UIs in clinical settings may lead to frustration among physicians [10]. In the context of CDSSs, the UI refers to the graphical or visual representation through which users interact with the system to access, input, and interpret information [63]. An effective UI should be intuitive, user-friendly, and tailored to health care professionals' specific needs and workflows [64].

Via integration with EHRs, well-designed interfaces within CDSSs can present a list of possible diagnoses associated with medical symptoms, the compatibility percentage of the patient's current medical status, and the factors contributing to the diagnosed illness [65,103]. In addition, effective data visualization enables quick comprehension of complex information, promoting time efficiency by enabling health care professionals to rapidly interpret data without the need for extensive analysis [66].

A responsive and interaction-enabled UI encompasses 2 distinct features: presentation and placement. In terms of presentation, elements such as simplicity, appropriate font size, meaningful colors, acceptable contrast, and bold text enhance readability and user engagement. Regarding placement, information should occupy a prominent position, be localized for easy access, and use multiple presentation layers to facilitate quick comprehension. Together, these interface features facilitate the seamless flow of essential medical data to decision makers [67].

In the interface realm, the physical layout concept refers to the arrangement and organization of screen objects with which users interact. This layout significantly impacts user control, allowing health care professionals to directly manipulate the screen's objects (such as buttons, sliders, and forms). When users have clear control over screen objects, they can more effectively navigate the system, customize their workflow, and respond to patient needs promptly, ultimately improving overall usability and satisfaction with the CDSS [68].

There are some facilitators and barriers relevant to the interface within CDSS environments. One of them is the content-based facilitator that enhances interface usability by presenting clear, concise, and relevant information. In this respect, tailoring content to meet the health care professionals' specific needs and preferences and providing real-time updates on clinical guidelines will enhance the interface. Conversely, inconsistent content within the data presentation and guidelines, potentially leading to confusion and misinterpretation, has been identified as an interface barrier [69].

One reviewed paper introduced the concept of "heuristic tags" to enhance CDSS interface design [70]. The paper describes a container comprising 14 HCI elements relevant to CDSS environments, referred to as semantic tags [70]. These elements are outlined in Table 2.

XSL•FO
RenderX

Table 2. Semantic tags.

Tag	Description
Consistency and standards	All system providers throughout the design must adhere to a unified protocol by maintaining consistency in terminology, sequence of actions, and data localization.
Visibility of the system status	Users should be notified of ongoing processes via suitable alerts and feedback mechanisms.
Match between system and world	Users' perceptions of the system should align with their mental models of how the system is expected to function.
Minimalist design	Any unnecessary information should be eliminated because it acts as a distraction and impedes efficiency.
Minimize memory load	Users should not be required to memorize extensive information to perform routine tasks.
Informative feedback	Users must receive informative feedback regarding their activities.
Flexibility and efficiency	The system should provide resiliency so that users can customize settings and expedite their tasks.
Good error messages	Messages should provide sufficient information for users to comprehend the nature of errors, learn from mistakes, and take corrective action.
Prevent errors	The design should deter errors by preventing incorrect actions.
Clear closure	Users must be explicitly informed when a task has been completed.
Reversible actions	Users should be allowed to recover from errors made by them.
Use user's language	The system language must be comprehensible to the targeted users.
Users in control	Users should not feel in control of the system without encountering unforeseen circumstances.
Help and documentation	The system should offer various forms of help, such as contextual assistance, mission-focused guidance, and alphabetically organized (lexicographically arranged) help topics for easier navigation.

Therefore, interface quality will influence the precision of medical decisions by facilitating intuitive access to relevant information and enhancing user interaction.

Visibility

Visibility was addressed in 7% (3/43) of the selected papers, emphasizing that quick access to information requires data to be readily accessible within a short time frame and ensuring that medical data are available at expected locations. This includes the clear and immediate visibility of possible errors or system warnings, allowing users to identify and address issues quickly [71]. Within the treatment cycles pertinent to CDSSs, visibility refers to the clarity and transparency of the workflow stages and progress, meaning that health care providers can easily track and understand the status of the treatment process, including diagnostic tests, medication administration, and patient outcomes [72]. In addition, by visually emphasizing critical information such as high-risk alerts or abnormal laboratory values through color coding, bold text, or other visual cues, clinicians can quickly identify and prioritize important information [73]. As visibility criteria provide a clear and transparent view of the treatment cycle workflow, they enable users to easily monitor the progress of patient care and make informed decisions based on the current stage of treatment, leading to the best possible decision [73].

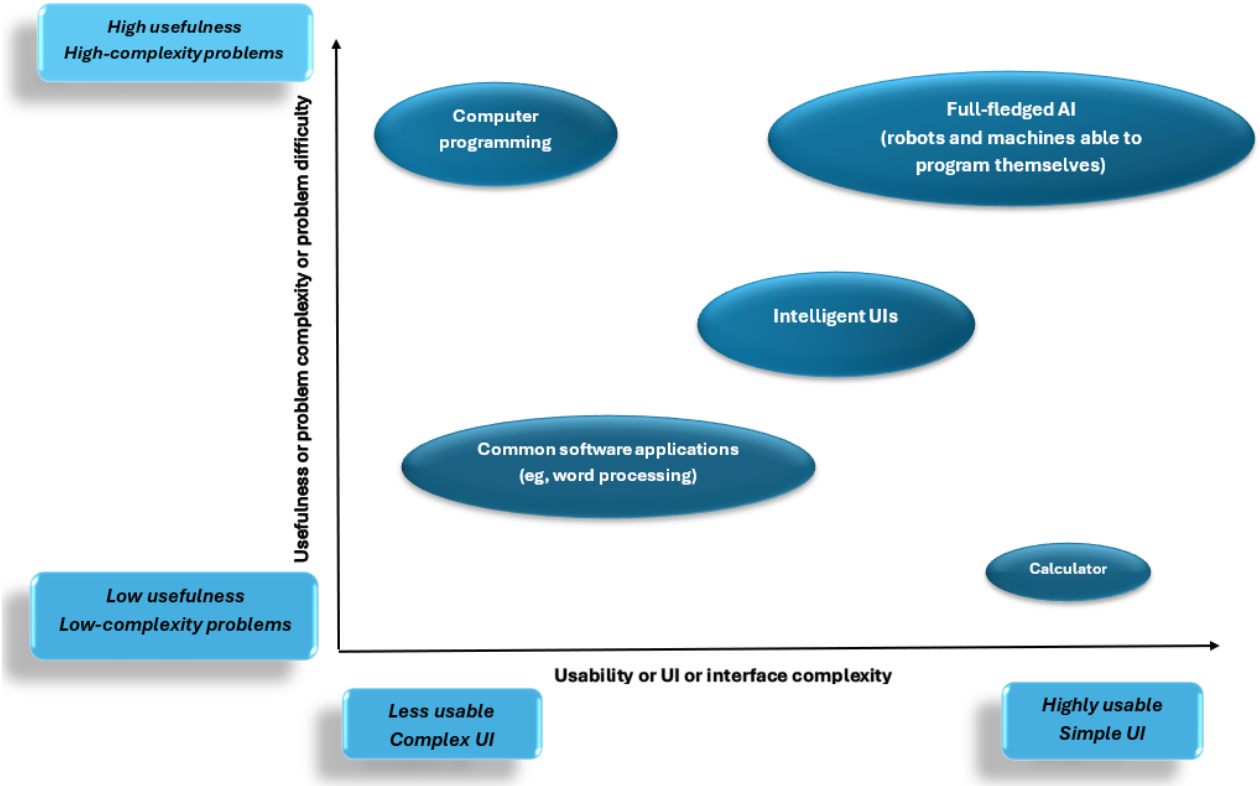
Simplification

Of the 43 selected papers, 2 (5%) addressed simplification elements. Simplification can be achieved by eliminating unnecessary complexities and providing an optimized workflow congruent with real-world practices [104]. Removing extraneous information and procedures is crucial for achieving this goal [74]. Pantazi et al [75] propounded a theory describing a paradox between simplification a medical system and its usefulness: systems with high usability typically can solve trivial problems, whereas solving intricate problems requires an appropriate level of complexity, which can reduce overall usability.

As shown in Figure 4 [75], some CDSS applications designed for intricate problems have complex UIs with lower usability levels, whereas sections with simpler structures within CDSSs, such as medical calculators, exhibit higher usability levels. Meanwhile, system designers are striving to develop medical AI systems that simultaneously enhance both usability and problem-solving capacity.

XSL•FO
RenderX

Figure 4. The relationship between system complexity and usability. AI: artificial intelligence; UI: user interface.



Nonetheless, while striving for complexity in medical procedures can be inherently beneficial, it is crucial to maintain a balance between usability and necessary intricacy [105]. Striking this balance ensures that procedures are not unnecessarily complex, saves time, and ultimately enables rapid emergency medical decision-making [106,107].

User Satisfaction

User satisfaction was discussed in 7% (3/43) of the selected papers. Recognized as a central focus in HCI research, user satisfaction is paramount to the effectiveness of CDSSs. This concept encompasses aspects such as system acceptability, learnability, memorability, and accessibility [76].

UCD methods are essential for achieving high user satisfaction [108]. By prioritizing the user's needs and perspectives throughout the development process, CDSSs can be tailored to meet their specific requirements [109]. Learnability facilitates users in quickly understanding how to operate the system, which enhances initial engagement [110]. Memorability ensures that the system's functions are easily recalled once learned, reducing the need for repeated training [111]. Accessibility ensures that the required medical information and history are promptly accessible in the system [77]. Incorporating these elements boosts overall user satisfaction, improves efficiency and adoption, and ultimately leads to better patient outcomes in clinical environments.

Conversely, the mismatch between meaningful-use criteria and physicians' expectations represents a significant barrier to user satisfaction within CDSSs. In these cases, the requirements set by system designers do not align with physicians' practical needs and processes, leading to frustration and decreased usability. This misalignment can hinder the effective adoption of CDSSs because physicians may find the systems cumbersome or irrelevant to their clinical workflow. Addressing this gap is essential for improving user satisfaction and ensuring that CDSSs effectively support health care professionals in delivering high-quality patient care [78].

Explainability

Explainability, as identified in 2% (1/43) of the selected papers, is a critical focus in the ongoing discourse surrounding CDSSs and medical AI [112]. Transparency, a cornerstone of explainability, fosters trust between the user and the system [113]. Once the CDSS provides recommendations and solutions, physicians can understand why and how the advice was generated. This understanding augments their clinical expertise and contributes to more informed decision-making over time [79].

Flexibility

Flexibility was articulated in 2% (1/43) of the selected papers. This element requires significantly more attention and could be leveraged to enhance CDSS functionality. Adjustable medical systems typically achieve higher adoption rates, particularly when they meet physicians' need for flexibility [114]. In this context, 3 diverse dimensions related to flexibility concerns have been identified [80]: (1) user tendency to return to previous stages to correct or modify information; (2) data entry in multiple different sequences; and (3) customizable data presentation.

Flexibility is a crucial attribute when designing digital pathways. By aligning flexibility with actual clinical workloads**,** treatment

XSL•FO
RenderX

routes can be tailored to meet physicians' and patients' specific needs and demands [115]. This alignment enhances medical decision accuracy, enhances user satisfaction, and ultimately improves patient outcomes [116].

User Control

User control was addressed in 2% (1/43) of the selected papers, emphasizing the importance of instructions that align with real-world workflows, aiding users in recovering from and preventing errors. Interfaces that are compatible with established practices help users rectify mistakes and reduce the recurrence of common errors, enhancing overall usability and efficiency [81]. In this context, an important framework is the unified theory of acceptance and use of technology (UTAUT). UTAUT is used to predict and understand individuals' acceptance and adoption of new technologies, integrating various factors such as performance expectancy, effort expectancy, social influence, and facilitating conditions. By enabling users to recover from mistakes and easily control their interactions, systems can enhance user satisfaction and increase the likelihood of technology adoption, as predicted by the UTAUT model [117].

Data Entry

Data entry was investigated in 2% (1/43) of the selected papers, with a focus on designing systems that minimize potential user errors [82]. Structured data entry management is fundamental in CDSSs, ensuring organized, standardized, and systematic data input. Emerging technologies such as voice recognition technology and natural language processing have facilitated and improved the data entry aspect in eHealth systems. These advancements minimize errors and inconsistencies, ultimately enhancing the reliability and effectiveness of CDSSs in supporting clinical decision-making [118].

Individuality

Individuality was presented in 2% (1/43) of the selected papers. In this area, Klumpp et al [83] discuss the HCI concept as a theory encompassing 4 key issues that shape individuals' interactions with computerized systems. The first factor is the identification of technology, including all technological features applied to the system [119]. The second factor is the assignment, referring to a defined task where the technology used varies depending on the satisfaction and aims of the task [120]. The third factor is context, which can vary depending on the geographic, corporate, or community conditions, meaning that the recommendations should be customized according to the geographic positions and other local conditions [121]. The fourth factor involves human dimensions, such as population characteristics, cognitive abilities, and the attitudes of individuals [122]. In other words, customizability should be based on the intended populations' approaches, attributes, and desires. Concentrating on system individualization and customization can enhance the accuracy of decision-making in CDSSs [83].

Mental Effort

Mental effort was identified as an HCI element in 2% (1/43) of the selected papers. The Rating Scale Mental Effort (RSME), first proposed by Militello et al [84], is designed to accurately measure perceived mental effort during task completion. In other words, it measures the extent to which individuals feel they have exerted mental effort to complete a task. The RSME can be used to gauge how effectively CDSSs aid medical practitioners without overly complicating their decision-making processes. A low RSME score indicates that the CDSS is well integrated into clinical workflows, and CDSS functionality is enhanced without imposing significant additional cognitive complexity on the user [123].

Discussion

Principal Findings

While several studies have explored HCI elements in CDSS environments, each has focused on specific aspects independently. This has left a research gap regarding a comprehensive analysis and categorization of these effective elements. This study attempts to present a scientific framework for CDSS design by identifying HCI elements that influence CDSS functionality across different dimensions. In this respect, after reviewing 43 selected papers in this SLR, we identified and stratified 12 key HCI elements, thereby distinguishing our investigation from the selected studies, which, although they explored the identification and classification of HCI elements, were experimental in nature, with a limited focus on specific, monitored elements.

A common thread among the selected studies is their focus on HCI elements relevant to CDSS applications, which can substantially influence CDSS functionality and performance. Throughout this SLR, the categorization of studies under specific HCI elements depended on the primary focus of each study. While some studies explicitly investigated a single HCI element, others addressed multiple aspects (such as heuristic semantic tags [70]) or offered broad recommendations. In such cases, the main HCI concern emphasized in the study was the selection criterion. This approach ensured that each study was categorized based on its core usability focus, maintaining consistency and relevance in the analysis.

Throughout this SLR, the categorization of studies under specific HCI elements depended on the primary focus of each study. While some studies explicitly investigated a single HCI element, others addressed multiple aspects (such as heuristic semantic tags [70]) or offered broad recommendations. In such cases, the main HCI concern emphasized in the study was the selection criterion. This approach ensured that each study was categorized based on its core usability focus, maintaining consistency and relevance in the analysis.

The interface is considered a distinct HCI element because it serves as the primary medium for user interaction [124]. Although it encompasses visibility, data presentation, and ease of use, its role transcends these components [125,126]. Indeed, the interface functions as the platform that enables the delivery and experience of these elements, making it a distinct and critical component of the HCI context [127]. This distinction is crucial for CDSS environments, where the interface must accommodate complex clinical workflows and user needs to ensure usability and patient safety. Hence, despite its broad scope, the interface has been examined as an independent HCI element in some

XSL•FO
RenderX

studies, allowing for a more focused and in-depth evaluation of its inherent contributions to UX and decision-making in CDSS environments [128]. Thus, we considered it a separate HCI element in this study.

The relationship between HCI and UI and UX design further underscores the importance of the interface as an independent element. While UI focuses on the visual and interactive components of a system (eg, buttons, menus, and layouts) and UX encompasses the user-centered experience (including usability, accessibility, and emotional satisfaction), the interface concept (within HCI) is the physical manifestation of these principles [129,130]. As Paneru et al [131] explain, the interface is a visual and interactive layer that enables users to interact with digital products. It is the medium through which all other HCI elements are experienced, making it a foundational component of HCI. However, the interface also extends beyond traditional UI and UX design by incorporating considerations such as workflow alignment, cognitive load reduction, and task efficiency, which are central to HCI. By contrast, Motlagh and Safaei [132] emphasized that, rather than relying solely on UI and UX considerations, HCI evaluation in health care systems should prioritize error prevention, cognitive effort, and information recall. This means that HCI elements promote interfaces that support accurate decision-making and clinical workflow integration.

According to ISO 9241-11, usability measures how effectively, efficiently, and satisfactorily users can achieve their goals with a system in a particular context. This highlights the importance of context and task dependency in evaluating usability because the relevance and effectiveness of HCI elements vary across different environments and user needs [133]. The HCI elements identified in this study align with these usability dimensions but manifest differently depending on task requirements and system context. In frequently used systems such as EHRs in busy hospitals, ease of use, simplification, and task efficiency are paramount HCI elements. Clinicians using these systems need interfaces that minimize cognitive load and streamline workflows because even small inefficiencies can accumulate, leading to frustration and errors. In this context, alerts must be designed to be highly visible but nonintrusive so that they support rather than disrupt workflow [48]. In these circumstances, data entry mechanisms should be optimized for speed and accuracy because frequent use requires efficiency. Conversely, in an infrequently used system (such as a specialized diagnostic tool used only occasionally), HCI elements such as explainability become more critical. Users of such systems will need more guidance and contextual support to navigate the interface effectively. Regarding user control and flexibility, although these elements align with the satisfaction component of the ISO 9241-11 standard, their impact is also context dependent. Experienced physicians may prefer more flexible customization to streamline interactions, while novice users benefit from structured interfaces with clear navigation paths and constraints to prevent errors [80,81]. As shown in Table 1, the identified HCI elements were categorized based on their primary usability contributions; nevertheless, their actual significance in improving CDSS usability depends on the task, frequency of use, and clinical environment.

Two closely related but distinct HCI elements—ease of use and simplification (simplification)—play unique roles in enhancing CDSS usability. Ease of use refers to the system's ability to provide an intuitive and seamless interaction experience, minimizing cognitive and physical effort for users. It encompasses aspects such as jargon avoidance, consistency of operations, and efficient task execution, ensuring that clinicians can navigate the system with minimal frustration [56,58-60]. By contrast, simplification focuses on reducing unnecessary complexity within the system, such as streamlining workflows, eliminating redundant steps, and minimizing information overload [74,75]. While ease of use improves UX, simplification optimizes system design by removing obstacles and extraneous procedures that could hinder task completion. Moreover, ease of use is closely tied to user satisfaction and IU. An easy-to-use system facilitates navigation, reduces cognitive efforts, and leads to a more positive UX, ultimately increasing IU and user satisfaction.

This study recognized another prevalent and critical challenge faced by CDSS designers: balancing the paradox between system simplification and maintaining its usefulness in solving complex problems, even if doing so leads to increased complexity. This balance is essential for developing systems with high usability and appeal, ensuring that users can effectively interact with the system without being overwhelmed by its complexity [75]. To reach this equilibrium, designers must adopt a user-centered approach that prioritizes the needs and capabilities of health care professionals. This involves continuous user testing and feedback loops to refine the interface and functionality [59]. In addition, incorporating adaptive interfaces can reduce complexity based on the user's expertise and the specific task to help manage this paradox, resulting in augmented overall satisfaction and higher adoption rates [65].

This study attempted to highlight the HCI elements dedicated to CDSSs, specifically those that address user requirements across various dimensions such as informative notifications [52], aesthetics [61,71], contextual facets [79], unification [70], and psychological aspects [84]. To illuminate all HCI aspects within CDSSs, this study also identified elements that represent challenges in this domain requiring mitigation, such as existing HCI barriers [78], usability flaws [43], user fatigue [7], and the phenomenon of alert override [49]. Although considering the identified HCI aspects is crucial, the final goal is to improve user interaction with the system, enhance ease of use, and anticipate and fulfill user expectations. This approach fosters user satisfaction and promotes a positive UX [74]. This focus on users as major "players" in the design process is now an established principle. Hence, UCD methods are fundamental to the development of CDSSs with high levels of user interaction [134]. Depending on the system's complexity and ease of use, both user interaction with the system and data accuracy can vary, consequently affecting the accuracy of medical decision-making [58].

In summary, this study provides comprehensive insights into the essential HCI elements applicable in the CDSS environment, answering the RQs. It contributes to the field by systematically categorizing these elements and developing a structured

XSL•FO
RenderX

framework aimed at improving usability and medical decision-making outcomes.

Conclusions

Of the 923 papers found in 3 databases (PubMed, Scopus, and Web of Science), 43 (4.7%) met the predefined criteria with regard to addressing HCI elements applicable in CDSS environments. This SLR indicates that attention to HCI elements tailored to CDSSs has increased since 2015. In this study, we identified 12 HCI elements. While elements such as "ease of use," "interface," and "alerts" have consistently garnered attention from experts, others such as "user satisfaction," "visibility," and "explainability" have recently received increased focus. This SLR has answered the RQs regarding the HCI elements that can affect CDSS functionality and described how these elements impact CDSS performance. Moreover, the study categorized the identified aspects according to their context and scientific perspectives.

This study identified a gap in current research: the lack of a comprehensive consideration of HCI elements within CDSS design. By leveraging the findings from this investigation, a set of applicable HCI elements for CDSS environments can be established. Consequently, these findings have the potential to enhance the synergistic effect between medical information systems and CDSSs, ultimately leading to improved overall performance.

This research provides the groundwork for generating more structured and reliable datasets to support medical decision-making. Systems designed with these HCI elements are expected to reduce missing data, minimize data redundancies, and improve data clarity. Accordingly, the accuracy and reliability of medical decisions derived from these systems will be noticeably enhanced. Future studies can leverage this framework to develop and examine CDSS systems that improve user interaction and ensure higher data integrity and precision in medical decision-making. This structured approach will advance medical informatics, leading to more efficient, accurate, and user-friendly CDSS solutions.

Acknowledgments

The authors appreciate all researchers' efforts regarding human-computer interaction concerns, especially within clinical decision support systems. The Spanish Ministry of Science and Innovation partially funded this research through the Analítica Visual y machine learning para toma de decisiones en ecosistemas de Salud (AVisSA; "Visual Analytics and Machine Learning for Decision-Making in Health Ecosystems") project (PID2020-118345RB-I00).

Authors' Contributions

AA and FJG-P contributed to conceptualization, methodology, and validation. AA contributed to formal analysis, investigation fulfillment, visualization, and writing—original draft preparation. FJG-P contributed to project administration and writing—review and editing. Both authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.
[PDF File (Adobe PDF File), 89 KB - humanfactors_v12i1e69333_app1.pdf]

Multimedia Appendix 2
The raw data from this study, including extracted papers from Scopus (506/923, 54.8%), Web of Science (188/923, 20.4%), and PubMed (229/923, 24.8%).
[XLSX File (Microsoft Excel File), 146 KB - humanfactors_v12i1e69333_app2.xlsx]

Multimedia Appendix 3
The final set of 43 papers selected for analysis.
[XLSX File (Microsoft Excel File), 26 KB - humanfactors_v12i1e69333_app3.xlsx]

References

1. Osheroff J, Teich J, Levick D, Saldana L, Velasco F, Sittig D, et al. Improving Outcomes with Clinical Decision Support: An Implementer's Guide. 2nd Edition. New York, NY: HIMSS Publishing; 2012.
2. Khajouei R, Peek N, Wierenga PC, Kersten MJ, Jaspers MW. Effect of predefined order sets and usability problems on efficiency of computerized medication ordering. Int J Med Inform 2010 Oct;79(10):690-698. [doi: 10.1016/j.ijmedinf.2010.08.001] [Medline: 20833104]

XSL•FO
RenderX

3. Roshanov PS, Fernandes N, Wilczynski JM, Hemens BJ, You JJ, Handler SM, et al. Features of effective computerised clinical decision support systems: meta-regression of 162 randomised trials. BMJ 2013 Feb 14;346(feb14 1):f657 [FREE Full text] [doi: 10.1136/bmj.f657] [Medline: 23412440]
4. Jaspers MW, Smeulers M, Vermeulen H, Peute LW. Effects of clinical decision-support systems on practitioner performance and patient outcomes: a synthesis of high-quality systematic review findings. J Am Med Inform Assoc 2011 May 01;18(3):327-334 [FREE Full text] [doi: 10.1136/amiajnl-2011-000094] [Medline: 21422100]
5. O'Donnell HC, Kaushal R, Barrón Y, Callahan MA, Adelman RD, Siegler EL. Physicians' attitudes towards copy and pasting in electronic note writing. J Gen Intern Med 2009 Jan;24(1):63-68 [FREE Full text] [doi: 10.1007/s11606-008-0843-2] [Medline: 18998191]
6. Antona M, Margetis G, Ntoa S, Degen H. Special issue on AI in HCI. Int J Hum Comput Interact 2023 Feb 21;39(9):1723-1726. [doi: 10.1080/10447318.2023.2177421]
7. Bolgova KV, Kovalchuk SV, Balakhontceva M, Zvartau N, Metsker O. Human computer interaction during clinical decision support with electronic health records improvement. In: Information Resources Management Association , editor. Research Anthology on Decision Support Systems and Decision Management in Healthcare, Business, and Engineering. New York, NY: IGI Global; 2021:1316-1330.
8. Ghaoui C. Encyclopedia of Human Computer Interaction. Cambridge, MA: Taxmann Publications; 2006.
9. Sønsterud H, Kirmess M, Howells K, Ward D, Feragen KB, Halvorsen MS. The working alliance in stuttering treatment: a neglected variable? Int J Lang Commun Disord 2019 Jul;54(4):606-619 [FREE Full text] [doi: 10.1111/1460-6984.12465] [Medline: 30866151]
10. Patel VL, Kannampallil TG, Kaufman DR. Cognitive Informatics for Biomedicine: Human Computer Interaction in Healthcare. Cham, Switzerland: Springer; 2015.
11. Khairat S, Marc D, Crosby W, Al Sanousi A. Reasons for physicians not adopting clinical decision support systems: critical analysis. JMIR Med Inform 2018 Apr 18;6(2):e24 [FREE Full text] [doi: 10.2196/medinform.8912] [Medline: 29669706]
12. Issa T, Isaias P. Usability and human computer interaction (HCI). In: Issa T, Isaias P, editors. Sustainable Design: HCI, Usability and Environmental Concerns. Cham, Switzerland: Springer; 2015:19-36.
13. Wassouf M. Physicians' expectations of future clinical decision support systems: exploring the expected user experience of physicians in interaction with future decision support systems. Halmstad University. URL: https://www.diva-portal.org/smash/get/diva2:1683751/FULLTEXT01.pdf [accessed 2024-04-29]
14. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. NPJ Digit Med 2020;3:17 [FREE Full text] [doi: 10.1038/s41746-020-0221-y] [Medline: 32047862]
15. Crowley J, Heavin C, Keenan P, Power D. CDSS and DSS: shared roots and divergent paths. J Decis Syst 2020 Aug 31;29(sup1):71-78. [doi: 10.1080/12460125.2020.1811446]
16. Zahabi M, Kaber DB, Swangnetr M. Usability and safety in electronic medical records interface design: a review of recent literature and guideline formulation. Hum Factors 2015 Aug;57(5):805-834. [doi: 10.1177/0018720815576827] [Medline: 25850118]
17. Blandford A. HCI for health and wellbeing: challenges and opportunities. Int J Hum Comput Interact 2019 Nov;131:41-51 [FREE Full text] [doi: 10.1016/j.ijhcs.2019.06.007]
18. García-Peñalvo F. Desarrollo de estados de la cuestión robustos: revisiones sistemáticas de literatura. Educ Knowl Soc 2022;14(7):e28600 [FREE Full text] [doi: 10.3390/jpm14070713] [Medline: 39063967]
19. Kitchenham B. Procedures for performing systematic reviews. Keele University. URL: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=29890a936639862f45cb9a987dd599dce9759bf5 [accessed 2024-04-29]
20. Kitchenham B, Charters S. Guidelines for performing systematic literature reviews in software engineering. Keele University. 2007. URL: https://legacyfileshare.elsevier.com/promis_misc/525444systematicreviewsguide.pdf [accessed 2024-04-29]
21. Kitchenham BA, Budgen D, Pearl Brereton O. Using mapping studies as the basis for further research – a participant-observer case study. Inf Soft Technol 2011 Jun;53(6):638-651 [FREE Full text] [doi: 10.1016/j.infsof.2010.12.011]
22. Pati D, Lorusso LN. How to write a systematic review of the literature. HERD 2018 Jan;11(1):15-30. [doi: 10.1177/1937586717747384] [Medline: 29283007]
23. Petticrew M, Roberts H. Systematic Reviews in the Social Sciences: A Practical Guide. Hoboken, NJ: John Wiley & Sons; 2006.
24. ISO 9241-210:2019 ergonomics of human-system interaction - part 210: human-centred design for interactive systems. International Organization for Standardization. URL: https://www.iso.org/standard/77520.html [accessed 2024-04-29]
25. Chou D. Health IT and patient safety: building safer systems for better care. JAMA 2012 Dec 05;308(21):2282. [doi: 10.1001/jama.308.21.2282-a]
26. Middleton B, Bloomrosen M, Dente MA, Hashmat B, Koppel R, Overhage JM, American Medical Informatics Association. Enhancing patient safety and quality of care by improving the usability of electronic health record systems: recommendations from AMIA. J Am Med Inform Assoc 2013 Jun;20(e1):e2-e8 [FREE Full text] [doi: 10.1136/amiajnl-2012-001458] [Medline: 23355463]

XSL•FO
RenderX

27. Howe JL, Adams KT, Hettinger AZ, Ratwani RM. Electronic health record usability issues and potential contribution to patient harm. JAMA 2018 Mar 27;319(12):1276-1278 [FREE Full text] [doi: 10.1001/jama.2018.1171] [Medline: 29584833]
28. Li X, Xu Y. Role of human-computer interaction healthcare system in the teaching of physiology and medicine. Comput Intell Neurosci 2022;2022:5849736 [FREE Full text] [doi: 10.1155/2022/5849736] [Medline: 35463243]
29. Rajeswari KS, Anantharaman RN. Role of human-computer interaction factors as moderators of occupational stress and work exhaustion. Int J Hum Comput Interact 2005 Sep;19(1):137-154. [doi: 10.1207/s15327590ijhc1901_9]
30. Hegde V. Role of human factors / usability engineering in medical device design. In: Proceedings of the 2013 Annual Conference on Reliability and Maintainability Symposium. 2013 Presented at: RAMS '13; January 28-31, 2013; Orlando, FL p. 1-5 URL: https://ieeexplore.ieee.org/document/6517650 [doi: 10.1109/rams.2013.6517650]
31. Adeola OA, Ayorinde O, Jerry S. Human factors and human-computer interaction. ResearchGate. URL: https://www.researchgate.net/publication/380728031_HUMAN_FACTORS_AND_HUMAN-COMPUTER_INTERACTION [accessed 2024-04-29]
32. Li Y, Oladimeji P, Monroy C, Cauchi A, Thimbleby H, Furniss D, et al. Design of interactive medical devices: feedback and its improvement. In: Proceedings of the 2011 IEEE International Symposium on IT in Medicine and Education. 2011 Presented at: ITiME '11; December 9-11, 2011; Guangzhou, China p. 204-208 URL: https://ieeexplore.ieee.org/document/6132022 [doi: 10.1109/itime.2011.6132022]
33. Eslami M, Firoozabadi M, Homayounvala E. User preferences for adaptive user interfaces in health information systems. Univ Access Inf Soc 2017 Aug 5;17(4):875-883. [doi: 10.1007/S10209-017-0569-1]
34. Kronenfeld MR, Bay RC, Coombs W. Survey of user preferences from a comparative trial of UpToDate and ClinicalKey. J Med Libr Assoc 2013 Apr;101(2):151-154 [FREE Full text] [doi: 10.3163/1536-5050.101.2.011] [Medline: 23646031]
35. Isaac T, Zheng J, Jha A. Use of UpToDate and outcomes in US hospitals. J Hosp Med 2012 Feb;7(2):85-90. [doi: 10.1002/jhm.944] [Medline: 22095750]
36. Benke K, Benke G. Artificial intelligence and big data in public health. Int J Environ Res Public Health 2018 Dec 10;15(12):25 [FREE Full text] [doi: 10.3390/ijerph15122796] [Medline: 30544648]
37. Jones C, Thornton J, Wyatt JC. Artificial intelligence and clinical decision support: clinicians' perspectives on trust, trustworthiness, and liability. Med Law Rev 2023 Nov 27;31(4):501-520 [FREE Full text] [doi: 10.1093/medlaw/fwad013] [Medline: 37218368]
38. Smith H, Fotheringham K. Artificial intelligence in clinical decision-making: rethinking liability. Med Law Int 2020 Aug 26;20(2):131-154. [doi: 10.1177/0968533220945766]
39. Kitchenham B, Brereton P. A systematic review of systematic review process research in software engineering. Inf Soft Technol 2013 Dec;55(12):2049-2075 [FREE Full text] [doi: 10.1016/j.infsof.2013.07.010]
40. krippendorff K. Computing Krippendorff's Alpha-reliability. Annenberg School for Communication. 2011. URL: https://www.asc.upenn.edu/sites/default/files/2021-03/Computing%20Krippendorff%27s%20Alpha-Reliability.pdf [accessed 2024-04-29]
41. Mengist W, Soromessa T, Legese G. Method for conducting systematic literature review and meta-analysis for environmental science research. MethodsX 2020;7:100777 [FREE Full text] [doi: 10.1016/j.mex.2019.100777] [Medline: 31993339]
42. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
43. Marcilly R, Ammenwerth E, Vasseur F, Roehrer E, Beuscart-Zéphir MC. Usability flaws of medication-related alerting functions: a systematic qualitative review. J Biomed Inform 2015 Jun;55:260-271. [doi: 10.1016/j.jbi.2015.03.006] [Medline: 25817918]
44. Richardson JE, Ash JS. A clinical decision support needs assessment of community-based physicians. J Am Med Inform Assoc 2011 Dec;18 Suppl 1(Suppl 1):i28-i35 [FREE Full text] [doi: 10.1136/amiajnl-2011-000119] [Medline: 21890874]
45. Cho I, Lee JH, Han H, Phansalkar S, Bates D. Evaluation of a Korean version of a tool for assessing the incorporation of human factors into a medication-related decision support system: the I-MeDeSA. Appl Clin Inform 2017 Dec 21;05(02):571-588. [doi: 10.4338/aci-2014-01-ra-0005]
46. Press A, McCullagh L, Khan S, Schachter A, Pardo S, McGinn T. Usability testing of a complex clinical decision support tool in the emergency department: lessons learned. JMIR Hum Factors 2015 Sep 10;2(2):e14 [FREE Full text] [doi: 10.2196/humanfactors.4537] [Medline: 27025540]
47. Gong Y, Kang H. Usability and clinical decision support. In: Berner ES, editor. Clinical Decision Support Systems: Theory and Practice. Cham, Switzerland: Springer; 2016:69-86.
48. Nishimura AA, Shirts BH, Salama J, Smith JW, Devine B, Tarczy-Hornoch P. Physician perspectives of CYP2C19 and clopidogrel drug-gene interaction active clinical decision support alerts. Int J Med Inform 2016 Feb;86:117-125 [FREE Full text] [doi: 10.1016/j.ijmedinf.2015.11.004] [Medline: 26642939]
49. Zachariah M, Phansalkar S, Seidling HM, Neri PM, Cresswell KM, Duke J, et al. Development and preliminary evidence for the validity of an instrument assessing implementation of human-factors principles in medication-related decision-support systems--I-MeDeSA. J Am Med Inform Assoc 2011 Dec 01;18 Suppl 1(Suppl 1):i62-i72 [FREE Full text] [doi: 10.1136/amiajnl-2011-000362] [Medline: 21946241]

XSL•FO
RenderX

50. Horsky J, Schiff GD, Johnston D, Mercincavage L, Bell D, Middleton B. Interface design principles for usable decision support: a targeted review of best practices for clinical prescribing interventions. J Biomed Inform 2012 Dec;45(6):1202-1216 [FREE Full text] [doi: 10.1016/j.jbi.2012.09.002] [Medline: 22995208]
51. Horsky J, Phansalkar S, Desai A, Bell D, Middleton B. Design of decision support interventions for medication prescribing. Int J Med Inform 2013 Jun;82(6):492-503. [doi: 10.1016/j.ijmedinf.2013.02.003] [Medline: 23490305]
52. Stehlik P, Bahmanpour A, Ahmet SY, Darziņš P, Marriott J. Fundamental elements identified for success of disease state management clinical decision support systems. Electron J Health Inform 2015;9:e6 [FREE Full text]
53. Garabedian PM, Gannon MP, Aaron S, Wu E, Burns Z, Samal L. Human-centered design of clinical decision support for management of hypertension with chronic kidney disease. BMC Med Inform Decis Mak 2022 Aug 13;22(1):217 [FREE Full text] [doi: 10.1186/s12911-022-01962-y] [Medline: 35964083]
54. Koskela T, Sandström S, Mäkinen J, Liira H. User perspectives on an electronic decision-support tool performing comprehensive medication reviews - a focus group study with physicians and nurses. BMC Med Inform Decis Mak 2016 Jan 22;16(1):6 [FREE Full text] [doi: 10.1186/s12911-016-0245-z] [Medline: 26801630]
55. Buenestado D, Elorz J, Pérez-Yarza EG, Iruetaguena A, Segundo U, Barrena R, et al. Evaluating acceptance and user experience of a guideline-based clinical decision support system execution platform. J Med Syst 2013 Apr 3;37(2):9910. [doi: 10.1007/s10916-012-9910-7] [Medline: 23377779]
56. Richardson S, Mishuris R, O'Connell A, Feldstein D, Hess R, Smith P, et al. "Think aloud" and "Near live" usability testing of two complex clinical decision support tools. Int J Med Inform 2017 Oct;106:1-8 [FREE Full text] [doi: 10.1016/j.ijmedinf.2017.06.003] [Medline: 28870378]
57. Richardson S, Feldstein D, McGinn T, Park LS, Khan S, Hess R, et al. Live usability testing of two complex clinical decision support tools: observational study. JMIR Hum Factors 2019 Apr 15;6(2):e12471 [FREE Full text] [doi: 10.2196/12471] [Medline: 30985283]
58. Lloyd S, Long K, Probst Y, Di Donato J, Oshni Alvandi A, Roach J, et al. Medical and nursing clinician perspectives on the usability of the hospital electronic medical record: a qualitative analysis. Health Inf Manag 2024 Sep;53(3):189-197 [FREE Full text] [doi: 10.1177/18333583231154624] [Medline: 36866778]
59. Shields C, Cunningham SG, Wake DJ, Fioratou E, Brodie D, Philip S, et al. User-centered design of a novel risk prediction behavior change tool augmented with an artificial intelligence engine (MyDiabetesIQ): a sociotechnical systems approach. JMIR Hum Factors 2022 Feb 08;9(1):e29973 [FREE Full text] [doi: 10.2196/29973] [Medline: 35133280]
60. Kushniruk AW, Patel VL. Cognitive and usability engineering methods for the evaluation of clinical information systems. J Biomed Inform 2004 Feb;37(1):56-76 [FREE Full text] [doi: 10.1016/j.jbi.2004.01.003] [Medline: 15016386]
61. Jones W, Drake C, Mack D, Reeder B, Trautner B, Wald H. Developing mobile clinical decision support for nursing home staff assessment of urinary tract infection using goal-directed design. Appl Clin Inform 2017 Dec 21;08(02):632-650. [doi: 10.4338/aci-2016-12-ra-0209]
62. Hoonakker P, Khunlertkit A, Tattersall M, Keevil J, Smith PD. Computer decision support tools in primary care. WORK 2012 Jan;41(S1):4474-4478. [doi: 10.3233/wor-2012-0747-4474]
63. Kulchak Rahm A, Walton NA, Feldman LK, Jenkins C, Jenkins T, Person TN, et al. User testing of a diagnostic decision support system with machine-assisted chart review to facilitate clinical genomic diagnosis. BMJ Health Care Inform 2021 May 07;28(1):e100331 [FREE Full text] [doi: 10.1136/bmjhci-2021-100331] [Medline: 33962988]
64. Vedanthan R, Blank E, Tuikong N, Kamano J, Misoi L, Tulienge D, et al. Usability and feasibility of a tablet-based decision-support and integrated record-keeping (DESIRE) tool in the nurse management of hypertension in rural western Kenya. Int J Med Inform 2015 Mar;84(3):207-219 [FREE Full text] [doi: 10.1016/j.ijmedinf.2014.12.005] [Medline: 25612791]
65. Porat T, Delaney B, Kostopoulou O. The impact of a diagnostic decision support system on the consultation: perceptions of GPs and patients. BMC Med Inform Decis Mak 2017 Jun 02;17(1):79 [FREE Full text] [doi: 10.1186/s12911-017-0477-6] [Medline: 28576145]
66. Rundo L, Pirrone R, Vitabile S, Sala E, Gambino O. Recent advances of HCI in decision-making tasks for optimized clinical workflows and precision medicine. J Biomed Inform 2020 Aug;108:103479 [FREE Full text] [doi: 10.1016/j.jbi.2020.103479] [Medline: 32561444]
67. Miller K, Mosby D, Capan M, Kowalski R, Ratwani R, Noaiseh Y, et al. Interface, information, interaction: a narrative review of design and functional requirements for clinical decision support. J Am Med Inform Assoc 2018 May 01;25(5):585-592 [FREE Full text] [doi: 10.1093/jamia/ocx118] [Medline: 29126196]
68. Akhloufi H, Verhaegh SJ, Jaspers MW, Melles DC, van der Sijs H, Verbon A. A usability study to improve a clinical decision support system for the prescription of antibiotic drugs. PLoS One 2019;14(9):e0223073 [FREE Full text] [doi: 10.1371/journal.pone.0223073] [Medline: 31553785]
69. Nair KM, Malaeekeh R, Schabort I, Taenzer P, Radhakrishnan A, Guenter D. A clinical decision support system for chronic pain management in primary care: usability testing and its relevance. J Innov Health Inform 2015 Aug 13;22(3):329-332 [FREE Full text] [doi: 10.14236/jhi.v22i3.149] [Medline: 26577423]
70. Zhang J, Johnson TR, Patel VL, Paige DL, Kubose T. Using usability heuristics to evaluate patient safety of medical devices. J Biomed Inform 2003;36(1-2):23-30 [FREE Full text] [doi: 10.1016/s1532-0464(03)00060-1] [Medline: 14552844]

XSL•FO
RenderX

71. Kruse CH, Bekker W, Bruce JL, Clarke DL. Striking a balance between usability and quality control in electronic health records. S Afr J Surg 2022 Sep;60(3):171-175. [doi: 10.17159/2078-5151/sajs3767]
72. Cho H, Keenan G, Madandola OO, Dos Santos FC, Macieira TG, Bjarnadottir RI, et al. Assessing the usability of a clinical decision support system: heuristic evaluation. JMIR Hum Factors 2022 May 10;9(2):e31758 [FREE Full text] [doi: 10.2196/31758] [Medline: 35536613]
73. Erturkmen GB, Yuksel M, Sarigul B, Arvanitis TN, Lindman P, Chen R, et al. A collaborative platform for management of chronic diseases via guideline-driven individualized care plans. Comput Struct Biotechnol J 2019;17:869-885 [FREE Full text] [doi: 10.1016/j.csbj.2019.06.003] [Medline: 31333814]
74. Tabla S, Calafiore M, Legrand B, Descamps A, Andre C, Rochoy M, et al. Artificial intelligence and clinical decision support systems or automated interpreters: what characteristics are expected by French general practitioners? Stud Health Technol Inform 2022:887-891. [doi: 10.3233/shti220207]
75. Pantazi SV, Kushniruk A, Moehr JR. The usability axiom of medical information systems. Int J Med Inform 2006 Dec;75(12):829-839. [doi: 10.1016/j.ijmedinf.2006.05.039] [Medline: 16870500]
76. Clausen C, Leventhal B, Nytrø Ø, Koposov R, Røst TB, Westbye OS, et al. Usability of the IDDEAS prototype in child and adolescent mental health services: a qualitative study for clinical decision support system development. Front Psychiatry 2023 Feb 23;14:1033724 [FREE Full text] [doi: 10.3389/fpsyt.2023.1033724] [Medline: 36911136]
77. Ghorayeb A, Darbyshire JL, Wronikowska MW, Watkinson PJ. Design and validation of a new Healthcare Systems Usability Scale (HSUS) for clinical decision support systems: a mixed-methods approach. BMJ Open 2023 Jan 30;13(1):e065323 [FREE Full text] [doi: 10.1136/bmjopen-2022-065323] [Medline: 36717136]
78. Friedberg MW, Chen PG, Van Busum KR, Aunon F, Pham C, Caloyeras J, et al. Factors affecting physician professional satisfaction and their implications for patient care, health systems, and health policy. Rand Health Q 2014;3(4):1 [FREE Full text] [Medline: 28083306]
79. Singh A, Schooley B, Floyd SB, Pill SG, Brooks JM. Patient preferences as human factors for health data recommender systems and shared decision making in orthopaedic practice. Front Digit Health 2023 Jun 20;5:1137066 [FREE Full text] [doi: 10.3389/fdgth.2023.1137066] [Medline: 37408539]
80. Gutenstein M, Pickering JW, Than M. Development of a digital clinical pathway for emergency medicine: lessons from usability testing and implementation failure. Health Informatics J 2019 Dec 15;25(4):1563-1571 [FREE Full text] [doi: 10.1177/1460458218779099] [Medline: 29905094]
81. Long J, Yuan MJ, Poonawala R. An observational study to evaluate the usability and intent to adopt an artificial intelligence-powered medication reconciliation tool. Interact J Med Res 2016 May 16;5(2):e14 [FREE Full text] [doi: 10.2196/ijmr.5462] [Medline: 27185210]
82. Bologva EV, Prokusheva D, Krikunov AV, Zvartau N, Kovalchuk SV. Human-computer interaction in electronic medical records: from the perspectives of physicians and data scientists. Procedia Comput Sci 2016;100:915-920 [FREE Full text] [doi: 10.1016/j.procs.2016.09.248]
83. Klumpp M, Hanelt A, Greve M, Kolbe LM, Tofangchi S, Böhrnsen F, et al. Accelerating the front end of medicine: three digital use cases and HCI implications. Healthcare (Basel) 2022 Oct 30;10(11):2176 [FREE Full text] [doi: 10.3390/healthcare10112176] [Medline: 36360517]
84. Militello LG, Diiulio JB, Borders MR, Sushereba CE, Saleem JJ, Haverkamp D, et al. Evaluating a modular decision support application for colorectal cancer screening. Appl Clin Inform 2017 Dec 20;26(01):162-179. [doi: 10.4338/aci-2016-09-ra-0152]
85. Smith A, Bannon L, Gulliksen J. Localising HCI practice for local needs. In: Proceedings of the 2010 HCI Conference on Interaction Design & International Development. 2010 Presented at: IHCI '10; March 20-24, 2010; New Delhi, India p. 123 URL: https://www.scienceopen.com/hosted-document?doi=10.14236/ewic/IHCI2010.15 [doi: 10.14236/ewic/ihci2010.15]
86. Middleton B, Sittig DF, Wright A. Clinical decision support: a 25 year retrospective and a 25 year vision. Yearb Med Inform 2016 Aug 02;Suppl 1(Suppl 1):S103-S116 [FREE Full text] [doi: 10.15265/IYS-2016-s034] [Medline: 27488402]
87. Chen Z, Liang N, Zhang H, Li H, Yang Y, Zong X, et al. Harnessing the power of clinical decision support systems: challenges and opportunities. Open Heart 2023 Nov 28;10(2):102 [FREE Full text] [doi: 10.1136/openhrt-2023-002432] [Medline: 38016787]
88. Novey M, Adali T, Roy A. A complex generalized gaussian distribution— characterization, generation, and estimation. IEEE Trans Signal Process 2010 Mar;58(3):1427-1433. [doi: 10.1109/tsp.2009.2036049]
89. Bevan N, Carter J, Harker S. ISO 9241-11 revised: what have we learnt about usability since 1998? In: Proceedings of the 17th International Conference on Human-Computer Interaction: Design and Evaluation. 2015 Presented at: HCI '15; August 2-7, 2015; Los Angeles, CA p. 143-151 URL: https://link.springer.com/chapter/10.1007/978-3-319-20901-2_13 [doi: 10.1007/978-3-319-20901-2_13]
90. Lee J, Han H, Ock M, Lee S, Lee S, Jo MW. Impact of a clinical decision support system for high-alert medications on the prevention of prescription errors. Int J Med Inform 2014 Dec;83(12):929-940 [FREE Full text] [doi: 10.1016/j.ijmedinf.2014.08.006] [Medline: 25256067]

XSL•FO
RenderX

91. Bauer NS, Carroll AE, Saha C, Downs SM. Experience with decision support system and comfort with topic predict clinicians' responses to alerts and reminders. J Am Med Inform Assoc 2016 Apr;23(e1):e125-e130 [FREE Full text] [doi: 10.1093/jamia/ocv148] [Medline: 26567326]
92. Wright A, Ash JS, Aaron S, Ai A, Hickman TT, Wiesen JF, et al. Best practices for preventing malfunctions in rule-based clinical decision support alerts and reminders: results of a Delphi study. Int J Med Inform 2018 Oct;118:78-85 [FREE Full text] [doi: 10.1016/j.ijmedinf.2018.08.001] [Medline: 30153926]
93. Murphy DR, Giardina TD, Satterly T, Sittig DF, Singh H. An exploration of barriers, facilitators, and suggestions for improving electronic health record inbox-related usability: a qualitative analysis. JAMA Netw Open 2019 Oct 02;2(10):e1912638 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.12638] [Medline: 31584683]
94. Ash JS, Sittig DF, Campbell EM, Guappone KP, Dykstra RH. Some unintended consequences of clinical decision support systems. AMIA Annu Symp Proc 2007 Oct 11;2007:26-30 [FREE Full text] [Medline: 18693791]
95. Backman R, Bayliss S, Moore D, Litchfield I. Clinical reminder alert fatigue in healthcare: a systematic literature review protocol using qualitative evidence. Syst Rev 2017 Dec 13;6(1):255 [FREE Full text] [doi: 10.1186/s13643-017-0627-z] [Medline: 29237488]
96. Cimino JJ, Patel VL, Kushniruk AW. Studying the human-computer-terminology interface. J Am Med Inform Assoc 2001;8(2):163-173 [FREE Full text] [doi: 10.1136/jamia.2001.0080163] [Medline: 11230384]
97. Chaffee BW, Zimmerman CR. Developing and implementing clinical decision support for use in a computerized prescriber-order-entry system. Am J Health Syst Pharm 2010 Mar 01;67(5):391-400. [doi: 10.2146/ajhp090153] [Medline: 20172991]
98. Feldstein A, Simon SR, Schneider J, Krall M, Laferriere D, Smith DH, et al. How to design computerized alerts to safe prescribing practices. Jt Comm J Qual Saf 2004 Nov;30(11):602-613. [doi: 10.1016/s1549-3741(04)30071-7] [Medline: 15565759]
99. Kevin L. The Science of Word Recognition; or how I learned to stop worrying and love the bouma. ResearchGate. 2005. URL: https://www.researchgate.net/publication/237657975_The_Science_of_Word_Recognition_or_how_I_learned_to_stop_worrying_and_love_the_bouma [accessed 2024-04-29]
100. Lamy JB, Venot A, Bar-Hen A, Ouvrard P, Duclos C. Design of a graphical and interactive interface for facilitating access to drug contraindications, cautions for use, interactions and adverse effects. BMC Med Inform Decis Mak 2008 Jun 02;8(1):21 [FREE Full text] [doi: 10.1186/1472-6947-8-21] [Medline: 18518945]
101. Abramson EL, Patel V, Malhotra S, Pfoh ER, Nena Osorio S, Cheriff A, et al. Physician experiences transitioning between an older versus newer electronic health record for electronic prescribing. Int J Med Inform 2012 Aug;81(8):539-548. [doi: 10.1016/j.ijmedinf.2012.02.010] [Medline: 22465355]
102. Hafsa S, Majid MA, Tawafak RM. Learnability factors of AR usage performance: validating through survey. In: Proceedings of the 2021 International Conference on Software Engineering & Computer Systems and 4th International Conference on Computational Science and Information Management. 2021 Presented at: ICSECS-ICOCSIM '21; August 24-26, 2021; Pekan, Malaysia p. 371-376 URL: https://ieeexplore.ieee.org/document/9537055 [doi: 10.1109/icsecs52883.2021.00074]
103. Kostopoulou O, Porat T, Corrigan D, Mahmoud S, Delaney BC. Diagnostic accuracy of GPs when using an early-intervention decision support system: a high-fidelity simulation. Br J Gen Pract 2017 Jan 30;67(656):e201-e208. [doi: 10.3399/bjgp16x688417]
104. Kurniawan E. Medical record to simplify the hospital decision making system. In: Proceedings of the 2nd Borobudur International Symposium on Science and Technology. 2020 Presented at: BIS-STE '00; November 18, 2020; Virtual Event p. 368-371 URL: https://www.atlantis-press.com/proceedings/bis-ste-20/125959961 [doi: 10.2991/aer.k.210810.064]
105. Koppel R, Metlay JP, Cohen A, Abaluck B, Localio AR, Kimmel SE, et al. Role of computerized physician order entry systems in facilitating medication errors. JAMA 2005 Mar 09;293(10):1197-1203. [doi: 10.1001/jama.293.10.1197] [Medline: 15755942]
106. Dorado-Díaz PI, Sampedro-Gómez J, Vicente-Palacios V, Sánchez PL. Applications of artificial intelligence in cardiology. The future is already here. Rev Esp Cardiol (Engl Ed) 2019 Dec;72(12):1065-1075. [doi: 10.1016/j.rec.2019.05.014] [Medline: 31611150]
107. Zheng C, Johnson TV, Garg A, Boland MV. Artificial intelligence in glaucoma. Curr Opin Ophthalmol 2019 Mar;30(2):97-103. [doi: 10.1097/ICU.0000000000000552] [Medline: 30562242]
108. Salinas E, Cueva R, Paz F. A systematic review of user-centered design techniques. In: Proceedings of the 9th International Conference, DUXU 2020, Held as Part of the 22nd HCI International Conference on Design, User Experience, and Usability, Interaction Design. 2020 Presented at: HCII '20; July 19-24, 2020; Copenhagen, Denmark p. 253-267 URL: https://link.springer.com/chapter/10.1007/978-3-030-49713-2_18 [doi: 10.1007/978-3-030-49713-2_18]
109. Stagg BC, Tullis B, Asare A, Stein JD, Medeiros FA, Weir C, et al. Systematic user-centered design of a prototype clinical decision support system for glaucoma. Ophthalmol Sci 2023 Sep;3(3):100279 [FREE Full text] [doi: 10.1016/j.xops.2023.100279] [Medline: 36970116]
110. Azadi A, García-Peñalvo F. Private health record system: improving the patient's medical knowledge with an e-learning approach. In: Proceedings of the 10th International Conference on Technological Ecosystems for Enhancing Multiculturality.

XSL•FO
RenderX

2022 Presented at: TEEM '22; October 19-21, 2022; Salamanca, Spain p. 182-191 URL: https://dl.acm.org/doi/10.1007/978-981-99-0942-1_18 [doi: 10.1007/978-981-99-0942-1_18]
111. Bainbridge WA. Memorability: how what we see influences what we remember. In: Federmeier KD, editor. Psychology of Learning and Motivation. New York, NY: Elsevier; 2019:1-27.
112. Amann J, Vetter D, Blomberg SN, Christensen HC, Coffee M, Gerke S, Z-Inspection initiative. To explain or not to explain?-Artificial intelligence explainability in clinical decision support systems. PLOS Digit Health 2022 Feb;1(2):e0000016 [FREE Full text] [doi: 10.1371/journal.pdig.0000016] [Medline: 36812545]
113. Larsson S. On the governance of artificial intelligence through ethics guidelines. Asian J Law Soc 2020 Oct 02;7(3):437-451. [doi: 10.1017/als.2020.19]
114. Laka M, Milazzo A, Merlin T. Factors that impact the adoption of clinical decision support systems (CDSS) for antibiotic management. Int J Environ Res Public Health 2021 Feb 16;18(4):36 [FREE Full text] [doi: 10.3390/ijerph18041901] [Medline: 33669353]
115. Hollnagel E, Wears R, Braithwaite J. From safety-I to safety-II: a white paper. ResearchGate. URL: https://www.researchgate.net/publication/282441875_From_Safety-I_to_Safety-II_A_White_Paper [accessed 2024-04-29]
116. Yao W, Kumar A. CONFlexFlow: integrating flexible clinical pathways into clinical decision support systems using context and rules. Decis Support Syst 2013 May;55(2):499-515 [FREE Full text] [doi: 10.1016/j.dss.2012.10.008]
117. Venkatesh V, Thong JY, Xu X. Unified theory of acceptance and use of technology: a synthesis and the road ahead. J Assoc Inf Syst 2016 May;17(5):328-376. [doi: 10.17705/1jais.00428]
118. Ismail L, Materwala H, Karduck AP, Adem A. Requirements of health data management systems for biomedical care and research: scoping review. J Med Internet Res 2020 Jul 07;22(7):e17508 [FREE Full text] [doi: 10.2196/17508] [Medline: 32348265]
119. Zeng L. Designing the user interface: strategies for effective human-computer interaction (5th Edition) by B. Shneiderman and C. Plaisant. Int J Hum Comput Interact 2009 Sep 25;25(7):707-708. [doi: 10.1080/10447310903187949]
120. Carroll JM, Rosson MB. Getting around the task-artifact cycle: how to make claims and design by scenario. ACM Trans Inf Syst 1992 Apr;10(2):181-212. [doi: 10.1145/146802.146834]
121. Zhang P, Li N. The intellectual development of human-computer interaction research: a critical assessment of the MIS literature (1990-2002). J Assoc Inf Syst 2005 Nov;6(11):227-292. [doi: 10.17705/1jais.00070]
122. Diederich S, Brendel AB, Kolbe LM. Designing anthropomorphic enterprise conversational agents. Bus Inf Syst Eng 2020 Mar 10;62(3):193-209. [doi: 10.1007/s12599-020-00639-y]
123. Ghanbary Sartang A, Ashnagar M, Habibi E, Sadeghi S. Evaluation of Rating Scale Mental Effort (RSME) effectiveness for mental workload assessment in nurses. J Occup Health Epidemiol 2016 Oct 01;5(4):211-217. [doi: 10.18869/acadpub.johe.5.4.211]
124. Cudd PA, Oskouie R. Combining HCI techniques for better user interfacing. In: Proceedings of the 1996 International Conference on Colloquium on Interfaces - The Leading Edge. 1996 Presented at: ILE '96; April 3, 1996; Dundee, UK p. 9 URL: https://ieeexplore.ieee.org/document/543401 [doi: 10.1049/ic:19960799]
125. Desolda G, Ardito C, Costabile MF, Matera M. End-user composition of interactive applications through actionable UI components. J Vis Lang Comput 2017 Oct;42:46-59. [doi: 10.1016/j.jvlc.2017.08.004]
126. Haxhixhemajli D. Visibility aspects importance of user interface reception in cloud computing applications with increased automation. School of Computing Blekinge Institute of Technology. 2012. URL: https://www.diva-portal.org/smash/get/diva2:832617/FULLTEXT01.pdf [accessed 2024-04-29]
127. Canny J. The future of human-computer interaction. Queue 2006 Jul;4(6):24-32. [doi: 10.1145/1147518.1147530]
128. Yuan MJ, Finley GM, Long J, Mills C, Johnson RK. Evaluation of user interface and workflow design of a bedside nursing clinical decision support system. Interact J Med Res 2013 Jan 31;2(1):e4 [FREE Full text] [doi: 10.2196/ijmr.2402] [Medline: 23612350]
129. Ramadani R, Mahdiana D. A systematic literature review of design thinking approach for user interface design. Int J Inform Visualization 2024;5(4):103-111 [FREE Full text]
130. Benyon D. Designing Interactive Systems: A Comprehensive Guide to HCI and Interaction Design. New York, NY: Pearson; 2014.
131. Paneru B, Paneru B, Poudyal R, Bikram Shah K. Exploring the Nexus of User Interface (UI) and User Experience (UX) in the context of emerging trends and customer experience, human computer interaction, applications of artificial intelligence. Int J Inform Inf Syst Compute Eng 2024 Mar 13;5(1):102-113 [FREE Full text]
132. Motlagh SJ, Safaei M. Enhancing human-computer interaction in healthcare: optimizing UI/UX design for electronic health records (EHR) systems. Int J Hum Comput Interact 2022;1(1):31-42 [FREE Full text]
133. Dietlein CS, Bock OL. Development of a usability scale based on the three ISO 9241-11 categories “effectiveness,” “efficacy” and “satisfaction”: a technical note. Accred Qual Assur 2019 Mar 8;24(3):181-189. [doi: 10.1007/S00769-018-01368-2]
134. Carroll C, Marsden P, Soden P, Naylor E, New J, Dornan T. Involving users in the design and usability evaluation of a clinical decision support system. Comput Methods Programs Biomed 2002 Aug;69(2):123-135. [doi: 10.1016/s0169-2607(02)00036-6] [Medline: 12100792]

XSL•FO
RenderX

Abbreviations

AI: artificial intelligence
CDSS: clinical decision support system
DSS: decision support system
EHR: electronic health record
EMR: electronic medical record
HCI: human-computer interaction
ISO: International Organization for Standardization
IU: intention to use
MQ: mapping question
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RQ: research question
RSME: Rating Scale Mental Effort
SLR: systematic literature review
UCD: user-centered design
UI: user interface
UTAUT: unified theory of acceptance and use of technology
UX: user experience

Edited by A Kushniruk; submitted 27.11.24; peer-reviewed by J Brooke, E Pisirir; comments to author 14.01.25; revised version received 10.02.25; accepted 22.03.25; published 06.05.25.

Please cite as:
Azadi A, García-Peñalvo FJ
Optimizing Clinical Decision Support System Functionality by Leveraging Specific Human-Computer Interaction Elements: Insights From a Systematic Review
JMIR Hum Factors 2025;12:e69333
URL: https://humanfactors.jmir.org/2025/1/e69333
doi:10.2196/69333
PMID:

©Ali Azadi, Francisco José García-Peñalvo. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 06.05.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Advancing Regional and Remote Health Care With Virtual Hospital Implementation: Rapid Review

Artika Archana Kumari[1*], BSc, MSc, MDH; Tafheem Ahmad Wani[1*], MIT, PhD; Michael Liem[1*], PhD; James Boyd[1*], PhD; Urooj Raza Khan[1*], PhD

Department of Public Health, School of Psychology and Public Health, La Trobe University, Bundoora, Australia

[*]all authors contributed equally

Corresponding Author:
Artika Archana Kumari, BSc, MSc, MDH
Department of Public Health
School of Psychology and Public Health
La Trobe University
Plenty Rd
Bundoora, 3083
Australia
Phone: 61 3 9479 1763
Email: artika.124@gmail.com

Abstract

Background: Disparities in health equity between metropolitan and rural areas are a global concern, especially in vast countries such as Australia, Canada, and the United States. Virtual care models in health care settings are promising in reducing inequalities, with virtual hospitals (VHs) potentially bridging the gap for isolated or underserved regions. However, evidence-based strategies and the complexities of VH implementation necessitate further research.

Objective: This rapid review aims to examine the role of VHs in enhancing regional and remote health care by focusing on accessibility, patient and health care provider experiences, and implementation barriers and facilitators. It provides tailored recommendations for large-scale implementation in communities with access issues, contributing to the discussion on equitable health care.

Methods: A rapid review was conducted in accordance with the World Health Organization guidelines. A systematic search was performed across PubMed, MEDLINE, CINAHL, and the La Trobe University Library for peer-reviewed articles published between January 2015 and March 2023. Additional gray literature was identified through Google searches and snowballing from relevant web articles. Studies were included if they focused on regional or remote populations and addressed VHs or virtual care. Studies that solely discussed hybrid models of care were excluded. Data were systematically extracted using a customized Microsoft Excel template. A mixed methods thematic analysis was conducted to identify recurring themes, barriers, facilitators, and recommendations related to VH implementation as well as patterns in clinical outcomes and stakeholder perspectives.

Results: A total of 35 articles were included in this review, comprising 23 (66%) peer-reviewed studies and 12 (34%) gray literature sources. Positive clinical outcomes were reported in 9 (26%) articles, highlighting outcomes such as reduced disease transmission, improved patient safety, fewer admissions and readmissions, lower mortality, shorter hospital stays, and better adherence to clinical best practices. Health system outcomes were identified in 15 (43%) articles, including reduced costs, enhanced patient experience and safety, improved care delivery and health care provider support, greater efficiency, broader geographic coverage, and better integration of services. Patient and health care provider perspectives were discussed in 12 (34%) articles, with positive views attributed to convenience, time and cost savings, and improved service quality. Barriers and facilitators were the most frequently discussed themes, appearing in 27 (77%) and 26 (74%) articles, respectively, with challenges and enablers commonly linked to people, processes, technology, and financial sustainability.

Conclusions: VHs have the potential to revolutionize regional and remote health care by overcoming barriers, using facilitators, and following recommended practices, leading to better clinical outcomes and increased satisfaction for patients and health care providers.

(JMIR Hum Factors 2025;12:e64582) doi:10.2196/64582

XSL•FO
RenderX

KEYWORDS

virtual hospitals; regional and remote health care; health care accessibility; patient experiences; health care provider perspectives; rural health care; telehealth; virtual care from home; telemedicine; health system outcomes; clinical outcomes; barriers and facilitators; recommendations

Introduction

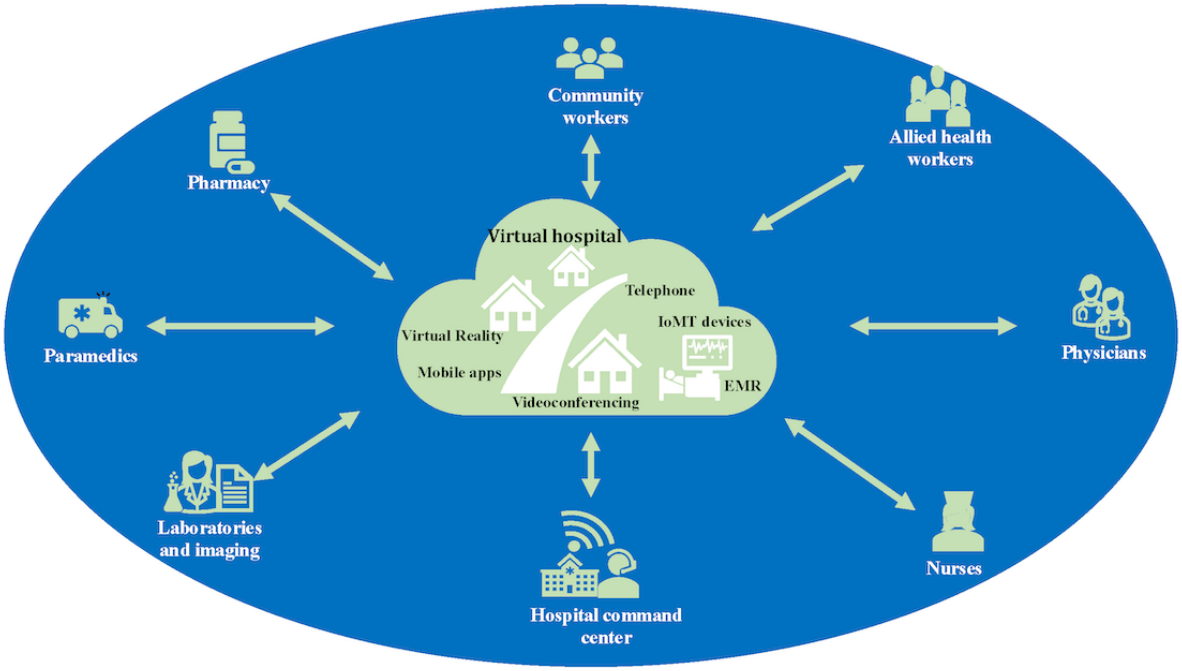
Background

Health care disparities between rural and remote regions and urban centers are a global challenge. Countries such as Australia, Canada, and the United States have documented significant disparities in health care access and outcomes [1-3]. Data from the Australian Institute of Health and Welfare reveal a 1.4-fold increased incidence of disease in remote and very remote regions compared to urban cities [2]. Furthermore, in very remote regions, the risk of premature mortality is elevated by 1.3 times for men and 1.5 times for women compared to metropolitan areas. This disparity aligns with findings from the Centers for Disease Control and Prevention, which highlights the challenges faced by rural Americans due to limited access to specialized and emergency health care, ultimately contributing to a higher prevalence of preventable deaths in these communities [4].

Traditional health care models often struggle to provide equitable services in geographically isolated regions due to shortages of medical facilities, specialized professionals, and advanced diagnostic equipment. Consequently, rural and remote populations experience higher rates of preventable diseases, avoidable hospitalizations, and mortality compared to urban counterparts [2,4,5]. In Australia, hospitalization rates for preventable conditions are 2.5 times higher [2]. Similarly, in the United States, rural populations have a 20% higher overall mortality rate compared to urban counterparts, and 1 in 5 rural adults experience multiple chronic conditions, leading to increased rates of preventable hospitalizations and deaths [5].

The emergence of virtual health offers a promising solution to bridge the health care gap in regional and remote locations. Virtual health, driven by digital technologies, has made considerable strides, transforming primary, secondary, and tertiary care services. Health care institutions have embraced virtual health solutions, such as virtual hospitals (VHs), to provide advanced treatments, surgical consultations, multidisciplinary team discussions, and hospital-at-home care, enhancing patient access to complex medical interventions. VHs leverage digital initiatives, such as telehealth and telemedicine technologies, to deliver a comprehensive range of remote health care services, including consultations, diagnostics, treatment plans, and hospital-level care at home [6]. These services are typically delivered through videoconferencing, web platforms, mobile apps, and other digital channels. Figure 1 further illustrates the concept of a VH.

Figure 1. Concept of a virtual hospital. IoMT: Internet of Medical Things; EMR: electronic medical record.



Unlike traditional hospitals with physical infrastructure, VHs operate primarily in the digital realm. Their significance lies in their ability to transcend geographic barriers, enhancing access to health care for patients in remote or underserved areas [7]. It is important to note that while the terms VH, virtual care, hospital at home (virtual wards), and telehealth are sometimes used interchangeably, there are key distinctions between them. Virtual care represents a broad spectrum of health care services delivered remotely, encompassing a range of synchronous and asynchronous technologies and applications. Talevski et al [8] states that virtual care refers to “the integration of digital technologies across the continuum of care to improve patient

XSL•FO
RenderX

outcomes." This definition covers existing health care programs, such as the Victorian Virtual Emergency Department [8], the Ochsner Connected Health Remote Patient Monitoring (RPM) Program [9], and the Cleveland Clinic Express Care Online [10]. Telehealth, on the other hand, is a subset of virtual care that primarily focuses on real-time communication between patients and providers using telecommunication technology [11]. An example of virtual care outside the scope of telehealth is RPM. RPM systems use devices, such as wearables or sensors, to collect and transmit health data, such as heart rate or glucose levels, to health care providers for analysis [12]. This occurs asynchronously, making it distinct from the real-time communication that characterizes telehealth. On the other hand, VH represents a more comprehensive approach to remote health care, offering a wide range of services beyond consultations, including remote diagnostics, treatment plan development, and virtual wards [13]. Hospital-at-home programs, a subset of VH or stand-alone initiatives, provide hospital-level care in the patient's home, incorporating remote monitoring and regular consultations [14].

Digital health technologies form the cornerstone of modern health care delivery systems, including virtual care, telehealth, VHs, and hospital-at-home programs, facilitating remote health care services through diverse technological platforms and applications. According to the World Health Organization [15], digital health encompasses "the field of knowledge and practice associated with the development and use of digital technologies to improve health". This definition encompasses the various terms often used when discussing the design and implementation of technology in health care, including virtual care, telehealth, and VHs. The Healthcare Information and Management Systems Society [12] further characterizes digital health as a comprehensive framework that incorporates various digital technologies, systems, and platforms designed to deliver health care, manage health information, and enhance patient care delivery.

The COVID-19 pandemic accelerated the adoption of VHs as a means to manage health care demands while mitigating infection risks. This rapid expansion led to demonstrated improvements in patient outcomes, including length of stay (LOS) and improved wait times [16-18]. Success stories, such as the Royal Prince Alfred VH [19], the Mayo Clinic Advanced Care at Home program [20], and the Victorian Virtual Emergency Department [21], further exemplify the capability of VHs to deliver effective health care remotely. The experience gained during the pandemic highlights the need for strategic planning and investment in VHs to address health care disparities in rural and remote areas [22,23].

The successful implementation of a VH is a complex endeavor, requiring careful planning, assessment, and integration of digital technologies [24]. Challenges such as inadequate guidance, poor governance, limited technological literacy, and concerns about safety and privacy can hinder the seamless adoption of VHs and related virtual care-from-home solutions [17,18]. To overcome these barriers and successfully implement VHs, evidence-based recommendations must be prioritized [25].

Research Focus and Aims

Despite the potential of VHs to transform health care delivery, there is a notable lack of evidence supporting context-specific and comprehensive implementation strategies of these systems [26]. Existing research often focuses on individual virtual care programs rather than the broader array of integrated health care technologies used in VHs [27]. To effectively deliver VH services, a deeper understanding of how different virtual care technologies contribute to overall patient care is essential. This knowledge is critical for optimizing the design, implementation, and scalability of VH services, particularly in rural and remote regions.

Current evidence highlights the potential of VHs to revolutionize health care delivery by offering improved clinical and health system outcomes, enhanced provider and patient experiences, and increased health care accessibility. However, this evidence is largely limited to the broader health care landscape, with a noticeable gap in understanding the specific application of VHs in regional and rural settings. Existing studies discuss the clinical effectiveness of VHs but primarily address the general population [28,29]. Similarly, research on the barriers, facilitators, and recommendations for implementing VHs does not align with the unique challenges of regional and remote communities, such as limited access to specialists, longer travel distances, and a higher prevalence of preventable conditions [17,18]. This gap in the literature necessitates research focused on tailored recommendations, barriers, and facilitators specific to VH implementation in regional and remote areas, enabling health care systems to fully harness the potential of this innovative approach.

This rapid review aims to synthesize existing evidence on the role of VHs in enhancing health care delivery in regional and remote areas. It seeks to analyze clinical and health system outcomes, explore patient and provider perspectives, identify implementation-related barriers and facilitators, and provide recommendations for its successful implementation. The findings will contribute to evidence-based policy making and decision-making to optimize the use of VHs in addressing health care disparities in challenging geographic settings.

Methods

Overview

This rapid review was conducted to assess the potential of VHs in addressing health care disparities in regional and remote areas. This approach aligns with the need to understand how VHs address health care disparities and improve access for remote and underserved populations in a timely manner. By analyzing a wide range of literature, a rapid review helped to identify the most current evidence and trends, providing a comprehensive overview of the benefits, challenges, and best practices associated with developing VHs to support regional and remote health care.

Using the World Health Organization rapid review guideline [30] and the 2020 PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement [31], a systematically defined, transparent, and concise approach was

XSL•FO
RenderX

used to identify relevant peer-reviewed articles and gray literature, extract relevant information, and synthesize the findings of this review. A completed PRISMA checklist is provided in Multimedia Appendix 1.

Protocol Development

The research protocol for this study was developed using the population, intervention, comparison, and outcome (PICO) framework [32]. The protocol, as presented in Table 1, guided the development of the search strategy (Multimedia Appendix 2). La Trobe librarians also contributed to the development of the protocol and search strategy.

Table 1. Population, intervention, comparison, and outcome protocol development.

Parameter	Description	Keywords
Population	• Patients living in regional and remote areas • Patients having limited access to health care due to distance and a lack of resources	Regional health, rural health, remote health, rural health services, and remote and regional area health
Intervention	• Building VH[a] using telemedicine, telehealth, or eHealth technologies	Virtual hospital, telemedicine, telehealth, telemonitoring, remote consultation, virtual care, virtual health, eHealth, mHealth, and health virtualization
Comparison	• Physical public and private hospitals (traditional health care delivery models) without VHs or digital health technologies	—[b]
Outcome	• Improved health care access, quality of care, patient satisfaction, health care provider satisfaction, and health outcomes and reduced health care costs	Health care delivery, health care access, health planning, hospital strategy, health care accessibility, clinical effectiveness, user experience, implementation, best practices, recommendations, enablers, strategies, benefits, facilitators, challenges, barriers, problems

[a]VH: virtual hospital.

[b]Not applicable.

Search Strategy

We searched for peer-reviewed articles from January 2015 to March 2023 across 3 key databases: PubMed, MEDLINE, and CINAHL. The La Trobe University Library database was also used. The time frame from January 2015 to March 2023 was intentionally chosen to capture the literature spanning both the prepandemic development of VH models and the significant innovations that emerged in response to the COVID-19 pandemic.

Our search strategy was designed to align with the PICO framework, using 2 distinct approaches. The first strategy targeted articles discussing enablers and facilitators of VH implementation, while the second strategy focused on identifying literature addressing barriers, challenges, and limitations. In addition, to ensure quality and minimize bias, both search techniques were reviewed collaboratively by all authors in consultation with La Trobe research librarians.

To identify relevant gray literature, the research topic "advancing regional and remote health care with virtual hospital implementation" was googled, and from the first 4 pages of the search result, policy documents and web articles that discussed the research questions were marked as relevant and selected for further analysis. Furthermore, snowballing was used to screen and select other gray literature cited in relevant web articles [33].

Inclusion and Exclusion Criteria

A set of inclusion and exclusion criteria was developed using the PICO framework (Textbox 1). This review included (1) studies addressing regional or remote populations or populations facing health care access difficulties; (2) studies discussing recommendations, strategies, facilitators, barriers, challenges, or limitations to VH or virtual care-from-home solutions, telehealth, telemedicine, eHealth, mobile health, and remote monitoring for hospital-level care at home; (3) studies discussing clinical or health system outcomes, user experience, and patient perspectives of the innovative technologies used in developing VHs and (4) studies conducted in a country such as Australia or countries with similar geographic landmasses, such as Canada, the United States, or India. Furthermore, only peer-reviewed studies published in English between January 2015 and March 2023 were included.

This review excluded (1) studies addressing urban or suburban population settings with comprehensive and timely health care access and (2) studies solely discussing hybrid models of care.

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Population: people of all age groups living in regional or remote regions and people with acute or chronic illness experiencing difficulties in accessing health care due to distance and lack of resources
- Intervention: virtual hospital (VH) or virtual care-from-home implementation; telemonitoring, telehealth, and telemedicine implementation supporting care-from-home services; eHealth or mobile health implementation; remote monitoring; interventions in a country such as Australia or countries with similar health care systems and geographic challenges; studies that discussed recommendations, opportunities, and challenges of implementing health care virtualization; and studies that discussed patient or provider experience or successful VH or virtual care-from-home implementation stories
- Comparison: an optional comparison group including people with no limitations to accessing health care but who preferred virtual consultation
- Outcomes: clinical outcomes of health care virtualization (improvements in health for people living in remote or regional areas or having limited access to resources); improved delivery of services; success strategies (stories) of VH implementation; program user experience (access, engagement, usability, experience, and satisfaction); and changes in health care resourcing
- Studies: published in peer-reviewed journals, conference proceedings, or health care services reports; published in the English language; and published between January 2015 and March 2023

Exclusion criteria

- Population: people living in urban or suburban areas and those having no limitation on health care access
- Intervention: traditional health care delivery models with minimum virtualization and virtualization in urban and suburban geographic regions
- Outcomes: studies discussing only the benefits and challenges of telehealth or telemedicine and studies discussing user perspectives only
- Studies: studies or reports from unverified sources that have not been peer reviewed, published in other languages, and published before January 2015 or after March 2023

Screening and Study Selection

The search results were initially filtered for peer-reviewed articles published in English between January 2015 and March 2023. Following the initial database search, 589 records were retrieved from 4 databases. These records were imported into Covidence Systematic Review Software (Veritas Health Innovation) [34], where 255 (43.29%) duplicate records were eliminated automatically. The remaining 334 (56.71%) articles underwent title and abstract screening based on predefined inclusion and exclusion criteria. This stage resulted in the exclusion of 159 (47.6%) articles. The initial screening was conducted by AAK, with URK verifying a subset of the records to ensure consistency and accuracy in applying the criteria.

The remaining 175 full-text articles were then assessed for eligibility using the same inclusion and exclusion criteria. During this phase, the snowballing technique was also used, where relevant peer-reviewed articles cited within these studies were identified and retrieved from Google Scholar for potential inclusion. After full-text screening and the additional retrieval through snowballing, 23 (13.1%) articles met all eligibility criteria and were included in the final analysis.

Data Extraction and Knowledge Synthesis

We used the Quality Criteria Checklist (Multimedia Appendix 3 [35]) to develop a Microsoft Excel template for quality assessment because our review included various study designs. This template helped us critically appraise the included studies and assess their potential bias.

In addition, a data extraction template was customized in Microsoft Excel (Multimedia Appendix 4), which was then used to systematically extract and record relevant information from the peer-reviewed studies and gray literature included. The template was prepared based on the extraction variables retrieved from the research question and objectives. Apart from the core characteristics, such as title, abstract, and author, the data extraction template included demographic location, health care conditions addressed, health care services provided virtually, virtual care technology used, health care outcomes, patient satisfaction, provider perspectives, barriers and facilitators, and recommendations and strategies for implementation. This approach helped ensure consistency and accuracy during the data extraction process. In addition, a concept matrix (Multimedia Appendix 5) was constructed to assess the relevance of each study to specific research objectives.

The extracted information from the included studies was then thematically analyzed to identify recurring patterns, emerging trends, and insights into common recommendations for developing VHs and their related barriers and facilitators. Data were also analyzed to identify common themes emerging from clinical and health care system outcomes and patient and provider perspectives. The thematic analysis approach allowed the identification and interpretation of key themes, enabling the synthesis and presentation of findings in a coherent and meaningful manner [36].

Results

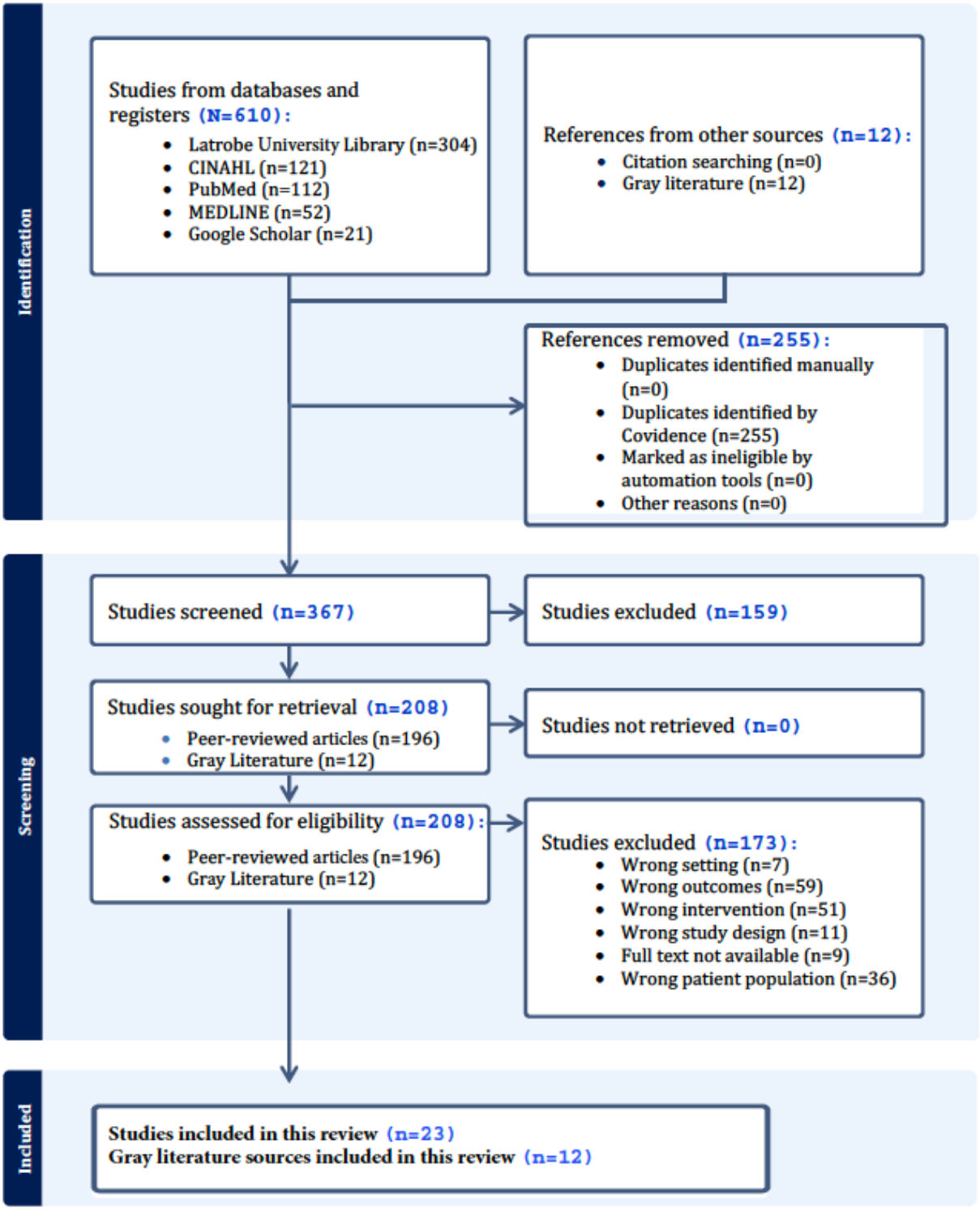
Study Characteristics

Overall, 23 peer-reviewed articles and 12 gray literature sources were included in our review. The selection process is summarized in the PRISMA flow diagram (Figure 2). Detailed characteristics of all included peer-reviewed articles and gray literature are presented in Tables 2 and 3. Among the 23 peer-reviewed articles, 12 (52%) focused on telehealth interventions [37-48], 6 (26%) discussed VHs or

XSL•FO
RenderX

hospital-at-home interventions [49-54], 2 (9%) examined telemedicine [55,56], 2 (9%) addressed RPM [57,58], and 1 (4%) evaluated a virtual pharmacy [59]. In the 12 gray literature sources, 8 (67%) discussed virtual care-from-home solutions comprehensively [60-68], while 4 (33%) articles specifically addressed the concept of a VH [62,69-71].

Figure 2. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) diagram.



XSL•FO
RenderX

Table 2. Included peer-reviewed articles.

Authors	Title	Country	Study design	Population	Health care conditions addressed	Intervention discussed
AlDossary et al [55], 2017	The development of a telemedicine planning framework based on needs assessment	United States	Case study	General rural population	Equitable access to specialty health care services	Telemedicine
Davis et al [38], 2020	Designing a multifaceted telehealth intervention for a rural population using a model for developing complex interventions in nursing	United States	Case study	Study: a team of HPML[a] faculty members, a graduate student, a program manager, 2 telehealth experts, and state stakeholders; stakeholders: the TMH[b] director, BMS[c] medical director, 2 BMS nurses, the program manager for the state traumatic brain injury waiver, the director of the WV[d] home- and community-based service program, 1 Medicaid waiver participant advocate, and the chief operations officer.	Multiple chronic conditions	Telehealth
DeHart et al [39], 2022	Benefits and challenges of implementing telehealth in rural settings: a mixed-methods study of behavioral medicine providers	United States	Mixed methods study	17 health care providers	Mental Health	Telehealth
Giroux et al [50], 2022	It's not one size fits all: a case for how equity-based knowledge translation can support rural and remote communities to optimise virtual health care	Canada	Commentary	General rural population	Equitable access to health care services	VH[e]: hospital in the home
Haque et al [40], 2021	Factors influencing telehealth implementation and use in frontier critical access hospitals: Qualitative study	United States	Qualitative study	36 stakeholders (8 critical access hospitals)	Frontier critical access care	Telehealth
Head et al [51], 2022	Virtual visits for outpatient burn care during the COVID-19 pandemic	United States	Quantitative study	143 patients	Burn care	VH: hospital in the home
Hirko et al [41], 2020	Telehealth in response to the COVID-19 pandemic: Implications for rural health disparities	United States	Commentary	General rural population	Ambulatory practices and triage before hospitalization	Telehealth

XSL•FO
RenderX

Authors	Title	Country	Study design	Population	Health care conditions addressed	Intervention discussed
Howland et al [42], 2021	Psychiatrist and psychologist experiences with telehealth and remote collaborative care in primary care: a qualitative study	United States	Qualitative study	10 telepsychiatrists and 4 telepsychologists	Mental Health	Telehealth
Jong et al [44], 2019	Enhancing access to care in northern rural communities via telehealth	Canada	Commentary	General rural population	General health care	Telehealth
LeBlanc et al [58], 2020	Patient and provider perspectives on eHealth interventions in Canada and Australia: A scoping review	Australia and Canada	Scoping review	General rural population	General health care	Remote patient monitoring
Thomas et al [48], 2023	Health workforce perceptions on telehealth augmentation opportunities	Australia	Qualitative study	53 health care professionals	Primary health care	Telehealth
Johnston et al [43], 2020	"From the technology came the idea": safe implementation and operation of a high quality teleradiology model increasing access to timely breast cancer assessment services for women in rural Australia	Australia	Qualitative study	55 service providers	Breast cancer screening	Telehealth
Correale et al [49], 2022	A model to implement standardised virtual care for low back pain amongst a large network of providers in urban and rural settings	Canada	Case study	12 focus groups	Management of lower back pain	VH: hospital in the home
Kocanda et al [45], 2021	Informing telehealth service delivery for cardiovascular disease management: exploring the perceptions of rural health professionals	Australia	Qualitative study	10 health care professionals	Cardiovascular disease	Telehealth
McPherson and Nahon [46], 2021	Telehealth and the provision of pelvic health physiotherapy in regional, rural and remote Australia	Australia	Qualitative study	54 participants	Pelvic health	Telehealth
Allan et al [59], 2021	"This is streets ahead of what we used to do": staff perceptions of virtual clinical pharmacy services in rural and remote Australian hospitals	Australia	Case study followed by qualitative analysis	15 focus groups (42 nurses, 8 physicians, 12 managers, 4 allied health staff, and 3 pharmacists)	Use of medications in rural and remote NSW[f] public hospitals	Virtual pharmacy

XSL•FO
RenderX

Authors	Title	Country	Study design	Population	Health care conditions addressed	Intervention discussed
Nataliansyah et al [47], 2022	Managing innovation: a qualitative study on the implementation of telehealth services in rural emergency departments	United States	Qualitative study	18 participants	Emergency department	Telehealth
Kuperman et al [52], 2018	The virtual hospitalist: a single site implementation bringing hospitalist coverage to critical access hospitals	United States	Case study	General rural population	Emergency and inpatient care at critical access hospitals	VH: hospital in the home
Sitammagari et al [53], 2021	Insights from rapid deployment of a virtual hospital as standard care during the COVID-19 pandemic	United States	Case study	1477 patients	COVID-19	VH: hospital in the home
Vindrola-Padros et al [54], 2021	Remote home monitoring (virtual wards) for confirmed or suspected COVID-19 patients: a rapid systematic review	United Kingdom, United States, Canada, China, the Netherlands, Ireland, Brazil, and Australia	Systematic review	General rural population	COVID-19 positive cases	VH: hospital in the home
Haleem et al [56], 2021	Telemedicine for health care: capabilities, features, barriers, and applications	India	Commentary	General rural population	General health care	Telemedicine
Gray et al [57], 2022	The rapid development of virtual care tools in response to COVID-19: case studies in three Australian health services	Australia	Qualitative study	13 participants	COVID-19	Remote patient monitoring
Bradford et al [37], 2016	Telehealth services in rural and remote Australia: a systematic review of models of care and factors influencing success and sustainability	Australia	Systematic review	General rural population	General health care	Telehealth

[a]HPML: health policy, management, and leadership.

[b]TMH: Take Me Home.

[c]BMS: Bureau for Medical Services.

[d]WV: West Virginia

[e]VH: virtual hospital.

[f]NSW: New South Wales.

XSL•FO
RenderX

Table 3. Included gray literature sources.

Authors or sources	Title	Country	Literature type	Health care conditions addressed	Intervention
Shaw and Wilson [71], 2022	RPA virtual hospital proof of concept trial evaluation report 2020–2021	Australia	Web report	Overall Hospital care	VH[a]
Hardy et al [67]	Improving equity of access in rural and regional health through hybrid and connected care	Australia	Web report	General health	Virtual care from home
Microsoft News Center [70], 2022	Remote Australian community harnesses mixed reality and space technologies to deliver better health care	Australia	Web report	General health	VH
Hospital and Healthcare [69], 2021	The virtual hospital	Australia	Web report	General health	VH
Agency for Clinical Innovation [60], 2021	Virtual care in practice	Australia	Web report	General health	Virtual care from home
Department of Health, Australian government [63], 2022	Future focused primary health care: Australia's Primary Health Care 10 Year Plan 2022-2032	Australia	National strategy	General health	Virtual care from home
Queensland government [66], 2021	Digital strategy for rural and remote healthcare - 10 year plan	Australia	State strategy	General health	Virtual care from home
Northern Territory government of Australia [65]	NT health virtual care strategy	Australia	State strategy	General health	Virtual care from home
Victoria Department of Health [68], 2022	Virtual care standard and guide	Australia	State strategy	General health	Virtual care from home
Department of Health, Australian government [61], 2016	National strategic framework for rural and remote health	Australia	National strategy	General health	Virtual care from home
Department of Infrastructure, Australian government [62], 2022	Better connectivity plan for regional and rural Australia	Australia	Web report	General health	VH
National Rural Health Alliance [64], 2021	Regional telecommunications review 2021	Australia	Telecommunications survey report	General health	Virtual care from home

[a]VH: virtual hospital.

The included 23 peer-reviewed articles were from 9 different countries, with 10 (43%) studies from the United States [38-42,47,51-53,55], 7 (30%) studies from Australia [37,43,45,46,48,57,59], 3 (13%) from Canada [44,49,50], and 1 (4%) from India [56]. Moreover, 1 (4%) study analyzed data from both Australia and Canada [58], while 1 (4%) was a multinational study where data were combined from the United Kingdom, the United States, Canada, China, the Netherlands, Ireland, Brazil, and Australia [54]. The 12 items of gray literature included in this review were all from Australia [60-71].

The selected academic articles covered a wide range of health care issues. Of the 23 peer-reviewed articles, 5 (22%) concentrated on general and primary health care [37,44,48,56,58], 3 (13%) addressed COVID-19 [53,54,57], 2 (9%) focused on emergency and inpatient care [47,52], and 2 (9%) discussed mental health [39,42]. In addition to this, there were 2 (9%) studies discussing equitable access to health care services [50,55] and 1 (4%) study each addressing ambulatory practices and triage [41], breast cancer assessment [43], burn care [51], cardiovascular disease [45], lower back pain [49], multiple chronic conditions [38], pelvic health [46], medication use [59], and frontier critical access care [40]. All 12 (100%) of the 12 items of gray literature addressed general health, focusing mainly on providing equitable health care services virtually [60-71].

In terms of digital technology used for VH implementation, most (20/23, 87%) of the peer-reviewed articles explored the use of videoconferencing [37-39,41-45,47-53,55-59], followed by telephone communication [38,40,41,49,50,53-55,58] and wearable sensors [4,39,41,43,44,54,56,57]. In total, 30% (7/23) of the studies focused on the use of online electronic medical record systems [39,40,49,53,54,57,59], 22% (5/23) of the studies described mobile apps [39,41,48,50,57], while 17% (4/23) of the studies highlighted text messaging for virtual care

XSL•FO
RenderX

[41,46,50,57]. Other technologies, such as store and forward (3/23, 13%) [37,43,44], email and chatbot (1/23, 4%) [57], and robotic ultrasonography (1/23, 4%) [44] were also included as supportive VH technologies. Among the 12 gray literature sources, only 2 (17%) reports explicitly addressed virtual care-from-home technologies. Shaw and Wilson [71] discussed the use of videoconferencing, Internet of Medical Things devices, and telephone technology for remote monitoring and directed self-care, while Microsoft News Center [70] concentrated on the use of mixed reality and space technology to support VH implementation. The remaining 10 (83%) gray literature sources holistically discussed the implementation and use of digital care technologies.

Study Design

Among the 23 peer-reviewed studies, 6 (26%) were case studies [38,49,52,53,55,59] and 1 (4%) was a mixed methods study [39]. In addition, there were 8 (35%) qualitative studies [40,42,43,45-48,57], 1 (4%) quantitative study [51], 1 (4%) scoping review [58], 2 (9%) systematic reviews [37,54], and 4 (17%) low-quality commentaries [41,44,50,56]. Among all included articles, only 2 (9%) reported selection bias [51,53]. Among the 12 gray literature sources, there were 2 (17%) national strategies [61,63], 3 (25%) state-level strategy documents [65,66,68], 6 (50%) web reports [60,62,67,69-71], and 1 (8%) telecommunication survey report [64].

Clinical Outcomes

Only 7 (30%) academic studies of the 23 peer-reviewed articles and 2 (17%) of the 12 gray literature sources assessed the clinical outcomes of the VHs or virtual care-from-home models.

Findings from existing literature demonstrated that VHs have contributed toward positive clinical outcomes, including specialist medication advice, better medication knowledge, reduced medication errors [59], reduced disease transmission [53,56], minimized disease exposure to staff [53], improved patient safety [57,59], reduced admission and readmission [52,71], lower mortality, shorter LOS, and maximized adherence to clinical best practices [52,54,69].

According to Allan et al [59], VHs facilitated enhanced medication management by enabling specialists to provide tailored advice based on individual patient needs. This personalized approach improved patient understanding and medication adherence. Furthermore, by facilitating direct communication between health care providers and specialists, VHs could help reduce medication errors because specialists could review prescriptions remotely, identify potential errors, and provide immediate feedback or corrections.

VHs also enabled remote consultations, diagnosis, and monitoring, reducing the need for physical interaction between patients and health care staff [53,56]. This reduced the risk of disease transmission, particularly in cases of infectious diseases. Moreover, by minimizing in-person visits, health care staff were exposed to fewer patients, reducing their risk of contracting contagious diseases.

Gray et al [57] and Allan et al [59] emphasized the role of continuous monitoring and timely interventions in enhancing patient safety. Remote monitoring devices and telemedicine consultations enabled early detection of complications, facilitating prompt medical interventions and reducing the risk of adverse events. Through remote monitoring and timely interventions, VHs could also manage chronic conditions more effectively, thereby reducing the frequency of hospital admissions and readmissions [52,71]. In addition, by providing patients with better access to health care resources and support, VHs could prevent exacerbations of chronic conditions, further reducing hospitalizations.

VHs have demonstrated potential in improving patient outcomes and reducing health care costs. By enabling continuous monitoring and timely interventions, VHs contributed to reduced mortality rates, shorter hospital stays, and lower readmission rates. Vindrola-Padros et al [54] reported that virtual wards for patients with COVID-19 had lower mortality rates, with only 29% admissions to the emergency department and <36% readmission rate. Likewise, Kuperman et al [52] reported that the pilot VH recorded a 6% decrease in patient transfers, a 17% decrease in admissions, and a 3.7% decrease in LOS. Furthermore, Head et al [51] discussed that virtual care-from-home service for patients with burn injury had similar clinical outcomes when compared to in-person care. None (0/23, 0%) of the studies reported negative clinical outcomes for VHs.

Despite diverse health care conditions, studies revealed similar clinical outcomes. For instance, Head et al [51] concentrated on burn care, Kuperman et al [52] discussed emergency and inpatient care, and Shaw and Wilson [71] addressed overall hospital care; however, all reported that VHs reduced admission, readmission, and transfer in their respective studies. Likewise, Gray et al [57], addressing virtual pharmacy, and Allan et al [59], discussing COVID-19, discovered that virtual care-from-home solutions improved patient safety.

Health System Outcomes

A total of 12 (52%) academic studies of the 23 peer-reviewed articles and 3 (25%) of the 12 gray literature sources discussed 9 positive health system outcomes.

The reported outcomes included improved access to care and efficient use of resources [39,52,54,71], reduced patient and health care costs [37,39,48,51,53,56,58], enhanced patient experience and safety [56-58,69], enhanced health care delivery and provider support through trusted workflow and virtual relocation [53,54,58,59], improved process outcomes and enhanced compliance [37,52,54,58,71], enhanced geographic coverage, and better integration of health care services [67,71].

VHs used telemedicine technologies to provide remote health care services, enabling patients to access health care professionals without physical travel and enhancing access to care in underserved areas [39,52,54,71]. Similarly, VHs enhanced health care efficiency by streamlining delivery processes, reallocating resources toward virtual care platforms, reducing physical infrastructure requirements, and resulting in cost savings and improved service delivery.

VHs also lowered patient and health care costs by minimizing the need for in-person visits, hospitalizations, and associated travel expenses. Telemedicine consultations and remote

XSL•FO
RenderX

monitoring enabled timely interventions, preventing costly complications and reducing the overall health care expenditure for patients and health care providers. According to Head et al [51], virtual visits resulted in significant savings, including an average reduction of 130 miles in travel distance, 164 minutes of travel time, US $104 in travel costs, and US $81 in forgone wages due to the time saved from not having to travel. None (0/23, 0%) of the articles highlighted negative health care system outcomes.

Furthermore, VHs could offer convenient and personalized health care experiences tailored to patients' needs because remote consultations and digital health tools empowered patients to actively participate in their care, leading to greater satisfaction and engagement [56-58,69]. In addition, virtual care models prioritized patient safety by minimizing exposure to health care–associated infections and promoting adherence to evidence-based practices. Similarly, by leveraging technology to optimize health care delivery processes and support health care providers in delivering high-quality care, VH facilitators had managed to enable seamless coordination among multidisciplinary teams, improving communication, collaboration, and patient outcomes [53,54,58,59]. The virtual support systems also enhanced provider efficiency and job satisfaction, ultimately enhancing the overall quality of care delivery.

Findings from the literature also revealed that VHs improved health care by standardizing workflows, promoting clinical adherence, and facilitating efficient data management [37,52,54,58,71]. The use of digital health solutions helped improve care coordination and patient outcomes while ensuring regulatory compliance and quality standards in health care delivery.

The findings were consistent across different health care conditions. While Head et al [51], DeHart et al [39], Kuperman et al [52], Shaw and Wilson [71], and Vindrola-Padros et al [54] focused on burn care, mental health, emergency and inpatient, general hospital care, and COVID-19, respectively, their studies collectively highlighted that health care virtualization significantly improved access to care. Furthermore, the Australian-based Royal Prince Alfred VH received positive feedback from patients for the convenience and stress-reducing benefits of receiving care from home [71].

In addition, Hardy et al [67] and Shaw and Wilson [71] reported that VHs offered health care services beyond traditional boundaries, improving accessibility and enabling broader geographic coverage. Telemedicine and remote monitoring enabled patients to have specialized care, reducing disparities in access. This care model also facilitated better integration of health care services by fostering collaboration among providers, improving care transitions, and promoting continuity of care across different settings.

Patient Perspectives

A total of 8 (35%) of the 23 peer-reviewed articles and 1 (8%) web report of the 12 gray literature sources examined patient perspectives regarding the implementation of VHs and virtual care-from-home solutions.

Allan et al [59] and Correale et al [49] highlighted high patient acceptance and satisfaction with virtual care-from-home services. Similarly, DeHart et al [39] noted increased satisfaction due to time, cost, and travel savings from telehealth services for rural patients. Likewise, Giroux et al [50] disclosed that quality virtual care-from-home enhanced empowerment, improved self-management, and expanded access to culturally appropriate care locally.

Patient involvement in remote care is associated with improved outcomes and satisfaction, as suggested by Vindrola-Padros et al [54]. VHs provided patients with the convenience of receiving treatment locally rather than having to relocate to a secondary care hospital, thus leading to an increased level of satisfaction, as highlighted by Gray et al [57] and Kuperman et al [52]. While none (0/23, 0%) of the studies discussed the negative patient perspectives, LeBlanc et al [58] stated that patient satisfaction was highly dependent on technological proficiency, health care provider readiness, and the quality of care provided virtually.

Health Care Provider Perspectives

A total of 8 (35%) of the 23 peer-reviewed articles and 1 (8%) of the 12 gray literature sources discussed the provider perspectives on the implementation of VHs and virtual care-from-home solutions.

Positive health care provider perspectives included increased provider satisfaction [43,59], enhanced service delivery [43,59], better collaboration and teamwork [39,49,50,60], and increased efficiency and accessibility [52,56,58,60]. Providers reported that VHs and hospital-in-the-home solutions offered them more flexibility in delivering care and managing their workload. They could engage in telemedicine consultations, remote monitoring, and virtual care delivery, which led to reduced burnout and increased job satisfaction. The convenience and efficiency of virtual care platforms contributed to a more satisfying work environment for health care professionals. Both peer-reviewed articles and gray literature also highlighted that providers reported that VHs improved health care delivery by enabling telemedicine consultations and remote monitoring, thus allowing them to reach a wider patient population, deliver timely interventions, and offer continuous support. Furthermore, the providers stated that the integration of digital health tools allowed for enhanced service quality, accessibility, and patient outcomes.

In addition, health care providers expressed that virtual care platforms allowed them to facilitate collaboration and communication among health care team members, regardless of their physical location [52,56,58,60]. Rural providers could easily consult with specialists, share patient information, and coordinate care plans in real time, promoting interdisciplinary collaboration and teamwork. This seamless exchange of information enhanced care coordination and improved patient outcomes.

Studies by Allan et al [59] and Johnston et al [43] indicated that enhanced service timeliness, greater accessibility, and reduced provider travel time increased provider satisfaction. They also stated that improved workflows and clinic processes and better

XSL•FO
RenderX

medication management further contributed to enhanced service delivery.

No patients or health care providers reported negative views of the results.

Barriers to VH Implementation

Barriers to VH implementation were reported in 21 (91%) of the 23 peer-reviewed articles and 6 (50%) web reports of the 12 gray literature sources.

Poor digital literacy and language barriers among rural patients emerged as significant obstacles, hindering successful VH adoption. Language barriers and a lack of digital skills presented challenges in connecting remote patients with virtual care solutions, contributing to a digital gap [48,50,53,58,71]. The lack of digital skills also presented a substantial obstacle in connecting remote patients with virtual care-from-home solutions [39,57,71]. Poor digital literacy and language barriers among patients often deterred them from engaging in virtual care services [57] and led to a lack of patient commitment to continue using virtual care-from-home services [39]. Furthermore, health care providers' low adoption and lack of provider knowledge or skills were reported, resulting in resistance to change and workforce constraints, indicating a need for targeted strategies to address these issues [37-40,44,48,53,68,71].

Technological barriers, such as poor technical infrastructure and connectivity [41,45,48-51,53,56,58,64,67,68], limited technical support [39,46,50,57,58,60,68,71], restricted access to digital equipment [38,43,46,49,53,57,60], and legacy and outdated technologies [47,68,71], were also identified as significant barriers to VH implementation for regional and remote communities.

In remote and rural areas, inadequate technical infrastructure and connectivity hindered the effective deployment of VH services [41,45,48-51,53,56,58,64,67,68]. Without reliable internet access and robust communication networks, health care providers and patients faced challenges in accessing telemedicine consultations, remote monitoring, and digital health platforms.

Similarly, the absence of adequate technical support exacerbated challenges associated with VH implementation [39,46,50,57,58,60,68,71]. Health care providers and patients encountered difficulties in troubleshooting technical issues, configuring digital devices, or navigating telemedicine platforms.

Another major technological barrier faced by individuals in remote and underserved communities was the lack of access to digital devices, such as tablets, webcams, or computers, which were essential for participating in virtual care consultations [38,43,46,49,53,57,60]. The digital divide exacerbated disparities in health care access and highlighted the need for initiatives aimed at providing digital equipment to marginalized populations.

Moreover, the presence of legacy and outdated technologies impeded the adoption and effectiveness of VH solutions [47,68,71]. Incompatible systems, outdated software, and obsolete hardware hindered interoperability, data exchange, and user experience.

Several studies reported barriers related to processes of coordination and communication among stakeholders [38,58], cultural neglect for rural communities [66,68], absence of robust governance and effective leadership [42,43,59,68,71], data privacy and security concerns [38,48,49,56,58,69], organizational challenges in relation to resource planning [54,59], operational processes [38,42,60], and limited policies supporting rural community inclusion [45,48,54,57].

According to Davis et al [38] and LeBlanc et al [58], effective coordination and communication among stakeholders were essential for the successful implementation of VHs. Barriers within these processes, such as fragmented communication channels or a lack of standardized protocols, impeded the seamless delivery of care and collaboration among health care providers, patients, and support staff.

In addition, the cultural neglect of rural communities highlighted disparities in health care access and delivery. Policies and initiatives that failed to consider the unique cultural, social, and geographic factors of rural populations exacerbated health care inequities [66,68]. Recognizing and addressing cultural diversity and rural-specific health care needs were essential for ensuring inclusive and patient-centered VH services.

Furthermore, inadequate governance frameworks and leadership support led to ambiguity, resistance to change, and lack of accountability, hindering progress and sustainability [42,43,59,68,71]. Strong leadership and governance mechanisms were essential for guiding strategic planning, resource allocation, and decision-making processes in VH initiatives, and these were fundamental for driving organizational change and ensuring the successful implementation of VHs.

Another critical consideration for VH implementation was data privacy and security concerns, particularly regarding the storage, transmission, and sharing of sensitive patient information [38,48,49,56,58,69]. Failure to address these concerns could impact patient trust, compromise confidentiality, and expose health care organizations to legal and regulatory risks. Robust data privacy policies, secure information systems, and adherence to industry standards were considered essential for safeguarding patient data and maintaining confidentiality.

Organizational challenges, such as inadequate resource planning and inefficient operational processes, impeded the scalability and sustainability of VHs [54,59]. Limited resources, workforce shortages, and inefficient workflows hindered the delivery of timely and quality care, affecting patient outcomes and satisfaction. Strategic resource allocation, process optimization, and continuous quality improvement were necessary for overcoming organizational barriers and enhancing operational efficiency in VHs.

In addition, limited policies supporting rural community inclusion also affected the implementation of equitable VH services [45,48,54,57]. The absence of supportive policies may have limited funding opportunities, regulatory incentives, and infrastructure development for VHs in rural areas. Advocacy for policies that prioritized rural community inclusion, addressed

XSL•FO
RenderX

health care disparities, and incentivized VH adoption were essential for promoting equitable access to health care services.

Furthermore, a few (4/35, 11%) studies identified financial sustainability challenges [47,57,60,69] and the unfair distribution of funds among different health care settings [42,47,57].

Facilitators of VH Initiatives

Facilitators for the successful implementation of VHs were identified in 19 (83%) of the 23 peer-reviewed articles and 7 (58%) web reports of the 12 gray literature sources.

The literature strongly suggested that key facilitators included clinical leadership and advocacy [37,48,71] together with staff and patient enrichment and training [40,42,43,46,47,52,54,59,71], emphasizing the importance of enhancing patient and provider capabilities for successful VH implementation. The Australian federal and state governments are investing in digital literacy upskilling programs to empower patients to self-manage symptoms and navigate VHs confidently [67,68].

Few other facilitators were recognized as a way to mitigate barriers such as collaboration and partnerships [43,48,57,59,71], effective communication [48,49,52,65], purposeful planning and development [37,60,65,71], a comprehensive governance structure [60,65,71], optimized organizational structure [38,43,49,53], iterative system implementation [59,65,71], and supportive government policies [37,51,67,69].

Collaboration and partnerships among stakeholders, including health care providers, government agencies, technology vendors, and community organizations, were instrumental in overcoming barriers and driving the adoption of VHs [43,48,57,59,71]. By fostering collaboration, organizations could leverage collective expertise, resources, and networks to address challenges; share best practices; and promote innovation in virtual care delivery. Likewise, effective communication strategies were essential for ensuring clear understanding, alignment, and engagement among stakeholders involved in VH implementation. Open and transparent communication channels facilitated knowledge sharing, stakeholder engagement, and problem-solving, enabling organizations to address concerns, manage expectations, and drive consensus toward common goals [48,49,52,65].

Purposeful planning and development also incorporated strategic initiatives aimed at designing, implementing, and optimizing VHs [37,60,65,71]. By adopting a systematic approach to planning, organizations could anticipate challenges, identify opportunities, and develop robust strategies for addressing barriers and achieving desired outcomes in virtual care delivery. Similarly, a comprehensive governance structure provided the framework for guiding decision-making, accountability, and oversight in VH initiatives [60,65,71]. Clear governance mechanisms, policies, and procedures ensured alignment with organizational goals, regulatory requirements, and industry standards while also promoting transparency, efficiency, and stakeholder engagement.

In addition, an optimized organizational structure will align roles, responsibilities, and workflows to support the effective delivery of VH services [38,43,49,53]. By streamlining processes, clarifying responsibilities, and fostering collaboration, organizations can enhance efficiency, agility, and responsiveness in virtual care delivery, ultimately improving patient outcomes and satisfaction. Likewise, iterative system implementation will involve continuous assessment, refinement, and optimization of VH services and processes. By adopting an iterative approach, organizations can adapt to evolving needs, address emerging challenges, and incorporate feedback from stakeholders, thereby enhancing the effectiveness, usability, and sustainability of virtual care solutions [59,65,71].

Moreover, supportive government policies created an enabling environment for VH implementation by providing regulatory intelligibility, funding support, and incentives for innovation. Policies that promoted telemedicine reimbursement, infrastructure development, workforce training, and data privacy enhanced the viability and scalability of virtual care initiatives while also fostering collaboration and alignment among stakeholders across the health care ecosystem [37,51,67,69].

These facilitators highlighted the importance of strategic planning, clear communication, and collaborative efforts involving various stakeholders to overcome process-related challenges and ensure successful VH deployment.

Investment in *technical infrastructure* [59,60,65,67,68,70], quality, user-friendly and affordable technology [53,59,60,67,68], and technical outreach and support services [43,58,68] were noted as enablers for VH initiatives.

Robust technical infrastructure formed the foundation for VHs, facilitating seamless communication, data exchange, and service delivery [59,60,65,67,68,70]. Hence, investment in infrastructure, such as high-speed internet connectivity, secure data storage, and interoperable systems, was essential for ensuring reliable and efficient virtual care delivery. Without adequate infrastructure, VHs would face connectivity issues, data security risks, and operational inefficiencies, hindering their effectiveness in reaching and serving rural patients.

Furthermore, access to quality, user-friendly, and affordable technology was critical for enhancing the accessibility and usability of VHs [53,59,60,67,68]. User-friendly interfaces, intuitive design, and affordability ensured that health care providers and patients could easily navigate and use virtual care platforms. Quality technology solutions also contributed to positive user experiences, engagement, and satisfaction, ultimately improving the effectiveness and adoption of VHs.

Similarly, technical outreach and support services played a vital role in assisting health care providers and patients in navigating and troubleshooting virtual care technology services [43,58,68]. These services might include training programs, help desks, and technical assistance teams that provide guidance, education, and troubleshooting support. By offering ongoing technical support, organizations could empower users to effectively use VH platforms, address technical issues, and maximize the benefits of remote care delivery.

These facilitators emphasized the importance of a robust technical foundation, user-friendly interfaces, and support services to overcome technological challenges and enhance the acceptance of VHs within regional and remote communities.

XSL•FO
RenderX

Sustainable financing also emerged as a crucial facilitator, emphasizing the importance of permanent insurance reimbursement solutions and grant funding to support and expand the reach of VH services [38,47,53].

Discussion

Principal Findings

This review demonstrated clinical and health system outcomes of VHs in remote and rural health settings. It addresses the current evidence gap regarding the impact of VHs for regional and rural populations on clinical, health system, patient, and provider outcomes, indicating that VHs significantly enhance clinical effectiveness, regardless of population demographics.

In terms of clinical effectiveness, this review reported positive outcomes against key indicators commonly used to evaluate health care interventions, such as mortality rates, readmission rates, LOS, and improvement in clinical indicators [71]. These findings are consistent with those of Norman et al [17], who reported similar benefits from hospital-at-home care for older individuals, such as shorter LOS and reduced readmission rates. However, it should be noted that this review does not examine the impact of VHs on chronic condition management, a critical area for evaluating the overall effectiveness of VHs. Further research is needed to explore how VHs can better support chronic disease management and improve long-term health outcomes.

Furthermore, this study highlights the health system benefits of VHs, noting their potential to reduce patient and health care expenditures for various conditions, including general health, mental health, burn care, and infectious diseases such as COVID-19. The analysis correlates with key metrics such as quality care, equitable access, cost management, and population health improvement, affirming that VHs significantly enhance health system outcomes in regional and remote communities.

Moreover, the health system outcomes identified in this review align with those observed in urban populations, as reported by Bidoli et al [13] and Snoswell et al [72]. Evidence indicates that VHs positively impact health system outcomes by improving efficiency, accessibility, and resource management, thus proving to be a valuable addition to contemporary health care delivery.

The findings also reveal that positive patient and provider perspectives are contingent on adequate internet connectivity, user-friendly technology, sufficient training, and enhanced collaboration between urban and rural health care providers. A comparison with the studies by Denny and Hill [73], who evaluated patient and health care worker perspectives of virtual care delivery for patients with cystic fibrosis in a suburban setting, and Babaei et al [18], who assessed barriers to and facilitators of virtual care, revealed similar findings. Both studies support the argument that high patient and provider satisfaction heavily depends on adequate internet connectivity, effective technology, and robust technical support systems.

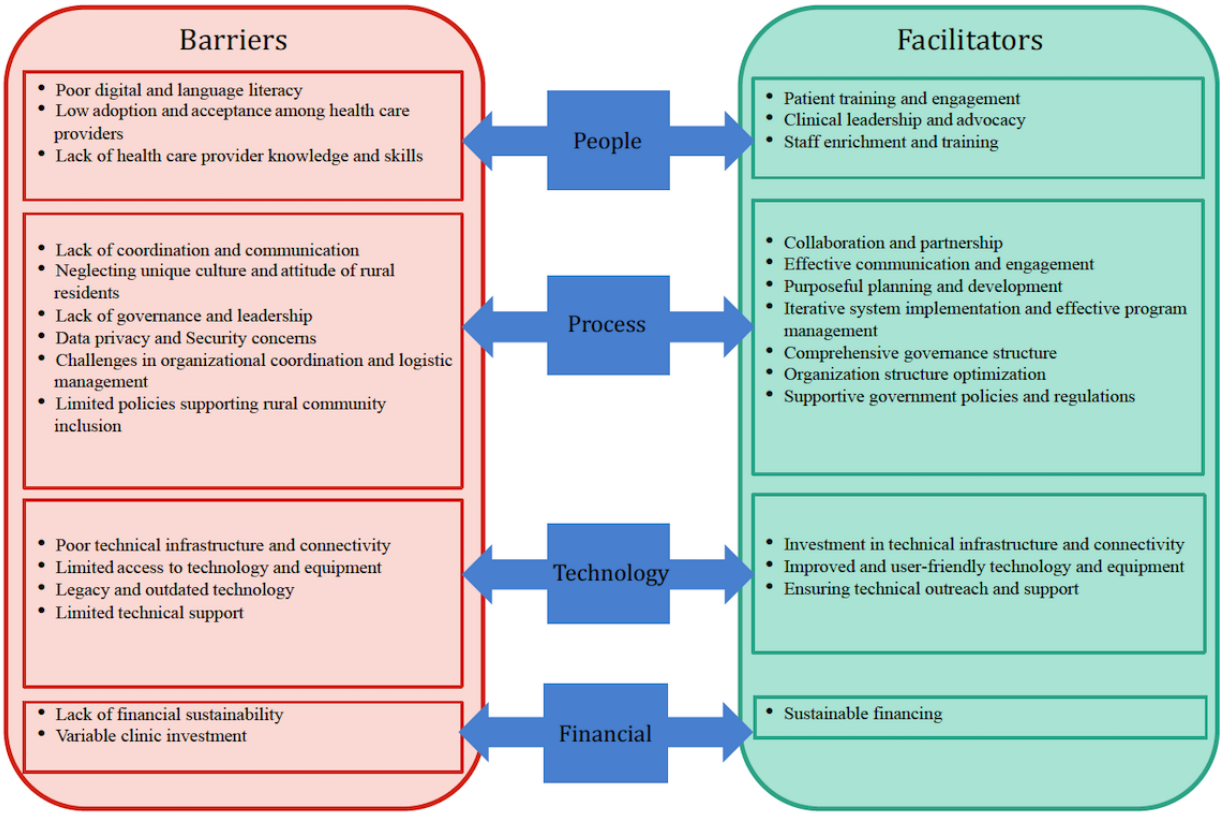
Beyond the findings of this study, it is worth noting that VHs also have the transformative potential in mitigating climate change while maintaining effective patient care. King et al [74] revealed a 99.37% reduction in carbon emissions per appointment compared to in-person visits. Likewise, Thiel et al [75] demonstrated that telemedicine reduced greenhouse gas emissions by nearly 17,000 metric tons in 2021, equivalent to the annual energy use of 2100 homes, by replacing in-person visits with virtual appointments.

Despite the numerous benefits of VHs for rural and remote health care, their implementation often faces significant challenges. These challenges can be categorized into 4 main themes: people, process, technology, and finance (Figure 3). Effectively addressing barriers in these areas requires a multifaceted approach, including digital literacy and enrichment, stakeholder collaboration, a patient-centered approach, sustainable funding structures, and the application of smart technologies. By implementing these strategies, regional and remote communities can gain better access to quality health care, potentially improving clinical and health system outcomes and enhancing satisfaction among rural patients and providers.

Addressing these multifaceted challenges is crucial for the successful implementation of VHs in regional and remote health care settings. The identified facilitators provide a road map for strategic interventions and investments to overcome barriers, thereby enhancing the effectiveness of VHs in improving health care accessibility and delivery in underserved areas.

XSL•FO
RenderX

Figure 3. Barriers to and facilitators of virtual hospital initiatives.



Recommendations for Developing VHs

The synthesis of studies on the development of VHs in regional and remote areas has led to 4 key recommendations. First, successful VH implementation relies heavily on stakeholder collaboration and patient-centered approaches. Engaging a diverse range of stakeholders, including community members and health care providers, ensures that VHs align with the needs and preferences of all involved. Key elements include collaborative engagement, health care professional advocacy, seamless communication policies, patient-centered design, and culturally responsive models. These strategies foster trust, improve coordination, and ensure inclusivity among and respect for diverse populations.

Second, achieving digital equity and securing sustainable funding are critical for the success of VHs. Implementing robust governance structures and advisory committees to oversee VH operations ensures coordination, safety, and quality standards. Policy-driven organizational transformation enables VHs to adapt to evolving health care needs while leveraging government grants to provide essential financial support, addressing infrastructure needs and operational deficits.

Third, investing in sustainable infrastructure and advancing health care technology are vital for enhancing VH services. Recommendations include developing resilient infrastructure to support VH technology, advancing integrated health care systems, empowering remote care staff with the necessary tools and training, building robust digital foundations through iterative development, and implementing stringent privacy and security measures to protect patient data.

Finally, enhancing digital literacy among health care providers and patients is crucial for the seamless implementation of VHs. Continuous training and involvement of health care providers build trust and competence in using new technologies, while educating patients about virtual care technologies improves acceptance and engagement.

Overall, successful VH implementation in regional and remote regions requires a holistic approach that considers stakeholder collaboration, a patient-centered approach, digital equity, smart technology application, and digital literacy enhancement. These recommendations provide a comprehensive framework for developing and sustaining VHs in regional and remote health care settings.

Limitations and Future Research Direction

This review encountered several limitations. First, the restricted search period, which concluded in March 2023, limited the capture of more recent developments and emerging insights. However, the unique focus on the remote and rural context offers timeless relevance to the baseline understanding of regional and remote VH implementation globally during the COVID-19 pandemic. This study provides a valuable foundation for future comparative research on VH implementation across rural and urban settings. Second, the absence of studies reporting negative outcomes restricts our understanding of the potential risks and drawbacks associated with the development of VHs for remote and regional communities. This lack of balanced reporting may lead to an overly optimistic view of VH implementation. In addition, as a rapid review, this study faced typical limitations, such as a constrained time frame and reliance on English language sources, potentially excluding relevant

XSL•FO
RenderX

studies and resulting in a less comprehensive literature search. While efficient, the rapid review methodology offers less depth compared to systematic reviews, potentially limiting the robustness of conclusions. Finally, this review insufficiently emphasized the broader environmental implications of VHs, particularly their demonstrated potential to reduce carbon emissions through decreased patient travel and resource consumption. This oversight limits the discussion of their potential as a tool for environmental sustainability. Finally, despite the authors' efforts to avoid using region-specific terms in the gray literature search, it is important to note that the use of the Google search engine in this study may have contributed to geographic bias, leading to an emphasis on Australian-focused gray literature sources.

Future research should focus on evaluating both positive and negative outcomes of VHs compared to inpatient care, using robust methodologies, such as randomized controlled trials or matched cohort studies. Such research would provide a comprehensive evaluation of the clinical and health system outcomes of VHs, facilitating direct comparisons to traditional inpatient care. Moreover, there is a pressing need for qualitative research involving interviews with key stakeholders engaged in successful VH projects. Analyzing these interviews would yield valuable insights into the challenges, successes, and lessons learned from specific projects, enabling the development of well-informed recommendations for future VH initiatives. In addition, research should investigate how implementing VH services in rural and remote areas could contribute to climate change mitigation through reduced health care–related travel and resource consumption.

Conclusions

Research into VHs is continuously evolving due to developments in regulations, digital advancements, communication tools, and remote monitoring capabilities. By bringing together information across different disciplines, including health care, technology, and policies, this evidence-based rapid review aids in effectively strategizing VH deployment so that the benefits can be realized by regional and remote communities.

This review suggests that VHs provide a promising solution to bridge health care gaps in regional and remote communities by enhancing clinical effectiveness, improving health system outcomes, and increasing patient and provider satisfaction. This review also supports the positive contribution of VHs on equitable health care access and addresses the existing evidence gap within the context of regional and rural health care.

To sum up, this rapid review highlights the barriers, mitigating facilitators, and recommendations that can serve as key strategic pillars to developing VHs so that clinical outcomes, health system performance, and satisfaction among patients and health care providers are enhanced within regional and remote communities.

Acknowledgments

The authors would like to thank Fiona Sturt and Umer Nisar for their guidance in directing this research. Their valuable time and advice made a huge contribution to this study, and the authors would like to thank them for this opportunity.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA checklist.
[DOCX File , 48 KB - humanfactors_v12i1e64582_app1.docx]

Multimedia Appendix 2
Search strategy.
[DOCX File , 25 KB - humanfactors_v12i1e64582_app2.docx]

Multimedia Appendix 3
Quality criteria checklist.
[DOCX File , 26 KB - humanfactors_v12i1e64582_app3.docx]

Multimedia Appendix 4
Data extraction table.
[XLSX File (Microsoft Excel File), 52 KB - humanfactors_v12i1e64582_app4.xlsx]

Multimedia Appendix 5
Overall concept matrix.
[XLSX File (Microsoft Excel File), 16 KB - humanfactors_v12i1e64582_app5.xlsx]

XSL•FO
RenderX

References

1. Dixit SK, Sambasivan M. A review of the Australian healthcare system: a policy perspective. SAGE Open Med 2018;6:2050312118769211 [FREE Full text] [doi: 10.1177/2050312118769211] [Medline: 29686869]
2. Rural and remote health. Australian Institute of Health and Welfare, Australian Government. URL: https://www.aihw.gov.au/reports/rural-remote-australians/rural-and-remote-health [accessed 2024-04-29]
3. Rechel B, Džakula A, Duran A, Fattore G, Edwards N, Grignon M, et al. Hospitals in rural or remote areas: an exploratory review of policies in 8 high-income countries. Health Policy 2016 Jul;120(7):758-769 [FREE Full text] [doi: 10.1016/j.healthpol.2016.05.011] [Medline: 27312144]
4. Thomas SL, Wakerman J, Humphreys JS. Ensuring equity of access to primary health care in rural and remote Australia - what core services should be locally available? Int J Equity Health 2015 Oct 29;14:111 [FREE Full text] [doi: 10.1186/s12939-015-0228-1] [Medline: 26510998]
5. About rural health. Centers for Disease Control and Prevention (CDC). URL: https://www.cdc.gov/ruralhealth/about.html#:~:text=Rural%20Americans%20are%20more%20likely,stroke%20than%20their%20urban%20counterparts [accessed 2024-05-29]
6. Ghafouri Fard M, Hasankhani HH. Virtual hospital: a new approach in education and treatment. J Med Edu Dev 2015;8(17):13 [FREE Full text]
7. Hesse BW, Ahern D, Ellison M, Aronoff-Spencer E, Vanderpool RC, Onyeije K, et al. Barn-raising on the digital frontier: the L.A.U.N.C.H. collaborative. J Appalach Health 2020;2(1):6-20 [FREE Full text] [doi: 10.13023/jah.0201.02] [Medline: 35769536]
8. Talevski J, Semciw AI, Boyd JH, Jessup RL, Miller SM, Hutton J, et al. From concept to reality: a comprehensive exploration into the development and evolution of a virtual emergency department. J Am Coll Emerg Physicians Open 2024 Aug;5(4):e13231 [FREE Full text] [doi: 10.1002/emp2.13231] [Medline: 39056087]
9. Katie A. How Ochsner Health's RPM pilot improved outcomes for patients with diabetes, hypertension. MedCity News. URL: https://medcitynews.com/2022/10/how-ochsner-healths-rpm-pilot-improved-outcomes-for-patients-with-diabetes-hypertension/?utm_source=chatgpt.com [accessed 2024-05-29]
10. Martinez KA, Deshpande A, Stanley E, Rothberg MB. Antibiotic prescribing for respiratory tract infections in urgent care: a comparison of in-person and virtual settings. Clin Infect Dis 2025 Feb 05;80(1):7-13. [doi: 10.1093/cid/ciae396] [Medline: 39078065]
11. WHO-ITU global standard for accessibility of telehealth services. World Health Organization. URL: https://iris.who.int/bitstream/handle/10665/356160/9789240050464-eng.pdf?sequence=1 [accessed 2024-05-29]
12. Digital health: a framework for healthcare transformation. Healthcare Information and Management Systems Society (HIMSS). URL: https://www.himss.org/sites/hde/files/media/file/2022/12/21/dhi-white-paper.pdf [accessed 2024-05-29]
13. Bidoli C, Pegoraro V, Dal Mas F, Bagnoli C, Bert F, Bonin M, et al. Virtual hospitals: the future of the healthcare system? An expert consensus. J Telemed Telecare 2025 Jan 24;31(1):121-133. [doi: 10.1177/1357633X231173006] [Medline: 37226478]
14. Gunsilius CZ, Heffner J, Bruinsma S, Corinha M, Cortinez M, Dalton H, et al. SOMAScience: a novel platform for multidimensional, longitudinal pain assessment. JMIR Mhealth Uhealth 2024 Jan 12;12:e47177 [FREE Full text] [doi: 10.2196/47177] [Medline: 38214952]
15. Digital health. World Health Organization. URL: https://www.who.int/health-topics/digital-health#tab=tab_1 [accessed 2024-05-29]
16. Catallo C, Chung-Lee L. How has COVID-19 changed the way we do virtual care? A scoping review protocol. Healthcare (Basel) 2022 Sep 23;10(10):1847 [FREE Full text] [doi: 10.3390/healthcare10101847] [Medline: 36292293]
17. Norman G, Bennett P, Vardy ER. Virtual wards: a rapid evidence synthesis and implications for the care of older people. Age Ageing 2023 Jan 08;52(1):afac319 [FREE Full text] [doi: 10.1093/ageing/afac319] [Medline: 36633298]
18. Babaei N, Zamanzadeh V, Valizadeh L, Lotfi M, Samad-Soltani T, Kousha A, et al. A scoping review of virtual care in the health system: infrastructures, barriers, and facilitators. Home Health Care Serv Q 2023 Jan 12;42(2):69-97. [doi: 10.1080/01621424.2023.2166888] [Medline: 36635987]
19. rpavirtual - a new way of caring. Sydney Local Health District. URL: https://www.health.nsw.gov.au/innovation/2020awards/Pages/rpavirtual.aspx [accessed 2024-05-29]
20. Duska A. Mayo Clinic launches advanced care at home model of care. Mayo Clinic. URL: https://newsnetwork.mayoclinic.org/discussion/mayo-clinic-launches-advanced-care-at-home-model-of-care/ [accessed 2024-05-29]
21. Virtual service expanding to relieve hospital pressures. Premier of Victoria. URL: https://www.premier.vic.gov.au/virtual-service-expanding-relieve-hospital-pressures [accessed 2024-05-29]
22. O'Kane G. Telehealth-Improving access for rural, regional and remote communities. Aust J Rural Health 2020 Aug 28;28(4):419-420 [FREE Full text] [doi: 10.1111/ajr.12663] [Medline: 32857466]
23. Australia's national digital health strategy. Australian Digital Health Agency. 2018. URL: https://tinyurl.com/z88492ec [accessed 2024-05-29]

XSL•FO
RenderX

24. Dhala A, Sasangohar F, Kash B, Ahmadi N, Masud F. Rapid implementation and innovative applications of a virtual intensive care unit during the COVID-19 pandemic: case study. J Med Internet Res 2020 Sep 03;22(9):e20143 [FREE Full text] [doi: 10.2196/20143] [Medline: 32795997]
25. Gajarawala SN, Pelkowski JN. Telehealth benefits and barriers. J Nurse Pract 2021 Feb;17(2):218-221 [FREE Full text] [doi: 10.1016/j.nurpra.2020.09.013] [Medline: 33106751]
26. Alison B. Rapid review of virtual care. Consumers Health Forum of Australia. 2021. URL: https://digitalhealthcrc.com/projects/reimagining-healthcare-in-australia/ [accessed 2024-05-29]
27. Armstrong CM, Wilck NR, Murphy J, Herout J, Cone WJ, Johnson AK, et al. Results and lessons learned when implementing virtual health resource centers to increase virtual care adoption during the COVID-19 pandemic. J Technol Behav Sci 2022;7(1):81-99 [FREE Full text] [doi: 10.1007/s41347-021-00227-1] [Medline: 34722860]
28. Richards D, Caldwell P. Improving health outcomes sooner rather than later via an interactive website and virtual specialist. IEEE J Biomed Health Inform 2018 Sep;22(5):1699-1706. [doi: 10.1109/jbhi.2017.2782210]
29. Summerfelt WT, Sulo S, Robinson A, Chess D, Catanzano K. Scalable hospital at home with virtual physician visits: pilot study. Am J Manag Care 2015 Oct;21(10):675-684 [FREE Full text] [Medline: 26633092]
30. Tricco AC, Langlois EV, Straus SE. Rapid reviews to strengthen health policy and systems: a practical guide. World Health Organization. 2017. URL: https://iris.who.int/bitstream/handle/10665/258698/9789241512763-eng.pdf [accessed 2024-05-29]
31. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
32. Huang X, Lin J, Demner-Fushman D. Evaluation of PICO as a knowledge representation for clinical questions. AMIA Annu Symp Proc 2006;2006:359-363 [FREE Full text] [Medline: 17238363]
33. Badampudi D, Wohlin C, Petersen K. Experiences from using snowballing and database searches in systematic literature studies. In: Proceedings of the 19th International Conference on Evaluation and Assessment in Software Engineering. 2015 Presented at: EASE '15; April 27-29, 2015; Nanjing, China p. 1-10 URL: https://dl.acm.org/doi/10.1145/2745802.2745818 [doi: 10.1145/2745802.2745818]
34. Covidence systematic review software. Veritas Health Innovation. URL: https://www.covidence.org/ [accessed 2024-05-29]
35. Duval D, Pearce-Smith N, Palmer JC, Sarfo-Annin JK, Rudd P, Clark R. Critical appraisal in rapid systematic reviews of COVID-19 studies: implementation of the Quality Criteria Checklist (QCC). Syst Rev 2023 Mar 27;12(1):55 [FREE Full text] [doi: 10.1186/s13643-023-02219-z] [Medline: 36973811]
36. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Med Res Methodol 2008 Jul 10;8:45 [FREE Full text] [doi: 10.1186/1471-2288-8-45] [Medline: 18616818]
37. Bradford NK, Caffery LJ, Smith AC. Telehealth services in rural and remote Australia: a systematic review of models of care and factors influencing success and sustainability. Rural Remote Health 2016;16(4):3808 [FREE Full text] [Medline: 27744708]
38. Davis SM, Jones A, Jaynes ME, Woodrum KN, Canaday M, Allen L, et al. Designing a multifaceted telehealth intervention for a rural population using a model for developing complex interventions in nursing. BMC Nurs 2020 Feb 04;19(1):9 [FREE Full text] [doi: 10.1186/s12912-020-0400-9] [Medline: 32042264]
39. DeHart D, King LB, Iachini AL, Browne T, Reitmeier M. Benefits and challenges of implementing telehealth in rural settings: a mixed-methods study of behavioral medicine providers. Health Soc Work 2022 Jan 31;47(1):7-18. [doi: 10.1093/hsw/hlab036] [Medline: 34910158]
40. Haque SN, DeStefano S, Banger A, Rutledge R, Romaire M. Factors influencing telehealth implementation and use in frontier critical access hospitals: qualitative study. JMIR Form Res 2021 May 05;5(5):e24118 [FREE Full text] [doi: 10.2196/24118] [Medline: 33949958]
41. Hirko KA, Kerver JM, Ford S, Szafranski C, Beckett J, Kitchen C, et al. Telehealth in response to the COVID-19 pandemic: implications for rural health disparities. J Am Med Inform Assoc 2020 Nov 01;27(11):1816-1818 [FREE Full text] [doi: 10.1093/jamia/ocaa156] [Medline: 32589735]
42. Howland M, Tennant M, Bowen DJ, Bauer AM, Fortney JC, Pyne JM, et al. Psychiatrist and psychologist experiences with telehealth and remote collaborative care in primary care: a qualitative study. J Rural Health 2021 Sep 06;37(4):780-787 [FREE Full text] [doi: 10.1111/jrh.12523] [Medline: 33022079]
43. Johnston K, Smith D, Preston R, Evans R, Carlisle K, Lengren J, et al. "From the technology came the idea": safe implementation and operation of a high quality teleradiology model increasing access to timely breast cancer assessment services for women in rural Australia. BMC Health Serv Res 2020 Nov 30;20(1):1103 [FREE Full text] [doi: 10.1186/s12913-020-05922-y] [Medline: 33256724]
44. Jong M, Mendez I, Jong R. Enhancing access to care in northern rural communities via telehealth. Int J Circumpolar Health 2019;78(2):1554174 [FREE Full text] [doi: 10.1080/22423982.2018.1554174] [Medline: 31066652]
45. Kocanda L, Fisher K, Brown LJ, May J, Rollo ME, Collins CE, et al. Informing telehealth service delivery for cardiovascular disease management: exploring the perceptions of rural health professionals. Aust Health Rev 2021 Mar 15;45(2):241-246. [doi: 10.1071/ah19231]

XSL•FO
RenderX

46. McPherson K, Nahon I. Telehealth and the provision of pelvic health physiotherapy in regional, rural and remote Australia. ANZ Cont J 2021 Aug;27(3):66-70. [doi: 10.33235/anzcj.27.3.66-70]
47. Nataliansyah MM, Merchant KA, Croker JA, Zhu X, Mohr NM, Marcin JP, et al. Managing innovation: a qualitative study on the implementation of telehealth services in rural emergency departments. BMC Health Serv Res 2022 Jul 02;22(1):852 [FREE Full text] [doi: 10.1186/s12913-022-08271-0] [Medline: 35780165]
48. Thomas LT, Lee CM, McClelland K, Nunis G, Robinson S, Norman R. Health workforce perceptions on telehealth augmentation opportunities. BMC Health Serv Res 2023 Feb 21;23(1):182 [FREE Full text] [doi: 10.1186/s12913-023-09174-4] [Medline: 36810089]
49. Correale MR, Soever LJ, Rampersaud YR. A model to implement standardized virtual care for low back pain amongst a large network of providers in urban and rural settings. J Prim Care Community Health 2022 Oct 27;13:21501319221130603 [FREE Full text] [doi: 10.1177/21501319221130603] [Medline: 36300425]
50. Giroux EE, Hagerty M, Shwed A, Pal N, Huynh N, Andersen T, et al. It's not one size fits all: a case for how equity-based knowledge translation can support rural and remote communities to optimize virtual health care. Rural Remote Health 2022 May;22(2):7252 [FREE Full text] [doi: 10.22605/RRH7252] [Medline: 35527567]
51. Head WT, Garcia D, Mukherjee R, Kahn S, Lesher A. Virtual visits for outpatient burn care during the COVID-19 pandemic. J Burn Care Res 2022 Mar 23;43(2):300-305. [doi: 10.1093/jbcr/irab202] [Medline: 34687201]
52. Kuperman EF, Linson EL, Klefstad K, Perry E, Glenn K. The virtual hospitalist: a single-site implementation bringing hospitalist coverage to critical access hospitals. J Hosp Med 2018 Nov 01;13(11):759-763. [doi: 10.12788/jhm.3061] [Medline: 30255859]
53. Sitammagari K, Murphy S, Kowalkowski M, Chou S, Sullivan M, Taylor S, et al. Insights from rapid deployment of a "virtual hospital" as standard care during the COVID-19 pandemic. Ann Intern Med 2021 Feb;174(2):192-199 [FREE Full text] [doi: 10.7326/M20-4076] [Medline: 33175567]
54. Vindrola-Padros C, Singh KE, Sidhu MS, Georghiou T, Sherlaw-Johnson C, Tomini SM, et al. Remote home monitoring (virtual wards) for confirmed or suspected COVID-19 patients: a rapid systematic review. EClinicalMedicine 2021 Jul;37:100965 [FREE Full text] [doi: 10.1016/j.eclinm.2021.100965] [Medline: 34179736]
55. AlDossary S, Martin-Khan MG, Bradford NK, Armfield NR, Smith AC. The development of a telemedicine planning framework based on needs assessment. J Med Syst 2017 May 20;41(5):74. [doi: 10.1007/s10916-017-0709-4] [Medline: 28321589]
56. Haleem A, Javaid M, Singh RP, Suman R. Telemedicine for healthcare: capabilities, features, barriers, and applications. Sens Int 2021;2:100117 [FREE Full text] [doi: 10.1016/j.sintl.2021.100117] [Medline: 34806053]
57. Gray K, Chapman W, Khan UR, Borda A, Budge M, Dutch M, et al. The rapid development of virtual care tools in response to COVID-19: case studies in three Australian health services. JMIR Form Res 2022 Apr 06;6(4):e32619 [FREE Full text] [doi: 10.2196/32619] [Medline: 35297765]
58. LeBlanc M, Petrie S, Paskaran S, Carson D, Peters P. Patient and provider perspectives on eHealth interventions in Canada and Australia: a scoping review. Rural Remote Health 2020 Sep;20(3):5754 [FREE Full text] [doi: 10.22605/RRH5754] [Medline: 32949485]
59. Allan J, Webster E, Chambers B, Nott S. "This is streets ahead of what we used to do": staff perceptions of virtual clinical pharmacy services in rural and remote Australian hospitals. BMC Health Serv Res 2021 Dec 04;21(1):1306 [FREE Full text] [doi: 10.1186/s12913-021-07328-w] [Medline: 34863164]
60. Virtual care in practice. Agency for Clinical Innovation. URL: https://aci.health.nsw.gov.au/__data/assets/pdf_file/0004/651208/virtual-care-in-practice.pdf [accessed 2024-05-29]
61. National strategic framework for rural and remote health. Department of Health, Australian Government. URL: https://www.health.gov.au/resources/publications/national-strategic-framework-for-rural-and-remote-health?language=en [accessed 2024-05-29]
62. Better connectivity plan for regional and rural Australia. Department of Infrastructure, Regional Development, Communications and the Arts, Australian Government. URL: https://www.infrastructure.gov.au/media-communications-arts/better-connectivity-plan-regional-and-rural-australia [accessed 2024-05-29]
63. Australia's primary health care 10 year plan 2022–2032. Department of Health, Australian Government. URL: https://www.health.gov.au/resources/publications/australias-primary-health-care-10-year-plan-2022-2032?language=en [accessed 2024-05-29]
64. Regional telecommunications review 2021. National Rural Health Alliance. URL: https://www.ruralhealth.org.au/policy/submission/regional-telecommunications-review-2021/ [accessed 2024-05-29]
65. NT health virtual care strategy. Northern Territory Government of Australia. URL: https://health.nt.gov.au/governance-strategies-committees/nt-health-strategies [accessed 2024-05-29]
66. Digital strategy for rural and remote healthcare - 10 year plan. Queensland Government. URL: https://www.health.qld.gov.au/__data/assets/pdf_file/0018/1141713/digital-strategy-for-rural-and-remote-healthcare.pdf [accessed 2024-05-29]
67. Hardy I, Rufus V, Jeyachandran A, Curran J. Improving equity of access in rural and regional health through hybrid and connected care. PwC. URL: https://www.pwc.com.au/health/improving-equity-of-access-through-hybrid-and-connected-care-in-rural-and-regional-health.html/1000 [accessed 2024-05-29]

XSL•FO
RenderX

68. Virtual care standard and guide. Victoria Department of Health. URL: https://www.health.vic.gov.au/virtual-care-standard-and-guide# [accessed 2024-05-29]
69. The virtual hospital. Hospital and Healthcare. URL: https://www.hospitalhealth.com.au/content/technology/article/the-virtual-hospital-234400308 [accessed 2024-05-29]
70. Microsoft News Center. Remote Australian community harnesses mixed reality and space technologies to deliver better healthcare. Microsoft Australia. URL: https://news.microsoft.com/en-au/features/remote-australian-community-harnesses-mixed-reality-and-space-technologies-to-deliver-better-healthcare/ [accessed 2024-05-29]
71. Shaw M, Wilson A. RPA virtual hospital proof of concept trial: evaluation report February 2020 to January 2021. rpavirtual. 2022. URL: https://ses.library.usyd.edu.au/bitstream/handle/2123/27474/90209%20RPA%20Virt%20Evaluation_A4_screen.pdf?sequence=3&isAllowed=y [accessed 2024-05-29]
72. Snoswell CL, Taylor ML, Comans TA, Smith AC, Gray LC, Caffery LJ. Determining if telehealth can reduce health system costs: scoping review. J Med Internet Res 2020 Oct 19;22(10):e17298 [FREE Full text] [doi: 10.2196/17298] [Medline: 33074157]
73. Denny CF, Hill U. P200 A virtual age? Evaluating the patient and healthcare worker perspective on virtual clinic delivery for patients with Cystic Fibrosis (CF) and non CF bronchiectasis (nCFB) at a specialist cardio-thoracic hospital. BMJ J 2022;77(Suppl 1):5 [FREE Full text] [doi: 10.1136/thorax-2022-btsabstracts.334]
74. King J, Poo SX, El-Sayed A, Kabir M, Hiner G, Olabinan O, GLINT Research Network, et al. Towards NHS Zero: greener gastroenterology and the impact of virtual clinics on carbon emissions and patient outcomes. A multisite, observational, cross-sectional study. Frontline Gastroenterol 2023 Nov 15;14(4):287-294 [FREE Full text] [doi: 10.1136/flgastro-2022-102215] [Medline: 37409339]
75. Thiel CL, Mehta N, Sejo CS, Qureshi L, Moyer M, Valentino V, et al. Telemedicine and the environment: life cycle environmental emissions from in-person and virtual clinic visits. NPJ Digit Med 2023 May 09;6(1):87 [FREE Full text] [doi: 10.1038/s41746-023-00818-7] [Medline: 37160996]

Abbreviations

LOS: length of stay
PICO: population, intervention, comparison, and outcome
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RPM: Remote Patient Monitoring
VH: virtual hospital

Edited by E Borycki; submitted 20.07.24; peer-reviewed by EP Greenup, E Opoku-Agyemang, J Hensel; comments to author 05.01.25; revised version received 30.01.25; accepted 17.04.25; published 03.06.25.

Please cite as:
Kumari AA, Wani TA, Liem M, Boyd J, Khan UR
Advancing Regional and Remote Health Care With Virtual Hospital Implementation: Rapid Review
JMIR Hum Factors 2025;12:e64582
URL: https://humanfactors.jmir.org/2025/1/e64582
doi:10.2196/64582
PMID:

©Artika Archana Kumari, Tafheem Ahmad Wani, Michael Liem, James Boyd, Urooj Raza Khan. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 03.06.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Toward Human-Centered Artificial Intelligence for Users' Digital Well-Being: Systematic Review, Synthesis, and Future Directions

Youngsoo Shin[1], PhD

Seidenberg School of Computer Science and Information Systems, Pace University, New York City, NY, United States

Corresponding Author:
Youngsoo Shin, PhD
Seidenberg School of Computer Science and Information Systems
Pace University
15 Beekman Street, Room 1120
New York City, NY, 10038
United States
Phone: 1 2123461888
Email: yshin@pace.edu

Abstract

Background: As information and communication technologies and artificial intelligence (AI) become deeply integrated into daily life, the focus on users' digital well-being has grown across academic and industrial fields. However, fragmented perspectives and approaches to digital well-being in AI-powered systems hinder a holistic understanding, leaving researchers and practitioners struggling to design truly human-centered AI systems.

Objective: This paper aims to address the fragmentation by synthesizing diverse perspectives and approaches to digital well-being through a systematic literature review. Using the stimulus-organism-response framework as a guiding lens, this study aims to develop a comprehensive model for designing human-centered AI systems that promote digital well-being.

Methods: A systematic review of 240 multidisciplinary publications was conducted to explore the intersection of AI, digital well-being, and human-centered design. The analysis involved identifying key themes, frameworks, and approaches, with the stimulus-organism-response model serving as an overarching perspective to organize findings and inform the model development.

Results: The review led to the development of a human-centered artificial intelligence model for digital well-being, a conceptual framework that consolidates current knowledge on designing AI systems to support digital well-being and positively influence human behavior. The proposed model integrates insights from cross-disciplinary research, providing a structured understanding of how AI system features (stimuli) affect users' internal states such as perceptions and emotions (organisms) and lead to attitudinal or behavioral changes (responses). Additionally, this paper highlights emerging challenges and opportunities, including ethical considerations, scalability, and practical guidelines for applying the model in long-term research and practice.

Conclusions: This study contributes to advancing the field by presenting an overarching framework for fostering digital well-being through human-centered AI systems. By addressing gaps in the fragmented literature and proposing a unifying model, the findings offer insights for researchers and practitioners. The human-centered artificial intelligence for digital well-being model serves as a foundation for future exploration and practical application in creating intelligent computing systems that improve users' digital well-being in everyday life.

(JMIR Hum Factors 2025;12:e69533) doi:10.2196/69533

KEYWORDS

digital well-being; digital health; human-centered AI; human-AI interaction; user experience; systematic literature review; artificial intelligence; mobile phone

Introduction

Digital well-being is a concept that describes the impact of information and communication technologies (ICTs) on people's physical and mental well-being in general [1-3]. Researchers and practitioners in many relevant fields, including computer science, health care, human-computer interaction (HCI), psychology, cognitive science, and data science, have emphasized that there are plenty of opportunities to leverage everyday ICTs in users' daily HCIs [4-14]. In this regard, daily intelligent computing systems have been developed and

XSL•FO
RenderX

investigated mainly with a focus on people's higher level of involvement in technology used to enhance their well-being and behavioral design [2,15].

Existing literature has provided diverse methods and approaches for building ICT-based computing systems that support users' daily quality of life by promoting users' positive attitudinal and behavioral changes [16]. For the past decade, the trend of increasing computing systems' complexity and interactivity has stimulated the introduction of a variety of artificial intelligence (AI) technologies (eg, chatbots, voice assistants, self-driving vehicles, or recommendation systems) that have penetrated far into people's everyday lives [17,18]. The application areas of these intelligent systems include (1) consumer applications such as recommender systems and social media platforms, (2) consequential applications in medical, legal, environmental, or financial systems, and (3) life-critical applications such as those in cars, airplanes, or military systems [19]. Furthermore, the previous COVID-19 pandemic has accelerated the use and adoption of these AI-powered intelligent systems in people's lifestyles [20]. For example, the social distancing restrictions in everyday life during the pandemic have increased the demand for AI-powered services and noncontact operations such as driverless delivery services, pre-entry wellness checking systems, or autonomous cleaning solutions [21,22].

However, several studies have pointed out negative aspects of prior approaches and perspectives that maximize users' engagement with these computing systems by attracting their attention and encouraging them to interact with the systems more frequently [1,23-26]. In this regard, large global technology companies, such as Google, Apple, Microsoft, and Meta, have reframed the meaning of digital well-being and have incorporated digital well-being into their business models. They have introduced various design and system features to help users manage their interactions with these computing systems and address these users' challenges in terms of being overloaded and distracted by the systems [27]. Many of these features have been developed as personal informatics tools (PITs) or digital self-control tools. With an emphasis on an instant and short-term well-being–enhancing intervention, these tools have mainly allowed users to monitor and reflect on their technology use with digital timers, usage dashboards, or lock-out features [16,25].

The concept of digital well-being has been interpreted in a broad sense among relevant fields with a strong consideration of daily HCI and recent human-AI interaction [28]. However, as aforementioned, several researchers and practitioners have also construed in a limited sense of the term by sharing a consensus that "excessive" and "frequent" use of systems could result in physical or mental health problems for users [11,29-34]. Despite this perspective gap, there is still a lack of studies that systematically reframe the concept of digital well-being with AI-powered computing systems. Further, the varying approaches developed in the related disciplines tend to focus on domain-, device-, or context-specific and fragmented aspects of human experience and interactions with computing systems, which can be limiting for researchers and practitioners as they search for a holistic understanding of digital well-being in daily HCIs and human-AI interactions [2,15,25,35,36].

In response, this paper sees the value in consolidating diverse but scattered initiatives regarding designing human-centered AI systems for users' digital well-being to support effective implementation by researchers and practitioners. More specifically, this paper addresses the research question as follows: "How can existing research on digital well-being with intelligent computing systems be reframed and synthesized toward an overarching model for human-centered AI?"

As part of the effort to develop such a cumulative body of knowledge, this paper conducts a multidisciplinary systematic literature review on intelligent computing systems for users' daily digital well-being with a consideration of user- and human-centered perspectives. The findings from the review serve as a conceptual foundation for the development of an overarching model: human-centered artificial intelligence for digital well-being (HCAI-DW) model. The results from this consolidation are also expected to enrich the current understanding of digital well-being in daily HCIs for further implementation of designing human-centered computing systems.

Methods

Theoretical Framework: The Stimulus-Organism-Response Model

To address the current knowledge gaps, this study developed an initial conceptual model of human-centered computing for digital well-being. In this process, this paper incorporated the stimulus-organism-response (SOR) model. According to the SOR model [37,38], an individual's responses to a stimulus are the result of a behavioral decision-making process that reflects interplay between internal aspects (eg, preference, personality, ability, motivation, etc) and external aspects of the specific context (eg, time, money, weather, etc). The SOR model states that either context or environment could be understood as a stimulus (S), which consists of a set of designed or nondesigned interactions that causes an internal cognitive process in individuals (O) and produces a response (R) [39,40].

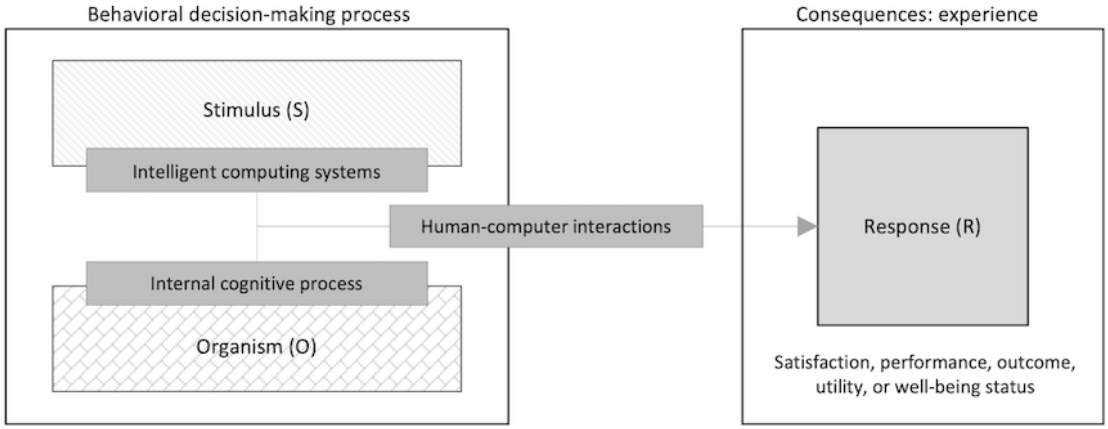
The SOR model has allowed researchers and practitioners in various fields to investigate how humans interact with nonhuman objects broadly and comprehensively [39]. In this regard, many have incorporated the model because of its broader range of possibilities for human behavior and objects and its structured way of focusing on behavioral mechanisms instead of behavioral consequences (eg, satisfaction, performance, outcome, utility, or increased well-being status). For this reason, this paper seeks to bridge the existing perspective gaps by using the basic SOR model, with complementary points of view for digital well-being and human-centered AI computing systems.

Through incorporating the SOR model as an initial viewpoint, it is expected that this study can investigate multilayered aspects of digital well-being with intelligent computing systems (Figure 1). Most previous perspectives on digital well-being in relevant fields mainly paid attention to the "consequences" of experiences, instead of focusing on the different ways people make behavioral decisions and their divergent experiential processes with intelligent systems and technologies. Researchers

XSL•FO
RenderX

and practitioners should consider behavioral decision-making mechanisms as well as a specific behavioral result targeted by designed interactions, including interface and system features. However, a systematic understanding of the multidimensional aspects of the behavioral decision-making process within the research boundary of design for digital well-being with intelligent computing systems is still lacking. Therefore, from the perspective of the SOR model, this study concentrates on individuals' behavioral decision-making mechanisms as a core concept for the unified conceptualization of digital well-being with intelligent computing systems.

Figure 1. Overarching viewpoint of digital well-being with intelligent computing systems.



Data Collection

Overview

This study uses a systematic literature review to gain a structured understanding of diverse perspectives on digital well-being within intelligent computing systems. Systematic literature reviews are recognized as rigorous and transparent methods for synthesizing existing evidence and addressing emerging interdisciplinary issues [41]. This review was guided by 2 complementary methodological frameworks: the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [42] (Multimedia Appendix 1) and the approach developed by Tranfield et al (2003) [43], which is widely adopted in evidence-informed management research. Following these frameworks, the review was conducted in four phases: (1) searching existing research, (2) screening relevant studies and approaches, (3) reviewing and evaluating focal studies on digital well-being with intelligent computing systems, and (4) conducting supplementary backward and forward searching.

In the initial phase, 4 primary databases—Web of Science, EBSCO, PubMed, and the Association for Computing Machinery (ACM) Digital Library—were used to identify relevant studies. The first publication search was conducted on August 15, 2022, with the final literature list updated on May 20, 2025, to ensure the inclusion of up-to-date studies.

The search strategy was designed by analyzing prior research and examining terminology used in the relevant scientific papers, which were critically reviewed in this paper's initial scoping review phase. For the searching process, the initial keyword string was designed to focus on the intersection of technology and well-being across multidisciplinary fields by constructing a query that combined multiple thematic dimensions:

- Core concepts: the query incorporated terms related to well-being, such as "wellbeing" and "well-being," to capture diverse terminology used across disciplines.
- Technological scope: keywords such as "system," "interface," "computer," and "technology" were included to ensure relevance to studies that involve digital and computational systems.
- Intelligence and interactivity: to refine the focus on intelligent and adaptive technologies, terms such as "artificial intelligence," "AI," "intelligent," "interactive," "smart," "autonomous," and "digital" were included.

Data Collection From the Web of Science

The Web of Science was chosen as the primary database for conducting a systematic literature search due to its comprehensive coverage of high-impact journals, conference proceedings, and other scholarly publications. The final query was adapted as follows:

TS=("wellbeing" OR "well-being") AND TS=("system" OR "interface" OR "computer" OR "technology") AND TS=("artificial intelligence" OR "AI" OR "intelligent" OR "interactive" OR "smart" OR "autonomous" OR "digital")

Boolean operators and wildcards were applied to broaden the search and accommodate variations in terminology. Advanced search filters were also used to focus on peer-reviewed publications, including journal papers, conference proceedings, and review papers. The systematic search yielded an initial pool of 5424 papers.

XSL•FO
RenderX

Data Collection From the EBSCO

The EBSCO database was used to conduct a comprehensive literature search, leveraging its extensive collection of resources spanning a wide array of academic disciplines. EBSCO was chosen for its strong interdisciplinary coverage, particularly in areas related to social sciences, computer science, and HCI, making it well-suited for exploring the digital aspects of well-being. The search string mirrored that of other databases:

("wellbeing" OR "well-being") AND ("system" OR "interface" OR "computer" OR "technology") AND ("artificial intelligence" OR "AI" OR "intelligent" OR "interactive" OR "smart" OR "autonomous" OR "digital")

The search was limited to peer-reviewed publications to ensure that only high-quality scholarly work was included. Additionally, in EBSCO, the search was restricted to the abstracts of papers to focus on studies that explicitly address the targeted themes. The initial query yielded a total of 6433 papers, encompassing a wide range of studies related to the HCI and digital well-being domains.

Data Collection From the PubMed

The PubMed database was chosen as a key resource for exploring studies at the intersection of digital well-being and intelligent systems due to its focus on digital health, biomedical research, and related fields. To align with the overarching research goals, the following search query was constructed to capture a wide range of relevant studies:

(("wellbeing" OR "well-being") AND ("system" OR "interface" OR "computer" OR "technology") AND ("artificial intelligence" OR "AI" OR "intelligent" OR "interactive" OR "smart" OR "autonomous" OR "digital"))

The initial query yielded 4073 papers. This output reflects the strength of PubMed in identifying multidisciplinary research that links intelligent systems with digital well-being, particularly within the contexts of health care and user interaction.

Data Collection From the ACM Digital Library

Finally, the ACM Digital Library was selected for its specialized focus on computing and information technology research, encompassing both theoretical advancements and practical insights from academia and industry. Given its prominence in the field of computing, the ACM Digital Library was anticipated to offer unique and highly relevant contributions to understanding the intersection of digital well-being and intelligent systems. The search was designed to explore themes of digital well-being, particularly in the context of HCI. To capture a broad range of studies, the following search query was used:

AllField:("wellbeing" OR "well-being" OR "digital wellbeing" OR "digital well-being")

To ensure high relevance and quality, 1 advanced filter was applied to limit the results to research papers published under the ACM SIGCHI (Special Interest Group on Computer-Human Interaction). SIGCHI represents a core community in HCI research, making it a critical source for studies at the nexus of well-being and intelligent systems. The initial search yielded a total of 5669 papers. This comprehensive set of results highlights the breadth of research contributions in the ACM Digital Library, particularly within the SIGCHI community, which has long been at the forefront of exploring the relationship between human-centered design (HCD), interaction, and user well-being.

Final Literature Sample and Data Exploration

From the initial search, a total of 21,599 papers were identified. These papers were manually screened by the researcher based on their titles, abstracts, and index keywords, resulting in a reduced set of 218 papers. The manual screening process was guided by 5 predefined inclusion criteria (Textbox 1), which ensured consistency and relevance in the selection process.

Textbox 1. The eligibility criteria for selecting the papers for the final review process.

- The paper should be related to the targeted research areas and topics of this study.
- The paper should be peer-reviewed and categorized as an academic paper.
- The paper should have well-defined research purposes to tackle the issue of design for digital well-being with intelligent computing systems in everyday contexts.
- The paper should point out the importance of investigating interactions between users and intelligent computing systems in their everyday behavioral decision-making contexts.
- The paper should contain practical aspects to create or evaluate the designed interactions for supporting users' digital well-being.

Subsequently, the remaining papers were reviewed in full. During this phase, backward and forward citation searches were conducted, yielding 12 additional papers. To further enhance the comprehensiveness and objectivity of the review, 10 relevant papers were also identified through consultations with external researchers and practitioners with expertise in HCI. These external consultations provided valuable third-party perspectives and helped mitigate potential bias in paper selection. After the review process, a total of 240 papers were selected for the final literature sample.

After the data collection and screening process, an inductive qualitative coding approach was conducted to systematically analyze the full text of each selected publication. The coding process was guided by the SOR model, which served as a conceptual lens to identify and organize how each paper addressed the roles of stimulus (ie, technology or system-level factors), organism (ie, user-level individual and contextual factors), and response (ie, behavioral outcomes or digital well-being effects). A coding scheme was iteratively developed and refined during this process to capture the range of constructs

XSL•FO
RenderX

and considerations described in the literature. Particularly, each paper was coded based on its entire content, allowing for the extraction of both explicit concepts and latent themes. The codes derived from this process were then used in an affinity diagramming procedure, wherein related codes were grouped and clustered to surface patterns across the corpus. This bottom-up synthesis approach enabled the researcher to identify conceptual similarities, contradictions, and gaps across the diverse literature on AI systems and digital well-being.

Through the affinity diagramming process [44], codes organically clustered into 3 major thematic areas. The first cluster focused on implementation domains of intelligent computing systems, capturing how technologies were designed to support overall well-being, facilitate behavior or habit change, or serve as behavioral intervention tools. The second cluster highlighted conceptualizations of human autonomy, including autonomy as supported through technology and autonomy from human control. The third cluster reflected each paper's approaches to system design and evaluation, with codes centering on the need to incorporate contextual adaptation (stimulus), individual differences (organism), and outcome-specific tailoring (response) into the development and assessment of AI-driven digital well-being interventions.

These 3 clusters, developed through iterative coding and visual grouping via affinity diagramming, provided the conceptual foundation for a model of human-centered AI computing systems for digital well-being presented in this study. All selected papers and their coding outcomes are described in the attached Multimedia Appendix 2, which is a review of 240 publications from 2008 to 2024.

Results

Sample Data Description

The review of 240 publications between 2008 and 2024 reveals an evolving landscape in which digital well-being and intelligent computing systems have become central themes of inquiry across multiple disciplines (Figure 2). The temporal distribution of studies demonstrates a marked increase in scholarly attention, with the number of relevant publications growing from just 2 papers in 2008 to 43 in 2024. This acceleration reflects the growing urgency—both academic and societal—to understand how intelligent systems can be designed and evaluated to support human well-being in digital contexts. Figure 3 illustrates the data collection process.

During the initial period from 2008 to 2016, research primarily focused on early digital interfaces and the psychological impact of technologies such as smartphones, social media, and ubiquitous computing platforms. These studies were largely rooted in exploratory research, with conceptual frameworks drawn from subjective well-being and affective computing. The dominant focus was on describing emotional and behavioral responses to digital engagement, often framed in terms of stress, distraction, or mood regulation. The technological systems under study were mostly nonadaptive and nonintelligent, and user autonomy was often treated as a passive construct.

An inflection point emerged in 2017. This period marked a transition from exploratory inquiries into more normative and design-oriented investigations. Triggered by public discourse around technology addiction, ethical technology, and digital detox movements, scholars began reframing digital well-being as a design goal rather than merely a behavioral outcome. This era coincided with a surge in industry initiatives (eg, Apple's Screen Time and Google's Digital Wellbeing suite), and research began to reflect growing societal concern about how technologies could undermine or support flourishing. Tools such as digital self-control tools and PITs became common subjects of investigation, signaling a broader shift toward proactive design strategies for well-being.

From 2022 onward, there was a notable rise in publications proposing frameworks, systems, and user studies that examined the role of human-centered AI in supporting digital well-being. Interdisciplinary contributions expanded, including papers from HCI, health informatics, behavioral science, and even management science. Concepts such as autonomy, persuasive design, ethical nudging, and emotion-aware interfaces were increasingly investigated through user-centered lenses. Many of these studies moved from simple heuristics to formal models, measurement tools, and empirical evaluations of intervention efficacy.

By 2024, the scope of the literature had broadened further to include the intersection of digital well-being with AI. This more recent body of work introduced sophisticated technologies, including generative AI, adaptive learning systems, hybrid human-AI collaboration platforms, and emotion recognition interfaces. These systems increasingly relied on real-time user data to personalize feedback and optimize interventions.

Overall, the sample data reinforces the need for design frameworks that synthesize empirical insights, respect user autonomy, and adapt to contextual and personal differences. As explored in the following findings and the model development section, the coding analysis helped identify key pain points and opportunities across system implementation, user experience (UX), and evaluation design.

XSL•FO
RenderX

Figure 2. Publication years of the selected papers on digital well-being and intelligent computing systems.

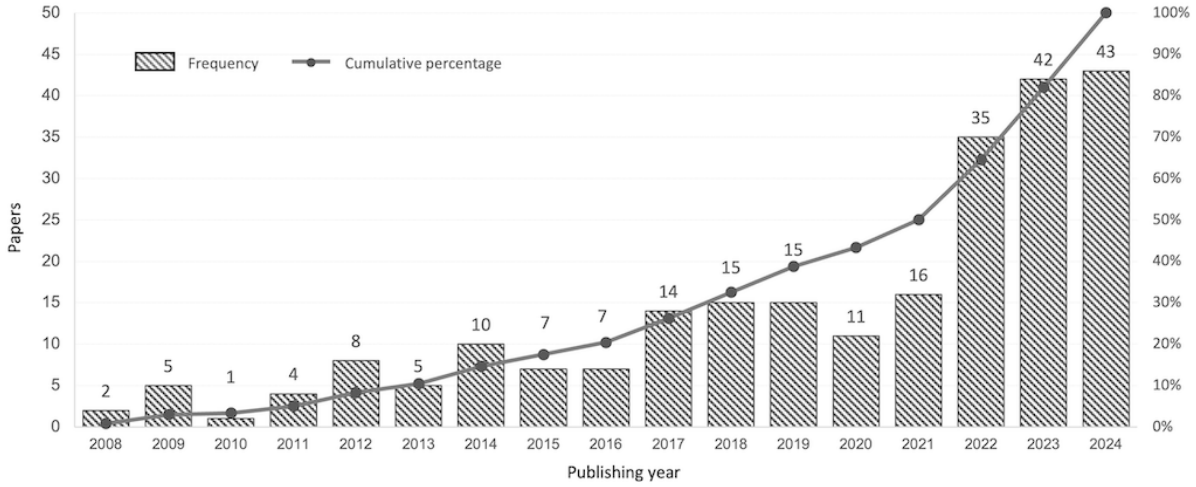
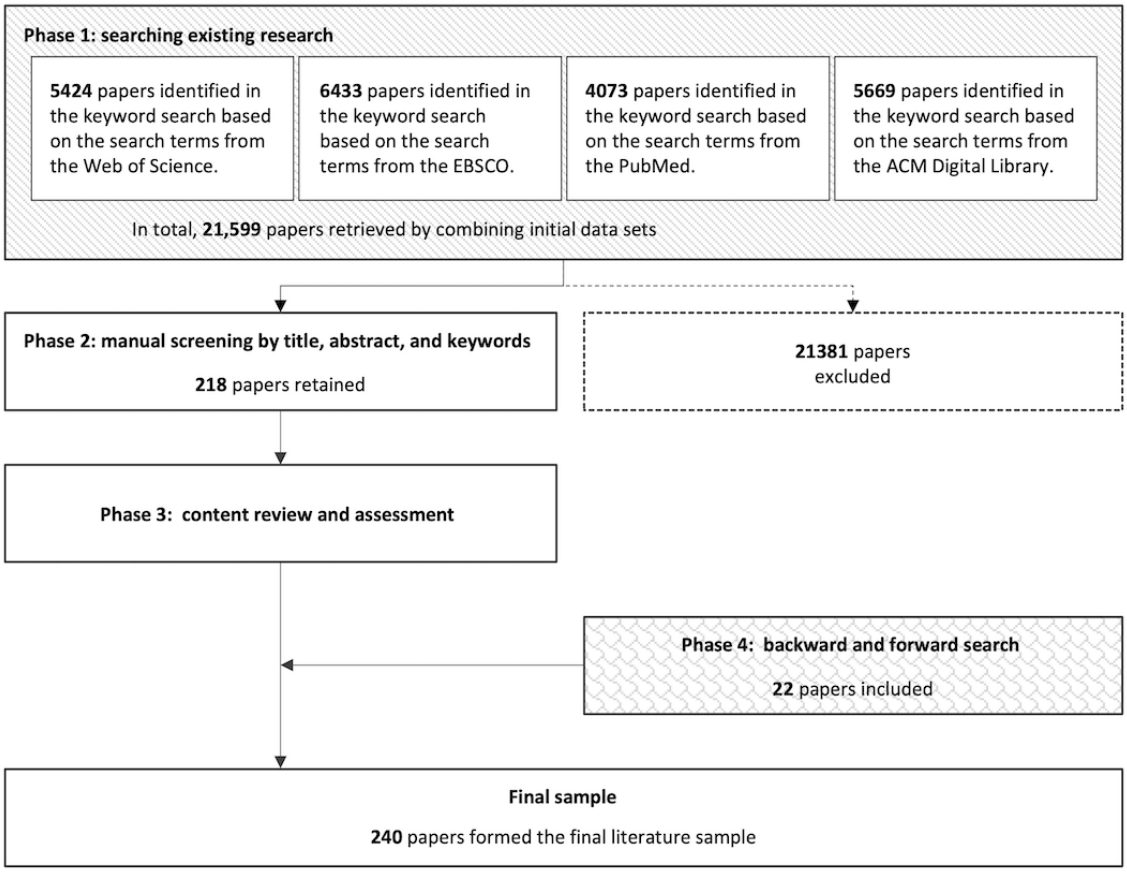


Figure 3. Overview of the data collection process. Databases: Web of Science, EBSCO, PubMed, and the Association for Computing Machinery (ACM) Digital Library.



Emerging Perspectives on Digital Well-Being and Intelligent Systems

Overview

From the in-depth analysis with the affinity diagramming process, 3 emerging perspectives were clustered, which can be incorporated into the development of a model of human-centered AI computing systems for digital well-being.

Finding 1. Three Implementation Areas of Intelligent Computing Systems for Digital Well-Being

As summarized in Table 1, the dataset revealed 3 major implementation areas—daily well-being in HCIs, design for behavioral change, and behavioral intervention technologies (BITs)—that form the basis for findings 1-3 in this study.

XSL•FO
RenderX

Table 1. Summary of 3 implementation areas in intelligent computing systems for digital well-being.

Implementation area	Description	Papers, n	Common theories or approaches	Key limitations
Daily well-being in HCIs[a]	Broad framing of well-being as a design goal, often informed by psychological theories	114	Self-determination theory and hedonic experience	Conceptual vagueness and limited operationalization
Design for behavioral change	Focus on changing user behaviors or habits via persuasive or adaptive systems	71	Dual-process theory, habit theory, and goal-setting theory	Lack of long-term efficacy and inconsistent success metrics
Behavioral intervention technology	Systematic approaches rooted in clinical settings using adaptive, data-driven interventions	55	BITs[b] and personal informatics	Limited practical integration and focus on high-income populations

[a]HCI: human-computer interaction.

[b]BIT: behavioral intervention technology.

Daily Well-Being in HCI

First, well-being in daily interactions with computing systems has been studied through a variety of lenses (ie, digital health, technology for mental health, or digital medicine) and under diverse human behavioral contexts over the past 2 decades. Among these various perspectives, "well-being in HCIs" is the broadest for considering technology as a tool for human flourishing and well-being [3]. Most related works aim to improve a user's psychological and subjective well-being by developing interactive design and technology interventions along 2 pathways. The first pathway involves formulating system design principles by using relevant theories from psychology, such as self-determination theory and hedonic experience [3,45,46]. The second pathway involves investigating certain well-being determinants derived from these theories—such as autonomy, competence, and relatedness—which typically serve as both intentions for computing system design and as criteria for evaluations in technology development processes [3].

The review of 240 papers confirmed that this conceptualization is foundational: 114 (48%) studies in the dataset treat well-being as a core design or evaluative objective. Many of these adopt an HCI-centric viewpoint, emphasizing positive emotional experiences, mental health, or life satisfaction as outcomes of well-designed user interactions. However, despite shared theoretical anchors—especially self-determination theory—there remains significant inconsistency in how well-being is defined, operationalized, and measured across studies. Most of the papers in this category relied on self-report measures (eg, user satisfaction surveys and qualitative feedback) to evaluate well-being outcomes. These findings reveal a methodological gap between theoretical aspiration and empirical implementation.

Previous studies have attempted to suggest tools for system design and development practices by grounding them in rigorous and evidence-based empirical research. However, the major criticism of the well-being perspective in HCIs has been the fragmented and unstructured nature of its implementation [3,47]. Furthermore, the notion of well-being is often used as a descriptive umbrella for current lifestyle trends or invoked conceptually to justify a human-centric design ethos, especially in discussions focused on the outcomes of system design [32]. As a result, many studies frame well-being as an aspirational goal but lack the definitional clarity or structured methodology required for robust evaluation.

Still, the literature reflects an evolution in the treatment of well-being. In earlier studies—particularly those published before 2016—the dominant approach was reactive: minimizing digital harms (eg, mitigating distraction, reducing screen time, or addressing smartphone addiction). More recent contributions, however, embrace a proactive stance. Instead of merely limiting technology's negative impacts, intelligent systems are now increasingly designed to scaffold reflection, self-regulation, emotional resilience, and long-term behavior change.

Despite these promising developments, the field still lacks integrative models that bridge disciplinary boundaries and connect technological features to measurable well-being determinants. Moreover, the challenge of balancing persuasive design with respect for user autonomy remains unresolved, as detailed in finding 2. While the well-being domain continues to expand in scope and influence, future research must prioritize the development of validated, multidimensional frameworks that align well-being theory with empirical evidence and real-world system design practices.

Design for Behavioral Changes

Second, the topic of design for behavioral change has been explored from a more defined and targeted angle by focusing on both short-term and long-term changes in users' behaviors [48]. In the context of digital well-being, designing for behavior change often involves addressing complex, subjective aspects of human experiences—elements that are inherently difficult to operationalize and quantify [49]. To address these challenges, researchers and practitioners have drawn extensively from behavioral science, using empirically validated frameworks that explain how human attitudes, cognitive patterns, and routines can be shaped through deliberate system design [28]. These behavioral foundations have been used to develop intelligent systems that aim to promote health-related behaviors and improve users' overall well-being.

The systematic review of 240 papers identified 71 (30%) studies that explicitly frame their system design objectives around catalyzing behavioral or habitual changes through intelligent computing. These studies span a wide range of behavioral domains, including digital overuse, emotional regulation, sleep

XSL•FO
RenderX

hygiene, and physical activity. What distinguishes this subset is its methodological rigor and theoretical alignment. Most of these studies ground their design logic in well-established behavioral science models, such as dual-process theory, habit formation theory, and implementation intentions. These theories provide a robust framework for creating interventions that influence users' decision-making processes across varying contexts and levels of awareness.

More recently, research and practice in this domain have shifted toward exploring the long-term consequences of technology-mediated behavior and its cumulative impact on well-being. For example, several studies [50,51] have sought to move beyond optimizing discrete moments of interaction (eg, interface usability and prompt timing) to designing systems that encourage enduring lifestyle changes. As a result, the field is beginning to transition from a narrow focus on behavior change to a more expansive framing around habit change. The review indicates a growing number of studies—particularly those integrating machine learning models—that aim to detect recurring behavior patterns and trigger interventions based on behavior-specific cues [52]. These approaches reflect a deeper ambition: to influence not only immediate choices but also the broader cognitive scaffolding that sustains long-term habits.

The growing interest in habit change has been accompanied by a series of conceptual and practical challenges. Niedderer et al [53] provided a macrolevel overview of behavior change design practices and noted major gaps in shared understanding, terminology, and implementation standards across sectors. This fragmentation is especially evident when comparing health-based interventions with HCI-focused strategies. For instance, while health researchers often emphasize measurable outcomes such as physical activity levels or sleep duration, HCI researchers tend to focus on experiential metrics such as emotional resonance or usability satisfaction. This inconsistency complicates the evaluation of intervention efficacy and limits the generalizability of findings.

Pinder et al [54] similarly noted the heterogeneity in how the concept of habit change is interpreted and operationalized across disciplines. In response, they proposed the "habit alternation model," which integrates 3 dominant theories—dual-process theory, modern habit theory, and goal-setting theory—into a single explanatory framework. This model has proven influential in reconciling divergent approaches and in guiding the design of more cohesive and context-sensitive interventions.

In parallel, a growing number of behavioral scientists have emphasized the importance of integrating HCD into behavior change systems [55]. HCD techniques—particularly participatory design methods—are increasingly used to ensure that interventions align with users' lived experiences, values, and goals. By involving users directly in the design and refinement process, these approaches seek to address the limitations of top-down, expert-driven interventions that may overlook individual motivation, digital literacy, or environmental constraints.

However, despite the growing sophistication of these approaches, the literature reveals several persistent issues. Chief among them is the challenge of sustaining user engagement over the long term and producing consistent results across diverse user populations [44,56,57]. While personalization features—such as content adaptation, feedback customization, and dynamic goal-setting—are widely used, the findings indicate that tailoring alone is insufficient. Systems that fail to account for users' deeper cognitive and emotional processes—such as internal motivation, self-regulation, and reflective goal alignment—often achieve only short-term success before engagement tapers off.

Another persistent concern is the lack of standardization in how success is measured. Some studies report efficacies in terms of quantitative behavior change (eg, reduced screen time and increased steps per day), while others emphasize qualitative improvements (eg, enhanced focus, better mood, or increased self-awareness). This lack of convergence around evaluative criteria undermines the comparability of results and highlights the need for more unified, theory-driven frameworks for assessing intervention outcomes.

In this regard, the literature points toward the need for deeper integration of behavioral theory with longitudinal user evaluation. The most promising studies combine short-term intervention efficacy with sustained reflection and feedback loops that help users internalize new behaviors. For example, systems that facilitate reflective journaling, goal adjustment, or feedback calibration were more likely to show lasting impact. These findings underscore the value of designing for not only behavioral compliance but also behavioral understanding and transformation.

In summary, the design for behavioral changes in the context of intelligent systems represents a maturing area of research—one that is increasingly grounded in theory, enriched by participatory methods, and informed by both real-time data and long-term evaluation. While notable progress has been made in creating adaptive, user-centered systems, the next frontier will require more holistic approaches that synthesize behavioral theory, co-design practices, and robust measurement strategies to enable durable and meaningful improvements in users' digital well-being.

About BIT

Third, the recent rise of BITs marks a significant and structured subdomain within the broader discourse of digital well-being, especially in health care and clinical settings. Defined by Mohr et al (2014) [58] as "the application of behavioral and psychological intervention strategies by using technological features to target behavioral, cognitive, and affective context and environment that support physical, behavioral, and mental health," BITs have gained traction through the development of computing systems designed to support health-related behavior change. This growth has enabled behavioral scientists to generate new empirical insights on how digital tools can promote physical, emotional, and psychological well-being [36,58,59].

The review of 240 papers revealed that 55 (23%) studies explicitly focus on BIT development and evaluation, making it one of the most clinically grounded areas in the literature. These systems typically emerge from interdisciplinary collaborations among behavioral scientists, HCI researchers, and system

XSL•FO
RenderX

developers, often underpinned by robust theoretical foundations. Common targets of BITs include mental health symptom reduction [60], emotional regulation, and adherence to behavioral protocols.

In recent years, BITs have evolved to incorporate technologies categorized under PITs and digital medicine. Several systems in this group exemplify the practical application of such theories through adaptive interfaces that personalize nudges, prompts, or feedback mechanisms in real time. These include systems that leverage decision science and data science to personalize interventions based on user behavior and preference data [24,61,62]. For instance, the deployment of Just-In-Time Adaptive Interventions (JITAIs)—where prompts are delivered based on biometric signals, app usage data, or behavioral cues—has become a dominant trend [63-66]. These JITAI-based systems aim to maximize user engagement by intervening at moments of high receptivity, demonstrating significant promise in both user satisfaction and short-term efficacy.

Despite their strengths, BITs continue to face several limitations. Traditionally, these technologies have been developed and evaluated within clinical contexts, emphasizing short-term efficacy over broader UX or long-term impact [67]. Researchers and practitioners have highlighted several persistent issues: (1) limited theoretical grounding in some system designs, (2) insufficient models for evaluating user interaction and engagement, (3) accessibility challenges and associated costs, (4) inconsistent user adherence, and (5) poor integration into broader health delivery infrastructures [48].

Moreover, there is a notable lack of diversity in the populations studied. Most BITs have been designed and tested in high-income settings, raising concerns about their generalizability and inclusivity. Addressing this gap, recent work has begun to explore the use of generative AI and low-cost sensing technologies to extend the reach of BITs in underserved or resource-constrained environments.

In response to these limitations, behavioral scientists have called for a shift toward more ethically grounded and human-centered BIT development. This includes integrating fragmented design perspectives into cohesive theoretical frameworks and emphasizing transparency, adaptability, and co-design practices [68]. Features such as explainable AI, user override functions, and participatory evaluation have been proposed to mitigate ethical concerns related to manipulation, autonomy, and privacy.

Synthesizing insights from BITs and broader literature on intelligent computing systems for digital well-being, 3 key observations emerge. First, much of the prior research has focused on the outcomes of behavioral decision-making without adequately considering the diversity of users. There is now increasing emphasis on individual differences—such as psychological profiles, health status, and behavioral tendencies—to inform more effective and equitable system design [53,69]. Second, while earlier work emphasized generic system usability, recent studies highlight the importance of well-structured, personalized interventions that account for contextual variables, such as environment and time of use [36,59,62,70]. Third, as AI technologies become more pervasive, researchers and practitioners are increasingly grappling with how to reconcile advancements in human autonomy research with digital well-being initiatives—particularly concerning personalized UXs and human-centered AI system design [3,67,68].

Finding 2. Conceptual Diversity on Human Autonomy With AI-Powered Intelligent Systems

The review results revealed that as AI-powered computing systems become increasingly integrated into our daily life—supporting decision-making, automating routine tasks, and influencing behavioral patterns—the concept of autonomy has emerged as a central design value [19,27,71]. However, the interpretation of autonomy varies widely across disciplines, leading to conceptual ambiguity and design inconsistencies. From the perspective of psychology and behavioral science, the concept of human autonomy is considered one of the basic psychological needs, particularly within the framework of self-determination theory, which views autonomy as freedom from external control or influence—including from other people and environmental factors [72]. From this perspective, psychologists and behavioral scientists have focused on behavioral interventions to change people's attitudes and behaviors in positive ways by investigating the relationship between human autonomy and the behavioral decision-making process [47].

The analysis of 240 papers found that 147 (61%) studies engaged directly with autonomy-related issues, with 2 dominant and sometimes competing perspectives emerging: (1) technological autonomy from human control and (2) human autonomy through technology.

The first perspective, technological autonomy from human control, is most prevalent in the fields of computer science and engineering. In this framing, autonomy refers primarily to the capabilities of AI systems themselves—their ability to function independently, adapt dynamically to changing conditions, and minimize the need for continuous human input [27,73]. This notion is often associated with intelligent automation, self-regulating algorithms, and predictive systems that optimize for performance and reduce cognitive load for users. Forty-six (19%) studies aligned with this perspective, emphasizing the design of systems that learn and act autonomously, such as context-aware recommendation engines, automated productivity tools, and self-adjusting health monitors. However, while this approach offers efficiency and scalability, it also raises concerns about diminishing user control, consent, and interpretability. Systems designed for technological autonomy may inadvertently undermine user agency by prioritizing automation over transparency or flexibility—especially when users are not able to understand, override, or adapt the system's decisions.

In contrast, the second and more human-centered perspective, human autonomy through technology, is emphasized in the disciplines of management, HCI, and HCD. This framing views technology not as a substitute for human agency but as an enabler of it—providing tools that amplify users' capacity to make intentional, informed, and self-endorsed decisions [61,74]. In total, 102 (43%) studies in the dataset foregrounded this approach, suggesting that AI systems should support user autonomy by offering decision-support dashboards, intelligent

XSL•FO
RenderX

filtering mechanisms, PITs, and customizable experiences [27,75,76]. In this view, autonomy is not merely preserved in the face of automation but is actively cultivated through interactions that respect the user's goals, values, and contexts. Systems aligned with this perspective often use noncoercive prompts, user-configurable settings, and ethical nudging strategies to support reflection and volition rather than compliance or behavioral control [61,74]. This divergence in perspectives is summarized in Table 2.

Table 2. Conceptualizations of human autonomy in AI[a] systems.

Autonomy perspective	Disciplinary origin	Papers, n	Core ideas	Key example use cases
Technological autonomy from human control	Computer science and engineering	46	AI systems operate independently, reducing user input	Predictive automation and context-aware recommender systems
Human autonomy through technology	Management, HCI[b], and HCD[c]	102	AI enhances user capacity and supports informed and volitional action	Personal informatics, decision-support interfaces, and nudging tools

[a]AI: artificial intelligence.

[b]HCI: human-computer interaction.

[c]HCD: human-centered design.

Despite widespread efforts to embed autonomy-supportive features into intelligent systems, many researchers and practitioners continue to struggle with translating these abstract values into coherent design strategies [17,71]. The conceptual gap between the system-centered and user-centered views of autonomy contributes to ongoing tensions in system development, particularly when trying to balance the benefits of automation with the preservation of user agency.

A recurring issue highlighted in the literature is the persuasion-autonomy dilemma, where AI systems intended to assist or influence behavior (eg, through nudges or reminders) risk being perceived as overly controlling, opaque, or misaligned with users' intentions. In practice, users often appreciate automation that reduces effort but become frustrated or resistant when systems overstep—such as enforcing goals that feel inappropriate, invasive, or difficult to override. For example, the review identified cases where health-monitoring applications encouraged physical activity without sensitivity to context (eg, illness), or algorithmic management systems in gig work environments reduced perceived autonomy and flexibility—prompting resistance among users [41,77].

These examples illustrate that perceived fairness, transparency, and reversibility often matter more to users than algorithmic sophistication alone. The review underscores that user trust and acceptance are closely linked to how much control and interpretability a system allows. In response to these concerns, a promising development is the emergence of "repairable AI"—systems that embed mechanisms for user revision, feedback, or override. These systems transform autonomy from a static feature into a relational and dynamic property, allowing users to negotiate control over time.

For instance, the TypeOut system by Xu et al [78], implemented self-affirmation prompts that allowed users to reflect on their values and motivations when trying to reduce smartphone overuse. This approach supported users' autonomy by reinforcing self-directed action rather than imposing strict limitations. Similarly, several studies explored hybrid intelligence systems that combine algorithmic recommendations with human coaching or collaborative decision-making. These systems demonstrated higher user satisfaction, engagement, and personalization, as they allowed users to maintain interpretive authority while benefiting from intelligent assistance.

Across the dataset, the most effective systems for supporting autonomy shared several characteristics: they offered clear explanations of system behavior, allowed users to customize settings or feedback modalities, and supported decision-making rather than unilateral automation. Autonomy was thus conceptualized not as a binary variable but as a continuum that must be calibrated to the user's needs, task context, and preferences.

In summary, the review suggests that autonomy in AI-powered systems should be approached as a relational, context-sensitive construct shaped through the dynamic interplay of system design, user interaction, and ethical affordances. While technological autonomy from human control may offer performance benefits, the long-term sustainability of human-centered AI depends on the extent to which these systems preserve and enhance human autonomy through technology. Future research and development should prioritize design practices that support transparency, reversibility, and participatory engagement, ensuring that users retain meaningful agency in an increasingly intelligent world.

Finding 3. Developing and Evaluating Process for Digital Well-Being With AI Systems

As AI systems become increasingly embedded in everyday life, the challenge of ensuring that these technologies genuinely support digital well-being has taken on new urgency. Existing initiatives have repeatedly highlighted a persistent gap between high-level theoretical ideals and the actionable methodologies required to guide the development and evaluation of human-centered AI systems. Two practical and recurring questions have emerged from these efforts: (1) How can researchers and practitioners identify applicable strategies for designing AI systems that enhance digital well-being? (2) How can they rigorously evaluate whether these systems are effective in achieving that goal?

Despite a growing body of literature advocating for human-centered values in AI, the systematic review of 240 papers revealed that few studies translate these values into

XSL•FO
RenderX

structured, repeatable design and assessment strategies. Instead, the current landscape remains fragmented—often disconnected from both theoretical frameworks and real-world user needs. In response to this problem, this paper synthesizes prior work through the lens of the SOR model, which offers a practical and conceptual bridge between research, design, and evaluation in the context of digital well-being.

As summarized in Table 3, the SOR-based framework articulates three key considerations for guiding human-centered AI system development: (1) using contextual differences for the stimulus component, (2) accounting for individual differences in the organism component, and (3) specifying consequential differences in the response and outcome components.

Table 3. Methodological focus in developing AI[a] systems for digital well-being.

SOR[b] component	Key research concerns	Papers, n	Representative methods	Challenges identified
Contextual difference (stimulus)	Leveraging real-time sensing and situational awareness	94	Wearables and ambient sensing	Reliability of context-detection
Individual differences (organism)	How to adapt systems to personal traits and contexts	79	Profiling, personalization, and HCD[c]	Often, surface-level traits are used only
Consequential difference (response and outcome)	Long-term changes in well-being and reflection processes	140	Multimodal evaluation and journaling tools	Rarely includes holistic well-being metrics

[a]AI: artificial intelligence.

[b]SOR: stimulus-organism-response.

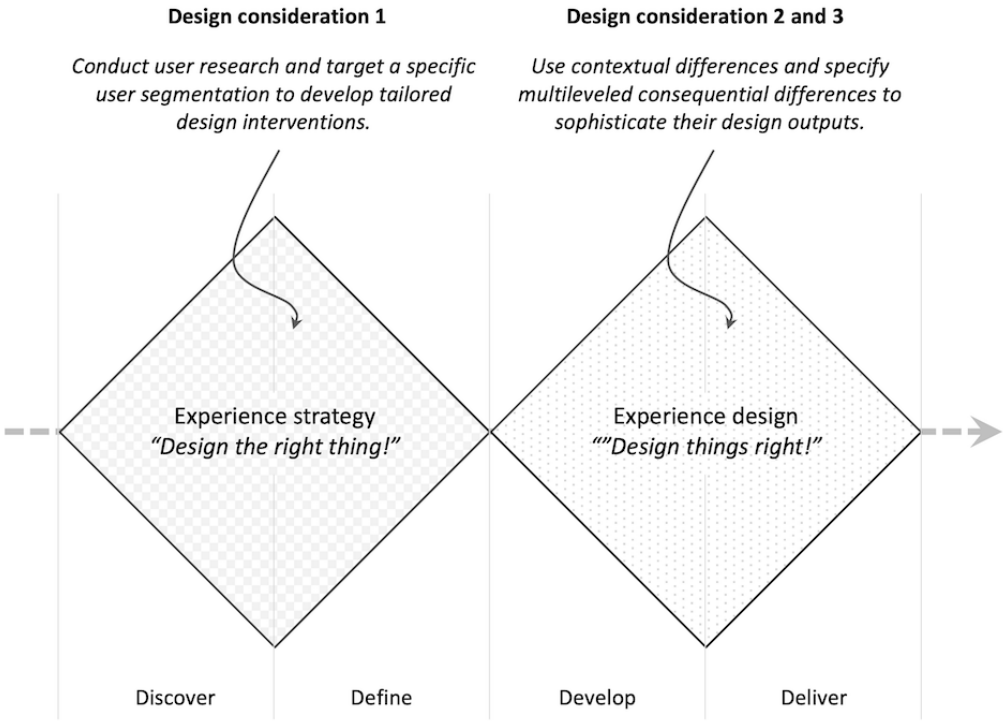
[c]HCD: human-centered design.

The SOR framework was well-reflected across the dataset. Ninety-four (39%) papers focused on the integration of contextual data into system design, leveraging information from wearable devices, mobile sensors, and behavioral logs to deliver personalized and context-aware interventions. These findings align with growing interest in BITs and JITAIs that respond dynamically to real-world environments and user states. Additionally, 79 (33%) papers emphasized the importance of addressing individual variability in AI systems for digital well-being. These studies adopted approaches such as adaptive personalization, participatory design, and segmentation based on psychological traits, user preferences, or digital literacy levels. Systems that integrated this type of individual profiling were consistently associated with higher levels of user engagement, perceived relevance, and intervention efficacy. However, the third design consideration—specifying consequential outcomes—remains underdeveloped across the literature. While 140 (58%) papers discussed the interaction between context, individual traits, and well-being outcomes, many stopped short of evaluating long-term psychological or behavioral change. Instead, outcome assessments often relied on narrow usability metrics (eg, user satisfaction or screen time reduction) or short-term behavioral proxies (eg, frequency of feature use), without considering deeper indicators such as habit formation, digital resilience, or emotional regulation.

Building on these insights, the development of digital well-being interventions through AI can be conceptualized in two phases, aligned with the revamped double diamond design model [79]: (1) experience strategy—designing the right thing, and (2) experience design—designing things right (Figure 4). In the experience strategy phase, researchers and practitioners conduct user research to understand core challenges and motivations. Here, individual differences (SOR: organism) are identified through profiling and segmentation—considering factors such as personality, digital habits, emotional needs, and cognitive preferences. These insights help define targeted user groups and establish design priorities.

XSL•FO
RenderX

Figure 4. The developing process for digital well-being with AI systems. AI: artificial intelligence.



The experience design phase focuses on implementation—developing system features that respond to both individual and contextual variation (SOR: stimulus). This includes creating adaptive interfaces, intelligent feedback loops, and multimodal interventions that nudge users toward meaningful behaviors. As part of this phase, teams are encouraged to define consequential outcomes (SOR: response and outcome) not just in terms of app usage or time-on-task, but through broader and more enduring metrics such as user autonomy, emotional well-being, and sustainable habit change.

Despite the clear utility of this model, the review found that many studies continued to rely on simplistic evaluation strategies, such as postuse surveys or single-time-point measurements. While these can offer useful feedback, they often fail to capture the temporal and emotional dynamics that characterize digital well-being. For instance, AI systems intended to reduce screen time were rarely evaluated for their effects on users' perceived autonomy or cognitive fatigue. Likewise, productivity tools often neglect to measure whether increased efficiency was accompanied by psychological strain or decision exhaustion.

Additionally, promising developments were observed in systems that integrate generative AI for personalized support in mindfulness, reflection, and journaling tasks. These tools showed early evidence of enhancing emotional regulation and user engagement. However, consistent frameworks for evaluating their effects—particularly on user autonomy and mental energy—remain lacking. Furthermore, despite increased technical sophistication, emotion recognition algorithms still struggle with reliability, and few studies incorporated robust longitudinal evaluations. Equity concerns also surfaced across the dataset. Very few evaluation frameworks accounted for variables such as socioeconomic status, neurodiversity, or differential access to digital tools, raising concerns that certain AI interventions may inadvertently reinforce existing disparities rather than reduce them.

In conclusion, the development and evaluation of AI systems for digital well-being must move beyond high-level rhetoric and toward deeply contextualized, empirically grounded processes. By applying frameworks such as the SOR model and drawing from findings in interdisciplinary literature, researchers and practitioners can more effectively align design goals with meaningful, measurable, and user-centered outcomes. The field's advancement depends not only on technical innovation but also on a more nuanced understanding of how intelligent systems intersect with users' daily decisions, emotional states, and broader societal contexts.

Development of HCAI-DW Model

Background

The preceding 3 findings collectively underscore a fragmented but evolving landscape in the design and evaluation of intelligent systems for digital well-being. First, implementation efforts tend to cluster around 3 key domains—supporting well-being in daily HCIs, designing for behavioral and habit change, and deploying BITs—but each varies in conceptual clarity, theoretical grounding, and methodological rigor. Second, the value of user autonomy is widely recognized, but interpreted through different disciplinary lenses, leading to tensions between automation and agency. Third, despite strong calls for human-centered development, the field still lacks coherent, actionable methodologies for designing and assessing AI systems that meaningfully support users' emotional, behavioral, and cognitive well-being.

In response to these interrelated challenges, this study proposes an integrative framework: the HCAI-DW model. Developed to

XSL•FO
RenderX

bridge the persistent gaps between conceptual ideals and real-world implementation, the HCAI-DW model synthesizes insights from empirical findings, behavioral theory, and interdisciplinary practice. It offers a structured lens to guide researchers and practitioners in designing, deploying, and evaluating AI-powered systems that promote human flourishing in digital environments.

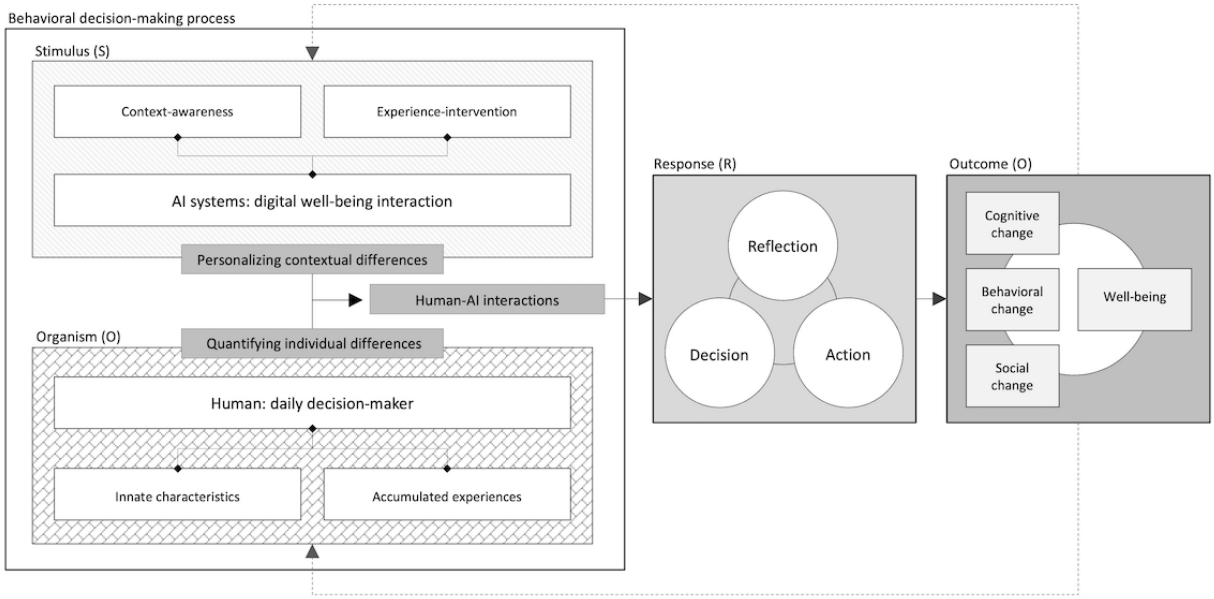
The model was developed through an iterative, participatory process that combined the systematic review of 240 research papers with expert feedback from 5 professors and 17 graduate students in HCI and HCD. These structured review sessions—conducted at both early and final stages of the modeling process—played a critical role in refining the model's conceptual clarity, usability, and relevance to current challenges. The goal was not only to reflect the state of the field but to cocreate a framework that is both theoretically robust and practically applicable.

A key insight that emerged throughout this process is that digital well-being is inherently multidimensional, shaped by the complex interplay of individual psychological traits, system-level design interventions, users' behavioral responses, and broader social or environmental outcomes. To capture this interplay, the HCAI-DW model adapts and extends the SOR model, a well-established paradigm in behavioral science, and reframes it for the context of human-AI interaction. This adaptation provides a structured but flexible blueprint for understanding how intelligent systems can shape and support digital well-being in everyday life.

Overview

As illustrated in Figure 5, the HCAI-DW model describes 4 essential and interrelated components that shape human-centered AI system development for digital well-being: (1) the organism component, which views users as daily behavioral decision makers; (2) the stimulus component, which frames AI system features as intentional behavioral interventions; (3) the response component, which introduces a cyclical model of user engagement involving decision, action, and reflection; and (4) the outcome component, which emphasizes a layered understanding of cognitive, behavioral, and social transformations. Together, these components offer a comprehensive pathway for designing intelligent systems that promote digital well-being in nuanced, personalized, and sustainable ways.

Figure 5. HCAI-DW Model. AI: artificial intelligence; HCAI-DW: human-centered artificial intelligence for digital well-being.



Organism: Daily Decision Makers

The organism component foregrounds the diversity of individual characteristics that influence how people interact with AI-powered systems. In the HCAI-DW model, these differences—including traits such as motivation, digital literacy, emotional state, and decision-making tendencies—are recognized as critical design inputs for tailoring interactions. As past research has demonstrated, individuals tend to make behavioral decisions based on their innate characteristics or accumulated experiences. Profiling user preferences and behaviors allows AI systems to deliver personalized support that is more effective and meaningful in fostering digital well-being [36,59]. The reviewed literature shows that most existing studies emphasized the importance of personalization, though only a fraction deployed adaptive systems that responded to individual differences in real time. By leveraging machine learning and deep learning algorithms, intelligent systems can move beyond generalized models to provide highly contextualized and evolving interventions [59,80,81]. This model suggests treating AI systems not merely as tools for behavior modification but as coregulatory agents that continuously adapt to users as they navigate complex decision-making environments. For this reason, the HCAI-DW model posits that seeing users as "daily behavioral decision makers" is essential for designing effective AI systems that address the nuances of digital well-being.

XSL•FO
RenderX

Stimulus: Digital Well-Being Interactions

The stimulus component redefines AI system features as intentional behavioral interventions that can promote or hinder digital well-being. While many previous initiatives have explored user interaction with AI as a design theme or an evaluation component [3,82], the HCAI-DW model differentiates between 2 interrelated stimulus phases: the context-awareness phase and the experience-intervention phase.

The context-awareness phase is derived from the quantified-self and personal informatics movements [59]. In this phase, AI systems track and interpret user data—including time, location, physical state, and social context—to deliver interventions that are sensitive to situational dynamics. Of the 240 reviewed studies, 94 specifically focused on the integration of such contextual variables into design. Wearable sensors, app usage logs, and mood-tracking tools were among the techniques used to improve the precision of interventions. The experience-intervention phase draws from BIT literature [70,83]. In this phase, AI systems deliver tailored and systematized interactions that aim to encourage behavioral shifts and preserve user autonomy. These interventions include nudges, reminders, content recommendations, and interface adjustments. Rather than enforcing a one-size-fits-all approach, the HCAI-DW model supports dynamic and transparent adaptation to individual feedback, reducing the risk of perceived manipulation or coercion.

Together, these 2 stimulus phases highlight the dual nature of AI system design: sensing and acting. Systems that only track without intervening, or intervene without contextual sensitivity, risk ineffectiveness or user resistance. The HCAI-DW model provides a roadmap for designing stimulus mechanisms that are integrated, adaptive, and user-centric.

Response: Decision-Action-Reflection Loop

The response component expands the understanding of user engagement beyond simplistic stimulus-response paradigms to include a full cycle of decision, action, and reflection. This loop is crucial for supporting sustainable behavior change and deeper cognitive engagement. Most previous studies have treated user responses as fixed behavioral outputs (eg, goal achievement and app usage). However, this approach overlooks the underlying reflective processes that support long-term well-being [16,36]. In contrast, the HCAI-DW model emphasizes the temporal and cyclical nature of engagement. AI systems should facilitate not only behavioral outcomes but also opportunities for users to reflect on their choices, understand system recommendations, and adjust their goals over time [17,19,53]. Such features are underdeveloped in the current literature; only few studies described systems that supported ongoing reflective practice, and even fewer implemented adaptive feedback based on longitudinal user behavior. Examples of effective reflective design include journaling prompts, feedback dashboards, and generative content that encourages users to process their experiences emotionally and cognitively. These elements help close the loop between system recommendation and user transformation, offering pathways for deeper and more enduring engagement.

Outcome: From Cognitive Change to Social Change

The outcome component identifies and categorizes the types of changes users experience after interacting with AI-powered systems [4,71]. Based on an analysis of prior research, the HCAI-DW model distinguishes 3 levels of outcome: cognitive change (eg, increased awareness and motivation), behavioral change (eg, reduced screen time and improved sleep), and social change (eg, enhanced family communication and digital equity). While many previous approaches target a single optimized outcome (eg, productivity or screen-time reduction), the HCAI-DW model encourages consideration of multilayered and user-defined metrics. It also addresses a common limitation of models such as Motivation, Engagement, and Thriving in User Experience model [27], which, despite offering a comprehensive structure for evaluating technology-mediated experience, struggle with implementation due to insufficient differentiation between experience levels and limited guidance on design implications.

This model fills that gap by suggesting a framework for systematically specifying consequential differences—that is, the variations in outcomes that result from users' unique behavioral decisions and reflective processes. Many prior interventions have focused on training or instructing users to modify their behavior [4,35] but have not always aligned these goals with users' lived experiences, values, or autonomy. The HCAI-DW model enables researchers and practitioners to avoid these pitfalls by distinguishing between process (decision-making) and product (outcome) and ensuring that both are addressed in a human-centered, contextualized manner.

In sum, the HCAI-DW model offers a synthesized and actionable framework that integrates the diverse findings of this review and provides a comprehensive guide for the design, development, and evaluation of intelligent systems that prioritize digital well-being. Its structure—rooted in the SOR model and refined through interdisciplinary and participatory input—reflects a growing need in the field for models that are not only conceptually rigorous but also practically useful. By highlighting individual differences, incorporating contextual adaptation, supporting reflective engagement, and mapping layered outcomes, the model promotes a vision of AI that respects human agency while optimizing support. As AI technologies become increasingly embedded in everyday life, models such as HCAI-DW will be critical in guiding their ethical, effective, and human-centered development.

Discussion

Principal Findings

This paper developed the HCAI-DW model to conceptualize how AI systems can support individuals' behavioral decision-making and foster behavioral changes that enhance their daily digital well-being. Grounded in evidence-based academic literature and established perspectives from relevant practical domains, this model reflects a human-centered approach to AI systems aimed at improving digital well-being. The HCAI-DW model is expected to enhance existing technology development practices by offering a systematic framework to encode and use core factors that drive digital

XSL•FO
RenderX

well-being in AI computing systems. Furthermore, the model was purposefully designed for intuitive and practical use, addressing the emerging needs of researchers and practitioners seeking to apply theoretical insights and design approaches to shape human experiences. As such, the model enables the creation of more effective AI system interactions by tailoring them to individual and contextual differences, particularly those associated with decision-making processes.

Limitations

This paper offers a timely investigation into human-centered AI systems designed to support users' digital well-being, incorporating insights from a systematic literature review. To provide a comprehensive and multidisciplinary overview of the most relevant research, 4 academic databases were used as the primary sources for data collection. This approach aimed to ensure the inclusion of diverse scholarly contributions from computing, design, behavioral science, and information systems. However, a key limitation lies in the fragmented nature of existing studies in this domain. Many initiatives related to digital well-being and intelligent computing systems are published under varied terminologies and conceptual framings. To address this, the initial phase of data collection used broadly defined search terms to capture a wide scope of literature. Despite this inclusive strategy, there is a possibility that some relevant studies—particularly those published under domain-specific or alternative terminologies—were not included. For example, literature from niche fields or interdisciplinary collaborations may have described similar phenomena without explicitly using the terms "digital well-being" or "AI systems."

Additionally, while the topic has gained increasing attention in both academic and industrial contexts, the number of studies explicitly integrating the concepts of AI and digital well-being remains relatively limited. This restricts the ability to generalize findings across all potential applications or use cases of AI systems designed for human well-being. Future systematic reviews should aim to capture these additional perspectives by refining search protocols and expanding the inclusion criteria to account for heterogeneity in terminologies, research methods, design strategies, usage contexts, and evaluative outcomes.

Furthermore, the HCAI-DW model proposed in this study should be understood as a preliminary step rather than a finalized theoretical framework. The model synthesizes key themes and design considerations from the existing body of work to propose an integrated perspective on designing for digital well-being with AI. However, the model has not yet been empirically validated, and its applicability across different cultural, technological, or organizational contexts remains uncertain. This limitation suggests that further empirical research—particularly studies conducted in real-world design and development settings—is necessary to test and refine the model's relevance, accuracy, and usability.

In summary, while this paper contributes a foundational understanding of how human-centered AI systems can support digital well-being, the evolving and interdisciplinary nature of the field calls for continued theoretical expansion and empirical validation. Follow-up works should also explore additional stakeholder perspectives and context-specific requirements to enhance the robustness and applicability of the HCAI-DW framework.

Future Directions

Overview

This study provides a foundation for advancing both theoretical and practical efforts in designing AI systems that promote digital well-being in daily HCIs. Based on insights from the systematic literature review, three key challenges emerge for future research and practice: (1) developing new evaluation methods for digital well-being in AI contexts, (2) addressing ethical considerations in AI-powered technology interventions, and (3) designing for hyper-personalization by accounting for individual decision-making contexts. The following sections elaborate on each of these challenges and outline opportunities to enrich the body of knowledge in this domain.

New Evaluation Methods for Digital Well-Being in AI Systems

Several studies have highlighted the limitations of existing evaluation tools for measuring subjective experiences of well-being in technology use. In particular, concerns have been raised regarding the reliability and validity of commonly used instruments [35,83]. From the perspective of the HCAI-DW model, digital well-being outcomes are shaped by a complex interplay of daily behavioral decisions and interactions with AI systems. However, a major challenge remains: there is a lack of well-validated and universally applicable measurement tools that can accurately assess the short-term and long-term effects of AI-based interventions.

Current tools are often adapted from adjacent domains and may not reflect the nuances of digital well-being in AI-enhanced contexts. As a result, the evaluation of AI system effectiveness remains inconsistent and difficult to standardize [84]. There is an urgent need to develop novel, context-sensitive evaluation methods that can capture the dynamic and individualized nature of well-being outcomes. These tools should account for both objective indicators and subjective experiences to ensure that the impact of AI systems on users' digital well-being can be meaningfully assessed.

Ethical Considerations in AI-Powered Interventions

While digital well-being aims to enhance users' lives through technology, intervention through AI systems must be approached with caution. Interventions, though potentially beneficial, are inherently value-neutral and may be used in manipulative or unethical ways [85]. For example, leveraging behavioral decision-making mechanisms could enable governments or corporations to influence individuals' choices in ways that serve organizational goals rather than personal well-being [54]. This raises significant ethical concerns, including the risk of undermining users' autonomy or reinforcing harmful behavioral patterns [86].

The HCAI-DW model calls for AI systems that are sensitive to individual needs and contexts. However, achieving this level of personalization often requires collecting sensitive behavioral data, which brings digital privacy into focus. Ethical design practices must address questions of informed consent, data

XSL•FO
RenderX

transparency, and user agency. Researchers and practitioners should ensure that AI interventions prioritize individual rights and safeguard against potential misuse or harm. Privacy-by-design principles and responsible data governance should be core components of future development strategies.

Designing for Hyper-Personalization Based on Individual Decision-Making Contexts

The advent of AI has fueled a growing trend toward hyper-personalization—tailoring content, services, and experiences to individual users based on real-time data and behavioral histories [17,18,62]. Although hyper-personalization holds promise for promoting sustained engagement and behavioral change, few studies have explored its implementation through the lens of digital well-being and human-centered AI [17,53,74].

The HCAI-DW model offers a conceptual foundation for understanding how personalization can support well-being by considering users' individual, contextual, and behavioral differences. It emphasizes that design should not only respond to users' past behaviors but also anticipate their goals, motivations, and environmental conditions. To fully realize the potential of hyper-personalization in AI systems, future research should identify actionable design principles that align AI-driven personalization with users' well-being outcomes. This includes developing frameworks for adaptive and context-aware personalization that respect user boundaries and preferences. By doing so, designers can increase the impact of AI systems on sustained behavioral change and help users make informed decisions that enhance their overall digital well-being.

Conclusions

This paper identifies everyday behavioral decision-making as a foundational concept for designing AI systems that enhance digital well-being. By synthesizing fragmented and often inconsistent perspectives from prior research, the study introduces the HCAI-DW model. This model offers a conceptual framework that links individual behaviors, AI system interactions, and well-being outcomes in everyday life. Through this model, this paper demonstrates how AI systems can be intentionally designed to support users' behavioral decisions in ways that promote meaningful and sustained improvements in digital well-being. Rather than treating digital well-being as an abstract or isolated construct, the HCAI-DW model emphasizes its dynamic and context-dependent nature, highlighting the critical role of personalized, ethical, and adaptive AI design. Ultimately, this work encourages researchers and practitioners to reconsider the evolving relationship between humans and AI in everyday contexts. It invites further exploration into what it means to be a "new type of human"—one whose daily life and well-being are increasingly mediated by intelligent technologies.

Acknowledgments

This project was supported by the NYC Scholarly Research Award from Pace University (F51338).

Conflicts of Interest

None declared.

Multimedia Appendix 1
The list of the reviewed 240 articles in the study.
[XLSX File (Microsoft Excel File), 46 KB - humanfactors_v12i1e69533_app1.xlsx]

Multimedia Appendix 2
PRISMA checklist.
[PDF File (Adobe PDF File), 109 KB - humanfactors_v12i1e69533_app3.pdf]

References

1. Burr C, Taddeo M, Floridi L. The ethics of digital well-being: a thematic review. Sci Eng Ethics 2020;26(4):2313-2343 [FREE Full text] [doi: 10.1007/s11948-020-00175-8] [Medline: 31933119]
2. Cecchinato ME, Rooksby J, Hiniker A, Munson S, Lukoff K, Ciolfi L, et al. Designing for digital wellbeing: a research and practice agenda. 2019 Presented at: Extended Abstracts of the SIGCHI Conference on Human Factors in Computing Systems; May 4 - 9, 2019; Glasgow, Scotland, UK p. 1-8. [doi: 10.1145/3290607.3298998]
3. Peters D, Ahmadpour N, Calvo RA. Tools for wellbeing-supportive design: features, characteristics, and prototypes. MTI 2020;4(3):40. [doi: 10.3390/mti4030040]
4. Calvo RA, Peters D. Positive Computing: Technology for Wellbeing and Human Potential. Cambridge, Massachusetts: MIT Press; 2014.
5. Desmet PM, Pohlmeyer AE. Positive design: an introduction to design for subjective well-being. Int J Des 2013;7(3):21-33.
6. Gaggioli A, Villani D, Serino S, Banos R, Botella C. Editorial: positive technology: designing E-experiences for positive change. Front Psychol 2019;10:1571 [FREE Full text] [doi: 10.3389/fpsyg.2019.01571] [Medline: 31338053]
7. Hassenzahl M, Eckoldt K, Diefenbach S, Laschke M, Lenz E, Kim J. Designing moments of meaning and pleasure: experience design and happiness. Int J Des 2013;7(3):21-31.

XSL•FO
RenderX

8. Kim Y, Jeon J, Choe E, Lee B, Kim K, Seo J. TimeAware: leveraging framing effects to enhance personal productivity. : ACM; 2016 Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; May 7 - 12, 2016; San Jose, CA p. 272-283. [doi: 10.1145/2858036.2858428]
9. Klapperich H, Laschke M, Hassenzahl M, Becker M, Cürlis D, Frackenpohl T, et al. Mind the gap: a social practice approach to wellbeing-driven design. In: Design for Wellbeing. Oxfordshire: Routledge; 2019:169.
10. Ko M, Choi S, Yatani K, Lee U. Lock n' LoL: group-based limiting assistance app to mitigate smartphone distractions in group activities. : ACM; 2016 Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; May 7 - 12, 2016; San Jose, CA p. 998-1010. [doi: 10.1145/2858036.2858568]
11. Turel O, Serenko A, Bontis N. Family and work-related consequences of addiction to organizational pervasive technologies. Inf Manage 2011;48(2-3):88-95. [doi: 10.1016/j.im.2011.01.004]
12. Okeke F, Sobolev M, Dell N, Estrin D. Good vibrations: can a digital nudge reduce digital overload? : ACM; 2018 Presented at: Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '18); Sept 3 - 6, 2018; Barcelona, Spain p. 1-12. [doi: 10.1145/3229434.3229463]
13. Pohlmeyer A. How Design Can (not) Support Human Flourishing. Berlin: Springer, Cham; 2017:235-255.
14. Thomas V, Azmitia M, Whittaker S. Unplugged: exploring the costs and benefits of constant connection. Comput Hum Behav 2016;63:540-548. [doi: 10.1016/j.chb.2016.05.078]
15. Monge RA, De RL. The race towards digital wellbeing: issues and opportunities. 2019 Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; May 4 - 9, 2019; Glasgow, Scotland, UK p. 1-14. [doi: 10.1145/3290605.3300616]
16. Shin Y, Yoon J. Towards designing human-centered time management interfaces: the development of 14 UX design guidelines for time-related experiences in mobile HCI. New York, NY, USA: Association for Computing Machinery; 2021 Presented at: Adjunct Publication of the 23rd International Conference on Mobile Human-Computer Interaction (MobileHCI '21); Sept 27 - Oct 1, 2021; Toulouse & Virtual France p. 1-7. [doi: 10.1145/3447527.3474861]
17. Amershi S, Weld D, Vorvoreanu M, Fourney A, Nushi B, Collisson P, et al. Guidelines for human-AI interaction. 2019 Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; May 4 - 9, 2019; Glasgow, Scotland, UK p. 1-13. [doi: 10.1145/3290605.3300233]
18. Howard J. Artificial intelligence: implications for the future of work. Am J Ind Med 2019;62(11):917-926. [doi: 10.1002/ajim.23037] [Medline: 31436850]
19. Shneiderman B. Int J Hum–Comput Interact 2020;36(6):495-504. [doi: 10.1080/10447318.2020.1741118]
20. Bonacini L, Gallo G, Scicchitano S. Working from home and income inequality: risks of a 'new normal' with COVID-19. J Popul Econ 2021;34(1):303-360 [FREE Full text] [doi: 10.1007/s00148-020-00800-7] [Medline: 32952308]
21. Baidu. How coronavirus is accelerating a future with autonomous vehicles. MIT Technology Review. 2020. URL: https://www.technologyreview.com/2020/05/18/1001760/how-coronavirus-is-accelerating-autonomous-vehicles/ [accessed 2025-08-07]
22. NA. Reopening: the tech-enabled office in a post-COVID world. CB Insights. URL: https://www.cbinsights.com/research/report/reopening-office-tech-work-post-covid/ [accessed 2025-08-07]
23. Lanaj K, Johnson RE, Barnes CM. Beginning the workday yet already depleted? Consequences of late-night smartphone use and sleep. Organ Behav Hum Decis Processes 2014;124(1):11-23. [doi: 10.1016/j.obhdp.2014.01.001]
24. Lee Y, Chang C, Lin Y, Cheng Z. The dark side of smartphone usage: Psychological traits, compulsive behavior and technostress. Comput Hum Behav 2014;31:373-383. [doi: 10.1016/j.chb.2013.10.047]
25. Roffarello AM, De Russis L. Designing technology that promotes users' digital wellbeing. XRDS 2021;28(1):14-18. [doi: 10.1145/3481823]
26. Satchell C, Dourish P. Beyond the user: usenon-use in HCI. : ACM; 2009 Presented at: Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7 (OZCHI '09); November 23 - 27, 2009; Melbourne, Australia p. 9-16. [doi: 10.1145/1738826.1738829]
27. Peters D, Calvo RA, Ryan RM. Designing for motivation, engagement and wellbeing in digital experience. Front Psychol 2018;9:797 [FREE Full text] [doi: 10.3389/fpsyg.2018.00797] [Medline: 29892246]
28. Smits M, Kim CM, van Goor H, Ludden GDS. From digital health to digital well-being: systematic scoping review. J Med Internet Res 2022;24(4):e33787 [FREE Full text] [doi: 10.2196/33787] [Medline: 35377328]
29. Ames MG. Managing mobile multitasking: the culture of iPhones on Stanford campus. : ACM; 2013 Presented at: Proceedings of the 2013 Conference on Computer Supported Cooperative Work; February 23 - 27, 2013; San Antonio, TX p. 1487-1498. [doi: 10.1145/2441776.2441945]
30. Chakraborty K, Basu D, Vijaya Kumar KG. Internet addiction: consensus, controversies, and the way ahead. East Asian Arch Psychiatry 2010;20(3):123-132 [FREE Full text] [Medline: 22348866]
31. Dabbish L, Mark G, González V. Why do I keep interrupting myself? Environment, habit, and self-interruption. : ACM; 2011 Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; May 7 - 12, 2011; Vancouver, British Columbia, Canada p. 3127-3130. [doi: 10.1145/1978942.1979405]
32. Douglas AC, Mills JE, Niang M, Stepchenkova S, Byun S, Ruffini C, et al. Internet addiction: meta-synthesis of qualitative research for the decade 1996–2006. Comput Hum Behav 2008;24(6):3027-3044. [doi: 10.1016/j.chb.2008.05.009]

XSL•FO
RenderX

33. Oulasvirta A, Rattenbury T, Ma L, Raita E. Habits make smartphone use more pervasive. Pers Ubiquit Comput 2011;16(1):105-114. [doi: 10.1007/s00779-011-0412-2]
34. Verduyn P, Lee DS, Park J, Shablack H, Orvell A, Bayer J, et al. Passive facebook usage undermines affective well-being: experimental and longitudinal evidence. J Exp Psychol Gen 2015;144(2):480-488. [doi: 10.1037/xge0000057] [Medline: 25706656]
35. Peters D, Ahmadpour N. Digital wellbeing through designvaluation of a professional development workshop on wellbeing-supportive design. : ACM; 2021 Presented at: Proceedings of the 32nd Australian Conference on Human-Computer Interaction (OzCHI '20); December 2 - 4, 2020; Sydney, NSW, Australia p. 148-157. [doi: 10.1145/3441000.3441008]
36. Shin Y, Kim C, Yoon J. Behavioural intervention technology in UX design: conceptual review, synthesis, and research direction. Singapore: Springer; 2022 Presented at: Congress of the International Association of Societies of Design Research; December 2–5, 2025; Taipei, Taiwan p. 450-465. [doi: 10.1007/978-981-19-4472-7_31]
37. Jacoby J. Stimulus‐organism‐response reconsidered: an evolutionary step in modeling (Consumer) behavior. J Consum Psychol 2008;12(1):51-57. [doi: 10.1207/s15327663jcp1201_05]
38. Mehrabian A, Russell JA. An Approach to Environmental Psychology. Cambridge, MA: MIT Press; 1974.
39. Jai T, O'Boyle MW, Fang D. Neural correlates of sensory‐enabling presentation: an fMRI study of image zooming and rotation video effects on online apparel shopping. J Consum Behav 2014;13(5):342-350. [doi: 10.1002/cb.1476]
40. Tai SH, Fung AM. Application of an environmental psychology model to in-store buying behaviour. Int Rev Retail, Distrib Consum Res 2011;7(4):311-337. [doi: 10.1080/095939697342914]
41. Lee J, Kim KJ, Lee S, Shin D. Can autonomous vehicles be safe and trustworthy? Effects of appearance and autonomy of unmanned driving systems. Int J Hum-Comput Interact 2015;31(10):682-691. [doi: 10.1080/10447318.2015.1070547]
42. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
43. Tranfield D, Denyer D, Smart P. Towards a methodology for developing evidence‐informed management knowledge by means of systematic review. Br J Manage 2003;14(3):207-222. [doi: 10.1111/1467-8551.00375]
44. Holtzblatt K, Wendell JB, Wood S. Rapid Contextual Design: A How-to Guide to Key Techniques for User-Centered Design. Amsterdam, Netherlands: Elsevier; 2004.
45. Fogg BJ. Persuasive technology. Ubiquity 2002;2002(December):5. [doi: 10.1145/764008.763957]
46. Keyes CLM, Annas J. Feeling good and functioning well: distinctive concepts in ancient philosophy and contemporary science. J Positive Psychol 2009;4(3):197-201. [doi: 10.1080/17439760902844228]
47. Klapperich H, Laschke M, Hassenzahl M. The positive practice canvas: gathering inspiration for wellbeing-driven design. : ACM; 2018 Presented at: Proceedings of the 10th Nordic Conference on Human-Computer Interaction; Sept 29 - Oct 3, 2018; Oslo Norway p. 74-81. [doi: 10.1145/3240167.3240209]
48. Consolvo S, McDonald DW, Landay JA. Theory-driven design strategies for technologies that support behavior change in everyday life. : ACM; 2009 Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; April 4 - 9, 2009; Boston MA USA p. 405-414. [doi: 10.1145/1518701.1518766]
49. Dodge R, Daly A, Huyton J, Sanders L. The challenge of defining wellbeing. Int J Wellbeing 2012;2(3):222-235. [doi: 10.5502/ijw.v2i3.4] [Medline: 8486778]
50. Axelsson M, Spitale M, Gunes H. "Oh, Sorry, I Think I Interrupted You": designing repair strategies for robotic longitudinal well-being coaching. : ACM; 2024 Presented at: HRI '24: Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction; March 11 - 15, 2024; Boulder, CO p. 13-22. [doi: 10.1145/3610977.3634948]
51. Chang Y, Yang C. Examining the mechanisms of virtual reality tourism's impact on the mental well-being of long-term care facility residents: perspectives on presence and flow. Behav Sci (Basel) 2024;14(9):781 [FREE Full text] [doi: 10.3390/bs14090781] [Medline: 39335996]
52. Torkamaan H, Ziegler J. Integrating behavior change and persuasive design theories into an example mobile health recommender system. : ACM; 2018 Presented at: UbiComp/ISWC '21 Adjunct: Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers; Sept 21 - 26, 2021; Virtual USA p. 218-225. [doi: 10.1145/3460418.3479330]
53. Niedderer K, Ludden G, Clune S, Lockton D, Mackrill J, Morris A, et al. Design for behaviour change as a driver for sustainable innovation: challenges and opportunities for implementation in the private and public sectors. Int J Des 2016;10(2):67-85.
54. Pinder C, Vermeulen J, Cowan BR, Beale R. Digital behaviour change interventions to break and form habits. ACM Trans Comput-Hum Interact 2018;25(3):1-66 [FREE Full text] [doi: 10.1145/3196830]
55. Stawarz K, Cox A. Designing for health behavior change: HCI research alone is not enough. 2015 Presented at: CHI'15 workshop: Crossing HCI and Health: Advancing Health and Wellness Technology Research in Home and Community Settings; April 18-23, 2015; Seoul, South Korea. [doi: 10.1145/2702613.2702652]
56. Direito A, Walsh D, Hinbarji M, Albatal R, Tooley M, Whittaker R, et al. Using the intervention mapping and behavioral intervention technology frameworks: development of an mHealth intervention for physical activity and sedentary behavior change. Health Educ Behav 2018;45(3):331-348. [doi: 10.1177/1090198117742438] [Medline: 29216765]

XSL•FO
RenderX

57. Sumner JA, Carey RN, Michie S, Johnston M, Edmondson D, Davidson KW. Using rigorous methods to advance behaviour change science. Nat Hum Behav 2018;2(11):797-799 [FREE Full text] [doi: 10.1038/s41562-018-0471-8] [Medline: 30931398]
58. Mohr DC, Schueller SM, Montague E, Burns MN, Rashidi P. The behavioral intervention technology model: an integrated conceptual and technological framework for eHealth and mHealth interventions. J Med Internet Res 2014;16(6):e146 [FREE Full text] [doi: 10.2196/jmir.3077] [Medline: 24905070]
59. Kersten-van Dijk ET, Westerink JH, Beute F, IJsselsteijn WA. Personal informatics, self-insight, and behavior change: a critical review of current literature. Hum–Comput Interact 2017;32(5-6):268-296. [doi: 10.1080/07370024.2016.1276456]
60. Bolpagni M, Pardini S, Gabrielli S. Human centered design of AI-powered digital therapeutics for stress prevention: perspectives from multi-stakeholders' workshops about the SHIVA solution. Internet Interv 2024;38:100775 [FREE Full text] [doi: 10.1016/j.invent.2024.100775] [Medline: 39314669]
61. Raff S, Wentzel D, Obwegeser N. Smart Products: conceptual review, synthesis, and research directions. J Prod Innov Manage 2020;37(5):379-404. [doi: 10.1111/jpim.12544]
62. Walch K. The seven patterns of AI. Forbes. 2019. URL: https://www.forbes.com/sites/cognitiveworld/2019/09/17/the-seven-patterns-of-ai/#3c6fbc1012d0 [accessed 2025-08-07]
63. Debackere F, Clavel C, Roren A, Tran V, Messai Y, Rannou F, et al. Design framework for the development of tailored behavior change technologies. 2023 Presented at: Adjunct Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization; June 26, 2023; Limassol, Cyprus p. 140-146. [doi: 10.1145/3563359.3596984]
64. Brons A, Wang S, Visser B, Kröse B, Bakkes S, Veltkamp R. Machine learning methods to personalize persuasive strategies in mHealth interventions that promote physical activity: scoping review and categorization overview. J Med Internet Res 2024;26:e47774 [FREE Full text] [doi: 10.2196/47774] [Medline: 39546334]
65. Hun Lee M, Siewiorek DP, Smailagic A, Bernardino A, Bermúdez I Badia S. Design, development, and evaluation of an interactive personalized social robot to monitor and coach post-stroke rehabilitation exercises. User Model User-adapt Interact 2023;33(2):545-569 [FREE Full text] [doi: 10.1007/s11257-022-09348-5] [Medline: 37123108]
66. Wang L, Miller LC. Just-in-the-moment adaptive interventions (JITAI): a meta-analytical review. Health Commun 2020;35(12):1531-1544. [doi: 10.1080/10410236.2019.1652388] [Medline: 31488002]
67. Dobkin BH, Dorsch AK. The evolution of personalized behavioral intervention technology: will it change how we measure or deliver rehabilitation? Stroke 2017;48(8):2329-2334 [FREE Full text] [doi: 10.1161/STROKEAHA.117.016620] [Medline: 28679855]
68. Shin Y. Supporting users' decision-making experiences through hyper-personalized human-technology interactions. New York, NY, USA: Association for Computing Machinery; 2022 Presented at: Proceedings of Designing Interactive Systems Conference (DIS '22 Companion); June 13 - 17, 2022; Virtual Event Australia p. 8-11. [doi: 10.1145/3532107.3532873]
69. Hermsen S, Renes R, Frost J. Persuasive by design: a model and toolkit for designing evidence-based interventions. 2014 Presented at: Chi Spark Nederland; April 03, 2014; The Hague University of Applied Sciences p. 74-77 URL: https://research.vu.nl/files/1162981/2014%20CHI%20Sparks%20Donker%20et%20al%20tiles.pdf
70. Russ TC, Woelbert E, Davis KAS, Hafferty JD, Ibrahim Z, Inkster B, MQ Data Science group. How data science can advance mental health research. Nat Hum Behav 2019;3(1):24-32 [FREE Full text] [doi: 10.1038/s41562-018-0470-9] [Medline: 30932051]
71. Calvo R, Peters D, Johnson D, Rogers Y. Autonomy in technology design. : ACM; 2014 Presented at: Extended Abstracts of the SIGCHI Conference on Human Factors in Computing Systems; Apr 26, 2014; Toronto, Ontario, Canada p. 37-40. [doi: 10.1145/2559206.2560468]
72. Ryan RM, Deci EL. On happiness and human potentials: a review of research on hedonic and eudaimonic well-being. Annu Rev Psychol 2001;52:141-166. [doi: 10.1146/annurev.psych.52.1.141] [Medline: 11148302]
73. Zanker M, Rook L, Jannach D. Measuring the impact of online personalisation: past, present and future. Int J Hum-Comput Stud 2019;131:160-168. [doi: 10.1016/j.ijhcs.2019.06.006]
74. Verganti R, Vendraminelli L, Iansiti M. Innovation and design in the age of artificial intelligence. J Prod Innov Manage 2020;37(3):212-227. [doi: 10.1111/jpim.12523]
75. Navarro-Alamán J, Lacuesta R, García-Magariño I, Lloret J. EmotIoT: An IoT system to improve users' wellbeing. Appl Sci 2022;12(12):5804. [doi: 10.3390/app12125804]
76. Chang TR, Kaasinen E, Kaipainen K. What influences users' decisions to take apps into use? A framework for evaluating persuasive and engaging design in mobile apps for well-being. 2012 Presented at: Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia; December 4 - 6, 2012; Ulm, Germany p. 1-10. [doi: 10.1145/2406367.2406370]
77. García-Sáez G, Rigla M, Martínez-Sarriegui I, Shalom E, Peleg M, Broens T, et al. Patient-oriented computerized clinical guidelines for mobile decision support in gestational diabetes. J Diabetes Sci Technol 2014;8(2):238-246 [FREE Full text] [doi: 10.1177/1932296814526492] [Medline: 24876573]
78. Xu X, Zou T, Xiao H, Li Y, Wang R, Yuan T, et al. TypeOut: Leveraging just-in-time self-affirmation for smartphone overuse reduction. 2022 Presented at: Proceedings of the CHI Conference on Human Factors in Computing Systems; April 29, 2022; New Orleans, LA p. 1-17. [doi: 10.1145/3491102.3517476]

XSL•FO
RenderX

79. Nessler D. How to Apply a Design Thinking, HCD, UX or Any Creative Process from Scratch. Medium. 2024. URL: https://medium.com/digital-experience-design/how-to-apply-a-design-thinking-hcd-ux-or-any-creative-process-from-scratch-b8786efbf812 [accessed 2025-08-07]
80. Hirsh JB, Kang SK, Bodenhausen GV. Personalized persuasion: tailoring persuasive appeals to recipients' personality traits. Psychol Sci 2012;23(6):578-581. [doi: 10.1177/0956797611436349] [Medline: 22547658]
81. Shin Y, Kim J. Data-centered persuasion: nudging user's prosocial behavior and designing social innovation. Comput Hum Behav 2018;80:168-178. [doi: 10.1016/j.chb.2017.11.009]
82. Harlow LL, Oswald FL. Big data in psychology: introduction to the special issue. Psychol Methods 2016;21(4):447-457 [FREE Full text] [doi: 10.1037/met0000120] [Medline: 27918177]
83. Law EL, van Schaik P, Roto V. Attitudes towards user experience (UX) measurement. Int J Hum-Comput Stud 2014;72(6):526-541. [doi: 10.1016/j.ijhcs.2013.09.006]
84. Hermes ED, Lyon AR, Schueller SM, Glass JE. Measuring the implementation of behavioral intervention technologies: recharacterization of established outcomes. J Med Internet Res 2019;21(1):e11752 [FREE Full text] [doi: 10.2196/11752] [Medline: 30681966]
85. Page R, Kray C. Ethics and persuasive technology: an exploratory study in the context of healthy living. 2010 Presented at: Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Service; Sept 7 - 10, 2010; Lisbon, Portugal p. 527-530 URL: https://ceur-ws.org/Vol-690/NIMD10Proceedings.pdf
86. Jun GT, Carvalho F, Sinclair N. Ethical issues in designing interventions for behavioural change. 2018 Presented at: Proceedings of Design Research Society; June 25-28, 2018; Limerick, Ireland. [doi: 10.21606/drs.2018.498]

Abbreviations

ACM: Association for Computing Machinery
AI: artificial intelligence
BIT: behavioral intervention technology
HCAI-DW: human-centered artificial intelligence for digital well-being
HCD: human-centered design
HCI: human-computer interaction
ICT: information and communication technology
JITAI: Just-In-Time Adaptive Intervention
PIT: personal informatics tool
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SIGCHI: Special Interest Group on Computer-Human Interaction
SOR: stimulus-organism-response
UX: user experience

Edited by E Borycki; submitted 02.12.24; peer-reviewed by CMDS Freitas, KJ Sileyew; comments to author 18.04.25; revised version received 09.06.25; accepted 30.07.25; published 10.09.25.

Please cite as:
Shin Y
Toward Human-Centered Artificial Intelligence for Users' Digital Well-Being: Systematic Review, Synthesis, and Future Directions
JMIR Hum Factors 2025;12:e69533
URL: https://humanfactors.jmir.org/2025/1/e69533
doi:10.2196/69533
PMID:40928842

©Youngsoo Shin. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 10.09.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Endometriosis Support and Development of Digital Technology–Based Interventions: Systematic Review

Tivizio Pavic[1*], MSc; Kévin Nadarajah[1,2*], PhD; Alain Somat[1], Prof Dr; Geneviève Cabagno[3], Prof Dr; Florence Terrade[4], Prof Dr

[1]Laboratoire de Psychologie: Cognition, Comportement, Communication (LP3C), Univ Rennes, Rennes, France

[2]Équipe Recherche Psychologie Appliquée (PsyCAP), Centre d'Études et d'Expertise sur les Risques, l'Environnement, la Mobilité et l'Aménagement (Cerema), Saint-Brieuc, France

[3]Valeurs, Innovations, Politiques, Socialisations et Sports (VIPS2), Univ Rennes, Rennes, France

[4]Centre d'études et de recherches en psychopathologie et psychologie de la santé (CERPPS), Toulouse, France

[*]these authors contributed equally

Corresponding Author:
Geneviève Cabagno, Prof Dr
Valeurs, Innovations, Politiques, Socialisations et Sports (VIPS2)
Univ Rennes
Avenue Charles Tillon
Rennes
France
Phone: 33 2 23 22 78 68
Email: genevieve.cabagno@univ-rennes2.fr

Abstract

Background: Endometriosis is a chronic disease that affects 1 in 10 women worldwide. The disease affects patients' daily life at physical, psychological, and social levels. In recent years, the management of this disease has evolved, thanks in particular to the emergence of digital technologies and associated interventions. However, despite their growing use, there seems to be no systematic review of their development, design, and efficacy.

Objective: A systematic review was conducted with the aim of characterizing the development process, design, and effectiveness of interventions using a digital tool for endometriosis.

Methods: A total of 7 databases (MEDLINE, APA PsycArticles, Academic Search Premier, Psychology and Behavioral Sciences Collection, APA PsycInfo, SocINDEX, and SPORTDiscus) were searched to identify relevant articles published between 2010 and 2024. The articles selected were analyzed using a methodological framework specific to the development of digital health interventions (Design and Evaluation of Digital Health Interventions [DEDHI]), consisting of 4 phases: preparation (phase 1, specific to application development), optimization (phase 2, dedicated to identifying the best intervention configurations), evaluation (phase 3, aiming to confirm the effectiveness of the intervention), and implementation (phase 4, implementing and updating the intervention on a large scale).

Results: A selection of 10 articles was made from the 381 studies retrieved from the databases. Among these 10 studies, 6 distinct digital health interventions were identified. The interventions based on digital devices produced physical and psychological benefits. Analysis using the DEDHI framework showed (1) a disparity in the responses to the different phases (ie, 9/10, 90% of studies responding to phase 1; 3/10, 30% to phase 2; 4/10, 40% to phase 3; and 2/10, 20% to phase 4) and (2) a variability in the completion of the evaluation criteria ranging from 10% (1/10) to 80% (8/10) in phase 1, 0% (0/13) to 77% (10/13) in phase 2, 0% (0/10) to 80% (8/10) in phase 3, and finally 0% (0/13) to 77% (10/13) in phase 4. The objectives of these digital interventions were to support pain management (5/6, 83%), to provide information about the disease and strategies for managing it (4/6, 67%), and to provide psychosocial support (2/6, 33%).

Conclusions: This systematic review highlights an emerging literature, limited regarding the use of digital technology in the management of endometriosis, and heterogeneous concerning the methodologies used. This variability limits the generalizability of the results and requires a nuanced interpretation of the available data. However, the results of this review have demonstrated the value of digital technology–based interventions to support endometriosis, while highlighting the importance of a methodological framework to structure their development to optimize patient support.

XSL•FO
RenderX

(JMIR Hum Factors 2025;12:e71859) doi:10.2196/71859

KEYWORDS

endometriosis; digital health interventions; technological intervention; endometriosis care; development framework; systematic review

Introduction

Endometriosis is a chronic disease characterized by the abnormal presence of an epithelium, stroma, or both, histologically similar to the endometrium, located outside the endometrium and myometrium, and frequently associated with an inflammatory state [1,2]. Its prevalence is estimated at around 10%; that is, approximately 190 million women of childbearing age worldwide [3]. Endometriosis is currently presented by major organizations such as the WHO, the European Parliament, the Office on Women's Health, and the scientific community as a major public health problem [4]. Although responsible for significant morbidity, endometriosis is a disease for which there is as yet no effective treatment [5].

Lack of treatment has a deleterious impact on patients' quality of life [6]. Thus, numerous physical constraints, such as chronic pelvic pain [7]; associated symptoms such as dysmenorrhea, dysuria, dyschezia, and dyspareunia [8]; and a greater predisposition to the risk of comorbidities [9] arise from this disease. As well as these physical impacts, endometriosis can also cause anxiety or depression [10], psychosomatic disorders [11], reduced self-esteem [12], or a negative perception of one's body [13]. Finally, endometriosis can also affect these women's social relationships, as it severely restricts their daily and professional lives [14], causes difficulties in their relationships and sexuality [15,16], and can, in some cases, lead to isolation [17]. The diversity of these symptoms and their varying degrees of severity in each woman can make universal management difficult. It is, therefore, important that the support offered to patients is based on a personalized approach (ie, a care pathway tailored to each patient), multidisciplinary (ie, close collaboration between different health care professionals), and multimodal (ie, several types and sources of support) [18].

Chronic disease support is embedded in a global context, where digital technology is playing an increasingly important role in health care [19]. Especially since the COVID-19 pandemic, digital technologies are receiving more and more attention as a relevant solution to improve chronic disease management [20,21]. Digital technologies can involve a wide range of aspects such as telemedicine, mobile health apps, internet, websites, virtual reality, artificial intelligence, distance education, etc [22-24]. The growth of these technologies offers new possibilities for digital interventions designed to support the development of knowledge, behavioral changes, etc [25]. For chronic disease, a dominant approach is to use these technologies to increase patients' autonomy in managing their symptoms [26]. They can be used to improve patients' quality of life [27,28] and allow them to be in charge of certain aspects of their health, such as symptom management [29].

Endometriosis has not escaped this digital wave, as several countries, including the Australian and French governments, have already introduced national plans based on these technologies to combat the disease. These plans aim to develop interventions to enhance the quality of life for women with endometriosis through the use of digital technology, both to raise awareness and to improve patient care. According to Giudice et al [30], there is an urgent need to invest in digital technologies (eg, mobile apps) dedicated to the management of endometriosis. Digital technologies can help to prevent the risks associated with the disease, involve different professionals in the care pathway, and adapt the care offered to the needs of each patient. Since 2013, enthusiasm for digital technologies addressing endometriosis has grown, with at least 26 health applications on platforms like the Apple iTunes Store or Google Play. These applications are designed to help patients manage their symptoms, provide information on diagnosis and treatment, act as a social network, and allow them to share their experiences or keep a health diary [31]. Although there is strong interest among researchers, politicians, and developers in interventions that use digital technologies to support chronic conditions, such as endometriosis [30,32,33], there is also evidence showing that these technologies do not necessarily deliver the expected positive health outcomes [34]. A lack of specific guidelines dedicated to the development and evaluation of technologies could explain this inefficiency [35,36]. Thus, methodological frameworks such as the iterative Design and Evaluation of Digital Health Interventions (DEDHI) [37] have been developed to support the long-term implementation of a digital technological innovation dedicated to health. Nevertheless, this type of framework still seems to have limited use [38,39], although it can be useful for guiding the development of innovations dedicated to newly studied pathologies, as is the case for endometriosis [4,37].

Scientists, politicians, and decision-makers are tending to promote the use of digital technologies, and applications dedicated to endometriosis are multiplying. However, to the authors' knowledge, there is no systematic review on digital and technological interventions to support the management (ie, pain, mental health, knowledge, behavior, etc) of endometriosis. This systematic review aims to examine how digital health interventions (DHIs) targeting endometriosis have been developed and designed and to evaluate their effectiveness, with a particular focus on the use of a structured methodological framework and reported outcomes related to physical and psychological health.

Methods

Protocol

This systematic review followed the recommendations of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist (Multimedia Appendix 1) [40,41]. The protocol of this review has been registered in the

XSL•FO
RenderX

International Prospective Register of Systematic Reviews (PROSPERO CRD42023461263).

Inclusion Criteria

The inclusion criteria for the studies are as follows: (1) participants must have endometriosis, (2) use digital technology solutions to manage endometriosis, (3) take place between 2010 and 2024, to reflect current technological development [42], and (4) be written in French or English.

Search Strategy

The following computer databases were searched: Academic Search Premier, MEDLINE, APA PsycArticles, Psychology and Behavioral Sciences Collection, APA PsycInfo, SocINDEX, and SPORTDiscus. A total of 3 filters were used when searching the databases: years (2010-2024), peer-reviewed academic journals, and the terms appearing in the title or abstract. In order to make the research equation as exhaustive as possible, particularly with regard to digital tools, it was decided to draw on equations from recent research specifically on technological and digital interventions in the field of mental health [23]. These databases were examined using the following search equation: “endometriosis” AND (websites OR “smartphone app*” OR wearable OR “virtual reality” OR “augmented reality” OR “immersive technology” OR platform OR mhealth OR “mobile health” OR ehealth OR “e-mental health” OR e-health OR internet OR mail OR chat OR SMS OR text message OR digital intervention OR technological intervention) NOT (systematic review or meta analysis or scoping review or literature review or umbrella review). Databases were searched using the same equation through the Rennes 2 University library portal, which provides access to a wide range of bibliographic resources. In July 2024, the search equation was run for the first time and identified 5 articles in the final selection. In line with best practice for systematic reviews (eg, see chapter 4 of the Cochrane Handbook) [43], which recommends rerunning the search before final analysis to capture all newly published studies, the same search equation was rerun in November 2024 before submission. This update ensured the completeness and timeliness of the review before submission. The updated search identified a further 5 articles.

Study Selection

As proposed by the PRISMA methodology [41], the screening stages were followed. Zotero software (Digital Scholar) [44] was used to extract references from databases. Rayyan software (developed by Ouzzani et al [45]) was used for the first selection steps, that is, duplicate identification and selection by title and abstract [46].

After removing duplicates, the titles and abstracts of all identified articles were examined for eligibility. Articles were excluded if (1) endometriosis was not mentioned, (2) participants were not ill, (3) it was a systematic review, exploratory review, or meta-analysis, (4) it did not mobilize digital technology or focus on an intervention mobilizing digital technology, (5) it was a research protocol, (6) it was written in a language other than French or English, and (7) it was written before 2010.

Potentially eligible articles were retrieved, and if the text was not available, the authors were contacted. All the texts retrieved were then examined by 2 reviewers (TP and KN). Any disagreements were resolved by discussion with a third author (GC).

Risk of Bias

The risk of bias of each study included in this review was assessed with the Mixed Methods Assessment Tool (MMAT) [47]. A percentage was calculated for each study as proposed by Pluye et al [48] using the following formula: (number of “yes” responses divided by the number of “relevant criteria”) × 100. “Yes” indicates that an item is satisfied. A higher score indicates a lower risk of bias. The risk of bias was analyzed by 2 reviewers. A general intraclass correlation coefficient was calculated to determine interrater reliability. As recommended by Hong et al [47], no study, not even one of low methodological quality, was excluded. Interjudge agreement was satisfactory (0.98, 95% CI 0.82-0.99). Percentage scores are reported in Table 1.

XSL•FO
RenderX

Table 1. Study characteristics for included studies (n=10).

Specific digital tool	Authors	Country	Methods	Diagnosis sample size, n	Digital technologies	Risk of bias[a], n (%)	Objective
Isolated initiative of an intervention test using a digital tool							
	Abdulai et al [49]	Canada	Mixed methods	Patients diagnosed with or clinically suspected of having endometriosis (n=12)	Website	4.5 (90)	Evaluate the usability of the website and assess for destigmatizing properties of sexual health–related web-based resources.
	Li et al [50]	Australia	Case study	N/A[b]	development of the information recommendation functionalities on the website	2 (40)	Presentation of a methodology for developing and implementing health information recommendation functions within web-based health applications, using the example of a website on endometriosis.
	Lutfi et al [51]	Australia	Quantitative pilot RCTs[c]	Endometriosis diagnosis (n=19; telehealth=7, VR[d]=8, control=4)	Telehealth (1 h) vs VR (1 h) vs control (continue with their activities of daily living)	3.5 (70)	Determine the immediate impact of a single session of "supervised" telehealth-delivered exercise compared to "self-managed" VR-delivered exercise on pelvic pain associated with endometriosis.
Testing an intervention using virtual reality (Endocare)							
	Merlot et al [52]	France	Quantitative RCTs	MRI[e] endometriosis Diagnosis (n=45; Endocare=23, control=22)	VR vs control (same situation but in 2D on a tablet)	3.5 (70)	Measure the immediate and 4-hour persisting effects of a single use 20-minute DTx[f] (Endocare) on pain in women experiencing pelvic pain due to endometriosis.
	Merlot et al [53]	France	Quantitative RCTs	Endometriosis diagnosis (n=102; Endocare=51, control=51)	VR vs control (same situation but in 2D on a tablet)	3.5 (70)	Assess the effects of repeated at-home administrations of a 20-minute (VR) solution (Endocare) compared with a sham condition on pain in women experiencing pelvic pain due to endometriosis.
Testing an intervention using a mobile app (Endo-App)							
	Rohloff et al [54]	Germany	Quantitative	Prior diagnosis of endometriosis (n=106; Endo-APP=64, no user=42)	App	3.5 (70)	Examine whether there is evidence of beneficial effects of the smartphone app "Endo-App" and whether a multicenter randomized controlled trial should be planned to substantiate these effects.
	Rohloff et al [55]	Germany	Quantitative pilot RCTs	Medical diagnosis of endometriosis (n=122)	App + standard care vs usual care	4.5 (90)	Examine the impact of the Endo-App on both disease-related quality of life and symptoms of endometriosis affecting it.

XSL•FO
RenderX

Specific digital tool	Authors	Country	Methods	Diagnosis sample size, n	Digital technologies	Risk of bias[a], n (%)	Objective
	Zugaj et al [56]	Germany	Qualitative	Patients diagnosed with endometriosis and who have received a prescription for the DiGa[g] application: "Endo-App" (n=10)	App	5 (100)	Investigates how a health care app can influence the subjective experience of illness in patients with endometriosis.
Testing an intervention using SMS text messaging (EndoSMS)							
	Sherman et al [57]	Australia	Mixed methods	Diagnosed with endometriosis (self-reported; n=17)	SMS text messaging	4 (80)	Co-design and evaluate the acceptability, readability, and quality of a bank of supportive SMS text messages (EndoSMS) for individuals with endometriosis.
	Sherman et al [58]	Australia	Mixed methods integrated into RCTs	Clinically diagnosed with endometriosis (n=225; SMS=110, waitlist=115)	SMS text messaging	4.5 (90)	Determine the feasibility and acceptability of EndoSMS, a psychologically focused SMS text messaging intervention designed to support individuals living with endometriosis.

[a]A higher percentage indicates a lower risk of bias.

[b]N/A: not available.

[c]RCT: randomized controlled trial.

[d]VR: virtual reality.

[e]MRI: magnetic resonance imaging.

[f]DTx: digital therapeutics.

[g]DiGa: digital health applications.

Analysis of Articles Using the Framework for Designing and Evaluating DHIs

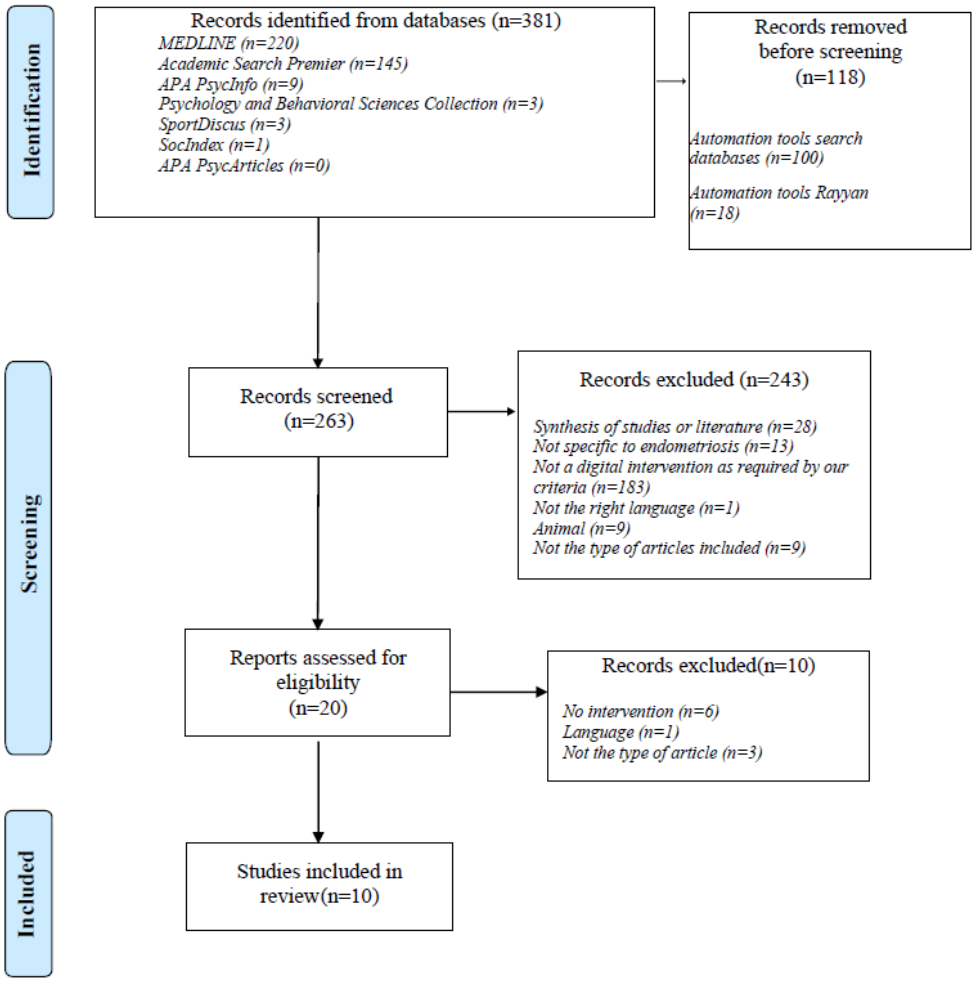
The DEDHI framework [37] was used to characterize the overall development process of digital technology–based interventions proposed in the studies of this corpus. Indeed, this framework seems particularly appropriate for the development of interventions based on digital technologies, specific guidelines dedicated to the design, and large-scale implementation of health-related innovations [59,60]. This framework proposes four stages: (1) the "preparation phase" describes the design phases of the innovation—review of the knowledge justifying an intervention, development of a conceptual model, and feasibility or acceptability study to test new components of the health intervention while identifying the optimization criterion offering the best result. This phase, therefore, requires a prototype offering the basic functionalities. (2) The "optimization phase" suggests building an optimized health care intervention using microtrials to identify the best configuration for the innovation, while respecting the optimization criteria. This phase requires a prototype offering all the functionalities set out in the conceptual model. (3) The "evaluation phase" is used to confirm the effectiveness of a health intervention, in particular, by carrying out a randomized controlled trial. At this stage, the prototype must be fully functional. And (4) the "implementation phase" offers a guide for the large-scale implementation and maintenance of an effective health intervention. For each phase, different criteria need to be addressed in order to optimize the long-term use of health innovations and recommended levels of technological maturity. They are organized according to 3 axes into each phase: goals and tasks (from 1 to 4 criteria, eg, feasibility and acceptability in the first phase), technical maturity (1 criterion, eg, elaborated research prototype in phase 2), and evaluation (from 5 to 10 criteria, eg, adherence, safety, and personalization). The 2 reviewers (TP and KN) independently examined the presence or absence of each criterion in the articles, and a third reviewer (GC) was called in in the event of disagreement. This made it possible to calculate a percentage of completion for each phase, determined by the ratio between the number of criteria present and the total number of criteria expected for that phase.

Results

Figure 1 illustrates the screening stages in the PRISMA flow diagram.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram.



Studies Design

Of the 10 studies included, 3 (30%) used mixed methods [49,57], one of which was a pilot randomized controlled trial [58]. One study was based on a qualitative analysis [56], and 5 (50%) studies were solely based on quantitative analyses [51-55], 4 of which were randomized controlled trials [51-53,55]. Finally, a study on the development of a methodology for designing a digital tool was based on a case study of endometriosis [50]. For this reason, due to its general methodology, this article will not be included in the following results sections: Sample Size, Age, and Participants.

Studies Characteristics

Sample Size

Among the 9 included studies, the smallest reported sample size was 10 [56] and the largest was 225 [58]. In quantitative studies, the number of participants ranged from 19 [51] to 122 [55]. In the mixed methods studies, the minimum number of participants was 12 [49] and the maximum was 225 [58]. Over the 9 studies, the average sample size was 73.1 (SD 72.4) individuals.

Age

Age is reported with varying degrees of accuracy depending on the methodologies used. The minimum age reported is 16 years [56], and the maximum is 63 years [49]. Based on the average ages reported by the studies [49,51-55,57,58], the total average age is 31.43 (SD 3.89) years.

Participants

The majority of studies (7/9, 78%) were based on samples of women diagnosed with endometriosis [51-56,58]. One study involved patients with a self-reported diagnosis [57], and one involved patients with a suspected diagnosis [49].

Countries

All the articles selected come from Western countries. Of these, the majority (5/10, 50%) were conducted in Europe, including 2 in France [52,53] and 3 in Germany [54-56]. Additionally, 4 studies took place in Australia [50,51,57,58]. Finally, 1 study has been carried out in Canada [49].

Types of DHIs

Of the 10 studies selected, the analysis revealed six distinct DHIs: (1) via a website [49], (2) health information recommendation features within web-based health applications [50], (3) an autonomous physical activity session using virtual reality video games for pain management [51], (4) a second intervention uses virtual reality and sensory stimuli for relaxation purposes to support patients in pain management (eg, Endocare) [52,53], (5) via a mobile application (ie, Endo-APP) [54-56], and (6) an SMS text messaging intervention (ie, EndoSMS) [57,58]. The first 3 DHIs were the subject of a single

XSL•FO
RenderX

publication [49-51]. The DHI proposed by Abdulai et al [49] is a website that aims to provide information on dyspareunia and its management (eg, treatment) and to combat the stigma associated with sexual pain. The DHI of Li et al [50] aims to propose personalized disease management strategies (eg, therapy or solutions to manage pain, dietary information, etc) based on symptoms (ie, reported in an initial questionnaire) and an algorithm. The third DHI [51] aims to reduce pelvic pain through physical activity performed either autonomously through virtual reality or supervised by telehealth in comparison with a control group.

The other 3 have been the subject of several studies and publications [52-58]. Of these last 3 DHIs, the first concerns a virtual reality device (ie, Endocare) compared with a 2D version, which was tested in the laboratory [52] and then autonomously at home in a second study [53]. The second DHI is a mobile app called Endo-APP, which was used in 3 studies [54-56]. Two of these studies were conducted by the same research team, although they did not cite each other [54,55]. This can be explained by the fact that the manuscript for the first study [54] was submitted in 2022 but accepted in 2024. The third study [56], which used Endo-APP, is a qualitative study focusing on patient experience following the prescription of this app by a health care professional. Finally, the third DHI was based on sending SMS text messages (ie, EndoSMS) and includes 2 studies [57,58]. Sherman et al [58] evaluated the acceptance and feasibility of a 3-month SMS text messaging intervention designed in a previous study [57].

Objectives, Interfaces, and Functionalities of the 6 DHIs Studied

The specific objectives pursued by these 6 DHIs vary from study to study. A total of 5 (83%) DHIs aimed to help manage pain or painful symptoms [49-56]. Among these 5 DHIs, 2 were based on a website but differ in their functionalities, with specific sections (eg, information on endometriosis, causes and treatments available for dyspareunia, and frequently asked questions) for Abdulai et al [49] and suggestions (eg, relaxation methods and dietary advice) based on initial responses to a questionnaire [50]. The other 2 DHIs used virtual reality but do not rely on the same functionalities. In the case of Lutfi et al [51], the interface and functionalities of the DHI depended on the participants' choices (eg, relaxation application and video games). For Merlot et al [52,53], Endocare's immersive intervention included auditory stimulation (ie, alpha and theta binaural beats and sounds inspired by nature) and visual stimulation by moving a laterally moving sphere. Finally, the Endo-APP mobile app was designed to help with this management. It features modules for monitoring symptoms, disease management (eg, educational content with articles and videos, exercise guides, and psychosocial support), or a personalized emergency plan for dealing with episodes of intense pain [54-56].

Overall, 4 (67%) DHIs aimed to provide information about endometriosis and strategies for managing the disease. They include the DHI proposed by Abdulai et al [49], that of Li et al [50], and the Endo-APP mobile app [54-56]. Another DHI targets this objective by being based on a unidirectional SMS text messaging broadcast, that is, with no possible response. The functionalities include the sending of 4 SMS text messages per week covering a variety of themes (eg, emotional health, social support, and information about the disease), and for some women, the possibility of adding a reminder to follow treatment. These authors also specify that this intervention can be personalized (ie, first name and time of receipt of messages) [57,58].

Furthermore, 2 (33%) interventions were intended to provide emotional and psychosocial support (ie, Endo-APP and EndoSMS). Finally, only the intervention proposed by Abdulai et al [49] sought to address the stigma associated with sexual pain. This information is summarized in Multimedia Appendix 2.

Patient Feedback on DHI Design and Deployment

Of the 6 DHIs studied in this systematic review, 3 were developed with the support of women with endometriosis [49,50,57]. For the other 3 DHIs (ie, the virtual reality and telehealth interventions of Lutfi et al, Endocare, and Endo-APP), the articles do not mention patient involvement during development. However, 4 DHIs studied in 5 articles [49,50,56-58] were evaluated by women with endometriosis. These were the 3 DHIs developed with patients [49,50,57,58] and 1 DHI evaluated after it was approved for prescribing to patients [56]. Overall, 3 DHIs were evaluated with quantitative analysis [49,50,57,58] and 3 were analyzed with qualitative analysis [49,56-58]. More precisely, the DHI proposed by Abdulai et al [49] was evaluated both quantitatively and qualitatively, as was the DHI EndoSMS [57,58]. Endo-APP was evaluated qualitatively in the study by Zugaj et al [56] and the DHI by Li et al [50] quantitatively. The other DHIs (ie, Endocare and Lutfi et al DHI) do not seem to have been evaluated by the participants.

Quantitative evaluations allow objectifying patients' user experience. A total of 3 DHIs were evaluated using this method. For example, the website proposed by Abdulai et al [49] is perceived as easy to use. Similarly, the study of Li et al [50] indicates that a quantitative evaluation of each proposed content was carried out. The details of each content are not reported. However, these authors noted a good user experience, although the number of returns was lower than expected. Similarly, Sherman et al [57] reported that the information provided in SMS text messages was useful and accessible to the majority of women with endometriosis, and so on. The same is true of the evaluation of this DHI under real-life conditions in terms of feasibility (eg, low attrition rate and few unsubscribes) and content acceptance [58].

Qualitative analysis identifies the limitations and potential for improvement of DHIs. A total of 3 DHIs were assessed using this methodology. For example, the study by Abdulai et al [49] enabled the research team to identify some minor problems (eg, comprehension of certain content such as graphics and icons), as well as features to be added (eg, search bar) or improved (eg, visibility of information on the home page). The study of Sherman et al [58] was positive overall, but some patients highlighted improvements (eg, disease information considered too general) and a greater need for personalization of the

XSL•FO
RenderX

intervention (eg, content, frequency of SMS text messages, etc). Similarly, although positive, the study by Zugaj et al [56] highlights some limitations of the app, its content, or functionality (eg, time required for correct and regular use, concerns about security and data use, and content deemed redundant). They also pointed out missing functionalities (eg, possibility of exchanging information with peers) and made suggestions for improvement (eg, integration of information on diseases concomitant with endometriosis). Other DHIs have not been evaluated in terms of their development. However, the study by Merlot et al [53] showed that patients who received the intervention were significantly more satisfied than women in the control group.

Efficiency to Manage Endometriosis

This systematic review highlights different types of benefits, such as physical, psychosocial, and self-management. A total of 5 studies [51-53,55,56] involving 3 different DHIs (ie, virtual reality and teleconsultation, Endocare, and Endo-APP) reported physical benefits. Regarding fatigue, the results are not homogeneous. Studies using the Endo-APP app [55] reported a significant reduction in fatigue compared to a control group. A DHI based on virtual reality (ie, Endocare) led to a reduction in fatigue, but there was no significant difference between the groups [53]. Finally, the other studies did not report any benefits in terms of fatigue. With regard to pain, these 5 studies reported a reduction in pain [51-53,55,56], regardless of the DHI used. It therefore seems that apps and virtual reality are relevant devices for managing physical symptoms, and specifically pain.

Regarding psychosocial benefits, 6 studies have reported benefits in this field. An improvement in quality of life was reported in 3 studies [53-55] based on 2 distinct DHIs (Endocare and Endo-APP). In the Endocare study (ie, virtual reality), there were no significant differences between the groups [53]. The other 2 studies reported significant improvements in Endo-APP users compared to the control group [54,55]. An improvement in self-efficacy in managing the disease and its symptoms was reported in 2 studies [55,56] involving a DHI (ie, Endo-APP). One study showed that patients found the website nonstigmatizing despite the subject matter [49]. Stress reduction was reported in 1 study [53] with no significant difference between groups, and 1 DHI has encouraged the adoption of active coping strategies [58]. Depressive symptoms were significantly reduced in 1 study [55], and an improvement in mood was reported in another [58]. A reduction in the nocebo effects associated with internet searches was observed in 1 study [56]. Better acceptance of the disease by certain women was reported in 1 study [56]. The same applies to the reduction of certain fears [56]. One study [58] reported a reduction in feelings of isolation and loneliness. The same applies to the increase in self-compassion [58].

In terms of the benefits of self-management, 3 studies report benefits, and 3 studies report on medication management. Overall, 3 types of DHIs (mobile app with personalized advice, EndoSMS, and Endo-APP) make it easier to access relevant information about the disease and change patients' perceptions of it [50,56,58]. Studies about medication management all report beneficial effects, whatever the DHI (ie, Endo-APP, Endocare, and EndoSMS), related to the use of analgesics, for example [53,55], or to remembering to take treatment [58]. This information is summarized in Multimedia Appendix 2.

Analysis of Selected Studies Based on the DEDHI Framework

Phase 1: Preparation

With the exception of Zugaj et al [56], all the studies mentioned at least 1 criterion corresponding to this first phase. However, there are major disparities between studies, with the criteria being met in 10%, with 1 in 10 criteria (eg, Rohloff et al [54]) to 80%, with 8 in 10 criteria (eg, Abdulai et al [49]) of cases. Among the criteria, a review of existing justificatory knowledge and feasibility were the most frequently reported in the studies. In contrast, criteria such as safety, privacy, and security were rarely reported.

Phase 2: Optimization

A total of 3 studies [50,51,54] reported criteria for the second phase. Li et al [50] and Rohloff et al [54] reported 54% (7/13) of the criteria, while the last study [51] reported 77% (10/13) of them. Criteria such as effectiveness, perceived benefits, and content quality were consistently reported. In contrast, perceived enjoyment was never mentioned, and criteria such as service quality, safety, privacy, and security were rarely taken into account. The other 7 studies [49,52,53,55-58] do not mention any elements relating to this second phase.

Phase 3: Evaluation

A total of 4 studies reported between 60% (6/10) and 80% (8/10) of the criteria for this third phase [50,52,53,55]. The other 6 studies [49,51,54,56-58] do not mention any of the criteria relating to this third phase of the DEDHI framework. As far as the evaluation criteria are concerned, effectiveness and perceived benefits are reported by all those meeting the criteria for this phase. Personalization was reported by only 1 of the 4 studies [50]. Moreover, privacy and security are rarely considered, as is the quality of service.

Phase 4: Implementation

A total of 2 articles met 77% (10/13) [50] and 62% (8/13) [56] of the criteria for the implementation phase of the DEDHI framework. The other 8 studies [49,51-55,57,58] did not mention any of the criteria for this phase. In these 2 studies, safety and service quality criteria are never mentioned. However, criteria such as personalization, perceived benefits, and content quality are the most regularly mentioned by studies responding to criteria from this fourth phase. The results of this section are summarized in Table 2.

XSL•FO
RenderX

Table 2. Analysis with Design and Evaluation of Digital Health Interventions (DEDHI) framework.

Phases, subphases, and subactivity	Abdulai et al [49]	Li et al [50]	Lutfi et al [51]	Merlot et al [52]	Merlot et al [53]	Rohloff et al [54]	Rohloff et al [55]	Sherman et al [57]	Sherman et al [58]	Zugaj et al [56]
Preparation phase (phase 1)										
n (%)[a]	8 (80)	8 (80)	2 (20)	1 (10)	1 (10)	1 (10)	1 (10)	7 (70)	7 (70)	0 (0)
Goals and tasks										
Review existing justificatory knowledge	Y[b] (p.[c]2)	Y (p.8)	Y (p.2)	Y (p.2)	Y (p.2)	Y (p.1158)	Y (p.1)	Y (p.2-3)	Y (p.2)	N[d]
Conceptual model	Y (p.2)	Y (p.8)	Y (p.4)	N	N	N	N	Y (p.7)	Y (p.2)	N
Feasibility and acceptability	Y (p.2)	Y (p.9)	N	N	N	N	N	Y (p.8)	Y (p.2)	N
Identify and optimization criterion	Y (p.5)	Y (p.9)	N	N	N	N	N	Y (p.8)	Y (p.7)	N
Technical maturity										
Research prototype that provides basic functionality	Y (p.4)	Y (p.9)	N	N	N	N	N	Y (p.8)	Y (p.2)	N
Evaluation criteria										
Ease of use	Y (p.7)	Y (p.9)	N	N	N	N	N	Y (p.9)	N	N
Adherence	Y (p.5)	N	N	N	N	N	N	N	Y (p.6)	N
Personalization	N	Y (p.8)	N	N	N	N	N	Y (p.8)	Y (p.2)	N
Safety	N	N	N	N	N	N	N	N	N	N
Privacy and security	Y (p.11)	Y (p.7)	N	N	N	N	N	N	N	N
Optimization phase (phase 2)										
n (%)	0 (0)	7 (54)	10 (77)	0 (0)	0 (0)	7 (54)	0 (0)	0 (0)	0 (0)	0 (0)
Goals and tasks										
Conduct optimization trials	N	Y (p.9)	Y (p.2)	N	N	Y (p.1158)	N	N	N	N
Identify the best DHI[e] configuration	N	N	Y (p.2)	N	N	N	N	N	N	N
Technical maturity										
Elaborated research prototype	N	Y (p.9)	Y (p.4)	N	N	Y (p.1158)	N	N	N	N
Evaluation criteria										
Effectiveness	N	Y (p.12)	Y (p.5)	N	N	Y (p.1161)	N	N	N	N
Perceived benefit	N	Y (p.9)	Y (p.5)	N	N	Y (p.1161)	N	N	N	N
Content quality	N	Y (p.9)	Y (p.4)	N	N	Y (p.1158)	N	N	N	N
Personalization	N	Y (p.9)	Y (p.4)	N	N	N	N	N	N	N
Perceived enjoyment	N	N	N	N	N	N	N	N	N	N
Aesthetics	N	N	Y (p.4)	N	N	N	N	N	N	N
Adherence	N	N	Y (p.4)	N	N	Y (p.1161)	N	N	N	N

XSL•FO
RenderX

Phases, subphases, and subactivity	Abdulai et al [49]	Li et al [50]	Lutfi et al [51]	Merlot et al [52]	Merlot et al [53]	Rohloff et al [54]	Rohloff et al [55]	Sherman et al [57]	Sherman et al [58]	Zugaj et al [56]
Service quality	N	N	N	N	N	Y (p.1158)	N	N	N	N
Safety	N	N	Y (p.5)	N	N	N	N	N	N	N
Privacy and security	N	Y (p.7)	N	N	N	N	N	N	N	N
Evaluation phase (phase 3)										
n (%)	0 (0)	6 (60)	0 (0)	8 (80)	7 (70)	0 (0)	6 (60)	0 (0)	0 (0)	0 (0)
Goals and tasks										
Confirm the effectiveness of an optimized DHI	N	N	N	Y (p.2)	Y (p.2)	N	Y (p.2)	N	N	N
Technical maturity										
Elaborated research prototype	N	Y (p.12)	N	Y (p.3)	Y (p.3)	N	Y (p.1)	N	N	N
Evaluation criteria										
Effectiveness	N	Y (p.12)	N	Y (p.8)	Y (p.7)	N	Y (p.5)	N	N	N
Perceived benefit	N	Y (p.12)	N	Y (p.8)	Y (p.7)	N	Y (p.5)	N	N	N
Adherence	N	N	N	Y (p.6)	Y (p.6)	N	Y (p.7)	N	N	N
Personalization	N	Y (p.12)	N	N	N	N	N	N	N	N
Service quality	N	N	N	Y (p.4)	N	N	N	N	N	N
Safety	N	N	N	Y (p.10)	Y (p.13)	N	N	N	N	N
Privacy and security	N	Y (p.7)	N	N	N	N	N	N	N	N
Accountability	N	Y (p.12)	N	Y (p.3)	Y (p.2)	N	Y (p.3)	N	N	N
Implementation phase (phase 4)										
n (%)	0 (0)	10 (77)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	8 (62)
Goals and tasks										
Develop a DHI product that can be implemented at a large scale in the health care market	N	Y (p.12)	N	N	N	N	N	N	N	N
Assess the long-term outreach and efficiency	N	Y (p.12)	N	N	N	N	N	N	N	Y (p.2254)
Update the DHI	N	Y (p.12)	N	N	N	N	N	N	N	Y (p.2260)
Technical maturity										
Elaborated research prototype	N	Y (p.12)	N	N	N	N	N	N	N	Y (p.2254)
Evaluation criteria										
Adherence	N	N	N	N	N	N	N	N	N	Y (p.2259)
Personalization	N	Y (p.12)	N	N	N	N	N	N	N	Y (p.2254)

XSL•FO
RenderX

Phases, subphases, and subactivity	Abdulai et al [49]	Li et al [50]	Lutfi et al [51]	Merlot et al [52]	Merlot et al [53]	Rohloff et al [54]	Rohloff et al [55]	Sherman et al [57]	Sherman et al [58]	Zugaj et al [56]
Perceived benefit	N	Y (p.12)	N	N	N	N	N	N	N	Y (p.2258-2259)
Content quality	N	Y (p.9)	N	N	N	N	N	N	N	Y (p.2254)
Ethics	N	Y (p.5)	N	N	N	N	N	N	N	N
Service quality	N	N	N	N	N	N	N	N	N	N
Safety	N	N	N	N	N	N	N	N	N	N
Privacy and security	N	Y (p.7)	N	N	N	N	N	N	N	N
Accountability	N	Y (p.12)	N	N	N	N	N	N	N	Y (p.2254)

[a]Percentage of completion for each phase, determined by the ratio between the number of criteria present and the total number of criteria expected for that phase.

[b]The criterion is identified, either explicitly stated or reasonably inferred from the content of the article.

[c]"p." represents the page of the article where the criterion is mentioned.

[d]The criterion is not identified or cannot be reasonably inferred.

[e]DHI: digital health intervention.

Discussion

Overview

The aim of this systematic review was to characterize the design, quality, and efficiency of the digital technologies used to support women with endometriosis. This review highlights four important points: (1) the literature on this topic is at an early stage, although it is developing rapidly, as is the case for other chronic diseases; (2) digital interventions show physical, psychological, and self-management benefits, but the small number of studies and methodological variability limit the reliability of current evidence; (3) few studies mobilized a methodological framework to develop their intervention; and (4) analysis of the corpus using the DEDHI framework revealed that the majority of the studies essentially reported criteria relating to the first phase (development), leaving the later phases as yet unexplored.

Results of this systematic review highlight recent literature that is embedded in the growth of digital technology. Indeed, the least recent article selected in this corpus was published in 2022. In addition, the studies are mainly based in Western countries that have implemented policies to address this disease (eg, France [61], Canada [62], and Australia [63]), for instance, by supporting the funding of targeted research projects. This geographical focus may be linked to the fact that the disease has received increasing media coverage in recent years [64]. This geographical disparity may therefore contribute to the lack of international literature. This may result in an incomplete understanding of patient needs on a global scale. At the same time, the recent increase in DHIs and related publications follows the trend observed in endometriosis research, which has been steadily increasing since 2017 [65]. Finally, as reported in the results section, most of the studies conducted and the DHIs designed have been developed by the same research teams, which may result in some limitations of the literature on the subject. This literature is also the result of the evolution in the recommendations concerning the management of the negative effects of illness [4,5,66,67]. These recommendations were aimed, in particular, at improving the operational efficiency (eg, time management) of health care professionals [32,68] and the optimization of care and the quality of service provided to patients [69]. A growing trend is the use of digital health technology to enable patients with chronic illness to manage their disease over a long period of time [70], while developing their disease management skills [71]. These technologies are regularly used because of the freedom they offer (eg, use at home, at any time, or on any day) [72], both for the practitioner providing the care and for the patients benefiting from it. This systematic review illustrates precisely this point, since a number of applications and technologies (eg, websites, mobile apps, SMS text messaging, and virtual reality) are being used to optimize support for endometriosis. The use of digital technology for this disease is part of a wider movement concerning chronic pathologies, reflecting a paradigm shift in which patients become the central actors in their own care [73]. Several studies in this review indicate that patients were involved from the tool's design stage [49,50,57,58] or during the evaluation phases of the tool designed to support them [49,50,56-58]. These studies show just how valuable it can be to involve patients from the earliest stages of device development—for example, to increase patient satisfaction [74]. The results of this study suggest that digital technologies can be useful in the treatment of endometriosis. However, they do not appear to be sufficient on their own to meet patients' needs and could even have deleterious effects. Indeed, many women already report a sense of isolation and lack of recognition, both socially and within the health care system [10,75], which could be accentuated through digital technology. In this context, DHI could be seen as a complement to traditional care as part of a hybrid care model. This approach would exploit the advantages

XSL•FO
RenderX

of digital tools while preserving the essential human dimension of care [76].

Studies in this systematic review reported physical, psychological, and self-management benefits of endometriosis. Indeed, some of the digital interventions in this corpus using digital technology [51-55] reported a reduction in pain, the main symptom of endometriosis [77]. Reduced pain can improve patients' overall quality of life [78] and, therefore, has beneficial repercussions on other areas of patients' lives, such as absenteeism from work [14] and social relationships [79]. Some studies in this corpus [54,55] reported a more global improvement in physical quality of life (eg, reduction in chronic pain) and psychological quality of life (eg, increased emotional well-being), while 1 study reported a reduction in feelings of loneliness and isolation [58]. Studies [53,55,58] have also shown a reduction in the use of analgesics. This could reduce treatment costs, which is particularly important in the context of endometriosis [80]. The DHIs presented in this systematic review seem to respond to the main challenges posed by endometriosis, whether related to the management of the disease or to its main consequences on physical, psychological, and social health, among others [81,82]. Furthermore, the articles reported an increase in patients' feeling of being able to manage their symptoms [55,56]. Yet, self-efficacy is a key variable, having been shown in the literature to be an important determinant of the adoption and maintenance of health-enhancing behaviors (eg, health belief model) [83,84]. Indeed, a person with a high degree of self-efficacy will be more inclined to respond to a perceived threat [85] and to adopt relevant strategies to manage the symptoms of endometriosis. For example, the patient will be more likely to adopt health behaviors that improve her quality of life, such as physical activity [86] or a balanced diet [87]. To promote behavior change and management of this disease, future DHIs could draw on psychological and behavioral theories. Self-determination theory [88], the transtheoretical model of change [89], and the taxonomy of behavior change techniques developed by Michie et al [90] could offer valuable leads for the design of DHIs. Such models can help identify motivating factors, stages of change, and techniques adapted to users' needs. Integrating these theoretical frameworks could improve the effectiveness and reproducibility of interventions. Finally, they provide a better understanding of the mechanisms of action underlying long-term user commitment and adherence. For instance, self-determination theory emphasizes autonomy, competence, and relationship as essential components of sustained motivation [91]. The transtheoretical model enables content to be adapted to the user's stage of change. In this way, each individual could have a personalized intervention [92], as requested by some patients (eg, [58]). Using the taxonomy of behavior change techniques would enable a structured approach to the selection and presentation of behavioral strategies.

Despite these advantages, a number of limitations need to be recognized. Most of the interventions examined were of short duration (up to 12 weeks [55]). It is therefore difficult to assess the durability of beneficial effects or the stability of user engagement over time. However, these elements appear to be crucial in understanding chronic diseases, such as endometriosis [93]. In addition, several studies [50,54-58] have proposed that intervention should be holistic, encompassing several dimensions of patient well-being (eg, dietary monitoring, stress management, and physical activity). While such an approach is recommended in the literature [94], this does not allow for the identification of the specific part played by the use of digital technology. Finally, as reported by Zugaj et al [56], the use of digital technologies does not seem to be suitable for all patients. This is why it seems necessary to pay particular attention to the personalization of devices [37,95] and take advantage of the variety of care modalities available in order to offer optimal support to each patient [94].

Among all the included studies and the 6 identified DHIs, 3 DHIs were developed with patient involvement [49,50,57]. Of these, only 2 articles [49,50] reported using a methodological framework to guide the development of a digital intervention and progressed to large-scale deployment. Nevertheless, 1 study [57] drew on various theories from social psychology (eg, theory of planned behavior [96]) to construct part of its intervention. After analyzing the studies using the DEDHI framework, it appears that almost all of them (9/10) focused on the criteria of phase 1 (preparation), leaving the subsequent phases unexplored. Furthermore, a detailed analysis of the phases revealed wide disparities in the way the criteria were addressed, particularly in this first phase. Thus, in the first phase, the criterion most often reported was the use of a literature review to justify technological choices and their positive effects on chronic pathologies. For example, Merlot et al [52] considered various meta-analyses on the use of virtual reality for chronic pain management in different pathologies, and Rohloff et al [54,55] did the same for the use of smartphones in pathologies with chronic pain. Nevertheless, even if the use of literature reviews is an essential step to support the development of intervention devices, it remains insufficient if criteria such as acceptability and acceptance of these devices are not taken into account because, as recommended by the DEDHI framework [37], they make it possible to adjust the intervention to the real expectations and needs of users. Moreover, considering "individual acceptability" would enable the adoption of a particular support strategy depending on the individual's specific relationship with the technology [97]. On the other hand, the studies do not seem to have adapted their tools to the various standards and regulations in force concerning the use of digital technology in health (eg, General Data Protection Regulation). One explanation may lie in the fact that the research teams are more often than not made up of health care professionals and do not always include specialists in the development and implementation of these digital tools (eg, engineering research teams, lawyers, psychologists, and ergonomists). Yet, interdisciplinarity is generally considered to be one of the essential elements in medical innovation, particularly to meet the different needs of the market [98]. The ever-increasing call for research into health care innovations [99] may lead research teams to neglect certain essential stages of development (eg, optimization, evaluation, and implementation), which can compromise the ownership and sustainability of the intervention device and associated benefits [37]. Li et al [50], who highlight the advantages of mobilizing a clearly defined methodological framework to think through the implementation of a technology

XSL•FO
RenderX

with potential users, support this view. Yet, some of the included studies collected patient perceptions during the development or trial phases of DHIs. For example, 3 studies [49,56,58] have highlighted that patient satisfaction and engagement were linked to perceptions of ease of navigation, clarity of content, and perceived credibility of information. Conversely, problems such as excessive text content [49], poor section visibility [50], or a lack of personalization [58] were seen as potential barriers to engagement with the intervention or continued use. These results, therefore, underline the need to systematically involve the various stakeholders in the development process. In addition, they also highlight the importance of user testing in order to limit identified obstacles and support potential long-term use.

Strengths and Limitations of This Review

This systematic review has 4 main strengths. First, the PRISMA methodology and its stages were applied [41]. Second, the quality of each study was assessed with an adjusted evaluation of the risk of bias according to the experimental design used [100]. Third, each article was evaluated by several researchers, a discussion was organized to harmonize assessments, and an intraclass correlation coefficient calculation [101] was performed for the risk of bias. Fourth, several databases were mobilized to improve the comprehensiveness of the search [100].

Several limitations can also be identified. First, a meta-analysis could not be carried out due to the heterogeneity of experimental designs implemented, the measurements performed, and the results observed in the review analysis [102]. This methodological variability has limited the possibility of producing a quantitative synthesis of results or drawing definitive conclusions about the effectiveness of specific digital interventions. Thus, future systematic reviews or meta-analyses on the effects of digital interventions will need to be conducted, particularly on variables of interest in the context of endometriosis (eg, perceived pain, quality of life, and anxiety). Second, the published literature often reports positive and significant results [103], and there is not always a control group to put the results into perspective. This bias may have influenced the overall impression of the effectiveness or feasibility of the DHIs presented in this review. As a result, these findings may provide an overestimated picture of the actual benefits or readiness of these tools for wider implementation, while underrepresenting the difficulties or limitations encountered in practice. Third, the growing interest in digital and associated technologies is such that the search equation used may not have identified some articles mobilizing terminologies different from ours [104]. Fourth, following the selection process, 10 studies were retained out of the 381 articles identified. While this highlights the current state of research, it also limits the scope of the conclusions of this systematic review and the generalizability of the results. However, the recent increase in the number of studies on the subject calls for clarification of the methodological framework used to develop interventions and the associated digital tools, which should help to improve the reproducibility of interventions. Finally, the studies in this review were mainly conducted in Western countries, which raises questions about the transferability of the results, particularly given the difficulties of accessing and using DHIs linked to cultural, economic, and physical factors [105]. It is possible that in some parts of the world, access to this type of technology and the internet is more limited, which could reduce the reach of these digital solutions. For example, without the support of health authorities, the financial cost of accessing these digital resources may be too high and restrict access. Unequal access to digital resources may also hinder the development of digital literacy (ie, digital skills and the skills to understand and use content available on the internet [106]). These elements are identified as potential barriers to the use of DHIs [107]. It therefore seems necessary for international research teams to focus on the development of digital health devices. The approach could be to develop or adapt existing frameworks to the cultural and contextual specificities of their implementation to ensure greater accessibility and usability for diverse groups.

Research Implications

In light of these challenges, future research could potentially focus on several key priorities. First, longitudinal studies appear necessary to better understand the mechanisms underlying the sustained use of digital health tools [94]. Second, more rigorous and interdisciplinary research (eg, randomized controlled trials) with larger samples and standardized outcome measures would also appear to be relevant. Carried out in this way, they would ensure reproducibility while strengthening the evidence base for the effectiveness of digital interventions. Third, qualitative studies exploring patients' experiences, expectations, and perceptions of digital health technologies could be carried out. This could complement the results of quantitative evaluations of content or user experience [108]. Fourth, cross-cultural studies could be conducted to assess the transferability of these interventions for different patient profiles (eg, age, sociodemographics, and digital literacy), as well as for different cultural, geographical, and socioeconomic contexts [109]. Fifth, future studies could explore hybrid models of care to better meet patients' needs and evaluate the effectiveness of this type of support. Sixth, future studies could be based on the study of behavior change theories or on the identification of behavior change techniques used by DHI. Finally, future studies should systematically apply methodological frameworks, such as DEDHI, to frame the development and implementation process. All relevant phases and criteria would then be transparently reported in their publications.

Conclusion

This systematic review highlights the current interest in using digital technology to support patients with endometriosis. It appears promising and offers advantages in terms of patient care. However, the development of digital devices is rarely based on a methodological framework that would direct each stage of development to maximize patient involvement and, ultimately, improve their overall quality of life.

Overall, this systematic review has highlighted a number of elements for the development of digital devices for the management of endometriosis, from the initial design phases through to market implementation and appropriation by patients. Three points seem crucial to develop interventions mobilizing digital technology and encourage patients with endometriosis

XSL•FO
RenderX

to appropriate these tools: (1) mobilizing a methodological framework tailored to the context and specific to the development of support with digital devices, (2) relying on a multidisciplinary team (ie, doctors, engineers, ergonomic psychologists, and lawyers) capable of addressing every aspect of the development of the digital device, and (3) involving all stakeholders in the management of endometriosis (ie, patients, algologists, gynecologists, psychologists, and physiotherapists) in the design. These few points could serve as a guideline for outlining the contours of a global research reflection on the development of interventions using digital tools in the context of endometriosis, aimed at improving the quality of patient care. Developed using this approach, these digital tools place patients at the heart of the interventions by offering them not only concrete tools to better manage their health but also a space to share their experiences with other patients concerned. These elements help to foster health empowerment and, ultimately, pain management [110]. These digital tools can help to change the way patients are supported, moving away from "traditional medical care" to "holistic care," which takes into account all the dimensions of a patient's health—physical, mental, and social.

Acknowledgments

This study is supported by the Agence Nationale de la Recherche (ANR) within the framework of the Programme d'Investissements d'Avenir, École Universitaire de Recherche DIGISPORT project (ANR-18-EURE-0022) and by the Bretagne region.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA checklist.
[DOCX File , 32 KB - humanfactors_v12i1e71859_app1.docx]

Multimedia Appendix 2
Details of studies included and results of interventions (n=10).
[DOCX File , 20 KB - humanfactors_v12i1e71859_app2.docx]

References

1. International Working Group of AAGL, ESGE, ESHREWES, Tomassetti C, Johnson NP, Petrozza J, Abrao MS, Einarsson JI, et al. An international terminology for endometriosis, 2021. Hum Reprod Open 2021;2021(4):hoab029 [FREE Full text] [doi: 10.1093/hropen/hoab029] [Medline: 34693033]
2. Lamceva J, Uljanovs R, Strumfa I. The main theories on the pathogenesis of endometriosis. Int J Mol Sci 2023;24(5):4254 [FREE Full text] [doi: 10.3390/ijms24054254] [Medline: 36901685]
3. Endometriosis. World Health Organization. 2023. URL: https://www.who.int/news-room/fact-sheets/detail/endometriosis [accessed 2024-02-27]
4. Ellis K, Munro D, Clarke J. Endometriosis is undervalued: a call to action. Front Glob Womens Health 2022;3:902371 [FREE Full text] [doi: 10.3389/fgwh.2022.902371] [Medline: 35620300]
5. Agarwal S, Foster WG, Groessl E. Rethinking endometriosis care: applying the chronic care model via a multidisciplinary program for the care of women with endometriosis. Int J Womens Health 2019;11:405-410. [doi: 10.2147/ijwh.s207373]
6. Szypłowska M, Tarkowski R, Kułak K. The impact of endometriosis on depressive and anxiety symptoms and quality of life: a systematic review. Front Public Health 2023;11:1230303 [FREE Full text] [doi: 10.3389/fpubh.2023.1230303] [Medline: 37744486]
7. Ashrafi M, Sadatmahalleh SJ, Akhoond MR, Talebi M. Evaluation of risk factors associated with endometriosis in infertile women. Int J Fertil Steril 2016;10(1):11-21 [FREE Full text] [doi: 10.22074/ijfs.2016.4763] [Medline: 27123195]
8. Zondervan KT, Becker CM, Missmer SA. Endometriosis. N Engl J Med 2020;382(13):1244-1256. [doi: 10.1056/NEJMra1810764] [Medline: 32212520]
9. Surrey ES, Soliman AM, Johnson SJ, Davis M, Castelli-Haley J, Snabes MC. Risk of developing comorbidities among women with endometriosis: a retrospective matched cohort study. J Womens Health (Larchmt) 2018;27(9):1114-1123. [doi: 10.1089/jwh.2017.6432] [Medline: 30070938]
10. Facchin F, Saita E, Barbara G, Dridi D, Vercellini P. "Free butterflies will come out of these deep wounds": a grounded theory of how endometriosis affects women's psychological health. J Health Psychol 2018;23(4):538-549 [FREE Full text] [doi: 10.1177/1359105316688952] [Medline: 28810386]
11. Laganà AS, Condemi I, Retto G, Muscatello MRA, Bruno A, Zoccali RA, et al. Analysis of psychopathological comorbidity behind the common symptoms and signs of endometriosis. Eur J Obstet Gynecol Reprod Biol 2015;194:30-33. [doi: 10.1016/j.ejogrb.2015.08.015] [Medline: 26319653]

XSL•FO
RenderX

12. van Barneveld E, Manders J, van Osch FHM, van Poll M, Visser L, van Hanegem N, et al. Depression, anxiety, and correlating factors in endometriosis: a systematic review and meta-analysis. J Womens Health (Larchmt) 2022;31(2):219-230. [doi: 10.1089/jwh.2021.0021] [Medline: 34077695]
13. Volker C, Mills J. Endometriosis and body image: comparing people with and without endometriosis and exploring the relationship with pelvic pain. Body Image 2022 Dec;43:518-522. [doi: 10.1016/j.bodyim.2022.10.014] [Medline: 36375365]
14. Soliman AM, Coyne KS, Gries KS, Castelli-Haley J, Snabes MC, Surrey ES. The effect of endometriosis symptoms on absenteeism and presenteeism in the workplace and at home. J Manag Care Spec Pharm 2017;23(7):745-754 [FREE Full text] [doi: 10.18553/jmcp.2017.23.7.745] [Medline: 28650252]
15. De Graaff AA, D'Hooghe TM, Dunselman GAJ, Dirksen CD, Hummelshoj L, WERF EndoCost Consortium, et al. The significant effect of endometriosis on physical, mental and social wellbeing: results from an international cross-sectional survey. Hum Reprod 2013;28(10):2677-2685 [FREE Full text] [doi: 10.1093/humrep/det284] [Medline: 23847114]
16. Fagervold B, Jenssen M, Hummelshoj L, Moen MH. Life after a diagnosis with endometriosis: a 15 years follow-up study. Acta Obstet Gynecol Scand 2009;88(8):914-919 [FREE Full text] [doi: 10.1080/00016340903108308] [Medline: 19568961]
17. Grogan S, Turley E, Cole J. 'So many women suffer in silence': a thematic analysis of women's written accounts of coping with endometriosis. Psychol Health 2018;33(11):1364-1378. [doi: 10.1080/08870446.2018.1496252] [Medline: 30306810]
18. Horne AW, Missmer SA. Pathophysiology, diagnosis, and management of endometriosis. BMJ 2022;379:e070750. [doi: 10.1136/bmj-2022-070750] [Medline: 36375827]
19. Stowell E, Lyson M, Saksono H, Wurth R, Jimison H, Pavel M, et al. Designing and evaluating mHealth interventions for vulnerable populations: a systematic review. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems Montreal QC Canada: ACM; 2018 Presented at: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems Montreal QC; 2018 June 10; Canada p. 1-17. [doi: 10.1145/3173574.3173589]
20. Blandford A, Wesson J, Amalberti R, AlHazme R, Allwihan R. Opportunities and challenges for telehealth within, and beyond, a pandemic. Lancet Glob Health 2020;8(11):e1364-e1365 [FREE Full text] [doi: 10.1016/S2214-109X(20)30362-4] [Medline: 32791119]
21. Quach S, Michaelchuk W, Benoit A, Oliveira A, Packham TL, Goldstein R, et al. Mobile heath applications for self-management in chronic lung disease: a systematic review. Netw Model Anal Health Inform Bioinform 2023;12(1):25 [FREE Full text] [doi: 10.1007/s13721-023-00419-0] [Medline: 37305790]
22. Alruwaili MM, Shaban M, Elsayed Ramadan OM. Digital health interventions for promoting healthy aging: a systematic review of adoption patterns, efficacy, and user experience. Sustainability 2023;15(23):16503 [FREE Full text] [doi: 10.3390/su152316503]
23. De Witte NAJ, Joris S, Van Assche E, Van Daele T. Technological and digital interventions for mental health and wellbeing: an overview of systematic reviews. Front Digit Health 2021;3:754337 [FREE Full text] [doi: 10.3389/fdgth.2021.754337] [Medline: 35005695]
24. Senbekov M, Saliev T, Bukeyeva Z, Almabayeva A, Zhanaliyeva M, Aitenova N, et al. The recent progress and applications of digital technologies in healthcare: a review. Int J Telemed Appl 2020;2020:8830200 [FREE Full text] [doi: 10.1155/2020/8830200] [Medline: 33343657]
25. Okumu M, Logie CH, Byansi W, Cohen F, Nyoni T, Nafula CN, et al. eHealth literacy and digital health interventions: key ingredients for supporting the mental health of displaced youth living in the urban slums of Kampala, Uganda. Comput Hum Behav 2025;162:108434. [doi: 10.1016/j.chb.2024.108434]
26. Fan K, Zhao Y. Mobile health technology: a novel tool in chronic disease management. Intell Med 2022;2(1):41-47. [doi: 10.1016/j.imed.2021.06.003]
27. Hou IC, Lin HY, Shen SH, Chang KJ, Tai HC, Tsai AJ, et al. Quality of life of women after a first diagnosis of breast cancer using a self-management support mHealth app in Taiwan: randomized controlled Trial. JMIR Mhealth Uhealth 2020;8(3):e17084 [FREE Full text] [doi: 10.2196/17084] [Medline: 32130181]
28. Khusial RJ, Honkoop PJ, Usmani O, Soares M, Simpson A, Biddiscombe M, myAirCoach study group. Effectiveness of myAirCoach: a mHealth self-management system in asthma. J Allergy Clin Immunol Pract 2020;8(6):1972-1979.e8 [FREE Full text] [doi: 10.1016/j.jaip.2020.02.018] [Medline: 32142961]
29. Farzandipour M, Nabovati E, Sharif R, Arani MH, Anvari S. Patient self-management of asthma using mobile health applications: a systematic review of the functionalities and effects. Appl Clin Inform 2017;8(4):1068-1081 [FREE Full text] [doi: 10.4338/ACI-2017-07-R-0116] [Medline: 29241254]
30. Giudice LC, Horne AW, Missmer SA. Time for global health policy and research leaders to prioritize endometriosis. Nat Commun 2023;14(1):8028 [FREE Full text] [doi: 10.1038/s41467-023-43913-9] [Medline: 38049392]
31. Giudice LC, Oskotsky TT, Falako S, Opoku-Anane J, Sirota M. Endometriosis in the era of precision medicine and impact on sexual and reproductive health across the lifespan and in diverse populations. FASEB J 2023;37(9):e23130 [FREE Full text] [doi: 10.1096/fj.202300907] [Medline: 37641572]
32. Kraus S, Schiavone F, Pluzhnikova A, Invernizzi AC. Digital transformation in healthcare: analyzing the current state-of-research. J Bus Res 2021;123:557-567. [doi: 10.1016/j.jbusres.2020.10.030]
33. Marques ICP, Ferreira JJM. Digital transformation in the area of health: systematic review of 45 years of evolution. Health Technol 2019;10(3):575-586. [doi: 10.1007/s12553-019-00402-8]

XSL•FO
RenderX

34. Iribarren SJ, Akande TO, Kamp KJ, Barry D, Kader YG, Suelzer E. Effectiveness of mobile apps to promote health and manage disease: systematic review and meta-analysis of randomized controlled trials. JMIR Mhealth Uhealth 2021;9(1):e21563 [FREE Full text] [doi: 10.2196/21563] [Medline: 33427672]
35. Nouri R, R Niakan Kalhori S, Ghazisaeedi M, Marchand G, Yasini M. Criteria for assessing the quality of mHealth apps: a systematic review. J Am Med Inform Assoc 2018;25(8):1089-1098 [FREE Full text] [doi: 10.1093/jamia/ocy050] [Medline: 29788283]
36. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
37. Kowatsch T, Otto L, Harperink S, Cotti A, Schlieter H. A design and evaluation framework for digital health interventions. Inf Technol (De Gruyter Oldenbourg) 2019;61(5-6):253-263. [doi: 10.1515/itit-2019-0019]
38. Jacob C, Lindeque J, Klein A, Ivory C, Heuss S, Peter MK. Assessing the quality and impact of eHealth tools: systematic literature review and narrative synthesis. JMIR Hum Factors 2023;10:e45143 [FREE Full text] [doi: 10.2196/45143] [Medline: 36843321]
39. Mascarenhas E, Abboute A, Wayne R, Kozlakidis Z. Digital healthcare: technologies, technical and design challenges. In: Kozlakidis Z, Muradyan A, Sargsyan K, editors. Digitalization of Medicine in Low- and Middle-Income Countries: Paradigm Changes in Healthcare and Biomedical Research. Cham: Springer International Publishing; 2024:207-217.
40. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med 2009;6(7):e1000100 [FREE Full text] [doi: 10.1371/journal.pmed.1000100] [Medline: 19621070]
41. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
42. Moses JC, Adibi S, Shariful Islam SM, Wickramasinghe N, Nguyen L. Application of smartphone technologies in disease monitoring: a systematic review. Healthcare (Basel) 2021;9(7):889 [FREE Full text] [doi: 10.3390/healthcare9070889] [Medline: 34356267]
43. Lefebvre C, Glanville J, Briscoe S, Featherstone R, Littlewood A, Metzendorf MI, et al. Chapter 4: searching for selecting studies. In: Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al, editors. Cochrane Handbook for Systematic Reviews of Interventions version 651. London: Cochrane; 2025.
44. Ivey C, Crum J. Choosing the right citation management tool: EndNote, Mendeley, RefWorks, or Zotero. J Med Libr Assoc 2018;106(3):399-403. [doi: 10.5195/jmla.2018.468]
45. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016;5(1):210 [FREE Full text] [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
46. Johnson N, Phillips M. Rayyan for systematic reviews. J Electron Resour Librariansh 2018;30(1):46-48. [doi: 10.1080/1941126x.2018.1444339]
47. Hong QN, Fàbregues S, Bartlett G, Boardman F, Cargo M, Dagenais P, et al. The mixed methods appraisal tool (MMAT) version 2018 for information professionals and researchers. EFI 2018;34(4):285-291. [doi: 10.3233/EFI-180221]
48. Pluye P, Gagnon MP, Griffiths F, Johnson-Lafleur J. A scoring system for appraising mixed methods research, and concomitantly appraising qualitative, quantitative and mixed methods primary studies in mixed studies reviews. Int J Nurs Stud 2009;46(4):529-546. [doi: 10.1016/j.ijnurstu.2009.01.009] [Medline: 19233357]
49. Abdulai AF, Howard AF, Yong PJ, Noga H, Parmar G, Currie LM. Developing an educational website for women with endometriosis-associated dyspareunia: usability and stigma analysis. JMIR Hum Factors 2022;9(1):e31317 [FREE Full text] [doi: 10.2196/31317] [Medline: 35238789]
50. Li W, O'Hara R, Hull ML, Slater H, Sirohi D, Parker MA, et al. Enabling health information recommendation using crowdsourced refinement in web-based health information Applications: user-centered design approach and EndoZone informatics case study. JMIR Hum Factors 2024;11:e52027. [doi: 10.2196/52027]
51. Lutfi M, Dalleck LC, Drummond C, Drummond M, Paparella L, Keith CE, et al. A single session of a digital health tool-delivered exercise intervention may provide immediate relief from pelvic pain in women with endometriosis: a pilot randomized controlled study. Int J Environ Res Public Health 2023;20(3):1665 [FREE Full text] [doi: 10.3390/ijerph20031665] [Medline: 36767032]
52. Merlot B, Dispersyn G, Husson Z, Chanavaz-Lacheray I, Dennis T, Greco-Vuilloud J, et al. Pain reduction with an immersive digital therapeutic tool in women living with endometriosis-related pelvic pain: randomized controlled trial. J Med Internet Res 2022;24(9):e39531 [FREE Full text] [doi: 10.2196/39531] [Medline: 36129733]
53. Merlot B, Elie V, Périgord A, Husson Z, Jubert A, Chanavaz-Lacheray I, et al. Pain reduction with an immersive digital therapeutic in women living with endometriosis-related pelvic pain: at-home self-administered randomized controlled trial. J Med Internet Res 2023;25:e47869 [FREE Full text] [doi: 10.2196/47869] [Medline: 37260160]
54. Rohloff N, Rothenhöfer M, Götz T, Schäfer SD. Observational pilot study on the influence of an app-based self-management program on the quality of life of women with endometriosis. Arch Gynecol Obstet 2024;310(2):1157-1170. [doi: 10.1007/s00404-024-07468-4] [Medline: 38869629]

XSL•FO
RenderX

55. Rohloff N, Götz T, Kortekamp SS, Heinze NR, Weber C, Schäfer SD. Influence of app-based self-management on the quality of life of women with endometriosis. Cureus 2024;16(8):e67655. [doi: 10.7759/cureus.67655] [Medline: 39314601]
56. Zugaj MR, Germeyer A, Kranz K, Züger A, Keßler J. Experiences of patients with endometriosis with a digital health application: a qualitative analysis. Arch Gynecol Obstet 2024;310(4):2253-2263. [doi: 10.1007/s00404-024-07651-7] [Medline: 39066952]
57. Sherman KA, Pehlivan MJ, Singleton A, Hawkey A, Redfern J, Armour M, et al. Co-design and development of EndoSMS, a supportive text message intervention for individuals living with endometriosis: mixed methods study. JMIR Form Res 2022;6(12):e40837 [FREE Full text] [doi: 10.2196/40837] [Medline: 36485029]
58. Sherman KA, Pehlivan MJ, Pereira C, Hawkey A, Singleton AC, Redfern J, et al. Randomised controlled pilot trial of the EndoSMS supportive text message intervention for individuals with endometriosis: feasibility and acceptability results. J Psychosom Res 2024;187:111929 [FREE Full text] [doi: 10.1016/j.jpsychores.2024.111929] [Medline: 39305835]
59. Holsteg S, Askeridis JM, Krajewski J, Mildner P, Freitag S, Müller T, et al. Virtual reality roleplays for patients with depression: a user experience evaluation. Internet Interv 2024;35:100713 [FREE Full text] [doi: 10.1016/j.invent.2024.100713] [Medline: 38318087]
60. Pérez-Aronsson A, Thell M, Lampa E, Löfving SG, Tökés A, Torakai N, et al. Adaptation of the trauma group intervention 'Teaching Recovery Techniques' for online delivery: a participatory design and usability study. Internet Interv 2022;30:100589 [FREE Full text] [doi: 10.1016/j.invent.2022.100589] [Medline: 36411786]
61. Stratégie nationale de lutte contre l'endométriose. Ministère du Travail, de la Santé, des Solidarités et des Familles. 2022. URL: https://sante.gouv.fr/archives/archives-presse/archives-dossiers-de-presse/article/strategie-nationale-de-lutte-contre-l-endometriose [accessed 2025-07-03]
62. Le gouvernement du Canada renforce l'accès des personnes vivant avec l'endométriose aux services de santé sexuelle et reproductive. Gouvernement du Canada. 2023. URL: https://www.canada.ca/fr/sante-canada/nouvelles/2023/09/le-gouvernement-du-canada-renforce-lacces-des-personnes-vivant-avec-lendometriose-aux-services-de-sante-sexuelle-et-reproductive.html [accessed 2025-07-03]
63. National action plan for endometriosis. Department of Health, Disability and Ageing, Australian Government. 2023. URL: https://www.health.gov.au/resources/publications/national-action-plan-for-endometriosis?language=en [accessed 2025-06-24]
64. Shiplo SN, Gholiof M, Sarin N, Leonardi M. Endometriosis influencers on Instagram: who are they and what are they posting? J Minim Invasive Gynecol 2025;32(8):693-700 [FREE Full text] [doi: 10.1016/j.jmig.2025.01.018] [Medline: 39922273]
65. Wettstein M, Imesch P, Witzel I, Metzler JM. Evolution of endometriosis research: a comprehensive analysis of research designs and topics over the last 15 years. J Endometr Uterine Disord 2024;7:100076. [doi: 10.1016/j.jeud.2024.100076]
66. Agarwal SK, Chapron C, Giudice LC, Laufer MR, Leyland N, Missmer SA, et al. Clinical diagnosis of endometriosis: a call to action. Am J Obstet Gynecol 2019;220(4):354.e1-354.e12 [FREE Full text] [doi: 10.1016/j.ajog.2018.12.039] [Medline: 30625295]
67. Missmer SA, Tu FF, Agarwal SK, Chapron C, Soliman AM, Chiuve S, et al. Impact of endometriosis on life-course potential: a narrative review. Int J Gen Med 2021;14:9-25 [FREE Full text] [doi: 10.2147/IJGM.S261139] [Medline: 33442286]
68. Karekla M, Kasinopoulos O, Neto DD, Ebert DD, Van Daele T, Nordgreen T, et al. Best practices and recommendations for digital interventions to improve engagement and adherence in chronic illness sufferers. Eur Psychol 2019;24(1):49-67. [doi: 10.1027/1016-9040/a000349]
69. Li Y, Bai C, Reddy CK. A distributed ensemble approach for mining healthcare data under privacy constraints. Inf Sci (NY) 2016;330:245-259 [FREE Full text] [doi: 10.1016/j.ins.2015.10.011] [Medline: 26681811]
70. Beerthuizen T, Rijssenbeek-Nouwens LH, van Koppen SM, Khusial RJ, Snoeck-Stroband JB, Sont JK. Internet-based self-management support after high-altitude climate treatment for severe Asthma: randomized controlled trial. J Med Internet Res 2020;22(7):e13145 [FREE Full text] [doi: 10.2196/13145] [Medline: 32706692]
71. Wang Y, Li B, Zhou Y, Wang Y, Han X, Zhang S, et al. Does endometriosis disturb mental health and quality of life? A systematic review and meta-analysis. Gynecol Obstet Invest 2021;86(4):315-335. [doi: 10.1159/000516517] [Medline: 34352799]
72. Taylor. Connected health-how digital technology is transforming health and social care. Mondaq. 2015. URL: https://www.mondaq.com/uk/healthcare/393714/connected-health---how-digital-technology-is-transforming-health-and-social-care [accessed 2024-03-05]
73. Nguyen AM, Rivera AM, Gualtieri L. A new health care paradigm: the power of digital health and e-patients. Mayo Clin Proc Digit Health 2023;1(3):203-209 [FREE Full text] [doi: 10.1016/j.mcpdig.2023.04.005] [Medline: 40206609]
74. Meskó B, deBronkart D. Patient design: the importance of including patients in designing health care. J Med Internet Res 2022;24(8):e39178 [FREE Full text] [doi: 10.2196/39178] [Medline: 36044250]
75. Culley L, Law C, Hudson N, Denny E, Mitchell H, Baumgarten M, et al. The social and psychological impact of endometriosis on women's lives: a critical narrative review. Hum Reprod Update 2013;19(6):625-639. [doi: 10.1093/humupd/dmt027] [Medline: 23884896]
76. Chen K, Huang JJ, Torous J. Hybrid care in mental health: a framework for understanding care, research, and future opportunities. NPP—Digit Psychiatry Neurosci 2024;2(1):16. [doi: 10.1038/s44277-024-00016-7]

XSL•FO
RenderX

77. Morotti M, Vincent K, Becker CM. Mechanisms of pain in endometriosis. Eur J Obstet Gynecol Reprod Biol 2017;209:8-13. [doi: 10.1016/j.ejogrb.2016.07.497] [Medline: 27522645]
78. Gete DG, Doust J, Mortlock S, Montgomery G, Mishra GD. Impact of endometriosis on women's health-related quality of life: a national prospective cohort study. Maturitas 2023;174:1-7 [FREE Full text] [doi: 10.1016/j.maturitas.2023.04.272] [Medline: 37182389]
79. Mellado BH, Falcone ACM, Poli-Neto OB, Rosa E Silva JC, Nogueira AA, Candido-Dos-Reis FJ. Social isolation in women with endometriosis and chronic pelvic pain. Int J Gynaecol Obstet 2016;133(2):199-201. [doi: 10.1016/j.ijgo.2015.08.024] [Medline: 26873129]
80. Sivajohan B, Lin T, Bedaiwy MA. Cost estimates associated with diagnosis treatment of endometriosis. In: Oral E, editor. Endometriosis and Adenomyosis: Global Perspectives Across the Lifespan. Cham: Springer International Publishing; 2022:361-395.
81. Olliges E, Bobinger A, Weber A, Hoffmann V, Schmitz T, Popovici R, et al. The physical, psychological, and social day-to-day experience of women living with endometriosis compared to healthy age-matched controls-a mixed-methods study. Front Glob Womens Health 2021;2:767114 [FREE Full text] [doi: 10.3389/fgwh.2021.767114] [Medline: 34977863]
82. Pickett C, Foster WG, Agarwal SK. Current endometriosis care and opportunities for improvement. Reprod Fertil 2023;4(3):e220091 [FREE Full text] [doi: 10.1530/RAF-22-0091] [Medline: 37402150]
83. Green EC, Murphy EM, Gryboski K. The health belief model. In: Sweeny K, Robbins ML, Cohen LM, editors. The Wiley Encyclopedia of Health Psychology. Hoboken, New Jersey: John Wiley & Sons, Ltd; 2020:211-214.
84. Nezami BT, Lang W, Jakicic JM, Davis KK, Polzien K, Rickman AD, et al. The effect of self-efficacy on behavior and weight in a behavioral weight-loss intervention. Health Psychol 2016;35(7):714-722 [FREE Full text] [doi: 10.1037/hea0000378] [Medline: 27183306]
85. Champion VL, Skinner CS. The health belief model. In: Glanz K, Rimer BK, Viswanath K, editors. Health Behavior and Health Education: Theory, Research, and Practice. San Francisco, CA: Jossey-Bass; 2008:45-65.
86. McKnight K, Omotosho O, Jassim S, Cotter A. Exercise and endometriosis-is there a promising future? A narrative review. Ir J Med Sci 2024;193(5):2375-2387. [doi: 10.1007/s11845-024-03733-2] [Medline: 38916808]
87. Barnard ND, Holtz DN, Schmidt N, Kolipaka S, Hata E, Sutton M, et al. Nutrition in the prevention and treatment of endometriosis: a review. Front Nutr 2023;10:1089891 [FREE Full text] [doi: 10.3389/fnut.2023.1089891] [Medline: 36875844]
88. Deci EL, Ryan RM. Intrinsic Motivation and Self-Determination in Human Behavior. New York: Kluwer Academic/Plenum Publishers; 1985.
89. Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: toward an integrative model of change. J Consult Clin Psychol 1983;51(3):390-395. [doi: 10.1037/0022-006X.51.3.390]
90. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013;46(1):81-95 [FREE Full text] [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
91. Deci EL, Ryan RM. The "what" and "why" of goal pursuits: human needs and the self-determination of behavior. Psychol Inq 2000;11(4):227-268. [doi: 10.1207/s15327965pli1104_01]
92. Klooster IT, Kip H, van Gemert-Pijnen L, Crutzen R, Kelders S. A systematic review on eHealth technology personalization approaches. iScience 2024;27(9):110771 [FREE Full text] [doi: 10.1016/j.isci.2024.110771] [Medline: 39290843]
93. Zangger G, Bricca A, Liaghat B, Juhl CB, Mortensen SR, Andersen RM, et al. Benefits and harms of digital health interventions promoting physical activity in people with chronic conditions: systematic review and meta-analysis. J Med Internet Res 2023;25:e46439 [FREE Full text] [doi: 10.2196/46439] [Medline: 37410534]
94. Petraglia F, Vannuccini S, Santulli P, Marcellin L, Chapron C. An update for endometriosis management: a position statement. J Endometriosis Uterine Disord 2024;6:100062. [doi: 10.1016/j.jeud.2024.100062]
95. Balaskas A, Schueller SM, Doherty K, Cox AL, Doherty G. Designing personalized mental health interventions for anxiety: CBT therapists' perspective. Int J Hum Comput Stud 2024;190:103319. [doi: 10.1016/j.ijhcs.2024.103319]
96. Ajzen I. The theory of planned behavior. Organ Behav Hum Decis Process 1991;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
97. Terrade F, Pasquier H, Reerinck-Boulanger J, Guingouain G, Somat A. Social acceptability: how social Determinent can influence analysis of technology system acceptability. Le Travail Hum 2009:383-395 [FREE Full text] [doi: 10.3917/th.724.0383]
98. Lepore D, Dolui K, Tomashchuk O, Shim H, Puri C, Li Y, et al. Interdisciplinary research unlocking innovative solutions in healthcare. Technovation 2023;120:102511. [doi: 10.1016/j.technovation.2022.102511]
99. Flessa S, Huebner C. Innovations in health care-a conceptual framework. Int J Environ Res Public Health 2021;18(19):10026 [FREE Full text] [doi: 10.3390/ijerph181910026] [Medline: 34639328]
100. Johnson BT, Hennessy EA. Systematic reviews and meta-analyses in the health sciences: best practice methods for research syntheses. Soc Sci Med 2019;233:237-251 [FREE Full text] [doi: 10.1016/j.socscimed.2019.05.035] [Medline: 31233957]
101. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016;15(2):155-163 [FREE Full text] [doi: 10.1016/j.jcm.2016.02.012] [Medline: 27330520]

XSL•FO
RenderX

102. Bartolucci AA, Hillegass WB. Overview, strengths, and limitations of systematic reviews and meta-analyses. In: Chiappelli F, editor. Evidence-Based Practice: Toward Optimizing Clinical Outcomes Berlin. Heidelberg: Springer Berlin Heidelberg; 2010:17-33.
103. Doleman B, Mathiesen O, Jakobsen JC, Sutton AJ, Freeman S, Lund JN, et al. Methodologies for systematic reviews with meta-analysis of randomised clinical trials in pain, anaesthesia, and perioperative medicine. Br J Anaesth 2021;126(4):903-911 [FREE Full text] [doi: 10.1016/j.bja.2021.01.004] [Medline: 33558052]
104. Smith KA, Cipriani A, Geddes JR. The usefulness and interpretation of systematic reviews. BJPsych Adv 2018;22(2):132-141. [doi: 10.1192/apt.bp.114.013128]
105. Nittas V, Daniore P, Chavez SJ, Wray TB. Challenges in implementing cultural adaptations of digital health interventions. Commun Med (Lond) 2024;4(1):7 [FREE Full text] [doi: 10.1038/s43856-023-00426-2] [Medline: 38182750]
106. Peng D, Yu Z. A literature review of digital literacy over two decades. Educ Res Int 2022;2022:1-8. [doi: 10.1155/2022/2533413]
107. Giebel GD, Speckemeier C, Abels C, Plescher F, Börchers K, Wasem J, et al. Problems and barriers related to the use of digital health applications: Scoping review. J Med Internet Res 2023;25:e43808 [FREE Full text] [doi: 10.2196/43808] [Medline: 37171838]
108. Tenny S, Brannan JM, Brannan GD. Qualitative study. In: StatPearls [Internet]. Treasure Island, FL: StatPearls Publishing; 2022.
109. Naderbagi A, Loblay V, Zahed IUM, Ekambareshwar M, Poulsen A, Song YJC, et al. Cultural and contextual adaptation of digital health interventions: narrative review. J Med Internet Res 2024;26:e55130 [FREE Full text] [doi: 10.2196/55130] [Medline: 38980719]
110. Akeel AU, Mundy D. Re-thinking technology and its growing role in enabling patient empowerment. Health Informatics J 2019;25(4):1278-1289 [FREE Full text] [doi: 10.1177/1460458217751013] [Medline: 29366367]

Abbreviations

DEDHI: Design and Evaluation of Digital Health Interventions
DHI: digital health intervention
PRISMA: Preferred Reporting Items for Systematic reviews and Meta-Analyses
MMAT: Mixed Methods Assessment Tool

Edited by C Jacob; submitted 28.01.25; peer-reviewed by S Hajebrahimi, J Ohan; comments to author 31.05.25; revised version received 04.07.25; accepted 07.08.25; published 14.10.25.

Please cite as:
Pavic T, Nadarajah K, Somat A, Cabagno G, Terrade F
Endometriosis Support and Development of Digital Technology–Based Interventions: Systematic Review
JMIR Hum Factors 2025;12:e71859
URL: https://humanfactors.jmir.org/2025/1/e71859
doi:10.2196/71859
PMID:

©Tivizio Pavic, Kévin Nadarajah, Alain Somat, Geneviève Cabagno, Florence Terrade. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 14.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Redesigning Telemedicine for Traditional Chinese Medicine: Service Design Approach to Digital Transformation

Arisara Jiamsanguanwong[1,2], DEng; Romanee Luo[1], BEng; Ratchapoom Kaingam[1], BEng; Oran Kittithreerapronchai[1,2], PhD; Waratta Authayarat[3], PhD

[1]Department of Industrial Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok, Thailand

[2]Human-Robot Collaboration and Systems Integration Research Unit, Chulalongkorn University, Bangkok, Thailand

[3]Department of Industrial Engineering, Faculty of Engineering, Burapha University, 169 Long-Had Bangsaen Road, Saensuk, Mueang, Chon Buri, Thailand

Corresponding Author:
Waratta Authayarat, PhD
Department of Industrial Engineering, Faculty of Engineering, Burapha University, 169 Long-Had Bangsaen Road, Saensuk, Mueang, Chon Buri, Thailand

Abstract

Background: With the rising global adoption of telemedicine, there is a crucial need to address inefficiencies and challenges in current service systems. This case study focused on enhancing the telemedicine service system of a traditional Chinese medicine clinic.

Objective: The primary objective was to identify and address pain points and inefficiencies in the existing telemedicine system with the aim of streamlining service operations for the benefit of both patients and service providers.

Methods: Through comprehensive service design analysis, including the creation of a customer journey map and a service blueprint, key areas for improvement were identified, and the service process was redesigned accordingly. A user-friendly web application was developed and evaluated using usability testing and satisfaction assessments. Participants took part voluntarily. Task testing was conducted using real-world scenarios, with index of item-objective congruence values ranging from 0.84 to 1.00. Participants were assigned role-specific tasks as either patients or service providers in a step-by-step format, followed by role-specific paper-based questionnaires.

Results: The redesigned system successfully streamlined operations by automating processes and reducing task complexity, resulting in improved time efficiency for both user groups. Participants included 6 patients (aged 23 - 54 years) and 7 service providers from various departments. Usability testing revealed a task success rate of 100% for all tasks among patients, coordinators, physicians, and finance personnel as well as 83.33% among pharmacists. Satisfaction outcomes were positive: patients reported a net promoter score of 67, whereas service providers reported a mean System Usability Scale score of 71.4 (SD 20.76).

Conclusions: This study highlights the transformative potential of telemedicine in health care delivery. For patients, consolidating services into a single digital platform improved accessibility and ease of use. For service providers, the system reduced repetitive tasks and facilitated more efficient task completion. These findings demonstrate the effectiveness of service design methodologies in enhancing telemedicine systems, ultimately contributing to improved health care quality and patient outcomes.

(JMIR Hum Factors 2025;12:e76752) doi:10.2196/76752

KEYWORDS

digital transformation; efficiency improvement; service design; telemedicine; traditional Chinese medicine

Introduction

Telemedicine, also known as telehealth or eHealth, was introduced in the 1970s as a means of delivering health care services over long distances using telecommunication and IT tools [1,2]. These tools include voice and video calls, SMS text messaging, remote monitoring, and virtual consultations [3]. The US Centers for Disease Control and Prevention has defined telemedicine as “healing at a distance” [4]. Today, telemedicine is used for a wide range of services, including diagnosis of common illnesses, follow-up consultations, psychotherapy, and chronic disease management [5].

The benefits of telemedicine are widely recognized. These include cost savings [6], improved access for patients with mobility limitations [7], and enhanced continuity of care while minimizing exposure during the COVID-19 pandemic [8,9]. The global health crisis in 2020 accelerated the adoption of telemedicine as health care systems sought to reduce viral transmission and maintain service continuity [4,10]. In the United States, telehealth visits increased by 50% in the first

XSL•FO
RenderX

quarter of 2020 compared to the same period in 2019 and surged by 154% by week 13 of 2020 [11]. Despite this rapid uptake, much of the telemedicine infrastructure at the time was still in its early stages of implementation, hindered by limited understanding of the virus and system unpreparedness [12].

While the urgency of the pandemic necessitated swift development and deployment of telemedicine platforms, numerous challenges persist. These include concerns about patient-provider relationships, digital literacy, technical reliability, privacy and data confidentiality, reimbursement policies, physical examination limitations, and acceptability among special populations [3,13,14].

This study focused on the development and improvement of a telemedicine system within a traditional Chinese medicine (TCM) clinic in Bangkok, Thailand. The clinic serves approximately 5000 patients monthly and presents unique service characteristics distinct from conventional modern medicine. TCM diagnostics rely on observation of the patient's physical appearance, vocal quality, pulse palpation, and tongue examination, whereas treatments often involve custom herbal preparations. These distinctive diagnostic and therapeutic practices introduce additional complexity in adapting telemedicine systems for TCM settings.

The clinic initially launched telemedicine services via telephone in April 2020, expanding to video consultations in March 2022. According to internal records, the number of telemedicine visits more than doubled from 1955 in 2020 to 3972 in 2021, with the age group from 41 to 50 years representing the largest proportion of users. Despite the growth, both patients and service providers reported difficulties. Patients experienced confusion regarding appointment scheduling and status updates, whereas staff encountered challenges related to fragmented communication channels, repetitive manual tasks, and workflow inefficiencies.

Given the known advantages of telemedicine [7] and the increasing demand for remote TCM services, this study aimed to systematically identify pain points in the existing telemedicine service system and redesign it to improve efficiency and user experience for both patients and service providers.

Methods

Investigation of Current Service Operation

Overview

The telemedicine service was investigated from both patient and service provider perspectives. Two hypothetical workflows were developed to visualize the end-to-end service interaction: a *customer journey map* representing the patient experience [15,16] and a *service blueprint* illustrating the provider-side operations [16,17]. These workflows were constructed using data collected from direct observations of the service operations, content review of the clinic's website, and semistructured interviews with service providers.

The purpose of these visual tools was to map the sequence of interactions between patients and the service system, as well as among service providers, to identify pain points and service gaps [18]. The preliminary workflows were subsequently validated through in-depth interviews: 4 current telemedicine patients validated the customer journey map, and representatives from each relevant department (eg, registration, consultation, and pharmacy) reviewed and confirmed the accuracy of the service blueprint.

Barriers From the Patient Perspective

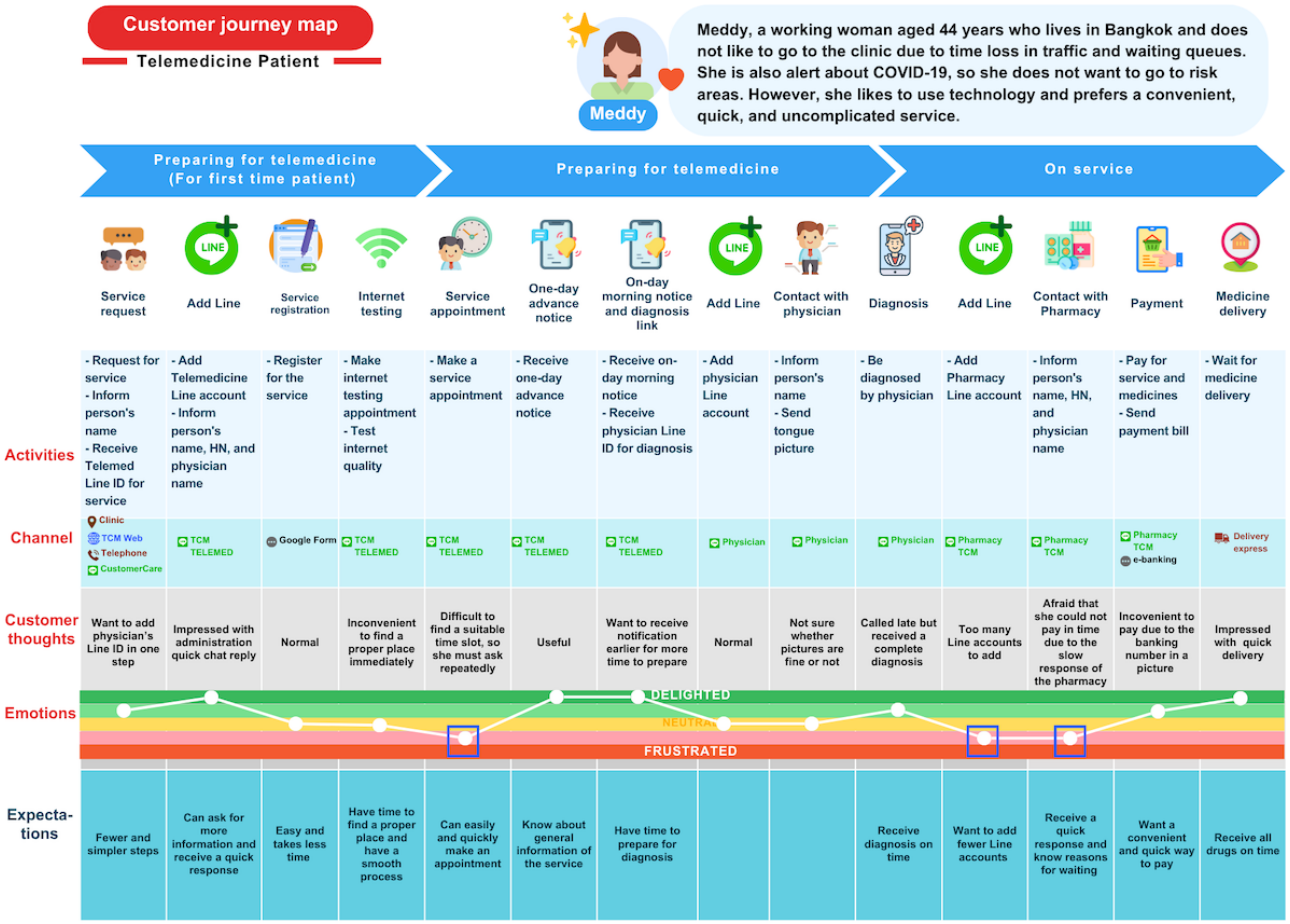
Figure 1 presents the validated customer journey map from the patients' point of view. The analysis revealed 3 key barriers that contributed to user dissatisfaction, indicated by negative emotional states on the emotion-tracking row of the map.

First, patients experienced difficulty in scheduling appointments. As shown in the "Customer thoughts" row in Figure 1, many users had to engage in multiple rounds of communication with the clinic to secure a suitable time slot, some reporting up to 5 separate calls. This resulted in excessive human effort and a high interaction cost. Patients expressed a clear expectation for a simplified and more efficient appointment scheduling process.

Second, the current system required patients to interact with multiple communication channels. Specifically, they had to register with 4 separate Line messaging accounts (CustomerCare TCM, TCM TELEMED, the physician's personal account, and Pharmacy TCM) to complete a single telemedicine session. This fragmentation added significant complexity and increased the interaction burden on patients. Many expressed the desire to minimize the number of communication accounts required for service access.

XSL•FO
RenderX

Figure 1. Customer journey map of the current telemedicine service (patient perspective). HN: health number; TCM: traditional Chinese medicine.



Third, there was a lack of transparency regarding service response times. Patients were not informed of estimated wait times for replies from service providers. Data logs revealed a wide variation in response time, ranging from 5 minutes to as long as 3 hours, with an average of 30 minutes. This uncertainty undermined patient confidence and satisfaction. Participants emphasized the need for real-time service tracking and automated notifications to improve the overall user experience.

Collectively, these issues of excessive communication effort, fragmented service channels, and uncertain wait times indicate critical inefficiencies in the current telemedicine system. From the patient perspective, three areas for improvement were identified: (1) simplification of the service process, (2) consolidation of communication channels, and (3) visibility and predictability of service status and wait times. Addressing these pain points through streamlined service design is expected to enhance patient satisfaction, reduce interaction costs, and improve the overall usability and efficiency of the telemedicine service.

Barriers From the Service Provider Perspective

Figure 2 presents the current service blueprint for the TCM telemedicine operation, highlighting 3 primary barriers reported by service providers.

First, providers faced challenges due to uncertain waiting times. They were required to monitor patient responses closely and ensure timely replies to maintain patient satisfaction (problem A in Figure 2). Patient response times varied widely, ranging from as little as 10 minutes to as long as 1 month, with an average of 2 days. This unpredictability increased provider workload and disrupted workflow. Digitizing the system to automate follow-up communication would mitigate this issue.

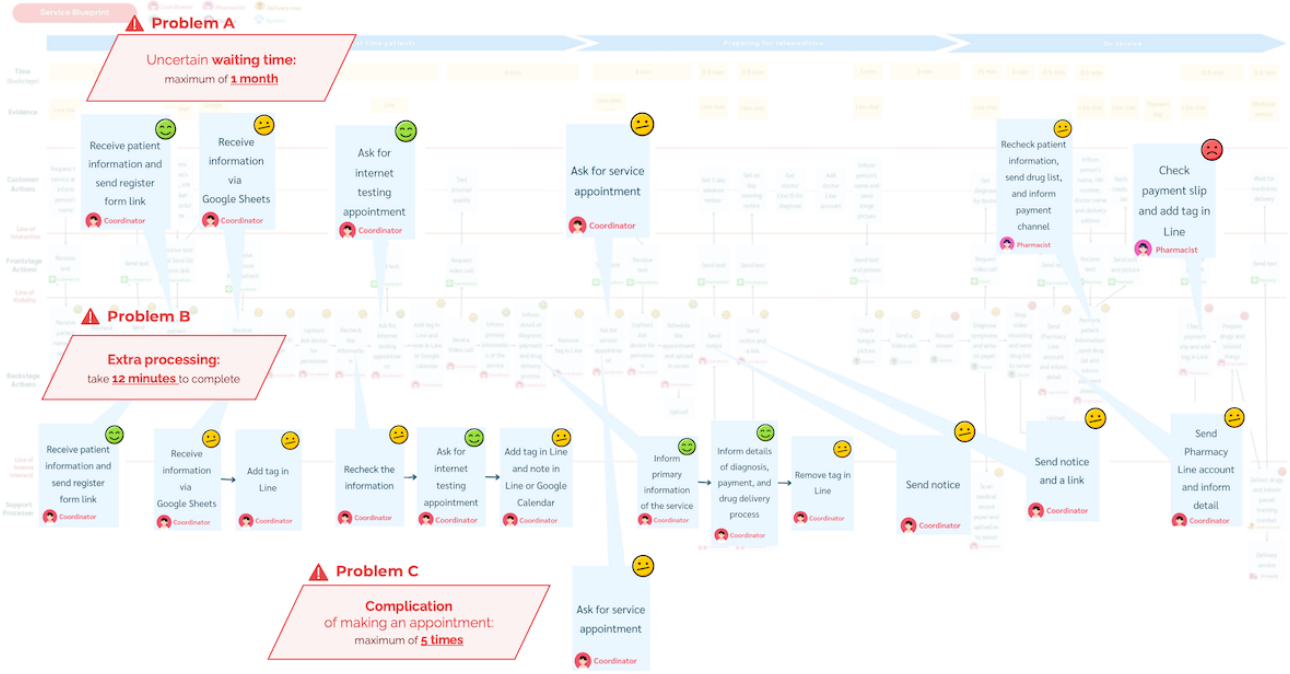
Second, service providers were burdened by numerous repetitive administrative tasks (problem B in Figure 2). These included sending registration forms, checking for completeness, updating status tags, issuing notifications, and sharing Line contact details. Collectively, these tasks consumed approximately 12 minutes per patient, representing wasted time and avoidable effort. Automating these processes could substantially reduce inefficiency.

Third, appointment scheduling was unnecessarily complicated (problem C in Figure 2). Providers often engaged in repeated communication with patients (up to 5 exchanges to finalize a booking due to mismatched availability). This process increased provider effort and interaction costs, diverting attention from more value-adding clinical activities.

Overall, these barriers illustrate inefficiencies that compromise the effectiveness and stability of the telemedicine service. From the provider perspective, three key areas required improvement: (1) reducing waiting times, (2) eliminating repetitive administrative tasks, and (3) simplifying the appointment scheduling process.

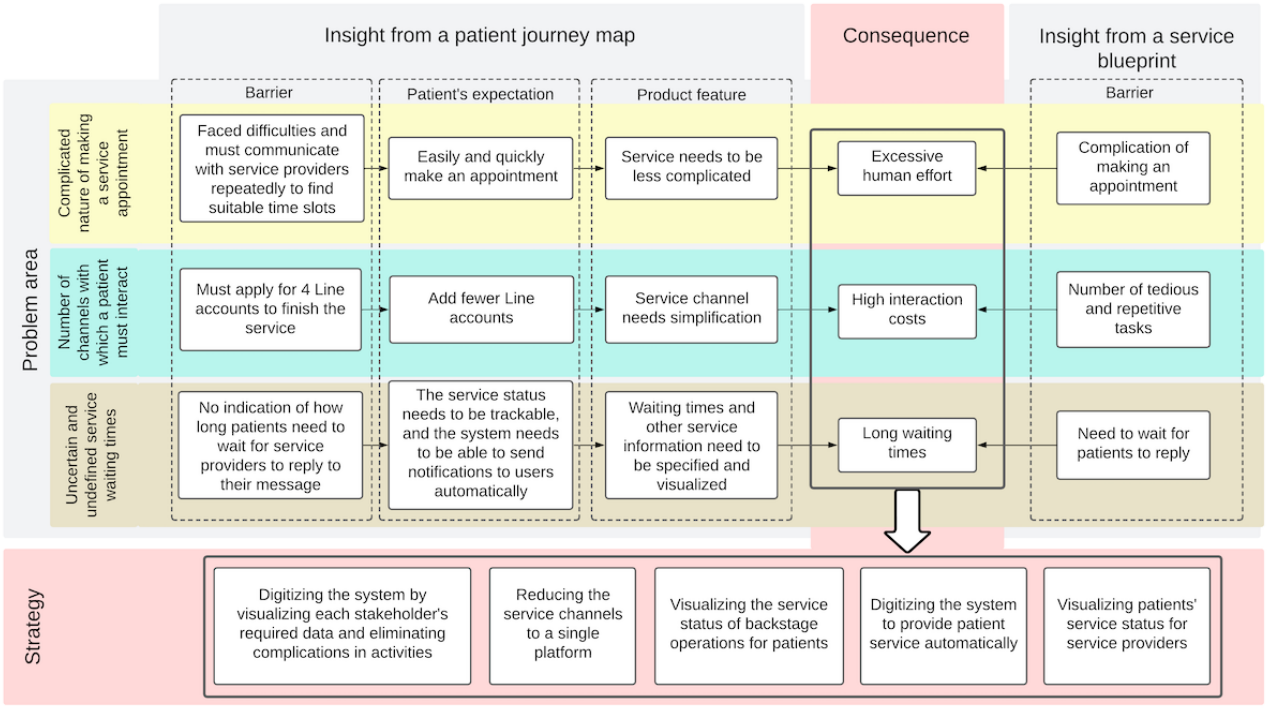
XSL•FO
RenderX

Figure 2. Service blueprint of the current traditional Chinese medicine telemedicine service (provider perspectives).



When considered alongside patient-reported barriers, a common set of challenges emerged for both groups: (1) excessive human effort, (2) high interaction costs, and (3) long or uncertain waiting times. As shown in Figure 3, addressing these problem areas is essential to improving operational efficiency and ensuring a better overall telemedicine experience for both patients and service providers.

Figure 3. Common problems and improvement strategies for patients and providers.



System Design

Redesigning Service Operations

Following the identification of barriers in the current service system, solutions were designed to eliminate inefficiencies and enhance user experience. A digital transformation strategy was adopted integrating information, computing, communication, and connectivity technologies [19,20]. This approach aimed to strengthen the telemedicine service's capabilities and address issues such as repetitive tasks and the difficulty of sharing real-time information between patients and providers. For example, dissatisfaction arising from uncertain waiting times could be mitigated by providing estimated wait times and visualizing them for users [21].

XSL•FO
RenderX

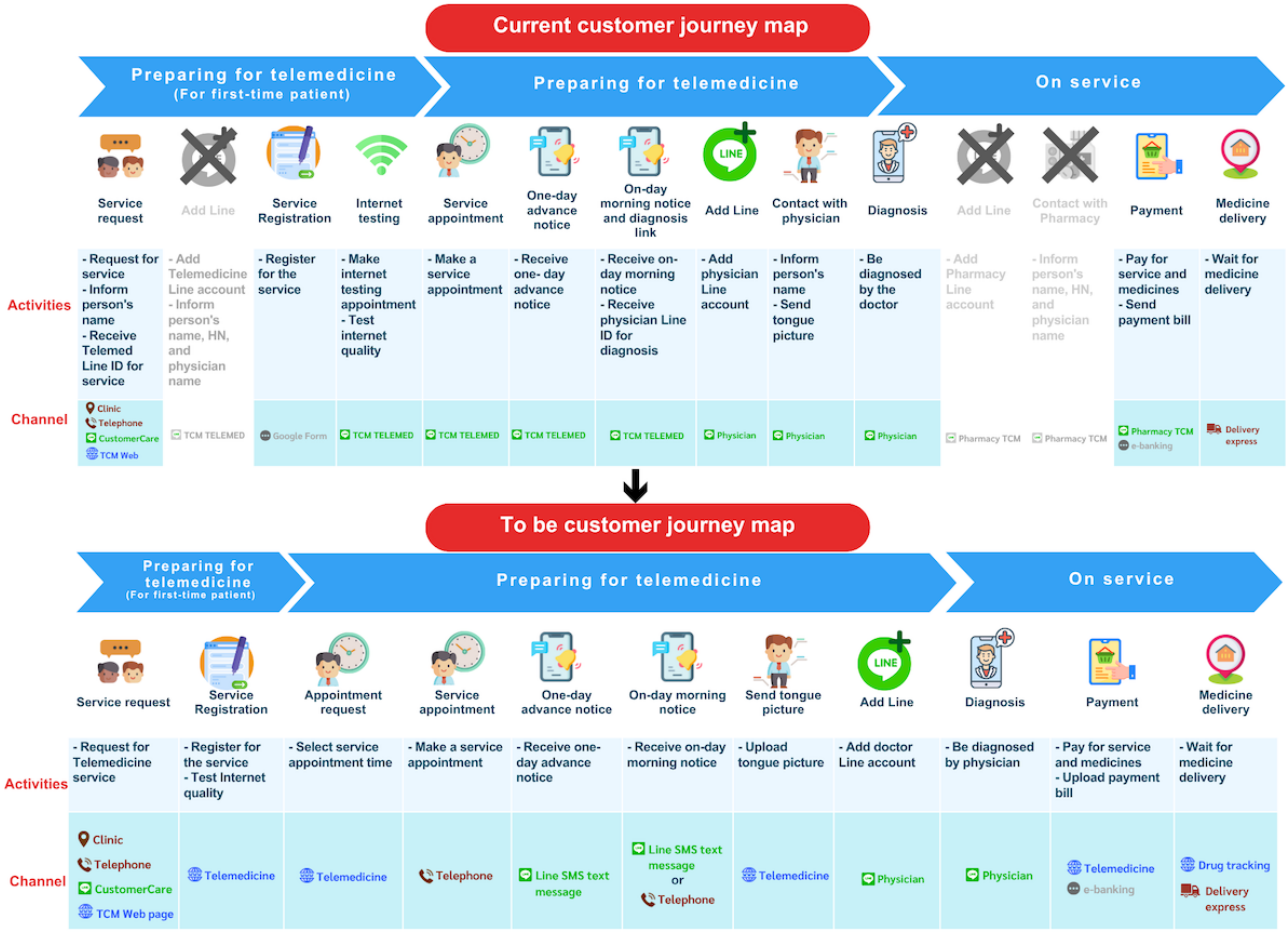
The success of such digital interventions also depended on staff skills and competencies. Therefore, comprehensive training for all system users was considered essential to ensure effective adoption of the redesigned processes.

To guide implementation, a "to be" customer journey map was developed (Figure 4) illustrating three major strategies: (1) *channel consolidation* (consolidating multiple service channels into a single integrated platform that supported all service activities, thereby lowering patient interaction costs), (2) *system digitization* (streamlining complex activities such as appointment scheduling by digitizing workflows and visualizing the required data for each stakeholder), and (3) *service status visualization* (providing patients with real-time updates on backstage operations to reduce uncertainty and dissatisfaction caused by undefined waiting times [22]).

In summary, the redesigned service model focused on consolidating communication channels, digitizing processes, and visualizing service status—strategies aimed at streamlining operations, improving efficiency, and delivering a more seamless telemedicine experience.

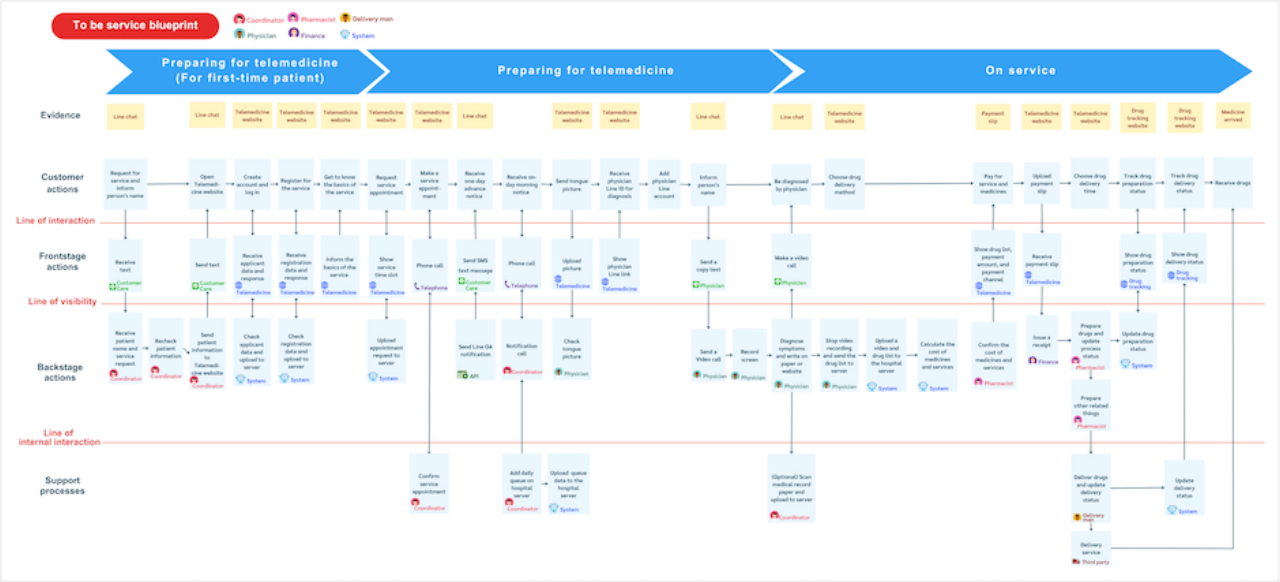
Moreover, a "to be" service blueprint was developed to enhance the service provider experience (Figure 5, or see the enlarged image in Multimedia Appendix 1). This blueprint emphasized 2 key strategies. First, the system was digitalized to automate patient-facing processes, thereby reducing provider workload associated with repetitive administrative tasks and ongoing patient follow-up. Automation also eliminated complex tasks such as appointment scheduling. Second, the blueprint introduced service status visualization for providers accompanied by automated notification alerts when action was required. This functionality helped reduce uncertainty during patient response delays and minimized the need for providers to manually review patient messages. In summary, streamlining provider operations was achieved through 2 strategies: system digitalization and visualization of service status.

Figure 4. "To be" customer journey map of the redesigned telemedicine service (patient perspective). TCM: traditional Chinese medicine.



XSL•FO
RenderX

Figure 5. "To be" service blueprint of the redesigned telemedicine service (provider perspective). API: application programming interface; OA: official account.



System Interface

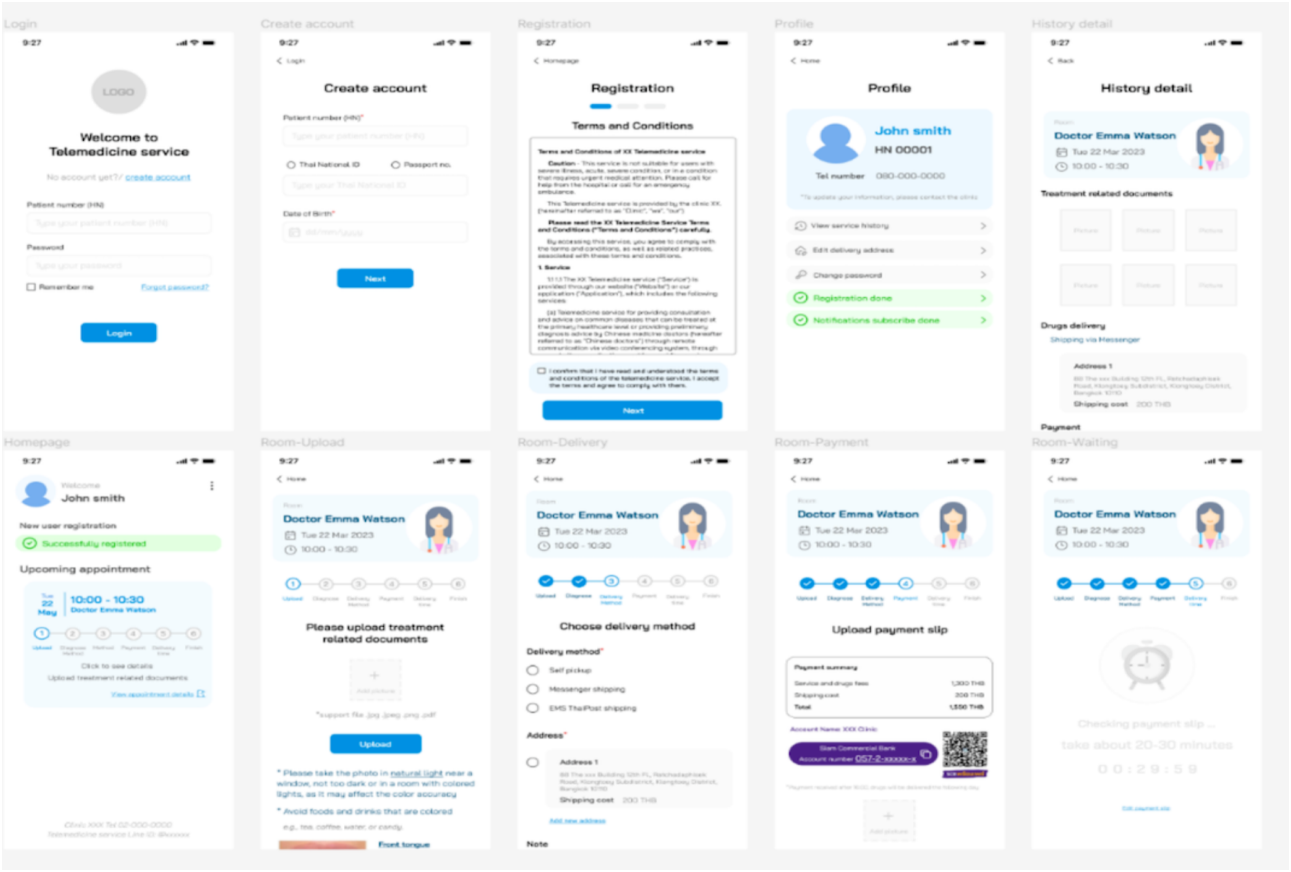
The system interface is a critical component of service improvement, functioning as the primary channel of interaction between patients and health care providers. To guide interface design, a use case diagram of the redesigned telemedicine web application was developed to define user roles and their corresponding system functions. For example, users with system authentication privileges could manage log-in and log-out processes, change account passwords, create new patient accounts, and edit other user accounts.

Interfaces for each role were prototyped using Figma, a digital design platform that enables realistic mock-ups and supports iterative development [23]. Interactive buttons and workflows were tested during the prototyping stage, and feedback from trial users was collected to refine the design through iterative improvement cycles.

The final interface was divided into 2 dedicated websites: one for patients and one for service providers. The patient website was optimized for mobile devices and designed to support all essential patient tasks, including log-in, registration, profile management, and appointment booking (Figure 6, or see the enlarged image in Multimedia Appendix 2). In contrast, the provider website was optimized for desktop use to reflect providers' routine work patterns. This platform supported role-specific functions for physicians, coordinators, pharmacists, and finance staff, facilitating streamlined management of patient services and system operations.

XSL•FO
RenderX

Figure 6. The patient web application interface optimized for mobile devices.



System Implementation

Following the design phase, the next step was the implementation of the new telemedicine web application system. This stage involved translating the design concepts into a functional system accessible to both patients and service providers. An overview of the telemedicine service system is shown in Figure 7.

To support implementation, an entity-relationship diagram was developed to represent the database structure and the relationships between its entities [24]. The entity-relationship diagram served as a blueprint for database design, ensuring that data were logically organized and efficiently structured to support the system's core functionalities.

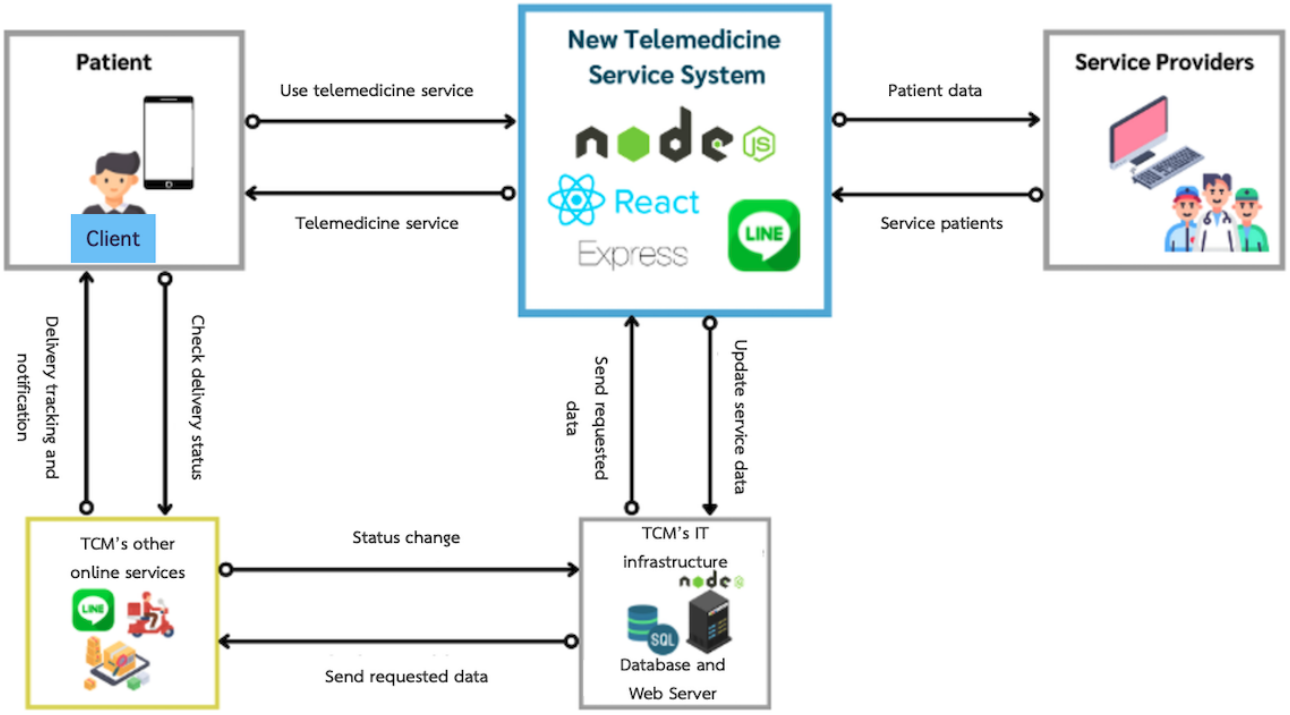
In developing the TCM clinic's telemedicine service system, the web applications described in the System Interface section were implemented using Node.js (version 10; OpenJS Foundation), a widely adopted runtime environment for web development. The system interfaces were designed to be dynamic and responsive, ensuring low latency and a positive user experience.

The implementation integrated several JavaScript libraries and packages. The front end was built using React (version 17), compatible with Node.js version 10, alongside additional libraries such as SweetAlert and Fortawesome to enhance interactivity and usability. For the back end, Express.js (OpenJS Foundation) served as the primary framework for building application programming interfaces that adhered to the Representational State Transfer architectural style, enabling communication between users and the clinic's Microsoft SQL Server 2000 database. Supporting packages included *mssql* for database connectivity, *cors* to enable cross-origin requests, *express-fileupload* for image uploading, *bcrypt* for secure password encryption, and *line/bot-sdk* for integrating notification services through the Line app.

System interactions followed a client-server architecture: users engaged with the front-end interface, which transmitted data requests to the back-end system. The back end processed the data, executed the required operations, and returned results to the front end for user display.

XSL•FO
RenderX

Figure 7. Overview of the redesigned telemedicine service system. TCM: traditional Chinese medicine.



Although the real-time communication module was developed independently from the back end, seamless integration allowed for direct connection between the web application and physicians' Line accounts. Furthermore, the new telemedicine service system was linked with the clinic's existing automated alert system, prescription tracking system [25], and medicine delivery service [26], ensuring compatibility with established workflows and continuity of care.

System Validation

A system validation process was conducted to ensure that the redesigned telemedicine platform met the requirements of both patients and service providers and to evaluate user satisfaction with its performance and functionality.

Participants

Two participant groups were recruited: patients and service providers. Recruitment flyers were posted on noticeboards throughout the hospital where the TCM clinic is located. The flyers included details of the study objectives, experimental procedures, and eligibility criteria for both groups. Interested individuals self-screened for eligibility and contacted the research team via telephone to register.

A purposive sampling strategy was used to allow for in-depth understanding of user experiences [27]. Initial eligibility interviews were conducted by 2 web application developers from the author team. Participants were intentionally selected based on criteria relevant to the research objectives. To avoid bias, the remaining author team members who conducted the data analysis were blinded to the recruitment process.

Tasks

Overview

Task testing was conducted using simulated real-world scenarios designed to cover all key functions that patients and service providers would perform within the system. Separate test scenarios were developed for each role to ensure that all system features were evaluated comprehensively.

Content validity of the test scenarios was assessed by 3 domain experts, with each task receiving an index of item-objective congruence score between 0.84 and 1.00, indicating strong validity and suitability for use [28].

Tasks for the Patient Role

The patient scenarios were designed to address challenges commonly faced by individuals visiting the TCM clinic in person, particularly those who lived at a distance and found travel burdensome. Six scenarios were developed to validate the telemedicine functions from the patient perspective:

1. Registration: creating a telemedicine account and completing a 2-step registration process, including agreement to the telemedicine service terms and opting in for service notifications
2. Appointment booking: requesting a telemedicine consultation by selecting a specific date and time and choosing a physician from the provider list
3. Telemedicine consultation: completing the end-to-end telemedicine process, including uploading treatment-related documents (eg, tongue photos), joining the scheduled consultation, selecting a medicine delivery method, making payment, and choosing a delivery time
4. Address entry: adding a medicine delivery address with complete details (street address, district, province, and postcode)
5. Service history: viewing consultation history, including appointment details, physician information, diagnosis, payment records, and medicine delivery status
6. Password reset: resetting the telemedicine account password if forgotten

XSL•FO
RenderX

Collectively, these scenarios ensured validation of all core patient-facing functions in the telemedicine system.

Tasks for the Coordinator Role

The coordinator role scenarios reflected coordinators' primary responsibilities, including reviewing service requests, managing appointments, and monitoring daily schedules. Four scenarios were developed to evaluate system functionality from the coordinators' perspective:

1. Appointment creation: reviewing patient appointment requests, confirming the appointment with the patient via phone, and creating the appointment record in the system
2. Daily schedule review: viewing all patient service appointments scheduled for the day, including health number, name, phone number, date, time, and attending physician
3. Service status review: entering a patient's health number in the search bar to access the profile page and view the current service status
4. Provider information management: editing a physician's information in the system, such as Line ID and Line URL

Tasks for the Physician Role

The physician role scenarios represented physicians' key responsibilities in reviewing patient appointments and treatment-related materials. Two scenarios were simulated:

1. Document reupload request: reviewing patient appointment details and treatment-related documents submitted via the system and then requesting patients to reupload documents when necessary
2. Service history review: accessing a patient's historical records by entering the health number in the search bar to view profile information and past service data.

Tasks for the Pharmacist Role

Three scenarios were developed to evaluate the pharmacist role, which primarily involved managing patient medication and related communications:

1. Expense notification: viewing the patient's health number, name, and delivery method and then entering service, medication, and shipping costs to inform the patient
2. Delivery detail review: accessing patient records to confirm delivery method and scheduled delivery time, including health number, name, phone number, and delivery type and time
3. Service status review: entering the patient's health number in the search bar to access the profile page and review the current service status

Tasks for the Finance Personnel Role

Two scenarios were created to reflect the finance personnel's responsibilities regarding payment verification and follow-up:

1. Payment reupload request: reviewing the "waiting for payment confirmation" list, assessing uploaded payment bills, and notifying patients when resubmission was required
2. Review of patients with unpaid consultations: viewing details of patients with unpaid consultations, including health number, name, phone number, and service date and time

Measurement

Three quantitative usability metrics were used to evaluate task performance: success rate, number of interactions, and time on task. The success rate measured the effectiveness of the system and represented the ultimate outcome of usability testing. The number of interactions was used to evaluate efficiency by capturing the steps required for task completion. Time on task measured the time needed to complete each task and served as a primary indicator of efficiency [29]. Together, these metrics provided a comprehensive assessment of task effectiveness and overall system usability.

User satisfaction was assessed through 2 role-specific paper-based questionnaires. The patient questionnaire consisted of 4 sections: understanding of service procedures, ease of service completion, overall satisfaction, and net promoter score (NPS). The NPS, first introduced in 2003, measures customer loyalty by asking participants how likely they are to recommend a service to others, with scores ranging from –100 to +100 [30]. All patient questionnaire items were rated on a scale from 0 to 10, with higher scores indicating greater satisfaction.

The service provider questionnaire consisted of 10 items adapted from the System Usability Scale (SUS) [31]. The SUS includes statements about system usability rated on a 5-point Likert scale from "strongly disagree" (1) to "strongly agree" (5). Scores were converted to a percentile scale ranging from 0 to 100, with higher scores indicating better usability [31].

Procedure

Testing was conducted using the provided mobile phone (participants could select either iOS or Android). Participants were given a scenario and a step-by-step list of tasks to complete, whereas their interactions with the interface were recorded. The focus of this study was to evaluate the usability of the telemedicine application, not the participants' skills or knowledge.

Upon completing all scenarios and tasks, participants filled out a role-specific satisfaction questionnaire reflecting their experience with the telemedicine system. Finally, participants were thanked for their involvement and dismissed.

Ethical Considerations

Ethics approval for this study was obtained from the Burapha University Ethics Committee for Human Research (approval HS 100/2567). All participants were comprehensively briefed in person regarding the procedural details and the data protection protocols implemented to ensure anonymity and subsequently provided informed consent before testing. The consent form stated that no personally identifiable information would be collected, testing would take approximately 30 minutes, and participants retained the right to withdraw at any time. It was explicitly communicated to participants that they retained the right to withdraw from the experiments at any moment without adverse consequences. Furthermore, participation was voluntary, with no compensation offered.

XSL•FO
RenderX

Results

System Design

One of the major advantages of the redesigned telemedicine system was its enhanced operational efficiency, achieved by applying insights from the customer journey map and service blueprint during the redesign process. By automating processes and reducing the number of required steps, the new system streamlined service operations and optimized task completion time for both patients and service providers.

From the patient perspective, the number of steps required to register for telemedicine services decreased from 3 to 1 (a 67% reduction). Similarly, the number of interactions required to book an appointment, make a payment, and enter delivery details decreased from 9 to 5 (a 44% reduction).

From the service provider perspective, efficiency gains were even more substantial. The number of steps required to register a patient, order testing online, and provide service details was reduced from 14 to 2 steps (an 86% reduction), saving approximately 10 minutes per patient. In addition, the number of steps required to schedule an appointment and prepare a patient for diagnosis decreased from 5 to 4 (a 20% reduction). The system also eliminated redundant tasks, such as adding patient tag statuses and sharing Line contact details of physicians, pharmacists, and finance personnel.

Collectively, these reductions enabled service providers to process requests more quickly and allocate more time to higher-value clinical tasks. Table 1 presents a comparison of interaction steps before and after system redesign from both the patient and provider perspectives.

Table . Improvement in interaction steps between users and the telemedicine system after redesign.

Tasks	Improvement in interaction steps		Reduction (%)
	Before redesign	After redesign	
Patient perspective			
Registering for telemedicine service	3	1	67
Communicating with the service provider during the pre- and postservice stages	9	5	44
Service provider perspective			
Registering, internet testing, and informing the patient of service details	14	2	86
Scheduling an appointment and preparing a patient for diagnosis	5	4	20
Adding the tag status of a patient in the system and sending the Line contacts of all stakeholders	5	0	100

System Validation

A comprehensive validation process was conducted to evaluate the usability of the redesigned telemedicine service system. Usability was defined as “the extent to which a product or service can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use” [32].

Three categories of quantitative usability metrics were assessed: effectiveness (success rate), efficiency (time on task), and satisfaction (overall satisfaction rating, NPS, and SUS score). As suggested by Eysenbach [33], validation encompassed both task performance and subjective evaluation. Figure 8 summarizes the number of participants in each role, along with details of their engagement, including assigned scenarios, completion of paper-based questionnaires, and the measurement metrics.

A total of 13 participants were involved in this study. The patient group comprised 6 individuals (n=5, 83% female and n=1, 17% male), with an age range from 23 to 54 years (mean 36.5, SD 9.95 years). The service provider group consisted of 7 individuals representing various departments involved in the telemedicine service system: 1 (14%) female coordinator, 2 (29%) physicians (n=1, 50% male and n=1, 50% female), 2 (29%) female pharmacists, and 2 (29%) female finance personnel.

During the usability testing phase, all participants completed role-specific tasks using the redesigned telemedicine web application. The tasks were designed to cover all functional aspects of the system. Six participants evaluated the patient-facing web application, whereas 7 participants evaluated the provider-facing management platform. Descriptive statistics from the usability testing are shown in Table 2 (patient usability) and Table 3 (service provider usability).

XSL•FO
RenderX

Figure 8. Number of participants in each role and details on their engagement. NPS: net promoter score; SUS: System Usability Scale.

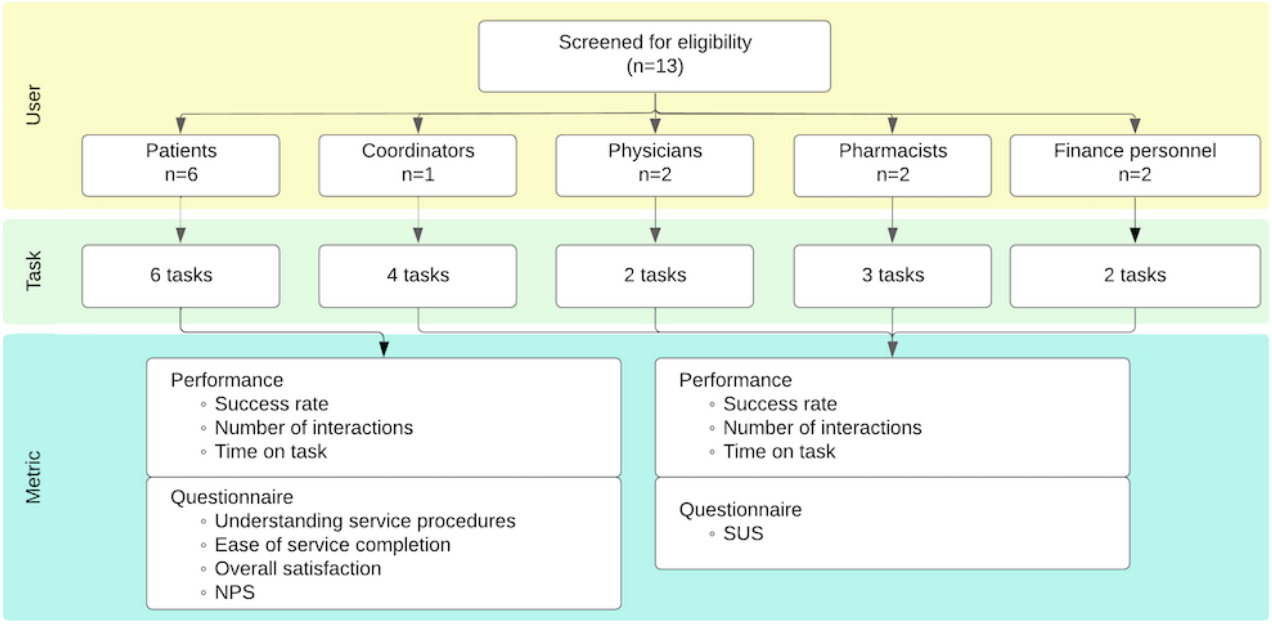


Table . Patient statistical usability testing results (N=6).

Scenarios	Number of interactions		Time on task (s)		Task completion						Success rate (%)
	Mean (SD)	Range	Mean (SD)	Range	Participants						
					#1	#2	#3	#4	#5	#6	
Registration	33.70 (0.52)	33 - 34	203.00 (61.96)	135 - 285	Pass	Pass	Pass	Pass	Pass	Pass	100
Appointment booking	4.00 (0.00)	4 - 4	30.20 (3.76)	27 - 37	Pass	Pass	Pass	Pass	Pass	Pass	100
Telemedicine consultation	21.20 (1.94)	20 - 25	204.80 (69.44)	107 - 294	Pass	Pass	Pass	Pass	Pass	Pass	100
Address entry	5.80 (1.33)	5 - 8	14.80 (6.82)	9 - 26	Pass	Pass	Pass	Pass	Pass	Pass	100
Service history	3.00 (0.00)	3 - 3	17.00 (10.18)	7 - 30	Pass	Pass	Pass	Pass	Pass	Pass	100
Password reset	6.70 (0.82)	6 - 8	115.20 (36.89)	72 - 155	Pass	Pass	Pass	Pass	Pass	Pass	100

For the patient-facing telemedicine web application, all patient participants successfully completed the full complement of six tasks. Variability was observed in the number of interactions and task completion times, which may reflect individual differences in participants' technological proficiency.

For the provider-facing telemedicine management platform, task success rates were similarly high, with 100% completion for most roles. The exception was the pharmacist role, which achieved a task success rate of 83.33%. The number of interactions and task completion times was not recorded for providers as the redesigned system allowed most tasks to be completed within a very short time. Variability in the pharmacist results may be attributed to individual differences in interface familiarity or task complexity (Table 3). Overall, the findings demonstrate that the provider platform was highly usable for most staff roles.

User satisfaction data are shown in Table 4. Among patients, the NPS was 67 based on feedback from 6 participants: 4 (67%) promoters, 2 (33%) passives, and no detractors. The mean satisfaction score was 9.2/10, indicating excellent user experience. For service providers, usability was evaluated using the SUS. The mean SUS score was 71.4 (SD 20.76) based on feedback from 7 participants, placing the system in the above-average usability range [34].

XSL•FO
RenderX

Table . Service provider statistical usability testing results (N=7).

Roles and tasks		Participants for each role		Success rate (%)
		1	2	
Coordinators (n=1)				100
	Task 1	Pass	—[a]	—
	Task 2	Pass	—	—
	Task 3	Pass	—	—
	Task 4	Pass	—	—
Physicians (n=2)				100
	Task 1	Pass	Pass	—
	Task 2	Pass	Pass	—
Pharmacists (n=2)				83.33
	Task 1	Fail	Pass	—
	Task 2	Pass	Pass	—
	Task 3	Pass	Pass	—
Financial personnel (n=2)				100
	Task 1	Pass	Pass	—
	Task 2	Pass	Pass	—

[a]Not applicable.

Table . Satisfaction scores of patients and System Usability Scale (SUS) results for service providers.

Item		Score, mean (SD)	Score, median (IQR)	Score range
Patients (n=6)				
	Knew the procedure well	8.7 (1.51)	8 (7.25 - 10)	7 - 10
	Completing the service felt easy	8.8 (0.98)	8 (8 - 9.75)	8 - 10
	Overall satisfaction with the service	9.2 (0.75)	9 (9 - 9.75)	8 - 10
Service providers (n=7)				
	SUS	71.4 (20.76)	65 (55 - 86.25)	52.5 - 100

Discussion

Principal Findings

This study developed and validated a new telemedicine service system for a TCM clinic centered on separate patient- and provider-facing web applications. The redesign was guided by service design tools, customer journey mapping, and service blueprinting that revealed inefficiencies in the legacy system, including fragmented communication channels, redundant steps, and undefined waiting times.

By reconfiguring these workflows into a logically harmonized digital design, the new system integrated patient-facing processes with provider operations in a coherent manner. Automation eliminated repetitive administrative tasks, communication was consolidated, and service notifications connected actions across roles. These features not only simplified the patient journey but also aligned provider tasks with patient needs in real time, reflecting a deliberate harmonization of activities across the digital service process.

Usability testing with 13 participants (n=6, 46% patients and n=7, 54% providers) demonstrated the effectiveness of this design. Patients achieved a 100% task success rate, with completion times within acceptable ranges, supporting learnability and intuitive system use [35,36]. Providers achieved similarly high success rates, with the exception of pharmacists (83.33%), whose tasks were inherently more complex. The observed errors were categorized as slips [37,38], suggesting that the design was fundamentally sound but required refinement in more information-intensive workflows, such as those for pharmacy.

The logically harmonized design also translated into high user satisfaction. Patients reported an average satisfaction score of 9.2/10 and an NPS of 67, both indicating excellent performance and strong acceptability [30,39,40]. This suggests that patients not only found the system easy to use but were also motivated

XSL•FO
RenderX

to recommend it, an important marker of adoption in digital health.

Among providers, the average SUS score was 71.4 (SD 20.76), above the average usability benchmark [34]. The variability in provider scores reflected differing levels of familiarity and adaptation to the new workflows. Importantly, this underscores a central challenge of digital transformation: harmonizing not only technical systems but also user practices and mindsets. With training and support tools (eg, video tutorials, workflow guides, and quick-reference materials) [35,36], providers can adapt to the harmonized processes more readily, leading to even higher usability and acceptance.

A distinctive contribution of this study is the introduction of service notifications as an orchestration mechanism in the harmonized design. Patients received automated alerts about next steps (eg, physician availability, payment confirmation, and delivery status), whereas providers were notified when specific actions were required (eg, reviewing documents, issuing receipts, and preparing medication). These notifications bridged the frontstage and backstage operations, ensuring that both patients and providers had a shared, transparent view of the service process. This alignment reduced uncertainty, supported timely action, and reinforced the logically harmonized structure of the digital service.

Patients highlighted improvements in convenience and accessibility, particularly for those with travel barriers, whereas providers reported efficiency gains from reduced redundant tasks and better cross-role coordination. Together, these results illustrate that harmonization of digital workflows not only improves efficiency but also strengthens the interconnectedness of patient and provider experiences in telemedicine.

This study contributes to the broader discourse on digital transformation in health care by demonstrating that success lies not only in digitizing existing processes but also in logically harmonizing them across all stakeholders. The integration of customer journey maps and service blueprints provided a structured pathway to identify misalignments and redesign them into a unified digital flow.

While service design tools remain underused in TCM and telemedicine contexts [41,42], our findings show that they can be instrumental in achieving harmonized digital transformation. The resulting framework could serve as a TCM telemedicine service design model or as guidelines for telehealth service design more broadly. By bridging service design thinking with digital health practice, this study highlights how logically harmonized design enables patient-centered care, reduces provider burden, and strengthens overall health care efficiency.

Limitations

This study faced several limitations. First, the case study clinic's IT infrastructure constrained the technical development of the web application. For example, while Node.js supported the system backbone, the integration of Web Real-Time Communication was limited due to the complexity of configuring Traversal Using Relays Around Network Address Translators and Session Traversal Utilities for Network Address Translators servers. These challenges affected the implementation of smooth real-time communication features.

Second, the patient sample (n=6) included participants aged 23 to 54 years, which does not fully represent the broader adult population typically served by the clinic, particularly older adults (with those aged 41-50 years being the most common age group in clinic data). As older adults are likely to benefit most from remote care solutions, future telemedicine research should prioritize this demographic, focusing on usability and accessibility tailored to older users.

Future Work

One of the distinctive challenges in TCM telehealth is the reliance on visual tongue examination for diagnosis. The current system addressed this by enabling patients to upload tongue photographs for physicians to review, as well as request resubmission if necessary. Future research could extend this by integrating machine learning and artificial intelligence to automate initial screening of tongue images. Automated feedback on image quality could reduce wasted time, decrease physician workload, and further improve diagnostic efficiency.

Beyond the TCM context, the system could be adapted for other traditional medical practices, such as Ayurveda. Although attempts to implement telemedicine in Ayurveda have been reported, barriers remain due to limited physician awareness and acceptability [43]. Applying the service design approach demonstrated in this study may support the integration of such services, particularly where workflows require specialized or customized patient data.

Finally, the use of service design tools (customer journey maps and service blueprints) in this study provides a replicable methodology for integrating new digital technologies into health care services. These tools can guide the redesign of service processes to optimize delivery, improve user experiences, and enhance both patient satisfaction and provider efficiency.

Conclusions

This study implemented strategies for both patients and service providers to improve the telemedicine service system and streamline its operations, with a focus on reducing human effort, interaction costs, and waiting times.

For patients, 3 strategies were central: consolidating service channels into a single platform, digitizing processes through web applications, and visualizing service status. These interventions lowered interaction costs by providing an all-in-one, easily accessible platform. Visualization of service status reduced dissatisfaction from undefined waiting times and improved patient understanding of service steps. Usability testing confirmed the positive effects of these improvements, with patients reporting an NPS of 67 and an average satisfaction score of 9.2, reflecting excellent performance.

For service providers, 2 strategies were central: digitizing tasks and visualizing service status. This reduces repetitive administrative work, facilitating tracking of patient responses and supporting more efficient communication across roles. While the system markedly reduced provider workload, structured training at the start of implementation is

XSL•FO
RenderX

recommended to support adoption, build confidence, and promote trust in the system.

Collectively, the results demonstrate that a logically harmonized digital transformation informed by service design methodologies can enhance telemedicine effectiveness and efficiency. By aligning redesigned workflows with patient and provider needs, the system improved usability and satisfaction while contributing to higher-quality health care delivery.

Acknowledgments

This research did not receive any specific funding from agencies in the public, commercial, or not-for-profit sectors.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to participants not providing informed consent for the data to be shared outside of this study but are available from the corresponding author on reasonable request.

Authors' Contributions

AJ: conceptualization, data analysis, data curation, methodology, project administration, system validation, visualization, and writing—review and editing. RL and RK: data collection and investigation, data analysis, system design, system validation, visualization, and writing—draft manuscript. OK: data analysis, system design, and system validation. WA: conceptualization, data analysis, data curation, methodology, system validation, visualization, and writing—review and editing. All authors agreed to be accountable for all aspects of the work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Enlarged image of the “to be” service blueprint of the redesigned telemedicine service (provider perspective).
[PNG File, 491 KB - humanfactors_v12i1e76752_app1.png]

Multimedia Appendix 2
Enlarged image of the patient web application interface optimized for mobile devices.
[PNG File, 3347 KB - humanfactors_v12i1e76752_app2.png]

References

1. Shirzadfar H, Lotfi F. The evolution and transformation of telemedicine. Int J Biosens Bioelectron 2017;3(4):303-306. [doi: 10.15406/ijbsbe.2017.03.00070]
2. Telemedicine: opportunities and developments in member states: report on the second global survey on ehealth. World Health Organization. URL: https://apps.who.int/iris/handle/10665/44497 [accessed 2023-05-04]
3. Ftouni R, AlJardali B, Hamdanieh M, Ftouni L, Salem N. Challenges of telemedicine during the COVID-19 pandemic: a systematic review. BMC Med Inform Decis Mak 2022 Aug 3;22(1):207. [doi: 10.1186/s12911-022-01952-0] [Medline: 35922817]
4. Strehle EM, Shabde N. One hundred years of telemedicine: does this new technology have a place in paediatrics? Arch Dis Child 2006 Dec;91(12):956-959. [doi: 10.1136/adc.2006.099622] [Medline: 17119071]
5. Loh PK, Donaldson M, Flicker L, Maher S, Goldswain P. Development of a telemedicine protocol for the diagnosis of Alzheimer’s disease. J Telemed Telecare 2007;13(2):90-94. [doi: 10.1258/135763307780096159] [Medline: 17359573]
6. Mishra V, Sharma MG. Telemedicine as frugal intervention to health care: a case of diabetes management. Int J Healthc Manag 2022 Apr 3;15(2):165-170. [doi: 10.1080/20479700.2020.1870348]
7. Watson S. Telehealth: the advantages and disadvantages. Harvard Health Publishing. 2020 Oct 12. URL: https://www.health.harvard.edu/staying-healthy/telehealth-the-advantages-and-disadvantages [accessed 2023-05-04]
8. Smith WR, Atala AJ, Terlecki RP, Kelly EE, Matthews CA. Implementation guide for rapid integration of an outpatient telemedicine program during the COVID-19 pandemic. J Am Coll Surg 2020 Aug;231(2):216-222. [doi: 10.1016/j.jamcollsurg.2020.04.030] [Medline: 32360960]
9. Kapoor S, Eldib A, Hiasat J, et al. Developing a pediatric ophthalmology telemedicine program in the COVID-19 crisis. J AAPOS 2020 Aug;24(4):204-208. [doi: 10.1016/j.jaapos.2020.05.008] [Medline: 32890736]
10. Anthony Jnr. B. Integrating telemedicine to support digital health care for the management of COVID-19 pandemic. Int J Healthc Manag 2021 Jan 2;14(1):280-289. [doi: 10.1080/20479700.2020.1870354]

XSL•FO
RenderX

11. Koonin LM, Hoots B, Tsang CA, et al. Trends in the use of telehealth during the emergence of the COVID-19 pandemic - United States, January-March 2020. MMWR Morb Mortal Wkly Rep 2020 Oct 30;69(43):1595-1599. [doi: 10.15585/mmwr.mm6943a3] [Medline: 33119561]
12. Seivert S, Badowski ME. The rise of telemedicine: lessons from a global pandemic. EMJ Innov 2021;5(1):64-69. [doi: 10.33590/emjinnov/20-00239]
13. Ma Q, Sun D, Tan Z, et al. Usage and perceptions of telemedicine among health care professionals in China. Int J Med Inform 2022 Oct;166:104856. [doi: 10.1016/j.ijmedinf.2022.104856] [Medline: 36037593]
14. Sajid M, Zakkariya KA, Peethambaran M. Predicting virtual care continuance intention in the post-COVID world: empirical evidence from an emerging economy. Int J Healthc Manag 2023 Jul 3;16(3):327-335. [doi: 10.1080/20479700.2022.2102119]
15. Rosenbaum MS, Otalora ML, Ramírez GC. How to create a realistic customer journey map. Bus Horiz 2017 Jan;60(1):143-150. [doi: 10.1016/j.bushor.2016.09.010]
16. Kalbach J. Mapping Experiences, 4th edition: O'Reilly Media; 2016.
17. Bitner MJ, Ostrom AL, Morgan FN. Service blueprinting: a practical technique for service innovation. Calif Manage Rev 2008 Apr;50(3):66-94. [doi: 10.2307/41166446]
18. Chokshi SK, Mann DM. Innovating from within: a process model for user-centered digital development in academic medical centers. JMIR Hum Factors 2018 Dec 19;5(4):e11048. [doi: 10.2196/11048] [Medline: 30567688]
19. Ingaldi M, Klimecka-Tatar D. Digitization of the service provision process - requirements and readiness of the small and medium-sized enterprise sector. Procedia Comput Sci 2022;200:237-246. [doi: 10.1016/j.procs.2022.01.222]
20. Vial G. Understanding digital transformation: a review and a research agenda. J Strateg Inf Syst 2019 Jun;28(2):118-144. [doi: 10.1016/j.jsis.2019.01.003]
21. Chu H, Westbrook RA, Njue-Marendes S, Giordano TP, Dang BN. The psychology of the wait time experience - what clinics can do to manage the waiting experience for patients: a longitudinal, qualitative study. BMC Health Serv Res 2019 Jul 8;19(1):459-467. [doi: 10.1186/s12913-019-4301-0] [Medline: 31286957]
22. Lallemand C, Gronier G. Enhancing user experience during waiting time in HCI: contributions of cognitive psychology. Presented at: Proceedings of the 2012 Designing Interactive Systems Conference (DIS '12); Jun 11-15, 2012; Newcastle, UK. [doi: 10.1145/2317956.2318069]
23. Figma design. Figma. URL: https://www.figma.com/design [accessed 2025-09-01]
24. Chen PPS. The entity-relationship model—toward a unified view of data. ACM Trans Database Syst 1976 Mar;1(1):9-36. [doi: 10.1145/320434.320440]
25. Jiamsanguanwong A, Ophaswongse C, Chansirinthorn C, Kitirattragarn N, Kittithreerapronchai O. Improving patients' experience concerning insufficient informational flow to patients during COVID-19 pandemic: case study of a traditional Chinese medicine clinic. Int J Healthc Manag 2023 Jan 2;16(1):51-60. [doi: 10.1080/20479700.2022.2071801]
26. Jiamsanguanwong A, Tientrakul P, Sookseng F, Ophaswongse C, Kittithreerapronchai O. System improvement of medicine delivery service: case study of traditional Chinese medicine. Int J Healthc Manag 2021 Oct 2;14(4):984-992. [doi: 10.1080/20479700.2020.1723945]
27. Etikan I, Musa SA, Alkassim RS. Comparison of convenience sampling and purposive sampling. Am J Theor Appl Stat 2016;5(1):1. [doi: 10.11648/j.ajtas.20160501.11]
28. Rovinelli RJ, Hambleton RK. On the use of content specialists in the assessment of criterion-referenced test item validity. Tijdschr Onderwijs Res 1977;2:49-60 [FREE Full text]
29. Frøkjær E, Hertzum M, Hornbæk K. Measuring usability: are effectiveness, efficiency, and satisfaction really correlated? Presented at: Proceedings of the 2000 SIGCHI conference on Human Factors in Computing Systems (CHI '00); Apr 1-6, 2000; New York, NY. [doi: 10.1145/332040.332455]
30. Reichheld FF. The one number you need to grow. Harv Bus Rev 2003 Dec;81(12):46-54. [Medline: 14712543]
31. Brooke J. SUS: a quick and dirty usability scale. In: Jordan PW, Thomas B, Weerdmeester B, McClelland AL, editors. Usability Evaluation in Industry: Taylor and Francis; 1996:189-194.
32. Common industry format for usability test reports, ANSI-NCITS 354–2001. National Institute of Standards and Technology. 2001. URL: https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=151449 [accessed 2025-10-15]
33. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126. [doi: 10.2196/jmir.1923] [Medline: 22209829]
34. Hyzy M, Bond R, Mulvenna M, et al. System usability scale benchmarking for digital health apps: meta-analysis. JMIR Mhealth Uhealth 2022 Aug 18;10(8):e37290. [doi: 10.2196/37290] [Medline: 35980732]
35. Nielsen J. Usability Engineering: Morgan Kaufmann Publishers; 1993.
36. Nielsen J. Enhancing the explanatory power of usability heuristics. Presented at: Proceedings of the 1994 SIGCHI Conference on Human Factors in Computing Systems (CHI' 94); Apr 24-28, 1994; Boston, MA. [doi: 10.1145/191666.191729]
37. Lewis C, Norman D. Designing for error. In: Norman DA, Draper SW, editors. User Centered System Design: New Perspectives on Human–Computer Interaction: Lawrence Erlbaum Associates; 1986:411-432. [doi: 10.1201/b15703-20]
38. PRüMPER J, Zapf D, Brodbeck FC, Frese M. Some surprising differences between novice and expert errors in computerized office work. Behav Inf Technol 1992 Nov;11(6):319-328. [doi: 10.1080/01449299208924353]

XSL•FO
RenderX

39. Bevan N, Kirakowski J, Maissel J. Human aspects in computing, design and use of interactive systems and work with terminals. Presented at: Proceedings of the 4th International Conference on Human–Computer Interaction; Sep 1-9, 1991; Stuttgart, Germany URL: https://www.hci.international/index.php?module=conference&CF_op=view&CF_id=17 [accessed 2025-05-05]
40. Australian Healthcare Index report November 2022. Australian Healthcare Index. 2022. URL: https://australianhealthcareindex.com.au/australian-healthcare-index-november-2022/ [accessed 2024-09-03]
41. Joseph AL, Kushniruk AW, Borycki EM. Patient journey mapping: current practices, challenges and future opportunities in healthcare. Knowl Manag e-Learn 2020 Dec 26;12(4):387-404. [doi: 10.34105/j.kmel.2020.12.021]
42. Joseph AL, Monkman H, Kushniruk A, Quintana Y. Exploring patient journey mapping and the learning health system: scoping review. JMIR Hum Factors 2023 Feb 27;10:e43966. [doi: 10.2196/43966] [Medline: 36848189]
43. Azeez K K NA, Krishnan A, H K S, et al. Telemedicine usage among ayurvedic physicians: current scenario - a cross-sectional study. Int J Exp Res Rev 2023;30:402-407. [doi: 10.52756/ijerr.2023.v30.037]

Abbreviations

NPS: net promoter score
SUS: System Usability Scale
TCM: traditional Chinese medicine

Edited by C Jacob; submitted 01.05.25; peer-reviewed by A Khamees, T Kanno; revised version received 28.09.25; accepted 07.10.25; published 28.10.25.

Please cite as:
Jiamsanguanwong A, Luo R, Kaingam R, Kittithreerapronchai O, Authayarat W
Redesigning Telemedicine for Traditional Chinese Medicine: Service Design Approach to Digital Transformation
JMIR Hum Factors 2025;12:e76752
URL: https://humanfactors.jmir.org/2025/1/e76752
doi:10.2196/76752

© Arisara Jiamsanguanwong, Romanee Luo, Ratchapoom Kaingam, Oran Kittithreerapronchai, Waratta Authayarat. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 28.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

An AI-Based Clinical Decision Support System for Antibiotic Therapy in Sepsis (KINBIOTICS): Use Case Analysis

Juliane Andrea Düvel[1], MSc; David Lampe[2], MSc; Maren Kirchner[2], BSc; Svenja Elkenkamp[2], MEd, MSc; Philipp Cimiano[3], Prof Dr; Christoph Düsing[3], MSc; Hannah Marchi[4], MSc; Sophie Schmiegel[4], MSc; Christiane Fuchs[4], Prof Dr; Simon Claßen[5], MD; Kirsten-Laura Meier[5], MD; Rainer Borgstedt[6], Dr; Sebastian Rehberg[6], Prof Dr; Wolfgang Greiner[2], Prof Dr

[1]Centre for electronic Public Health Research (CePHR), School of Public Health, Bielefeld University, P.O. Box 10 01 31, Bielefeld, Germany

[2]AG 5 - Health Economics and Health Care Management, School of Public Health, Bielefeld University, Bielefeld, Germany

[3]AG Semantic Computing, Technical Faculty, Bielefeld University, Bielefeld, Germany

[4]Data Science Group, Faculty of Business Administration and Economics, Bielefeld University, Bielefeld, Germany

[5]Department of Anaesthesiology, Surgical Intensive Care, Emergency Medicine, and Pain Therapy, Hospital Bielefeld, University Hospital Bielefeld, Bielefeld, Germany

[6]Department of Anaesthesiology, Intensive Care, Emergency Medicine, Transfusion Medicine and Pain Therapy, Campus Bielefeld-Bethel, University Hospital Bielefeld, Bielefeld, Germany

Corresponding Author:
Juliane Andrea Düvel, MSc
Centre for electronic Public Health Research (CePHR), School of Public Health, Bielefeld University, P.O. Box 10 01 31, Bielefeld, Germany

Abstract

Background: Antimicrobial resistances pose significant challenges in health care systems. Clinical decision support systems (CDSSs) represent a potential strategy for promoting a more targeted and guideline-based use of antibiotics. The integration of artificial intelligence (AI) into these systems has the potential to support physicians in selecting the most effective drug therapy for a given patient.

Objective: This study aimed to analyze the feasibility of an AI-based CDSS pilot version for antibiotic therapy in sepsis patients and identify facilitating and inhibiting conditions for its implementation in intensive care medicine.

Methods: The evaluation was conducted in 2 steps, using a qualitative methodology. Initially, expert interviews were conducted, in which intensive care physicians were asked to assess the AI-based recommendations for antibiotic therapy in terms of plausibility, layout, and design. Subsequently, focus group interviews were conducted to examine the technology acceptance of the AI-based CDSS. The interviews were anonymized and evaluated using content analysis.

Results: In terms of the feasibility, barriers included variability in previous antibiotic administration practices, which affected the predictive ability of AI recommendations, and the increased effort required to justify deviations from these recommendations. Physicians' confidence in accepting or rejecting recommendations depended on their level of professional experience. The ability to re-evaluate CDSS recommendations and an intuitive, user-friendly system design were identified as factors that enhanced acceptance and usability. Overall, barriers included low levels of digitization in clinical practice, limited availability of cross-sectoral data, and negative previous experiences with CDSSs. Conversely, facilitators to CDSS implementation were potential time savings, physicians' openness to adopting new technologies, and positive previous experiences.

Conclusions: Early integration of users is beneficial for both the identification of relevant context factors and the further development of an effective CDSS. Overall, the potential of AI-based CDSSs is offset by inhibiting contextual conditions that impede its acceptance and implementation. The advancement of AI-based CDSSs and the mitigation of these inhibiting conditions are crucial for the realization of its full potential.

(JMIR Hum Factors 2025;12:e66699) doi:10.2196/66699

KEYWORDS

CDSS; use case analysis; technology acceptance; sepsis; infection; infectious disease; antimicrobial resistance; clinical decision support system; decision-making; clinical support; machine learning; ML; artificial intelligence; AI; algorithm; model; analytics; predictive models; deep learning; early warning; early detection

XSL•FO
RenderX

Introduction

Sepsis infections caused by pathogens with antimicrobial resistance (AMR) represent a significant global challenge in health care [1,2]. In 2017, there were 48.9 million new cases of sepsis and 11 million deaths related to sepsis, accounting for 19.7% of all global deaths [3]. In Germany, sepsis incidence increased by an average of 5.7% per year, from 280 cases in 2010 to 370 cases in 2015 per 100,000 individuals [4]. A recent meta-analysis indicated that the 30-day mortality rate for sepsis in Germany was estimated to be 26.5%, which is consistent with the observed rates in North America and Europe [5]. Furthermore, over 1.27 million deaths per year are attributed to AMR worldwide. In 2019, there were 9650 deaths attributable to AMR (mortality rate of 5 per 100,000) and 45,700 deaths associated with AMR (mortality rate of 22 per 100,000) in Germany [6]. Recent data from Germany show heterogeneous trends in AMR proportions of infected patients underscoring the urgent need for enhanced infection prevention measures to limit AMR spread [7]. Inappropriate antibiotic prescribing represents a significant contributing factor [8]. In particular, the prolonged or improper use of nonspecific, broad-spectrum antibiotics is highly problematic, as these antibiotics provide symptomatic relief but also facilitate the development of resistance in other bacteria [9,10].

Clinical decision support systems (CDSSs) provide a solution for promoting targeted and guideline-based antibiotic prescribing [11-13]. CDSSs adopt a variety of forms and are integrated into routine practice. Despite their diverse nature, they share a common objective: to assist medical professionals in identifying the most appropriate form of therapy for each patient, based on existing data and established guidelines, through the use of programmed algorithms or artificial intelligence (AI) [14,15]. In the context of CDSSs, previous studies have demonstrated a decrease in antibiotic prescription rates, more rapid initiation of appropriate antibiotic therapy for patients, and improved clinical outcomes, such as reduced mortality, increased antibiotic-free days, and fewer medical complications [16-19]. However, CDSSs do not always improve clinical practice [20]. For example, there is still insufficient evidence to conclusively show positive effects on therapy duration, dosage, or adherence to clinical guidelines [11,16,21].

Despite the growing body of research on AI-based CDSSs in intensive care units (ICUs) [22,23], there is a notable gap in the implementation of AI tools in routine care in general [24] and the availability of AI-based CDSSs for sepsis in the German health care system. The KINBIOTICS (translated as “AI-based decision support for antibiotic therapy”) project, funded by the German Federal Ministry of Health, aimed to train AI-based algorithms and improve the prediction of suitable antibiotics for sepsis infections on the basis of a comprehensive data set. In addition to the development of an AI-based CDSS, a cross-sectoral resistance observatory and a new rapid test for sequencing the antibiotic genome were developed [25].

Engaging clinicians in the design and development of CDSSs is often suggested as a strategy to enhance the alignment between the system and the needs of its users [26-28]. Interviews and expert groups are approaches used for preimplementation clinician involvement [29]. Insights into clinicians’ views provide valuable information on the barriers and facilitators that affect their willingness to adopt and use CDSSs in practice [30]. By examining clinicians’ perceptions, the study seeks to inform targeted strategies to improve the CDSS’s design, usability, and relevance, thereby promoting more effective and widespread adoption in clinical practice.

This study aimed to perform a use case analysis of a pilot version of the AI-based CDSS within an ICU setting. The objective was to analyze the clinical decision-making processes, and adopt a more comprehensive perspective on the factors that facilitate or hinder the future implementation of such a system.

Methods

Reporting is based on the Standards for Reporting Qualitative Research (SRQR) [31]. The SRQR checklist is provided in Multimedia Appendix 1.

Study Design and Setting

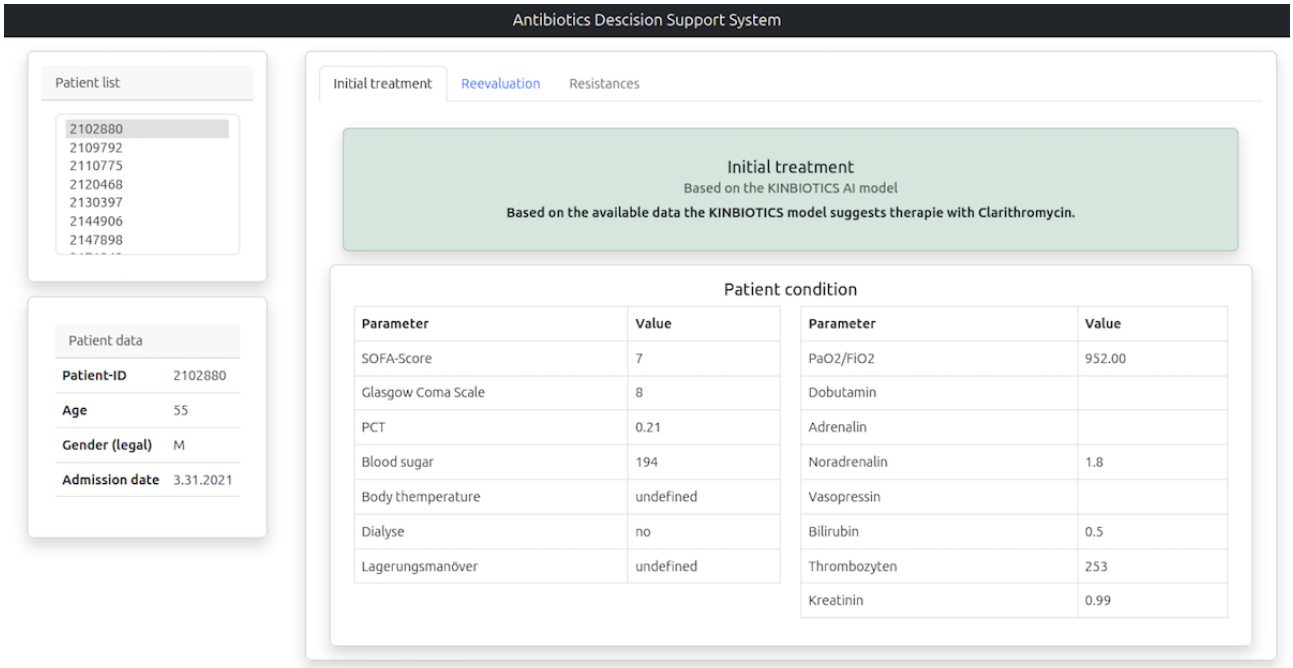
A 2-stage base model (random forest at both stages) was developed for the initial therapy using patient data, laboratory data, and clinical data. Subsequently, the model was optimized using a variety of parameters, including the number of decision trees and tree depth (the model specifications and results will be published separately once the AI-based CDSS model has been finalized). The developed model was trained and tested on data from 1 of the 3 participating clinics and then evaluated for robustness using data from a second clinic. Although the performance of the underlying AI model was insufficient in terms of accuracy, specificity and sensitivity at the time of the evaluation, it was crucial at this time to gain insights into the decision-making mechanisms of the physicians. This was undertaken in order to scrutinize the variables included in the initial model and, where necessary, to supplement or adjust them. Accordingly, the preliminary prototype was not modified during the interview phase. However, subsequent to the interviews, the model underwent alterations and is still under development.

The pilot version of the AI-based CDSS was evaluated in a 2-stage semistructured qualitative process [32]. With regard to feasibility, physicians were each shown 5 exemplary sepsis patient cases in a desktop version of the CDSS. Cases were selected randomly from the evaluation dataset and were identical for each interviewee. In addition to basic patient information, this contained relevant vital parameters, and the treatment recommendation determined by the AI-based CDSS (Figure 1). This antibiotic therapy recommendation was presented to the physicians both for the admission situation to the ICU (initial treatment; the start of antibiotic therapy) and for a possible therapy correction at the time the microbiological findings were available (re-evaluation; usually 48-72 hours after start of antibiotic therapy). The physicians were informed during the individual interviews that the CDSS is still in a pilot status and that the recommendations are likely not yet reliable. The physicians were asked to assess and evaluate the intensive care situation at both time points as well as the AI-based model

XSL•FO
RenderX

recommendation for antibiotic therapy. In addition to the plausibility of the content of the therapy recommendations, the design and layout of the pilot version were evaluated.

Figure 1. Screenshot of the pilot version of the artificial intelligence–based clinical decision support system.



Following the individual interviews, physicians from each clinic were interviewed again as part of a focus group interview. The theoretical focus here was on technology acceptance of AI-based CDSSs and the identification of challenges and conductive conditions with regard to future implementation.

Expert Selection

Expert interviews were conducted with intensive care physicians from the 3 clinical centers involved in the project at the University Medical Center East Westphalia-Lippe. The interviews were conducted on a voluntary basis and the potential participants were selected by the respective project partners in the 3 clinical centers.

Data Collection

Between October and November 2023, the face-to-face expert interviews and focus groups were conducted. The interview guideline (Textbox 1) for the focus group interviews was developed on the basis of the extended Unified Theory of Acceptance and Use of Technology (UTAUT 2) [33,34]. The UTAUT has recently being used to explain technology adoption among health care practitioners' intention to use AI-based CDSSs [35].

XSL•FO
RenderX

Textbox 1. Interview guideline for the focus groups (based on the Unified Theory of Acceptance and Use of Technology 2).

- Opening question
- What were the objectives of today's meeting?
- Questions about the tool
- Please describe your impressions of the KINBIOTICS user interface.
- What information may have been lacking in your medical decision-making process that led to the prescription of antibiotics?
- Please indicate which aspects you consider to be beneficial for a potential application.
- What additional functions would be beneficial to include in the KINBIOTICS user interface to ensure effective utilization in routine clinical practice?
- Questions regarding the use of artificial intelligence (AI) in general
- To what extent do digital tools contribute to the processes undertaken in your day-to-day work?
- Please indicate which digital applications you currently use in your professional activities.
- Please indicate which digital health applications you use in your private life.
- Please indicate your opinion on the utilization of AI, and provide your expectations regarding the implementation of AI in healthcare, both in general and on your ward.
- Please describe your previous experience of contact with AI and clinical decision support systems.
- What is the prevailing attitude in your clinical setting, on the ward and among your professional colleagues with regard to the utilization of artificial intelligence in everyday medical procedures?
- In the hypothetical scenario of utilizing such a tool like KINBIOTICS in tomorrow's clinical decision-making process, what potential challenges might be identified?
- Questions about future utilization of a clinical decision support system
- Please describe the benefits you perceive in the utilization of AI and clinical decision support systems in your everyday professional practice.
- In order to apply such a system in medical decision-making, what prerequisites must be met?
- Please indicate whether you identify any additional aspects that should be addressed for use in everyday clinical practice.

The interviews were recorded by audio and subsequently transcribed. After complete transcription, the audio recordings were irrevocably deleted in accordance with the data protection policy.

Data Analysis

The data analysis of the transcribed interview material was conducted deductively and inductively based on categories according to Kuckartz (2018) [36]. The procedure for the individual interviews primarily followed evaluative content analysis, while the focus groups were analyzed in terms of content structure. To gain as much knowledge as possible, the transcripts of the individual interviews were also analyzed with regard to possible content-structuring findings. The transcripts were analyzed anonymously using the qualitative analysis software MAXQDA (version 2022, VERBI) [37]. The interview analyses and category assignment were carried out independently by 2 researchers (JAD and MK). In the case of discrepancies, a third researcher (DL) was consulted. Categories were adapted or reformulated by the interdisciplinary research team (JAD, MK, DL, SE, and WG).

Ethical Considerations

Informed consent was obtained before the interviews. A positive ethics vote (November 22, 2021) from the ethics committee of the Medical Association of Westphalia-Lippe and the Westphalian Wilhelms University of Münster was received for the entire project (No. 2021-699-f-S). The interview transcripts were anonymized. The interviewees did not receive any compensation; participation was on a voluntary basis.

Results

Sample

A total of 19 individual interviews and 3 focus group interviews were conducted. The distribution of respondents with regard to gender, age, and professional experience was relatively balanced, with 10 (53%) of the respondents identifying as male. In each case, 7 (37%) respondents were aged 30-40 years, and 40-50 years. In total, 5 (26%) respondents had less than 5 years or 11 - 20 years of professional experience, 4 (21%) had between 5 and 10 years, and 3 (16%) respondents indicated that they had more than 20 years of professional medical experience. The duration of the expert interviews ranged from 17 to 43

XSL•FO
RenderX

minutes, while the average duration of the focus group interviews was 1 hour.

Feasibility of the AI-Based Clinical Decision Support System

The approval or rejection of the AI-based antibiotic recommendations was heterogeneous and primarily dependent on 2 factors: first, the plausibility of the AI-based individual therapy (change) recommendations (refer to quotes 1 and 2). It should be noted that not every detailed recommendation was already plausible or medically advisable due to the pilot status of the CDSS respective example cases.

> ***Quote 1:****The system now suggests piperacillin/tazobactam. Of course, this is a very far-fetched suggestion for primary antibiotics when it comes to designing a calculated antibiotic therapy. For me, this would not be the first drug of choice for patients who arrive at the hospital as primary patients, who have been treatment-naïve so far and have not yet received antibiotic therapy [...].* [Expert 5, Clinical Center 3]

> ***Quote 2:****So, [piperacillin/tazobactam]. Well, difficult [...] with a completely normal PCT [Procalcitonin], I would find it difficult [...] to start directly with such a broad antibiotic.* [Expert 8, Clinical Center 2]

> ***Quote 3:****[...] So, [...] such a recommendation, I'm very sure, will be malpractice in the next three to five years [...]. This reflects the reality, but this is the worst thing that can happen, to treat a Staphylococcus aureus bacteremia with [piperacillin/tazobactam], that is an absolute no-go.* [Expert 4, Clinical Center 1]

Second, the evaluation of the recommendations is dependent upon the level of professional experience (quote 3). Accordingly, the more advanced the professional experience, the more confident a recommendation was to be approved or rejected. However, the data indicate that the majority of respondents expressed consensus regarding the CDSS recommendation in each case study (Table 1).

Table . Assessments of the experts surveyed on the feasibility of the pilot version of the artificial intelligence–based clinical decision support system (N=19).

	CDSS recommendation for initial treatment		CDSS recommendation for re-evaluation	
	Agreeing with the recommendation, n (%)	Rejecting the recommendation, n (%)	Agreeing with the recommendation, n (%)	Rejecting the recommendation, n (%)
Use case 1	14 (78)	4 (22)	12 (67)	6 (33)
Use case 2	11 (61)	7 (39)	1 (5)	18 (95)
Use case 3	3 (16)	16 (84)	2 (11)	16 (89)
Use case 4	15 (83)	3 (17)	14 (78)	4 (22)
Use case 5	13 (76)	4 (24)	12 (71)	5 (29)

A suggestion was made that the pilot CDSS should be expanded to include the integration of combination therapies. The vital parameters and patient information should be included as variables in the statistical model for predicting therapy to include the suspected focus of infection. In addition, it was noted that the data provided for the development of the model and the case studies in the pilot version were incomplete in many instances (Table 1). The physicians encountered significant challenges in evaluating the CDSS recommendation when the fields for data that could be viewed in the application were left empty (quote 4).

> ***Quote 4:****Here again the PCT [Procalcitonin] is missing, that would make it easier for us. We only have [...] indirect information that we have an infection here [...] measured by the thrombocytopenia, and then [...] I can get along with actually using piperacillin and tazobactam primarily here when it comes to the calculated antibiotic therapy [...].* [Expert 5, Clinical Center 3]

> ***Quote 5:****The [application] is more explanatory than my phone. So, everything is fine in that respect.* [Expert 3, Clinical Center 3]

From the physicians' perspective, the primary factor influencing the acceptance of the AI-based CDSS was the time required for completion of the process. If the system is able to produce an evident medication recommendation in a shorter time than the physicians can achieve, a high willingness to use it was reported. There was unanimous feedback that the design and layout of the CDSS are pleasant and easy to understand (quote 5). Furthermore, the majority of respondents indicated a preference for a mobile version of the system for use in everyday clinical practice. This preference was again justified by the availability of digital resources.

Facilitating and Inhibiting Factors for the Implementation of an AI-Based Clinical Decision Support Systems

The focus group interviews showed that while there is already strong foundational support for digital tools and applications in diagnostics and therapy as part of routine clinical practice, the use of more advanced applications, especially those based on AI, was limited to isolated cases and primarily on a project basis. Previous experiences with CDSSs had a significant impact on expectations regarding the quality of the CDSS, the effort required for integration into everyday medical practice, and the intention to potentially use such a system (quotes 6 and 7).

> ***Quote 6:****I'm thinking about the echocardiography, [...] where the AI has become so good that we let it*

XSL•FO
RenderX

take over. We do the examination ourselves. And yet the module helps us with that. That's how I imagine it [here] too. [Expert 2, focus group 2]

Quote 7:*Yes, for the colonoscopy [...] we once had [a CDSS]. But the human was faster than the AI. So in that respect [...] rather not.* [Expert 3, focus group 3]

Nevertheless, the majority of respondents indicated a high level of willingness to adopt AI-based CDSSs (quotes 8 and 9). The option of re-evaluation included in the application was considered to be particularly beneficial.

Quote 8:*I actually found the idea compelling. Because in the early days of my clinical work, even more so than now, the question always arises: am I doing it right now?* [Expert 1, focus group 2]

Quote 9:*[T]he field is so complex [...] that's why I think a support tool that works well is very helpful.* [Expert 3, focus group 1]

As previously noted, the acceptance and implementation of digital solutions is largely dependent on the amount of time available. In addition, the majority of respondents indicated that they are consistently seeking tools that facilitate efficient support in routine clinical practice. Consequently, the respondents exhibited an intrinsic motivation to use CDSSs, which was observed across all age groups and genders, but was particularly pronounced in the "30 - 40 years" and "40 - 50 years" age groups.

In contrast with the physician's technology openness as a beneficial factor, the level of digitization in the clinics was found to be heterogeneous. Furthermore, the physicians' assessment indicated that there was significant potential for improvement in all 3 hospitals (quote 10). The statements indicated that the level of digital infrastructure, including stable internet coverage, exhibited considerable variation not only between the clinics but also between individual wards within the clinics. Overall, there was a consensus that better equipment in terms of digital devices and applications was needed (quote 11). The lack of comprehensive digital and cross-sectoral availability of relevant patient health information was also identified as a barrier to the implementation of AI-based CDSSs and efficient medical treatment in general. This demonstrated that the degree of digitization in everyday clinical practice affects the perceived feasibility of implementing the CDSS and is the most significant barrier to the beneficial factors.

Quote 10:*So, it's really funny that you can still secure good pens. Because we do so much paper based.* [Expert 2, focus group 2]

Quote 11:*The problem is that there are far too many individual components that all [...] communicate with each other via interfaces. And in the process, values are lost or the data transfer doesn't work. And then you end up standing there and yes, the computer hangs.* [Expert 4, focus group 2]

Furthermore, the existing data basis for the AI-based CDSS model was identified as a significant challenge. The participants were aware that, in addition to the aforementioned lack of digitized patient information, the heterogeneous quality of previous antibiotic administration practice also had an influence on the predictive ability of the recommendation from the CDSS. Furthermore, an increased documentation and justification effort in the event of potential deviation from the CDSS recommendation was also seen as a barrier. However, all participants were aware that such a system would be implemented as a supportive and not a replacement measure, and this was not criticized.

Two relevant recommendations for improvement were derived from the interviews. First, the majority of respondents indicated that applicable guidelines on antibiotic therapy should be incorporated into the specification of the underlying statistical models. Second, there was a recurring suggestion that, alongside the further development of the presented system, the capability to diagnose sepsis infections should also be integrated into a CDSS. Once again, the time to correct diagnosis and subsequent appropriate treatment was identified as the main driver. Suggestions for adapting the tested CDSS comprised the inclusion of quantities and suggestions for the duration of antibiotic administration, in addition to the therapeutic agents themselves (quote 12).

Quote 12:*So, in an ideal world, [...] [AI] can be a huge step forward as a decision-making tool. But [A] the right information has to flow in. And [B] the decisive recommendation [...] is already anticipated here [...]. So, the first step: can this be sepsis? That is an important step. [...] In the second step, a recommendation for initial therapy is helpful [...], yes.* [Expert 4, focus group 1]

Discussion

Principal Findings

Overall, this use case analysis shows various barriers and facilitators for the implementation of AI-based CDSSs for antibiotic therapy in sepsis (Textbox 2). The findings of the expert and focus group interviews indicate that the potential of digitization and, in particular, AI-based tools is expected and viewed in a favorable light by the majority of the physicians. This finding is in line with other research which indicates that clinicians have a predominantly positive perception of AI systems [38-40]. Similarly, the willingness to use these tools is high [41]. In this context, the plausibility of the prediction and the potential time savings are of particular importance. It is evident that the existing contextual conditions represent the most significant obstacle to future implementation. In addition to the basic technical equipment in the clinics, other factors, such as a stable, comprehensive internet supply and the necessary digital availability of data, were also identified as potential barriers to future implementation. Accordingly, the question arises as to what extent clinics are equipped with the digital infrastructure to fully leverage the benefits of AI-based CDSSs. At this point it has to be noted that the digital maturity of German hospitals appears to be comparatively limited [42]. This can be attributed, at least in part, to a lack of financial resources [24,43]. It remains to be seen to what extent hospitals use the transformation fund provided by the German Hospital

XSL•FO
RenderX

Future Act and how this can enhance the conditions for the implementation of AI systems.

Textbox 2. Summary of barriers, facilitators, and recommendations for artificial intelligence–based clinical decision support systems' feasibility and implementation.

Barriers

- Overall
- Low degree of digitization in everyday clinical practice.
- Lack of comprehensive digital and cross-sectoral data availability.
- Negative previous experiences with clinical decision support systems.
- Use Cases
- Heterogeneous quality of previous antibiotic administration practice limits predictive ability of the artificial intelligence–based recommendations.
- Increased effort to document and justify potential deviations from the clinical decision support system's recommendation.
- The confidence to approve or reject artificial intelligence–based antibiotic recommendations depends on the level of professional experience.

Facilitators

- Overall
- Potential time savings.
- Physician's technology openness or intrinsic motivation.
- Positive previous experiences with clinical decision support systems.
- Use Cases
- Re-evaluation of clinical decision support system recommendations is considered to be particularly beneficial.
- Pleasant design and layout of the clinical decision support system (easy to understand).

Recommendations

- Incorporation of guidelines on antibiotic therapy into the specification of the statistical models.
- Clinical decision support systems should include the integration of combination therapies.
- Statistical model for predicting antibiotic recommendations should include vital parameters and further patient information to identify the suspected focus of infection.
- Besides sepsis treatment, the capability to diagnose sepsis infections should also be integrated into a clinical decision support system.
- Preference for a mobile version of the clinical decision support system.

The medication recommendations by the CDSS were not yet perceived as reliable due to the limited and biased data available for training the AI system. Often, a very broad standard therapy (piperacillin/tazobactam) was recommended. This was mainly due to the fact that this medication occurred considerably more frequently within the learning data set than all other antibiotics. According to the experts, this approach is not absolutely mandatory and sometimes maybe counterproductive in terms of calculated antibiotic therapy. Younger physicians in particular tended to agree with the recommendations. This is a sign that a CDSS for the most common "standard medication" might not be beneficial if not detrimental. This can be confirmed by the physicians' statements that they would like decision support primarily for medication combinations and for rare antibiotic agents or rare sources of infection. In view of the goal of avoiding the development of resistance by reducing the use of broad-spectrum antibiotics, the current pilot version of the AI system would possibly even promote this risk of further resistance. To avoid this, current guidelines should be incorporated into the model design in addition to retrospective primary data.

The respondents' call for the upstream integration of sepsis diagnostics in addition to the further development of the existing CDSS appears to be purposeful. A recent study has confirmed initial successes in this regard [44]. Although no clear correlations between sociodemographic characteristics and the willingness to use CDSSs were found in the present evaluation, differences in the approval or rejection of the AI-based CDSS recommendations depending on professional experience were identified. This finding is supported by Lambert et al [45]. It is recommended that future studies investigate the degree to which professional experience and other sociodemographic variables exert an influence on technology acceptance and the willingness to use CDSSs. In this regard, using the UTAUT would be appreciated [35].

XSL•FO
RenderX

Generally, the following challenges associated with the use of AI in health care should not be overlooked. For example, automation bias, in which health care providers may develop an overreliance on AI recommendations, potentially reducing their clinical vigilance and leading to oversights in patient care, needs to be considered [46]. Furthermore, AI-driven CDSSs may unintentionally propagate biases present in training datasets, resulting in discriminatory recommendations that may disadvantage certain demographic groups [47]. Another key concern is the ambiguity and complexity of liability; it remains unclear who is legally responsible if an AI-influenced decision contributes to a negative patient outcome [48]. Furthermore, previous research has highlighted the potential negative impact of a lack of user involvement in the development of CDSSs on the willingness to use them, even when the system quality is optimal [26]. One of the key strengths of the study is the early involvement of future users. The participatory approach highlights the need for involvement. Consequently, given the significant lack of engagement with CDSS interventions reported within the literature [49], future studies should consider the involvement of users from the earliest stages of development [26,28,50]. The findings of the present study confirm that in addition to personal factors, the plausibility of the recommendations is the most decisive factor determining the willingness to use the tool [51]. Further development of the pilot version or similar systems is therefore crucial for implementing and exploiting the potential benefits of AI. Furthermore, the feedback from the physicians participating in this study provided valuable insights into additional variables that were subsequently integrated into the further development of the CDSS, following the completion of the interviews.

Limitations

Several limitations must be acknowledged in the evaluation of the use case presented in this study. First, the efficacy of the CDSS was clearly limited due to the rather small amount of data used for training. The findings suggest that the CDSS may offer less benefit for physicians during the initial therapy phase compared with its use in the context of therapy re-evaluation. At this stage, it was not feasible to accurately predict the most appropriate antibiotic, one that would minimize treatment changes and side effects, based on the currently available data. The acceptance of AI-based interventions heavily depends on the plausibility and relevance of therapy recommendations, which ultimately impacts their efficacy. Therefore, further development and refinement of AI-based CDSSs are essential, along with an expansion of the digital data infrastructure.

In addition, it is important to consider that the evaluation was based on a qualitative interview study, which comes with inherent limitations. The voluntary nature of participation may introduce self-selection bias, potentially leading to an overestimation of the participants' openness to technology and their acceptance of AI-based CDSSs. This limitation suggests that the findings may not be fully generalizable to the broader population of health care professionals.

Conclusion

This pilot study of an AI-based CDSS indicates that further development is needed to achieve the original goals of minimizing switches in antibiotic prescriptions and reducing the reliance on broad-spectrum antibiotics. Relying solely on retrospective data from past care practices does not seem to be an effective strategy for meeting these objectives. In addition, it is essential to address the key challenges identified in this study to enable the successful integration of AI into clinical settings. While the respondents generally showed a positive attitude toward the use of AI-based CDSSs, a more comprehensive evaluation of technology acceptance among health care professionals can only be conducted when a version with realistic, guideline-compliant recommendations is available. Early engagement of users in the development process proves beneficial, not only for refining the CDSS but also for fostering acceptance among medical practitioners.

Acknowledgments

The authors would like to thank all members of the KINBIOTICS project and in particular the contact persons at the 3 hospitals (Evangelisches Klinikum Bethel, Klinikum Bielefeld Mitte, and Klinikum Lippe) for their support in conducting the interviews. Special thanks go to all the physicians who took part in the interviews and to Inga Heißenberg, who helped to prepare and conduct the interviews. Furthermore, we would like to thank Dr Timothy Mc Call for his support in translating the original quotations. The KINBIOTICS project was funded by the German Federal Ministry of Health (grant ZMVI1-2520DAT930). The funding organization was not involved in the data collection, evaluation, analysis, or writing of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Standards for Reporting Qualitative Research checklist.
[PDF File, 121 KB - humanfactors_v12i1e66699_app1.pdf]

References

1. Kumar NR, Balraj TA, Kempegowda SN, Prashant A. Multidrug-resistant sepsis: a critical healthcare challenge. Antibiotics (Basel) 2024 Jan 4;13(1):46. [doi: 10.3390/antibiotics13010046] [Medline: 38247605]

XSL•FO
RenderX

2. GBD 2021 Antimicrobial Resistance Collaborators. Global burden of bacterial antimicrobial resistance 1990-2021: a systematic analysis with forecasts to 2050. Lancet 2024 Sep 28;404(10459):1199-1226. [doi: 10.1016/S0140-6736(24)01867-1] [Medline: 39299261]
3. Rudd KE, Johnson SC, Agesa KM, et al. Global, regional, and national sepsis incidence and mortality, 1990-2017: analysis for the Global Burden of Disease Study. Lancet 2020 Jan 18;395(10219):200-211. [doi: 10.1016/S0140-6736(19)32989-7] [Medline: 31954465]
4. Fleischmann-Struzek C, Mikolajetz A, Schwarzkopf D, et al. Challenges in assessing the burden of sepsis and understanding the inequalities of sepsis outcomes between National Health Systems: secular trends in sepsis and infection incidence and mortality in Germany. Intensive Care Med 2018 Nov;44(11):1826-1835. [doi: 10.1007/s00134-018-5377-4] [Medline: 30284637]
5. Bauer M, Groesdonk HV, Preissing F, Dickmann P, Vogelmann T, Gerlach H. Mortality in sepsis and septic shock in Germany. Results of a systematic review and meta-analysis. Anaesthesist 2021 Aug;70(8):673-680. [doi: 10.1007/s00101-021-00917-8] [Medline: 33559687]
6. The burden of antimicrobial resistance in G7 countries and globally: an urgent call for action. Robert Koch-Institut. 2022. URL: https://edoc.rki.de/handle/176904/9864;jsessionid=BB6B616866626A2E3B56758E5A1C5C8B?locale-attribute=en [accessed 2025-02-15]
7. Rödenbeck M, Ayobami O, Eckmanns T, Pletz MW, Bleidorn J, Markwart R. Clinical epidemiology and case fatality due to antimicrobial resistance in Germany: a systematic review and meta-analysis, 1 January 2010 to 31 December 2021. Euro Surveill 2023 May;28(20):2200672. [doi: 10.2807/1560-7917.ES.2023.28.20.2200672] [Medline: 37199987]
8. Zilberberg MD, Shorr AF, Micek ST, Vazquez-Guillamet C, Kollef MH. Multi-drug resistance, inappropriate initial antibiotic therapy and mortality in Gram-negative severe sepsis and septic shock: a retrospective cohort study. Crit Care 2014 Nov 21;18(6):596. [doi: 10.1186/s13054-014-0596-8] [Medline: 25412897]
9. Pradipta IS, Sodik DC, Lestari K, et al. Antibiotic resistance in sepsis patients: evaluation and recommendation of antibiotic use. N Am J Med Sci 2013 Jun;5(6):344-352. [doi: 10.4103/1947-2714.114165] [Medline: 23923107]
10. Kalın G, Alp E, Chouaikhi A, Roger C. Antimicrobial multidrug resistance: clinical implications for infection management in critically ill patients. Microorganisms 2023 Oct 16;11(10):2575. [doi: 10.3390/microorganisms11102575] [Medline: 37894233]
11. Neugebauer M, Ebert M, Vogelmann R. A clinical decision support system improves antibiotic therapy for upper urinary tract infection in a randomized single-blinded study. BMC Health Serv Res 2020 Mar 6;20(1):185. [doi: 10.1186/s12913-020-5045-6] [Medline: 32143630]
12. Van Dort BA, Penm J, Ritchie A, Baysari MT. The impact of digital interventions on antimicrobial stewardship in hospitals: a qualitative synthesis of systematic reviews. J Antimicrob Chemother 2022 Jun 29;77(7):1828-1837. [doi: 10.1093/jac/dkac112] [Medline: 35388414]
13. Xu R, Wu L, Wu L, Xu C, Mu T. Effectiveness of decision support tools on reducing antibiotic use for respiratory tract infections: a systematic review and meta-analysis. Front Pharmacol 2023;14(1253520):1253520. [doi: 10.3389/fphar.2023.1253520] [Medline: 37745052]
14. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. NPJ Digit Med 2020;3(17):17. [doi: 10.1038/s41746-020-0221-y] [Medline: 32047862]
15. Graafsma J, Murphy RM, van de Garde EMW, et al. The use of artificial intelligence to optimize medication alerts generated by clinical decision support systems: a scoping review. J Am Med Inform Assoc 2024 May 20;31(6):1411-1422. [doi: 10.1093/jamia/ocae076] [Medline: 38641410]
16. Rittmann B, Stevens MP. Clinical decision support systems and their role in antibiotic stewardship: a systematic review. Curr Infect Dis Rep 2019 Jul 24;21(8):29. [doi: 10.1007/s11908-019-0683-8] [Medline: 31342180]
17. Rutten JJS, van Buul LW, Smalbrugge M, et al. An electronic health record integrated decision tool and supportive interventions to improve antibiotic prescribing for urinary tract infections in nursing homes: a cluster randomized controlled trial. J Am Med Dir Assoc 2022 Mar;23(3):387-393. [doi: 10.1016/j.jamda.2021.11.010] [Medline: 34896069]
18. Hartman EAR, Groen WG, Heltveit-Olsen SR, et al. Implementation of a tailored multifaceted antibiotic stewardship intervention to improve antibiotic prescribing for urinary tract infections in frail older adults (ImpresU) in four European countries: a process evaluation alongside a pragmatic cluster randomized controlled trial. Trials 2024 Oct 18;25(1):691. [doi: 10.1186/s13063-024-08545-4] [Medline: 39425170]
19. Laka M, Milazzo A, Merlin T. Can evidence-based decision support tools transform antibiotic management? A systematic review and meta-analyses. J Antimicrob Chemother 2020 May 1;75(5):1099-1111. [doi: 10.1093/jac/dkz543] [Medline: 31960021]
20. Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. BMJ 2005 Apr 2;330(7494):765. [doi: 10.1136/bmj.38398.500764.8F] [Medline: 15767266]

XSL•FO
RenderX

21. Catho G, Sauser J, Coray V, et al. Impact of interactive computerised decision support for hospital antibiotic use (COMPASS): an open-label, cluster-randomised trial in three Swiss hospitals. Lancet Infect Dis 2022 Oct;22(10):1493-1502. [doi: 10.1016/S1473-3099(22)00308-5] [Medline: 35870478]
22. Hong N, Liu C, Gao J, et al. State of the art of machine learning-enabled clinical decision support in intensive care units: literature review. JMIR Med Inform 2022 Mar 3;10(3):e28781. [doi: 10.2196/28781] [Medline: 35238790]
23. Komorowski M, Celi LA, Badawi O, Gordon AC, Faisal AA. The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care. Nat Med 2018 Nov;24(11):1716-1720. [doi: 10.1038/s41591-018-0213-5] [Medline: 30349085]
24. Weinert L, Müller J, Svensson L, Heinze O. Perspective of information technology decision makers on factors influencing adoption and implementation of artificial intelligence technologies in 40 German hospitals: descriptive analysis. JMIR Med Inform 2022 Jun 15;10(6):e34678. [doi: 10.2196/34678] [Medline: 35704378]
25. KINBIOTICS [Website in German]. URL: https://kinbiotics.de [accessed 2024-11-07]
26. Khairat S, Marc D, Crosby W, Al Sanousi A. Reasons for physicians not adopting clinical decision support systems: critical analysis. JMIR Med Inform 2018 Apr 18;6(2):e24. [doi: 10.2196/medinform.8912] [Medline: 29669706]
27. Ji M, Yu G, Xi H, Xu T, Qin Y. Measures of success of computerized clinical decision support systems: an overview of systematic reviews. Health Policy Technol 2021 Mar;10(1):196-208. [doi: 10.1016/j.hlpt.2020.11.001]
28. Carroll C, Marsden P, Soden P, Naylor E, New J, Dornan T. Involving users in the design and usability evaluation of a clinical decision support system. Comput Methods Programs Biomed 2002 Aug;69(2):123-135. [doi: 10.1016/s0169-2607(02)00036-6] [Medline: 12100792]
29. Newton N, Bamgboje-Ayodele A, Forsyth R, Tariq A, Baysari MT. Does involving clinicians in decision support development facilitate system use over time? a systematic review. Stud Health Technol Inform 2023 Jun 22;304:11-15. [doi: 10.3233/SHTI230359] [Medline: 37347561]
30. Westerbeek L, Ploegmakers KJ, de Bruijn GJ, et al. Barriers and facilitators influencing medication-related CDSS acceptance according to clinicians: a systematic review. Int J Med Inform 2021 Aug;152:104506. [doi: 10.1016/j.ijmedinf.2021.104506] [Medline: 34091146]
31. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251. [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
32. Döring N. Datenanalyse. In: Döring N, editor. Forschungsmethoden Und Evaluation in Den Sozial- Und Humanwissenschaften [Book in German]: Berlin, Heidelberg: Springer Berlin Heidelberg; 2023:587-766. [doi: 10.1007/978-3-662-64762-2_12]
33. Venkatesh V, Thong JYL, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Q 2012;36(1):157. [doi: 10.2307/41410412]
34. Venkatesh V, Thong J, Xu X, University of Arkansas, Hong Kong University of Science and Technology, The Hong Kong Polytechnic University. Unified theory of acceptance and use of technology: a synthesis and the road ahead. JAIS 2016;17(5):328-376. [doi: 10.17705/1jais.00428]
35. Dingel J, Kleine AK, Cecil J, Sigl AL, Lermer E, Gaube S. Predictors of health care practitioners' intention to use AI-enabled clinical decision support systems: meta-analysis based on the unified theory of acceptance and use of technology. J Med Internet Res 2024 Aug 5;26:e57224. [doi: 10.2196/57224] [Medline: 39102675]
36. Kuckartz U. Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung 4. Auflage. Weinheim [Book in German]: Basel: Beltz Juventa; 2018.
37. MAXQDA. URL: https://www.maxqda.com/ [accessed 2025-02-20]
38. van der Meijden SL, de Hond AAH, Thoral PJ, et al. Intensive care unit physicians' perspectives on artificial intelligence-based clinical decision support tools: preimplementation survey study. JMIR Hum Factors 2023 Jan 5;10:e39114. [doi: 10.2196/39114] [Medline: 36602843]
39. Maassen O, Fritsch S, Palm J, et al. Future medical artificial intelligence application requirements and expectations of physicians in German university hospitals: web-based survey. J Med Internet Res 2021 Mar 5;23(3):e26646. [doi: 10.2196/26646] [Medline: 33666563]
40. Pecqueux M, Riediger C, Distler M, et al. The use and future perspective of artificial Intelligence-a survey among German surgeons. Front Public Health 2022;10:982335. [doi: 10.3389/fpubh.2022.982335] [Medline: 36276381]
41. Hennrich J, Kauffmann AL, Buck C, Eymann T. Künstliche intelligenz in der hausärztlichen versorgung. In: Pfannstiel MA, editor. Künstliche Intelligenz Im Gesundheitswesen: Entwicklungen, Beispiele Und Perspektiven [Book in German] 2022:413-430. [doi: 10.1007/978-3-658-33597-7_18]
42. Geissler A, Hollenbach J, Haring M, Amelung VE, Thun S, Haering A. A nationwide digital maturity assessment of hospitals – results from the German DigitalRadar. Health Policy Technol 2024 Sep;13(4):100904. [doi: 10.1016/j.hlpt.2024.100904]
43. Burmann A, Fischer B, Brinkkötter N, Meister S. Managing directors' perspectives on digital maturity in German hospitals-a multi-point online-based survey study. Int J Environ Res Public Health 2022 Aug 6;19(15):9709. [doi: 10.3390/ijerph19159709] [Medline: 35955066]
44. Steinbach D, Ahrens PC, Schmidt M, et al. Applying machine learning to blood count data predicts sepsis with ICU admission. Clin Chem 2024 Mar 2;70(3):506-515. [doi: 10.1093/clinchem/hvae001] [Medline: 38431275]

XSL•FO
RenderX

45. Lambert SI, Madi M, Sopka S, et al. An integrative review on the acceptance of artificial intelligence among healthcare professionals in hospitals. NPJ Digit Med 2023 Jun 10;6(1):111. [doi: 10.1038/s41746-023-00852-5] [Medline: 37301946]
46. Goddard K, Roudsari A, Wyatt JC. Automation bias: a systematic review of frequency, effect mediators, and mitigators. J Am Med Inform Assoc 2012;19(1):121-127. [doi: 10.1136/amiajnl-2011-000089] [Medline: 21685142]
47. Gianfrancesco MA, Tamang S, Yazdany J, Schmajuk G. Potential biases in machine learning algorithms using electronic health record data. JAMA Intern Med 2018 Nov 1;178(11):1544-1547. [doi: 10.1001/jamainternmed.2018.3763] [Medline: 30128552]
48. Price WN 2nd, Gerke S, Cohen IG. Potential liability for physicians using artificial intelligence. JAMA 2019 Nov 12;322(18):1765-1766. [doi: 10.1001/jama.2019.15064] [Medline: 31584609]
49. Rawson TM, Moore LSP, Hernandez B, et al. A systematic review of clinical decision support systems for antimicrobial management: are we failing to investigate these interventions appropriately? Clin Microbiol Infect 2017 Aug;23(8):524-532. [doi: 10.1016/j.cmi.2017.02.028] [Medline: 28268133]
50. Neff MC, Schaaf J, Noll R, et al. Initial user-centred design of an AI-based clinical decision support system for primary care. Stud Health Technol Inform 2024 Jan 25;310:1051-1055. [doi: 10.3233/SHTI231125] [Medline: 38269975]
51. Panicker RO, George AE. Adoption of automated clinical decision support system: a recent literature review and a case study. Archives of Medicine and Health Sciences 2023;11(1):86-95. [doi: 10.4103/amhs.amhs_257_22]

Abbreviations

AI: artificial intelligence
AMR: antimicrobial resistance
CDSS: clinical decision support system
ICU: intensive care unit
SRQR: Standards for Reporting Qualitative Research
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by A Kushniruk; submitted 02.10.24; peer-reviewed by D Sharma, GD Giebel, T Bommhardt; revised version received 22.11.24; accepted 31.12.24; published 04.03.25.

Please cite as:
Düvel JA, Lampe D, Kirchner M, Elkenkamp S, Cimiano P, Düsing C, Marchi H, Schmiegel S, Fuchs C, Claßen S, Meier KL, Borgstedt R, Rehberg S, Greiner W
An AI-Based Clinical Decision Support System for Antibiotic Therapy in Sepsis (KINBIOTICS): Use Case Analysis
JMIR Hum Factors 2025;12:e66699
URL: https://humanfactors.jmir.org/2025/1/e66699
doi:10.2196/66699

© Juliane Andrea Düvel, David Lampe, Maren Kirchner, Svenja Elkenkamp, Philipp Cimiano, Christoph Düsing, Hannah Marchi, Sophie Schmiegel, Christiane Fuchs, Simon Claßen, Kirsten-Laura Meier, Rainer Borgstedt, Sebastian Rehberg, Wolfgang Greiner. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 4.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Impact of Human-Robot Collaboration Levels on Postural Stability During Working Tasks Performed While Standing: Experimental Study

Daniele Bibbo, PhD; Giovanni Corvini, PhD; Maurizio Schmid, PhD; Simone Ranaldi, PhD; Silvia Conforto, PhD

Department of Industrial, Electronic and Mechanical Engineering, Roma Tre University, Via Vito Volterra, 62 - Corpo B, Roma, Italy

Corresponding Author:
Daniele Bibbo, PhD
Department of Industrial, Electronic and Mechanical Engineering, Roma Tre University, Via Vito Volterra, 62 - Corpo B, Roma, Italy

Abstract

Background: The integration of collaborative robots (cobots) in industrial settings has the potential to enhance worker safety and efficiency by improving postural control and reducing biomechanical risk. Understanding the specific impacts of varying levels of human-robot collaboration on these factors is crucial for optimizing cobot use.

Objective: This study aims to investigate the biomechanical effects of different levels of human-robot collaboration on postural stability and control during simulated working tasks.

Methods: A total of 14 participants performed simulated cashier working activities under 4 different collaboration modalities, with increasing levels of cobot assistance: full (Fu), half robot touch (HRT), half robot (HRb), and full robot (FRb). Center of pressure trajectories were extracted from 2 force plates' data to calculate 4 posturography parameters—mean distance (MDIST), mean velocity (MVELO), 95% confidence ellipse area (AREA-CE), and sway area (AREA-SW)—which were analyzed to assess the impact of cobot intervention on postural control.

Results: Nonparametric tests showed significance in the effect of the collaboration modalities on the 4 analyzed parameters. Post hoc tests revealed that FRb modality led to the greatest enhancement in postural stability, with a reduction in MDIST (4.2, SD 1.3 cm in Fu vs 1.6, SD 0.5 cm in FRb) and MVELO (16.3, SD 5.2 cm/s in Fu vs 7.9, SD 1.1 cm/s in FRb). AREA-CE and AREA-SW also decreased significantly with higher levels of cobot assistance (AREA-CE: 134, SD 91 cm^2 in Fu vs 22, SD 12 cm^2 in FRb; AREA-SW: 16.2, SD 8.4 cm^2/s in Fu vs 4.0, SD 1.6 cm^2/s in FRb). Complete assistance of the cobot significantly reduced interindividual variability of all center of pressure parameters. In FRb modality, as compared with all other conditions, removing the weight of the object during loading or unloading phases caused a significant decrease in all parameter values.

Conclusions: Increased cobot assistance significantly enhances postural stability and reduces biomechanical load on workers during simulated tasks. Full assistance from cobots, in particular, minimizes postural displacements, indicating more consistent postural control improvements across individuals. However, high levels of cobot intervention also reduced the natural variation in how people balanced themselves. This could potentially lead to discomfort in the long run. Midlevel cobot assistance modalities can thus be considered as a good compromise in reducing biomechanical risks associated with postural stability at the same time granting a satisfactory level of user control.

(JMIR Hum Factors 2025;12:e64892) doi:10.2196/64892

KEYWORDS

human-robot collaboration; cobot assistance; postural control; biomechanical risk; ergonomics; collaborative robot

Introduction

Industrial manufacturing is transitioning from well-established production procedures toward more flexible and intelligent manufacturing systems (Industry 4.0, as initially described by Kagermann et al [1]). This evolution aims to develop innovative, sustainable solutions that create new business models, improve working conditions, increase plant productivity, and enhance product quality [2]. Robotics plays a crucial role in this context, with industrial robots being widely adopted due to their ability to relieve humans from repetitive, unhealthy, or dangerous tasks [3]. The field of robotics has evolved to not only allow humans to share the same workspace with robots but also use them as assistants, thereby enhancing human-robot collaboration (HRC) [4].

Collaborative robots (cobots) are at the forefront of this revolution, offering increased productivity, flexibility, versatility, and safety compared with traditional industrial robots. Cobots are designed to work alongside humans, sharing the same workspace without the need for safety cages and

XSL•FO
RenderX

providing greater mobility and flexibility [5]. The integration of cobots in industrial environments has been shown to significantly enhance operational efficiency [6] and worker satisfaction by enabling safer and more ergonomic working conditions [7,8]. However, the integration of cobots impacts various human factors, which include psychological aspects [9], emotions [10], and biomechanical effects [11]. Consequently, it is important to consider both cognitive ergonomics and biomechanical safety to optimize the overall effectiveness and well-being of human workers [12]. For instance, Gualtieri et al [13] have investigated the cognitive elements involved in human interaction with cobots, revealing significant insights into how mental workload and task complexity can affect worker performance and satisfaction [14]. Similarly, improper collaboration with cobots can expose workers to biomechanical risks, even though these robots are intended to alleviate heavy and tedious tasks. Previous research studies [15,16] emphasized the importance of ergonomic considerations in HRC to prevent musculoskeletal disorders (MSDs) among workers.

Thus, the right choice of cobot collaboration modality can reduce biomechanical overload on workers, ensuring their safety.

Recent advancements in sensor technology and data analytics have enabled detailed assessments of human biomechanics during HRC [17,18]. Studies have shown that real-time monitoring and analysis of physiological data can provide valuable insights into the physical strain experienced by workers and help in optimizing collaborative processes [19]. For example, the use of wearable sensors and motion capture systems could effectively measure and analyze the biomechanical load on workers during different collaborative tasks with robots [20]. Other recent studies investigated the effects of different HRC modalities on physiological human parameters such as trunk oscillations [21] and muscle coactivation [22,23]. Such studies can also be profitably taken into consideration for the optimization of workplace design and task allocation.

When work tasks are performed in quasi-static conditions, direct measures of posture could reveal insights into the stability and balance of workers and can thus be used to improve ergonomics during HRC, as they can help identify postural adjustments and biomechanical loads, providing a more comprehensive understanding of HRC ergonomic impact [24].

Following this principle, this study will evaluate and analyze posturography data to assess the biomechanical risks in HRC working contexts where cobots are used in different modalities. Through this evaluation, the study seeks to determine whether cobot assistance can effectively reduce biomechanical risks: by examining how different collaborative modalities influence workers' stability and balance, thus providing insights into the effects of cobot assistance on postural control, as this latter one indirectly reflects safety, efficiency, and ergonomic well-being of human workers in collaborative environments.

Methods

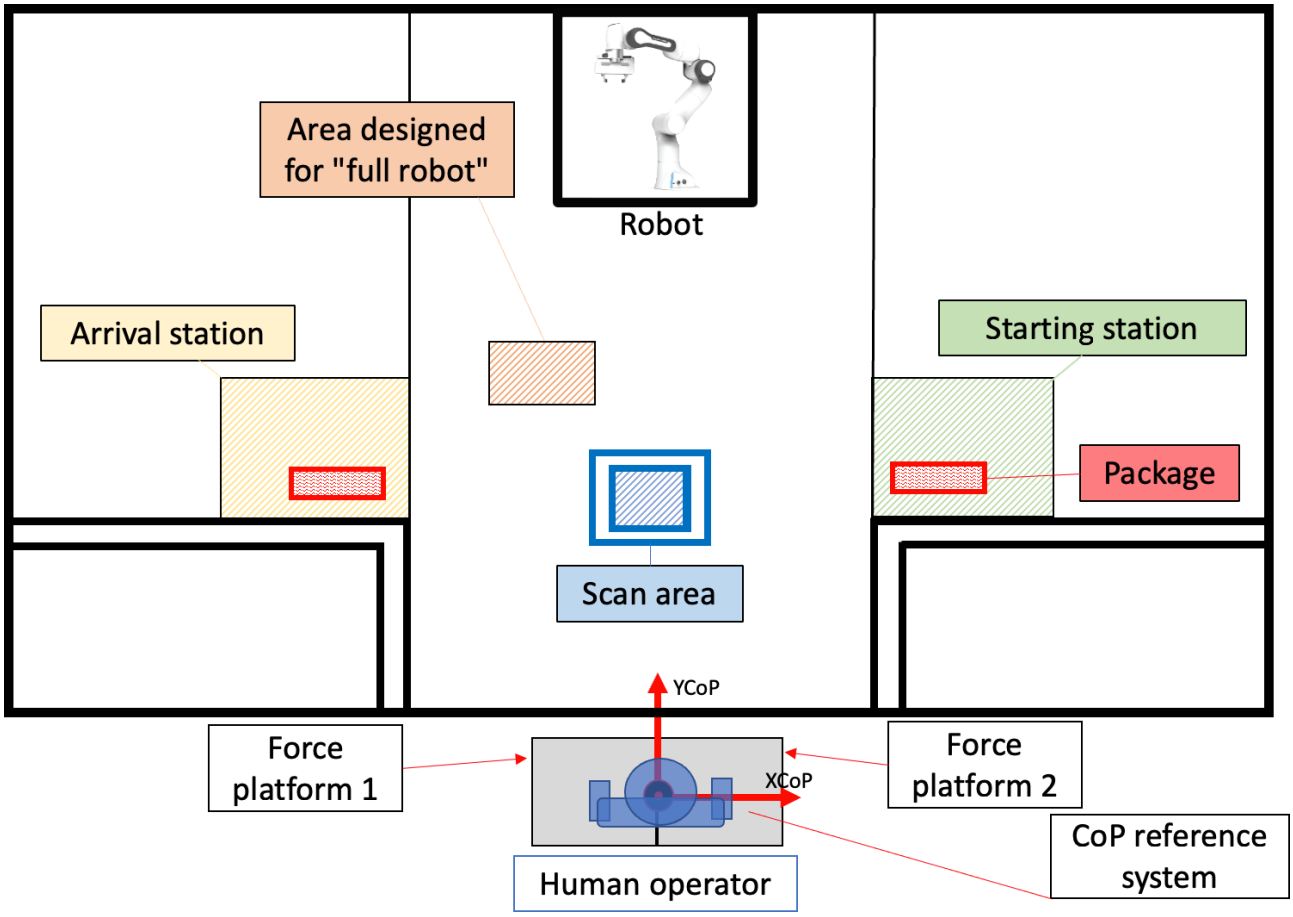
Experimental Setup

Overview

The setup was designed to compare a simulated work activity performed with and without cobot assistance. A specific workbench was created by positioning the cobot on one of the longer sides of a 2 × 1 m table, with the operator standing upright on the opposite side (Figure 1). The operator's standing position was fixed relative to the edge of the table, defined by the position of his or her feet and the projection onto the ground of the long side of the table. A starting station (ie, the initial position to the right of the operator) and an arrival station (ie, the final position to the left) were defined, each 1m apart from the other and 30 cm away from the operator in the direction of the cobot. A scanning area was defined in the center of the table in front of the operator. This workbench arrangement allowed the operator to perform all the activities (or part of them) independently, while the cobot could collaborate on some or none of the tasks without interfering with the operator's movements. This configuration was defined to minimize possible hindering of the operator's movements.

XSL•FO
RenderX

Figure 1. Experimental setup dimensioning for the simulation of load-moving activities on a workbench. CoP: center of pressure.



The activities involved the handling and relocating of rectangular packages, each measuring 14 × 12 × 5 cm and weighing 1 kg, bearing a QR code on one side.

The tests were conducted with a group of 14 male volunteers aged between 25 and 48 years (weight: 79, SD 7 kg; height: 180, SD 5 cm). All volunteers were healthy and had no evident or declared mobility issues. Moreover, given the design of the setup, only right-handed subjects have been involved in the tests. Participants were instructed to (1) grasp the packages from the starting station, (2) identify the QR code and scan it (ie, facing the QR code to the scan area) in the central area, and then (3) place the package in the arrival station. Different levels of collaboration were implemented by removing the first part of these activities (ie, grasping the package and bringing it to the center of the table) when in partial assistance, and removing also the last part (ie, bringing the package to the arrival station) when in full cobot assistance. Each participant was asked to perform the task as naturally as possible. This task was designed to simulate the work activity of a cashier in a supermarket, a highly relevant example in terms of strenuous activity [25], where an increase in the level of cobot assistance is hypothesized to reduce the associated biomechanical risk.

In the test, each participant performed consecutive repetitions of the same task, for a 5-minute duration, with different levels of cobot assistance.

The frequency of processing (packages per minute [PPM]) to be handled was not defined a priori, and participants were allowed to choose their own pace in performing all the activities. When collaborating with the cobot, the PPM frequency was however influenced by the cobot's package handling activity, as described below.

Each participant was asked to perform the described activity under 4 different collaboration modalities, performed in random order, with a 5-minute break between each test.

Full (Fu)

In this baseline scenario, the operator performs the tasks independently, without cobot collaboration. The operator picks up the package from the starting station with the right hand, examines it with both hands to locate the QR code, brings it close to the scanner simulation area, and finally places the package in the arrival station with the left hand. The processing rate was freely self-selected by each participant.

Half Robot Touch (HRT)

In this condition, the cobot task is to pick up the package from the starting station and wait for the operator to extend their right arm toward the same area, and to touch it to start. When touched, the cobot moves bringing the package in front of the operator who grabs and examines it with both hands to locate the QR code. The operator then brings it close to the scanner simulation area and finally places the package in the arrival station with the left hand. The task is kinematically like the "Full" case, but there is no lifting of the package from the starting station toward the scanning area. Moreover, the operator drives the activation of the cobot with a touch, and therefore the number of PPM is decided by everyone, taking into account the cobot speed.

XSL•FO
RenderX

Half Robot (HRb)

In this modality, the cobot's task is to pick up the package from the starting station and autonomously transport it in front of the operator, who grabs and examines it with both hands to locate the QR code. The operator then brings it close to the scanner simulation area and finally places the package in the arrival station with the left hand. The operator is not required to initially reach the right arm toward the starting station to ask the cobot for help, so the task is different in terms of kinematic features compared with the previous cases. Packages are moved continuously by the cobot while the operator is performing the required tasks, therefore, the cobot determines the number of PPM.

Full Robot (FRb)

In this condition, the cobot's task is to pick up the package from the starting station and autonomously transport it continuously in front of the operator, who grabs and examines it with both hands to locate the QR code. The operator then brings it close to the scanner simulation area and finally places the package in a closer area with respect to the arrival station (Figure 1). Meanwhile, the cobot picks up the package left by the operator and takes it to the arrival station. The cobot intervenes significantly, fully collaborating in the activity. However, the time taken by the cobot to move the packages always exceeds the time that the operator needs to perform the activity. Therefore, the operator must wait for the cobot, and the motor activity is discontinuous, with the number of PPM defined by the cobot kinematics.

Ethical Considerations

The experimental protocol was approved by the Commissione Etica dell'Università degli Studi Roma Tre (BRIC2019-BRISK approval r.01-10/06/2021). Informed consent was obtained from all participants prior to their inclusion in the study, and they were informed of their right to withdraw at any time without consequences. All data collected were anonymized to protect participants' privacy and confidentiality. No personally identifiable information was retained, and strict security measures were implemented to safeguard the data. Participants did not receive any financial or material compensation for their involvement in this study.

Data Acquisition and Processing

To assess the dynamic interaction of the operator with the ground, two 6-component force platforms (BTS P-6000) were positioned under the participants' feet. These platforms allow for measuring the ground reaction of each participant's lower limbs (force and moment components), from which the center of pressure (CoP) position of each foot is extracted during the execution of performed tasks. Data were acquired from force platforms using a sampling frequency of 500Hz (fixed by the HW system). An optoelectronic system (SMART DX 6000, 8 infrared cameras, 2 Mpixel@300 Hz) was employed to segment and identify each task repetition, and each phase of the repetition. Data from the optoelectronic system were acquired using a sampling frequency of 250Hz. Force platforms data were undersampled at 250 Hz to synchronize those with the optoelectronic system ones.

Dynamic Parameters Extraction

Trajectories of the CoP for each of the lower limbs (ie, CoP-right and CoP-left) during the execution of each task were extracted from the force plates data [26]. The total CoP is obtained as the weighted average between CoP-right and CoP-left and its displacement is defined by both the medial-lateral and the anterior-posterior coordinates with respect to a reference system centered between the 2 force plates (Figure 1).

Four CoP-derived parameters were calculated to evaluate the postural stability of the operator and quantify alterations in balance: for each of the collaboration modalities, the mean distance (MDIST), mean velocity (MVELO), 95% confidence ellipse area (AREA-CE), and sway area (AREA-SW) were calculated, according to the definitions reported in [27], for each cycle of the task repetition. For each parameter, mean values (as a measure of central tendency) and SD across repetitions (as a measure of intraindividual variability) were also calculated and underwent inferential statistics.

Cycles were identified using data extracted from motion capture data, which were synchronized with force plate data: in particular, the time location of the maximum excursion to the left in the medio-lateral direction of the marker placed on the right wrist was used to segment each cycle, as it corresponds to the instant when each repetition finishes.

Statistical Analysis

For each CoP-derived parameter thus obtained, the normality of the data, using the Shapiro-Wilk test, was assessed. When all distributions were normal, the homogeneity of the variances was checked with the Leneve test. If these conditions were not satisfied, the Kruskal-Wallis nonparametric test was applied to examine the effect of the collaboration modality. When significant, a post hoc test with a Bonferroni correction was conducted to investigate the differences among the 4 examined collaboration modalities. The α level was set to .05 and .01 (denoted by * and **, respectively).

Results

Rate of Package Processing

The average rate of package processing chosen by the operators in the absence of cobot assistance (Fu) was 15 (IQR 14-16) PPM. The different cobot collaboration modalities yielded median rates of 8 (IQR 7-8.5) for PPM for half robot touch (HRT), 12 (IQR 11.5-13) PPM for half robot (HRb), and 4 (IQR 3.9-4.1) PPM for full robot (FRb).

CoP Trajectories

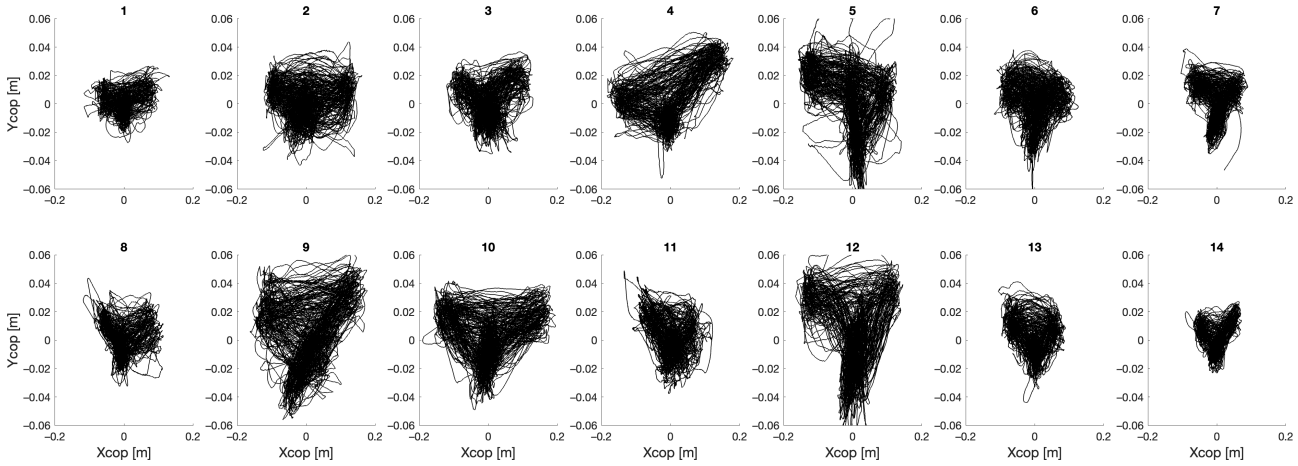
In the following, the different trajectories of the CoP, for all 14 participants and for the 4 different collaboration modalities (ie, full [Fu], HRB, HRb, and FRb) are shown, with Xcop and Ycop denoting respectively the medium-lateral and the antero-posterior coordinate.

Considering the Fu modality (Figure 2), observed CoP paths can be classified into 2 main groups: participants (1,2,3,6,7,8,11,13,14) displayed a compact size of the CoP path, spread in both anterior-posterior (AP) and medial-lateral (ML)

XSL•FO
RenderX

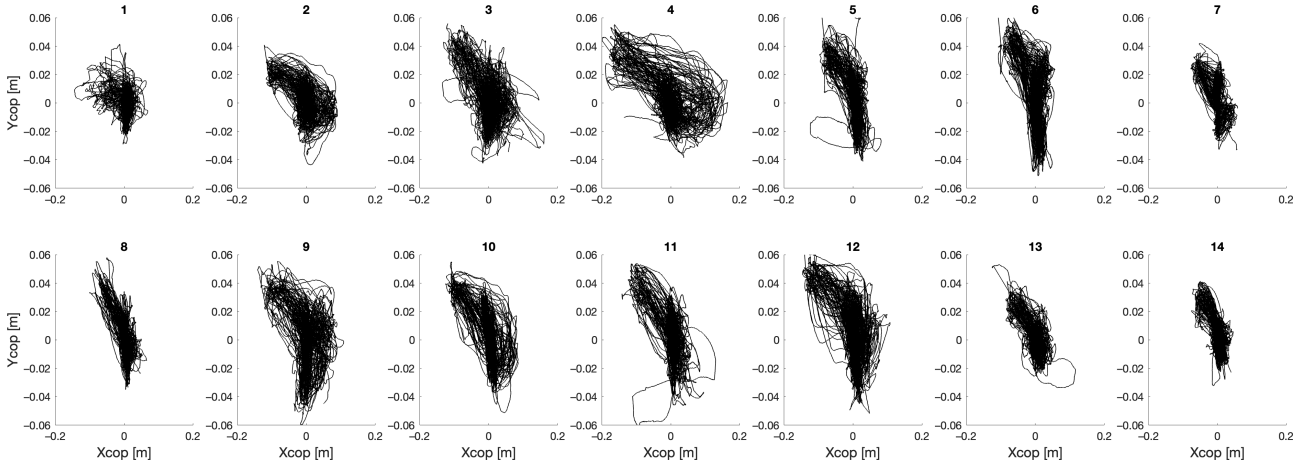
directions, while participants (4,5,9,10,12) drew wider areas, with a remarkable trajectory from the central backward position to the right forward position. However, all the CoP paths present quite diverse behavior, with no visible common aspects. Negative values for Ycop denote backward displacements.

Figure 2. Center of pressure (cop) trajectories for each participant during the full collaboration modality.



Considering the HRT collaboration modality (Figure 3), CoP path can be classified into 2 main groups also in this case: participants (1,2,7,8,13,14) maintained a compact size of the CoP path, more spread along the AP direction, while participants (3,4,5,6,9,10,11,12) drew a wider area of the CoP path. In this case, all CoP trajectories denoted a bias from the central backward position to the left forward position. Displayed CoP paths still exhibit rather diverse behavior, but with more commonalities as compared with Fu modality.

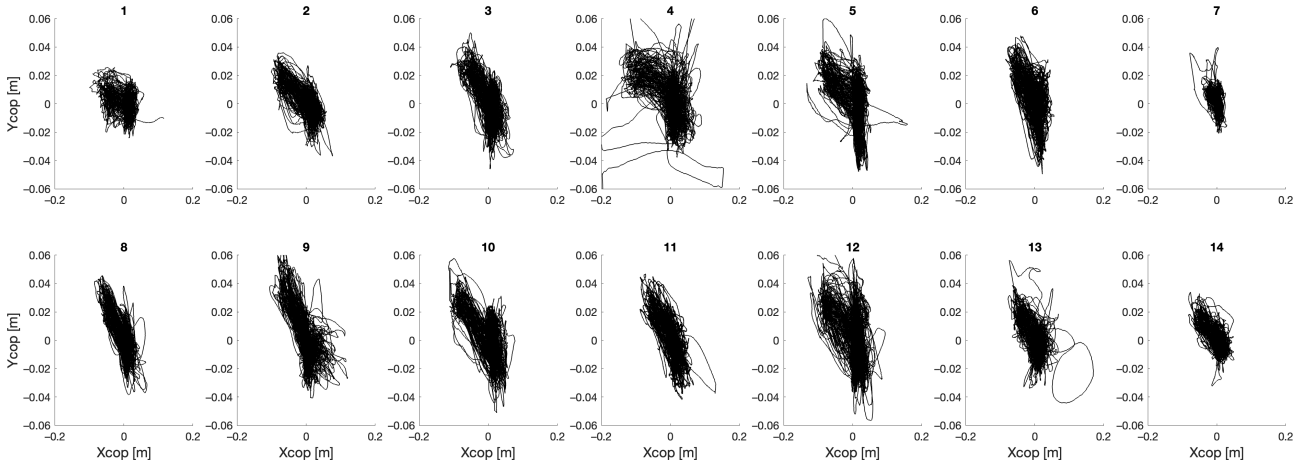
Figure 3. Center of pressure (cop) trajectories for each participant during the half robot touch collaboration modality.



For the HRb collaboration modality (Figure 4), also in this case CoP path can be classified into 2 main groups: for participants (1,2,7,13,14) CoP path was maintained compact, differently from participants (3,4,5,6,8,9,10,11,12), especially considering the AP coordinate. Common behavior of CoP excursion across subjects was instead present along the ML coordinate, while the CoP bias from the central backward position to the left forward position observed in the HRT case is less evident.

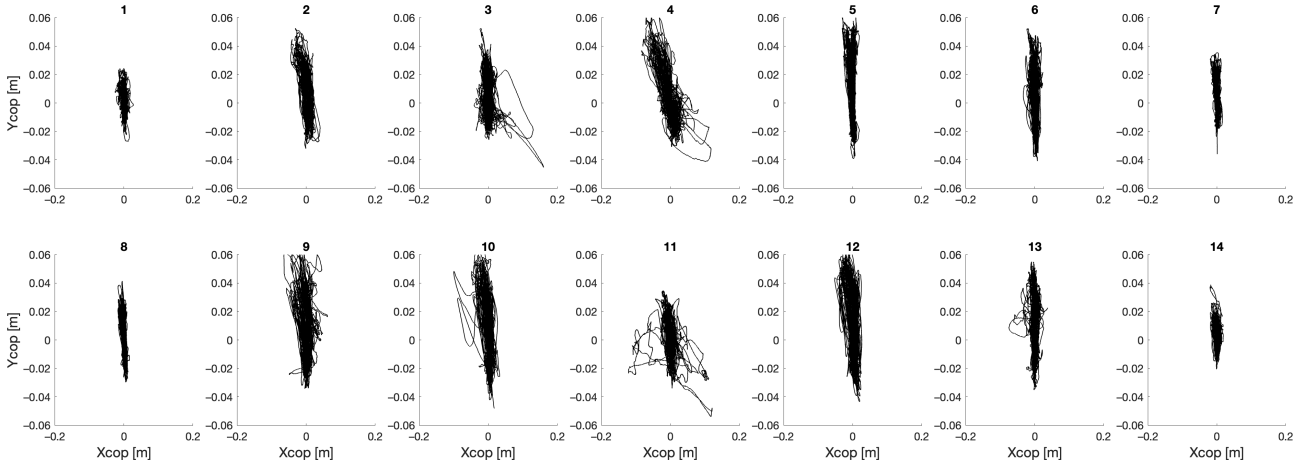
XSL•FO
RenderX

Figure 4. Center of pressure (cop) trajectories for each participant under the half robot collaboration modality.



In the FRb collaboration modality (Figure 5), commonalities of CoP behavior are more relevant. While some participants still display lower excursions along the AP direction, the characteristics of CoP trajectories are mostly similar, with no bias both in the right and left forward positions. The amount of excursion in the ML direction is similar across participants.

Figure 5. Center of pressure (cop) trajectories for each participant during the full robot collaboration modality.



Normality Tests

The Shapiro-Wilk test indicated that some conditions of the parameters did not follow a normal distribution. For the MDIST values, a *P* value <.05, indicating rejection of the null hypothesis of normal distribution, was observed for the Fu condition (*P*=.048). In contrast, a normal distribution (*P* value >.05) was found for all the other conditions (HRT: *P*=.44; HRb: *P*=.57; FRb: *P*=.60).

Similar results were obtained for both the AREA-CE and AREA-SW parameters. The Fu condition showed a nonnormal distribution (*P*=.02 for AREA-CE and *P*=.04 for AREA-SW), while all other conditions were normally distributed (for AREA-CE: HRT, *P*=.18; HRb, *P*=.25; FRb, *P*=.33; for AREA-SW: HRT, *P*=.08; HRb, *P*=.86; FRb, *P*=.25).

Since at least 1 of the distributions for each parameter was not normal, nonparametric statistical analyses were applied.

For the MVELO parameter, data were normally distributed across all conditions. However, Levene test revealed that the variances among the different conditions were not homogeneous (*P* value <.05). Consequently, a nonparametric test was also used for this parameter.

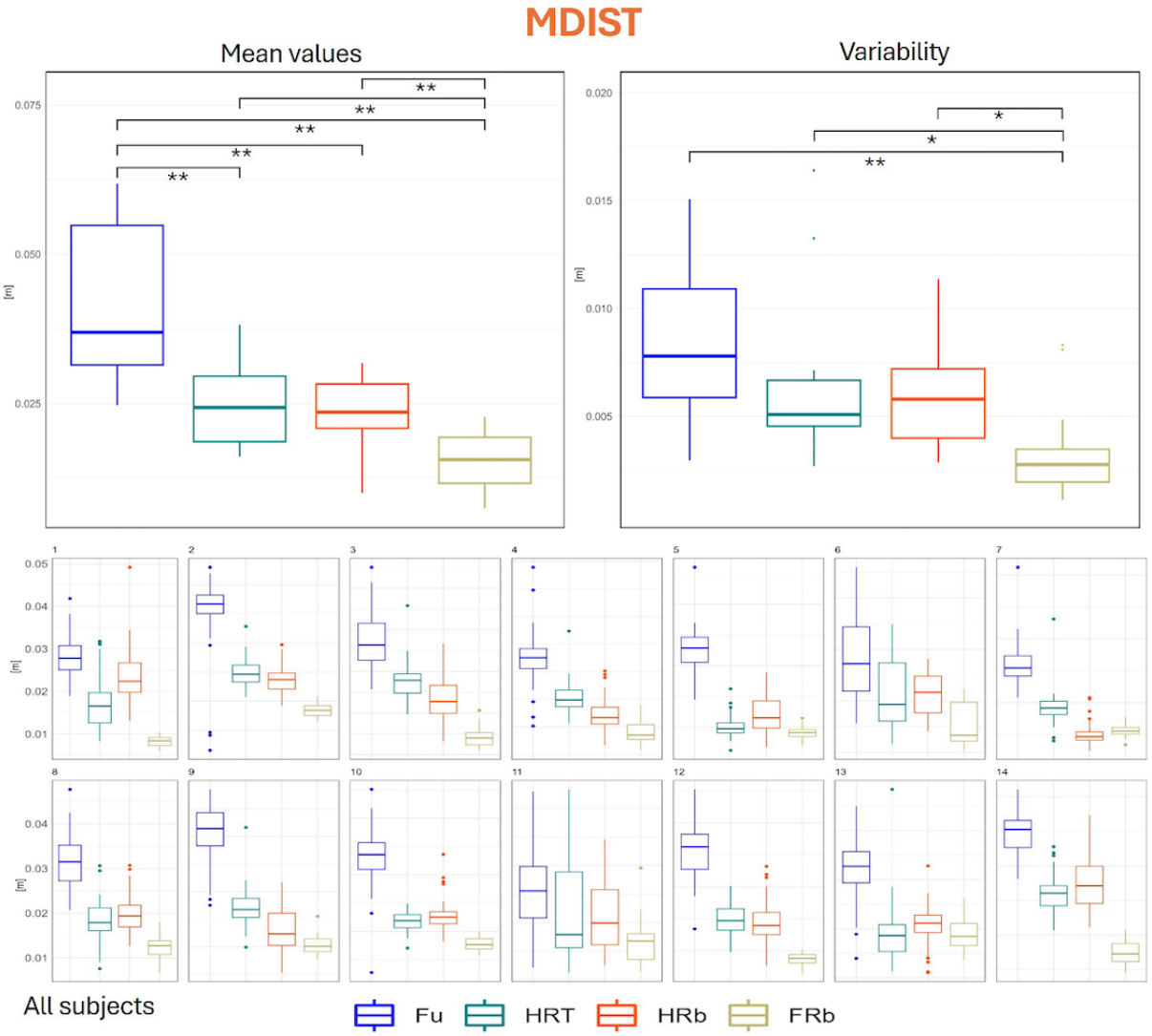
CoP-Derived Parameters

Descriptive statistics at the individual level, and descriptive and inferential statistics at the group level for MDIST, MVELO, AREA-CE, and AREA-SW, as obtained across the 4 collaboration modalities, are given in the following figures.

MDIST, an overall measure of excursion of the CoP from its equilibrium position, was clearly and significantly higher in Fu condition (4.21, SD 1.30 cm) than in all other conditions, and this was confirmed also at the individual level for almost every participant. In addition, as shown in Figure 6, FRb condition caused significantly lower values (1.55, SD 0.50 cm) than both HRT (2.50, SD 0.67 cm) and HRb (2.38, SD 0.59 cm), and this appeared also in all individuals but 2. No significant difference was observed between HRT and HRb conditions, as the result of different behavior across participants. Regarding intraindividual variability, we observed a reduced variability of MDIST across most participants in FRb condition (0.34, SD 0.22 cm) as compared with HRb (0.59, SD 0.23 cm), HRT (0.65, SD 0.37 cm), and Fu (0.85, SD 0.36 cm).

XSL•FO
RenderX

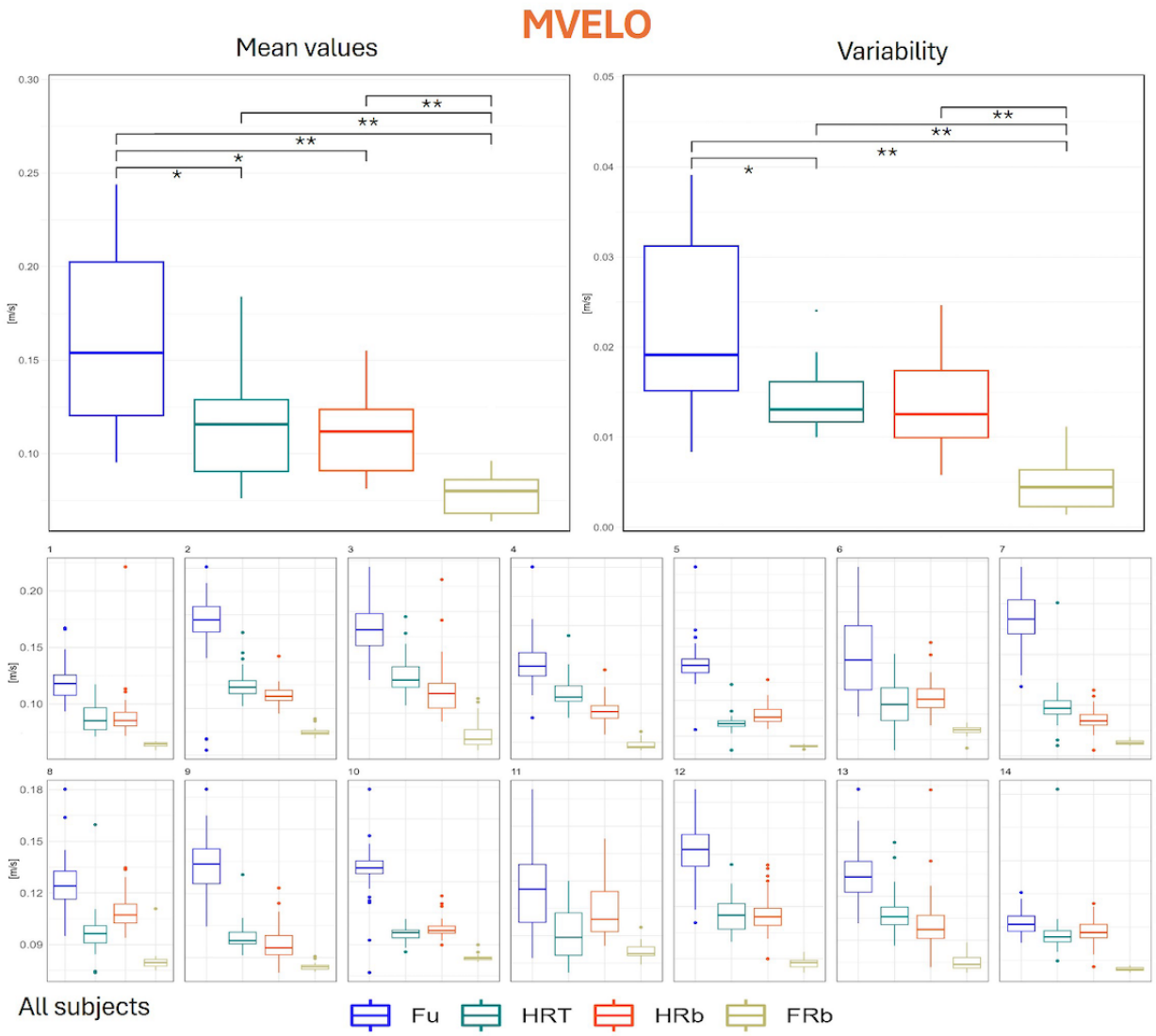
Figure 6. Distribution of MDIST values. Upper panel, left: group mean values across the 4 collaboration modalities; upper panel, right: distribution of intraindividual variability (SD across repetitions); lower panels: distribution of values for each participant. * denotes $P<.05$ as obtained from inferential statistics; ** denotes $P<.01$ as obtained from inferential statistics. FRb: full robot; Fu: full; HRb: half robot; HRT: half robot touch; MDIST: mean distance.



MVELO, an indicator of the amount of movement described by the CoP trajectory, in general confirms what already reported for MDIST: Fu condition determined higher values (16.3, SD 5.2 cm/s) than all other collaboration conditions, and in FRb, MVELO was significantly lower (7.9, SD 1.1 cm/s) than in all other conditions (11.1, SD 2.2 cm/s for HRb; 11.3, SD 2.9 cm/s for HRT). As depicted in Figure 7, these figures were confirmed also at the individual level for all but 2 participants. In terms of variability, while we observed that it was mostly reduced in the FRb condition (0.5, SD 0.3 cm/s) as compared with all the other conditions (2.3, SD 1.0 cm/s for Fu; 1.4, SD 0.5 cm/s for HRb), a significant decrease of the variability was present also when moving from Fu to HRT (1.4, SD 0.4 cm/s). Other variations were not significant at the group level.

XSL•FO
RenderX

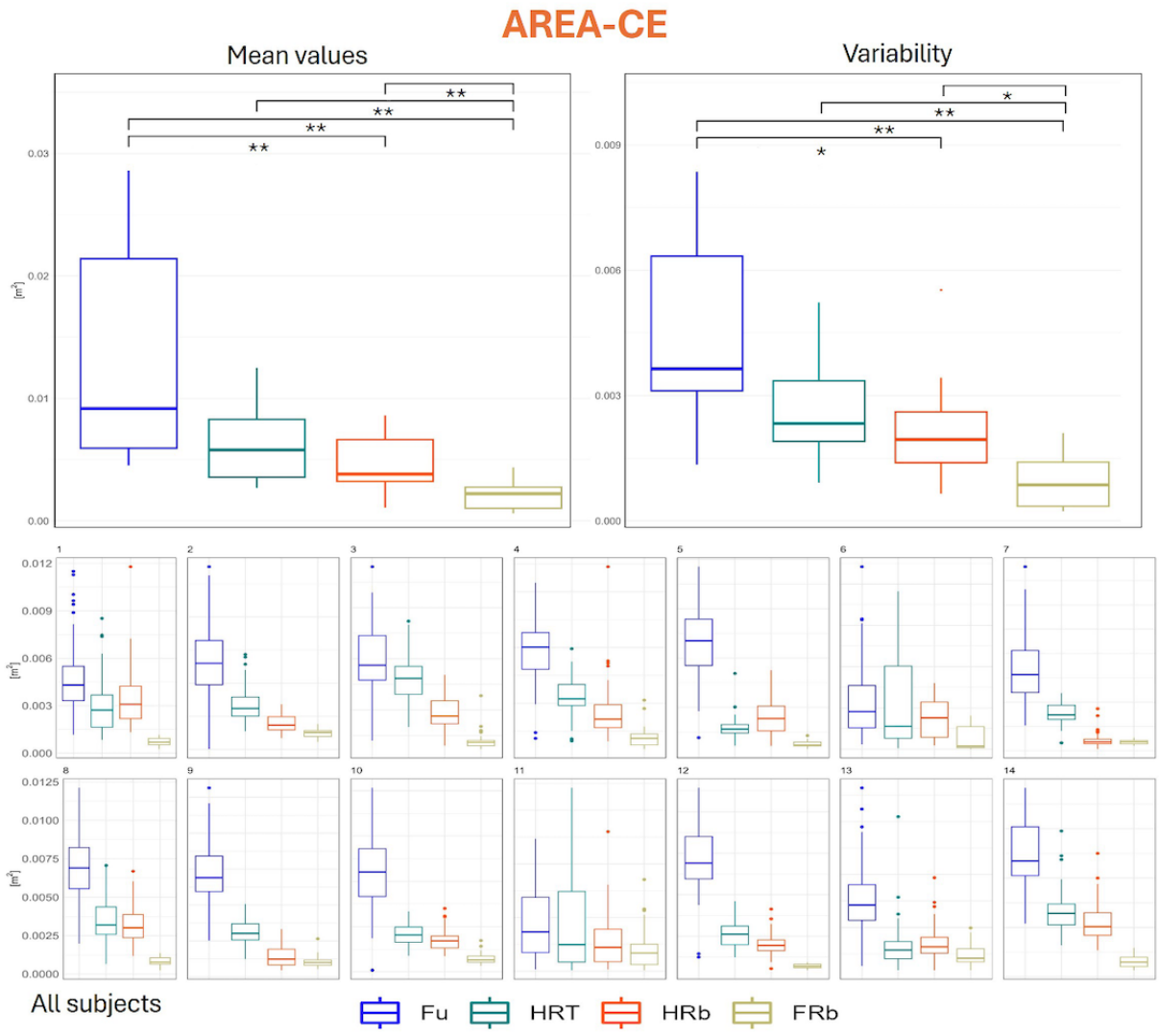
Figure 7. Distribution of MVELO values. Upper panel, left: group mean values across the 4 collaboration modalities; upper panel, right: distribution of intraindividual variability (SD across repetitions); lower panels: distribution of values for each participant. * denotes *P*<.05 as obtained from inferential statistics; ** denotes *P*<.01 as obtained from inferential statistics. FRb: full robot; Fu: full; HRb: half robot; HRT: half robot touch; MVELO: mean velocity.



AREA-CE, which quantifies the planar extent of coverage of the CoP, displayed a behavior similar to what was seen for MVELO: values resulted significantly higher in Fu (134, SD 91 cm²), and lower in FRb (22, SD 12 cm²) than in all the other conditions (64, SD 32 cm² for HRT; 47, SD 24 cm² for HRb), respectively (Figure 8). Again, FRb intraindividual variability was lower in FRb (10, SD 7 cm²) than in all other conditions, and we saw a reduced variability when passing from Fu (45, SD 23 cm²) to HRb (21, SD 12 cm²).

XSL•FO
RenderX

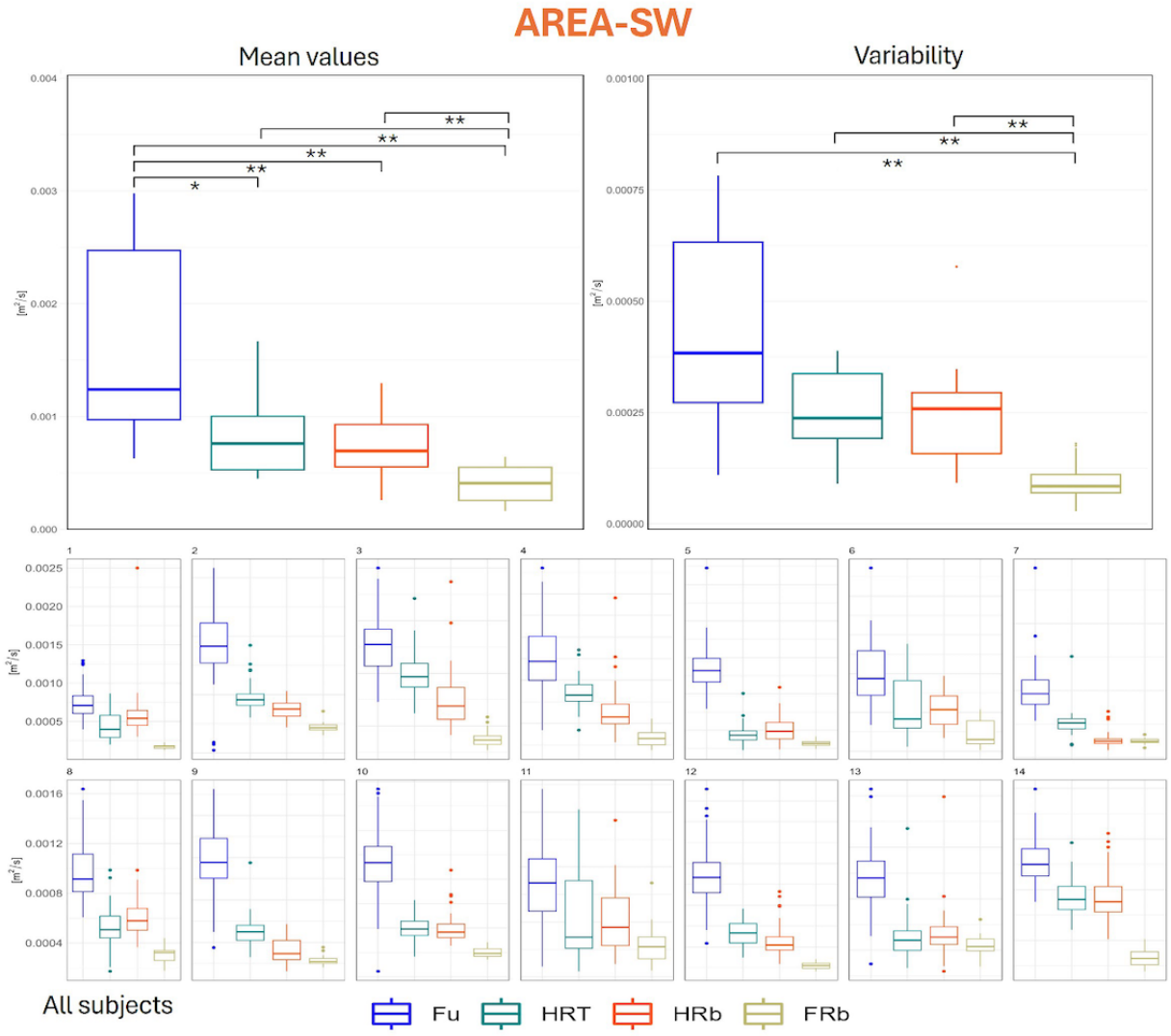
Figure 8. Distribution of AREA-CE values. Upper panel, left: group mean values across the 4 collaboration modalities; upper panel, right: distribution of intraindividual variability (SD across repetitions); lower panels: distribution of values for each participant. * denotes *P*<.05 as obtained from inferential statistics; ** denotes *P*<.01 as obtained from inferential statistics. AREA-CE: confidence ellipse area; FRb: full robot; Fu: full; HRb: half robot; HRT: half robot touch.



Finally, AREA-SW, a planar measure that combines CoP extension and velocity, displayed behavior similar to what obtained with MDIST: Fu caused significantly higher AREA-SW values (16.2, SD 8.4 cm²/s) than all the other conditions (8.4, SD 3.7 cm²/s and 7.3, SD 2.8 cm²/s for HRT and HRb, respectively), and again FRb brought to values lower (4.0, SD 1.6 cm²/s) than all the other collaboration modalities (Figure 9). Both these results were reflected at the individual level for all participants but one. Intraindividual variability of AREA-SW was significantly lower in FRb (0.9, SD 0.4 cm²/s) than in all other conditions (4.4, SD 2.2 cm²/s; 2.5, SD 0.8 cm²/s; and 2.5, SD 1.2 cm²/s for Fu, HRT, and HRb, respectively).

XSL•FO
RenderX

Figure 9. Distribution of AREA-SW values. Upper panel, left: group mean values across the 4 collaboration modalities; upper panel, right: distribution of intraindividual variability (SD across repetitions); lower panels: distribution of values for each participant. * denotes *P*<.05 as obtained from inferential statistics; ** denotes *P*<.01 as obtained from inferential statistics. AREA-SW: sway area; FRb: full robot; Fu: full; HRb: half robot; HRT: half robot touch.



Discussion

Principal Findings

The analysis of the CoP trajectories shows that considering data obtained in tests going from the absence of cobot intervention (Fu) to the ones with maximum cobot assistance (FRb), the area of the CoP trajectory significantly decreases, thus possibly indicating a lower amount of oscillations of the body when the cobot intervenes.

As a general consideration, CoP trajectories, when no action of the cobot is present, tend to show greater intensity toward the loading area, indicating a higher amount of sway. This suggests that lifting affects balance control strategies. On the other hand, in both tests involving partial cooperation with the cobot, the CoP draws trajectories mainly toward the arrival station, even in the HRT modality, where the participant's arm moves toward the starting station direction to ask the cobot for help. This aspect suggests that the movement of the arm while carrying the load notably influences the CoP trajectory, while the arm movement, required to ask the cobot for help occurring in the HRT condition, does not produce a significant influence on the CoP trajectory, as compared with the modality where the cobot independently carries the load to the scanning zone (ie, HRb collaboration modality).

Regarding the detailed analysis of the specific CoP trajectories of all subjects and for all test types, results show that during tasks performed alone, CoP paths can be classified into two types: (1) concentrated with symmetrical movements toward both the starting and arrival stations (right and left), and (2) more scattered with a tendency to oscillate between the central rest zone and the arrival station (left). The variability of CoP trajectories across individuals may confirm the adoption of different independent control strategies, as also identified by the 2 types listed before. When instead maximum assistance by the cobot is present, areas drawn by CoP are smaller, and with a reduced interindividual variability. This homogenization effect can be explained by the fact that the cobot sets the activity's pace, causing the subject to adopt a more uniform behavior and timing. This is both apparent at the individual level (ie, each

XSL•FO
RenderX

subject tends to maintain a more uniform CoP trajectory across repetitions), and at the group level.

Results coming from the qualitative description of CoP trajectories were also confirmed by statistical analysis of CoP-derived parameters: specifically, CoP tends to travel more (MVELO and, to a certain extent, AREA-SW) and further (MDIST and AREA-CE) if the level of intervention from the cobot increases. In general, smaller excursions can be associated with improved balance: this means a higher tendency to remain in a smaller (and generally more stable) area during task repetitions. A possible interpretation is that while the cobot influences task execution, it also makes the operator move less with reduced postural oscillation. This aspect is relevant as an increased excursion from the equilibrium position and increased movement of the CoP are associated with an increased risk of falling [28], and incorrect postures in the working scenario are agreed to lead to the development of MSDs [29]. As seen from this perspective only, the modality of full assistance by the cobot brings to the lowest risk, as it clearly leads to minimal excursions and movements of the body.

However, we also observed a marked reduction of intraindividual variability of these parameters when full assistance from the cobot is present: one hypothesis for this evidence is that the presence of the cobot reduces autonomy in task execution and may thus impose a control strategy which is not necessarily beneficial (or chosen) by the operator. If a reduction of variability of such parameters from the baseline condition may be seen as an indicator of long-term discomfort, this needs to be appropriately considered when designing the most beneficial collaboration modalities. In this regard, intermediate collaboration modalities seem to incorporate a tradeoff between the advantages in the carrying task compensation given by the cobot on the load and the possibility of granting autonomy to the operator during the task execution.

Regarding the 2 collaboration modalities that maintain a material operator intervention, statistical analysis at the group level was rather inconclusive: in excursion-based CoP parameters, no difference was found, and this is counterintuitive, considering that one of the modalities required participants to raise the arm and kinematically reach out to ask for assistance. This was not reflected in most CoP-derived parameters, and the possible explanation for this absence of differences may be ascribed to the inherent nature of the CoP variable: it is only indirectly associated with whole body movement: coming from the ground reaction forces, it is a measure of where the foot pressure is concentrated [30]. In static conditions, CoP basically tracks the body barycenter projection to the ground (and dynamically overtakes it in specific conditions) to counteract balance disequilibrium. According to the inverted pendulum hypothesis, the body barycenter projection is thus a filtered version of the CoP trajectory. However, with self-initiated body part movements, the central nervous system is able to predict future whole-body excursions [31], and the position of the CoP may not necessarily reflect the excursions of the body barycenter. If this were the case, activation of the relevant muscles (typically, tibialis anterior) would appear in anticipation of the upper limb movement.

Regarding the inconclusive results appearing between the 2 modalities where collaboration was midlevel (ie, HRb and HRT), we tend to ascribe this process to a rather diverse behavior observed at the individual level: decreases of measures of CoP excursion and movement were in some individuals clear when decreasing the amount of user intervention (ie, passing from HRT to HRb), however this effect disappears in a nonnegligible share of individuals; as it was apparent from the qualitative description of CoP trajectories, 2 different group behaviors seemed to emerge: in one group, CoP is maintained compact despite the lesser level of cobot intervention as compared with FRb, while the second group displays higher amounts of excursion and movement. We were not able to ascribe this phenomenon to specific individual characteristics; whole body kinematics and electromyography activity analysis may help shed light on this, in particular, to check if this evidence might be associated with task role-taking, which may be a key factor in HRC activities [32].

Reduced postural variability, as evidenced by the decreased CoP trajectory area and reduced inter and intraindividual differences under maximum cobot assistance, can have significant long-term implications for musculoskeletal health. A prolonged state of limited variability in posture and movement, while potentially stabilizing in the short term, may contribute to the development of MSDs over time. This is because variability in postural strategies allows the redistribution of mechanical loads across different tissues and joints, reducing the risk of overuse injuries or chronic strain on specific musculoskeletal structures. Conversely, homogenized and repetitive postures imposed by external systems, such as cobots, may lead to localized fatigue, reduced muscular engagement, and impaired adaptability of the neuromuscular system. Consequently, these factors can heighten the susceptibility to cumulative trauma disorders, particularly in dynamic occupational scenarios. These findings highlight the importance of designing collaborative robotic systems that balance task support with opportunities for natural and varied movement patterns to mitigate potential long-term health risks for operators.

The implications of these findings for industrial settings may be interesting. The integration of cobots can significantly enhance worker safety and productivity by reducing physical strain and improving postural control. By enabling safer and more ergonomic working conditions, cobots can play a crucial role in creating healthier work environments, reducing the incidence of work-related MSDs, and improving worker satisfaction.

In general, referring to the obtained results, it must be remarked that the tasks analyzed were simulated rather than conducted in real-world contexts, so further analysis could be conducted to generalize the findings. However, the tests conducted in this study, even if simulating a cashier task, represent real work tasks themselves, and this supports the hypothesis that similar results can be obtained in the real world. Additionally, the sample consisted exclusively of young male participants, so a direct validation on female subjects and older ones was not conducted. Rather than a limit, this might be seen as a validation for the examined population, although it can be extended to other worker categories.

XSL•FO
RenderX

Conclusions

This study investigated the biomechanical impact of various levels of HRC on postural control during working tasks while standing, providing substantial insights into the benefits of cobot assistance for workers' safety and ergonomics. Through an analysis of CoP trajectories and related parameters, it was found that higher levels of cobot collaboration, particularly in the HRT, HRb, and FRb modalities, significantly improved postural stability and reduced biomechanical load on workers. These findings suggest a decreased risk of MSDs, as evidenced by lower values in CoP-derived parameters when cobot assistance increases. The study also highlighted that the intraindividual variability of CoP parameters diminished when levels of cobot involvement were increased: this indicates on one side that cobot integration promotes more uniform ergonomic benefits across different individuals, but at the same time it may tend to impose a similar behavior in balance control which can lead subjects to adopt less ecological strategies.

Moreover, while participants can set autonomously their repetition pace, HRC activities will determine an increased homogenization of repetition rhythm (or parts of it) across workers, and this may potentially lead to decrease ergonomics, as subjects can feel forced to perform the required tasks based on timings dictated by HRC. Intermediate collaboration modalities, such as HRT, can provide both advantages in terms of reduction of CoP-derived parameters—that can be indirectly associated with a reduced biomechanical risk—with the possibility of selecting a more comfortable rhythm in executing tasks.

However, further research is needed to explore the long-term effects of cobot assistance on worker health and productivity. In particular, future studies should include other sensing technologies, able to capture the neuromechanics of HRC. In addition, application to diverse populations and real-world industrial settings to generalize the results and provide a more comprehensive understanding of the benefits and potential challenges of cobot integration. Additionally, exploring the psychological and social impacts of working with cobots on employees would provide a holistic view of the implications of collaborating scenarios in industrial environments: if properly integrated into working environments cobots can potentially revolutionize industrial work by enhancing safety, efficiency, and ergonomic well-being, facilitating the development of more intelligent and flexible workrooms and working conditions.

Acknowledgments

This work was partially supported by National Institute for Insurance against Accidents at Work (INAIL): BRIC2019-BRISK and BRIC2022-CoRoMan (CUP: F83C22002780001). Activities of MS, SR, and SC were also supported by Project ECS0000024 Rome Technopole (CUP: F83B22000040006).

Authors' Contributions

DB, GC, MS, and SC contributed to the conceptualization of the study. DB developed the methodology and designed the experimental setup. GC was responsible for software development. DB, GC, and MS performed validation. Formal analysis and visualization were conducted by DB. Data curation was carried out by DB, GC, and SR. DB prepared the original draft, while DB, GC, MS, SR, and SC participated in reviewing and editing the manuscript. Supervision was provided by DB and SC. MS and SC oversaw project administration, and SC secured funding for the study.

Conflicts of Interest

None declared.

References

1. Kagermann H, Helbig J, Wahlster W. Recommendations for implementing the strategic initiative INDUSTRIE 4.0: securing the future of german manufacturing industry. : Forschungsunion; 2013 URL: https://www.din.de/resource/blob/76902/e8cac883f42bf28536e7e8165993f1fd/recommendations-for-implementing-industry-4-0-data.pdf [accessed 2025-02-17]
2. Todescato M, Braholli O, Chaltsev D, et al. Sustainable manufacturing through application of reconfigurable and intelligent systems in production processes: a system perspective. Sci Rep 2023 Dec 16;13(1):22374. [doi: 10.1038/s41598-023-49727-5] [Medline: 38104211]
3. Ajoudani A, Albrecht P, Bianchi M, et al. Smart collaborative systems for enabling flexible and ergonomic work practices [industry activities]. IEEE Robot Automat Mag 2020;27(2):169-176. [doi: 10.1109/MRA.2020.2985344]
4. Kadir BA, Broberg O. Human-centered design of work systems in the transition to industry 4.0. Appl Ergon 2021 Apr;92:103334. [doi: 10.1016/j.apergo.2020.103334] [Medline: 33264676]
5. Keshvarparast A, Battini D, Battaia O, Pirayesh A. Collaborative robots in manufacturing and assembly systems: literature review and future research agenda. J Intell Manuf 2024 Jun;35(5):2065-2118. [doi: 10.1007/s10845-023-02137-w]
6. Villani V, Pini F, Leali F, Secchi C. Survey on human–robot collaboration in industrial settings: safety, intuitive interfaces and applications. Mechatronics (Oxf) 2018 Nov;55:248-266. [doi: 10.1016/j.mechatronics.2018.02.009]
7. Realyvásquez-Vargas A, Cecilia Arredondo-Soto K, Luis García-Alcaraz J, Yail Márquez-Lobato B, Cruz-García J. Introduction and configuration of a collaborative robot in an assembly task as a means to decrease occupational risks and

XSL•FO
RenderX

increase efficiency in a manufacturing company. Robot Comput Integr Manuf 2019 Jun;57:315-328. [doi: 10.1016/j.rcim.2018.12.015]
8. Cherubini A, Navarro B, Passama R, et al. Interdisciplinary evaluation of a robot physically collaborating with workers. PLoS ONE 2023;18(10):e0291410. [doi: 10.1371/journal.pone.0291410] [Medline: 37819889]
9. Johanson DL, Ahn HS, MacDonald BA, et al. The effect of robot attentional behaviors on user perceptions and behaviors in a simulated health care interaction: randomized controlled trial. J Med Internet Res 2019 Oct 4;21(10):e13667. [doi: 10.2196/13667] [Medline: 31588904]
10. Fiorini L, D'Onofrio G, Sorrentino A, et al. The role of coherent robot behavior and embodiment in emotion perception and recognition during human-robot interaction: experimental study. JMIR Hum Factors 2024 Jan 26;11:e45494. [doi: 10.2196/45494] [Medline: 38277201]
11. Yerebakan MO, Gu Y, Gross J, Hu B. Evaluation of biomechanical and mental workload during human-robot collaborative pollination task. Hum Factors 2025 Feb;67(2):100-114. [doi: 10.1177/00187208241254696] [Medline: 38807491]
12. Lorenzini M, Lagomarsino M, Fortini L, Gholami S, Ajoudani A. Ergonomic human-robot collaboration in industry: a review. Front Robot AI 2022;9:813907. [doi: 10.3389/frobt.2022.813907] [Medline: 36743294]
13. Gualtieri L, Fraboni F, De Marchi M, Rauch E. Development and evaluation of design guidelines for cognitive ergonomics in human-robot collaborative assembly systems. Appl Ergon 2022 Oct;104:103807. [doi: 10.1016/j.apergo.2022.103807] [Medline: 35763990]
14. Gualtieri L, Fraboni F, Brendel H, Pietrantoni L, Vidoni R, Dallasega P. Updating design guidelines for cognitive ergonomics in human-centred collaborative robotics applications: an expert survey. Appl Ergon 2024 May;117:104246. [doi: 10.1016/j.apergo.2024.104246] [Medline: 38354552]
15. Ajoudani A, Zanchettin AM, Ivaldi S, Albu-Schäffer A, Kosuge K, Khatib O. Progress and prospects of the human–robot collaboration. Auton Robot 2018 Jun;42(5):957-975. [doi: 10.1007/s10514-017-9677-2]
16. Merlo E, Lamon E, Fusaro F, et al. An ergonomic role allocation framework for dynamic human–robot collaborative tasks. J Manuf Syst 2023 Apr;67:111-121. [doi: 10.1016/j.jmsy.2022.12.011]
17. Guideline for introducing and implementing real-time instrumental-based tools for biomechanical risk assessment. : European Committee for Standardization (CEN); 2023.
18. Ranavolo A, Ajoudani A, Cherubini A, et al. The sensor-based biomechanical risk assessment at the base of the need for revising of standards for human ergonomics. Sensors (Basel) 2020 Oct 10;20(20):5750. [doi: 10.3390/s20205750] [Medline: 33050438]
19. McDevitt S, Hernandez H, Hicks J, et al. Wearables for biomechanical performance optimization and risk assessment in industrial and sports applications. Bioengineering (Basel) 2022 Jan 13;9(1):33. [doi: 10.3390/bioengineering9010033] [Medline: 35049742]
20. Lind CM, Abtahi F, Forsman M. Wearable motion capture devices for the prevention of work-related musculoskeletal disorders in ergonomics-an overview of current applications, challenges, and future opportunities. Sensors (Basel) 2023 Apr 25;23(9):4259. [doi: 10.3390/s23094259] [Medline: 37177463]
21. Ranaldi S, Bibbo D, Corvini G, Schmid M, Conforto S. Modalities of sequential human robot collaboration trigger different modifications of trunk oscillations. Front Neurorobot 2023;17:1183164. [doi: 10.3389/fnbot.2023.1183164] [Medline: 37425334]
22. Corvini G, Ajoudani A, Conforto S, et al. Assessing biomechanical risks in human-robot collaboration: analysis of muscle activity with different intervention conditions. Gait Posture 2023 Oct;105:S18-S19. [doi: 10.1016/j.gaitpost.2023.07.308]
23. Varrecchia T, Chini G, Tarbouriech S, et al. The assistance of BAZAR robot promotes improved upper limb motor coordination in workers performing an actual use-case manual material handling. Ergonomics 2023 Dec;66(12):1950-1967. [doi: 10.1080/00140139.2023.2172213] [Medline: 36688620]
24. Winter D. Human balance and posture control during standing and walking. Gait Posture 1995 Dec;3(4):193-214. [doi: 10.1016/0966-6362(96)82849-9]
25. Draicchio F, Trebbi M, Mari S, et al. Biomechanical evaluation of supermarket cashiers before and after a redesign of the checkout counter. Ergonomics 2012;55(6):650-669. [doi: 10.1080/00140139.2012.659762] [Medline: 22455556]
26. Schmid M, Conforto S, Camomilla V, Cappozzo A, D'Alessio T. The sensitivity of posturographic parameters to acquisition settings. Med Eng Phys 2002 Nov;24(9):623-631. [doi: 10.1016/S1350-4533(02)00046-2]
27. Prieto TE, Myklebust JB, Hoffmann RG, Lovett EG, Myklebust BM. Measures of postural steadiness: differences between healthy young and elderly adults. IEEE Trans Biomed Eng 1996 Sep;43(9):956-966. [doi: 10.1109/10.532130] [Medline: 9214811]
28. Piirtola M, Era P. Force platform measurements as predictors of falls among older people - a review. Gerontology 2006;52(1):1-16. [doi: 10.1159/000089820] [Medline: 16439819]
29. Gómez-Galán M, Pérez-Alonso J, Callejón-Ferre Á, López-Martínez J. Musculoskeletal disorders: OWAS review. Ind Health 2017 Aug 8;55(4):314-337. [doi: 10.2486/indhealth.2016-0191] [Medline: 28484144]
30. Gage WH, Winter DA, Frank JS, Adkin AL. Kinematic and kinetic validity of the inverted pendulum model in quiet standing. Gait Posture 2004 Apr;19(2):124-132. [doi: 10.1016/S0966-6362(03)00037-7]

XSL•FO
RenderX

31. Kaewmanee T, Aruin AS. The role of predictability of perturbation in control of posture: a scoping review. Motor Control 2022 Jan 1;26(1):97-143. [doi: 10.1123/mc.2021-0074] [Medline: 34891127]
32. Sacheli LM, Tidoni E, Pavone EF, Aglioti SM, Candidi M. Kinematics fingerprints of leader and follower role-taking during cooperative joint actions. Exp Brain Res 2013 May;226(4):473-486. [doi: 10.1007/s00221-013-3459-7] [Medline: 23503771]

Abbreviations

AP: anterior-posterior
AREA-CE: confidence ellipse area
AREA-SW: sway area
cobot: collaborative robot
CoP: center of pressure
FRb: full robot
Fu: full
HRb: half robot
HRC: human-robot collaboration
HRT: half robot touch
MDIST: mean distance
ML: medial-lateral
MSD: musculoskeletal disorder
MVELO: mean velocity
PPM: packages per minute

Edited by A Kushniruk; submitted 30.07.24; peer-reviewed by A Ranavolo, J Walsh; revised version received 05.12.24; accepted 20.12.24; published 27.02.25.

Please cite as:
Bibbo D, Corvini G, Schmid M, Ranaldi S, Conforto S
The Impact of Human-Robot Collaboration Levels on Postural Stability During Working Tasks Performed While Standing: Experimental Study
JMIR Hum Factors 2025;12:e64892
URL: https://humanfactors.jmir.org/2025/1/e64892
doi:10.2196/64892

© Daniele Bibbo, Giovanni Corvini, Maurizio Schmid, Simone Ranaldi, Silvia Conforto. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 27.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Feasibility of a 12-Week, Therapist-Independent, Smartphone-Based Biofeedback Treatment for Episodic Migraine in Adults: Single-Center, Open-Label, 1-Armed Trial

Amalie Christine Poole[1,2], MD; Ingunn Grøntveit Winnberg[2], MD; Melanie Rae Simpson[3], PhD; Anker Stubberud[1,4], PhD; Kjersti Grøtta Vetvik[1,5], PhD; Marte-Helene Bjørk[1,6,7], PhD; Lise Rystad Øie[1,2], PhD; Petter Holmboe[8], MSc; Alexander Olsen[1,9,10], PhD; Erling Tronvik[1,2], PhD; Tore Wergeland[1,2,4], PhD

[1]Norwegian Centre for Headache Research (NorHead), NTNU Norwegian University of Science and Technology, Trondheim, Norway

[10]Department of Physical Medicine and Rehabilitation, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway

[2]Department of Neurology and Neurophysiology, St. Olavs University Hospital, Trondheim, Norway

[3]Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, NTNU Norwegian University of Science and Technology, Trondheim, Norway

[4]Department of Neuromedicine and Movement Science, NTNU Norwegian University of Science and Technology, Trondheim, Norway

[5]Department of Neurology, Akershus University Hospital, Lørenskog, Norway

[6]Department of Neurology, Haukeland University Hospital, Bergen, Norway

[7]Department of Clinical Medicine, University of Bergen, Haukeland, Norway

[8]Nordic Brain Tech AS, Oslo, Norway

[9]Department of Psychology, NTNU Norwegian University of Science and Technology, Trondheim, Norway

Corresponding Author:
Amalie Christine Poole, MD
Norwegian Centre for Headache Research (NorHead), NTNU Norwegian University of Science and Technology, Trondheim, Norway

Abstract

Background: Biofeedback is an established treatment principle for migraine, but home-based therapy with proven efficacy is not available.

Objective: This study aims to assess the feasibility, usability, and safety of 12 weeks of daily use of a novel medical device (Cerebri; Nordic Brain Tech AS) for therapist-independent multimodal biofeedback preventative treatment in adults with episodic migraine.

Methods: In this open-label, one-armed trial, 20 adult participants with episodic migraine used Cerebri for 12 weeks. The primary outcome was the feasibility of the Cerebri system, measured by the level of adherence to daily biofeedback and electronic headache diary (eDiary) entries. Secondary outcomes were safety, usability, and efficacy.

Results: Initial adherence to biofeedback was high (16/20, 80% in weeks 1 - 4), declining to 20% (4/20) by weeks 9 - 12. eDiary adherence remained high (15/20, 75% in weeks 9 - 12). Reduction in migraine days was not significant (–0.6, 95% CI –2.4 to 1.1 days; *P*=.47). App usability was impacted by software issues. No safety concerns were reported.

Conclusions: Cerebri demonstrates potential in self-managed migraine treatment, with strong initial engagement and high safety. Usability issues, including technical bugs, were identified as the most important modifiable cause for the decline in adherence. This highlights the need for further app refinement to sustain user engagement.

Trial Registration: Clinicaltrials.gov NCT05454319; https://clinicaltrials.gov/study/NCT05454319

(JMIR Hum Factors 2025;12:e59622) doi:10.2196/59622

KEYWORDS

mHealth; mobile health; app; biofeedback; therapist-independent; behavioral therapy; headache; migraine; neurology; psychophysiology; complementary; psychology; feasibility; safety; usability; adherence; mobile phone

XSL•FO
RenderX

Introduction

Migraine is a complex and multifactorial disease and includes autonomic dysregulation and alterations in stress-related behavior [1-3]. Notably, patients with both chronic and episodic migraine have reduced heart rate variability (HRV) in comparison to healthy controls [4,5]. Biofeedback uses technology to monitor physiological processes that are typically modulated unconsciously [6-8]. Humans can learn to modulate these physiological processes by receiving continuous feedback from them [6-8]. Biofeedback training has been linked to a range of positive health effects, including reduced severity of migraine symptom burden [7,9]. A recent study has shown biofeedback's ability to alter HRV in a migraine population [10], and meta-analytical evidence suggests that biofeedback-based interventions are effective in the prophylaxis of migraine [11]. However, despite its efficacy, biofeedback is not widely used. An observational study from a specialized headache care center found that patients with migraine referred for biofeedback therapy face various obstacles to pursuing treatment, including issues related to time constraints, financial obstacles (including, but not limited to, health care insurance), and limited provider availability [12]. Traditional biofeedback training requires a trained therapist and costly and cumbersome equipment, thus limiting use to specialized clinics. Digital therapeutics and mobile health (mHealth) provide great potential for increased access to self-administered nonpharmacological treatments [13,14]. Despite this potential, we are not aware of any market-available, therapist-independent migraine biofeedback treatment with proven efficacy [15].

To fill this gap, we have developed a multimodal biofeedback system for independent home-based use. Two development and usability studies on progenitor versions of the system have already been published, but with short test periods (4 and 2 wk home testing) [16,17]. A small sham-controlled trial in adolescents tested a previous version of the medical device for 8 weeks [18]. The trial was prematurely halted due to the COVID-19 pandemic and included only 16 participants (4 sham). Poor adherence (40% in wk 5 - 8) hampered the interpretation of the trial results. The purpose of this study is to assess the feasibility, usability, and safety of 12 weeks of daily use of a home-based biofeedback system as a preventative in adults with episodic migraine. The findings of the study are intended to lead to refinements in the feasibility and usability of Cerebri (version 0.1.0; Nordic Brain Tech AS) and guide study design choices for a future randomized controlled trial.

Methods

Study Design and Participants

The study was a single-center, open-label, one-armed interventional trial conducted at St. Olav's University Hospital in Trondheim, Norway, from May 2022 to September 2022. The target population was adults with episodic migraine from all of Norway. The total duration of study participation for each participant was 12 weeks.

Inclusion criteria were (1) 18 years of age or older, (2) episodic migraine with or without aura diagnosed by a neurologist or physician per the International Classification of Headache Disorders, Third Edition, (3) having kept a headache diary with at least 80% adherence as part of routine clinical care in the last 28-days prior to inclusion, (4) 4 to 14 migraine attacks per 28-day period in the 3 months prior to screening, (5) at least three months of experience with smartphone and access to an iOS or Android phone at home, (6) onset of migraine before the age of 50 years, and (7) signed written informed consent.

Exclusion criteria were (1) having a continuous background headache that never disappears completely; (2) more than 14 days of headache (all types) per 28-day period; (3) participants diagnosed with trigeminal autonomic cephalalgias and neuralgias or (4) secondary headache conditions including medication overuse headache according to the International Classification of Headache Disorders 3rd edition; (5) participants with pathologies that inhibit use of the device according to the instructions for use (eg, blindness and deafness); (6) participants currently using migraine prophylaxis (pharmacological or nonpharmacological); (7) participants who have previously failed 3 prophylactic pharmacological treatments; (8) opioid use (>3 d per month) or use of barbiturates at time of screening; (9) alcohol overuse according to The International Classification of Diseases, Tenth Revision, Clinical Modification Code F10.1 or illicit drug use; (10) participation in another clinical investigation; and (11) participants unlikely to follow Clinical Investigation Plan or where treatment seems futile in the opinion of the Investigator or have demonstrated an inability to sufficiently adhere to headache diary entries (<80%).

Interventions

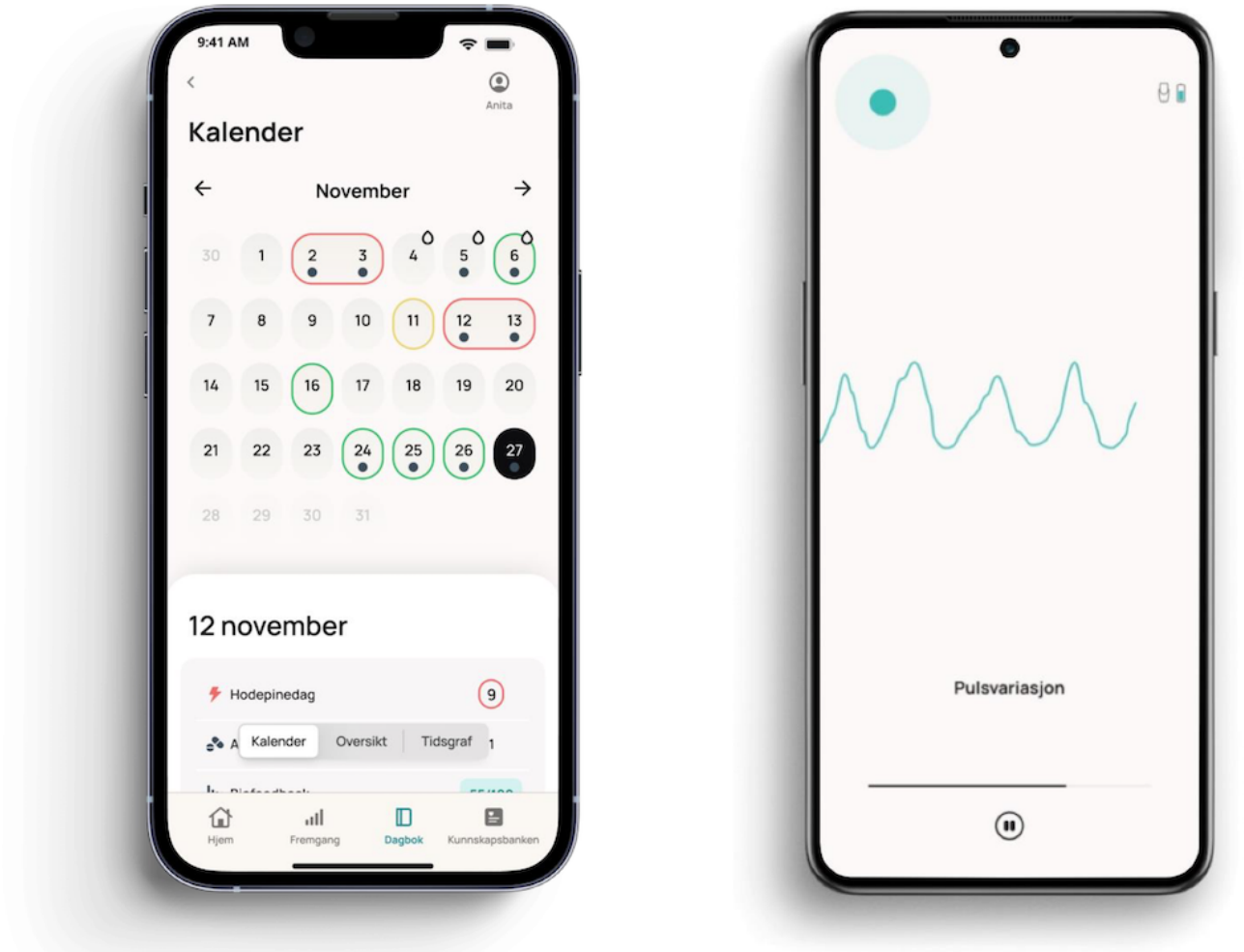
The medical device tested in this study is a therapist-independent biofeedback system (Cerebri) consisting of a smartphone app (Figure 1) and 2 noninvasive wireless sensors (Figure 2). One sensor measures neck muscle tension as surface electromyographic (EMG) voltage from the skin over the upper trapezius muscle. The other sensor is placed on one index finger to measure HRV by photoplethysmography and peripheral skin temperature. As part of the product development, validation of sensor performance showed agreement with simultaneous control measurements made with stationary neurophysiological equipment [19]. The 12-week treatment phase with Cerebri consisted of daily biofeedback therapy sessions of a minimum of 10 minutes duration each.

Since the device is intended for independent use without the need for any supervision, no instructional sessions were offered upon receipt of the device or during the course of the study. A limited, written step-by-step on-screen guidance was provided to onboard the participant prior to the first 3 biofeedback sessions. The participants were introduced to 3 biofeedback modalities with real-time on-screen feedback: temperature biofeedback during the first session, HRV biofeedback in the second session, and EMG biofeedback in the third session. During the remaining sessions, the participants could freely alternate between the 3 modalities at their own discretion. Additional user instructions were provided in a user manual and in a library in the app, where the users could read more about biofeedback and relaxation techniques. During the

XSL•FO
RenderX

biofeedback sessions, an adjustable breathing pacer was accessible to users.

Figure 1. Cerebri smartphone app interface. Screenshot of eDiary and HRV biofeedback interface. eDiary: electronic headache diary; HRV: heart rate variability.



XSL•FO
RenderX

Figure 2. Cerebri sensors: illustrations and photograph of the finger sensor (HRV/temperature) and sEMG sensor (muscle tension). HRV: heart rate variability; sEMG: surface electromyographic.

XSL•FO
RenderX

Study Follow-Up

All study visits were conducted remotely, and the sensor kit was sent by postal service. Participants accessed the app through a unique link provided via email, compatible with both Android and iOS devices. The app required an initial internet connection for download, but could subsequently operate without an active internet connection. Follow-up included 2 video consultations with a neurologist (IGW) at inclusion and the 12-week final visit; and 3 telephone follow-ups (wk 2, 4, and 8) with a study nurse. Participants were required to keep the daily eDiary integrated with the smartphone app throughout the study participation. Notifications for performing daily biofeedback sessions and diary entries were available for iPhone users, but did not work for Android users in the trial. Android users were therefore encouraged by study personnel to add a reminder manually in the integrated phone calendar to ensure this. The issue was persistent throughout the trial. Although participants were asked to register daily, the eDiary allowed registration within a 3-day interval (present day, and the 2 previous days). eDiary users were first asked if they had a headache or not. If a headache was reported, the app prompted the participant to answer additional questions about the type of headache (migraine or other), as well as the duration, intensity, and medication use. Questionnaires regarding usability and efficacy were collected via Helsenorge.

Outcome Measures

Overview

The primary objective of this study was to investigate the feasibility of the Cerebri biofeedback system as a preventative treatment in adults with episodic migraine. Key secondary objectives were to (1) assess usability, (2) assess the user experience, (3) assess the safety and tolerability, and (4) evaluate data on the efficacy of the intervention, recognizing that the study was not specifically designed or powered for an efficacy analysis.

Assessment of Primary Outcome

The feasibility of Cerebri was evaluated by assessing the level of adherence to daily biofeedback sessions and daily eDiary entries. Adherence was defined as the percentage of completed sessions and eDiary entries per 28-day period. On a given day, adherence was defined as the completion of a 10-minute biofeedback session or a completed eDiary entry, respectively.

Assessment of Secondary Outcomes

Usability, satisfaction, and ease of use of the device were assessed using a Norwegian translation of the mHealth App Usability Questionnaire (MAUQ) [20]. Questionnaires were collected at weeks 1 and 12. The MAUQ contains 18 items distributed across three subscales: (1) ease of use (5 items), (2) interface and satisfaction (7 items), and (3) usefulness (6 items). Responses for each item were recorded on a 7-point scale, ranging from 1=disagree to 7=agree, with higher scores indicating better usability. Users were allowed to select the option “not applicable” for each item. Such answers were considered missing data and excluded from the analysis.

Semistructured usability interviews in a subset of participants were used to gather qualitative data about the user experience of Cerebri for product development purposes. All study participants were asked if they wished to be considered for the interview, and 5 participants declined. From a pool of 15 potential participants, 6 individuals were selected based on a set of predefined criteria (Multimedia Appendix 1). Interviews were conducted midway (at wk 6) and upon completion of the study (at wk 12).

Data on adverse events (AEs) and device deficiencies were collected at phone consultations and the final visit. All treatment-emergent AE were collected throughout study participation.

Treatment efficacy was assessed using self-reported migraine and headache days that were collected prospectively in the integrated eDiary. The number of days in each 28-day period was compared to that collected from the past 28 days prior to inclusion.

Change in self-reported, migraine-related disability was measured using the Migraine-Specific Quality of Life Questionnaire version 2.1 (MSQ) from before treatment to end-of-treatment [21]. The questionnaire contains 14 items that measure the quality of life impacts of migraine in 3 domains: “role function-restrictive,” “role function-preventive,” and “emotional function.” Participants rate their experiences on a scale of 1=none of the time to 6=all of the time. The scores for each domain in the MSQ were calculated in accordance with the license holder’s instructions, with each domain receiving a score out of 100, where a higher score indicates more favorable functioning.

The participants’ perceived change in overall health was measured by the Patient Global Impression of Change (PGI-C) at weeks 4, 8, and 12. Participants were asked to rate their current condition compared to their condition before receiving the treatment intervention on a 7-point scale, ranging from a score of 1=very much improved to a score of 7=very much worse. For descriptive purposes, participants may be classified into 4 categories: much improved (score of 1 or 2), minimally improved (score of 3), unchanged (score of 4) or disease deterioration (score of 5, 6, or 7).

Data Analytic Strategies

Descriptive data were summarized as means with SDs or medians with IQRs according to normality distributions. Assessment of normality was based on visual inspection of histograms. Adherence rates to the eDiary and biofeedback program were calculated for each 28-day period separately. Summary statistics for MAUQ were calculated for weeks 1 and 12. Changes in the number of days with migraine and headache in each 28-day period compared to the last 28 days prior to inclusion were analyzed with a mixed logistic regression model with participant ID included as a random variable. The model estimated the odds ratio for each 28-day period. This odds ratio represented the estimated daily odds of migraine or headache in each 28-day period compared to the baseline period. The estimated number of migraines or headaches in each 28-day period and the difference from baseline were calculated based

XSL•FO
RenderX

on the regression coefficients using the nlcom command in Stata (version 17.0; StataCorp LLC). The use of a mixed logistic regression model allowed the inclusion of all available daily data, also from participants who had not completed all eDiary entries during each 28-day period. The change in MSQ domain scores was analyzed using mixed linear regression models with the mean change in each domain score estimated from these models.

Semistructured interviews were held through video or phone and recorded. Transcripts were stored on the Miro (RealtimeBoard Inc) platform during analysis, where they were grouped, mixed, and sorted in several rounds to identify patterns of themes that were relevant to the usability experience.

Ethical Considerations

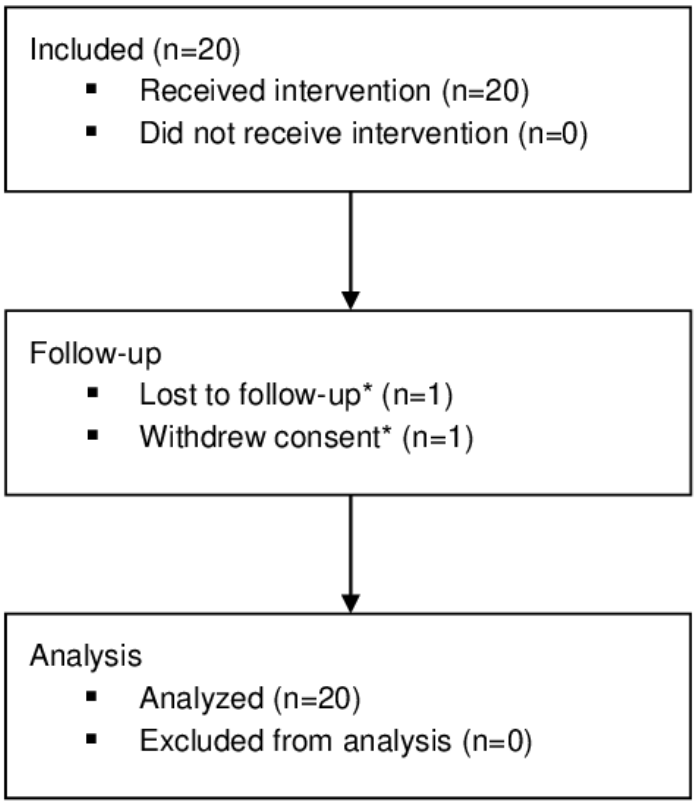
The study was approved by the regional ethics committee (457654) and the Norwegian Medical Products Agency (NoMA; 22/04658 - 4; Eudamed: CIV-NO-22-03-039126), and registered at ClinicalTrials.gov (NCT05454319) prior to initiation. Trial information was disseminated via multiple communication channels. Interested parties made contact by email. All participants received written information via email prior to the first study visit. Eligibility assessment and a review of the study information were performed by a consultant neurologist (IGW) with headache expertise in a video meeting. If eligible, consent was collected by digital signature via the national government-operated electronic data entry application Helsenorge (Norsk helsenett SF). Data collection for the study was initiated after electronic consent. Data were deidentified through pseudonymization. No payment was made to the study participants for their participation.

Results

Overview

A total of 20 participants were enrolled in the study intervention. All 20 participants contributed to the evaluation of the feasibility of the Cerebri biofeedback system. One participant withdrew consent, and one was lost to follow-up (Figure 3). The majority of participants were female and in full-time employment, with a mean of 6.7 (SD 2.8) migraine days per month (Table 1). One participant was pregnant during the entire study period.

Figure 3. CONSORT flow diagram. Intervention: Cerebri biofeedback treatment. *Data collected from participants who withdrew or were lost to follow-up were included in the analysis. CONSORT: Consolidated Standards of Reporting Trials.



XSL•FO
RenderX

Table . Baseline characteristics.

	Value (N=20)
Continuous variables, mean (SD)	
Age (years)	40.4 (9.3)
BMI (kg/m^2)	24.8 (3.1)
Baseline headache information[a]	
Number of headache days	7.7 (2.7)
Number of migraine days	6.7 (2.8)
Number of days with acute medication use	6.5 (3.1)
Categorical variables, n (%)	
Sex (female)	19 (95)
Pregnancy	1 (5)
Work status[b]	
Student	1 (5)
Full-time (≥100%) paid employment	15 (75)
Part-time (<100%) paid employment	1 (5)
Partial sick or medical leave	1 (5)
Sick or medical leave (100%)	1 (5)
Work assessment allowance	2 (10)
Previous use of preventive medication	4 (20)
Propranolol	1 (5)
Candesartan	1 (5)
Amitriptyline	2 (10)

[a]Participants all kept a headache diary for a minimum of 28 days prior to inclusion.

[b]One participant reported both full-time paid employment and "other governmental benefits" (parental leave). The student also received work assessment allowance and is thus registered twice.

Primary Outcome: Feasibility

Out of 20 participants, the majority continued to regularly record information in the daily eDiary throughout the study period, with 15 (75%) participants still recording information on at least 70% of days (20 d) in weeks 9 to 12 (Table 2). The proportion of participants completing at least 70% of biofeedback sessions decreased throughout the study, with 16 (80%) participants during weeks 1 to 4, and 4 (20%) participants during the final 4 weeks (Table 2).

XSL•FO
RenderX

Table . Adherence to eDiary[a] and biofeedback in the 3 intervals, respectively[b].

Period	Mean (SD)	Median (IQR)	Proportion with ≥20 days recorded of all included (N=20), n (%)
eDiary			
Day 1 - 28 (n=20)[c]	25.7 (3.8)	27 (15 - 28)	18 (90)
Day 29 - 56 (n=19)	25.5 (4.3)	27 (10 - 28)	18 (90)
Day 57 - 84 (n=18)	24.2 (4.8)	26 (12 - 28)	15 (75)
Biofeedback			
Day 1 - 28 (n=20)	20.7 (5.0)	22 (5 - 27)	16 (80)
Day 29 - 56 (n=19)	20.3 (5.6)	22 (6 - 26)	13 (65)
Day 57 - 84 (n=18)	16.7 (4.8)	17.5 (6-26)	4 (20)

[a]eDiary: electronic headache diary.

[b]The proportion of participants with adherence to intervention (eDiary and biofeedback, respectively) >20 days (=70% adherence) per 28-day period includes all participants (N=20).

[c]n is the number of participants in each period

Usability

The Cerebri app received favorable user perceptions, with MAUQ summary score for the ease of use subscale at week 1 (median score 6.2, IQR 5.6-6.6) and at week 12 (median score 6.0, IQR 5.2-6.8; Figure 4). Summary scores for the interface and satisfaction and usefulness subscales were 6.1 (IQR 5.7-6.7) and 5.0 (IQR 4.2-5.6) at week 1, and 5.8 (IQR 5.3-6.6) and 4.8 (IQR 3.5-5.2) at week 12, respectively.

XSL•FO
RenderX

Figure 4. MAUQ weeks 1 (n=19) and 12 (n=17). Median scores and IQRs from each question (1-18) with median summary scores for each of the 3 domains are presented. N is the number of responses to the question, the response “not applicable” is excluded. MAUQ: mHealth App Usability Questionnaire; mHealth: mobile health.

XSL•FO
RenderX

The app was regarded as both easy to learn and to use (questions 1 - 5), with stable median scores of 6 or 7 for almost all of these questions at both weeks 1 and 12. The exception was a notable decline in median scores from 6 (IQR 4-7) at week 1 to 4 (IQR 2.5-7) at week 12 for participants' ability to recover after having made a mistake using the app (question 5).

MAUQ questions concerning the app's interface and user satisfaction (questions 6 - 12) also received favorable scores. However, there was a noticeable decline in the median score from 7 (IQR 6-7) at week 1 to 5 (IQR 5-7) at week 12 for question 11, which assessed users' likelihood of using the app again. Nevertheless, overall satisfaction with the app (question 12) remained high with a median score of 7 (IQR 6-7) at week 1, and 6 (IQR 6-7) at week 12.

Furthermore, the scores for MAUQ question 13, which assessed the app's perceived usefulness for health and well-being, remained consistently high with a median score of 6 (IQR 4-7) and 6 (IQR 4-6) at weeks 1 and 12, respectively. Meanwhile, MAUQ question 18, which evaluated whether the app provides an acceptable way to receive health care services, showed a median score of 5.5 (IQR 4-6) at week 1, decreasing slightly to 5 (IQR 4-6) by week 12.

MAUQ question 17, which addressed whether the app could be used without or with a poor internet connection, maintained a median score of 4 (IQR 4-6) at week 1 and 4 (IQR 3-5) at week 12. Notably, this item had fewer respondents compared to other questions, with 9 participants responding at week 1 and 11 participants at week 12, suggesting that not all users encountered situations where the app's functionality under these conditions was tested.

Usability Experience

A total of 6 participants were recruited to participate in the semistructured interviews. Five midway user interviews were completed in total. Three of these interviewees also partook in the final user interview. All interviewees reported that Cerebri was easy to both learn and use. Finding time for daily training sessions and implementing this as a daily routine activity seemed to be the most challenging part. However, despite this challenge, interviewees expressed a sense of satisfaction in being able to dedicate 10 minutes to achieving a state of complete relaxation, which they found rewarding in itself.

A concern was identified that the app focused excessively on achievements and earning high scores. Participants feared that they might not experience the intended treatment effects unless they attained the top achievements. When probed further about their reasoning, participants attributed significance to such score achievements, assuming a correlation between achieving them and experiencing a reduction in migraine symptoms.

The finger temperature parameter was reported to be preferred among users, as it was the easiest to understand whilst also providing the most concrete goal during biofeedback sessions. Output user data from all completed biofeedback sessions from all study participants identified the same preference, with a total collective screen time of 75 hours used within the finger temperature modality across a total of 132 hours in 795 sessions. In comparison, 22 and 35 hours were used in the EMG and HRV modalities, respectively. The EMG parameter appeared to provide little feedback for participants to observe and improve upon, except for one individual who reported having tense muscles and experiencing a strong connection between muscle tension and migraines. All 5 participants expressed difficulties understanding the goal of HRV measurement, how to achieve it, or both.

Several bugs in the app occurred during the second half of the study, resulting in difficulties launching the app or starting biofeedback sessions. Notably, all iPhone users were involuntarily logged out of the app at one point due to these issues. All 3 participants partaking in the final user interview mentioned bugs as a reason for reduced motivation. In total, 13 unique bugs were reported by one or more participants during the study (not limited to the interviewees). These were perceived as disruptive to the user experience and led to disruptions in use patterns, ultimately impacting user motivation and compliance. The degree to which the 3 participants experienced a drop in motivation due to technical errors was not quantified specifically during the interview, but addressed for the entire population in questions 4 and 5 in the MAQU questionnaire.

Safety and Tolerability

Three AEs were recorded during the trial period (Multimedia Appendix 2). All AEs were assessed as nonrelated to the medical device. Thus, no adverse device effects and no serious adverse events occurred. One participant was pregnant during trial participation. No AE were registered for that participant, and the child was born healthy at term.

Exploratory Data on Efficacy

We observed a statistically nonsignificant reduction in the number of migraine days compared to baseline in all three 28-day periods, but no clear change in the number of headache days (Table 3). At all timepoints, approximately 50% (8/16) of participants described overall health conditions as "unchanged" on the PGI-C questionnaire (Figure 5). However, 22% (4/18) and 24% (4/17) of participants described their condition as "much improved" in week 8 and 12, respectively. No reports of worsening conditions were reported after week 4. There was a trend toward improvements in all MSQ domains throughout the study, reaching statistical significance for the domain "role function-restrictive" in week 4 and the domain "emotional function" in week 12 (Table 4).

XSL•FO
RenderX

Table . Number of days with migraine or headache in each 28-day period and estimated change from baseline.

Period	Observed	Estimated from mixed logistic model			
	Mean[a] (SD)	OR[b,c]	Estimated mean (95% CI)	Estimated difference from baseline (95% CI)	*P* value
Migraine days					
Baseline (n=20)	6.7 (2.8)	reference	6.2 (5.1 to 7.4)	reference	N/A[d]
Week 1 - 4 (n=20)	5.7 (4.5)	0.80	5.2 (3.4 to 7.0)	–1.0 (–2.3 to 0.3)	.13
Week 5 - 8 (n=19)	5.3 (4.0)	0.80	5.2 (3.4 to 7.1)	–1.0 (–3.0 to 0.9)	.31
Week 9 - 12 (n=18)	5.9 (4.2)	0.87	5.6 (3.8 to 7.4)	–0.6 (–2.4 to 1.1)	.47
Headache days					
Baseline (n=20)	7.7 (2.7)	reference	7.5 (6.3 to 8.7)	reference	N/A
Week 1 - 4 (n=20)	8.1 (4.2)	1.07	7.9 (6.1 to 9.6)	0.4 (–0.9 to 1.6)	.58
Week 5 - 8 (n=19)	8.0 (4.1)	1.04	7.7 (5.9 to 9.6)	0.2 (–1.7 to 2.2)	.81
Week 9 - 12 (n=18)	7.4 (4.1)	0.97	7.3 (5.5 to 9.1)	0.2 (–2.1 to 1.7)	.85

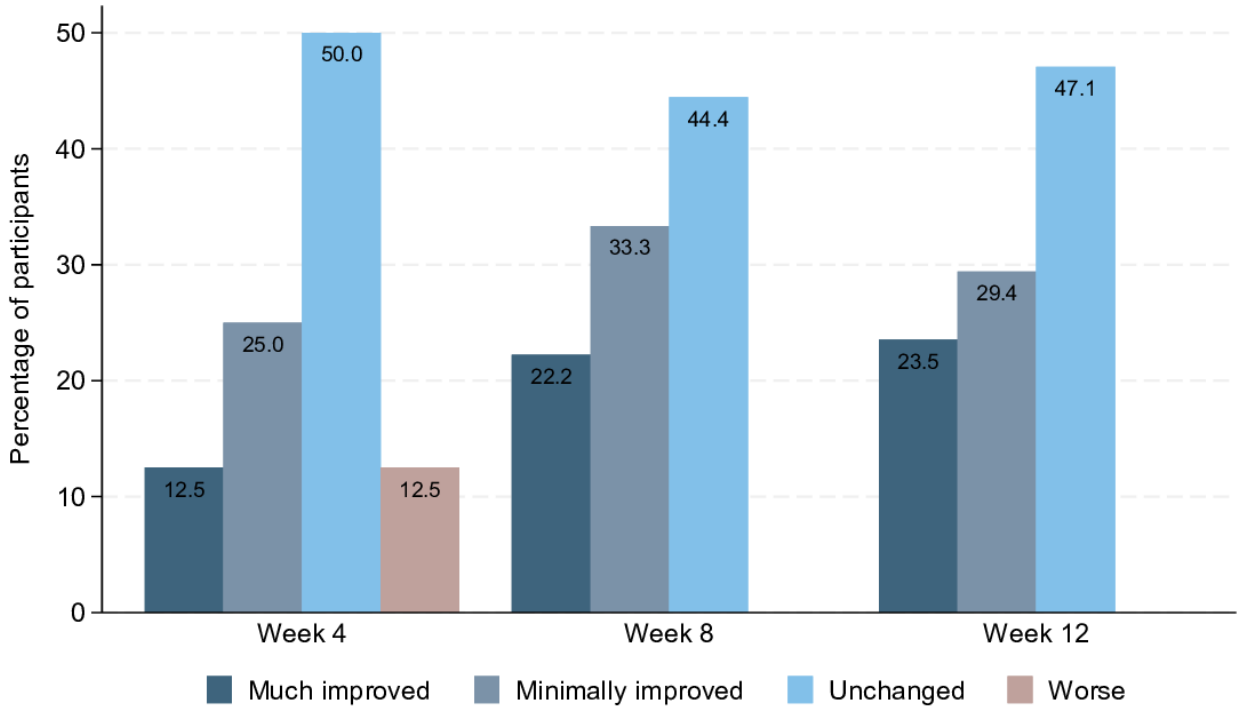
[a]Mean number of days based on available information in each 28-day period. For participants who did not complete the headache diary on all 28 days, the number of migraine or headache days was divided by the number of days with complete headache diary entries multiplied by 28 to obtain the number of days the participant would have been expected to have a migraine or headache.

[b]OR: odds ratio.

[c]A mixed logistic regression model was used to estimate the OR for each 28-day period. This OR represented the estimate OR compared to the daily odds of migraine and headache for each 28-days period. The estimated 28-day migraine and headache differences was calculated based on the regression coefficients and using the nlcom command in Stata.

[d]Not applicable.

Figure 5. Distribution of responses to the PGI-C question at weeks 4 (n=16), 8 (n=18), and 12 (n=17). The corresponding average (SD) for weeks 4, 8, and 12 were 3.6 (0.9), 3.2 (0.8), and 3.2 (0.8), respectively. PGI-C: Patient Global Impression of Change.



XSL•FO
RenderX

Table . Migraine-specific quality of life questionnaire scores at different timepoints and differences from baseline: presented according to the 3 domains.

	Observed		Estimated difference from baseline	
Week	Mean (SD)	Median (IQR)	Mean difference (95% CI)	*P* value[a]
Role function—restrictive domain				
Baseline (n=19)	61.2 (14.2)	60.0 (51.4-74.3)	reference	N/A[b]
Week 4 (n=16)	68.0 (10.2)	68.6 (60.0-77.1)	6.8 (0.6 to 13)	.03[a]
Week 8 (n=18)	69.8 (15.3)	70.0 (60.0-77.1)	8.6 (–0.2 to 17.4)	.06
Week 12 (n=17)	67.7 (17.9)	62.9 (60.0-80.0)	6.5 (–4.4 to 17.4)	.24
Role function—preventive domain				
Baseline (n=19)	75.8 (17.2)	75 (60-95.0)	reference	N/A
Week 4 (n=16)	81.6 (13.3)	85 (77.5-90.0)	5.7 (–1.8 to 13.2)	.14
Week 8 (n=18)	85.0 (15.9)	95 (65.0-95.0)	8.9 (–0.9 to 18.7)	.07
Week 12 (n=17)	80.6 (14.6)	80 (75.0-95.0)	4.6 (–5.8 to 15)	.39
Emotional function domain				
Baseline (n=19)	66.0 (20.2)	73.3 (53.3-80.0)	reference	
Week 4 (n=16)	72.9 (18)	76.7 (63.3-86.7)	5.7 (–1 to 12.5)	.10
Week 8 (n=18)	74.4 (18)	76.7 (60.0-86.7)	7.6 (–1.7 to 16.9)	.11
Week 12 (n=17)	76.9 (18.3)	80.0 (66.7-93.3)	9.9 (0.3 to 19.5)	.04[a]

[a]Indicate statistical significance ($P<.05$).

[b]Not applicable.

Discussion

Overview

This study was designed to address the feasibility of the home-based, therapist-independent biofeedback system Cerebri as a preventive treatment in adults with episodic migraine to serve as a preparatory step toward a clinical efficacy trial. Additionally, major emphasis was given to assessing the usability and the safety of the invention.

Principal Findings

Daily biofeedback session adherence declined throughout the course of the study from 80% in weeks 1 - 4 to 20% in weeks 9 - 12. The decline in diary adherence was much less pronounced, with a reduction from 90% to 75% in the same time interval. Low adherence is a common concern for nonpharmacological headache treatments [22]. This is also true for biofeedback treatment, as seen in a previous pilot on a precursor to the Cerebri system [18], as well as in another mHealth app for biofeedback intervention in migraine [23]. In this study, several causes for low adherence were identified.

First, software bugs in the last part of the study period hampered adherence, causing at one point all iPhone users to be logged off the app. Second, as eDiary adherence was better, software malfunction cannot be the sole cause for the declining biofeedback adherence. Usability assessments, including MAUQ scores, identified several other software elements that should be improved. Details were elucidated in semistructured interviews held in a subset of participants for product development purposes. The achievement features conveyed unrealistic goals for some participants, and failure to reach them was misinterpreted as a failed session by some. In addition, the goal of the HRV biofeedback training was difficult to understand.

This study was not designed to assess efficacy, and outcomes must be interpreted with caution. There was an observed tendency for improvements in several headache parameters, including quality of life measures. Notably, a not statistically significant average reduction in migraine days was observed across all three 28-day intervals compared to baseline, with an estimated reduction of 1 migraine day in weeks 1 - 4 and weeks 5 - 8. On PGI-C scoring, 24% (4/17) of participants described their overall health condition as “much improved” at week 12.

The intervention was safe, with no device-related or serious AE.

Interpretation

This study highlights the potential for therapist-independent, multimodal biofeedback systems to address unmet needs in migraine management, despite challenges in adherence. Adherence challenges, observed in this and prior studies [18,23], emphasize the need for robust technology platforms and user-centered design to secure long-term engagement, particularly in home-based interventions.

A possible explanation for the decreasing adherence to the biofeedback sessions, despite consistently high eDiary adherence, is that users experienced some sort of improvement from the therapy, and therefore, adjusted the frequency of treatment sessions while retaining the treatment effect. However, it may also be an expression of a lack of perceived effect, and

XSL•FO
RenderX

either way, user interviews did not identify this as a major mechanism.

There is no consensus on the optimal treatment duration or frequency of home-based mHealth behavioral interventions in migraine. Traditional therapist-dependent biofeedback is typically provided as hour-long sessions once or twice weekly, sometimes with intermittent home-training of shorter duration. In a recent pilot study examining a home-based single-modality surface EMG biofeedback intervention for episodic migraine, participants engaged in weekly sessions lasting 40 - 45 minutes under the remote guidance of a biofeedback therapist over a 6-week period [24]. Participants were instructed to do additional sessions of 30 minutes 3 times a week and keep track of their home practice. Participants attended on average 4.8 (SD 1.2) of the 6 sessions, whereas adherence to the therapist-independent sessions was not reported in the paper.

It can be hypothesized that increased treatment duration yields larger treatment effects. For home-based therapist-independent treatment, long sessions may hamper adherence, whereas short sessions may reduce potential efficacy. In a pilot study on a combined HRV-biofeedback-virtual reality device, participants attended an initial in-person session, after which they were instructed to perform sessions at home a minimum of 3 times per week, with each lasting at least 10 minutes [10]. There was an attrition rate of 9 out of 25 participants in the intervention group, compared to 3 out of 25 in the control group. No statistically significant reduction in headache days between the groups was observed. Results from this study suggest the possibility of significant dropout even with less frequent sessions. Therefore, increased session duration or frequency may not necessarily be determining factors for adherence to the intervention. This pilot study did identify adherence challenges, though other underlying mechanisms seem more important than session duration and frequency itself. Future work should empirically investigate optimal treatment duration and frequency.

Electronic behavioral interventions are dependent on sufficient user satisfaction to secure adherence. The overall positive perceptions of the app's ease of use and utility underscore Cerebri's potential to serve as an accessible, self-administered treatment option. However, the decline in median scores for certain usability aspects, particularly users' likelihood of using the app again, could signal the need to revise and improve certain elements of the app's design. On the other hand, a lack of perceived clinical effect could also explain such a finding. Regardless, addressing concerns related to usability features is crucial to maintaining patient compliance. To that end, revisions related to app achievement features should be a priority to mitigate misunderstandings that may affect participants' motivation and perceived treatment effectiveness. Other strategies to reinforce habit formation should be explored.

Furthermore, our results suggest that the app interface of the HRV biofeedback modality should be revised to facilitate the correct performance of training within this modality. This could be achieved by including a visual guiding line matching the optimal sinus curve for the selected breathing frequency. In addition, improving HRV-related textual information in the app may increase user accessibility.

Limitations

This pilot study has several limitations. First, the sample size of 20 participants makes the study underpowered to make claims regarding efficacy. Additionally, the absence of a control group limits our ability to attribute effects specifically to the intervention. A notable limitation of our qualitative analysis is the small sample size in the final user interviews, where only 3 participants participated. Additionally, the sequential selection of interview participants during the week 4 follow-up appointments could have introduced a selection bias as they were not randomly drawn. However, we aimed to include a diverse representation of age, gender, and adherence, and thus, opted for a sequential selection.

Participants were biofeedback naïve, but had already kept a headache diary with a minimum of 80% adherence in the 4 weeks prior to inclusion as part of routine clinical care, and were experienced smartphone users. Thus, study results may not be representative of a smartphone and eDiary naïve population.

Furthermore, as the study included only individuals with episodic migraine, caution is advised when inferring results to other patient groups, including chronic migraine or other primary headache disorders.

Conclusions

Cerebri is a unique multimodal, app-based, and therapist-independent biofeedback intervention that, in this study, underwent a rigorous feasibility and usability evaluation. The main feasibility obstacle was declining adherence to daily biofeedback sessions throughout a 12-week treatment period. On the other hand, eDiary adherence was consistently high. Usability issues and software bugs were identified and will be addressed in the final product development of Cerebri prior to a pivotal randomized controlled trial.

Acknowledgments

This clinical trial was funded by public research grants from the Eurostars project (E! 115549 – CEREBRI) and the Oslo Regional Research Fund.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

XSL•FO
RenderX

Conflicts of Interest

The sponsor of this trial, Nordic Brain Tech AS, as well as Norwegian University of Science and Technology (NTNU) and the Central Norway Regional Health Authority, may benefit financially from the commercialization of the proposed treatment. A commercial license agreement exists between Nordic Brain Tech AS and NTNU Technology Transfer AS on the commercialization of the treatment. ET, AO, and AS are cofounders and shareholders of Nordic Brain Tech AS. ET, AO, and AS hold a patent related to the Cerebri invention described in this study. In addition, ET, AO, and AS are coinventors of the proposed treatment in this study and may benefit financially from a license agreement between Nordic Brain Tech AS and NTNU Technology Transfer AS. PH is employed by Nordic Brain Tech AS and holds stock options in the company. ACP, IGW, MRS, KGV, MHB, LRØ, and TW have no ownership interest and do not own stocks of Nordic Brain tech AS or other companies developing biofeedback treatment. ACP has received travel grants from TEVA, AbbVie, and Pfizer. AS has received lecture honoraria from TEVA. KGV has received honoraria for lectures or advisory boards from AbbVie, TEVA, Novartis, Pfizer, Roche, Eli Lilly, and Lundbeck. MHB received speaker honoraria or served on scientific advisory boards for Teva, Eisai, AbbVie, Pfizer, Novartis, Lundbeck, Angelini Pharma, Jazz Pharmaceuticals, and Lilly over the last 5 years. LRØ has received honoraria for advisory boards from Roche. ET has received honoraria for lectures or advisory boards from Amgen, Novartis, Eli Lilly, AbbVie, TEVA, Roche, Lundbeck, Pfizer, Biogen. ET also holds stocks and intellectual property in Man & Science and Keimon Medical AS and serves as a consultant for Man & Science. TW has received honoraria from TEVA, Roche, Lilly, Lundbeck, and holds stocks in 2 medtech companies: Vilje Bionics AC and Keimon Medical AS.

Multimedia Appendix 1
Predefined criteria for semistructured usability interviews.
[DOCX File, 18 KB - humanfactors_v12i1e59622_app1.docx]

Multimedia Appendix 2
Registered adverse event (AE).
[DOCX File, 17 KB - humanfactors_v12i1e59622_app2.docx]

References

1. Shechter A, Stewart WF, Silberstein SD, Lipton RB. Migraine and autonomic nervous system function: a population-based, case-control study. Neurology (ECronicon) 2002 Feb 12;58(3):422-427. [doi: 10.1212/wnl.58.3.422] [Medline: 11839842]
2. Gazerani P, Cairns BE. Dysautonomia in the pathogenesis of migraine. Expert Rev Neurother 2018 Feb;18(2):153-165. [doi: 10.1080/14737175.2018.1414601] [Medline: 29212396]
3. Stubberud A, Buse DC, Kristoffersen ES, Linde M, Tronvik E. Is there a causal relationship between stress and migraine? Current evidence and implications for management. J Headache Pain 2021 Dec 20;22(1):155. [doi: 10.1186/s10194-021-01369-6] [Medline: 34930118]
4. Chuang CH, Li JY, King JT, et al. Abnormal heart rate variability and its application in predicting treatment efficacy in patients with chronic migraine: an exploratory study. Cephalalgia 2023 Oct;43(10):3331024231206781. [doi: 10.1177/03331024231206781] [Medline: 37851663]
5. Zhang L, Qiu S, Zhao C, Wang P, Yu S. Heart rate variability analysis in episodic migraine: a cross-sectional study. Front Neurol 2021;12:647092. [doi: 10.3389/fneur.2021.647092]
6. Kropp P, Niederberger U. [Biofeedback for headaches]. Pain 2010 Jun;24(3):279-289. [doi: 10.1007/s00482-010-0892-4]
7. Frank DL, Khorshid L, Kiffer JF, Moravec CS, McKee MG. Biofeedback in medicine: who, when, why and how? Ment Health Fam Med 2010 Jun;7(2):85-91. [Medline: 22477926]
8. Everly GS, Lating JM. Biofeedback in the treatment of the stress response. In: A Clinical Guide to the Treatment of the Human Stress Response: Springer New York; 2013:267-291. [doi: 10.1007/978-1-4614-5538-7_14]
9. Kropp P, Meyer B, Meyer W, Dresler T. An update on behavioral treatments in migraine—current knowledge and future options. Expert Rev Neurother 2017 Nov;17(11):1059-1068. [doi: 10.1080/14737175.2017.1377611] [Medline: 28877611]
10. Cuneo A, Yang R, Zhou H, et al. The utility of a novel, combined biofeedback-virtual reality device as add-on treatment for chronic migraine: a randomized pilot study. Clin J Pain 2023 Jun 1;39(6):286-296. [doi: 10.1097/AJP.0000000000001114] [Medline: 37026763]
11. Nestoriuc Y, Martin A. Efficacy of biofeedback for migraine: a meta-analysis. Pain 2007 Mar;128(1-2):111-127. [doi: 10.1016/j.pain.2006.09.007] [Medline: 17084028]
12. Minen MT, George A, Cuneo AZ. Factors associated with patient adherence to biofeedback therapy referral for migraine: an observational study. Appl Psychophysiol Biofeedback 2024 Jun;49(2):281-289. [doi: 10.1007/s10484-024-09622-9] [Medline: 38386246]
13. Wang C, Lee C, Shin H. Digital therapeutics from bench to bedside. NPJ Digital Med 2023 Mar 10;6(1):38. [doi: 10.1038/s41746-023-00777-z] [Medline: 36899073]
14. Rowland SP, Fitzgerald JE, Holme T, Powell J, McGregor A. What is the clinical value of mHealth for patients? NPJ Digital Med 2020;3(1):4. [doi: 10.1038/s41746-019-0206-x] [Medline: 31970289]

XSL•FO
RenderX

15. Noser AE, Gibler RC, Ramsey RR, Wells RE, Seng EK, Hommel KA. Digital headache self-management interventions for patients with a primary headache disorder: a systematic review of randomized controlled trials. Headache 2022 Oct;62(9):1105-1119. [doi: 10.1111/head.14392] [Medline: 36286601]
16. Stubberud A, Tronvik E, Olsen A, Gravdahl G, Linde M. Biofeedback treatment app for pediatric migraine: development and usability study. Headache 2020 May;60(5):889-901. [doi: 10.1111/head.13772] [Medline: 32052871]
17. Ingvaldsen SH, Tronvik E, Brenner E, et al. A biofeedback app for migraine: development and usability study. JMIR Form Res 2021 Jul 28;5(7):e23229. [doi: 10.2196/23229] [Medline: 34319243]
18. Stubberud A, Linde M, Brenner E, et al. Self‐administered biofeedback treatment app for pediatric migraine: a randomized pilot study. Brain Behav 2021 Feb;11(2):e01974. [doi: 10.1002/brb3.1974]
19. Stubberud A, Omland PM, Tronvik E, Olsen A, Sand T, Linde M. Wireless surface electromyography and skin temperature sensors for biofeedback treatment of headache: validation study with stationary control equipment. JMIR Biomed Eng 2018;3(1):e1. [doi: 10.2196/biomedeng.9062]
20. Zhou L, Bao J, Setiawan IMA, Saptono A, Parmanto B. The mHealth App Usability Questionnaire (MAUQ): development and validation study. JMIR Mhealth Uhealth 2019 Apr 11;7(4):e11500. [doi: 10.2196/11500] [Medline: 30973342]
21. Martin BC, Pathak DS, Sharfman MI, et al. Validity and reliability of the migraine-specific quality of life questionnaire (MSQ Version 2.1). Headache 2000 Mar;40(3):204-215. [doi: 10.1046/j.1526-4610.2000.00030.x] [Medline: 10759923]
22. Matsuzawa Y, Lee YSC, Fraser F, et al. Barriers to behavioral treatment adherence for headache: an examination of attitudes, beliefs, and psychiatric factors. Headache 2019 Jan;59(1):19-31. [doi: 10.1111/head.13429] [Medline: 30367821]
23. Minen MT, Corner S, Berk T, et al. Heartrate variability biofeedback for migraine using a smartphone application and sensor: a randomized controlled trial. Gen Hosp Psychiatry 2021;69:41-49. [doi: 10.1016/j.genhosppsych.2020.12.008] [Medline: 33516964]
24. Lazaridou A, Paschali M, Bernstein C, Curiel M, Moore S, Edwards RR. sEMG biofeedback for episodic migraines: a pilot randomized clinical trial. Appl Psychophysiol Biofeedback 2024 Jun;49(2):271-279. [doi: 10.1007/s10484-023-09615-0] [Medline: 38280149]

Abbreviations

AE: adverse event
EMG: electromyographic
HRV: heart rate variability
MAUQ: mHealth App Usability Questionnaire
mHealth: mobile health
MSQ: Migraine-Specific Quality of Life Questionnaire
PGI-C: Patient Global Impression of Change

Edited by A Kushniruk; submitted 19.04.24; peer-reviewed by AR Gantenbein, R Marshall; revised version received 12.02.25; accepted 04.03.25; published 09.06.25.

Please cite as:
Poole AC, Winnberg IG, Simpson MR, Stubberud A, Vetvik KG, Bjørk MH, Øie LR, Holmboe P, Olsen A, Tronvik E, Wergeland T Feasibility of a 12-Week, Therapist-Independent, Smartphone-Based Biofeedback Treatment for Episodic Migraine in Adults: Single-Center, Open-Label, 1-Armed Trial
JMIR Hum Factors 2025;12:e59622
URL: https://humanfactors.jmir.org/2025/1/e59622
doi:10.2196/59622

© Amalie Christine Poole, Ingunn Winnberg, Melanie Rae Simpson, Anker Stubberud, Kjersti Grøtta Vetvik, Marte-Helene Bjørk, Lise Rystad Øie, Petter Holmboe, Alexander Olsen, Erling Tronvik, Tore Wergeland. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 9.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Usability Evaluation of an Electrically Powered Orthopedic Exerciser: Focus Group Interview and Satisfaction Survey Study

Seojin Hong[1], MS; Hyun Choi[2], PhD; Hyosun Kweon[1], PhD

[1]Department of Clinical Rehabilitation Research, Rehabilitation Research Institute, Korea National Rehabilitation Center, 58 Samgaksan-ro, Gangbuk-gu, Seoul, Republic of Korea

[2]Department of Healthcare and Public Health, Rehabilitation Research Institute, Korea National Rehabilitation Center, Seoul, Republic of Korea

Corresponding Author:
Hyosun Kweon, PhD
Department of Clinical Rehabilitation Research, Rehabilitation Research Institute, Korea National Rehabilitation Center, 58 Samgaksan-ro, Gangbuk-gu, Seoul, Republic of Korea

Abstract

Background: Musculoskeletal disorders significantly impair physical function and quality of life, necessitating systematic rehabilitation. Electrically powered orthopedic exercisers, such as continuous passive motion devices, are widely used to enhance joint mobility and muscle recovery. However, existing devices often lack advanced functionalities and user-specific adaptability, limiting their effectiveness. To address these shortcomings, the Rebless Pro was developed as a novel device supporting active and passive exercises with personalized treatment programs.

Objective: This study aimed to conduct a formative usability evaluation of the Rebless Pro prototype using focus group interviews (FGIs) and satisfaction surveys with health care professionals specializing in rehabilitation medicine. The goal was to identify areas for improvement to enhance the safety, usability, and information clarity of the device.

Methods:: Usability evaluation was performed at the National Rehabilitation Center with 10 participants (5 physiatrists and 5 physical therapists) who had prior experience using similar devices. FGIs were conducted to collect qualitative insights into user experiences, while satisfaction surveys provided quantitative data on ease of use of the user interface and identifiability and understanding of information. Data collection focused on identifying risk factors and usability challenges.

Results: Three key areas for improvement were identified: (1) product upgrades to ensure patient safety, including adjustments to exercise speed and resistance; (2) hardware and software improvements to improve usability, including adjustments to the location of the emergency button and improvements to the graphical user interface elements; and (3) improvements to the user manual, including detailed contraindications, patient criteria, and clearer operating instructions. Although the mean score of physiatrists (mean 4.463, SD 0.298) was higher than that of physical therapists (mean 4.114, SD 0.829) in terms of the ease of use of the user interface, the difference was not statistically significant (P=.69). Similarly, in the category of identifiability and understanding of information, higher scores were again reported by physiatrists (mean score 4.467, SD 0.506) than by physical therapists (mean 3.733, SD 0.894), but this difference was also not statistically significant (P=.22).

Conclusions: Usability evaluation provided actionable insights into improving the Rebless Pro's safety, usability, and information clarity. To further refine the device, iterative usability evaluations involving both health care professionals and patients are recommended. These efforts are expected to contribute to the development of a safe, effective, and user-friendly electrically powered orthopedic exerciser suitable for commercialization.

(JMIR Hum Factors 2025;12:e60607) doi:10.2196/60607

KEYWORDS

usability evaluation; usability; electrically powered orthopedic exerciser; focus group interview; satisfaction survey

Introduction

Background

Musculoskeletal (MSK) disorders, resulting from diseases or injuries affecting muscles, tendons, ligaments, and bones, are a major cause of reduced physical function and quality of life [1]. Patients with MSK disorders often experience difficulties in performing independent daily activities due to reduced joint mobility, weakened muscle strength, and chronic pain, necessitating systematic and continuous rehabilitation therapy [2]. In clinical rehabilitation, electrically powered orthopedic exercisers play a crucial role in facilitating joint tissue regeneration and muscle strength recovery [3].

Continuous passive motion (CPM) devices are representative examples of exercisers; they are widely used to prevent joint stiffness and promote tissue healing, particularly after orthopedic

XSL•FO
RenderX

surgeries [4,5]. However, despite their therapeutic benefits, conventional CPM devices have several limitations that hinder optimal rehabilitation outcomes. First, they typically provide only passive range of motion (ROM) exercises without accommodating more dynamic modes, such as active-assisted or resisted exercises, which are essential for patients requiring progressive muscle strengthening or functional restoration [6,7]. Second, these devices lack user-specific customization features. Most CPM units operate on fixed protocols that do not account for individual differences in pain tolerance, ROM progression, or rehabilitation goals [8,9]. As a result, patients may experience discomfort or disengagement due to inappropriate exercise intensity or duration. Third, many CPM devices are ergonomically unsuitable for certain populations, such as individuals with obesity, severe joint deformities, or an inability to maintain a supine position. Since most CPM devices are designed for supine use, they may also limit usability and patient compliance in nonhospital settings [10,11].

To address these issues, alternative models, such as sitting-type CPM devices, have been developed. These devices allow patients to perform rehabilitation exercises in a seated position, which can enhance comfort, accessibility, and safety. Some studies have reported that sitting-type CPM devices offer comparable efficacy to conventional models while improving patient satisfaction [9]. Nevertheless, these alternatives are still relatively underused in clinical practice, partly due to limited availability and familiarity among clinicians.

In response to the need for more adaptive and patient-centered solutions, the Rebless Pro was developed as a next-generation orthopedic exerciser supporting both passive and active-assisted exercises. The device was designed to allow personalized treatment protocols and improve usability in various clinical settings.

Electrically Powered Orthopedic Exerciser

The Rebless Pro (H Robotics Inc) is an electrically powered orthopedic exerciser developed to support full ROM exercises and facilitate active and passive rehabilitation of the knee and ankle in individuals with neurological or MSK disorders (Figure 1). During the initial design phase, user requirements were collected from clinicians to inform the development of device functionalities and interface elements. Currently in its prototype stage, the Rebless Pro comprises a main unit and a tablet-based controller operated via the Rebless Pro app.

Figure 1. Electrically powered orthopedic exerciser.



The app features an intuitive graphical user interface (UI) and is structured around five core functional modules: (1) passive ROM, which enables therapist-initiated joint mobilization; (2) active-assisted ROM, providing adjustable support for patient-initiated movement; (3) evaluation ROM measurements, allowing objective tracking of joint range; (4) resisted ROM, designed to improve muscular strength and motor control through applied resistance; and (5) interactive custom setup, which enables therapists to personalize treatment protocols according to patient-specific needs.

The device operation workflow begins with patient registration and profile setup. Therapists select exercise modes and set individualized parameters—such as ROM extent, resistance, and assistance levels—based on therapeutic goals. During sessions, the app provides real-time feedback on joint movement, resistance, and exercise performance. Upon completion, session data are automatically stored, supporting longitudinal progress monitoring and outcome evaluation.

The Rebless Pro is intended for use by physical therapists and physiatrists, with the primary patient population being individuals with neurological or MSK impairments, including lower-extremity muscle paralysis or joint contractures, as well as people undergoing postsurgical rehabilitation. These users often require guided motor recovery and muscle strengthening to improve functional outcomes.

Usability Evaluation

Usability issues in medical devices can directly affect patient safety and treatment outcomes, often arising from deficiencies in UI design [12]. To mitigate such risks, the International Electrotechnical Commission (IEC) introduced the IEC 62366 standard, which emphasizes the integration of usability engineering throughout the device development process, distinguishing between formative and summative evaluations [13-15]. Formative evaluations, conducted during the early stages of design, aim to identify UI-related strengths, weaknesses, and potential use errors. These evaluations are iterative and typically involve direct input from intended users to guide design refinement [16], such as from focus group interviews (FGIs), which are widely used during the formative phase to collect user feedback on prototype usability [17,18].

This study conducted a formative usability evaluation of the Rebless Pro, which is currently in its implementation phase. Using FGIs and user satisfaction surveys with rehabilitation medicine professionals—the device's intended users—this study aimed to identify areas for improvement to enhance the safety and convenience of the device, in accordance with IEC 62366.

XSL•FO
RenderX

Methods

In this study, FGIs and satisfaction surveys were conducted to collect quantitative and qualitative data on an electrically powered orthopedic exerciser.

Ethical Considerations

This study was approved by the institutional review board of the National Rehabilitation Center (NRC; NRC-2023-03-020). All participants provided voluntary consent to participate after receiving a comprehensive explanation of the study objectives and methods. No compensation was provided for participating in the study. For the protection of personal information and the confidentiality of the participants, all data were anonymized.

Participants

The evaluation included 10 NRC-affiliated individuals, including 5 physiatrists and 5 physical therapists, who were recruited through peer referral and a bulletin board announcement at the NRC Usability Testbed.

The inclusion criteria were being a licensed physical medicine and rehabilitation or physical therapist with experience usingelectrically powered orthopedic exercisers and sufficient proficiency in Korean and English to comprehend software terminology. Individuals lacking experience in electrically powered orthopedic exercisers were excluded from the study. The participants provided voluntary consent after receiving a comprehensive explanation of the study objectives, methods, and procedures. All participants had at least 1 year of clinical experience using electrically powered orthopedic exercisers, such as CPM devices. The full formative assessment process, including the FGIs, was conducted separately for the physiatrist group and the physical therapist group.

Usability Evaluation Procedure

A usability evaluation was conducted at the NRC Usability Testbed. The evaluation was carried out in 2 separate sessions, one each for the physiatrist group and the physical therapist group, with each session lasting 90 minutes. The evaluator provided an introduction to the assessment to familiarize participants who may have been unfamiliar with usability evaluations, explaining the objectives, methods, and procedures of the product evaluation. Following this, participants were required to confirm their understanding of the evaluation details, after which informed consent was obtained for voluntary participation as well as for video and audio recordings. The FGIs commenced with questions regarding the general characteristics of the participants. Subsequently, the evaluator confirmed that the participants had reviewed the relevant information on the Rebless Pro and inquired about their prior experiences with similar devices, including the specific names of those devices. A demonstration of the use scenarios (Multimedia Appendix 1) was then conducted in conjunction with the FGIs using the Rebless Pro. Upon completion of the interviews, a satisfaction survey was administered to gather participants' feedback.

The usability evaluations were conducted in environments that closely resembled real-world use scenarios; such environments have been reported to accurately capture user experience and satisfaction, thereby playing a pivotal role in enhancing the commercial viability of medical devices [19]. In alignment with past findings, this study implemented formative evaluations within a simulated environment designed to replicate actual use conditions. The usability evaluations included a simulated physical therapy environment in which the Rebless Pro was used; the device was set up on a Bobath table (Figure 2). The luminous intensity, temperature, and relative humidity of the evaluation sites were measured before usability evaluation. The luminous intensity was 550 (SD 100) lux, temperature was 24ºC (SD 2ºC), relative humidity was 50% (SD 10%), and ambient noise was 50 (SD 5) dBA. Scenes of the evaluation are shown in Figure 3.

Figure 2. Evaluation environment.



XSL•FO
RenderX

Figure 3. Scenes of the evaluation.



FGI Procedure

FGIs were used as a qualitative data collection method to evaluate the usability of the Rebless Pro prototype. Each FGI session was conducted once per group, with 1 moderator, 1 observer for recording, and 5 participants present in the evaluation space. The moderator posed identical questions to each participant sequentially, ensuring that all participants—from the first to the fifth—had an opportunity to respond. The FGI questionnaire was structured following the guidelines of Krueger and Casey [17], incorporating introductory, transitional, key, and end questions (Multimedia Appendix 2). To gather feedback on the Rebless Pro, product demonstrations were conducted according to predefined use scenarios. Scenarios for 8 tasks were designed based on the user manual. Participants observed the device during these demonstrations and provided insights into predictable or identified risk factors and areas for UI improvement. The core questions focused on collecting participants' opinions regarding the UI of the Rebless Pro after observing its functionality and use. Additionally, participants offered suggestions for mitigating risks and improving usability based on their observations during the product demonstrations. Content analysis was conducted on the data obtained from the FGIs [20,21]. Microsoft Excel was used throughout the processes of data transcription, coding, and thematic derivation.

Satisfaction Surveys

After the FGIs, satisfaction surveys for the Rebless Pro were conducted to collect quantitative data. The satisfaction survey consisted of 51 items on the ease of use of the UI (Cronbach α=0.987) and 3 items on identifiability and understanding of information (Cronbach α=0.617). The survey items were rated on a 5-point Likert scale ranging from 1 (very difficult) to 5 (very easy). The difference in satisfaction between the physiatrist group and the physical therapist group was analyzed using the Mann-Whitney *U* test in SPSS (version 25; IBM Corp). The satisfaction survey results are presented as means and SDs for each item by group (Multimedia Appendices 3 and 4).

Results

Participant Characteristics

In the physiatrist group, clinical experience ranged from 1 year and 3 months to 2 years and 3 months. Experience in using similar medical devices ranged from 3 months to more than 1 year. In the physical therapist group, clinical experience ranged from 7 to 28 years, and experience using similar medical devices was more than 1 year for all members (Table 1). All members of both groups had experience using electrically powered orthopedic exercisers, such as MOTOmed (RECK-Technik GmbH) and other CPM devices, but all were using the Rebless Pro for the first time.

XSL•FO
RenderX

Table . General participant characteristics.

Number	Sex	Age (years)	Experience	Occupation	Experience using similar medical devices	Similar medical device
1	Female	29	2 yr 3 mo	Physiatrist	1 yr, >once a day	MOTOmed
2	Female	27	2 yr 3 mo	Physiatrist	1 mo, once a day	CPM[a] device
3	Male	26	2 yr 3 mo	Physiatrist	6 mo, once a day	MOTOmed
4	Female	28	1 yr 3 mo	Physiatrist	4 mo, once a day	MOTOmed
5	Male	25	1 yr 3 mo	Physiatrist	3 mo, twice a day	MOTOmed
6	Female	51	28 yr 10 mo	Physical therapist	1 yr, 3 times a day	MOTOmed
7	Male	40	14 yr	Physical therapist	1 yr, once a day	MOTOmed
8	Female	36	9 yr 10 mo	Physical therapist	1 yr, once a day	MOTOmed
9	Female	32	9 yr 6 mo	Physical therapist	1 yr, once a day	MOTOmed
10	Female	29	7 yr 3 mo	Physical therapist	1 yr, 3 times a day	MOTOmed, CPM device

[a]CPM: continuous passive motion.

FGI Results

The FGIs revealed 3 key factors and 7 subfactors representing possible improvements to the Rebless Pro. The three key factors were as follows: (1) product upgrades to ensure safety, (2) hardware and software improvements for convenience of use, and (3) improvement of the manual for better identifiability and understanding of the information (Table 2).

Table . Outcome of content analysis.

Factor	Suggested solution
1. Product upgrades to ensure safety	• Improve the exterior • Fasten the device to the table • Adjust the motor for exercise speed and resistance
2. Improvements for convenience of use	• Improve the hardware user interface • Improve the software user interface
3. Improvement to the manual for better identifiability and understanding of the information	• Input details about the intended patient group and medical indications • Input details about device use and operation

Product Upgrades to Ensure Safety

The participants prioritized patient safety and suggested the following potential risks of and improvements to the Rebless Pro:

> *The device should be fastened to the table, but force applied by the therapist should allow it to slide or move.*
>
> *The exterior requires improvement as the device has a complex structure and accidents may be caused through items getting caught on its exterior.*
>
> *The device should be secured using an anti-slip pad.*
>
> *Patients may complain of pain because the speed of the basic exercise may be too fast. The initial position angle should be adjusted while the device is being fitted.*
>
> *Movement speed needs to be more finely adjusted. The operating exercise speed should be shown on the controller display.*
>
> *The basic resistance intensity of the device may be too high for the patient to counteract.*
>
> *There is a high likelihood of errors occurring during ROM measurement, and measurement range errors may occur.*

Hardware and Software Improvements for Convenience of Use

For user convenience and ease of use, the following hardware and software improvements to the Rebless Pro were suggested:

> *The power (on, off) button should be repositioned.*
>
> *The emergency stop button should be repositioned, and separate buttons should be provided for the patient and therapist.*
>
> *The material of the fastening strap where the device is secured to the patient should be changed.*
>
> *Noise from the motor may interfere with the treatment of other patients. The motor noise should be adjusted.*

XSL•FO
RenderX

As this device is provided to the patient by medical staff, the wording of the instructions should be improved to make it easier for patients to understand (instead of using medical terminology).

The information about intensity and stiffness sensitivity is provided only in the manual. It is cumbersome to have to refer to the manual while using the device.

A pop-up window with a summary of each function should be added to the controller display.

Control icons, such as up and down arrows, should be presented more clearly on the display.

Improvement of the Manual for Better Identifiability and Understanding of the Information

Participants suggested the following improvements to the Rebless Pro manual to allow better identifiability and understanding of information by users:

Detailed information about the contraindications for intended patient groups, stages of life and age groups, specific criteria for applicable patients, and the patient's position when fitting the device should be provided.

With respect to use of the device, calf/thigh measurement criteria, exercise intensity criteria based on Manual Muscle Test or Modified Ashworth Scale standards, and clear exercise speed criteria should be provided.

Satisfaction Survey Results

To identify the difference in satisfaction between physiatrists and physical therapists, a Mann-Whitney *U* test was conducted. When comparing the overall ease of use of the UI, the satisfaction of the physiatrists (mean score 4.463, SD 0.298) was higher than that of the physical therapists (mean score 4.114, SD 0.829), but this was not a statistically significant difference (U=10.000; Z=−0.522; P=.69). In addition, when comparing the overall information identifiability, the satisfaction of the physiatrists (mean score 4.467, SD 0.506) was higher than that of the physical therapists (mean score 3.733, SD 0.894), but this was not a statistically significant difference (U=6.000; Z=−1.375; P=.22).

Discussion

Principal Results

A formative evaluation is a usability evaluation performed to improve the design of a medical device under development. To gain qualitative and quantitative insights into the course of the development of the Rebless Pro, an electrically powered orthopedic exerciser, we conducted a formative evaluation. This was achieved using FGIs and satisfaction surveys with rehabilitation medicine professionals who had experience using other electrically powered orthopedic exercisers.

FGIs enable individuals to freely express their own experiences and engage in discussions using not only their own knowledge but that of others, thereby providing diverse perspectives and detailed information on a specific topic [17]. Moreover, FGIs are used to identify and analyze expert responses as experts share information and experience in using a particular product or service. Therefore, FGI findings may be used to develop or improve products and services [22].

Motorized medical devices can cause unintended movements by the patient or user, potentially leading to discomfort or injury [23]. Because the operator of the Rebless Pro is a medical professional and the user is a patient, the participants in this study prioritized patient safety and presented opinions on improvement of the exterior structure of the device and upgrades to the device fastening method, exercise speed, and resistance strength. These points suggest that patient and operator safety is crucial when designing a rehabilitation device and that the product should be upgraded to ensure safety.

A previous study concluded that devices using robotic technology should be equipped with at least one emergency stop button that is easily accessible to the user [24], and another study reported that an emergency stop system should be installed to prioritize user and patient safety in the event of device malfunction or user error [12]. On the basis of their own experience using similar medical devices and the product demonstration, the participants in our study indicated the need for the emergency button to be redesigned and repositioned so that it would be more visible and easily accessible to medical staff or patients in the event of an emergency. Consequently, we determined that the emergency stop system of a medical device has a direct impact on the ability to prevent device malfunction and that it plays an important role in ensuring device safety and efficacy.

The Rebless Pro is a novel device currently under research, meaning that all users, including medical staff and patients, can be considered novices in its operation. As the device is primarily operated by medical staff but worn by patients, it is crucial that the UI design and user manual prioritize intuitiveness and accessibility. This ensures that both medical staff and patients can easily learn to operate the device. Consistent with previous findings, simple terms should be used in the manual to facilitate quick and accurate operation, allowing users to interact with the device intuitively and consistently [25].

From an ergonomic perspective, the medical device design should consider not only convenience and usability but also the prevention of possible injuries caused by repeated strain and the physical defects of individual people [26]. The key to CPM devices is their ability to deliver the same movement as the actual human body through human lower-extremity ROM and accurate alignment [27]. As in previous studies, marking the joint ROM and axes on the surface of the medical device was suggested in this study as an area of improvement for electrically powered orthopedic exercisers, considering the individual physical factors of each patient.

The user manual of a medical device is important not only for usability but also for identifying and understanding information [28]. The user manual explains device operation, maintenance, and troubleshooting, which can potentially affect patient safety [29]. The participants in this study also requested specific information regarding specific operating standards and clear

XSL•FO
RenderX

information about the intended patient groups, including contraindications and applicable age. Based on previous and the present studies, the development of a high-quality user manual is essential for the effective use of the device.

The high average scores (≥4) for ease of use of the UI were likely influenced by the fact that participants assessed the device during a controlled product demonstration rather than real clinical use. The findings also showed that the mean satisfaction for information identifiability and understanding of information were lower in the physical therapist group than in the physiatrist group. In rehabilitation medicine, physiatrists and physical therapists form a collaborative, multidisciplinary team; however, physiatrists are responsible for establishing the overall treatment plan, whereas physical therapists specialize in assessing motor problems and providing rehabilitation training [30-32]. The satisfaction scores may have been lower for physical therapists than for physiatrists because the physical therapists assigned scores for identification and understanding of the Rebless Pro on the basis of their own experience. These findings confirm that when developing a medical device, the intended users should be clearly defined so that they can satisfy the information identifiability and understanding requirements.

Because continuous formative evaluations of improved medical devices are important to ensure their efficacy [33], repeated formative evaluations are required for the commercialization of our high-quality electrically powered orthopedic exerciser. Improving the UI to reduce errors that may cause serious harm is recommended to enhance device efficacy and safety.

Limitations

Owing to the nature of rehabilitation devices, both medical staff who operate the device and patients who receive therapy are considered users. However, this study focused on evaluating the usability of the Rebless Pro prototype from the perspective of rehabilitation professionals, specifically examining tasks performed by medical staff, such as operating the tablet-based controller and assisting patients in wearing the device. Although we did not include patients as direct participants, we obtained valuable insights from health care professionals specializing in rehabilitation medicine, including perspectives from the patient's viewpoint during the evaluation process. Also, standardized tools such as the System Usability Scale were not used for quantitative data collection in this study. To ensure further refinement of the device through iterative formative evaluations, future studies should incorporate standardized tools to collect quantitative data.

Conclusions

If improvements are made to the safety of the electrically powered orthopedic exerciser, the identifiability and comprehensibility of information related to it, and its ease of use based on identified risk factors and areas for UI enhancement, the device can be refined to better meet user needs. This study primarily identified physiatrists and physical therapists as key user groups for evaluating the electrically powered orthopedic exerciser, given their expertise in device operation and patient safety. Future evaluations should further integrate patient feedback to ensure usability and safety from both clinical and end-user perspectives. Iterative formative evaluations involving these specific user groups, alongside the systematic implementation of necessary modifications, will facilitate a summative evaluation of the final product. These efforts are expected to contribute to the commercialization of an electrically powered orthopedic exerciser that prioritizes both safety and usability for its intended users.

Acknowledgments

This study was supported by the National Rehabilitation Center Rehabilitation Research & Development Support Program funded by the Ministry of Health and Welfare (NRCRSP-23TB01 and NRCRSP-25TB01).

Authors' Contributions

SH contributed to the planning and execution of the study and writing (original draft); HC and HK performed the overall supervision and observation of the study, project administration, and writing (review and editing).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Use scenarios for rehabilitation medical staff.
[DOCX File, 27 KB - humanfactors_v12i1e60607_app1.docx]

Multimedia Appendix 2
Question guide for focus group interviews.
[DOCX File, 26 KB - humanfactors_v12i1e60607_app2.docx]

Multimedia Appendix 3
Survey results on ease of use of the user interface.
[DOCX File, 19 KB - humanfactors_v12i1e60607_app3.docx]

XSL•FO
RenderX

Multimedia Appendix 4
Survey results on identifiability and understanding of information.
[DOCX File, 14 KB - humanfactors_v12i1e60607_app4.docx]

References

1. Perna A, Proietti L. Editorial on: musculoskeletal rehabilitation: current challenges and new perspectives. J Clin Med 2023 Jun 12;12(12):3981. [doi: 10.3390/jcm12123981] [Medline: 37373675]
2. Favre J, Cantaloube A, Jolles BM. Rehabilitation for musculoskeletal disorders: the emergence of serious games and the promise of personalized versions using artificial intelligence. J Clin Med 2023 Aug 15;12(16):5310. [doi: 10.3390/jcm12165310] [Medline: 37629350]
3. Medical Devices Act. Ministry of Food and Drug Safety (South Korea). 2024. URL: https://www.law.go.kr [accessed 2025-05-12]
4. White L. CPM machine protocol for knee replacements. The University of Texas at Tyler. 2020. URL: https://scholarworks.uttyler.edu/nursing_msn/63 [accessed 2025-05-12]
5. Salter RB, Hamilton HW, Wedge JH, et al. Clinical application of basic research on continuous passive motion for disorders and injuries of synovial joints: a preliminary report of a feasibility study. J Orthop Res 1984;1(3):325-342. [doi: 10.1002/jor.1100010313] [Medline: 6481515]
6. Yang Y, Wang J, Zhang XY, Dong L, Liu AF, Li CY. The effect of continuous passive motion in patients treated with total knee arthroplasty for osteoarthritis: a systematic review and meta-analysis of randomized controlled trials. Am J Phys Med Rehabil 2021 Dec 1;100(12):1160-1169. [doi: 10.1097/PHM.0000000000001718] [Medline: 33605578]
7. Jia Z, Zhang Y, Zhang W, Xu C, Liu W. Efficacy and safety of continuous passive motion and physical therapy in recovery from knee arthroplasty: a systematic review and meta-analysis. J Orthop Surg Res 2024 Jan 13;19(1):68. [doi: 10.1186/s13018-024-04536-y] [Medline: 38218933]
8. Wang X, Liu Q, Tan J, et al. Optimizing continuous passive motion duration following arthroscopic release of elbow contracture: a retrospective study. BMC Musculoskelet Disord 2024 Nov 26;25(1):965. [doi: 10.1186/s12891-024-08108-3] [Medline: 39593033]
9. Lee BC, Moon CW, Choi WS, et al. Clinical evaluation of usefulness and effectiveness of sitting-type continuous passive motion machines in patients with total knee arthroplasty: a dual-center randomized controlled trial. BMC Musculoskelet Disord 2024 Dec 19;25(1):1039. [doi: 10.1186/s12891-024-08062-0] [Medline: 39702172]
10. Hamdan WK, Ali SM, Saleh ZH, Awad DM, Tareq FZ, Radi TM. Design and fabricating of continuous passive motion (CPM) machine for physical knee pain therapy. In: 4th International Scientific Conference of Alkafeel University (ISCKU 2022): AIP Publishing; 2023. [doi: 10.1063/5.0182136]
11. Liao CD, Huang YC, Chiu YS, Liou TH. Effect of body mass index on knee function outcomes following continuous passive motion in patients with osteoarthritis after total knee replacement: a retrospective study. Physiotherapy 2017 Sep;103(3):266-275. [doi: 10.1016/j.physio.2016.04.003] [Medline: 27647443]
12. Miclăuş T, Valla V, Koukoura A, et al. Impact of design on medical device safety. Ther Innov Regul Sci 2020 Jul;54(4):839-849. [doi: 10.1007/s43441-019-00022-4] [Medline: 32557299]
13. Santiago GB, Aiquipa WA, Santiago MB, et al. Electromagnetic compatibility tests and informed consent, in the development of research on biomedical equipment design. In: Proceedings of the 2019 5th International Conference on Communication and Information Processing, ICCIP 2019: Association for Computing Machinery; 2019:134-138. [doi: 10.1145/3369985.3371674]
14. Whitworth E, Lewis JA, Boian R, et al. Formative evaluation of a virtual reality telerehabilitation system for the lower extremity. In: Proceedings of the 2nd International Workshop on Virtual Rehabilitation (IWVR2003): IEEE; 2003:21-22.
15. IEC 62366-1:2015: Medical Devices; Part 1: application of usability engineering to medical devices. International Organization for Standardization. URL: https://www.iso.org/standard/63179.html [accessed 2025-05-12]
16. Ravizza D, Sternini F, Lantada A, et al. Techniques for usability risk assessment during medical device design. In: Proceedings of the 12th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2019): Science and Technology Publications:207-214. [doi: 10.5220/0007483102070214]
17. Krueger RA, Casey MA. Focus Groups: A Practical Guide for Applied Research: Sage Publications; 2015.
18. Porras-Martínez E, López-Matas H, Lobo-Prat J, et al. Safety, performance and user satisfaction study of the new ABLE lower-limb exoskeleton for individuals with SCI: the result of a co-creation process with clinicians and patients. Research Square. Preprint posted online on Jul 28, 2023. [doi: 10.21203/rs.3.rs-3199890/v1]
19. Hinricher N, König S, Schröer C, Backhaus C. Effects of virtual reality and test environment on user experience, usability, and mental workload in the evaluation of a blood pressure monitor. Front Virtual Real 2023 Jun 12;4:1151190. [doi: 10.3389/frvir.2023.1151190]
20. Creswell JW, Poth CN. Qualitative Inquiry and Research Design: Choosing Among Five Approaches: SAGE Publications; 1998.

XSL•FO
RenderX

21. Park SJ, Moon HS, Lee EW, Youn JH, Lee HJ. A study on the experience of art therapy about emotional labor stress for nurse managers: focusing on focus group interviews. J Arts Psychother 2021 Dec 31;17(4):33-61. [doi: 10.32451/KJOAPS.2021.17.4.033]
22. Kim DW, Woo JH, Jang WH, Lee YH, Kim J. A qualitative study on the user experience of manual/electric hybrid wheelchairs in leisure activity participation: focusing on focus group interview. J Korean Soc Occup Ther 2021 Dec 30;29(4):165-179. [doi: 10.14519/kjot.2021.29.4.11]
23. Baek S, Park SW. Parallel safety control system for upper-limb rehabilitation robot considering user safety. Trans Korean Inst Electr Eng P 2022 Sep 30;71(3):175-183. [doi: 10.5370/KIEEP.2022.71.3.175]
24. Ahn SJ, Yuk SW, Yu WJ, Moon IH. A study on standard for safety requirements for care robots. Resko 2022 Nov 30;16(4):190-195. [doi: 10.21288/resko.2022.16.4.190]
25. Kim AJ, Kweon OS, Lee BJ. An IEC 62366-based case study of a user interface design process for a rehabilitation device. University of Cincinnati. 2017 Oct 31. URL: https://scholar.uc.edu/concern/articles/th83kz33c?locale=en [accessed 2025-05-06]
26. Shin T, Chung D. A study on usability of rehabilitation medical equipment ; focused on stationary training equipment. Ind Des Res 2017 Mar 31;39:85-94. [doi: 10.37254/ids.2017.03.39.09.85]
27. Kim KJ, Kang MS, Choi YS, Han J, Han C. Conceptualization of an exoskeleton continuous passive motion (CPM) device using a link structure. IEEE Int Conf Rehabil Robot 2011;2011:5975494. [doi: 10.1109/ICORR.2011.5975494] [Medline: 22275691]
28. Andrade FO, Nascimento LN, Wood GA, et al. Applying heuristic evaluation on medical devices user manuals. In: Jaffray DA, editor. World Congress on Medical Physics and Biomedical Engineering: Springer International Publishing:1515.
29. Abbate AJ, Throckmorton AL, Bass EJ. Using formal task analytic models to support user manual development: an LVAD case study. Proc Int Symp Hum Factors Ergon Healthc 2015 Jun;4(1):114-117. [doi: 10.1177/2327857915041039]
30. Tederko P, Mycielski J, Angerova Y, et al. Role of diversity in assembling of rehabilitation teams in Central Europe. Eur J Phys Rehabil Med 2020 Apr;56(2):131-141. [doi: 10.23736/S1973-9087.20.05904-3] [Medline: 31939266]
31. Momsen AM, Rasmussen JO, Nielsen CV, Iversen MD, Lund H. Multidisciplinary team care in rehabilitation: an overview of reviews. J Rehabil Med 2012 Nov;44(11):901-912. [doi: 10.2340/16501977-1040] [Medline: 23026978]
32. Singh R, Küçükdeveci AA, Grabljevec K, Gray A. The role of interdisciplinary teams in physical and rehabilitation medicine. J Rehabil Med 2018 Aug 22;50(8):673-678. [doi: 10.2340/16501977-2364] [Medline: 29944171]
33. Markiewicz K. Health technology assessment of medical devices during development [thesis]. University of Twente. 2017. URL: https://research.utwente.nl/en/publications/health-technology-assessment-of-medical-devices-during-developmen [accessed 2025-05-12]

Abbreviations

CPM: continuous passive motion
FGI: focus group interview
IEC: International Electrotechnical Commission
MSK: musculoskeletal
NRC: National Rehabilitation Center
ROM: range of motion
UI: user interface

Edited by C Jacob; submitted 27.05.24; peer-reviewed by JH Lee, JL Candiotti; revised version received 18.04.25; accepted 22.04.25; published 30.05.25.

Please cite as:
Hong S, Choi H, Kweon H
Usability Evaluation of an Electrically Powered Orthopedic Exerciser: Focus Group Interview and Satisfaction Survey Study
JMIR Hum Factors 2025;12:e60607
URL: https://humanfactors.jmir.org/2025/1/e60607
doi:10.2196/60607

© Seojin Hong, Hyun Choi, Hyosun Kweon. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 30.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information,

XSL•FO
RenderX

a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Factors Shaping Public Perceptions of a Range of Robotic Technologies in Surgery: Cross-Sectional Web-Based Survey

Sarek Shen[1], MD, MS; Deborah Xie[1], MD; Andy Ding[1], MD, MSE; Lisa Zhang[2], MD; Francis Creighton[1], MD

[1]Department of Otolaryngology-Head and Neck Surgery, Johns Hopkins School of Medicine, 601 N Caroline Dr, Baltimore, MD, United States
[2]Department of Otolaryngology-Head and Neck Surgery, Ohio State University School of Medicine, Columbus, OH, United States

Corresponding Author:
Sarek Shen, MD, MS
Department of Otolaryngology-Head and Neck Surgery, Johns Hopkins School of Medicine, 601 N Caroline Dr, Baltimore, MD, United States

Abstract

Background: Within the surgical field, there has been an evolution in the application of robotic technology. Fully automatic robotic systems and augmented visualization tools are being introduced and may eventually replace existing surgical extenders such as the da Vinci surgical system. The literature on public perception of robotic surgery is growing, though specific drivers of these attitudes remain under investigation.

Objective: The aim of this study is to investigate the underlying motivators of public perceptions toward robotic surgeries with varying levels of autonomy through a formal technology acceptance model.

Methods: An online survey was distributed via the Amazon Mechanical Turk platform. Survey participants were provided definitions of a continuum of robotic technologies: robotic surgical extenders (technology without independent actions), semiautonomous robotic surgery (technology that provides guidance to the surgeon and requires surgeon input), and fully autonomous robotic surgery (technology that performs tasks autonomously without direct human interaction). The survey assessed overall attitudes toward each application of robotic technology in surgery and included questions delineating specific receptivity based on (1) perceived usefulness, (2) social risk, (3) time risk, (4) personal risk, and (5) reliability. A technology acceptance model was built to identify associations between these factors and overall attitudes toward robotic and semiautonomous surgeries.

Results: A total of 1221 survey responses were recorded (mean age 38, SD 12 y; females: n=635, 52%). Individuals were more willing to accept robotic surgical extenders and semiautonomous robotic surgery compared to autonomous robotic surgery. Higher levels of education and better self-reported health were correlated with more positive attitudes toward autonomous robotic surgery. Perceptions of these technologies were not associated with age, gender, or income. Overall, attitudes toward robotic technologies in surgery were driven by views on the reliability, safety, and efficiency of the procedures. There was less concern regarding time risk and social risk associated with robotic and semirobotic surgeries.

Conclusions: The public is more accepting of semiautonomous surgery and surgical extenders than fully autonomous surgery. General perceptions of the reliability, safety, and efficiency of these technologies drive variations in attitude. Time and social risk do not appear to have a significant impact on receptivity. Understanding these perspectives can help guide education within an advancing surgical field.

(JMIR Hum Factors 2025;12:e64224) doi:10.2196/64224

KEYWORDS

perception; robotic surgery; semiautonomous robotic surgery; surgical extenders; safety; reliability; structural equation model; technology; acceptance; model; public; perceptions; cross-sectional; web-based; assessment; robot-assisted; surgery; robot; robotic technology; medical care

Introduction

Applications of robotic and artificial intelligence technologies in surgery are rapidly developing [1-6]. The predominant robotic surgical technology currently available is surgical extenders such as the da Vinci (Intuitive)—these robotic technologies move only in response to surgeon input and provide nonaugmented visualization to the surgeon. Following surgical extenders, several semiautonomous systems have emerged in various surgical fields. At present, these technologies generally involve the incorporation of image guidance or augmented reality to identify anatomic structures and navigate an operative field [7,8]. Within orthopedics, semiautonomous systems are being used for improved hardware placement during spine surgery and hip surgery [9-13]. Looking further, researchers have developed fully autonomous robotic systems capable of performing tasks such as cochlear implantation [14-16] and

XSL•FO
RenderX

laparoscopic bowel anastomosis [17,18]. Although not yet commercially available, these systems may emerge as standard complements to the surgical workflow.

Despite the promising developments within robotic surgery technologies, there remain social, financial, and regulatory barriers to these systems becoming commonplace. One obstacle is a nascent understanding of the public's perception and tolerance of these technologies. Currently available robotic systems have been shown to improve operating room efficiency and surgical planning [19-23], and the public's attitudes toward these technologies continue to evolve. There is a growing body of research investigating the relationship between sociodemographic factors and views on autonomous technologies within medicine [24,25]. However, the underlying drivers of these perceptions are not fully explored. Given the growing prevalence of autonomous and semiautonomous technologies in our day-to-day lives, we sought to investigate the public's perspectives of the spectrum of robotic systems in surgery. Using a technology acceptance model, we investigate the public's attitudes, and drivers of those attitudes, toward a range of robotic surgery technologies.

Methods

Ethical Considerations

This study was reviewed and approved by the Johns Hopkins institutional review board (IRB00267594). A survey was designed on Qualtrics (Qualtrics LLC) and distributed through Amazon Mechanical Turk (MTurk) in October 2020 (Multimedia Appendix 1). MTurk is a crowdsourcing platform that allows individuals to complete assignments online in an anonymous fashion. All study data were deidentified after collection was complete. A US $0.25 compensation was provided for those who completed the survey and correctly answered all attention check questions.

Participant Recruitment

Respondents were included if they were over 18 years of age and a US resident. The survey included multiple attention check questions to ensure active engagement (eg, "Please select the number 3"). Respondents were excluded if their primary language was not English or if they failed to answer attention check questions correctly. We do not have data on the number of people who viewed but elected not to partake in the survey; therefore, we cannot calculate traditional participation rates. The completion rate of those who started the survey was 97.2%.

Survey Instrument

The survey instrument was designed to assess perceptions and attitudes toward applications of robotic technology in surgery. Each survey started by providing the participant with definitions of 3 levels of robotic technology in surgery:

- Robotic surgical extenders: robot controlled via surgeon, does not perform actions independently without the input of the surgeon.
- Semiautonomous robotic surgery: provides guidance to surgeon based on a patient's information. The robot does not perform actions independently without surgeon input.
- Autonomous robotic surgery: operates and performs tasks on its own, without direct interaction with the surgeon.

Questions in the survey were adopted from previously validated surveys assessing acceptance of technology [26-28]. Respondents rated statements on a 7-point Likert scale from "strongly disagree" to "strongly agree." Higher scores indicated a more positive or favorable opinion toward the technology being assessed. A respondent's overall attitude (A) was interrogated directly via questions in the survey. Further questions were designed to assess one of the following latent variables:

- Perceived usefulness (PU): The ability for robotic technology to improve the care that one receives in surgery.
- Social risk (SR): The possibility that using robotic technology in surgery would lead to disapproval from one's friends, family, or social group.
- Time risk (TR): The loss of time and inconvenience that would result from the use of robotic technology in surgery.
- Personal risk (PR): The physical safety of robotic technology in surgery.
- Reliability (R): The ability of robotic technology in surgery to work reliably over time.

Each latent variable was evaluated by multiple questions, abbreviated by the question number and the latent variable assessed (eg, PU1 and PU2 were the two questions assessing perceived usefulness). All questions were completed for each application of robotic technology in surgery (ie, each latent variable was assessed for robotic surgical extenders, semiautonomous robotic surgery, and autonomous robotic surgery). The relationship between the latent variables and the overall attitude was determined with structural equation models (SEMs) based on the technology acceptance model [29] and theory of planned behavior [30,31]. Participants were also asked to rate their personal comfort and trust with technology, as well as their own health on a 5-point Likert scale. Demographic data of each respondent were also collected. The survey instrument is included in the Multimedia Appendix 1.

Data Analysis

Descriptive analyses, including frequency distribution and central tendency, were performed to illustrate the demographic characteristics and participants' questionnaire responses. Participant education and self-perceived health levels were further stratified into quartiles. ANOVA, *t* tests, and Pearson correlation coefficient were used to compare participants' attitudes and intentions among the surgical extender types, as well as differences in viewpoints among demographic variables.

Confirmatory factor analysis was used to measure the relationship between the measured variables (question responses) and the latent variables (key factors) through factor loading. Multiple SEMs were built to identify associations between the latent variables and overall attitude. Standard regression coefficients calculated by the model represent the ability of our survey questions to capture data on its latent variable. The assumption of multivariate normality was assessed and confirmed with skewness and kurtosis tests. Cronbach α was computed for each key factor to determine the internal

XSL•FO
RenderX

consistency of the research construct. Values of >0.7 indicate high reliability, between 0.5 and 0.7 indicate moderate reliability, and <0.5 indicate low reliability [32].

Root mean square error of approximation and X^2 tests were used to measure the discrepancy between the model-based and observed correlation matrices. The final structural model was built and standard regression coefficients were calculated. Data analyses were performed with R (version 4.1.2; R Foundation for Statistical Computing), with SEM calculations performed using the *lavaan* package. All α levels and error probabilities were set at the standard .05 level.

Results

Background Characteristics

A total of 1221 surveys were included for analysis. The mean age of participants was 37.7 (SD 12.1) years. Females made up 51.7% (n=631) of respondents, and the majority (n=940, 77%) were White. Full demographics are listed in Table 1.

Table . Participant sociodemographics and personal health ratings.

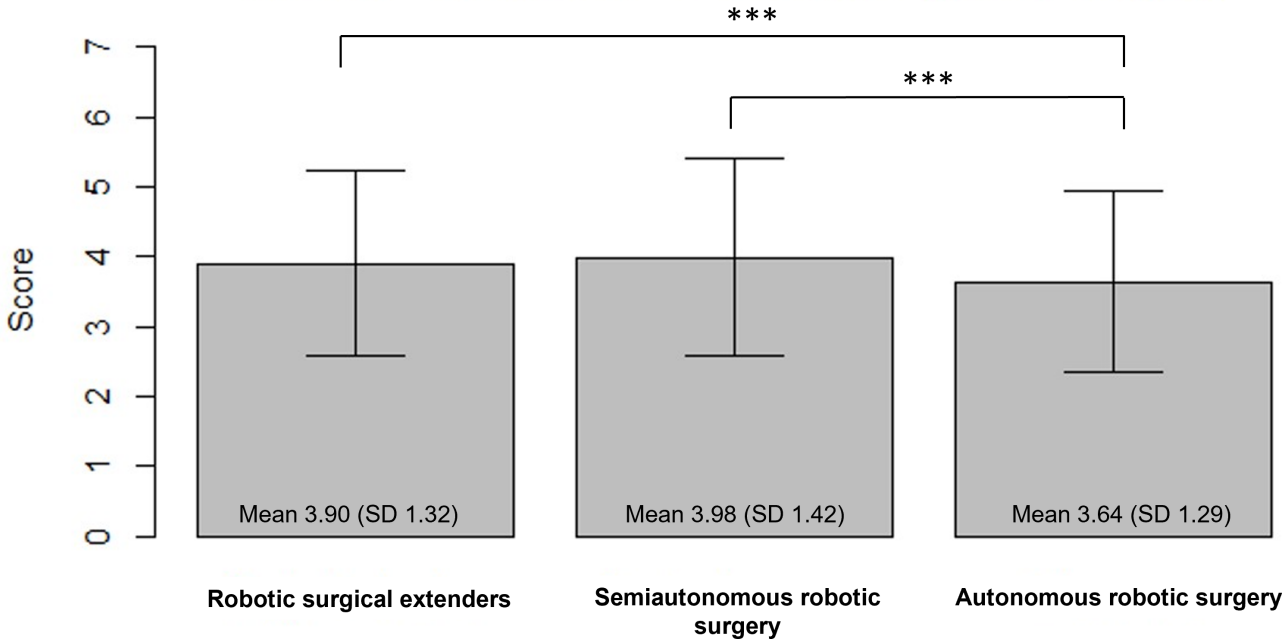
Characteristics	Values
Age (years), mean (SD)	37.7 (12.1)
Gender, n (%)	
Male	580 (47.5)
Female	631 (51.7)
Nonbinary	10 (0.8)
Race/ethnicity, n (%)	
White	940 (77)
African/African American	124 (10.2)
American Indian/Alaska Native	37 (3)
Asian	91 (7.5)
Other	29 (2.4)
Highest level of education, n (%)	
High school graduate or less	78 (6.4)
Some college	226 (18.5)
College degree	671 (55)
Advanced degree (Master's, PhD, MD, etc)	246 (20.1)
Personal rating of health, n (%)	
Excellent	266 (21.8)
Very good	533 (43.7)
Good	334 (27.4)
Fair	73 (6)
Poor	15 (1.2)

Sociodemographic Associations

The surveyed cohort expressed more positive attitudes toward robotic surgical extenders and semiautonomous robotic surgery compared to autonomous robotic surgery (Figure 1). There was no significant correlation between age and attitude toward robotic surgical extenders (r=0.024, 95% CI −0.032 to 0.080), semiautonomous robotic surgery (r=0.017, 95% CI −0.039 to 0.073), or autonomous robotic surgery (r=−0.0124, 95% CI −0.069 to 0.044). Similarly, there was no significant correlation between gender and race relative to attitude toward all applications of robotic technologies in surgery. There was a positive correlation between comfort with technology and overall attitudes for robotic surgical extenders (r=0.090, 95% CI 0.034 to 0.145), semiautonomous robotic surgery (r=0.088, 95% CI 0.032 to 0.143), and autonomous robotic surgery (r=0.118, 95% CI 0.062 to 0.173).

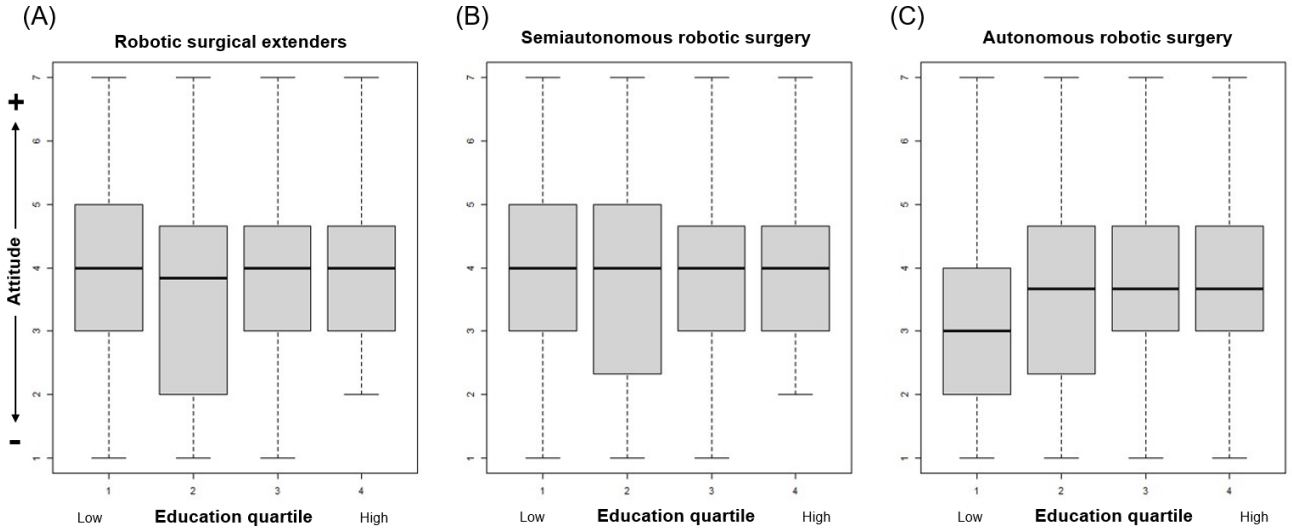
XSL•FO
RenderX

Figure 1. Comparison of overall attitudes toward different applications of robotic technology in surgery. ***P<.001.



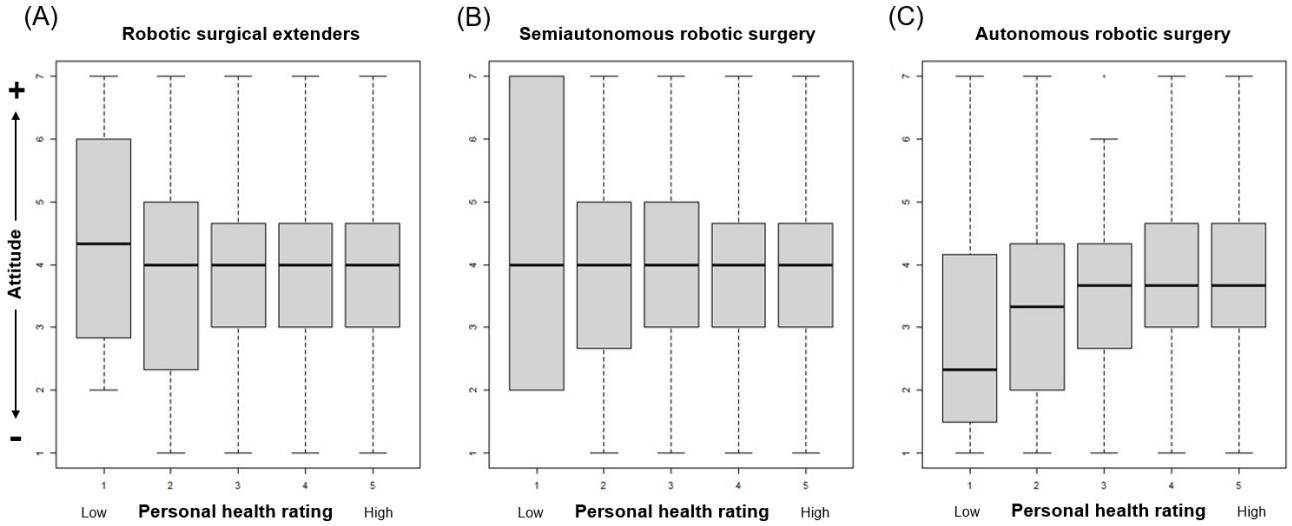
Attitudes toward robotic surgical extenders (P=.24) and semiautonomous robotic surgery (P=.57) did not vary based on level of education. However, respondents with higher levels of education did have more positive attitudes toward autonomous robotic surgery (P<.001, Figure 2). Similarly, there was no significant difference in attitudes toward robotic surgical extenders (P=.60) and semiautonomous robotic surgery (P=.29) across all levels of personal ratings of health. Those with higher ratings of health held more favorable attitudes toward autonomous robotic surgery (P=.003, Figure 3). There was no significant difference in attitudes toward all applications of robotic technology between individuals with or without a surgical history (all P>.05).

Figure 2. Relationships between education and overall attitudes toward different applications of robotic technology in surgery.



XSL•FO
RenderX

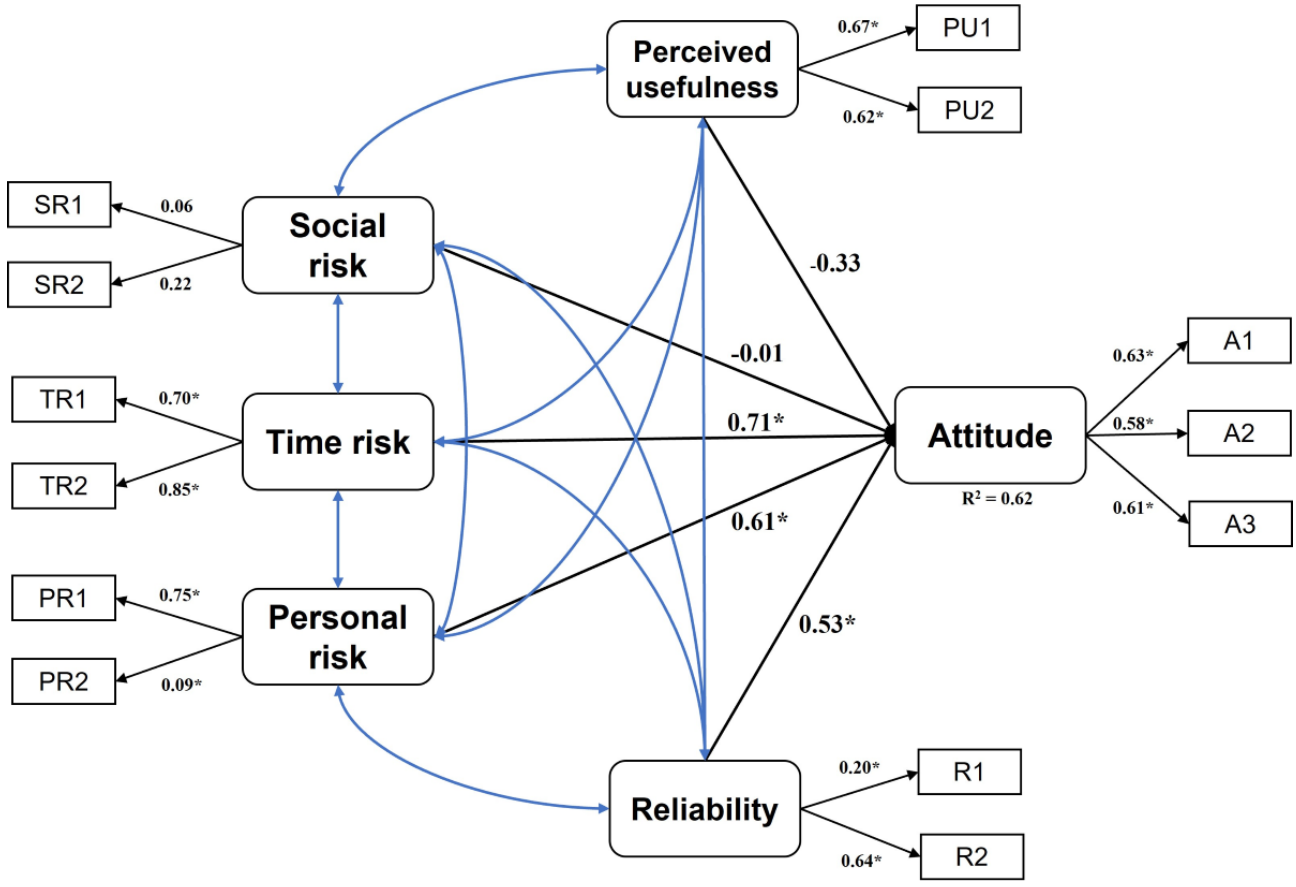
Figure 3. Relationships between personal ratings of health and overall attitudes toward different applications of robotic technology in surgery.



Technology Acceptance Model

An SEM was built to explore the relationships between the latent variables (PU, SR, TR, PR, and R) and overall attitude (Figure 4). The correlation between survey questions for each latent variable was calculated using Cronbach α (Table 2). The low coefficient magnitude (SR1: 0.06, SR2: 0.22) combined with the lack of statistical significance suggest that the concept of "social risk" was poorly characterized by our survey questions. Attitudes toward robotic technologies in surgery were predicted by TR (β=0.71, P<.05), PR (β=0.61, P<.05), and R (β=0.53, P<.05). Together, these variables explain 62% of the variance of attitudes (R^2=0.62, coefficient of determination). SR (β=−0.01, P>.05) and PU (β=−0.33, P>.05) did not significantly affect attitude.

Figure 4. Results of structural equation modeling analysis. A: attitude; PR: personal risk; PU: perceived usefulness; R: reliability; SR: social risk; TR: time risk. *P<.05.



XSL•FO
RenderX

Table . Cronbach α of the survey questions assessing the included latent variables (perceived usefulness, social risk, time risk, personal risk, reliability) and overall attitude.

	Robotic surgical extenders	Semiautonomous	Autonomous	Total
Perceived usefulness	0.81	0.46	0.26	0.59
Social risk	0.01	0.09	0.15	0.02
Time risk	0.76	0.8	0.68	0.75
Personal risk	0.01	0.16	0.34	0.22
Reliability	0.24	0.28	0.38	0.26
Attitude	0.46	0.52	0.45	0.63

Discussion

Principal Findings

Implementation of robotic technology in the operating room setting has the potential to improve the delivery of health care. From a surgeon's perspective, the technical benefits of robotic surgery include improved visualization, better maneuverability, and favorable ergonomics [33,34]. As automated technologies become increasingly integrated into medicine, there has been a growing body of literature on how the public perceives the continuum of robotic procedures [35]. Palmisciano et al [19] found that patients are more favorable toward the use of autonomous technology in noninvasive applications, such as operative planning and preoperative imaging, compared to surgical procedures. In our study, we similarly found that the public is more open to surgical extenders and semiautonomous technologies than fully autonomous systems. Based on our technology acceptance model, the differences in these attitudes appear to be driven primarily by concerns involving safety and reliability, more so than social pressures or apprehension about the usefulness of the technologies. These findings may help guide efforts in patient education as these robotic systems become more commonplace in surgery.

Sociodemographic Correlations

Our study found that individuals with higher levels of education, as well as those with higher personal ratings of health, tend to be more receptive toward fully autonomous surgeries. Education has been well-studied within the technology acceptance model; numerous investigations have demonstrated increased acceptance and utilization of technology among those with secondary or tertiary degrees [36-38]. Werner et al [39] have postulated that this is associated with a relative increase in access to and familiarity with new technologies, while Chen et al [40] have noted that these cohorts are more self-efficacious and may have less technology-related anxiety. From a health perspective, we found that participants with poorer self-ratings of health were associated with a more negative attitude toward fully robotic surgeries. This finding is congruent with studies investigating patterns of use of mobile health devices; for example, Chandrasekaran et al [41] showed that healthier patients are more likely to adopt new health care–associated technology [42]. However, these associations are complex and likely depend on the specific technology in question. Further study is necessary to elucidate the complicated relationship between the perceived health and technology acceptance of surgical patients.

Interestingly, we did not find a correlation between age and attitude toward fully autonomous surgeries. Recent studies investigating technology acceptance within aging populations, such as in Japan and South Korea, have similarly demonstrated a positive perception of robotic technology in an older cohort [43,44]. This is in part driven by the growing ubiquity of semiautonomous technology within day-to-day life, which increases familiarity, builds trust, and changes beliefs and attitudes toward new technology [45].

Drivers of Public Perception

In investigating the drivers of overall patient attitude toward robotic surgeries, we found that risk to patient, reliability of technology, and perceived time risk were the primary concerns. There was less consideration for usefulness or social risk. These results are similar to those found in the autonomous automobile industry [46,47] and for assistive technology in patients undergoing physical rehabilitation [48]. Other studies have also found that apprehension toward adoption of robotic surgeries is driven by fears of equipment error or malfunction [34,49]. The evolution of attitudes toward autonomous driving may be particularly informative. Studies in this space suggest that increased system transparency, which refers to improved user understanding of the operational predictability, is positively associated with acceptance of the technology [46]. Within surgery, an increase in system transparency may begin with addressing misconceptions about these robotic surgical technologies. The colloquial definition and usage of the term "robot" tends to imply a programmable machine that can replicate the actions of a person. For those not familiar with robotic technology in health care, the nuances of its uses may not be readily obvious. In 2016, Boys et al [34] published a survey-based study specifically investigating the perceptions of robotic surgical extenders; interestingly, 21% of their 747 respondents believed that robotic surgical extenders had autonomous function during surgery. To address these misconceptions, the creation of accessible patient education materials, specifically focusing on reliability, safety, and efficiency, may be the best way to impact overall attitudes toward robotic technologies in surgery.

Limitations

There are several limitations to this study. First, while questions in our instrument had been previously validated [26-28], the

XSL•FO
RenderX

consistency of responses related to the latent variables, specifically the social risk, was lower than expected. This may reflect inherent difficulty in summarizing a broad variable within 2 - 3 questions, particularly within health care. However, the drivers of patient attitude that our group identified are similar to those implicated in other areas with a growing autonomous technological presence, including online banking and self-driving cars [26,27,50]. This concordance is reassuring and suggestive of a similar pattern of technological acceptance between fields. Second, we understand that there may be a slight selection bias in the population sampled, as we used an online survey platform, which may have a younger and more technologically savvy population compared to the prototypical patient. However, it is notable that the mean age of patients undergoing the most common robotic procedure in the United States is around 41.5 years, consistent with our survey population. The racial and ethnic breakdown of patients undergoing robotic surgeries is similar to our survey cohort [51]. Additionally, prior studies have also shown that the Amazon MTurk survey results are comparable to those collected by more traditional methods [52-54]. Third, our survey was not able to capture additional socioeconomic variables, such as income or employment status, which may be hidden confounders of our findings. To address some of these limitations, future studies will include an in-person investigation into treatment decision-making and patient preference within robotic surgeries.

Conclusion

Current public perception toward surgical extenders and semiautonomous robotic surgery is more positive than that toward autonomous robotic surgery. Overall attitudes appear to be driven by concerns about the reliability, safety, and efficiency of these technologies. As robotic technologies emerge and continue to evolve within surgery, these findings can inform efforts in directed education. Comprehensive preoperative discussions targeted at the reliability and efficacy—rather than the usefulness or social risk—of these technologies may be more impactful. However, further work is needed in exploring potential obstacles to implementation of robotic systems within surgical practices.

Funding

This work was funded in part by the National Institute on Deafness and Other Communication Disorders (NIDCD) grant number 5T32DC000027-33.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Survey instrument
[PDF File, 125 KB - humanfactors_v12i1e64224_app1.pdf]

References

1. Bankar GR, Keoliya A. Robot-assisted surgery in gynecology. Cureus 2022 Sep;14(9):e29190. [doi: 10.7759/cureus.29190] [Medline: 36259016]
2. Lira RB, Kowalski LP. Robotic head and neck surgery: beyond TORS. Curr Oncol Rep 2020 Jul 9;22(9):88. [doi: 10.1007/s11912-020-00950-7] [Medline: 32643128]
3. Rassweiler JJ, Autorino R, Klein J, et al. Future of robotic surgery in urology. BJU Int 2017 Dec;120(6):822-841. [doi: 10.1111/bju.13851] [Medline: 28319324]
4. Shlobin NA, Huang J, Wu C. Learning curves in robotic neurosurgery: a systematic review. Neurosurg Rev 2022 Dec 12;46(1):14. [doi: 10.1007/s10143-022-01908-y] [Medline: 36504244]
5. Walshaw J, Huo B, McClean A, et al. Innovation in gastrointestinal surgery: the evolution of minimally invasive surgery-a narrative review. Front Surg 2023;10:1193486. [doi: 10.3389/fsurg.2023.1193486] [Medline: 37288133]
6. Yamout T, Orosz LD, Good CR, Jazini E, Allen B, Gum JL. Technological advances in spine surgery: navigation, robotics, and augmented reality. Orthop Clin North Am 2023 Apr;54(2):237-246. [doi: 10.1016/j.ocl.2022.11.008] [Medline: 36894295]
7. Tamaki A, Rocco JW, Ozer E. The future of robotic surgery in otolaryngology - head and neck surgery. Oral Oncol 2020 Feb;101:104510. [doi: 10.1016/j.oraloncology.2019.104510] [Medline: 31841882]
8. Chan JYK, Holsinger FC, Liu S, Sorger JM, Azizian M, Tsang RKY. Augmented reality for image guidance in transoral robotic surgery. J Robot Surg 2020 Aug;14(4):579-583. [doi: 10.1007/s11701-019-01030-0] [Medline: 31555957]
9. Lieberman IH, Kisinde S, Hesselbacher S. Robotic-assisted pedicle screw placement during spine surgery. JBJS Essent Surg Tech 2020;10(2):e0020. [doi: 10.2106/JBJS.ST.19.00020] [Medline: 32944411]

XSL•FO
RenderX

10. Hu X, Scharschmidt TJ, Ohnmeiss DD, Lieberman IH. Robotic assisted surgeries for the treatment of spine tumors. Int J Spine Surg 2015;9:1. [doi: 10.14444/2001] [Medline: 25709885]
11. Kochanski RB, Lombardi JM, Laratta JL, Lehman RA, O'Toole JE. Image-guided navigation and robotics in spine surgery. Neurosurgery 2019 Jun 1;84(6):1179-1189. [doi: 10.1093/neuros/nyy630] [Medline: 30615160]
12. Loomans L, Leirs G, Vandenneucker H. Operating room efficiency after the implementation of MAKO robotic-assisted total knee arthroplasty. Arch Orthop Trauma Surg 2023 Sep;143(9):5501-5506. [doi: 10.1007/s00402-023-04834-w] [Medline: 36943503]
13. Perazzini P, Trevisan M, Sembenini P, et al. The Mako robotic arm-assisted total hip arthroplasty using direct anterior approach: surgical technique, skills and pitfals. Acta Biomed 2020 May 30;91(4-S):21-30. [doi: 10.23750/abm.v91i4-S.9659] [Medline: 32555073]
14. Caversaccio M, Wimmer W, Anso J, et al. Robotic middle ear access for cochlear implantation: first in man. PLoS One 2019;14(8):e0220543. [doi: 10.1371/journal.pone.0220543] [Medline: 31374092]
15. Labadie RF, Noble JH, Dawant BM, Balachandran R, Majdani O, Fitzpatrick JM. Clinical validation of percutaneous cochlear implant surgery: initial report. Laryngoscope 2008 Jun;118(6):1031-1039. [doi: 10.1097/MLG.0b013e31816b309e] [Medline: 18401279]
16. Warren FM, Balachandran R, Fitzpatrick JM, Labadie RF. Percutaneous cochlear access using bone-mounted, customized drill guides: demonstration of concept in vitro. Otol Neurotol 2007 Apr;28(3):325-329. [doi: 10.1097/01.mao.0000253287.86737.2e] [Medline: 17414037]
17. Saeidi H, Le HND, Opfermann JD, et al. Autonomous laparoscopic robotic suturing with a novel actuated suturing tool and 3D endoscope. IEEE Int Conf Robot Autom 2019 May;2019:1541-1547. [doi: 10.1109/icra.2019.8794306] [Medline: 33628614]
18. Saeidi H, Opfermann JD, Kam M, et al. Autonomous robotic laparoscopic surgery for intestinal anastomosis. Sci Robot 2022 Jan 26;7(62):eabj2908. [doi: 10.1126/scirobotics.abj2908] [Medline: 35080901]
19. Palmisciano P, Jamjoom AAB, Taylor D, Stoyanov D, Marcus HJ. Attitudes of patients and their relatives toward artificial intelligence in neurosurgery. World Neurosurg 2020 Jun;138:e627-e633. [doi: 10.1016/j.wneu.2020.03.029] [Medline: 32179185]
20. Titano JJ, Badgeley M, Schefflein J, et al. Automated deep-neural-network surveillance of cranial images for acute neurologic events. Nat Med 2018 Sep;24(9):1337-1341. [doi: 10.1038/s41591-018-0147-y] [Medline: 30104767]
21. Bennett CC, Hauser K. Artificial intelligence framework for simulating clinical decision-making: a Markov decision process approach. Artif Intell Med 2013 Jan;57(1):9-19. [doi: 10.1016/j.artmed.2012.12.003] [Medline: 23287490]
22. Kim JS, Merrill RK, Arvind V, et al. Examining the ability of artificial neural networks machine learning models to accurately predict complications following posterior lumbar spine fusion. Spine (Phila Pa 1986) 2018;43(12):853-860. [doi: 10.1097/BRS.0000000000002442]
23. Hashimoto DA, Rosman G, Rus D, Meireles OR. Artificial intelligence in surgery: promises and perils. Ann Surg 2018 Jul;268(1):70-76. [doi: 10.1097/SLA.0000000000002693] [Medline: 29389679]
24. Torrent-Sellens J, Jiménez-Zarco AI, Saigí-Rubió F. Do people trust in robot-assisted surgery? Evidence from Europe. Int J Environ Res Public Health 2021 Nov 28;18(23):12519. [doi: 10.3390/ijerph182312519] [Medline: 34886244]
25. McDonnell C, Devine M, Kavanagh D. The general public's perception of robotic surgery - a scoping review. Surgeon 2025 Apr;23(2):e49-e62. [doi: 10.1016/j.surge.2024.12.001] [Medline: 39658498]
26. Lee MC. Factors influencing the adoption of internet banking: an integration of TAM and TPB with perceived risk and perceived benefit. Electron Commer Res Appl 2009 May;8(3):130-141. [doi: 10.1016/j.elerap.2008.11.006]
27. Lee C, Ward C, Raue M, D'Ambrosio L, Coughlin JF. Age differences in acceptance of self-driving cars: a survey of perceptions and attitudes. In: Zhou J, Salvendy G, editors. Human Aspects of IT for the Aged Population Aging, Design and User Experience. ITAP 2017. Lecture Notes in Computer Science. [doi: 10.1007/978-3-319-58530-7_1]
28. Kim J, Park HA. Development of a health information technology acceptance model using consumers' health behavior intention. J Med Internet Res 2012 Oct 1;14(5):e133. [doi: 10.2196/jmir.2143] [Medline: 23026508]
29. Fishbein M, Ajzen I. Belief, Intention and Behavior: An Introduction to Theory and Research: Longman Higher Education; 1975.
30. Ajzen I. Perceived behavioral control, self‐efficacy, locus of control, and the theory of planned behavior. J Applied Social Pyschol 2002 Apr;32(4):665-683. [doi: 10.1111/j.1559-1816.2002.tb00236.x]
31. Ajzen I. The theory of planned behavior. Organ Behav Hum Decis Process 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
32. Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Educ 2011 Jun 27;2:53-55. [doi: 10.5116/ijme.4dfb.8dfd] [Medline: 28029643]
33. Lee N. Robotic surgery: where are we now? Lancet 2014 Oct 18;384(9952):1417. [doi: 10.1016/S0140-6736(14)61851-1] [Medline: 25390313]
34. Boys JA, Alicuben ET, DeMeester MJ, et al. Public perceptions on robotic surgery, hospitals with robots, and surgeons that use them. Surg Endosc 2016 Apr;30(4):1310-1316. [doi: 10.1007/s00464-015-4368-6] [Medline: 26173543]

XSL•FO
RenderX

35. Muaddi H, Zhao X, Leonardelli GJ, et al. Fear of innovation: public's perception of robotic surgery. Surg Endosc 2022 Aug;36(8):6076-6083. [doi: 10.1007/s00464-022-09011-5] [Medline: 35059838]
36. McMurray J, Strudwick G, Forchuk C, et al. The importance of trust in the adoption and use of intelligent assistive technology by older adults to support aging in place: scoping review protocol. JMIR Res Protoc 2017 Nov 2;6(11):e218. [doi: 10.2196/resprot.8772] [Medline: 29097354]
37. Jarvis MA, Sartorius B, Chipps J. Technology acceptance of older persons living in residential care. Information Development 2020 Sep;36(3):339-353. [doi: 10.1177/0266666919854164]
38. Pan S, Jordan-Marsh M. Internet use intention and adoption among Chinese older adults: from the expanded technology acceptance model perspective. Comput Human Behav 2010 Sep;26(5):1111-1119. [doi: 10.1016/j.chb.2010.03.015]
39. Werner JM, Carlson M, Jordan-Marsh M, Clark F. Predictors of computer use in community-dwelling, ethnically diverse older adults. Hum Factors 2011 Oct;53(5):431-447. [doi: 10.1177/0018720811420840] [Medline: 22046718]
40. Chen K, Chan AHS. Gerontechnology acceptance by elderly Hong Kong Chinese: a senior technology acceptance model (STAM). Ergonomics 2014;57(5):635-652. [doi: 10.1080/00140139.2014.895855] [Medline: 24655221]
41. Chandrasekaran R, Katthula V, Moustakas E. Patterns of use and key predictors for the use of wearable health care devices by US adults: insights from a national survey. J Med Internet Res 2020 Oct 16;22(10):e22443. [doi: 10.2196/22443] [Medline: 33064083]
42. Fowe IE, Boot WR. Understanding older adults' attitudes toward mobile and wearable technologies to support health and cognition. Front Psychol 2022;13:1036092. [doi: 10.3389/fpsyg.2022.1036092] [Medline: 36571059]
43. Cavallo F, Esposito R, Limosani R, et al. Robotic services acceptance in smart environments with older adults: user satisfaction and acceptability study. J Med Internet Res 2018 Sep 21;20(9):e264. [doi: 10.2196/jmir.9460] [Medline: 30249588]
44. D'Onofrio G, Fiorini L, Hoshino H, et al. Assistive robots for socialization in elderly people: results pertaining to the needs of the users. Aging Clin Exp Res 2019 Sep;31(9):1313-1329. [doi: 10.1007/s40520-018-1073-z] [Medline: 30560429]
45. McCreadie C, Tinker A. The acceptability of assistive technology to older people. Ageing Soc 2005 May;25(1):91-110. [doi: 10.1017/S0144686X0400248X]
46. Koul S, Eydgahi A. Utilizing Technology Acceptance Model (TAM) for driverless car technology adoption. Journal of Technology Management & Innovation 2018;13(4):37-46. [doi: 10.4067/S0718-27242018000400037]
47. Nastjuk I, Herrenkind B, Marrone M, Brendel AB, Kolbe LM. What drives the acceptance of autonomous driving? An investigation of acceptance factors from an end-user's perspective. Technol Forecast Soc Change 2020 Dec;161:120319. [doi: 10.1016/j.techfore.2020.120319]
48. Debeuf R, Fobelets M, Vaneyghen J, et al. Healthcare professionals' perspectives on development of assistive technology using the comprehensive assistive technology model. Assist Technol 2024 Jan 2;36(1):51-59. [doi: 10.1080/10400435.2023.2202713] [Medline: 37115650]
49. Szabó B, Őrsi B, Csukonyi C. Robots for surgeons? Surgeons for robots? Exploring the acceptance of robotic surgery in the light of attitudes and trust in robots. BMC Psychol 2024 Jan 24;12(1):45. [doi: 10.1186/s40359-024-01529-8] [Medline: 38267998]
50. Hewitt C, Politis I, Amanatidis T, Sarkar A. Assessing public perception of self-driving cars: the autonomous vehicle acceptance model. Presented at: IUI '19: Proceedings of the 24th International Conference on Intelligent User Interfaces; Mar 17-20, 2019; Marina del Ray, CA p. 518-527. [doi: 10.1145/3301275.3302268]
51. Sheetz KH, Claflin J, Dimick JB. Trends in the adoption of robotic surgery for common surgical procedures. JAMA Netw Open 2020 Jan 3;3(1):e1918911. [doi: 10.1001/jamanetworkopen.2019.18911] [Medline: 31922557]
52. Mortensen K, Hughes TL. Comparing Amazon's Mechanical Turk platform to conventional data collection methods in the health and medical research literature. J Gen Intern Med 2018 Apr;33(4):533-538. [doi: 10.1007/s11606-017-4246-0] [Medline: 29302882]
53. Chandler J, Shapiro D. Conducting clinical research using crowdsourced convenience samples. Annu Rev Clin Psychol 2016;12:53-81. [doi: 10.1146/annurev-clinpsy-021815-093623] [Medline: 26772208]
54. Buhrmester MD, Talaifar S, Gosling SD. An evaluation of Amazon's Mechanical Turk, its rapid rise, and its effective use. Perspect Psychol Sci 2018 Mar;13(2):149-154. [doi: 10.1177/1745691617706516] [Medline: 29928846]

Abbreviations

A: attitude
MTurk: Mechanical Turk
PR: personal risk
PU: perceived usefulness
R: reliability
SEM: structural equation model
SR: social risk
TR: time risk

XSL•FO
RenderX

Edited by A Kushniruk; submitted 11.07.24; peer-reviewed by CK Tan, J Odumegwu, L Philpot, RS Gomaa Mahmoud; revised version received 12.07.25; accepted 15.07.25; published 20.11.25.

Please cite as:
Shen S, Xie D, Ding A, Zhang L, Creighton F
Factors Shaping Public Perceptions of a Range of Robotic Technologies in Surgery: Cross-Sectional Web-Based Survey
JMIR Hum Factors 2025;12:e64224
URL: https://humanfactors.jmir.org/2025/1/e64224
doi:10.2196/64224

© Sarek Shen, Deborah Xie, Andy Ding, Lisa Zhang, Francis Creighton. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 20.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluating the Usability of an HIV Prevention Artificial Intelligence Chatbot in Malaysia: National Observational Study

Zhao Ni[1,2], PhD; Sunyoung Oh[1], MSN; Rumana Saifi[3,4], PhD; Iskandar Azwa[3,5], MD; Frederick L Altice[2,3,6,7], MD

[1]School of Nursing, Yale University, 400 West Campus Drive, Orange, CT, United States

[2]Center for Interdisciplinary Research on AIDS (CIRA), Yale University, New Haven, CT, United States

[3]Centre of Excellence for Research in Infectious Diseases & AIDS (CERiA), Universiti Malaya, Kuala Lumpur, Malaysia

[4]Department of Social and Preventive Medicine, Universiti Malaya, Kuala Lumpur, Malaysia

[5]Infectious Disease Unit, Faculty of Medicine, Universiti Malaya, Kuala Lumpur, Malaysia

[6]Section of Infectious Disease, Department of Internal Medicine, Yale School of Medicine, New Haven, CT, United States

[7]Division of Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven, CT, United States

Corresponding Author:
Zhao Ni, PhD
School of Nursing, Yale University, 400 West Campus Drive, Orange, CT, United States

Abstract

Background: Malaysia, an upper middle-income country in the Asia-Pacific region, has an HIV epidemic that has transitioned from needle sharing to sexual transmission, mainly in men who have sex with men (MSM). MSM are the most vulnerable population for HIV in Malaysia. In 2022, our team developed a web-based artificial intelligence (AI) chatbot and tested its feasibility and acceptability among MSM in Malaysia to promote HIV testing. To enhance the usability of the AI chatbot, we made it accessible to the public through the website called MYHIV365 and tested it in an observational study.

Objective: This study aimed to test the usability of an AI chatbot in promoting HIV testing among MSM living in Malaysia.

Methods: This observational study was conducted from August 2023 to March 2024 among 334 MSM. Participants were recruited through community outreach and social-networking apps using flyers. The interactions between participants and the AI chatbot were documented and retrieved from the chatbot developer's platform. Data were analyzed following a predefined metrics using R software (Posit Software, PBC).

Results: The AI chatbot interacted with 334 participants, assisting them in receiving free HIV self-testing kits, offering information on HIV, pre-exposure prophylaxis (PrEP), and mental health, and providing details of 220 MSM-friendly clinics, including their addresses, phone numbers, and operating hours. After the study, 393 human-chatbot interactions were documented on the chatbot developer's platform. Most participants (304/334, 91.0%) interacted with the AI chatbot once, 30 (9.0%) engaged 2 or more times at different intervals. Participants' interaction time with the chatbot varied, ranging from 1 to 31 minutes. The AI chatbot properly addressed most participants' questions (362/393, 92.1%) about HIV and PrEP. However, in 31 interactions, participants posed additional questions to the chatbot that were not programmed into the chatbot algorithms, resulting in unanswered interactions.

Conclusions: The web-based AI chatbot demonstrated high usability in delivering HIV self-testing kits and providing clinical information on HIV testing, PrEP, and mental health services. To enhance its usability in community and clinical settings, the chatbot must offer personalized health information and precise interaction, powered by sophisticated machine learning algorithms. In addition, establishing an effective connection between the AI chatbot and health care systems to eliminate stigma and discrimination toward MSM is crucial for the future implementation of AI chatbots.

(JMIR Hum Factors 2025;12:e70034) doi:10.2196/70034

KEYWORDS

artificial intelligence; chatbot; feasibility; HIV prevention; HIV testing; men who have sex with men; MSM; mHealth; pre-exposure prophylaxis; PrEP; usability; mobile health

Introduction

Men who have sex with men (MSM) are the most vulnerable population for HIV globally, accounting for 44% of new HIV cases in the Asia-Pacific region [1]. Malaysia, an upper middle-income country in this region, is experiencing one of the fastest growing HIV epidemics in MSM [2,3], characterized by a shift from needle sharing to sexual transmission [4]. The vulnerability of MSM to HIV is associated with various factors, including the stigma and discrimination against HIV and same-sex sexual behaviors [3,5,6]. In Malaysia, same-sex sexual behaviors are criminalized by both secular and Shariah law

XSL•FO
RenderX

[3,7], and discrimination against MSM has extended to clinical settings in physicians and medical students [5]. As a result, MSM in Malaysia must navigate through individual, clinical, and policy barriers that perpetuate stigma and discrimination to minimally access HIV-prevention services, including HIV testing and pre-exposure prophylaxis (PrEP). To increase access to quality HIV-prevention services where stigma and discrimination against MSM persist, leveraging cutting-edge technologies is crucial for developing culturally tailored interventions. Therefore, we propose using an artificial intelligence (AI) chatbot to promote HIV testing and PrEP uptake in Malaysia, aiming to counteract the discriminatory environment MSM face and increase their likelihood of accessing high-quality health care.

An AI chatbot is a software technology embedded with machine learning algorithms that can simulate human conversations through text or voice interactions [8]. Since the emergence of ChatGPT (OpenAI Inc), the world has witnessed the popularity of AI chatbot use. In the literature, numerous studies have discussed the potential of AI chatbots in promoting the efficiency of health care and improving people's access to quality health care services [8]. In the health care industry, however, evidence of testing AI chatbots is urgently needed for guiding the future research and implementation of AI chatbot technology to complement current health care systems, particularly in countries and regions where certain vulnerable populations have limited access to high-quality nonstigmatizing health care services. Studying innovative technologies, such as AI, machine learning, and chatbots, is essential because a major barrier to the implementation of conventional digital health interventions in community and clinical settings is that they often require high-intensity and sustained human inputs, management, and maintenance [8-10]. As AI technology advances, AI chatbot–powered digital interventions are poised to overcome the barrier as they simplify the process of obtaining responses by eliminating the need for users to manually sift through search results as required in web-based searches [8] or navigating complex app interfaces among various app features. With advancements in the AI chatbot algorithms, the search for health information is likely to evolve into a streamlined Q&A process in the future.

AI chatbots applied in health care, however, must be sophisticated enough to provide precise and personalized information and update their responses based on users' inputs without human interventions. This feature is particularly important and promising in intensively busy health care systems, where a significant amount of human labor has traditionally been required to provide health care services regularly and manually. The development and implementation of AI chatbots in health care is still in its early stages. Our team has been conducting research on AI chatbots since 2020, and we have created an AI chatbot and tested its feasibility and acceptability for promoting HIV testing among MSM in Malaysia. We found that such an AI chatbot is acceptable by MSM [3,6]. To further enhance the algorithms and features of the AI chatbot, it is essential to test its usability among a large number of end-users using predefined metrics. Therefore, we conducted this observational study to evaluate the usability of our AI chatbot for promoting HIV testing among MSM in Malaysia.

Methods

Study Design and Participants

The AI chatbot was created and embedded into the website MYHIV365.com. The description of the chatbot and the website has been published elsewhere [6]. In this study, participants were recruited from August 28, 2023, to March 9, 2024, through community outreach and social-networking apps. Three research assistants in Kuala Lumpur, Malaysia, distributed recruitment flyers to MSM-friendly clinics and nonprofit organizations. They also posted the flyers on social-networking apps commonly used by MSM in Malaysia, including WhatsApp, Twitter, Telegram, Facebook, TikTok, Grindr, and Instagram. The eligibility criteria for participants include: (1) aged 18 years or older, (2) able to speak and read Bahasa Malaysia or English, and (3) self-identified as MSM. Participants who reside outside Malaysia were excluded because this study aimed to create an AI chatbot that is culturally tailored for MSM living in Malaysia. We recruited 334 participants, who generated 393 interactions with the AI chatbot. An interaction was defined as a contact initiated by a user with the chatbot.

The 3 research assistants collected research data on human-chatbot interactions through the AI chatbot developer's platform using predefined metrics (Table 1), including satisfaction, recommendation score, number of speech bubbles in each interaction, unanswered questions, interaction time, topic selection, HIV self-testing kit order, depression screening, and clinic searching. A speech bubble was defined as a unit of conversation, which is a visual element commonly used in chat interfaces to represent a chatbot's or user's inputs. For instance, if a user accessed the AI chatbot once and asked 10 questions, then there would be 1 interaction and 10 speech bubbles.

XSL•FO
RenderX

Table . The chatbot usability metrics.

Metrics	Description	Data type
Satisfaction with the chatbot		
Satisfaction level	Do you think this chatbot is useful? "Thumb up" or "Thumb down"	Ordinal
Frequency of the satisfaction evaluation	The frequency with which the chatbot evaluates participants' satisfaction level.	Categorical
Recommendation score	"How likely are you to recommend this chatbot to others?" On the scale, 0 stands for extremely unlikely, and 10 stands for extremely likely.	Numerical (discrete)
Number of speech bubbles	A speech bubble is a visual element commonly used in chat interfaces to represent a chatbot's or user's inputs. Higher counts of speech bubbles reflect greater levels of engagement.	Numerical (discrete)
Unanswered questions		
Additional questions asked by participants	Additional questions asked by participants that were not included in the chatbot algorithms.	String (text data)
Frequency of unanswered questions	Participant's questions to which the chatbot responded with "I cannot understand your question."	Categorical
Interaction time		
Duration	Total time duration of a human-chatbot interaction.	Numerical (discrete)
Contact frequency	Total number of contacts from each participant. Higher frequencies reflect greater engagement with the chatbot.	Categorical
Topic selection	The frequency of topic choices (ie, HIV testing, PrEP, and mental health) selected by participants.	Categorical
HIV self-testing kit order	The number of participants who ordered an HIV self-testing kit.	Numerical (discrete)
Depression screening	The frequency of screening depression using the chatbot.	Categorical
Clinic searching	The frequency of searching for details of MSM-friendly clinics, including their addresses, phone numbers, and operating hours.	Categorical

Statistical Analyses

Descriptive statistics were calculated to summarize the variables listed in the chatbot usability metrics. Categorical variables, such as language used, satisfaction level, satisfaction evaluation, unanswered questions, contact frequency, topic selection, depression screening, and clinic searching were reported as frequency (n) and percentage (%). Numerical variables including recommendation score, number of speech bubbles, duration, and HIV self-testing kit order were reported as mean and SD. All analyses were performed using R (version 4.3.3) software (Posit Software, PBC).

Ethical Considerations

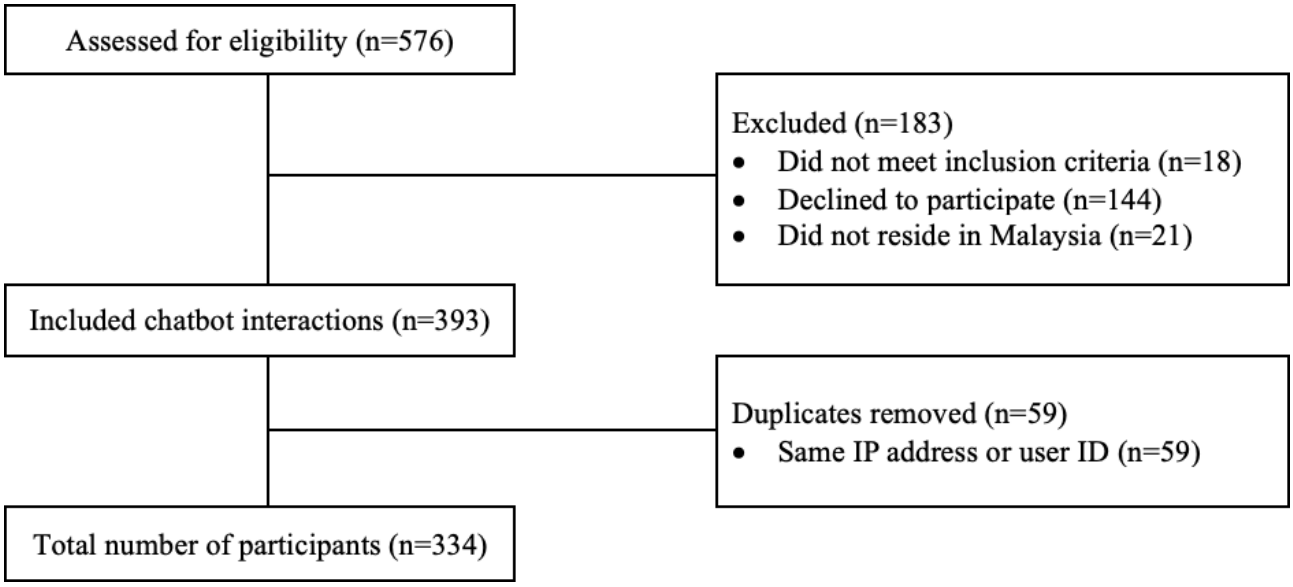
This study received ethical approval from the Yale University Institutional Review Board (Protocol ID: 2000030718) and the Medical Research Ethics Committee of University of Malaya Medical Centre. Participation was voluntary and data were anonymized, thus no compensations were provided. Prior to participation, all participants were provided with a Participant Information Sheet detailing the study. Participants provided electronic consent before beginning the study.

Results

As mentioned, we recruited 334 participants for this study (Figure 1).

XSL•FO
RenderX

Figure 1. The flowchart of participants' recruitment.



Language

In this study, 52.2% (205/393) of the interactions were in English, 47.1% (185/393) were in Bahasa Malaysia, and 0.8% (3/393) were in both languages.

Participants' Satisfaction

The AI chatbot automatically and randomly assessed participants' satisfaction in 127 (32.3%) human-chatbot interactions out of a total of 393. The frequency of evaluations during a single interaction varied randomly; there were occasions when the chatbot assessed participants' satisfaction multiple times but on different topics (eg, HIV testing, PrEP, and mental health). Specifically, the frequency of satisfaction evaluations about the usefulness of the information on HIV testing, PrEP, and mental health was 53, 52, and 32, respectively. We calculated all participants' chatbot evaluations using a scoring system where thumbs up were assigned a value of +1 and thumbs down were assigned a value of −1, resulting in a mean satisfaction score of 1.3 (SD 1.1) (Table 2). Out of the 127 participants who evaluated the AI chatbot, only 14 (11%) provided negative "Thumbs down" feedback. The chatbot randomly assessed the recommendation scores of 43 participants (10.9%) using the Net Promoter Scale, yielding a mean score of 8.4 (SD 2.3).

XSL•FO
RenderX

Table . A summary of the human-chatbot interactions.

Variables	Interactions (N=393)	
	Value, n (%)	Mean (SD)
Language		—[a]
English	205 (52.2)	
Bahasa Malaysia	185 (47.1)	
Both	3 (0.8)	
Satisfaction level[b]	200 (100)	1.3 (1.1)
Thumb up	185 (92.5)	
Thumb down	15 (7.5)	
Frequency of satisfaction evaluation[c]	393	—
0	266 (67.7)	
1	77 (19.6)	
2	38 (9.7)	
3	7 (1.8)	
4	3 (0.8)	
6	1 (0.2)	
8	1 (0.2)	
Recommendation score[d]	43 (10.9)	8.4 (2.3)
Number of speech bubbles	14,262 (100)	
Chatbot	9665 (67.8)	24.6 (14.7)
Participants	4597 (32.2)	11.7 (8.4)
Frequency of unanswered questions	31 (8%)	—
1	25 (6.4)	
2	6 (1.6)	
Duration (minutes)		—
< 1	21 (5.3)	
1-5	303 (77.1)	
6-10	42 (10.7)	
>10	27 (6.9)	
Contact frequency[e]		—
1	304 (91.0)	
2	24 (7.2)	
3	4 (1.2)	
4	2 (0.6)	
Topic selection[f]		—
HIV testing	313 (79.6)	
PrEP[g]	101 (25.7)	
Mental health	49 (12.5)	
HIV self-testing kit order	260 (66.2)	—
Depression screening (PHQ-9)[h]	27 (6.9)	13.4 (7.8)
Clinic searching	93 (23.7)	—

XSL•FO
RenderX

Variables	Interactions (N=393)	
	Value, n (%)	Mean (SD)
HIV testing	25 (5.9)	
PrEP	65 (16.5)	
Mental health	16 (4.1)	

[a]Not applicable.

[b]In the scoring system, thumb up was assigned as +1, and thumb down was -1.

[c]The chatbot randomly evaluated 127 participants' satisfaction score. In every evaluation, the chatbot may randomly assess more than one topic selected from HIV testing, PrEP, and mental health.

[d]The chatbot randomly evaluated 43 participants' recommendation score.

[e]The contact frequency of 334 participants.

[f]Among the 393 interactions, some were related to more than one topic.

[g]PrEP: pre-exposure prophylaxis.

[h]PHQ-9: Patient Health Questionnaire-9.

Unanswerable Questions

The AI chatbot properly responded to 92% of participants' questions that have already been embedded or trained in the chatbot algorithms. However, in 31 human-chatbot interactions, the AI chatbot failed to respond or provide information for participants' questions. For example, a participant requested to interact with an actual human agent through the AI chatbot, but the AI chatbot failed to do so. Another example is that a user inquired about the existence of NGO clinics in Sepang District, Malaysia, but the AI chatbot failed to provide information.

Chat Bubble Count

We analyzed the number of speech bubbles generated during the human-chatbot interactions. In total, the AI chatbot and participants produced 9665 and 4597 speech bubbles, respectively. On average, the AI chatbot produced 24.6 (SD 14.7) speech bubbles, ranging from 2 to 107, while users contributed an average of 11.7 (SD 8.6) speech bubbles, with a range of 1 to 71. Notably, when participants inquired about all 3 topics (HIV testing, PrEP, and mental health), the number of speech bubbles increased.

Duration of Human-Chatbot Interaction, Contact Frequency, and Topics

The participants' interaction time with the AI chatbot varied, ranging from less than 1 minute to 31 minutes. Among the 393 interactions, 21 (5.3%) interacted with the AI chatbot for less than 1 minute. A total of 303 (77.1%) interactions were conducted over 1 minute to 5 minutes, while 42 (10.7%) interactions took place between 6 to 10 minutes, and 27 (6.9%) interactions extended beyond 10 minutes.

Most participants (91%) interacted with the AI chatbot only once, while 24 (7.2%) engaged with the AI chatbot twice at different times during the study. Four participants (1.2%) interacted with the AI chatbot 3 times, and 2 participants (0.6%) engaged with the AI chatbot 4 times during the study. The most selected topics provided by the AI chatbot were HIV testing (79.6%), followed by PrEP (25.7%) and mental health (12.5%). In this study, 260 (66.2%) orders for HIV self-testing kits were submitted.

Depression

Only 27 (6.9%) participants were assessed for depression, with a mean Patient Health Questionnaire-9 (PHQ-9) score of 13.4 (SD 7.8). Based on the PHQ-9 scoring system, 6 participants were identified as having mild depression, 2 as moderate, 2 as moderately severe, and 8 as severely depressed.

Clinic Information

The AI chatbot provided participants with information on 3 types of clinics–HIV, PrEP, and mental health–including their addresses, phone numbers, and operating hours. In the 393 interactions, 93 (23.7%) participants requested clinics' information through the AI chatbot. Of the 393 interactions, 65 (16.5%) participants inquired about PrEP clinics, 25 (5.9%) inquired about HIV testing clinics, and 16 (4.1%) requested mental health counselors' information.

Discussion

Principal Findings

Our AI chatbot has the capacity to reach a large number of MSM, assisting them in receiving HIV self-testing kits and providing them with information on MSM-friendly clinics to facilitate access to services on HIV testing, PrEP, and mental health. Among all chatbot features, offering free HIV self-testing kits was the most popular, followed by inquiries about MSM-friendly clinics' information, including addresses, phone numbers, and operating hours. The mean satisfaction score of 1.3 (SD 1.1) and mean recommendation score of 8.4 (SD 2.3) reveal participants' favorable view of our AI chatbot. This study demonstrates that an AI chatbot embedded with culturally tailored health services and information has the potential to complement the current health care system by offering automatic and algorithm-based educational materials, consulting services, and free medical tools to facilitate and empower health care consumers. However, a significant barrier to increasing the usability of this AI chatbot is the lack of sophisticated algorithms to facilitate precise and personalized interactions with participants. For example, our AI chatbot showed limitations in addressing questions that were relevant but not directly embedded or trained in its algorithms. To integrate this AI

XSL•FO
RenderX

chatbot into community and clinical settings that offer HIV prevention services to MSM in Malaysia, its algorithms must be improved by integrating more sophisticated AI algorithms and by contextualizing it for specific settings. Without robust algorithms, AI chatbots will not be able to fulfill their potential to streamline processes and improve patient care. They will only be able to undertake simple tasks, such as appointment scheduling and prescription refill reminders, but will be incapable of handling more complex tasks in real clinical settings, including symptom checking, patient triage, mental health support, and answering complex health queries [11]. Not to mention providing empathetic consultation or building trust with users. To develop sophisticated AI chatbot algorithms, emerging techniques such as preference-based fine-tuning, retrieval-augmented generation, and quantum machine learning can enable researchers to build more dynamic, context-aware, and customizable chatbots. These approaches warrant systematic testing and incorporation into health care applications.

In addition to the fundamental role of technical sophistication in supporting AI chatbots' capacities, evidence from the literature demonstrates that researchers worldwide have reached a consensus that the primary factors influencing user engagement with AI chatbots include privacy protection, trustworthiness, and accuracy of information [3,8,11,12]. Despite the promising future of integrating large language models into chatbots as AI advances, testing the efficacy of AI chatbots in HIV prevention and addressing associated ethical and practical concerns remains inadequate [12]. So far, only a limited number of studies have been conducted to evaluate the efficacy of AI chatbots in HIV prevention. These studies are mainly focused on high-income countries and regions, such as the United States, Canada [13], and Hong Kong [14]. There is a scarcity of evidence from low-income nations [15], and very few studies have investigated the feasibility of leveraging AI chatbots to facilitate self-management among people living with HIV [13]. In addition, research comparing the effectiveness of AI chatbots with conventional digital tools, such as short message service (SMS) and mobile apps, remains limited. This gap underscores the importance of conducting this study in Malaysia, as it will provide essential user feedback to enhance our AI chatbot before future efficacy testing against other digital health interventions.

Given that the ChatGPT-4 omni (ChatGPT-4o) has robust and sophisticated algorithms capable of facilitating effective human-chatbot conversations, equipping our AI chatbot with ChatGPT's open-source algorithms might be a solution to the concern that our AI chatbot is not precise and personalized enough for use in real-world health care settings. Therefore, in future studies, we will explore the feasibility of integrating the open-sourced ChatGPT-3 algorithms (ChatGPT-4 is not open-sourced yet) with our AI chatbot to test whether an updated version would be able to overcome this issue. While this integration might be a solution, it has significant barriers. First, the information entered on ChatGPT-3 is documented and could potentially be used for other purposes by the OpenAI Inc. or other third parties; therefore, privacy is a major concern that must be addressed by researchers and chatbot developers. To maintain the confidentiality and security of the information entered, we propose that future AI chatbots equipped with self-developed algorithms and the ChatGPT-3's open-source algorithms should be hosted on a safe and secure server owned by a responsible stakeholder, such as a university's cloud environment, where the evaluation of information security has been completed and ethical approval has been received. For researchers who do not have such resources, we recommend defensive measures to avoid data breaches, including backing up data locally, avoiding storage of sensitive information on the cloud, encrypting all cloud data, installing antivirus software to protect encrypted data, creating strong passwords, changing passwords routinely, and requiring at least 2 identification questions to confirm the authorized users.

Second, ChatGPT's algorithms are used for the public, not specifically designed for HIV prevention and MSM; therefore, after upgrading the algorithms, testing the AI chatbot is a key. The focus of the chatbot testing should be on the chatbot's capability to address contextualized problems and on customized training to improve its usability based on end-users' feedback. The testing and training will be an interactive process, during which health scientists and chatbot engineers need to closely collaborate to tailor the AI chatbot's programming language and codes. More importantly, researchers and engineers interested in developing AI chatbots for health purposes must identify the real-world needs in health care settings. They must create user-friendly AI chatbot solutions to meet the expectations of health care professionals and fulfill the needs of health care consumers. In this study, the frequency of unanswered questions was 31 (8%), which could potentially be addressed by improving the AI chatbot's algorithms. Although this study used an observational design, the contact frequency of 304 (91%) indicates that most participants interacted with the AI chatbot only once. This finding suggests that long-term engagement remains a concern that should be explored and tested in future research. The HIV self-testing kit order rate of 260 (66.2%) indicates that most participants primarily used the AI chatbot to order testing kits. Since participants were able to order free HIV self-testing kits through our AI chatbot, this service appeared to be a key incentive for its use. Providing additional free services and integrating them into the AI chatbot would likely increase sustained user engagement. For future implementation of AI chatbots, the priority is to establish an effective connection between the AI chatbot and trusted health care systems to reduce stigma and discrimination toward MSM, while also offering convenient and free services.

Despite the rapid evolution of research on AI chatbots in health care, the application of AI chatbots in health care systems is still in an early stage. There is a need for clear regulations and standards governing the use of AI chatbots in health care, particularly focusing on privacy protection, accuracy of chatbot outputs, cost-effectiveness, and continuous adaptation. For example, depression is a common health concern among MSM globally, and the use of AI chatbots in providing mental health support has gained attention. In our study, the AI chatbot randomly assessed a small proportion of MSM for depression and identified some as having depression. However, the rationale and rules for the random selection of participants are unknown, and after identifying participants with depression, the next step

XSL•FO
RenderX

of linking them with trusted health care systems is unclear. Future research should be conducted to address these questions.

Linking AI chatbots with health care systems to promote high-quality and easy-access health care services is the ultimate goal of our chatbot project. Given that the design of this study is centered on observing users' interaction with the chatbot, we were unable to evaluate the impact of chatbot use on actual HIV prevention behaviors. Our next step is to conduct a randomized controlled trial to test the efficacy of the AI chatbot in promoting HIV testing in the Malaysian context. We have conducted a series of formative studies, including focus group interviews and alpha and beta tests of AI chatbots [3,6,8]. Combining prior study findings with the results of this national observational study, we aim to create an AI chatbot equipped with our own chatbot algorithms and the open-source ChatGPT-3 algorithms in the social-networking app Telegram, which is a widely accepted app for embedding an AI chatbot among MSM in Malaysia, particularly among young MSM [6]. To test the efficacy of the AI chatbot, we will compare it with an attention chatbot in a longitudinal study aimed at improving HIV testing.

Limitations

This study has contributed important information to the usability of AI chatbots in Malaysia, but it has several limitations. First, participants self-reported whether they met the study's inclusion criteria. Since this is a web-based national observational study, we were unable to validate all participants' eligibility. Second, we counted the number of participants using their internet protocol (IP) address. While this method is accurate in calculating the number of participants who used the same IP address to interact with the AI chatbot, it might not capture the same participant who used different IP addresses. Therefore, the actual sample size of the participants might be less than the calculated total sample size. To address these limitations, in our future randomized controlled trial, we will assign a unique study identification number to every participant along with a unique hyperlink to access the AI chatbot. In addition, since the AI chatbot randomly assessed participants' satisfaction and recommendation scores, this approach might have introduced selection bias, as respondents may have been more engaged or satisfied users. Finally, patients with moderate to severe depression were provided with clinical information, but whether they were linked to care or referred to psychiatrists or mental health specialists for treatment and support was not clear to our study team. Future research should include this follow-up information on participants' engagement with clinical care.

Conclusions

The AI chatbot demonstrated high feasibility in delivering HIV self-testing kits and providing clinical information on HIV testing, PrEP, and mental health services. To enhance its usability in community and clinical settings, the AI chatbot must offer personalized health information and precise interaction, powered by sophisticated machine learning algorithms. In addition, establishing an effective connection between the AI chatbot and health care systems to eliminate stigma and discrimination toward MSM is crucial for the future implementation of the AI chatbot.

Acknowledgments

This work was supported by research grants from the National Institutes of Health (R21TW011663 and R33TW011663 for ZN and FA) and the Yale University School of Nursing Faculty Research Support Fund for ZN.

Conflicts of Interest

None declared.

References

1. Men who have sex with men. World Health Organization. URL: https://www.who.int/teams/global-hiv-hepatitis-and-stis-programmes/populations/men-who-have-sex-with-men [accessed 2025-06-14]
2. Baral S, Turner RM, Lyons CE, et al. Population size estimation of gay and bisexual men and other men who have sex with men using social media-based platforms. JMIR Public Health Surveill 2018 Feb 8;4(1):e15. [doi: 10.2196/publichealth.9321] [Medline: 29422452]
3. Peng ML, Wickersham JA, Altice FL, et al. Formative evaluation of the acceptance of HIV prevention artificial intelligence chatbots by men who have sex with men in Malaysia: focus group study. JMIR Form Res 2022 Oct 6;6(10):e42055. [doi: 10.2196/42055] [Medline: 36201390]
4. Global AIDS monitoring: country progress report - Malaysia. Ministry of Health Malaysia. 2022. URL: https://www.moh.gov.my/moh/resources/Penerbitan/Laporan/Umum/MYS_country_report_2021.pdf [accessed 2025-06-14]
5. Ni Z, Shrestha R, Earnshaw VA, et al. Exploring Malaysian physicians' intention to discriminate against gay, bisexual, and other men who have sex with men patients. LGBT Health 2023;10(2):169-175. [doi: 10.1089/lgbt.2021.0452] [Medline: 36251945]
6. Cheah MH, Gan YN, Altice FL, et al. Testing the feasibility and acceptability of using an artificial intelligence chatbot to promote HIV testing and pre-exposure prophylaxis in Malaysia: mixed methods study. JMIR Hum Factors 2024 Jan 26;11:e52055. [doi: 10.2196/52055] [Medline: 38277206]

XSL•FO
RenderX

7. Ellis-Petersen H. Malaysia accused of 'state-sponsored homophobia' after LGBT crackdown. The Guardian. 2018. URL: https://www.theguardian.com/world/2018/aug/22/malaysia-accused-of-state-sponsored-homophobia-after-lgbt-crackdown [accessed 2025-06-14]
8. Ni Z, Peng ML, Balakrishnan V, et al. Implementation of chatbot technology in health care: protocol for a bibliometric analysis. JMIR Res Protoc 2024 Feb 15;13:e54349. [doi: 10.2196/54349] [Medline: 38228575]
9. Ni Z, Liu C, Wu B, Yang Q, Douglas C, Shaw RJ. An mHealth intervention to improve medication adherence among patients with coronary heart disease in China: development of an intervention. Int J Nurs Sci 2018 Oct 10;5(4):322-330. [doi: 10.1016/j.ijnss.2018.09.003] [Medline: 31406843]
10. Ni Z, Wu B, Yang Q, Yan LL, Liu C, Shaw RJ. An mHealth intervention to improve medication adherence and health outcomes among patients with coronary heart disease: randomized controlled trial. J Med Internet Res 2022 Mar 9;24(3):e27202. [doi: 10.2196/27202] [Medline: 35262490]
11. Altamimi I, Altamimi A, Alhumimidi AS, Altamimi A, Temsah MH. Artificial intelligence (AI) chatbots in medicine: a supplement, not a substitute. Cureus 2023 Jun;15(6):e40922. [doi: 10.7759/cureus.40922] [Medline: 37496532]
12. van Heerden A, Bosman S, Swendeman D, Comulada WS. Chatbots for HIV prevention and care: a narrative review. Curr HIV/AIDS Rep 2023 Dec;20(6):481-486. [doi: 10.1007/s11904-023-00681-x] [Medline: 38010467]
13. Ma Y, Achiche S, Tu G, et al. The first AI-based chatbot to promote HIV self-management: a mixed methods usability study. HIV Med 2025 Feb;26(2):184-206. [doi: 10.1111/hiv.13720] [Medline: 39390632]
14. Chen S, Zhang Q, Chan CK, et al. Evaluating an innovative HIV self-testing service with web-based, real-time counseling provided by an artificial intelligence chatbot (HIVST-Chatbot) in increasing HIV self-testing use among Chinese men who have sex with men: protocol for a noninferiority randomized controlled trial. JMIR Res Protoc 2023 Jun 30;12:e48447. [doi: 10.2196/48447] [Medline: 37389935]
15. Laymouna M, Ma Y, Lessard D, Schuster T, Engler K, Lebouché B. Roles, users, benefits, and limitations of chatbots in health care: rapid review. J Med Internet Res 2024 Jul 23;26:e56930. [doi: 10.2196/56930] [Medline: 39042446]

Abbreviations

AI: artificial intelligence
IP: internet protocol
MSM: men who have sex with men
PHQ-9: Patient Health Questionnaire-9
PrEP: pre-exposure prophylaxis

Edited by A Kushniruk; submitted 13.12.24; peer-reviewed by J Tao, Z Wang; revised version received 19.05.25; accepted 27.05.25; published 15.07.25.

Please cite as:
Ni Z, Oh S, Saifi R, Azwa I, Altice FL
Evaluating the Usability of an HIV Prevention Artificial Intelligence Chatbot in Malaysia: National Observational Study
JMIR Hum Factors 2025;12:e70034
URL: https://humanfactors.jmir.org/2025/1/e70034
doi:10.2196/70034

© Zhao Ni, Sunyoung Oh, Rumana Saifi, Iskandar Azwa, Frederick L Altice. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 15.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

User Experience of a Bespoke Videoconferencing System for Web-Based Family Visitation for Patients in an Intensive Care Unit: 1-Year Cross-Sectional Survey of Nursing Staff

Aoife Murray[1,2], MBChB, MSc; Irial Conroy[1], BEng, MSc; Frank Kirrane[3], BSc, MSc; Leonie Cullen[4], CNM; Hemendra Worlikar[1], MSc; Derek T O'Keeffe[2,5], BEng, MBChB, MSc, MD, PhD

[1]Health Innovation Via Engineering Laboratory, School of Medicine, University of Galway, University Road, Galway, Ireland

[2]School of Medicine, College of Medicine Nursing and Health Sciences, University of Galway, Galway, Ireland

[3]Department of Medical Physics and Clinical Engineering, University Hospital Galway, Galway, Ireland

[4]Critical Care Department, University Hospital Galway, Galway, Ireland

[5]Lero Science Foundation Ireland Centre for Software Research, University of Galway, Galway, Ireland

Corresponding Author:
Aoife Murray, MBChB, MSc
Health Innovation Via Engineering Laboratory, School of Medicine, University of Galway, University Road, Galway, Ireland

Abstract

Background: During the COVID-19 pandemic, in-person visitation within hospitals was restricted and sometimes eliminated to reduce the risk of transmission of SARS-CoV-2. Many health care professionals created novel strategies that were deployed to maintain a patient-centered approach. Although pandemic-related restrictions have eased, these systems, including videoconferencing or web-based bedside visits, remain relevant for visitors who cannot be present due to other reasons (lack of access to transport, socioeconomic restraints, geographical distance, etc).

Objective: The aims of this study were (1) to report the experience of intensive care nursing staff using a bespoke videoconferencing system called ICU FamilyLink; (2) to examine the scenarios in which the nursing staff used the system; and (3) to assess the future use of videoconferencing systems to enhance communication with families.

Methods: A modified Telehealth Usability questionnaire was administered to the nursing staff (N=22) of an intensive care unit in a model 4 tertiary hospital in Ireland 1 year after implementing the bespoke videoconferencing system.

Results: In total, 22 nurses working in the intensive care department at University Hospital Galway, Ireland, responded to the survey. A total of 23% (n=5) of participants were between the ages of 25 and 34 years, 54% (n=12) were between 35 and 44 years, and 23% (n=5) were between 45 and 54 years. Most (n=15, 68%) of the participants reported never using videoconferencing in the intensive care setting to communicate with family members before March 2020. The modified Telehealth Usability Questionnaire showed overall satisfaction scores for each subcategory of ease of use and learnability, interface quality, interaction quality, reliability, satisfaction and future use, and usefulness. In total, 21 (95%) participants agreed or strongly agreed with the statement, “I would use the ICU FamilyLink system in future circumstances in which family members cannot be physically present (ie, pandemics, abroad, inability to travel, etc),” and 1 participant responded neutrally. One participant highlighted a common scenario in intensive care settings in which a videoconferencing system can be used “Even without COVID, web-based communication is important when patients become unexpectedly ill and when families are abroad.”

Conclusions: This study provides valuable insights into health care professionals’ experience using a videoconferencing system to facilitate web-based visits for families. We conclude that videoconferencing systems when appropriately tailored to the environment with the users in mind can be an acceptable solution to maintain communication with family members who cannot be physically present at the bedside. The bespoke videoconferencing system had an overall positive response from 22 nursing staff who interacted with the system at varying frequency levels.

(JMIR Hum Factors 2025;12:e54560) doi:10.2196/54560

KEYWORDS

telemedicine; health; telehealth; videoconferencing; web-based; usability; intensive care; critical care; communication; COVID-19; SARS-COV-2; intensive care unit; ICU; cross-sectional survey; nursing; transmission; transmission risk; usability questionnaire; questionnaire; reliability; satisfaction; usefulness; family

XSL•FO
RenderX

Introduction

Background

In recent years, many intensive care departments have adopted a patient- and family-centered care approach. This includes having close family members physically present at the bedside during day-to-day activities and attending care meetings. Previous studies have highlighted the psychological and physical benefits of this approach for patients and their families during an individual's inpatient stay in intensive care [1-3]. However, having family members physically present at the patient's bedside can be difficult or even impossible due to geographical distance, financial challenges, and other family obligations. A significant barrier to bedside visits in the recent past was the COVID-19 pandemic.

The pandemic imposed severe and often complete restrictions on visits and forced health care professionals to change their traditional way of delivering patient care [4]. Prior to COVID-19, critical care nursing staff had been striving toward involving relatives in the intensive care unit (ICU) setting. However, many ICUs suspended these practices with the imposition of visitor restrictions [5]. These restrictive policies were introduced to decrease the risk of transmission of COVID-19 in health care settings [4]. These policies have been shown to be extremely isolating for patients, distressing for family members, and even hampering clinical care provided by health care professionals [1,6]. In particular, patients who are critically ill or vulnerable are often reliant on family for support, and at times, families are surrogate decision makers for patients. Health care professionals quickly realized the impact of visitor restrictions on family-centered care in intensive care settings. They called urgently for ways to maintain communication with family members who were no longer allowed to be physically present at the bedside [7]. At the beginning of the pandemic, hospitals quickly adopted technological solutions, including videoconferencing systems, to maintain or reopen communication channels. Previous studies have shown that adopting telemedicine solutions has numerous barriers, including the user's acceptance and experience with the technology [8-10]. Other studies have shown positive patient experiences using videoconferencing systems for web-based family support during patient rounds. However, other studies identify challenges for health care professionals, including additional workload and difficulty learning and integrating new technology into clinical practice [11,12].

Prior to the pandemic, videoconferencing in the intensive care department was limited to tele-ICU and rarely used for remote family communication. At the start of the COVID-19 pandemic, ICUs worldwide adopted various modes of maintaining communication with family members who were not allowed to be physically present at the bedside due to the risk of transmission of COVID-19. Intensive care units often serve wide geographical areas, so videoconferencing systems have the potential to improve communication with patients' families and reduce the psychological effects of intensive care admission on both patients and families. Videoconferencing systems can enable family members who cannot be physically present to be more involved during a patient's intensive care journey.

Aim of the Study

The aims of this study were (1) to report the experience of intensive care nursing staff using a bespoke videoconferencing system designed for patients and staff to communicate with patients' families remotely during hospitalization in an intensive care department of a tertiary referral hospital that had no prior video-calling system in place; (2) to examine the scenarios that nursing staff used the videoconferencing system; and (3) to assess the future use of videoconferencing systems to enhance family communication in intensive care settings using a modified version of the validated Telehealth Usability Questionnaire (mTUQ) [13].

Methods

Overview of ICU FamilyLink

A bespoke videoconferencing system called "ICU FamilyLink" was developed for the intensive care environment to allow for ad hoc web-based bedside visits available 24/7 with 1 or more close family members who may be in separate households [14]. The system was designed to be easy to use and to allow health care professionals to maintain appropriate control to maintain the security and privacy required for a hospital setting. The technology was chosen to deliver a reliable connection with high-quality audio and video. Several key requirements were previously identified during the iterative process of development [14]. This system used a 23-inch touch-screen video end point mounted on a mobile unit stand with a commercial cloud videoconferencing platform (Figure 1).

XSL•FO
RenderX

Figure 1. Video end point mounted on a mobile unit stand at an intensive care unit bedside.



A simple menu was designed for ease of use and maximum video-call security. Functions such as recording, chat, and screen sharing were disabled for security and privacy. The home screen mirrored the department's naming scheme for patient bed spaces and had the minimum number of clicks to connect. Three mobile video units were available for bedside web-based visits, and a stationary unit was available in a quiet room for medical staff to meet with families digitally.

A suite of documents, email templates, and multimedia was also created to support staff training, use, and troubleshooting. In addition, volunteer IT professionals operated a helpline for family members to ensure clinical staff were not burdened by technical queries outside their clinical training scope. A Data Protection Impact Assessment was completed, and the system was in keeping with hospital protocol and General Data Protection Regulation compliant.

The system was developed and rapidly introduced, with the first call within 2 weeks and the complete system rollout in under 3 weeks. Despite this rapid rollout, significant emphasis was placed on staff training, completion of relevant hospital risk assessments, and adherence to hospital policies. Focus on these areas was essential to ensure that staff, who were the main facilitators of the system, would find the system acceptable.

Survey Development

The validated Telehealth Usability questionnaire (TUQ) was used with minor modifications to capture relevant information to the specific use case [13]. The modified questionnaire (mTUQ) was developed with input from clinicians, engineers, a medical physicist, and a senior intensive care nurse.

The survey consisted of 40 questions in total. It included 22 questions from the TUQ, using a 7-point Likert scale, to assess the ease of use and learnability, interface quality, interaction quality, reliability, satisfaction and future use, and usefulness. The remaining questions ascertained participants' demographics and details on the respondent's use of the videoconferencing system in the ICU. Free-text comment boxes were available to gather additional qualitative data.

XSL•FO
RenderX

Recruitment

The mTUQ was distributed in paper form and via a web-based survey 1 year after the ICU FamilyLink was deployed. This allowed the system to become established in the department and minimized any bias that may have been created at the peak of the pandemic. All ICU nursing staff were eligible to participate in this study. Nursing staff were the primary users and facilitators of the ICU FamilyLink system. The system was mainly used for web-based bedside visits, rather than formal medical updates from the ICU medical team, although it was available if needed. The survey was open to any ICU staff member. Staff were recruited via emails, verbal reminders at nursing handover, and paper copies of surveys left in break rooms and main staff areas.

Data Analysis

Data were exported from the web-based survey and manually inputted from the paper surveys into a Microsoft Excel spreadsheet for analysis.

Ethical Considerations

Ethics approval for this study was obtained from the Galway Hospital Clinical Research Ethics Committee (ref C.A. 2674), and written informed consent, including web-based signatures for web-based participants was obtained from all respondents. All respondents participated voluntarily, could opt out at any point, and did not receive financial compensation. All data were anonymized.

Results

A total of 22 nurses who had been working in the intensive care unit at University Hospital Galway responded to the survey. All the respondents had used the ICU FamilyLink system at least once, although prior usage was not a requirement for participation in this study.

Characteristics of Participants

All participants were nursing staff, of these, 18% (n=4) were nurse managers. In total, 23% (n=5) of participants were between the ages of 25 and 34 years, 54% (n=12) were between 35 and 44, and 23% (n=5) were between 45 and 54 years. A total of 68% (n=15) of participants reported that they had never used videoconferencing in the intensive care setting to communicate with family members prior to March 2020 (Table 1).

During the initial introduction of the system, short training sessions were held, “superusers” were trained, and an instruction manual and video training were created. The survey showed that 36% (n=8) of participants watched a colleague before independent use and 77% (n=17) asked a colleague to teach them.

Table . Baseline characteristics of respondents.

	Respondents (N=22), n (%)
Age categories (years)	
25-34	5 (23)
35-44	12 (54)
45-55	5 (23)
Role	
Staff nurse	18 (82)
Nurse manager	4 (18)
Prior experience with any videoconferencing software in the critical care setting to communicate with patient’s family members	
Yes	7 (32)
No	15 (68)
Frequency of use (number of video calls)	
1-2	6 (27)
3-5	3 (14)
6-10	4 (18)
10-20	4 (18)
>20	5 (23)

Participants’ Use of ICU FamilyLink

The system’s usage was examined based on the frequency of video calls made (Table 1) and various scenarios (Table 2). In total, 27% (n=6) of participants used the system only 1-2 times, while 23% (n=5) of participants used the system over 20 times. Participants most commonly used the system for family time with patients while the staff were present (n=19, 86%), family time with the patient without clinical staff present (n=15, 68%), and providing daily progress updates (n=10, 45%). In total, 27% (n=6) used the system for remote family support while 1 family member was physically present at the bedside. While 9% (n=2) of respondents used the system to deliver bad news, 14% (n=3) used the system for an end-of-life scenario; 1 respondent

XSL•FO
RenderX

reported using the system for a formal care team meeting with the family. In addition, respondents were asked what purposes they felt the system was not suitable for (Table 3). A total of 45% (n=10) of respondents felt that delivering bad news and end of life were not suitable scenarios for using the videoconferencing system (Table 3). One respondent commented that, "Ideally, you would give bad news in person, but it could be used if not possible."

Table . Use of the ICU FamilyLink system.

Scenarios that the respondent has used the ICU FamilyLink system	Respondents (N=22), n (%)
Family time with the patient (while clinical staff carry out routine care)	19 (86)
Family time with the patient (without clinical staff present)	15 (68)
Daily progress updates	10 (45)
Remote family support for a designated in-person family member	6 (27)
End of life	3 (14)
Giving bad news	2 (9)
Formal care team meeting or family meeting	1 (4)
Education session for family	0 (0)

Table . Scenarios that participants felt were not suitable for the ICU FamilyLink system.

Clinical scenarios	Respondents (N=22), n (%)
Giving bad news	10 (45)
End of life	10 (45)
All are suitable	7 (32)
Formal care team meeting or family meeting	3 (14)
Education session for family	2 (9)
Daily progress updates	1 (4)
Other	1 (4)
Family time with the patient (without clinical staff present)	1 (4)
Family time with the patient (while clinical staff carry out routine care)	1 (4)
Remote family support for a designated in-person family member	0 (0)

Usability Attributes of the Videoconferencing System Using the mTUQ

The mTUQ showed an overall positive satisfaction score for all subcategories, including ease of use and learnability, interface quality, interaction quality, reliability, satisfaction and future use, and usefulness (Table 4).

XSL•FO
RenderX

Table . Responses to the modified Telehealth Usability questionnaire (N=22).

	Strongly disagree, n (%)	Disagree, n (%)	Somewhat disagree, n (%)	Neutral, n (%)	Somewhat agree, n (%)	Agree, n (%)	Strongly agree, n (%)
Ease of use and learnability							
1. It was easy to learn how to set up and use the system.	0 (0)	0 (0)	2 (9)	3 (14)	6 (27)	3 (14)	8 (36)
2. It was simple to use the system.	0 (0)	0 (0)	3 (14)	2 (9)	8 (36)	2 (9)	7 (32)
3. Compared to facilitating and managing in-person visits, the FamilyLink system was a more time-efficient way to engage with families.	0 (0)	2 (9)	1 (4)	4 (18)	7 (32)	6 (27)	2 (9)
Interface quality							
4. I am able to navigate setup, initiate, and complete calls without difficulty.	0 (0)	1 (4)	2 (9)	4 (18)	4 (18)	3 (14)	8 (36)
5. The onscreen menu for FamilyLink was intuitive to navigate.	0 (0)	1 (4)	0 (0)	2 (9)	5 (23)	6 (27)	8 (36)
6. The ICU FamilyLink system could do everything I wanted it to do.	0 (0)	0 (0)	1 (4)	2 (9)	4 (18)	5 (23)	10 (45)
Interaction quality							
7. The video quality was good and provided a clear 2-way conversation between me (and the patient, when able) and the family members.	0 (0)	1 (4)	0 (0)	0 (0)	3 (14)	4 (18)	14 (64)
8. The audio quality was good and provided clear 2-way conversation between me (and the patient, if able) and their family members.	0 (0)	0 (0)	1 (4)	1 (4)	1 (4)	5 (23)	14 (64)
9. I prefer the ICU FamilyLink monitor on a stand rather than a handheld device or smaller tablet.	0 (0)	0 (0)	0 (0)	6 (27)	1 (4)	5 (23)	10 (45)

XSL•FO
RenderX

	Strongly disagree, n (%)	Disagree, n (%)	Somewhat disagree, n (%)	Neutral, n (%)	Somewhat agree, n (%)	Agree, n (%)	Strongly agree, n (%)
10. I was able to express myself effectively using the FamilyLink system.	0 (0)	0 (0)	0 (0)	1 (4)	6 (27)	5 (23)	10 (45)
11. Compared to telephone conversations, the ICU FamilyLink was better to communicate and expressing important messages.	0 (0)	1 (4)	2 (9)	2 (9)	3 (14)	6 (27)	8 (36)
Reliability							
12. The system was reliable and consistently facilitated video calls.	0 (0)	0 (0)	1 (4)	2 (9)	4 (18)	4 (18)	11 (50)
13. I feel the FamilyLink system facilitates private and secure communication.	2 (9)	2 (9)	0 (0)	3 (14)	3 (14)	8 (36)	4 (18)
14. Whenever I made a mistake using the ICU FamilyLink system, I could quickly recover.	0 (0)	0 (0)	0 (0)	4 (18)	4 (18)	10 (45)	4 (18)
15. Family members were able to follow the emailed instructions and connect to a web-based visit call without additional technical assistance from me or other health care staff.	0 (0)	1 (4)	0 (0)	2 (9)	7 (32)	6 (27)	6 (27)
Satisfaction and future use							
16. I feel comfortable communicating with my patient's family members using the FamilyLink system.	0 (0)	0 (0)	0 (0)	1 (4)	5 (23)	8 (36)	8 (36)
17. The ICU FamilyLink system is an acceptable way to communicate while visitors are restricted.	0 (0)	0 (0)	0 (0)	0 (0)	2 (9)	9 (41)	11 (50)

XSL•FO
RenderX

	Strongly disagree, n (%)	Disagree, n (%)	Somewhat disagree, n (%)	Neutral, n (%)	Somewhat agree, n (%)	Agree, n (%)	Strongly agree, n (%)
18. I would use the ICU FamilyLink system in future circumstances in which family members cannot be physically present (ie, pandemics abroad and inability to travel).	0 (0)	0 (0)	0 (0)	1 (4)	0 (0)	6 (27)	15 (68)
19. I will use the ICU FamilyLink system in addition to in-person visits in the future.	0 (0)	0 (0)	1 (4)	2 (9)	2 (9)	7 (32)	10 (45)
20. I would recommend the ICU FamilyLink system in other health care settings.	0 (0)	0 (0)	0 (0)	1 (4)	1 (4)	4 (18)	16 (73)
Usefulness							
21. The ICU FamilyLink system provides for continuity of communication while visiting is limited.	0 (0)	0 (0)	0 (0)	0 (0)	2 (9)	5 (23)	15 (68)
22. The ICU FamilyLink system has positively impacted the overall care of patients in critical care during the visitor restrictions.	0 (0)	0 (0)	0 (0)	1 (4)	3 (14)	5 (23)	13 (59)

Ease of Use and Learnability

The majority of nursing staff responded positively to ease of use and learnability questions. Interestingly, when specifically asked to compare the time efficiency of organizing in-person visits versus web-based visits (question 3), 18% (n=4) of participants responded neutrally, 14% (n=3) provided negative responses, and 68% (n=15) gave positive responses (Table 4, questions 1-3).

Two comments were made about the process of sending the videoconferencing link to family members, indicating the need to improve the process.

The equipment itself was very easy to use, it can be a pain to log out of the computer and relog in, an app on the computer at the nurses station would be better.

[Respondent who used the system 10-20 times]

I think it would be beneficial if the email sent to the families to download the software could be sent to every computer in the unit.

[Respondent who used the system more than 20 times]

These comments indicate an area for improvement to reduce the burden of work for staff members and highlight the frustration staff commonly feel when newly introduced technology adds extra tasks. There are limitations to fully integrating technology; however, decreasing the administration time for the clinical staff could lead to better technology uptake. Despite this feedback, the 2 participants reported frequent use of the system.

Interface Quality

Nursing staff most frequently responded with “strongly agree” to the interface quality questions, question 4 (n=8; 36%), question 5 (n=8; 36%), and question 6 (n=10; 45%), respectively (Table 4).

XSL•FO
RenderX

Free text comments from 2 participants included “very simple and easy to use” and “good quality.”

Interaction Quality

A total of 64% (n=14) participants strongly agreed that the video and audio quality was good and provided clear 2-way conversation. When asked about their preference for the ICU FamilyLink monitor on a stand rather than a handheld device or smaller tablet, the majority of participants (n=16, 73%) responded in the positive, only 6 (27%) participants were neutral, and none responded in the negative. Most of the respondents (n=17, 77%) felt that compared to telephone conversations, the ICU FamilyLink was better for communicating and expressing important messages (Table 4, questions 7-11).

One participant highlighted the importance of network coverage for family members, though beyond the hospital’s control, was an important factor: “Very dependent on internet coverage for the family.” Another participant indicated that the visual aspect of the system was better than phone communication: “It helped build a connection with families that isn’t possible via phone.”

Reliability

Overall, respondents more frequently responded positively to reliability (Table 4, questions 12-15). However, 4 (18%) respondents disagreed or strongly disagreed with the statement, “I felt the ICU FamilyLink system facilitates private and secure communication.”

Satisfaction and Future Use

All participants responded positively to the statement, “The ICU FamilyLink system is an acceptable way to communicate while visitors are restricted.” In total, 21 (95%) participants agreed or strongly agreed with the statement, “I would use the ICU FamilyLink system in the future circumstances in which family members cannot be physically present (i.e., pandemics, abroad, inability to travel, etc),” and 1 participant responded neutrally (Table 4, questions 16-20).

Qualitative data also indicate health care professionals’ positive feedback and wish to use the system in the future:

Even without COVID, online communication is important when patients become unexpectedly ill, and families are abroad.

Extremely valued by family members, I have had limited exposure to its use due to working mostly night shifts.

Familylink has been a lifesaver during the pandemic; in my opinion, it is very easy to use and great for patients to be able to see their relatives onscreen while visiting is restricted. I will use it at any opportunity.

Usefulness

In total, 21 (95%) participants responded positively to questions about usefulness (Table 4, questions 21 and 22), except 1 neutral response to question 22: “The ICU FamilyLink system has positively impacted the overall care of patients in intensive care during the visitor restrictions.”

Discussion

Overview

With an overall shift toward a patient- and family-centered approach, health care professionals can harness technology to enable and optimize this approach. However, with the introduction of any technology into a new environment, particularly one as complex as health care delivery, careful design, assessment, and research should be completed to ensure acceptability, usability, and impact on all users. While health care professionals are trained to a high degree in health and medicine, their training, understanding, and use of technology vary widely.

Strengths, Limitations, and Future Directions

This study included participants who had only interacted with the ICU FamilyLink system a handful of times and also very frequent users. It also was inclusive for all age ranges and had both staff grade and senior nurses. In usability studies, it is vital to capture input from all spectrums of technical users.

One limitation of this study was that it did not interview or survey patients or their families. While this was considered, due to limited staff and stressful circumstances for patients and families, it was not included as a component of the study. There have been studies recently published examining family and patient experiences in using both phone and video for web-based visits [15-17]. Staff acceptance is critical for implementing systems like the ICU FamilyLink into regular practice. ICU staff are usually required to facilitate these calls due to the dependency needs of the ICU patients.

This study, like many similar studies, was conducted under pandemic conditions. The TUQ was modified to account for these conditions. It is encouraging that responses to questions on usefulness and future use are overall positive, with 68% (n=15) of respondents “strongly agreeing” that they would use the ICU FamilyLink system in future circumstances in which family members cannot be physically present (question 18). While there are positive indications for postpandemic use and its potential value is noted by the staff members, confirming these results outside of pandemic conditions would be important.

Conclusions

The findings of this study provide valuable insights into health care professionals’ experiences using a videoconferencing system to facilitate web-based bedside visits for family members. We conclude that when appropriately tailored to the environment and with the users in mind, videoconferencing systems can be an acceptable solution for maintaining communication with family members who cannot be physically present at the bedside. Further studies are needed to better understand usability factors for technology in order to enhance and augment communication with ICU patients, staff, and patients’ family members.

XSL•FO
RenderX

Acknowledgments

The authors would like to thank all the staff and volunteers who supported the development and deployment of the ICU FamilyLink system in Galway University Intensive Care Unit as well as the participants of this study for their valuable insight and feedback about the system. This publication has resulted from research supported in part by a grant from Science Foundation Ireland and the European Regional Development Fund (grants 13/RC/2073_P2 to AM and DTOK and 13/RC/2094_P2 to DTOK).

Conflicts of Interest

None declared.

References

1. Ashana DC, Cox CE. Providing family-centered intensive care unit care without family presence-human connection in the time of COVID-19. JAMA Netw Open 2021 Jun 1;4(6):e2113452. [doi: 10.1001/jamanetworkopen.2021.13452] [Medline: 34152421]
2. van Mol MM, Boeter TG, Verharen L, Kompanje EJ, Bakker J, Nijkamp MD. Patient- and family-centred care in the intensive care unit: a challenge in the daily practice of healthcare professionals. J Clin Nurs 2017 Oct;26(19-20):3212-3223. [doi: 10.1111/jocn.13669] [Medline: 27875001]
3. Vincent JL. Critical care—where have we been and where are we going? Crit Care 2013;17(Suppl 1):S2. [doi: 10.1186/cc11500] [Medline: 23514264]
4. Cucinotta D, Vanelli M. WHO declares COVID-19 a pandemic. Acta Biomed 2020 Mar 19;91(1):157-160. [doi: 10.23750/abm.v91i1.9397] [Medline: 32191675]
5. McConnell B, Moroney T. Involving relatives in ICU patient care: critical care nursing challenges. J Clin Nurs 2015 Apr;24(7-8):991-998. [doi: 10.1111/jocn.12755] [Medline: 25597494]
6. Kentish-Barnes N, Cohen-Solal Z, Morin L, Souppart V, Pochard F, Azoulay E. Lived experiences of family members of patients with severe COVID-19 who died in intensive care units in France. JAMA Netw Open 2021 Jun 1;4(6):e2113355. [doi: 10.1001/jamanetworkopen.2021.13355] [Medline: 34152418]
7. Wakam GK, Montgomery JR, Biesterveld BE, Brown CS. Not dying alone—modern compassionate care in the Covid-19 pandemic. N Engl J Med 2020 Jun 11;382(24):e88. [doi: 10.1056/NEJMp2007781] [Medline: 32289215]
8. Gajarawala SN, Pelkowski JN. Telehealth benefits and barriers. J Nurse Pract 2021 Feb;17(2):218-221. [doi: 10.1016/j.nurpra.2020.09.013] [Medline: 33106751]
9. Scott Kruse C, Karem P, Shifflett K, Vegi L, Ravi K, Brooks M. Evaluating barriers to adopting telemedicine worldwide: a systematic review. J Telemed Telecare 2018 Jan;24(1):4-12. [doi: 10.1177/1357633X16674087] [Medline: 29320966]
10. Zhang T, Mosier J, Subbian V. Identifying barriers to and opportunities for telehealth implementation amidst the COVID-19 pandemic by using a human factors approach: a leap into the future of health care delivery? JMIR Hum Factors 2021 Apr 9;8(2):e24860. [doi: 10.2196/24860] [Medline: 33779566]
11. Østervang C, Vestergaard LV, Dieperink KB, Danbjørg DB. Patient rounds with video-consulted relatives: qualitative study on possibilities and barriers from the perspective of healthcare providers. J Med Internet Res 2019 Mar 25;21(3):e12584. [doi: 10.2196/12584] [Medline: 30907746]
12. Petersson NB, Jørgensen AL, Danbjørg DB, Dieperink KB. Video-consulted rounds with caregivers: the experience of patients with cancer. Eur J Oncol Nurs 2020 Apr 21;46:101763. [doi: 10.1016/j.ejon.2020.101763] [Medline: 32447267]
13. Parmanto B, Lewis AN Jr, Graham KM, Bertolet MH. Development of the Telehealth Usability Questionnaire (TUQ). Int J Telerehabil 2016;8(1):3-10. [doi: 10.5195/ijt.2016.6196] [Medline: 27563386]
14. Conroy I, Murray A, Kirrane F, Cullen L, Anglim P, O'Keeffe D. Key requirements of a video-call system in a critical care department as discovered during the rapid development of a solution to address COVID-19 visitor restrictions. JAMIA Open 2021 Oct;4(4):ab091. [doi: 10.1093/jamiaopen/ooab091] [Medline: 34805775]
15. Galazzi A, Binda F, Gambazza S, et al. The end of life of patients with COVID-19 in intensive care unit and the stress level on their family members: a cross-sectional study. Nurs Crit Care 2023 Jan;28(1):133-140. [doi: 10.1111/nicc.12783] [Medline: 35599352]
16. Piscitello GM, Fukushima CM, Saulitis AK, et al. Family meetings in the intensive care unit during the coronavirus disease 2019 pandemic. Am J Hosp Palliat Care 2021 Mar;38(3):305-312. [doi: 10.1177/1049909120973431] [Medline: 33207937]
17. Zante B, Erne K, Jeitziner MM. Video calls did not reduce PTSD symptoms in relatives during restricted ICU visits in the COVID-19 pandemic. Sci Rep 2022 Aug 24;12(1):14405. [doi: 10.1038/s41598-022-18616-8] [Medline: 36002566]

Abbreviations

ICU: intensive care unit
mTUQ: modified Telehealth Usability Questionnaire
TUQ: Telehealth Usability Questionnaire

XSL•FO
RenderX

Edited by P Santana-Mancilla; submitted 14.12.23; peer-reviewed by N Ahmadi, PD Reyes, R Marshall; revised version received 31.07.24; accepted 12.08.24; published 21.03.25.

Please cite as:
Murray A, Conroy I, Kirrane F, Cullen L, Worlikar H, O'Keeffe DT
User Experience of a Bespoke Videoconferencing System for Web-Based Family Visitation for Patients in an Intensive Care Unit: 1-Year Cross-Sectional Survey of Nursing Staff
JMIR Hum Factors 2025;12:e54560
URL: https://humanfactors.jmir.org/2025/1/e54560
doi:10.2196/54560

© Aoife Murray, Irial Conroy, Frank Kirrane, Leonie Cullen, Hemendra Worlikar, Derek T O'Keeffe. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 21.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Feasibility of Long-Term Physical Activity Measurement With a Wearable Activity Tracker in Patients With Axial Spondyloarthritis: 1-Year Longitudinal Observational Study

Emil Eirik Kvernberg Thomassen[1,2], MSc; Anne Therese Tveter[1,3], Prof Dr; Inger Jorid Berg[1], Dr med, PhD; Eirik Klami Kristianslund[1], Dr med, PhD; Andrew Reiner[4], MSc; Sarah Hakim[1], MD, PhD; Laure Gossec[5,6], Dr med, Prof Dr; Gary J Macfarlane[7], Dr med, Prof Dr; Annette de Thurah[8,9], MPH, Prof Dr; Nina Østerås[1,2], Prof Dr

[1]Centre for Treatment of Rheumatic and Musculoskeletal Diseases, Diakonhjemmet Hospital, Diakonveien 12, Oslo, Norway

[2]Faculty of Medicine, University of Oslo, Oslo, Norway

[3]Department of Rehabilitation Science and Health Technology, OsloMet – Oslo Metropolitan University, Oslo, Norway

[4]Oslo Centre for Biostatistics & Epidemiology, Oslo University Hospital, Oslo, Norway

[5]INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique, Sorbonne Université, Paris, France

[6]Rheumatology Department Pitié-Salpêtrière Hospital, Assistance Publique – Hôpitaux de Paris, Paris, France

[7]Aberdeen Centre for Arthritis and Musculoskeletal Health (Epidemiology Group), University of Aberdeen, Aberdeen, United Kingdom

[8]Department of Rheumatology, Aarhus University Hospital, Aarhus, Denmark

[9]Department of Clinical Medicine, Aarhus University, Aarhus, Denmark

Corresponding Author:
Emil Eirik Kvernberg Thomassen, MSc
Centre for Treatment of Rheumatic and Musculoskeletal Diseases, Diakonhjemmet Hospital, Diakonveien 12, Oslo, Norway

Abstract

Background: Using wearable activity trackers shows promise in measuring physical activity in patients with axial spondyloarthritis (axSpA). However, little is known regarding the feasibility of long-term use.

Objectives: This study aimed to explore the feasibility of recording physical activity using a wearable activity tracker and describe wear-time patterns among patients with axSpA.

Methods: Data from a randomized controlled trial (NCT: 05031767) were analyzed. Patients with axSpA and low disease activity were recruited from an outpatient clinic and asked to wear a Garmin vívosmart 4 activity tracker for 1 year. The activity tracker measured steps and heart rate. Trial feasibility (eligibility, inclusion rate, and patient characteristics), technical feasibility (data recorded, tracker adherence, ie, days worn, and missing data), and operational feasibility (synchronization reminders and tracker replacements) were analyzed. Tracker adherence was calculated as the percentage of recorded minutes of the maximum possible minutes. Unsupervised hierarchical clustering was used to explore tracker wear-time patterns.

Results: Of the 160 patients screened, 75 (47%) agreed to use the tracker and 64 (85%) were analyzed (11 had insufficient data). The median activity tracker adherence over 1 year was 66% (IQR 30 - 86). There was 30% missing step and 0.01% heart rate data in the physical activity dataset. A median of 18 (IQR 9 - 25) reminders per patient to synchronize activity data were distributed. Analysis of wear-time patterns resulted in 3 groups: Adherent (33/64, 51% of patients), Minimal Use (17/64, 27%), and Intermittently adherent (14/64, 22%).

Conclusions: Trial feasibility was low, while technical and operational feasibility were acceptable. Only 51% of the patients were highly adherent. Activity trackers, though trendy, have low to moderate feasibility over 1 year in patients with axSpA. Automated synchronization and adherence barriers should be further explored.

(JMIR Hum Factors 2025;12:e68645) doi:10.2196/68645

KEYWORDS

physical activity; wearables; axial spondyloarthritis; activity trackers; feasibility

Introduction

Recent technological advancements enable novel ways of managing patients with chronic disease by using wearable devices, such as activity trackers [1]. In addition to measuring physical activity, wearable activity trackers can act as motivators to increase levels of physical activity [2-4]. axSpA is a chronic inflammatory joint disease, primarily characterized by sacroiliitis, back pain, and stiffness [5]. First-line management

XSL•FO
RenderX

includes nonsteroidal anti-inflammatory drugs and regular exercise [6]. However, there are indications that patients with axSpA engage in lower levels of physical activity compared with healthy people [7-9] and that they report more barriers to engaging in physical activity compared to controls [10]. Identifying physically inactive patients allows for optimized treatment by supporting patients' self-management of their disease and potentially enhancing exercise adherence [11,12]. The passive collection of data using wearables has been highlighted as a goal within remote monitoring in rheumatology, as it may ease the monitoring of disease activity besides using electronic patient-reported outcomes [13,14]. In addition, continuous measurement with an activity tracker has the potential to provide further insight into how the physical activity levels of patients with axSpA are affected by their disease [15-17].

In other patient groups, such as osteoarthritis and gout, research has shown low to acceptable adherence to the use of activity trackers during 3- and 6-month periods, with declining adherence toward the end of the studies [18,19]. However, further investigation is warranted, given the lack of evidence on the feasibility and adherence to long-term use of wearable activity trackers among patients with axSpA [20-24].

The aim of this study was to explore the trial, technical, and operational feasibility of measuring physical activity using commercially available wearable activity trackers over 1 year among patients with axSpA. Second, the study aimed to analyze wear-time patterns and compare patients' demographic and clinical characteristics between the wear-time clusters.

Methods

Study Design and Setting

This study includes a post hoc analysis of a randomized controlled trial, titled Remote Monitoring in Specialist Health Care Study (ReMonit, ClinicalTrials.gov: NCT05031767). The primary results of the trial are reported elsewhere [25]. The ReMonit Study was a 3-armed randomized controlled trial comparing remote monitoring and patient-initiated care to usual care (prescheduled regular hospital visits) among patients with axSpA [26]. Patients were recruited from an outpatient clinic at Diakonhjemmet Hospital, Oslo, Norway, and randomized 1:1:1 ratio to receive usual care, remote monitoring, or patient-initiated care. Patients randomized to the remote monitoring and patient-initiated groups were asked to use the activity tracker for 1 year. This was a commercially available activity tracker (Garmin vívosmart 4). Physical activity data recorded by the vívosmart 4 was wirelessly transferred manually by patients each week via Garmin Software Development Kit to the MyDignio app [27].

Patients

In the ReMonit study, patients with axSpA with low disease activity (Axial Spondyloarthritis Disease Activity C-Reactive Protein Score (ASDAS-CRP) <2.1) and stable treatment with a tumor necrosis factor inhibitor (TNFi) over the past 6 months were included [26]. Since we used the Garmin Software Development Kit for the present study, patients who already used a Garmin device and the Garmin Connect app could not be included, as this interfered with synchronization of data to the MyDignio app. In addition, due to privacy regulations, direct downloading of data from the patients' private Garmin devices was not allowed. Therefore, patients could not use their private devices in the present study.

Data Collection

At baseline, patients completed a digital questionnaire including age, sex, education level, and working status. BMI was calculated based on self-reported body height and weight, and information on years since diagnosis of axSpA was obtained from their medical records. The patients completed the recommended disease-specific questionnaires [28] such as Bath ankylosing spondylitis disease activity index (BASDAI) (0 - 10, 10 being the worst score), Bath ankylosing spondylitis functional index (BASFI) (0 - 10, 10 being the worst score), and patient global assessment (0 - 10, 10 being the worst score) [29,30]. The ASDAS-CRP was calculated based on patients' self-reported disease activity-related questions and measurement of C-reactive protein [31]. The Work Productivity and Activity Impairment item 6 (WPAI) (0 - 10, 10 being the worst score) was used for measuring the impact of the disease on patients' everyday life [32]. In addition, eHealth literacy was measured by 4 scales from the eHealth Literacy Questionnaire [33]: Using technology to process health information (Scale 1), Ability to actively engage with digital services (Scale 3), Feel safe and in control (Scale 4), and Motivated to engage with digital services (Scale 5). For measuring self-reported physical activity and exercise, we used 3 items from a population-based study measuring the frequency, duration, and intensity of exercise [34]. We then further stratified the patients into 3 groups based on whether the level of physical activity was below recommended (below 150 min of moderate-intensity exercise per week), at recommended (150 moderate or 60 vigorous minutes each week), or above recommended [35].

Feasibility

To assess the feasibility of using activity trackers for measuring physical activity, we categorized feasibility into 3 novel subcategories inspired by a previous feasibility study on activity trackers [18].

Trial Feasibility

This included the inclusion rate (number of patients included), eligibility of patients (characteristics of patients who declined and those included), proportion of patients with recorded data, and differences in patient characteristics between patients in the clusters. For evaluating the trial feasibility, we considered the inclusion rate and proportion of patients with recorded data corresponding to <50% as low, 50 - 70% as moderate, and >70% as high.

Technical Feasibility

This included the number of physical activity minutes recorded by patients, adherence to wearing the activity tracker, and missing data on steps or heart rate. Adherence to the use of the activity tracker was defined and calculated as the number of minutes with recorded data by each patient (per minute) divided by the total number of minutes during daytime (16 h; 348,000

XSL•FO
RenderX

min × 100). Adherence was classified as either low (<50%), moderate (50 - 70%) or high (>70%). An acceptable level of missing data was determined to be below 40% [36].

Operational Feasibility

This included the number of automatic reminders sent to the patients (in comparison to the total maximum possible number of 3900 reminders, eg, 52 weeks × 75 patients), and number of replaced activity trackers during the 1-year period. Equal to or less than a median of 26 reminders per patient (eg, in 50% of the weeks) and<15% replacements of activity trackers in the study sample were deemed acceptable.

Measurement of Physical Activity

The patients received oral and written instructions stating that the activity tracker should be worn on the non-dominant wrist for at least 10 hours a day. It should not be worn on the outside of garments but could be used while swimming and showering. Instructions were also given regarding charging, connecting, and synchronizing the device with the app. Contact information for the study team was provided in the event of technical issues.

The Garmin vívosmart 4 activity tracker (Olathe) was integrated into the MyDignio app by the company Dignio. Physical activity data were recorded at a minute level, measuring both steps per minute and average heart rate during the concurrent minute. The Garmin vívosmart 4 uses an accelerometer for measurement of physical movement and activity and a photoplethysmography sensor for measuring heart rate [37].

In order to reduce data noise from the activity tracker, a filter was applied by Dignio, which filtered out heart rate values below 20 beats per minute. The native Garmin motivational messages and notifications were muted. Patients were instructed to weekly synchronize physical activity data manually from the device using the MyDignio app due to limited internal memory of the activity tracker. If patients failed to synchronize in time, the MyDignio app automatically sent a push notification as a reminder. If patients still did not synchronize after receiving the push notification, the study team followed up with an SMS text message reminder. The number of manual SMS text message reminders sent out was not registered. Any technical issues regarding app connectivity or the activity tracker were primarily resolved by the study team, or if necessary, by developers at Dignio.

Data Analyses

Median values with IQR or mean values with SD were used for describing the demographical and clinical variables. Data regarding the trial, technical, and operational feasibility were described by either percentage, mean with SD, or median with IQR. Differences between patients agreeing to use a wearable compared to those declining were assessed by the Mann-Whitney *U* test or Student's *t*-test.

To assess wear-time patterns based on the data returned from the activity tracker, we used the hourly level data, which were summed up on a weekly level. Each patient's first week of recording was assigned as their baseline week. Unsupervised hierarchical clustering using the pheatmap package in RStudio was used for assessing different wear-time patterns based on the number of hours of physical activity data recorded [38]. The optimal numbers of predefined clusters were based on the visual inspection of the clustered heatmap and internal validation from the R-package clValid [39]. The cluster groups were named according to their observable patterns that occurred within each group. Patient characteristics in the clustering groups for wear-time patterns were compared and tested for significant differences using the Kruskal-Wallis test, and Dunn test if the initial Kruskal-Wallis test had a *P*-value below 0.05. The Fisher exact test was used for testing differences in the distribution of sexes between the groups. Preparation and analysis of data were conducted using RStudio and Stata (version 18.0; StataCorp LLC).

Ethical Considerations

The study was conducted according to the Helsinki Declaration. All patients provided written consent to participate and were informed that they could stop using the activity tracker at any given time during the study. Data were de-identified and stored at a secure research server. The Regional Committees for Medical and Health Research Ethics South-Eastern Norway approved the study (ref: 229187).

Patient and Public Involvement

In total, 2 patient research partners were involved in the planning of the study. They also pilot-tested the activity tracker and gave feedback on the instructions to the patients and on the study logistics. One of the patient research partners (SH) also contributed to analyzing and discussing the results and is a co-author of this manuscript.

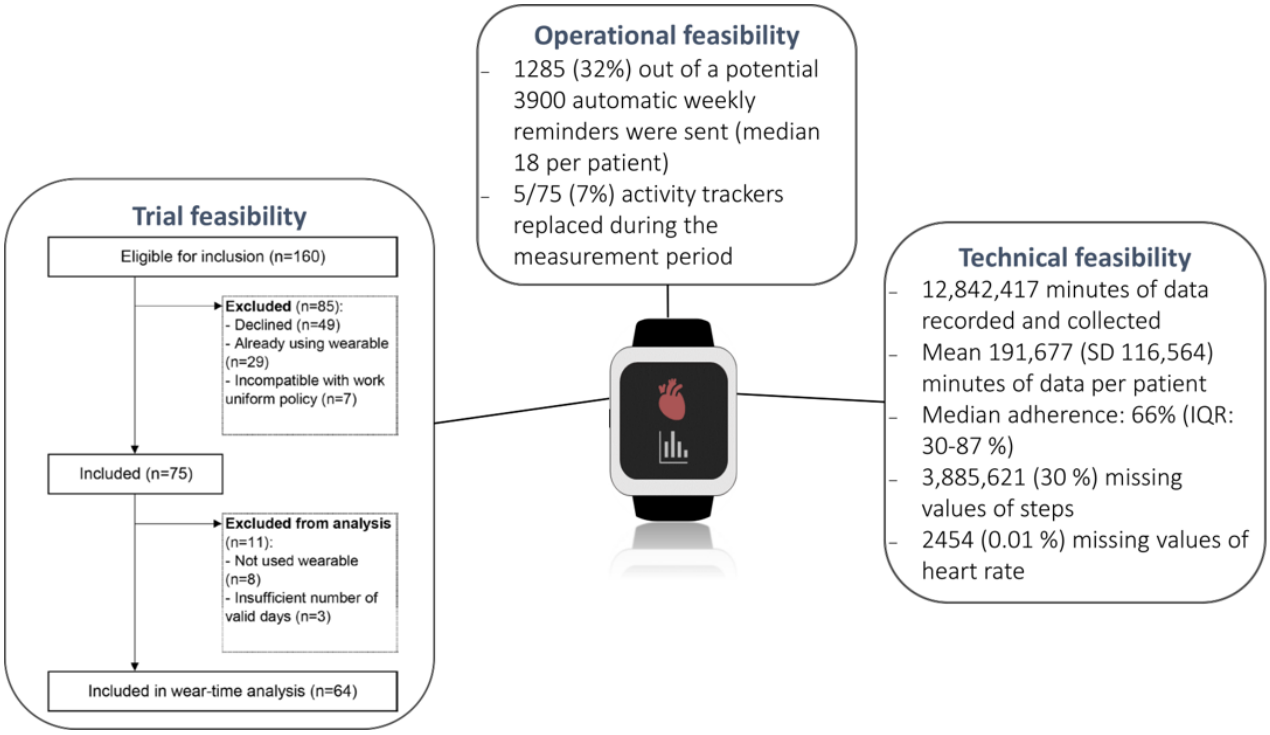
Results

Trial Feasibility

A total of 75 (47%) out of 160 patients agreed to use the activity tracker for 1 year and were included in this study (Figure 1). Among the 85 non-participating patients, 49 (31%) declined to wear an activity tracker, 29 (18%) were already users of a Garmin activity tracker device, and 7 (4%) could not wear an activity tracker due to uniform regulations at their workplace (Figure 1). Baseline characteristics were mostly similar between patients who agreed and those who declined to use the activity tracker, with minor numerical variances in self-reported physical activity levels (Table 1). Patients who declined had significantly lower eHealth literacy scores compared with those who agreed to use the activity tracker, but the numeric differences in mean scores were small (Table 1).

XSL•FO
RenderX

Figure 1. Trial, technical, and operational feasibility of using a wearable activity tracker to record physical activity over 1 year.



XSL•FO
RenderX

Table . Baseline characteristics of the total sample of patients asked to use the wearable activity tracker for 1 year and subsamples.

Characteristics	Total sample asked to use activity trackers (N=160)	Agreed to use the activity tracker (n=75)	Declined to use the activity tracker (n=49)	Ineligible[a] to use the activity tracker (n=36)
Age, years, median (IQR); min-max	42 (34 - 51); 22 - 70	43 (34 - 53); 24 - 70	42 (39 - 50); 22 - 68	41 (32 - 50); 24 - 63
Males, n (%)	123 (77)	57 (76)	39 (77)	28 (78)
Education level, n (%)				
Primary level	33 (21)	19 (25)	9 (18)	5 (14)
University level	127 (79)	56 (75)	40 (82)	31 (86)
Working status, n (%)				
Full-time paid work	126 (79)	59 (79)	39 (80)	28 (78)
Age retired or disability pension	13 (8)	9 (12)	3 (6)	1 (3)
Other[b]	21 (13)	7 (9)	7 (14)	7 (19)
BMI (kg/m^2), median (IQR)	24.9 (22.8 - 27.2)	25.2 (23.0 - 27.5)	24.6 (23.1 - 27.4)	23.8 (21.8 - 26.1)
ASDAS-CRP[c], median (IQR)	0.9 (0.6 - 1.4)	0.9 (0.6 - 1.5)	0.9 (0.6 - 1.2)	0.8 (0.6 - 1.3)
BASDAI[d], median (IQR)	1.0 (0.3 - 2.0)	1.2 (0.5 - 2.3)	1.0 (0.7 - 2.0)	0.6 (0.1 - 1.8)
Fatigue, median (IQR)	1.0 (0.0 - 3.0)	2.0 (1.0 - 4.0)	2.0 (0.0 - 3.0)	1.0 (0.0 - 2.0)
Morning stiffness, median (IQR)	1.0 (0.0 - 2.5)	1.0 (0.0 - 2.5)	1.0 (0.0 - 2.0)	1.0 (0.0 - 1.5)
PGA[e], median (IQR)	1 (0.0 - 2.0)	1.0 (1.0 - 3.0)	1.0 (1.0 - 2.0)	1.0 (0.0 - 2.0)
BASFI[f], median (IQR)	0.3 (0.0 - 1.2)	0.4 (0.1 - 1.7)	0.2 (0.0 - 0.9)	0.1 (0.0 - 0.8)
Years since axSpA diagnosis, median (IQR)	12 (6 - 20)	12 (6 - 23)	12 (8 - 19)	10 (5 - 17)
Self-reported physical activity level[g], n[h] (%)				
Below recommended	10 (11)	3 (7)	5 (21)	2 (11)
Recommended	25 (28)	12 (26)	7 (29)	6 (33)
Above recommended	53 (60)	31 (67)	12 (50)	10 (56)
eHealth literacy, mean (SD)				
eHLQ Scale 1	3.3 (0.5)	3.3 (0.5)	3.1 (0.6)	3.3 (0.6)
eHLQ Scale 3	3.6 (0.4)	3.6 (0.4)	3.5 (0.5)	3.5 (0.6)
eHLQ Scale 4	3.4 (0.4)	3.4 (0.4)	3.3 (0.4)	3.4 (0.5)
eHLQ Scale 5	3.3 (0.5)	3.4 (0.4)	3.1 (0.5)	3.2 (0.6)

[a]Ineligible: due to uniform regulations at work place or already owning a Garmin activity tracker.

[b]Other: receiving social benefits, on sick leave, student/housekeeping, part-time work and unemployed, ASDAS-CRP: Ankylosing spondylitis Disease Activity Score, BASDAI: Bath ankylosing spondylitis disease Activity Index (0-10, 10 being worst), BASFI: Bath ankylosing spondylitis functional index (0-10, 10 being worst).

[c]ASDAS-CRP: Axial Spondyloarthritis Disease Activity C-Reactive Protein Score.

[d]BASDAI: Bath ankylosing spondylitis disease activity index.

[e]PGA: patient global assessment.

[f]BASFI: Bath ankylosing spondylitis functional index.

[g]Self-reported using the HUNT physical activity questionnaire based on World health organization recommendations on physical activity.

[h]n: lower number of participants due to missing.

[i]eHLQ: eHealth literacy Questionnaire (0-4, 4 being best): eHLQ Scale 1: Using technology to process health information, eHLQ Scale 3: Ability to actively engage with digital services, eHLQ Scale 4: Feel safe and in control, eHLQ Scale 5: Motivated to engage with digital services.

XSL•FO
RenderX

Technical Feasibility

Among the 75 patients who agreed to use the activity tracker, 64 (85%) had valid physical activity data and a median adherence to use the activity tracker of 66% (IQR 30 - 86) (Figure 1). Over the 1-year measurement period, 8 out of 75 patients never recorded any data. In total, 3 patients returned less than 1000 minutes of data recorded (corresponding to approximately a total of 16 h of physical activity data) and were subsequently excluded from the wear-time pattern analyses during aggregation of data (Figure 1). A total of 12 million datapoints were returned, corresponding to a mean of 191,677 (SD 116,564) minutes of data per patient (Figure 1). Analyses showed that the proportion of missing data was higher for the steps (30%) compared to heart rate data (0.01%).

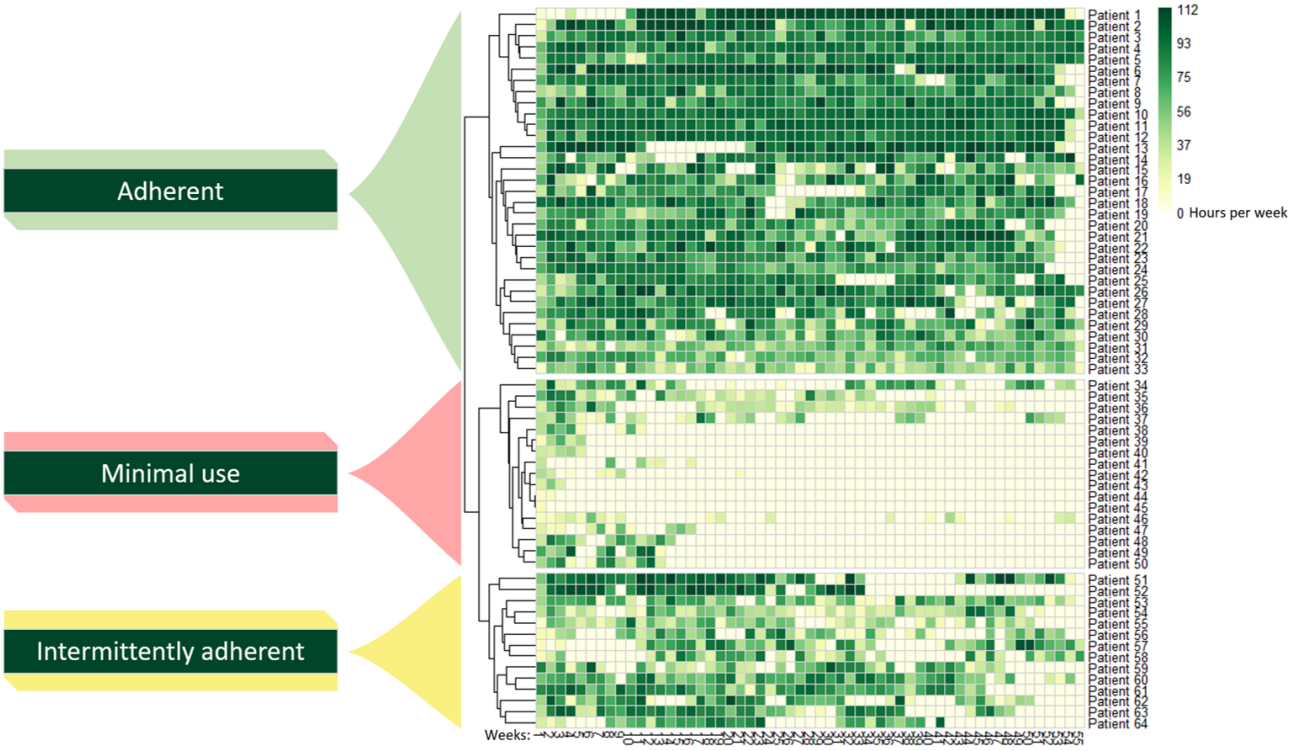
Operational Feasibility

To remind patients to synchronize the activity tracker data, a total of 1285 (32%) out of a potential 3900 automatic push notifications were distributed, with a median of 18 (IQR 9 - 25) notifications per patient. In total, 5 activity trackers required replacement during the study, of which 4 were caused by connectivity issues with the app and one due to a loss of the activity tracker (Figure 1).

Wear-Time Patterns

Data on wear time were grouped in clusters according to the visual wear-time patterns, and 3 cluster groups were found to be the optimal number of clusters. The 3 cluster groups were categorized as “Adherent” (n=33), “Minimal use” (n=17) and “Intermittently adherent” (n=14) according to their visible wear-time patterns (Figure 2). The patients included in the Adherent group had a steady recording of physical activity data, the Minimal use group showed longer periods of no recorded data, and the Intermittently adherent group presented with a notably larger variation in the amount of recorded data (Figure 2). We found that the Adherent group had a median of 85% (IQR 72 - 89) adherence to the use of the activity tracker, while the Minimal use and Intermittently adherent groups had lower adherence with 7% (IQR 4 - 16) and 49% (IQR 43 - 59), respectively (Table 2).

Figure 2. Hierarchical clustered heatmap containing 3 cluster based on wear-time patterns, expressed as hours of physical activity data per week per patient (n=64). The x-axis shows the weeks for each patient who recorded data, with the y-axis showing the number of hours of recorded data per week. Dendrograms at the left show the clustering.



XSL•FO
RenderX

Table . Age, sex, self-reported physical function, impact on daily activities, and disease activity across 3 cluster groups with different wear-time patterns.

	Adherent (n=33)	Minimal use (n=17)	Intermittently adherent (n=14)
Age, years median (IQR)	46 (42 - 55)	35 (32 - 50)	40 (33 - 53)
Sex, males n (%)	23 (70)	16 (94)	10 (71)
BASFI[a], median (IQR)	1.1 (0.2 - 1.9)	0.2 (0.0 - 0.4)	0.7 (0.3 - 1.9)
WPAI[b], median (IQR)	1.0 (0.0 - 3.0)	0.0 (0.0 - 1.0)	1.0 (0.0 - 2.0)
ASDAS[c], median (IQR)	1.2 (0.7 - 1.8)	0.8 (0.6 - 1.2)	1.1 (0.7 - 1.7)
Adherence[d], % median (IQR)	85 (72 - 89)	7 (4-16)	49 (43 - 59)

[a]BASFI: Bath ankylosing spondylitis functional index (0-10, 10 being worst).

[b]Work Productivity and Activity Impairment, Ability to perform daily activities (WPAI Item 6, NRS 0-10, 10 being worst).

[c]ASDAS-CRP: Ankylosing spondylitis Disease Activity Score.

[d]Adherence to use the activity tracker was calculated as data returned divided by maximum amount of day time data.

The Minimal use group had a significantly (*P*=.019) lower median age (35 y, IQR 32 - 50) compared to the Adherent (46 y, IQR 42 - 55) and Intermittently adherent groups (40 y, IQR 33 - 53) (Table 2). There were nonsignificant between-group differences (*P*=.107) in the proportion of sex, with 23 out of 33 (70%) males in the Adherent group, 16 out of 17 (94%) in the Minimal use group, and 10 out of 14 (71%) in the Intermittently adherent group. There were no differences between the groups in disease activity, self-reported physical function, or the impact on daily activities.

Discussion

Principal Findings and Comparison With Previous Works

This study explored the feasibility and wear-time pattern of long-term use of a commercially available activity tracker over 1 year among patients with axSpA. The results indicated a low trial feasibility and acceptable technical and operational feasibility. In total, 3 different cluster groups were identified, with the largest group demonstrating adherence to using the wearable activity tracker.

In the assessment of trial feasibility, we found a low inclusion rate for the activity tracker, with less than half of the 160 eligible patients agreeing to wear the activity tracker for 1 year. This might indicate that the willingness to use a wearable activity tracker over a long period may not be present among the majority of patients with axSpA with low disease activity. However, we believe that a higher inclusion rate could have been achieved if direct data downloads from patients' private Garmin devices had been possible. Implementing such solutions may require specially adapted software and the use of "proxy-users" [40], potentially leading to complex data management.

We observed that a small proportion of patients either did not record any data or recorded less data than the cutoff for a valid day (>10 hours). An earlier study on patients with gout reported similar findings, showing that 33 of 44 patients had valid data, and 40% of the total data was missing [19]. Missing data, when measuring at this high-level granularity, are expected but yet remain a challenge considering the use of activity trackers in clinical settings [41]. A possible reason for the high proportion of missing steps data might be due to differences between Android and iOS phones in the interpretation of the step data, where Android phones misinterpreted inactivity as missing data instead of providing the numerical value of 0. The low number of missing heart rate data may be due to the filter that deleted observations of heart rates below 20 beats per minute, possibly resulting in a skewed representation of the number of missing heart rate data. Additionally, the discontinuation of using the activity tracker has been discussed in earlier studies, showing that perceived usefulness and inaccuracy of data could be a reason for discontinuing the use of activity trackers among patients with osteoarthritis and for healthy people [18,42]. Patients' perception of data inaccuracy may also play a role in our study. Since we only recorded steps and heart rate as measures of physical activity, we effectively limited our recordings to step-related activities, thereby excluding popular Norwegian activities such as bicycling, cross-country skiing, and swimming.

The operational feasibility showed that a median of 18 reminders were sent out per patient, indicating a fairly low number based on the fact that reminders were sent out on a weekly level. It was also observed that the majority of tracker replacements were due to connectivity issues, which could lead to longer periods without recording physical activity.

The hierarchical clustering analysis showed that the largest cluster group was patients who showed adherence to using the activity tracker. Since this represents 33 out of a possible 160 patients, it could be argued that using activity trackers is only feasible for this specific group of patients. The between-group comparisons of patient characteristics revealed numerically small differences between the 3 wear-time cluster groups regarding age and proportion of sex. However, all 3 groups had a small sample size, which limits trust in the between-group comparisons.

Considering the potential benefits of physical activity monitoring, activity trackers may prove valuable in guiding and motivating patients to engage in higher levels of physical activity [2]. Longitudinal data on physical activity can potentially offer

XSL•FO
RenderX

new insights into how axSpA impacts patients' activity levels and can aid in developing targeted exercise and physical activity interventions. The collection of longitudinal physical activity may potentially support physiotherapists and other health care professionals in prescribing lifestyle interventions for patients with axSpA [2]. However, challenges such as low adherence and technical issues must be addressed. A potential increase in adherence might have been achieved if health care professionals had provided personalized guidance on each patient's physical activity level and had set specific goals. These strategies warrant further research to explore their impact on enhancing adherence to using activity trackers.

Strengths and Limitations

The strengths of this study are the long measurement period of 1 year. This allowed for longitudinal analysis of how many patients recorded physical activity and how many discontinued their recordings. This insight further explains how activity trackers may function in clinical settings. Further strengths include the aspect of technical and operational feasibility, showing that the use of wearable activity trackers may involve an increased workload for the health care providers and require complex data management.

Limitations in the study include that we did not set cutoff values as to what constituted a valid week with physical activity measurement. This was decided because we wanted to explore the adherence to using the activity tracker. However, using >10 hours of valid day measurement has also been used by previous studies [22,40]. The hierarchical clustering analysis with the different wear-time patterns has a notable limitation, as it was conducted on a limited sample. This reduces the external validity of our study's findings. Furthermore, the low number of patients per cluster group makes the comparison between these groups prone to biases and skewed data. Lastly, the definition of adherence used in this article may be misleading, as it uses the total amount of physical activity data that was returned. For example, technical issues related to synchronization may have led to missing physical activity data, thus not reflecting the actual use of the activity tracker.

Implications

In order to increase the feasibility of wearable activity trackers, some optimizations are needed. First, data transfer should be automated, allowing for a "passive" collection of data. In the present study, patients had to use the MyDignio app and manually transfer data each week, which might have increased the burden on patients. Second, providing additional measurements beyond steps and heart rate would provide a more comprehensive overview of physical activity levels with time spent in different intensities and perhaps would increase adherence to using the activity tracker. Lastly, using wearable activity trackers, as demonstrated in this study, results in a vast amount of data. Considering the ethical perspectives of storing massive amounts of data, a structured and well-thought-out plan for how data are to be used should be implemented before physical activity data collection is initiated.

Future research should be based upon the recently published Wearable Activity Tracker Checklist for Healthcare (WATCH), which disposes a 12-point list of aspects to consider when implementing a wearable activity tracker in health care [43]. The checklist should be used in studies, preferably conducted in a real-world clinical setting. Additionally, "bring your own device" clinical studies may also hold promise as study design by representing a more accurate resemblance of real-world settings in which some patients already are owners of activity trackers [44]. Future research should incorporate qualitative methods to explore contributing factors behind the variations in wear-time patterns and adherence to using activity trackers.

Conclusions

Despite the trendiness in both research on and commercial use of wearable activity trackers, long-term use of activity trackers had low to moderate feasibility over 1 year in patients with axSpA in low disease activity. The trial feasibility of using wearable activity trackers was low, while technical and operational feasibility were acceptable. Based on the wear-time patterns, we found that only 51% had consistently high activity tracker adherence. Future research should aim to ensure automated synchronization and investigate motivational factors influencing tracker adherence.

Acknowledgments

The authors thank all patients who participated in the ReMonit trial. The study was funded by The South-Eastern Norway Regional Health Authority (2021062) and The Research Council of Norway (328657).

Authors' Contributions

Study conception and design were done by EEKT, NO, IJB, and ATT. EEKT, NO, IJB, ATT, and EKK involved in the acquisition of data. Analysis and interpretations of data were done by EEKT, NO, IJB, ATT, EKK, AR, SH, LG, GM, and AT. Drafting the work or revising was done by EEKT, NO, IJB, ATT, EKK, AR, SH, LG, GM, and AT. EEKT, NO, IJB, ATT, EKK, AR, SH, LG, GM, and AT approved the version of the manuscript to be published and are accountable for all aspects of the work.

Conflicts of Interest

LG has received grants from AbbVie, Biogen, Lilly, Novartis, and UCB. LG has received consulting fees from AbbVie, BMS, Celltrion, Janssen, Novartis, Pfizer, and UCB. LG has received lecture honoraria from AbbVie, Amgen, BMS, Celltrion, Janssen, Lilly, MSD, Novartis, Pfizer, Stada, and UCB. LG has received support for meetings from Biogen and Pfizer. LG has participated

XSL•FO
RenderX

in advisory boards at Janssen, Pfizer, and UCB. LG is EULAR-elected board member. LG has received medical writing support from AbbVie, Amgen, Janssen, Pfizer, and UCB.

References

1. Guillotin T, Saraux A. Role of technology-based innovation in chronic disease management in rheumatology. RMD Open 2024 May 28;10(2):e004264. [doi: 10.1136/rmdopen-2024-004264] [Medline: 38806191]
2. Ferguson T, Olds T, Curtis R, et al. Effectiveness of wearable activity trackers to increase physical activity and improve health: a systematic review of systematic reviews and meta-analyses. Lancet Digit Health 2022 Aug;4(8):e615-e626. [doi: 10.1016/S2589-7500(22)00111-X] [Medline: 35868813]
3. Matcham F, Hotopf M, Galloway J. Mobile apps, wearables and the future of technology in rheumatic disease care. Rheumatology (Oxford) 2019 Jul 1;58(7):1126-1127. [doi: 10.1093/rheumatology/key391] [Medline: 30535022]
4. McGagh D, Song K, Yuan H, et al. Digital health technologies to strengthen patient-centred outcome assessment in clinical trials in inflammatory arthritis. Lancet Rheumatol 2025 Jan;7(1):e55-e63. [doi: 10.1016/S2665-9913(24)00186-3] [Medline: 39089297]
5. Navarro-Compán V, Sepriano A, El-Zorkany B, van der Heijde D. Axial spondyloarthritis. Ann Rheum Dis 2021 Dec;80(12):1511-1521. [doi: 10.1136/annrheumdis-2021-221035] [Medline: 34615639]
6. Ramiro S, Nikiphorou E, Sepriano A, et al. ASAS-EULAR recommendations for the management of axial spondyloarthritis: 2022 update. Ann Rheum Dis 2023 Jan;82(1):19-34. [doi: 10.1136/ard-2022-223296] [Medline: 36270658]
7. Swinnen TW, Scheers T, Lefevre J, Dankaerts W, Westhovens R, de Vlam K. Physical activity assessment in patients with axial spondyloarthritis compared to healthy controls: a technology-based approach. PLoS ONE 2014;9(2):e85309. [doi: 10.1371/journal.pone.0085309] [Medline: 24586239]
8. O'Dwyer T, O'Shea F, Wilson F. Physical activity in spondyloarthritis: a systematic review. Rheumatol Int 2015 Mar;35(3):393-404. [doi: 10.1007/s00296-014-3141-9] [Medline: 25300728]
9. Jacquemin C, Servy H, Molto A, et al. Physical activity assessment using an activity tracker in patients with rheumatoid arthritis and axial spondyloarthritis: prospective observational study. JMIR Mhealth Uhealth 2018 Jan 2;6(1):e1. [doi: 10.2196/mhealth.7948] [Medline: 29295810]
10. Fongen C, Sveaas SH, Dagfinrud H. Barriers and facilitators for being physically active in patients with ankylosing spondylitis: a cross-sectional comparative study. Musculoskeletal Care 2015 Jun;13(2):76-83. [doi: 10.1002/msc.1088] [Medline: 25370164]
11. Damgaard AJ, Primdahl J, Esbensen BA, Latocha KM, Bremander A. Self-management support needs of patients with inflammatory arthritis and the content of self-management interventions: a scoping review. Semin Arthritis Rheum 2023 Jun;60:152203. [doi: 10.1016/j.semarthrit.2023.152203] [Medline: 37068398]
12. Kiefer D, Braun J, Kiltz U, et al. Patients' awareness towards physical activity in the treatment of axial spondyloarthritis. Joint Bone Spine 2023 Sep;90(5):105585. [doi: 10.1016/j.jbspin.2023.105585] [Medline: 37127259]
13. Knitza J, Gupta L, Hügle T. Rheumatology in the digital health era: status quo and quo vadis? Nat Rev Rheumatol 2024 Dec;20(12):747-759. [doi: 10.1038/s41584-024-01177-7] [Medline: 39482466]
14. Gossec L, Guyard F, Leroy D, et al. Detection of flares by decrease in physical activity, collected using wearable activity trackers in rheumatoid arthritis or axial spondyloarthritis: an application of machine learning analyses in rheumatology. Arthritis Care Res (Hoboken) 2019 Oct;71(10):1336-1343. [doi: 10.1002/acr.23768] [Medline: 30242992]
15. Davergne T, Rakotozafiarison A, Servy H, Gossec L. Wearable activity trackers in the management of rheumatic diseases: where are we in 2020? Sensors (Basel) 2020 Aug 25;20(17):17. [doi: 10.3390/s20174797] [Medline: 32854412]
16. Gill JM, Chico TJ, Doherty A, et al. Potential impact of wearables on physical activity guidelines and interventions: opportunities and challenges. Br J Sports Med 2023 Oct;57(19):1223-1225. [doi: 10.1136/bjsports-2023-106822] [Medline: 37549997]
17. Ocagli H, Agarinis R, Azzolina D, et al. Physical activity assessment with wearable devices in rheumatic diseases: a systematic review and meta-analysis. Rheumatology (Oxford) 2023 Mar 1;62(3):1031-1046. [doi: 10.1093/rheumatology/keac476] [Medline: 36005834]
18. Beukenhorst AL, Howells K, Cook L, et al. Engagement and participant experiences with consumer smartwatches for health research: longitudinal, observational feasibility study. JMIR Mhealth Uhealth 2020 Jan 29;8(1):e14368. [doi: 10.2196/14368] [Medline: 32012078]
19. Elmagboul N, Coburn BW, Foster J, et al. Physical activity measured using wearable activity tracking devices associated with gout flares. Arthritis Res Ther 2020 Aug 3;22(1):181. [doi: 10.1186/s13075-020-02272-2] [Medline: 32746893]
20. Davergne T, Pallot A, Dechartres A, Fautrel B, Gossec L. Use of wearable activity trackers to improve physical activity behavior in patients with rheumatic and musculoskeletal diseases: a systematic review and meta-analysis. Arthritis Care Res (Hoboken) 2019 Jun;71(6):758-767. [doi: 10.1002/acr.23752] [Medline: 30221489]
21. Soulard J, Carlin T, Knitza J, Vuillerme N. Wearables for measuring the physical activity and sedentary behavior of patients with axial spondyloarthritis: systematic review. JMIR Mhealth Uhealth 2022 Aug 22;10(8):e34734. [doi: 10.2196/34734] [Medline: 35994315]

XSL•FO
RenderX

22. Chan A, Chan D, Lee H, Ng CC, Yeo AHL. Reporting adherence, validity and physical activity measures of wearable activity trackers in medical research: a systematic review. Int J Med Inform 2022 Apr;160:104696. [doi: 10.1016/j.ijmedinf.2022.104696] [Medline: 35121356]
23. Benavent D, Fernández-Luque L, Núñez-Benjumea FJ, et al. Monitoring chronic inflammatory musculoskeletal diseases mixing virtual and face-to-face assessments-results of the digireuma study. PLOS Digit Health 2022 Dec;1(12):e0000157. [doi: 10.1371/journal.pdig.0000157] [Medline: 36812651]
24. Triantafyllidis A, Kondylakis H, Katehakis D, et al. Smartwatch interventions in healthcare: a systematic review of the literature. Int J Med Inform 2024 Oct;190:105560. [doi: 10.1016/j.ijmedinf.2024.105560] [Medline: 39033723]
25. Berg IJ, Sexton J, Kristianslund E, et al. LBA0004 remote monitoring and patient-initiated care compared to regular face-to-face outpatient visits in axial spondyloarthritis: results from a randomized non-inferiority trial. Ann Rheum Dis 2024 Jun;83(Suppl 1):234-235. [doi: 10.1136/annrheumdis-2024-eular.LBA75]
26. Berg IJ, Tveter AT, Bakland G, et al. Follow-up of patients with axial spondyloarthritis in specialist health care with remote monitoring and self-monitoring compared with regular face-to-face follow-up visits (the ReMonit Study): protocol for a randomized, controlled open-label noninferiority trial. JMIR Res Protoc 2023 Dec 27;12:e52872. [doi: 10.2196/52872] [Medline: 38150310]
27. Dignio. URL: https://www.dignio.com [accessed 2024-10-10]
28. Navarro-Compán V, Boel A, Boonen A, et al. Instrument selection for the ASAS core outcome set for axial spondyloarthritis. Ann Rheum Dis 2023 Jun;82(6):763-772. [doi: 10.1136/annrheumdis-2022-222747] [Medline: 35680390]
29. Calin A, Garrett S, Whitelock H, et al. A new approach to defining functional ability in ankylosing spondylitis: the development of the Bath Ankylosing Spondylitis Functional Index. J Rheumatol 1994 Dec;21(12):2281-2285. [Medline: 7699629]
30. Garrett S, Jenkinson T, Kennedy LG, Whitelock H, Gaisford P, Calin A. A new approach to defining disease status in ankylosing spondylitis: the Bath Ankylosing Spondylitis Disease Activity Index. J Rheumatol 1994 Dec;21(12):2286-2291. [Medline: 7699630]
31. Lukas C, Landewé R, Sieper J, et al. Development of an ASAS-endorsed disease activity score (ASDAS) in patients with ankylosing spondylitis. Ann Rheum Dis 2009 Jan;68(1):18-24. [doi: 10.1136/ard.2008.094870] [Medline: 18625618]
32. Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. Pharmacoeconomics 1993 Nov;4(5):353-365. [doi: 10.2165/00019053-199304050-00006] [Medline: 10146874]
33. Kayser L, Karnoe A, Furstrand D, et al. A multidimensional tool based on the eHealth literacy framework: development and initial validity testing of the eHealth Literacy Questionnaire (eHLQ). J Med Internet Res 2018 Feb 12;20(2):e36. [doi: 10.2196/jmir.8371] [Medline: 29434011]
34. Kurtze N, Rangul V, Hustvedt BE, Flanders WD. Reliability and validity of self-reported physical activity in the Nord-Trøndelag Health Study: HUNT 1. Scand J Public Health 2008 Jan;36(1):52-61. [doi: 10.1177/1403494807085373] [Medline: 18426785]
35. Moholdt T, Wisløff U, Lydersen S, Nauman J. Current physical activity guidelines for health are insufficient to mitigate long-term weight gain: more data in the fitness versus fatness debate (The HUNT study, Norway). Br J Sports Med 2014 Oct;48(20):1489-1496. [doi: 10.1136/bjsports-2014-093416] [Medline: 24782484]
36. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials - a practical guide with flowcharts. BMC Med Res Methodol 2017 Dec 6;17(1):162. [doi: 10.1186/s12874-017-0442-1] [Medline: 29207961]
37. Kim S, Xiao X, Chen J. Advances in photoplethysmography for personalized cardiovascular monitoring. Biosensors (Basel) 2022 Oct 12;12(10):863. [doi: 10.3390/bios12100863] [Medline: 36290999]
38. Kolde R. Pheatmap: pretty heatmaps. R package version. 2019. URL: https://cran.r-project.org/web/packages/pheatmap/pheatmap.pdf [accessed 2024-05-24]
39. Brock G, et al. clValid: an R package for cluster validation. J Stat Softw 2008;25(4):1-22. [doi: 10.18637/jss.v025.i04]
40. Henriksen A, Sand AS, Deraas T, Grimsgaard S, Hartvigsen G, Hopstock L. Succeeding with prolonged usage of consumer-based activity trackers in clinical studies: a mixed methods approach. BMC Public Health 2020 Aug 27;20(1):1300. [doi: 10.1186/s12889-020-09406-w] [Medline: 32854671]
41. Curtis O. 5 The limitations of using commercial wearable activity trackers, such as fitbits, for the clinical monitoring of patient activity levels. . 2022 Nov p. A3-A4. [doi: 10.1136/bmjhci-2022-FCIASC.5]
42. Attig C, Franke T. Abandonment of personal quantification: a review and empirical study investigating reasons for wearable activity tracking attrition. Comput Human Behav 2020 Jan;102:223-237. [doi: 10.1016/j.chb.2019.08.025]
43. Szeto K, Arnold J, Maher C. The Wearable Activity Tracker Checklist for Healthcare (WATCH): a 12-point guide for the implementation of wearable activity trackers in healthcare. Int J Behav Nutr Phys Act 2024 Mar 13;21(1):30. [doi: 10.1186/s12966-024-01567-w] [Medline: 38481238]
44. Demanuele C, Lokker C, Jhaveri K, et al. Considerations for conducting Bring Your Own “Device” (BYOD) clinical studies. Digit Biomark 2022;6(2):47-60. [doi: 10.1159/000525080] [Medline: 35949223]

XSL•FO
RenderX

Abbreviations

ASDAS-CRP: Axial Spondyloarthritis Disease Activity C-Reactive Protein Score
axSpA: axial spondyloarthritis
BASDAI: Bath ankylosing spondylitis disease activity index
BASFI: Bath ankylosing spondylitis functional index
TNFi: tumor necrosis factor inhibitor

Edited by P Santana-Mancilla; submitted 12.11.24; peer-reviewed by J Knitza, V Marozas; revised version received 20.03.25; accepted 26.03.25; published 07.05.25.

Please cite as:
Thomassen EEK, Tveter AT, Berg IJ, Kristianslund EK, Reiner A, Hakim S, Gossec L, J Macfarlane G, de Thurah A, Østerås N Feasibility of Long-Term Physical Activity Measurement With a Wearable Activity Tracker in Patients With Axial Spondyloarthritis: 1-Year Longitudinal Observational Study
JMIR Hum Factors 2025;12:e68645
URL: https://humanfactors.jmir.org/2025/1/e68645
doi:10.2196/68645

© Emil Eirik Kvernberg Thomassen, Anne Therese Tveter, Inger Jorid Berg, Eirik Klami Kristianslund, Andrew Reiner, Sarah Hakim, Laure Gossec, Gary Macfarlane, Annette de Thurah, Nina Østerås. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 7.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Building and Beta-Testing Be Well Buddy Chatbot, a Secure, Credible and Trustworthy AI Chatbot That Will Not Misinform, Hallucinate or Stigmatize Substance Use Disorder: Development and Usability Study

Adam Jerome Salyers[1], BA; Sheana Bull[1], MPH, PhD; Joshva Silvasstar[1], MSIS, BS; Kevin Howell[2], MSIS; Tara Wright[2], PhD; Farnoush Banaei-Kashani[3], PhD

[1]Clinic Chat, LLC, 2950 Arkins Ct, Unit 605, Denver, CO, United States

[2]University of Texas Health Sciences at San Antonio, San Antonio, TX, United States

[3]Department of Computer Science and Engineering, University of Colorado, Denver, CO, United States

Corresponding Author:
Sheana Bull, MPH, PhD
Clinic Chat, LLC, 2950 Arkins Ct, Unit 605, Denver, CO, United States

Abstract

Background: Artificially intelligent (AI) chatbots that deploy natural language processing and machine learning are becoming more common in health care to facilitate patient education and outreach; however, generative chatbots such as ChatGPT face challenges, as they can misinform and hallucinate. Health care systems are increasingly interested in using these tools for patient education, access to care, and self-management, but need reassurances that AI systems can be secure and credible.

Objective: This study aimed to build a secure system that people can use to send SMS with questions about substance use, and which can be used to screen for substance use disorder (SUD). The system will rely on data transfer via third party vendors and will thus require reliable and trustworthy encryption of protected health information .

Methods: We describe the process and specifications for building an AI chatbot that users can access to gain information on and screen for SUD from Be Well Texas, a clinical provider affiliated with the University of Texas Health Sciences Center at San Antonio.

Results: The AI chatbot system uses natural language processing and machine learning to classify expert-curated content related to SUD. It illustrates how we can comply with best practices in HIPPA (Health Insurance Portability and Accountability Act) compliance in data encryption for data transfer and data at rest, while still offering a state-of-the-art system that uses dynamic, user-driven conversation to dialogue about SUD, screen for SUD and access SUD treatment services.

Conclusions: Recent calls for attention to user-friendly design concerning user rights that honor digital rights and regulations for digital substance use offerings suggest that this study is timely and appropriate while still advancing the field of AI.

Trial Registration: Not Applicable

(JMIR Hum Factors 2025;12:e69144) doi:10.2196/69144

KEYWORDS

artificial intelligence; chatbot; infrastructure; substance use disorder; digital health; health communication; conversational agent; HIPAA; AI; Healthcare Insurance Portability and Accountability Act

Introduction

Given the ongoing opioid epidemic in the United States [1], consideration of how and whether we can use digital tools to facilitate a response is warranted. One way that digital solutions may be deployed is by using them to scale screening for substance use disorder (SUD). Screening for SUD is critical to identify people at risk [2], as it can lead to referrals for services including outpatient as well as medically managed intensive inpatient services [3]. Although Screening, Brief Intervention, and Referral to Treatment (SBIRT) programs have been widely used to facilitate screening [4], and these programs are available electronically [5], there is still a major gap in screening, given current estimates indicate fewer than 10% of those at risk for SUD ever screened. This exacerbates an already costly response to the opioid epidemic, given that the United States spends more than $271 billion annually to address SUD-related health issues, crime, and lost productivity [6]. If we could enroll more people at-risk in treatment, we would save $4 for every treatment dollar expended [7].

XSL•FO
RenderX

Artificially intelligent (AI) chatbots that deploy natural language processing (NLP) and machine learning (ML) are becoming more common in health care to facilitate patient education and outreach [8,9]. These systems advance earlier generation rule-based chatbots that rely on "fixed state" messages that force users to choose from a predetermined set of responses to an interactive, user-driven system where people can initiate conversations on any number of topics and chatbots can generate answers on-the-fly. The current AI chatbots, particularly generative chatbots such as ChatGPT [10], rely on access to a large language model (LLM) for training bots to correctly classify and respond to queries. These LLMs represent a much more sophisticated way to interact with people that allows for rapid exploration of vast data sources to learn how to correctly understand the intention of the queries that people make, to make inferences from data, and generate relevant answers in response to user queries.

There is an ample body of evidence from investigations of earlier, fixed-state message systems showing they can be used effectively to positively impact health behaviors and health outcomes [11]. Although we do not yet know the impact of moving from fixed-state chatbots to conversational and generative AI systems, health care systems that have adopted these AI systems are optimistic that they will offer ample return on investment [12]. Since the more advanced AI systems are nascent, we have not yet established standards for best practices in their design. System users and developers of AI chatbots, including generative chatbots such as ChatGPT, have raised concerns related to their use that warrant attention for health-related applications [13] Ethical concerns include observations that systems could discriminate against stereotype or stigmatize users, and could also compromise privacy and data sovereignty [14,15]. Researchers have documented tendencies for generative chatbots to misinform [16], hallucinate [17], and obscure information about how data are being accessed and used [18].

Recent reviews and published literature have highlighted the critical challenges that are inherent in maintaining security for AI chatbots. They highlight chronic concerns with data breaches and malicious input and called for standards such as end-to-end encryption, organizational control, and adversarial training (ie, purposefully attempting to confuse a system during development to train it to recognize potentially malicious inputs) to mitigate these [19,20]. One researcher has highlighted additional critical concerns that are specific to health care organizations that seek to use AI chatbots, including a need to align chatbot design and security with Healthcare Insurance Portability and Accountability Act (HIPAA) regulations that govern patient protections in care delivery [21].

Optimizing digital tools for substance use is warranted. However, we currently do not have digital tools that focus explicitly on user rights, including privacy; are evidence-based; user friendly; easily accessible, person-centered [22]; and can be delivered without generating or reinforcing stigma.

In this paper, we present the infrastructure and technical specifications used to design Be Well Buddy, an AI chatbot focused on raising awareness and access to screening and treatment for SUD. It intentionally addresses the security and ethical concerns identified here. We also present findings from a beta test of the system whose goals were to verify system security and functionality.

Methods

This work represents a partnership between Clinic Chat, LLC, a health technology start-up company, and The University of Texas Health Sciences Center at San Antonio (UT Health San Antonio), a US-based university with a robust clinical program called Be Well Texas, focused on screening, treatment, and social support for SUD. The principals from Clinic Chat have a background in scientific research related to health as well as the technical skill to design and deploy AI chatbots; UT Health San Antonio and Be Well Texas have a wide-reaching health education initiative focused on substance use prevention and treatment. The partnership was established to adapt health AI chatbots to focus specifically on raising awareness and providing access to screening for and treatment of SUD within the university program.

System Specifications

Given concerns that generative chatbots can hallucinate and misinform, along with ample evidence that careful attention to the design of messages for health communication can be more impactful than generic messaging, we determined it appropriate to develop and curate messages specific to SUD that could be delivered by a closed-domain AI chatbot system (ie, one where messages returned in response to user queries would only come from this specific library of messages). This approach allows us to avoid a common pitfall of generative AI such as ChatGPT that can reinforce biases or misinform when seeking responses to user queries. By using a closed library of responses that have been reviewed for accuracy, empathy and tone, our system can only choose from a limited group of options that are medically correct and consistent with clinical guidelines for care. Guided by literature documenting approaches to health communication that can increase engagement with messages [23], we generated content that allowed for tailoring to individual users to increase content relevance (eg, by prefacing each message with their first name or by inviting them to name and explore preferred topics) [24]; that could offer intuitive suggestions for behavioral decision-making [25,26]; that offered narrative, emotional messaging, status-enhancing and skill-building content [27]; and that worked to destigmatize substance use [28]. These message design strategies have consistently been shown to have a positive impact on health behavior [29-31]. We developed an initial library of messages and then reviewed the message content in a formative process with members of the intended audience of the chatbot to obtain their feedback on the message content, tone, acceptability, and capacity to engage people in dialogue. One key outcome from this process was the name "Be Well Buddy," which we gave to the system, generated from a participant in this formative message development process (Sarah Mumby, MPH, email August 15, 2024).

Beta Testing

Once the system was designed and built, we conducted a one-day beta test to ensure system functionality with members

XSL•FO
RenderX

of the study team from Clinic Chat and UT Health San Antonio. All participants were encouraged to try and push the system to its limits. Specifically, they asked expected questions (eg, What is a substance use disorder? What is medication-assisted therapy? How much does treatment cost?) and asked multiple questions in a short period of time (ie, five-six questions in a minute). They made queries in different languages and completed all the screeners embedded in the system. They deliberately worked to confuse the system with questions that were nonsensical or unrelated to SUD. Our goals for the beta test were to (1) identify errors in sending and receiving messages and use this to determine the system's precision (ie, whether responses sent by the system appropriately matched the intent behind the user queries); (2) identify any problems with encryption; and (3) identify any delays in message delivery.

Ethical Considerations

The work we describe here was reviewed by UT Health San Antonio Institutional Review Board, who deemed these activities as preliminary to research conducted exclusively by paid research staff and therefore exempted from requirements for human subjects' approval. Identifying data for persons beta-testing the system included telephone numbers, which were encrypted once outside the UT Health San Antonio firewall. The data presented in this paper informed an observational study that has been reviewed and approved by the UT Health San Antonio Institutional Review Board (protocol number 20230662H).

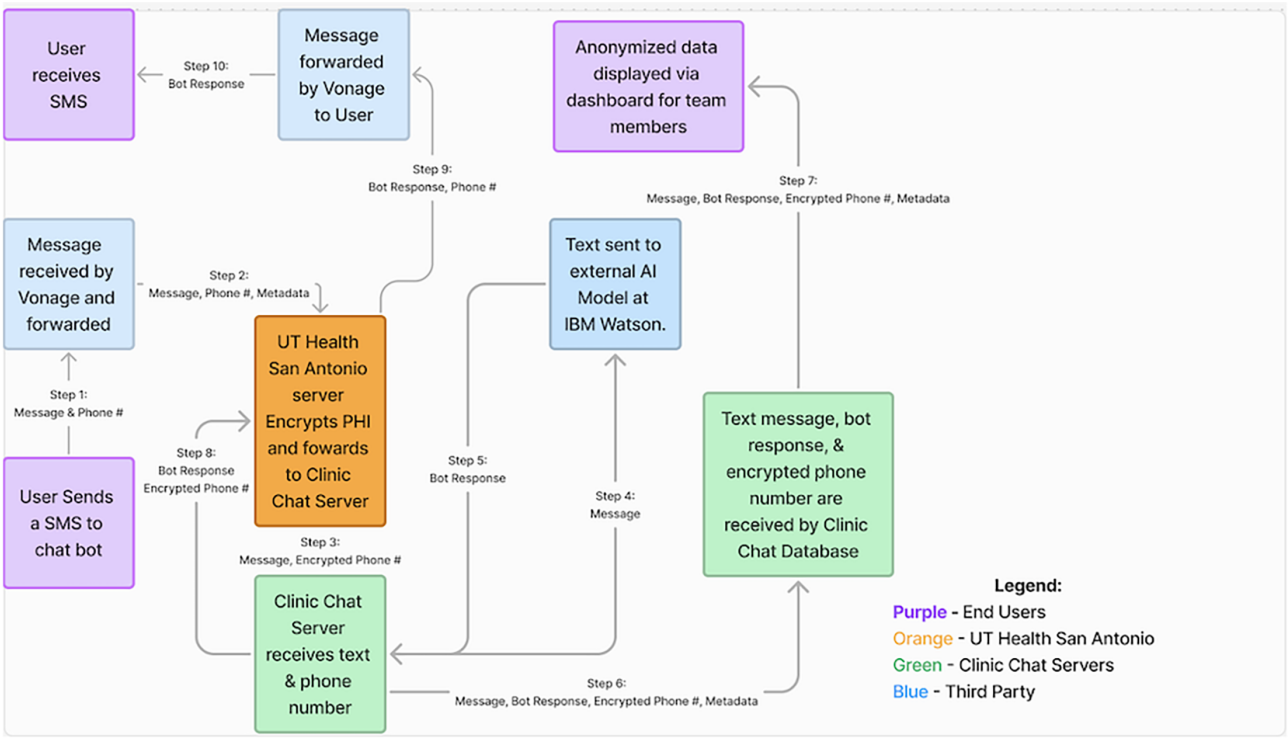
A description of system specifications and functionality, and results of the beta test is given below.

Results

System Specifications

Figure 1 illustrates how the system retrieves and shares information, along with a description of the security steps that are in place to ensure user data are protected. In this figure, purple boxes represent end users of the system, ie, people who send and receive text messages via SMS. The blue boxes represent any third-party vendor that retrieves or sends information from the user to the Clinic Chat AI Chatbot, called "Be Well Buddy." The orange box represents the firewall for UT Health San Antonio, home to Be Well Texas, the clinical entity whose patients and prospective patients are the intended users of Be Well Buddy.

Figure 1. System workflow. AI: artificial intelligent; PHI: protected health information.



Data exchanges between users and the AI chatbot are described as steps in Figure 1. In Step 1, Be Well Buddy initiates a dialogue with users when it sends an initial SMS message to them after receiving information on their first name and telephone number from UT Health San Antonio, which they provide upon enrolling a user in the feasibility study. All SMS messages—incoming and outgoing—are handled by our third-party provider, Vonage Telecommunications Company (Step 2).

System users can respond to any message from Be Well Buddy. When this is done, the data are forwarded to the UT Health San Antonio Server (Step 3), where we host a Flask application designed for minimal impact on system performance. This Flask app uses a webhook to ingest the information efficiently. The lightweight nature of Flask, as opposed to heavier frameworks like Django, is crucial in maintaining fast processing speeds behind the secure firewall. The information received by the Flask app includes one piece of protected health information (PHI)— the phone number—which is now encrypted using SHA-256 bit encryption. This robust encryption method ensures

XSL•FO
RenderX

that the data cannot be read or altered without the encryption key and remains encrypted anywhere outside the firewall, including Clinic Chat Amazon Web Server (AWS).

The message content and metadata without identifiers (eg, message time, type), along with the encrypted phone number are sent to a Clinic Chat-owned and managed AWS, known as EC2 instance that hosts a Django application to authenticate and retrieve information from the Be Well Buddy chatbot; only whitelisted IP's can send requests to the AWS EC2 instance.

At this point, a request is made to our IBM Watson AI (version 10.2.0) assistant that includes only the message body from the SMS message (Step 4). IBM Watson currently serves as our NLP AI system to classify message content. Prior to forwarding messages to Watson, our Django app (version 5.2) performs a variety of preprocessing tasks using regex and custom code to effectively manage number inputs, survey and screener responses, and any content that does not constitute a direct user query. While IBM Watson effectively classifies message content and returns a response (Step 5), we plan to transition to an internally hosted LLaMA integration in the future to leverage the impressive capabilities of LLMs. This response, along with the encrypted phone number, is sent back to a Flask app—a lightweight application optimized for faster processing—hosted behind the university server before being forwarded to the user through Vonage.

The response, message, metadata, and encrypted phone number are sent to an Amazon Web Server (AWS) Relational Database Service (RDS) instance for storage (Step 6). A dashboard hosted on the Clinic Chat subdomain (through another AWS EC2 instance) can then send requests for the system to retrieve data so team members can view large or small trends in anonymized message data (Step 7).

The system relies on data transfers with multiple third-party vendors, including Vonage that aggregates message delivery via multiple cell phone providers; IBM Watson, which classifies message content; and AWS that stores data. Following is a description of how each of these providers operate with the Be Well Buddy chatbot.

Vonage is our third-party telecommunications provider. They manage all outgoing and incoming messages. As they handle phone numbers, which are considered PHI that is regulated under the HIPAA, they have executed a Business Associate Agreement with Clinic Chat. This Business Associate Agreement ensures that both Vonage and Clinic Chat remain compliant with HIPAA regulations related to PHI by redacting all phone numbers and messages from Vonage logs. To retrieve the phone numbers and message information, the UT Health San Antonio server has generated a whitelist of IP addresses from Vonage (ie, the IP addresses for computers used by Clinic Chat staff) that will limit JSON data intakes exclusively from these IP addresses. All data will be sent to Vonage and received from Vonage through HTTPS, a secure data transfer protocol that encrypts data during transfer.

IBM Watson is used t0 store data to train our AI models, allowing the system to more precisely interpret incoming queries from users so the system can then respond with a message from our curated library that corresponds to the intent behind each user query. IBM Watson will never receive any PHI in any form from messages—they only receive user queries as the text message body, which is used to match to one of our answers in our library of responses before returning it to the Clinic Chat AWS instance [32] (Step 8).

We are using two AWS EC2 server instances—one to host the backend server responsible for connecting the UT Health San Antonio server, IBM Watson and the database, and a second to host the front-end team interaction and database. Both are configured according to AWS HIPAA guidelines, although unencrypted PHI will never reach AWS. The backend server will only connect to the UT Health San Antonio server via HTTPS connection, IBM Watson, and AWS RDS. We whitelist the UT Health San Antonio server IP (and a few developer IPs for testing purposes), but no other IPs will be able to hit the server. The front end is also whitelisted to limited team member IPs and contains a login and user authentication system to view anonymized data.

We also use an AWS RDS instance to store our data for the project. All PHI are encrypted when entering AWS and are stored in encrypted form in RDS; they are never decrypted in AWS. The RDS database is connected to EC2 instances. The backend will only write data to the database, and the front end will only pull data.

The UT Health San Antonio server hosts a Flask app to process and encrypt PHI before passing it on to AWS, and process and decrypt phone numbers when sending it to Vonage. The app has a webhook that interprets a series of whitelisted IPs from Vonage, which will send JSON data including phone number and message content. Phone numbers are the only PHI the system receives, so the Flask app encrypts the phone number using Python Fernet encryption before sending the message content and encrypted phone number to our AWS EC2 backend instance. Additionally, the Flask app has a webhook that receives incoming JSON data from the same AWS EC2 instance. The AWS EC2 instance is whitelisted, and along with Vonage, these are the only servers that are whitelisted (outside of a dedicated IP for a developer to fix any issues). The Flask app will take in an encrypted phone number and response message from the EC2 instance, decrypt the phone number, and pass the phone number and response message back to Vonage to be sent to the user (Steps 9 and 10).

In anticipation of possible system attacks, we also built in protocols to reduce their impact should they occur. We exposed our model to anticipated queries that are adversarial (eg, "I don't want to screen for SUD, leave me alone," or "Your answers are stupid") to train for resilience and respond professionally (eg, "I'm sorry you are disappointed with my response. I would be happy to try again or discuss a different topic"). When a user asks the system a question, the response is chosen from the library using NLP and then relies on probabilistic models to determine how likely it is that a response from our library will match the intent behind the user query. We monitor all logs to determine if system users and system content are behaving as expected and to identify any suspicious interactions.

XSL•FO
RenderX

Beta Testing

The beta test was conducted over the course of one day with six users. Users sent 426 messages to the system (for an average of 71 messages per user, with a range of 25 - 110 messages), and the system responded 800 times. The additional system responses included intentional follow-up prompts (eg, "Ask me about something else! I can answer your questions about medication-assisted therapy" or "Are you interested in screening for SUD?") that were appended to responses. We documented two instances where users received responses in Spanish, although the content in the library was only available in English. Upon investigation, we determined that the system was retrieving messages from a different library within the Clinic Chat system focused on chronic illness self-management. We unlinked the libraries between the SUD and chronic conditions content to avoid this error in the future. These two errors were the only ones we documented when the system returned an incorrect response (ie, 2 of 426 or <1%). When a user sent a nonsensical query or one that did not align with SUD, the system correctly responded with "I'm sorry, I am still learning and did not understand your question. Will you please ask again in different words?" Thus, we established a >99% level of system precision overall. We documented three instances where message return was quite slow (ie, a >90 seconds to return a response). In all other instances, responses were sent within the first 10 seconds of message receipt. We recognized this as a problem with Vonage and reached out to report the message delays to them. We identified a problem with the responses returned to users who screened for anxiety, depression, and SUD, where those with high risk were not alerted to their risk and given a referral, necessitating a review and correction of the referral algorithm. The algorithm for screening and responses related to screening were incorrect and fixed. We did not identify any problems with system encryption or data security in transferring queries via SMS. Given the short time frame and small sample of beta testers, we did not experience any adversarial attacks to our system during the beta test and cannot report on the robustness of our protocol to address these attacks in this paper.

Discussion

In this study, we present the technical specifications included in Be Well Buddy to facilitate access to information and screening for SUD for Be Well Texas, a clinical organization affiliated with the UTHSA. Our system adheres to HIPAA regulations for the protection of PHI while avoiding pitfalls of current generative AI chatbots by using curated chat content that does not misinform or hallucinate. Users of Be Well Buddy can get credible and complete information about SUD and can access care on a 24/7 basis. The beta test uncovered and resolved two functional errors; ie,(1) delivery of messages unrelated to SUD and (2) an incorrect algorithm for screening feedback. After correcting these errors, we verified that the system functions as intended and has a high level of precision [33].

This study is not without limitations. Our objective was to ensure that the system worked without error and was secure; therefore, using a small number of beta testers was appropriate to achieve these findings [34]. However, the small number of beta testers did not allow us to delve deeply into how well the content resonated for users. Further, as noted above, given the short time frame and small sample of beta testers, we did not experience any adversarial attacks to our system during the beta test and thus cannot report on the robustness of our protocol in such scenarios. These topics would be appropriate for a subsequent trial of system use.

While we stand behind our approach of using a closed library system to avoid challenges with misinformation and hallucination, this model requires regular updates to content to ensure it remains consistent with medical and professional guidelines related to SUD, SUD screening and treatment referrals. This task may evolve to become cumbersome if library updates are frequent or extensive.

AI chatbots can struggle with delineating between nuanced concepts, making it difficult to support complex questions from users. As we rely on IBM Watson for classification and other language models, this challenge persists and will require careful attention to identify inaccuracies to allow for appropriate reclassification of content when systems do not respond with precision.

Finally, we recognize that the datasets used to train AI chatbots have inherent biases, which we risk reproducing when scaling this system. This risk can be particularly challenging for the Be Well Buddy chatbot whose explicit goal is to reduce stigma. Extra effort is recognized to ensure that the content does not produce or reinforce stereotypes about substance use or people who use substances [35].

According to a recent McKinsey report [12], a large group of surveyed health care leaders indicated that their organizations are eager to use generative AI to enhance operations—particularly patient engagement processes—but most are still adopting a wait-and-see approach. While traditional rule-based chatbots have already proven effective, the AI-driven chatbots, particularly those that are driven by generative models such as ChatGPT, have yet to be proven safe for health care applications. In particular, a number of ethical concerns, such as privacy, confidentiality, bias and fairness, transparency, accountability, regulatory compliance, risk anticipation and copyright complications remain the subject of active research and mitigation in generative models. Additionally, the risks of misinformation, perpetuation of bias, or hallucination continue to dampen enthusiasm for widespread deployment of AI Chatbots in health care [8].

Until effective and proven solutions are developed to address these concerns, widespread use of AI-driven chatbots in health care applications is not anticipated.

In this paper, we introduced the Be Well Buddy chatbot, a novel AI-driven chatbot that adopts a secure and reliable approach to engage with patients with SUD. Our system circumvents challenges with misinformation, bias, and hallucination, while securely delivering and supporting access to SUD screening and treatment options.

Our research is consistent with recent recommendations for optimized digital substance use interventions, such as solutions

XSL•FO
RenderX

that emphasize digital rights of privacy and confidentiality, accessibility, and user-friendliness [15]. In the future, we will further experiment with Be Well Buddy to evaluate and report its efficacy in promoting self-screening and referral for SUD. We anticipate integration with treatment providers who wish to accelerate referrals for care by inviting persons living in their catchment areas to use the system and for self-screening. This system may also be integrated into other organizations adjacent to SUD treatment, such as the criminal justice system where care may be mandated, or the parole system where people may seek to closely monitor their own or clients' risk for SUD.

Acknowledgments

The authors gratefully acknowledge the participants who were willing to engage with the system and offer their feedback. This work was supported with funding from the National Institutes of Health, National Institute of Drug Abuse (NIDA) under grant number 1R41DA059275-01.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

SB, AJS, and JS are affiliated with the organization Clinic Chat LLC that developed and deployed the Be Well Buddy system described in this study.

References

1. Salazar CI, Huang Y. The burden of opioid-related mortality in Texas, 1999 to 2019. Ann Epidemiol 2022 Jan;65:72-77. [doi: 10.1016/j.annepidem.2021.09.004] [Medline: 34560252]
2. Overdose death rates.: National Institute of Drug Abuse; 2024. URL: https://nida.nih.gov/research-topics/trends-statistics/overdose-death-rates [accessed 2022-12-14]
3. Medicaid innovation accelerator program reducing substance use disorders. high intensity learning collaborative fact sheet. Medicaid Innovation Accelerator Program. 2014. URL: https://www.medicaid.gov/state-resource-center/innovation-accelerator-program/iap-downloads/learn-hilciap.pdf [accessed 2022-07-11]
4. Saitz R. "SBIRT" is the answer? Probably not. Addiction 2015 Sep;110(9):1416-1417. [doi: 10.1111/add.12986] [Medline: 26223169]
5. Olmstead TA, Yonkers KA, Ondersma SJ, Forray A, Gilstad-Hayden K, Martino S. Cost-effectiveness of electronic- and clinician-delivered screening, brief intervention and referral to treatment for women in reproductive health centers. Addiction 2019 Sep;114(9):1659-1669. [doi: 10.1111/add.14668] [Medline: 31111591]
6. Trends & statistics: costs of substance abuse. NIH National Institute on Drug Abuse. 2020 Jun. URL: https://nida.nih.gov/research-topics/trends-statistics
7. Substance Abuse and Mental Health Services Administration (US); Office of the Surgeon General (US). Vision for the future: A public health approach. In: Facing Addiction in America: The Surgeon General's Report on Alcohol, Drugs, and Health: US Department of Health and Human Services; 2016. URL: https://www.ncbi.nlm.nih.gov/books/NBK424861/ [accessed 2024-12-15]
8. Giansanti D. Artificial intelligence in public health: current trends and future possibilities. Int J Environ Res Public Health 2022 Sep 21;19(19):11907. [doi: 10.3390/ijerph191911907] [Medline: 36231208]
9. Tudor Car L, Dhinagaran DA, Kyaw BM, et al. Conversational agents in health care: scoping review and conceptual analysis. J Med Internet Res 2020 Aug 7;22(8):e17158. [doi: 10.2196/17158] [Medline: 32763886]
10. OpenAI. ChatGPT (Mar 14 version) Large language model. 2023. URL: https://chat.openai.com/chat [accessed 2024-04-15]
11. Singh B, Olds T, Brinsley J, et al. Systematic review and meta-analysis of the effectiveness of chatbots on lifestyle behaviours. NPJ Digit Med 2023 Jun 23;6(1):118. [doi: 10.1038/s41746-023-00856-1] [Medline: 37353578]
12. Lamb J, Israelstam G, Agarwal R, Bakster S. Generative AI in healthcare: adoption trends and what's next. : McKinsey & Company; 2024. URL: https://www.mckinsey.com/industries/healthcare/our-insights/generative-ai-in-healthcare-adoption-trends-and-whats-next [accessed 2024-11-11]
13. Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. Healthcare (Basel) 2023 Mar 19;11(6):887. [doi: 10.3390/healthcare11060887] [Medline: 36981544]
14. Miner AS, Laranjo L, Kocaballi AB. Chatbots in the fight against the COVID-19 pandemic. NPJ Digit Med 2020;3:65. [doi: 10.1038/s41746-020-0280-0] [Medline: 32377576]
15. Hamdoun S, Monteleone R, Bookman T, Michael K, Michael K. AI-based and digital mental health apps: balancing need and risk. IEEE Technol Soc Mag 2023;42(1):25-36. [doi: 10.1109/MTS.2023.3241309]

XSL•FO
RenderX

16. Menz BD, Modi ND, Sorich MJ, Hopkins AM. Health disinformation use case highlighting the urgent need for artificial intelligence vigilance: weapons of mass disinformation. JAMA Intern Med 2024 Jan 1;184(1):92-96. [doi: 10.1001/jamainternmed.2023.5947] [Medline: 37955873]
17. Aljamaan F, Temsah MH, Altamimi I, et al. Reference hallucination score for medical artificial intelligence chatbots: development and usability study. JMIR Med Inform 2024 Jul 31;12:e54345. [doi: 10.2196/54345] [Medline: 39083799]
18. Hosseini M, Rasmussen LM, Resnik DB. Using AI to write scholarly publications. Account Res 2024 Oct;31(7):715-723. [doi: 10.1080/08989621.2023.2168535] [Medline: 36697395]
19. Yang J, Chen YL, Por LY, Ku CS. A systematic literature review of information security in chatbots. Appl Sci (Basel) 2023;13(11):6355. [doi: 10.3390/app13116355]
20. Sebastian G. Privacy and data protection in ChatGPT and other AI chatbots: strategies for securing user information. International Journal of Security and Privacy in Pervasive Computing 2023;15:1-14. [doi: 10.4018/IJSPPC.325475]
21. Li J. Security implications of AI chatbots in health care. J Med Internet Res 2023 Nov 28;25:e47551. [doi: 10.2196/47551] [Medline: 38015597]
22. Scheibein F, Caballeria E, Taher MA, et al. Optimizing digital tools for the field of substance use and substance use disorders: backcasting exercise. JMIR Hum Factors 2023 Dec 12;10:e46678. [doi: 10.2196/46678] [Medline: 38085569]
23. Bull S, Ezeanochie N. From Foucault to Freire through Facebook: toward an integrated theory of mHealth. Health Educ Behav 2016 Aug;43(4):399-411. [doi: 10.1177/1090198115605310] [Medline: 26384499]
24. Collins-McNeil J, Edwards CL, Batch BC, Benbow D, McDougald CS, Sharpe D. A culturally targeted self-management program for African Americans with type 2 diabetes mellitus. Can J Nurs Res 2012 Dec;44(4):126-141 [FREE Full text] [Medline: 23448079]
25. Tversky A, Kahneman D. The framing of decisions and the psychology of choice. Science 1981 Jan 30;211(4481):453-458. [doi: 10.1126/science.7455683] [Medline: 7455683]
26. Thaler RH, Sunstein CR. Nudge: Improving Decisions About Health, Wealth, and Happiness Revised & Expanded Edition: Penguin Books; 2009.
27. Berger J. Contagious: Why Things Catch On: Simon & Schuster; 2013. URL: https://jonahberger.com/books/contagious/ [accessed 2025-04-25]
28. Earnshaw VA. Stigma and substance use disorders: a clinical, research, and advocacy agenda. Am Psychol 2020 Dec;75(9):1300-1311. [doi: 10.1037/amp0000744] [Medline: 33382299]
29. Ybarra ML, Prescott TL, Phillips GL, Bull SS, Parsons JT, Mustanski B. Pilot RCT results of an mHealth HIV prevention program for sexual minority male adolescents. Pediatrics 2017 Jul;140(1):e20162999. [doi: 10.1542/peds.2016-2999] [Medline: 28659456]
30. Bull S, Devine S, Schmiege SJ, Pickard L, Campbell J, Shlay JC. Text messaging, teen outreach program, and sexual health behavior: a cluster randomized trial. Am J Public Health 2016 Sep;106(S1):S117-S124. [doi: 10.2105/AJPH.2016.303363] [Medline: 27689478]
31. Domek GJ, Contreras-Roldan IL, Bull S, et al. Text message reminders to improve infant immunization in Guatemala: a randomized clinical trial. Vaccine (Auckl) 2019 Sep 30;37(42):6192-6200. [doi: 10.1016/j.vaccine.2019.08.046] [Medline: 31492475]
32. IBM. Watson privacy, compliance & security proof of value (POV) documentation. 2018 URL: https://www.ibm.com/watson/assets/duo/pdf/Watson-Privacy-and-Security-POV_final_062819_tps.pdf [accessed 2024-12-15]
33. Williamson SM, Prybutok V. The era of artificial intelligence deception: unraveling the complexities of false realities and emerging threats of misinformation. Information 2024;15(6):299. [doi: 10.3390/info15060299]
34. Nielsen J, Landauer TK. A mathematical model of the finding of usability problems. CHI '93: Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems 1993 May 1:206-213. [doi: 10.1145/169059.169166]
35. Beattie H, Watkins L, Robinson WH, Rubin A, Watkins S. Measuring and mitigating bias in AI-chatbots. 2022 Presented at: 2022 IEEE International Conference on Assured Autonomy (ICAA); Fajardo, PR, USA p. 117-123. [doi: 10.1109/ICAA52185.2022.00023]

Abbreviations

AI: artificial intelligence
AWS: Amazon Web Server
HIPAA: Healthcare Insurance Portability and Accountability Act
LLM: large language model
ML: machine learning
NLP: natural language processing
PHI: protected health information
SUD: substance use disorder

XSL•FO
RenderX

Edited by A Choudhury; submitted 22.11.24; peer-reviewed by M Krishnapatnam, S Kumar; revised version received 04.03.25; accepted 06.03.25; published 07.05.25.

Please cite as:
Salyers AJ, Bull S, Silvasstar J, Howell K, Wright T, Banaei-Kashani F
Building and Beta-Testing Be Well Buddy Chatbot, a Secure, Credible and Trustworthy AI Chatbot That Will Not Misinform, Hallucinate or Stigmatize Substance Use Disorder: Development and Usability Study
JMIR Hum Factors 2025;12:e69144
URL: https://humanfactors.jmir.org/2025/1/e69144
doi:10.2196/69144

© Adam Jerome Salyers, Sheana Bull, Joshva Silvasstar, Kevin Howell, Tara Wright, Farnoush Banaei-Kashani. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 7.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The mChoice App, an mHealth Tool for the Monitoring of Preexposure Prophylaxis Adherence and Sexual Behaviors in Young Men Who Have Sex With Men: Usability Evaluation

Fabiana Cristina Dos Santos[1], MSN, RN, PhD; Maeve Brin[1], BA; Mary R Tanner[2], MD; Carla A Galindo[2], MPH; Rebecca Schnall[1], MPH, RN, PhD

[1]School of Nursing, Columbia University, 560 W 168th St, New York, NY, United States

[2]Division of HIV Prevention, Centers for Disease Control and Prevention, Atlanta, GA, United States

Corresponding Author:
Fabiana Cristina Dos Santos, MSN, RN, PhD
School of Nursing, Columbia University, 560 W 168th St, New York, NY, United States

Abstract

Background: Mobile health (mHealth) apps provide easy and quick access for end users to monitor their health-related activities. Features such as medication reminders help end users adhere to their medication schedules and automatically record these actions, thereby helping manage their overall health. Due to insufficient mHealth tools tailored for HIV preventive care in young men who have sex with men (MSM), our study evaluated the usability of the mChoice app, a tool designed to enhance preexposure prophylaxis (PrEP) adherence and promote sexual health (eg, encouraging the use of condoms and being aware of the partner's HIV status and PrEP use).

Objective: This study aimed to apply systematic usability evaluations to test the mChoice app and to refine the visualizations to better capture and display patient-reported health information.

Methods: Usability testing involved heuristic evaluations conducted with 5 experts in informatics and user testing with 20 young MSM who were taking or were eligible to take PrEP.

Results: End users demonstrated satisfaction with the appearance of the mChoice app, reporting that the app has an intuitive interface to track PrEP adherence. However, participants highlighted areas needing improvement, including chart titles and the inclusion of "undo" and "edit" buttons to improve user control when recording PrEP use.

Conclusions: Usability evaluations involving heuristic experts and end users provided valuable insights into the mChoice app's design. Areas for improvement were identified, such as enhancing chart readability and providing additional user controls. These findings will guide iterative refinements, ensuring that future versions of the app better address the needs of its target audience and effectively support HIV prevention.

(JMIR Hum Factors 2025;12:e59780) doi:10.2196/59780

KEYWORDS

HIV prevention; data visualization; patient-reported health information; mHealth; digital health; usability; human immunodeficiency virus; preexposure prophylaxis; men who have sex with men; apps; HIV; PrEP

Introduction

The incidence of HIV, although having declined modestly overall in the United States, remains a significant public health problem for men who have sex with men (MSM) [1,2]. Notably, young MSM aged 13 to 34 years represent 58% of the estimated HIV infections in 2021, indicating a pressing need for targeted intervention in this group [2]. Preexposure prophylaxis (PrEP) is a highly effective biomedical prevention strategy to reduce HIV incidence and curb the HIV epidemic [3-10]. When taken as prescribed (daily or at least 4 times per week), PrEP reduces HIV transmission by up to 99% among MSM [7-9,11-13]. While awareness of PrEP has increased among MSM [14,15], those who are most disproportionately affected have been found to be less aware of PrEP [15], and overall, widespread uptake among those with the greatest indications for PrEP remains low [14-17].

To overcome this clinical and public health challenge, technological interventions, such as mobile health (mHealth) apps, have emerged to support public health care initiatives [18]. However, despite the potential of mHealth for delivering sexual health and HIV prevention awareness, mHealth is underutilized for supporting PrEP adherence [19,20]. In response, our study team developed mChoice, an innovative PrEP adherence monitoring app, which combines the CleverCap smart pill bottle with WiseApp, allowing end users to

XSL•FO
RenderX

self-monitor and manage their medication adherence. The CleverCap pill bottle dispenses the prescribed medication and interfaces with WiseApp for real-time tracking. The app was created by Compliance Meds Technology and has been used in prior studies among persons living with HIV [10,21-25]. It was adapted for the mChoice study. The app's key functionalities include monitoring PrEP adherence, visualizing data trends, and documenting sexual behavior, particularly among young MSM from diverse backgrounds.

A notable issue in the proliferation of mHealth apps is the development of this technology with minimal input from end users, leading to poor design, inadequate consideration of user needs, and ultimately, poor usability [26]. Poorly designed apps, lacking in usability, are prone to misuse, underutilization, and failure to achieve their intended objectives.

To ensure the best utilization of mHealth apps, it is essential to understand their usability, keeping in mind target end users (eg, young MSM), tasks (eg, PrEP adherence management and sexual health tracking), and cultural contexts (eg, language and beliefs). Adherence to PrEP is crucial for its effectiveness in preventing HIV, yet many young MSM face challenges such as inconsistent medication schedules and limited access to tailored support tools. The mChoice app was designed to address these barriers by providing intuitive features, including medication reminders, adherence tracking, and functions to support sexual health decision-making. In this study, we sought to apply systematic usability evaluations to test the mChoice app and to refine the functions to better capture and display patient-reported health information.

Methods

Ethical Considerations

The Columbia University Institutional Review Board reviewed and approved all study activities (approval number AAAT8812). Participants signed informed consent forms prior to participating in usability testing. In our dataset, each participant was assigned an identifier, and information was deidentified at the point of analysis.

Heuristic Evaluation

Sample

Usability testing of the mChoice app included heuristic evaluation conducted with experts in informatics and end-user testing conducted with young MSM. Five experts were recruited as evaluators in accordance with Nielsen's recommendation to include 3 to 5 experts in usability testing [27]. Of the 5 experts, 4 had a PhD in Nursing with expertise in human-computer interaction, interface design, and usability testing. The experience of the experts varied between 7 and 23 years, with several impactful publications in the field of informatics.

Procedure

The experts were provided with the mChoice app and asked to complete a session in the app through case scenarios that represented the main functions of the system and think-aloud methodology [28]. The mChoice app required a password, and all study data were encrypted and stored on secure HIPAA-compliant servers at Columbia University. The experts were asked to describe what they were thinking, seeing, and doing as they completed the following 10 scenarios: (1) log in to the CleverCap app; (2) complete a sexual activity log; (3) mark your PrEP dose for today as "Taken"; (4) view tomorrow's pending PrEP dose; (5) edit your sexual activity log; (6) delete your sexual activity log; (7) view information on PrEP dosing over the past month; (8) send a chat to the study team; (9) watch a video about an HIV story; and (10) find PrEP information (see Figures 1 and 2 illustrating app functions and Figure 3 demonstrating the CleverCap smart pill bottle). Following their use of the app, the experts completed an online Heuristic Evaluation Checklist to assess how the system adhered to Nielsen's 10 usability principles [27] using Qualtrics XM, an online survey software. Each question ranged from 0 (not a usability problem) to 4 (usability catastrophe). A member of the research team analyzed the experts' comments regarding usability problems to identify areas of usability concern that could be targeted for improvement. The mean usability problem severity scores were calculated for each of the 10 usability heuristics.

XSL•FO
RenderX

Figure 1. Sexual activity log. A confidential and secure feature for users to record their sexual activities, such as gender of partner, whether a condom was used, and the HIV status of their partner. Users of the on-demand/intermittent (2-1-1) preexposure prophylaxis (PrEP) regimen can use this feature to trigger subsequent PrEP dose reminders following sexual activity (ie, one pill 24 hours and one pill 48 hours after sexual activity).

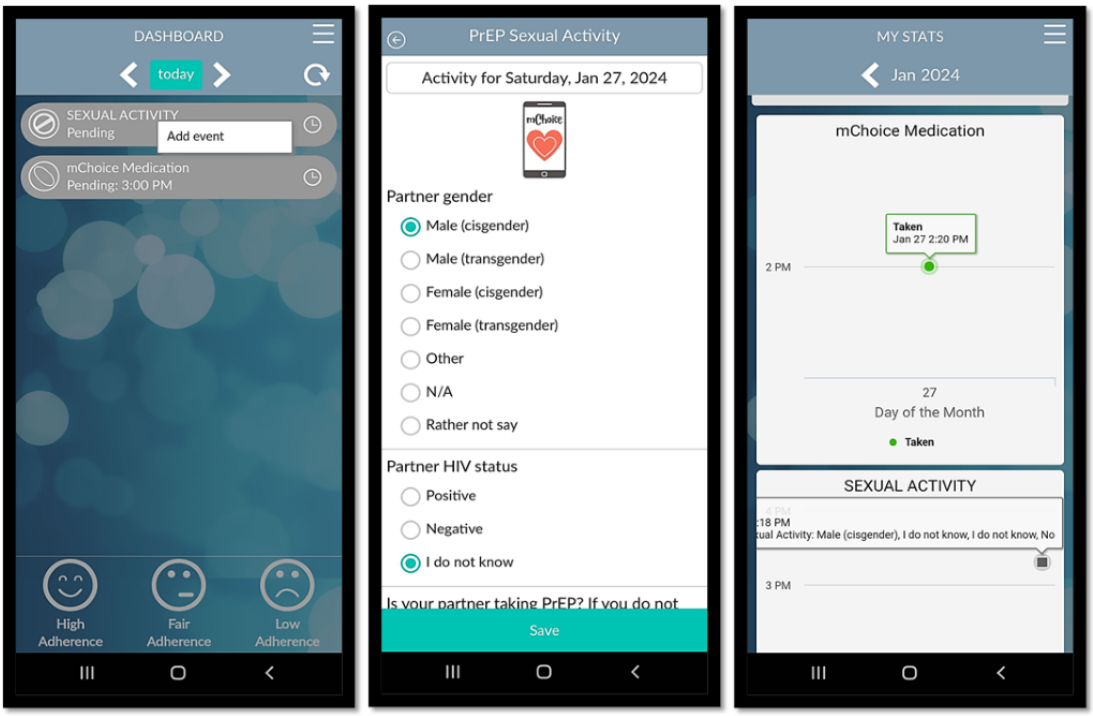
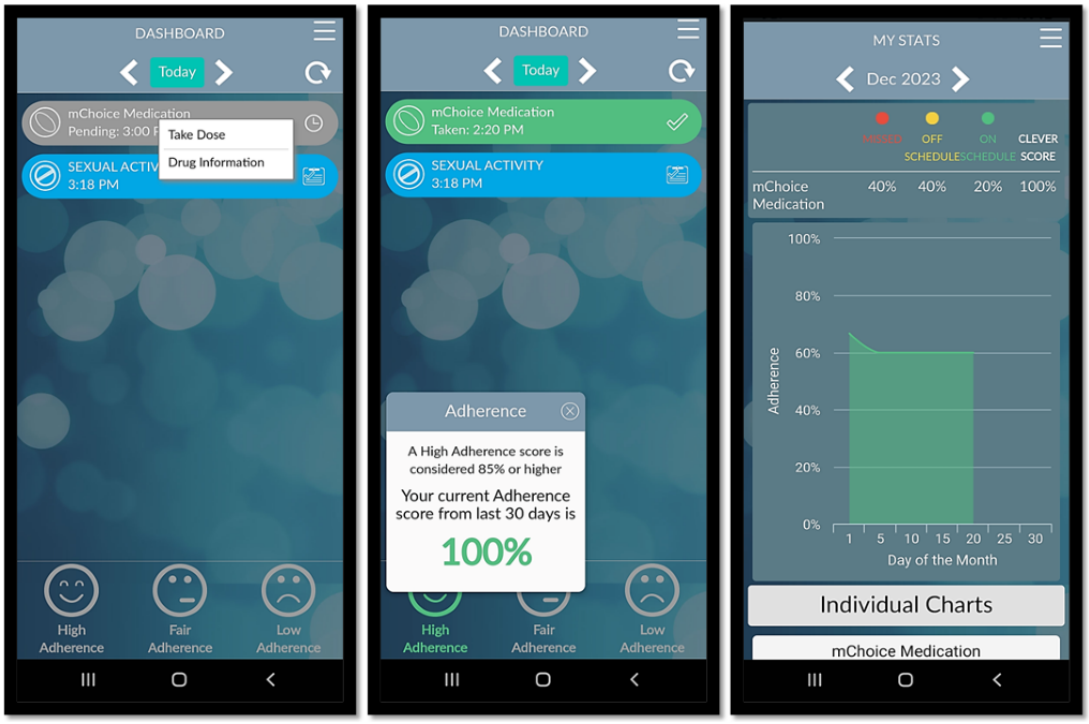


Figure 2. Preexposure prophylaxis (PrEP) adherence tracking. A simple interface for users to track PrEP intake, facilitating adherence monitoring over time. This feature is linked with the CleverCap device, a smart pill bottle that tracks when users take their medication. PrEP tracking is also customizable for users on different PrEP regimens (ie, 2-1-1, injectable, or daily oral dose) to ensure the timely intake of PrEP medication. For instance, users on injectable PrEP will be alerted of their upcoming appointment or need for their PrEP injection.



XSL•FO
RenderX

Figure 3. CleverCap device. A smart pill bottle linked with the mChoice app to track preexposure prophylaxis (PrEP) medication.



End-User Testing

Sample

Twenty end users were recruited for the usability testing. The eligibility criteria were young MSM between 18 and 39 years old, English speaking, living in New York or Alabama, and taking or eligible to take PrEP. The inclusion criteria were designed to ensure the mChoice app was evaluated by its intended target audience, ensuring a diverse usability evaluation. Potential end users were contacted using a research database and invited to participate in the study.

Procedure

During usability testing, the end users were provided with a brief explanation of the mChoice app and its functions. The end users were directed to complete the same 10 scenarios completed by the experts in the heuristic evaluation. Sessions took place remotely via Zoom meetings (Zoom Communications, Inc), and a member of the team was present during sessions to provide guidance when an end user was unable to move through a task independently and to take notes. Following their use of the app, the end users completed a survey using the Qualtrics XM software.

The survey included demographics and two validated measures of usability. The first usability measure was the self-reported ease of use measured by using the Health Information Technology Usability Evaluation Scale (Health-ITUES) [29,30]. This 20-item tool is designed to support customization at the item level to match the health information technology while retaining standardization at the construct level. It has been demonstrated to be useful for evaluating the usability of mHealth [31]. It is scored on a 5-point Likert scale ranging from 5 (strongly agree) to 1 (strongly disagree), where higher scores indicate a system that is easier to use. The second measure was the short version of the Post-Study System Usability Questionnaire (PSSUQ), a 16-item survey that assesses users' perceived satisfaction with a system [32]. Scoring is based on a 7-point Likert scale ranging from 1 (strongly agree) to 7 (strongly disagree). Lower scores indicate satisfaction with the system. Through validated measures, the end users contributed feedback on the app's functionality and user experience.

Data analysis was performed using R statistical software (version 4.1.2; R Foundation for Statistical Computing) to analyze the usability measures. The mean scores were calculated for each survey.

Results

Heuristic Evaluation

Overall, the severity scores ranged from 0.8 to 2.2 for the 10 items on the heuristic checklist, where scores closest to 0 indicate a more usable system. The mean severity scores for each heuristic item are presented in Table 1. The area identified as the most in need of improvement was "user control and freedom," where the experts identified that the app did not have an "undo" button or "edit" button for the PrEP medication entry.

The next most identified area for improvement was the "visibility of the system status," for which experts identified the lack of one "add" button to add a new sexual event. To improve the "visibility of the system status," the experts suggested we provide one "edit" button for modifying recorded sexual encounters and incorporate a separate "add" button for logging new encounters, especially in instances of multiple events. Additionally, "consistency and standards" was an area of concern with difficult terminology in charts. "Error prevention" was another concern, and experts highlighted the importance of one "edit" button to edit medication time as an issue that might impede system usability. Experts provided favorable feedback on the app's user experience. One expert (Expert 2) reported, "the app is very simple and intuitive interface…and it's nice that the button changes color with pleasant colors."

The rating score ranged from 0 (ie, best) to 4 (ie, worst), with a score of 0 indicating no usability problem, 1 indicating a cosmetic problem only, 2 indicating a minor usability problem, 3 indicating a major usability problem, and 4 indicating a usability catastrophe.

XSL•FO
RenderX

Table . Mean severity scores for items on the heuristic evaluation.

Heuristic principles	Mean (SD) severity score	Experts' comments to identify the areas of usability concerns
User control and freedom	2.2 (1.3)	• "No undo or edit for the medication entry. This is a big problem." (Expert 5)
Visibility of the system status	1.6 (1.1)	• "…When adding a sexual encounter record, it is not clear whether it edits or adds a new one, so there should be two buttons, one for edit and one to add a new one. The screen when you use edit also shows the old entry, which implies you are editing it, not starting a new one, but actually it is starting a new one. That is confusing." (Expert 3)
Consistency and standards	1.6 (1.1)	• "Not clear if the clever score is just adherence or if it means something else." (Expert 4) • "…I think in the menu it says stats. Rather than stats, it'd be adherence stats or adherence tracking over time. You know something that makes it really clear what people are clicking on." (Expert 2)
Error prevention	1.6 (1.1)	• "When entering the time of a medication, it might be easy to miss the AM/PM change. I don't think there's any design change to be made for this, but rather the user needs to be able to edit their entry if they forget to adjust this." (Expert 4)
Match between the system and the real world	1.4 (1.3)	• "Would like to see one section specific to PrEP information." (Expert 5)
Recognition rather than recall	1.4 (0.9)	• "Some of the main features can be moved to more obvious places." (Expert 1) • "Organization for videos could be improved, especially if you plan to add more content." (Expert 4)
Help and documentation	1.4 (1.3)	• "Information buttons and an introduction to the app would be helpful." (Expert 1)
Help users recognize, diagnose, and recover from errors	1.2 (1.1)	• "When the dose was missed, no way to edit if it was a mistype." (Expert 5)
Aesthetic and minimalist design	1.2 (1.1)	• "Graphs need clearer tiles and more labels." (Expert 1) • "Need more medication information or guidance if adherence is low." (Expert 5)
Flexibility and efficiency of use	0.8 (1.1)	• "…Search function for videos. Tags for topics would be good, too." (Expert 4)

End-User Testing

The end-user group for evaluating the mChoice app's usability consisted entirely of men, with the majority identifying as homosexual (14/20, 70%). The mean (SD) age was 28 (3.4) years. Half of the end users were White, and 75% (15/20) were non-Hispanic or Latino. Considering education, 50% (10/20) had a college degree. Two specific measures were employed for the usability assessment: the Health-ITUES and the PSSUQ, as detailed in Table 2.

The Health-ITUES scale, where a higher score indicates better usability, showed that "perceived ease of use" scored the highest with a mean (SD) score of 4.6 (0.5), suggesting a favorable user assessment. However, "user control" received the lowest score in this category, with a mean (SD) score of 3.8 (0.9), corroborating the heuristic evaluator's ratings.

Conversely, the PSSUQ, where a lower score indicates better usability, reflected different aspects of the mChoice performance. "System quality" scored the lowest with a mean

XSL•FO
RenderX

(SD) of 1.7 (1.2), while "interface quality" had the highest score with a mean (SD) score of 3.0 (1.8). These results suggested a positive user experience with notable ease of use and system quality strengths. However, there were areas for improvement, particularly in enhancing user control and interface quality.

Table . Usability measures.

Measures	Mean (SD) score
Health-ITUES[a]	
Perceived ease of use	4.6 (0.5)
Impact	4.1 (0.8)
Perceived usefulness	4.0 (0.8)
User control	3.8 (0.9)
Overall Health-ITUES score	4.2 (0.4)
PSSUQ[b]	
System quality	1.7 (1.2)
Information quality	2.4 (1.3)
Interface quality	3.0 (1.8)
Overall PSSUQ score	2.3 (0.6)

[a]Health-ITUES: Health Information Technology Usability Evaluation Scale (rating the score from 5 being the best score to 1 being the worst score).

[b]PSSUQ: Post-Study System Usability Questionnaire (rating the score from 1 being the best score to 7 being the worst score).

Discussion

Principal Results

Our study evaluated the mChoice app's usability in enhancing PrEP adherence for young MSM. Overall, heuristic experts and end users demonstrated satisfaction with the mChoice app, reporting that it is a simple and intuitive interface for tracking PrEP adherence. Graphs, charts, and icons are the main features of mHealth apps that help end users track their goals, habits, and achievements [33,34]. The mChoice app combines iconography (eg, emoticons) and color-coded charts to provide clear and immediate feedback on PrEP adherence. These diverse visualizations allow end users to interpret their data in several formats. The app includes emoticons in different colors (green for high adherence, red for missed doses, and yellow for off-schedule doses) to represent adherence status. For instance, a sad emoticon in red for low adherence might highlight medication adherence challenges and prompt self-reflection on areas needing improvement. Such immediate visual feedback can motivate end users to improve their medication routine. In addition, graphs showed the trend of PrEP medication adherence over time, including evidence-based information on the percentage of missed, off-schedule, and on-time doses.

While prior research has indicated that varied data visualization options help match different end users' preferences [33] and color-coded visual cues can aid in quick decision-making and interpretation of health data [35-38], our study did not directly assess end-user interpretations of these visual elements. The mChoice app incorporates these evidence-based design principles, but further research would be needed to evaluate their impact on user comprehension and medication adherence behaviors.

Our findings suggest that the visual elements available in the mChoice app offer a positive end-user experience and function accessibility that could assist in timely medication intake and monitor adherence. However, our heuristic evaluation highlighted areas needing improvement, particularly in the need for clearer chart titles and more labels to exemplify each component (eg, Clever score), underscoring the importance of refinement of the mChoice design for optimized end-user experiences.

The accuracy of the information collected through the mChoice app was another area of attention that heuristic experts raised concerns, particularly highlighting the absence of "undo," "edit," and "add" buttons. These limitations restricted user control, potentially affecting the accuracy of patients' health information input. For instance, the absence of "undo" and "edit" buttons could lead to inaccuracies if end users mistakenly enter the time of their PrEP dose intake. Additionally, the lack of an "add" button for recording multiple sexual events was seen as an important barrier, restricting end users from documenting their sexual event entries accurately. These findings underscore the need for enhanced app capabilities to ensure accurate and comprehensive user input and customizations. Similarly, other studies [39,40] reported concerns about user experiences and data accuracy in mobile apps, reinforcing the findings of our study.

The actionable finding from this study is to reevaluate the functions of the mChoice app after integrating end-user feedback and addressing heuristic violations. Iterative usability evaluations should be considered as a best practice in the app development process to ensure the effectiveness of the app's future iterations.

XSL•FO
RenderX

Limitations

A limitation of our study was the convenience sample, recruited from a database that included participants who had participated in past research and expressed interest in future studies. Additionally, while our sample targeted young MSM, English-speaking men, and those living in New York or Alabama, this demographic representation might limit our ability to understand how the app might perform for medication adherence across diverse populations. Given that PrEP users come from varied cultural, linguistic, and socioeconomic backgrounds, our findings may not be generalizable to groups whose cultural beliefs, language preferences, and health care engagement patterns differ from those of our study population. While end users were central for evaluating the usability of the mChoice app through validated quantitative measures, qualitative data were collected only from experts during heuristic evaluations. In this study, we did not evaluate the artificial intelligence or algorithmic components that could enhance personalization and medication adherence predictions. Despite these limitations, we identified areas of improvement to refine the mChoice app. In addition, this study enriched the body of research on evaluating the mHealth usability for HIV prevention. This was achieved by employing end users and heuristic experts to assess this innovative app designed for PrEP adherence and HIV prevention.

Conclusions

Our usability study of the mChoice app involved rigorous evaluations through interactive heuristic evaluations and end-user testing. The results demonstrated the app's simplicity and user-friendly interface, showcasing its potential in monitoring health-related activities such as PrEP adherence. Further research could explore how these usability enhancements might influence user engagement and behavior change to support HIV prevention efforts.

Acknowledgments

Research reported in this study was funded by the Centers for Disease Control and Prevention through grant #U01PS005229. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Conflicts of Interest

None declared.

References

1. HIV & AIDS trends and U.S. statistics overview. HIV.gov. URL: https://www.hiv.gov/hiv-basics/overview/data-and-trends/statistics [accessed 2024-02-01]
2. Estimated HIV incidence and prevalence in the United States, 2017–2021. Centers for Disease Control and Prevention. 2023. URL: http://www.cdc.gov/hiv/library/reports/hiv-surveillance.html [accessed 2023-11-12]
3. Carnegie NB, Goodreau SM, Liu A, et al. Targeting pre-exposure prophylaxis among men who have sex with men in the United States and Peru: partnership types, contact rates, and sexual role. J Acquir Immune Defic Syndr 2015 May 1;69(1):119-125. [doi: 10.1097/QAI.0000000000000555] [Medline: 25942463]
4. Juusola JL, Brandeau ML, Owens DK, Bendavid E. The cost-effectiveness of preexposure prophylaxis for HIV prevention in the United States in men who have sex with men. Ann Intern Med 2012 Apr 17;156(8):541-550. [doi: 10.7326/0003-4819-156-8-201204170-00001] [Medline: 22508731]
5. Kessler J, Myers JE, Nucifora KA, et al. Evaluating the impact of prioritization of antiretroviral pre-exposure prophylaxis in New York. AIDS 2014 Nov 28;28(18):2683-2691. [doi: 10.1097/QAD.0000000000000460] [Medline: 25493594]
6. Zablotska IB. Likely impact of pre-exposure prophylaxis on HIV epidemics among men who have sex with men. Sex Health 2017 Feb;14(1):97-105. [doi: 10.1071/SH16153] [Medline: 27883309]
7. McCormack S, Dunn DT, Desai M, et al. Pre-exposure prophylaxis to prevent the acquisition of HIV-1 infection (PROUD): effectiveness results from the pilot phase of a pragmatic open-label randomised trial. Lancet 2016 Jan 2;387(10013):53-60. [doi: 10.1016/S0140-6736(15)00056-2] [Medline: 26364263]
8. Grant RM, Anderson PL, McMahan V, et al. Uptake of pre-exposure prophylaxis, sexual practices, and HIV incidence in men and transgender women who have sex with men: a cohort study. Lancet Infect Dis 2014 Sep;14(9):820-829. [doi: 10.1016/S1473-3099(14)70847-3] [Medline: 25065857]
9. Volk JE, Marcus JL, Phengrasamy T, et al. No new HIV infections with increasing use of HIV preexposure prophylaxis in a clinical practice setting. Clin Infect Dis 2015 Nov 15;61(10):1601-1603. [doi: 10.1093/cid/civ778] [Medline: 26334052]
10. Dos Santos FC, Batey DS, Kay ES, et al. The effect of a combined mHealth and community health worker intervention on HIV self-management. J Am Med Inform Assoc 2025 Mar 1;32(3):510-517. [doi: 10.1093/jamia/ocae322] [Medline: 39798152]
11. Effectiveness of prevention strategies to reduce the risk of acquiring or transmitting HIV. Centers for Disease Control and Prevention. 2022. URL: https://www.cdc.gov/hiv/risk/estimates/preventionstrategies.html#anchor_1562942347 [accessed 2024-05-03]

XSL•FO
RenderX

12. Liu AY, Cohen SE, Vittinghoff E, et al. Preexposure prophylaxis for HIV infection integrated with municipal- and community-based sexual health services. JAMA Intern Med 2016 Jan;176(1):75-84. [doi: 10.1001/jamainternmed.2015.4683] [Medline: 26571482]
13. Marcus JL, Hurley LB, Nguyen DP, Silverberg MJ, Volk JE. Redefining human immunodeficiency virus (HIV) preexposure prophylaxis failures. Clin Infect Dis 2017 Oct 30;65(10):1768-1769. [doi: 10.1093/cid/cix593] [Medline: 29020235]
14. Grov C, Rendina HJ, Whitfield THF, Ventuneac A, Parsons JT. Changes in familiarity with and willingness to take preexposure prophylaxis in a longitudinal study of highly sexually active gay and bisexual men. LGBT Health 2016 Aug;3(4):252-257. [doi: 10.1089/lgbt.2015.0123] [Medline: 27183232]
15. Mayer KH, Agwu A, Malebranche D. Barriers to the wider use of pre-exposure prophylaxis in the United States: a narrative review. Adv Ther 2020 May;37(5):1778-1811. [doi: 10.1007/s12325-020-01295-0] [Medline: 32232664]
16. Dawit R, Predmore Z, Raifman J, et al. Identifying HIV PrEP attributes to increase PrEP use among different groups of gay, bisexual, and other men who have sex with men: a latent class analysis of a discrete choice experiment. AIDS Behav 2024 Jan;28(1):125-134. [doi: 10.1007/s10461-023-04131-y] [Medline: 37474623]
17. Dubov A, Ogunbajo A, Altice FL, Fraenkel L. Optimizing access to PrEP based on MSM preferences: results of a discrete choice experiment. AIDS Care 2019 May;31(5):545-553. [doi: 10.1080/09540121.2018.1557590] [Medline: 30554519]
18. Global strategy on digital health 2020 - 2025. World Health Organization. 2021. URL: https://iris.who.int/handle/10665/344249 [accessed 2024-02-01]
19. Liu AY, Laborde ND, Coleman K, et al. DOT diary: developing a novel mobile app using artificial intelligence and an electronic sexual diary to measure and support PrEP adherence among young men who have sex with men. AIDS Behav 2021 Apr;25(4):1001-1012. [doi: 10.1007/s10461-020-03054-2] [Medline: 33044687]
20. Whiteley L, Mena L, Craker LK, Healy MG, Brown LK. Creating a theoretically grounded gaming app to increase adherence to pre-exposure prophylaxis: lessons from the development of the viral combat mobile phone game. JMIR Serious Games 2019 Mar 27;7(1):e11861. [doi: 10.2196/11861] [Medline: 30916652]
21. Schnall R, Rojas M, Bakken S, et al. A user-centered model for designing consumer mobile health (mHealth) applications (apps). J Biomed Inform 2016 Apr;60:243-251. [doi: 10.1016/j.jbi.2016.02.002] [Medline: 26903153]
22. Aliabadi N, Carballo-Dieguez A, Bakken S, et al. Using the information-motivation-behavioral skills model to guide the development of an HIV prevention smartphone application for high-risk MSM. AIDS Educ Prev 2015 Dec;27(6):522-537. [doi: 10.1521/aeap.2015.27.6.522] [Medline: 26595265]
23. Schnall R, Travers J, Rojas M, Carballo-Diéguez A. eHealth interventions for HIV prevention in high-risk men who have sex with men: a systematic review. J Med Internet Res 2014 May 26;16(5):e134. [doi: 10.2196/jmir.3393] [Medline: 24862459]
24. Schnall R, Kuhns LM, Pearson C, et al. Efficacy of MyPEEPS mobile, an HIV prevention intervention using mobile technology, on reducing sexual risk among same-sex attracted adolescent males: a randomized clinical trial. JAMA Netw Open 2022 Sep 1;5(9):e2231853. [doi: 10.1001/jamanetworkopen.2022.31853] [Medline: 36129712]
25. Beauchemin M, Gradilla M, Baik D, Cho H, Schnall R. A multi-step usability evaluation of a self-management app to support medication adherence in persons living with HIV. Int J Med Inform 2019 Feb;122:37-44. [doi: 10.1016/j.ijmedinf.2018.11.012] [Medline: 30623782]
26. McCurdie T, Taneva S, Casselman M, et al. mHealth consumer apps: the case for user-centered design. Biomed Instrum Technol 2012;Suppl:49-56. [doi: 10.2345/0899-8205-46.s2.49] [Medline: 23039777]
27. Nielsen J, Molich R. Heuristic evaluation of user interfaces. Presented at: CHI90: Conference on Human Factors in Computing; Apr 1-5, 1990; Seattle, Washington, United States p. 249-256 URL: https://dl.acm.org/doi/10.1145/97243.97281 [accessed 2025-02-19] [doi: 10.1145/97243.97281]
28. Bertini E, Gabrielli S, Kimani S, Catarci T, Santucci G. Appropriating and assessing heuristics for mobile computing. Presented at: AVI06: The International Conference on Advanced Visual Interfaces; May 23-26, 2006; Venezia, Italy URL: https://dl.acm.org/doi/10.1145/1133265.1133291 [accessed 2025-02-19] [doi: 10.1145/1133265.1133291]
29. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
30. Schnall R, Cho H, Liu J. Health Information Technology Usability Evaluation Scale (Health-ITUES) for usability assessment of mobile health technology: validation study. JMIR Mhealth Uhealth 2018 Jan 5;6(1):e4. [doi: 10.2196/mhealth.8851] [Medline: 29305343]
31. Brown W 3rd, Yen PY, Rojas M, Schnall R. Assessment of the Health IT Usability Evaluation Model (Health-ITUEM) for evaluating mobile health (mHealth) technology. J Biomed Inform 2013 Dec;46(6):1080-1087. [doi: 10.1016/j.jbi.2013.08.001] [Medline: 23973872]
32. Lewis JR. Psychometric evaluation of the PSSUQ using data from five years of usability studies. Int J Hum Comput Interact 2002 Sep;14(3-4):463-488. [doi: 10.1080/10447318.2002.9669130]
33. Caldeira C, Chen Y, Chan L, Pham V, Chen Y, Zheng K. Mobile apps for mood tracking: an analysis of features and user reviews. AMIA Annu Symp Proc 2017;2017:495-504. [Medline: 29854114]

XSL•FO
RenderX

34. Carroll LN, Au AP, Detwiler LT, Fu TC, Painter IS, Abernethy NF. Visualization and analytics tools for infectious disease epidemiology: a systematic review. J Biomed Inform 2014 Oct;51:287-298. [doi: 10.1016/j.jbi.2014.04.006] [Medline: 24747356]
35. Albers EAC, Fraterman I, Walraven I, et al. Visualization formats of patient-reported outcome measures in clinical practice: a systematic review about preferences and interpretation accuracy. J Patient Rep Outcomes 2022 Mar 3;6(1):18. [doi: 10.1186/s41687-022-00424-3] [Medline: 35239055]
36. Kuijpers W, Giesinger JM, Zabernigg A, et al. Patients' and health professionals' understanding of and preferences for graphical presentation styles for individual-level EORTC QLQ-C30 scores. Qual Life Res 2016 Mar;25(3):595-604. [doi: 10.1007/s11136-015-1107-3] [Medline: 26353905]
37. Smith KC, Brundage MD, Tolbert E, et al. Engaging stakeholders to improve presentation of patient-reported outcomes data in clinical practice. Support Care Cancer 2016 Oct;24(10):4149-4157. [doi: 10.1007/s00520-016-3240-0] [Medline: 27165054]
38. Dos Santos FC, Yao Y, Macieira TGR, Dunn Lopez K, Keenan GM. Nurses' preferences for the format of care planning clinical decision support coded with standardized nursing languages. J Am Med Inform Assoc 2023 Oct 19;30(11):1846-1851. [doi: 10.1093/jamia/ocad093] [Medline: 37257882]
39. Gu J, Mackin S, Zheng Y. Making sense: an innovative data visualization application utilized via mobile platform. Presented at: 2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS); Jun 28-30, 2018; Exeter, United Kingdom p. 1105-1109. [doi: 10.1109/HPCC/SmartCity/DSS.2018.00184]
40. Katz D, Dalton N, Holland S, O'Kane A, Price BA. Questioning the reflection paradigm for diabetes mobile apps. In: Giokas K, Bokor L, Hopfgartner F, editors. eHealth 360°: Springer, Cham; 2017:315-326. [doi: 10.1007/978-3-319-49655-9_39]

Abbreviations

Health-ITUES: Health Information Technology Usability Evaluation Scale
mHealth: mobile health
MSM: men who have sex with men
PrEP: preexposure prophylaxis
PSSUQ: Post-Study System Usability Questionnaire

Edited by A Kushniruk; submitted 16.05.24; peer-reviewed by C Grov, D Chrimes, R Ma; revised version received 19.01.25; accepted 19.01.25; published 28.02.25.

Please cite as:
Dos Santos FC, Brin M, Tanner MR, Galindo CA, Schnall R
The mChoice App, an mHealth Tool for the Monitoring of Preexposure Prophylaxis Adherence and Sexual Behaviors in Young Men Who Have Sex With Men: Usability Evaluation
JMIR Hum Factors 2025;12:e59780
URL: https://humanfactors.jmir.org/2025/1/e59780
doi:10.2196/59780

© Fabiana Cristina Dos Santos, Maeve Brin, Mary R Tanner, Carla A Galindo, Rebecca Schnall. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 28.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Use of Wearable Transdermal Alcohol Sensors for Monitoring Alcohol Consumption After Detoxification With Contingency Management: Pilot Randomized Feasibility Trial

Eileen Brobbin[1], PhD; Colin Drummond[1], MD; Stephen Parkin[2], PhD; Paolo Deluca[1], PhD

[1]King's College London, London, United Kingdom

[2]Department of Public Health, Environments and Society at London School of Hygiene and Tropical Medicine, King's College London, London, United Kingdom

Corresponding Author:
Eileen Brobbin, PhD
King's College London
4 Windswor Walk
Denmark Hill
London, SE5 8BB
United Kingdom
Phone: 44 07758287265
Email: eileen.brobbin@kcl.ac.uk

Abstract

Background: Wearable transdermal alcohol sensor (TAS) devices generate continuous data on alcohol consumption through the indiscernible sweat vapors on the skin. This continuous alcohol monitoring capability could provide a new method for alcohol services to monitor service users at various stages of their alcohol treatment.

Objective: We aimed to assess the feasibility of using a TAS as part of alcohol treatment with alcohol service users using the device with or without contingency management (CM).

Methods: A feasibility study was conducted of a convenience sample of 29 current service users from 3 South London alcohol services. Participants were randomized into either a control (treatment as usual) or CM group (treatment as usual+CM). We assessed the feasibility of enrollment, participation, device tampering and return, and device wearability and the accuracy of data capture. These data were reported descriptively where appropriate, the groups were compared, and alcohol self-report data were compared to the transdermal alcohol concentration to assess accuracy.

Results: A total of 34 individuals were approached, and 32 (94%) were enrolled and randomized (n=17, 53% to the control group and n=15, 47% to the CM group) over 5 months. In total, 3 participants withdrew (n=2, 67% from the control group and n=1, 33% from the CM group). There was a total of 203 meetings arranged (29 participants × 7 meetings), and 185 (91.1%) were attended. Only 1 of the 29 participants (3%) admitted to turning the TAS off to avoid monitoring. There were some issues with the TAS not functioning properly and not being able to be cleaned. Removals were recorded, but the definition of TAS removal may need to be improved for future trials. There was a high TAS return rate (28/29, 97% of the participants returned the TAS). Secondary outcomes suggest that the BACtrack Skyn remains an accurate tool to monitor alcohol consumption compared to self-report data and that it is acceptable to wearers over 2 weeks, with many participants (27/28, 96%) answering that they would wear it again and for longer but that the CM procedure could be made clearer.

Conclusions: The delivery of CM via a TAS was feasible in this study, but recommendations for a future larger trial include that the study design should be changed to provide an operationalized rather than manual method of checking whether TAS data meet CM criteria. This would reduce researcher burden and researcher and participant time. Current recruitment and research meeting design seem suitable for a future larger trial.

Trial Registration: International Standard Randomised Controlled Trial Number (ISRCTN) ISRCTN46845361; https://www.isrctn.com/ISRCTN46845361

(JMIR Hum Factors 2025;12:e64664) doi:10.2196/64664

XSL•FO
RenderX

KEYWORDS

alcohol; alcohol dependence; alcohol monitoring; alcohol treatment; contingency management; transdermal alcohol sensor; transdermal technology; wearable; wearable alcohol biosensor

Introduction

Background

Various wearable transdermal alcohol sensor (TAS) devices have been developed. These devices measure alcohol consumption from alcohol vapors in the skin via sweat, known as transdermal alcohol concentration (TAC), and can record at automated regular intervals. One potential use of TASs is as a tool for objective alcohol measurement in clinical alcohol treatment.

TASs could be used in specialist community alcohol services for monitoring alcohol consumption before detoxification to determine drinking levels and patterns, during detoxification, or after detoxification to be combined with sessions with key workers and other psychosocial or medication treatment. The recorded alcohol data could be used to consider triggers and provide evidence of abstinence or alcohol reduction and proof of engagement, which may be useful as evidence for funded treatment, or TASs could be used to implement contingency management (CM) for alcohol reduction. However, before TASs can be implemented into treatment, their accuracy, acceptability, and feasibility in this setting must be validated.

There have been several studies that have investigated the various TAS device brands on TAC data accuracy compared to self-report, blood alcohol concentration, and breath alcohol content data [1-17]. Most of this previous research was conducted with healthy adult volunteers [18,19]. Various brands of TASs have been developed and are in differing stages of validation and research. There have only been a few studies conducted with a specific focus on exploring TAS acceptability and feasibility measures [12,20-26] and even fewer specifically on TAS implementation of CM rewards [1,2,26-31]. While most of the TAS research reports on the SCRAM monitor, there is a growing number of individuals using a newer generation of TASs, the BACtrack Skyn [7,12-14,32-42].

CM has been evidenced as an effective treatment for substance use. It is an established treatment recommended by the National Institute for Health and Care Excellence [43] and has proven to be effective for a range of substance use treatments [44-47]. Although initially developed to be used with alcohol use disorder (AUD) treatment, it has had limited use in this area [48,49]. This is due to the nature of alcohol metabolism and its short detectability within the body [50,51].

Currently used methods for alcohol measurement include breath, blood, and urine tests, which have a relatively short time frame to detect alcohol [52]. Thus, to accurately implement CM in alcohol treatment, the individual would require frequent, multiple breath, blood, or urine tests daily to prove abstinence or alcohol reduction and accurately achieve CM rewards [47,49,53,54]. This is not always feasible with staff time and resources and increases the burden on both service users and staff. The portable mobile phone–linked breathalyzer has also tried to address this limitation and appears to work well in tracking alcohol treatment progress [55-58]. However, it still requires repeated daily compliance from the individual completing each breathalyzer test.

The development of TASs has the potential to address CM implementation with alcohol use and can provide a low-burden, low-intensity solution [1,2,26-31]. Previous literature has started to explore how TASs can implement CM [1,2,26-31]. These studies have found the SCRAM monitor successful in implementing the CM procedure and found that the CM intervention was able to significantly reduce alcohol consumption [1,2,27,29]. Of these 8 studies, none recruited alcohol-dependent participants. In total, 2 involved recent driving while intoxicated offenders with differing criteria on alcohol consumption, 1 with an Alcohol Use Disorders Identification Test score of ≥4 [28] and the other with an Alcohol Use Disorders Identification Test score of ≥8 [30]; 2 recruited HIV-diagnosed individuals with higher levels of alcohol consumption [26,31]; and the other 4 classified participants as risky or heavy drinkers with varying ways to define this [1,2,27,29]. The length of TAS wear periods and CM length ranged from 1 to 4 months, and all studies were conducted in the United States and used the SCRAM monitor.

This study followed a previous study that used the Skyn with individuals in alcohol treatment wearing the device for 1 week and interviewed staff on their thoughts [36,59,60]. From this previous study, we found that most of the individuals were willing to wear the device for longer and that staff want patients wearing the device for longer than a week for it to be used in alcohol treatment. Therefore, for this study, we extended the wear time for 2 weeks and added the CM component. The findings of this study will contribute to the knowledge on TAS feasibility in alcohol research and providing CM in alcohol services in South London.

Objectives

This study aimed to explore the feasibility, strengths, and limitations of using a TAS to monitor alcohol consumption in individuals in treatment for AUD with or without CM to promote abstinence or low-level alcohol consumption.

Methods

This was a randomized controlled trial with a 1:1 allocation ratio to the control and CM groups.

Patient and Public Involvement and Staff Consultation

To ensure that the recruitment, study design, and outcomes were appropriate, when designing this study, we conducted monthly patient and public involvement groups and staff consultation for 6 months. Service users and staff from 3 South London alcohol services were invited to attend. There were 2 meetings held each month, one for service users (in person) and one for

XSL•FO
RenderX

staff (on Microsoft Teams). In these groups, we discussed study aspects and the participants' thoughts on potential challenges.

Setting

In total, 3 drug and alcohol services were recruited from South London and Maudsley National Health Service Foundation Trust. The 3 services were the Pier Road Project (Erith), Wandsworth Community Drug and Alcohol Service (Wandsworth and Richmond), and the Assertive Alcohol Outreach Team (Camberwell). All services were willing to be involved in this study and had recruitment occur at their facilities. At each service, after referral, patients typically completed a community detoxification, were prescribed any medication if appropriate, and started to attend group meetings and regular one-to-one meetings with their key worker. All participants that were recruited for this study were currently attending groups and one-to-one sessions.

Participants and Sample Size

Participants were referred by service staff and by the researcher attending patient groups to discuss the study. One of the aims of this trial was to investigate the feasibility of conducting a larger trial. Therefore, this study conducted no formal statistical sample size calculation but aimed to recruit 30 participants (15 in each arm).

The inclusion criteria were (1) reception of alcohol treatment for an AUD in one of the participating South London alcohol services, (2) age of ≥18 years, (3) ability to speak English competently, (4) ability to meet throughout the study period, (5) no current participation in any other research trials, and (6) willingness to provide informed consent to participate. Exclusion criteria were (1) current use of any illegal or addictive substances (excluding cannabis), (2) age of <18 years, and (3) inability to speak English without a translator.

The study inclusion criteria were intentionally kept broad, enabling individuals receiving any treatment for an AUD (provided that they met the other criteria) to participate.

Randomization

Participants were randomized via dedicated software and sealed envelopes by an independent statistician. The team member recruiting participants (EB) was not aware of the randomization allocation sequence until the sealed numbered envelope was opened at each participant's first research meeting. EB enrolled all participants and assigned participants to their allocated group. The researcher and participants were not blind to the allocation; both were aware of whether they were receiving CM rewards.

Procedure

At enrollment, participants were randomized into one of two groups: (1) treatment as usual+wearing a TAS (control group) or (2) treatment as usual+wearing a TAS+CM for low or no alcohol consumption as measured using the TAS (CM group).

Each participant had 7 research meetings arranged across 15 days. At the first meeting, the participant was trained in using the TAS and provided with a quick leaflet guide to take home with them, had the study protocol explained, and was randomized into a group. Meetings 2 to 6 were for TAS data download and timeline followback (TLFB) completion. At the final meeting, meeting 7, participants also had TAS data downloaded and completed the TLFB as well as a postwear survey. This procedure is shown in Table 1.

Table 1. Study procedure—an example of the research meetings if the first meeting happened on a Monday.

	Meeting description
Week 1	
Monday	Meeting 1—talk participants through the study and train them in using the TAS[a]
Wednesday	Meeting 2—TAS data download and TLFB[b] (+CM[c])
Friday	Meeting 3—TAS data download and TLFB (+CM)
Week 2	
Monday	Meeting 4—TAS data download and TLFB (+CM)
Wednesday	Meeting 5—TAS data download and TLFB (+CM)
Friday	Meeting 6—TAS data download and TLFB (+CM)
Week 3	
Monday	Meeting 7—TAS data download, TLFB (+CM), and postwear survey

[a]TAS: transdermal alcohol sensor.

[b]TLFB: timeline followback.

[c]CM: contingency management.

If the participant was randomized into the CM group, in addition to the aforementioned steps, at each meeting (meetings 2-7), the TAS data were checked to confirm whether the participant had been abstinent or consumed an amount of alcohol below our set threshold and whether they had been wearing the TAS. If they met these 2 criteria, they were then rewarded with the corresponding amount for the days of abstinence or low level of drinking.

The reason for the research meeting design was for regular TAS data download to ensure that the TAS did not start overwriting data. At the time of this study, the Skyn could store up to 72 hours before overwriting data. Participants also completed a

XSL•FO
RenderX

TLFB with the researcher at every meeting, so they only had to recall alcohol consumption for the previous 2 to 3 days [61].

CM Intervention

Participants who were randomized to the CM group could also earn rewards by being abstinent or for low drinking as measured using the TAS and wearing the TAS consistently (we specified that removal of no more than 60 minutes per day was allowed for a shower or bath). The CM reward was a £5 (US $6.25) voucher for each day that the target behavior occurred. There were also bonuses for consecutive days of the target behavior occurring. At each meeting (meetings 2-7), the researcher checked their TAC data and provided any earned CM rewards since the previous meeting. If the participant met the target behavior every single day for the study period, there was an additional bonus given at the end (£35 [US $43.75]). Therefore, a total of £180 (US $225.02) over the study period could be given in CM vouchers. If there was a day in which this behavior did not occur, then participants received no CM for that day and were not eligible for that bonus. The CM plan is shown in Table 2.

Table 2. Contingency management (CM) plan—an example of the CM plan for a Monday start. Participants were ineligible for CM if the device was removed (skin temperature of <30 °C for >1 hour). A 1-time 1-hour removal was allowed per day.

Day number	Day	CM for 1-day abstinence (£5 [US $6.25] per day)[a]	CM bonus for consecutive-day abstinence (£5 [US $6.25] per day)[b]	CM bonus for 14 consecutive days of abstinence (£35 [US $43.75] for 14 days)[c]
1	Monday	£5 (US $6.25)	—[d]	—
2	Tuesday	£5 (US $6.25)	—	—
3	Wednesday	£5 (US $6.25)	£10 (US $12.50; second meeting)	—
4	Thursday	£5 (US $6.25)	—	—
5	Friday	£5 (US $6.25)	£10 (US $12.50; third meeting)	—
6	Saturday	£5 (US $6.25)	—	—
7	Sunday	£5 (US $6.25)	—	—
8	Monday	£5 (US $6.25)	£15 (US $18.75; fourth meeting)	—
9	Tuesday	£5 (US $6.25)	—	—
10	Wednesday	£5 (US $6.25)	£10 (US $12.50; fifth meeting)	—
11	Thursday	£5 (US $6.25)	—	—
12	Friday	£5 (US $6.25)	£10 (US $12.50; sixth meeting)	—
13	Saturday	£5 (US $6.25)	—	—
14	Sunday	£5 (US $6.25)	—	—
15	Monday	£5 (US $6.25)	£15 (US $18.75; seventh meeting)	£35 (US $43.75; seventh meeting)

[a]Total of £75 (US $93.76).

[b]Total of £70 (US $87.51).

[c]Total of £35 (US $43.75).

[d]Not applicable.

Measures

BACtrack Skyn

The TAS used in this study was the BACtrack Skyn (model T15). It was worn on the participants' preferred wrist, but they could change which wrist they wore it on during the study period if that was comfortable for them. The Skyn continuously measured the participant's TAC while being worn, as well as skin temperature (°C). Output was analyzed at 1-minute intervals. The participants could remove the TAS at any time if they did not wish to wear it and were required to remove it for bathing as it is not waterproof. The CM group were told that they could remove it once a day for up to 60 consecutive minutes to bathe and still be eligible for their CM reward. If it was removed for longer than an hour, they would no longer be able to receive the CM reward. If they wore the TAS according to this and the TAC did not increase above our set threshold of 115.660 µg/L, they were eligible for the CM reward for that day. Our set threshold was based on previous work by the research team [35].

TLFB Method

A TLFB was completed at meetings 2 to 7 to assess self-reported alcohol consumption and compare it against TAC. The TLFB is a calendar-based measure to record self-reported substance use. A day was considered from midnight to 11:59 PM.

Postwear Surveys

Participants completed a postwear survey on their experience of wearing the Skyn at their last meeting. This survey was adapted from the work by Alessi et al [20]. If participants were randomized to the CM group, they also completed a survey on their CM experience. This survey was adapted from the work by Miguel et al [62].

XSL•FO
RenderX

Outcomes

Feasibility of the Trial

This primary outcome was the feasibility of this study design. Feasibility was defined by enrollment, participation, device tampering, removals, adjustments, malfunction rates, and the number of TASs returned.

Enrollment

Participants who were identified, approached, eligible, and enrolled were recorded. Participants who were approached but did not wish to participate were asked about their main reason for this. Participant safety was recorded via adverse events for each participant. The number of participants enrolled on each service was also recorded.

Participation

The number of attended meetings by enrolled participants, withdrawal rate, compliance with wearing the TAS, and reasons for incomplete participant data (nonattendance, TAS data overwriting, and TAS technical faults) were recorded. Participants were asked what their main reason was for nonattendance or for not complying with TAS wear.

Removals

TAS removals were defined as a skin temperature of <30 °C for >2 minutes. Removals were recorded for each participant.

Tampering

Participants were asked about tampering with the TAS if there was no clear reason for missing TAC data in the output.

Malfunction

TAS errors, missing data due to a technical fault, and charging or syncing issues were recorded, as well as how much data were missing for these reasons.

TAS Return

The number of Skyn devices that were returned to the research team at the end of each participant's study period was recorded.

CM Delivery

The feasibility of TASs to measure CM target behavior was a combination of the factors mentioned previously, including compliance, removals, TAS tampering, malfunction, and accuracy.

Secondary Outcomes

Acceptability

This was measured using the postwear survey.

Accuracy

Accuracy outcomes were measured using the TLFB and TAC data. The TLFB was self-reported [63] and was used to determine alcohol-drinking days and how many units were consumed (in the United Kingdom, 8 g or 10 mL of pure ethanol=1 unit). The TAS (BACtrack Skyn) continuously measured TAC and skin temperature (°C), and from this, Skyn removal and then reported drinking days based on the TAC (µg/L) were determined. The TLFB and TAC data were analyzed to determine TAS accuracy compared to self-report data.

To note, the alcohol-drinking day defined using TAC is different to the CM intervention criteria, which allowed for a low amount of drinking and had a higher TAC criterion of 115.660 µg/L.

Data Handling

There is currently no guidance from BACtrack to determine drinking event criteria. Courtney et al [34] described their procedure for processing Skyn output to identify drinking episodes, and we used these guidelines when processing participant data. We replaced any negative values recorded with 0 in the data output. Missing data were classified as any minutes not reported. Removed data were those reporting a temperature of <30 °C for >2 minutes.

An alcohol event was based on TAC greater than a specific value (µg/L) for more than a set number of minutes. These criteria were chosen due to previous work carried out by this research team [35,59].

Analysis

Descriptive statistics were reported for baseline and demographic variables. All statistical analyses were conducted using SPSS (version 28; IBM Corp).

The feasibility outcomes reported included enrollment and recruitment rate, participant attendance, response and compliance, removals, TAS tampering and error, and TAS return. Appropriate summary statistics were reported, and independent-sample 2-tailed *t* tests were used to compare means between the control and CM groups when appropriate.

Summary statistics were reported for the secondary outcomes to be explored for a possible future larger trial. The postwear survey answers were reported. TAS accuracy was determined by analyzing the TAS data compared to self-report TLFB. The analysis focused on the sensitivity, specificity, positive predictive value, negative predictive value, and percentage accuracy in classification of TAC compared to TLFB as the gold standard. Recorded drinking and abstinent days were assessed using Spearman rank correlations. Sensitivity in detecting alcohol events and specificity in classifying an alcohol-drinking day versus a non–alcohol-drinking day were assessed using different TAC criteria: TAS 15 (TAC>15 µg/L for >15 min), TAS 60 (TAC>15 µg/L for >60 min), and TAS 90 (TAC>15 µg/L for >90 min).

Ethical Considerations

This study was approved by the Cornwall and Plymouth Research Ethics Committee (reference: 23/SW/0066) and registered on the Open Science Framework and International Standard Randomised Controlled Trial Number (reference: ISRCTN46845361). Informed written consent was obtained from all participants and their data were anonymized. All participants received a £5 (US $6.25) voucher at every meeting they attended (total for all 7 meetings=£35 [US $43.75]), a voucher for the return of the TAS on the last meeting (£10 [US $12.50]), and reimbursement for travel costs to each meeting.

XSL•FO
RenderX

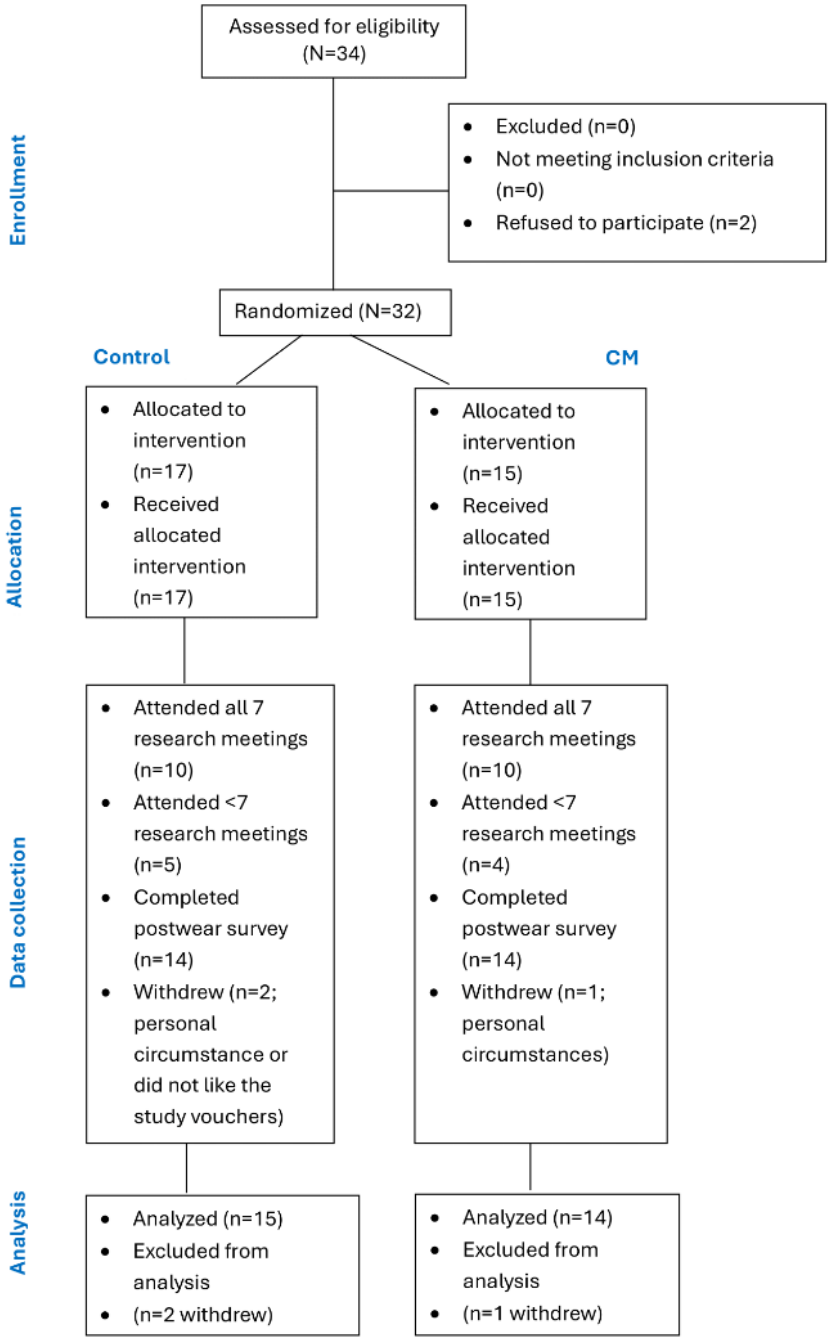
Therefore, each participant could have received up to £45 (US $56.25) for their participation plus travel costs.

Results

Participants

A total of 32 healthy adult participants (n=10, 31% female and n=22, 69% male) enrolled, and a total of 29 (91%) completed the study. In total, 3 participants withdrew (n=2, 67% from the control group and n=1, 33% from the CM group). Of these 3 withdrawals, 2 (67%) happened during the study period (P8 [control group] on day 3 and P28 [CM group] on day 12), and 1 (33%) occurred at the end of the first meeting (control group) due to the specific type of voucher, which could not be used at the participant's local supermarket. P8 and P28 withdrew because of personal circumstances changing. Therefore, a total of 29 participants were included in the analysis (n=10, 34% female and n=19, 66% male; Figure 1). Table 3 shows the characteristics of the 29 participants included in the analysis split by group (control vs CM).

Figure 1. Flowchart of participant recruitment and retention. CM: contingency management.



XSL•FO
RenderX

Table 3. Differences in group characteristics.

Characteristic	Control group (n=15)	CM[a] group (n=14)	Group intervention difference
Sex[b], n (%)			χ^2_1=0.4; *P*=.52
Female	6 (40)	4 (28.6)	
Male	9 (60)	10 (71.4)	
Age (y), mean (SD; range)	44.4 (13.78; 30-75)	51.5 (9.77; 35-60)	*P*=.12
BMI (kg/m^2), mean (SD; range)[c]	27.3 (6.24; 20.8-42.3)	28.7 (6.35; 22.2-42.5)	*P*=.58
Height (cm), mean (SD; range)	168.17 (6.40; 162-177)	171.29 (7.49; 162-185)	*P*=.24
Weight (kg), mean (SD; range)	77.2 (17.16; 53-112.5)	84.63 (19.75; 59-126.33)	*P*=.31
Ethnicity[b], n (%)			χ^2_{16}=43.3; *P*<.001
Black African	3 (20)	0 (0)	
Black British	2 (13.3)	0 (0)	
Caribbean	0 (0)	1 (7.1)	
Hispanic	1 (6.7)	0 (0)	
White British	7 (46.7)	9 (64.3)	
White Irish	1 (6.7)	3 (21.4)	
White Polish	0 (0)	1 (7.1)	
White Russian	1 (6.7)	0 (0)	
Units of alcohol consumed, mean (SD; range)	117.14 (105.83; 0-316.69)	53.33 (93.21; 0-294.69)	*P*=.14
Days in which alcohol was consumed (TLFB[d]), mean (SD)	8.14 (6.44)[e]	4.50 (6.62)[f]	*P*=.08
Average units consumed over days of alcohol consumption, mean (SD)	117.14[g] (105.83)	53.33[h] (93.21)	—[i]
TAS[j] removals (min), n (%)	60,879 (20.8)[k]	30,103 (11)[l]	*P*=.08
TAS missing data (min), n (%)	28,213 (9.6)[k]	28,591 (10.4)[l]	*P*=.80
Meetings attended, n (%)	94 (89.5)[m]	91 (92.9)[n]	*P*=.57
Participants who completed the postwear survey, n (%)	14 (93.3)	14 (100)	*P*=.34
Participants who returned the TAS, n (%)	14 (93.3)	14 (100)	*P*=.34

[a]CM: contingency management.

[b]Ethnicity is represented as the frequency of individuals in each group identifying as Black African, Black British, Caribbean, Hispanic, White British, White Irish, White Polish, and White Russian and was compared across sexes using a chi-square analysis.

[c]One female and one male participant requested not to be weighed, so BMI could not be calculated (both in the control group).

[d]TLFB: timeline followback.

[e]Total of 114 days.

[f]Total of 63 days.

[g]1640.01 units over 114 days.

[h]746.67 units over 63 days.

[i]No group comparison was possible.

[j]TAS: transdermal alcohol sensor.

[k]n=293,186.

[l]n=274,218.

[m]n=105.

[n]n=98.

XSL•FO
RenderX

Feasibility of the Trial

Enrollment

A total of 34 potentially eligible participants were approached. Only 6% (2/34) who were identified by staff declined to participate after discussing the study with the researcher. Their reasons were (1) that they did not like wearing a watch so did not think that they would like to wear a TAS and (2) that they were very busy and could not commit to the regular meetings. No adverse events occurred. The services involved were also willing to help with recruitment, and all recruited at least 4 participants over the 5-month recruitment period.

Participation

It was feasible to enroll more than our target sample size (30 participants) within 5 months. The first participant was enrolled on July 5, 2023, and the final participant (N=32) was enrolled on December 6, 2023. These data could be used to inform the sample size power calculation for a future larger trial.

Among the 29 participants who completed the study period and did not withdraw, a total of 203 meetings were arranged (29 × 7 visits). Of these 203 meetings, 185 (91.1%) were attended. Reasons for nonattendance included illness, their partner being ill, being double booked with another health appointment, a broken phone so the researcher could not be contacted, and hospitalization. The research team decided to book research meetings one at a time, so at meeting 1, meeting 2 was agreed upon, and so on. The day before each meeting, a reminder SMS text message was sent out with the time and date and asking the participants to let the research team know whether this needed to be changed. This design seems feasible to maintain a high attendance rate.

Only 1 of the 29 participants (3%) spoke to the research team about removing their TAS early and removed it at the sixth meeting (2 days early) because they did not want to wear it anymore as they felt that “it is messing with my head” (P22).

Among the 3 services involved, the Assertive Alcohol Outreach Team recruited 53% (17/32) of the participants, the Pier Road Project recruited 34% (11/32) of the participants, and Wandsworth Community Drug and Alcohol Service recruited 12% (4/32) of the participants.

Removals

We defined TAS removal as >2 minutes during which the skin temperature was of <30 °C. We reminded participants to wear the TAS as much as possible, including while asleep. However, the TAS is removable by the wearer, and they could choose not to wear it if they wanted to and must remove it for water activities (bathing and swimming). Of the total minutes collected from the 29 participants (464,324 minutes of data, approximately 322 days), there was a total of 19.59% (90,982/464,324) of minutes of removals (approximately 63 days).

The control group had a total of 20.21% (60,879/301,179) of minutes of removals. The CM group had a total of 10.79% (30,103/278,862) of minutes of removals (t_{27}=1.843; P=.08; d=0.685).

Tampering

One participant (P1) admitted to turning the Skyn off while wearing it when they did not want to be monitored. No other evidence or reports of tampering from any other participant was recorded. This participant had 65.65% (13,173/20,067) of minutes of their participation time successfully recorded and 34.35% (6893/20,067) of missing data. Of these missing data, approximately half could be due to not attending meetings and data being overwritten (3386/6893, 49.12%). However, all other meetings were attended, so the 50.88% (3507/6893) of minutes of missing data could be due to the TAS being turned off or technology error and additional data overwriting (approximately 2.5 days’ worth of data).

Malfunction

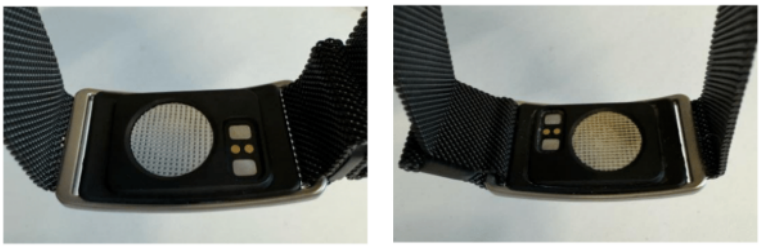
A total of 16 Skyns were ordered in phases during the study period. The first order of 6 Skyns had 2 issues: one Skyn had trouble pairing for the first 2 weeks, and another would not charge or pair with any research iPhone. BACtrack sent 3 Skyn replacements, and 2 of these would not pair with a research iPhone at first (they did 1 month later). We ordered a second batch of 7 Skyns halfway through data collection, and there were no issues with any devices from this order. BACtrack were unsure about the reasons for the issue with pairing and charging. Other issues experienced included syncing taking >40 minutes on occasions (the usual time is approximately 5 minutes). This was most likely due to poor connection or Wi-Fi.

Apart from 1 Skyn that arrived and could not be charged, there were no charging issues. The Skyn website states that the battery lasts for approximately 10 days; however, we found that it lasted for >2 weeks. Participants were provided with a fully charged Skyn at the first meeting. There were 2 participants whose battery decreased to a low level within their 2-week study period, so the warning red light started flashing. As the TASs were typically able to retain the charge for the entire study period, we did not provide the charging cable to participants. If the TASs will be used for >2 weeks in a future study, either the charging cable will need to be provided or there will need to be planned charging time with the researchers.

There was evidence of Skyn degradation after 2 months, which is partly why it was decided to order another batch of Skyns halfway through data collection [33]. Another reason for this order was their state of cleanliness (Figure 2). BACtrack do provide guidelines on how to clean the sensor; however, after following these instructions, the sensor could not be cleaned well.

XSL•FO
RenderX

Figure 2. These 2 photos of the BACtrack Skyn sensor show a brand-new, unworn Skyn (left) and a worn Skyn (right).



TAS Return

All but 1 Skyn were returned. This was due to the participant disengaging from the service and not being contactable by the researcher or key worker. One other participant did not attend their last research meeting, but they arranged an additional meeting to return the TAS.

CM and Intervention Delivery

CM rewards were delivered if 2 criteria were met: (1) the TAC for the day (midnight-11:59 PM) had a peak of <115.660 µg/L and (2) there were no removals lasting >60 minutes (temperature of <30 °C).

The TAS data were downloaded at each meeting (meetings 2 to 7) and checked, and any CM vouchers were provided for the previous 2 to 3 days. This was considered the best way to implement CM as close to those days in which the behavior occurred as possible. Table 4 details the amount of CM rewards earned by each participant and any reason why CM rewards were not earned.

XSL•FO
RenderX

Table 4. Contingency management (CM) rewards earned by participants in the CM group.

Participant ID	Days in which CM rewards were earned[a]	Total amount earned[b]	Days in which CM rewards were not earned[c]	Reason why CM rewards were not earned[d]
P4	0	£0	15	15 days=TAC[e] over the limit
P5	13	£105 (US $131.26)	2	1 day=TAC over the limit; 1 day=removal for >60 min
P6	12	£95 (US $118.76)	3	1 day=TAC over the limit; 2 days=removal for >60 min
P7	10	£85 (US $106.26)	5	1 day=TAC over the limit; 4 days=removal for >60 min
P10	12	£90 (US $112.51)	3	1 day=TAC over the limit; 2 days=removal for >60 min
P11	0	£0	15	15 days=TAC over the limit
P16	2	£10 (US $12.50)	13	6 days=removal for >60 min; 7 days=DNA[f] appointment and data overwritten
P17	5	£25 (US $31.25)	10	6 days=TAC over the limit; 2 days=removal for >60 min; 2 days=DNA appointment and data overwritten
P18	9	£55 (US $68.75)	6	6 days=removal for >60 min
P19	9	£75 (US $93.76)	6	3 days=removal for >60 min; 3 days=DNA appointment and data overwritten
P21	13	£110 (US $137.51)	2	2 days=removal for >60 min
P22	1	£5 (US $6.25)	11	9 days=TAC over the limit; 2 days=removal for >60 min; handed TAS[g] back 2 days early
P25	3	£15 (US $18.75)	12	2 days=TAC over the limit; 10 days=removal for >60 min
P29	11	£85 (US $106.26)	4	4 days=removal for >60 min

[a]Total: 48.3% (100/207) of days.

[b]Total: £755 (US $943.82).

[c]Total: 51.7% (107/207) of days.

[d]Total: 24.6% (51/207) of days with transdermal alcohol concentration over the limit, 21.3% (44/207) of days of removals, and 5.8% (12/207) of days of data overwritten.

[e]TAC: transdermal alcohol concentration.

[f]DNA: did not attend.

[g]TAS: transdermal alcohol sensor.

Control group participants attended 89.5% (94/105) of the meetings, and the CM group participants attended 92.9% (91/98) of the meetings (t_{27}=−0.573; *P*=.57; *d*=−0.213). The control group self-reported 48.6% (107/220) of abstinent days, and the CM group self-reported 70% (147/210) of abstinent days (t_{27}=−1.403; *P*=.17; *d*=−0.522). The proportion of days in which the Skyn data reported <1 hour of removals was 15.9% (35/220) for the control group and 24.8% (52/210) for the CM group (t_{27}=−1.326; *P*=.20; *d*=−0.493). The difference in the amount of units consumed between groups was not statistically significant (control: mean 109.33, SD 27.47; CM: mean 53.33, SD 24.91; t_{27}=1.503; *P*=.14; *d*=0.559).

Acceptability and Postwear Survey

A total of 97% (28/29) of the participants completed the postwear survey. One participant did not complete it because they did not attend their final research meeting and became unreachable. When asked about physical comfort, most participants (11/28, 39%) rated it to be quite to very comfortable, with the average score being 8.57/10 (SD 0.71; 10=very comfortable). Social comfort (how they felt wearing it in public) was also rated highly, with an average score of 9.63/10 (SD 0.00; 10=very socially comfortable). When asked how often they noticed the Skyn on their wrist, only 11% (3/28) of the participants said that they never noticed it when wearing it. Most (18/28, 64%) noticed it once or twice a day to every hour, but 25% (7/28) did report noticing it several times per hour. When asked to rate its interference with various activities

XSL•FO
RenderX

(exercise, mood, normal work, enjoyment of life, ability to concentrate, social life, and clothing choices), these were all rated with an average score of <2 (1=no interference at all; general activity: 1.75, SD 0.00; exercise: 1.07, SD 0.00; mood: 1.11 SD 0.00; work: 1.07, SD 0.00; enjoyment of life: 1.21, SD 0.00, ability to concentrate: 1.43, SD 0.00; social life: 1, SD 0.00; choice of clothing: 1.24, SD 0.00). The only activity that ranked higher was sleeping, with an average score of 2.71 (SD 2.83; 1=no interference at all).

A total of 32% (9/28) of the participants reported experiencing a mark on their skin from the TAS, with the other 68% (19/28) reporting never experiencing a mark or side effect. These side effects included itching (mean 1.95, SD 0.00), sweating (mean 1.88, SD 0.00), soreness (mean 1.07, SD 0.00), and irritation (mean 1.50, SD 0.00), with the scale being from 1=not noticeable to 10=unbearable. Participants were asked whether they would continue to wear the TAS for longer than the 2-week study period, and 96% (27/28) reported *yes*, and only 4% (1/28) reported *no*.

The following statements—"the device is too uncomfortable to wear for any longer," "I wouldn't want to wear it any longer because I am embarrassed," and "I want to wear short sleeves but won't while wearing the device"—were answered by all participants with "not true." Statements asking about changing their clothing choices, wanting to remove the TAS, and not liking the regular download visits all had an average score of 0.25 (SD 0.00 for all these statements) (on a scale of 0-4 where 0=*not true* and 4=*very true*). Only 4% (1/28) of the participants stated that they were tired of explaining the device to people. In total, 4% (1/28) stated that they were "ready to stop wearing the device because I am just ready to be done," and another participant (1/28, 4%) said that the "financial compensation for wearing the device any longer than this would not be worth it." A total of 7% (2/28) of the participants said that it was true that "I would not continue wearing the device because I do not like having to do the downloads at a specific time."

When asked about their drinking, 11% (3/28) self-reported that they were able to completely reduce their drinking, 25% (7/28) reported that they reduced it quite a bit, 14% (4/28) reported that they reduced it somewhat, 7% (2/28) reported that they tried but did not reduce it, and 11% (3/28) reported that they did not reduce or try to reduce it. In total, 32% (9/28) of the participants were abstinent throughout their study period.

Many participants (12/28, 43%) liked the in-person vouchers that were given at each meeting; however, 7% (2/28) would have preferred an e-voucher, 21% (6/28) would have preferred a bank transfer, and 18% (5/28) would be happy with any of these options (in-person voucher, e-voucher, or bank transfer). In total, 7% (2/28) reported preferring only an e-voucher or bank transfer, and 4% (1/28) preferred an in-person voucher or a bank transfer. Finally, participants were asked about the ease of meeting the researcher for the meetings, and 89% (25/28) reported that it was *not at all difficult*, whereas the other 11% (3/28) reported that it was *a little bit difficult*.

CM Survey

The 48% (14/29) of participants who were allocated the CM group completed another survey specifically on their experience of the CM rewards. All participants in the CM group (14/14, 100%) completed this survey. When asked about how easy it was to understand the CM reward procedure, most (9/14, 64%) said that it was very easy, but 21% (3/14) rated it as very difficult. When asked whether the CM had any effect on their response to treatment, half (7/14, 50%) said that it helped them either a lot or a little, 43% (6/14) said that it did not make a difference, and 7% (1/14) said that it had a negative impact. When asked whether they liked the CM rewards, all (14/14, 100%) said that they liked them either a lot or a little. When asked when CM rewards would help other people who seek substance treatment for alcohol dependence, 86% (12/14) said that it would help them a lot or a little, and 14% (2/14) said that it would depend on the person.

Accuracy

Participants wore the TAS for a cumulative total of 580,040 minutes. A total of 80.05% (464,324/580,040) of minutes of this participation was successfully recorded by the TAS. Of these 464,324 minutes of recorded data, there was a total of 90,982 (19.59%) minutes that suggested removal, a total of 56,858 (12.25%) minutes that were missing for unknown reasons, and a total of 58,986 (12.7%) minutes that were missing due to being overwritten as a result of not attending appointments. The Skyn is not waterproof, so participants were required to remove it for showering, baths, and any other water activities (eg, swimming). No fit adjustments were required.

The TAS recorded data for a total of 388 days. Of these 388 days, 337 (86.9%) had >60 minutes of removals or missing data. There was a total of 68.6% (266/388) of days in which there were >300 minutes of removed or missing data (not necessarily a consecutive 5-hour period).

TLFB Versus Skyn TAC

Overview

TLFB drinking days were recorded via the participant reporting any alcohol consumed. TAS drinking days were recorded according to the TAC criteria. We considered 3 different criteria for detecting alcohol-drinking days: TAC of >15 µg/L for >15 minutes, TAC of >15 µg/L for >60 minutes, and TAC of >15 µg/L for >90 minutes. The agreement between the TLFB and these criteria is reported in Table 5.

XSL•FO
RenderX

Table 5. Timeline followback (TLFB)– and transdermal alcohol sensor (TAS)–reported alcohol-drinking and nondrinking days.

Criteria	Alcohol-drinking days detected	Alcohol-drinking days in agreement (n=388), n (%)	Non–alcohol-drinking days in agreement (n=388), n (%)	Days reported by the TAS but not by TLFB as alcohol-drinking days (n=388), n (%)	Days reported by TLFB but not by the TAS as alcohol-drinking days (n=388), n (%)
TAC[a] >15 µg/L for >15 min		164 (42.3)	145 (37.4)	65 (16.8)	14 (3.6)
TLFB	185				
TAS	227				
TAC>15 µg/L for >60 min		146 (37.6)	171 (44.1)	39 (10.1)	32 (8.2)
TAS	185				
TLFB	185				
TAC>15 µg/L for >90 min		139 (35.8)	179 (46.1)	31 (8)	39 (10.1)
TAS	185				
TLFB	170				

[a]TAC: transdermal alcohol concentration.

Criteria for a Drinking Event of TAC>15 µg/L for >15 Minutes

The TLFB and TAS agreed on 42.3% (164/388) of days as alcohol-drinking days and on 37.4% (145/388) of days as abstinent days. There were 3.6% (14/388) of days in which the TLFB reported an alcohol-drinking day and the TAS did not and 16.8% (65/388) of days in which the TAS reported an alcohol-drinking day and the TLFB did not.

When splitting the groups and conducting a correlation between TAS- and TLFB-reported drinking days, both were positively significantly correlated, but the CM group had a stronger correlation effect (control group: r_{15}=0.625 and P=.01; CM group: r_{14}=0.836 and P<.001).

Criteria for a Drinking Event of TAC>15 µg/L for >60 Minutes

The TLFB and TAS agreed on 37.6% (146/388) of days as alcohol-drinking days and on 44.1% (171/388) of days as abstinent days. There were 8.2% (32/388) of days in which the TLFB reported an alcohol-drinking day and the TAS did not and 10.1% (39/388) of days in which the TAS reported an alcohol-drinking day and the TLFB did not.

When splitting the groups and conducting a correlation between TAS- and TLFB-reported drinking days, both were positively significantly correlated, but the CM group had a stronger correlation effect (control group: r_{15}=0.764 and P<.001; CM group: r_{14}=0.895 and P<.001).

Criteria for a Drinking Event of TAC>15 µg/L for >90 Minutes

The TLFB and TAS agreed on 35.8% (139/388) of days as alcohol-drinking days and on 46.1% (179/388) of days as abstinent days. There were 10.1% (39/388) of days in which the TLFB reported an alcohol-drinking day and the TAS did not and 8% (31/388) of days in which the TAS reported an alcohol-drinking day and the TLFB did not.

When splitting the groups and conducting a correlation between TAS- and TLFB-reported drinking days, both were positively significantly correlated, but the CM group had a stronger correlation effect (control group: r_{15}=0.805 and P<.001; CM group: r_{14}=0.913 and P<.001).

Skyn TAC

Reasons why the TAS did not detect an alcohol event but it was reported in the TLFB could include a large amount of missing data; TAS removal; a low amount of alcohol self-reported; or it being the last day of participation, which may mean that there was not enough time for alcohol to appear in the TAC output. There were other days in which there was no obvious reason why the TAS did not detect alcohol consumption.

Reasons why the TAS reported an alcohol-drinking day but the participant did not self-report alcohol consumption could include a sudden TAC spike, which suggests contact with an alcohol-containing product (eg, aftershave, deodorant, or hand sanitizer); the skin temperature suggesting that the TAS was not being worn at the time of TAC event detection, so the TAS may have been on a table near an alcohol spill or alcohol-containing product; and the TAS appearing to detect an alcohol event, so the self-report was inaccurate. There were other days in which there was no obvious reason why the participant did not self-report alcohol consumption.

In Figure 3, a visual presentation of the data for each TAC criteria is provided. It shows TLFB and the TAS criteria of TAS 15 (TAC>15 µg/L for >15 min), TAS 60 (TAC>15 µg/L for >60 min), and TAS 90 (TAC>15 µg/L for >90 min). The presence of the color corresponding to each of these TAC criteria indicates that an alcohol-drinking day was detected.

XSL•FO
RenderX

Figure 3. Visual representation of participants' Timeline Followback (TLFB)– and transdermal alcohol sensor (TAS)–detected alcohol-drinking days for (A) the control group and (B) the contingency management (CM) group. The figure includes 3 different TAS criteria for detecting a drinking day: TAS 15 (transdermal alcohol concentration [TAC]>15 µg/L for >15 min), TAS 60 (TAC>15 µg/L for >60 min), and TAS 90 (TAC>15 µg/L for >90 min).

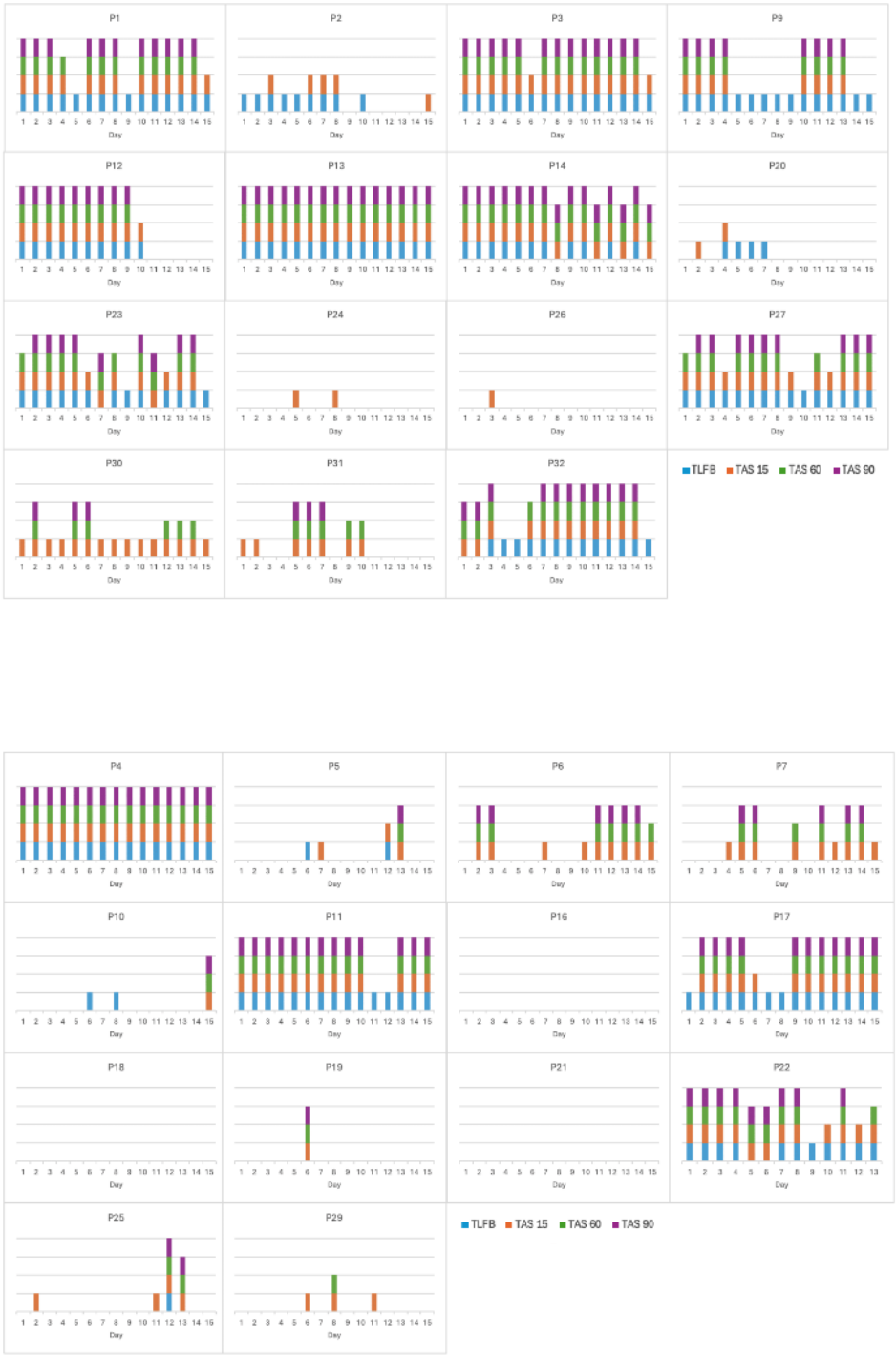
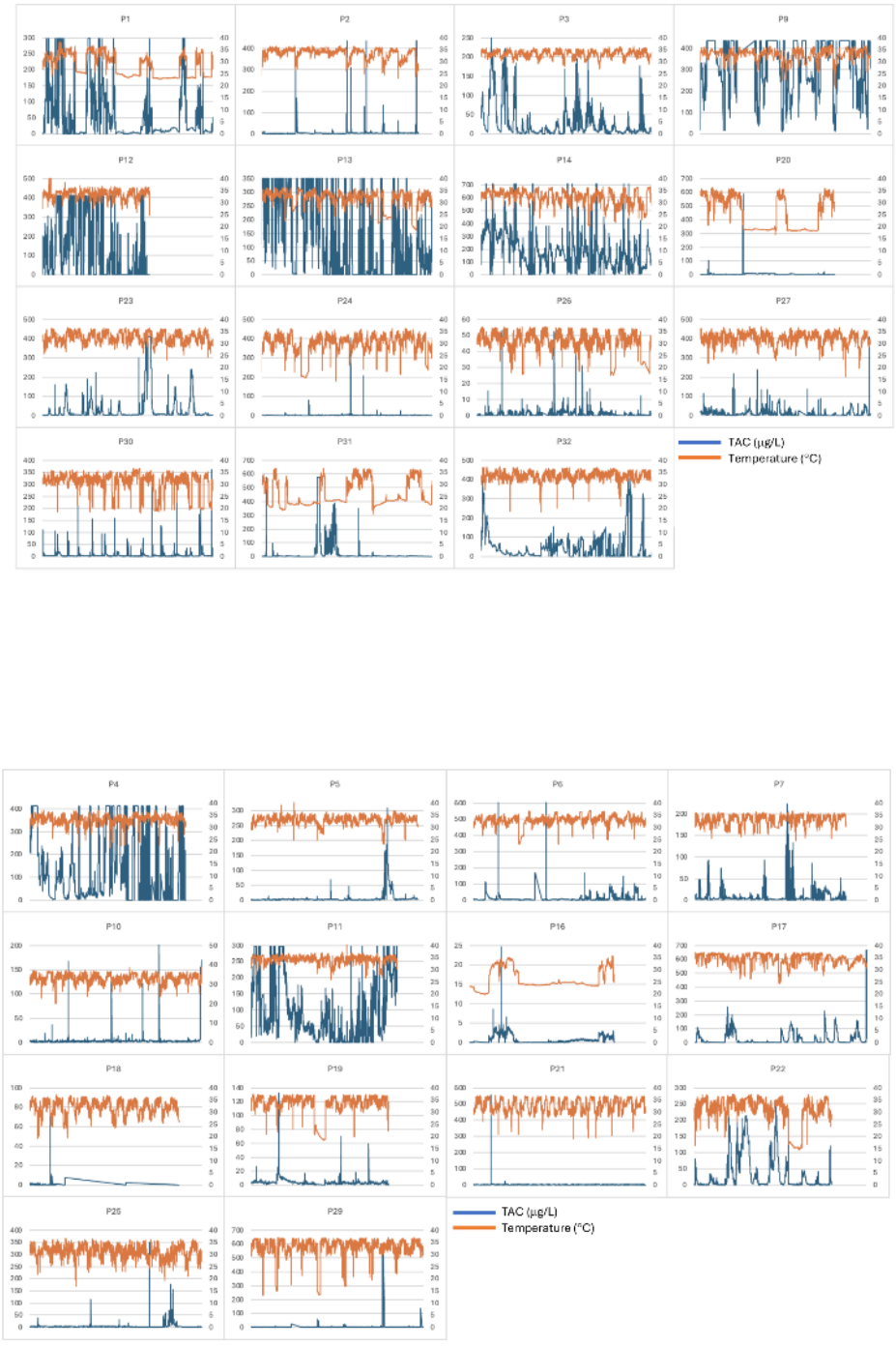


Figure 4 shows the data for TAC and the skin temperature of all participants who wore the TAS for the entire study period. There were participants whose TAC (blue line) peaked on separate occasions, suggesting multiple alcohol events, for example, P5 or P17. In contrast, other participant data suggest more frequent alcohol consumption, for example, P9 and P13. The orange lines depict the skin temperature recorded; the skin temperature for P1 and P20 suggests regular, long TAS removal.

XSL•FO
RenderX

Figure 4. Individual participant Skyn data (excluding withdrawn participants). The primary axis shows the transdermal alcohol concentration (TAC; µg/L) data (blue line), and the secondary axis shows the skin temperature (°C) data (orange line).



Sensitivity, Specificity, and Correlations

We calculated the sensitivity, specificity, positive predictive value, negative predictive value, and percentage accuracy in classification for 3 drinking event thresholds: TAS 15 (TAC>15 µg/L for >15 min), TAS 60 (TAC>15 µg/L for >60 min), and TAS 90 (TAC>15 µg/L for >90 min), shown in Table 6.

XSL•FO
RenderX

Table 6. Transdermal alcohol concentration (TAC) criteria—sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and percentage accuracy classification (PAC).

Criteria	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	PAC
TAC>15 µg/L for 15 min	92.7	68.29	71.74	91.5	0.80
TAC>15 µg/L for 60 min	82.49	80.5	78.92	83.85	0.81
TAC>15 µg/L for 90 min	78.09	85.31	81.76	82.19	0.82

A Spearman correlation found a significant relationship between TLFB- and Skyn-measured alcohol-drinking days at the TAC>15 µg/L for >15 minutes threshold (r_s=0.695, bias–corrected and accelerated [BCa] 95% CI 0.417-0.883; *P*<.001) [29]. A Spearman correlation between the number of standard alcohol units self-reported and the average TAC values was also found to be significant (r_s=0.770, BCa 95% CI 0.632-0.852; *P*<.001) [29].

A Spearman correlation found a significant relationship between TLFB- and Skyn-measured alcohol-drinking days at the TAC>15 µg/L for >60 minutes threshold (r_s=0.773, BCa 95% CI 0.599-0.861; *P*<.001) [29]. A Spearman correlation between the total number of standard alcohol units self-reported and the TAS-reported drinking days at the TAC>15 µg/L for >60 minutes threshold was also found to be significant (r_s=0.740, BCa 95% CI 0.439-0.891; *P*<.001) [29].

A Spearman correlation found a significant relationship between TLFB- and Skyn-measured alcohol-drinking days at the TAC>15 µg/L for >90 minutes threshold (r_s=0.784, BCa 95% CI 0.625-0.868; *P*<.001) [29]. A Spearman correlation between the total number of standard alcohol units self-reported and the TAS-reported drinking days at the TAC>15 µg/L for >90 minutes threshold was also found to be significant (r_s=0.745, BCa 95% CI 0.460-0.892; *P*<.001) [29].

Discussion

Principal Findings

This study aimed to explore the feasibility, strengths, and limitations of using a TAS to monitor alcohol consumption in individuals in treatment for AUD with or without CM to promote abstinence or low-level alcohol consumption. The findings suggest that TAS-delivered CM to encourage alcohol reduction was feasible and acceptable to participants. There was a high correlation between TAS-recorded alcohol-drinking days and self-reported alcohol units, suggesting that the TAS was accurate in measuring the desired behavior. There was also high recruitment, attendance, and compliance of participants. Good rates of meeting attendance and data completeness were achieved. It was feasible to deliver CM using the TAS; however, key features of this process were identified for improvement. This is the first study of TAS-delivered CM in the United Kingdom and the first targeted at alcohol service users for alcohol treatment and reduction.

Participants found the wear of the TAS acceptable and comfortable even for social occasions. They found that it did not interfere with daily activities, and most (19/28, 68%) did not experience any side effects (such as a rash or skin irritation) from the strap. There was a high attendance rate and willingness to participate from those approached. Participants liked the reminder SMS text messages about each meeting, and there were few issues when arranging meetings. Only 3% (1/29) of the participants lost contact with the researcher during participation and did not return the TAS.

However, the answers to the CM survey suggest that this process may need to be improved. Some participants (3/28, 11%) reported it as difficult to understand, and half (7/14, 50%) said that it made no difference or did not help their treatment. These findings suggest that there should be more consideration on how the CM procedure is described and presented to participants at the start. Most (12/14, 86%) said that they did think that CM rewards would help other people who seek treatment for alcohol dependence a little to a lot, with the other 14% (2/14) of the participants saying that it would probably help but only if the patient was motivated.

To deliver the CM, TAS data were downloaded and reviewed at meetings 2 to 7. This was feasible; however, there were issues that arose that would need to be dealt with before scaling up the study size. Considering other options for automated TAC data interpretation will reduce the time-consuming manual data checking, which would only increase the researchers' time burden if conducted on a larger trial. In addition, creating an automated system could increase consistency and avoid human error.

The recruitment and follow-up rates from this study were the same as or higher than the recruitment and follow-up rates of other TAS feasibility or pilot studies. Previous studies using TASs had sample sizes for analysis of 5 to 13 participants [2,12,34,64]. In total, 2 of these studies used the SCRAM monitor [2,64], and 2 used the BACtrack Skyn [12,34], but 3 [2,34,64] mentioned discomfort with the TAS as key feedback, with Courtney et al [34], who used the BACtrack, having one participant drop out due to discomfort.

Overall, the findings support the feasibility of implementing and delivering CM using a TAS over a 2-week wear period with alcohol service users currently receiving alcohol treatment. The TAS used, BACtrack Skyn, was well liked and rated as comfortable by participants, with little daily interference. The recruitment rates were high and are encouraging for a larger trial. TAS delivery of CM has been shown to be effective in the United States using the SCRAM monitor [1,2,26-31,65], but this pilot shows promise for translation to other TAS brands delivering CM in the United Kingdom, with these findings highlighting aspects to address and improve before scaling to a larger study design. On the basis of these results, future work could explore the possibility of solutions to these challenges and a cost-effectiveness assessment.

XSL•FO
RenderX

In the future, we would also recommend collecting initial data from the participants to determine each participant's baseline skin temperature. This could be done with the researcher present to ensure correct TAS wear. Some participants were observed by the researcher to be wearing the Skyn during the meeting but the Skyn temperature was reported as <30 °C. Therefore, it may be better to establish a baseline for each participant to then use for their Skyn data for better accuracy at determining wear and removals.

Strengths and Limitations

The strengths of this study are that it was able to demonstrate the feasible use of the BACtrack Skyn over 2 weeks with individuals currently diagnosed with alcohol dependence and receiving alcohol treatment. Participants wore the TAS in their natural settings and unsupervised, consistent with how it should be worn. While the objective of this study was to assess the feasibility of a larger trial, the data collected as part of this study were able to provide more evidence of how this population wears, uses, and experiences a TAS. There was a high meeting attendance and TAS return rate for this study and no issues with participants being unsure or needing additional training for using the TAS after the baseline training. This study continues to support the use of TASs among the population.

The limitations of this study are similar to those of the design of all the studies conducted as part of this PhD—participants were only recruited if they were willing to wear the TAS from the start of the study period. This means that no participants were recruited who were not willing to attempt wearing the TAS. While this is not a limitation in some considerations as the TAS would be a voluntary treatment option to service users if implemented in services, it does mean that the postwear survey findings may be skewed more positively. Only those who were interested in and willing to wear the TAS had the chance to complete the postwear survey at the end of the 2 weeks. However, participants being willing to wear the TAS for the study does not necessarily mean that they were positive about the technology or that they had a good experience of wearing it; therefore, while we note that the study design does exclude those who are not initially willing to attempt wearing a TAS, this may reflect a truer view of service users who would try wearing the TAS as part of alcohol treatment if TASs were to be implemented in clinical settings.

Conclusions

To conclude, if planning a future larger trial for TAS delivery of CM, the proposed design changes from this study are (1) changes to how TAC data are checked for CM rewards from a manual to an automatic process, (2) a clearer explanation of the CM procedure, (3) consideration of the times available for participant meetings to improve participant availability and reduce missed appointments, (4) consideration of other TAS brands that have longer data storage, (5) consideration of whether participants should be provided with an iOS device and trained to sync their data or whether participants should have this option with their personal iOS device if using the BACtrack Skyn (which has been deemed feasible by wearers), and (6) consideration of the assessment of each participant's baseline skin temperature. These points could be assessed via an internal pilot study.

Acknowledgments

This is a summary of independent research funded by the National Institute for Health and Care Research (NIHR) Applied Research Collaboration South London. The views expressed are those of the authors and not necessarily those of the National Health Service (NHS), the NIHR, or the Department of Health and Social Care. CD and PD were supported by the NIHR Maudsley Biomedical Research Centre at the South London and Maudsley NHS Foundation Trust and King's College London. They were also supported by the NIHR Collaboration for Leadership in Applied Health Research and Care at King's College Hospital NHS Foundation Trust and the NIHR Applied Research Collaboration South London at King's College Hospital NHS Foundation Trust. CD was supported by an NIHR Senior Investigator Award. SP has been partly funded by income from research grants obtained from Mundipharma Research Ltd and Camurus AB. SP was a Research Fellow at King's College London at the time of writing. The authors would also like to acknowledge Andrea Solar-Reyes and Noel Kehoe for their data collection work alongside EB for this study and Simon Coulton for his help in participant randomization.

Authors' Contributions

EB led on the conceptualization, methodology, analysis, investigation, resources, data curation, project administration, visualization, and drafting and revision of the manuscript. CD supported the conceptualization, methodology, analysis, supervision, and revision of the manuscript. SP supported the supervision and revision of the manuscript. PD supported the conceptualization, methodology, analysis, supervision, and revision of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT (Consolidated Standards of Reporting Trials) checklist.
[DOC File , 227 KB - humanfactors_v12i1e64664_app1.doc]

XSL•FO
RenderX

References

1. Barnett NP, Celio MA, Tidey JW, Murphy JG, Colby SM, Swift RM. A preliminary randomized controlled trial of contingency management for alcohol use reduction using a transdermal alcohol sensor. Addiction 2017 Jun 22;112(6):1025-1035 [FREE Full text] [doi: 10.1111/add.13767] [Medline: 28107772]
2. Barnett NP, Tidey J, Murphy JG, Swift R, Colby SM. Contingency management for alcohol use reduction: a pilot study using a transdermal alcohol sensor. Drug Alcohol Depend 2011 Nov 01;118(2-3):391-399 [FREE Full text] [doi: 10.1016/j.drugalcdep.2011.04.023] [Medline: 21665385]
3. Rash CJ, Petry NM, Alessi SM, Barnett NP. Monitoring alcohol use in heavy drinking soup kitchen attendees. Alcohol 2019 Dec;81:139-147 [FREE Full text] [doi: 10.1016/j.alcohol.2018.10.001] [Medline: 30308286]
4. Karns-Wright TE, Dougherty DM, Hill-Kapturczak N, Mathias CW, Roache JD. The correspondence between transdermal alcohol monitoring and daily self-reported alcohol consumption. Addict Behav 2018 Oct;85:147-152 [FREE Full text] [doi: 10.1016/j.addbeh.2018.06.006] [Medline: 29910035]
5. Barnett NP, Meade EB, Glynn TR. Predictors of detection of alcohol use episodes using a transdermal alcohol sensor. Exp Clin Psychopharmacol 2014 Feb;22(1):86-96 [FREE Full text] [doi: 10.1037/a0034821] [Medline: 24490713]
6. Dougherty DM, Charles NE, Acheson A, John S, Furr RM, Hill-Kapturczak N. Comparing the detection of transdermal and breath alcohol concentrations during periods of alcohol consumption ranging from moderate drinking to binge drinking. Exp Clin Psychopharmacol 2012 Oct;20(5):373-381 [FREE Full text] [doi: 10.1037/a0029021] [Medline: 22708608]
7. Fairbairn CE, Kang D. Temporal dynamics of transdermal alcohol concentration measured via new-generation wrist-worn biosensor. Alcohol Clin Exp Res 2019 Oct 30;43(10):2060-2069 [FREE Full text] [doi: 10.1111/acer.14172] [Medline: 31469451]
8. Hill-Kapturczak N, Lake SL, Roache JD, Cates SE, Liang Y, Dougherty DM. Do variable rates of alcohol drinking alter the ability to use transdermal alcohol monitors to estimate peak breath alcohol and total number of drinks? Alcohol Clin Exp Res 2014 Oct 21;38(10):2517-2522 [FREE Full text] [doi: 10.1111/acer.12528] [Medline: 25335857]
9. Bond JC, Greenfield TK, Patterson D, Kerr WC. Adjustments for drink size and ethanol content: new results from a self-report diary and transdermal sensor validation study. Alcohol Clin Exp Res 2014 Dec 07;38(12):3060-3067 [FREE Full text] [doi: 10.1111/acer.12589] [Medline: 25581661]
10. Karns-Wright TE, Roache JD, Hill-Kapturczak N, Liang Y, Mullen J, Dougherty DM. Time delays in transdermal alcohol concentrations relative to breath alcohol concentrations. Alcohol Alcohol 2017 Jan 13;52(1):35-41 [FREE Full text] [doi: 10.1093/alcalc/agw058] [Medline: 27522029]
11. Swift RM, Martin CS, Swette L, LaConti A, Kackley N. Studies on a wearable, electronic, transdermal alcohol sensor. Alcohol Clin Exp Res 1992 Aug 11;16(4):721-725. [doi: 10.1111/j.1530-0277.1992.tb00668.x] [Medline: 1530135]
12. Rosenberg M, Ludema C, Kianersi S, Luetke M, Jozkowski K, Guerra-Reyes L, et al. Wearable alcohol monitors for alcohol use data collection among college students: feasibility and acceptability. Alcohol 2023 Sep;111:75-83. [doi: 10.1016/j.alcohol.2023.05.007] [Medline: 37295566]
13. Wang Y, Fridberg DJ, Shortell DD, Leeman RF, Barnett NP, Cook RL, et al. Wrist-worn alcohol biosensors: applications and usability in behavioral research. Alcohol 2021 May;92:25-34 [FREE Full text] [doi: 10.1016/j.alcohol.2021.01.007] [Medline: 33609635]
14. Fairbairn CE, Kang D, Bosch N. Using machine learning for real-time BAC estimation from a new-generation transdermal biosensor in the laboratory. Drug Alcohol Depend 2020 Nov 01;216:108205 [FREE Full text] [doi: 10.1016/j.drugalcdep.2020.108205] [Medline: 32853998]
15. Alessi SM, Barnett NP, Petry NM. Objective continuous monitoring of alcohol consumption for three months among alcohol use disorder treatment outpatients. Alcohol 2019 Dec;81:131-138 [FREE Full text] [doi: 10.1016/j.alcohol.2019.01.008] [Medline: 30710610]
16. Roache JD, Karns TE, Hill-Kapturczak N, Mullen J, Liang Y, Lamb RJ, et al. Using transdermal alcohol monitoring to detect low-level drinking. Alcohol Clin Exp Res 2015 Jul 18;39(7):1120-1127 [FREE Full text] [doi: 10.1111/acer.12750] [Medline: 25988708]
17. Simons JS, Wills TA, Emery NN, Marks RM. Quantifying alcohol consumption: self-report, transdermal assessment, and prediction of dependence symptoms. Addict Behav 2015 Nov;50:205-212 [FREE Full text] [doi: 10.1016/j.addbeh.2015.06.042] [Medline: 26160523]
18. Brobbin E, Deluca P, Hemrage S, Drummond C. Acceptability and feasibility of wearable transdermal alcohol sensors: systematic review. JMIR Hum Factors 2022 Dec 23;9(4):e40210 [FREE Full text] [doi: 10.2196/40210] [Medline: 36563030]
19. Brobbin E, Deluca P, Hemrage S, Drummond C. Accuracy of wearable transdermal alcohol sensors: systematic review. J Med Internet Res 2022 Apr 14;24(4):e35178 [FREE Full text] [doi: 10.2196/35178] [Medline: 35436239]
20. Alessi SM, Barnett NP, Petry NM. Experiences with SCRAMx alcohol monitoring technology in 100 alcohol treatment outpatients. Drug Alcohol Depend 2017 Sep 01;178:417-424 [FREE Full text] [doi: 10.1016/j.drugalcdep.2017.05.031] [Medline: 28709081]
21. Caluzzi G, Pennay A, Cook M, Wright C, Norman T, Kuntsche E. Transdermal monitors to assess alcohol consumption in real-time and real-life – a qualitative study on user-experience. Addict Res Theory 2019 Jan 09;27(4):354-361. [doi: 10.1080/16066359.2018.1530765]

XSL•FO
RenderX

22. Croff JM, Hartwell ML, Chiaf AL, Crockett EK, Washburn IJ. Feasibility and reliability of continuously monitoring alcohol use among female adolescents and young adults. Drug Alcohol Rev 2021 Nov 17;40(7):1143-1154 [FREE Full text] [doi: 10.1111/dar.13045] [Medline: 32185847]
23. Goodall CA, Neville FG, Williams DJ, Donnelly PD. Preliminary research informing policy on remote alcohol monitoring in criminal justice: the Scottish experience. Int J Public Health 2016 Nov 13;61(8):865-872. [doi: 10.1007/s00038-016-0886-9] [Medline: 27624623]
24. Neville FG, Williams DJ, Goodall CA, Murer JS, Donnelly PD. An experimental trial exploring the impact of continuous transdermal alcohol monitoring upon alcohol consumption in a cohort of male students. PLoS One 2013 Jun 25;8(6):e67386 [FREE Full text] [doi: 10.1371/journal.pone.0067386] [Medline: 23825656]
25. Norman T, Peacock A, Ferguson SG, Kuntsche E, Bruno R. Combining transdermal and breath alcohol assessments, real-time drink logs and retrospective self-reports to measure alcohol consumption and intoxication across a multi-day music festival. Drug Alcohol Rev 2021 Nov 10;40(7):1112-1121. [doi: 10.1111/dar.13215] [Medline: 33174260]
26. Villalba K, Cook C, Dévieux JG, Ibanez GE, Oghogho E, Neira C, et al. Facilitators and barriers to a contingency management alcohol intervention involving a transdermal alcohol sensor. Heliyon 2020 Mar;6(3):e03612 [FREE Full text] [doi: 10.1016/j.heliyon.2020.e03612] [Medline: 32258468]
27. Dougherty DM, Hill-Kapturczak N, Liang Y, Karns TE, Cates SE, Lake SL, et al. Use of continuous transdermal alcohol monitoring during a contingency management procedure to reduce excessive alcohol use. Drug Alcohol Depend 2014 Sep 01;142:301-306. [doi: 10.1016/j.drugalcdep.2014.06.039] [Medline: 25064019]
28. Mathias CW, Hill-Kapturczak N, Karns-Wright TE, Mullen J, Roache JD, Fell JC, et al. Translating transdermal alcohol monitoring procedures for contingency management among adults recently arrested for DWI. Addict Behav 2018 Aug;83:56-63 [FREE Full text] [doi: 10.1016/j.addbeh.2018.01.033] [Medline: 29397211]
29. Dougherty DM, Karns TE, Mullen J, Liang Y, Lake SL, Roache JD, et al. Transdermal alcohol concentration data collected during a contingency management program to reduce at-risk drinking. Drug Alcohol Depend 2015 Mar 01;148:77-84 [FREE Full text] [doi: 10.1016/j.drugalcdep.2014.12.021] [Medline: 25582388]
30. Averill F, Brown TG, Robertson RD, Tchomgang A, Berbiche D, Nadeau L, et al. Transdermal alcohol monitoring combined with contingency management for driving while impaired offenders: a pilot randomized controlled study. Traffic Inj Prev 2018 Jul 04;19(5):455-461. [doi: 10.1080/15389588.2018.1448079] [Medline: 29543499]
31. Richards VL, Wang Y, Porges EC, Gullett JM, Leeman RF, Zhou Z, et al. Using alcohol biosensors and biomarkers to measure changes in drinking: associations between transdermal alcohol concentration, phosphatidylethanol, and self-report in a contingency management study of persons with and without HIV. Exp Clin Psychopharmacol 2023 Dec;31(6):991-997. [doi: 10.1037/pha0000637] [Medline: 36649152]
32. Ariss T, Fairbairn CE, Bosch N. Examining new-generation transdermal alcohol biosensor performance across laboratory and field contexts. Alcohol Clin Exp Res (Hoboken) 2023 Jan;47(1):50-59 [FREE Full text] [doi: 10.1111/acer.14977] [Medline: 36433786]
33. Ash GI, Gueorguieva R, Barnett NP, Wang W, Robledo DS, DeMartini KS, et al. Sensitivity, specificity, and tolerability of the BACTrack Skyn compared to other alcohol monitoring approaches among young adults in a field-based setting. Alcohol Clin Exp Res 2022 May;46(5):783-796 [FREE Full text] [doi: 10.1111/acer.14804] [Medline: 35567595]
34. Courtney JB, Russell MA, Conroy DE. Acceptability and validity of using the BACtrack skyn wrist-worn transdermal alcohol concentration sensor to capture alcohol use across 28 days under naturalistic conditions - a pilot study. Alcohol 2023 May;108:30-43 [FREE Full text] [doi: 10.1016/j.alcohol.2022.11.004] [Medline: 36473634]
35. Brobbin E, Deluca P, Coulton S, Drummond C. Accuracy of transdermal alcohol monitoring devices in a laboratory setting. Alcohol Alcohol 2024 Jan 11;59(1):agad068. [doi: 10.1093/alcalc/agad068] [Medline: 37873967]
36. Brobbin E, Parkin S, Deluca P, Drummond C. A qualitative exploration of the experiences of transdermal alcohol sensor devices amongst people in receipt of treatment for alcohol use disorder. Addict Behav Rep 2024 Jun;19:100544 [FREE Full text] [doi: 10.1016/j.abrep.2024.100544] [Medline: 38596194]
37. Didier NA, King AC, Polley EC, Fridberg DJ. Signal processing and machine learning with transdermal alcohol concentration to predict natural environment alcohol consumption. Exp Clin Psychopharmacol 2024 Apr;32(2):245-254. [doi: 10.1037/pha0000683] [Medline: 37824232]
38. Gunn RL, Merrill JE, Haines AM, Fernandez ME, Souza T, Berey BL, et al. Use of the BACtrack Skyn alcohol biosensor: practical applications for data collection and analysis. Addiction 2023 Aug;118(8):1586-1595 [FREE Full text] [doi: 10.1111/add.16207] [Medline: 37060272]
39. Kianersi S, Ludema C, Agley J, Ahn YY, Parker M, Ideker S, et al. Development and validation of a model for measuring alcohol consumption from transdermal alcohol content data among college students. Addiction 2023 Oct 30;118(10):2014-2025 [FREE Full text] [doi: 10.1111/add.16228] [Medline: 37154154]
40. Richards VL, Rajendran S, Cook RL, Leeman RF, Wang Y, Prins C, et al. Identifying desired features that would be acceptable and helpful in a wrist-worn biosensor-based alcohol intervention: interview study among adults who drink heavily. J Med Internet Res 2023 Feb 02;25:e38713 [FREE Full text] [doi: 10.2196/38713] [Medline: 36729576]

XSL•FO
RenderX

41. Richards VL, Barnett NP, Cook RL, Leeman RF, Souza T, Case S, et al. Correspondence between alcohol use measured by a wrist-worn alcohol biosensor and self-report via ecological momentary assessment over a 2-week period. Alcohol Clin Exp Res (Hoboken) 2023 Feb 21;47(2):308-318 [FREE Full text] [doi: 10.1111/acer.14995] [Medline: 36507857]
42. Wang Y, Fridberg DJ, Leeman RF, Cook RL, Porges EC. Wrist-worn alcohol biosensors: strengths, limitations, and future directions. Alcohol 2019 Dec;81:83-92 [FREE Full text] [doi: 10.1016/j.alcohol.2018.08.013] [Medline: 30179709]
43. Drug misuse in over 16s: psychosocial interventions clinical guideline. National Institute for Health and Care Excellence. URL: https://www.nice.org.uk/guidance/cg51 [accessed 2024-04-29]
44. Drake RE, O'Neal EL, Wallach MA. A systematic review of psychosocial research on psychosocial interventions for people with co-occurring severe mental and substance use disorders. J Subst Abuse Treat 2008 Jan;34(1):123-138. [doi: 10.1016/j.jsat.2007.01.011] [Medline: 17574803]
45. Dutra L, Stathopoulou G, Basden SL, Leyro TM, Powers MB, Otto MW. A meta-analytic review of psychosocial interventions for substance use disorders. Am J Psychiatry 2008 Feb;165(2):179-187. [doi: 10.1176/appi.ajp.2007.06111851] [Medline: 18198270]
46. Prendergast M, Podus D, Finney J, Greenwell L, Roll J. Contingency management for treatment of substance use disorders: a meta-analysis. Addiction 2006 Nov 10;101(11):1546-1560. [doi: 10.1111/j.1360-0443.2006.01581.x] [Medline: 17034434]
47. Petry NM. A comprehensive guide to the application of contingency management procedures in clinical settings. Drug Alcohol Depend 2000 Feb 01;58(1-2):9-25. [doi: 10.1016/s0376-8716(99)00071-x] [Medline: 10669051]
48. Petry NM, Martin B, Cooney JL, Kranzler HR. Give them prizes, and they will come: contingency management for treatment of alcohol dependence. J Consult Clin Psychol 2000 Apr;68(2):250-257. [doi: 10.1037//0022-006x.68.2.250] [Medline: 10780125]
49. Alessi SM, Petry NM. A randomized study of cellphone technology to reinforce alcohol abstinence in the natural environment. Addiction 2013 May 30;108(5):900-909 [FREE Full text] [doi: 10.1111/add.12093] [Medline: 23279560]
50. Swift R. Direct measurement of alcohol and its metabolites. Addiction 2003 Dec 30;98 Suppl 2(s2):73-80. [doi: 10.1046/j.1359-6357.2003.00605.x] [Medline: 14984244]
51. Paton A. Alcohol in the body. BMJ 2005 Jan 08;330(7482):85-87 [FREE Full text] [doi: 10.1136/bmj.330.7482.85] [Medline: 15637372]
52. Litten RZ, Bradley AM, Moss HB. Alcohol biomarkers in applied settings: recent advances and future research opportunities. Alcohol Clin Exp Res 2010 Jun 25;34(6):955-967. [doi: 10.1111/j.1530-0277.2010.01170.x] [Medline: 20374219]
53. McDonell MG, Leickly E, McPherson S, Skalisky J, Srebnik D, Angelo F, et al. A randomized controlled trial of ethyl glucuronide-based contingency management for outpatients with co-occurring alcohol use disorders and serious mental illness. Am J Psychiatry 2017 Apr 01;174(4):370-377 [FREE Full text] [doi: 10.1176/appi.ajp.2016.16050627] [Medline: 28135843]
54. Oluwoye O, Reneau H, Herron J, Alcover KC, McPherson S, Roll J, et al. Pilot study of an integrated smartphone and Breathalyzer contingency management intervention for alcohol use. J Addict Med 2020;14(3):193-198 [FREE Full text] [doi: 10.1097/ADM.0000000000000553] [Medline: 31567597]
55. Lauckner C, Taylor E, Patel D, Whitmire A. The feasibility of using smartphones and mobile breathalyzers to monitor alcohol consumption among people living with HIV/AIDS. Addict Sci Clin Pract 2019 Nov 26;14(1):43 [FREE Full text] [doi: 10.1186/s13722-019-0174-0] [Medline: 31771655]
56. You CW, Chen YC, Chen CH, Lee CH, Kuo PH, Huang MC, et al. Smartphone-based support system (SoberDiary) coupled with a Bluetooth breathalyser for treatment-seeking alcohol-dependent patients. Addict Behav 2017 Feb;65:174-178. [doi: 10.1016/j.addbeh.2016.10.017] [Medline: 27816043]
57. Hämäläinen MD, Zetterström A, Winkvist M, Söderquist M, Karlberg E, Öhagen P, et al. Real-time monitoring using a breathalyzer-based eHealth system can identify lapse/relapse patterns in alcohol use disorder patients. Alcohol Alcohol 2018 Jul 01;53(4):368-375. [doi: 10.1093/alcalc/agy011] [Medline: 29590325]
58. Goldfine C, Lai JT, Lucey E, Newcomb M, Carreiro S. Wearable and wireless mHealth technologies for substance use disorder. Curr Addict Rep 2020 Sep 01;7(3):291-300 [FREE Full text] [doi: 10.1007/s40429-020-00318-8] [Medline: 33738178]
59. Brobbin E, Deluca P, Coulton S, Parkin S, Drummond C. Comparison of transdermal alcohol concentration and self-reported alcohol consumption in people with alcohol dependence attending community alcohol treatment services. Drug Alcohol Depend 2024 Mar 01;256:111122 [FREE Full text] [doi: 10.1016/j.drugalcdep.2024.111122] [Medline: 38367536]
60. Brobbin E, Parkin S, Deluca P, Drummond C. A qualitative exploration of the experiences of transdermal alcohol sensor devices amongst alcohol service practitioners (South London, UK). Addict Res Theory 2025;33(1):36-47 [FREE Full text] [doi: 10.1080/16066359.2024.2309869] [Medline: 39906322]
61. Merrill JE, Fan P, Wray TB, Miranda Jr R. Assessment of alcohol use and consequences: comparison of data collected via timeline followback interview and daily reports. J Stud Alcohol Drugs 2020 Mar;81(2):212-219 [FREE Full text] [doi: 10.15288/jsad.2020.81.212] [Medline: 32359051]
62. Miguel AQ, Madruga CS, Cogo-Moreira H, Yamauchi R, Simões V, da Silva CJ, et al. Contingency management is effective in promoting abstinence and retention in treatment among crack cocaine users in Brazil: a randomized controlled trial. Psychol Addict Behav 2016 Aug;30(5):536-543 [FREE Full text] [doi: 10.1037/adb0000192] [Medline: 27442691]

XSL•FO
RenderX

63. Sobell L, Sobell M. Timeline follow-back. In: Litten RZ, Allen JP, editors. Measuring Alcohol Consumption: Psychosocial and Biochemical Methods. Totowa, NJ: Humana Press; 1992:41-72.
64. van Egmond K, Anderson-Luxford D, Kuntsche E, Wright C, Caluzzi G, Pennay A. Measuring alcohol consumption while watching sport events: a feasibility and validity study comparing ecological momentary assessments and transdermal alcohol monitors. Journal of Substance Use 2023 Apr 20;29(5):689-695. [doi: 10.1080/14659891.2023.2199850]
65. Dougherty DM, Moon TJ, Liang Y, Roache JD, Lamb RJ, Mathias CW, et al. Effectiveness of contingency management using transdermal alcohol monitoring to reduce heavy drinking among driving while intoxicated (DWI) arrestees: a randomized controlled trial. Alcohol Clin Exp Res (Hoboken) 2023 Oct 13;47(10):1989-2001 [FREE Full text] [doi: 10.1111/acer.15180] [Medline: 37864527]

Abbreviations

AUD: alcohol use disorder
BCa: bias–corrected and accelerated
CM: contingency management
TAC: transdermal alcohol concentration
TAS: transdermal alcohol sensor
TLFB: timeline followback

Edited by A Kushniruk; submitted 24.07.24; peer-reviewed by R Zhang, O Zhabenko; comments to author 14.10.24; revised version received 18.10.24; accepted 20.12.24; published 14.03.25.

Please cite as:
Brobbin E, Drummond C, Parkin S, Deluca P
Use of Wearable Transdermal Alcohol Sensors for Monitoring Alcohol Consumption After Detoxification With Contingency Management: Pilot Randomized Feasibility Trial
JMIR Hum Factors 2025;12:e64664
URL: https://humanfactors.jmir.org/2025/1/e64664
doi:10.2196/64664
PMID:

©Eileen Brobbin, Colin Drummond, Stephen Parkin, Paolo Deluca. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 14.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Requirement Analysis for Data-Driven Electroencephalography Seizure Monitoring Software to Enhance Quality and Decision Making in Digital Care Pathways for Epilepsy: A Feasibility Study from the Perspectives of Health Care Professionals

Pantea Keikhosrokiani[1,2], PhD; Johanna Annunen[3,4*], MD, PhD; Jonna Komulainen-Ebrahim[3,5*], MD, PhD; Jukka Kortelainen[6,7], PhD; Mika Kallio[3,5], MD, PhD; Päivi Vieira[3,5], MD, PhD; Minna Isomursu[1,2], PhD; Johanna Uusimaa[3,8], MD, PhD

[1]Empirical Software Engineering in Software Systems and Services, Faculty of Information Technology and Electrical Engineering, University of Oulu, Oulu, Finland

[2]Research Unit of Health Sciences and Technology, Faculty of Medicine, University of Oulu, Oulu, Finland

[3]University of Oulu, Oulu, Finland

[4]Oulu University Hospital, Medical Research Center, Neurocenter (Member of ERN EpiCARE), Oulu, Finland

[5]Oulu University Hospital, Oulu, Finland

[6]Center for Machine Vision and Signal Analysis, Faculty of Information Technology and Electrical Engineering, University of Oulu, Oulu, Finland

[7]Cerenion Ltd, Oulu, Finland

[8]Research Unit of Clinical Medicine, Faculty of Medicine, University of Oulu, Oulu, Finland

[*]these authors contributed equally

Corresponding Author:
Pantea Keikhosrokiani, PhD
Empirical Software Engineering in Software Systems and Services
Faculty of Information Technology and Electrical Engineering
University of Oulu
Pentti Kaiteran katu 1
Linnanmaa
Oulu, 90014
Finland
Phone: 1 358514410809
Email: pantea.keikhosrokiani@oulu.fi

Related Article:

This is a corrected version. See correction statement: https://humanfactors.jmir.org/2025/1/e79484/

Abstract

Background: Abnormal brain activity is the source of epileptic seizures, which can present a variety of symptoms and influence patients' quality of life. Therefore, it is critical to track epileptic seizures, diagnose them, and provide potential therapies to manage people with epilepsy. Electroencephalography (EEG) is helpful in the diagnosis and classification of the seizure type, epilepsy, or epilepsy syndrome. Ictal EEG is rarely recorded, whereas interictal EEG is more often recorded, and the results can be abnormal or normal even in the case of epilepsy. The current digital care pathway for epilepsy (DCPE) lacks the integration of data-driven seizure detection, which could potentially enhance epilepsy treatment and management.

Objective: This study aimed to determine the requirements for integrating data-driven medical software into the DCPE to meet the project's goals and demonstrate practical feasibility regarding resource availability, time constraints, and technological capabilities. This adjustment emphasized ensuring that the proposed system is realistic and achievable. Perspectives on the feasibility of data-driven medical software that meets the project's goals and demonstrates practical feasibility regarding resource availability, time constraints, and technological capabilities are presented.

XSL•FO
RenderX

Methods: A 4-round Delphi study using focus group discussions was conducted with 7 diverse panels of experts from Oulu University Hospital to address the research questions and evaluate the feasibility of data-driven medical software for monitoring individuals with epilepsy. This collaborative approach fostered a thorough understanding of the topic and considered the perspectives of various stakeholders. In addition, a qualitative study was carried out using semistructured interviews.

Results: Drawing from the findings of the thematic analytics, a detailed set of guidelines was created to facilitate the seamless integration of the proposed data-driven medical software for EEG seizure monitoring into the DCPE. These guidelines encompass system requirements, data collection and analysis, and user training, offering a comprehensive road map for the effective implementation of the software.

Conclusions: The study outcome presents a comprehensive strategy for improving the quality of care, providing personalized solutions, managing health care resources, and using artificial intelligence and sensor technology in clinical settings. The potential of artificial intelligence and sensor technology to revolutionize health care is exciting. The study identified practical strategies, such as real-time EEG seizure monitoring, predictive modeling for seizure occurrence, and data-driven analytics integration to enhance decision-making. These strategies were aimed at reducing diagnostic delays and providing personalized care. We are actively working on integrating these features into clinical workflows. However, further case studies and pilot implementations are planned for future studies. The results of this study will guide system developers in the meticulous design and development of systems that meet user needs in the DCPE.

(JMIR Hum Factors 2025;12:e59558) doi:10.2196/59558

KEYWORDS

ubiquitous health; uHealth; data-driven software; requirement analysis; epilepsy seizure monitoring; digital care pathway; electroencephalography; Delphi study; focus group; artificial intelligence; AI

Introduction

Background

Epilepsy is a neurological condition that is characterized by recurrent and unprovoked seizures [1]. Epilepsy is one of the most common neurological conditions, affecting millions of people worldwide [2]. Epileptic seizures occur due to abnormal brain electrical activity, leading to a wide variety of symptoms and manifestations. This condition can substantially impact patients' quality of life. Therefore, to effectively manage and provide support for people with epilepsy, it is essential to understand epilepsy, including its causes, diagnosis, available treatments, and comorbidities [3]. One of the most critical aspects of adequate care is knowing thoroughly about the seizure types, etiology, and provoking factors [4,5].

Electroencephalography (EEG) data play a crucial role in detecting epileptic seizures by capturing distinct patterns and anomalies associated with epileptic seizures by recording the brain's electrical activity. Sharp waves and spikes, slow waves, rhythmic patterns, and interictal epileptiform discharges are some of the possible EEG findings of people with epilepsy [4,6-10]. Experienced neurologists or automated algorithms can visually identify and analyze these abnormal patterns. The possibilities of EEG data processing for seizure detection have been substantially enhanced by recent developments in machine learning and artificial intelligence (AI) [11,12]. It is possible to train algorithms to recognize and categorize seizure occurrences automatically, giving individuals or health care professionals real-time notifications.

On the basis of the discussion with health care professionals, the current digital care pathway for epilepsy (DCPE) faces several issues and shortcomings. Despite therapeutic advances, the classification, diagnosis, and treatment of epilepsy remain complex. Current medications only manage seizures in around two-thirds of people with epilepsy, which emphasizes the need for more potent therapies [13]. The current DCPE lacks data-driven seizure detection integration, which could enhance differential diagnosis and treatment in drug-resistant patients. The addition of data-driven technologies could offer significant improvements, including more accurate seizure detection and prediction, personalized treatment plans, better management of medication-resistant epilepsy, and shared decision-making. However, implementing AI solutions such as machine learning for seizure detection and monitoring is not yet a standard component of DCPE. Integrating a new data-driven medical software for EEG seizure monitoring into the current DCPE requires a thorough understanding of the user's needs and system requirements. Furthermore, robust requirement engineering and precise stakeholder requirements definitions are needed to design and develop data-driven medical software to ensure patient safety, privacy, compliance with regulatory standards, functional accuracy and relevance, interpretability, scalability, and sustainability [14-16].

This Study

This study aimed to define the system requirements for developing and integrating data-driven medical software for EEG seizure monitoring (EEG-EpiDigi) in a clinical setting. The system requirement analysis (SRA) focused on assessing user needs, gathering those requirements, analyzing them, and prioritizing their importance. While patient-centered care is a key focus of this system, the initial phase of the study emphasized collecting insights from health care professionals about system requirements. Future research will include patient perspectives and experiences, particularly regarding home monitoring systems, to ensure the software is user friendly. This study explicitly investigated health care professionals' opinions to identify the requirements for implementing data-driven medical software for EEG seizure monitoring within hospitals,

XSL•FO
RenderX

thereby assisting in their decision-making processes. The main research questions of this study were as follows:

- What are the user needs from EEG-EpiDigi in DCPE?
- What are the system requirements for developing and integrating the EEG-EpiDigi system in the current DCPE?
- What are the priorities of the EEG-EpiDigi system in the current DCPE?

The study aimed to determine the feasibility of integrating data-driven medical software in the DCPE. After addressing the research questions and assessing the proposed system requirements, a 4-round Delphi study using focus group discussions with health care professionals from Oulu University Hospital, including neurologists, digital health specialists, pediatric neurologists, experts in EEG signal processing, and clinical neurophysiologists, was conducted from May to November 2023. Semistructured interviews were conducted, and the data were analyzed qualitatively. The final round of the Delphi study directs the discussion to the system requirements confirmation as the first step of the new EEG-EpiDigi system development life cycle. The remainder of the paper is organized as follows: the details of the proposed methodology, such as study design, participants, interview guide, and data analysis, are discussed. Afterward, the results and analysis are added, which include the data exploration, study mind map, and the thematic analysis of the system requirements for EEG-EpiDigi. The paper is wrapped up with a Discussion section, including the futuristic architecture for EEG-EpiDigi and the proposed SRA-EpiDigi.

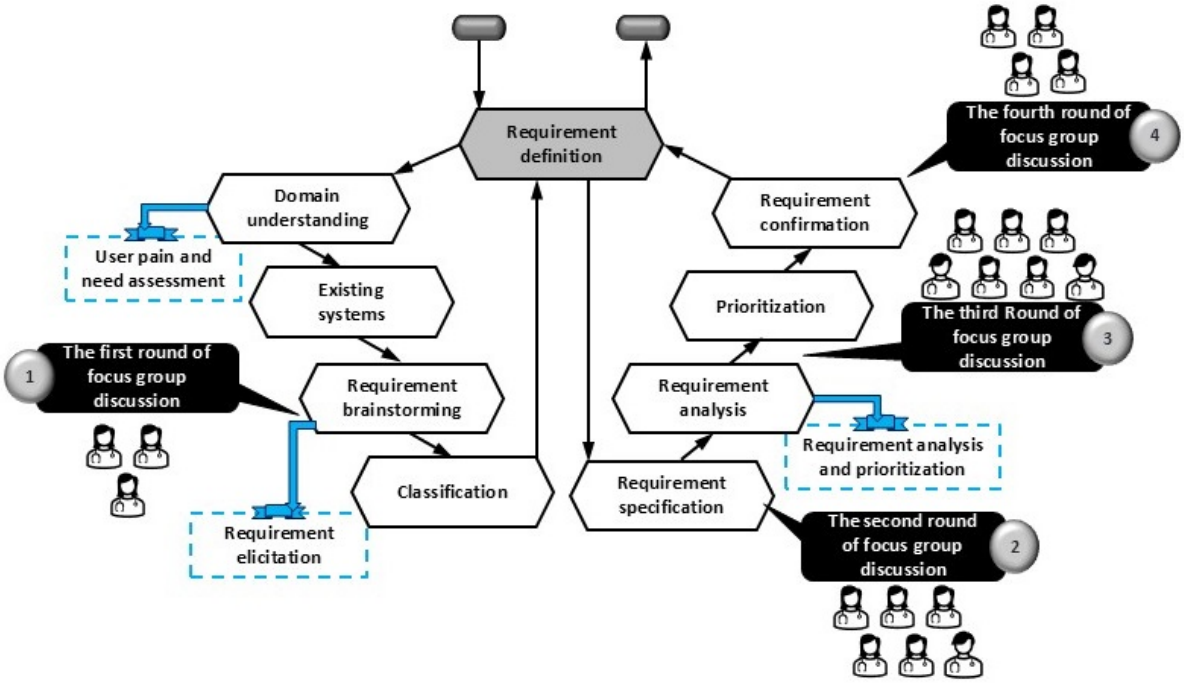
Methods

Overview

This study focused on expert opinions to determine the requirements of implementing and integrating data-driven medical software for EEG seizure monitoring in the current DCPE within hospitals and to assist health care professionals in decision-making. Although the proposed system emphasizes patient-centered solutions, the primary goal of this initial research was to collect only health care professionals' opinions on the system's needs. This study was done to implement the proposed system within the hospital only, whereas future research focuses more on home monitoring systems to take patient viewpoints into account.

A 4-round Delphi study, widely used in health care research to gather and refine expert opinions, was conducted using focus group discussions with health care professionals from May to December 2023. The aim was to identify the system requirements for integrating data-driven medical software for EEG seizure monitoring in DCPE. Several steps were taken while conducting the Delphi study to define the system requirements, leading us to form a process, as shown in Figure 1. The Delphi approach was selected for its ability to collect expert opinions and refine the system requirements iteratively and systematically. Each round had distinct objectives as follows:

Figure 1. The process of system requirement definitions for data-driven medical software for electroencephalography seizure monitoring in the digital care pathway for epilepsy.



XSL•FO
RenderX

Round 1 of the Delphi study was dedicated to identifying the current issues in the DCPE and user needs. The focus was on integrating data-driven medical software for EEG seizure monitoring, emphasizing the user-centric approach. This round was designed to ensure that the user needs were at the forefront of the study, making the audience feel that their needs were being prioritized.

Round 2 aimed at defining system requirements based on expert feedback from the first round.

Round 3 of the Delphi study was a pivotal stage where the system requirements were prioritized and refined through expert discussions. This thorough process was designed to ensure that the system aligned with clinical needs, instilling confidence in the audience about its alignment with their requirements.

Round 4 confirmed and validated the system requirements, incorporating final expert feedback for development.

The requirement analysis process began with the initial definition of system requirements, where user needs and pain points were assessed. On the basis of the domain understanding and comparison of the existing system, a guideline was prepared and presented to the health care professionals in the first round of the Delphi study. With their extensive experience and knowledge, these professionals were instrumental in starting the requirement elicitation. The idea of data-driven medical software for EEG seizure monitoring integration in DCPE was presented, and requirement brainstorming was started. The main point of this round of the Delphi, according to the health care professionals, was as follows:

> *The main issue is when the patients are alone during their seizure, or no one is around them to monitor their seizure. It would be nice if we can have the track of the seizures using EEG as well as the video and recording of the seizure to check the symptoms, the type of the seizure.* [Experts 1 and 2]

Therefore, the user pain point is the lack of monitoring and assistance for patients during their seizures, especially when no one is around them. The user needs a system that can track and record the patient's EEG and a video of the seizure to assess symptoms and seizure type and reduce the risk of any dangerous incident during the seizure. Currently, the hospital has video EEG recording systems in Finland that can track different types of seizures for several days up to 2 weeks. Monitoring and analyzing patients' historical seizure data can help predict and prevent future incidents. On the basis of reviewing existing systems and the first round of Delphi study discussions, the monitoring systems for epilepsy were classified into five main categories based on their aims: (1) self-management and medication adherence, (2) EEG-based systems for seizure detection, (3) cardiac-based systems, (4) non-EEG seizure detection systems using wearables, and (5) video-tracking systems. The system requirements were updated based on the study in the second category, EEG-based systems for seizure detection.

After updating the system requirement, the second round of the Delphi study was started with the aim of a requirement specification. Six health care professionals attended the session for 60 minutes. The main points of this session were as follows:

> *The type of sensors used for seizure remote monitoring are challenging. We can start with commercially available devices that have already been validated. The system can be specified for Continuous EEG monitoring at the ICU first. It would be nice to add more electrodes with additional channels into the current validated devices.* [Experts 1 and 2]

Subsequently, SRA and prioritization were initiated, and the third round of the Delphi study was conducted. Seven health care professionals, including an EEG signal processing specialist, worked closely with the industry to create EEG sensors. This session lasted 55 minutes. The experts discussed the preferences and system requirement prioritization in this session:

> *Probably, it is good to have some preferences considering how the proposed system fits our plans. For example, jumping right into the home remote monitoring and putting much effort into it might be not easy. Still, we can bind the project around developing the University Hospital care pathway for status epilepticus. EEG monitoring for the patients with status epilepticus or sudden unconsciousness due to unknown etiology at the emergency department, neurological department, or intensive care unit, could be the first target of the system, especially for the neurological department or intensive care unit.* [Expert 1]
>
> *The commercially available device at the current form requires a server that is installed inside the hospital network. But since we already have this kind of server in use in the intensive care unit.* [Expert 2]
>
> *It is better to start with the adults at the emergency and ICU department. After the system is validated, it can be specialized for patients of different age groups, preferably above 12 years old. The system can be used for patients with different ages and demographics.* [Experts 1, 2, and 3]

On the basis of the discussion, the development of data-driven medical software for EEG seizure monitoring needs to be started for the emergency department at the hospital targeting adult patients. If the system is adopted successfully, it can be extended to a home monitoring system, and a video recording system can be added as an additional feature. The home monitoring system is part of a plan that requires a requirements analysis study and co-design sessions with patients as the primary users of the system, along with other stakeholders.

Compatibility with existing hospital systems, such as electronic health records and EEG monitoring infrastructure, was a primary consideration. The EEG-EpiDigi system is designed to integrate with current systems without requiring extensive modifications. However, the hospital is preparing the new infrastructure to accommodate the modern version of EEG-EpiDigi.

XSL•FO
RenderX

The final interview round was done in December 2023 with 4 experts to summarize and confirm the system requirements and the discussion of the previous rounds. The interview lasted around 55 minutes.

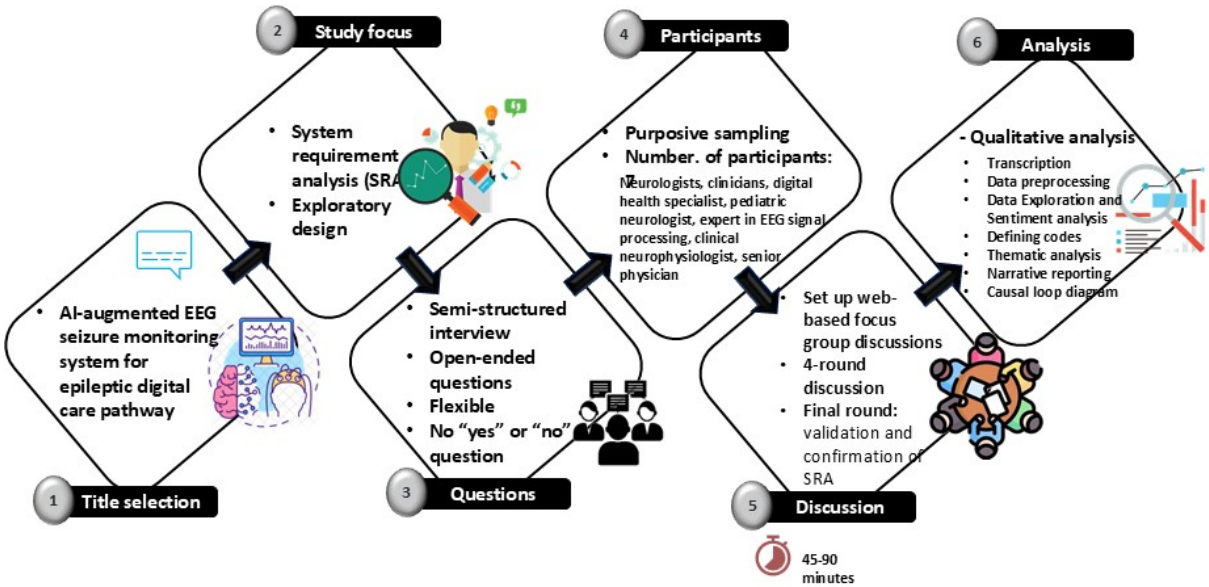
Study Design

To address the research questions, a 4-round Delphi study using focus group discussions was chosen to systematically gather expert opinions and iteratively refine the system requirements for data-driven medical software designed for EEG seizure monitoring (EEG-EpiDigi) in the DCPE (Figure 2).

The first round concentrated on identifying current issues within the DCPE and understanding user needs, focusing on integrating data-driven medical software for EEG-based seizure monitoring. In the second round, we aimed to define the system requirements based on the feedback received from experts in the first round. The third round prioritized and refined these requirements through further expert discussions to ensure the system met clinical needs. Finally, the fourth round confirmed and validated the system requirements, incorporating additional expert feedback to guide development.

After reviewing the existing systems and assessing the user needs and the shortcomings of the current DCPE based on discussion with health care professionals, data-driven medical software for EEG seizure monitoring was proposed, and the initial requirements were defined. The qualitative Delphi study using a focus group was centered on SRA for the proposed EEG-EpiDigi using exploratory design. A semistructured interview was conducted using a flexible, open-ended questionnaire. Each round of the Delphi study used a focus group that lasted 45 to 90 minutes and included 3 to 7 participants. The initial round of the Delphi study was conducted face-to-face, while the subsequent 3 rounds took place on the web via Teams (Microsoft Corp). This mixed method approach allowed for greater flexibility in expert participation, although it had a minimal impact on data collection and analysis. Thematic consistency was maintained across both formats, and the online discussions generated richer data due to the ease of participation. The first round focused on identifying user needs and the current issues related to DCPE. In contrast, the second to fourth rounds concentrated on defining system requirements, prioritizing them, and confirming their validity. The system requirements process was guided by the latest advancements in the field [17-19] and industry standards, including IEC 62304 for medical device software [20,21] and ISO 14971 for risk management [22-24]. These standards ensured that our approach aligns with best practices in medical software development. We analyzed the collected data qualitatively, coding the results to extract and present themes narratively, followed by interpretation.

Figure 2. The study design for a 4-round Delphi study using focus group discussions. AI: artificial intelligence; EEG: electroencephalography; SRA: system requirement analysis.



Participants

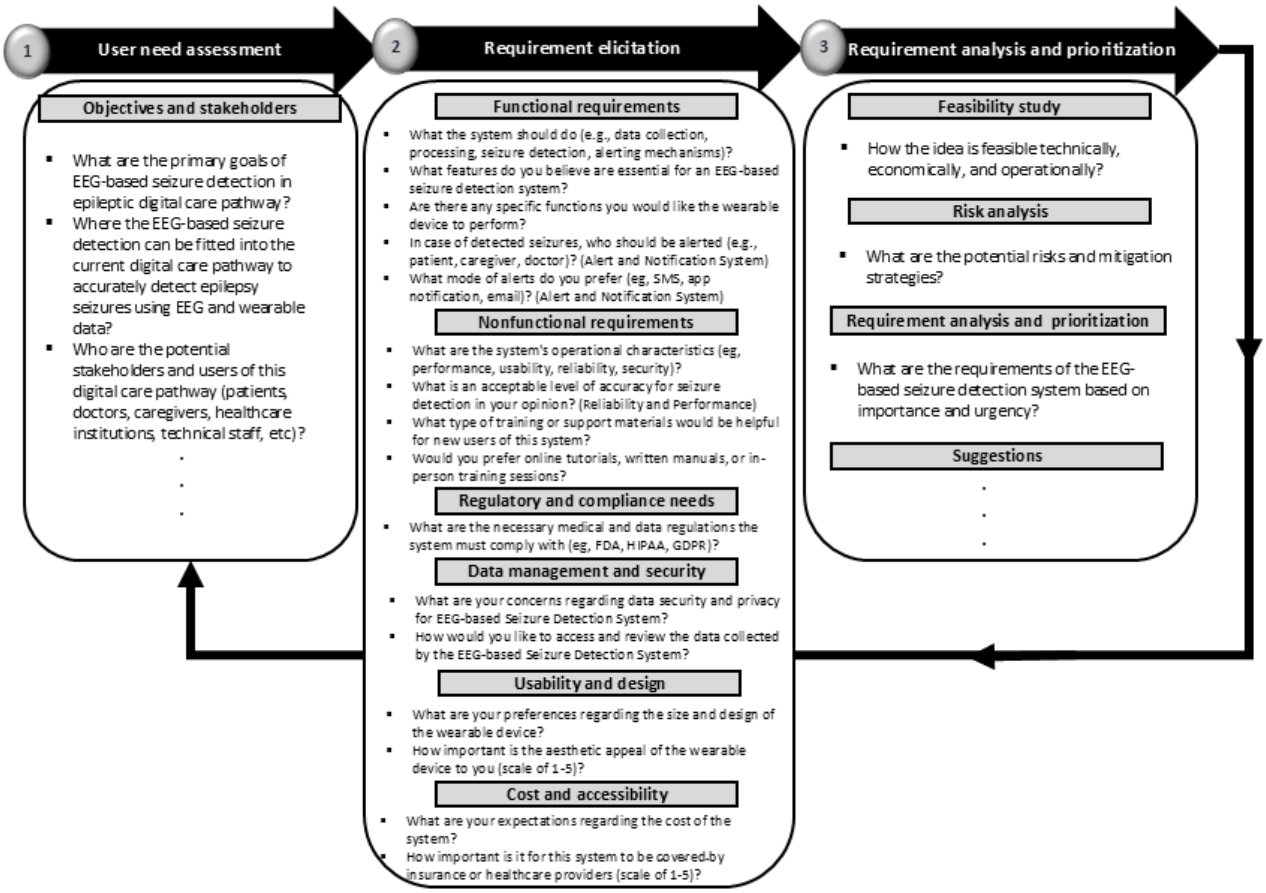
Purposive sampling aligns the sample with the goals and objectives of the study, thereby enhancing its rigor and the reliability of the data and findings [25]. For this study, purposive sampling was chosen to ensure that the expert panel is representative and diverse concerning epilepsy care and EEG monitoring. A group of health care professionals from Oulu University Hospital were selected, including neurologists, digital health specialists, pediatric neurologists, experts in EEG signal processing from the University of Oulu and industry, clinical neurophysiologists, and senior physicians. This diverse composition allowed for a comprehensive coverage of clinical, technical, and research perspectives. According to Burch et al [26], focus group discussions can be conducted with 4 to 8 participants for a minimum of 45 minutes. In this study, each round of the Delphi study was designed to include 4 to 7 participants and lasted between 45 and 90 minutes.

XSL•FO
RenderX

Interview Guideline

The Delphi study used focus group discussion and aimed at SRA for data-driven medical software for EEG seizure monitoring in DCPE. It was conducted using a semistructured interview guideline, as shown in Figure 2. To provide proper orientation, the interviewer was partially engaged in the discussion management. According to the studies by Keikhosrokiani [17], Catanio [27], Keikhosrokiani [28], and Grady [29], the interview guideline followed an iterative process from user need assessment to requirement elicitation, analysis, and prioritization. The iterative process encompassed 10 main topics: objectives and stakeholders, functional requirements, nonfunctional requirements, regulatory and compliance needs, data management and security, usability and design, cost and accessibility, feasibility study, risk analysis, analysis, and prioritization. The iterative process started from the initial round of Delphi and was enhanced for the next rounds. The interview questions were carefully designed based on a thorough review of relevant literature and insights from an expert panel, ensuring they aligned with the research questions. A pilot study was conducted before the primary research to improve the validity and clarity of the questionnaire. Feedback from this pilot study was essential in refining the questions, ensuring they effectively addressed the research objectives while remaining clear and accessible to participants. This iterative process significantly enhanced the reliability and relevance of the data collected from the Delphi study. Figure 3 illustrates the final version of the interview guideline, which was used for the fourth round of Delphi study discussions for requirement confirmation, composed in the previous rounds.

Figure 3. The interview guideline for a semistructured Delphi study using focus groups for requirement confirmation. EEG: electroencephalography; FDA: Food and Drug Administration; GDPR: General Data Protection Regulation; HIPAA: Health Insurance Portability and Accountability Act.



Data Analysis

The data analysis steps for the focus group discussion are shown in the final step of Figure 1. The transcribed data were initially read before the data analysis process to ensure the data's general content and scope. The data analysis began with data preprocessing and data cleaning to remove unnecessary words, symbols, etc, and the data were organized based on the topics. Next, data exploration using visualizations and sentiment analysis was performed to overview the story of the data at a glance. Afterward, data coding was done using qualitative tools to examine patterns, keywords, and concepts throughout the text. The broader themes representing larger text sections were defined based on the initial codes. The themes were reviewed to ensure they were coherent, distinct, and effectively represented the dataset. After explaining each theme, articulating what each cover was and how they relate to the content was done. Finally, the results were reported narratively with interpretation and quotations. A causal loop diagram was created for further analysis and discussion to complement the results.

Ethical Considerations

Ethical approval from the regional medical research ethics committee of the Wellbeing Services County of North

XSL•FO
RenderX

Ostrobothnia was granted for the EpiDigi project to involve human participants who are health care professionals and patients (approval number EETTMK: 59/2023). Verbal informed consent was taken from the participants. The names of the participants, products, and the company were anonymized. No compensation was provided to the participants.

Results

Data Exploration

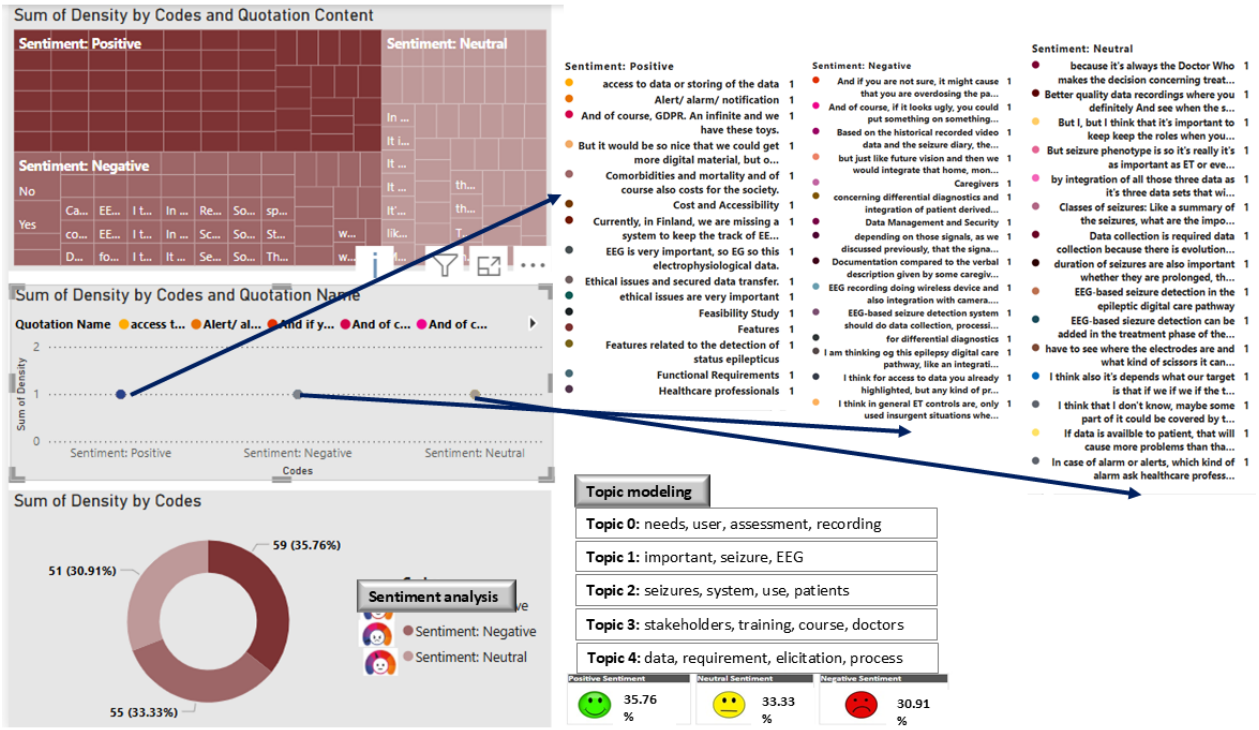
Data were explored from the interview transcripts before the thematic analysis to understand the overall data. Word frequency analysis, topic modeling, and sentiment analysis were used for text exploration, as shown in Figure 4.

The most frequent words in the word cloud indicated that health care professionals preferred wearable video EEG seizure detection systems at the hospital to provide accurate and timely data on seizure types for epilepsy classification and treatment decision-making. They requested courses to train them in using the proposed system and were interested in expanding the system for home monitoring using wearable devices.

On the basis of the topic modeling results, 5 main topics were highlighted from the data. Topic 0 focused on user needs and assessment, whereas topic 1 centered around data, seizures, EEG, and related concepts. Topic 2 discussed seizures, systems, use, and patients, while topic 3 concerned stakeholders, their views, and course of action. Finally, topic 4 was related to data requirements and elicitation processes.

Sentiment analysis was also performed on the data collected from the interviews. As shown in Figure 4, most sentiments toward the data-driven medical software for EEG seizure monitoring were positive, followed by neutral, and negative, respectively.

Figure 4. Data exploration for the interview using word frequency, topic modeling, and sentiment analysis. EEG: electroencephalography.



Study Mind Map for the Qualitative Data Analysis

The study mind map was prepared based on transcribed data to be used as a guide for SRA. As shown in Figure 5, SRA consists of 3 steps: user needs assessment, requirements elicitation, and requirements analysis and prioritization. The needs assessment evaluated the system's aims, objectives, and stakeholders. In contrast, functional requirements, nonfunctional requirements, regulatory and compliance needs, usability design, data management, and security were reviewed in the requirements elicitation and feasibility study, and risk analysis and requirement prioritization were considered in the requirements analysis and prioritization step. The main keywords for each subevaluation are addressed in Figure 5.

As outlined in Table 1, SRA consisted of 3 steps: user needs assessment, requirements elicitation, and requirements analysis and prioritization. The needs assessment step evaluated the system's aims, objectives, and stakeholders. In contrast, functional requirements, nonfunctional requirements, regulatory and compliance needs, usability design, data management, and security were reviewed in the requirements elicitation and feasibility study, and risk analysis and requirement prioritization were considered in the requirements analysis and prioritization step. The main keywords for each subevaluation are addressed in Table 1.

XSL•FO
RenderX

Figure 5. The study mind map for system requirement analysis.

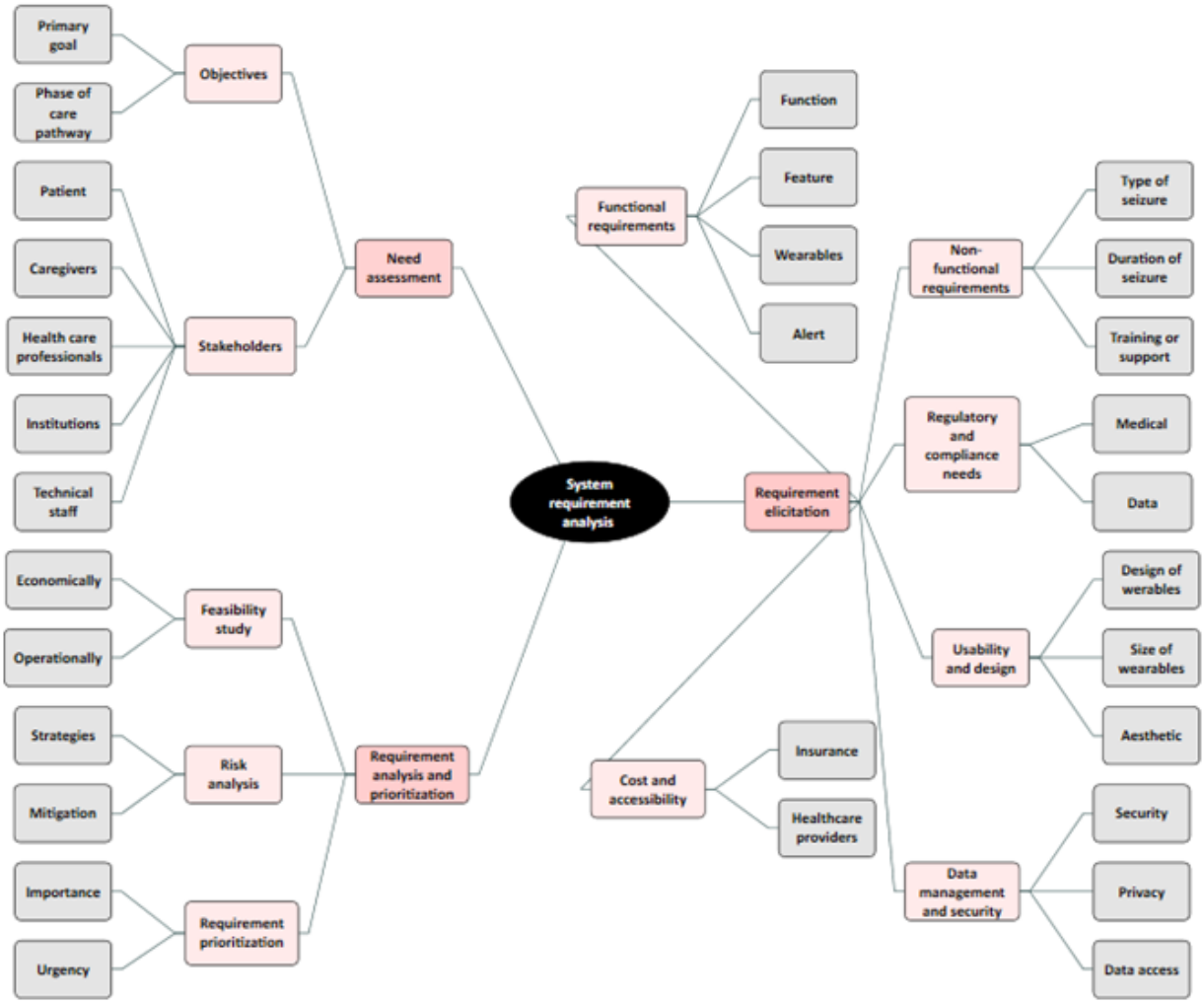


Table 1. The study mind map for system requirement analysis.

System requirements and subrequirements	Items
No 1. Need assessment	
Objectives	Primary goals; the phase of the care pathway
Stakeholders	Patients; caregivers; health care professionals; institutions; technical staff
No 2. Requirement elicitation	
Functional requirements	Function; feature; wearables; alert
Nonfunctional requirements	Type of seizure; duration of seizure; training and support
Regulatory compliance and needs	Medical; data
Usability and design (Method section)	Design of wearables; size of wearables; aesthetic
Data management and security	Security; privacy; data access
Cost and accessibility	Insurance; health care providers
No 3. Requirement analysis and prioritization	
Feasibility study	Economical; operational
Risk analysis	Strategies; mitigation
Requirement prioritization	Importance; urgency

XSL•FO
RenderX

Users Need Assessment

Primary Goals

After data were transcribed, coding was done to assist us in the thematic analysis. A coding network was generated for each part of the mind map, as shown in Figure 5. A sample of a network of coding for the primary goal of data-driven medical software for EEG seizure monitoring is illustrated in Figure 6.

The focus group discussion on user need assessment led us to some themes and subthemes, as summarized in Table 2. For instance, the experts mentioned the following:

> *EEG-based seizure detection is especially required in the emergency department for the status epilepticus phase. It is also helpful because we cannot get the 24/7 EEG recording from all those patients...*
>
> *[W]e need continuous electrocardiogram (ECG) monitoring all the time in the ICU. We can proceed with adult patients first, then use it for children...*

Thus, the primary goal was to implement EEG-based seizure detection specifically for emergencies, such as prolonged seizures or status epilepticus, starting from the emergency room to intensive care units (ICUs). A critical theme is the need to minimize *diagnosis time* during severe epileptic events, emphasizing rapid detection and treatment initiation. This discussion led us to the subtheme *enhancing emergency seizure management.*

Furthermore, the discussion emphasizes leveraging *remote detection technology to identify seizures*, potentially integrating AI-based analytics to enhance time and cost-efficiency. In addition, the discussion that "EEG-based wearable seizure detection system should be available and provided for doctors for differential diagnostics and treatment decision-making, especially if the seizure is prolonged..." and "Currently, we would need a recording, especially in the in the emergency in the status epilepticus phase" suggests the need for long-duration EEG recording, especially for status epilepticus.

Health care professionals continued the discussion by stating the following:

> *This wearable EEG-based seizure detection could help health care professionals in decision-making...thus, in this pilot process, we are focusing on emergency where you have a prolonged seizure, especially status epilepticus as a neurological emergency...We could start from the emergency room to the ICU intensive care unit.*

EEG-based seizure detection is mentioned as a tool to aid health care professionals in deciding whether to admit patients to *intensive care*, especially in nonconvulsive status epilepticus or nonepileptic seizures. EEG-based seizure detection in the DCPE treatment is proposed to improve patient management.

Moreover, the experts highlighted that "there is a need for EEG analytical models to assist the neurologists in differential diagnostics, especially in the evenings and weekends..." It was argued that there is a need for the most suitable type of wireless devices and wearables for seizure recording and video systems to detect seizures, highlighting the importance of technology in monitoring and data storage. Predictive analytics based on historical recorded video data and seizure diaries can be used to predict future seizures and their types, indicating a need for advanced analytic tools in health care.

The experts further mentioned, "Since the EEG analysis is time-consuming, if we could utilize AI-based preliminary analytics before the decision-making of health care professionals, it would be very efficient and cost-effective." This kind of solutions emphasizes the need for cost-effective solutions, suggesting that wireless EEG devices could aid in differential diagnostics and reduce the burden on clinical neurophysiologists. Furthermore, they emphasized the importance of integrating patient-derived data, device data, and electronic medical records for decision-making.

Future integration of home monitoring as part of the DCPE is touched upon as an extension of current hospital-based approaches.

XSL•FO
RenderX

Figure 6. A sample network of codes for user needs assessment. AI: artificial intelligence; EEG: electroencephalography; ICU: intensive care unit.

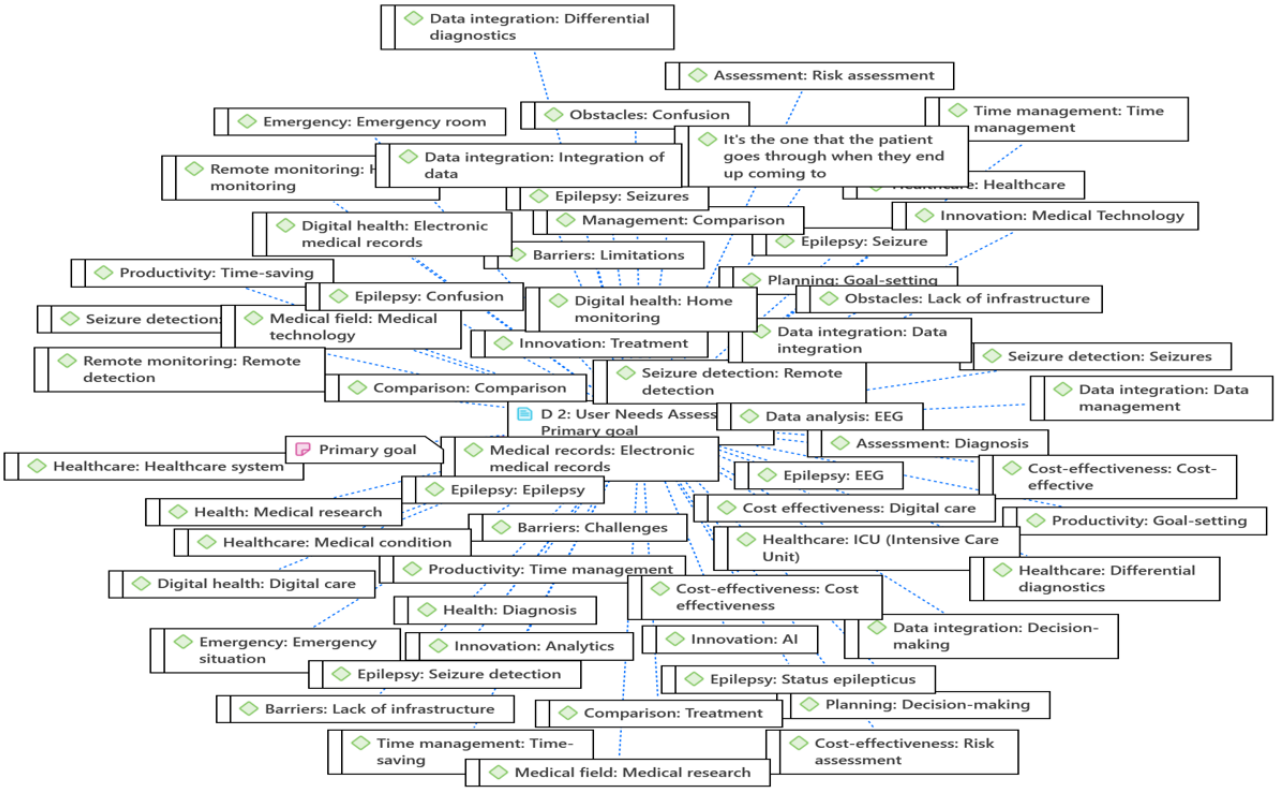


Table 2. A Summary of themes, subthemes, and codes for user need assessment.

Themes and subthemes	Codes extracted from quotations
Primary goal	
Enhancing emergency seizure management	Emergency, ICU[a], intensive care, seizure detection, management, and effective productivity
EEG[b]-based wearable seizure detection technology	EEG, seizure detection, and technology
Health care decision-making and treatment	Treatment, decision-making, health care, and digital health
Technological needs in monitoring and predictive analysis	Technology, seizure monitoring, remote monitoring, home monitoring, analytics, diagnosis, differential diagnostics, and predictive analytics
Cost and resource efficiency	Cost-effectiveness and timesaving
Future directions and needs	Limitations, challenges, barriers, medical field, and a lack of infrastructure
Phases of digital care pathway for epilepsy	
Development of an updated digital care pathway for epilepsy	Planning, updated care, and emergency care
Differentiation in seizure types	Differential diagnostics, seizure types, and status epilepticus
Integration of wearable EEG–based detection	EEG, seizure detection, integration, and data integration
Stakeholders	
Accessibility and use of medical data	Health care data, data analysis, and access
Decision-making in health care	Decision-making, medical decision-making, health monitoring, treatment decisions, and expert opinion
Data privacy and management	Confidentiality, patient privacy, patient engagement, patient data, and patient role
Role of data in treatment planning	Data analysis, health care data, treatment planning, and disagreement

[a]ICU: intensive care unit.

[b]EEG: electroencephalography.

XSL•FO
RenderX

Phases of DCPE

As highlighted in the primary goals mentioned subsequently, the primary aim is to implement the proposed system for hospital environments first, specifically for the *treatment phase* on the current DCPE:

> *Developing a system for hospital environments and then proceeding with remote monitoring systems from home...differentiating between epileptic and non-epileptic seizures...plans to incorporate EEG recordings into the DCPE, focusing on emergency and treatment phases for patients, especially for those with drug-resistant epilepsy, can assist health care professionals...*

If the patients successfully adopt the system, expanding it to the home remote monitoring system using suitably validated wearables is recommended.

The system's initial development within a hospital environment would be feasible, emphasizing hospitals' foundational role in the care pathway system. It mentions the subsequent *plan* to proceed with a remote monitoring system from the home, indicating a phased approach to implementation and an expansion of care beyond the hospital setting.

The importance of *epileptic versus nonepileptic seizures* was stated in the discussion to point out the need to separate seizures into epileptic and nonepileptic categories, suggesting a focus on specificity in diagnosis and treatment within the DCPE.

Plans to incorporate *EEG-based detection* into the DCPE were discussed, indicating a direction toward more advanced, technology-integrated solutions for monitoring and diagnosis. It was added that EEG-based detection is primarily used in emergencies and for follow-up treatment effects, but it is specifically aimed at health care professionals rather than patients. This highlights a focus on enhancing professional decision-making.

The pathway discusses a phased approach, starting with hospital-based systems and expanding to home monitoring. This indicates a comprehensive care strategy that extends beyond immediate hospital care.

Stakeholders

This section presents the user needs assessment focusing on various stakeholders, particularly relating to epilepsy management and the *use of data* in DCPE. The stakeholders identified include patients, caregivers, technical staff, institutions, and health care professionals. The document discusses the use of EEG recording data, seizure recordings, and other types of medical data.

The discussion suggests that the analysis of EEG recording should not be shown to patients or informal caregivers because it might not be beneficial and could lead to more problems than benefits. The assertion is that patients are generally not involved in decision-making related to diagnostics and thus might not need direct access to these data. In contrast, health care professionals are seen as requiring access to these data to make informed decisions about diagnosis, treatment, and care. The data can be transported or analyzed by another physician from a different hospital, indicating a need for data sharing and interoperability among health care institutions.

The discussion emphasizes the health care providers' role in making treatment decisions, often during inpatient consultations. It highlights the importance of keeping roles defined, with the physicians and patients or caregivers working together based on the information provided by the health care providers. While not explicitly mentioned, the concern over patient *access to data* and the need for health care professionals to share and analyze data suggest underlying themes *of data privacy, security, and effective data management* in health care. The discussion iterates that health care providers use data to inform treatment options and discuss these with patients and caregivers, depending on the situation. Data are central to *treatment planning* and the patient–care provider relationship. Furthermore, the following implications for each stakeholder were highlighted:

- For patients, improved understanding and communication about why specific data are not shared with them might be necessary.
- For health care providers, there is a need for practical tools and systems to manage, share, and analyze patient data securely and efficiently.
- For technical staff and institutions, ensuring technological solutions' reliability, usability, and security for data recording and predictive analysis is crucial.

Themes Relationships

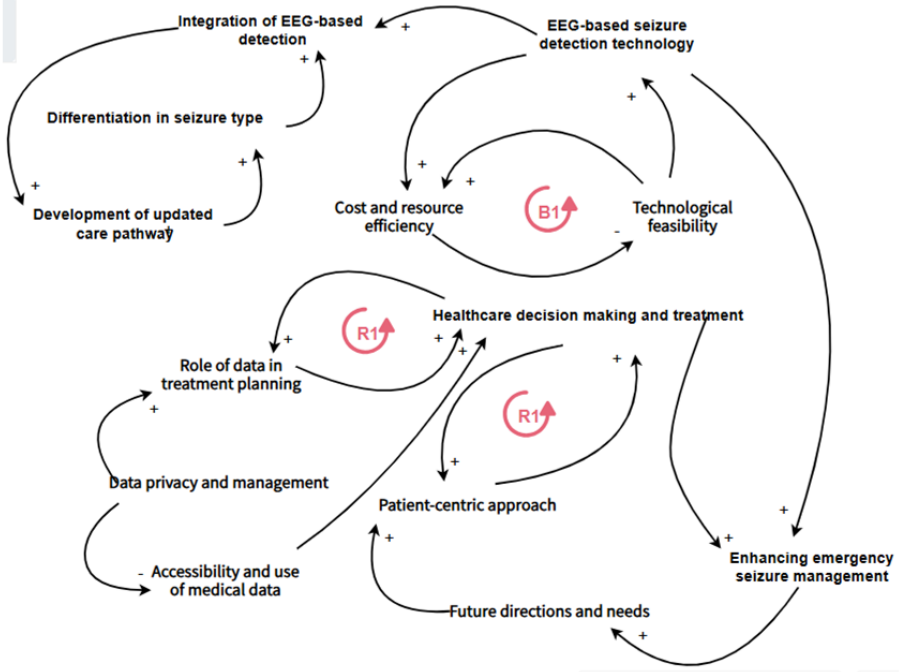
A causal relationship was created between the extracted themes, as shown in Figure 7, that shows the enhancement of emergency seizure management is significantly affected by the quality and accuracy of the EEG-based seizure detection technology. Furthermore, its improvement leads to more refined future directions and user needs, reflected in better patient-centric approaches. EEG-based seizure detection technology reduces the time and resources required for the system while improving cost and resource efficiency. Technological feasibility is the primary influence on the quality of EEG-based seizure detection technology. Better decision-making leads to improved patient outcomes, enhancing overall emergency seizure management. Improving cost-efficiency and resource efficiency enables further investment and improvement in the health care system. Balancing feedback might occur if costs reduce the ability to invest in newer technologies. Focusing on individual patient needs and outcomes (patient-centric approach) ensures the quality of emergency seizure management, and it is influenced by health care decision-making and treatment as better strategies and future technologies are developed.

As DCPE are updated and improved, there is a natural progression toward refining the understanding and categorization of conditions treated within those pathways, such as epilepsies. Once there is an established need or desire to differentiate seizure types more effectively, integrating more precise diagnostic tools such as EEG becomes a priority. With the integration of innovative technologies and diagnostic methods, the care pathway itself may be updated to incorporate these advancements, completing the loop and potentially starting a new cycle of updates and improvements.

XSL•FO
RenderX

A positive link between accessibility and the use of medical data and health care decision-making indicates that better accessibility improves decision-making capabilities. Moreover, informed decision-making positively influences how data are used in planning treatment. There is a negative link between data privacy and management and accessibility and the use of medical data, suggesting that inflexible data privacy policies might limit the accessibility of medical data. In contrast, adequate data privacy and management practices enhance the role of data in treatment planning. Finally, a positive feedback link between the role of data in treatment planning and health care decision-making illustrates that the effective use of data in treatment planning feeds back into improving decision-making in health care.

Figure 7. A causal loop diagram for the main objective. EEG: electroencephalography.



Requirement Elicitation

Functional Requirements

On the basis of the focus group discussion with health care professionals, the coding network is shown in Figure 8, and the themes and subthemes are listed in Table 3—functional requirements emphasize the functions of the systems, features, wearables, and alert systems.

The conversation "the functional requirements are probably very different for home monitoring and hospital use..." notes the variation in functional requirements between home monitoring and hospital use, emphasizing the need for *contextual adaptable* solutions tailored to different settings. According to the discussion, "It would be good to integrate all types of data concerning the seizures in the system, such as motions, movement, sleep, video," the system should *integrate various data types* (eg, motions, movements, sleep, and video) along with the number of seizures and device data to comprehensively understand the patient's condition.

Efficiency in seizure detection, data processing, and *alerting* mechanisms was mentioned as a vital component. Furthermore, the challenge of differentiating between epileptic seizures and nonepileptic events was highlighted. The requirement emphasized the need for accurate diagnosis to prevent overtreatment and adverse events.

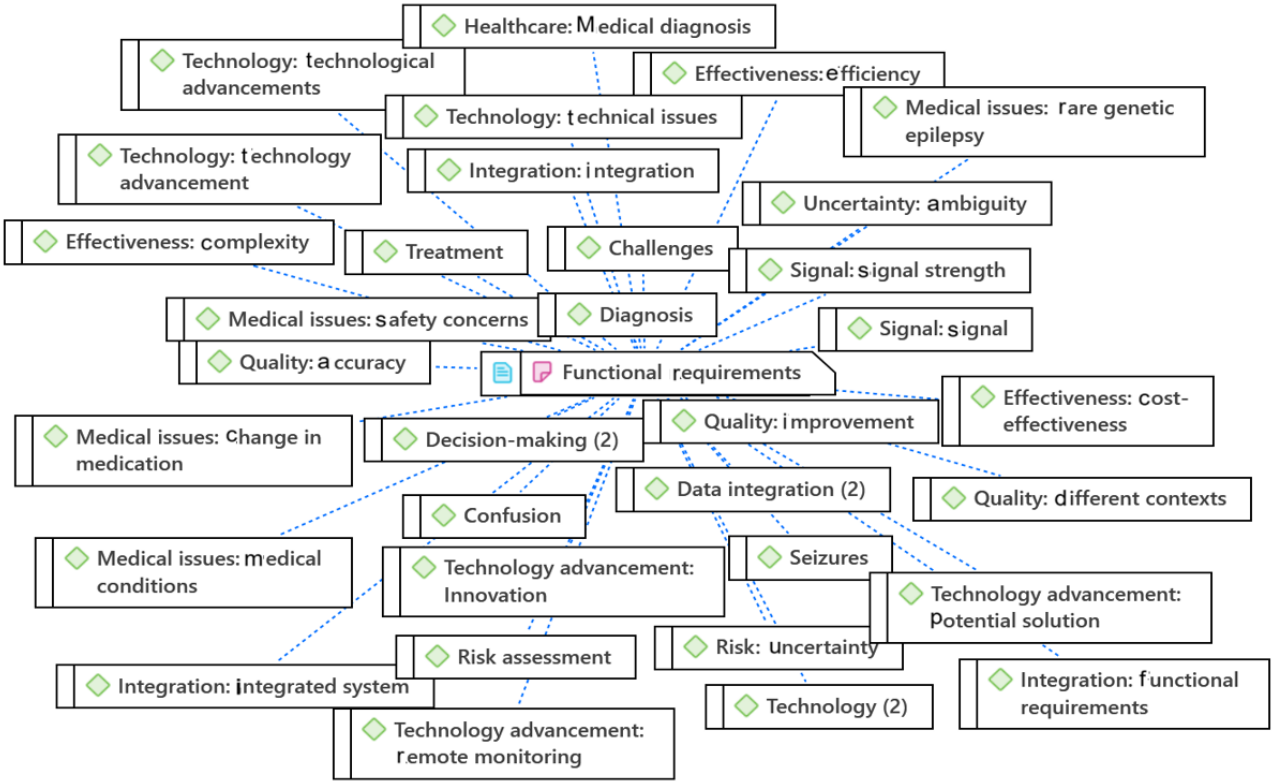
The evaluation of remote monitoring systems is closely linked to privacy and security risk concerns. Furthermore, real-time monitoring and the quality of service continue to pose challenges for remote monitoring services. Health care professionals identified several key issues with these systems, particularly concerning the types of wearable devices used for seizure detection, the accuracy of algorithms that classify seizures, and the overall trustworthiness of the system from the patient's perspective.

One major challenge highlighted by health care professionals was the need for real-time access to patients' seizure activity. Interviews conducted emphasized the evolving nature of epilepsy and stress the importance of continuous data collection and seizure phenotyping to adjust treatments and improve diagnostic accuracy. High-quality data recordings are crucial, especially for accurately capturing the onset of symptoms.

Integrating various data types, including video recordings, is essential for effective remote monitoring. The discussion also explored the potential of home remote monitoring systems that combine cameras, wireless EEG recording, and other technologies, envisioning a future where continuous and comprehensive monitoring is achievable. Wearable devices, which have already been tested, can be further refined for these remote monitoring systems.

XSL•FO
RenderX

Figure 8. A sample network of codes for functional requirements.



The need for *user compliance*, especially in scenarios such as detecting prolonged seizures at night, was mentioned. The system should *effectively* alarm caregivers and integrate smoothly with hospital protocols. Furthermore, the discussion outlined various *features* or requirements for a system, possibly related to health care or patient monitoring, given the references to EEG, electrocardiogram, and ICU monitoring.

The discussion about *quality and usability concerns* started by highlighting the need for quality and ease of use in devices, particularly those related to brain monitoring (eg, EEG systems). This suggests a priority on user-friendly and high-quality *wearable devices* for monitoring purposes. There was a specific mention of using "good quality wearables and devices for monitoring." This might refer to the need for reliable and accurate devices in a clinical setting, particularly for *continuous monitoring*.

The interview mentioned the need for continuous EEG monitoring in the ICU, indicating an emphasis on uninterrupted patient monitoring. This requirement is critical, starting with adult patients and extending to children. Moreover, the health care professionals discussed using *alarms* to detect nonconvulsive seizures. Emphasis on the placement of electrodes around the head and the reference to seizures spreading to the hemispheres suggested a focus on detailed and responsive monitoring systems that could detect and alarm medical staff about patient conditions. *Age-specific considerations were one of the critical points of discussion for the new system.*

The highlight of starting with adults and then extending monitoring capabilities to children indicates a phased, careful approach to implementing these technologies, likely due to the different physiological and anatomical considerations between adults and children. Prioritizing the features based on the *target demographics*, such as age group, and understanding which features are essential and desirable, requires stakeholder consultation.

For each requirement, technical feasibility and potential impact on patient care and health care workflow must be assessed [30]. Given the technical and clinical nature of the requirements, consulting with health care professionals and technology experts would be crucial in further refining and validating these needs. Additional steps would involve deeper stakeholder engagement, technical assessment of the proposed features, and a structured approach to integrating these requirements into a cohesive system design.

A health care professional stated the following:

> *We start with adults first, then we could implement a research project focusing on children...If we choose a device, it should be suitable for home use, minimize artifacts, and be user-friendly for patients or parents. It should be sensitive and specific...The device should inform users about seizures and be capable of detecting different kinds of seizures...We need to consider electrode placement and the types of seizures it can detect, focusing on devices that capture generalized seizures, particularly tonic-clonic seizures, while minimizing false alarms...Look into developing methods for ICU use, leveraging artificial intelligence...Determine the device's suitability for children of certain ages.*

XSL•FO
RenderX

Table 3. A summary of themes, subthemes, and codes for requirement elicitation.

Themes and subthemes	Codes extracted from quotations
Functional requirements	
Contextual adaptability	Adaptable solution tailored to different settings
Comprehensive data integration	Data type (eg, EEG[a], motions, movements, sleep, and video)
Alerting and detection efficiency	Alarm, alerting mechanism, and seizure detection
Differential diagnostics	Differential diagnosis, epileptic seizures, and nonepileptic events
Evolutionary data collection	Seizure phenotyping, treatment, diagnosis, and diagnostic precision
Quality of data recording	Quality, data recording, capturing, and video recording
Remote monitoring and technological integration	Home monitoring, remote monitoring, and wearables
User compliance and system effectiveness	Prolonged seizures, night, alarm, and caregivers
Quality and usability concerns	Ease of use, wearable device, user friendly, brain monitoring, and high-quality
Use of wearables and devices	Good quality, wearable device, clinical setting, and continuous monitoring
Continuous monitoring	Continuous EEG monitoring, uninterrupted patient monitoring, adult, and children
Alarm for seizure detection	Health care professionals, alarms, responsive monitoring systems, and nonconvulsive seizures
Age-specific considerations	Adults, children, age group, and features
Target demographic	Differing physiological and developmental needs, adults, children, specific age, and design
Usability and user friendly	Noninvasive, easy to use, user friendly, home care routines, simplicity, and intuitive design
Technical specifications	Sensitivity, specificity, customization, and tonic-clonic seizures
Information delivery	Inform, seizures, and communication
Innovation and development	Innovative, artificial intelligence, and predictive analytics
Nonfunctional requirements	
Types of seizures	Duration, nonepileptic, epileptic, nonconvulsive seizures, and tonic-clonic seizures
EEG importance	EEG, vital, and seizure
Seizure phenotypes	Types, variety, seizure, symptom, and abnormal
Duration of seizures	Duration and seizure
Materials and contents for training and support	Training, educational materials, on the web, in person, documentation, and digital content
Regulatory and compliance needs and data management and security	
Access and storing of data	Complexity, sensitivity, handling patient data, storing, and access
GDPR[b] compliance	GDPR, regulations, and storage
Ethical issues	Ethical considerations and authorized personnel
Standardization and legal issues	Lack of standardization and regulatory frameworks
Case study—Nelli system	Nelli system, developed by Tampere University and Tampere University Hospital
Usability and design	
Size and aesthetic considerations for patient devices	Device size, patient movement, younger patients, customizable elements, colorful caps, children, less intimidating, user friendly, and hiding
Designing for demographics and user-centered design in medical devices	Children, tailored design, personalized medical devices, patient's age, user-centered design, and physical psychological needs
Cost and accessibility	
Cost of implementation by hospital	Accessibility, affordable services, and brand-new hospital
Cost of services by insurance	Insurance and services
All public and private hospitals	Public and private

XSL•FO
RenderX

Themes and subthemes	Codes extracted from quotations
Rural areas	Remote and rural areas

[a]EEG: electroencephalography.

[b]GDPR: General Data Protection Regulation.

On the basis of the aforementioned discussion, the themes related to wearables were extracted. The device was initially targeted at adults and then adapted for children. This indicates a phased approach, likely due to differing physiological and developmental needs. *Age specificity* is another crucial factor when selecting a device. The device's suitability for children of certain ages is to be determined, indicating the need for customization or scalability in design. It was further discussed that the device should be *suitable for home use*, implying it needs to be noninvasive, *easy to use*, and potentially integrated with home care routines. It should be easy for the patients and parents to use, emphasizing simplicity and intuitive and *user-friendly* design. Regarding the *technical specifications*, the device must accurately detect seizures (high sensitivity) and not overreport nonseizure events (high specificity). It should minimize the recording of artifacts, which are irrelevant signals that could obscure seizure detection. Considering electrode placement and seizure types, it should be able to detect seizures, focusing on generalized and tonic-clonic seizures, and be amenable to customization or adaptable to various seizure types.

The *device should effectively inform users about seizures for information delivery*, indicating a need for clear alerts or communication mechanisms. The discussion about using current commercially available devices suggests leveraging existing innovative technology platforms as a starting point or benchmark. Considering the development of ICU use and leveraging AI, ambitions for advanced monitoring capabilities and predictive analytics are suggested.

Nonfunctional Requirements

The nonfunctional requirements were mainly emphasized in terms of seizure phenotypes, the importance of EEG, the duration of seizures, and training and support. While discussing the variety and importance of understanding different seizures, it was pointed out that although a seizure might be broadly categorized, the specific type and duration were critical. In addition, it emphasized the role of EEG in seizure detection and analysis, acknowledging its limitations since not all seizure types exhibited noticeable changes in EEG patterns depending on electrode placement. The significance of observable characteristics of seizures, such as facial expressions, eye movements, and other physical symptoms that are crucial for proper identification and need to be correlated with EEG data were highlighted. The importance of seizure duration, particularly in terms of the potential risks associated with prolonged episodes, such as status epilepticus was noted.

Regarding the types of seizures, the interview suggested a nuanced view of seizures, urging the need to understand different types and durations. For nonfunctional requirements, this theme emphasized the need for systems to accommodate a variety of seizure presentations and durations. EEG is recognized as a vital tool in seizure detection but is also noted for its limitations. Thus, there is a need for systems incorporating EEG data, but it is also possible to suggest integrating additional data sources and testing awareness. Moreover, the discussion described phenotypes as critical in seizure analysis, perhaps even more than EEG in some cases. This emphasizes the importance of comprehensive observational data in seizure monitoring software, possibly advocating for multimodal approaches. Duration of seizure was highlighted as a significant factor, especially concerning the risks associated with prolonged seizures. Systems must be sensitive to the duration of seizures, possibly incorporating alerts or interventions for protracted episodes.

The interview listed educational materials, including documentation, web-based tutorials, user manuals, in-person training sessions, and digital health parts. This theme explored the diversity and scope of materials available for epilepsy education and awareness. The materials were designed for different audiences, such as teenagers, caregivers, and patients.

A real-world application of EEG monitoring in a pediatric setting where a wearable device is used for seizure detection and continuous monitoring can be considered a practical example of this application. Incorporating case studies, such as a child with epilepsy who uses a wearable device that tracks seizures and alerts caregivers, could illustrate current technologies' effectiveness and limitations. This example can clarify how the device recognizes seizure patterns and integrates with mobile apps for parents and health care providers.

However, some challenges remain unresolved in integrating data-driven software into the DCPE, such as data privacy and technical integration. Data privacy remains a critical issue in health care technology. A practical challenge is how an EEG device processes patient data while ensuring compliance with General Data Protection Regulation (GDPR). There is a potential risk of unauthorized access if role-based access is not strictly enforced, and the design must include secure data transmission methods, such as encrypted channels. For technical integration, exploring how the EEG device communicates with other medical systems could highlight potential barriers, such as interoperability issues, where devices from different manufacturers need help to share data effectively. Incorporating perspectives from various stakeholders, including health care professionals, patients, and regulatory bodies, can provide a more comprehensive view. For example, engaging physicians in discussions about the importance of EEG data alongside observable seizure characteristics can help identify what features are essential for them. Gathering feedback from patients and caregivers about their experiences with existing devices can emphasize the user-centered design approach. Moreover, including insights from regulatory bodies can shed light on the compliance challenges they face, addressing gaps in current frameworks for epilepsy monitoring technology. By integrating these elements, we can create a more robust discussion that

XSL•FO
RenderX

addresses nonfunctional requirements and acknowledges the complexities of implementing health care technologies in a practical and user-friendly manner.

Regulatory and Compliance Needs and Data Management and Security

On the basis of the interview discussion, the SRA revealed several key themes surrounding regulatory and compliance needs in health care, particularly regarding data management, GDPR compliance, ethical considerations, and specific case studies or systems.

The critical nature of where and how data are accessed and stored was mentioned in the discussion, especially in hospital systems and for research purposes, highlighting the complexity and sensitivity of handling patient data. In addition, GDPR is mentioned as an overarching and strict regulation affecting the storage and handling of data. The health care professionals suggested that GDPR compliance is a significant challenge that requires continuous discussion and resolution at the ministry level. Ethical considerations in accessing patient data were highlighted, emphasizing that only authorized personnel (eg, physicians and nurses involved in patient care) should access patient data. The importance of role-based access and strict data access logging to maintain privacy and accountability was underlined. The discussion further suggested a lack of standardization, particularly for patients with epilepsy, indicating a need for consistent legal and regulatory frameworks across different diagnostic groups. Finally, a specific system called the Nelli system at Tampere University was described [31]. It details a program for handling patient data, including AI scanning, anonymous analysis, and specialized handling by trained nurses or technicians. The system appears to be a part of a digital health care initiative, possibly relating to epilepsy treatment, as it includes familiarity with EEG techniques.

Usability and Design

The "usability and design" discussion discusses the importance of size and aesthetics in medical devices, particularly those used by patients for home care or remote care in hospitals. On the basis of the extracted theme of *size and aesthetic considerations for patient devices, it was highlighted that* devices should be functional and stable without restricting patient movement, especially for younger patients. The devices should not cause discomfort or pain to the patient. Aesthetics can be improved by adding customizable elements such as colorful caps, particularly to appeal to children. The idea is to make the device less intimidating and more user friendly, hiding its medical nature.

> *To focus on designing for demographics and user-centered design in medical devices, making the devices less intimidating and more appealing to children implies a tailored design approach where the patient's age and preferences are considered. This suggests a move toward more personalized medical devices catering to specific demographics. The emphasis on comfort, lack of restriction in movement, and aesthetic appeal point to a user-centered design approach. It reflects an understanding that medical devices should be designed to meet the end user's physical and psychological needs beyond mere functionality. Human-centered design will be beneficial in the design of medical devices by reducing resources in software-related processes. It provides interactive and user-friendly solutions by focusing on user needs and requirements. User-centered design must involve software design, testing, and development phases* [32]

The discussion on size, aesthetics, and the balance between functionality and comfort in medical devices reflected broader themes in the medical device industry. These include the importance of user-centered design, the need for personalization in medical equipment, and the recognition that the success of a medical device is not solely dependent on its functional effectiveness, but also on its usability and acceptance by patients. Moreover, the health care providers mentioned the necessity of balancing technical requirements with human factors, advocating for designs that consider the patient's physical comfort, psychological well-being, and overall acceptance of the medical device.

Cost and Accessibility

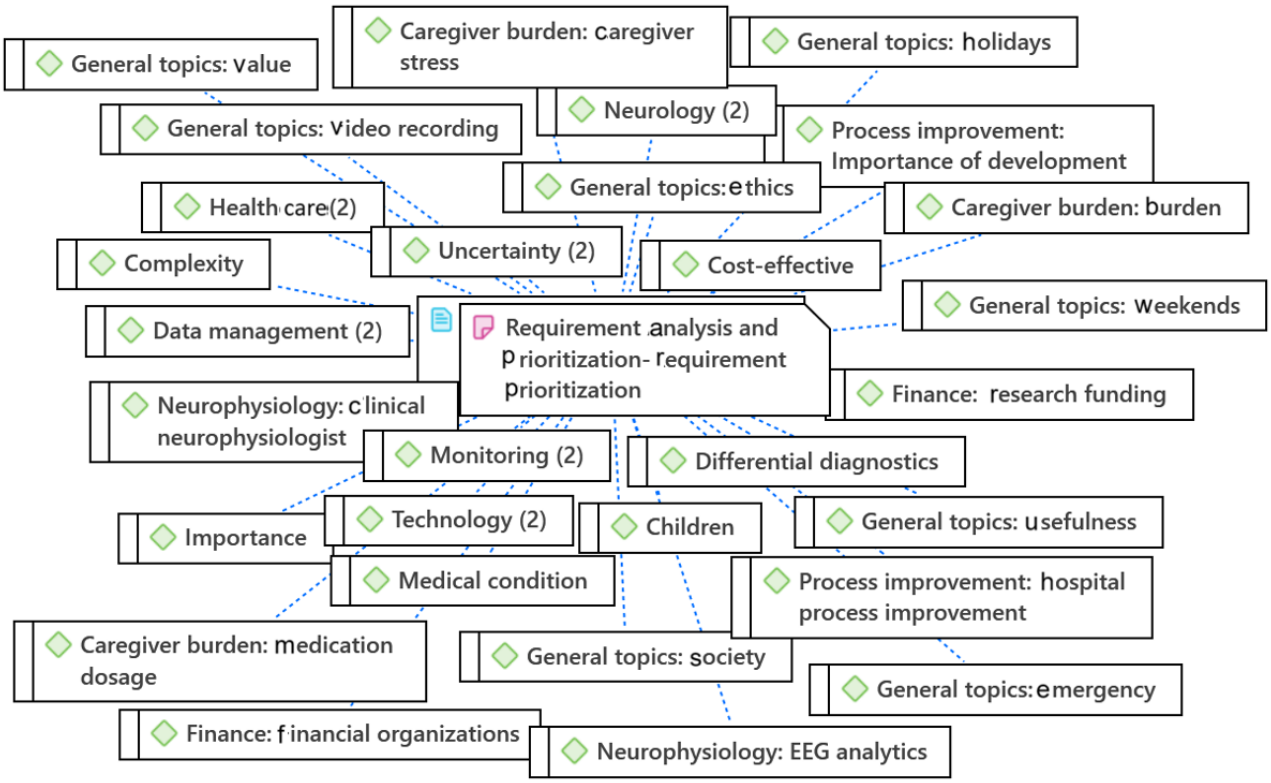
Regarding the cost and accessibility of the proposed system, the interviewees mentioned that "maybe some part of the cost of the services provided by the new system can be covered by the patient or their insurance systems, but the costs should not be like prohibiting the use of the system...however, the hospitals can bear the cost of developing and implementing the system...the system needs to be accessible to all public and private hospitals...also in the rural areas..." Moreover, they mentioned, "The system would be essential and cost-effective for the whole society, and we cannot estimate the value at the individual level." They concluded the interview by highlighting "Comorbidities and mortality and, of course, also costs for the society."

Requirement Analysis and Prioritization

As discussed in the interview, some priorities were set for developing the EEG-EpiDigi. For instance, the system will be implemented in the hospital environment with a target on the emergency unit and ICU. The system will be used for adults first and expanded for other age groups. If the system is adopted successfully, remote home monitoring will be added as an extra module. The relevant themes were extracted based on the network of codes for requirement prioritization illustrated in Figure 9. The list of themes, subthemes, and codes extracted from the quotations for SRA and prioritization are shown in Table 3.

XSL•FO
RenderX

Figure 9. An example network of codes for prioritizing requirements. EEG: electroencephalography.



Feasibility Study

In the focus group discussion, the health care professionals emphasized, "It would be more feasible to start from the hospital monitoring system and then expand it to a home remote monitoring system in the future." Therefore, the following options must be considered to assess the feasibility of integrating data-driven medical software for EEG seizure monitoring into the DCPE. On the basis of the themes and subthemes outlined in Table 3, the *demand for hospital or home remote monitoring systems* needs to be carefully evaluated. This demand was assessed through expert feedback, focusing on current clinical needs and future potential for home monitoring systems. The idea of a home monitoring system is the next step after implementing it in the hospital. Therefore, data collection for requirement analysis and co-design sessions with patients will be performed after successfully implementing the proposed system in the ICU. The experts prioritized hospital-based monitoring (ICU and emergency departments) before expanding to home-based systems. Evaluation criteria included feasibility, scalability, and clinical relevance. Furthermore, *technological resources need to be carefully planned.* Before starting the project, the technology and infrastructure available for monitoring systems must be examined. In addition, the impact of monitoring systems on patient care and *health outcomes* needs to be determined. The financial implications of implementing and maintaining monitoring systems must be estimated. The feasibility of EEG-EpiDigi is high for hospital-based use, particularly in ICU settings. However, challenges include data privacy, integration with existing hospital infrastructure, and ensuring AI interpretability for clinical decision-making. Further research is required to address these challenges, especially for home-based monitoring.

The feasibility of the EEG-EpiDigi system, designed for managing epilepsy through digital tools, hinges on several critical factors outlined in Textbox 1, including technological integration, user acceptance, regulatory compliance, and data management.

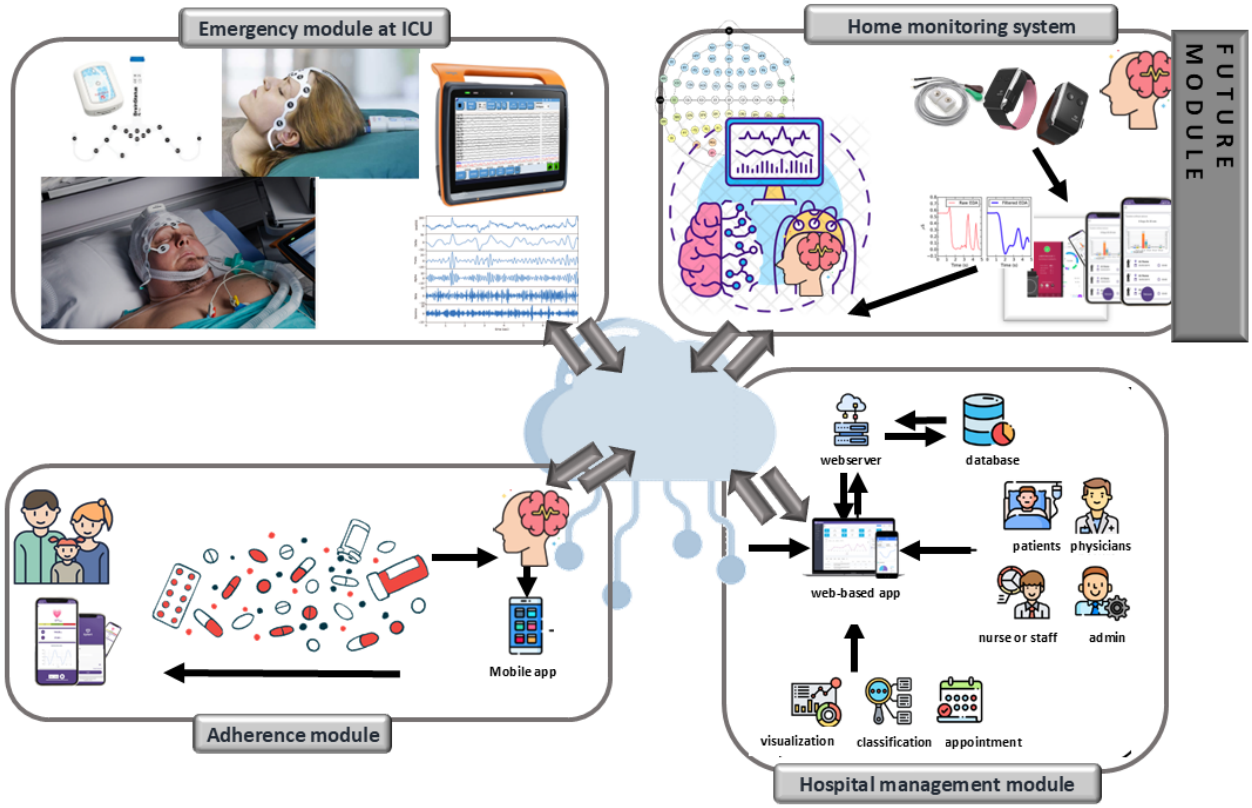
Based on the study's results and assessing the feasibility, developing and integrating a system capable of monitoring EEG recording is required and feasible in DCPE (Figure 10).

XSL•FO
RenderX

Textbox 1. Key factors influencing the feasibility of the electroencephalography (EEG)-EpiDigi system.

1. Technological integration: the successful implementation of EEG-EpiDigi will require seamless communication between EEG devices and other medical systems. Challenges include interoperability issues, where devices from different manufacturers need help to share data effectively. Ensuring that machine learning algorithms can accurately analyze the data from diverse sources is also vital. Creating standardized protocols and frameworks that enable this integration would be necessary while preserving data integrity.
2. User acceptance: engaging users, including patients, caregivers, and health care professionals, is crucial for the system's success. The system must be designed with a user-centered approach, addressing the needs and preferences of different demographics. For instance, if the design does not consider the comfort and aesthetics of devices used by children or young adults, it may face resistance from these groups. Continuous feedback loops should be established to refine user experiences based on real-world inputs.
3. Regulatory compliance: navigating the complexities of regulations such as General Data Protection Regulation (GDPR) is a significant challenge. The system must ensure that all patient data are stored, accessed, and transmitted promptly. This includes implementing strict role-based access controls and maintaining comprehensive data access records. Moreover, as regulations evolve, the system may need to adapt to new compliance requirements, which could entail ongoing updates and modifications.
4. Data management and security: the EEG-EpiDigi system would need robust mechanisms for data management to handle sensitive patient information. Security features such as encrypted data transmission and secure storage are paramount to mitigate the risks of unauthorized access. In addition, ensuring the ethical use of data, including informed consent from patients, is necessary to build trust and acceptance among users.
5. Continuous monitoring and adaptation: the nature of machine learning systems requires ongoing training and validation of algorithms to ensure their effectiveness over time. This includes continuously updating the system based on new research findings and user feedback. The challenge lies in establishing systems for routine monitoring that can keep pace with the rapid advancements in AI and epilepsy research.
6. Stakeholder collaboration: collaboration across various stakeholders, including health care professionals, patients, and regulatory bodies, will be essential to address the multifaceted challenges of the EEG-EpiDigi system. Building a platform for dialogue can help ensure that the system meets clinical needs while adhering to compliance and ethical standards. In summary, while the EEG-EpiDigi system has the potential to revolutionize epilepsy management through enhanced data integration and user engagement, careful consideration of technological, regulatory, and user-centered challenges is necessary for successful practical applications. Engaging diverse stakeholders and implementing robust data management practices will be crucial for navigating these challenges and maximizing the system's potential benefits.

Figure 10. The proposed system architecture for EEG-EpiDigi in the digital care pathway for epilepsy. ICU: intensive care unit.



Risk Analysis

The risk analysis of the proposed system in DCPE depends on the treatment effectiveness, health care system coverage, diagnostic tools, and resources of health care professionals. This involves understanding treatment, management, and factors impacting their efficacy. *The health care system coverage* theme suggested what aspects of treatment or diagnosis are covered by health care systems. Finally, *diagnostic tools and indications*

XSL•FO
RenderX

must be clarified to show how different diagnostic tools, such as EEG, are used and under what circumstances.

Requirement Prioritization

The "requirement prioritization" discussion focused on prioritizing requirements for the data-driven medical software for EEG seizure monitoring for epilepsy, specifically in a hospital environment. On the basis of Table 4, the discussion emphasized the importance of a *step-by-step* approach to developing the system. Furthermore, it was mentioned that there is a specific focus on patients with *urgent neurological emergency* events, such as status epilepticus. *The* interview discussed the need for *EEG analytics* and video recording in hospitals, particularly during evenings and weekends, for differential diagnostics. *Children* were highlighted as a unique group for consideration due to their inability to make decisions or consent themselves. Moreover, the importance of data *storage and privacy*, especially in EEG-based systems, was mentioned.

The system is mainly aimed at helping patients with *drug-resistant seizures*. The system's *utility during holidays* and weekends for private and government hospitals was noted, along with its voluntary nature. The discussion further underscored the importance of the system in terms of *cost-effectiveness* and its *value to society*. Moreover, it was highlighted that the implications of monitoring in chronic stages should be considered to avoid adding burdens on patients, families, and health care professionals. The discussion pointed out the need to carefully consider who the right person was to use the system on, which suggested the theme of decision-making *and responsibility*. The need to prioritize developing this system and securing funding from financial organizations was further emphasized. The system is intended to have a broad application across demographics for many users, including children, adults, teenagers, and older adults. The system must be successfully implemented for adults first, then expanded for other age groups. From this analysis, it is clear that the discussion was centered around developing and implementing a health care system with a strong focus on neurological emergencies, EEG analytics, and patient-centric considerations, especially regarding data privacy, accessibility, and cost-effectiveness.

Table 4. A summary of themes, subthemes, and codes for requirement analysis and prioritization.

Theme and subthemes	Codes extracted from quotations
Feasibility	
Demand for monitoring systems	Hospital, home, and remote monitoring systems
Technological resources	Technology, infrastructure, and start
Health care outcomes	Patient care and health outcome
Costs	Cost, implementation, maintenance, and implications
Risk analysis	
Treatment effectiveness	Treatment, effective, and efficacy
Health care system coverage	Aspects, treatment, and diagnosis
Diagnostic tools and indications	Different diagnostic tools, EEG[a], and circumstances
Requirement analysis and prioritization	
Step-by-step process	Step-by-step, developing
Focus on urgent neurological cases	Urgent, neurological, emergency events, and status epilepticus
EEG analytics in hospitals	EEG, analytics, video recording, and differential diagnostics
Special consideration for children	Children and special groups who are unable to make decisions give consent
Data storing and privacy	Data, storing, and privacy
Drug-resistant seizures	Drug-resistant seizures help patients
Accessibility and utility during off hours	Holidays, weekends, and voluntary
Cost-effectiveness and societal value	Cost-effective and society value
Chronic stages monitoring	Chronic stages, burden, patients, families, and medical care
Decision-making and responsibility	Careful decision, right person, and the use of the system
Priority in development and funding	Secure funding, organizations, priority, hospital environment, and emergency
Broad application across demographics	Wide range, different uses, children, adults, teenagers, and older people

[a]EEG: electroencephalography.

XSL•FO
RenderX

Quality of Care

The paper outlines several strategies aimed at improving the quality of care and resource management within the context of the proposed EEG seizure monitoring system. The identified strategies and their development process are outlined in Textbox 2.

The strategies were developed through a systematic 4-round Delphi study involving focus group discussions with health care experts. Details on the 4 rounds are as follows:

1. Identifying current issues within the DCPE and gather user needs
2. Defining system requirements based on the initial feedback
3. Prioritizing and refining these requirements through further expert discussions
4. Confirming and validating the specific system requirements for development

This iterative approach ensures the strategies are well-informed and aligned with clinical realities and user demands.

While the paper does not provide specific case studies, the methods used (eg, focus group discussions and expert panels) are grounded in recognized best practices for medical software development. The emphasis on user-centered design and compatibility with existing technologies reflects practices successfully used in various health care settings to support implementing new systems.

Furthermore, the intention to start with adults in emergency and ICU settings is supported by existing literature highlighting the importance of targeted, evidence-backed implementations in enhancing patient outcomes and optimizing resource use. Future cases could arise once the EEG-EpiDigi system is operational, as real-world applications will provide concrete examples of how these strategies improve care quality and resource management.

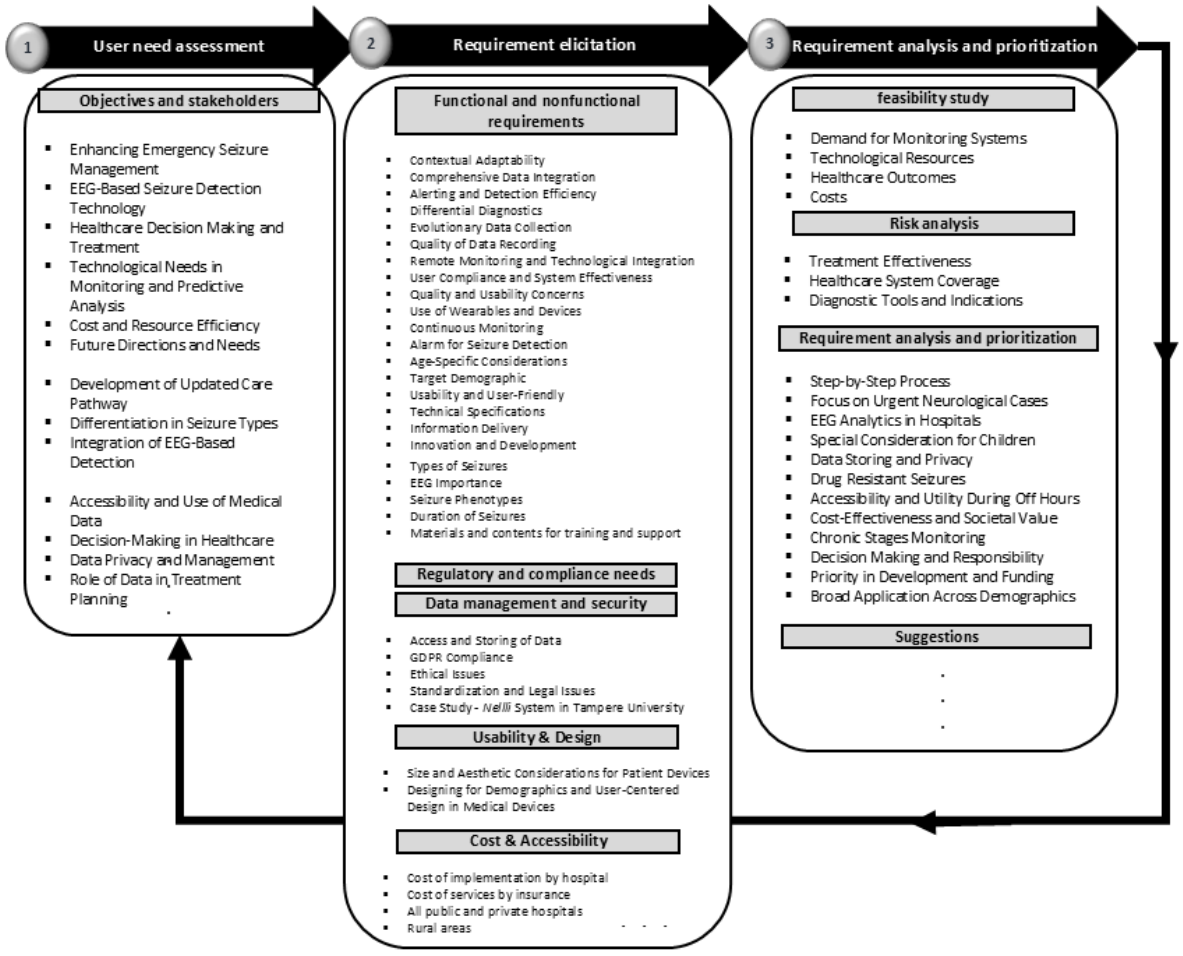
Based on the focus group discussion and thematic analysis results, this study proposed an SRA-EpiDigi for data-driven medical software to enhance EEG seizure monitoring quality in DCPE, as shown in Figure 11. SRA-EpiDigi addressed all research questions of this study.

Textbox 2. Strategies for improving care and resource management in electroencephalography (EEG) seizure monitoring.

1. Integration with existing systems: ensuring compatibility with current hospital systems, such as electronic health records and EEG monitoring procedures, is vital for seamless implementation. This approach minimizes disruptions and enhances workflow efficiency.
2. User-centered design: the development of the EEG-EpiDigi system incorporates co-design sessions with primary users, including patients and health care professionals, to gather insights and preferences. This helps tailor the system to meet user needs, enhancing user satisfaction and care quality.
3. Phased implementation: focusing the system on adult patients in emergency and intensive care settings allows for more controlled and effective validation. Once the system proves effective, it can be expanded to cater to pediatric populations.
4. Focus on data-driven decision-making: using real-time EEG data aid clinicians in making informed decisions, improving responsiveness and treatment accuracy.
5. Feedback loops and iterative development: the Delphi study method ensures that expert feedback is continuously integrated into the design and development process. Each round serves to refine and adjust system requirements based on expert input.

XSL•FO
RenderX

Figure 11. The proposed system requirement analysis (SRA-EpiDigi) for data-driven medical software for electroencephalography seizure monitoring in digital care pathway for epilepsy. EEG: electroencephalography; GDPR: General Data Protection Regulation.



Discussion

Principal Findings

The World Health Organization has stated that epilepsy is among chronic diseases causing the most comorbidity and mortality, and that it also incurs significant costs for society. It affects people of all ages, and it is one of the most common neurological diseases worldwide [2]. This highlights the severity and broader impact of epilepsy beyond the individual, affecting public health systems and society's economic burden. As mentioned earlier, the health care professionals reviewed the current DCPE's shortcomings in the focus group discussion. It was emphasized that despite several sophisticated electrocardiograms monitoring heart problems for decades, there are no comparable methods for EEG signal monitoring. The determination of seizure type is critical for choosing an appropriate medication, but predicting patient response to drugs is another challenge. About one-third of people with epilepsy have incomplete seizure control with current medicines, highlighting the need for more effective treatments [13]. Furthermore, the health care professionals highlighted the following:

> *Currently, the time taken to diagnose epilepsy is too long, and there is a need to minimize this duration. It would be a real risk to the patient, as the seizure could potentially prolong further.*

Therefore, developing and integrating a system capable of monitoring EEG recording is required in DCPE (Figure 10). The focus group discussion outlines a need for advanced wearable wireless EEG–based seizure detection tools in emergency care for epilepsy, emphasizing quick diagnosis, efficient treatment, and cost-effective technology integration. However, implementing widespread sensor-based EEG seizure monitoring software, especially with AI integration, requires assessing technological feasibility and reliability. The solution would be cost-efficient, implying that budget and resource allocation will be key considerations. The patient-centric approach was suggested for further improvements in DCPE, especially in emergency contexts, pointing toward a more patient-centric approach focusing on rapid and personalized care. The study analysis revealed a comprehensive approach to enhancing care for epileptic seizures, suggesting significant implications for patient treatment, health care resource management, and technology integration in medical practices.

The proposed integration of EEG-based detection signifies an inclination toward the influence of technology to improve diagnostic accuracy and treatment effectiveness. This implies a broader digital transformation within health care systems and reflects an essential theme of enhancing professional capabilities while hinting at a need for more patient-centered tools in future developments.

XSL•FO
RenderX

The interview further outlines a focused, data-driven approach to developing a sustainable DCPE [32,33]. It prioritizes professional health care delivery, technological integration, and a phased expansion from hospital to home settings. Each theme reflects a commitment to improving care through specificity in diagnosis (differentiating seizure types), advanced monitoring technologies (EEG integration), and targeted user groups (health care professionals). The discussion suggests an evolving health care model that is adaptive, technologically advanced, and professional-centric. As health care continues to grow, such systems are crucial in guiding the development of more efficient, accurate, and comprehensive DCPE.

The functional requirements revolve around a versatile, accurate, and comprehensive seizure detection system that can adapt to different settings, integrate various data types, ensure high-quality recordings, and evolve with patient needs. There is an emphasis on balancing technological innovation with user needs and medical requirements to improve care for patients with epilepsy. Adopting and improving upon existing technologies and exploring data-driven systems, AI, and ICU applications show a commitment to cutting-edge solutions and broad applicability. The discussion, while brief, brings forward several essential aspects to consider in the context of nonfunctional requirements for seizure-related systems or analysis. It highlights the diversity and complexity of seizures, the roles and limitations of EEG, the importance of phenotypic observations, and the critical nature of seizure duration. Each of these themes has implications for the design and function of seizure monitoring, detection, and analysis systems, suggesting a need for comprehensive, adaptable, and sensitive approaches beyond traditional EEG data. This thematic analysis suggests directions for further research or development in enhancing seizure-related health care or monitoring technologies.

Furthermore, the discussion indicates the critical role of data in health care, the importance of clear roles and communication between patients and health care providers, and the potential of technology to improve patient care through advanced monitoring and predictive analytics. It also highlights the need to consider data accessibility and privacy concerns carefully. The study provides a glimpse into the complexities of health care regulation, especially concerning data handling, privacy, GDPR compliance, ethical considerations, and the need for specialized systems and training for health care professionals.

The study discusses the balance between making a functional and comfortable device for the patient. This includes not restricting patient movement, particularly for younger patients, and ensuring that the device does not cause discomfort or pain. Although functionality is paramount, aesthetics also plays a significant role in patient acceptance and comfort. The discussion suggests that aesthetics can be improved with customizable elements such as colorful caps, especially to make devices less intimidating for children. This shows a clear understanding that while the primary goal of a medical device is to serve a health-related function, its acceptance by patients, especially younger ones, can be significantly enhanced through thoughtful design that considers aesthetic appeal.

Conclusions

In conclusion, this study incorporates health care professionals' opinions, which can contribute significantly to a sustainable care pathway by providing insights and data-driven approaches that optimize resource allocation, improve patient outcomes, and integrate long-term solutions for patient care. Moreover, it offers valuable insights into the complex interplay among functionality, aesthetics, and user acceptance in medical system design. This research supports a thorough approach that considers patients' physical and emotional needs, aiming to enhance the overall user experience and acceptance of medical software in various care settings. This study seems to be a concise set of suggestions (Figure 8) regarding improved health care technology for care pathways for epilepsy. It draws attention to the advancements in epilepsy health monitoring to highlight the gap and potential in neurological monitoring. It touches on public health themes, technology disparity, and the call for future-focused health care innovations.

The proposed SRA-EpiDigi for data-driven medical software for EEG seizure monitoring in DCPE requires ongoing improvements. One major limitation identified is the lack of patient involvement in the focus group discussions, which primarily included health care professionals. Since the system will initially be implemented within a hospital setting to aid health care providers in decision-making, it is essential to incorporate patient perspectives. To enhance the study's future direction, emphasis should be placed on requirement analysis and co-design sessions, particularly for home monitoring systems tailored for people with epilepsy. The scope of development should not only include the EEG seizure monitoring software but also expand to real-time seizure video analytics.

In addition, these co-design sessions can effectively engage system developers and health care professionals in selecting optimal, user-friendly wearables and sensors for improved seizure management. This approach will ensure the system remains relevant and user friendly for health care providers and patients, thus enhancing its overall effectiveness and acceptance.

Acknowledgments

MI and JU contributed equally to this work. The authors thank the Oulu University Hospital, the health care professionals, and the related industry for their precious time and valuable feedback. Furthermore, the authors acknowledge using the generative artificial intelligence tool Grammarly [34] to edit the paper and address the reviewer's comments.

The authors express their heartfelt gratitude for the immense support from the 6G-enabled Sustainable Society (6GESS), Oulu University Hospital, healthcare professionals, and industry partners for their valuable time and feedback. Additionally, we note that some authors are members of the European Reference Network on Rare and Complex Epilepsies (EpiCARE). Lastly, the

XSL•FO
RenderX

authors acknowledge using the generative artificial intelligence tool, Grammarly, to edit the paper and address the reviewers' comments.

Conflicts of Interest

None declared.

References

1. Sander JW. Epileptic seizures. In: Aminoff MJ, Daroff RB, editors. Encyclopedia of the Neurological Sciences (Second Edition). Cambridge, MA: Academic Press; 2014.
2. Epilepsy. World Health Organization. 2024 Feb 7. URL: https://www.who.int/news-room/fact-sheets/detail/epilepsy [accessed 2025-04-02]
3. Keikhosrokiani P, Polus M, Guardado Medina S, Isomursu M. The effectiveness of medical adherence mobile health solutions for individuals with epilepsy: protocol for a systematic review. JMIR Res Protoc 2024 Aug 06;13:e55123 [FREE Full text] [doi: 10.2196/55123] [Medline: 39106484]
4. Jeppesen J, Christensen J, Johansen P, Beniczky S. Personalized seizure detection using logistic regression machine learning based on wearable ECG-monitoring device. Seizure 2023 Apr;107:155-161 [FREE Full text] [doi: 10.1016/j.seizure.2023.04.012] [Medline: 37068328]
5. Alzamanan MZ, Lim KS, Akmar Ismail MA, Abdul Ghani N. Self-management apps for people with epilepsy: systematic analysis. JMIR Mhealth Uhealth 2021 May 28;9(5):e22489 [FREE Full text] [doi: 10.2196/22489] [Medline: 34047709]
6. Jibon FA, Miraz MH, Khandaker MU, Rashdan M, Salman M, Tasbir A, et al. Epileptic seizure detection from electroencephalogram (EEG) signals using linear graph convolutional network and DenseNet based hybrid framework. J Radiat Res Appl Sci 2023 Sep;16(3):100607 [FREE Full text] [doi: 10.1016/j.jrras.2023.100607]
7. Liu Y, Jiang B, Feng J, Hu J, Zhang H. Classification of EEG signals for epileptic seizures using feature dimension reduction algorithm based on LPP. Multimed Tools Appl 2020 Jun 24;80(20):30261-30282 [FREE Full text] [doi: 10.1007/s11042-020-09135-7]
8. Rasheed K, Qayyum A, Qadir J, Sivathamboo S, Kwan P, Kuhlmann L, et al. Machine learning for predicting epileptic seizures using EEG signals: a review. IEEE Rev Biomed Eng 2021;14:139-155. [doi: 10.1109/RBME.2020.3008792] [Medline: 32746369]
9. Wong S, Simmons A, Rivera-Villicana J, Barnett S, Sivathamboo S, Perucca P, et al. EEG datasets for seizure detection and prediction- a review. Epilepsia Open 2023 Jun;8(2):252-267 [FREE Full text] [doi: 10.1002/epi4.12704] [Medline: 36740244]
10. Rahman R, Varnosfaderani SM, Makke O, Sarhan NJ, Asano E, Luat A. Comprehensive analysis of EEG datasets for epileptic seizure prediction. In: Proceedings of the IEEE International Symposium on Circuits and Systems. 2021 Presented at: ISCAS 2021; May 22-28, 2021; Daegu, Korea. [doi: 10.1109/iscas51556.2021.9401766]
11. Chen Z, Lu G, Xie Z, Shang W. A unified framework and method for EEG-based early epileptic seizure detection and epilepsy diagnosis. IEEE Access 2020;8:20080-20092 [FREE Full text] [doi: 10.1109/access.2020.2969055]
12. Tveit J, Aurlien H, Plis S, Calhoun VD, Tatum WO, Schomer DL, et al. Automated interpretation of clinical electroencephalograms using artificial intelligence. JAMA Neurol 2023 Aug 01;80(8):805-812 [FREE Full text] [doi: 10.1001/jamaneurol.2023.1645] [Medline: 37338864]
13. Privitera M. Current challenges in the management of epilepsy. Am J Manag Care 2011 Jun;17 Suppl 7:S195-S203. [Medline: 21761951]
14. Özcan-Top Ö, McCaffery F. A hybrid assessment approach for medical device software development companies. J Softw Evol Process 2018;30(7):e1929 [FREE Full text] [doi: 10.1002/smr.1929]
15. Rahanu H, Georgiadou E, Siakas K, Ross M. Health, safety, and well‐being through ethical behaviours in computer and information systems development and deployment. J Softw Evol Proc 2019 Aug 16;32(3):e2225 [FREE Full text] [doi: 10.1002/smr.2225]
16. Ouhbi S, Fernández‐Alemán JL, Toval A, Rivera Pozo J, Idri A. Sustainability requirements for connected health applications. J Softw Evol Proc 2017 Nov 28;30(7):e1922 [FREE Full text] [doi: 10.1002/smr.1922]
17. Keikhosrokiani P. Perspectives in the Development of Mobile Medical Information Systems: Life Cycle, Management, Methodological Approach and Application. Cambridge, MA: Academic Press; 2019.
18. Grady JO. 2 - Requirements foundation. In: System Requirements Analysis (Second Edition). Amsterdam, The Netherlands: Elsevier; 2014:93-150.
19. Dooley JF, Kazakova VA. Software Development, Design, and Coding: With Patterns, Debugging, Unit Testing, and Refactoring. Berkeley, CA: Apress; 2024.
20. IEC 62304:2006 medical device software — software life cycle processes. International Organization for Standardization. 2006. URL: https://www.iso.org/standard/38421.html [accessed 2025-02-18]
21. Jordan PMDSSLP. InIET Seminar on Software for Medical devices. pp 41? 2025:47. [doi: 10.1049/ic:20060141]

XSL•FO
RenderX

22. Hegde V. Case study — risk management for medical devices (based on ISO 14971). In: Proceedings - Annual Reliability and Maintainability Symposium. 2011 Presented at: RAMS 2011; January 24-27, 2011; Lake Buena Vista, FL. [doi: 10.1109/rams.2011.5754492]
23. Chan T, Tong RK. ISO 14971: application of risk management to medical devices. In: Handbook of Medical Device Regulatory Affairs in Asia. Boca Raton, FL: CRC Press; 2018.
24. Flood D, McCaffery F, Casey V, McKeever R, Rust P. A roadmap to ISO 14971 implementation. J Softw Evol Proc 2015 Apr 14;27(5):319-336 [FREE Full text] [doi: 10.1002/smr.1711]
25. Campbell S, Greenwood M, Prior S, Shearer T, Walkem K, Young S, et al. Purposive sampling: complex or simple? Research case examples. J Res Nurs 2020 Dec;25(8):652-661 [FREE Full text] [doi: 10.1177/1744987120927206] [Medline: 34394687]
26. Burch J, Taylor C, Wilson A, Norton C. "You're just on your own": exploring bowel symptom management needs after rectal cancer surgery through patient and clinician focus groups. Eur J Oncol Nurs 2023 Dec;67:102406 [FREE Full text] [doi: 10.1016/j.ejon.2023.102406] [Medline: 37804751]
27. Catanio JT. Requirements analysis: a review. In: Proceedings of the Advances in Systems, Computing Sciences and Software Engineering. 2005 Presented at: SCSS 2005; 2005; Online. [doi: 10.1007/1-4020-5263-4_64]
28. Keikhosrokiani P. Chapter 2 - Mobile Medical Information System (mMIS) requirement. In: Perspectives in the Development of Mobile Medical Information Systems: Life Cycle, Management, Methodological Approach and Application. Amsterdam, The Netherlands: Academic Press; 2020:25-35.
29. Grady J. System Requirements Analysis. New York, NY: McGraw-Hill Education; 1993.
30. Keikhosrokiani P. Chapter 9 - Feasibility study for Mobile Medical Information Systems (mMISs). In: Perspectives in the Development of Mobile Medical Information Systems: Life Cycle, Management, Methodological Approach and Application. Amsterdam, The Netherlands: Academic Press; 2020:135-145.
31. Basnyat P, Mäkinen J, Saarinen JT, Peltola J. Clinical utility of a video/audio-based epilepsy monitoring system Nelli. Epilepsy Behav 2022 Aug;133:108804 [FREE Full text] [doi: 10.1016/j.yebeh.2022.108804] [Medline: 35753111]
32. Keikhosrokiani P, Isomursu M, Uusimaa J, Kortelainen J. A sustainable artificial-intelligence-augmented digital care pathway for epilepsy: automating seizure tracking based on electroencephalogram data using artificial intelligence. Digit Health 2024;10:20552076241287356 [FREE Full text] [doi: 10.1177/20552076241287356] [Medline: 39381810]
33. Khan S, Keikhosrokiani P, Uusimaa J, Annunen J, Komulainen-Ebrahim J, Polus M, et al. Design sustainability goals for digital care pathway for epilepsy: a healthcare professionals' perspective. In: Proceedings of the First Nordic Conference on Digital Health and Wireless Solutions. 2024 Presented at: NCDHWS 2024; May 7-8, 2024; Oulu, Finland. [doi: 10.1007/978-3-031-59080-1_20]
34. Grammarly homepage. Grammarly. URL: https://www.grammarly.com/ [accessed 2025-02-18]

Abbreviations

AI: artificial intelligence
DCPE: digital care pathway for epilepsy
EEG: electroencephalography
GDPR: General Data Protection Regulation
ICU: intensive care unit
SRA: system requirement analysis

Edited by A Kushniruk; submitted 15.04.24; peer-reviewed by F Xu, O El-Gayar; comments to author 14.10.24; revised version received 07.12.24; accepted 31.12.24; published 30.05.25.

Please cite as:
Keikhosrokiani P, Annunen J, Komulainen-Ebrahim J, Kortelainen J, Kallio M, Vieira P, Isomursu M, Uusimaa J
Requirement Analysis for Data-Driven Electroencephalography Seizure Monitoring Software to Enhance Quality and Decision Making in Digital Care Pathways for Epilepsy: A Feasibility Study from the Perspectives of Health Care Professionals
JMIR Hum Factors 2025;12:e59558
URL: https://humanfactors.jmir.org/2025/1/e59558
doi:10.2196/59558
PMID:

©Pantea Keikhosrokiani, Johanna Annunen, Jonna Komulainen-Ebrahim, Jukka Kortelainen, Mika Kallio, Päivi Vieira, Minna Isomursu, Johanna Uusimaa. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 30.05.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License

XSL•FO
RenderX

(https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Integrating a Large Language Model Into a Socially Assistive Robot in a Hospital Geriatric Unit: Two-Wave Comparative Study on Performance, Engagement, and User Perceptions

Lauriane Blavette[1,2,3], MSc; Sébastien Dacunha[1,2,3], MSc; Xavier Alameda-Pineda[4], PhD; Jeanne Cattoni[1,2,3], MSc; Anne-Sophie Rigaud[1,2,3], Prof Dr Med; Maribel Pino[1,2,3], PhD

[1]Institut National de la Santé et de la Recherche Médicale, Optimisation Thérapeutique en Pharmacologie OTEN Unité Mixte de Recherche-Santé 1144, Université Paris Cité, Paris, France

[2]Service Gériatrie 1&2, Centre Mémoire de Ressources et Recherches Île-de-France-Broca, Assistance Publique – Hôpitaux de Paris, Paris, Île-de-France, France

[3]Broca Living Lab, Hôpital Broca, Paris, France

[4]Centre Institut National de Recherche en Informatique et en Automatique de l'Université Grenoble Alpes, Grenoble, Auvergne-Rhône-Alpes, France

Corresponding Author:
Lauriane Blavette, MSc
Institut National de la Santé et de la Recherche Médicale, Optimisation Thérapeutique en Pharmacologie OTEN Unité Mixte de Recherche-Santé 1144
Université Paris Cité
4, Avenue de l'Observatoire
Paris, 75006
France
Phone: 33 660125406
Email: Lauriane.blavette@aphp.fr

Abstract

Background: Addressing the complex medical and psychosocial needs of older adults is increasingly difficult in resource-limited care settings. In this context, socially assistive robots (SARs) provide support and practical functions such as orientation and information delivery. Integrating large language models (LLMs) into SAR dialogue systems offers opportunities to improve interaction fluency and adaptability. Yet, in real-world use, acceptability also depends on minimizing both technical and conversational errors, ensuring successful user interactions, and adapting to individual user characteristics.

Objective: This study aimed to evaluate the impact of integrating an LLM into a SAR dialogue system in a hospital geriatric unit by (1) comparing system performance and interaction success across 2 experimental waves, (2) examining the links between robot errors, interaction success, and multidimensional user engagement, and (3) exploring how user characteristics influence performance and perceptions of acceptability and usability.

Methods: Over an 8-month period, 28 older adults (≥60 years of age) attending a geriatric day care hospital (Paris, France) participated in a single-session evaluation of a SAR. Interactions took place in the day care hospital and were video-recorded across 2 waves: wave 1 (basic dialogue system) and wave 2 (LLM-based system). From the recordings, system performance (error types and interaction success) and user engagement (verbal, physical, and emotional dimensions) were coded. Acceptability and usability were measured using the Acceptability E-Scale and the System Usability Scale. Sociodemographic data were collected, and quantitative results were supplemented with a thematic analysis of qualitative observations.

Results: Following LLM integration, error-free interactions increased from 27.8% (10/36) to 70.2% (66/94; P<.001), comprehension failures decreased from 47.2% (17/36) to 17% (16/94; P<.001), and interaction success rose from 25% (9/36) to 74.5% (70/94; P<.001). Acceptability (Acceptability E-Scale: 12.8 vs 20.8; P=.003) and usability (System Usability Scale: 40.0 vs 60.4; P=.04) were significantly higher in wave 2. Engagement scores did not differ significantly between waves, though emotional engagement correlated positively with interaction success (r=0.28; P=.008), and age was negatively associated with both physical engagement (r=–0.30; P<.001) and acceptability (r=–0.20; P=.03).

Conclusions: Behavioral engagement with a SAR in geriatric care is shaped by both system performance and individual user characteristics. Improvements in dialogue quality observed in wave 2, coinciding with the integration of the LLM, were associated with higher interaction success and enhanced user experience. Nevertheless, other contextual or group-related factors may also

XSL•FO
RenderX

have contributed to this outcome. These findings highlight the importance of combining multimodal behavioral analysis with self-reported measures to inform the iterative, user-centered design of socially responsive robots in clinical contexts.

(JMIR Hum Factors 2025;12:e81936) doi:10.2196/81936

KEYWORDS

human-robot interaction; socially assistive robot; older adults; behavioral engagement; multimodal analysis; geriatric care; large language models

Introduction

The rapid aging of the population worldwide, combined with a shortage of health care professionals, is placing growing pressure on care systems, particularly in geriatric hospitals and long-term care settings [1]. This strain is amplified when caring for older adults with neurocognitive disorders, such as Alzheimer disease and related dementias, whose needs extend beyond medical supervision to include sustained social, emotional, and cognitive support. However, many care institutions, including hospitals, long-term care facilities, and other residential settings, often lack the resources and structures needed to address these complex needs, resulting in care experiences characterized by unmet psychosocial demands, ineffective communication strategies, and increased vulnerability to distress and disorientation [2-4].

Socially assistive robots (SARs) have emerged as a promising technological complement to human-delivered care in such contexts. SARs are designed primarily for social rather than physical interaction, using speech, gesture, and expressive behaviors to engage users and deliver cognitive, emotional, or informational support [5,6]. Building on their social capabilities, SARs have shown potential in enhancing emotional well-being, reducing anxiety and apathy, and promoting social interaction among older adults, particularly in institutional settings, including but not limited to those living with dementia [7-9]. Beyond therapeutic applications, SARs are increasingly explored as tools to assist health care professionals in nonclinical tasks, such as welcoming patients, providing information, or facilitating recreational activities [10-12].

Institutional applications often encompass reception, visitor guidance, and wayfinding. For instance, studies have shown that robots acting as medical receptionists or lobby greeters can facilitate orientation, manage patient flow, and enhance the perceived quality of service, particularly in high-traffic environments [13,14]. A systematic review by González-González et al [15] further suggests that SARs in hospital contexts often adopt hybrid roles, combining informational support with emotional engagement, especially in waiting rooms or ambulatory care settings. Such use cases broaden the vision of SARs beyond therapeutic agents to include communication and orientation facilitators [16], aligning with broader efforts to improve patient experience where staff availability is limited.

However, despite advances in robotics, including improved sensor technologies, natural language processing, and adaptive user interfaces, the integration of SARs into real-world care environments remains limited, with technical limitations representing one of the most significant barriers [17]. Health care providers have reported frustration with issues such as complex operational steps, short battery life, slow system responses, and rigid dialogue systems [18,19]. These limitations may discourage routine use, particularly in dynamic care settings where efficiency is essential. In addition, persistent challenges in personalization and speech recognition further limit deployment, as current systems often fail to accommodate the wide range of capabilities, preferences, and interaction styles found among geriatric care users [20,21].

From the end user's perspective, errors in human-robot interaction (HRI), including comprehension failures, incoherent or incomplete responses, and misaligned task execution, can erode trust in the robot and reduce engagement [22]. For older adults, even minor conversational breakdowns (eg, misrecognitions or incorrect answers) may trigger confusion, frustration, and diminished trust, particularly when the task is perceived as important [23]. As shown by Kim [24], older adults can also struggle to adapt their speech to the rigid input formats required by conversational systems, sometimes realizing only after delivering a lengthy command that the system had failed to process their request. Such breakdowns not only cause irritation but also expose the gap between the intended "natural" interaction and the constrained, one-directional exchanges these systems often afford. Similar patterns have been observed elsewhere: Khosla et al [25] in a 3-month deployment of a SAR with 5 older adults reported that negative emotions (eg, anger, sadness, and anxiety) occasionally emerged during interactions with the robot, most often when it failed to respond as expected.

Recent advances in large language models (LLMs) offer the potential to mitigate some of these conversational limitations by improving fluency, contextual relevance, and adaptability in robot dialogue systems [26-28]. While early results from other domains (eg, customer service, museum guidance, and educational robotics) are promising, there is limited empirical evidence on how integrating LLMs into SARs affects real-time interactions, error rates, and user experience in real-world geriatric care contexts.

Another critical factor in SAR evaluation is user engagement, which is inherently multidimensional, encompassing verbal behaviors (eg, speech production and turn-taking), physical involvement (eg, gestures and posture), and emotional signals (eg, facial expressions or affective cues) [29]. In HRI, engagement not only serves as a proxy for interaction quality but also predicts subsequent outcomes, such as willingness to re-engage with the robot and the perceived value of the system [30]. Research in care environments, including work with older adults by Hebesberger et al [19], shows that sustained engagement is essential for acceptance and that both technical

XSL•FO
RenderX

reliability and interaction fluency shape the depth and duration of participation. Despite its theoretical and practical importance, little is known about how these engagement dimensions are influenced by robot performance, error frequency, or task success in geriatric care contexts, representing a notable gap in the literature.

Finally, user characteristics, including age, socioeducational background, and cognitive functioning, are likely to influence both observable HRI behaviors and subjective evaluations of SARs [31-34]. For instance, older age may be associated with reduced physical expressiveness, while higher cognitive functioning could support more complex conversational exchanges. Understanding these relationships is essential for informing adaptive robot design and deployment strategies that are inclusive and responsive to diverse user needs.

Given these gaps, this study investigated the deployment of a SAR in a hospital geriatric unit, focusing on three main objectives (1) to assess changes in system performance and interaction success following the integration of an LLM into the robot's dialogue system, comparing 2 experimental waves; (2) to examine the relationships between robot errors, interaction success, and user engagement, considering verbal, physical, and emotional dimensions; and (3) to explore how user characteristics relate to system performance and subjective evaluations of acceptability and usability.

By combining quantitative measures of performance, engagement, and user ratings with qualitative analysis of participant perceptions, this study provides a real-world grounded perspective on the opportunities and challenges of SAR deployment in aging care contexts.

Methods

Participants

Overview

The study involved older adults attending consultations at the geriatric day care hospital (DCH) of Broca Hospital (Assistance Publique—Hôpitaux de Paris), France. Inclusion criteria were (1) being 60 years and older of age; (2) participating in a scheduled DCH consultation; (3) having a Mini-Mental State Examination (MMSE) [35] score above 10, indicating the absence of severe cognitive impairment; (4) having no current symptoms of altered reality (eg, delusions or hallucinations); and (5) having fluent comprehension and expression in French. No exclusion criteria were applied based on sex, socioeconomic backgrounds, or ethnicity.

Recruitment

Recruitment was carried out using the DCH database, and participants were prescreened prior to enrollment, contacted over the phone, and invited to participate in the study the day of their next consultation. An information letter was sent by post, and informed consent was collected onsite.

A total of 41 older adults were recruited for the study. However, this analysis focuses on a subset of 28 participants who met the following criteria: (1) provided consent to be filmed; (2) completed a full interaction session with the robot; and (3) generated usable video data, defined by adequate visibility and recording quality. The remaining participants were excluded due to 1 or more of the following reasons: refusal to engage with the robot, technical malfunctions during the session, or withdrawal of consent for video recording.

Setting

The study was conducted between May 2023 and December 2023 in the DCH of a geriatric hospital in Paris (France). The DCH provides specialized outpatient care for older adults with physical or cognitive impairments, offering a wide range of consultations, including neurology, oncology, cardiology, psychiatry, and memory assessments.

The data collection occurred in 2 waves during this period, with distinct participant samples recruited for each phase. All interactions were carried out in a quiet, dedicated room located near the DCH waiting area. This space, typically used for rest and informal activities, was chosen to provide a calm and comfortable environment for testing, while maintaining proximity to the clinical setting.

Study Design

This cross-sectional observational study was conducted as part of a broader research protocol evaluating the integration of a SAR in geriatric care [36,37]. This analysis focuses on real-world HRI, combining behavioral observations with user-reported data. Each participant engaged in a single, nonscripted interaction session with the robot, followed by questionnaires. Data were collected across 2 experimental waves. The first wave involved a baseline version of the robot without LLM integration, while the second wave used an updated LLM-enabled version. Each wave included a different group of older adults.

Material

ARI Robot

The ARI robot developed by Pal Robotics (Spain) is a 1.65-m (5 ft 5 in) wheeled humanoid platform equipped with a touchscreen, cameras, microphones, animated eyes, and articulated arms (Figure 1). In this study, ARI was used as a socially assistive agent in a geriatric hospital setting.

XSL•FO
RenderX

Figure 1. Front and side views of the ARI robot (photo credit: PAL Robotics).



For this study, ARI was programmed with interaction modules developed as part of the European H2020 SPRING (Socially Pertinent Robots in Gerontological Healthcare) project, aimed at enabling SAR in real-world clinical environments [36]. Two different configurations of ARI's conversational system were tested during this study, as detailed in the System Evolution section. As ARI had not yet been commercially deployed in care institutions at the time of the study, its use in this context was part of a controlled, exploratory evaluation conducted within the hospital environment.

System Evolution

To assess the impact of iterative improvements to ARI's interaction capabilities, we conducted 2 experimental waves, each following a major system update informed by participant feedback.

- Wave 1 (May to July 2023) deployed a first version of ARI featuring a modular dialogue system based on traditional rule-based intent handling, retrieval-based responses, and basic open-domain generation. This initial architecture was supported by core diagnostic tools, on-screen speech transcription, and perceptual modules (eg, person tracking and facial recognition).
- Wave 2 (September to December 2023) introduced a redesigned conversational module centered on an LLM architecture. Specifically, the system was upgraded to integrate Vicuna-13B-v1.5, a 13-billion parameter LLM optimized for offline use. A custom prompt in French was developed to align with hospital-based use cases, ensuring context-sensitive, coherent interactions. Other system improvements included a refined transcription interface and more efficient vision module processing. All data were securely stored on servers dedicated exclusively to the project. This setup ensured robust data control, as the language model was hosted locally, giving the project partners full control over processing, storage, and use.

A detailed technical overview of these updates is available in SPRING Deliverable D1.6 [38].

Assessment Instruments

Demographic Data Collection

Basic demographic information was collected using a standardized paper-based form prior to the start of the interaction session. Participants were asked to report their age, sex, and level of education (years of formal education). The level of cognitive function (MMSE) was completed beforehand by the researcher.

System Performance

To assess system performance, all interactions were recorded and annotated based on robot behavior. Each interaction was defined as a conversational unit, beginning when the participant

XSL•FO
RenderX

initiated a verbal input and ending when the robot responded or failed to do so. For example, if a participant asked, “Where is the restroom?” and the robot replied, “To your right as you exit this room,” the exchange was considered complete, regardless of its duration. Participants could then initiate a new exchange, continue the dialogue, or end the session. This approach produced interactions of widely varying length, from brief question-answer sequences to longer, multiturn conversations, depending on participant intent and conversational flow.

Robot behaviors were annotated using a predefined classification scheme consisting of four categories:

1. Comprehension failure: The robot fails to interpret or meaningfully respond to the user’s input, typically resulting in silence or a fallback message (eg, “I don’t understand”).
2. Inappropriate verbal response: The robot produces a reply that is unrelated, incoherent, or socially inappropriate in context (eg, to the question “What time is my appointment?” the robot replies “Of course. The hospital is open every day from 9 a.m. to 5.30 p.m.”).
3. Incomplete utterances: The robot produces an unfinished or abruptly cutoff response, resulting in a message that lacks necessary information (eg, when asked “Can you tell me where the toilets are?” the robot replies “The toilets are on your left as you come out of ...”).
4. Technical error: System-level malfunctions such as speech synthesis failure, audio dropout, or interface freezing.

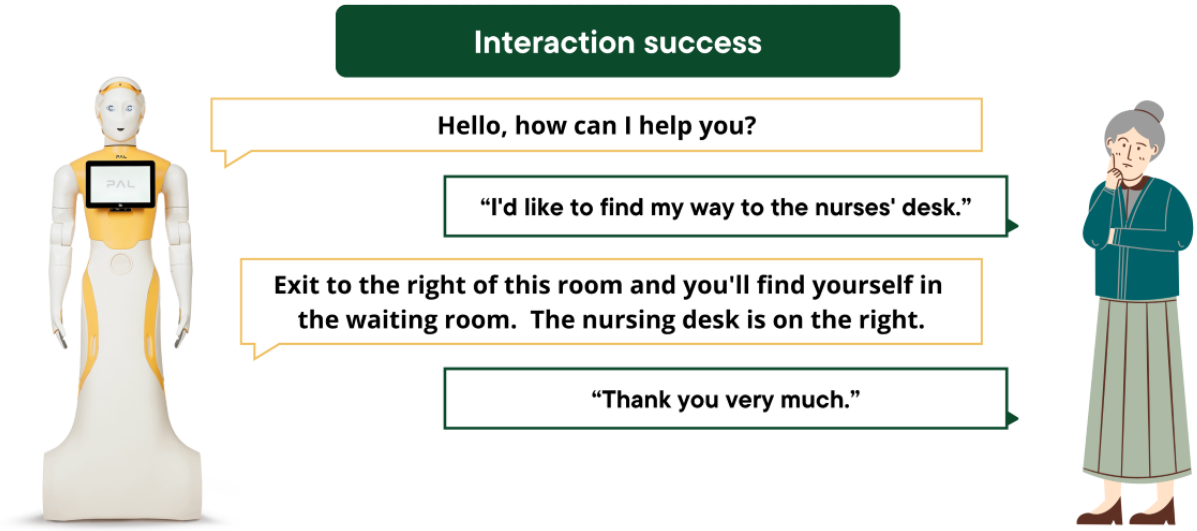
The first 3 robot behaviors, comprehension failures, inappropriate verbal responses, and incomplete utterances, were classified as verbal-related errors. This categorization allowed for exploratory correlation analyses between robot behavior and user engagement, enabling a distinction between technical malfunctions and conversational breakdowns.

Multiple error types could be assigned to a single interaction when applicable (eg, a comprehension failure accompanied by a technical issue). This allowed for the analysis of both error frequencies and error co-occurrence patterns as well as comparisons across the 2 experimental waves to examine changes in system performance over time.

Interaction Success

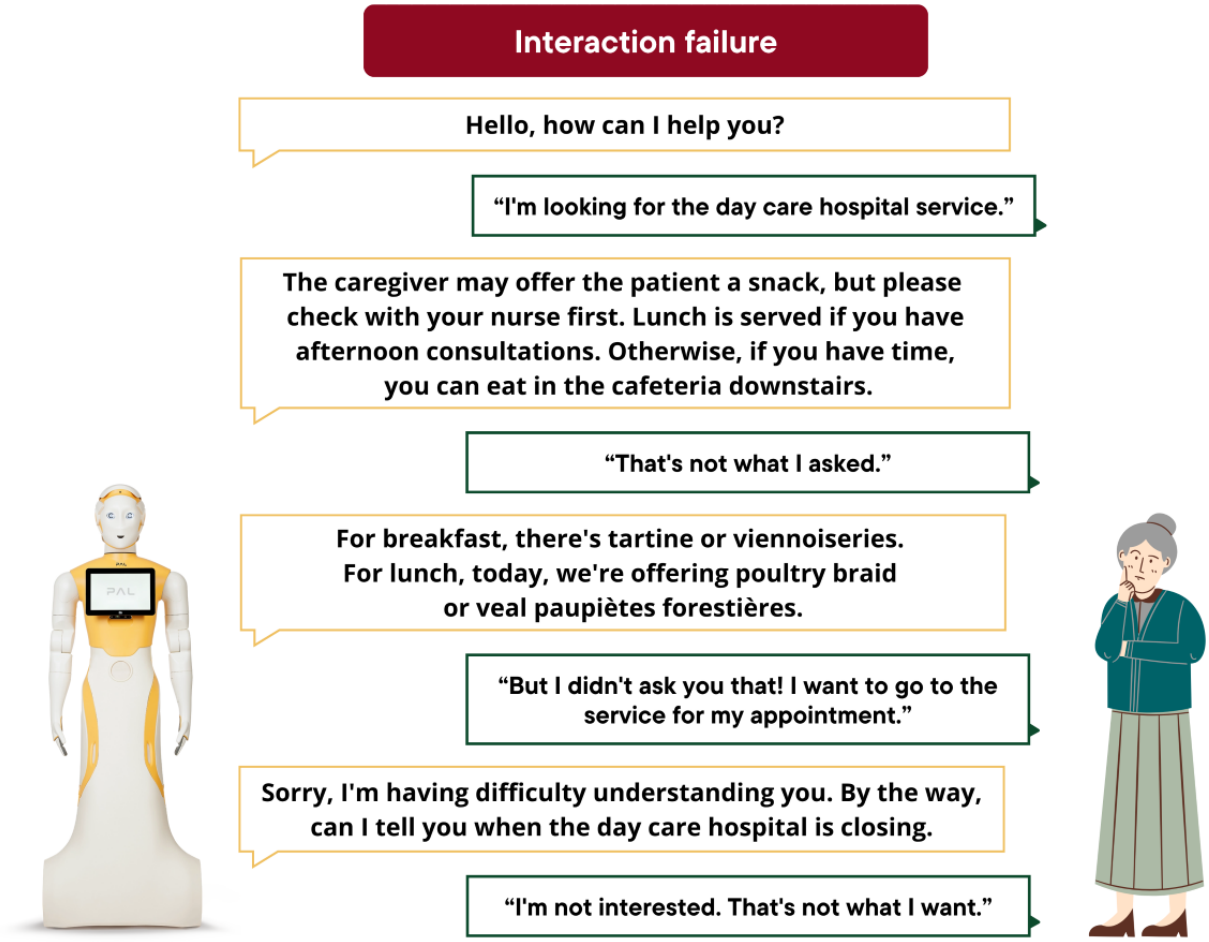
Interaction success was defined as the robot providing a relevant and coherent response that appropriately addressed the user’s request or corresponded to the information provided. Examples of successful and failed interactions are illustrated in Figures 2 and 3.

Figure 2. Example of a successful interaction between the ARI robot and a participant.



XSL•FO
RenderX

Figure 3. Example of a failed interaction between the ARI robot and a participant.



User Engagement Coding Framework

User engagement was assessed using a structured behavioral coding scheme (Table 1), developed to capture 3 dimensions: verbal, physical, and emotional engagement. The verbal dimension reflected speech production and conversational responsiveness; the physical dimension encompassed posture, gesture use, body orientation, and other forms of physical involvement; and the emotional dimension captured visible emotional expressivity. Each dimension was rated on a 5-point scale, ranging from 1=minimal engagement or rejection to 5=high involvement and expressivity.

XSL•FO
RenderX

Table 1. User engagement dimensions and rating scale.

Dimension and score	Description
Verbal engagement	
1	Persistent silence or clearly disengaged verbal behavior (eg, hostile or rejecting remarks).
2	Minimal verbal output (eg, 1-word replies such as “yes” or “no”), delivered in a flat tone, with no spontaneous elaboration or follow-up.
3	Brief, functional responses that are appropriate and clearly articulated but emotionally neutral and lacking initiative.
4	Active participation through contextually appropriate answers and spontaneous follow-up contributions.
5	Fluent, sustained conversation featuring regular verbal initiative, coconstructed dialogue, and spontaneous, personally meaningful comments or anecdotes.
Physical engagement	
1	Defensive or avoidant posture (eg, crossed limbs and leaning away), evasive gaze, and rejecting or dismissive gestures.
2	Passive physical presence, with infrequent or minimal gestures and body oriented away from the robot.
3	Upright, neutral posture with limited physical expressivity; gaze may be directed toward the robot but lacks clear engagement; few gestures.
4	Open body orientation toward the robot, accompanied by expressive gestures and visibly responsive physical behavior (eg, nodding and leaning in).
5	Strong physical involvement, including spontaneous gestures, mimicry of the robot’s actions, or voluntary physical contact (eg, touching the robot).
Emotional engagement	
1	Clear negative emotional expression (eg, anger, frustration, and disgust), often accompanied by signs of interaction breakdown or withdrawal.
2	Very low emotional reactivity, flat facial affect, and visible signs of fatigue, boredom, or disinterest.
3	Neutral but attentive facial expression, stable gaze toward the robot, with limited or absent emotional expressivity.
4	Moderate social-affective signals, such as smiling, nodding, or short expressive reactions, indicating active but contained emotional engagement.
5	Strong affective display, including laughter, animated facial or physical expressions, and moments of emotional synchrony with the robot.

The coding grid was designed by the research team, grounded in prior work on social signal processing [39], and iteratively refined through preliminary testing to ensure clarity, consistency, and applicability in naturalistic HRI settings. It incorporated both affective and interactional markers, such as speech spontaneity, posture alignment, gesture expressiveness, and emotional display.

All interactions were independently coded by 2 trained researchers (LB and MP) using a detailed coding manual to reduce subjectivity inherent in qualitative behavioral analysis [40]. In cases of disagreement, coders discussed the item to reach consensus; if consensus could not be achieved, a third researcher (JC) acted as an arbitrator to determine the final score. This multicoder approach aligns with best practices in HRI research, which emphasizes the use of multiple raters to mitigate individual bias and enhance intersubjective reliability [39,41].

Emotional engagement scores were not assigned for 28 of 130 interactions due to limited facial visibility caused by surgical masks, which impeded the reliable interpretation of affective expressions. These interactions involved 5 participants, 3 who remained masked throughout and 2 who removed their masks partway. While excluded from emotional engagement analysis, these sessions were retained for all other measures.

In addition to the quantitative ratings, coders were encouraged to document any ambiguous, unexpected, or contextually meaningful behaviors using an open comment field. These qualitative annotations served to refine and contextualize the interpretation of engagement patterns observed in the coded data.

Assessment Scales

Two standardized self-report questionnaires were administered immediately after the interaction to assess participants’ perceptions of the robot’s acceptability and usability. The Acceptability E-Scale (AES), adapted from Heerink et al [42] and translated into French by Micoulaud-Franchi et al [43], measures perceived acceptability across 6 dimensions: trust, perceived usefulness, enjoyment, sociability, ease of use, and intention to use. The scale includes 6 items, each rated on a 5-point Likert scale, yielding a total score ranging from 6 to 30. AES scores above 25.81 of 30 are considered to reflect high perceived acceptability, while lower scores indicate limited acceptability. The adapted version of the scale used in this study is provided in Multimedia Appendix 1.

XSL•FO
RenderX

The System Usability Scale (SUS) [44] is a widely used, validated 10-item questionnaire designed to assess perceived usability. Items are rated on a 5-point Likert scale, and the final score is calculated according to standard scoring procedures, yielding a composite usability score between 0 and 100. According to standard interpretive guidelines, SUS scores below 50.9 are considered poor, scores between 51 and 71.4 are rated as OK to good, and scores above 71.4 reflect excellent perceived usability [45]. The adapted version of the scale used in this study is provided in Multimedia Appendix 1.

Assessment Procedure

Each session (~45 minutes) followed a standardized sequence consisting of a short reminder of the objectives of the study, informed consent, interaction with the robot, and postinteraction assessments. Participants were first introduced to the robot's general capabilities (eg, hospital orientation, entertainment, and appointment assistance) and informed that they could ask the robot questions freely related to these functions. No detailed interaction instructions or scripted prompts were provided, in order to preserve a naturalistic interaction context.

A fixed-position camera recorded the entire session for behavioral coding. A researcher (LB) was present throughout the entire session to provide assistance if needed, while minimizing interference.

Following the interaction, participants completed 2 standardized self-report questionnaires, the AES and the SUS, with support from the research team when necessary.

Data Analysis

Quantitative analyses were conducted using the jamovi software (version 2.4.11; The Jamovi Project) and focused on 4 dimensions: sociodemographic data, system performance, user engagement, and acceptability and usability.

For sociodemographic characteristics (age, sex, socioeducational level, and MMSE scores), descriptive statistics (mean, SD, minimum, and maximum) were first computed. Comparisons across experimental waves were then conducted using the Mann-Whitney *U* test, due to violations of normality and homogeneity of variance assumptions, as assessed by the Shapiro-Wilk and Levene tests.

For the analysis of HRIs, the unit of analysis was the individual interaction. Interaction duration was calculated for each HRI session based on video recordings. Mean durations were computed separately for each experimental wave, and the total cumulative duration of all recorded interactions was calculated to quantify the overall volume of video data analyzed. The definition of an interaction used in this study is provided in the System Performance section.

A descriptive analysis was first conducted for each wave, covering these elements: robot behavior categorized by 4 error types (comprehension failures, inappropriate verbal responses, incomplete responses, and technical errors) and the proportion of interactions with and without errors.

Subsequently, between-wave comparisons were performed. Robot behavior was compared across waves based on the average error rate per interaction. Analyses were first conducted considering all error types, followed by a focused analysis on the 3 verbal-related errors: comprehension failures, inappropriate verbal responses, and incomplete responses. These comparisons were conducted using Mann-Whitney *U* tests, due to nonnormal distributions and violations of homogeneity of variance in the error frequency data.

User engagement was then analyzed using multimodal behavioral coding across 3 dimensions: verbal, physical, and emotional engagement. A descriptive analysis was first conducted for each dimension. Subsequently, engagement scores were compared between experimental waves using Kruskal-Wallis tests, due to violations of normality and homogeneity of variance assumptions.

Interaction success was operationalized as the completion of a communicative goal, specifically, when the user successfully obtained the information requested from the robot. Success rates were compared between experimental waves using a chi-square test of independence to assess whether system upgrades influenced the likelihood of a successful exchange.

The dimensions of acceptability and usability, assessed using the AES and SUS questionnaires, respectively, were first examined through descriptive analysis. Subsequently, scores were compared between experimental waves using independent samples 2-tailed *t* tests, following confirmation of normality (Shapiro-Wilk test) and homogeneity of variance (Levene test).

To complement group-level comparisons, additional analyses were conducted to explore relationships between the dimensions of engagement, system performance, user experience, and participant characteristics. Pearson correlation analyses were used to examine associations between user engagement scores and interaction outcomes (ie, successful vs unsuccessful interactions) as well as between engagement scores and perceived acceptability (AES) and usability (SUS). Correlation analyses were also performed on the full sample to investigate relationships between sociodemographic characteristics (age, socioeducational level, and MMSE), system performance indicators (total and conversational error counts), user engagement dimensions (verbal, physical, and emotional), and subjective evaluations of the robot (AES and SUS). These analyses aimed to determine whether individual differences were associated with user experience, robot performance, or behavioral engagement during interactions.

A short qualitative analysis was conducted using thematic classification [46] of spontaneous, marginal comments made by participants during their interactions with the robot. These comments were not obtained through interviews or explicit discussions about the HRI but occurred naturally during the exchanges. Comments were categorized under emerging themes to complement the quantitative findings and provide additional insight into user engagement and interaction dynamics. Coding was performed independently by 2 observers (LB and MP), who then compared and discussed their categorizations to reach consensus on the final themes. Because of the small amount of data, no formal interrater reliability statistic (eg, Cohen κ) was calculated.

XSL•FO
RenderX

Ethical Considerations

The study was approved by the French National Ethics Committee ("Comité de Protection des Personnes, CPP Ouest II [21.02.03.45524 2021/20], Maison de la Recherche Clinique – CHU Angers [2020-A02643-36]"; and complied with the General Data Protection Regulation. Data processing was registered with the Data Protection Officer (reference 20210114153645) in the Assistance Publique—Hôpitaux de Paris registry. The study did not involve randomization or a clinical intervention. Informed consent was obtained from all participants, and they were informed that they could withdraw from the study at any time. The original consent included approval for secondary analyses without requiring additional consent. All participant data were anonymized, and no compensation was provided. From an ethical standpoint, the use of SARs raises several issues, particularly regarding confidentiality and transparency. Because these systems collect and process sensitive data (vocal, visual, or interactional), their use and potential reuse must be explicitly stated. Ethical guidelines also demand that users and participants are clearly informed, so they understand how their data are managed. With the integration of LLMs, this obligation further includes clarifying whether experimental results will be used for model training and under which conditions.

Results

Participant Characteristics and Baseline Equivalence

The sample included 28 older adults, with 10 participants in wave 1 and 18 in wave 2. The mean age was 78.2 (SD 6.25; range 67-93) years, and most participants were women (n=20). The average MMSE score was 26.3 of 30 (SD 3.73; range 18-30), with scores ≥26 generally indicating normal cognitive function. The mean socioeducational level was 12.8 (SD 1.94) years, ranging from 9 (completion of lower secondary education or vocational training) to 14 or more years (university-level qualifications such as a bachelor's degree or higher).

Baseline equivalence between groups was assessed using Mann-Whitney *U* tests for MMSE scores and socioeducational level, as both variables were nonnormally distributed. No significant differences were found between wave 1 and wave 2 for MMSE (U=49.5; P=.05; r=0.45) or socioeducational level (U=79.5; P=.57; r=0.12), supporting the comparability of the 2 experimental groups. The inclusion of effect sizes indicates that the observed differences were small to moderate in magnitude, further supporting baseline equivalence despite minor variability in MMSE. Participant characteristics by experimental wave are summarized in Table 2.

Table 2. Sociodemographic characteristics of participants by experimental wave.

Experimental wave	Sex		Age (years), mean (SD)	Education[a] (years), mean (SD)	MMSE[b] score, mean (SD)
	Male, n (%)	Female, n (%)			
Wave 1 (n=10)	3 (20)	7 (80)	78.6 (7.6)	13.1 (1.7)	28.0 (1.9)
Wave 2 (n=18)	6 (33.3)	12 (66.7)	77.9 (5.6)	12.6 (2.1)	25.3 (4.1)
Total (N=28)	9 (32.1)	19 (67.9)	78.2 (6.3)	12.8 (1.9)	26.3 (3.7)

[a]Years of education refer to the total years of formal schooling completed.

[b]MMSE: Mini-Mental State Examination (range 0-30); scores≥26 are generally considered within the normal cognitive range.

Interaction Duration and Dataset Composition

A total of 130 HRIs were analyzed (wave 1=36 HRIs and wave 2=94 HRIs), comprising the full dataset used for multimodal behavioral coding. Interaction durations ranged from 7 seconds to 3 minutes and 14 seconds, with a mean duration of 52 (SD 0.00038) seconds: wave 1=40 (SD 0.00037) seconds and wave 2=58 (SD 0.00037) seconds. In total, the recordings resulted in 1 hour, 53 minutes, and 55 seconds of video data. The experimental setup is shown in Figure 4, which depicts a participant interacting with the SAR during a session.

XSL•FO
RenderX

Figure 4. Experimental context: participant interacting with the socially assistive robot.



Robot Error Patterns and System Performance Across Waves

Overview

Table 3 presents the distribution of robot error types across the 2 experimental waves. Comprehension failures were the most frequent error, occurring in 25.4% (33/130) of all interactions, with a notably higher rate in wave 1 (17/36, 47.2%) compared to wave 2 (16/94, 17%). Incomplete utterances were absent in wave 1 but appeared in 14.9% (14/94) of interactions in wave 2. Inappropriate verbal responses were also more common in wave 1 (13/36, 36.1%) than in wave 2 (14/94, 14.9%), contributing to a combined rate of 20.8% (27/130). Technical errors remained relatively stable between conditions, occurring in 11.1% (4/36) of interactions in wave 1 and 11.7% (11/94) in wave 2. For analysis, each error type was binary-coded per interaction (0=absent and 1=present).

Table 3. Distribution of robot error types by experimental wave.

Robots' behaviors	Wave 1, n (%)	Wave 2, n (%)	Total, n (%)
Comprehension failures	17 (47.2)	16 (17)	33 (25.4)
Incomplete utterances	0 (0)	14 (14.9)	14 (10.8)
Inappropriate verbal responses	13 (36.1)	14 (14.9)	27 (20.8)
Technical error	4 (11.1)	11 (11.7)	15 (11.5)

Robot error types were compared across the 2 experimental waves based on total error counts. Verbal-related errors, including comprehension failures (U=989; P<.001; r=0.42), incomplete utterances (U=1440; P=.015; r=0.15), and inappropriate responses (U=1333; P=.008; r=0.21), differed significantly between wave 1 and wave 2. A higher number of verbal-related errors were recorded in wave 1, reflecting reduced conversational performance in this condition and suggesting improved system behavior following the updates implemented in wave 2.

An additional analysis was conducted at the level of individual interactions to compare error rates between experimental waves. Results showed higher error rates per interaction in wave 1 than in wave 2: comprehension failures occurred on average 2.03 times per interaction in wave 1 compared to 0.61 in wave 2; inappropriate responses occurred at a rate of 0.36 in wave 1

XSL•FO
RenderX

versus 0.15 in wave 2; and incomplete utterances, which were absent in wave 1, appeared with a mean frequency of 0.16 per interaction in wave 2. In contrast, technical error rates did not differ significantly between waves (U=1682; P=.93; r=0.01), indicating stable hardware performance across conditions.

Finally, the proportion of interactions without any system errors increased significantly in wave 2 following the introduction of the LLM (U=995; P<.001; r=0.41), indicating improved system stability and interaction reliability after the update. Only 27.8% (10/36) of interactions in wave 1 were error-free, compared to 70.2% (28/94) in wave 2. This substantial increase in error-free interactions suggests that the integration of the LLM contributed positively to the overall robustness and consistency of the system's performance.

User Engagement

Participant engagement was evaluated using multimodal behavioral coding across 3 dimensions: verbal, physical, and emotional. Each dimension was scored on a 5-point scale, with higher scores indicating greater levels of engagement.

The mean verbal engagement score was 3.64 (SD 0.64) in wave 1 and 3.47 (SD 0.77) in wave 2. A Kruskal-Wallis test revealed no significant difference in verbal engagement between the 2 waves (χ^2_1=2.5; P=.11).

For physical engagement, scores remained stable across waves, with a mean of 3.03 (SD 0.77) in wave 1 and 2.99 (SD 0.73) in wave 2. No significant difference was observed (χ^2_1=0.06; P=.80).

Emotional engagement showed a slight increase, from a mean score of 2.92 (SD 1.02) in wave 1 to 3.33 (SD 0.86) in wave 2. No significant difference was observed (χ^2_1=3.4; P=.06). Descriptive statistics by wave are presented in Table 4.

Table 4. Comparison of acceptability and usability scores, user engagement metrics, and interaction success across waves.

Dimension	Wave 1	Wave 2	Difference, *P* value
AES[a] score out of 30, mean (SD)	12.8 (4.58)	20.8 (6.52)	<.001
SUS[b] score out of 100, mean (SD)	40.0 (24.04)	60.4 (23.11)	<.001
Duration of interaction (seconds), mean (SD)	40 (32)	58 (32)	<.001
Verbal engagement mean score out of 5 (SD)	3.64 (0.64)	3.47 (0.77)	.24
Physical engagement mean score out of 5 (SD)	3.03 (0.77)	2.99 (0.73)	.79
Emotional engagement mean score out of 5 (SD)	2.92 (1.02)	3.33 (0.86)	.05
Interaction success, n/N (%)	9/36 (25)	70/94 (74.5)	<.001

[a]AES: Acceptability E-Scale.

[b]SUS: System Usability Scale.

Interaction Success and Perceived Acceptability and Usability

Interaction success, defined as the completion of a communicative goal without breakdown, improved significantly across waves: 25% (9/36) in wave 1 versus 74.5% (70/94) in wave 2 (χ^2_1=26.7; P<.001).

Participants who experienced a successful interaction reported significantly higher ratings of both acceptability and usability. For the AES, mean scores were 21.6 of 30 (SD 5.84) in successful interactions, compared to 17.4 of 30 (SD 7.15) in failed ones (U=1261.5; P<.001; r=0.35). SUS scores followed the same pattern, with a mean of 63.3 of 100 (SD 22.5) for successful exchanges versus 51.1 of 100 (SD 26.5) for unsuccessful ones (U=1456.5; P=.01; r=0.25).

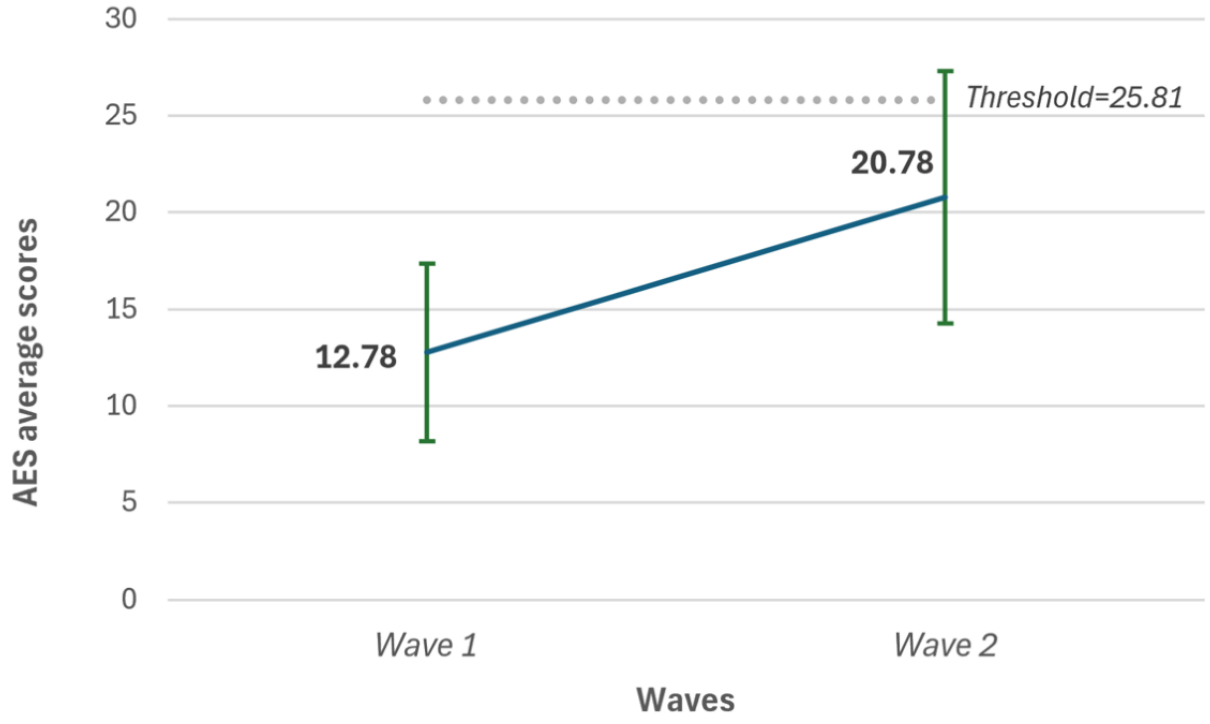
Importantly, no significant differences were found between participants who experienced successful versus unsuccessful interactions in terms of age (U=1898.5; P=.66; r=0.05), years of education (U=1915.5; P=.58; r=0.06), or MMSE scores (U=1921; P=.65; r=0.05), suggesting that interaction success was not dependent on participants' demographic or cognitive profiles.

Acceptability Assessment

In wave 1, acceptability scores ranged from 9 to 23, with a mean of 12.8 (SD 4.58). In wave 2, scores ranged from 7 to 30, with a mean of 20.8 (SD 6.52). The AES ranges from 6 to 30, with higher scores indicating greater acceptability; a commonly used threshold of 25.81 denotes high acceptability. An independent sample 2-tailed t test revealed a significant difference between waves (t_{25}=–3.28; P=.003), with higher acceptability observed in wave 2 compared to wave 1. Figure 5 illustrates the distribution of AES scores across the 2 experimental waves, showing the increase in mean acceptability from wave 1 to wave 2.

XSL•FO
RenderX

Figure 5. AES scores across the 2 experimental waves. AES: Acceptability E-Scale.

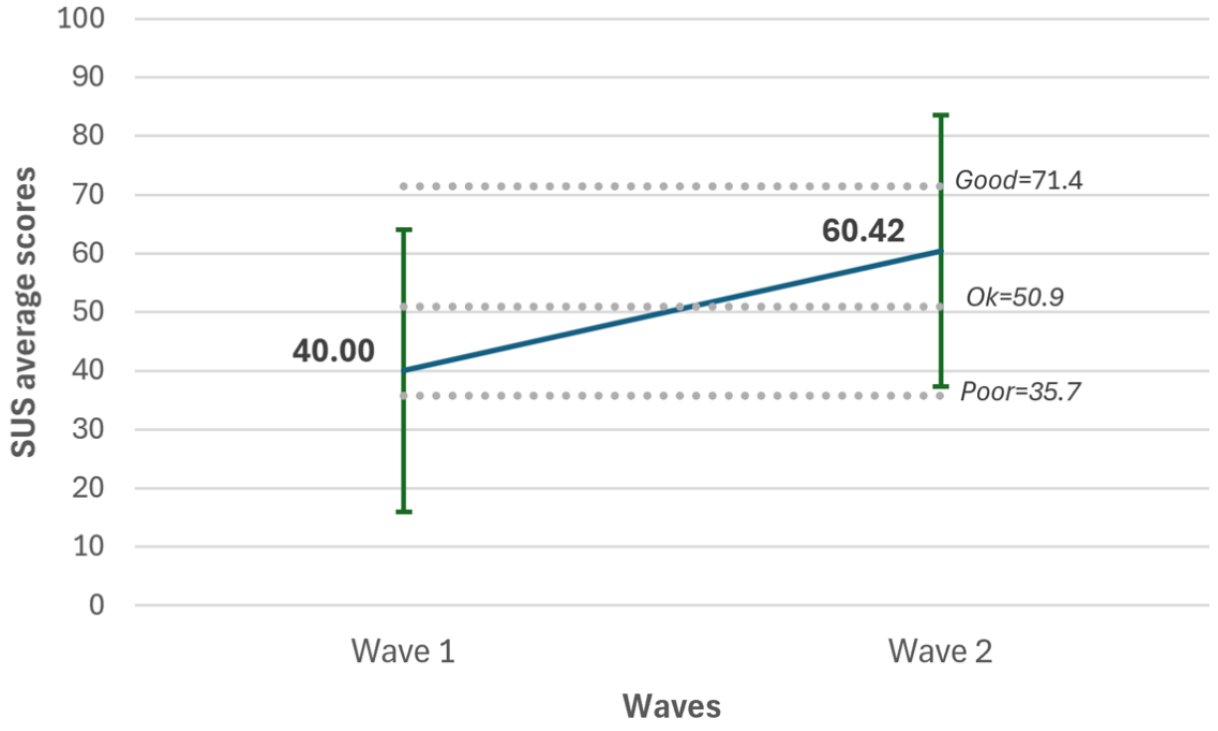


Usability Assessment

Overview

In wave 1, SUS scores ranged from 12.5 to 67.5, with a mean of 40.00 (SD 24.04). In wave 2, scores ranged from 2.5 to 92.5, with a mean of 60.42 (SD 23.11). The SUS ranges from 0 to 100, with 50.9 commonly cited as the minimum threshold for acceptable usability; scores below this indicate insufficient usability. An independent sample 2-tailed *t* test revealed a significant difference between waves (t_{25}=–2.14; *P*=.04), indicating greater perceived usability in wave 2 compared to wave 1. Figure 6 shows the increase in mean SUS scores between the 2 experimental waves, indicating improved perceived usability in wave 2.

Figure 6. SUS scores across the 2 experimental waves. SUS: System Usability Scale.



XSL•FO
RenderX

Associations Among Engagement, Outcomes, and User Characteristics

A Pearson correlation analysis was conducted to explore the relationships between user engagement (verbal, physical, and emotional), interaction outcomes, system evaluations, and participant characteristics (see Multimedia Appendix 2 for the full correlation matrix). Emotional and physical engagement were strongly positively correlated (r=0.63; P<.001), indicating that increased emotional expressiveness was closely associated with greater physical involvement. Emotional engagement was also positively associated with interaction success (r=0.28; P=.008), suggesting that successful interactions tended to elicit higher emotional responsiveness.

A moderate positive correlation was observed between physical and verbal engagement (r=0.23; P=.009), indicating that participants who were more physically expressive also tended to engage in more sustained or contextually responsive verbal behavior. Conversely, verbal engagement was negatively correlated with perceived acceptability (AES; r=–0.27; P=.002), suggesting that more frequent or proactive verbal contributions were associated with lower acceptability ratings.

Among participant characteristics, age showed a moderate negative correlation with physical engagement (r=–0.30; P<.001), indicating that older participants tended to be less physically expressive during interactions. Age was also negatively associated with acceptability ratings (r=–0.20; P=.03), suggesting slightly reduced system acceptability among older users. No significant correlations were found between cognitive functioning, as measured by MMSE scores, or socioeducational level, and the other engagement or evaluation measures.

Qualitative Analysis of Spontaneous Verbalizations

Overview

During the interactions, participants occasionally produced spontaneous comments in reaction to specific events, often triggered by the robot's conversational or technical errors. These remarks provide insight into how older adults evaluated the exchange, revealing patterns of frustration, adaptation, humor, and expectation management. Thematic analysis identified 5 recurrent themes (Table 5).

Table 5. Thematic analysis of participants' spontaneous verbalizations.

Theme	Description	Exemplar quotes
Silence, confusion, and humor as coping	Errors or nonresponses from the robot often led participants to react with humor, irony, or playful comments, softening frustration.	• "We're going round in circles here. [laughs]" [Wave 1]. • "Does it [the robot] have blocked ears? [laughs]" [Wave 1). • "Yet [the robot] understood (points to the screen with the transcript) but it [the robot] doesn't answer my question" [Wave 1].
Self-attribution of blame	Communication breakdowns were sometimes attributed by participants to their own speech, hearing, or contextual factors.	• "I can't hear very well, as I have a lot of trouble with my ears" [Wave 1]. • "It [the robot] is having trouble understanding me. Is it the mask?" [Wave 2]. • "Am I speaking loud enough [for the robot to understand]?" [Wave 1].
Expectation-response mismatch and technology comparisons	Participants noted discrepancies between expected and actual robot responses, sometimes comparing them to earlier or limited technologies.	• "Doesn't it understand anything? [the robot did not give an accurate answer]" [Wave 1]. • "We're not going to upset it [laughs]" [Wave 1]. • "It reminds me of when they first introduced voice word processing. We had some absolutely incredible things written down. It didn't make any sense at all" [Wave 2].
Seeking external guidance	Interaction difficulties prompted participants to turn to the experimenter for clarification or support.	• "How can I find out what questions I can ask the robot? [to the experimenter]" [Wave 2]. • "Do I have to talk to it [the robot]?" [Wave 2]. • "So now I can start again? Do I have to restart from zero or should I step out of the field and come back? [to the experimenter]" [Wave 2].
Positive reactions to robot-initiated exchanges	When the robot took initiative, participants expressed surprise or delight, highlighting novelty and engagement.	• "It's impressive [the robot]!" [Wave 2]. • "Oh! It [the robot] talks to me first, just like a person!" [Wave 2]. • "So if I don't say anything, it [the robot] will keep asking me things?" [Wave 2].

Silence, Confusion, and Humor as Coping

When the robot failed to respond or gave irrelevant answers, participants often reacted with a mix of confusion and humor. Some expressed mild frustration, "Well, yes, but it's not answering me" (Wave 1), while others turned to playful remarks, joking that the robot had "blocked ears" or even deserved a slap. Sarcasm and irony also emerged as ways to reframe disappointment, with comments like "[laughs] Wasn't that question planned [in the database]?" (Wave 1) or "We asked it a question it's not very good at" (Wave 1). Participants also relied on the robot's transcript display to make their point: "It

XSL•FO
RenderX

understood (points to the screen) but it doesn't answer my question." These responses show how older adults shifted between frustration and collaborative problem-solving. Some participants therefore chose humor rather than disengagement, demonstrating resilience when interacting with a new technology such as SAR.

In addition, participants recurrently described conversational "loops" when repeated prompts or ARI's lack of uptake brought dialogue to a standstill. The interaction frequently felt circular: the robot kept repeating information and failed to build on prior exchanges. This resulted in participants asking clarifying questions like "Shall I continue?" (Wave 2), signaling a breakdown in mutual understanding. Affect often shifted from amused to impatient, captured in remarks such as "We're going round in circles here [laughs]" (Wave 1) and "Wake up ARI!" (Wave 1), yielding interactional impasses marked by circularity, uncertainty about mutual understanding, and mounting frustration.

Self-Attribution of Blame

Instead of attributing breakdowns to the robot's limitations, participants often questioned their own role in communication difficulties. Comments such as "I can't hear very well, as I have a lot of trouble with my ears" (Wave 1) or "It's having trouble understanding me. Is it the mask?" (Wave 2) show how challenges were internalized, sometimes linked to environmental factors or speech clarity. This self-scrutiny extended to interaction etiquette, "Do I speak fast enough?" (Wave 1), "Am I speaking loud enough?" (Wave 1), and even basic rules, as in "Ah! Because I have to ask a question?" (Wave 2). Such attributions highlight how older adults often personalize technological breakdowns, with uncertainty toward novel systems manifesting as self-questioning rather than solely blaming the robot.

Expectation-Response Mismatch and Technology Comparisons

Participants at times noted a gap between their expectations and the robot's actual capabilities. Remarks such as "It's a catering robot, not a robot for guiding patients" or "Doesn't it [the robot] understand anything?" (Wave 1) revealed disappointment or confusion about its function. Others responded more playfully or with emotional distance, "We're not going to upset it [the robot] [laughs]" and "It [the robot] looks at me with strange eyes" (Wave 1). Some also drew comparisons with earlier technologies, as in "It [the robot] reminds me of when they first introduced voice word processing. We had some absolutely incredible things written down, it didn't make any sense at all" (Wave 2). Such expressions reflect a mix of critique and humor, showing how unmet expectations often evoked both skepticism and familiarity with past technological shortcomings.

Seeking External Guidance

When conversational flow was disrupted or when uncertainties about the interaction emerged, several participants turned to the experimenter for clarification or support. Requests like "Come along please [to the experimenter]" (Wave 2) and "How can I find out what questions I can ask the robot?" (Wave 2) point to participants' desire for guidance on possible actions or system boundaries. This tendency was also evident in more basic queries, "Do I have to talk to it [the robot]?" and logistical questions, for example, "So now I can start again? Do I have to restart from zero or should I step out of the field and come back? [to the experimenter]" (Wave 2). These excerpts highlight not only a search for reassurance but also the need for clearer cues and support mechanisms within the human-robot interface, especially in scenarios where expectations and interaction rules may not be fully explicit.

Positive Reactions to Robot-Initiated Exchanges

Participants sometimes reacted with surprise or curiosity when the robot initiated the conversation. Exclamations such as "It's impressive! [the robot]" (Wave 2) or "Oh! It [the robot] talks to me first, just like a person!" (Wave 2) show an appreciation for its apparent autonomy and proactive engagement. This behavior was perceived as a sense of dynamism to the interaction. The realization that the robot could sustain the exchange on its own, "So if I don't say anything, it will keep asking me things?" (Wave 2), further stimulated engagement.

Discussion

Principal Findings

This study evaluated the integration of a SAR into the routine activities of a geriatric hospital unit through a 2-wave, real-world comparative design. In both waves, older adults interacted freely with the robot within its functional domains (eg, hospital orientation, entertainment, and appointment assistance), without scripted prompts, in order to preserve the spontaneity. Quantitative measures included system performance indicators (error types and rates), user engagement (verbal, physical, and emotional dimensions), interaction success, and subjective evaluations of acceptability and usability; these were complemented by a thematic analysis of postinteraction interviews focusing on perceptions of robot errors.

Results showed that the introduction of an LLM between waves was associated with marked improvements in conversational performance, including a substantial reduction in verbal errors, an increase in error-free interactions, and higher rates of task success. These gains in system reliability were mirrored by higher acceptability (AES) and usability (SUS) ratings in the second wave. Yet, these gains, concerning HRI and user experience, should not be viewed as solely the product of the LLM; aspects of the study design (eg, different samples) and contextual factors (eg, the kind of questions users asked the robot in each wave) may also have shaped the outcomes.

On another level, qualitative analysis showed that participants' spontaneous reactions to the robot were shaped by both system performance and their own expectations. Breakdowns often prompted humor, self-blame, or requests for clarification, while successful or unexpected robot-initiated exchanges could elicit surprise and positive engagement. These findings suggest that emotional and behavioral responses in SAR-older adult interactions are influenced as much by expectation management and novelty by technical reliability.

These observations are consistent with prior work on how humans perceive and respond to erroneous robots. Mirnig et al

XSL•FO
RenderX

[47] found that user reactions to robot errors are not exclusively negative but often include adaptive strategies such as humor, indulgence, or self-attribution of blame. Minor errors may even strengthen engagement by rendering the robot more "human-like," provided that such errors remain occasional and do not critically hinder task achievement. In line with this, several of our participants coped with breakdowns through irony or laughter, suggesting that, in geriatric care contexts, error management is not only a technical challenge but also a social and affective process shaping the overall interaction climate.

The following discussion interprets these findings in light of existing literature, examining how the observed changes in performance, engagement, and user perceptions contribute to current knowledge on SAR deployment in geriatric care and highlighting implications for future system design and implementation.

LLM Integration and Conversational Reliability

The results suggest that the integration of the LLM, introduced between wave 1 (without LLM) and wave 2 (with LLM), enhanced the SAR's ability to process diverse user input and produce contextually relevant responses. This improvement was evident not only in lower rates of comprehension failures and incoherent replies but also in greater interactional fluidity, characteristics essential for sustaining user engagement. The marked improvement in interaction success across waves (wave 1: 9/36, 25% vs wave 2: 70/94, 74.5%) further underscores the practical benefits of LLM integration for enabling communicative goals to be achieved without breakdowns. These gains align with the observed reductions in verbal errors and the higher proportion of error-free interactions, indicating that technical reliability is closely tied to the system's capacity to support smooth exchanges. In addition, successful interactions were associated with higher acceptability (AES) and usability (SUS) ratings, reinforcing prior evidence that perceived system competence is a key determinant of user satisfaction and trust in SARs.

Such outcomes align with initial evidence from other domains, including customer service [48], museum guidance [49], and the use of SARs for education [50,51], where LLMs have been reported to improve language comprehension, contextual adaptation, and response coherence [52]. Compared to rule-based or task-specific systems, systems equipped with LLM seem to offer more fluid, adaptive, and context-sensitive interactions [53,54], which is consistent with our findings, showing a significant improvement in interaction success and perceived acceptability and usability after LLM integration.

Recent work with older adults further demonstrates the potential of LLM-powered systems to support health and well-being in real-life care contexts. For example, conversational agents have been used in home environments to monitor safety risks, verify symptoms, and initiate alerts during emergencies [55]. Other studies have leveraged LLMs to collect richer health information with less provider effort [56] or to deliver cognitive stimulation through structured dialogue, resulting in improved task performance, social engagement, and high acceptance among older adults [57].

System and User Constraints in LLM-Enhanced Older Adult-Robot Interaction

However, LLM-based applications, whether implemented in clinical or everyday settings involving older adults, also reveal limitations that can arise from both the system and the use. On the system side, latency from cloud-based speech recognition and LLM processing, as observed in Lima et al [57], can interrupt conversational flow and disrupt turn-taking, sometimes being perceived as inattentiveness. These challenges are consistent with broader observations in spoken dialogue systems for robotics, where developers must balance trade-offs between accuracy and latency [58]. While cloud-based systems often deliver higher accuracy, they introduce delays that can break the rhythm of interaction; conversely, on-premises systems avoid latency but are typically less accurate and more limited in vocabulary. Furthermore, pretrained recognizers are usually optimized on datasets that differ from the spontaneous, fragmentary, and context-dependent speech, common in real-world robot use, making adaptation resource-intensive. In our study, similar conversational breakdowns occurred in verbal-related errors, such as incomplete, irrelevant, or incoherent responses, which interrupted the exchange and occasionally led to repeated, circular interactions. We believe this might be due to a combination of several factors. First, even if we adapted the speech recognition model with French data, our on-premises solution has its limitations in terms of accuracy and confuses some words. Second, even if the LLM solution operated on a partner's cloud with sufficient resources, it could not always provide accurate and appropriate answers, even when the speech recognition worked flawlessly. Third, the system used a half-duplex audio, listening to the user only when the system was not speaking, which reduced interaction fluidity. Finally, even if these limitations are mild individually, their combination could cause misinterpretation on the robot's side and generate user frustration, which did not ease the next steps of the interaction.

A further system-related limitation is that current LLMs are not specifically trained for interactions with older adults. As noted by Diaz et al [59] and Chu et al [60], speech and interaction data from older adults are scarce in artificial intelligence (AI) training corpora, and existing datasets often contain age-related biases. As a result, characteristic features of older adult communication, such as changes in prosody, vocabulary, conversational pacing, and the presence of hesitations or fragmented discourse in noisy environments, are underrepresented. This underrepresentation may reduce the ability of LLM-based systems to accurately interpret older adult speech, anticipate their communicative needs, and adapt to their interaction style in real-world settings.

From the user perspective, common age-related factors such as slower speech tempo, pauses, and fluctuations in volume can pose difficulties for speech recognition, particularly in the presence of cognitive decline [61]. Beyond these physiological and cognitive aspects, there is also an expectation gap in how older adults perceive conversational systems. Mahmood et al [62] describe how older adults often approach such technologies through the lens of human conversational norms, anticipating richer and more contextually adaptive exchanges than the system

XSL•FO
RenderX

can actually deliver. This tendency is reinforced, as noted by Mahmood et al [63] and Liu et al [64], by two recurring influences: (1) the way many voice interfaces are designed and promoted to mirror human conversational styles and (2) the inherent affordance of speech as an interaction modality, which can implicitly suggest that open-ended, human-like dialogue is possible when, in practice, the system's scope is more constrained and task-oriented.

These perceptions can be resistant to change, even with repeated exposure, and may lead to frustration when the robot's responses fall short of these implicit promises [65,66]. In our study, this was evident in qualitative data: some participants persisted in treating the robot as a human interlocutor, even after repeated conversational errors. For example, one participant humorously asked, "Does it [the robot] have blocked ears?" (Wave 1), while another, after receiving no reply, remarked, "Well, yes, but it [the robot] is not answering me" (Wave 1). Others attributed breakdowns to their own performance, asking, "Am I speaking loud enough [for the robot to understand]?" (Wave 1) or "Do I speak fast enough [for the robot to understand]?" (Wave 1). In some cases, repeated failures prompted disengagement or the need for external guidance, as illustrated by the query, "How can I find out what questions I can ask the robot?" (Wave 2).

User Engagement and Its Relationship With Interaction Outcomes

Verbal, physical, and emotional engagement scores in our study remained relatively stable between waves, with no statistically significant differences. Although wave 2 showed an improvement in interaction success, acceptability, and usability, these gains did not directly translate into measurable increases in engagement behaviors.

Emotional engagement showed a modest upward trend from wave 1 to wave 2, potentially reflecting more fluid exchanges with the LLM-enhanced system, but this change did not reach statistical significance. Because participants in the 2 waves were different individuals, the pattern is unlikely to reflect individual-level adaptation over time and may instead be related to situational factors or differences in interaction dynamics. It is also possible that more pronounced differences in engagement were not detected because the observed interactions were brief (approximately 52 seconds on average) or because the 5-point coding scale used to assess engagement may not have been sufficiently sensitive to capture subtle variations. Moreover, many of the analyzed exchanges were primarily functional in nature, where high levels of observable emotional engagement would not necessarily be expected, even successful interactions can occur with minimal emotional display (eg, "Where are the toilets?" (User), "At the end of the corridor, on your left" (Robot)).

Our findings also revealed strong correlations between emotional and physical engagement, indicating that when users are emotionally responsive, they also tend to be more physically or behaviorally expressive. However, unlike system performance indicators, these engagement dimensions were not directly associated with interaction success. For instance, participants sometimes laughed or joked at robot errors.

Interestingly, verbal engagement was negatively correlated with acceptability, echoing findings by Mahmood et al [63] that higher verbal activity in older adult-robot interactions can sometimes stem from the need to overcome conversational breakdowns rather than from enjoyment. In our context, this may indicate that some users spoke more in an attempt to repair misunderstandings, potentially lowering their subjective evaluations of the system.

Accessibility Challenges in Multimodal Robot-Older Adult Interaction

Across conversational HRI with SARs in geriatrics, particularly those powered by AI-based dialogue models, key accessibility concerns span model-training quality (representativeness and bias), user diversity and inclusion (cognitive, linguistic, sensory, and cultural variability), and interface ergonomics (voice or visual or physical modalities and interaction load).

First, studies highlight the underrepresentation of older adults in generative AI training databases, leading to persistent generational, ethnic, and cultural biases and limiting the recognition of nicknames, idioms, or nonstandard expressions in HRI [67-70].

Second, age-related changes in voice, such as reduced intelligibility, rhythm fluctuations, or speech disorders (dysarthria and aphasia), further complicate dialogue quality and voice processing [71,72]. Although targeted adaptation of acoustic models can reduce error rates, most recognition systems remain unable to accurately integrate the modulation and heterogeneity of vocal productions related to age and pathologies [73].

Third, in terms of interfaces, using the ARI tablet to display transcripts is a way of compensating for hearing or cognitive impairments, but literature on sensory impairment shows that true accessibility requires multimodal support (visual, tactile, and auditory). This includes optimized font size and contrast, pictograms, and gestural instructions for different levels of visual impairment [74-76].

Finally, the study did not assess the impact of specific pathologies (eg, aphasia and Parkinson disease) on speech recognition or multimodal adaptation. Future SAR development should rely on more diverse datasets, advanced multimodal calibration, and adaptive features (volume, articulation, and gesture sensitivity) to enhance inclusion and responsiveness in geriatric care.

Limitations of the Study

This study has several limitations. First, it was conducted in a single geriatric hospital unit, which limits the generalizability of the findings to other care contexts or to community-dwelling older adults. The relatively small sample size may have reduced the statistical power to detect certain associations, particularly regarding user characteristics. In addition, the 2-wave, cross-sectional design, where different participants were recruited for each wave, limits our ability to isolate the impact of the LLM from potential differences in participant profiles. Because the individuals differed between samples, variations in their characteristics may partly explain the observed

XSL•FO
RenderX

differences in HRI quality. A within-subjects design would allow a more precise assessment of the LLM's specific contribution to the outcomes.

Second, this study focused on stationary, face-to-face interactions and did not explore the effect of robot mobility (eg, autonomous navigation) or expressive gestures on user engagement. Prior research has shown that these dimensions can significantly enhance social presence and interactional involvement in SAR-older adult encounters, particularly in stimulating physical engagement and sustained attention [30].

Third, the analysis focused on short-term, functional interactions. It did not capture the relational and identity-based dimensions of conversation, which are especially meaningful for older adults. For many older adults, dialogue is not only just about obtaining information but also about sustaining social ties and sharing narratives of lived experiences. By overlooking these deeper aspects, our findings may underestimate key factors influencing long-term engagement with SARs. Future work should explore how LLM-enhanced dialogue systems can use biographical cues, narrative prompts, and adaptive strategies to better reflect the relational nature of older adults' communication.

Finally, the short interaction duration (52 seconds) provided only a snapshot of spontaneous use, limiting conclusions about long-term engagement and sustained acceptance. Longitudinal deployments are needed to capture how perceptions, trust, and interaction patterns evolve over extended use.

Recommendations for Future Development and Implementation of Conversational SARs Using LLMs

This study highlights the potential of conversational SARs powered by LLMs to substantially improve conversational reliability with older adults. Future development should build on these gains while addressing key limitations and contextual needs observed in our findings.

Adapt AI Architectures for Older Adult Use, Including Future Vision-Language Model and Vision-Language-Action Model Applications

Beyond LLMs, future conversational SARs are likely to incorporate vision-language models and vision-language-action models, enabling richer multimodal understanding and more context-aware behaviors. However, all such models should be specifically adapted for older adults, particularly those with cognitive impairment, whose speech patterns, vocabulary, cognitive processing, and sensory capabilities differ from the datasets typically used to train general-purpose systems. Responsible adaptation must comply with regulatory and ethical frameworks to ensure safe, equitable, and trustworthy use in care settings.

Address the Limited Prior Experience of Older Adults With Conversational SARs Through Structured Familiarization Periods

Many older adults, particularly those without prior exposure to conversational agents, may require time and guided support to understand the robot's capabilities, limitations, and "rules of engagement." Future implementations should include a facilitator-led familiarization phase, where older adults can safely explore functions, interaction styles, and question formats without performance pressure. This early stage can build confidence, reduce uncertainty, and lay the groundwork for more meaningful and sustained engagement once independent use begins.

Strengthen the Facilitator's Role During Familiarization

For both novice and older adults with cognitive impairment, a human facilitator can play a critical role in bridging early HRI challenges. This includes demonstrating effective input styles, explaining the robot's capabilities and limitations, and supporting older adults during breakdowns. Incorporating a structured familiarization period may accelerate learning, reduce frustration, and foster more positive first impressions.

Implement Adaptive Error-Handling Strategies

Conversational SARs should detect and repair interaction breakdowns early, such as prolonged silences, repeated input, or off-topic responses, by offering clarifying prompts, rephrased questions, or multimodal alternatives (eg, visual selection and touch inputs). This can maintain engagement and reduce the negative emotional impact of errors.

Enhance HRI Personalization Across Sensory, Cognitive, and Motivational Dimensions

Personalization should include tuning speech recognition thresholds, adjusting interaction pacing, adapting vocabulary complexity, and aligning content with individual preferences and abilities. For older adults with varying cognitive, sensory, and mobility profiles, this is critical for both accessibility and engagement.

Integrate Proactive and Context-Aware Conversational Strategies

Initiative-taking behaviors, such as relevant follow-up questions, personalized activity suggestions, and reference to prior interactions, can help sustain older adult interest. Care should be taken to ensure that such behaviors remain contextually appropriate and nonintrusive.

Evaluate Future Systems With Multidimensional Frameworks

Assessment should combine technical measures (eg, error rates and latency), behavioral indicators (eg, engagement scores and turn-taking dynamics), and experiential outcomes (eg, trust and perceived usefulness). Such an approach ensures that improvements target both system performance and user experience in real-world contexts.

Promote Secure and Transparent Deployment

Beyond technical performance, deploying LLM-enhanced SARs in geriatric care requires attention to ethics and practice. In our study, only consenting participants were recorded, and all data were pseudonymized and securely stored in compliance with the General Data Protection Regulation. While these measures reduce privacy risks, future use will demand strict safeguards and transparency on secondary data use. Caution is also needed against overtrust, as older adults may attribute undue authority

XSL•FO
RenderX

to robots. Clear system limits, caregiver oversight, staff training, and assessment of costs are vital to ensure safe, equitable, and sustainable integration in care.

Conclusions

In this real-world, 2-wave study in a geriatric hospital day care unit, integrating an LLM into a conversational SAR improved interaction success, conversation quality, and perceived system performance. Comparing a baseline dialogue system (wave 1) with an LLM-enhanced system (wave 2) showed that advanced AI can produce more coherent, contextually relevant exchanges. These improvements were reflected in higher acceptability and usability ratings from older adults. Conducting the study in a naturalistic care environment was key to capturing spontaneous interactions and real-world constraints, factors often absent in laboratory settings.

Our findings confirm the potential of LLMs to enhance HRI in care contexts. However, they also highlight the need to better understand how interaction success relates to engagement quality and to user sociodemographic characteristics. To maximize these benefits, future systems should be tuned to older adults' speech, cognitive, and sensory profiles. Strategies such as structured familiarization periods, adaptive error recovery, and personalized interactions could help maintain engagement over time. Aligning technical capabilities with the needs and preferences of older adults will be essential for making conversational SARs trusted and effective tools for geriatric care.

Acknowledgments

The authors would like to sincerely thank all the participants who took part in this study as well as all the health care professionals whose involvement contributed to the successful conduct of the Socially Pertinent Robots in Gerontological Healthcare project.

Funding

This research was funded by the EU H2020 program (grant 871245) [77].

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Study conception and design: LB (lead), XA-P (equal), ASR (equal), MP (equal)
Data collection: LB
Analysis and interpretation of results: LB (lead), MP (equal), JC (supporting)
Writing—original draft: LB
Writing—review and editing: LB (lead), MP (equal), ASR (equal), SD (equal)
All authors reviewed the results and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Evaluation scale Acceptability E-Scale and System Usability Scale (translated from French).
[DOCX File , 18 KB - humanfactors_v12i1e81936_app1.docx]

Multimedia Appendix 2
Correlations among behavioral, successful, and subjective measures.
[DOCX File , 281 KB - humanfactors_v12i1e81936_app2.docx]

References

1. Ageing and health. World Health Organization. 2025. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2025-08-12]
2. Loukissa D. Understanding and addressing aggressive and related challenging behaviors in individuals with dementia. AJH 2018;5(4):245-264. [doi: 10.30958/ajh.5-4-1]
3. Abley C. Developing a dementia care leaders' toolkit for older patients with cognitive impairment. Nurs Older People 2025;37(5) [FREE Full text] [doi: 10.7748/nop.2020.e1250]
4. Hendlmeier I, Bickel H, Heßler-Kaufmann JB, Schäufele M. Care challenges in older general hospital patients: impact of cognitive impairment and other patient-related factors. Z Gerontol Geriatr 2019;52(Suppl 4):212-221 [FREE Full text] [doi: 10.1007/s00391-019-01628-x] [Medline: 31628613]

XSL•FO
RenderX

5. Feil-Seifer D, Mataric M. Defining socially assistive robotics. 2005 Presented at: 9th International Conference on Rehabilitation Robotics, 2005. ICORR 2005; June 28, 2005; Chicago, IL, United States p. 465-468. [doi: 10.1109/icorr.2005.1501143]
6. Tapus A, Mataric A, Scassellati B. Socially assistive robotics [grand challenges of robotics]. IEEE Robot Autom Mag 2007;14:35-42 [FREE Full text] [doi: 10.1109/mra.2007.339605]
7. Hsu T, Hsieh C. Socially assistive robots assisting older adults in an internet and smart healthcare era: a literature review. 2022 Presented at: 2022 IEEE 4th Eurasia Conference on IOT. Communication and Engineering (ECICE) 2022.?594; October 28-30, 2022; Yunlin, Taiwan p. 591-594. [doi: 10.1109/ecice55674.2022.10042947]
8. Abdollahi H, Mahoor MH, Zandie R, Siewierski J, Qualls SH. Artificial emotional intelligence in socially assistive robots for older adults: a pilot study. IEEE Trans Affect Comput 2023;14(3):2020-2032 [FREE Full text] [doi: 10.1109/taffc.2022.3143803] [Medline: 37840968]
9. Dosso JA, Bandari E, Malhotra A, Guerra GK, Hoey J, Michaud F, et al. User perspectives on emotionally aligned social robots for older adults and persons living with dementia. J Rehabil Assist Technol Eng 2022;9:20556683221108364 [FREE Full text] [doi: 10.1177/20556683221108364] [Medline: 35782883]
10. Rigaud AS, Dacunha S, Harzo C, Lenoir H, Sfeir I, Piccoli M, et al. Implementation of socially assistive robots in geriatric care institutions: healthcare professionals' perspectives and identification of facilitating factors and barriers. J Rehabil Assist Technol Eng 2024;11:20556683241284765 [FREE Full text] [doi: 10.1177/20556683241284765] [Medline: 39324049]
11. Aymerich-Franch L, Ferrer I. Socially assistive robots' deployment in healthcare settings: a global perspective. Int J Human Robot 2023;20(1):2350002. [doi: 10.1142/s0219843623500020]
12. Graf B, Eckstein J. Service robots and automation for the disabled and nursing home care. In: Nof SY, editor. Springer Handbook of Automation. Cham: Springer International Publishing; 2023:1331-1347.
13. Sutherland C, Ahn B, Brown B, Lim J, Johanson D, Broadbent E, et al. The doctor will see you now: could a robot be a medical receptionist? 2019 Presented at: International Conference on Robotics and Automation (ICRA); May 20-24, 2019; Montreal, Quebec, Canada p. 4310-4316. [doi: 10.1109/icra.2019.8794439]
14. Babu A, Srikanth K. Empowering elderly care with social robots. Recent Res Rev J 2024;3(2):410-423 [FREE Full text] [doi: 10.36548/rrrj.2024.2.008]
15. González-González CS, Violant-Holz V, Gil-Iranzo RM. Social robots in hospitals: a systematic review. Appl Sci 2021;11(13):5976. [doi: 10.3390/app11135976]
16. Romero-Garcés A, Bandera J, Marfil R, González-García M, Bandera A. CLARA: building a socially assistive robot to interact with elderly people. Designs 2022;6(6):125. [doi: 10.3390/designs6060125]
17. Wong KLY, Hung L, Wong J, Park J, Alfares H, Zhao Y, et al. Adoption of artificial intelligence-enabled robots in long-term care homes by health care providers: scoping review. JMIR Aging 2024;7:e55257 [FREE Full text] [doi: 10.2196/55257] [Medline: 39190455]
18. Huisman C, Kort H. Two-year use of care robot Zora in Dutch nursing homes: an evaluation study. Healthcare (Basel) 2019;7(1) [FREE Full text] [doi: 10.3390/healthcare7010031] [Medline: 30791489]
19. Hebesberger D, Koertner T, Gisinger C, Pripfl J. A long-term autonomous robot at a care hospital: a mixed methods study on social acceptance and experiences of staff and older adults. Int J Soc Robotics 2017;9(3):417-429. [doi: 10.1007/s12369-016-0391-6]
20. Ghafurian M, Hoey J, Dautenhahn K. Social robots for the care of persons with dementia: a systematic review. ACM Trans Hum-Robot Interact 2021;10(4):1-31. [doi: 10.1145/3469653]
21. Lima MR, Wairagkar M, Gupta M, Rodriguez Y Baena F, Barnaghi P, Sharp DJ, et al. Conversational affective social robots for ageing and dementia support. IEEE Trans Cogn Dev Syst 2022;14(4):1378-1397. [doi: 10.1109/tcds.2021.3115228]
22. Wald S, Puthuveetil K, Erickson Z. Do mistakes matter? Comparing trust responses of different age groups to errors made by physically assistive robots. ArXiv Preprint posted online on August 23, 2024 [FREE Full text] [doi: 10.1109/ro-man60168.2024.10731271]
23. Giorgi I, Tirotto F, Hagen O, Aider F, Gianni M, Palomino M, et al. Friendly but faulty: a pilot study on the perceived trust of older adults in a social robot. IEEE Access 2022;10:92084-92096. [doi: 10.1109/access.2022.3202942]
24. Kim S. Exploring how older adults use a smart speaker-based voice assistant in their first interactions: qualitative study. JMIR Mhealth Uhealth 2021;9(1):e20427 [FREE Full text] [doi: 10.2196/20427] [Medline: 33439130]
25. Khosla R, Chu M, Khaksar SMS, Nguyen K, Nishida T. Engagement and experience of older people with socially assistive robots in home care. Assist Technol 2021;33(2):57-71. [doi: 10.1080/10400435.2019.1588805] [Medline: 31063044]
26. Miyake T, Wang Y, Yang PC, Sugano S. Feasibility study on parameter adjustment for a humanoid using LLM tailoring physical care. 2023 Presented at: 15th International Conference, ICSR 2023; December 3-7, 2023; Doha, Qatar p. 230-243. [doi: 10.1007/978-981-99-8715-3_20]
27. Esteban-Lozano I, Castro-González Á, Martínez P. Using a LLM-based conversational agent in the social robot Mini. In: Artificial Intelligence in HCI. Cham: Springer Nature Switzerland; 2024:15-26.
28. Pinto-Bernal M, Biondina M, Belpaeme T. Designing social robots with LLMs for engaging human interaction. Appl Sci 2025 Jun 05;15(11):6377 [FREE Full text] [doi: 10.3390/app15116377]

XSL•FO
RenderX

29. Oertel C, Castellano G, Chetouani M, Nasir J, Obaid M, Pelachaud C, et al. Engagement in human-agent interaction: an overview. Front Robot AI 2020;7:92 [FREE Full text] [doi: 10.3389/frobt.2020.00092] [Medline: 33501259]
30. Del Duchetto F, Baxter P, Hanheide M. Are you still with me? Continuous engagement assessment from a robot's point of view. Front Robot AI 2020;7:116 [FREE Full text] [doi: 10.3389/frobt.2020.00116] [Medline: 33501282]
31. Rossi S, Conti D, Garramone F, Santangelo G, Staffa M, Varrasi S, et al. The role of personality factors and empathy in the acceptance and performance of a social robot for psychometric evaluations. Robotics 2020;9(2):39. [doi: 10.3390/robotics9020039]
32. Kuo IH, Rabindran JM, Broadbent E, Lee YI, Kerse N, Stafford RMQ, et al. Age and gender factors in user acceptance of healthcare robots. 2009 Presented at: RO-MAN 2009—The 18th IEEE International Symposium on Robot and Human Interactive Communication; September 27-October 2, 2009; Toyama, Japan p. 214-219. [doi: 10.1109/roman.2009.5326292]
33. Flandorfer P. Population ageing and socially assistive robots for elderly persons: the importance of sociodemographic factors for user acceptance. Int J Popul Res 2012;2012:1-13. [doi: 10.1155/2012/829835]
34. Lee SH, Kim JS, Yu S. The impact of care robots on older adults: a systematic review. Geriatr Nurs 2025;65:103507. [doi: 10.1016/j.gerinurse.2025.103507] [Medline: 40669132]
35. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
36. Alameda-Pineda X, Addlesee A, Hernández García D, Reinke C, Arias S, Arrigoni F, et al. Socially pertinent robots in gerontological healthcare. Int J Soc Robot 2025 Nov 01. [doi: 10.1007/s12369-025-01330-6]
37. Blavette L, Dacunha S, Alameda-Pineda X, Hernández García D, Gannot S, Gras F, et al. Acceptability and usability of a socially assistive robot integrated with a large language model for enhanced human-robot interaction in a geriatric care institution: mixed methods evaluation. JMIR Hum Factors 2025;12:e76496 [FREE Full text] [doi: 10.2196/76496] [Medline: 40750072]
38. SPRING D1.6 ' user feedback from the final validation in relevant environments. SPRING Consortium. 2024. URL: https://spring-h2020.eu/wp-content/uploads/2024/07/SPRING_D1.6_-User-feedback-from-the-final-validation-relevant-environments_vFinal_31.05.2024.pdf [accessed 2025-08-13]
39. Anzalone SM, Boucenna S, Ivaldi S, Chetouani M. Evaluating the engagement with social robots. Int J Soc Robot 2015;7(4):465-478. [doi: 10.1007/s12369-015-0298-7]
40. Hemmler VL, Kenney AW, Langley SD, Callahan CM, Gubbins EJ, Holder S. Beyond a coefficient: an interactive process for achieving inter-rater consistency in qualitative coding. Qual Res 2020;22(2):194-219. [doi: 10.1177/1468794120976072]
41. Indurkhya X, Venture G. Lessons in developing a behavioral coding protocol to analyze in-the-wild child–robot interaction events and experiments. Electronics 2024;13(7):1175. [doi: 10.3390/electronics13071175]
42. Heerink M, Krose B, Evers V, Wielinga B. The influence of a robot's social abilities on acceptance by elderly users. 2006 Presented at: ROMAN 2006—The 15th IEEE International Symposium on Robot and Human Interactive Communication; September 6-8, 2006; Hatfield, United Kingdom p. 521-526. [doi: 10.1109/roman.2006.314442]
43. Micoulaud-Franchi JA, Sauteraud A, Olive J, Sagaspe P, Bioulac S, Philip P. Validation of the French version of the Acceptability E-scale (AES) for mental E-health systems. Psychiatry Res 2016;237:196-200. [doi: 10.1016/j.psychres.2016.01.043] [Medline: 26809367]
44. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the System Usability Scale. Int J Human-Comput Interact 2008;24(6):574-594. [doi: 10.1080/10447310802205776]
45. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. JUX J User Exp 2025:114-123 [FREE Full text] [doi: 10.7717/peerj-cs.1918/supp-1]
46. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2008;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
47. Mirnig N, Stollnberger G, Miksch M, Stadler S, Giuliani M, Tscheligi M. To err is robot: how humans assess and act toward an erroneous social robot. Front Robot AI 2017;4:1-15. [doi: 10.3389/frobt.2017.00021]
48. Peddinti SR, Katragadda SR, Pandey BK, Tanikonda A. Utilizing large language models for advanced service management: potential applications and operational challenges. J Sci Technol 2023;4(2) [FREE Full text] [doi: 10.2139/ssrn.5119925]
49. Vasic I, Fill HG, Quattrini R, Pierdicca R. LLM-aided museum guide: personalized tours based on user preferences. 2024 Presented at: International Conference, XR Salento 2024; September 4-7, 2024; Lecce, Italy p. 249-262. [doi: 10.1007/978-3-031-71710-9_18]
50. Li C, Xing W, Leite W. Building socially responsible conversational agents using big data to support online learning: a case with Algebra Nation. Br J Educ Technol 2022 May 10;53(4):776-803 [FREE Full text] [doi: 10.1111/bjet.13227]
51. Voultsiou E, Vrochidou E, Moussiades L, Papakostas GA. The potential of large language models for social robots in special education. Prog Artif Intell 2025;14(2):165-189. [doi: 10.1007/s13748-025-00363-2]
52. Chen J, Liu Z, Huang X, Wu C, Liu Q, Jiang G, et al. When large language models meet personalization: perspectives of challenges and opportunities. World Wide Web 2024;27(4):42. [doi: 10.1007/s11280-024-01276-1]
53. Wang C. Comparing Rule-Based and LLM-Based Methods to Enable Active Robot Assistant Conversations. 2025. URL: https://www.semanticscholar.org/paper/Comparing-Rule-based-and-LLM-based-Methods-to-Robot-Wang/54bffd8096740732ce4bf6486b513a261f05e313 [accessed 2025-08-12]

XSL•FO
RenderX

54. Kim C, Lee C, Mutlu B. Understanding large-language model (LLM)-powered human-robot interaction. 2024 Presented at: Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction; March 11-15, 2024; Boulder, CO, United States p. 371-380. [doi: 10.1145/3610977.3634966]
55. Lima MR, Su T, Jouaiti M, Wairagkar M, Malhotra P, Soreq E, et al. Discovering behavioral patterns using conversational technology for in-home health and well-being monitoring. IEEE Internet Things J 2023;10(21):18537-18552. [doi: 10.1109/jiot.2023.3290833]
56. Yang Z, Xu X, Yao B, Rogers E, Zhang S, Intille S, et al. Talk2Care: an LLM-based voice assistant for communication between healthcare providers and older adults. 2024 Presented at: Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies; October 9 2024; UbiComp/ISWC p. 1-35 URL: https://dl.acm.org/doi/10.1145/3659625 [doi: 10.1145/3659625]
57. Lima M, O'Connell A, Zhou F, Nagahara A, Hulyalkar A, Deshpande A, et al. Promoting cognitive health in elder care with large language model-powered socially assistive robots. 2025 Presented at: CHI '25: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems; April 26-May 1, 2025; Yokohama, Japan p. 1-22. [doi: 10.1145/3706598.3713582]
58. Marge M, Espy-Wilson C, Ward NG, Alwan A, Artzi Y, Bansal M, et al. Spoken language interaction with robots: recommendations for future research. Comput Speech Lang 2022;71:101255. [doi: 10.1016/j.csl.2021.101255]
59. Diaz M, Johnson I, Lazar A, Piper A, Gergle D. Addressing age-related bias in sentiment analysis. 2018 Presented at: CHI '18: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems; April 21-26, 2018; Montreal, Quebec, Canada p. 1-14. [doi: 10.1145/3173574.3173986]
60. Chu CH, Nyrup R, Leslie K, Shi J, Bianchi A, Lyn A, et al. Digital ageism: challenges and opportunities in artificial intelligence for older adults. Gerontologist 2022;62(7):947-955 [FREE Full text] [doi: 10.1093/geront/gnab167] [Medline: 35048111]
61. Kobayashi M, Kosugi A, Takagi H, Nemoto M, Nemoto K, Arai T, et al. Effects of age-related cognitive decline on elderly user interactions with voice-based dialogue systems. 2019 Presented at: IFIP Conference on Human-Computer Interaction; September 2-6, 2019; Paphos, Cyprus p. 53-74. [doi: 10.1007/978-3-030-29390-1_4]
62. Mahmood A, Cao S, Stiber M, Antony VN, Huang C. Voice assistants for health self-management: designing for and with older adults. 2025 Presented at: CHI '25: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems; April 26-May 1, 2025; Yokohama, Japan p. 1-22. [doi: 10.1145/3706598.3713839]
63. Mahmood A, Wang J, Huang CM. Situated understanding of errors in older adults' interactions with voice assistants: a month-long, in-home study. ArXiv Preprint posted online on September 23, 2024. [doi: 10.48550/arXiv.2403.02421]
64. Liu YC, Chen CH, Lin YS, Chen HY, Irianti D, Jen TN, et al. Design and usability evaluation of mobile voice-added food reporting for elderly people: randomized controlled trial. JMIR Mhealth Uhealth 2020;8(9):e20317 [FREE Full text] [doi: 10.2196/20317] [Medline: 32985999]
65. Kontogiorgos D, Tran M, Gustafson J, Soleymani M. A systematic cross-corpus analysis of human reactions to robot conversational failures. 2020 Presented at: Proceedings of the 2021 International Conference on Multimodal Interaction; October 18-22, 2021; Montréal, Qubec, Canada URL: https://dl.acm.org/doi/10.1145/3462244.3479887 [doi: 10.1145/3462244.3479887]
66. Liu S, Parreira M, Ju W. “I'm Done”: describing human reactions to successive robot failure. 2025 Presented at: 20th ACM/IEEE International Conference on Human-Robot Interaction (HRI); March 4-6, 2025; Melbourne Convention Center, Melbourne, Australia p. 1458. [doi: 10.1109/hri61500.2025.10974098]
67. Park JS, Bernstein MS, Brewer RN, Kamar E, Morris MR. Understanding the describing human reactions to successive robot failure. 2021 Presented at: AIES '21: AAAI/ACM Conference on AI, Ethics, and Society; May 19-21, 2021; Virtual Event p. 834-842. [doi: 10.1145/3461702.3462590]
68. Salomé S, Monfort E. Révolution numérique et âgisme: les enjeux éthiques de l’intelligence artificielle pour les personnes âgées [The digital revolution and ageism: the ethical challenges of artificial intelligence for older people]. NPG Neurol Psych Gériatr 2023;23(138):383-387. [doi: 10.1016/j.npg.2023.09.004]
69. Ali R, Tang OY, Connolly ID, Abdulrazeq HF, Mirza FN, Lim RK, et al. Demographic representation in 3 leading artificial intelligence text-to-image generators. JAMA Surg 2024 Jan 01;159(1):87-95 [FREE Full text] [doi: 10.1001/jamasurg.2023.5695] [Medline: 37966807]
70. Bozkurt S, Cahan E, Seneviratne MG, Sun R, Lossio-Ventura JA, Ioannidis JPA, et al. Reporting of demographic data and representativeness in machine learning models using electronic health records. J Am Med Inform Assoc 2020;27(12):1878-1884 [FREE Full text] [doi: 10.1093/jamia/ocaa164] [Medline: 32935131]
71. Aman F, Vacher M, Rossato S, Dugheanu R, Portet F, le Grand J, et al. Etude de la performance des modèles acoustiques pour des voix de personnes âgées en vue de l'adaptation des systèmes de RAP [Assessment of the acoustic models performance in the ageing voice case for ASR system adaptation]. 2012 Presented at: Proceedings of the Joint Conference JEP-TALN-RECITAL 2012, Volume 1: JEP; June 2012; Grenoble, France URL: https://aclanthology.org/F12-1089/
72. Vigouroux N, Truillet P, Privat R. Etude de l¿effet du vieillissement sur les productions langagières et sur les performances en reconnaissance automatique de la parole. 2004. URL: https://www.semanticscholar.org/paper/

XSL•FO
RenderX

Etude-de-l%C2%BFeffet-du-vieillissement-sur-les-et-sur-Vigouroux-Truillet/1b861e7349333a62466a27e7cad81f6ea4fc7404 [accessed 2025-10-03]
73. Klaassen W, Dijk BV, Spruit M. A review of challenges in speech-based conversational AI for elderly care. Stud Health Technol Inform 2024;327:858-862 [FREE Full text] [doi: 10.48550/arXiv.2412.07388]
74. Merenda T, Depasse F, Patris S. Déficience visuelle: développement et validation de recommandations sur les aides techniques applicables dans la pratique officinale et détermination de la population cible [Visual impairment: development and validation of recommendations on technical aids applicable in pharmacy practice and determination of the target population]. Le Pharm Clin 2024;59(2):e118-e119. [doi: 10.1016/j.phacli.2024.04.197]
75. Simonnet M, Bothorel C, Maximiano LF, Thépaut A. GeoTablet, une application cartographique pour les personnes déficientes visuelles. 2012. URL: https://www.semanticscholar.org/paper/GeoTablet%2C-une-application-cartographique-pour-les-Simonnet-Bothorel/88d477be711b72f914f692676ef3525e9848b31f [accessed 2025-10-03]
76. Gonsalves C, Pichora-Fuller MK. The effect of hearing loss and hearing aids on the use of information and communication technologies by community-living older adults. Can J Aging 2008;27(2):145-157. [doi: 10.3138/cja.27.2.145] [Medline: 18845510]
77. SPRING: Socially Pertinent Robots in Gerontological Healthcare. 2025. URL: https://spring-h2020.eu/ [accessed 2025-08-13]

Abbreviations

AES: Acceptability E-Scale
AI: artificial intelligence
DCH: day care hospital
HRI: human-robot interaction
LLM: large language model
MMSE: Mini-Mental State Examination
SAR: socially assistive robot
SPRING: Socially Pertinent Robots in Gerontological Healthcare
SUS: System Usability Scale

Edited by E Borycki; submitted 13.08.25; peer-reviewed by C González-González, DC Muchaluat-Saade, R Singh; comments to author 13.09.25; revised version received 27.10.25; accepted 29.10.25; published 03.12.25.

Please cite as:
Blavette L, Dacunha S, Alameda-Pineda X, Cattoni J, Rigaud AS, Pino M
Integrating a Large Language Model Into a Socially Assistive Robot in a Hospital Geriatric Unit: Two-Wave Comparative Study on Performance, Engagement, and User Perceptions
JMIR Hum Factors 2025;12:e81936
URL: https://humanfactors.jmir.org/2025/1/e81936
doi:10.2196/81936
PMID:

©Lauriane Blavette, Sébastien Dacunha, Xavier Alameda-Pineda, Jeanne Cattoni, Anne-Sophie Rigaud, Maribel Pino. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 03.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Enhancing Safety in Clinical Practice: Cross-Sectional, Within-Subjects, Simulated-Use Compliance Study Using BD PosiFlush SafeScrub

Kyle De Boer[1], RN, BSN, MS-HCM; John Simpson[1], MS; Purnima Sharma[1], PhD; Trevor J Steele[1], PhD

Medical Affairs, BD Biosciences (United States), Franklin Lakes, NJ, United States

Corresponding Author:
Kyle De Boer, RN, BSN, MS-HCM
Medical Affairs
BD Biosciences (United States)
1 Becton Drive
Franklin Lakes, NJ, 07417
United States
Phone: 1 2018476800
Email: kyle.de.boer@bd.com

Abstract

Background: Needleless access devices are essential for intravenous therapy but can be a source of contamination and catheter-related bloodstream infections (CRBSIs) if not disinfected properly. The BD PosiFlush SafeScrub (Becton, Dickinson and Company) is designed to support aseptic nontouch technique (ANTT) by incorporating a built-in reminder to “scrub-the-hub” before flushing. This feature can help improve compliance with disinfection practices and may reduce the risk of microbial contamination.

Objective: This study aimed to evaluate compliance with scrubbing before flushing using the BD PosiFlush SafeScrub in a simulated clinical environment compared with standard disinfection and flushing practices (alcohol swabs and prefilled saline syringes).

Methods: A cross-sectional, within-subjects, simulated-use compliance study was conducted with health care professionals familiarized with BD PosiFlush SafeScrub (a prefilled BD PosiFlush Syringe with an integrated disinfecting unit). Compliance was defined according to the disinfection procedure specified for each scenario; participants were considered compliant with standard practice if they followed their own institutional policy (ranging from 5 to 30 seconds or based on stroke counts), while compliance with BD PosiFlush SafeScrub required scrubbing for at least 10 seconds with a minimum of 8 clockwise and 8 counterclockwise rotations, in accordance with the instructions for use (IFU). Compliance with disinfection was monitored and recorded for both the BD PosiFlush SafeScrub and standard disinfection and flushing practice.

Results: Compliance with catheter hub disinfection was assessed among 60 participants for BD PosiFlush SafeScrub and 57 participants for standard practice. During preaccess procedures (Flush 1), BD PosiFlush SafeScrub achieved 46% compliance versus 21% with standard practice, representing an absolute improvement of 25% and a 119% relative improvement (P<.001). During the postmedication procedure (Flush 2), compliance was 22% with BD PosiFlush SafeScrub compared with 13% for standard practice, corresponding to a 9% absolute improvement and 69% relative improvement, although not statistically significant (P=.12). Overall, the compliance rate was 34% (81/240 interactions) in the BD PosiFlush SafeScrub group compared with 17% (39/228 interactions) in the standard practice group, representing an absolute improvement of 17% and a relative improvement of 100% (P<.001).

Conclusions: The BD PosiFlush SafeScrub, with its integrated disinfection unit, yielded approximately double the scrub-the-hub compliance (34%) before flushing compared to the standard practice of alcohol pads and prefilled saline syringes (17%), supporting its role in facilitating adherence to ANTT, which may reduce microbial growth.

(JMIR Hum Factors 2025;12:e78967) doi:10.2196/78967

KEYWORDS

BD PosiFlush SafeScrub; needleless access device disinfection; disinfection compliance; simulation study; scrub-the-hub; scrubbing; flushing

XSL•FO
RenderX

Introduction

Background

Vascular access devices (VADs) and needleless access devices (NADs) are essential tools in modern health care, enabling the administration of medications, fluids, and nutrition. NADs—such as needle-free connectors, Y-sites, and stopcocks—are connected to the VAD hub or extension set to support intermittent or continuous infusion. These devices help maintain a closed system, reducing the risk of needlestick injuries and contamination, and are now considered integral to infection control practices [1,2].

Due to their routine use, safe disinfection and flushing practices are critical. Among the most significant complications associated with these devices are catheter-related bloodstream infections (CRBSIs) and central-line-associated bloodstream infections (CLABSIs), which are known to increase patient mortality, hospital length of stay, and health care costs. Recent data show that CLABSIs are associated with a 3.5-fold increased risk of mortality and can extend hospital stays by up to 24.9 days, with associated costs reaching as high as US $90,000 per case [3]. To mitigate these risks, updated guidelines strongly emphasize adherence to standardized insertion bundles, education, and site care and maintenance practices, including the use of antimicrobial NADs and active or passive disinfection [1].

The 2024 revision to the Infusion Therapy Standards of Practice recommends the use of active or passive disinfection to "disinfect the connection surface and sides of the needleless connector attached to any VAD" [1]. In addition, the standard recommends performing active disinfection by vigorous mechanical scrubbing containing 70% isopropyl alcohol or alcohol-based chlorhexidine suitable for use with medical devices. Buetti et al [4], emphasized the importance of hub disinfection using alcohol-based antiseptics for 15 seconds, recommending the use of disinfection caps or active scrubbing techniques to further reduce contamination and biofilm formation. These practices form an essential part of evidence-based bundles aimed at the prevention of CLABSIs.

Despite efforts to educate health care providers and improve disinfection agents, studies indicated that compliance with active disinfection of access ports remains low. A systematic review by Moureau et al [5], found that approximately 50% of central VAD infections stem from bacterial colonization of NADs due to a lapse in aseptic technique and inadequate disinfection. Despite this known risk, adherence to NAD disinfection protocols remains suboptimal, with observed compliance rates as low as 10%. Various factors have been identified that contribute to this noncompliance, including a lack of standardized protocols, high workload and time constraints, forgetfulness, and limited availability of necessary supplies, such as alcohol wipes at the bedside.

Devices like the BD PosiFlush SafeScrub (Becton, Dickinson and Company), which integrates disinfection and flushing into a single-use syringe, offer workflow efficiency and improved adherence to aseptic techniques. These innovations aim to reduce contamination risk and enhance patient safety through simplified, standardized practices. This study sought to evaluate health care providers' compliance with scrubbing practices when using the BD PosiFlush SafeScrub versus standard methods (alcohol wipes and prefilled syringes) to determine its effectiveness in promoting aseptic vascular access.

Problem Statement

This study investigated whether the BD PosiFlush SafeScrub can improve compliance with disinfection practice, before flushing compared to standard practice, which may reduce microbial growth.

Methods

Study Design

This cross-sectional, within-subjects, simulated-use compliance study aimed to evaluate compliance with disinfection protocols using the BD PosiFlush SafeScrub compared with BD PosiFlush prefilled saline syringes with alcohol swabs. "BD PosiFlush saline syringe with alcohol swabs" is hereafter referred to as the "standard practice." Sixty registered nurses from varied clinical specialties, including emergency department, intensive care unit (ICU), critical care unit, outpatient infusion, home health care, and medical-surgical units, were enrolled from acute care facilities. The recruitment and scheduling of participants for this simulation, survey-based study was done in a usability lab setting at Delve (Delve Research), a design consultancy and research services provider. The study was conducted in two geographical locations, Delve's Philadelphia office and Boston office, from June to July 2024.

Study Participants

All the study participants were healthy individuals. The registered nurses were from the emergency room, ICU, critical care unit, outpatient infusion center, home health care, and medical-surgical units. No participant had any hand injuries, latex sensitivity, or other conditions that might affect scrubbing behavior. The participants recruited for the study had prior experience with the care and maintenance of VADs and NADs.

Participants were recruited from acute care facilities with more than 50 beds. Inclusion criteria required participants to have experience with VADs and disinfection practices. Eligible individuals needed to be at least 18 years of age and proficient in understanding written and spoken English. Additionally, participants were required to have a minimum of 2 years of related nursing experience. They must not have participated in any usability studies related to flushing with prefilled syringes within 6 months of their participation in this study. This structured approach ensured that the study included a diverse and experienced group of health care professionals, providing a comprehensive evaluation of the BD PosiFlush SafeScrub in various clinical scenarios.

Materials

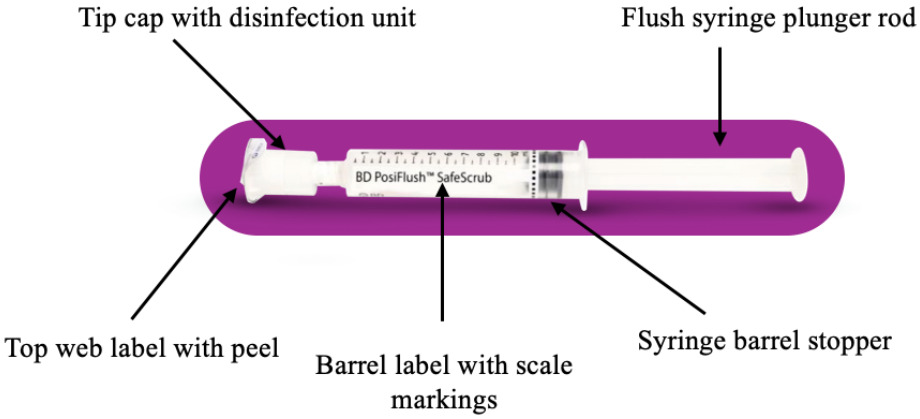
The materials used in this study included the BD PosiFlush SafeScrub Prefilled Saline Flush Syringe (Becton Dickinson and Co), standard alcohol swabs, and the BD PosiFlush Prefilled Flush Syringe (Becton Dickinson and Co). A mannequin was positioned in a simulated hospital bed or stretcher, with one

XSL•FO
RenderX

arm outfitted with a peripherally inserted central catheter (PICC; Becton Dickinson and Co) and the other with a peripheral intravenous catheter (PIVC; Becton Dickinson & Co). The study consisted of 60 simulated-use sessions, each lasting 60 minutes. Participants were observed through four simulated push medication scenarios, using both the standard practices and the BD PosiFlush SafeScrub. They were provided with all necessary materials typically found in a clinical setting, including gloves, alcohol swabs with 70% isopropyl alcohol, flush syringes filled with saline, prefilled simulated medication syringes (filled with colored saline), and other ancillary components. All participants were asked, before beginning the study, if the environment was similar to what they would see in their clinical areas of practice. This was done to ensure that the way they would respond to the scenario would be very close to what would occur in a real clinical setting.

The BD PosiFlush SafeScrub combines a prefilled saline syringe with an integrated disinfection cap, simplifying and streamlining the scrubbing and flushing workflow (Figure 1). It is a device designed to improve the disinfection process of NADs in vascular access systems. The dual-purpose design aims to simplify the disinfection process and improve compliance among health care providers with an all-in-one scrub and flush solution, ensuring consistent and effective disinfection before each flush. As a single-use product, it minimizes the risk of cross-contamination, making it a valuable tool in promoting aseptic nontouch technique (ANTT) in clinical settings [1]. By supporting consistent and effective disinfection, such devices play a vital role in reducing the contaminants that can cause bloodstream infections and enhance overall patient care.

Figure 1. BD PosiFlush SafeScrub labeled image.



Study Protocol

Following consent and study enrollment, participants were informed that their hospital adopted a new catheter hub disinfection product. They were then asked, “What would you do?” Based upon their responses, participants either received the instructions for use (IFU) or were shown a demonstration video (along with the IFU), which were the only 2 training formats provided to ensure consistency. A total of 60 simulated-use sessions, each lasting approximately 60 minutes in duration, were conducted. They were instructed to follow aseptic technique throughout the medication administration process to minimize the potential for touch contamination. Each participant completed 4 simulated intravenous push medication scenarios, 2 using the standard practice and 2 using the BD PosiFlush SafeScrub. The order for using each syringe and scenario was randomized. The randomization determined both the sequence of device use (standard practice vs SafeScrub) and the catheter type (peripherally inserted central catheter [PICC] vs peripheral intravenous catheter [PIVC]) for each scenario. A 4-factorial forced randomization design was used. Participants were randomly assigned to one of the several pregenerated scenario sequences using a computer-based randomization algorithm. Each participant received their assigned sequence upon enrollment, ensuring a balanced distribution of scenario order and device type.

Each medication administration scenario included a flush before medication delivery and a flush postdelivery; however, the medication delivery step was solely for contextual purposes and not included in data analysis. The data were collected on compliance with scrubbing before flushing within the medication administration process. Throughout the study, the focus was on the scrubbing and flushing process, comparing the standard practice with that of the BD PosiFlush SafeScrub.

Three pilot studies were conducted to identify any potential issues and provide the most realistic setup and scenarios for the current study. The protocol, which was chosen after the pilot studies, included nonemergent clinical scenarios and no planned distractions for the participants. This protocol was then applied to the full data collection for all 60 participants.

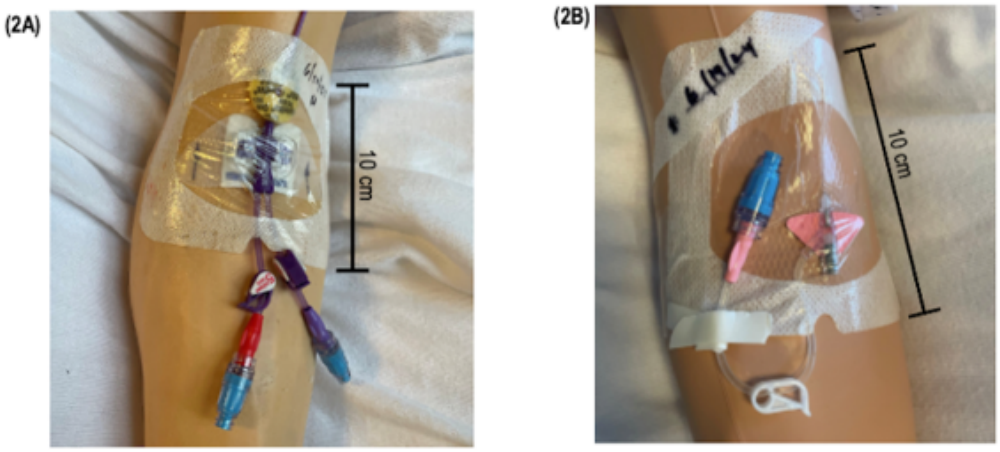
These panels illustrate the simulated clinical environments developed for evaluating catheter hub disinfection practices. In each setup, the disinfection site (hub) is depicted to show the scrub-the-hub step performed before flushing. Incorporating both PICC and PIVC scenarios allowed the study design to reflect routine vascular access procedures and provided a standardized framework for assessing compliance with hub disinfection across intervention and control groups.

A study session was set up as follows, a mannequin was positioned in a simulated hospital bed or stretcher, with one arm outfitted with a PICC and the other with a PIVC. Both catheters were equipped with NADs. The arrangement and

XSL•FO
RenderX

location of the PICC line setup and the PIVC line setup on the mannequin are shown in Figure 2A and 2B, respectively. These images accurately represent a clinical PICC or PIVC setup, with both lines connected to a NAD. The NAD was not covered with a passive disinfecting cap, reflecting certain clinician practices that may not require scrubbing before accessing the NAD after a passive cap has been removed.

Figure 2. Simulated catheter set-ups demonstrating scrub-the-hub disinfection workflow. (A) Peripherally inserted central catheters. (B) Peripherally intravenous catheter.



The study protocol included several key steps outlined as follows:

- Flush 1: this was the intended scrub and flush before accessing the line for the medication delivery. It ensured that the line was free of occlusion before medication was delivered.
- Medication delivery: after Flush 1, the medication was administered. It is important to note that there was no compliance scoring for the scrubbing done immediately before this step.
- Flush 2: after the medication was delivered, there was a second scrub to disinfect the hub, followed by a flush step. This was intended to disinfect the hub before a final flush, which ensures no medication residue remains in the line.
- Compliance scoring: participants were evaluated based on their adherence to the protocols for Flush 1 and Flush 2. Absolute improvement in compliance was calculated as the difference in compliance rates between the BD PosiFlush SafeScrub group and the standard practice group, {(absolute improvement) = [(compliance of SafeScrub) – (compliance of standard practice)]}. Relative improvement in compliance was calculated as the absolute improvement expressed as a percentage of the standard practice compliance rate, {[relative improvement] = [(compliance SafeScrub) – (compliance standard) / (compliance standard)] times 100]}.

In summary, the protocol is designed to ensure proper disinfection or cleaning during the administration of flushing and medication administration via an intravenous line, with specific compliance measures for the disinfecting steps before and after the medication delivery. The study compared the standard practice with the BD PosiFlush SafeScrub.

Data Collection

Data collected during the simulation study included user demographics, current practices, and hospital policies on scrubbing and flushing. Specific metrics recorded were the frequency and duration of scrubbing before flushing for both the standard practice and BD PosiFlush SafeScrub, as well as the drying time allowed before accessing the NAD. For each location, a facilitator was in the room with the participant, and note takers and observers viewed from a separate room or remotely by way of streaming video or poststudy video review (Multimedia Appendix 1). Participants' overall impressions and reactions to each method were also collected. Data on compliance was collected through direct observation and video recording of the disinfection and flushing workflow process. For standard practice scenarios, participants were asked what their hospital policy was for scrubbing prior to flushing, and they were considered "compliant" if they followed their own individual hospital policy. Policies ranged from scrubbing 5-30 seconds, and some involved the number of times scrubbing back and forth rather than duration. For BD PosiFlush SafeScrub scenarios, participants were considered "compliant" if they scrubbed for at least 10 seconds, for a minimum of eight repetitions clockwise alternating with eight repetitions counterclockwise, in accordance with the IFU. The time was recorded using a stopwatch on a phone. We acknowledge that defining compliance according to the BD PosiFlush SafeScrub IFU (ie, ≥ 10-second scrub with at least eight clockwise and eight counterclockwise repetitions) while using institution-specific scrubbing policies for the standard practice arm introduces methodological asymmetry. However, this approach was intentionally chosen to preserve external validity and generalizability by reflecting the diverse, real-world practices of health care providers in different institutional contexts. In pragmatic study designs, such variation in comparator definitions is often embraced to capture how an intervention performs under routine operational conditions without artificially standardizing comparator practices [6,7].

Moreover, our goal was to assess whether the SafeScrub device promotes compliance beyond prevailing norms, rather than to apply a uniform compliance threshold across both conditions. By allowing participants to follow their own institutional standards in the comparator arm, we both minimized bias from

XSL•FO
RenderX

altering habitual workflows and maintained the relevance of our findings to everyday practice.

Thus, the differing compliance definitions enhance the external validity of our results and reinforce the practical utility of the BD PosiFlush SafeScrub device within diverse clinical settings.

Additional data collection methods included prestudy screener criteria, manual observations by note takers, video recordings and images for poststudy analysis, and reviews by data analysts to finalize results and uncover additional insights. After completing all 4 scenarios, participants were asked about their hospital policies and what they stated regarding scrubbing catheter hubs. After participants used the BD PosiFlush SafeScrub in the study, they were asked to complete a user acceptance survey, which was comprised of 15 questions. The survey incorporated both 1-5 and 0-10 Likert-type scales, intentionally selected based on the nature of the information being gathered. The 1-5 scale was used for most survey items to assess agreement with specific usability and compliance statements related to the device (eg, ease of use, learning curve, and reinforcement of guidelines). This scale is standard for capturing general attitudes and perceptions and satisfaction levels with sufficient sensitivity while minimizing respondent fatigue. The 0-10 scale was selectively applied to 2 questions that evaluated broader behavioral intentions–likelihood to recommend and inclination to use the device in clinical practice. These items benefit from greater granularity and are commonly used in health care research. The expanded range enables differentiation among high-positive responses (eg, 8 vs 10), which is important for assessing clinical acceptability and forecasting real-world implementation.

To aid interpretability and comparison across all survey items, responses were summarized using both average ratings (mean [SD]) and the percentage of respondents selecting a high response (defined as ≥4 for 1-5 scale items and ≥8 for 0-10 scale items).

Prior to enrollment, we conducted a priori power analysis based on the three pilot studies for a 2-tailed independent samples *t* test to determine the required sample size, targeting 80% power at α=.05 to detect a meaningful between-group effect.

Data Analysis

Data were analyzed using R version 4.3.0 (R Foundation for Statistical Computing). Descriptive statistics summarize the data, and inferential statistics, such as chi-square tests, were used to determine the significance of the differences observed. The results were considered significant for all analyses at a 5% level of significance (P<.05).

Ethical Considerations

No patient data were collected in this study, as patients were not included. In accordance with the ethical standards of the Declaration of Helsinki, written informed consent was obtained from all study participants before study initiation [8,9]. Participants were assured that all identifying information would remain confidential and would not be disclosed in any study-related publications. All data were deidentified before analysis to maintain participant anonymity.

Participants were informed of the potential risks associated with the procedure. Furthermore, the intervention involved a commercially available, marketed product containing substances already used in routine clinical practice, specifically isopropyl alcohol and 0.9% sodium chloride (normal saline). Although formal IRB approval was not required due to the nature of the simulation and the minimal risk posed to participants who were only exposed to the test device and preinserted catheters, this simulated use study was conducted in accordance with clinical and ethical best practices under the guidance of Delve (Delve Research).

Results

Study Participants

Sixty registered nurses participated in the study, representing a diverse clinical background and varying levels of professional experience. The demographics of the participants are shown in Table 1. All participants completed the simulation and provided complete demographic and clinical information. Participants were employed across multiple health care settings, including adult emergency departments (n=11), adult intensive care or critical care units (n=16), outpatient infusion centers (n=15), home health care (n=5), and medical-surgical units (n=13). The average overall clinical experience among participants was 15.87 years, with an average of 9.09 years of experience in their current primary department.

Table 1. Demographic and clinical background information of participants.

Participants number (N=60)	Variables
Primary department of occupation, n	
Adult emergency room nurse	11
Adult critical care nurse	16
Outpatient infusion center nurse	15
Home health care (HHC) nurse	5
Med-surg nurse	13
Overall years of experience (years), mean (SD)	15.87 (10.23)
Years of experience in the current primary department (years), mean (SD)	9.09 (7.17)

XSL•FO
RenderX

Hospital Policy for Scrubbing Time

Participants reported a wide range of institutional policies related to catheter hub disinfection practices before flushing. Details on hospital policy counts and ranges are presented in Table 2. The most cited policy specified a scrubbing duration of 10 seconds, reported by 22 participants. A smaller number reported (n=9) scrubbing durations of 15 seconds, 15-30 seconds, and 20-30 seconds. One participant each reported scrubbing durations of 5-10 seconds, 5-15 seconds, or 25-30 seconds.

Variations in policy also included combined time and technique directives. Two participants noted a requirement of 10 scrubbing rotations, and one specified 10 rotations using 2 separate swabs. One participant reported different time-based policies for peripheral versus central lines (15 and 30 seconds, respectively), and another reported a general scrub for peripheral lines with a specified 15-second scrub for central lines. Three participants reported not knowing their hospital disinfection policy, with one stating that no disinfection policy was in place. A few participants (n=4) indicated policies in which only the initial flush required scrubbing, with no scrubbing required after medication administration.

Table 2. Reported hospital policies and ranges.

Reported hospital policy	Participants with hospital policy, n (%)
3-5 seconds first flush, no scrubbing second flush	1 (1.6)
5 seconds scrubbing time	1 (1.6)
10 seconds scrubbing time	22 (36.6)
15 seconds scrubbing time	9 (15)
15 seconds first flush, no scrubbing second flush	2 (3.3)
10 seconds first flush, no scrubbing second flush	1 (1.6)
30 seconds scrubbing time	6 (10)
5-10 seconds scrubbing time	1 (1.6)
5-15 seconds scrubbing time	1 (1.6)
10-15 seconds scrubbing time	2 (3.3)
15-30 seconds scrubbing time	2 (3.3)
25-30 seconds scrubbing time	1 (1.6)
20-30 seconds scrubbing time	2 (3.3)
10 rotations scrubbing time	2 (3.3)
10 rotations, 2 swabs	1 (1.6)
Peripheral just scrub, no time requirement Central 15 seconds	1 (1.6)
15 seconds peripheral, 30 seconds central	1 (1.6)
No requirement – as stated by the participant	1 (1.6)
Does not know – participants not familial with hospital policy	3 (5)

Compliance

Compliance with catheter hub disinfection protocols was assessed during both preaccess (Flush 1) and postmedication (Flush 2) procedures using either standard practices or the BD PosiFlush SafeScrub. For the BD PosiFlush SafeScrub scenarios, compliance was defined as scrubbing for ≥ 10 seconds with a minimum of 8 clockwise and 8 counterclockwise repetitions, as per the IFU. A total of 60 participants were evaluated in the SafeScrub group. In contrast, data from 57 participants were analyzed for standard practice due to incomplete hospital policy reporting from three individuals, rendering their standard practice compliance scoring invalid. Table 3 summarizes the compliance data for the participants using the standard practice and BD PosiFlush SafeScrub. Compliance was assessed across both the Flush 1 (preaccess) and Flush 2 (postmedication) procedures, and the total number of compliant interactions (overall compliance) was calculated for each method.

XSL•FO
RenderX

Table 3. Compliance data for the participants using the standard practice and BD PosiFlush SafeScrub.

Category	Standard practice (N=57)	BD PosiFlush SafeScrub (N=60)
Flush 1 (preaccess)		
PICC[a], n (%)	13 (23)	27 (45)
PIVC[b], n (%)	11 (19)	28 (47)
Total, n/N (%)	24/114 (21)	55/120 (46)
Flush 2 (postmedication)		
PICC, n (%)	6 (11)	12 (20)
PIVC, n (%)	9 (16)	14 (23)
Total, n/N (%)	15/114 (13)	26/120 (22)
Overall compliance, n/N (%)	39/228 (17)	81/240 (34)
P value	—[c]	<.001

[a]PICC: peripherally inserted central catheter.

[b]PIVC: peripheral intravenous catheter.

[c]Not available.

Flush 1 (Preaccess)

The BD PosiFlush SafeScrub achieved a compliance rate of 46% (55/120), significantly higher than the 21% (24/114) observed with standard practice. This corresponds to an absolute improvement of 25% (*P*<.001) and a relative improvement of 119% in preaccess disinfection compliance when using the BD PosiFlush SafeScrub.

Flush 2 (Postmedication)

Post-medication compliance was also higher with BD PosiFlush SafeScrub, at 22% (26/120) interactions, compared to just 13% (15/114) interactions with the standard practice. This represents an absolute improvement of 9% and a relative improvement of 69%; however, the difference did not reach statistical significance (*P*=.12).

Overall Compliance

Across all interactions, the overall compliance rate was 34% (81/240) interactions in the BD PosiFlush SafeScrub group compared with 17% (39/228) interactions in the standard practice group, representing an absolute improvement of 17% and a relative improvement of 100% (*P* <.001).

An exploratory post-hoc analysis applying a 10-second threshold in participants using the standard practice and BD PosiFlush SafeScrub was also conducted. A 2-sample proportion test demonstrated a statistically significant difference between groups (*P*<.000) such that the BD PosiFlush SafeScrub group had a higher overall compliance when compared with the standard practice group.

Summary of Root Causes

Participants often thought that they followed the IFU for scrubbing the BD PosiFlush SafeScrub for at least 10 seconds, for a minimum of 8 repetitions clockwise alternating with 8 repetitions counterclockwise, and many noted that they were counting times back and forth rather than the seconds in their head. For BD PosiFlush SafeScrub and standard practice, most participants reported one of the following reasons for not following hospital policy or IFU.

- The participants believed that they were following their hospital policy.
- The participant reported that they knew they were in a study and that no actual patient would be at risk for infection if they did not scrub the entire required time (study artifact).
- The participant reported that clinicians generally do not follow the hospital policy for scrubbing.

User Acceptance Survey Results

The results of the user acceptance survey revealed consistently high levels of agreement with positive usability statements, indicating acceptance of the device within clinical practice. The user acceptance survey responses are summarized in Table 4.

XSL•FO
RenderX

Table 4. User acceptance survey results for the BD PosiFlush SafeScrub device.

Parameters	Agreement (respondents selecting ≥4 on 1-5 scale or ≥8 on 0-10 scale), %	Average rating, mean (SD)
SafeScrub Device (1-5 rating)		
Ease of use	96.6	4.6 (0.62)
Learning curve	96.6	4.7 (0.58)
Impact on patient safety	91.6	4.6 (0.64)
Compliance with best practices	88.3	4.4 (0.75)
Reinforcement of guidelines	96.6	4.7 (0.52)
Ease of compliance	93.3	4.6 (0.55)
Labeling and integration impact	85	4.5 (0.62)
Efficiency improvement	58.3	4.4 (0.83)
Frequency of distractions	86	3.1 (0.98)
Impact of distractions	85	4.1 (0.76)
Overall sentiment	91.6	4.4 (0.67)
Clinical acceptability/impact (0-10 rating)		
Inclination to use in practice	43.3 (≥8 on scale)	8.7 (1.93)
Likelihood to recommend	78.3 (≥8 on scale)	8.5 (1.75)

For ease of use and learning curve, a total of 96.6% of participants agreed that the BD PosiFlush SafeScrub device is easy to use, with an average rating of 4.6 on a 5-point scale. Similarly, 96.6% indicated that most clinicians would learn to use the device quickly, yielding an average learning curve rating of 4.7. These findings reflect minimal training burden and favorable ergonomics.

For the impact of safety and compliance regarding patient safety, 91.6% of respondents reported that the device contributed to reducing the risk of microbial growth and potential CRBSIs, with an average rating of 4.6. In terms of clinical compliance, 88.3% believed that the device improves adherence to evidence-based best practices—such as ANTT and “scrub-the-hub”—relative to their current practice, with a rating of 4.4. Furthermore, 96.6% reported that the device design reinforces compliance with “scrub-the-hub” guidelines (rating: 4.7), and 93.3% indicated that becoming compliant with its use would be easy (rating: 4.6).

Device labeling and integration features, including the peel tab and built-in disinfection unit, were positively received by 85% of participants, with an average rating of 4.5. A total of 58.3% of participants felt that the elimination of the need to search for alcohol wipes improved workflow efficiency (rating: 4.4), and 43.3% of participants indicated a high inclination to adopt the device in daily practice, with a mean rating of 8.7 on a 10-point scale. Upon further investigation, we observed that around 30% did not select a response, leading to a low inclination for adoption of the device. Notably, 86% of respondents acknowledged frequent distractions during catheter access tasks, and 85% agreed that such distractions may cause clinicians to rush the disinfection process, both rated at 4.1.

The overall sentiment toward the BD PosiFlush SafeScrub was favorable, with 91.6% of participants rating their experience 4 or above on a 5-point scale (mean 4.4, SD 0.67). Additionally, 78.3% expressed a high likelihood of recommending the device to colleagues, with a recommendation score of 8.5 out of 10.

These findings suggest that the BD PosiFlush SafeScrub is perceived as a user-friendly and effective tool that enhances compliance with infection prevention protocols and supports safer, more effective clinical workflows.

Discussion

Principal Findings

The purpose of this study was to evaluate compliance with scrubbing before flushing using the BD PosiFlush SafeScrub in a simulated clinical environment compared with standard disinfection and flushing practices (alcohol swabs and prefilled saline syringes). The results revealed that the BD PosiFlush SafeScrub, with its integrated disinfection unit, yielded approximately double the scrub-the-hub compliance (34%) before flushing compared to the standard practice of alcohol pads and prefilled saline syringes (17%). This improvement aligns with prior compliance-focused interventions. In adult acute care, Samanta et al [10] demonstrated that structured interventions, including daily audits, bedside teaching, and reinforced scrub-the-hub protocols, raised compliance rates for insertion checklists (83.3%), full barrier precautions (78.4%), and scrub-the-hub technique (65%), underscoring the value of both behavioral and system-based measures in high-risk environments [10]. A recent review reports that the Institute for Safe Medication Practices suggests that educational interventions are the easiest to implement, but they are the least effective approaches [11]. Further, the authors stated that the “‘forcing functions,’ which are safety design features that necessitate actions occurring in the same way each time, are the most e ective type of intervention, as they can eliminate the risk

XSL•FO
RenderX

of error (ie, a high leverage strategy based on system reliability), although they are more complex to introduce."

Bundle approaches have also been investigated to improve patient safety. Devrim et al [12], reported that implementation of a central line bundle—including split-septum devices, needleless connectors, and single-use prefilled flushing syringes—reduced CLABSI incidence in pediatric oncology patients by 59% with an estimated cost saving of approximately US $208,977 [12]. Similarly, Samanta et al [10], observed a 43% reduction in CLABSI rates in an adult polytrauma ICU, with 49 consecutive CLABSI-free days, reinforcing the link between compliance and infection prevention. In a neonatal ICU, another study demonstrated that routine scrubbing of catheter hubs reduced CLABSI incidence from 1.89 to 0.23 per 1000 central line days [13]. Collectively, these findings confirm that interventions ranging from bedside timing aids to institutional bundles and device-integrated solutions can deliver both clinical and economic benefits across diverse populations.

Compared with system-level programs that rely on continuous auditing and retraining, the BD PosiFlush SafeScrub offers a device-level behavioral nudge that integrates disinfection into the flushing step. By reducing reliance on memory, product availability, and accurate time estimation—factors vulnerable to erosion under workload stress—the device helps standardize practice in real-world, high-pressure environments such as emergency departments or ICUs. Its integrated design not only simplifies workflow but also encourages habit formation and consistent adherence, even when time is constrained.

These findings align with another study, which showed that implementing a universal disinfection cap program reduced CLABSI rates by over 40% with annual cost savings of US $282,840, reinforcing the association between disinfection compliance and economic benefits [14]. By embedding disinfection directly into practice, BD PosiFlush SafeScrub provides a low-barrier complement to larger institutional initiatives, helping to close the persistent gap between policy awareness and actual compliance.

According to the Infusion Nurses Society Infusion Therapy Standards of Practice (Ninth Edition, 2024), disinfection of the needleless connector is required before each access, including before flushing and locking after medication administration [1]. While this study found that the BD PosiFlush SafeScrub device improved preflush compliance substantially, the same effect was not observed following administration of medication (Flush 2 compliance). This endpoint was powered a priori to detect such differences, suggesting that improvements in post-medication scrub compliance may require interventions beyond device integration.

In addition, participant feedback further highlighted variability in adherence. Several participants believed they were following the BD PosiFlush SafeScrub IFU (at least 10 seconds with a minimum of 8 clockwise and eight counterclockwise repetitions) but admitted to counting motions rather than measuring time. In standard practice scenarios, deviations were often attributed to perceived protocol familiarity, low perceived risk during study participation, or inconsistent hospital policies ranging from 3 to 30 seconds. These observations emphasize the challenges of ensuring consistent adherence to disinfection protocols and reinforce the need for education, standardization, and behavioral support tools [11].

Impact Statement

This study demonstrated a clinically significant increase in disinfection compliance following the use of BD PosiFlush SafeScrub. Enhanced compliance with disinfection practices is critical for helping to reduce the microbial burden, which may lead to health care–associated infections, providing better health care, and lowering associated costs. By promoting a culture of accountability and standardizing device disinfection, this intervention could serve as a model for improving patient safety and quality of care in various health care settings.

Limitations of the Study

This study has several limitations. The simulated nature of the design, while valuable for the controlled exploration of device performance, cannot fully capture the complexities, time pressures, and interruptions present in real-world clinical environments across different health care settings. Simulated conditions may inadvertently inflate compliance rates due to factors such as the Hawthorne effect, simplification of tasks, scripted scenarios, and the absence of real patient risk, which could alter participant behavior. For instance, clinicians might demonstrate more meticulous disinfection in a simulation when no actual patient is in danger. These limitations affect ecological validity and underscore the need for further research to validate these findings under actual clinical conditions, where time pressures, competing priorities, and workflow disruptions are present.

In addition, the generalizability of results may be limited due to all participants being based in the United States. Further, not all aspects of compliance and disinfection were measured, leaving some critical elements unevaluated, and compliance improvement was only assessed over a limited time frame, preventing conclusions about long-term adherence. Another limitation is that the intervention group received structured training (via IFU or demonstration video), whereas the control group followed existing hospital procedures without equivalent refresher training. This methodological asymmetry introduces a confounding effect, as training alone can increase adherence regardless of device performance. Further studies should consider designs that isolate training effects, such as crossover trials or parallel group designs, to better determine device-specific impact.

Additionally, given the current health care environment, cost-conscious decision-making, and prevailing standard practices, we recognize that the unit cost of an alcohol swab and a prefilled flush may be lower than that of the study product, BD PosiFlush SafeScrub. However, the true value of the device lies in its ability to enhance compliance, as demonstrated in this study. Its integrated design serves as a visual and functional prompt for clinicians to scrub the access site prior to flushing, helping reduce the risk of introducing contaminants that could lead to CRBSIs. Further studies should therefore evaluate cost-effectiveness, implementation feasibility, and sustainability dimensions in real-world settings.

XSL•FO
RenderX

Finally, three pilot studies informed the main study design. Based on the clinician's feedback, the control group was evaluated against their own hospital policies rather than a ≥10-second threshold to preserve external validity by reflecting real-world practice. We acknowledge that this approach may introduce methodological asymmetry. To address this concern, we conducted an exploratory post hoc analysis applying a 10-second compliance threshold to both groups. This analysis demonstrated a statistically significant difference in favor of the intervention, indicating that the findings are robust regardless of the compliance definition applied.

Conclusions

The BD PosiFlush SafeScrub device approximately doubled the scrub-the-hub compliance before flushing compared to alcohol pads and prefilled saline syringes, simplifying workflow by embedding disinfection into the flushing step. Further real-world studies are needed to confirm cost-effectiveness and environmental impact across diverse health care settings.

Acknowledgments

We would like to acknowledge the staff at Delve for facilitation of the study, assisting with study design, collection of data, summarizing data, and data analysis. Corey Cook for obtaining quotes; securing funding, selecting a study partner; supporting the study protocol and methodology. Christophe Van Laethem for his supervision throughout clinical study. Sharath Kandhi for preparing the first draft of manuscript.

Funding

The funding of this study was provided by Becton Dickinson and Co (BD).

Authors' Contributions

KDB: writing – review & editing, conceptualization, data curation, funding acquisition, methodology, resources, supervision, validation. JS: writing – review & editing, conceptualization, data curation, funding acquisition, methodology, resources, supervision, validation. PS: writing – original draft, writing – review & editing, project administration, formal analysis, visualization. TJS: writing – review & editing, project administration, supervision, visualization.

Conflicts of Interest

The authors declare the following financial interest and personal relationships which may be considered as potential competing interests: KDB is an employee of and owns stocks in Becton, Dickinson and Company. JS reports that he is a committee member (alternate) of the AAMI standards group "HE - Human Factors Engineering Committee." He reports being an employee of and owns stock in Becton, Dickinson and Company. PS and TJS report that they are employees of Becton Dickinson. BD designed the study and interpreted the results. Delve independently conducted the study and statistical analysis of the data.

Multimedia Appendix 1
Simulated room set-up and Illustration of BD PosiFlush SafeScrub Prefilled Saline Syringe scrubbing technique.
[PPTX File , 610 KB - humanfactors_v12i1e78967_app1.pptx]

References

1. Nickel B, Gorski L, Kleidon T, Kyes A, DeVries M, Keogh S, et al. Infusion Therapy Standards of Practice, 9th Edition. J Infus Nurs 2024;47(1S Suppl 1):S1-S285. [doi: 10.1097/NAN.0000000000000532] [Medline: 38211609]
2. Pittiruti M, Van Boxtel T, Scoppettuolo G, Carr P, Konstantinou E, Ortiz Miluy G, et al. European recommendations on the proper indication and use of peripheral venous access devices (the ERPIUP consensus): A WoCoVA project. J Vasc Access 2023 Jan;24(1):165-182. [doi: 10.1177/11297298211023274] [Medline: 34088239]
3. Hasan B, Bechenati D, Bethel HM, Cho S, Rajjoub NS, Murad ST, et al. A Systematic Review of Length of Stay Linked to Hospital-Acquired Falls, Pressure Ulcers, Central Line-Associated Bloodstream Infections, and Surgical Site Infections. Mayo Clin Proc Innov Qual Outcomes 2025 Jun;9(3):100607 [FREE Full text] [doi: 10.1016/j.mayocpiqo.2025.100607] [Medline: 40256621]
4. Buetti N, Marschall J, Drees M, Fakih MG, Hadaway L, Maragakis LL, et al. Strategies to prevent central line-associated bloodstream infections in acute-care hospitals: 2022 Update. Infect Control Hosp Epidemiol 2022 May;43(5):553-569 [FREE Full text] [doi: 10.1017/ice.2022.87] [Medline: 35437133]
5. Moureau NL, Flynn J. Disinfection of Needleless Connector Hubs: Clinical Evidence Systematic Review. Nurs Res Pract 2015;2015:796762 [FREE Full text] [doi: 10.1155/2015/796762] [Medline: 26075093]
6. Jeong NY, Kim S, Lim E, Choi N. Pragmatic Clinical Trials for Real-World Evidence: Concept and Implementation. Cardiovasc Prev Pharmacother 2020;2(3):85-98. [doi: 10.36011/cpp.2020.2.e12]
7. Ford I, Norrie J. Pragmatic trials. N Engl J Med 2016;375(5):454-463 [FREE Full text] [doi: 10.1056/NEJMra1510059] [Medline: 27518663]

XSL•FO
RenderX

8. World Medical Association. World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Participants. JAMA 2025 Jan 07;333(1):71-74. [doi: 10.1001/jama.2024.21972] [Medline: 39425955]
9. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
10. Samanta P, Soni KD, Parveen R, Thakur AK, Chandra Das B, Tluanpuii V, et al. A quality improvement initiative to reduce central line-associated bloodstream infection rates in the polytrauma intensive care unit of a tertiary care trauma hospital in India. Infection Prevention in Practice 2025 Jul:100478-100478. [doi: 10.1016/j.infpip.2025.100478]
11. Gilmore NT, Metz T. Prevention of Catheter-Related Infections and Complications: A Narrative Literature Review of Vascular Care and Maintenance. Int J Vasc Med 2025;2025:1427129 [FREE Full text] [doi: 10.1155/ijvm/1427129] [Medline: 40785780]
12. Devrim I, Özkul MT, Çağlar İ, Oruç Y, Demiray N, Tahta N, et al. Central line bundle including split-septum device and single-use prefilled flushing syringes to prevent port-associated bloodstream infections: a cost and resource-utilization analysis. BMC Health Serv Res 2020 Apr 21;20(1):336 [FREE Full text] [doi: 10.1186/s12913-020-05221-6] [Medline: 32316939]
13. Fitzsimons KM, Speekman J, Senior T, Curtis K, Cochrane-Davis A, Barnes R. An observational study of the securement of central venous access devices with a subcutaneous anchor device in a paediatric population at a tertiary level hospital. J Vasc Access 2020 Nov;21(6):959-962. [doi: 10.1177/1129729820918460] [Medline: 32370686]
14. Merrill KC, Sumner S, Linford L, Taylor C, Macintosh C. Impact of universal disinfectant cap implementation on central line-associated bloodstream infections. Am J Infect Control 2014 Dec;42(12):1274-1277. [doi: 10.1016/j.ajic.2014.09.008] [Medline: 25465256]

Abbreviations

ANTT: aseptic nontouch technique
CLABSI: central-line-associated bloodstream infection
CRBSI: catheter-related bloodstream infection
ICU: intensive care unit
IFU: instructions for use
NAD: needleless access device
PICC: peripherally inserted central catheter
PIVC: peripheral intravenous catheter
VAD: vascular access device

Edited by A Kushniruk; submitted 18.06.25; peer-reviewed by J Llego, F Al Dhabbari, M Al Zoubi; comments to author 27.07.25; revised version received 20.09.25; accepted 20.11.25; published 15.12.25.

Please cite as:
De Boer K, Simpson J, Sharma P, Steele TJ
Enhancing Safety in Clinical Practice: Cross-Sectional, Within-Subjects, Simulated-Use Compliance Study Using BD PosiFlush SafeScrub
JMIR Hum Factors 2025;12:e78967
URL: https://humanfactors.jmir.org/2025/1/e78967
doi:10.2196/78967
PMID:41269868

©Kyle De Boer, John Simpson, Purnima Sharma, Trevor J Steele. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 15.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Optimizing SloMo, a Digitally Supported Therapy Targeting Paranoia, for Implementation: Inclusive, Human-Centered Design Study

Thomas Gant[1,2], BSc, MSc; Kathryn M Taylor[1,2], BSc, DClinPsy; Thomas Ward[1,2], MSci, DClinPsy, PhD; Philippa Garety[1], BA (Hons), MPhil, MA, CPsychol, PhD; Amy Hardy[1,2], BSc (Hons), DClinPsy, PhD

[1]Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom
[2]South London and Maudsley NHS Foundation Trust, London, United Kingdom

Corresponding Author:
Thomas Gant, BSc, MSc
Department of Psychology
Institute of Psychiatry, Psychology and Neuroscience
King's College London
De Crespigny Park
London, SE5 8AF
United Kingdom
Phone: 44 20 7848 0002
Email: thomas.1.gant@kcl.ac.uk

Abstract

Background: Despite the promise of digital therapeutics in providing scalable interventions for psychosis, translating them from clinical trials to routine care is challenging. SloMo is an evidence-based, digitally supported cognitive behavioral therapy for psychosis comprising a web-based therapy platform and mobile app. The therapy encourages individuals to slow down for a moment, to address fast-thinking habits fueling paranoia. SloMo has received a NICE Early Value Assessment recommendation for use in the National Health Service to address evidence gaps related to its use in the real world, and an implementation study is underway.

Objective: This study aimed to optimize the SloMo software for implementation by addressing limitations of the first release, reducing technology complexity, and improving user experience, to increase equitable outcomes.

Methods: An inclusive, human-centered design methodology was used to optimize SloMo. The redesign sought to reduce the technology's complexity and improve the user experience for diverse patients and therapists. The Double Diamond framework structured the iterative redesign, integrating insights from patient and public involvement consultants, therapists, and a transdisciplinary co-design team. The Double Diamond process was facilitated through 24 transdisciplinary workshops. These were supported by the following methods: identifying implementation barriers through desk research of SloMo's evidence and qualitative interviews with experts by experience (n=2); redefining user needs; iteratively developing solutions through user testing sessions with service user consultants (n=32); and validating the minimum viable product through think-aloud testing sessions with therapist (n=10) and service user (n=11) consultants.

Results: Users wanted a form of cognitive behavioral therapy for psychosis that was usable, trustworthy, enjoyable, personalized, normalizing, and memorable. The redesign, therefore, included a minimalist user interface, more diverse lived experience vignette content, enhanced agency over data, greater representation of protected characteristics and their intersectionality, and intuitive navigation. Feedback from a purposively sampled patient and therapist sample validated the redesign as it was associated with a strong user experience, particularly in relation to usability and usefulness.

Conclusions: The study produced a redesign of the SloMo software optimized for real-world use, whilst retaining fidelity to the therapeutic content of the previous version. Through an inclusive, human-centered approach, the optimized design of SloMo addresses barriers to adoption by reducing complexity and fostering accessibility. This study underscores the value of integrating lived experience involvement into digital therapeutics to support scalable, equitable, and sustainable mental health care solutions.

(JMIR Hum Factors 2025;12:e75377) doi:10.2196/75377

XSL•FO
RenderX

KEYWORDS

co-design; design thinking; digital health; user-centered design; participatory design; mHealth; eHealth; schizophrenia; psychosis; user experience; engagement; therapy

Introduction

Digital therapeutics present an opportunity for providing scalable, usable, and targeted evidence-based interventions for people with psychosis. Emerging evidence suggests digital technologies are acceptable and effective at addressing a range of treatment targets in this population [1,2]. With device ownership and levels of technology confidence on the increase, investment and interest in digital therapeutics for psychosis are growing. However, despite their promise, implementing digital therapeutics within real-world clinical settings following clinical trials is difficult, and a notable evidence-practice gap exists [3].

Successful implementation of digital technology in psychosis is multifactorial, with technology complexity being a particularly important barrier [4]. People from minoritized backgrounds are more likely to experience additional barriers to technology access and digital literacy, and if this is not given careful consideration, there is the risk that technology might inadvertently magnify health inequalities [5]. The National Health Service (NHS) Inclusive Digital Healthcare Framework [6] lists accessibility and ease of technology use as a key domain of digital inclusion. Inclusive, human-centered design (iHCD) is a method that seeks to ensure the design of a product aligns with the needs and capabilities of the people for whom it is intended [7]. iHCD uses ethnography to understand users and their contexts, with purposive recruitment to ensure representation across relevant user identities and characteristics (ie, gender, age, ethnicity, cognitive abilities, use of technology, and attitudes to therapy) and the inclusion of seldom-heard groups [8,9].

SloMo is a digitally supported psychological therapy targeting paranoia in psychosis. It was developed to address the limitations of conventional psychological therapies for psychosis; namely, issues in access, experience, and outcomes [10-12]. SloMo works by targeting fast thinking habits that fuel worries and supports people to slow down for a moment to find ways of feeling safer and living well. Specifically, SloMo adopts a causal-interventionist approach [13] to enhance belief flexibility (“slow thinking”) and build awareness of reasoning biases (“fast thinking”) that are characteristic of paranoia [14,15].

The technology comprises a web-based therapy platform that augments face-to-face individual therapy sessions. The platform offers interactive features such as animated psychoeducation videos, which encourage reflection, games promoting belief flexibility, and personalized content, including a psychological formulation, session summaries, and between-session task plans. Key content is synchronized with a mobile app that supports self-management in daily life. The mobile app supports users to notice worries and connect to alternative, safer thoughts through interactive features to promote slowing down, reminders of session content, and a record of between-session tasks. SloMo uses technology to support the visualization of thoughts and thinking habits. People can interact with their personalized SloMo thought bubbles, altering speed and size to reflect thinking habits and distress. SloMo tips support people to slow down and notice new information, helping to shrink fast-spinning gray, worry bubbles and grow colorful, slow-spinning helpful thoughts to promote well-being [16].

iHCD was used for the design of the first release of SloMo [16,17]. Ethnographic research insights indicated users with paranoia needed a form of cognitive behavioral therapy for psychosis (CBTp) that was usable, trustworthy, enjoyable, personalized, normalizing, and memorable, and the design solution was iteratively developed with them. The therapy design and usability of SloMo was validated by a comprehensive mixed methods approach [15,18-20]. Notably, although baseline technology use and confidence were lower in groups known to be disproportionately affected by the digital divide (Black people and older adults [21]), these baseline characteristics were not associated with differences in outcomes or user experience. This suggests the SloMo therapy design may bridge the “digital divide” and support accessibility, as intended.

In a multi-site randomized controlled trial (RCT) (N=362), SloMo demonstrated improved paranoia, self-concept, and well-being outcomes over 6 months compared with treatment as usual, with small to moderate effects [18]. SloMo received a NICE Early Value Assessment recommendation [22] for use in the NHS whilst real world data is collected to address identified evidence gaps, and accordingly is now being tested in a type II hybrid implementation-effectiveness study (ClinicalTrials.gov identifier: NCT06568081). McGinty et al [23] recommend that the evidence-practice gap be approached with a complexity science lens, including adaptation of the intervention to the implementation context. Consistent with this, we sought to optimize the SloMo therapy software for routine care prior to commencing our implementation-effectiveness study. The Non-adoption, Abandonment, Scale-up, Spread, Sustainability (NASSS) framework, an evidence-based implementation science framework for technologies, informed the work [24,25]. This mixed methods approach encourages the minimization of complexity in key domains (target problem, technology, adopters, organization, and broader system) to reduce the likelihood of nonadoption and abandonment, and increase the likelihood of scale-up, spread, and sustainability. The NASSS model has been used to plan and evaluate the implementation of SloMo and was chosen as it is an implementation science framework developed specifically for health care technologies. This paper reports on work focused on reducing complexity in the technology domain. Research to address other NASSS domains relevant to SloMo’s implementation will be reported separately. Intervention complexity has been identified as a key barrier to the implementation of digital technologies in psychosis and, therefore, warrants attention [4]. It is crucial we design digital technology to meet the needs of the end user, particularly in the

XSL•FO
RenderX

context of paranoia, where concerns about privacy and technology are often elevated [26].

In summary, the study aimed to optimize the SloMo software for implementation in routine care by addressing limitations of the first release, reducing technology complexity, and improving the user experience for service users and therapists, relative to the first version of the software. We anticipated that through iHCD methods, the likelihood of successful implementation would be enhanced. This will be evaluated further in an ongoing implementation-effectiveness study.

Methods

Design

Consistent with the design of the first version of SloMo [16,17] the Double Diamond [8], a human-centered design framework, was used to optimize SloMo for implementation, using inclusive design principles [9]. The Double Diamond is the most widely used co-design framework in the United Kingdom. This framework was selected because of its alignment with inclusive design principles and because it was used to develop the previous release of SloMo [16,17].

Ethical Considerations

This work was undertaken as a quality improvement project within South London and Maudsley NHS Foundation Trust and, according to institutional policy, did not require review or approval by a research ethics board. Participants provided verbal consent to participate in user-testing sessions and focus groups. Participation was voluntary, and all data were collected anonymously. No identifiable personal or health information was recorded, stored, or analyzed. The project was carried out in compliance with relevant ethical standards for quality improvement initiatives. Participants received gift vouchers for their contributions in line with National Institute for Health and Care Research payment rates [27].

Software

SloMo (release 1; R1) was self-certified as a class I medical device. We anticipate SloMo release 2 (R2) to be a Class I medical device under Rule 12 (All other active devices; MEDDEV Guideline 2.4/1 Rev 9), as the criteria for Rules 9, 10, and 11 for active devices are not met. The software was programmed by Bitam Ltd, which met ISO (International Organization for Standardization) 13485 and ISO14971 standards (ISO13485 is an international standard for medical device quality management systems; ISO14971 is an international standard for risk management of medical devices). The SloMo therapy platform architecture is compliant with ISO/IEC 27001:2022, 27001:201, ISO/IEC 27017:2015 and 27018:201.

Sample

To address the risk inherent in participatory design that the most willing, able, and vocal users are more likely to be involved, neglecting the needs of underserved groups, we purposively sampled people from a wide range of backgrounds (ie, across diverse gender, age, ethnicity, technology confidence, and attitudes to therapy). The group was representative of service users across the 3 NHS Trusts, with respect to ethnicity and age (Table 1). Confidence in using technology was self-reported by service user consultants at the start of testing sessions on a scale of “0” (not confident at all) to “100” (very confident). A wide range of technology confidence was self-reported, skewed towards people with higher levels. This is in line with recent literature indicating digital literacy is improving amongst people with psychosis [28].

Therapist patient and public involvement (PPI) consultants currently working with people with psychosis were purposively sampled in relation to gender, age, ethnicity, and experience of delivering psychological interventions for psychosis. Ten therapist PPI consultants participated in the deliver phase. The mean age was 34.9 years (range 26-54) and comprised 7 women and 3 men. One therapist identified as Bengali, 2 as Black British-African, 2 as Black British-Caribbean, 4 as White British, and 1 as White (other). Seven therapists were clinical psychologists or counseling psychologists, one was a high-intensity CBT therapist, two were trainee clinical psychologists, and one was an assistant psychologist. The mean technology confidence score was 86.2 out of 100 (range 70 to 100).

XSL•FO
RenderX

Table 1. Demographic characteristics of service user consultants for the develop and deliver phases (N=32).

Characteristics	Value
Gender, n (%)	
Man	18 (56.3)
Woman	13 (40.6)
Nonbinary	1 (3.1)
Age (years), n (%)	
18-25	1 (3.1)
26-35	12 (37.5)
36-45	5 (15.6)
46-60	11 (34.4)
60+	3 (9.4)
Ethnicity, n (%)	
Asian or Asian British	3 (9.4)
Black British, Caribbean or African	10 (31.3)
Mixed or multiple ethnic groups	3 (9.4)
White British	15 (46.9)
Any other White background	1 (3.1)
Technology confidence (0-100), mean (range)	82 (10-100)

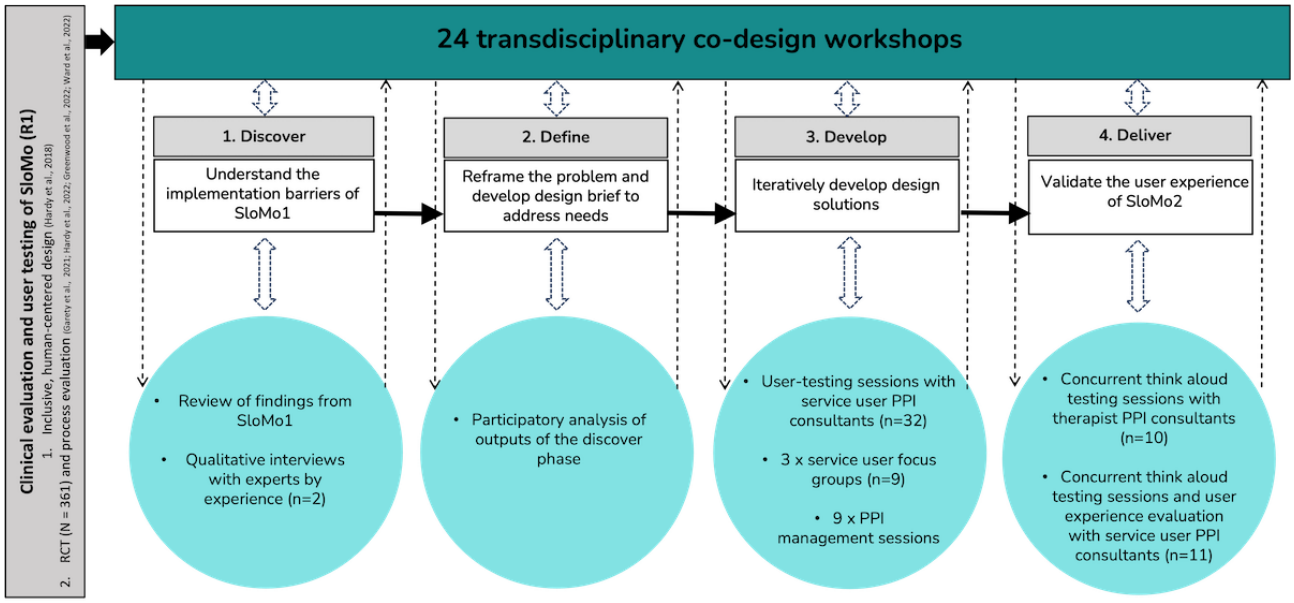
Procedure

Transdisciplinary co-design workshops (referred to as "workshops" here forward) were central to the redesign of the new version of SloMo. Our project team included experts by experience (EBEs), designers, clinicians, academics, software developers, video producers, and illustrators (Multimedia Appendix 1 contains further details of the co-design team). Workshops were initially held in person (to support group cohesion), then switched to remote (to facilitate engagement of people from geographically diverse locations). Attendees were informed of the workshop's focus in advance of the meeting, which was prespecified according to the phase of development. Sessions were recorded so that the discussion could be reviewed, and members who were unable to attend could review and offer feedback via email. A PPI management team oversaw the project.

The Double Diamond is organized into 4 convergent and divergent phases: discover (understand the problem; divergent), define (reframe the problem and develop a design brief; convergent), develop (develop potential solutions; divergent), and deliver (test out solutions at a small scale and iterate; convergent). It should be noted that divergent phases were necessarily constrained by the previous version of SloMo and its clinical evaluation, and so we adopted a "waterfall-agile" approach [16]. Figure 1 provides an overview of the Double Diamond phases used to optimize SloMo for implementation, which are explained in more detail below.

XSL•FO
RenderX

Figure 1. Optimizing SloMo for implementation—design research methods used across the 4 phases of the Double Diamond framework. PPI: patient and public involvement.



Discover

The aim of the discover phase was to develop an understanding of possible implementation barriers for the SloMo software. This first involved reviewing the outputs of the SloMo RCT [15,18-20] in the transdisciplinary co-design workshops, to explore key technology problem areas that could potentially be targeted for the next release of SloMo. In addition, the designers conducted qualitative interviews with 2 EBEs who had received the first release of SloMo. Interviews focused on exploring what was required from a psychological therapy for psychosis, building on insights from Hardy et al [17], as well as exploring how SloMo could be improved.

Define

The define phase synthesized the insights gathered in the transdisciplinary workshops. The EBEs who provided feedback during the discover phase also took part in the next round of workshops to assist in synthesis. The information from the discover phase was sent to co-design group members to review ahead of time to allow for the opportunity to familiarize and reflect on the materials. Workshops involved reviewing the insights as a group, noting patterns, and sorting into themes, using a group consensus approach. Multiple workshops were run until a clear and workable design brief was developed, in line with identified user needs.

Develop

The develop phase aimed to generate potential design solutions that helped to address the redefined user needs outlined in the preceding phase. Guided by clinician and EBE recommendations, the designers initially explored content relevant to SloMo (ie, existing apps available on the market, including for mental health, as well as apps in parallel markets, such as meditation, nutrition, and habit-building). The designers fed back their design solutions in co-design workshops, and group consensus determined which ideas were taken forward for iterative development.

Low-fidelity digital prototypes of initial design concepts were developed (Multimedia Appendix 2). Preferred concepts, selected through workshops, were taken for one-to-one user testing with service user PPI consultants. Data from the user testing sessions informed the iterative development of design solutions, leading to the development of higher fidelity digital prototypes of adopted concepts. Prototypes were further iteratively tested and developed through one-to-one service user testing, focus groups with service user and clinician consultants, and workshops. Once a prototype had been validated, it was added to the optimized SloMo minimum viable product (MVP) software, which was taken forward to the deliver phase.

Deliver

The convergent deliver phase aimed to validate the user experience of the optimized version of SloMo. Given the redesign of SloMo focused on 2 software components (a web-based therapy platform and native mobile app), the deliver phase aimed to validate both through consultation with the applicable user groups. For the web-based therapy platform, therapist PPI consultants were chosen as this would be used by therapists to augment the face-to-face therapy sessions they deliver. Service user PPI consultants were chosen for mobile app feedback, as they were the main intended users of the app.

Concurrent think aloud testing sessions were carried out with therapist PPI consultants. This is a form of usability testing where participants are asked to verbalize their thoughts in real time as they interact with a user interface (UI; ie, modules on the SloMo web-based therapy platform) [29]. Following this, additional unstructured qualitative feedback was gathered. A similar approach was taken for the mobile app. Service user PPI consultants who were involved in the develop phase took part in think-aloud testing sessions. Service users were asked to complete different core functionalities on the app (eg, adding

XSL•FO
RenderX

a new worry, slowing down a worry, viewing session summary cards) whilst thinking aloud. Interactions were observed, and qualitative feedback was elicited. Following this, a user experience survey (UES) [19] was completed. This is a 10-item questionnaire (adapted from Lobban et al [30]) with subscales relating to enjoyment, usability, and usefulness. Each item is rated from 0 to 10 and summed for the subscale total scores, which were calculated as percentages, with higher percentages indicating a more positive experience. Following user experience validation, an MVP of SloMo, optimized for implementation, was finalized for release to Apple and Android app stores for piloting clinical use.

Results

SloMo Redesign Outputs

The key insights and outputs from each phase of the Double Diamond are presented in Table 2 and expanded on below. Additional details on participants, methods, and outputs can be found in Multimedia Appendix 3 [15,17-20].

XSL•FO
RenderX

Table 2. Summary of SloMo redesign using the double diamond framework.

SloMo (R1) RCT[a] original design solution	Discover: implementation barriers to design solution from SloMo (R1)	Define: design solutions for SloMo (R2) implementation	Develop and deliver: SloMo design outputs
Usable	Poorer usability in men compared with women [15]	UI[b] that is sufficiently visually appealing	• Minimalist redesign of UI • Character aesthetics made more realistic
Usable	Lower adherence with less tech confidence [15]	App navigation should feel intuitive	• App home screen limited to key functions • Embedded onboarding guidance
Usable	Participants disliked carrying 2 phones [20]	Convenient access to mobile app	• App availability through Android and iOS app stored
Usable	Therapists struggled pacing sessions[c]	Feedback available on session progress	• Session progress bar • Burger menu to support navigation
Usable	Difficulty generating goals and safer thoughts[c]	To feel supported in generating content	• Co-produced "community" goals and safer thoughts
Trustworthy	Lack of clarity on data transfer[c]	Greater agency over data sharing	• Notifications to accept/reject sharing content • Ability to select specific content to share/not share • "Lock thoughts" function to keep entries private
Enjoyable	Worries on home screen increased distress[c]	To feel supported when opening app	• Home screen displays safer thoughts as default • Active navigation required to view worries
Personalized	In session progress too slow for some [15,20]	Therapist to tailor sessions to person	• Burger menu to tailor navigation • Optional in session and between-session tasks
Personalized	Lack of diverse content in vignettes [20]	Manage more paranoia related worries	• Co-design of novel vignettes
Normalizing	Negative connotations of "slow" in SlowMo[c]	For SloMo to feel like a positive experience	• Therapy name updated to "SloMo"
Normalizing	Lack of intersectional diversity in videos[c]	To feel represented and normalized	• Videos conveying a range of protected characteristics • Customizable avatar representing the SloMo user
Memorable	Users wanted memory aides post-therapy[c]	SloMo journey to continue after completion	• Function to print a "therapy blueprint" • Catch-up screen at start of each session • Increased repetition of key learning messages

[a]RCT: randomized controlled trial.

[b]UI: user interface.

[c]Barrier identified by patient and public involvement (PPI) consultants in transdisciplinary co-design workshops.

Discover

Key implementation barriers in the technology domain for the first software version were identified through review of the SloMo RCT outputs and interviews with EBEs. This included reporting of poorer software usability by men and people with lower levels of technology confidence. In addition, qualitative data revealed service users disliked carrying 2 mobile handsets (the first release of SloMo was only available on a study handset). Therapists reported difficulties in pacing the sessions, as there were no signifiers indicating how much session content remained, which aligned with service user feedback indicating progress was at times too slow. Therapists and service users reported finding the content of psychoeducational vignettes

XSL•FO
RenderX

limited and instead used SloMo for a greater range of paranoia-related worries than was covered in the software. Other barriers included a lack of clarity around when data was being transferred between devices (ie, the web-based therapy platform and mobile app), an overemphasis on worries instead of safety on the app home screen, difficulties in generating personalized goals and safer thoughts, a lack of diversity across psychoeducational vignettes, and forgetting content covered in therapy sessions once these had ended. Outputs at the end of the discover phase included a list of technology-related implementation barriers. These barriers were reviewed during transdisciplinary workshops and taken forward to the define phase.

Define

The define phase reaffirmed the importance of SloMo's design brief of users needing a form of CBTp that was usable, trustworthy, enjoyable, personalized, normalizing, and memorable [17]. The problems identified during the discover phase were mapped onto these domains, indicating these were areas that could be targeted in the next iteration to further deliver on this design brief and enhance user experience.

The issues outlined in the discover phase were reframed as user needs through co-design workshops. Service users needed a visually appealing UI, an intuitive mobile app navigation, and convenient access to support usability. To foster a sense of trust, service users required greater agency over data sharing to and from the web-based therapy platform and mobile app. Service users needed to feel psychologically supported when opening the mobile app, for SloMo to feel like a positive experience, and to feel represented throughout. There was a need for learning ways of managing a wider range of paranoia-related worries, and for the SloMo journey to continue once face-to-face sessions had ceased. For therapists, it was important that they were provided feedback on session progress so they could pace sessions efficiently. There was a need for therapists to be able to use the web-based therapy platform flexibly, allowing them to tailor sessions to the client. The outputs at the end of the define phase included a design brief of the solution, which specified the user needs to be addressed in relation to the original design definition, which informed the subsequent development phase.

Develop

Thirty-two 1:1 service user testing sessions and 3 service user focus groups (n=9) were carried out. Initial user testing focused on design concepts, and later stages focused on interactive digital prototypes. Initial design concepts taken to 1:1 user testing sessions included (1) aesthetics of bubbles used to convey worry and safer thoughts, (2) design of the avatar representing the SloMo user on their journey, (3) SloMo character illustrations, and (4) aesthetics of physical tip cards used as a memory aid. Other concepts (eg, lived experience vignettes, UI changes, data transfer updates) were iteratively developed in co-design workshops. Multimedia Appendix 4 displays the updated concepts for the new version of SloMo compared with the previous one used in the RCT.

High-fidelity interactive digital prototypes developed and taken out for individual user testing included (1) user interactions with thought bubbles, (2) user interactions for building a customizable avatar, (3) character illustrations featured in audio-visual psychoeducational vignettes, and (4) the content of newly developed psychoeducational vignettes. User testing feedback in both the concept and prototype development phases was discussed in co-design workshops and iteratively developed until there was a group consensus that the prototype should be adopted. Figures 2 to 4 provide an overview of the development of specific SloMo features from the discover to deliver phase. Multimedia Appendix 5 provides additional data collected during the develop phase.

Outputs at the end of the develop phase included iteratively developed digital design solutions addressing the user needs and design brief from the define phase. In addition, early-stage demonstration and interaction testing of design solutions with user testing consultants provided preliminary validation of specific design solutions. These were then developed into the SloMo (R2) MVP during the deliver phase.

XSL•FO
RenderX

Figure 2. Development of SloMo characters featuring in psychoeducational vignettes. PPI: patient and public involvement.

Discover

- Lack of intersectional diversity in characters featuring in psychoeducational vignettes.

Define

- Service users need to feel represented across a range of protected characteristics.

Develop

- Low fidelity prototypes using different illustrations developed: semiabstract (A) and semirealistic, reflecting different sociodemographic realities.



A B

- User testing indicated preference was divided. Qualitative feedback indicated a majority would have opted for B If diversity in skin tones was shown.
- Co-design workshops determined subsequent user testing should focus on different shading styles:



A B C

- There was a majority preference for C, with qualitative feedback indicating that this style promoted a normalising message, that these types of difficulties can impact anyone.
- High fidelity digital prototypes of our three SloMo characters were taken for user testing:



- Prototype testing indicated that the visual representation of the three characters worked well for a purposively sampled group of PPI consultants.
- Characters were added and taken forward to the deliver phase.

XSL•FO
RenderX

Figure 3. Co-designing psychoeducational vignettes on anomalous experiences and paranoia. EBE: experts by experience; MVP: minimum viable product; PPI: patient and public involvement.

Discover

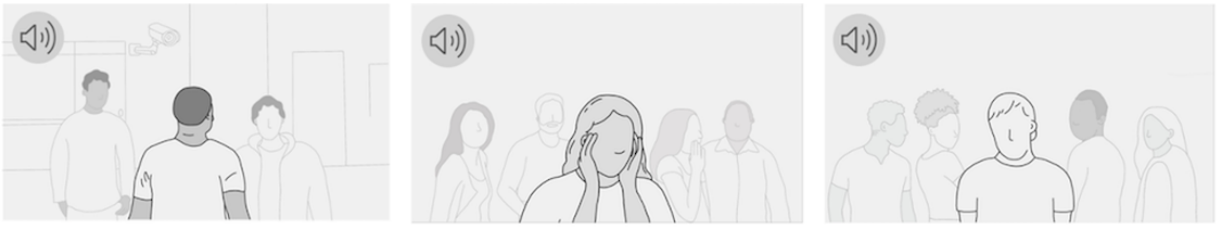
- Participants wanted more psychoeducational vignettes.

Define

- Learn ways of managing a wider range of paranoia-related worries.

Develop

- Audio recordings of initial vignettes developed in co-design workshops were taken for user testing feedback.



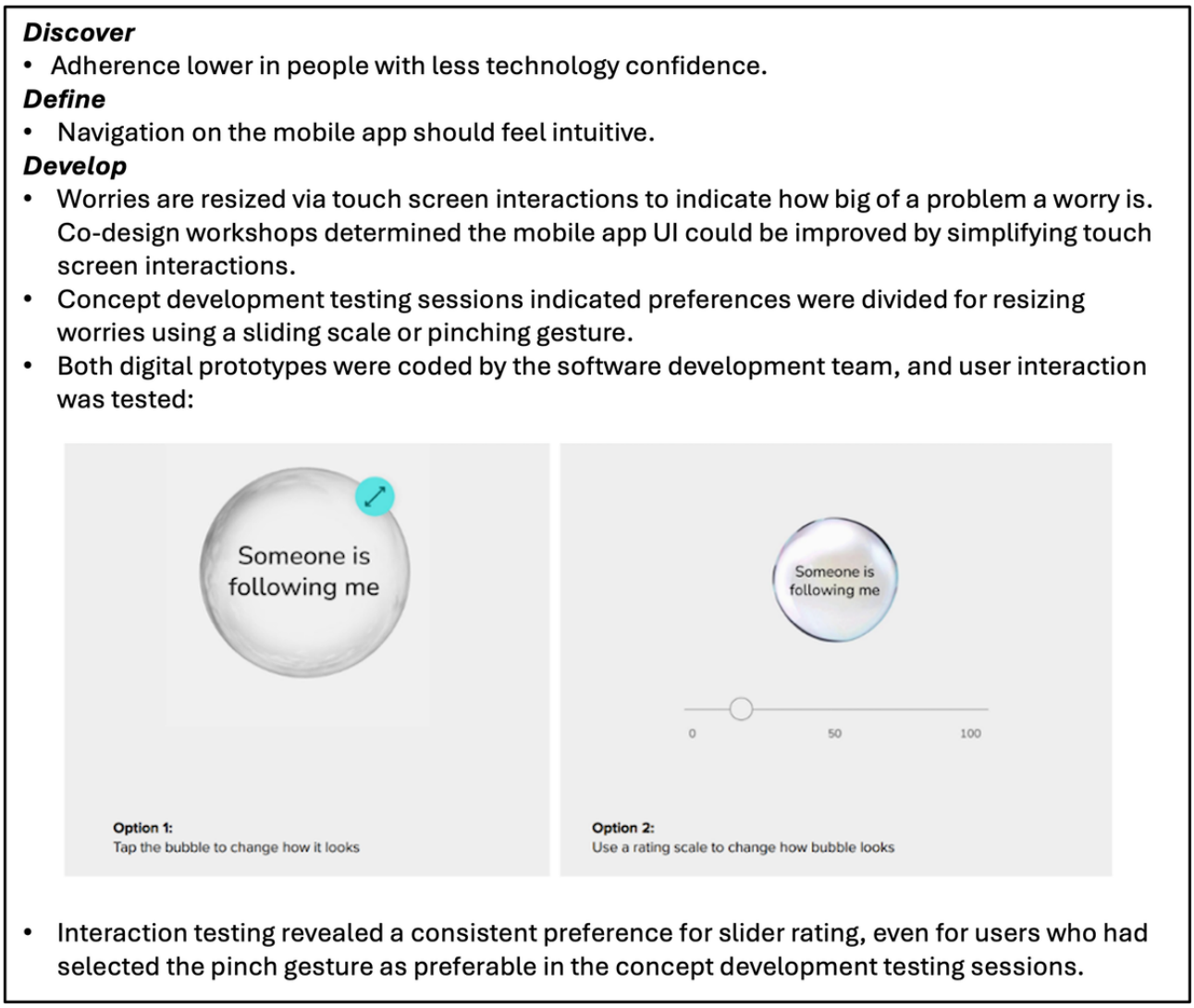
- Feedback indicated vignettes were interesting, understandable, and relatable.
- Some feedback suggested wording changes should be made to increase authenticity.
- During testing of audio recordings, several users commented that the module title (which was initially ”unusual experiences”) felt stigmatising. This was fed back in co-design workshops, where an EBE suggested “when the worlds doesn’t feel right”.
- This was taken for further user testing. Nine out of nine PPI consultants preferred this title.
- High fidelity audio-visual prototypes of vignettes adapted from previous feedback were taken out for further rounds of user testing.



- Feedback was sought on story content, illustrations, and narrators.
- The vignettes were well received and added to the optimised SloMo MVP.

XSL•FO
RenderX

Figure 4. Bubble interactions on the mobile app to indicate levels of distress.



Deliver

Table 3 presents quotes from the therapist PPI consultants during the think-aloud testing sessions. Overall, the web-based therapy platform received positive feedback. Therapists appreciated the platform's simplicity and visual appeal, believing it would enhance the enjoyment and engagement of sessions. They noted therapeutic benefits, such as providing more structure to sessions and making therapy feel more personalized. Additionally, therapists found the platform engaging and flexible, allowing for adaptation based on various therapeutic needs. The addition of progress bars was positively commented on as a helpful feature for therapists to manage the session time more effectively.

During the think-aloud sessions, several areas for improvement were identified. Some therapists expressed hesitation about navigating back and forth due to concerns that personalized content entered into the platform might be lost. Others suggested that enjoyment could be further enhanced by incorporating more gamification elements, such as badges or achievements, to boost engagement. Additionally, there were recommendations to increase inclusivity by adding different languages, enabling access for users whose first language is not English.

Eleven service user PPI consultants who had previously taken part in user testing feedback sessions in the development develop phase took part in the concurrent think-aloud testing sessions for the deliver phase. Think-aloud consultants were purposively sampled to be representative of the psychosis population across the respective areas of the 3 trusts. The sample included 7 women and 4 men, whose ages ranged from 25 to 60 years, and 4 consultants identified as Black British, 1 as South American, and the remaining 6 as White British.

The UES findings for each subscale are presented in Table 4. The ratings suggest a good user experience of the SloMo demonstration, with a total UES mean of 83.4%. Notably, the mean score for enjoyment was around 10% lower than the domains of usability and usefulness, as well as showing a greater range of scores. This suggested that improved enjoyment may be an area requiring further development.

Table 5 displays qualitative feedback captured during think-aloud user testing sessions with service user PPI consultants. The feedback aligned with the UES ratings, indicating areas across the 3 domains of user experience that were well received. Interestingly, there were suggestions around how enjoyment could be enhanced, possibly providing insight into lower UES scores for this domain. For instance, haptic

XSL•FO
RenderX

feedback was seen as a feature that could boost enjoyment, and improving touch screen interactions could further enhance usability. Outputs at the end of the develop phase included an MVP of SloMo (R2) that had received early-stage user experience validation from service users and therapist consultants. The user experience of the software will be further evaluated in an ongoing implementation-effectiveness study. Figure 5 [16,31] illustrates the prototype delivered from the develop phase.

Table 3. Quotes from the therapist consultants (n=10) during think-aloud testing.

Theme	Quote
Enjoyment	• I like the visuals… It is attractive to use • The gamified elements make it more enjoyable • This is definitely more enjoyable and accessible than conventional CBT for psychosis
Usability	• I like how simple it is… It's a nice scaffolding tool • I like the bar along the bottom to help with time management • I feel more confident to be flexible given my experience level... I worry that low intensity therapists would struggle to deliver in a flexible manner
Usefulness	• The structure scaffolds good CBT… you can be more focused on the person in the room • The formulation feels personalized • It gives such a clear focus for both therapists and client • I think it makes [the therapy process] less intense for people who struggle with conversations for an hour

Table 4. User experience survey (UES) ratings of SloMo from service user consultants (n=11).

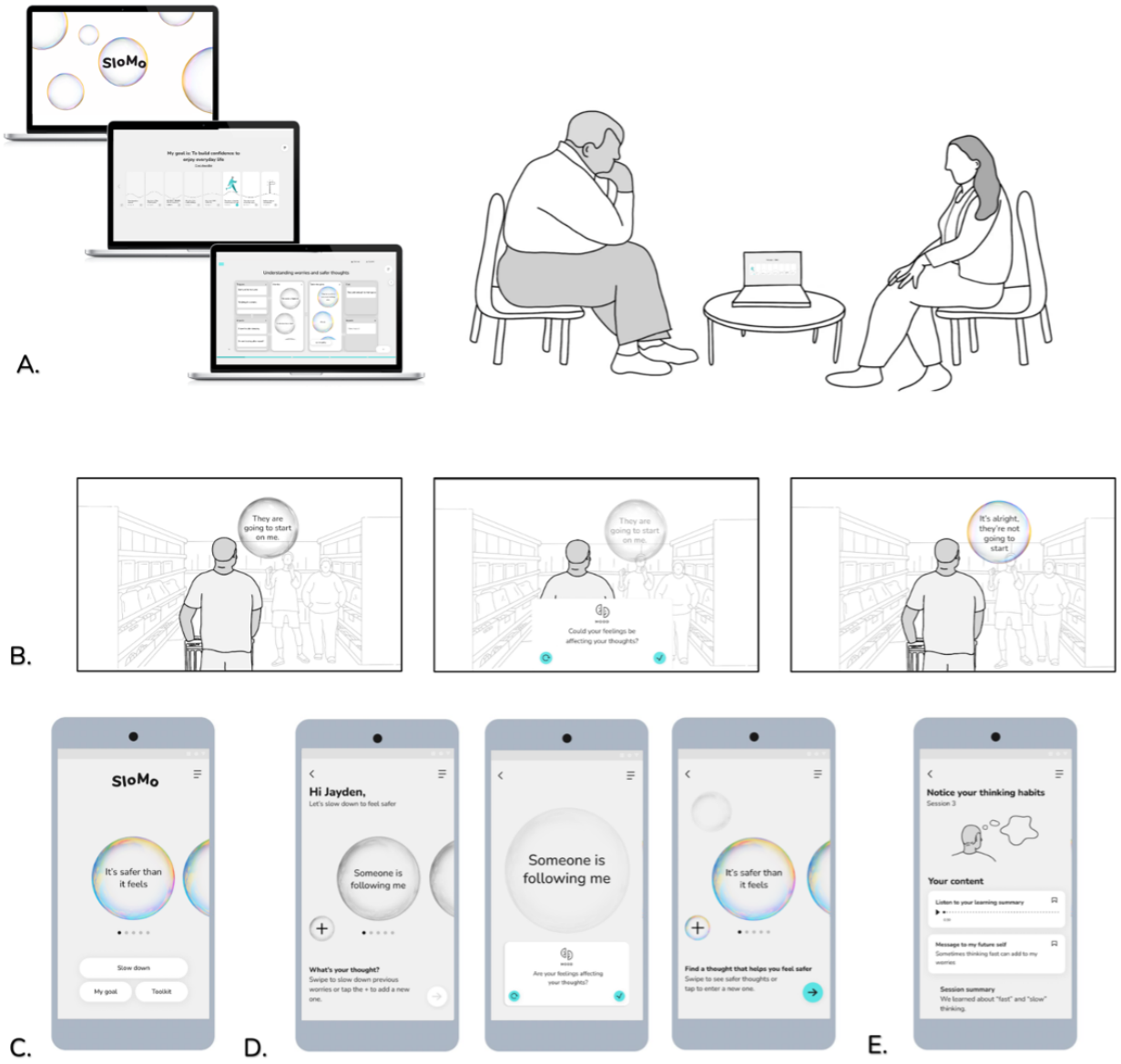
UES variable	Values
Enjoyment	
Mean % (SD)	76.1 (17.3)
Range	47.5-100.0
Usability	
Mean % (SD)	85.8 (8.5)
Range	72.5-100.0
Usefulness	
Mean % (SD)	87.0 (8.5)
Range	67.5-100.0
Total UES	
Mean % (SD)	83.4 (9.7)
Range	69.2-95.8

Table 5. Quotes from service user consultants (n=11) during think aloud testing sessions.

Theme	Quote
Enjoyment	• I really liked the safer thoughts and the colours on the positive safer thoughts • Using haptics would be good… adds another layer • The swiping [screen interaction] is difficult in places. I would find this frustrating
Usability	• Ease of access is the biggest positive • It's not difficult to add worries and sometimes people are afraid of technology, but I think it's not difficult to use this app • I feel like I can navigate around it and haven't spent much time on in… It's pretty intuitive
Usefulness/Acceptability	• I like you don't need to share all of your thoughts with the therapist …sometimes you don't want to show your thoughts to other people • I like that it's really personalised, I can't count the amount of times that I have been given something that has not been in my own words in therapy

XSL•FO
RenderX

Figure 5. SloMo sessions with a therapist (in person or online) are supported by a (A) web-based therapy platform which promotes (B) understanding and managing worries through lived experience stories and interactive tasks. Personalized thought bubbles are visualized and synchronized with the (C) SloMo mobile app, (D) then "slowed down" to find ways of feeling safer. Additional mobile app features can be accessed via a toolkit, which includes (E) personalized learning summaries of therapy sessions (reproduced from Hardy et al [16], which is published under Creative Commons Attribution 4.0 International License [31]).



Discussion

Principal Results

This study optimized the SloMo therapy software for implementation in routine NHS care, through reducing technology complexity and boosting user experience [24,25]. An iHCD was used, involving transdisciplinary co-design workshops, focus groups, one-to-one user testing, and think-aloud testing sessions. From this research, we have produced an optimized MVP of SloMo, which is now being tested in the SloMo2 hybrid implementation-effectiveness study (ClinicalTrials.gov identifier: NCT06568081). User testing indicated that users need a form of CBTp that is usable, trustworthy, enjoyable, personalized, normalizing, and memorable, in line with earlier work from our team [17]. Design targets for improving the SloMo software were identified to deliver this design brief and address user needs. For reducing complexity, key targets included simplified app interfaces to highlight core functionalities, improved navigation, and, for boosting user experience, this included providing users with greater agency over data, and improving the representation of protected characteristics, including their intersectionality. The user experience of this optimized version of SloMo may be improved relative to the first version that was tested in the SloMo RCT, indicated by higher UES scores compared with those reported by Hardy et al [19].

A distinctive aspect of our product development approach is the integration of lived experience involvement in all stages of the design work. Traditional approaches to intervention development in clinical research often follow a convergent design process using existing theory and evidence to determine intervention content and structure. This approach can miss novel

XSL•FO
RenderX

opportunities to improve the user experience of therapies. In contrast, we adopted a "waterfall-agile" approach to facilitate the co-design of SloMo as an evidence-based digital therapeutic [16]. This allowed for agility in iterating specific features (eg, UI, psychoeducational vignettes) while maintaining fidelity to the previous UI and therapy protocol [18]. To support this, we embraced divergence, revisiting and challenging earlier design assumptions in line with iHCD principles. An example of this was the service user feedback that "unusual experiences" could be perceived as stigmatizing, despite this wording being widely used in CBTp [32] and the identification of a term describing these types of experiences, which felt more relevant and acceptable—"when the world doesn't feel right." Despite this divergent approach, we were necessarily constrained in certain respects. The need to build on a preexisting evidence base meant that some design decisions were necessarily limited by prior research findings and the parameters of the therapeutic model. For instance, some users suggested moving away from visualizing thoughts and thinking habits as bubbles, which was not possible, as this is a core component of SloMo and its previous clinical evaluation.

Limitations

Resource constraints restricted our ability to optimally develop this new version of SloMo. This included time pressure, as there was a need to produce an MVP in line with our implementation-effectiveness study milestones. Another constraint was workforce access, such as access to PPI consultants (therapists and service users), designers, and software developers. These limitations meant that pragmatic decisions had to be made for the redesign, for instance, regarding the types of software changes that could be made and what features were selected for user testing. Whilst decisions were made collaboratively within the transdisciplinary co-design workshops, it is possible that different design decisions may have been made with additional resources. These constraints may be viewed as a barrier to reducing complexity in the technology domain of the NASSS framework [24,25]. Ideally, future projects could be based in "living labs," whereby lived experience, clinical, technological, industry, regulatory, and research expertise is brought together in an organizational structure, with coproduction underpinning the co-design of digital health [33].

A limitation of the PPI consultant validation is that the service user and therapist experience of the optimized version of SloMo was based on a relatively brief exposure to the software, and so may not be directly comparable to the user experience findings from the SloMo RCT. Further, the PPI consultants involved in the delivery phase had been involved in the SloMo design work for an extended period and may have been less independent than the RCT sample. This may limit the conclusions we can draw from this early-stage validation. However, more rigorous data will be provided in our ongoing implementation study, examining SloMo when delivered by frontline therapists.

Validation of the SloMo (R2) MVP with service user consultants found enjoyment was rated lower than other domains on the UES [17], indicating an aspect of user experience that warrants attention in future development. Qualitative feedback highlighted specific areas for improvement, such as incorporating haptic feedback and improving user interaction. The digital therapy engagement literature could inform additional design strategies [34]. Nonetheless, lower enjoyment scores may reflect comparisons with nontherapeutic commercial apps, where higher levels of enjoyment are likely expected. It will be important to monitor this throughout the implementation-effectiveness study, and iterate on the software as needed; for example, monitoring engagement may act as a proxy for enjoyment.

Conclusions and Future Directions

This study demonstrates the value of using iHCD to optimize DMHIs for implementation in routine mental health care. By adopting an iterative, transdisciplinary approach that prioritized collaboration with service users, therapists, and technical experts, the SloMo redesign addressed critical usability barriers, fostered inclusivity, and enhanced the therapy's ability to meet the diverse needs of its target population. This methodology offers a valuable model for future mental health innovation, where the needs of diverse user groups are addressed, and therapeutic technologies are designed to be impactful and sustainable in routine care settings. We are now evaluating the implementation and effectiveness of the optimized version of SloMo when delivered by routine care therapists. We envisage SloMo will continue to be iterated based on learning from this work, in line with the principles of implementation science [35].

Acknowledgments

The first author (TG) is in receipt of a PhD studentship funded by the National Institute for Health and Care Research (NIHR) Maudsley Biomedical Research Centre (BRC).

Funding

AH and TW acknowledge funding from the Maudsley Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The authors acknowledge therapy development funding from the Maudsley Charity, the Helen Hamlyn Trust, and PG and Professor Elizabeth Kuipers National Institute for Health and Care Research (NIHR) Senior Investigator awards. The authors acknowledge funding for the SlowMo RCT from the Efficacy and Mechanism Evaluation Programme, a Medical Research Council (MRC) and NIHR partnership Project, ref. 15/48/21. The EME Programme is funded by the MRC and NIHR, with contributions from the Chief Scientist Office in Scotland, National Institute for Social Care and Health Research in Wales, and the HSC R and D Division, Public Health Agency, in Northern Ireland. The views expressed in this publication

XSL•FO
RenderX

are those of the authors and not necessarily those of the MRC, NHS, NIHR, or the Department of Health and Social Care. This work was supported by the Wellcome Trust (109,586/Z/15/Z) and (225,816/Z/22/Z).

Conflicts of Interest

AH, PG, and TW are co-founders of SloMo therapy.

Multimedia Appendix 1
Codesign team involved across Double Diamond phases.
[DOCX File , 28 KB - humanfactors_v12i1e75377_app1.docx]

Multimedia Appendix 2
Digital prototypes used in user testing sessions during develop phase.
[DOCX File , 282 KB - humanfactors_v12i1e75377_app2.docx]

Multimedia Appendix 3
Participants, methods, and outputs for subphases of the Double Diamond redesign process.
[DOCX File , 29 KB - humanfactors_v12i1e75377_app3.docx]

Multimedia Appendix 4
Comparison of design outputs from the develop phase, compared to SloMo (release 1).
[DOCX File , 826 KB - humanfactors_v12i1e75377_app4.docx]

Multimedia Appendix 5
Examples of qualitative feedback from service user PPI consultants during the develop phase.
[DOCX File , 19 KB - humanfactors_v12i1e75377_app5.docx]

References

1. Morales-Pillado C, Fernández-Castilla B, Sánchez-Gutiérrez T, González-Fraile E, Barbeito S, Calvo A. Efficacy of technology-based interventions in psychosis: a systematic review and network meta-analysis. Psychol Med 2023;53(13):6304-6315 [FREE Full text] [doi: 10.1017/S0033291722003610] [Medline: 36472150]
2. Clarke S, Hanna D, Mulholland C, Shannon C, Urquhart C. A systematic review and meta-analysis of digital health technologies effects on psychotic symptoms in adults with psychosis. Psychosis 2019;11(4):362-373. [doi: 10.1080/17522439.2019.1632376]
3. Theobald S, Brandes N, Gyapong M, El-Saharty S, Proctor E, Diaz T, et al. Implementation research: new imperatives and opportunities in global health. Lancet 2018;392(10160):2214-2228. [doi: 10.1016/S0140-6736(18)32205-0] [Medline: 30314860]
4. Aref-Adib G, McCloud T, Ross J, O'Hanlon P, Appleton V, Rowe S, et al. Factors affecting implementation of digital health interventions for people with psychosis or bipolar disorder, and their family and friends: a systematic review. Lancet Psychiatry 2019;6(3):257-266. [doi: 10.1016/S2215-0366(18)30302-X] [Medline: 30522979]
5. Bond RR, Mulvenna MD, Potts C, O'Neill S, Ennis E, Torous J. Digital transformation of mental health services. Npj Ment Health Res 2023;2(1):13 [FREE Full text] [doi: 10.1038/s44184-023-00033-y] [Medline: 38609479]
6. Inclusive digital healthcare: a framework for NHS action on digital inclusion. NHS England. URL: https://www.england.nhs.uk/long-read/inclusive-digital-healthcare-a-framework-for-nhs-action-on-digital-inclusion/ [accessed 2025-03-03]
7. The design of everyday things - Don Norman.pdf. Google Docs. URL: https://drive.google.com/file/d/0B5tR1YhNBlD2dVBKSlRSSGM4bUE/view?usp=sharing&usp=embed_facebook [accessed 2025-03-03]
8. Framework for innovation. Design Council. URL: https://www.designcouncil.org.uk/our-resources/framework-for-innovation/ [accessed 2025-03-03]
9. John Clarkson P, Coleman R. History of inclusive design in the UK. Appl Ergon 2015;46 Pt B:235-247. [doi: 10.1016/j.apergo.2013.03.002] [Medline: 23570838]
10. Jenkins P. Schizophrenia commission report. Ment Health Today 2013:7. [Medline: 24354212]
11. Vial S, Boudhraâ S, Dumont M. Human-centered design approaches in digital mental health interventions: exploratory mapping review. JMIR Ment Health 2022;9(6):e35591 [FREE Full text] [doi: 10.2196/35591] [Medline: 35671081]
12. Morris RM, Sellwood W, Edge D, Colling C, Stewart R, Cupitt C, et al. Ethnicity and impact on the receipt of cognitive-behavioural therapy in people with psychosis or bipolar disorder: an English cohort study. BMJ Open 2020;10(12):e034913 [FREE Full text] [doi: 10.1136/bmjopen-2019-034913] [Medline: 33323425]
13. Kendler KS, Campbell J. Interventionist causal models in psychiatry: repositioning the mind-body problem. Psychol Med 2009;39(6):881-887. [doi: 10.1017/S0033291708004467] [Medline: 18845010]

XSL•FO
RenderX

14. Ward T, Garety PA. Fast and slow thinking in distressing delusions: a review of the literature and implications for targeted therapy. Schizophr Res 2019;203:80-87 [FREE Full text] [doi: 10.1016/j.schres.2017.08.045] [Medline: 28927863]
15. Ward T, Hardy A, Holm R, Collett N, Rus-Calafell M, Sacadura C, et al. SlowMo therapy, a new digital blended therapy for fear of harm from others: an account of therapy personalisation within a targeted intervention. Psychol Psychother 2022;95(2):423-446 [FREE Full text] [doi: 10.1111/papt.12377] [Medline: 35019210]
16. Hardy A, Taylor KM, Grant A, Christie L, Walsh L, Gant T, et al. Co-designing technology to improve psychological therapy for psychosis: SloMo, a blended digital therapy for fear of harm from others. Schizophr Res 2024;274:526-534 [FREE Full text] [doi: 10.1016/j.schres.2024.11.004] [Medline: 39579660]
17. Hardy A, Wojdecka A, West J, Matthews E, Golby C, Ward T, et al. How inclusive, user-centered design research can improve psychological therapies for psychosis: development of SlowMo. JMIR Ment Health 2018;5(4):e11222 [FREE Full text] [doi: 10.2196/11222] [Medline: 30518514]
18. Garety P, Ward T, Emsley R, Greenwood K, Freeman D, Fowler D, et al. Digitally supported CBT to reduce paranoia and improve reasoning for people with schizophrenia-spectrum psychosis: the SlowMo RCT. Effic Mech Eval 2021;8(11):1-90. [Medline: 34398537]
19. Hardy A, Ward T, Emsley R, Greenwood K, Freeman D, Fowler D, et al. Bridging the digital divide in psychological therapies: observational study of engagement with the SlowMo mobile app for paranoia in psychosis. JMIR Hum Factors 2022;9(3):e29725 [FREE Full text] [doi: 10.2196/29725] [Medline: 35776506]
20. Greenwood KE, Gurnani M, Ward T, Vogel E, Vella C, McGourty A, SlowMo Patient, Public Involvement (PPI) team. The service user experience of SlowMo therapy: a co-produced thematic analysis of service users' subjective experience. Psychol Psychother 2022;95(3):680-700 [FREE Full text] [doi: 10.1111/papt.12393] [Medline: 35445520]
21. Robotham D, Satkunanathan S, Doughty L, Wykes T. Do we still have a digital divide in mental health? a five-year survey follow-up. J Med Internet Res 2016;18(11):e309 [FREE Full text] [doi: 10.2196/jmir.6511] [Medline: 27876684]
22. Digital health technologies to help manage symptoms of psychosis and prevent relapse in adults and young people: early value assessment. NICE. 2024. URL: https://www.nice.org.uk/guidance/HTE17/chapter/1-Recommendations [accessed 2025-03-03]
23. McGinty EE, Alegria M, Beidas RS, Braithwaite J, Kola L, Leslie DL, et al. The Lancet Psychiatry Commission: transforming mental health implementation research. Lancet Psychiatry 2024;11(5):368-396. [doi: 10.1016/S2215-0366(24)00040-3] [Medline: 38552663]
24. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
25. Greenhalgh T, Abimbola S. The NASSS framework - a synthesis of multiple theories of technology implementation. Stud Health Technol Inform 2019;263:193-204. [doi: 10.3233/SHTI190123] [Medline: 31411163]
26. Rajapakse T, Garcia-Rosales A, Weerawardene S, Cotton S, Fraser R. Themes of delusions and hallucinations in first-episode psychosis. Early Interv Psychiatry 2011:254-258 [FREE Full text] [doi: 10.1111/j.1751-7893.2011.00281.x]
27. Payment guidance for researchers and professionals. Learning for Involvement. URL: https://www.learningforinvolvement.org.uk/content/resource/nihr-payment-guidance-for-researchers-and-professionals/ [accessed 2025-12-06]
28. Eisner E, Berry N, Bucci S. Digital tools to support mental health: a survey study in psychosis. BMC Psychiatry 2023;23(1):726 [FREE Full text] [doi: 10.1186/s12888-023-05114-y] [Medline: 37803367]
29. Nielsen J. Usability Engineering 1st Edition. San Francisco, California: Morgan Kaufmann; 1994.
30. Lobban F, Dodd A, Dagnan D, Diggle P, Griffiths M, Hollingsworth B, et al. Feasibility and acceptability of web-based enhanced relapse prevention for bipolar disorder (ERPonline): trial protocol. Contemp Clin Trials 2015;41:100-109 [FREE Full text] [doi: 10.1016/j.cct.2015.01.004] [Medline: 25602581]
31. Attribution 4.0 International (CC BY 4.0). Creative Commons. URL: https://creativecommons.org/licenses/by/4.0/ [accessed 2025-12-07]
32. Dodgson G, Aynsworth C, Mitrenga KJ, Gibbs C, Patton V, Fernyhough C, et al. Managing unusual sensory experiences: a feasibility trial in an at risk mental states for psychosis group. Psychol Psychother 2021;94(3):481-503 [FREE Full text] [doi: 10.1111/papt.12323] [Medline: 33320425]
33. Mangyoku M, Vérilhac I, Pallot M, Aragall F, Serra A, Sanchez L. IDeALL: Investigating design-for-all and living-lab methods for engaging users in value co-creation. 2014 Presented at: International Conference on Engineering, Technology and Innovation (ICE); 2014 June 23-25; Bergamo, Italy p. 1-8.
34. Eisner E, Faulkner S, Allan S, Ball H, Di Basilio D, Nicholas J, et al. Barriers and facilitators of user engagement with digital mental health interventions for people with psychosis or bipolar disorder: systematic review and best-fit framework synthesis. JMIR Ment Health 2025;12:e65246 [FREE Full text] [doi: 10.2196/65246] [Medline: 39832352]
35. Bauer MS, Kirchner J. Implementation science: what is it and why should I care? Psychiatry Res 2020;283:112376 [FREE Full text] [doi: 10.1016/j.psychres.2019.04.025] [Medline: 31036287]

XSL•FO
RenderX

Abbreviations

CBTp: cognitive behavioral therapy for psychosis
EBE: experts by experience
iHCD: inclusive, human-centered design
ISO: International Organization for Standardization
MVP: minimum viable product
NASSS: Non-adoption, Abandonment, Scale-up, Spread, Sustainability
NHS: National Health Service
PPI: patient and public involvement
RCT: randomized controlled trial
UES: user experience survey
UI: user interface

Edited by A Kushniruk; submitted 02.04.25; peer-reviewed by C Rauschenberg, B Wolfe-Piggott; comments to author 26.06.25; revised version received 26.09.25; accepted 20.11.25; published 22.12.25.

Please cite as:
Gant T, Taylor KM, Ward T, Garety P, Hardy A
Optimizing SloMo, a Digitally Supported Therapy Targeting Paranoia, for Implementation: Inclusive, Human-Centered Design Study
JMIR Hum Factors 2025;12:e75377
URL: https://humanfactors.jmir.org/2025/1/e75377
doi:10.2196/75377
PMID:

©Thomas Gant, Kathryn M Taylor, Thomas Ward, Philippa Garety, Amy Hardy. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 22.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Patient-Derived Design Principles for Technology-Enabled Healing at Home Following Hospital Discharge: Mixed Methods Study

Lindsey M Philpot, PhD; Abhinav Singla, MD; Sagar B Dugani, PhD; Rachel E Canning, BS; Christina M Smith, BUS; Meredith A DeZutter, BFA; Priya Ramar, MPH; Jennifer M P Hovell, MHA; Jon O Ebbert, MD, MSc

Department of Medicine, Mayo Clinic, 200 First Street SW, Rochester, MN, United States

Corresponding Author:
Lindsey M Philpot, PhD
Department of Medicine, Mayo Clinic, 200 First Street SW, Rochester, MN, United States

Abstract

Background: As more patients transition from hospital to home for postacute care, a growing interest exists in leveraging technology to support recovery, yet a limited understanding exists on how to design these tools to align with patient and caregiver needs and preferences.

Objective: This study aimed to explore the perceptions, attitudes, and beliefs of recently discharged patients to develop user-centered design principles for digital tools that support safe and effective transitions from hospital to home.

Methods: A vignette-based, mixed methods survey was conducted, grounded in the technology acceptance model, to explore patient perceptions of digital tools supporting postdischarge care. A random sample of 1000 recently discharged adult patients received a survey featuring validated vignettes and technology acceptance model–informed questions, with both quantitative and qualitative items. Open-ended responses were analyzed using grounded theory to derive design principles that inform the development and implementation of patient-centered digital health tools. Quantitative items were descriptive and are summarized as counts (n) and frequencies (%).

Results: Of the 967 eligible patients contacted, 116 completed the survey (response rate of 2%), with respondents having a median age of 71 (IQR 61-78) years, high rates of chronic illness, and access to smartphones (98/116, 84.5%) and in-home internet (111/116, 95.7%). Qualitative analysis revealed 6 key themes—connection to care, technical ease of use, solution usability, human connection, cost, and privacy—informing 3 patient-centered design principles focused on user experience, affordability, and transparent communication to guide future technology-supported hospital discharge interventions. Respondents reported the following factors as highly important: reassurance that a care team member would reach out if something seemed wrong (107/116, 92.2%), responsiveness to patient needs (95/116, 81.9%), ability to see their data (95/116, 81.9%), and out-of-pocket cost (94/116, 81.0%). Less important factors included duration of device use (22/116, 19.0%) and battery life (21/116, 18.0%).

Conclusions: Grounded in patient perspectives, this study identified the 3 core design principles of user experience and accessibility, cost and privacy, and communication and transparency that should guide the development and implementation of digital tools to support safe, effective, and human-centered transitions from hospital to home.

(JMIR Hum Factors 2025;12:e72913) doi:10.2196/72913

KEYWORDS

hospital discharge; technology; user-centered design; patient transitions; digital health; remote patient monitoring; ecological momentary assessments; accelerometers

Introduction

The quality of postacute care impacts both short- and long-term outcomes for patients discharging from the hospital [1,2]. In the United States, 2 in 5 hospital discharges of Medicare patients result in admission to a postacute care facility [1], despite evidence that most patients and caregivers have a strong preference for home-based postacute care [3]. In addition to patient preference for healing at home, limitations in hospital bed availability, restricted access to postacute care facilities, the high cost of care within these settings, and endorsement for in-home programs such as Medicare's Acute Hospital at Home have shifted rehabilitation and recovery to patient homes [4,5].

Although most patients feel relief returning home after hospital discharge, this relief is often accompanied by anxiety and confusion about managing recovery [6]. Telephone-based postdischarge outreach by primary care is recognized as a best practice for safe transitions of care and is supported by Medicare's transitional care management services, but can be limited due to the operational burden associated with nursing time and lack of longitudinal contact through the recovery period [7]. Technology can help facilitate patient transitions from

XSL•FO
RenderX

hospital care to home by addressing challenges experienced postdischarge, including continuity of care, patient adherence to therapy plans, patient and caregiver engagement, and timely escalation of needs to prevent adverse events or complications. Automated communications via standardized text messaging (eg, ecological momentary assessments) are the most widely reported upon technology, which can be used to schedule follow-up appointments, encourage treatment adherence, and monitor symptoms [7-9]. In addition to SMS-based text messaging, other active technology support for patient transitions from hospital settings to home includes remote monitoring of conditions such as acute kidney injury [10] and heart failure [11]. In these models, patients transfer daily vitals and patient-reported outcomes back to care teams, while interactive care plans guide them through discrete tasks to help stabilize a new or evolving medical condition [12,13]. Passive technology-based support, which is enabled without direct input from the patient or caregiver, is also emerging as an option to support patients transitioning from hospital care to home, but is less studied. Accelerometers and GPSs are common elements found in modern cellular phones and smartwatches and hold the potential to detect patient movement patterns, assessing for changes in physical activity levels, gait, or falls as proxies for recovery [14-16]. These passive modes of technology support are nascent and have only been studied in specific populations [14,16]. A gap exists in the scientific literature about how to design technologies to optimize patient and caregiver acceptability [9]. This will be increasingly important as intelligent monitoring systems based on Internet of Things technologies are being developed and proposed as possible health care solutions [17,18].

Hybrid models of technological connection between patients and care team members continue to evolve and be shared in the academic literature, while additional insights into implementation and scaling of successful interventions, and opportunities to address digital disparities experienced by patients, continue to emerge [7,19]. User-centered design (UCD) approaches can optimize the human experience with technologies, facilitating increased acceptance and uptake of new tools and thereby supporting implementation and scaling of successful interventions [20]. UCD focuses on a deep understanding of the needs, preferences, and barriers experienced by end users of design tools, services, and interventions addressing the needs of people [21]. Although major funding organizations and patient advocacy groups are encouraging participatory design for digital health interventions, the principles and approaches leveraging UCD relating to technology support for transitions from hospital care to home are uncommonly described in the academic literature. Design thinking can be deployed to create design principles in the build of new tools and services grounded in a patient-centered format. These design principles are critical constructs to inform product design and build, as well as implementation strategies to support adoption and sustainability.

The objective of our study was to elicit the perceptions, attitudes, and beliefs of patients recently discharged from the hospital to create design principles to enable the design and implementation of digital tools. We present our methodology, results of our vignette-based qualitative inquiry and survey-based quantitative findings, and propose a set of design principles to be used in technology design and clinical implementation.

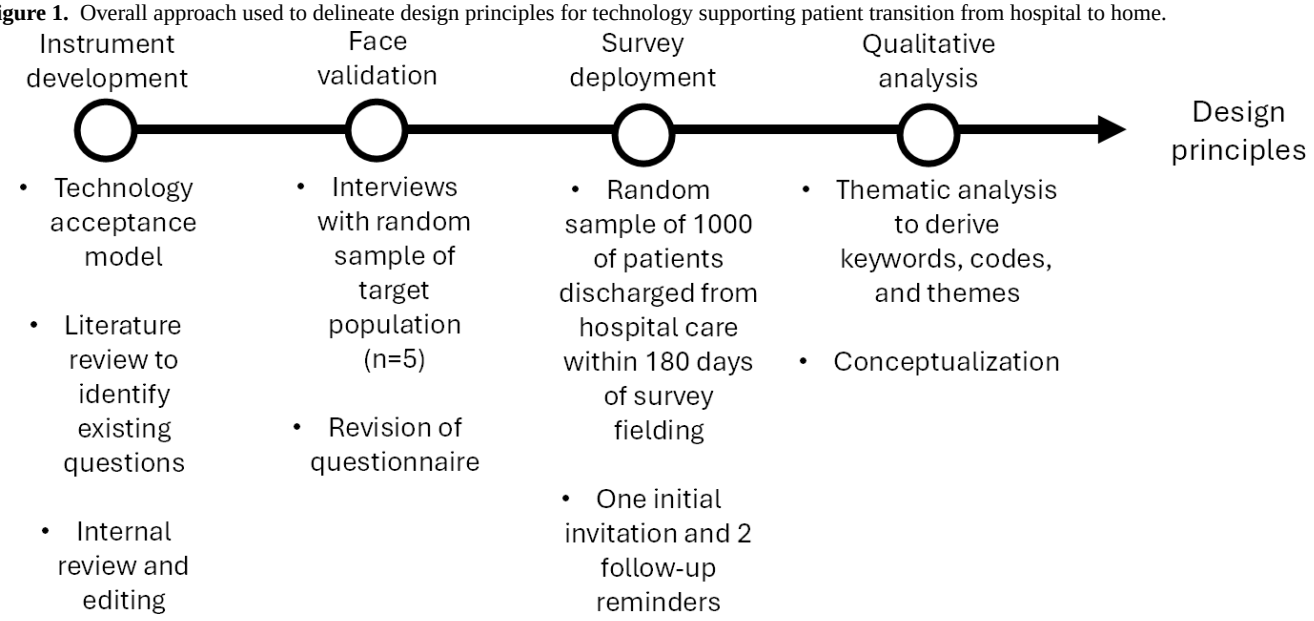
Methods

Study Design, Setting, and Population

This study was a vignette-based, mixed methods survey design grounded in the technology acceptance model (TAM) to investigate patient-perceived barriers and facilitators to the use of digital tools to support patient connection back to care teams following hospital discharge. The TAM is a widely used theoretical framework for understanding when, how, and why individuals use technology [22]. Incorporating TAM constructs into the development of questionnaires ensures that all aspects known to influence technology use and acceptance are assessed within the survey instrument [22]. Vignette-based studies within the qualitative research landscape allow for an in-depth understanding of patient attitudes on potential changes to their health care environment and are commonly used in the human-centered design process to facilitate the development of design principles [23]. An overview of the approach is displayed in Figure 1. A Checklist for Reporting Results of Internet E-Surveys (CHERRIES) form is provided as Multimedia Appendix 3.

XSL•FO
RenderX

Figure 1. Overall approach used to delineate design principles for technology supporting patient transition from hospital to home.



Mayo Clinic is a multispecialty medical center with campuses in Rochester, Minnesota; Jacksonville, Florida; and Phoenix, Arizona, and a health care system spanning 44 communities in southern Minnesota, western Wisconsin, and northern Iowa. The Rochester, Minnesota, campus has 2 hospitals with over 2000 beds. The target population was 1000 unique patients discharged from either of the 2 campuses in Rochester, Minnesota. To ensure the relevancy and plausibility of our findings, we focused on patients discharged within 180 days before survey distribution. A simple random sample was performed among patients aged 18 years or younger at the time of discharge, residing within the hospital's catchment area, who had agreed to participate in research per state statute, were discharged to home, and were not deceased according to our medical records. The target sample size was 1000 patients to gain responses from at least 200 individuals, based on the 20% response rate typically observed in this population. The survey was tested, deployed, and managed by the Mayo Clinic Survey Research Center, as described in the later section "Instrument fielding."

Instrument Development and Testing

Instrument development involved a multistep process: development of vignettes, face validation, and survey development. Vignettes focused on 2 technologies under consideration for use by hospital practice: watch-based accelerometers intended to help hospital teams understand patient movement as a proxy for recovery, and text message–based ecological momentary assessments to check in on patients using short sets of questions. Initial vignettes were drafted by the principal investigator (LMP) in collaboration with the Mayo Clinic Center for Digital Health, responsible for the implementation of digital tools for patients as part of routine clinical care. The first review was performed by members of the research team (PR and REC) and internal medicine physicians specializing in hospital care (SBD and AS) and primary care (JOE).

Once initial vignettes were completed, they entered a phase of face validation, whereby 5 individuals fitting the criteria for our target population were interviewed one-on-one with a research team member (CMS) trained in qualitative research methods. A semistructured interview format was used, whereby each participant was provided an overview of the intent of the study and provided a printed copy of each draft vignette, as well as had the vignette read aloud by the interviewer (CMS). Each participant was asked about the clarity of the vignette through a semistructured interview. Participants were offered the option to be interviewed remotely (telephone or video) or in person, and all interviews were recorded for analysis purposes. Based on face validation, the technology vignettes were revised before launch. The interview guide is provided as Multimedia Appendix 1.

The complete survey instrument included the revised technology vignettes along with 2 open-ended questions: "What initial impressions do you have when thinking about (using an accelerometer/receiving ecologic momentary assessment) after a hospital stay?" and "What questions come to mind that you would want answered before agreeing to (wear an accelerometer/receive ecologic momentary assessment) after a hospital stay?" Following each vignette, respondents were asked to evaluate the level of importance of literature-derived or hypothesized design principles related to the use of each technology on a 5-point Likert scale. Questions grounded in the TAM that have been previously used and tested by the study team [24,25] were used to assess general use of technology such as the ability to access the internet at home or other locations, access to devices (eg, personal computers or laptops, and smartphones), and patient interest in and ability to use technology. Patient demographic factors, medical comorbidities, and patient-reported quality of life using the EuroQol 5-Dimension Scale (EQ-5D) [26] were also assessed. The survey instrument is provided as Multimedia Appendix 2.

Instrument Fielding

The finalized instrument was built by the Mayo Clinic Survey Research Center staff within Qualtrics survey software (Qualtrics LLC) for electronic deployment, sent to patient-provided email addresses. Pilot testing of the finalized

XSL•FO
RenderX

instrument led the study team to anticipate that respondents would need 20 - 25 minutes to complete the survey. Based on the inclusion criteria, 1000 patients were sampled to complete the online survey on Thursday, October 31, 2024. In total, 33 invitation emails bounced upon delivery. A total of 2 reminders were sent to nonresponders on November 3 and November 8, 2024, and the survey was closed on Wednesday, November 14, 2024, with 116 total responses.

Thematic Review and Derivation of Design Principles

Thematic review of participant-provided responses was completed using grounded theory. Open coding was performed to identify initial themes within the open-ended questions following the technology vignettes. Specifically, the input provided to the following questions was used: (1) "What initial impressions do you have when thinking about (using an accelerometer/receiving text messages) after a hospital stay?" and (2) "What questions come to mind that you would want answered before agreeing to (wear an accelerometer/answering questions from your care team via text message) after a hospital stay?" Keywords were derived from responses via close examination of responses, whereby recurring patterns and terms arose. Coding was performed to capture core messages provided by survey respondents, allowing the data to be transformed into generalizable, theoretical elements related to the main research questions. Design themes were derived from coded values, whereby design principles were created via the group's conceptualization of keywords, codes, and themes. The final design principles are shared as the final product to help inform the design of future technological tools to support patients transitioning from hospital to home. This process was conducted independently by 2 individuals trained in research methods with the following characteristics: Analyst A is a first-generation Asian American woman with training in public health science and epidemiology with extensive experience within health care delivery systems, while Analyst B is a White woman with training in human-centered design with experience in designing new products and services to enhance patient experiences with health care. Analyst A's experience navigating systemic barriers in health care for older family members informed her sensitivity to issues of access and equity, whereas Analyst B's 20-year experience designing digital solutions shaped her sensitivities to digital barriers. While Analyst A shares lived experience as a caregiver to the study population, she remains mindful of the varying challenges and psychosocial factors that patients of different ages, socioeconomic levels, and cultural backgrounds may face upon hospital discharge, thereby allowing her to approach data with a mindset open to understanding different perspectives. Analyst B, as a trained designer who is practiced in contextual inquiry and eliciting people's behaviors, goals, and motivations, approached this work with a trained perspective of understanding patient needs through qualitative research methods and analysis. The collaboration was grounded in mutual respect and critical dialogue, allowing us to challenge each other's assumptions and enrich our interpretations. The team recognized that positionalities are not fixed and committed to ongoing reflexivity throughout the research process.

Analysis of Quantitative Survey Responses

Quantitative survey response data were used to reinforce the design principles derived via our qualitative approach and were descriptive. Likert-based survey responses are described using frequency (n) and proportion. All data management and analyses were performed using SAS (version 9.2; SAS Institute). Response visualizations were created using R statistical software (version 4.1.2; R Foundation for Statistical Computing) and the ggplot2 package.

Ethical Considerations

This study was reviewed by the Mayo Clinic Institutional Review Board and deemed exempt from Institutional Review Board approval under 45 CFR 46.104(d), category/subcategory 2(i) (24 - 005650; principal investigator LMP). Confidential or identifying information was not collected for study participants, and all data were shared in aggregate form. Remuneration for study participation was not provided. As protected health information was not requested, written consent and Health Insurance Portability and Accountability Act (HIPAA) authorization were not required following 45 CFR 160.103.

Results

Overview

We had 116 completed responses to our vignette-based investigation (116/967, 12% response rate; excluding 33 bounced emails; Table 1). Our responding population had a median age of 71 (IQR 61-78) years, was predominantly of White race (112/116, 97%), and reported having chronic obstructive pulmonary disease (103/116, 88.8%), diabetes (87/116, 75.0%), cancer (65/116, 56.0%), osteoarthritis/chronic joint pain (63/116, 54.3%), and predominantly "good" (48/116, 41.4%) or "fair" (29/116, 25.0%) health. Most respondents reported having access to a smartphone (98/116, 84.5%), personal computer/laptop (85/116, 73.3%), and in-home internet (111/116, 95.7%).

XSL•FO
RenderX

Table . Descriptive characteristics of respondents to vignette-based survey regarding technology to support hospital discharge (N=116).

Characteristic	Overall
Age (years), median (IQR)	71 (61-78)
Race, n (%)	
Black, African, or African American	1 (0.9%)
White	112 (96.6%)
Mixed races	1 (0.9%)
Prefer not to disclose	1 (0.9%)
Missing	1 (0.9%)
Comorbidities, n (%)	
Chronic obstructive pulmonary disease	103 (88.8%)
Diabetes (type 1 or type 2)	87 (75.0%)
Cancer	65 (56.0%)
Osteoarthritis/chronic joint pain	63 (54.3%)
High cholesterol	57 (49.1%)
Hypertension/high blood pressure	42 (36.2%)
Self-rated health, n (%)	
Excellent	3 (2.6%)
Very good	26 (22.4%)
Good	48 (41.4%)
Fair	29 (25.0%)
Poor	10 (8.6%)
Technology access (select all that apply), n (%)	
Smartphone	98 (84.5%)
Personal computer/laptop	85 (73.3%)
Tablet	47 (40.5%)
Smartwatch	14 (12.1%)
None	1 (0.9%)
Are you able to access the internet from your smartphone when you are not at home or logged into an internet/Wi-Fi service?, n (%)	
The patient does not have a smartphone	18 (15.5%)
Yes	88 (89.8%)
No	8 (8.2%)
I do not know	2 (2.0%)
Do you access the internet from your home?, n (%)	
Yes	111 (95.7%)
No	4 (3.4%)
I do not know	1 (0.9%)
Are you satisfied with your ability to access the internet?, n (%)	
Patient does not have in-home internet	5 (4.31)
Satisfied	91 (82.0%)
Neutral	18 (16.2%)
Dissatisfied	2 (1.8%)

XSL•FO
RenderX

Themes Associated With Technology-Supported Hospital Discharge

Several themes arose during qualitative analysis. The themes are outlined below with representative patient quotes.

Connection and Response

Most participants felt positively toward having a continuous connection with Mayo Clinic once discharged and thought they would benefit from sharing their data and receiving recommendations or other support to manage their health. One respondent wrote,

> *If it helps my providers gain needed information that can improve my overall outcome, I am in favor.*

while another stated,

> *I really like the idea. It would be able to track how a person is generally doing instead of relying 100% on the patient.*

Technical Ease of Use

Participants felt comfortable with the technology associated with these concepts, although some expressed concerns with management regarding difficulties if the technology proved challenging. One participant stated,

> *I am afraid I wouldn't understand how to use the device or if there was a glitch, I wouldn't be able to figure out how to fix it.*

and another less tech-savvy patient mentioned wanting "an alternative way to keep in touch because my techno skills are non-existent."

Solution Usability

Participants had questions or concerns with the wearability, simplicity, or frequency of use/interaction associated with the solution. This was the most popular theme, as participants recognized the impact of daily use. Several participants were wondering how long they would need to wear the device. One asked,

> *Do I wear it 24 hours a day?*

Regarding simplicity, another stated,

> *I don't look at my phone much sometimes 1 time a day.*

suggesting they may need to change their behavior if more frequent interaction is needed.

Human Connection

Participants mentioned a concern that these technologies could take away from the human connection, replacing it with an impersonal solution that might not recognize their specific needs. One participant mentioned,

> *It's okay but I feel it is better to talk to someone so I can express my feelings about my recovery.*

and another stated,

> *I would not want these text messages to replace the after-hospital discharge phone call from a hospital staff member.*

Cost

Participants were concerned about any costs associated with the product and expected not to incur any extra costs. One participant summed up the general feelings of the group by stating,

> *I would not expect to be charged for the device or access.*

Privacy and Security

Participants were concerned about data privacy and the security of the information shared. One participant mentioned,

> *What exactly does it monitor and who has access to the data?*

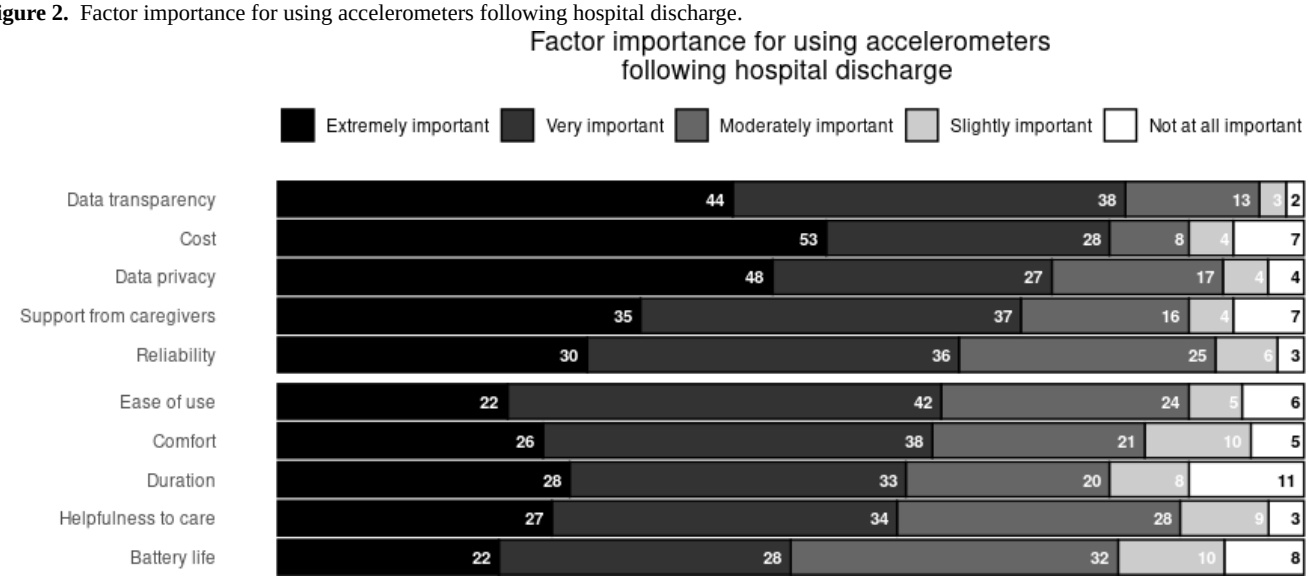
Another stated, "I would want to make sure it is the doctor and not spam," and still another mentioned, "it seems a bit invasive."

Patient-Reported Technology Attribute Importance to Support Hospital Discharge

Regarding use of an accelerometer following a hospital stay, respondents reported that "the ability to see my own accelerometer data" (95/116, 81.9% rated as "Very important" or "Extremely Important") and "how much I would need to pay" (94/116, 81.0% rated as "Very Important" or "Extremely Important") had the highest cumulative levels of importance impacting willingness to use an accelerometer following a hospital stay. These were followed closely by "reassurance that my data is private between me and my provider team" (87/116, 75% rated as "Very Important" or "Extremely Important") and "how supportive my caregivers are to my use" (84/116, 72.4% rated as "Very Important" or "Extremely Important"). Respondents noted that "how long I would need to wear an accelerometer" (22/116, 19% rated as "Not Important at all" or "Slightly Important") and "How long the battery would last" (21/116, 18.1% rated as "Not Important at all" or "Slightly Important") were cumulative of least importance regarding willingness to use the device (Figure 2).

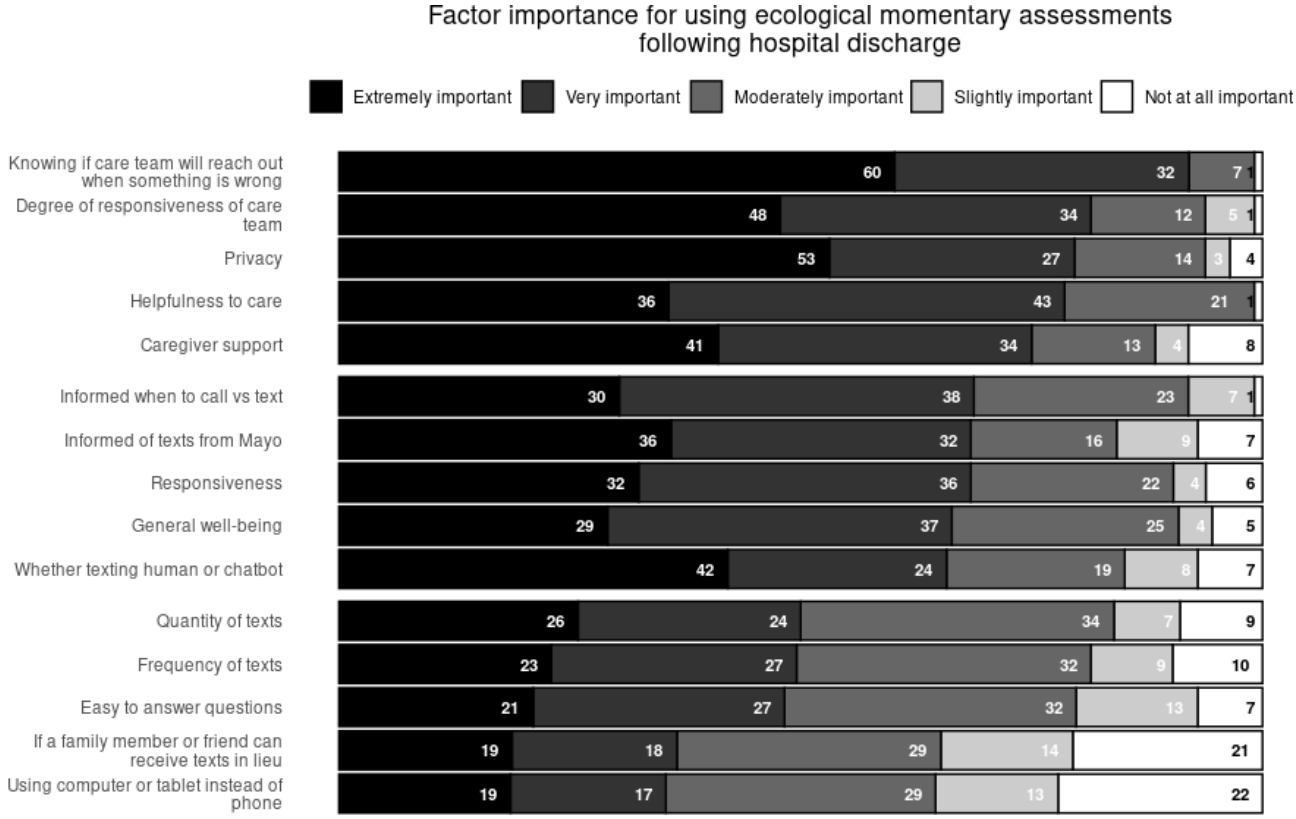
XSL•FO
RenderX

Figure 2. Factor importance for using accelerometers following hospital discharge.



When asked about factor importance in respondents' willingness to receive ecological momentary assessments following a hospital stay, the factors with the highest cumulative ranked importance were: "Whether a member of the care team would reach out if something was wrong" (107/116, 92.2% rated as "Very Important" or "Extremely important"), "How responsive the care team is to my answers" (95/116, 81.9% rated as "Very Important" or "Extremely important"), and "How helpful I feel the questions will be to my care" (92/116, 79.3% rated as "Very Important" or "Extremely important"; Figure 3). Respondents indicated "Whether I could have a family member or friend receive the messages instead of me" (24/116, 20.7% rated as "Not at all Important") and "Whether I can use my computer or tablet to answer instead of a cell phone" (25/116, 21.6% rated as "Not at all Important") had most responses indicating this factor was of no importance at all.

Figure 3. Factor importance for using ecological momentary assessments following hospital discharge.



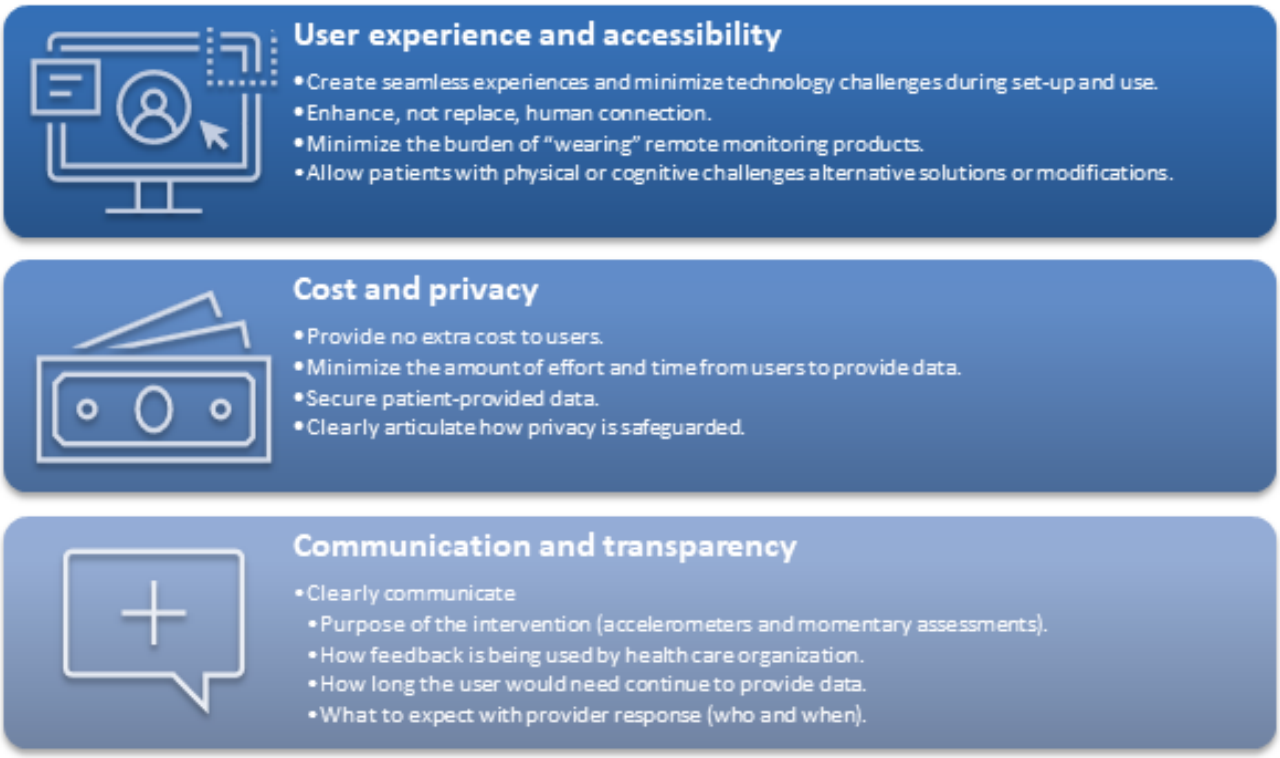
Design Principles for Technology-Supported Hospital Discharge

Using the themes as guidance, 3 design principles were derived to highlight how concepts should perform (Figure 4). The first design principle describes user experience and accessibility by patients and includes constructs such as a seamless experience for end users (patients) during initial engagement with and use of the technology, how technology support should enhance the connection patients feel to care teams and not replace the human

XSL•FO
RenderX

connection, how technology needs to minimize the amount of effort needed by patients to enable the end goal of supportive healing at home following hospital discharge, and how modifications are needed to support patients with physical or cognitive limitations, such as patients reliant on wheelchairs or caregivers for cognitive support.

Figure 4. Design principles for technology to support patients during hospital discharge.



The second design principle derived focuses on consideration of cost to the patient when deploying new technology-based support for patients transitioning from hospital to home care. Cost is broadly defined as financial, including copayments and out-of-pocket costs, in addition to time and effort needed to participate in postdischarge, technology-based care models. Patients also identify the need for data security to ensure their information is private between the patient and the care team, and clear communication on how data privacy will be assured.

The third design principle emphasizes the need for communication about the purpose of the intervention and how the technology will be part of their care experience and used by the care team. Expectations regarding the duration of expected engagement with the technology were also important to patients, as was a clear understanding of which member of the care team would be interacting with the technology, and what the anticipated response content and timeline may be.

Discussion

Principal Findings

In this study, we derived 3 categories of design principles to aid in the design, development, and deployment of technology to support patients as they transition from hospital to home. We observed that patients value technology fitting seamlessly into their care experience, extending but not replacing human connection. Within our study, nearly every respondent indicated high importance that care teams be responsive to information received via technology, particularly if something seems to be going wrong. Patients were sensitive to additional costs that could be incurred based on use of these new technologies and placed high importance on data privacy and security in their willingness to engage with technology to support hospital discharge. Communication and transparency of the "why, what, who, and when" is important when engaging with technology to support transitions from hospital care to home. While the technical aspects of using technology are a consideration for patients, battery life, duration of use, and ability to use alternative mechanisms such as tablets or desktop computers for participation were of lower importance in patients' willingness to engage.

Comparison With Previous Work

Several studies have demonstrated a high degree of acceptance of digital technology to support hospital-to-home transition [7,9,27,28]. Studies deploying SMS-based text messaging had high patient satisfaction rates, and patients rated programs high on convenience and responsiveness during hospital-to-home transitions [7,27]. In addition, patients have reported appreciation for the flexibility enabled via automation and the asynchronous nature of technology solutions compared with traditional in-person care models [29]. Within our design exercise, we observed that patients value human connection supplemented by, but not replaced by, technology, building upon these aspects reported in the literature. Technology supporting human connection has been demonstrated in the hosting of virtual birthdays, funerals, and graduations, enabling shared human experiences across physical distances, which is particularly helpful for individuals experiencing physical

XSL•FO
RenderX

disabilities, frailty, or limited resources to support travel. In health care, technology has enabled human connection to mental health support that is accessible and practical, particularly for certain populations, such as adolescents [30] and those residing in areas with decreased availability of support services [31]. The multimodal delivery of mental health support, including video- and telephone-enabled care, mobile apps, and therapeutic gaming, supports the ability to foster human connection via technology [32].

Research has shown that bidirectional communication is considered more valuable than automated reminders by patients [9,19]. Both within our study and as reported elsewhere [33], patients indicate the desire to know that the interactions they are having with a device are being monitored or overseen by a human being, with the ability to escalate needs should a patient's condition warrant increased levels of care [34]. This reassurance has been reported to increase patient and caregiver confidence in the transition from institutionalized care to home care [34]. Use of technology to support hospital-to-home transitions is a mechanism to mitigate costs experienced by health care systems and payers alike via improved patient outcomes and decreased use [7,35]. However, the technology needed to support transitions from hospital to home may consist of personal computers, laptops, or tablets for inputting information, and in-home monitoring supplies such as glucose or blood pressure monitoring. Payment models to support these costs vary greatly, from coverage by payers to hospital systems to patients and families themselves. In addition to being cognizant of the costs associated with these new technologies, health care institutions need to address patient privacy and communicate those privacy considerations to patients upon consent for participation. Devices and platforms must be chosen that follow federal requirements (ie, HIPAA compliance within the United States), which require end-to-end data encryption, access controls, and audit trails. In addition, health care systems must require role-based access controls to ensure that only pertinent members of the care team can access patient information. All of these requirements must be supported by robust informed consent processes that ensure that patients understand how their data will be collected, used, and shared, and how patients and caregivers can protect themselves via secured network access, locking of personal devices, and avoidance of phishing.

The ability to incorporate supportive technology into transitions from hospital-based care to home as part of the care team, and not a separate entity, was emphasized as a need by our study participants. Technologies such as SMS-based text messaging, accelerometers, and other tools could be used to facilitate seamless transitions between our traditional care settings of hospital-based care, postacute care, and home with return to primary care. Patients with complex medical needs can experience frequent care transitions between our traditional care settings. However, care must be taken to ensure that safe and effective care is being maintained, particularly among those in a medically vulnerable state, such as following hospital discharge. Accuracy of supportive technology must be ensured so that care teams responding to in-home monitoring have reliable information, and devices must be usable by patients and caregivers. One opportunity is the movement toward "Bring Your Own Device" to help ensure usability; however, variability exists in the accuracy of direct-to-consumer monitoring devices, such as for blood pressure [36]. Health care systems may opt to designate a set of acceptable monitoring devices, allowing patients to use the device they choose within safety parameters outlined by health care providers. Grounding new tools in user experience and accessibility, cost and privacy, and communication and transparency can help clinical teams and developers improve the incorporation of tools into routine clinical care, thereby advancing seamless transitions during an anxiety-filled and uncertain time for patients.

Strengths and Limitations

Our study had limitations. Although our sample would be considered large for a qualitative assessment (n=116), the overall response rate was low (12%). We believe that lack of incentivization for study participation, use of patient-provided email as opposed to use of the clinic-provided patient portal, and survey availability in a single language (English) may have contributed to our low overall response rate. In addition, our response population was predominantly older adults (median age 71 years, IQR 61-78) and White (112/116, 97%) respondents from a single academic medical center, which may limit our ability to generalize to other populations. The older age of our participants may have influenced the digital literacy and access of our population, as research shows that older adults are more likely to experience decreased access to digital tools and services [37], and decreased digital literacy serves as an independent stressor to the well-being of older adults [38]. In addition, individuals of racial and ethnic minority groups are less likely to feel trust in the health care system and may be less willing to provide surveillance data directly back to their health care provider. Our response population was generally favorable about the use of movement tracking and text-based assessments following hospital discharge, but these results may not be consistent with other populations. Our investigation focused on 2 technologies (ecological momentary assessments via text message and accelerometers), which may not encompass the nuances of new tools that could be created to support all care transitions. One of the tools explored (accelerometers) is considered an emerging technology, which may not be validated for use in clinical settings. Our study had several strengths. First, this is one of the few investigations published to incorporate design thinking and UCD practices to help inform the design, development, and implementation of technology to support patients transitioning from hospital to home, focusing on the practical implications of new tools or services. Second, we used a vignette-based, mixed methods inquiry to derive open-ended feedback from a large qualitative sample of individuals who recently experienced hospital-to-home discharge. We used qualitative methodology to derive design constructs that can be used in multiple design settings and validate findings with quantitative assessment derived through the face validation exercise of patient interviews.

Future Directions

As health care systems increasingly rely on digital technologies to support transitions from hospital to home, future research must prioritize equity, inclusivity, and adaptability, building on

XSL•FO
RenderX

and testing the design principles presented here. One critical area for exploration is how these design principles perform across diverse cultural, racial, and demographic groups, given the limited diversity in our present sample. In addition, opportunities exist to evaluate emerging technologies—such as artificial intelligence–driven care coordination platforms, wearable sensors, and voice-assisted devices—for their usability, accessibility, and clinical impact. These innovations hold promise for improving postdischarge outcomes, but their success depends on aligning with patients' real-world needs, preferences, and capabilities. Finally, interdisciplinary collaboration among clinicians, social scientists, caregivers, and patients themselves will be essential to design and implement user-centered, culturally responsive digital solutions. By addressing these gaps, future work can help ensure that digital health technologies truly enhance recovery and reduce disparities as patients transition from hospital to home.

Conclusion

Our study identified key design principles for developing technology to support patients transitioning from hospital to home, emphasizing the importance of seamless integration into care, minimizing patient and caregiver burden, and ensuring data privacy and security. Patients value human connection supplemented by technology and prefer transparent communication about the use and purpose of these tools. Future efforts should focus on incorporating these principles to enhance patient experience and outcomes during postacute care transitions.

Acknowledgments

We thank the Mayo Clinic Survey Research Center for their help in planning, testing, and implementing the survey. This publication was made possible by the Clinical and Translational Science Award (CTSA; UL1 TR002377) from the National Center for Advancing Translational Sciences (NCATS), a component of the National Institutes of Health (NIH). Author (SBD) was also supported by the National Institute on Minority Health and Health Disparities (NIH; K23 MD016230). The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

LMP, AS, SBD, PR, JMPH, and JOE contributed to conceptualization. REC, CMS, MAD, and PR handled data curation. REC, CMS, MAD, and PR conducted formal analysis. LMP, AS, SBD, and JOE contributed to funding acquisition. LMP, REC, CMS, MAD, and PR managed investigation. LMP, CMS, MAD, and PR handled methodology. LMP and JMPH managed resources. REC, CMS, MAD, and PR worked on the software. LMP and JOE managed supervision. LMP, AS, SBD, and JOE contributed to validation. LMP, MAD, PR, and JOE wrote the original draft. LMP, AS, SBD, REC, CMS, MAD, PR, JMPH, and JOE contributed to the review and editing of the manuscript.

Conflicts of Interest

LMP and SBD have received funding from the US National Institutes of Health that does not conflict with this manuscript. All authors report no conflicts of interest related to this manuscript.

Multimedia Appendix 1
Semistructured interview guide.
[DOCX File, 30 KB - humanfactors_v12i1e72913_app1.docx]

Multimedia Appendix 2
Deployed survey instrument.
[DOCX File, 41 KB - humanfactors_v12i1e72913_app2.docx]

Checklist 1
Checklist for Reporting Results of Internet E-Surveys (CHERRIES) checklist.
[PDF File, 106 KB - humanfactors_v12i1e72913_app3.pdf]

References

1. Hospital inpatient and outpatient services. In Report to the Congress: Medicare and the health care delivery system (Chapter 5). : Medicare Payment Advisory Commission; 2018.

XSL•FO
RenderX

2. Shi SM, Kosar CM, Gouskova N, Berry S. Post-acute care rehabilitation services and outcomes in skilled nursing facilities before and during the COVID-19 pandemic. JAMA Health Forum 2023 Mar 3;4(3):e230019. [doi: 10.1001/jamahealthforum.2023.0019] [Medline: 36867421]
3. Geng F, McGarry BE, Rosenthal MB, Zubizarreta JR, Resch SC, Grabowski DC. Preferences for postacute care at home vs facilities. JAMA Health Forum 2024 Apr 5;5(4):e240678. [doi: 10.1001/jamahealthforum.2024.0678] [Medline: 38669031]
4. Madara J. Letter of Support: Telehealth Modernization Act of 2024. : American Medical Association; 2024.
5. Janke AT, Venkatesh AK. Understanding and addressing the US hospital bed shortage-build, baby, build. JAMA Netw Open 2025 Feb 3;8(2):e2460652. [doi: 10.1001/jamanetworkopen.2024.60652] [Medline: 39969886]
6. Hudon C, Aubrey-Bassler K, Chouinard MC, et al. Better understanding care transitions of adults with complex health and social care needs: a study protocol. BMC Health Serv Res 2022 Feb 15;22(1):206. [doi: 10.1186/s12913-022-07588-0] [Medline: 35168628]
7. Bressman E, Long JA, Honig K, et al. Evaluation of an automated text message-based program to reduce use of acute health care resources after hospital discharge. JAMA Netw Open 2022 Oct 3;5(10):e2238293. [doi: 10.1001/jamanetworkopen.2022.38293] [Medline: 36287564]
8. Fakhoury H, Trochez R, Kripalani S, Choma N, Blessinger E, Nelson LA. Patient engagement with an automated postdischarge text messaging program for improving care transitions. J Hosp Med 2024 Jun;19(6):513-517. [doi: 10.1002/jhm.13334] [Medline: 38497416]
9. Bruce C, Pinn-Kirkland T, Meyers A, et al. Investigating patient engagement associations between a postdischarge texting programme and patient experience, readmission and revisit rates outcomes. BMJ Open 2024 Mar 13;14(3):e079775. [doi: 10.1136/bmjopen-2023-079775] [Medline: 38485169]
10. Charkviani M, Barreto EF, Pearson KK, et al. Development and implementation of an acute kidney injury remote patient monitoring program: research letter. Can J Kidney Health Dis 2023;10:20543581231192746. [doi: 10.1177/20543581231192746] [Medline: 37577175]
11. Nundy S, Razi RR, Dick JJ, et al. A text messaging intervention to improve heart failure self-management after hospital discharge in a largely African-American population: before-after study. J Med Internet Res 2013 Mar 11;15(3):e53. [doi: 10.2196/jmir.2317] [Medline: 23478028]
12. Holbrook A, Perri D, Levine M, et al. Improving medication prescribing-related outcomes for vulnerable elderly in transitions on high-risk medications (IMPROVE-IT HRM): a pilot randomized trial protocol. Pilot Feasibility Stud 2024 Apr 10;10(1):60. [doi: 10.1186/s40814-024-01484-6] [Medline: 38600599]
13. Lindemann CH, Burst V, Völker LA, et al. Personalized, interdisciplinary patient pathway for cross-sector care of multimorbid patients (eliPfad trial): study protocol for a randomized controlled trial. Trials 2024 Mar 11;25(1):177. [doi: 10.1186/s13063-024-08026-8] [Medline: 38468319]
14. Xu C, Ishii Y, Mizutani M, et al. Predictors of gait variability in older inpatients: an exploratory study among orthopedic patients. Cureus 2024 Oct;16(10):e71733. [doi: 10.7759/cureus.71733] [Medline: 39552952]
15. Fessler L, Tessitore E, Craviari C, et al. Motivational and emotional correlates of physical activity and sedentary behavior after cardiac rehabilitation: an observational study. BMC Sports Sci Med Rehabil 2024 Oct 3;16(1):209. [doi: 10.1186/s13102-024-00997-0] [Medline: 39363352]
16. Suffoletto B, Kim D, Toth C, et al. Feasibility of measuring smartphone accelerometry data during a weekly instrumented timed up-and-go test after emergency department discharge: prospective observational cohort study. JMIR Aging 2024 Sep 4;7:e57601. [doi: 10.2196/57601] [Medline: 39258924]
17. Sergi I, Montanaro T, Shumba AT, et al. An iot-aware system for remote monitoring of patients with chronic heart failure. Presented at: 2023 8th International Conference on Smart and Sustainable Technologies (SpliTech); Jun 20-23, 2023; Split/Bol, Croatia. [doi: 10.23919/SpliTech58164.2023.10193338]
18. Philpot LM, Dugani SB, Singla A, DeZutter M, Ebbert JO. Digital care horizon: a framework for extending health care through digital transformation. Mayo Clin Proc Digit Health 2023 Sep;1(3):210-216. [doi: 10.1016/j.mcpdig.2023.05.005] [Medline: 37601768]
19. Ahn A, Morgan AU, Burke RE, et al. Postdischarge needs identified by an automated text messaging program: a mixed-methods study. J Hosp Med 2024 Dec;19(12):1138-1146. [doi: 10.1002/jhm.13466] [Medline: 39051626]
20. Tellioğlu H. User-centered design. In: Xiang Z, editor. Handbook of E-Tourism: Springer International Publishing; 2020:1-19 URL: https://link.springer.com/referencework/10.1007/978-3-030-05324-6 [accessed 2025-08-14]
21. Saragosa M, Nizzer S, McKay S, Kuluski K. The hospital-to-home care transition experience of home care clients: an exploratory study using patient journey mapping. BMC Health Serv Res 2023 Aug 31;23(1):934. [doi: 10.1186/s12913-023-09899-2] [Medline: 37653515]
22. AlQudah AA, Al-Emran M, Shaalan K. Technology acceptance in healthcare: a systematic review. Appl Sci (Basel) 2021;11(22):10537. [doi: 10.3390/app112210537]
23. Hawkins M, Massuger W, Cheng C, et al. Codesign and implementation of an equity-promoting national health literacy programme for people living with inflammatory bowel disease (IBD): a protocol for the application of the Optimising Health Literacy and Access (Ophelia) process. BMJ Open 2021 Aug 26;11(8):e045059. [doi: 10.1136/bmjopen-2020-045059] [Medline: 34446478]

XSL•FO
RenderX

24. Philpot LM, Ramar P, Roellinger DL, McIntee MA, Ebbert JO. Digital health literacy and use of patient portals among Spanish-preferred patients in the United States: a cross-sectional assessment. Front Public Health 2024;12:1455395. [doi: 10.3389/fpubh.2024.1455395] [Medline: 39720810]
25. Philpot LM, Ramar P, Roellinger DL, Njeru JW, Ebbert JO. Individual-level digital determinants of health and technology acceptance of patient portals: cross-sectional assessment. JMIR Form Res 2024 Jun 10;8:e56493. [doi: 10.2196/56493] [Medline: 38695754]
26. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec;20(10):1727-1736. [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]
27. Rosenberg K. Automated texting program reduces use of resources after discharge. Am J Nurs 2023;123(2):61-61. [doi: 10.1097/01.NAJ.0000919752.18752.9c]
28. Donati M, Bacchillone T, Fanucci L, Saponara S, Costalli F. Operating protocol and networking issues of a telemedicine platform integrating from wireless home sensors to the hospital information system. Journal of Computer Networks and Communications 2013;2013(1):1-12. [doi: 10.1155/2013/781620]
29. Keng CJS, Goriawala A, Rashid S, et al. Home to stay: an integrated monitoring system using a mobile app to support patients at home following colorectal surgery. J Patient Exp 2020 Dec;7(6):1241-1246. [doi: 10.1177/2374373520904194] [Medline: 33457571]
30. Punukollu M, Marques M. Use of mobile apps and technologies in child and adolescent mental health: a systematic review. Evid Based Ment Health 2019 Nov;22(4):161-166. [doi: 10.1136/ebmental-2019-300093] [Medline: 31358537]
31. McCord C, Ullrich F, Merchant KAS, et al. Comparison of in-person vs. telebehavioral health outcomes from rural populations across America. BMC Psychiatry 2022 Dec 10;22(1):778. [doi: 10.1186/s12888-022-04421-0] [Medline: 36496352]
32. Ferreira A, Moreira MT, Rodrigues S, Lima A, Ferreira S, Fernandes CS. Use of technologies in promoting adolescent mental health: a scoping review. Creat Nurs 2025 May;31(2):111-118. [doi: 10.1177/10784535251313554] [Medline: 40371452]
33. Liss DT, Serrano E, Wakeman J, et al. “The doctor needs to know”: acceptability of smartphone location tracking for care coordination. JMIR Mhealth Uhealth 2018 May 4;6(5):e112. [doi: 10.2196/mhealth.9726] [Medline: 29728349]
34. Sanger PC, Hartzler A, Han SM, et al. Patient perspectives on post-discharge surgical site infections: towards a patient-centered mobile health solution. PLoS One 2014;9(12):e114016. [doi: 10.1371/journal.pone.0114016] [Medline: 25436912]
35. Zhang Y, Peña MT, Fletcher LM, Lal L, Swint JM, Reneker JC. Economic evaluation and costs of remote patient monitoring for cardiovascular disease in the United States: a systematic review. Int J Technol Assess Health Care 2023 Apr 28;39(1):e25. [doi: 10.1017/S0266462323000156] [Medline: 37114456]
36. Ho A, Quick O. Leaving patients to their own devices? Smart technology, safety and therapeutic relationships. BMC Med Ethics 2018 Mar 6;19(1):18. [doi: 10.1186/s12910-018-0255-8] [Medline: 29510750]
37. Yang R, Gao S, Jiang Y. Digital divide as a determinant of health in the U.S. older adults: prevalence, trends, and risk factors. BMC Geriatr 2024 Dec 21;24(1):1027. [doi: 10.1186/s12877-024-05612-y] [Medline: 39709341]
38. Zheng Q, Wang W. The relationship between the digital divide and the well-being of older adults: the mediating role of learned helplessness and the moderating role of growth mindset. Curr Psychol 2024 Jul;43(25):21547-21556. [doi: 10.1007/s12144-024-05974-x]

Abbreviations:

CHERRIES: Checklist for Reporting Results of Internet E-Surveys
EQ-5D: EuroQol 5-Dimension Scale
HIPAA: Health Insurance Portability and Accountability Act
TAM : technology acceptance model
UCD: user-centered design

Edited by A Kushniruk; submitted 20.02.25; peer-reviewed by C Br, J Ogunsakin, MV Ebrahim, MAA Mamun, R Tripathi; revised version received 06.06.25; accepted 23.06.25; published 20.08.25.

Please cite as:
Philpot LM, Singla A, Dugani SB, Canning RE, Smith CM, DeZutter MA, Ramar P, Hovell JMP, Ebbert JO
Patient-Derived Design Principles for Technology-Enabled Healing at Home Following Hospital Discharge: Mixed Methods Study
JMIR Hum Factors 2025;12:e72913
URL: https://humanfactors.jmir.org/2025/1/e72913
doi:10.2196/72913

XSL•FO
RenderX

© Lindsey M Philpot, Abhinav Singla, Sagar B Dugani, Rachel E Canning, Christina M Smith, Meredith A DeZutter, Priya Ramar, Jennifer M P Hovell, Jon O Ebbert. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 20.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Patient Experiences With a Mobile Self-Care Solution for Low-Complex Orthopedic Injuries: Mixed Methods Study

Jelle Spierings[1*], MD; Gijs Willinge[2*], MD; Marike Kokke[1*], MD; Sjoerd Repping[3*], PhD; Wendela de Lange[4], BSc; Thijs Geerdink[2*], MD, PhD; Ruben van Veen[2*], MD, PhD; Detlef van der Velde[1*], MD, PhD; J Carel Goslings[2*], MD, PhD; Bas Twigt[2*], MD, PhD; Collaboration Group[6]

[1]Department of Trauma Surgery, St. Antonius Ziekenhuis Utrecht, Koekoekslaan 1, Nieuwegein, The Netherlands

[2]Department of Trauma Surgery, OLVG Hospital, Amsterdam, The Netherlands

[3]University of Amsterdam, Amsterdam University Medical Centers, Amsterdam, The Netherlands

[4]The Healthcare Innovation Center, Julius Centrum, University Medical Center Utrecht, Utrecht, The Netherlands

6

[*]these authors contributed equally

Corresponding Author:
Jelle Spierings, MD
Department of Trauma Surgery, St. Antonius Ziekenhuis Utrecht, Koekoekslaan 1, Nieuwegein, The Netherlands

Related Article:

This is a corrected version. See correction statement: https://humanfactors.jmir.org/2025/1/e75812

Abstract

Background: The Dutch acute health care system faces challenges with limited resources and increasing patient numbers. To reduce outpatient follow-up, direct discharge (DD) has been implemented in over 30 out of 80 Dutch hospitals. With DD, no routine follow-up appointments are scheduled after the emergency department (ED) visit for low-complex, isolated, and stable musculoskeletal injuries. This policy is supported by information leaflets, a smartphone app, and a telephone helpline with human support. Growing evidence shows that DD is satisfactory, safe, and effective in reducing secondary health care use, but thorough patient experiences are lacking.

Objective: The aim of this study was to explore the experiences of patients with DD to ensure durable adoption and to improve the treatment protocol.

Methods: A mixed method study was conducted parallel to the implementation of DD in 3 hospitals. Data were collected through a survey directly after the ED visit, a survey 3 months post injury, and semistructured interviews. Quantitative data were reported descriptively, and qualitative data used thematic analysis. Outcomes included the Bowen feasibility parameters: implementation, acceptance, preliminary efficacy, and demand. All patients who consented to the study face-to-face with one of the 12 low-complex musculoskeletal injuries were included in the study during the implementation period.

Results: Of the 429 patients who started the primary survey, 138 patients completed both surveys. A total of 18 semistructured interviews were conducted and analyzed. Patients reported a median treatment satisfaction score of 7.8 (IQR 6.6-8.8) on a 10-point scale of DD at the ED. Information quality was experienced as good (106/138, 77%), and most preferred DD over face-to-face follow-up (79/138, 59%). Patient information demands and app use varied among patients, with a median frequency of use of 4 times (ranging from 1 to 30).

Conclusions: This study shows that patients consider DD a feasible and safe alternative to traditional treatment, with a favorable perception of its acceptability, efficacy, applicability, and demand. Nevertheless, response rates were relatively low, and personal nuances and preferences must be considered when implementing DD. Clinicians and policy makers can use the insights to improve DD and work towards the integration of DD into clinical practice and future guidelines.

(JMIR Hum Factors 2025;12:e53074) doi:10.2196/53074

KEYWORDS

self-care application; mHealth; experience; traumasurgery; orthopedic surgery; virtual fracture clinic; patient perspective; direct discharge; musculoskeletal injury; mobile self-care; method study; health care system; hospital; mobile health; app; smartphone; satisfactory; effectiveness; treatment; virtual clinic; virtual care; digital health

XSL•FO
RenderX

Introduction

The Dutch acute health care system faces substantial challenges due to limited resources and a rising number of patients requiring in-hospital care [1,2]. To alleviate this pressure, virtual fracture clinics have been introduced for Orthopedic and Trauma surgery patients [3]. Direct discharge (DD) is the most basic part of a virtual fracture clinic, concerning solely low-complex, isolated, stable musculoskeletal injuries. With DD, patients with these injuries are discharged directly from the emergency department (ED) without routine outpatient follow-up. Patients receive a removable orthosis or sling and are given information summarized in a mobile self-care app, the Virtual Fracture Care (VFC) app.

Growing evidence shows that DD is a safe and effective alternative to "traditional" care with routine follow-up [4,5]. DD reduces secondary health care use without causing a shift to primary health care use, while patient-reported outcomes (eg, functional outcome and satisfaction scores) and adverse outcomes (eg, complications and persistent complaints) are comparable [3,4]. These results and the usefulness of this method during COVID-19 social distancing measurements have led to a rapid uptake of DD in the Netherlands for 12 frequent injuries at the ED [6]. Since the first introduction in 2019, over 25 out of the 80 Dutch hospitals have implemented DD as the standard of care, adding to over 100 virtual fracture clinics and DDs worldwide [7]. While this reorganization of care has yielded beneficial outcomes in terms of reducing resources with comparable patient outcomes, there is a lack of research on the experiences of patients and their relatives with DD.

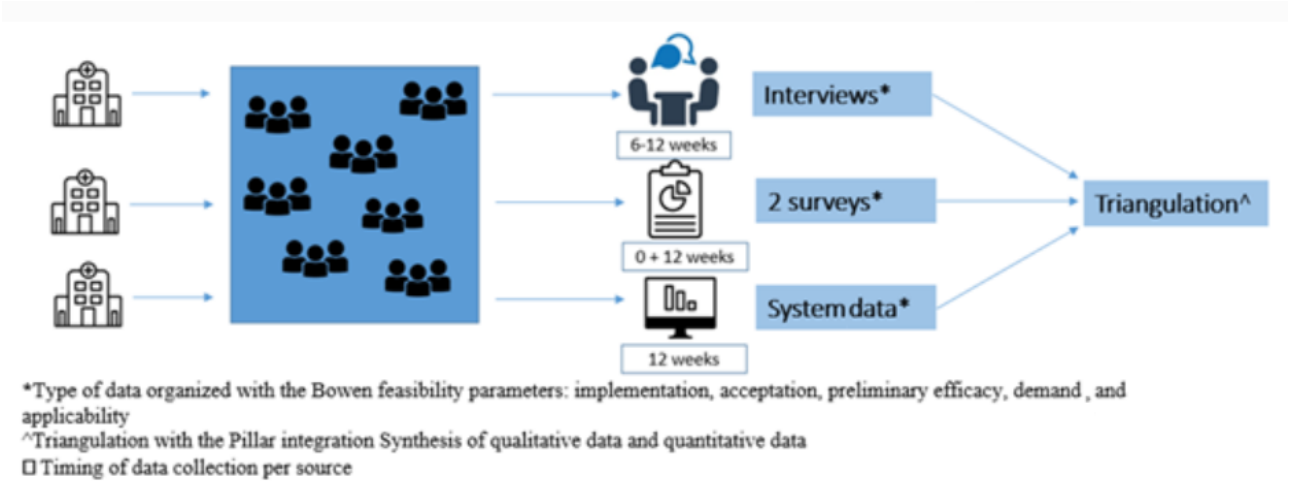
Performing an in-depth evaluation from an end user perspective deepens the insight into the quality of care by gaining a deeper understanding of user experiences and exploring reasons for nonadoption or abandonment, ensuring sustainable adoption of DD (inter)nationally. The aim of this study was therefore to explore the experiences of patients with DD to ensure durable adoption and to improve the protocol.

Methods

Design

An observational mixed method study was conducted among patients and parents of patients younger than 12 years who sustained low-complex, isolated, stable musculoskeletal injuries parallel to the implementation of DD in 3 Dutch level-2 trauma centers from September 2021 to July 2022 with an inclusion period of 3 months per hospital (Figure 1). Quantitative data and qualitative data were collected and analyzed separately by a quantitative team (GW and JS) and a qualitative team (WL and E Mathijsen). The Bowen feasibility framework was used to organize the data within the following parameters: implementation, acceptation, preliminary efficacy, and demand [8]. After separate analyses, quantitative data and qualitative data were triangulated with the Pillar Integration Process [9]. This study was reported according to the Good Reporting on a Mixed Methods Study (GRAMMS) criteria (Multimedia Appendix 1) [10], and according to the "Improving and Standardizing Evaluation Reports of Web-based and Mobile Health Interventions from the CONSORT-EHEALTH" [11].

Figure 1. Summary of study procedures and models used to evaluate the direct discharge protocol among patients.



Context

All participating centers were urban, level-2 trauma centers and teaching hospitals with up to 3 locations per hospital. All locations per hospital have an ED and implemented DD at the same time. Each hospital treats between 1200 and 1800 patients with low-complex, stable, isolated musculoskeletal injuries annually.

Traditional Treatment

Before DD was implemented, patients were treated according to local trauma protocols. These protocols consisted of immobilization or support with either a cast, sling, bandage, or splint and brief information about the injury at the ED. At least 1 outpatient follow-up appointment was scheduled at the plaster room or in the outpatient clinic within 2 weeks after the injury for review, extensive information, and definitive management planning.

XSL•FO
RenderX

Direct Discharge Protocol

This protocol was derived from the British model of a virtual fracture clinic and adapted to the Dutch health care setting in 2018. In its Dutch adaptation, DD includes 12 treatment protocols for low-complex, stable traumatic orthopedic injuries with additional injury-related criteria (Multimedia Appendix 2) [6]. Patients who met the injury-related inclusion criteria in Multimedia Appendix 2 and spoke Dutch or English fluently were included. No further predefined restrictions (eg, age or comorbidity) were used. Patients were excluded from the protocol if they had initial treatment in another hospital, follow-up in another hospital (eg, closer to home), multiple injuries, the reason for follow-up other than the injury (eg, social-care reasons), Eye/Motor/Verbal-score<15 at presentation, or intoxication. With DD, patients were discharged directly from the ED without routine outpatient follow-up. They receive a removable orthosis or a sling (eg, brace instead of a cast), and extensive information at the ED, summarized in a mobile self-care app, the VFC app. Apart from the criteria in Multimedia Appendix 2 and Dutch or English language skills (level B1), the protocols did not contain any predefined restrictions (eg, age or comorbidity). If, however, the ED staff believed physical follow-up was the most suitable treatment, outpatient follow-up was scheduled accordingly.

Patient eligibility for the protocol was re-evaluated on the next workday (within 24 h) by a team consisting of an (orthopedic) trauma surgeon and a radiologist. Patients who were incorrectly discharged directly were contacted by phone and scheduled for a face-to-face appointment.

VFC App

The VFC app provides self-care assistance through information, videos, and a helpline and can be downloaded for free at the Google Play Store and iOS App Store (Figure 2). Injury-specific leaflets with recovery information, treatment rules, and red flags were included. Furthermore, frequently asked questions, audiovisual exercise-, immobilization-, and analgesic instructions were included to assist patients. If patients required human contact in addition to the information, a helpline operated by an employee (eg, plaster technician) was available during working hours. The VFC app aimed to increase self-management and self-care during recovery and to substitute face-to-face follow-up. No reminders were sent, and use was voluntary. The app was developed by OLVG to reduce health care use and built by medical doctors with IT experience, usability testing was performed with peers. Due to its success, it was implemented in other hospitals. No major changes occurred during the study period.

XSL•FO
RenderX

Figure 2. English and Dutch in-app screenshots of the Virtual Fracture Care app used in the direct discharge protocol.





Study Population

Patients who met inclusion criteria and consented to participate in the surveys in the VFC app were included. Based on the annual number of patients (1200 - 1800 per hospital), and the inclusion time of 3 months per hospital, the estimated number of eligible patients was between 900 and 1350. Parents were asked to fill out the surveys if patients were 12 years or younger. If patients were between 13 and 16 years of age, patients and parents were allowed to complete the surveys. Patients older than 16 years could fill out the surveys alone. The exclusion criteria for this study were the same as the previously mentioned exclusion criteria for DD.

Sampling and Recruitment

Before discharge from the ED, in a face-to-face setting, eligible patients were asked to download and open the VFC app. An in-app pop-up asked for informed consent to participate in 2 surveys and an interview. After informed consent, patients were redirected to a Research Electronic Data Capture (REDCap) environment, a web-based survey system, to fill out the primary survey and were given an opt-out form by the treating physician, [12]. Additional information was given about the study and withdrawal methods. In the REDCap environment, all participants who started the survey were invited to participate in the interview by clicking a button and providing their email addresses. Age, sex, type of injury, and hospital were used to select a purposive sample among the quantitative population for interview patients. Eligible patients were contacted a

XSL•FO
RenderX

minimum of 6 weeks after injury to schedule a semistructured interview. During data collection, authors WDL and E Mathijsen considered whether the qualitative data had gained sufficient depth to perform a thorough analysis. Patients were reminded by email to complete the survey, and the second survey was sent 3 months post injury. Additional patient data were collected from electronic patient records.

Data Collection

Data were collected from surveys, semistructured interviews, and system data.

Surveys

Two surveys, 1 directly after the ED visit and one 3 months after the ED visit, with 63 questions, including close-ended questions, multiple-choice questions, 5-point Likert Scales, visual analog scales, and free-text questions, measured 5 Bowen feasibility parameters (Multimedia Appendix 3) [8]. As no golden standard for the evaluation of innovations exists, the surveys and topic list were developed by 4 researchers (JS, GW, BT, and THG) and checked by 2 experts on relevance: a professor in trauma surgery (JC Goslings) and an associate professor in-process evaluations of health care innovations (JCA Trappenburg). We pretested the survey with 5 patients to improve clarity.

Semistructured Interviews

Two independent researchers specialized in qualitative research, not involved in daily clinical practice or the VFC research team, conducted digital semistructured interviews to minimize social-desirability bias. The interviews were held within 6 and 10 weeks post injury to warrant an optimal recall. The interviews were guided by a topic list based on literature, including previously mentioned Bowen feasibility parameters (Multimedia Appendix 4) [8]. The research team piloted the topic list for clarity and completeness and modified it during data collection.

System Data

Quantitative data were extracted from the electronic patient record after 3 months of follow-up. The patient characteristics included compliance to therapy, complications (yes/no), type of complications, follow-up (yes/no), type of follow-up, and imaging. Downtime from the app was extracted from the log record of the VFC app.

Data Analysis

Quantitative data were analyzed using the SPSS (version 27; IBM Corporation) [13]. Baseline characteristics were reported descriptively using numbers and proportions for categorical variables and mean with SD or median with IQR as appropriate. The normal distribution of continuous data was assessed with visual analysis. Discrete variables were reported as numbers (percentages of the total population). The paired *t* test or the Mann-Whitney *U* test was used to determine the statistical significance of parametric variables for normally and nonnormally distributed data.

Qualitative data were analyzed according to the principles of the 6 phases of thematic analysis by Braun and Clark [14]. We have used an inductive, categorical approach because the triangulation process started after 4 phases. Data analysis started after the first 5 interviews. Interviews were audiotaped, transcribed verbatim, and analyzed using the software program NVivo (QSR International) [14]. One researcher (WL) independently analyzed the data, and another researcher (E Mathijsen) reviewed the analysis. Two researchers (WL and E Mathijsen) used several methods to ensure reliability and validity [15-17]. Discrepancies and remarks were discussed until they reached a consensus about data interpretations. Memos were made to track research decisions during analysis. Code saturation was reached when no new categories or themes emerged from the new raw data [15,16]. Instead of steps 5 and 6 according to Braun and Clark [14], we organized the data per theme during the triangulation session. The final themes were used to describe the value and feasibility of DD from the perspective of patients.

Triangulation

After the separate quantitative and partly qualitative analyses, findings were triangulated with a simplified approach of the Pillar Integration Process technique [9]. This approach uses a transparent and rigorous 4-stage technique for integrating and presenting qualitative and quantitative findings in a joint display Microsoft Excel, 2018 [18]. One of the researchers presented the quantitative findings (JS) per study parameter, and another the qualitative findings (WDL). (Dis)similarities and self-contained themes were objectified. One of the researchers (E Mathijsen) merged these themes into a meaningful narrative (the pillar), reviewed by researchers JS and WDL.

Ethical Considerations

This study, including the process analysis, was reviewed and approved by the Medical Ethical Committee of Utrecht, Netherlands (W21.261). Patients provided consent for participation in the research and could opt out at any time after request by email. The original consent and institutional review board approval covers secondary analysis without additional consent. A data key is stored at the local hospitals in a secured map and coded file. This is only accessible to JS and GW. The accessible data have been deidentified as far as possible (eg, age in years instead of the date of birth). Patients received no compensation to participate in this research.

Results

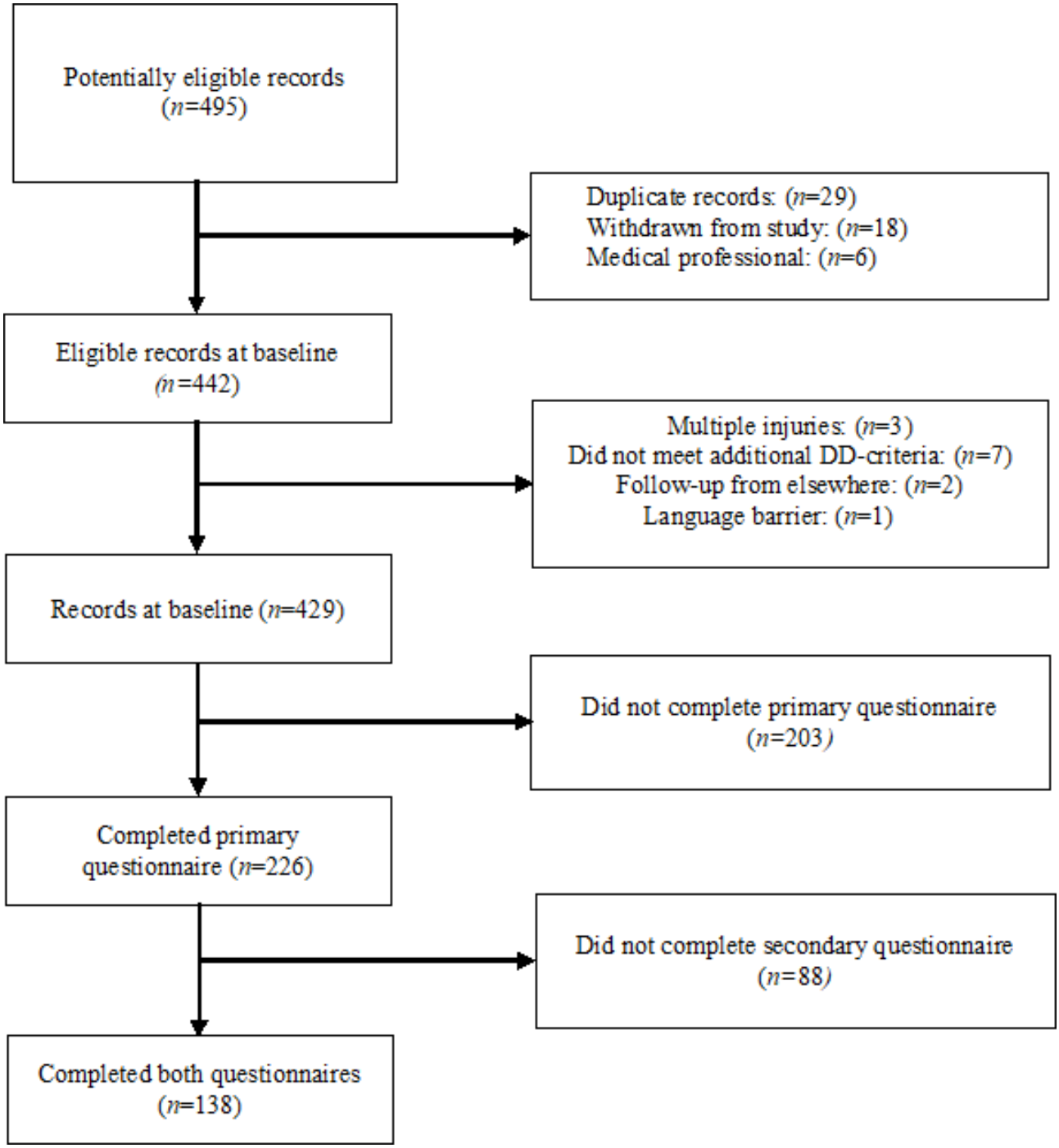
Demographics

Of the 429 patients who started the primary survey, 203 did not provide any data or contact details and did not complete the first survey and 88 did not complete both surveys (Figure 3). Of the 138 patients that completed both surveys (response rate: 138/429, 32%), 83 out of 138 (60%) were female, and the median age was 50 years (IQR 12 to 61) (Table 1). Patients who provided contact details at baseline varied significantly from responders regarding age (*P*=.01) but not sex (*P*=.14). Most patients were native Dutch speakers, who had attended primary school in the Netherlands (136/198, 98%), and over half had a minimum of a bachelor's degree (82/198, 59%) (Table 1). In addition, 18 patients sampled from the quantitative source participated in a web-based semistructured interview, of which

XSL•FO
RenderX

15 (83%) patients were female and 6 (33%) patients were parents of children (Table 2).

Figure 3. Flow diagram of eligible patients to evaluate the DD protocol. DD: direct discharge.



XSL•FO
RenderX

Table . Baseline characteristics of included patients in quantitative analysis of direct discharge.

Characteristic	Included in quantitative analysis (n=138)
Age (years), median (IQR)	50 (12 - 61)
Age younger than 18 years, n (%)	51 (37)
Sex (female), n (%)	83 (60)
Country of primary school, n (%)	
Netherlands	136 (98)
Philippines	1 (1)
Italy	1 (1)
Primary language, n (%)	
Dutch	136 (98)
Filipino	1 (1)
Italian	1 (1)
Highest level of education, n (%)	
Elementary school	4 (3)
High school	21 (15)
Vocational school	28 (20)
College	50 (36)
University	32 (23)
Other	3 (2)
Hospital treated, n (%)	
Center A	54 (39)
Center B	26 (19)
Center C	58 (42)

XSL•FO
RenderX

Table . Demographics of interviewed patients or parents regarding experiences with the direct discharge protocol.

ID	Sex	Type of injuries	Year of birth
#1	Female	Mallet finger	1969
#2	Female	Weber A Fx[a] or avulsion Fx	1971
#3	Female	Weber A Fx or avulsion Fx	2000
#4	Male	Fifth metatarsal Fx	1956
#5	Female	Radial head or neck Fx	1957
#6	Female	Ankle distortion	1992
#7	Female	Radial head or neck Fx	1963
#8	Female	Torus Fx of the radius	2011
#9	Male	Radial head or neck Fx	1959
#10	Female	Hallux Fx	1964
#11	Male	Fifth metatarsal Fx	1980
#12	Female	Hallux Fx	2010
#13	Female	Spoke injury	2017
#14	Female	Greenstick Fx	2014
#15	Female	Radial head or neck Fx	1974
#16	Female	Torus Fx	2014
#17	Female	Torus Fx	2011
#18	Female	Weber A Fx or avulsion Fx	1956

[a]Fx: fracture.

Implementation

Both data sources indicated that most patients were satisfied with their ED visit and the introduction of the app, despite some mentioning the hectic ED environment and difficulty downloading the app due to poor Wi-Fi connection. The physiological distress, and the hectic ED environment, resulted in an inability to recollect all the information provided by the ED physician. Several patients pointed out that the information in the VFC app was a valuable supplement to their ED visit (quote 1). Mobile app stability was excellent, with only 1 patient reporting issues accessing the app. The helpline could not be reached in 4 cases, resulting in 2 patients contacting the ED directly. The app was available 99.6% of the time, with the only downtime (5 hours in 3 months) caused by a software problem that was fixed by the app builder.

> *Because… it always goes quickly. It is always busy. But still, you know, this was completely clear; what I could expect and that I could download the app.* [quote 1; ID #10]

Acceptation

The median satisfaction with treatment was 7.8 (IQR 6.6-8.8) on a 10-point scale. Qualitative data complemented quantitative data, as most patients mentioned that they were satisfied with DD (quotes 2 and 3).

> *I would give DD a 7 or 8. Yeah, let's say 7, because I do feel that it might be difficult for older people. Especially because they don't always understand technology, you know.* [quote 2; ID #15]

> *To be honest, I think it's better. You know, often it's like, you go to a hospital, and it takes half a day just to get there, and come back, and all those things, and then it's just like: "O, it looks fine."* [quote 3; ID #5]

VFC App Acceptation

Most patients (106/138, 77%) reported that the quality of the information in the app was good and reported that the information answered the questions they had during their recovery (82/138, 58%). Additionally, 59% (79/138) would prefer the app over face-to-face follow-up if they were to have a similar injury. However, in qualitative data, some patients expressed concerns that the app may not be suitable for people with limited digital skills. Even though none of the patients identified themselves as "vulnerable," many reported that the app may not be appropriate for vulnerable individuals. Patients expressed that DD should not be mandatory for all in order to protect potentially vulnerable patients for whom DD would not be suitable.

Brace Acceptation

A total of 122 patients out of 138 (89%) used the brace for the prescribed period. Most patients (116/138, 84%) removed the brace or the sling during showering, at night (118/138, 86%), and (72/138, 52%) during non–weight bearing exercises. Most patients (97/138, 71%) found the brace comfortable, and 94 out of 138 patients (68%) found the brace convenient for these types of injuries (Multimedia Appendix 5). These findings were

XSL•FO
RenderX

complementary to the qualitative data. Patients preferred the brace to a cast as it allowed for better mobility because it was less bulky and less rigid compared with a cast (quote 4). However, the increased mobility made it tempting to overexert oneself, resulting in increased pain and insecurity for some patients (quote 5). The smaller size of the braces made some injuries seem less serious and burdensome, reportedly leading to social pressure to return to work earlier than advised by the doctor or app. Additionally, the less rigid nature of the brace resulted in increased skin friction, causing a superficial ulcer in 1 patient in the qualitative data.

> *I really like the brace because it allowed me to move. So, I was not stuck with a bulky cast on my leg, but rather, I had a lot of flexibility. I could actually determine what I wanted to do or not. So, I found that very enjoyable.* [quote 4; ID #2]

> *The brace was sometimes the reason I went over my limit. I could move more and was in the process of moving to a new home. But then the pain came back. I found it very difficult to 'guard' my limits.* [quote 5; ID #18]

Helpline and Perceived Safety

Most patients found the helpline important (116/138, 84%) and that the helpline offered them a sense of security (86/138, 62%). For some patients, the helpline was a way of checking if their recovery was still on track (quote 6). Most patients (94/138, 68%) in both data sources expressed that treatment with a brace and a helpline is safe for these types of injuries. However, some expected that if injuries were more severe, they would need more assistance than a brace and a smartphone app (quote 7).

Self-Empowerment

An increase in perceived self-empowerment was reported in 67 out of 138 (49%) of all patients, 51 (37%) patients reported neutral results, and 20 (14%) patients reported no increase. Enhanced treatment engagement was reported in 58 (42%) patients, while 50 (36%) patients reported neutral results, and 31 out of 138 (23%) patients reported no enhanced engagement (Multimedia Appendix 6). In the qualitative data, some of the interviewed patients stated that this type of treatment provided more self-empowerment and therefore, more control in their recovery (quote 8).

> *After about a week or three, I had a setback. I couldn't find this in the app, so I decided to call the helpline. The pain came back, and I was afraid I had broken something or something like that. They reassured me that it could not happen so quickly and that I just needed to rest for 24 hours h. They were right! It was nice to be able to check this.* [quote 6; ID #11]

> *Yes, I believe it is safe, given the circumstances. Because it wasn't that serious. Yes. I do have the confidence that if it is something serious I won't receive this type of treatment.* [quote 7; ID #17]

> *It was nice that I could read what I was allowed to do and what I was not. I think that gave me more control over my recovery. I knew what I could do myself in terms of exercises, and that was very helpful.* [quote 8; ID #2]

Preliminary Efficacy

Secondary Health Care Use

A total of 10 out of 138 (7%) patients had a face-to-face follow-up, and 9 (7%) patients by phone. Two patients attended the ED again after discharge due to anxiety and pain at the fracture site and were scheduled for outpatient follow-up. Two patients received follow-up for a wound check and 2 received follow-up as decided by the medical specialist due to severe pain complaints at the ED.

Functional Outcome

More than half of the patients had fully recovered 3 months after injury in terms of daily activities, sports, and work (84/138, 61%). Between 31% and 40% of patients were limited in functional outcomes once or twice a week. Approximately one in 10 patients were limited in physical activities more than 3 times a week. (Multimedia Appendix 7).

Pain

Most (92/138, 67%) patients have used painkillers in the first 3 weeks of recovery. Of these patients, 70 out of 92 (76%) patients have used acetaminophen, 17 (12%) patients have used nonsteroidal anti-inflammatory drugs, and 5 (4%) patients have used other undefined analgesics. Few patients (7/92, 5%) used cooling of the injury site to reduce pain. Analgesic use in the second week was lower (35/92, 38%) and further declined in the third week (9/92, 10%) and in the fourth week or after (9/92, 10%). Qualitative data showed limited pain complaints after immobilization of the initial trauma. Pain complaints in the first weeks were treated with analgesics. After the immobilization, a few patients had persistent pain symptoms. Most of these patients sustained a Weber A ankle fracture or ankle distortion. Those patients were most limited in their daily activities, specifically during more intense physical activities (eg, labor and sports) (quote 9).

> *The eight weeks of recovery are over, I believe. So now I should be able to start exercising again, well I do karate but I'll wait a little longer for that.* [quote 9; ID #4]

Demand

Almost all patients used the app during recovery (133/138, 95%) with a median of 4 times (IQR 2-6.5, range 0 - 30). Reasons to use the app included checking recovery exercises (91/138, 68%), treatment rules (69/133, 51%), the phase of recovery (68/133, 51%), seeking helpline information (57/133, 42%), and analgesics (10/133, 8%). Qualitative data supported the quantitative data. App use was focused on the first weeks post injury, and parents consulted the app more than children. They occasionally showed it to their child, predominantly if the information contained photos and videos (quote 10). Among the interviewed patients, almost all patients reported not requiring face-to-face follow-up for these types of injuries. The biggest advantage of DD is that it is time-saving on a personal level. Nevertheless, a few patients expressed a preference for outpatient follow-up. Especially if they had persistent complaints

XSL•FO
RenderX

or when the recovery was slower than expected. Patients expressed that this led to insecurity, which was also expressed by patients who used the app more frequently or later in the recovery phase (quote 11).

A minority of patients reported a lack of human contact and the physicians' reassurance of adequate recovery. Subsequently, as a minimal substitute, these patients suggested a feedback system (eg, pain scores or communication tools) to assist them further.

> *As for the app, he did see it, but you know, whatever! The only thing they find really interesting are pictures. What I liked that in the app is that they actually showed what was going on and what it looks like.* [quote 10; ID #17]

> *I was constantly looking for confirmation in the app, online, or at the fracture helpline.* [quote 11; ID #11]

Discussion

Principal Findings

This study shows that patients consider DD a feasible and safe alternative to traditional treatment, with a favorable perception of its acceptability, efficacy, applicability, and demand. Nevertheless, response rates were relatively low, and personal nuances and preferences must be considered when implementing DD.

Comparison With Literature

Patient satisfaction and perceived safety with DD aligned with previous studies [3,19,20]. Most patients responded positively to the introduction of DD at the ED, brace treatment, and assistance with the VFC app and helpline. The VFC app proved valuable in overcoming low recall of verbal information due to the chaotic ED environment and psychological distress. This finding is consistent with a systematic review highlighting the benefits of additional visual and written information in enhancing recall and injury knowledge [21]. Adequate understanding of the injury is crucial for properly following self-care protocols and monitoring red flags, this aligns with the finding that almost all patients used the app at least once. The braces used with DD were well-received, offering advantages such as easy removal (eg, during or at night), and improved daily living activities due to the early weight bearing. Braces seem to affect patients positively, as the lack of these advantages has been reported as most burdensome during cast immobilization [22]. However, the perceived decreased severity of the injury may pose a risk of overexertion and require further research.

Preliminary efficacy results align with previous research, demonstrating low complication rates, secondary health care use, comparable functional outcomes, and patient satisfaction with treatment [4,5,23]. Most patients preferred the VFC app over face-to-face follow-up for these injuries. While using apps in orthopedic and trauma surgery is not new, adding a self-care app to DD and virtual fracture clinics is a novel approach [24-26]. Patients found the information quality good and appreciated the time-saving component (eg, reducing work absenteeism). However, in the case of more severe injuries with wounds or complex fractures, some patients would prefer face-to-face interaction for additional reassurance. As previously suggested, digital self-care could be combined with face-to-face follow-up or used for preappointment education in patients with more severe injuries [27,28].

Although a good fit for most, some (young) patients noted that the app might not be suitable for older individuals or those with limited digital skills, potentially increasing health inequities in an already digitally oriented world [29,30]. However, it is important to note that DD targets relatively young patients, who are considered capable by the treating physician.

While previous studies have highlighted the potential for increased self-empowerment and patient engagement with eHealth, this study did not explicitly confirm it [31-33]. Despite positive study results, individual nuances in patients and injury types require a continuous assessment to ensure personalized patient care.

Information demand and app use varied among patients, implying different levels of demand among users. A minority of patients expressed a desire for more human contact and reassurance. A possible suggested solution was developing an in-app numerical feedback system, such as a questionnaire or communication tool.

Strengths and Limitations

This study has several strengths. To our knowledge, it is the first to provide in-depth interviews with patients treated using DD, providing valuable insights. Second, the research team's multidisciplinary involvement ensured a comprehensive evaluation and analysis of the data from multiple perspectives. Thirdly, the mixed method approach, including the separate collection of both data sources until data saturation, combined with the triangulation, increased the likelihood of realistic and rigorous results. Additionally, using a validated framework provided a structured insight into each feasibility parameter.

However, limitations include potential responder and selection bias due to a younger sample of patients with higher education levels and Dutch literacy rates than the general Dutch population. This age difference could be caused by the 12 selected injuries, of which half only occur in pediatric patients. Furthermore, the response rate of this study was 32%. Response rate, literacy, and education may limit the generalizability of the findings to a broader Dutch population, especially those with lower health literacy or digital skills. Furthermore, the timing to measure functional outcomes might have been suboptimal in these patients, as for some injuries (eg, mallet finger injuries), the immobilization period had just ended, resulting in limitations in daily activities.

Implications for Clinicians and Policy Makers, and Future Perspectives

DD has emerged as a promising approach to reduce outpatient follow-up while maintaining positive effects on primary health care use, patient satisfaction, complications, and functional outcomes [4,5]. By delivering follow-up care close to home, associated health care costs and societal costs decrease [34]. In addition, it limits unnecessary patient travel to the hospital thereby reducing the environmental impact of health care.

XSL•FO
RenderX

Including patients' perspectives in evaluating new care pathways, whether digitally assisted or not, is crucial for sustainable adoption and quality of care. Clinicians, researchers, and policy makers should prioritize patient involvement throughout the design, prototype, pilot, and evaluation phases. "Design thinking," a validated approach widely recognized in user experience and implementation literature, can be used to design these pathways. For example, replacing hospital care with DD involves changes in location and care deliverers, which presents new needs, challenges, and opportunities suitable to solve with design thinking. This study has identified areas for improvement of DD in terms of functions and features, and adjusted language to lower literacy. Future studies should focus on co-designing in-app feedback systems that address patient and health care professional needs for reassurance and monitoring like communication tools or questionnaires.

Conclusion

Patients consider DD a feasible and safe alternative to traditional treatment, with a favorable perception of its acceptability, efficacy, applicability, and demand. Nevertheless, response rates were relatively low, and personal nuances and preferences must be considered when implementing DD. Clinicians and policy makers can use the insights to improve DD and work towards the integration of DD into clinical practice and future guidelines.

Acknowledgments

The authors wish to thank Dr JCA Trappenburg, associate professor at the Julius Center from the University Medical Center of Utrecht, and Drs WL and E Mathijsen from the Healthcare Innovation Center (THINC) for the assistance in the development of the surveys and collecting and analyzing of the qualitative data. The institutions of the authors affiliated with St Antonius Hospital and OLVG Hospital have received funding to support this work through an unrestricted grant from the Dutch organization for health care sciences and health care innovation, ZonMw, the Hague, the Netherlands (grant 516012524).

The Colloboration Group is: Dr N Sosef, Dr T L Groenesteege, and Dr J G Ten Brinke.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Good Reporting on a Mixed Methods Study (GRAMMS) criteria for the study: "Patient Experiences With a Mobile Self-care Solution for Low-complex Orthopedic Injuries: Mixed Methods study".
[DOCX File, 26 KB - humanfactors_v12i1e53074_app1.docx]

Multimedia Appendix 2
Additional criteria and immobilization for treatment of low-complex, traumatic orthopedic injuries with the direct discharge protocol.
[DOCX File, 13 KB - humanfactors_v12i1e53074_app2.docx]

Multimedia Appendix 3
Surveys used to evaluate direct discharge among health care professionals.
[DOCX File, 163 KB - humanfactors_v12i1e53074_app3.docx]

Multimedia Appendix 4
Topic list for health care professionals the direct discharge protocol.
[DOCX File, 15 KB - humanfactors_v12i1e53074_app4.docx]

Multimedia Appendix 5
5-point Likert scale distribution of acceptance-related outcomes regarding experiences with a brace.
[DOCX File, 129 KB - humanfactors_v12i1e53074_app5.docx]

Multimedia Appendix 6
Self-empowerment and perceived safety of care with direct discharge.
[DOCX File, 73 KB - humanfactors_v12i1e53074_app6.docx]

Multimedia Appendix 7
Frequency of limitation per week in physical function, activities of daily living, and school or work in patients treated with direct discharge at 3 months follow-up.

XSL•FO
RenderX

[DOCX File, 15 KB - humanfactors_v12i1e53074_app7.docx]

References

1. Goslings JC, Gorzeman MP, M. Offeringa-Klooster JB. Brandbrief 'Regionale spoedzorg–de rek is er uit [Report in Dutch]. : TraumaNet AMC; 2016.
2. van der Linden C, Reijnen R, Derlet RW, et al. Emergency department crowding in the Netherlands: managers' experiences. Int J Emerg Med 2013 Oct 24;6(1):41. [doi: 10.1186/1865-1380-6-41] [Medline: 24156298]
3. Murphy EP, Fenelon C, Murphy RP, et al. Are virtual fracture clinics during the COVID-19 pandemic a potential alternative for delivering fracture care? A systematic review. Clin Orthop Relat Res 2020 Nov;478(11):2610-2621. [doi: 10.1097/CORR.0000000000001388] [Medline: 32657810]
4. Geerdink TH, Verbist J, van Dongen JM, Haverlag R, van Veen RN, Goslings JC. Direct discharge of patients with simple stable musculoskeletal injuries as an alternative to routine follow-up: a systematic review of the current literature. Eur J Trauma Emerg Surg 2022 Aug;48(4):2589-2605. [doi: 10.1007/s00068-021-01784-z] [Medline: 34529086]
5. Khan SA, Asokan A, Handford C, Logan P, Moores T. How useful are virtual fracture clinics? Bone Joint Open 2020 Nov 1;1(11):683-690. [doi: 10.1302/2633-1462.111.BJO-2020-0107.R1]
6. Geerdink TH, Haverlag R, van Veen RN, Bouwmeester OVA, Goslings JC. Direct discharge from the ED for patients with simple stable injuries: a Dutch pilot study. Ned Tijdschr Geneeskd 2020 Feb 25;164:D4604. [Medline: 32391996]
7. Fracture Clinic Redesign Programme. URL: http://www.fractureclinicredesign.org/visit/ [accessed 2023-03-01]
8. Bowen DJ, Kreuter M, Spring B, et al. How we design feasibility studies. Am J Prev Med 2009 May;36(5):452-457. [doi: 10.1016/j.amepre.2009.02.002] [Medline: 19362699]
9. Johnson RE, Grove AL, Clarke A. Pillar integration process: a joint display technique to integrate data in mixed methods research. J Mix Methods Res 2019 Jul;13(3):301-320. [doi: 10.1177/1558689817743108]
10. O'Cathain A, Murphy E, Nicholl J. The quality of mixed methods studies in health services research. J Health Serv Res Policy 2008 Apr;13(2):92-98. [doi: 10.1258/jhsrp.2007.007074] [Medline: 18416914]
11. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126. [doi: 10.2196/jmir.1923] [Medline: 22209829]
12. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381. [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
13. IBM SPSS Statistics. IBM. URL: https://www.ibm.com/products/spss-statistics [accessed 2025-03-06]
14. NVivo 14 Windows. URL: https://help-nv.qsrinternational.com/14/win/Content/welcome.htm [accessed 2025-02-25]
15. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
16. Morse JM, Barrett M, Mayan M, Olson K, Spiers J. Verification strategies for establishing reliability and validity in qualitative research. Int J Qual Methods 2002 Jun;1(2):13-22. [doi: 10.1177/160940690200100202]
17. Boeije H. Analysis in Qualitative Research: Sage Publishing LTD; 2009.
18. Microsoft Excel. Microsoft. URL: https://office.microsoft.com/excel [accessed 2025-02-25]
19. Geerdink TH, Spierings JF, Schuijt HJ, et al. Healthcare utilization and satisfaction with treatment before and after direct discharge from the Emergency Department of simple stable musculoskeletal injuries in the Netherlands. Eur J Trauma Emerg Surg 2022 Jun;48(3):2135-2144. [doi: 10.1007/s00068-021-01835-5] [Medline: 34997258]
20. Mackenzie SP, Carter TH, Jefferies JG, et al. Discharged but not dissatisfied: outcomes and satisfaction of patients discharged from the Edinburgh Trauma Triage Clinic. Bone Joint J 2018 Jul;100-B(7):959-965. [doi: 10.1302/0301-620X.100B7.BJJ-2017-1388.R2]
21. Hoek AE, Anker SCP, van Beeck EF, Burdorf A, Rood PPM, Haagsma JA. Patient discharge instructions in the emergency department and their effects on comprehension and recall of discharge instructions: a systematic review and meta-analysis. Ann Emerg Med 2020 Mar;75(3):435-444. [doi: 10.1016/j.annemergmed.2019.06.008] [Medline: 31439363]
22. McKeown R, Kearney RS, Liew ZH, Ellard DR. Patient experiences of an ankle fracture and the most important factors in their recovery: a qualitative interview study. BMJ Open 2020 Feb 4;10(2):e033539. [doi: 10.1136/bmjopen-2019-033539] [Medline: 32024789]
23. Davey MS, Coveney E, Rowan F, Cassidy JT, Cleary MS. Virtual fracture clinics in orthopaedic surgery - a systematic review of current evidence. Injury 2020 Dec;51(12):2757-2762. [doi: 10.1016/j.injury.2020.11.001] [Medline: 33162011]
24. Dittrich F, Back DA, Harren AK, et al. A possible mobile health solution in orthopedics and trauma surgery: development protocol and user evaluation of the Ankle Joint app. JMIR Mhealth Uhealth 2020 Feb 26;8(2):e16403. [doi: 10.2196/16403] [Medline: 32130171]
25. Houwen T, Vugts MAP, Lansink KWW, et al. Developing mHealth to the context and valuation of injured patients and professionals in hospital trauma care: qualitative and quantitative formative evaluations. JMIR Hum Factors 2022 Jun 20;9(2):e35342. [doi: 10.2196/35342] [Medline: 35723928]

XSL•FO
RenderX

26. Van Reijen M, Vriend II, Zuidema V, van Mechelen W, Verhagen EA. The implementation effectiveness of the "Strengthen your ankle" smartphone application for the prevention of ankle sprains: design of a randomized controlled trial. BMC Musculoskelet Disord 2014 Jan 7;15:2. [doi: 10.1186/1471-2474-15-2] [Medline: 24393146]
27. Timmers T, Janssen L, van der Weegen W, et al. The effect of an app for day-to-day postoperative care education on patients with total knee replacement: randomized controlled trial. JMIR Mhealth Uhealth 2019 Oct 21;7(10):e15323. [doi: 10.2196/15323] [Medline: 31638594]
28. Timmers T, Janssen L, Kool RB, Kremer JAM. Educating patients by providing timely information using smartphone and tablet apps: systematic review. J Med Internet Res 2020 Apr 13;22(4):e17342. [doi: 10.2196/17342] [Medline: 32281936]
29. Yao R, Zhang W, Evans R, Cao G, Rui T, Shen L. Inequities in health care services caused by the adoption of digital health technologies: scoping review. J Med Internet Res 2022 Mar 21;24(3):e34144. [doi: 10.2196/34144] [Medline: 35311682]
30. Lau AYS, Piper K, Bokor D, Martin P, Lau VSL, Coiera E. Challenges during implementation of a patient-facing mobile app for surgical rehabilitation: feasibility study. JMIR Hum Factors 2017 Dec 7;4(4):e31. [doi: 10.2196/humanfactors.8096] [Medline: 29217504]
31. Latulippe K, Hamel C, Giroux D. Social health inequalities and eHealth: a literature review with qualitative synthesis of theoretical and empirical studies. J Med Internet Res 2017 Apr 27;19(4):e136. [doi: 10.2196/jmir.6731] [Medline: 28450271]
32. Lupton D. Digitized health promotion: risk and personal responsibility for health and illness in the Web 2.0 era. In: To Fix Or To Heal: Patient Care, Public Health, and the Limits of Biomedicine: NYU Press; 2016:152-176. [doi: 10.18574/nyu/9781479878246.003.0007]
33. Risling T, Martinez J, Young J, Thorp-Froslie N. Evaluating patient empowerment in association with eHealth technology: scoping review. J Med Internet Res 2017 Sep 29;19(9):e329. [doi: 10.2196/jmir.7809] [Medline: 28963090]
34. Geerdink TH, Geerdink NJ, van Dongen JM, et al. Cost-effectiveness of direct discharge from the emergency department of patients with simple stable injuries in the Netherlands. Trauma Surg Acute Care Open 2021;6(1):e000763. [doi: 10.1136/tsaco-2021-000763] [Medline: 34722930]

Abbreviations

DD: direct discharge
ED: emergency department
GRAMMS: Good Reporting on a Mixed Methods Study
REDCap: Research Electronic Data Capture
VFC: Virtual Fracture Care

Edited by A Kushniruk; submitted 25.09.23; peer-reviewed by C Tang, EVDV Isselt; revised version received 28.06.24; accepted 08.11.24; published 14.03.25.

Please cite as:
Spierings J, Willinge G, Kokke M, Repping S, de Lange W, Geerdink T, van Veen R, van der Velde D, Goslings JC, Twigt B, Collaboration Group
Patient Experiences With a Mobile Self-Care Solution for Low-Complex Orthopedic Injuries: Mixed Methods Study
JMIR Hum Factors 2025;12:e53074
URL: https://humanfactors.jmir.org/2025/1/e53074
doi:10.2196/53074

© Jelle Spierings, Gijs Willinge, Marike Kokke, Sjoerd Repping, Wendela de Lange, Thijs Geerdink, Ruben van Veen, Detlef van der Velde, J Carel Goslings, Bas Twigt, Collaboration Group. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 14.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effect of SMS Ward Round Notifications on Inpatient Experience in Acute Medical Settings: Retrospective Cohort Study

Jongchan Lee[1], MD; Soyeon Ahn[2], PhD; Jung Hun Ohn[1], MD, PhD; Eun Sun Kim[1], MD, PhD; Yejee Lim[1], MD, PhD; Hye Won Kim[1], MD, PhD; Hee-Sun Park[1], MD; Jae Ho Cho[3], MD; Sun-wook Kim[1], MD; Jiwon Ryu[1], MD; Jihye Kim[1], MD; Hak Chul Jang[1], MD, PhD; Nak-Hyun Kim[1], MD, PhD

[1]Department of Internal Medicine, Seoul National University Bundang Hospital, Seoul National University College of Medicine, 82, Gumi-ro 173 Beon-gil, Bundang-gu, Gyeonggi-do, Seongnam-si, Republic of Korea

[2]Department of Statistics, Seoul National University Bundang Hospital, Seoul National University College of Medicine, Seongnam-si, Republic of Korea

[3]Department of Hospital Medicine, Seoul St. Mary's Hospital, Seoul, Republic of Korea

Corresponding Author:
Nak-Hyun Kim, MD, PhD
Department of Internal Medicine, Seoul National University Bundang Hospital, Seoul National University College of Medicine, 82, Gumi-ro 173 Beon-gil, Bundang-gu, Gyeonggi-do, Seongnam-si, Republic of Korea

Abstract

Background: Ward rounds are an essential component of inpatient care. Patient participation in rounds is increasingly encouraged, despite the occasional complicated circumstances, especially in acute care settings.

Objective: This study aimed to evaluate the effect of real-time ward round notifications using SMS text messaging on the satisfaction of inpatients in an acute medical ward.

Methods: Since January 2021, a service implementing real-time ward round notifications via text messaging (WR-SMS) has been operational at a tertiary-care medical center in Korea. To assess its impact, we conducted a retrospective cohort study of patients admitted to the acute medical unit who participated in a patient experience survey. Patient satisfaction was compared between patients admitted in 2020 (pre–WR-SMS group) and 2021 (post–WR-SMS group).

Results: From January 2020 to December 2021, a total of 100 patients were enrolled (53 patients in the pre–WR-SMS group and 47 patients in the post–WR-SMS group). Compared with the pre–WR-SMS group, the post–WR-SMS group showed significantly greater satisfaction about being informed about round schedules (mean 3.43, SD 0.910 vs mean 3.89, SD 0.375; $P<.001$) and felt more emotionally supported during admission (mean 3.49, SD 0.800 vs mean 3.87, SD 0.397; $P<.001$). Regarding other questionnaire scores, the post–WR-SMS group showed an overall, although statistically insignificant, improvement compared with the pre–WR-SMS group.

Conclusions: Real-time round notifications using a user-friendly SMS may improve inpatient satisfaction effectively.

(JMIR Hum Factors 2025;12:e57470) doi:10.2196/57470

KEYWORDS

rounds; round-time notification; text messaging; patient experience assessment; patient experiences; patient-centeredness; patient participation

Introduction

Inpatient ward rounds constitute a salient core activity of the daily care of hospitalized patients and present a critical opportunity to deliver patient-centered care [1,2]. Ward rounds are a complex process that involves organizing the clinical care of inpatients, including diagnostic assessment and therapeutic planning by the care team [3]. For the patient, inpatient ward rounds provide an opportunity for direct patient-clinician communication, sharing of information, and participation in their own care planning [3,4]. Despite the widely recognized importance of ward rounds, an environment conducive to conducting such rounds and effectively communicating with patients is not always achievable, resulting in ineffective communication with and dissatisfaction of patients [5]. Especially in acute care settings, the participation of patients in ward rounds is likely to be limited not only due to their acute illness and fatigue [6] but also by the acute care context itself and time constraints of the ward staff [7,8].

With the emphasis on patient-centeredness, the Korean Ministry of Health and Welfare (MOHW) and the Health Insurance Review and Assessment Service (HIRA) developed the Korean Patient Experience Survey, to gain insight into the needs of inpatients and improve inpatient care quality. This survey has

XSL•FO
RenderX

been conducted biennially by telephone since 2017 among patients discharged from hospitals, and the results have repeatedly revealed that patients remain least satisfied with inpatient ward rounds [9,10]. In line with global trends, studies from the United States and the United Kingdom highlight the importance of improving patient communication and providing clear, predictable information to enhance patient satisfaction [11-13]. Furthermore, evidence suggests that interventions focusing on real-time communication [14], such as the use of SMS text messaging, can help alleviate patient dissatisfaction by improving the timeliness and transparency of ward rounds.

As part of the efforts to improve patients' satisfaction, Seoul National University Bundang Hospital (SNUBH), a 1300-bed tertiary referral hospital in the Republic of Korea, has implemented the real-time ward round notifications via text messaging (WR-SMS) service since January 2021 to inform the patients and their carers about the start of daily ward rounds, and thereby alleviating arbitrary waiting for rounds.

In 2015, a hospitalist-run acute medical unit (AMU) was established to enhance patient safety and care efficiency for patients with acute medical conditions. Although the AMU improved patient outcomes and efficiency of care [15,16], patients' perceptions of inpatient care were found to be inconsistent, which could be attributed to the acuteness and complexity of clinical conditions, the hectic hospital environment, and heightened levels of anxiety.

As mobile devices continue to advance, SMS text messaging has been suggested to enhance clear and efficient communication in various clinical contexts [14,17]. However, there remains a gap in research regarding the potential impact of informing patients about round times via SMS text messages on communication effectiveness and patient satisfaction. This study aimed to determine the effect of real-time WR-SMS service on the satisfaction of AMU inpatients.

Methods

Study Design and Clinical Setting

This retrospective cohort study was conducted at the AMU of SNUBH, a hospitalist-run 46-bed ward caring for patients with acute medical conditions, who were admitted from the emergency or outpatient departments for active acute medical care, unless being hemodynamically unstable requiring invasive monitoring and critical care or terminally ill with cancer requiring only palliative care. After receiving active acute care, patients were either discharged or transferred to a specialty ward for further treatment. The AMU was operated 24 hours a day and 7 days a week by 10 board-certified medical hospitalists on a weekly rotation basis.

The SNUBH patient-experience survey has been routinely performed by telephone since 2006. Patients eligible for the survey were randomly selected within 2 to 56 days (8 wk) after discharge. Patients unable to communicate effectively due to conditions resulting in severe cognitive or communication impairments were excluded from the survey. Following the introduction of the official Korean Patient Experience Survey developed by the Korean MOHW/HIRA in 2017, the SNUBH patient-experience survey used the very same questionnaire to assess patients' hospital experiences.

Study Participants

Adult patients ≥19 years of age who were admitted to the AMU of SNUBH between January 2020 and December 2021 for ≥1 day and who participated in the SNUBH patient-experience survey after discharge were included. Their demographic data were retrospectively collected from the electronic medical record system, and their satisfaction was assessed based on the SNUBH patient-experience survey results. To determine whether patient satisfaction improved, we compared the scores from 2020 (before the implementation of the WR-SMS, pre–WR-SMS group) with those of 2021 (after the implementation of the WR-SMS post–WR-SMS group).

Real-Time Ward Round Notification Service With Text Messaging

The WR-SMS service was first developed in September 2020. It was actively implemented in January 2021, after an introductory period of 3 months (October to December 2020) during which the staff was educated about and encouraged to use the service via campaigns and email or SMS notifications.

The WR-SMS service was embedded into the SNUBH electronic health record system, enabling the doctors to send short text messages (80 to 160 bytes; ie, SMS) to their inpatients directly from the electronic health record system with a few clicks. From their patient list, doctors were able to select or deselect the patients who would receive the SMS text message. The default content of the SMS text message read, "Ward rounds by Dr. OOO will begin shortly. Visiting times may vary depending on the room." Further adjustments to the content of the message could be made as needed by the doctor. The notification could also be sent simultaneously to other relevant medical staff. The doctors were encouraged to send this WR-SMS to their patients just before starting their ward rounds, usually right after reviewing the patients' records.

Patient-Experience Survey

The SNUBH patient-experience survey questionnaire, which uses the same questions as the official Korean Patient Experience Survey developed by the Korean MOHW/HIRA, comprises 19 questions organized into 5 domains about the patients' hospital experience (4 questions pertaining to "Services from nurses," 4 pertaining to "Services from physicians," 5 pertaining to "Medication and treatment processes," 2 pertaining to "Hospital environment," 4 pertaining to "Ensuring patients' rights", and 4 assessing patients' demographic characteristics) (detailed questionnaire is presented in Multimedia Appendix 1). Each question was scored on a 4-point Likert scale (1=very dissatisfied, 2=dissatisfied, 3=satisfied, and 4=very satisfied) to assess patients' subjective satisfaction with the quality of inpatient services. Except for the question about providing information about plans after discharge, answers were either "Yes" or "No." For this study, the responses to the "Services from physicians" and "Medication and treatment processes" domain, which are closely related to inpatient ward rounds, were considered.

XSL•FO
RenderX

Statistical Analysis

Data reflecting patients' baseline characteristics were expressed as a median with ranges for continuous variables and as frequencies and percentages for categorical variables. Data from Likert scales were expressed as means and SDs, as well as median with ranges (IQRs). Comparisons were performed using the Mann-Whitney *U* test, Student *t* test, and χ^2 test. All tests were 2-sided and performed at a significance level of .05. Statistical analyses were performed using IBM SPSS v. 21.0 (IBM Corp).

Ethical Considerations

The study was conducted in accordance with the Declaration of Helsinki, revised in 2013. The SNUBH institutional review board approved this study design and waived the need for obtaining informed consent from the participants due to the retrospective design (B-2203-746-101). Participant confidentiality and anonymity were strictly maintained throughout the research process, and data were handled in a secure manner to protect their privacy. In addition, the study ensured that no undue harm or burden was placed on participants and that their rights, safety, and well-being were prioritized at all times.

Results

Baseline Characteristics of the Participants

A total of 100 patients were enrolled from January 2020 to December 2021 (53 in the pre–WR-SMS group before WR-SMS implementation and 47 in the post–WR-SMS group after WR-SMS implementation). The median age of the patients was 59.5 (range 25 - 81) years, and of the total, 59 were female. A total of 74 patients were hospitalized because of malignancy and 26 patients were admitted for non-cancer illnesses. A total of 44 and 56 patients were admitted from the emergency and outpatient departments, respectively. The median length of hospital stay was 7 (range 3 - 30) days. No statistically significant difference of patients' baseline characteristics between the pre- and post–WR-SMS groups was observed (Table 1).

Table . Baseline demographic and clinical characteristics of study participants.

Characteristics		Pre–WR-SMS group[a] (n=53)	Post–WR-SMS group[b] (n=47)	*P* value
Age (year), median (range)		58 (25-81)	60 (38-76)	.90
Sex, n (%)				.36
	Male	19 (36)	22 (47)	
	Female	34 (64)	25 (53)	
Disease, n (%)				.43
	Malignancy	37 (70)	37 (79)	
	Non-cancer illness	16 (30)	10 (21)	
Hospitalization route, n (%)				.94
	Emergency department	24 (45)	20 (43)	
	Outpatient department	29 (55)	27 (57)	
Education, n (%)				.25
	≤High school	29 (55)	32 (68)	
	≥College	24 (45)	15 (32)	
Length of stay (days), median (range)		7 (3-30)	7 (3-23)	.76

[a]Pre–WR-SMS group: patients hospitalized at the acute medical unit in 2020, before the implementation of ward round notifications via text messaging service.

[b]Post–WR-SMS group: patients hospitalized at the acute medical unit in 2021, after the implementation of ward round notifications via text messaging.

Effects of Real-Time Round Notification on Satisfaction of Hospitalized Patients

After the WR-SMS implementation in 2021, the score for the question pertaining to whether the patient had received information about round schedules improved significantly compared with 2020. Differences in the Likert scale for this question are illustrated in Figure 1. The proportion of patients who assigned a rating of 4 points (very satisfied) increased from 66% (35/53) in the pre–WR-SMS group to 92% (43/47) in the post–WR-SMS group, while the proportion of those who assigned a rating of 1 point (very dissatisfied) decreased from 6% (3/53) in the pre–WR-SMS group to 0% in in the post–WR-SMS group.

XSL•FO
RenderX

Figure 1. Changes in Likert scale scores regarding the provision of patient information on round schedules before and after the implementation of ward round notifications via text messaging. Pre–WR-SMS group: patients hospitalized at the acute medical unit in 2020, before the implementation of ward round notifications via text messaging service; Post–WR-SMS group: patients hospitalized at the acute medical unit in 2021, after the implementation of ward round notifications via text messaging service.

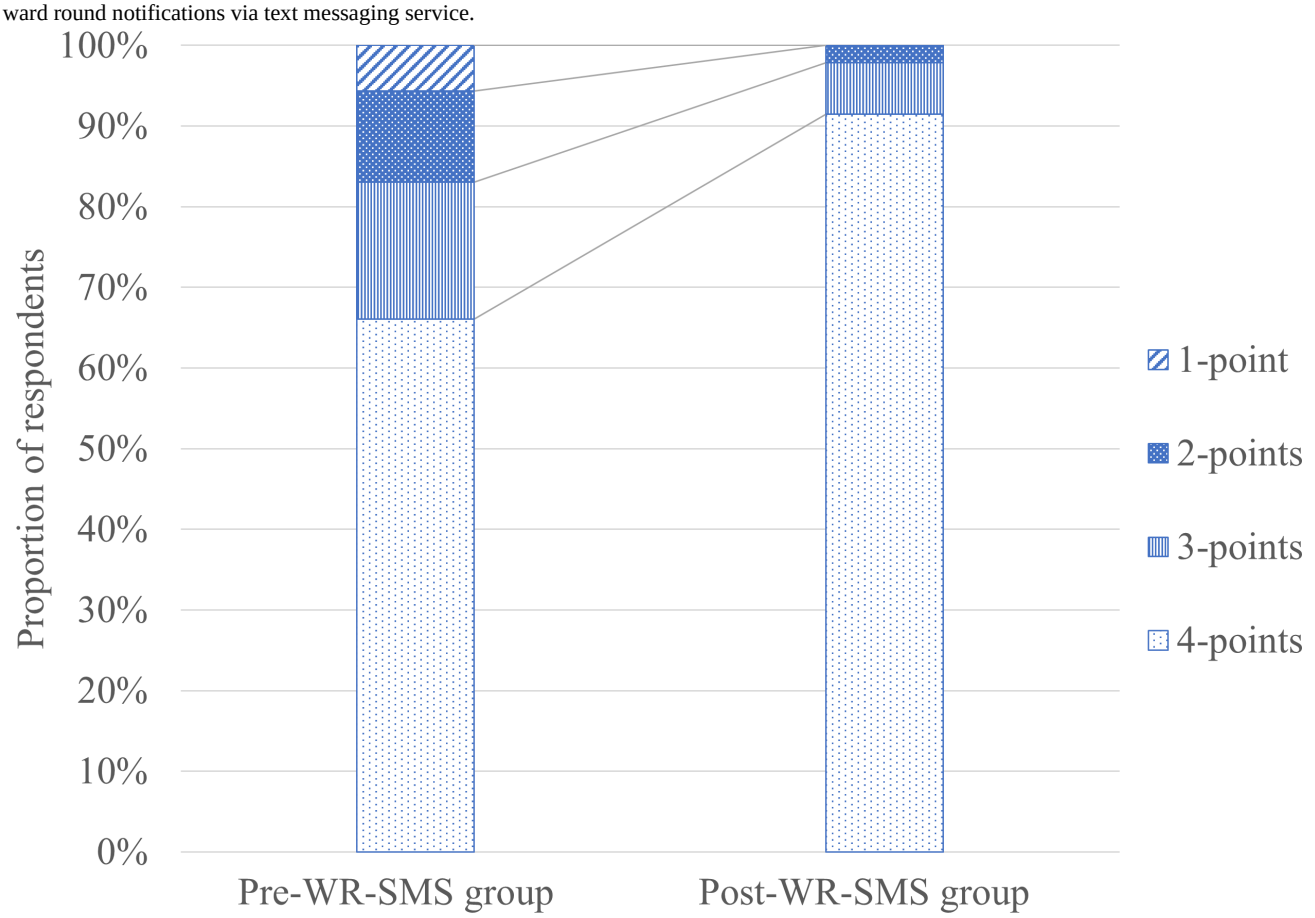


Table 2 summarizes the responses for the "Services from physicians" and "Medication and treatment processes" domains. Naturally, patients in the post–WR-SMS group showed significantly higher satisfaction about being informed about round schedules (mean 3.43, SD 0.910 vs mean 3.89, SD 0.375; *P*<.001) compared with patients in the pre–WR-SMS group. Interestingly, patients in the post–WR-SMS group also demonstrated significantly higher satisfaction with the emotional support (comfort and empathy) they received during admission (mean 3.49, SD 0.800 vs mean 3.87, SD 0.397; *P*<.001). Although not statistically significant, they were also satisfied with the description of the treatment processes (mean 3.62, SD 0.713 vs mean 3.83, SD 0.481; *P*=.09) and the explanations of possible side effects during treatment (mean 3.60, SD 0.743 vs mean 3.85, SD 0.465; *P*=.05). For all other questions in the "Services from physicians" and "Medication and treatment processes" domain, the post–WR-SMS group recorded a higher overall score compared to the pre–WR-SMS group, although it was not statistically significant (Multimedia Appendix 2).

XSL•FO
RenderX

Table . Survey results on patient satisfaction before and after the implementation of ward round notification using SMS text messaging service[a].

Parameter		Pre–WR-SMS group[b] (n=53)	Post–WR-SMS group[c] (n=47)	*t* test (df)	*P* value
Services from physicians					
	Courtesy and respect for patients			–0.93 (83.47)	
	Median (IQR)	4 (4-4)	4 (4-4)		.65
	Mean (SD)	3.81 (0.557)	3.89 (0.312)		.36
	Careful listening to patients			–0.47 (88.02)	
	Median (IQR)	4 (4-4)	4 (4-4)		.88
	Mean (SD)	3.83 (0.545)	3.87 (0.337)		.64
	Opportunities for patients to meet their doctors			–1.58 (94.08)	
	Median (IQR)	4 (3-4)	4 (4-4)		.15
	Mean (SD)	3.58 (0.663)	3.77 (0.476)		.12
	Receipt of information about round schedules			–3.37 (70.90)	
	Median (IQR)	4 (3-4)	4 (4-4)		.002
	Mean (SD)	3.43 (0.910)	3.89 (0.375)		<.001
Medication and treatment process					
	Detailed description of treatment processes			–1.72 (91.73)	
	Median (IQR)	4 (3-4)	4 (4-4)		.07
	Mean (SD)	3.62 (0.713)	3.83 (0.481)		.09
	Easy-to-understand explanation of potential side effects			–2.02 (88.60)	
	Median (IQR)	4 (3-4)	4 (4-4)		.045
	Mean (SD)	3.60 (0.743)	3.85 (0.465)		.047
	Appropriate pain relief management			–1.15 (61.49)	
	Median (IQR)	4 (4-4)	4 (4-4)		.23
	Mean (SD)	3.86 (0.462)	3.95 (0.216)		.26
	Comfort and sympathy regarding illness			–3.08 (78.06)	
	Median (IQR)	4 (3-4)	4 (4-4)		.003
	Mean (SD)	3.49 (0.800)	3.87 (0.397)		<.001
	Receipt of information on postdischarge care, n (%)	48 (90.6)	42 (89.4)	—[d]	.89

[a]Values in the first row for each item represent medians with IQRs and *P* values from the Mann-Whitney test. Values in the second row for each item represent means with SDs and *P* values from the Student *t* test (2-tailed).

[b]Pre–WR-SMS group: patients hospitalized at the acute medical unit in 2020, before the implementation of ward round notifications via text messaging service.

[c]Post–WR-SMS group: patients hospitalized at the acute medical unit in 2021, after the implementation of ward round notifications via text messaging service.

[d]Not applicable.

Discussion

Principal Findings

The WR-SMS implementation resulted in significant improvement of inpatient satisfaction among patients hospitalized in the AMU. Besides the direct improvement in scores pertaining to the receipt of round schedule information (mean 3.43, SD 0.910 vs mean 3.89, SD 0.375; *P*<.001), patients felt more emotionally supported and involved in the care process (mean 3.49, SD 0.800 vs mean 3.87, SD 0.397; *P*<.001), and also expressed higher satisfaction with the explanations provided

XSL•FO
RenderX

about treatment side effects (mean 3.60, SD 0.743 vs mean 3.85, SD 0.465; *P*=.047). These significant findings emphasize the value of real-time communication on both procedural and emotional aspects of care. Our results indicate that by enabling patients to participate more actively in routine ward rounds, WR-SMS fosters a greater sense of involvement and satisfaction, reaffirming the critical role of patient-centered communication in inpatient care.

Patient-centeredness is an essential component of high-quality health care, contributing to improved experience and clinical outcomes [18]. It is defined by the Institute of Medicine as the establishment of a partnership between the patient and health care providers to ensure that the wishes, needs, and preferences of the former are respected in the shared decision-making process [19]. Since the early 2000s, many countries have implemented patient-reported experience measures (such as standardized surveys) at the national level. The data are publicly reported, in order to assess and monitor patients' health care experiences and improve the quality of care [11,12]. For example, the United States Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) survey, first implemented in 2006, measures patients' perspectives on the quality of care they receive during their hospital stay, with its scores influencing hospital payment [13]. The United Kingdom and the Netherlands also conduct national-level surveys to assess patient satisfaction, which are used for various purposes related to improving health care services [20,21]. Previous studies revealed that quality improvement activities were initiated after patient care performance data were publicly released [22], with even hospitals that exhibited high levels of patient satisfaction delivering higher-quality clinical care [13].

In South Korea, the Korean Patient Experience Survey has been regularly conducted by HIRA biennially since 2017 by telephone among patients discharged from hospitals. It was developed based on the American HCAHPS, but further modified to suit the Korean context and to promote improvement of current Korean health care service quality [23]. A 2015 survey by the Korea Institute for Health and Social Affairs revealed that 30.5% of hospitalized patients claimed not to have been adequately informed about ward rounds and 20.4% reported not being included in the care process [24]. Furthermore, qualitative research during the Korean Patient Experience Survey development process revealed that inpatients were frequently unsatisfied with the communication with their doctors, especially regarding the lack of advance notice and unpredictability of ward round times [23]. To gauge these concerns, a direct question about whether the patient has been informed about the ward round time was included in the official Korean Patient Experience Survey.

Despite the increasing emphasis on patient-centeredness, patients are often still excluded from their own care process [5,6]. The results of the previous Korean Patient Experience Surveys (first to third, conducted from 2017 to 2021) reveal that the scores for the "Ensuring patients' rights" and "Services from physicians" domains remain lowest [9,10,25], suggesting that communication between patients and doctors remains unsatisfactory and patients do not feel sufficiently involved in the care process. Among several factors influencing patient participation and communication, such as the ward staff's attitude, time constraints [7], and health literacy [26], the care environment context itself has been identified as critical [8]. Previous Korean patient surveys revealed that patients felt dissatisfied and even frustrated by the medical staff's impolite, indifferent, or authoritarian attitudes and faced difficulties in communicating with their doctors who always appeared busy, and took issue with unpredictable and irregular round times [23]. Particularly in acute care environments, the patients' ability to participate is further diminished by their vulnerability, illness, and anxiety; furthermore, even the ward atmosphere is often disorderly and far from suitable for effective communication. In such situations, patient involvement is notably infrequent [6].

Patients in our AMU were mostly afflicted with complex, acute illness and frequently in need of interventional procedures, imaging studies, and even surgical procedures. Therefore, these patients and their carers often missed routine ward rounds during planned procedures or did not notice that ward rounds were taking place. Furthermore, the acuteness of the patients' conditions and lack of awareness about ward rounds likely led to frustration, anxiety, and dissatisfaction of patients and their carers toward their care. Real-time notifications provided through the WR-SMS addressed these issues by increasing awareness and predictability of ward rounds, which may help mitigate the emotional burden associated with the uncertainty of acute care. This aligns with studies suggesting that structured and predictable communication mechanisms, such as SMS notifications, create a more conducive environment for patient participation in care processes. A study conducted by Redley et al [6] at an acute inpatient ward indicated that only 18% (20/52) of rounds involved patient participation in clinical decisions, although even the patients who participated in rounds indicated low preferences for participation upon arrival at the hospital. The authors suggest that the patients' preferences for participation could change during the course of their hospital stay and it is also likely that passive participation may have been driven by limited opportunities to participate [6]. Similarly, Walton et al [27] revealed that familiarity with the hospital environment tended to positively affect participation in ward rounds, with several participants highlighting the uncertainty regarding the timing of rounds as a source of difficulty and anxiety.

With the development of communication devices such as cellular phones and smartphones, text messaging via SMS is increasingly used to enhance effective communication with patients and carers [14,17]. In our institution, the WR-SMS was developed and introduced to inform inpatients about the start of ward rounds in real time to improve their opportunities to participate in them. As physicians typically review their patients' medical records right before they begin ward rounds, the WR-SMS was embedded into our electronic health record system to enhance its use instead of using a separate platform. Following its implementation in January 2021, its utilization rate rapidly increased with long-term promotional activities, reaching an average of 65% in 2021. With its use, although only a statistically insignificant increase in patients' perceptions of opportunities to meet their doctors was noted (possibly owing

XSL•FO
RenderX

to the small sample size and the nature of the ward, which is staffed by hospitalists who are primarily based within the unit), inpatient satisfaction significantly improved with regard to "providing information about round-time" and the "Medication and treatment processes" domain. However, scores for other aspects in the "Services from physicians" domain did not improve significantly, suggesting that while real-time SMS notifications enhance patients' involvement and satisfaction with specific aspects of care, they may need to be supplemented with other interventions to address broader communication gaps.

Limitations

To our knowledge, this is the first study to investigate the effect of real-time round notification via SMS on inpatients' satisfaction. Nevertheless, it has some limitations. First, it was conducted at a single referral center in South Korea. As of 2019, the penetration rate of mobile phones in South Korea was 100%, with smartphones accounting for 95%. Therefore, our results may not be reproduced in other countries with lower mobile phone penetration rates. Second, as we used information from subjects who participated in the SNUBH patient-experience survey, instead of conducting a new survey, the number of patients included in this study is relatively small, precluding further adjustments and analyses based on clinical characteristics or comorbidities. However, the clinical setting remained constant throughout the study period and the subjects for the SNUBH patient-experience survey were randomly selected, which should minimize bias. Third, since this study included only patients capable of communication as survey respondents, it was not possible to assess the impact of WR-SMS on caregiver satisfaction.

Conclusion

Effective communication between physicians and patients, particularly during inpatient ward rounds, is a crucial aspect of inpatient care. This study proposes that real-time round notifications delivered through a user-friendly SMS service could improve inpatient satisfaction, likely by providing predictability and fostering an improved perception of emotional support. In the contemporary landscape, with personal communication devices becoming increasingly ubiquitous, the strategic utilization of such devices is anticipated to significantly influence patient satisfaction.

Acknowledgments

We would like to thank Editage for the English-language editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Questionnaire that was used for the Seoul National University Bundang Hospital patient-experience survey and the official Korean patient-experience survey developed by the Korean Ministry of Health and Welfare/Health Insurance Review and Assessment Service.
[DOCX File, 18 KB - humanfactors_v12i1e57470_app1.docx]

Multimedia Appendix 2
Changes in Likert scale survey scores on patient satisfaction before and after the implementation of ward round notifications via text messaging. Pre–WR-SMS group: patients hospitalized at the acute medical unit in 2020, before the implementation of ward round notifications via text messaging service; Post–WR-SMS group: patients hospitalized at the acute medical unit in 2021, after the implementation of ward round notifications via text messaging service.
[PNG File, 758 KB - humanfactors_v12i1e57470_app2.png]

References

1. Hulland O, Farnan J, Rabinowitz R, et al. What's the purpose of rounds? A qualitative study examining the perceptions of faculty and students. J Hosp Med 2017 Nov;12(11):892-897. [doi: 10.12788/jhm.2835] [Medline: 29091976]
2. Stickrath C, Noble M, Prochazka A, et al. Attending rounds in the current era: what is and is not happening. JAMA Intern Med 2013 Jun 24;173(12):1084-1089. [doi: 10.1001/jamainternmed.2013.6041] [Medline: 23649040]
3. Ward Rounds in Medicine: Principles for Best Practice: Royal College of Physicians; 2012.
4. Reddin G, Davis NF, Donald KM. Ward stories: lessons learned from patient perception of the ward round. Ir J Med Sci 2019 Nov;188(4):1119-1128. [doi: 10.1007/s11845-019-01975-z] [Medline: 30712243]
5. Aronsson L, Frithiof A, Röstedal A, Rudberg C, Ekström W. The ward round: patients' perceptions of a patient-centered approach and their suggestions for improved participation. Patient Prefer Adherence 2022;16:2203-2211. [doi: 10.2147/PPA.S343955] [Medline: 36003801]
6. Redley B, McTier L, Botti M, et al. Patient participation in inpatient ward rounds on acute inpatient medical wards: a descriptive study. BMJ Qual Saf 2019 Jan;28(1):15-23. [doi: 10.1136/bmjqs-2017-007292] [Medline: 29475980]

XSL•FO
RenderX

7. Moore L, Britten N, Lydahl D, Naldemirci Ö, Elam M, Wolf A. Barriers and facilitators to the implementation of person-centred care in different healthcare contexts. Scand J Caring Sci 2017 Dec;31(4):662-673. [doi: 10.1111/scs.12376] [Medline: 27859459]
8. McCormack B, Karlsson B, Dewing J, Lerdal A. Exploring person-centredness: a qualitative meta-synthesis of four studies. Scand J Caring Sci 2010 Sep;24(3):620-634. [doi: 10.1111/j.1471-6712.2010.00814.x] [Medline: 21050249]
9. Medical service patient experience survey results, 2017 (1st survey). : Health Insurance Review and Assessment Service; 2018.
10. Medical service patient experience survey results, 2019 (2nd survey). : Health Insurance Review and Assessment Service; 2020.
11. Goldstein E, Farquhar M, Crofton C, Darby C, Garfinkel S. Measuring hospital care from the patients' perspective: an overview of the CAHPS Hospital Survey development process. Health Serv Res 2005 Dec;40(6 Pt 2):1977-1995. [doi: 10.1111/j.1475-6773.2005.00477.x] [Medline: 16316434]
12. Fujisawa R, Klazinga NS. Measuring patient experiences (PREMS): progress made by the OECD and its member countries between 2006 and 2016. OECD. 2017. URL: https://www.oecd.org/content/dam/oecd/en/publications/reports/2017/12/measuring-patient-experiences-prems_b6354242/893a07d2-en.pdf [accessed 2025-02-26]
13. Jha AK, Orav EJ, Zheng J, Epstein AM. Patients' perception of hospital care in the United States. N Engl J Med 2008 Oct 30;359(18):1921-1931. [doi: 10.1056/NEJMsa0804116] [Medline: 18971493]
14. Wesevich A, Key-Solle M, Kandakatla A, Feeney C, Pollak KI, LeBlanc TW. Utilization of text messages to supplement rounding communication: a randomized feasibility study. J Gen Intern Med 2022 Sep;37(12):2991-2997. [doi: 10.1007/s11606-021-07285-4] [Medline: 35212877]
15. Ohn JH, Kim NH, Kim ES, et al. An acute medical unit in a Korean tertiary care hospital reduces the length of stay and waiting time in the emergency department. J Korean Med Sci 2017 Dec;32(12):1917-1920. [doi: 10.3346/jkms.2017.32.12.1917] [Medline: 29115071]
16. Kim HJ, Kim J, Ohn JH, Kim NH. Impact of hospitalist care model on patient outcomes in acute medical unit: a retrospective cohort study. BMJ Open 2023 Aug 3;13(8):e069561. [doi: 10.1136/bmjopen-2022-069561] [Medline: 37536969]
17. Souza-Silva MVR, Domingues MLDP, Chagas VS, et al. Implementation of a text messaging intervention to patients on warfarin therapy in Brazilian primary care units: a quasi-experimental study. BMC Prim Care 2022 Mar 23;23(1):54. [doi: 10.1186/s12875-022-01647-5] [Medline: 35321654]
18. Siouta E, Olsson U. Patient centeredness from a perspective of history of the present: a genealogical analysis. Glob Qual Nurs Res 2020;7:2333393620950241. [doi: 10.1177/2333393620950241] [Medline: 32944591]
19. Committee on Quality of Health Care in America. Crossing the Quality Chasm: A New Health System for the 21st Century: National Academies Press (US); 2001. [doi: 10.17226/10027] [Medline: 25057539]
20. NHS patient surveys. : NHS England & the Department of Health and Social Care; 2024. URL: https://nhssurveys.org [accessed 2025-02-26]
21. Delnoij DMJ, Rademakers J, Groenewegen PP. The Dutch consumer quality index: an example of stakeholder involvement in indicator development. BMC Health Serv Res 2010 Apr 6;10:88. [doi: 10.1186/1472-6963-10-88] [Medline: 20370925]
22. Fung CH, Lim YW, Mattke S, Damberg C, Shekelle PG. Systematic review: the evidence that publishing patient care performance data improves quality of care. Ann Intern Med 2008 Jan 15;148(2):111-123. [doi: 10.7326/0003-4819-148-2-200801150-00006] [Medline: 18195336]
23. Do YK, Kim JE, Lee JY, H-y L, Cho MW. Developmental research on models to evaluate patient-centeredness. : Health Insurance Review and Assessment Service; 2015.
24. Kang HJ, Lee SK, Kim SE, Hong JS, Lee KS. Korea healthcare quality report: achievements and challenges of Korea health care system. : Korea Institute for Health and Social Affairs; 2015.
25. Medical Service Patient Experience Survey results, 2021 (3rd survey). : Health Insurance Review and Assessment Service; 2022.
26. Edwards M, Davies M, Edwards A. What are the external influences on information exchange and shared decision-making in healthcare consultations: a meta-synthesis of the literature. Patient Educ Couns 2009 Apr;75(1):37-52. [doi: 10.1016/j.pec.2008.09.025] [Medline: 19036550]
27. Walton V, Hogden A, Long JC, Johnson JK, Greenfield D. Patients, health professionals, and the health system: influencers on patients' participation in ward rounds. Patient Prefer Adherence 2019;13:1415-1429. [doi: 10.2147/PPA.S211073] [Medline: 31692552]

Abbreviations

AMU: acute medical unit
HCAHPS: Hospital Consumer Assessment of Healthcare Providers and Systems
HIRA: Health Insurance Review and Assessment Service
MOHW: Ministry of Health and Welfare
SNUBH: Seoul National University Bundang Hospital

XSL•FO
RenderX

WR-SMS: ward round notifications via text messaging

Edited by A Kushniruk; submitted 17.02.24; peer-reviewed by A Hassan, H Poulsen, JH Lee, R Bidkar; revised version received 10.01.25; accepted 14.01.25; published 12.03.25.

Please cite as:
Lee J, Ahn S, Ohn JH, Kim ES, Lim Y, Kim HW, Park HS, Cho JH, Kim SW, Ryu J, Kim J, Jang HC, Kim NH
Effect of SMS Ward Round Notifications on Inpatient Experience in Acute Medical Settings: Retrospective Cohort Study
JMIR Hum Factors 2025;12:e57470
URL: https://humanfactors.jmir.org/2025/1/e57470
doi:10.2196/57470

© Jongchan Lee, Soyeon Ahn, Jung Hun Ohn, Eun Sun Kim, Yejee Lim, Hye Won Kim, Hee-Sun Park, Jae Ho Cho, Sun-wook Kim, Jiwon Ryu, Jihye Kim, Hak Chul Jang, Nak-Hyun Kim. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 12.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Empowering Community Health Workers With Scripted Medicine: Design Science Research Study

Dario Staehelin[1,2], PhD; Damaris Schmid[2], MSc; Felix Gerber[3,4,5], MD; Mateusz Dolata[2,6], Prof Dr; Gerhard Schwabe[2], Prof Dr

[1]Department for Information and Process Management, Eastern Switzerland University of Applied Sciences, St. Gallen, Switzerland

[2]Department of Informatics, University of Zurich, Zurich, Switzerland

[3]Division of Clinical Epidemiology, Department of Clinical Research, University Hospital Basel, Basel, Switzerland

[4]Swiss Tropical and Public Health Institute, University of Basel, Basel, Switzerland

[5]Swiss Tropical and Public Health Institute, Allschwil, Switzerland

[6]Department of Political and Social Sciences, Zeppelin University, Friedrichshafen, Germany

Corresponding Author:
Dario Staehelin, PhD
Department for Information and Process Management
Eastern Switzerland University of Applied Sciences
Rosenbergstrasse 59
St. Gallen, 9001
Switzerland
Phone: 41 582573106
Email: staehelin@ifi.uzh.ch

Abstract

Background: The World Health Organization anticipates a shortage of 14 million health workers by 2030, particularly affecting the Global South. Community health workers (CHWs) may mitigate the shortages of professional health care workers. Recent studies have explored the feasibility and effectiveness of shifting noncommunicable disease (NCD) services to CHWs. Challenges, such as high attrition rates and variable performance, persist due to inadequate organizational support and could hamper such efforts. Research on employee empowerment highlights how organizational structures affect employees' perception of empowerment and retention.

Objective: This study aims to develop Scripted Medicine to empower CHWs to accept broader responsibilities in NCD care. It aims to convey relevant medical and counseling knowledge through medical algorithms and ThinkLets (ie, social scripts). Collaboration engineering research offers insights that could help address the structural issues in community-based health care and facilitate task shifting.

Methods: This study followed a design science research approach to implement a mobile health–supported, community-based intervention in 2 districts of Lesotho. We first developed the medical algorithms and ThinkLets based on insights from collaboration engineering and algorithmic management literature. We then evaluated the designed approach in a field study in the ComBaCaL (Community Based Chronic Disease Care Lesotho) project. The field study included 10 newly recruited CHWs and spanned over 2 weeks of training and 12 weeks of field experience. Following an abductive approach, we analyzed surveys, interviews, and observations to study how Scripted Medicine empowers CHWs to accept broader responsibilities in NCD care.

Results: Scripted Medicine successfully conveyed the required medical and counseling knowledge through medical algorithms and ThinkLets. We found that medical algorithms predominantly influenced CHWs' perception of structural empowerment, while ThinkLets affected their psychological empowerment. The different perceptions between the groups of CHWs from the 2 districts highlighted the importance of considering the cultural and economic context.

Conclusions: We propose Scripted Medicine as a novel approach to CHW empowerment inspired by collaboration engineering and algorithmic management. Scripted Medicine broadens the perspective on mobile health–supported, community-based health care. It emphasizes the need to *script* not only essential medical knowledge but also *script* counseling expertise. These scripts allow CHWs to embed medical knowledge into the social interactions in community-based health care. Scripted Medicine empowers CHW to accept broader responsibilities to address the imminent shortage of medical professionals in the Global South.

XSL•FO
RenderX

(JMIR Hum Factors 2025;12:e57545) doi:10.2196/57545

KEYWORDS

community-based health care; community health workers; collaboration engineering; algorithmic management; Scripted Medicine; task shifting; empowerment; mobile health; mHealth; digital health

Introduction

Background

The World Health Organization projects a health worker shortage of 14 million by 2030 [1]. Many countries struggle to train and retain enough medical professionals [1], with the Global South disproportionately affected by this shortage [2]. In response to this situation, community health workers (CHWs) have been integral in alleviating the consequences of staff shortages and extending the health systems in many countries [3]. CHWs are laypeople residing within their communities who receive basic medical training to provide essential health services in their communities and to link patients to professional facility-based health care [3]. Whereas physicians are often rare, expensive, and concentrated in urban hubs, CHWs are available in vast numbers and comparably inexpensive. While CHWs have traditionally focused on maternal, neonatal, and infectious disease services, recent studies have investigated the feasibility of expanding CHWs' responsibilities to tackle the rapidly growing noncommunicable disease (NCD) pandemic [4,5]. In task shifting, specific tasks are shifted from medical professionals to CHWs [6,7]. For example, many community-based health programs in the Global South shift screening and monitoring tasks for infectious diseases, such as HIV and tuberculosis, from medical professionals to CHWs. Task shifting may be a cost-effective approach to increasing the reach of health care delivery systems. It may reduce important access barriers by allowing for more decentralized health care delivery while freeing medical professionals' resources [8].

Despite the increasing popularity of community-based health programs, attrition rates among CHWs are generally high, and studies often report high variability in CHW performance [9-13]. These persisting challenges are linked to inadequate organizational support [12-14], including insufficient training and economic opportunities [15,16] and a lack of appreciation and support from supervisors [9,14]. Further, CHWs reported other social challenges, such as a lack of support from their families and insufficient trust, linked to low uptake rates of services by the community [14].

Empowering CHWs to Accept Broader Responsibilities

The empowerment literature reported high attrition rates and performance variability due to a lack of organizational support, especially in nurses [17-19]. Studies show that an organization needs to transfer power from supervisors to employees to create ideal work conditions. This transfer of power is called structural empowerment [20,21]. It emphasizes the necessity of providing employees with a work environment that enables them to carry out their work. Structural empowerment is achieved by providing employees *access to information*, *resources*, *support*, and *opportunities* [21,22]. The act of empowerment positively influences the perception of empowerment. Psychological empowerment builds on the idea that perceived control increases intrinsic motivation. It draws on prior research on self-efficacy [23] and self-determination [24] and is measured through the perception of *meaning*, *competence*, *autonomy*, and *impact* [25]. Empowering employees increases retention rates and job satisfaction and decreases job strain [18,22,26]. Employees feel disempowered when lacking access to critical information (eg, through training), resources (eg, adequate pay), or support from peers and supervisors. Disempowered employees are less productive and more likely to change jobs.

Unsurprisingly, CHW empowerment is a growing topic of interest. However, CHW empowerment is a poorly defined concept despite its recent popularity. Researchers often use empowerment as a synonym for facilitating or enabling CHWs to carry out specific tasks without diving into the depth of its meaning. For example, CHWs are empowered to treat diabetes [27] or provide counseling for family planning [28]. Consequently, they often miss addressing the persisting challenges of inadequate organizational support that lead to high attrition rates and performance variability. Only a few articles introduce CHW empowerment conceptualizations that consider the sociotechnical aspects of community-based health care. They highlight how the lack of organizational support disempowers CHWs with social, economic, and cultural consequences [29-32].

Mobile Health for CHWs

Mobile technologies often play a central role in efforts to empower CHWs in their work [33-35]. Recent studies demonstrate the potential of mobile health (mHealth) to address the challenges in CHW retention and performance [36], for example, through decision support [11,37-39]. Clinical decision support systems provide CHWs with the necessary clinical knowledge through medical algorithms to ensure the proper implementation of guidelines [37,40-42]. These explorations in empowering CHWs with decision-support tools align well with the premise of algorithmic management. Algorithmic management is a growing research topic in human-computer interaction and information systems research. It argues that management becomes a process embedded in technology rather than a human routine [43,44]. Studies have shown how algorithmic management can increase productivity without increasing job strain by managing the user's work process [45,46].

The 5RSM framework describes 7 forms of algorithmic management for process restriction: *algorithmicrecommending*, *restricting*, *requiring*, *rating*, *rewarding*, *sanctioning*, and *monitoring* [47]. Table 1 provides the original definition for each form of algorithmic management from Hirsch et al [47]. We complement the table with CHW-related examples.

XSL•FO
RenderX

Table 1. Forms and definitions of algorithmic management [47], with community health toolkit (CHT)–related examples.

Form	Definition (how the algorithm aligns worker behavior with organizational goals)	Example (CHT application)
Recommending	The algorithm automatically suggests specific behavior to the worker that takes the form of either explicit recommendations for action or implicit forms (such as nudges).	Medical algorithms propose a task prioritization that CHWs[a] can adopt to organize their workflow.
Restricting	The algorithm withholds information or limits the scope of action for the worker.	Medical algorithms hide recommendations for a drug prescription in case of reported intolerance.
Requiring	The algorithm instructs workers to perform specific actions that they cannot reject without negative consequences.	Medical algorithms require CHWs to perform tasks by a specific due date.
Rating	The algorithm ranks workers or allows third parties to rate work behavior and outcome.	Medical algorithms display leader boards comparing the number of clients CHWs have counseled.
Rewarding	The algorithm assigns rewards for work behaviors that align with organizational goals.	Medical algorithms issue nonmonetary (eg, digital badges) and monetary rewards (eg, performance-based compensation).
Sanctioning	The algorithm issues penalties for work behaviors that do not align with the organizational goals.	Medical algorithms reduce performance-based compensation if tasks are not completed in due time.
Monitoring	The algorithm enables supervision of the work process by the worker or a third party by collecting, aggregating, and displaying data on working behavior.	Medical algorithms provide performance dashboards with performance indicators such as the number of delayed task completions per month.

[a]CHW: community health workers.

While algorithmic management offers valuable tools for empowering CHW with mHealth, we believe 2 dimensions require specific consideration for community-based health care. *Recommending* might only be applied in areas where patient safety is not at stake. For example, algorithms must *require* instead of *recommending* CHWs to perform specific tasks such as distributing medication on time to ensure patient safety. Work termination, the strongest form of *sanctioning*, should not be performed by algorithms. Work termination has far-reaching consequences as communities would temporarily be left without their CHW.

Digital CHW empowerment is associated with the improvement of patient knowledge and attitude [48], disease surveillance and health coverage [33], and identification of health priorities in the communities [49]. These efforts often aim to address the root causes of CHW disempowerment [35]: the lack of organizational support and appreciation (eg, by supervisors) [9,14]; inadequate training and learning opportunities [15]; and insufficient monetary compensation [16].

However, the growing CHW empowerment literature mentioned earlier criticizes the unilateral focus on clinical knowledge to capacitate CHWs [29,30,50]. Research often abstracts from the context in which the clinical knowledge is applied [51]. It studies how technology *empowers* CHWs to become better and more reliable at diagnosing and monitoring diseases in the field. Consequently, the importance of a functioning CHW-client interaction is often omitted [52,53]. A recent study highlights the persisting, often implicit, view of CHWs as sensors for the health system because of this unilateral consideration [54]. CHWs are often deployed to *sense* the health status of their community by diagnosing and monitoring specific diseases (eg, HIV). However, they are not autonomous actors who offer treatment services. This perception stands against the proclaimed active role CHWs should claim to address predicted staff shortages [4,50]. Consequently, research lacks a comprehensive approach that embeds clinical knowledge in its application context to empower CHWs to accept broader responsibilities.

Collaboration Engineering

Current research often abstracts from the social interactions—the collaboration between CHWs and patients—in community-based health care. However, we argue that community-based health care is not the outcome of an individual's effort (ie, CHW) or technology. It is a collaborative process between CHWs, their clients, and the health system. Seen as a collaborative process, community-based health care would likely benefit from insights from collaboration engineering research. Collaboration engineering originates from research on group support systems (GSSs) [55]. This research studies how technology enhances group collaboration by reducing process losses, such as conformity pressure or the conversation domination of 1 or a few collaborators [56]. Despite the substantial benefits, GSS adoption struggled as they require professional facilitators who design and conduct GSS-supported group collaborations [56,57]. However, facilitators are a rare and expensive resource, requiring extensive training and experience [58].

The premise of collaboration engineering is to drastically reduce costs and training efforts by enabling virtually anyone to design and conduct group collaboration. Briggs et al [59] described this as the "facilitator in a box." The core assumption is that high value and recurring tasks can be predesigned to empower nonexperts through process restrictiveness [57,60]. This restriction quickly allows nonexperts to carry out expert tasks without extensive training and experience. However, collaboration engineering stresses that they must benefit users and offer adequate flexibility to avoid adverse effects [59,60]. The literature describes enabling nonexperts to fruitfully collaborate by restricting them in the process through (1) a carefully designed series of activities, (2) the configuration of technology to support behaviors within each activity, and (3)

XSL•FO
RenderX

guidance on the behaviors (including constraints) they should perform in each activity [59]. Collaboration engineering research proposes ThinkLets to guide the behaviors in these activities. ThinkLets are tools to "codify expertise such that it would be easy for [nonexperts] to learn and reuse." Similar to a cooking recipe, they consist of explicit instructions on collaboration behaviors, including instructions on configuring and using support tools [57]. ThinkLets formalize the required expertise for nonexperts to reach predictable and repeatable outcomes, increasing collaboration performance and outcomes [61].

On the basis of these insights, we argue that collaboration engineering is a promising approach for CHW empowerment as its underlying premise closely resembles global struggles in health care. The fundamental problem that collaboration engineering aims to resolve is that of scarce, valuable resources. Facilitators are replaced with inexpensive and plentiful nonexperts, who receive tools to facilitate collaboration in a specific context successfully. Collaboration engineering breaks processes into repeatable and formalized pieces, preconfigured technology, and guidance on required behaviors. From this perspective, collaboration engineering and community-based health care adopt a similar approach. Medical professionals facilitate their medical consultations based on their formal training and a repertoire of scripts grown by experience. This training is time-consuming and expensive. Likewise, CHWs are the nonexpert collaborators in medical consultations. They are available in large numbers and at comparably low costs. They lack critical knowledge and expertise that must be compensated through process restrictions. Thus, medical professionals are trained facilitators, and CHWs are nonexperts facilitating health care within a limited domain.

Following this analogy, we claim collaboration engineering offers efficient and effective task shifting tools that "standardize the performance and interpretation of certain tasks" [6]. Consequently, CHWs could be empowered to accept broader responsibilities and perform tasks beyond diagnosing and monitoring patients. In turn, empowered CHWs might be retained longer, reducing attrition rates. Despite this potential value, there is a lack of research studying how collaboration engineering can formalize crucial knowledge to facilitate task shifting from medical professionals to CHWs.

Research Objective

This study addresses this research gap by developing Scripted Medicine, an approach for CHW empowerment inspired by research on collaboration engineering and algorithmic management. The insights from algorithmic management align well with restricting collaboration processes through (1) a carefully designed series of activities and (2) the configuration of technology to support behaviors within each activity. In community-based health care, mHealth combines both principles as medical algorithms represent the series of activities and provide the user with the functionalities required to perform these activities. Collaboration engineering offers ThinkLets as a tool to guide CHWs' behaviors throughout the activities. They convey expert knowledge in predictable and repeatable recipe-like scripts that make the knowledge easily accessible for nonexperts. Accordingly, we formulate the design goal and solution objectives demonstrated in Textbox 1.

Textbox 1. Design goal and solution objectives.

Design goal

- Empower community health workers with Scripted Medicine to accept broader responsibilities in noncommunicable disease care

Solution objectives

- "Medical Knowledge"
  - Convey medical knowledge through a carefully designed series of activities and preconfigured technology by drawing on insights from algorithmic management
- "Counseling Knowledge"
  - Convey knowledge through guidance on specific behaviors in the activities in ThinkLets

Methods

Overview

We followed the design science research methodology proposed by Peffers et al [62] to develop the Scripted Medicine approach for CHW empowerment. Design science research is especially suitable as it aims to create artifacts that address a practical problem [63]. Such artifacts include software designs, theoretical models, or approaches [62]. While originating from information systems and computer science, design science research is increasingly applied in medical informatics to study emerging phenomena and generate new knowledge through a design approach [64-68]. The following subsections introduce the ComBaCaL (Community Based Chronic Disease Care Lesotho) project and describe our implementation of the design science research methodology. This manuscript focuses on the diabetes pilot of the ComBaCaL project.

Ethical Considerations

The pilot is approved by the Ministry of Health Research and Ethics Committee Lesotho (ID 176-2021). All participants gave verbal informed consent to participate in the study. The obtained data was de-identified. The participants did not receive any compensation for participating in the study.

XSL•FO
RenderX

ComBaCaL: Community-Based NCD Care

ComBaCaL is an innovative initiative to address the escalating NCD epidemic in Lesotho. The 5-year project used a community-based model, leveraging CHWs and mHealth to enhance NCD prevention, screening, diagnosis, and care.

One key aspect of ComBaCaL is its emphasis on empowering CHWs to play a pivotal role in delivering NCD care at the community level. Lesotho successfully decreased HIV transmission and AIDS-related deaths mainly due to the broad rollout of HIV testing and antiretroviral treatment. Previous studies explored the potential of decentralized, home-based service delivery for HIV care in the setting [69-71]. The project aimed to continue expanding the responsibilities of CHWs horizontally (ie, other diseases) and vertically (ie, offer treatment).

ComBaCaL is among the first projects to expand CHWs' responsibilities to provide treatment for NCDs through task shifting. The project strategically focuses on 2 critical NCDs: type 2 diabetes and hypertension, with a comprehensive care package beyond screening and monitoring. ComBaCaL follows a more comprehensive approach to community-based care that includes lifestyle counseling and prescription of first-line treatments for uncomplicated type 2 diabetes and hypertension. The goal was to provide accessible NCD care within the community, thus minimizing the need for travel to the often distant health care facilities without compromising patient safety and care quality.

To achieve this goal, ComBaCaL used mHealth, specifically a community health toolkit (CHT), to capacitate CHWs. These tools guide CHWs through clinical algorithms on tablets, enabling them to effectively screen and monitor NCDs, prescribe first-line treatments, and provide necessary care within the community.

The CHT is an open-source toolkit for community-based health care that reflects the needs of the local context [72]: a community-based perspective, an offline-first approach, and flexibility in configuring medical algorithms. These tools should streamline CHW activities and facilitate decision support based on entered data, ensuring effective and accurate care provision.

ComBaCaL is a multidisciplinary implementation research project that aims to improve care for chronic diseases in rural Lesotho by developing and implementing innovative community-based health service delivery models. The project consortium includes SolidarMed, a local nonprofit organization; the Division of Clinical Epidemiology at the University Hospital Basel; the National University of Lesotho; the Lesotho Ministry of Health; and the Department of Informatics at the University of Zurich.

Design and Development

An interdisciplinary and multicultural project team adopted a human-centered approach to gather extensive requirements for the envisioned community-based health program with support from students in Lesotho and Switzerland [73,74]. This approach allowed the project team to include all relevant stakeholders and their needs. The program was developed in 2 streams between February 2021 and April 2022. In the first stream, the project team developed medical algorithms instantiated in the CHT, named the ComBaCaL app. In collaboration with medical experts, informatics students from a university in Switzerland build medical algorithms that contain step-by-step guidance in registering, screening, diagnosing, and (if necessary) treating clients for diabetes. The algorithms were developed using agile methods and tested before the intervention. In the second stream, a group of 4 researchers iteratively codeveloped ThinkLets. The researchers developed 9 ThinkLets covering the social interactions associated with the medical algorithm's guidance. They especially considered the interrelation between medical procedures and social interactions between CHWs and their patients. The ThinkLets were revised and finalized after the CHW training to incorporate learning from practice sessions.

Demonstration and Evaluation

The concept was evaluated with 10 newly recruited CHWs in 2 districts in Lesotho in Spring 2022. The communities elected the CHWs according to the Lesotho Village Health Policy. The scope of the evaluation covers 2 weeks of training and approximately 12 weeks of field experience, during which CHWs provided diabetes counseling in their community.

The first week of training focused on clinical knowledge. Medical experts educated CHWs on diabetes risk factors, pathophysiology, diagnosis, and treatment. Practical sessions on medical procedures, such as capillary blood glucose measurement, complemented the lectures and allowed CHWs to practice with peers. The second training week focused on embedding the gained clinical knowledge into the interactions between CHWs and their clients. The practice-oriented counseling training was organized in modules along the counseling process. Each module was built on the previous one to cover the whole interaction when combined. The modules all followed the same approach. First, the trainers introduced the relevant ThinkLet to the plenum. They explained the aim, reason, and process of each ThinkLet and answered questions. A trainer demonstrated how a ThinkLet intends to support the CHWs in handling the medical algorithm on a tablet app and their interaction with clients. A total of 2 CHWs then applied the ThinkLet in a roleplay, with the remaining CHWs watching and asking questions. Second, the CHWs practiced on each other in groups of 5 by applying the clinical knowledge in the algorithms in its social context supported by ThinkLets. Third, the groups met again in the plenum to discuss what they had learned and to answer open questions. This approach was repeated until the CHWs counseled their peers, from welcoming to saying goodbye. The training concluded with 2 days in a nearby village where the CHWs applied their knowledge and skills in a real-world setting with people closely resembling their future clients.

Data Collection and Analysis

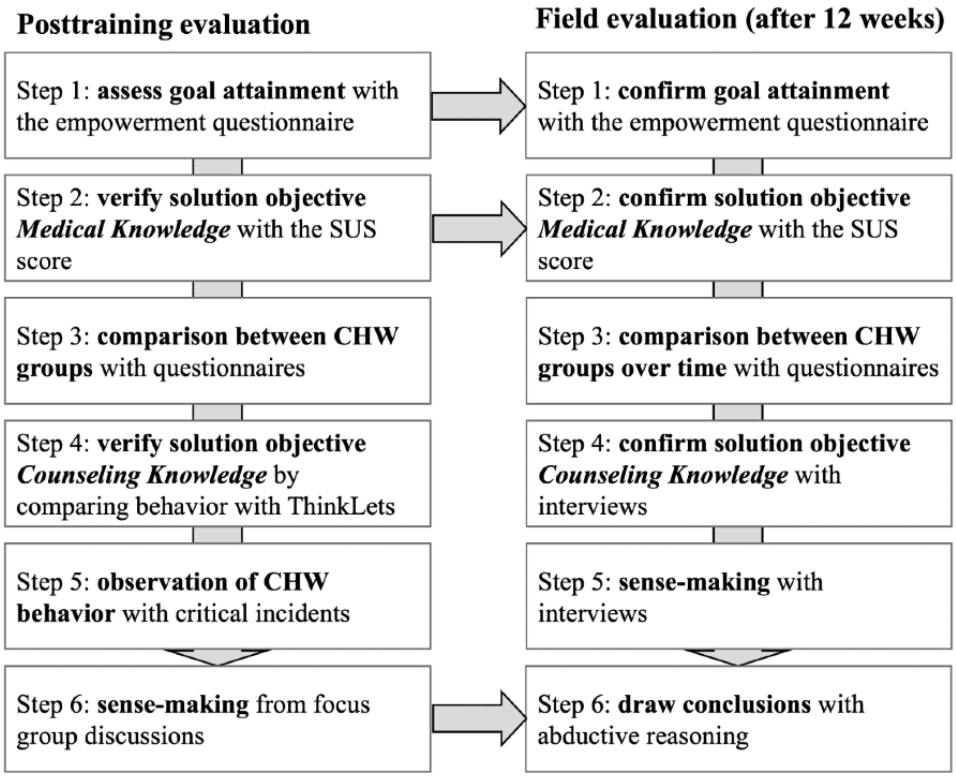
Overview

We followed an abductive approach to study how Scripted Medicine empowers CHWs to accept broader responsibilities in NCD care. Deductive and inductive approaches aim to test a theory objectively (ie, deductive) or generate a theory from

XSL•FO
RenderX

interpretation (ie, inductive). Compared to these approaches, abductive research balances empirical data and existing theory [75]. This balance makes abductive reasoning suitable for design science research and this study as it considers a theory (ie, empowerment) and empirical data to generate new knowledge [76]. We conducted our abductive approach in 2 rounds, as depicted in Figure 1. Throughout the data collection, the CHWs always could speak Sesotho to express themselves in the best way possible. Native speaking Basotho either participated in the focus group discussion or conducted the interviews if preferred. Questionnaires were in Sesotho and English languages.

Figure 1. An abductive analysis approach for the posttraining and field evaluation. CHW: community health worker; SUS: System Usability Scale.



Posttraining Evaluation

We collected data during the CHWs' training. We captured video recordings during the training sessions to study the CHWs' learning curve. At the end of the training, the CHWs participated in a focus group discussion to reflect on their training experience and expectations toward their new role. We divided the CHWs into 2 groups of 5 to give each CHW enough space to speak. Finally, each CHW completed a questionnaire on structural and psychological empowerment [17,25] and rated the ComBaCaL app on the System Usability Scale (SUS) [77] (see Multimedia Appendix 1 [17] and Multimedia Appendix 2 [77] for the questionnaires).

We conducted the data analysis in 6 steps. First, we began our analysis by focusing on the questionnaire data to assess goal attainment regarding the CHWs' perception of structural and psychological empowerment. We calculated the median and mean for each item and item group (where applicable). Second, we calculated the SUS score for the ComBaCaL app. The SUS evaluates the quality of the process restrictions instantiated in the ComBaCaL app through (1) the designed sequence of activities and (2) the configuration of the app, and, therefore, verifies solution objective 1. Third, we compared the results between the Mokhotlong and Butha Buthe groups (5 CHWs each). This step elicited interesting insights to guide the subsequent observation and interview data analysis. We then analyzed the observations (ie, video recordings). In the fourth step, we compared the CHWs' behaviors with the counseling ThinkLets to assess their adoption (solution objective 2). Fifth, we applied the critical incidents technique to identify outstandingly effective or ineffective behaviors [78]. Finally, we analyzed the 2 focus group discussions following a theory-driven coding approach based on the empowerment literature [20,25] paired with in-vivo coding for emerging themes. Analyzing the focus group discussions allowed us to understand the insights from the previous steps. The initial codes were *access to information*, *support*, *resources*, and *opportunities* (structural empowerment) and *meaning*, *impact*, *competence*, and *autonomy* (psychological empowerment).

Field Evaluation

We collected a second round of data after the CHWs worked in the field for approximately 12 weeks. The CHWs again completed the same questionnaire on empowerment and system usability. Further, they provided insights into their work through semistructured interviews (individual interviews).

We first analyzed the empowerment questionnaire and SUS score (steps 1 and 2). In step 3, we calculated changes between the initial and second questionnaires in addition to the median and mean. These changes yielded interesting insights into the dynamics of empowerment that we further explored in the interviews (step 5). We analyzed the interviews based on the

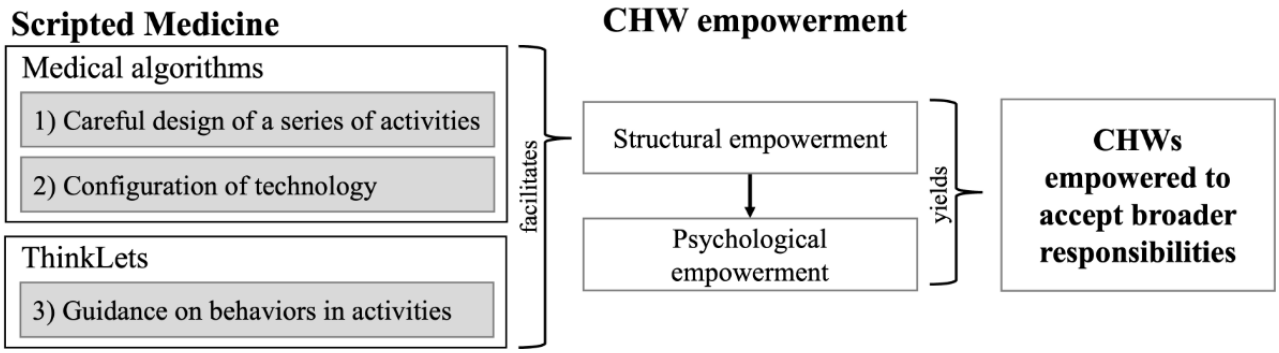
XSL•FO
RenderX

same coding approach to find explanations for the identified changes. We further analyzed the interviews regarding the continued use or adaptation of the ThinkLets (step 4). Finally, we combined survey data, interviews, and field observations to draw our conclusions. We could generate a rich and multidimensional understanding of Scripted Medicine through careful triangulation. Our abductive approach sensitized us to balance the empirical findings and theoretical grounding in collaboration engineering, algorithmic management, and empowerment. Finally, we formulated design principles as proposed in a study by Gregor et al [79] to synthesize our contribution and inform designers.

Artifact Description

Scripted Medicine was developed based on the 3 principles of process restriction proposed in collaboration engineering: (1) careful design of a series of activities; (2) configuration of technology to support behaviors within each activity; and (3) guidance on the behaviors (including constraints) performed in each activity [59]. Insights from algorithmic management guided the design of the series of activities and the configuration of technology, while ThinkLets formalize the behaviors in the activities. Scripted Medicine interlinks medical and counseling expertise conveyed by algorithms and ThinkLets to structurally and psychologically empower CHWs. The act and perception of empowerment allow CHWs to accept broader responsibilities in NCD care. Figure 2 depicts this conceptual model for Scripted Medicine.

Figure 2. A conceptual model for community health worker (CHW) empowerment with Scripted Medicine based on process restrictions through medical algorithms and ThinkLets.



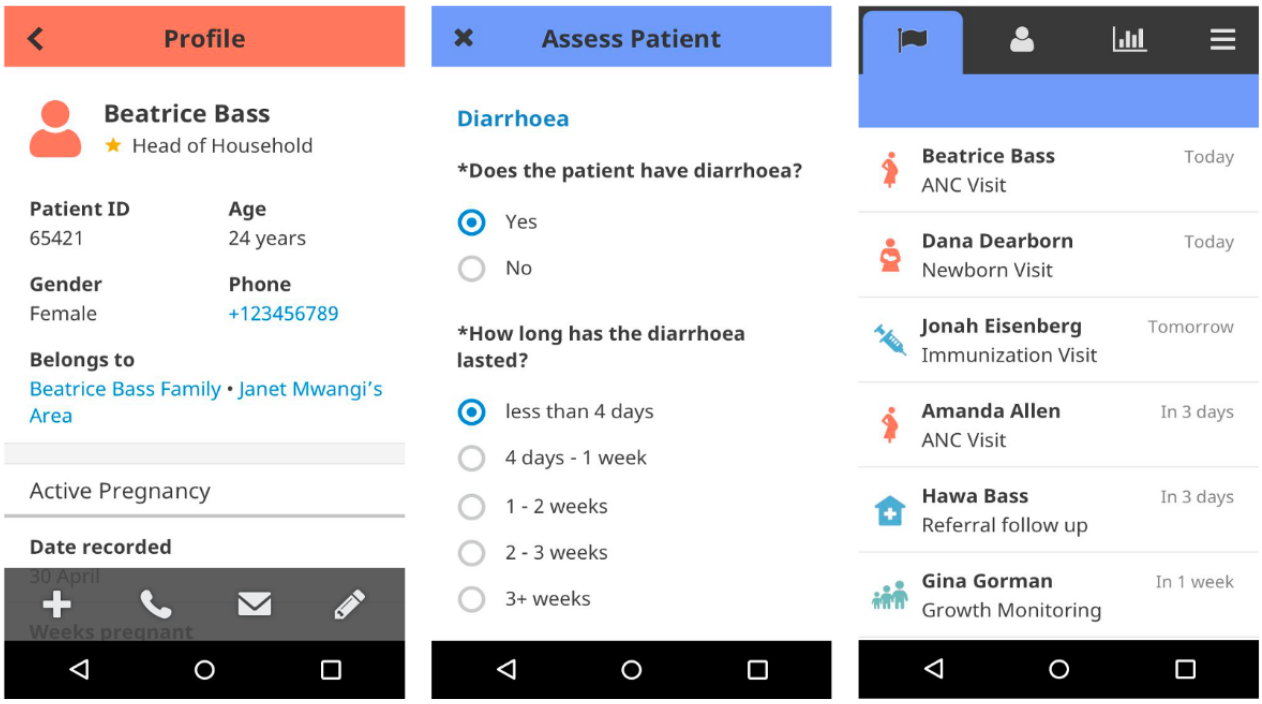
Formalizing Clinical Knowledge in ComBaCaL

Medical algorithms formalize the clinical knowledge necessary to offer diagnosis, monitoring, and treatment services for NCDs. These algorithms represent a carefully designed series of activities that follow existing medical guidelines. They constrain CHWs in their actions to ensure the proper application of care procedures for patient safety. These constraints facilitate the expansion of CHWs' responsibilities and have been shown to empower CHWs [35]. For example, they offer step-by-step guidance to support CHWs in implementing the proper medical procedures in the correct order. Further, mHealth automates documentation, which frees up time for CHWs. Finally, such tools give CHWs access to relevant medical history information collected previously.

The medical algorithms were implemented in CHT and called the ComBaCaL app. The toolkit offers a "shell" with basic functionalities and allows for rapid deployment of medical algorithms. The ComBaCaL app is structured in 3 tabs: tasks, care guides, and profiles (Figure 3 [72]). The medical algorithms influence each tab by defining open tasks and their due date, storing, and displaying information in patient profiles, and providing step-by-step guidance in the care guides. For the study, the medical algorithm covered medical procedures for (1) *patient and household enumeration* (ie, registration and obtaining informed consent); (2) *screening for diabetes*, using capillary blood sugar measurements after prescreening using age and body mass; (3) *diagnosing diabetes*; and (4) *initiating treatment* (ie, medication and lifestyle advice, if applicable). In a diagnosis, the medical algorithm would initiate screening for signs and symptoms of advanced disease and other criteria that would require a referral of the patient to a health facility. In an uncomplicated disease, the algorithm would start a treatment cycle with regular follow-ups to adjust medication until an acceptable blood glucose level is reached. If acceptable blood glucose levels cannot be reached through the treatment provided by the CHW, the algorithm would then suggest referral to the health facility for further care. Guidelines for medical professionals are confined to essential clinical information. More guidance is required in community-based health care as CHWs do not have this formal education. In ComBaCaL, the medical algorithms contained relevant clinical questions, notes on preparations (eg, to ensure that all necessary materials, such as protective gloves are available), information to provide to clients (eg, possible side effects of medication), and alerts, if a client must be sent to a medical professional. However, not all relevant knowledge can be conveyed in algorithms, as they would become cluttered and overly lengthy.

XSL•FO
RenderX

Figure 3. The community health toolkit structure in the tabs: (A) profile; (B) care guides; and (C) tasks. ANC: antenatal care.



Formalizing Counseling Knowledge in ComBaCaL

Accordingly, Scripted Medicine formalized counseling knowledge in ThinkLets. Medical scripts (ie, medical algorithms) guided the CHWs through a series of activities. This way, they offered the information basis for the social interactions (eg, a diagnosis a CHW must explain to the patient). ThinkLets formalized counseling expertise to guide CHWs in these social interactions throughout the consultation process. While initially developed to guide behaviors in large group collaboration, ThinkLets are not limited to such collaborations. They have been successfully applied in dyadic collaboration. For example, ThinkLets can empower service providers to use technology to adapt work practices for value cocreation in service encounters [80]. They can also facilitate knowledge transfer between expert advisers and nonexpert clients in burglary prevention [81].

In Scripted Medicine, they contained the counseling knowledge required to interact meaningfully with clients and build quality relationships. Counseling ThinkLets embed the clinical knowledge instantiated in algorithms into social interactions. Such ThinkLets can range from high-level steps to microinteractions. For example, a ThinkLet could provide instructions on approaching new clients and introducing the health program to start counseling. Another ThinkLet on comforting clients could offer CHWs a tool to handle patients' emotional reactions after a diagnosis.

In ComBaCaL, social interactions along the medical algorithm's guidance were scripted in 9 ThinkLets. The ThinkLets covered the steps of *greeting and consent*, *enumeration*, *screening*, *confirmation*, *counseling*, *referral*, *medication*, *dosing*, *closing with diabetes*, and *closing no diabetes*. These high-level steps are intended as a proof-of-concept that should be broken down into more nuanced ThinkLets with growing experience and insights. Table 2 shows the *greeting and consent* ThinkLet as an example. Italicized texts indicate posttraining adaptations made due to insights from the training. For example, Basotho customs require people entering the property to announce themselves with "Koko." The ThinkLets indicates the inputs required and expected outputs for the interaction. The setup includes instructions, for example, on how to position oneself or organize the working space. The steps and "say this..." offer guidance on how to structure the interaction. In the subsequent example, the ThinkLet also refers to additional materials to be used (ie, "according to the introduction cheat sheet"). Finally, notes offer insights for CHWs and others into the context and mechanisms of a ThinkLet.

XSL•FO
RenderX

Table 2. An example ThinkLet greeting and consent, with posttraining adaptations in italics.

Dimension	Description
Name	• Greeting and consent
Inputs	• None
Outputs	• Signed informed consent form and client who knows the CHW[a] and ComBaCaL[b]
Setup	• *Say loudly "koko" while entering the property and repeat until you can be certain that the owner heard you if they are around*
Steps	1. Greet clients 2. Briefly present yourself 3. Briefly present ComBaCaL cohort study 4. Wait for permission to take a seat or ask for it (outdoor or indoor) 5. Take tablet and the *introduction cheat sheet*
Say this	1. *According to the introduction cheat sheet:* • present yourself • present ComBaCaL • explain informed consent form 2. Let all clients sign the informed consent form (including children) 3. *Start creating the household in the app* 4. *Answer all required questions while creating the household* 5. *Submit*
Notes	• Basotho houses in rural areas differ from houses in western countries in that you already enter a more private part of the property, even if you have not yet stepped through a front door. Because many houses consist of only 1 room, one already enters the bedroom when entering the house interior. The properties are fenced and when you pass through the fence gate, however, you cannot knock. The observations during the trainings have shown that the CHWs replace the ringing or knocking at the door with a loud "Koko." This should be loud enough so that it can he heard over the entire property and the owner will be aware of it if they are home.

[a]CHW: community health worker.

[b]ComBaCaL: Community Based Chronic Disease Care Lesotho.

Results

Overview

In this section, we report on the evaluation outcomes of Scripted Medicine in the field study in ComBaCaL. We evaluated the attainment of the design goal (ie, empower CHWs to accept broader responsibilities in NCD care) in 2 rounds. First, we examined Scripted Medicine's effect on CHW empowerment directly after the training. Second, we highlighted changes in the CHWs' perception after approximately 3 months in the field.

CHW Empowerment After the Training

The video recordings captured the CHWs' application of the counseling ThinkLets. While the interactions in practice sessions at the beginning appeared artificial, the CHWs felt noticeably more comfortable interacting with clients in the nearby villages. The analysis showed that most CHWs closely followed the ThinkLets. However, they had already started adjusting the scripts to their liking. For example, instead of enumerating a household with everyone present, they opted to enumerate each member individually and asked others to leave. Later, the CHWs explained that due to the personal questions in the general health questionnaire (eg, HIV status), they wanted to create a more secure environment.

The medical algorithms in the ComBaCaL app scored 74.25 on the SUS. This indicated that the CHWs perceive the medical algorithms and the ComBaCaL app as useful in counseling clients. The analysis of the video recordings supports this indication. Most CHWs navigated the app without difficulties at the end of the practice-oriented training.

The empowerment questionnaire showed that the training empowered the CHWs. On average, the CHWs rated their structural empowerment at 4.5 on a 5-point Likert scale (Table 3; Multimedia Appendix 3 provides further details). In the focus group discussions, many referred to the clear structure provided by the medical algorithms and ThinkLets. One CHW said the following:

> *I thought it's going to be hard to go through [all the procedures]. But then, when we proceeded, everything became clearer. The app reminds me, and things come easy.*

XSL•FO
RenderX

Table 3. Evaluation results of community health worker empowerment on a 5-point Likert scale and System Usability Scale (range 0-100).

Measure and dimension	After training	After field experience	Trend
Structural empowerment (5-point Likert scale)			
Access to opportunities	4.5	4.6	↑
Access to information	4.6	4.6	—[a]
Access to support	4.3	4	↓
Access to resources	4.6	4.3	↓
Psychological empowerment (5-point Likert scale)			
Meaning	4.9	4.8	↓
Competence	4.6	4.5	↓
Autonomy	4.1	4.2	↑
Impact	3.8	4	↑
System Usability Scale (range 0-100)			
Total score	74.25	76.5	↑

[a]Not available.

The training approach also resonated well with the CHWs. They could make mistakes in a safe environment, where they received feedback from trainers and peers. Further, the Butha Buthe CHWs agreed in the focus group discussion that the last days in the nearby village greatly benefited them in interacting with strangers. Back in their community, their clients mostly will not be strangers but people they already know to a certain degree, which would make the interactions even easier. However, some CHWs, particularly those from Mokhotlong (ie, the more rural district), raised infrastructure concerns such as the lack of electricity to charge their tablets. Consequently, the Mokhotlong CHWs rated their access to resources as 14% lower (4.27 compared to 4.87) and access to support as 17% lower than their Butha Buthe colleagues (3.93 compared to 4.6).

The CHWs also felt empowered as they rated their combined psychological empowerment at 4.35 out of 5. Notably, impact is the only dimension rated <4, at 3.8. Again, the Mokhotlong CHWs rated the perceived impact noticeably lower (26%, 3.33 compared to 4.2). In the focus group discussions, some feared that people might refuse their services due to skepticism, as NCDs "are for the rich people" and would only affect "people who are older."

Conversely, both groups perceived an exceptional meaning in their future work (4.9 out of 5), which one CHW explained with this:

> *We don't want to see an increased number of patients diagnosed with diabetes and hypertension, and also, we don't want to see people who already have hypertension or diabetes die due to complications of those diseases.*

Another CHW highlighted the benefits of the counseling ThinkLets. She described herself as shy, "but this time, I've seen that I'm not shy anymore, and I can speak freely to everyone." The Mokhotlong CHWs all rated meaning with the maximum score and, thereby, slightly higher than the Butha Buthe CHWs. Further, all CHWs felt confident with their new skills and roles. One CHW said the following:

> *I don't think there will be challenges. The people in my village are waiting for me.*

However, the CHWs also raised some initial concerns. The concerns were primarily related to their acceptance, as community members might not be willing to accept their services. Both groups would prefer a project representative to introduce them to their village to gain initial credibility. Some mentioned that weather conditions influencing their ability to travel were a concern, especially in winter. These concerns often evolved around the loss of control, which could explain the CHWs' lower perception of autonomy compared to the other dimension (4.1 out of 5 on a 5-point Likert scale).

When asked what they expect upon returning, one CHW emphasizes the collaboration with patients as follows:

> *We're expecting to work hand in hand with patients. And being kind to them, to speak to each other. I think the most important thing is the idea of ComBaCaL to come through, and this is going to happen by the way we'll be working. We are going to be sure that each and everyone will be happy with our services.*

Despite raising some concerns, the CHWs felt well-prepared for their new responsibilities when returning to their communities after the training.

CHW Empowerment After Field Experience

After approximately 12 weeks in the field, the CHWs still felt empowered and ranked the ComBaCaL app slightly higher at 76.5 points on the SUS (Table 3).

Over the 3 months, CHW empowerment showed some minor changes. As the CHWs are distributed across 2 districts in Lesotho, they appeared to perceive less access to support (from 4.3 to 4 out of 5). During the training, they had access to supervisors, experts, and peers to ask questions and receive feedback. After returning to their communities, the CHWs

XSL•FO
RenderX

mainly worked alone. In the interview, one CHW proposed adding "a chatting platform, where people are ready to help in case of seeking for help; on which our instructors are able to communicate with us." Further, they had fewer opportunities to acquire new skills than during the training (from 4.6 to 4.3). Some mentioned in the interview that adjusting from a training setting to working back in their community took some time.

The CHWs perceived the same or even slightly better access to information (from 4.5 to 4.6) and resources (remained at 4.6) compared to the training. One CHW refers to the tablet as the source of information and resources as follows:

> *It helps a lot since it shows all the activities we were taught in training. It shows which people to attend the following day and how many people I have offered services to. It...helps me to keep track of people who are served and others to be served. If I happen to make a mistake in my activity, it reminds me because it shows that something is wrong.*

The CHWs' psychological empowerment also underwent some changes. Despite their high confidence in their skills, most CHWs report challenges they encountered in their community. Accordingly, perceived meaning and competence slightly decreased. A CHW mentioned convincing young adults as a challenging task as follows:

> *The main challenge is that the youngest adults refuse to agree when we test for them. But after talking to them about the importance of this, then they agree.*

Interestingly, most CHWs reported similar experiences of persuading reluctant clients when asked about their greatest success so far. For example, another CHW said: "what made me happy is the joy of being able to convince a person, or generally making a person have interest in matters of health." They could persuade clients to accept their services despite the initial skepticism. This success might have increased perceived autonomy and impact, as "those one I helped are very impressed since ComBaCaL is taking care of their lives door to door, and they even gained some information on how to take care of themselves." After translating their training into practice, the CHWs could see the impact of their work. This gave them confidence in their capabilities and resulted in more trust from their patients, as one CHW summarized:

> *Most of my clients see the confidence in me when I walk; when I touch anything. So, when you enter the door of the household, you should look confident and say clear things so clients are able to trust you. [The training] was basic because it helped me overcome my fear. That was my first action when I started; to see how to work with a patient. So, I learned there, and I was able to move forward with the knowledge.*

Interestingly, the interviews showed that medical algorithms dominated the perception of structural empowerment. However, psychological empowerment is often attributed to the ThinkLets that embed clinical knowledge in client interactions. As in the example earlier, the CHW refers to "work with a patient" and their confidence in their interactions with the patient. Hence, combining medical and counseling knowledge empowered CHWs to accept their responsibilities in ComBaCaL.

Discussion

Principal Findings

This study aimed to design Scripted Medicine, which is an approach to empower CHWs to accept broader responsibilities in NCD care. We grounded the approach in research on collaboration engineering and algorithmic management. Restricting CHWs in the process facilitated the transfer of power (ie, structural empowerment) that led to psychological empowerment. As hypothesized, the empowered CHWs could accept the broadened responsibilities and offer more holistic NCD care to their community. These results allow us to make a 2-fold contribution.

Scripted Medicine

Scripted Medicine is a novel approach inspired by existing research on collaboration engineering and algorithmic management. It is the product of scripting the necessary medical and counseling expertise. Scripted Medicine combines previously separated knowledge spheres to create significant empowerment potential in community-based health care. Process restrictions empower CHWs by conveying the relevant knowledge required for their work.

Medical algorithms empower CHWs to perform medical procedures [40-42]. Scripted Medicine broadens the perspective on medical algorithms by drawing on algorithmic management literature [43,44,47]. The 5RSM framework provides a nuanced distinction of algorithmic management forms that influence how researchers design and study medical algorithms [47]. For example, our evaluation highlights the necessity for carefully considering algorithmic requiring and recommending. CHWs might feel disempowered when requiring noncritical activities instead of providing certain flexibility in work planning. On the contrary, previous research reports seemingly contradictory observations in CHW empowerment. CHWs perceived algorithmic restricting as empowering instead of a loss of autonomy [54]. This observation aligns with collaboration engineering research [59,60]. Accordingly, Scripted Medicine explains how restricting CHWs in the process (eg, restricting unsafe actions) while allowing enough flexibility for users to adapt to their context is empowering rather than disempowering. It also emphasizes the "well-designed" when instantiating clinical knowledge in mHealth tools. Consequently, we formulate the design principle of scripted medical activities (Textbox 2 [79]).

XSL•FO
RenderX

Textbox 2. The design principle of scripted medical activities following the study by Gregor et al.

Title

- Design principle of scripted medical activities

Aim, implementer, and user

- For designers and researchers (implementers) to efficiently and effectively convey medical knowledge (aim; solution objective [SO] medical knowledge) by drawing on the algorithmic recommending, restricting, requiring, rating, rewarding, sanctioning, and monitoring (5RSM) framework to guide community health workers (CHWs; users) through carefully designing a series of medical activities

Context

- Community-based health care heavily relies on medical algorithms to enable CHWs to perform medical procedures (eg, diagnostics) safely and independently [40-42]. While requiring CHWs to follow a specific path creates more reliability, it ignores the individual preferences of CHWs and might lead to conflicts with other responsibilities (eg, families). This could result in dropouts [13,14].

Mechanism

- Provide medical knowledge to CHWs through scripts based on algorithms (eg, decision trees), allowing them to have agency over noncritical medical activities

Rationale

- Collaboration engineering allows nonexperts to perform activities by restricting their process flexibility [59]. While this has an empowering effect, certain forms of algorithmic management (ie, recommending and requiring) can have opposite effects [47]. Our results have shown that requiring a specific order in which tasks are completed feels too restrictive. Allowing CHWs to plan their work independently, within limits, offers them an agency that is perceived as empowering. Accordingly, the 5RSM framework offers more nuanced forms of algorithmic management that acknowledge the need for reliability and safety as well as independence and agency.

Scripted Medicine further argues that formalizing clinical knowledge alone is insufficient for sustainable community-based health care. While medical algorithms provide the foundation to expand CHWs' responsibilities, they do not adequately address all relevant aspects of empowering CHWs. Medical algorithms cater to 2 out of 3 aspects of process restrictions as follows: (1) careful design of a series of activities and (2) configuration of technology to support behaviors within each activity. However, they lack the final ingredient that translates the clinical knowledge into its application context: (3) guidance on the behaviors performed in each activity to maneuver CHWs and their clients through the counseling process.

Scripted Medicine addresses a significant shortcoming in current research that often assumes that medical algorithms alone empower CHWs. This limited perspective results in unilateral solutions that solely focus on the technical dimension of CHW empowerment [29,30,54]. However, CHW acceptance and overall health outcomes also depend on the quality of the social interactions between CHWs and their clients [52,53]. Yet, counseling knowledge is seldom in scope in community-based health programs. Scripted Medicine extends the existing work on CHW empowerment as it embeds clinical knowledge into the context where it is applied. Our evaluation indicated differences between Mokhotlong and Butha Buthe CHWs. These differences highlight the importance of considering the social context. Both CHW groups received the same scripted expertise and training but showed different perceptions of structural and psychological empowerment. While the survey only offered indications, our observations and interviews allow us to argue that the slightly different contexts between the 2 districts cause the indicated differences. People in Mokhotlong might be more skeptical of Western medicine in general. Hence, CHWs might believe the community will be reluctant to accept their services and perceive less impact of their work. Further, Mokhotlong CHWs raise the issue of lacking access to electricity or cellular signals, while their Butha Buthe colleagues do not. Infrastructure differences also alter empowerment perceptions. Mokhotlong CHWs perceive less access to resources (ie, the tablet could run out of battery) and support (ie, no means to call or text a supervisor). Scripted Medicine obliges designers and researchers to consider the context within which they operate carefully.

Counseling ThinkLets serve multiple purposes in Scripted Medicine. First, ThinkLets were useful training modules for counseling clients with the support of medical algorithms. Trainers can structure training along ThinkLets to maintain a logical order. ThinkLets can first be discussed and demonstrated in the plenum before moving to practice-oriented training sessions. In these sessions, CHWs apply ThinkLets to build expertise in a safe environment and receive feedback from trainers and peers. Second, ThinkLets are individual learning modules beyond official training sessions. CHWs can always revert to them in case they feel insecure. For example, CHWs might not perform specific tasks regularly and want to refresh their memory before counseling clients. Third, ThinkLets are intended to provide a foundation for social interactions in counseling. While the medical algorithm provides continuous constraints for care procedures, ThinkLets should "kickstart" the CHWs' counseling practices that they adapt over time as their expertise grows. Finally, ThinkLets allow designers and researchers to study the evolving routines of individual CHWs when performing their work. These insights can be formalized and shared with other CHWs to make knowledge accessible within the program.

Scripted Medicine emphasizes the sociotechnical context, which is often lacking in existing literature on technology-supported

XSL•FO
RenderX

community–based health care. It allows research to go beyond medical outcomes and consider CHWs' difficult position at the margin of different communities (eg, their families and the health system), often without fully belonging to any of them [29]. In combination, counseling ThinkLets and medical algorithms facilitate the structural empowerment of CHWs as they gain access to information, resources, support, and opportunities. Scripted Medicine postulates a more comprehensive view of structural empowerment, as optimizing one aspect might not cater to all dimensions. For example, medical algorithms might reduce or even prevent access to opportunities. CHWs cannot deviate from the algorithm's predetermined workflow to ensure patient safety. On the other hand, counseling ThinkLets open plenty of opportunities for CHWs to develop their personal way of interacting with clients. They empower CHWs to transfer medical procedures from isolation into the social interactions of health services. On the basis of these insights, we formulate the design principle of scripted counseling (Textbox 3 [79]).

Textbox 3. The design principle of scripted counseling following the study by Gregor et al.

Title

- Design principle of scripted counseling

Aim, implementer, and user

- For designers and researchers (implementers) to efficiently and effectively convey counseling knowledge (aim; solution objective [SO] counseling knowledge) to support community health workers (CHWs; users) in the creation of their own, culturally adapted counseling practices

Context

- Digital health tools for community-based health care primarily convey medical knowledge to CHWs [29,30,54]. While medical knowledge is crucial, health care is always an interaction between people. Accordingly, existing research does not support CHWs enough to transfer this medical knowledge into action.

Mechanism

- Provide counseling knowledge to CHWs through ThinkLets (ie, communication scripts), allowing them to scaffold their advice and to develop their counseling practices in accordance with their preferences and the cultural context

Rationale

- CHW acceptance and health outcomes significantly rely on social interactions [52,53]. Counseling ThinkLets require designers to consider the context for which they design and allow CHWs to develop their own practices that suit their preferences and strengths. Further, it allows tailoring to nuances within specific cultural contexts. Our findings indicated differences between the CHWs from the more urban Butha Buthe and the rural Mokhotlong.

Scripted Medicine is more than stringing together medical and counseling scripts. These scripts are interdependent. Explicating medical knowledge in the form of guidelines has long been established. Counseling knowledge, as established in the introduction, is often implicit and gained with years of training and experience. Accordingly, guidelines for medical professionals focus on essential clinical information. They assume medical professionals possess the counseling experience to translate clinical information into the consultation context.

The literature often neglects that CHWs require more guidance than medical professionals. Scripted Medicine is an approach to formalizing both areas of expertise, as clinical knowledge alone is insufficient. For example, CHWs can counsel people if they can rely on proper medical guidance and access relevant clinical information about patients and medical procedures. However, access to that data and being able to implement the medical procedures depends on a good working relationship with patients. Establishing and maintaining such a relationship requires guidance on counseling practices.

Scripted Medicine considers the symbiosis of medical and counseling knowledge and their interdependence. Accordingly, it is more than merely a combination of clinical and counseling guidance. For example, medical algorithms might include different media types for CHWs to educate patients about their condition after a diagnosis. Further, ThinkLets must consider the process restrictions implemented in algorithms to avoid unnatural interactions. This combination of formalized knowledge in Scripted Medicine is the foundation for psychological empowerment. While clinical knowledge itself might enable CHWs to conduct medical procedures, CHWs perceive the embedding in social interactions as empowering. It allows CHWs to see the meaning in their work, feel competent in counseling clients, autonomously plan their work, and observe their work's impact on their community. Thus, we formulate the design principle of Scripted Medicine (Textbox 4 [79]).

XSL•FO
RenderX

Textbox 4. The design principle of Scripted Medicine following the study by Gregor et al.

Title

- Design principle of scripted medicine

Aim, implementer, and user

- For designers and researchers (implementers) to meaningfully intertwine formalized medical and counseling knowledge (aim) to empower community health workers (CHWs; users) to accept broader responsibilities in noncommunicable disease care

Context

- On one side, CHWs depend on medical algorithms to offer health services. On other side, medical algorithms depend on the data inserted by CHWs during consultations. Thus, CHWs require a tool that carefully aligns the designed medical activities with counseling practices to carry out more complex tasks successfully.

Mechanism

- Integrate formalized medical and counseling knowledge visually and algorithmically to empower CHWs

Rationale

- Collaboration engineering is an approach that carefully considers the effects of restricting and guiding users through activities and interactions [57,59]. IT-supported community–based health care is a sociotechnical system in which people, tasks, systems, and organizational structures (communities) are inherently interdependent [54]. Thus, CHWs require access to information, resources, support, and opportunities that consider the social and technical dimensions of community-based health care.

Becoming Agents of Social Change

Existing research provides in-depth insights into the design of mHealth to capacitate CHWs to perform actions they could not perform otherwise, such as diabetes screening [5,34,37,41]. While this work is fundamental for expanding CHWs' responsibilities, it fosters a unilateral focus on conveying medical knowledge through mHealth. It has been argued that most studies view CHWs as sensors of the health system, leading to a reductionist perspective on community-based health care [54]. In other words, existing research almost exclusively focuses on the technical domain in a sociotechnical system and omits the social dimension. This situation was mentioned in a critique by Engel [82] on biomedical medicine as follows:

> *The crippling flaw of the [biomedical] model is that it does not include the patient and [their] attributes as a person, as a human being. Yet, in the everyday work of the physician, the prime object of study is a person.*

Similarly, existing research does not consider community-based health care as a collaborative process between CHWs and their clients. However, multiple studies highlight the importance of social interactions in community-based health care for patient experience and treatment adherence [52,53]. Scripted Medicine offers a lens that emphasizes the collaborative nature of community-based health care. While the immense benefits of community-based health care are undoubted, we criticize the often reductionist consideration of health outcomes only. mHealth might increase CHW performance and retention or capacitate them to perform more complex tasks guided by medical algorithms [36]. However, they do not necessarily address the root causes of low performance and high attrition (eg, lack of organizational support and appreciation [9,14]). Accordingly, mHealth could even worsen the situation for CHWs. Their responsibilities continuously broaden and become more complex, while their work conditions do not necessarily improve at the same rate. From an empowerment perspective, current research focuses on shifting responsibilities and often forgets shifting power. This could lead to CHWs feeling more disempowered and increase instead of decrease attrition rates over time.

Scripted Medicine aims to address such structural issues. It accentuates the interdependence between the 2 dimensions of health care: the medical and outcome-oriented procedure (health) and the interpersonal, collaborative character (care). For example, existing reports on health outcomes of community-based health initiatives are limited to people willing to accept CHWs' services. Reluctant people do not benefit from improved access to health care and, consequently, are usually not considered in such studies. Our evaluation indicates that Scripted Medicine might empower CHWs to better persuade this population to accept their services. By considering both dimensions of health care, we believe CHWs become better service providers and become the "agents of social change, beyond their role as links between the community and the health system." [50]

Conclusions and Limitations

Task shifting to CHWs is a promising remedy to the imminent shortage of medical professionals in the Global South [4,5]. However, community-based health programs often face performance variability and high attrition rates [9-13]. As a result, task-shifting efforts might be built on unstable foundations if the root causes of the existing challenges are not resolved.

Scripted Medicine provides a more comprehensive perspective to empower CHWs to accept broader responsibilities in NCD care. Scripted Medicine "scripts" clinical and counseling knowledge that promises to quickly empower CHWs to accept broader responsibilities and become active members of the

XSL•FO
RenderX

health system. It draws on research on collaboration engineering and algorithmic management to restrict CHWs in the counseling process in a way that empowers them to offer more complex health services, such as first-line treatment for uncomplicated type 2 diabetes and hypertension.

Scripted Medicine positions CHWs as active members and not sensors of the health system. Scripting medical and counseling knowledge transfers the power necessary to offer high-quality health services to CHWs. In turn, CHWs feel empowered to accept new responsibilities and, thereby, increase the accessibility of health care. Scripted Medicine highlights the need for a deeper understanding of the local context and the design of contextually adapted community-based health programs to support CHWs in their work.

However, this study does not come without limitations. Scripted Medicine broadens the unilateral focus on the technical dimension by emphasizing social interactions. However, more research is needed to study other potential dimensions influencing CHW empowerment. Empowerment is a dynamic concept that requires a longitudinal assessment of changing CHW demands [35]. Finally, the rich dataset allowed us to triangulate Scripted Medicine from empirical data. Future work is needed to validate these empirical observations at a larger scale.

Acknowledgments

This study is part of the Community Based Chronic Care Project Lesotho (ComBaCaL). The ComBaCaL project is funded by the TRANSFORM grant of the Swiss Agency for Development and Cooperation (project 7F-10345.01.01) and a grant by the World Diabetes Foundation (WDF-1778) to SolidarMed. The authors would like to thank the CHWs interviewed in this study and all involved ComBaCaL staff members, especially Thabo Lejone, who organized the community health workers' training and assisted with the interviews.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The empowerment questionnaire as administered in the study with Sesotho translation, adapted from a study by Spence Laschinger et al.
[PDF File (Adobe PDF File), 86 KB - humanfactors_v12i1e57545_app1.pdf]

Multimedia Appendix 2
The System Usability Scale as administered in the study with Sesotho translation, adapted from a study by Bangor et al.
[PDF File (Adobe PDF File), 55 KB - humanfactors_v12i1e57545_app2.pdf]

Multimedia Appendix 3
Survey results for the System Usability Scale and structural and psychological empowerment per community health worker.
[XLSX File (Microsoft Excel File), 11 KB - humanfactors_v12i1e57545_app3.xlsx]

References

1. Global strategy on human resources for health: workforce 2030. World Health Organization. 2016. URL: https://iris.who.int/bitstream/handle/10665/250368/9789241511131-eng.pdf [accessed 2025-04-03]
2. Boniol M, Kunjumen T, Nair TS, Siyam A, Campbell J, Diallo K. The global health workforce stock and distribution in 2020 and 2030: a threat to equity and 'universal' health coverage? BMJ Glob Health 2022 Jun 27;7(6):e009316 [FREE Full text] [doi: 10.1136/bmjgh-2022-009316] [Medline: 35760437]
3. Perry HB, Zulliger R, Rogers MM. Community health workers in low-, middle-, and high-income countries: an overview of their history, recent evolution, and current effectiveness. Annu Rev Public Health 2014;35:399-421. [doi: 10.1146/annurev-publhealth-032013-182354] [Medline: 24387091]
4. Fernández LG, Firima E, Robinson E, Ursprung F, Huber J, Amstutz A, et al. Community-based care models for arterial hypertension management in non-pregnant adults in sub-Saharan Africa: a literature scoping review and framework for designing chronic services. BMC Public Health 2022 Jun 04;22(1):1126 [FREE Full text] [doi: 10.1186/s12889-022-13467-4] [Medline: 35658850]
5. Firima E, Gonzalez L, Huber J, Belus JM, Raeber F, Gupta R, et al. Community-based models of care for management of type 2 diabetes mellitus among non-pregnant adults in sub-Saharan Africa: a scoping review protocol. F1000Res 2021 Jul 5;10:535. [doi: 10.12688/f1000research.52114.1]
6. Joshi R, Thrift AG, Smith C, Praveen D, Vedanthan R, Gyamfi J, et al. Task-shifting for cardiovascular risk factor management: lessons from the Global Alliance for Chronic Diseases. BMJ Glob Health 2018 Nov 08;3(Suppl 3):e001092 [FREE Full text] [doi: 10.1136/bmjgh-2018-001092] [Medline: 30483414]

XSL•FO
RenderX

7. Lehmann U, van Damme W, Barten F, Sanders D. Task shifting: the answer to the human resources crisis in Africa? Hum Resour Health 2009 Jun 21;7(1):49 [FREE Full text] [doi: 10.1186/1478-4491-7-49] [Medline: 19545398]
8. Optimizing community health worker programmes for HIV services: a guide for health policy and system support. World Health Organization. 2021 Nov 30. URL: https://www.who.int/publications/i/item/9789240040168 [accessed 2025-04-03]
9. Rassi C, Gore-Langton GR, Gidudu Walimbwa B, Strachan CE, King R, Basharat S, et al. Improving health worker performance through text messaging: a mixed-methods evaluation of a pilot intervention designed to increase coverage of intermittent preventive treatment of malaria in pregnancy in West Nile, Uganda. PLoS One 2018;13(9):e0203554 [FREE Full text] [doi: 10.1371/journal.pone.0203554] [Medline: 30188956]
10. Rowe AK, de Savigny D, Lanata CF, Victora CG. How can we achieve and maintain high-quality performance of health workers in low-resource settings? Lancet 2005 Sep;366(9490):1026-1035. [doi: 10.1016/s0140-6736(05)67028-6] [Medline: 16168785]
11. Aninanya GA, Williams JE, Williams A, Otupiri E, Howard N. Effects of computerized decision support on maternal and neonatal health-worker performance in the context of combined implementation with performance-based incentivisation in Upper East Region, Ghana: a qualitative study of professional perspectives. BMC Health Serv Res 2022 Dec 26;22(1):1581 [FREE Full text] [doi: 10.1186/s12913-022-08940-0] [Medline: 36567357]
12. Ngugi AK, Nyaga LW, Lakhani A, Agoi F, Hanselman M, Lugogo G, et al. Prevalence, incidence and predictors of volunteer community health worker attrition in Kwale County, Kenya. BMJ Glob Health 2018 Jul 31;3(4):e000750 [FREE Full text] [doi: 10.1136/bmjgh-2018-000750] [Medline: 30105093]
13. Nkonki L, Cliff J, Sanders D. Lay worker attrition: important but often ignored. Bull World Health Org 2011 Dec 01;89(12):919-923. [doi: 10.2471/blt.11.087825]
14. Olang'o CO, Nyamongo IK, Aagaard-Hansen J. Staff attrition among community health workers in home-based care programmes for people living with HIV and AIDS in western Kenya. Health Policy 2010 Oct;97(2-3):232-237 [FREE Full text] [doi: 10.1016/j.healthpol.2010.05.004] [Medline: 20807685]
15. Rawal L, Jubayer S, Choudhury SR, Islam SM, Abdullah AS. Community health workers for non-communicable diseases prevention and control in Bangladesh: a qualitative study. Glob Health Res Policy 2020 Dec 24;6(1):1 [FREE Full text] [doi: 10.1186/s41256-020-00182-z] [Medline: 33407942]
16. de Vries DH, Pool R. The influence of community health resources on effectiveness and sustainability of community and lay health worker programs in lower-income countries: a systematic review. PLoS One 2017 Jan 17;12(1):e0170217 [FREE Full text] [doi: 10.1371/journal.pone.0170217] [Medline: 28095475]
17. Spence Laschinger HK, Leiter M, Day A, Gilin D. Workplace empowerment, incivility, and burnout: impact on staff nurse recruitment and retention outcomes. J Nurs Manag 2009 Apr;17(3):302-311. [doi: 10.1111/j.1365-2834.2009.00999.x] [Medline: 19426367]
18. Cicolini G, Comparcini D, Simonetti V. Workplace empowerment and nurses' job satisfaction: a systematic literature review. J Nurs Manag 2014 Oct;22(7):855-871. [doi: 10.1111/jonm.12028] [Medline: 25298049]
19. Orgambídez A, Almeida H. Exploring the link between structural empowerment and job satisfaction through the mediating effect of role stress: a cross-sectional questionnaire study. Int J Nurs Stud 2020 Sep;109:103672. [doi: 10.1016/j.ijnurstu.2020.103672] [Medline: 32569828]
20. Laschinger HK. A theoretical approach to studying work empowerment in nursing: a review of studies testing Kanter's theory of structural power in organizations. Nurs Adm Q 1996;20(2):25-41. [doi: 10.1097/00006216-199602020-00006] [Medline: 8700383]
21. Kanter RM. Men and Women of the Corporation: New Edition. New York, NY: Basic Books; 1977.
22. Laschinger HK, Finegan J, Shamian J, Wilk P. Impact of structural and psychological empowerment on job strain in nursing work settings: expanding Kanter's model. J Nurs Adm 2001 May;31(5):260-272. [doi: 10.1097/00005110-200105000-00006] [Medline: 11388162]
23. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977 Mar;84(2):191-215. [doi: 10.1037//0033-295x.84.2.191] [Medline: 847061]
24. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000 Jan;55(1):68-78. [doi: 10.1037//0003-066x.55.1.68] [Medline: 11392867]
25. Spreitzer GM. Psychological empowerment in the workplace: dimensions, measurement, and validation. Acad Manag J 1995 Oct 01;38(5):1442-1465. [doi: 10.2307/256865]
26. Monje-Amor A, Xanthopoulou D, Calvo N, Abeal Vázquez JP. Structural empowerment, psychological empowerment, and work engagement: a cross-country study. Eur Manag J 2021 Dec;39(6):779-789. [doi: 10.1016/j.emj.2021.01.005]
27. Duffy S, Svenson J, Chavez A, Kelly M, Wise P. Empowering community health workers with mobile technology to treat diabetes. Ann Fam Med 2019 Mar 11;17(2):176 [FREE Full text] [doi: 10.1370/afm.2361] [Medline: 30858263]
28. Linda I, Santoso H, Lubis Z, Siregar MF. Empowering village health workers to influence preconception behavior and increase utilization of preconception health services in the Deli serdang regency. Open Access Maced J Med Sci 2022 Jan 12;10(E):191-197. [doi: 10.3889/oamjms.2022.8219]

XSL•FO
RenderX

29. Ismail A, Kumar N. Empowerment on the margins: the online experiences of community health workers. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 2019 Presented at: CHI '19; May 4-9, 2019; Glasgow, UK. [doi: 10.1145/3290605.3300329]
30. Nyanja N, Nyamu N, Nyaga L, Chabeda S, Lusambili A, Temmerman M, et al. Application of the Ultra-Poverty Graduation Model in understanding community health volunteers' preferences for socio-economic empowerment strategies to enhance retention: a qualitative study in Kilifi, Kenya. Hum Resour Health 2021 Aug 28;19(1):104 [FREE Full text] [doi: 10.1186/s12960-021-00645-5] [Medline: 34454506]
31. Altobelli LC. Sharing histories-a transformative learning/teaching method to empower community health workers to support health behavior change of mothers. Hum Resour Health 2017 Aug 23;15(1):54 [FREE Full text] [doi: 10.1186/s12960-017-0231-2] [Medline: 28835240]
32. Becker J, Kovach AC, Gronseth DL. Individual empowerment: how community health workers operationalize self‐determination, self‐sufficiency, and decision‐making abilities of low‐income mothers. J Community Psychol 2004 Apr;32(3):327-342. [doi: 10.1002/jcop.20000]
33. Stanton M, Molineux A, Mackenzie C, Kelly-Hope L. Mobile technology for empowering health workers in underserved communities: new approaches to facilitate the elimination of neglected tropical diseases. JMIR Public Health Surveill 2016 Jan 14;2(1):e2 [FREE Full text] [doi: 10.2196/publichealth.5064] [Medline: 27227155]
34. Rodrigues SM, Kanduri A, Nyamathi A, Dutt N, Khargonekar P, Rahmani AM. Digital health-enabled community-centered care: scalable model to empower future community health workers using human-in-the-loop artificial intelligence. JMIR Form Res 2022 Apr 06;6(4):e29535 [FREE Full text] [doi: 10.2196/29535] [Medline: 35384853]
35. Staehelin D, Greve M, Schwabe G. Empowering community health workers with mobile health: learnings from two projects on noncommunicable disease care. 2023 Presented at: The 32nd European Conference on Information Systems: People First: Constructing Digital Futures Together; June 13-19, 2024; Paphos, Cyprus URL: https://research.cbs.dk/en/publications/empowering-community-health-workers-with-mobile-health-learnings-
36. Källander K, Tibenderana JK, Akpogheneta OJ, Strachan DL, Hill Z, ten Asbroek AH, et al. Mobile health (mHealth) approaches and lessons for increased performance and retention of community health workers in low- and middle-income countries: a review. J Med Internet Res 2013 Jan 25;15(1):e17 [FREE Full text] [doi: 10.2196/jmir.2130] [Medline: 23353680]
37. Zaman SB, Evans RG, Singh R, Singh A, Singh P, Singh R, et al. Feasibility of community health workers using a clinical decision support system to screen and monitor non-communicable diseases in resource-poor settings: study protocol. Mhealth 2021;7:15 [FREE Full text] [doi: 10.21037/mhealth-19-258] [Medline: 33634198]
38. Bakibinga P, Kamande E, Omuya M, Ziraba AK, Kyobutungi C. The role of a decision-support smartphone application in enhancing community health volunteers' effectiveness to improve maternal and newborn outcomes in Nairobi, Kenya: quasi-experimental research protocol. BMJ Open 2017 Jul 20;7(7):e014896 [FREE Full text] [doi: 10.1136/bmjopen-2016-014896] [Medline: 28729309]
39. Kateera F, Riviello R, Goodman A, Nkurunziza T, Cherian T, Bikorimana L, et al. The effect and feasibility of mHealth-supported surgical site infection diagnosis by community health workers after cesarean section in rural Rwanda: randomized controlled trial. JMIR Mhealth Uhealth 2022 Jun 08;10(6):e35155 [FREE Full text] [doi: 10.2196/35155] [Medline: 35675108]
40. Silveira DV, Marcolino MS, Machado EL, Ferreira CG, Alkmim MB, Resende ES, et al. Development and evaluation of a mobile decision support system for hypertension management in the primary care setting in Brazil: mixed-methods field study on usability, feasibility, and utility. JMIR Mhealth Uhealth 2019 Mar 25;7(3):e9869 [FREE Full text] [doi: 10.2196/mhealth.9869] [Medline: 30907740]
41. Marcolino MS, Oliveira JA, Cimini CC, Maia JX, Pinto VS, Sá TQ, et al. Development and implementation of a decision support system to improve control of hypertension and diabetes in a resource-constrained area in Brazil: mixed methods study. J Med Internet Res 2021 Jan 11;23(1):e18872 [FREE Full text] [doi: 10.2196/18872] [Medline: 33427686]
42. Cimini CC, Maia JX, Pires MC, Ribeiro LB, Pinto VS, Batchelor J, et al. Pandemic-related impairment in the monitoring of patients with hypertension and diabetes and the development of a digital solution for the community health worker: quasiexperimental and implementation study. JMIR Med Inform 2022 Mar 29;10(3):e35216 [FREE Full text] [doi: 10.2196/35216] [Medline: 35191842]
43. Schildt H. Big data and organizational design – the brave new world of algorithmic management and computer augmented transparency. Innovation 2016 Nov 08;19(1):23-30. [doi: 10.1080/14479338.2016.1252043]
44. Tarafdar M, Page X, Marabelli M. Algorithms as co‐workers: human algorithm role interactions in algorithmic work. Inf Syst J 2022 May 07;33(2):232-267. [doi: 10.1111/isj.12389]
45. Cronin P, Kimball A. Success of automated algorithmic scheduling in an outpatient setting. Am J Manag Care 2014 Jul;20(7):570-576 [FREE Full text] [Medline: 25295403]
46. Bai B, Dai H, Zhang DJ, Zhang F, Hu H. The impacts of algorithmic work assignment on fairness perceptions and productivity: evidence from field experiments. Manuf Serv Oper Manag 2022 Nov;24(6):3060-3078. [doi: 10.1287/msom.2022.1120]

XSL•FO
RenderX

47. Hirsch F, Alizadeh A, Wiener M, Cram AW. Algorithmic control in platform and traditional work settings: an updated conceptual framework. In: Proceedings of the European Conference on Information System. 2023 Presented at: ECIS 2023; June 11-16, 2023; Kristiansand, Norway URL: https://www.researchgate.net/publication/371694869_Algorithmic_Control_in_Platform_and_Traditional_Work_Settings_An_Updated_Conceptual_Framework
48. Dewi A, Supriyatiningsih S, Sundari S, Sugiyo D, Dewi DT. Community health worker empowerment through collaborative models with community midwifery. Int J Public Health Sci 2023 Jun 01;12(2). [doi: 10.11591/ijphs.v12i2.22431]
49. Cupertino AP, Saint-Elin M, de los Rios JB, Engelman KK, Greiner KA, Ellerbeck EF, et al. Empowering promotores de salud as partners in cancer education and research in rural southwest Kansas. Oncol Nurs Forum 2014 Dec 26;42(1):15-22. [doi: 10.1188/15.onf.15-22]
50. Kane S, Kok M, Ormel H, Otiso L, Sidat M, Namakhoma I, et al. Limits and opportunities to community health worker empowerment: a multi-country comparative study. Soc Sci Med 2016 Sep;164:27-34 [FREE Full text] [doi: 10.1016/j.socscimed.2016.07.019] [Medline: 27459022]
51. Staehelin D, Miscione G, Dolata M. From solution trap to solution patchwork: tensions in digital health in the global context. In: Proceedings of the 2023 International Conference on Information Systems. 2023 Presented at: ICIS 2023; December 10-13, 2023; Hyderabad, India URL: https://tinyurl.com/y47ea9rh
52. Jaskiewicz W, Tulenko K. Increasing community health worker productivity and effectiveness: a review of the influence of the work environment. Hum Resour Health 2012 Sep 27;10:38 [FREE Full text] [doi: 10.1186/1478-4491-10-38] [Medline: 23017131]
53. McCollum R, Gomez W, Theobald S, Taegtmeyer M. How equitable are community health worker programmes and which programme features influence equity of community health worker services? A systematic review. BMC Public Health 2016 May 20;16(1):419 [FREE Full text] [doi: 10.1186/s12889-016-3043-8] [Medline: 27207151]
54. Staehelin D, Dolata M, Peyer N, Gerber F, Schwabe G. Algorithmic management for community health worker in sub-Saharan Africa: curse or blessing? In: Proceedings of the Human-Computer Interaction – INTERACT 2023. 2023 Presented at: INTERACT 2023; August 28-September 1, 2023; York, UK. [doi: 10.1007/978-3-031-42286-7_6]
55. Nunamaker JF, Dennis AR, Valacich JS, Vogel D, George JF. Electronic meeting systems. Commun ACM 1991 Jul;34(7):40-61. [doi: 10.1145/105783.105793]
56. Dennis AR, Wixom BH, Vandenberg RJ. Understanding fit and appropriation effects in group support systems via meta-analysis. MIS Q 2001 Jun;25(2):167-193. [doi: 10.2307/3250928]
57. Briggs RO, de Vreede GJ, Nunamaker JFJ. Collaboration engineering with ThinkLets to pursue sustained success with group support systems. J Manag Inf Syst 2014 Dec 23;19(4):31-64. [doi: 10.1080/07421222.2003.11045743]
58. Randrup NL, Briggs RO. Evaluating the performance of collaboration engineers. In: Proceedings of the 48th Hawaii International Conference on System Sciences. 2018 Presented at: HICSS 2018; January 5-8, 2015; Kauai, HI. [doi: 10.1109/hicss.2015.78]
59. Briggs RO, Kolfschoten GL, de Vreede GJ, Lukosch S, Albrecht CC. Facilitator-in-a-box: process support applications to help practitioners realize the potential of collaboration technology. J Manag Inf Syst 2014 Dec 08;29(4):159-194. [doi: 10.2753/mis0742-1222290406]
60. de Vreede GJ, Briggs RO. A program of collaboration engineering research and practice: contributions, insights, and future directions. J Manag Inf Syst 2019 Mar 31;36(1):74-119. [doi: 10.1080/07421222.2018.1550552]
61. Briggs RO, de Vreede GJ, Nunamaker JF, Tobey D. ThinkLets: achieving predictable, repeatable patterns of group interaction with group support systems (GSS). In: Proceedings of the 34th Annual Hawaii International Conference on System Sciences. 2001 Presented at: HICSS 2001; January 6, 2001; Maui, HI. [doi: 10.1109/hicss.2001.926238]
62. Peffers K, Tuunanen T, Rothenberger MA, Chatterjee S. A design science research methodology for information systems research. J Manag Inf Syst 2014 Dec 08;24(3):45-77. [doi: 10.2753/MIS0742-1222240302]
63. Simon HA. The Sciences of the Artificial. Cambridge, MA: The MIT Press; 1996.
64. Subramanian H. A decentralized marketplace for patient-generated health data: design science approach. J Med Internet Res 2023 Feb 27;25:e42743 [FREE Full text] [doi: 10.2196/42743] [Medline: 36848185]
65. Lapão LV, Peyroteo M, Maia M, Seixas J, Gregório J, Mira da Silva M, et al. Implementation of digital monitoring services during the COVID-19 pandemic for patients with chronic diseases: design science approach. J Med Internet Res 2021 Aug 26;23(8):e24181 [FREE Full text] [doi: 10.2196/24181] [Medline: 34313591]
66. Subramanian H, Subramanian S. Improving diagnosis through digital pathology: proof-of-concept implementation using smart contracts and decentralized file storage. J Med Internet Res 2022 Mar 28;24(3):e34207 [FREE Full text] [doi: 10.2196/34207] [Medline: 35343905]
67. Miranda A, Miah SJ. Design and evaluation of a contextual model for information retrieval from web-scale discovery services to improve evidence-based practice by health care practitioners: mixed methods study. J Med Internet Res 2019 Aug 21;21(8):e12621 [FREE Full text] [doi: 10.2196/12621] [Medline: 31436167]
68. Walters LE, Scott RE, Mars M. A teledermatology scale-up framework and roadmap for sustainable scaling: evidence-based development. J Med Internet Res 2018 Jun 20;20(6):e224 [FREE Full text] [doi: 10.2196/jmir.9940] [Medline: 29925492]
69. Amstutz A, Lejone TI, Khesa L, Kopo M, Kao M, Muhairwe J, et al. Offering ART refill through community health workers versus clinic-based follow-up after home-based same-day ART initiation in rural Lesotho: the VIBRA cluster-randomized

XSL•FO
RenderX

clinical trial. PLoS Med 2021 Oct 21;18(10):e1003839 [FREE Full text] [doi: 10.1371/journal.pmed.1003839] [Medline: 34673765]
70. Amstutz A, Lejone TI, Khesa L, Muhairwe J, Bresser M, Vanobberghen F, et al. Home-based oral self-testing for absent and declining individuals during a door-to-door HIV testing campaign in rural Lesotho (HOSENG): a cluster-randomised trial. Lancet HIV 2020 Nov;7(11):e752-e761. [doi: 10.1016/s2352-3018(20)30233-2] [Medline: 33045193]
71. Labhardt ND, Motlomelo M, Cerutti B, Pfeiffer K, Kamele M, Hobbins MA, et al. Home-based versus mobile clinic HIV testing and counseling in rural Lesotho: a cluster-randomized trial. PLoS Med 2014 Dec;11(12):e1001768 [FREE Full text] [doi: 10.1371/journal.pmed.1001768] [Medline: 25513807]
72. The community health toolkit's core framework: an overview. Squarespace. URL: https://static1.squarespace.com/static/5bd25eea65a707ad54c1e8ca/t/5db502b9628b613d418ce7be/1572143837269/Core+Framework+Overview.pdf [accessed 2023-01-27]
73. Bannon L. Reimagining HCI: toward a more human-centered perspective. interactions 2011 Jul;18(4):50-57. [doi: 10.1145/1978822.1978833]
74. Holeman I, Kane D. Human-centered design for global health equity. Inf Technol Dev 2019 Sep 29;26(3):477-505 [FREE Full text] [doi: 10.1080/02681102.2019.1667289] [Medline: 32982007]
75. Timmermans S, Tavory I. Theory construction in qualitative research: from grounded theory to abductive analysis. Sociol Theory 2012 Sep 10;30(3):167-186. [doi: 10.1177/0735275112457914]
76. Lee JS, Pries-Heje J, Baskerville R. Theorizing in design science research. In: Proceedings of the 6th International Conference on Service-Oriented Perspectives in Design Science Research. 2011 Presented at: DESRIST 2011; May 5-6, 2011; Milwaukee, WI. [doi: 10.1007/978-3-642-20633-7_1]
77. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Hum Comput Interact 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]
78. Flanagan JC. The critical incident technique. Psychol Bull 1954 Jul;51(4):327-358. [doi: 10.1037/h0061470] [Medline: 13177800]
79. Gregor S, Chandra Kruse L, Seidel S. Research perspectives: the anatomy of a design principle. J Assoc Inf Syst 2020 Nov 01;21(6). [doi: 10.17705/1jais.00649]
80. Giesbrecht T, Schwabe G, Schenk B. Service encounter thinklets: how to empower service agents to put value co‐creation into practice. Inf Syst J 2016 Mar 10;27(2):171-196. [doi: 10.1111/isj.12099]
81. Schwabe G, Briggs RO, Giesbrecht T. Advancing collaboration engineering: new ThinkLets for dyadic problem solving and an application for mobile advisory services. In: Proceedings of the 49th Hawaii International Conference on System Sciences. 2016 Presented at: HICSS 2016; January 05-08, 2016; Koloa, HI. [doi: 10.1109/hicss.2016.102]
82. Engel GL. The clinical application of the biopsychosocial model. Am J Psychiatry 1980 May 01;137(5):535-544. [doi: 10.1176/ajp.137.5.535] [Medline: 7369396]

Abbreviations

5RSM: algorithmic recommending, restricting, requiring, rating, rewarding, sanctioning, and monitoring
CHT: community health toolkit
CHW: community health worker
ComBaCaL: Community Based Chronic Disease Care Lesotho
GSS: group support system
mHealth: mobile health
NCD: noncommunicable disease
SUS: System Usability Scale

Edited by Y Gong; submitted 20.02.24; peer-reviewed by A Hassan, A Hassan, M Hirschi, A Hevner; comments to author 24.12.24; revised version received 18.02.25; accepted 17.03.25; published 23.04.25.

Please cite as:
Staehelin D, Schmid D, Gerber F, Dolata M, Schwabe G
Empowering Community Health Workers With Scripted Medicine: Design Science Research Study
JMIR Hum Factors 2025;12:e57545
URL: https://humanfactors.jmir.org/2025/1/e57545
doi:10.2196/57545
PMID:

XSL•FO
RenderX

©Dario Staehelin, Damaris Schmid, Felix Gerber, Mateusz Dolata, Gerhard Schwabe. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 23.04.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Knowledge Sharing Maturity Model for Medical Imaging Departments: Development Study

Maryam Almashmoum[1,2], BA, MSc; James Cunningham[1], MSc, PhD; John Ainsworth[1], MSc, PhD

[1]Division of Informatics Imaging and Data Sciences, School of Health Sciences Faculty of Biology, Medicine, and Health, University of Manchester, Manchester, United Kingdom

[2]Nuclear Medicine Department, Faisal Sultan Bin Eissa, Kuwait Cancer Control Center, Kuwait city, Kuwait

Corresponding Author:
Maryam Almashmoum, BA, MSc
Division of Informatics Imaging and Data Sciences
School of Health Sciences Faculty of Biology, Medicine, and Health
University of Manchester
Manchester Academic Health Science Centre
Vaughan House, Portsmouth St
Manchester, M13 9GB
United Kingdom
Phone: 44 07949062010
Email: Maryam_jassem@yahoo.com

Abstract

Background: Knowledge sharing in medical imaging departments is driven by the need to improve health care services, develop health care professionals' skills, and reduce repetitive mistakes. It is considered an important step in the implementation of knowledge management solutions. By following a maturity model of knowledge sharing, knowledge-sharing practices can be improved.

Objective: This study aimed to develop a maturity model for knowledge sharing in the medical imaging department to help managers to assess the level of maturity of knowledge-sharing practices. In modern health care institutions, improvements in health care professionals' skills and health care services are often driven through practicing knowledge-sharing behaviors. Managers can follow the indicators of maturity model of knowledge sharing and its measurements to identify the current level and move to the next level.

Methods: This study was conducted in 4 stages: an *overview stage* that highlighted the factors that affect knowledge-sharing practices in medical imaging departments; an *analysis factor stage* that was designed to assess the factors that affect knowledge sharing using a concurrent mixed methods approach (questionnaires and semistructured interviews) in 2 medical imaging departments; *a structuring maturity model knowledge sharing stage*, where a maturity model of knowledge sharing was developed based on the findings of the first and second stages; and finally, an *assessment of reliability and validity stage*, where a modified Delphi method was used to obtain consensus among experts on model components to be ready for implantation.

Results: The model presented in this study includes 17 indicators, divided into 11 components. Those components were derived from the findings of the questionnaires and semistructured interviews that were applied in the medical imaging departments. It consisted of 5 maturity levels: initial, aware, defined, managed, and optimized. In each level, measurements were included to help managers assess the current level by answering the questions. On the basis of reliability, the experts reached a consensus agreement on the model's components in 2 rounds with SD <1.

Conclusions: This maturity model of knowledge sharing in medical imaging departments allows managers and policy makers to measure the maturity level of knowledge sharing in those departments. Although the model has been applied to medical imaging departments, it could easily be modified for application in other institutions.

(JMIR Hum Factors 2025;12:e54484) doi:10.2196/54484

KEYWORDS

knowledge management; knowledge sharing; medical imaging departments; cancer centers; the Christie; Kuwait Cancer Control Center; KCCC; maturity model; factors; indicators; measurement

XSL•FO
RenderX

Introduction

Background

Knowledge is an essential asset in achieving successful institutional practices, and sharing knowledge gives a sustainable competitive advantage to modern institutions [1,2]. Accessing knowledge is the main step in problem-solving and decision-making [3]. Knowledge is a mixture of experiences, thoughts, ideas, values, and information. In the context of institutions, knowledge can be shared and transformed among employees within institutions to create new experiences and information that did not exist before [4]. It consists of 2 main categories: tacit and explicit knowledge [5]. Tacit knowledge is any knowledge that exists in human minds, such as thoughts and ideas. This kind of knowledge is difficult to document and share. In contrast, explicit knowledge is the knowledge that exists in policies, letters, documents, and manuals. This kind of knowledge is easy to document and share with others using several mechanisms [6]. Both kinds of knowledge are not separate from each other. The dynamic process between tacit and explicit knowledge helps share knowledge efficiently [5,7].

Health care institutions, for example, hospitals and specialized centers, must apply a knowledge management system to build a proper and effective network among all health care providers [4]. The reasons for implementing a knowledge management system are mainly the complexity and huge number of knowledge-based resources that need to be managed [8]. Implementing knowledge management depends on understanding its processes (knowledge creation, knowledge capture, knowledge sharing, and knowledge applications). Moreover, identifying a clear framework to follow is vital to targeting any weaknesses in each step and creating a good environment for health care professionals. As a result, health care services and patient outcomes can be improved [8].

Knowledge sharing is considered an important step in implementing knowledge management. Without a knowledge-sharing process, knowledge management will be difficult to implement because it relies on the dynamic process of tacit and explicit knowledge sharing among employees. Therefore, creating a culture of knowledge sharing among them, and understanding the resources and factors that affect knowledge sharing, are important for any knowledge management system. Looking at health care services, knowledge sharing is an essential step for successful knowledge management within institutions, as it can lead to improved health care settings performance, save them time, reduce costs, and increase health education levels [9]. It plays an important role in providing great accountability and establishing good practices in health planning and policy making. Abzari et al [10] stated that knowledge sharing is considered an essential factor in successful institutional performance. It refers to the act of sharing both tacit and explicit knowledge, such as thoughts, ideas, and experiences from one person, group, or institution to another to formulate new knowledge [11]. Knowledge sharing is one of the challenges for health care institutions due to the variety of resources and knowledge. Health care institutions are complex environments given the variety of specialties in each department and who operates institutional resources in each department [12].

Medical imaging departments are essential in any health care institution due to the importance of performing important procedures for the pretreatment plan and interpreting results. On the basis of the observation at the cancer center, there are several problems facing knowledge-sharing practices in the medical imaging department, such as a lack of awareness of the importance of knowledge sharing, difficulty in sharing knowledge among health care professionals, and a lack of implementation of information and communication technology (ICT) [13]. Creating a communication environment is important to control the amount of knowledge shared among health care professionals in their institutions [14]. In addition, dealing with health care institutional issues is directly related to human capital [15]. Sharing knowledge among health care professionals is critical to apply that knowledge in their daily work and allow them to create new knowledge that is needed for developing the institutional process. Both tacit and explicit knowledge are important to enhance the knowledge-sharing process among health care professionals, because they are directly related to the health care professional's experience and skills [16]. Therefore, identifying a clear maturity model (MM) for knowledge sharing in medical imaging is important to help policy makers and managers use it to enhance knowledge-sharing practices among health care professionals.

An Overview of the MMs

Maturity in this context refers to the degree to which technology, institutional process, or frameworks evolve over time [17,18]. In institutions, an MM can be methodically used to define operations and identify stages, which can lead to policy plans [19]. The concept of the MM was developed in the early 1970s [20]. It is increasingly applied in information systems [21] and is established in many fields, such as knowledge management, information management, and software performance and management [22-25]. The purpose of MMs is to provide a clear model based on an institution's capabilities in a certain managerial area using a set of criteria and related evaluation methodologies [26]. It can be a powerful tool that helps identify strengths and weaknesses in a specific area [27]. Bititci et al [28] indicated that MMs have a positive impact on improving institutional performance and respond to many challenges by describing each step and stage properly.

Health care institutions face challenges in achieving best practices in knowledge management system implementation [29]. Maturity in knowledge management is defined as the degree to which knowledge assets are effectively managed and controlled within institutions [30]. There have been several MMs for evaluating and describing these areas of management that have been proposed in different studies [17-19,31-33]. Some other applications of MMs have been applied to health care intuition information systems [34].

On the basis of the definition of maturity in knowledge management, there is a strong link with the knowledge-sharing definition, which is sharing experience and knowledge with each other within an institution. This knowledge, either tacit or explicit, needs to be managed to be able to be shared. In

XSL•FO
RenderX

addition, to share any kind of knowledge, it should be managed by human experience; this kind of management is required for implementing the knowledge management process because it is *dynamic and active* [35]. Therefore, there is a strong relationship between the 2 definitions, as both deal with the assets that need to be managed to obtain the objective. To share the knowledge, they must understand the types of knowledge and the factors that control the process of knowledge sharing, all of which need to be managed and controlled within institutions. On the basis of this context, to improve the knowledge management process, we need to improve the process of sharing the knowledge.

There is a continuing need to structure and develop a new knowledge sharing MM (KSMM) for enhancing knowledge-sharing practices among employees in the medical imaging departments [36]. Structuring a KSMM helps decision makers to achieve institutional tasks and improve the quality of patient outcomes [37]. There are several models of knowledge sharing [38-41]. Despite the variety of these models, few of them focus on knowledge sharing in health care in general and in the medical imaging department in particular [42]. There are several factors that affect knowledge sharing among health care professionals, including the MM of multidisciplinary team (MDT) meetings, community of practice, ICT, social media networks, picture archiving and communication system (PACS), telemedicine, and digital library [43-49]. Creating a MM for knowledge sharing that consists of all factors that affect their behaviors is vital to implementing a knowledge-sharing environment among health care professionals. The KSMM should be an umbrella that covers all the factors.

To date, there is no MM to assess knowledge-sharing practices among health care professionals in general hospitals, or in medical imaging departments specifically, covering all the factors that affect knowledge sharing in medical imaging departments. Ipe [50] represented the model for the relation of the knowledge sharing between individuals, factors, and the relation between them. Therefore, developing an MM to evaluate the level of maturity of knowledge-sharing behaviors is a challenge due to the interactions of health care professionals' behaviors with each other on the one hand and with ICT on the other. In addition, this model is dependent on human beings, cultural environments, and technological facilitators [50].

Aim and Research Questions

The aims of this paper were to present the steps of the development of a KSMM for understanding how health care professionals share their knowledge; define the proper stages of achieving the best knowledge practices among health care professionals in medical imaging departments; and use the Delphi method to examine the reliability and validity among experts on that model [51].

Therefore, the key research questions were as follows: (1) what are the main components that structure the KSMM? (2) what are the stages of MM that managers in medical imaging departments need to follow to evaluate and enhance knowledge-sharing practices among health care professionals? (3) what are the indicators that control the KSMM, and how can managers measure them? and (4) how to evaluate the proposed KSSM regarding validity and reliability?

Objectives

This paper had various objectives. The first objective is to identify the main stages and components that relate to the KSMM. The second objective is to explain the main characteristics of each stage involved in the MM in depth. The third objective is to create a proposal for an MM that includes the knowledge-sharing influencing indicators that affect knowledge-sharing practices among health care professionals in medical imaging departments, and to present the measurements for each indicator that form the components of the KSMM. Finally, the last objective is to generate consensus among the experts in knowledge management and senior managers of the medical imaging departments in the United Kingdom and Kuwait on the components of the MM along with its indicators and measurements to be implemented in all departments.

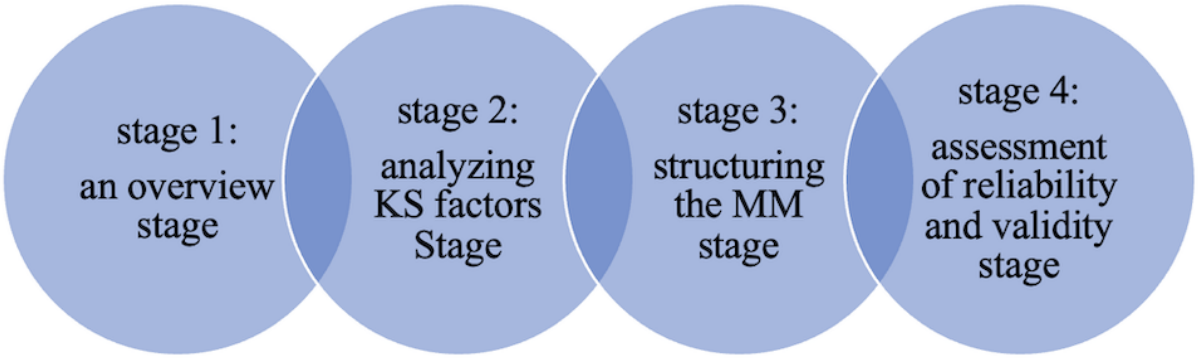
Methods

Overview

This paper illustrates the development of an MM for knowledge sharing in medical imaging departments. There were four stages conducted in this study, as shown in Figure 1: (1) an *overview stage*, consisting of a review of previous literature that identified factors that affect knowledge-sharing practices in health care institutions in general and medical imaging departments specifically; (2) the *analysis of factors* stage, which consists of applying several methodologies to examine those factors; (3) the *structuring the MM* stage; and (4) the *assessment of reliability and validity* stage. Hence, the initial MM for knowledge sharing in the medical imaging department was developed.

XSL•FO
RenderX

Figure 1. The 4 stages conducted in this study. KS: knowledge sharing; MM: maturity model.



Ethical Considerations

This study followed the ethical guidelines outlined by the Christie hospital and the Kuwait Cancer Control Center (KCCC). Stage 1 was a systematic review, which did not require ethics approval. Stage 2 assessed potential factors with the involvement of health care professionals, and stage 3 analyzed the data from previous stages. The Christie hospital, following the University of Manchester regulations, did not require ethics approval for these two stages because no sensitive or personal question were asked; the KCCC, following regulation from the Ministry of Health, Kuwait, provided the approval (3797). Stage 4 used the Delphi method, which also does not require ethics approval because all information based on the respondents will be kept anonymous.

All respondents were asked to fill out the consent form electronically, as a mandatory request on the Qualtrics survey tool, before participating in the study. A brief description of the study was provided at the start of the electronic survey, followed by the consent form. Additionally, participants had a right to withdraw at any point of the survey based on the consent form provided. All the data are stored in the University of Manchester Qualtrics platform and kept anonymous, only accessible by the authors for the purposes of this study.

Stage 1: An Overview

An overview stage illustrated the factors that affect knowledge sharing in medical imaging departments at the general hospital and cancer centers specifically. There were several studies that identified knowledge-sharing factors, which were divided into 3 categories: individual factors, administrative factors, and technological factors [42,52-64]. Any factors that have a positive impact on enhancing knowledge sharing are called facilitators. In contrast, any factors that hinder knowledge-sharing practices are called barriers.

Stage 2: Analysis of Factors

In this stage, to examine and analyze those factors, mixed methods were used in this study. The mixed methods consist of questionnaires and semistructured interviews [65]. A concurrent, cross-sectional, triangulation, mixed methods design was used in this study in 2 cancer centers (the Christie hospital and KCCC). The questionnaire was structured based on the previous studies, and the semistructured interview questions were formed based on the factors in the previous stage to examine those factors by distributing questionnaires and conducting semistructured interview among health care professionals who are working in the medical imaging department at the 2 cancer centers [66-68]. We used self-selection sampling for the questionnaires and snowball sampling for the semistructured sampling techniques. The questionnaire was divided into 3 parts: the demographic section, knowledge-sharing practices, and factors that affect knowledge-sharing practices in the medical imaging departments. Multimedia Appendix 1 shows the survey questions. The questionnaires were distributed electronically using the Qualtrics (Qualtrics International Inc) survey tool. For the health care professionals working at the KCCC, a WhatsApp group (Meta Platforms) was used for distributing questionnaires, whereas the internal page was used as an official web-based tool for those questionnaires. The semistructured interviews were conducted via Microsoft Teams for the health care professionals in the medical imaging department of 2 cancer centers: the Christie hospital and KCCC.

The collection of the data was conducted between February 2023 and July 2023. After collecting the data, an analysis for the quantitative data were performed using Qualtrics Experience Management software package. The duration of the semistructured interviews was approximately 25 to 45 minutes. The thematic analysis was conducted to analyze the transcripts. All the codes and themes were analyzed using the NVivo software (Lumivero).

Stage 3: Structuring the MM

In this stage, an MM for knowledge sharing was developed. On the basis of the identification and evaluation of factors that affect knowledge-sharing practices in medical imaging departments in previous rounds, a clear vision was developed to build a KSMM in a proper manner, helping the managers and policy makers to follow these steps and understand how these factors are related to each other in different stages. Furthermore, it helped them either to implement a knowledge-sharing environment or to identify the weak points of knowledge-sharing practices.

Stage 4: Assessment of Reliability and Validity

In this stage, after the KSMM was developed, it was further assessed for reliability to validate the model. Without validation, the KSSM lacked reliability in its components. To apply this

XSL•FO
RenderX

model, experts and senior managers should agree on its use. Therefore, this stage aimed to generate consensus among the experts in knowledge management and senior managers of the medical imaging departments in the United Kingdom and Kuwait on the components of the MM along with its indicators and measurements to be implemented in all departments. Consensus on the model contributes to updating and correcting missing information before reaching out to whom it concerns.

Our target was to recruit 7 to 10 participants (experts in knowledge management and senior managers in the medical imaging department). This is the minimum number of participants required to conduct this kind of study [69,70]. In addition, Cyphert and Gant [71] illustrated that the best results might be obtained by a homogenous group of 10 to 15 experts. There are no rigid rules in the Delphi method regarding selecting the size of the expert's panel. However, the number of experts depends on the purpose of the study [72]. Delphi studies often use a sample size adjusted to the topic. By selecting experts with comparable broad understanding of the subject, a smaller sample size can be used [73]. We used the following four ways to choose participants: (1) we searched for the experts of knowledge management through the official university websites to find the senior lecturer in that field, especially at Kuwait University and the University of Manchester; (2) we searched through LinkedIn (Microsoft Inc) to find stakeholders for the National Health Service who have a role in knowledge management practices; (3) throughout the previous literature in knowledge management, there were several potential articles in this field with the authors name and contacts details, so we contacted them through this information to request their participation in this study; and (4) we asked the senior managers in the medical imaging department at the Christie hospital and KCCC to participate and help us to add their perspectives based on their professional experience. All participants remained anonymous during the research.

The modified Delphi method survey questions were generated based on the indicators and the measurements that structured the KSMM. The experts who participated in this study were informed of the aim, methods, and main components that structured the model to give them an overview of the study. The consensus survey questionnaire consisted the consent form (5 statements) and 3 sections: demographic questions (7 multiple-choice questions), a Likert scale with 5 points (disagree to agree) for the 17 indicators, and open-ended questions for it measurements to allow the experts to add their comments for updating the model. Study data were collected electronically using the Quartics survey tool and distributed via email. The use of email has a potential benefit in making it easier to maintain the anonymity of the respondents. Delphi studies ensure the anonymity of suggested comments and assessment to prevent group dynamics, strong personalities, or group conformism from influencing personal relationships between participants [74]. Descriptive statistics were used to analyze the Delphi studies, including SD, mean, and the intraclass correlation coefficient (ICC). We computed the mean, ICC, and SD based on the Likert scale of each indicator, where 1 indicates the least agreement and 5 indicates the highest agreement. In Delphi studies, IQRs are widely used to identify the level of group censuses [75,76]. However, we used mean, ICC, and SD in this study as they might be enough to identify the level of agreement for a model that was already structured.

Results

Stage 1: An Overview

On the basis of the previous literature, the factors that affect knowledge sharing in medical imaging departments are the same whether in a general hospital or a cancer center. However, they are different in terminology because the nature of the cancer center is mainly concerned with treating cancer cases, which require more than one specialty to end up with an appropriate protocol to treat specific cases [77]. Therefore, documenting the factors that affect knowledge sharing in medical imaging departments from the previous literature is important to test these factors in the next stage [77]. These factors are shown in Table 1 [77].

XSL•FO
RenderX

Table 1. Factors that affect knowledge sharing in the medical imaging departments in general hospitals versus cancer centers.

Types of factors	Facilitators terminologies in medical imaging departments in general hospitals	Facilitators terminologies in medical imaging departments in cancer centers
Individual factors	• Trust • Positive attitudes • Awareness • Experience • Intrinsic motivation • Personality • Self-esteem • Self-efficacy	• Trust • Positive attitudes • Awareness • Experience • Intrinsic motivation • Personality • Self-esteem • Self-efficacy
Departmental factors	• CoP[a] and interprofessional collaboration • Leadership • Culture • Teamwork • Extrinsic motivation • Learning and training • Physician rounds • Departmental arrangements	• MDT[b] and community of oncologists • Leadership • Culture • Teamwork • Extrinsic motivation • Learning and training • Physician rounds • Departmental arrangements
Technological factors	• ICT[c] (PACS[d], social media, intranet, extranet, telemedicine, and teleradiology) • Network • Digital library	• ICT (PACS, social media, intranet, extranet, telemedicine, and teleradiology) • Network • Digital library

[a]CoP: communities of practice.

[b]MDT: multidisciplinary team.

[c]ICT: information and communication technology.

[d]PACS: picture archiving and communication system.

Stage 2: Analysis of Factors

A total of 85 responses were received (56 from the KCCC and 29 from the Christie hospital), and the response rate based on Quadrics was 100%. Qualtrics Experience Management software was used to analyze the quantitative data. The factors were divided into 3 categories based on the systematic review: individual factors, departmental factors, and technological factors [77]. The demographics characteristics of the 2 centers are shown in Table 2. The results of the questionnaires for the 2 cancer centers are provided in Multimedia Appendix 2.

The semistructured interviews were conducted online using Microsoft Teams with 13 health care professionals who were working in the medical imaging departments at the 2 cancer centers: 11 (85%) were from the KCCC and only 2 (15%) were from the Christie hospital. Thematic analysis was used to analyze the qualitative data. Codes were organized using NVivo software. This was done to validate the quantitative methods, understand their views, understand how knowledge-sharing practices were going on in these centers, and assess if there was a clear policy to adopt knowledge-sharing practices. Multimedia Appendix 3 shows the interview questions and consent form. All these methods were applied according to the ethical considerations at both canters.

XSL•FO
RenderX

Table 2. Demographic characteristics.

Characteristic	KCCC[a] (n=56), n (%)	The Christie (n=29), n (%)
Sex		
Male	26 (46)	7 (24)
Female	29 (52)	20 (69)
Prefer not to say	0 (0)	2 (7)
Age group (y)		
<20	0 (0)	1 (3)
20-30	3 (5)	8 (28)
30-40	19 (34)	7 (24)
40-50	25 (45)	8 (28)
50-60	7 (12)	3 (10)
>60	2 (4)	2 (7)
Educational level		
Diploma	7 (12)	4 (14)
First degree (Bachelor)	26 (47)	12 (41)
Master's degree	11 (20)	8 (27)
Doctorate degree	9 (16)	0 (0)
Other	3 (5)	5 (17)
Work experience (y)		
<10	18 (32)	22 (76)
10-20	29 (51)	6 (21)
20-30	7 (12)	1 (3)
>30	2 (4)	0 (0)

[a]KCCC: Kuwait Cancer Control Center.

The results revealed that there were 11 components derived from 5 categories, as shown in Table 3. These categories started with awareness, which is considered the main step to adopting knowledge-sharing practices in the department. Awareness of the importance of knowledge sharing in developing health care professionals' skills and improving health care services is important to increase their willingness to share their knowledge. The results of the quantitative method showed that health care professionals in both cancer centers have a high level of awareness of the importance of knowledge sharing. The second category was related to the types of knowledge sharing. It is divided into 2 categories: tacit and explicit knowledge. Tacit knowledge is the most dominant type of knowledge in the medical imaging department because several meetings and specialized meetings (MDT meetings, community of practices, and community of oncologists) occur in these departments. In contrast, explicit knowledge exists in the endorsed documents, protocols, policies, and procedures manuals. Therefore, understanding both types of knowledge and organizing them is important to increase knowledge-sharing practices.

The third, fourth, and fifth categories were related to the factors that affect knowledge-sharing practices: individual, departmental, and technological factors. The mean scores of the factors fell between somewhat agree and strongly agree, which reveals that health care professionals believe in the importance of these factors in enhancing knowledge-sharing behaviors. The individual factors consisted of 2 components: communication among health professionals, and personality and positive attitudes. Building trusting relationships among health care professionals is important to allow them to share their experiences. In addition, most of the respondents agreed that intrinsic motivation has a positive impact on sharing knowledge by increasing their self-efficacy and self-esteem by giving them opportunities to practice their abilities and experiences. There is a strong relationship between personality, positive attitudes, and increased knowledge-sharing practices. The respondents indicated that knowledge-sharing behaviors occur among individuals. Each indivdual has a specific personality and attitude toward knowledge-sharing practices. Some of them like to share their knowledge in large groups, while others prefer when they share it in small groups.

XSL•FO
RenderX

Table 3. Categories, components, and indicators that build maturity model of knowledge sharing for the medical imaging departments.

Categories and components	Indicators
Awareness	
• Awareness and willingness	• (1/17) Awareness and willingness toward KS[a]
Types of knowledge sharing: tacit knowledge and explicit knowledge	
• Understanding and organizing a knowledge-sharing repository	• (2/17) Structured and collected both types of knowledge
Individual factors: trust, positive attitudes, experience, intrinsic motivation, personality, self-esteem, and self-efficacy	
• Communication among HCPs[b]	• (3/17) Building trust among HCPs and sharing their experience • (4/17) Increased intrinsic motivation (self-efficacy, and self-esteem)
• Personality and positive attitudes	• (5/17) Personality and communication among HCPs
Departmental factors: MDT[c] and community of oncologists, leadership; culture, teamwork, extrinsic motivation, learning and training, physician rounds, and departmental arrangements	
• Leadership and culture	• (6/17) Structured leadership and creating culture • (7/17) Handover policy
• Achieving departmental tasks	• (8/17) Creating teamwork
• Continuous education and develop HCP skills	• (9/17) Organizing (learning lectures, workshops, training sessions, physician rounds, and participation in conferences)
• Decision-making	• (10/17) Regular meeting • (11/17) MDT and CoP[d] decision-making
• Infrastructure and workforce	• (12/17) Meeting room and office layout • (13/17) Enhanced extrinsic motivation • (14/17) Organized work process
Technological factors: ICT[e] (PACS[f], social media, intranet, extranet, telemedicine, and teleradiology), network, and digital library	
• Stored and shared patient data electronically	• (15/17) Strong network • (16/17) Implementation information, communication technology, and maintenance
• Access to the electronical databases	• (17/17) Digital library

[a]KS: knowledge sharing.

[b]HCP: health care professional.

[c]MDT: multidisciplinary team.

[d]CoP: communities of practice.

[e]ICT: information and communication technology.

[f]PACS: picture archiving and communication system.

The departments have a big responsibility to enhance knowledge sharing based on respondents' thoughts. Departmental factors consisted of 5 components: leadership and culture, achieving departmental task, continuous education, decision-making, and infrastructure and workforce. The participants agreed that a leader has a responsibility to create a culture of communication that allows health care professionals to practice knowledge-sharing behaviors. Developing health care professionals' skills require setting a clear plan for practicing continuing education activities, such as attending conferences, lectures, training sessions, and workshops. Offering them space and time is vital for allowing them to practice knowledge-sharing activities by organizing tasks among them and giving them empty spaces. To achieve departmental tasks, working within a team is crucial to increase the number of tasks achieved. Decision-making regarding patient treatment is one of the important tasks in the department, which is usually undertaken in highly specialized meetings, such as MDTs and communities of practices.

Technological factors consisted of several technological modalities, such as PACS, social media, the intranet, the extranet, telemedicine, and teleradiology, that require high-speed networking. These factors consisted of 2 components: storing and sharing electronic data electronically, and access to electronic databases. The respondents illustrated that using

XSL•FO
RenderX

technology requires skills for using it efficiently. Moreover, accessing the databases is vital to expanding health care professionals' knowledge about up-to-date treatment plans that are used to treat patients with cancer.

Stage 3: Structuring the MM

The development of the MM was derived from a systematic review that identified the factors affecting knowledge-sharing behaviors in the medical imaging department, as well as concurrent mixed methods that evaluated those factors in 2 cancer centers. The nonprobability sampling technique was used for mixed methods: the self-selection sampling technique for the quantitative method and the snowball sampling technique for the qualitative method. The results indicated that the KSMM consisted of 17 indicators divided into 11 components. These components represented the following 5 main categories: awareness, knowledge-sharing repository, individual factors, departmental factors, and technological factors. Table 3 shows the details of the MM indicators for knowledge sharing. Knowledge sharing is a dynamic process among health care professionals to manage the institutional process and interaction with ICT infrastructures. As a consequence, these components were incorporated into the 5 maturity levels. The levels were adopted from the work of Pee and Kankanhalli [19] that indicated that there are 5 levels of maturity in knowledge management (initial, aware, defined, managed, and optimized) based on people, process, and technology. The KSMM for the medical imaging department is presented in Multimedia Appendix 4. For each indicator, there are measurement questions that help managers and policy makers assess each indicator.

Stage 4: Assessment of Reliability and Validity

A total of 9 different participants were recruited in this study: 9 in the first round and 9 in the second part. They were from 2 countries: the United Kingdom and Kuwait. Table 4 shows the demographic characteristics of the participants. Most of respondents were female (n=7, 78%), and they were mainly from Kuwait (n=7, 78%). Most (n=6, 67%) of the respondents had a doctorate degree and the rest (n=3, 34%) had first degree.

Table 4. Demographic characteristics of the participants in the Delphi method (N=9).

Demographic characteristics	Participants, n (%)
Gender	
Female	7 (78)
Male	2 (22)
Age group (y)	
40-50	7 (78)
50-60	2 (22)
Highest education level	
First degree	3 (33)
Doctorate degree	6 (67)
Background	
Senior lecturer in knowledge managements	3 (33)
Senior manager in medical imaging departments	6 (67)
Experience (y)	
<10	1 (11)
10-20	4 (44)
20-30	4 (44)
Country	
United Kingdom	2 (22)
Kuwait	7 (78)

In the first round, most of the indicators and their measurements that formed the KSMM reached consensus with a narrow SD (<1); Multimedia Appendix 5 shows the level of consensus in both rounds. In contrast, indicators 1 (SD 1.25) and 17 (SD 1.15), which were about awareness and the digital libraries, respectively, did not reach to the consensus level with a large SD (>1). Therefore, the SD was more obvious and sensitive for this study to assess the level of consensus, compared to other measurements [74,78,79]. Awareness is the first component that forms the KSMM because it has a crucial role in enhancing knowledge-sharing practices among health care professionals. On the basis of their comments, there is a difference between the willingness to share knowledge and awareness. The willingness to share their knowledge comes from an awareness of the importance of knowledge sharing and believing in the value and benefits of their shared knowledge [80]. To increase awareness about knowledge sharing, the managers should set a clear policy to identify the benefits of sharing practices in

XSL•FO
RenderX

their daily work to increase patient outcomes and reduce medical errors. In contrast, their willingness to share knowledge is directly related to their personality and how their intrinsic and extrinsic motivations encourage them to share their knowledge among other peers. On the basis of indicator 17, most (4/9, 44%) of the respondents indicated that the digital resources are more comprehensive among those indicators. The digital resources include digital databases that contain libraries that play an important role in enhancing their knowledge in any field and therefore increasing knowledge-sharing practices. Regarding other measurements, they showed less sensitivity compared to the SD because they did not indicate any disagreements compared to the SD. Therefore, the SD is a more sensitive measurement compared to others [79].

Besides the analysis of the Likert scale, we received an important comment that helped to update the KSMM based on their experiences. Multimedia Appendix 6 shows the comments of the participants and our response to update the model.

In round 2, all the components of KSMM met the consensus SD level (<1.0). This indicates that all the respondents agreed with the updated version of the model that kept all their comments and considerations. Therefore, there was no follow-up round for this study because consensus on all the indicators along with their measurements of the KSMM was met. The final KSMM in the medical imaging department is presented in Multimedia Appendix 4.

Discussion

Principal Findings

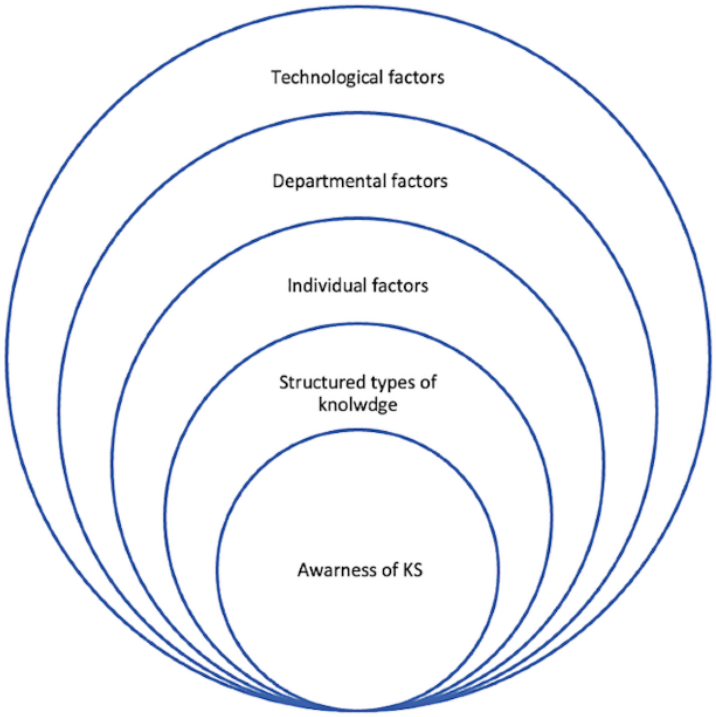
Despite the considerable attention given to knowledge management in health care institutions and the significance of the MM in managing health care resources, knowledge sharing in hospitals has been less focused on developing the MM as a tool for assessing knowledge-sharing practices or a road map to adopting knowledge-sharing behaviors. This is the essence of this work. This study aimed to develop a KSMM in the medical imaging department by assessing the factors that affect knowledge-sharing practices. Several MMs have been developed for social media, health systems, digital libraries, ICT, PACS, and telemedicine [43,45,47-49]. These models were used to assess each factor that affects knowledge-sharing practices in benchmarking efforts and to develop progressive strategies that might improve its activities. In addition, Liu et al [46] shed light on a MM for multidisciplinary cancer teams. Their model consisted of 17 indicators that were used to measure health care professionals' performance and monitor the quality of performance at cancers centers over time.

Figure 2 highlights the 5 components that affect the KSMM in the medical imaging department and classifies each component in terms of its influence on knowledge-sharing behaviors. Awareness of the importance of knowledge management is one of the core components that contribute to the adoption of knowledge-sharing practices among health care professionals. Most of the respondents showed a high level of awareness of the importance of knowledge sharing in developing their skills, increasing health care services, and reducing medical errors. Therefore, results showed that more than half of respondents in both cancer centers (the Christie: 17/29, 59%; KCCC: 32/56, 57%) participated daily in knowledge-sharing activities that were available in their department. In addition, one of the articles showed that without awareness of the importance of knowledge sharing, there were no knowledge-sharing practices in the medical imaging department [13]. The next step was structuring the types of knowledge (tacit and explicit). Understanding the types of knowledge available in the department and how to capture, document, and share it is vital to accelerating knowledge-sharing practices in the department. Tacit knowledge appears in the medical imaging department as a dominant type among health care professionals that is considered a tool for sharing knowledge in lectures, conferencing, and meetings. In contrast, explicit knowledge takes several forms, such as documents, policies, procedures, and manuals. This form allows workers to reach it anytime in an easy way [81]. From the third to the fifth step, factors that affected knowledge sharing were divided into 3 categories: individual factors, followed by departmental factors, and finally technological factors [13,42,52-55,57,59-62].

The structure of the KSMM relies on the deep interpretation of several methods. The KSMM was built based on the interpretation of the systematic review and the concurrent mixed methods. That model is lacking in validation. Applying the modified Delphi method helps to validate this model by involving experts. In the first round, the results revealed that indicators 1 and 17 had a large SD (>1).

XSL•FO
RenderX

Figure 2. Schematic description of the 5 components that affect the knowledge sharing maturity model in medical imaging departments. KS: knowledge sharing.



The first round provided several comments and suggestions to modify the model. There were significant comments that needed to be considered and updated before being shared with the participants in the second round. Some of the respondents (R6 and R7) indicated that tacit knowledge is different from explicit knowledge in optimization, so the updated model indicated the way to optimize tacit knowledge in a significant way. Tacit knowledge is intangible, which is difficult to codify and capture [82]. To optimize and capture tacit knowledge, it needs to be part of an externalization process that focuses on transforming the tacit knowledge into explicit knowledge [83]. There are several methods of externalization, for example, recording the meeting and writing a summary note to make it easy to share it with health care professionals. On the basis of indicator 3, the respondents (R1, R6, and R7) illustrated that there are 2 types of trust that we should be aware of. Trust in their abilities and other expertise helps them perform more tasks in the working area, and benevolent trust helps to reduce conflicts among them and, therefore, increases knowledge sharing. The trust itself among group members might be increased throughout social activities more than task communication [84]. In addition, Chen and Hung [85] illustrated that both types of trust help increase the desire to share their knowledge, resulting in enhancing the level of positive outcomes in shared groups. Therefore, leaders and managers should be aware of social activities that help increase their trust, which is defined as a level of confidence among members. Regarding individual facilitators, a respondent (R7) illustrated that there is a difference between self-efficacy and self-esteem. Chan and Hung [85] have defined knowledge-sharing self-efficacy as an individual's confidence in their knowledge, abilities, and experience to achieve several tasks that might be helpful for other peers. A high level of self-efficacy is directly related to a higher level of performance on tasks [85]. In contrast, self-esteem is defined to play a key role in emotional and behavioral adjustment, as well as academic achievements [86]. Leaders and senior managers should encourage their employees to share their knowledge, helping them to understand the importance of sharing knowledge among them and how it has direct positive outcomes—by understanding their abilities and increasing their awareness of knowledge sharing, and not by forcing them. Giving them positive feedback on their abilities directly increases their self-efficacy and self-esteem [80,87].

Regarding applying knowledge-sharing practices, the respondents (R1, R5, R6, and R7) indicated that the senior managers and leaders should be aware of the knowledge-sharing models. There are several models for knowledge sharing, and each of them has a certain characteristic [36,88-90]. In addition, they must be able to create a knowledge-sharing model based on the resources that are available in the department. On the basis of the indicator 7, a respondent (R5) indicated that the handover policy should be applied and used not only when health care professionals decide to leave the department. They must share their knowledge with other peers in routine work to allow them to cover their work when they are absent. Such behavior helps to make knowledge-sharing practices more active in daily work. On the basis of teamwork, a respondent (R7) illustrated that to maximize the level of knowledge sharing, leaders and managers should be aware of the diversity of the teamwork in the workplace and the diversity of the member professionals in one teamwork [91]. The diversity is directly related to the individual differences in their knowledge and the abilities that make them unique from other peers [91]. The diversity of the member professionals in one teamwork helps to increase their performance by bringing new knowledge and experiences. However, diversity in one teamwork has several

XSL•FO
RenderX

challenges, such as increasing conflicts and a lack of trust. In contrast, the diversity of teamwork who are performing the task has a positive impact on sharing knowledge.

On the basis of indicator 10, which illustrates meetings, the respondents (R1, R5, R6, and R7) indicated that there are formal and informal meetings in the department, and informal meetings have a significant role in increasing knowledge-sharing behaviors. Hutchinson and Qunitas [92] have defined informal meetings as meetings that indicate knowledge-sharing practices without focusing on the task and the target that is directed by the organization. However, the leader should be aware of the difference between formal and informal meetings and strike a balance between them to enhance knowledge-sharing practices.

Indicator 15 illustrated the strong network. A strong network plays a vital role in enhancing knowledge sharing by introducing new technology to store and share data, anytime and anywhere. A respondent (R7) identified that artificial intelligence (AI) is an important technology, playing an important role in increasing efficiency, effectiveness, and productivity, and therefore increasing knowledge sharing [93]. It is defined as a set of ICTs that mimic human intellect with the goal of enhancing jobs, increasing efficiency, and boosting economic growth [94]. In addition, Arakpogun et al [94] indicated that the benefit of using AI is limited, because new knowledge needs to be introduced as part of the AI learning process, and therefore, more effective collaboration with knowledge-sharing systems is needed.

The Delphi method has limitations, including the use of nonrandomized samples; subjectivity and bias from expert panel composition; and a lack of consensus recommendations for the number of participants, rounds, and reporting criteria [69]. In this study, a significant limitation was the small size of the panel in 2 rounds and only in 2 countries. Despite these limitations, the modified Delphi method is an effective tool for validating several models.

Through this study, the key research questions were answered. The model can be used as a scoring tool to assess knowledge-sharing practices, with each MM scoring 1, 2, 3, 4, and 5, respectively. Moreover, this model can help managers and policy makers find opportunities for improvements and ways to achieve them. In addition, applying the KSMM helps health care institutions increase health services and patient outcomes, reduce errors, and solve the problem in a practical way. The Delphi study identified several comments and suggestions that helped in improving the KSMM. All these comments were kept in consideration for updating the model because they were based on participants' experiences in the workplace as either a senior manager in knowledge managements or a lecturer in the knowledge management field. Multimedia Appendix 4 shows the new KSMM with its updates.

Conclusions

Knowledge sharing is considered a core step in the implementation of knowledge management. Health care institutions have a responsibility to adopt knowledge-sharing practices to manage knowledge that is either tacit or explicit. A medical imaging department is crucial to any health care institution. Therefore, creating the KSMM is important to develop knowledge-sharing practices. The model proposed in this study allows managers to measure the maturity level of knowledge sharing in the medical imaging department. By providing the road map, the KSSM allows policy makers and managers in health care to appraise knowledge-sharing practices and adopt a culture of knowledge sharing to achieve departmental tasks and improvements. In addition, it could help managers assess knowledge-sharing practices in the medical imaging department and find the weak points that have a negative impact on those behaviors. A range of factors were addressed in our previous work, and we then evaluated those factors in 2 medical imaging departments in 2 cancer centers. The factors were further divided into 5 categories. Therefore, the KSMM consisted of 17 indicators that were divided into 11 components and presented in 5 categories. The most important of these indicators is awareness of the importance of knowledge sharing, presumably because it allows health care professionals to develop their skills and perform several tasks. The KSMM went through 2 rounds to ensure consensus among the experts who participated in the Delphi method. Regarding the measurements, all the indicators reached the consensus level with SD <1. Therefore, the model is ready to be applied in the medical imaging departments, and with minor modifications based on their resources, it might be applied in any institution. The model presented might also be used as a reference for improving knowledge-sharing practices. The measurement of each indicator helps the managers assess what level they are at. If the answer to the first level is no, then they must work on their current level until they achieve it and then move on to the next level. It enables them to share their experience in knowledge-sharing practices and serve as a lesson learned for another department with lower KSMM results. In the future, we are looking forward to comparing knowledge-sharing practices before and after implementing the KSMM in the medical imaging department. In general, knowledge-sharing practices are important in health care institutions to avoid repetitive errors, improve health care services, improve collaboration and communication among staff, and therefore encourage them to come up with new ideas together.

Acknowledgments

This study was completed as a part of doctoral studies funded by Ministry of Health, Kuwait, and University of Manchester, United Kingdom.

The authors would like to acknowledge all the professionals who participated in this study.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
The survey questions and consent form.
[DOCX File , 47 KB - humanfactors_v12i1e54484_app1.docx]

Multimedia Appendix 2
Results of the factors that affect knowledge sharing in both cancer centers.
[DOCX File , 30 KB - humanfactors_v12i1e54484_app2.docx]

Multimedia Appendix 3
The interview questions and consent form.
[DOCX File , 18 KB - humanfactors_v12i1e54484_app3.docx]

Multimedia Appendix 4
The details of the maturity model for knowledge sharing in medical imaging departments.
[DOCX File , 40 KB - humanfactors_v12i1e54484_app4.docx]

Multimedia Appendix 5
The level of consensus in both rounds.
[DOCX File , 22 KB - humanfactors_v12i1e54484_app5.docx]

Multimedia Appendix 6
The respondents' comments with our response.
[DOCX File , 22 KB - humanfactors_v12i1e54484_app6.docx]

References

1. Kogut B, Zander U. What firms do? Coordination, identity, and learning. Organ Sci 1996 Oct;7(5):502-518. [doi: 10.1287/orsc.7.5.502]
2. Mc Evoy PJ, Ragab MA, Arisha A. The effectiveness of knowledge management in the public sector. Knowl Manag Res Pract 2018 Nov 07;17(1):39-51. [doi: 10.1080/14778238.2018.1538670]
3. Islam QM. Innovation in primary healthcare in the twenty-first century. J Health Manag 2021 Mar 31;23(1):129-142 [FREE Full text] [doi: 10.1177/0972063421994987]
4. Rego I, Pereira L, Dias Á, Gonçalves R, da Costa RL, Costa A. Knowledge management maturity in healthcare service. Int J Knowl Learn 2023;16(1):17. [doi: 10.1504/ijkl.2023.127335]
5. Nonaka I. A dynamic theory of organizational knowledge creation. Organ Sci 1994 Feb;5(1):14-37. [doi: 10.1287/orsc.5.1.14]
6. Collins HM. Tacit knowledge, trust and the Q of sapphire. Soc Stud Sci 2001 Feb 01;31(1):71-85. [doi: 10.1177/030631201031001004]
7. Alavi M, Leidner DE. Review: knowledge management and knowledge management systems: conceptual foundations and research issues. MIS Q 2001 Mar;25(1):107. [doi: 10.2307/3250961]
8. Singh Sandhu M, Kishore Jain K, Umi Kalthom bte Ahmad I. Knowledge sharing among public sector employees: evidence from Malaysia. Int J Public Sect Manag 2011 Apr 05;24(3):206-226. [doi: 10.1108/09513551111121347]
9. Al-Kurdi OF, El-Haddadeh R, Eldabi T. The role of organisational climate in managing knowledge sharing among academics in higher education. Int J Inf Manage 2020 Feb;50:217-227. [doi: 10.1016/j.ijinfomgt.2019.05.018]
10. Abzari M, Shahin A, Abasaltian A. Studying the impact of personality constructs on employees' knowledge sharing behavior through considering the mediating role of intelligent competencies in project-oriented organizations. Mod Appl Sci 2016 Apr 16;10(6):194. [doi: 10.5539/mas.v10n6p194]
11. Vuori V, Okkonen J. Knowledge sharing motivational factors of using an intra‐organizational social media platform. J Knowl Manag 2012;16(4):592-603. [doi: 10.1108/13673271211246167]
12. Radević I, Dimovski V, Lojpur A, Colnar S. Quality of healthcare services in focus: the role of knowledge transfer, hierarchical organizational structure and trust. Knowl Manag Res Pract 2021 Jun 15;21(3):525-536. [doi: 10.1080/14778238.2021.1932623]
13. Al Mashmoum MJ, Hamade SN. Knowledge sharing among employees of Kuwait cancer control center. Int J Knowl Manag Pract 2019;7(1):15-25 [FREE Full text]
14. Abidi SS. Knowledge management in healthcare: towards 'knowledge-driven' decision-support services. Int J Med Inform 2001 Sep;63(1-2):5-18. [doi: 10.1016/S1386-5056(01)00167-8]

XSL•FO
RenderX

15. Sharma K, Sharma V. Evaluating knowledge management practices in Indian manufacturing and service industry: an overview. Int J Knowl Manag Stud 2018;9(3):222. [doi: 10.1504/ijkms.2018.094212]
16. Dorow P, Medeiros C, Camozzato T, Silva C, Vargas F, Huhn A. Knowledge sharing process between radiologist. Int J Adv Res 2018 Jul 31;6(7):431-436. [doi: 10.21474/IJAR01/7380]
17. Klimko G. Knowledge management and maturity models: building common understanding. In: Proceedings of the 2nd European Conference on Knowledge Management. 2001 Presented at: ECKM '01; November 8-9, 2001; Honolulu, HI p. 264 URL: https://tinyurl.com/47rkab3w
18. Serenko A, Bontis N, Hull E. An application of the knowledge management maturity model: the case of credit unions. Knowl Manag Res Pract 2017 Dec 19;14(3):338-352. [doi: 10.1057/kmrp.2014.37]
19. Pee LG, Kankanhalli A. A model of organisational knowledge management maturity based on people, process, and technology. J Inf Knowl Manag 2011 Nov 21;08(02):79-99. [doi: 10.1142/S0219649209002270]
20. Gibson CF. Managing the four stages of EDP growth. Harv Bus Rev 1974;52(1):76-88 [FREE Full text]
21. Paulk M, Curtis B, Chrissis M, Weber C. Capability maturity model, version 1.1. IEEE Softw 1993 Jul;10(4):18-27. [doi: 10.1109/52.219617]
22. Aho M. What is your PMI? A model for assessing the maturity of performance management in organizations. In: Proceedings of the 2012 Conference on Performance Management: From Strategy to Delivery. 2012 Presented at: PMA '12; July 11-13, 2012; Cambridge, UK p. 1-22 URL: https://tinyurl.com/5m4fksts
23. Becker MC, Salvatore P, Zirpoli F. The impact of virtual simulation tools on problem-solving and new product development organization. Res Policy 2005 Nov;34(9):1305-1321. [doi: 10.1016/j.respol.2005.03.016]
24. van Aken EM, Letens G, Coleman G, Farris J, van Goubergen D. Assessing maturity and effectiveness of enterprise performance measurement systems. Int J Product Perform Manag 2005;54:400-418. [doi: 10.1108/17410400510604557]
25. Wettstein T, Kueng P. A maturity model for performance measurement systems. Manag Inf Syst 2002;26:102 [FREE Full text]
26. Maier AM, Moultrie J, Clarkson PJ. Assessing organizational capabilities: reviewing and guiding the development of maturity grids. IEEE Trans Eng Manage 2012 Feb;59(1):138-159. [doi: 10.1109/tem.2010.2077289]
27. Jääskeläinen A, Roitto JM. Designing a model for profiling organizational performance management. Int J Product Perform Manag 2015;64(1):5-27. [doi: 10.1108/ijppm-01-2014-0001]
28. Bititci US, Garengo P, Ates A, Nudurupati SS. Value of maturity models in performance measurement. Int J Prod Res 2014 Oct 27;53(10):3062-3085. [doi: 10.1080/00207543.2014.970709]
29. Yan Q, Xiang F, Shi X, Zhu Q. Implementation of knowledge management in Chinese hospitals. Curr Med Sci 2018 Apr 30;38(2):372-378. [doi: 10.1007/s11596-018-1888-z] [Medline: 30074199]
30. Šajeva S, Jucevičius R. The model of knowledge management system maturity and its approbation in business companies. Soc Sci 2010(3):57-68.
31. Teah HY, Pee LG, Kankanhalli A. Development and application of a general knowledge management maturity model. In: Proceedings of the 2006 Pacific Asia Conference on Information Systems. 2006 Presented at: PACIS '06; July 6-9, 2006; Kuala Lumpur, Malaysia p. 12 URL: https://www.researchgate.net/publication/221228902_Development_and_Application_of_a_General_Knowledge_Management_Maturity_Model
32. Kulkarni U, St Louis R. Organizational self assessment of knowledge management maturity. In: Proceedings of the 9th Americas Conference on Information Systems. 2023 Presented at: AMCIS '03; August 4-6, 2003; Tampa, FL p. 2003 URL: https://www.researchgate.net/publication/220892849_Organizational_Self_Assessment_of_Knowledge_Management_Maturity
33. Spanellis A, MacBryde J, D rfler V. A dynamic model of knowledge management in innovative technology companies: a case from the energy sector. Eur J Oper Res 2021 Jul;292(2):784-797. [doi: 10.1016/j.ejor.2020.11.003]
34. Carvalho JV, Rocha Á, van de Wetering R, Abreu A. A Maturity model for hospital information systems. J Bus Res 2019 Jan;94:388-399. [doi: 10.1016/j.jbusres.2017.12.012]
35. Serna M E. Knowledge management maturity model in the interpretativist perspective. Int J Inf Manage 2012 Aug;32(4):365-371. [doi: 10.1016/j.ijinfomgt.2011.12.001]
36. Arif M, Al Zubi M, Gupta AD, Egbu C, Walton RO, Islam R. Knowledge sharing maturity model for Jordanian construction sector. Eng Constr Archit Manag 2017 Jan 16;24(1):170-188. [doi: 10.1108/ecam-09-2015-0144]
37. Mettler T, Rohner P. Situational maturity models as instrumental artifacts for organizational design. In: Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology. 2009 Presented at: DESRIST '09; May 7-8, 2009; Philadelphia, PA p. 1-9 URL: https://dl.acm.org/doi/10.1145/1555619.1555649 [doi: 10.1145/1555619.1555649]
38. Lodhi SA. Culture based knowledge sharing model [Dissertation]. National College of Business Administration & Economics, Lahore. 2005 Nov. URL: https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=LODHI+SA.+National+College+of+Business+Administration+%26+Economics+Lahore.+2005.&btnG= [accessed 2024-04-29]
39. Sayed S, Mutasa R, Kaaya E, Mudenda V, Rajiv E, Vuhahula E, et al. Establishing the college of pathologists of east, central and Southern Africa - the regional east central and Southern Africa college of pathology. Afr J Lab Med 2020 Jun 03;9(1):979 [FREE Full text] [doi: 10.4102/ajlm.v9i1.979] [Medline: 32537427]

XSL•FO
RenderX

40. Taylor W, Wright G. Organizational readiness for successful knowledge sharing: challenges for public sector managers. Inf Resour Manag J 2004;17(2):22-37. [doi: 10.4018/irmj.2004040102]
41. Supar N, Ibrahim A, Mohamed Z, Yahya M, Abdul M. Home page. In: editors. Factors affecting knowledge sharing and its effects on performance: a study of three selected higher academic institutions. International Conference on Knowledge Management (ICKM); 2005 Presented at: CHI '19; April 21-26, 2018; San Diego, CA p. a.
42. Lema B. A framework to support knowledge sharing practice among health care professionals at Yekatit 12 hospital medical College. Addis Ababa University. 2017. URL: https://www.academia.edu/65659988/A_Framework_to_Support_Knowledge_Sharing_Practice_among_Health_Care_Professionals_at_Yekatit_12_Hospital_Medical_College [accessed 2024-04-29]
43. Fitterer R, Rohner P. Towards assessing the networkability of health care providers: a maturity model approach. Inf Syst E-Bus Manage 2009 Aug 19;8(3):309-333. [doi: 10.1007/S10257-009-0121-9]
44. Ekuobase GO, Olutayo VA. Study of Information and Communication Technology (ICT) maturity and value: the relationship. Egypt Inform J 2016 Nov;17(3):239-249. [doi: 10.1016/j.eij.2015.12.001]
45. Sheikhshoaei F, Naghshineh N, Alidousti S, Nakhoda M. Design of a digital library maturity model (DLMM). Electron Libr 2018 Sep 19;36(4):607-619. [doi: 10.1108/el-05-2017-0114]
46. Liu Y, Evans L, Kwan T, Callister J, Poon S, Byth K, et al. Developing a maturity model for cancer multidisciplinary teams. Int J Med Inform 2021 Dec;156:104610 [FREE Full text] [doi: 10.1016/j.ijmedinf.2021.104610] [Medline: 34649110]
47. van Dyk L, Schutte CS. The telemedicine service maturity model: a framework for the measurement and improvement of telemedicine services. Telemedicine 2013;5:217-238. [doi: 10.5772/56116]
48. van de Wetering R, Batenburg R. A PACS maturity model: a systematic meta-analytic review on maturation and evolvability of PACS in the hospital enterprise. Int J Med Inform 2009 Feb;78(2):127-140. [doi: 10.1016/j.ijmedinf.2008.06.010] [Medline: 18752988]
49. Jami Pour M, Jafari SM. Toward a maturity model for the application of social media in healthcare. Online Inf Rev 2019 Jun 10;43(3):404-425. [doi: 10.1108/oir-02-2018-0038]
50. Ipe M. Knowledge sharing in organizations: a conceptual framework. Hum Resour Dev Rev 2003 Dec 01;2(4):337-359. [doi: 10.1177/1534484303257985]
51. Dalkey N, Helmer O. An experimental application of the DELPHI method to the use of experts. Manag Sci 1963 Apr;9(3):458-467. [doi: 10.1287/mnsc.9.3.458]
52. Monazam Tabrizi N. Models for describing knowledge sharing practices: the case study of UK hospitals [PhD thesis]. The University of Manchester. 2017. URL: https://research.manchester.ac.uk/en/studentTheses/models-for-describing-knowledge-sharing-practices-the-case-study- [accessed 2024-04-29]
53. Lee HS, Hong SA. Factors affecting hospital employees' knowledge sharing intention and behavior, and innovation behavior. Osong Public Health Res Perspect 2014 Jun;5(3):148-155 [FREE Full text] [doi: 10.1016/j.phrp.2014.04.006] [Medline: 25180147]
54. Shahmoradi L, Safdari R, Piri Z, Dehghani Mahmodabadi A, Shahmoradi S, Farzaneh Nejad A. Knowledge sharing as a powerful base for management: barriers and solutions. Health Care Manag (Frederick) 2017;36(2):176-183. [doi: 10.1097/HCM.0000000000000150] [Medline: 28379919]
55. Gider Ö, Ocak S, Top M. Perceptions of physicians about knowledge sharing barriers in Turkish health care system. J Med Syst 2015 May 3;39(5):42. [doi: 10.1007/s10916-015-0230-6] [Medline: 25732078]
56. Kim S, Park M. Leadership, knowledge sharing, and creativity: the key factors in nurses' innovative behaviors. J Nurs Adm 2015 Dec;45(12):615-621. [doi: 10.1097/NNA.0000000000000274] [Medline: 26565640]
57. Imran MK, Fatima T, Aslam U, Muhammad Javed Iqbal S. Exploring the benefits of social media towards knowledge sharing among doctors. Pak J Psychol Res 2019 Jul 23;34(2):331-352. [doi: 10.33824/PJPR.2019.34.2.18]
58. Yip K. Exploring barriers to knowledge sharing: a case study of a virtual community of practice in a Swedish multinational corporation [Bachelor's thesis]. Linnėuniversitetet. 2011. URL: https://www.diva-portal.org/smash/get/diva2:457132/FULLTEXT01.pdf [accessed 2025-04-30]
59. Ariati N, Indra Sensuse D, Wuri Handayani P. Factors affecting knowledge sharing capability of doctors in Palembang. J Phys Conf Ser 2020 Apr 01;1500(1):012092. [doi: 10.1088/1742-6596/1500/1/012092]
60. Lin TC, Lai MC, Yang SW. Factors influencing physicians' knowledge sharing on web medical forums. Health Informatics J 2016 Sep 26;22(3):594-607 [FREE Full text] [doi: 10.1177/1460458215576229] [Medline: 25888432]
61. MamoMulate E, Gojeh LA. Current status and factors affecting knowledge sharing practices among health professionals in Hiwot Fana specialized university hospital in Ethiopia. Univers J Manag 2020 Jul;8(4):160-174. [doi: 10.13189/ujm.2020.080405]
62. Demsash A, Chakilu B, Mazengia A. Knowledge sharing practice and its associated factors among healthcare providers at University of Gondar Comprehensive Specialized Hospital, North West Ethiopian: cross-sectional study. Res Sq 2021:39. [doi: 10.21203/rs.3.rs-407759/v1]
63. Almeida M, Ramos A, Sousa M, Santos C, Fontes A. Organizational factors affecting knowledge sharing capabilities: a study with Portuguese health professionals. In: Proceedings of the 21st European Conference on Knowledge Management.

XSL•FO
RenderX

2023 Presented at: ECKM '23; September 7-8, 2023; London, UK p. 46 URL: https://www.researchgate.net/publication/371256729_Organizational_Factors_Affecting_Knowledge_Sharing_Capabilities_A_Study_with_Portuguese_Health_Professionals
64. Pun JK, Matthiessen CM, Murray KA, Slade D. Factors affecting communication in emergency departments: doctors and nurses' perceptions of communication in a trilingual ED in Hong Kong. Int J Emerg Med 2015 Dec 15;8(1):48 [FREE Full text] [doi: 10.1186/s12245-015-0095-y] [Medline: 26667242]
65. Creswell J, Creswell J. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. Thousand Oaks, CA: Sage Publications; 2017.
66. Alhalhouli Z, Hassan Z, Abualkishik A. An updated model to enhance knowledge sharing among stakeholders in Jordanian hospitals using social networks. Middle East J Sci Res 2013;18(8):1089-1098 [FREE Full text]
67. Noor NM, Salim J. Factors influencing employee knowledge sharing capabilities in electronic government agencies in Malaysia. Int J Comput Sci Issue 2011;8(4):106 [FREE Full text]
68. Lee HS. Knowledge management enablers and process in hospital organizations. Osong Public Health Res Perspect 2017 Feb 28;8(1):26-33 [FREE Full text] [doi: 10.24171/j.phrp.2017.8.1.04] [Medline: 28443221]
69. Linstone H, Turoff M. Delphi Method: Techniques and Applications. Cambridge, MA: Addison-Wesley Educational Publishers; 1975.
70. Fowles J, Fowles R. Handbook of Futures Research. New York, NY: Greenwood Press; 1978.
71. Cyphert FR, Gant WL. The delphi technique: a case study. Phi Delta Kappan 1971;52(5):272-273 [FREE Full text]
72. Akins RB, Tolson H, Cole BR. Stability of response characteristics of a Delphi panel: application of bootstrap data expansion. BMC Med Res Methodol 2005 Dec 01;5(1):37 [FREE Full text] [doi: 10.1186/1471-2288-5-37] [Medline: 16321161]
73. Gordon T. The Delphi method. In: Glenn JC, Gordon TJ, editors. Futures Research Methodology. Version 3. New York, NY: The Millennium Project; 1994:1.
74. Rüetschi U, Olarte Salazar CM. An e-Delphi study generates expert consensus on the trends in future continuing medical education engagement by resident, practicing, and expert surgeons. Med Teach 2019 Dec 27;42(4):444-450. [doi: 10.1080/0142159x.2019.1704708]
75. Faherty V. Continuing social work education: results of a Delphi survey. J Educ Soc Work 1979 Jan;15(1):12-19. [doi: 10.1080/00220612.1979.10671539]
76. Raskin MS. The Delphi study in field instruction revisited: expert consensus on issues and research priorities. J Educ Soc Work 2013 May 31;30(1):75-89. [doi: 10.1080/10437797.1994.10672215]
77. Almashmoum M, Cunningham J, Alkhaldi O, Anisworth J. Factors that affect knowledge-sharing behaviors in medical imaging departments in cancer centers: systematic review. JMIR Hum Factors 2023 Jul 12;10:e44327 [FREE Full text] [doi: 10.2196/44327] [Medline: 37436810]
78. Naughton B, Roberts L, Dopson S, Brindley D, Chapman S. Medicine authentication technology as a counterfeit medicine-detection tool: a Delphi method study to establish expert opinion on manual medicine authentication technology in secondary care. BMJ Open 2017 May 06;7(5):e013838 [FREE Full text] [doi: 10.1136/bmjopen-2016-013838] [Medline: 28478398]
79. Chia-Chien H, Sandford BA. The Delphi technique: making sense of consensus. Pract Assess Res Eval 2007 Sep;19(7):10. [doi: 10.7748/paed.19.7.10.s18]
80. Cabrera A, Cabrera EF. Knowledge-sharing dilemmas. Organ Stud 2002 Sep 01;23(5):687-710. [doi: 10.1177/0170840602235001]
81. Adeyelure T, Kalema B, Motlanthe B. An empirical study of knowledge sharing: a case of South African healthcare system. Knowl Manag E-Learn 2019;11(1):114-128. [doi: 10.34105/j.kmel.2019.11.007]
82. Sigala M, Chalkiti K. Improving performance through tacit knowledge externalisation and utilisation. Int J Product Perform Manag 2007 Jun 26;56(5/6):456-483. [doi: 10.1108/17410400710757141]
83. Nonaka I, von Krogh G, Voelpel S. Organizational knowledge creation theory: evolutionary paths and future advances. Organ Stud 2006 Aug 01;27(8):1179-1208. [doi: 10.1177/0170840606066312]
84. Pinjani P, Palvia P. Trust and knowledge sharing in diverse global virtual teams. Inf Manag 2013 Jun;50(4):144-153. [doi: 10.1016/j.im.2012.10.002]
85. Chen C, Hung S. To give or to receive? Factors influencing members' knowledge sharing and community promotion in professional virtual communities. Inf Manag 2010 May;47(4):226-236. [doi: 10.1016/j.im.2010.03.001]
86. Miller D, Moran T. Self-Esteem: A Guide for Teachers. Thousand Oaks, CA: Sage Publications; 2012.
87. Kong DT. Ostracism perception as a multiplicative function of trait self-esteem, mindfulness, and facial emotion recognition ability. Pers Individ Dif 2016 Apr;93:68-73. [doi: 10.1016/j.paid.2015.08.046]
88. Ismail M, Yusof Z. Knowledge sharing models: do they really fit public organizations? In: Proceedings of the 2008 International Symposium on Information Technology. 2008 Presented at: ITSIM '08; August 26-28, 2008; Kuala Lumpur, Malaysia p. 1-9 URL: https://ieeexplore.ieee.org/document/4631661 [doi: 10.1109/itsim.2008.4631661]
89. Lin W. The effect of knowledge sharing model. Expert Syst Appl 2008 Feb;34(2):1508-1521. [doi: 10.1016/j.eswa.2007.01.015]
90. Ayestarán S, Gómez D, Martínez-Moreno E, Lira E, Da CS. A model of knowledge-sharing for the 21st century organizations. J Work Organ Psychol 2023;38(3):175-187. [doi: 10.5093/jwop2022a21]

XSL•FO
RenderX

91. Suryani N, Rustiana A, Muhsin M, Rahmaningtyas W. Diversity perspectives in teamwork: is the difference importance? Int J Res Bus Soc Sci 2019 Dec 30;9(1):24-30. [doi: 10.20525/ijrbs.v9i1.607]
92. Hutchinson V, Quintas P. Do SMEs do knowledge Management? Int Small Bus J 2008 Apr 01;26(2):131-154. [doi: 10.1177/0266242607086571]
93. Olan F, Ogiemwonyi Arakpogun E, Suklan J, Nakpodia F, Damij N, Jayawickrama U. Artificial intelligence and knowledge sharing: contributing factors to organizational performance. J Bus Res 2022 Jun;145:605-615. [doi: 10.1016/j.jbusres.2022.03.008]
94. Arakpogun EO, Elsahn Z, Olan F, Elsahn F. Artificial intelligence in Africa: challenges and opportunities. In: Hamdan A, Hassanien AE, Razzaque A, Alareeni B, editors. The Fourth Industrial Revolution: Implementation of Artificial Intelligence for Growing Business Success. Cham, Switzerland: Springer; 2021:375-388.

Abbreviations

AI: artificial intelligence
ICT: information and communication technology
KCCC: Kuwait Cancer Control Center
KSMM: knowledge sharing maturity model
MDT: multidisciplinary team
MM: maturity model
PACS: picture archiving and communication system

Edited by A Choudhury; submitted 11.11.23; peer-reviewed by T Liaw, X Jing, V Kumar, E Borycki; comments to author 21.04.24; revised version received 22.08.24; accepted 12.02.25; published 06.05.25.

Please cite as:
Almashmoum M, Cunningham J, Ainsworth J
Knowledge Sharing Maturity Model for Medical Imaging Departments: Development Study
JMIR Hum Factors 2025;12:e54484
URL: https://humanfactors.jmir.org/2025/1/e54484
doi:10.2196/54484
PMID:

©Maryam Almashmoum, James Cunningham, John Ainsworth. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 06.05.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Auxiliary Teaching and Student Evaluation Methods Based on Facial Expression Recognition in Medical Education

Xueling Zhu, MSE; Roben A Juanatas, PhD

College of Computing and Information Technologies, National University, 551 Mariano Fortunato Jhocson Street, Sampaloc, Manila, Philippines

Corresponding Author:
Xueling Zhu, MSE
College of Computing and Information Technologies, National University, 551 Mariano Fortunato Jhocson Street, Sampaloc, Manila, Philippines

Abstract

Traditional medical education encounters several challenges. The introduction of advanced facial expression recognition technology offers a new approach to address these issues. The aim of the study is to propose a medical education–assisted teaching and student evaluation method based on facial expression recognition technology. This method consists of 4 key steps. In data collection, multiangle high-definition cameras record students' facial expressions to ensure data comprehensiveness and accuracy. Facial expression recognition uses computer vision and deep learning algorithms to identify students' emotional states. The result analysis stage organizes and statistically analyzes the recognized emotional data to provide teachers with students' learning status feedback. In the teaching feedback stage, teaching strategies are adjusted according to the analysis results. Although this method faces challenges such as technical accuracy, device dependency, and privacy protection, it has the potential to improve teaching effectiveness, optimize personalized learning, and promote teacher-student interaction. The application prospects of this method in medical education are broad, and it is expected to significantly enhance teaching quality and students' learning experience.

(JMIR Hum Factors 2025;12:e72838) doi:10.2196/72838

KEYWORDS

medical education; facial expression recognition; auxiliary teaching; student assessment; teaching innovation

Introduction

Under the traditional mode of medical education, although rich teaching experience and knowledge system have been accumulated for a long time, some inherent problems have gradually emerged with the advancement of technology and changes in social needs [1,2]. First, in terms of teaching methods, traditional medical education generally adopts a teacher-led teaching mode, which often neglects students' initiative and participation, and limits the cultivation of their independent thinking and practical abilities [3]. Second, in terms of evaluation mechanism, traditional medical education mainly relies on written examinations and clinical internships, but this evaluation method often cannot fully and accurately reflect students' learning outcomes and abilities. Especially in nonacademic evaluations, such as emotions, attitudes, and values, traditional methods seem inadequate [4]. In addition, with the continuous updating of medical knowledge and the rapid development of technology, traditional medical education often lags behind the latest scientific research achievements and practical experience in teaching content and methods, which makes it difficult for students to adapt to the rapidly changing medical environment after graduation. Therefore, it is necessary to introduce new teaching methods and technological means to make up for the shortcomings of traditional medical education and improve the quality and effectiveness of medical education.

Facial expression recognition technology is a technique that recognizes and understands an individual's emotional state by capturing and analyzing facial expressions [5-8]. The development history of facial expression recognition technology can be traced back to the 1970s when American psychologists Ekman and Friesen pioneered the definition of 6 basic human expressions and established the Facial Action Coding System, laying the foundation for subsequent research on facial expression recognition [9-11]. This milestone work not only clarifies the framework for facial expression classification but also provides standardized analytical tools for subsequent research [12].

Entering the 21st century, with the vigorous development of computer vision and artificial intelligence technology, facial expression recognition technology has ushered in new opportunities for development. In early research, researchers mainly relied on traditional image processing and machine learning algorithms such as support vector machines (SVMs) and principal component analysis to process and analyze facial expressions [13,14]. However, the recognition accuracy of these methods is limited by manually designed feature extraction methods and limited training data [15-17].

In recent years, the rise of deep learning technology has brought revolutionary changes to the field of facial expression recognition. Deep learning models such as convolutional neural

XSL•FO
RenderX

networks (CNNs) and recurrent neural networks have achieved significant results in facial expression recognition [18-20]. For example, AlexNet proposed by Krizhevsky et al has made breakthrough progress in ImageNet image classification tasks, providing strong support for the application of deep learning in expression recognition [21]. Subsequently, researchers began to apply deep learning models to facial expression recognition tasks, continuously improving the accuracy and robustness of facial expression recognition by constructing deeper network structures and introducing attention mechanisms and other methods.

In specific research work, there are several aspects that deserve special attention. First, multimodal fusion technology has become a research hotspot in recent years. Researchers have found that in addition to facial expressions, a person's emotional state can also be expressed through other modalities such as speech, posture, and physiological signals. Therefore, integrating information from multiple modalities can further improve the accuracy and reliability of facial expression recognition. For example, some studies have achieved more accurate emotion recognition by integrating facial expressions and voice information [22].

Second, the application of self-supervised learning in facial expression recognition is gradually receiving attention. Due to the high cost of obtaining labeled data, researchers have begun exploring the use of unlabeled data for pretraining to improve the model's generalization ability. This self-supervised learning method can fully use a large amount of unlabeled data resources and extract useful feature representations by designing appropriate pretraining tasks. For example, some studies have achieved significant performance improvements by designing self-supervised learning tasks based on facial movements to pretrain models [23].

In addition, with the continuous development of computer vision technology, some new technologies have also been introduced into facial expression recognition. For example, 3D facial expression recognition technology can more accurately recognize facial expressions by capturing 3D information of the face. This technology can overcome issues such as lighting and posture in 2D images and improve the robustness of facial expression recognition. In addition, some researchers have also explored expression synthesis techniques based on generative adversarial networks, which can generate facial images with specific expressions, providing more training data and testing samples for expression recognition [24-26].

Introducing facial expression recognition technology to assist teaching and student assessment in medical education can bring significant advantages [27]. First, this technology can capture students' facial expressions in real time, providing teachers with intuitive and immediate feedback, thereby helping them more accurately grasp students' learning status, adjust teaching strategies in a timely manner, and improve teaching effectiveness. Second, facial expression recognition technology can help students better understand teaching content. Through students' facial feedback, teachers can promptly answer students' questions and enhance students' learning motivation and participation [28]. In addition, facial expression recognition technology can also help achieve personalized teaching. Based on the emotional reactions of different students, teachers can develop more targeted teaching plans to meet their personalized needs [29].

Facial expression recognition technology has unique advantages in solving traditional medical education problems. Traditional medical education often relies on written examinations and clinical internships to evaluate students' learning status, but this evaluation method is difficult to comprehensively and accurately reflect students' learning outcomes and abilities. Especially for nonacademic evaluations of students' emotions, attitudes, and values, traditional methods seem inadequate. Facial expression recognition technology can capture students' emotional reactions, providing teachers with a more comprehensive evaluation basis, thereby more accurately assessing students' comprehensive qualities and potential. In addition, traditional medical education often adopts a teacher-led teaching mode in teaching methods, lacking student participation and interaction. Facial expression recognition technology can obtain real-time feedback from students, making teaching more interactive and targeted and improving students' learning outcomes.

Facial Expression Recognition Technology and Its Application in Medical Education

Principles of Facial Expression Recognition Technology

Facial expression recognition technology is a technique that uses image processing and analysis to recognize and understand human facial expressions. It is mainly based on algorithms, such as machine learning and deep learning, and achieves automatic recognition and classification of facial expressions through training and learning on a large number of facial expression images [30-33]. The facial expression recognition technology mainly relies on computer vision and pattern recognition technology and achieves automatic recognition and classification of facial expressions through feature extraction and classification of facial images [34]. This process can be summarized as the following steps:

- Image input: first, an image containing facial expressions needs to be input. These images can come from various sources, such as real-time video streams captured by cameras and static image libraries.
- Facial detection: using facial detection algorithms to accurately locate facial regions in the input image. This is one of the key steps in facial expression recognition technology, as only by accurately detecting the face can facial expression features be further extracted.
- Image preprocessing: preprocessing the detected facial images, including grayscale, normalization, denoising, and other operations. These preprocessing steps help eliminate redundant information and interference factors in the image, improving the accuracy of subsequent feature extraction.
- Feature extraction: in the preprocessed image, use feature extraction algorithms to extract key features that can characterize facial expressions. These features can include geometric features of the face (such as the shape and

XSL•FO
RenderX

position of the eyes and mouth), texture features (such as skin texture and wrinkles), and dynamic features (such as the movement of facial muscles). Feature extraction is one of the core steps in facial expression recognition technology, and the quality of the extracted features directly affects the accuracy of subsequent facial expression recognition.

- Pattern classification: after extracting facial expression features, machine learning or deep learning algorithms are used to classify these features. These algorithms can automatically recognize different facial expression categories by training and learning from a large number of labeled facial expression images. Common machine learning algorithms include SVM, decision tree, and random forest; deep learning algorithms include CNN, recurrent neural networks, and so on. These algorithms continuously optimize model parameters during the training process to improve the accuracy and generalization ability of facial expression recognition.

Potential Applications of Facial Expression Recognition in Medical Education

In medical education, facial expression recognition technology can be applied in 2 aspects: assisting teaching and student assessment [35]. First, teachers can analyze students' facial expressions to understand their understanding and mastery of the teaching content and adjust teaching strategies and methods accordingly. Second, student assessment can also achieve more objective and accurate evaluation results through facial expression recognition technology. For example, students' participation, interest, and learning status can be evaluated by analyzing their facial expressions during classroom discussions, experimental operations, and other processes.

Auxiliary Teaching: Optimizing Teaching Strategies Through Facial Expression Recognition

In medical education, teachers can observe students' facial expressions to understand their understanding and mastery of the teaching content [36]. Facial expression recognition technology can provide effective support for this. By analyzing students' facial expressions in real time, teachers can quickly identify their emotional states such as confusion, interest, or fatigue. Based on this information, teachers can adjust teaching strategies and methods in a timely manner to better meet students' learning needs.

Specifically, when teachers find that students are confused about a certain knowledge point, they can pause the explanation and clarify it through questioning or giving examples; when students show a strong interest, they can further expand their relevant knowledge points to stimulate their learning interest; and when students show fatigue, they can adjust the teaching pace appropriately and relieve their fatigue through rest or small games.

Student Assessment: Achieving Objective and Accurate Assessment Through Facial Expression Recognition

In addition to assisting teaching, facial expression recognition technology can also be used for student assessment. Traditional student assessment methods often rely on students' self-reports or teachers' subjective judgments, which have certain subjectivity and errors. Facial expression recognition technology can evaluate students' participation, interest, and learning status by analyzing their facial expressions, providing more objective and accurate evaluation results.

For example, in classroom discussions, teachers can evaluate students' engagement by analyzing their facial expressions. When students actively participate in discussions, they usually show a positive and focused expression; when student engagement is low, they may exhibit expressions of boredom, fatigue, or confusion. By collecting and analyzing these facial expression data, teachers can more accurately evaluate students' classroom performance.

In addition, facial expression recognition technology can also be used to evaluate students' learning status in experimental operations. When students encounter difficulties or challenges, they usually show expressions of tension, anxiety, or confusion; when students successfully solve problems, they will show excited and satisfied expressions. By analyzing these facial expression data, teachers can understand students' learning status during experimental operations and provide timely guidance and assistance.

Auxiliary Teaching and Student Evaluation Methods Based on Facial Expression Recognition in Medical Education

Overview of Methods

The auxiliary teaching and student evaluation method based on facial expression recognition in medical education proposed in this section mainly include the following 4 steps: data collection, facial expression recognition, result analysis, and teaching feedback.

Data Collection

In the medical education–assisted teaching and student evaluation method based on facial expression recognition, the data collection process is the foundation and key of the entire process. The primary task of data collection is to obtain high-quality video and image data, which mainly comes from the facial expression records of students during the teaching process. To ensure the comprehensiveness and accuracy of data, multiple high-definition cameras are usually installed in the classroom to capture students' facial expressions from different angles. These cameras need to have good resolution and frame rate to ensure clear and detailed images. In addition, portable devices such as smartphones and tablets can be used for supplementary shooting to meet the flexible data collection needs in special situations.

The collected data should not only include static facial images but also dynamic video recordings. This is because the changes in facial expressions are often continuous, and a single static image may not fully reflect the emotional state of students. Therefore, the collection of video data is crucial, as it can capture the changes in students' expressions at different time points, providing richer information for subsequent expression

XSL•FO
RenderX

recognition and analysis. At the same time, video data can also record classroom interactions, including scenarios such as students' questions, answers, and discussions. These behavioral data also have important reference value for evaluating teaching effectiveness.

During the data collection process, special attention should be paid to privacy protection and data security. Due to the personal privacy involved in the collected facial expression data, it is necessary to strictly comply with relevant ethical norms and laws and regulations. Prior to data collection, informed consent from students should be obtained, clearly informing them of the purpose, scope of use, and protective measures of the data collection. At the same time, encryption technology is used in the process of data storage and transmission to prevent data leakage and ensure that student privacy is not violated. In addition, in order to further improve the confidentiality of the data, anonymization can be performed on the collected data, that is, personal identity information can be separated from facial expression data through technical means, thereby ensuring the effectiveness of data analysis while maximizing the protection of student privacy.

Finally, in order to ensure the high quality and practicality of the data, detailed standards and specifications need to be developed in the data collection process. These standards include the installation position and angle of the camera, control of lighting conditions, and selection of shooting time. For example, cameras should be installed at angles that can clearly capture students' faces and avoid adverse lighting conditions such as backlighting. The shooting time should cover the entire teaching process, including explanations, interactions, and breaks, to comprehensively reflect the changes in students' expressions at different stages of teaching. By formulating and implementing these standards and specifications, the accuracy and reliability of data collection can be maximized, providing a solid foundation for subsequent facial expression recognition and result analysis.

Facial Expression Recognition

In medical education–assisted teaching and student evaluation methods based on facial expression recognition, the facial expression recognition process is a key step in converting the collected raw data into meaningful emotional information. The process of facial expression recognition first relies on advanced computer vision and machine learning technologies, which can automatically detect and analyze key feature points of students' faces. Specifically, facial detection algorithms locate facial regions by processing video frames or images and further extract key feature points including eyes, eyebrows, nose, and mouth. The precise localization of these feature points is the foundation for facial expression recognition, ensuring that the algorithm can accurately capture subtle changes in the face.

After extracting key feature points, the expression recognition system will use the trained classifier to recognize different expression categories. Common classifiers include SVM and CNN. These classifiers can learn and distinguish feature patterns corresponding to various expressions by training on a large amount of labeled data. Especially, CNNs have shown superior performance in facial expression recognition, as they can automatically extract high-level feature representations, significantly improving the accuracy and robustness of facial expression recognition. In the recognition process, the system not only needs to distinguish basic expressions, such as happiness, sadness, and surprise, but also needs to combine contextual information to recognize complex emotional states, such as confusion, focus, or boredom.

To further improve the accuracy of facial expression recognition, multimodal fusion is usually required, which combines multiple sources of information such as facial expressions, speech emotions, and posture changes. Multimodal fusion can make up for the limitations of a single mode and provide more comprehensive and accurate emotion recognition results. For example, when students write with their heads down, their facial expressions may not fully capture their emotional state. In this case, combining speech intonation and body movement information can enhance the accuracy and reliability of emotion recognition. Through the fusion processing of multimodal data, the system can better understand the true emotional state of students and provide rich emotional data for subsequent result analysis.

In addition, the accuracy of facial expression recognition also depends on the personalized modeling of different individuals. Due to individual differences in facial features and expressions among students, a unified recognition model may not be able to adapt to all individuals. Therefore, introducing personalized modeling mechanisms into facial expression recognition systems and establishing personalized facial expression recognition models for each student can significantly improve recognition accuracy. This process typically involves learning and adapting individual features, dynamically updating the model using previously collected data to continuously optimize recognition performance.

The final output of the facial expression recognition process is a detailed record of the emotional state of each student throughout the entire teaching process. These records include time series data of emotional states as well as emotional fluctuations in different teaching stages. By organizing and analyzing these emotional data, comprehensive emotional feedback can be provided, laying a solid foundation for subsequent result analysis and teaching feedback. The facial expression recognition stage is not only the initial processing stage of emotional data but also the core stage of the entire auxiliary teaching and student evaluation method. Its accuracy and effectiveness directly affect the final quality of teaching feedback.

Result Analysis

In the medical education–assisted teaching and student evaluation method based on facial expression recognition, the result analysis stage plays a bridging role, which is a key step in transforming the identified emotional data into a deep understanding of teaching effectiveness and student learning status. The result analysis first involves the organization and preprocessing of emotional data, including smoothing of time series, detection and removal of outliers, and statistical analysis of the distribution of different emotional categories. Through preliminary processing of emotional data, the emotional change

XSL•FO
RenderX

curves of each student at different time periods can be obtained, which provides a foundation for further in-depth analysis.

Based on the organized emotional data, statistical analysis and data mining techniques are further used to reveal the relationship between students' emotional states and teaching activities. Specifically, through regression analysis, correlation analysis, and other methods, the impact of different teaching stages, content, and methods on students' emotional states can be explored. For example, analyzing the changes in students' concentration, confusion, and participation in different teaching stages, such as teacher lectures, interactive questioning, and experimental operations, in order to evaluate the teaching effectiveness of these stages. In addition, through cluster analysis and pattern recognition, student groups with similar emotional change patterns can be identified, providing a basis for personalized teaching and hierarchical guidance.

To gain a deeper understanding of the teaching significance behind students' emotional data, result analysis should also be combined with the specific context of classroom teaching and the teacher's teaching design. By synchronously analyzing emotional data and teaching videos, it is possible to identify which specific teaching behaviors have caused students' emotional fluctuations. For example, when students exhibit general confusion or anxiety at a certain point in time, it is possible to trace back to the teaching content and the teacher's explanation during that period and identify the difficult issues or unclear explanations that may have caused confusion for students. This situational analysis not only helps diagnose problems in teaching but also provides specific reference suggestions for teachers to improve their teaching.

In the result analysis stage, special attention should be paid to the handling of individual differences. Due to significant differences in emotional performance and learning styles among each student, a unified analysis method may not fully reflect an individual's learning status. Therefore, the result analysis should be combined with personalized emotional data for targeted analysis. For example, by longitudinal tracking and comparative analysis of a student's emotional change data, it is possible to evaluate their emotional response under different courses and teaching methods and identify the most suitable teaching strategy for their learning style. In this way, not only can it improve the accurate assessment of individual learning status, but it can also provide a scientific basis for personalized teaching.

Finally, the output of the result analysis stage is a comprehensive evaluation report of the entire teaching process and students' emotional state. These reports include statistical results of emotional data, trend analysis of emotional changes, emotional feedback in teaching processes, and personalized teaching recommendations. Through these detailed analysis reports, teachers can comprehensively understand the teaching effectiveness and students' learning status, adjust teaching strategies and methods in a targeted manner, and thus achieve continuous improvement in teaching quality. Result analysis is not only a simple processing of emotional data but also a data-driven teaching improvement process, and its accuracy and scientificity directly affect the effectiveness and operability of teaching feedback.

Teaching Feedback

In the medical education–assisted teaching and student evaluation method based on facial expression recognition, the teaching feedback link is the terminal link of the entire process, aiming to transform the conclusions drawn from the result analysis into specific teaching improvement measures. Teaching feedback is not only feedback on students' learning status and teaching effectiveness but also guidance for future teaching design and strategies. Medical education has its unique characteristics, including highly specialized knowledge content, rigorous practical skills training, and comprehensive requirements for students' overall quality, which makes the teaching feedback link particularly important in medical education.

First, in the teaching feedback stage, the emotional data and behavioral patterns identified in the result analysis need to be fed back to the teacher. This process not only includes the aggregation and presentation of emotional data but also requires providing detailed explanations and analysis reports in conjunction with specific teaching contexts and classroom activities. For example, in the result analysis, it is found that some students exhibit obvious confusion or anxiety in specific teaching stages. Teaching feedback should provide a detailed explanation of the possible reasons for these emotional changes and provide corresponding teaching improvement suggestions. For medical education, this meticulous feedback helps teachers to explain complex medical concepts or skill operations in a more targeted manner in subsequent teaching, improving teaching effectiveness.

Second, the teaching feedback process should also focus on providing individual feedback and guidance to students. Students in medical education not only need to master a large amount of theoretical knowledge but also need to possess solid clinical skills and good professional ethics. Personalized feedback can help students understand their weak areas in learning and provide specific improvement suggestions. For example, through facial expression recognition and emotion analysis of students in simulated clinical operations, feedback can point out the possible tension or lack of confidence that students may have during the operation process and suggest that they improve their operational skills and psychological qualities through more practice or psychological counseling. Such personalized feedback not only helps students make progress in their studies but also assists them in better coping with various challenges in their future careers.

In addition, the teaching feedback process should also take into account the characteristics of team learning and collaborative practice in medical education. Medical education not only focuses on individual learning but also emphasizes the importance of teamwork and collective decision-making. Therefore, teaching feedback should include emotional and behavioral analysis of student teamwork, identifying potential communication barriers or collaboration issues within the team. For example, through facial expression recognition technology, analyzing students' emotional states during team discussions

XSL•FO
RenderX

or collaborative tasks can identify which members may have issues with insufficient participation or poor communication and provide suggestions for improving the team's collaboration methods. In this way, not only can it improve the efficiency of team learning, but it can also cultivate students' cooperation ability and team spirit, which are key qualities in medical education.

Finally, the teaching feedback process should focus on long-term effectiveness evaluation and continuous improvement. Medical education is a long-term process, and the cultivation of students' learning outcomes and professional qualities requires continuous attention and improvement. Through regular emotional data analysis and teaching feedback, teachers can dynamically adjust teaching strategies and continuously optimize teaching content and methods. For example, regularly reviewing and summarizing students' emotional data and learning performance, evaluating the effectiveness of teaching improvement measures, and proposing new teaching improvement plans based on this. In this way, teaching feedback is not only a reflection and improvement of current teaching activities but also a continuous and dynamic process of improving teaching quality.

In summary, the feedback loop in teaching is of great significance in medical education–assisted teaching and student evaluation methods based on facial expression recognition. Through meticulous analysis of emotional data and personalized feedback, not only can teaching effectiveness be improved and students overcome learning difficulties, but also team collaboration and long-term improvement of teaching quality can be promoted. The particularity of medical education requires that the teaching feedback process pays more attention to individual differences, teamwork, and continuous improvement in order to comprehensively enhance the quality of medical education and the comprehensive quality of students.

Advantages and Limitations of the Method

Advantages of Methods

The medical education–assisted teaching and student evaluation methods based on facial expression recognition have multiple advantages, which are particularly significant in the special context of medical education. First, through real-time facial expression recognition technology, it is possible to objectively and instantly capture and analyze students' emotional states in the classroom. This real-time data acquisition and analysis enables teachers to timely understand students' emotional changes during the teaching process, especially their understanding of complex or abstract medical concepts. When students show negative emotions, such as confusion or anxiety, teachers can immediately adjust teaching strategies, such as providing more detailed explanations or adding interactive elements, to alleviate students' emotions and improve teaching effectiveness. This real-time feedback mechanism significantly enhances the dynamic adjustment ability of classroom teaching, ensuring that every student can keep up with the teaching progress.

Second, methods based on facial expression recognition can provide refined personalized teaching feedback. The content of medical education is complex and covers a wide range of professional knowledge, and students' understanding and mastery vary greatly. By analyzing students' facial expression data in different teaching stages, specific difficulties and emotional fluctuations of individuals in the learning process can be accurately identified. For example, when studying anatomy, if a student shows persistent confusion in a specific chapter, the facial expression recognition system can capture and record this phenomenon and provide targeted learning suggestions for the student in teaching feedback. This personalized feedback mechanism can not only help students overcome obstacles in learning but also improve their learning efficiency and effectiveness, thereby achieving the goal of teaching according to their aptitude.

In addition, this method has unique advantages in cultivating students' clinical practice skills and comprehensive qualities. Medical education not only emphasizes the mastery of theoretical knowledge but also pays more attention to the training of clinical skills and the cultivation of professional qualities. Through facial expression recognition technology, students' emotional states can be monitored and analyzed in real time during simulated clinical operations and actual clinical training. For example, during simulated surgical procedures, the system can capture students' emotional changes, such as tension, anxiety, or confidence, and use these data to evaluate their proficiency and psychological resilience. This evaluation method is more objective and comprehensive than traditional manual observation, which helps teachers to strengthen students' practical training and psychological development in teaching feedback and improve their clinical and emergency response abilities.

In addition, the evaluation method based on facial expression recognition has also demonstrated significant advantages in cultivating teamwork and communication skills. Many tasks and projects in medical education require students to collaborate as a team, and effective communication and collaboration among team members are key to success. Facial expression recognition technology can monitor and analyze the emotional states of team members during collaboration, identifying potential communication barriers or collaboration issues. For example, in team discussions or collaborative tasks, if the system detects insufficient participation or low mood among certain members, it can promptly provide feedback to the teacher or team leader, prompting them to take measures to improve the team's collaborative atmosphere. This team evaluation and feedback based on emotional data not only improves the efficiency of team learning but also cultivates students' communication skills and teamwork spirit.

Finally, the teaching evaluation method based on facial expression recognition has the advantage of continuous improvement. Medical education is a long-term and dynamic process that requires continuous adjustment and optimization of teaching strategies. By regularly collecting and analyzing emotional data, teachers can continuously evaluate teaching effectiveness and identify and improve deficiencies in teaching. For example, by tracking students' emotional changes and learning performance over the long term, evaluating the effectiveness of different teaching methods and content, and

XSL•FO
RenderX

continuously optimizing and adjusting them in subsequent teaching. This data-driven continuous improvement mechanism ensures the high quality of medical education and the efficiency of student learning, meeting the needs of modern medical education for refined and personalized teaching.

Limitations of the Method

Although facial expression recognition–based methods for medical education auxiliary teaching and student evaluation offer significant advantages, such as real-time feedback and objective assessment, their limitations, including issues related to data privacy, cross-cultural recognition accuracy, and the complexity of emotional interpretation, cannot be overlooked. First, the accuracy and reliability of facial expression recognition technology face challenges in practical applications. Facial expressions are influenced by various factors, including lighting, shooting angle, and facial occlusion, which may result in inaccurate recognition results. In addition, there are significant differences in facial expressions and emotional expression styles among individuals, which limits the applicability of universal expression recognition models in personalized applications. Especially in medical education, students may have more subtle emotional expressions when faced with complex professional content, which further increases the difficulty of facial expression recognition.

Second, the dependency of facial expression recognition technology is an important issue. This technology relies on high-quality video data and advanced computing devices, and in practical teaching environments, especially with limited resources, it is difficult to ensure that every classroom is equipped with sufficient cameras and computing power. This not only increases implementation costs but may also lead to incomplete or low-quality data collection, thereby affecting the accuracy and reliability of evaluation results. In addition, prolonged video surveillance may cause students' resistance and affect their natural performance, thereby interfering with the authenticity of emotional data.

Furthermore, facial expression recognition technology mainly focuses on facial expressions, and emotions and psychological states are complex and multidimensional. Relying solely on facial expression data is difficult to fully reflect students' true emotions and learning status. In medical education, students' stress, anxiety, and cognitive load are often complex and varied, and these internal states may not be fully expressed through facial expressions. Therefore, relying solely on facial expression recognition for emotional assessment has limitations and may lead to a one-sided understanding of students' learning status. In addition, the interpretation of emotional data is also subjective, and different teachers may have different interpretations of the same emotional data, which affects the consistency and effectiveness of teaching feedback.

In addition, privacy protection and ethical issues are major challenges based on facial expression recognition technology. Collecting and analyzing students' facial expression data involve personal privacy and must strictly comply with relevant laws, regulations, and ethical guidelines. In practical operation, how to ensure the security, anonymity, and transparency of data use is a complex and sensitive issue. If not handled properly, it may cause concerns for students and parents and even lead to legal disputes. In addition, long-term monitoring and data analysis may bring psychological pressure, affecting students' learning experience and mental health.

Finally, there is uncertainty in the application effectiveness of facial expression recognition–based evaluation methods in actual teaching improvement. Although emotional data can provide some reference, improving teaching effectiveness involves multiple factors, including teaching content, teaching methods, and teacher quality. Relying solely on emotional data is difficult to comprehensively improve teaching quality. Therefore, in the application process, it is necessary to combine other evaluation methods and means and comprehensively analyze and improve teaching strategies to ensure the comprehensiveness and scientificity of teaching feedback.

Conclusions and Prospect

In this paper, we explore the application of facial expression recognition technology in medical education and its potential in assisting teaching and student assessment. Through in-depth research and practical application analysis of facial expression recognition technology, we have found that this technology can significantly enhance teacher-student interaction, improve teaching effectiveness, and provide strong support for personalized education. Specifically, facial expression recognition technology can capture and analyze students' facial expressions in real time, helping teachers understand students' emotional states and understanding in a timely manner, thereby adjusting teaching strategies and providing more targeted guidance.

In addition, facial expression recognition technology can also serve as an effective student assessment tool, providing detailed feedback data for teachers by quantitatively analyzing students' emotional reactions and participation in the classroom. These data can not only be used to evaluate students' learning outcomes but also to improve curriculum design and teaching methods, thereby achieving a comprehensive improvement in educational quality.

Looking ahead to the future, with the continuous development and improvement of facial expression recognition technology, its application prospects in medical education will be even broader. We foresee that future medical education will become more intelligent and personalized, and facial expression recognition technology will be combined with other advanced technologies such as artificial intelligence and big data analysis to jointly build a comprehensive and dynamic education ecosystem. At the same time, with the popularization of technology and the accumulation of application experience, we also need to pay attention and solve the ethical and privacy issues that technology may face in practical applications, ensuring the sustainable development of technology in the field of education.

In short, facial expression recognition technology, as an innovative tool, can not only improve the teaching effectiveness and evaluation efficiency of medical education but also provide new ideas and directions for the future development of medical

XSL•FO
RenderX

education. We believe that with the deepening of research and the promotion of application, facial expression recognition technology will bring more innovation and change to medical education and help cultivate more outstanding medical talents.

Acknowledgments

This research was supported by the China National University Student Innovation & Entrepreneurship Development Program (S202212216023, AH202112216200, AH202112216202, and S202312216029), School-level Scientific Research Teams of Anhui Xinhua University (kytd202203), and Natural Science Foundation of Anhui Xinhua University (2024zr003).

Authors' Contributions

XZ wrote the first draft of the manuscript. All authors reviewed and edited the manuscript and approved the final version of the manuscript.

Conflicts of Interest

None declared.

References

1. Kaufman DM. Teaching and learning in medical education: how theory can inform practice. In: Understanding Medical Education: Evidence, Theory, and Practice: Wiley; 2018:37-69. [doi: 10.1002/9781119373780]
2. Maudsley G, Strivens J. Promoting professional knowledge, experiential learning and critical thinking for medical students. Med Educ 2000 Jul;34(7):535-544. [doi: 10.1046/j.1365-2923.2000.00632.x] [Medline: 10886636]
3. Frenk J, Chen L, Bhutta ZA, et al. Health professionals for a new century: transforming education to strengthen health systems in an interdependent world. Lancet 2010 Dec;376(9756):1923-1958. [doi: 10.1016/S0140-6736(10)61854-5]
4. Holmboe ES, Hawkins RE. Methods for evaluating the clinical competence of residents in internal medicine: a review. Ann Intern Med 1998 Jul 1;129(1):42-48. [doi: 10.7326/0003-4819-129-1-199807010-00011] [Medline: 9652999]
5. Barrett LF, Adolphs R, Marsella S, Martinez AM, Pollak SD. Emotional expressions reconsidered: challenges to inferring emotion from human facial movements. Psychol Sci Public Interest 2019 Jul;20(1):1-68. [doi: 10.1177/1529100619832930] [Medline: 31313636]
6. Russell JA. Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. Psychol Bull 1994 Jan;115(1):102-141. [doi: 10.1037/0033-2909.115.1.102] [Medline: 8202574]
7. Tian YL, Kanade T, Cohn JF. Recognizing action units for facial expression analysis. IEEE Trans Pattern Anal Mach Intell 2001 Feb;23(2):97-115. [doi: 10.1109/34.908962] [Medline: 25210210]
8. Choi S, Gao Y, Jin Y, et al. PPGface: like what you are watching? Earphones can “feel” your facial expressions. In: Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies: Association for Computing Machinery; 2022:1-32. [doi: 10.1145/3534597]
9. Ambadar Z, Ekman P. Observer-based measurement of facial expression with the facial action coding system. In: The Handbook of Emotion Elicitation and Assessment: Oxford University Press; 2007:203-221. [doi: 10.1093/oso/9780195169157.003.0014]
10. An AH. A comparative analysis on facial expression in advertisements—by utilising Facial Action Coding System (FACS). J Korea Contents Assoc 2019;19(3):61-71. [doi: 10.5392/JKCA.2019.19.03.061]
11. Kring AM, Sloan DM. The Facial Expression Coding System (FACES): development, validation, and utility. Psychol Assess 2007 Jun;19(2):210-224. [doi: 10.1037/1040-3590.19.2.210] [Medline: 17563202]
12. Kopalidis T, Solachidis V, Vretos N, Daras P. Advances in facial expression recognition: a survey of methods, benchmarks, models, and datasets. Information 2024;15(3):135. [doi: 10.3390/info15030135]
13. Mohammed AA, Minhas R, Jonathan Wu QM, Sid-Ahmed MA. Human face recognition based on multidimensional PCA and extreme learning machine. Pattern Recognit DAGM 2011 Oct;44(10-11):2588-2597. [doi: 10.1016/j.patcog.2011.03.013]
14. Pantic M. Machine analysis of facial expressions. In: Face Recognition: I-Tech Education and Publishing:377-416. [doi: 10.5772/4847]
15. Due Trier Ø, Jain AK, Taxt T. Feature extraction methods for character recognition—a survey. Pattern Recognit DAGM 1996 Apr;29(4):641-662. [doi: 10.1016/0031-3203(95)00118-2]
16. Weimer D, Scholz-Reiter B, Shpitalni M. Design of deep convolutional neural network architectures for automated feature extraction in industrial inspection. CIRP Annals 2016;65(1):417-420. [doi: 10.1016/j.cirp.2016.04.072]
17. Chen Y, Jiang H, Li C, Jia X, Ghamisi P. Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. IEEE Trans Geosci Remote Sensing 2016;54(10):6232-6251. [doi: 10.1109/TGRS.2016.2584107]
18. Jain N, Kumar S, Kumar A, Shamsolmoali P, Zareapoor M. Hybrid deep neural networks for face emotion recognition. Pattern Recognit Lett 2018 Nov;115:101-106. [doi: 10.1016/j.patrec.2018.04.010]

XSL•FO
RenderX

19. Khan AR. Facial emotion recognition using conventional machine learning and deep learning methods: current achievements, analysis and remaining challenges. Information 2022;13(6):268. [doi: 10.3390/info13060268]
20. Ebrahimi Kahou S, Michalski V, Konda K, Memisevic R, Pal C. Recurrent neural networks for emotion recognition in video. Presented at: ICMI '15: Proceedings of the 2015 ACM on International Conference on Multimodal Interaction; Nov 9, 2015; Seattle, Washington, United States. [doi: 10.1145/2818346.2830596]
21. Al-Saffar AAM, Tao H, Talab MA. Review of deep convolution neural network in image classification. Presented at: 2017 International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications (ICRAMET); Oct 23-24, 2017; Jakarta. [doi: 10.1109/ICRAMET.2017.8253139]
22. Busso C, Deng Z, Yildirim S, et al. Analysis of emotion recognition using facial expressions, speech and multimodal information. Presented at: ICMI '04: Proceedings of the 6th International Conference on Multimodal interfaces; Oct 13, 2004; State College, PA, United States. [doi: 10.1145/1027933.1027968]
23. Pourmirzaei M, Montazer GA, Esmaili F. Using self-supervised auxiliary tasks to improve fine-grained facial representation. arXiv. Preprint posted online on Aug 8, 2022. [doi: 10.48550/arXiv.2105.06421]
24. Nonis F, Dagnes N, Marcolin F, Vezzetti E. 3D approaches and challenges in facial expression recognition algorithms—a literature review. Appl Sci (Basel) 2019;9(18):3904. [doi: 10.3390/app9183904]
25. Fang T, Zhao X, Ocegueda O, Shah SK, Kakadiaris IA. 3D facial expression recognition: a perspective on promises and challenges. Presented at: 2011 IEEE International Conference on Automatic Face & Gesture Recognition (FG); Mar 21-25, 2011; Santa Barbara, CA, United States. [doi: 10.1109/FG.2011.5771466]
26. Sandbach G, Zafeiriou S, Pantic M, Rueckert D. Recognition of 3D facial expression dynamics. Image Vis Comput 2012 Oct;30(10):762-773. [doi: 10.1016/j.imavis.2012.01.006]
27. Wang W, Xu K, Niu H, Miao X. Emotion recognition of students based on facial expressions in online education based on the perspective of computer simulation. Complexity 2020 Sep 11;2020(1):1-9. [doi: 10.1155/2020/4065207]
28. Liu M, Pang H. Design and experimentation of face recognition technology applied to online live class. Sens Mater 2023;35(12):4307. [doi: 10.18494/SAM4436]
29. Tomlinson CA, Brighton C, Hertberg H, et al. Differentiating instruction in response to student readiness, interest, and learning profile in academically diverse classrooms: a review of literature. J Educ Gift 2003 Dec;27(2-3):119-145. [doi: 10.1177/016235320302700203]
30. Ge H, Zhu Z, Dai Y, Wang B, Wu X. Facial expression recognition based on deep learning. Comput Methods Programs Biomed 2022 Mar;215:106621. [doi: 10.1016/j.cmpb.2022.106621] [Medline: 35164903]
31. Zhao X, Shi X, Zhang S. Facial expression recognition via deep learning. IETE Techn Rev 2015 Sep 3;32(5):347-355. [doi: 10.1080/02564602.2015.1017542]
32. Martinez B. Advances, challenges, and opportunities in automatic facial expression recognition. In: Advances in Face Detection and Facial Image Analysis: Springer International Publishing; 2016:63-100. [doi: 10.1007/978-3-319-25958-1_4]
33. Caroppo A, Leone A, Siciliano P. Comparison between deep learning models and traditional machine learning approaches for facial expression recognition in ageing adults. J Comput Sci Technol 2020 Oct;35(5):1127-1146. [doi: 10.1007/s11390-020-9665-4]
34. Takalkar M, Xu M, Wu Q, Chaczko Z. A survey: facial micro-expression recognition. Multimed Tools Appl 2018 Aug;77(15):19301-19325. [doi: 10.1007/s11042-017-5317-2]
35. Chang LC, Yeh CC, Lin HR. Integration of facial expression detection into the design and assessment of programming teaching materials. Educ Inf Technol . [doi: 10.1007/s10639-024-13126-8]
36. ten Cate O, Kusurkar RA, Williams GC. How self-determination theory can assist our understanding of the teaching and learning processes in medical education. AMEE Guide No. 59. Med Teach 2011 Dec;33(12):961-973. [doi: 10.3109/0142159X.2011.595435]

Abbreviations

CNN: convolutional neural network
SVM: support vector machine

Edited by A Kushniruk; submitted 19.02.25; peer-reviewed by C Fang, T Cao; accepted 23.03.25; published 22.05.25.

Please cite as:
Zhu X, Juanatas RA
Auxiliary Teaching and Student Evaluation Methods Based on Facial Expression Recognition in Medical Education
JMIR Hum Factors 2025;12:e72838
URL: https://humanfactors.jmir.org/2025/1/e72838
doi:10.2196/72838

XSL•FO
RenderX

© Xueling Zhu, Roben A Juanatas. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 22.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Viewpoint

Human-Centered Design and Digital Transformation of Mental Health Services

William Fleming[1], PhD; Adam Coutts[2], PhD; Diane Pochard[3], MBA, MPA; Daksha Trivedi[4], PhD; Kristy Sanderson[5], PhD

[1]Wellbeing Research Centre, University of Oxford, Oxford, United Kingdom

[2]NIHR Applied Research Collaboration East of England, University of Cambridge, Cambridge, United Kingdom

[3]Centre for Science and Policy, University of Cambridge, Cambridge, United Kingdom

[4]Centre for Research in Public Health and Community Care, School of Health, Medicine and Life Sciences, University of Hertfordshire, Hatfield, United Kingdom

[5]NIHR Applied Research Collaboration East of England, University of East Anglia, Norwich, United Kingdom

Corresponding Author:
William Fleming, PhD
Wellbeing Research Centre
University of Oxford
Harris Manchester College
Mansfield Road
Oxford, OX1 3TD
United Kingdom
Phone: 44 01865271006
Email: william.fleming@hmc.ox.ac.uk

Abstract

Mental health services face a multitude of challenges, such as increasing demand, underfunding, and limited workforce capacity. The accelerated digital transformation of public services is positioned by government, the private sector, and some academic researchers as the solution. Alongside this, human-centered design has emerged as a guiding paradigm for this transformation to ensure user needs are met. We define what digital transformation and human-centered design are, how they are implemented in the UK policy context, and their role within the evolving delivery of mental health services. The involvement of one of our coauthors (DP) in the design and delivery of these policies over the past 5 years provides unique insights into the decision-making process and policy story. We review the promises, pitfalls, and ongoing challenges identified across a multidisciplinary literature. Finally, we propose future research questions and policy options to ensure that services are designed and delivered to meet the mental health needs of the population.

(JMIR Hum Factors 2025;12:e66040) doi:10.2196/66040

KEYWORDS

digital mental health; digital transformation; human-centred design; mental health policy; mental health services; patient and public involvement

Introduction

Background

Mental health services are under strain in the United Kingdom [1,2]. Demands for assessment, diagnosis, and treatment are increasing [3], while funding and infrastructure struggle to keep up and are widely seen to be in crisis [4]. Waiting lists and prolonged referral times are commonplace, with disastrous consequences for the well-being of service users [5]. Frontline organizations struggle to recruit, retain, and train staff, with the mental health of those in position already at risk [6,7]. The United Kingdom and others across Europe and the Organisation for Economic Cooperation Development countries are seeking innovative and ambitious policy solutions [8].

Digital transformation is widely touted as the answer to these challenges, promising affordability, accessibility, and rapid scalability [9,10]. In the United Kingdom, widely regarded as a leader in digital transformation across policy domains [11,12], a digital-first delivery model is now embedded in the long-term policy for the health service [4,13,14]. Special attention is placed on the expansion and reform of mental health care [15,16], with a view to supporting financial cuts. Globally, the digital delivery

XSL•FO
RenderX

of mental health services is viewed as vital to meet demand [17-20], especially for the poorest and the hardest-to-reach populations, whether culturally, geographically, or economically [21,22].

Despite national and international enthusiasm from policy constituencies and academic researchers, digital transformation remains an ambiguous concept [23] and is poorly defined on the ground. Digital transformation has been ongoing in health care for decades as various information and communication technologies (eg, emails, mobile phones, and electronic health records) were integrated into service delivery, at times through professional and patient discretion. Still, recent calls for acceleration and increased policy discourse demand a robust definition moving forward. Furthermore, despite optimism, digital transformation risks being an inequitable and ineffective process if it fails to meet the needs and requirements of service users, health care providers, and professionals [24]. In response, service-user involvement is increasingly advocated in the form of human-centered design (HCD) [17,25,26].

Outline

We begin by defining digital transformation and HCD for mental health services. We contextualize these ideas and processes in a personal view section, including a case study of the Mental Health Act (MHA) in UK policy. These policy insights draw directly from our experiences implementing HCD digital transformation in the UK Department for Health and Social Care (DHSC).

In the second half of the paper, we integrate academic knowledge on these themes, bringing researchers and practitioners together on policy and research developments. We first discuss opportunities and challenges in digital therapeutics before emphasizing the need for service-user involvement in therapeutics. We then move to digital transformation of services more broadly. We argue that HCD is essential to the digital transformation of mental health. In so doing, we warn against technosolutionism in digital transformation [17]. Finally, we summarize ongoing service challenges, avenues for future research, and policy recommendations. Overall, the paper contributes to an ongoing debate on the digital transformation of essential public services.

This review is a product of a National Institute for Health and Care Research Applied Research Collaboration East of England project, which holds knowledge exchange to be central. We brought together researchers, service delivery experts in the DHSC, the National Health Service (NHS), the Office for Health Improvement and Disparities, as well as local coordinators and frontline staff in the east of England. The review is informed by the ongoing exchange, including interviews with practitioners and a training workshop we hosted on "human-centered design for mental health services" [27]. Among frontline professionals we found enthusiasm for HCD, alongside many practical constraints to implementation. The search strategy and selection criteria are as follows: we identified relevant articles in English through structured searches on Web of Science and unstructured searches on Google Scholar using key terms, such as "mental health," "human-centered design," "co-production," "digital," and "services." Our disciplinary focus was journals in psychiatry and public health as well as relevant literature from public policy and design. We also searched the key terms in *The Lancet Psychiatry*, the *Journal of Medical Internet Research*, and *JMIR Mental Health*. Relevant policy gray literature is included to define the context, sourced from professional networks, and project stakeholder engagement.

Defining Digital Transformation

Digital transformation is a poorly defined concept, with multiple uses found across practitioner and academic literature. For example, the UK DHSC does not offer any definition in their policy paper "A plan for digital health and social care" [13], despite central claims that "the long-term sustainability of health and social care is dependent on having the right digital foundations in place, and so digital transformation must be the linchpin upon which all of these reforms are based." Ambiguity is repeated in government guidance for UK NHS trusts [28] and in global settings, with the World Health Organization neglecting to define digital transformation, despite its "Global strategy on digital health 2020-2025" [18].

In a systematic review of the digital transformation literature, Vial [23] offers a synthesized definition: "a process that aims to improve an entity by triggering significant changes to its properties through combinations of information, computing, communication, and connectivity technologies." In the context of increased discourse around the digitalization of mental health and the need to understand implementation and outcomes, a domain-specific definition of digital transformation is required. We adapt and define mental health digital transformation as "a process that aims to improve the provision of mental health services by triggering significant changes to its administration, coordination, treatments and therapeutics, through combinations of information, computing, communication, and connectivity technologies."

Defining HCD

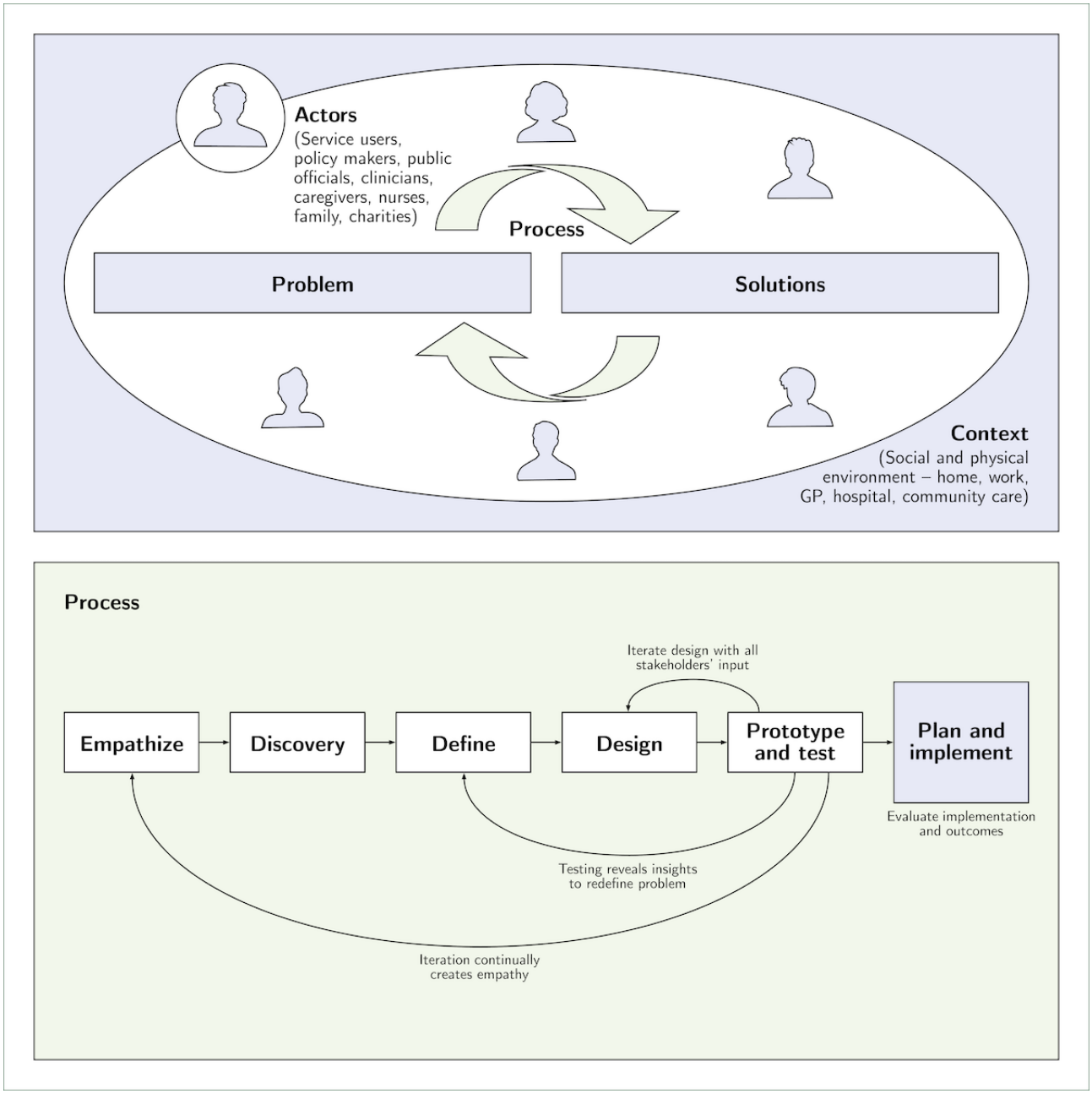
The concept of HCD originated in the fields of computer science, engineering, and ergonomics. The technical origin is captured in the International Organization for Standardization definition of HCD [29], but the emphasis in early technology-oriented definitions centers the relationship between individual users and physical objects. The conceptual scope of HCD expanded as the social potential was realized: "today's human centred design is based on the use of techniques which communicate, interact, empathise and stimulate people involved, obtaining an understanding of their needs, desires and experiences" [30]. HCD principles can guide value and well-being creation in the delivery of products, systems, and services. We define HCD in mental health services and research as "a practical, iterative approach to the design, development and reform of mental health systems, services and products that is achieved through communication, interaction and empathy with users' needs, desires and experiences."

Figure 1 summarizes the HCD approach. HCD is applied in an iterative process and includes 4 dimensions: the problems and coevolving solutions (*objects*), social and physical environment

XSL•FO
RenderX

(*context*), structure and dynamic of the activity (*process*), and those involved (*actors*) [30,31]. In the process, HCD deploys a range of qualitative techniques, such as service logic modeling, semistructured interviews, focus groups, and ethnographic research. This qualitative view creates *empathy* in the *discovery* phase to *define* the issue, including identifying user needs, the interdependencies and interactions between different user groups, and how users navigate and interact with services and products. The aim is to understand the overall context and provide a map of pathways through which a service or product user moves. The new qualitative understanding then guides the *design* of the product and service to be *tested* iteratively in a prototype phase, continuing to redefine the issue and provide new insight. Final *implementation* is enacted while quantitatively evaluating processes and outcomes.

Figure 1. Human-centered design process for mental health services. GP: general practitioner.



Accompanying the expansion of HCD is increasing conceptual confusion, with related terms such as user-centered design, user experience design, stakeholder-centered design, design thinking, and participatory design [17,31-33]. HCD is the most theoretically developed and commonly applied in policy. We propose HCD can be used as an umbrella term for iterative end user involvement in product and service delivery that emphasizes empathy, discovery, and testing. Synonymous terms can be understood as types of HCD, deployed in different contexts with varying aims.

Coproduction is a crucial related idea that dominates discussion of user involvement in public services. Coproduction is a broader umbrella concept that also entails citizen and user involvement to varying degrees [34-37]. In contrast to HCD, coproduction emphasizes sharing power (decision-making) and collaboratively delivering a service or product. With HCD,

XSL•FO
RenderX

decision-making and delivery remain with the service or product provider while strongly involving user input in the iterative process outlined above. HCD is a lighter approach to citizen involvement. Figure 1 visualizes the HCD process and demonstrates the underlying principles, which are more important than semantic debates [37,38].

Personal View: Untold Stories From the United Kingdom

Overview

In the United Kingdom health care system, digital transformation methodologies were formally enacted with the creation of "NHSX" [39]. However, the origins and design of the "whole of government" approach were first articulated by the review by Lane Fox [40]. The report emphasized the need for a user-centered digital transformation and a policy culture that is "putting the needs of citizens ahead of those of departments" [40].

The Government Digital Service (GDS) unit within the Cabinet Office was created with the mandate to digitally transform the whole of government and public services more broadly, while introducing new ways of working, such as agile delivery and HCD [41]. While GDS standards, resources, and practices offered a baseline, fragmented governance, dispersed responsibilities, and split accountability among multiple agencies in the health care system often led to inefficient transformation and processes losing sight of the end user.

NHSX was embedded with the mandate of user-centered design, with "X" standing for "user experience." NHSX was the convening body between DHSC (which designs policy), NHS England (which currently designs policy implementation but is soon to be subsumed in DHSC), and NHS Digital (which delivers technology), alongside arm's-length bodies of the NHS [42]. In the June 2021 health ministerial transition, an independent review was conducted to assess progress in digital transformation, how to build on what had already been delivered under NHSX's mandate, and to inform future structuring and investment decisions for NHSX as well as its consequences for other NHS institutions [43]. Therefore, NHSX was dissolved, and a new transformation directorate was created, hosted by NHS England.

Digitizing the MHA

An example of an "untold story" [44] on an NHSX project is the application of HCD and digital transformation to reforming the MHA. The HCD discovery phase was conducted internally in the UK government and until now has not featured publicly in academic or policy debate.

The MHA is the main legislation covering the assessment, treatment, and rights of people with mental health disorders. As the review of the MHA was being commissioned, the intention was to "make sure those suffering from mental health issues are treated with dignity and respect, with their liberty and autonomy respected" [45,46]. MHA reform began following recommendations arising from another independent review, "Modernising the Mental Health Act" [47]. The recommendations, accepted by the DHSC [48], were to modernize the act, digitize processes, and design a more human-centered system, ensuring "a person's wishes and preferences will have much more impact and be harder to ignore than before" (47: 186).

DHSC commissioned and delivered a series of discoveries on the act. Discovery is the first qualitative stage of exploratory work, outlined earlier, to identify problems and user needs. One area of focus was advance choice documents (ACDs) [49], which allow a person to set out preferences on future care and treatment, especially in cases of detention. This exploratory work examined the various service pathways people experience the application of the MHA, such as the psychological state that people are in at engagement and where "pinch points" are for service users and staff. Through HCD methods, different users of the ACD system were identified [49]. Discovery primarily involved semistructured interviews with these users, including service users or those at risk of detention, their support network, approved mental health professionals, responsible clinicians, charge nurses, and care coordinators. Exploration also included questionnaires, secondary research, and further focus groups with mental health professionals. The findings provided several recommendations and insights for all types of users.

From the user research conducted, we learned that in moments of acute crisis, there is often no time to meaningfully incorporate patient preferences. People with lived experience described feeling that the most distressing part of being detained is the absence of agency in decision-making. They expressed the need to feel listened to, treated with respect, and "not just a number." From professionals' perspective, user involvement emphasized the importance of understanding why a patient would or would not seek specific treatments. Approved mental health professionals we spoke to also expressed concerns about the practical limitations they face in accessing up-to-date medical information and about the crucial challenge of ensuring ACDs reflect the most up-to-date picture of a patient's needs and circumstances. Nurses also shared concerns about how staff-to-patient ratios directly affect the ability to engage in person-centered care.

Further noted challenges were the lack of information interoperability between care coordinators' systems and external health care professionals. Overall, the process was felt to have created empathy and illuminated implementation barriers as well as specific challenges and contexts from different users' perspectives. This HCD process emphasized the crucial role of human relationships in improving care, rather than specific needs for digital services and tools. The follow-up recommendations from the study were to test prototypes for ACDs through a pilot program designed to understand uptake, ease-of-use, and effectiveness across a range of roles and to identify challenges and improvements. Unfortunately, government prioritization decisions meant that recommendations were not put into action.

XSL•FO
RenderX

Digital Mental Health Research

Digital Therapeutics

In most discussions of digital mental health, the emphasis is on "apps, social media, chatbots, and virtual reality" [50]. Digital interventions like these are seen as the solution for growing demands for mental health support. Benefits appear to be affordability, accessibility, quick response times, potential integration of resources, and creation of large amounts of data [26,50-53]. The largest benefits are expected for young people who are digitally native [50,51], underrepresented racial and ethnic minority groups [54-56], and low- to middle-income countries where mental health care is minimal to absent [21]. Digital interventions are viewed as especially helpful during acute personal crisis such as suicide prevention [56,57], and external crisis, such as pandemics and natural disasters [58-60]. Researchers are generally positive about the promise of these digital therapeutics [57,61,62], especially through self-help [57,63-65]. However, there are difficulties for the practical widespread adoption, with several reviewers unwilling to recommend adoption and expansion of digital interventions, judging the current evidence inadequate [66,67]. Claims that digital transformation will address health inequalities are also yet to be substantiated, with existing evidence for health care broadly indicating they can exacerbate inequalities through unequal access and outcomes [24].

There are further difficulties for digital interventions. There are a vast number of digital mental health tools, with, for example, estimates of over 15,000 mobile mental health apps [68,69]. Therefore, continued use of the apps is low, with high attrition [68,70,71]. Preventing uptake and upkeep are concerns around data transparency, lack of regulation, and inadequate links with existing services for both patients and clinicians [50,53,69,71]. There is a polarization between low-use, evidence-based practices and high-use, less substantial platforms, accelerated by overcommercialization reducing the quality of products [72]. Furthermore, transplanting digital interventions from trial settings into real-world contexts is a major hurdle [73]. Other difficulties are known too, such as little engagement from those with severe mental illness [74], supply-side technical issues, and a lack of personalization [75]. Clinician involvement in design and implementation is found to produce better effects, but this is not always practiced [76].

Some of these problems arise from limitations in underlying research, resulting in researchers, practitioners, and policy makers being unable to make clear judgments. Trials often target less complex mental illnesses with more standardized solutions, such as generalized anxiety disorders and cognitive behavioral therapy (CBT), acceptance and commitment therapy, and mindfulness [61]. Witt et al [67] state that "most studies are biased" and "it is unclear whether reductions would be clinically meaningful."

As one of the more controversial examples, chatbot "conversational agent" therapies can exemplify the difficulties. Chatbots are one artificial intelligence–driven therapy enthusiastically proposed as cheap and effective [77]. Reviewers have argued that trials are of low-quality study design overall, with small samples, high attrition, negligible effects, and a high risk of bias [78]. However, new evidence is frequently published, pointing to the benefits of generative artificial intelligence therapeutic solutions for clinical depression [79]. Further user involvement is needed for assessing efficacy [17].

While chatbots exemplify the need for user engagement where therapeutic efficacy is not assured, the same is cited for proven digital interventions. Togetherall is an illuminating example of a prominent peer support platform for mental health [80,81]. Across multiple studies, partnered researchers show that service can address gaps in provision, offer support for those on waiting lists, and provide self-management opportunities. These studies indicate positive signs but reveal an ongoing need for rigorous and independent evaluation.

HCD and the Need for User Involvement

As in policy reviews, academic commentators continually call for more user involvement in the design and implementation of digital interventions [25,26,51,82], proposing coproduction as the "standard way of working" in digital mental health research [83]. This desire is echoed by practitioners and patients as well [84], all recognizing the need for co-designed solutions for effective and ethical care [85]. Therefore, it should not be surprising that HCD emerged as a guiding paradigm alongside digital transformation. Indeed, commentators in the design literature cite the digital transformation of public services as the reason why HCD has become so popular [86,87].

Vial et al [31] review the prevalence of HCD principles in the development of digital mental health interventions. They find 22 studies mentioning HCD or specific design methods, offering some promise for the integration of these ideas into digital mental health products. Since then, other interesting case studies have been published. The best example is perhaps from Creswell et al [88]. The "digitally augmented" platform they co-designed with users [89,90] provides online training and therapist support for parents to deliver themselves in-person CBT treatment for their child's anxiety. They find this parent-led, real-world application to be as effective as regular therapist-delivered CBT and more cost-effective. The strengths of the intervention stem from their HCD approach, ensuring fit with all user needs and professional support throughout. Our examples show that digital technologies can support the improvement and expansion of mental health services, with HCD principles enhancing development to meet service-user needs. These results suggest that social interactions remain key to effective treatment, while people can be supported with digital tools in the delivery rather than replaced.

Digital Mental Health Services

Overview

Within digital mental health research and policy, there is an overwhelming focus on the delivery of therapeutic interventions. This can be seen across many of the systematic reviews [63] and policy documents on the topic. However, this limits the potential of digital transformation and hinders the stated aim of improving the effectiveness, affordability, and scalability of mental health services for all. Further uses are evident in the studies by Naslund et al [91] and Stein et al [92].

XSL•FO
RenderX

Data Collection and Sharing

Digital technologies offer opportunities for the collection, communication, and analysis of patient data. Internal electronic health records are an early example of digital transformation in health services. A more recent patient-focused avenue is patient-accessible electronic health records (PAEHRs). In recent years, health services in North America and Europe have allowed patients to view health data and clinical notes (eg, the NHS App). The aim is to improve transparency, patient engagement, patient-clinician communication, and health literacy [93,94]. Service-user involvement in the design of PAEHRs shows that different mental illnesses reveal different attitudes toward the practice. Those with more severe and chronic mental concerns decline open PAEHRs more often and are more critical of systems, and in cases of severe mental illness, records are often concealed [95]. Systems can be more equitable with HCD applied in the development and ongoing improvement of services [94,95].

Linkage and interoperability are ongoing data challenges, highlighted in our discussion of the MHA [96,97]. Data sharing is vital for the delivery of services, with overlapping policy domains of health, employment, housing, and education. In our stakeholder engagement we learned of difficulties faced by asylum-seekers and refugees in the United Kingdom who would be "bounced" between immigration, housing, health, and employment systems with minimal communication or knowledge sharing between organizations [27]. Considered adoption of HCD could develop these insights for improving information sharing.

Assessment

Assessment and diagnostics appear fruitful for the implementation of digital technologies [51]. Multiple studies propose digital screening in general practice offices to identify mental health issues [98,99], alleviating resource constraints on professionals. While promising, there have been too few attempts to seriously involve those being assessed in the design of digital screening tools. One exception is the depression monitor by Nickels et al [100], which correlated electronic pulse surveys with smartphone data to explore its applicability in place of targeted questionnaires. Another radical form of assessment is the use of smart home sensing by Tiersen et al [101] for those with dementia, finding HCD enhanced applicability and relevance.

Promotion and Primary Prevention

Mental health promotion is frequently proposed as a fruitful domain for digital transformation [10,17,102-104]. Assessment and diagnostics is a key component of prevention to develop early intervention, but promotion and prevention can take on a variety of forms. Some interventions include education (to follow), mood management, diaries, peer support, or community engagement. Some level of co-design is almost always argued in the development of prevention interventions to ensure either underrepresented voices are included [103], engagement with products is maintained [102,104], or there is trust generally between users, designers, and health professionals [102].

Education

Digital transformation and HCD can come together on a smaller scale than full system change. Education is one such area where research shows digitalization can assist and be effectively designed for information provision for either clinicians or citizens [105,106]. For example, Manning et al [107] show a digital training program for nurses can improve professional effectiveness when co-designed by young people who have self-harmed. Digital learning can also benefit communities through mental health literacy, especially in contexts with previous low levels [108].

Rehabilitation

Finally, digital technologies may help with the rehabilitation process. Engdahl et al [109] show that their co-designed, digital return-to-work scheme is an effective way of supporting those with common mental illnesses. "mWorks" operates as a supplementary self-management tool for existing return-to-work services to prevent relapse. The participatory aspects were central to the success of the platform. Such an initiative shows that successful digital mental health tools may supplement rather than replace core services.

"The Solution May or May Not Be Digital": HCD in Mental Health Services

We have highlighted areas that indicate benefits from HCD in digital mental health. However, by only considering HCD principles in the design of digital therapeutic tools, researchers, practitioners, and policy makers would fail to embody the ideals of evolving services to meet citizen needs. HCD promotes solutions tailored to user needs discovered through qualitative engagement. It may be that appropriate solutions are not digital but instead other service reforms or improved financial, staffing, and technical resources. An implicit "technosolutionism" [17] in discussions of HCD in psychiatry can be seen by the emphasis placed on using it to enhance acceptability of these technologies [110]. Similarly, assuming digital transformation will aid co-creation methods risks ignoring that they may instead be used to bypass democratic oversight and citizen involvement [111]. These contradictions demand a more thorough integration of HCD principles into the design of mental health services as a whole.

HCD appears beneficial for the reform of health services—digital or not digital [87,112]—while lessons can be learned from the previous application of HCD in the public sector outside of mental health. HCD and coproduction are proposed by public policy scholars as ways to create public value in services, enhance trust in the face of democratic deficit, create active citizens and communities, and enhance the legitimacy and communication of science [35,36]. Bason and Austin [113] review various service and policy redesigns in areas such as family services, housing, and education, finding multiple advantages in HCD projects, including the ability to identify previously unknown barriers and iteratively improve services. They suggest HCD can produce a model of public governance that is more relational, networked, and reflexive. Several commentators outline these same empowering benefits

XSL•FO
RenderX

in health and care services but cite several challenges for embedding HCD as best practice [114,115].

The value of HCD for service delivery is demonstrated beyond academic research teams. One case highlighted through stakeholder engagement was conducted by the mental health charity, Mind. They co-designed perinatal support for Muslim women in the east of England [116], a population that is typically considered difficult to reach. Mind also publicly detailing their commitments to human-centered service design and offering guidance for others who are interested in these methods [117]. The charity Mind shows that HCD is relevant for third-sector actors as well as for public policy, academics, or commercial developers.

Ongoing Challenges: Directions for Research and Policy

Research

Digital transformation and HCD ideals can offer great potential for expanding and improving mental health services. However, we see several ongoing challenges for researchers, practitioners, policy makers, and service users.

First, many of the ideas within HCD principles are not new to health care delivery, with elements found in, for instance, “no decision about me, without me” [118], patient and public involvement, triangle of care, and person-centered care. HCD brings with it structured and working methodologies as well as more expansive considerations, but progress from previous citizen involvement agendas must not be forgotten or lost amid calls for the latest trending policy paradigm.

As discussed, technosolutionism can limit the impact of HCD if digital transformation is assumed to be the required resource for meeting needs. Next, a common concern in HCD and coproduction is continued power asymmetries [119,120]. For many, “the gap between principles and practice is challenging” [121], with dangers of practitioners, researchers, and designers acting “as if” power boundaries are blurred [122], when in reality the resources of current and future users of service are not enhanced [123]. Relatedly, service-user involvement must cover underrepresented populations [124]. This is most likely to be those with more severe mental conditions, exposed to multiple deprivations, and from discriminated groups. Difficulties, including the hardest-to-reach populations, emphasize that HCD will not guarantee success in the delivery of mental health services.

In addition, there are tensions between medical models of health care and the person, patient, or citizen emphasis of HCD. Traditional models of health care assume that expertise lies with professionals who have extensive training and clinical experience. Indeed, this is especially relevant in severe cases of mental illness; for example, following our discussion of the MHA, a person may be detained against their wishes. Mental health care generally requires classification and standardized protocols such as the MHA. In contrast, HCD assumes that the user is the expert [125], and in the case of mental health, where the experiences are intensely subjective and often context-specific, it poses domain-specific difficulties for developing universally or nationally appropriate services. This epistemic and political incompatibility is a common challenge elsewhere in coproduction, leading to the ideal of coproduction being undermined, proving tokenistic or “pretend” [122,126,127]. In HCD, where the ultimate delivery remains with the service provider, these limitations are less crucial. Service provision can be designed with strong involvement of service users, without appearing to dispense with professional health expertise. Nevertheless, as with all involvement practices, a central challenge remains which is achieving successful, meaningful engagement with service users [126,128,129].

To evaluate these challenges, HCD digital transformation must embed implementation evaluation and success metrics from the start and include user involvement in this design process [31,68,130]. This includes developing theories of change [81,131], identifying process mechanisms, and safeguarding those who may be negatively impacted by interventions [67].

Finally, thorough HCD can be time intensive, requires staffing commitments and funding, and can be difficult to standardize [131]. New policy structures, regulatory and commissioning practices, and funding are needed if HCD is to be a long-term solution [115,131], especially when facing the need for cost-effectiveness of digital solutions. Recent regulation of digital mental health technology from government agencies is a welcome protection against these latter harms [132]. In our own stakeholder engagement, the cost and time for co-designing were cited as limitations, but so too was the feeling that these ideas and digitalization were a top-down policy approach putting pressure on service delivery instead of providing funding. We learned of the gaps between, on one hand, those designing policy plans for digital transformation and ideals of HCD, and on the other, the realities of underresourced services. The future research agenda is given in Textbox 1.

XSL•FO
RenderX

Textbox 1. Future research agenda.

- How effective is large-scale digital transformation for meeting service-user needs and health outcomes?
- What are the barriers to engagement in human-centered design (HCD) and coproduction processes?
- How can user involvement be better integrated into data sharing, linking, and interoperability systems?
- How are government digital transformation policy mandates received by practitioners in health service delivery and local authorities? How are HCD ideas translated and received across levels of governance?
- How does prospective regulation of artificial intelligence and other emergent technologies link to digital mental health?
- What is the political economy of digital transformation and HCD in mental health services, including political and economic incentives at multiple levels, funding arrangements, and private capital involvement?
- Where do "policy ideas" on mental health emerge from?

Policy Recommendations

Overview

Our review of the literature reveals many opportunities, but also challenges, for the digital transformation of mental health services, and especially the contradictory promise of HCD. We urge caution on technosolutionism, with increased institutional support and attention to the tensions between technocratic and HCD models as well as implementation and evaluation.

Similarly, insights from various NHSX projects, in the personal view section, reveal systemic issues that hinder the design and provision of HCD: the absence of agency for patients, staff shortages, operational pressures that prioritize risk-based approaches, fractured data systems, lack of systems' interoperability, and outdated data and digital infrastructure. These realities highlight the need for structural, policy, and capability reforms that go beyond individual behavior change to reshape decision-making systems.

The recent political decision to abolish NHS England marks a significant shift in UK health care, offering a unique opportunity to redesign the system for greater efficiency and responsiveness. The elimination of duplicative governance structures and clearer decision-making pathways bring modernization possibilities for operations and realignment with user needs.

Therefore, in conclusion, we advance several policy recommendations for mental health in the United Kingdom, developed from our review, stakeholder engagement, and professional experience of digital transformation and HCD.

Independent and Cross-Departmental Digital Transformation Commission

The digital transformation directorate currently within NHS England should be externalized to establish an autonomous unit tasked with leading digital transformation across the health care ecosystem. A common strength of GDS and NHSX lies in their external positioning vis-à-vis "existing routines" and political constraints [11]. An external directorate would provide more effective convening powers, enabling better collaboration and system-wide alignment of incentives, efforts, and decisions across national and local public institutions and arm's-length bodies.

Agreed Definitions of Methodologies and Outcomes

In the course of this paper, we have developed definitions of digital transformation and HCD with respect to mental health services, and these should be incorporated into the digital transformation directorate. We hope and intend for these to be used in future policy work. Further research is needed to agree on the criteria for successful transformation, standardized processes, cost guidelines, capability assessment and needs guidance, and impact evaluation methods.

Digital Standards Collaboratively Defined

Similarly to the first recommendation, HCD approaches should be leveraged to develop data standards and operability with collaboration across systems for common terminologies, protocols, and exchange standards, as recommended by DHSC [133] and NHSX [134]. Standards and operability must function across policy domains and government departments, linking NHS mental health services with other government departments.

Integrate HCD in Other Public Policy Domains and at Both Local and National Levels in Health Services

Health care is viewed as the most promising and important avenue for the adoption of HCD principles [112,135]. Still, HCD and coproduction have been shown to be beneficial in an array of policy settings [36,113]. To facilitate the success of HCD in mental health, extending the application of HCD across public services and at multiple policy levels will normalize and strengthen the relevant social and operational systems required for success. Widespread use of HCD can bring standardization and capacity building and connect it to other models of involvement to ensure meaningful service-user engagement.

Conflicts of Interest

None declared.

References

XSL•FO
RenderX

1. Cardoso F, McHayle Z. The economics and social costs of mental ill health: review of methodology and update of calculations. Centre for Mental Health. 2024. URL: https://www.centreformentalhealth.org.uk/wp-content/uploads/2024/03/CentreforMH_TheEconomicSocialCostsofMentalIllHealth.pdf [accessed 2025-05-29]
2. Dyer C. NHS will continue to fail people with mental health conditions without urgent reforms, warns ombudsman. BMJ 2024 Feb 01;384:q272. [doi: 10.1136/bmj.q272] [Medline: 38302145]
3. Baker C, Kirk-Wade E. Mental health statistics: prevalence, services and funding in England (Report No: CBP-06988). House of Commons Library. 2024 Mar 01. URL: https://researchbriefings.files.parliament.uk/documents/SN06988/SN06988.pdf [accessed 2025-05-29]
4. Health and Social Care Committee. Evaluation of government commitments made on the digitisation of the NHS: fourth special report of session 2022-23 (Report No: HC 780). House of Commons. 2023 Feb 07. URL: https://committees.parliament.uk/publications/33979/documents/186799/default/ [accessed 2025-05-29]
5. Committee of Public Accounts. Progress in improving NHS mental health services: sixty-fifth report of session 2022-23 (Report No 1000). House of Commons. 2023 Jul 10. URL: https://committees.parliament.uk/publications/40960/documents/199502/default/ [accessed 2025-05-29]
6. Harvey SB, Epstein RM, Glozier N, Petrie K, Strudwick J, Gayed A, et al. Mental illness and suicide among physicians. Lancet 2021 Sep 04;398(10303):920-930 [FREE Full text] [doi: 10.1016/S0140-6736(21)01596-8] [Medline: 34481571]
7. Murthy VH. Confronting health worker burnout and well-being. N Engl J Med 2022 Aug 18;387(7):577-579. [doi: 10.1056/NEJMp2207252] [Medline: 35830683]
8. Health at a glance 2023. Organisation for Economic Co-operation and Development. URL: https://www.oecd.org/en/publications/2023/11/health-at-a-glance-2023_e04f8239.html [accessed 2025-05-29]
9. Blanco C, Wall MM, Olfson M. Implications of telepsychiatry for cost, quality, and equity of mental health care. JAMA Psychiatry 2022 Dec 01;79(12):1147-1148. [doi: 10.1001/jamapsychiatry.2022.3330] [Medline: 36260304]
10. Maximising the potential of digital in mental health. NHS Confederation. 2023 Sep 07. URL: https://www.nhsconfed.org/publications/maximising-potential-digital-mental-health [accessed 2025-05-29]
11. Kattel R, Takala V. The case of the UK's government digital service: the professionalisation of a paradigmatic public digital agency. Digit Gov Res Pract 2023 Dec 27;4(4):1-15. [doi: 10.1145/3630024]
12. Mergel I. Digital service teams in government. Gov Inf Q 2019 Oct;36(4):101389. [doi: 10.1016/j.giq.2019.07.001]
13. A plan for digital health and social care. Department for Health and Social Care. URL: https://archive.ph/hL5QX [accessed 2025-05-29]
14. Fit for the future: 10 year health plan for England. National Health Service, England. 2025. URL: https://assets.publishing.service.gov.uk/media/6888a0b1a11f859994409147/fit-for-the-future-10-year-health-plan-for-england.pdf [accessed 2025-05-29]
15. Mental health and wellbeing plan: discussion paper. Department for Health and Social Care. 2022. URL: https://www.gov.uk/government/calls-for-evidence/mental-health-and-wellbeing-plan-discussion-paper-and-call-for-evidence/mental-health-and-wellbeing-plan-discussion-paper [accessed 2025-05-29]
16. Topol E. The Topol review: preparing the healthcare workforce to deliver the digital future. Department for Health and Social Care. 2019. URL: https://topol.hee.nhs.uk/ [accessed 2025-05-29]
17. Bond RR, Mulvenna MD, Potts C, O'Neill S, Ennis E, Torous J. Digital transformation of mental health services. Npj Ment Health Res 2023 Aug 22;2(1):13 [FREE Full text] [doi: 10.1038/s44184-023-00033-y] [Medline: 38609479]
18. Global strategy on digital health 2020-2025. World Health Organization. URL: https://www.who.int/docs/default-source/documents/gs4dhdaa2a9f352b0445bafbc79ca799dce4d.pdf [accessed 2025-05-29]
19. Comprehensive mental health action plan 2013-2030. World Health Organization. URL: https://iris.who.int/bitstream/handle/10665/345301/9789240031029-eng.pdf?sequence=1 [accessed 2025-05-29]
20. World mental health report: transforming mental health for all. World Health Organization. URL: https://iris.who.int/bitstream/handle/10665/356119/9789240049338-eng.pdf?sequence=1 [accessed 2025-05-29]
21. Fu Z, Burger H, Arjadi R, Bockting CL. Effectiveness of digital psychological interventions for mental health problems in low-income and middle-income countries: a systematic review and meta-analysis. Lancet Psychiatry 2020 Oct;7(10):851-864 [FREE Full text] [doi: 10.1016/S2215-0366(20)30256-X] [Medline: 32866459]
22. Naslund JA, Aschbrenner KA, Araya R, Marsch LA, Unützer J, Patel V, et al. Digital technology for treating and preventing mental disorders in low-income and middle-income countries: a narrative review of the literature. Lancet Psychiatry 2017 Jun;4(6):486-500 [FREE Full text] [doi: 10.1016/S2215-0366(17)30096-2] [Medline: 28433615]
23. Vial G. Understanding digital transformation: a review and a research agenda. J Strateg Inf Syst 2019 Jun;28(2):118-144. [doi: 10.1016/j.jsis.2019.01.003]
24. Yao R, Zhang W, Evans R, Cao G, Rui T, Shen L. Inequities in health care services caused by the adoption of digital health technologies: scoping review. J Med Internet Res 2022 Mar 21;24(3):e34144 [FREE Full text] [doi: 10.2196/34144] [Medline: 35311682]
25. Hill C, Martin JL, Thomson S, Scott-Ram N, Penfold H, Creswell C. Navigating the challenges of digital health innovation: considerations and solutions in developing online and smartphone-application-based interventions for mental health disorders. Br J Psychiatry 2017 Aug 02;211(2):65-69. [doi: 10.1192/bjp.bp.115.180372] [Medline: 28522435]

XSL•FO
RenderX

26. Hollis C, Morriss R, Martin J, Amani S, Cotton R, Denis M, et al. Technological innovations in mental healthcare: harnessing the digital revolution. Br J Psychiatry 2015 Apr;206(4):263-265 [FREE Full text] [doi: 10.1192/bjp.bp.113.142612] [Medline: 25833865]
27. Human-centred design for mental health services: workshop. NIHR Applied Research Collaboration, England. URL: https://arc-eoe.nihr.ac.uk/sites/default/files/documents/Human_Centred_Design_for_mental_health_services_workshop_report_250523-3.pdf [accessed 2025-05-29]
28. Your guide to digital transformation in the NHS. Crown Commercial Service. 2022. URL: https://assets.crowncommercial.gov.uk/wp-content/uploads/CCS-digital-transformation-in-the-NHS-guide-v2.pdf [accessed 2025-05-29]
29. ISO 9241-210:2019: ergonomics of human-system interaction: part 210: human-centred design for interactive systems. International Organization for Standardization. 2019. URL: https://www.iso.org/standard/77520.html [accessed 2025-05-29]
30. Giacomin J. What is human centred design? Des J 2015 Apr 28;17(4):606-623. [doi: 10.2752/175630614X14056185480186]
31. Vial S, Boudhraâ S, Dumont M. Human-centered design approaches in digital mental health interventions: exploratory mapping review. JMIR Ment Health 2022 Jun 07;9(6):e35591 [FREE Full text] [doi: 10.2196/35591] [Medline: 35671081]
32. Altman M, Huang TT, Breland JY. Design thinking in health care. Prev Chronic Dis 2018 Sep 27;15:E117 [FREE Full text] [doi: 10.5888/pcd15.180128] [Medline: 30264690]
33. Lyon AR, Koerner K. User-centered design for psychosocial intervention development and implementation. Clin Psychol (New York) 2016 Jun;23(2):180-200 [FREE Full text] [doi: 10.1111/cpsp.12154] [Medline: 29456295]
34. Nabatchi T, Sancino A, Sicilia M. Varieties of participation in public services: the who, when, and what of coproduction. Public Adm Rev 2017 Mar 30;77(5):766-776. [doi: 10.1111/puar.12765]
35. Osborne SP, Nasi G, Powell M. Beyond co‐production: value creation and public services. Public Adm 2021 Feb 14;99(4):641-657. [doi: 10.1111/padm.12718]
36. Osborne SP, Radnor Z, Strokosch K. Co-production and the co-creation of value in public services: a suitable case for treatment? Public Manag Rev 2016 Mar 09;18(5):639-653. [doi: 10.1080/14719037.2015.1111927]
37. Masterson D, Areskoug Josefsson K, Robert G, Nylander E, Kjellström S. Mapping definitions of co-production and co-design in health and social care: a systematic scoping review providing lessons for the future. Health Expect 2022 Jun 23;25(3):902-913 [FREE Full text] [doi: 10.1111/hex.13470] [Medline: 35322510]
38. Co-production: what it is and how to do it. Social Care Institute for Excellence. 2022. URL: https://www.scie.org.uk/co-production/what-how/ [accessed 2025-05-29]
39. NHSX: new joint organisation for digital, data and technology. Department of Health and Social Care. 2019. URL: https://archive.ph/U5Rud [accessed 2025-05-29]
40. Lane Fox M. Directgov 2010 and beyond: revolution not evolution, a report by Martha Lane Fox. Cabinet Office, Efficiency and Reform Group and Government Digital Service. 2010. URL: https://www.gov.uk/government/publications/directgov-2010-and-beyond-revolution-not-evolution-a-report-by-martha-lane-fox [accessed 2025-05-29]
41. Government digital strategy. UK Cabinet Office. 2012. URL: https://assets.publishing.service.gov.uk/media/5a7ba0bfed915d41476219cd/Government_Digital_Stratetegy_-_November_2012.pdf [accessed 2025-05-29]
42. The Department of Health and Social Care remit to NHS digital for 2018/19. Department of Health and Social Care. 2018. URL: https://assets.publishing.service.gov.uk/media/5b76b09be5274a44c0636197/DHSC_Remit_to_NHS_Digital_2018-19.pdf [accessed 2025-05-29]
43. Wade-Gerry L. Putting data, digital and tech at the heart of transforming the NHS. Department for Health and Social Care. URL: https://archive.ph/piwhl [accessed 2025-05-29]
44. Cornet VP, Toscos T, Bolchini D, Rohani Ghahari R, Ahmed R, Daley C, et al. Untold stories in user-centered design of mobile health: practical challenges and strategies learned from the design and evaluation of an app for older adults with heart failure. JMIR Mhealth Uhealth 2020 Jul 21;8(7):e17703 [FREE Full text] [doi: 10.2196/17703] [Medline: 32706745]
45. Government commits to reform the Mental Health Act. Department for Health and Social Care. 2018. URL: https://archive.ph/gmwCr [accessed 2025-05-29]
46. Press release: measures to end unequal mental health treatment kickstarted by PM. Prime Minister's Office. 2019. URL: https://archive.ph/UQjOc [accessed 2025-05-29]
47. Wessely S. Modernising the Mental Health Act: increasing choice, reducing compulsion. Department for Health and Social Care. 2018. URL: https://assets.publishing.service.gov.uk/media/5c6596a7ed915d045f37798c/Modernising_the_Mental_Health_Act_-_increasing_choice__reducing_compulsion.pdf [accessed 2025-05-29]
48. Reforming the Mental Health Act: government response. Department for Health and Social Care. 2021. URL: https://archive.ph/4f1S0 [accessed 2025-05-29]
49. Stephenson L, Gergel T, Owen G, Keene AR, Pollitt A, Wilkinson B. The future of advance decision making in the Mental Health Act. Kings College, London. 2019. URL: https://www.kcl.ac.uk/policy-institute/assets/adm-mental-health-act.pdf [accessed 2025-05-29]
50. Torous J, Bucci S, Bell IH, Kessing LV, Faurholt-Jepsen M, Whelan P, et al. The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality. World Psychiatry 2021 Oct 09;20(3):318-335 [FREE Full text] [doi: 10.1002/wps.20883] [Medline: 34505369]

XSL•FO
RenderX

51. Bucci S, Schwannauer M, Berry N. The digital revolution and its impact on mental health care. Psychol Psychother 2019 Jun 28;92(2):277-297. [doi: 10.1111/papt.12222] [Medline: 30924316]
52. Tal A, Torous J. The digital mental health revolution: opportunities and risks. Psychiatr Rehabil J 2017 Sep;40(3):263-265. [doi: 10.1037/prj0000285] [Medline: 28891658]
53. Torous J, Kiang MV, Lorme J, Onnela JP. New tools for new research in psychiatry: a scalable and customizable platform to empower data driven smartphone research. JMIR Ment Health 2016 May 05;3(2):e16 [FREE Full text] [doi: 10.2196/mental.5165] [Medline: 27150677]
54. Friis-Healy EA, Nagy GA, Kollins SH. It is time to REACT: opportunities for digital mental health apps to reduce mental health disparities in racially and ethnically minoritized groups. JMIR Ment Health 2021 Jan 26;8(1):e25456 [FREE Full text] [doi: 10.2196/25456] [Medline: 33406050]
55. Kozelka EE, Jenkins JH, Carpenter-Song E. Advancing health equity in digital mental health: lessons from medical anthropology for global mental health. JMIR Ment Health 2021 Aug 16;8(8):e28555 [FREE Full text] [doi: 10.2196/28555] [Medline: 34398788]
56. Tighe J, Shand F, Ridani R, Mackinnon A, De La Mata N, Christensen H. Ibobbly mobile health intervention for suicide prevention in Australian Indigenous youth: a pilot randomised controlled trial. BMJ Open 2017 Jan 27;7(1):e013518 [FREE Full text] [doi: 10.1136/bmjopen-2016-013518] [Medline: 28132007]
57. Torok M, Han J, Baker S, Werner-Seidler A, Wong I, Larsen ME, et al. Suicide prevention using self-guided digital interventions: a systematic review and meta-analysis of randomised controlled trials. Lancet Digit Health 2020 Jan;2(1):e25-e36 [FREE Full text] [doi: 10.1016/S2589-7500(19)30199-2] [Medline: 33328037]
58. Javakhishvili J, Makhashvili N, Winkler P, Votruba N, van Voren R. Providing immediate digital mental health interventions and psychotrauma support during political crises. Lancet Psychiatry 2023 Sep;10(9):727-732. [doi: 10.1016/S2215-0366(23)00120-7] [Medline: 37392753]
59. Stein DJ, Naslund JA, Bantjes J. COVID-19 and the global acceleration of digital psychiatry. Lancet Psychiatry 2022 Jan;9(1):8-9 [FREE Full text] [doi: 10.1016/S2215-0366(21)00474-0] [Medline: 34921798]
60. Torous J, Jän Myrick K, Rauseo-Ricupero N, Firth J. Digital mental health and COVID-19: using technology today to accelerate the curve on access and quality tomorrow. JMIR Ment Health 2020 Mar 26;7(3):e18848 [FREE Full text] [doi: 10.2196/18848] [Medline: 32213476]
61. Pauley D, Cuijpers P, Papola D, Miguel C, Karyotaki E. Two decades of digital interventions for anxiety disorders: a systematic review and meta-analysis of treatment effectiveness. Psychol Med 2023 Jan 28;53(2):567-579 [FREE Full text] [doi: 10.1017/S0033291721001999] [Medline: 34047264]
62. Philippe TJ, Sikder N, Jackson A, Koblanski ME, Liow E, Pilarinos A, et al. Digital health interventions for delivery of mental health care: systematic and comprehensive meta-review. JMIR Ment Health 2022 May 12;9(5):e35159 [FREE Full text] [doi: 10.2196/35159] [Medline: 35551058]
63. De Witte NA, Joris S, Van Assche E, Van Daele T. Technological and digital interventions for mental health and wellbeing: an overview of systematic reviews. Front Digit Health. URL: https://www.frontiersin.org/articles/10.3389/fdgth.2021.754337 [accessed 2025-05-29]
64. Firth J, Torous J, Nicholas J, Carney R, Pratap A, Rosenbaum S, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. World Psychiatry 2017 Oct 21;16(3):287-298 [FREE Full text] [doi: 10.1002/wps.20472] [Medline: 28941113]
65. Torok M, Han J, McGillivray L, Wong Q, Werner-Seidler A, O'Dea B, et al. The effect of a therapeutic smartphone application on suicidal ideation in young adults: findings from a randomized controlled trial in Australia. PLoS Med 2022 May 31;19(5):e1003978 [FREE Full text] [doi: 10.1371/journal.pmed.1003978] [Medline: 35639672]
66. Lehtimaki S, Martic J, Wahl B, Foster KT, Schwalbe N. Evidence on digital mental health interventions for adolescents and young people: systematic overview. JMIR Ment Health 2021 Apr 29;8(4):e25847 [FREE Full text] [doi: 10.2196/25847] [Medline: 33913817]
67. Witt K, Spittal MJ, Carter G, Pirkis J, Hetrick S, Currier D, et al. Effectiveness of online and mobile telephone applications ('apps') for the self-management of suicidal ideation and self-harm: a systematic review and meta-analysis. BMC Psychiatry 2017 Aug 15;17(1):297 [FREE Full text] [doi: 10.1186/s12888-017-1458-0] [Medline: 28810841]
68. Bear HA, Ayala Nunes L, DeJesus J, Liverpool S, Moltrecht B, Neelakantan L, et al. Determination of markers of successful implementation of mental health apps for young people: systematic review. J Med Internet Res 2022 Nov 09;24(11):e40347 [FREE Full text] [doi: 10.2196/40347] [Medline: 36350704]
69. Torous J, Roberts LW. Needed innovation in digital health and smartphone applications for mental health: transparency and trust. JAMA Psychiatry 2017 May 01;74(5):437-438. [doi: 10.1001/jamapsychiatry.2017.0262] [Medline: 28384700]
70. Connolly SL, Kuhn E, Possemato K, Torous J. Digital clinics and mobile technology implementation for mental health care. Curr Psychiatry Rep 2021 May 07;23(7):38 [FREE Full text] [doi: 10.1007/s11920-021-01254-8] [Medline: 33961135]
71. Lipschitz JM, Van Boxtel R, Torous J, Firth J, Lebovitz JG, Burdick KE, et al. Digital mental health interventions for depression: scoping review of user engagement. J Med Internet Res 2022 Oct 14;24(10):e39204 [FREE Full text] [doi: 10.2196/39204] [Medline: 36240001]

XSL•FO
RenderX

72. Larsen ME, Huckvale K, Nicholas J, Torous J, Birrell L, Li E, et al. Using science to sell apps: evaluation of mental health app store quality claims. NPJ Digit Med 2019 Mar 22;2(1):18 [FREE Full text] [doi: 10.1038/s41746-019-0093-1] [Medline: 31304366]
73. Mohr DC, Riper H, Schueller SM. A solution-focused research approach to achieve an implementable revolution in digital mental health. JAMA Psychiatry 2018 Feb 01;75(2):113-114. [doi: 10.1001/jamapsychiatry.2017.3838] [Medline: 29238805]
74. Aref-Adib G, McCloud T, Ross J, O'Hanlon P, Appleton V, Rowe S, et al. Factors affecting implementation of digital health interventions for people with psychosis or bipolar disorder, and their family and friends: a systematic review. Lancet Psychiatry 2019 Mar;6(3):257-266. [doi: 10.1016/S2215-0366(18)30302-X] [Medline: 30522979]
75. Borghouts J, Eikey E, Mark G, De Leon C, Schueller SM, Schneider M, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. J Med Internet Res 2021 Mar 24;23(3):e24387 [FREE Full text] [doi: 10.2196/24387] [Medline: 33759801]
76. Leung C, Pei J, Hudec K, Shams F, Munthali R, Vigo D. The effects of nonclinician guidance on effectiveness and process outcomes in digital mental health interventions: systematic review and meta-analysis. J Med Internet Res 2022 Jun 15;24(6):e36004 [FREE Full text] [doi: 10.2196/36004] [Medline: 35511463]
77. D'Alfonso S. AI in mental health. Curr Opin Psychol 2020 Dec;36:112-117. [doi: 10.1016/j.copsyc.2020.04.005] [Medline: 32604065]
78. Abd-Alrazaq AA, Rababeh A, Alajlani M, Bewick BM, Househ M. Effectiveness and safety of using chatbots to improve mental health: systematic review and meta-analysis. J Med Internet Res 2020 Jul 13;22(7):e16021 [FREE Full text] [doi: 10.2196/16021] [Medline: 32673216]
79. Heinz MV, Mackin DM, Trudeau BM, Bhattacharya S, Wang Y, Banta HA, et al. Randomized trial of a generative AI chatbot for mental health treatment. NEJM AI 2025 Mar 27;2(4):e220. [doi: 10.1056/aioa2400802]
80. Deng D, Rogers T, Naslund JA. The role of moderators in facilitating and encouraging peer-to-peer support in an online mental health community: a qualitative exploratory study. J Technol Behav Sci 2023 Feb 16;8(2):128-139 [FREE Full text] [doi: 10.1007/s41347-023-00302-9] [Medline: 36810998]
81. Thomson M, Henderson G, Rogers T, Locke B, Vines J, MacBeth A. Digital mental health and peer support: building a theory of change informed by stakeholders' perspectives. PLOS Digit Health 2024 May 30;3(5):e0000522 [FREE Full text] [doi: 10.1371/journal.pdig.0000522] [Medline: 38814923]
82. Jabir AI, Lin X, Martinengo L, Sharp G, Theng Y, Tudor Car L. Attrition in conversational agent-delivered mental health interventions: systematic review and meta-analysis. J Med Internet Res 2024 Feb 27;26:e48168 [FREE Full text] [doi: 10.2196/48168] [Medline: 38412023]
83. Strudwick G, Impey D, Torous J, Krausz RM, Wiljer D. Advancing e-mental health in Canada: report from a multistakeholder meeting. JMIR Ment Health 2020 Apr 30;7(4):e19360 [FREE Full text] [doi: 10.2196/19360] [Medline: 32330114]
84. Murphy C, Thorpe L, Trefusis H, Kousoulis A. Unlocking the potential for digital mental health technologies in the UK: a Delphi exercise. BJPsych Open 2020 Jan 28;6(1):e12 [FREE Full text] [doi: 10.1192/bjo.2019.95] [Medline: 31987060]
85. Jo E, Kouaho WJ, Schueller SM, Epstein DA. Exploring user perspectives of and ethical experiences with teletherapy apps: qualitative analysis of user reviews. JMIR Ment Health 2023 Sep 22;10:e49684 [FREE Full text] [doi: 10.2196/49684] [Medline: 37738085]
86. Scholten H, Granic I. Use of the principles of design thinking to address limitations of digital mental health interventions for youth: viewpoint. J Med Internet Res 2019 Jan 14;21(1):e11528 [FREE Full text] [doi: 10.2196/11528] [Medline: 31344671]
87. Veldmeijer L, Terlouw G, Van Os J, Van Dijk O, Van 't Veer J, Boonstra N. The involvement of service users and people with lived experience in mental health care innovation through design: systematic review. JMIR Ment Health 2023 Jul 25;10:e46590 [FREE Full text] [doi: 10.2196/46590] [Medline: 37490326]
88. Creswell C, Taylor L, Giles S, Howitt S, Radley L, Whitaker E, et al. Digitally augmented, parent-led CBT versus treatment as usual for child anxiety problems in child mental health services in England and Northern Ireland: a pragmatic, non-inferiority, clinical effectiveness and cost-effectiveness randomised controlled trial. Lancet Psychiatry 2024 Mar;11(3):193-209 [FREE Full text] [doi: 10.1016/S2215-0366(23)00429-7] [Medline: 38335987]
89. Green I, Reardon T, Button R, Williamson V, Halliday G, Hill C, et al. Increasing access to evidence-based treatment for child anxiety problems: online parent-led CBT for children identified via schools. Child Adolesc Ment Health 2023 Feb 13;28(1):42-51 [FREE Full text] [doi: 10.1111/camh.12612] [Medline: 36514819]
90. Hill C, Reardon T, Taylor L, Creswell C. Online Support and Intervention for Child Anxiety (OSI): development and usability testing. JMIR Form Res 2022 Apr 13;6(4):e29846 [FREE Full text] [doi: 10.2196/29846] [Medline: 35416781]
91. Naslund JA, Marsch LA, McHugo GJ, Bartels SJ. Emerging mHealth and eHealth interventions for serious mental illness: a review of the literature. J Ment Health 2015 May 28;24(5):321-332 [FREE Full text] [doi: 10.3109/09638237.2015.1019054] [Medline: 26017625]
92. Stein DJ, Fineberg NA, Chamerlain SR. Mental Health in a Digital World: A volume in Global Mental Health in Practice. New York, NY: Academic Press; 2021.

XSL•FO
RenderX

93. Durocher K, Shin HD, Lo B, Chen S, Ma C, Strudwick G. Understanding the role of patient portals in fostering interprofessional collaboration within mental health care settings: mixed methods study. JMIR Hum Factors 2023 Jul 19;10:e44747 [FREE Full text] [doi: 10.2196/44747] [Medline: 37467024]
94. Schwarz J, Bärkås A, Blease C, Collins L, Hägglund M, Markham S, et al. Sharing clinical notes and electronic health records with people affected by mental health conditions: scoping review. JMIR Ment Health 2021 Dec 14;8(12):e34170 [FREE Full text] [doi: 10.2196/34170] [Medline: 34904956]
95. Schwarz J, Meier-Diedrich E, Neumann K, Heinze M, Eisenmann Y, Thoma S. Reasons for acceptance or rejection of online record access among patients affected by a severe mental illness: mixed methods study. JMIR Ment Health 2024 Feb 05;11:e51126 [FREE Full text] [doi: 10.2196/51126] [Medline: 38315523]
96. Campmas A, Iacob N, Simonelli F. How can interoperability stimulate the use of digital public services? An analysis of national interoperability frameworks and e-Government in the European Union. Data Policy 2022 Jun 13;4:e19. [doi: 10.1017/dap.2022.11]
97. The European Interoperability Framework in detail Internet. European Commission. 2024. URL: https://joinup.ec.europa.eu/collection/nifo-national-interoperability-framework-observatory/european-interoperability-framework-detail [accessed 2025-05-29]
98. Martel R, Shepherd M, Goodyear-Smith F. Implementing the routine use of electronic mental health screening for youth in primary care: systematic review. JMIR Ment Health 2021 Nov 19;8(11):e30479 [FREE Full text] [doi: 10.2196/30479] [Medline: 34807833]
99. Whitton AE, Hardy R, Cope K, Gieng C, Gow L, MacKinnon A, et al. Mental health screening in general practices as a means for enhancing uptake of digital mental health interventions: observational cohort study. J Med Internet Res 2021 Sep 16;23(9):e28369 [FREE Full text] [doi: 10.2196/28369] [Medline: 34528896]
100. Nickels S, Edwards MD, Poole SF, Winter D, Gronsbell J, Rozenkrants B, et al. Toward a mobile platform for real-world digital measurement of depression: user-centered design, data quality, and behavioral and clinical modeling. JMIR Ment Health 2021 Aug 10;8(8):e27589 [FREE Full text] [doi: 10.2196/27589] [Medline: 34383685]
101. Tiersen F, Batey P, Harrison MJ, Naar L, Serban A, Daniels SJ, et al. Smart home sensing and monitoring in households with dementia: user-centered design approach. JMIR Aging 2021 Aug 11;4(3):e27047 [FREE Full text] [doi: 10.2196/27047] [Medline: 34383672]
102. Löchner J, Carlbring P, Schuller B, Torous J, Sander L. Digital interventions in mental health: an overview and future perspectives. Internet Interv 2025 Jun;40:100824 [FREE Full text] [doi: 10.1016/j.invent.2025.100824] [Medline: 40330743]
103. Bergin AD, Vallejos EP, Davies EB, Daley D, Ford T, Harold G, et al. Preventive digital mental health interventions for children and young people: a review of the design and reporting of research. NPJ Digit Med 2020 Oct 15;3(1):133 [FREE Full text] [doi: 10.1038/s41746-020-00339-7] [Medline: 33083568]
104. Smith KA, Blease C, Faurholt-Jepsen M, Firth J, Van Daele TV, Moreno C, et al. Digital mental health: challenges and next steps. BMJ Ment Health 2023 Feb;26(1):e300670 [FREE Full text] [doi: 10.1136/bmjment-2023-300670] [Medline: 37197797]
105. Grim K, Rosenberg D, Svedberg P, Schön UK. Development and usability testing of a web-based decision support for users and health professionals in psychiatric services. Psychiatr Rehabil J 2017 Sep;40(3):293-302. [doi: 10.1037/prj0000278] [Medline: 28737415]
106. Honary M, Fisher NR, McNaney R, Lobban F. A web-based intervention for relatives of people experiencing psychosis or bipolar disorder: design study using a user-centered approach. JMIR Ment Health 2018 Dec 07;5(4):e11473 [FREE Full text] [doi: 10.2196/11473] [Medline: 30530457]
107. Manning JC, Carter T, Latif A, Horsley A, Cooper J, Armstrong M, et al. 'Our care through our eyes'. Impact of a co-produced digital educational programme on nurses' knowledge, confidence and attitudes in providing care for children and young people who have self-harmed: a mixed-methods study in the UK. BMJ Open 2017 May 04;7(4):e014750 [FREE Full text] [doi: 10.1136/bmjopen-2016-014750] [Medline: 28473515]
108. Bevan Jones R, Thapar A, Rice F, Beeching H, Cichosz R, Mars B, et al. A web-based psychoeducational intervention for adolescent depression: design and development of MoodHwb. JMIR Ment Health 2018 Feb 15;5(1):e13 [FREE Full text] [doi: 10.2196/mental.8894] [Medline: 29449202]
109. Engdahl P, Svedberg P, Lexén A, Tjörnstrand C, Strid C, Bejerholm U. Co-design process of a digital return-to-work solution for people with common mental disorders: stakeholder perception study. JMIR Form Res 2023 Jan 18;7:e39422 [FREE Full text] [doi: 10.2196/39422] [Medline: 36652285]
110. Harst L, Lantzsch H, Scheibe M. Theories predicting end-user acceptance of telemedicine use: systematic review. J Med Internet Res 2019 May 21;21(5):e13117 [FREE Full text] [doi: 10.2196/13117] [Medline: 31115340]
111. Lember V, Brandsen T, Tõnurist P. The potential impacts of digital technologies on co-production and co-creation. Public Manag Rev 2019 Jun 06;21(11):1665-1686. [doi: 10.1080/14719037.2019.1619807]
112. Ku B, Lupton E. Health Design Thinking: Creating Products and Services for Better Health. Cambridge, MA: MIT Press; 2020.
113. Bason C, Austin RD. Design in the public sector: toward a human centred model of public governance. Public Manag Rev 2021 Jun 09;24(11):1727-1757. [doi: 10.1080/14719037.2021.1919186]

XSL•FO
RenderX

114. Bazzano AN, Martin J, Hicks E, Faughnan M, Murphy L. Human-centred design in global health: a scoping review of applications and contexts. PLoS One 2017 Nov 1;12(11):e0186744 [FREE Full text] [doi: 10.1371/journal.pone.0186744] [Medline: 29091935]
115. Needham C, Carr S. SCIE research briefing 31: co-production: an emerging evidence base for adult social care transformation. Social Care Institute for Excellence (SCIE). 2009. URL: https://lx.iriss.org.uk/sites/default/files/resources/briefing31.pdf [accessed 2025-05-29]
116. Perinatal mental health support for Muslim communities. Mind Cambridgeshire, Peterborough and South Lincolnshite. 2021. URL: https://www.cpslmind.org.uk/resource/starting-well-perinatal-mental-health-support-for-muslim-communities/ [accessed 2025-05-29]
117. Influence and participation toolkit. Mind. URL: https://archive.ph/QnpYS [accessed 2025-05-29]
118. Liberating the NHS: no decision about me, without me. Department of Health. 2012. URL: https://assets.publishing.service.gov.uk/media/5a7c80cde5274a2674eab180/Liberating-the-NHS-No-decision-about-me-without-me-Government-response.pdf [accessed 2025-05-29]
119. Brotherdale R, Berry K, Branitsky A, Bucci S. Co-producing digital mental health interventions: a systematic review. Digit Health 2024 Apr 25;10:20552076241239172 [FREE Full text] [doi: 10.1177/20552076241239172] [Medline: 38665886]
120. Smith H, Budworth L, Grindey C, Hague I, Hamer N, Kislov R, et al. Co-production practice and future research priorities in United Kingdom-funded applied health research: a scoping review. Health Res Policy Syst 2022 Apr 02;20(1):36 [FREE Full text] [doi: 10.1186/s12961-022-00838-x] [Medline: 35366898]
121. Farr M, Davies P, Andrews H, Bagnall D, Brangan E, Davies R. Co-producing knowledge in health and social care research: reflections on the challenges and ways to enable more equal relationships. Humanit Soc Sci Commun 2021 May 06;8(1):e510. [doi: 10.1057/S41599-021-00782-1]
122. Kirkegaard S, Andersen D. Co-production in community mental health services: blurred boundaries or a game of pretend? Sociol Health Illn 2018 Jun 25;40(5):828-842. [doi: 10.1111/1467-9566.12722] [Medline: 29479721]
123. Needham C. Realising the potential of co-production: negotiating improvements in public services. Social Policy & Society 2008 Apr 01;7(2):221-231. [doi: 10.1017/s1474746407004174]
124. Jakobsson CE, Genovesi E, Afolayan A, Bella-Awusah T, Omobowale O, Buyanga M, et al. Co-producing research on psychosis: a scoping review on barriers, facilitators and outcomes. Int J Ment Health Syst 2023 Aug 30;17(1):25 [FREE Full text] [doi: 10.1186/s13033-023-00594-7] [Medline: 37644476]
125. Blandford A, Gibbs J, Newhouse N, Perski O, Singh A, Murray E. Seven lessons for interdisciplinary research on interactive digital health interventions. Digit Health 2018 May 03;4:2055207618770325 [FREE Full text] [doi: 10.1177/2055207618770325] [Medline: 29942629]
126. Williams O, Sarre S, Papoulias SC, Knowles S, Robert G, Beresford P, et al. Lost in the shadows: reflections on the dark side of co-production. Health Res Policy Syst 2020 May 07;18(1):43 [FREE Full text] [doi: 10.1186/s12961-020-00558-0] [Medline: 32380998]
127. Rose D, Kalathil J. Power, privilege and knowledge: the untenable promise of co-production in mental "health". Front Sociol 2019;4:57 [FREE Full text] [doi: 10.3389/fsoc.2019.00057] [Medline: 33869380]
128. Steen T, Brandsen T, Verschuere B. The dark side of co-creation and co-production. In: Brandsen T, Verschuere B, Steen T, editors. Co-Production and Co-Creation: Engaging Citizens in Public Services. London, UK: Routledge; 2018:284-292.
129. Ocloo J, Matthews R. From tokenism to empowerment: progressing patient and public involvement in healthcare improvement. BMJ Qual Saf 2016 Aug 18;25(8):626-632 [FREE Full text] [doi: 10.1136/bmjqs-2015-004839] [Medline: 26993640]
130. Lipschitz J, Hogan TP, Bauer MS, Mohr DC. Closing the research-to-practice gap in digital psychiatry. J Clin Psychiatry 2019 May 14;80(3):12659. [doi: 10.4088/jcp.18com12659]
131. Cross S, Nicholas J, Mangelsdorf S, Valentine L, Baker S, McGorry P, et al. Developing a theory of change for a digital youth mental health service (moderated online social therapy): mixed methods knowledge synthesis study. JMIR Form Res 2023 Nov 03;7:e49846 [FREE Full text] [doi: 10.2196/49846] [Medline: 37921858]
132. Digital mental health technology - regulation and evaluation for safe and effective products. Medicines & Healthcare products Regulatory Agency. URL: https://assets.publishing.service.gov.uk/media/6866572fadfe29730ea3a9d5/MHRA_guidance_on_DMHT_-_Device_characterisation_regulatory_qualification_and_classification.pdf [accessed 2025-05-29]
133. Data saves lives: reshaping health and social care with data. Department for Health and Social Care. URL: https://archive.ph/aaj35 [accessed 2025-05-29]
134. Bolychevsky I. An introduction to standards and interoperability at NHSX. NHS England. URL: https://archive.ph/zzwIF [accessed 2025-05-29]
135. Veldmeijer L, Terlouw G, Veer, van Os J, Boonstra N. Design for mental health: can design promote human-centred diagnostics? Design Health 2023 Feb 14;7(1):5-23. [doi: 10.1080/24735132.2023.2171223]

XSL•FO
RenderX

Abbreviations

ACD: advance choice document
CBT: cognitive behavioral therapy
DHSC: Department for Health and Social Care
GDS: Government Digital Service
HCD: human-centered design
MHA: Mental Health Act
NHS: National Health Service
PAEHR: patient-accessible electronic health record

Edited by A Kushniruk; submitted 02.09.24; peer-reviewed by G Terlouw, C Kealy; comments to author 09.04.25; revised version received 22.05.25; accepted 09.07.25; published 11.08.25.

Please cite as:
Fleming W, Coutts A, Pochard D, Trivedi D, Sanderson K
Human-Centered Design and Digital Transformation of Mental Health Services
JMIR Hum Factors 2025;12:e66040
URL: https://humanfactors.jmir.org/2025/1/e66040
doi:10.2196/66040
PMID:

©William Fleming, Adam Coutts, Diane Pochard, Daksha Trivedi, Kristy Sanderson. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 11.08.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Device Failures and Adverse Events Associated With Rhinolaryngoscopes: Analysis of the Manufacturer and User Facility Device Experience (MAUDE) Database

Shao-Hsuan Chang[1*], PhD; Daishi Chen[2*], PhD, MD; Chi-Sheng Chen[1], MEng; Dong Zhou[1], PhD; Lung-Kun Yeh[3,4], MD, PhD

[1]Department of Biomedical Engineering, Chang Gung University, Taoyuan, Taiwan

[2]Department of Otolaryngology, Shenzhen People's Hospital, Shenzhen, China

[3]Department of Ophthalmology, Linkou Chang Gung Memorial Hospital, No. 5, Fuxing St, Guishan Dist, Taoyuan, Taiwan

[4]College of Medicine, Chang Gung University, Taoyuan, Taiwan

[*]these authors contributed equally

Corresponding Author:
Lung-Kun Yeh, MD, PhD
Department of Ophthalmology, Linkou Chang Gung Memorial Hospital, No. 5, Fuxing St, Guishan Dist, Taoyuan, Taiwan

Abstract

Background: Rhinolaryngoscopes are one of the most widely used tools by otolaryngologists and speech-language pathologists in current clinical practice. However, there is limited data on adverse events associated with or caused by the use of rhinolaryngoscopes.

Objective: In this study, we used the Manufacturer and User Facility Device Experience (MAUDE) database with the aim of providing insights that may assist otolaryngologists in better understanding the limitations of these devices and selecting appropriate procedures for their specific clinical setting.

Methods: We characterized complications associated with the postmarket use of rhinolaryngoscope devices from the US Food and Drug Administration MAUDE database from 2016 through 2023.

Results: A total of 2591 reports were identified, including 2534 device malfunctions, 56 injuries, and 1 death, from 2016 through 2023. The most common device problem with rhinolaryngoscopes was breakage (n=1058 reports, 40.8%), followed by fluid leaks (n=632 reports, 24.4%). The third most common problem was poor image quality (n=467 reports, 18%). Other device issues included contamination or device reprocessing problems (n=127 reports, 4.9%), material deformation or wear (n=125 reports, 4.8%), and device detachment (n=73 reports, 2.8%). Of the 63 reported adverse events, the most common patient-related adverse event was hemorrhage or bleeding, accounting for 18 reports, with the root causes including material deformation or wear, breakage, wrinkled rubber, or improper operation.

Conclusions: Our study offers valuable insights for endoscopists and manufacturers to recognize potential issues and adverse events associated with the use of rhinolaryngoscopes. It emphasizes the need for improving device reliability, training, and procedural protocols to enhance patient safety during diagnostic procedures.

(JMIR Hum Factors 2025;12:e67036) doi:10.2196/67036

KEYWORDS

medical device; device malfunction; rhinolaryngoscope; adverse event; MAUDE; Manufacturer and User Facility Device Experience

Introduction

Rhinolaryngoscopes are currently available in various forms, including rigid, flexible, and video versions. They are essential tools routinely used by otolaryngologists and speech-language pathologists [1,2]. Given the substantial proportion of patients who require a rhinolaryngoscopic procedure as part of a head and neck examination, rhinolaryngoscopy has become an indispensable diagnostic procedure in otolaryngological practice [3]. Modern rhinolaryngoscopes can have a distal diameter as small as 2 mm, are equipped with lighting, are flexible, and have photo and video capabilities, enabling direct visualization of the nose, throat, and airway in diverse clinical settings, including emergency, inpatient, and outpatient scenarios. The range of indications for rhinolaryngoscopy includes visualizing polyps, tumors, and sources of epistaxis in the nasal cavity and aiding in the identification of suspected tumors or adenoidal hypertrophy [4]. In the oropharynx or laryngopharynx,

XSL•FO
RenderX

rhinolaryngoscopy can be instrumental in evaluating foreign bodies and potential airway obstructions from etiologies such as neoplasm and epiglottitis, obstructive sleep apnea, dysphagia, dysphonia, tonsillar hypertrophy, glossoptosis, laryngomalacia, and vocal fold lesions [5,6]. Rhinolaryngoscopy may also assist in assessing the severity of angioedema. An alternative visualization tool is the flexible fiber-optic laryngoscope [7]. While laryngoscopes may be less expensive and simpler to use than rhinolaryngoscopes, fiber-optic rhinolaryngoscopy provides clearer visualization and better access to the larynx anatomy. In general, imaging quality, including illumination, color fidelity, resolution, and accurate length representation, plays a pivotal role in visualizing abnormalities. Therefore, the choice of instrument is often determined by clinician preference or equipment capabilities [8].

Rhinolaryngoscopes are classified as a moderate-risk medical device and require a 510(k) submission for premarket review by the US Food and Drug Administration (FDA; regulation number 21 CFR 874.4760). Although rhinolaryngoscopy is considered a generally safe procedure with rare serious adverse events, complications such as mucosal tearing, damage to anatomic structures (particularly with rigid scopes), bleeding, and laryngospasm have been reported [9]. However, there is limited data regarding device failures for both rigid and flexible rhinolaryngoscopes, incidence rates, and adverse events associated with these devices. The FDA's Manufacturer and User Facility Device Experience (MAUDE) database serves as a repository for adverse events and malfunction reports related to medical devices, and it has been used as a data source to study device-related adverse events [10,11]. In addition, capturing user experiences and integrating them into the design during medical device development has become an essential component for ensuring patient safety and device effectiveness. In this study, we used the MAUDE database to characterize postmarket complications associated with the use of rhinolaryngoscope devices from 2016 through 2023, with the aim of providing insights that may assist otolaryngologists in better understanding the limitations of these devices and selecting appropriate procedures for their specific clinical setting.

Methods

Search and Selection

MAUDE, a comprehensive postmarket surveillance database from the FDA, was chosen because of its extensive collection of medical device reports (MDRs) from both mandatory and voluntary reporters related to FDA-cleared medical devices [11]. Mandatory reporters comprise manufacturers or device user facilities, whereas voluntary reporters include physicians, patients, or other device consumers. The MDRs from MAUDE include device malfunctions and potential patient harms associated with device failures during procedures. Therefore, the event classifications used in this analysis include device malfunctions and adverse events among patients.

The MAUDE database was queried by searching for the product category "rhinolaryngoscope (Flexible or Rigid)," which included MDRs from January 2016 through December 2023. This search encompassed terms such as "rhino-laryngo videoscope," "rhino-laryngo fiberscope," "telescope," and "single use endoscope." Each MDR was then logged in a Microsoft Excel spreadsheet. Among the reports, items that may have been incorrectly classified, such as cystoscopes and arthroscopes, were manually removed. These misclassifications accounted for only a small portion of the reports and resulted from errors made by the reporters. The MDRs underwent manual review, including assessment of the date of the event, event type (device malfunction, injury, or death), and the root causes of the adverse event. Device failures were defined as reported instances of rhinolaryngoscopes not functioning as expected during a procedure.

Data Extraction

The device failure problems were classified as (1) breakage, (2) fluid leakage, (3) poor image quality, (4) contamination, (5) material deformation, (6) device detachment, (7) unintended movement, (8) mechanical problems, (9) unidentified device or use problems, (10) device overheating, (11) improper or incorrect procedures, (12) moisture damage, (13) packing problems, and (14) electrical shorting. To minimize potential bias and ensure consistency in categorization, 2 reviewers (SHC and DC) independently screened and categorized the reported events. Adverse events among patients included all reported complications resulting in injury, harm, or death during procedures associated with rhinolaryngoscopes. Following the categorization of the failure problems and adverse events, statistical analysis was conducted. The frequency of each failure type and adverse event was calculated, and the results were expressed as the number and percentage of occurrences within each category.

Additionally, to assess the context of these device failures in clinical practice, we analyzed the US Medicare population using *Current Procedural Terminology* (*CPT*) codes associated with rhinolaryngoscope diagnostic procedures (excluding surgeries). Specifically, we selected *CPT* codes 31231 (nasal endoscopy, diagnostic; unilateral or bilateral), 31575 (laryngoscopy, flexible; diagnostic), and 92511 (nasopharyngoscopy with endoscope), which may correspond to diagnostic procedures involving rhinolaryngoscopes. The data were accessed under the "Research, Statistics, Data & Systems" section. The Centers for Medicare and Medicaid Services Part B National Summary Data File was used to obtain annual procedure data based on *CPT* codes 31231, 31575, and 92511. These codes were used to compare the number of procedure claims with the number of device failure reports, allowing us to explore the relationship between the frequency of rhinolaryngoscope use (as indicated by procedure billing) and the reported failure incidents. This comparison provide a picture of the relative incidence of device failures in clinical practice compared to the frequency of procedures conducted, shedding light on potential areas for improvement in device performance or clinical training.

Ethical Considerations

This study involved querying the FDA MAUDE database and US *CPT* codes. These activities did not involve direct interaction with human subjects, nor did they involve the collection or analysis of personal health data. Therefore, such data queries typically do not require approval from an institutional review

XSL•FO
RenderX

board. Publicly available databases like the FDA MAUDE and *CPT* codes generally do not raise ethical concerns. We analyzed publicly available data and did not collect personal data or perform clinical trials, so institutional review board approval was not required.

Results

Device-Related Problems

A total of 2591 reports were identified, including 2534 device malfunctions, 56 injuries, and 1 death among patients from 2016 through 2023. The number of reports submitted significantly increased in 2021, 2022, and 2023 (n=347, n=1147, and n=918 reports, respectively), constituting greater than 90% of all reports submitted during the study period. Device malfunctions and their incidence rates are presented in Table 1. The most common device problem with rhinolaryngoscopes was breakage (n=1058 reports, 40.8%), followed by fluid leakage (n=632 reports, 24.4%). The third most common problem was poor image quality (n=467 reports, 18%). Other device issues included contamination or device reprocessing (n=127 reports, 4.9%), material deformation or wear (n=125 reports, 4.8%), device detachment (n=73 reports, 2.8%), and unintended movement (n=55 reports, 2.2%), with the remaining issues each accounting for less than 1%. Mechanical problems, device overheating, packaging issues, and electrical shorting were relatively infrequent problems, as summarized in Table 2.

Table . Event types reported to the US Food and Drug Administration Manufacturer and User Facility Device Experience (MAUDE) database for rhinolaryngoscopes per year.

Year	Deaths (n=1), n (%)	Injuries (n=56), n (%)	Malfunctions (n=2534), n (%)
2023	0 (0)	14 (25)	918 (36.2)
2022	1 (100)	7 (12.5)	1147 (45.3)
2021	0 (0)	2 (3.6)	347 (13.7)
2020	0 (0)	10 (17.8)	37 (1.5)
2019	0 (0)	14 (25)	38 (1.5)
2018	0 (0)	2 (3.6)	21 (0.8)
2017	0 (0)	5 (8.9)	13 (0.5)
2016	0 (0)	2 (3.6)	13 (0.5)

Table . Device problems reported for rhinolaryngoscopes.

Device problems	Reports (n=2591), n (%)
Breakage	1058 (40.8)
Fluid leakage	632 (24.4)
Poor image quality; optical distortion	467 (18)
Contamination; device reprocessing problem	127 (4.9)
Material deformation; wear	125 (4.8)
Detachment of device	73 (2.8)
Unintended movement	55 (2.2)
Mechanical problem	12 (0.5)
Unidentified device or use problem	11 (0.4)
Overheating of device	10 (0.4)
Improper or incorrect procedure	6 (0.2)
Moisture damage	6 (0.2)
Packaging problem	5 (0.2)
Electrical shorting; flare or flash	4 (0.2)

We conducted a comprehensive analysis of device failures, examining variations across specific manufacturers, including Olympus, Storz, Pentax, and Gyrus. A high incidence of breakage and fluid leakage reports was associated with Pentax. In comparison, Olympus devices showed a breakage rate of 4.8% (50/1028). However, Olympus devices were also associated with the highest frequency of contamination reports, followed by Pentax. These higher failure rates could be linked to the relatively larger market share of these manufacturers. Additionally, we scrutinized the frequency of nasal endoscopies performed in the US Medicare population using *CPT* codes 31231, 31575, and 92511. In 2022, a total of 1,196,520

XSL•FO
RenderX

procedures were performed. Meanwhile, the MAUDE database reported 1155 device failures. This results in an estimated failure rate of approximately 1 per 1000 procedures.

Patient-Related Adverse Events

Adverse events or complications associated with device problems were evaluated based on the MAUDE database. Adverse events in patients in reports included all instances of complications resulting in injury, harm, or death associated with the use of rhinolaryngoscopes. Adverse events in patients were categorized as (1) hemorrhage/bleeding, (2) a foreign body in the patient, (3) burns or thermal burns, (4) discomfort or pain, (5) anaphylactic shock, (6) edema, (7) laryngospasm and dyspnea, (8) unspecified tissue injury, (9) allergic reaction, and (10) death. The specified root cause of the adverse event included material and mechanical problems, device chemical contamination, improper operation, fluid leakage, overheating, and display issues, among others.

The adverse events in patients and their specified root causes are presented in Table 3. A total of 63 reported adverse events were identified, including hemorrhage/bleeding (n=18 reports, 28%), a foreign body in the patient (n=12 reports, 19%), discomfort/pain (n=10 reports, 16%), thermal burns (n=9 reports, 14%), anaphylactic shock (n=4 reports, 6%), edema (n=3 reports, 5%), laryngospasm and dyspnea (n=3 reports, 5%), unspecified tissue injury (n=2 reports, 3%), allergic reaction (n=1 report, 2%) and death (n=1 report, 2%). The most common patient-related adverse event was hemorrhage/bleeding, accounting for 18 reports, with the root causes including material deformation or wear, breakage, wrinkled rubber, or improper operation. One death was reported, with an unspecified cause. Material breakage, material wear, and improper operation were identified as the major root causes for the reported adverse events. In 2019, 1 case of edema was attributed to improper reprocessing and device contamination with chemicals. The manufacturer's investigation determined that the customer had insufficiently wiped phthalaldehyde-based disinfectants off the endoscope during reprocessing, resulting in the possibility of anaphylactic shock. Table 4 displays the numbers of adverse events from 2016 through 2023, including thermal burns, foreign bodies in patients, and discomfort, reported in 2023.

Table . Adverse events in patients from rhinolaryngoscope device problems.

Adverse event	Root cause	Events (n=63), n (%)
Hemorrhage/bleeding	Material deformation or wear; breakage; rubber wrinkling; improper operation	18 (28)
Foreign body in patient	Detachment of device or device component; breakage; material fragmentation; wear; improper operation	12 (19)
Discomfort/pain	Material wear; breakage; material invagination; mechanical problems	10 (16)
Burns/thermal burns	Scratched material; cut or torn material; burst battery; device overheating; use of device problem	9 (14)
Anaphylactic shock	Material perforation; material cracks or holes	4 (6)
Edema	Unidentified device or use problem; allergic reaction due to improper reprocessing and device chemical contamination	3 (5)
Laryngospasm and dyspnea	Display problem/poor image; unspecified issue	3 (5)
Unspecified tissue injury	Fluid/blood leakage; material wear	2 (3)
Allergic reaction	Reprocessing agent	1 (2)
Death	Unspecified issue	1 (2)

XSL•FO
RenderX

Table . Number of adverse events based on year and type.

	Events, n							
	2023	2022	2021	2020	2019	2018	2017	2016
Hemorrhage/bleeding	0	0	0	2	12	0	3	1
Foreign body in patient	4	0	0	0	2	0	5	1
Discomfort/pain	5	2	0	1	2	0	0	0
Burns/thermal burns	3	1	1	1	0	1	1	1
Anaphylactic shock	0	0	0	4	0	0	0	0
Edema	0	0	0	0	1	0	2	0
Laryngospasm and dyspnea	0	0	0	0	3	0	0	0
Unspecified tissue injury	0	2	0	0	0	0	0	0
Allergic reaction	0	0	1	0	0	0	0	0
Death	0	1	0	0	0	0	0	0

Discussion

Principal Findings

This study examined adverse events associated with the use of rhinolaryngoscopes based on the MAUDE database, analyzing 2591 common device issues and 63 patient-related adverse events from January 2016 through June 2023. Breakage (40.83%) was the most common device issue, whereas hemorrhage (28.57%) was the most common patient-related adverse event, primarily attributed to material deformation or breakage. These findings underscore the overall low occurrence of adverse events associated with the use of rhinolaryngoscopes. Previous studies have highlighted mucosal tearing and bleeding as the most common complications, with laryngospasm occurring in less than 1% of procedures [12-14]. To prevent these complications, adequate nasal decongestion and limited force are recommended. However, our results show that improper operation or use problems frequently contributed as the root cause of adverse events, including hemorrhages, foreign bodies in patients, thermal burns, pain, and edema. Operator experience with the device may also influence the occurrence of adverse events. One customer reported a broken eyepiece on a rigid scope (report number 9610773-2023-03626). Following investigation, the manufacturer attributed the issue to wear and tear coupled with excessive force. Similarly, another report detailed a damaged lens at the distal end of a scope (report number 9610773-2023-02419). Upon examination, the manufacturer concluded that this damage stemmed from user error, improper handling, and the application of excessive force. Therefore, considerations of usability and ergonomics are imperative to enhance patient care [15-17].

Despite gathering 125 reports of material wear from the MAUDE database, which accounted for one-ninth of the total breakage reports, it is noteworthy that the majority of the reported failures were not attributed to routine wear and tear, as determined by the manufacturer's investigation. This observation underscores the importance of scrutinizing reported failures to distinguish between typical wear-related issues and potentially more serious underlying concerns. It is evident that device failures may arise from various root causes beyond routine wear and tear, which manufacturers may not consistently highlight in their manuals. For instance, in the manual for the TJF-Q180V (Olympus) endoscope, there is an acknowledgment that repeated use and reprocessing of the endoscope and its accessories can lead to gradual wear and tear. This underscores the importance of thorough understanding of and adherence to manufacturer instructions, as well as recognizing potential factors contributing to device failures, beyond typical wear-related issues.

Moreover, it was observed that many cases of device failure could be linked to improper reprocessing practices. In this study, 127 device failures caused by rhinolaryngoscope contamination were reported. Although manufacturers provide recommendations for product use and reprocessing, Biadsee et al [18] reported that less than 20% of physicians adhere to the recommended decontamination process outlined by manufacturers. Another study, from Jiang et al [19], of rhinolaryngoscope device failure due to contamination using the MAUDE database from 2013 to 2019 revealed associations with laryngeal edema, rather than infection, highlighting 1 injury resulting from improper reprocessing procedures (insufficient disinfectant removal from the endoscope). Anaphylaxis resulting from the use of phthalaldehyde-based disinfectants has been reported in cases involving the cleaning of rhinolaryngoscopes and cystoscopes [20,21]. In this case (report number 9610877-2019-00238), the risk of injury from microbial

XSL•FO
RenderX

infection may be less than the potential harm caused by the disinfectant solution itself. In response to this incident, 5 members of the otolaryngology department were identified and retrained as the personnel at the hospital did not properly reprocess the endoscope. Regarding the compatibility of disinfectants, the manufacturer of the endoscope recommends that facilities adhere strictly to the instructions provided by the disinfectant manufacturer, including specific parameters such as concentration, temperature, and exposure time. Furthermore, it is crucial to ensure thorough rinsing of internal channels, external endoscope surfaces, and components with clean water to remove any residual detergent solution, thereby minimizing the risk of adverse reactions or complications associated with disinfectant use. Our results identified 1 allergic adverse event associated with a reprocessing agent, and 3 cases of edema due to improper reprocessing or device use, confirming previous literature regarding rhinolaryngoscope contamination [22,23]. Compared to other endoscope types, such as bronchoscopes and duodenoscopes, the contamination rate of rhinolaryngoscopes is less commonly associated with patient harm or death [18]. This indicates a comparatively lower risk associated with rhinolaryngoscopes in contrast to other endoscopic examinations.

This study reviewed the MAUDE database and identified 9 cases of thermal burns in patients who underwent rhinolaryngoscope examinations over 8 years. Scope overheating was reported as a root cause, likely because of prolonged procedure times with the light source on or the brightness setting of the lamp. Several studies assessed the heat effects of endoscopes in otorhinolaryngology and recommended keeping light sources at the lowest effective intensity [24-26]. MacKeith et al [27] evaluated the amount of heat produced by endoscopes and showed that larger-diameter endoscopes attain a higher temperature. Chitnavis [28] demonstrated that even momentary proximity could cause a thermal burn to a patient's skin, without generating smoke or fire. The risk of thermal injury may be associated with the light source, endoscope caliber, and angulation [25,27,29,30]. Consequently, available cameras are often equipped with a regulation system capable of automatic gain control in poor lighting situations [31].

Furthermore, it is noteworthy that the number of MDRs underwent a significant surge in 2021, 2022, and 2023, with 347, 1147, and 918 reports, respectively. This notable increase underscores the importance of continued vigilance and thorough investigation into the factors driving these trends to ensure the safety and efficacy of medical devices. The growth in frequency can be attributed to several major factors, including the increased prevalence of laryngeal diseases, heightened use of laryngoscopes in airway management, guideline recommendations, elevated demand for respiratory products, and the impact of the COVID-19 pandemic. However, there was no significant rise in the number of nasal endoscopies performed in the US Medicare population with *CPT* codes 31231, 31575, and 92511 from 2016 to 2022. Another plausible explanation for the increased MDRs could stem from new product launches and strategic activities by manufacturers, which have had a moderate impact on the laryngoscope market, influencing the growth in the frequency of MDRs.

The statistical analysis conducted using the FDA's MAUDE database carries significant implications for clinical practice. These data not only help identify trends and patterns in device failures but also provide critical safety information for clinical health care. By analyzing the MAUDE data, we can identify failure types and frequencies associated with specific devices, such as rhinolaryngoscopes, thereby understanding their potential impact on patient safety. The data reveal the sources of device failures, including variations across manufacturers and models, which can aid clinical decision-makers in selecting devices based on more informed criteria. Furthermore, by examining the time trends of device failures in the MAUDE database, we can correlate these trends with medical policies or other external factors. This trend analysis can assist health care institutions in adjusting operational processes and training programs to address the growing issue of device failures, especially during times of infectious disease outbreaks or when new devices are introduced. The data derived provide empirical evidence that can help in developing more effective clinical guidelines and training plans, ultimately contributing to improved quality of health care services [32]. These findings highlight the importance of timely adjustments to clinical operational processes and training programs to address the growing issue of device failures and provide guidance for future device improvements and clinical recommendations.

Limitations

A major limitation of this study is its reliance on reports from the FDA's MAUDE database, which may not account for comorbidities involved in reported device failures. Furthermore, the information may be incomplete or limited because MDRs can be submitted by health care professionals, patients, or manufacturers. Reporting variations between private-practice clinics and large academic medical centers may also affect data accuracy and which brands are used. Therefore, establishing causative associations for some reported adverse events is challenging. However, it is important to note that manufacturers are required to investigate reported adverse events and device failures, and their findings or follow-up reports are included in the MAUDE database. The manufacturer reports help to mitigate potential bias arising from self-reported data. Since the MAUDE database primarily focuses on device-related issues, we recognize the possibility of underreporting or inconsistent reporting practices. Nonetheless, the inclusion of both user-reported and manufacturer-investigated information adds a level of rigor to the analysis, allowing for a more nuanced understanding of the issues related to device failures. Despite these limitations, our study provides important insights for endoscopists and manufacturers to recognize potential issues and adverse events associated with the use of rhinolaryngoscopes, emphasizing the importance of patient safety.

Conclusions

Although our findings underscore the overall low occurrence of adverse events associated with the use of rhinolaryngoscopes, the results indicate that improper operation frequently contributed to adverse events, including hemorrhage, foreign bodies in patients, thermal burns, pain, and edema. Operator

XSL•FO
RenderX

experience with the device may also influence the occurrence of adverse events. Results from this study will be important for endoscopists and manufacturers to have a thorough understanding of the equipment and its limitations. Future research should assess the broader organizational impact, including otolaryngology teams, documentation practices, clinician training, and patient perspectives.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' Contributions

Supervision: SHC and LKY
Conceptualization: SHC and DC
Methodology: SHC and CSC
Formal analysis: SHC and DZ
Writing—original draft preparation: SHC
Writing—review and editing: DC

Conflicts of Interest

None declared.

References

1. Nankivell PC, Pothier DD. Nasal and instrument preparation prior to rigid and flexible nasendoscopy: a systematic review. J Laryngol Otol 2008 Oct;122(10):1024-1028. [doi: 10.1017/S0022215108002533] [Medline: 18492307]
2. Rosen CA, Amin MR, Sulica L, et al. Advances in office-based diagnosis and treatment in laryngology. Laryngoscope 2009 Nov;119 Suppl 2:S185-S212. [doi: 10.1002/lary.20712] [Medline: 19856405]
3. Zacharias SRC, Weinrich B, Brehm SB, et al. Assessment of vibratory characteristics in children following airway reconstruction using flexible and rigid endoscopy and stroboscopy. JAMA Otolaryngol Head Neck Surg 2015 Oct;141(10):882-887. [doi: 10.1001/jamaoto.2015.1740] [Medline: 26402578]
4. Aremu SK. Flexible nasopharyngolaryngoscopy: evaluation and appraisal of its effectiveness and diagnostic yield, the Nigerian experience. J Family Med Prim Care 2019 Oct;8(10):3399-3403. [doi: 10.4103/jfmpc.jfmpc_489_19] [Medline: 31742176]
5. Bentsianov BL, Parhiscar A, Azer M, Har-El G. The role of fiberoptic nasopharyngoscopy in the management of the acute airway in angioneurotic edema. Laryngoscope 2000 Dec;110(12):2016-2019. [doi: 10.1097/00005537-200012000-00007] [Medline: 11129012]
6. Zeleník K, Walderová R, Kučová H, Jančatová D, Komínek P. Comparison of long-term voice outcomes after vocal fold augmentation using autologous fat injection by direct microlaryngoscopy versus office-based calcium hydroxylapatite injection. Eur Arch Otorhinolaryngol 2017 Aug;274(8):3147-3151. [doi: 10.1007/s00405-017-4600-1] [Medline: 28478500]
7. Milstein CF, Charbel S, Hicks DM, Abelson TI, Richter JE, Vaezi MF. Prevalence of laryngeal irritation signs associated with reflux in asymptomatic volunteers: impact of endoscopic technique (rigid vs. flexible laryngoscope). Laryngoscope 2005 Dec;115(12):2256-2261. [doi: 10.1097/01.mlg.0000184325.44968.b1] [Medline: 16369176]
8. Cohen SM, Pitman MJ, Noordzij JP, Courey M. Evaluation of dysphonic patients by general otolaryngologists. J Voice 2012 Nov;26(6):772-778. [doi: 10.1016/j.jvoice.2011.11.009] [Medline: 22285452]
9. Wrigley SR, Black AE, Sidhu VS. A fibreoptic laryngoscope for paediatric anaesthesia. A study to evaluate the use of the 2.2 mm Olympus (LF-P) intubating fibrescope. Anaesthesia 1995 Aug;50(8):709-712. [doi: 10.1111/j.1365-2044.1995.tb06100.x] [Medline: 7645703]
10. Wilson N, Abdallah M, Jaber F, et al. Analysis of reported adverse events associated with over-the-scope endoscopic suturing system: an FDA MAUDE database study. Obes Surg 2023 Apr;33(4):1253-1258. [doi: 10.1007/s11695-023-06525-2] [Medline: 36810809]
11. Gurtcheff SE. Introduction to the MAUDE database. Clin Obstet Gynecol 2008 Mar;51(1):120-123. [doi: 10.1097/GRF.0b013e318161e657] [Medline: 18303506]
12. Ricchetti A, Becker M, Dulguerov P. Internal carotid artery dissection following rigid esophagoscopy. Arch Otolaryngol Head Neck Surg 1999 Jul;125(7):805-807. [doi: 10.1001/archotol.125.7.805] [Medline: 10406322]
13. Jones JW, Perryman M, Judge P, et al. Resident education in laryngeal stroboscopy and perceptual voice evaluation: an assessment. J Voice 2020 May;34(3):442-446. [doi: 10.1016/j.jvoice.2018.11.016] [Medline: 30545492]
14. Ngan JH, Fok PJ, Lai EC, Branicki FJ, Wong J. A prospective study on fish bone ingestion. Experience of 358 patients. Ann Surg 1990 Apr;211(4):459-462. [doi: 10.1097/00000658-199004000-00012] [Medline: 2322040]
15. Gupta SP, Pidgeon A. An analytical approach to identifying potential use-related issues concerning a medical device under development. J Med Eng Technol 2016;40(3):61-71. [doi: 10.3109/03091902.2015.1132785] [Medline: 26795586]

XSL•FO
RenderX

16. van der Peijl J, Klein J, Grass C, Freudenthal A. Design for risk control: the role of usability engineering in the management of use-related risks. J Biomed Inform 2012 Aug;45(4):795-812. [doi: 10.1016/j.jbi.2012.03.006] [Medline: 22466009]
17. Gao X, Wen Q, Duan X, et al. A hazard analysis of class I recalls of infusion pumps. JMIR Hum Factors 2019 May 3;6(2):e10366. [doi: 10.2196/10366] [Medline: 31066695]
18. Biadsee A, Crosby L, Chow W, Sowerby LJ. Cost minimization analysis of nasopharyngoscope reprocessing in community practice. J Otolaryngol Head Neck Surg 2023 Feb 8;52(1):8. [doi: 10.1186/s40463-022-00610-9] [Medline: 36750881]
19. Jiang R, Kasle DA, Alzahrani F, Kohli N, Lerner MZ. A manufacturer and user facility device experience analysis of upper aerodigestive endoscopy contamination: is flexible laryngoscopy different? Laryngoscope 2021 Mar;131(3):598-605. [doi: 10.1002/lary.28826] [Medline: 32558941]
20. Sokol WN. Nine episodes of anaphylaxis following cystoscopy caused by Cidex OPA (ortho-phthalaldehyde) high-level disinfectant in 4 patients after cytoscopy. J Allergy Clin Immunol 2004 Aug;114(2):392-397. [doi: 10.1016/j.jaci.2004.04.031] [Medline: 15316522]
21. Suzukawa M, Komiya A, Koketsu R, et al. Three cases of ortho-phthalaldehyde-induced anaphylaxis after laryngoscopy: detection of specific IgE in serum. Allergol Int 2007 Sep;56(3):313-316. [doi: 10.2332/allergolint.C-06-51] [Medline: 17582212]
22. Kramer A, Kohnen W, Israel S, et al. Principles of infection prevention and reprocessing in ENT endoscopy. GMS Curr Top Otorhinolaryngol Head Neck Surg 2015;14:Doc10. [doi: 10.3205/cto000125] [Medline: 26770284]
23. Mistry R, Russell RV, Walker N, Ofo E. The single-use rhinolaryngoscope: an evaluation and cost comparison. J Laryngol Otol 2020 Sep 28;134(9):1-8. [doi: 10.1017/S0022215120001656] [Medline: 32985399]
24. Bottrill I, Perrault DF Jr, Poe D. In vitro and in vivo determination of the thermal effect of middle ear endoscopy. Laryngoscope 1996 Feb;106(2 Pt 1):213-216. [doi: 10.1097/00005537-199602000-00020] [Medline: 8583856]
25. Lewis T, Levin M, Sommer DD. Too hot to handle-quantifying temperature variations in the nasal endoscope ocular assembly and light post. Am J Rhinol Allergy 2020 Mar;34(2):262-268. [doi: 10.1177/1945892419892182] [Medline: 31779480]
26. Yavuz Y, Skogås JG, Güllüoğlu MG, Langø T, Mårvik R. Are cold light sources really cold? Surg Laparosc Endosc Percutan Tech 2006 Oct;16(5):370-376. [doi: 10.1097/01.sle.0000213711.32805.15] [Medline: 17057587]
27. MacKeith SAC, Frampton S, Pothier DD. Thermal properties of operative endoscopes used in otorhinolaryngology. J Laryngol Otol 2008 Jul;122(7):711-714. [doi: 10.1017/S0022215107000734] [Medline: 17942007]
28. Chitnavis J. Silent burn: the hidden danger and effects of bright light from fibre-optic cables in arthroscopic knee surgery. J Surg Case Rep 2020 Apr;2020(4):rjaa068. [doi: 10.1093/jscr/rjaa068] [Medline: 32280444]
29. Prasad N, Tavaluc R, Harley E. Thermal injury to common operating room materials by fiber optic light sources and endoscopes. Am J Otolaryngol 2019;40(5):631-635. [doi: 10.1016/j.amjoto.2019.05.007] [Medline: 31178283]
30. Souza E Silva TX, Nicolau ABF, Antunes ML. Thermal variation in human temporal bone using rigid endoscope. Braz J Otorhinolaryngol 2024;90(3):101381. [doi: 10.1016/j.bjorl.2023.101381] [Medline: 38364523]
31. Modlin IM, Kidd M, Lye KD. From the lumen to the laparoscope. Arch Surg 2004 Oct;139(10):1110-1126. [doi: 10.1001/archsurg.139.10.1110] [Medline: 15492154]
32. Hoffmann C, Avery K, Macefield R, et al. Usability of an automated system for real-time monitoring of shared decision-making for surgery: mixed methods evaluation. JMIR Hum Factors 2024 Apr 10;11:e46698. [doi: 10.2196/46698] [Medline: 38598276]

Abbreviations

***CPT*:** *Current Procedural Terminology*
FDA: Food and Drug Administration
MAUDE: Manufacturer and User Facility Device Experience
MDR: medical device report

Edited by A Kushniruk; submitted 30.09.24; peer-reviewed by J Walsh, LE Galvao Martins; revised version received 11.01.25; accepted 14.01.25; published 05.02.25.

Please cite as:
Chang SH, Chen D, Chen CS, Zhou D, Yeh LK
Device Failures and Adverse Events Associated With Rhinolaryngoscopes: Analysis of the Manufacturer and User Facility Device Experience (MAUDE) Database
JMIR Hum Factors 2025;12:e67036
URL: https://humanfactors.jmir.org/2025/1/e67036
doi:10.2196/67036

XSL•FO
RenderX

© Shao-Hsuan Chang, Daishi Chen, Chi-Sheng Chen, Dong Zhou, Lung-Kun Yeh. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 5.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Completion Rates for Ecological Momentary Assessments of Food Intake During Pregnancy and Post Partum: Descriptive Study

Sarah Annalise Sanders[1], MPH, PhD; Serwaa Omowale[2], LMSW, MPH, PhD; Andrea Casas[1], MPH, PhD; Alexis Kiyanda[3], MS; Abigail Smith Kosbie[4], MPH; Yu-Hsuan Lai[1], MSPH, PhD; Meredith Wallace[5], PhD; Stephen Rathbun[6], MSc, PhD; Tiffany Gary-Webb[1], MHS, PhD; Esa Davis[7], MPH, MD; Lora Burke[8], MPH, PhD; Dara Méndez[1], MPH, PhD

[1]Department of Epidemiology, University of Pittsburgh, 130 De Soto Street, Pittsburgh, PA, United States

[2]Department of Management, Policy, and Community Health, University of Texas Health Science Center at Houston, Houston, TX, United States

[3]Department of Environmental and Occupational Health, University of Pittsburgh, Pittsburgh, PA, United States

[4]Department of Behavioral & Community Health Sciences, University of Pittsburgh, Pittsburgh, PA, United States

[5]Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA, United States

[6]Department of Epidemiology & Biostatistics, University of Georgia, Athens, GA, United States

[7]Family and Community Medicine, School of Medicine, University of Maryland, Baltimore, MD, United States

[8]School of Nursing, University of Pittsburgh, Pittsburgh, PA, United States

Corresponding Author:
Dara Méndez, MPH, PhD
Department of Epidemiology, University of Pittsburgh, 130 De Soto Street, Pittsburgh, PA, United States

Abstract

Background: The collection of dietary behavior data is crucial in childbearing populations. In addition to observed inequities in perinatal dietary intake and quality, burdensome assessment methods (eg, 24-h dietary recall) may limit research participation for some groups. Ecological momentary assessment (EMA) is associated with reduced recall bias and participant convenience, but there is a dearth of studies with diverse cohorts.

Objective: Our aim is to describe participant completion of food intake items in EMA surveys, overall and across individual characteristics (eg, prepregnancy BMI).

Methods: Using secondary EMA data from participants in a longitudinal study, we report average completion rates of survey items regarding dietary behavior (eg, number of meals eaten in a day) across individual demographic variables (eg, age) and combined strata (eg, race+age) during late pregnancy and throughout 12 months post partum.

Results: In our analytic sample (N=310), the average completion rate was 52.4% (SD 27.8%) during pregnancy, rising to 59.1% (SD 22.0%) after giving birth. Participants who were older (>30 y), overweight before pregnancy, self-identified as White, working, or earning higher annual income (>US $50,000) had higher average completion rates than their counterparts. Examining combined strata, we found some variation in survey completion within racial groups. Black participants using a study phone had higher average completion rates during pregnancy and post partum, but this relationship was reversed for White participants.

Conclusions: Our secondary analysis showed relatively stable engagement with EMA surveys in a childbearing cohort across 15 months. Increased completion rates among privileged groups (eg, White, higher income) may demonstrate the impact of socioeconomic advantages on individual health behaviors. Investigators should consider how intersections between race and other factors (eg, employment) may impact participation and data collection.

(JMIR Hum Factors 2025;12:e67081) doi:10.2196/67081

KEYWORDS

dietary behavior; ecological momentary assessment; pregnancy; post partum; dietary; diet; dietary intake; survey; longitudinal study

Introduction

The collection of dietary behavior data is integral to understanding how diet and eating patterns shape health for individuals and communities [1,2]. In childbearing populations, studies show that dietary intake patterns during pregnancy are critical to the health of both the mother and infant [3]. Inadequate nutrition during pregnancy can increase the risk of cardiometabolic comorbidities (eg, hypertension disorders of pregnancy and gestational diabetes) during pregnancy and of

XSL•FO
RenderX

low birthweight from fetal growth restriction and neurodevelopmental delays [4,5]. Additionally, research shows significant disparities in perinatal diet quality across social determinants of health (education, food security, income) [6-9], obliging studies to examine dietary behaviors across these strata. Racial inequities also exist, where Black birthing populations report higher food insecurity [6] and lower-quality diets [7-9] and increased prevalence of perinatal and postnatal comorbidities compared to White populations [10].

Commonly used dietary assessment methods include 24-hour dietary recalls, food frequency questionnaires, and food diaries, each presenting advantages and disadvantages in their administration and the quality of data collected [11]. The 24-hour dietary recall provides detailed information on dietary intake and has proven to be acceptable across diverse populations including pregnant individuals [12-15]. The major disadvantages of the 24-hour dietary recall measure are that it is prone to recall bias, it is limited by its reliance on an individual's episodic memory, there is potential for intentional misreporting, and it requires the greatest amount of time to complete [11,13]. Food frequency questionnaires also rely on participants' memory and lack detailed collection of specific food intake. Food diaries and the 24-hour dietary recall are similar in recording detailed food intake, increasing the burden on study participants and potentially influencing their eating behavior on the days being recorded [16].

Ecological momentary assessment (EMA) is a sampling method used to repeatedly collect information in "real time" [17] and in one's natural setting; it has been used to understand eating behaviors and food intake, typically via digital devices (eg, smartphones) [18]. EMA reduces retrospective bias and promotes reliability and ecological validity. Compared to more time-intensive methods, EMA reduces the burden of reporting and often the time lapse between a behavior and its recording. Furthermore, the repeated sampling design facilitates a broader behavioral assessment over time [17-22]. Recent research using EMA to assess dietary behaviors has been conducted in different populations, including children [23,24], adolescents [25,26], mother-child dyads [27-29], individuals with eating disorders [30], and postbariatric surgery patients [31]. The few EMA studies examining dietary behavior in perinatal populations report homogenous samples and poor adherence to EMA completion, making a strong case for more demographically diverse cohorts in evaluating the usability of these assessment tools [32,33].

Since EMA is a novel approach to collecting dietary data, particularly over long periods of time (eg, >3 mo), it is critical to assess participant adherence to the EMA protocol for reporting food intake. It is important to note that EMA surveys are designed to be briefer, making them ideal for studying the context of eating behaviors, rather than nutritional content, which may be measured more precisely in recalls and diaries. Leveraging EMA data from a diverse childbearing cohort within a longitudinal, observational investigation (approximately 15 mo), this secondary analysis aimed to describe participant completion of dietary intake items in EMA surveys, overall and across individual characteristics (eg, prepregnancy BMI).

Methods

Study Design

This secondary analysis uses participant-reported data regarding dietary intake from the Postpartum Mothers Mobile Study (PMOMS), a longitudinal, observational investigation [34] designed to examine factors associated with racial disparities in postpartum weight retention and cardiometabolic health. Pregnant people were recruited between 18 and 32 weeks' gestation and followed through 12 months post partum. Participants completed daily EMA surveys via smartphones, weighed themselves weekly using Bluetooth-enabled scales, and attended follow-up visits for anthropometric measurements and biospecimen collection. The study applied EMA methods to better understand participants' experiences and exposures in their natural environment via real-time measurements of psychosocial (eg, stress, discrimination), behavioral (eg, eating behavior, consumption of foods with added sugars, physical activity), and contextual (eg, geographic location) factors.

Brief EMA surveys were administered at the beginning of day (BOD) and end of day (EOD), representing time-contingent responses, and at random times throughout the day (signal-contingent responses). Participants selected times for receiving BOD and EOD surveys, with at least 9 hours between the 2 surveys. Links for BOD and EOD surveys were active for 30 minutes (60 min for random surveys) to ensure that participant-reported data reflected real-time information. Each participant contributed approximately 15 months of EMA data during the pregnancy and postpartum periods. The protocol for PMOMS is described in more detail in a previous publication [34].

Ethical Considerations

The study was approved by the institutional review board within the University of Pittsburgh Human Research Protection Office (#PRO16100117).

Participants and Recruitment

PMOMS was an ancillary study to a randomized clinical trial, the Comparison of Two Screening Strategies for Gestational Diabetes (GDM^2) [35,36], conducted at a women's hospital in southwestern Pennsylvania. The clinical trial visits provided 2 in-person opportunities for our staff to approach and recruit potential participants and determine eligibility between 18 and 32 weeks' gestation. Participants were recruited this way between December 2017 and March 2019. After the GDM^2 study ended recruitment, PMOMS staff directly recruited participants from December 2019 to March 2020, when study activities were interrupted by the COVID-19 pandemic.

After providing their written, informed consent at a baseline visit (between 18 and 32 weeks' gestation), participants completed demographic questionnaires and anthropometric measures. At this same visit, a research assistant facilitated setting up participants' smartphones for receiving daily EMA surveys, which were scheduled to begin the following day [34]. Individuals using their own phone for study participation were compensated US $35 each month for using their data (up to US

XSL•FO
RenderX

$525 throughout the study). For those without a suitable device for participating in the study, PMOMS offered a smartphone with a paid, unlimited data plan for the duration of the study. Upon completing the study, these participants had the option of keeping the smartphone provided by the study as an additional incentive.

All participants were compensated monthly if their EMA completion rate reached 60% or more during that month. Those who reached a completion rate of 80% or higher were entered into a monthly drawing for additional compensation. In addition, PMOMS staff conducted routine check-ins (via phone or text) with all participants falling below the 60% threshold to identify and address any barriers to increasing their completion rate.

Pregnant individuals were excluded from participating in PMOMS if they indicated any of the following: plan to deliver at a different hospital from the study recruitment site, multiple pregnancy, plan to move from the area before 12 months post partum, history of type 1 or type 2 diabetes diagnosis, gestational diabetes diagnosis before 20 weeks' gestation in current pregnancy, consumption of oral glucocorticoids in last 30 days, history of "dumping syndrome" following gastric bypass surgery, hypertension diagnosis or medication for high blood pressure, and diagnosis of severe liver disease (eg, cirrhosis). Further details about the PMOMS protocol can be found elsewhere [34].

Measures

Although EOD surveys were delivered to participants every day, eating habits and intake of added-sugar foods were only assessed in these surveys on 10 randomly selected weekdays and 4 randomly selected weekend days for every 28-day block of time after EMA prompts began [34]. Participants were asked to report the number of meals and snacks they ate during the day. They were also asked to report whether they had any food or drinks with added sugar. If they answered yes to either, they were asked to report the number of foods with added sugar in 4 categories (cookies, cake, pie, and donuts; candy; ice cream and frozen treats; and other sugary foods), as well as the number of drinks with added sugar in 4 categories (regular soda pop, juice drinks, coffee shop drink and frozen treats, and other sugary drinks). These EMA survey items, along with dozens of others, were developed for the purposes of our study and were tested in a pilot study [37]. All EMA surveys were designed to be brief (1 - 2 min in duration), and those including dietary intake items did not ask participants to specify sizes or types of added sugar items consumed.

Upon enrollment, we collected data on maternal age, education, race, ethnicity, employment status, marital status, and annual household income. We dichotomized some demographic variables for descriptive statistics: age as ≤30 years and >30 years (based on the sample median), employment status as working (part-time or full-time) versus not working (unemployed or on disability), and annual household income as ≤US $50,000 per year versus >US $50,000 per year (based on the sample median). In this study, we also examined EMA completion between participants using their own phone and those using a smartphone provided by PMOMS. Study staff measured participant height at the first clinical trial visit, and participants self-reported their prepregnancy weight at the second clinical trial visit. We calculated prepregnancy BMI by dividing participants' weight in kilograms by their height in meters squared and categorized participants as normal/healthy (BMI 18.0 - 24.9) and overweight (BMI≥25.0) [38].

Data Analysis

We determined our final sample for this secondary analysis (N=310) after excluding participants only if they were determined to be ineligible for PMOMS (n=3) at the study enrollment phase because they withdrew from either PMOMS or GDM^2 (parent study). We report frequencies and percentages for all demographic and anthropometric variables.

Completion rates were calculated for all participants included in the analytic sample by dividing the number of responses to dietary intake questions by the number of EOD surveys delivered during both pregnancy and postpartum periods. Given the sampling design described above, we only considered EOD surveys with the dietary intake questions in these completion calculations. We then stratified the study sample across demographic and clinical measures, describing average completion rates for each pair. In addition to the subgroups described above (by age, employment status upon study enrollment, household income), we examined average completion rates for participants by prepregnancy BMI (healthy: 18.0 - 24.9 vs overweight: ≥25.0), study phone usage, and race (Black vs White). Lastly, we looked at average completion rates across combined sample strata (eg, by race and age). We focused specifically on Black/African American and White participants in describing completion rates in subgroups due to sample size limitations in other racial groups.

Results

Table 1 provides an overall description of the analytic sample (N=310), which attempted a total of 36,687 EOD surveys that included dietary intake items, with a mean of 121.5 (SD 70.7) EOD surveys attempted per person (median 2128.0, IQR 60.3-185.8) and a range of 1.0 - 243.0 EOD surveys attempted per participant over the study period. Only 3 participants were excluded from these analyses due to being deemed ineligible to participate in PMOMS after enrolling in the study (eg, withdrew from parent study).

XSL•FO
RenderX

Table . Descriptive statistics of participant demographics and average end of day completion rates in the analytic sample (N=310).

	Sample, n (%)
Using study phone	87 (28.1)
≤30 y old	172 (55.5)
Healthy prepregnancy BMI (18.5 - 24.9)	81 (26.1)
Employment	
Full-time	172 (55.5)
Part-time	46 (14.8)
Disability	5 (1.6)
Unemployed/not working	87 (28.1)
Annual household income (US $)	
<20,000	77 (24.8)
21,000 - 30,000	41 (13.2)
31,000 - 40,000	17 (5.5)
41,000 - 50,000	19 (6.1)
51,000 - 60,000	22 (7.1)
61,000 - 70,000	12 (3.9)
71,000 - 80,000	19 (6.1)
>81,000	103 (33.2)
Education	
<High school	16 (5.2)
High school (or equivalent)	57 (18.4)
Some college	65 (21.0)
College degree	82 (26.5)
Master'ss degree	55 (17.7)
Doctoral degree or higher	35 (11.3)
Marital status	
Single (never married)	122 (39.4)
Separated/divorced	11 (3.6)
Married	175 (56.5)
Other	2 (1.0)
Hispanic/Latina	18 (5.8)
Race	
Asian	12 (3.9)
NHOPI[a]	1 (0.3)
Black/African American	87 (28.1)
White	187 (60.3)
More than one race	11 (3.6)
Other	10 (3.2)

[a]NHOPI: Native Hawaiian or other Pacific Islander.

Compared to the overall demographics of the county where this research was conducted, participants in the present analytic sample are more likely to be Black, Hispanic, or multiracial; be in the labor force; have higher educational attainment; and earn less income annually [39].

Table 2 shows the average completion rates of dietary intake items during the pregnancy (third trimester) and postpartum

XSL•FO
RenderX

periods (up to 12 mo after giving birth), as well as across sample strata (eg, study phone vs own phone). We note that values are missing for 6 (1.9%) participants during pregnancy and 11 (3.5%) participants post partum, due to no EOD surveys being delivered to these participants during that time. Rate calculations for the 4 dietary intake items in EOD surveys did not vary between the items, so the completion rates for just 1 item (meals) are presented as proxy rates for overall completion of these items.

Table . Average completion rates for dietary items during pregnancy and post partum in the overall cohort (N=310), as well as by subgroups and combined strata.

	Average completion rate during pregnancy		Average completion rate post partum	
	Percent, mean (SD)	Sample, n	Percent, mean (SD)	Sample, n
Overall	52.4 (27.8)	299	59.1 (22.0)	304
Study phone	46.7 (29.2)	86	59.2 (20.7)	85
Own phone	54.7 (26.9)	213	59.0 (22.5)	219
BMI<25.0	49.2 (27.5)	77	56.7 (23.6)	80
BMI≥25.0	53.5 (27.8)	222	59.9 (21.4)	224
≤30 y	44.0 (28.6)	163	54.5 (24.2)	169
>30 y	62.5 (23.2)	136	64.9 (17.3)	135
Black/AA[a]	37.3 (27.6)	85	51.3 (24.7)	85
White	58.5 (25.6)	179	62.8 (19.8)	183
Working	43.2 (27.9)	88	59.4 (22.0)	214
Not working	56.3 (26.8)	211	58.4 (22.2)	90
≤US $50,000/y	41.9 (28.4)	148	55.1 (23.5)	150
>US $50,000/y	62.7 (23.0)	151	63.0 (19.7)	154
Black/AA+another variable				
Study phone	40.9 (28.2)	46	57.3 (22.2)	44
Own phone	33.1 (26.7)	39	44.9 (25.9)	41
BMI<25.0	31.1 (24.8)	24	45.9 (25.2)	24
BMI≥25.0	39.8 (28.5)	61	53.4 (24.4)	61
≤30 y	34.3 (27.7)	65	50.5 (25.7)	65
>30 y	47.2 (25.6)	20	53.9 (21.8)	20
Working	39.3 (29.9)	45	46.9 (25.1)	47
Not working	35.2 (25.1)	40	56.8 (23.5)	38
≤US $50,000/y	37.1 (27.7)	79	51.0 (25.3)	79
>US $50,000/y	41.0 (28.9)	6	55.2 (15.6)	6
White+another variable				
Study phone	52.3 (29.3)	28	58.5 (20.0)	29
Own phone	59.6 (24.7)	151	63.6 (19.8)	154
BMI<25.0	56.6 (24.2)	42	60.3 (23.0)	45
BMI≥25.0	59.0 (26.0)	137	63.6 (18.7)	138
≤30 y	51.3 (27.6)	80	57.6 (23.4)	85
>30 y	64.3 (22.3)	99	67.2 (14.8)	98
Working	60.5 (24.6)	142	63.3 (19.8)	143
Not working	50.5 (28.0)	37	60.8 (20.3)	40
≤US $50,000/*y*	46.1 (28.5)	53	60.4 (20.8)	54
>US $50,000/y	63.7 (22.3)	126	63.7 (19.4)	129

[a]AA: African American.

XSL•FO
RenderX

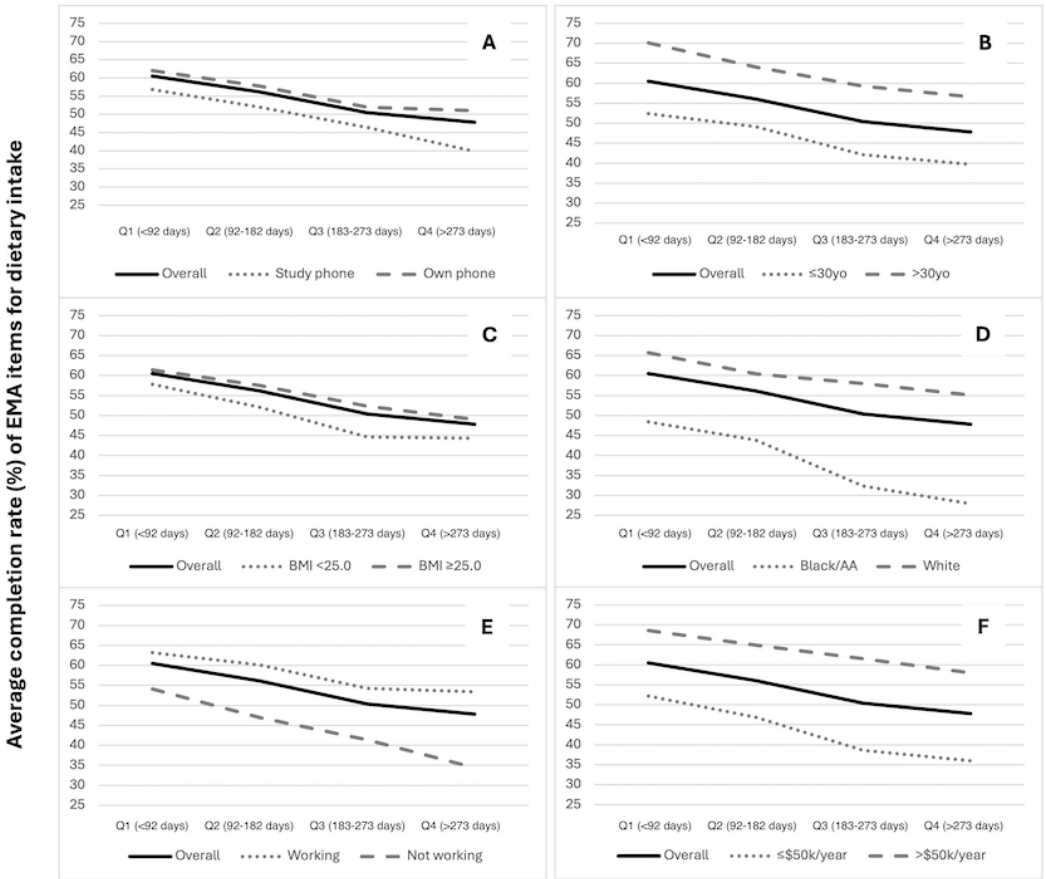
Overall, participants had higher average completion rates for dietary intake items during the postpartum period (59.1%, SD 22.0%) compared to the pregnancy period (52.4%, SD 27.8%). Looking at subgroup strata, participants using a study phone had lower completion rates on average (46.7%, SD 29.2%) compared to those using their own phone (54.7%, SD 26.9%) during pregnancy, but this difference was nearly eliminated after giving birth (59.2%, SD 20.7% vs 59.0%, SD 22.5%, respectively). Across both pregnancy and postpartum periods, participants who had an overweight BMI before pregnancy (>24.9), were older (>30 years), self-identified as White, were working, or earned a higher income (>US $50,000 annually) had higher average completion rates of dietary intake items.

We also examined completion rates within combined strata, which consisted of race (Black or White) and another variable (eg, phone use), to have a better understanding of variation in responding to EMA surveys within racial groups (Table 2). We observed increased average completion rates from pregnancy to post partum across all strata of Black participants, with the largest increases found for those not working at the time of study enrollment (61%), those with a healthy prepregnancy BMI (48%), and those 30 years old or younger (47%). Black participants using the study phone had considerably higher completion rates on average compared to those using their own phone in both the pregnancy and postpartum periods. During pregnancy, Black participants who were not working at the time of study enrollment (eg, unemployed, on disability) had lower average completion rates (35.2%) compared to those who were working (39.3%), but this was reversed in the postpartum period (56.8% vs 46.9%, respectively).

Similarly, we observed increased average completion rates from pregnancy to post partum across all strata of White participants, although with considerably less variation between time periods. The largest increases were found for White participants earning more than US $50,000 annually (31%) and those who were not working at the time of study enrollment (20%).

Lastly, given the length of the postpartum period (12 mo), we reported completion rates for each postpartum quarter (approximately 3 mo or 91 d) in the overall sample and across subgroup strata (Figure 1). With very few exceptions, completion rates generally decreased over time during the postpartum period, with the lowest rates for all subgroups in the fourth quarter (>273 d after giving birth). We saw similar declining trends in average completion rates for the same combined strata across postpartum quarters (Table S1, Multimedia Appendix 1).

Figure 1. Line graphs depicting average completion rates (%) of EMA items for dietary intake across four postpartum quarters (~91 d) for overall sample, compared against individual characteristics: study phone usage (A); age at study enrollment (B); prepregnancy BMI (C); self-reported race (D); employment status at study enrollment (E); and annual household income (F). AA: African American; EMA: ecological momentary assessment.



XSL•FO
RenderX

Discussion

Principal Findings

Our secondary study is innovative in its examination of overall completion rates for EMA surveys of dietary behavior in a longitudinal, childbearing cohort (~15 mo, from late pregnancy to 1 y post partum), as well as across individual characteristics. Overall, the analytic sample (N=310) had average completion rates of 52.4% during pregnancy (approximately 3 mo) and 59.1% during the first postpartum year (12 mo). This relationship generally held true across subgroups, with average completion rates increasing during the postpartum period, compared to the pregnancy period.

The average completion rates examined here are lower than ideal for EMA research, where at least 80% is often cited as the minimum threshold for "adequate" levels of compliance [40]. Most dietary behavior studies using EMA report adherence rates close to this level or higher [19]. Even so, very few of these studies employed EMA for more than a month, with many of them having study durations of just 1 week [19,41,42]. In our secondary analysis, we observed average completion rates over periods of about 3 months, potentially obscuring higher adherence during shorter intervals. Furthermore, the compensation structure for PMOMS participants was based on monthly rates of EMA adherence, rather than individual survey completion, which may have influenced participant responses.

Our findings partially contradict previous research that demonstrates reduced adherence to EMA protocols over time [19,20,43]. We observed an increase in average completion rates from pregnancy to post partum for all strata examined, potentially attributable to added availability for responding to EMA prompts (ie, maternity leave), reduced stress related to pregnancy symptoms, or increased interest in the study's objectives regarding postpartum health.

Even so, during the postpartum period, we observed a gradual decline in survey completion across the cohort (Table S1, Multimedia Appendix 1), in agreement with previous literature. We interpret these reductions to survey burden, as well as changing parental obligations, which may impact some populations more than others. In the last postpartum quarter (>273 d after giving birth), our overall cohort had an average completion rate of 47.8%. It is also important to highlight previous evidence demonstrating increased EMA adherence in the morning [20,43,44], whereas the dietary intake items being examined here appeared on the EOD survey.

We observed considerable disparities between subgroups in average completion rates, some of which provide insights into how racial and socioeconomic privilege may impact participation in research applying EMA. Participants who were >30 years old, self-identified as White, and earned >US $50,000 annually had higher completion rates than their counterparts across both prenatal and postpartum periods. Although research examining demographic differences in EMA adherence is limited, a 9-week smoking cessation study found greater declines in adherence for participants who were younger, working full-time, and not White [44]. These differences in EMA completion rates may be attributable to the socioeconomic advantages associated with privileged groups, such as increased access to resources (eg, employment, childcare) and reduced exposure to stressors (eg, food insecurity, housing instability). Chronic exposure to stressors, particularly related to caregiving, may exacerbate EMA burden for participants, even from one day to the next [42].

During the first postpartum year, average completion rates increased by 38%, 35%, 32%, and 24% for participants who self-identified as Black, reported not working (upon study enrollment), had lower income (≤US $50,000 annually), and were 30 years old or younger, respectively. These increases were more considerable than those observed among the privileged subgroups, which ranged from 0.5% to 7%. Beyond the experience of survey burden for participants, these findings suggest that the prenatal period carries additional burdens for childbearing people from racially and socioeconomically oppressed groups, potentially inhibiting their participation in research.

Examining combined strata here provided additional insights that would not have been afforded to us if we had ceased the analysis at single-identity categories. Black participants using the study phone had average completion rates approximately 20% higher than those using their own phone, across the study period, contrasting with results from both the overall study cohort and strata of White participants. We consider this a novel finding in EMA research, where study designs typically leverage the use of the participant's current mobile device, potentially excluding those who lack compatible devices or access to quality phone or internet service [41].

Strengths and Limitations

While this study offers novel, scientific insights into demographic and socioeconomic factors associated with completion of EMA surveys regarding food intake in a longitudinal, childbearing cohort, it is not without its limitations. First, PMOMS is an ancillary study to a randomized clinical trial, which had specific exclusionary criteria (eg, history of diabetes diagnosis) that may have ultimately biased EMA survey responses within the cohort. Second, examining combined strata (eg, race and income) resulted in some small samples, making our study underpowered to conduct comparative analyses or report statistical significance. For this reason, we decided to provide descriptive results only. Future investigations of variation in EMA participation across sociodemographic factors would benefit from a larger, representative sample to apply comparative methods (ie, linear regression). Next, we conducted our assessment of employment status upon study enrollment (during late pregnancy). We asked about employment status in an exit survey within an administrative supplement of the study, meaning only a subset of the population had employment information at the conclusion of the study. Lastly, we would also like to qualify that our study findings do not include information about any technical issues (eg, device malfunction, survey delivery failure), which would likely impact participants' EMA completion. Although the PMOMS team documented these issues throughout the study period to assist participants

XSL•FO
RenderX

and troubleshoot, these data were not prepared to be included in these analyses.

Public Health Implications

The findings presented here are novel and valuable to perinatal health research, particularly that which applies EMA or telehealth technologies to collect data in diverse, representative cohorts for longer periods (eg, several weeks, months). Although PMOMS is considerably longer than most EMA studies, here we demonstrated relatively stable engagement (average completion rate of 59%) in a childbearing population over the course of 15 months. Given evidence linking contextual factors (eg, stress) to parental behavior, including dietary behaviors [41], EMA may serve as a better methodological tool in maternal health research compared to more burdensome dietary recalls.

Our descriptive analysis of EMA completion rates in a childbearing cohort has important implications for future research and intervention designs. PMOMS remains unique in its application of EMA over more than 12 months, though there is room for improving completion rates in longer EMA studies. Immediately actionable strategies include reducing the number of EMA surveys per day or week and increasing opportunities for compensation among participants. Another approach would be increasing participant engagement through an online dashboard where they can see their own response data, as well as aggregate data for the entire cohort.

We recommend that investigators consider the intersections between race and other factors (eg, employment) in relation to study participation and develop strategies to improve EMA completion (eg, study-provided smartphone) among participants who are likely to face more socioeconomic burdens.

Acknowledgments

The authors thank the study participants who generously gave their time and effort. The authors would also like to thank Terri Washington, study coordinator; Maura Dugan, Marquita Smalls, and Gabriella Mendez for study recruitment and follow-up; Er Wang, Pedram Gharani, Meirman Syzdykbayev, and Bradley Wheeler, and Dr Hassan Karimi for technology and app development and support; and John Gianakas for database development and management. Finally, the authors thank all the staff and investigators of the Comparison of Two Screening Strategies for Gestational Diabetes trial.

Conflicts of Interest

ED is a member of the United States Preventive Services Task Force. This article does not necessarily represent the views and policies of the United States Preventive Services Task Force. MW is a consultant for Noctem, but the author's work is not related to this project.

Multimedia Appendix 1
Average completion rates (%) for dietary intake items by postpartum quarter (~91 d) across combined strata (race+another variable).
[DOCX File, 16 KB - humanfactors_v12i1e67081_app1.docx]

References

1. Bailey RL. Overview of dietary assessment methods for measuring intakes of foods, beverages, and dietary supplements in research studies. Curr Opin Biotechnol 2021 Aug;70:91-96. [doi: 10.1016/j.copbio.2021.02.007] [Medline: 33714006]
2. Mousa A, Naqash A, Lim S. Macronutrient and micronutrient intake during pregnancy: an overview of recent evidence. Nutrients 2019 Feb 20;11(2):443. [doi: 10.3390/nu11020443] [Medline: 30791647]
3. Mitran AM, Gherasim A, Ni ă O, et al. Exploring lifestyle and dietary patterns in pregnancy and their impact on health: a comparative analysis of two distinct groups 10 years apart. Nutrients 2024 Jan 27;16(3):377. [doi: 10.3390/nu16030377] [Medline: 38337662]
4. Marshall NE, Abrams B, Barbour LA, et al. The importance of nutrition in pregnancy and lactation: lifelong consequences. Am J Obstet Gynecol 2022 May;226(5):607-632. [doi: 10.1016/j.ajog.2021.12.035] [Medline: 34968458]
5. Cortés-Albornoz MC, García-Guáqueta DP, Velez-van-Meerbeke A, Talero-Gutiérrez C. Maternal nutrition and neurodevelopment: a scoping review. Nutrients 2021 Oct 8;13(10):3530. [doi: 10.3390/nu13103530] [Medline: 34684531]
6. Shriver LH, Eagleton SG, Hosseinzadeh M, Buehler C, Wideman L, Leerkes EM. Associations among eating behaviors, food security status, and dietary intake during pregnancy. Appetite 2023 Dec 1;191:107062. [doi: 10.1016/j.appet.2023.107062] [Medline: 37742786]
7. Yee LM, Silver RM, Haas DM, et al. Quality of periconceptional dietary intake and maternal and neonatal outcomes. Am J Obstet Gynecol 2020 Jul;223(1):121. [doi: 10.1016/j.ajog.2020.01.042] [Medline: 31981510]
8. Harris A, Chilukuri N, West M, et al. Obesity-related dietary behaviors among racially and ethnically diverse pregnant and postpartum women. J Pregnancy 2016;2016:9832167. [doi: 10.1155/2016/9832167] [Medline: 27298738]
9. Bodnar LM, Simhan HN, Parker CB, et al. Racial or ethnic and socioeconomic inequalities in adherence to national dietary guidance in a large cohort of US pregnant women. J Acad Nutr Diet 2017 Jun;117(6):867-877. [doi: 10.1016/j.jand.2017.01.016] [Medline: 28320597]

XSL•FO
RenderX

10. Howell EA. Reducing disparities in severe maternal morbidity and mortality. Clin Obstet Gynecol 2018 Jun;61(2):387-399. [doi: 10.1097/GRF.0000000000000349] [Medline: 29346121]
11. Naska A, Lagiou A, Lagiou P. Dietary assessment methods in epidemiological research: current state of the art and future prospects. F1000Res 2017;6:926. [doi: 10.12688/f1000research.10703.1] [Medline: 28690835]
12. Arab L, Wesseling-Perry K, Jardack P, Henry J, Winter A. Eight self-administered 24-hour dietary recalls using the internet are feasible in African Americans and Whites: the energetics study. J Am Diet Assoc 2010 Jun;110(6):857-864. [doi: 10.1016/j.jada.2010.03.024] [Medline: 20497774]
13. DeBiasse MA, Bowen DJ, Quatromoni PA, Quinn E, Quintiliani LM. Feasibility and acceptability of dietary intake assessment via 24-hour recall and food frequency questionnaire among women with low socioeconomic status. J Acad Nutr Diet 2018 Feb;118(2):301-307. [doi: 10.1016/j.jand.2017.08.011] [Medline: 29102422]
14. Savard C, Lemieux S, Lafrenière J, Laramée C, Robitaille J, Morisset AS. Validation of a self-administered web-based 24-hour dietary recall among pregnant women. BMC Pregnancy Childbirth 2018 Apr 23;18(1):112. [doi: 10.1186/s12884-018-1741-1] [Medline: 29685127]
15. Karvetti RL, Knuts LR. Validity of the 24-hour dietary recall. J Am Diet Assoc 1985 Nov;85(11):1437-1442. [doi: 10.1016/S0002-8223(21)03818-9] [Medline: 4056262]
16. Thompson FE, Subar AF. Ecological momentary assessment. Annu Rev Clin Psychol 2017:1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415]
17. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
18. Maugeri A, Barchitta M. A systematic review of ecological momentary assessment of diet: implications and perspectives for nutritional epidemiology. Nutrients 2019 Nov 7;11(11):2696. [doi: 10.3390/nu11112696] [Medline: 31703374]
19. Battaglia B, Lee L, Jia SS, Partridge SR, Allman-Farinelli M. The use of Mobile-Based Ecological Momentary Assessment (mEMA) methodology to assess dietary intake, food consumption behaviours and context in young people: a systematic review. Healthcare (Basel) 2022 Jul 18;10(7):1329. [doi: 10.3390/healthcare10071329] [Medline: 35885855]
20. Spook JE, Paulussen T, Kok G, Van Empelen P. Monitoring dietary intake and physical activity electronically: feasibility, usability, and ecological validity of a mobile-based Ecological Momentary Assessment tool. J Med Internet Res 2013 Sep 24;15(9):e214. [doi: 10.2196/jmir.2617] [Medline: 24067298]
21. Schembre SM, Liao Y, O'Connor SG, et al. Mobile ecological momentary diet assessment methods for behavioral research: systematic review. JMIR Mhealth Uhealth 2018 Nov 20;6(11):e11170. [doi: 10.2196/11170] [Medline: 30459148]
22. Tanweer A, Khan S, Mustafa FN, Imran S, Humayun A, Hussain ZUN. Improving dietary data collection tools for better nutritional assessment—a systematic review. Computer Methods Programs Biomedicine Update 2022;2:100067. [doi: 10.1016/j.cmpbup.2022.100067]
23. Campbell KL, Babiarz A, Wang Y, Tilton NA, Black MM, Hager ER. Factors in the home environment associated with toddler diet: an ecological momentary assessment study. Public Health Nutr 2018 Jul;21(10):1855-1864. [doi: 10.1017/S1368980018000186] [Medline: 29526170]
24. O'Reilly GA, Huh J, Schembre SM, Tate EB, Pentz MA, Dunton G. Association of usual self-reported dietary intake with ecological momentary measures of affective and physical feeling states in children. Appetite 2015 Sep;92:314-321. [doi: 10.1016/j.appet.2015.05.032] [Medline: 26032196]
25. Borgogna N, Lockhart G, Grenard JL, Barrett T, Shiffman S, Reynolds KD. Ecological momentary assessment of urban adolescents' technology use and cravings for unhealthy snacks and drinks: differences by ethnicity and sex. J Acad Nutr Diet 2015 May;115(5):759-766. [doi: 10.1016/j.jand.2014.10.015] [Medline: 25482855]
26. Grenard JL, Stacy AW, Shiffman S, et al. Sweetened drink and snacking cues in adolescents: a study using ecological momentary assessment. Appetite 2013 Aug;67:61-73. [doi: 10.1016/j.appet.2013.03.016] [Medline: 23583312]
27. Dunton GF, Dzubur E, Huh J, et al. Daily associations of stress and eating in mother-child dyads. Health Educ Behav 2017 Jun;44(3):365-369. [doi: 10.1177/1090198116663132] [Medline: 27531169]
28. O'Connor SG, Koprowski C, Dzubur E, Leventhal AM, Huh J, Dunton GF. Differences in mothers' and children's dietary intake during physical and sedentary activities: an Ecological Momentary Assessment Study. J Acad Nutr Diet 2017 Aug;117(8):1265-1271. [doi: 10.1016/j.jand.2017.02.012] [Medline: 28392348]
29. Mason TB, O'Connor SG, Schembre SM, Huh J, Chu D, Dunton GF. Momentary affect, stress coping, and food intake in mother-child dyads. Health Psychol 2019 Mar;38(3):238-247. [doi: 10.1037/hea0000714] [Medline: 30762403]
30. Levinson CA, Sala M, Fewell L, Brosof LC, Fournier L, Lenze EJ. Meal and snack-time eating disorder cognitions predict eating disorder behaviors and vice versa in a treatment seeking sample: a mobile technology based ecological momentary assessment study. Behav Res Ther 2018 Jun;105:36-42. [doi: 10.1016/j.brat.2018.03.008] [Medline: 29614379]
31. Thomas JG, Bond DS, Ryder BA, et al. Ecological momentary assessment of recommended postoperative eating and activity behaviors. Surg Obes Relat Dis 2011;7(2):206-212. [doi: 10.1016/j.soard.2010.10.007] [Medline: 21130703]
32. Most J, Vallo PM, Altazan AD, et al. Food photography is not an accurate measure of energy intake in obese, pregnant women. J Nutr 2018 Apr 1;148(4):658-663. [doi: 10.1093/jn/nxy009] [Medline: 29659958]

XSL•FO
RenderX

33. Mason TB, Wang WL, Bastain T, et al. Bi-directional associations of affective states and diet among low-income Hispanic pregnant women using ecological momentary assessment. Psychiatry Int 2022;3(4):273-285. [doi: 10.3390/psychiatryint3040022]
34. Mendez DD, Sanders SA, Karimi HA, et al. Understanding pregnancy and postpartum health using ecological momentary assessment and mobile technology: protocol for the postpartum mothers mobile study. JMIR Res Protoc 2019 Jun 26;8(6):e13569. [doi: 10.2196/13569] [Medline: 31244478]
35. Davis EM, Abebe KZ, Simhan HN, et al. Perinatal outcomes of two screening strategies for gestational diabetes mellitus: a randomized controlled trial. Obstet Gynecol 2021 Jul 1;138(1):6-15. [doi: 10.1097/AOG.0000000000004431] [Medline: 34259458]
36. Abebe KZ, Scifres C, Simhan HN, et al. Comparison of Two Screening Strategies for Gestational Diabetes (GDM 2) trial: design and rationale. Contemp Clin Trials 2017 Nov;62:43-49. [doi: 10.1016/j.cct.2017.08.012] [Medline: 28823926]
37. Mendez D, DeVaughn S, Duell J. The pregnancy ecological momentary assessment pilot study: applications of mobile technology. Integr Gyn Obstet J 2019 Feb 26;2(1):1-6. [doi: 10.31038/IGOJ.2019215]
38. Adult BMI calculator. Centers for Disease Control and Prevention. 2022 Jun 3. URL: https://www.cdc.gov/bmi/adult-calculator/index.html [accessed 2025-09-29]
39. Quick facts. US Census Bureau. 2024. URL: https://www.census.gov/quickfacts/fact/table/alleghenycountypennsylvania/BZA115223 [accessed 2025-09-29]
40. Stone AA, Shiffman S. Capturing momentary, self-report data: a proposal for reporting guidelines. Ann Behav Med 2002;24(3):236-243. [doi: 10.1207/S15324796ABM2403_09] [Medline: 12173681]
41. Trofholz A, Tate A, Janowiec M, et al. Ecological momentary assessment of weight-related behaviors in the home environment of children from low-income and racially and ethnically diverse households: development and usability study. JMIR Res Protoc 2021 Dec 1;10(12):e30525. [doi: 10.2196/30525] [Medline: 34855612]
42. Tate AD, Fertig AR, de Brito JN, et al. Momentary factors and study characteristics associated with participant burden and protocol adherence: ecological momentary assessment. JMIR Form Res 2024 Apr 24;8:e49512. [doi: 10.2196/49512] [Medline: 38656787]
43. Comulada WS, Swendeman D, Koussa MK, et al. Adherence to self-monitoring healthy lifestyle behaviours through mobile phone-based ecological momentary assessments and photographic food records over 6 months in mostly ethnic minority mothers. Public Health Nutr 2018 Mar;21(4):679-688. [doi: 10.1017/S1368980017003044] [Medline: 29199630]
44. Tonkin S, Gass J, Wray J, et al. Evaluating declines in compliance with ecological momentary assessment in longitudinal health behavior research: analyses from a clinical trial. J Med Internet Res 2023 Jun 22;25:e43826. [doi: 10.2196/43826] [Medline: 37347538]

Abbreviations

BOD: beginning of day
EMA: ecological momentary assessment
EOD: end of day
GDM^2: Comparison of Two Screening Strategies for Gestational Diabetes
PMOMS: Postpartum Mothers Mobile Study

Edited by A Kushniruk; submitted 01.10.24; peer-reviewed by T Baranowski, TM Kennedy; revised version received 19.06.25; accepted 10.07.25; published 08.10.25.

Please cite as:
Sanders SA, Omowale S, Casas A, Kiyanda A, Kosbie AS, Lai YH, Wallace M, Rathbun S, Gary-Webb T, Davis E, Burke L, Méndez D
Completion Rates for Ecological Momentary Assessments of Food Intake During Pregnancy and Post Partum: Descriptive Study
JMIR Hum Factors 2025;12:e67081
URL: https://humanfactors.jmir.org/2025/1/e67081
doi:10.2196/67081

© Sarah Annalise Sanders, Serwaa Omowale, Andrea Casas, Alexis Kiyanda, Abigail Smith Kosbie, Yu-Hsuan Lai, Meredith Wallace, Stephen Rathbun, Tiffany Gary-Webb, Esa Davis, Lora Burke, Dara Méndez. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 8.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete

XSL•FO
RenderX

bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Text Messaging Between Patients With Inflammatory Rheumatic Diseases and Pharmacists to Solve Drug-Related Problems: Prospective Feasibility Study

Lex L Haegens[1], MSc; Charlotte L Bekker[2], PhD; Marcel Flendrie[3], PhD; Bart J F van den Bemt[1,2,4], Prof Dr; Victor J B Huiskes[2,4], PhD

[1]Department of Research, Sint Maartenskliniek, Hengstdal 3, Ubbergen, The Netherlands

[2]Department of Pharmacy, Radboudumc, Nijmegen, The Netherlands

[3]Department of Rheumatology, Sint Maartenskliniek, Ubbergen, The Netherlands

[4]Department of Pharmacy, Sint Maartenskliniek, Ubbergen, The Netherlands

Corresponding Author:
Lex L Haegens, MSc
Department of Research, Sint Maartenskliniek, Hengstdal 3, Ubbergen, The Netherlands

Abstract

Background: Patients with inflammatory rheumatic diseases often experience drug-related problems (DRPs). As these can result in negative health consequences, DRPs should be identified and addressed in a timely manner. Text messaging between patients and pharmacists at the initiative of the patient has the potential to deliver support with DRPs more continuously, increase accessibility and efficiency, and enhance patient involvement in the process of identifying and solving DRPs.

Objective: This study aimed to assess the feasibility of text messaging from both the patients' and health care practitioners' perspectives before a large-scale implementation.

Methods: Adult patients using a disease-modifying antirheumatic drug were given access to text messaging with pharmacists to discuss DRPs for a period of 8 weeks. Patients received a response from a pharmacist within 4 working hours. Feasibility was evaluated based on five domains of Bowen's framework for designing feasibility studies: (1) demand: actual use, expressed interest (user version of the Mobile Application Rating Scale – section E), and factors impacting future use; (2) limited efficacy: number of DRPs solved, DRPs resulting in follow-up, and DRPs warranting involvement of health care provider; (3) implementation: degree of execution (number of conversations answered within service level) and resources needed (pharmacists' time investment per conversation); (4) acceptability: satisfaction and appropriateness (theoretical framework of acceptability); and (5) practicality: ability to carry out intervention activities (System Usability Scale). Data were collected by means of usage data and a questionnaire.

Results: In total, 45 patients (median age 57, IQR 52-65 y; n=31, 69% female) and 5 pharmacists (median age 41, IQR 26-47 y; n=1, 20% female) actively participated in this study. In the demand domain, 158 unique DRPs were raised in 133 conversations, with a median of 3 (IQR 2-4) unique DRPs per patient. Expressed interest was rated high by patients (median 4, IQR 4-5), and 90% (37/41) of patients would recommend text messaging to others. In the limited-efficacy domain, all DRPs were solved, and 77% (122/158) of DRPs warranted involvement of a health care provider. In the implementation domain, 87% (116/133) of conversations were answered within the promised timeframe with a median time investment of 4:15 (IQR 2:21-7:27) minutes per conversation. Acceptability was rated high by patients (median 4, IQR 4-5) and pharmacists (median 5, IQR 4-5). Finally, in the practicality domain, System Usability Scale was scored above average for patients (mean 72, SD 18) and pharmacists (mean 81, SD 16).

Conclusion: Text messaging with pharmacists at the initiative of patients with rheumatic diseases seems feasible for discussing DRPs in terms of limited efficacy, implementation, acceptability, demand, and practicality for patients and pharmacists.

(JMIR Hum Factors 2025;12:e66514) doi:10.2196/66514

KEYWORDS

drug-related problems; text messaging; disease-modifying antirheumatic drug; inflammatory rheumatic diseases; digital health; feasibility; inflammatory; rheumatic diseases; DRPs; questionnaires; mHealth; mobile health

XSL•FO
RenderX

Introduction

Pharmacotherapy is the cornerstone of the treatment of many inflammatory rheumatic diseases, which involves patients often using multiple drugs chronically. Disease-modifying antirheumatic drugs (DMARDs) are effective in decreasing disease activity and disability in patients with inflammatory rheumatic diseases and can even decrease or prevent joint damage [1]. For optimal treatment outcomes and patient safety, it is essential that DMARDs are used as agreed upon between a patient and health care provider, also called adherence [2]. However, nonadherence is common in this population [3], which is partly caused by drug-related problems (DRPs) experienced by patients with rheumatic diseases [4]. These DRPs encompass side effects, difficulties with medication management (eg, problems regarding opening medication packaging, administering and storing medication, or traveling with medication), a lack of knowledge, or concerns about lack of effect and unwanted effects of long-term medication use, among others. DRPs can potentially lead to negative health consequences, such as increased morbidity, which in turn can result in increased health care costs and decreased quality of life [5,6]. Thus, DRPs should be identified and resolved as soon as possible (ie, at the earliest stage of DRP development possible) to be able to prevent or minimize these negative consequences. However, during regular health care provider consultations, patients either do not always report all DRPs or health care providers do not act upon DRPs reported by patients [7]. In addition, patient-provider consultations occur only a few times a year, while DRPs can occur continuously, often in between these contact moments. Ideally, for adequate management of DRPs, patients should be able to contact their health care provider as soon as a DRP occurs, and health care providers should be able to react in a timely manner. To this end, one could argue that more continuous accessible contact between patients and health care providers is needed.

eHealth, which entails information and communication technology in health care that can benefit both patients and health care providers, could facilitate such contact by, for example, making health care more efficient and more easily accessible because of its time- and place-independent possibilities [8,9]. In addition, eHealth can increase patient involvement in the process of identifying and solving DRPs, for example, by lowering the threshold for patients to contact health care providers. Previous research has shown that patient involvement is important in identifying and solving clinically relevant DRPs [10,11], underlining the key role of patients in this process.

One such eHealth instrument is text messaging with a pharmacist regarding DRPs. Text messaging has the potential to enable patients to discuss DRPs in a timely manner with a health care provider, as it can provide faster contact between patients and health care providers without waiting or travel time. In addition, text messaging can be an addition to the already available communication channels, such as contact via telephone. Research demonstrated that offering multiple communication channels is important in providing patients with a suitable channel for discussing individual DRPs [12]. Finally, text messaging seems more efficient compared with phone calls, as shown by diabetes self-management support via text messaging requiring less staff time than via phone [13]. Furthermore, text messaging has been shown to be effective and feasible as an intervention to manage, for instance, blood pressure [14], adherence in type 2 diabetes [15], and postoperative pain [16]. However, these studies evaluated text messaging only as a channel for communication at the initiative of the health care provider to instruct patients (eg, by sending reminders regarding adherence) and provide health-related information to patients, and in most cases, only encompassed one-way texting from health care provider to patient (ie, patients could not respond to text messages).

As patients fulfill a key role in timely identifying potentially clinically relevant DRPs, this study opts for text messaging at the initiative of the patient, in which patients can contact their pharmacist when DRPs arise to facilitate timely resolution. This potentially provides a personalized and quick channel to raise, discuss, and resolve DRPs effectively. As patient-initiated text messaging with a pharmacist to address and resolve DRPs has not been evaluated, it is important to first study the feasibility of this functionality from the perspectives of both patients and pharmacists before its effect on reducing DRPs can be studied and large-scale implementation can be considered. Therefore, this study aims to assess the feasibility of text messaging between patients with inflammatory rheumatic diseases and pharmacists at the initiative of the patient when experiencing DRPs.

Methods

Study Design

A prospective feasibility study was conducted at the outpatient pharmacy of the Sint Maartenskliniek in the Netherlands between August 2023 and February 2024.

Ethical Considerations

The Medical Research Ethics Committee of Arnhem-Nijmegen, the Netherlands, reviewed this study and waived official ethical approval (case 2023 - 16580) as this study was deemed not subject to the Medical Research Involving Human Subjects Act. All participants gave informed consent digitally before commencement of the study (Multimedia Appendix 1). All participant data were anonymized to protect participant privacy and were stored on a secure environment only accessible to the researchers. Participants did not receive compensation for participation.

Participants

Adult patients (≥18 y) with an inflammatory rheumatic disease receiving at least 1 DMARD from the outpatient pharmacy were eligible for inclusion. In addition, pharmacists covering pharmacy day shifts during the study period were eligible for participation. Pharmacists conducted study activities as part of their daily routine and gave consent verbally.

As the number and types of DRPs are likely dependent on the type of DMARDs patients use, as well as on the duration of their treatment [17], purposive sampling was performed using

XSL•FO
RenderX

the following criteria: (1) biological disease-modifying antirheumatic drugs (bDMARDs), conventional synthetic DMARDS or targeted synthetic DMARDS; (2) patients initiating a new DMARD (new users) or patients using a specific DMARD for at least 6 months before inclusion (existing users). Although previous studies regarding sample sizes in feasibility studies advise sample sizes ranging from 12 to 35 participants per group [18,19], we aimed to include 60 participants (30 new and 30 existing users) distributed evenly among the 3 groups of DMARDs.

New users were selected from the pharmacy information system and were defined as patients who got their DMARD dispensed for the first time in the 2 weeks before inclusion. For bDMARDs, only bDMARD-naïve patients (ie, patients starting their first bDMARD ever) were eligible for participation, as switching between bDMARDs is common and DRPs reported by patients usually differ minimally between bDMARDs.

Existing users were selected from all patients who received a DMARD from the outpatient pharmacy in the last 3 months before the study started and those who had been using that specific DMARD for at least 6 months before the study started. A random sample was then taken from which patients were purposively approached for inclusion.

Inclusion was performed by telephone. During this phone call, patients received a verbal explanation of the study content. Patients willing to participate received written study information and a digital informed consent form. Participation started after informed consent was given.

Intervention

The intervention used in this study was text messaging with a pharmacist at the initiative of the patient. This enables patients to report DRPs directly to a pharmacist. Participants were able to access the text messaging functionality by securely logging in to the hospital's patient portal either via the hospital's website or a smartphone app before each use. Participants were asked to discuss questions or problems regarding their antirheumatic drugs with a pharmacist using text messaging for a period of 8 weeks and received reminders at the start of the study and every 2 weeks thereafter, to ensure exposure to the intervention. There was no restriction on the number of conversations participants were allowed to initiate. The service level (ie, the aimed maximum time between patients reporting a DRP and pharmacists responding) was set at 4 hours from the moment the question was asked, within working hours and days (Monday to Friday between 9 AM and 5 PM).

Feasibility Evaluation: Outcome Measures

Overview

Feasibility was assessed using the framework for designing feasibility studies by Bowen et al [20]. This widely applied model describes feasibility in 8 areas of focus. For this study, five of these eight domains were assessed: (1) demand, (2) limited-efficacy testing, (3) implementation, (4) acceptability, and (5) practicality. Furthermore, 3 domains (adaptation, integration, and expansion) were not included in this study as these focus on implementation and scale-up of the intervention, which was not yet relevant for the intervention under investigation. The domains of the framework by Bowen et al [20] were assessed using the outcomes of interest and outcome measures that are presented in Table 1.

XSL•FO
RenderX

Table . Outcome measures per domain of Bowen's framework for designing feasibility studies.

Bowen's domain, outcome of interest, and outcome measures	Subject[a]
Demand	
Actual use	
D[b]: Number of interactions	—[c]
D: Number of unique DRPs[d]	—
D: Type of DRPs	—
Expressed interest	
Q[e]: Subjective quality of text messaging (user-version Mobile Application Rating Scale - section E) [21]	Pt[f] and Ph[g]
Additional questions	
Q: Factors impacting future use of text messaging	Pt
Q: Factors ensuring future use of text messaging	Pt
Limited-efficacy testing	
Effects on key variables	
D: Number of DRPs solved	—
D: Number of DRPs in need of follow-up	—
D: Number of DRPs warranting HCP[h] involvement	—
Implementation	
Degree of execution	
D: Response within service level	—
Resources needed to implement	
D: HCP's time investment per conversation	—
Acceptability	
Satisfaction and appropriateness	
Q: Acceptability of text messaging (theoretical framework acceptability) [22]	Pt and Ph
Additional questions	
Q: Benefits and disadvantages of text messaging	Pt and Ph
Q: Types of DRPs acceptable or not acceptable to discuss using text messaging	Pt and Ph
Q: Preferred service level	Pt and Ph
Q: Preferred form of address	Pt
Q: Preferred alternative if text messaging was not available	Pt
Q: Fit of text messaging within current care	Pt and Ph
Q: Added value of text messaging to current care	Pt
Practicality	
Ability to carry out intervention activities	—
Q: System usability (Dutch version of the System Usability Scale) [23]	Pt and Ph
Additional questions	
Q: Positive and negative properties regarding ease of use of text messaging	Pt and Ph

[a]Outcome measures part of the questionnaire were answered by patients, pharmacists, or both.

[b]Outcome measures with prefix 'D' were collected by means of usage data.

XSL•FO
RenderX

[c]Not applicable.

[d]DRP: drug-related problem.

[e]Outcome measures with prefix 'Q' were collected by means of questionnaire data.

[f]Pt: patients.

[g]Ph: pharmacists.

[h]HCP: health care provider.

Demand

To measure demand for the intervention, the actual use of the intervention in terms of conversations and DRPs was assessed. In addition, the extent to which the intervention was likely to be used was assessed by determining participants' expressed interest in using the intervention, as well as factors impacting future use.

Limited-Efficacy Testing

To measure the efficacy of the intervention, the effect of the intervention on relevant key variables was measured. This was by measuring the extent to which discussed DRPs could be solved, the number of DRPs that needed follow-up from a health care provider other than a pharmacist, and the number of DRPs warranting health care provider involvement (ie, did an individual DRP need to be discussed with a pharmacist, or could this be solved by, for example, a frequently asked questions [FAQ] page).

Implementation

To measure implementation, the degree of execution was assessed by determining protocol fidelity, defined as the extent to which pharmacists responded to conversations within the service level. In addition, the resources needed to execute the intervention were assessed by calculating pharmacists' time investment per conversation.

Acceptability

To measure the acceptability of the intervention, satisfaction and appropriateness of text messaging with a pharmacist were assessed as part of the questionnaire sent to participants. In addition, participants' intention to (continue) using the intervention, satisfaction, preferred service level, preferred way patients wish to be addressed, preferred alternative if text messaging was not available, and added value of the text messaging to current care were assessed.

Practicality

To measure practicality, participants' ability to carry out intervention activities (ie, using the intervention) was assessed. For patients, this meant using text messaging to initiate conversations with pharmacists, and for pharmacists, this meant answering conversations initiated by patients.

Measurement Instruments

Study Sample

The patient characteristics that were retrieved from the electronic health record were age, sex, diagnosis, disease duration, number of drugs prescribed by a rheumatologist, and type of DMARDs at the start of the study.

Demand

Actual use was measured from usage data by counting the number of interactions initiated by the participants, the number of DRPs raised (as multiple DRPs can be raised in a single conversation), and the types of DRPs. To this end, individual DRPs were categorized according to the DRP classification as used by Haegens et al [4]. One researcher categorized DRPs (LH), which were reviewed by a second researcher (VH), after which discrepancies were discussed until consensus was reached. In case patients had not responded to the question of whether a DRP had been answered satisfactorily, one researcher (LH) determined if the DRP was resolved, after which a second researcher (VH) reviewed these cases, and discrepancies were discussed until consensus was reached. Expressed interest was assessed by incorporating the questions of the validated User version of the Mobile App Rating Scale [21] (uMARS) section E regarding subjective app quality into the questionnaire sent to the participants, consisting of 4 questions answered on 5-point Likert scales (Multimedia Appendix 2), which are reported individually. Other sections of the uMARS were not incorporated in this questionnaire, as they aim to assess factors other than demand for the intervention or other domains that are not of interest in this study. In addition, participants were asked about factors that positively or negatively impacted their future use of text messaging and how future use of text messaging can be ensured.

Limited-Efficacy Testing

DRPs were regarded as solved either when participants answered a confirmatory question at the end of every conversation positively or, in the case participants did not answer this question. In total, 2 researchers (LLH and VJBH) determined if DRPs were solved separately, after which discrepancies were discussed until consensus was reached. The number of DRPs that involved follow-up with a health care provider other than the pharmacist, or contained advice to do so, was counted from usage data. For each DRP, it was determined if raised DRPs warranted health care provider involvement by 2 researchers (LLH and VJBH), determining if individual DRPs needed to be discussed with a pharmacist or could be solved by, for example, an FAQ page independently, after which discrepancies were discussed until consensus was reached.

Implementation

Degree of execution was assessed by determining if pharmacists responded to individual conversations within the predefined service level of 4 hours, as well as pharmacists' average response time. To measure this, only working hours (ie, time between 9 AM and 5 PM on Monday to Friday) were counted, and hours outside working hours were not. Pharmacists' time investment per conversation (ie, including time for researching answers and contacting other health care providers) was

XSL•FO
RenderX

registered by the pharmacist responding to conversations using a stopwatch.

Acceptability

Acceptability was measured using the questionnaire from the theoretical framework of acceptability [22], which consists of questions targeting 8 acceptability constructs answered on 5-point Likert scales (Multimedia Appendix 2) and are reported individually. In addition, open-ended questions regarding the intention to (continue) using the intervention, participants' satisfaction, the preferred service level (ie, within what timeframe do patients expect to receive an answer from a pharmacist), patients' preferred form of address, what alternatives patients would use instead of text messaging, and the added value of the text messaging to current care were constructed based on the framework by Bowen et al [20] and its description of the acceptability domain. All questions regarding acceptability were constructed by one researcher (LH) and discussed among coauthors (LH, BB, and VH) until consensus was reached.

Practicality

Participants' ability to carry out intervention activities was assessed by using the validated Dutch translation of the System Usability Scale (SUS) [23], which is a 10-item questionnaire scored on a 5-point Likert scale from "strongly disagree" to "strongly agree" (Multimedia Appendix 2) that results in a score between 0 and 100. These questions were supplemented with open-ended questions regarding the main benefits and disadvantages of text messaging with a pharmacist, the main positive and negative properties regarding the ease of use of text messaging, and the fit of text messaging with current care.

Data Collection

Data on the feasibility of the intervention were collected either via a web-based questionnaire sent to all participants (both patients and pharmacists) or from usage data from the text messaging app. Furthermore, 2 separate questionnaires were constructed for participating patients and pharmacists. These questionnaires consisted of the same domains and predominantly differed in wording (eg, asking questions vs answering questions). Some questions, however, were not included in the questionnaire for pharmacists due to nonrelevance (Multimedia Appendix 2). Participants received a notification from an electronic data capture system via email with the invitation to complete the questionnaire and were reminded after 1 week if needed. Usage data were collected from the system that the text messaging is part of, by extracting relevant data regarding individual conversations into a separate database.

Data Analysis

Participating patients were regarded as "active participants" if at least 1 conversation was initiated via text messaging and were otherwise regarded as "nonactive participants" and thus excluded from the analyses. Usage data and data from questionnaires were descriptively analyzed using STATA 17 (StataCorp LLC) [24]. Proportions were presented as percentages. Parametric data were presented as means with SDs; nonparametric data were presented as medians with IQRs. Questionnaires with standardized analysis were scored and reported accordingly.

Results

Population Characteristics

Of the 108 patients that approached for participation, 56% (n=60) of patients were included and signed informed consent with a median age of 58 (IQR 53-65) years, and 70% (n=42) were female (Table 2). The most frequent reasons reported by patients for not participating were not interested (n=13) or too busy (n=8). Of the 60 included patients, 75% (n=45) of patients (median age 57, IQR 52-65 y; and n=31, 69% female) raised at least 1 DRP via text messaging (hereinafter referred to as active participants) and received the questionnaire. Reasons for not actively participating were that participants forgot to ask questions or were unable to ask questions due to personal reasons. Of the active participants, 87% (n=39) completed the questionnaire fully (hereinafter referred to as responders), 4% (n=2) of active participants completed the questionnaire partly, of whom only the fully completed domains are reported. Furthermore, 5 pharmacists (median age 41, IQR 26-47 y; and n=1, 20% female) participated in this study by responding to the text messages.

XSL•FO
RenderX

Table . Population characteristics.

Characteristic	Active participants[a] (n=45)	Nonactive participants[b] (n=15)
Sex, n (%)		
Female	31 (70)	10 (64)
Age (y), median (IQR)	57 (52 - 65)	61 (54 - 65)
Diagnosis, n (%)		
Rheumatoid arthritis	24 (53)	11 (73)
Psoriatic arthritis	10 (22)	1 (7)
Axial spondyloarthritis	6 (13)	1 (7)
Other	5 (11)	1 (7)
Disease duration (y), median (IQR)	6 (2-10)	9 (2-15)
Drugs prescribed by rheumatologist, median (IQR)	3 (4-5)	3 (5-6)
DMARD[c] group of inclusion, n (%)		
csDMARD[d]	16 (36)	7 (47)
tsDMARD[e]	11 (24)	5 (33)
bDMARD[f]	18 (40)	2 (13)
Experience category, n (%)		
New user	23 (51)	7 (47)
Experienced user	22 (49)	8 (53)

[a]Active participants are defined as included patients that discussed at least one drug-related problem during the study using text messaging.

[b]Nonactive participants are defined as included patients that did not raise a drug-related problem during the study.

[c]DMARD: disease-modifying antirheumatic drug.

[d]csDMARD: conventional synthetic disease-modifying antirheumatic drug.

[e]tsDMARD: targeted synthetic disease-modifying antirheumatic drug.

[f]bDMARD: biological disease-modifying antirheumatic drug.

Demand

During the study, 45 participants initiated 133 conversations via text messaging containing 158 unique DRPs (ie, DRPs that had not been raised before by individual participants), with a median of 3 (IQR 2-4) unique DRPs per participant. DRPs reported concerned information needs (n=42, 27%), (suspected) side effects (n=36, 23%), problems regarding medication management (n=34, 22%), logistics (n=14, 9%), medication concerns (n=13, 8%), medication effectiveness (n=9, 6%), contraindications (n=5, 3%), and others (n=5, 3%).

In total, 41 patients (91%) and 5 pharmacists (100%) completed the questions regarding demand. Demand was measured as the expressed interest or intention to use text messaging (uMARS section E). Both patients and pharmacists scored text messaging overall with a median of 4 (IQR 4-5) out of 5 stars. Out of 41 patients, 37 (90%) patients would recommend text messaging to others who could benefit from this service, and all 5 pharmacists would recommend text messaging to others. Furthermore, 63% (26/41) of patients would use text messaging at least once a year, and 32% (13/41) would consider paying for text messaging.

To ensure future use, 39% (16/41) of patients mentioned promotion of text messaging as a means to discuss DRPs, 20% (8/41) providing good instructions regarding and support with text messaging, and 17% (7/41) reducing the number of steps to navigate to the text messaging within the online patient portal as improvements. Furthermore, 34% (14/41) of patients mentioned quick answers, 27% (11/41) mentioned text messaging being easy to use, and 15% (6/41) receiving clear answers to questions as factors that would positively impact future use. Factors negatively impacting future use included text messaging being impractical (2/41, 5%) and impersonal (2/41, 5%).

Limited-Efficacy Testing

Of the 158 unique DRPs raised by patients, 58% (n=92) were solved according to patient confirmation, and the remaining 42% were solved according to pharmacist determination. In total, 5 reported DRPs were followed up in consultation with other health care providers: 3 with a rheumatologist, 1 with a specialized rheumatology nurse, and 1 with a pharmacy technician. In addition, 17 DRPs resulted in pharmacists advising follow-ups, mainly with the patient's rheumatologist (n=12). It was determined that 77% (122/158) of DRPs required a health care provider, and 23% (36/158) could have been solved

XSL•FO
RenderX

without health care provider involvement. Within DRPs that required health care provider involvement, the most common types of DRPs were suspected side effects (32/122, 26%), medication management (30/122, 25%), and information needs (16/122, 13%). DRPs that did not require health care provider involvement consisted of information needs (26/36, 72%), suspected side effects (4/36, 11%), and medication management (4/36, 11%). All types of DRPs were determined to require health care provider involvement for the majority of DRPs within their type (ranging from 30/34, 88% to N=13, 100%), except for DRPs regarding information needs, for which 2 out of 3 DRPs did not require health care provider involvement.

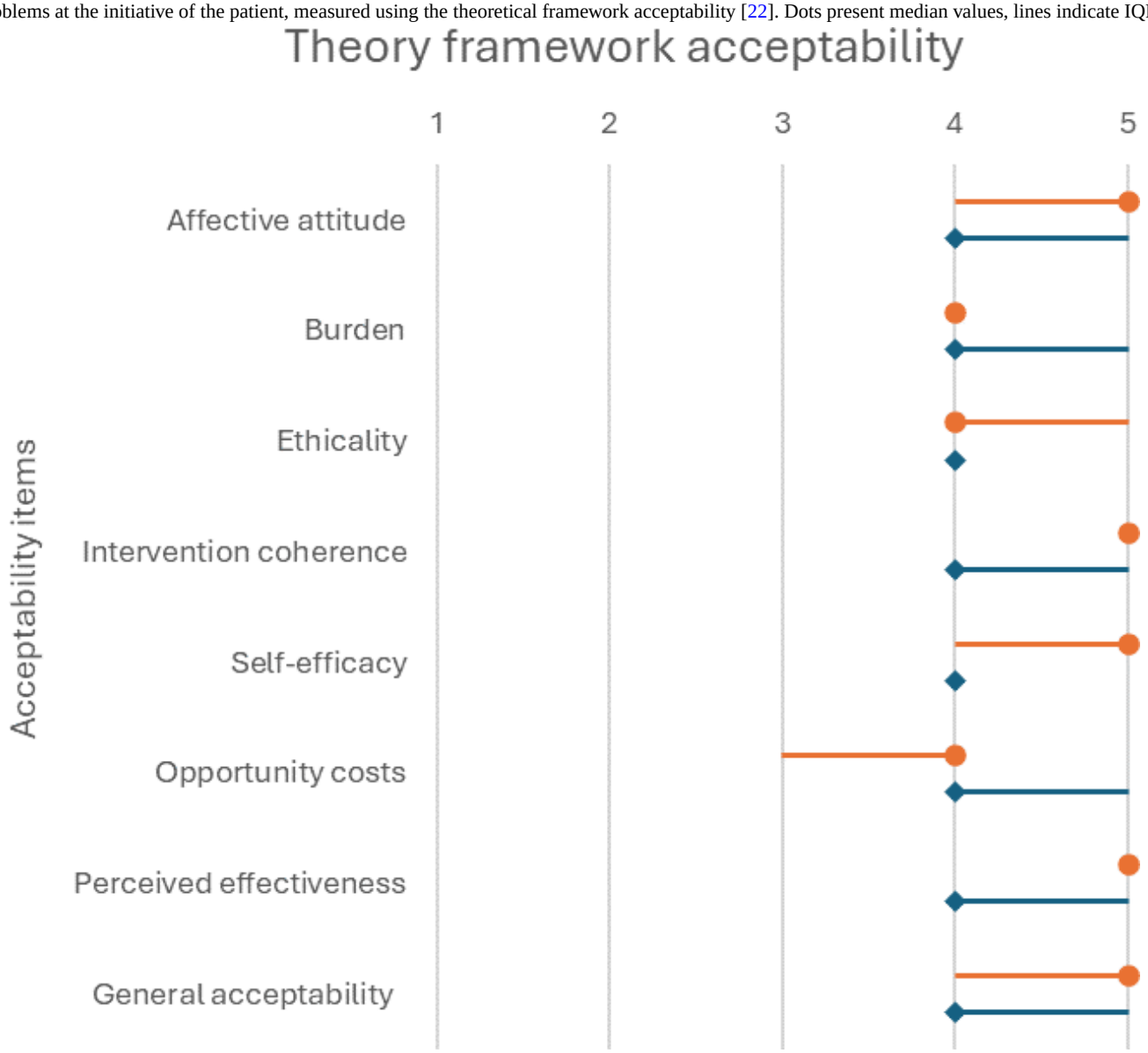
Implementation

Of the 133 conversations initiated by patients, 116 (87%) conversations were answered within the predefined service level of 4 hours. The median time until response was 80 (IQR 34-153) minutes with a total range of 0 minutes to 39.2 hours within opening hours of the pharmacy. The pharmacists' median time investment per conversation was 4:21 (IQR 2:15-7:41) minutes for DRPs that required health care provider involvement and 3:32 (IQR 2:39-6:41) minutes for DRPs that did not require health care provider involvement, which did not differ significantly (P=.10).

Acceptability

From the 45 active participants, 39 patients (87%) and all 5 pharmacists completed the questions regarding acceptability. Patients scored acceptability of the intervention with a median score of 4 (IQR 4-5) out of 5, and pharmacists with a median of 5 (IQR 4-5) (Figure 1).

Figure 1. Patients' (orange) and pharmacists' (blue) acceptability scores for text messaging between patients and pharmacists to discuss drug-related problems at the initiative of the patient, measured using the theoretical framework acceptability [22]. Dots present median values, lines indicate IQRs.



The top 3 most-mentioned advantages and disadvantages regarding the acceptability of text messaging mentioned by participating patients and pharmacists, as well as types of DRPs that are acceptable and not acceptable to discuss via text messaging, are presented in Table 3.

XSL•FO
RenderX

Table . Top 3 (where available) most-mentioned responses regarding advantages and disadvantages of text messaging, as well as types of DRPs acceptable and not acceptable for discussing via text messaging.

Questionnaire item, respondent, and response	Value, n (%)
Advantages regarding acceptability of text messaging	
Patients (n=39)	
Quick answers	21 (54)
Clear answers	9 (17)
No waiting times	7 (13)
Pharmacists (n=5)	
Accessibility of contact between patient and pharmacist	4 (80)
Sufficient time to formulate answer	3 (60)
Disadvantages regarding acceptability of text messaging	
Patients (n=39)	
Difficulties with logging in	4 (10)
Too many steps to reach text messaging	2 (5)
Less personal than contact via telephone	1 (3)
Pharmacists (n=5)	
Answering messages not yet embedded in daily routine	2 (40)
Contact becoming too accessible might lead to unnecessary contact	1 (20)
Questions need to be carefully answered, as these are in writing	1 (20)
Types of DRPs acceptable for discussing via text messaging	
Patients (n=39)	
Side effects	9 (23)
Everything regarding medication	4 (10)
Medication intake	3 (8)
Pharmacists (n=5)	
Factual questions regarding information	3 (60)
(Mild) side effects	2 (40)
Practical problems such as medication intake	2 (40)
Types of DRPs not acceptable for discussing via text messaging	
Patients (n=39)	
Questions for treating physician	1 (3)
Questions in need of in-person answer	1 (3)
Pharmacists (n=5)	
Insufficient medication effect	1 (20)
Allergies	1 (20)
Acute side effects	1 (20)
Emotionally-loaded questions	1 (20)

In total, 37 out of 45 patients (82%) and all pharmacists (N=5, 100%) completed the additional service-related questions. Regarding the preferred service level (ie, within what timeframe do patients prefer to receive a response to their text message), the majority of patients preferred a response within 4 (17/37, 46%) or 24 hours (12/37, 32%), with 2 patients (5%) indicating that the service level is dependent on individual DRPs. Pharmacists preferred a service level of 24 hours (2/5, 40%), 8 hours (1/5, 20%), or 4 hours (1/5, 20%), with 2 pharmacists (40%) suggesting letting the patient indicate the service level per DRP. Patients preferred to be addressed informally (15/37, 41%), formally (6/37, 16%), or did not have a preference (14/37, 38%). The majority of patients (24/37, 65%) would have called the pharmacy if text messaging were not available, and 11%

XSL•FO
RenderX

(4/37) indicated having contacted their health care provider but did not specify in which way. Another 11% (4/37) of patients indicated they would wait until the next consultation to discuss their DRP, and another 5% (2/37) would not discuss their DRP at all. Patients scored the extent to which text messaging fits into current care with a median of 4 (IQR 4-5), whereas pharmacists scored fit with a median of 4 (IQR 3-4). Pharmacists mentioned the inability to redirect questions to other health care providers and the fact that text messaging is not well-integrated into the existing information technology systems currently in use as disadvantages. In general, the added value of text messaging according to patients included quick answer to questions (7/37, 19%), having an additional way to ask questions (3/37, 8%), easy accessibility of contacting a pharmacist (2/37, 5%), and having an additional option for support when starting with new medication (2/37, 5%).

Practicality

Out of 45 patients, 39 (87%) and all 5 pharmacists (100%) completed the questions regarding practicality. The mean SUS score was 72 (SD 18) among patients, with 22 participants (57%) rating usability as good, excellent, or best imaginable. Mean SUS score was 81 (SD 16) among pharmacists, with 4 pharmacists (80%) rating usability as good, excellent, or best imaginable. The most-mentioned responses to the additional open-ended questions regarding practicality can be found in Table 4.

Table . Most-mentioned responses to additional open-ended questions regarding practicality.

Questionnaire item, respondent, and response	Value, n (%)
Properties positively influencing ease of use of text messaging	
Patients (n=39)	
Quick answer	14 (36)
Time- and place-independent	7 (18)
Easy to use	5 (13)
No waiting time	5 (13)
Pharmacists (n=5)	
Easy to use	3 (60)
Privacy-friendly	1 (20)
Time to formulate answers	1 (20)
Properties negatively influencing ease of use of text messaging	
Patients (n=39)	
Too many steps to reach text messaging within online patient-portal	8 (21)
Difficulties with logging in	6 (15)
Image size too small on mobile	3 (8)
Pharmacists (n=5)	
Absence of notifications of new messages	2 (40)
Less personal contact	1 (20)
Uncertainty as to whether DRP has been solved	1 (20)

Discussion

Principal Findings

In this study, the feasibility of text messaging at the initiative of the patient to discuss DRPs directly with pharmacists was examined. Text messaging between patient and pharmacist resulted in 158 unique DRPs in 133 conversations over an 8-week period. Pharmacists responded to all messages with advice to solve the problem or an answer to the question. All DRPs were solved, and most (116/133, 87%) were within the promised service level, and 14% (22/158) of DRPs resulted in follow-up or advised follow-up by health care providers other than pharmacists.

Both patients and pharmacists mentioned good accessibility of information and contact between patient and pharmacist as important factors for the acceptability of text messaging, which is also confirmed by other studies where accessibility was regarded as an important factor in communication between patient and pharmacist [25,26] as well as a driver for patients, preferring more accessible channels [12,27]. However, 1 pharmacist in our study noted that too easily accessible contact between patient and health care provider could lead to unnecessary contact (ie, contact regarding DRPs that do not warrant health care provider involvement) and thus an increased workload, which is a concern found in several other studies regarding the growing application of eHealth and its relation to health care providers' workload [28,29]. However, data on the

XSL•FO
RenderX

actual magnitude of this effect is lacking [30] and should thus be further researched.

In a study by Turcotte et al [25], patients mentioned the feeling of pharmacists being able to devote more time to answering their questions, which is in line with the pharmacists in our study mentioning having sufficient time to answer individual questions from patients when compared with current care. In a study by Zafar et al [26], patients were unsure about what type of questions were appropriate to ask via text messaging. Although patients in our study named several types of DRPs they find suitable to discuss, some differences exist between what patients and pharmacists view as appropriate DRPs to discuss via text messaging, with pharmacists making a distinction between the severity of side effects that patients do not make, which is important to align in managing patients' and pharmacists' expectations.

Although comparable studies did not employ the SUS, usability was rated "good" or "best imaginable" by the majority of patients and pharmacists, and both scored above the average SUS value of 68 [31]. Regarding the ease of using text messaging, which was in general mentioned as a positive factor driving acceptability, patients mentioned time- and place-independence and shorter waiting times of text messaging as positive properties. The SUS scores possibly find their origin in the comparability between the text messaging applied in this study with online and app-based consumer text messaging services. As some downsides regarding the practicality of the text messaging app used in this study were mentioned, it might be beneficial to organize and offer text messaging in a comparable manner to consumer apps to increase ease of use and thus acceptability, for example, by easing or removing the login procedure.

The top 3 most common types of patient-reported DRPs were information needs (42/158, 27%), (suspected) side effects (36/158, 23%), and medication management (34/158, 22%). These numbers are comparable with previous research regarding types of DRPs discussed with patients with inflammatory rheumatic diseases via 4 biweekly phone calls, in which (suspected) side effects accounted for 28% (85/308) and medication management for 26% (80/308) [4]. In our study, 27% (42/158) of DRPs were related to information needs, which is considerably higher than the 4% (13/308) of DRPs in the aforementioned study, which could be explained by the fact that DRPs in this study were patient-initiated, while the referred study assessed DRPs in health care provider–initiated conversations. In addition, the conceivability that patients regard information-related DRPs as more suitable to discuss via text messaging than via telephone might also have contributed to this. Although one might conclude that DRPs related to information needs could be solved by channels without health care provider involvement as this mainly pertains to information provision, 77% (122/158) of all DRPs were determined to warrant health care provider involvement, including 38% (16/42) of all DRPs related to information needs which mostly concerned needs for information regarding individual patient's treatment (eg, questions regarding safety of alternative medicine in combination with current therapy).

In practice, questions and DRPs can be answered and solved through various communication channels, such as during consultations, telephone calls, or with an FAQ page, for example. In this context, text messaging is a valuable addition to currently available channels, offering patients with a wide array of channels to discuss and resolve DRPs with high quality and efficiency, corresponding with patients' individual needs [12].

Implications

Patients in this study were generally positive regarding the feasibility of text messaging with pharmacists to discuss DRPs. The fact that 65% (24/37) of patients indicated that they would have raised their DRPs via telephone if text messaging had not been available implies that patients view text messaging as a suitable channel for raising DRPs in which health care provider involvement is warranted from the patient's perspective. In addition, 16% (6/37) of patients indicated that they would wait for the next consultation to raise a DRP if text messaging were not available, indicating DRPs possibly remain unraised and thus unsolved for a longer period of time. Although presumably part of these DRPs can wait until the next consultation, another part can lead to negative health consequences if not solved in a timely manner. Offering text messaging in addition to the currently available channels could lead to more DRPs raised, as well as in an earlier stage, which might therefore prevent negative health consequences associated with DRPs, as well as save time during consultations normally dedicated to such problems. Although all DRPs raised by patients were solved, either by patient confirmation or expert determination, the effect of this on patients' health outcomes was not assessed. Therefore, future research should aim to quantitatively assess the impact of the intervention on patients' clinical outcomes, such as medication adherence or improvement of health, as pharmacist interventions aimed at lowering DRPs seem to be able to improve clinical outcomes in other settings [32].

For health care providers, implementing text messaging can have several implications. First, the actual demand of text messaging in a large population will have an impact on its feasibility, as pharmacists must be able to incorporate the workload associated with text messaging into their daily routine. To this end, real-life demand needs to be further researched. As demand for text messaging is also related to the demand for other available channels, this should be researched in the context of all available channels rather than for text messaging alone. Second, although patients mentioned some advantages of text messaging over contacting a pharmacist via telephone, the extent to which text messaging can reduce the number of contacts via telephone or other channels and save time and resources is still unknown and should be researched in a trial with sufficient follow-up and a subsequent cost-benefit analysis. Third, although text messaging might initially increase pharmacists' workload, the potential of text messaging to save time and resources further down the chain (eg, by saving time during consultations) should be researched to determine the net decrease (or increase) in workload. Fourth, to avoid offering text messaging from leading to an additional workload, it is important to prevent (intentional or unintentional) inappropriate use of text messaging (ie, raising DRPs that could be solved

XSL•FO
RenderX

without health care provider involvement, for example, using an FAQ), which could be applicable for approximately 1 in 4 DRPs raised in this study that did not warrant health care provider involvement. Therefore, it is important to communicate to patients what DRPs are suitable to discuss via text messaging and what DRPs better suit different channels, or direct patients to the most suitable channel based on the subject of their DRP, for which DRPs warranting health care provider involvement in this study could be a first indication.

In terms of technological feasibility, several implications need to be considered to improve implementability and future uptake. Patients indicated that ease of use should be optimal for them to use text messaging most optimally. To this end, minor improvements should be made (eg, to the logging in procedure and the number of steps patients need to take) to be able to raise a DRP within the online patient portal. However, overall feasibility was scored as high, and 90% (37/41) of participants would recommend text messaging to others. Importantly, although pharmacists scored system usability above average and higher than patients, the fit with current care was scored lower than patients. Therefore, it is essential to look further into seamless integration of text messaging into the existing pharmacy systems and workflow, as well as ensuring seamless communication with other health care providers, to maximize its potential and ensure future use.

Limitations

Some limitations should be acknowledged. First, patients were reminded to ask a question at least once every 2 weeks to ensure user experience before completing the questionnaire. This could have led to DRPs being made up for study purposes, and thus we cannot conclude on the real-life demand that exists for text messaging from the patients' side in the researched population and setting. Second, 2 out of 5 pharmacists who answered patient messages were also involved in the conception of this study and as coauthors of this paper. This could have led to bias regarding outcome measures that included pharmacists' involvement, such as answering DRPs within the promised service level. However, as feasibility was researched in the context of current practice, the involvement of participating pharmacists in the conception of outcome measures was an added value, as they were able to include metrics that are needed for assessing feasibility and deciding future research and implementation. Third, selection bias might have been present in the study sample; for example, more digitally inclined or eHealth-literate patients could have agreed to participate than digitally-averse or eHealth-illiterate patients. This possibly led to an overestimation of the feasibility of text messaging from the patient's perspective, as these patients might have a more positive attitude toward digital tools and eHealth in general. Although 1 in 4 patients did not actively participate, these patients did not significantly differ from active participants in terms of sex, age, diagnosis, disease years, or prescribed DMARD group. Fourth, the response rate to the question of whether a DRP was answered sufficiently was low, with a nonresponse rate of 42%, resulting in a large number of DRPs determined to be answered sufficiently by the pharmacist's determination. Despite meticulous determination by 2 researchers (LLH and VJBH), this determination might not be as precise as patient confirmation. The relatively low response rate might be caused by the fact that the need to indicate if a question was answered satisfactorily after the conversation is concluded might not be intuitive and is normally not present in face-to-face conversations. Finally, the generalizability of the study is limited. As this study was performed in a specialized center rather than a general hospital and on a limited number of chronic conditions, standards of current (digital) care might differ from those in general hospitals and other chronic conditions. This can result in a different point of reference and thus make this study's findings not fully generalizable to other settings.

Conclusion

To conclude, 2-way text messaging between patients and pharmacists for discussing DRPs seems feasible from the patients' and pharmacists' perspectives. Furthermore, participants used text messaging primarily for raising DRPs, warranting health care provider involvement, which shows the potential of text messaging as an efficient addition to currently available information and communication channels if correctly integrated and intuitively usable. However, the impact of actual demand in a larger population on the feasibility for health care providers, as well as the impact on prevention of possible negative health consequences of DRPs, should be researched before full-scale implementation.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Blank informed consent form in Dutch.
[DOCX File, 14 KB - humanfactors_v12i1e66514_app1.docx]

Multimedia Appendix 2
Questionnaires.
[DOCX File, 17 KB - humanfactors_v12i1e66514_app2.docx]

References

XSL•FO
RenderX

1. Abbasi M, Mousavi MJ, Jamalzehi S, et al. Strategies toward rheumatoid arthritis therapy; the old and the new. J Cell Physiol 2019 Jul;234(7):10018-10031. [doi: 10.1002/jcp.27860] [Medline: 30536757]
2. Vrijens B, De Geest S, Hughes DA, et al. A new taxonomy for describing and defining adherence to medications. Br J Clin Pharmacol 2012 May;73(5):691-705. [doi: 10.1111/j.1365-2125.2012.04167.x] [Medline: 22486599]
3. Gil-Guillen VF, Balsa A, Bernárdez B, et al. Medication non-adherence in rheumatology, oncology and cardiology: a review of the literature of risk factors and potential interventions. Int J Environ Res Public Health 2022 Sep 23;19(19):12036. [doi: 10.3390/ijerph191912036] [Medline: 36231341]
4. Haegens LL, Huiskes VJB, Smale EM, Bekker CL, van den Bemt BJF. Drug-related problems reported by patients with rheumatic diseases: an observational study. BMC Rheumatol 2023 Apr 18;7(1):7. [doi: 10.1186/s41927-023-00326-x] [Medline: 37069634]
5. Leendertse AJ, Van Den Bemt PMLA, Poolman JB, Stoker LJ, Egberts ACG, Postma MJ. Preventable hospital admissions related to medication (HARM): Cost analysis of the HARM study. Value Health 2011 Jan;14(1):34-40. [doi: 10.1016/j.jval.2010.10.024]
6. Magdelijns FJH, Stassen PM, Stehouwer CDA, Pijpers E. Direct health care costs of hospital admissions due to adverse events in the Netherlands. Eur J Public Health 2014 Dec;24(6):1028-1033. [doi: 10.1093/eurpub/cku037] [Medline: 24699427]
7. Huiskes VJB, Cramer-van der Welle CM, van den Ende CHM, et al. Communication about drug-related problems (DRPs) during patients' visits to Dutch physicians and pharmacies. Health Commun 2020 Feb;35(2):201-208. [doi: 10.1080/10410236.2018.1551301] [Medline: 30526088]
8. Mathijssen EG, Vriezekolk JE, Eijsbouts AM, van den Hoogen FH, van den Bemt BJ. Support needs for medication use and the suitability of eHealth technologies to address these needs: a focus group study of older patients with rheumatoid arthritis. Patient Prefer Adherence 2018;12:349-358. [doi: 10.2147/PPA.S152759] [Medline: 29563778]
9. Benefits of eHealth. Government of the Netherlands. URL: https://www.government.nl/topics/ehealth/benefits-of-ehealth [accessed 2024-05-05]
10. Kwint HF, Faber A, Gussekloo J, Bouvy ML. The contribution of patient interviews to the identification of drug-related problems in home medication review. J Clin Pharm Ther 2012 Dec;37(6):674-680. [doi: 10.1111/j.1365-2710.2012.01370.x] [Medline: 22861493]
11. Kari H, Kortejärvi H, Airaksinen M, Laaksonen R. Patient involvement is essential in identifying drug ‐ related problems. Brit J Clinical Pharma 2018 Sep;84(9):2048-2058. [doi: 10.1111/bcp.13640]
12. Haegens LL, Huiskes VJB, van der Ven J, van den Bemt BJF, Bekker CL. Factors influencing preferences of patients with rheumatic diseases regarding telehealth channels for support with medication use: qualitative study. JMIR Form Res 2023 Jul 20;7:e45086. [doi: 10.2196/45086] [Medline: 37471137]
13. Greenwood DA, Hankins AI, Parise CA, Spier V, Olveda J, Buss KA. A comparison of in-person, telephone, and secure messaging for type 2 diabetes self-management support. Diabetes Educ 2014 Jul;40(4):516-525. [doi: 10.1177/0145721714531337] [Medline: 24742540]
14. Zhai P, Hayat K, Ji W, et al. Efficacy of text messaging and personal consultation by pharmacy students among adults with hypertension: randomized controlled trial. J Med Internet Res 2020 May 20;22(5):e16019. [doi: 10.2196/16019] [Medline: 32432556]
15. Gautier JF, Boitard C, Michiels Y, Raymond G, Vergez G, Guedon G. Impact of personalized text messages from pharmacists on medication adherence in type 2 diabetes in France: A real-world, randomized, comparative study. Patient Educ Couns 2021 Sep;104(9):2250-2258. [doi: 10.1016/j.pec.2021.02.022] [Medline: 33750593]
16. Buck C, Keweloh C, Bouras A, Simoes EJ. Efficacy of short message service text messaging interventions for postoperative pain management: systematic review. JMIR Mhealth Uhealth 2021 Jun 16;9(6):e20199. [doi: 10.2196/20199] [Medline: 34132646]
17. Grove ML, Hassell AB, Hay EM, Shadforth MF. Adverse reactions to disease-modifying anti-rheumatic drugs in clinical practice. QJM 2001 Jun 1;94(6):309-319. [doi: 10.1093/qjmed/94.6.309]
18. Julious SA. Sample size of 12 per group rule of thumb for a pilot study. Pharm Stat 2005 Oct;4(4):287-291. [doi: 10.1002/pst.185]
19. Teare MD, Dimairo M, Shephard N, Hayman A, Whitehead A, Walters SJ. Sample size requirements to estimate key design parameters from external pilot randomised controlled trials: a simulation study. Trials 2014 Jul 3;15(1):264. [doi: 10.1186/1745-6215-15-264] [Medline: 24993581]
20. Bowen DJ, Kreuter M, Spring B, et al. How we design feasibility studies. Am J Prev Med 2009 May;36(5):452-457. [doi: 10.1016/j.amepre.2009.02.002] [Medline: 19362699]
21. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the user version of the Mobile Application Rating Scale (uMARS). JMIR Mhealth Uhealth 2016 Jun 10;4(2):e72. [doi: 10.2196/mhealth.5849] [Medline: 27287964]
22. Sekhon M, Cartwright M, Francis JJ. Development of a theory-informed questionnaire to assess the acceptability of healthcare interventions. BMC Health Serv Res 2022 Dec;22(1). [doi: 10.1186/s12913-022-07577-3]
23. Ensink CJ, Keijsers NLW, Groen BE. Translation and validation of the System Usability Scale to a Dutch version: D-SUS. Disabil Rehabil 2024 Jan;46(2):395-400. [doi: 10.1080/09638288.2022.2160837] [Medline: 36573399]

XSL•FO
RenderX

24. Stata statistical software: release 17. : StataCorp LLC; 2021.
25. Turcotte V, Chagnon A, Guénette L. Experience and perspectives of users and non-users of the Ask your pharmacist teleconsultation platform. Explor Res Clin Soc Pharm 2021 Jun;2:100031. [doi: 10.1016/j.rcsop.2021.100031] [Medline: 35481130]
26. Zafar SN, Hazlewood G, Dhiman K, et al. “How are you?” Perspectives from patients and health care providers of text messaging to support rheumatoid arthritis care: a thematic analysis. ACR Open Rheumatol 2024 May;6(5):276-286. [doi: 10.1002/acr2.11652] [Medline: 38376004]
27. van der Ven J, van den Bemt BJF, van Dijk L, et al. Preferences of patients with musculoskeletal disorders regarding the timing and channel of eHealth and factors influencing its use: mixed methods study. JMIR Hum Factors 2023 Sep 27;10:e44885. [doi: 10.2196/44885] [Medline: 37756049]
28. MacLure K, Stewart D, Strath A. A systematic review of medical and non-medical practitioners’ views of the impact of eHealth on shared care. Eur J Hosp Pharm 2014 Feb;21(1):54-62. [doi: 10.1136/ejhpharm-2013-000337]
29. Ariens LF, Schussler-Raymakers FM, Frima C, et al. Barriers and facilitators to eHealth use in daily practice: perspectives of patients and professionals in dermatology. J Med Internet Res 2017 Sep 5;19(9):e300. [doi: 10.2196/jmir.7512] [Medline: 28874336]
30. Stevens WJM, van der Sande R, Beijer LJ, Gerritsen MG, Assendelft WJ. eHealth apps replacing or complementing health care contacts: scoping review on adverse effects. J Med Internet Res 2019 Mar 1;21(3):e10736. [doi: 10.2196/10736] [Medline: 30821690]
31. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009;4(3):114-123 [FREE Full text]
32. Ayhan YE, Aksoy M, Abdulrahman Y, et al. Evaluation of the effect of management of drug-related problems on clinical outcomes of pulmonary embolism outpatients: a randomized controlled trial. J Clin Med 2025 Feb 12;14(4):1202. [doi: 10.3390/jcm14041202] [Medline: 40004736]

Abbreviations

bDMARD: biological disease-modifying antirheumatic drug
DMARD: disease-modifying antirheumatic drug
DRP: drug-related problem
FAQ: frequently asked question
SUS: System Usability Scale
uMARS: user-version of the Mobile Application Rating Scale

Edited by A Kushniruk; submitted 15.09.24; peer-reviewed by G Whelan, G Li; revised version received 23.04.25; accepted 05.05.25; published 08.10.25.

Please cite as:
Haegens LL, Bekker CL, Flendrie M, van den Bemt BJF, Huiskes VJB
Text Messaging Between Patients With Inflammatory Rheumatic Diseases and Pharmacists to Solve Drug-Related Problems: Prospective Feasibility Study
JMIR Hum Factors 2025;12:e66514
URL: https://humanfactors.jmir.org/2025/1/e66514
doi:10.2196/66514

© Lex L Haegens, Charlotte L Bekker, Marcel Flendrie, Bart J F van den Bemt, Victor J B Huiskes. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 8.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Intention to Use Automated Diagnosis and Clinical Risk Perceptions Among First Contact Clinicians in Resource-Poor Settings: Questionnaire-Based Study Focusing on Acute Burns

Constance Boissin[1,2], PhD; Lisa Blom[2], PhD; Zara Taha[2], MSc; Lee Wallis[3], MBChB, MD; Nikki Allorto[4], MBChB, Mmed, FCS; Lucie Laflamme[2,5], PhD

[1]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

[2]Department of Global Public Health, Karolinska Institutet, Tomtebodavagen 18A, Widerstromska Huset, Stockholm, Sweden

[3]Division of Emergency Medicine, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa

[4]Pietermaritzburg Burn Service, University of Kwa-Zulu Natal, Durban, South Africa

[5]Institute for Social and Health Sciences, University of South Africa, Pretoria, South Africa

Corresponding Author:
Constance Boissin, PhD
Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

Abstract

Background: Burn automated diagnosis may be instrumental for accurate and timely decision-making at point-of-care, helping to ensure that the right patients are triaged to burns centers. This is particularly important in resource-poor settings.

Objective: We studied the intention of nonspecialized clinicians to engage in automated diagnosis in burn care as well as their perceptions toward clinical risks.

Methods: A self-administered survey was used among a purposive sample of first contact clinicians (n=56) and burns specialists (n=35). The survey had 2 main parts: 1 measuring the intention to use automated diagnosis as per 7 constructs of the Automation Acceptance Model (yielding 8 hypotheses) and 1 on clinical risk perceptions (likelihood and severity of 7 risks). Structural Equation Modelling was used to test the hypotheses among first contact clinicians, and the Mann-Whitney U test was used to measure differences in risk perceptions between the two clinical groups.

Results: Many first contact clinicians would intend to use automated diagnosis for burns should the technology be made available in their departments (41/56, 73%). The Automation Acceptance Model concepts contributed moderately to explain what the intention to use automated diagnosis rests on (R^2=0.432), with 5 out of 8 hypotheses being supported. The intention to use automated diagnosis was associated with perceived usefulness but not with attitudes toward using it. Of the 7 risks studied, the 1 that was most often considered as high risk of occurring was that of complex burns not being recognized (n=23, 29%). The 2 groups differed significantly in their concern regarding both the likelihood of happening and the severity of 2 risks: the undermanagement of severe burns and the overmanagement of minor burns. Specifically, a larger proportion of first contact clinicians were more concerned than burns specialists (n=13, 27% versus 6% and n=11, 23% versus 6% for undermanagement and overmanagement, respectively).

Conclusions: Almost three-quarters of first contact clinicians were inclined to seek automated advice for burn diagnosis. The proposed model contributes to explaining the intention to use with 5 hypotheses supported. When seeking additional determinants, clinical risk perception is a dimension that should be considered in any artificial intelligence implementation process, to help ensure sustainability.

(JMIR Hum Factors 2025;12:e56300) doi:10.2196/56300

KEYWORDS

artificial intelligence; burns; clinical decision support systems; clinical risk perception; diagnosis, computer-assisted; first contact clinicians; image-based diagnostics; intention to use; telemedicine; triage

Introduction

Image-based automated diagnosis is developing rapidly, not least with the access to deep neural networks and even more with recent methods that can improve performance, such as foundation models or vision transformers [1-3]. These algorithms can outperform clinicians in medical specialties such as cancer pathology [4], ophthalmology [5], dermatology [6], and burns [7]. However, implementation of the technology into the clinical setting has been relatively low and is dependent on

XSL•FO
RenderX

the medical specialty. For instance, results in the radiology field emphasized the need for grounded, well-structured, and hospital-wide implementation processes [8]. In surgical pathology, rapid and accurate results provided by automated diagnosis enabled pathologists to focus on more complex consultative tasks [9]. Further, maxillofacial surgeons, ophthalmologists, dermatologists, and radiation oncologists acknowledged the need to be trained before implementation and to have well-defined responsibilities among all users involved in the implementation process, especially if errors were to happen [10,11].

In Sub-Saharan Africa, the potential for artificial intelligence (AI) is big, as it could assist with early detection of infectious diseases [12,13], but also assist with the lack of specialists in underserved areas [14]. As an example, a solution for tuberculosis detection in chest x-rays has been proposed in South Africa [15].

Despite those promises, AI has not unleashed its full potential in the continent, with the uptake of this technology being relatively low at first point of care, perhaps because of limited involvement from health care specialists at the development stage [16,17]. Users' trust and task seamlessness [18] have been well described as uptake barriers since the introduction of telemedicine [19,20] and mHealth [21]. Gaining knowledge on first contact clinicians' intention to adopt automated diagnosis procedures is an essential milestone in their development and implementation process, allowing for user-sensitive transformations of diagnostic decision making in the clinical setting.

The visual nature of burns, the wide uptake of mHealth [22], the progress made in the technology [16], and the scarcity of specialists, not least in low resource settings [23], are important arguments in favor of implementing automated diagnosis in this clinical domain. This current study follows prior work showing that the technology performs well [16] and investigates the factors influencing the willingness of first contact clinicians to use automated diagnostic assistance in burn care. It adds to the body of knowledge generated from implementation science studies, setting the stage for translation from knowledge to practice and informing on end users' perceptions in view of a change in clinical procedures.

Methods

Aim

The primary aim of this study was to determine the factors influencing the intention to use an automated burn diagnosis tool among first contact clinicians. An additional aim was to assess whether the views of those clinicians are like those of burns specialists, in particular, how they perceive the likelihood and severity of certain clinical risks.

Study Design

This was a questionnaire-based cross-sectional study. The CHERRIES checklist for reporting results of e-surveys was used [24], with results presented in Table S1 in Multimedia Appendix 1.

Data Collection

In 2019, two consecutive international congresses took place in South Africa and gathered clinicians involved in burn injury care at different levels, both first-contact care and specialized care delivery. Those were judged good opportunities to reach out to several professionals from different parts of Sub-Saharan Africa and other parts of the world that, otherwise, would have been difficult to reach. The survey took place in the context of the development of an algorithm for image-based automated burn diagnosis [23]. The meetings were the Pan African and South African Burn Congress 2019 and the EMSSA (7th Emergency Medicine Society of South Africa) International Conference. At both events, participants were informed about this study taking place when they registered, and pamphlets containing information about the survey were distributed in sessions. Using a QR code provided, those who wished to participate could go on and fill in the survey on their own device or using one of the tablets provided by the research team.

Questionnaire

A questionnaire developed built on previous publications [20,21,25] was used to assess the intention to use an automated burn diagnosis tool among the participants based on the Automation Acceptance Model (AAM) [25]. For that purpose, the questionnaire was divided into 3 parts. Following some information on the context of automated burn diagnosis tool development, and with a specific description of what such a tool would consist of, the first section included demographic questions about the participant and its practice in burn care. The second part included 23 questions representing the 7 constructs from AAM [25], as well as anxiety [21,26,27] to reflect the intention to use an automated burn diagnosis tool in resource-poor settings in Africa (Table 1). The 7 constructs were behavioral intention (1 measurement item), attitude (3 items), perceived usefulness (8 items), perceived ease of use (2 items), trust (2 items), and compatibility (2 items) all directly inspired from the AAM and adapted to burns diagnosis and anxiety (5 items) which was shown to be predictive in similar populations for the introduction of telehealth. All constructs were measured on a 7-point Likert scale from 1 "strongly disagree" to 7 "strongly agree." The third part of the questionnaire evaluated the participant's perceptions of the risks linked with the use of an automated burn diagnosis tool. Seven risks were specifically evaluated: (1) that severe burns are missed, leading to undermanagement; (2) that minor burns are diagnosed as severe, leading to overmanagement; (3) that clinicians feel less responsible for the care they provide; (4) that the emotional well-being of the patient will be neglected; (5) that the patient confidentiality will be breached; (6) that patients will be less empowered in decisions around their care; (7) that some complex burns may not be recognized by the algorithm. For each risk, the participant had to evaluate the likelihood of happening on a 3-point scale from low or very low to high or very high; and the severity of the risk would it happen on a 4-point scale (insignificant, minor, moderate, or major).

XSL•FO
RenderX

Table . Measurement items used in the questionnaire in relation to the defined construct they measure.

Construct (definition)	Measurement items
Behavioral intention: The motivational factors that determine an individual's readiness to use the technology, which is assumed to be an immediate antecedent of behavior [28]	· I intend to use automated diagnosis for burns as soon as it becomes available in my department
Attitude: An individual's positive or negative feelings about using the technology [28]	· I believe that using automated diagnosis for burns is a good idea · Using automated diagnosis for burns would make work more interesting · I believe that using automated diagnosis for burns is a bad idea
Perceived usefulness: The degree to which an individual believes using the technology will assist in performing a task [29]	· It would make it easier to do my job when managing patients with burns · It could improve the care I give to my patients with burns · It will reduce the probability of medical errors in burn care · Using it will improve the quality of my burn-related work · It will be useful in my burn-related work · It will be useful in improving patient safety · I would use it for burn wounds that I normally would hesitate to consult a specialist about · Using it in my burn-related work would make me feel more independent
Perceived ease of use: The degree to which an individual believes using the technology will be effortless [29]	· Learning to operate a tool like this would be easy for me · It would be easy for me to become skilled at using such a tool
Trust: The generalized expectancy to which an individual can rely on the technology [25,30]	· I believe that the diagnostic ability of automated diagnosis is superior to the clinical experience of human doctors · I would have difficulties trusting an automated diagnosis
Compatibility: The degree to which the technology is believed to be consistent with an individual's existing values, needs, and experiences [25]	· Using a tool like this would fit well with the way I like to work · Using an app like this would be compatible with most aspects of my work
Anxiety: The degree to which emotional reactions are evoked when using the technology [27]	· I feel apprehensive about using such a tool · Using automated diagnosis for burns would make me feel uncomfortable · I am concerned about possible liability issues associated with its use · Working with such a tool would make me feel anxious · Using it would make me feel less anxious about dealing with burns patients

The questionnaire was pilot-tested before the data collection to ensure comprehensibility, content, and length of time.

Study Population

Overview

Two study groups were included in this study, all of whom were familiar with cases from Sub-Saharan Africa: first contact clinicians, as our focus, and burn specialists, as a basis of comparison.

First Contact Clinicians

First contact clinicians treating burn injuries at the front line in Sub-Saharan African countries were targeted because they would be the ones using such an image-based automated diagnostic assistance in burn care. The inclusion criteria were therefore that the participant was a doctor or a nurse who may need regular or occasional advice for referral, diagnosis, or treatment of burn injuries in their practice. Another eligibility criterion was that one had to work in an emergency or trauma center (where patients with acute burns are brought) or at a specialized surgical unit that manages patients with burns. Furthermore, the participant had to be practicing, or have practiced in a Sub-Saharan African country.

Burn Specialists

Those who took part in this study were burn specialists who have an expert role in their setting when it comes to burns assessment and treatment; they completed a limited part of the questionnaire (see below), where they were asked about potential risks that may accompany the use of automated diagnosis for burns.

A total of 122 participants started the survey, of which 31 were excluded because: they did not complete the consent form (n=6), they did not complete the survey (n=18), they only practiced in a high income country (n=3), or they were prehospital paramedics (n=4). The analyses were performed on the remaining 91 participants.

Data Analyses

Intention to Use Automated Diagnosis for Burn Care

To explain the intention to use automated diagnosis for burn care, 8 hypotheses (presented in Table 2) were proposed based on the 7 constructs previously defined. To evaluate the accuracy and consistency of the developed measurement items with regard to their respective constructs, Cronbach α and composite reliability were measured. Any constructs that had a Cronbach

XSL•FO
RenderX

α lower than 0.7 or a composite reliability value lower than 0.6 were excluded from further analyses [31]. The results are presented in Table 3 and show that all constructs are consistent and reliable except for trust, for which the 2 measurement items do not yield an appropriate Cronbach α. This construct was therefore excluded from further analyses and the model.

Table . Hypotheses tested, definitions, and results obtained.

Hypothesis	Definition	Supported by the results
1	Perceived compatibility is positively related to the perception of the usefulness of automated diagnosis for burn care.	Yes
2	Perceived compatibility is positively related to the perception of the ease of use of automated diagnosis for burn care.	Yes
3	Perceived ease of use of automated diagnosis for burn care is positively related to the perception of its usefulness.	No
4	Perceived usefulness of automated diagnosis for burn care is positively related to attitudes toward using it.	Yes
5	Perceived ease of use of automated diagnosis for burn care is positively related to attitudes toward using it.	No
6	Anxiety is negatively related to attitudes toward using automated diagnosis for burn care.	Yes
7	Perceived usefulness of automated diagnosis for burn care is positively related to the behavioral intention to use it.	Yes
8	Positive attitudes toward using automated diagnosis for burn care are positively related to the behavioral intention of using it.	No

Table . Constructs summary (number of items and mean), consistency statistic (Cronbach α), and composite reliability measure.

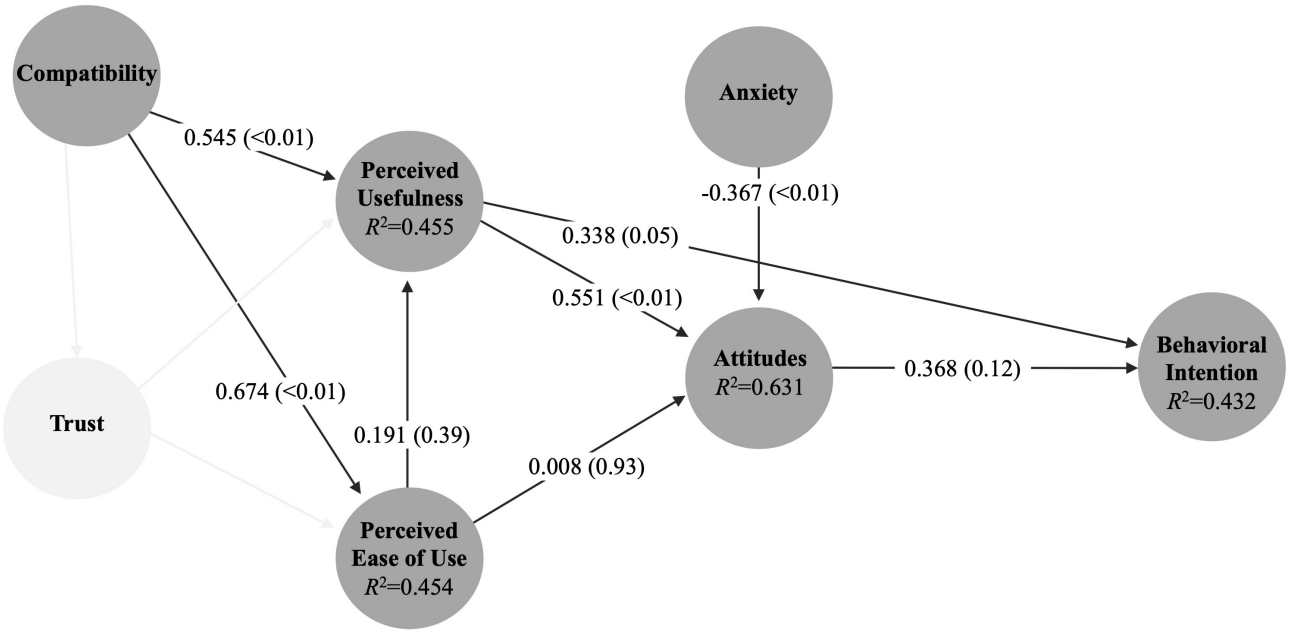
Construct	Number of items	Mean (SD)	Cronbach α	Composite reliability
Behavioral Intention	1	5.4 (1.6)	—[a]	—
Attitudes	3	5.4 (1.1)	0.732	0.846
Perceived usefulness	8	5.3 (1)	0.918	0.934
Perceived ease of use	2	5.7 (1.2)	0.866	0.937
Trust	2	4 (1)	−0.658	0.332
Compatibility	2	5.4 (1.3)	0.886	0.946
Anxiety	5	4.7 (1.1)	0.701	0.803

[a]Not applicable.

We then used Partial Least Squares Structural Equation Modelling to examine the proposed model (Figure 1) and test the hypotheses. Partial Least Squares Structural Equation Modelling is a multivariate statistical analysis that allows for testing models with an established theoretical foundation and with latent variables [32]. Bias-corrected and accelerated bootstrapping using 5000 subsamples was used to measure statistical significance. We used a hard cutoff to define whether a hypothesis was supported or not, which was an obtained value higher than 0.2 with a P value <.05 [20]. Analyses were performed using SmartPLS (version 4.0).

XSL•FO
RenderX

Figure 1. Seven hypotheses generated by the obtained AAM model and their respective path coefficients. AAM: Automation Acceptance Model.



Risks That May Occur With Implementation

Both first contact clinicians and burns specialists completed the survey section on potential risks. For each of these risks, the proportion of answers was reported in 3 categories. For the severity of each risk, the categories of insignificant or minor risk were combined and collapsed into one. A nonparametric Mann-Whitney U test was used to assess the differences in opinions between the 2 groups. For that, all data points for the 2 groups are ranked in ascending order, and the sum of the ranks (R1 and R2) are then calculated for both groups separately (first contact clinicians and burns specialists). The U statistic is then calculated for each group:

$U1=n1 \cdot n2+n1n1+12-R1$ for group 1 and $U2=n1 \cdot n2+n22+12-R2$ for group 2

where n1 and n2 are the sizes of the 2 groups. The smaller U value is then used to assess significance with a defined *P* value threshold of .05.

All analyses were performed using Stata BE (version 17.0; StataCorp LLC).

Ethical Considerations

This study was approved in South Africa by the Human Research Ethics Committee of the University of Cape Town (Dnr 823/2019) and in Sweden by the Ethics Committee (Etikprövningsmyndigheten [Dnr 2019 - 05122]). Informed consent was provided by participants prior to filling in the online survey. The data are deidentified and no compensation was provided for participating.

Results

Of the 91 respondents who completed at least 1 section of the survey, 62% (56/91) were first contact clinicians and 38% (35/91) were burn specialists. Table 4 presents the respective characteristics of the 2 groups with regard to their demographics and experience with burn care. Overall, these were mostly men (49/91, 54%), although the opposite was observed at first contact care, with over half being women (31/56, 55%). A total of 52% (29/56) of first contact clinicians managed patients with burns at no or few occasions per week, whereas 20% (7/35) of burns specialists manage patients at least once a week.

XSL•FO
RenderX

Table . Characteristics of participants by clinical expertise.

Characteristic	All (n=91), n (%)	First contact clinicians (n=56), n (%)	Burns specialists (n=35), n (%)
Sex			
Men	49 (53.9)	25 (44.6)	24 (68.6)
Women	42 (46.2)	31 (55.4)	11 (46.2)
Age (in years)			
<35	22 (24.2)	14 (25)	8 (22.9)
35 - 39	27 (29.7)	21 (37.5)	6 (17.1)
40 - 44	21 (23.1)	15 (26.8)	6 (17.1)
≥45	21 (23.1)	6 (10.7)	15 (42.9)
Occupation			
Nurse	17 (18.7)	13 (23.2)	4 (11.4)
Medical officer	20 (22)	12 (21.4)	8 (22.9)
Registrar	21 (23.1)	16 (28.6)	5 (14.3)
Specialist	33 (36.7)	15 (26.8)	18 (51.4)
Experience with burn treatment			
None or minimal	18 (19.8)	17 (30.4)	1 (2.9)
Moderate	55 (60.4)	34 (60.7)	21 (60)
Extensive	18 (19.8)	5 (8.9)	13 (37.1)
Frequency of managing patients with burns			
No or few occasions	27 (31.9)	29 (51.8)	0 (0)
A few per month	52 (55)	22 (39.3)	28 (80)
A few per week or daily	12 (13.2)	5 (8.9)	7 (20)
Confident in own burn management			
Not at all	4 (4.4)	4 (7.1)	0 (0)
Somewhat	55 (60.4)	39 (69.6)	16 (45.7)
Very	32 (35.2)	13 (23.2)	19 (54.3)

A total of 41 of the 56 (73%) first contact clinicians replied that they agreed or strongly agreed that they would use automated diagnosis for burns if it becomes available in their department. Figure 1 presents the results of the tested model for path coefficients. While perceived usefulness was positively and significantly related to the intention to use, attitudes toward the practice were not. Together, these 2 constructs explain 43% of the variance in intention to use for clinicians at the point-of-care. Almost two-thirds of the variance in attitudes toward the practice (R^2=0.631) could be explained by the 3 constructs of anxiety, perceived usefulness, and perceived ease of use. For its part, compatibility was positively related to the perception of both usefulness and ease of use. Overall, out of the 8 tested hypotheses, 5 were supported in the model (Table 2).

Table 5 indicates the position of the 2 groups of professionals concerning the likelihood of different risks. All respondents aggregated, the risk of complex burns not being recognized and that of clinicians feeling less responsible for the care they provide are the 2 risks that most often were considered as high or very high (23/79, 29% and 19/85, 22%, respectively). There are 2 risks where the opinions of the 2 groups differ significantly, and they concern the undermanagement of severe burns and the overmanagement of minor burns. In both instances, the risk is seen as high or very high by more first contact clinicians than burn specialists.

XSL•FO
RenderX

Table . Likelihood of each risk to happen as perceived by clinicians at first contact and burn specialists.

Risks assessed for likelihood	Do not know, n	Low or very low, n (%)	Moderate, n (%)	High or very high, n (%)	Mann-Whitney *P* value
That severe burns are missed, leading to undermanagement					.03
First contact clinicians	5	15 (31.3)	20 (41.7)	13 (27.1)	
Burn specialists	0	16 (45.7)	17 (48.6)	2 (5.7)	
All participants	5	31 (37.4)	37 (44.6)	15 (18.1)	
That minor burns are diagnosed as severe, leading to overmanagement					<.01
First contact clinicians	4	14 (28.6)	24 (49)	11 (22.5)	
Burn specialists	1	21 (61.8)	11 (32.4)	2 (5.9)	
All participants	5	35 (42.2)	35 (42.2)	13 (15.7)	
That clinicians feel less responsible for the care they provide					.11
First contact clinicians	3	25 (50)	11 (22)	14 (28)	
Burn specialists	0	23 (65.7)	7 (20)	5 (14.3)	
All participants	3	48 (56.5)	18 (21.2)	19 (22.4)	
That the emotional well-being of the patient will be neglected					.98
First contact clinicians	1	35 (67.3)	11 (21.2)	6 (11.5)	
Burn specialists	1	23 (67.7)	7 (20.6)	4 (11.8)	
All participants	2	58 (67.4)	18 (20.9)	10 (11.6)	
That the patient confidentiality will be breached					.38
First contact clinicians	3	31 (62)	16 (32)	3 (6)	
Burn specialists	0	20 (57.1)	8 (22.9)	7 (20)	
All participants	3	51 (60)	24 (28.2)	10 (11.8)	
That patients will be less empowered in decisions around their care					.22
First contact clinicians	2	35 (68.6)	10 (19.6)	6 (11.8)	
Burn specialists	1	27 (79.4)	6 (17.7)	1 (2.9)	
All participants	3	62 (72.9)	16 (18.8)	7 (8.2)	
That some complex burns may not be recognized by the algorithm					.2
First contact clinicians	8	6 (13.3)	24 (53.3)	15 (33.3)	
Burn specialists	1	8 (23.5)	18 (52.9)	8 (23.5)	
All participants	9	14 (17.7)	42 (53.2)	23 (29.1)	

Table 6 presents the perceived severity of each risk for both groups of respondents. For all risks, the severity is perceived as major to a higher proportion of first contact clinicians than burn specialists. Just as for the likelihood, all respondents aggregated, the risk for complex burns not being recognized is the risk for which severity is most often considered as major (45/82, 55%). A large proportion of clinicians (46/85, 54%) also rated as major the severity of the risk for severe burns being missed, although this is true especially for those at point of care (32/51, 63%). While almost 80% (27/34) of burns specialists considered the severity of minor burns being diagnosed as minor or insignificant, only 40% (21/52) of first contact clinicians did.

XSL•FO
RenderX

Table . Severity of each of the risks would it be perceived by first contact clinicians and burn specialists.

Risks assessed for severity	Do not know n	Insignificant or minor, n (%)	Moderate, n (%)	Major, n (%)	Mann-Whitney *P* value
That severe burns are missed, leading to undermanagement					.02
First contact clinicians	2	6 (11.8)	13 (25.5)	32 (62.8)	
Burn specialists	1	11 (32.4)	9 (26.5)	14 (41.2)	
All participants	3	17 (20)	22 (25.9)	46 (54.1)	
That minor burns are diagnosed as severe, leading to overmanagement					<.01
First contact clinicians	1	21 (40.4)	24 (46.2)	7 (13.5)	
Burn specialists	1	27 (79.4)	7 (20.6)	0 (0)	
All participants	2	48 (55.8)	31 (36.1)	7 (8.1)	
That clinicians feel less responsible for the care they provide					.07
First contact clinicians	2	17 (33.3)	17 (33.3)	17 (33.3)	
Burn specialists	1	17 (50)	11 (32.4)	6 (17.7)	
All participants	3	34 (40)	28 (32.9)	23 (27.1)	
That the emotional well-being of the patient will be neglected					.14
First contact clinicians	4	19 (38.8)	17 (34.7)	13 (26.5)	
Burn specialists	1	19 (55.9)	9 (26.5)	6 (17.7)	
All participants	5	38 (45.8)	26 (31.3)	19 (22.9)	
That the patient confidentiality will be breached					.05
First contact clinicians	1	21 (40.4)	12 (23.1)	19 (36.5)	
Burn specialists	1	20 (58.8)	8 (23.5)	6 (17.7)	
All participants	2	41 (47.7)	20 (23.3)	25 (29.1)	
That patients will be less empowered in decisions around their care					.03
First contact clinicians	3	22 (44)	18 (36)	10 (20)	
Burn specialists	1	23 (64.7)	11 (32.4)	1 (2.9)	
All participants	4	44 (52.4)	29 (34.5)	11 (13.1)	
That some complex burns may not be recognized by the algorithm					.096

XSL•FO
RenderX

Risks assessed for severity	Do not know n	Insignificant or minor, n (%)	Moderate, n (%)	Major, n (%)	Mann-Whitney *P* value
First contact clinicians	3	7 (14)	11 (22)	32 (64)	
Burn specialists	3	4 (12.5)	15 (46.9)	13 (40.6)	
All participants	6	11 (13.4)	26 (31.7)	45 (54.9)	

Discussion

Main Findings

This study is an attempt to bridge the gap between what is known about the potential performance of automated diagnosis for burn care and the readiness of first contact clinicians to endorse this as a practice. We showed that a majority of first contact clinicians from resource-poor African settings would be willing to use automated diagnosis for burns should the technology be made available in their departments. The concepts embedded in the AAM contributed moderately to explaining what the intention to use such a technology rests on. Of special interest is the association between the intention to use and perceived usefulness, but not with attitudes toward the technology. It is also of note that 5 of the 8 hypotheses derived from the model were supported.

When reflecting upon the risks that such technologies could imply, first contact clinicians showed more concerns than burns specialists. They were typically concerned with the consequences of poor diagnostic accuracy (ie, undermanagement of severe burns and overmanagement of minor burns) and, like their specialized peers, feared that clinicians' sense of responsibility would be reduced and that complex burns might not be identified. The latter was regarded as the risk with the most severe consequences across all clinicians, alongside that of severe burns not being identified. By contrast, first contact clinicians were less concerned with the risk of minor burns being diagnosed as minor or insignificant.

That first contact clinicians show a widespread acceptance of the intention to use automated burn diagnosis is well in line with previous research on the use of clinical AI in general, where 77% of the respondents from different parts of the world indicated being willing to use clinical AI if needed [33]. Perceived usefulness has already been suggested as a strong predictor of both behavioral intention and technology acceptance. This is true for technologies such as smart health care [34,35], whereas in the context of burns, it has been shown for the acceptance of mHealth [21].

As the model used does not pay attention to risk perception per se in the determination of intention to use, our results can shed light on it being a potentially underlying source of influence on the intention to use. Indeed, first contact clinicians' appraisal of the risk is more multifaceted than that of their specialized peers. A first explanation could be self-confidence, doubting their ability to interpret the output of an automated burn diagnosis [33]. Another one could be minimal experience of mHealth for burn diagnosis, while experience of the like is known to be associated with a decrease in perceived risks in the use of smart health care, thanks to technology transfer knowledge [35]. A third one could be a sense of loss in personal usefulness, first contact clinicians being more prone to answer positively to the fact that AI would replace them.

Strengths and Limitations

This is the first study performed in resource-poor settings specifically on the topic of acceptance of automated diagnostic acceptance among health care professionals. This can complement the work done on a similar population on the acceptance of mHealth for burn injuries [21], but also results in the acceptance of AI-based diagnosis among medical students, also performed in resource-poor settings [36]. Studies on intention to use are important to be performed for specific contexts and populations of interests as it has been shown previously that physicians from resource-poor settings may have different opinions regarding clinical AI than their counterparts from other regions of the world [33]. The perceptions might also depend on the qualification of the surveyed professional [21,33,35]. The development of new methodologies to analyze the obtained data (such as machine learning techniques) might also assist in the future in targeting more specifically what might influence the intention to use new technologies by survey participants. The study also focused on 7 different risks that are relevant for the diagnosis of burn care and for which the knowledge around clinicians' perceptions is important both for the development of algorithms as well as for implementation. It is likely that burns specialists, sharing a familiarity with the condition and aware of the progress made in assisted diagnosis are more inclined to trust the technology [37] and that, by contrasts to first contact clinicians, are more heterogeneous as a group as regards their knowledge and experience about burns, their skills and competences [21]. The sample was conveniently selected among clinicians attending 2 specialized conferences in South Africa for their interest in burn care and likelihood of using automated diagnosis would it be developed. This, however, comes with the limitation that the pool of people who can be included is relatively small, and the results cannot be generalized to other specialties or settings where the viewpoints of the professionals might differ. Furthermore, there might have been respondent bias with those accepting to fulfill the questionnaire having stronger opinions in favor of or against automated diagnosis for burns than those who did not respond. This study was cross-sectional and only assessed the intention to use a potential tool rather than

XSL•FO
RenderX

effectively measuring the clinicians' behaviors once automated diagnosis was implemented. While this information will be interesting to investigate in the future, the knowledge gathered in this study can further assist in the development of the algorithms themselves to respond most appropriately to the stakeholders' demands. It could also be of interest to assess users' perceptions with regard to different methodologies with varying performances, in particular when current advances in the technological domain are extremely promising.

Future Research

This study has implications for future research as the AAM model contributes poorly to informing what the determinants of the intention to use are beyond the perceived usefulness. There is a need to assess the input of clinical risk perceptions: this should be an integral part of the model in the future, and it should receive attention during the evaluation and implementation process. Furthermore, with the body of evidence growing on the potential for automated diagnosis in burn care and its potential benefit in resource-poor settings, research on the attitudes of medical professionals with regard to this technology is necessary for a successful implementation. Finally, it would be interesting to investigate, when implementing such a technology in practice, whether it would enhance the clinical abilities of first-contact clinicians, as well as improve diagnostic efficiency, given the current state of resources and the lack of specialist clinicians.

Conclusions

The results show that most first contact clinicians were inclined to seek automated advice for burn diagnosis if it was available. The conceptual model proposed contributes moderately to explaining what influences intention to use beyond perceived usefulness, with implications for clinical practice where the perceived usefulness makes little doubt. This is encouraging as there are good reasons to believe that the performance of automated diagnostic support will continuously improve. When seeking additional determinants, greater attention should be placed on clinical risk perceptions as well as negative attitudes due to bad experiences or unrealistic expectations. These dimensions should also naturally belong to any AI implementation process in practice, not least to ensure sustainability.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Table S1. CHERRIES checklist. CHERRIES: Checklist for Reporting Results of Internet E-Surveys.
[PDF File, 114 KB - humanfactors_v12i1e56300_app1.pdf]

References

1. Farhad M, Masud MM, Beg A, Ahmad A, Ahmed LA, Memon S. A data-efficient zero-shot and few-shot Siamese approach for automated diagnosis of left ventricular hypertrophy. Comput Biol Med 2023 Sep;163:107129. [doi: 10.1016/j.compbiomed.2023.107129] [Medline: 37343469]
2. Xu K, Huang S, Yang Z, et al. Automatic detection and differential diagnosis of age-related macular degeneration from color fundus photographs using deep learning with hierarchical vision transformer. Comput Biol Med 2023 Dec;167:107616. [doi: 10.1016/j.compbiomed.2023.107616] [Medline: 37922601]
3. Yeasmin MN, Al Amin M, Joti TJ, Aung Z, Azim MA. Advances of AI in image-based computer-aided diagnosis: A review. Array 2024 Sep;23:100357. [doi: 10.1016/j.array.2024.100357]
4. Bejnordi BE, Veta M, van Diest PJ, et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. JAMA 2017 Dec 12;318(22):2199-2210. [doi: 10.1001/jama.2017.14585] [Medline: 29234806]
5. Ting DSW, Cheung CYL, Lim G, et al. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. JAMA 2017 Dec 12;318(22):2211-2223. [doi: 10.1001/jama.2017.18152] [Medline: 29234807]
6. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature New Biol 2017 Feb 2;542(7639):115-118. [doi: 10.1038/nature21056] [Medline: 28117445]
7. Boissin C, Laflamme L, Fransén J, et al. Development and evaluation of deep learning algorithms for assessment of acute burns and the need for surgery. Sci Rep 2023 Jan 31;13(1):1794. [doi: 10.1038/s41598-023-28164-4] [Medline: 36720894]
8. Strohm L, Hehakaya C, Ranschaert ER, Boon WPC, Moors EHM. Implementation of artificial intelligence (AI) applications in radiology: hindering and facilitating factors. Eur Radiol 2020 Oct;30(10):5525-5532. [doi: 10.1007/s00330-020-06946-y] [Medline: 32458173]
9. Ahmad Z, Rahim S, Zubair M, Abdul-Ghafar J. Artificial intelligence (AI) in medicine, current applications and future role with special emphasis on its potential and promise in pathology: present and future impact, obstacles including costs and acceptance among pathologists, practical and philosophical considerations. A comprehensive review. Diagn Pathol 2021 Dec;16(1):24. [doi: 10.1186/s13000-021-01085-4]

XSL•FO
RenderX

10. Eschert T, Schwendicke F, Krois J, Bohner L, Vinayahalingam S, Hanisch M. A survey on the use of artificial intelligence by clinicians in dentistry and oral and maxillofacial surgery. Medicina (Kaunas) 2022 Aug 5;58(8):1059. [doi: 10.3390/medicina58081059] [Medline: 36013526]
11. Scheetz J, Rothschild P, McGuinness M, et al. A survey of clinicians on the use of artificial intelligence in ophthalmology, dermatology, radiology and radiation oncology. Sci Rep 2021 Mar 4;11(1):5193. [doi: 10.1038/s41598-021-84698-5] [Medline: 33664367]
12. Tonnang HEZ, Tchouassi DP, Juarez HS, Igweta LK, Djouaka RF. Zoom in at African country level: potential climate induced changes in areas of suitability for survival of malaria vectors. Int J Health Geogr 2014 May 7;13(1):12. [doi: 10.1186/1476-072X-13-12] [Medline: 24885061]
13. Holmström O, Stenman S, Suutala A, et al. A novel deep learning-based point-of-care diagnostic method for detecting Plasmodium falciparum with fluorescence digital microscopy. PLOS ONE 2020;15(11):e0242355. [doi: 10.1371/journal.pone.0242355] [Medline: 33201905]
14. Alaran MA, Lawal SK, Jiya MH, et al. Challenges and opportunities of artificial intelligence in African health space. Digit HEALTH 2025;11:20552076241305915. [doi: 10.1177/20552076241305915] [Medline: 39839959]
15. Rajakumar MP, Sonia R, Maheswari BU, Karuppiah SP. Tuberculosis detection in chest X-ray using Mayfly-algorithm optimized dual-deep-learning features. J Xray Sci Technol 2021;29(6):961-974. [doi: 10.3233/XST-210976] [Medline: 34459432]
16. Boissin C, Laflamme L. Accuracy of image-based automated diagnosis in the identification and classification of acute burn injuries. a systematic review. EBJ 2021;2(4):281-292. [doi: 10.3390/ebj2040020]
17. Craglia M, Nepelski D, et al. AI Watch: AI Uptake in Health and Healthcare: Publications Office of the European Union; 2020. [doi: 10.2760/948860]
18. Wallis L, Hasselberg M, Barkman C, et al. A roadmap for the implementation of mHealth innovations for image-based diagnostic support in clinical and public-health settings: a focus on front-line health workers and health-system organizations. Glob Health Action 2017 Jun;10(sup3):1340254. [doi: 10.1080/16549716.2017.1340254] [Medline: 28838310]
19. Ruxwana NL, Herselman ME, Conradie DP. ICT applications as e-health solutions in rural healthcare in the Eastern Cape Province of South Africa. HIM J 2010 Mar;39(1):17-29. [doi: 10.1177/183335831003900104]
20. Kifle M, Payton FC, Mbarika V, Meso P. Transfer and adoption of advanced information technology solutions in resource-poor environments: the case of telemedicine systems adoption in Ethiopia. Telemed J E Health 2010 Apr;16(3):327-343. [doi: 10.1089/tmj.2009.0008] [Medline: 20406120]
21. Klingberg A, Sawe HR, Hammar U, Wallis LA, Hasselberg M. m-Health for burn injury consultations in a low-resource setting: an acceptability study among health care providers. Telemed J E Health 2020 Apr;26(4):395-405. [doi: 10.1089/tmj.2019.0048] [Medline: 31161967]
22. Blom L. mHealth for image-based diagnostics of acute burns in resource-poor settings: studies on the role of experts and the accuracy of their assessments. Glob Health Action 2020 Dec 31;13(1):1802951. [doi: 10.1080/16549716.2020.1802951] [Medline: 32814518]
23. Boissin C. Clinical decision-support for acute burn referral and triage at specialized centres - contribution from routine and digital health tools. Glob Health Action 2022 Dec 31;15(1):2067389. [doi: 10.1080/16549716.2022.2067389] [Medline: 35762795]
24. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34. [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
25. Ghazizadeh M, Lee JD, Boyle LN. Extending the technology acceptance model to assess automation. Cogn Tech Work 2012 Mar;14(1):39-49. [doi: 10.1007/s10111-011-0194-3]
26. Compeau DR, Higgins CA. Computer self-efficacy: development of a measure and initial test. MIS Q 1995 Jun;19(2):189. [doi: 10.2307/249688]
27. Venkatesh V. Determinants of perceived ease of use: integrating control, intrinsic motivation, and emotion into the technology acceptance model. Inf Syst Res 2000 Dec;11(4):342-365. [doi: 10.1287/isre.11.4.342.11872]
28. Collins SE, Witkiewitz K, Larimer ME. The theory of planned behavior as a predictor of growth in risky college drinking. J Stud Alcohol Drugs 2011 Mar;72(2):322-332. [doi: 10.15288/jsad.2011.72.322] [Medline: 21388605]
29. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. Manage Sci 1989 Aug;35(8):982-1003. [doi: 10.1287/mnsc.35.8.982]
30. Rotter JB. Interpersonal trust, trustworthiness, and gullibility. Am Psychol 1980;35(1):1-7. [doi: 10.1037//0003-066X.35.1.1]
31. Bagozzi R, University of Michigan. The legacy of the technology acceptance model and a proposal for a paradigm shift. JAIS 2007;8(4):244-254. [doi: 10.17705/1jais.00122]
32. Hair J, Hult GT, Ringle C, Sarstedt M. A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM), 3rd edition: Sage; 2022. [doi: 10.1007/978-3-030-80519-7]
33. Chen M, Zhang B, Cai Z, et al. Acceptance of clinical artificial intelligence among physicians and medical students: a systematic review with cross-sectional survey. Front Med (Lausanne) 2022;9:990604. [doi: 10.3389/fmed.2022.990604] [Medline: 36117979]

XSL•FO
RenderX

34. Zhai H, Yang X, Xue J, et al. Radiation oncologists' perceptions of adopting an artificial intelligence-assisted contouring technology: model development and questionnaire study. J Med Internet Res 2021 Sep 30;23(9):e27122. [doi: 10.2196/27122] [Medline: 34591029]
35. Pan J, Ding S, Wu D, Yang S, Yang J. Exploring behavioural intentions toward smart healthcare services among medical practitioners: a technology transfer perspective. Int J Prod Res 2019 Sep 17;57(18):5801-5820. [doi: 10.1080/00207543.2018.1550272]
36. Tran AQ, Nguyen LH, Nguyen HSA, et al. Determinants of intention to use artificial intelligence-based diagnosis support system among prospective physicians. Front Public Health 2021;9:755644. [doi: 10.3389/fpubh.2021.755644] [Medline: 34900904]
37. Blom L, Laflamme L, Alvesson HM. Expectations of medical specialists about image-based teleconsultation - a qualitative study on acute burns in South Africa. PLOS ONE 2018;13(3):e0194278. [doi: 10.1371/journal.pone.0194278] [Medline: 29543847]

Abbreviations

AAM: Automation Acceptance Model
AI: artificial intelligence
EMSSA: Emergency Medicine Society of South Africa

Edited by A Kushniruk; submitted 12.01.24; peer-reviewed by H Chen, J Sung, J Walsh; revised version received 19.03.25; accepted 23.03.25; published 03.06.25.

Please cite as:
Boissin C, Blom L, Taha Z, Wallis L, Allorto N, Laflamme L
Intention to Use Automated Diagnosis and Clinical Risk Perceptions Among First Contact Clinicians in Resource-Poor Settings: Questionnaire-Based Study Focusing on Acute Burns
JMIR Hum Factors 2025;12:e56300
URL: https://humanfactors.jmir.org/2025/1/e56300
doi:10.2196/56300

© Constance Boissin, Lisa Blom, Zara Taha, Lee Wallis, Nikki Allorto, Lucie Laflamme. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 3.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Mobile-Based Cognitive Behavioral Therapy for Health Care Workers' Mental Health in Ecuador: Quasi-Experimental Study

Sandra Lorena Muñoz-Ortega[1], PhD; Rubén Vladimir Alvarado Muñoz[2], MD; Daniela Santamaria Guayaquil[1]; Jade Pluas-Borja[1]; Marco Faytong-Haro[1], MA

[1]Facultad de Ciencias de la Salud, Universidad Espiritu Santo, Km 2.5 vía a Samborondón, Samborondón, Ecuador

[2]Centro Interdisciplinario de estudios en Salud (CIESAL), Departamento de Salud Pública, Escuela de Medicina, Universidad de Valparaiso, Valparaiso, Chile

Corresponding Author:
Sandra Lorena Muñoz-Ortega, PhD
Facultad de Ciencias de la Salud, Universidad Espiritu Santo, Km 2.5 vía a Samborondón, Samborondón, Ecuador

Abstract

Background: Mental health challenges, including depression, anxiety, and burnout, have become increasingly prevalent among health care workers, who face high-stress environments, limited resources, and long working hours. The COVID-19 pandemic has intensified these issues, especially in regions like Latin America, where health care professionals experience heightened anxiety and depression. The urgent need for mental health support has prompted the development of mobile health (mHealth) solutions. These tools offer accessible, confidential interventions that help reduce stigma and encourage engagement. The "Psicovida" mobile app was designed to provide cognitive behavioral therapy (CBT)–based activities tailored to health care workers, supporting them in managing stress, anxiety, and depression.

Objective: This study aims to evaluate the effectiveness of Psicovida, a mobile app that delivers CBT-based interventions, in reducing depressive symptoms and emotional distress among health care workers over a 3-month period.

Methods: A quasi-experimental, nonrandomized controlled study was conducted with health care workers at a public hospital in Ecuador. Participants were recruited offline and assigned to either an intervention group that used the Psicovida app or a control group that received no intervention. The app provided weekly CBT-based tasks focused on stress management, cognitive restructuring, and emotional regulation. Data collection included demographic information, with mental health outcomes assessed pre- and postintervention using the Patient Health Questionnaire-9 (PHQ-9) to measure depression and the General Health Questionnaire-12 to assess overall psychological well-being.

Results: A total of 211 health care workers participated, with 88 in the intervention group and 96 in the control group, and 29 participants dropped out. Among the intervention group, adherence varied: 34% (30/88) used the app consistently for 10 - 12 weeks, 42% (37/88) for 7 - 9 weeks, and 24% (21/88) for fewer than 6 weeks. Significant improvements in mental health outcomes were observed among app users. The intervention group exhibited a statistically significant reduction in depressive symptoms, with PHQ-9 scores decreasing significantly (P<.001; 95% CI 6.17-9.36). Within this group, 20% (18/88) achieved complete remission of depressive symptoms (PHQ-9 scores <5), 32% (28/88) showed mild symptoms (PHQ-9 scores=5 - 9), and 48% (42/88) remained in the range requiring treatment referral (PHQ-9 scores ≥10). General Health Questionnaire-12 scores similarly showed substantial improvement in psychological well-being (P<.001; 95% CI 3.99-5.58).

Conclusions: The Psicovida mobile app demonstrates promise as an accessible, effective tool for reducing depression and anxiety among health care workers through CBT-based interventions. This study highlights the potential of mHealth technology to deliver targeted mental health support, especially in resource-limited settings. Future research should focus on evaluating long-term impacts and broader applications in varied health care environments.

Trial Registration: ClinicalTrials.gov NCT06650449; https://clinicaltrials.gov/study/NCT06650449

(JMIR Hum Factors 2025;12:e58943) doi:10.2196/58943

KEYWORDS

mental health; depression; mobile apps; mental disorders; effectiveness; depressive symptoms; emotional distress; cognitive behavioral therapy; experimental study; design; Ecuador; psychological well-being; psychological; tool; health care worker; accessible; support system

XSL•FO
RenderX

Introduction

Background

Mental health has emerged as a global challenge affecting millions of individuals across geographical and cultural boundaries. Mental disorders, including depression and anxiety, now affect over 450 million people globally, according to the World Health Organization, posing serious consequences for quality of life and productivity. Events like the COVID-19 pandemic have only exacerbated these challenges, highlighting the urgent need for innovative and accessible solutions [1].

Health care workers, especially doctors and nurses, are among those most affected by the mental health crisis, facing high rates of depression, anxiety, and burnout. Research shows that nearly half of junior doctors report symptoms of depression, while over 63% experience significant anxiety. Physicians also face burnout rates 20% - 30% higher than those of other professions and are 1.4 to 2.3 times more likely to die by suicide [2]. Despite these concerning statistics, many health care workers hesitate to seek mental health support due to stigma and fear of professional consequences [3]. This hesitation highlights the critical need for accessible, confidential mental health resources that encourage help-seeking within health care environments, fostering a culture where mental health support is accepted and valued.

In Latin America, health care professionals contend with additional mental health challenges, worsened by systemic issues like insufficient health care infrastructure, extended work hours, and direct exposure to the COVID-19 crisis. Anxiety affects between 10.6% and 76.5% of health care professionals, with depression impacting up to 81% in some regions. In Ecuador, for example, 46.4% of health care workers report experiencing anxiety, while 34.5% report depression. Among female obstetricians in Peru, over half exhibit depressive symptoms [4]. Frontline workers, particularly young and female health care professionals, face heightened vulnerability due to high workloads, inadequate personal protective equipment, and ethical dilemmas, leading to both physical and mental exhaustion [5]. The prevalence of these mental health issues points to an urgent need for and met the inclusion criteria effective psychological support and systemic improvements in working conditions.

Addressing these challenges requires immediate attention to systemic reforms that include accessible and confidential mental health services. In this context, mHealth solutions have emerged as promising tools. Recent studies have assessed the effectiveness of mHealth interventions targeting various mental health issues, including depression [6,7], suicide [8], schizophrenia [9], substance use disorders [10], and psychosis [11], among others [12]. Recent systematic reviews focused on assessing smartphone apps for mental health demonstrated their ability to generate significant reductions in anxiety [13] and depression [13]. mHealth apps have been proven particularly effective in breaking down barriers to help-seeking, such as stigma and time constraints, and are shown to improve user engagement when co-designed with health care workers [14]. Additionally, mHealth tracking apps have been recognized for increasing self-awareness and enabling users to manage their symptoms by identifying triggers and patterns. This kind of proactive self-management can be highly beneficial for health care workers facing depression and burnout [15]. These digital tools, if implemented widely, could play a pivotal role in addressing the ongoing mental health crisis among health care workers.

Objective

This study aims to bridge the gap in understanding the effectiveness of mHealth interventions in mitigating mental health issues among health care workers. By conducting a quasi-experimental, nonrandomized controlled trial among health care workers in a public hospital in Ecuador, this research seeks to evaluate the impact of the “Psicovida” mobile app [16], a digital tool designed to support mental health through cognitive behavioral therapy (CBT)–based activities, on depression and emotional distress symptoms. The app offers structured interventions aimed at alleviating symptoms associated with mental health disorders, providing users with accessible and guided support. The unique challenges faced by health care personnel in Latin America, especially in regions like Ecuador, where systematic monitoring of mental health issues is lacking, underscore the importance of this research.

Methods

Study Design and Setting

This study used a quasi-experimental, nonrandom, parallel, controlled trial with pre- and postintervention assessments. It involved 2 groups with similar baseline characteristics: an intervention group receiving access to the Psicovida app and a control group without access. The study started in January 2023, coinciding with a surge in COVID-19 cases in Ecuador. This wave significantly impacted the health care system, creating heightened stress among health care workers [17]. The timing of the study underscores the critical need for effective mental health support mechanisms for this population.

The trial adheres to the Consolidated Standards of Reporting Trials (CONSORT) eHealth guidelines, and the detailed study protocol and statistical analysis plan are publicly available on ClinicalTrials.gov (NCT06650449) [18,19].

Rationale for Quasi-Experimental Design

The decision to use a quasi-experimental approach was driven by the desire to achieve homogeneity in key demographic characteristics (age, gender, marital status, education level, and profession) between the intervention and control groups. To achieve this, participants were purposefully assigned to 2 cohorts: those who used the Psicovida app and those who did not. This intentional grouping aimed to control for potential confounding variables, acknowledging the limitation that true random assignment was not used.

Participants

The study population comprised health care workers (doctors and nurses) from various departments of a public hospital in Quito, Ecuador. Participants were purposefully assigned to 2 groups—those who used the Psicovida app and those who did

XSL•FO
RenderX

not—to control for potential confounding variables, acknowledging the limitation that true random assignment was not used. At the request of hospital administrators, some staff from high-acuity wards, whose heavy workloads left no additional time for training, were kept in the control group, whereas personnel from outpatient clinics with more flexible schedules were placed in the intervention group. Inclusion criteria were (1) employment as health care personnel at the designated hospital for at least 12 months, (2) ownership of a smartphone, and (3) provision of informed consent. Participants included both men and women aged 18 years and older, regardless of marital status. These criteria ensured engagement with the Psicovida app and effective completion of study tasks.

Individuals not meeting any of the above inclusion criteria were excluded from participation.

Recruitment and Data Collection

Recruitment and data collection occurred primarily through the Psicovida app between October 2022 and January 2023, using 3 strategies. First, informational meetings were conducted by the research team to explain the study and the Psicovida app and to obtain digital consent. Second, from an initial pool of 296 health care professionals who expressed interest and met the inclusion criteria, 211 eligible participants were selected based on EPIDAT software (EPIDETAL, developed by the Epidemiology Service of the Dirección Xeral de Saúde Pública, Consellería de Sanidade (Xunta de Galicia), with support from the Pan American Health Organization (PAHO/WHO) and CES University of Colombia) recommendations and assigned to either the intervention or control group. A double-entry table was used during the assignment to ensure balanced distribution across key variables such as profession and sex. Finally, baseline data were collected, including personal information (email, age, gender, and education level) for participant profiling, and psychological evaluations using the Patient Health Questionnaire (PHQ-9) and the General Health Questionnaire (GHQ-12) to assess emotional distress and depressive symptoms.

During the study design phase, a plan was implemented to account for potential dropout rates, including strategies to manage missing data and maintain sample integrity. The study was conducted remotely via the Psicovida app, with minimal direct interaction with the research team, except when participants opted to initiate communication.

Intervention

The Psicovida app is a free tool designed to support health care workers' mental health by providing CBT-based activities aimed at alleviating symptoms of depression and emotional distress. The app includes a structured program of 12 weekly activities, each taking around 2 to 10 minutes to complete, focusing on introspection, stress management, and the restructuring of negative thoughts. The modules cover various aspects of mental health and personal growth, delivered through interactive tasks that address emotional, cognitive, and physical dimensions.

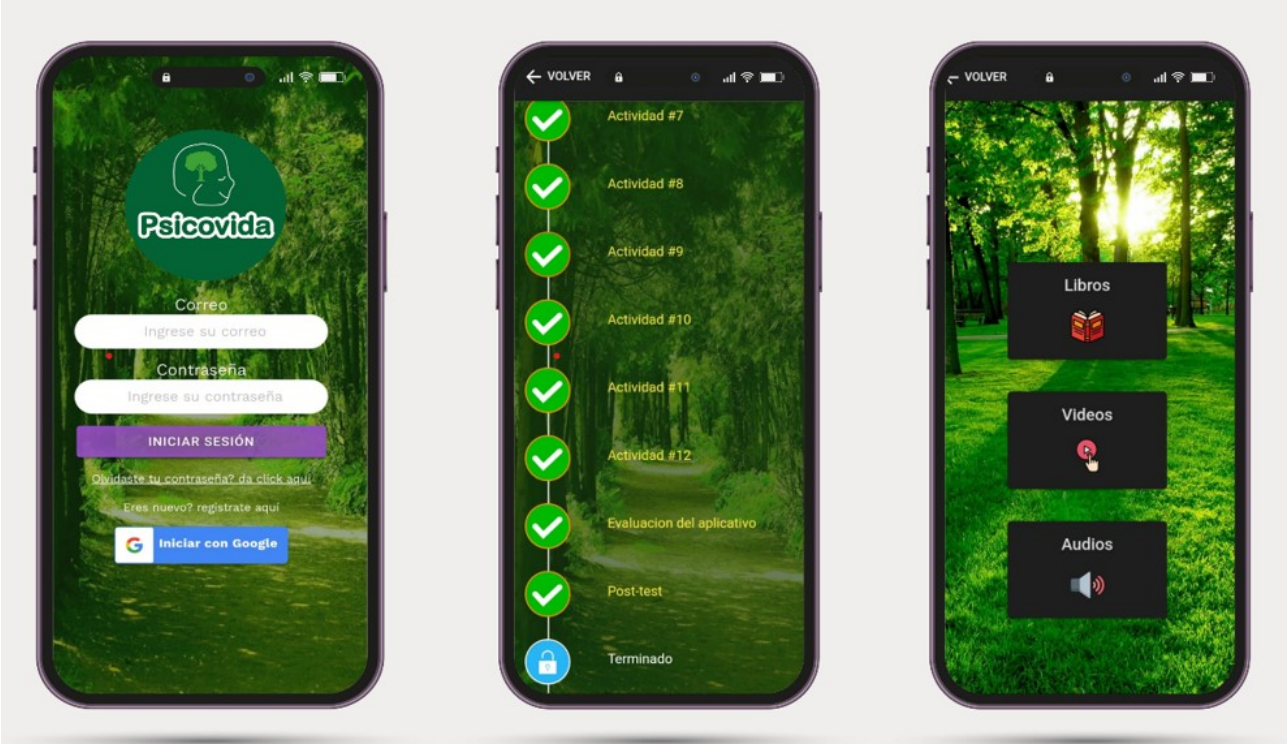
The app was available to all interested users for download from the Google Play Store, without the requirement to belong to a specific group to access it. Participants did not pay for access, nor did they receive any financial compensation for their participation. During the onboarding phase, participants were encouraged to engage with the app regularly, completing the assigned activities over a period of up to 3 months. The app sent daily reminder emails to participants from Monday to Friday, encouraging consistent participation and improving user engagement. Users could adjust the settings for reminders, allowing flexibility in how they interact with the app. Users could adjust the settings for reminders, allowing flexibility in how they interact with the app. You can see the app interface in Figure 1.

Since its initial release, significant updates have been made to improve user experience, including expanding system bandwidth to address downtimes and enhancing the app's user interface based on feedback from previous users. These changes, validated by clinical psychology experts, ensure that the therapeutic quality of the activities remains intact, offering personalized support to health care workers. Developed as Discord Psicovida Version 1 by principal author (SMO), the program underwent rigorous evaluations to align with the needs of its user base. For digital preservation and access, the Psicovida app can be accessed via a dedicated link on the Open Science Framework repository [20]. Upon completing the program, participants received a notification thanking them for their participation and were provided with general recommendations for maintaining long-term mental health. The description of each activity is presented in Multimedia Appendix 1.

At the end of the app use period, participants provided feedback through a series of structured questions. This feedback will help guide improvements for future versions of the app, based on user experiences and suggestions.

XSL•FO
RenderX

Figure 1. Psicovida mobile app.



Assessments

Depression

The PHQ-9 was carefully implemented as a digital assessment tool to ensure data reliability and participant engagement. Participants provided informed consent online after being fully informed about the purpose of the PHQ-9, confidentiality measures, and their data privacy rights. The questionnaire maintained its standard, validated format with nine items, each required to be answered to ensure complete data collection. The questions were presented in a sequential, user-friendly interface that preserved the questionnaire's established structure without adaptive questioning or alterations.

Participant completion was monitored at both baseline and postintervention to assess adherence and track changes in depression severity over time. Unique identifiers were used to ensure that each participant completed the PHQ-9 only once at each time point, eliminating the risk of duplicate entries. Data security was prioritized through encryption and secure storage of responses. Although specific online validation was not conducted, the PHQ-9's high internal consistency (Cronbach α>0.80) and proven sensitivity and specificity in various settings support its reliability for this digital deployment [21].

Anxiety

The GHQ-12 was implemented digitally to assess emotional distress while upholding rigorous standards for data quality and user experience. Before beginning the GHQ-12, participants reviewed and agreed to informed consent details, which clarified the study's objectives, the role of the GHQ-12, and data protection practices. All 12 items were mandatory, ensuring comprehensive data for each participant and maintaining the GHQ-12's validated structure. The questionnaire was presented in a fixed, linear format, allowing participants to complete it with ease and consistency, without adaptive questioning or alterations.

Completion rates were monitored both at the beginning and end of the study to evaluate participant engagement and the tool's effectiveness in capturing changes in emotional distress. Each participant was assigned a unique identifier to track usage and prevent multiple entries, supporting data integrity. Responses were securely encrypted and stored, safeguarding participant confidentiality. While no digital-specific validation of the GHQ-12 was mentioned, its established reliability (Cronbach α ranging from 0.70 to 0.90) and consistent performance across studies reinforce its suitability for digital application [22].

Demographic and Lifestyle Variables

Information on age, gender, marital status, education level, profession, and toxic habits (alcohol and tobacco consumption) was collected to describe the study population and adjust for potential confounders.

Procedure

The Psicovida quasi-experimental trial followed a structured process to ensure participant engagement and data collection. After recruitment and eligibility confirmation in January 2023, participants were randomly assigned to either the intervention group, which used the Psicovida app, or the control group, which did not receive app access.

Both groups completed baseline assessments, including demographic data collection and psychological evaluations using validated tools (PHQ-9 for depression and GHQ-12 for emotional distress). Following this, the intervention group started app use, while the control group received standard mental health support information. Outcome assessments were

XSL•FO
RenderX

conducted at two time points: (1) baseline (start of the study) and (2) study completion (end of 3 months). Participants received periodic email reminders to encourage adherence to the study protocol and completion of assessments. The intervention group was encouraged to engage with the app weekly for up to 12 weeks, with each activity designed to take 2 - 10 minutes. Upon completing the study, participants had the option to continue using the app, subject to updates or maintenance periods.

The study ran from January 2023 to March 2023, spanning three months.

Data Management and Monitoring

All study data were collected and recorded electronically through the Psicovida mobile app. The data, including identifiable information such as participants' email addresses and demographic details, was stored in a secure, encrypted database managed by the research team. Identifiers were removed during data processing to ensure confidentiality before analysis. Weekly backups of all data were conducted by the research team during the data collection phase, and backups were securely stored in a password-protected environment. Only approved members of the research team and designated technical staff had access to the data.

Trial Harms and Risk Management

Given the sensitive nature of mental health content, it was possible that some participants could experience distress while using the Psicovida app. Although no serious adverse events were anticipated based on previous trials of similar apps, the research team recorded and monitored any adverse events reported by participants. Adverse events were defined as any reported psychological distress or health impacts resulting from app use. The principal investigator documented and notified the ethics committee within 24 hours of any adverse events. If necessary, a data safety monitoring board consisting of independent clinical reviewers assessed the situation and recommended whether the study should continue. The app also provided direct access to mental health support services via a "reach out" button on the app's home screen, allowing users to contact help if needed.

Data Analysis

To determine the appropriate sample size for this study, EPIDAT, a specialized statistical software, was used with a one-sided hypothesis anticipating improved outcomes in the intervention group. Based on a 95% CI and a 5% margin of error, EPIDAT recommended reducing the initial pool of 296 health care personnel to a final sample size of 211 participants. Of these, 96 were allocated to the control group and 115 to the intervention group. However, there was an approximate loss of 29 participants, representing about 15% of the recommended sample size, due to personnel leaving the hospital or not completing the study.

As the study required active participant engagement with the Psicovida app, blinding participants to group allocation was not feasible. However, outcome assessors and data analysts were blinded to group assignments to reduce bias during data collection and analysis. Fortunately, there was no attrition during the study period, ensuring that all participants completed both baseline and postintervention assessments, thereby preserving the integrity of the data and the validity of the findings.

Descriptive statistics were used to summarize demographic characteristics, with chi-square tests for categorical variables and *t* tests (2-tailed) for continuous variables to compare distributions between groups. The impact of the Psicovida app on depression and emotional distress was analyzed using paired *t* tests for within-group comparisons and independent *t* tests to assess changes between groups. Effect size calculations were performed to evaluate the clinical significance of these findings.

To further understand the intervention's impact, binary logistic regression models were constructed to examine predictors of depression (PHQ-9 scores) and emotional distress (GHQ-12 scores) at the end of the study, considering factors such as app use, baseline psychopathology, and demographic variables. All statistical analyses were conducted using SPSS software (IBM Corp), with significance set at $P<.05$, based on data sourced from the Psicovida app database as of March 2023.

Ethical Considerations

This study received approval from the Human Research Ethics Committee of the Carlos Andrade Marín Hospital (registration number IESS-HCAM-CEISH-2022 - 0007) on April 21, 2022. All procedures adhered to the ethical standards of the institutional and national research committee, as well as the principles of the Declaration of Helsinki. No exemptions from ethical review were sought, as the study involved human participants.

Participants provided informed consent digitally before enrolling in the study. The consent process included detailed information on the study's purpose, objectives, potential benefits, risks, and confidentiality measures. Participants confirmed their voluntary agreement by checking a digital consent form. They were informed of their right to withdraw from the study at any time without repercussions.

To ensure participant privacy and confidentiality, all data collected through the Psicovida mobile app were anonymized prior to analysis. Identifiable information, such as email addresses, was encrypted and stored in a secure database accessible only to the research team. Unique identifiers were used to link participant data while preserving anonymity during analysis. No personally identifiable information was included in the final dataset.

Participants were not provided with financial or material compensation for their participation in this study. Instead, they were thanked for their contribution to advancing mental health research and given general recommendations for maintaining long-term mental health upon completing the study.

Results

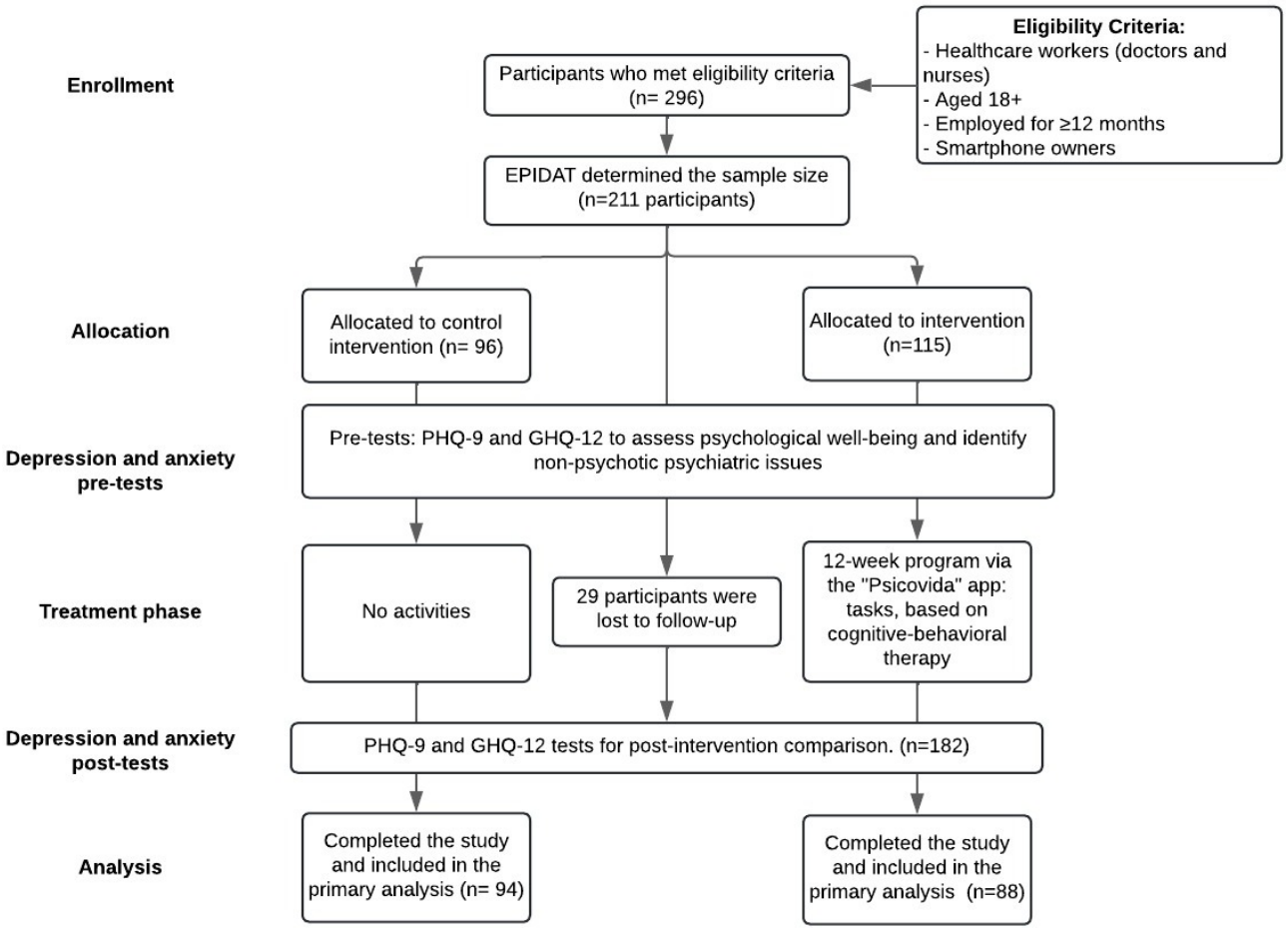
Overview

Between January 2023 and March 2023, 296 health care workers expressed interest in participating in the Psicovida study. Based on EPIDAT's recommendations, the sample size was reduced

XSL•FO
RenderX

to 211 eligible participants who provided informed consent and were assigned to either the Psicovida intervention group (n=115) or the control group (n=96). However, due to a 15% dropout rate, approximately 29 participants discontinued their involvement. As a result, the final distribution included 94 participants in the control group and 88 in the intervention group, who were evaluated for the primary outcome. The procedure for the Psicovida quasi-experimental trial is presented in Figure 2.

Figure 2. Psicovida study participation flowchart based on the CONSORT (Consolidated Standards of Reporting Trials) 2011 statement recommendations. GHQ-12: General Health Questionnaire-12; PHQ-9: Patient Health Questionnaire-9.



In terms of app use, 34% (30 participants) in the intervention group used the app for the full duration of 10 - 12 weeks, 42% (37 participants) used it for 7 - 9 weeks, and 24% (21 participants) used it for fewer than 6 weeks. This variation in usage was taken into account in the analysis to understand the impact of exposure levels on primary outcomes.

Table 1 reveals a distinct age group distribution within the app use group, primarily due to the higher percentage of individuals aged between 18 and 25 years. Both groups exhibited a uniform distribution of sex and marital status, with females and single individuals prevailing. Similarly, the predominant education level was undergraduate studies, consistent with the most common age group. As for profession, despite having inversely proportional frequencies in the study groups, there was no statistically significant difference. Overall, less than 11% (n=19) of the studied participants engaged in alcohol or tobacco consumption.

XSL•FO
RenderX

Table . Distribution of demographic variables in the study groups.[a]

Demographic variables	Intervention group	Group without intervention	*P* value
Age (years), n (%)			<.001
18 - 25	54 (61.4)	46 (48.9)	
26 - 35	10 (11.4)	18 (19.2)	
36 - 55	20 (22.7)	20 (21.3)	
Over 55	4 (4.5)	10 (10.6)	
Sex, n (%)			.13
Female	60 (68.2)	54 (57.4)	
Male	28 (31.8)	40 (42.6)	
Marital status, n (%)			.23
Single	43 (48.9)	46 (48.9)	
Divorced	8 (9.1)	12 (12.8)	
Free union	19 (21.6)	13 (13.8)	
Married	18 (20.4)	23 (24.5)	
Educational level, n (%)			.38
Undergraduate	77 (87.5)	86 (91.5)	
Postgraduate	11 (12.5)	8 (8.5)	
Profession, n (%)			.11
Doctors	48 (54.5)	40 (42.6)	
Nurses	40 (45.5)	54 (57.4)	
Toxic habits, n (%)			
Alcohol consumption	15 (17.0)	4 (4.3)	.005
Tobacco consumption	5 (5.7)	19 (20.2)	.004

[a]Source: Database of the Psicovida mobile app, March 2023.

Evaluation of Depression Symptoms at the Beginning and End of the Groups With and Without Use of the App (With the PHQ-9 Test)

Table 2 shows the scores in both groups at the start and end of the research. We categorized the PHQ-9 variable to discern the proportion of subjects with associated psychopathology according to the applied test evaluation.

The group using the app showed significant improvements (*P*<.001), with a 95% CI, with a difference of 6.17 - 9.36. This suggests that in a sample similar to the one studied, the application of the app is anticipated to result in a decrease in the PHQ-9 test score from 6 to 9 points with 95% CI.

At the end of the application, 20% (18/88) of the participants met the criteria for complete remission or no symptoms of depression (PHQ-9 scores <5), 32% (28/88) met the criteria for recovery or mild symptoms (PHQ-9 scores=5 - 9) and 48% (42/88) continued to meet the referral criteria for treatment (PHQ-9 ≥10). In the group that did not use the app, although a significant improvement was observed in some participants and there was a significant decrease in the mean in the final measurement (*P*=.03), the percentage that met the referral criteria for treatment (PHQ-9 scores ≥10) increased by 4%.

XSL•FO
RenderX

Table . Comparison of the results of the PHQ-9[a] test in the study groups at the beginning and end of the research.[b]

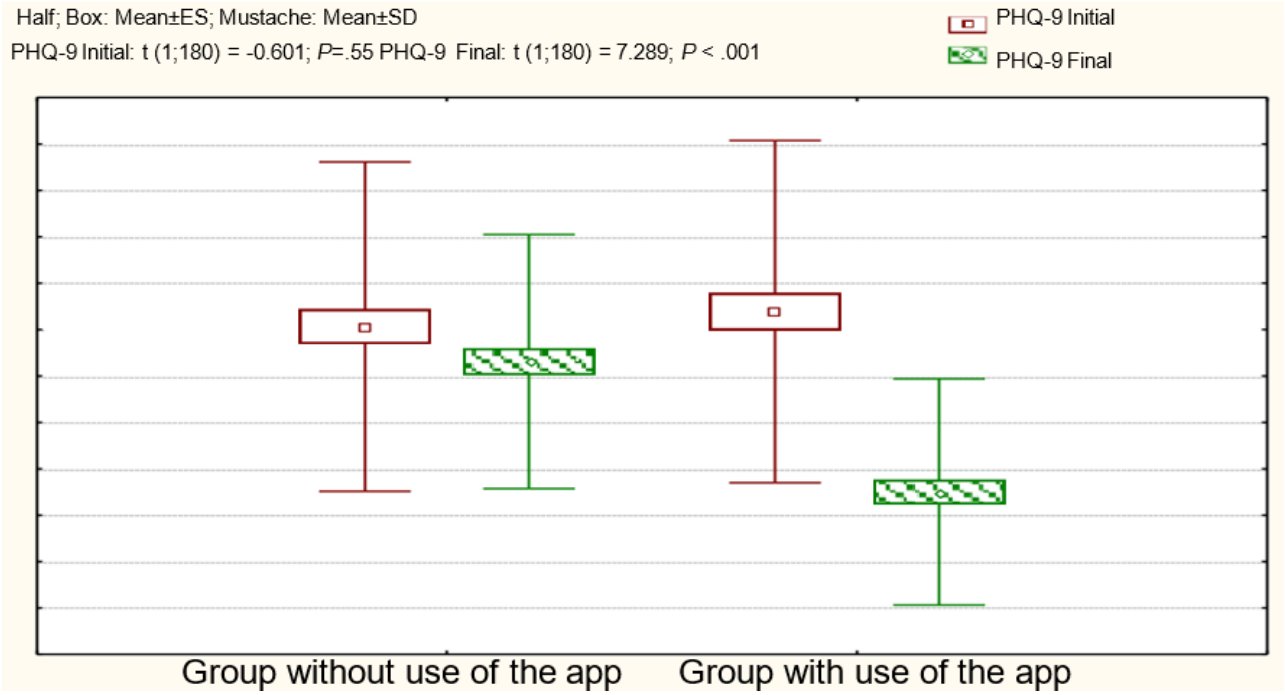
	Start	End	*P* value
PHQ-9, mean (SD)			
Cluster with use of the app	16.77 (7.37)	9.00 (4.88)	<.001
Group without use of the app	16.13 (−7.1)	14.62 (5.48)	.03
PHQ-9 ≥10, n (%)			
Cluster with use of the app	78 (89)	42 (48)	<.001
Group without use of the app	79 (84)	83 (88)	.34

[a]PHQ-9: Patient Health Questionnaire-9.

[b]Source: Database of the Psicovida mobile app, March 2023.

The following graph (Figure 3) shows the differences between the groups. The initial scores on the PHQ-9 test showed no statistically significant differences between the groups, but significant disparities emerged in the final scores. In the app user group, a decrease in the average was observed, and a large effect size was calculated (d=0.95).

Figure 3. Comparison of the results of the PHQ-9 test in the 2 study groups. ES: estimated standard error; PHQ-9: Patient Health Questionnaire-9.



Evaluation of the Symptoms of Emotional Distress at the Beginning and End of the Research Groups (With the GHQ-12 Test)

Similarly, for the GHQ-12 test (Table 3), we see significant improvements in the group using the app (*P*<.001), with a 95% CI for the difference of 3.99 - 5.58. This means that in a similar sample, using the app is expected to result in an average decrease in the GHQ-12 test score between 4 and 5 points, about 95% of the time.

XSL•FO
RenderX

Table . Comparison of the results of the GHQ-12[a] test in the study groups at the beginning and end of the research.[b]

	Start	End	*P* value
GHQ-12 , mean (SD)			
Cluster with use of the app	14.53 (7.22)	9.75 (5.24)	<.001
Group without use of the app	15.15 (5.95)	17.04 (5.57)	<.001
GHQ-12 ≥16, n (%)			
Group with use of the app	55 (57.29)	10 (11)	<.001
Group without use of the app	41 (44.71)	54 (57)	.008

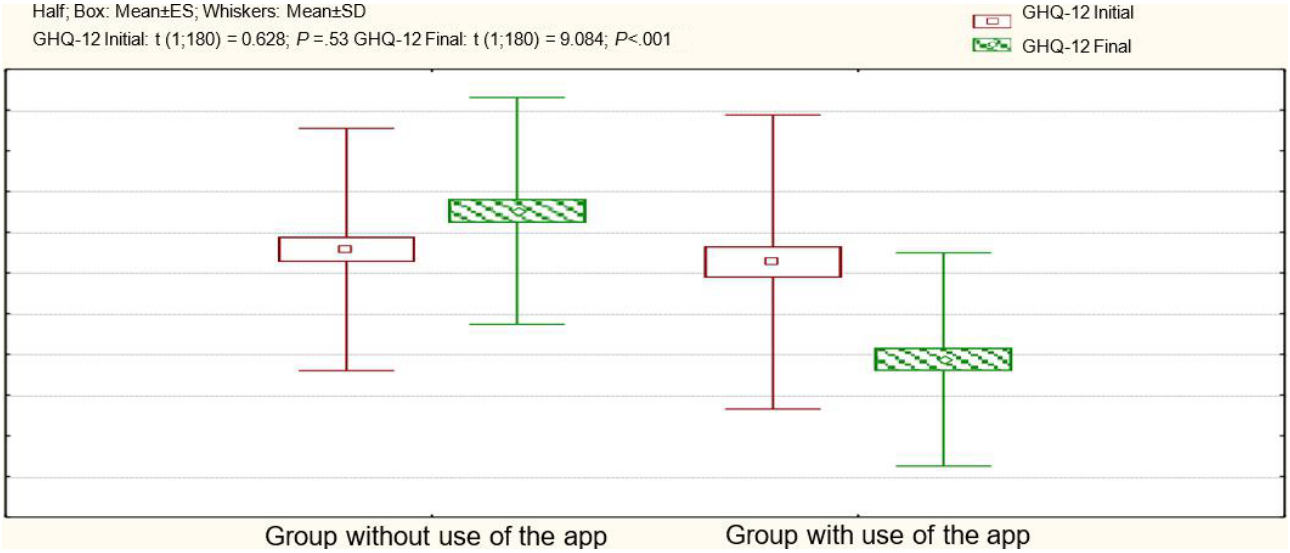
[a]GHQ-12: General Health Questionnaire-12.

[b]Source: Database of the Psicovida mobile app, March 2023.

In this group, 89% (78/88) of the participants met the criteria for complete remission or showed no distress symptoms (GHQ-12 scores <16), and only 11% (10/94) still met the referral criteria for treatment (GHQ-12 scores ≥16). Meanwhile, for the group without app use, the percentage of participants with psychopathology increased, and this difference became statistically significant compared to the initial proportion.

Figure 4 shows the GHQ-12 test results. Initially, both groups had similar scores. Towards the end, there are significant differences favoring the app users, with a noticeable effect (d=1.12). The group without app use also exhibits an increase in final scores.

Figure 4. Comparison of the results of the GHQ-12 test in the 2 study groups. ES: estimated standard error; GHQ-12: General Health Questionnaire-12.



Binary Logistic Regression of the PHQ-9 Test

The PHQ-9 results in our study were dichotomized to categorize participants into 2 distinct groups based on their depression severity levels. Individuals scoring 10 or above were classified as having moderate to severe depressive symptoms, indicating a need for further clinical evaluation. Those with scores below 10 were considered to have minimal to mild depressive symptoms, suggesting lower levels of depression. This dichotomization allowed for a clearer analysis of the relationship between demographic factors such as sex, age group, marital status, educational level, or profession, and depression severity.

Despite this categorization, our findings revealed no significant relationships between these demographic variables and the final PHQ-9 results.

The model (Table 4) had the presence of psychopathologies at the end of the research (PHQ-9f≥10) as the dependent variable. It includes 2 independent variables: the use of the Psicovida app (with the use of the app) and the variable (PHQ-9i≥10) representing the presence of psychopathologies at the beginning of the investigation, determined by the PHQ-9 test itself. This model shows a significance of less than .05 in the omnibus test, indicating that it contributes to explaining the observed event.

XSL•FO
RenderX

Table . Variables in the equation, PHQ-9[a]. Variables introduced in step 1: with use of the app and PHQ-9i≥10.

Variables	b	SE	Wald chi-square (*df*)	*P* value	Odds ratio (95% CI)
With use of the app[b]	−2.269	0.398	32.513 (1)	<.001	0.103 (0.047-0.226)
Step 1a PHQ-9i≥10[c]	2.196	0.498	19.458 (1)	<.001	8.987 (3.388-23.841)
Constant[d]	−0.166	0.419	0.157 (1)	.69	0.847 (N/A)[e]

[a]PHQ-9: Patient Health Questionnaire-9.

[b]With use of the app denotes a binary indicator that captures whether participants engaged with the mobile app during the intervention period (coded 1 for use, 0 for nonuse).

[c]Step 1a PHQ-9≥10 is a baseline severity flag identifying individuals whose initial PHQ-9 score met or exceeded the clinical threshold of 10, indicating moderate-to-severe depressive symptoms (coded 1 for ≥10, 0 otherwise).

[d]Constant represents the model intercept—the log-odds of screening positive when both predictors are set to zero (that is, for nonapp users with a baseline PHQ-9 <10).

[e]N/A: not available.

The independent variables of the model explain up to 37% of the dependent variable, and the correct classification level of the model is 72%. The first independent variable, the use of the Psicovida app, has a negative coefficient, an odds ratio (OR) <1, and therefore represents a protective factor, contrary to what happens with the initial presence of psychopathologies, which shows a positive coefficient, an OR>1; therefore, it is considered a risk factor.

Binary Logistic Regression of the GHQ-12 Test

There were no significant relationships between sex, age group, marital status, educational level, or profession in the final GHQ-12 results. As well, no association was observed between the initial or final values obtained in the PHQ-9 test (Table 5).

Table . Variables in the equation, GHQ-12[a]. Variables introduced in step 1: with use of the app and GHQ-12i≥16.

Variables	b	SE	Wald chi-square (*df*)	*P* value	Odds ratio (95% CI)
With use of the app[b]	−3.859	0.612	39.697 (1)	<.001	0.021 (0.006-0.070)
Step 1a GHQ-12i≥16[c]	3.093	0.588	27.695 (1)	<.001	22.054 (6.968-69.798)
Constant[d]	−0.778	0.292	7.085 (1)	.008	0.459 (N/A)[e]

[a]GHQ-12: General Health Questionnaire-12

[b]With use of the app denotes a binary indicator that captures whether participants engaged with the mobile app during the intervention period (coded 1 for use, 0 for nonuse).

[c]Step 1a GHQ-12i≥16 is a baseline severity flag identifying individuals whose initial GHQ-12 score met or exceeded the clinical threshold of 16, indicating moderate-to-severe anxiety symptoms (coded 1 for ≥16, 0 otherwise).

[d]Constant represents the model intercept—the log-odds of screening positive when both predictors are set to zero (that is, for nonapp users with a baseline GHQ-12 <16).

[e]N/A: not available.

The model used a dependent variable representing the presence of psychopathologies at the conclusion of the study (GHQ-12*f*≥16), and as independent variables, the use of the Psicovida app and the variable (GHQ-12i≥16), which represents the presence of psychopathologies at the beginning of the research (determined by the GHQ-12 test itself). The model demonstrates a *P* value below .05 in the omnibus test, suggesting it significantly contributes to understanding the event in question.

The independent variables of the model explain up to 54% of the dependent variable, and the correct classification level of the model is 83%. The first independent variable, the use of Psicovida application, has a negative coefficient, an OR <1, and therefore represents a protective factor, contrary to what happens with the initial presence of psychopathologies, which shows a positive coefficient, an OR >1; therefore, it is considered a risk factor.

Discussion

Principal Findings

In this study, we evaluated the effectiveness of Psicovida, a mobile app–based CBT intervention, in reducing depressive symptoms and emotional distress among health care workers. The results demonstrated its efficacy as a mental health treatment option, with high user engagement, positive feedback, and significant reductions in depression severity—nearly half of the participants showed improvement, and 20% achieved full remission after three months. These findings, comparable to those from face-to-face CBT, indicate that Psicovida could be a scalable solution in resource-limited settings.

Moreover, the data revealed that demographic factors such as age, sex, education, profession, and marital status did not significantly influence the intervention's success, particularly in urban areas like Quito. This suggests that mHealth

XSL•FO
RenderX

applications like Psicovida can be effective across diverse user profiles, making it a versatile tool for addressing mental health challenges in health care workers.

Perspectives

A pivotal aspect of this investigation centers on the efficacy of Psicovida in ameliorating symptoms of anxiety and depression. The initial findings indicated that 16.77% of app users reported persistent symptoms of anxiety and depression, a figure that notably declined to 9% following the app's use. This contrasted with the control group, which exhibited a negligible shift in symptomatology, underscoring the app's potential therapeutic value. This outcome resonates with the results observed in other mHealth interventions, such as the MHapp app, which amalgamates three components (Moodkit, MoodPrism, and MoodMission) rooted in CBT principles, achieving similar reductions in anxiety and depression symptoms [23].

Similarly, the Mindset for Depression app, which combines structured cognitive behavioral tasks with therapist guidance, has demonstrated comparable reductions in depression severity [24]. In contrast, user-driven apps focusing on self-management and flexible health-tracking features, as highlighted in another study, show the importance of personalization in enhancing user engagement and outcomes. These parallels further underscore Psicovida's potential in reducing depression and emotional distress, even without direct therapist involvement, demonstrating the scalability and accessibility of its design in supporting mental health among health care workers.

Intellicare, another app developed with Intervention Technology and incorporating diverse media such as videos, audio files, and animations [25], mirrors Psicovida's approach in behavior and cognition modification. A distinctive feature of Intellicare is its provision of brief individualized assistance by professionals, a facet absent in Psicovida. Despite this difference, Intellicare's impact on reducing anxious and depressive symptoms, as assessed by the PHQ-9 and Generalized Anxiety Disorder-7 questionnaires, showcases significant parallels with Psicovida, reinforcing the potential of varied mHealth designs to effectively address mental health challenges.

The discourse extends beyond symptom reduction to the enhancement of positive psychological attributes, as evidenced by apps like OneUS and My Coping Plan [26]. These apps aim to cultivate positive emotions and thoughts, thereby elevating general well-being. The OneUS app, for instance, demonstrated a significant uplift in psychological well-being among its users, suggesting that mHealth interventions can play a crucial role in not only mitigating mental health symptoms but also in promoting positive mental states. This dual capability underscores the transformative potential of mHealth solutions in both therapeutic and preventive mental health contexts [27].

Writing exercises have long been recognized as a valuable tool in digital therapeutic interventions for their ability to foster self-reflection and cognitive restructuring. In Psicovida, these activities serve as a core component, encouraging users to explore their thoughts and emotions in a structured manner to facilitate changes in thinking patterns [28]. Compared to apps like Wysa, which use artificial intelligence (AI) for real-time, adaptive responses to users' concerns, Psicovida emphasizes a guided and introspective process. However, integrating AI into Psicovida could enhance interactivity by offering personalized feedback or tailored guidance during writing tasks. This hybrid model would merge the depth of structured exercises with the adaptability of AI, potentially increasing user engagement and broadening its appeal [29].

Despite the encouraging trajectories observed in Psicovida and similar mHealth apps [30], the research landscape in Latin America, and indeed globally, reveals a pronounced scarcity in scientifically rigorous studies validating the clinical efficacy of these digital interventions. The prevailing emphasis on app usability and feasibility studies, while critical, underscores a nascent field that is yet to fully embrace the rigorous evaluation of clinical effectiveness [31]. This gap highlights an urgent need for comprehensive, long-term studies that can provide a robust evidence base to inform the integration of mHealth solutions into standardized mental health care practices, especially in Latin America [32].

Psicovida's adaptability is further demonstrated by its potential to serve diverse user groups. For instance, adaptations inspired by Stay Strong, an app used to support the emotional well-being of Australian Indigenous groups [33], could make Psicovida effective in addressing stress and anxiety in incarcerated populations. Similarly, features from Intellect, which aids medical students in managing subclinical obsessive-compulsive disorder symptoms [34], could be incorporated to support those facing high stress levels. In the aftermath of global crises, such as the COVID-19 pandemic, Psicovida could adopt elements from apps like Foundations to provide structured and accessible mental health support [35].

The success of mHealth apps like Psicovida, however, depends on addressing both patient needs and mental health professionals' recommendations. Psychiatrists often cite barriers such as a lack of individualization, inadequate data privacy protections, and the perception of these tools as supplements rather than replacements for in-person therapy, especially in complex cases [36]. Psicovida addresses many of these concerns by using CBT principles, offering an intuitive design, personalized content, and flexible scheduling, which optimize user experiences and complement traditional treatments. Nevertheless, maximizing the impact of such tools requires respecting the therapeutic relationship and reinforcing clinical work, ensuring their acceptance by both patients and professionals [37].

In synthesizing these insights, it becomes evident that mHealth applications like Psicovida hold considerable promise in transforming mental health interventions. Their capacity to transcend traditional demographic limitations, coupled with their efficacy in reducing mental health symptoms and enhancing positive psychological well-being, positions them as valuable assets in the mental health care continuum [38-41]. However, the journey towards their full integration into mainstream mental health services necessitates a concerted effort to bolster the scientific rigor of efficacy studies, ensuring these digital tools can be deployed with confidence in clinical and preventive mental health settings. As this field continues

XSL•FO
RenderX

to evolve [42-44], the intersection of technology and psychology heralds a new era in mental health care, characterized by accessibility, personalization, and a holistic approach to well-being.

Limitations

Despite the significant improvements observed in mental health outcomes for health care workers using the 'Psicovida' mobile app, it is important to acknowledge and critically assess several limitations inherent to these findings, aligning with the best guidelines for eHealth practices [37]. First, the quasi-experimental design presents limitations. While the study ensured comparable characteristics between groups, a fully randomized controlled trial (RCT) would enhance internal validity by further reducing the likelihood of confounding variables influencing outcomes.

Second, this study was conducted in a single public hospital in Quito, Ecuador, which may limit the generalizability of the findings to other health care settings or regions. Factors such as organizational culture, health care policies, and available resources vary across institutions and countries, potentially impacting the app's effectiveness in other contexts. Expanding future studies to include multiple institutions in diverse geographic locations would provide more robust evidence of the app's applicability.

Third, the study intervention lasted only 3 months, making it challenging to assess the long-term sustainability of the app's impact on mental health outcomes. Mental health interventions often require ongoing support to maintain improvements, and it remains unclear whether participants' mental health would regress postintervention without continued app use or alternative support mechanisms. Additionally, participant engagement with the app varied, with 34% of participants using it for 10 - 12 weeks, 42% (n=37) for 7 - 9 weeks, and 24% (n=21) for fewer than 6 weeks. However, the study did not explore how these variations in engagement levels might have influenced outcomes. Understanding the relationship between frequency and intensity of app use and mental health improvements would offer valuable insights into optimal usage patterns.

Another limitation involves the reliance on self-reported measures, such as the PHQ-9 and GHQ-12, which may introduce bias as participants could underreport or overreport symptoms due to social desirability or personal perception. While these instruments are widely validated, incorporating objective measures of mental health, like clinical interviews or physiological data (eg, sleep patterns or stress biomarkers), could provide more comprehensive and unbiased assessments.

Finally, typical limitations in eHealth trials apply here, including the lack of blinding for participants, which could introduce expectation bias. Additionally, the multiplicity of outcomes assessed increases the risk of type I errors. Privacy concerns and technical issues, though minimal, were not fully explored in this study but should be considered in future research, as unintended effects in eHealth interventions may also include both negative and positive incidental findings.

In this non-RCT, structured reminders and prompts were used to encourage regular app engagement, which may have positively influenced adherence and outcomes. However, in a routine application setting, these elements may be absent, potentially impacting user adoption and consistent usage. Additionally, the level of technical support and structured monitoring available during the trial might not be feasible in typical usage scenarios. The absence of these supportive features in everyday settings could lead to decreased engagement, highlighting the need for integrating alternative engagement strategies, such as in-app notifications or occasional follow-up messages, to maintain effectiveness outside the structured environment of a clinical trial.

Generalizability

The findings of this trial have implications for similar health care settings but may have limited generalizability to broader populations. Given that this study focused specifically on health care workers, the results may not directly apply to a general internet population or patients outside of the study setting. Health care workers face unique stressors and mental health challenges, which may not be reflective of other groups, limiting the applicability of the intervention's effects in different contexts.

Moreover, certain elements of the non-RCT, such as periodic reminders and structured monitoring, may have positively influenced user engagement and adherence. In a routine application setting, the absence of these supportive elements could reduce both app usage and its effectiveness. To improve generalizability, future implementations should consider integrating support features, such as reminders or optional human guidance, which may enhance engagement and outcomes outside of a controlled trial environment. Such adaptations could facilitate the adoption and sustained use of eHealth interventions, making them more practical and effective in broader health care and community contexts.

Further Research

Future research should address these limitations to strengthen the evidence supporting the Psicovida app and similar mHealth interventions. A fully RCT design could enhance internal validity by minimizing biases that may arise from a quasi-experimental approach. Expanding the study across multiple health care institutions and regions, particularly in Latin America, is essential. The lack of robust mental health research in this region, especially among health care workers, underscores the need for broader evaluations to assess the app's effectiveness in diverse and underrepresented contexts.

As this study concluded, Psicovida serves primarily as an adjuvant to traditional mental health treatments rather than a standalone solution. Future research should explore how app-based interventions like Psicovida can be integrated with conventional therapeutic approaches to maximize effectiveness. This is particularly relevant for individuals with more severe mental health conditions, where professional support and structured treatments are indispensable.

Longitudinal studies are needed to evaluate the long-term effects of the Psicovida app. While this study demonstrated promising short-term outcomes, further investigation is required to determine whether the observed mental health improvements

XSL•FO
RenderX

persist over time and to identify the optimal frequency or duration of app use. These studies should also explore the potential need for supplementary interventions to sustain benefits, ensuring the app remains effective as a supportive resource.

Comparative studies between Psicovida, other mHealth apps, and alternative therapeutic interventions would provide valuable insights into best practices for app design and implementation. Such research could help refine strategies to address the unique challenges faced by health care workers and adapt interventions to specific cultural or occupational contexts.

Additionally, future research should examine the app's effectiveness across various health care roles (eg, doctors, nurses, and administrative staff) and investigate the influence of workplace stressors or social support. Understanding how user engagement impacts mental health outcomes is critical, as it can inform strategies to optimize adherence and enhance effectiveness. Incorporating objective mental health assessments, such as clinical interviews or physiological measures, alongside self-reported outcomes would further bolster the evidence base.

By addressing these research gaps, future studies can help refine and expand the application of Psicovida and similar tools, making them more practical and effective in broader health care and community contexts.

Conclusions

The Psicovida app represents a significant advancement in the accessibility and convenience of mental health resources, offering a valuable tool for users to access support anytime and anywhere. This is particularly beneficial for individuals facing barriers to traditional mental health services due to geographic, social, or economic constraints. The app incorporates a variety of therapeutic approaches, including CBT and mindfulness, and provides interactive tools for emotional support, therapeutic guidance, and progress monitoring.

Despite its benefits, Psicovida should not be seen as a replacement for face-to-face therapy, but rather as a complementary resource. The effectiveness of the app can vary based on the user's specific condition, the app's design quality, and the user's engagement. Therefore, it's crucial to have realistic expectations based on scientific evidence and to view such apps as part of a broader mental health strategy.

Acknowledgments

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The study was financially supported by the principal author (SLM-O).

Data Availability

Research data supporting this publication are available from Open Society Foundations (OSF) repository [19].

Authors' Contributions

SLM-O conceived and designed the study, curated and analyzed the data, developed the methodological framework and software tools, led the investigation, managed project resources and administration, supervised the research team, validated and visualized the results, drafted the original manuscript, and coordinated all subsequent revisions. RVAM refined the methodological approach and contributed substantial edits during manuscript revision. DSG provided critical editorial feedback and improved the clarity of the manuscript. JP-B reviewed the manuscript and enhanced its coherence and style through targeted edits. MF-H performed additional formal analyses and worked on manuscript revisions to ensure analytic accuracy. All authors reviewed and approved the final manuscript.

Conflicts of Interest

One of the authors, SLM-O, is the developer and owner of the "Psicovida" app used in this study, which may present a potential conflict of interest. However, the research was conducted with transparency, and all findings were independently analyzed to ensure objectivity. No additional commercial or financial relationships influenced the outcomes of this study.

Multimedia Appendix 1
Additional figures and tables.
[DOCX File, 11448 KB - humanfactors_v12i1e58943_app1.docx]

References

1. WHO highlights urgent need to transform mental health and mental health care. World Health Organization. URL: https://www.who.int/news/item/17-06-2022-who-highlights-urgent-need-to-transform-mental-health-and-mental-health-care [accessed 2024-10-05]
2. Lewis M, Palmer VJ, Kotevski A, et al. Rapid design and delivery of an experience-based co-designed mobile app to support the mental health needs of health care workers affected by the COVID-19 pandemic: Impact Evaluation Protocol. JMIR Res Protoc 2021 Mar 9;10(3):e26168. [doi: 10.2196/26168] [Medline: 33635823]

XSL•FO
RenderX

3. McDuff D, Barakat A, Winbush A, et al. The Google Health Digital Well-Being Study: protocol for a digital device use and well-being study. JMIR Res Protoc 2024 May 14;13(1):e49189. [doi: 10.2196/49189] [Medline: 38743938]
4. Muñoz-Ortega S, Santamaría-Guayaquil D, Pluas-Borja J, et al. Mental health in healthcare workers post-COVID-19: a Latin American review and insights into personalized management strategies. J Pers Med 2024 Jun 25;14(7):680. [doi: 10.3390/jpm14070680] [Medline: 39063934]
5. Rosales Vaca KM, Cruz Barrientos OI, Girón López S, et al. Mental health of healthcare workers of Latin American countries: a review of studies published during the first year of COVID-19 pandemic. Psychiatry Res 2022 May;311:114501. [doi: 10.1016/j.psychres.2022.114501] [Medline: 35320759]
6. Imamura K, Kawakami N, Tsuno K, Tsuchiya M, Shimada K, Namba K. Effects of web-based stress and depression literacy intervention on improving symptoms and knowledge of depression among workers: a randomized controlled trial. J Affect Disord 2016 Oct;203:30-37. [doi: 10.1016/j.jad.2016.05.045] [Medline: 27280960]
7. Reins JA, Boß L, Lehr D, Berking M, Ebert DD. The more I got, the less I need? Efficacy of Internet-based guided self-help compared to online psychoeducation for major depressive disorder. J Affect Disord 2019 Mar 1;246:695-705. [doi: 10.1016/j.jad.2018.12.065] [Medline: 30611913]
8. Larsen ME, Nicholas J, Christensen H. A systematic assessment of smartphone tools for suicide prevention. PLoS ONE 2016;11(4):e0152285. [doi: 10.1371/journal.pone.0152285] [Medline: 27073900]
9. Firth J, Torous J, Nicholas J, Carney R, Rosenbaum S, Sarris J. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. J Affect Disord 2017 Aug 15;218(15–22):15-22. [doi: 10.1016/j.jad.2017.04.046] [Medline: 28456072]
10. Steinkamp JM, Goldblatt N, Borodovsky JT, et al. Technological interventions for medication adherence in adult mental health and substance use disorders: a systematic review. JMIR Ment Health 2019 Mar 12;6(3):e12493. [doi: 10.2196/12493] [Medline: 30860493]
11. Killikelly C, He Z, Reeder C, Wykes T. Improving adherence to web-based and mobile technologies for people with psychosis: systematic review of new potential predictors of adherence. JMIR Mhealth Uhealth 2017 Jul 20;5(7):e94. [doi: 10.2196/mhealth.7088] [Medline: 28729235]
12. Miralles I, Granell C, Díaz-Sanahuja L, et al. Smartphone apps for the treatment of mental disorders: systematic review. JMIR Mhealth Uhealth 2020 Apr 2;8(4):e14897. [doi: 10.2196/14897] [Medline: 32238332]
13. Drissi N, Ouhbi S, Marques G, de la Torre Díez I, Ghogho M, Janati Idrissi MA. A systematic literature review on e-mental health solutions to assist health care workers during COVID-19. Telemed J E Health 2021 Jun;27(6):594-602. [doi: 10.1089/tmj.2020.0287] [Medline: 32970532]
14. Lai L, Sanatkar S, Mackinnon A, et al. Testing the effectiveness of a mobile smartphone app designed to improve the mental health of junior physicians: protocol for a randomized controlled trial. JMIR Res Protoc 2024 Sep 19;13(1):e58288. [doi: 10.2196/58288] [Medline: 39298756]
15. Polhemus A, Simblett S, Dawe-Lane E, et al. Health tracking via mobile apps for depression self-management: qualitative content analysis of user reviews. JMIR Hum Factors 2022 Nov 23;9(4):e40133. [doi: 10.2196/40133] [Medline: 36416875]
16. Psicovida. URL: https://psicovida.netlify.app [accessed 2025-07-30]
17. Monitoreo del coronavirus covid-19 en Ecuador [Web page in Spanish]. Coronavirus Ecuador. URL: https://www.covid19ecuador.org/ecuador [accessed 2024-10-05]
18. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126. [doi: 10.2196/jmir.1923] [Medline: 22209829]
19. Haro MAF. APP: mental health intervention model for healthcare personnel. 2024 Oct URL: https://clinicaltrials.gov/study/NCT06650449 [accessed 2025-07-30]
20. Muñoz S. Aplicativo Psicovida. OSF. 2024 Feb 8. URL: https://osf.io/zt8a9/ [accessed 2024-11-08]
21. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613. [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
22. Hankins M. The reliability of the twelve-item general health questionnaire (GHQ-12) under realistic assumptions. BMC Public Health 2008 Oct 14;8(1):355. [doi: 10.1186/1471-2458-8-355] [Medline: 18854015]
23. Rickard N, Arjmand HA, Bakker D, Seabrook E. Development of a mobile phone app to support self-monitoring of emotional well-being: a mental health digital innovation. JMIR Ment Health 2016 Nov 23;3(4):e49. [doi: 10.2196/mental.6202] [Medline: 27881358]
24. Wilhelm S, Bernstein EE, Bentley KH, et al. Feasibility, acceptability, and preliminary efficacy of a smartphone app–led cognitive behavioral therapy for depression under therapist supervision: open trial. JMIR Ment Health 2024 Apr 9;11:e53998. [doi: 10.2196/53998] [Medline: 38592771]
25. Mohr DC, Tomasino KN, Lattie EG, et al. IntelliCare: an eclectic, skills-based app suite for the treatment of depression and anxiety. J Med Internet Res 2017 Jan 5;19(1):e10. [doi: 10.2196/jmir.6645] [Medline: 28057609]
26. Dahl CJ, Wilson-Mendenhall CD, Davidson RJ. The plasticity of well-being: a training-based framework for the cultivation of human flourishing. Proc Natl Acad Sci U S A 2020 Dec 22;117(51):32197-32206. [doi: 10.1073/pnas.2014859117] [Medline: 33288719]

XSL•FO
RenderX

27. Uribe FAR, Favacho MFM, Moura PMN, Patiño DMC, da Silva Pedroso J. Effectiveness of an app-based intervention to improve well-being through cultivating positive thinking and positive emotions in an adult sample: study protocol for a randomized controlled trial. Front Psychol 2023 Jul 19;14. [doi: 10.3389/fpsyg.2023.1200960]
28. Thakkar A, Gupta A, De Sousa A. Artificial intelligence in positive mental health: a narrative review. Front Digit Health 2024;6:1280235. [doi: 10.3389/fdgth.2024.1280235] [Medline: 38562663]
29. Chang CL, Sinha C, Roy M, Wong JCM. AI-led mental health support (Wysa) for health care workers during COVID-19: service evaluation. JMIR Form Res 2024 Apr 19;8(1):e51858. [doi: 10.2196/51858] [Medline: 38640476]
30. Frías Á, Palma C, Salvador A, et al. B·RIGHT: Usability and satisfaction with a mobile app for self-managing emotional crises in patients with borderline personality disorder. Australas Psychiatry 2021 Jun;29(3):294-298. [doi: 10.1177/1039856220924321] [Medline: 32438869]
31. Karyotaki E, Efthimiou O, Miguel C, et al. Internet-Based Cognitive Behavioral Therapy for Depression: A Systematic Review and Individual Patient Data Network Meta-analysis. JAMA Psychiatry 2021 Apr 1;78(4):361-371. [doi: 10.1001/jamapsychiatry.2020.4364] [Medline: 33471111]
32. Garay CJ, Celleri M. Aplicaciones móviles en salud mental: percepción y perspectivas en Argentina [Article in Spanish]. PSICODEBATE 2022 Jun 1;22(1):38-48. [doi: 10.18682/pd.v22i1.4869]
33. Perdacher E, Kavanagh D, Sheffield J, Dale P, Heffernan E. The use of a digital well-Being app (Stay Strong App) with indigenous people in prison: randomized controlled trial. JMIR Ment Health 2024 Dec 6;11(1):e53280. [doi: 10.2196/53280] [Medline: 39642362]
34. Lee Yoon Li M, Lee Si Min S, Sündermann O. Efficacy of the mHealth app intellect in improving subclinical obsessive-compulsive disorder in university students: randomized controlled trial with a 4-week follow-up. JMIR Mhealth Uhealth 2024 Dec 16;12(1):e63316. [doi: 10.2196/63316] [Medline: 39680884]
35. Catuara-Solarz S, Skorulski B, Estella-Aguerri I, et al. The efficacy of "Foundations," a digital mental health app to improve mental well-being during COVID-19: proof-of-principle randomized controlled trial. JMIR Mhealth Uhealth 2022 Jul 1;10(7):e30976. [doi: 10.2196/30976] [Medline: 34978535]
36. Kidman PG, Curtis RG, Watson A, Maher CA. When and why adults abandon lifestyle behavior and mental health mobile apps: scoping review. J Med Internet Res 2024 Dec 18;26(1):e56897. [doi: 10.2196/56897] [Medline: 39693620]
37. Hildebrand AS, Planert J, Machulska A, Margraf LM, Roesmann K, Klucken T. Exploring psychotherapists' attitudes on internet- and mobile-based interventions in Germany: thematic analysis. JMIR Form Res 2024 Nov 7;8:e51832. [doi: 10.2196/51832] [Medline: 39510514]
38. Fitzpatrick KK, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. JMIR Ment Health 2017 Jun 6;4(2):e19. [doi: 10.2196/mental.7785] [Medline: 28588005]
39. Talkspace. URL: https://www.talkspace.com/ [accessed 2024-10-05]
40. Fonseca Pedrero E, Bobes-Bascarán M, Gimeno-Peón A, et al. Tratamientos psicológicos empíricamente apoyados para adultos: una revisión selectiva [Article in Spanish]. Psicothema 2021 May 1;2(33):188-197. [doi: 10.7334/psicothema2020.426]
41. Bégin C, Berthod J, Martinez LZ, Truchon M. Use of mobile apps and online programs of mindfulness and self-compassion training in workers: a scoping review. J Technol Behav Sci 2022;7(4):477-515. [doi: 10.1007/s41347-022-00267-1] [Medline: 36091081]
42. Fiol-DeRoque MA, Serrano-Ripoll MJ, Jiménez R, et al. A mobile phone–based intervention to reduce mental health problems in health care workers during the COVID-19 pandemic (PsyCovidApp): randomized controlled trial. JMIR Mhealth Uhealth 2021 May 18;9(5):e27039. [doi: 10.2196/27039] [Medline: 33909587]
43. Graser J, Stangier U. Compassion and loving-kindness meditation: an overview and prospects for the application in clinical samples. Harv Rev Psychiatry 2018;26(4):201-215. [doi: 10.1097/HRP.0000000000000192] [Medline: 29975338]
44. Liu C, Chen H, Zhang A, et al. The effects of short video app-guided loving-kindness meditation on college students' mindfulness, self-compassion, positive psychological capital, and suicide ideation. Psicol Refl Crít 2023 Oct 30;36(1):32. [doi: 10.1186/s41155-023-00276-w]

Abbreviations

AI: artificial intelligence
CBT: cognitive behavioral therapy
CONSORT: Consolidated Standards of Reporting Trials
GHQ-12: General Health Questionnaire-12
OR: odds ratio
PHQ-9: Patient Health Questionnaire-9
RCT: randomized controlled trial

XSL•FO
RenderX

Edited by A Kushniruk; submitted 28.03.24; peer-reviewed by A Eisingerich, D Patel; revised version received 29.05.25; accepted 03.06.25; published 05.08.25.

Please cite as:
Muñoz-Ortega SL, Alvarado Muñoz RV, Santamaria Guayaquil D, Pluas-Borja J, Faytong-Haro M
Mobile-Based Cognitive Behavioral Therapy for Health Care Workers' Mental Health in Ecuador: Quasi-Experimental Study
JMIR Hum Factors 2025;12:e58943
URL: https://humanfactors.jmir.org/2025/1/e58943
doi:10.2196/58943

© Sandra Lorena Muñoz Ortega, Rubén Vladimir Alvarado Muñoz, Daniela Santamaria Guayaquil, Jade Pluas-Borja, Marco Faytong-Haro. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 5.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Finnish Registry-Based Protocol for Screening and Management of Fatigue and Cognitive Problems in Multiple Sclerosis: Observational Study

Päivi Hämäläinen[1,2,3], PhD; Elina Lämsä[4], PhD; Matias Viitala[1], MSc; Hanna Kuusisto[3,5], MD, PhD; Marja Niiranen[6], MD, PhD; Sari Avikainen[7], MD; Juha Puustinen[8], MD, PhD; Mervi Ryytty[9,10], MD, PhD; Juhani Ruutiainen[2,11], MD, PhD; Merja Soilu-Hänninen[11,12], MD, PhD

[1]StellarQ Ltd, Eerikinkatu 3 A, Turku, Finland
[10]Research Unit of Clinical Medicine, University of Oulu, Oulu, Finland
[11]Clinical Neurosciences, University of Turku, Turku, Finland
[12]Neurocenter and Division of Clinical Neurosciences, Turku University Hospital, Turku, Finland
[2]Finnish Neuro Society, Masku, Finland
[3]Department of Health and Social Management, University of Eastern Finland, Kuopio, Finland
[4]Novartis Finland, Espoo, Finland
[5]Department of Neurology, Tampere University Hospital, Tampere, Finland
[6]Neuro Center, Neurology, Kuopio University Hospital, Kuopio, Finland
[7]Neurology, Nova Hospital in Central Finland, Jyväskylä, Finland
[8]Unit of Neurology, Satasairaala Central Hospital, Satakunta Wellbeing Services County, Pori, Finland
[9]Medical Research Center and Neurocenter, Oulu University Hospital, Oulu, Finland

Corresponding Author:
Päivi Hämäläinen, PhD
StellarQ Ltd, Eerikinkatu 3 A, Turku, Finland

Abstract

Background: Digital patient registries are actively used to monitor long-term diseases. However, their potential in symptom management remains underused.

Objective: This study aimed to report on the Finnish registry-based protocol to screen and manage cognitive symptoms and fatigue in multiple sclerosis (MS). Data on a sample collected during the first 2 years are presented.

Methods: At the beginning of 2021, a Finnish protocol to screen and manage patient-perceived concerns related to cognition and fatigue, together with self-assessment of disease severity, symptoms, and quality of life (QoL) annually, was introduced. The Symbol Digit Modalities Test (SDMT), the Multiple Sclerosis Neuropsychological Questionnaire (MSNQ), the Fatigue Scale for Motor and Cognitive Functions (FSMC), as well as the Patient-Reported Expanded Disability Status Scale, the Visual Analog Scales, and the Euro QoL-5 Dimension were implemented into the Finnish MS registry. To support symptom management, patients were offered feedback reports based on the results of the FSMC and the MSNQ. The implementation of the protocol was evaluated in 5 Finnish wellbeing services counties.

Results: Our sample from the beginning of 2021 to the end of 2022 includes data on 430 patients. A total of 86 (20%) patients have been assessed with the SDMT, whereas 329 (76.5%) patients have filled out the FSMC, and 172 (40.0%) patients have completed the MSNQ. The mean SDMT score is 49.0 (SD 13.56), MSNQ score is 35.3 (SD 9.39), total FSMC score is 63.0 (SD 22.49), and subscores for motor and cognitive fatigue are 31.6 (SD 11.43) and 31.5 (SD 11.68), respectively. The SDMT did not correlate with the MSNQ or the FSMC. Instead, the SDMT, MSNQ, and the FSMC correlated significantly with QoL.

Conclusions: Fatigue and cognitive problems have an effect on QoL. In our preliminary sample, patient reports of cognitive problems and especially fatigue were conducted more frequently than the objective evaluation of processing speed. Although the Finnish MS registry offers a digital platform for the systematic screening of fatigue and cognitive problems, further education is needed to support the implementation of the protocol.

(JMIR Hum Factors 2025;12:e67990) doi:10.2196/67990

KEYWORDS

multiple sclerosis; electronic health records; digital health; patient-generated data; cognition; fatigue

XSL•FO
RenderX

Introduction

Multiple sclerosis (MS) registries are in active use in many European countries [1], and the need to develop registry-based data collection is recognized worldwide [2]. Registries help with patients' follow-up and care [3,4]. There is a need to increase patient involvement and the collection of patient-generated data. Tools to identify and manage the wide range of patient-perceived MS-related concerns are needed. Moreover, there is a special need for early recognition of symptoms and factors that affect productivity and quality of life (QoL).

According to a mapping survey, there are 19 MS registries in Europe [1]. Patient-reported outcomes (PROs) are included only in 7 registries. One of those is the Finnish MS registry with a unique patient interface. The Finnish MS registry was launched in 2014 to enable digital treatment monitoring, systematic follow-up of incidence and prevalence, as well as the course of the disease [5]. At present, the registry is used in 18 out of 21 of Finland's wellbeing services counties, including all 5 university hospitals. A patient interface for PROs was created in 2017, allowing patients to contribute health data on their symptoms, QoL, and overall well-being to inform health care professionals (HCPs) and to support patient-centered treatment decisions [6,7]. The PROs implemented in the patient interface include validated patient-reported outcome measures (PROMs) as well as more informal PRO data.

Fatigue has been presented as the most common symptom of MS, with a rate of up to 80% [8-10], whereas approximately 50% of patients with MS have been shown to have cognitive deficits [11]. The negative effects of cognitive problems and fatigue on working ability have been shown in registry-based studies [12,13]. Self-perceived cognitive problems and fatigue have shown a linear relationship with days on sick leave and a decrease in weekly working hours and self-perceived working ability [12]. Furthermore, impaired processing speed as evaluated by the Symbol Digit Modalities Test (SDMT) has been found to be related to decreased work productivity as measured by days on sick leave and total annual income [13]. Cognitive deficits and fatigue, as invisible symptoms, often remain underdiagnosed, and treatment options are considered too late to promote working ability. Despite these findings, patient reports on fatigue or cognitive problems are not systematically included in MS registries [1]. Because the underlying mechanisms of fatigue are multifaceted and no clear definition for the symptom exists, the assessment is particularly challenging and relies heavily on patient report [8]. The fact that fatigue and cognitive concerns have been reported to be confounded with depression and problems with mood further complicates the assessment of the symptoms [14]. At the same time, the detrimental effects of fatigue and cognitive concerns on employment call for early diagnostics and treatment.

In 2020, a Finnish initiative called Invisible Symptoms in MS was presented to meet the need to consider cognitive symptoms and fatigue systematically and in a cost-effective way already at the early phases of the disease. The primary aims of the initiative were to promote early and holistic evaluation and treatment of MS and to offer patients tools to evaluate and manage fatigue and cognitive problems by themselves. The objective of this paper is to describe the protocol to screen and manage fatigue and cognitive problems by using the Finnish MS registry. Furthermore, descriptive statistics are reported on the data collected during the first 2 years following the implementation of the procedure, based on a sample from 5 well-being services counties.

Methods

Finnish MS Registry

The Finnish MS registry is a browser-based registry for public health care organizations [5]. The software service is provided by StellarQ Ltd. The use of the registry is voluntary, and each well-being services county has decided individually whether to acquire it. At the time of the study, 17 out of 21 counties used the registry with data on over 12,500 patients with MS. This is estimated to cover approximately 90% of all patients with MS in Finland [6]. The content of the registry is described in detail elsewhere [5]. All patients registered in the clinician-based registry can use the patient interface, called MyMS (patient interface of the Finnish MS registry), which can be accessed with the national Finnish electronic authorization. A total of 1201 patients with MS were reported to have recorded data into MyMS by the beginning of 2023 [6]. The content of the patient interface is described in our previous studies [6,7]. MyMS has existed since 2017 with the possibility to record background and lifestyle factors, suspected relapses, medications, need for assistance and social support, rehabilitation, and other diseases. Furthermore, 3 self-rating questionnaires, the Fatigue Severity Scale, the Multiple Sclerosis Impact Scale-29, and the 15D QoL scale have been available for users to assess their conditions since 2017 [6].

Registry-Based Protocol for Screening of Fatigue and Cognitive Problems

The protocol for registry-based evaluation of cognitive problems and fatigue, along with the patient-perceived disability and QoL, was accepted by the Expert Academic Advisory Board of the Finnish MS registry. The protocol has been available since the beginning of 2021. The annual protocol includes the SDMT [15] as well as the following PROMs: the Patient-Reported Expanded Disability Status Scale (PREDSS) [16,17], the Visual Analog Scales (VASs) on neurological symptoms, the Euro QoL-5 dimensions [18], the Multiple Sclerosis Neuropsychological Questionnaire (MSNQ) [19,20], and the Fatigue Scale for Motor and Cognitive Functions (FSMC) [20,21]. The questionnaires are scored according to the standardized procedures of the instruments. The content of the protocol is described in Table 1.

XSL•FO
RenderX

Table . The annual protocol for screening and management of fatigue and cognitive problems by using the Finnish multiple sclerosis (MS) registry.

Assessment tool	Domain assessed	Assessment method / scoring	Role in the protocol
Symbol Digit Modalities Test (SDMT) in clinical MS registry	Cognitive processing speed	Assessed by a nurse; the patient matches symbols with digits in 90 seconds.	Screens for cognitive impairment. Neuropsychological assessment is recommended if the score declines by more than 10% or 4 points compared with baseline or previous assessment.
Patient-Reported Expanded Disability Status Scale (PREDSS) in MyMS[a]	Disease severity	Patient-reported evaluation of disease severity on a scale from 0 to 9, corresponding to the clinician-rated Expanded Disability Status Scale (EDSS).	Provides a perspective on patient-perceived disease severity.
Visual Analog Scales on Neurological Symptoms (VAS) in MyMS	Neurological symptoms	Patient-reported evaluation of symptom severity using a 0 to 100 scale, where 100 indicates the worst possible symptom.	Provides perspective on patient-perceived symptom severity. VAS mood score can be used to control the effect of mood on fatigue and cognition.
Euro Quality of Life-5 Dimension (EQ-5D) in MyMS	Quality of life	Patient-reported QoL evaluation using an index score from 0 to 1, where 1 indicates perfect QoL.	Provides perspective on patient-perceived quality of life.
Fatigue Scale for Motor and Cognitive Functions (FSMC) in MyMS	Motor, cognitive, and overall fatigue	Patient-reported; consists of 10 items on motor fatigue and 10 on cognitive fatigue, scored from 1 (totally disagree) to 5 (totally agree). Motor and cognitive subscores up to 50 each, total score up to 100.	Provides perspective on patient-perceived fatigue. Offers feedback to the patient including score interpretations and practical guidance for managing motor and cognitive fatigue in daily life.
Multiple Sclerosis Neuropsychological Questionnaire (MSNQ) in MyMS	Cognitive problems	Patient-reported; consists of 15 items assessing cognitive problems, each scored from 0 (does not occur) to 4 (very often/very disruptive), yielding a total score ranging from 0 to 60.	Provides perspective on patient-perceived cognitive problems. Offers feedback to the patient including score interpretation and practical guidance for managing cognitive difficulties in daily life.

[a]MyMS: patient interface of the Finnish MS registry.

According to the Invisible Symptoms protocol, patients with MS were familiarized with MyMS by the MS nurses in the hospital clinics. Educational material was offered for the nurses. Patients were provided with a leaflet containing instructions on how to use MyMS (Multimedia Appendix 1).

Study Sample and Statistical Analysis

A preliminary sample collected from 5 well-being services counties between the beginning of 2021 and the end of 2022 was analysed. Date imputation, as the middle of the month or year, was used for clinical variables with partial dates. Date imputation was not needed for patient-reported data. The date of the first data entry in the patient interface was considered the index date to calculate age and disease duration. Descriptive statistics for demographics and PRO measures included means, SDs, medians, and IQRs for continuous variables, and counts and proportions for categorical variables. In addition, the proportion of missing data was reported.

For the correlation analysis, each pair of samples was created by matching the earliest dates of 2 PRO recordings that occurred within 28 days of one another. Correlation coefficients and testing were based on Pearson's product-moment correlation. For controlling and checking the false discovery rate, the Benjamini-Hochberg procedure was used as a correction for multiple comparisons. *P* values under .05 after the correction were considered significant. All data analyses were done using RStudio (version 2021.09.1; Posit, PBC).

Ethical Considerations

According to Finnish legislation, the approval of an ethical committee was not required as the study was a noninterventional registry study in which patients were not contacted. Patients permit the use of the data for research purposes including the secondary use when they log into the registry. Patients did not get any compensation for participation in the study. All used data were permitted via Finnish Social and Health Data Permit Authority Findata, which is responsible for data permits considering secondary use of data with multiple data sources and registry keepers (permission numbers THL/1809/14.02.00/2022 and THL/2104/14.05.00/2023 findata-rems-2023). Clinical data such as age at diagnosis, course of disease, disease duration, and Expanded Disability Status Scale (EDSS) [22] were extracted from the clinical MS registry and PRO data from the patient interface. All the data was analyzed in pseudonymized form, and in audited secure processing environment. Only anonymized information can be exported from the secure processing environment, and patient-level data is not publicly available to either authors or third-parties.StellarQ Ltd is the data processor for all data extracted.

XSL•FO
RenderX

Results

The Invisible Symptoms protocol was available in 17 out of 21 Finnish well-being services counties at the time of the analysis. For the purposes of the study, we evaluated a sample from 5 Finnish well-being services counties: Southwest Finland, Satakunta, Pirkanmaa, Central Finland, North Savo. The population in these 5 counties represents approximately 30% of the total population in Finland and approximately 40% of the national MS population. The 5 counties represent areas with different population density and prevalence of MS. The implementation of the protocol was voluntary and depended on the resources the clinics had to administer the SDMT and introduce MyMS to the patients.

Altogether, 430 patients in the 5 well-being services counties had recorded data related to the Invisible Symptoms protocol into MyMS. The demographic and clinical variables of the protocol sample are listed in Table 2. The mean age of the patients at the first PRO recording was 42.5 years, the median disease duration was 4.1 years, and the median EDSS score was 2.5. Over 80% of patients were women (357 out of 430) and had a relapsing-remitting course of the disease (354 out of 430).

Table . Demographic and disease-related variables of the patients with data related to the Invisible Symptoms protocol.

Variable	Protocol sample (n=430)
Age at MS[a] onset (years)	
Mean (SD)	33.1 (10.18)
Missing data, n (%)	108 (25.1)
Age at MS diagnosis (years), mean (SD)	35.5 (10.13)
Age at first PRO[b] recording (years), mean (SD)	42.5 (10.28)
Sex (female), n (%)	357 (83.0)
Disease course, n (%)	
RRMS[c]	354 (82.3)
SPMS[d]	20 (4.7)
PPMS[e]	22 (5.1)
UNS[f]	34 (7.9)
Disease duration (years), median (Q1-Q3)	4.1 (0.6-11.3)
EDSS[g]	
Median (IQR)	2.5 (2.0-3.5)
Missing data, n (%)	50 (12)
Education (total years)	
Mean (SD)	14.3 (2.73)
Missing data, n (%)	118 (27.4)
Smoking (yes), n (%)	
Number	39 (10)
Missing data	24 (6)

[a]MS: multiple sclerosis.
[b]PRO: patient-reported outcome.
[c]RRMS: relapsing-remitting MS.
[d]SPMS: secondary progressive MS.
[e]PPMS: primary progressive MS.
[f]UNS: unspecified disease course.
[g]EDSS: Expanded Disability Status Scale.

Of the 430 patients, 349 (81.2%) patients have filled out the PREDSS with a mean score of 2.9 (SD 1.67) (Table 3). Table 3 summarizes the coverage of data on the SDMT as well as the PROMs related to the Invisible Symptoms protocol. The SDMT assessment was performed on 86 (20%) patients. The PROM on fatigue, the FSMC, was completed by 76.5% (329 out of 430) of the patients, whereas the PROM on cognitive problems, the MSNQ, was completed by 40.0% (172 out of 430). When a patient completes the FSMC, they receive a feedback report summarizing the results. The report also includes information about fatigue, along with practical tips for managing it in daily life. An example of the feedback report is presented as the

XSL•FO
RenderX

Multimedia Appendix 2. A similar report is available after completing the MSNQ, summarizing the questionnaire results and providing tips for managing cognitive problems.

Table . Results on the Symbol Digit Modalities Test (SDMT) and the patient-reported measures of disability, quality of life, fatigue, cognition, and mood (n=430).

Variable	Total, n (%)	Mean (SD)	Median (IQR)
Patient Reported Expanded Disability Status Scale score (PREDSS)[a]	349 (81.2)	2.9 (1.67)	3.0 (2.0-4.0)
SDMT score	86 (20)	49.0 (13.56)	49.0 (40.8-56.8)
VAS[b] scores, neurological symptoms			
Fatigue	319 (74.2)	45.2 (27.78)	42.0 (20.0-70.0)
Cognition	295 (68.6)	34.1 (25.96)	30.0 (10.5-52.0)
Mood	264 (61.4)	25.9 (26.10)	15.0 (6.8-37.2)
Euro Quality of Life-5D score[c]	297 (69.1)	0.8 (0.19)	0.8 (0.7-0.9)
FSMC[d] scores			
Total	329 (76.5)	63.0 (22.49)	68.0 (47.0-80.0)
Motor fatigue	337 (78.4)	31.6 (11.43)	34.0 (24.0-41.0)
Cognitive fatigue	331 (77.0)	31.5 (11.68)	32.0 (23.0-41.0)
MSNQ[e] scores			
Total score	172 (40.0)	35.3 (9.39)	35.0 (28.0-41.2)
Memory and learning	215 (50.0)	2.2 (0.74)	2.3 (1.7-2.7)
Attention	291 (67.7)	2.4 (0.85)	2.3 (1.8-3.0)
Verbal, problem-solving, and behavior	199 (46.3)	2.0 (0.64)	2.0 (1.5-2.3)

[a]PREDSS 0 - 9 where 0 stands for no disability and 9 for bedridden most of the time.

[b]VAS: Visual Analog Scale. VAS scores where 0 stands for no problems and 100 for worst possible problems.

[c]EQ-5D: Euro Quality of Life–5 Dimension. Higher scores refer to better quality of life in EQ-5D (range 0 - 1).

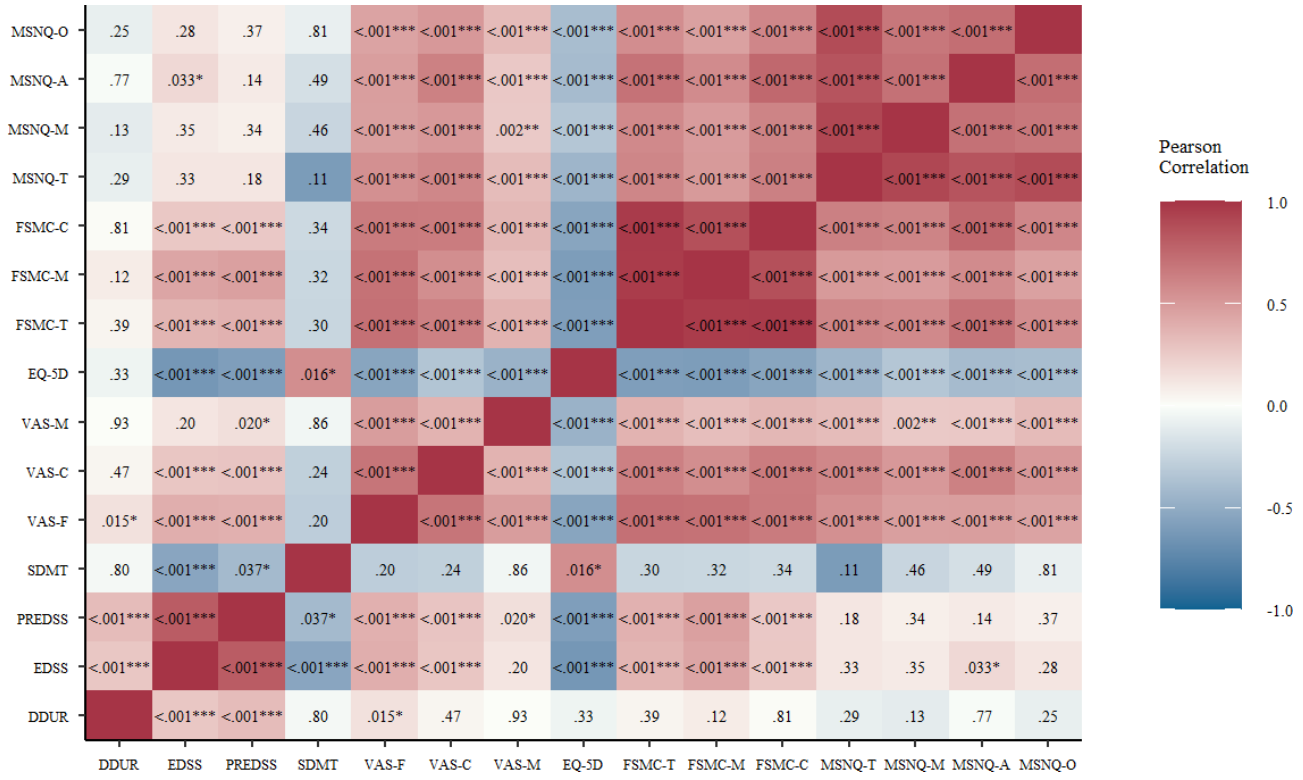
[d]FSMC: Fatigue Scale for Motor and Cognitive Functions. Higher scores refer to more severe symptoms/problems in the FSMC (total score range 20 - 100, subscore range 10 - 50).

[e]MSNQ: Multiple Sclerosis Neuropsychological Questionnaire. Higher scores refer to more severe/frequent cognitive problems in MSNQ (total score range 0 - 60, subscores are means and range from 0 to 4).

In the present sample, a significant relationship was observed between disease duration and both the EDSS and the PREDSS (Figure 1). Disease duration did not show a significant relationship with any of the measures on mood, fatigue, cognition, or QoL except for the VAS-F. The EDSS correlated significantly with the SDMT, the VAS scores on fatigue and cognition, QoL, FSMC scores, and attentional items of the MSNQ. The PREDSS showed a correlation pattern similar to that of the EDSS. In contrast to the EDSS, it also correlated significantly with the VAS-M, but not with the MSNQ scores. The SDMT did not correlate significantly with the self-reports on mood, fatigue, or cognition. Instead, a significant correlation was observed between the SDMT and QoL. The VAS scores on fatigue, cognition, and mood correlated consistently with each other and the Euro Quality of Life–5 Dimension. Both the FSMC and MSNQ scores showed significant correlations with self-reports on fatigue, cognition, mood, and QoL.

XSL•FO
RenderX

Figure 1. Correlation analysis of DDUR, disability (EDSS and PREDSS), SDMT performance, and PROMs on quality of life, fatigue, cognition, and mood. DDUR: disease duration; EDSS: Expanded Disability Status Scale; EQ-5D: Euro Quality of Life–5 Dimension; FSMC: Fatigue Scale for Motor and Cognitive Functions; MSNQ: Multiple Sclerosis Neuropsychological Questionnaire; PREDSS: Patient Reported Expanded Disability Status Scale; SDMT: Symbol Digit Modalities Test; VAS: Visual Analog Scale. *P<.05, **P<.01, ***P<.001.



Discussion

The Finnish invisible symptom initiative presents a protocol for registry-based screening and management of fatigue and cognitive concerns. It includes an annual assessment of processing speed using the SDMT, as well as patient-reported measures of disease severity, QoL, and symptoms, with an emphasis on cognitive symptoms and fatigue. For symptom management purposes, patients are provided with a feedback report on self-perceived fatigue based on the FSMC and on self-perceived cognitive problems based on the MSNQ. Our sample from 5 Finnish well-being services counties shows that the implementation of the protocol has been gradual, and patient-reported assessments have been conducted more frequently than objective assessments of processing speed. The descriptive statistics of the present sample support earlier findings on the important role of fatigue and cognitive symptoms in QoL [7,10,12,13,23].

Our protocol for the screening and management of problems and fatigue aligns with the recommendations by Kalb et al [24] for baseline screening and annual follow-up of cognitive performance in MS using the SDMT. The SDMT has been introduced as the most sensitive measure of cognitive performance and processing speed in MS [25]. The Finnish validation of the brief international cognitive assessment in MS showed that the SDMT was the most sensitive measure, also in Finnish patients with MS [20]. In the same study, the MSNQ and the FSMC were found to be suitable for self-evaluation of fatigue and cognitive symptoms, provided that the effects of mood state are adequately controlled. To assess patient perspective, the MSNQ and the FSMC were thus implemented into the patient interface of the Finnish MS registry. The VAS score on mood was used to control the effects of mood on fatigue and cognitive concerns.

Only 20% (86 out of 430) of the patients in the present sample were assessed with the SDMT. The mean SDMT score in the present registry-based sample was higher than that observed in the Finnish validation study with older and more disabled patients [20] but in line with international studies with patients of similar age and EDSS [26-29]. Over 50% (291 out of 430) of the patients of the present sample had completed the MSNQ, and over 70% (337 out of 430) of the patients had completed the FSMC, at least partially. Based on the FSMC [21], the patients of the present sample reported severe overall fatigue and moderate motor and cognitive fatigue. Based on the MSNQ [19], they also perceived at least occasional problems with cognitive functions. The patients of the present sample reported more fatigue and cognitive concerns than the patients in the Finnish validation study [20]. Further, the FSMC scores of the present sample were higher than those in another Finnish study involving patients with relapsing-remitting MS [10]. The present sample is likely overrepresented by patients with pronounced fatigue and cognitive symptoms, as these individuals were probably more often instructed to complete the questionnaires to support symptom management.

In the present sample, physical disability as measured by either the EDSS or the PREDSS correlated with all the self-report measures used, except the MSNQ. On the contrary, disease duration did not have a significant relationship with the self-reports, certifying the earlier findings that invisible

XSL•FO
RenderX

symptoms may emerge at any phase of the disease [30]. The SDMT correlated with both the EDSS and the PREDSS. The SDMT did not show a significant relationship with the self-report of depression. On the contrary, self-reports on fatigue and cognitive problems correlated significantly with those on depression. These results are in line with earlier findings that concerns with fatigue, cognition, and mood are interrelated and should all be considered in the screening and search for optimal treatment methods [11,14,19,20].

In the Invisible Symptoms initiative, feedback reports based on the patient's answers in the MSNQ and the FSMC were developed. Digital self-management tools could be further developed into web-based rehabilitation programs that integrate assessments, feedback, and strategies with individual goal setting and professional counseling. Positive effects on the management of cognitive symptoms in MS have been demonstrated not only through neuropsychological and cognitive rehabilitation, but also through the use of digital applications [31-33]. Especially cost-effective, individualized, and holistic intervention with an input from HCP has been called for [11].

This study has several limitations. The procedure for screening and managing fatigue and cognitive concerns has been in use for only 2 years and is not yet fully implemented across all regions, resulting in a relatively small sample limited to 5 of Finland's 21 wellbeing services counties—though these account for approximately 40% of the national MS population. Adoption of the protocol has likely been slowed by limited resources for conducting the SDMT and challenges in interpreting and integrating PROM data into routine clinical decision-making. Additionally, the patient interface has so far been available only as a web-based tool, which may have reduced usability; a mobile app is currently under development in response to patient feedback, particularly from younger individuals. The sample may also over-represent individuals with fatigue and cognitive symptoms, as participation is voluntary and allows patients to choose which questionnaires to complete. It is also unknown whether patients have engaged with the feedback provided. Despite these limitations, the model offers clinics a structured means to guide annual assessment and symptom management through the digital registry. Ideally, this approach would be applied systematically and include both PROMs and the SDMT, which is internationally recommended for evaluating cognitive function in people with MS.

Comprehensive management of invisible symptoms requires effective communication between HCPs and patients [34,35]. The Finnish Invisible Symptoms initiative provides a systematic approach to addressing patient-perceived cognitive problems and fatigue. Cognitive performance and patient-perceived symptoms, including fatigue and cognitive concerns, can be assessed annually using validated tools, with results accessible to patients via the patient interface and to HCPs via the clinician interface. According to a European survey published in 2019 [1], only 7 out of 19 identified MS registries include patient-derived measures. Digital registries can serve as platforms for patient-reported data, empowering individuals to take an active role in managing their disease, adhering to treatment, and making beneficial lifestyle choices. As stated by Lakin et al [34], by pairing clinical knowledge with an understanding and consideration of the patient perspective, HCPs are equipped to foster patient-centered dialogue that encourages shared decision-making and high-quality care. The need for innovative interventions addressing invisible symptoms is increasingly recognized. Positive experiences with digital solutions have already been reported, highlighting the importance of continued development and evaluation [31,33,36].

Acknowledgments

The authors gratefully acknowledge the contribution of the StellarQ IT team for building and continuously developing the Finnish MS Register IT platform. All members of the Steering Committee of the Finnish MS registry, as well as patients who have participated in the development of the patient interface of the Finnish MS registry (MyMS), are gratefully acknowledged. Building up the Finnish MS registry and the patient interface has received funding from Biogen Idec, Merck, Novartis, Roche, Sanofi Genzyme, Teva, and Business Finland. The Invisible Symptoms initiative received additional funding from Novartis. The preparation of the manuscript was in part funded by the Strategic Research Council (SRC) established within the Academy of Finland (funding number 358415).

Data Availability

The datasets generated or analyzed during this study are not publicly available due to Finnish legislation. Data are analyzed based on the Act on the Secondary Use of Social and Health Data, which restricts patient-level data from being handled only for a granted purpose, in pseudonymized form, and in an audited secure processing environment. Only anonymized information can be exported from the secure processing environment, and patient-level data are not publicly available to either authors or third parties.

Conflicts of Interest

PH is an employee of the Finnish Neuro Society and StellarQ. PH has served as an adviser or speaker for Merck, Novartis, and Sanofi. EL is an employee of Novartis Finland Oy. HK has served as an adviser or speaker for Merck, Biogen, Celgene (BMS), Novartis, Roche, Jansen Cilag, Teva, Sandoz, and Sanofi, and has received support for congress participation from Merck. MN has served as an adviser or speaker for Novartis, Roche, Merck, Sanofi, and Teva; and received support for congress participation from Novartis, Merck, Roche, and Sanofi. MR has received honoraria for lectures, advisory boards, congress visits, or for serving

XSL•FO
RenderX

as an investigator for clinical trials from Abbvie, Biogen, Merck, Novartis, Roche, Sandoz, and Sanofi. MS-H has served as an adviser or speaker for Biogen, Celgene (BMS), Novartis, Roche, Sanofi, and Teva; has received institutional research grants for clinical research from Bayer, Biogen, Merck, Novartis, and Roche; and support for congress participation from Biogen, Celgene (BMS), Novartis, Roche, Sanofi, and Teva. MV, JP, and JR declare no conflicts of interest.

Multimedia Appendix 1
Instruction leaflet describing the use of MyMS.
[PNG File, 680 KB - humanfactors_v12i1e67990_app1.png]

Multimedia Appendix 2
An example of the feedback report the patient receives after filling out the Fatigue Scale for Motor and Cognitive Functions (FSMC).
[PNG File, 407 KB - humanfactors_v12i1e67990_app2.png]

Checklist 1
STROBE checklist.
[DOCX File, 33 KB - humanfactors_v12i1e67990_app3.docx]

References

1. Glaser A, Stahmann A, Meissner T, et al. Multiple sclerosis registries in Europe - an updated mapping survey. Mult Scler Relat Disord 2019 Jan;27:171-178. [doi: 10.1016/j.msard.2018.09.032] [Medline: 30384204]
2. Bebo BF Jr, Fox RJ, Lee K, Utz U, Thompson AJ. Landscape of MS patient cohorts and registries: recommendations for maximizing impact. Mult Scler 2018 Apr;24(5):579-586. [doi: 10.1177/1352458517698250] [Medline: 28279128]
3. Ziemssen T, Hillert J, Butzkueven H. The importance of collecting structured clinical information on multiple sclerosis. BMC Med 2016 May 31;14:81. [doi: 10.1186/s12916-016-0627-1] [Medline: 27246898]
4. Trojano M, Bergamaschi R, Amato MP, et al. The Italian multiple sclerosis register. Neurol Sci 2019 Jan;40(1):155-165. [doi: 10.1007/s10072-018-3610-0] [Medline: 30426289]
5. Laakso SM, Viitala M, Kuusisto H, et al. Multiple sclerosis in Finland 2018-data from the national register. Acta Neurol Scand 2019 Nov;140(5):303-311. [doi: 10.1111/ane.13145] [Medline: 31271648]
6. Hämäläinen P, Viitala M, Kuusisto H, Ruutiainen J, Soilu-Hänninen M. MyMS: an interface for patient-reported outcomes for finnish individuals with multiple sclerosis. Int J MS Care 2024 Oct;26(Q4):273-280. [doi: 10.7224/1537-2073.2023-082] [Medline: 39386272]
7. Hämäläinen P, Viitala M, Soilu-Hänninen M, et al. Evaluation of patient-perceived fatigue in multiple sclerosis utilizing the Finnish MS registry. Mult Scler J Exp Transl Clin 2025;11(1):20552173251325098. [doi: 10.1177/20552173251325098] [Medline: 40110484]
8. Krupp L. Fatigue is intrinsic to multiple sclerosis (MS) and is the most commonly reported symptom of the disease. Mult Scler 2006 Aug;12(4):367-368. [doi: 10.1191/135248506ms1373ed] [Medline: 16900749]
9. Broch L, Simonsen CS, Flemmen H, et al. High prevalence of fatigue in contemporary patients with multiple sclerosis. Mult Scler J Exp Transl Clin 2021;7(1):2055217321999826. [doi: 10.1177/2055217321999826] [Medline: 33796331]
10. Ahvenjärvi H, Niiranen M, Simula S, et al. Fatigue and health-related quality of life depend on the disability status and clinical course in RRMS. Mult Scler Relat Disord 2023 Sep;77:104861. [doi: 10.1016/j.msard.2023.104861] [Medline: 37442075]
11. Sumowski JF, Benedict R, Enzinger C, et al. Cognition in multiple sclerosis: state of the field and priorities for the future. Neurology (ECronicon) 2018 Feb 6;90(6):278-288. [doi: 10.1212/WNL.0000000000004977] [Medline: 29343470]
12. Kobelt G, Langdon D, Jönsson L. The effect of self-assessed fatigue and subjective cognitive impairment on work capacity: the case of multiple sclerosis. Mult Scler 2019 Apr;25(5):740-749. [doi: 10.1177/1352458518769837] [Medline: 29663869]
13. Kavaliunas A, Tinghög P, Friberg E, et al. Cognitive function predicts work disability among multiple sclerosis patients. Mult Scler J Exp Transl Clin 2019;5(1):2055217318822134. [doi: 10.1177/2055217318822134] [Medline: 30729025]
14. Chalah MA, Riachi N, Ahdab R, Créange A, Lefaucheur JP, Ayache SS. Fatigue in multiple sclerosis: neural correlates and the role of non-invasive brain stimulation. Front Cell Neurosci 2015;9:460. [doi: 10.3389/fncel.2015.00460] [Medline: 26648845]
15. Smith A. Symbol digit modalities test: manual. : Western Physiological Services; 1982.
16. Kobelt G, Berg J, Lindgren P, Jönsson B. Costs and quality of life in multiple sclerosis in Europe: method of assessment and analysis. Eur J Health Econ 2006 Sep;7 Suppl 2:S5-13. [doi: 10.1007/s10198-006-0365-y] [Medline: 17310344]
17. Ruutiainen J, Viita AM, Hahl J, Sundell J, Nissinen H. Burden of illness in multiple sclerosis (DEFENSE) study: the costs and quality-of-life of Finnish patients with multiple sclerosis. J Med Econ 2016;19(1):21-33. [doi: 10.3111/13696998.2015.1086362] [Medline: 26360615]

XSL•FO
RenderX

18. EuroQol Group. EuroQol - a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [doi: 10.1016/0168-8510(90)90421-9]
19. Benedict RHB, Munschauer F, Linn R, et al. Screening for multiple sclerosis cognitive impairment using a self-administered 15-item questionnaire. Mult Scler 2003 Feb;9(1):95-101. [doi: 10.1191/1352458503ms861oa] [Medline: 12617275]
20. Hämäläinen P, Leo V, Therman S, Ruutiainen J. Validation of the Finnish version of the brief international cognitive assessment for multiple sclerosis (BICAMS) and evaluation of the applicability of the multiple sclerosis neuropsychological questionnaire (MSNQ) and the fatigue scale for motor and cognitive functions (FSMC). Brain Behav 2021 Jun;11(6):e02087. [doi: 10.1002/brb3.2087] [Medline: 33960700]
21. Penner IK, Raselli C, Stöcklin M, Opwis K, Kappos L, Calabrese P. The fatigue scale for motor and cognitive functions (FSMC): validation of a new instrument to assess multiple sclerosis-related fatigue. Mult Scler 2009 Dec;15(12):1509-1517. [doi: 10.1177/1352458509348519] [Medline: 19995840]
22. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology (ECronicon) 1983 Nov;33(11):1444-1452. [doi: 10.1212/wnl.33.11.1444] [Medline: 6685237]
23. Moore H, Nair KPS, Baster K, Middleton R, Paling D, Sharrack B. Fatigue in multiple sclerosis: a UK MS-register based study. Mult Scler Relat Disord 2022 Aug;64:103954. [doi: 10.1016/j.msard.2022.103954] [Medline: 35716477]
24. Kalb R, Beier M, Benedict RH, et al. Recommendations for cognitive screening and management in multiple sclerosis care. Mult Scler 2018 Nov;24(13):1665-1680. [doi: 10.1177/1352458518803785] [Medline: 30303036]
25. Benedict RH, DeLuca J, Phillips G, et al. Validity of the symbol digit modalities test as a cognition performance outcome measure for multiple sclerosis. Mult Scler 2017 Apr;23(5):721-733. [doi: 10.1177/1352458517690821] [Medline: 28206827]
26. Walker LAS, Osman L, Berard JA, et al. Brief international cognitive assessment for multiple sclerosis (BICAMS): Canadian contribution to the international validation project. J Neurol Sci 2016 Mar 15;362:147-152. [doi: 10.1016/j.jns.2016.01.040] [Medline: 26944137]
27. Filser M, Schreiber H, Pöttgen J, Ullrich S, Lang M, Penner IK. The brief international cognitive assessment in multiple sclerosis (BICAMS): results from the German validation study. J Neurol 2018 Nov;265(11):2587-2593. [doi: 10.1007/s00415-018-9034-1] [Medline: 30171410]
28. Costers L, Gielen J, Eelen PL, et al. Does including the full CVLT-II and BVMT-R improve BICAMS? evidence from a Belgian (Dutch) validation study. Mult Scler Relat Disord 2017 Nov;18:33-40. [doi: 10.1016/j.msard.2017.08.018] [Medline: 29141818]
29. O'Connell K, Langdon D, Tubridy N, Hutchinson M, McGuigan C. A preliminary validation of the brief international cognitive assessment for multiple sclerosis (BICAMS) tool in an Irish population with multiple sclerosis (MS). Mult Scler Relat Disord 2015 Nov;4(6):521-525. [doi: 10.1016/j.msard.2015.07.012] [Medline: 26590658]
30. Chiaravalloti ND, DeLuca J. Cognitive impairment in multiple sclerosis. Lancet Neurol 2008 Dec;7(12):1139-1151. [doi: 10.1016/S1474-4422(08)70259-X] [Medline: 19007738]
31. Podda J, Tacchino A, Ponzio M, et al. Mobile health app (DIGICOG-MS) for self-assessment of cognitive impairment in people with multiple sclerosis: instrument validation and usability study. JMIR Form Res 2024 Jun 20;8:e56074. [doi: 10.2196/56074] [Medline: 38900535]
32. Chen MH, Chiaravalloti ND, DeLuca J. Neurological update: cognitive rehabilitation in multiple sclerosis. J Neurol 2021 Dec;268(12):4908-4914. [doi: 10.1007/s00415-021-10618-2] [Medline: 34028615]
33. Tacchino A, Pedullà L, Bonzano L, et al. A new app for at-home cognitive training: description and pilot testing on patients with multiple sclerosis. JMIR Mhealth Uhealth 2015 Aug 31;3(3):e85. [doi: 10.2196/mhealth.4269] [Medline: 26323749]
34. Lakin L, Davis BE, Binns CC, Currie KM, Rensel MR. Comprehensive approach to management of multiple sclerosis: addressing invisible symptoms-a narrative review. Neurol Ther 2021 Jun;10(1):75-98. [doi: 10.1007/s40120-021-00239-2] [Medline: 33877583]
35. Brichetto G, Zaratin P. Measuring outcomes that matter most to people with multiple sclerosis: the role of patient-reported outcomes. Curr Opin Neurol 2020 Jun;33(3):295-299. [doi: 10.1097/WCO.0000000000000821] [Medline: 32324704]
36. Thomas S, Pulman A, Dogan H, et al. Creating a digital toolkit to reduce fatigue and promote quality of life in multiple sclerosis: participatory design and usability study. JMIR Form Res 2021 Dec 9;5(12):e19230. [doi: 10.2196/19230] [Medline: 34889744]

Abbreviations

EDSS: Expanded Disability Status Scale
FSMC: Fatigue Scale for Motor and Cognitive Functions
FSS: Fatigue Severity Scale
HCP: health care professional
MS: multiple sclerosis
MSNQ: Multiple Sclerosis Neuropsychological Questionnaire
MyMS: patient-interface of the Finnish multiple sclerosis registry
PREDSS: Patient Reported Expanded Disability Status Scale

XSL•FO
RenderX

PRO: patient-reported outcome
PROM: patient-reported outcome measure
QoL: quality of life
SDMT: Symbol Digit Modalities Test
VAS: Visual Analog Scale

Edited by A Kushniruk; submitted 25.10.24; peer-reviewed by J Podda, J Zeng; revised version received 09.05.25; accepted 27.05.25; published 22.08.25.

Please cite as:
Hämäläinen P, Lämsä E, Viitala M, Kuusisto H, Niiranen M, Avikainen S, Puustinen J, Ryytty M, Ruutiainen J, Soilu-Hänninen M
Finnish Registry-Based Protocol for Screening and Management of Fatigue and Cognitive Problems in Multiple Sclerosis: Observational Study
JMIR Hum Factors 2025;12:e67990
URL: https://humanfactors.jmir.org/2025/1/e67990
doi:10.2196/67990

© Päivi Hämäläinen, Elina Lämsä, Matias Viitala, Hanna Kuusisto, Marja Niiranen, Sari Avikainen, Juha Puustinen, Mervi Ryytty, Juhani Ruutiainen, Merja Soilu-Hänninen. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 22.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Health Care Providers' Experiences and Perceptions With Telehealth Tools in a Hospital-at-Home Program: Mixed Methods Study

Shi Yun Low[1], BSc, MPH; Stephanie Qianwen Ko[1,2], MBBS, MMed, MPH; Ian Yi Han Ang[3], BS, MA, PhD

[1]NUHS@Home, National University Health System, Singapore, Singapore

[2]Division of Advanced Internal Medicine, Department of Medicine, National University Hospital, 1E Kent Ridge Road, Singapore, Singapore

[3]Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore

Corresponding Author:
Stephanie Qianwen Ko, MBBS, MMed, MPH
NUHS@Home, National University Health System, Singapore, Singapore

Abstract

Background: The growing demand for hospital-based care, driven by aging populations and constrained resources, has accelerated the adoption of telehealth tools such as teleconsultations and remote monitoring in hospital-at-home (HaH) programs. Despite their increasing use in delivering acute care at home, studies exploring health care providers' experiences and perceptions of these tools within HaH settings remain limited.

Objective: This study aimed to understand the experiences and perspectives of health care providers toward teleconsultations and vital signs monitoring systems within a HaH program in Singapore to optimize effectiveness and address challenges in future implementation.

Methods: A convergent mixed methods approach that combines qualitative in-depth interviews with an electronic survey designed based on the 5 domains (usefulness, ease of use, effectiveness, reliability, and satisfaction) of the Telehealth Usability Questionnaire was used.

Results: In total, 37 surveys and 20 interviews were completed. Participants responded positively to the use of both teleconsultation and vital signs monitoring with a mean total score of each method being 4.55 (SD 0.44) and 4.52 (SD 0.42), respectively. Significantly higher mean ratings were observed among doctors compared with other health care providers for usefulness (P=.03) and ease of use (P=.047) in teleconsultations. Health care providers with fewer years of clinical experience also perceived the use of vital signs monitoring to be more effective (P=.02) and more usable (P=.04) than those with more years of experience. Qualitative analysis identified four themes: (1) benefits of telehealth for health care providers such as improved work convenience, efficiency, and satisfaction; (2) challenges of telehealth implementation relating to communication and technology; (3) perspectives on telehealth impact; and (4) enablers for successful implementation. Comparing both datasets, qualitative findings were aligned with and confirmed quantitative results.

Conclusion: This study highlighted the benefits and usability of telehealth among health care providers. However, challenges relating to patient communication, technological issues, and delivery of care were also discussed along with enablers for successful implementation. These insights can inform strategies to optimize future implementation of telehealth in HaH.

(JMIR Hum Factors 2025;12:e56860) doi:10.2196/56860

KEYWORDS

telehealth usability; hospital-at-home; health care provider experience; virtual consultation; vital signs monitoring; mixed-methods study; health care provider; experience; perception; telehealth tools; telehealth; e-consultations; teleconsultation; hospital-based; home-based; mobile phone

Introduction

As health care systems worldwide face growing challenges from aging populations, the demand for hospital-based care continues to rise, putting increasing pressure on limited health care resources [1,2]. In response, innovative care models such as hospital-at-home (HaH) have been developed to meet these demands [3]. Although the concept of HaH is not novel, its adoption has gained momentum in recent years driven by advancements in digital health technologies, with successful implementations in the United States, Europe, Australia, and Singapore [4,5]. By leveraging these technologies, HaH aims to optimize health care delivery, allowing more patients who require hospital admission to receive treatment at home and supporting the early discharge of patients whose care can safely

XSL•FO
RenderX

continue outside the hospital [6,7]. Likewise, HaH played a critical role in supporting health care systems during the COVID‐19 pandemic [8-11], drawing attention to its potential as a promising care model that extends coverage for pandemic and nonpandemic times.

The integration of digital health tools, such as remote monitoring systems and telecommunication technologies in HaH, allow health care providers to remotely assess patients via video or telephone consultations and monitor self-reported vital signs and symptoms, creating a seamless and efficient care experience [10-12]. Studies have consistently shown that these approaches are both safe and effective, with low mortality and return to hospital admissions rates [10,11,13,14], and high patient satisfaction levels, as patients feel reassured knowing their conditions are remotely monitored [13-17].

Despite these promising outcomes, there remains a lack of research exploring the experiences and perspectives of health care providers using telehealth and remote monitoring tools in the HaH setting. Existing studies that have explored health care provider experiences with telehealth use have primarily focused on routine follow-up in outpatient clinics [18-20], long-term care [21,22], and palliative setting [23] rather than in the HaH setting caring for acutely ill patients. Understanding these experiences can help design appropriate training for health care workers looking to work in HaH settings and develop appropriate support systems to ensure health care workers' well-being, the fourth domain of the Quadruple Aim [24] .

To address this gap, we aimed to understand the experiences and perspectives of health care providers toward teleconsultations and vital signs monitoring systems within an HaH program.

Methods

Study Design

This study adopted a convergent mixed methods design. Quantitative data from an electronic telehealth usability survey and qualitative data from in-depth interviews were collected simultaneously during a single data collection phase. These datasets were then analyzed separately before being interpreted collectively [25]. The mixed methods approach was chosen for its capacity to enhance data interpretation by leveraging the complementary strengths of each method. The quantitative survey data offer statistical evidence and a general overview of health care providers' perceptions regarding the usability of telehealth in HaH. On the other hand, qualitative interviews provide valuable contextual information and deeper insights. By integrating both datasets, a more comprehensive understanding of health care providers' experiences and perspectives can be achieved [26].

Setting and Sample Frame

The National University Health System (NUHS) is one of Singapore's 3 public health care clusters, responsible for taking care of the Western population of Singapore. The HaH program at NUHS, known as NUHS@Home, offers home-based hospitalization services to 3 acute hospitals within the health care cluster [4]. In September 2021, NUHS@Home expanded its services to include patients with COVID-19 to manage the surge in cases. To accommodate the larger volume of patients, patients' daily clinician review transitioned from in-person home visits to teleconsultations (both video and audio) with the care team that consists of doctors, nurses, pharmacists, and allied health professionals.

Patients admitted in NUHS@Home during the study period were primarily older adults with COVID-19 infection. Most were proficient in smartphone use or have younger caregivers comfortable with using technology to manage patient care at home. Patients received daily regular scheduled teleconsultations using their own mobile devices and reported their home-monitored vitals through a chatbot with an online reporting form. For patients less proficient with technology, automated vital sign monitoring kits were provided, which transmitted home-monitored vitals to the care team and included a tablet for teleconsultations. Instructions on remote vital signs monitoring and frequency of teleconsultations were provided to patients or their caregivers before their admission to NUHS@Home.

Purposive sampling was used to select participants for this study. All health care providers who were involved in delivering care within NUHS@Home between September 2021 and December 2022 were sampled. Those that have used teleconsultation, vital signs monitoring, or both telehealth methods to provide care in NUHS@Home, were eligible for the study.

Data Collection

Overview

Participants were recruited via an email invitation sent out by the primary investigator of the study (SQK) between November 2022 and January 2023. Recruitment for the study concluded when no new survey responses were received, and data saturation was achieved for the qualitative component of the study.

Surveys

We chose to adopt the Telehealth Usability Questionnaire (TUQ), which assesses usability across 5 domains, that are usefulness, ease of use, effectiveness, reliability, and satisfaction [27]. The TUQ was selected as it has been widely used to capture user perspectives on telehealth utilization [28]. It has demonstrated strong content validity and internal consistency, with Cronbach α values ranging from 0.81 to 0.93 for each domain [27]. The electronic survey included 21 TUQ questions specific to teleconsultation and the same set of 21 questions related to vital signs monitoring. Demographic information of health care providers, their roles in NUHS@Home, previous telehealth experience, and years of clinical experience were also collected. Survey questions were presented in English, and responses to TUQ questions were rated on a 5-point Likert scale (ranging from strongly disagree [−1] to strongly agree [−5]).

In-Depth Interviews

Participants who completed the electronic survey were also invited to participate in an additional in-depth interview. A semistructured interview guide was developed based on the TUQ domains and existing literature on health care providers'

XSL•FO
RenderX

telehealth experiences [29-32]. The interview guide was pilot tested with 2 other health care providers before the actual interviews. All interviews were conducted online in English via Zoom (Zoom Communications, Inc) video call and audio recorded for analysis. Each interview lasted approximately between 30 and 90 minutes. The interviews were conducted and transcribed by the first author of this paper (SYL), a Master of Public Health graduate student with qualitative training. Given the comprehensive nature of the interview data collected, no repeat interviews were conducted, and transcripts were not returned to participants for clarification.

Data Analysis

Survey data, including participant characteristics and TUQ scores, were analyzed and summarized using descriptive statistics. Percentages of responses for each Likert scale point were calculated for each question. The TUQ total score was obtained by calculating the mean of responses for all 21 questions. TUQ subscale scores were calculated with the mean of the responses for the questions in each of the 5 domains. Means and SDs of TUQ total and subscale scores are presented. As the TUQ total and subscale scores did not follow a normal distribution, the Mann-Whitney *U* test was used to examine differences in scores between professional groups, years of clinical experience, and previous telehealth experience. The significance level was set at α=.05. Statistical analysis was conducted using the R statistical software version 4.2.2 (R Core Team).

All audio-recorded interviews were transcribed verbatim, excluding any local Singaporean slang, and participant identifiers were omitted during the transcription. The transcripts were analyzed using an inductive reflexive thematic approach outlined by Braun and Clarke [33] and Byrne [34]. Transcripts were coded line-by-line using ATLAS.ti Web software v5.16.1-2023-10-14 (Lumivero) by SYL with no second coder, prioritizing the researcher's reflexive and thoughtful engagement with both the data and the analytic process. To ensure the rigor and trustworthiness of emerging subthemes and themes, the meanings of codes were compared and discussed monthly among all authors of this study in an iterative and collaborative process.

Upon completing analysis for both sets of data, qualitative findings were compared with quantitative results by TUQ domains to assess confirmation, expansion, or discordance between the datasets and to draw meta-inferences [35]. Data integration was completed by SYL, and the Good Reporting of a Mixed Methods Study criteria [36] guided the reporting of results to ensure reporting quality (Multimedia Appendix 1).

Ethical Considerations

This study was approved by the National Healthcare Group Domain Specific Review Board (Ref 2021/00896). Since the survey was anonymous, the requirement for informed consent from survey participants was waived. To maintain anonymity for the survey, respondents who were interested in participating in the interviews were asked to voluntarily provide their contact information in a separate, optional section of the survey. This ensured that survey responses remained anonymous and unlinked to identifying information unless the respondent explicitly chose to participate in the interview. Informed consent was only obtained from participants who expressed interest in participating in the in-depth interviews. All data collected and analyzed were anonymized. Each interviewed participant was also compensated for their time and involvement with SGD $20 (SGD $1=US $0.75).

Results

Participant Demographics

A total of 49 health care providers were invited to participate in the study. Among the 49 eligible health care providers, 37 responded and completed the electronic survey, resulting in a response rate of 76%. Out of the 37 survey participants, 20 also took part in the interviews. While 59% (22/37) of the participants were doctors, the remaining participants represented a diverse mix of health care providers of various professions such as nurses, pharmacists, and allied health professionals (Table 1). In total, 76% (28/37) of participants reported using both teleconsultation and vital signs monitoring in NUHS@Home. The participants had a median of 7 (IQR 4‐10) years of clinical experience, but 35% (13/37) of them did not have any previous experience with telehealth.

XSL•FO
RenderX

Table . Characteristics of participants.

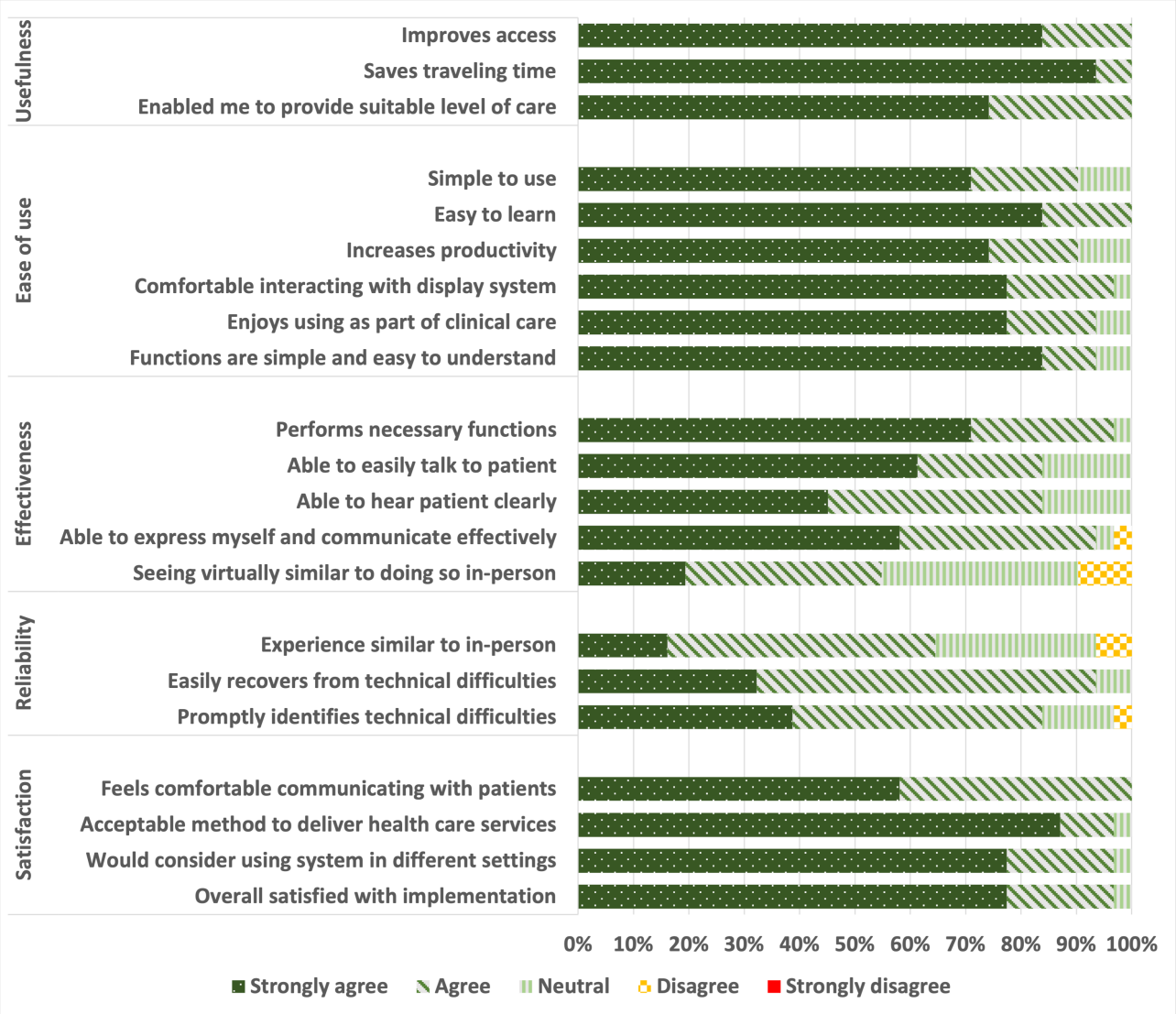
Characteristics		Survey participants (n=37)	Interview participants (n=20)
Age group (years), n (%)			
	20 - 29	9 (24)	4 (20)
	30 - 39	22 (59)	13 (65)
	40 - 49	5 (14)	2 (10)
	50 - 59	1 (3)	1 (5)
Sex, n (%)			
	Male	15 (41)	9 (45)
	Female	22 (59)	11 (55)
Ethnicity, n (%)			
	Chinese	33 (89)	18 (90)
	Malay	1 (3)	1 (5)
	Indian	1 (3)	0 (0)
	Other	2 (5)	1 (5)
Profession, n (%)			
	Doctor	22 (59)	9 (45)
	Nurse	5 (14)	4 (20)
	Allied health professional	7 (19)	5 (25)
	Pharmacist	3 (8)	2 (10)
Telehealth methods used, n (%)			
	Teleconsultation	3 (8)	2 (10)
	Vital signs monitoring	6 (16)	5 (25)
	Both methods	28 (76)	13 (65)
Days per week involved in NUHS@Home, median (IQR)		4 (2-5)	4 (2-5)
Years of clinical experience, median (IQR)		7 (4-10)	8 (5-13)
Previous experience with telehealth, n (%)			
	No experience	13 (35)	6 (30)
	<3 months	4 (11)	3 (15)
	3 - 6 months	4 (11)	1 (5)
	6 - 9 months	5 (14)	2 (10)
	9 - 12 months	3 (8)	0 (0)
	>12 months	8 (22)	8 (40)

Survey Results

In total, 31 responses were collected for the use of teleconsultation in NUHS@Home. The TUQ total score for teleconsultation ranged from 3.4 to 5, with a mean score of 4.55 (SD 0.44). For teleconsultation, the mean TUQ subscale scores for each domain, from highest to lowest scores, were as follows: “usefulness” (mean 4.84, SD 0.30), “satisfaction” (mean 4.73, SD 0.43), “ease of use” (mean 4.72, SD 0.49), “effectiveness” (mean 4.31, SD 0.61), and “reliability” (mean 4.06, SD 0.57) (Figure 1). Furthermore, 34 responses were collected for vital signs monitoring. The TUQ total scores for vital signs monitoring ranged from 3.5 to 5, with a mean score of 4.52 (SD 0.42). For vital signs monitoring, the TUQ subscale scores for “usefulness” (mean 4.8, SD 0.36), “satisfaction” (mean 4.75, SD 0.35), and “ease of use” (mean 4.7, SD 0.43) were high (Figure 2). However, “effectiveness” (mean 4.29, SD 0.59) and “reliability” (mean 3.98, SD 0.87) scored slightly lower. Further breakdown of TUQ scores is available in Multimedia Appendix 2.

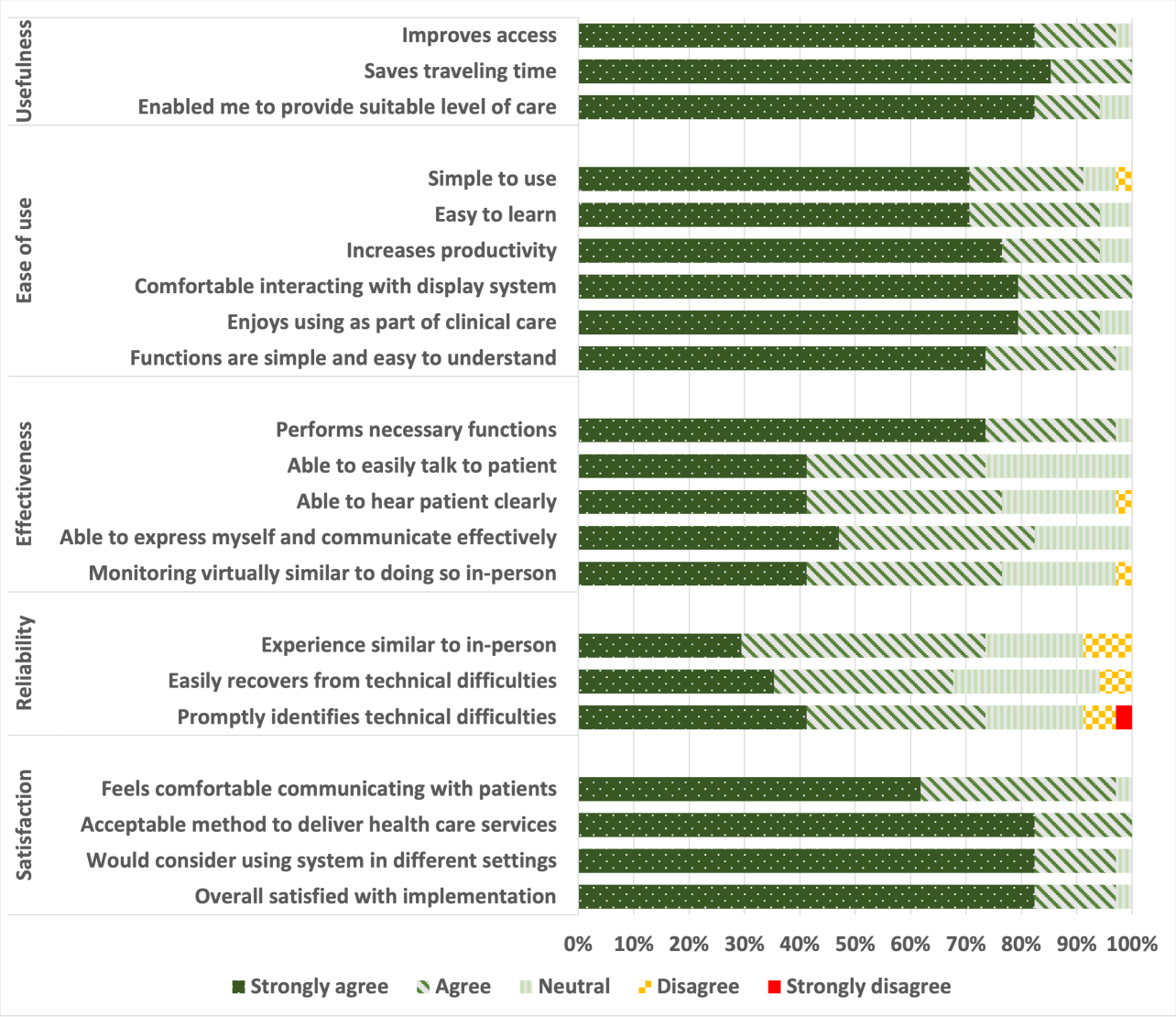
XSL•FO
RenderX

Figure 1. Telehealth Usability Questionnaire responses for teleconsultation.



XSL•FO
RenderX

Figure 2. Telehealth Usability Questionnaire responses for vital signs monitoring.



Doctors perceived teleconsultation to be more useful compared with other health care providers (mean usefulness 4.91 vs 4.67; *P*=.03). Similarly, doctors found teleconsultation easier to use compared with other health care providers (mean ease of use 4.87 vs 4.35; *P*=.047). No significant differences in TUQ subscale scores between doctors and other health care providers were observed for other TUQ domains of teleconsultation or vital signs monitoring. Health care providers with fewer years of clinical experience perceived the use of vital signs monitoring to be more effective (mean effectiveness 4.48 vs 4.04; *P*=.02) and more usable (mean 4.63 vs 4.39; *P*=.04) than those with more years of clinical experience. No significant differences in TUQ scores were observed for teleconsultation or other TUQ domains of vital signs monitoring based on years of clinical experience. Furthermore, no significant differences in TUQ scores were found when comparing health care providers based on previous telehealth experience for both teleconsultation and vital signs monitoring (Multimedia Appendix 3).

Interview Results

Thematic analysis of the interview transcripts yielded 62 unique codes, which were grouped into 4 overarching themes and 8 subthemes. The 4 themes identified are benefits of telehealth tools for health care providers (23 codes and 135 quotes), challenges of telehealth implementation (16 codes and 133 quotes), perspectives on telehealth impact (14 codes and 84 quotes), and enablers for successful implementation (9 codes and 66 quotes). A summary of themes, subthemes, and examples of codes and participant quotes is presented in Table 2.

XSL•FO
RenderX

Table . Summary of participants' themes, subthemes, codes, quotes.

Themes, subthemes, and codes			Examples of quotes
Benefits of telehealth for healthcare providers			
	Improved work convenience, efficiency, and productivity		
		Easy to learn and use vital signs monitoring systems	"…it's a very straightforward system that is very easy to learn to navigate, very easy to use." [ID9, male doctor]
		Reduced need for traveling	"So I think teleconsult has it's good because we actually cut down traveling and received a lot of time in return. The patient is just a phone call or a video call away… You just have to have a set of working equipment in front of you, and you can just sit in front of that and provide care for your patients." [ID15, female nurse]
		Ability to use telehealth methods to effectively identify signs of deterioration	"…through the patients submitting their vitals, actually see and catch any red flags immediately. You know, like persistent high blood pressure, then we know there's something going on. So we can provide treatment immediately." [ID1, female allied health professional]
	Enhanced work satisfaction and confidence		
		Improved teamwork and collaboration	"…because it's all tele, so the team can be physically close together. That means the doctors, the nurses, the pharmacists, all in like one room… I think it's good for like collaboration or when they have issues or when we have questions, it's much easier to just reach them…" [ID5, male pharmacist]
		Facilitated physical distancing and protecting care providers against infectious diseases	"We don't need to gown up and wear all the PPE. Yeah, so in terms of the safety of the staff and with infection control, this is definitely a plus point." [ID20, female pharmacist]
		Improved care provider's confidence in using telehealth	"…I come from a time where physical consultation is extremely important… So at the start, I don't think my confidence was very high to do telemedicine. But after going through [HaH], the interactions with patients have been mostly pleasant… I think it has increased my competence and confidence in managing patients at home using telemedicine." [ID13, male doctor]
Challenges of telehealth implementation			
	Difficulties in patient communication and engagement		

XSL•FO
RenderX

Themes, subthemes, and codes			Examples of quotes
		Difficulty assessing nonverbal information and conducting physical assessments	"You don't get to physically examine your patients. On top of that, because you are not seeing them fully, you can't read their body language, you can't assess a lot of things. So you end up having to talk a lot more to glean more information from a patient to make a clinical decision, as compared to in a ward or in a clinic where you can size up patient pretty fast." [ID9, male doctor]
		Difficulty conveying information to patients	"I guess it's a lot of repetition. Sometimes they don't understand, or they can't hear clearly over video [call]…. Or they can't really operate the video calls very well." [ID11, female doctor]
		Difficulty contacting noncompliant patients and patients at certain time of the day	"You get a few [patients] who just don't want to comply with the rules like they don't want to submit their vitals, they don't want to turn on video because they're sleeping, or they don't want to pick up their phone call." [ID17, female doctor]
	Technological inconveniences related to telehealth systems		
		Care providers encountering technical issues with vital signs monitoring systems	"Audio call, sometimes… including the video calls as well, sometimes it can be quite choppy. We then need to make the calls multiple times. So it tends to get a bit frustrating also." [ID7, female nurse]
		Slow and tedious process to recover and manage technical issues	"Actually, we have feedback to the [vital signs monitoring system vendors] team. But somehow the team also can't quite troubleshoot it when there's problem." [ID15, female nurse]
		Inconveniences encountered due to lack of interoperability of information technology platforms	"…you need to ensure you enter the patient's details correctly. Because [vital signs monitoring system] itself, the patient details is not like an auto flow from [electronic medical records]. So if you enter wrongly or when you create the entry wrongly, then the data error sort of carries forward." [ID3, male doctor]
	Operational issues		
		Increased risk of compromising patient privacy during virtual consults or unconducive environment to conduct virtual calls	"…it got a bit noisy, which is not that good and also can potentially compromise confidentiality because the guy next to you might be talking very loudly, asking for the NRIC number of his or her patients, and your patient might potentially hear the whole conversation." [ID16, male doctor]

XSL•FO
RenderX

Themes, subthemes, and codes			Examples of quotes
		Labor intensive for care providers to set up automated vital signs monitoring system	"The use of the automated kind of system where the vitals can automatically transmit into the system… involves a bit more resource setup, because someone needs to go to the patient's house and set up the system there." [ID14, male doctor]
Perspectives on telehealth impact			
	Perceived patient benefits		
		Perceived improved patient experience due to increased convenience and comfort	"…a shorter hospital stay, convalescing in a very comfortable home environment. So in terms of the experience for [the patients], it's a lot better." [ID12, male doctor]
		Improves patient empowerment	"[Patients] will have more self-awareness of their health. I think they will be more in control of their health when they monitor the blood pressure themselves." [ID6, female nurse]
	Potential benefits to health care systems		
		Reducing need for hospital admissions and saving hospital resources	"…during COVID, one of the big issues was a lack of beds. So [telehealth] definitely helps to free up a lot of beds, especially for patients who are technically stable, who didn't really need to be in [the hospital]." [ID17, female doctor]
		Increased capacity to care for more patients	"We can manage a large group of patients at once. We don't need to visit all of them physically." [ID10, male doctor]
	Potential risks and challenges		
		Concerns with accuracy of patient vitals submission	"Technically, it's very difficult to find out if someone is actually, you know like faking the vitals reading… I think the readings that are taken at home will have some degree of lower reliability compared to in the hospital because the operators are the patients themselves and the caregivers rather than the nurses or the healthcare assistants." [ID14, male doctor]
		Concerns of "overmonitoring"	"So we get patients who are very tired of being chased after to measure [vital signs] three times a day because they wouldn't have otherwise be required to monitor their health at all… Also because of the timing that we would like to see the vital signs, we also ask them to do it at 7am, which is not their usual wake up time." [ID7, female nurse]
Enablers to successful implementation			

XSL•FO
RenderX

Themes, subthemes, and codes		Examples of quotes
	Suitability of using telehealth to manage patients with COVID-19	"…for COVID in particular, majority of the patients are fairly stable… there's no reason to believe that they might deteriorate in front of your eyes. And I honestly tell you, medicine and monitoring are probably the preferred and sensible route… So honestly patients just need to go home…and can be monitored at home safely." [ID12, male doctor]
	Importance of care provider-patient trust	"I think trust is implicit, we just have to trust that the numbers they are submitting is truly what they're doing." [ID13, male doctor]

Integrating Survey and Interview Results

Quantitative and qualitative findings were summarized and meta-inferences were drawn in Table 3. Overall, high levels of usefulness were attributed to perceptions of increased convenience, efficiency, and usefulness. High satisfaction rates were attributed to a positive work environment, teamwork and collaboration, protection from infectious diseases, and high levels of acceptance described by participants. Systems were generally found to be easy to learn and use; however, lack of interoperability with electronic health records and noncompliant patients were cited as key challenges. The comparatively lower scores for effectiveness were likely associated with challenges relating to patient communication and engagement. Finally, the lower scores in the "reliability" domain were associated with technical inconveniences, such as connectivity issues and having a lengthy turnaround time to resolve technical issues.

XSL•FO
RenderX

Table . Integration of findings and the mixed methods meta-inference.

TUQ[a] domain	Quantitative results	Qualitative results	Meta-inference
Usefulness	Mean TUQ subscale scores of 4.84 and 4.8 for the use of both teleconsultation and vital signs monitoring respectively. 100% and 98% of participants rated positively[b] to use of virtual consultation and vital signs monitoring, respectively.	Benefits of telehealth: • Improved convenience due to reduced need for traveling • Increased flexibility and ease of reaching out to their patients • Efficient identification of early signs of deterioration Perceptions: • Perceived patient benefits (eg, patient empowerment, patient comfort, and convenience) • Potential benefits to health care system (eg, saving hospital resources and increased capacity to care for more patients) Enablers: • Suitability of using telehealth to care for patients with COVID-19	• Confirmation: participants' positive experiences relating to improved convenience and efficiency, as well as their perception of the suitability of using telehealth to provide care for patients with COVID-19, align with the high scores observed in the "usefulness" domain. • Expansion: participants also shared on their perceived benefits of using telehealth to both patients and the health care system, thus expanding on the concept of usefulness of telehealth not just for health care providers but also for patients and health care system.
Satisfaction	Mean TUQ subscale scores of 4.73 and 4.75 for the use of both teleconsultation and vital signs monitoring, respectively. 98% and 97% of participants rated positively[b] to use of virtual consultation and vital signs monitoring, respectively.	Benefits of telehealth: • More opportunity for collaboration and sharing of expertise • Improvement in work conditions • Enjoyable virtual interactions • Increased confidence and acceptance of telehealth	• Confirmation: improvements in work conditions, protection from infectious diseases, improved teamwork and collaboration, and increased acceptance of telehealth described by participants were able to explain and confirm the relatively high scores rated for the "satisfaction" domain.
Ease of Use	Mean TUQ subscale scores of 4.72 and 4.70 for the use of both teleconsultation and vital signs monitoring, respectively. 94% and 95% of participants rated positively[b] to use of virtual consultation and vital signs monitoring, respectively.	Benefits of telehealth: • Easy to learn and use vital signs monitoring systems Challenges of telehealth: • Difficulty dealing noncompliant patients • Inconveniences due to lack of interoperability between systems • Labor intensive to set up automated systems Enablers: • Using familiar platforms for virtual consultation	• Confirmation: the majority of participants reported that the vital signs monitoring systems were easy to learn and use, with the utilization of familiar platforms for teleconsultations as an enabler. These findings corroborate the high scores observed in the "ease of use" domain. • Expansion: a handful of the negative experiences encountered by the participants relating to inconveniences caused by the lack of interoperability between systems and setting up of automated systems, and difficulty dealing with noncompliant patients, could explain the lower percentage of participants who rated positively for the "ease of use" domain in telehealth use.

XSL•FO
RenderX

TUQ[a] domain	Quantitative results	Qualitative results	Meta-inference
Effectiveness	Mean TUQ sub-scale scores of 4.31 and 4.29 for the use of both virtual consultation and vital signs monitoring, respectively. 83% and 81% of participants rated positively[b] to use of teleconsultation and vital signs monitoring, respectively.	Challenges of telehealth: • Inability to assess nonverbal information • Difficulty communicating with older adult patients • Difficulty engaging and establishing rapport • Unconducive environment Perceptions: • Accuracy of patient vital signs submissions • Fear of making wrong clinical decisions	• Confirmation: the negative experiences reported by participants, specifically regarding challenges in patient communication and engagement, align with, and provide an explanation for the lower scores observed in the "effectiveness" domain.
Reliability	Mean TUQ subscale scores of 4.06 and 3.98 for the use of both teleconsultation and vital signs monitoring, respectively. 81% and 72% of participants rated positively[b] to use of teleconsultation and vital signs monitoring, respectively.	Challenges of telehealth: • Technical issues with vital signs monitoring system • Connectivity issues during virtual consultations • Process of resolving such issues was often slow and cumbersome	• Confirmation: recounts of negative experiences associated with the technical inconveniences related to telehealth systems, align with and confirm the lower scores observed in the "reliability" domain.

[a]Telehealth Usability Questionnaire.

[b]Rating either "strongly agree" or "agree."

Discussion

Principal Findings and Comparison With Previous Work

In this mixed methods study exploring health care providers' engagement in teleconsultations and vital signs monitoring for acute care at home, "usefulness," "satisfaction," and "ease of use" received the highest ratings, while "effectiveness" and "reliability" received lower ratings. These findings were corroborated during interviews, where overall positive perceptions were highlighted alongside key challenges related to patient communication and engagement, and technological issues.

Our study contributes both fresh insights and confirmation of existing literature on health care providers' experiences with telehealth. Consistent with previous research [37], participants emphasized the convenience and flexibility of teleconsultations and vital signs monitoring, particularly in optimizing health care resources during the COVID-19 pandemic. This marks a shift from previous reports favoring telehealth use for chronic diseases [38] or palliative care [39] to acute care settings, possibly due to the unique demands of the pandemic. Challenges of managing noncompliant patients in HaH setting differed from stable chronic conditions in outpatient virtual settings, due to the acuity and risk of deterioration of these patients [40]. Concerns regarding vital signs accuracy and making clinical judgments using telehealth tools were also novel observations, possibly attributed to the adjustment from face-to-face consultations to telehealth adoption, especially among more experienced health care providers. Another noteworthy finding is that telehealth use in a multidisciplinary HaH setting brought health care providers closer together, fostering collaboration and shared learning opportunities.

On the other hand, participants who stressed the importance of interoperability between telehealth systems and electronic medical records were aligned with existing literature advocating for seamless integration and data exchange for successful telehealth implementation [41]. Reports of communication difficulties and patient engagement issues also mirror previous findings, underscoring the need for specific communication skills development among health care providers [42]. In addition, our study reaffirmed the impact of technological issues on telehealth reliability [43], emphasizing the need for robust technical support to ensure system stability, though participants mainly struggled with the slow resolution of problems rather than frequent connectivity or technical difficulties. Finally, participants' high satisfaction with telehealth use echoes previous research [37,39,44], highlighting its acceptance and benefits across diverse care settings.

Health care providers' real-world experiences offer vital insights that can help policy makers understand practical implications and identify policy gaps. Considerations for the real-world implementation and context can aid in crafting practical telehealth and HaH guidelines that reflect the realities of these settings. Health care systems can also learn from these experiences to create adaptable telehealth-ready systems to meet the needs of an aging population and improve health care accessibility, efficiency, and quality. Based on our findings, we recommend 3 key actions. First, strengthen health care provider training in developing skills for telecommunication, remote assessment, and decision-making, along with early exposure to digital health care tools. Second, ensure the reliability of telehealth systems through regular checks, updates, and robust technical support, and additionally integrate systems with electronic medical records to ensure efficient data exchange. Finally, promote multidisciplinary collaboration by designing

XSL•FO
RenderX

physical spaces that facilitate effective communication and allocate private areas for teleconsultations to ensure patient confidentiality.

Limitations

Our study has several limitations. First, the health care providers sampled were from a single HaH program, and their perspectives may not fully represent those of providers in other programs that use different telehealth tools or operate in varying contexts. Nonetheless, we included a diverse range of health care professions in both the survey and interviews to ensure a broad representation of experiences.

Second, the sample size in this study may be considered small, potentially limiting the statistical power of the quantitative analysis. However, the integration of qualitative data through a mixed methods approach mitigates this limitation by providing a more comprehensive and nuanced understanding of trends observed in the quantitative findings.

Third, there is a potential for recall bias among participants who were no longer actively staffing the service during the data collection period. In addition, the qualitative data analysis was conducted by a single coder and not independently in duplicate as per best practice. While this approach may introduce bias, the reflexive thematic approach used during analysis emphasizes the researcher's interpretative engagement with the data over consensus or reliability measures [34]. This enables a richer, more cohesive analysis that upholds reflexivity and interpretative rigor. Nevertheless, future studies could enhance richer and deeper interpretations of the data by incorporating multiple coders in a collaborative and reflexive manner.

Furthermore, the clinical context during the data collection period was exclusively shaped by COVID-19, which presents unique challenges and care requirements compared with nonpandemic conditions typically managed in HaH programs. Despite this, our findings offer valuable insights into the use of telehealth tools, which remain relevant as HaH programs continue to expand and scale.

Finally, while the TUQ has demonstrated strong validity and reliability for assessing telehealth usability, it was originally developed for general telehealth users rather than providers. Nonetheless, we selected the TUQ for its comprehensiveness in evaluating core usability dimensions that are also relevant to health care providers. The comprehensive triangulation of data underscores the novelty of our study, as it demonstrates the adaptability of the TUQ beyond traditional end user settings and highlights its effectiveness in capturing the multidimensional experiences of health care providers engaging with telehealth within HaH.

Conclusions

In summary, this study suggests that this generation of health care providers, though not considered to be formally trained in the use of telehealth tools, were generally adaptable and held positive perceptions regarding the usability of telehealth. Our findings further emphasize the equal importance of the health care providers as key stakeholders and end users (in addition to patients) in the development of telehealth tools. As HaH programs continue to scale internationally, further studies are needed to design and evaluate telehealth training programs for HaH healthcare providers and explore the influence of physical workspace design on health care provider effectiveness when using telehealth tools.

Acknowledgments

Generative artificial intelligence was not employed to develop ideas or informational content for this manuscript. However, ChatGPT was used in the proofreading process to improve the clarity of the writing. The authors would like to acknowledge the doctors, nurses, allied health professionals, pharmacists, and operations staff who have contributed to the development and implementation of the NUHS@Home program. This work was supported by the Health Services Research Grant from the National Medical Research Council HSRGMS20nov-0004 and the Practicum Fund from the National University of Singapore, Saw Swee Hock School of Public Health, Master of Public Health program.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to institutional data sharing policies. However, they may be obtained from the corresponding author on reasonable request.

Authors' Contributions

This study was conceived by SQK. All authors contributed to the design of the study. SYL and SQK were involved in the data collection process. All authors participated in the analysis of data collected. SYL drafted the manuscript while SQK and IYHA edited the manuscript. All authors reviewed the final manuscript.

Conflicts of Interest

SQK is the lead of NUHS@Home on which the study setting is based. This relationship however did not influence the study design, data collection, analysis, or interpretation of the findings.

Multimedia Appendix 1

XSL•FO
RenderX

Good reporting of a mixed methods study criteria.
[DOCX File, 15 KB - humanfactors_v12i1e56860_app1.docx]

Multimedia Appendix 2
Detailed breakdown of Telehealth Usability Questionnaire scores.
[DOCX File, 22 KB - humanfactors_v12i1e56860_app2.docx]

Multimedia Appendix 3
Summary of subgroup analyses.
[DOCX File, 19 KB - humanfactors_v12i1e56860_app3.docx]

References

1. Ageing and health. World Health Organization. 2022. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2025-04-11]
2. Van den Heede K, Bouckaert N, Van de Voorde C. The impact of an ageing population on the required hospital capacity: results from forecast analysis on administrative data. Eur Geriatr Med 2019 Oct;10(5):697-705. [doi: 10.1007/s41999-019-00219-8] [Medline: 34652701]
3. Leff B. Defining and disseminating the hospital-at-home model. CMAJ 2009 Jan 20;180(2):156-157. [doi: 10.1503/cmaj.081891] [Medline: 19153385]
4. Ko SQ, Goh J, Tay YK, et al. Treating acutely ill patients at home: data from Singapore. Ann Acad Med Singap 2022 Jul;51(7):392-399. [doi: 10.47102/annals-acadmedsg.2021465] [Medline: 35906938]
5. Leong MQ, Lim CW, Lai YF. Comparison of hospital-at-home models: a systematic review of reviews. BMJ Open 2021 Jan 29;11(1):e043285. [doi: 10.1136/bmjopen-2020-043285] [Medline: 33514582]
6. Gonçalves-Bradley DC, Iliffe S, Doll HA, et al. Early discharge hospital at home. Cochrane Database Syst Rev 2017 Jun 26;6(6):CD000356. [doi: 10.1002/14651858.CD000356.pub4] [Medline: 28651296]
7. Shepperd S, Iliffe S, Doll HA, et al. Admission avoidance hospital at home. Cochrane Database Syst Rev 2016 Sep 1;9(9):CD007491. [doi: 10.1002/14651858.CD007491.pub2] [Medline: 27583824]
8. Schiff R, Oyston M, Quinn M, Walters S, McEnhill P, Collins M. Hospital at Home: another piece of the armoury against COVID-19. Future Healthc J 2022 Mar;9(1):90-95. [doi: 10.7861/fhj.2021-0137] [Medline: 35372768]
9. Nogués X, Sánchez-Martinez F, Castells X, et al. Hospital-at-home expands hospital capacity during COVID-19 pandemic. J Am Med Dir Assoc 2021 May;22(5):939-942. [doi: 10.1016/j.jamda.2021.01.077] [Medline: 33639115]
10. Ko SQ, Kumar SK, Jacob J, et al. Technology-enabled virtual ward for COVID management of the elderly and immunocompromised in Singapore: a descriptive cohort. BMC Infect Dis 2023 Feb 21;23(1):102. [doi: 10.1186/s12879-023-08040-2] [Medline: 36809977]
11. Tan MWJ, Arciga MGA, Arba'in JB, et al. Outcomes of a "hospital at home" programme for the supervised home recovery of COVID-19 patients in Singapore. Proc Singap Healthc 2023 Dec;32:20101058231152049. [doi: 10.1177/20101058231152049]
12. Levine DM, Paz M, Burke K, et al. Remote vs in-home physician visits for hospital-level care at home: a randomized clinical trial. JAMA Netw Open 2022 Aug 1;5(8):e2229067. [doi: 10.1001/jamanetworkopen.2022.29067] [Medline: 36040741]
13. Chung-Lee L, Catallo C. A new approach to digital health? Virtual COVID-19 care: a scoping review. Digit Health 2023;9:20552076231152171. [doi: 10.1177/20552076231152171] [Medline: 36798886]
14. Vindrola-Padros C, Singh KE, Sidhu MS, et al. Remote home monitoring (virtual wards) for confirmed or suspected COVID-19 patients: a rapid systematic review. EClinicalMedicine 2021 Jul;37:100965. [doi: 10.1016/j.eclinm.2021.100965] [Medline: 34179736]
15. Raffan F, Anderson T, Sinclair T, et al. The virtual care experience of patients diagnosed with COVID-19. J Patient Exp 2021;8:23743735211008310. [doi: 10.1177/23743735211008310] [Medline: 34179437]
16. Ko SQ, Chua CMS, Koh SH, Lim YW, Shorey S. Experiences of patients and their caregivers admitted to a hospital-at-home program in Singapore: a descriptive qualitative study. J Gen Intern Med 2023 Feb;38(3):691-698. [doi: 10.1007/s11606-022-07765-1] [Medline: 36008593]
17. Shaw A, Moodley M, McSporran K, et al. P18 Safety and effectiveness of an integrated, telehealth-led supported discharge service for Covid-19. Thorax 2021 Nov:A76-A7A. [doi: 10.1136/thorax-2021-BTSabstracts.128]
18. Banks J, Corrigan D, Grogan R, et al. LoVE in a time of CoVID: Clinician and patient experience using telemedicine for chronic epilepsy management. Epilepsy Behav 2021 Feb;115:107675. [doi: 10.1016/j.yebeh.2020.107675] [Medline: 33342712]
19. Gold KJ, Laurie AR, Kinney DR, Harmes KM, Serlin DC. Video visits: family physician experiences with uptake during the COVID-19 pandemic. Fam Med 2021 Mar;53(3):207-210. [doi: 10.22454/FamMed.2021.613099] [Medline: 33723819]

XSL•FO
RenderX

20. Björndell C, Premberg Å. Physicians' experiences of video consultation with patients at a public virtual primary care clinic: a qualitative interview study. Scand J Prim Health Care 2021 Mar;39(1):67-76. [doi: 10.1080/02813432.2021.1882082] [Medline: 33650941]
21. Tan AJ, Rusli KD, McKenna L, Tan LL, Liaw SY. Telemedicine experiences and perspectives of healthcare providers in long-term care: a scoping review. J Telemed Telecare 2024 Feb;30(2):230-249. [doi: 10.1177/1357633X211049206] [Medline: 34666535]
22. MacNeill V, Sanders C, Fitzpatrick R, et al. Experiences of front-line health professionals in the delivery of telehealth: a qualitative study. Br J Gen Pract 2014 Jul;64(624):e401-e407. [doi: 10.3399/bjgp14X680485] [Medline: 24982492]
23. Lundereng ED, Nes AAG, Holmen H, et al. Health care professionals' experiences and perspectives on using telehealth for home-based palliative care: scoping review. J Med Internet Res 2023 Mar 29;25:e43429. [doi: 10.2196/43429] [Medline: 36989024]
24. Bodenheimer T, Sinsky C. From triple to quadruple aim: care of the patient requires care of the provider. Ann Fam Med 2014;12(6):573-576. [doi: 10.1370/afm.1713] [Medline: 25384822]
25. Creswell JW, Clark VLP. Designing and Conducting Mixed Methods Research: Sage Publications; 2007.
26. Wasti SP, Simkhada P, van Teijlingen ER, Sathian B, Banerjee I. The growing importance of mixed-methods research in health. Nepal J Epidemiol 2022 Mar;12(1):1175-1178. [doi: 10.3126/nje.v12i1.43633] [Medline: 35528457]
27. Parmanto B, Lewis AN, Graham KM, Bertolet MH. Development of the Telehealth Usability Questionnaire (TUQ). Int J Telerehabil 2016;8(1):3-10. [doi: 10.5195/ijt.2016.6196] [Medline: 27563386]
28. Hajesmaeel-Gohari S, Bahaadinbeigy K. The most used questionnaires for evaluating telemedicine services. BMC Med Inform Decis Mak 2021 Feb 2;21(1):36. [doi: 10.1186/s12911-021-01407-y] [Medline: 33531013]
29. Pogorzelska K, Marcinowicz L, Chlabicz S. A qualitative study of primary care physicians' experiences with telemedicine during the COVID-19 pandemic in North-Eastern Poland. Int J Environ Res Public Health 2023 Jan 20;20(3):1963. [doi: 10.3390/ijerph20031963] [Medline: 36767329]
30. White J, Byles J, Walley T. The qualitative experience of telehealth access and clinical encounters in Australian healthcare during COVID-19: implications for policy. Health Res Policy Syst 2022 Jan 15;20(1):9. [doi: 10.1186/s12961-021-00812-z] [Medline: 35033107]
31. Turner K, Bobonis Babilonia M, Naso C, et al. Health care providers' and professionals' experiences with telehealth oncology implementation during the COVID-19 pandemic: a qualitative study. J Med Internet Res 2022 Jan 19;24(1):e29635. [doi: 10.2196/29635] [Medline: 34907900]
32. Davoodi NM, Chen K, Zou M, et al. Emergency physician perspectives on using telehealth with older adults during COVID-19: a qualitative study. J Am Coll Emerg Physicians Open 2021 Oct;2(5):e12577. [doi: 10.1002/emp2.12577] [Medline: 34693401]
33. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
34. Byrne D. A worked example of Braun and Clarke's approach to reflexive thematic analysis. Qual Quant 2022 Jun;56(3):1391-1412. [doi: 10.1007/s11135-021-01182-y]
35. Fetters MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs-principles and practices. Health Serv Res 2013 Dec;48(6 Pt 2):2134-2156. [doi: 10.1111/1475-6773.12117] [Medline: 24279835]
36. O'Cathain A, Murphy E, Nicholl J. The quality of mixed methods studies in health services research. J Health Serv Res Policy 2008 Apr;13(2):92-98. [doi: 10.1258/jhsrp.2007.007074] [Medline: 18416914]
37. Lampickienė I, Davoody N. Healthcare professionals' experience of performing digital care visits-a scoping review. Life (Basel) 2022 Jun 17;12(6):913. [doi: 10.3390/life12060913] [Medline: 35743944]
38. Johnsen TM, Norberg BL, Kristiansen E, et al. Suitability of video consultations during the COVID-19 pandemic lockdown: cross-sectional survey among Norwegian general practitioners. J Med Internet Res 2021 Feb 8;23(2):e26433. [doi: 10.2196/26433] [Medline: 33465037]
39. Zhang H, Cha EE, Lynch K, et al. Radiation oncologist perceptions of telemedicine from consultation to treatment planning: a mixed-methods study. Int J Radiat Oncol Biol Phys 2020 Oct 1;108(2):421-429. [doi: 10.1016/j.ijrobp.2020.07.007] [Medline: 32890525]
40. Hinman RS, Nelligan RK, Bennell KL, Delany C. "Sounds A Bit Crazy, But It Was Almost More Personal:" a qualitative study of patient and clinician experiences of physical therapist-prescribed exercise for knee osteoarthritis via Skype. Arthritis Care Res (Hoboken) 2017 Dec;69(12):1834-1844. [doi: 10.1002/acr.23218] [Medline: 28217864]
41. Thomas EE, Haydon HM, Mehrotra A, et al. Building on the momentum: sustaining telehealth beyond COVID-19. J Telemed Telecare 2022 May;28(4):301-308. [doi: 10.1177/1357633X20960638] [Medline: 32985380]
42. Jarva E, Oikarinen A, Andersson J, et al. Healthcare professionals' perceptions of digital health competence: a qualitative descriptive study. Nurs Open 2022 Mar;9(2):1379-1393. [doi: 10.1002/nop2.1184] [Medline: 35094493]
43. Ftouni R, AlJardali B, Hamdanieh M, Ftouni L, Salem N. Challenges of telemedicine during the COVID-19 pandemic: a systematic review. BMC Med Inform Decis Mak 2022 Aug 3;22(1):207. [doi: 10.1186/s12911-022-01952-0] [Medline: 35922817]

XSL•FO
RenderX

44. Saiyed S, Nguyen A, Singh R. Physician perspective and key satisfaction indicators with rapid telehealth adoption during the coronavirus disease 2019 pandemic. Telemed J E Health 2021 Nov;27(11):1225-1234. [doi: 10.1089/tmj.2020.0492] [Medline: 33513045]

Abbreviations

HaH: hospital-at-home
NUHS: National University Health System
TUQ: Telehealth Usability Questionnaire

Edited by A Kushniruk; submitted 28.01.24; peer-reviewed by E Jarva, G Lee, H Alshawaf; revised version received 06.03.25; accepted 07.03.25; published 17.04.25.

Please cite as:
Low SY, Ko SQ, Ang IYH
Health Care Providers' Experiences and Perceptions With Telehealth Tools in a Hospital-at-Home Program: Mixed Methods Study
JMIR Hum Factors 2025;12:e56860
URL: https://humanfactors.jmir.org/2025/1/e56860
doi:10.2196/56860

© Shi Yun Low, Stephanie Q Ko, Ian Yi Han Ang. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 17.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Assessing the Impact on Electronic Health Record Burden After Five Years of Physician Engagement in a Canadian Mental Health Organization: Mixed-Methods Study

Tania Tajirian[1,2]; Brian Lo[2,3,4]; Gillian Strudwick[2,3]; Adam Tasca[2,5]; Emily Kendell[1,2]; Brittany Poynter[2,5]; Sanjeev Kumar[2,5]; Po-Yen (Brian) Chang[1,2]; Candice Kung[2,5]; Debbie Schachter[2,5]; Gwyneth Zai[2,5]; Michael Kiang[2,5]; Tamara Hoppe[2,5]; Sara Ling[2,5,6]; Uzma Haider[2]; Kavini Rabel[2]; Noelle Coombe[2]; Damian Jankowicz[2,3,4]; Sanjeev Sockalingam[2,5]

[1]Department of Family and Community Medicine, University of Toronto, Toronto, ON, Canada

[2]Centre for Addiction and Mental Health, Office 6168G, 100 Stokes Street, Toronto, ON, Canada

[3]Institute for Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada

[4]Information Technology, Unity Health Toronto, Toronto, ON, Canada

[5]Department of Psychiatry, University of Toronto, Toronto, ON, Canada

[6]Lawrence Bloomberg Faculty of Nursing, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Tania Tajirian
Department of Family and Community Medicine, University of Toronto, Toronto, ON, Canada

Abstract

Background: The burden caused by the use of electronic health record (EHR) systems continues to be an important issue for health care organizations, especially given human resource shortages in health care systems globally. As physicians report spending 2 hours documenting for every hour of patient care, there has been strong interest from many organizations to understand and address the root causes of physician burnout due to EHR burden.

Objective: This study focuses on evaluating physician burnout related to EHR usage and the impact of a physician engagement strategy at a Canadian mental health organization 5 years after implementation.

Methods: A cross-sectional survey was conducted to assess the perceived impact of the physician engagement strategy on burnout associated with EHR use. Physicians were invited to participate in a web-based survey that included the Mini-Z Burnout questionnaire, along with questions about their perceptions of the EHR and the effectiveness of the initiatives within the physician engagement strategy. Descriptive statistics were applied to analyze the quantitative data, while thematic analysis was used for the qualitative data.

Results: Of the 254 physicians invited, 128 completed the survey, resulting in a 50% response rate. Among the respondents, 26% (33/128) met the criteria for burnout according to the Mini-Z questionnaire, with 61% (20/33) of these attributing their burnout to EHR use. About 52% of participants indicated that the EHR improves communication (67/128) and 38% agreed that the EHR enables high-quality care (49/128). Regarding the physician engagement strategy initiatives, 39% (50/128) agreed that communication through the strategy is efficient, and 75% (96/128) felt more proficient in using the EHR. However, additional areas for improvement within the EHR were identified, including (1) medication reconciliation and prescription processes; (2) chart navigation and information retrieval; (3) longitudinal medication history; and (4) technology infrastructure challenges.

Conclusions: This study highlights the potential impact of EHRs on physician burnout and the effectiveness of a unique physician engagement strategy in fostering positive perceptions and improving EHR usability among physicians. By evaluating this initiative in a real-world setting, the study contributes to the broader literature on strategies aimed at enhancing physician experience following large-scale EHR implementation. However, the findings indicate a continued need for system-level improvements to maximize the value and usage of EHRs. The physician engagement strategy demonstrates the potential to enhance physicians' EHR experience. Future efforts should prioritize system-level advancements to increase the EHR's impact on quality of care and develop standardized approaches for engaging physicians on a broader Canadian scale.

(JMIR Hum Factors 2025;12:e65656) doi:10.2196/65656

KEYWORDS

electronic health records; health informatics; documentation burden; physician burnout; engagement strategy; evaluation; EHR; burden; physician; cross-sectional survey

XSL•FO
RenderX

Introduction

Background

The burden caused by the use of electronic health record (EHR) systems continues to be an important issue for health care organizations, especially given human resource shortages in health care systems globally [1-4]. As physicians report spending 2 hours documenting for every hour of patient care, there has been strong interest from many organizations to understand and address the root causes of physician burnout due to EHR burden [5]. This is particularly important given the growing digital maturity of many health systems coupled with broader health care challenges, specifically an overburdened health system and a health care workforce crisis [2-4]. Over the last few years, many organizations, including the American Medical Informatics Association [6], the Canadian Medical Association [2-4], and the American Medical Association [7], have recognized and advocated for strategies to target these challenges .

As a result, there have been many initiatives focused on mitigating the impacts of EHR burden on physicians. For example, the 25 × 5 initiative, led by American Medical Informatics Association and including a task force representing health professionals or systems, policy or advocacy, impact, and technology requirements, is focused on developing a multi-level approach (vendor, health systems, and government) to reducing documentation burden by 25% in 5 years in the United States [6,8]. Through these work streams, a toolkit has been developed that provides recommendations and next steps at the organization and system levels [8]. In addition, others have explored unique initiatives to identify and address EHR-related burden within their own organizations [9]. An example of this is the Getting Rid of Stupid Stuff initiative from the University of Colorado, which focuses on removing redundant processes from the EHR [10]. Other initiatives include more agile delivery of personalized training and system optimization targeting the pain points of the end-users [11-13].

However, a key challenge that has impacted the ability to develop strong recommendations on approaches to address EHR-related issues has been limited insight into the impact of these initiatives [9,14,15]. In particular, most of these evaluations have been limited to a single-site pilot with a limited duration and small participant size. Moreover, since these studies examined the intervention immediately after implementation, there is little information about the long-term impact of these initiatives on EHR end-user experience [16]. To better understand the impact of initiatives to reduce EHR burden, it is essential to evaluate them in large-scale real-world environments with numerous users, as studies of digital tools indicate that effectiveness must be tested across diverse settings. Overall, this study focuses on evaluating physician burnout related to EHR usage and the impact of a physician engagement strategy at a Canadian mental health organization 5 years post-implementation.

Physician Engagement Strategy

The physician engagement strategy [17] was developed in 2018 through feedback from an organizational benchmark survey, which was conducted to understand the contributors of EHR-related burden and burnout rate among physicians across a large Canadian academic hospital [18]. In the benchmark survey, full-time and part-time physicians at the organization were invited to provide their perceptions and concerns in order to co-develop the first iteration of the physician engagement survey. Briefly, from the benchmark survey, about 75% (155/208) of physicians with symptoms of burnout reported that the EHR was a major contributor to burnout. In addition, participants highlighted process and infrastructure issues that can be improved through an organization-wide strategy.

As such, the physician engagement strategy was created and implemented and consists of 4 key components.

The "SWAT" Initiative

One of the main challenges has been an approach to rapidly identify and address bottlenecks and issues related to the EHR. As a result, we developed the "SWAT" initiative, which focuses on bringing an interdisciplinary team to rapidly triage and address issues related to the EHR in an agile manner [19].

Physician Think Tanks

Given the ongoing challenges of ensuring that decisions made for the EHR are aligned with the needs of end-users, a medical staff subcommittee was convened to engage in discussions on EHR and digital health matters. These think tanks meet on a monthly basis, and a representative from each medical division attends to provide feedback and discuss relevant issues.

Communication, Education, and Informatics Strategies

In order to address the outstanding issues from the benchmark survey, numerous initiatives were implemented. Several peer education videos to provide peer support on common issues that physicians encounter in using the EHR. In addition, a monthly newsletter was implemented to communicate new changes that have been implemented to the EHR. Lastly, speech recognition technology (SRT) was procured and implemented to improve efficiency in documentation. SRT allows physicians to dictate their documentation directly into the EHR instead of typing [20].

Using EHR Usage Log Data to Guide Further Optimizations

In order to assess the impact of these interventions on the EHR, back-end usage log data are being collected and reviewed. The solution is included as part of the EHR system and includes metrics such as documentation time per patient and after-hours use of the EHR (pajama time).

While smaller evaluations have been conducted on individual components [19-21], there remains a gap in understanding how the overall strategy has supported improvements in EHR use over the past 5 years. It is vital to explore opportunities for continuously advancing the strategy as new technologies, such as generative artificial intelligence, are integrated into the environment [22-28]. This is particularly useful given the growing issue of physician burnout and the rapid growth of digital tool adoption in health care organizations outside of Canada [29-32]. The core contribution of this manuscript is to demonstrate the impact and lessons learned from implementing

XSL•FO
RenderX

a multi-pronged organizational strategy focused on reducing EHR burnout in a real-world environment, as well as some of the outstanding areas of focus. This work helps support our limited understanding of these initiatives. This is particularly relevant to other countries and hospitals as they gain more advanced, comprehensive EHRs such as the ones being used at this hospital. EHR-related burnout is a global issue, and these initiatives can be translated to other organizations facing similar challenges.

Objectives

In this study, we present our findings from a survey conducted 5 years after the implementation of the physician engagement strategy. The primary objective is to understand the impact and perceptions of the strategy's 4 components and identify opportunities for further improvement. In order to do so, a survey was developed that asks questions related to the effectiveness of each of the components of the strategy. These findings may be valuable for Canadian health care organizations seeking to adopt a long-term strategy to enhance physician experience and reduce EHR-related burden.

The specific aims of this study were to evaluate overall burnout rates and the extent to which they are linked to the EHR 5 years after implementing the physician engagement strategy; assess the perceived effectiveness of the different components within the physician engagement strategy; and pinpoint key areas for future improvement and optimization of the physician engagement strategy to further alleviate EHR-related burden.

Methods

Overview

A cross-sectional survey was conducted in January 2023 for 6 weeks, to evaluate the impact and challenges in delivering the physician engagement strategy at a large, urban, Canadian mental health hospital.

Participants and Settings

In order to identify physicians who are eligible for the survey, we obtained a list from the Medical Affairs department. The inclusion criteria for physicians are as follows: a staff physician, resident, or fellow at the organization having a full-time or part-time appointment at the organization as of November 2022, and having a psychiatry, addictions, or hospitalist role that provides care across inpatient, outpatient, emergency, outreach, and community settings.

Physicians who joined after November 2022 were not invited to participate in the survey due to their limited exposure and use of the EHR prior to completing the survey.

In terms of setting, the large academic mental health hospital is located in Toronto, Ontario, and provides care to over 34,000 patients a year. The EHR system was implemented in 2014 and is being used for all facets of patient care. In 2017, the organization achieved Healthcare Information Management Systems Society (HIMSS) stage 7 on the Electronic Medical Record Adoption Model (ENRAM) [33], which is used to identify gaps in EHR usage, guide digital transformation, improve data access and patient outcomes, and reduce grunt work. Achieving the highest level, stage 7, of this model is a testament to the high adoption and use of the EHR for clinical care and requires our organization to use our EHR in dynamic ways and leverage analytics and insights.

Data Collection Approach

A web-based survey was deployed to physicians via the REDCap survey tool (developed by Vanderbily University) [34,35]. The survey was developed using a co-design approach with members of the Physician Think Tanks and was informed by the Benefits Evaluation Framework, benchmark survey data, and relevant literature. The survey underwent testing with the Physician Think Tank members to refine its topics, content, and wording. It was available from January to March 2023, with biweekly reminders sent to encourage participation. The Chief Medical Officer (SS) and the Division Chiefs encouraged physicians to participate in the survey. The survey has 4 components, which can be found in Multimedia Appendix 1. The first comprises demographics, which include age, gender, appointment, and years of experience in the organization. The second component focuses on burnout, which includes the Mini-Z Burnout questionnaire [36], as well as a question on the perceived contribution of the EHR to burnout. The third section contains questions related to the perceived impact of the initiatives within the physician engagement strategy, as well as emerging issues of interest outlined by the Physician Think Tanks. The last section is an open-ended question for participants to outline opportunities for improvement in terms of the use of the EHR and initiatives within the physician engagement strategy.

Data Analysis

Both quantitative and qualitative approaches were used to analyze the data. For Likert scale questions, descriptive statistics (eg, mean, standard deviation) were generated. In addition, one-way ANOVA and Chi-squared tests were used to compare burnout rates from the benchmark survey. For qualitative data, a content analysis approach was used to identify the main themes [37,38]. The analysis was led by authors TT and BL. TT conducted a review of the qualitative responses and developed an initial codebook. BL then reviewed and identified any discrepancies. Any discrepancies were reviewed and discussed with the rest of the project team. Based on the revised codebook, TT completed the coding of the qualitative responses. BL then reviewed the coding to identify and discuss any additional discrepancies. To ensure the trustworthiness of the data, 2 reviewers have analyzed the data using the approaches outlined by Elo et al [39]

Ethical Considerations

This work has received approval by the Quality Project Ethics Review (QPER) Board at the Centre for Addiction and Mental Health (QPER 2018 - 2019 #014) as it was a quality improvement project and did not require approval from our Research Ethics Board. Informed consent was obtained for physicians participating in the project. Numerous safeguards were also implemented to ensure the confidentiality of participants. First, data were collected anonymously. Participants were instructed not to share identifying information in their

XSL•FO
RenderX

responses. All survey data were de-identified before analysis if individuals entered identifying information in the survey. Second, findings were reported only in aggregate form to both internal and external audiences. Compensation was not provided to participants.

Results

Participants

Of 254 staff physicians eligible to complete the survey, 128 participants completed it and were included in the analysis (50%). There were 18 individuals who opened the link to the survey but did not submit the survey. While we had originally invited residents and staff fellows to the survey, despite our extensive efforts in recruitment, we had received very few responses resulting in a negligible response rate. Hence, they were not included in the analysis of the survey. The findings of the survey are described below.

Respondent Characteristics

The demographics reported by participants who completed the survey (n=128) are outlined in Table 1. Respondents were primarily staff physicians with outpatient duties and provided psychiatric care to adult populations. There was a range in years of experience of respondents, with 27% (35/128) of participants, reporting 0 - 5 years of experience. In terms of gender identity and ethnicity, there was about equal responses from participants identifying as a man (45%, 58/128) or woman (42%, 54/128), and 53% of individuals (68/128) identified with a white ethnic group.

XSL•FO
RenderX

Table . Demographic characteristics of respondents.

Characteristic category	Number of participants (%) (n=128)
Clinical activities (multiple options were allowed to be selected)	
Inpatient physician	46 (36)
Outpatient physician	93 (73)
Emergency department	22 (17)
Community-based care	10 (8)
Outreach	8 (6)
Primary division	
Addictions	23 (18)
Adult neurodevelopment and geriatric psychiatry	12 (9)
Child and youth psychiatry	11 (9)
Forensics	8 (6)
General adult and health systems psychiatry	46 (36)
Hospital medicine	15 (11)
Schizophrenia	13 (10)
Frequency in providing care	
Full-time (≥0.6FTE)	106 (83)
Part-time (<0.6FTE)	22 (17)
Experience in practicing as a physician	
0 - 5 Years	35 (27)
6 - 10 Years	26 (20)
11 - 15 Years	24 (19)
16 - 20 Years	12 (9)
21 - 25 Years	12 (9)
26+ Years	18 (14)
Prefer not to answer	1 (1)
Age group	
<30 years old	3 (2)
31 - 40 years old	46 (36)
41 - 50 years old	39 (30)
51 - 60 years old	19 (15)
61+ years old	13 (10)
Prefer not to answer or did not respond	8 (6)
Gender identity	
Man	58 (45)
Non-binary	2 (2)
Woman	54 (42)
Prefer not to answer or not reported	14 (11)
Ethnicity	
East Asian	11 (9)
Latin American	2 (2)
Middle Eastern	9 (7)
South Asian	10 (8)

XSL•FO
RenderX

Characteristic category	Number of participants (%) (n=128)
Southeast Asian	2 (2)
White	68 (53)
Indian-Caribbean	1 (1)
Indigenous	1 (1)
Mixed ethnicity	4 (3)
Identity not listed	2 (2)
Prefer not to answer or did not respond	18 (14)

Burnout and Perceptions of the EHR

The level of burnout reported by participants using the Mini Z questionnaire [32] is outlined in Table 2. About 26% (33/128) of individuals reported a description that meets the threshold of burnout. However, only 4% (5/128) of individuals reported being at the highest and second-highest levels of burnout. Of those who indicated a level that meets the threshold of burnout, 61% (20/33) of the respondents indicated that the EHR contributed to their burnout often or always. There was no relationship between burnout levels and age, ethnicity, or years of experience.

Table . Reported levels of burnout for benchmark survey and follow-up survey.

	Burnout, n (%)	
Characteristic category	Baseline survey (N=208) [18]	Follow-up survey (N=128)
I enjoy my work. I have no symptoms of burnout.	45 (22)	27 (21)
Occasionally I am under stress, and I don't always have as much energy as I once did, but I don't feel burned out	111 (53)	68 (53)
I am definitely burning out and have one or more symptoms of burnout, such as physical and emotional exhaustion	35 (17)	28 (22)
The symptoms of burnout that I'm experiencing won't go away. I think about frustration at work a lot.	15 (7)	4 (3)
I feel completely burned out and often wonder if I can go on. I am at the point where I may need some changes or may need to seek some sort of help.	1 (1)	1 (1)

There were no significant differences in burnout level (t(283)=.191, P=.85) between the current survey and the previously conducted benchmark survey [18]. Of those who reported burnout, there is a considerable reduction in the number of individuals who reported that the EHR contributes to burnout all or almost all the time (t(63)=−1.255, P=.21).

With regard to the perceptions towards the EHR, over half the participants (74/128) agreed or strongly agreed that the EHR adds to their daily frustration (Table 3). From a benefits perspective, a similar proportion of participants also felt that the EHR improved communication within their circle of care (67/128), enabled the delivery of high-quality care (49/128), and kept patients safe (50/128).

XSL•FO
RenderX

Table . Perceived effectiveness of the various components of the physician engagement strategy.

Characteristic category	Participants, N=128, n (%)
Do you think [EHR][a] contributes to your symptoms of burnout? (only for those who answered having at least one or more symptoms of burnout (n=33))	
Rarely	1 (3)
Sometimes	12 (36)
Often	14 (42)
Always	6 (18)
[EHR] adds to my daily frustration.	
Strongly disagree	2 (2)
Disagree	20 (16)
Neutral	32 (25)
Agree	51 (40)
Strongly agree	23 (18)
[EHR] improves communication within the circle of care.	
Strongly disagree	2 (2)
Disagree	25 (20)
Neutral	34 (27)
Agree	55 (43)
Strongly agree	12 (9)
[EHR] enables me to deliver high-quality care.	
Strongly disagree	8 (6)
Disagree	18 (14)
Neutral	53 (41)
Agree	43 (34)
Strongly agree	6 (5)
[EHR] helps keep my patients safe.	
Strongly disagree	4 (3)
Disagree	19 (15)
Neutral	55 (43)
Agree	44 (34)
Strongly agree	6 (5)

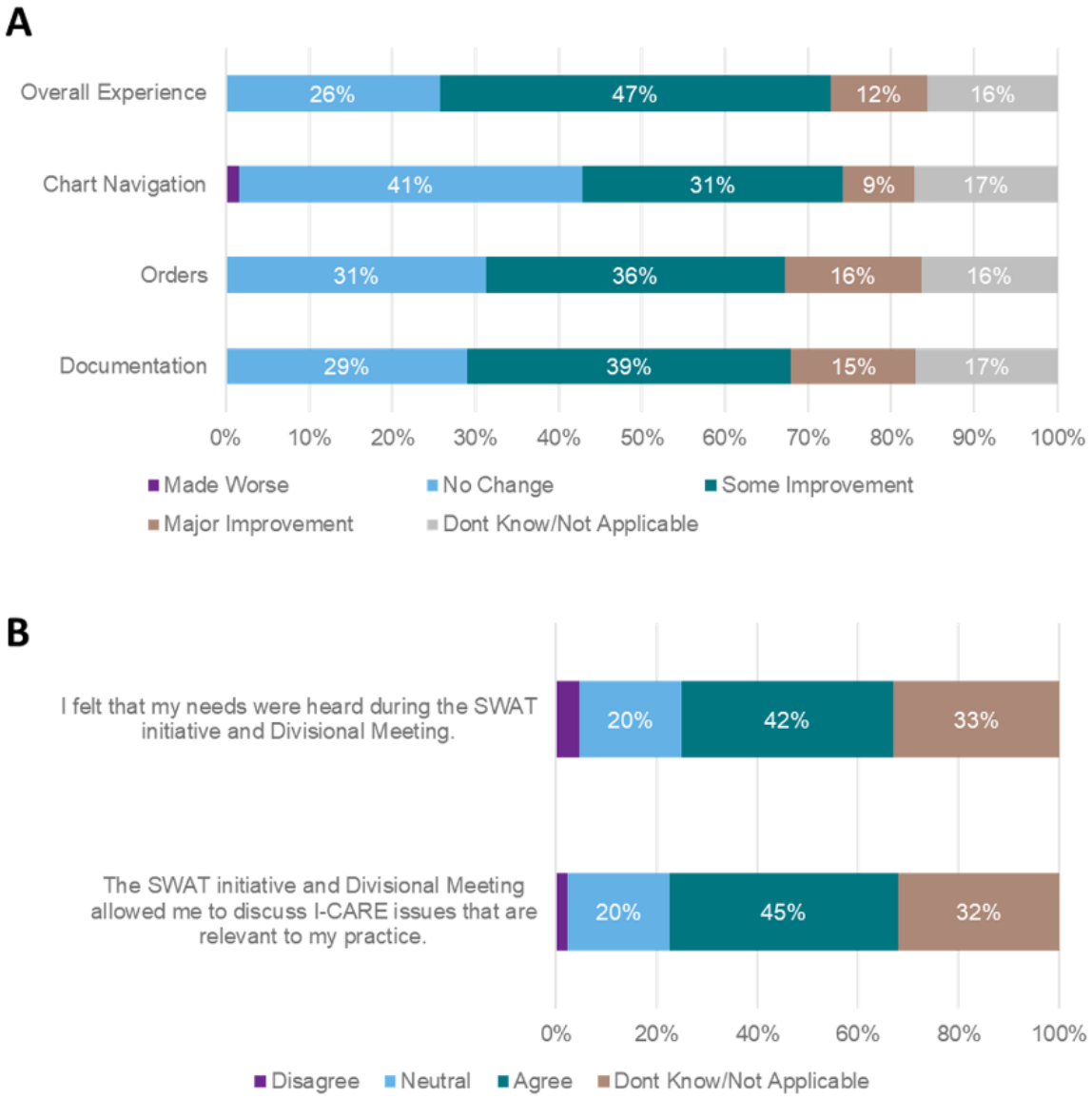
[a]EHR: electronic health record.

SWAT Initiative—Rapid Handling of EHR Issues

In the survey, participants indicated whether the SWAT initiative improved documentation, order entry, and chart navigation within the EHR (Figure 1A). Of those who provided a response, about 54% of participants reported that the SWAT initiative had made some or major improvements across documentation (69/128) and order entry. For chart navigation, the numbers of respondents were marginally less, with 40% of participants indicating that SWAT has made some or major improvements to the functionalities related to chart navigation (51/128).

XSL•FO
RenderX

Figure 1. Impact of SWAT (agile approach for resolving electronic health record (EHR)-related issues) initiative on EHR burden (n=128).



In terms of overall experience (Figure 1A), approximately 59% (75/128) of respondents reported that the SWAT initiative has made some or major improvement to their use of the EHR. Similarly, about 42% (54/128) of participants agreed or strongly agreed that the initiative made them feel their needs were acknowledged and 45% (58/128) agreed that it provided a platform to discuss issues pertinent to their practice (Figure 1B).

Communication and Education

With regard to communication, about 77% (98/128) of respondents recall reading the monthly newsletter that is sent out from the strategy (Figure 2A). Of those who have read the newsletter, about half (50/98) suggested that the communication of EHR changes is efficient through this method.

XSL•FO
RenderX

Figure 2. Impact of communication and education initiatives (n=128). EHR: electronic health record.



For education, about 64% (82/128) of individuals believe they are proficient with the use of the EHR, yet only 39% (50/128) of individuals thought that the education support provided by the organization is available and sufficient (Figure 2B). Regarding peer education videos (Figure 2C), about 32% (41/128) of individuals reported watching 1 of the 3 peer education videos released to date. For those who did not watch a video, about 57% (73/128) of respondents were not aware that peer education videos existed. Among the 41 individuals who watched a video, 68% (28/41) of individuals reported that the video was useful, and 66% (27/41) of individuals believed that the length and details were appropriate.

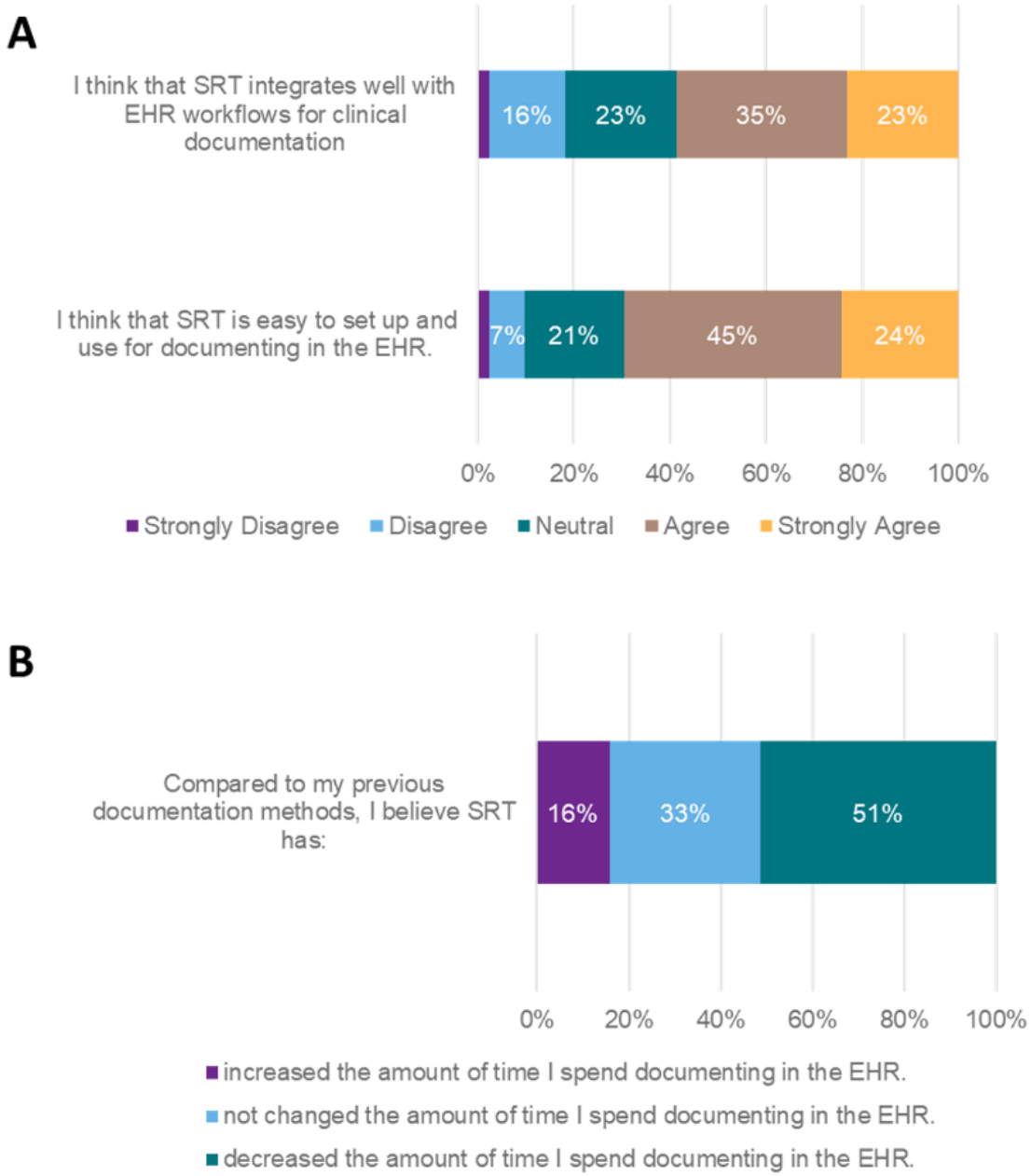
Informatics Tools (SRT Documentation and Continuity of Care)

With regard to documentation, about 64% (82/128) of respondents reported trying to use SRT for documentation (Figure 3A). Of those who indicated using SRT, 69% (57/82)

XSL•FO
RenderX

of individuals reported that it was easy to set up and use, and 58% (48/82) of respondents indicated that it integrates well with the EHR (Figure 3). About 51% (42/82) of respondents believed that the use of SRT decreased their documentation time (Figure 3B).

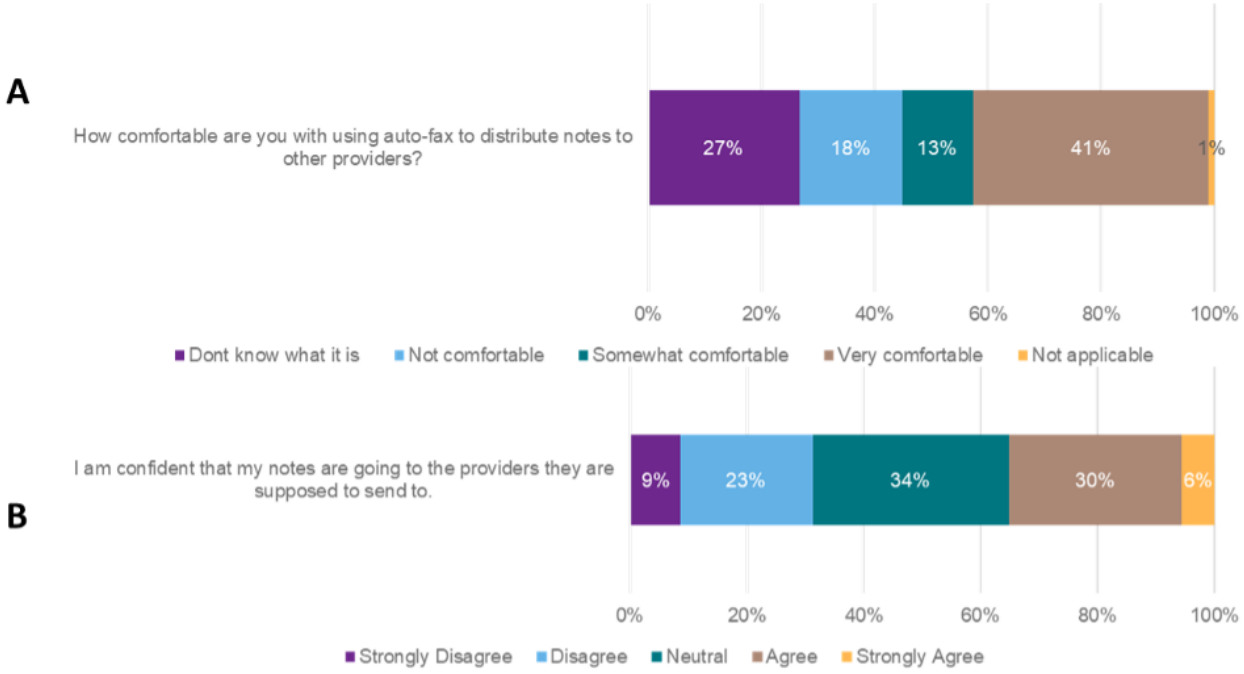
Figure 3. Impact of speech recognition technology on documentation in the electronic health record (EHR; N=128). SRT: speech recognition technology.



For continuity of care, participants were asked to comment on their confidence level regarding delivery of notes to external health care providers. Only 35% (45/128) of individuals agreed that they are confident that notes are sent to the provider as intended (Figure 4A). With automated fax transmission being the default approach for document delivery within the EHR, 54% (69/128) of respondents indicated that they are somewhat or very comfortable with sending their notes through the current automated fax transmission process instead of sending notes through health records distribution process (Figure 4B).

XSL•FO
RenderX

Figure 4. Perceptions of continuity of care in the electronic health record (n=128).



Areas of Focus

From the qualitative findings, 4 themes were identified. These include (1) medication reconciliation and prescriptions processes; (2) chart navigation and information retrieval; (3) longitudinal medication history; and (4) technology infrastructure challenges. These themes are detailed below.

Medication Reconciliation and Prescriptions Processes

Physicians had concerns about the current capabilities of the EHR with regard to medication reconciliation and transmission of prescriptions. While medication reconciliation and communication of prescriptions were identified as highly important for supporting patient care, the processes involved were identified as cumbersome and time-consuming. As one physician outlines:

> *Outdated and clunky interface - navigation is clumsy and could be streamlined. More like 2010 than 2023 software…Med Reconciliation is a slow and laborious chore - great when pharmacists do it.* [Psychiatrist, SC-FT-246]

Another physician also highlights the lack of intuitive search functionality with the current EHR system:

> *Discharge med rec is frequently broken and unsafe. Search function is desperately needed. For example, cognitive testing hard to find or not accessible, not organized, not uploaded.* [Geriatrics Psychiatry Physician, GE-APP-47]

Collectively, these limitations of the system can make the process unsafe for patient care. Furthermore, the impact on continuity of care back to the community also affects the process at discharge. Fax is still commonly used for transmission of prescriptions and documents to community providers and pharmacies, but the limitations of this technology represent a significant burden for clinicians. This is particularly outlined from a physician who engages in telepsychiatry practice:

> *The prescription function on [the EHR] is awful for telemedicine prescriptions. First, I have to enter the prescription on the software. Next, I have to download the PDF, find that, and sign manually. Finally, I have to open separate software [company] to fax the prescription. This whole process takes 5 - 10 minutes, whereas with the other EMR software I can send a prescription in<1 min. Since I send many prescriptions per day, I can waste anywhere between 30 - 60 min of a clinical day on this cumbersome prescription faxing process.* [Telemedicine Psychiatrist, AD-AP-22]

Collectively, the administrative burden spent on faxing prescriptions can take a significant toll on physicians. More importantly, the lack of feedback from the process hinders the ability of the physician to know if the process worked for the patient:

> *Sending external lab requisitions and prescriptions, especially when delivering virtual care is non-obvious, tedious, time-consuming and error prone. For instance, for nearly 3 months my prescriptions had a return fax number that no one could tell me if and where returned faxes from pharmacies were going to.* [Addiction Psychiatrist, AP-AD-6]

Given the growing adoption of team-based care within the community, continuity of care through current medication reconciliation processes and faxing significantly hinders the value and utility of the EHR.

XSL•FO
RenderX

Chart Navigation and Information Retrieval

Given the complexity and chronicity of the patient population, chart navigation was another critical theme outlined by physicians. For example, several physicians highlighted issues about being able to view other notes while documenting:

> *It is horrific that one cannot simultaneously see a document and type into a note. I don't know who designed such a system, they do not understand physician work at all. Opening an ED note locks you out of seeing other documentation.* [Addiction Psychiatrist, AD-FT-33]

> *Being unable to quickly click through the pages, for example have a tab to click one month back or to be able to skip an inpatient admission. Being unable to have a note open while I am also looking at other applications. Have to write the draft note somewhere else and paste it. [General Adult Psychiatrist, GA-FT-171]*

As such, being able to review and generate new documentation can be burdensome and can result in developing workarounds that are not aligned with best practices. Another key limitation of the system is the limited ability to search for relevant documentation. Given that some patients have many previous consultations, searching through prior documentation efficiently is critical. One physician highlighted that it:

> *It can be hard to find needed documentation if the chart is extensive because the filtering system isn't great/doesn't let you just search for the things you need.* [General Adult Psychiatrist, GA-FT-158]

As such, this can hinder the ability for physicians to understand the broader clinical picture in a timely manner through the EHR.

Longitudinal Medication History

A review of past medication history often informs decision-making regarding which medication regimen to trial next for a particular patient. However, currently, there is a limited ability to review the medication histories of patients in a longitudinal way:

> *[The EHR] is not well set up for longitudinal care. Medication histories, for example, are not well captured, nor are key pieces of information easily parked somewhere that can be referenced later without having to scroll through many progress notes.* [Schizophrenia Physician, SZ-AP-220]

Having to review numerous progress notes can make it difficult for physicians to grasp the patient's history and develop an evidence-based treatment plan.

Challenges with Technology Infrastructure

Since the implementation of the EHR 10 years ago, physicians have reported growing issues with regard to the challenges with IT infrastructure. For example, one psychiatrist highlights how the application slows down the system when open and takes excessive amounts of time to complete tasks within the EHR.

> *The program freezes frequently and then it is a process to restart it. [The EHR] often has a lag and this slows down work for outpatient work, there is no easy way to fax prescriptions. What I do is create a PDF, open it in program (Premium which I pay for personally), sign it, save it, then open the [fax] program, and then send it. The fax process itself takes 5+ minutesand isn't an effective use of physician time.* [General Adult Psychiatrist, GA-FT-184]

This process becomes especially challenging when physicians must act under pressure, such as during an emergency code:

> *To work on multiple units when on call- Duty Doctor is very time consuming as clinicians are required to start up their [EHR] on a new computer on each unit and the initializing process to get into the platform to access [EHR] takes 5 - 10 minutes sometimes which is very delayed when responding to a code where a clinician is required to order meds/restraints etc.* [General Adult Psychiatrist, GA-FT-114]

Since IT infrastructure is fundamental to EHR usability, it is crucial that it remain up to date to ensure the system operates efficiently and effectively over time, particularly in critical situations.

Discussion

Principal Findings and Comparison With Previous Work

This study examines the current findings from the physician engagement strategy at a Canadian mental health hospital in Toronto, Canada. Although extensive literature has investigated the causes of EHR-related physician burnout and proposed various approaches to address these issues [6], there is limited research on how physician engagement strategies and their components interact strategically to mitigate burnout through a health informatics perspective. To our knowledge, this work is among the first to involve a large number of Canadian physicians in assessing the impact of these initiatives within a cohesive and supportive environment. The findings from this work are critical given that many of the current studies are pilot or feasibility studies in a small setting or clinic. This work provides preliminary insights into how these interventions can be expanded to an organization and their impact. This is particularly important given the growing adoption of advanced EHRs with decision support, artificial intelligence, and other capabilities [40-44]. Overall, the physician engagement strategy has positively impacted perceptions and value of the EHR among physicians. However, there are still many opportunities for improvement, particularly at the system level. The implications of these findings are described below.

We found that while the physician overall burnout rate remained unchanged from the benchmark survey in 2018, the perceptions of the EHR's impact on physician burnout have considerably improved. Physicians who used the physician engagement initiatives reported enhancements in EHR usage, communication, education, and documentation, which are key components of the overall strategy [18]. These findings are consistent with the growing literature on the role of perceptions of physician burnout. Several studies have highlighted how

XSL•FO
RenderX

perceptions can influence overall satisfaction and utility of the EHR, despite its impact from a quantitative (objective) perspective [45]. Studies that take a sociotechnical perspective suggest that characteristics of the individual and the environment can have a profound impact on burnout rate and EHR usage patterns [46]. With this in mind, it is likely that understanding EHR-related burnout requires a broader consideration of the environment and the system at play.

Since the evaluation was conducted 5 years post-implementation of the strategy, improvements seem to be sustained over time. This provides promising support for the long-term impact of initiatives such as a physician engagement strategy, which is an area that has not been widely investigated [6].

Several areas for improvement have been identified since the strategy's inception, many of which necessitate system-level interventions. For example, through the qualitative SWAT interviews, issues related to the transmission of prescriptions and documentation continue to be critical EHR-related barriers to continuity of care. In the Canadian setting, initiatives supporting electronic prescribing continue to progress at a slow pace, despite the high interest and discussion over the last decade [47]. In addition, functional interoperability could be used to support better continuity of care and the sunsetting of aging technology such as faxing [48]. As such, while these initiatives have made improvements to the use of the EHR, there continues to be a significant requirement for novel, system-level solutions that fall outside the scope of a single organization or single digital tool. As a next step from this work, the project team will focus the efforts of the strategy on addressing these challenges.

Finally, this work makes a significant contribution to the international literature on interventions to reduce EHR burden for physicians. To date, most of the literature, including the 25 × 5 toolkit and the Getting Rid of Stupid Stuff initiative, has a strong focus on the US context [10]. However, these constraints do not necessarily apply to the Canadian and international context. As such, the impact observed in this study provides suggestive evidence of the broad impact these tools can have, irrespective of clinical and administrative environments. Future work should focus on examining the impact of similar initiatives and strategies in other geographical and cultural contexts.

This work has broad-reaching implications for health care organizations and health systems. At the organization level, clinicians and health administrators can consider building and evaluating similar strategies within their own organizations. By leveraging similar evaluation approaches and methodologies (eg, Mini-Z for burnout), it will enable comparison and optimization of the initiatives outlined in this strategy. For researchers, it would be useful to explore how theoretical frameworks from implementation science and change management can be used to further understand the impact and key success factors in rolling out these initiatives across complex health care environments. At the health policy and system level, policy makers and health care organizations can leverage the findings from this study to inform the development of system-level recommendations and initiatives to address the root causes that are consistent across the health care system (eg, interoperability).

Limitations

The study's findings should be interpreted with caution due to several limitations. Foremost, the strategy is currently being evaluated in a single health care organization for staff physicians only. Despite our efforts to recruit residents and fellows through engagement of leadership and communications, the rapid turnover and limited time at the organization make it difficult to recruit learners to complete the survey. Future work should focus on exploring ways to enhance reach and engage these groups in EHR-related burden. Moreover, due to the limited number of time points available in the evaluation, this study fails to examine how the broader environment and the temporal nature of events can influence the impact of this strategy on EHR burden. In addition, there are currently limited validated instruments to examine the effectiveness of these interventions. As such, the questions of the survey were based on the feedback from the liaison members of the Physician Think Tank and the members of the project team. Future work should focus on developing validated survey instruments to assess initiatives for reducing EHR-related burden [6]. Lastly, this study did not examine objective data sources like EHR usage logs, which could provide a more comprehensive understanding of the initiatives' impacts.

Conclusions

This study evaluates the impact of a physician engagement strategy on improving physician experience and reducing EHR burden at a Canadian mental health organization. The study highlights the ongoing challenges and perceptions of staff physicians regarding the EHR system, emphasizing its impact on burnout, workflow efficiency, and patient care. While physicians recognized that the EHR improved communication and supported high-quality care, the system also played a significant role in daily frustration for many physicians, with many citing the usability of the system as a major challenge. However, the cross-sectional survey suggests significant improvements in EHR use since the inception of the strategy in 2018. In addition, the findings from the survey demonstrate the feasibility and preliminary effectiveness of these initiatives on reducing EHR burden at an organizational level. Despite some positive findings, additional interventions and system-level changes are needed to further reduce EHR burden and enhance physician experience and patient care. Gaps continue to persist in areas such as medication reconciliation, chart navigation, and the reliability of IT infrastructure. Targeted improvements such as streamlining prescription processes, enhancing search functionalities, and addressing system inefficiencies could alleviate some administrative burdens and improve physician satisfaction. Addressing these concerns is essential to optimizing the usability of the EHR and supporting physician well-being in clinical practice. Future work should also focus on exploring how these system-level changes can be adopted and explored in other health systems.

XSL•FO
RenderX

Acknowledgments

The authors of this study would like to acknowledge the support from the Centre for Addiction and Mental Health and the physicians, clinicians, clinical informatics nurses, and IT experts, involved in providing their feedback in the survey and the strategy. The authors would also like to acknowledge Jessica Kemp for her contributions to support the development and revisions required for this manuscript. This work is funded by the Centre for Addiction and Mental Health and informally supported by grants from the Canadian Institutes for Health Research (FRN # PCS-185454). BL is supported by a Frederick Banting & Charles Best Canada Graduate Scholarships-Doctoral Research Award (FRN# FBD-177877) from the Canadian Institutes for Health Research.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Physician Survey.
[PDF File, 58 KB - humanfactors_v12i1e65656_app1.pdf]

References

1. Arnsten AFT, Shanafelt T. Physician distress and burnout: the neurobiological perspective. Mayo Clin Proc 2021 Mar;96(3):763-769. [doi: 10.1016/j.mayocp.2020.12.027] [Medline: 33673923]
2. Association (CMA) CM. A struggling system: understanding the health care impacts of the pandemic. The Association. URL: https://digitallibrary.cma.ca/link/digitallibrary7 [accessed 2025-05-01]
3. Physician burnout nearly doubles during pandemic |. CMA. URL: https://www.cma.ca/about-us/what-we-do/press-room/physician-burnout-nearly-doubles-during-pandemic [accessed 2025-01-21]
4. Health care groups call on premiers to make canada's collapsing health system their top priority. Canadian Medical Association. URL: https://www.cma.ca/about-us/what-we-do/press-room/health-care-groups-call-premiers-make-canadas-collapsing-health-system-their-top-priority [accessed 2024-11-04]
5. Johnson KB, Neuss MJ, Detmer DE. Electronic health records and clinician burnout: a story of three eras. J Am Med Inform Assoc 2021 Apr 23;28(5):967-973. [doi: 10.1093/jamia/ocaa274] [Medline: 33367815]
6. 25 By 5: Symposium to Reduce Documentation Burden on US Clinicians by 75% by 2025. URL: https://www.dbmi.columbia.edu/25x5/ [accessed 2022-04-07]
7. Measuring and addressing physician burnout. : American Medical Association; 2024. URL: https://www.ama-assn.org/practice-management/physician-health/measuring-and-addressing-physician-burnout [accessed 2025-01-21]
8. AMIA 25x5. AMIA - American Medical Informatics Association. URL: https://amia.org/about-amia/amia-25x5 [accessed 2024-11-04]
9. Lo B, Kemp J, Cullen C, Tajirian T, Jankowicz D, Strudwick G. Electronic health record-related burnout among clinicians: practical recommendations for Canadian healthcare organizations. Healthc Q 2020 Oct;23(3):54-62. [doi: 10.12927/hcq.2020.26332] [Medline: 33243367]
10. Otokiti AU, Craven CK, Shetreat-Klein A, Cohen S, Darrow B. Beyond getting rid of stupid stuff in the electronic health record (Beyond-GROSS): protocol for a user-centered, mixed-method intervention to improve the electronic health record system. JMIR Res Protoc 2021 Mar 16;10(3):e25148. [doi: 10.2196/25148] [Medline: 33724202]
11. Sieja A, Markley K, Pell J, et al. Optimization sprints: improving clinician satisfaction and teamwork by rapidly reducing electronic health record burden. Mayo Clin Proc 2019 May;94(5):793-802. [doi: 10.1016/j.mayocp.2018.08.036] [Medline: 30824281]
12. Sieja A, Whittington MD, Patterson VP, et al. The influence of a Sprint optimization and training intervention on time spent in the electronic health record (EHR). JAMIA Open 2021 Jul;4(3):ooab073. [doi: 10.1093/jamiaopen/ooab073] [Medline: 34435176]
13. Chen J, Chi WN, Ravichandran U, et al. Sprint-inspired one-on-one post-go-live training session (mini-sprint) improves provider electronic health record efficiency and satisfaction. Appl Clin Inform 2024 Mar;15(2):313-319. [doi: 10.1055/s-0044-1786368] [Medline: 38657955]
14. Strudwick G, Jeffs L, Kemp J, et al. Identifying and adapting interventions to reduce documentation burden and improve nurses' efficiency in using electronic health record systems (The IDEA Study): protocol for a mixed methods study. BMC Nurs 2022 Aug 4;21(1):213. [doi: 10.1186/s12912-022-00989-w] [Medline: 35927701]
15. Ontario Medical Association. Healing the healers: system-level solutions to physician burnout recommendations of the ontario medical association burnout task force. Ontario Medical Association. 2021 URL: https://www.oma.org/siteassets/oma/media/pagetree/advocacy/issues/burnout/burnout-paper.pdf [accessed 2025-05-01]

XSL•FO
RenderX

16. Weisel KK, Fuhrmann LM, Berking M, Baumeister H, Cuijpers P, Ebert DD. Standalone smartphone apps for mental health-a systematic review and meta-analysis. NPJ Digit Med 2019;2(118):118. [doi: 10.1038/s41746-019-0188-8] [Medline: 31815193]
17. Tajirian T, Jankowicz D, Lo B, et al. Tackling the burden of electronic health record use among physicians in a mental health setting: physician engagement strategy. J Med Internet Res 2022 Mar 8;24(3):e32800. [doi: 10.2196/32800] [Medline: 35258473]
18. Tajirian T, Stergiopoulos V, Strudwick G, et al. The influence of electronic health record use on physician burnout: cross-sectional survey. J Med Internet Res 2020 Jul 15;22(7):e19274. [doi: 10.2196/19274] [Medline: 32673234]
19. Sequeira L, Almilaji K, Strudwick G, Jankowicz D, Tajirian T. EHR "SWAT" teams: a physician engagement initiative to improve Electronic Health Record (EHR) experiences and mitigate possible causes of EHR-related burnout. JAMIA Open 2021 Apr;4(2):ooab018. [doi: 10.1093/jamiaopen/ooab018] [Medline: 33898934]
20. Lo B, Sequeira L, Strudwick G, et al. Accuracy of physician electronic health record usage analytics using clinical test cases. Appl Clin Inform 2022 Oct;13(5):928-934. [doi: 10.1055/s-0042-1756424] [Medline: 36198309]
21. Lo B, Almilaji K, Jankowicz D, Sequeira L, Strudwick G, Tajirian T. Application of the i-PARIHS framework in the implementation of speech recognition technology as a way of addressing documentation burden within a mental health context. AMIA Annu Symp Proc 2021;2021:803-812. [Medline: 35308937]
22. Asch DA. An Interview With ChatGPT About Health Care Catalyst Non-Issue Content: Massachusetts Medical Society; 2023, Vol. 4. [doi: 10.1056/CAT.23.0043]
23. Cascella M, Montomoli J, Bellini V, Bignami E. Evaluating the feasibility of ChatGPT in healthcare: an analysis of multiple clinical and research scenarios. J Med Syst 2023 Mar 4;47(1):33. [doi: 10.1007/s10916-023-01925-4] [Medline: 36869927]
24. Nguyen J, Pepping CA. The application of ChatGPT in healthcare progress notes: a commentary from a clinical and research perspective. Clin Transl Med 2023 Jul;13(7):e1324. [doi: 10.1002/ctm2.1324] [Medline: 37394880]
25. Sajjad M, Saleem R. Evolution of healthcare with ChatGPT: a word of caution. Ann Biomed Eng 2023 Aug;51(8):1663-1664. [doi: 10.1007/s10439-023-03225-x] [Medline: 37219697]
26. Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. Healthcare (Basel) 2023 Mar 19;11(6):887. [doi: 10.3390/healthcare11060887] [Medline: 36981544]
27. Vaishya R, Misra A, Vaish A. ChatGPT: Is this version good for healthcare and research? Diabetes Metab Syndr 2023 Apr;17(4):102744. [doi: 10.1016/j.dsx.2023.102744] [Medline: 36989584]
28. Wang C, Liu S, Yang H, Guo J, Wu Y, Liu J. Ethical considerations of using ChatGPT in health care. J Med Internet Res 2023 Aug 11;25(1):e48009. [doi: 10.2196/48009] [Medline: 37566454]
29. Melnick ER, West CP, Nath B, et al. The association between perceived electronic health record usability and professional burnout among US nurses. J Am Med Inform Assoc 2021 Jul 30;28(8):1632-1641. [doi: 10.1093/jamia/ocab059] [Medline: 33871018]
30. Physician burnout: a global crisis. Lancet 2019 Jul 13;394(10193):93. [doi: 10.1016/S0140-6736(19)31573-9] [Medline: 31305255]
31. Ortega MV, Hidrue MK, Lehrhoff SR, et al. Patterns in physician burnout in a stable-linked cohort. JAMA Netw Open 2023 Oct 2;6(10):e2336745. [doi: 10.1001/jamanetworkopen.2023.36745] [Medline: 37801314]
32. Rotenstein LS, Torre M, Ramos MA, et al. Prevalence of burnout among physicians. JAMA 2018 Sep 18;320(11):1131. [doi: 10.1001/jama.2018.12777]
33. Sulkers H, Tajirian T, Paterson J, et al. Improving inpatient mental health medication safety through the process of obtaining HIMSS Stage 7: a case report. JAMIA Open 2019 Apr;2(1):35-39. [doi: 10.1093/jamiaopen/ooy044] [Medline: 31984343]
34. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381. [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
35. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208. [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
36. Olson K, Sinsky C, Rinne ST, et al. Cross-sectional survey of workplace stressors associated with physician burnout measured by the Mini-Z and the Maslach Burnout Inventory. Stress Health 2019 Apr;35(2):157-175. [doi: 10.1002/smi.2849] [Medline: 30467949]
37. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
38. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
39. Elo S, Kääriäinen M, Kanste O, Pölkki T, Utriainen K, Kyngäs H. Qualitative content analysis: a focus on trustworthiness. SAGE Open SAGE Publications 2014;4(1):2158244014522633. [doi: 10.1177/2158244014522633]
40. Park EH, Watson HI, Mehendale FV, O'Neil AQ, Clinical Evaluators. Evaluating the impact on clinical task efficiency of a natural language processing algorithm for searching medical documents: prospective crossover study. JMIR Med Inform 2022 Oct 26;10(10):e39616. [doi: 10.2196/39616] [Medline: 36287591]

XSL•FO
RenderX

41. Luu HS, Filkins LM, Park JY, et al. Harnessing the electronic health record and computerized provider order entry data for resource management during the COVID-19 pandemic: development of a decision tree. JMIR Med Inform 2021 Oct 18;9(10):e32303. [doi: 10.2196/32303] [Medline: 34546942]
42. Allam A, Feuerriegel S, Rebhan M, Krauthammer M. Analyzing patient trajectories with artificial intelligence. J Med Internet Res 2021 Dec 3;23(12):e29812. [doi: 10.2196/29812] [Medline: 34870606]
43. Yang X, Mu D, Peng H, et al. Research and application of artificial intelligence based on electronic health records of patients with cancer: systematic review. JMIR Med Inform 2022 Apr 20;10(4):e33799. [doi: 10.2196/33799] [Medline: 35442195]
44. Zou X, He W, Huang Y, et al. AI-driven diagnostic assistance in medical inquiry: reinforcement learning algorithm development and validation. J Med Internet Res 2024 Aug 23;26(1):e54616. [doi: 10.2196/54616] [Medline: 39178403]
45. Denton CA, Soni HC, Kannampallil TG, et al. Emergency physicians' perceived influence of EHR use on clinical workflow and performance metrics. Appl Clin Inform 2018 Jul;9(3):725-733. [doi: 10.1055/s-0038-1668553] [Medline: 30208497]
46. Lanham HJ, Sittig DF, Leykum LK, Parchman ML, Pugh JA, McDaniel RR. Understanding differences in electronic health record (EHR) use: linking individual physicians' perceptions of uncertainty and EHR use patterns in ambulatory care. J Am Med Inform Assoc 2014;21(1):73-81. [doi: 10.1136/amiajnl-2012-001377] [Medline: 23698256]
47. Motulsky A, Liang M, Moreault MP, Borycki E, Kushniruk A, Sicotte C. Evaluation of a nationwide e-prescribing system. Stud Health Technol Inform 2019 Aug 21;264:714-718. [doi: 10.3233/SHTI190316] [Medline: 31438017]
48. Gheorghiu B, Hagens S. Measuring interoperable EHR adoption and maturity: a Canadian example. BMC Med Inform Decis Mak 2016 Jan 25;16(1):8. [doi: 10.1186/s12911-016-0247-x] [Medline: 26810606]

Abbreviations

EHR: electronic health record
QPER: Quality Project Ethics Review
SRT: speech recognition technology
SWAT: agile approach for resolving EHR-related issues

Edited by A Kushniruk; submitted 21.08.24; peer-reviewed by S Bidmon, S Thawillarp; revised version received 26.02.25; accepted 04.04.25; published 09.05.25.

Please cite as:
Tajirian T, Lo B, Strudwick G, Tasca A, Kendell E, Poynter B, Kumar S, Chang PY(, Kung C, Schachter D, Zai G, Kiang M, Hoppe T, Ling S, Haider U, Rabel K, Coombe N, Jankowicz D, Sockalingam S
Assessing the Impact on Electronic Health Record Burden After Five Years of Physician Engagement in a Canadian Mental Health Organization: Mixed-Methods Study
JMIR Hum Factors 2025;12:e65656
URL: https://humanfactors.jmir.org/2025/1/e65656
doi:10.2196/65656

© Tania Tajirian, Brian Lo, Gillian Strudwick, Adam Tasca, Emily Kendell, Brittany Poynter, Sanjeev Kumar, Po-Yen (Brian) Chang, Candice Kung, Debbie Schachter, Gwyneth Zai, Michael Kiang, Tamara Hoppe, Sara Ling, Uzma Haider, Kavini Rabel, Noelle Coombe, Damian Jankowicz, Sanjeev Sockalingam. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 9.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Swedish Version of the System Usability Scale: Translation, Adaption, and Psychometric Evaluation

Helene Åvik Persson[1], RNCS, PhD; Charlotte Castor[1], RNCS, PhD; Nilla Andersson[1], OTR, PhD; Mia Hylén[2,3], RNCS, PhD

[1]Department of Health Sciences, Lund University, Box 117, Sölvegatan 19, Lund, Sweden

[2]Department of Care Science, Faculty of Health and Society, Malmo University, Malmö, Sweden

[3]Department of Intensive and Perioperative Care, Skane University Hospital, Malmö, Sweden

Corresponding Author:
Helene Åvik Persson, RNCS, PhD
Department of Health Sciences, Lund University, Box 117, Sölvegatan 19, Lund, Sweden

Abstract

Background: The Swedish health care system is undergoing a transformation. eHealth technologies are increasingly being used. The System Usability Scale is a widely used tool, offering a standardized and reliable measure for assessing the usability of digital health solutions. However, despite the existence of several translations of the System Usability Scale into Swedish, none have undergone psychometric validation. This highlights the urgent need for a validated and standardized Swedish version of the System Usability Scale to ensure accurate and reliable usability evaluations.

Objective: The aim of the study was to translate and psychometrically evaluate a Swedish version of the System Usability Scale.

Methods: The study utilized a 2-phase design. The first phase translated the System Usability Scale into Swedish and the second phase tested the scale's psychometric properties. A total of 62 participants generated a total of 82 measurements. Descriptive statistics were used to visualize participants' characteristics. The psychometric evaluation consisted of data quality, scaling assumptions, and acceptability. Construct validity was evaluated by convergent validity, and reliability was evaluated by internal consistency.

Results: The Swedish version of the System Usability Scale demonstrated high conformity with the original version. The scale showed high internal consistency with a Cronbach α of .852 and corrected item-total correlations ranging from 0.454 to 0.731. The construct validity was supported by a significant positive correlation between the System Usability Scale and domain 5 of the eHealth Literacy Questionnaire (P=.001).

Conclusions: The Swedish version of the System Usability Scale demonstrated satisfactory psychometric properties. It can be recommended for use in a Swedish context. The positive correlation with domain 5 of the eHealth Literacy Questionnaire further supports the construct validity of the Swedish version of the System Usability Scale, affirming its suitability for evaluating digital health solutions. Additional tests of the Swedish version of the System Usability Scale, for example, in the evaluation of more complex eHealth technology, would further validate the scale.

Trial Registration: ClinicalTrials.gov NCT04150120; https://clinicaltrials.gov/study/NCT04150120

(JMIR Hum Factors 2025;12:e64210) doi:10.2196/64210

KEYWORDS

application; Swedish; System Usability Scale; usability; validation

Introduction

In the rapidly evolving landscape of global health care, the advent of eHealth technologies has emerged as a transformative force that promises innovative solutions to the multifaceted challenges faced by health care systems worldwide [1,2]. The Swedish health care system is currently transforming along these lines. The use of digital applications and other digital contact methods, collectively described under the term eHealth, is increasing. The World Health Organization (WHO) defines eHealth as "the use of information and communication technologies (ICT) for health." These technologies include a wide range of systems interventions, applications, and devices such as mobile health and telehealth [3]. There is compelling evidence for the increasing influence of eHealth on the provision of health care globally today and how it is enhancing the efficiency and responsiveness of health systems to meet people's needs and expectations [3]. It is essential to ensure that health technologies are designed appropriately to meet the needs of end users before deploying them as health interventions [4].

XSL•FO
RenderX

Employing robust evaluation methods to ensure high-level usability has been recognized as a crucial component of good practice for achieving this goal [5].

Usability is defined as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use" [6]. To determine the potential usability of a digital solution, the System Usability Scale (SUS) has been widely adopted as a standardized evaluative device. In 1996, Brooke [7] published the SUS as an instrument that could easily measure usability. Since then, it has been used to evaluate computer systems, applications, and other digital solutions in a wide range of areas [8-10]. The SUS is a questionnaire, consisting of 10 items each scored on a 5-point Likert scale from "strongly disagree" to "strongly agree." It is easy to administer and has been shown to generate results with good reliability and validity [7,8,11]. The instrument is free, and no fees are required to use it [7]. With an international reach, it is the most widely used standardized questionnaire for the assessment of perceived usability today [10]. The SUS can be used to evaluate a wide range of usability in products, including digital applications, mobile devices, and web pages [8-10].

However, as psychometric properties are sample dependent, it is essential to evaluate the psychometric properties when using patient-reported outcome measure in new settings or populations [12]. The SUS has been translated into numerous languages such as Chinese [13], Finnish [14], French [15], Hindi [15], Indonesian [16], and Polish [17]. It has undergone psychometric validation [18], including in Arabic [19], Danish [20], Dutch [21], German [22], Italian [23], Malay [24], Persian [25], Portuguese [26], Slovene [27], and Spanish [28]. The psychometric properties that have resulted from these studies show that adapted versions of the SUS are a reliable tool for usability assessments. However, a number of these studies adopted a general focus and examined only the total sum in the test. Only a small number of studies have tested the instrument at an item level [19-28], with none of them in Swedish. This emphasizes the need for a comprehensive testing of a Swedish version of the SUS on an item level. Determining the robustness is critical for ensuring that the measurement instrument has sufficient validity in the proposed context.

Lewis [10] outlines the essential requirements for translating the SUS into multiple languages and conducting validations across diverse countries. In Sweden, one translation of the SUS has been published in a scientific journal [29], although several unpublished versions exist [9]. However, the linguistic discrepancies between these translations give room for ambiguities, and a united translation with a rigorous psychometric testing in a Swedish context is warranted.

In 2016, an initial vision was for Sweden to be "the best in the world" in eHealth by 2025. However, a more realistic view today is that Sweden has made significant progress in this area. Within the European context, Sweden demonstrates a distinct approach to digitalization in health care, emphasizing collaboration and innovation to address specific challenges and opportunities [30].

With this in mind, there is a great need for a nationally united SUS that has been rigorously tested and proven effective. A robust process of evaluating the psychometric properties of a Swedish SUS will foster participatory usability research and ultimately improve the quality of health care services on a broader scale. Although there are different Swedish translations of the SUS, there is still a lack of psychometric testing showing their robustness. Before any Swedish SUS can be recommended for use, both translation and psychometric evaluation of the instrument are necessary. Therefore, the aim of this study was to translate and psychometrically evaluate a Swedish version of the SUS.

Methods

Overview

This study consists of 2 steps. In step 1, the SUS was translated and adapted into Swedish. In step 2, the psychometric properties of the Swedish version were tested: data quality, scaling assumptions, acceptability, convergent validity, and internal consistency. The translation, adaptation process, and psychometric evaluation adhered to the COSMIN (Consensus-Based Standards for the Selection of Health Status Measurement Instruments) checklist [31].

The SUS Instrument

The SUS instrument consists of 10 items (statements). The items are divided equally into positively and negatively worded statements. Requests for a response are graded on a scale of agreement ranging from strongly disagree (1) to strongly agree (5). The score is calculated as follows: for positively worded items (1, 3, 5, 7, and 9), the score is the position on the scale (1-5) minus 1, and for negatively worded items (2, 4, 6, 8, and 10), the score is the position on the scale minus 5. Individual item scores, therefore, can range from 0 - 4. The sum of all 10 scores is then multiplied by 2.5, resulting in a total score ranging from 0 - 100, with higher numbers representing greater usability [7,8].

To date, factor analysis has not been able to show conclusively whether the SUS consists of one factor (usability) or two (usability and learnability) [11]. Efforts to replicate these findings have led to the conclusion that addressing the instrument as 2 dimensional has no practical or theoretical interest. This study, therefore, treats the SUS as a unidimensional instrument of perceived usability [32].

Translation and Adaption

According to Brooke's [7] original formulation of the SUS, no formal permission is needed for translation, and it can be used free of charge. Further, Brooke [7] allows for the possibility to, in any version, exchange the wording of the scale to a word or expression suitable for the situation. The translation process adopted in this study was inspired by Beaton et al [33]. First, the original instrument was translated from English to Swedish by one of the researchers (CC), proceeding from the versions presented by Bangor et al [34] and Lewis [10]. During the process, perceived difficulties and uncertainties were noted. This first version was then reviewed and discussed by the group, consisting of researchers familiar with eHealth, until consensus

XSL•FO
RenderX

was reached. Second, an authorized translator, naive to the research field, carried out a back translation on the instrument and the notes taken during the research process. Finally, the research group reviewed all versions of the translation, with all notes attached, and finalized a second version of a Swedish SUS. This version was then compared to other SUS translations within the Scandinavian countries for content validity and the original version for expressions and conceptions that could have been culturally influenced. Following this last step, the group then decided upon a final version by consensus. The response options were structured in the same way as the original questionnaire, although the phrases “strongly disagree” and “strongly agree” were exchanged for the Swedish equivalents of “totally disagree” and “totally agree.”

Psychometric Evaluation

Sample

The evaluation of the psychometric properties was carried out in conjunction with a larger intervention study evaluating eHealth, the eChildHealth tablet study [35]. Parents of children with a range of illnesses and health conditions who were patients in the pediatric department of a level-3 hospital in the south of Sweden were invited to take part. A total of 66 parents to 52 unique children gave informed consent and were included in the eChildHealth tablet study. Of these parents, 62 provided information on the SUS, resulting in a total of 82 measurements.

Data Collection

Parents were introduced to an app on a tablet computer through which they could communicate with health care staff after their child had been discharged from hospital. The app made it possible to continue to communicate with hospital staff whom the parents knew well, through chat messages, sending photos, video calls, and predesigned questionnaires. Using a questionnaire, data on various aspects of eHealth were collected for each parent. The SUS was one aspect, and the eHealth Literacy Questionnaire (eHLQ) [36] was another. A study-specific questionnaire was used to collect demographic data such as age and level of education. Between October 2022 and October 2023, the 66 parents were included in study. Data were collected after 1 - 2 weeks of use and at a second time point for those participants who used a tablet for more than 1 month. These parents constitute the eligible participants for this study. A sufficient adequate sample size of approximately 80 measurements was based on recommendations from COSMIN [31] and Beaton et al [33].

Data Analysis

The psychometric properties of the SUS were analyzed with IBM SPSS Statistics for Windows (version 28.0, IBM Corp). Descriptive statistics (mean, SD, and percentage) were used to visualize participants’ characteristics along with data quality, scaling assumptions, and acceptability. Construct validity was evaluated by convergent validity, and reliability was evaluated by internal consistency.

For data quality, use within the clinical setting was determined by item nonresponse and missing scale scores, as they reflect the acceptance and understanding of a measure [37]. Data quality was determined as high if the percentage of missing data per item was low (<10% acceptable). In this study, there were missing data in 5 items (ranging from 1/81, 1% to 3/81, 4%), representing high data quality. Participants with more than 3 unanswered questions were excluded (n=1), while mean imputation was carried out for those with 1 or 2 missing items (n=6, all missing 1 item) [38]. This resulted in a total of 81 measurements being included in the final analysis.

Regarding scaling assumptions, the dimensionality of the SUS has been evaluated previously [32]. In accordance with these studies, this study assumes that the SUS is to be treated as unidimensional (all items measure the same construct). Instruments composed of Likert-scale items can be summarized if they have similar means and SDs. Furthermore, item-total correlations (the correlation between each individual item score and the total score) would indicate if all items contribute equally to the total score. In line with Hobart et al [37], item-total correlations with values of $r \geq 0.3$ were regarded as sufficient for summing up the items to a total sum score.

To evaluate the acceptability of score distributions, ceiling and floor effects along with skewness were calculated. Ceiling and floor effects were regarded as present if they exceeded 90%, that is, the percentage of responses for the lowest and highest scores. Skewness statistics should preferably be within the range of −1 to +1 [37].

Construct validity was explored through convergent validity. It was evaluated with the correlation between the high total sum of the SUS and one of the domains of the eHLQ [36], with the hypothesis that it would correlate with the total score of the SUS. Both instruments were distributed to the parents at the same time points. The eHLQ has already been translated, adapted, and validated within a Swedish context [39]. The eHLQ consists of 7 domains across 35 items, with each domain being extractable and treatable as a separate scale. Responses to each domain on the eHLQ are recorded on a Likert scale from 1 (strongly disagree) to 4 (strongly agree). The eHLQ was designed to be used to understand and evaluate people’s interaction with digital health services [36,39]. The research group independently reviewed the eHLQ and, after discussion, decided to use domain 5, measuring the motivation to engage with digital services. The correlation between the instruments was assessed using the nonparametric correlation coefficient of Kendall Tau-b. With regard to the comparative instrument not being used in total (only 1 domain), the correlation was regarded as acceptable if it was moderate or greater (>0.3) [40] and with the level of significance being $P<.05$.

Internal consistency—how items are related to each other—was explored according to the indicators recommended by Hobart et al [37]: corrected item-total correlations and Cronbach α. The cutoffs were >.04 for acceptable corrected item-total correlations and >.8 for Cronbach α [37,41]. The internal consistency was completed with an SEM to analyze the measurement error of the instrument. The SEM represents the smallest difference in scores and indicates a change on a group level. SEM was analyzed with $SD_{baseline} \times \sqrt{1}$ – reliability and complemented with a CI 95% [12].

XSL•FO
RenderX

Ethical Considerations

This study was conducted in accordance with the Declaration of Helsinki [42] and approved by the Swedish Ethical Review Authority (2021 - 05077). The invited parents were recruited through gatekeepers at the pediatric department. Information regarding the study was initially provided orally and then followed up by written information, before written informed consent was obtained by a study nurse. Data were handled confidentially, and participating parents were able to quit without any explanation or impact on the care their children received. Participation in the study was voluntary, and no financial or other form of compensation was provided to the participants.

Results

Translation and Adaption

Both translations (forward and backward) were similar and did not differ substantially. Overall, the conformity of phrasing was high, with some discrepancy for phrasing in statements 2 (I found the system unnecessarily complex) and 8 (I found the system very cumbersome to use). In statement 2, the word “complicated” was suggested in the back translation from the Swedish word “komplex.” In statement 8, the word “awkward” was suggested instead of the original word “cumbersome.”

All such discrepancies were discussed within the research group and a consensus was reached for the final version. The final step, comparing the Swedish version to other Scandinavian versions, generated no further changes. Table 1 shows the original English version and the final Swedish version of the SUS.

Table . Original English version [7] and the proposed Swedish version of the System Usability Scale, including (1) response alternatives and (2) statements.

Component	English version	Swedish version
Response alternatives	• 1 - Strongly disagree • 2 • 3 • 4 • 5 - Strongly agree	• 1 - Instämmer inte alls • 2 • 3 • 4 • 5 - Instämmer helt
Statements	1. I think that I would like to use this system frequently 2. I found the system unnecessarily complex 3. I thought the system was easy to use 4. I think that I would need the support of a technical person to be able to use this system 5. I found the various functions in this system were well integrated 6. I thought that there was too much inconsistency in this system 7. I would imagine that most people would learn to use this system very quickly 8. I found the system very cumbersome to use 9. I felt very confident using the system 10. I needed to learn a lot of things before I could get going with this system	1. Jag tror att jag skulle vilja använda denna applikation ofta 2. Jag uppfattar denna applikation som onödigt komplex 3. Jag tycker att denna applikation är enkel att använda 4. Jag tror att jag skulle behöva stöd för att kunna använda denna applikation 5. Jag upplever att de olika funktionerna i denna applikation var väl integrerade 6. Jag tycker att applikationen är inkonsekvent 7. Jag föreställer mig att de flesta personer skulle lära sig att använda denna applikation väldigt snabbt 8. Jag upplever denna applikation som krånglig 9. Jag känner mig trygg med att använda denna applikation 10. Jag behövde lära mig många saker innan jag kunde komma igång med denna applikation

[a]In this study, the Swedish instruction was as follows: "Markera det alternativ som bäst beskriver din reaktion för applikation i surfplatta idag," meaning “Please indicate your agreement with the following statements, one at a time.”

Psychometric Evaluation of the Swedish Translation

A total of 62 individuals were included in this study. Of these, 20 answered the questionnaire, including the SUS, twice. This resulted in a total of 81 measurements included in the analysis. Demographic data for the participants are shown in Table 2.

XSL•FO
RenderX

Table . Demographic data of participants included in the study (N=62).

Demographic data	Values
Age (years), median (range)	33 (22 - 52)
Gender, n (%)	
Female	41 (66)
Male	20 (32)
Unknown	1 (2)
Marital status, n (%)	
Married	33 (53)
Living together	28 (45)
Divorced or separated	1 (2)
Education level, n (%)	
High school	18 (29)
College or university	42 (68)
Other	2 (3)
Born in Sweden, n (%)	
Yes	52 (84)
No	6 (10)
Unknown	4 (6)
First language	
Swedish	56 (90)
Other	6 (10)

Data Quality

The percentage of missing data per item was low (ranging from 0/81, 0% to 3/81, 4%) across all items. A tendency could be seen for a higher percentage within the 3 highest steps of the scale, resulting in 5 items having 0% in the lowest points of the 5-point scale (Table 3).

Table . Missing data (n and %) and item frequency distribution (%) of answers per response alternative in each question of the Swedish System Usability Scale (n=81). The item “0” equals the response of “1” on the scale, etc.

Item	Missing data, n (%)	Item frequency distribution, n (%)				
		0	1	2	3	4
1	0 (0)	1 (1)	5 (6)	21 (26)	36 (44)	19 (23)
2	1 (1)	0 (0)	4 (5)	16 (20)	9 (11)	52 (63)
3	0 (0)	1 (1)	0 (0)	10 (12)	24 (29)	47 (58)
4	2 (2)	0 (0)	0 (0)	6 (7)	16 (20)	58 (71)
5	0 (0)	2 (2)	2 (2)	21(26)	32 (39)	25 (31)
6	3 (4)	2 (2)	3 (4)	21 (25)	13 (16)	40 (49)
7	1 (1)	0 (0)	1 (1)	4 (5)	34 (42)	42 (51)
8	2 (2)	1 (1)	1(1)	5 (6)	10 (13)	63 (77)
9	0 (0)	0 (0)	0 (0)	6 (7)	28 (34)	48 (59)
10	0 (0)	0 (0)	1 (1)	4 (5)	14 (17)	63 (77)

Scaling Assumptions

Item means ranged from 2.82 to 3.70 (Table 4). The item-total correlations showed that each item contributed substantially to the total score with correlations ranging from 0.454 to 0.731 (Table 5), thus indicating that the scale can be summarized. The total sum for the SUS in the data ranged between 50 and 100 (mean 84, SD 13).

XSL•FO
RenderX

Table . Item descriptive statistics for the Swedish version of the System Usability Scale (n=81).

Item	Score, mean (SD)	Skewness
1	2.82 (0.904)	−.550
2	3.35 (0.964)	−1.094
3	3.41 (0.800)	−1.490
4	3.65 (0.618)	−1.585
5	2.93 (0.940)	−.766
6	3.09 (1.076)	−.876
7	3.44 (0.652)	−1.038
8	3.66 (0.762)	−2.682
9	3.51 (0.633)	−.943
10	3.70 (0.622)	−2.203

Table . Item-total statistics for the Swedish version of the System Usability Scale (n=81).

Item	Scale mean if item deleted	Corrected item-total correlations	Cronbach α if item deleted
1	30.80	0.463	.847
2	30.30	0.731	.820
3	30.22	0.534	.840
4	30.00	0.657	.833
5	30.72	0.479	.847
6	30.57	0.463	.852
7	30.21	0.454	.846
8	29.99	0.687	.827
9	30.12	0.675	.831
10	29.96	0.585	.837

Acceptability

As indicated by the item mean score (range 2.82-3.70; Table 4) and item-frequency distribution (Table 3), the instrument showed acceptability. For 3 items (4, 8, and 10), the item frequency was above 70%, which is still within the acceptable range for the absence of a ceiling effect. Skewness statistics were below or near the acceptable range of −1 for a total of 7 items, thus indicating that the distribution was excessively skewed (Table 4).

Convergent Validity

Convergent validity was evaluated with the correlation between the total sum of the SUS and the total sum of domain 5 in the eHLQ. As expected, there was a positive correlation between the total score of the instruments (correlation coefficient 0.305), which was significant (*P*=.001) and supported the construct validity of the Swedish version of the SUS.

Internal Consistency and Measurement Error

Cronbach α for the scale was .852. Corrected item-total correlations were between 0.454 and 0.731, as shown in Table 5, indicating internal consistency for the different items. For all items except item 6 (I thought there was too much inconsistency in this system), the α value if the item deleted was lower than the Cronbach α for the scale. The SEM was 5.05 (95% CI –4.84 to 14.954) points for the Swedish version of the SUS.

Discussion

Principal Findings

This study presents a new Swedish version of the SUS that is psychometrically tested. This study seeks to establish the new Swedish version of the SUS as a reliable and valid instrument for assessing system usability. Overall, the psychometric testing showed high data quality, good scaling assumptions, high internal consistency, and fair convergent validity. Together, these analyses support the validity and reliability of the new Swedish version of the SUS.

The translation process of the new Swedish version of the SUS was executed incrementally, both within and outside the research team, involving an authorized translator who conducted a back translation. This method facilitated thorough scrutiny and comparison of the translation from multiple perspectives. The approach also helped reduce the risk of bias and improved the scale's validity by incorporating multiple viewpoints [43].

The psychometric testing regarding internal consistency showed that the Cronbach α values were satisfactory and that all items

XSL•FO
RenderX

contributed to the instrument's total score. This indicates that the Swedish translation of the SUS is a stable instrument to use in a Swedish context.

Convergent validity showed fair correlation between the SUS and domain 5 of the eHLQ, which measured the motivation to engage with digital services (I find that digital technology support me in taking care of my health). This supported our hypothesis that a generally positive attitude toward digital solutions would correlate with a high usability score. This hypothesis is also in line with previous research, which shows that positive expectations directed toward a product generate positive subjective usability ratings [14].

There was an indication of a ceiling effect for 3 items on the Swedish version of the SUS (items 4, 8, and 10). They presented above 70% in item frequency distribution and a high mean total score, which is in line with studies indicating a generally high score for the SUS [10]. Previous studies have also shown ceiling effects of the SUS but only for the total score [15,44], thus the responses on the different items can not be justly compared between studies. This study, however, explores the SUS in much more detail and in line with recommended psychometric evaluations, as it explores the psychometric properties for each item separately [12,37]. Further, the distribution was excessively skewed. This could be an expected result of the nonnormal distribution, since negatively skewed variables are assumed to have a ceiling effect [37,45]. Also, the robustness of skewed distribution as the sole indicator of ceiling effect has been questioned [45].

However, based on the skewness result of the items, the question arises as to whether the range of the response options is wide enough (5-point scale) and if the wording in the response options ("strongly disagree" to "strongly agree") is a sufficient description. It could be interesting to expand the number of response options but, arguably, there could be obstacles with revising such a widely used and widespread instrument.

Regarding the high mean scores for both the total sum and the items of the SUS, it should be noted that the population had a mean age of 33 years and are, therefore, used to digital solutions. This could have influenced their experience of the app. Therefore, the high item scores of the SUS could also indicate a product that is perceived to have usability for this group. A recommendation for future studies would be to test the instrument in other contexts and with different age groups. This app was designed to support users who were in exactly the same situation as the participants in this study: parents in a specific situation. Going home with a child after hospitalization can be stressful; the parents in this study reported this eHealth solution as supportive [46]. The SUS has previously been tested for its sensitivity in different digital solutions and has shown a high scoring (ceiling effect) within the best-of-class products [44].

The Swedish government has envisioned Sweden as a leading country of eHealth in the near future. It has declared the need for individual users to act as the cocreators of such solutions [30]. Within eHealth, there has already been a call for a participatory design to increase the equality of access to digital solutions [47]. Even though digital technology and eHealth are intended to enhance access to health care, for example, despite geographical distances. It may also result in the opposite, as people have different knowledge in using digital solutions. For example, younger people tend to use and access digital solutions differently than older people, which is why solutions should be adapted to the intended end users [48]. In this study, the population was younger, as previously discussed, which could have influenced the result. To have a project design of cocreation often means that end users are involved from the start of the project as collaborators. Coproduction, on the other hand, often involves the end users during the implementation phase [49]. Arguably, the end users are invaluable in all steps of the process, and the usability of products needs to be evaluated where instruments such as the SUS can be useful [50]. This study could therefore be regarded as enhancing the possible participation of end users in the future development of digital solutions and eHealth in a Swedish context.

In conclusion, this study indicates that this developed and psychometrically tested Swedish version of the SUS can be recommended for use within the Swedish adult context.

Limitations

There are some methodological challenges in this study. First, the sample could be regarded as small (N=62), although it was sufficient according to the recommendations [31,33]. A larger sample might enable different methods of analysis, such as Rasch analysis, which might offer deeper psychometric insights.

Second, the 1-year interval for data collection could be considered long in a rapidly changing world. This length of time could influence how the app was perceived. That said, the digital product was not changed to any great extent during this period, so the perception of usability should not have been influenced to any great extent. Additional data collection was carried out after COVID-19 restrictions were essentially withdrawn in Sweden. There is, however, always the possibility of perceptions fluctuating, as the product could be used in various ways according to the circumstances of the parents.

Third, there was an intention to pursue a data quality test for reproducibility as some of the participants answered the SUS twice. Regrettably, the test-retest sample (n=20) was regarded as too small in this study, and such analyses need to be considered in future studies. In addition, the use of imputation can be discussed. There was a low number of missing data, with a noncomplete frequency of only 9 responses in total. The percentage of missing data per item was also low across all items. This resulted in only 6 imputed scores across the items, and based on mean score for the item, this was regarded as not influencing the result noticeably.

Conclusions

This study presents a new Swedish version of the SUS. It is the first study to carry out a psychometric evaluation of a Swedish SUS, to establish the Swedish version of the SUS as a reliable and valid instrument for assessing system usability. Overall, the psychometric testing showed high data quality, good scaling assumptions, high internal consistency, and fair convergent validity, all of which support the validity and reliability of the new Swedish version. The results from this study are promising. They raise the possibility that the Swedish version of the SUS

XSL•FO
RenderX

could be used to evaluate digital health solutions. To further strengthen the usability of the scale, we suggest additional analysis on data that evaluate more complex eHealth technology and include a wider participant age group, both younger and older.

Acknowledgments

We would like to express our sincere gratitude to the participants in the study. We would like to acknowledge Inger Kristensson Hallström, who passed away in 2024, for her extensive work, deep commitment to the field of pediatric nursing, and expertise dedicated to the eChildHealth program. Additionally, we wish to express our appreciation to Magnus C. Persson, for the design of the questionnaires in REDCap (Research Electronic Data Capture) and assisting in data transfer, as well as Desirée Lemmen and other gatekeepers for their assistance with participant inclusion and the distribution of questionnaires to the study participants. This study was funded by the Swedish Research Council for Health, Working Life and Welfare, Sweden (grant 2018-01399).

Authors' Contributions

All authors have made substantial contributions to all of the following: (1) the conception and design of the study, the acquisition of data, or the analysis and interpretation of data; (2) the drafting of this paper or its critical revision for important intellectual content; and (3) final approval of the manuscript version to be submitted.

Conflicts of Interest

None declared.

References

1. Jacob C, Lindeque J, Klein A, Ivory C, Heuss S, Peter MK. Assessing the quality and impact of eHealth tools: systematic literature review and narrative synthesis. JMIR Hum Factors 2023 Mar 23;10:e45143. [doi: 10.2196/45143] [Medline: 36843321]
2. Lennon MR, Bouamrane MM, Devlin AM, et al. Readiness for delivering digital health at scale: lessons from a longitudinal qualitative evaluation of a national digital health innovation program in the United Kingdom. J Med Internet Res 2017 Feb 16;19(2):e42. [doi: 10.2196/jmir.6900] [Medline: 28209558]
3. Digital health. World Health Organization. 2024. URL: https://www.who.int/westernpacific/health-topics/e-health [accessed 2024-07-01]
4. Brown W III, Yen PY, Rojas M, Schnall R. Assessment of the Health IT Usability Evaluation Model (Health-ITUEM) for evaluating mobile health (mHealth) technology. J Biomed Inform 2013 Dec;46(6):1080-1087. [doi: 10.1016/j.jbi.2013.08.001]
5. Zapata BC, Fernández-Alemán JL, Idri A, Toval A. Empirical studies on usability of mHealth apps: a systematic literature review. J Med Syst 2015 Feb;39(2):1. [doi: 10.1007/s10916-014-0182-2] [Medline: 25600193]
6. Dix A. Human-Computer Interaction: Pearson; 2003. URL: https://www.pearson.com/en-gb/subject-catalog/p/human-computer-interaction/P200000008804/9780130461094 [accessed 2025-01-07]
7. Brooke J. SUS: a “quick and dirty” usability. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation In Industry: CRC Press; 1996:189-194 URL: https://www.taylorfrancis.com/chapters/edit/10.1201/9781498710411-35/sus-quick-dirty-usability-scale-john-brooke [accessed 2025-01-07]
8. Brooke J. SUS: a retrospective. J User Exp 2013;8(2):29-40 [FREE Full text]
9. Hägglund M, Scandurra I. User evaluation of the Swedish patient accessible electronic health record: System Usability Scale. JMIR Hum Factors 2021 Jul 27;8(3):e24927. [doi: 10.2196/24927] [Medline: 34313596]
10. Lewis JR. The System Usability Scale: past, present, and future. Int J Hum-Comput Interact 2018 Jul 3;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]
11. Lewis JR, Sauro J. The factor structure of the System Usability Scale. In: Kurosu M, editor. Human Centered Design. HCD 2009. Lecture Notes in Computer Science, vol 5619: Springer; 2009:94-103. [doi: 10.1007/978-3-642-02806-9_12]
12. Hobart J, Cano S. Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods. Health Technol Assess 2009 Feb;13(12):iii, ix-x, 1-iii, ix-x177. [doi: 10.3310/hta13120] [Medline: 19216837]
13. Wang Y, Lei T, Liu X. Chinese System Usability Scale: translation, revision, psychological measurement. Int J Hum-Comput Interact 2020 Jun 14;36(10):953-963. [doi: 10.1080/10447318.2019.1700644]
14. Raita E, Oulasvirta A. Too good to be bad: favorable product expectations boost subjective usability ratings. Interact Comput 2011 Jul;23(4):363-371. [doi: 10.1016/j.intcom.2011.04.002]
15. Gao M, Kortum P, Oswald FL. Multi-language toolkit for the System Usability Scale. Int J Hum-Comput Interact 2020 Dec 13;36(20):1883-1901. [doi: 10.1080/10447318.2020.1801173]
16. Sharfina Z, Santoso HB. An Indonesian adaptation of the System Usability Scale (SUS). 2016 Presented at: 2016 International Conference on Advanced Computer Science and Information Systems (ICACSIS); Malang, Indonesia p. 145-148. [doi: 10.1109/ICACSIS.2016.7872776]

XSL•FO
RenderX

17. Borkowska A, Jach K, Borzemski L, Grzech A, Świątek J, Wilimowska Z, editors. Pre-testing of Polish translation of System Usability Scale (SUS). In: Information Systems Architecture and Technology: Proceedings of the 37th International Conference on Information Systems Architecture and Technology–ISAT 2016–Part I. Advances in Intelligent Systems and Computing, vol 521: Springer; 2017:143-153. [doi: 10.1007/978-3-319-46583-8_12]
18. Streiner DL, Norman GR, Cairney J. Health Measurement Scales: A Practical Guide to Their Development and Use: Oxford University Press; 2024. URL: https://global.oup.com/academic/product/health-measurement-scales-9780199685219?cc=us&lang=en&amp [accessed 2025-01-07]
19. AlGhannam BA, Albustan SA, Al-Hassan AA, Albustan LA. Towards a standard Arabic System Usability Scale: psychometric evaluation using communication disorder app. Int J Hum-Comput Interact 2018 Sep 2;34(9):799-804. [doi: 10.1080/10447318.2017.1388099]
20. Hvidt JCS, Christensen LF, Sibbersen C, Helweg-Jørgensen S, Hansen JP, Lichtenstein MB. Translation and validation of the System Usability Scale in a Danish mental health setting using digital technologies in treatment interventions. Int J Hum-Comput Interact 2020 May 8;36(8):709-716. [doi: 10.1080/10447318.2019.1680922]
21. Ensink CJ, Keijsers NLW, Groen BE. Translation and validation of the System Usability Scale to a Dutch version: D-SUS. Disabil Rehabil 2024 Jan;46(2):395-400. [doi: 10.1080/09638288.2022.2160837] [Medline: 36573399]
22. Moorthy P, Weinert L, Harms BC, Anders C, Siegel F. German version of the mHealth App Usability Questionnaire in a cohort of patients with cancer: translation and validation study. JMIR Hum Factors 2023 Nov 1;10:e51090. [doi: 10.2196/51090] [Medline: 37910144]
23. Borsci S, Federici S, Lauriola M. On the dimensionality of the System Usability Scale: a test of alternative measurement models. Cogn Process 2009 Aug;10(3):193-197. [doi: 10.1007/s10339-009-0268-9] [Medline: 19565283]
24. Mohamad Marzuki MF, Yaacob NA, Yaacob NM. Translation, cross-cultural adaptation, and validation of the Malay version of the System Usability Scale questionnaire for the assessment of mobile apps. JMIR Hum Factors 2018 May 14;5(2):e10308. [doi: 10.2196/10308] [Medline: 29759955]
25. Dianat I, Ghanbari Z, AsghariJafarabadi M. Psychometric properties of the Persian language version of the System Usability Scale. Health Promot Perspect 2014 Jul 12;4(1):82-89. [doi: 10.5681/hpp.2014.011] [Medline: 25097841]
26. Martins AI, Rosa AF, Queirós A, Silva A, Rocha NP. European Portuguese validation of the System Usability Scale (SUS). Proc Comput Sci 2015;67:293-300. [doi: 10.1016/j.procs.2015.09.273]
27. Blažica B, Lewis JR. A Slovene translation of the System Usability Scale: the SUS-SI. Int J Hum Comput Interact 2015 Feb;31(2):112-117. [doi: 10.1080/10447318.2014.986634]
28. Sevilla-Gonzalez MDR, Moreno Loaeza L, Lazaro-Carrera LS, et al. Spanish version of the System Usability Scale for the assessment of electronic tools: development and validation. JMIR Hum Factors 2020 Dec 16;7(4):e21161. [doi: 10.2196/21161] [Medline: 33325828]
29. SUS Svensk: System Usability Scale in Swedish. Rosenfeld Media. 2011 Jul 4. URL: https://rosenfeldmedia.com/sus-svensk-system-usability-scale-in-swedish/ [accessed 2024-05-01]
30. Ministry of Health and Social Affairs, Swedish Association of Local Authorities and Regions. Vision for eHealth 2025 – common starting points for digitisation of social services and health care. e-Hälsa 2025. URL: https://ehalsa2025.se/wp-content/uploads/2021/02/vision-for-ehealth-2025.pdf [accessed 2024-10-01]
31. Mokkink LB, Prinsen C, Patrick DL, et al. COSMIN study design checklist for patient-reported outcome measurement instruments. BMJ. 2019 Jul. URL: https://gut.bmj.com/content/gutjnl/70/1/139/DC1/embed/inline-supplementary-material-1.pdf [accessed 2025-01-07]
32. Lewis JR. Measuring perceived usability: SUS, UMUX, and CSUQ ratings for four everyday products. Int J Hum-Comput Interact 2019 Sep 14;35(15):1404-1419. [doi: 10.1080/10447318.2018.1533152]
33. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine (Phila Pa 1976) 2000 Dec 15;25(24):3186-3191. [doi: 10.1097/00007632-200012150-00014] [Medline: 11124735]
34. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the System Usability Scale. Int J Hum Comput Interact 2008 Jul 29;24(6):574-594. [doi: 10.1080/10447310802205776]
35. Kristensson Hallström I, Castor C, Hansson H, et al. eHealth as an aid for facilitating and supporting self-management in families with long-term childhood illness; development, evaluation, and implementation in clinical practice – a research programme. Clin Health Promot 2018;13(1):e23002 [FREE Full text]
36. Kayser L, Karnoe A, Furstrand D, et al. A multidimensional tool based on the eHealth Literacy Framework: development and initial validity testing of the eHealth Literacy Questionnaire (eHLQ). J Med Internet Res 2018 Feb 12;20(2):e36. [doi: 10.2196/jmir.8371] [Medline: 29434011]
37. Hobart JC, Riazi A, Lamping DL, Fitzpatrick R, Thompson AJ. Improving the evaluation of therapeutic interventions in multiple sclerosis: development of a patient-based measure of outcome. Health Technol Assess 2004 Mar;8(9):iii, 1-iii,48. [doi: 10.3310/hta8090] [Medline: 14982653]
38. Huisman M. Imputation of missing item responses: some simple techniques. Qual Quant 2000;34(4):331-351. [doi: 10.1023/A:1004782230065]

XSL•FO
RenderX

39. Sjöström AE, Hajdarevic S, Hörnsten Å, Kristjánsdóttir Ó, Castor C, Isaksson U. The Swedish version of the eHealth Literacy Questionnaire: translation, cultural adaptation, and validation study. J Med Internet Res 2023 Apr 12;25(1):e43267. [doi: 10.2196/43267] [Medline: 37043268]
40. Abma IL, Rovers M, van der Wees PJ. Appraising convergent validity of patient-reported outcome measures in systematic reviews: constructing hypotheses and interpreting outcomes. BMC Res Notes 2016 Apr 19;9:226. [doi: 10.1186/s13104-016-2034-2] [Medline: 27094345]
41. Samuelsson M, Wennick A, Bengtsson M, Lydrup ML, Jakobsson J. Translation, cultural adaptation, and psychometric testing of the Supportive Care Needs Survey for partners and caregivers for Swedish family members of persons diagnosed with colorectal cancer. J Patient Rep Outcomes 2023 Oct 11;7(1):100. [doi: 10.1186/s41687-023-00636-1] [Medline: 37819416]
42. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human participants. JAMA 2024 Oct 19. [doi: 10.1001/jama.2024.21972] [Medline: 39425955]
43. Polit DF, Beck CT. Nursing Research: Generating and Assessing Evidence for Nursing Practice, th edition: Wolters Kluwer; 2021.
44. Kortum PT, Bangor A. Usability ratings for everyday products measured with the System Usability Scale. Int J Hum Comput Interact 2013 Jan;29(2):67-76. [doi: 10.1080/10447318.2012.681221]
45. Ho AD, Yu CC. Descriptive statistics for modern test score distributions: skewness, kurtosis, discreteness, and ceiling effects. Educ Psychol Meas 2015 Jun;75(3):365-388. [doi: 10.1177/0013164414548576] [Medline: 29795825]
46. Lindkvist RM, Sjöström-Strand A, Landgren K, Johnsson BA, Stenström P, Hallström IK. "In A Way We Took the Hospital Home"-a descriptive mixed-methods study of parents' usage and experiences of eHealth for self-management after hospital discharge due to pediatric surgery or preterm birth. Int J Environ Res Public Health 2021 Jun 15;18(12):6480. [doi: 10.3390/ijerph18126480] [Medline: 34203985]
47. Kirklies PC, Neumann O, Hohensinn L. Promoting digital equality in co-production: the role of platform design. Gov Inf Q 2024 Mar;41(1):101903. [doi: 10.1016/j.giq.2023.101903]
48. Frennert S. Hitting a moving target: digital transformation and welfare technology in Swedish municipal eldercare. Disabil Rehabil Assist Technol 2021 Jan;16(1):103-111. [doi: 10.1080/17483107.2019.1642393] [Medline: 31348681]
49. Lember V, Brandsen T, Tõnurist P. The potential impacts of digital technologies on co-production and co-creation. Pub Manag Rev 2019 Nov 2;21(11):1665-1686. [doi: 10.1080/14719037.2019.1619807]
50. van Velsen L, Evers M, Bara CD, Op den Akker H, Boerema S, Hermens H. Understanding the acceptance of an eHealth technology in the early stages of development: an end-user walkthrough approach and two case studies. JMIR Form Res 2018 Jun 15;2(1):e10474. [doi: 10.2196/10474] [Medline: 30684434]

Abbreviations

COSMIN: Consensus-Based Standards for the Selection of Health Status Measurement Instruments
eHLQ: eHealth Literacy Questionnaire
SUS: System Usability Scale
WHO: World Health Organization

Edited by A Kushniruk, E Borycki; submitted 13.07.24; peer-reviewed by JS Berner, J Brooke; revised version received 04.12.24; accepted 08.12.24; published 16.01.25.

Please cite as:
Åvik Persson H, Castor C, Andersson N, Hylén M
Swedish Version of the System Usability Scale: Translation, Adaption, and Psychometric Evaluation
JMIR Hum Factors 2025;12:e64210
URL: https://humanfactors.jmir.org/2025/1/e64210
doi:10.2196/64210

© Helene Åvik Persson, Charlotte Castor, Nilla Andersson, Mia Hylén. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 16.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluating the Development, Reliability, and Validation of the Tele-Primary Care Oral Health Clinical Information System Questionnaire: Cross-Sectional Questionnaire Study

Rosnah Sutan[1*], MPH, MD, PhD; Shahida Ismail[1*], MPH, MD; Roszita Ibrahim[1*], BSc, MSc, PhD

Department of Public Health Medicine, Faculty of Medicine, Universiti Kebangsaan Malaysia, Cheras Kuala Lumpur, Malaysia

[*]all authors contributed equally

Corresponding Author:
Rosnah Sutan, MPH, MD, PhD
Department of Public Health Medicine
Faculty of Medicine
Universiti Kebangsaan Malaysia
Jalan Yaakob Latiff, Bandar Tun Razak Cheras
Cheras Kuala Lumpur, 56000
Malaysia
Phone: 60 391459587
Email: rosnah.sutan@hctm.ukm.edu.my

Abstract

Background: Evaluating digital health service delivery in primary health care requires a validated questionnaire to comprehensively assess users' ability to implement tasks customized to the program's needs.

Objective: This study aimed to develop, test the reliability of, and validate the Tele-Primary Care Oral Health Clinical Information System (TPC-OHCIS) questionnaire for evaluating the implementation of maternal and child digital health information systems.

Methods: A cross-sectional study was conducted in 2 phases. The first phase focused on content item development and was validated by a group of 10 experts using the content validity index. The second phase was to assess its psychometric testing for reliability and validity.

Results: A structured questionnaire of 65 items was constructed to assess the TPC-OHCIS delivery for primary health care use based on literature and has been validated by 10 experts, and 319 respondents answered the 65-item TPC-OHCIS questionnaire, with mean item scores ranging from 1.99 (SD 0.67) to 2.85 (SD 1.019). The content validity, reliability, and face validity showed a scale-level content validity index of 0.90, scale-level content validation ratio of 0.90, and item-level face validity index of 0.76, respectively. The internal reliability was calculated as a Cronbach α value of 0.90, with an intraclass correlation coefficient of 0.91. Scales were determined by the scree plot with eigenvalues >1, and 13 subscales were identified based on principal component analysis. The Kaiser-Meyer-Olkin value was 0.90 (P<.049). The total variance explained was 76.07%, and factor loading scores for all variables were >0.7. The Bartlett test of sphericity, determining construct validity, was found to be significant (P<.049).

Conclusions: The TPC-OHCIS questionnaire is valid to be used at the primary health care level to evaluate the TPC-OHCIS implementation. It can assess health care workers' work performance and job acceptance and improve the quality of care.

(JMIR Hum Factors 2025;12:e53630) doi:10.2196/53630

KEYWORDS

telehealth; electronic health; eHealth; public health information system; psychometric analysis

Introduction

Background

Digital health is mostly used interchangeably with eHealth, telehealth, or mobile health in the literature [1,2]. It requires integrated and interdisciplinary sector involvement to use knowledge information and communication technology in health (eg, medicine, public health, pharmaceutical, dentistry, and health management) [1-8]. It enables the national health system to ensure population access to health services and the ability to monitor and evaluate health system delivery performance [1,2,9]. Electronic health records are created from digital health systems for case management monitoring and can be shared

XSL•FO
RenderX

across health care settings [9]. The data sharing platform allows access to integration interfaces that include electronic medical records, appointments, electronic prescriptions, e-commerce, public health surveillance, system monitoring such as vaccination, environmental health, institutional health management, and an online platform for teaching and learning among health care workers (HCWs) [10-16].

Malaysian Digital Health

Malaysia began integrating digital health into its health care system in 1998 [4,5]. The Telemedicine Act 1997 (Act 564) was enacted to regulate telemedicine practices in the country [4,5,17-25]. Later, the telemedicine blueprint was created to outline the government's vision for digital health implementation and align it with the 7 National Multimedia Super Corridor flagship applications [19-22]. The initiative was referred to as telemedicine and later restated as telehealth. The telehealth system serves as a platform for digital health services in Malaysia [20]. The National Telehealth Policy was launched to support the Vision 2020 agenda, focusing on four key components: (1) lifetime health plan, (2) mass customized health information and education, (3) continuous medical education, and (4) teleconsultation application [1]. The policy was formulated during the Eighth Malaysia Plan and managed by the National Health Informatics Centre Division [4,17,19-22]. The lifetime health plan covers health services from womb to tomb [19]. The telehealth system also includes the clinical support system and health information management and support services, and it encompasses the hospital information system and the Tele-Primary Care Oral Health Clinical Information System (TPC-OHCIS) [18-24]. The TPC-OHCIS is a comprehensive electronic medical record system for primary health care [22-24]. It was initially developed for outpatient services at primary health care clinics (PHCs), and later, it incorporated maternal and child health (MCH), oral health, and other life stage health services to ensure continuity of care [21,22]. The development of TPC-OHCIS was a collaborative effort between the Ministry of Science, Technology, and Innovation; the Ministry of Health, Malaysia; and the Malaysian Institute of Microelectronic Systems, Berhad [21-24]. The system was first tested in the PHCs of Seremban District, Negeri Sembilan, and later expanded to 34 PHCs across 3 states (ie, Perlis, Sarawak, and Selangor) [17,21]. To date, the TPC-OHCIS has been implemented in 108 PHCs across 7 additional states [22,25]. The TPC-OHCIS is a web-based platform that allows HCWs to enter data during clinic services or home visits, automatically updating when connected to the internet [23,24].

Role of Digital Health in Service Quality Performance

The recent COVID-19 pandemic revealed a significant public health issue in health care system delivery to provide comprehensive quality care [2,26-28]. Malaysia has experienced various health service delivery disruptions at PHCs during the critical phase of the COVID-19 pandemic [29]. MCH services include a wide range of services covering school health programs that experienced substantial disruptions due to closing and movement control orders, thus preventing mothers and children from receiving adequate health care services [22]. During the COVID-19 pandemic, many countries have improved health care delivery through digital technology, enhanced resource coordination, and facilitated universal health coverage [29,30].

The HCWs at the PHCs worked on the frontlines, assessing risks, monitoring care treatment, and promoting health empowerment in the community [17]. Most of the administration work related to patient care is recorded manually. Even with the implementation of TPC-OHCIS, many facilities still need to record data manually and enter it into the system because of internet instability at some PHCs [22]. At present, there is no specific policy published by the Ministry of Health to completely replace manual recording of patient information monitoring with digital health. The HCWs monitor patients during home visits and conduct outreach activities to cover areas inaccessible to health facilities [17,22]. Therefore, it is important to continue patient using manual data recording when services are provided offline. The TPC-OHCIS is an electronic medical record system used as a daily operating system and in "real time" at PHCs [22-24]. However, the TPC-OHCIS implemented in selected facilities in Malaysian PHCs is used mostly to record data and monitor patient care only [22-24].

Specific Study Measurement Tool

The conceptualization of this study diverges from earlier studies primarily in its integration of advanced technological modalities and emphasis on patient-centered care delivery, unlike traditional telehealth frameworks, which often focus on providing remote consultations or basic medical services [1,3,18,25]. The TPC-OHCIS incorporates elements of comprehensive primary care delivery, leveraging telehealth technologies to facilitate longitudinal patient-provider relationships, care coordination, and proactive health management. Moreover, this study's conceptualization places a heightened emphasis on the integration of patient health data, wearable devices, and digital health platforms to enhance care delivery and patient engagement. The TPC-OHCIS can remotely monitor patient health metrics, deliver personalized interventions, and empower individuals to take an active role in their health management [24]. The importance of interdisciplinary collaboration and team-based care is crucial in the implementation of the TPC-OHCIS at PHCs. Hence, it is important to assess the implementation of TPC-OHCIS by focusing on technology, organization, environment, or human resource components, as suggested in the literature [3-5,9]. Earlier studies were conducted to evaluate the effectiveness of the TPC-OHCIS in improving health care quality; however, there is limited evidence on assessing HCWs' perceived usefulness and ease of use [9,18,25]. A comprehensive study of HCWs' perspectives is important to provide a shred of extensive knowledge and evidence on the TPC-OHCIS implementation in PHCs.

Therefore, this study aimed at developing, testing reliability, and validating the TPC-OHCIS questionnaire, which was designed as a survey instrument to collect data on the implementation of the TPC-OHCIS at the PHC setting among HCWs related to various aspects of primary care services, including MCH services. It aims to ensure that the TPC-OHCIS questionnaire is a robust and effective tool for assessing the

XSL•FO
RenderX

implementation of MCH services, which are the core services for PHCs. The research question formulated was as follows: "Is a customized questionnaire on digital health information system (TPC-OHCIS) able to assess the HCWs' perception of its implementation process in the delivery of MCH services at the primary care level?" A thorough design of the TPC-OHCIS questionnaire validation helps increase the questionnaire's relevance and usefulness for decision-making purposes, as it is designed to facilitate monitoring and evaluating the TPC-OHCIS operability among the HCWs working at PHCs. To the best of our knowledge, there is no valid questionnaire available to measure the implementation evaluation of any digital health information system for the primary care level that focuses on MCH services, which is a priority service component of PHC and which was monitored regularly for sustainable development goals performance achievement [17].

Methods

Study Design

This study involved several steps for questionnaire development (phase 1), reliability, and validation (phase 2).

Phase 1: Questionnaire Development

The questionnaire is adapted from literature and document review [4,5,9,17-25,31-36]. The TPC-OHCIS questionnaire was developed based on various theoretical models, which may address multiple aspects of remote primary care and patient information systems. The questionnaire was created based on a combination of the technology-organization-environment (TOE) theory [31], the technology acceptance model (TAM) theory [32,33], the human organization technology-fit (HOT-fit) model [34,35], and the diffusion of innovation (DOI) theory [36]. There was a 65-item questionnaire with a 4-point Likert scale (1=highly disagree, 2=disagree, 3=agree, and 4=highly agree) developed based on the aforementioned theories [31-36]. The score scale is created according to the Likert scale that indicates the following: 1=strongly agree, 2=agree, 3=disagree, and 4=strongly disagree. In this study, the researcher did not put a neutral on a scale of 3 to avoid respondent bias [37]. The 4-point Likert scale does not impact the reliability and validity of the questionnaire [2,37].

The development of the questionnaire items was partly adapted from various questionnaires available from previous studies conducted in Malaysia, based on selected theoretical models, using a 4-point scale questionnaire: (1) technology, (2) organization, (3) environment, and (4) human [31-36] (Multimedia Appendix 1 [4,5,9,17-25,31-36]). Therefore, we classified the 65-item TPC-OHCIS questionnaire into 4 scales (Multimedia Appendix 2), described below.

- Domain A, technology: this comprised 17 items and four subscales that include (1) relative advantage (items 1-5), (2) compatibility (items 6-9), (3) complexity (items 10-13), and (4) security concern (items 14-17).
- Domain B, organization: this contained 18 items and four subscales that include (1) the presence of a specified liaison officer (BCHAMP: items 18-22), (2) infrastructure (BINFRA: items 23-26), (3) top management support (BTP: items 27-31), and (4) financial resources (BFIN: items 32-35).
- Domain C, environment: this focuses on vendor support (CVEN: items 36-39).
- Domain D, human: this contained 26 items and six subscales that include (1) staff competency in information systems (DPT: items 40-44), (2) knowledge of the TPC-OHCIS system (DEISK: items 45-50), (3) clinical information technology competency (DCIT: items 51-54), (4) perceived innovativeness of the IT officer (DCIO: items 55-57), (5) perceived ease use (DPEU: 58-61), and (6) perceived usefulness (DPU: items 62-65).

The 65-item TPC-OHCIS questionnaire was developed based on the requirement of the TPC-OHCIS implementation plan for the PHCs in Malaysia. A total of 15 items (item numbers 10, 11, 12, 13, 33, 40, 41, 42, 48, 49, 50, 53, 54, 58, and 39) were written as negative items. The Likert scale score of negative items was reversed for scoring analysis before data were entered into the SPSS software (version 26.0; IBM Corp).

Phase 2: Questionnaire Validation

On the basis of the steps suggested by Boateng et al [38], the second phase involved scale development, which consists of pretesting questions, sampling and survey administration, item reduction, and extraction of latent factors. The scale evaluation requires tests of dimensionality, reliability, and validity. A cross-sectional study was conducted to assess the psychometric properties of the questionnaire.

Study Sample

The sample for each step was calculated and listed in Table 1.

The sample-to-item ratio was decided based on the number of items in the questionnaire. The ratio per item was determined using 5 to 20 samples per item [39,40]. In this study, the survey included 65 items. Therefore, the total calculated sample size required was 325 respondents.

Table 1. Calculated sample size for each validation step.

Steps and aims	Sample size required, n	Reference
Pilot study and face validity: to test the adequacy of instrumentation in which the outcome is in the form of a scale	40	[3]
Test-retest reliability: to test the degree of consistency exhibited when a measurement is repeated under identical conditions	12	[4]
Field survey: to test the psychometric analysis of the questionnaire based on tests of dimensionality, reliability, and validity	325	[5]

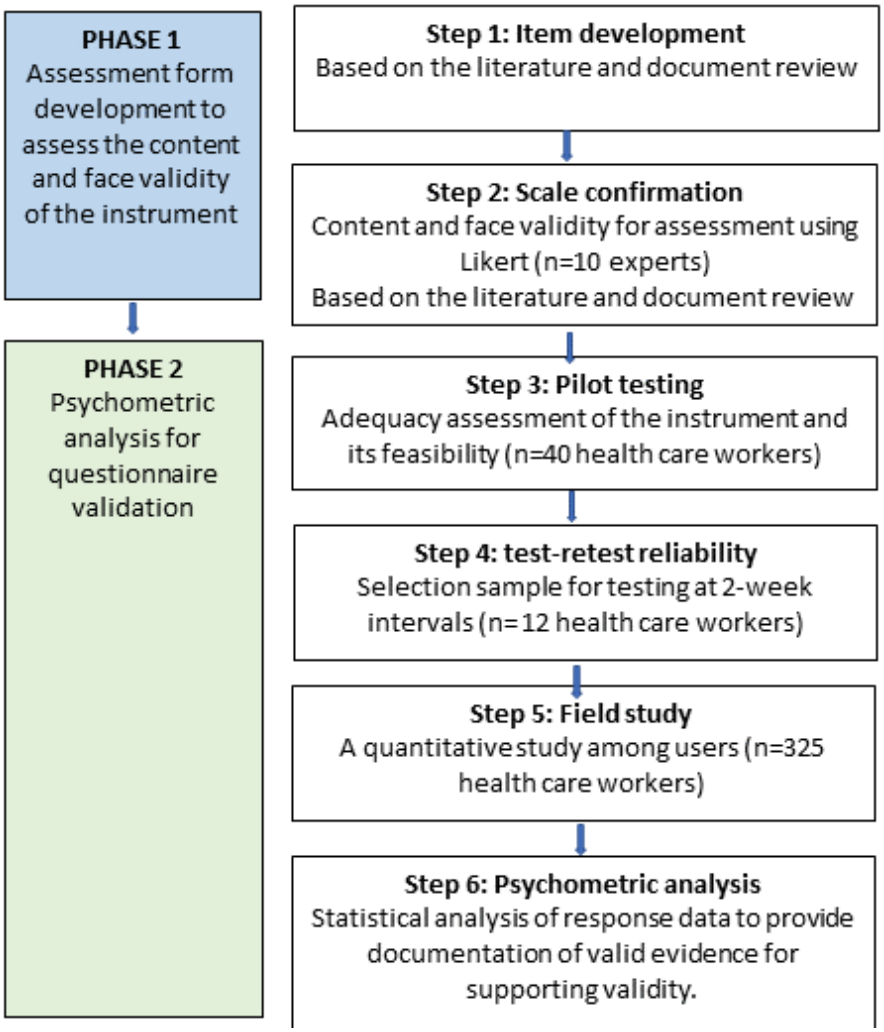
XSL•FO
RenderX

Content Validity Index

The flow process of questionnaire validation was conducted based on Figure 1.

On the basis of Figure 1, a total of 6 steps were taken to perform content validity for the questionnaire.

Figure 1. Flow process for questionnaire validation.



Step 1: Assessment Form Development for Content and Face Validity

In our literature review search [37,39-41], we listed items to determine the content and face validity. We used the clarity scale of 1 to 3 to indicate the relevance of each item (1=not clear, 2=revision needed, and 3=very clear). The essentiality of the questionnaire was identified using a 1 to 3 scale (1=not essential, 2=useful but not essential, and 3=essential). A similar form was used to assess face validity in evaluating the language, feasibility, readability, and style formatting consistency [37,41].

Step 2: Content and Face Validity (Expert Review)

Face validity was conducted among the experts who were involved in designing and using the system at the primary care level. Ten experts were invited to review the 65-item Tele-Primary Care Oral Health Clinical Information System (TPC-OHCIS) questionnaire. All experts have extensive experience and knowledge related to the TPC-OHCIS. Of the 10 experts, 3 (30%) were from the TPC-OHCIS division, 3 (30%) were public health researchers, 2 (20%) were from the maternal and child health unit services in the Selangor State Health Department, and finally, 2 (20%) were individuals who worked directly with the TPC-OHCIS system and also served as system liaison officers at health clinics. The adaptation of questionnaires from earlier studies [41-43] was checked for content suitability and appropriateness before forward translation was performed. Scale evaluation was performed by experts using a developed form. Owing to time constraints and lockdown because of the COVID-19 pandemic, a face-to-face method for collecting data from the experts was impossible. A Google form (Google LLC) link was created and sent via WhatsApp (Meta Platforms, Inc) to assess the face and content validity. The selected 10 experts were asked to evaluate the scale for content and face validation. All 10 experts' comments were used to improve the questionnaire. Experts were encouraged to provide both verbal and written comments regarding the questionnaire items. Verbal comments were collected via phone call, and written comments were collected via the Google form.

XSL•FO
RenderX

Step 3: Pilot Study

A pilot study was conducted to assess the adequacy of the instrumentation, which measured outcomes to ensure that the instruments used were appropriate and effective for the intended purpose. Through the pilot study, face validity was evaluated to confirm that the 65-item TPC-OHCIS questionnaire developed was constructed accurately and comprehensively. We purposively invited health care workers (HCWs) who were involved in the initial pilot-tested version of the TPC-OHCIS in Seremban District, Negeri Sembilan [17]. The questionnaire was sent after a thorough face and content validity assessment by the experts. Correction and fine-tuning were conducted by the researchers based on their feedback.

Step 4: Test-Retest Reliability

We selected 12 HCWs who were familiar with the TPC-OHCIS involved in the pilot study mentioned above to test the degree of consistency exhibited when a measurement is repeated under identical conditions. The same questionnaire was sent to the same respondents 2 weeks after the first time the survey was conducted.

Step 5: Field Study

A cross-sectional study was performed to assess the item reduction and extraction factors among the HCWs of the TPC-OHCIS public health care clinics (PHCs) in selected states in Malaysia. The TPC-OHCIS is a system used in the PHCs. The PHC is headed by a family medicine specialist who supervises staff from various disciplines (eg, physicians, nurses, medical assistants, pharmacists, laboratory technicians, and record officers).

Step 6: Performing an Analysis for Content Validity, Face Validity, and Test-Retest Reliability

The item-level content validity index (I-CVI) and scale-level content validity index (S-CVI) were used. I-CVI is defined as the proportion of content experts giving items a relevance rating of 3 and 4. S-CVI was calculated by the proportion of understanding by experts (universal agreement [UA] of S-CVI), which gave ratings of 3 and 4, or by average scores given (average of S-CVI). The average of S-CVI is defined as the average of the I-CVI scores for all items on the scale or the average of proportion relevance judged by all experts. The UA of S-CVI is the proportion of items on the scale that achieve a relevance scale of 3 or 4 by all experts. UA score is given as 1 when the item achieved 100% agreement by all experts; otherwise, the UA score is given as 0 [39]. First, we calculated the experts in agreement. Second, a score of 1 was given to items that achieved 100% agreement by all experts, while a score of 0 was given to items that did not achieve 100% agreement. Third, experts in the agreement were divided by the total number of experts (calculated for I-CVI). The fourth step was to divide the results obtained in the third step by 65, which was the total number of items in the questionnaire (calculated for average of S-CVI). Finally, we divided the UA result from step 2 by 65 items (calculated for UA of S-CVI). The face validity index was used to evaluate the form of clarity and comprehensibility of language and instructions used in the questionnaire [41]. The respondents were requested to rate the comprehensibility of each item using a scale of 1 to 4 (1=not understandable, 2=somewhat understandable, 3=understandable, and 4=very understandable). The item-level face validity index was computed for each item by dichotomizing the 4-point scale, where items scoring either 1 or 2 were recorded as 0 and items scoring 3 and 4 were recorded as 1. The values later were added up according to each item, and the total values were divided by the total number of respondents. Test-retest reliability was used to define whether the questionnaire was answered by respondents due to chance. The internal reliability of questionnaires was assessed using the Cronbach α coefficient. The Pearson correlation and intraclass correlation coefficients of the scores of the 2 tests were calculated. The test-retest reliability is achieved when the value of intraclass correlations is 0.6 to 0.8 (good reliability), and values >0.8 indicate excellent reliability, which means the higher the correlation, the higher the test-retest reliability, with values close to 0 indicating low reliability [38].

Forward-Backward Translation

The questionnaire was translated from English to the Malay language, then underwent content assessment, and its cross-cultural validity was evaluated by 3 experts representing Malaysia's diverse ethnic backgrounds (Malay, Chinese, and Indian), proficient in both English and Malay. After the questionnaire had been reviewed, the forward-backward translation was sent to Proofreaders United company for translation and proofreading.

Ethical Considerations

Ethics approval was granted by the National University of Malaysia (FF-2021-124) and the Ministry of Health Malaysia National Medical Research Registry (NMRR-21-599-58521, investigatory initiated research). Meanwhile, the researcher also needed to obtain verbal permission and a signed consent letter from the respondent as evidence of consent to participate in the study. Respondents were informed that their participation was voluntary and they would not receive any payment. All research information collected was treated as confidential. The data are displayed anonymously without the name, address, or any identity that describes the respondent when presented as the study output.

Statistical Analysis

The internal reliability was assessed using Cronbach α. The validity index was calculated based on the content validity index (CVI). The Bartlett test of sphericity was used to determine the construct validity. The Kaiser-Meyer-Olkin measure and total variance explained (TVE) were evaluated. Factor analysis was performed using varimax rotation in principal component analysis (PCA) to verify construct validity (ie, discriminant and convergent validity). Items loading >0.60 were considered for further analysis [38].

Results

CVI Evaluation

Out of 10 experts evaluating the 65-item TPC-OHCIS questionnaire (TPC-OHCIS questionnaire in Multimedia

XSL•FO
RenderX

Appendix 2), the scale-level CVI (S-CVI) average for item 1 of domain A, relative advantage, to item 65 of domain D, perceived usefulness, was 0.9. The average of S-CVI (based on item-level CVI) was 7.415, the average of S-CVI (based on proportion) was 0.742, and the universal agreement of S-CVI was 0.045. The universal agreement of S-CVI was calculated based on the proportion of items on the scale that obtained a relevance value of 3 or 4 from all the experts (Multimedia Appendix 3).

Face Validation

The face validity index was assessed for clarity, language comprehensibility, and instructions used in the questionnaire. The calculated item face validity index was 0.785 and fulfilled the face validity criteria, which was >0.6 [41].

Test-Retest Reliability

The 65-item survey's internal reliability score for Cronbach α was 0.90 for the whole instrument and exceeded the suggested minimum value of Cronbach α=0.70 [38]. Pearson correlation coefficient was 0.90, and the intraclass correlation coefficient was 0.91.

Descriptive Analysis

There were 319 respondents (Table 2) who answered the 65-item TPC-OHCIS questionnaire, with mean item scores ranging from 1.99 to 2.85 (Table 3). Our study found that a high mean score was achieved in all subscales except for the item 17 in domain A, security concerns (*how much can I trust the vendor)*. The subscale mean scores were lowest at 1.93 (SD 0.751) and 1.99 (SD 0.832) for item 63 in domain D, perceived usefulness (*TPC-OHCIS application improved the quality of my work, respectively).* The mean score for communication was the highest for item 38 in domain C, vendor, with a mean score of 3.45 (SD 0.819), and item 39 in domain C, vendor, with a mean score of 3.39 (SD 0.805).

Table 2. Respondent profile for field study (N=319).

Variables and category	Frequency, n (%)
Occupation	
Family medicine specialist	5 (1.6)
Medical officer	69 (21.6)
Matron	2 (0.6)
Head nurse	17 (5.3)
Nurse	138 (43.3)
Community nurse	88 (27.6)
Sex	
Male	6 (1.9)
Female	313 (98.1)
Race	
Chinese	7 (2.2)
Indian	21 (6.6)
Malay	285 (89)
Others	6 (1.9)

XSL•FO
RenderX

Table 3. Descriptive statistics for item mean score of the Tele-Primary Care Oral Health Clinical Information System (TPC-OHCIS) questionnaire.

Item question	Question statement	Scores, mean (SD)
ARA[a] 1	Using the TPC-OHCIS application enables me to do my work quickly.	2.05 (0.865)
ARA 2	Using the TPC-OHCIS application improves the quality of my work.	1.99 (0.832)
ARA 3	Using the TPC-OHCIS application enhances my effectiveness on the job.	2.01 (0.864)
ARA 4	Using the TPC-OHCIS application increases my productivity.	2.04 (0.825)
ARA 5	Using the TPC-OHCIS application makes my job easier.	2.02 (0.795)
ACOM[b] 6	TPC-OHCIS application can be easily accessed across multiple platforms (laboratory results, x-ray, and other related patient data).	2.53 (0.941)
ACOM 7	TPC-OHCIS user interfaces provide transparent access to all platforms (e-notification and VEKPRO[c]).	2.62 (0.931)
ACOM 8	Data received from other devices (tablet/laptop/ smartphone) outside health facilities in the TPC-OHCIS application can be easily merged into the database for analysis.	2.74 (0.932)
ACOM 9	Information is shared seamlessly across our organization regardless of location.	2.74 (0.839)
ACOMPLEX[d] 10	I do not know enough about the TPC-OHCIS application to handle my job satisfactorily.	2.86 (0.870)
ACOMPLEX 11	I need a long time to understand and get familiar with the TPC-OHCIS application.	2.82 (0.961)
ACOMPLEX 12	I do not find enough time to study and upgrade my technology skills before using the TPC-OHCIS system.	2.71 (0.870)
ACOMPLEX 13	I often find it too complex for me to understand and use the TPC-OHCIS application.	2.99 (1.019)
ASEC[e] 14	I feel secure in using the TPC-OHCIS application, keying in patients' data, and sharing it across my organization.	2.05 (0.855)
ASEC 15	I would feel safe using the TPC-OHCIS application to retrieve patient data.	2.80 (0.754)
ASEC 16	I am concerned about data patient leakage.	2.07 (0.785)
ASEC 17	I am concerned about how much I can trust the vendor.	1.93 (0.751)
BCHAMP[f] 18	A specified liaison officer will provide useful information to top managers and vendors about the TPC-OHCIS application faulty.	2.02 (0.736)
BCHAMP 19	A specified liaison officer plays a role in upgrading the TPC-OHCIS application for users' needs.	2.01 (0.760)
BCHAMP 20	A specified liaison officer has a good relationship with both vendors and top managers.	2.22 (0.851)
BCHAMP 21	A specified liaison officer can bring staff to use the TPC-OHCIS application well.	2.41 (0.972)
BCHAMP 22	A specified liaison officer provides training/courses for the users a few times a year.	2.54 (0.895)
BINFRA[g] 23	We have enough computers for the TPC-OHCIS application use.	2.27 (0.743)
BINFRA 24	We have a reliable computer network in our use.	2.11 (0.767)
BINFRA 25	Appropriate hardware, software, and network infrastructure were in place before TPC-OHCIS implementation.	1.90 (0.749)
BINFRA 26	We have integrated IS[h] applications encompassing different functional areas (laboratory and pharmacy).	2.05 (0.775)
BTP[i] 27	Top management always supports and encourages the use of the TPC-OHCIS application for job-related tasks.	2.15 (0.785)
BTP 28	Top management provides most of the necessary help and resources to enable people to use the TPC-OHCIS system.	2.17 (0.746)
BTP 29	Top management provides good access to hardware when staff need it.	2.15 (0.732)
BTP 30	Top management gives feedback to vendors on every dismayed or unsatisfactory comment from staff.	2.44 (0.822)

XSL•FO
RenderX

Item question	Question statement	Scores, mean (SD)
BTP 31	Top management provides good access to the TPC-OHCIS application when staff need it.	2.55 (0.842)
BFIN[j] 32	There is enough financial aid from the organization for the coordination of the system implementation.	2.44 (0.901)
BFIN 33	I find difficulties in using the TPC-OHCIS application because it cannot be upgraded due to not having enough budget.	2.75 (0.814)
BFIN 34	Enough computers are available to access the TPC-OHCIS application.	2.29 (0.743)
BFIN 35	We easily obtain obsolete computer replacements.	2.28 (0.729)
CVEN[k] 36	Vendors entertain each of our complaints dutifully.	2.41 (0.704)
CVEN 37	The vendor can upgrade the TPC-OHCIS according to our needs.	2.41 (0.771)
CVEN 38	The system vendor attends technical meetings quite frequently.	3.45 (0.819)
CVEN 39	I have a platform to voice out problems regarding the TPC-OHCIS application directly to the vendor.	3.39 (0.805)
DPTC[l] 40	I do not know how to use computers.	3.19 (0.874)
DPTC 41	I never used to work online.	2.11 (0.849)
DPTC 42	I need people's help to use a computer.	2.15 (0.859)
DPTC 43	I like to work using the online system.	2.29 (0.836)
DPTC 44	The TPC-OHCIS application is easy to operate.	2.33 (0.794)
DEISK[m] 45	I have enough training before working with the TPC-OHCIS application.	2.25 (0.836)
DEISK 46	It took me only a few days before I could master the TPC-OHCIS application well.	2.85 (0.920)
DEISK 47	The TPC-OHCIS facilitates task management.	2.63 (0.912)
DEISK 48	The TPC-OHCIS application is hard to use.	2.61 (0.900)
DEISK 49	I have to open many interfaces just to key in 1 patient's data.	2.05 (0.730)
DEISK 50	The TPC-OHCIS application takes much time because I have to open so many interfaces.	2.50 (0.715)
DCIT[n] 51	I have confidence in my ability to operate the TPC-OHCIS application.	2.60 (0.725)
DCIT 52	I have the expertise regarding Information technology to provide valuable knowledge to the organization.	2.23 (0.737)
DCIT 53	It does not make any difference whether I add/share knowledge with others related to the use of the TPC-OHCIS application.	2.22 (0.688)
DCIT 54	I feel that other employees can provide more valuable knowledge about the system's use.	2.22 (0.688)
DCIO[o] 55	The ITO[p] is actively considering the introduction of new technology to solve to organization's problem.	2.34 (0.672)
DCIO 56	The ITO tries to keep a technological leading edge by adopting new technology early.	2.84 (0.823)
DCIO 57	The ITO tends to take risks in the decision-making of new technology introduction.	2.80 (0.847)
DPEU[q] 58	I often become confused every time I use the TPC-OHCIS application.	2.11 (0.810)
DPEU 59	Interacting with the TPC-OHCIS application is frequently frustrating.	2.18 (0.768)
DPEU 60	I find that the TPC-OHCIS application makes my job easier.	2.28 (0.846)
DPEU 61	The TPC-OHCIS application provides useful guidance in performing tasks.	2.06 (0.764)
DPU[r] 62	My job would be hard to perform without the TPC-OHCIS application.	2.13 (0.835)
DPU 63	Using the TPC-OHCIS application improves my job performance.	2.18 (0.767)
DPU 64	Using the TPC-OHCIS application saves me time.	2.13 (0.851)

XSL•FO
RenderX

Item question	Question statement	Scores, mean (SD)
DPU 65	Using the TPC-OHCIS application supports critical aspects of my job (e.g, retrieving patients with missed treatment).	2.17 (0.782)

[a]ARA: domain A, relative advantage.
[b]ACOM: domain A, compatibility.
[c]VEKPRO: vector program for reporting vector-borne diseases and outbreaks.
[d]ACOMPLEX: domain A, complexity.
[e]ASEC: domain A, security.
[f]BICHAMP: domain B, champion.
[g]BINFRA: domain B, infrastructure.
[h]IS: information system.
[i]BTP: domain B, top management support.
[j]BFIN: domain B, financial support.
[k]CVEN: domain C, vendor.
[l]DPTC: domain D, perceived technical competence.
[m]DEISK: domain D, employee information system knowledge.
[n]DCIT: domain D, competency of employee's IT.
[o]DCIO: domain D, chief information officer innovativeness.
[p]ITO: IT officer.
[q]DPEU: domain D, perceived ease of use.
[r]DPU: domain D, perceived usefulness.

Statistical Analysis

The 65-item TPC-OHCIS questionnaire was created based on a mix of items adapted from various sources in the development of the questionnaire. It was adapted by a combination of the TOE theory [31], the HOT-fit model [34,35], the DOI theory [36], and the TAM theory [32,33]. The content validity was assessed by experts according to the 4 scales (ie, technology, organization, environment, and human), as highlighted in the previous studies [33-35]. In this study, we explored the constructs generated from this initial group of 65 items, which were developed based on their multidimensional nature. An exploration of the subdimensions based on the 65-item TPC-OHCIS questionnaire at the initial phase found that there were 13 factors, without being restricted to 4 domains. Exploratory factor analysis (EFA) determined the quantity of components (or themes) that emerged for the questionnaire items. The process involved grouping the measurements of comparable themes. The scree plot reveals that there were 13 components in a 65-item questionnaire (Figure 2).

A PCA with varimax rotation yielded 13 components with an eigenvalue of >1. The eigenvalue for the first component was 17.71 and accounted for 27% of the variance. The difference between the first and second components was 11.98, while the subsequent eigenvalues were small (4.91, 4.64, and 3.40). The Kaiser-Meyer-Olkin value was 0.908, which showed sample adequacy. The Bartlett test of sphericity value for the approximate chi-square was χ^2_{2080}=18,219.9; this was significant (P=.001).

Table 4 shows that the eigenvalues from 1.01 to 17.71, corresponding to the extraction sums of squared loadings of 13 subscales of the questionnaire. The highest eigenvalue accounted for 17.71% of the total variance, while the lowest eigenvalue explained 1.01%. The result also shows that the extraction eigenvalue does not vary much from the rotation eigenvalue. Therefore, 65 items loaded strongly on 13 factors at the extraction level.

Table S1 in Multimedia Appendix 4 provides the scale construct and its item. All 13 components had acceptable internal reliability (Cronbach α>0.7). All construct eigenvalues were >1. The deleted items were 14 and 15 of domain A, relative advantage, and 56 and 57 of domain D, chief information officer innovativeness, and the deletions were done because their factor loadings were <0.6. After discussing with the experts, questions for items 56 and 57 of domain D, chief information officer innovativeness, were rephrased for clarity. Question for item 15 of domain A, relative advantage, was omitted, and question for item 14 of the same domain remained as it was.

Table S1 in Multimedia Appendix 4 provides a detailed breakdown of various items and their corresponding domain scores, reflecting their relative contribution to each domain. The items are identified by numbers and grouped under distinct domains. For instance, items 18, 19, 22, 20, 21, 16, 17, and 14 display scores ranging from 0.493 to 0.854, indicating their relevance within a particular domain. Similarly, items 63, 65, 62, 64, 61, 60, 59, and 58 exhibit scores from 0.636 to 0.807 and, therefore, were grouped in another domain. The third domain was categorized based on items 40, 42, 44, 43, 41, 38, and 39, with scores ranging from 0.701 to 0.927. The fourth domain includes items 48, 46, 49, 47, 45, and 50, with scores between 0.0715 and 0.829, and the last domain covers items 27, 31, 30, 29, and 28, with scores ranging from 0.773 to 0.822.

XSL•FO
RenderX

Figure 2. Scree plot revealing the 13-factor components extracted from principal component analysis with varimax rotation.

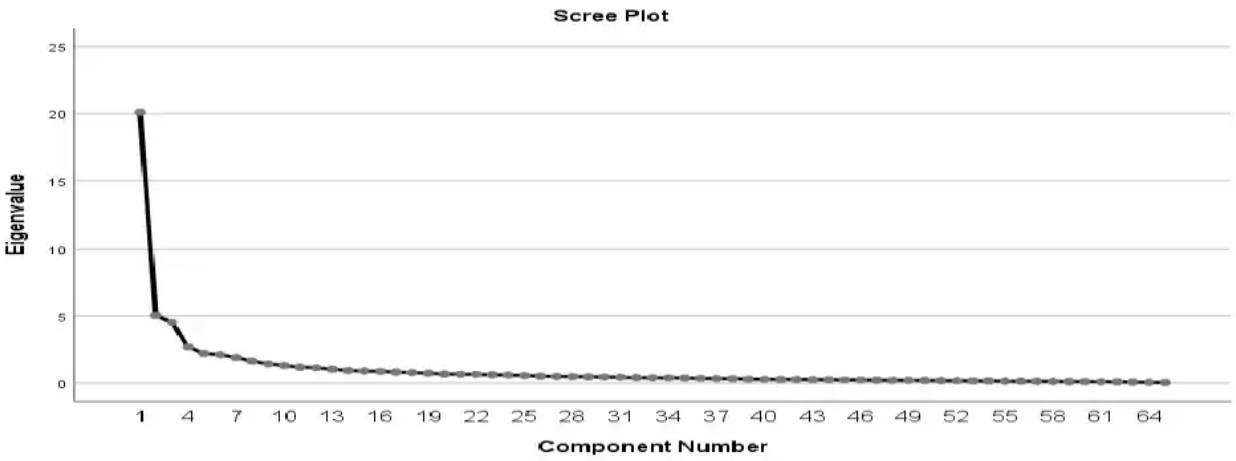


Table 4. The total variance explained.

Factor	Initial eigenvalues			Extraction sums of squared loadings		
	Total, n	Variance (%)	Cumulative (%)	Total, n	Variance (%)	Cumulative (%)
1	18	27.2	27.2	18	27.2	27.2
2	6	8.8	36.1	6	8.8	36.1
3	5	7.6	43.6	5	7.6	43.6
4	5	7.1	50.7	5	7.1	50.7
5	3	5.2	55.9	3	5.2	55.9
6	3	4.3	60.3	3	4.3	60.3
7	2	2.9	63.2	2	2.9	63.2
8	2	2.7	65.9	2	2.7	65.9
9	2	2.6	68.6	2	2.6	68.6
10	1	2.2	70.7	1	2.2	70.7
11	1	2.1	72.8	1	2.1	72.8
12	1	1.7	74.5	1	1.7	74.5
13	1	1.6	76.1	1	1.6	76.1

Discussion

Principal Findings

This study aimed to validate a specific tool used in health care delivery in Malaysia's health care system, a new technology telehealth application called TPC-OHCIS, making this evaluation timely. The 65-item TPC-OHCIS questionnaire demonstrated a comprehensive and reliable tool ready to be used for the assessment of the implementation of the digital application system for monitoring the health service program at the PHCs. Monitoring any health care service performance requires a periodic and ongoing operation to ensure that it can be delivered as planned and with good tracking of progress. Evaluation of how the program should be implemented needs to have a valid tool that can give relevant and appropriate results. A validated tool can be used to replicate similar studies in the future for evaluation and comparison.

Developing and validating a locally customized questionnaire for specific program service delivery is important. Therefore, this study used 4 theories, including the HOT-fit model, DOI theory, TOE theory, and TAM theory. The TPC-OHCIS questionnaire was proven valid based on content, face, and construct validities, as well as good psychometric analysis reliability. A total of 4 scales created using PCA did not align with a local study [17] that assessed the hospital information system implementation in Malaysian hospitals. Measuring implementation of technology in different service set-ups and levels of care, at either hospitals or PHCs, should use a customized tool. No similar research paper is available to check for disagreement or benchmarks on other tools, to compare them with the TPC-OHCIS questionnaire validation. Our findings showed that the 13 scales were validated based on the local population working in the PHC with the TPC-OHCIS. In comparison, an earlier study conducted local testing on the hospital information system and identified 4 scales [18]. Our

XSL•FO
RenderX

analysis focuses on the MCH services, which are the routine and regularly monitoring health services delivered at PHCs.

A CVI score of 0.90 is generally considered very high and indicative of strong content validity for a measurement tool, particularly in the academic context. Content validity refers to the extent to which the items in a measurement instrument represent the entire range of content that the instrument is intended to measure [39]. A CVI of 0.90 provides strong evidence that the measurement tool has undergone rigorous development and validation processes.

Face validity has been widely used not only in questionnaire development but also in questionnaires that were adopted from other researchers' tools, which required local forward-backward translation [37-41]. Experts in the field who were recruited to determine face and content validity in our study gave feedback regarding content relevance, wording, quantity of items, the amount of information, and other related issues. The analysis results for the CVI from 10 experts indicated good quality (>0.9), which complies with the recommendations in the literature [41].

Internal reliability assessment using Cronbach α revealed how close the selected items are to one another when measuring the construct [39]. The Cronbach α value of the 65-item TPC-OHCIS questionnaire was 0.90, which exceeded the suggested minimum value of 0.70 [39]. By measuring Cronbach α, the tool's accuracy and reliability have been verified. Our findings demonstrated that the variables are correlated with the components as constructed through PCA, and it proved that they are internally consistent.

The EFA also determined the TVE for the construct. The TVE illustrated the precision with which the measuring objects and their constituent parts estimate the construct. This study found that the TVE was 76.08%, which exceeded the required minimum of 60%, indicating that the overall variance explained was satisfactory [39]. PCA is the most common EFA method for dimension reduction. The EFA results using varimax rotation with maximum likelihood showed that the TPC-OHCIS questionnaire has 13 subscales.

The study may provide evidence-based recommendations for policy makers and health care stakeholders regarding the implementation and optimization of the TPC-OHCIS in the MCH service delivery monitoring. By identifying key factors influencing the TPC-OHCIS implementation and effectiveness, the study offers actionable insights into policy development, resource allocation, and quality improvement efforts aimed at strengthening MCH care systems. The study may contribute methodological innovations to the field of TPC-OHCIS evaluation, such as novel approaches to questionnaire development, validation, and implementation. By documenting the methodological processes and challenges encountered in assessing the TPC-OHCIS questionnaire validation, the study adds to the methodological toolkit available to researchers and practitioners working in the field of digital health.

Overall, the TPC-OHCIS study offers novel insights into the evaluation of the TPC-OHCIS, which focuses on the MCH services, highlighting the importance of tailored assessment tools, cross-cultural adaptation, policy considerations, and methodological innovations. These insights contribute to a deeper understanding of the complexities and opportunities associated with the integration of digital health technologies into health care delivery and provide a foundation for future research and practice in this area. The TPC-OHCIS questionnaire is validated for evaluating other health information systems, as it takes into account the outcome aspects of health services, technology, organizations, vendors, and human resources. Future researchers can replicate studies to validate the findings of the TPC-OHCIS questionnaire in different populations, settings, and contexts. By replicating studies using a valid and reliable tool with diverse samples, researchers can assess the generalizability and robustness of the questionnaire across various demographic, cultural, and organizational contexts.

Suggestions to Stakeholders

There is a need to integrate machine learning and artificial intelligence into the TPC-OHCIS so that data can be extracted faster and can be used in line with precision medicine. The validated TPC-OHCIS questionnaire can help stakeholders make effective evidence-based decisions in managing patient-centered care. The system needs to be more stable, easier to use, and faster in terms of data collection, without the need for complicated training, and easy to download from the system itself. Awareness related to the importance of complete and accurate data from an organizational perspective needs to be understood at the level of clinic users and data administrators because they will then enter, process, and ensure that the data entered are complete and accurate. These crucial issues were addressed in our study during the content and face validity assessment of the tool.

The COVID-19 pandemic made us see the need for the TPC-OHCIS to be rolled out to wider PHCs in Malaysia. With <10% of the TPC-OHCIS currently implemented, the vision to having a digital health system in Malaysia by the year 2030 needs to be facilitated. The TPC-OHCIS should be developed as a comprehensive package that includes embedded training and health education to facilitate its delivery. The issue of data security can be compared to the concept of banking system apps in Malaysia, where users only need to download the app on their mobile phones to share information with other health care providers, unlike the TPC-OHCIS, which ensures continuity of care. Otherwise, printing a duplicate report is sufficient for the time being while the country works on transitioning all clinics into a digital system.

Study Strength

The study demonstrates the development of a tailored questionnaire specifically designed to assess the implementation process and the effectiveness of the TPC-OHCIS in the context of MCH health services at the primary care level. This tailored approach acknowledges the unique challenges and complexities of MCH service delivery and highlights the importance of designing measurement tools that are sensitive to the needs of this population. The study determined comprehensive items to measure the integration implementation of the TPC-OHCIS, which is one of the digital health technologies focusing on

XSL•FO
RenderX

priority services, such as MCH service delivery at the primary care level. The study sheds light on the practical considerations, barriers, and facilitators associated with the adoption and use of digital health solutions in MCH care settings. The study offers insights into the cross-cultural adaptability assessment tools, particularly in diverse multiethnicity population settings. By evaluating the linguistic and cultural appropriateness of the TPC-OHCIS questionnaire, the study highlights the importance of considering cultural context and language preferences in developing and validating measurement instruments for MCH services.

Limitations of This Study

This study was conducted in 2021 when movement control orders were implemented in Malaysia. The study was conducted entirely online, and all surveys were distributed to the respondents' superiors without proper briefing about the survey and the overall idea of the research except for study information highlighted at the beginning of the online survey questionnaire form. Therefore, each respondent answered as they saw fit, without recognizing the true meaning of the survey. In addition, the deployment of staff during the COVID-19 pandemic could have caused fatigue among them, which was likely to cause them to answer the questions lightly without focus. Meanwhile, the distribution of questionnaires through Google Forms contributed to the risk of bias among respondents when there was no supervision or explanation that could be given as a guide to answering the questionnaire.

Conclusions

The psychometric validation process of the questionnaire was done comprehensively for all iterative stages, including initial reliability tests, potential modifications based on these tests, and integration with existing digital health assessment frameworks. The novelty of the study lies in its unique approach to conceptualizing the implementation of the TPC-OHCIS. This may offer new insights into the integration of life stage records, with a focus on priority services at PHCs, such as MCH services, for its development. Cross-cultural adaptability was considered to ensure wider applicability, and the rigor of the process was demonstrated in part by high CVI scores, which is important in an academic research setting. Likert scale points were carefully chosen to capture nuanced responses, and while the TPC-OHCIS has certain academic limitations, its robustness paves the way for future research on policy implications and telehealth. Comparison with existing literature establishes its validity and reliability, showing its potential impact on policy development.

Acknowledgments

The authors would like to especially thank the Selangor Health State Department, Kuala Lumpur and Putrajaya Health State Department, and Pahang Health State Department for their support in the study. The authors would also like to thank the Director General of Health Malaysia for his permission to conduct the study and publish this paper. This work was supported by a research grant from the Ministry of Higher Education (FRGS/1/2021/SKK04/UKM/02/1) and by the Medical Faculty of Universiti Kebangsaan Malaysia, which contributed to the publication costs of this paper.

Authors' Contributions

SI, RS, and RI conceived the scope of the review to meet the study objectives. Literature search, data collection, analysis, and interpretation were conducted by SI and verified by RS. SI drafted the first iteration of the manuscript, which was revised and improvised by RS. All authors made substantial contributions to the critical review, editing, and revision of the manuscript. All authors approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Evidence table.
[PDF File (Adobe PDF File), 151 KB - humanfactors_v12i1e53630_app1.pdf]

Multimedia Appendix 2
Tele-Primary Care Oral Health Clinical Information System questionnaire.
[DOCX File , 42 KB - humanfactors_v12i1e53630_app2.docx]

Multimedia Appendix 3
Expert review analysis.
[XLSX File (Microsoft Excel File), 22 KB - humanfactors_v12i1e53630_app3.xlsx]

Multimedia Appendix 4
Domain construct according to item question.
[DOCX File , 25 KB - humanfactors_v12i1e53630_app4.docx]

XSL•FO
RenderX

References

1. Yang K, Hu Y, Qi H. Digital health literacy: bibliometric analysis. J Med Internet Res 2022 Jul 06;24(7):e35816 [FREE Full text] [doi: 10.2196/35816] [Medline: 35793141]
2. Marcolino MS, Diniz CS, Chagas BA, Mendes MS, Prates R, Pagano A, et al. Synchronous teleconsultation and monitoring service targeting COVID-19: leveraging insights for postpandemic health care. JMIR Med Inform 2022 Dec 22;10(12):e37591 [FREE Full text] [doi: 10.2196/37591] [Medline: 36191175]
3. Abd Ghani MK, Mohamed MA, Mostafa SA, Mustapha A, Aman H, Jaber MM. The design of flexible telemedicine framework for healthcare big data. Int J Eng Technol 2018;7:461-468. [doi: 10.14419/ijet.v7i3.30.19096]
4. Abidi SS, Goh A, Yusoff Z. Telemedicine and medical informatics in the multimedia super corridor: the Malaysian vision. Stud Health Technol Inform 1998;52 Pt 2:1282-1286. [Medline: 10384666]
5. Mudaris IS. Digitalised health records: does Malaysia need it? Khazanah Research Institute. 2023 Mar 27. URL: https://www.krinstitute.org/Views-@-Digitalised_Health_Records-;_Does_Malaysia_Need_It%5E.aspx [accessed 2024-05-24]
6. Cresswell K, Williams R, Sheikh A. Developing and applying a formative evaluation framework for health information technology implementations: qualitative investigation. J Med Internet Res 2020 Jun 10;22(6):e15068 [FREE Full text] [doi: 10.2196/15068] [Medline: 32519968]
7. Bhattacharya AA, Umar N, Audu A, Felix H, Allen E, Schellenberg JR, et al. Quality of routine facility data for monitoring priority maternal and newborn indicators in DHIS2: a case study from Gombe State, Nigeria. PLoS One 2019 Jan 25;14(1):e0211265 [FREE Full text] [doi: 10.1371/journal.pone.0211265] [Medline: 30682130]
8. Al-Shorbaji N. Improving healthcare access through digital health: the use of information and communication technologies. In: Agrawal A, Kosgi S, editors. Healthcare Access. London, UK: IntechOpen; 2022.
9. Ibrahim AA, Ahmad Zamzuri M, Ismail R, Ariffin AH, Ismail A, Muhamad Hasani MH, et al. The role of electronic medical records in improving health care quality: a quasi-experimental study. Medicine (Baltimore) 2022 Jul 29;101(30):e29627 [FREE Full text] [doi: 10.1097/MD.0000000000029627] [Medline: 35905245]
10. Fryer K, Delgado A, Foti T, Reid CN, Marshall J. Implementation of obstetric telehealth during COVID-19 and beyond. Matern Child Health J 2020 Sep;24(9):1104-1110 [FREE Full text] [doi: 10.1007/s10995-020-02967-7] [Medline: 32564248]
11. Yuen J, Pike S, Khachikyan S, Nallasamy S. Telehealth in ophthalmology. In: Linwood SL, editor. Digital Health. Brisbane, Australia: Exon Publications; 2022.
12. Samiei V, Wan Puteh SE, Abdul Manaf MR, Abdul Latip K, Ismail A. Are Malaysian diabetic patients ready to use the new generation of health care service delivery? A telehealth interest assessment. Malays J Med Sci 2016 Mar;23(2):44-52 [FREE Full text] [Medline: 27547114]
13. Akiyama H, Hasegawa Y. A trial case of medical treatment for primary headache using telemedicine. Medicine (Baltimore) 2018 Mar;97(9):e9891 [FREE Full text] [doi: 10.1097/MD.0000000000009891] [Medline: 29489688]
14. Dehnavieh R, Haghdoost A, Khosravi A, Hoseinabadi F, Rahimi H, Poursheikhali A, et al. The District Health Information System (DHIS2): a literature review and meta-synthesis of its strengths and operational challenges based on the experiences of 11 countries. Health Inf Manag 2019 May 13;48(2):62-75. [doi: 10.1177/1833358318777713] [Medline: 29898604]
15. Faujdar DS, Sahay S, Singh T, Jindal H, Kumar R. Public health information systems for primary health care in India: a situational analysis study. J Family Med Prim Care 2019 Nov;8(11):3640-3646. [doi: 10.4103/jfmpc.jfmpc_808_19] [Medline: 31803666]
16. Beheshti L, Kalankesh LR, Doshmangir L, Farahbakhsh M. Telehealth in primary health care: a scoping review of the literature. Perspect Health Inf Manag 2022 Jan 1;19(1):1n [FREE Full text] [Medline: 35440933]
17. World Health Organization. Regional Office for the Western Pacific. Malaysia health system review. World Health Organization Regional Office for the Western Pacific. 2012. URL: https://iris.who.int/handle/10665/206911 [accessed 2024-05-24]
18. Marzuki NM, Ismail S, Abdul Mohsein NA, Ehsan FZ. Evaluation of telehealth implementation in government primary health clinics - a study protocol. In: Proceedings of the International Conference on Green and Ubiquitous Technology. 2012 Presented at: GUT 2012; July 7-8, 2012; Bandung, Indonesia. [doi: 10.1109/gut.2012.6344169]
19. Malaysia's telemedicine blueprint: leading healthcare into the information age. Ministry of Health Malaysia. 1997 Jul 25. URL: https://www.moh.gov.my/moh/resources/auto%20download%20images/5ca1b20928065.pdf [accessed 2024-04-24]
20. Mohan J, Razali Raja Yaacob R. The Malaysian Telehealth Flagship Application: a national approach to health data protection and utilisation and consumer rights. Int J Med Inform 2004 Mar 31;73(3):217-227. [doi: 10.1016/j.ijmedinf.2003.11.023] [Medline: 15066550]
21. Maon SN, Edirippulige S. An overview of the national telehealth initiative in Malaysia. Stud Health Technol Inform 2010;161:95-103. [Medline: 21191162]
22. Teleprimary Care – Oral Health Clinical Information System (TPC-OHCIS). Bahagian Pembangunan Kesihatan Keluarga. URL: https://hq.moh.gov.my/bpkk/index.php/sistem-tpc-ohcis/sistem-tpc-ohcis [accessed 2024-04-24]
23. Teleprimary care and oral health clinical information system. MIMOS. 2020. URL: https://www.mimos.my/wp-content/uploads/2020/09/TPC-OHCIS.pdf [accessed 2024-04-24]
24. Teleprimary Care and Oral Health Clinical Information System (TPC-OHCIS). MIMOS. URL: https://www.mimos.my/main/teleprimary-care-and-oral-health-clinical-information-system-TPC-OHCIS [accessed 2024-04-24]

XSL•FO
RenderX

25. Marzuki N, Ismail S, Al-Sadat N, Ehsan FZ, Chan CK, Ng CW. Integrating information and communication technology for health information system strengthening: a policy analysis. Asia Pac J Public Health 2015 Nov;27(8 Suppl):86S-93S. [doi: 10.1177/1010539515590180] [Medline: 26085477]
26. Aziz A, Zork N, Aubey JJ, Baptiste CD, D'Alton ME, Emeruwa UN, et al. Telehealth for high-risk pregnancies in the setting of the COVID-19 pandemic. Am J Perinatol 2020 Jun;37(8):800-808 [FREE Full text] [doi: 10.1055/s-0040-1712121] [Medline: 32396948]
27. Wu H, Sun W, Huang X, Yu S, Wang H, Bi X, et al. Online antenatal care during the COVID-19 pandemic: opportunities and challenges. J Med Internet Res 2020 Jul 22;22(7):e19916 [FREE Full text] [doi: 10.2196/19916] [Medline: 32658860]
28. Parajuli R, Bohara D, Kc M, Shanmuganathan S, Mistry SK, Yadav UN. Challenges and opportunities for implementing digital health interventions in Nepal: a rapid review. Front Digit Health 2022;4:861019 [FREE Full text] [doi: 10.3389/fdgth.2022.861019] [Medline: 36120714]
29. Faizal AM, Muhammad Azrai A, Abdul Kadir AK. The impact of COVID-19 pandemic on oncofertility services in advanced reproductive centre UKM Medical Centre Malaysia. Asian Pac J Cancer Care 2020 Jul 01;5(S1):165-166. [doi: 10.31557/apjcc.2020.5.s1.165-166]
30. Lindberg MH, Venkateswaran M, Abu Khader K, Awwad T, Ghanem B, Hijaz T, et al. eRegTime, efficiency of health information management using an electronic registry for maternal and child health: protocol for a time-motion study in a cluster randomized trial. JMIR Res Protoc 2019 Aug 07;8(8):e13653 [FREE Full text] [doi: 10.2196/13653] [Medline: 31392962]
31. Khobi JA, Mtebe JS, Mbelwa JT. Factors influencing district health information system usage in Sierra Leone: a study using the technology-organization-environment framework. Electron J Inf Syst Dev Ctries 2020 Jun 08;86(6):e12140. [doi: 10.1002/isd2.12140]
32. Alsharo M, Alnsour Y, Alabdallah M. How habit affects continuous use: evidence from Jordan's national health information system. Inform Health Soc Care 2020 Jan;45(1):43-56. [doi: 10.1080/17538157.2018.1540423] [Medline: 30457025]
33. Salleh MI, Abdullah R, Zakaria N. Evaluating the effects of electronic health records system adoption on the performance of Malaysian health care providers. BMC Med Inform Decis Mak 2021 Feb 25;21(1):75 [FREE Full text] [doi: 10.1186/s12911-021-01447-4] [Medline: 33632216]
34. Ahmadi H, Nilashi M, Ibrahim O. Organizational decision to adopt hospital information system: an empirical investigation in the case of Malaysian public hospitals. Int J Med Inform 2015 Mar;84(3):166-188. [doi: 10.1016/j.ijmedinf.2014.12.004] [Medline: 25612792]
35. Ahmadi H, Nilashi M, Shahmoradi L, Ibrahim O, Sadoughi F, Alizadeh M, et al. The moderating effect of hospital size on inter and intra-organizational factors of Hospital Information System adoption. Technol Forecast Soc Change 2018 Sep;134:124-149. [doi: 10.1016/j.techfore.2018.05.021]
36. Zhang X, Yu P, Yan J, Ton A M Spil I. Using diffusion of innovation theory to understand the factors impacting patient acceptance and use of consumer e-health innovations: a case study in a primary care clinic. BMC Health Serv Res 2015 Feb 21;15(1):71 [FREE Full text] [doi: 10.1186/s12913-015-0726-2] [Medline: 25885110]
37. Taherdoost H. What is the best response scale for survey and questionnaire design; review of different lengths of Rating Scale / Attitude Scale / Likert Scale. Int J Acad Res Manag 2019;8(1):1-10 [FREE Full text] [doi: 10.4135/9781412963947.n273]
38. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL. Best practices for developing and validating scales for health, social, and behavioral research: a primer. Front Public Health 2018;6:149 [FREE Full text] [doi: 10.3389/fpubh.2018.00149] [Medline: 29942800]
39. Taherdoost H. Validity and reliability of the research instrument; how to test the validation of a questionnaire/survey in a research. SSRN 2016 Aug 10 [FREE Full text] [doi: 10.2139/ssrn.3205040]
40. Suhr D. Exploratory or confirmatory factor analysis? In: Proceedings of the 31st Annual SAS Users Group International Conference. 2006 Presented at: SUGI 2006; March 26-29, 2006; San Francisco, CA.
41. Yusoff MS. ABC of response process validation and face validity index calculation. Educ Med J 2019 Oct 31;11(3):55-61. [doi: 10.21315/eimj2019.11.3.6]
42. Yusoff MS. Construct validity, internal consistency and normative data of the USMaP-i in a sample of medical students. Int Med J 2013 Feb;20(1):1-7 [FREE Full text]
43. Svedbo Engström M, Leksell J, Johansson UB, Eeg-Olofsson K, Borg S, Palaszewski B, et al. A disease-specific questionnaire for measuring patient-reported outcomes and experiences in the Swedish National Diabetes Register: development and evaluation of content validity, face validity, and test-retest reliability. Patient Educ Couns 2018 Jan;101(1):139-146. [doi: 10.1016/j.pec.2017.07.016] [Medline: 28736071]

Abbreviations

CVI: content validity index
DOI: diffusion of innovation
EFA: exploratory factor analysis

XSL•FO
RenderX

HCW: health care worker
HOT-fit: human organization technology-fit
MCH: maternal and child health
PCA: principal component analysis
PHC: primary health care clinic
S-CVI: scale-level content validity index
TAM: technology acceptance model
TOE: technology-organization-environment
TPC-OHCIS: Tele-Primary Care Oral Health Clinical Information System
TVE: total variance explained

Edited by A Choudhury; submitted 12.10.23; peer-reviewed by F Medina, Y Chu, D Chrimes; comments to author 15.03.24; revised version received 14.05.24; accepted 25.11.24; published 29.01.25.

Please cite as:
Sutan R, Ismail S, Ibrahim R
Evaluating the Development, Reliability, and Validation of the Tele-Primary Care Oral Health Clinical Information System Questionnaire: Cross-Sectional Questionnaire Study
JMIR Hum Factors 2025;12:e53630
URL: https://humanfactors.jmir.org/2025/1/e53630
doi:10.2196/53630
PMID:

©Rosnah Sutan, Shahida Ismail, Roszita Ibrahim. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 29.01.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Creation of Text Vignettes Based on Patient-Reported Data to Facilitate a Better Understanding of the Patient Perspective: Design Study

Sofia Charlotte Kelly[1*], MSc; Emilie Kauffeldt Wegener[1*], MA, MSc; Lars Kayser[1*], MD, PhD

Department of Public Health, University of Copenhagen, Copenhagen, Denmark

[*]all authors contributed equally

Corresponding Author:
Sofia Charlotte Kelly, MSc
Department of Public Health
University of Copenhagen
Øster Farimagsgade 5
Copenhagen, 1353
Denmark
Phone: 45 61722415
Email: sofiakelly95@gmail.com

Abstract

Background: Patient-reported outcome (PRO) data refer to information systematically reported by patients, or on behalf of patients, without the influence of health care professionals. It is a focal point of the health care system's ambition toward becoming more involving and personalized. It is recognized that PROs provide valuable data. However, despite this recognition, there are challenges related to both patients' and clinicians' accurate interpretations of the quantitative data. To overcome these challenges, this study explores text vignettes as a representation of PROs.

Objective: This study aimed to develop data-informed text vignettes based on data from the Readiness and Enablement Index for Health Technology (READHY) instrument as another way of representing PRO data and to examine how these are perceived as understandable and relevant for both patients and clinicians.

Methods: The text vignettes were created from participant responses to the READHY instrument, which encompasses health literacy, health education, and eHealth literacy. The text vignettes were created from 13 individual text strings, each corresponding to a scale in the READHY instrument. This study consisted of 3 sequential parts. In part 1, individuals with chronic obstructive pulmonary disease completed the READHY instrument, providing data to be used to create vignettes based on cluster profiles from the READHY instrument. Part 2 focused on the development of scale-based strings representing all READHY dimensions, which were evaluated through iterative cognitive interviews. In part 3, clinicians and patients assessed the understanding and relevance of the text vignettes.

Results: Clinicians and patients both understood and related to the text vignettes. Patients viewed the text vignettes as an accurate reflection of their PRO responses, and clinicians perceived the text vignettes as aligned with their understanding of patients' experiences.

Conclusions: Text vignettes can be developed using PRO instruments, with individual scales as input strings. This provides an opportunity to present numeric values in a text format that is understandable and recognizable to most patients and clinicians. Challenges with the vignette's language and layout require customization and clinician training to ensure meaningful interpretation. Findings also support the need to expand the study and enhance clinical relevance with alternative or contextually relevant text vignettes.

(JMIR Hum Factors 2025;12:e58077) doi:10.2196/58077

KEYWORDS

patient-reported outcome; text vignette; data representation; Readiness and Enablement Index for Health Technology; understanding; health care system; data analysis; clinical training; clinician; health professional

XSL•FO
RenderX

Introduction

Background

Patient-reported outcome (PRO) data represent patients' self-assessed health status in terms of physical and mental health and do not involve health professionals [1]. PRO data can be used as a screening tool before outpatient visits as a supplement to in-person or virtual consultations, serving as a stratification tool to reduce the number of visits needed. PRO data can also be used during consultations to screen for symptoms or serve as a dialogue and decision tool. The use of PRO data during consultations may increase the patient's self-awareness and health literacy, facilitate joint decision-making, enable self-management, and lead to better health outcomes [1,2].

Health Literacy

Competence in health literacy and digital knowledge is fundamental for an accurate understanding of PRO data as well as knowledge about health behavior and health practices and how patients understand and interpret health information [3]. Challenges arise when data accessibility does not consider differences in health literacy. While aggregate-level understanding may tolerate some variations, individual patient pathways require a common understanding to capture nuances [4,5]. Therefore, it is crucial to account for patients' health literacy and cognition in the clinical use of PROs.

Interpretation of PRO Data

When incorporating PROs into individual patient pathways, it becomes crucial to ensure that patients fully understand the questions, fostering a shared understanding. In working with PRO data, an assumption is often made that people have the same needs, preferences, perceptions, and experiences and will attach similar importance to items in the questionnaire, thus leaving no room to capture different additions of value or uncover nuances [4]. The quantitative part of PRO data operates within a set of assumptions about the world, which states that the standardized numerical score is the objective representation of the respondents' health status and health-related quality of life, independent of an individual's unique life story [4].

For both clinicians and patients, it may be difficult to understand the meaning of PRO data when they are presented as values. Patients often prefer visual representations, such as bar charts and line charts, which prove useful for quick comprehension and comparisons over time [6]. However, the accuracy of interpretation is closely tied to the individuals' prior knowledge of these graphical representations [6].

Graphical presentations may be a challenge. Grossman et al [7] highlighted that individuals lacking prior knowledge of bar charts struggled to interpret them accurately. Furthermore, health literacy plays a crucial role in correct interpretation, with studies suggesting that patients with lower health literacy may find it challenging to grasp the longitudinal nature of PRO data [8]. There is a notable correlation between educational level and accurate interpretation, along with age-related difficulties in correctly interpreting graphic representations of PRO data [6,9]. Patients generally, regardless of educational level, prefer less-complex presentations of PRO data, but a correlation has been demonstrated between age and accurate interpretation, where patients aged >65 years had substantially more difficulty interpreting graphic representations of PRO data correctly [6,9]. In Denmark, 1 in 5 individuals is aged >65 years [10], which means that a relatively large proportion of all individuals will potentially have difficulty interpreting PRO data.

The use of traffic light colors, that is, green, yellow, and red, in representing PRO data adds another layer of complexity. While this method increases accuracy in interpretation across various graphical representations [6,11], it is not foolproof and 93% of patients in a study population were able to correctly interpret the meaning of traffic light colors [6]. Despite this visual aid, there remains a need for additional descriptive explanations to enhance understanding [9].

Patients' and health care professionals' perceived understanding of graphically presented PRO data often exceeds their actual comprehension [9]. In addition, studies indicate that there may be a discrepancy between patients' preferred form of visualization and the form of visualization they can interpret with the highest accuracy [9].

Patients may also view the use of PRO data in treatment as impersonal and nonbeneficial for the relationship between themselves and their clinicians because they experience that the focus shifts from their own perspective to quantitative goals [12]. This perspective highlights the significance of maintaining a patient-centered approach.

A Common Understanding of PROs

When patients and clinicians collaborate on data interpretation, potential errors may arise due to questionnaire misinterpretation. This divergence in understanding creates a need for a common language or understanding of the representations of colors, figures, or text. A challenge is that patients and clinicians possess distinct cognitive frameworks influencing their actions and comprehension [13]. Data interpretation involves understanding available data, extracting information, and supplementing it for comprehensive comprehension [13]. Clinicians are shaped by a positivist background and may differ from patients in cognitive approaches, impacting data interpretation. When PRO data are translated into quantitative targets, it may be unclear what data have been made available and what information can be extracted. Patients provide a wide range of data, but data visualization often only includes part of these data.

Despite psychometric validation, there is a need to address discordance in individual interpretations of PRO data. Discordance reasons include word interpretation mismatches, evolving patient circumstances, and differing expectations [14]. Awareness of these differences is crucial in interpreting PRO data [14]. PRO tools vary in complexity, leading to further potential errors if patients lack an accurate understanding. A study found that only 11% of 59 PRO questionnaires are readable by the average UK adult [15], emphasizing the need for a patient-clinician dialogue to enhance understanding and decision-making [5]. Numerical scores derived from PROs can represent a person's health status or health-related quality of life to a limited extent because some of the narrative is lost

XSL•FO
RenderX

along the way. In the representation of PRO data, there is a need to be able to give meaning to individual experiences [4]. Positivism argues that knowledge should be based on the objective world via empirically verified causal explanations based on "positive" facts and not on abstract inferences; however, according to Meadow [4], we cannot understand the patient if we do not also try to understand subjective and emotional meaning.

PROs as Text Vignettes

One approach to a more understandable and personalized representation of data about a group of patients or individuals is vignettes created based on a combination of clustered data, which are enriched with sociodemographic characteristics and interviews with individuals representing these clusters. An example of this is the Optimizing Health Literacy and Access (OPHELIA) process, where issues related to health care access and engagement are identified and addressed [16,17]. This way of creating data-informed text vignettes is a powerful tool to provide clinicians and policy makers with insights into a more detailed description of archetypes of patients [18]. In this way, text vignettes play a pivotal role in providing a nuanced understanding of the perceived status of various subgroups of patients in a specific context or area. The primary aim is to bring to life the "person behind the numbers," referring to PRO scores, addressing the challenge of visualizing the individuals behind the statistical data. To the best of our knowledge, this has not yet been done at the level of individuals by creating person-specific text vignettes.

In this study, inspired by the OPHELIA process, we will explore whether text vignettes can be developed from more structured text inputs and how this work can be advanced to create text vignettes that can be used at the individual level, rather than at the group level, to present personalized texts that reflect each patient's perception of their condition. The study uses the Readiness and Enablement Index for Health Technology (READHY) [19], which covers 3 key themes—self-management, social capital, and digital health literacy. This instrument was selected because, as its developers, we have extensive experience with its performance, and its design makes it well suited as a PRO tool for telehealth services. The study demonstrates the process of converting psychometric instrument scales into data-informed text vignettes, highlighting the challenges and implications to consider for ensuring understanding and relevance for both patients and clinicians.

Objectives

This study aimed to develop data-informed text vignettes based on data from the READHY instrument as another way of representing PRO data and to examine how these are perceived as understandable and relevant for both patients and clinicians.

Part 1 Introduction: Generating Vignettes Based on READHY Clusters Enriched With Sociodemographic Data

This part is a proof of concept, where cluster-based profiles were created to help develop representatives of characteristics of a patient population divided into subgroups. This enabled us to work with materials based on real-life data varying with respect to the 13 scale scores. This has provided insights into how dimension names can be phrased for text vignettes without a systematic approach, while also helping us establish a shared understanding of the OPHELIA [16] principles. These principles were operationalized in a data-informed manner, bridging the previously reported qualitative approach with a more quantitative-based methodology.

Part 2 Introduction: Creation of Text Strings and Testing of Meaning in Vignettes

In part 2, scale-based text strings describing the different dimensions measured in READHY were developed through a series of iterations. These were then combined into comprehensive text vignettes profiling the participants. For this process, the upper threshold was adjusted from 2.50 to 2.70 based on data from the study by Kayser et al [20]. The vignettes were continuously refined through cognitive interviews to evaluate and improve the participants' understanding and ability to comprehend the content.

Part 3 Introduction: Applicability of Text Vignettes in a Clinical Setting

The text strings iteratively created and validated in part 2 served as the foundation for the work in part 3. In part 3, the focus shifted to testing the applicability of individual text vignettes within a clinical setting. The emphasis was on how meaningful and relevant these text vignettes were in a setting involving both patients and clinicians within a telehealth service, where the READHY instrument was suitable as a PRO instrument. The clinical setting included here was the same as the one used to gather data for part 1.

Methods

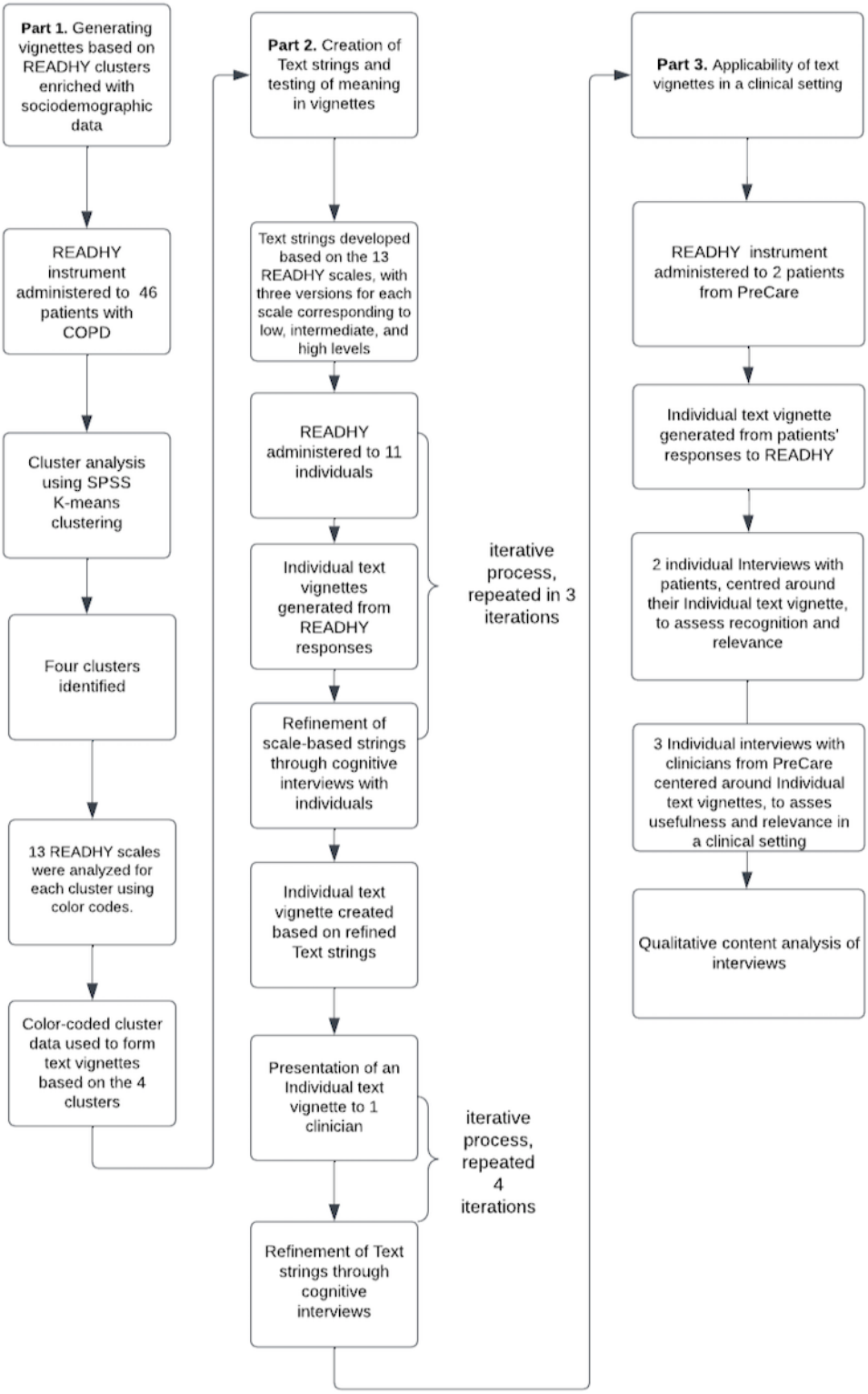
Overview

The study comprises 3 parts, each building upon the previous one. In part 1, which was inspired by the OPHELIA process [16], the READHY instrument was administered to individuals living with chronic obstructive pulmonary disease (COPD). Data gathered from this group of participants served as the foundation for generating 4 clusters, which were characterized by high, intermediate, and low levels of the 13 READHY scales, which together with another Health Literacy Questionnaire (HLQ) scale, "actively managing my health and sociodemographics", were used to develop a proof of concept of how data-informed text vignettes could be used to illustrate the quantitative data qualitatively to better understand profiles of user groups. The text vignettes in part 2 were crafted as narratives, emphasizing the group's high or low scores. This approach created profiles that highlighted their resources and barriers. In part 2, we expanded on the idea by constructing scale-specific strings for each of the 3 levels for each scale in READHY. The text strings were created to constitute a list of variables, which in a later process could be combined based on an algorithm to form the text vignettes. This computer-based combination of the text strings is beyond the scope of this study that only addresses the development of our concept up to a technology readiness level of 3 [21]. Text vignettes are consequently created by manually combining the text strings.

XSL•FO
RenderX

In this process, the text strings were evaluated for clarity and understanding through multiple iterations of cognitive interviews with a diverse set of participants. In part 3, the refined scale-based strings were combined manually, informed by the 2 patients' READHY scores, into individual text vignettes, which were presented to clinicians and patients for assessment of their understanding, relevance, and recognizability (Figure 1).

Figure 1. Flowchart: the 3-part process for examining the use of text vignettes. COPD: chronic obstructive pulmonary disease; READHY: Readiness and Enablement Index for Health Technology.



Setting and Population

The setting for the reported activities was in part 1 and part 3 of a telehealth service, PreCare [22], which provides 24-7 access to a response and coordination center manned with registered nurses and backed up by an e-doctor. PreCare is organized based on the Epital care model [22,23]. The participants in part 2 were

XSL•FO
RenderX

recruited broadly among individuals in the eastern part of Denmark and clinicians not currently working in telehealth.

Participants

Part 1

The participants were 46 adults living with COPD, with a mean age of 74.4 (SD 6.8; range 59-91) years. They were patients recruited by convenience sampling from the European Commission funded project Smart Inclusive Living Environments, were enrolled in the PreCare project in Region Zealand, and were diagnosed with different severity levels of COPD [24]. The inclusion criteria were a COPD diagnosis and being active in the PreCare project. The patients should be able to understand and communicate in Danish.

Part 2

Individuals

A total of 11 individuals were recruited using a snowball technique in the Capital Region of Denmark and Region Zealand. Initially, 3 members of interest were identified and approached within their local areas. These individuals were then asked to refer others in their community who might be willing to participate. With their consent, we received the necessary contact information and subsequently reached out to the referred participants by phone to arrange their involvement in the study. The inclusion criteria for individuals' participation was that they had to be aged >18 years, able to understand Danish, capable of giving informed consent, and able to complete the READHY instrument, with assistance if needed (ie, the ability to read and write independently was not a criterion). The first 3 recruited individuals resided in the metropolitan area, had medium-length higher education (International Standard Classification of Education level 6-7) [25], and had an average age of 53 (SD 12.36) years. This group was expanded to include 8 participants, aged between 36 and 86 years, from the capital region, Lolland, and Jutland with educational backgrounds ranging from International Standard Classification of Education level 4 to 6 [25] to attain a more diverse and comprehensive perspective.

Clinicians

A total of 4 clinicians were also recruited using a snowballing technique. Initially, 3 clinicians were approached, but not all expressed interest in participating. These contacts referred to others, and with their consent, we obtained their contact information and contacted them. The clinicians who participated in the interviews were 1 registered nurse, 2 physicians, and 1 professional with a master's and PhD degree in biomedicine. Of the 4 physicians, 2 (50%) had prior knowledge of READHY as active researchers using the instrument. All physicians were involved in professional settings related to digital health in patient care, and they were all familiar with PROs.

Part 3

A total of 3 registered nurses and 2 patients from the PreCare clinic were interviewed about the understanding, recognizability, and relevance of individual text vignettes in a clinical setting.

The 2 patients, a woman aged 75 years and a man aged 80 years, lived in the region where the PreCare project is based. They were selected by PreCare staff based on the following inclusion criteria: a COPD diagnosis, active participation in PreCare, the ability to understand and communicate in Danish, clinic enrollment at the time of data collection, and prior consent through PreCare as well as informed consent specifically for this study. A total of 3 (75%) of the 4 nurses in PreCare were available for interviews.

Application of the READHY Instrument

READHY is a validated PRO instrument that combines several aspects of PROs to illuminate experienced support, digital health literacy, and coping with everyday life in relation to one's health [19]. The tool focuses on aspects of knowledge and skills related to health; self-care, disease, and health-related mindset; experiences with health technology and the individual understanding of these; as well as the extent to which users feel supported by relatives and health care professionals [19].

READHY has been used to identify different health technology readiness profiles within different patient groups, such as patients living with type 2 diabetes and patients with cancer [26,27].

READHY consists of 65 items comprising 13 dimensions or scales. The 13 scales are selected from 3 PRO instruments. Four scales are from the Health Education Impact Questionnaire (heiQ), which evaluates the effect of health education interventions on self-care [28]. These address aspects of self-management with relevance for a digital health context, such as insights and attitudes. Two scales from the HLQ, which assesses health literacy, including support from relatives, peers, and health professionals [29], and all 7 scales from the eHealth Literacy Questionnaire (eHLQ), which address patients' digital health literacy, were included [30]. The eHLQ informs about knowledge, skills, motivation, perceptions, and experiences in relation to digital health. The 13 dimensions of READHY, which correspond to its 13 scales, are presented in Textbox 1.

XSL•FO
RenderX

Textbox 1. The 13 dimensions of the Readiness and Enablement Index for Health Technology framework.

- Health Education Impact Questionnaire (heiQ) 3: self-monitoring and insight
- heiQ4: constructive attitudes and approaches
- heiQ5: skill and technique acquisition
- heiQ8: emotional distress
- Health Literacy Questionnaire (HLQ) 1: feel understood and supported by health care providers
- HLQ4: social support for health
- eHealth Literacy Questionnaire (eHLQ) 1: using technology to process health information
- eHLQ2: understanding of health concepts and language
- eHLQ3: ability to actively engage with digital services
- eHLQ4: feel safe and in control
- eHLQ5: motivated to engage with digital services
- eHLQ6: access to digital services that work
- eHLQ7: digital services that suit individual needs

The READHY instrument uses a 4-point response scale for all items in the instrument [19], where 1 represents strongly disagree, 2 represents disagree, 3 represents agree, and 4 represents strongly agree. The results are 13 scale scores. For the scale heiQ8 (emotional distress), which includes 6 items, the scale is reversed by subtracting the scored value from 5. This adjustment simplifies interpretation, as higher scores now correspond to less impact.

Ethical Considerations

Written informed consent was obtained from all participants after providing information in accordance with the Helsinki II Declaration, including their right to opt out. Participants did not receive any salary or reimbursement of costs. Health science questionnaire surveys and interview studies that do not involve human biological material (section 14(2) of the Danish Act on Committees) do not require reporting or approval from the Danish National Center for Ethics [31].

Part 1 Methods

Data Sampling

The READHY instrument was handed out and collected with a combination of mail using prestamped envelopes by EKW and a researcher in PreCare in May 2021. A total number of 120 surveys including the READHY instrument were distributed.

Cluster Analysis

In managing the READHY data from the 46 patients, clusters were created using the K-means clustering function in IBM's software package SPSS Statistics (version 24) [32]. Three models were created with 3, 4, and 5 clusters. On the basis of an evaluation of which cluster model would be most meaningful from a clinical perspective, a 4-cluster model was selected. The final cluster constituted 41 valid cases and 5 missing cases out of the 46 responses. The number of cases in each cluster was as follows: 16 in cluster 1, 5 in cluster 2, 11 in cluster 3, and 9 in cluster 4.

The cluster data were then used to form the group-based text vignettes. For each cluster, the 13 scales were initially labeled with a color indicating whether the level was low, intermediate, or high. We used arbitrary thresholds informed by a regional survey, Sundhedsprofilen 2021 [33], which is the region where the PreCare project is based. All values <2.00 were assigned red, all values between 2.01 and 2.50 were assigned yellow, and values >2.50 were given green. This helped us create text strings that were specific for each of the colors indicating high or low average scores for each scale and assigning them to the cluster text vignette. The text strings were created from the dimension name and informed by the items in the scale. Additional data from an additional HLQ scale, "actively managing my health," and sociodemographic variables, such as educational level, experience living with, and number of chronic conditions, were added to the profiles to enrich them and create personas that were recognizable to the clinicians.

Part 2 Methods

Overview

The overall process for constructing the text vignettes involved several steps. First, text strings were developed based on the 13 READHY scales, with 3 versions for each scale corresponding to low, intermediate, and high levels. Next, participants completed the instrument. The average score for each scale was then manually calculated from their responses, and this average score was matched with the corresponding text string for that scale score. Finally, the 13 text strings were combined into an individual text vignette. The text strings were combined in a coherent order, prioritizing readability and clarity, rather than strictly following the numerical order of the scales. In part 2, the focus was only on the READHY scales and text vignettes; therefore, part 2 did not include sociodemographic information, as these are anticipated to be retrieved from other data sources such as electronic health records as they are specific at the individual level.

XSL•FO
RenderX

Constructing Data-Informed Text Vignettes

The text vignettes were created by combining 13 text strings into a cohesive PRO profile. Each sentence in the text vignette corresponds to a specific scale in READHY, describing the knowledge, skills, and attitudes that the scale items aim to capture.

Initially, 39 text strings were constructed—3 for each of the 13 READHY scales, representing different levels of knowledge, skills, or attitudes within each scale. These initial text strings were closely aligned with the scale names, as the scale names summarized the experiences the items were designed to encompass. For example, the scale "self-monitoring and insight" (heiQ3) assesses an individual's ability to monitor their health and their insight into their health [19]. An initial sentence for this scale might read as follows: "Has good insight into and control over his/her health and has a constructive attitude and approach to it."

The levels are determined based on the average score for each scale, with 3 categories: level 1 indicates problematic knowledge, skills, and attitudes; level 2 corresponds to limited or insufficient knowledge, skills, and attitudes; and level 3 signifies adequate knowledge, skills, and attitudes. To be classified as level 3, the average score must be >2.7, indicating agreement or strong agreement with the statement (Table 1). The cutoffs are estimates defined by the authors and have not been validated by other parties. However, they have been used and found relevant in other studies using HLQ and eHLQ scores [20,33].

Table 1. Breakdown of knowledge, skills, and attitude into levels based on average scores.

Average scale score[a]	Level of knowledge, skills, and attitudes
<2.00	Level 1
2 to 2.7	Level 2
>2.7	Level 3

[a]The cutoffs are author-defined estimates, unvalidated by others but found relevant in studies using the Health Literacy Questionnaire and eHealth Literacy Questionnaire [20,33].

These 39 text strings are organized in Table 2 with the 13 scales as rows and the 3 levels as columns. To construct a text vignette, the appropriate text strings were selected from the table based on the participant's average scale score for each scale. The 13 selected text strings were then manually combined into a coherent text vignette. The initial text strings were revised throughout part 2 of the study in an iterative process using cognitive interviews for accuracy and relevance.

XSL•FO
RenderX

Table 2. Refined text strings based on scores.

Scale	Scores <2.00	Scores between 2.00 and 2.70	Scores >2.70
heiQ3[a]: self-monitoring and insight	He/she experiences having insufficient insight into his/her state of health, with insufficient knowledge of how to deal with health problems when they arise.	He/she experiences having limited insight into his/her state of health, with limited knowledge of how to deal with health problems when they arise.	He/she experiences having good insight into his/her state of health, with sufficient knowledge of how to deal with health problems when they arise.
heiQ4: constructive attitudes and approaches	He/she finds that his/her health problems limit him/her from enjoying life.	He/she finds that his/her health problems limit him/her from enjoying life to some extent.	He/she does not find that his/her health problems limit him/her from enjoying life.
heiQ5: skill and technique acquisition	He/she experiences having insufficient skills and techniques to deal with health problems when they arise.	He/she experiences having limited skills and techniques to deal with health problems as they arise.	He/she experiences having sufficient skills and techniques to deal with health problems as they arise.
heiQ8: emotional distress	He/she generally experiences a high degree of emotional distress due to his/her health.	He/she generally experiences a bit of emotional distress due to his/her health.	He/she generally experiences no emotional distress due to his/her health.
HLQ4[b]: social support for health	He/she does not feel understood or supported by his/her social network.	He/she feels only partially understood and supported by his/her social network.	He/she feels adequately understood and supported by his/her social network.
HLQ1: feel understood and supported by health care providers	He/she experiences a lack of understanding and support from the health care professionals he/she has contact with and access to.	He/she experiences limited understanding and support from the health care professionals he/she has contact with and access to.	He/she experiences adequate understanding and support from the health care professionals she/he has contact with and access to.
eHLQ2[c]: understanding of health concepts and language	He/she experiences having trouble understanding information about health and illness and may experience insufficient knowledge to have conversations with others about this.	He/she may experience problems understanding information about health and illness and may experience having limited knowledge to have conversations with others about this.	He/she does not experience having problems understanding information about health and illness and has enough knowledge to have conversations with others about this.
eHLQ1: using technology to process health information	He/she finds that he/she has have insufficient knowledge about how he/she can use technology to take care of his/her health.	He/she finds that he/she has limited knowledge about how he/she can use technology to take care of his/her health.	He/she has sufficient knowledge about how he/she can use technology to take care of his/her health.
eHLQ3: ability to actively engage with digital services	He/she has insufficient knowledge of how to use technology to navigate health care systems.	He/she has limited knowledge of how to use technology to navigate health systems.	He/she has good knowledge of how to use technology and can navigate health systems easily.
eHLQ6: access to digital services that work	He/she experiences insufficient access to digital health systems that work and can be accessed by himself/herself and others who need it.	He/she experiences having limited access to digital health systems that work and can be accessed by himself/herself and others who need it.	He/she experiences having sufficient access to digital health systems that work and can be accessed by himself/herself and others who need it.
eHLQ7: digital services that suit individual needs	He/she believes that digital services are insufficiently able to adapt to his/her needs.	He/she believes that digital services can adapt to his/her needs to a limited extent.	He/she believes that digital services can adequately adapt to his/her needs.
eHLQ5: motivated to engage with digital services	He/she finds that technology is insufficiently helpful and helps him/her keep up with and take care of his/her health.	He/she finds that technology is helpful to a limited extent and helps him/her keep up with and take care of his/her health.	He/she finds that technology is sufficiently useful and helps them keep up with and take care of his/her health.
eHLQ4: feel safe and in control	He/she feels uneasy about how his/her health data are used by others.	He/she feels partly uncomfortable with how his/her health data are used by others.	He/she feels comfortable with how his/her health data are used by others.

[a]heiQ: Health Education Impact Questionnaire.

[b]HLQ: Health Literacy Questionnaire.

[c]eHLQ: eHealth Literacy Questionnaire.

Data Sampling and Analysis

Data collection involved conducting cognitive interviews with each participant to validate the individual text vignettes developed from the individuals' READHY responses. These interviews allowed for iterative feedback and refinement of the text string to ensure their coherence and relevance. The analysis focused on identifying and addressing any issues in understanding and meaningfulness of the text strings and vignettes.

Iterative Validation of Data-Informed Text Vignettes

The text strings were validated through cognitive interviews with individuals and clinicians in an iterative process. Cognitive

XSL•FO
RenderX

interviews, often used for validating questionnaires, involve participants "thinking out loud" by verbally reporting their mental activity while engaging with an item in a questionnaire [34,35]. These interviews are conducted using a semistructured approach, with in-depth sessions where participants read a text vignette aloud and provide feedback by sharing their thoughts and insights throughout the process.

The interviews were conducted in 2 stages: first with individuals and then with clinicians. Before each individual's interview, they completed READHY, which was then used to create individual text vignettes. Each individual was presented with a full vignette, created based on their previous responses, to gather feedback on both individual text strings and the vignette as a coherent whole, evaluating its clarity and personal relevance.

After each interview, the data were analyzed individually to identify issues in the understanding and meaningfulness of the text vignettes. These individual findings were then aggregated across interviews to identify recurring problems, leading to adjustments of the text vignettes.

The iterative process included multiple stages of feedback incorporation. The first iteration considered feedback from 3 participants, with subsequent iterations incorporating feedback from additional participants—3 in the second and third iterations and 2 in the fourth.

Identified problems were categorized informally into 4 groups: scale-specific wording changes, need for further specification, issues with text vignette setup, and difficulties in answering the instrument [34].

Once all individual interviews were completed and the text vignettes and algorithm were adjusted, clinicians were introduced to the revised vignettes. The assignment of vignettes to clinicians was random, as they had no prior knowledge of or relationship with the individuals. Clinicians provided feedback on these vignettes, which led to further refinements. Each revised version was then presented to subsequent clinicians for additional input, ensuring that the final text was coherent and well understood.

Part 3 Methods

Data Sampling

First, 2 patients completed the instrument. From their responses, 2 individual text vignettes were constructed using the same method as applied in part 2. The average scale scores were calculated for each scale, and the scale score was then matched with the final adjusted corresponding sentence (Table 2). Afterward, the 13 text strings were combined to create an individual text vignette. These text vignettes were then presented to the 3 clinicians in 3 individual semistructured interviews designed to explore the text vignettes within an actual clinical environment and gain the clinicians' perspective. Each clinician was presented with 1 text vignette. Because all clinicians were somewhat familiar with booth patients, the specific vignette used in the interviews was randomly chosen based on which vignette was completed at the time of the interview.

Patients were also interviewed individually to assess their recognition of themselves in their text vignettes and whether it made sense for them in a clinical environment. However, the primary focus in part 3 was on the clinicians. Patient interviews were conducted via phone at the request of both patients. The interviews were recorded with consent and later transcribed.

The unique aspect lies in its exploration of text vignettes within clinical contexts, focusing on clinicians who interact with the patients described, and incorporating insights from their practical experience.

Analysis

Two interview guides were prepared, 1 for patient interviews and 1 for clinician interviews (Multimedia Appendix 1). These questions were grounded in the central inquiry: How can text vignettes generated by PRO data, as an alternative or supplement to traditional quantitative representations, improve dialogue and understanding between clinicians and patients regarding PROs?

All interviews were analyzed using qualitative content analysis using inductive coding to identify and describe patterns, categories, and meanings within the data [36]. The analysis was conducted using NVivo software (version 12.4.0; Lumivero) [37]. From the coding process, 4 main categories emerged: "recognition and confirmation," "challenges with context and conceptual understanding," "one size does not fit all," and "reflections on using text vignettes."

An example of a text vignette, as presented to both clinicians and patients, is shown in Textbox 2 (translated by the authors). This vignette is based on one of the patient's responses to the READHY instrument.

Textbox 2. Text vignette based on Readiness and Enablement Index for Health Technology.

She experiences having limited insight into her state of health, with limited knowledge of how to deal with health problems when they arise.

She finds that her health issues to some extent limit her from enjoying life.

She experiences having sufficient skills and techniques to deal with health problems as they arise and generally, experiences no emotional distress due to her health.

She feels adequately understood and supported by her social network and experiences adequate understanding and support from the healthcare professionals she has contact and access to.

She experiences having trouble understanding information about health and illness and may experience insufficient knowledge to have a conversation with others about this.

XSL•FO
RenderX

She finds that she has insufficient knowledge about how she can use technology to take care of her health and has insufficient knowledge of how to use technology to navigate healthcare systems.

She experiences having sufficient access to digital health systems that work and can be accessed by herself and others who need it, and that digital services can adequately adapt to her needs.

She finds that technology is helpful to a limited extent and helps her keep up with and take care of her health.

She feels comfortable with how her health data is used by others.

Results

Part 1

Four clusters with various levels of high, intermediate, and low average scores of scales within self-management, social capital, and digital health literacy were created (Table 3).

The mean value for each scale was presented for each of the 4 clusters that were informed. Values considered to represent a problematic level were indicated in red (<2.00), values considered to be below a sufficient level were indicated in yellow (2.00-2.50), and values considered to represent a sufficient level were represented in green (>2.50).

Table 3. Stratification using the Readiness and Enablement Index for Health Technology.

Metric (mean values)	Cluster 1 (n=16)	Cluster 2 (n=5)	Cluster 3 (n=11)	Cluster 4 (n=9)
heiQ3[a]: self-monitoring and insight	2.98	3.66	3.13	3.11
heiQ4: constructive attitudes and approaches	2.81	3.28	2.76	3.46
heiQ5: skill and technique acquisition	2.65	3.15	2.84	3.11
heiQ8: emotional distress	2.50	2.76	2.13	3.12
HLQ1[b]: feel understood and supported by health care providers	2.78	3.70	3.40	3.44
HLQ3: actively managing my health	2.68	3.40	2.92	2.68
HLQ4: social support for health	2.83	3.59	2.90	3.33
eHLQ1[c]: using technology to process health information	2.78	3.92	1.89	2.03
eHLQ2: understanding of health concepts and language	2.85	3.96	2.82	2.77
eHLQ3: ability to actively engage with digital services	2.95	3.96	1.99	2.65
eHLQ4: feel safe and in control	2.97	4.00	2.95	3.18
eHLQ5: motivated to engage with digital services	2.85	3.92	2.11	2.60
eHLQ6: access to digital services that work	2.86	3.96	2.42	2.97
eHLQ7: digital services that suit individual needs	2.78	3.80	2.04	2.68

[a]heiQ: Health Education Impact Questionnaire.

[b]HLQ: Health Literacy Questionnaire.

[c]eHLQ: eHealth Literacy Questionnaire.

On the basis of the thresholds, a focused text vignette was constructed almost as a narrative for each cluster together with key sociodemographic characteristics to create 4 text vignettes, wherein each represented an archetype of the profile—termed a "persona" (Multimedia Appendix 2). To simplify the information in this initial phase, only scales with high (initial green color) or low (initial red) average scores (>2.50 or <2.0) were included in the text vignettes, thereby creating focused text vignettes highlighting potential areas of barriers and resources.

EKW and LK then presented these text vignettes to colleagues in the Smart Inclusive Living Environments project. On the basis of the insights gained from the proof of concept with text vignettes at group or personal level, it was decided to proceed to part 2, focusing on the development of structured text strings that encompass high, intermediate, and low average score levels.

Part 2

A total of 39 initial text strings were created, with 3 strings for each of the 13 READHY scales, reflecting different levels of knowledge, skills, or attitudes within each scale. These text strings were designed to align with the scale names, which summarize the relevant experiences. The initial text strings are included in Multimedia Appendix 3.

Immediate feedback from the first 3 participants indicated that they could easily recognize themselves in the text vignettes, found them meaningful, and saw how their responses were reflected in them.

In the first iteration of text vignette validation, participants largely recognized themselves but identified issues related to the structure of the text vignettes and the completion of the instrument. Feedback included suggestions to revise lengthy sentences and reduce academic language, with specific terms

XSL•FO
RenderX

such as “health professionals” and “health-related concepts” being difficult to understand. Participants also struggled with the term “technology” used in the text vignette, such as in the description of scale eHLQ3.

On the basis of cognitive interviews, “healthcare providers” was changed to “healthcare professionals” and “health-related concepts” was simplified to “information of health and illness.” The order in which the scales were presented was also slightly changed. The order of HLQ4 and HLQ1 was adjusted to create a more natural flow when reading through the vignette. Similarly, the eHLQ items were reordered to enhance the coherence of the vignette.

In subsequent validation iterations, feedback was generally positive, with individuals finding the text vignettes understandable and recognizable. However, they struggled more with areas where they rated themselves as limited or insufficient. Some participants mentioned difficulties with the questionnaire due to the lack of a “don’t know” option and noted that terms such as “insufficient” felt too academic and unrepresentative of their self-view. The term “digital services” was also revised to “digital health services” to clarify its meaning. Participants suggested that discrepancies in text vignette descriptions might stem from specific episodes rather than their general experiences, and some participants requested more precise wording to better reflect their experiences. During iterations and refinement based on feedback, the understandability and recognizability of the vignettes improved, with individuals having an overall positive view of the text vignettes after final refinement.

The clinicians had a generally positive view about the generated READHY text vignettes. Feedback emphasized the need for clarity in highlighting that the text vignettes were based on a PRO instrument. The length of the text vignettes had mixed reviews. Feedback suggested categorizing content into taxonomic groups. The language was too academic, and adjustments had to be made to align with the instrument.

Clinicians see the value of using the text vignettes at the population level, allowing for generalization and description of specific populations based on competencies and characteristics. However, the current format may not be suitable for individual-level use, as it offers a simplified representation for personalized interaction. On the basis of input received from all participants, the strings were refined. The refined text strings are presented in Table 3 (translated from Danish to English by the authors). In the context of this study, we have used gender pronouns, but they should be adjusted to ensure relevance and significance within the context and setting.

Part 3

There was a positive attitude among the clinicians and patients toward the text vignette format and this way of presenting and communicating PROs. The clinicians experienced being able to recognize patients in the text vignettes, and patients experienced being able to recognize themselves in their text vignettes. Although the text vignette format does not necessarily meet the specific needs, the textual presentation makes the information from PROs more accessible in an understandable way for both clinicians and patients.

There are different preferences for how dynamic the tool should be. Some prefer to select only specific dimensions, while others believe that the strength of the text vignettes comes from them providing the full picture insights and exposing possibly overlooked aspects. Clinicians who favor the text vignette suggest that color coding as a supplement to the text vignette would be helpful and that a schematic summary would help with readability, indicating that flexibility in the profile presentation is needed to accommodate different preferences. Some prefer text-based representations, while others find visual presentations more intuitive.

Clinicians may face language and comprehension challenges due to ambiguity in words and abstract concepts within the text vignettes. Ambiguity in words and concepts, such as “to some extent,” leaves clinicians uncertain about the severity of the patient’s problems. Abstract concepts such as “health” and “technology” also pose challenges in interpretation. The text vignettes highlight new angles in patients’ self-perception, health, motivation, and emotional distress, which clinicians find crucial.

Opinions differ on whether to highlight strengths or weaknesses. Some prefer insights into both aspects, while others prefer focusing solely on the weaknesses that are most likely to be approved. Clinicians acknowledge the value of using text vignettes at a population or group level for generalization but point out some limitations for individual-level use due to simplified representations. Emphasis on personal dialogue with patients remains crucial for addressing unresolved questions and determining personalized interactions.

From the patient’s viewpoint, there is self-recognition and understanding; however, engaging in discussions about text vignette representation proves challenging. It is essential to note that clinicians report based on their experience of interacting with patients, highlighting the dynamic nature of this interaction in understanding and interpreting the text vignettes. The clinician’s insights provide a valuable perspective for understanding the broader implications of text vignettes in a clinical setting.

Discussion

Principal Findings

This study presents a process for developing text vignettes based on PRO data, resulting in the creation of individual vignettes that incorporate all 13 scales of the READHY framework. The motivation for the development of these text vignettes was the challenges associated with presenting PRO data as numeric values or graphical representations, aiming to provide a clearer and more engaging way to present and communicate PRO information.

This findings of this study suggest an overall positive attitude toward the PRO data–informed text vignette format and this way of presenting and communicating PROs. Clinicians experienced being able to recognize patients in the text vignettes,

XSL•FO
RenderX

and the patient's experienced being able to recognize themselves in and understand the text vignettes describing themselves.

Considering the background emphasizing the challenges associated with graphical representations of PRO data, the findings suggest that PRO-informed text vignettes are understandable for clinicians and patients. While the text format may not perfectly align with all the needs and interests of clinicians and patients, it presents PRO data in a comprehendible manner, with the text vignettes presenting a recognizable description.

Patients' Perspective

Patients found elements of self-recognition in the text vignettes, although discussing their representation could be challenging. It is important to note that text vignettes also face challenges related to health literacy and differences in understanding, as highlighted in part 3. These challenges can pose significant barriers for individuals with low health literacy or those who cannot read. In such cases, regardless of the presentation format, additional support, such as reading the vignettes aloud, may be necessary. However, the written format generally makes the data more accessible and less complex.

Another potential challenge is that the threshold values used to select appropriate text strings may result in choices that are not recognized by the informants. For example, individuals scoring between 2.00 and 2.70 might not feel limited, or they may perceive issues in areas not fully captured by the items in a scale. This issue requires further exploration in future studies to determine whether scales can be used effectively or whether a focus on single items is more appropriate, particularly for those scoring at the extremes. This may be necessary as we work in the intersection of presenting data from psychometric valid scales reporting on latent variables and presenting numeric values as an interpretation of experiences. The patients' experiences and feelings may sometimes be undermined or misinterpreted in this process.

Clinicians' Perspective

The data-informed text vignettes add a candid dimension to the information, making it directly engaging and allowing for the capture of new perspectives on patients that clinicians would not otherwise have focused on, although not all perspectives are equally relevant. It is especially the aspects of experienced support and patients' own knowledge about health and illness that are important for clinicians to understand and gain perspective on. There are different preferences among clinicians regarding the format of the text vignettes; some prefer text-based representations, while others prefer visual presentations. Likewise, there are different preferences regarding how dynamic the text-based tool should be. Some experience a need to be able to select only specific dimensions, while others believe it is part of the strength of the vignettes that it presents the full picture and thus insight into possibly overlooked aspects. This indicates that there is a need for flexibility in the text vignette presentation to accommodate different preferences.

Clinical Relevance

Clinicians using text vignettes stress the importance of technology that effectively addresses practical and treatment-specific needs. It remains imperative that the information within text vignettes aligns with the treatment being offered.

Clinicians emphasize that the central focus should be on understanding the condition and enabling patients to lead fulfilling lives. Reading the text vignettes prompts clinicians to reflect on aspects that may have been previously overlooked, serving as a reminder of the critical need for clinical relevance in any work involving PROs. It underlines that for PROs, and indeed text vignettes, to realize its full potential in creating value, it must align with the clinical context and be relevant to clinical practice.

A More Straightforward Representation of PROs

In addressing some of the challenges with the quantitative representation of PROs, the use of text vignettes allows for a deeper understanding of PROs. Text vignettes recognize the complexities of patients' experiences and provide a broader perspective. This approach prevents reducing patients to just numerical data, offering instead a more comprehensive view of an individual's health. The text vignettes are presented in an understandable and consistent format, with transparency about how the collected information is applied. This transparency is crucial for patients' effective use of PROs [38]. By transforming complicated numerical data into a clearer and more coherent text format, text vignettes offer a direct representation that ensures patients and clinicians can grasp details without relying on prior knowledge of graphical presentations. This textual format encourages conversation, minimizes ambiguity, and fosters dialogue for a better understanding of individual experiences.

Promoting Dialogue for Mutual Understanding

A data-informed text vignette is not a universal solution, but rather an approach with both advantages and disadvantages that must be weighted and adapted to individual preferences and needs and that requires a more dynamic approach to meet needs. Recognizing these challenges suggests a need for meaningful conversations between patients and clinicians about aligning their understanding at a conceptual level. Such discussions offer a chance to clarify word meanings and ensure a common understanding of the patient's experiences. This approach is essential for accurate and meaningful communication between patients and clinicians, ultimately improving the quality of care and treatment. Clinicians and patients may have different perspectives on what the patient reports in the PRO instrument. An example is how HLQ scores were not perceived in the same way by patients and their physicians, with a disconcordance ranging between 20% and 44% in the 9 HLQ scales [14]. Translating the data into text vignettes may ease the conversation and help identify these disconcordances for other PRO data also, ease the alignment of expectations, and provide a common understanding of the patient perspective.

XSL•FO
RenderX

Bias Concern

Clinicians are generally receptive to text vignettes and the insights they provide. Many clinicians feel that the text vignettes offer a strong sense of the patients before any direct contact is made. However, clinicians also express concerns that the text vignette may make them biased in their encounter with the patient, which may affect their ability to understand and treat the patient objectively. It emphasizes the necessity of acquiring proficiency in tool use, particularly when presenting information in a text format. Being mindful of potential biases becomes central, especially in a context where the text vignette format might intensify the impact. This highlights the importance of a nuanced and skillful application of the tool, recognizing that biases are not exclusive to text vignette representation but should be carefully navigated, especially when conveyed in a confrontational textual manner.

Strengths and Limitations

A strength of this study lies in the stepwise process of creating a dataset using a validated instrument, followed by data triangulation. This process is based on iteratively developed text lines that can be combined into cohesive text vignettes, forming complete paragraphs. The cocreation with individuals, people living with a chronic condition, municipalities, hospital-based clinicians, and 1 general practitioner helps ensure the validity and meaningfulness of the proposed solution.

The number of participants in each of the 3 parts may be considered a limitation. In part 1, convenience sampling was used among participants active in a telehealth service and was limited by availability. We do not find this to be a problem for the presented results, as we were able to create 4 profiles and personas that are meaningful and can serve to illustrate the principles used as background for parts 2 and 3. The number of participants in part 2 was determined through concurrent evaluation during data sampling. When the inclusion of additional informants no longer contributed new information, the sampling process was considered complete. By including data from 2 distinct sociodemographic areas, we aimed to capture a range of perspectives. On the basis of these arguments, we found the number to have sufficient information power for a proof-of-concept study [39]. This is also supported by most of the participants finding the text vignettes understandable and recognizable.

In particular, in part 3, the number of patients is likely to be too small, but we have included this part to illustrate how clinicians who are familiar with the patients respond to the text vignettes describing them. The findings need to be assessed in this light. Moreover, the results may differ from those that will be obtained using PRO data instruments that are developed for clinical purposes, such as European Organisation For Research And Treatment Of Cancer [40], The World Health Organization-Five Well-Being Index [41], or European Quality of life–5 Dimensions [42].

Perspective

The study has used the READHY instrument to demonstrate how data-based text vignettes can help to understand patients and align expectations about what the data tell. During the field studies and in discussions afterward, we have received expressions of interest in setting up algorithms to convert other established PRO data results into text vignettes. This calls for abundant work in this field in collaboration with several PRO instrument owners and distributors, and we welcome collaboration in this field. It also calls for atomization programming interfaces that build on the proposed structures with filters based on cutoff values combined with generative artificial intelligence to refine the created text vignettes.

Conclusions

The study demonstrates how text vignettes can be developed using PRO instruments, with individual scales as input strings. This provides an opportunity to present numeric values in a text format that is understandable and recognizable to most patients and clinicians. Some clinicians prefer text vignettes to other forms of presentations and find them especially useful in establishing a better understanding of a patient before initial contact. However, language and the arrangement of text vignettes can present challenges, underscoring the need for adaptation to specific needs and clinician training to ensure accurate interpretation of the text vignettes. Different preferences for text vignette formats highlight the need for a more dynamic and adaptable tool that can cater to individual needs and preferences.

Acknowledgments

The authors would like to thank all participants for their contribution. The authors would like to specially thank Torsten Risør, Sine Rossen, Rikke Terp, Thyge Lynghøj Nielsen, and the PreCareKlinikken for providing feedback throughout the process and Mette Krog Andreasen and Morten Brandrup for facilitating access to informants.

The authors thank the European Union (EU) Horizon 2020 and all partners in Smart Inclusive Living Environments (SMILE) for making this work possible. This research was funded by the EU Horizon 2020 Research and Innovation Program under grant 101016848.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guides.

XSL•FO
RenderX

[DOCX File , 17 KB - humanfactors_v12i1e58077_app1.docx]

Multimedia Appendix 2
Group-based text vignettes.
[DOCX File , 640 KB - humanfactors_v12i1e58077_app2.docx]

Multimedia Appendix 3
Initial text strings based on scores.
[PDF File (Adobe PDF File), 63 KB - humanfactors_v12i1e58077_app3.pdf]

References

1. Anvendelse af PRO-data. Danske Patienter. URL: https://danskepatienter.dk/vibis/inddragelse-i-eget-forloeb/pro-data/anvendelse-af-pro-data [accessed 2023-02-02]
2. Sandø N. Sundhedskompetence i det danske sundhedsvæsen. Copenhagen, Demark: Sundhedsstyrelsen; Aug 10, 2022.
3. Sundheds kompetence i det danske sundhedsvæsen - En vej til mere lighed. Sundhedsstyrelsen. URL: https://www.sst.dk/da/viden/Sundhedsvaesen/Ulighed-i-sundhed/Sundhedskompetence [accessed 2023-03-22]
4. Meadows K. Do patient-reported outcome measures tell us the full story? Clin Nurs Res 2022 Feb;31(2):159-162. [doi: 10.1177/10547738221078335] [Medline: 35125032]
5. Kwon JY, Russell L, Coles T, Klaassen RJ, Schick-Makaroff K, Sibley KM, et al. Patient-reported outcomes measurement in radiation oncology: interpretation of individual scores and change over time in clinical practice. Curr Oncol 2022 Apr 27;29(5):3093-3103 [FREE Full text] [doi: 10.3390/curroncol29050251] [Medline: 35621641]
6. Albers EA, Fraterman I, Walraven I, Wilthagen E, Schagen SB, van der Ploeg IM, et al. Visualization formats of patient-reported outcome measures in clinical practice: a systematic review about preferences and interpretation accuracy. J Patient Rep Outcomes 2022 Mar 03;6(1):18 [FREE Full text] [doi: 10.1186/s41687-022-00424-3] [Medline: 35239055]
7. Grossman LV, Feiner SK, Mitchell EG, Masterson Creber RM. Leveraging patient-reported outcomes using data visualization. Appl Clin Inform 2018 Jul;9(3):565-575 [FREE Full text] [doi: 10.1055/s-0038-1667041] [Medline: 30068012]
8. Liu LH, Garrett SB, Li J, Ragouzeos D, Berrean B, Dohan D, et al. Patient and clinician perspectives on a patient-facing dashboard that visualizes patient reported outcomes in rheumatoid arthritis. Health Expect 2020 Aug;23(4):846-859 [FREE Full text] [doi: 10.1111/hex.13057] [Medline: 32270591]
9. Kuijpers W, Giesinger JM, Zabernigg A, Young T, Friend E, Tomaszewska IM, et al. Patients' and health professionals' understanding of and preferences for graphical presentation styles for individual-level EORTC QLQ-C30 scores. Qual Life Res 2016 Mar;25(3):595-604 [FREE Full text] [doi: 10.1007/s11136-015-1107-3] [Medline: 26353905]
10. Redegørelse Om Ældre Området 2020. Sundheds- og Ældreministeriet. 2020. URL: https://aeldremin.dk/Media/637666126708148318/AE_Redegoerelse_om_aeldreomraadet_2020.pdf [accessed 2024-04-29]
11. Loth FL, Holzner B, Sztankay M, Bliem HR, Raoufi S, Rumpold G, et al. Cancer patients' understanding of longitudinal EORTC QLQ-C30 scores presented as bar charts. Patient Educ Couns 2016 Dec;99(12):2012-2017 [FREE Full text] [doi: 10.1016/j.pec.2016.08.004] [Medline: 27506581]
12. Campbell R, Ju A, King MT, Rutherford C. Perceived benefits and limitations of using patient-reported outcome measures in clinical practice with individual patients: a systematic review of qualitative studies. Qual Life Res 2022 Jun;31(6):1597-1620. [doi: 10.1007/s11136-021-03003-z] [Medline: 34580822]
13. Munzner T. Visualization Analysis and Design: Opportunities for Domestic Investments in Water and Sanitation for the Poor. Boca Raton, FL: CRC Press; 2015.
14. Hawkins M, Gill SD, Batterham R, Elsworth GR, Osborne RH. The Health Literacy Questionnaire (HLQ) at the patient-clinician interface: a qualitative study of what patients and clinicians mean by their HLQ scores. BMC Health Serv Res 2017 Apr 27;17(1):309 [FREE Full text] [doi: 10.1186/s12913-017-2254-8] [Medline: 28449680]
15. El-Daly I, Ibraheim H, Rajakulendran K, Culpan P, Bates P. Are patient-reported outcome measures in orthopaedics easily read by patients? Clin Orthop Relat Res 2016 Jan;474(1):246-255 [FREE Full text] [doi: 10.1007/s11999-015-4595-0] [Medline: 26472587]
16. Beauchamp A, Dodson S, Batterham R, Asproloupos D, Richard OC, Cavalli S, et al. Ophelia manual how to apply the Ophelia health literacy process for improving health outcomes. Swinburne University of Technology. 2017. URL: https://www.nswmentalhealthcommission.com.au/sites/default/files/2021-06/swinburne-_ophelia_manual_orange.pdf [accessed 2024-04-29]
17. Batterham RW, Hawkins M, Collins PA, Buchbinder R, Osborne RH. Health literacy: applying current concepts to improve health services and reduce health inequalities. Public Health 2016 Mar;132:3-12. [doi: 10.1016/j.puhe.2016.01.001] [Medline: 26872738]
18. Bakker MM, Putrik P, Aaby A, Debussche X, Morrissey J, Råheim BC, et al. Acting together – WHO national health literacy demonstration projects (NHLDPs) address health literacy needs in the European region. World Health Organization.

XSL•FO
RenderX

2019. URL: https://iris.who.int/bitstream/handle/10665/327059/php-5-2-3-233-243-eng.pdf?sequence=1&isAllowed=y [accessed 2024-04-29]
19. Kayser L, Rossen S, Karnoe A, Elsworth G, Vibe-Petersen J, Christensen JF, et al. Development of the multidimensional readiness and enablement index for health technology (READHY) tool to measure individuals' health technology readiness: initial testing in a cancer rehabilitation setting. J Med Internet Res 2019 Feb 12;21(2):e10377 [FREE Full text] [doi: 10.2196/10377] [Medline: 30747717]
20. Kayser L, Karnoe A, Duminski E, Jakobsen S, Terp R, Dansholm S, et al. Health professionals' eHealth literacy and system experience before and 3 months after the implementation of an electronic health record system: longitudinal study. JMIR Hum Factors 2022 Apr 29;9(2):e29780 [FREE Full text] [doi: 10.2196/29780] [Medline: 35486414]
21. Mankins JC. Technology readiness assessments: a retrospective. Acta Astronautica 2009 Nov;65(9-10):1216-1223. [doi: 10.1016/j.actaastro.2009.03.058]
22. Hvad er PreCareKlinikken? PreCareKlinikken. URL: https://precareklinikken.dk/om/ [accessed 2023-05-21]
23. Phanareth K, Dam AL, Hansen MA, Lindskrog S, Vingtoft S, Kayser L. Revealing the nature of chronic obstructive pulmonary disease using self-tracking and analysis of contact patterns: longitudinal study. J Med Internet Res 2021 Oct 19;23(10):e22567 [FREE Full text] [doi: 10.2196/22567] [Medline: 34665151]
24. Home page. SMILE eHealth. URL: https://smileehealth.eu/home [accessed 2024-01-13]
25. International standard classification of education ISCED 2011. UNESCO Institute for Statistics. URL: https://uis.unesco.org/en/topic/international-standard-classification-education-isced [accessed 2024-04-29]
26. Rossen S, Kayser L, Vibe-Petersen J, Ried-Larsen M, Christensen JF. Technology in exercise-based cancer rehabilitation: a cross-sectional study of receptiveness and readiness for e-Health utilization in Danish cancer rehabilitation. Acta Oncol 2019 May;58(5):610-618. [doi: 10.1080/0284186X.2018.1562213] [Medline: 30698060]
27. Terp R, Kayser L, Lindhardt T. Older patients' competence, preferences, and attitudes toward digital technology use: explorative study. JMIR Hum Factors 2021 May 14;8(2):e27005 [FREE Full text] [doi: 10.2196/27005] [Medline: 33988512]
28. Osborne RH, Elsworth GR, Whitfield K. The Health Education Impact Questionnaire (heiQ): an outcomes and evaluation measure for patient education and self-management interventions for people with chronic conditions. Patient Educ Couns 2007 May;66(2):192-201. [doi: 10.1016/j.pec.2006.12.002] [Medline: 17320338]
29. Sørensen K, Van den Broucke S, Pelikan JM, Fullam J, Doyle G, Slonska Z, HLS-EU Consortium. Measuring health literacy in populations: illuminating the design and development process of the European Health Literacy Survey Questionnaire (HLS-EU-Q). BMC Public Health 2013 Oct 10;13(1):948 [FREE Full text] [doi: 10.1186/1471-2458-13-948] [Medline: 24112855]
30. Kayser L, Karnoe A, Furstrand D, Batterham R, Christensen KB, Elsworth G, et al. A multidimensional tool based on the eHealth literacy framework: development and initial validity testing of the eHealth literacy questionnaire (eHLQ). J Med Internet Res 2018 Feb 12;20(2):e36 [FREE Full text] [doi: 10.2196/jmir.8371] [Medline: 29434011]
31. Overview of mandatory reporting | Danish research ethics committees. The Danish National Center for Ethics. URL: https://researchethics.dk/information-for-researchers/overview-of-mandatory-reporting [accessed 2024-03-05]
32. IBM SPSS statistics for Windows. IBM Corp. URL: https://www.ibm.com/products/spss-statistics [accessed 2024-04-29]
33. Poulsen HS, Eiriksson SD, Christiansen AS, Wingstrand A. Sundhedsprofil 2021 for Region Sjælland og kommuner: »Hvordan har du det?«. Region Sjælland. URL: https://pure-portal.regsj.dk/da/publications/sundhedsprofil-2021-for-region-sj%C3%A6lland-og-kommuner-hvordan-har-d [accessed 2024-04-29]
34. Willis GB. Analyzing and documenting cognitive interview results. In: Willis GB, editor. Cognitive Interviewing. Thousand Oaks, CA: Sage Publications; 2011.
35. Conrad F, Blair J. From impressions to data: Increasing the objectivity of cognitive interviews. ResearchGate. 1996. URL: https://www.researchgate.net/publication/237601397_From_impressions_to_data_Increasing_the_objectivity_of_cognitive_interviews [accessed 2024-04-29]
36. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
37. NVivo. Lumivero. URL: https://lumivero.com/products/nvivo/ [accessed 2024-04-29]
38. Carfora L, Foley CM, Hagi-Diakou P, Lesty PJ, Sandstrom ML, Ramsey I, et al. Patients' experiences and perspectives of patient-reported outcome measures in clinical care: a systematic review and qualitative meta-synthesis. PLoS One 2022 Apr 21;17(4):e0267030 [FREE Full text] [doi: 10.1371/journal.pone.0267030] [Medline: 35446885]
39. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res 2016 Nov;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
40. Wintner LM, Sztankay M, Aaronson N, Bottomley A, Giesinger JM, Groenvold M, EORTC Quality of Life Group. The use of EORTC measures in daily clinical practice-a synopsis of a newly developed manual. Eur J Cancer 2016 Nov;68:73-81 [FREE Full text] [doi: 10.1016/j.ejca.2016.08.024] [Medline: 27721057]
41. Topp CW, Østergaard SD, Søndergaard S, Bech P. The WHO-5 well-being index: a systematic review of the literature. Psychother Psychosom 2015;84(3):167-176 [FREE Full text] [doi: 10.1159/000376585] [Medline: 25831962]
42. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [doi: 10.1016/0168-8510(90)90421-9] [Medline: 10109801]

XSL•FO
RenderX

Abbreviations

COPD: chronic obstructive pulmonary disease
eHLQ: eHealth Literacy Questionnaire
heiQ: Health Education Impact Questionnaire
HLQ: Health Literacy Questionnaire
OPHELIA: Optimizing Health Literacy and Access
PRO: patient-reported outcome
READHY: Readiness and Enablement Index for Health Technology

Edited by A Kushniruk, D Kaufman; submitted 05.03.24; peer-reviewed by J Yang, C Cheng, Y Chu; comments to author 04.08.24; revised version received 04.09.24; accepted 31.12.24; published 05.02.25.

Please cite as:
Kelly SC, Wegener EK, Kayser L
Creation of Text Vignettes Based on Patient-Reported Data to Facilitate a Better Understanding of the Patient Perspective: Design Study
JMIR Hum Factors 2025;12:e58077
URL: https://humanfactors.jmir.org/2025/1/e58077
doi:10.2196/58077
PMID:

©Sofia Charlotte Kelly, Emilie Kauffeldt Wegener, Lars Kayser. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 05.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Development of a Web-Based Intervention for Middle Managers to Enhance Resilience at the Individual, Team, and Organizational Levels in Health Care Systems: Multiphase Study

Eva Gil-Hernández[1], MSc; Irene Carrillo[2], PhD; Jimmy Martin-Delgado[3,4], MD, PhD; Daniel García-Torres[1], MSc; José Joaquín Mira[1,2,5], MPH, PhD

[1]ATENEA Research Group, Foundation for the Promotion of Health and Biomedical Research of Valencia Region (FISABIO), Alicante, Spain

[2]Health Psychology Department, Universidad Miguel Hernández, Elche, Spain

[3]Junta de Beneficencia de Guayaquil, Hospital de Especialidades Alfredo Paulson, Portoviejo, Ecuador

[4]Instituto de Investigación e Innovación en Salud Integral, Facultad de Ciencias de la Salud, Universidad Católica de Santiago de Guayaquil, Guayaquil, Ecuador

[5]Conselleria de Sanidad, Health Centre Hospital Provincial-Pla, Alicante-Sant Joan Health District, Alicante, Spain

Corresponding Author:
José Joaquín Mira, MPH, PhD
Health Psychology Department
Universidad Miguel Hernández
Avinguda de la Universitat d'Elx
Elche, 03202
Spain
Phone: 34 966658984
Email: jose.mira@umh.es

Abstract

Background: Health care institutions face high systemic risk due to the inherent uncertainty and complexity of their operations. This often leads to stressful incidents impacting the well-being of health care professionals, which can compromise the effectiveness of health care systems. Enhancing resilience among health care professionals is essential for maintaining high-quality care and ensuring patient safety. The role of middle managers is essential to ensure the response capacity of individuals and teams.

Objective: This study aims to develop a web-based intervention aimed at middle management to enhance individual, team, and organizational resilience.

Methods: An observational study was conducted in 3 phases: design, validation, and pilot study. The study was initiated in February 2022 and concluded in June 2023. Phase 1 involved designing the content for the web-based tool based on a comprehensive review of critical elements around resilience. Phase 2 included validation by an international panel of experts who reviewed the tool and rated it according to a structured grid. They were also encouraged to highlight strengths and areas for improvement. Phase 3 involved piloting the tool with health care professionals in Ecuador to refine the platform and assess its effectiveness. A total of 458 people were invited to participate through the Institutional Course on Continuous Improvement in Health Care Quality and Safety offered by the Ministry of Public Health of Ecuador.

Results: The tool, eResiliencia, was structured into 2 main blocks: individual and team resilience and organizational resilience. It included videos, images, PDFs, and links to dynamic graphics and additional texts. Furthermore, 13 (65%) of the 20 experts validated the tool, rating content clarity at an average of 4.5 (SD 0.7) and utility at an average of 4.7 (SD 0.5) out of 5. The average overall satisfaction was 9.3 (SD 0.6) out of 10 points, and feedback on improvements was implemented. A total of 362 health care professionals began the intervention, of which 218 (60.2%) completed preintervention and postintervention questionnaires, with significant knowledge increases (P<.001). Of the 362 health care professionals, 146 (40.3%) completed the satisfaction questionnaire, where overall satisfaction was rated at an average of 9.4 (SD 1.1) out of 10 points.

Conclusions: The eResiliencia web-based platform provides middle managers with resources to enhance resilience among their teams and their components, promoting better well-being and performance, even under highly stressful events. Future research should focus on long-term impacts and practical applications in diverse clinical settings.

(JMIR Hum Factors 2025;12:e67263) doi:10.2196/67263

XSL•FO
RenderX

KEYWORDS

resilience; health care professionals; web-based intervention; middle management; well-being; patient safety

Introduction

Systemic Risk

Health care institutions are considered high-reliability organizations due to the inherent systemic risk associated with the activities they conduct [1]. Uncertainty and complexity, inherent in health care delivery along with other factors, contribute to the occurrence of higher stressful incidents that sometimes result in harm (adverse events) [2]. Health care professionals are exposed almost daily to stressful situations that can deteriorate their occupational well-being, thus compromising the effectiveness of health care systems [3]. In addition, the impact and demands stemming from unexpected health outcomes and health care crises further exacerbate this situation. Ultimately, the post–COVID-19 pandemic period has underscored a reality that has always existed but has not been adequately addressed until now: the well-being of health care professionals is a necessary condition to ensure the effectiveness of health care systems and the quality of care [4].

Professional Well-Being

In clinical contexts, situations with the potential to deteriorate the well-being of professionals are varied. These situations can range from those related to patient safety, such as near-misses, poor prognoses, unexpected health outcomes, a patient's deteriorating condition, preventable and unavoidable adverse events, and the unexpected death of a patient, to those more associated with working conditions or structural factors. These include a lack of resources, high care pressure, elevated staff turnover, increased demand, violence against professionals, and moral injury. They can have an emotional impact on professionals, which affects their habits, practices, or clinical decisions [5]. When this behavior change occurs in the undesired direction, we may encounter defensive practices, burnout, or detachment that often entail assuming unnecessary risks for the patient and additional costs for the system [6,7].

Vicious Cycle

Even in the absence of behavioral change, the loss of professional well-being is accompanied by an exponential increase in the risk of human error in patient care, leading to lower performance and a deterioration in the quality of care provided to patients [8]. This worsening of outcomes can, in turn, contribute to the occurrence of new errors and safety incidents. Thus, we find workers in a vicious cycle where the loss of well-being, the inability to regain balance and face stressful situations, and the quality of care are intimately related [5,9]. Hence, preserving and enhancing the resilience of professionals is a key element in ensuring the proper functioning of health care systems, the quality of care, patient safety, and achieving better health outcomes [10].

Resilience

Resilience refers to the ability of individuals, teams, and organizations to adapt, recover, and grow in the face of stress, adversity, and challenging circumstances [11]. In the context of health care, it enables professionals to maintain their well-being and continue to provide high-quality care despite the inevitable challenges they encounter. This not only benefits the individuals themselves but also enhances the safety and effectiveness of care provided by health care teams [12].

Individual Differences

In addition to the inherent risk in health care activities, other cultural aspects can also be a source of distress for health care professionals and may even constitute a barrier to the process of recovering well-being once it has been compromised and are also directly related to values. From an individual perspective, health care professionals share a set of ideals focused on doing good and prioritizing the well-being of others. While these ideals are desirable for good performance, their rigid application in critical situations can intensify distress or hinder constructive coping and professional recovery [13].

Organizational Culture

At the organizational level, the institution's culture plays a key role in determining how professionals feel, interpret, and cope with stressful situations. The prevailing professional culture in the health care sector is characterized by strong ideals of perfection that often lead to a negative and unconstructive view of human response, considering someone weak for asking for help or being unable to cope with a stressful situation. Understanding these causes allows for designing and implementing barriers that prevent or reduce the likelihood of succumbing to stressful situations occurring in the future.

Challenge

The World Health Organization has incorporated resilience in health care systems into the Global Patient Safety Action Plan 2021-2030 through strategy 2.4: providing a strong human factor and ergonomics-based perspective and contributing to strengthening the resilience of health care organizations and clinical practices [14]. This initiative aims to address a basic need for the proper functioning of health care systems, which is to take care of those who care.

We intend to address this need not only from a reactive approach when things go wrong but also by creating resilient teams and providing middle managers with tools to reinforce the resilience of professionals, teams, and the health care organization to ensure their ability to respond to critical situations. We believe that middle managers, due to their strategic position in the institution's structure, are key agents for the development of resilience at all levels; therefore, the materials of this project are specifically aimed at this group.

Methods

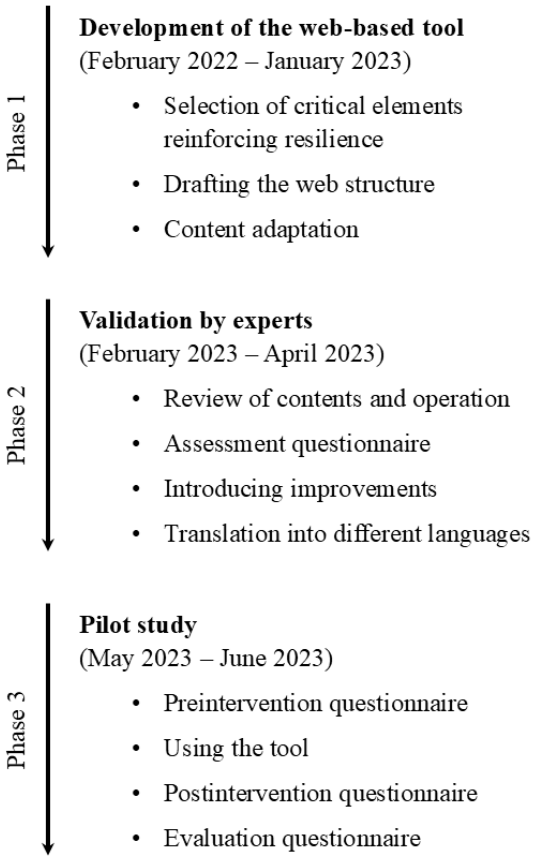
Study Design

An observational study was developed in 3 phases: development of the web-based tool (phase 1), validation by experts (phase 2), and piloting with professionals (phase 3; Figure 1). Primary care and hospital were defined as usual settings for the new tool

XSL•FO
RenderX

named eResiliencia. The established 8 principles for learning tools that aim to support the translation of resilience into practice by Haraldseid-Driftland et al [15] were used. In this study, resilience in health care was defined as the capacity to adapt to challenges and changes at different system levels to maintain high-quality care [12].

Figure 1. Flowchart of project phases and steps.



The study was developed by a core group involving a multidisciplinary team supported by a panel of international experts in health administration, occupational safety, and patient safety. It was initiated in February 2022 and concluded in June 2023.

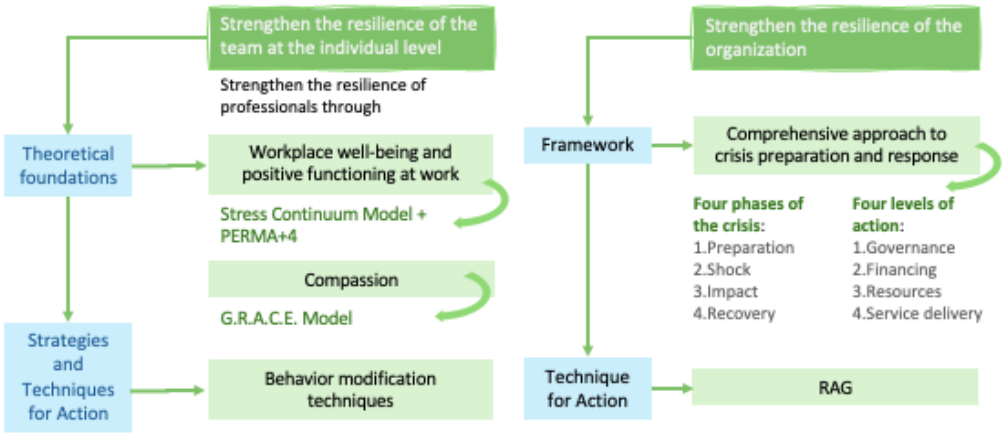
Phase 1: Development of the Web-Based Tool

To design the content for the web-based tool (phase 1), a comprehensive search of critical elements reinforcing resilience was conducted, culminating in a preliminary selection from among the most frequently cited elements. This included the revision of reviews and systematic reviews (after 2019) on the concept of resilience, experimental studies measuring the effectiveness of recent programs designed to enhance the resilience of health care professionals, and the content of different scales for measuring resilience. In particular, this search was aimed at identifying experiences and successful strategies [16].

The eResiliencia intervention is built on a robust theoretical framework that integrates principles of resilience, human behavior, and organizational adaptability to address the complex challenges faced by health care systems (Figure 2). This approach emphasizes strengthening resilience through the development of 5 essential traits: self-control, which involves regulating emotions and behavior under pressure; adaptability or the flexibility to adjust to changes and demands; optimism, characterized by a positive outlook and confidence in overcoming challenges; self-sufficiency, reflecting the ability to independently manage difficulties; and persistence, the sustained effort to persevere in the face of adversity.

XSL•FO
RenderX

Figure 2. eResiliencia theoretical framework. PERMA+4: PERMA+4 Well-Being model; RAG: Resilience Analysis Grid.



The intervention draws from the 5-tier model proposed by Seys et al [17], which combines preventive and supportive measures to build resilience. Levels 1 and 2 focus on proactive strategies such as prevention and self-care, while levels 3 to 5 offer structured peer and clinical support to address resilience at both individual and organizational levels. This multilevel approach ensures that resilience-building efforts are comprehensive and adaptable to the diverse contexts within health care settings.

eResiliencia operates across 3 interconnected levels: individual, team, and organizational. It uses both proactive and reactive strategies to enhance resilience. Proactive measures aim to build capacity before crises occur, using tools such as stress management models and workplace well-being frameworks. Reactive strategies focus on recovery and learning following critical events, ensuring that individuals and organizations can adapt effectively to challenges.

The intervention is informed by several evidence-based models, including the stress continuum model [18], which helps monitor and manage stress; the PERMA+4 Well-Being model [19], which promotes workplace well-being through its multidimensional framework; the G.R.A.C.E. model [20], which supports health care workers in cultivating compassion and managing burnout; and the Resilience Analysis Grid [21], which was applied to assess and enhance team response capacity in unexpected events. In addition, eResiliencia incorporates a set of behavior change techniques [22] designed to equip middle managers with targeted strategies for promoting workplace well-being. These techniques are tailored to the unique cultural and organizational contexts of health care systems, ensuring their relevance and practicality.

Following this, a second step involved drafting the web structure to accommodate the intended contents effectively. Subsequently, a process of content adaptation to the health care context ensued, wherein various support materials for the tool were developed. These materials underwent 2 rounds of internal review before arriving at version 1.0, ensuring their alignment with the intended objectives. The first version of the web platform was developed in Spanish and hosted on the Miguel Hernández University of Elche (Spain) servers, which comply with security requirements and regulations. A total of 11 months were spent to complete this phase, commencing in February 2022.

Phase 2: Validation by Experts

Version 1.0 underwent a review and validation process by an international panel of experts between February and April 2023. A total of 20 professionals integrated this panel, comprising 14 (70%) from Spain, 2 (10%) from Colombia, 1 (5%) from Argentina, 1 (5%) from Ecuador, 1 (5%) from Brazil, and 1 (5%) from Chile. The experts' identification was made using the snowball approach. All experts were informed of the pursued objective and the approximate time commitment and were requested voluntary, unpaid participation. Of the 20 professionals, 11 (55%) were women and 9 (45%) were men, and all met the criterion of a minimum of 10 years of professional experience. It was ensured that no personal evaluation would be disclosed, and communications with them were conducted via telephone or email.

Following a comprehensive examination of all contents and their structural organization, experts engaged were required to respond to a rubric designed to evaluate various aspects. This rubric used a scale ranging from 1 to 5 to assess the clarity of the content and its utility and applicability, grading from completely disagree to completely agree on whether the materials facilitated learning strategies to foster the resilience of professionals or teams under their purview, their likelihood to recommend the tool to other colleagues, and their intention to apply learned tools or strategies. In addition, experts were encouraged to identify strengths and areas for improvement and to rate their overall satisfaction with the web-based intervention and its contents on a scale from 1 to 10, where the higher the score, the greater the satisfaction. The level of agreement among evaluators was assessed to inform decisions regarding aspects requiring changes. All feedback and suggested improvements were meticulously implemented to enhance the quality of the content. Subsequently, the refined version of the tool was translated into English, Portuguese, and Valencian. Careful attention was given to ensuring equivalence across all versions through a process of back translation. Necessary adjustments were made to ensure the adequacy and coherence of the system across languages. An adaptation was also made to the idioms

XSL•FO
RenderX

and expressions commonly used in Latin American Spanish. This procedure also ensured that social and cultural aspects specific to each country were considered.

Phase 3: Pilot Study

From May 1 to June 30, 2023, a pilot study was conducted in a real setting in Ecuador. The participants' feedback served to refine the platform and the tool. A broader cohort of 458 health care professionals from Ecuador was invited to participate. They served as executives or managers in one of the country's public health care networks. This invitation was extended through the Institutional Course on Continuous Improvement in Health Care Quality and Safety offered by the Ministry of Public Health of Ecuador with the support of the Miguel Hernández University, Spain. The invitation was formally extended to all professionals from the management teams of public health care institutions in the country, including physicians, nurses, and other professional profiles. They assessed whether they wished to participate in this training, ensuring their data and performance were completely anonymized. Participants were required to complete a preintervention questionnaire at the moment of registration to access the contents of the web-based tool. This questionnaire comprised 20 questions assessing the knowledge areas to be addressed. In Multimedia Appendix 1, the topics addressed by each question can be consulted. After 7 days, a postintervention questionnaire containing the same questions was made available to assess content assimilation. Participants had the option to retake this postintervention questionnaire up to a maximum of 10 times, with a minimum score of 8 out of 10 required to qualify for a proficiency certificate. Following the completion of the postintervention questionnaire, participants were presented with an evaluation questionnaire consisting of the same questions posed to the experts during phase 2.

Data Analysis

In this study, both quantitative and qualitative analyses were conducted. The profile of participants who completed all phases of the iteration was compared with those who dropped out and did not finish to identify potential differences. For phase 2, descriptive analyses were performed on the rubric scores, and content analyses were conducted for open-ended responses. Fleiss κ was used to analyze the level of agreement among evaluators. The response options were recoded into 2 categories: "agree" (values 4 and 5) and "disagree" (values 1, 2, and 3), evaluating the overall agreement among experts across the 5 variables (clarity, utility, intention to recommend, intention to apply, and eResiliencia).

Similarly, in phase 3, the same analyses as in phase 2 were performed for the evaluation questionnaire. Descriptive statistics were conducted for both the preintervention and postintervention questionnaires. To assess differences between the means of such questionnaires, the Wilcoxon signed rank test was used. In addition, the McNemar statistic was calculated to compare the number of correct responses before and after the use of the web-based tool for each question. Qualitative analysis also considered the frequency of positive comments and improvement suggestions made by participants at the end of the structured survey. These comments, expressed in natural language, were categorized by 2 researchers (EG-H and IC) to facilitate their documentation. Finally, sociodemographic data were analyzed using descriptive statistics.

Ethical Considerations

In accordance with the Ecuadorian Ministerial Agreement AM00005-2022 and the Spanish Law on Biomedical Research (14/2007), studies conducted through anonymous surveys that do not collect health-related data are exempt from ethical review. This study was fully compliant with all regulatory standards for personal data protection. Ethical principles for medical research involving human participants included in the guidelines set by the Helsinki Declaration were followed. Informed consent for study participation was obtained at the time of registration on the platform, whereby individuals were required to select the corresponding checkbox, with instructions provided regarding the process for revoking their participation. No form of financial compensation was provided for participation.

Results

Phase 1: Development of the Web-Based Tool

A total of 42 documents were deemed relevant and were analyzed in depth to extract lessons learned and proposals for the design of eResiliencia (Table 1).

XSL•FO
RenderX

Table 1. Documents considered when designing the content of eResiliencia.

Name	Study
Resilience scales (n=7)	
Resilience at Work Scale	Winwood et al [23], 2013
50-Item Resilience Questionnaire	The Psychometric Project [24], 2013
Brief Resilience Scale	Smith et al [25], 2008
Connor-Davidson Resilience Scale	Connor and Davidson [26], 2003
Resilience Scale for Adults	Friborg et al [27], 2003
Resilience Scale	Wagnild and Young [28], 1993
Dispositional Resilience Scale	Bartone et al [29], 1989
Resilience concept (n=11)	
Psychological Resilience: An Affect-Regulation Framework	Troy et al [30], 2023
A Simultaneous Concept Analysis of Resilience, Coping, Posttraumatic Growth, and Thriving	Bowling et al [31], 2022
An Approach to the Unified Conceptualization, Definition, and Characterization of Social Resilience	Moya and Goenechea [32], 2022
Integrative Review of the Recent Literature on Human Resilience: From Concepts, Theories, and Discussions Towards a Complex Understanding	Métais et al [33], 2022
Resilience: An Integrated Review	Daly [34], 2020
Psychological Resilience as an Emergent Characteristic for Well-Being: A Pragmatic View	Tay and Lim [35], 2020
What Does Resilience Signify? An Evaluation of Concepts and Directions for Future Research	Infurna [36], 2020
Nurse Resilience: A Concept Analysis	Cooper et al [37], 2020
What Are the Factors Affecting Resilience in Health Professionals? A Synthesis of Systematic Reviews	Huey and Palaganas [38], 2020
Toward a Transversal Definition of Psychological Resilience: A Literature Review	Sisto et al [11], 2019
Resilience as a Multimodal Dynamic Process	Stainton et al [39], 2019
Resilience in health care systems (n=9)	
Exploring the Nature of Adaptive Capacity for Resilience in Healthcare Across Different Healthcare Contexts; a Metasynthesis of Narratives	Lyng et al [40], 2022
Resilience in Organization-Related Research: An Integrative Conceptual Review Across Disciplines and Levels of Analysis	Raetze et al [41], 2022
Shifting Focus from Burnout and Wellness toward Individual and Organizational Resilience	Vercio et al [42], 2021
Health System Resilience: a Literature Review of Empirical Research	Biddle et al [43], 2020
Measuring the Resilience of Health Systems in Low- and Middle-Income Countries: a Focus on Community Resilience	Bhandari and Alonge [44], 2020
Maintaining capacity in the health care system during the COVID‐19 pandemic by reinforcing clinicians' resilience and supporting second victims	Strametz et al [45], 2020
COVID-19: Peer Support and Crisis Communication Strategies to Promote Institutional Resilience	Wu et al [46], 2020
Strengthening Health Systems Resilience: Key Concepts and Strategies	Thomas et al [47], 2020
Conceptual Analysis of Health Systems Resilience: A Scoping Review	Turenne et al [48], 2019
Existing intervention programs to enhance resilience (n=6)	
Psychosocial Interventions for Building Resilience of Informal Carers of People Living with Stroke: a Systematic Review	Qureshi et al [49], 2023
Building Personal Resilience following an online Resilience Training Program for BScN Students	Stoliker et al [50], 2022

XSL•FO
RenderX

Name	Study
A Resilience-Building App to Support the Mental Health of Health Care Workers in the COVID-19 Era: Design Process, Distribution, and Evaluation	Golden et al [51], 2021
The Effectiveness of Charge Nurse Training on Leadership Style and Resiliency	Spiva et al [52], 2020
Comprehensive Meta-analysis of Resilience Interventions	Liu et al [53], 2020
Interventions to Improve Resilience in Physicians Who have Completed Training: A Systematic Review	Venegas et al [54], 2019
Useful tools (n=9)	
PERMA+4: A Framework for Work-Related Wellbeing, Performance and Positive Organizational Psychology 2.0	Donaldson et al [19], 2022
Effectiveness of a One Day Self-Compassion Training for Pediatric Nurses' Resilience	Franco and Christie [55], 2021
The Positive Functioning at Work Scale: Psychometric Assessment, Validation, and Measurement Invariance	Donaldson and Donaldson [56], 2021
The PERMA-Profiler: A Brief Multidimensional Measure of Flourishing	Butler and Kern [57], 2016
The Behavior Change Technique Taxonomy (v1) of 93 Hierarchically Clustered Techniques: Building an International Consensus for the Reporting of Behavior Change Interventions	Michie et al [22], 2013
G.R.A.C.E. for Nurses: Cultivating Compassion in Nurse/Patient Interactions	Halifax [20], 2014
RAG—Resilience Analysis Grid	Hollnagel et al [21], 2011
US Marine Corps and Navy Combat and Operational Stress Continuum Model: A Tool for Leaders	Nash [18], 2016
Caring for Our Own: Deploying a Systemwide Second Victim Rapid Response Team	Scott et al [58], 2010

The working team outlined a presentation structure for eResiliencia and chose the most suitable appearance from several options. With the information gathered and the ideas from the working group itself, the first draft of the content was structured in the beta version of the website that hosts eResiliencia [59].

The content was divided into 2 main blocks (Figure 3), one aimed at building resilience at the individual and team levels, and another at the organizational level, preceded by a conceptual framework. Platform navigation was conducted through both the menu bar and the side panel (Figure 4). Within each block, there were subdivisions as shown in Table 2. In the tab dedicated to the conceptual framework, an explanation of the course context, an introduction to resilience, several examples of situations where it is crucial, and a section with definitions of the most relevant concepts of the course were provided. The first block, tools to strengthen resilience at the individual and team levels, offered the necessary strategies to enhance the resilience of team members at an individual level, placing particular emphasis on the importance of workplace well-being. Finally, the block on tools to reinforce resilience at the organizational level provided strategies to assess and achieve resilient institutions in the face of crises and unexpected events. In addition, a road map was included at the end of the course to aid in selecting the most useful strategies based on the objectives.

XSL•FO
RenderX

Figure 3. The home screen of the eResiliencia web-based platform.

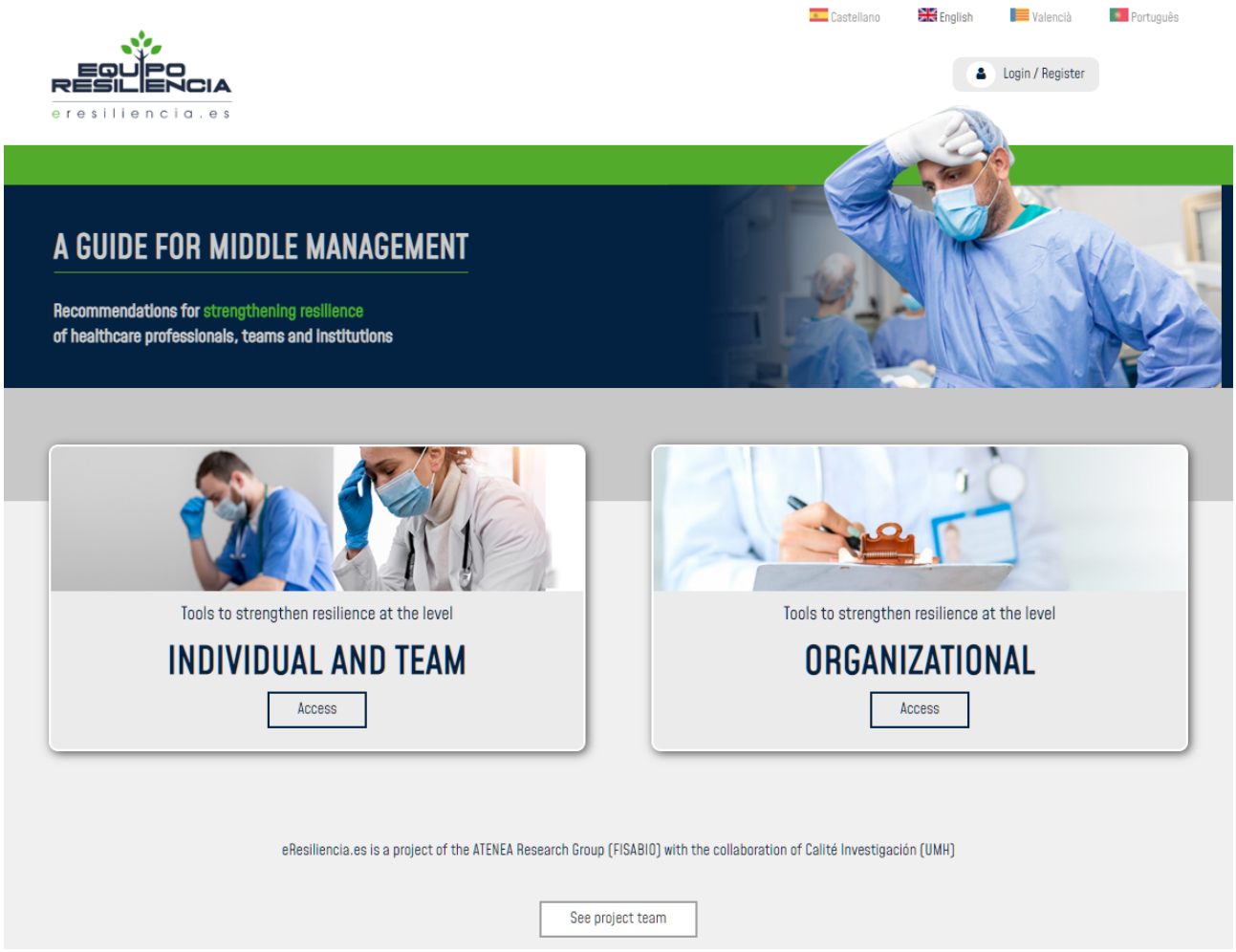
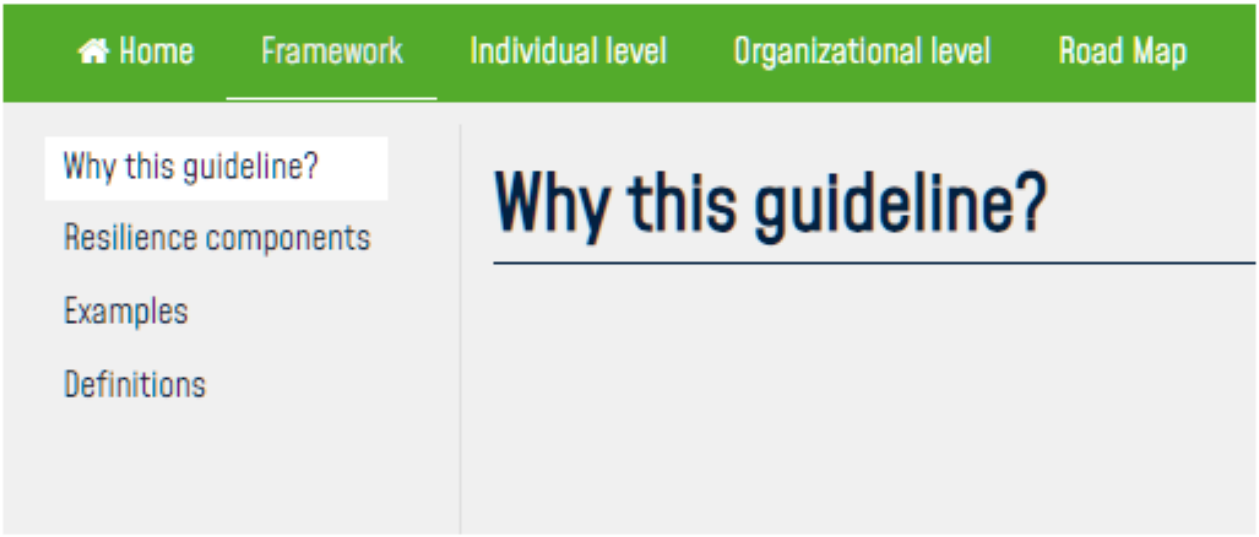


Figure 4. Navigation panel of the eResiliencia web-based platform.



XSL•FO
RenderX

Table 2. Structure of the eResiliencia web-based platform.

Structure	Items
Framework	• Why this guideline? • Resilience components • Examples • Definitions
Individual level	• Stress Continuum Model • Enhancing resilience at the individual level • G.R.A.C.E. model • PERMA+4 Well-Being model for occupational wellness • Behavior change technique for building work-related well-being
Organizational level	• Enhancing resilience at the organizational level • Resilience Analysis Grid

In total, it comprised 17 videos, 11 figures and images, 4 PDF files, and references for further information expansion, among which downloadable materials or video, graph, and text links were also available. Upon accessing their personal account, users were provided with a section containing pertinent information regarding the platform's functionality, along with a video elucidating navigation and content. In addition, users had access to a support resource and the option to change their password.

Phase 2: Validation by Experts

Of the 20 experts, 13 (65%) assessed the tool and provided feedback. They rated the clarity of the content at an average of 4.5 (SD 0.7) and the utility and applicability at an average of 4.7 (SD 0.5), out of 5 in both cases. Regarding whether the materials facilitated learning strategies to foster the resilience of professionals or teams under their purview, 62% (8/13) of the experts indicated they *completely agree*, 31% (4/13) of the experts indicated they *agree*, and 8% (1/13) of the experts indicated they *completely disagree*. Regarding their likelihood to recommend the tool to other colleagues, 92% (12/13) of the experts responded *definitely yes*, and 8% (1/13) of the experts responded *probably yes*. The same results were obtained for their intention to apply learned tools or strategies. The level of agreement among experts for the combined set of the 5 evaluative variables of eResiliencia was 0.79 (95% CI 0.65-0.92).

Overall satisfaction was rated with an average of 9.3 (SD 0.6) out of 10 points. Some of the strengths highlighted were the chosen content and the ease of navigation and presentation of the web-based platform. In contrast, areas for improvement included making the videos more engaging and the possibility of incorporating practical exercises or a forum for participant interaction. As a suggestion, an introductory video of the content and the functionality of the web-based platform was included (Multimedia Appendix 2) [60]. The full list of comments made by experts is presented in Multimedia Appendix 3.

Phase 3: Pilot Study

Of the 458 professionals invited to participate, 362 (79%) registered on the platform. The preintervention and postintervention questionnaires of the pilot study were completed by 218 (60.2%) of these 362 health care professionals exercising leadership roles who began the intervention. Moreover, 146 (40.3%) health care professionals responded to the experience evaluation questionnaire.

Most of them were physicians (90/218, 41.3%) and nurses (59/218, 27.1%), with a mean experience of 10.3 (SD 8.8) years. Their primary work setting was primary care (113/218, 51.8%), hospital (79/218, 36.2%), quality area (9/218, 4.1%), or other management positions (7/218, 3.2%). The profile of those who did not complete the entire educational intervention was similar to those who did, with 42.4% (61/144 being physicians, 24.3% (35/144) being nurses, and 45.1% (65/144) working in primary care. Those who did not complete the intervention had slightly average experience, with 11.4 years compared with 10 years.

The scores obtained in the preintervention questionnaire averaged 42.0 (SD 16.6) of 100, while for the postintervention questionnaire, they averaged 74.7 (SD 22.7), with significant differences between them (P<.001; Z=–11.728). The increase in the percentage of correct answers for each question is represented in Table 3.

XSL•FO
RenderX

Table 3. Increase in percentage of correct answers for each question (n=218).

Questions[a]	Correct responses on the preintervention questionnaire, n (%)	Correct responses on the postintervention questionnaire, n (%)	Increase of correct answers (%)	McNemar test	*P* value
1	148 (67.9)	192 (88.1)	20.2	25.681	<.001
2	73 (33.5)	148 (67.9)	34.4	53.165	<.001
3	102 (46.8)	155 (71.1)	24.3	27.876	<.001
4	32 (14.7)	134 (61.5)	46.8	87.940	<.001
5	103 (47.2)	165 (75.7)	28.4	37.969	<.001
6	66 (30.3)	176 (80.7)	50.5	99.008	<.001
7	92 (42.2)	162 (74.3)	32.1	52.900	<.001
8	110 (50.5)	180 (82.6)	32.1	48.582	<.001
9	94 (43.1)	153 (70.2)	27.1	30.862	<.001
10	82 (37.6)	159 (72.9)	35.3	55.010	<.001
11	139 (63.8)	189 (86.7)	22.9	29.280	<.001
12	117 (53.7)	169 (77.5)	23.8	28.900	<.001
13	44 (20.2)	165 (75.7)	55.5	111.628	<.001
14	110 (50.5)	181 (83)	32.5	49.495	<.001
15	29 (13.3)	111 (50.9)	37.6	64.324	<.001
16	86 (39.4)	145 (66.5)	27.1	33.307	<.001
17	138 (63.3)	194 (89)	25.7	34.375	<.001
18	114 (52.3)	154 (70.6)	18.3	20.554	<.001
19	84 (38.5)	162 (74.3)	35.8	55.934	<.001
20	68 (31.2)	163 (74.8)	43.6	73.025	<.001

[a]Because this is active training, the questions have been masked to avoid spoilers. They are available upon reasonable request.

Moreover, 58.3% (127/218) of the health care professionals passed the postintervention questionnaire without requiring any retries. However, 14.2% (31/218) of the health care professionals needed 1 retry, 8.3% (18/218) needed 2 retries, 6% (13/218) needed 3 retries, and 3.7% (8/218) needed between 4 and 7 retries. Furthermore, 9.6% (21/218) of the health care professionals exhausted all 10 retry attempts without achieving the required score.

Regarding the perception of the 146 people who completed the evaluation questionnaire, the clarity of the content received an average score of 4.7 (SD 0.6) on a 1 to 5 scale, while the utility or applicability scored 4.8 (SD 0.5). The results of their opinion are depicted in Table 4.

Global satisfaction was rated 9.4 (SD 1.1) out of 10 points on average. Among the strengths underscored were the comprehensiveness of the content and the interesting techniques explained. Conversely, identified improvement areas consisted of incorporating more illustrative examples and providing downloadable materials.

XSL•FO
RenderX

Table 4. Results of the nonnumerical scale questions from the assessment questionnaire (n=146).

Question	Completely disagree, n (%)	Disagree, n (%)	Neither agree nor disagree, n (%)	Agree, n (%)	Completely agree, n (%)	Definitely not, n (%)	Probably not, n (%)	Not sure, n (%)	Probably yes, n (%)	Definitely yes, n (%)
The course content has enabled me to learn strategies for reinforcing the resilience of the professionals or teams under my supervision.	7 (4.8)	1 (0.7)	6 (4.1)	54 (37)	78 (53.4)	—[a]	—	—	—	—
I will recommend the course to other colleagues.	—	—	—	—	—	0 (0)	0 (0)	6 (4.1)	22 (15.1)	118 (80.8)
I will apply some of the tools or strategies learned to strengthen individual, team, or organizational resilience in my work environment.	—	—	—	—	—	0 (0)	2 (1.4)	2 (1.4)	28 (19.2)	114 (78.1)

[a]Not applicable.

Qualitative Analysis of Open-Ended Comments

Strengths

The most repeated topics among the strengths were the learning of new techniques or specific models (34/146, 23.3%), the web-based methodology that allowed access at any time (14/146, 9.6%), the organizational approach (10/146, 6.8%), the clarity of the content (9/146, 6.1%), stress management (6/146, 4.1%), and the examples provided (4/146, 2.7%).

Opportunities for Improvement

In terms of aspects to be improved, the incorporation of more examples (14/146, 9.6%) and the expansion of the content (14/146, 9.6%) were generally highlighted. In addition, improving the quality of the videos (12/146, 8.2%), including the possibility of interacting with other students in a more practical way (9/146, 6.2%), enabling downloadable content (9/146, 6.2%), reducing the length of the questionnaires (6/146, 4.1%), including web-based tutorials for resolving doubts with the teachers (5/146, 3.4%), and adding a progress bar to visualize the course as a whole (4/146, 2.7%) were emphasized.

Discussion

Principal Findings

Our developed web-based approach operates as a comprehensive tool designed primarily for middle managers, aiming to foster resilience at both individual and organizational levels within the staff at their charge. By using multimedia resources, this approach has been shown to enhance the knowledge about resilience of health care professionals across diverse fields and backgrounds.

eResiliencia was highly rated by the panel of experts regarding clarity and usefulness, indicating that the tool is perceived as comprehensible and relevant for application in professional contexts. High ratings were also obtained for the willingness to recommend the tool to colleagues and the intention to apply the learned strategies, demonstrating strong acceptance and confidence in the resource. For most (12/13, 92%) participants, the materials were potentially effective in strengthening team resilience, except for 1 (8%) expert who expressed complete disagreement. This highlights the need to consider diverse perspectives and possible improvements in adapting the content to various contexts. Nonetheless, overall satisfaction was high, with more than 9 points out of 10.

Implementing eResiliencia in Real-World Settings

The eResiliencia platform features a user-friendly interface, and in most cases, managerial staff possesses the necessary skills to use such digital tools effectively. However, some systemic limitations may impact its implementation. First, the increasing measures to prevent cyberattacks and data ransom incidents in health care institutions have led to stricter internet access controls, even in areas where connectivity was previously unrestricted. This decision may hinder access to eResiliencia within some organizations. In addition, in certain health care networks, managers lack the resources to access the platform outside their offices. Incorporating offline functionality could address these limitations and facilitate broader access to the platform.

Second, resistance from health care organizations, often rooted in established workflows or skepticism about the benefits of new interventions, represents a traditional barrier. In this case, the integration of eResiliencia was facilitated by embedding it within a national patient safety and quality training program aimed at launching a global institutional strengthening plan. A similar strategy could help overcome organizational resistance in other settings.

XSL•FO
RenderX

Third, ensuring flexibility and accessibility is essential to meet the diverse needs of users. The use of asynchronous learning options was key in this instance, allowing participants to decide when and how to engage with the platform. This flexibility significantly increases accessibility and is particularly relevant in environments with high workloads. However, allocating dedicated time during work hours could further enhance participation rates by reducing competing demands on participants' schedules.

Comparison With Prior Studies

It is known that health care professionals face stressful situations in their professional work. The frequency of burnout or the existence of defensive practices is also not uncommon. It is no surprise that the role of team leaders influences the team's ability to respond and recover after experiencing adverse situations [61]. The COVID-19 pandemic brought to the attention of managers, executives, and society in general concerns about the resilience capacity of these collectives, both at an individual and collective level, to face what has undoubtedly been the greatest global challenge for health care professionals. This study originates from that period and delves into both aspects, offering an easily accessible training format for those responsible for clinical teams. In this regard, it is based on the observation that most programs designed to increase resilience were developed to operate at the individual level and that it was necessary to create interventions focused on the work group [30].

eResiliencia incorporates agreed-upon resilience attributes [37] and skills that leaders must possess to strengthen their teams' resilience, which include self-efficacy, self-control, the ability to provide support and help, and learning from difficulties [62].

Uncertainty, resource shortages, constant protocol variations, the risk of infection and transmission, and critical decision-making with ethical implications, among other factors, were prevalent during the COVID-19 pandemic and contributed to fear, anxiety, stress, moral injury, or compassion fatigue among health care professionals [63,64]. These situations have highlighted the need to reconsider the role of middle management, as they must address the needs of their teams to face challenges. Therefore, their resilience is a crucial competence required for their work.

The scores obtained in this study by the professionals who underwent the training show that after the use of the tool, the results were better both overall and for each individual question. However, the preintervention questionnaire also revealed that participants' previous knowledge was relatively low in many areas, with correct response rates ranging from 13.3% (question 15; behavior change technique) to 67.9% (question 1; definition of resilience). This underscores the necessity and effectiveness of the educational intervention implemented. The greatest increases were in questions with initially low correct response rates (question 4, question 6, and question 13), suggesting the acquisition of new concepts. In contrast, questions such as question 1 and question 18, which initially had a higher percentage of correct responses, showed a less pronounced increase compared to those with lower initial performance. The increase in scores on the final questionnaire compared to the initial questionnaire is consistent with the findings of other studies, which show that both in-person [52,65-67] and on the web [50,68] interventions increase resilience capacity. In addition, it has been observed that they also protect the mental health of professionals [69].

Solutions to assist professionals were implemented in different countries and were similar, including the availability of psychological services, support hotlines, and websites with resource materials [70]. Focusing on tools available for Spanish speakers, an example is the digital platform and mobile app, Be+Against COVID, which was used in Argentina, Brazil, Colombia, Chile, Ecuador, and Spain [71].

Although middle managers have usually received training in occupational risk prevention and conflict management, many times there is a lack of specific and detailed training in fundamental aspects of team management and resilience. Traditionally, team management has focused on organizational and clinical aspects to ensure that patients' needs are adequately met. Nowadays, middle managers are expected to efficiently manage the resources of the professional teams they lead, particularly human talent, to adequately respond to health challenges.

Implications of Findings

In the postpandemic era, there is a growing consensus that among the functions of executives and middle managers (such as the organization of responses to care demands, task allocation, conflict management within teams, and the strengthening of teams to meet care demands effectively), there should be an emphasis on maintaining and strengthening the resilience capacity of their teams of professionals, because their ability to cope with highly stressful situations determines the quality of care provided and the outcomes for patients. This eResiliencia program follows this trend and takes a proactive approach by offering a framework for executives and middle managers to use some or all the resources to achieve this objective and prepare their team for everyday challenges and future crises.

Although most (197/218, 90.4%) of the participants were able to pass the training without any difficulties, some (21/218, 9.6%) individuals exhausted all reattempts without achieving the minimum required score. While these data suggest that the material and teaching methods are effective for most, it would be beneficial to delve into the reasons behind the difficulties faced by this small percentage to address and mitigate the problem.

Resilience training is presented as a crucial element in the professional development of middle managers in the health care sector. Promoting and reinforcing resilience is not only essential for the individual well-being of professionals but also strengthens the cohesion and effectiveness of the team as a whole. However, for this to be feasible, it is essential that clinical management evolves and provides effective tools to middle managers, facilitating resource management, including resilience training.

In this era, there is a growing consensus that the existing "find and fix" solutions should be replaced by those that are more forward-thinking [15]. Moreover, the organizational culture is

XSL•FO
RenderX

critical to ensuring patients receive adequate care and workers' resilience when things go wrong [72]. Health care organizations are evolving toward the principles of Safety II framework here [73], which involves recognizing the challenge of complexity and uncertainty in clinical practice to anticipate errors. Promoting the resilience of workers at both individual and team levels contributes to realizing the implementation of Safety II [74].

Fragmentation in the assignment of professionals poses a significant challenge to teamwork. Thus, cohesion among middle medical and nursing managers is essential to overcoming the barriers created by this fragmentation and fostering a more collaborative and efficient work environment. The training of middle managers in the health care organization must be a top-down strategy, driven by leadership. Health care executives must communicate these priorities to middle managers and provide the necessary support for implementation. Only through strong and consistent leadership from the upper echelons can we ensure that middle managers are equipped with the competencies required to effectively manage their teams and promote a resilient work environment.

The response rate indicates that 4 (40%) of the 10 professionals who registered on the platform were eventually not interested in the training on how to foster resilience in their teams and collaborators. This could be because they already possess the personal resources to tackle the task or because eResiliencia did not meet their expectations or failed to motivate them sufficiently. This aspect could be addressed, for example, with a more practical approach through group dynamics and cooperative tasks to generate greater interest and engagement, so participants can see a more direct application of the training tools.

Strengths

This study addresses the 5 key questions of a resilience study in the health care sector [75]. It proposes an alternative approach to individual-focused training by emphasizing the role of clinical team leaders and how to strengthen the resilience of their teams and team members.

Limitations

Although international experts share a common cultural background, they represent a group of Spanish-speaking countries. The generalization of content to other countries should not be done directly. Participants in the pilot study phase worked in various public health networks in Ecuador and took part in this experience as part of a quality and patient safety management course, which emphasized the need to move away from a blame culture. This aspect may have influenced their engagement with the eResiliencia content. Participants had a set amount of time to complete the training on the eResiliencia web-based platform. In a different self-directed learning context without this limitation, trainee behavior might be different. Because this is a study without a control group, the changes observed in participants' knowledge may not be exclusively attributable to eResiliencia. Finally, it should be noted that although the participants were enrolled in training organized by the Ministry of Public Health of Ecuador, it was emphasized that they could freely choose whether to complete the eResiliencia training and respond to the satisfaction survey. Answering the questionnaire did not provide any advantage and the responses were not disclosed, ensuring anonymity at all times, which was clearly communicated to the participants. This approach was taken to avoid a reporting bias.

Cultural Factors Influencing Resilience-Building Capacity

Ecuador's health care system is characterized by hierarchical structures and centralized decision-making, typical of high-power distance cultures. eResiliencia capitalizes on the influential role of middle managers, who act as intermediaries between upper management and frontline workers. By equipping middle managers with resilience-building tools, the intervention empowers them to cascade these strategies throughout their teams, reinforcing resilience at all levels of the organization. The networks engaged in this study operate with limited resources, requiring flexibility and improvisation to address challenges. eResiliencia provides structured yet adaptable strategies, offering tools for middle managers to tailor resilience-building efforts according to the specific needs and constraints of their teams. Although Ecuador faces challenges in long-term health care planning, eResiliencia introduces structured frameworks that emphasize proactive resilience building. These frameworks align with preventive measures to prepare health care professionals for crises, fostering a shift from reactive to proactive management.

Many of these aspects are common to countries in this region. However, the outcomes of eResiliencia in Asia, Europe, or North America are likely to differ significantly due to the cultural factors influencing resilience-building capacity. Cultural dimensions such as collectivism, power distance, and approaches to uncertainty, among others, play a crucial role in shaping how resilience strategies are developed and applied across different contexts.

Future Research

It is necessary to delve into a common definition of what is meant by resilience [11,75] and, by extension, the role of managers in the resilience of their team members. A critical aspect for future development involves the incorporation of mechanisms to monitor the utility and long-term impact of the knowledge acquired through the web-based tool, particularly considering the less optimistic data provided by some studies on nursing students [76]. In later phases, this intervention will be complemented by the development of improvement and action plans based on the application of the eResiliencia tools by health care teams and middle management interested in their centers. This will allow a more effective monitoring of the tool, its usefulness, and applicability. These approaches could also be extended to companies that provide home care, to train those who coordinate the work of informal caregivers in the home [77], in light of the growing caregiving economy. Finally, as these are preliminary data, this educational intervention should be tested in the context of other organizational models and diverse cultural settings to validate its effectiveness, acceptability, and utility.

XSL•FO
RenderX

Conclusions

The proposal to evaluate each professional, identify signs of stress, take appropriate measures, and manage team cohesion represents a training initiative that could be implemented clearly and explicitly. However, it is essential to first analyze the organizational context where the intervention will be applied and identify potential adjustments to improve its acceptability based on national and local organizational factors. This tailored approach ensures that the intervention aligns with the specific needs and characteristics of each setting, enhancing its effectiveness and relevance. This training would enable middle managers not only to manage conflicts but also to prevent them and foster a healthy and productive work environment. The developed tool proved to be well received by both the experts and the health care professionals who participated in this pilot phase.

Acknowledgments

The authors would like to acknowledge the experts who participated in the evaluation and validation of the web-based platform as well as all the participants from Ecuador who took part in the pilot study, thereby contributing to an improved experience with the platform. The authors also extend their gratitude to the National Directorate of Quality, Patient Safety, and Health Control of the Ministry of Public Health of Ecuador for their support and dissemination efforts in the initial recruitment process. Clara Pérez contributed to the statistical analyses.

The project was funded by Conselleria de Sanitat Universal i Salut Pública (Generalitat Valenciana, Spain) and the EU Operational Program of the European Regional Development Fund for the Valencian Community 2014-2020, within the framework of the Recovery Assistance for Cohesion and the Territories of Europe program, as the European Union’s response to the COVID-19 pandemic; Ser+ Contra COVID (UGP-21-408). The maintenance of the web-based platform is made possible by the funds allocated to the preparatory action in the UniSalut Program 2023 (subprogram ILISABIO, reference ILISABIO22_AP09). Throughout the execution of this study and composition of this manuscript, JJM benefited from an augmented research activity contract granted by the Carlos III Health Institute (reference INT22/00012).

Data Availability

The datasets generated analyzed during this study are available from the corresponding author on reasonable request. Topics addressed in each question are displayed in Multimedia Appendix 1.

Authors' Contributions

All the authors meet the International Committee of Medical Journal Editors criteria for authorship. JJM was responsible for the design of the study. JJM and IC conducted the search and the preliminary selection of contents related to resilience. IC and EGH drafted the web-based platform and developed the materials. JMD was responsible for the enrollment of people and the acquisition of data. DGT performed the statistical analysis and interpretation of the results. EGH developed the first version of the manuscript. All authors revised the paper critically for important intellectual content and read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Topics of the preintervention and postintervention questionnaire questions.
[DOC File , 132 KB - humanfactors_v12i1e67263_app1.doc]

Multimedia Appendix 2
eResiliencia introductory video.
[MP4 File (MP4 Video), 189408 KB - humanfactors_v12i1e67263_app2.mp4]

Multimedia Appendix 3
List of comments received from the experts during the validation phase.
[DOC File , 152 KB - humanfactors_v12i1e67263_app3.doc]

References

1. Cantu J, Gharehyakheh A, Fritts S, Tolk J. Assessing the HRO: tools and techniques to determine the high-reliability state of an organization. Saf Sci 2021 Feb;134:105082. [doi: 10.1016/j.ssci.2020.105082]
2. Schulman PR. Reliability, uncertainty and the management of error: new perspectives in the COVID‐19 era. Conting Crisis Mgmt 2021 Mar 12;30(1):92-101. [doi: 10.1111/1468-5973.12356]

XSL•FO
RenderX

3. Jacobs B, McGovern J, Heinmiller J, Drenkard K. Engaging employees in well-being: moving from the triple aim to the quadruple aim. Nurs Adm Q 2018;42(3):231-245. [doi: 10.1097/NAQ.0000000000000303] [Medline: 29870489]
4. Søvold LE, Naslund JA, Kousoulis AA, Saxena S, Qoronfleh MW, Grobler C, et al. Prioritizing the mental health and well-being of healthcare workers: an urgent global public health priority. Front Public Health 2021 May 07;9:679397 [FREE Full text] [doi: 10.3389/fpubh.2021.679397] [Medline: 34026720]
5. Schiess C, Schwappach D, Schwendimann R, Vanhaecht K, Burgstaller M, Senn B. A transactional "second-victim" model-experiences of affected healthcare professionals in acute-somatic inpatient settings: a qualitative metasynthesis. J Patient Saf 2021 Dec 01;17(8):e1001-e1018. [doi: 10.1097/PTS.0000000000000461] [Medline: 29384831]
6. Zheng J, Lu Y, Li W, Zhu B, Yang F, Shen J. Prevalence and determinants of defensive medicine among physicians: a systematic review and meta-analysis. Int J Qual Health Care 2023 Dec 07;35(4):mzad096. [doi: 10.1093/intqhc/mzad096] [Medline: 38060672]
7. Bourne T, Shah H, Falconieri N, Timmerman D, Lees C, Wright A, et al. Burnout, well-being and defensive medical practice among obstetricians and gynaecologists in the UK: cross-sectional survey study. BMJ Open 2019 Nov 25;9(11):e030968 [FREE Full text] [doi: 10.1136/bmjopen-2019-030968] [Medline: 31767585]
8. Schwappach DL, Boluarte TA. The emotional impact of medical error involvement on physicians: a call for leadership and organisational accountability. Swiss Med Wkly 2009 Jan 10;139(1-2):9-15 [FREE Full text] [doi: 10.4414/smw.2009.12417] [Medline: 18951201]
9. Mira J, Carillo I, Tella S, Vanhaecht K, Panella M, Seys D, et al. The European Researchers' Network Working on Second Victim (ERNST) policy statement on the second victim phenomenon for increasing patient safety. Public Health Rev 2024 Sep 18;45:1607175 [FREE Full text] [doi: 10.3389/phrs.2024.1607175] [Medline: 39360222]
10. Bodenheimer T, Sinsky C. From triple to quadruple aim: care of the patient requires care of the provider. Ann Fam Med 2014;12(6):573-576 [FREE Full text] [doi: 10.1370/afm.1713] [Medline: 25384822]
11. Sisto A, Vicinanza F, Campanozzi LL, Ricci G, Tartaglini D, Tambone V. Towards a transversal definition of psychological resilience: a literature review. Medicina (Kaunas) 2019 Nov 16;55(11):745 [FREE Full text] [doi: 10.3390/medicina55110745] [Medline: 31744109]
12. Wiig S, Aase K, Billett S, Canfield C, Røise O, Njå O, et al. Defining the boundaries and operational concepts of resilience in the resilience in healthcare research program. BMC Health Serv Res 2020 Apr 19;20(1):330 [FREE Full text] [doi: 10.1186/s12913-020-05224-3] [Medline: 32306981]
13. Watson P, Westphal R. Stress first aid for health care workers. National Center for PTSD. 2020. URL: https://www.ptsd.va.gov/professional/treat/type/SFA/docs/SFA_HCW_Manual_508.pdf [accessed 2025-01-13]
14. Global patient safety action plan 2021-2030: towards eliminating avoidable harm in health care. World Health Organization. 2021. URL: https://www.google.co.in/books/edition/Global_patient_safety_action_plan_2021_2/csZqEAAAQBAJ?hl=en [accessed 2025-01-13]
15. Haraldseid-Driftland C, Lyng HB, Guise V, Waehle HV, Schibevaag L, Ree E, et al. Learning does not just happen: establishing learning principles for tools to translate resilience into practice, based on a participatory approach. BMC Health Serv Res 2023 Jun 16;23(1):646 [FREE Full text] [doi: 10.1186/s12913-023-09653-8] [Medline: 37328864]
16. Haraldseid-Driftland C, Dombestein H, Le AH, Billett S, Wiig S. Learning tools used to translate resilience in healthcare into practice: a rapid scoping review. BMC Health Serv Res 2023 Aug 23;23(1):890 [FREE Full text] [doi: 10.1186/s12913-023-09922-6] [Medline: 37612671]
17. Seys D, Panella M, Russotto S, Strametz R, Joaquín Mira J, van Wilder A, et al. In search of an international multidimensional action plan for second victim support: a narrative review. BMC Health Serv Res 2023 Jul 31;23(1):816 [FREE Full text] [doi: 10.1186/s12913-023-09637-8] [Medline: 37525127]
18. Nash WP. US Marine Corps and Navy combat and operational stress continuum model: a tool for leaders. In: Combat and Operational Behavioral Health. North Charleston, SC: CreateSpace Independent Publishing Platform; 2016.
19. Donaldson SI, van Zyl LE, Donaldson SI. PERMA+4: a framework for work-related wellbeing, performance and positive organizational psychology 2.0. Front Psychol 2022 Jan 24;12:817244 [FREE Full text] [doi: 10.3389/fpsyg.2021.817244] [Medline: 35140667]
20. Halifax J. G.R.A.C.E. for nurses: cultivating compassion in nurse/patient interactions. J Nurs Educ Pract 2014;4(1):121-128. [doi: 10.5430/JNEP.V4N1P121]
21. Hollnagel E, Pàries J, Woods DD, Wreathall J. Resilience Engineering in Practice: A Guidebook. Farnham, UK: Ashgate Publishing; 2011.
22. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95 [FREE Full text] [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
23. Winwood PC, Colon R, McEwen K. A practical measure of workplace resilience: developing the resilience at work scale. J Occup Environ Med 2013 Oct;55(10):1205-1212. [doi: 10.1097/JOM.0b013e3182a2a60a] [Medline: 24064782]
24. Resilience test. Psychometric Test. 2013 Apr 29. URL: https://www.psychometrictest.org.uk/resilience-test/ [accessed 2022-05-04]

XSL•FO
RenderX

25. Smith BW, Dalen J, Wiggins K, Tooley E, Christopher P, Bernard J. The brief resilience scale: assessing the ability to bounce back. Int J Behav Med 2008;15(3):194-200. [doi: 10.1080/10705500802222972] [Medline: 18696313]
26. Connor KM, Davidson JR. Development of a new resilience scale: the Connor-Davidson Resilience Scale (CD-RISC). Depress Anxiety 2003;18(2):76-82. [doi: 10.1002/da.10113] [Medline: 12964174]
27. Friborg O, Hjemdal O, Rosenvinge JH, Martinussen M. A new rating scale for adult resilience: what are the central protective resources behind healthy adjustment? Int J Methods Psychiatr Res 2003;12(2):65-76 [FREE Full text] [doi: 10.1002/mpr.143] [Medline: 12830300]
28. Wagnild GM, Young HM. Development and psychometric evaluation of the Resilience Scale. J Nurs Meas 1993;1(2):165-178. [Medline: 7850498]
29. Bartone PT, Ursano RJ, Wright KM, Ingraham LH. The impact of a military air disaster on the health of assistance workers. A prospective study. J Nerv Ment Dis 1989 Jun;177(6):317-328. [doi: 10.1097/00005053-198906000-00001] [Medline: 2723619]
30. Troy AS, Willroth EC, Shallcross AJ, Giuliani NR, Gross JJ, Mauss IB. Psychological resilience: an affect-regulation framework. Annu Rev Psychol 2023 Jan 18;74:547-576 [FREE Full text] [doi: 10.1146/annurev-psych-020122-041854] [Medline: 36103999]
31. Bowling J, Vercruysse C, Krinner LM, Greene T, Bello-Ogunu F, Webster C. A simultaneous concept analysis of resilience, coping, posttraumatic growth, and thriving. Nurs Forum 2022 Sep;57(5):905-919. [doi: 10.1111/nuf.12754] [Medline: 35643843]
32. Moya J, Goenechea M. An approach to the unified conceptualization, definition, and characterization of social resilience. Int J Environ Res Public Health 2022 May 09;19(9):5746 [FREE Full text] [doi: 10.3390/ijerph19095746] [Medline: 35565141]
33. Métais C, Burel N, Gillham JE, Tarquinio C, Martin-Krumm C. Integrative review of the recent literature on human resilience: from concepts, theories, and discussions towards a complex understanding. Eur J Psychol 2022 Feb 25;18(1):98-119 [FREE Full text] [doi: 10.5964/ejop.2251] [Medline: 35330859]
34. Daly LM. Resilience: an integrated review. Nurs Sci Q 2020 Oct;33(4):330-338. [doi: 10.1177/0894318420943141] [Medline: 32940160]
35. Tay PK, Lim KK. Psychological resilience as an emergent characteristic for well-being: a pragmatic view. Gerontology 2020;66(5):476-483. [doi: 10.1159/000509210] [Medline: 32784301]
36. Infurna FJ. What does resilience signify? An evaluation of concepts and directions for future research. Gerontology 2020;66(4):323-331. [doi: 10.1159/000507365] [Medline: 32408298]
37. Cooper AL, Brown JA, Rees CS, Leslie GD. Nurse resilience: a concept analysis. Int J Ment Health Nurs 2020 Aug;29(4):553-575. [doi: 10.1111/inm.12721] [Medline: 32227411]
38. Huey CW, Palaganas JC. What are the factors affecting resilience in health professionals? A synthesis of systematic reviews. Med Teach 2020 May;42(5):550-560. [doi: 10.1080/0142159X.2020.1714020] [Medline: 31984844]
39. Stainton A, Chisholm K, Kaiser N, Rosen M, Upthegrove R, Ruhrmann S, et al. Resilience as a multimodal dynamic process. Early Interv Psychiatry 2019 Aug;13(4):725-732. [doi: 10.1111/eip.12726] [Medline: 30126047]
40. Lyng HB, Macrae C, Guise V, Haraldseid-Driftland C, Fagerdal B, Schibevaag L, et al. Exploring the nature of adaptive capacity for resilience in healthcare across different healthcare contexts; a metasynthesis of narratives. Appl Ergon 2022 Oct;104:103810 [FREE Full text] [doi: 10.1016/j.apergo.2022.103810] [Medline: 35635941]
41. Raetze S, Duchek S, Maynard MT, Wohlgemuth M. Resilience in organization-related research: an integrative conceptual review across disciplines and levels of analysis. J Appl Psychol 2022 Jun;107(6):867-897. [doi: 10.1037/apl0000952] [Medline: 34766796]
42. Vercio C, Loo LK, Green M, Kim DI, Beck Dallaghan GL. Shifting focus from burnout and wellness toward individual and organizational resilience. Teach Learn Med 2021;33(5):568-576 [FREE Full text] [doi: 10.1080/10401334.2021.1879651] [Medline: 33588654]
43. Biddle L, Wahedi K, Bozorgmehr K. Health system resilience: a literature review of empirical research. Health Policy Plan 2020 Oct 01;35(8):1084-1109 [FREE Full text] [doi: 10.1093/heapol/czaa032] [Medline: 32529253]
44. Bhandari S, Alonge O. Measuring the resilience of health systems in low- and middle-income countries: a focus on community resilience. Health Res Policy Syst 2020 Jul 17;18(1):81 [FREE Full text] [doi: 10.1186/s12961-020-00594-w] [Medline: 32680527]
45. Strametz R, Raspe M, Ettl B, Huf W, Pitz A. Recommended actions: reinforcing clinicians' resilience and supporting second victims during the COVID-19 pandemic to maintain capacity in the healthcare system [Article in German]. Zentralbl Arbeitsmed Arbeitsschutz Ergon 2020;70(6):264-268 [FREE Full text] [doi: 10.1007/s40664-020-00405-7] [Medline: 32905028]
46. Wu AW, Connors C, Everly GSJ. COVID-19: peer support and crisis communication strategies to promote institutional resilience. Ann Intern Med 2020 Jun 16;172(12):822-823 [FREE Full text] [doi: 10.7326/M20-1236] [Medline: 32251512]
47. Thomas S, Sagan A, Larkin J, Cylus J, Figueras J, Karanikolos M. Strengthening Health Systems Resilience: Key Concepts and Strategies [Internet]. Copenhagen, Denmark: European Observatory on Health Systems and Policies; 2020.

XSL•FO
RenderX

48. Turenne CP, Gautier L, Degroote S, Guillard E, Chabrol F, Ridde V. Conceptual analysis of health systems resilience: a scoping review. Soc Sci Med 2019 Jul;232:168-180. [doi: 10.1016/j.socscimed.2019.04.020] [Medline: 31100697]
49. Qureshi A, Hargest C, Swain N, Aldabe D, Hale L. Psychosocial interventions for building resilience of informal carers of people living with stroke: a systematic review. Disabil Rehabil 2023 May;45(9):1419-1432. [doi: 10.1080/09638288.2022.2063419] [Medline: 35468030]
50. Stoliker BE, Vaughan AD, Collins J, Black M, Anderson GS. Building personal resilience following an online resilience training program for BScN students. West J Nurs Res 2022 Aug;44(8):755-764. [doi: 10.1177/01939459211017240] [Medline: 33998338]
51. Golden EA, Zweig M, Danieletto M, Landell K, Nadkarni G, Bottinger E, et al. A resilience-building app to support the mental health of health care workers in the COVID-19 era: design process, distribution, and evaluation. JMIR Form Res 2021 May 05;5(5):e26590 [FREE Full text] [doi: 10.2196/26590] [Medline: 33872189]
52. Spiva L, Davis S, Case-Wirth J, Hedenstrom L, Hogue V, Box M, et al. The effectiveness of charge nurse training on leadership style and resiliency. J Nurs Adm 2020 Feb;50(2):95-103. [doi: 10.1097/NNA.0000000000000848] [Medline: 31977947]
53. Liu JJ, Ein N, Gervasio J, Battaion M, Reed M, Vickers K. Comprehensive meta-analysis of resilience interventions. Clin Psychol Rev 2020 Dec;82:101919. [doi: 10.1016/j.cpr.2020.101919] [Medline: 33045528]
54. Venegas CL, Nkangu MN, Duffy MC, Fergusson DA, Spilg EG. Interventions to improve resilience in physicians who have completed training: a systematic review. PLoS One 2019 Jan 17;14(1):e0210512 [FREE Full text] [doi: 10.1371/journal.pone.0210512] [Medline: 30653550]
55. Franco PL, Christie LM. Effectiveness of a one day self-compassion training for pediatric nurses' resilience. J Pediatr Nurs 2021;61:109-114. [doi: 10.1016/j.pedn.2021.03.020] [Medline: 33839602]
56. Donaldson SI, Donaldson SI. The positive functioning at work scale: psychometric assessment, validation, and measurement invariance. J Well Being Assess 2021 Jan 11;4:181-215. [doi: 10.1007/s41543-020-00033-1]
57. Butler J, Kern ML. The PERMA-Profiler: a brief multidimensional measure of flourishing. Int J Wellbeing 2016;6(3):1-48. [doi: 10.5502/ijw.v6i3.526]
58. Scott SD, Hirschinger LE, Cox KR, McCoig M, Hahn-Cover K, Epperly KM, et al. Caring for our own: deploying a systemwide second victim rapid response team. Jt Comm J Qual Patient Saf 2010 May;36(5):233-240. [doi: 10.1016/s1553-7250(10)36038-7] [Medline: 20480757]
59. Una guía para mandos intermedios: recomendaciones para reforzar la resiliencia de los profesionales, los equipos y las instituciones sanitarias. Equipo Resiliencia. URL: https://www.eresiliencia.es/es/ [accessed 2025-01-13]
60. Presentació eResiliencia. YouTube. URL: https://www.youtube.com/watch?v=8P_p3e6MyVU [accessed 2025-01-13]
61. Herrin S. Examining relationships among leadership behaviors, demographic and program factors, and resilience in physical therapy professional education program directors. J Phys Ther Educ 2023 Dec 01;37(4):342-350 [FREE Full text] [doi: 10.1097/JTE.0000000000000310] [Medline: 38478790]
62. Howe A, Smajdor A, Stöckl A. Towards an understanding of resilience and its relevance to medical training. Med Educ 2012 Apr;46(4):349-356. [doi: 10.1111/j.1365-2923.2011.04188.x] [Medline: 22429170]
63. Martin-Delgado J, Viteri E, Mula A, Serpa P, Pacheco G, Prada D, et al. Availability of personal protective equipment and diagnostic and treatment facilities for healthcare workers involved in COVID-19 care: a cross-sectional study in Brazil, Colombia, and Ecuador. PLoS One 2020 Nov 11;15(11):e0242185 [FREE Full text] [doi: 10.1371/journal.pone.0242185] [Medline: 33175877]
64. Martin-Delgado J, Poblete R, Serpa P, Mula A, Carrillo I, Fernández C, et al. Contributing factors for acute stress in healthcare workers caring for COVID-19 patients in Argentina, Chile, Colombia, and Ecuador. Sci Rep 2022 May 19;12(1):8496 [FREE Full text] [doi: 10.1038/s41598-022-12626-2] [Medline: 35589975]
65. Turan N, Canbulat Ş. The effectiveness of the training program on accepting and expressing emotions on the psychological resilience and depression levels of nurses: a two-year follow-up study. Arch Psychiatr Nurs 2023 Jun;44:1-7. [doi: 10.1016/j.apnu.2023.03.002] [Medline: 37197852]
66. Burton NW, Pakenham KI, Brown WJ. Feasibility and effectiveness of psychosocial resilience training: a pilot study of the READY program. Psychol Health Med 2010 May;15(3):266-277. [doi: 10.1080/13548501003758710] [Medline: 20480432]
67. Rogers D. Which educational interventions improve healthcare professionals' resilience? Med Teach 2016 Dec;38(12):1236-1241. [doi: 10.1080/0142159X.2016.1210111] [Medline: 27573430]
68. Henshall C, Ostinelli E, Harvey J, Davey Z, Aghanenu B, Cipriani A, et al. Examining the effectiveness of web-based interventions to enhance resilience in health care professionals: systematic review. JMIR Med Educ 2022 Sep 06;8(3):e34230 [FREE Full text] [doi: 10.2196/34230] [Medline: 36066962]
69. Mahaffey BL, Mackin DM, Rosen J, Schwartz RM, Taioli E, Gonzalez A. The disaster worker resiliency training program: a randomized clinical trial. Int Arch Occup Environ Health 2021 Jan;94(1):9-21 [FREE Full text] [doi: 10.1007/s00420-020-01552-3] [Medline: 32448931]

XSL•FO
RenderX

70. López-Pineda A, Carrillo I, Mula A, Guerra-Paiva S, Strametz R, Tella S, et al. Strategies for the psychological support of the healthcare workforce during the COVID-19 pandemic: the ERNST study. Int J Environ Res Public Health 2022 May 02;19(9):5529 [FREE Full text] [doi: 10.3390/ijerph19095529] [Medline: 35564924]
71. Mira JJ, Vicente MA, Lopez-Pineda A, Carrillo I, Guilabert M, Fernández C, et al. Preventing and addressing the stress reactions of health care workers caring for patients with COVID-19: development of a digital platform (Be + against COVID). JMIR Mhealth Uhealth 2020 Oct 05;8(10):e21692 [FREE Full text] [doi: 10.2196/21692] [Medline: 32936769]
72. Pacutova V, Geckova AM, de Winter AF, Reijneveld SA. Opportunities to strengthen resilience of health care workers regarding patient safety. BMC Health Serv Res 2023 Oct 19;23(1):1127 [FREE Full text] [doi: 10.1186/s12913-023-10054-0] [Medline: 37858175]
73. Smith AF, Plunkett E. People, systems and safety: resilience and excellence in healthcare practice. Anaesthesia 2019 Apr;74(4):508-517 [FREE Full text] [doi: 10.1111/anae.14519] [Medline: 30585298]
74. Aase K, Guise V, Billett S, Sollid SJ, Njå O, Røise O, et al. Resilience in Healthcare (RiH): a longitudinal research programme protocol. BMJ Open 2020 Oct 26;10(10):e038779 [FREE Full text] [doi: 10.1136/bmjopen-2020-038779] [Medline: 33109657]
75. Brown L, Cohen B, Costello R, Brazhnik O, Galis ZS. Next steps: operationalizing resilience research. Stress Health 2023 Sep;39(S1):62-66. [doi: 10.1002/smi.3256] [Medline: 37184110]
76. Ejaz H, Sultan B, Pienaar AJ, Froelicher ES. Effectiveness of a resilience-focused educational program for promoting resilience in nursing students: a systematic review and meta-analysis. Nurse Educ Pract 2024 Jul;78:104014. [doi: 10.1016/j.nepr.2024.104014] [Medline: 38879909]
77. Palacio GC, Krikorian A, Gómez-Romero MJ, Limonero JT. Resilience in caregivers: a systematic review. Am J Hosp Palliat Care 2020 Aug;37(8):648-658. [doi: 10.1177/1049909119893977] [Medline: 31830813]

Edited by A Kushniruk; submitted 07.10.24; peer-reviewed by D Bolivar, M Munikumar; comments to author 23.12.24; revised version received 30.12.24; accepted 31.12.24; published 05.02.25.

Please cite as:
Gil-Hernández E, Carrillo I, Martin-Delgado J, García-Torres D, Mira JJ
Development of a Web-Based Intervention for Middle Managers to Enhance Resilience at the Individual, Team, and Organizational Levels in Health Care Systems: Multiphase Study
JMIR Hum Factors 2025;12:e67263
URL: https://humanfactors.jmir.org/2025/1/e67263
doi:10.2196/67263
PMID:

©Eva Gil-Hernández, Irene Carrillo, Jimmy Martin-Delgado, Daniel García-Torres, José Joaquín Mira. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 05.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Web-Based Lifestyle-Related Course for People Living With Multiple Sclerosis: Quantitative Evaluation of Course Completion, Satisfaction, and Lifestyle Changes Among Participants Enrolled in a Randomized Controlled Trial

Maggie Yu[1], MPH, MPsy; Sandra Neate[1], MBBS; Steve Simpson-Yap[1,2,3], PhD; Rebekah Davenport[4], MCP; William Bevens[5], PhD; George Jelinek[1], MBBS, MD; Jeanette Reece[1], PhD

[1]Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Australia

[2]Florey Institute of Neuroscience and Mental Health, The University of Melbourne, Melbourne, Australia

[3]MS Research Flagship, Menzies Institute for Medical Research, University of Tasmania, Hobart, Australia

[4]Melbourne School of Psychological Sciences, The University of Melbourne, Melbourne, Australia

[5]Department of Psychiatry, University of California, San Diego, IN STEP Children's Mental Health Research Center, Child & Adolescent Services Research Center (CASRC), La Jolla, CA, United States

Corresponding Author:
William Bevens, PhD
Department of Psychiatry, University of California, San Diego
IN STEP Children's Mental Health Research Center
Child & Adolescent Services Research Center (CASRC)
9500 Gilman Drive
La Jolla, CA, 92093-0812
United States
Phone: 1 949 824 3358
Email: wbevens@health.ucsd.edu

Abstract

Background: Web-based health courses providing lifestyle-related information can potentially increase knowledge, facilitate behavior change, and improve health outcomes for people living with multiple sclerosis (MS). Despite the low engagement with web-based programs by this population, few studies have evaluated factors influencing engagement. This study evaluated engagement with our 6-week lifestyle-related course (Multiple Sclerosis Online Course; MSOC) by participants enrolled in a large, international randomized controlled trial, as well as preliminary outcomes.

Objective: This study aimed to quantitatively assess engagement with the MSOC (the intervention course [IC] and standard-care course [SCC]), motivators of and barriers to participants' course completion, course satisfaction, engagement with the community forum, and intentions to implement lifestyle changes.

Methods: We collected data via a baseline survey before course commencement and an evaluation survey 1 month after the 6-week course. Course completers were queried on motivators of completion, course satisfaction, previous knowledge, forum participation, and intentions to adopt lifestyle changes. Noncompleters were queried on barriers to course completion. Differences between the 2 study arms were examined using chi-square and 2-tailed *t* tests. Multivariable linear regression models assessed factors (sociodemographic and course and health related) associated with participants' intentions to adopt lifestyle changes adjusting for baseline lifestyle factors. Moderation analyses were conducted to test group differences.

Results: Of the 857 participants, 442 (51.6%) completed the MSOC (IC: n=218, 49.3%; SCC: n=224, 50.7%), and 291 (34%) completed the evaluation survey (n=254, 87.3% course completers; n=37, 12.7% noncompleters). Key motivators of course completion included an interest in participating in MS research, optimizing health, course flexibility, and relevant and useful course content. Barriers to course completion included time constraints and technical issues. Most course completers rated the MSOC as "excellent/very good" (IC: 92/126, 73%; SCC: 78/128, 60.9%; *P*=.17). Engagement with the facilitator-led community forum was higher in the IC than in the SCC (56/126, 44.4% vs 32/128, 25%; *P*=.003). More IC completers versus SCC completers expressed their intention to adopt dietary changes (89/125, 71.2% vs 74/127, 58.3%; *P*=.04), increase their sun exposure (82/124, 66.1% vs 62/124, 50%; *P*=.01), supplement with omega-3 (84/125, 67.2% vs 60/126, 47.6%; *P*=.004), and practice meditation

XSL•FO
RenderX

(85/124, 68.5% vs 66/126, 52.4%; *P*=.009). Forum engagement, course satisfaction, new course content, and an interest in receiving additional course content were associated with intentions to adopt lifestyle changes across both study arms.

Conclusions: The web-based lifestyle IC provided new and satisfactory content and facilitated intentions to adopt lifestyle changes. Positive associations between engagement with the community forum and intentions to implement lifestyle changes and identifying barriers to completion such as time constraints provide important insights to inform the design of future digital health interventions for people living with MS and possibly other chronic conditions.

Trial Registration: Australian New Zealand Clinical Trials Registry (ANZCTR) ACTRN12621001605886; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=382778&isReview=true

(JMIR Hum Factors 2025;12:e59363) doi:10.2196/59363

KEYWORDS

multiple sclerosis; randomized controlled trial; lifestyle; multimodal; digital intervention; Multiple Sclerosis Online Course; course engagement

Introduction

Background

Multiple sclerosis (MS) is a chronic inflammatory disease of the central nervous system that affects almost 3 million people worldwide [1]. It is the most common cause of nontraumatic disability in young and middle-aged adults [2] and manifests with symptoms such as fatigue, depression, and anxiety, as well as reduced quality of life (QoL) [3,4].

Although treatments for MS such as disease-modifying therapies (DMTs) can ameliorate disease progression and reduce the risk of relapse (transient exacerbations of new or recurring neurological dysfunction) [5], people living with MS have expressed a strong interest in self-management strategies [6]. Self-management strategies, such as modification of lifestyle-related risk factors (eg, diet, exercise, and stress management), have been shown to be beneficial in managing MS-related symptoms and improving individuals' sense of control over their disease as an adjunct to pharmacological treatments [7-10]. More recently, web-based interventions have demonstrated effectiveness in promoting self-management strategies to improve MS-related knowledge, self-efficacy, and symptom management [11,12]. Web-based health interventions, particularly lifestyle-related ones, could improve accessibility issues and reduce barriers of geography and mobility, thereby increasing knowledge translation to people living with MS [12-15].

Despite the potential advantages of web-based health interventions for people living with MS in terms of availability and cost-effectiveness, these interventions have encountered issues related to high attrition [16,17] and low engagement rates [18]. Previous studies evaluating participation in web-based health-related courses for people living with MS have also focused primarily on course satisfaction [18], knowledge gain [18], and self-management of fatigue [11,12]. Few studies have comprehensively examined engagement (particularly course commencement and completion) with web-based health interventions to facilitate self-management strategies for people living with MS, including whether knowledge gain through interventions translates into changes in attitudes and lifestyle.

In addition, the influence of social support and interaction on the completion of self-paced web-based interventions for people living with MS remains unclear. While chat forums, online message groups, and web-based coaches have been found to increase course satisfaction [19], there is a noticeable gap in the literature regarding social engagement and the learning outcomes of web-based lifestyle interventions. Given the rapid shift toward web-based health care platforms and their potential benefits for people living with MS, studies exploring engagement with web-based interventions by populations with chronic illnesses such as MS remain a high priority.

We developed the Multiple Sclerosis Online Course (MSOC), a web-based program delivering evidence-based lifestyle recommendations for people living with MS with the aim of improving health and well-being outcomes in this population. After confirming the feasibility, acceptability, and learnability of the MSOC [14], a large randomized controlled trial (RCT) was designed and is currently being conducted to investigate the effectiveness of a multimodal lifestyle intervention course (IC) in improving health outcomes and QoL in people living with MS compared to a standard-care course (SCC) [15].

Objectives

This study aimed to evaluate engagement with the web-based lifestyle program, the MSOC, by RCT participants (an international cohort of people living with MS) and the impact of the MSOC in facilitating intentions to adopt lifestyle changes *after* participants were provided with access to the MSOC (IC or SCC; Textbox 1). This study used data collected from an evaluation survey 1 month after the delivery of the MSOC and specifically aimed to determine motivators of course completion, levels of course satisfaction, engagement with a facilitator-led community forum, and barriers to engagement among course noncompleters (Multimedia Appendix 1). In addition, we examined intentions to adopt lifestyle changes across the IC and SCC.

XSL•FO
RenderX

Textbox 1. A comparison of aspects of engagement with the Multiple Sclerosis Online Course across the intervention course and standard-care course in course completers and noncompleters.

Aspects examined among course completers
Motivators of course completion
Course satisfaction
Previous knowledge of course content
Participation in the facilitator-run community forum
Intention to adopt recommended lifestyle modifications
Sociodemographic and course- and health-related factors associated with participants' intention to adopt recommended lifestyle modifications
Aspects examined among course noncompleters
Barriers to course completion

Notably, this study complements our previously published quantitative study from the same RCT, which specifically aimed to determine sociodemographic, clinical, and lifestyle-related factors associated with course commencement and completion using participant baseline data [20]. That is, baseline data were collected before participants had access to the MSOC, whereas this study used data collected 1 month after the 6-week course. In another 2 *qualitative* studies from the same RCT, we examined how people living with MS (N=38 course completers) seek information online [21] and participants' perceptions of their MS and their perceived sense of control over their disease [22].

Overall, this study provides unique insights into factors influencing engagement and lifestyle change intentions and examines the role of course design elements such as facilitator-led forums. The study findings can be used to improve the design of future digital health interventions for people living with MS and potentially other chronic conditions.

Methods

Overview

In this ongoing RCT, the IC, based largely on information from the Overcoming MS lifestyle program [23], is being tested for its effectiveness in improving health outcomes in a large international cohort of people living with MS compared to the SCC, which provides general lifestyle recommendations from reputable MS websites, as described fully in the protocol paper [24] (Multimedia Appendices 2 and 3).

Enrolled participants (people living with MS recruited online from the United Kingdom and Ireland, North America, and Australia and New Zealand) were allocated to 1 of 2 courses using simple randomization: the IC or the SCC. The web-based courses were freely delivered worldwide in English via a laptop or PC connected to the internet. Course content was presented in videos, animations, text, and images and structured into seven modules covering the following topics: (1) introduction, (2) eating well (diet recommendations), (3) sunlight and vitamin D, (4) exercise, (5) meditation and the mind-body connection, (6) medication and family prevention, and (7) conclusion (Table S1 in Multimedia Appendix 4). Briefly, the 7 learning modules were delivered biweekly over 4 weeks followed by a 2-week catch-up period (total course length was 6 weeks). Each course also included a facilitator-led community forum.

In the feasibility RCT, while the MSOC was found to be both acceptable and useful by people living with MS, 45% of participants did not complete the course, and there was no engagement with the community forum [25]. As a result, facilitators and researchers were added to the community forum in the MSOC effectiveness RCT [24] to promote dialogue by introducing each module topic, promoting discussion within the module by asking questions related to the topic, and answering participants' questions with the intent of increasing engagement and, subsequently, course completion.

In total, 5 rounds of recruitment took place via international MS websites and social media: August 1, 2022, to September 12, 2022; October 19, 2022, to November 21, 2022; November 21, 2022, to December 31, 2022; March 27, 2023, to May 7, 2023; and July 24, 2023, to September 3, 2023.

Ethical Considerations

All participants provided informed consent for their data to be used for research. This study was approved by the University of Melbourne Human Research Ethics Subcommittee (ID: 22140). The trial was registered with the Australian New Zealand Clinical Trials Registry (ACTRN12621001605886). All methods were carried out in accordance with relevant guidelines and regulations. Participants were provided with a detailed plain-language statement explaining the study objectives, procedures, potential risks, and benefits. This study was administered and conducted solely by the Neuroepidemiology Unit at the Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, the University of Melbourne. The single-site nature of this study ensured that all participant data were securely stored in 1 location, on the University of Melbourne server. Participants did not receive monetary compensation for their involvement.

MSOC Completion

Course completion was defined as completing modules 1 to 6 of the MSOC as module 7 comprised a closing session and did not provide lifestyle-related information. Course noncompletion was defined as completing module 1 but failing to complete module 6. At the end of each run of the 6-week IC and SCC, all participants (including those who did not complete the

XSL•FO
RenderX

course) were emailed a voluntary postcourse evaluation survey to complete. Invited participants who did not complete the voluntary evaluation survey were then sent a further email reminder 2 weeks later.

Participant Sociodemographic and Health Characteristics

Participants submitted a baseline survey before course commencement via the online course platform. The baseline survey collected a range of sociodemographic characteristics (eg, sex, age, highest attained educational level, employment status, partnership status, and country of residence) and health data, including MS type (categorized into nonprogressive [benign or relapsing-remitting MS] and progressive [primary progressive, secondary progressive, or progressive-relapsing MS]); MS duration (in years), derived from the year of diagnosis and the year of survey completion; BMI, calculated from weight and height squared and categorized into underweight (<18.5 kg/m^2), normal weight (18.5-24.9 kg/m^2), overweight (25.0-29.9 kg/m^2), and obese (≥30.0 kg/m^2) according to World Health Organization guidelines; treated comorbidities queried via the Self-Administered Comorbidity Questionnaire and then dichotomized (≤1 or >1); DMT use (no or yes); and self-reported ongoing symptoms of relapse in the preceding 30 days (no or yes). Participants were also queried about whether they were taking medications for their MS.

Patient-Reported Outcome Measures

Fatigue was measured using the 9-item Fatigue Severity Scale, with mean Fatigue Severity Scale scores of >5 indicating clinically significant fatigue [26]. Disability was assessed using the Patient-Determined Disease Steps [27,28], a self-reported assessment of ambulatory disability that has been validated against the Expanded Disability Status Scale and is classified into 3 categories: normal or mild (0-2), moderate (3-5), and severe (6-8). The 14-item Hospital Anxiety and Depression Scale was used to assess anxiety and depressive symptoms [29,30]. This tool comprises 7 questions each for anxiety and depression rated on a Likert scale from 0 to 3 and has been validated in people living with MS [31]. Health-related QoL was assessed using the psychometrically validated Multiple Sclerosis Quality of Life–54 (MSQOL-54) instrument [32]. The Physical Health Composite and Mental Health Composite were derived from relevant subscales per the MSQOL-54 guidelines. Although minimal clinically important differences have not been established for the MSQOL-54 composite scores, differences of ≥5 points have previously been determined as the minimum clinically meaningful change in a health-related QoL measure [33,34].

Motivators of and Barriers to Course Completion

Participants completing either the IC or SCC (completers) were queried on motivating factors that facilitated course completion, with multiple-choice options offered: “Convenience/flexible time and location,” “Opportunity to participate in MS research,” “Topics that were relevant and/or important to me,” “The course was interesting,” “Connection and interaction with other people with MS,” “Engagement with facilitators,” “To optimize my health,” and “Other (specify in text).” Participants could select more than one response to this question.

Participants not completing either the IC or SCC (noncompleters) were asked to identify obstacles or issues they encountered that prevented course completion. Multiple-choice options offered included “Technical issues (e.g., computer, internet connection, website),” “Problems with navigating the modules,” “Lack of time or inability to schedule conveniently,” “Irrelevant/unimportant course content/topics,” “Low quality/attractiveness of course pages/videos/materials,” “Health related to MS,” “Other health issues,” “Family or work issues,” “Participation in another course or intervention,” “Did not receive reminder emails,” and “Other (specify in text).” Participants could select more than one response to this question.

Participants’ Levels of Satisfaction With the MSOC

Completers were queried regarding their satisfaction with the corresponding courses. This included two questions related to (1) overall course experience, measured on a 6-point Likert scale from “excellent” to “very poor” (responses were categorized into “Excellent/very good,” “Good/average,” and “Poor/very poor”); and (2) how likely they would be to recommend the course to a peer or family member with MS, measured on a 5-point Likert scale ranging from “Extremely likely” to “Extremely unlikely” (responses were categorized into “Extremely/Somewhat likely,” “Neither,” and “Extremely/Somewhat unlikely”).

Familiarity With the Course Content and Additional Information Seeking

Completers were asked to rate their familiarity with the course content before commencing the course using a 5-point Likert scale with options from “All of the content was new to me” to “I was familiar with all of the content.” Responses were categorized into “Familiar with most/all the content,” “Some of the content was new,” and “All/most of the content was new.”

Completers were also asked to select from a list of topics related to each module that they would like to receive additional information on (eg, “What is MS,” “Sunlight and vitamin D,” and “Exercise”; *yes* or *no*).

Community Forum Engagement

We assessed engagement with the facilitator-run community forum by querying completers on whether they participated in the forum during the course (yes or no) and, if so, whether they found the community forum beneficial (yes or no), with text options provided to expand on these responses.

Intention to Adopt Lifestyle Change Recommendations

To assess whether course engagement could facilitate lifestyle modification, completers were asked to rate their likelihood of adopting the lifestyle recommendations of the IC or SCC. The following question was asked for each topic or module: “As a result of taking the course, how likely are you to change the following lifestyle behaviour? (diet, smoking cessation, omega-3 supplementation, sun exposure or vitamin D supplementation, physical activity, meditation, and other stress reduction activities).” Responses were rated on a 5-point Likert scale: “extremely unlikely,” “moderately unlikely,” “neither,”

XSL•FO
RenderX

"moderately likely," and "extremely likely," resulting in a total score ranging from 0 to 5. Participants selected "not applicable" if they already adhered to the lifestyle recommendation (eg, nonsmoking).

Statistical Analysis

Participant Characteristics and Group Engagement

Demographic characteristics of the study participants at baseline were examined as sample sizes and percentages (categorical variables) and means and SDs (continuous variables). Motivators of and barriers to course completion, course satisfaction levels, engagement with the community forum, interest in seeking additional information on course topics, and intention to adopt lifestyle recommendations of the IC and SCC were examined as percentages. Differences between the IC and SCC groups were assessed using the 2-tailed *t* test for continuous variables and chi-square tests for categorical variables.

Assessing Factors of Intention to Implement Lifestyle Changes

We examined whether participant- or course-related factors were associated with participants' intention to adopt individual lifestyle recommendations among course completers. Multivariate linear regression models were conducted to estimate adjusted regression coefficients and 95% CIs adjusting for baseline engagement with lifestyle recommendations and participation in another lifestyle or intervention program. To assess differences between the IC and SCC, we conducted moderation analyses by adding an interaction term (group × lifestyle behavior) to the regression models. All statistical analyses were conducted using Stata (version 17.0; StataCorp). Statistical significance was set at $P<.05$.

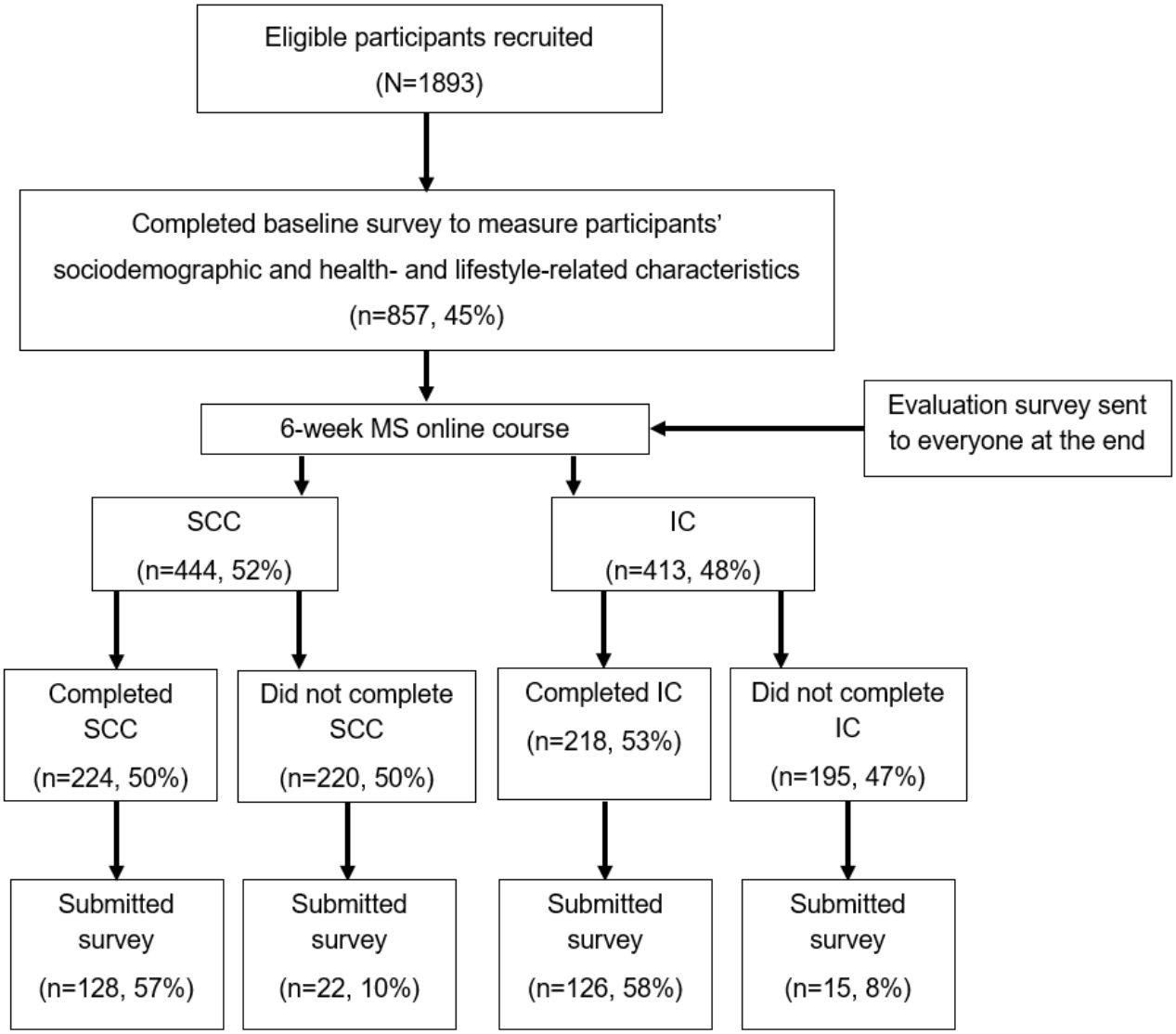
Results

Participant Characteristics

In total, 857 participants (aged ≥18 years with a confirmed clinical diagnosis of MS) enrolled in the MSOC effectiveness RCT and completed the baseline survey (IC: n=413, 48.2%; SCC: n=444, 51.8%; Figure 1). The analysis cohort comprised 291 participants (n=254, 87.3% course completers and n=37, 12.7% course noncompleters) who completed the baseline survey and postcourse evaluation survey.

XSL•FO
RenderX

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flow diagram. IC: intervention course; MS: multiple sclerosis; SCC: standard-care course.



There were similar numbers of participants (141/291, 48.5% IC participants and 150/291, 51.5% SCC participants) across both study arms (Table 1). Participants were 87.6% (255/291) women, and the mean age was approximately 50 years (SCC: mean 50.6, SD 10.7 years; IC: mean 49.8, SD 11.5 years). On the basis of visual inspection alone, participants' sociodemographic characteristics appeared similar across both study arms except for BMI—the percentage of participants with obesity was lower in the IC (28/141, 19.9%) than in the SCC (48/150, 32%). A large proportion of participants held a university degree (187/291, 64.3%) and were partnered (210/291, 72.2%). Approximately half (141/291, 48.5%) of the participants were employed. The most frequent regions of residence were North America and Australia and New Zealand (Table S2 in Multimedia Appendix 4).

XSL•FO
RenderX

Table 1. Descriptive statistics of the study sample (N=291)[a].

	SCC[b] (n=150)	IC[c] (n=141)
Sociodemographic characteristics		
Gender (men), n (%)	15 (10.1)[d]	21 (14.9)
Age (y), mean (SD)	50.6 (10.7)	49.8 (11.5)
Educational level (lower than university), n (%)	58 (38.9)[d]	45 (31.9)
University education, n (%)	91 (61.1)[d]	96 (68.1)
Employment status (working), n (%)	67 (47.9)[e]	74 (56.9)[f]
Partnered (married or de facto partnership), n (%)	109 (73.7)[g]	101 (71.6)
Country of residence, n (%)		
Australia or New Zealand	44 (29.5)[d]	41 (29.1)
North America (United States or Canada)	55 (36.9)[d]	49 (34.8)
United Kingdom	20 (13.4)[d]	20 (14.2)
Other	30 (20.1)[d]	31 (22.0)
Health characteristics		
MS[h] type (nonprogressive), n (%)	105 (74.5)[i]	100 (76.9)[f]
MS duration (y), mean (SD)	9.6 (8.43)	10.0 (8.87)
Comorbidities (≥1), n (%)	83 (58.9)[i]	77 (57.9)[j]
BMI, n (%)		
Underweight or normal weight	65 (43.6)[d]	82 (58.6)[e]
Overweight	36 (24.2)[d]	30 (21.4)[e]
Obese	48 (32)[d]	28 (19.9)[k]
Taking MS-related medication, n (%)	106 (71.6)[g]	88 (63.3)[l]
Patient-reported outcome measures , n (%)		
Disability (PDDS[m])		
Normal or mild	72 (48.3)[d]	72 (51.1)
Moderate	63 (42.3)[d]	52 (36.9)
Severe	14 (9.4)[d]	17 (12.1)
Clinically significant fatigue (mean FSS[n]>5)	80 (54.1)[g]	79 (56.8)[l]
Depression (HADS-D[o])		
Normal	90 (62.1)[p]	93 (67.4)[q]
Borderline	32 (22.1)[p]	24 (17.4)[q]
Abnormal	23 (15.9)[p]	21 (15.2)[q]
Anxiety (HADS-A[r])		
Normal	69 (47.3)[s]	68 (49.3)[q]
Borderline	30 (20.6)[s]	40 (29)[q]
Abnormal	47 (32.2)[s]	30 (21.7)[q]

[a]Percentages are based on available data; totals may be smaller due to missing values.

[b]SCC: standard-care course.

XSL•FO
RenderX

[c]IC: intervention course.
[d]n=149.
[e]n=140.
[f]n=130.
[g]n=148.
[h]MS: multiple sclerosis.
[i]n=141.
[j]n=133.
[k]*P*<.05.
[l]n=139.
[m]PDDS: Patient-Determined Disease Steps.
[n]FSS: Fatigue Severity Scale.
[o]HADS-D: Hospital Anxiety and Depression Scale–Depression subscale.
[p]n=145.
[q]n=138.
[r]HADS-A: Hospital Anxiety and Depression Scale–Anxiety subscale.
[s]n=146.

Most participants had nonprogressive MS types, and the mean time since diagnosis was approximately 10 years (SCC: mean 9.6, SD 8.43 years; IC: mean 10.0, SD 8.87 years). Approximately half of the participants reported no or mild disability (144/291, 49.5%) or clinically significant fatigue (159/291, 54.6%). On average, participants reported at least one comorbidity that was either physical or mental health–related. Approximately two-thirds of the participants reported symptoms of depression (183/291, 62.9%) and were taking medication to manage their MS (194/291, 66.7%), and 47.1% (137/291) had symptoms of anxiety within the normal range.

Motivation for and Barriers to Completion

Among the 857 participants who completed the baseline survey, course completion rates were 50.5% (224/444) for the SCC and 52.8% (218/413) for the IC (Figure 1). Among the course completers, 57.1% (128/224) and 57.8% (126/218) of SCC and IC participants, respectively, completed the postcourse evaluation survey (Table 2). The main motivation for course completion was similar between both groups regarding the "opportunity to participate in MS research" (SCC: 104/128, 81.3%; IC: 96/126, 76.2%). Other commonly reported motivators of course completion included "topics were relevant" (SCC: 88/128, 68.8%; IC: 89/126, 70.6%), "convenience/flexible time" (SCC: 88/128, 68.8%; IC: 83/126, 65.9%), "to optimize my health" (SCC: 77/128, 60.2%; IC: 86/126, 68.3%), and "the course was interesting" (SCC: 67/128, 52.3%; IC: 68/126, 54%). Less common motivators of completion were "connection and interaction with others" (SCC: 7/128, 5.5%; IC: 20/126, 15.9%) and "engagement with facilitators" (SCC: 8/128, 6.3%; IC: 14/126, 11.1%).

Of the 417 participants who did not complete the course, 37 (8.9%) completed the evaluation survey (22/222, 9.9% SCC noncompleters and 15/195, 7.7% IC noncompleters; Figure 1). Among noncompleters, the most commonly reported barriers to course completion were "lack of time or inability to schedule conveniently" (SCC: 9/22, 41%; IC: 7/15, 47%) and "technological issues" (SCC: 7/22, 32%; IC: 2/15, 13%). Other barriers to completion were "health issues" and "family or work-related issues," and low quality or attractiveness (ie, the course content lacked visual appeal or was not interesting) was reported by 14% (3/22) of SCC participants but no IC participants.

XSL•FO
RenderX

Table 2. Most common reasons for course completion and noncompletion[a].

	SCC[b], n (%)	IC[c], n (%)
Motivators among course completers		
Opportunity to participate in MS[d] research	104 (81.3)[e]	96 (76.2)[f]
Topics were relevant	88 (68.8)[e]	89 (70.6)[f]
Convenient or flexible time	88 (68.8)[e]	83 (65.9)[f]
To optimize their health	77 (60.2)[e]	86 (68.3)[f]
The course was interesting	67 (52.3)[e]	68 (54)[f]
Connection and interaction with others	7 (5.5)[e]	20 (15.9)[f,g]
Engagement with facilitators	8 (6.3)[e]	14 (11.1)[f]
Other	10 (7.8)[e]	10 (7.9)[f]
Barriers to starting or completing the course among noncompleters		
Lack of time or inability to schedule time for the course conveniently	9 (40.9)[h]	7 (46.7)[i]
Technological issues	7 (31.8)[h]	2 (13.3)[i]
Health issues	3 (13.6)[h]	4 (26.7)[i]
Family or work-related issues	1 (4.5)[h]	4 (26.7)[i]
Low quality or attractiveness of the course	3 (13.6)[h]	0 (0)[i]
Irrelevant or unimportant content or topics	2 (9.1)[h]	1 (6.7)[i]
Problems navigating the modules	1 (4.5)[h]	1 (6.7)[i]
Participation in another course or intervention	1 (4.5)[h]	1 (6.7)[i]
Did not receive reminder emails	0 (0)[h]	1 (6.7)[i]
Other	2 (9.1)[h]	0 (0)[i]

[a]Differences between the standard-care course and the intervention course were assessed using the chi-square test.

[b]SCC: standard-care course.

[c]IC: intervention course.

[d]MS: multiple sclerosis.

[e]n=128.

[f]n=126.

[g]$P<.05$.

[h]n=22.

[i]n=15.

Participant Satisfaction

Most course completers were satisfied with the course across both study arms (78/128, 60.9% of SCC completers vs 92/126, 73% of IC completers rated the course as “excellent/very good,” and 48/128, 37.5% of SCC completers vs 32/126, 25.4% of IC completers rated it as “good/average”; Table 3). Less than 2% of completers across both study arms (SCC: 2/128, 1.6%; IC: 2/126, 1.6%) rated the course as “poor/very poor.” Most course completers (SCC: 106/128, 82.8%; IC: 113/126, 89.7%) were “extremely likely” or “somewhat likely” to recommend the course to friends or family members with MS, whereas approximately 4% (SCC: 5/128, 3.9%; IC: 5/126, 4%) were “extremely unlikely” or “somewhat unlikely” to recommend the course.

XSL•FO
RenderX

Table 3. Course completers' evaluation of the intervention course (IC) and standard-care course (SCC) in terms of course satisfaction, course recommendation, and familiarity with the contenta.

Survey query	SCC (n=128), n (%)	IC (n=126), n (%)
Overall satisfaction		
Excellent or very good	78 (60.9)	92 (73)
Good or average	48 (37.5)	32 (25.4)
Poor or very poor	2 (1.6)	2 (1.6)
Recommendation to friend or family member with MS[b]		
Extremely or somewhat likely	106 (82.8)	113 (89.7)
Neither	17 (13.3)	8 (6.3)
Extremely or somewhat unlikely	5 (3.9)	5 (4)
Familiarity with the content before MSOC[c]		
Familiar with most of or all the content	102 (79.7)	87 (69)[d]
Some of the content was new	22 (17.2)	27 (21.4)[d]
All or most of the content was new	4 (3.1)	12 (9.5)[d]

[a]Differences between the SCC and IC were assessed using the chi-square test.

[b]MS: multiple sclerosis.

[c]MSOC: Multiple Sclerosis Online Course.

[d]$P<.05$.

Familiarity With Course Content and Information Seeking

More SCC completers (102/128, 79.7%) compared with IC completers (87/126, 69%) were familiar with some or all the course content before taking the course, and only 3.1% (4/128) of SCC completers and 9.5% (12/126) of IC completers found that all or most of the course content was new (Table 3). When course completers were queried as to whether they would like to receive additional information on each course topic, the most common topics selected by ≥24% of SCC and IC completers were "diet," "physical activity (exercise)," "stress management," "medication and family prevention," and "review and consolidation" (Table 4). There was less interest in receiving additional information on "what is MS" (SCC: 12/128, 9.4%; IC: 16/125, 12.8%) and sun exposure (SCC: 24/128, 18.8%; IC: 17/125, 13.6%). Key topics of interest that SCC completers expressed greater interest in receiving additional information on than IC completers included physical activity (SCC: 49/128, 38.3%; IC: 30/125, 24%; $P=.01$) and "review and consolidation," which was part of the conclusion module (SCC: 63/128, 49.2%; IC: 32/125, 25.6%; $P<.001$).

Table 4. Topics that course completers would have liked additional information on[a].

Topic	SCC[b] (n=128), n (%)	IC[c] (n=125), n (%)
What is MS[d]	12 (9.4)	16 (12.8)
Diet	56 (43.8)	48 (38.4)
Sunlight and vitamin D	24 (18.8)	17 (13.6)
Physical activity	49 (38.3)	30 (24)[e]
Stress management	62 (48.4)	46 (36.8)
Medication and family prevention	41 (32)	32 (25.6)
Review and consolidation	63 (49.2)	32 (25.6)[f]

[a]Differences between the intervention course and the standard-care course groups were assessed using chi-square tests.

[b]SCC: standard-care course.

[c]IC: intervention course.

[d]MS: multiple sclerosis.

[e]$P<.01$.

[f]$P<.001$.

XSL•FO
RenderX

Community Forum Engagement

A lower proportion of participants in the SCC (32/128, 25%) than in the IC (55/126, 43.7%) engaged with the community forum within the course. Furthermore, most completers who engaged with the forum found it useful (SCC: 23/32, 72%; IC: 44/55, 80%), with reasons including "appreciated the opportunity to gain insights from others' experiences," "highlighted the benefits of receiving answers to their questions," "provided valuable resources from facilitators," and "appreciated the ability to share their own perspective and experiences." In contrast, the 2 main reasons why completers did not find the forum useful included "the limited ability to directly reply to and engage with other course members" and "the low number of participants in the forums."

Participants' Intention to Make Lifestyle Modifications

A large proportion of IC and SCC completers intended to adopt lifestyle recommendations, particularly diet, sun exposure, physical activity, and stress management (Table 5). However, a smaller proportion of SCC versus IC completers were "extremely likely" or "moderately likely" to make changes in their diet based on the corresponding course content (74/128, 57.8% vs 89/125, 71.2%; *P*=.04). SCC completers were also less likely to increase their sun exposure (62/124, 50% vs 82/124, 66.1%; *P*=.01), supplement with omega-3 (60/126, 47.6% vs 84/125, 67.2%; *P*=.004), and practice meditation (66/126, 52.4% vs 85/124, 68.5%; *P*=.009). Of note, the SCC did not contain specific recommendations regarding sun exposure, omega-3 supplementation, or meditation practice (Table S1 in Multimedia Appendix 4).

Table 5. Intentions of participants in the intervention course (IC) and the standard-care course (SCC) to implement lifestyle changes.

Lifestyle change	SCC (n=128)[a], n (%)			IC (n=126)[b], n (%)			*P* value
	Extremely likely or moderately likely	Neither, extremely unlikely, or moderately unlikely	Not applicable	Extremely likely or moderately likely	Neither, extremely unlikely, or moderately unlikely	Not applicable	
Diet	74 (58.3)[c]	45 (35.4)[c]	8 (6.3)[c]	89 (71.2)[d]	26 (20.8)[d]	10 (8.0)[d]	.04
Vitamin D supplementation	49 (38.6)[c]	31 (24.4)[c]	47 (37.0)[c]	64 (52.5)[e]	20 (16.4)[e]	38 (31.2)[e]	.07
Sun exposure	62 (50.0)[f]	45 (36.3)[f]	17 (13.7)[f]	82 (66.1)[f]	25 (20.2)[f]	17 (13.7)[f]	.01
Omega-3 supplementation	60 (47.6)[g]	43 (34.1)[g]	23 (18.3)[g]	84 (67.2)[d]	22 (17.6)[d]	19 (15.2)[d]	.004
Physical activity	84 (67.2)[d]	25 (20.0)[d]	16 (12.8)[d]	85 (69.1)[h]	20 (16.3)[h]	18 (14.6)[h]	.72
Meditation	66 (52.4)[g]	46 (36.5)[g]	14 (11.1)[g]	85 (68.6) [f]	24 (19.4) [f]	15 (12.1) [f]	.009
Other stress reduction activities [i]	84 (67.2)[d]	30 (24.0)[d]	11 (8.8)[d]	92 (75.4)[e]	23 (18.9) [e]	7 (5.7)[e]	.34
Nonsmoking	12 (9.5)[g]	14 (11.1)[g]	100 (79.4) [g]	14 (11.5)[e]	8 (6.6)[e]	100 (82.0)[e]	.42

[a]Percentages are based on available data only; subgroup totals may be smaller than 128 due to missing data.

[b]Percentages are based on available data only; subgroup totals may be smaller than 126 due to missing data.

[c]n=127.

[d]n=125.

[e]n=122.

[f]n=124.

[g]n=126.

[h]n=123.

[i]Stress reduction activities other than meditation.

Factors Associated With Participants' Intentions for Lifestyle Modifications

Multivariate analysis found associations between several course-related factors and completers' intentions to adopt the lifestyle recommendations of each course (Table 6).

In particular, course satisfaction was associated with an intention to adopt vitamin D supplementation among SCC completers, whereas among IC completers, course satisfaction was associated with an intention to adopt dietary recommendations, vitamin D supplementation or sun exposure, and meditation practice. For example, IC completers who rated the course as "very good" or "excellent" had a 1.69-point (95% CI 0.24-3.13) higher intention to adopt a healthy diet due to the course. Engagement with the community forum, new course content, and seeking additional information on certain topics were also associated with intentions to change certain factors, with no group differences between the SCC and IC.

XSL•FO
RenderX

Table 6. Course-related, sociodemographic, and health-related factors associated with intentions to change certain lifestyle behaviors in standard-care course (SCC) and intervention course (IC) completers[a].

Factor and course	aβ[b] (95% CI)					
	Healthy diet	Vitamin D supplementation	Sun exposure	Omega-3 supplementation	Physical activity	Meditation
Course related						
High satisfaction[c]						
SCC	0.02 (−2.10 to 2.14)	0.86 (0.21 to 1.50)[d]	−0.31 (−2.41 to 1.80)	*−0.05 (−2.48 to 2.38)*[e]	0.18 (−1.95 to 2.32)	−0.45 (−2.84 to 1.94)
IC	1.69 (0.24 to 3.13)[d]	0.99 (0.49 to 1.50)[f]	3.39 (1.41 to 5.36)[g]	*−0.21 (−1.64 to 1.21)*	1.81 (−0.01 to 3.64)	1.54 (0.01 to 3.07)[d]
Participated in forum						
SCC	0.98 (0.26 to 1.69)[g]	0.80 (−0.55 to 2.15)	0.56 (−0.49 to 1.61)	*0.67 (−0.47 to 1.80)*	1.20 (0.42 to 1.99)[g]	0.77 (−0.14 to 1.67)
IC	0.28 (−0.48 to 1.04)	−0.20 (−1.31 to 0.91)	1.02 (0.16 to 1.88)[d]	*0.11 (−0.76 to 0.98)*	0.67 (−0.26 to 1.60)	0.87 (0.14 to 1.60)[d]
Learning new content[h]						
SCC	0.15 (−0.35 to 0.64)	0.06 (−0.72 to 0.84)	0.26 (−0.15 to 0.66)	*0.09 (−0.49 to 0.68)*	0.63 (0.06 to 1.19)[d]	0.58 (0.20 to 0.97)[g]
IC	0.56 (0.16 to 0.97)[g]	0.23 (−0.26 to 0.73)	0.52 (0.06 to 0.98)[d]	*0.42 (−0.00 to 0.84)*	0.33 (−0.31 to 0.97)	0.46 (0.04 to 0.88)[d]
Seeking information						
SCC	0.55 (0.16 to 0.94)[g]	0.94 (0.24 to 1.65)[d]	0.22 (−0.19 to 0.64)	*0.50 (0.03 to 0.96)* [d]	0.45 (0.06 to 0.85)[d]	0.00 (−0.48 to 0.49)
IC	0.55 (0.17 to 0.93)[g]	−0.05 (−0.75 to 0.65)	0.44 (−0.01 to 0.89)	*0.22 (−0.24 to 0.69)*	0.52 (0.10 to 0.95)[d]	0.42 (−0.07 to 0.90)
Sociodemographic						
Age of >54 y						
SCC	−0.02 (−0.72 to 0.68)	0.43 (−0.77 to 1.63)	−0.30 (−0.99 to 0.40)	*−0.03 (−0.80 to 0.11)*	−0.85 (−1.51 to −0.16)[d]	−0.51 (−1.28 to 0.27)
IC	−0.52 (−1.26 to 0.21)	−0.38 (−1.25 to 0.49)	0.39 (−1.08 to 0.31)	*0.87 (0.10 to 1.64)* [d]	−0.38 (−1.08 to 0.31)	−0.05 (−0.83 to 0.72)
Working						
SCC	−0.20 (−0.62 to 0.22)	−0.25 (−0.91 to 0.41)	0.29 (−0.15 to 0.73)	−0.09 (−0.60 to 0.42)	0.02 (−0.40 to 0.45)	*−0.53 (−0.98to−0.07)*[d]
IC	0.20 (−0.22 to 0.62)	0.25 (−0.26 to 0.76)	0.13 (−0.31 to 0.57)	0.06 (−0.38 to 0.50)	0.39 (−0.03 to 0.81)	*0.45 (0.01 to 0.90)* [d]
University education						
SCC	*0.23 (−0.18 to 0.63)*	−0.45 (−1.10 to 0.19)	−0.01 (−0.43 to 0.41)	*−0.09 (−0.59to0.41)*	−0.16 (−0.57 to 0.26)	0.23 (−0.24 to 0.70)
IC	*−0.50 (−0.90to−0.07)*[d]	−0.35 (−0.87 to 0.17)	0.08 (−0.40 to 0.55)	*−0.58 (−1.02to−0.15)*[g]	−0.34 (−0.77 to 0.10)	−0.35 (−0.82 to 0.12)
Health related						
Long MS[i] duration						
SCC	0.65 (0.05 to 1.26)[d]	*0.97 (0.03 to 1.91)* [d]	−0.27 (−0.90 to 0.35)	−0.15 (−0.91 to 0.61)	−0.19 (−0.83 to 0.46)	−0.06 (−0.74 to 0.63)
IC	−0.06 (−0.60 to 0.49)	*0.12 (−0.54 to 0.79)*	0.13 (−0.45 to 0.72)	0.03 (−0.51 to 0.57)	−0.12 (−0.64 to 0.40)	−0.40 (−0.97 to 0.17)

XSL•FO
RenderX

Factor and course	aβ[b] (95% CI)					
	Healthy diet	Vitamin D supplementation	Sun exposure	Omega-3 supplementation	Physical activity	Meditation
Used DMTs[j]						
SCC	0.10 (−0.33 to 0.53)	0.04 (−0.64 to 0.72)	0.40 (−0.05 to 0.84)	0.26 (−0.27 to 0.78)	−0.18 (−0.63 to 0.27)	−0.44 (−0.93 to 0.04)
IC	0.44 (0.03 to 0.85)[d]	0.12 (−0.39 to 0.63)	0.03 (−0.41 to 0.46)	0.21 (−0.22 to 0.64)	0.26 (−0.14 to 0.67)	−0.19 (−0.64 to 0.26)
≥1 comorbidity						
SCC	0.03 (−0.38 to 0.43)	0.10 (−0.67 to 0.67)	−0.21 (−0.65 to 0.23)	−0.14 (−0.66 to 0.36)	0.10 (−0.42 to 0.44)	−0.05 (−0.53 to 0.44)
IC	0.31 (−0.09 to 0.71)	0.06 (−0.45 to 0.58)	0.01 (−0.43 to 0.44)	0.50 (0.10 to 0.89)[d]	0.13 (−0.27 to 0.53)	0.33 (−0.13 to 0.78)

[a]We assessed associations between sociodemographic and course- and health-related factors and participants' intention to making changes in the healthy lifestyle behavior (range from 0=extremely unlikely to 4=extremely likely). Multiple linear regression models were used to estimate adjusted regression coefficients and 95% CIs adjusting for baseline lifestyle behaviors and involvement in other lifestyle interventions or programs. Differences between the SCC and IC were assessed using interaction analyses.

[b]aβ: adjusted regression coefficient.

[c]High satisfaction: participants who rated their overall satisfaction with the course as "very good" or "excellent."

[d]*P*<.05 between factors and intentions for lifestyle change within the SCC or IC.

[e] Group difference (*P*<.05) from SCC.

[f]*P*<.001 between factors and intentions for lifestyle change within the SCC or IC.

[g]*P*<.01 between factors and intentions for lifestyle change within the SCC or IC.

[h]Learning new content: participants who rated the course content as "most," "all," or "some" being new.

[i]MS: multiple sclerosis.

[j]DMT: disease-modifying therapy.

Sociodemographic factors associated with intentions for lifestyle modifications included age (being aged >54 years was associated with intentions to increase physical activity and supplement with omega-3, with the latter being significantly lower in SCC vs IC completers), employment status (being employed was associated with a higher intention to practice meditation among SCC completers and a lower intention among IC completers), educational level (a university degree was associated with a greater intention for diet changes and omega-3 supplementation among SCC vs IC completers).

Regarding health-related factors, longer MS duration (>15 years since diagnosis) was associated with a greater intention for dietary modifications and vitamin D supplementation among SCC versus IC completers. The use of DMTs and the presence of comorbidities were associated with intentions to adopt certain lifestyle modifications, with no significant differences between study arms.

Poorer participant-reported health outcomes at baseline were linked to higher intentions to adopt various lifestyle modifications, particularly among SCC completers (Table 7). Severe anxiety, depression and disability, and clinically significant fatigue were associated with greater intentions to adopt dietary modifications, vitamin D supplementation, increased sun exposure, and physical activity among SCC versus IC completers. However, higher physical and mental QoL was associated with lower intentions to modify vitamin D intake, omega-3 supplementation, and physical activity among SCC completers.

XSL•FO
RenderX

Table 7. Participant-reported health outcomes associated with intentions to change certain lifestyle behaviors among standard-care course (SCC) and intervention course (IC) completers[a].

Outcome and course	aβ[b] (95% CI)					
	Healthy diet	Vitamin D supplementation	Sun exposure	Omega-3 supplementation	Physical activity	Meditation
Health outcomes						
Severe anxiety[c]						
SCC	0.26 (−0.19 to 0.72)	0.51 (−0.20 to 1.23)	0.45 (−0.12 to 1.02)	0.65 (0.09 to 1.20)[d]	0.50 (−0.03 to 1.02)	*0.59 (0.08 to 1.10)* [d,e]
IC	0.45 (−0.00 to 0.90)	−0.16 (−0.77 to 0.45)	0.19 (−0.29 to 0.67)	0.11 (−0.43 to 0.66)	0.08 (−0.37 to 0.52)	*0.06 (−0.50 to 0.62)*
Severe depression[c]						
SCC	0.79 (0.25 to 1.34)[f]	0.97 (0.10 to 1.85)[d]	0.10 (−0.43 to 0.63)	0.33 (−0.26 to 0.92)	0.51 (0.01 to 1.01)[d]	0.12 (−0.45 to 0.68)
IC	0.42 (−0.15 to 1.00)	0.21 (−0.49 to 0.90)	0.13 (−0.39 to 0.66)	−0.04 (−0.56 to 0.47)	0.19 (−0.31 to 0.68)	0.06 (−0.61 to 0.72)
Fatigue[g]						
SCC	0.05 (−0.35 to 0.46)	0.62 (0.03 to 1.24)[d]	0.38 (−0.04 to 0.79)	0.19 (−0.29 to 0.67)	0.20 (−0.21 to 0.61)	0.14 (−0.31 to 0.60)
IC	−0.01 (−0.41 to 0.40)	0.08 (−0.40 to 0.57)	0.02 (−0.40 to 0.43)	−0.07 (−0.48 to 0.34)	−0.18 (−0.58 to 0.22)	0.03 (−0.41 to 0.48)
Severe disability[h]						
SCC	*0.80 (0.10 to 1.50)[d]*	*1.51 (0.43 to 2.59)* [f]	0.48 (0.03 to 0.92)[d]	0.21 (−0.63 to 1.05)	*0.49 (0.06 to 0.91)* [d]	0.78 (−0.06 to 1.62)
IC	*−0.42 (−0.83to−0.00)*[d]	*−0.51 (−1.47to0.35)*[d]	−0.14 (−0.58 to 0.30)	−0.03 (−0.70 to 0.64)	*−0.18 (−0.60 to 0.23)*	−0.31 (−1.15 to 0.54)
Physical HRQoL[i]						
SCC	−0.00 (−0.02 to 0.01)	*−0.03 (−0.04 to −0.01)*[f]	0.01 (−0.00 to 0.02)	−0.01 (−0.02 to 0.00)	−0.01 (−0.02 to −0.00)[d]	−0.01 (−0.02 to 0.00)
IC	−0.01 (−0.01 to 0.01)	*0.01 (−0.01 to 0.02)*	−0.01 (−0.02 to 0.00)	0.00 (−0.01 to 0.01)	−0.00 (−0.01 to 0.01)	−0.00 (−0.02 to 0.01)
Mental HRQoL[i]						
SCC	−0.00 (−0.01 to 0.01)	*−0.02 (−0.04 to −0.01)*[f]	−0.01 (−0.02 to 0.00)	−0.01 (−0.02 to −0.00)[d]	−0.01 (−0.02 to −0.00)[d]	−0.01 (−0.02 to 0.00)
IC	−0.01 (−0.02 to 0.00)	*0.01 (−0.01 to 0.01)*	0.00 (−0.01 to 0.01)	−0.01 (−0.02 to 0.00)	−0.00 (−0.01 to 0.01)	0.00 (−0.01 to 0.01)

[a]We assessed associations between health outcomes and participants' intention of making changes to the healthy lifestyle behavior (range from 0=extremely unlikely to 4=extremely likely) in each course. Multiple linear regression models were used to estimate adjusted regression coefficients and 95% CIs adjusting for baseline lifestyle behaviors and involvement in other lifestyle interventions or programs. Differences between the SCC and IC were assessed using interaction analyses.

[b]aβ: adjusted regression coefficient.

[c]Hospital Anxiety and Depression Scale.

[d]*P*<.05 between health outcomes and intentions for lifestyle change within the SCC or IC.

[e] Group difference (*P*<.05) from SCC.

[f]*P*<.01 between health outcomes and intentions for lifestyle change within the SCC or IC.

[g]Clinically significant fatigue (Fatigue Severity Scale).

[h]Patient-Determined Disease Steps.

[i]HRQoL: health-related quality of life.

XSL•FO
RenderX

Discussion

Principal Findings

Despite the potential advantages of digital health interventions for people living with MS in terms of accessibility and cost-effectiveness, their engagement with web-based interventions is low [18]. This study comprehensively evaluated factors associated with engagement with a novel web-based lifestyle-related program, the MSOC, by people living with MS and examined short-term outcomes following course completion nested within the flagship RCT. Findings revealed that participants in both study groups (SCC and IC) found the course content interesting and useful and reported high levels of course satisfaction. Importantly, there was evidence that the course promoted intentions to implement lifestyle behavior changes, and key factors associated with these intentions were identified across both study arms. These findings provide new insights for designing and implementing web-based lifestyle interventions tailored to the MS community, addressing critical gaps in usability, engagement, and learnability of digital health solutions.

Course Completion

We observed completion rates of 23.3% (442/1893) for enrolled participants and 51.6% (442/857) among those who completed the baseline survey. These completion rates are noteworthy considering that participants were required to complete a 166-question baseline survey, and a previous scoping review reported 5% to 15% completion rates of health education massive open online courses (MOOCs) by the general population [35]. Given the documented desire of people living with MS to receive lifestyle-related information to improve their future health prospects [25], this may explain the higher rates of completion in this study. Similarly, this cohort enrolled in the RCT may represent a subgroup of highly motivated people living with MS.

Motivators of and Barriers to Completion

Primary factors contributing to course completion included the "opportunity to participate in MS research" and "to optimize my health." These findings align with those of the MSOC feasibility study [26], where many people living with MS reported that these factors were crucial to their engagement with the course. The intention of people living with MS to contribute to MS research is consistent with previous findings and represents one of the most common reasons why individuals engage in any research [27]. In addition, participants considered relevant course content and the flexibility to engage with the course in their own time as factors that influence course completion.

Primary barriers to completion were time constraints and technical issues. Other completion barriers included health-, family-, and work-related commitments. While MS-related studies examining noncompletion of web-based courses are limited, a recent mixed methods study of a 6-week MOOC called Understanding MS aimed at increasing understanding and awareness of MS also found that time constraints and MS-related symptoms were common completion barriers [36]. Other studies have also identified barriers to the uptake of web-based health interventions, such as older age, lower educational level, and MS symptoms [37].

Course Satisfaction, Familiarity, and Seeking of Additional Information

While course satisfaction was high across both study arms, satisfaction levels were higher in the IC arm than in the SCC arm (IC: 92/126, 73%; SCC: 78/128, 60.9%). Due to the quantitative nature of our evaluative study, we can only speculate why participants found the course satisfactory. Plausible reasons for high course satisfaction may be participants' strong interest in obtaining information on healthy lifestyle recommendations or self-management strategies to improve their future health prospects, as found in the qualitative studies of this trial, where participants expressed that they sought new knowledge to empower them to self-manage their MS [21,22].

The positive tone of the MSOC may have further contributed to the high satisfaction rates as this was found to be an important factor in the qualitative aspect of the MSOC feasibility RCT [38,39]. Similarly, "keep it positive" was 1 of the 3 core principles for developing the 6-week Understanding MS MOOC [40].

Participants in the SCC arm were more familiar with the course content and expressed greater interest in receiving additional lifestyle-related information than those in the IC. This is not surprising due to the SCC course content being sourced from publicly available MS society websites, whereas the content of the IC was novel for many IC participants. People living with chronic illnesses such as MS often exhibit a strong desire for new information to enhance their understanding and management of their condition [41]. Seeking additional information may reflect participants' motivation and engagement with digital health interventions aimed at disease self-management. Similarly, digital interventions for diabetes and rheumatoid arthritis have shown that reliable, professional guidance is particularly valued, especially by newly diagnosed individuals [42,43].

Interestingly, while the SCC provided information from readily available sources such as international MS websites, only a small proportion of participants (SCC: 2/128, 1.6%; IC: 2/126, 1.6%) were dissatisfied with the course, and a substantial proportion (SCC: 106/128, 82.8%; IC: 113/126, 89.7%) indicated that they would recommend the course to others. These findings are consistent with those of the MSOC feasibility study [39], where the MSOC was found to meet participants' expectations regardless of their familiarity with the course materials.

While new information that addresses knowledge gaps is critical for engaging participants in digital interventions, current findings suggest that even those familiar with the material can benefit when interventions provide fresh perspectives and dynamic delivery formats. Digital health interventions that offer updated, evidence-based content tailored to diverse participant needs—ranging from foundational knowledge for beginners to

XSL•FO
RenderX

advanced updates for experienced users—are likely to achieve higher satisfaction and engagement.

Community Forum Engagement

In the previous MSOC feasibility RCT, there was no engagement with the forum [39]. Although participants expressed positivity regarding the advantages of the forum in terms of connecting with other people with MS, they felt nervous and were reluctant to initiate discussions. This limitation was anticipated to potentially affect course completion as a pilot RCT of a web-based self-management program for fatigue for people living with MS found that lower interaction and support were associated with increased study dropout [12]. To address this issue in the current MSOC effectiveness RCT, a researcher facilitator was introduced into the community forum in each course. The facilitator, who was familiar with the evidence-based recommendations and course content, was responsible for initiating conversations, creating discussions, and answering participants' queries to encourage participant interaction and, ultimately, course completion. Subsequently, we found that 25% (32/128) and 43.7% (55/126) of participants in the SCC and IC study arms, respectively, engaged with the forum in this study.

Notably, the engagement with the facilitator-led community forum did not significantly improve course completion rates (IC: 218/413, 52.8%; SCC: 224/444, 50.5%) compared to the feasibility study (59% in the IC and 50% in the SCC). Furthermore, facilitators were cited as a motivating factor for course completion by only a small proportion of participants (SCC: 8/128, 6.3%; IC: 14/126, 11.1%), although limited forum engagement, with less than half of the participants actively taking part, likely constrained its impact on course completion. Key barriers to forum engagement identified by the participants included the inability to directly reply to one another (restricted to only replying in a group thread) and the small number of forum participants. As shown in previous digital health interventions, participants with neurological disorders particularly value emotional support from others with similar conditions [44]. Addressing these restrictions could create more opportunities for meaningful interaction and enhance the potential benefits of peer support in future iterations.

Interestingly, among those who did participate in the forum, most found it helpful (SCC: 23/32, 72%; IC: 44/55, 80%), particularly for accessing shared resources and gaining insights from others. These findings align with the findings of the 2 qualitative studies related to this RCT [21,22] and existing evidence [45] suggesting that participants are more likely to engage with digital health interventions that promote both social connectedness and information sharing.

Overall, while facilitator-led forums hold promise, the presence of a facilitator alone may not be sufficient to drive higher course completion rates. Introducing structured activities or prompts could further encourage engagement and foster dynamic conversations that extend beyond a small subset of participants.

Participants' Intentions to Undertake Lifestyle Modifications

The overarching aim of the MSOC was to improve QoL and health outcomes in people living with MS. Preliminary evaluation found that course completers across both study arms intended to initiate lifestyle modifications following the course. According to the theory of planned behavior [46], intentions serve as crucial predictors for behavior change, thereby providing support for the potential of the MSOC to encourage lifestyle modification in people living with MS. This will be assessed in detail by analyzing primary and secondary outcomes of the RCT at the 6-, 12-, 24-, and 30-month follow-ups.

Across both study arms, physical activity and stress reduction activities that did not include meditation practice were the most common lifestyle recommendations that participants intended to adopt. Notably, significantly more IC completers intended to increase their sun exposure, supplement with omega-3, and practice meditation than SCC completers. This finding was expected considering that no specific recommendations regarding these lifestyle modifications were provided in the SCC. In contrast, the SCC provided more general dietary information, emphasizing the importance of a balanced diet and providing information on the vast range of MS-related diets adopted by people living with MS.

A recent systematic review of web-based health education interventions found that most included studies focused on knowledge gain as the main outcome, whereas the impact of knowledge gain on other outcomes such as lifestyle modification was less clear [18]. However, the Understanding MS MOOC study did examine lifestyle modification in people living with MS 8 to 10 weeks after course completion and found that 43% to 52% of participants had made improvements in their diet quality, increased their physical activity, and initiated supplementation with vitamin D [47]. Of note, these modifications were observed despite only 1 of the 6 MOOC learning modules providing information on lifestyle-related risk factors. Collectively, the aforementioned study combined with our study findings demonstrates that web-based learning can potentially facilitate lifestyle modification in people living with MS. The findings are consistent with meta-analytic evidence of the effects of web-based lifestyle interventions in improving diet quality and increasing physical activity in survivors of cancer [48] and people with cardiovascular disease [49].

Factors Associated With Intentions to Undertake Lifestyle Modifications

Understanding factors that may influence individuals' experiences with web-based lifestyle interventions is important in optimizing the effectiveness of tailored programs. In this study, course-related factors were associated with participants' immediate learning outcomes. For instance, higher course satisfaction and an interest in seeking new information correlated with a greater intention to implement lifestyle modifications across both study arms, consistent with previous reports of links between knowledge gain and increased motivation [50].

Engagement with the forum was also associated with participants' intentions to change certain lifestyle behaviors,

XSL•FO
RenderX

consistent with previous study findings indicating that social connections made with peers during educational programs are important for engagement [51,52] and can mediate lifestyle changes [53]. In particular, a recent systematic review found that human support was a crucial component in the effectiveness of web-based behavior change interventions [19]. Furthermore, engagement with health care professionals in an MS-related face-to-face lifestyle program was found to increase participants' sense of control and agency in adopting lifestyle changes [52,54]. Overall, these results highlight the importance of facilitator-led community forums as integral components of digital health interventions. With tailored design and implementation, these features could be optimized to promote increased health-related knowledge and support lifestyle changes.

Of interest, a large proportion of SCC completers expressed their intention to change their lifestyle despite being familiar with the course content. Therefore, the SCC may have served to refresh or reinforce their existing knowledge. Engagement with information presented in a different setting or format (eg, animation and videos) may also have been a motivating factor for facilitating intentions for lifestyle changes. Alongside the observed high satisfaction levels, these results highlight the perceived value of the MSOC irrespective of participants' previous familiarity with the course materials, translating acquired or refreshed knowledge into motivational changes. This insight offers a key recommendation for future digital health design: developing modules that cater to both beginners and those with previous knowledge. For participants already familiar with the content, modules can serve as a refresher or provide new perspectives through web-based and visually engaging formats such as animations, videos, or quizzes.

Interestingly, other participant-related factors were also associated with intentions for lifestyle change. For instance, higher educational level was associated with lower intention for lifestyle change, consistent with studies indicating that participants with lower educational levels find course content more informative and applicable [17,55]. This highlights the importance of using clear, practical messaging and simplifying language and concepts to ensure accessibility for a broad audience. Disability was associated with an intention for diet change, vitamin D supplementation, and increased physical activity in the SCC but not in the IC. This may be attributed to the content of the SCC as it may not be possible for someone with a severe disability to adopt the restrictive diet or physical activity recommendations of the IC, especially if they rely on carers. As a result, it is important to consider course content and its applicability to people living with MS of all abilities.

In addition, SCC participants with severe mental health issues and lower QoL reported higher intentions to adopt lifestyle modifications. However, interestingly, our analysis of baseline characteristics and course completion in this RCT found that those with poor health status at baseline were less likely to complete the course [20]. While participants with poorer health may face physical or psychological challenges that reduce their likelihood of completing web-based courses, collectively, our results suggest that participants who complete web-based courses have a higher likelihood of implementing behavior change. Moreover, these findings suggest that future digital health interventions need to be designed to attempt to overcome course noncompletion, possibly by implementing tailored and motivational strategies such as progress tracking, personalized feedback, and interactive support tools to help overcome barriers to engagement to maximize the benefits of these programs.

Strengths and Limitations

This study sample represents a large and diverse international cohort of people living with MS, with characteristics similar to those of other international cohorts [7,37]. The findings expand on previously documented qualitative insights from this trial [21,22], providing a more comprehensive understanding of user experiences and engagement with web-based educational programs. The MSOC's utility as a web-based program demonstrates the potential of digital interventions to reach large, geographically diverse populations. This scalability is particularly beneficial for other chronic conditions for which health care access may be limited.

Various limitations should be considered. First, the study sample was restricted to RCT participants who completed the baseline and evaluation surveys. As research on web-based health interventions for chronic illnesses is likely to face challenges in recruiting and retaining participants [16,55], our study sample may be biased toward healthier individuals or those highly motivated to acquire lifestyle-related knowledge and adopt lifestyle changes. Second, while efforts were undertaken to recruit participants through various channels worldwide, this study also primarily involved participants from English-speaking countries with convenient internet able to access the recruitment sites. Consequently, the generalizability of the results may be limited, warranting caution in their interpretation. Third, the sample size of noncompleters across study arms may be considered small. This increases the risk of selection bias, potentially limiting the generalizability of the findings to a broader population of noncompleters. Nevertheless, identifying barriers to course completion, such as time limitations and technical and health-related issues, provides important insights for informing the development of web-based programs. For example, extending the course to >6 weeks may allow more participants to complete it, a strategy supported by findings from the Understanding MS MOOC study [36]. Other considerations may include presenting participants with incentives to complete the MSOC, such as completion certificates [56]. Furthermore, it is important to consider health disparities within the MS community to ensure that course content is appropriate for web-based interventions.

Conclusions

The evaluation of the MSOC in this study provided new and valuable insights for the development of future web-based programs for people living with MS. Participants expressed high levels of satisfaction with the MSOC and strong intentions to adopt lifestyle modifications, even among those already familiar with the course content. Notably, course completers with poorer health had stronger intentions to adopt recommended lifestyle changes upon completing the course, underscoring the potential value of web-based interventions among those most likely to benefit. Identifying motivators and

XSL•FO
RenderX

barriers to web-based course engagement could inform future course development to facilitate increased engagement and potentially achieve intervention objectives. Subsequent follow-up studies of the RCT's primary and secondary outcomes will examine the effectiveness of the MSOC intervention in supporting the adoption and maintenance of lifestyle changes and achieving the primary outcome of the study (clinically significant improvements in QoL). Overall, tailored, accessible, and innovative web-based interventions offer an important platform to assist people living with MS in self-managing their MS and related symptoms and potentially improving their future health prospects.

Acknowledgments

The authors express their sincere gratitude to all participants of the Multiple Sclerosis Online Course for their generous contribution of time and spirit. Their participation was invaluable, and without it, this study could not have been possible. They also acknowledge the role of their industry partner, JMAcreative, in the redevelopment of the courses. The authors extend special thanks to the project manager, Grace Suhartanto, for her support, guidance, and expertise in redesigning, testing, and implementing the courses. The authors also express their gratitude to the Overcoming MS charity (United Kingdom), who financially supported the technical development of the Multiple Sclerosis Online Course via JMAcreative. This work was supported by the Neuroepidemiology Unit at the University of Melbourne, which is funded by philanthropic donors.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

MY contributed to conceptualization, methodology, data curation, formal analysis, writing of the original draft, and writing—review and editing. SS-Y contributed to methodology, formal analysis, and writing—review and editing. WB contributed to investigation and writing—review and editing. RD contributed to methodology and writing—review and editing. GJ contributed to funding acquisition, investigation, and writing—review and editing. SN contributed to study design, investigation, and writing—review and editing. JR contributed to study design, project administration, conceptualization, methodology, writing of the original draft, investigation, and writing—review and editing.

Conflicts of Interest

GJ and SN previously received remuneration for facilitating lifestyle modification workshops. GJ receives royalties from his book Overcoming Multiple Sclerosis: The 7 Step Recovery Program. GJ and SN receive royalties from their book The Overcoming Multiple Sclerosis Handbook: Roadmap to Good Health.

Multimedia Appendix 1
Postcourse evaluation questions for participants.
[PDF File (Adobe PDF File), 165 KB - humanfactors_v12i1e59363_app1.pdf]

Multimedia Appendix 2
Microsoft PowerPoint (Microsoft Corp) presentation of the conference talk about the Multiple Sclerosis Online Course.
[PPTX File , 5584 KB - humanfactors_v12i1e59363_app2.pptx]

Multimedia Appendix 3
Screenshots of the Multiple Sclerosis Online Course website.
[PPTX File , 249 KB - humanfactors_v12i1e59363_app3.pptx]

Multimedia Appendix 4
Content of the standard-care course and intervention course and countries of residence of Multiple Sclerosis Online Course participants.
[DOCX File , 24 KB - humanfactors_v12i1e59363_app4.docx]

Multimedia Appendix 5
CONSORT-eHEALTH (V 1.6.1) checklist.
[PDF File (Adobe PDF File), 1098 KB - humanfactors_v12i1e59363_app5.pdf]

References

XSL•FO
RenderX

1. Walton C, King R, Rechtman L, Kaye W, Leray E, Marrie RA, et al. Rising prevalence of multiple sclerosis worldwide: insights from the Atlas of MS, third edition. Mult Scler 2020 Dec 11;26(14):1816-1821 [FREE Full text] [doi: 10.1177/1352458520970841] [Medline: 33174475]
2. Schirmer L, Velmeshev D, Holmqvist S, Kaufmann M, Werneburg S, Jung D, et al. Neuronal vulnerability and multilineage diversity in multiple sclerosis. Nature 2019 Sep 17;573(7772):75-82 [FREE Full text] [doi: 10.1038/s41586-019-1404-z] [Medline: 31316211]
3. Boeschoten RE, Braamse AM, Beekman AT, Cuijpers P, van Oppen P, Dekker J, et al. Prevalence of depression and anxiety in multiple sclerosis: a systematic review and meta-analysis. J Neurol Sci 2017 Jan 15;372:331-341. [doi: 10.1016/j.jns.2016.11.067] [Medline: 28017241]
4. Strober LB, Arnett PA. An examination of four models predicting fatigue in multiple sclerosis. Arch Clin Neuropsychol 2005 Jul;20(5):631-646. [doi: 10.1016/j.acn.2005.04.002] [Medline: 15894455]
5. Tettey P, Siejka D, Simpson S, Taylor B, Blizzard L, Ponsonby A, et al. Frequency of comorbidities and their association with clinical disability and relapse in multiple sclerosis. Neuroepidemiology 2016 Jan 20;46(2):106-113. [doi: 10.1159/000442203] [Medline: 26784322]
6. Bandura A. Health promotion by social cognitive means. Health Educ Behav 2004 Apr 01;31(2):143-164. [doi: 10.1177/1090198104263660] [Medline: 15090118]
7. Simpson-Yap S, Neate SL, Nag N, Probst YC, Yu M, Jelinek GA, et al. Longitudinal associations between quality of diet and disability over 7.5 years in an international sample of people with multiple sclerosis. Eur J Neurol 2023 Oct;30(10):3200-3211. [doi: 10.1111/ene.15980] [Medline: 37433564]
8. Simpson-Yap S, Nag N, Probst Y, Reece JC, Jelinek GA, Neate S. Prospective associations of better quality of the diet with improved quality of life over 7.5 years in people with multiple sclerosis. Mult Scler Relat Disord 2022 Apr;60:103710 [FREE Full text] [doi: 10.1016/j.msard.2022.103710] [Medline: 35219239]
9. Stephens S, Shams S, Lee J, Grover SA, Longoni G, Berenbaum T, et al. Benefits of physical activity for depression and fatigue in multiple sclerosis: a longitudinal analysis. J Pediatr 2019 Jun;209:226-32.e2. [doi: 10.1016/j.jpeds.2019.01.040] [Medline: 30878208]
10. Briones-Buixassa L, Milà R, Mª Aragonès J, Bufill E, Olaya B, Arrufat FX. Stress and multiple sclerosis: a systematic review considering potential moderating and mediating factors and methods of assessing stress. Health Psychol Open 2015 Jul 04;2(2):2055102915612271 [FREE Full text] [doi: 10.1177/2055102915612271] [Medline: 28070374]
11. Moss-Morris R, McCrone P, Yardley L, van Kessel K, Wills G, Dennison L. A pilot randomised controlled trial of an internet-based cognitive behavioural therapy self-management programme (MS Invigor8) for multiple sclerosis fatigue. Behav Res Ther 2012 Jun;50(6):415-421 [FREE Full text] [doi: 10.1016/j.brat.2012.03.001] [Medline: 22516321]
12. van Kessel K, Wouldes T, Moss-Morris R. A New Zealand pilot randomized controlled trial of a web-based interactive self-management programme (MSInvigor8) with and without email support for the treatment of multiple sclerosis fatigue. Clin Rehabil 2016 May 07;30(5):454-462. [doi: 10.1177/0269215515584800] [Medline: 25952587]
13. De Angelis M, Lavorgna L, Carotenuto A, Petruzzo M, Lanzillo R, Brescia Morra V, et al. Digital technology in clinical trials for multiple sclerosis: systematic review. J Clin Med 2021 May 26;10(11):2328 [FREE Full text] [doi: 10.3390/jcm10112328] [Medline: 34073464]
14. Heesen C, Lühmann D. Reply: mobile health interventions in multiple sclerosis: a systematic review. Mult Scler 2024 Apr 29;30(4-5):619-620. [doi: 10.1177/13524585241234777] [Medline: 38420930]
15. Wills OC, Probst YC. Understanding lifestyle self-management regimens that improve the life quality of people living with multiple sclerosis: a systematic review and meta-analysis. Health Qual Life Outcomes 2022 Nov 25;20(1):153 [FREE Full text] [doi: 10.1186/s12955-022-02046-1] [Medline: 36434609]
16. Bevens W, Weiland T, Gray K, Jelinek G, Neate S, Simpson-Yap S. Attrition within digital health interventions for people with multiple sclerosis: systematic review and meta-analysis. J Med Internet Res 2022 Feb 09;24(2):e27735 [FREE Full text] [doi: 10.2196/27735] [Medline: 35138262]
17. Claflin SB, Campbell JA, Doherty K, Farrow M, Bessing B, Taylor BV. Evaluating course completion, appropriateness, and burden in the understanding multiple sclerosis massive open online course: cohort study. J Med Internet Res 2021 Dec 07;23(12):e21681 [FREE Full text] [doi: 10.2196/21681] [Medline: 34878985]
18. Claflin SB, Klekociuk S, Fair H, Bostock E, Farrow M, Doherty K, et al. Assessing the impact of online health education interventions from 2010-2020: a systematic review of the evidence. Am J Health Promot 2022 Jan 12;36(1):201-224. [doi: 10.1177/08901171211039308] [Medline: 34382448]
19. Santarossa S, Kane D, Senn CY, Woodruff SJ. Exploring the role of in-person components for online health behavior change interventions: can a digital person-to-person component suffice? J Med Internet Res 2018 Apr 11;20(4):e144 [FREE Full text] [doi: 10.2196/jmir.8480] [Medline: 29643048]
20. Reece J, Yu M, Bevens W, Simpson-Yap S, Davenport R, Jelinek G, et al. Sociodemographic, health, and lifestyle-related characteristics associated with the commencement and completion of a web-based lifestyle educational program for people with multiple sclerosis: randomized controlled trial. J Med Internet Res 2024 Aug 28;26:e58253 [FREE Full text] [doi: 10.2196/58253] [Medline: 39197163]

XSL•FO
RenderX

21. Bevens WD, Davenport R, Neate S, Yu M, Jelinek P, Jelinek GA, et al. Web-based health information seeking by people living with multiple sclerosis: qualitative investigation of the multiple sclerosis online course. J Med Internet Res 2024 Feb 09;26:e53372 [FREE Full text] [doi: 10.2196/53372] [Medline: 38335016]
22. Davenport R, Bevens W, Neate S, Jelinek P, Yu M, Jelinek G, et al. Perceptions of illness control, coherence, and self-efficacy following a web-based lifestyle program for multiple sclerosis: a qualitative analysis of semistructured interviews. J Med Internet Res 2024 Nov 29;26:e60240 [FREE Full text] [doi: 10.2196/60240] [Medline: 39612211]
23. Jelinek GA. Overcoming Multiple Sclerosis: The Evidence-Based 7 Step Recovery Program. 2nd edition. Melbourne, Australia: Allen & Unwin; 2016.
24. Reece JC, Yu M, Bevens W, Simpson-Yap S, Jelinek G, Jelinek P, et al. Study protocol for an online lifestyle modification education course for people living with multiple sclerosis: the multiple sclerosis online course (MSOC). BMC Neurol 2023 Jun 29;23(1):249 [FREE Full text] [doi: 10.1186/s12883-023-03298-0] [Medline: 37386385]
25. Bevens W, Weiland TJ, Gray K, Neate SL, Nag N, Simpson-Yap S, et al. The feasibility of a web-based educational lifestyle program for people with multiple sclerosis: a randomized controlled trial. Front Public Health 2022 Apr 27;10:852214 [FREE Full text] [doi: 10.3389/fpubh.2022.852214] [Medline: 35570898]
26. Learmonth YC, Motl RW, Sandroff BM, Pula JH, Cadavid D. Validation of patient determined disease steps (PDDS) scale scores in persons with multiple sclerosis. BMC Neurol 2013 Apr 25;13(1):37 [FREE Full text] [doi: 10.1186/1471-2377-13-37] [Medline: 23617555]
27. Hohol MJ, Orav EJ, Weiner HL. Disease steps in multiple sclerosis: a simple approach to evaluate disease progression. Neurology 1995 Feb;45(2):251-255. [doi: 10.1212/wnl.45.2.251] [Medline: 7854521]
28. Rizzo MA, Hadjimichael OC, Preiningerova J, Vollmer TL. Prevalence and treatment of spasticity reported by multiple sclerosis patients. Mult Scler 2004 Oct 01;10(5):589-595. [doi: 10.1191/1352458504ms1085oa] [Medline: 15471378]
29. Jones KH, Ford DV, Jones PA, John A, Middleton RM, Lockhart-Jones H, et al. A large-scale study of anxiety and depression in people with multiple sclerosis: a survey via the web portal of the UK MS Register. PLoS One 2012 Jul 30;7(7):e41910 [FREE Full text] [doi: 10.1371/journal.pone.0041910] [Medline: 22860028]
30. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983 Jun 23;67(6):361-370. [doi: 10.1111/j.1600-0447.1983.tb09716.x] [Medline: 6880820]
31. Watson TM, Ford E, Worthington E, Lincoln NB. Validation of mood measures for people with multiple sclerosis. Int J MS Care 2014;16(2):105-109 [FREE Full text] [doi: 10.7224/1537-2073.2013-013] [Medline: 25061435]
32. Vickrey BG, Hays RD, Harooni R, Myers LW, Ellison GW. A health-related quality of life measure for multiple sclerosis. Qual Life Res 1995 Jun;4(3):187-206. [doi: 10.1007/bf02260859]
33. Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. J Clin Oncol 1998 Jan;16(1):139-144. [doi: 10.1200/JCO.1998.16.1.139] [Medline: 9440735]
34. Taphoorn MJ, Bottomley A. Health-related quality of life and symptom research in glioblastoma multiforme patients. Expert Rev Pharmacoecon Outcomes Res 2005 Dec 09;5(6):763-774. [doi: 10.1586/14737167.5.6.763] [Medline: 19807618]
35. Longhini J, Rossettini G, Palese A. Massive open online courses for nurses' and healthcare professionals' continuous education: a scoping review. Int Nurs Rev 2021 Mar 15;68(1):108-121. [doi: 10.1111/inr.12649] [Medline: 33855697]
36. Claflin S, Klekociuk S, Campbell J, Taylor B. Reasons for non-completion of a massive open online course about multiple sclerosis: a mixed methods study. Mult Scler Relat Disord 2022 Nov;67:104092. [doi: 10.1016/j.msard.2022.104092] [Medline: 35977440]
37. Marrie RA, Leung S, Tyry T, Cutter GR, Fox R, Salter A. Use of eHealth and mHealth technology by persons with multiple sclerosis. Mult Scler Relat Disord 2019 Jan;27:13-19. [doi: 10.1016/j.msard.2018.09.036] [Medline: 30296732]
38. Neate SL, Bevens W, Jelinek PL, Gray KM, Weiland TJ, Nag N, et al. A multiple sclerosis lifestyle behavior online course: qualitative analysis of participants' motivations, expectations and experiences. Front Public Health 2022 Dec 7;10:1022185 [FREE Full text] [doi: 10.3389/fpubh.2022.1022185] [Medline: 36568793]
39. Bevens W, Reece J, Jelinek P, Weiland T, Nag N, Simpson-Yap S, et al. The feasibility of an online educational lifestyle program for people with multiple sclerosis: a qualitative analysis of participant semi-structured interviews. Digit Health 2022 Sep 04;8:20552076221123713 [FREE Full text] [doi: 10.1177/20552076221123713] [Medline: 36081754]
40. Claflin SB, Gates R, Maher M, Taylor BV. Building a successful massive open online course about multiple sclerosis: a process description. J Med Internet Res 2020 Jul 29;22(7):e16687 [FREE Full text] [doi: 10.2196/16687] [Medline: 32723726]
41. van der Smissen D, Rietjens JA, van Gemert-Pijnen L, van Dulmen S, van der Heide A, Korfage IJ. Information needs of patients with chronic diseases and their relatives for web-based advance care planning: a qualitative interview study. BMC Palliat Care 2021 May 30;20(1):77 [FREE Full text] [doi: 10.1186/s12904-021-00770-x] [Medline: 34053435]
42. Hawkes RE, Benton JS, Cotterill S, Sanders C, French DP. Service users' experiences of a nationwide digital type 2 diabetes self-management intervention (healthy living): qualitative interview study. JMIR Diabetes 2024 Jul 18;9:e56276 [FREE Full text] [doi: 10.2196/56276] [Medline: 39024002]
43. Weidmann AE, Sonnleitner-Heglmeier A, Dartsch DC. Effect of patient education videos on modifying medication-related health behaviours: a systematic review using the behaviour intervention functions. Patient Educ Couns 2023 Dec;117:107992 [FREE Full text] [doi: 10.1016/j.pec.2023.107992] [Medline: 37788537]

XSL•FO
RenderX

44. Hossain SN, Jaglal SB, Shepherd J, Perrier L, Tomasone JR, Sweet SN, et al. Web-based peer support interventions for adults living with chronic conditions: scoping review. JMIR Rehabil Assist Technol 2021 May 25;8(2):e14321 [FREE Full text] [doi: 10.2196/14321] [Medline: 34032572]
45. Borghouts J, Eikey E, Mark G, De Leon C, Schueller SM, Schneider M, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. J Med Internet Res 2021 Mar 24;23(3):e24387 [FREE Full text] [doi: 10.2196/24387] [Medline: 33759801]
46. Ajzen I. The theory of planned behavior. Organ Behav Hum Decis Process 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
47. Claflin SB, Mainsbridge C, Campbell J, Klekociuk S, Taylor BV. Self-reported behaviour change among multiple sclerosis community members and interested laypeople following participation in a free online course about multiple sclerosis. Health Promot J Austr 2022 Jul 13;33(3):768-778. [doi: 10.1002/hpja.559] [Medline: 34807490]
48. Roberts AL, Fisher A, Smith L, Heinrich M, Potts HW. Digital health behaviour change interventions targeting physical activity and diet in cancer survivors: a systematic review and meta-analysis. J Cancer Surviv 2017 Dec 4;11(6):704-719. [doi: 10.1007/s11764-017-0632-1] [Medline: 28779220]
49. Akinosun AS, Polson R, Diaz-Skeete Y, De Kock JH, Carragher L, Leslie S, et al. Digital technology interventions for risk factor modification in patients with cardiovascular disease: systematic review and meta-analysis. JMIR Mhealth Uhealth 2021 Mar 03;9(3):e21061. [doi: 10.2196/21061] [Medline: 33656444]
50. Kim S, Sargent-Cox KA, Anstey KJ. A qualitative study of older and middle-aged adults' perception and attitudes towards dementia and dementia risk reduction. J Adv Nurs 2015 Jul 20;71(7):1694-1703. [doi: 10.1111/jan.12641] [Medline: 25702823]
51. Heikkinen M. Power and support from the net: usability and sociability on an internet-based rehabilitation course for people with multiple sclerosis. Int J Web Based Communities 2009;5(1):83. [doi: 10.1504/ijwbc.2009.021563]
52. Barnard E, Brown CR, Weiland TJ, Jelinek GA, Marck CH. Understanding barriers, enablers, and long-term adherence to a health behavior intervention in people with multiple sclerosis. Disabil Rehabil 2020 Mar 30;42(6):822-832 [FREE Full text] [doi: 10.1080/09638288.2018.1510550] [Medline: 30375258]
53. Barrera Jr M, Toobert DJ, Angell KL, Glasgow RE, Mackinnon DP. Social support and social-ecological resources as mediators of lifestyle intervention effects for type 2 diabetes. J Health Psychol 2006 May 01;11(3):483-495. [doi: 10.1177/1359105306063321] [Medline: 16774900]
54. O'Kearney EL, Brown C, Jelinek G, Neate S, Taylor K, Bevens W, et al. Mastery is associated with greater physical and mental health-related quality of life in two international cohorts of people with multiple sclerosis. Mult Scler Relat Disord 2020 Feb;38:101481. [doi: 10.1016/j.msard.2019.101481] [Medline: 31756608]
55. Farrow M, Fair H, Klekociuk SZ, Vickers JC. Educating the masses to address a global public health priority: the Preventing Dementia Massive Open Online Course (MOOC). PLoS One 2022 May 4;17(5):e0267205 [FREE Full text] [doi: 10.1371/journal.pone.0267205] [Medline: 35507576]
56. Ogunyemi AA, Quaicoe JS, Bauters M. Indicators for enhancing learners' engagement in massive open online courses: a systematic review. Comput Educ Open 2022 Dec;3:100088. [doi: 10.1016/j.caeo.2022.100088]

Abbreviations

DMT: disease-modifying therapy
IC: intervention course
MOOC: massive open online course
MS: multiple sclerosis
MSOC: Multiple Sclerosis Online Course
MSQOL-54: Multiple Sclerosis Quality of Life–54
QoL: quality of life
RCT: randomized controlled trial
SCC: standard-care course

XSL•FO
RenderX

Edited by A Kushniruk; submitted 18.04.24; peer-reviewed by R Haase; comments to author 28.11.24; revised version received 22.01.25; accepted 04.04.25; published 26.05.25.

Please cite as:
Yu M, Neate S, Simpson-Yap S, Davenport R, Bevens W, Jelinek G, Reece J
A Web-Based Lifestyle-Related Course for People Living With Multiple Sclerosis: Quantitative Evaluation of Course Completion, Satisfaction, and Lifestyle Changes Among Participants Enrolled in a Randomized Controlled Trial
JMIR Hum Factors 2025;12:e59363
URL: https://humanfactors.jmir.org/2025/1/e59363
doi:10.2196/59363
PMID:

©Maggie Yu, Sandra Neate, Steve Simpson-Yap, Rebekah Davenport, William Bevens, George Jelinek, Jeanette Reece. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 26.05.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Patients' Perceptions of Using a Digital Previsit Tool in Outpatient Settings (Part 2): Qualitative Study

Petra Pohl[1*], PhD; Emma K Kjörk[2,3,4*], PhD

[1]Department of Health and Rehabilitation, Sahlgrenska Academy, Institute of Neuroscience and Physiology, University of Gothenburg, Gothenburg, Sweden

[2]Department of Clinical Neuroscience, Institute of Neuroscience and Physiology, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

[3]Centre for Person-Centred Care (GPCC), University of Gothenburg, Gothenburg, Sweden

[4]Sahlgrenska University Hospital, Gothenburg, Sweden

[*]all authors contributed equally

Corresponding Author:
Emma K Kjörk, PhD
Department of Clinical Neuroscience
Institute of Neuroscience and Physiology, Sahlgrenska Academy
University of Gothenburg
Rehabiliteringsmedicin, Vita stråket 12, vån 4
Sahlgrenska universitetssjukhuset
Gothenburg, 41345
Sweden
Phone: 46 31 3422391
Email: emma.kjork@neuro.gu.se

Abstract

Background: Patients with long-term conditions, such as stroke, require regular follow-up visits to health care professionals to identify changes in symptoms. The digital previsit tool Strokehälsa (Strokehealth) has been designed to encourage individuals with stroke to reflect on stroke-related health concerns before a follow-up visit, thereby potentially enhancing their engagement during the visit. Strokehealth has previously been evaluated using a patient satisfaction survey (part 1), but there remains a need to further explore patients' perceptions and needs to optimize its functionality before broader implementation.

Objective: The overall aim was to attain deeper insights into patients' views and experiences of using the digital previsit tool Strokehealth before a follow-up visit. A secondary aim was to identify potential improvements to the tool based on these insights.

Methods: For this qualitative study, patients who had used Strokehealth version 1.0 before a follow-up visit were recruited through the previous survey between November 2020 and June 2021. Individual semistructured interviews were conducted, and data were analyzed using reflexive thematic analysis. Subsequent workshops were held with people with firsthand experiences of stroke, other stakeholders (including health professionals and researchers), and a web consultant to finalize decisions regarding adjustments to be implemented in Strokehealth version 2.0.

Results: Interviews were conducted with 33 participants (23 men and 10 women), with a median age of 67 (IQR 55-76) years. Analysis of the data regarding participants' experiences of using Strokehealth revealed three overarching themes: (1) a supporting tool for preparing dialogue and identifying needs, (2) how Strokehealth is introduced and communicated affects perceived usability, and (3) the wording and structure of Strokehealth influences the response process. The findings captured various aspects of receiving and using the digital previsit tool, highlighting its simplicity and purpose. Overall, Strokehealth was well received and contributed to a sense of being well cared for. Participants generally not only found Strokehealth easy to use but also shared suggestions on how to better address stroke-related issues, such as mental fatigue or pain. Examples of changes that have been implemented in Strokehealth version 2.0, based on participant feedback, include improved explanatory texts and expanded opportunities for free text.

Conclusions: The findings indicate that the freely available digital previsit tool Strokehealth was generally well received by patients with stroke who were scheduled for follow-up visits in outpatient settings.

Trial Registration: Researchweb 275135; https://www.researchweb.org/is/vgr/project/275135

XSL•FO
RenderX

(JMIR Hum Factors 2025;12:e73477) doi:10.2196/73477

KEYWORDS

stroke rehabilitation; post-stroke checklist; person-centered care; follow-up; eHealth; health literacy; patient engagement

Introduction

Health care systems are increasingly adopting digital support to address future challenges and to transition toward more person-centered care [1]. This shift has prompted a significant increase in the development and use of digital services and tools for self-care among community-dwelling people [2]. Advancements in medical care and improvements in health care services are leading to rising numbers of individuals living with conditions that require long-term care, such as stroke [3]. In patients with such conditions, a lack of adequate self-care or health care interventions can lead to serious consequences [3]. Continuous access to health care services is essential to ensure optimal self-management, psychological and emotional support [4,5], and treatment for patients, regardless of their condition [6]. The integration of digital technology can potentially enhance easy access to health care, patient engagement, and self-management [2,7]. However, individuals managing long-term conditions commonly show low eHealth literacy, potentially resulting in their exclusion from digital health resources [8]. Therefore, it is imperative that digital tools be tailored specifically for individuals with diverse health conditions and be designed with enhanced usability and accessibility.

Regular follow-up visits enable caregivers to identify potential health problems, detect symptom changes or worsening at an early stage, and take preventive actions to minimize their impact [9]. Follow-up visits also serve as opportunities to provide patients with support, education, and counseling regarding their condition, treatment options, and symptom management. This helps patients better understand their condition and empowers them to take a more active role in their health care. During follow-up visits, meaningful conversations can be facilitated using tools such as checklists or dialogue aids, for example, the Post Stroke Checklist [10]. However, patients have emphasized the importance of advance preparation to enhance engagement during discussions with health care professionals [11]. Such preparation can be achieved through previsit tools [12,13], which encourage patients to reflect in advance, and thereby make interactions with health care providers more efficient. Accordingly, a previsit tool for individuals with stroke—Strokehälsa ("Strokehealth"), based on the Post Stroke Checklist—was developed in a participatory co-design process with a patient partner, stakeholders, and a service designer [14]. The aim was to create a tool that would enable people—including those with various disabilities—to take an active part in their decision-making process by providing them with targeted questions and time to reflect on stroke-related concerns before their scheduled visit to an outpatient clinic specialized in stroke care [14]. Other previsit tools exist but most are designed to collect physical health data, such as blood glucose levels, or are highly comprehensive (eg, the stroke-related "Rehabkompassen" [15]), and require substantial time and effort from the user. In contrast, Strokehealth was designed to support reflection and communication, rather than extensive data collection.

In short, the content of Strokehealth and the procedure for its use in health care can be described as follows. First, the patient is introduced to Strokehealth version 1.0 through a brief introductory text before answering 14 questions (Multimedia Appendix 1) covering various health-related domains. Each question includes a short explanatory text to clarify and provide examples. To indicate any health concerns, the response choices are "yes," "no," or "choose not to answer." Upon completing Strokehealth, patients have the opportunity to provide additional open-ended comments or information before their scheduled visit. In addition, patients are offered further advisory information about stroke and self-management strategies, which is accessible via a clickable link. In clinical practice, health care providers send Strokehealth to patients via the Swedish national patient portal, prompting them by email or text message to log in and complete the tool. The health care provider—usually a stroke-specialized nurse—then has the option to review the patient's responses in advance and, if necessary, consult other professionals—for example, a physician or an occupational therapist—prior to the appointment, in order to proactively address the patient's concerns, although this step is optional.

In the development and evaluation of complex interventions, contextual factors—including organizational and local conditions—play a decisive role [16]. Therefore, Strokehealth was evaluated in a real-world setting involving patients who were discharged from a stroke unit and scheduled for follow-up visits. A previous study (part 1) was conducted using a web-based patient satisfaction survey, revealing that Strokehealth was perceived to provide a structured framework that facilitated the consideration of important issues in conjunction with health care visits [17]. A wide range of health problems was identified, and 56 out of 58 (96.6%) participants expressed confidence that Strokehealth effectively captured their health-related concerns. Analysis of free-text responses revealed that Strokehealth was viewed as a valuable tool for structured follow-up and was described as enabling increased patient engagement. Participants also noted that Strokehealth provided valuable information, offering enhanced insights into their condition and the available support. Suggestions for improvement included extended response options, the opportunity to develop one's answers, and clearer explanatory texts [17]. Sustainable development and implementation of the previsit tool require evaluating feasibility from different perspectives, especially from the patients' perspective as end users. Thus, it was considered necessary to conduct individual interviews to explore the participants' survey responses in depth and to allow them to elaborate on their suggestions for improvements. The present interview study is an important step in the development and evaluation of Strokehealth before its implementation in health care.

XSL•FO
RenderX

In this study, we aimed to gain deeper insights into patients' experiences with and views on using the Strokehealth digital previsit tool before follow-up visits. A secondary aim was to identify further ways to improve the previsit tool based on these findings.

Methods

Study Design

A qualitative study was conducted [18] as part of a larger research project aimed at developing and evaluating the Strokehealth previsit tool, using quantitative and qualitative methods and following the Medical Research Council framework [16]. The research paradigm guiding this study, as part of the larger project, is inspired by the principles of realistic evaluation based on scientific realism, in which various data sources are used in a pragmatic and flexible manner to explore "what works, for whom, under what circumstances?" [19]. This study involved the interpretative analysis of semistructured interviews to attain insights regarding how different contexts and mechanisms influence outcomes and to inform changes to the previsit tool. In this study, the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist was used to ensure explicit and comprehensive reporting (Multimedia Appendix 2) [20].

Ethical Considerations

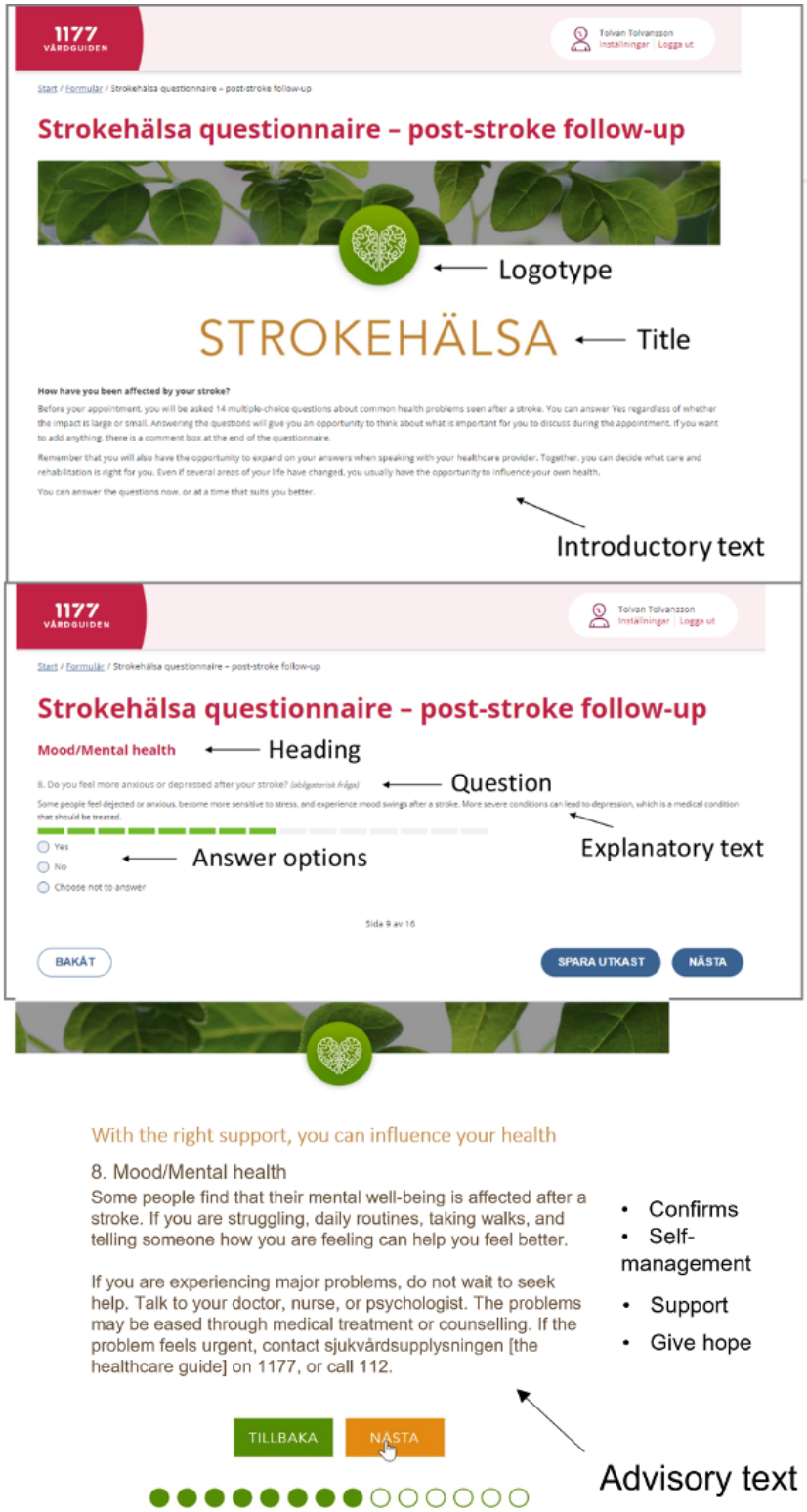
Ethical approval was obtained from the Swedish Ethical Review Authority (556-17 and 2020-03324). After logging into the patient portal, patients were presented with information regarding the research project and asked whether they would consent to participate. Only those who provided informed consent were able to proceed with the survey (part 1) [17]. The introductory information made it clear that participation was voluntary and would not impact their medical treatment. No incentives were offered for participation. At the end of the survey, participants were asked a follow-up question about participating in an interview. All transcripts were anonymized before analysis and treated confidentially because they contained health-related information. Only the research team had access to the interview recordings, transcripts, and consent forms. No prior relationship was established between the researchers and participants before the interviews.

Setting and Participants

This study was conducted in the Västra Götaland region of Sweden. The study participants were community-dwelling individuals who had experienced a stroke, had recently used Strokehealth (Figure 1), completed a subsequent patient satisfaction survey, and volunteered to share their more in-depth experiences regarding the tool. Initial participant recruitment was performed consecutively between November 2020 and June 2021 from 3 hospitals in the region. The inclusion criteria were admission to a stroke unit, being scheduled for a follow-up visit to a stroke team nurse after hospital discharge, and having a personal account on the Swedish national patient portal. No exclusion criteria were set. Strokehealth was sent approximately 3 weeks to 3 months after discharge from the hospital, depending on the standard routines of the outpatient clinic.

XSL•FO
RenderX

Figure 1. A screenshot of the elements in the previsit tool Strokehälsa (English version; reproduced from Kjörk et al [14], which is published under Creative Commons Attribution 4.0 International License [21].



A web-based survey was sent to all patients who had used the tool (n=80). The survey was completed by 73% (58/80; 39 men; age: median 67, IQR 56-75 years). The survey included information about its purpose: to gain a better understanding of their experiences using the previsit tool Strokehealth in preparation for their follow-up visit, with a focus on its usability and functionality. At the end of the survey, respondents were asked whether they consented to future contact from the research team regarding a more in-depth interview. It was emphasized that the interviews would contribute to improving follow-up care supported by the tool. A total of 36 patients indicated their interest in further collaboration and were contacted via telephone by the last author (EKK), of whom 33 agreed to participate in individual interviews.

XSL•FO
RenderX

Data Collection

Data were collected using a semistructured interview guide, including questions about experiences and aspects of using Strokehealth (Multimedia Appendix 3). The interview guide was pilot-tested, and no adjustments were deemed necessary; therefore, the pilot interview was included in the present analysis. All interviews were conducted in Swedish over the telephone at prearranged dates and times. At the time of the interviews, only a few participants had completed their follow-up visits. No relationship had been established between the interviewer and the participants before the interviews. The participants' next of kin were allowed to be present and to participate when appropriate. During 2 interviews, a spouse was present for support.

All interviews were audio recorded. The study aims were repeated, and the participants orally confirmed their consent. To ensure high credibility, all interviews began with the same introduction, emphasizing the importance of the participants' experiences and viewpoints and that there were no right or wrong answers. Prompts such as "Could you elaborate on your thoughts?" were used. The duration of the interviews ranged from 8 to 40 minutes (median 24 minutes, IQR 21-28 minutes). All interviews were transcribed verbatim by a person who was not part of the study, using MS Word. The transcripts comprised 289 pages of single-spaced text. No member checks were conducted, and participants were not asked to provide feedback on the results.

Data Analysis

A reflexive thematic analysis [22] was conducted, aligning with an interpretivist paradigm that emphasizes understanding individuals' subjective experiences [18]. Reflexive thematic analysis was deemed an appropriate analytical approach based on the study's focus on the unique experiences of the patients who had used Strokehealth. Furthermore, reflexive thematic analysis entails a flexible process that facilitates the identification, analysis, and reporting of patterns of meaning within interview data [22]. In reflexive thematic analysis, both coding and theme development are flexible processes that evolve throughout the analysis. An inductive approach was used, that is, the analysis was data driven and based on the participants' subjective experiences. Analysis was performed by 2 authors (EKK and PP) using the NVivo software (version 12; Lumivero).

The analysis was performed using the 6-stage framework described by Braun and Clarke (Textbox 1) [22]. Throughout the analysis process, conscious reflections were made, and reflexive thematic analysis facilitated the authors' reflective engagement.

Textbox 1. The 6-stage framework of reflexive thematic analysis.

Stage 1. Both authors independently read and reread all transcripts to become familiar with the data.
Stage 2. Important features related to the research question were identified within the dataset and coded according to their content.
Stage 3. The codes were examined to identify significant patterns of meaning and clustered into potential themes and subthemes.
Stage 4. Themes and subthemes were reviewed against the dataset, and redefined in an iterative process, with continuous back-and-forth between the main texts and the codes.
Stage 5. The themes were analyzed in detail to refine their scope and focus, and informative names were assigned.
Stage 6. Analytical narrative and data extracts were contextualized in relation to existing literature, and the manuscript was produced.

After the completion of the analysis, some of the suggested changes that emerged from the interviews were implemented in Strokehealth, resulting in version 2.0. This process included consultation with members of the advisory board who were involved in the previous study [14]. In addition, a linguist reviewed the text to improve its clarity and structure. Finally, the proposed amendments to Strokehealth were discussed in formal decision-making workshops with the project's research team in collaboration with the patient partner. After the final decision workshop, adjustments to Strokehealth in the patient portal were made through close collaboration between EKK and a web designer. These workshops guided the final decisions about the design of Strokehealth and addressed issues such as clarifying the language by adding or removing text and answer options.

Research Team and Reflexivity

The interviews were conducted by EKK, a female who holds a PhD degree in occupational therapy and is a senior lecturer. She is an established researcher with previous experience in qualitative research and is well published in the fields of stroke rehabilitation and eHealth. Transcripts were cross-checked by PP, a female who holds a PhD degree in physical therapy. She is an assistant professor with previous experience in qualitative research and is well published in the fields of neurological rehabilitation and eHealth.

Results

Participants

The 33 study participants included 23 men and 10 women, with a median age of 67 (range 43-87) years (Table 1). Among these participants, 28 (85%) reported using the web multiple times per day, 3 (9%) reported using it weekly, and 2 (6%) reported using it once per year. Four participants (12%) received assistance in completing Strokehealth.

XSL•FO
RenderX

Table 1. Participant characteristics.

Characteristics	Participants in interviews (n=33)
Age (years)	
All patients, median (minimum-maximum)	67 (43-87)
Sex, n (%)	
Male	23 (70)
Female	10 (30)
Education (highest level), n (%)	
Mandatory	10 (30)
High school	8 (24)
University	14 (42)
Other	1 (3)
Source of income at inclusion, n (%)	
Work	13 (39)
Sick leave	2 (6)
Retirement	17 (52)
Other	1 (3)
Prestroke living conditions, n (%)[a]	
Without assisted care in own home	32 (100)
Living alone	5 (15)
Independent in activities of daily living	30 (94)
Stroke characteristics (onset)	
Cerebral infarct, n (%)	31 (94)
Intracerebral hemorrhage, n (%)	2 (6)
Previous stroke, n (%)	5 (15)
Stroke severity, NIHSS[b], median (minimum-maximum)	1 (0-7)
Stroke-related outcomes	
Length of hospital stay in days median (minimum-maximum)	5.5 (1-35)
Discharged to own home, n (%)	33 (100)

[a]Missing data: without assisted care in own home (n=1), independent (n=1).

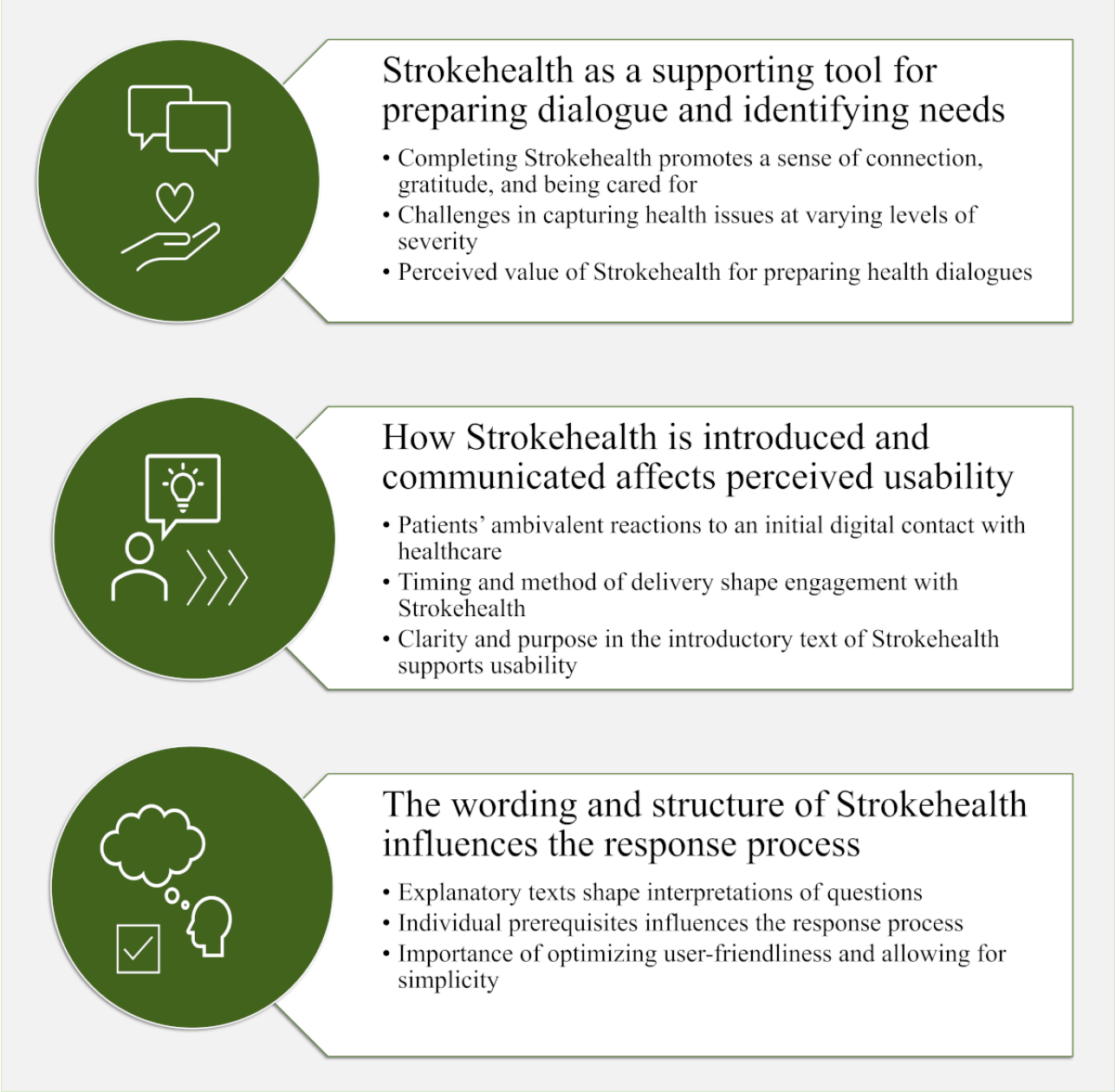
[b]NIHSS: National Institutes of Health Stroke Scale; measured ≤24 hours of admission, normal values 0-42. Values are presented as numbers and valid percentages unless stated otherwise.

Themes Created Based on the Individual Interviews

The analysis resulted in the identification of 3 main themes, reflecting the participants' overall experiences with receiving and using the previsit tool before their planned follow-up visit to the nurse. The three main themes were as follows: (1) Strokehealth as a supporting tool for preparing dialogue and identifying needs, (2) how Strokehealth is introduced and communicated affects perceived usability, and (3) the wording and structure of Strokehealth influences the response process (Figure 2).

XSL•FO
RenderX

Figure 2. Themes and subthemes identified from interviews regarding the digital previsit tool Strokehealth.



Strokehealth as a Supporting Tool for Preparing Dialogue and Identifying Needs

This main theme represents aspects of how effectively Strokehealth captured the participants' health issues and how it was perceived in preparing patients for dialogue with health care professionals as well as family members.

Completing Strokehealth Promotes a Sense of Connection, Gratitude, and Being Cared for

The use of Strokehealth as preparation for a visit was perceived to contribute to the sense of being well cared for by health care. Some participants felt reassured by encountering familiar information within the tool, as this validated their existing knowledge and aligned with previous information received from health care providers. This recognition created a sense of connection and reassurance that they were not alone, and that experiencing such issues after a stroke was both expected and normal. Others expressed that the information provided within Strokehealth helped to fill knowledge gaps, supplemented their existing understanding, and brought a sense of hopefulness.

> Participant 2: It's really nice, what I'm looking at right now [advisory text]; it looks really good. It's short, concise, pleasant, and gives hope [laughs]. No, but I think so. I don't really have anything to object to. You can't do more than this, really. It looks nice, good, and informative. [Male, 61 years old]

Completing Strokehealth also evoked feelings of gratitude for being taken seriously by the health care system and for receiving good care. It provided participants with an opportunity to actively reflect on what it meant to have survived a stroke. They felt grateful for not being more severely affected by the stroke and for the confirmation that it could have been much worse, as they read about different problem areas—one question after another.

XSL•FO
RenderX

Challenges in Capturing Health Issues at Varying Levels of Severity

Most participants felt that the questions in Strokehealth were relevant and adequately addressed their health issues, including mental fatigue. However, individuals experiencing very mild symptoms perceived that the questions or the given examples did not fully capture their own experience. This prompted reflections on whether Strokehealth might be better suited for individuals with more severe strokes. One participant experiencing mild aphasia found that none of the questions adequately addressed this specific problem.

Participant 15: It also depends on what level I'm answering from. If I'm answering from an everyday perspective—can I go to the store, talk to my family and friends, participate in Teams meetings, or write messages, emails, or texts—then I have no problems. But if I want to do it at a higher level, workwise, maybe speak another language, and so on, then that's a completely different level, and perhaps a completely different answer to the question. [Male, 43 years old]

Another participant mentioned that none of the questions captured the challenges of everyday tasks that were previously effortless, such as emptying the dishwasher. Now, she had to consider each step while maintaining balance, which made the task exhausting. Participants suggested that more generous open-response areas would allow for detailed explanations of such issues. Participants also found that the questions in Strokehealth were insufficient to capture psychological factors, such as worrying thoughts about the future. They noted that the question related to this topic did not precisely reflect what participants wanted to express.

Perceived Value of Strokehealth for Preparing Health Dialogues

Participants found that completing Strokehealth not only helped them prepare for their follow-up visit but also provided them with a valuable opportunity to reflect on their situation and thoroughly consider their individual needs. They also felt that the tool helped make them aware of certain issues that might not have otherwise been addressed, such as their relationship with loved ones. Some participants even expressed that Strokehealth was valuable for explaining their problems to family members and highlighting common poststroke issues. They suggested that a previsit tool like Strokehealth could be beneficial for structuring and optimizing any health care visit that is often time-limited.

Participant 4: It makes things easier; it becomes like a preparation beforehand. I guess that's the case with all meetings—the more prepared you are, the better it is. If you know what it will be about and can reflect on it beforehand, it's always beneficial. So, I think that's a good thing. [Male, 58 years old]

Participants also reflected on the potential value to the staff if patients filled out Strokehealth, speculating that it might make the meeting more efficient. However, not everyone believed that Strokehealth was necessary; some felt confident about directly articulating their concerns and believed that the nurse was already informed about their problems.

Furthermore, Strokehealth also introduced new vocabulary that inspired them to seek further information in advance—for example, through search engines such as Google—to learn more about their condition and possible interventions. Consequently, this process of independently seeking information further enhanced their preparedness for the upcoming health care encounter.

How Strokehealth Is Introduced and Communicated Affects Perceived Usability

This main theme emphasizes that the way Strokehealth is introduced and communicated to patients plays a crucial role in how the tool is perceived and used. The analysis underscores the significance of offering patients various options for appointment preparation and enabling them to make active choices.

Patients' Ambivalent Reactions to an Initial Digital Contact With Health Care

Participants described different reactions upon receiving the digital notification, that is, a text message on their smartphone indicating that there is a form to complete (Strokehealth). Many participants who were already familiar with the patient portal became curious and felt confident that the short text message was sent from a trustworthy source. However, the meaning of the text message was not always clear to participants who had never previously received a notification from the patient portal. Some participants felt a sense of suspicion, thinking that an authority wanted to poke around in their private life or had concerns about the risk of fraud, as often reported in the media. In some cases, the sender (the nurse from the stroke team) would follow up with a telephone call. These verbal reminders increased patients' motivation to log in and complete Strokehealth. However, some participants mentioned that they would have preferred the initial contact after their stroke to be with a real person rather than an anonymous text message prompting them to answer questions using a digital tool, while others were fully content with the "digital-first" approach.

Participant 19: Well, I'm a bit critical and thinking like this—I was at the stroke unit about 14 days ago, and this is the first contact, about ten days later, that I receive from the healthcare side, and I find that a bit contradictory. On the one hand, maybe one would have wanted to talk to a person and not fill out a form because, after all, it's a transformative experience to have a stroke. [Male, 63 years old]

Participant 30: I felt that it's really good. I think all follow-up feels reassuring. It's very intense when it happens, and then the healthcare system kicks in, and suddenly, you're back home. I think all follow-up is good. And digital tools are totally fine, especially in combination with actually talking to someone, like in this case [Strokehealth combined with a dialogue with a nurse]. So, no, I'm in favor. [Male, 45 years old]

Several participants expressed that they would have preferred to be informed about Strokehealth during their hospital stay rather than only through a digital notification. However, they also acknowledged that this information might have been overlooked due to the many other concerns during their hospitalization.

XSL•FO
RenderX

Timing and Method of Delivery Shape Engagement With Strokehealth

Participants had various thoughts on how to personalize the process of when and how to complete Strokehealth. For instance, they reflected on the time span between hospital discharge and receipt of Strokehealth. It was generally considered that one needed some time to settle back home after discharge before the follow-up visit was scheduled and Strokehealth was delivered.

Participants also reflected on the timing of receiving the text message with a prompt to open the tool. Most considered the forenoon to be appropriate because they still had some of their often-limited energy left. If the message was sent during the night, the signal was likely to wake the recipient. Participants also emphasized that they wanted the flexibility to choose the most convenient time of day to engage with the tool. Approximately 1 week of preparation time was generally considered sufficient to allow participants to select the best time of day to engage with Strokehealth.

Participant 11: And then you could complete it at your own pace as well, you didn't have to rush through the questions, so it's quite nice, so yeah, I think it was a good thing, a really good thing. [Female, 53 years old]

Participants generally appreciated that they could log in and complete Strokehealth using a device of their own choice. The text was considered sufficiently visible on any device—including computers, tablets, and smartphones. However, the prompt to complete Strokehealth digitally caused stress for participants with less experience using technology. Additionally, the web-based system of the patient portal was perceived as difficult to navigate, adding to the digital stress. Some participants also experienced stress related to clicking the prompt to open Strokehealth and expressed concern about clicking the wrong button. However, once participants successfully managed to navigate the patient portal, most considered it easy to complete Strokehealth. For participants with severe technological issues, it was considered advantageous to have someone assist them. Additionally, some participants supported the idea of being able to choose between the digital previsit tool and a paper version on which they could make notes with a pen and then refer to during the follow-up visit.

Clarity and Purpose in the Introductory Text of Strokehealth Supports Usability

The participants' perception of the initial introductory text had an impact on their willingness to fill out Strokehealth. Most participants appreciated that the text was short and clear. Based on the contents of the introductory text, all participants clearly understood the purpose of answering the questions before their follow-up visit. Participants appreciated that the introductory text explicitly stated that there were only 14 questions, estimated to take a reasonable amount of time. Specific information that participants found helpful included that the answering process could be initiated at any preferred time and place and that the upcoming follow-up visit would provide an opportunity to elaborate on any of the indicated symptoms (ie, "yes" responses in Strokehealth).

Participant 25: Well, if it [the introductory text] says that your nurse will ask you about answers, so that you can explain how you have thought, then it makes you answer [the question]. Well, then I answer "yes" when it comes to swallowing because then I know that I can explain that later. And pain, I have that, but I can explain how much pain it is. [Female, 73 years old]

However, participants suggested providing the information that one could take a break at any point and making it clear that they could revisit and review their answers before submitting Strokehealth. On the other hand, participants expressed concerns that adding too much new information to the introductory text could impact readability.

Despite previous adjustments extending the introductory text, several participants did not notice the link to the advisory text at the end of the tool (or forgot it while answering the questions), suggesting a need for further improvements in future versions. Others acknowledged seeing the link to the advisory text but opted not to read it, either because they no longer had symptoms or because they believed that they had already received sufficient information.

The Wording and Structure of Strokehealth Influences the Response Process

This theme highlights how the linguistic formulation and structural design of Strokehealth affect users' understanding, interpretation, and response behavior. The response process was also perceived as dependent on individual prerequisites, highlighting the importance of how information is presented, the clarity of instructions, and the perceived user-friendliness of the interface.

Explanatory Texts Shape Interpretations of Questions

Most participants saw no disadvantages to the explanatory texts, and some even found them crucial for interpreting the questions and understanding their meaning. The provided contextual information—including various examples related to specific health issues—prompted participants to make new associations related to their own impairments. These reflections helped participants provide more accurate and informed answers to the questions. On the other hand, one participant expressed dissatisfaction with the comprehensive list of possible ailments in the explanatory text. He was concerned that selecting "yes" would inaccurately suggest that he experienced all listed ailments. Another participant noted that although some questions, for example, "Do you have more difficulty walking or moving safely after a stroke?" could include minor symptoms, such as slightly impaired balance, these symptoms were not mentioned within the explanatory text. He therefore found this question difficult to respond to.

Participants expressed overall satisfaction with the explanatory texts, finding them easy to comprehend. The texts were deemed concise and straightforward, and not disruptive when answering the questions. Participants emphasized the importance of maintaining simplicity and noted that excessive length could reduce response rates. However, participants identified certain problem areas that could benefit from more detailed explanations. For instance, participants found the question about pain difficult due to the complexity of the condition.

XSL•FO
RenderX

Participant 4: Perhaps the definition of pain is a bit subjective too. I don't have severe pain; it's more like a pressure sensation. But of course, that could also be considered pain, depending on how one defines pain—whether it's sharp pain, severe pain, discomfort, or something like that. I'm not sure if I would have answered 'pain' even though I had this slight discomfort. It didn't hurt; it's more of a discomfort if anything, especially if it's a milder sensation that doesn't feel like direct pain. [Male, 58 years old]

Not everyone felt a need to read the explanatory texts. Some participants chose to read the attached text only if the question sparked their interest in the topic, was unclear, or prompted deeper reflection on their remaining symptoms.

Participant 22: In those cases, my answer was obvious, I didn't read all of them thoroughly. I didn't go into the details or the fine print—I just moved on to the next one. I didn't read everything all the time; if the question was crystal clear to me, I just went to the next one without any comments. Yes, there were certain questions—not all of them—but some where I thought, "What the hell do they mean by that?" And of course, in those cases, I had to read more carefully to understand what it actually said. [Male, 65 years old]

Individual Prerequisites Influence the Response Process

During the interview, several participants reported experiencing communication or memory problems, including signs of aphasia. For many, these impairments impacted their ability to recall details, understand, and express their thoughts effectively within Strokehealth. Some participants mentioned that having a supportive spouse was advantageous when completing Strokehealth. Others mentioned that they answered the questions quickly, either due to their natural tendency to complete questionnaires quickly or because they were at work when the notification arrived. Responding rapidly may have reduced the time available for reflection on the questions; however, participants did not spontaneously raise this concern during the interviews.

Participant 6: It could be that I took a break from work when I did it, and maybe I felt like I didn't really have time to continue or look into it further, so I felt done. That's often how it feels—when I've completed a survey, I just want to get on with my life [laughs]. [Male, 46 years old]

Participants commonly mentioned the problem of fluctuating mental fatigue since their stroke. When participants experienced mental fatigue during their completion of Strokehealth, it influenced every answer. Participants felt that it was misleading to answer "yes" to questions about difficulties with activities of daily living if they were typically able to manage but were currently hindered by fatigue. They suggested including a specific question about fluctuating mental fatigue or adding a free-text area where participants could indicate "sometimes" instead of being limited to the binary "yes" or "no" responses. Another suggestion was to move the question about "life after stroke" from the bottom to the top of the tool, since that question was related to fatigue and hence impacted their responses to the subsequent questions. Notably, the presence of mental fatigue significantly affected participants' ability to focus during the response process. One participant described feeling mentally sluggish, as if their thoughts were moving "as slow as toffee." Accordingly, participants highlighted the need for flexibility to take breaks, without fear of losing previous responses.

Importance of Optimizing User-Friendliness and Allowing for Simplicity

Participants provided valuable insights regarding how to improve the tool's design to enhance the response process. Most participants found the structure, layout, and language of Strokehealth to be attractive, educational, and easy to comprehend. The number of questions, response options, wording, and sequencing of questions were all generally deemed user-friendly and straightforward. No language was considered offensive. In particular, participants with mild poststroke symptoms found the tool easy to navigate.

Regarding the response options, participants speculated that binary choices ("yes" or "no") might be insufficient, suggesting a need for more nuanced options.

Participant 6: I mean... I answered "yes" for those where I had even the slightest symptom. I did, yes. That's how I interpreted it: if I experienced even the slightest symptom, I would answer "yes," and if I didn't, I would answer "no."

Interviewer: So it wasn't hard for you to understand that this is how you should answer, at least?

Participant: No, no. But the feeling could still be that this "yes" might not convey the full truth. [Male, 46 years old]

Additionally, they emphasized the importance of providing a free-text area below each question to allow for personalized responses and improve user-friendliness. Participants felt that a single free-text area at the end of the tool was insufficient because they often wanted to specify directly after each question. The need for a free-text area was particularly evident after the question "Life after stroke," which was generally considered quite broad, with the ambition of capturing many aspects of importance.

Participant 2: I just think I would have liked to answer "yes," but then define it. I would want to specify right under that question... I would like to specify them a bit more, either by breaking that question into several or by allowing a free-text option right after, where you can specify what you mean. [Male, 61 years old]

Not all participants shared this perspective. Some found that the combination of the question itself and a single free-text area at the end was sufficient—especially if they found it challenging to write due to linguistic or motor impairments. Participants also reflected on the answer option "choose not to answer," which some viewed as somewhat peculiar, while others saw it as a "way out." No specific suggestions were made for alternative wording to enhance clarity, but some reflected that this response option could be supplemented with a free-text area to allow individuals to explain why they chose not to answer.

Participants appreciated that a summary appeared when they completed the response process before submitting Strokehealth.

XSL•FO
RenderX

Some suggested that the "submit" button should be made more prominent and visible. After submitting Strokehealth, it disappeared from the screen, which was seen as a disadvantage if someone wanted to review their answers before their follow-up visit. Similarly, it was suggested that it should be made easier to print the summary of their responses.

Development and Amendments to Version 2.0

Strokehealth was amended and altered based on the findings from the individual interviews and the theoretical framework—for example, aspects of person-centered care and health literacy (Multimedia Appendix 4). Several important changes were made in version 2.0, including a revision of the introductory text to better align with expressed preferences. Additionally, the answer options were extended to include a free-text box below each question and additional space in the general free-text box at the end of the tool. Moreover, the explanatory and advisory texts were improved to exhibit greater consistency in structure and content. These included confirming experiences, suggestions of what one can do by oneself (self-management), suggestions of potential interventions, and a prompt to seek support, if needed. Notably, the platform for the national portal has some limitations regarding layout options, predefined typography, and colors that cannot be changed.

Discussion

Principal Findings and Comparison With Prior Work

User Experience and Perceived Usefulness

The Strokehealth digital previsit tool, when used before a scheduled follow-up visit after stroke, was perceived as a supportive aid for facilitating dialogue and identifying patient needs. It was also considered to be useful and user-friendly. Themes identified from the interviews emphasized the importance of patient comprehension and motivation to initiate the response process. The real-world experiences captured in this qualitative study, combined with the findings from the initial patient satisfaction survey (part 1 [17]), provided the foundation for further improvements to the tool. While the initial survey provided useful data on general user satisfaction and perceived usability, the subsequent qualitative evaluation offered deeper insight into patients' experiences and contextual factors affecting their interaction with the tool. It captured emotional responses such as feeling acknowledged and supported, practical barriers related to digital literacy, and the need for more nuanced response options. These findings informed concrete improvements to the tool and highlighted the importance of considering patients' individual circumstances when implementing digital previsit tools in routine care. Based on user feedback, minor adjustments were recommended in version 2.0 to better support the follow-up process from initiation to identifying patient needs. Additionally, participants suggested greater flexibility in the timing and method of introducing Strokehealth to patients to enhance its accessibility and usability.

Supporting Patient Engagement and Self-Management

The current findings indicate that the information presented in Strokehealth fosters increased patient engagement and potentially supports meaningful dialogue with both health care professionals and family members. The stroke-specific information addresses common knowledge gaps that are often experienced by people with stroke and their families [4,23], especially after first-time strokes [24]. Participants reported that the information reassured them that their experiences were normal and, in some cases, motivated them to seek additional resources on the web or take health-related actions. Patients require relevant, concise, and easily accessible information to be able to make informed decisions and support their self-management [4,23,25]. The concise information within Strokehealth aims to foster shared responsibility for health—aligning with the broader health care shift toward self-management [26]; however, simply providing information is insufficient to elicit meaningful behavior changes without active patient engagement [25] or effective communication with health care providers [4]. Patient health engagement can be further encouraged by combining Strokehealth with a follow-up visit (whether on-demand or in-person) during which patients can revisit information and discuss it with health care professionals.

Self-Reflection and Response Design

Participants in this study emphasized the importance of self-reflection during the response process, highlighting various aspects of its significance. First, the response process encouraged participants to consider issues they might not have otherwise reflected upon. Somewhat surprisingly, participants also expressed that the response process fostered a sense of gratitude about not being more severely affected, as well as feeling cared for by the health care system, rather than abandoned—a sentiment that aligns with previous research findings [6]. Second, the self-reflections and responses were influenced by how participants interpreted the explanatory texts attached to each question. Most participants found these texts to be sufficient; however, some expressed that they did not fully address subtle impairments. Third, self-reflection was influenced by the design of the response options. Participants expressed their desire to provide more nuanced responses than were allowed by the binary response options, which reinforced findings from previous evaluations of the co-design process [14]. While this limitation was acknowledged, participants with firsthand experience of a stroke indicated that the response options were generally effective when complemented by a free-text area with sufficient space to record their reflections. Following the evaluation, restructuring the questions to include additional response options aimed at capturing milder symptoms was considered. However, overly complex or ambiguous questions could hinder engagement. Therefore, this approach was ultimately deemed to introduce unnecessary complexity, potentially compromising the goal of developing a user-friendly tool. Based on these insights, Strokehealth version 2.0 was modified to incorporate free-text areas below each question and revised text to better capture the nuances of patients' experiences and support the response process.

Accessibility and Digital Literacy

The present results suggest that Strokehealth can be accessible to a broad range of patients with stroke if choices are offered regarding when and how the tool is introduced. Some

XSL•FO
RenderX

participants described experiences of "digital stress" when navigating the patient portal, although 48 out of 58 participants (82%) used the web on a daily basis. This highlights the need for Strokehealth to be accessible to individuals with limited experience using digital tools. It is recommended that health care providers consider patients' eHealth literacy levels [27] and implement alternative communication methods as needed. For patients with limited digital literacy—including low digital skills, low confidence, or reluctance to use digital tools—health care providers could encourage seeking support from friends or family to alleviate digital anxiety. Our results indicate that some participants successfully completed the tool with support from spouses. Additionally, alternatives to the digital version should be readily available to help build patients' confidence in their ability to self-manage. Therefore, Strokehealth is also offered as a paper version [28], which can be sent by a health care professional to a patient who chooses not to interact with the patient portal after receiving the initial text message. Such efforts can improve accessibility and organizational health literacy, ultimately supporting patients in more effectively navigating the health care system [29].

Implementation Considerations

Importantly, evaluations of a digital health tool such as Strokehealth must consider not only the functionality of the tool itself but also the contextual factors [16] related to the follow-up process. Effective introduction is crucial for engaging patients in the proactive management of their health [30], including clear communication of the purpose of a notification and building trust in the sender [31]. In this study, the timing and method of delivery played a key role in engagement with the tool, with some participants initially questioning the legitimacy of the digital notification. Mentioning Strokehealth upon hospital discharge could raise awareness about the tool, enhance trust, and help patients perceive the follow-up process as seamless and integrated. Ideally, patients could be informed at the time of hospital discharge that they will later receive a digital previsit tool, once they have spent some time at home. This information should include the purpose of Strokehealth, and if they choose to opt out, a paper version could be offered at that later stage. Engagement was further influenced by factors such as fatigue and cognitive impairment, underscoring the need for well-timed notifications, ideally in the morning when patients' energy levels are higher. Strokehealth version 2.0 addresses these challenges by promoting flexibility, allowing patients to complete the tool at their convenience and to take control of their follow-up process.

Implications for Digitally Supported, Person-Centered Care

Our findings align with previous calls for digitally supported, person-centered care in long-term conditions such as stroke [1,2]. By providing insights into how patients experience the use of a digital previsit tool, this study contributes to the growing evidence base on how eHealth interventions can promote reflection, engagement, and preparedness ahead of follow-up visits.

Strengths and Limitations

This study involved patients who were consecutively recruited by the selected health care professionals; therefore, bias due to selection cannot be excluded. However, this sample—the majority of whom had a mild stroke—is representative of the Swedish population. The comprehensive description of the context and participants in this study, complemented by part 1 [17], supports the transferability of the findings to other organizations. In accordance with the co-design method, the researchers were involved in all parts of the design process [32,33], including the evaluation elements. However, the researchers' involvement could have made participants feel hesitant to express drawbacks, despite encouragement to openly share thoughts. Some of the interviews were relatively short in duration, which may have limited the depth of participants' responses. It is possible that conducting the interviews face-to-face, rather than by telephone, could have facilitated richer and more nuanced accounts by allowing for nonverbal communication and a more personal connection between the interviewer and the participant. Nevertheless, the critical feedback and rich data in this study demonstrate participants' willingness to share their experiences and contribute to improving the tool. To enhance reflexivity, reflections were systematically noted. Importantly, the researchers' preunderstanding is considered a factor that influences the data analysis and interpretation [34]. To ensure confirmability, the impact of this preunderstanding was considered throughout the analysis, together with continuous return to the coded data and transcripts in NVivo. Moreover, credibility was strengthened by conducting the interviews in a familiar environment, using a telephone interview. However, the use of the telephone could have influenced participant recruitment. Interviews were conducted by telephone due to the COVID-19 pandemic; however, this was found to be a surprisingly well-functioning method of collecting data in this sample.

The Strokehealth digital tool has now been tested within a clinical care pathway to assess whether patient needs and preferences are being appropriately addressed in accordance with prior recommendations [16]. Accordingly, Strokehealth has been implemented within the Swedish health care system, allowing health care professionals to access patient responses via the national patient portal prior to clinical visits. However, the broader implications and potential benefits from a health care system perspective warrant further exploration. Following guidelines for evaluating complex interventions, the next step would be to gather the perspectives of health care professionals to understand how Strokehealth could support them in the delivery of care. Together, our studies (parts 1 and 2) indicate that the tool's overarching purpose—to encourage reflection and preparation for a planned health care visit—remains broadly applicable among patients with stroke.

Conclusions

The digital previsit tool Strokehealth was generally well received, with participants describing it as supportive, easy to use, and valuable for fostering reflection and preparing for meaningful dialogue. Patients' feedback led to concrete refinements in Strokehealth version 2.0, including improved

XSL•FO
RenderX

explanatory texts, added flexibility in the timing and format of delivery, and expanded opportunities for free-text input. Importantly, the findings highlight that digital tools must account for individual circumstances, including eHealth literacy and cognitive or emotional challenges, in order to ensure equitable access. This study contributes to the evidence base for implementing person-centered, digitally supported follow-up care in stroke rehabilitation.

Acknowledgments

The authors sincerely appreciate the participation of individuals affected by stroke and their willingness to share their experiences. The authors also extend their gratitude to the health care professionals for managing the previsit tool alongside their regular duties. Furthermore, the authors would like to thank the members of the advisory board, the linguist, the web designers, and the patient partner and experts who participated in the decision workshops. This study was funded in part by grants from the Swedish state under an agreement among the Swedish government and the county councils, the ALF agreement (ALFGBG-965653); the Local Research and Development Board for Gothenburg and Södra Bohuslän (VGFOUGSB-76712); Sahlgrenska Hospitals Foundation (SU-892141 and SU-793481); the Swedish Stroke Association; the Swedish Neuro Association; an unconditional grant from Allergan; the Hjalmar Svenssons Foundation; Göteborg Foundation for Neurological Research; Rune and Ulla Amlövs Foundation; Per-Olof Ahls Foundation; John and Brit Wennerströms Foundation; the Innovation foundation at region Västra Götaland (VGRINN-968305); and the Greta and Einar Askers Foundation.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Strokehealths' 14 questions.
[DOCX File , 38 KB - humanfactors_v12i1e73477_app1.docx]

Multimedia Appendix 2
COREQ checklist.
[PDF File (Adobe PDF File), 423 KB - humanfactors_v12i1e73477_app2.pdf]

Multimedia Appendix 3
Semistructured interview guide.
[DOCX File , 16 KB - humanfactors_v12i1e73477_app3.docx]

Multimedia Appendix 4
List of amendments to Strokehealth version 2.0.
[DOCX File , 23 KB - humanfactors_v12i1e73477_app4.docx]

References

1. Leonardsen AL, Bååth C, Helgesen AK, Grøndahl VA, Hardeland C. Person-Centeredness in Digital Primary Healthcare Services-A Scoping Review. Healthcare (Basel) 2023 May 01;11(9):1296 [FREE Full text] [doi: 10.3390/healthcare11091296] [Medline: 37174838]
2. Rosenlund M, Kinnunen U, Saranto K. The use of digital health services among patients and citizens living at home: scoping review. J Med Internet Res 2023;25:e44711 [FREE Full text] [doi: 10.2196/44711] [Medline: 36972122]
3. GBD 2019 Stroke Collaborators. Global, regional, and national burden of stroke and its risk factors, 1990-2019: a systematic analysis for the global burden of disease study 2019. Lancet Neurol 2021;20(10):795-820 [FREE Full text] [doi: 10.1016/S1474-4422(21)00252-0] [Medline: 34487721]
4. Pearce G, Pinnock H, Epiphaniou E, Parke HL, Heavey E, Griffiths CJ, et al. Experiences of self-management support following a stroke: a meta-review of qualitative systematic reviews. PLoS One 2015;10(12):e0141803 [FREE Full text] [doi: 10.1371/journal.pone.0141803] [Medline: 26657458]
5. Guo Y, Zhang Z, Lin B, Mei Y, Liu Q, Zhang L, et al. The unmet needs of community-dwelling stroke survivors: a systematic review of qualitative studies. Int J Environ Res Public Health 2021;18(4):2140. [doi: 10.3390/ijerph18042140] [Medline: 33671734]
6. Spiers G, Boulton E, Corner L, Craig D, Parker S, Todd C, et al. What matters to people with multiple long-term conditions and their carers? Postgrad Med J 2023;99(1169):159-165. [doi: 10.1136/postgradmedj-2021-140825] [Medline: 37222056]
7. Barello S, Triberti S, Graffigna G, Libreri C, Serino S, Hibbard J, et al. eHealth for patient engagement: a systematic review. Front Psychol 2015;6:2013 [FREE Full text] [doi: 10.3389/fpsyg.2015.02013] [Medline: 26779108]

XSL•FO
RenderX

8. Neter E, Brainin E. eHealth literacy: extending the digital divide to the realm of health information. J Med Internet Res 2012;14(1):e19 [FREE Full text] [doi: 10.2196/jmir.1619] [Medline: 22357448]
9. Norrving B, Barrick J, Davalos A, Dichgans M, Cordonnier C, Guekht A, et al. Action plan for stroke in Europe 2018-2030. Eur Stroke J 2018;3(4):309-336 [FREE Full text] [doi: 10.1177/2396987318808719] [Medline: 31236480]
10. Philp I, Brainin M, Walker MF, Ward AB, Gillard P, Shields AL, Global Stroke Community Advisory Panel. Development of a poststroke checklist to standardize follow-up care for stroke survivors. J Stroke Cerebrovasc Dis 2013;22(7):e173-e180 [FREE Full text] [doi: 10.1016/j.jstrokecerebrovasdis.2012.10.016] [Medline: 23265778]
11. Kjörk EK, Carlsson G, Sunnerhagen KS, Lundgren-Nilsson. Experiences using the poststroke checklist in Sweden with a focus on feasibility and relevance: a mixed-method design. BMJ Open 2019;9(5):e028218 [FREE Full text] [doi: 10.1136/bmjopen-2018-028218] [Medline: 31072862]
12. Stacey D, Lewis KB, Smith M, Carley M, Volk R, Douglas EE, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2024;1(1):CD001431. [doi: 10.1002/14651858.CD001431.pub6] [Medline: 38284415]
13. Holt JM, Cusatis R, Winn A, Asan O, Spanbauer C, Williams JS, et al. Impact of pre-visit contextual data collection on patient-physician communication and patient activation: a randomized trial. J Gen Intern Med 2021;36(11):3321-3329 [FREE Full text] [doi: 10.1007/s11606-020-06583-7] [Medline: 33559067]
14. Kjörk EK, Sunnerhagen KS, Lundgren-Nilsson, Andersson AK, Carlsson G. Development of a digital tool for people with a long-term condition using stroke as a case example: participatory design approach. JMIR Hum Factors 2022;9(2):e35478 [FREE Full text] [doi: 10.2196/35478] [Medline: 35657650]
15. Hu X, Jonzén K, Lindahl OA, Karlsson M, Norström F, Lundström E, et al. Digital graphic follow-up tool (Rehabkompassen) for identifying rehabilitation needs among people after stroke: randomized clinical feasibility study. JMIR Hum Factors 2022;9(3):e38704 [FREE Full text] [doi: 10.2196/38704] [Medline: 35904867]
16. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: update of medical research council guidance. BMJ 2021;374:n2061 [FREE Full text] [doi: 10.1136/bmj.n2061] [Medline: 34593508]
17. Pohl P, Klerfors K, Kjörk EK. Evaluation of a digital previsit tool for identifying stroke-related health problems before a follow-up visit (Part 1): survey study. JMIR Hum Factors 2024;11:e55852. [doi: 10.2196/55852] [Medline: 39226546]
18. Lincoln Y, Guba E. Naturalistic inquiry. In: The Blackwell Encyclopedia of Sociology. Beverly Hills, CA: Sage; 1985.
19. Pawson R, Tilley N. An introduction to scientific realist evaluation. In: Chelimsky E, Shadish WR, editors. Evaluation for the 21st Century: A Handbook. Thousand Oaks, CA: Sage Publications, Inc; 1999:405-418.
20. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
21. Attribution 4.0 International (CC BY 4.0). Creative Commons. URL: http://creativecommons.org/licenses/by/4.0/ [accessed 2025-09-04]
22. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2008;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
23. Helbach J, Hoffmann F, Hecht N, Heesen C, Thomalla G, Wilfling D, et al. Information needs of people who have suffered a stroke or TIA and their preferred approaches of receiving health information: a scoping review. Eur Stroke J 2025;10(1):5-21 [FREE Full text] [doi: 10.1177/23969873241272744] [Medline: 39188106]
24. Chen L, Xiao LD, De Bellis A. First-time stroke survivors and caregivers' perceptions of being engaged in rehabilitation. J Adv Nurs 2016;72(1):73-84. [doi: 10.1111/jan.12819] [Medline: 26399942]
25. Hibbard JH. Patient activation and the use of information to support informed health decisions. Patient Educ Couns 2017;100(1):5-7. [doi: 10.1016/j.pec.2016.07.006] [Medline: 27432014]
26. Strategi för genomförande av e-hälsa vision 2025 [Strategies for the implementation of the eHealth vision 2025]. e-Hälsa. 2021 Mar 21. URL: https://ehalsa2025.se [accessed 2025-02-13]
27. Verma A, Towfighi A, Brown A, Abhat A, Casillas A. Moving towards equity with digital health innovations for stroke care. Stroke 2022;53(3):689-697 [FREE Full text] [doi: 10.1161/STROKEAHA.121.035307] [Medline: 35124973]
28. Strokehälsa. University of Gothenburg. 2023. URL: https://www.gu.se/en/neuroscience-physiology/strokehalsatm [accessed 2025-02-13]
29. Wang X, Luan W. Research progress on digital health literacy of older adults: a scoping review. Front Public Health 2022;10:906089 [FREE Full text] [doi: 10.3389/fpubh.2022.906089] [Medline: 35991040]
30. Spreadbury JH, Young A, Kipps CM. A comprehensive literature search of digital health technology use in neurological conditions: review of digital tools to promote self-management and support. J Med Internet Res 2022;24(7):e31929 [FREE Full text] [doi: 10.2196/31929] [Medline: 35900822]
31. Norgaard O, Furstrand D, Klokker L, Karnoe A, Batterham R, Kayser L. The e-health literacy framework: a conceptual framework for characterizing e-health users and their interaction with e-health systems. Knowledge Manage E-Learn 2015;7(4):522-540. [doi: 10.34105/j.kmel.2015.07.035]
32. Sanders EB, Stappers PJ. Co-creation and the new landscapes of design. CoDesign 2008;4(1):5-18. [doi: 10.1080/15710880701875068]

XSL•FO
RenderX

33. Cowdell F, Dyson J, Sykes M, Dam R, Pendleton R. How and how well have older people been engaged in healthcare intervention design, development or delivery using co-methodologies: a scoping review with narrative summary. Health Soc Care Community 2022;30(2):776-798. [doi: 10.1111/hsc.13199] [Medline: 33103313]
34. Malterud K. Qualitative research: standards, challenges, and guidelines. Lancet 2001;358(9280):483-488. [doi: 10.1016/S0140-6736(01)05627-6] [Medline: 11513933]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research

Edited by A Kushniruk; submitted 05.03.25; peer-reviewed by AM Lunde Huseboe, N Khan, S Ivynian; comments to author 27.05.25; revised version received 15.07.25; accepted 28.08.25; published 06.10.25.

Please cite as:
Pohl P, Kjörk EK
Patients' Perceptions of Using a Digital Previsit Tool in Outpatient Settings (Part 2): Qualitative Study
JMIR Hum Factors 2025;12:e73477
URL: https://humanfactors.jmir.org/2025/1/e73477
doi:10.2196/73477
PMID:

©Petra Pohl, Emma K Kjörk. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 06.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Pursuit of Digital Innovation in Psychiatric Data Handling Practices in Ireland: Comprehensive Case Study

Rana Zeeshan[1]; John Bogue[2]; Amna Gill[3]; Mamoona Naveed Asghar[1]

[1]School of Computer Science, University of Galway, University Road, New Castle, Galway, Ireland

[2]School of Psychology, University of Galway, Galway, Ireland

[3]Mental Health Services, Ballymun Civic Centre, Dublin North, Dublin, Ireland

Corresponding Author:
Rana Zeeshan
School of Computer Science, University of Galway, University Road, New Castle, Galway, Ireland

Abstract

Background: Ireland is ranked among the most disadvantageous European countries in terms of mental health challenges. Contrary to general health services that primarily focus on diagnosis and treatment, the mental health sector in Ireland deals with highly sensitive psychiatric case notes based on patient-doctor conversations. Such data, therefore, must be collected, analyzed, and stored with an approach customized specifically for psychiatry.

Objective: This study's objective involves examining the state of data handling practices in the Irish Mental Health Services (MHS), identifying the shortcomings regarding privacy, security, and usability of psychiatric case notes, and proposing an innovative technological solution that addresses most of the surfaced challenges.

Methods: The study was conducted using a comprehensive methodology. Our approach involved a thorough literature review, ethics approval, web-based surveys with mental health professionals as participants, interviews of psychiatrists, interactions with mental health organizations, analysis of inspection reports by the Ireland Mental Health Commission, and comparative evaluation of existing IT solutions. The thoroughness of our adopted research methodology instills confidence in the reliability and validity of our findings.

Results: Our study revealed outdated data management, heavy reliance on paperwork resulting in serious repercussions, parallel workload, alarmingly low readability of notes, and a nonviable setup that hinders research and analytical examination. Our survey reported an average score of 4.37 of 10 (SD 1.25) given by participants in terms of technology use. Regarding privacy measures, 75% (n=12) of participants mentioned that staff members are allowed to keep their phones while accessing psychiatric case notes. Similarly, 80% (n=13) of submissions highlighted that multiple staff members can access sensitive notes and patients' contact information. On the other hand, Mental Health Commission reports showed that their inspections are limited to evaluating physical privacy only. Regarding technological comparative analysis, we observed that conventional IT solutions are vulnerable against cyberattacks and fall short in addressing multiple challenges simultaneously. Therefore, an innovative convergence of different technologies is needed. Our research supports speech-to-text transcription for data collection, interactive artificial intelligence for data analysis, and permissioned blockchain for data storage and retrieval. Our survey participants also estimated the proposed solution to optimize their workload by an average of 35%.

Conclusions: Irish MHS seem to be handling psychiatric data under polycrisis circumstances; therefore, a single-dimensional digitization of records would not be sufficient in addressing the wide range of concerns. In addition to highlighting intertwined challenges in Irish psychiatry and validating the need for innovation in data handling practices in Irish MHS, this study culminated in the proposal of an innovative technological solution that offers a significant contribution to a considerably improved, efficient, and compliant service delivery in mental health care.

(JMIR Hum Factors 2025;12:e64919) doi:10.2196/64919

KEYWORDS

mental health records; Irish Mental Health Services; mental health care reforms; speech-to-text transcription; blockchain application in health care; AI application in psychiatry; Ireland; psychiatric; mental disorders; behavior disorders; mental illness; mental health professionals; psychiatrists; data management; surveys; data; psychiatric data; case study design; comprehensive case study; digital; innovation; technology use

XSL•FO
RenderX

Introduction

Mental health care is a critical component of public health systems, playing a pivotal role in addressing the challenges of mental disorders and promoting overall well-being. The importance of mental health care extends beyond individual health outcomes, encompassing significant social and economic implications.

According to the third annual Mental State of the World (MSW) report, Ireland ranks among the countries with the lowest mental health score globally. Specifically, Ireland's Mental Health Quotient places it as the second worst in Europe, with the United Kingdom ranking worst, as shown in Table 1. Furthermore, Ireland is notable for having a significant proportion of respondents classified as distressed or struggling, ranging from 30% to 36% (Table 1). These findings were based on data collected from over 400,000 respondents across 64 countries in 2022, providing a comprehensive assessment of the mental well-being of included nations [1].

Table . Relevant statistics from the third annual Mental State of the World report 2022.

Countries	Mental health quotient score	Percentage of distressed population
Belgium	72.1	18.9
France	74.1	19
Spain	70.2	22.6
Germany	65.9	24.5
Ireland	56	30.4
United Kingdom	46.2	35.7

Another study conducted by scholars from Maynooth University, National College of Ireland, and Trinity College Dublin reveals that more than 40% of Irish adults experience a mental health disorder [2]. In terms of economic burden, mental health issues in Ireland are reported to be costing the country over 8.2 billion euros annually (a currency exchange rate of 1 EUR=US $1.1201 is applicable) [3]. As highlighted in the Organisation for Economic Co-operation and Development report from 2018, the cost of mental health problems accounts for 3.2% of Ireland's gross domestic product [4]. These financial implications underscore the urgency of addressing mental health challenges within the country and investing in robust mental health care infrastructure.

Mental health services (MHS) in Ireland are responsible for handling the sensitive and private data of patients, including their medical history, diagnoses, treatment plans, and personal information. When it comes to psychiatric data, the stakes get much higher. There have been several instances of mental health data breaches, abuses, and privacy violations within Ireland and across Europe, highlighting the need for improved security measures. With the growing reliance on digital health records, the risk of data-related threats continues to increase. In 2023 alone, European Union countries reported 309 major cybersecurity incidents in the health care sector, the highest among all critical sectors [5]. We have also listed some relevant incidents of data violations in Table 2.

XSL•FO
RenderX

Table . List of relevant data violation incidents.

Country	Incident	Year
France	Two French health care payment providers, Viamedis and Almerys, experienced data breaches affecting over 33 million people, exposing personal details and insurance information [6].	2024
Romania	A ransomware attack on Romania's health care management system has forced 100 hospitals to take their systems offline [7].	2024
Ireland	The Health Service Executive of Ireland experienced a serious ransomware cyberattack, which led to all its IT systems nationwide to be halted. The attackers demanded a ransom for not publishing private health data [8].	2021
Finland	A hack of psychotherapy records affected thousands of patients. Many of them reported instances of receiving emails with ransom demands paired with warning of disclosing their private conversations with the therapists [9].	2020
Ireland	A doctor was reported for the abuse of his patient's contact information by sending sexually motivated text messages [10].	2018
Ireland	Irish High Court ordered the suspension of a psychiatrist over allegations of sexual misconduct with vulnerable patients based on their mental health records [11].	2013

These incidents underscore the critical necessity of comprehending the deficiencies in data handling practices within Irish MHS and the pressing need to devise an appropriate solution that effectively addresses different concerns. Considering such challenges, it was important to evaluate different technological methods that are commonly adopted for securing psychiatric data and then compare those practices with the ones in Ireland MHS. When it comes to the state of technological advancements in psychiatry, the most common solutions are electronic health records (EHRs) and personal health records (PHRs). While EHRs and PHRs are useful solutions, they come with their own set of security risks [12,13]. Similarly, there is a wide range of technological solutions involving cloud, blockchain, and conventional databases that have been adopted by health care sector all over the world. eHealth or e-therapy is another recent form of counseling that has been arising during the last few years [14,15]. The World Psychiatry Association also promotes digital means in mental health care worldwide. Moreover, it has a dedicated working group on digitalization in mental health [16,17]. Since our research picked Ireland's MHS as a case study, we finalized our research objectives as follows:

- System insights: To gain insight into psychiatric procedures and evaluate how sensitive data are handled within the Irish mental health system.
- Stakeholders engagement: Interact and collaborate with different stakeholders, such as psychiatrists, psychologists, and mental health organizations. Document their concerns and establish a feedback loop.
- Technical proposal: Document the research findings and propose a solution that addresses the surfaced challenges after getting it validated by the stakeholders.

Furthermore, to validate and document the existing state of data handling practices in Irish MHS, this study makes a noteworthy research contribution toward technological advancement in the mental health care sector. Several methods were adopted to confirm the consistency of surfaced shortcomings. Moreover, different stakeholders of Irish MHS were engaged in research participation, public and patient involvement contribution, and critical analysis of proposed solutions. In addition to interacting with public MHS, the scope of this study was extended to relevant third-party organizations as well. Such a multifaceted approach helped us map out an innovative and novel technological solution based on consultations with the most pertinent stakeholders of the loop. This involves the convergence of different technologies, that is, speech-to-text transcription, artificial intelligence (AI), and permissioned blockchain. These technologies were selected for specific purposes with proper justifications, following an in-depth comparative evaluation against alternatives.

The remaining sections of this paper are organized as follows. The Methods section presents the methodology with pertinent details, involving interviews, surveys, third-party interactions, and technological evaluation. The Results section includes results and findings related to interviews, surveys, and descriptive comparative evaluation of different technologies. The Discussion section discusses a proposed solution, along with its applicability and novelty compared to the existing state. It also discusses the foreseen challenges and limitations of the presented work. The Conclusions section contains concluding remarks and future scope.

XSL•FO
RenderX

Methods

Study Design

An extensive mixed methodology was adopted to perform a thorough case study regarding existing data handling practices in Irish MHS. This research intends to document the current practices and propose an IT solution to further secure and improve existing service delivery or address any identified shortcomings. The following steps were taken for preliminary investigations, which led to the new technical proposal for Irish MHS.

Interviews With Psychiatrists

Detailed interviews were conducted with psychiatrists working in MHS in Ireland, that is, Galway, Dublin, Cavan, and Donegal. This helped us develop a basic understanding of the Irish psychiatry sector, its challenges, and any scope for improvement.

Case Notes in Psychiatry

It was brought to our knowledge that a large amount of paperwork is generated each day in the Irish MHS in the form of "case notes." Psychiatry case notes are used by doctors and nurses to plan, implement, and evaluate treatment for mental health patients. They are also used to track progress, exchange information among other health care practitioners, and ensure that care is being delivered safely, ethically, and effectively [18]. Furthermore, these notes are highly useful in the analysis of different cases. Clinicians often study these notes to identify what went wrong at what stage. Therefore, the authenticity and quality of psychiatry notes are of particular significance.

Reliance on Cognitive Abilities

Another aspect brought to our notice was that the existing practice relies on the human ability of a doctor to write down the conversation with a patient without skipping any important information. Such reliance on doctors' cognitive abilities results in a lack of uniformity regarding a patient's case record [19].

Moreover, a large percentage of these handwritten notes have quite a low readability. One of the psychiatrists claimed that 70% of the contents of psychiatric case notes are not readable. This is quite concerning, as this low-readability problem would also hinder the potential of research and analysis of such notes.

Legal Aspects

One interesting application of these psychiatry notes is connected to occasional legal matters. Evaluation of a suspect's mental vulnerability during criminal proceedings is quite common in Europe [20]. Under some circumstances, the solicitors are granted access to the psychiatry notes of a patient by the courts. The solicitors then use the documented conversation and assessments to establish their points whether in favor of proceedings or against it. This is quite a serious and sensitive domain. If we connect this application with the readability and uniformity challenges described earlier, those shortcomings could potentially manipulate or affect the outcome of a legal battle.

We came across a public example quoted by solicitor Sarah Grace in an interview to the RTE Upfront program in 2023 [21], which validated this aspect.

In her words: "In Ireland, in any sexual offence case, the therapy notes, or counseling records of the victims can be used against them to disprove what they are saying in trial."

She also pointed out that the entire defense team of solicitors as well as the accused can read those notes: "After the violation of your body, you also go through the violation of your privacy and dignity."

Technology Adoption

When it comes to the existing state of technology use in the Irish MHS, psychiatrists mentioned that so far, the technology has not actually helped them, instead, it has built a parallel system (paperwork+digital tasks) and increased the workload. As articulated by a psychiatrist:

> *Technology was supposed to REPLACE our workload, not ADD to it.*

A bit of variation was noticed in the extent of technology use among different MHS across Ireland. For example, unlike Cavan MHS, which relied largely on paperwork, the MHS in Letterkenny adopted a tool named T-Pro for the digital dictation of nursing notes, letters, and other documents [22]. Although this tool reduces time consumption in some specific tasks, it still does not replace the overall practice.

On the other hand, the digitization of mental health records should be considered beyond workload optimization, as it offers numerous benefits like enhanced case analysis and improved readability. Thus, workload reduction should be seen as an added advantage rather than a necessary requirement.

Patient-Doctor Equation

Psychiatrists also mentioned that several patients are not fully confident about the privacy of these notes. Some patients get uncomfortable, watching their clinician making notes about them while they are talking. They are, therefore, reluctant to share their thoughts freely, which hinders diagnosis. We also came across a similar observation made by psychiatrists in Switzerland [23], where the patients were worried about being judged or mistreated by medical staff if their report mentions their mental ailment. For some clinicians in the Irish MHS, this often results in maintaining 2 versions of notes, that is, an official version for the institution and a private version to keep track of highly confidential information related to their patients.

Quest for Solution

These insights put emphasis on the need for an automated technological solution to collect such sensitive data. The process should be automated to some extent, as asking doctors to replace "writing" with "typing" will only solve the readability issue. Still, it will not solve all the other challenges discussed earlier, that is, reliance on cognitive ability, workload, uncomfortable patients, authenticity, and uniformity.

We then asked, what if we record such treatment sessions (audio)? Although that could potentially solve many problems, we were informed that recording conversations would increase

XSL•FO
RenderX

the sensitivity of these data by multiple folds. This discussion then brought us to speech-to-text transcription as the most reasonable potential solution, which will be discussed later in this paper.

Community Engagement—Web-Based Survey

Overview

A web-based survey was also conducted to validate the preliminary findings and explore the respective scope of improvement. The target participants for our study were mental health practitioners working at Irish MHS, that is, psychiatrists, psychologists, psychiatric nurses, and senior consultants.

The inclusion criteria were mental health practitioners working at Irish MHS, who are involved in handling psychiatric case notes in any capacity.

A participant gets excluded if he or she is not a mental health practitioner or clinician, he or she has not experienced working in the Irish MHS, and he or she was not involved in handling mental health notes or patient's data.

Survey Sections

The survey comprised 9 sections, with 3 to 5 multiple-choice questions each. The questionnaire did not ask for any confidential information related to any institution. The summarized breakdown of the different sections of the survey is given below:

1. Participant's experience of working at Irish MHS (years).
2. Admission: Queries regarding the process of data handling during a patient's admission.
3. Treatment: Queries regarding the process of data handling during a patient's treatment.
4. Procedures: Data handling during an exchange of information between different departments.
5. Privacy and security: Measures in place to ensure privacy and security of the patient's data.
6. Technology: Extent of technology use and familiarity with modern digital practices.
7. Workload: Queries regarding the workload of patients on practitioners based on existing practices.
8. Assumed cases: Queries regarding participant's opinions about different assumed situations.
9. Suggestions: Submit any useful suggestion that could aid the research objectives.

Third-Party Interactions

Overview

In addition to evaluating the public and mainstream MHS in Ireland, we found it important to involve the insights from relevant third parties, such as government regulatory bodies for psychiatry and private sector. We had three reasons for this outreach: (1) to paint a bigger picture that involves diverse perspectives in the case study, so we may not miss out on something significant; (2) to check and ask around whether the shortcomings surfaced during our research have already been identified or documented by some other stakeholder or not; and (3) to go through the progress made by different players in psychiatry in case the solution we are looking for might already be in practice. And if yes, to what extent?

Brothers of Charity Services Galway

A local mental health organization named Brothers of Charity Services (BoCS) Galway is known among mental health clinicians for having quite advanced data management software. Upon our request for a demonstration, they were considerate enough to schedule a meeting with their IT team. Some important findings and insights are highlighted in the Results section.

Mental Health Commission

This study also involved interacting with Ireland's Mental Health Commission, which is responsible for promoting, encouraging, and fostering high standards and good practices in the delivery of MHS in Ireland [24]. They conduct annual inspections of MHS in Ireland and publish their reports [25]. Therefore, the Mental Health Commission's perspective was an obvious choice to be considered and included in this case study.

Technological Evaluation

Progressive digitization of all conventional processes and systems results in a constant generation of data at quite a rapid rate. All these data are produced, collected, processed, stored, and shared for various purposes. Many data science technologies are being adopted for health-related data handling. However, the areas of data security and privacy are not up to the mark yet [26]. We evaluated different IT solutions for mental health care data management and found that conventional data security practices are still prone to violations of different sorts.

Ethical Considerations

Before carrying out the survey and interviews, ethics approval was obtained from the research ethics committee of the University of Galway (2023.08.024). All participants signed a consent form and were well aware of their right to opt out at any stage. No financial compensation was offered or provided to the participants. The study does not involve confidential or sensitive data collection; instead, it entails validation of routine standard practice. It is important to note that any mental health patient data were not collected, stored, or used in the study.

Results

We have compiled results and findings from different components of the research methodology. It was interesting to notice how one problem was connected to many more.

Interview Findings

Overview

The key points from our interviews with mental health practitioners working in the Irish MHS have been summarized and listed as follows: excessive reliance on paperwork and physical privacy; in addition to record-keeping, psychiatry notes are quite useful in case analysis; psychiatry notes have occasional application in legal matters; handwritten notes have alarmingly low readability and uniformity; there is a dual system

XSL•FO
RenderX

in place, that is, paperwork and digital tasks; patients are often uncomfortable with this note-taking activity; patients are reluctant in sharing information to bypass the notes; patients often ask clinicians to skip some information to avoid stigma; and clinicians often maintain 2 versions of notes, that is, official and private.

Nonstatutory Services Report

Although BoCS's IT solution and dashboard were comprehensive and all-inclusive, it still lacked case note–specific data handling that we sought. However, it was quite an impressive system developed over 2 decades. This is why we picked it as a reference for our investigation regarding data handling practices in the private sector of Irish MHS. Here are some highlights:

- Around 2000 staff members use their IT solution referred to as "OLIS," that is, online information system.
- OLIS is not merely a mental health record management system, it is an extensive enterprise resource planning system with around 18 merged modules that support largely all aspects of their services. They have developed this IT system for over 20 years.
- Regarding security, the BoCS team mentioned that they have not experienced any serious cyberattack yet; however, they reported occasional phishing attacks.
- Their system does not store or manage psychiatry notes digitally. Instead, they have general fields for comments.

It is important to understand that if we remove the psychiatry notes aspect from the equation, mental health records would get the same treatment as general health records. Psychiatry notes are what make mental health records more sensitive and novel as compared to general health records.

Survey Findings

A total of 16 submissions by mental health professionals have been collected and analyzed. A breakdown of the participants' profiles is provided in Table 3.

Overall, the results confirm several concerns brought to our attention during the interviews. In one section, we asked all 16 participants to rank their Mental Health Department's processes out of 10 regarding IT use and modern digital practices. The distribution of responses was as follows: a total of 2 participants rated the processes at 7 of 10, a total of 4 participants gave a score of 5 of 10, a total of 6 participants rated them 4 of 10, and a total of 4 participants assigned a score of 3 of 10. The average score of the submissions is 4.37 of 10 (SD 1.25). This is alarming enough to establish the need for digitization. Moreover, some relevant findings regarding technology use, privacy measures, and workload at Irish MHS are shared in Tables 4-6, respectively.

Table . Survey participants' profile.

Profile	Values, n (%)[a]
Participants' role	
Nonconsultant hospital doctors	6 (40)
Senior psychiatrists	5 (30)
Psychiatric nurses	3 (20)
Health officers	2 (10)
Participants' role	
Working experience <5 years	50
Working experience >5 years	35
Working experience >10 years	15

[a]Percentages are rounded to the nearest tenth (10%).

Table . Survey findings regarding "technology use" at Irish Mental Health Services.

Question	Majority or average response	Values, n (%) [a]
How does a general physician usually refer a patient to your Mental Health Services?	General physician sends a letter by post.	15 (90)
What information is collected at the time of admission? Is that information stored electronically or just in a patient's paper file or folder?	Patient's identity, contact, and residence information are collected and stored both electronically and in paperwork.	12 (75)
When patients describe their private thoughts or mental challenges to a psychologist or psychiatrist, how is that information stored?	Notes on a paper-based chart or form.	15 (90)
To what extent, are the modern technology and digital practices being used by staff members regarding mental health records?	Staff members use technology to the extent of sending emails, faxes, and scans.	10 (65)

[a]Percentages are rounded to the nearest multiple of 5%.

XSL•FO
RenderX

Table . Survey findings regarding "privacy measures" at Irish Mental Health Services.

Question	Majority or average response	Values, n (%)[a]
How many staff members are involved in the process of discussing a patient's case?	5 to 10 including nonmedical support staff.	9 (60)
How often does a patient get assigned to alternative staff members, based on their availability, or shift timing?	50% split response between calling it a "routine practice" and "no such changes are made."	N/A[b]
Which staff members are authorized to access mental health patient files?	Any staff member can access files, if needed.	13 (80)
To what extent are the staff members aware of the privacy laws and how strict are the privacy policies?	All staff members are fully aware of GDPR[c] and relevant privacy laws, with repeated reminders or updates.	15 (90)
Are staff members allowed to keep or use their cell phones while they access patients' records or files?	Yes, staff members are allowed to keep cell phones.	12 (75)
Could assigned staff members access the contact information of an active patient being treated by them? That is, phone number and address.	Yes, assigned staff can access patient's contact information.	16 (100)

[a]Percentages are rounded to the nearest multiple of 5%.

[b]N/A: not applicable.

[c]GDPR: General Data Protection Regulation.

Table . Survey findings regarding "workload" at Irish Mental Health Services.

Question	Majority or average response	Values, n (%)[a]
On average, how many mental health patients does a psychiatrist or psychologist attend in a full working day?	5 - 10	11 (70)
On average, how many hours do psychiatrists or psychologists work in a full day (officially)?	7 - 9	15 (90)
On average, how much time is allocated for or consumed by each mental health patient? That is, consultation, diagnosis, report writing, and prescription.	30 minutes to 1 hour	13 (80)
On average, how early can a patient get an appointment with a psychiatrist or psychologist?	Wide range of responses; ranging between 2 days, a week, and a month.	N/A[b]

[a]Percentages are rounded to the nearest multiple of 5%.

[b]N/A: not applicable.

We shared a summarized version of the IT solution plan with our survey participants and asked them how much reduction in workload do they estimate if this plan gets implemented? Results are shown in Table 7.

From a different perspective, if we optimize the existing system, we can expect approximately 35% more patients to be accommodated each day. In our assessment, some improvements could be made by simple changes in the procedural protocols or policies, such as restricting the contact information to administrative staff only.

Table . Expected reduction in workload (survey results).

Range of reduction in workload	Participants, n
0%	0
10% - 25%	7
25% - 50%	4
50% - 60%	2
60% - 75%	3
>75%	0

XSL•FO
RenderX

Technological Evaluation Findings

The relevant use of technology was evaluated from 3 perspectives, that is, digitization of psychiatric case notes, the scope of AI, and data storage and retrieval.

Digitization of Psychiatric Case Notes

Overview

Numerous challenges pertaining to data handling practices in Irish MHS can be effectively addressed through the digitization of psychiatric case notes. However, certain limitations persist. These concerns encompass the human limitations associated with memory retention and the time-consuming nature of manual typing during patient sessions. In light of these usability considerations, we explored the implementation of a speech-to-text transcription approach. This approach involves facilitating psychiatrists with a speech-to-text transcription tool during mental health treatment sessions.

State of the Art

Several digital tools have been developed for the facilitation of note-taking process. Some well-known tools include Evernote, Microsoft OneNote, Google Keep, and Apple Notes [27]. Recent progress in this industry involves the development of several audio or video recording–based software that converts the audio or video input to text format. Some notables are T-Pro, CarePatron, VoiceBoxMD, Mentalyc, Athreon, Transcription City, Upheal, etc [22,28-33]. However, such tools rely on audio or video-recorded input, which has more serious implications of its own.

This leaves us with transcription tools that process live audio data to text form without storing the sound file. Most of the tools developed for this purpose are based on different speech-to-text application programming interfaces. Some notable speech-to-text application programming interfaces are being offered by Amazon, Assembly AI, Wav2Letter, DeepSpeech, SpeechBrain, JavaScript, IBM Watson, OpenAI, Google, Intel PyTorch, etc [34-43].

Scope of AI

Upon the successful transcription of the entire treatment session conversation between the psychiatrist and the mental health patient, we propose leveraging AI for two primary purposes: (1) to generate a summary of the treatment session and (2) to identify potential disorders and recommend appropriate treatment options.

AI has become an increasingly active area of research over the past few years in precision mental health care [44,45]. AI is expected to significantly improve psychotherapy research in the coming years. It has a lot of potential in predictive modeling [46]. Since we are dealing with the summarization of psychiatric treatment session, sentiment analysis of patient's statements could also be quite helpful in identifying various disorders. Sentiment analysis includes natural language processing, text analysis, and computational linguistics for identifying and extracting subjective information in source materials [47].

We propose a human-supervised, query-based summarization AI model for verifying factual consistency and identifying conflicts between source documents and a generated summary. Simply put, we would keep the psychiatrist or psychologist involved in the process to verify and direct the AI-generated content before submission.

During this research, we studied the shortcomings of unsupervised automatic speech recognition [48], which refers to the challenge of generating text transcriptions from raw speech during a conversation between a psychiatrist and a patient. Another interesting component is automatic text summarization, which is becoming quite important because of the large amount of textual content that grows exponentially on the internet and various archives [49]. Automatic text summarization approaches are either extractive, abstractive, or hybrid [50]. Moreover, these days, there is a need for summarizing text based on queries.

Data Storage and Retrieval

In response to the risks associated with EHRs and PHRs, more advanced IT solutions have been introduced. Cloud-based storage is another solution that has gained popularity in recent years. It allows for the secure storage and sharing of data over the internet [51]. Although cloud-based data handling is quite convenient and user-friendly, it could still make the records vulnerable to unauthorized access and loss of data [52].

This brought us to the doorsteps of blockchain technology, which is a decentralized digital ledger that stores data in a secure and tamper-proof manner [53]. Each block in the chain contains a timestamp and a cryptographic hash of the previous block, making it nearly impossible to alter or delete data without being detected.

Two well-known blockchain solutions were evaluated as potential platforms for our project, that is, Hyperledger Fabric and Ethereum [54]. We first compared these blockchain solutions in terms of the nature of their technology being decentralized or centralized. Although Ethereum can be categorized as decentralized and several security analysis tools have been developed for its smart contracts, they mostly have a command line interface with low usability [55]. Moreover, we found that using Ethereum could potentially violate the privacy of mental health records; therefore, Hyperledger Fabric seems more practical for us, as it is a private permissioned blockchain and is more suitable for private data [56].

Then, we compared these technologies based on their consensus mechanisms and how fast they achieve confirmations. This again resulted in Hyperledger Fabric being more appropriate because it has fewer confirmation nodes compared to Ethereum [57]. Finally, we differentiated Ethereum and Hyperledger Fabric based on throughput. This too ended in Hyperledger Fabric's favor, as it can process far more transactions per second compared to Ethereum [57].

Comparative Evaluation—Findings

We have categorized the findings of technological evaluation into 3 parent aspects, that is, data collection, data analysis, and data storage or retrieval. Table 8 describes what sort of solutions were available to us for each aspect of psychiatric data handling,

XSL•FO
RenderX

which ones did we choose, and what are the justifications of our choices or recommendations.

Table . Comparative evaluation of different existing technological solutions.

Data handling aspects	Available solutions	Suitable choice	Justification
Data collection	Handwritten notes, tools for typing notes (Evernote, OneNote, Google Keep, Apple Notes, etc), audio or video recording tools (Voice-BoxMD, T-Pro, Upheal, Mentalyc, Athreon, CarePatron, etc), speech-to-text transcription tools.	Speech-to-text transcription tools (OpenAI, Google API[a], JavaScript API, AWS transcribe, Assembly AI, Intel PyTorch, etc).	Digital note-taking tools would only resolve readability and digitization concerns. Whereas speech-to-text transcription would also resolve the workload, cognitive reliance, and patient reluctance concerns.
Data analysis	Multidisciplinary team meetings, written or typed reports and charts, statistical analysis tools, artificial intelligence (AI).	AI-enabled statistical analysis program with "Human in the Loop," trained with psychiatry-specific data for general statistics, identification of patterns, and recommendations.	With the exception of handwritten notes or charts, a convergence of all available solutions is needed. AI must be involved to reduce workload and facilitate the process of analysis, diagnosis, and treatment.
Data storage or retrieval	Patient folders with case notes in cabinets, scanned or typed files in PCs, conventional databases (SQL, etc), cloud (AWS, Google, etc), blockchain (Ethereum, Hyperledger Fabric, etc).	Permissioned blockchain with hashed storage. For example: Hyperledger Fabric.	Cabinet storage does not resolve any challenge. Conventional databases and cloud services are vulnerable against cyberattacks. Blockchain is secure, distributed, and tamper-proof, which would also address authenticity, in legal aspects.

[a]API: application programming interface.

Discussion

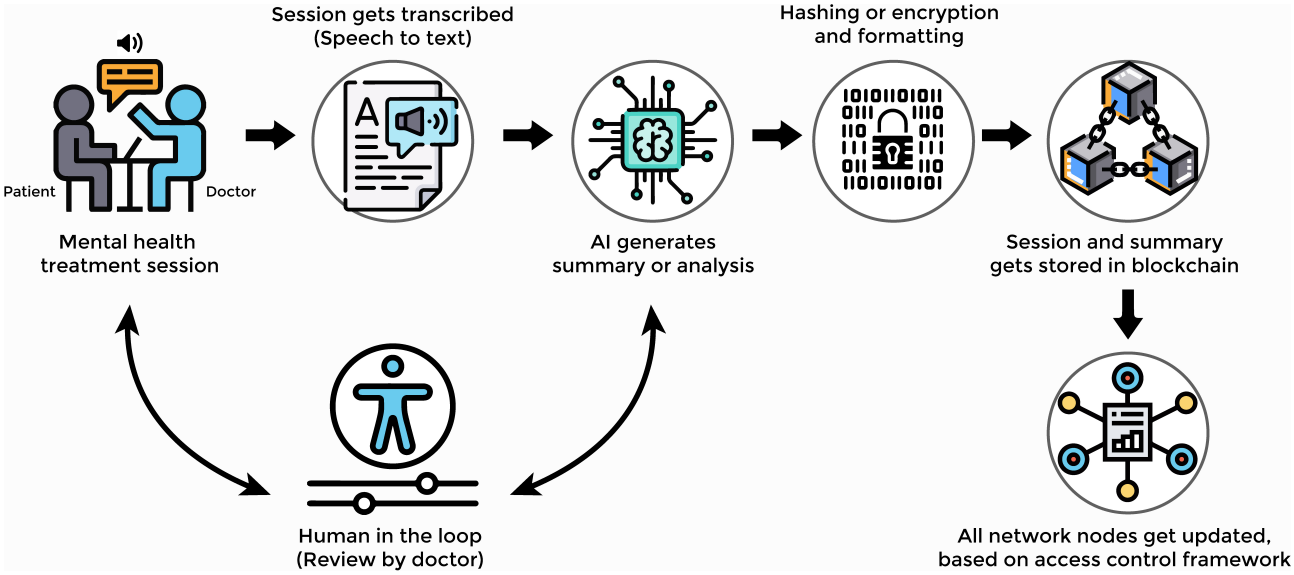
After compiling results and insights from our mixed methods, we started framing a technological solution and kept some psychiatrists in the loop to validate our technological proposals. This section summarizes the thought process behind reaching the proposed solution and how does it offer novelty.

Proposed Approach

The infographic overview of the proposed solution with its operations is given in Figure 1. If we further break the workflow into steps, it would be as follows:

- A mental health treatment session takes place, involving a patient and a doctor.
- During consultation, the conversation gets transcribed into digital text format.
- AI is then used to generate a summarized version or highlights of the treatment session. This step holds the potential to be advanced to probable diagnosis and treatment suggestions and insightful analysis by the AI tool.
- The AI-generated summary and analysis would then be reviewed by the doctor for any corrections or additions.
- Once approved by the doctor, the entire data would then get encrypted for security and hashed to reduce its size.
- The hashed data would then get stored to a permissioned blockchain.
- All nodes of the network would get updated based on an access control framework, that is, which staff members will have access to which sections of each patient record.

Figure 1. Overview of the proposed solution. AI: artificial intelligence.



XSL•FO
RenderX

Proposed Novelty

In Table 9, we have summarized 12 comparative aspects of the existing state of data handling practices in Irish MHS. It shows how our proposed solution would address, innovate, or improve each aspect.

Table . Comparative aspects between the current state and proposed solution.

Comparative aspects	Current state	Proposed solution
Digitization	Psychiatric case notes are largely based on paperwork in handwritten format.	Paperwork would get replaced by digitized psychiatric case notes.
Readability	Handwritten psychiatric case notes lack readability and hinder case study.	Digitized case notes would enable 100% readability and support the case study.
Cognitive reliance	Heavy reliance on the human ability to remember and write or type conversation.	Allows one to skip recalling or writing and focus on psychiatric evaluation instead.
Workload	Parallel workload (paperwork+digital tasks). If not write, one must type.	Speech-to-text transcription reduces workload and automates the note-making.
Expression reluctance	Note-writing makes patients reluctant and uncomfortable in sharing thoughts.	Automated transcription enables patients to feel comfortable while expressing.
Data storage	Patient's data and case are written in paper files and stored in cabinets.	Data would get stored in a permissioned blockchain in a hashed format.
Data privacy	Patient's contact data and case are explicitly described in files with stickers.	Data would be categorized in privacy terms and kept in encrypted format.
Access authorization	Unauthorized staff may access and read patient files and take photographs.	Role-based access authorization would be implemented using smart contracts.
Data tampering	Signatures are used for alteration in notes, with no digital tracking in place.	Blockchain would ensure data immutability and traceable digital footprints of all actions.
Data security	Outdated security measures for data storage and transfer, prone to cyberattacks.	Distributed yet semicentralized blockchain network would prevent cyberattacks.
Research and analysis	Handwritten notes obstruct the potential of research and psychiatric data analysis.	Digitized Psy-Notes enables psychiatric data analysis and diagnostic or treatment research.
Uniform data collection	Existing practice may result in variations in collected data by different doctors.	Transcribed treatment sessions would ensure uniformity in data for evaluation.

Challenges and Limitations

It is important to acknowledge that the initial phase of execution of such an advanced and reform-driven technological solution is not going to be a smooth process. Each aspect of this proposed plan would involve different sorts of challenges and demands, which are as follows: lack of familiarity with emerging technologies like blockchain and AI, training needs of the medical staff regarding the adopted technological solution, need for advanced usability and human-computer interaction standards, setting up the environment and distributed network of nodes, securing investment for a required equipment and routine maintenance, employment of skillful IT professionals for development and cybersecurity, employment of a dedicated research team for tracking progress and sharing insights, reliance on speech-to-text transcription tool for accommodating different Irish accents, the accuracy of the content generated by AI, integration of the proposed solution with existing systems, and amendments to the existing policies, where necessary, such as access control protocol.

The limitations related to AI and speech-to-text transcription were scaled down by adopting a human-in-the-loop approach. However, these technologies must be improved and upgraded over time to the extent where human involvement would be reduced to formality or validation. When it comes to blockchain applications, most people are concerned about the energy consumption associated with it. It is important to clarify that the energy-related drawbacks are related to those blockchain platforms that are based on a proof-of-work consensus mechanism. Our technological evaluation has recommended a permissioned blockchain platform, such as Hyperledger Fabric, which does not involve a proof-of-work mechanism.

Conclusions

Ireland's mental health challenges place it among the worst in Europe. This encouraged us to conduct a thorough case study regarding data handling practices in Irish MHS, in order to contribute toward an improved service delivery in the Irish psychiatry sector. We adopted a mixed methodology to collect data from various sources for comparative analysis and cross-validation. Our study involved literature review, interviews, web-based surveys, interaction with different organizations, analysis of some relevant inspection reports, comparative technological evaluation, and a feedback loop with stakeholders.

It was noteworthy that all results were complimenting and supporting each other. However, the list of surfaced shortcomings in Irish MHS from different aspects turned up way longer than our expectations. The main concerns highlighted by results included lack of digitization, reliance on

XSL•FO
RenderX

paperwork and doctors' cognitive abilities for writing case notes, low readability, reluctance of patients in sharing their thoughts due to confidentiality concerns, sensitive legal applications, data violations by the staff, and parallel workload.

Upon brainstorming and consultation, we finalized a solution that holds the potential to resolve multiple intertwined challenges collectively. It involves collecting psychiatric data using speech-to-text transcription, then performing analysis on it using AI, which upon doctor's review gets stored in a distributed permissioned blockchain in categorized and hashed format for retrieval based on access control rules. We shared our results and the respective proposed solution with psychiatrists working at Irish MHS, who estimated that it could reduce an average of 35% of their existing workload.

Although this proposed solution has some technical and practical challenges regarding development and usability, it is still innovative and worthy enough to be considered and executed, as it offers many positive implications. This is a novel convergence of different existing technologies that could get used smartly to address a wide range of flaws in the existing state of data handling in Irish MHS. Upon implementation, such a solution would not only facilitate the mental health practitioners in their diagnosis and treatment process but also optimize the workload and cost expenditure in the long term.

Acknowledgments

The authors would like to thank and acknowledge all mental health practitioners who facilitated their study with their participation and insightful feedback. Additionally, the authors appreciate the role played by organizations like the Mental Health Commission and the Brothers of Charity Services in shaping this case study. Finally, the authors recognize the resources and platform provided to the research team by the College of Science and Engineering and the School of Psychology in the University of Galway.

Conflicts of Interest

None declared.

References

1. The Mental State of the World Report. 2022. URL: https://mentalstateoftheworld.report/ [accessed 2025-06-16]
2. Hyland P, Vallières F, Shevlin M, et al. State of Ireland's mental health: findings from a nationally representative survey. Epidemiol Psychiatr Sci 2022 Jul 1;31:e47. [doi: 10.1017/S2045796022000312] [Medline: 35773999]
3. Mental health issues cost State €8.2bn each year, says Harris. The Irish Times. URL: https://www.irishtimes.com/news/health/mental-health-issues-cost-state-8-2bn-each-year-says-harris-1.4001595 [accessed 2023-10-08]
4. Health at a glance: Europe 2022: state of health in the EU cycle. Organisation for Economic Co-operation and Development. URL: https://www.oecd.org/health/health-at-a-glance-europe-23056088.htm [accessed 2024-01-29]
5. Bolstering the cybersecurity of the healthcare sector. European Commission. URL: https://commission.europa.eu/news/bolstering-cybersecurity-healthcare-sector-2025-01-15_en [accessed 2025-02-08]
6. Data breaches and cyber attacks in Europe in February 2024—50,884,709 records breached. IT Governance Europe. 2024 Mar 21. URL: https://www.itgovernance.eu/blog/en/data-breaches-and-cyber-attacks-in-europe-in-february-2024-50884709-records-breached [accessed 2025-02-08]
7. Ransomware attack forces 100 Romanian hospitals to go offline. Bleeping Computer. URL: https://www.bleepingcomputer.com/news/security/ransomware-attack-forces-100-romanian-hospitals-to-go-offline/ [accessed 2025-02-08]
8. Health Service Executive ransomware attack. Wikipedia. 2023. URL: https://en.wikipedia.org/w/index.php?title=Health_Service_Executive_ransomware_attack&oldid=1177522371 [accessed 2023-10-08]
9. 'Shocking' hack of psychotherapy records in Finland affects thousands. The Guardian. 2020 Oct 26. URL: https://www.theguardian.com/world/2020/oct/26/tens-of-thousands-psychotherapy-records-hacked-in-finland [accessed 2023-05-15]
10. McCarthaigh S. Locum GP gave 'sexually motivated' massage to female patient in cork, inquiry told. Irish Examiner. 2023 Mar 13. URL: https://www.irishexaminer.com/news/munster/arid-41092198.html [accessed 2023-10-08]
11. High Court orders suspension of psychiatrist over sex-with-patients allegations. Irish Independent. URL: https://www.independent.ie/irish-news/courts/high-court-orders-suspension-of-psychiatrist-over-sex-with-patients-allegations/29122901.html [accessed 2023-10-08]
12. Chawki M. Security and privacy in the era of electronic health records (EHRs). RAIS J Soc Sci 2021;5(1):1-12 [FREE Full text]
13. Keshta I, Odeh A. Security and privacy of electronic health records: concerns and challenges. Egyptian Informatics Journal 2021 Jul;22(2):177-183. [doi: 10.1016/j.eij.2020.07.003]
14. Barnett P, Goulding L, Casetta C, et al. Implementation of telemental health services before COVID-19: rapid umbrella review of systematic reviews. J Med Internet Res 2021 Jul 20;23(7):e26492. [doi: 10.2196/26492] [Medline: 34061758]
15. Bass E, Garabrant J, Salyers MP, Patterson S, Iwamasa GY, McGuire AB. eHealth use on acute inpatient mental health units: implementation processes, common practices, and barriers to use. Adm Policy Ment Health 2023 Jul;50(4):603-615. [doi: 10.1007/s10488-023-01262-1] [Medline: 36943598]

XSL•FO
RenderX

16. Mucic D, Shore J, Hilty DM. The World Psychiatric Association telepsychiatry global guidelines. J Technol Behav Sci 2023;9(3):572-579. [doi: 10.1007/s41347-023-00339-w]
17. Ramalho R, Volpe U, Gaebel W. WPA digitalization in mental health and care: empirical report on action plan and related activities. J Technol Behav Sci 2023;9(3):567-571. [doi: 10.1007/s41347-023-00352-z]
18. Instefjord MH, Aasekjær K, Espehaug B, Graverholt B. Assessment of quality in psychiatric nursing documentation—a clinical audit. BMC Nurs 2014;13(1):32. [doi: 10.1186/1472-6955-13-32] [Medline: 25349532]
19. Giacco D, Bird VJ, Ahmad T, et al. The same or different psychiatrists for in- and out-patient treatment? A multi-country natural experiment. Epidemiol Psychiatr Sci 2020;29:e10. [doi: 10.1017/S2045796018000732]
20. Mergaerts L. Defence lawyers' views on and identification of suspect vulnerability in criminal proceedings. International Journal of the Legal Profession 2022 Sep 2;29(3):281-301. [doi: 10.1080/09695958.2021.1982719]
21. Upfront with katie hannon [@rteupfront]: 'In ireland, in any sexual offence case - the therapy notes of the victim can be used against them to disprove what they are saying in trial'. solicitor sarah grace outlines why certain processes make the justice system traumatic for survivors of sexual assault. X. 2023 Nov 13. URL: https://twitter.com/RTEUpfront/status/1724212708231287243 [accessed 2024-01-30]
22. HQ TP: medical dictation, speech recognition & transcription software. T-Pro. URL: https://info.tpro.io [accessed 2023-08-06]
23. Chivilgina O, Elger BS, Benichou MM, Jotterand F. "What's the best way to document information concerning psychiatric patients? I just don't know"—a qualitative study about recording psychiatric patients notes in the era of electronic health records. PLOS ONE 2022;17(3):e0264255. [doi: 10.1371/journal.pone.0264255] [Medline: 35239698]
24. Mental Health Commission Ireland. Mental Health Commission. URL: https://www.mhcirl.ie/ [accessed 2024-01-30]
25. Mental Health Commission publishes nine inspection reports for seven approved centres. Mental Health Commission. URL: https://www.mhcirl.ie/news/mental-health-commission-publishes-nine-inspection-reports-seven-approved-centres [accessed 2024-01-30]
26. Hathaliya JJ, Tanwar S. An exhaustive survey on security and privacy issues in Healthcare 4.0. Comput Commun 2020 Mar;153:311-335. [doi: 10.1016/j.comcom.2020.02.018]
27. Pitura J. Digital note-taking for writing. In: Kruse O, Rapp C, Anson CM, et al, editors. Digital Writing Technologies in Higher Education: Theory, Research, and Practice: Springer International Publishing; 2023. [doi: 10.1007/978-3-031-36033-6_7]
28. Carepatron medical dictation. Carepatron. URL: https://www.carepatron.com/medical-dictation [accessed 2023-08-06]
29. VoiceboxMD software for psychiatrists. VoiceboxMD. URL: https://voiceboxmd.com/medical-dictation/psychiatrist/ [accessed 2023-08-06]
30. Mentalyc. URL: https://www.mentalyc.com/ [accessed 2023-08-06]
31. Athreon psychology. Athreon. URL: https://www.athreon.com/healthcare/medical-specialties/psychology/ [accessed 2023-08-06]
32. Transcription City. URL: https://transcriptioncity.co.uk/ [accessed 2023-08-07]
33. AI-powered progress notes. Upheal. URL: https://www.upheal.io/ [accessed 2023-08-06]
34. Amazon Web Services, Inc. URL: https://aws.amazon.com/transcribe/ [accessed 2023-08-06]
35. AssemblyAI. URL: https://www.assemblyai.com/ [accessed 2023-08-06]
36. Wav2Letter. GitHub. URL: https://github.com/flashlight/wav2letter [accessed 2023-08-07]
37. DeepSpeech. GitHub. URL: https://github.com/mozilla/DeepSpeech [accessed 2023-08-07]
38. SpeechBrain. GitHub. URL: https://github.com/speechbrain/speechbrain [accessed 2023-08-06]
39. How to convert voice to text using JavaScript. AssemblyAI. 2022 Mar 8. URL: https://www.assemblyai.com/blog/voice-to-text-javascript/ [accessed 2023-08-07]
40. IBM Watson Speech to Text. IBM. URL: https://www.ibm.com/products/speech-to-text [accessed 2023-08-06]
41. OpenAI Platform. URL: https://platform.openai.com [accessed 2023-08-06]
42. Transcribe API. Google Cloud. URL: https://cloud.google.com/speech-to-text/docs/transcribe-api [accessed 2023-08-06]
43. Speech-to-text AI for healthcare providers. Intel. URL: https://www.intel.com/content/www/us/en/developer/articles/reference-kit/ai-transcription-for-behavioral-health.html [accessed 2023-08-06]
44. Atzil-Slonim D, Penedo JMG, Lutz W. Leveraging novel technologies and artificial intelligence to advance practice-oriented research. Adm Policy Ment Health 2024 May;51(3):306-317. [doi: 10.1007/s10488-023-01309-3] [Medline: 37880473]
45. Bickman L. Improving mental health services: a 50-year journey from randomized experiments to artificial intelligence and precision mental health. Adm Policy Ment Health 2020 Sep;47(5):795-843. [doi: 10.1007/s10488-020-01065-8] [Medline: 32715427]
46. Horn RL, Weisz JR. Can artificial intelligence improve psychotherapy research and practice? Adm Policy Ment Health 2020 Sep;47(5):852-855. [doi: 10.1007/s10488-020-01056-9] [Medline: 32715430]
47. Ridzwan Yaakub M, Iqbal Abu Latiffi M, Safra Zaabar L. A review on sentiment analysis techniques and applications. IOP Conf Ser: Mater Sci Eng 2019 Aug 1;551(1):012070. [doi: 10.1088/1757-899X/551/1/012070]
48. Aldarmaki H, Ullah A, Ram S, Zaki N. Unsupervised automatic speech recognition: a review. Speech Commun 2022 Apr;139:76-91. [doi: 10.1016/j.specom.2022.02.005]

XSL•FO
RenderX

49. El-Kassas WS, Salama CR, Rafea AA, Mohamed HK. Automatic text summarization: a comprehensive survey. Expert Syst Appl 2021 Mar;165:113679. [doi: 10.1016/j.eswa.2020.113679]
50. Kryscinski W, McCann B, Xiong C, Socher R. Evaluating the factual consistency of abstractive text summarization. arXiv. Preprint posted online on Oct 29, 2019. [doi: 10.48550/arXiv.1910.12840]
51. Mahajan HB, Rashid AS, Junnarkar AA, et al. Integration of Healthcare 4.0 and blockchain into secure cloud-based electronic health records systems. Appl Nanosci 2023;13(3):2329-2342. [doi: 10.1007/s13204-021-02164-0] [Medline: 35136707]
52. Al-Issa Y, Ottom MA, Tamrawi A. eHealth cloud security challenges: a survey. J Healthc Eng 2019;2019:7516035. [doi: 10.1155/2019/7516035] [Medline: 31565209]
53. Chukwu E, Garg L. A systematic review of blockchain in healthcare: frameworks, prototypes, and implementations. IEEE Access 2020;8:21196-21214. [doi: 10.1109/ACCESS.2020.2969881]
54. What is Ethereum? Ethereum. URL: https://ethereum.org [accessed 2023-07-02]
55. Zeeshan R, Tal I. Technical usability assessment of security analysis tools for ethereum based smart contracts. Presented at: 2022 IEEE 22nd International Conference on Software Quality, Reliability, and Security Companion (QRS-C); Dec 5-9, 2022; Guangzhou, China p. 87-95. [doi: 10.1109/QRS-C57518.2022.00021]
56. Androulaki E, Barger A, Bortnikov V, et al. Hyperledger fabric: a distributed operating system for permissioned blockchains. Presented at: Proceedings of the Thirteenth EuroSys Conference EuroSys '18; Apr 23-26, 2018; Porto, Portugal p. 1-15. [doi: 10.1145/3190508.3190538]
57. Zhao Z. Comparison of hyperledger fabric and ethereum blockchain. Presented at: 2022 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC); Apr 14-16, 2022; Dalian, China p. 584-587. [doi: 10.1109/IPEC54454.2022.9777292]

Abbreviations

AI: artificial intelligence
BoCS: Brothers of Charity Services
EHR: electronic health record
MHS: mental health services
PHR: personal health record

Edited by A Kushniruk; submitted 30.07.24; peer-reviewed by S Khan, S Malik; revised version received 13.02.25; accepted 14.02.25; published 24.06.25.

Please cite as:
Zeeshan R, Bogue J, Gill A, Asghar MN
Pursuit of Digital Innovation in Psychiatric Data Handling Practices in Ireland: Comprehensive Case Study
JMIR Hum Factors 2025;12:e64919
URL: https://humanfactors.jmir.org/2025/1/e64919
doi:10.2196/64919

© Rana Zeeshan, John Bogue, Amna Gill, Mamoona Naveed Asghar. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 24.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Usability and Quality of the JoyPop App: Prospective Evaluation Study

Ishaq Malik, MA; Teagan Neufeld, MSc; Aislin Mushquash, PhD

Department of Psychology, Lakehead University, 955 Oliver Road, Thunder Bay, ON, Canada

Corresponding Author:
Aislin Mushquash, PhD
Department of Psychology, Lakehead University, 955 Oliver Road, Thunder Bay, ON, Canada

Abstract

Background: Mental health difficulties are increasing among Canadian postsecondary students, and many face barriers to accessing mental health care. Mobile health smartphone apps for mental health reduce common barriers to care and improve student mental health outcomes. However, students' engagement and use of mental health apps is low. Evaluating the usability and quality of mental health apps is essential not only for user engagement but also for safety and overall utility. Few mental health apps have undergone usability and quality evaluations, especially with measures explicitly designed for these apps. The JoyPop app is a resilience-building mental health app with evidence supporting its effectiveness for student mental health. It has yet to be evaluated using standardized measures of mental health app usability and quality, and the influence of usability and quality on use is unknown.

Objective: We evaluated the usability and quality of the JoyPop app and the predictive importance of usability and quality, compared to other relevant user characteristics, in predicting intentions to use the app in the future (usage intentions).

Methods: Participants (N=183) completed preapp measures assessing demographics and personality traits, then used the app for 1 week, and then completed postapp measures assessing the usability, quality, and use of the JoyPop app. Usability (overall; and subscales: ease of use, interface and satisfaction, and usefulness) and quality (objective, subjective, and perceived impact) were assessed with descriptive statistics. Multiple regression analyses tested the predictive importance of usability and quality on usage intentions after controlling for other user characteristics.

Results: Participants rated the JoyPop app's overall usability as "very good" (mean 5.63, SD 0.85). Participants rated the JoyPop app's overall objective quality as "excellent" (mean 4.06, SD 0.54). Subjective quality ratings were good, with many participants (135/183, 73.8%) indicating they would recommend the app to others. Participants rated the app as having a moderate and helpful impact on their mental health and coping skills (mean 3.48, SD 0.88). In each regression model, usability (β=.56, P<.001) and quality (β=.52, P<.001) were the strongest predictors and predicted usage intentions over and above other user characteristics.

Conclusions: Results align with prior research evaluating the JoyPop app and maintain that it is an engaging and high-quality mental health app that can support students. Findings provide important insight into the optimal design of mental health apps for students and inform adaptations to future iterations of the JoyPop app.

(JMIR Hum Factors 2025;12:e65472) doi:10.2196/65472

KEYWORDS

usability; quality; mobile health; mental health app; postsecondary students; uMARS; MAUQ; evaluation study; mental health; app; resilience; Canada; student; usefulness; satisfaction; regression analyses; coping; smartphone app; mHealth app; mHealth; users; user version of the Mobile Application Rating Scale; Mobile Application Usability Questionnaire

Introduction

Background

Mental health difficulties among Canadian postsecondary students are increasing [1-3]. Unaddressed mental health difficulties can exacerbate existing challenges and negatively impact academic and long-term health outcomes [4,5]. Concerningly, students face a multitude of barriers (eg, long wait times and financial constraints) to access mental health support [4,6], resulting in prolonged distress, difficulties for campus mental health services, and low future help-seeking behaviors [7,8]. Novel approaches that increase students' access to mental health support are needed [9,10].

Mobile health smartphone apps designed to improve mental health are promising tools to support students [10,11]. Mental health apps also expand service capacity, improve availability and access to support, and reduce stigma, cost, time, and travel-related barriers for students [12-14]. Mental health apps are especially promising for students because they use smartphones frequently and are open to using apps for

XSL•FO
RenderX

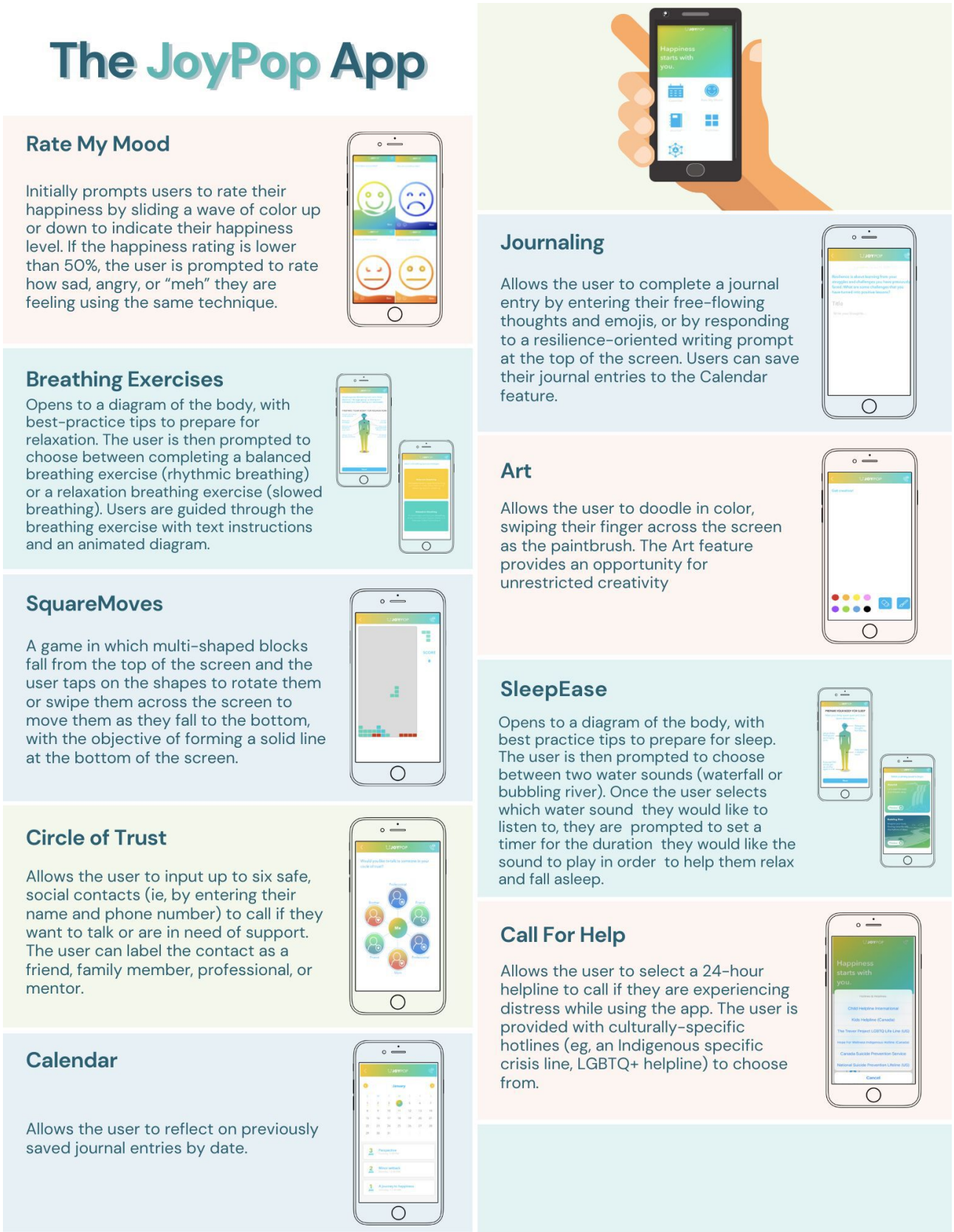
health-related reasons [15,16]. Research examining the effectiveness and efficacy of mental health apps among students shows that apps can improve well-being and reduce stress, depression, and anxiety [17,18].

The JoyPop App

The JoyPop app is a mental health app designed to improve resilience through evidence-based features that promote emotion regulation and adaptive coping among youths and emerging adults [19,20]. The JoyPop app was co-designed by youths, researchers, and service providers [20]. The app's transdiagnostic focus makes it a valuable addition to usual mental health supports and applicable to a diverse range of students [21,22]. The app is particularly relevant to students as many are undergoing critical life transitions, and timely interventions that promote resilience can have a long-lasting positive impact on well-being and mental health [23,24]. Details of each feature are provided in Figure 1, and the rationale for the inclusion of each feature is documented elsewhere [19,25].

XSL•FO
RenderX

Figure 1. Summary and highlights of features in the JoyPop app. LGBTQ+: lesbian, gay, bisexual, transgender, and queer.



There is an increasing multimethod evidence base supporting the effectiveness and value of the JoyPop app as a complement to mental health services across diverse users in clinical and nonclinical settings [19,25-29]. Among students, using the JoyPop app is associated with improved emotion regulation, improved depressive symptoms, and reduced stress [19,29]. Additionally, students perceive that the app improves emotion regulation and facilitates opportunities to implement adaptive coping skills [28]. Multiple randomized controlled trials (with students more generally [30] and with youths waiting for mental

XSL•FO
RenderX

health services [31]) evaluating the app's effectiveness in improving emotion regulation and mental health outcomes are ongoing.

Although evidence of a mental health app's effectiveness is essential, it does not guarantee users will engage with or enjoy their experience with an app [15,32]. Student engagement and long-term uptake of mental health apps are typically low [33-35] and can reduce the overall positive outcomes associated with app use [32,36]. Engagement, safety, and overall utility are affected by how usable an app is perceived to be and a user's perspective on the quality of an app [15,34,36]. Despite this, the majority of mental health apps available in major marketplaces have not undergone usability and quality evaluations [35,37,38].

Usability

Usability is defined as a "quality attribute that assesses how easy interfaces are to use" [39]. Evaluating mental health app usability can reveal factors (eg, functionality, feature usefulness, layout, and readability) that reduce engagement and use [35,36]. Mental health app usability studies show that students value mental health apps with a convenient and intuitive design, customization, anonymity, privacy, peer involvement, games, and options to access professional help [40,41].

Most studies assessing usability include measures designed for general technologies [42,43], such as the System Usability Scale [44]. Applying general technology usability measures to mental health apps can provide some valuable information, but they miss identifying factors unique to mobile health apps, such as limited computation power, security and privacy challenges, and characteristics associated with small screens and portable devices [43,45]. Measures designed specifically for apps, such as the Mobile Application Usability Questionnaire (MAUQ) [43], are rarely used [42,46]. Usability evaluations with validated and standardized measures developed for apps are recommended [43,46].

Quality

Evaluating the quality (eg, effectiveness, credibility, reliability, and safety) of mental health apps is also essential. Users are more likely to use a mental health app if it is provided by a credible source, positively impacts their mental health, and has evidence of safety and effectiveness [32]. Without proof of quality, users may rely on subjective star ratings, which can be inconsistent with an app's actual quality and clinical utility [47-49]. Claims of mental health app quality are not consistently assessed by regulatory bodies [49,50], and most apps lack evidence of quality and effectiveness, clinical input from professionals, and evidence-based features [51,52]. Poor-quality mental health apps can therefore pose significant risks to users, such as misdiagnosis, unknown side effects, and physical and mental harm [53].

Several resources are now available to locate, assess, and regulate the quality of mental health apps, such as app rating guidelines and platforms [37,54,55]. For example, the Mobile Health App Database (MHAD) rating platform provides quality ratings for a wide variety of mental health apps to help guide users and health care providers in selecting quality apps [54,56]. The Mobile Application Rating Scale (MARS) [49] is the most popular app rating guideline [42] and allows experts in mental health to rate apps based on quality indicators from health-related fields [49]. The user version of the Mobile Application Rating Scale (uMARS) is also available to help consumers without training and expertise to rate the quality of mobile health apps [55].

This Study

In this study, we aimed to (1) evaluate the usability and quality of the JoyPop app and (2) test the relative predictive importance of usability and quality on usage intentions, compared to other user characteristics. We used questionnaires explicitly designed to evaluate mobile health apps [42,46] to ensure that important facets of usability and quality unique to these apps were captured [43,57]. We chose age, gender, neuroticism, and agreeableness as relevant user characteristics based on past research highlighting their influence on mental health app use [58-61]. Because of the extensive literature highlighting the importance of usability and quality in mental health app use [32,37,45], we hypothesized that usability and quality would significantly predict JoyPop usage intentions over and above the other user characteristics.

Methods

Design

We used a 1-week prospective design to ensure students had sufficient time to use the app thereby facilitating more informed app evaluations.

Ethical Considerations

The study was reviewed and approved by the Lakehead University research ethics board (Romeo #1470023). All participants provided informed consent before beginning the study, were under no obligation to participate, and were free to withdraw at any time without penalty. All data were deidentified. We compensated participants for their time via cash or e-transfer or bonus points toward an eligible psychology course for completing part 1 and part 3 of the study (CAD $10 [US $7.28] or 1 bonus point for part 1, CAD $20 [US $14.57] or 1.5 bonus points for part 3).

Participants and Procedure

Participants were eligible for the study if they were students at Lakehead University who spoke and read fluently in English. We recruited students throughout the fall (2023) and winter (2024) terms (N=183). To increase the representativeness of our sample, we used multiple recruitment methods (eg, posters and class emails), had no exclusionary criteria based on student characteristics (eg, year of study), and provided refurbished iPhones with the JoyPop app for students who did not have a suitable device of their own. The study had 3 parts. In part 1, students attended an orientation session where they provided informed consent, learned about downloading and using the app, and completed preapp measures (ie, demographics and personality traits). In part 2, the research team reminded participants (via text or email) to use the app twice daily for 1

XSL•FO
RenderX

week. In part 3, participants were emailed or texted a link to complete postapp measures (ie, usability and quality).

Measures

Demographics

To describe the characteristics of the sample, we captured participant demographics (13 items) along with students' experience with mental health apps (5 items) by adapting items from prior questionnaires assessing technology experience [62,63]. We used the age and gender of participants as control variables in our analysis.

The Big Five Aspects Scale

We used the Big Five Aspects Scale (BFAS) [64] to gather information on participant personality traits. The BFAS consists of 100 items rated on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Participants rate their level of agreement with how well brief statements generally describe how they are. The BFAS provides scores for the traditional Big Five domains and 2 aspects for each domain. We used scores for the neuroticism and agreeableness domains as control variables in our analysis. The BFAS shows excellent internal consistency and test-retest reliability in postsecondary samples [64-66]. The internal consistency of domains used in this study was excellent (neuroticism, α=.9; agreeableness, α=.82).

MAUQ

The MAUQ [43] assesses the usability of mobile health apps. We used the 18-item standalone patient version of the MAUQ in this study as it best suits the context and design of the JoyPop app. Overall usability includes 3 subscales: ease of use (5 items), interface and satisfaction (7 items), and usefulness (6 items). Participants rate their level of agreement with brief statements on a 7-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree). We calculated the overall usability score by computing the mean across all items. We derived subscale scores by computing the mean of scores for relevant subscale items. Higher scores indicate greater usability. Consistent with prior research [67,68], we used descriptors based on quartiles derived from a maximum possible score of 7 on the MAUQ scales to facilitate reporting and interpreting results: 0 - 1.75 (poor), 1.76 - 3.5 (moderate), 3.51 - 5.25 (good), and 5.26 - 7 (very good). The MAUQ has adequate psychometric properties across samples, including among postsecondary students [43]. The internal consistency of the overall scale score and subscale scores in this study was strong (overall scale, α=.92; ease of use, α=.79; interface and satisfaction, α=.86; usefulness, α=.85).

uMARS

The uMARS is a simple and reliable tool for end users to comprehensively evaluate mobile health app quality via indicators of objective and subjective quality, and perceived impact on a target health behavior [55]. We used the 27-item uMARS to evaluate the quality of the JoyPop app from the user's perspective. The 16-item "objective quality" scale of the uMARS allows users to rate app quality based on indicators derived from research and expert and clinician recommendations on a 5-point Likert scale ranging from 1 (inadequate) to 5 (excellent). The "objective quality" scale has 4 subscales: engagement (5 items), functionality (4 items), aesthetics (3 items), and information (4 items). We derived the overall and subscale scores by calculating the mean of relevant items (items rated as N/A were omitted). Based on item ratings and past research [69,70], we used descriptors for ranges of scores: 0 - 1 (inadequate), 1 - 2 (poor), 2 - 3 (acceptable), 3 - 4 (good), and 4 - 5 (excellent), to facilitate the interpretation of results and comparisons with other mental health apps. The overall scale and subscale scores have good psychometric properties [55]. The internal consistencies in this study were good (objective quality, α=.89; engagement, α=.77; functionality, α=.75; aesthetics, α=.78; information, α=.79).

The 4-item "subjective quality" subscale of the uMARS allows users to rate different indicators (eg, star ratings, whether it is worth recommending) of a mobile health app's subjective quality on a 5-point Likert scale ranging from 1 to 5 (anchors vary across questions). The 6-item "perceived impact" subscale of the uMARS allows users to rate its impact on user knowledge, awareness, intention to change, help-seeking, behavior change, and attitudes on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). This subscale was designed so that the wording of the items could be adapted to fit the function (eg, to impact mental health) of an app. We calculated the subjective quality and perceived impact subscale scores by calculating the mean of the relevant items (higher scores indicate stronger quality and impact). The subjective quality scale and perceived impact scale have good psychometric properties [55]. The internal consistency in this study was good (subjective quality, α=.79; perceived impact, α=.93).

Usage Intentions

We used an item ("How many times do you think you would use the JoyPop app in the next 12 months if it was relevant to you") from the subjective quality scale on the uMARS to assess participants' intentions to use the app in the future (usage intentions). Participants responded on a 5-point scale ranging from 1 (none) to 5 (>50 times). (Although we tracked the number of days participants used the app across the study period, we did not include this as our primary app use outcome variable as there was insufficient variation; that is, the majority of participants used the app every day; see descriptive results below.)

Statistical Analysis

We used SPSS (version 29, IBM Corporation) for the analyses. We calculated and examined the descriptive statistics for demographics and personality traits to characterize the sample. We calculated descriptive statistics for each scale and item of the MAUQ and uMARS to evaluate the usability and quality of the JoyPop app. To test our hypotheses that usability and quality would predict app use over and above other user characteristics, we conducted 2 multiple regression analyses. The first model included overall usability as a predictor; the second model included overall objective quality as a predictor. Both models included age, gender, agreeableness, and neuroticism as covariates. We included all predictors together for each model because our focus was to examine the relative importance of quality and usability in predicting use.

XSL•FO
RenderX

Results

User Characteristics

We present the characteristics of the sample in Table 1. On average, participants were aged 22.6 (SD 7.08; range 16 - 56) years. Most participants were women (150/183, 82%), White (100/183, 54.6%) or Black (33/183, 18%), full-time students (156/183, 85.2%), and in their first (81/183, 44.3%) or second (54/183, 29.5%) year of university. The majority of participants reported no current or past use of mental health apps (148/183, 80.9%), and most did not currently have a mental health app installed on their phones/devices (158/183, 86.3%). Among participants who used mental health apps in the past (35/183, 19.1%), most used them for 0 - 6 months (115/183, 62.8%). Descriptive statistics of relevant personality traits (ie, neuroticism: mean 3.18, SD 0.63; agreeableness: mean 4, SD 0.42) were consistent with prior research among students [64,66]. Most participants used the app every day of the study period (mean 4.69, SD 0.81).

Table . User characteristics (N=183).

	Value, n (%)
Gender	
Women	150 (82)
Men	33 (18)
Ethnicity	
White	100 (54.6)
Black	33 (18)
South Asian	20 (10.9)
Indigenous	10 (5.5)
Southeast Asian	7 (3.8)
Other (eg, Middle Eastern, South American, and Asian Canadian)	13 (7.1)
Country of birth	
Canada	116 (63.4)
Nigeria	26 (14.2)
India	14 (7.7)
Pakistan	4 (2.2)
Philippines	3 (1.6)
Other (eg, Italy, Barbados, Vietnam, Sri Lanka, Uganda, and China)	20 (10.9)
Program	
Nursing	66 (36.1)
Psychology	62 (33.9)
Education	10 (5.5)
Social work	7 (3.8)
Computer science	7 (3.8)
Biology	7 (3.8)
Other (eg, kinesiology, business, outdoor recreation, and political science)	24 (13.1)
Year of university	
1	81 (44.3)
2	54 (29.5)
3	26 (14.2)
4	19 (10.4)
>5	3 (1.6)

XSL•FO
RenderX

Usability

Participants rated the JoyPop app's overall usability as "very good" (mean 5.63, SD 0.85; range 2.22 - 7), and 69.9% (128/183) of the sample had scores in the "very good" range. Participants rated the ease of using the app as "very good" (mean 6.37, SD 0.67; range 3.2 - 7), and 92.3% (169/183) of the sample reported scores in the "very good" range. Participants rated the interface and satisfaction with the app as "very good" (mean 5.7, SD 0.99; range 1.57 - 7), and 68.9% (126/183) had scores within the "very good" range. Participants rated the usefulness of the app as "good" (mean 4.93, SD 1.17; range 1.5 - 7), with 42.1% (77/183) of scores in the "good" range and 38.3% (70/183) of scores in the "very good" range. We found that all item scores (see Table 2) on the ease of use and interface and satisfaction subscales fell within the "very good" range, and all item scores on the usefulness subscale were in the "good" range.

Table . Means and SDs of MAUQ[a] subscale items.

Subscale items	Mean (SD)
Ease of use	
1. The app was easy to use.	6.58 (0.66)
2. It was easy for me to learn to use the app.	6.64 (0.55)
3. The navigation was consistent when moving between screens.	6.37 (0.94)
4. The interface of the app allowed me to use all the functions (such as entering information, responding to reminders, and viewing information) offered by the app.	6.32 (1.04)
5. Whenever I made a mistake using the app, I could recover easily and quickly.	5.96 (1.20)
Interface and satisfaction	
6. I like the interface of the app.	5.75 (1.23)
7. The information in the app was well organized, so I could easily find the information needed.	6.29 (0.92)
8. The app adequately acknowledged and provided information to let me know the progress of my action.	5.50 (1.45)
9. I feel comfortable using this app in social settings.	5.84 (1.24)
10. The amount of time involved in using this app has been fitting for me.	5.59 (1.42)
11. I would use this app again.	5.42 (1.50)
12. Overall, I am satisfied with this app.	5.55 (1.50)
Usefulness	
13. The app would be useful for my health and well-being.	5.19 (1.58)
14. The app improved my access to health care services.	4.63 (1.62)
15. The app helped me manage my health effectively.	4.52 (1.55)
16. This app has all the functions and capabilities I expected it to have.	4.85 (1.71)
17. I could use the app even when the internet connection was poor or not available.	5.16 (1.31)
18. This app provided an acceptable way to receive health care services, such as accessing educational materials, tracking my own activities, and performing self-assessments.	5.22 (1.45)

[a]MAUQ: Mobile Application Usability Questionnaire.

Quality

Participants rated the overall objective quality of the app as excellent (mean 4.06, SD 0.54; range 1.96 - 5), with 60.7% (111/183) of participants rating app quality as "excellent." Participants rated the app's functionality (mean 4.47, SD 0.52; range 2.75 - 5; 136/183, 74.3% in the "excellent range"), aesthetics (mean 4.17, SD 0.67; range 1 - 5; 101/183, 55.2% "excellent"; 41/183, 21.4% "good"), and information (range 1 - 5; 107/183, 58.5% "excellent"; 39/183, 21.3% "good") as "excellent." Participants rated the app's engagement as "good" (range 1 - 5; 74/183, 40.4% "good"; 34/183, 18.6% "excellent").

Regarding subjective quality (see Table 3), we found that many participants (144/183, 78.7%) would recommend the app to others. For our measure of app use, which asked how often participants would use the app in the next 12 months, 24% (44/183) reported 3 - 10 times, 42.1% (77/183) reported 10 - 50 times, and 19.1% (35/183) reported >50 times. Most participants (132/183, 72.1%) did not want to pay for the app. The most

XSL•FO
RenderX

frequent star rating of the app was 4 stars (92/183, 50.3%), followed by 3 stars (62/183, 33.9%).

Table . Means, SDs, and frequencies of "subjective quality" item responses (N=183).

Item and response options	Responses, n (%)	Responses, mean (SD)
Would you recommend the JoyPop app to people who might benefit from it?		3.69 (1.22)
Not at all; I would not recommend this app to anyone	7 (3.8)	
There are very few people I would recommend this app to	32 (17.5)	
Maybe, there are several people I would recommend this app to	33 (18)	
There are many people I would recommend this app to	49 (26.8)	
Definitely, I would recommend this app to everyone	62 (33.9)	
How many times do you think you would use the JoyPop app in the next 12 months if it was relevant to you?		3.55 (1.61)
None	19 (10.4)	
1 to 2	8 (4.4)	
3 to 10	44 (24)	
10 to 50	77 (42.1)	
>50	35 (19.1)	
Would you pay for this app?		1.9 (0.98)
Definitely not	82 (44.8)	
Probably not	50 (27.3)	
Unsure	40 (21.9)	
Probably yes	9 (4.9)	
Definitely yes	2 (1.1)	
What is your overall (star) rating of the app?		3.54 (0.80)
One	3 (1.6)	
Two	13 (7.1)	
Three	62 (33.9)	
Four	92 (50.3)	
Five	13 (7.1)	

The overall perceived impact of the app on participants' mental health and coping skills (see Table 4) was moderate (mean 3.48, SD 0.88; range 1 - 5). Most participants rated the app as having a moderate impact (55/183, 30.1% of scores between ≥3 and <4) or a strong impact (45/183, 24.6% of scores ≥4). Many participants agreed or strongly agreed that the app had a positive impact on their awareness (107/183, 58.5%), behaviors (103/183, 56.3%), knowledge and understanding (103/183, 56.3%), intentions and motivation (106/183, 57.9%), and attitudes (100/183, 54.6%) related to mental health and coping skills. Most agreed or strongly agreed that the app had a helpful impact on their willingness to seek further help for mental health and coping skills if needed (115/183, 62.8%).

XSL•FO
RenderX

Table . Descriptive statistics for perceived impact subscale items (N=183).

Items	Response option frequencies, n (%)					Response, mean (SD)
	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree	
Increased my awareness of the importance of addressing mental health and coping skills	7 (3.8)	28 (15.3)	41 (22.4)	78 (42.6)	29 (15.8)	3.51 (1.05)
Increased my knowledge and understanding of mental health and coping skills	10 (5.5)	26 (14.2)	44 (24)	82 (44.8)	21 (11.5)	3.42 (1.04)
Changed my attitudes toward improving mental health and coping skills	9 (4.9)	32 (17.5)	42 (23)	79 (43.1)	21 (11.5)	3.39 (1.06)
Increased my intentions and motivation to address mental health and coping skills	7 (3.8)	26 (14.2)	44 (24)	80 (43.7)	26 (14.2)	3.49 (1.02)
Would encourage me to seek further help to address mental health and coping skills	6 (3.3)	20 (10.9)	42 (23)	81 (44.2)	34 (18.6)	3.64 (1.01)
Will improve my mental health and coping skills	10 (5.5)	19 (10.4)	51 (27.9)	86 (47)	17 (9.3)	3.44 (0.99)
Total score	—[a]	—	—	—	—	3.48 (0.88)

[a]Not applicable.

Usability, Quality, and Usage Intentions

We present the results of each regression model examining the relative predictive importance of usability and quality on usage intentions in Table 5. We found evidence of heteroskedasticity for age, so we used robust standard errors to obtain unbiased estimates of the regression coefficients. Our first regression model tested the predictive importance of usability on usage intentions. Age (β=.15; *P*=.008) and gender (β=.12; *P*=.04) were both significant and positive predictors of usage intentions. Neuroticism and agreeableness were not significant predictors of usage intentions. In line with our hypothesis, usability significantly and positively predicted usage intentions over and above other user characteristics, and usability was the strongest predictor in the model ($F_{5,177}$=15.69; *P*<.001; $R^2_{adjusted}$=0.29; β=.56; *P*<.001). Our second regression model tested the predictive importance of quality on usage intentions. Age was a significant and positive predictor of usage intentions (β=.17; *P*=.008). Gender, neuroticism, and agreeableness were not significant predictors of usage intentions. Consistent with our hypothesis, quality significantly and positively predicted usage intentions over and above other user characteristics and was the strongest predictor in the model ($F_{5,177}$=12.58; *P*<.001; $R^2_{adjusted}$=0.24; β=.52; *P*<.001).

XSL•FO
RenderX

Table . Multiple regression analysis predicting intention to use the JoyPop app in the future.

Predictors	B[a]	Robust SE	β	*P* value
Model 1				
Constant	0.14	1.05	—[b]	.90
Gender	–0.38	0.18	.12	.04
Age	0.03	0.01	.15	.01
Neuroticism	–0.02	0.13	–.01	.86
Agreeableness	–0.26	0.18	–.10	.15
Usability	0.76	0.10	.56	<.001
Model 2				
Constant	0.26	1.07	—	.81
Gender	–0.36	0.20	.12	.07
Age	0.03	0.01	.17	.01
Neuroticism	–0.06	0.14	–.03	.66
Agreeableness	–0.34	0.18	–.12	.06
Quality	1.12	0.17	.52	<.001

[a]Unstandardized coefficient.

[b]Not applicable.

Discussion

Usability and Quality of the JoyPop App

We evaluated the usability and quality of the JoyPop app and assessed their impact on intentions to use the app in the future after controlling for other user characteristics. Participants rated the overall usability as “very good” and all indicators comprising usability subscales were in the “good” to “very good” range. The cumulative evidence strongly supports that JoyPop app users are comfortable and satisfied with using it, find it easy to learn and use, and view it as a helpful tool to support mental health and well-being in different contexts (eg, clinical and academic settings) [25-28]. Usability ratings for the JoyPop app from this study are also similar to ratings of other popular mental health apps [71]. These findings will support organizations and consumers who are considering offering or using the JoyPop app.

Participants rated the JoyPop app’s overall objective quality, functionality, aesthetics, and quality of information as “excellent” and the app’s engagement as “good.” Results are consistent with past research on the JoyPop app, highlighting it as an engaging, high-quality mental health app with strong functionality, visual design, and helpful evidence-based content [25-28]. Results also indicate that the JoyPop app falls within the upper ranges for objective quality when compared to popular and research-supported mental health apps (eg, Calm, PTSD Coach, SuperBetter, and Destressify) rated by experts [54,72]. For example, across approximately 50 mental health apps rated using the MARS, the average objective quality was 3.54 (range 1.63 - 4.75) [54]. Among 19 popular research-supported mental health apps, the average objective quality score was 3.52 (SD 0.71), and average subscale scores were 3.98 (SD 0.82) for engagement, 3.42 (SD 0.80) for functionality, 3.23 (SD 0.90) for aesthetics, and 3.47 (SD 0.69) for information [72].

Response frequencies to items on the subjective quality scale also support the quality of the JoyPop app. Although these are subjective ratings, their importance is highlighted by the consistency between subjective consumer ratings of mental health apps and objective expert reviews of mental health apps [72]. Subjective quality results also showed that most participants would not pay for the app. This is not surprising, as research consistently shows that people are less likely to use apps that have associated costs [32,73]. Taken together, the high objective and subjective quality ratings of the JoyPop app across indicators of mobile health app quality, combined with its quality ratings compared to mental health apps reviewed by experts, strongly support its safety, credibility, helpfulness, and use for students. Future research will benefit from having experts evaluate the quality of the JoyPop app (eg, using the MARS) to validate this study’s findings and by compiling end-user quality evaluations of mental health apps to facilitate comparisons.

Finally, more than half of the participants perceived that the JoyPop app positively impacted their mental health and coping skills in all areas queried. This finding further supports the overall quality of the app. The broad range of areas impacted by the JoyPop app is consistent with past research demonstrating its effectiveness and benefits among diverse samples of users and supports its overall ability to deliver users an effective experience [19,25-29]. Results from rigorously controlled trials examining effectiveness are still needed. However, results on the app’s perceived impact on mental health and coping skills can delineate whether future outcomes are meaningful and impactful, provide researchers with an understanding of the potential mechanisms and reasons why the app might be

XSL•FO
RenderX

effective, and inform outcome measures that can be assessed in rigorously controlled trials (eg, measures of help-seeking and coping skills).

Influence of Usability and Quality on Usage Intentions

In our regression models, overall usability and objective quality were the strongest predictors of usage intentions and predicted usage intentions over and above age, gender, neuroticism, and agreeableness. Although we did not objectively measure app usage (eg, number of days used), behavioral intentions are a strong predictor of actual behavior, and user intentions are consistently shown to predict the use of digital health technologies [74-76]. Consequently, our results align with research demonstrating the substantial impact of quality and usability on engagement, long-term uptake, and use of mental health apps [32,35,37]. The importance of usability and quality in predicting participants' intentions to use the JoyPop app in the future, along with insight into indicators influencing its quality and usability, provides integral information for JoyPop and other mental health app development teams in optimal app design to improve user experience. An iterative process can be used in which development teams have users rate the usability and quality of the app, determine specific indicators requiring improvement, adapt the app based on these indicators, and then re-evaluate the usability and quality among similar populations.

Limitations

This study has some important limitations. First, we used a convenience sample primarily comprised of women, which may reduce the generalizability of results to men and the broader student population. However, the demographics and diversity of our sample help support the generalizability of the results to students because this study's demographics are similar to those found among students across Ontario [77]. Second, in our study, we requested that participants use the app twice a day over 7 days. This likely contributed to the limited variability in the number of days participants used the app and necessitated our use of a subjective measure of usage intention instead of actual use. To better capture objective usage metrics, studies conducted over longer periods without prompting participants to use the app are needed to establish the relative predictive importance of usability and quality on app use [57].

Conclusions

Mental health apps are promising tools for students' mental health. This study evaluated the usability and quality of a mental health app (JoyPop) designed to improve resilience while examining the relative predictive importance of usability and quality in predicting intentions to use the app in the future. Participants rated the JoyPop app as having "very good" usability. Participants rated the app as having high overall quality because of its "excellent" objective quality, good subjective quality, and helpful impact on their mental health and coping skills. After accounting for user characteristics, usability and quality were the strongest predictors of participants' intentions to use the JoyPop app in the future. Findings facilitate consistency and comparisons of mental health app usability and quality and contribute to the growing evidence base supporting the JoyPop app as an engaging and high-quality mental health app that can support students.

Conflicts of Interest

The senior author (AM) recently acquired intellectual property ownership rights for the JoyPop app (in June 2024). Prior to this time, intellectual property rights were owned by another researcher and institution. AM did not own any intellectual property rights when the study was developed, when data were collected, when data were analyzed, or when the initial draft of the manuscript was developed.

References

1. Canadian reference group—executive summary, spring 2022. : American College Health Association; 2022 URL: https://www.acha.org/wp-content/uploads/2024/07/NCHA-III_SPRING_2022_CANADIAN_REFERENCE_GROUP_EXECUTIVE_SUMMARY.pdf [accessed 2025-06-05]
2. Linden B, Stuart H. Post-secondary stress and mental well-being: a scoping review of the academic literature. Can J Commun Ment Health 2020;39(1):1-32. [doi: 10.7870/cjcmh-2020-002]
3. Linden B, Boyes R, Stuart H. Cross-sectional trend analysis of the NCHA II survey data on Canadian post-secondary student mental health and wellbeing from 2013 to 2019. BMC Public Health 2021 Mar 25;21(1):590. [doi: 10.1186/s12889-021-10622-1] [Medline: 33765965]
4. Moghimi E, Stephenson C, Gutierrez G, et al. Mental health challenges, treatment experiences, and care needs of post-secondary students: a cross-sectional mixed-methods study. BMC Public Health 2023 Apr 6;23(1):655. [doi: 10.1186/s12889-023-15452-x] [Medline: 37020282]
5. Wood D, Crapnell T, Lau L, Bennett A, Lotstein D, Ferris M, et al. Emerging adulthood as a critical stage in the life course. In: Halfon N, Forrest C, Lerner R, Faustman E, editors. Handbook of Life Course Health Development: Springer; 2018:123-143. [doi: 10.1007/978-3-319-47143-3_7]
6. Dunley P, Papadopoulos A. Why is it so hard to get help? Barriers to help-seeking in postsecondary students struggling with mental health issues: a scoping review. Int J Ment Health Addiction 2019 Jun;17(3):699-715. [doi: 10.1007/s11469-018-0029-z]

XSL•FO
RenderX

7. Jaworska N, De Somma E, Fonseka B, Heck E, MacQueen GM. Mental health services for students at postsecondary institutions: a national survey. Can J Psychiatry 2016 Dec;61(12):766-775. [doi: 10.1177/0706743716640752] [Medline: 27310230]
8. Ogrodniczuk JS, Kealy D, Laverdière O. Who is coming through the door? A national survey of self‐reported problems among post‐secondary school students who have attended campus mental health services in Canada. Couns Psychother Res 2021 Dec;21(4):837-845. [doi: 10.1002/capr.12439]
9. Kazdin AE. Addressing the treatment gap: a key challenge for extending evidence-based psychosocial interventions. Behav Res Ther 2017 Jan;88:7-18. [doi: 10.1016/j.brat.2016.06.004] [Medline: 28110678]
10. e-Mental health in Canada: transforming the mental health system using technology. : Mental Health Commission of Canada; 2014 URL: https://www.mentalhealthcommission.ca/wp-content/uploads/drupal/MHCC_E-Mental_Health-Briefing_Document_ENG_0.pdf [accessed 2025-06-21]
11. Supporting mental health and well-being for post-secondary students. : Mental Health Commission of Canada; 2020 URL: https://mentalhealthcommission.ca/wp-content/uploads/2021/05/PSS_Standard_Leaflet_eng.pdf [accessed 2025-06-05]
12. Becker TD, Torous JB. Recent developments in digital mental health interventions for college and university students. Curr Treat Options Psych 2019 Sep;6(3):210-220. [doi: 10.1007/s40501-019-00178-8]
13. Price M, Yuen EK, Goetter EM, et al. mHealth: a mechanism to deliver more accessible, more effective mental health care. Clin Psychol Psychother 2014;21(5):427-436. [doi: 10.1002/cpp.1855] [Medline: 23918764]
14. Wiljer D, Shi J, Lo B, et al. Effects of a mobile and web app (Thought Spot) on mental health help-seeking among college and university students: randomized controlled trial. J Med Internet Res 2020 Oct 30;22(10):e20790. [doi: 10.2196/20790] [Medline: 33124984]
15. Bautista J, Schueller SM. Understanding the adoption and use of digital mental health apps among college students: secondary analysis of a national survey. JMIR Ment Health 2023 Mar 22;10:e43942. [doi: 10.2196/43942] [Medline: 36947115]
16. Selected online activities by gender, age group and highest certificate, diploma or degree completed. : Statistics Canada; 2023 Jul 20. [doi: 10.25318/2210013701-eng]
17. Choudhury A, Kuehn A, Shamszare H, Shahsavar Y. Analysis of mobile app-based mental health solutions for college students: a rapid review. Healthcare (Basel) 2023 Jan 16;11(2):272. [doi: 10.3390/healthcare11020272] [Medline: 36673640]
18. Oliveira C, Pereira A, Vagos P, Nóbrega C, Gonçalves J, Afonso B. Effectiveness of mobile app-based psychological interventions for college students: a systematic review of the literature. Front Psychol 2021;12:647606. [doi: 10.3389/fpsyg.2021.647606] [Medline: 34045994]
19. MacIsaac A, Mushquash AR, Mohammed S, Grassia E, Smith S, Wekerle C. Adverse childhood experiences and building resilience with the JoyPop app: evaluation study. JMIR Mhealth Uhealth 2021 Jan 4;9(1):e25087. [doi: 10.2196/25087] [Medline: 33393908]
20. Wekerle C. JoyPop: a resilience intervention for youth mental health. PageSuite. 2021. URL: https://edition.pagesuite-professional.co.uk/html5/reader/production/default.aspx?pubname=&edid=b43e8d51-b6f4-48ca-9e31-d7936ed7a950 [accessed 2025-06-05]
21. Bakker D, Kazantzis N, Rickwood D, Rickard N. Mental health smartphone apps: review and evidence-based recommendations for future developments. JMIR Ment Health 2016 Mar 1;3(1):e7. [doi: 10.2196/mental.4984] [Medline: 26932350]
22. Jadhakhan F, Blake H, Hett D, Marwaha S. Efficacy of digital technologies aimed at enhancing emotion regulation skills: literature review. Front Psychiatry 2022;13:809332. [doi: 10.3389/fpsyt.2022.809332] [Medline: 36159937]
23. Howard AL, Galambos NL, Krahn HJ. Paths to success in young adulthood from mental health and life transitions in emerging adulthood. Int J Behav Dev 2010 Nov;34(6):538-546. [doi: 10.1177/0165025410365803]
24. Schulenberg JE, Sameroff AJ, Cicchetti D. The transition to adulthood as a critical juncture in the course of psychopathology and mental health. Dev Psychopathol 2004;16(4):799-806. [doi: 10.1017/s0954579404040015] [Medline: 15704815]
25. Malik I, Perez A, Toombs E, et al. Female youth and mental health service providers' perspectives on the JoyPop™ app: a qualitative study. Front Digit Health 2023;5:1197362. [doi: 10.3389/fdgth.2023.1197362] [Medline: 37829596]
26. Au-Yeung A, Marfatia D, Beers K, et al. Exploring the feasibility of a mental health application (JoyPop™) for Indigenous youth. Front Psychiatry 2023;14:1269347. [doi: 10.3389/fpsyt.2023.1269347] [Medline: 37867769]
27. Kim K, Au-Yeung A, Dagher D, et al. Exploring the relevance of a psychology-based resilience app (JoyPop™) for Indigenous youth. Child Abuse Negl 2024 Feb;148:106343. [doi: 10.1016/j.chiabu.2023.106343] [Medline: 37451896]
28. Mushquash AR, Pearson ES, Waddington K, et al. User perspectives on a resilience-building app (JoyPop): qualitative study. JMIR Mhealth Uhealth 2021 Jul 8;9(7):e28677. [doi: 10.2196/28677] [Medline: 34255696]
29. Maurer K, Kimyaci M, Konyk K, Wekerle C. Building resilience through daily smartphone app use: results of a pilot study of the JoyPop app with social work students. Front Digit Health 2023;5:1265120. [doi: 10.3389/fdgth.2023.1265120] [Medline: 38053917]
30. MacIsaac A, Mann V, Toombs E, et al. Promoting mental health and wellbeing among post-secondary students with the JoyPop app: study protocol for a randomized controlled trial. Trials 2024;25:576. [doi: 10.1186/s13063-024-08424-y]

XSL•FO
RenderX

31. MacIsaac A, Neufeld T, Malik I, et al. Increasing access to mental health supports for 18- to 25-year-old Indigenous youth with the JoyPop mobile mental health app: study protocol for a randomized controlled trial. JMIR Res Protoc 2025 Jan 30;14:e64745. [doi: 10.2196/64745] [Medline: 39883939]
32. Jacob C, Sezgin E, Sanchez-Vazquez A, Ivory C. Sociotechnical factors affecting patients' adoption of mobile health tools: systematic literature review and narrative synthesis. JMIR Mhealth Uhealth 2022 May 5;10(5):e36284. [doi: 10.2196/36284] [Medline: 35318189]
33. Linardon J, Fuller-Tyszkiewicz M. Attrition and adherence in smartphone-delivered interventions for mental health problems: a systematic and meta-analytic review. J Consult Clin Psychol 2020 Jan;88(1):1-13. [doi: 10.1037/ccp0000459] [Medline: 31697093]
34. Melcher J, Camacho E, Lagan S, Torous J. College student engagement with mental health apps: analysis of barriers to sustained use. J Am Coll Health 2022;70(6):1819-1825. [doi: 10.1080/07448481.2020.1825225] [Medline: 33048626]
35. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. Evid Based Ment Health 2018 Aug;21(3):116-119. [doi: 10.1136/eb-2018-102891] [Medline: 29871870]
36. Alqahtani F, Orji R. Usability issues in mental health applications. In: UMAP'19 Adjunct: Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization: Association for Computing Machinery; 2019:343-348. [doi: 10.1145/3314183.3323676]
37. Alon N, Torous J. Current challenges for evaluating mobile health applications. J Am Med Inform Assoc 2023 Feb 16;30(3):617-624. [doi: 10.1093/jamia/ocac244] [Medline: 36484621]
38. Messner EM, Probst T, O'Rourke T. mHealth applications: potentials, limitations, current quality and future directions. In: Montag C, Baumeister H, editors. Digital Phenotyping and Mobile Sensing: Springer; 2023:321-334. [doi: 10.1007/978-3-030-98546-2_18]
39. Nielsen J. Usability 101: introduction to usability. Nielsen Norman Group. 2012 Jan 3. URL: https://www.nngroup.com/articles/usability-101-introduction-to-usability/ [accessed 2025-06-05]
40. Montagni I, Tzourio C, Cousin T, Sagara JA, Bada-Alonzi J, Horgan A. Mental health-related digital use by university students: a systematic review. Telemed J E Health 2020 Feb;26(2):131-146. [doi: 10.1089/tmj.2018.0316] [Medline: 30888256]
41. Oti O, Pitt I. Online mental health interventions designed for students in higher education: a user-centered perspective. Internet Interv 2021 Dec;26:100468. [doi: 10.1016/j.invent.2021.100468] [Medline: 34703772]
42. Hajesmaeel-Gohari S, Khordastan F, Fatehi F, Samzadeh H, Bahaadinbeigy K. The most used questionnaires for evaluating satisfaction, usability, acceptance, and quality outcomes of mobile health. BMC Med Inform Decis Mak 2022 Jan 27;22(1):22. [doi: 10.1186/s12911-022-01764-2] [Medline: 35081953]
43. Zhou L, Bao J, Setiawan IMA, Saptono A, Parmanto B. The mHealth App Usability Questionnaire (MAUQ): development and validation study. JMIR Mhealth Uhealth 2019 Apr 11;7(4):e11500. [doi: 10.2196/11500] [Medline: 30973342]
44. Brooke J. SUS: a "quick and dirty" usability scale. In: Jordan P, Thomas B, Weerdmeester B, editors. Usability Evaluation in Industry: Taylor & Francis; 1996:189-194.
45. Harrison R, Flood D, Duce D. Usability of mobile applications: literature review and rationale for a new usability model. J Interact Sci 2013;1(1):1. [doi: 10.1186/2194-0827-1-1]
46. Inal Y, Wake JD, Guribye F, Nordgreen T. Usability evaluations of mobile mental health technologies: systematic review. J Med Internet Res 2020 Jan 6;22(1):e15337. [doi: 10.2196/15337] [Medline: 31904579]
47. Kuehnhausen M, Frost VS. Trusting smartphone apps? To install or not to install, that is the question. In: 2013 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA): IEEE; 2013:30-37. [doi: 10.1109/CogSIMA.2013.6523820]
48. Singh K, Drouin K, Newmark LP, et al. Many mobile health apps target high-need, high-cost populations, but gaps remain. Health Aff (Millwood) 2016 Dec;35(12):2310-2318. [doi: 10.1377/hlthaff.2016.0578]
49. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015;3(1):e27. [doi: 10.2196/mhealth.3422]
50. Huckvale K, Nicholas J, Torous J, Larsen ME. Smartphone apps for the treatment of mental health conditions: status and considerations. Curr Opin Psychol 2020 Dec;36:65-70. [doi: 10.1016/j.copsyc.2020.04.008] [Medline: 32553848]
51. Kertz SJ, MacLaren Kelly J, Stevens KT, Schrock M, Danitz SB. A review of free iPhone applications designed to target anxiety and worry. J Technol Behav Sci 2017 Jun;2(2):61-70. [doi: 10.1007/s41347-016-0006-y]
52. Sucala M, Cuijpers P, Muench F, et al. Anxiety: there is an app for that. A systematic review of anxiety apps. Depress Anxiety 2017 Jun;34(6):518-525. [doi: 10.1002/da.22654] [Medline: 28504859]
53. Albrecht UV, von Jan U. mHealth apps and their risks: taking stock. Stud Health Technol Inform 2016;226:225-228. [doi: 10.3233/978-1-61499-664-4-225] [Medline: 27350511]
54. Neary M, Schueller SM. State of the field of mental health apps. Cogn Behav Pract 2018 Nov;25(4):531-537. [doi: 10.1016/j.cbpra.2018.01.002] [Medline: 33100810]
55. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the user version of the Mobile Application Rating Scale (uMARS). JMIR Mhealth Uhealth 2016;4(2):e72. [doi: 10.2196/mhealth.5849]

XSL•FO
RenderX

56. Stach M, Kraft R, Probst T, et al. Mobile Health App Database – a repository for quality ratings of mHealth apps. In: 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS): IEEE; 2020:427-432. [doi: 10.1109/CBMS49503.2020.00087]
57. Ng MM, Firth J, Minen M, Torous J. User engagement in mental health apps: a review of measurement, reporting, and validity. Psychiatr Serv 2019 Jul 1;70(7):538-544. [doi: 10.1176/appi.ps.201800519] [Medline: 30914003]
58. Antezana G, Venning A, Smith D, Bidargaddi N. Do young men and women differ in well-being apps usage? Findings from a randomised trial. Health Informatics J 2022;28(1). [doi: 10.1177/14604582211064825] [Medline: 35128952]
59. Ervasti M, Kallio J, Määttänen I, Mäntyjärvi J, Jokela M. Influence of personality and differences in stress processing among Finnish students on interest to use a mobile stress management app: survey study. JMIR Ment Health 2019 May 13;6(5):e10039. [doi: 10.2196/10039] [Medline: 31094358]
60. Gitlow L. Technology use by older adults and barriers to using technology. Phys Occup Ther Geriatr 2014 Sep;32(3):271-280. [doi: 10.3109/02703181.2014.946640]
61. Smail-Crevier R, Powers G, Noel C, Wang J. Health-related internet usage and design feature preference for e-mental health programs among men and women. J Med Internet Res 2019 Mar 18;21(3):e11224. [doi: 10.2196/11224] [Medline: 30882361]
62. Jabour AM, Rehman W, Idrees S, Thanganadar H, Hira K, Alarifi MA. The adoption of mobile health applications among university students in health colleges. J Multidiscip Healthc 2021;14:1267-1273. [doi: 10.2147/JMDH.S310539] [Medline: 34103927]
63. Venkatesh V, Thong JY, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Q 2012;36(1):157. [doi: 10.2307/41410412]
64. DeYoung CG, Quilty LC, Peterson JB. Between facets and domains: 10 aspects of the Big Five. J Pers Soc Psychol 2007 Nov;93(5):880-896. [doi: 10.1037/0022-3514.93.5.880] [Medline: 17983306]
65. DeYoung CG, Carey BE, Krueger RF, Ross SR. Ten aspects of the Big Five in the personality inventory for DSM-5. Personal Disord 2016 Apr;7(2):113-123. [doi: 10.1037/per0000170] [Medline: 27032017]
66. Ka L, R E, K W, G J, Lje B. Associations between facets and aspects of Big Five personality and affective disorders: a systematic review and best evidence synthesis. J Affect Disord 2021 Jun;288:175-188. [doi: 10.1016/j.jad.2021.03.061]
67. Mortezaei S, Rabiei R, Asadi F, Emami H. Development and usability evaluation of a mHealth application for albinism self-management. BMC Med Inform Decis Mak 2023 Jun 13;23(1):106. [doi: 10.1186/s12911-023-02202-7] [Medline: 37312174]
68. Rezaee R, Ghaffari M, Rabiei R, Kavousi A, Rakhshanderou S. Design and usability evaluation of a mobile application for self-care among Iranian adolescents. BMC Public Health 2024 Mar 25;24(1):892. [doi: 10.1186/s12889-024-18341-z] [Medline: 38528452]
69. Hoffmann A, Faust-Christmann CA, Zolynski G, Bleser G. Gamification of a stress management app: results of a user study. In: Marcus A, Wang W, editors. Design, User Experience, and Usability Application Domains HCII 2019 Lecture Notes in Computer Science: Springer; 2019. [doi: 10.1007/978-3-030-23538-3_23]
70. LeBeau K, Huey LG, Hart M. Assessing the quality of mobile apps used by occupational therapists: evaluation using the user version of the mobile application rating scale. JMIR Mhealth Uhealth 2019 May 1;7(5):e13019. [doi: 10.2196/13019] [Medline: 31066712]
71. Baumel A, Faber K, Mathur N, Kane JM, Muench F. Enlight: a comprehensive quality and therapeutic potential evaluation tool for mobile and web-based eHealth interventions. J Med Internet Res 2017 Mar 21;19(3):e82. [doi: 10.2196/jmir.7270] [Medline: 28325712]
72. Lau N, O'Daffer A, Yi-Frazier JP, Rosenberg AR. Popular evidence-based commercial mental health apps: analysis of engagement, functionality, aesthetics, and information quality. JMIR Mhealth Uhealth 2021 Jul 14;9(7):e29689. [doi: 10.2196/29689] [Medline: 34259639]
73. Huang HY, Bashir M. Users' adoption of mental health apps: examining the impact of information cues. JMIR Mhealth Uhealth 2017 Jun 28;5(6):e83. [doi: 10.2196/mhealth.6827] [Medline: 28659256]
74. Khechine H, Lakhal S, Ndjambou P. A meta - analysis of the UTAUT model: eleven years later. Can J Adm Sci 2016 Jun;33(2):138-152. [doi: 10.1002/cjas.1381]
75. McEachan RRC, Conner M, Taylor NJ, Lawton RJ. Prospective prediction of health-related behaviours with the theory of planned behaviour: a meta-analysis. Health Psychol Rev 2011 Sep;5(2):97-144. [doi: 10.1080/17437199.2010.521684]
76. McEachan R, Taylor N, Harrison R, Lawton R, Gardner P, Conner M. Meta-analysis of the reasoned action approach (RAA) to understanding health behaviors. Ann Behav Med 2016 Aug;50(4):592-612. [doi: 10.1007/s12160-016-9798-4] [Medline: 27169555]
77. School attendance by visible minority and immigrant status: Canada, provinces and territories, census metropolitan areas and census agglomerations with parts. : Statistics Canada; 2022 Nov 30. [doi: 10.25318/9810043301-eng]

Abbreviations

BFAS: Big Five Aspects Scale

XSL•FO
RenderX

MARS: Mobile Application Rating Scale
MAUQ: Mobile Application Usability Questionnaire
MHAD: Mobile Health App Database
uMARS: user version of the Mobile Application Rating Scale

Edited by A Kushniruk; submitted 16.08.24; peer-reviewed by A O'Daffer, J Hu; revised version received 13.04.25; accepted 18.04.25; published 04.07.25.

Please cite as:
Malik I, Neufeld T, Mushquash A
Usability and Quality of the JoyPop App: Prospective Evaluation Study
JMIR Hum Factors 2025;12:e65472
URL: https://humanfactors.jmir.org/2025/1/e65472
doi:10.2196/65472

© Ishaq Malik, Teagan Neufeld, Aislin Mushquash. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 4.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Usability and Usefulness of Occupational Health Care Patient Portals: Patient-Based Cross-Sectional Study

Sari Nissinen, PhD; Pauliina Toivio, MSc; Erja Sormunen, PhD

Finnish Institute of Occupational Health, Topeliuksenkatu 41b, Helsinki, Finland

Corresponding Author:
Sari Nissinen, PhD
Finnish Institute of Occupational Health, Topeliuksenkatu 41b, Helsinki, Finland

Abstract

Background: Patient portals are a crucial part of modern health care services, as they provide patients with access to their personal information and enable communication with health care professionals. The usability and usefulness of these portals are decisive factors in their adoption. There is a lack of previous research on the use of patient portals in occupational health care (OHC).

Objective: This study aims to examine patients' experiences with the usability and usefulness of OHC patient portals and to identify the factors that influence their perceived usability and usefulness.

Methods: A cross-sectional study was conducted through a web-based survey in April 2024 in Finland. Of the 3072 respondents, usability was assessed using 12 statements, and usefulness was evaluated with 9 statements. Responses were collected on a 5-point Likert scale. The survey also gathered respondents' background information, including age, gender, education, information and communication technology (ICT) skills, satisfaction, frequency of use, and data privacy concerns. Data analysis was performed using SPSS Statistics (version 29; IBM Corp), applying frequency analysis and a general linear model.

Results: The results showed that 75.1% (1895/2523) of respondents agreed that the portal was easy to learn and use, 70.5% (1774/2517) felt it supported collaboration with OHC, while 52.4% (1316/2511) reported that it provided a good overview of their work ability, and 35.5% (891/2505) felt it offered a good overview of their working conditions. The perceived usability of OHC patient portals was significantly associated with several factors in the adjusted model: fear of unauthorized access to data (F_4=4.49, P=.001, η^2=0.007), need for guidance (F_4=52.2, P<.001, η^2=0.080), frequency of use (F_1=34.3, P<.001, η^2=0.014), satisfaction with the portal (F_1=577.1, P<.001, η^2=0.193), perceived ICT skills (F_1=12.2, P<.001, η^2=0.005), and age (F_1=10.8, P<.001, η^2=0.004), with younger users (≤50 years) reporting better usability. Perceived usefulness was significantly influenced by frequency of use (F_1=9.80, P=.002, η^2=0.004) and satisfaction (F_1=548.0, P<.001, η^2=0.183), while other factors such as fear of unauthorized access (F_4=0.41, P=.80), need for guidance (F_4=1.52, P=.20), ICT skills, education, gender, and age were not statistically significant.

Conclusions: OHC patient portals must enhance their capacity to provide information on work ability and working conditions. Improved documentation of work-related hazards, risks, and workload factors in medical records is needed. These enhancements could raise patient awareness of work well-being factors.

(JMIR Hum Factors 2025;12:e73197) doi:10.2196/73197

KEYWORDS

patient portal; patient care; occupational health; work ability data; online questionnaire; Finland; descriptive statistics; general linear regression

Introduction

Generally, a patient portal is a website or mobile app that provides patients with access to their personal health information using a username and password. Through the portal, patients can view health records documented by health care professionals, such as prescribed medications, administered vaccinations, identified allergies, and laboratory test results. Additionally, the portal might enable secure communication with health care professionals [1-3].

Usability and usefulness are key factors in the adoption of patient portals [4-7]. Positive experiences are often associated with the portal's usability and useful features, such as the ability to request prescription renewals and view visit documentation [2,6,8]. Real-time access to information is also an essential aspect of a good user experience [5,6]. Some patients find portals easy to use and useful, while others encounter challenges due to complex interfaces [5,9], lack of user training [10], doubts about the portal's usefulness [11,12], and comprehensibility of medical terminology [8]. Concerns about security and privacy

XSL•FO
RenderX

[2,5,12,13] or lack of professional recommendations may also reduce willingness to use portals [1,14].

The impact of patient portal use on health outcomes has been variable. Some studies [15-17] have found no significant effect of portal use on health outcomes. Conversely, some studies have reported improvements in areas such as quality of care [18], health behavior [19], and understanding own health [5,20]. Many other studies indicate that patient portals enhance communication between patients and health care professionals and increase patient engagement in their own care [5,8,21]. Furthermore, patient portals have helped patients identify errors in their health records, enhancing patient safety [5,22,23].

My Kanta is the national patient portal accessible to all citizens in Finland, and citizens are widely using My Kanta, which allows them to view their medical records and prescriptions, as well as check laboratory and radiology results [24,25]. Alongside My Kanta, occupational health care (OHC) providers in Finland have developed patient portals for their clients [26], likely to strengthen customer relationships by offering their own portals [27].

In Finland, legislation mandates that all employers provide preventive OHC services to their employees to prevent work-related health issues and promote employees' health, safety, and work ability. In addition to the mandatory preventive services, employers also have the option to offer primary health care services at the general practitioner level. Thus, a significant user group for OHC portals consists of working-age patients, as around 91% of Finnish employees are covered by OHC services [28].

Similar to My Kanta, OHC patient portals are used to view medical records, prescriptions, vaccinations, and to check laboratory and radiology results. However, OHC portals, available as both a mobile app and a web service, expand on these functions by also allowing appointment booking, viewing sick leave certificates, and the designated OHC team, as well as enabling remote consultations and secure messaging [6,27].

Information in the portal is based on clinical data entered by OHC professionals and automatically transferred from the electronic health record system. This data is presented as read-only to patients. Additionally, patients are able to input their own health questionnaire responses, self-measurements, and treatment records from other health care providers. However, this patient-generated data is not visible to professionals without the patient's explicit consent. Access rights are clearly defined. Clinical data in the portal can be updated by OHC professionals, while patients are only able to manage their own entries. Furthermore, the portal does not display records from other health care providers, such as My Kanta [25].

A previous study has shown that in OHC, a significant amount of work ability data is generated through patient work [29], as a central objective of OHC is to prevent occupational illnesses and accidents, as well as to maintain and promote work ability [30].

Despite the significant role of patient portals in health care, there is a notable gap in research regarding the experiences of OHC patients with these portals. Only 1 study has explored the experiences of patients from a single OHC provider; however, the patients' ages ranged between 56 and 70 years, and the focus was on using patient portals for health promotion [3]. According to the study, patients were willing to use portals if the information is readily accessible and the functionalities are seamless.

Therefore, this study aims to examine the experiences of OHC patients concerning the usability and usefulness of patient portals. We are also interested in identifying the factors that influence the usability and perceived usefulness of these portals, thereby addressing the existing research gap.

The research questions (RQs) are:

RQ1: How do patients perceive the usability and usefulness of OHC patient portals?

RQ2: Which patient factors are associated with the usability and usefulness of OHC patient portals?

Methods

Participants and Data Collection

The study was conducted through an online survey in April 2024. Two organizations participated in the study: a trade union and a patient expert organization, both representing OHC patients. Each organization appointed a contact person who was responsible for distributing the survey. This contact person sent the survey link via email invitation to all members of their organization, with 2 reminders sent to encourage responses. Before the data collection phase, 5 individuals with experience using OHC patient portals reviewed the survey content for readability and comprehensibility. Their feedback was used to prepare the final version of the survey.

The email invitation to the potential respondents explained that this study aimed to gather insights into users' experiences with digital patient portals in OHC to help develop customer-oriented digital services. It also noted that the survey was part of a project by the Finnish Institute of Occupational Health, supported by the Finnish Work Environment Fund. Participation was voluntary, and the estimated survey completion time was provided. The invitation emphasized respondent anonymity, assuring that personal data and organizational affiliations would not be disclosed. It also mentioned that the results would be published in a scientific paper. Additionally, the invitation included contact information, a research information sheet, and a link to the project website.

A total of 3100 individuals completed the survey. Although the exact response rate could not be determined due to the lack of precise information on how many individuals received the survey (we estimated the size of this study's population to comprise approximately 46,000 individuals), it is known that 8372 people opened the survey link. Based on those who opened the link, the response rate can be estimated at approximately 37% (3100/8372).

XSL•FO
RenderX

Ethical Considerations

The Ethics Committee of the Finnish Institute of Occupational Health approved this study on March 1, 2024 (approval number ETR 022024, ID154008), and the study followed the Finnish research ethics practice. All procedures followed were per the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2000.

Participants were not asked to provide written consent for this study. In Finland, completing the survey is generally seen as an indication of consent to participate in the research and to permit the use of the data collected. Every participant received an official cover letter detailing this study's background, emphasizing that participation is entirely voluntary and anonymous, and noting that answering the questionnaire is considered informed consent to join this study. Also, the cover letter included the purpose of this study, the estimated time required to complete the survey, who the researchers are, the data collection methods, and the duration of data storage. The survey included a request for consent to use the responses for scientific research purposes. All responses were collected anonymously, and no identifying information was gathered. Data were stored securely as per institutional data protection policies to ensure participant privacy and confidentiality. Participants did not receive any compensation for their participation.

Measurements

Dependent Variables

The perceived usability of OHC patient portals was assessed using 12 statements developed per attributes of usability: learnability, efficiency, memorability, errors, and satisfaction [31]. The perceived usefulness of OHC patient portals was evaluated with 9 statements, which were based on researchers' experiences with OHC services and also used elements from the Technology Acceptance Model [32]. Responses to these statements were collected using a 5-point Likert scale, categorized as follows: "totally disagree," "somewhat disagree," "neither agree nor disagree," "somewhat agree," and "totally agree."

All survey variables that investigated the usability (12 statements) and usefulness (9 statements) of OHC patient portals were combined into 2 separate sum variables. Before forming these sum variables, the internal consistency of the items was assessed using the Cronbach α coefficient, resulting in values of 0.95 for the usability sum variable and 0.93 for the usefulness sum variable.

Independent Variables

The survey used in this study included questions on the respondents' background information: gender, age, education, information and communication technology (ICT) skills, satisfaction with OHC patient portals, frequency of use, a need for guidance, and fear of unauthorized persons seeing the data. These were used only in statistical analysis as adjustments.

For analyses, the variables were recoded to facilitate interpretation and enhance clarity. Gender was originally categorized as 1=female, 2=male, and 3=other, or do not want to say. For analysis purposes, these categories were recoded into 1=women and 2 - 3=other. Age was initially divided into 5 groups: 1=up to 30 years, 2=31 - 40 years, 3=41 - 50 years, 4=51 - 60 years, and 5=61 years or older. These were recoded into 2 categories: 1 - 3=50 years or younger and 4 - 5=51 years or older. Education was originally classified with response options of 1=basic education degree (primary school), 2=matriculation examination, 3=vocational college, 4=University or University of Applied Sciences, and 5=other. For analysis, these categories were retained except the "other" category, which was excluded. Perceived ICT skills were initially rated as 1=weak, 2=moderate, and 3=good and were recoded into 1=weak and 2 - 3=at least moderate for analysis. The satisfaction variable was originally measured using response options: 1=totally unsatisfied, 2=somewhat unsatisfied, 3=neither agree nor disagree, 4=somewhat satisfied, and 5=totally satisfied. These were recoded into 2 categories for analysis: 1 - 3=unsatisfied and 4 - 5=satisfied. Lastly, variables, need for guidance, and fear of unauthorized persons seeing the data were retained as per the originals. Additionally, the response option "neither agree nor disagree" is referred to as "neutral."

Data Analysis

The survey data were analyzed descriptively using frequency analysis in SPSS Statistics (version 29; IBM Corp) software. This study's respondent group consisted of a total of 3100 participants who answered at least 1 survey question. Of these, 639 reported regular use of the OHC patient portal, 1906 used it randomly, and 539 had never used it. Additionally, 16 did not respond to this question. The respondents' characteristic variables were described by the absolute (n) and relative (%) frequencies. The number of respondents varies depending on the background variable. No imputation was performed for missing data.

Before analyzing this study's data, participants who reported using the OHC patient portal were included. Respondents' (2545/3084) attitude statements regarding the perceived usability and perceived usefulness of OHC patient portals were described by the relative (%) frequencies (RQ1). However, the number of respondents varies because not all participants answered all attitude statements regarding usability and usefulness. Missing data was not imputed.

A general linear model (GLM) was used to examine the adjusted associations of respondent characteristics with the perceived usability and perceived usefulness of OHC patient portals (RQ2). In the GLM, only those respondents who reported using the patient portal were included in the analysis (2545/3084). The number of respondents varies because only those who both used the portal and responded to the relevant question were analyzed. Additionally, the differences between the crude and adjusted models arise because, in the adjusted model, only respondents who provided complete responses to all variables included in the model were analyzed. No imputation was performed for missing data. Descriptive statistics are presented alongside GLM results to aid interpretation of effect directions and group-level differences.

XSL•FO
RenderX

In the crude models, dichotomous variables were not treated as continuous to facilitate easier interpretation and comparison of differences. Using dichotomous variables in their original form allows for more straightforward analysis and clearer understanding of the results. In the adjusted model, dichotomous variables were treated as continuous covariates to enhance the model's stability. A model with numerous categorical variables could become unstable, leading to less reliable results.

Model selection was guided by a combination of theoretical relevance, model fit indices (eg, Akaike Information Criterion), and the principle of parsimony. Variables that did not improve model performance or interpretability were excluded, even if they were statistically significant in univariate analyses.

First, separate analyses (crude model) were conducted to examine the association of each independent variable with the dependent variable. Second, a fully adjusted multivariable model (adjusted model) was performed to measure the joint effects of all the statistically significant independent variables as covariates in the multivariable model. Collinearity diagnostics were performed, and variance inflation factor values were checked to ensure no significant multicollinearity was present among the independent variables. We chose a statistical significance level of <0.001 to apply a stricter criterion for assessing the significance of the results and to reduce the risk of false-positive results because of the large sample size. This helps to obtain more reliable and practically significant results.

Results

Characteristics of Respondents

Regular use of the OHC patient portal was reported by 20.7% (634/3060) of all respondents. A total of 23.3% (102/438) of men and 20.3% (528/2606) of women reported regular use. Among respondents aged 30 years or younger, 23.4% (22/94) reported regular use. Of those who were totally satisfied with the portal, 39.7% (95/239) reported regular use. Respondents with good ICT skills included 23.2% (264/1140) regular users. Among those who totally disagreed with concerns about unauthorized access, 31.2% (170/545) reported regular use. Additionally, 31.9% (273/855) respondents who did not feel a need for additional guidance reported regular use. Respondents' background characteristics are presented in Table 1.

XSL•FO
RenderX

Table . Respondents' characteristics.

Characteristics	Use of the OHC[a] portal by respondents			
	Regularly, n (%)	Occasionally, n (%)	Never, n (%)	Overall
Gender	634 (20.7)	1893 (61.9)	533 (17.4)	3060 (100)
Women	528 (20.3)	1635 (62.7)	443 (17.0)	2606 (100)
Men	102 (23.3)	249 (56.8)	87 (19.9)	438 (100)
Other, or do not want to say	4 (25)	9 (56.3)	3 (18.8)	16 (100)
Age (years)	633 (20.7)	1894 (61.9)	535 (17.5)	3062 (100)
Up to 30	22 (23.4)	49 (52.1)	23 (24.5)	94 (100)
31 - 40	62 (20)	202 (65.2)	46 (14.8)	310 (100)
41 - 50	135 (19.8)	432 (63.3)	116 (17)	683 (100)
51 - 60	311 (21.7)	882 (61.6)	239 (16.7)	1432 (100)
61 or older	103 (19)	329 (60.6)	111 (20.4)	543 (100)
Education	636 (20.7)	1901 (61.9)	535 (17.4)	3072 (100)
Basic education (primary school)	45 (18.8)	137 (57.3)	57 (23.8)	239 (100)
Matriculation examination	32 (23.2)	81 (58.7)	25 (18.1)	138 (100)
Vocational college	512 (20.6)	1551 (62.3)	428 (17.2)	2491 (100)
University or University of Applied Sciences	46 (23.1)	129 (64.8)	24 (12.1)	199 (100)
Other	1 (20)	3 (60)	1 (20)	5 (100)
Information and communication technology skills	634 (20.6)	1899 (61.8)	538 (17.5)	3071 (100)
Weak	26 (12.2)	109 (51.2)	78 (36.6)	213 (100)
Moderate	344 (20)	1070 (62.3)	304 (17.7)	1718 (100)
Good	264 (23.2)	720 (63.2)	156 (13.7)	1140 (100)
Satisfaction with occupational health care patient portals	633 (21.3)	1871 (63)	466 (15.7)	2970 (100)
Totally unsatisfied	14 (11.9)	55 (46.6)	49 (41.5)	118 (100)
Somewhat unsatisfied	40 (16.1)	177 (71.4)	31 (12.5)	248 (100)
Neither dissatisfied nor satisfied	104 (10.5)	579 (58.2)	312 (31.4)	995 (100)
Somewhat satisfied	380 (27.7)	931 (68)	59 (4.3)	1370 (100)
Totally satisfied	95 (39.7)	129 (54)	15 (6.3)	239 (100)
Guidance needed	632 (21.5)	1860 (63.2)	451 (15.3)	2943 (100)
Totally disagree	273 (31.9)	537 (62.8)	45 (5.3)	855 (100)
Somewhat disagree	112 (24.9)	314 (69.8)	24 (5.3)	450 (100)
Neither disagree nor agree	150 (14.7)	608 (59.7)	260 (25.5)	1018 (100)
Somewhat agree	68 (16.9)	276 (68.7)	58 (14.4)	402 (100)
Totally agree	29 (13.3)	125 (57.3)	64 (29.4)	218 (100)
Fear of unauthorized persons seeing the data	630 (21.3)	1872 (63.2)	458 (15.5)	2960 (100)
Totally disagree	170 (31.2)	344 (63.1)	31 (5.7)	545 (100)
Somewhat disagree	147 (23.9)	422 (68.7)	45 (7.3)	614 (100)
Neither disagree nor agree	151 (15.2)	571 (57.6)	270 (27.2)	992 (100)
Somewhat agree	121 (21.5)	382 (67.7)	61 (10.8)	564 (100)

XSL•FO
RenderX

Characteristics	Use of the OHC[a] portal by respondents			
	Regularly, n (%)	Occasionally, n (%)	Never, n (%)	Overall
Totally agree	41 (16.7)	153 (62.4)	51 (20.8)	245 (100)

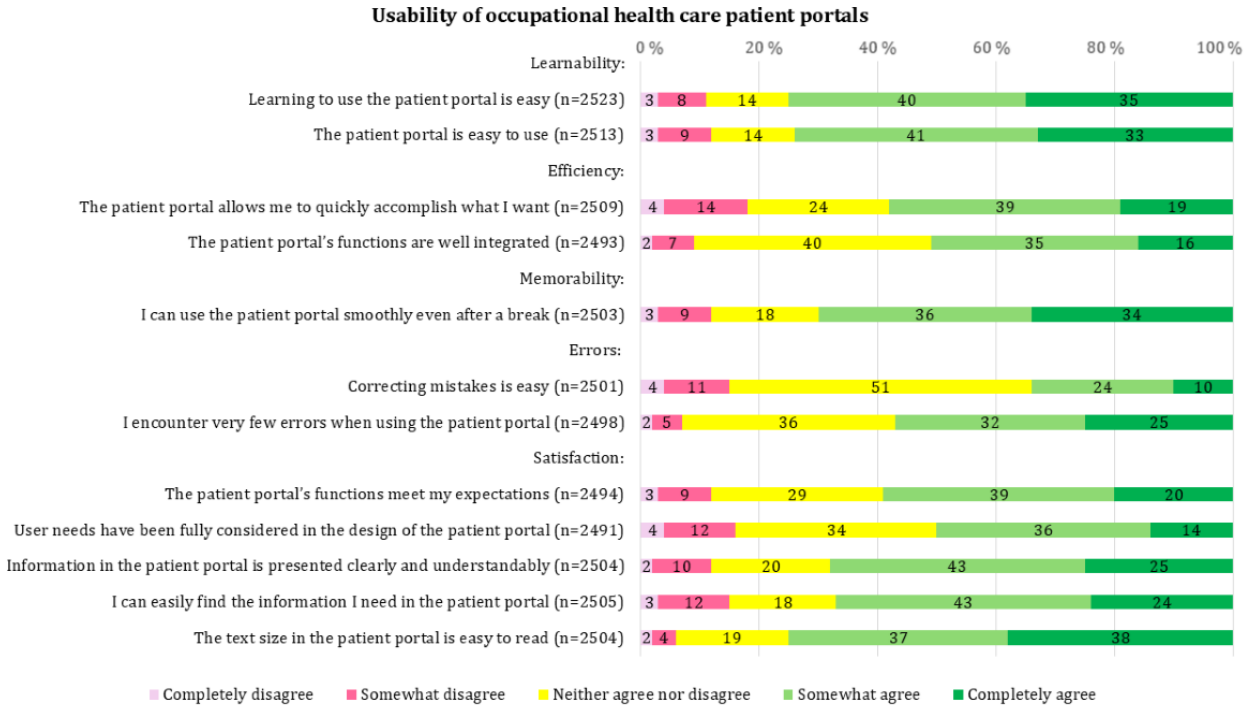
[a]OHC: Occupational Health Care.

The Usability of OHC Patient Portals

According to the results, respondents most positively evaluated aspects related to learnability and ease of use. A large majority (1895/2523, 75.1%) felt that learning to use the patient portal was easy, and most (1861/2513, 74%) experienced the portal as easy to use. Additionally, many respondents (1875/2504, 75%) considered the text size in the portal easy to read. In contrast, the highest levels of disagreement were related to efficiency and user-centered design. Nearly 1 in 5 (451/2509, 18%) respondents disagreed that the portal allows them to quickly accomplish what they want, and 16% (399/2491) disagreed that user needs have been fully considered in the design of the portal. Figure 1 presents respondents' attitudes toward the usability of OHC patient portals, displayed as relative percentages (without decimals).

The analysis of the GLM assessed the impact of various factors on the usability of OHC patient portals. Satisfaction had the strongest association with the perceived usability of OHC patient portals. Need for guidance and frequency of use were also significantly associated, though to a lesser extent. Fear of unauthorized persons seeing the data, perceived ICT skills, and age were statistically significant but had small effects. The full results are presented in Table 2.

Figure 1. Respondents' attitude statements regarding the perceived usability of occupational health care patient portals.



XSL•FO
RenderX

Table . Results of general linear model analysis on perceived usability of occupational health care patient portals: crude and adjusted models.

	Crude model					Adjusted model				
Variable	Patients, n	Mean (SD)	*F* test (df)	*P* value	η^2	Patients, n	Mean (SD)	*F* test (df)	*P* value	η^2
Fear of unauthorized persons seeing the data	2492	—	33.3 (4)	<.001	0.051	2427	—	4.49 (4)	.001	0.007
Totally disagree	513	4.0 (0.90)				499	4.0 (0.89)			
Somewhat disagree	565	3.8 (0.73)				558	3.8 (0.73)			
Neutral	720	3.5 (0.73)				694	3.5 (0.73)			
Somewhat agree	501	3.6 (0.70)				489	3.6 (0.70)			
Totally agree	193	3.7 (1.0)				187	3.7 (1.0)			
Need for guidance	2482	—	107.4 (4)	<.001	0.148	2427	—	52.2 (4)	<.001	0.08
Totally disagree	808	4.1 (0.77)				789	4.1 (0.77)			
Somewhat disagree	425	3.8 (0.69)				418	3.8 (0.68)			
Neutral	754	3.4 (0.70)				734	3.4 (0.69)			
Somewhat agree	343	3.4 (0.68)				338	3.4 (0.68)			
Totally agree	152	3.4 (1.1)				148	3.4 (1.1)			
Usage frequency[ab]	2519	—	110.5 (1)	<.001	0.042	2427	—	34.3 (1)	<.001	0.014
Regularly	635	4.0 (0.77)								
Occasionally	1884	3.6 (0.84)								
Satisfaction[ab]	2391	—	730.9 (1)	<.001	0.227	2427	—	577.1 (1)	<.001	0.193
Dissatisfied	965	3.2 (0.77)								
Satisfied	1526	4.0 (0.70)								
Perceived ICT[c] skills[ab]	2507	—	41.3 (1)	<.001	0.016	2427	—	12.2 (1)	<.001	0.005
Weak	133	3.2 (0.93)								
At least moderate	2374	3.7 (0.82)								
Education	2519	—	2.77 (3)	.04	0.003	—	—	—	—	—
Basic education degree (primary school)	180	3.6 (0.89)								
Matriculation examination	112	3.8 (0.79)								

XSL•FO
RenderX

	Crude model					Adjusted model				
Variable	Patients, n	Mean (SD)	*F* test (df)	*P* value	η^2	Patients, n	Mean (SD)	*F* test (df)	*P* value	η^2
Vocational college	2052	3.7 (0.77)								
University or University of Applied Sciences	175	3.7 (0.77)								
Gender	2501	—	0.95 (1)	.33	0	—	—	—	—	—
Women	2138	3.7 (0.84)								
Other	363	3.6 (0.82)								
Age (years)[ab]	2502	—	35.8 (1)	<.001	0.014	2427	—	10.8 (1)	<.001	0.004
50 or younger	896	3.8 (0.80)								
51 or older	1606	3.6 (0.84)								

[a]As a continuous covariate in the adjusted model; group means not reported.

[b]As a dichotomic variable in the crude model.

[c]ICT: information and communication technology.

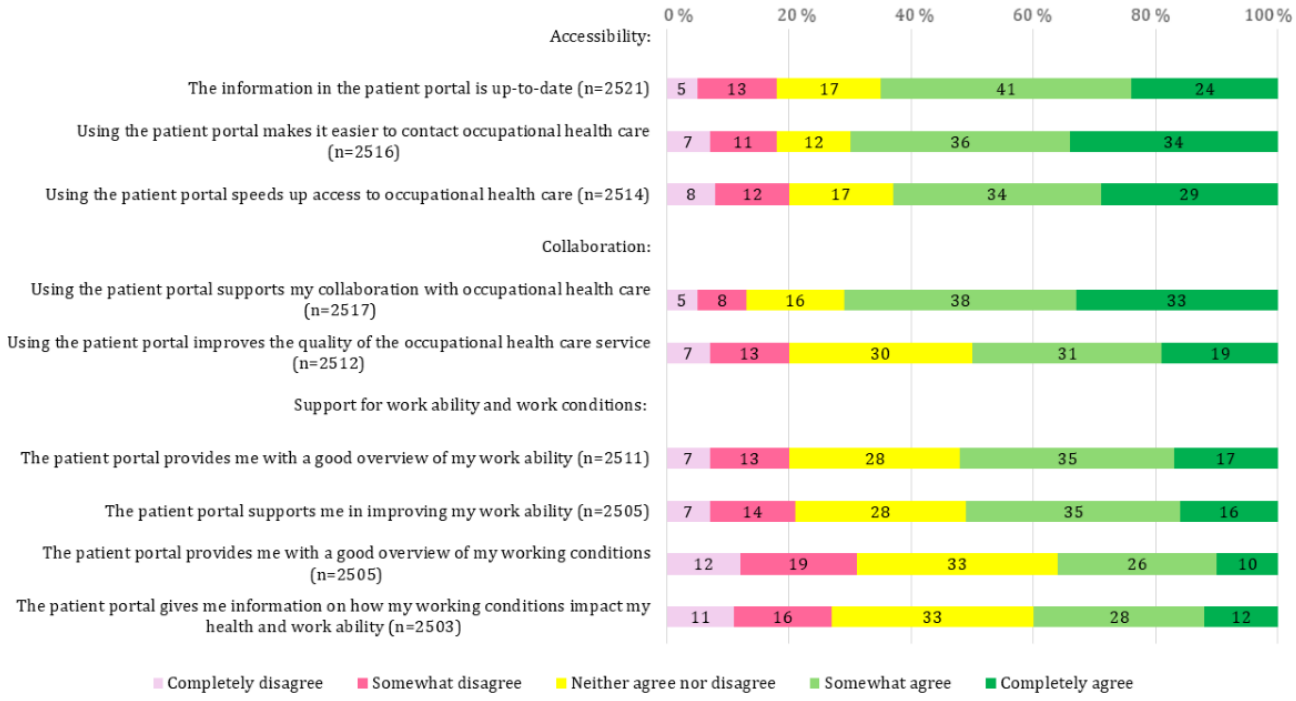
The Usefulness of OHC's Patient Portals

According to the results, the most positively received aspects of the OHC patient portal were its ability to facilitate easier contact with OHC and to support collaboration. A total of 70.1% (1765/2516) of respondents agreed that using the portal makes it easier to contact OHC, and 70.5% (1774/2517) believed that the portal supports their collaboration with OHC. In contrast, only 52.4% (1316/2511) of respondents felt that the portal provides a good overview of their work ability, and just 51.5% (1291/2505) felt that it supports them in improving it. Regarding working conditions, only 35.5% (891/2505) of respondents felt that the portal offers a good overview of their working conditions, and 39.9% (1000/2503) agreed that it provides information about how those conditions impact their health and work ability. Figure 2 presents respondents' attitudes toward the usefulness of OHC patient portals, displayed as relative percentages (without decimals).

According to the GLM analysis, satisfaction had the strongest association with the perceived usefulness of OHC patient portals. Frequency of use was also significantly associated with perceived usefulness, though to a lesser extent. The full results are presented in Table 3.

XSL•FO
RenderX

Figure 2. Respondents' attitude statements regarding the perceived usefulness of occupational health care patient portals.



XSL•FO
RenderX

Table . Results of general linear model analysis on perceived usefulness of occupational health care patient portals: crude and adjusted models.

	Crude model					Adjusted model				
Variable	Patients, n	Mean (SD)	*F* test (df)	*P* value	η^2	Patients, n	Mean (SD)	*F* test (df)	*P* value	η^2
Fear of unauthorized persons seeing the data	2490	—	5.49 (4)	<.001	0.009	2451	—	0.41 (4)	.80	0.001
Totally disagree	513	3.3 (0.96)				503	3.3 (0.95)			
Somewhat disagree	564	3.2 (0.76)				563	3.2 (0.76)			
Neutral	720	3.1 (0.77)				704	3.1 (0.76)			
Somewhat agree	500	3.2 (0.84)				489	3.2 (0.84)			
Totally agree	193	3.3 (1.1)				192	3.3 (1.0)			
Need for guidance	2480	—	6.58 (4)	<.001	0.011	2451	—	1.52 (4)	.20	0.002
Totally disagree	808	3.3 (0.89)				798	3.3 (0.89)			
Somewhat disagree	425	3.2 (0.77)				422	3.2 (0.78)			
Neutral	752	3.1 (0.80)				741	3.1 (0.80)			
Somewhat agree	343	3.2 (0.82)				340	3.2 (0.82)			
Totally agree	152	3.3 (1.1)				150	3.3 (1.1)			
Usage frequency[ab]	2523	—	48.08 (1)	<.001	0.019	2451	—	9.80 (1)	.002	0.004
Regularly	634	3.4 (0.87)								
Occasionally	1889	3.1 (0.86)								
Satisfaction[ab]	2491	—	603.3 (1)	<.001	0.195	2451	—	548.0 (1)	<.001	0.183
Dissatisfied	965	2.7 (0.78)								
Satisfied	1526	3.5 (0.78)								
Perceived ICT[c] skills[a]	2511	—	2.64 (1)	.10	0.001					
Weak	134	3.1 (0.97)								
At least moderate	2377	3.2 (0.87)								
Education	2523	—	1.41 (3)	.24	0.002					
Basic education (primary school)	181	3.3 (0.98)								
Matriculation examination	112	3.3 (0.81)								
Vocational college	2055	3.2 (0.87)								

XSL•FO
RenderX

	Crude model					Adjusted model				
Variable	Patients, n	Mean (SD)	*F* test (df)	*P* value	η^2	Patients, n	Mean (SD)	*F* test (df)	*P* value	η^2
University or University of Applied Sciences	175	3.1 (0.80)								
Gender	2505	—	0.001 (1)	.98	0					
Women	2143	3.2 (0.87)								
Other	362	3.2 (0.86)								
Age[a] (years)	2506	—	2.70 (1)	.1	0.001					
50 or younger	896	3.2 (0.84)								
51 or older	1610	3.6 (0.84)								

[a]As a continuous covariate in the adjusted model; group means not reported.

[b]As a dichotomic variable in the crude model.

[c]ICT: information and communication technology.

Discussion

Principal Findings

This study assessed the usability and usefulness of OHC patient portals using a survey. The analysis focused on respondents who reported using the portal either regularly or occasionally, excluding nonusers. Thus, the results reflect the experiences of actual OHC portal users.

Usability was generally rated positively: respondents highlighted ease of use, readability, and ease of learning as the most favorable aspects. Lower ratings were given to task completion speed and the extent to which user needs were considered in the portal's design. Age and ICT skills were statistically significant factors influencing usability ratings, with respondents aged 50 years or younger and those with strong ICT skills providing more positive assessments.

Usefulness received more mixed feedback: slightly more than half of the respondents felt the portal provided a good overview of their work ability, and just over a third considered it to offer a comprehensive view of their working conditions. Many respondents indicated that the portal did not sufficiently support them in improving their work ability or provide clear information on how working conditions might affect their health and work capacity. In contrast to usability, neither age nor ICT skills significantly influenced how useful the portal was perceived to be.

Additionally, user satisfaction with the OHC patient portal was strongly associated with both its perceived usability and usefulness, with regular users giving it significantly higher ratings than occasional users.

Comparison With Previous Studies

The findings of this study are consistent with previous research that has identified younger users and those with stronger digital skills as perceiving OHC digital services to be more user-friendly and intuitive [33]. Contrary to earlier studies that highlighted the significance of ICT skills, age, and gender in shaping perceptions of usefulness in digital health services [6,34,35], our study found that these factors did not significantly influence how users perceived the usefulness of digital health services. Moreover, the other studies suggest that both older [36] and younger users [37] generally find OHC patient portals useful, particularly for their clear overview of medical history and treatment plans. The results also showed that satisfied portal users valued its usability more, and those who used the portal regularly perceived its usability as better than occasional users. These results align with previous studies on user satisfaction [3,6,38].

Privacy concerns are a common issue with patient portals in Finland [12,13]. Despite these concerns, patients typically do not restrict the use or disclosure of their information [2]. Interestingly, users of OHC patient portals experience fewer privacy concerns (809/2960, 27.3%) compared to users of the national My Kanta portal (1265/3731, 33.9%) [12]. During the pandemic, privacy concerns among My Kanta users increased even more, with up to half of the respondents expressing concern (1923/3823, 50.3%) [13]. This topic warrants further examination, as patient portals are intended to serve as an essential tool for information exchange and communication between health care professionals and patients.

In this study, only slightly over half of the respondents felt that the portal gave a good overview of their work ability, while only a third found it provided a comprehensive view of their working conditions. Patients generally value access to personal health information [3], and when work-related data—such as the connection between diagnoses and occupational factors—is visible in the portal, it can increase awareness of occupational illnesses [39] and help patients understand how their work environment affects their health and work ability [40]. The patient data displayed in OHC patient portals originates from professionals' entries in the medical records. Previous studies

XSL•FO
RenderX

have found that work-related patient data is rarely documented, even when these issues are discussed with the patient [39,41,42].

Implications in Occupational Health

In OHC, work ability refers to an individual's capacity to perform their job, influenced by health, functional capacity, and working conditions [43,44]. OHC visits often focus on evaluating these aspects [45]. To enable OHC patient portals to provide access to this information, the documentation practices must be changed [46,47], particularly toward a more systematic recording of work-related patient data in medical records.

The observed lack of information on work ability and working conditions contrasts with the general expectation that digital services would improve access to information and increase users' awareness of their own situation [5,6]. When patients have access to information about the hazard, risk, and workload factors of their work and work environments, they can better understand the impact of working conditions on their well-being. This information could enable patients to make informed decisions to provide better opportunities to improve work ability, facilitate discussions with supervisors and OHC professionals on needs for work modifications, and promote the implementation of preventive measures.

To enhance the usability of OHC patient portals, it is important to ensure that the information they contain is comprehensive and easily understandable for patients, and that the portal's functions are useful to them. For example, the portal could remind users of upcoming periodic occupational health examinations or other job-related health tests, such as spirometry for those exposed to respiratory hazards in the workplace. Additionally, it could allow users to fill out symptom diaries to investigate symptom variation between work and nonwork periods, as well as exposure diaries to monitor the accumulation of exposure [48].

Work, working conditions, and work ability are central areas of OHC [30]. Additionally, OHC patient portals are widely available to many working citizens [28]. Future research should explore how these digital services can be developed to enhance collaboration between patients and OHC professionals.

Strengths and Limitations

This was a cross-sectional study that examined OHC patients' experiences with the usability and usefulness of patient portals, aiming to identify factors influencing these aspects and address existing research gaps. The theoretical framework of this study provided a solid foundation for designing the questionnaire and studying the usefulness and ease of use of OHC portals. Although the questionnaire used in this study was self-developed, its reliability is supported by the good Cronbach α values of the composite variables measuring usability and usefulness. This indicates that the questionnaire measured these concepts consistently and strengthens the reliability of the results. Thus, the research methods used in this study were well-suited and provided answers to the RQs while generating new information in line with the research objectives. However, this study has both strengths and limitations.

The strength of this study lies in its uniqueness. To our knowledge, no previous study has systematically assessed patient perceptions of the usability and usefulness of OHC patient portals. Another strength is the large sample size, which ensures a substantial number of observations for analysis and makes the data highly representative. This feature increases confidence in the accuracy and generalizability of the results. The third strength is its timeliness, considering the rapid growth of digital health services. Additionally, calculating the effect size was useful in this study as it provided an objective measure of the magnitude of observed differences. Effect size enabled comparisons between different variables and helped identify clinically significant associations.

The first limitation of this study is the dominant participation of 1 gender; thus, the sample is heavily skewed toward women (2606/3060, 85.2%). This gender imbalance may influence the results and limit the applicability of the findings to a broader population. Future research should aim for a more balanced gender representation to enhance generalizability.

Second, the sample likely includes individuals with higher digital literacy, as the survey was conducted online. This may result in an overrepresentation of participants who are comfortable with digital technologies, potentially excluding those with limited ICT skills. Consequently, the findings may not fully capture the experiences and perspectives of individuals with lower digital literacy.

The third limitation is that this study relies on self-reported data collection. However, certain valuable information can only be obtained through surveys, as was the case in this study. The fourth limitation is the lack of available information about the respondents. As participation in this study was anonymous, we do not know the characteristics of those who declined to participate or did not return the questionnaire. Moreover, despite the statistically significant results obtained in this study, caution should be exercised in interpreting the findings due to the low explanatory power of some results.

Conclusions

The ability of OHC patient portals to provide information on the work ability and the healthiness and safety of working conditions needs to be improved. In particular, more comprehensive documentation in medical records related to work-related hazards, risks, and workload factors is needed. This could strengthen the role of portals in promoting work ability and in preventing work-related illnesses and accidents, while also increasing patients' awareness of factors affecting work well-being. Furthermore, improving patient satisfaction with the usability and usefulness of OHC patient portals may lead to increased adoption of portals and, consequently, more equitable access to digital OHC services.

XSL•FO
RenderX

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

References

1. Osborn CY, Mayberry LS, Wallston KA, Johnson KB, Elasy TA. Understanding patient portal use: implications for medication management. J Med Internet Res 2013 Jul 3;15(7):e133. [doi: 10.2196/jmir.2589] [Medline: 23823974]
2. Sääskilahti M, Ojanen A, Ahonen R, Timonen J. Benefits, problems, and potential improvements in a nationwide patient portal: cross-sectional survey of pharmacy customers' experiences. J Med Internet Res 2021 Nov 3;23(11):e31483. [doi: 10.2196/31483] [Medline: 34730542]
3. Valkonen P, Kujala S, Hörhammer I, Savolainen K, Helminen R, Vartia I. Health self-management of older employees: identifying critical peak experiences of a patient portal. Finn J eHealth eWelfare 2023;15(2):114-129. [doi: 10.23996/fjhw.126837]
4. Scheckel B, Schmidt K, Stock S, Redaèlli M. Patient portals as facilitators of engagement in patients with diabetes and chronic heart disease: scoping review of usage and usability. J Med Internet Res 2023 Aug 25;25:e38447. [doi: 10.2196/38447] [Medline: 37624629]
5. Eriksson-Backa K, Hirvonen N, Enwald H, Huvila I. Enablers for and barriers to using My Kanta - a focus group study of older adults' perceptions of the National Electronic Health Record in Finland. Inform Health Soc Care 2021 Dec 2;46(4):399-411. [doi: 10.1080/17538157.2021.1902331] [Medline: 33787438]
6. Simola S, Hörhammer I, Xu Y, et al. Patients' experiences of a national patient portal and its usability: cross-sectional survey study. J Med Internet Res 2023 Jun 30;25:e45974. [doi: 10.2196/45974] [Medline: 37389909]
7. Kujala S, Simola S, Wang B, et al. Benchmarking usability of patient portals in Estonia, Finland, Norway, and Sweden. Int J Med Inform 2024 Jan;181:105302. [doi: 10.1016/j.ijmedinf.2023.105302] [Medline: 38011806]
8. Kujala S, Hörhammer I, Väyrynen A, et al. Patients' experiences of web-based access to electronic health records in Finland: cross-sectional survey. J Med Internet Res 2022 Jun 6;24(6):e37438. [doi: 10.2196/37438] [Medline: 35666563]
9. Hägglund M, Scandurra I. Usability of the Swedish accessible electronic health record: qualitative survey study. JMIR Hum Factors 2022 Jun 23;9(2):e37192. [doi: 10.2196/37192] [Medline: 35737444]
10. Akyirem S, Wagner J, Chen HN, et al. Recommendations to address barriers to patient portal use among persons with diabetes seeking care at community health centers: interview study with patients and health care providers. JMIR Diabetes 2024 Sep 16;9:e58526. [doi: 10.2196/58526] [Medline: 39284181]
11. Vanderhout S, Taneja S, Kalia K, Wodchis WP, Tang T. Patient experiences and perspectives when MyChart is introduced in a large community hospital: mixed methods study. J Med Internet Res 2025 Jan 23;27:e66353. [doi: 10.2196/66353] [Medline: 39847431]
12. Kyytsönen M, Vehko T, Jylhä V, Kinnunen UM. Privacy concerns among the users of a national patient portal: a cross-sectional population survey study. Int J Med Inform 2024 Mar;183:105336. [doi: 10.1016/j.ijmedinf.2023.105336] [Medline: 38183787]
13. Kainiemi E, Vehko T, Kyytsönen M, et al. The factors associated with nonuse of and dissatisfaction with the national patient portal in Finland in the era of COVID-19: population-based cross-sectional survey. JMIR Med Inform 2022 Apr 22;10(4):e37500. [doi: 10.2196/37500] [Medline: 35404831]
14. Moqbel M, Hewitt B, Nah FFH, McLean RM. Sustaining patient portal continuous use intention and enhancing deep structure usage: cognitive dissonance effects of health professional encouragement and security concerns. Inf Syst Front 2022;24(5):1483-1496. [doi: 10.1007/s10796-021-10161-5] [Medline: 34177359]
15. Griffin A, Skinner A, Thornhill J, Weinberger M. Patient portals: who uses them? What features do they use? And do they reduce hospital readmissions? Appl Clin Inform 2016;7(2):489-501. [doi: 10.4338/ACI-2016-01-RA-0003] [Medline: 27437056]
16. Aljabri D, Dumitrascu A, Burton MC, et al. Patient portal adoption and use by hospitalized cancer patients: a retrospective study of its impact on adverse events, utilization, and patient satisfaction. BMC Med Inform Decis Mak 2018 Jul 27;18(1):70. [doi: 10.1186/s12911-018-0644-4] [Medline: 30053809]
17. Dumitrascu AG, Burton MC, Dawson NL, et al. Patient portal use and hospital outcomes. J Am Med Inform Assoc 2018 Apr 1;25(4):447-453. [doi: 10.1093/jamia/ocx149] [Medline: 29300961]
18. Lear R, Freise L, Kybert M, Darzi A, Neves AL, Mayer EK. Perceptions of quality of care among users of a web-based patient portal: cross-sectional survey analysis. J Med Internet Res 2022 Nov 17;24(11):e39973. [doi: 10.2196/39973] [Medline: 36394922]
19. Huang J, Chen Y, Landis JR, Mahoney KB. Difference between users and nonusers of a patient portal in health behaviors and outcomes: retrospective cohort study. J Med Internet Res 2019 Oct 7;21(10):e13146. [doi: 10.2196/13146] [Medline: 31593546]

XSL•FO
RenderX

20. Kinney AP, Sankaranarayanan B. Effects of patient portal use on patient satisfaction: survey and partial least squares analysis. J Med Internet Res 2021 Aug 27;23(8):e19820. [doi: 10.2196/19820] [Medline: 34448712]
21. Zanaboni P, Kummervold PE, Sørensen T, Johansen MA. Patient use and experience with online access to electronic health records in norway: results from an online survey. J Med Internet Res 2020 Feb 7;22(2):e16144. [doi: 10.2196/16144] [Medline: 32031538]
22. Wang B, Kristiansen E, Fagerlund AJ, et al. Users' experiences with online access to electronic health records in mental and somatic health care: cross-sectional study. J Med Internet Res 2023 Dec 25;25:e47840. [doi: 10.2196/47840] [Medline: 38145466]
23. Freise L, Neves AL, Flott K, et al. Assessment of patients' ability to review electronic health record information to identify potential errors: cross-sectional web-based survey. JMIR Form Res 2021 Feb 26;5(2):e19074. [doi: 10.2196/19074] [Medline: 33635277]
24. Jormanainen V. Over 89% adoption rate of the nationwide online patient portal in Finland. Stud Health Technol Inform 2022 May 25;294:589-593. [doi: 10.3233/SHTI220537] [Medline: 35612157]
25. Jormanainen V, Lindgren M, Keskimäki I, Kaila M. Use of My Kanta in Finland 2010-2022. Stud Health Technol Inform 2023 Jun 29;305:448-451. [doi: 10.3233/SHTI230528] [Medline: 37387062]
26. Vehko T, Kaihlanen AM, Kainiemi E, Kyytsönen M, Nadav J, Saukkonen P. Suomalaisten hyvinvointi [Article in Finnish]. : Helsinki: Terveyden ja hyvinvoinnin laitos; 2022 URL: https://urn.fi/URN:ISBN:978-952-343-996-2 [accessed 2025-07-17]
27. Ruotanen R, Kangas M, Tuovinen T, Keränen N, Haverinen J, Reponen J. Finnish e-health services intended for citizens – national and regional development. FinJeHeW 2021;13(3):283-301. [doi: 10.23996/fjhw.109778]
28. Occupational health care statistics of kela 2023 [Article in Finnish]. Official Statistics of Finland. 2025. URL: https://urn.fi/URN:NBN:fi-fe2025061668492 [accessed 2025-07-17]
29. Nissinen SP, Soini S, Hakulinen H. Kirjatun työkykytiedon tärkeys ja hyödyllisyys työterveyshuollossa – kyselytutkimus työterveyshuollon ammattilaisille [Article in Finnish]. FinJeHeW 2021;13(1):19-31. [doi: 10.23996/fjhw.99513]
30. Translation from Finnish legally binding only in Finnish and Swedish. Occupational Healthcare Act 1383/2001. : Finlex; 2021 URL: https://finlex.fi/api/media/statute-foreign-language-translation/212871/mainPdf/main.pdf?timestamp=2001-12-21T00%3A00%3A00.000Z [accessed 2025-07-30]
31. Nielsen J. Usability Engineering: San Diego: Academic Press; 1993.
32. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319-340. [doi: 10.2307/249008]
33. Nissinen S, Pesonen S, Toivio P, Sormunen E. Exploring the use, usefulness and ease of use of digital occupational health services: a descriptive correlational study of customer experiences. Digit Health 2024;10:20552076241242668. [doi: 10.1177/20552076241242668] [Medline: 38601187]
34. Woods SS, Forsberg CW, Schwartz EC, et al. The association of patient factors, digital access, and online behavior on sustained patient portal use: a prospective cohort of enrolled users. J Med Internet Res 2017 Oct 17;19(10):e345. [doi: 10.2196/jmir.7895] [Medline: 29042345]
35. Clarke MA, Fruhling AL, Lyden EL, Tarrell AE, Bernard TL, Windle JR. The role of computer skills in personal health record adoption among patients with heart disease: multidimensional evaluation of users versus nonusers. JMIR Hum Factors 2021 Jul 26;8(3):e19191. [doi: 10.2196/19191] [Medline: 34309574]
36. Huvila I, Rexhepi H, Moll J, et al. Affordance trajectories and the usefulness of online records access among older adults in Sweden. Digit Health 2024;10:20552076241287354. [doi: 10.1177/20552076241287354] [Medline: 39444731]
37. Hagström J, Blease C, Scandurra I, et al. Adolescents' reasons for accessing their health records online, perceived usefulness and experienced provider encouragement: a national survey in Sweden. BMJ Paediatr Open 2024 Mar 9;8(1):e002258. [doi: 10.1136/bmjpo-2023-002258] [Medline: 38460965]
38. Yin R, Law K, Neyens D. Examining how internet users trust and access electronic health record patient portals: survey study. JMIR Hum Factors 2021 Sep 21;8(3):e28501. [doi: 10.2196/28501] [Medline: 34546182]
39. Soini S, Ryynänen KR, Nissinen S, Miettunen J, Ala-Mursula L. Assessing the work relatedness of diagnoses in occupational health primary care appointments: a 3-year review of electronic medical records. Occup Environ Med 2025 Jul 9;82(4):176-182. [doi: 10.1136/oemed-2024-109991] [Medline: 40393725]
40. Nissinen SP, Soini S, Tarvainen K, Kangas P, Leino T. Työterveyshuollon kirjaamiskäytännöt sairauden liittymisestä työhön ja vaikutuksesta työkykyyn [Article in Finnish]. FinJeHeW 2021;13(1):7-18. [doi: 10.23996/fjhw.95610]
41. Weiste E, Vehviläinen S, Leino T, Laitinen J. Ohjauksen vuorovaikutuskäytänteet työterveystarkastuksissa: vertailu ennen koulutusinterventiota ja sen jälkeen [Article in Finnish]. Psykologia 2018;53(5-6):328-344. [doi: 10.62443/psykologia.v53i5-6.70050]
42. Lappalainen K, Leino T, Rautio M, Nissinen S. An evaluation of personal health plans in occupational health check-ups. Int J Occup Health Public Health Nurs 2020;6(1):1-4 [FREE Full text]
43. Ilmarinen J. Work ability--a comprehensive concept for occupational health research and prevention. Scand J Work Environ Health 2009 Jan;35(1):1-5. [doi: 10.5271/sjweh.1304] [Medline: 19277432]
44. Tengland PA. The concept of work ability. J Occup Rehabil 2011 Jun;21(2):275-285. [doi: 10.1007/s10926-010-9269-x] [Medline: 21052807]

XSL•FO
RenderX

45. Ikonen A. Primary care visits in the Finnish occupational health services and their connections to prevention and work-related factors. : Studies in Social Security and Health; 2012 URL: https://helda.helsinki.fi/server/api/core/bitstreams/14435da3-446a-462b-b1a4-9cb662595e89/content [accessed 2025-07-17]
46. Meier-Diedrich E, Lyckblad C, Davidge G, et al. Impact of patient online record access on documentation: scoping review. J Med Internet Res 2025 Feb 20;27:e64762. [doi: 10.2196/64762] [Medline: 39977853]
47. Kariotis T, Prictor M, Gray K, Chang S. Patient-accessible electronic health records and information practices in mental health care contexts: scoping review. J Med Internet Res 2025 Feb 7;27:e54973. [doi: 10.2196/54973] [Medline: 39918859]
48. Fazen LE, Martin BE 4th, Isakari M, et al. Occupational electronic health records: recommendations for the design and implementation of information systems in occupational and environmental medicine practice-ACOEM Guidance Statement. J Occup Environ Med 2024 Nov 1;66(11):e614-e627. [doi: 10.1097/JOM.0000000000003236] [Medline: 39378367]

Abbreviations

GLM: general linear model
ICT: information and communication technology
OHC: occupational health care
RQ: research question

Edited by A Kushniruk; submitted 27.02.25; peer-reviewed by H Kitamura, V Jormanainen; revised version received 17.06.25; accepted 26.06.25; published 14.08.25.

Please cite as:
Nissinen S, Toivio P, Sormunen E
Usability and Usefulness of Occupational Health Care Patient Portals: Patient-Based Cross-Sectional Study
JMIR Hum Factors 2025;12:e73197
URL: https://humanfactors.jmir.org/2025/1/e73197
doi:10.2196/73197

© Sari Nissinen, Pauliina Toivio, Erja Sormunen. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 14.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Comprehensive Profiling System Integrating Myers-Briggs Type Indicator (MBTI) and Dominance, Influence, Steadiness, and Conscientiousness (DISC) for Personalized Health Training: Correlational Analysis and Usability Evaluation

Donghyun Kim[1], MS; Dong Hun Lee[2], MS; Mi Kyung Hwang[3], PhD

[1]Department of Industrial and Systems Engineering, Dongguk University, Seoul, Republic of Korea

[2]KOI Healthcare Co., Ltd., Seoul, Republic of Korea

[3]Department of Health and Medical Administration, Tongwon University, 268, Jungang-ro,, Gwangju-si, Gyeonggi-do, Republic of Korea

Corresponding Author:
Mi Kyung Hwang, PhD
Department of Health and Medical Administration, Tongwon University, 268, Jungang-ro,, Gwangju-si, Gyeonggi-do, Republic of Korea

Abstract

Background: This study proposes an integrated approach to developing personalized health behavior change programs by combining personality traits and behavior types. Existing tools, such as Myers-Briggs Type Indicator (MBTI) and Dominance, Influence, Steadiness, and Conscientiousness (DISC), have limitations: MBTI reflects internal tendencies but lacks behavioral insights, while DISC highlights behavior but overlooks deeper personality aspects. To address these gaps, the study integrates MBTI and DISC to create a comprehensive profiling system.

Objective: The goal of this research is to design a novel profiling system that merges MBTI and DISC for personalized health management. This system aims to link personality traits with behavior patterns to enhance the effectiveness of tailored health behavior change programs.

Methods: The study involved 3 phases: administering MBTI and DISC tests to 130 participants to analyze correlations, developing an integrated survey for health behavior analysis, and testing its usability with 20 experts for validation.

Results: Significant correlations were observed between MBTI and DISC indicators, including a notable negative correlation between Thinking-Feeling (T/F) and Dominance (D), suggesting an inverse relationship between decision-making preferences and assertiveness. Usability testing results indicated high participant satisfaction, with an average SUS (System Usability Scale) score of 86.0. The SUS is a widely used questionnaire for measuring subjective assessments of usability. This score exceeded industry benchmarks for system usability. Expert evaluations further reinforced the system's practical applicability, highlighting its potential to enhance user engagement through personalized behavioral insights.

Conclusions: This study presents a combined MBTI and DISC profiling system, offering both theoretical insights and practical tools for health behavior change programs. Future research should validate its effectiveness with larger samples and explore broader applications in various health domains.

(JMIR Hum Factors 2025;12:e73397) doi:10.2196/73397

KEYWORDS

health behavior change; personalized programs; integrated profiling; personality types; Myers-Briggs Type Indicator; MBTI; Dominance, Influence, Steadiness, and Conscientiousness; DISC; personality assessment

Introduction

Overview

In modern society, health behavior change has become increasingly important. According to the World Health Organization, adopting healthy lifestyle habits is essential for preventing chronic diseases such as obesity and cardiovascular disorders, as well as for improving overall quality of life [1,2]. While health-promoting activities, such as exercise and dietary management, can yield positive outcomes, standardized approaches that fail to account for individual personality traits and behavioral tendencies often face limitations in sustainability and effectiveness [3,4]. For instance, the same exercise program may feel burdensome to individuals with introverted tendencies but serve as a motivating factor for those with extroverted traits [5,6].

XSL•FO
RenderX

Several personality assessment models, including the Myers-Briggs Type Indicator (MBTI), Dominance, Influence, Steadiness, and Conscientiousness (DISC), Big Five, HEXACO, and Sixteen Personality Factor (16PF) Questionnaire, are widely used to analyze personality and behavior in fields such as psychology and health sciences. The MBTI categorizes personality into 16 types based on 4 dichotomies: Extraversion-Introversion, Sensing-Intuition, Thinking-Feeling, and Judging-Perceiving [7,8]. DISC focuses on 4 behavioral dimensions—Dominance, Influence, Steadiness, and Conscientiousness—emphasizing interpersonal interactions. The Big Five (OCEAN) assesses Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism, with Conscientiousness and Neuroticism predicting health regimen adherence and emotional barriers. HEXACO adds Honesty-Humility to similar dimensions, offering insights into ethical health decisions. The 16PF Questionnaire measures 16 factors, such as emotional stability, refining health intervention strategies. The MBTI provides cognitive insights through its 16 personality types but lacks behavioral specificity. DISC focuses on 4 behavioral dimensions, Dominance, Influence, Steadiness, and Conscientiousness, but overlooks deeper psychological motivations. The Big Five measures broad traits such as Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism, with Conscientiousness and Neuroticism predicting health regimen adherence and emotional barriers, yet it lacks detailed granularity. HEXACO adds Honesty-Humility, offering insights into ethical health decisions, but is less established. The 16PF Questionnaire refines health intervention strategies through 16 factors, such as emotional stability, but can be complex to administer.

We examined various analytical tools, and while each operates within its unique framework, no research has been identified that provides specialized solutions for complex and integrative domains, such as health behavior change programs [9].

This study proposes a novel approach that combines the strengths of MBTI and DISC to overcome these limitations. By integrating the analysis of individual personality traits and behavioral tendencies, the study aims to design personalized programs for health behavior change. Specifically, it seeks to identify correlations between the 2 tools and develop a methodology that maximizes the effectiveness of health behavior change programs, contributing both practical and academic value.

Literature Review on MBTI and DISC and Their Integration

The MBTI, developed in the 1940s by Catherine Cook Briggs and Isabel Briggs Myers, is a psychological tool designed to understand and explain individual personality types [10]. It categorizes personalities into 16 types based on cognitive preferences and information processing styles, using 4 dichotomies: Extraversion (E)–Introversion (I), Sensing (S)–Intuition (N), Thinking (T)–Feeling (F), and Judging (J)–Perceiving (P). MBTI is widely used to understand interpersonal relationships, decision-making styles, and job performance [10]. However, its focus on static personality classifications can limit its ability to reflect situational flexibility [9].

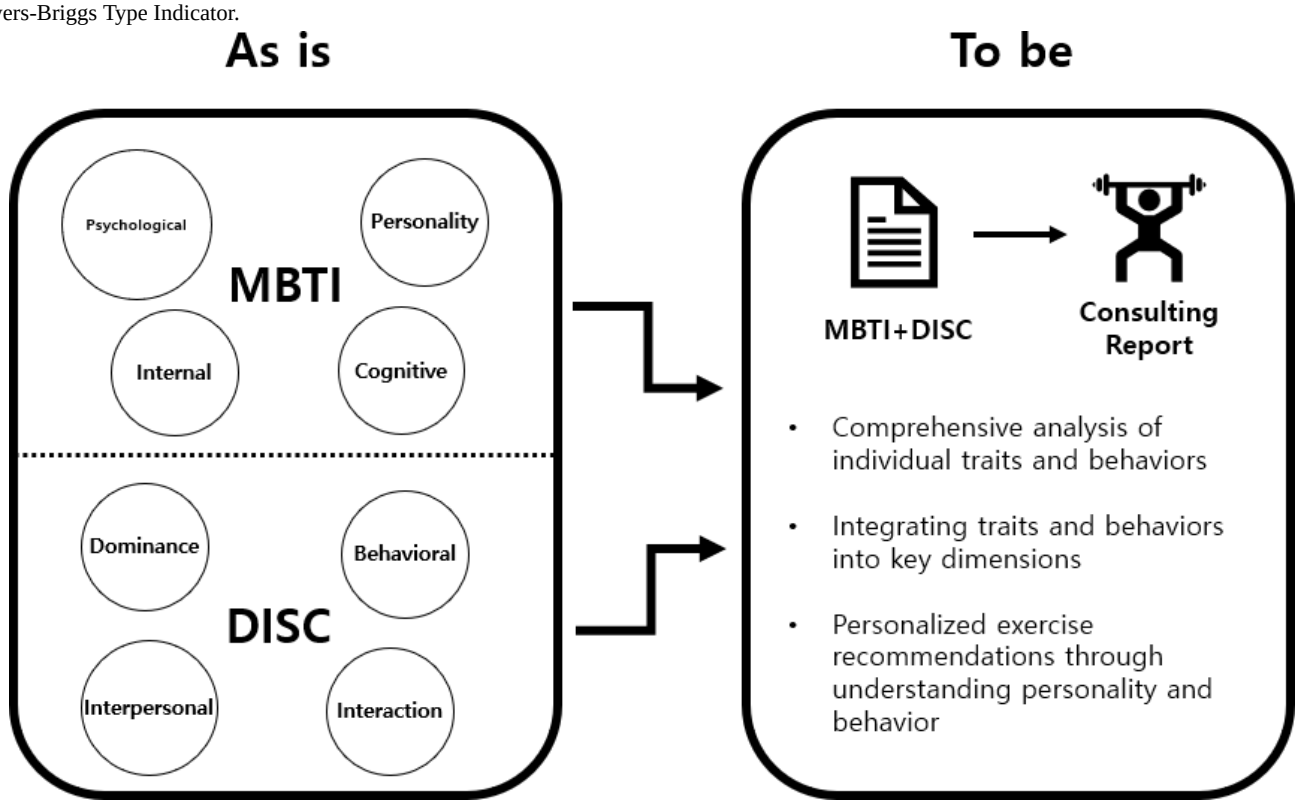
DISC, developed from the theories of William Moulton Marston in the 1920s, analyzes behavioral styles based on 4 dimensions: Dominance (D), Influence (I), Steadiness (S), and Compliance (C). It is especially effective in understanding interpersonal dynamics, improving teamwork, and managing workplace interactions [11]. While DISC excels in analyzing external behavior and interaction patterns, it lacks the depth to fully capture internal psychological factors [12].

Although MBTI emphasizes psychological preferences and DISC focuses on external behaviors, both tools are valuable for understanding individual tendencies [13]. Integrating their strengths allows for a comprehensive understanding of both psychological and behavioral dimensions. This integrated approach is particularly suited for addressing complex issues such as health behavior change, where both internal and external factors play critical roles [14].

Prior to initiating this study, we recognized that the cognitive insights of the MBTI and the behavioral analysis of DISC are essential for health behavior change programs. However, no related studies combining MBTI and DISC to analyze health consulting or behavior change were identified. Accordingly, this study proposes an integration of the 2 models, aiming to analyze correlations between their core dimensions and develop a novel profiling system tailored for designing personalized health behavior change programs (Figure 1).

XSL•FO
RenderX

Figure 1. The methodology for generating consulting reports integrates insights from both MBTI and DISC frameworks. The MBTI model contributes psychological, internal, and personality-based dimensions (eg, personality and cognitive preferences), while the DISC model incorporates behavioral traits such as dominance, interaction styles, and action-oriented tendencies. DISC: Dominance, Influence, Steadiness, and Conscientiousness; MBTI: Myers-Briggs Type Indicator.



Methods

Verification of the Correlation Between MBTI and DISC

The purpose of this study is to systematically verify the correlation between MBTI and DISC by administering both surveys to the same group of participants. Through this process, the research aims to establish a foundation for an integrated understanding of personality and behavior and to provide fundamental data for developing an integrated personality analysis model.

The correlation analysis between MBTI and DISC centers on clearly identifying the similarities and differences that each tool provides and proposing a new approach that can leverage the strengths and compensate for the weaknesses of both tools. This study also explores the possibility of overcoming the limitations of existing personality assessment tools and developing a more efficient and practical personality profiling system.

For this research, 4 elements were set as the main criteria: interpersonal relationships, information-gathering style, decision-making style, and planning tendency. These criteria were selected to evaluate the potential applicability of this approach to health behavior change strategies, such as exercise and diet coaching. For example, the Dominance (D) type in DISC is predicted to have a high correlation with the Extroversion (E) and Thinking (T) traits in MBTI, suggesting that it may be suitable for goal-oriented methods. Meanwhile, the Influence (I) type is associated with the Feeling (F) trait, indicating a preference for a more friendly and motivational approach.

This study investigated the correlation between MBTI and DISC by surveying 130 Korean students. Demographic data (eg, age and gender) of the participants were collected to strengthen the validity of personality trait analysis, with detailed results summarized in Table 1.

XSL•FO
RenderX

Table . Demographic analysis of the survey subjects.

Variables and categories	Frequency, n (%)
Year of birth	
2004-2000	100 (76.9)
1999 - 1990	19 (14.6)
1989 - 1980	11 (8.5)
Total	130 (100)
Sex	
Man	38 (29.2)
Woman	92 (70.8)
Total	130 (100)

To analyze personality types using MBTI and DISC, each tool's official survey was administered to ensure reliable responses from the students. Based on the survey results, this study examined the correlation between the 2 personality assessments and identified patterns in traits that appeared repeatedly. Graphs and charts were used to visualize these correlations and patterns, allowing for a clearer analysis of the relationship between MBTI and DISC.

Refinement of an Integrated Trait Analysis Methodology and Proposed Survey Items

The aim of this study is to develop an integrated personality analysis model that concisely and practically uses personality traits by examining the correlations between the 4 DISC types and the 16 MBTI types.

DISC categorizes human behavioral styles into 4 types: Dominance (D) for goal-oriented and challenging individuals; Influence (I) for optimistic and social people; Steadiness (S) for calm and patient individuals; and Conscientiousness (C) for analytical and logical thinkers.

MBTI explains psychological preferences in perceiving and judging the world through 4 dichotomies. The 'Attitude' (Extraversion/Introversion) dichotomy relates to how we engage with others during exercise; for instance, questions gauge if one easily socializes in new exercise groups (Extraversion) or prefers solitary workouts for focus (Introversion). The "Perceiving Function" (Sensing/Intuition) focuses on how we gather information when choosing an exercise, assessing a preference for detailed rules (Sensing) versus imagining overall potential (Intuition). The "Judging Function" (Thinking/Feeling) examines how we make exercise decisions, involving questions about prioritizing objective data (Thinking) versus considering emotional satisfaction (Feeling). Finally, "Lifestyle" (Judging/Perceiving) looks at how we plan exercise, determining whether one strictly adheres to plans (Judging) or prefers flexible, spontaneous changes (Perceiving).

Based on the key indicators of these existing tools, this study will categorize core traits into 3 - 8 types to design a model suitable for personalized coaching programs and behavior change strategies. We developed 40 survey items based on DISC and MBTI correlations, covering interpersonal relationships, information-gathering, decision-making, and planning tendencies, all adapted to specific contexts such as exercise and dieting. This survey will collect individual trait data, serving as a valuable source for designing customized health behavior programs.

Customer Health Care Profiling Application Based on Survey Items

An expert group of 20 was selected for analysis. This sample was chosen to evaluate the practicality and validity of the integrated trait analysis model. Based on the survey results, a profiling system was developed to provide tailored health care advice and personality descriptions suited to each customer's traits.

In developing the system, the 20 participants' data were used to review the validity of the integrated model and 2 experts were consulted to assess its practicality.

This system can be used in various health care settings such as hospitals, health and fitness centers, and wellness counseling services, with an emphasis on delivering services, products, and personalized coaching optimized for each individual's traits. For example, depending on the customer's personality type, the system can propose effective exercise plans, diet methods, and stress management strategies, providing professional and customized health management solutions.

Ethical Considerations

All information was collected anonymously. The data collected pertained to MBTI and DISC profiles, which are not individually identifiable and do not constitute sensitive personal information that could pinpoint individuals. At the time of the survey, participants were informed about the purpose of the study and their right to withdraw from participation at any time. Therefore, this study was deemed not to fall within the scope requiring institutional review board review. Specifically, this study was exempt from ethics review under Article 13(2) of the Enforcement Regulations of the Bioethics and Safety Act of Korea. This regulation exempts studies where researchers do not directly face identifiable individuals, and where "sensitive information" as defined by Article 23 of the Personal Information Protection Act is neither collected nor recorded.

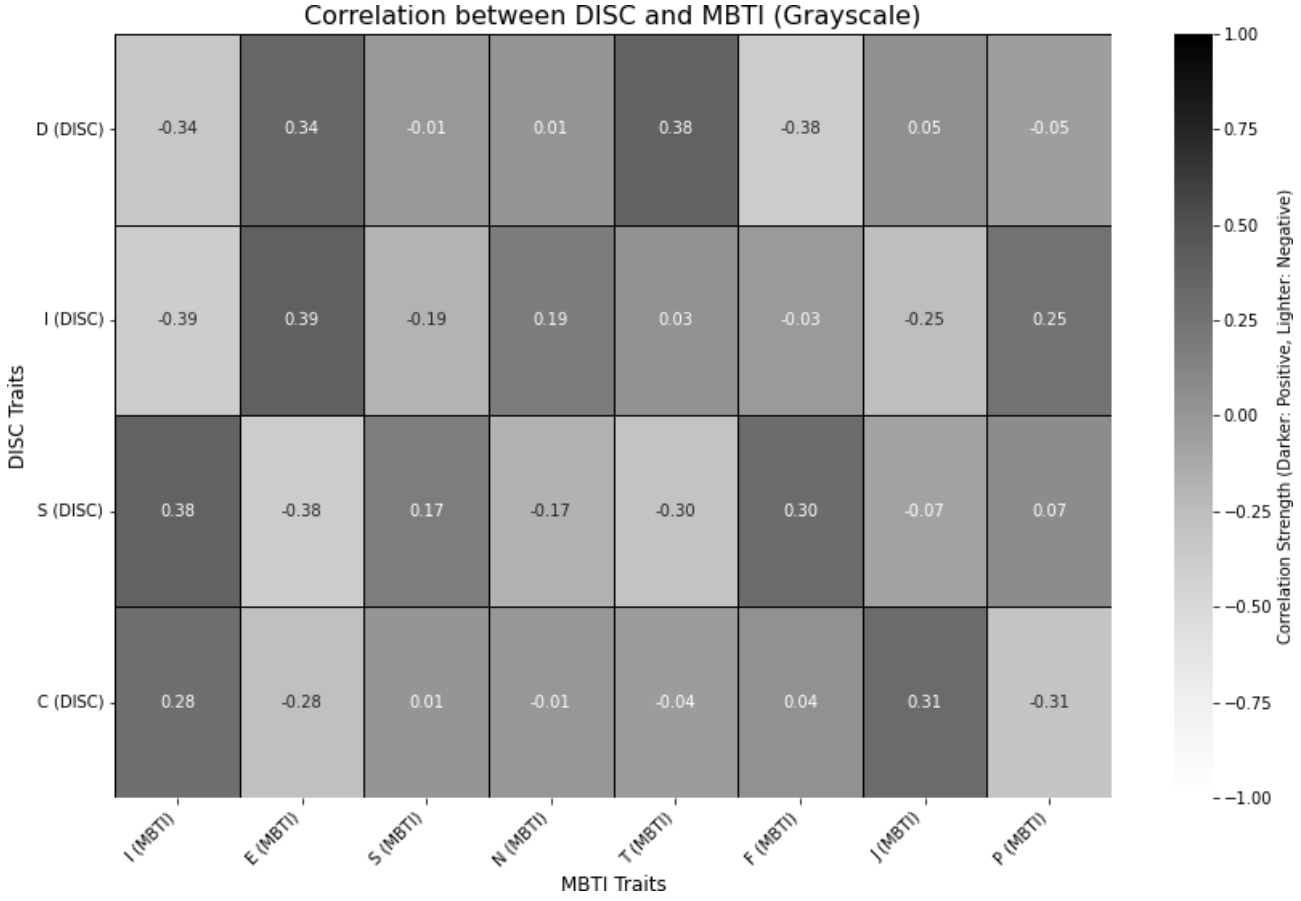
XSL•FO
RenderX

Results

Analysis of the Correlation Between MBTI and DISC

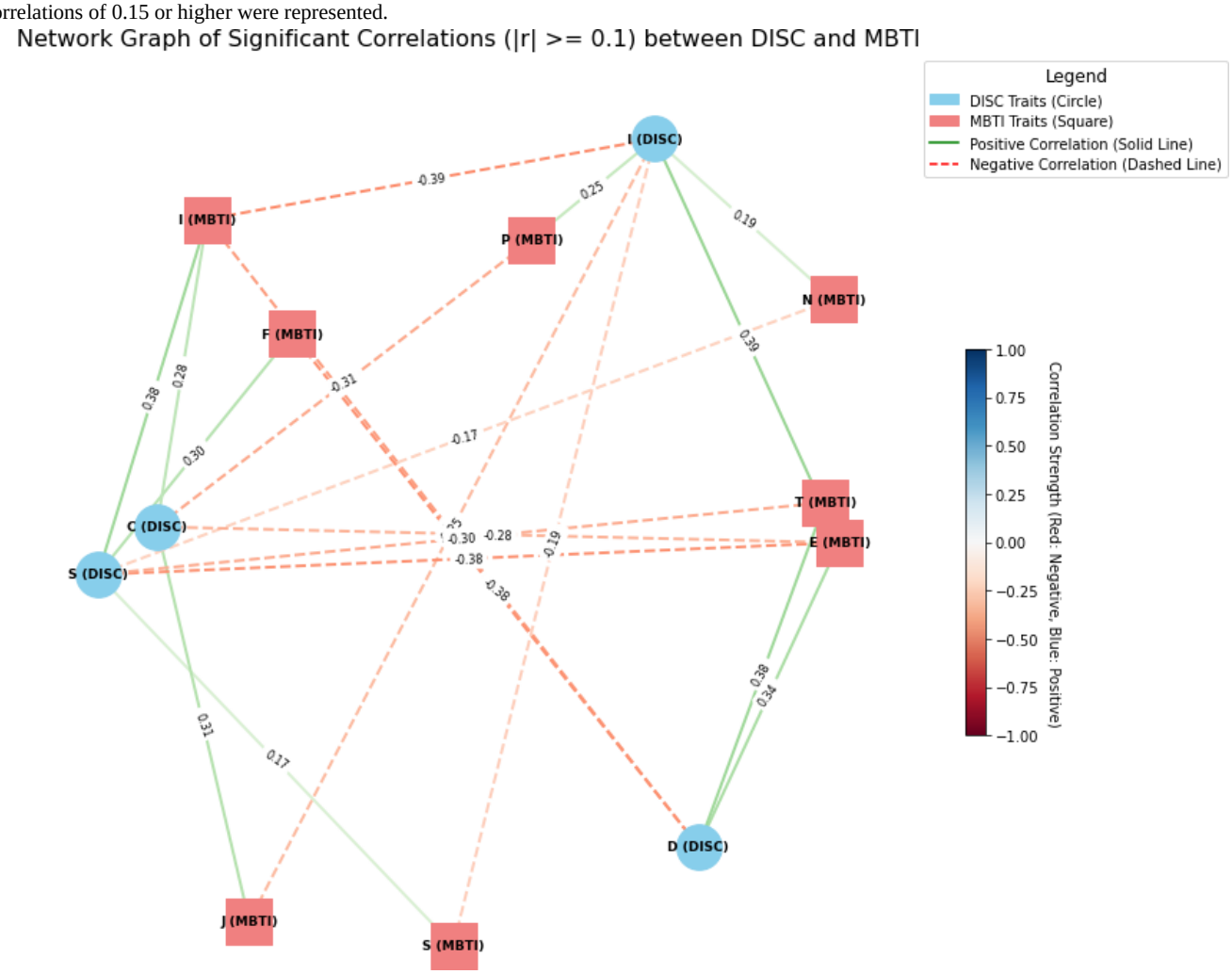
The correlation analysis results for identifying the relationship between the 2 personality assessment tools, MBTI and DISC, are presented in Figures 2 and 3.

Figure 2. Summary of the analysis results of the correlation between MBTI and DISC, derived from the comparison table. Darker shades indicate stronger positive correlations and lighter shades represent stronger negative correlations. Only values with an absolute value of 0.15 or higher are displayed. DISC: Dominance, Influence, Steadiness, and Conscientiousness; MBTI: Myers-Briggs Type Indicator.



XSL•FO
RenderX

Figure 3. A network diagram of Myers-Briggs Type Indicator (MBTI)—comprising Extraversion (E)–Introversion (I), Sensing (S)–Intuition (N), Thinking (T)–Feeling (F), and Judging (J)–Perceiving (P)—and Dominance, Influence, Steadiness, and Conscientiousness (DISC) based on nodes and edges. Circles denote DISC and squares denote MBTI. Solid lines denote positive correlations and dashed lines denote negative correlations. Only correlations of 0.15 or higher were represented.



Through these findings, we aimed to determine the specific connections between MBTI's personality type indicators (E/I, N/S, T/F, P/J) and DISC's behavioral types (D, I, S, C), and to provide basic data for developing an integrated personality analysis model. The analysis was conducted using SPSS software (IBM Corp) and Pearson correlation coefficients were used to examine the relationships between the data. The sample consisted of 130 college and graduate students in their 20s and 30s.

First, both Extraversion (E) and Introversion (I) in MBTI showed significant correlations with all DISC types (D, I, S, C) (P<.001). Introversion (I) demonstrated positive relationships with Stability (S) and Conscientiousness (C), with a particularly stronger correlation with S (correlation coefficients 0.38, P<.001). In contrast, Extraversion (E) showed negative correlations with Dominance (D) and Influence (I), with a stronger negative correlation with I (correlation coefficients −0.38, P<.001).

Both Intuition (N) and Sensing (S) in MBTI showed significant correlations with Influence (I) and Stability (S) in DISC (P<.001). Intuition (N) had a weak negative correlation with Influence (I) (correlation coefficient −0.19, P=.03), whereas Sensing (S) had a weak positive correlation with Stability (S) (correlation coefficient 0.17, P<.04). Thinking (T) and Feeling (F) also exhibited significant relationships with Dominance (D) and Stability (S) in DISC (P<.001). Thinking (T) showed a strong negative correlation with Dominance (D) (correlation coefficient −0.38, P<.001), and Feeling (F) showed a positive correlation with Stability (S) (correlation coefficient 0.30, P<.001).

Finally, Perceiving (P) and Judging (J) in MBTI had significant correlations with Influence (I) and Conscientiousness (C) in DISC (P<.001). Perceiving (P) showed a negative relationship with Influence (I) (correlation coefficient −0.25, P<.001), while Judging (J) exhibited a positive relationship with Conscientiousness (C) (correlation coefficient 0.31, P<.001).

These results indicate that the correlation between MBTI and DISC is clearly evident in each type. For instance, according to the figure, MBTI's Introversion (I) had a strong positive relationship with DISC's Stability (S), while MBTI's Thinking (T) showed a strong negative relationship with DISC's Dominance (D). Despite MBTI and DISC each analyzing different aspects of personality—psychological preferences and behavioral patterns—these findings suggest that they can be mutually complementary.

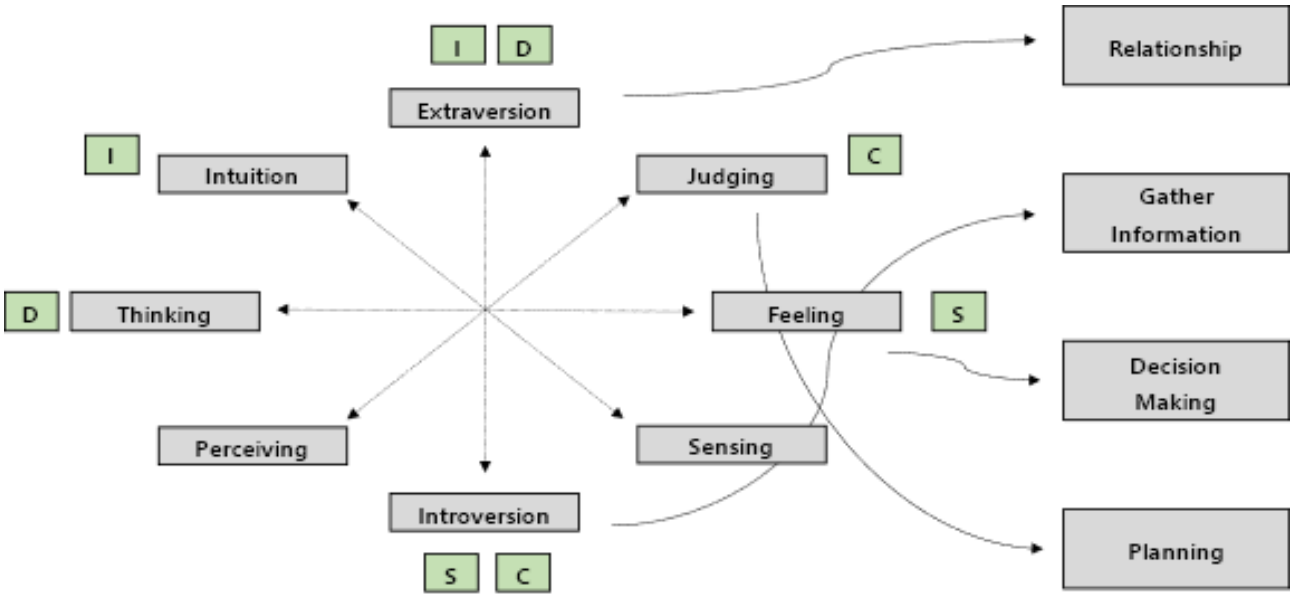
XSL•FO
RenderX

Proposing an Integrated Trait Analysis Model and Developing Survey Items

The survey items were designed based on the DISC-MBTI correlation analysis results, narrowing down the classification into 4 types. A model was developed to categorize individuals into these 4 types, and the survey items were structured accordingly as Figure 4.

Based on the designed survey items, exercise prescriptions have been summarized in Table 2.

Figure 4. Type exploration and classification model based on MBTI-DISC. Considering the correlation between MBTI and DISC, 4 types (Relationship, Gather Information, Decision-Making, and Planning) were categorized. Clustering is represented using arrows. DISC: Dominance, Influence, Steadiness, and Conscientiousness; MBTI: Myers-Briggs Type Indicator.



XSL•FO
RenderX

Table . Exercise prescription types based on the Myers-Briggs Type Indicator (MBTI)–Dominance, Influence, Steadiness, and Conscientiousness (DISC) survey. The table systematically classifies exercise prescription types by organizing them into key categories, including type, description, recommended exercises, and coaching strategies.

Categories and types	Description	Recommended exercises	Coaching, support, and planning strategies
Relationship			
Group-oriented	Finds motivation in activities with others, prefers social interaction.	Group fitness (Zumba, aerobics), team sports	Provide positive feedback, enhance collaboration with peers.
Individual-oriented	Prefers solo activities, focuses on achieving personal goals.	Solo running, gym equipment exercises	Emphasize individual performance, provide an independent environment.
Expert-dependent	Requires guidance and structured plans from trainers or experts, prefers systematic feedback.	Personal training, customized exercise plans	Offer professional, data-driven feedback.
Competition-oriented	Gains motivation through competition or performance comparison, values winning.	Competition-based sports (tennis, boxing, swimming)	Visualize performance comparisons with rivals, provide rewards for victories.
Gather information			
Experience-seeker	Chooses exercises based on reviews or others' experiences.	Trendy programs, popular classes	Provide review-based materials, emphasize user reviews and ratings.
Principle-seeker	Logically selects exercises based on scientific evidence and principles.	Systematic, structured exercise programs	Explain exercise effects and provide scientific data, offer professional resources.
Decision-making			
Logical goal-oriented	Sets specific goals and approaches exercise systematically to achieve them.	HIIT[a], strength training	Set performance-based goals, motivate with concrete data.
Emotion or relationship-oriented	Values emotional satisfaction or building relationships, prioritizes fun and experience.	Dance classes, group sports	Provide emotional support, foster positive relationship building.
Planning			
Structured or detail-oriented	Consistently follows a systematic and detailed plan.	Regular yoga, Pilates	Provide detailed schedules, set step-by-step goals.
Flexible and people-oriented	Adjusts plans flexibly according to circumstances, prefers interaction with others.	Trekking, diverse activity experiences	Emphasize autonomy, offer engaging and appealing options.

[a]HIIT: high-intensity interval training.

Based on the DISC-MBTI correlation analysis, the survey items were designed to narrow the classification down to 4 types. A model was then developed to categorize individuals into these 4 types, and corresponding survey items were created accordingly. Finally, the exercise prescriptions derived from these survey items are summarized.

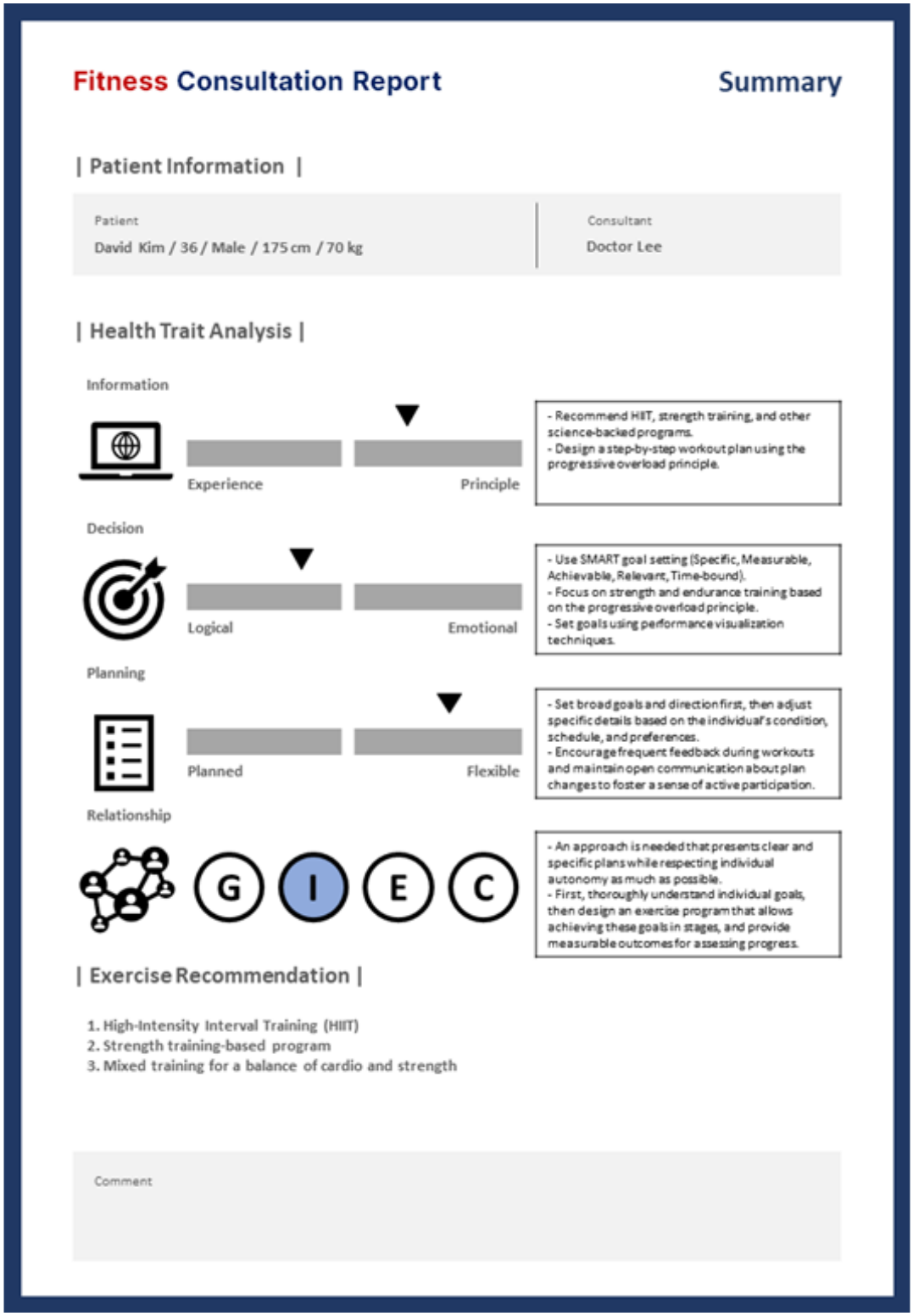
In this study, we developed survey items to classify exercise tendencies by measuring individuals' relational tendencies. Each item assesses aspects such as the individual's level of activity, expression styles, and relationship preferences in interpersonal contexts. This approach identifies preferred environments and motivational factors during exercise. For the purposes of this study, we focused exclusively on methods for categorizing relationships with others during exercise.

Implementation and Validation of an Exercise Prescription Reporting Service

A practical profiling web application was implemented to input survey items and verify whether the results matched actual patient information. The web application was developed based on Flask and calculated the total scores by summing up the scores for each of the 4 survey items. The results of this process are presented in Figure 5.

XSL•FO
RenderX

Figure 5. Example of an exercise prescription consulting report for 4 categories: relationship, information-gathering, decision-making, and planning.



A pilot test was conducted to validate the practicality and reliability of the proposed integrated trait analysis model. This process assessed whether the proposed profiling system effectively analyzed individuals' personality and behavioral traits and whether it could be practically used. The validation process consisted of a pilot test, result analysis and interpretation, and a practicality evaluation through expert interviews.

The pilot test involved applying the profiling system, developed based on MBTI and DISC survey data, to a group of 20 experts. The experts responded to the survey items, and the system analyzed their personality and behavioral types to generate customized profiling results. This process confirmed whether the profiling system could accurately classify and interpret types based on the survey data.

XSL•FO
RenderX

System Usability Scale Score Result for Prescription Reporting Service

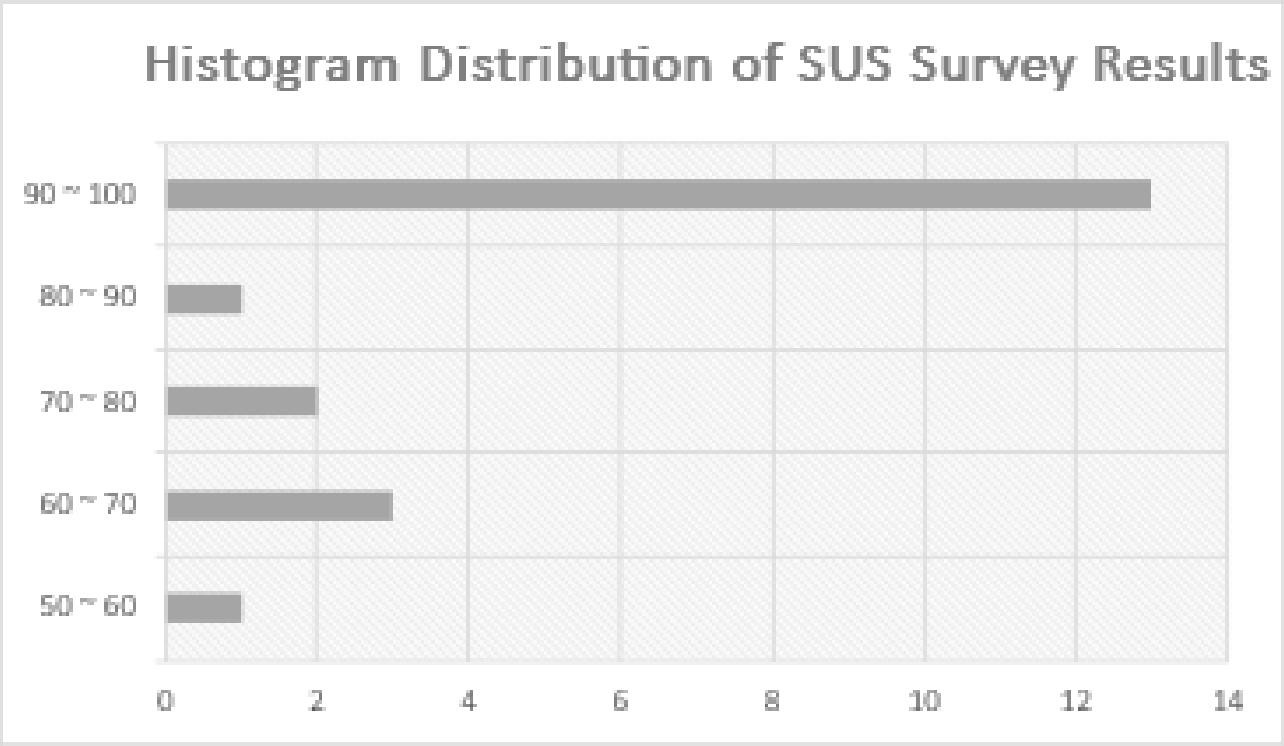
To evaluate the usability of the proposed system, a System Usability Scale (SUS) assessment was conducted with 20 experts, as Table 3.

Table . Survey participants by occupation. The survey was primarily conducted with trainers, physical therapists, and professors.

Occupation	Participants, n
Trainer	8
Physical therapist	4
Exercise specialist	4
Professor in related fields	3
Aerobics instructor	1

The SUS, originally developed by Brooke [15], is a widely used and reliable tool for measuring system usability across various domains. The questionnaire consists of 10 items rated on a 5-point Likert scale, designed to assess both positive and negative aspects of usability. Odd-numbered items measure positive usability attributes (eg, ease of use and system integration), while even-numbered items measure negative usability attributes (eg, complexity and inconsistency). The standard scoring methodology was applied, where responses to odd-numbered items were adjusted by subtracting 1, and even-numbered items were adjusted by subtracting the response from 5. The final SUS score was computed by summing the adjusted values and multiplying by 2.5, following the established methodology described by Lewis and Sauro [16]. The resulting SUS score of 86.0 indicates a high level of usability and user satisfaction, as shown in Figure 6.

Figure 6. Histogram distribution of System Usability Scale (SUS) survey results. The majority of scores clustered in the 90 - 100 range, although a few respondents provided lower ratings.



Trainers, physical therapists, exercise specialists, professors in related fields, and aerobics instructors responded positively, highlighting that the system is intuitive, effectively integrates various functions, and is easy to navigate. In addition, participants across clinical, academic, and fitness backgrounds stressed that a highly personalized approach, tailored to individual user characteristics, is key to enhancing engagement and usability. Some experts, particularly trainers and exercise specialists, noted that the system has the potential to serve as an effective tool for continuous health care management by adapting to users' needs. However, some participants emphasized the need for more detailed exercise coaching programs and suggested conducting further real-world validation to ensure the system's effectiveness in practical settings. These findings, supported by feedback from professionals across diverse fields, including trainers, sports dance instructors, and professors, confirm the system's practical applicability and its potential for real-world implementation.

XSL•FO
RenderX

Discussion

Principal Results

This study identified the integrated correlation between MBTI and DISC, addressing the limitations of existing personality theories and enabling theoretical expansion. MBTI focuses on internal psychological preferences, while DISC emphasizes external behavioral patterns. By combining the strengths of these 2 tools, this study proposed a new analytical framework for comprehensively understanding individuals' personality and behavior.

The integrated trait analysis model, based on the correlation between MBTI and DISC, clarified the relationship between personality and behavioral types and validated its theoretical soundness through data-driven analysis. This research provides a new direction for personality analysis and behavior prediction studies in academic fields and serves as foundational data for future research.

In practical applications, this model can be used as a valuable tool for health care and health behavior change programs. For example, personalized health programs can be designed by optimizing exercise plans, diet strategies, and stress management techniques according to individual personality and behavioral types. For instance, individuals with DISC's Stability (S) and MBTI's Introversion (I) may benefit from quiet, structured exercise environments, whereas those with DISC's Influence (I) and MBTI's Extraversion (E) may thrive in sociable and dynamic activities.

In addition, the findings contribute to the development of personalized coaching programs that motivate behavior change based on a deep understanding of individual tendencies and behavior patterns. This system can be applied in diverse practical environments, such as hospitals, fitness centers, and counseling institutions, offering high user satisfaction while enhancing service quality.

Limitations

This study offers valuable insights into the correlation between MBTI and DISC, but it has several limitations. First, our sample size was limited to 130 Korean college and graduate students in their 20s and 30s, which may restrict the generalizability of our findings. The sample was also somewhat skewed toward female students and those born in the 2000s. The reliance on self-reported survey data also introduces potential biases, such as social desirability or self-perception. Future research should include larger, more diverse populations and incorporate additional data collection methods, like observational studies or third-party assessments.

We recognize that the cultural context of South Korea significantly influenced our survey results, particularly among the 20s demographic, which is sensitive to social norms. For example, Korea's collectivistic culture may modulate the expression of Extraversion (MBTI) and emphasize diligence for Conscientiousness (DISC C-type) or harmony for Steadiness (DISC S-type) in responses.

Furthermore, our study focused solely on MBTI and DISC. Integrating additional frameworks, such as the Big Five Personality Traits, could offer a more comprehensive understanding. The practical applications discussed, like healthcare and behavior change programs, also require further real-world validation to confirm their effectiveness. Finally, we simplified complex traits into 4 main categories for usability, potentially overlooking nuanced individual differences and not accounting for longitudinal changes in personality. Future studies should explore these dynamics and validate the profiling system in broader technological and organizational contexts.

Conclusions

In conclusion, this study successfully identified the integrated correlation between MBTI and DISC, combining their strengths to propose a novel analytical framework for understanding personality and behavior. By validating the integrated trait analysis model with data-driven methods, the research highlighted its potential to clarify the relationship between psychological preferences and behavioral patterns. This model offers both theoretical contributions to personality analysis and practical applications, such as personalized health behavior programs, which can optimize exercise plans, diet strategies, and stress management based on individual traits. While further validation in diverse settings is needed, the findings pave the way for future innovation in health care, counseling, and other fields that benefit from tailored behavior change strategies.

Conflicts of Interest

None declared.

References

1. Noncommunicable diseases country profiles 2018. World Health Organization. 2018. URL: https://www.who.int/publications/i/item/ncd-country-profiles-2018 [accessed 2025-06-13]
2. DiPietro L, Al-Ansari SS, Biddle SJH, et al. Advancing the global physical activity agenda: recommendations for future research by the 2020 WHO physical activity and sedentary behavior guidelines development group. Int J Behav Nutr Phys Act 2020 Nov 26;17(1):143. [doi: 10.1186/s12966-020-01042-2] [Medline: 33239105]
3. Ma JK, Floegel TA, Li LC, et al. Tailored physical activity behavior change interventions: challenges and opportunities. Transl Behav Med 2021 Dec 14;11(12):2174-2181. [doi: 10.1093/tbm/ibab106]

XSL•FO
RenderX

4. Kremers SPJ, de Bruijn GJ, Visscher TLS, van Mechelen W, de Vries NK, Brug J. Environmental influences on energy balance-related behaviors: a dual-process view. Int J Behav Nutr Phys Act 2006 May 15;3(1):9. [doi: 10.1186/1479-5868-3-9] [Medline: 16700907]
5. Wilson KE, Dishman RK. Personality and physical activity: a systematic review and meta-analysis. Pers Individ Dif 2015 Jan;72:230-242. [doi: 10.1016/j.paid.2014.08.023]
6. Courneya KS, Hellsten LAM. Personality correlates of exercise behavior, motives, barriers and preferences: an application of the five-factor model. Pers Individ Dif 1998 May;24(5):625-633. [doi: 10.1016/S0191-8869(97)00231-6]
7. Judge TA, Bono JE, Ilies R, Gerhardt MW. Personality and leadership: a qualitative and quantitative review. J Appl Psychol 2002 Aug;87(4):765-780. [doi: 10.1037/0021-9010.87.4.765] [Medline: 12184579]
8. Gardner WL, Martinko MJ. Using the Myers-Briggs Type Indicator to study managers: a literature review and research agenda. J Manage 1996;22(1):45-83. [doi: 10.1016/S0149-2063(96)90012-4]
9. Pittenger DJ. Cautionary comments regarding the Myers-Briggs Type Indicator. Consult Psychol J 2005 Jan 1;57(3):210-221. [doi: 10.1037/1065-9293.57.3.210]
10. Myers IB, McCaulley MH. Manual: A Guide to the Development and Use of the Myers-Briggs Type Indicator: Consulting Psychologists Press; 1985. URL: http://ci.nii.ac.jp/ncid/BA19325536 [accessed 2025-07-11]
11. Slowikowski MK. Using the DISC behavioral instrument to guide leadership and communication. AORN J 2005 Nov;82(5):835-838. [doi: 10.1016/s0001-2092(06)60276-7] [Medline: 16355940]
12. Jamjoom H, Alshareef W, Alotaibi N. Association between personality type and academic achievement in undergraduate dental students in King Abdulaziz University. Ann Dent Spec 2021 Jan 1;9(4):65-72. [doi: 10.51847/gWuAXEVdLY]
13. Idris N, Utami E, Hartanto AD, Raharjo S. Systematic literature review of profiling analysis personality from social media. J Phys: Conf Ser 2021 Mar 1;1823(1):012115. [doi: 10.1088/1742-6596/1823/1/012115]
14. Brooke J. SUS: a “quick and dirty” usability scale. In: Usability Evaluation in Industry: CRC Press; 1996:207-212. [doi: 10.1201/9781498710411-35]
15. Lewis JR. The System Usability Scale: past, present, and future. Int J Hum-Comput Interact 2018 Jul 3;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]
16. Lewis JR, Sauro J. The factor structure of the System Usability Scale. In: Human Centered Design: Springer; 2009:94-103. [doi: 10.1007/978-3-642-02806-9_12]

Abbreviations

16PF : Sixteen Personality Factor
DISC: Dominance, Influence, Steadiness, and Conscientiousness
MBTI: Myers-Briggs Type Indicator
SUS: System Usability Scale

Edited by D Gooch; submitted 03.03.25; peer-reviewed by C Udensi, J Abua; revised version received 21.05.25; accepted 21.05.25; published 18.07.25.

Please cite as:
Kim D, Lee DH, Hwang MK
A Comprehensive Profiling System Integrating Myers-Briggs Type Indicator (MBTI) and Dominance, Influence, Steadiness, and Conscientiousness (DISC) for Personalized Health Training: Correlational Analysis and Usability Evaluation
JMIR Hum Factors 2025;12:e73397
URL: https://humanfactors.jmir.org/2025/1/e73397
doi:10.2196/73397

© Donghyun Kim, Dong Hun Lee, Mi Kyung Hwang. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 18.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Personalized Digital Care Pathways Enable Enhanced Patient Management as Perceived by Health Care Professionals: Mixed-Methods Study

David Rodrigues[1,2,3], MD; Clara Jasmins[2,3], MD; Ricardo Ladeiras-Lopes[3,4], MD, PhD; Luis Patrao[3,5], MD; Eduardo Freire Rodrigues[3], MD

[1]Comprehensive Health Research Centre, NOVA Medical School, Lisbon, Portugal

[2]Family Medicine Department, NOVA Medical School, Lisbon, Portugal

[3]UpHill Health, SA, Lisbon, Portugal

[4]Cardiovascular R&D Centre-UnIC@RISE, Department of Surgery and Physiology, Faculty of Medicine, University of Porto, R. Dr. Plácido da Costa, Porto, Portugal

[5]Faculty of Health Sciences, University of Beira Interior, Covilhã, Portugal

Corresponding Author:
Ricardo Ladeiras-Lopes, MD, PhD
UpHill Health, SA, Lisbon, Portugal

Abstract

Background: Clinical decision support systems are known to improve adherence to clinical practice guidelines and patient outcomes by providing clinicians with timely, accurate, and appropriate knowledge.

Objective: This study investigates the perceived usefulness and practical implementation of UpHill Route v3, a personalized digital care pathway (PDCP) system, in enhancing clinical decision-making and patient management across various clinical settings.

Methods: A mixed-methods retrospective study was conducted among medical doctors and nurses from four National Health System-Local Health Units in Portugal. Data were collected from May 2023 to April 2024. The primary data source was an anonymous questionnaire assessing health care professionals' perceptions of UpHill Route v3's usefulness using the Likert scale ranging from 0 (do not agree) to 10 (totally agree). Secondary analysis involved quantifying decisions across heart failure, multimorbidity, diabetes, and colorectal and breast cancer clinical pathways. These data were collected from user interactions with UpHill Route v3 as well as from its internal database. Descriptive and bivariate statistics were used to analyze the data.

Results: A total of 22 health care professionals with mean age 44.7 (SD 10.6) years, including 15 (68%) female participants and 9 (41%) physicians were included in the study. High ratings for adherence to clinical protocols, mean score 8.06 (SD 1.73); clinical decision support, mean score 8.05 (SD 1.73); patient care improvement, mean score 7.63 (SD 2.22); and confidence in patient management, mean score 8.26 (SD 1.56) were reported. Secondary analysis showed that across 3574 patients, 25,741 clinical decisions were informed, and 9254 actions were performed with the assistance of the PDCP tool.

Conclusions: The UpHill Route v3 PDCP tool is highly valued by health care professionals for its ability to support clinical decision-making and improve operational efficiency across various clinical settings. Our findings suggest that this tool can effectively bridge the gap between clinical guidelines and real-world practice.

(JMIR Hum Factors 2025;12:e68581) doi:10.2196/68581

KEYWORDS

care integration; clinical pathway; digital; heart failure

Introduction

Computerized clinical decision support systems (CDSS) are designed to enhance patient care by providing clinicians and patients with timely, accurate, and appropriate knowledge in the context of specific medical interactions or encounters. These systems have gained significant traction for their potential to improve adherence to clinical practice guidelines, bridging the gap between knowledge and its application in diverse clinical contexts [1,2]. Systematic reviews and studies have been conducted to evaluate the effectiveness of CDSS, showing that CDSS can lead to improvements in process adherence, thus leading to higher diagnosis accuracy and comprehensive disease management [3,4]. There are multiple types of CDSS. Traditional CDSS primarily offer static recommendations, whereas personalized CDSS integrates patient-specific data

XSL•FO
RenderX

from an electronic health record (EHR) to tailor clinical recommendations. Personalized digital care pathway (PDCP) represents a further advancement, extending CDSS capabilities beyond single encounters by ensuring continuity of care across multiple medical interactions. PDCP systems facilitate decision-making and action execution across multidisciplinary teams, while maintaining compliance with best practice protocols and continuously adapting based on patient-specific data. These systems have demonstrated benefits for patients, health care professionals (HCP), and institutions by contributing to improved service quality, workforce efficiency, and patient outcomes. This expands the scope of CDSS to the entire continuum of care, enabling the interaction among multiple multidisciplinary teams through a shared care pathway. [4,5].

PDCPs ensure that each HCP delivers tailored care aligned with standard protocols and guidelines, especially relevant for chronic conditions requiring frequent assessments and timely therapy adjustments across different contexts [6]. Current management strategies for chronic patient care typically rely on traditional methods such as face-to-face consultations, periodic monitoring, and manual documentation of patient records. Care plans are often manually developed based on clinical guidelines and personal clinical judgment, with minimal decision support integration. Most EHRs lack advanced decision support capabilities, decision characterisation, and patient communication features. Overall, current management strategies can be time-consuming and prone to variability in care delivery.

For example, CDSS have been shown to improve patient outcomes in heart failure management by enhancing monitoring, management, and optimization of therapies [7]. As more clinical practice guidelines emphasize the necessity of digital solutions for early disease detection and patient self-care [8], multidisciplinary chronic disease management programs powered by CDSS tools have yielded better quality of life, lower readmission rates, and reduced mortality [3]. Despite these benefits, there is a lack of evidence regarding the perceived usefulness of CDSS among HCPs using these systems and on the quantification and characterization of this usage on the number and types of decisions and actions usually performed by CDSS in chronic disease management [3,9].

This study hypothesizes that HCPs consider PDCPs, such as Uphill Route v3, to be more valuable than current management strategies for chronic patient management. As a PDCP and a subtype of CDSS, UpHill Route v3 aids medical decision-making and improves patient management by increasing compliance with locally validated clinical pathways. The software creates personalized care pathways by capturing patient health and demographic data from hospital EHRs, manual inputs by HCPs, or self-reported patient questionnaires (sent via text messages, web forms, or voice interaction capabilities). This approach further supports HCP clinical decisions and tasks throughout the clinical pathway. Thus, the main aim of this study was to characterize the perceived usefulness of a PDCP in chronic patient management. The secondary aim was to describe the number and types of decisions supported by the software, thus quantifying them as a proxy for the system's impact on decision support and patient management efficiency.

Methods

Study Design

This was a mixed-methods retrospective study, conducted to evaluate the usefulness of PDCP as a CDSS for HCPs. For the primary aim, qualitative data was collected through a questionnaire directed at HCPs using UpHill Route v3, to assess the perceived usefulness of PDCP and compare it with current management practices for chronic patient care. For the secondary aim, quantitative data on decisions and tasks performed during care pathways was retrieved from UpHill Route v3. This design allowed for a comprehensive understanding of the impact of UpHill Route v3 on clinical decision-making and patient management.

Setting

Medical doctors and nurses from four National Health System Local Health Units (LHUs) in Portugal, who were pilot-testing UpHill Route v3, were invited to participate. These professionals were trained in using UpHill Route v3 as part of the software implementation framework, as described in the software technical documentation and instructions for use. As part of the implementation process, clinical content embedded in the software was validated and approved, meaning that clinical leaders from the enrolled institutions approved the protocols reflected on the software.

The inclusion of patients in UpHill Route v3 was the exclusive responsibility of HCPs, who were instructed and trained to use the software in accordance with the technical documentation for noncritical patients in non-critical clinical settings.

The study encompassed the central, capital, and southern coastal regions of Portugal and conducted from May 2023 to April 2024. The LHUs—Litoral Alentejano, Amadora Sintra, Coimbra, and São José—serve a combined population of over 1.40 million people across 17,660 km². This vast region includes urban and suburban populations, and the clinical journeys in these four LHUs include heart failure, multimorbidity (ie, combined heart failure, chronic obstructive pulmonary disease, and diabetes), diabetes, and colorectal and breast cancer, which represent a diverse spectrum of chronic diseases with frequent acute care interventions.

Participants

To be eligible for this study, participants were required to be part of the HCP user base at the previously described health units that were pilot testing UpHill Route v3 in their daily practice and to have completed the training sessions in the proper use of the software. The instruction to use UpHill Route v3 only in noncritical settings was directly related to the targeted diseases, which were chronic or elective conditions integrated into predefined care pathways. The HCPs were briefed on the appropriate use of the system, and adherence to this criterion was monitored through participant feedback and usage data.

It is important to note that the device does not autonomously provide care to patients, and all the content and behavior are based on the locally validated clinical pathway. All participants

XSL•FO
RenderX

provided informed consent to be included in each care journey, ensuring their voluntary participation.

Endpoints and Variables

The primary aim was to study the perceived usefulness of this PDCP captured through a questionnaire administered to HCPs piloting UpHill Route v3. This questionnaire was developed so that the constructs under each question contributed to understanding the perceived usefulness of UpHill Route v3: (1) adherence to clinical protocols (ease of compliance with protocols; (2) clinical decision support (measure of recommendation relevance); (3) improvement in patient care (perceived value of the presented recommendations to the patient); (4) task management (time-saving benefits); and (5) confidence in patient management (reliability of support for delivery of exhaustive and proper care). Each question was answered using a Likert scale ranging from 0 to 10, indicating the participant's level of agreement with the presented component concerning PDCP usefulness. The questionnaire was written in Portuguese and distributed to participants via email. It is available in the supplementary materials (Multimedia Appendix 1).

The secondary aim was to describe the number and type of decisions supported by the software, by analyzing human-computer interactions made on the UpHill Route v3 graphical user interface. Decisions made with the support of the software are encapsulated in decision gateways, which are elements within UpHill Route v3 that prompt decision-making based on specific criteria. These criteria may be related to the patient's condition, signs and symptoms, or clinical results. Multimedia Appendix 2 represents an example interface for UpHill Route v3 graphical interface for the type 2 diabetes mellitus clinical pathway of a test patient. In this interface, HCPs are presented with patient-specific actions and decisions in the center panel. On the right panel, supportive complementary information related to the selected action or decision in the central panel is presented to the HCP. Data were collected from user interactions with UpHill Route v3, as well as from its internal database. Participants were characterized according to demographic variables (eg, sex, age) and years of clinical experience. Additionally, information on academic activity as a binary variable indicating whether participants were involved in academic roles alongside their clinical duties was collected.

Data Sources/Measurement

Primary objective data were collected through an online questionnaire. The questionnaire (in Portuguese) was delivered through Microsoft Forms and sent to the participants through email and are available in Multimedia Appendix 1. The analysis of the data was done using Microsoft Excel and STATA (version 14; StataCorp).

Secondary objective data were collected through queries from UpHill Route v3 internal database.

Data Processing Protocol

Data were collected from the software's database to describe the number and type of decisions supported by the software. The database recorded each option clicked on by the user for each pathway element. Then, these data were aggregated by pathway elements (ie, actions and decisions), each with a unique identification number. To analyze these elements and their selected options, queries on the database were performed, aggregating the same pathway element across all patient records.

Decisions were categorized into one of the ten clinical decision types as defined in the Decision Identification and Classification Taxonomy for Use in Medicine (DICTUM) [10]. These includes gathering additional information, evaluating test results, defining problems, drug-related decisions, therapeutic procedure-related decisions, legal and insurance-related decisions, contact-related decisions, advice, and precaution, setting treatment goals, and deferment.

Actions were categorized as clinical if assigned to nurses or physician profiles, or as nonclinical if assigned to administrative profiles. For example, clinical tasks included the recommendations to request a mandatory diagnostic test, while nonclinical tasks involved appointment scheduling. All data were aggregated by pathway elements, combining all the patient records without any identifiable patient variable.

Study Size

For the primary objective, all 48 potentially eligible HCPs who were actively piloting UpHill Route v3 at the time of study recruitment were invited via email to answer the questionnaire.

For the secondary objectives, internal data were retrieved anonymously for HCPs and patient data. Therefore, data from the activities of all 48 professionals was analyzed.

Statistical Methods

Descriptive statistics were used to summarize the demographic characteristics of the participants. Continuous variables were summarized with central tendency and dispersion measures, namely mean and standard deviation. Categorical variables were reported as relative and absolute frequencies. Bivariate analysis was conducted with a χ^2 test for categorical variables and a 2-tailed Student *t* test for continuous variables.

To address the primary objective, responses collected from the questionnaire were analyzed by calculating mean scores and measures of dispersion (standard deviation) for each question. Higher mean scores indicated greater perceived usefulness of UpHill Route v3, while low dispersion reflected stronger agreement among participants regarding the software's usefulness.

To address the secondary objective, we categorized and counted the types of decisions involving human-computer interaction, which were presented as absolute and relative frequencies. Additionally, we analyzed the tasks suggested by UpHill Route v3 within each care pathway, reporting these tasks as percentages.

Ethical Considerations

This research project was submitted and approved by NOVA Medical School under the main author's PhD research project "Characterizing uncertainty in decision making with complex multimorbidity patients," (No.108/2021/CEFCM) (Multimedia Appendix 3). The study involved secondary analysis of existing

XSL•FO
RenderX

data for which informed consent was obtained at the time of initial collection, and the original consent and institutional review board approval covered secondary analysis without requiring additional consent from the participants. All data used in this study were fully anonymized to ensure privacy and confidentiality. No compensation was provided to participants.

Results

A total of 22 participants answered the questionnaire (response rate, 46%). The mean age of the participants was 44.7 (SD 10.6) years, while mean years of experience was 20.6 (SD 11.0) years. Of the 22 participants, 15 (68%) were female; 9 (41%) were medical doctors, and the remaining 13 (59%) were nurses. Additionally, 2 (9%) of participants were involved in academic activities.

Bivariate analysis showed no significant associations between gender and medical doctor status (P=.30), gender and nurse status (P=.13), or gender and involvement in academic activities (P=.66). Similarly, t tests for continuous variables indicated no significant differences in age (P=.54) or years of experience (P=.44) between male and female participants.

Perceived Usefulness of UpHill Route V3

Participants rated the statement "UpHill Route v3 makes it easier to manage patients according to the clinical protocols of my institution" with a mean score of 8.06 (SD 1.73), a statement that aimed to assess the ability to easily comply with protocols. Focusing on clinical decision support, the statement "the UpHill Route v3 provides me with relevant clinical recommendations at the right time" received a mean score of 8.05 (SD 1.73). For the statement "the UpHill Route v3 helps me to improve my patient care", which evaluated the sense of patient care improvement, the mean score was 7.63 (SD 2.22). Finally, for confidence in patient management, "the UpHill Route v3 gives me more confidence that my patients are being properly cared for" received a mean score of 8.26 (SD 1.56). The results of the questionnaire are presented in Table 1.

Table . Results of the questionnaire applied to HCPs[a] to assess the perceived usefulness of the UpHill Route v3[b].

Question	Mean score (SD)
UpHill Route v3 makes it easier to manage patients according to the clinical protocols of my institution	8.06 (1.73)
The UpHill Route v3 provides me with relevant clinical recommendations at the right time	8.05 (1.73)
The UpHill Route v3 helps me to improve my patient care	7.63 (2.22)
The UpHill Route v3 gives me more confidence that my patients are being properly cared for	8.26 (1.56)

[a]HCP: health care practitioner.

[b]Responses collected from the questionnaire were analyzed by calculating mean scores and measures of dispersion (standard deviation) for each question using the Likert scale ranging from 0 (do not agree) to 10 (totally agree) indicated the participant's level of agreement with UpHill Route v3..

There were no significant differences in the perceptions of effectiveness of the UpHill Route v3 between physicians and nurses, as well as between genders.

Overall, the demographic characteristics or perceived usefulness of the UpHill Route v3 did not significantly differ based on gender or professional role (ie, medical doctor vs. nurse).

Decisions Supported by the Software

Data were collected from heart failure, diabetes, and cancer (colorectal and breast) care pathways across the four different institutions (ie, LHUs).

Forty-eight HCPs used UpHill Route for supporting decisions of a total of 3574 patients, with 25,741 decisions captured. HCPs answered 427 decision gateways, performed 9254 actions, and the system generated 2429 clinical notes. Overall, there were 37,534 activities (ie, sum of decisions, actions, and notes) performed in UpHill Route v3. These results are summarized in Table 2.

XSL•FO
RenderX

Table . Aggregated data of human-computer interactions across care pathways.

Local health unit (Care pathway)	Patients (n=3574)	Gateways (n=427)	Decisions (n=25,741)	Actions (n=9254)	Clinical notes (n=2429)	Total activities per pathway, n (activities per patient)
Litoral Alentejano (heart failure)	609	76	8249	2570	0	10,819 (19.5)
Litoral Alentejano (multimorbidity)	668	207	10,680	2350	0	13,030 (17.8)
Coimbra (colorectal cancer)	355	3	685	1356	0	2041 (5.8)
São José (colorectal cancer)	202	25	883	856	144	1883 (9.3)
São José (breast cancer)	223	32	878	1079	107	2064 (9.3)
Amadora/Sintra (diabetes)	1517	84	4466	1053	2178	7697 (5.1)
Total	—[a]	—	—	—	—	37,534 (10.5)

[a]Not applicable.

The Litoral Alentejano multimorbidity care pathway presented the highest number of activities per patient, with an average of 19.5 activities, followed by the Litoral Alentejano heart failure care pathway which exhibited a high level of clinical activity, with 17.8 activities per patient. São José colorectal and breast cancer pathways had an average of 9.3 activities per patient. Coimbra colorectal cancer and Amadora/Sintra diabetes care pathways averaged 5.8 and 5.1 activities per patient, respectively. Overall, the mean number of clinical activities per patient performed in the four health institutions using the UpHill Route v3 was 10.5.

As part of the secondary objective, the categorization of clinical decisions showed that the majority of decisions (9215/25,841, 35.7%) were related to “defining problem”, followed by “gathering additional information” (n=5759, 22.3%) and “treatment goal” (n=4845, 18.7%). The distribution of these decisions varied across different care pathways, with Litoral Alentejano multimorbidity and heart failure accounting for the largest proportions, representing 41.3% (n=10,678) and 31.9% (n=8251) of the total decisions, respectively. Other decision types, such as “contact-related” and “therapeutic procedure–related,” accounted for smaller proportions, 7.0% (n=1814) and 5.0% (n=1280), respectively. This distribution highlights the diverse range of decision types supported by UpHill Route v3, tailored to the specific needs of different clinical contexts. The results of the decision categorization, presented by clinical journey are summarized in Multimedia Appendices 4 and 5 .

While analyzing decision types, a total of 1814 (n=7.0%) decisions were found to be contact-related. These involved admitting or discharging patients, scheduling follow-ups, and referring patients to other parts of the health care system. There were 460 (1.8%) deferment decisions, which involved actively delaying or rejecting a decision on a problem presented by the patient. The highest number of decisions were defining-problem decisions, totaling 9215/25,841 (35.7%). These complex, interpretative assessments defined what the problem was, including diagnostic conclusions, evaluations of health status, etiological inferences, and prognostic judgments. A total of 1415 (5.5%) decisions were related to drug therapy, which involved starting, stopping, altering, or maintaining a drug regimen. There were 1053 (4.1%) decisions related to the evaluation of test results, involving simple, normative assessments of clinical findings and tests. A total of 5759 (22.3%) decisions were aimed at gathering additional information, which involved obtaining information from sources other than patient interviews, physical examinations, and patient charts. There were 1280 (5.0%) decisions related to therapeutic procedures; these involve planning, performing, or refraining from therapeutic interventions. A total of 4845 (18.7%) decisions were related to setting treatment goals, which involved defining specific objectives for patient care. The comprehensive details of each decision type’s categorization can be found in Multimedia Appendix 6.

Discussion

Principal Findings

This study evaluated the perceived usefulness and practical implementation of the UpHill Route v3 as a PDCP in various clinical settings. The primary outcome focused on the perceived usefulness of the system, as assessed by HCPs through a structured questionnaire. The results demonstrated high scores across multiple dimensions, including adherence to clinical protocols, clinical decision support, and confidence in patient management. Specifically, participants rated the system highly in the ease of adhering to clinical protocols by providing relevant clinical recommendations, demonstrating that HCPs perceive UpHill Route v3 as a valuable tool in clinical settings. Our analysis found no significant differences in the perceived usefulness of UpHill Route v3 between genders or professional roles (medical doctors vs nurses). This lack of disparity indicates that UpHill Route v3 is perceived to be equally beneficial across different demographic and professional groups within health care settings. Overall, this posits that the software with its

XSL•FO
RenderX

decision support capabilities proves valuable in adhering to clinical guidelines, providing relevant clinical guidance, enhancing patient care, and instilling confidence in the provided care.

The high scores across multiple dimensions of perceived usefulness indicate that HCPs find the UpHill Route v3 system valuable for clinical decision support. The system's ability to integrate clinical protocols and provide timely recommendations appears to enhance both clinical efficiency and the quality of patient care.

Concerning secondary outcomes, the data indicated significant usage of clinical activities being supported by UpHill Route v3, averaging 10.5 activities per patient. There were expected variations in clinical activities per patient across different care pathways. Litoral Alentejano multimorbidity and heart failure cases required the most activities per patient, suggesting highly complex pathways with numerous decision points. These care pathways naturally require more complex clinical activity for the management of these conditions. In contrast, Coimbra Colorectal Cancer and Amadora Sintra Diabetes involved fewer activities per patient, reflecting a more focused and shorter clinical pathway in those users. The distribution of decisions supported by UpHill Route v3 varied significantly, with defining–problem decisions and gathering additional information being the most common. This finding is consistent with the complex nature of clinical diagnostics and the need for comprehensive information gathering to inform clinical decisions in the follow-up phase of symptomatic chronic diseases such as heart failure or multimorbidity [11]. The system's ability to support these critical decision types highlights its relevance and utility in clinical practice. Overall, the categorization of decisions and tasks demonstrates the system's versatility in supporting various aspects of patient management, from diagnosis and treatment to follow-up care. These findings suggest that the UpHill Route v3 system can effectively bridge the gap between clinical guidelines and real-world practice, potentially leading to improved patient management.

The study's findings on UpHill Route v3 align with existing literature on the effectiveness of CDSS in healthcare [12]. Our findings emphasize the importance of CDSS, particularly PDCP, in providing timely and accurate decision support, through personalization and longitudinal continuity, which is crucial for clinicians in complex and dynamic environments [13]. This aligns with findings from multiple systematic reviews and studies. Although many CDSS enhance practitioner performance, evidence of their impact on patient outcomes remains limited and inconsistent, indicating a need for further research[14]. A recently published systematic review and meta-analysis focused on the primary prevention of cardiovascular disease demonstrated that CDSS were effective in improving blood pressure clinical target attainments. However, multilayered barriers affecting the uptake, longer-term use and active engagement from clinicians and patients may hinder their full potential in achieving other quality-of-care outcomes [15]. While digital care pathways have shown promise in optimizing the efficiency of care delivery, effective implementation requires training, resource allocation, support for information management, and optimizing alert algorithms to manage information overload [16,17].

The positive feedback on UpHill Route v3's decision support capabilities reinforces the concept that a PDCP as a CDSS can enhance clinical decision-making and patient management, and the subsequent step for PDCPs is likely to be task automation of predictable and mandatory tasks. Previous findings in the literature highlight its benefits. For instance, a systematic review by Jones et al [18] highlighted how task automation through CDSS can streamline workflows, reduce administrative burdens, and minimize human errors, allowing HCPs to focus more on direct patient care.

Future research should explore the long-term impacts of task automation on clinical outcomes and examine how the integration of PDCPs with other health care technologies can provide efficient orchestration of patient health care.

Several limitations and potential sources of bias need to be acknowledged in this study. First, only 22 HCPs replied to the questionnaire, a small number that may limit the generalizability of the findings, although they worked at multiple health care centers. Additionally, the study relied on self-reported data to assess the perceived usefulness of the system, which could be subject to response bias. Despite anonymization, participants may have provided favorable responses due to social desirability or perceived expectations of the study organizers. However, a strength of the study is the inclusion of aggregated data from diverse patient populations (eg, heart failure, diabetes, breast and colorectal cancer) across various clinical settings (ie, primary and secondary care) and contexts (ie, urban and rural).

Although all authors are affiliated with UpHill Health, S.A., the study design, data collection, and analysis were conducted using standardized methodologies to ensure objectivity. The findings were interpreted transparently and align with prior research on clinical decision support systems, reinforcing their validity.

While the study's findings are promising, the small sample size limits the generalization of the results. Future research should aim to include a larger and more diverse cohort across multiple health care settings (ie, various specialties and practice settings) and geographic regions to enhance external validity and applicability to broader clinical contexts. We speculate that variations in health care system infrastructure, clinical workflows, cultural factors, and user experience levels may influence the adoption and perceived effectiveness of PDCP.

Additionally, this study primarily focuses on short-term implementation outcomes. Future research should adopt a longitudinal approach to evaluate sustained adherence to protocols, patient outcomes, and cost efficiency, allowing for a more comprehensive assessment of the long-term impact of the UpHill Route v3 system on clinical outcomes and health care efficiency. Despite these limitations, the positive qualitative feedback and the objective quantitative interaction with the software by HCPs suggest that the UpHill Route v3 system is a valuable tool in various clinical settings, supporting decision-making and enhancing patient care.

XSL•FO
RenderX

Conclusion

This study demonstrates that HCPs highly value PDCPs as CDSSs for enhancing adherence to clinical protocols, providing timely recommendations, and improving patient care. The system showed significant activity, with several decisions supported across various clinical journeys and settings.

Despite the study's small sample size, the anonymous nature of the questionnaire and objective secondary outcome data provide robust evidence of the system's effectiveness across different clinical scenarios such as chronic, oncological, and surgical conditions. Further research with larger and more diverse populations is needed to confirm these findings and explore the system's long-term impact.

In summary, the PDCP systems support clinical decision-making and operational efficiency and promote the accurate execution of a wide range of clinical activities in patient management in health care settings.

Acknowledgments

This study was supported by UpHill, S.A., which provided funding as part of the main author's PhD studies. The funder's role was limited to facilitating the research activities and data collection. UpHill, S.A. was not involved in the study design, data analysis or interpretation, writing of the manuscript, or the decision to submit the article for publication.

Conflicts of Interest

All authors of this study are employees of UpHill, S.A. This affiliation may represent a potential conflict of interest. Efforts were made to ensure the study's design, data collection, analysis, and interpretation were conducted impartially and without undue influence from the author's employment status. The findings and conclusions presented in this manuscript are based on objective data and aim to contribute to the scientific understanding of clinical decision support systems.

Multimedia Appendix 1
Health Care Professionals Questionnaire.
[DOCX File, 65 KB - humanfactors_v12i1e68581_app1.docx]

Multimedia Appendix 2
UpHill Route V3 graphical interface.
[DOCX File, 222 KB - humanfactors_v12i1e68581_app2.docx]

Multimedia Appendix 3
Ethics committee approval.
[DOCX File, 403 KB - humanfactors_v12i1e68581_app3.docx]

Multimedia Appendix 4
Decision categorisation presented by care pathway.
[PNG File, 158 KB - humanfactors_v12i1e68581_app4.png]

Multimedia Appendix 5
Aggregated clinical pathway supported decisions categorized as by DICTUM taxonomy.
[DOCX File, 67 KB - humanfactors_v12i1e68581_app5.docx]

Multimedia Appendix 6
Table with detailed information on decision categorization presented by care pathway.
[DOCX File, 22 KB - humanfactors_v12i1e68581_app6.docx]

References

1. El Asmar ML, Dharmayat KI, Vallejo-Vaz AJ, Irwin R, Mastellos N. Effect of computerised, knowledge-based, clinical decision support systems on patient-reported and clinical outcomes of patients with chronic disease managed in primary care settings: a systematic review. BMJ Open 2021 Dec 22;11(12):e054659. [doi: 10.1136/bmjopen-2021-054659] [Medline: 34937723]
2. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. NPJ Digit Med 2020;3(17):17. [doi: 10.1038/s41746-020-0221-y] [Medline: 32047862]
3. Chen Z, Liang N, Zhang H, et al. Harnessing the power of clinical decision support systems: challenges and opportunities. Open Heart 2023 Nov 28;10(2):e002432. [doi: 10.1136/openhrt-2023-002432] [Medline: 38016787]

XSL•FO
RenderX

4. Rotter T, Kinsman L, James E, et al. Clinical pathways: effects on professional practice, patient outcomes, length of stay and hospital costs. Cochrane Database Syst Rev 2010 Mar 17;17(3):CD006632. [doi: 10.1002/14651858.CD006632.pub2] [Medline: 20238347]
5. Allen D, Gillen E, Rixson L. The effectiveness of integrated care pathways for adults and children in health care settings: a systematic review. JBI Libr Syst Rev 2009;7(3):80-129. [doi: 10.11124/01938924-200907030-00001] [Medline: 27820426]
6. Rotter T, Jong RB, Evans-Lacko S, Ronellenfitsch U, Kinsman L, et al. Clinical pathways as a quality strategy. In: Busse R, Klazinga N, Panteli D, editors. Improving Healthcare Quality in Europe: Characteristics, Effectiveness and Implementation of Different Strategies [Internet]: European Observatory on Health Systems and Policies; 2019:12 URL: https://www.ncbi.nlm.nih.gov/books/NBK549262/ [accessed 2025-05-09]
7. McBeath KCC, Angermann CE, Cowie MR. Digital technologies to support better outcome and experience of care in patients with heart failure. Curr Heart Fail Rep 2022 Jun;19(3):75-108. [doi: 10.1007/s11897-022-00548-z] [Medline: 35486314]
8. McDonagh TA, Metra M, Adamo M, et al. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. Eur Heart J 2021 Sep 21;42(36):3599-3726. [doi: 10.1093/eurheartj/ehab368] [Medline: 34447992]
9. Chen W, O'Bryan CM, Gorham G, et al. Barriers and enablers to implementing and using clinical decision support systems for chronic diseases: a qualitative systematic review and meta-aggregation. Implement Sci Commun 2022 Jul 28;3(1):81. [doi: 10.1186/s43058-022-00326-x] [Medline: 35902894]
10. Ofstad EH, Frich JC, Schei E, Frankel RM, Gulbrandsen P. What is a medical decision? A taxonomy based on physician statements in hospital encounters: a qualitative study. BMJ Open 2016 Feb 11;6(2):e010098. [doi: 10.1136/bmjopen-2015-010098] [Medline: 26868946]
11. Román-Villarán E, Alvarez-Romero C, Martínez-García A, et al. A personalized ontology-based decision support system for complex chronic patients: retrospective observational study. JMIR Form Res 2022 Aug 2;6(8):e27990. [doi: 10.2196/27990] [Medline: 35916719]
12. Funer F, Schneider D, Heyen NB, et al. Impacts of clinical decision support systems on the relationship, communication, and shared decision-making between health care professionals and patients: multistakeholder interview study. J Med Internet Res 2024 Aug 23;26:e55717. [doi: 10.2196/55717] [Medline: 39178023]
13. Huang S, Liang Y, Li J, Li X. Applications of clinical decision support systems in diabetes care: scoping review. J Med Internet Res 2023 Dec 8;25:e51024. [doi: 10.2196/51024] [Medline: 38064249]
14. Garg AX, Adhikari NKJ, McDonald H, et al. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. JAMA 2005 Mar 9;293(10):1223-1238. [doi: 10.1001/jama.293.10.1223] [Medline: 15755945]
15. Buzancic I, Koh HJW, Trin C, et al. Do clinical decision support tools improve quality of care outcomes in the primary prevention of cardiovascular disease: a systematic review and meta-analysis. Am J Prev Cardiol 2024 Dec;20(100855):100855. [doi: 10.1016/j.ajpc.2024.100855] [Medline: 39416379]
16. Connell A, Black G, Montgomery H, et al. Implementation of a digitally enabled care pathway (part 2): qualitative analysis of experiences of health care professionals. J Med Internet Res 2019 Jul 15;21(7):e13143. [doi: 10.2196/13143] [Medline: 31368443]
17. Connell A, Raine R, Martin P, et al. Implementation of a digitally enabled care pathway (part 1): impact on clinical outcomes and associated health care costs. J Med Internet Res 2019 Jul 15;21(7):e13147. [doi: 10.2196/13147] [Medline: 31368447]
18. Jones SS, Rudin RS, Perry T, Shekelle PG. Health information technology: an updated systematic review with a focus on meaningful use. Ann Intern Med 2014 Jan 7;160(1):48-54. [doi: 10.7326/M13-1531] [Medline: 24573664]

Abbreviations

CDSS: clinical decision support systems
EHR: electronic health record
HCP: health care professional
PDCP: personalized digital care pathway

Edited by A Parush; submitted 09.11.24; peer-reviewed by A Panahi, R Vidal-Perez; revised version received 03.04.25; accepted 06.04.25; published 15.05.25.

Please cite as:
Rodrigues D, Jasmins C, Ladeiras-Lopes R, Patrao L, Rodrigues EF
Personalized Digital Care Pathways Enable Enhanced Patient Management as Perceived by Health Care Professionals: Mixed-Methods Study
JMIR Hum Factors 2025;12:e68581
URL: https://humanfactors.jmir.org/2025/1/e68581
doi:10.2196/68581

XSL•FO
RenderX

© David Rodrigues, Clara Jasmins, Ricardo Ladeiras-Lopes, Luis Patrao, Eduardo Freire Rodrigues. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 15.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

ChatGPT-Delivered Physical Activity Intervention for Children With Autism Spectrum Disorder: Pre-Post Feasibility Study

Uğur Aydemir, PhD

Faculty of Sport Sciences, Bayburt University, 26 Muteber Street, Bayburt, Turkey

Corresponding Author:
Uğur Aydemir, PhD
Faculty of Sport Sciences, Bayburt University, 26 Muteber Street, Bayburt, Turkey

Abstract

Background: The use of digital technologies, such as mobile apps, Zoom (Zoom Communications), virtual reality, and video games, to promote physical activity in individuals with autism spectrum disorder (ASD) has been increasing. However, there are no studies using ChatGPT (OpenAI), a popular tool in recent years, for promoting physical activity in children with ASD.

Objective: This study aimed to evaluate the feasibility and potential effectiveness of ChatGPT-delivered physical activity interventions in children with ASD.

Methods: A total of 26 families (parent-child dyads) participated in the study. Families were randomly assigned to an application group (n=13) and a control group (n=13). In the application group, parents implemented physical activities recommended by ChatGPT for their children with ASD. Data were collected using the Leisure Time Exercise Questionnaire (LTEQ) and a feasibility questionnaire.

Results: Parents reported that ChatGPT-delivered physical activities were a feasible intervention to increase physical activity levels in children with ASD. They also found the activity content suggested by ChatGPT to be interesting and useful. LTEQ measurements corroborated these findings, showing a significant increase in the physical activity levels of children in the intervention group after the intervention.

Conclusions: The results suggest that ChatGPT-delivered physical activities could be a promising intervention to enhance physical activity in children with ASD. Further investigation is warranted.

(JMIR Hum Factors 2025;12:e71119) doi:10.2196/71119

KEYWORDS

autism spectrum disorder; physical activity; parent; ChatGPT; child

Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder that commonly emerges in early childhood and is characterized by the deficiencies in social interaction and communication skills and the restricted-repetitive behavioral patterns [1]. ASD negatively affects the behavioral, social, cognitive, and mental functions of children diagnosed with this disorder [2], as well as their physical activity levels [3]. Recent studies have revealed that children with ASD are less physically active than their typically developing peers [4], and that these children tend to lead sedentary lives [5]. Many individuals with disabilities, including children with ASD, spend a lot of time in sedentary behaviors such as watching television [6]. High levels of inactivity, a concern for all children, is an important issue to address since children with ASD have fewer opportunities to participate in physical activity [7,8].

Participation in physical activity provides many benefits to children with ASD in terms of physical and mental health [9-11]. Studies examining physical activities in children with ASD have revealed that physical activities can have a positive effect on manipulative and locomotor skills [12-14]. At the same time, some studies have found that physical activity helps increase the self-confidence and self-efficacy of children with ASD [10,14]. Furthermore, some studies have shown that stress and anxiety levels of children with ASD, who participate in physical activity, decrease [15] while their social and communication skills improve [16,17].

Despite the numerous health and social benefits of regular physical activity, there are many potential barriers that limit the participation of children with ASD in physical activity [9]. To explain the barriers that cause physical activity deficiencies in children with ASD, some studies focus on characteristic features of ASD, such as inadequate motor skills (joint flexibility, balance, speed, etc), biophysical behaviors (inattention and hyperactivity), and communication disorders [18-20]. Other studies have drawn attention to the limitations in the sustainability of current physical activity interventions for children with ASD [21]. Current physical activity interventions aimed at increasing the physical activity level of children with

XSL•FO
RenderX

ASD often require specialist support (teacher, medical clinician, or physiotherapist), costly specialized equipment, and unnatural environments (school, horse, pool, or equine therapy center, etc) [21]. In most cases, access to existing physical activity programs is limited to people living in areas where specialists provide services [22], and especially during global crises such as COVID-19 pandemic, children with ASD may have limited access to physical activity [3]. For this reason, it is stated that there is a need for an alternative physical activity presentation that children with ASD can do at home with their parents and reach a wider audience.

Recent studies have suggested that internet technologies, programs, and apps may have significant potential in promoting physical activity in children with ASD and creating a more sustainable physical activity presentation [21,23,24]. Among these apps, conversational agents using large language models (LLMs), which have recently become widely used worldwide, have attracted attention as powerful information search tools for children with ASD [25]. Their ability to produce contextually relevant and consistent text positions conversational agents using LLM as potential agents, especially for educational content [26]. Extensive information can be easily accessed anytime and anywhere by simply typing a query about individuals with ASD into conversational agents, usually exemplified by ChatGPT (OpenAI) [27,28]. Conversational agents using LLM provide useful information about daily living for individuals with ASD, such as daily tasks, personal care, and social interaction [29,30].

Although there are different conversational agents such as ChatGPT, BERT (Google), Cipherbot, and DeepSeek in the literature, it has been seen that ChatGPT is frequently used for children with ASD. These studies [31-35] included important information about application protocols and ChatGPT prompts. For this reason, this study used ChatGPT to deliver physical activities to parents of children with ASD. Released by OpenAI in November 2022, ChatGPT is an open AI language model that produces human-like responses to text-based prompts [36]. ChatGPT, which can run on a phone, tablet, or computer, can understand responses in a variety of languages, write stories of different types and lengths, summarize information in complex texts, provide explanations on various topics, and even refuse to respond to inappropriate prompts [37]. ChatGPT is also a highly advanced app in that it can provide a continuous dialogue by remembering what the user has said before in the conversation thread [38]. Although limited in number, studies [31-35] have revealed that ChatGPT can be a useful tool for improving the social skills of children with ASD or for providing information to parents about children with ASD. Although it is a new topic, there is no study evaluating the effectiveness and feasibility of ChatGPT-delivered physical activities in children with ASD. Considering this gap in the literature, the aim of this feasibility study was to examine the effectiveness and feasibility of ChatGPT-delivered physical activities to increase the physical activity level of children with ASD.

Methods

Study Design

This study was designed to evaluate the feasibility of ChatGPT-delivered physical activities to increase the physical activity level of children with ASD. A feasibility study examines whether an intervention is suitable for further testing [39] and allows the intervention procedure to be used in future studies [21]. The feasibility process of the study was based on previous studies that used feasibility measures to evaluate a web-based intervention [21,22,24,40-42] and consisted of 5 stages (Textbox 1).

Textbox 1. Feasibility process of ChatGPT-delivered physical activity.

- Stage 1 (assessment for participant eligibility): The researcher collaborated with a local special education and rehabilitation center in Ankara to recruit parents of children with autism spectrum disorder as participants.
- Stage 2 (ChatGPT training for parents): Parents attended three 40-minute training sessions that included information about ChatGPT and physical activity.
- Stage 3 (preintervention measurements): A week before the intervention, the physical activity level of children with autism spectrum disorder.
- Stage 4 (intervention [4 weeks]): Parents practiced the physical activities recommended by ChatGPT for 40 minutes, 3 days a week for 4 weeks.
- Stage 5 (postintervention measurements): A week after the intervention, the physical activity level of children with autism spectrum disorder and parents' opinions on ChatGPT-delivered physical activities were evaluated.

Participants

The participants of the study were recruited from Ankara province. The criterion sampling method was used in determining the participants [43]. Inclusion criteria were individuals agreeing to participate in the study voluntarily, having a child diagnosed with ASD according to the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* criteria, and parents and children not having any health problems that prohibit physical activity. In line with these criteria, support was received from a special education and rehabilitation center to identify parents. With the permission of the director of the association, the primary researcher interviewed the participant candidates on the phone and explained the aim of the study and participation process to them. Consent forms were received from all parents who voluntarily participated in the study via email. A total of 26 parents (9 mothers and 17 fathers) and their children with ASD participated in the study. All parents filled out the demographic information form and physical activity questionnaire (pretest) before the intervention and the feasibility measures and physical activity questionnaire (posttest) after the intervention. The characteristics of the parents and their children with ASD are presented in Table 1.

XSL•FO
RenderX

Table . The characteristics of the participants.

Characteristics	Groups	
	Application (n=13)	Control (n=13)
Children (n=26)		
Sex, n (%)		
Female	6 (46)	8 (62)
Male	7 (54)	5 (38)
Age (years), mean (SD)	14.38 (3.06)	13.92 (3.52)
Diagnosis, n (%)		
ASD[a]	13 (100)	13 (100)
Additional comorbidities, n (%)		
SD	1 (8)	1 (8)
ID	11 (84)	7 (54)
N	1 (8)	5 (38)
Parents (n=26)		
Gender, n (%)		
Female	5 (38)	4 (31)
Male	8 (62)	9 (69)
Age (years), mean (SD)	42.84 (7.70)	44.07 (7.95)
Education, n (%)		
Secondary school	5 (38)	3 (23)
High school	4 (31)	7 (54)
University	4 (31)	3 (23)
Income, n (%)		
Low	3 (23)	—[b]
Middle	7 (54)	8 (62)
High	3 (23)	5 (38)

[a]ASD: autism spectrum disorder.

[b]Not available.

Intervention

ChatGPT Training for Parents

In this study, parents played an important role by actively participating in ChatGPT-delivered physical activities and helping their children with ASD. Parental participation was an integral part of ChatGPT-delivered physical activities. To prepare parents for this role, the researcher conducted three 40-minute training sessions. Each session lasted until parents demonstrated proficiency. The first session focused on developing parents' skills in using ChatGPT. The researcher explained to parents how to download ChatGPT and how to get physical activity recommendations using ChatGPT. Then, the researcher asked all parents to give ChatGPT the following command: "Create a 4-week, 3-day, 40-minute physical activity program that can be implemented at home for my child who is…years old and has ASD." The researcher discussed the program created by ChatGPT with parents and told them to give the following command to ChatGPT for physical activities they did not understand: "What is…activity? Show me in detail." The researcher emphasized the importance of all parents using the commands suggested by the researcher in order to optimize the prompts. The second session included informing parents about the parts of a physical activity (warm-up, cool-down, and main part). The third session addressed strategies that parents can use when providing physical activity to children with ASD. These strategies generally included information about preparing children with ASD for physical activities (wearing sports clothes, social stories, and preview of activities and equipment), organizing the environment where physical activities will take place (safety, ventilation of the environment and sports pictures), including other members of the family in physical activities (parent and sibling participation), and reward is given. A special WhatsApp (Meta) group was created to answer parents' questions during the implementation of the ChatGPT-delivered physical activities. Parents' questions regarding the implementation of the physical activities and the use of the ChatGPT software were answered instantly via the WhatsApp

XSL•FO
RenderX

group. In addition, at the end of the daily physical activity session, parents chatted in the WhatsApp group to evaluate the effectiveness of the session and explore the experiences of other parents.

ChatGPT-Delivered Physical Activities

All parents, who completed the ChatGPT training, used ChatGPT-delivered physical activities to increase the physical activity levels of their children with ASD. One week before the intervention, the researcher asked each parent to give ChatGPT an appropriate command so that the physical activities were age-appropriate for children with ASD and could be easily carried out in the home environment. Although the content of ChatGPT varied, it generally suggested physical activities to parents consisting of 3 parts: warm-up, main part, and cool-down (Table 2). The researcher asked parents to ask ChatGPT again about the activities they did not understand about the activities suggested by ChatGPT until the intervention began. In addition, the researcher supported the parents by explaining the activities they did not understand in the WhatsApp group. The parents implemented the intervention for 40 minutes, 3 days a week for 4 weeks. The parents and the children with ASD participated in the activities together. During the implementation process, the researcher immediately answered the parents' questions about the activities and strategies in the WhatsApp group. In addition, at the end of the daily physical activity session, the parents chatted in the WhatsApp group to evaluate the effectiveness of the session and prepare them for the next session. In the interview, the researcher asked the parents questions about the extent to which physical activities were performed, who participated in physical activities, and what the benefits of physical activities were. In this way, the researcher tried to verify to what extent and how physical activities were carried out. All parents and children with ASD completed the 4-week intervention. The recommended physical activities in a typical week are shown in Table 2.

Table . Physical activities in a typical week recommended by ChatGPT.

Day and physical activity	Content
First day (movement and balance)	
Warming up (10 min)	Opening and closing the arms up and to the sides; slow walking with the knees pulled in; slow running in place.
Main exercise	
Balance line (5 min)	Draw a line on the ground and have your child walk along this line.
One leg on pillow (5 min)	Try to keep your balance on the pillow.
Rolling with ball (5 min)	Sit on the floor with your child and roll a large ball to each other.
Jumping (5 min)	Use a ring or cushion to bounce in and out.
Cooling down (10 min)	Slow walking and stretching movements (stretching the arms and stretching the legs).
Second day (coordination and movement)	
Warming up (10 min)	Stretching arms and legs; animal walks (bear walk, frog leap).
Main exercise	
Throwing and catching balls (10 min)	You and your child throw and catch a large ball to each other.
Rolling on the mat (5 min)	Rolling back and forth on a large cushion.
Imitation games (5 min)	Exercises of standing on one leg and finding balance while turning slightly.
Cooling down (10 min)	Slow breathing and yawning.
Third day (game day and free movement)	
Warming up (10 min)	Free dance and simple movements with rhythmic music.
Main exercise	
Balloon games (10 min)	Play a game with your child by holding balloons in the air.
Obstacle course (10 min)	Set up a simple track with cushions and chairs inside the house and have your child walk through the track (crawling, walking, and jumping).
Cooling down (10 min)	Slow paced walking and light stretching.

Data Collection

Overview

Data were obtained with the following 2 tools: (1) Feasibility Questionnaire and (2) Leisure Time Exercise Questionnaire (LTEQ). Measurements were made 1 week before and after the intervention. Data were collected via Google forms. The researcher shared the Google form links in the WhatsApp group and asked all parents to complete the forms in the links within 1 week. A personal information form was used to obtain

XSL•FO
RenderX

information about the demographic characteristics of parents (age, gender, and level of education) and children with ASD (age, gender, additional comorbidities, and number of people in the household).

Feasibility Questionnaire

Parents rated their ChatGPT-delivered physical activities via the Feasibility Questionnaire after the 4-week intervention. Questionnaire questions were designed considering previous studies investigating the physical activity level of children with ASD [3,21]. Questionnaire questions included the following: (1) how would you rate your overall experience with ChatGPT-delivered physical activities, (2) how interesting were the ChatGPT-delivered physical activities, (3) how beneficial were the ChatGPT-delivered physical activities for your child's physical activity level, and (4) to what extent did you learn about physical activity through ChatGPT? The Feasibility Questionnaire was rated on a 5-point Likert-type scale (very satisfied to not very satisfied, very to a little, very useful to not useful at all).

Leisure Time Exercise Questionnaire

To assess the physical activity level of children with ASD before and after the intervention, the researcher used the LTEQ developed by Godin and Shephard [44]. The LTEQ has been frequently used in previous studies [3,21,22,45-47] to determine the leisure time physical activity level of individuals with ASD. Memari et al [46] showed that the LTEQ has a good test-retest reliability score that can be used to determine the physical activity level of children with ASD. The questionnaire includes questions about at least 15 minutes of leisure time physical activity performed in the last 7 days and aims to determine the number of times of strenuous-intensity physical activities, moderate-intensity physical activities, and mild-intensity physical activities performed in the last week. To calculate the total score of the scale, high-intensity activities are multiplied by 9, moderate-intensity activities by 5, and light activities by 3, and all are added together. The formula is as follows: weekly leisure time activity score = (9 × strenuous intensity) + (5 × moderate intensity) + (3 × mild intensity). The calculated values are added up and generally evaluate the individual's leisure time activity. In this evaluation, scores of 24 and above are classified as "Active," 14 - 23 as "Moderately active," and 13 and below as "Not sufficiently active" [48]. The Turkish adaptation study of the LTEQ was conducted by Yerlisu-Lapa et al [49]. As a result of the exploratory factor analysis conducted to determine the factor structure of the questionnaire, the total correlations of the items were determined as 0.80, 0.76, and 0.65 for all 3 items, respectively, and it was found that 55% of the total variance was explained and gathered under a single factor. Test-retest reliability analysis yielded a Pearson correlation coefficient of r = 0.84 for the entire Leisure-Time Exercise Questionnaire (LTEQ), and 0.80, 0.76, and 0.72 for its vigorous, moderate, and mild intensity subscales, respectively. To assess equivalent form reliability, a comparison was made between the LTEQ and the International Physical Activity Questionnaire–Short Form (IPAQ-SF). The Spearman correlation coefficient for the relationship between the two forms was found to be ρ = 0.92, indicating a strong positive correlation [49].

Statistical Analysis

Descriptive statistics (mean, SD, percentage, and frequency) were used in the study for feasibility measurements, demographic information, and frequency of participation in physical activities. A 2-way mixed ANOVA (analysis of variance; 2 groups × 2 time points) was used to determine the effects of ChatGPT-delivered physical activities (from baseline to intervention) on the physical activity level of children with ASD. All these statistical processes were carried out with the SPSS package program (version 25.0; IBM Corp).

Ethical Considerations

The research was approved by the Bayburt University noninterventional clinical research ethics committee (number 209) on July 23, 2024. This study was approved by the Ethics Committee of Bayburt University (no. E-15604681-100-215816). Informed consent was obtained from the legal guardians of the participants. Participation was voluntary, and the participants had the right to withdraw from the study at any time. For secondary analyses using existing data with primary consent, the original consent or institutional review board approval covers secondary analyses without additional consent. Data have been anonymized and deidentified (pseudonyms were used in the study to protect the identities of the participants. Participant names were listed as P1, P2, P3, and so on. No financial compensation was given to the participants, but small symbolic rewards (stickers, reward cards, etc.) were given.

Results

Parental Opinions and Feedback on Feasibility of ChatGPT-Delivered Physical Activity

All parents expressed that they were very happy to participate in the ChatGPT-delivered physical activities. According to the parents, the ChatGPT-delivered physical activity content was interesting and very useful. Parents stated that the ChatGPT-delivered physical activities positively affected the physical activity levels of their children with ASD. Parents reported that they gained important information about physical activity through ChatGPT. In total, 92% (12/13) of the parents participated in the discussions in the WhatsApp group and shared pictures and videos of their children in the group.

Physical Activity Level

In order to determine the effect of ChatGPT-delivered physical activities on the physical activity level of children with ASD, a 2-way mixed ANOVA (2 groups × 2 time points) test was performed (Table 3). While the mean physical activity score of children with ASD in the application group was 6.69 (not sufficiently active) before the intervention, their physical activity score increased to 34 (active) after the intervention. In contrast, while the mean physical activity score of children with ASD in the control group was 7.46 (not sufficiently active) before the intervention, their physical activity score decreased to 7.23 (not sufficiently active) after the intervention. Regarding the mean

XSL•FO
RenderX

physical activity scores, 2-way mixed ANOVA revealed a significant group effect ($F_{1, 24}$=108.769; P<.05, ηp^2=.819), a significant time effect ($F_{1, 24}$=341.333; P<.05, ηp^2=.934) and a significant group × time interaction ($F_{1, 24}$=353.069; P<.05, ηp^2=.936) on the physical activity level of the participants in the application and control groups.

Table . Analysis of variance results related to the pretest-posttest physical activity scores.

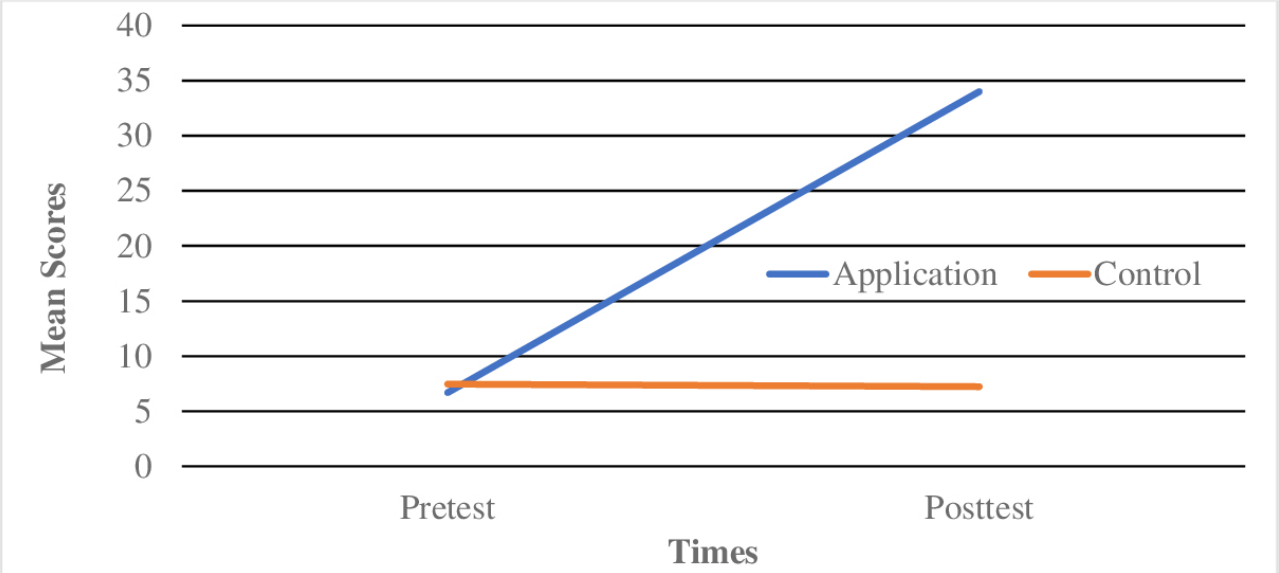
Source	Type III sum of squares	Mean square	*F* test (*df*)	*P* value	Partial eta squared (ηp^2)
Between-Subjects	9969.231				
Group (A/C)	2197.000	2197.000	108.769 (1, 24)	<.001[a]	.819
Error	484.769	20.199	—[b]	—	—
Within-subjects	5014.999				
Time (pretest-posttest)	2382.769	2382.769	341.333 (1, 24)	<.001[a]	.934
Time group	2464.692	2464.692	353.069 (1, 24)	<.001[a]	.936
Error	167.538	6.981	—	—	—

[a]Significant at P<.05.

[b]Not applicable.

Since ANOVA results reveal a significant group × time interaction, the adjusted Bonferroni value was used to determine whether there was a change in application and control groups over time. As reported in Figure 1, pair-wise comparisons revealed the significant increase in the preintervention physical activity scores of children with ASD in the application group compared with their postintervention physical activity scores (adjusted Bonferroni*: P*<.05, difference: +27.308); this was not the case in the control group (adjusted Bonferroni*: P*>.05, difference: −0.231).

Figure 1. Mean physical activity scores of the groups in preintervention and postintervention.



Discussion

Principal Findings

The aim of this study was to examine the effects of ChatGPT-delivered physical activity intervention on the physical activity level of children with ASD. The results revealed that there was a positive and significant increase in the physical activity level of children with ASD in the application group after ChatGPT-delivered physical activities compared with children with ASD in the control group. Current physical activity interventions that aim to increase the physical activity level of children with ASD generally include physical activity programs such as karate [50], swimming [51], summer camp [52], fitness [53], and therapeutic horseback riding [54]. Systematic reviews suggest that participation in existing physical activity programs has a positive effect on the developmental domains of children with ASD [10,12]. Despite these promising findings, there are some limitations in the sustainability of current physical activity interventions [24]. Current physical activity interventions are often disrupted due to dependency on professionals, unnatural environments, and high costs [21]. In most cases, access to current physical activity programs is limited to people living in

XSL•FO
RenderX

areas where professionals provide services [22]. In fact, in situations such as school closures due to COVID-19 pandemic, children with ASD have completely restricted access to physical activity [55]. Realizing this, researchers have recently begun to show interest in alternative physical activity interventions that can reach more people and that children with ASD can do at home with their families [3]. Various studies have been conducted on sustainable physical activity interventions, especially for children with ASD who cannot access physical activity [21,24,56]. The results of the studies revealed that physical activities provided remotely via WhatsApp and Facebook (Meta) are effective interventions to increase physical activity in children with ASD.

As seen in the literature, alternative physical activity interventions for children with ASD have been focused on social media platforms. Although it is a new topic, no study has been found using ChatGPT to improve the physical activity level of children with ASD. To our knowledge, this is the first study to evaluate the feasibility of ChatGPT-delivered physical activities to increase the physical activity level of children with ASD. In the study, it was taken into consideration that parents have a critical role in the participation of their children with ASD in physical activity [57,58], and parents assumed an important responsibility in the entire process. The parents who participated in the study reported that ChatGPT-delivered physical activities were a feasible and useful intervention to increase the physical activity level of children with ASD. Despite these positive findings, the parents did not understand some of the physical activity content in the WhatsApp group and consulted the researcher about the content. The researcher asked the parents to reprompt ChatGPT regarding the content they did not understand, and the problem was solved in this way. In conclusion, this study demonstrated that ChatGPT-delivered physical activities effectively increased the physical activity level of children with ASD. The findings provided preliminary evidence that ChatGPT-delivered physical activities could be an alternative physical activity that can be easily implemented in the home environment by parents.

Limitations of the Study

Although the study is a pioneering study examining the effectiveness of ChatGPT-delivered physical activities for children with ASD, it has some limitations. Since the study was conducted as a feasibility study, ChatGPT-delivered physical activities were implemented for only 4 weeks. In addition, a follow-up test was not conducted for the physical activity level of children with ASD. Therefore, it may be difficult to interpret the durability of the effects of ChatGPT-delivered physical activities on the physical activity level of children with ASD. Since the sample consisted of 26 families (parent and child dyads), the generalizability of the survey results regarding the physical activity level of children with ASD may be limited. ChatGPT training was organized to prepare the parents in the intervention group to implement physical activities. However, no interviews were conducted regarding the ChatGPT training to determine the perceived usefulness of this training. We relied on the self-report measures of the parents regarding the diagnoses and physical activity level of the children with ASD who participated in the study, did not make any observations, and excluded the perceptions of the children with ASD about the activity program.

Implications for Future Research

Future studies may use measurements such as those obtained via observation or pedometers, in addition to the LTEQ, to assess the physical activity levels of children with ASD during ChatGPT-delivered physical activities. This study does not provide any information on comparing ChatGPT-delivered physical activities with other physical activity programs. Future studies may compare the effects of ChatGPT-delivered and face-to-face physical activities. Future studies may include interviews with parents about the benefits of ChatGPT training organized for parents. We used ChatGPT version 4.0 in the study. Since this version only included verbal information, parents did not understand some of the content related to physical activity. Future studies may provide a more understandable process for parents by using different versions of ChatGPT that include image and video support. Furthermore, future studies can use different conversational agents such as BERT, Cipherbot, and DeepSeek to deliver physical activities to parents of children with ASD.

Acknowledgments

The author would like to thank the individuals with autism spectrum disorders and their parents who participated in the study. The author also expresses gratitude to the Special Education and Rehabilitation Center administrators and personnel for their trust, openness, and cooperation.

Data Availability

The participants of this study did not give written consent for their data to be shared publicly, so due to the sensitive nature of the research, supporting data are not available.

Authors' Contributions

UA contributed to conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, software, supervision, validation, visualization, writing – original draft, and writing – review and editing.

XSL•FO
RenderX

Conflicts of Interest

None declared.

References

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 5th edition: American Psychiatric Association; 2013.
2. Newschaffer CJ, Croen LA, Daniels J, et al. The epidemiology of autism spectrum disorders. Annu Rev Public Health 2007;28:235-258. [doi: 10.1146/annurev.publhealth.28.021406.144007] [Medline: 17367287]
3. Yarımkaya E, Esentürk OK, İlhan EL, Kurtipek S, Işım AT. Zoom-delivered physical activities can increase perceived physical activity level in children with autism spectrum disorder: a pilot study. J Dev Phys Disabil 2023;35(2):189-207. [doi: 10.1007/s10882-022-09854-9] [Medline: 35677592]
4. Lee D, Cothran DJ, Kennedy JM, et al. Association between physical activity, sedentary time, and technology use in autistic adults. J Dev Phys Disabil 2024:1-19. [doi: 10.1007/s10882-024-09983-3]
5. Thompson C, Brook M, Hick S, Miotti C, Toong R, McVeigh JA. Physical activity, sedentary behaviour and their correlates in adults with autism spectrum disorder: a systematic review. Rev J Autism Dev Disord 2023 Sep;10(3):546-562. [doi: 10.1007/s40489-022-00305-x]
6. Must A, Phillips SM, Curtin C, et al. Comparison of sedentary behaviors between children with autism spectrum disorders and typically developing children. Autism 2014 May;18(4):376-384. [doi: 10.1177/1362361313479039] [Medline: 24113339]
7. Garcia-Pastor T, Salinero JJ, Theirs CI, Ruiz-Vicente D. Obesity status and physical activity level in children and adults with autism spectrum disorders: a pilot study. J Autism Dev Disord 2019 Jan;49(1):165-172. [doi: 10.1007/s10803-018-3692-9] [Medline: 30043355]
8. Pan CY, Tsai CL, Hsieh KW. Physical activity correlates for children with autism spectrum disorders in middle school physical education. Res Q Exerc Sport 2011 Sep;82(3):491-498. [doi: 10.1080/02701367.2011.10599782] [Medline: 21957708]
9. Healy S, Marchand G, Williams E. “I’m not in this alone” the perspective of parents mediating a physical activity intervention for their children with autism spectrum disorder. Res Dev Disabil 2018 Dec;83:160-167. [doi: 10.1016/j.ridd.2018.08.014] [Medline: 30218986]
10. Lang R, Koegel LK, Ashbaugh K, Regester A, Ence W, Smith W. Physical exercise and individuals with autism spectrum disorders: a systematic review. Res Autism Spectr Disord 2010 Oct;4(4):565-576. [doi: 10.1016/j.rasd.2010.01.006]
11. Petrus C, Adamson SR, Block L, Einarson SJ, Sharifnejad M, Harris SR. Effects of exercise interventions on stereotypic behaviours in children with autism spectrum disorder. Physiother Can 2008;60(2):134-145. [doi: 10.3138/physio.60.2.134] [Medline: 20145777]
12. LaLonde KB, MacNeill BR, Eversole LW, Ragotzy SP, Poling A. Increasing physical activity in young adults with autism spectrum disorders. Res Autism Spectr Disord 2014 Dec;8(12):1679-1684. [doi: 10.1016/j.rasd.2014.09.001]
13. Menear KS, Neumeier WH. Promoting physical activity for students with autism spectrum disorder: barriers, benefits, and strategies for success. Journal of Physical Education, Recreation & Dance 2015 Mar 24;86(3):43-48. [doi: 10.1080/07303084.2014.998395] [Medline: 18858293]
14. Sowa M, Meulenbroek R. Effects of physical exercise on autism spectrum disorders: a meta-analysis. Res Autism Spectr Disord 2012 Jan;6(1):46-57. [doi: 10.1016/j.rasd.2011.09.001]
15. Liamputtong P. Qualitative Research Methods: Oxford University Press; 2009.
16. Gregor S, Bruni N, Grkinic P, et al. Parents’ perspectives of physical activity participation among Canadian adolescents with autism spectrum disorder. Res Autism Spectr Disord 2018 Apr;48:53-62. [doi: 10.1016/j.rasd.2018.01.007]
17. Yarımkaya E, İlhan EL, Karasu N. An investigation of the changes in the communication skills of an individual with autism spectrum disorder participating in peer mediated adapted physical activities. Ankara Univ Educ Sci Fac Spec Educ J 2017;18(2):225-252. [doi: 10.21565/ozelegitimdergisi.319423]
18. Bandini LG, Gleason J, Curtin C, et al. Comparison of physical activity between children with autism spectrum disorders and typically developing children. Autism 2013 Jan;17(1):44-54. [doi: 10.1177/1362361312437416] [Medline: 22807562]
19. Macdonald M, Esposito P, Ulrich D. The physical activity patterns of children with autism. BMC Res Notes 2011 Oct 18;4:422. [doi: 10.1186/1756-0500-4-422] [Medline: 22008607]
20. Pan CY, Frey GC. Physical activity patterns in youth with autism spectrum disorders. J Autism Dev Disord 2006 Jul;36(5):597-606. [doi: 10.1007/s10803-006-0101-6] [Medline: 16652237]
21. Healy S, Marchand G. The feasibility of project CHASE: a Facebook-delivered, parent-mediated physical activity intervention for children with autism. Intl J Disabil Dev Educ 2020 Mar 3;67(2):225-242. [doi: 10.1080/1034912X.2019.1597968]
22. Marchand G, Healy S. Implementation of project CHASE (children with autism supported to exercise): a Facebook-delivered, parent-mediated physical activity intervention. Palaestra 2019;33(3):32-39.
23. Garcia JM, Cathy BS, Garcia AV, et al. Transition of a judo program from in-person to remote delivery during COVID-19 for youth with autism spectrum disorder. Adv Neurodev Disord 2021;5(2):227-232. [doi: 10.1007/s41252-021-00198-7] [Medline: 33718591]

XSL•FO
RenderX

24. Yarımkaya E, Esentürk OK, İlhan EL, Karasu N. A WhatsApp-delivered intervention to promote physical activity in young children with autism spectrum disorder. Int J Dev Disabil 2022;68(5):732-743. [doi: 10.1080/20473869.2021.1887436] [Medline: 36210901]
25. Hoover A, Spengler S. For some autistic people, ChatGPT is a lifeline. Wired. 2023 May 30. URL: https://www.wired.com/story/for-some-autistic-people-chatgpt-is-a-lifeline/ [accessed 2025-06-03]
26. Hang CN, Wei Tan C, Yu PD. MCQGen: a large language model-driven MCQ generator for personalized learning. IEEE Access 2024;12:102261-102273. [doi: 10.1109/ACCESS.2024.3420709]
27. Kasneci E, Sessler K, Küchemann S, et al. ChatGPT for good? On opportunities and challenges of large language models for education. Learn Individ Differ 2023 Apr;103:102274. [doi: 10.1016/j.lindif.2023.102274]
28. Razniewski S, Yates A, Kassner N, Weikum G. Language models as or for knowledge bases. arXiv. Preprint posted online on Oct 10, 2021. [doi: 10.48550/arXiv.2110.04888]
29. Krauss MW, Seltzer MM, Jacobson HT. Adults with autism living at home or in non-family settings: positive and negative aspects of residential status. J Intellect Disabil Res 2005 Feb;49(Pt 2):111-124. [doi: 10.1111/j.1365-2788.2004.00599.x] [Medline: 15634320]
30. Salminen J, Jung SG, Medina J, et al. Using Cipherbot: an exploratory analysis of student interaction with an LLM-based educational chatbot. Presented at: Proceedings of the Eleventh ACM Conference on Learning@Scale; Jul 18-20, 2024; Atlanta, GA, USA p. 279-283. [doi: 10.1145/3657604.3664690]
31. Xu X, Gao T, Chen X, Zhang R, Lin S. Exploring emotion recognition in children with autism spectrum disorder using ChatGPT. Presented at: 2024 6th International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI); Nov 1-3, 2024; Hangzhou, China p. 280-284. [doi: 10.1109/MLBDBI63974.2024.10823703]
32. Li X, Hu C, Li W, et al. Exploiting ChatGPT for diagnosing autism-associated language disorders and identifying distinct features. Res Square. Preprint posted online on May 21, 2024. [doi: 10.21203/rs.3.rs-4359726/v1]
33. Rakap S, Balikci S. Enhancing IEP goal development for preschoolers with autism: a preliminary study on ChatGPT integration. J Autism Dev Disord 2024 Apr 16. [doi: 10.1007/s10803-024-06343-0] [Medline: 38625490]
34. McFayden TC, Bristol S, Putnam O, Harrop C. ChatGPT: artificial intelligence as a potential tool for parents seeking information about autism. Cyberpsychol Behav Soc Netw 2024 Feb;27(2):135-148. [doi: 10.1089/cyber.2023.0202] [Medline: 38181176]
35. Moraiti I, Drigas A. AI tools like ChatGPT for people with neurodevelopmental disorders. Int J Onl Eng 2023;19(16):145-155. [doi: 10.3991/ijoe.v19i16.43399]
36. Pocock K. What is ChatGPT? Why you need to care about GPT-4. PC Guide. URL: https://www.pcguide.com/apps/what-is-chat-gpt/ [accessed 2024-01-27]
37. Kocoń J, Cichecki I, Kaszyca O, et al. ChatGPT: Jack of all trades, master of none. Information Fusion 2023 Nov;99:101861. [doi: 10.1016/j.inffus.2023.101861]
38. Jiao W, Wang W, Huang JT, Wang X, Shi S, Tu Z. Is ChatGPT a good translator? Yes with GPT-4 as the engine. arXiv. Preprint posted online on Nov 2, 2023. [doi: 10.48550/arXiv.2301.08745]
39. Bowen DJ, Kreuter M, Spring B, et al. How we design feasibility studies. Am J Prev Med 2009 May;36(5):452-457. [doi: 10.1016/j.amepre.2009.02.002] [Medline: 19362699]
40. Muntaner-Mas A, Vidal-Conti J, Borràs PA, Ortega FB, Palou P. Effects of a Whatsapp-delivered physical activity intervention to enhance health-related physical fitness components and cardiovascular disease risk factors in older adults. J Sports Med Phys Fitness 2017;57(1-2):90-102. [doi: 10.23736/S0022-4707.16.05918-1] [Medline: 26364690]
41. Newton RL Jr, Marker AM, Allen HR, et al. Parent-targeted mobile phone intervention to increase physical activity in sedentary children: randomized pilot trial. JMIR Mhealth Uhealth 2014 Nov 10;2(4):e48. [doi: 10.2196/mhealth.3420] [Medline: 25386899]
42. Wójcicki TR, Grigsby-Toussaint D, Hillman CH, Huhman M, McAuley E. Promoting physical activity in low-active adolescents via Facebook: a pilot randomized controlled trial to test feasibility. JMIR Res Protoc 2014 Oct 30;3(4):e56. [doi: 10.2196/resprot.3013] [Medline: 25357008]
43. Creswell JW. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches, 3rd edition: SAGE; 2009.
44. Godin G, Shephard RJ. A simple method to assess exercise behavior in the community. Can J Appl Sport Sci 1985 Sep;10(3):141-146. [Medline: 4053261]
45. Hamm J, Yun J. The motivational process for physical activity in young adults with autism spectrum disorder. Disabil Health J 2018 Oct;11(4):644-649. [doi: 10.1016/j.dhjo.2018.05.004] [Medline: 29934100]
46. Memari AH, Panahi N, Ranjbar E, et al. Children with autism spectrum disorder and patterns of participation in daily physical and play activities. Neurol Res Int 2015;2015:531906. [doi: 10.1155/2015/531906] [Medline: 26171247]
47. Russell S, Healy S, Braithwaite RE. Hobby preferences and physical activity participation among children with and without autism spectrum disorder. Eur J Adapt Phys Act 2019 Jan 25;11(2):8-8. [doi: 10.5507/euj.2018.008]
48. Godin G. The Godin-Shephard leisure-time physical activity questionnaire. The Health & Fitness Journal of Canada 2011;4(1). [doi: 10.14288/hfjc.v4i1.82]
49. Yerlisu-Lapa T, Certel Z, Kaplan K, Yağar G. Validity and reliability study of the leisure time exercise questionnaire for adolescent students. Journal of Research in Education and Teaching 2016 Jun;5(5):1-9 [FREE Full text]

XSL•FO
RenderX

50. Bahrami F, Movahedi A, Marandi SM, Sorensen C. The effect of karate techniques training on communication deficit of children with autism spectrum disorders. J Autism Dev Disord 2016 Mar;46(3):978-986. [doi: 10.1007/s10803-015-2643-y] [Medline: 26577688]
51. Pan CY. Effects of water exercise swimming program on aquatic skills and social behaviors in children with autism spectrum disorders. Autism 2010 Jan;14(1):9-28. [doi: 10.1177/1362361309339496] [Medline: 20124502]
52. Brookman L, Boettcher M, Klein E, Openden D, Koegel RL, Koegel LK. Facilitating social interactions in a community summer camp setting for children with autism. J Posit Behav Interv 2003 Oct;5(4):249-252. [doi: 10.1177/10983007030050040801]
53. Pitetti KH, Rendoff AD, Grover T, Beets MW. The efficacy of a 9-month treadmill walking program on the exercise capacity and weight reduction for adolescents with severe autism. J Autism Dev Disord 2007 Jul;37(6):997-1006. [doi: 10.1007/s10803-006-0238-3] [Medline: 17151799]
54. Gabriels RL, Agnew JA, Holt KD, et al. Pilot study measuring the effects of therapeutic horseback riding on school-age children and adolescents with autism spectrum disorders. Res Autism Spectr Disord 2012 Apr;6(2):578-588. [doi: 10.1016/j.rasd.2011.09.007]
55. Esentürk OK. Parents' perceptions on physical activity for their children with autism spectrum disorders during the novel Coronavirus outbreak. Int J Dev Disabil 2021;67(6):446-457. [doi: 10.1080/20473869.2020.1769333] [Medline: 34925775]
56. Esentürk OK, Yarımkaya E. WhatsApp-based physical activity intervention for children with autism spectrum disorder during the novel coronavirus (COVID-19) pandemic: a feasibility trial. Adapt Phys Activ Q 2021 Oct 1;38(4):569-584. [doi: 10.1123/apaq.2020-0109] [Medline: 34193630]
57. Baranowski T. Families and health actions. In: Gochman DS, editor. Handbook of Health Behavior Research I: Personal and Social Determinants: Plenum Press; 1997:179-206.
58. Sallis JF, Prochaska JJ, Taylor WC. A review of correlates of physical activity of children and adolescents. Med Sci Sports Exerc 2000 May;32(5):963-975. [doi: 10.1097/00005768-200005000-00014] [Medline: 10795788]

Abbreviations

ANOVA: analysis of variance
ASD: autism spectrum disorder
LLM: large language model
LTEQ: Leisure Time Exercise Questionnaire

Edited by A Kushniruk; submitted 10.01.25; peer-reviewed by CN Hang; revised version received 29.03.25; accepted 04.05.25; published 11.06.25.

Please cite as:
Aydemir U
ChatGPT-Delivered Physical Activity Intervention for Children With Autism Spectrum Disorder: Pre-Post Feasibility Study
JMIR Hum Factors 2025;12:e71119
URL: https://humanfactors.jmir.org/2025/1/e71119
doi:10.2196/71119

© Uğur Aydemir. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 11.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Using a Mobile Health App (ColonClean) to Enhance the Effectiveness of Bowel Preparation: Development and Usability Study

Hui-Yu Chen[1,2*], RN, MS; Ming-Hsiang Tu[2*], RN, PhD; Miao-Yen Chen[2*], RN, PhD

[1]Endoscopy Center for Diagnosis and Treatment, Taipei Veterans General Hospital, Taipei City, Taiwan

[2]School of Nursing, National Taipei University of Nursing and Health Sciences, Room B631, No. 365, Ming-te Road, Peitou District, Taipei City, Taiwan

[*]all authors contributed equally

Corresponding Author:
Ming-Hsiang Tu, RN, PhD
School of Nursing, National Taipei University of Nursing and Health Sciences, Room B631, No. 365, Ming-te Road, Peitou District, Taipei City, Taiwan

Abstract

Background: Colonoscopy is the standard diagnostic method for colorectal cancer. Patients usually receive written and verbal instructions for bowel preparation (BP) before the procedure. Failure to understand the importance of BP can lead to inadequate BP in 25%-30% of patients. The quality of BP impacts the success of colonoscopy in diagnostic yield and adenoma detection. We developed the "ColonClean" mobile health (mHealth) app for Android devices. It incorporates visual representations of dietary guidelines, steps for using bowel cleansing agents, and observations of the last bowel movement. We used the Technology Acceptance Model to investigate whether the use of the ColonClean mHealth app can improve users' attitudes and behaviors toward BP.

Objective: This study aims to validate the effectiveness of the ColonClean app in enhancing user behavior and improving BP, providing safe and cost-effective outpatient colonoscopy guidance.

Methods: This study uses a structured questionnaire to assess perceived usefulness, perceived ease of use, and users' attitudes and behaviors toward BP regarding the ColonClean mHealth app. A total of 40 outpatients who were physically and mentally healthy and proficient in Chinese were randomly chosen for this study. The data were analyzed using SPSS 25.0, and we used Pearson product-moment correlation and simple regression analysis to predict the perception of ColonClean.

Results: The results showed that 75% (30/40) of participants achieved an "excellent" or "good" level of BP according to the Aronchick Bowel Preparation Scale. Perceived usefulness and perceived ease of use of the ColonClean mHealth app were positively correlated with users' attitudes and behaviors (P<.05).

Conclusions: The ColonClean mHealth app serves as an educational reference and enhances the effectiveness of BP. Users expressed their willingness to use the app again in the future and recommend it to family and friends, highlighting its effectiveness as an educational guide for BP.

(JMIR Hum Factors 2025;12:e58479) doi:10.2196/58479

KEYWORDS

mobile health app; bowel preparation; nursing guidance; technology acceptance model; mHealth; mobile health

Introduction

Colorectal cancer (CRC) is the third most prevalent cancer and the second leading cause of cancer-related deaths worldwide [1-3]. Colonoscopy is the standard diagnostic method for colorectal cancer, and adequate bowel preparation (BP) is a necessary and crucial step to effectively examine the entire intestinal mucosa [4,5]. Currently, BP before colonoscopy is primarily provided by nurses through paper-based nursing instructions and oral explanations of relevant precautions [6,7]. This involves dietary restrictions and the use of bowel cleansing agents to remove feces from the colon and facilitate visual examination by physicians [8,9].

The explanations for BP are often lengthy, complex, and difficult to understand and remember. As a result, the proportion of inadequate BP ranges from 25% to 30% [10]. Poor BP could lead to 42% of adenomas and 27% of advanced adenomas not being diagnosed [11,12]. Inadequate BP also leads to a missed detection of 10.55% for sessile polyps [13], which can develop into CRC and are particularly difficult to identify as they lie flat on the mucosal layer of the colon [9]. Incomplete

XSL•FO
RenderX

colonoscopy increases the risk of cancer and requires repreparation, adding to health care resource costs [14-16].

The European Society of Gastrointestinal Endoscopy guidelines emphasize that good BP is essential for the diagnosis, treatment, and removal of tumors and precancerous lesions, as well as for the reduction of CRC incidence and its mortality rate. It ensures the quality, safety, and effectiveness of colonoscopy in the clinical examination environment. Adequate BP provides optimal visualization for physicians during the examination, facilitating the detection of polyps or other lesions [10,17].

During the COVID-19 pandemic in 2020, the need for a mobile health (mHealth) app became evident [18,19]. Currently, nursing instructions for BP mainly rely on paper-based materials accompanied by oral explanations, multimedia instruction CDs, or telephone interviews [20]. Using mobile apps can enhance patient education for bowel cleansing [21,22]. Therefore, our research team conducted a literature review and assessed current needs to design and build the "ColonClean" mHealth app for the Android operating system. The purpose of this study is to validate the effectiveness of the ColonClean mHealth app in enhancing users' attitudes and behaviors toward BP through the Technology Acceptance Model (TAM) by examining the correlations among perceived usefulness, perceived ease of use, usage attitudes, and actual usage behaviors. We aim for the ColonClean mHealth app to serve as a comprehensive and effective outpatient colonoscopy guidance tool, ultimately improving the effectiveness of BP.

Methods

Ethical Considerations

This study is an interventional investigation conducted in accordance with research ethics regulations. Approval from the Institutional Review Board of Taipei Veterans General Hospital was obtained prior to the commencement of the study (approval number 2021-10-007AC). The study was conducted from October 27, 2021, to December 31, 2022. Before enrolling participants, the purpose and details of the study were explained to them, and they provided informed consent by signing a consent form. Throughout the research process, all participants' privacy rights were strictly protected, and the collected data were used solely for research purposes. Participants were not compensated for their participation in this study.

Study Participants

G*Power 3.1.9.7 for Windows was used to analyze the questionnaire data, considering a power (1-β) of 0.8, α value of .05, a medium effect size of 0.3, and an estimated dropout rate of 20%. As a result, a sample size of 40 participants was determined. A total of 40 outpatients from the gastroenterology and endoscopy center of the medical center were recruited. The inclusion criteria were patients aged 20 years or older, recommended by physicians to undergo colonoscopy, without visual or hearing impairments, and without mental illnesses. After explaining the purpose and content of the study and obtaining written informed consent, paper-based nursing instructions were provided along with verbal explanations of ColonClean, and the BP was conducted. For easy access to ColonClean, the participants downloaded the ColonClean mHealth app on their mobile devices after scanning a QR code. The nursing instruction process involved 1-on-1 teaching, with an explanation lasting approximately 20 minutes in the independent waiting area outside the outpatient visit room. Each research participant was asked to submit feedback on their experience with ColonClean within 3 days of their colonoscopy.

mHealth App (ColonClean)

Individuals aged ≥50 years are the target group for CRC screening, making them the primary users of the ColonClean mHealth app. Considering the physiological and cognitive decline in older adults, it is crucial to design a user-friendly interface to avoid difficulties in usage. The interface design for older adults incorporates intuitive elements, clear and concise text, and straightforward operational steps [23]. Contrast colors were used for interface color schemes. The operational methods are simple and clear, with functions executed through clicks or straight-line swiping gestures [24,25]. All features were designed to operate within a single page, avoiding menu structures with more than 2 layers of information. This design reduces the psychological pressure on older adults, who may be afraid of making mistakes due to an unfamiliarity with technology; this allows them to focus and read the important information comfortably. Handwritten or voice inputs were preferred for filling in the forms. Clear return or cancel buttons were placed at the top or bottom of the screen. Considering the characteristics of the users, the emphasis was on interaction and feedback rather than simple page browsing to enhance the impression of BP.

The ColonClean mHealth app was developed using the Android 11 operating system due to its low development cost and popularity among the general public [26]. The user interface and content were designed using the Android app package, the Java programming language, and a built-in database.

The screen display of ColonClean is divided into 3 sections: top, middle, and bottom in a single-page layout for the main screen. The top section displays the "Examination Date and Time." The middle section features 6 key topics arranged in a zigzag pattern based on their importance. From left to right, these topics include "Dietary Guidelines," explaining the necessary dietary education starting 3 days before the examination as shown in Figure 1, and "Laxative Use," explaining the use of bowel cleansing medication. The "Fluid Intake" screen is for recording the amount of fluid consumed (Figure 2), while the "Stool Image Recognition" screen is for providing images to identify the type of the last bowel movement (Figure 3).

XSL•FO
RenderX

Figure 1. Representative images of “Dietary Options” in ColonClean.

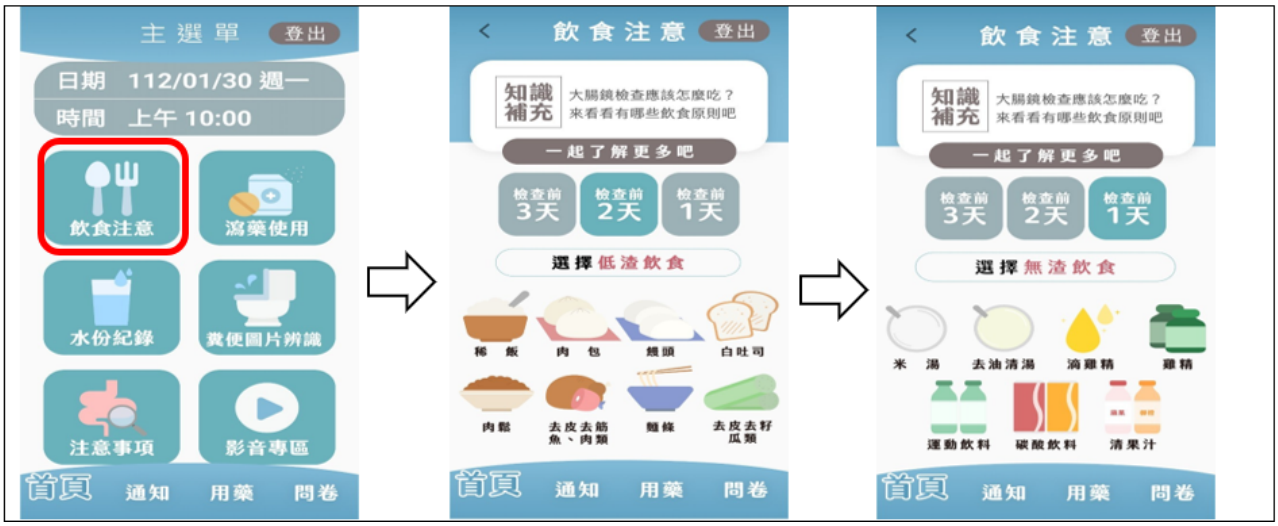


Figure 2. The “Hydration Intake Record” in ColonClean. Bowel preparation requires plenty of fluid intake to facilitate the passage of feces from the bowel.

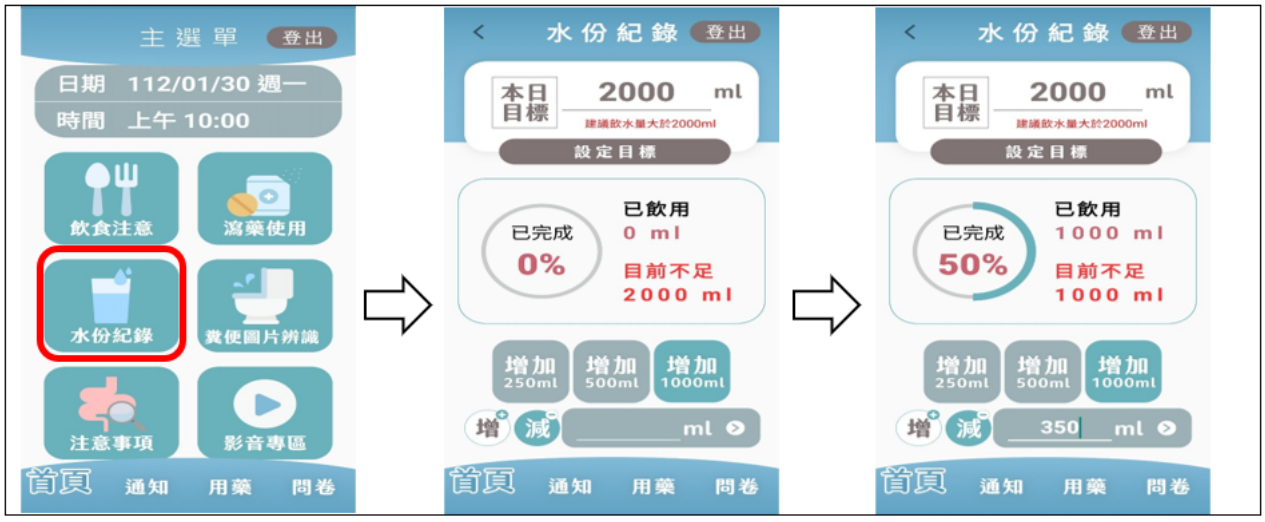
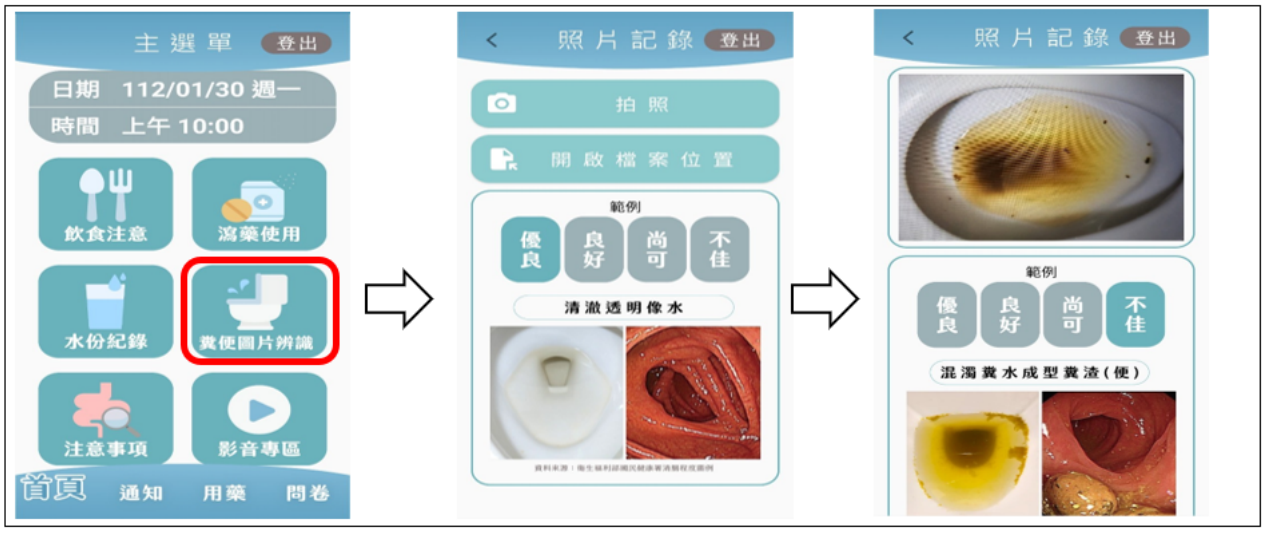


Figure 3. Images of the “Last Bowel Movement Type” section of ColonClean.



“Precautions” explains important points to keep in mind. The bottom section of the screen, from left to right, includes the “Home” button to return to the main screen; “Notifications,” which displays text messages; “Medication,” which provides

XSL•FO
RenderX

settings and modifications for regular medication or bowel cleansing agents; and "Questionnaire," which allows users to provide feedback on their experience using the ColonClean mHealth app.

The ColonClean app sent reminders and notifications to participants in the experimental group starting 3 days before their procedure. These notifications advised participants to begin dietary restrictions, such as avoiding high-fiber fruits and vegetables. Two days before the procedure, the app reminded users to switch to a low-residue diet, focusing on foods like congee or plain white noodles. The day before the procedure, the app prompted users to follow a clear liquid diet, including oil-free broth and sports drinks, and provided instructions on taking BP medications.

The app utilizes tap and swipe gestures for navigation; for user registration, it supports built-in Chinese phonetic input or handwriting recognition on mobile devices. The interface features a fresh color palette, primarily a warm lake green. Key information is highlighted in red, and unnecessary content is removed to offer concise summaries. Unique icons are designed as entry pointers to access relevant content. Tapping on an icon triggers entry into the main menu, which includes dietary guidance, instructions on using bowel cleansing agents (via images and videos), precautions for a colonoscopy, and fluid intake records, and provides references to the type of last bowel movement for understanding the differences compared to that during endoscopy. The design allows users to navigate intuitively, freely scrolling, tapping, and browsing related content between pages [27,28]. Push notifications are recommended to include instructions or prompts. Therefore, through push notifications, users are informed about relevant BP information to avoid any execution errors. Considering the user groups are middle-aged and older adults, complex actions are avoided. Instead, the key content is displayed directly in the push notification message. When a notification is received, users can click on the pushed message, which leads them to the relevant page with clear instructions or recommendations to view more detailed content. This approach reduces the chance of users misunderstanding the message. The user-centered push notification feature for BP steps enhances user engagement [29,30].

Aronchick Bowel Preparation Scale

After completing a colonoscopy, the physician submits a rating of the BP level using the Aronchick Bowel Preparation Scale (ABPS) in the endoscopy report system [1,31,32]. The ABPS is one of the most comprehensive and commonly used scales to describe the visualization of the colon and assess the percentage of stool covering the entire colon mucosa to explain the state of BP [33]. Based on this assessment, the adequacy of bowel cleansing is graded as excellent, good, fair, poor, or inadequate [11,33-35]. In Taiwan, domestic quality indicators for colonoscopy align with ABPS standards, targeting a 90% or above rating for both "excellent" and "good" categories. This ensures a high standard of care in colorectal screening, enhancing procedural accuracy and improving patient outcomes [36,37].

TAM

The TAM explains the behavior of users accepting new information technology in the field of computer technology [38]. Perceived usefulness and perceived ease of use affect users' attitudes toward adopting new technology, thereby influencing willingness and usage behavior. The perceived ease of use, perceived usefulness, attitudes, and behavior of outpatient endoscopy patients toward the ColonClean mHealth app were evaluated to assess their acceptance of using it. The questionnaire adopted a 5-point Likert scale. The 4 dimensions assessed by outpatient endoscopy patients after using ColonClean were perceived ease of use, perceived usefulness, attitudes, and usage behavior. The response options ranged from 1="strongly disagree," 2="disagree," 3="neutral," 4="agree," to 5="strongly agree." Higher scores indicated a higher level of agreement, while lower scores indicated a lower level of agreement. The questionnaire was initially developed based on literature references [39]. Three nursing informatics experts and scholars were then invited to evaluate the applicability and clarity of each question and to provide suggestions for improvement.

The Cronbach α coefficient was used to assess the internal consistency of the questionnaire items in this study. For the perceived usefulness dimension, consisting of 6 items, Cronbach α was 0.971. For the perceived ease of use dimension, consisting of 6 items, Cronbach α was 0.950. For the attitude dimension, consisting of 5 items, Cronbach α was 0.930. For the usage behavior dimension, consisting of 6 items, Cronbach α was 0.976. The questionnaire comprised a total of 23 items, and the Cronbach α coefficient for the scale was 0.986, indicating high internal consistency, meeting the requirement of a Cronbach α coefficient exceeding 0.7.

The structured questionnaire was designed and validated for reliability and validity by 3 nursing informatics experts. Participants accessed the questionnaire through a built-in link within the app, with the collected data stored in the backend management center for subsequent analysis.

Analysis

The data were analyzed using SPSS 25.0 (IBM Corp), with the utilization of the Pearson product-moment correlation and simple regression analysis to predict the perception of ColonClean.

Results

Demographic Characteristics

The demographic characteristics of the participants were as follows: 17 males (43%) and 23 females (58%); 27 participants aged ≥50 years (68%); and 35 participants with a high school/vocational education or above (88%).

ColonClean

After completing their colonoscopy, each research participant provided feedback on their experience with the ColonClean app within 3 days.. The feedback highlighted several key aspects of the app's design that contributed to user satisfaction and engagement:

XSL•FO
RenderX

- User-centered design: Participants appreciated the intuitive user interface, which maintained consistency in themes and main menu content. The clear and simple instructions, presented through a combination of text and visuals, facilitated user-friendly navigation. This consistency helped users feel more confident using the app.
- Examination date reminder: The feature displaying the examination date at the top of the main menu was well-received, as users felt it helped them stay organized and aware of the timing, potentially increasing compliance with preparation protocols.
- Graphic representations: Users found the unique graphic representations of dietary restrictions particularly helpful, as they enhanced understanding compared to traditional text descriptions. This visual approach was seen as effective in clarifying preexamination dietary options.
- Medication instructions: The clear, red-highlighted instructions for using bowel cleansing agents, along with drug images, dosage information, and timing, received positive feedback. Users appreciated the visual cues that simplified the medication process.
- Hydration tracking: Participants valued the hydration tracking feature, which encouraged them to monitor their water intake. Reminders to increase consumption when below recommended levels and positive feedback for meeting goals were highlighted as motivational elements that enhanced their commitment to the BP process.
- Bowel movement comparison: The option to take a photo of their last bowel movement and compare it with reference images was particularly beneficial. Users reported that this feature provided clarity on the required level of bowel cleanliness, enhancing their understanding of preparation requirements.
- Medication reminders and push notifications: The integration of medication reminders, particularly for chronic medications and anticoagulant discontinuation, was praised for preventing potential issues during the colonoscopy process. Users appreciated the timely push notifications that kept them informed about important tasks related to their preparation.

Overall, the research results indicated that users found the design of the ColonClean app effective in enhancing their understanding of and compliance with BP instructions. The app's focus on intuitive navigation, visual aids, and timely reminders contributed to a positive user experience, ultimately improving the effectiveness of BP for colonoscopy.

ABPS

After completing the colonoscopy, the BP level is reported by the endoscopist using the ABPS: 10 individuals (25%) achieved a "fair" level, 26 individuals (65%) achieved a "good" level, and 4 individuals (10%) achieved an "excellent" level. There were no cases of "poor" or insufficient preparation leading to interruption or inability to perform the examination.

TAM

Perceived Usefulness

The details of perceived usefulness are shown in Table 1.

Table . Perceived usefulness of the ColonClean app.

Perceived usefulness of the ColonClean app	Mean (SD)
Using ColonClean helps me quickly understand the tasks required for bowel preparation.	4.60 (0.59)
Using ColonClean provides easy access to relevant information about bowel preparation.	4.55 (0.68)
I believe that using ColonClean can improve the accuracy of my bowel preparation.	4.65 (0.53)
Using ColonClean enables me to effectively complete the required bowel preparation tasks.	4.65 (0.48)
ColonClean provides clear reminders about bowel preparation tasks, such as dietary restrictions and monitoring the type of my last bowel movement.	4.60 (0.59)
I find the educational information provided by ColonClean sufficient for completing my bowel preparation tasks.	4.65 (0.53)

The table presents users' perceptions of the usefulness of the ColonClean app in managing BP tasks. The mean scores, averaging around 4.62 of 5, suggest that users generally find the app very helpful. Below is an interpretation of the key findings:

- Understanding tasks (mean 4.60, SD 0.59): Users feel that the ColonClean app helps them quickly understand the requirements for BP, indicating clear and effective communication of instructions.
- Access to relevant information (mean 4.55, SD 0.68): The app is seen as a convenient tool for accessing important information about the preparation process. However, this metric has a slightly lower score, suggesting minor room for improvement in delivering information.
- Increased accuracy (mean 4.65, SD 0.53): Users believe that the app enhances the accuracy of their BP, reflecting a high level of trust in the app's guidance to meet medical standards.
- Effective completion of tasks (mean 4.65, SD 0.48): The app is viewed as highly effective in helping users complete BP tasks correctly, suggesting that it provides practical support throughout the process.

XSL•FO
RenderX

- Clear reminders (mean 4.60, SD 0.59): The app is viewed as highly effective in helping users complete BP tasks correctly, suggesting that it provides practical support throughout the process.
- Sufficient educational information (mean 4.65, SD 0.53): The app provides enough educational material for users to feel confident about completing their preparation, further supporting its educational role.

In summary, the ColonClean app was perceived as highly useful across various aspects of the BP process, with particularly strong scores for accuracy, task completion, and educational sufficiency. These results highlight the app's effectiveness in improving the user experience and promoting adherence to BP protocols.

Perceived Ease of Use

The details of perceived ease of use are shown in Table 2.

Table . Perceived ease of use of the ColonClean app.

Perceived ease of use of the ColonClean app	Mean (SD)
I think the interface design of ColonClean is clear and easy to understand.	4.63 (0.54)
I find the process of using ColonClean to be smooth.	4.63 (0.54)
It is easy to navigate and locate the desired functions in ColonClean.	4.63 (0.54)
I think the data presentation in ColonClean is quick and stable.	4.65 (0.58)
I find the font size in the ColonClean interface appropriate and easy to read.	4.53 (0.75)
The information provided by ColonClean makes it easier to understand the bowel preparation tasks.	4.68 (0.53)

Table 2 presents users' perceptions of the ease of use of the ColonClean app, highlighting their experiences with various interface features. The mean scores indicate a generally high level of satisfaction, with most ratings above 4.62 of 5, suggesting that users find the app intuitive and user-friendly. Below is an interpretation of the key findings:

- Clear interface (mean 4.63, SD 0.54): Users feel that the interface design is intuitive and user-friendly. This clarity in design is crucial for users' initial engagement with the app, suggesting that the developers prioritized usability.
- Smooth user experience (mean 4.63, SD 0.54): The app operates smoothly, without lag or issues, which is essential for encouraging continued use and satisfaction with digital health tools.
- Easy navigation (mean 4.63, SD 0.54): The ease of navigation suggests that users can quickly find the features they need, reducing frustration and enhancing overall satisfaction.
- Quick and stable data presentation (mean 4.65, SD 0.58): Users appreciate the app's rapid and reliable presentation of information. This reliability enhances user trust, which is particularly important in health care applications where accuracy is critical.
- Appropriate font size (mean 4.53, SD 0.75): Although the font size received a slightly lower rating, it still indicates that most users find it suitable. The slightly lower score suggests that some users might prefer larger text or different font styles, which could be an area for improvement.
- Enhanced understanding of BP (mean 4.68, SD 0.53): Users find that the information provided by ColonClean significantly aids their understanding of BP tasks, which is essential for ensuring compliance with medical procedures.

In summary, the high average scores across these metrics indicate that users generally find ColonClean to be an easy-to-use app that effectively supports BP. The highlighted areas demonstrate strengths in design, functionality, and educational support, with only minor adjustments needed in font size to further enhance the user experience. These findings suggest that ColonClean is a valuable tool for improving patient preparation and potentially contributing to better clinical outcomes.

Attitudes Toward Usage

The details of attitudes toward usage are shown in Table 3.

Table . Attitudes toward usage of the ColonClean app.

Attitudes toward usage of the ColonClean app	Mean (SD)
I believe ColonClean is an excellent choice for assisting with bowel preparation before an examination.	4.63 (0.54)
I am currently satisfied with the benefits provided by using ColonClean.	4.65 (0.53)
I would be willing to use ColonClean again in the future if I need to undergo bowel preparation.	4.63 (0.59)
I find the information provided by ColonClean to be very useful.	4.60 (0.59)
I prefer using ColonClean over paper-based nursing instructions (single sheets).	4.35 (0.77)

XSL•FO
RenderX

Table 3 summarizes users' attitudes toward the use of the ColonClean app, reflecting generally positive perceptions of its utility and effectiveness in assisting with BP tasks. The mean scores indicated a high level of satisfaction, with most ratings above 4.57 of 5, suggesting a strong endorsement of the app's features and benefits. Below is an interpretation of the key findings:

- Good choice for assistance (mean 4.63, SD 0.54): This high mean score reflects a strong consensus among users that ColonClean is an effective tool for supporting BP. The low SD suggests consistent agreement on its usefulness in this context.
- Satisfaction with benefits (mean 4.65, SD 0.53): Users' satisfaction indicates that the app meets their expectations and provides meaningful benefits during BP. The relatively low SD further shows that most users feel positively about the app's contributions.
- Willingness to use again (mean 4.63, SD 0.59): Users expressed a strong intention to use ColonClean for future BPs, suggesting that their initial positive experiences built trust in the app and that they see it as reliable.
- Usefulness of information (mean 4.60, SD 0.59): The perception that ColonClean provides highly useful information reinforces its role as a valuable educational resource, effectively aiding users in understanding the preparation process.
- Preference over paper instructions (mean 4.35, SD 0.77): Although users generally prefer ColonClean over traditional paper-based instructions, the slightly lower score and higher SD suggest that some users may still value printed materials or have mixed feelings about fully transitioning to digital formats. This variability presents an opportunity for further enhancements in the app's usability and content delivery.

In summary, the high mean scores (above 4.57) suggested that users had a positive attitude toward ColonClean, seeing it as an effective and beneficial tool for BP. The consistent responses (low SDs) indicated broad agreement among users on its usefulness. Meanwhile, the variation in preference over paper instructions may highlight areas for potential improvement. This feedback can guide future updates and enhancements to maximize user satisfaction and efficacy.

Usage Behavior

The details of usage behavior are shown in Table 4.

Table . Usage behavior of the ColonClean app.

Usage behavior of the ColonClean app	Mean (SD)
I am willing to use ColonClean as an aid for bowel preparation.	4.73 (0.45)
I am highly satisfied with the ability to access knowledge through ColonClean during bowel preparation.	4.70 (0.52)
I would recommend ColonClean to friends and family in the future.	4.63 (0.63)
I would use ColonClean to improve the effectiveness of bowel preparation.	4.70 (0.52)
I believe ColonClean provides useful information during the bowel preparation period.	4.68 (0.53)
Overall, I find using ColonClean to be satisfactory.	4.73 (0.51)

Table 4 summarizes the usage behavior of the ColonClean app and reveals overwhelmingly positive user sentiment. The overall mean score of 4.69 indicates high satisfaction with the experience of using ColonClean. This average score, combined with a relatively low SD (0.50), suggests a strong consensus among users regarding the app's effectiveness and utility. Below is an analysis of the meaning behind the scores:

- Willingness to use (mean 4.73, SD 0.45): The high mean score indicates that users are very inclined to use ColonClean as a supportive tool for BP. The low SD signifies that this sentiment is widely shared, suggesting strong confidence in the app's benefits.
- Satisfaction with knowledge access (mean 4.70, SD 0.52): Users expressed a high level of satisfaction with the knowledge accessible through the app during BP. This indicates that ColonClean effectively meets their informational needs, enhancing their preparation experience.
- Recommendation to others (mean 4.63, SD 0.63): The willingness to recommend ColonClean to friends and family reflects users' trust in the app's efficacy. Although this score remained positive, the slightly higher SD suggests some variability in users' willingness to recommend it, possibly due to differing experiences or expectations.
- Improvement of effectiveness (mean 4.70, SD 0.52): Users believed that ColonClean enhanced the effectiveness of their BP, reinforcing the app's perceived value as a helpful resource in the preparation process.
- Useful information obtained (mean 4.68, SD 0.53): The belief that the app provides useful information further underscores its utility as an educational tool. Users recognized that the app offers relevant guidance throughout the preparation period.
- Overall satisfaction (mean 4.73, SD 0.51): The high overall satisfaction score indicates that users felt positively about their experience with ColonClean. The consistency of this sentiment, reflected by the low SD, suggests that many users share a similar level of satisfaction.

In summary, the high mean scores across all categories reflect users' positive attitudes and satisfaction, underscoring ColonClean as a valuable tool for assisting with BP. The low SDs across most items indicate consistency in user experiences, suggesting that ColonClean effectively meets users' needs and

XSL•FO
RenderX

expectations. This feedback is critical for developers, as it reinforces the app's strengths and highlights areas where further enhancements could be beneficial.

Analysis

Pearson Correlation Coefficients

Pearson correlation coefficients were used to analyze the relationships among users' attitudes, perceived usefulness, perceived ease of use, and usage behavior of the ColonClean mHealth app (Table 5). The results indicated a significant positive correlation between users' attitudes and perceived usefulness (r_{38}=.907, P<.001), users' attitudes and perceived ease of use (r_{38}=.825, P<.001), users' attitudes and usage behavior (r_{38}=.835, P<.001), perceived usefulness and perceived ease of use (r_{38}=.894, P<.001), perceived usefulness and usage behavior (r_{38}=.933, P<.001), and perceived ease of use and usage behavior (r_{38}=.958, P<.001). Positive correlations were demonstrated between perceived usefulness and perceived ease of use, usage attitude, and usage behavior.

Table . Correlations among perceived usefulness, perceived ease of use, usage attitude, and usage behavior.

		Perceived usefulness	Perceived ease of use	Usage attitude	Usage behavior
Perceived usefulness					
	Pearson correlation analysis	1			
	Significance level (2-tailed)				
	Number	40			
Perceived ease of use					
	Pearson correlation analysis	0.894	1		
	Significance level (2-tailed)	<.001			
	Number	40	40		
Usage attitude					
	Pearson correlation analysis	0.907	0.825	1	
	Significance level (2-tailed)	<.001	<.001		
	Number	40	40	40	
Usage behavior					
	Pearson correlation analysis	0.933	0.958	0.835	1
	Significance level (2-tailed)	<.001	<.001	<.001	
	Number	40	40	40	40

Simple Regression Analysis

We used simple regression analysis to predict the correlation among the perceived usefulness, perceived ease of use, usage attitude, and usage behavior in the context of ColonClean. The correlation between perceived usefulness and usage attitude, with usage attitude as the criterion variable and perceived usefulness as the predictor variable, yielded B=0.907, t_{38}=13.3. This model explained 82.3% of the variance in usage attitude (R^2=0.823), and the test result with P<.001 indicated that perceived usefulness was a significant predictor variable for usage attitude. This suggests that when users perceive the app as providing substantial assistance, their attitude toward using it improves significantly. This highlights the importance of enhancing the app's practical functionality to foster a positive user experience. Similarly, the correlation between perceived usefulness and usage behavior, with usage behavior as the criterion variable and perceived usefulness as the predictor variable, resulted in B=0.933, t_{38}=16.0. This model explained 87.1% of the usage behavior (R^2=0.871), and the test result with P<.001 demonstrated that perceived usefulness significantly predicted usage behavior. When users believe that the app helps improve their preparation process, they are more likely to engage actively with it. The correlation between perceived ease of use and usage attitude, with usage attitude as the criterion variable and perceived ease of use as the predictor variable, resulted in B=0.825, t_{38}=9.0. This explained 68% of the variance in attitude toward use (R^2=0.68), with P<.001, indicating perceived ease of use as a significant predictor of attitude toward use. When users find the app easy to use, their attitude becomes more positive. This reflects their comfort and satisfaction during use, which in turn influences their willingness to engage with the

XSL•FO
RenderX

app. When determining the correlation between perceived ease of use and usage behavior, with usage behavior as the criterion variable and perceived ease of use as the predictor variable, the result was B=.958, t_{38}=20.5. This model explained 91.7% of the variance in usage behavior (R^2=.917), with P<.001, indicating that perceived ease of use significantly predicted usage behavior. When users perceive the app as easy to operate, they are more likely to engage actively. This underscores the importance of simplifying the app's interface and functionality to enhance user engagement and satisfaction.

In summary, the results indicate that both perceived usefulness and perceived ease of use have a significant positive impact on user attitudes and behaviors. This suggests that enhancing the perceived usefulness and ease of use of the ColonClean app during its design may encourage more positive user attitudes and behaviors, thereby improving health management outcomes. These findings not only validate the app's usability but also provide empirical support for future improvements.

Discussion

Principal Findings

Mobile health apps are increasingly used for illness prevention, education, and promoting healthy lifestyles [40]. However, the requirements for BP are often complex and challenging for patients to navigate and remember. To help outpatient colonoscopy patients adhere to dietary guidelines, manage bowel cleansing agents, and maintain regular medication schedules, push notifications were incorporated into the app's design. The app, protected under the patent titled "ColonClean Mobile Management System" (patent number M639094, Republic of China), features a user-centric interface that is both intuitive and visually appealing. The well-designed interface, content, push notifications, and multimedia section of the ColonClean mHealth app provide users with a convenient and helpful experience when using their mobile devices for health management. Users expressed their willingness to use the app again in the future and recommend it to family and friends, highlighting its effectiveness as an educational guide for BP.

User Attitudes and Preference for Paper-Based Instructions

Despite the overall positive reception of the ColonClean app, some users indicated a preference for paper-based instructions, as reflected in the lower mean score (mean 4.350.77) on this aspect. This finding aligns with existing literature suggesting that certain users may find physical formats more reliable or easier to reference than digital alternatives [41]. Factors contributing to this preference may include concerns about technology accessibility, device compatibility, or personal comfort with digital apps for health management.

Additionally, the preference for paper-based instructions may arise from a desire for tangible resources that can be easily annotated or highlighted [42]. Some users may feel that printed materials provide a more straightforward and immediate reference during the preparation process, particularly in situations where mobile devices are not accessible or convenient [43]. Therefore, while the digital format of the ColonClean app offers numerous advantages, such as interactivity and real-time updates, it is essential to recognize the value that traditional paper-based resources continue to hold for certain individuals.

Limitations and Recommendations

The ColonClean app operates on the Android system and functions as an mHealth app for mobile devices. However, it is important to note that during the acceptance process, instances may arise where individuals with non-Android operating systems or outdated device models are unable to be included in this study. Those unable to download ColonClean by scanning the QR code were excluded. It is advisable to consider optimizing future mHealth apps so that they can be integrated into mobile devices of various operating systems. This enhancement would contribute to more inclusive nursing guidance for BP, ensuring accessibility across platforms.

In the current era of smart health care, it is recommended to incorporate artificial intelligence recognition capabilities into the ColonClean app. Specifically, for the section where users capture images of their last bowel movement, artificial intelligence can automatically analyze and determine the quality of the stool, categorizing it as "excellent," "good," "fair," or "poor." This improvement would facilitate a more accurate assessment of the final bowel movement status. Furthermore, optimizing the ColonClean mHealth app for other individuals who require BP, such as those undergoing colorectal surgery or hospitalized patients preparing for colonoscopy, would ensure the successful completion of bowel cleansing.

Additionally, in light of the COVID-19 pandemic, when individuals undergoing screening needed to reschedule their appointments due to a positive diagnosis or being a close contact, or when they encountered issues like poor bowel movements, a consultation chat window was recommended as a future addition to address user inquiries and provide real-time solutions, thus maintaining user engagement and satisfaction. To ensure the continuity of examinations and achieve the objectives of examination, diagnosis, and treatment, it is crucial to minimize ineffective examinations or schedule cancellations that may impact individuals in need of these examinations. This approach aims to maintain a steady number of examinations and fulfill the purpose of providing necessary medical assessments and subsequent care.

Conclusions

In summary, while the ColonClean app shows promise in supporting BP, it is essential to recognize user attitudes toward traditional paper-based instructions. Enhancing the app with broader accessibility, artificial intelligence capabilities, and additional user support features could significantly improve its effectiveness and reach, ultimately contributing to better health outcomes for patients requiring BP.

XSL•FO
RenderX

Acknowledgments

We would like to express our heartfelt gratitude to the patients from the outpatient department of the gastroenterology and endoscopy center for their valuable participation in this study.

Data Availability

Data from this study are available from the corresponding author upon reasonable request.

Authors' Contributions

HYC conceptualized the study, participated in data curation and formal analysis, completed the methodology, validated the data, and drafted the manuscript. MHT supervised the study and reviewed the manuscript. MYC reviewed the manuscript. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

References

1. Chiu YT, Kuo CY, Lee FJ, Chang CY. Dedicated staff for patient education improves bowel preparation quality and reduces the cecal intubation time of colonoscopy: a single institution retrospective study. Medicine (Balt) 2022 Jul 22;101(29):e29437. [doi: 10.1097/MD.0000000000029437] [Medline: 35866774]
2. Mazidimoradi A, Tiznobaik A, Salehiniya H. Impact of the COVID-19 pandemic on colorectal cancer screening: a systematic review. J Gastrointest Cancer 2022 Sep;53(3):730-744. [doi: 10.1007/s12029-021-00679-x] [Medline: 34406626]
3. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021 May;71(3):209-249. [doi: 10.3322/caac.21660] [Medline: 33538338]
4. Hung SY, Chen HC, Chen WTL. A randomized trial comparing the bowel cleansing efficacy of sodium picosulfate/magnesium citrate and polyethylene glycol/bisacodyl (The Bowklean Study). Sci Rep 2020 Mar 27;10(1):5604. [doi: 10.1038/s41598-020-62120-w] [Medline: 32221332]
5. Patel N, Kashyap S, Mori A. Bowel Preparation: StatPearls; 2021. URL: https://www.ncbi.nlm.nih.gov/books/NBK535368 [accessed 2024-12-31]
6. Bechtold ML, Mir F, Puli SR, Nguyen DL. Optimizing bowel preparation for colonoscopy: a guide to enhance quality of visualization. Ann Gastroenterol 2016;29(2):137-146. [doi: 10.20524/aog.2016.0005] [Medline: 27065725]
7. Wen CC, Jao SW, Hsiao CW. A modified bowel preparation regimen for colonoscopy providing the patients' satisfaction and convenience. Med Sci Monit 2017 Jun 26;23:3123-3129. [doi: 10.12659/msm.905431] [Medline: 28649124]
8. Walter J, Francis G, Matro R, et al. The impact of diet liberalization on bowel preparation for colonoscopy. Endosc Int Open 2017 Apr;5(4):E253-E260. [doi: 10.1055/s-0043-101694] [Medline: 28382323]
9. Zad M, Do CN, Heffernan A, Johnston L, Al-Ansari M. Factors affecting bowel preparation adequacy and procedural time. JGH Open 2020 Apr;4(2):206-214. [doi: 10.1002/jgh3.12241] [Medline: 32280766]
10. Ye Z, Chen J, Xuan Z, Gao M, Yang H. Educational video improves bowel preparation in patients undergoing colonoscopy: a systematic review and meta-analysis. Ann Palliat Med 2020 May;9(3):671-680. [doi: 10.21037/apm.2020.03.33] [Medline: 32312062]
11. Sharma P, Burke CA, Johnson DA, Cash BD. The importance of colonoscopy bowel preparation for the detection of colorectal lesions and colorectal cancer prevention. Endosc Int Open 2020 May;8(5):E673-E683. [doi: 10.1055/a-1127-3144] [Medline: 32355887]
12. Sulz MC, Kröger A, Prakash M, Manser CN, Heinrich H, Misselwitz B. Meta-analysis of the effect of bowel preparation on adenoma detection: early adenomas affected stronger than advanced adenomas. PLoS One 2016;11(6):e0154149. [doi: 10.1371/journal.pone.0154149] [Medline: 27257916]
13. Chang JY, Moon CM, Lee HJ, et al. Predictive factors for missed adenoma on repeat colonoscopy in patients with suboptimal bowel preparation on initial colonoscopy: a KASID multicenter study. PLoS ONE 2018;13(4):e0195709. [doi: 10.1371/journal.pone.0195709] [Medline: 29698398]
14. Hassan C, East J, Radaelli F, et al. Bowel preparation for colonoscopy: European Society of Gastrointestinal Endoscopy (ESGE) Guideline – Update 2019. Endoscopy 2019 Aug;51(8):775-794. [doi: 10.1055/a-0959-0505] [Medline: 31295746]
15. Rex DK, Boland CR, Dominitz JA, et al. Colorectal cancer screening: recommendations for physicians and patients from the U.S. Multi-Society Task Force on Colorectal Cancer. Gastroenterology 2017 Jul;153(1):307-323. [doi: 10.1053/j.gastro.2017.05.013] [Medline: 28600072]
16. Urban G, Tripathi P, Alkayali T, et al. Deep learning localizes and identifies polyps in real time with 96% accuracy in screening colonoscopy. Gastroenterology 2018 Oct;155(4):1069-1078. [doi: 10.1053/j.gastro.2018.06.037] [Medline: 29928897]

XSL•FO
RenderX

17. Di Leo M, Iannone A, Arena M, et al. Novel frontiers of agents for bowel cleansing for colonoscopy. World J Gastroenterol 2021 Dec 7;27(45):7748-7770. [doi: 10.3748/wjg.v27.i45.7748] [Medline: 34963739]
18. Greenhalgh T, Koh GCH, Car J. Covid-19: a remote assessment in primary care. BMJ 2020 Mar 25;368:m1182. [doi: 10.1136/bmj.m1182] [Medline: 32213507]
19. Thulesius H. Increased importance of digital medicine and eHealth during the Covid-19 pandemic. Scand J Prim Health Care 2020 Jun;38(2):105-106. [doi: 10.1080/02813432.2020.1770466] [Medline: 32484725]
20. Liu Z, Zhang MM, Li YY, Li LX, Li YQ. Enhanced education for bowel preparation before colonoscopy: a state-of-the-art review. J Dig Dis 2017 Feb;18(2):84-91. [doi: 10.1111/1751-2980.12446] [Medline: 28067463]
21. Hernández G, Gimeno-García AZ, Quintero E. Strategies to improve inadequate bowel preparation for colonoscopy. Front Med (Lausanne) 2019;6:245. [doi: 10.3389/fmed.2019.00245] [Medline: 31781565]
22. Walter B, Frank R, Ludwig L, et al. Smartphone application to reinforce education increases high-quality preparation for colorectal cancer screening colonoscopies in a randomized trial. Clin Gastroenterol Hepatol 2021 Feb;19(2):331-338. [doi: 10.1016/j.cgh.2020.03.051] [Medline: 32240835]
23. Salman HM, Wan Ahmad WF, Sulaiman S. Usability evaluation of the smartphone user interface in supporting elderly users from experts' perspective. IEEE Access 2018;6:22578-22591. [doi: 10.1109/ACCESS.2018.2827358]
24. Darvishy A, Hutter HP. Recommendations for age-appropriate mobile application design. Stud Health Technol Inform 2017;242:676-686. [Medline: 28873871]
25. Johnson J, Finn K. Designing User Interfaces for an Aging Population: Towards Universal Design: Morgan Kaufmann; 2017. [doi: 10.1016/C2015-0-01451-4]
26. Samhi J, Allix K, Bissyandé TF, Klein J. A first look at Android applications in Google Play related to COVID-19. Empir Softw Eng 2021;26(4):57. [doi: 10.1007/s10664-021-09943-x] [Medline: 33903798]
27. Bidargaddi N, Pituch T, Maaieh H, Short C, Strecher V. Predicting which type of push notification content motivates users to engage in a self-monitoring app. Prev Med Rep 2018 Sep;11:267-273. [doi: 10.1016/j.pmedr.2018.07.004] [Medline: 30109172]
28. Freyne J, Yin J, Brindal E, Hendrie GA, Berkovsky S, Noakes M. Push notifications in diet apps: influencing engagement times and tasks. Int J Hum-Comput Interact 2017 Oct 3;33(10):833-845. [doi: 10.1080/10447318.2017.1289725]
29. Back SY, Kim HG, Ahn EM, et al. Impact of patient audiovisual re-education via a smartphone on the quality of bowel preparation before colonoscopy: a single-blinded randomized study. Gastrointest Endosc 2018 Mar;87(3):789-799. [doi: 10.1016/j.gie.2017.09.007] [Medline: 28939500]
30. van der Zander QEW, Reumkens A, van de Valk B, Winkens B, Masclee AAM, de Ridder RJJ. Effects of a personalized smartphone app on bowel preparation quality: randomized controlled trial. JMIR mHealth uHealth 2021 Aug 19;9(8):e26703. [doi: 10.2196/26703] [Medline: 34420924]
31. Sharma RS, Rossos PG. A review on the quality of colonoscopy reporting. Can J Gastroenterol Hepatol 2016;2016:9423142. [doi: 10.1155/2016/9423142] [Medline: 27446877]
32. Nam SJ, Kim YJ, Keum B, et al. Impact of diet restriction on bowel preparation for colonoscopy. Medicine (Balt) 2018 Oct;97(41):e12645. [doi: 10.1097/MD.0000000000012645] [Medline: 30313052]
33. Kastenberg D, Bertiger G, Brogadir S. Bowel preparation quality scales for colonoscopy. World J Gastroenterol 2018 Jul 14;24(26):2833-2843. [doi: 10.3748/wjg.v24.i26.2833] [Medline: 30018478]
34. Manem N, Donovan K, Miller D, et al. Open-access colonoscopy quality indicators and patient perception using split-dose bowel preparation. JGH Open 2021 May;5(5):563-567. [doi: 10.1002/jgh3.12532] [Medline: 34013055]
35. Donovan K, Manem N, Miller D, et al. The impact of patient education level on split-dose colonoscopy bowel preparation for CRC prevention. J Cancer Educ 2022 Aug;37(4):1083-1088. [doi: 10.1007/s13187-020-01923-x] [Medline: 33405208]
36. Quality Indicators for Screening Colonoscopy [Website in Chinese]. Health Promotion Administration, Ministry of Health and Welfare. 2021. URL: https://www.hpa.gov.tw/File/Attach/1093/File_19179.pdf [accessed 2024-12-31]
37. May FP, Shaukat A. State of the science on quality indicators for colonoscopy and how to achieve them. Am J Gastroenterol 2020 Aug;115(8):1183-1190. [doi: 10.14309/ajg.0000000000000622] [Medline: 32759620]
38. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
39. Lai YH, Huang FF, Yang SH. Development and related research of the Chinese version of the Technology Acceptance Model scale. J Orient Inst Techn 2015;35(12):201-221 [FREE Full text]
40. Lee YL, Cui YY, Tu MH, Chen YC, Chang P. Mobile health to maintain continuity of patient-centered care for chronic kidney disease: content analysis of apps. JMIR mHealth uHealth 2018 Apr 20;6(4):e10173. [doi: 10.2196/10173] [Medline: 29678805]
41. Ferraz LT, Santos AJT, Lorenzi LJ, Frohlich DM, Barley E, Castro PC. Design considerations for the migration from paper to screen-based media in current health education for older adults: a scoping review. BMJ Open 2024 Apr 10;14(4):e078647. [doi: 10.1136/bmjopen-2023-078647] [Medline: 38604627]
42. Schorr EN, Gepner AD, Dolansky MA, et al. Harnessing mobile health technology for secondary cardiovascular disease prevention in older adults: a scientific statement from the American Heart Association. Circ Cardiovasc Qual Outcomes 2021 May;14(5):e000103. [doi: 10.1161/HCQ.0000000000000103] [Medline: 33793309]

XSL•FO
RenderX

43. Pischke CR, Voelcker-Rehage C, Ratz T, et al. Web-based versus print-based physical activity intervention for community-dwelling older adults: crossover randomized trial. JMIR mHealth uHealth 2022 Mar 23;10(3):e32212. [doi: 10.2196/32212] [Medline: 35319484]

Abbreviations

ABPS: Aronchick Bowel Preparation Scale
BP: bowel preparation
CRC: colorectal cancer
mHealth: mobile health
TAM: Technology Acceptance Model

Edited by Y Gong; submitted 18.03.24; peer-reviewed by L Weinert, YL Lee; revised version received 08.11.24; accepted 26.11.24; published 08.01.25.

Please cite as:
Chen HY, Tu MH, Chen MY
Using a Mobile Health App (ColonClean) to Enhance the Effectiveness of Bowel Preparation: Development and Usability Study
JMIR Hum Factors 2025;12:e58479
URL: https://humanfactors.jmir.org/2025/1/e58479
doi:10.2196/58479

© Hui-Yu Chen, Ming-Hsiang Tu, Miao-Yen Chen. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 8.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

NutriDiary, a Smartphone-Based Dietary Record App: Description and Usability Evaluation

Linda Klasen[1], MSc; Stefanie Anna Julia Koch[1], PhD; Maike Elena Benz[1], PhD; Johanna Conrad[2], PhD; Ute Alexy[1], PhD; Konrad Blaszkiewicz[2,3], MSc; Ionut Andone[2,3], MSc; Ute Nöthlings[1], DrPH

[1]Nutritional Epidemiology, Institute of Nutrition and Food Sciences, University of Bonn, Friedrich-Hirzebruch-Allee 7, Bonn, Germany

[2]German Nutrition Society (DGE), Bonn, Germany

[3]Department of Computer Science, University of Bonn, Bonn, Germany

Corresponding Author:
Ute Nöthlings, DrPH
Nutritional Epidemiology, Institute of Nutrition and Food Sciences, University of Bonn, Friedrich-Hirzebruch-Allee 7, Bonn, Germany

Abstract

Background: Repeated applications of short-term dietary assessment instruments are recommended for estimating usual dietary intake. For this purpose, NutriDiary, a smartphone app for collecting weighed dietary records (WDRs) in the German population, was developed.

Objective: We aim to describe NutriDiary and evaluate its usability and acceptability.

Methods: NutriDiary was developed as a WDR, allowing users to enter food items via text search, barcode scanning, or free text entry. The sample for the evaluation study included 74 participants (n=51, 69% female, aged 18 - 64 years), including 27 (37.5%) experts and 47 (63.5%) laypersons (including n=22, 30%, nutrition students). Participants completed a 1-day WDR and entered a predefined sample meal (n=17 foods) the following day by using NutriDiary. An evaluation questionnaire was answered from which the system usability scale (SUS) score (0 - 100) was calculated. A backward selection procedure (PROC REG in SAS; SAS Institute) was used to identify potential predictors for the SUS score (age, sex, status [expert or laypersons], and operating system [iOS or Android]).

Results: The median SUS score of 75 (IQR 63 - 88) indicated good usability. Age was the only characteristic identified as a potential predictor for a lower SUS score (P<.001). The median completion time for an individual WDR was 35 (IQR 19 - 52) minutes. Older participants took longer to enter the data than younger ones (18 - 30 y: median 1.5, IQR 1.1 - 2.0 min/item vs 45 - 64 y: median 1.8, IQR 1.3 - 2.3 min/item). Most participants expressed a preference for NutriDiary over the traditional paper-based method.

Conclusions: Good usability and acceptability make NutriDiary promising for use in epidemiological studies.

(JMIR Hum Factors 2025;12:e62776) doi:10.2196/62776

KEYWORDS

dietary assessment; food record; barcode scanning; app; mobile phone

Introduction

Usual dietary intake, the long-term average daily intake of a nutrient or food**,** is the relevant exposure when studying diet-health relationships in nutritional epidemiology. To estimate usual dietary intake, repeated applications of short-term dietary assessment instruments are recommended [1]. Self-reported dietary intake is the most commonly used method in large-scale studies through its rapid and cost-effective use. Mostly, instruments such as 24-hour dietary recalls, food frequency questionnaires or dietary records (DRs) are used, with each of these methods having individual limitations and strengths [2]. Repeated use of these traditional dietary assessment instruments is costly and burdensome for both participants and researchers [3,4]. Innovative technologies such as web- or smartphone-based tools have the potential to facilitate self-reported dietary assessment by reducing the costs and time effort of data collection and postprocessing while achieving higher acceptance in study participants [5-7]. Smartphone technology seems particularly promising among those innovative approaches. First, smartphones are widely available to a large proportion of the population. According to a survey conducted in 2021, over 95% of people aged older than 13 years in Germany already use a smartphone [8]. Second, due to the advantage of portability and the fact that most people always carry their smartphone with them, smartphone-based tools are well suited for real-time recording of food intake [1]. Third, smartphone apps can facilitate food entry through the supplementary use of an

XSL•FO
RenderX

integrated barcode scanner or by using the camera function, which reduces the need to look up every food item in a database or enter it manually [9]. Nowadays, a large number of commercial nutrition apps for self-tracking or nutritional advice are freely available [10,11]. However, because of their limited scope and questionable quality of nutrient information, those apps tend to be unsuitable for dietary assessment in epidemiological studies [12,13]. Thus, several DR apps have been developed specifically for use in epidemiological studies. In a systematic review by König et al [14], 5 core assessment features to collect data on dietary intake in scientific studies by smartphone apps were identified: photo-based assessment, assessment of serving or portion sizes, free-text description of food intake, selection from a food database, and classification systems. These features are used either alone or in combination. Thereby, the combination of photo-based recording with free-text descriptions of the consumed foods or the joint use of a food database and the assessment of serving or portion sizes were the most popular methods [14].

To provide a digital alternative for paper-based dietary assessment in epidemiological studies, we developed NutriDiary, an app for conducting weighed dietary records (WDRs) with an integrated barcode scanner. This paper aims to describe the current version of NutriDiary and to report on its usability and acceptability in laypersons and experts.

Methods

NutriDiary and the NutriDiary Database

Development and Functions of NutriDiary

NutriDiary was developed as a smartphone app to conduct WDRs within nutritional epidemiological studies. The app is available in common app stores and study participants (hereafter referred to as users) can use the app on their smartphones with personal login data (Figure 1A). Users can start their WDR after agreeing to the data protection regulations (Figure 1B). When entering a new eating occasion, users are initially asked to enter the date, time, and place of consumption (Figure 1C). Then, the app offers three ways to enter food, beverages, and supplements: (1) a text search and subsequent selection from the underlying database, (2) barcode scanning in the underlying database, or (3) free text entry.

Figure 1. The NutriDiary app (android version): (A) screen for entering access data, (B) welcome screen, and (C) food entry.

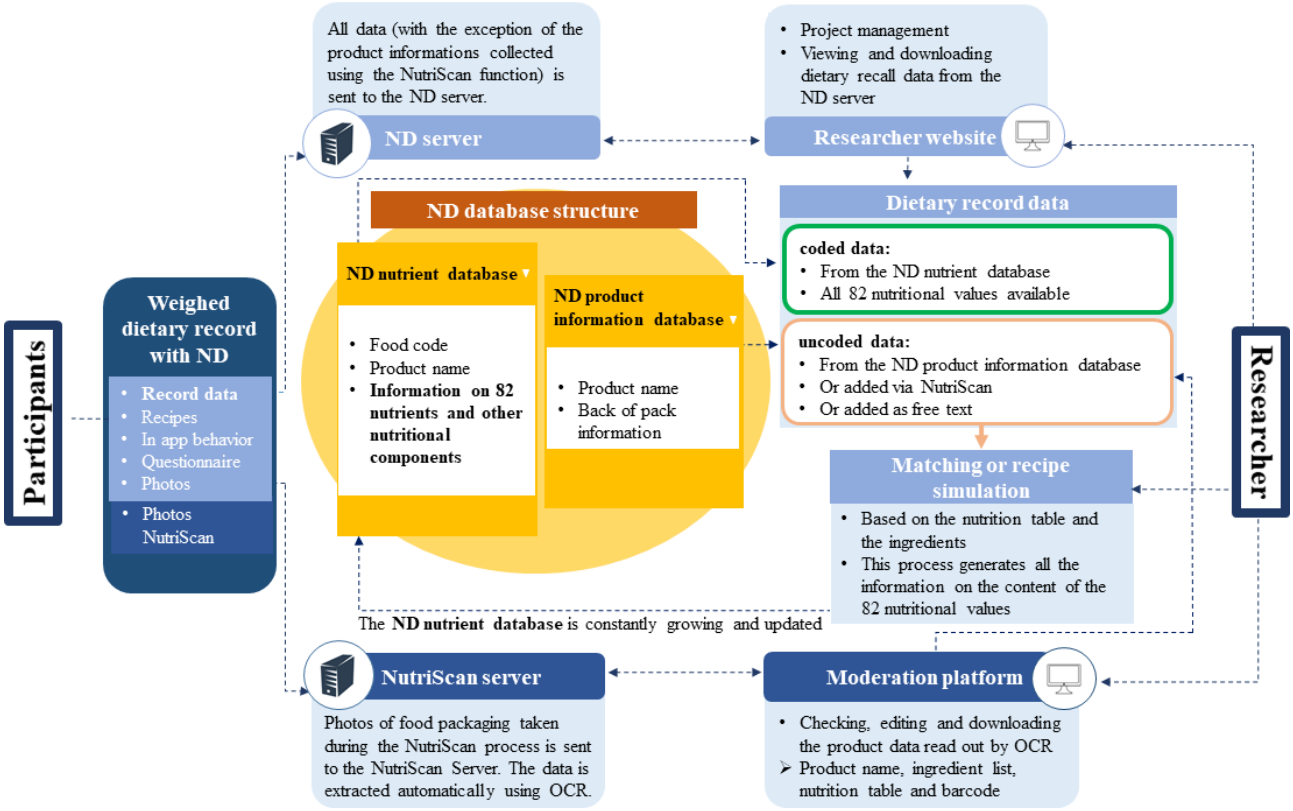


If a food item cannot be identified via barcode scanning, the user is guided through a standardized process for collecting all relevant product information (hereinafter referred to as the NutriScan process). Thereby, the user is asked to take photos of the brand and product name, the barcode, the ingredient list and the nutrient table following step-by-step instructions (Multimedia Appendix 1). This data is then sent to the NutriScan server and automatically read out using optical character reading. Researchers can access, edit, and download this data via a moderation platform. Based on the packaging information, dieticians can match detailed nutrient data from a similar product in the database or estimate detailed nutrient values by recipe simulation in order to continuously update and expand the underlying database (Figure 2). The recipe simulation is carried out manually using the list of ingredients and the nutritional table and is described in detail elsewhere [15]. Finally, users enter the weighed amount consumed, preparation method, and quantity of potential leftovers (Figure 1C). If weighing is not possible, users are provided with a range of specified options in the drop-down menu to select an estimated portion size, such

XSL•FO
RenderX

as teaspoon or slice. After all foods and beverages have been entered, users are redirected to the main screen, where all entered eating occasions are displayed for final review. To further ease the process of recording, NutriDiary offers some usability features. A recipe editor allows entry of custom recipes, which will be added to the user's personal databases. An integrated help mode provides immediate and problem-specific assistance to users by simply touching the screen element they need help with. In addition, the app includes a photo function for collecting information, for example, on meals consumed out-of-home and entered via free text. After finishing the WDR, data are submitted to a server of the University of Bonn (NutriDiary server, Figure 2). This server also provides an administration tool to researchers (researcher website) where scientific personnel can select project-specific settings, for example, study name, study duration or the number of recording days. The app also offers the option to integrate an individual questionnaire at the end of the recording period.

Figure 2. Overview of the data structure and data flow of the NutriDiary app; ND: NutriDiary; OCR: optical character reading.



Description of the NutriDiary Database

The complex database structure of NutriDiary is shown in Figure 2. It currently contains more than 150,000 items (approximately 25% are duplicates due to different packaging sizes and different barcodes) which come from various sources. The core database is divided into the NutriDiary nutrient database and the NutriDiary product information database. The basis of the NutriDiary nutrient database is an adaptation of the in-house food and nutrient database of the DONALD (Dortmund Nutritional and Anthropometric Longitudinally Designed) study named LEBTAB (LEBensmittelTABelle, food table) [15-17]. LEBTAB currently contains around 19,000 generic and branded food items (as of June 2023) with corresponding contents of energy, 82 nutrients, and other nutritional components. Data entries for generic foods are based predominantly on the German national standard food database "Bundeslebensmittelschlüssel" (version 3.02 [18]; Federal Ministry of Food and Agriculture). The energy and nutrient contents of branded foods are predominantly estimated by recipe simulation using labeled ingredients and declared nutrient contents [15,17]. When the participant enters a food item from the NutriDiary nutrient database, this item is automatically coded and the corresponding values for energy and 82 nutrients are available. To enhance user-friendliness and allow barcode scanning, additional branded food products and barcode information were added to the NutriDiary database structure by:

- requesting product information (product name, barcode, ingredient list, and nutrition table) from food manufacturers and others (Atrify [19], a subsidiary of GS1 Germany [1WorldSync] and DATA NatuRe eG [20], a cooperative to create a central data pool for organically produced foods) to set up the NutriDiary product information database, and
- adding barcode information of already included branded products to the NutriDiary nutrient database by searching in open databases such as Open Food Facts [21], codecheck [22] (Producto Check GmbH), and Open EAN/GTIN Database (European Article Number [EAN]/Global Trade Item Number [GTIN]) [23].

For the branded products obtained through step 1 (NutriDiary product information database), information only on the nutritional values indicated on the packaging (big 7) and the ingredient list are available (Figure 2). When a participant reports one of these products in a WDR, the trained dieticians

XSL•FO
RenderX

match extended nutrient values of an equivalent food from the NutriDiary nutrient database or carry out a recipe simulation as already mentioned above [15]. This step generates all 82 nutrients for these products and they are then transferred to the NutriDiary nutrient database, which is constantly growing.

The NutriDiary Evaluation Study

Study Design

A layperson and expert evaluation was conducted. Participants were asked to keep an individual WDR with NutriDiary for 1 day and to enter a predefined sample meal on the following day. The sample meal was identical for all participants and included 4 meals (breakfast, snack, lunch, and dinner) and was provided as a digital presentation sent to study participants via email. The sample meal comprised both, generic (n=15, presented as text) and branded food items (n=3, presented as pictures of the packaging including barcode), all labeled with a hypothetical quantity of consumption, preparation method, and type of entry (text entry vs barcode scanning). Participants were generally instructed to enter branded products via scanning the barcode. Further, 1 product was intentionally not available in both NutriDiary databases, requiring participants to complete the NutriScan process in order to add the food item to the record. Furthermore, participants were tasked to correct a logged entry and to enter a hypothetical leftover without further instructions but with reference to the NutriDiary website [24] to test whether the aids provided there (frequently asked questions and help videos) fulfill their purpose (ie, helping users to help themselves). After participants completed their WDR and entered the sample meal, they were asked to answer an app-integrated evaluation questionnaire on usability and acceptability of NutriDiary on day 3. As the WDRs were not analyzed at the nutrient level, no scales were handed out for weighing the food on day 1. The participants were asked to use scales from their households.

All participants were provided with a short video giving key instructions on how to use NutriDiary before starting the WDR. The video gave a brief overview of how to use the help mode, enter foods (via text search, barcode scanning, or free text entry) and navigate the recipe book. Furthermore, participants were informed about the NutriDiary website [24], where frequently asked questions and help videos for both, iOS and Android can be found. Beyond this, participants did not undergo any additional training in using the app.

Recruitment of This Study's Population

We aimed to recruit a minimum of 51 participants, based on the assumptions outlined by Lewis and Sauro [25]. This recommendation takes into account an SD of 17.7, which is typical for system usability scale (SUS) scores [25], and ensures sufficient precision to achieve a 95% CI with a margin of error of ±5 points. This study recruited both laypersons and experts. Experts (trained nutritionists with experience in the field of dietary assessment) were recruited via direct invitations (n=28). Laypersons (n=52) were recruited via oral advertisement in lectures, mailing, and personal contact by students studying nutrition and food science at the University of Bonn, and as part of a student project. The latter mainly targeted participants between 30 and 60 years of age (in the personal environment of the students and in 2 gyms) in order to increase the number of participants in middle age and older in the group of laypersons. All participants had to be fluent in the German language, have a functioning smartphone and a valid email address. Written informed consent from all participants was obtained before enrollment.

Usability and Acceptability Assessment

The questionnaire on usability and acceptability included 14 questions on 3 different categories (usability, acceptance, and technical issues). The usability of NutriDiary was assessed by using the SUS by Brooke [26,27], which allows for comparison between similar systems or products. In short, the SUS is a Likert scale consisting of 5 positive statements (odd-numbered) and 5 negative statements (even-numbered) to which respondents indicate their degree of agreement on a scale from 1 (strong disagreement) to 5 (strong agreement). For odd items, 1 is subtracted from the user response, and for even-numbered items, 5 is subtracted from the user response, added up and multiplied by 2.5 to convert the score ranging from 0 to 100, whereas higher scores indicate better usability. According to Bangor et al [28], an SUS score below 50 is considered as "not acceptable," a score between 50 - 70 years as "marginal," and a score above 70 as "acceptable" [28,29]. In our study, an appropriate German translation of the SUS developed by SAP usability professionals of a German software corporation (SAP SE; Systems, Applications & Products in Data Processing Societas Europaea) was used and integrated into the usability questionnaire [30].

Age and sex of the participants were assessed within the evaluation questionnaire. Status (expert or layperson) was already categorized during the recruitment process. Information on the operating system was automatically recorded when NutriDiary was used and sent to the server together with the questionnaire data. To find out whether participants would prefer the app to a traditional paper-based WDR, 2 additional questions in the same structure of the SUS were added. In order to assess technical problems participants were asked whether technical errors occurred (yes or no) and, if yes, to describe the error in more detail (free text entry).

Furthermore, in-app behavior of users was recorded and evaluated. The completion time for a WDR with NutriDiary was determined by using the activity protocol of the app, in which all actions were recorded and time-stamped. For this, the time of all input activities was summed up. Interruptions of more than 5 minutes were counted as breaks and excluded from summation. As the time effort of a WDR depends on the complexity of the meals and the number of foods eaten, the relative completion time (completion time divided by the number of items) was additionally calculated. Furthermore, the percentage of estimated household measurements was of interest as well as the percentage of automatically coded WDR entries.

Statistical Analyses

Results and participants' characteristics are presented as median with their lower IQRs for continuous variables or as relative frequencies (%) for categorical variables. A backward selection

XSL•FO
RenderX

procedure (PROC REG in SAS) was used to identify characteristics of participants that were potential predictors for the SUS score. The following variables were tested: sex (male or female), age (years), status (expert or layperson), and operating system (iOS or Android). All statistical analyses were conducted using SAS (version 9.4). The significance level was set at $P<.05$.

Ethical Considerations

All examinations were carried out with written informed consent from study participants. The NutriDiary evaluation study was approved by the Ethics Committee of the University of Bonn (project identification 445/23). All data collected in this study has been pseudonymized to protect participant privacy. No reimbursement was provided to participants for their involvement in this study.

Results

From overall 80 study participants, 74 (27 experts) completed the NutriDiary evaluation study according to this study protocol and answered the evaluation questionnaire (Table 1). Most participants were female (51/74, 69%) and younger than 30 years of age (41/74, 55%).

The overall age ranged from 18 to 64 years and the median age was 29 years (IQR 25 - 45). Overall, 54% (40/74) of the participants owned a smartphone with an iOS operating system. Table 2 shows the SUS score for NutriDiary for the total study sample and stratified by sex, age groups, status, and operating system. In the overall study sample, the SUS score for NutriDiary ranged from 43 to 100 (data not shown). The median SUS score of 75 (IQR 63 - 88) indicates a good usability of NutriDiary. The mean SUS score was 74 (SD 15; 95% CI 70-77; data not shown). The median SUS score was higher in women than in men (80, IQR 65 - 88, vs 70, IQR 55 - 78) and higher in the group of experts than in the group of laypersons (80, IQR 65 - 88 vs 73, IQR 58 - 88), with female laypersons having a higher SUS score than female experts (83, IQR 66 - 88 vs 80, IQR 60 - 85). Looking at the group of laypersons after excluding nutrition students (labeled in Tables 1 and 2 as “others,” median age 55, IQR 29 - 58, years), the median SUS score was 63 (IQR 50 - 73).

Table . Characteristics of study participants of the NutriDiary evaluation study (N=74).

	Total	Experts	Laypersons		
			Total	Nutrition students	Others
Participants, n (%)	74 (100)	27 (36.5)	47 (63.5)	22 (47)	25 (53)
Sex (group), n (%)					
Women	51 (69)	23 (85)	28 (60)	19 (86)	9 (36)
Age, median (IQR)	29 (25 - 45)	35 (28 - 40)	27 (23 - 56)	23 (22 - 26)	55 (29 - 58)
Age (years), n (%)					
Age group 1: 18-≤30	41 (55.4)	12 (44.4)	29 (61.7)	22 (100)	7 (28)
Age group 2:>30-≤45	15 (20.3)	11 (40.7)	4 (8.5)	0 (0)	4 (16)
Age group 3:>45 - 65	18 (24.3)	4 (14.8)	14 (29.8)	0 (0)	14 (56)
Operating system, n (%)					
iOS	40 (54)	12 (44.4)	28 (59.6)	12 (54.5)	16 (64)
Android	34 (46)	15 (55.6)	19 (40.4)	10 (45.5)	9 (36)

XSL•FO
RenderX

Table . System usability scale (SUS) score for the NutriDiary app presented as median (IQR) values.

	Total		Women		Men	
	n	Median (IQR)	n	Median (IQR)	n	Median (IQR)
SUS score	74	75 (63 - 88)	51	80 (65 - 88)	23	70 (55 - 78)
Stratified by age group						
Age group 1: 18 -≤30 years	41	80 (70 - 88)	31	83 (70 - 88)	10	71 (63 - 85)
Age group 2: >30 -≤45 years	15	80 (68 - 90)	11	85 (65 - 93)	4	74 (70 - 78)
Age group 3: >45 - 65 years	18	58 (50 - 70)	9	58 (48 - 65)	9	65 (50 - 70)
Stratified by status						
Experts	27	80 (65 - 88)	23	80 (60 - 85)	4	83 (70 - 90)
Laypersons (total)	47	73 (58 - 88)	28	83 (66 - 88)	19	70 (55 - 73)
Nutrition students	22	85 (75 - 88)	19	88 (78 - 90)	3	73 (70 - 85)
Others	25	63 (50 - 73)	9	55 (48 - 65)	16	66 (53 - 73)
Stratified by operating system						
iOS	40	78 (65 - 88)	26	80 (68 - 90)	14	66 (55 - 73)
Android	34	74 (58 - 85)	25	75 (58 - 85)	9	73 (70 - 85)

Among the age groups 1 (18-≤30 y) and 2 (>30-≤45 y), the median SUS score was noticeably higher (80, IQR 70 - 88 and 80, IQR 68 - 90) than in age group 3 (58, IQR 50 - 70). The results of the backward selection procedure showed that age was the only characteristic identified as a potential predictor for the SUS score in the examined sample (P<.001).

Figure 3 shows the individual statements of the SUS questionnaire and a summary of the answers given by the 74 participants, presented as box plots. The figure shows that agreement tends to be high for the positive (odd-numbered) statements and tends to be low for the negative (even-numbered) statements, as is a prerequisite for a higher SUS score. Among the positive statements, the participants showed the lowest level of agreement with statement 1 “I think that I would like to use NutriDiary frequently.” For the negative statements, agreement was highest for statement 8 “I found NutriDiary very cumbersome to use.”

Figure 3. User rating (N=74) of the individual statements of the SUS evaluating the usability and acceptability of NutriDiary, shown as box plots. The dot represents the mean and the thick stripe the median. SUS: system usability scale.

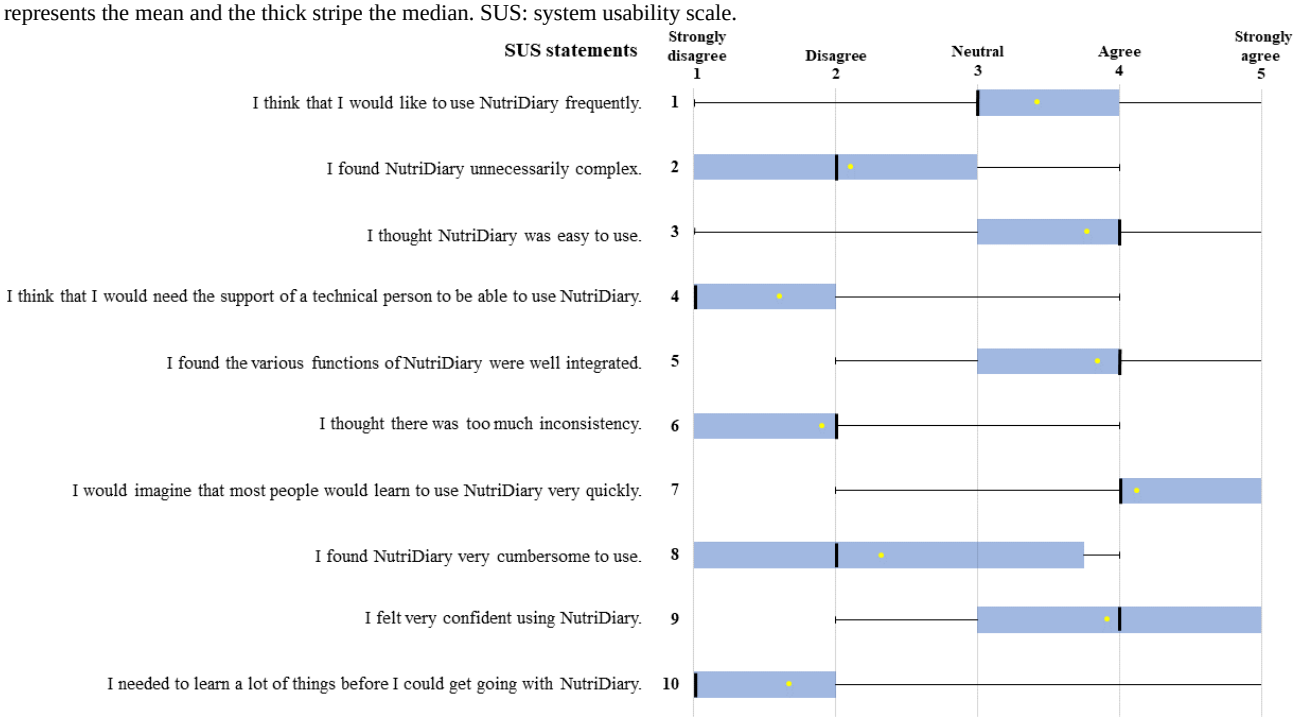


Table 3 shows the results of the evaluated in-app behavior logs and WDRs. The calculated median completion time for an individual WDR (day 1) with NutriDiary was 35 (IQR 19 - 52) minutes. Participants needed a median of 1.6 (IQR 1.2 - 2)

XSL•FO
RenderX

minutes to enter an item. This relative completion time was slightly higher in age group 3 (>45 - 63 y) compared to younger participants (1.8, IQR 1.3 - 2.3, min/item vs 1.5, IQR 1.1 - 2, min/item). The median time it took participants to enter the sample meal in NutriDiary on day 2 was 15 (IQR 12 - 18) minutes. The median relative completion time for entering the sample meal was 0.8 (IQR 0.7 - 0.9) min/item. Participants in age group 3 took slightly longer than younger participants. The median proportion of already coded items was 68% (IQR 55% - 82%) and the median proportion of estimated quantities (not weighed) was 8% (IQR 0% - 29%). Younger participants (age group 1) used more estimated household measurements instead of weighing than participants in age groups 2 and 3 (0% - 26%, 8% vs 0% - 31%, 7% and 0% - 23%, and 4%, respectively).

Table . Completion time and proportion of items already coded for the total sample (N=74) and stratified by age groups.

	Total	Age group 1: 18 - 30 y	Age group 2: 31 - 45 y	Age group 3: 46 - 65 y
Participants, n (%)	74 (100)	41 (55.4)	15 (20.3)	18 (24.3)
Number of entered items day 1[a]	21 (15 - 31)	22 (16 - 31)	21 (19 - 32)	21 (13 - 30)
Total completion time day 1 (min)[a]	35 (19 - 52)	35 (19 - 45)	44 (19 - 55)	36 (17 - 53)
Relative completion time day 1 (min/item)[a]	1.6 (1.2 - 2)	1.5 (1.1 - 2)	1.5 (1.1 - 2)	1.8 (1.3 - 2.3)
Total completion time sample meal (min)[a]	15 (12 - 18)	14 (11 - 17)	17 (11 - 19)	17 (13 - 23)
Relative completion time sample meal (min/item)[a]	0.8 (0.7 - 0.9)	0.8 (0.6 - 0.9)	0.9 (0.6 - 1.1)	0.9 (0.7 - 1.3)
Proportion of items already coded (%)[a]	68 (55 - 82)	63 (56 - 79)	77 (62 - 86)	71 (50 - 86)
Proportion of estimated household quantities (%)[a]	8 (0 - 29)	8 (0 - 26)	7 (0 - 31)	4 (0 - 23)

[a]Median (IQR).

In total, 24% (18/74) of the participants reported technical issues. In 2 cases, the app crashed, but could be reopened afterward and no data was lost. In 5 cases, participants stated that they had problems finding newly added products (this process is sometimes delayed in rare cases). Further, 4 participants reported that the selection from the drop-down menu did not work properly. In the remaining cases, participants described issues rather associated with this study design than with the technical functions of NutriDiary or gave general comments. For example, 4 participants reported that a barcode presented in the sample meal could not be scanned from the screen (probably due to varying screen-brightness and screen-resolution).

Overall, 85% (63/74) of the participants were able to enter the sample meal correctly. Of the 11 participants who made mistakes when entering the sample meal, 5 were in age group 1, 2 in age group 2, and 4 in age group 3, indicating that age did not affect the accuracy of data entry. In 2 cases, food items were missing. Further, 3 participants had difficulties in editing a WDR entry and entering hypothetical leftovers. Furthermore, 6 participants entered a different amount of food than that presented in the sample meal. When participants were asked whether they would prefer NutriDiary to the traditional paper method, a total of 77% (57/74) agreed. Only 9.5% (7/74) could imagine, that it would be easier to keep a WDR with pen-and-paper than to use NutriDiary (Multimedia Appendix 2).

Discussion

Principal Findings

To our knowledge, an app-based WDR system with barcode scanning function for scientific studies in Germany does not exist so far. The popularity and widespread use of smartphones make NutriDiary a promising alternative to the traditional pen-and-paper approach. Different ways for entering food items (text search, barcode scanning, and free text entry) and the NutriScan function enable users to record the products they consume in a very detailed manner. The digital data output has the potential to reduce the burden for researchers also. In this study, we evaluated NutriDiary in a convenience sample of experts and laypersons and found good usability and acceptability.

The median SUS score of 75 indicated good usability of NutriDiary. The SUS score has been developed as a means to measure the overall perceived usability of a system [26,27]. However, technology-based dietary assessment instruments for scientific purposes are very specific tools that are not designed for everyday use but to generate scientifically useful data. The collection of these data is often challenging for the participants. The first statement of the SUS questionnaire is “I think that I would like to use this system frequently.” This statement is unlikely to be agreed upon by many people in the case of a WDR due to the burden on participants by the method itself.

XSL•FO
RenderX

Nevertheless, the advantage of the SUS score is that it offers the possibility to compare systems used in the same context [27]. The median SUS score of 75 for NutriDiary (mean SUS score: 74, SD 15, for comparison) is comparable with SUS scores for other similar technology-based food records such as "METADIETA-web" (mean SUS score: 68, SD 15, n=26) [31], the "Eat and Track app" (mean SUS score: 69, n=15) [32] or the "Traqq app" (mean SUS score: 79, SD 15, n=22) [33].

Age has been discussed as a major limiting factor for the use of technology-based systems [29] or innovative dietary assessment instruments [31,34,35]. Consistently, the results of the backward selection procedure showed that age was identified as a potential predictor for a lower SUS score in the examined sample. Further, older participants took slightly longer to complete a WDR with NutriDiary than younger ones. Feasibility testing of the digital food record METADIETA-web by Vitale et al [31] also showed that the preference for using the digital tool instead of the traditional pen-and-paper method decreased with increasing age [31]. Older generations did not grow up with smartphones and computers and are therefore likely to be less intuitive with apps and web-based tools than younger generations. However, longer completion times could also be explained by higher accuracy in entering items and greater patience [36], which would be in line with the observation that older participants used fewer estimated household measurements instead of weighing than younger ones in our study (0% - 23%, 4% for age group 3 vs 0% - 31%, 7% and 0% - 26%, 8% for age group 2 and 1, respectively). Nevertheless, an understandable introduction to the use of the instrument might be particularly important in the group of older adults. For practical reasons, only an introductory video was sent out in this study. Our experience in training users to use the NutriDiary app shows that older people tend to ask more questions. This can only be addressed in a face-to-face conversation where it can be ensured that all relevant information is conveyed. Therefore, we recommend a personal introduction when using NutriDiary in epidemiological studies, if feasible.

The median SUS score was higher in women than in men and higher in the group of experts than in the group of laypersons. This may partly be explained by the fact that the proportion of older people was higher in men than in women and higher in the group of laypersons than in the group of experts. The group of laypersons also included students of nutritional science. Although this group certainly cannot be described as experts, the students may have some background knowledge that could have influenced the outcome. Whether this makes them more critical or less critical in their judgment remains questionable.

WDRs have the potential to provide the most accurate description of the types and amounts of the foods consumed over a specified period of time, but they are also considered to be one of the most burdensome and elaborate dietary assessment methods. When conducting a WDR, participants have to weigh and write down everything they eat and drink and always carry their kitchen scales, the record sheet, and a pen with them. This process is exhausting and time-consuming and requires a high level of cooperation from the participants [37]. The median completion time for conducting a WDR with NutriDiary was 35 minutes (1.6 min/item, Table 3) and can be rated as acceptable. When looking at the median relative completion time of the sample meal (here, the amount of the presented foods was predefined), it can be assumed that the single entering process (without weighing) takes around 0.8 minutes per item. This result suggests that weighing accounts for about half the completion time. Keeping this in mind, the single entering process is roughly comparable with the reported average completion time of other text-based DR apps, for example, with "My Meal Mate." Here, participants needed on average 22 minutes per day to complete a record with estimated (not weighed) portion sizes [38].

Whether the NutriDiary app significantly reduces the completion time compared to the traditional pen-and-paper method is questionable, because the digital version of a WDR does not change the fact that weighing is required for this dietary assessment method. However, NutriDiary offers some advantages that can make its use more attractive than the traditional method. First, most people always carry their smartphone with them and recording "on the go" is more practical than using pen-and-paper. Presumably, this makes it less likely that participants forget their records and need to add the meal at a later time. Second, features such as the recipe book and the barcode scanner can make recording easier, as frequently consumed food combinations and recipes can be stored and retrieved later, and consumed products can be added more easily. Furthermore, integrated standard household measurements make it easier to estimate the quantity if weighing is not possible, for example, when eating out of home. The photo function allows participants to add a photo of the meal, which can help the postprocessing. To figure out whether NutriDiary is more attractive to study participants than the traditional pen-and-paper method, we added 2 more questions to the SUS questionnaire (, Multimedia Appendix 2). The result clearly showed that the vast majority would prefer using the NutriDiary app instead of a paper-based food record. To avoid errors caused by application problems using technology-based instruments, an understandable introduction and information structure are necessary to prevent frustration and enable participants to help themselves quickly if they have difficulties. For this purpose, we designed a website where all information and support materials are centralized in 1 location [24].

When using traditional DRs in epidemiological studies, the postprocessing of the DRs is very time-consuming. The data needs to be digitized and manually coded by the study staff. In this study, 68% (IQR 55% - 82%) of the food entered in NutriDiary was already coded. Furthermore, the data was already available in digital form, suggesting less costs and time in data postprocessing, compared to the pen-and-paper method. NutriDiary was designed as an app for keeping WDRs, meaning that participants are asked to weigh all the food and drinks they consume. However, as experience showed that this is not always possible, the app also offers the choice of standard household measurements (eg, teaspoon, glass, or portion) to estimate the quantity in situations where weighing is not possible. Considering the burden of weighing all foods, this selection option might tempt participants to estimate rather than weigh. However, the proportion of estimated quantities was rather low and can be considered manageable in this study.

XSL•FO
RenderX

Within the development process of NutriDiary, building an appropriate food database structure and the collection, standardization, and integration of barcode and food packaging information to enable food entry via barcode scanning was one of the most challenging tasks. The underlying database determines the users' success in searching for food items and is essential for the functionality and accuracy of technology based, self-administered dietary assessment tools [3,39]. Therefore, we aimed to make the database as complete as possible. Due to a high number of available branded foods and the frequently changing food market, this is a major challenge [12,40-43]. According to the Food Federation Germany [44], there are more than 170,000 food products available on the German food market. Every year, about 40,000 new products are launched and just as many disappear [44]. It also happens that manufacturers change food recipes, which can also change the ingredient lists and nutritional values. These circumstances mean that the database needs to be constantly updated. To make this possible, the NutriScan function described above was developed and integrated. If the app is used regularly, the database is regularly updated with new products recorded by users.

Strengths and Limitations

The NutriDiary evaluation study had some strengths and limitations. As NutriDiary was developed for scientific use, practicing nutritionists were also recruited as experts. Although experts are not intended users of the app, their evaluation allows for a professional view of the usability of NutriDiary. Furthermore, experts decide on the use of the assessment instrument in scientific studies, which is why their opinion on usability is of particular importance in this context. To assess user-friendliness in older age groups as well, we specifically recruited older participants. However, this study population was not representative of the general German population. Most participants were female, young, and highly educated. This limited the generalizability of our results. Nevertheless, it is important to acknowledge that participants in epidemiological studies frequently possess a higher educational status compared to the general population. In addition, participants were aware of the purpose of this study, which was to evaluate the usability and acceptability of NutriDiary, making it difficult to assess the quality and transferability of the data on the completion time and proportion of items already coded. This study aimed to assess the usability and acceptability of NutriDiary and does not contain any data on the validity of the NutriDiary app. However, providing information on the app's validity and psychometric properties is a critical prerequisite for its effective application. Therefore, a validation study has already been initiated and will give insights into the validity and quality of the assessed nutritional data.

Conclusion

NutriDiary is the first smartphone based WDR app with integrated barcode scanning function for scientific purposes in Germany. The evaluation by experts and laypersons indicated an acceptable completion time, good usability and acceptability on the users' side, whereby younger experts and laypersons tended to rate the app better than older ones. Future research will give insights into the validity and feasibility of NutriDiary in different study populations.

Acknowledgments

The authors would like to thank all study participants for the provision of their data. Special thanks go to the students Sofie Frey, BSc, and Franziska Oidtmann, BSc, for their support with data collection. This work was supported by Diet-Body-Brain (DietBB) and the Competence Cluster in Nutrition Research funded by the Federal Ministry of Education and Research (FKZ: 01EA1410A).

Authors' Contributions

LK and UN did the conceptualization. IA, KB, JC, SAJK, LK, MEB, and UN developed the app. LK and MEB handled data curation and worked on the formal analyses. LK wrote the original draft. SAJK, MEB, UA, JC, IA, and UN reviewed and edited the writing. UN supervised the study.

Conflicts of Interest

The authors were involved in the development of the NutriDiary app.

Multimedia Appendix 1
The NutriScan process.
[PNG File, 73 KB - humanfactors_v12i1e62776_app1.png]

Multimedia Appendix 2
User rating (n=74) on participants' preferences for NutriDiary compared to a traditional WDR (pen -and -paper method).
[JPEG File, 90 KB - humanfactors_v12i1e62776_app2.jpeg]

References

1. Conrad J, Nöthlings U. Innovative approaches to estimate individual usual dietary intake in large-scale epidemiological studies. Proc Nutr Soc 2017 Aug;76(3):213-219. [doi: 10.1017/S0029665116003025] [Medline: 28162110]

XSL•FO
RenderX

2. Cade JE. Measuring diet in the 21st century: use of new technologies. Proc Nutr Soc 2017 Aug;76(3):276-282. [doi: 10.1017/S0029665116002883] [Medline: 27976605]
3. Koch SAJ, Conrad J, Hierath L, et al. Adaptation and evaluation of Myfood24-Germany: a web-based self-administered 24-h dietary recall for the German adult population. Nutrients 2020 Jan 6;12(1):160. [doi: 10.3390/nu12010160] [Medline: 31935885]
4. Park Y, Dodd KW, Kipnis V, et al. Comparison of self-reported dietary intakes from the automated self-administered 24-h recall, 4-d food records, and food-frequency questionnaires against recovery biomarkers. Am J Clin Nutr 2018 Jan 1;107(1):80-93. [doi: 10.1093/ajcn/nqx002] [Medline: 29381789]
5. Sharp DB, Allman-Farinelli M. Feasibility and validity of mobile phones to assess dietary intake. Nutrition 2014;30(11-12):1257-1266. [doi: 10.1016/j.nut.2014.02.020] [Medline: 24976425]
6. Illner AK, Freisling H, Boeing H, Huybrechts I, Crispim SP, Slimani N. Review and evaluation of innovative technologies for measuring diet in nutritional epidemiology. Int J Epidemiol 2012 Aug;41(4):1187-1203. [doi: 10.1093/ije/dys105] [Medline: 22933652]
7. Thompson FE, Subar AF, Loria CM, Reedy JL, Baranowski T. Need for technological innovation in dietary assessment. J Am Diet Assoc 2010 Jan;110(1):48-51. [doi: 10.1016/j.jada.2009.10.008] [Medline: 20102826]
8. Tenzer F. Smartphone-Besitz bei Kinder und Jugendliche in Deutschland im Jahr 2021 nach Altersgruppe. Statista. 2022. URL: https://de.statista.com/statistik/daten/studie/1106/umfrage/handybesitz-bei-jugendlichen-nach-altersgruppen/ [accessed 2025-01-22]
9. Mortazavi BJ, Gutierrez-Osuna R. A review of digital innovations for diet monitoring and precision nutrition. J Diabetes Sci Technol 2023 Jan;17(1):217-223. [doi: 10.1177/19322968211041356] [Medline: 34467803]
10. Holzmann SL, Pröll K, Hauner H, Holzapfel C. Nutrition apps: quality and limitations. An explorative investigation on the basis of selected apps. Ernahrungs Umschau 2017;64(5):80-89. [doi: 10.4455/eu.2017.018]
11. Ferrara G, Kim J, Lin S, Hua J, Seto E. A focused review of smartphone diet-tracking apps: usability, functionality, coherence with behavior change theory, and comparative validity of nutrient intake and energy estimates. JMIR mHealth uHealth 2019 May 17;7(5):e9232. [doi: 10.2196/mhealth.9232] [Medline: 31102369]
12. Maringer M, Wisse-Voorwinden N, Geelen A. Food identification by barcode scanning in the Netherlands: a quality assessment of labelled food product databases underlying popular nutrition applications. Public Health Nutr 2019 May;22(7):1215-1222. [doi: 10.1017/S136898001800157X] [Medline: 29962361]
13. Fallaize R, Zenun Franco R, Pasang J, Hwang F, Lovegrove JA. Popular nutrition-related mobile apps: an agreement assessment against a UK reference method. JMIR mHealth uHealth 2019 Feb 20;7(2):e9838. [doi: 10.2196/mhealth.9838] [Medline: 30785409]
14. König LM, Van Emmenis M, Nurmi J, Kassavou A, Sutton S. Characteristics of smartphone-based dietary assessment tools: a systematic review. Health Psychol Rev 2022 Dec;16(4):526-550. [doi: 10.1080/17437199.2021.2016066] [Medline: 34875978]
15. Sichert-Hellert W, Kersting M, Chahda C, Schäfer R, Kroke A. German food composition database for dietary evaluations in children and adolescents. J Food Compost Anal 2007 Feb;20(1):63-70. [doi: 10.1016/j.jfca.2006.05.004]
16. Kroke A, Manz F, Kersting M, et al. The DONALD Study. History, current status and future perspectives. Eur J Nutr 2004 Feb;43(1):45-54. [doi: 10.1007/s00394-004-0445-7] [Medline: 14991269]
17. Perrar I, Alexy U, Nöthlings U. Cohort profile update-overview of over 35 years of research in the Dortmund Nutritional and Anthropometric Longitudinally Designed (DONALD) study. Eur J Nutr 2024 Apr;63(3):727-740. [doi: 10.1007/s00394-023-03290-x] [Medline: 38151532]
18. Bundeslebensmittelschnlüssel (BLS), version 3.02. Federal Ministry of Food and Agriculture. URL: https://www.blsdb.de/ [accessed 2025-01-22]
19. GDSN plattform für GS1 member organisationen und communities. 1WorldSync. 2023. URL: https://www.atrify.com/partner/gs1-member-organizations/?gclid=CjwKCAjwqZSlBhBwEiwAfoZUIJ8s4-oKxmTJ2Q86jycpue1mf7O2__wWg6zVBsrmh0KQ_kBbV0gloRoCPXQQAvD_BwE [accessed 2025-01-22]
20. DATA NatuRe eG. 2023. URL: https://datanature.de/ [accessed 2025-01-22]
21. Open food facts. 2023. URL: https://de.openfoodfacts.org/discover [accessed 2025-01-22]
22. Codecheck. 2023. URL: https://www.codecheck.info/ [accessed 2025-01-22]
23. Open EAN/GTIN Database. 2023. URL: https://opengtindb.org/ [accessed 2025-01-22]
24. Universität Bonn. NutriDiary. 2024. URL: https://www.nutridiary-app.de [accessed 2025-01-22]
25. Lewis J, Sauro J. Sample sizes for a SUS score. MeasuringU. 2022. URL: https://measuringu.com/sample-sizes-for-sus-ci [accessed 2025-01-22]
26. Brooke J. SUS-a quick and dirty usability scale. URL: https://hell.meiert.org/core/pdf/sus.pdf [accessed 2025-01-22]
27. Brooke J. SUS: a retrospective. J Usability Stud 2013;8(2):29-40 [FREE Full text] [doi: 10.5555/2817912.2817913]
28. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J User Exper 2009;4(3):114-123 [FREE Full text] [doi: 10.5555/2835587.2835589]

XSL•FO
RenderX

29. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Hum Comput Interact 2008 Jul 29;24(6):574-594. [doi: 10.1080/10447310802205776]
30. System usability scale – jetzt auch auf deutsch. SAP Website. 2023. URL: https://blogs.sap.com/2016/02/01/system-usability-scale-jetzt-auch-auf-deutsch/ [accessed 2025-01-22]
31. Vitale M, Bruno V, D'Abbronzo G, et al. Evaluation of eating habits by 7-day food record: web-PC vs. traditional paper format. Int J Food Sci Nutr 2023;74(4):580-587. [doi: 10.1080/09637486.2023.2236808] [Medline: 37455370]
32. Wellard-Cole L, Potter M, Jung JJ, Chen J, Kay J, Allman-Farinelli M. A tool to measure young adults' food intake: design and development of an australian database of foods for the eat and track smartphone app. JMIR mHealth uHealth 2018 Nov 7;6(11):e12136. [doi: 10.2196/12136] [Medline: 30404768]
33. Lucassen DA, Brouwer-Brolsma EM, van de Wiel AM, Siebelink E, Feskens EJM. Iterative development of an innovative smartphone-based dietary assessment tool: Traqq. J Vis Exp 2021 Mar 19(169). [doi: 10.3791/62032] [Medline: 33818566]
34. Gilsing A, Mayhew AJ, Payette H, et al. Validity and reliability of a short diet questionnaire to estimate dietary intake in older adults in a subsample of the canadian longitudinal study on aging. Nutrients 2018 Oct 17;10(10):1522. [doi: 10.3390/nu10101522] [Medline: 30336568]
35. Carter MC, Albar SA, Morris MA, et al. Development of a UK online 24-h dietary assessment tool: myfood24. Nutrients 2015 May 27;7(6):4016-4032. [doi: 10.3390/nu7064016] [Medline: 26024292]
36. Kirkpatrick SI, Gilsing AM, Hobin E, et al. Lessons from studies to evaluate an online 24-hour recall for use with children and adults in canada. Nutrients 2017 Jan 31;9(2):100. [doi: 10.3390/nu9020100] [Medline: 28146125]
37. Rutishauser IHE. Dietary intake measurements. Public Health Nutr 2005 Oct;8(7A):1100-1107. [doi: 10.1079/phn2005798] [Medline: 16277822]
38. Carter MC, Burley VJ, Nykjaer C, Cade JE. "My Meal Mate" (MMM): validation of the diet measures captured on a smartphone application to facilitate weight loss. Br J Nutr 2013 Feb 14;109(3):539-546. [doi: 10.1017/S0007114512001353] [Medline: 22717334]
39. Kupis J, Johnson S, Hallihan G, Olstad DL. Assessing the usability of the Automated Self-Administered Dietary Assessment Tool (ASA24) among low-income adults. Nutrients 2019 Jan 10;11(1):132. [doi: 10.3390/nu11010132] [Medline: 30634581]
40. Ocké MC, Westenbrink S, van Rossum CTM, Temme EHM, van der Vossen-Wijmenga W, Verkaik-Kloosterman J. The essential role of food composition databases for public health nutrition – experiences from the Netherlands. J Food Compost Anal 2021 Aug;101:103967. [doi: 10.1016/j.jfca.2021.103967]
41. Westenbrink S, van der Vossen-Wijmenga W, Toxopeus I, Milder I, Ocké M. LEDA, the branded food database in the Netherlands: data challenges and opportunities. J Food Compost Anal 2021 Sep;102:104044. [doi: 10.1016/j.jfca.2021.104044]
42. Pravst I, Hribar M, Žmitek K, Blažica B, Seljak BK, Kušar A. Branded foods databases as a tool to support nutrition research and monitoring of the food supply: insights from the Slovenian composition and labeling information system. Front Nutr 2021;8:798576. [doi: 10.3389/fnut.2021.798576] [Medline: 35059426]
43. Cade JE, Warthon-Medina M, Albar S, et al. DIET@NET: Best Practice Guidelines for dietary assessment in health research. BMC Med 2017 Nov 15;15(1):202. [doi: 10.1186/s12916-017-0962-x] [Medline: 29137630]
44. 170.000 Lebensmittelprodukte auf dem deutschen Markt. Lebensmittelverband Deutschland. 2023. URL: https://www.lebensmittelverband.de/de/aktuell/20200731-tag-der-lebensmittelvielfalt [accessed 2025-01-22]

Abbreviations

DONALD study: Dortmund Nutritional and Anthropometric Longitudinally Designed study
DR: dietary record
EAN: European Article Number
GTIN: Global Trade Item Number
LEBTAB: LEBensmittelTABelle, food table
SAP SE: Systems, Applications & Products in Data Processing Societas Europaea
SUS: system usability scale
WDR: weighed dietary record

Edited by A Kushniruk; submitted 31.05.24; peer-reviewed by C Collins, D Budiningsari; revised version received 19.12.24; accepted 22.12.24; published 10.02.25.

Please cite as:
Klasen L, Koch SAJ, Benz ME, Conrad J, Alexy U, Blaszkiewicz K, Andone I, Nöthlings U
NutriDiary, a Smartphone-Based Dietary Record App: Description and Usability Evaluation
JMIR Hum Factors 2025;12:e62776
URL: https://humanfactors.jmir.org/2025/1/e62776
doi:10.2196/62776

XSL•FO
RenderX

© Linda Klasen, Stefanie Anna Julia Koch, Maike Elena Benz, Johanna Conrad, Ute Alexy, Konrad Blaszkiewicz, Ionut Andone, Ute Nöthlings. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 10.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Think-Aloud Testing of a Companion App for Colonoscopy Examinations: Usability Study

Christine Jacob[1], PhD; Roman Müller[1], MSc; Sonja Schüler[1], PhD; Alix Rey[2,3], PhD; Guillaume Rey[2], PhD; Berj Armenian[4*], MD; Alain Vonlaufen[4*], MD, PhD; Michael Drepper[4*], MD; Marius Zimmerli[5*], MD

[1]FHNW University of Applied Sciences and Arts Northwestern Switzerland, Riggenbachstrasse 16, Olten, Switzerland

[2]Gimini Biosciences SàRL, Geneva, Switzerland

[3]International Institute for Management Development, Lausanne, Switzerland

[4]GGHA Geneva Gastroenterology & Hepatology Associates SA, Geneva, Switzerland

[5]Clarunis University Digestive Health Care Center, Basel, Switzerland

[*]these authors contributed equally

Corresponding Author:
Christine Jacob, PhD
FHNW University of Applied Sciences and Arts Northwestern Switzerland, Riggenbachstrasse 16, Olten, Switzerland

Abstract

Background: Colonoscopies are vital for initial screening, follow-ups, surveillance of neoplasia, and assessing symptoms such as rectal bleeding. Successful colonoscopies require thorough colon preparation, but up to 25% fail due to poor preparation. This can lead to longer procedures, repeat colonoscopies, inconvenience, poorer health outcomes, and higher costs. eHealth tools can enhance bowel preparation and potentially reduce the need for repeat procedures.

Objective: This usability study aimed to identify strengths and weaknesses in a prototype companion app for colonoscopy examinations. The objective was to obtain in-depth insights into the app's usability, ease of use, and content comprehension, with the aim of refining the tool to effectively fulfill its intended purpose, guided by feedback from potential users.

Methods: From February to August 2024, we conducted a qualitative study using the think-aloud procedure. Each session involved 6 tasks and a semistructured interview to delve deeper into participants' task experiences. All think-aloud sessions and interviews were recorded. Quantitative usability questions were analyzed using Microsoft Excel, while qualitative data underwent coding and analysis based on thematic analysis principles.

Results: In total, 17 individuals, all smartphone users, participated in this study. Participants were recruited from 1 hospital, 1 private clinic, and 1 patient organization in Switzerland. The study found that participants rated the app's usability metrics positively, with an overall mean rating of ease of use at 4.29 (SD 0.59), usefulness at 4.53 (SD 0.72), and comprehensibility at 4.29 (SD 0.92). For the individual features, the mean ratings for ease of use were between 4 and 4.65, usefulness ranged from 4.35 to 4.82, and comprehensibility received ratings between 4.29 and 4.53, all measured on a 5-point scale, where 1 represented low agreement and 5 indicated high agreement. Additionally, 100% of participants indicated they will or may use the app if they require a colonoscopy examination. Participants highlighted the need for reminders and alerts in the week leading up to the colonoscopy, along with tailored content, simplified language, and visual aids.

Conclusions: The app prototype demonstrated favorable results with the majority of participants, and the testing process enabled the prompt identification and resolution of usability issues. The next phase will prioritize and assess potential improvements based on urgency and feasibility to guide a focused development plan. Usability testing highlighted features such as push notifications and personalized content as top priorities for participants, making them key areas for immediate attention. Moving forward, the app has the potential to function effectively as a companion app for colonoscopy examinations. To achieve this, further studies with a larger sample in real-world settings will be crucial.

(JMIR Hum Factors 2025;12:e67043) doi:10.2196/67043

KEYWORDS

eHealth; mobile health; mHealth; digital health; technology assessment; technology adoption; technology implementation; usability study; colonoscopy; app; application; examinations; smartphone; usability

XSL•FO
RenderX

Introduction

Background

Colonoscopies are widely recognized as the most reliable method for detecting colorectal issues; their effectiveness and safety hinge on the thoroughness of bowel preparation beforehand [1,2]. Ensuring adequate bowel preparation is crucial for achieving clear visualization of the colon's inner lining during the procedure. Poor bowel preparation is linked to risks such as missed significant lesions, procedural challenges, longer operation times, higher rates of interval colorectal cancers, and increased health care expenses [3]. However, a colonoscopy is an invasive procedure that demands extensive preparation. This includes taking a laxative, restricting food and liquid intake, and stopping certain medications in the week before the appointment.

Research indicates that up to 11% of individuals miss their colonoscopy appointments [4,5], and among those who do attend, many have insufficient bowel preparation, hindering clear colonic visualization [6-12]. A recent study investigated the efficacy of various bowel preparation regimens (4L, 2L, and ≤1L) for colorectal cancer screening, focusing on key quality indicators such as bowel cleanliness, cecal intubation rate, adenoma detection rate, and polyp detection rate, all aligned with the performance standards set by the European Society of Gastrointestinal Endoscopy (ESGE) [13]. While all regimens met the ESGE's minimum quality thresholds, the adequacy of bowel preparation varied significantly between volumes [13]. Ultralow-volume preparations achieved an adequacy rate of 79%, notably lower than the 86.4% seen with high-volume preparations [13]. In particular, bowel preparation with sodium picosulfate and magnesium citrate (SPMC) and 1L polyethylene glycol with ascorbic acid (1L-PEGA) was adequate in only 75.2% and 82.9% of cases, respectively, highlighting the need for careful consideration when selecting a preparation method based on patient needs and procedural goals [13].

Misunderstanding dietary guidelines and cleansing instructions, along with noncompliance, significantly contributes to inadequate bowel preparation [14]. Ineffective bowel preparation can lead to several adverse outcomes, including reduced adenoma detection rates, extended procedure times, lower cecal intubation rates, increased electrocautery risks, and more frequent examination intervals [15,16]. To enhance patient adherence to colonoscopy procedures, various educational strategies have been used. Tools such as booklets, cartoons, and SMS text messaging have proven effective in increasing follow-up rates compared to standard care [17]. Furthermore, smartphone-based strategies have been developed to assist patients in preparing for colonoscopy [18-24]. Research indicates that these smartphone interventions generally lead to better outcomes, such as higher bowel cleansing quality scores, compared to usual care control groups [18-20,25,26]. However, there is limited evidence that these tools were designed with input from their intended users [27], which may reduce their effectiveness. Engaging potential users in the app development process is likely to enhance usability by ensuring the app's content and features match user needs and preferences [28-30].

Objectives

Considering that user research can significantly enhance a tool's adoption and adherence rates post launch [28], this usability study was conducted to pinpoint strengths and weaknesses in the prototype of a companion app for colonoscopy examinations and to provide detailed insights into its quality regarding usefulness, ease of use, and content comprehension. The goal is to refine the product based on evidence gathered from potential users, ensuring it meets its intended purpose effectively.

Methods

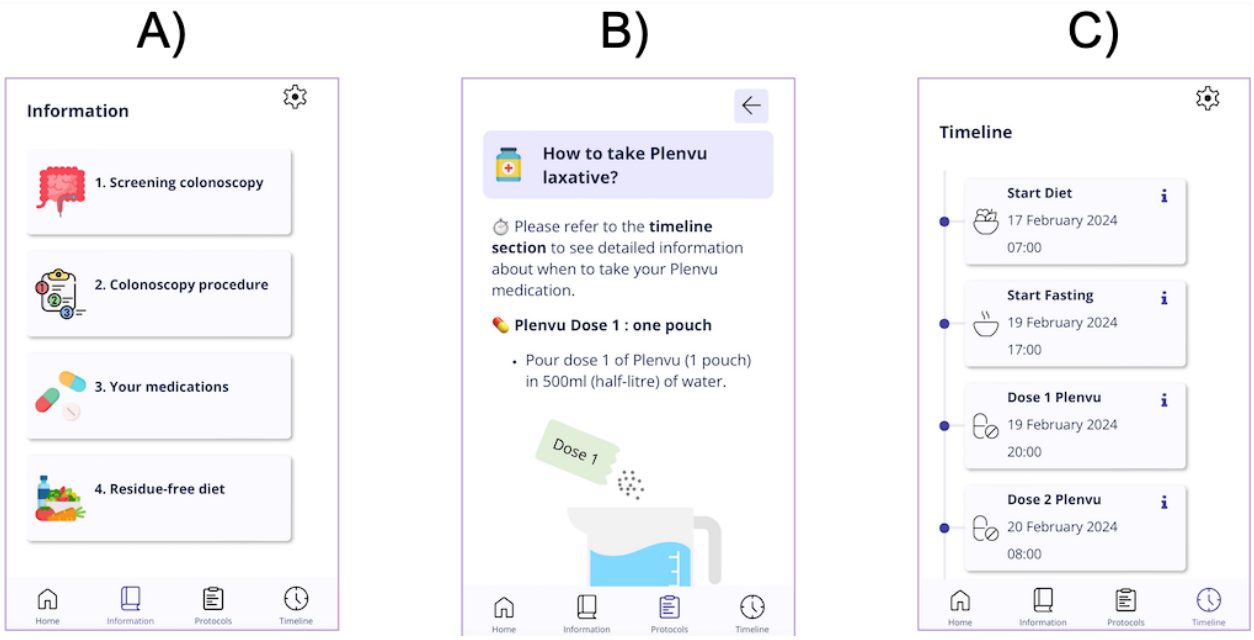
App Prototype

The health care technology company Gimini Biosciences SàRL is developing digital health solutions to empower patients undergoing complex medical interventions. Their mission is to ensure every patient has access to the necessary information for a successful medical examination.

The companion app for colonoscopy examinations, their first use case, features three main sections: (1) educational content about colonoscopy examinations; (2) digital protocols detailing diet, fasting, and laxative instructions; and (3) a personalized timeline based on the user's examination date and time to guide them on diet, fasting, and laxative schedules. Figure 1 illustrates the design of the prototype, showcasing these sections.

XSL•FO
RenderX

Figure 1. App prototype overview, including (A) patient information, (B) digital protocols, and (C) timeline and schedule.



Study Design

The study involved conducting a qualitative interview with each participant using the standardized think-aloud method [31], which is described in detail below. Participants were guided using a semistructured test script to articulate their cognitive process in real time while completing a list of predefined tasks. Similar to many qualitative studies, this research used purposive sampling to gather in-depth insights [32]. Participants were selected based on their capacity to offer detailed and firsthand insights into the research topic, ensuring they could effectively articulate their real-life experiences [32,33].

Participant Recruitment

The inclusion criteria covered individuals aged 40 - 65 years who use a smartphone, have undergone a colonoscopy examination, have access to Wi-Fi and email, are comfortable using teleconferencing tools (eg, Microsoft Teams), and are capable of screen-sharing during the testing session. Three institutions played a key role in recruiting participants: Clarunis (the University Digestive Health Care Center of St. Clara Hospital and University Hospital Basel), GGHA (Geneva Gastroenterology & Hepatology Associates SA), and EUPATI Switzerland (European Patients' Academy on Therapeutic Innovation).

Participants were directly approached and recruited by the collaborating institutions. Once a participant agreed to participate and signed the consent form, the respective institution forwarded the signed consent form along with the participant's contact details to the core research team at the University of Applied Sciences and Arts Northwest Switzerland (FHNW). From that point, the research team took over the coordination and management of the participant's involvement in the study.

Participants had the option to conduct the testing session in English, French, or German. The English version of the participant information sheet and consent form can be found in Multimedia Appendix 1. The information was also provided in both German and French, giving participants the freedom to select the language they felt most comfortable using, and the recruitment process spanned from February to July 2024. The researchers aimed to recruit a sufficient number of participants to achieve saturation, indicating that enough data had been collected when new information no longer provided additional insights [32,33].

Think-Aloud Procedure

The think-aloud method is extensively used in app development as a popular tool for assessing usability [34]. In a think-aloud test, participants are asked to use the system and verbalize their thoughts continuously as they navigate through the user interface. This technique allows researchers to gain deeper insights into user misconceptions, which often lead to practical recommendations for redesign: when users misunderstand design elements, those elements may require modification. It also sheds light on why users make incorrect assumptions about certain aspects of the design and why they find other parts intuitive. The method is used to make cognitive processes, which would otherwise remain implicit and unspoken, more visible. By having participants verbalize their thoughts while performing a task, researchers can gain insights into thought processes, decision-making strategies, and individual patterns of interpretation.

Participants evaluated a web-based prototype of the app. They were instructed to explore the prototype while vocalizing their actions and observations, and to offer feedback on features, navigation, and perceived usefulness as they interacted with it. Researchers reminded participants to maintain a continuous stream of thoughts when needed and observed their behavior throughout the test tasks. All sessions were recorded, capturing participants' screen interactions with the prototype. Detailed notes were taken during feedback sessions and stored using the research platform Tivian.

During each test session, participants were asked to complete 6 tasks, during which we observed their behavior and

XSL•FO
RenderX

interactions with the tool. Additionally, they provided subjective assessments for each task, evaluating its usefulness, ease of use, and comprehensibility of the content. Textbox 1 displays the list of tasks participants were required to complete during the test sessions. The English version of the complete test script can be found in Multimedia Appendix 2.

Textbox 1. The list of tasks participants were required to complete during the test sessions.

Task 1: Customization and language choice
Task 2: First screen tour
Task 3: Accessing and navigating the background information section
Task 4: Accessing and navigating the fasting and food instructions
Task 5: Accessing and navigating the laxative instructions
Task 6: Accessing and navigating the meal and exercise recommendations

The think-aloud test was designed to encompass every possible task a user might perform within the app, covering all available functions and interactions. By including all tasks, we ensured a thorough capture of user feedback across the app's entire functionality, providing a comprehensive assessment of the user experience. Below is a high-level summary of what each task involved:

- Task 1: Customization and language selection—In this task, participants were prompted to select their preferred language and adjust settings such as font size or content complexity, allowing them to personalize the app to their needs.
- Task 2: First screen tour—This task aimed to assess the intuitiveness of the app's initial screen, encouraging participants to reflect on the clarity of the navigation and their understanding of the content available in each section.
- Task 3: Exploring background information—Here, participants were guided to access and navigate the section on background information about the colonoscopy examination, supporting their understanding of the procedure as they prepared for their appointment.
- Task 4: Reviewing fasting and dietary instructions—This task directed participants to find and review guidelines on what they can and cannot eat before the examination, including details on when to begin the special diet and the required fasting period.
- Task 5: Reviewing laxative instructions—Participants were asked to locate and understand instructions on taking the laxative, including information on dosage timing for effective preparation.
- Task 6: Reviewing meal and exercise recommendations—This task involved participants exploring the app's suggestions for meals and exercises that could help them prepare optimally for the examination.

Analysis

After compiling all notes from the testing sessions, CJ, RM, and SS collectively reviewed and synthesized the findings. These were then presented to the app development team for discussion, focusing on identifying features and functions requiring modification. The data encompassed audio recordings of the think-aloud sessions, observations noted by CJ, RM, and SS, and responses to usability questions.

Quantitative usability data, such as user ratings of ease of use and usefulness of the different features, were aggregated and analyzed using Microsoft Excel for Mac 2021 (version 16.86) to compute totals, percentages, means, and standard deviations. Qualitative participant comments were translated into English as needed for coding purposes. NVivo version 1.7.2 (QSR International), a qualitative data analysis software, was used for coding and categorizing the qualitative data. The data underwent thematic analysis to capture the depth and the unique interpretative contributions of individual researchers [35]. To ensure coding reliability, the first 3 authors, who conducted testing sessions in 3 different languages, engaged in collaborative discussions. The initial codebook was organized around core eHealth app development components: user interface (including navigation and visual design), user experience design (such as personalization), functionality (including notifications and reminders), and patient engagement and support (such as educational content). Initial coding was carried out by the first author (CJ), followed by a review from RM. Any coding discrepancies were addressed through discussions with SS until a consensus was reached.

Ethical Considerations

The ethics committee of Northwest and Central Switzerland determined that ethics approval was not needed for this study, according to the Federal Act on Research Involving Human Beings, article 2, paragraph 1 (reference number Req-2023 - 01506). All participants were briefed about the research background and signed a consent form agreeing to participate. Participants did not receive payment but were offered the opportunity for early and free access to the app upon its launch.

Results

Sample Characteristics

In total, 17 participants from 4 institutions tested the app prototype (3 of them were pilot tests, 1 in each of the 3 test languages). Participants were mostly male (13/17, 76%) and aged 40 - 50 (8/17, 47%) years. The gender discrepancy among participants is largely due to recruitment challenges and the voluntary nature of the study. Since participation was optional, it may have led to a self-selection bias, where individuals more comfortable with technology or who have a specific interest in health applications were more likely to take part, resulting in a less balanced gender distribution among participants. Tests were conducted in German (6/17, 35%), French (6/17, 35%), and

XSL•FO
RenderX

English (5/17, 29%). Table 1 presents the demographics and characteristics of the sample.

Table . Sample characteristics (N=17).

Characteristic		Values, n (%)
Gender		
	Male	13 (76)
	Female	4 (24)
Age (years)		
	<40	1 (6)[a]
	40 - 50	8 (47)
	51 - 60	5 (29)
	>60	3 (18)
Language		
	German	6 (35)
	French	6 (35)
	English	5 (29)
Referring institution		
	Clarunis[b]	9 (53)
	GGHA[c]	4 (24)
	FHNW[d] (pilot testers)	3 (18)
	EUPATI Switzerland[e]	1 (6)

[a]One of the pilot tests, hence age inclusion criteria were not applied.

[b]Clarunis: The University Digestive Health Care Center of St. Clara Hospital and University Hospital Basel.

[c]GGHA: Geneva Gastroenterology & Hepatology Associates SA.

[d]FHNW: University of Applied Sciences and Arts, Northwestern Switzerland.

[e]EUPATI Switzerland: European Patients' Academy on Therapeutic Innovation, Switzerland.

Usability Metrics

Usability metrics are specific measurements used to evaluate a digital product's usability. These metrics typically assess how quickly users complete tasks, how often they make mistakes, and their overall satisfaction with the tool. By analyzing various usability metrics, we can gain a comprehensive understanding of the user's experience and the tool's overall usability.

During the testing sessions, we incorporated satisfaction metrics, which are subjective measures based on users' self-assessments. These metrics evaluated the ease of use, usefulness, and content comprehensibility of different sections of the app, as well as their overall impression of the app. Responses were rated on a Likert scale from 1 (low agreement) to 5 (high agreement). Table 2 provides an overview of these subjective measures for the various tasks and the app as a whole. Some measures were not applicable for certain tasks. For example, in task 1 (customization and language selection), there was no text content to assess for comprehensibility, so this measure is marked as N/A (not applicable) for that task. In task 2 (first screen tour), participants were only asked to reflect on whether the initial screen was intuitive, rather than completing an action. Therefore, measures such as ease of use, usefulness, and comprehensibility were not rated for this task.

XSL•FO
RenderX

Table . Subjective usability measures (N=17).

Task	Ease of use[a], mean (SD)	Usefulness[a], mean (SD)	Comprehensibility[a], mean (SD)
Task 1: Customization and language choice	4.41 (0.71)	4.35 (1.11)	N/A[b]
Task 2: First screen tour	N/A	N/A	N/A
Task 3: Accessing and navigating the background information section	4.41 (0.94)	4.76 (0.44)	4.53 (0.87)
Task 4: Accessing and navigating the fasting and food instructions	4 (1)	4.41 (0.87)	4.29 (0.99)
Task 5: Accessing and navigating the laxative instructions	4.53 (0.62)	4.82 (0.39)	4.35 (0.93)
Task 6: Accessing and navigating the meal and exercise recommendations	4.65 (0.7)	4.59 (0.62)	4.53 (0.8)
Overall satisfaction with the app as a whole	4.29 (0.59)	4.53 (0.72)	4.29 (0.92)

[a]User satisfaction of usability attributes was rated on a scale of 1 (low agreement) to 5 (high agreement).

[b]N/A: not applicable.

To provide a comprehensive assessment, test moderators also recorded observations on three additional usability metrics: (1) completion metrics (these measure whether users can successfully complete or partially complete tasks, indicating the tool's effectiveness); (2) duration metrics (these track the average time users take to perform a task, reflecting the design's complexity and the efficiency of user navigation); and (3) error metrics (these refer to actions users take that do not lead to the expected outcome, highlighting areas of confusion in the user interface or challenges with functionality). Table 3 presents an overview of these observed measures. Certain measures in this table were not applicable for task 2 (first screen tour), as participants were only asked to assess the intuitiveness of the initial screen rather than perform any specific action. As a result, metrics such as completion rate and error rate were not relevant for this task, since there was no actionable step for participants to complete or errors to quantify.

Table . Observed usability measures (N=17).

Task	Completion, n (%)	Error rate, n (%)	Duration (seconds), mean (SD)
Task 1: Customization and language choice	16 (94)	1 (6)	44.47 (32.39)
Task 2: First screen tour	N/A[a]	N/A	57.41 (45.94)
Task 3: Accessing and navigating the background information section	15 (88)	7 (41)	66.53 (46.32)
Task 4: Accessing and navigating the fasting and food instructions	16 (94)	7 (41)	75.41 (46.07)
Task 5: Accessing and navigating the laxative instructions	15 (88)	3 (18)	54.88 (47.89)
Task 6: Accessing and navigating the meal and exercise recommendations	17 (100)	3 (18)	56.76 (50.49)

[a]N/A: not applicable.

When asked if they would use the app for a colonoscopy examination once it becomes available, out of 17 participants, 14 (82%) said yes, 3 (18%) said maybe, and 0 (0%) said no.

Qualitative Feedback

Participants provided comments and qualitative feedback on the 6 tasks they performed during the test sessions. While their feedback concerning the comprehensibility of the content, the user-friendliness, and the general user benefits of the application was mostly positive, they also expressed some confusion about certain content or features. Additionally, they offered suggestions for improvements in areas such as design and visualization as well as user guidance to address the gaps or issues they identified in the prototype. Overall, participants stressed the importance of receiving reminders and alerts in the week leading up to their colonoscopy. They also preferred tailored content, simplified language, and visual aids to enhance their understanding. Table 4 summarizes their qualitative feedback, organized into 4 key themes: app content, design, guidance, and features. These themes collectively illustrate the

XSL•FO
RenderX

participants' preferences and priorities, providing valuable insights for the development and improvement of the application.

XSL•FO
RenderX

Table . Key themes that emerged from the thematic analysis of the qualitative feedback.

Themes and subthemes		Qualitative feedback
Content		
	Content accuracy	Participants expressed a need for greater accuracy in the app's content. Specifically, they suggested: • Allowing the selection of examination times down to the minute, rather than in 15-minute intervals. • Clearly stating that tea and hot drinks should not contain milk, as specified in the paper-based guidance. • Ensuring the app accurately reflects the strict fasting phase, as some participants experienced longer fasting periods than indicated.
	Content clarity	Some content has been found to be unclear or confusing. For instance: • The term "residue-free diet" may be difficult for laypeople to understand; "bowel cleansing" is suggested as a more straightforward alternative. • The term "protocols" is ambiguous, and "preparation" is recommended as a clearer option. • Instead of using the word "option" for meal examples, provide explicit examples to avoid misleading users into thinking these are the only choices available. • For dosage instructions, emphasize that the liquid should be consumed in sips rather than all at once. • The timeline entry "pickup medication" is unclear and potentially misleading; a more descriptive label is needed. • More information on physical exercise should be included, such as its effects and importance in the preparation process.
	Content completeness	It has been noted that certain content is currently lacking in the app, namely: • Expanded details on the colonoscopy procedure should be provided, with careful attention to wording. Starting the information section with cancer detection details can be perceived as alarming. It is essential to introduce the procedure in a reassuring manner, highlighting its benefits and emphasizing preventive care. • Information on the effects of laxatives. • Vegetarian menu options. • Post-colonoscopy care and potential adverse events. • Any assistance with laxative consumption would be greatly appreciated, as it can be quite unpleasant. For example, mixing the laxative with clear syrup to improve taste. Another tip is to sip water alternately with the laxative instead of consuming them sequentially, which can make the experience less difficult. Additional advice on making the process more manageable could include using a straw to drink the laxative or drinking tea beforehand to mitigate the salty taste.

XSL•FO
RenderX

Themes and subthemes		Qualitative feedback
	FAQs[a]	Participants have suggested including a section with FAQs to provide clear and concise answers to common concerns. Suggested questions include: • Why is a colonoscopy performed? Include an explanation of its purpose and benefits, and note any differences in the procedure or considerations for men and women. • Can I drive after taking the medication? • When can I resume a normal diet? • What are the advantages and disadvantages of various examination methods? Offer a comparison of different diagnostic options, such as stool examinations, to help users understand their relative benefits and limitations.
	Translation accuracy	A few minor translation inaccuracies have been identified: • “Dose 1” was incorrectly translated into German; the correct term is “Dosis.” • Under “fasting time,” “Untersuchung” is a more appropriate translation than “Prüfung.”
Design		

XSL•FO
RenderX

Themes and subthemes		Qualitative feedback
	Navigation clarity	Participants have proposed several improvements to enhance the clarity of the app's navigation. Suggested enhancements include: • Adjust the color scheme of the language selection navigation to make it more prominent. • The button for confirming language selection should be available in multiple languages. • Change the name of the navigation section labeled "Information" to better reflect that it pertains to the procedure rather than app-related information. • The back arrow should be made more prominent by increasing its size and repositioning it centrally for better visibility and ease of use. • There is uncertainty about the type of information the sections labeled "Protocols" and "Timeline" contain.
	Navigation structure	There were several suggestions to improve the app's navigation structure: • Increase the size of the customization icon and consider positioning it at the same level as the other main navigation icons to attract more attention. • Keep content concise and allow users to click for more detailed information when needed. • Display the timeline immediately after scheduling an appointment, as it is highly relevant. Ideally, place the timeline in the same section where laxative information is provided. • Clearly mark the section for taking the laxative in the protocol and integrate the "how" with the "when" to provide a comprehensive guide. • Distinguish and prioritize personalized content related to individual preparation for the colonoscopy from general informational content about the procedure, such as examination details and menu suggestions.
Guidance		

XSL•FO
RenderX

Themes and subthemes		Qualitative feedback
	Step-by-step guidance	Participants have reported some confusion and expressed a need for clearer guidance on navigating the app and understanding required actions. Specific recommendations include: • Add a note indicating that users can change the language and the date of their intervention, as this option is not immediately apparent. • Clearly explain why users need to enter their appointment details. • Provide a clear explanation of the next steps immediately after entering appointment information, as some users were unsure where to click. • Indicate that some content is customized based on the user's appointment time to clarify how the app personalizes information. • Include preliminary information about the general procedure at the start of the app, such as an overview page, to give users a better understanding of what to expect.
Features		
	Personalized notifications	Nearly all participants emphasized the critical need for personalized reminders and alerts, tailored to each patient based on the specific date and time of their colonoscopy appointment.
	Sharing and printing info	The importance of having the ability to print or share information directly from the app was emphasized.
	Tailored content	Participants noted that the app's capability to customize content for individual users makes it more favorable compared to other methods of information delivery, such as paper. Specific suggestions include: • Set the app to automatically select the language based on the user's browser settings. • Allow direct integration with the user's calendar for streamlined timeline management. • Ensure that the link directs users to the correct hospital right from the start. • Add the medical center's phone number or provide a direct link for users to easily contact the center.
	Visual aids	Participants recommended integrating visual aids to improve the clarity and comprehension of the information. Suggestions include: • Add illustrations to assist in understanding the content more easily. • Use visual cues to highlight the fasting time more prominently. • Present information about meal preparation and dishes in a more visual format, including photos of menu items. • Integrate photos and potentially a video to visually explain the procedure.

[a]FAQs: frequently asked questions.

XSL•FO
RenderX

Discussion

Principal Findings and Implications for App Improvement

Our findings indicate that participants were generally willing to use the companion app for colonoscopy examinations and found it mostly useful and easy to navigate. Usability is crucial for implementation, as technologies that are user-friendly are more likely to be consistently used over time. This is why leading technology acceptance frameworks such as the Technology Acceptance Model (TAM) and the Unified Theory of Acceptance and Use of Technology (UTAUT) emphasize ease of use and usefulness as key predictors for adopting new technologies [36,37]. The inclusion of subjective measures in assessments has been debated in the literature due to potential variability introduced by users' subjective views. Despite this challenge, many scholars advocate for the inclusion of subjective criteria, such as ease of use and visual appeal, because they are fundamental drivers of adoption [38-41]. Therefore, incorporating subjective criteria, such as perceived ease of use, into the testing process could enhance tool adherence and improve health outcomes [42]. However, it is important to recognize that moderating factors such as a person's age, education, and digital skills can influence their assessment of a technology's perceived ease of use and overall usefulness [38].

Although participants rated the app highly for both ease of use and usefulness, they also offered several suggestions for improvement. For instance, they recognized that preparing for a colonoscopy involves multiple complex steps during the week leading up to the appointment. They expressed a desire for the app to send timely reminders to guide users throughout the process, aligning with findings from previous similar research [27]. Reminders are likely crucial for the effectiveness of the app. In fact, a randomized controlled trial found that without reminders the app group had the same quality of bowel cleanliness as the control group [21]. Previous studies have demonstrated that personalized reminders and alerts enhance the adoption of electronic health technologies, effectively modifying various health behaviors [43].

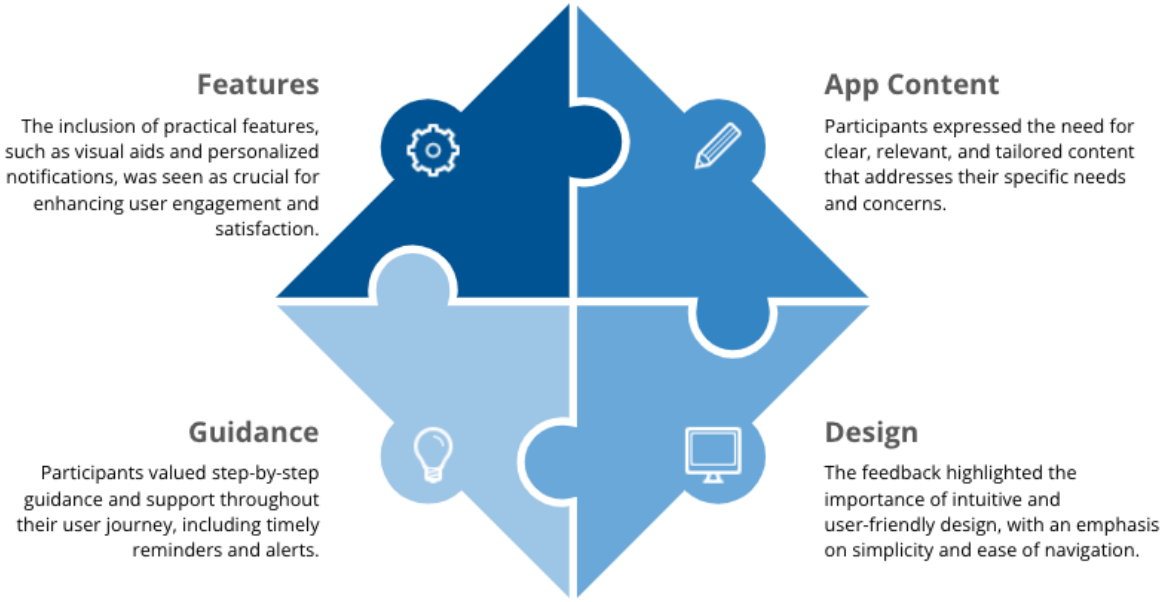
The app content was generally well received for its clarity and ease of understanding. However, there were a few requests for improved clarity or accuracy in certain areas, as detailed in the Results section. Additionally, users identified some content gaps in the tested prototype. Specifically, they requested more comprehensive guidance on how to take the laxative and minimize the unpleasantness of the experience. Furthermore, there was a need for information addressing post-colonoscopy issues. This latter finding aligns with similar research; for instance, Sewitch et al [27] also highlighted the perceived importance of post-colonoscopy information among potential users. Their study similarly underscored the need for detailed guidance on what to expect and how to manage any subsequent issues after the procedure. This is especially pertinent considering that statistics reveal approximately 25% of patients experience a minor adverse event within 48 hours of a colonoscopy, and 0.5% encounter a serious adverse event within 30 days [44].

Participants indicated a need for more detailed, step-by-step guidance on using the app. They sought clear instructions on the use of entering specific information, such as their procedure date and time, and how to effectively use the app to prepare for their examination. This feedback aligns with the extensive literature emphasizing the importance of user training and guidance to enhance adoption rates [38]. Introducing an app tour with guided notes detailing each section, coupled with clear instructions on features such as adjusting font size, could greatly enhance the overall user experience. This initial guidance would make it easier for users to understand how to navigate the app and maximize its use right from the start. Research suggests that users' perception of ease of use can be significantly improved with high-quality training materials that guide them on how to effectively optimize these technologies [38,45,46].

Participants expressed a strong preference for more visual and interactive content, as well as a navigation design that prioritizes personalized information, such as the preparation timeline. They expressed a desire for information to be presented clearly at every stage of the colonoscopy preparation process. This feedback is consistent with previous research, which underscores the critical role of design elements and personalization in achieving successful patient adoption of digital health tools [38]. Several studies have specifically highlighted the importance of personalization. For example, a lack of customization options to meet individual needs can result in lower adoption rates or even abandonment of the tool [38,47]. Figure 2 provides a summary of our findings regarding user preferences for app design, content, features, and guidance. These insights will help boost the app's usability and functionality, ultimately improving user adoption.

XSL•FO
RenderX

Figure 2. Key areas for app improvement.



The next step will involve a thorough prioritization and impact assessment, where improvements according to urgency and feasibility will be ranked to guide a targeted development plan. For example, the usability testing results indicate that features such as push notifications and personalized content are top priorities for the participants. These will be addressed at the start of the next development cycle to ensure they are optimized before moving into real-world testing.

Limitations and Future Research

While the practical insights from participants regarding the app prototype were valuable and guided us in making important modifications, it is important to acknowledge several limitations. Using the think-aloud methodology, there is a potential for reporting bias where participants may shape their responses to align with perceived researcher expectations; additionally, participants might selectively verbalize thoughts that align with what they believe interviewers want to hear, known as social desirability bias [48]. To address these potential biases, moderators were trained to actively encourage participants to provide verbal feedback and to collect observational data. These observations were used alongside participants' subjective feedback during data analysis.

While the sample size of this study may appear small, previous research indicates that 80% - 90% of usability issues in websites and apps can be identified with 5 to 9 participants [49,50]. During the individual think-aloud sessions, recurring themes emerged, suggesting data saturation was achieved [32]. However, it is possible that additional participants could have uncovered different usability issues and provided diverse perspectives. Looking ahead, the app has promising potential as a companion for colonoscopy examinations. However, conducting larger-scale studies in real-world environments will be essential to validate and optimize its value and effectiveness.

Conclusions

Overall, participants expressed satisfaction with the app's usability. The think-aloud sessions provided real-time insights into the app's appeal, relevance, and use. Minor adjustments to the prototype's functionality were identified as necessary to enhance usability. Feedback and suggestions from participants have been integrated into the final app design. The initial findings from this usability study indicate that the app holds promising potential as a companion for colonoscopy examinations. This work establishes the foundation for further research to evaluate usability and feasibility among a larger, real-world population.

Acknowledgments

This study was sponsored by Innosuisse (Swiss Innovation Agency, grant 70325.1 INNO-ICT). The role of Innosuisse is to promote science-based innovation in the interests of industry and society in Switzerland. We extend our gratitude to the 3 institutions that were instrumental in recruiting the study participants: Clarunis (the University Digestive Health Care Center of St. Clara Hospital and University Hospital Basel), GGHA (Geneva Gastroenterology & Hepatology Associates SA), and EUPATI Switzerland (European Patients' Academy on Therapeutic Innovation).

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request. Restrictions may apply to the availability of some of these data due to, for example, privacy and ethical reasons.

XSL•FO
RenderX

Conflicts of Interest

AR and GR are cofounders of Gimini Biosciences SàRL, the technology provider developing the application being studied. CJ is an editorial board member of *JMIR Human Factors* at the time of this publication. The other authors have no conflicts of interest to declare.

Multimedia Appendix 1
Participant information sheet and consent form.
[PDF File, 139 KB - humanfactors_v12i1e67043_app1.pdf]

Multimedia Appendix 2
Think-aloud usability test script.
[PDF File, 244 KB - humanfactors_v12i1e67043_app2.pdf]

References

1. Beck DE. Bowel preparation for colonoscopy. Clin Colon Rectal Surg 2010 Feb;23(1):10-13. [doi: 10.1055/s-0030-1247851] [Medline: 21286285]
2. Corley DA, Jensen CD, Marks AR, et al. Adenoma detection rate and risk of colorectal cancer and death. N Engl J Med 2014 Apr 3;370(14):1298-1306. [doi: 10.1056/NEJMoa1309086] [Medline: 24693890]
3. Lebwohl B, Kastrinos F, Glick M, Rosenbaum AJ, Wang T, Neugut AI. The impact of suboptimal bowel preparation on adenoma miss rates and the factors associated with early repeat colonoscopy. Gastrointest Endosc 2011 Jun;73(6):1207-1214. [doi: 10.1016/j.gie.2011.01.051] [Medline: 21481857]
4. Dalton ARH. Incomplete diagnostic follow-up after a positive colorectal cancer screening test: a systematic review. J Public Health 2018;40(1):e46-e58. [doi: 10.1093/pubmed/fdw147]
5. Chopra D, Hookey LC. Comorbid illness, bowel preparation, and logistical constraints are key reasons for outpatient colonoscopy nonattendance. Can J Gastroenterol Hepatol 2016;2016:2179354. [doi: 10.1155/2016/2179354] [Medline: 27478818]
6. Sulz MC, Kröger A, Prakash M, Manser CN, Heinrich H, Misselwitz B. Meta-analysis of the effect of bowel preparation on adenoma detection: early adenomas affected stronger than advanced adenomas. PLoS One 2016;11(6):e0154149. [doi: 10.1371/journal.pone.0154149] [Medline: 27257916]
7. Govani SM, Elliott EE, Menees SB, et al. Predictors of suboptimal bowel preparation in asymptomatic patients undergoing average-risk screening colonoscopy. World J Gastrointest Endosc 2016 Sep 16;8(17):616-622. [doi: 10.4253/wjge.v8.i17.616] [Medline: 27668072]
8. Yee R, Manoharan S, Hall C, Hayashi A. Optimizing bowel preparation for colonoscopy: what are the predictors of an inadequate preparation? Am J Surg 2015 May;209(5):787-792. [doi: 10.1016/j.amjsurg.2014.12.018] [Medline: 25796095]
9. Dik VK, Moons LMG, Hüyük M, et al. Predicting inadequate bowel preparation for colonoscopy in participants receiving split-dose bowel preparation: development and validation of a prediction score. Gastrointest Endosc 2015 Mar;81(3):665-672. [doi: 10.1016/j.gie.2014.09.066] [Medline: 25600879]
10. Kherad O, Restellini S, Martel M, Barkun AN. Polyethylene glycol versus sodium picosulfalte bowel preparation in the setting of a colorectal cancer screening program. Can J Gastroenterol Hepatol 2015 Oct;29(7):384-390. [doi: 10.1155/2015/350587] [Medline: 26301330]
11. Menees SB, Kim HM, Wren P, et al. Patient compliance and suboptimal bowel preparation with split-dose bowel regimen in average-risk screening colonoscopy. Gastrointest Endosc 2014 May;79(5):811-820. [doi: 10.1016/j.gie.2014.01.024] [Medline: 24631492]
12. Kurlander JE, Sondhi AR, Waljee AK, et al. How efficacious are patient education interventions to improve bowel preparation for colonoscopy? A systematic review. PLoS One 2016;11(10):e0164442. [doi: 10.1371/journal.pone.0164442] [Medline: 27741260]
13. Theunissen F, Lantinga MA, Ter Borg PCJ, et al. Efficacy of different bowel preparation regimen volumes for colorectal cancer screening and compliance with European Society of Gastrointestinal Endoscopy performance measures. United European Gastroenterol J 2023 Jun;11(5):448-457. [doi: 10.1002/ueg2.12386] [Medline: 37190897]
14. Froehlich F, Wietlisbach V, Gonvers JJ, Burnand B, Vader JP. Impact of colonic cleansing on quality and diagnostic yield of colonoscopy: the European Panel of Appropriateness of Gastrointestinal Endoscopy European multicenter study. Gastrointest Endosc 2005 Mar;61(3):378-384. [doi: 10.1016/s0016-5107(04)02776-2] [Medline: 15758907]
15. Alvi H, Rasheed T, Shaikh MA, Ali FS, Zuberi BF, Samejo AA. Impact of bowel preparation on caecal intubation time during colonoscopy. Pak J Med Sci 2019;35(6):1516-1519. [doi: 10.12669/pjms.35.6.1031] [Medline: 31777485]
16. Kluge MA, Williams JL, Wu CK, et al. Inadequate Boston Bowel Preparation Scale scores predict the risk of missed neoplasia on the next colonoscopy. Gastrointest Endosc 2018 Mar;87(3):744-751. [doi: 10.1016/j.gie.2017.06.012] [Medline: 28648575]

XSL•FO
RenderX

17. Liu Z, Zhang MM, Li YY, Li LX, Li YQ. Enhanced education for bowel preparation before colonoscopy: a state-of-the-art review. J Dig Dis 2017 Feb;18(2):84-91. [doi: 10.1111/1751-2980.12446] [Medline: 28067463]
18. Cho J, Lee S, Shin JA, Kim JH, Lee HS. The impact of patient education with a smartphone application on the quality of bowel preparation for screening colonoscopy. Clin Endosc 2017 Sep;50(5):479-485. [doi: 10.5946/ce.2017.025] [Medline: 28669148]
19. Walter B, Schmid R, von Delius S. A smartphone app for improvement of colonoscopy preparation (ColoprAPP): development and feasibility study. JMIR Mhealth Uhealth 2017 Sep 20;5(9):e138. [doi: 10.2196/mhealth.7703] [Medline: 28931498]
20. Walter BM, Klare P, Neu B, Schmid RM, von Delius S. Development and testing of an automated 4-day text messaging guidance as an aid for improving colonoscopy preparation. JMIR Mhealth Uhealth 2016;4(2):e75. [doi: 10.2196/mhealth.5289]
21. Sharara AI, Chalhoub JM, Beydoun M, et al. A customized mobile application in colonoscopy preparation: a randomized controlled trial. Clin Transl Gastroenterol 2017 Jan 5;8(1):e211. [doi: 10.1038/ctg.2016.65] [Medline: 28055031]
22. Jung JW, Park J, Jeon GJ, et al. The effectiveness of personalized bowel preparation using a smartphone camera application: a randomized pilot study. Gastroenterol Res Pract 2017;2017:4898914. [doi: 10.1155/2017/4898914] [Medline: 28928768]
23. Kang X, Zhao L, Leung F, et al. Delivery of instructions via mobile social media app increases quality of bowel preparation. Clin Gastroenterol Hepatol 2016 Mar;14(3):429-435. [doi: 10.1016/j.cgh.2015.09.038] [Medline: 26492848]
24. Back SY, Kim HG, Ahn EM, et al. Impact of patient audiovisual re-education via a smartphone on the quality of bowel preparation before colonoscopy: a single-blinded randomized study. Gastrointest Endosc 2018 Mar;87(3):789-799. [doi: 10.1016/j.gie.2017.09.007]
25. van der Zander QEW, Reumkens A, van de Valk B, Winkens B, Masclee AAM, de Ridder RJJ. Effects of a personalized smartphone app on bowel preparation quality: randomized controlled trial. JMIR Mhealth Uhealth 2021 Aug 19;9(8):e26703. [doi: 10.2196/26703] [Medline: 34420924]
26. Aksan F, Tanriverdi LH, Figueredo CJ, Barrera LC, Hasham A, Jariwala SP. The impact of smartphone applications on bowel preparation, compliance with appointments, cost-effectiveness, and patients' quality of life for the colonoscopy process: a scoping review. Saudi J Gastroenterol 2023;29(2):71-87. [doi: 10.4103/sjg.sjg_207_22] [Medline: 36073572]
27. Sewitch MJ, Fallone CA, Ghali P, Lee GE. What patients want in a smartphone app that supports colonoscopy preparation: qualitative study to inform a user-centered smartphone app. JMIR Mhealth Uhealth 2019 Jul 2;7(7):e12242. [doi: 10.2196/12242] [Medline: 31125310]
28. Jacob C, Bourke S, Heuss S. From testers to cocreators—the value of and approaches to successful patient engagement in the development of eHealth solutions: qualitative expert interview study. JMIR Hum Factors 2022 Oct 6;9(4):e41481. [doi: 10.2196/41481] [Medline: 36102548]
29. Aguayo GA, Goetzinger C, Scibilia R, et al. Methods to generate innovative research ideas and improve patient and public involvement in modern epidemiological research: review, patient viewpoint, and guidelines for implementation of a digital cohort study. J Med Internet Res 2021 Dec 23;23(12):e25743. [doi: 10.2196/25743] [Medline: 34941554]
30. Harte R, Glynn L, Rodríguez-Molinero A, et al. A human-centered design methodology to enhance the usability, human factors, and user experience of connected health systems: a three-phase methodology. JMIR Hum Factors 2017;4(1):e8. [doi: 10.2196/humanfactors.5443]
31. Ericsson KA, Simon HA. How to study thinking in everyday life: contrasting think-aloud protocols with descriptions and explanations of thinking. Mind Cult Act 1998 Jul;5(3):178-186. [doi: 10.1207/s15327884mca0503_3]
32. Braun V, Clarke V. Successful Qualitative Research: A Practical Guide for Beginners: SAGE Publications; 2013.
33. Creswell JW, Poth CN. Qualitative Inquiry and Research Design: Choosing Among Five Approaches: SAGE Publications; 2017.
34. Nielsen J. Thinking aloud: the #1 usability tool. Nielsen Norman Group. 2012 Jan 15. URL: https://www.nngroup.com/articles/thinking-aloud-the-1-usability-tool/ [accessed 2025-01-17]
35. Braun V, Clarke V. What can "thematic analysis" offer health and wellbeing researchers? Int J Qual Stud Health Well-being 2014;9(1):26152. [doi: 10.3402/qhw.v9.26152] [Medline: 25326092]
36. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425. [doi: 10.2307/30036540]
37. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
38. Jacob C, Sezgin E, Sanchez-Vazquez A, Ivory C. Sociotechnical factors affecting patients' adoption of mobile health tools: systematic literature review and narrative synthesis. JMIR Mhealth Uhealth 2022 May 5;10(5):e36284. [doi: 10.2196/36284] [Medline: 35318189]
39. Jacob C, Lindeque J, Klein A, Ivory C, Heuss S, Peter MK. Assessing the quality and impact of eHealth tools: systematic literature review and narrative synthesis. JMIR Hum Factors 2023;10:e45143. [doi: 10.2196/45143]
40. Lagan S, Sandler L, Torous J. Evaluating evaluation frameworks: a scoping review of frameworks for assessing health apps. BMJ Open 2021 Mar 19;11(3):e047001. [doi: 10.1136/bmjopen-2020-047001] [Medline: 33741674]
41. Jacob C, Lindeque J, Müller R, et al. A sociotechnical framework to assess patient-facing eHealth tools: results of a modified Delphi process. NPJ Digit Med 2023 Dec 15;6(1):232. [doi: 10.1038/s41746-023-00982-w] [Medline: 38102323]

XSL•FO
RenderX

42. Uncovska M, Freitag B, Meister S, Fehring L. Rating analysis and BERTopic modeling of consumer versus regulated mHealth app reviews in Germany. NPJ Digit Med 2023 Jun 21;6(1):115. [doi: 10.1038/s41746-023-00862-3] [Medline: 37344556]
43. Marcolino MS, Oliveira JAQ, D'Agostino M, Ribeiro AL, Alkmim MBM, Novillo-Ortiz D. The impact of mHealth interventions: systematic review of systematic reviews. JMIR Mhealth Uhealth 2018 Jan 17;6(1):e23. [doi: 10.2196/mhealth.8873] [Medline: 29343463]
44. Sewitch MJ, Azalgara VM, Sing MF. Screening indication associated with lower likelihood of minor adverse events in patients undergoing outpatient colonoscopy. Gastroenterol Nurs 2018;41(2):159-164. [doi: 10.1097/SGA.0000000000000308] [Medline: 29596130]
45. Louissaint J, Lok AS, Fortune BE, Tapper EB. Acceptance and use of a smartphone application in cirrhosis. Liver Int 2020 Jul;40(7):1556-1563. [doi: 10.1111/liv.14494] [Medline: 32343467]
46. Blonigen D, Harris-Olenak B, Kuhn E, Humphreys K, Timko C, Dulin P. From "step away" to "stand down": tailoring a smartphone app for self-management of hazardous drinking for veterans. JMIR Mhealth Uhealth 2020 Feb 13;8(2):e16062. [doi: 10.2196/16062] [Medline: 32053118]
47. Meyerowitz-Katz G, Ravi S, Arnolda L, Feng X, Maberly G, Astell-Burt T. Rates of attrition and dropout in app-based interventions for chronic disease: systematic review and meta-analysis. J Med Internet Res 2020 Sep 29;22(9):e20283. [doi: 10.2196/20283] [Medline: 32990635]
48. Güss CD. What is going through your mind? Thinking aloud as a method in cross-cultural psychology. Front Psychol 2018;9:1292. [doi: 10.3389/fpsyg.2018.01292] [Medline: 30150948]
49. Faulkner L. Beyond the five-user assumption: benefits of increased sample sizes in usability testing. Behav Res Methods Instrum Comput 2003 Aug;35(3):379-383. [doi: 10.3758/BF03195514]
50. Zapata BC, Fernández-Alemán JL, Idri A, Toval A. Empirical studies on usability of mHealth apps: a systematic literature review. J Med Syst 2015 Feb;39(2):1. [doi: 10.1007/s10916-014-0182-2] [Medline: 25600193]

Abbreviations

1L-PEGA: 1L polyethylene glycol with ascorbic acid
Clarunis: the University Digestive Health Care Center of St. Clara Hospital and University Hospital Basel
ESGE: European Society of Gastrointestinal Endoscopy
EUPATI Switzerland: European Patients' Academy on Therapeutic Innovation
FHNW: University of Applied Sciences and Arts Northwest Switzerland
GGHA: Geneva Gastroenterology & Hepatology Associates SA
SPMC: sodium picosulfate and magnesium citrate
TAM: Technology Acceptance Model
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by A Mavragani; submitted 30.09.24; peer-reviewed by LAK Breuch, MM Kessler; revised version received 08.12.24; accepted 08.12.24; published 12.02.25.

Please cite as:
Jacob C, Müller R, Schüler S, Rey A, Rey G, Armenian B, Vonlaufen A, Drepper M, Zimmerli M
Think-Aloud Testing of a Companion App for Colonoscopy Examinations: Usability Study
JMIR Hum Factors 2025;12:e67043
URL: https://humanfactors.jmir.org/2025/1/e67043
doi:10.2196/67043

© Christine Jacob, Roman Müller, Sonja Schüler, Alix Rey, Guillaume Rey, Berj Armenian, Alain Vonlaufen, Michael Drepper, Marius Zimmerli. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 12.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Comparison of a Novel Machine Learning–Based Clinical Query Platform With Traditional Guideline Searches for Hospital Emergencies: Prospective Pilot Study of User Experience and Time Efficiency

Hamza Ejaz[1], MD; Hon Lung Keith Tsui[2*], MD; Mehul Patel[2], MD; Luis Rafael Ulloa Paredes[2], MSc; Ellen Knights[1], BSc; Shah Bakht Aftab[3*], MD; Christian Peter Subbe[4*], Dr med

[1]Ysbyty Gwynedd, Clinical School, Clinical Research Office, Bangor, United Kingdom

[2]Medwise.ai Ltd, Leeds, United Kingdom

[3]Russells Hall Hospital, Dudley, United Kingdom

[4]North Wales Medical School, Bangor University, Brigantia Building, Bangor, United Kingdom

[*]these authors contributed equally

Corresponding Author:
Christian Peter Subbe, Dr med
North Wales Medical School, Bangor University, Brigantia Building, Bangor, United Kingdom

Abstract

Background: Emergency and acute medicine doctors require easily accessible evidence-based information to safely manage a wide range of clinical presentations. The inability to find evidence-based local guidelines on the trust's intranet leads to information retrieval from the World Wide Web. Artificial intelligence (AI) has the potential to make evidence-based information retrieval faster and easier.

Objective: The aim of the study is to conduct a time-motion analysis, comparing cohorts of junior doctors using (1) an AI-supported search engine versus (2) the traditional hospital intranet. The study also aims to examine the impact of the AI-supported search engine on the duration of searches and workflow when seeking answers to clinical queries at the point of care.

Methods: This pre- and postobservational study was conducted in 2 phases. In the first phase, clinical information searches by 10 doctors caring for acutely unwell patients in acute medicine were observed during 10 working days. Based on these findings and input from a focus group of 14 clinicians, an AI-supported, context-sensitive search engine was implemented. In the second phase, clinical practice was observed for 10 doctors for an additional 10 working days using the new search engine.

Results: The hospital intranet group (n=10) had a median of 23 months of clinical experience, while the AI-supported search engine group (n=10) had a median of 54 months. Participants using the AI-supported engine conducted fewer searches. User satisfaction and query resolution rates were similar between the 2 phases. Searches with the AI-supported engine took 43 seconds longer on average. Clinicians rated the new app with a favorable Net Promoter Score of 20.

Conclusions: We report a successful feasibility pilot of an AI-driven search engine for clinical guidelines. Further development of the engine including the incorporation of large language models might improve accuracy and speed. More research is required to establish clinical impact in different user groups. Focusing on new staff at beginning of their post might be the most suitable study design.

(JMIR Hum Factors 2025;12:e52358) doi:10.2196/52358

KEYWORDS

artificial intelligence; machine learning; information search; emergency care; developing; testing; information retrieval; hospital care; training; clinical practice; clinical experience; user satisfaction; clinical impact; user group; users; study design; mobile phone

Introduction

In making decisions about patient care, clinicians frequently are faced with queries and are often unable to retrieve answers to them in a timely fashion. A systematic review [1] estimates that the per-patient frequency of queries raised by clinicians ranges from 0.4 to 0.8 per patient and that about two-thirds of these queries are left unanswered. This picture has been fairly stable over time despite the broad availability of web-based evidence resources that can answer these questions. Unanswered

XSL•FO
RenderX

questions may lead to suboptimal patient care decisions and are missed opportunities for timely learning and practice improvement.

Medwise.ai is a solution codeveloped with Betsi Cadwaladr University Health Board), which helps clinicians find answers quickly from local guidelines and provide seamless and just-in-time access to high-quality evidence in the context of patient care decision-making in the clinical environment. A previous study [2] has demonstrated that smartphone apps can increase the speed to access guidelines when compared to using desktop computers. However, access to guidelines does not equate to finding answers to clinical questions quickly and effectively at the point of care. The Medwise.ai platform can collate local clinical guidelines and break them down into chunks of content that can be retrieved using natural language processing and information retrieval technologies as bite-sized answers for clinician's questions. We conducted a time-motion analysis, comparing cohorts of junior doctors using an artificial intelligence (AI)–supported search engine and the traditional hospital intranet, to examine the impact of the AI-supported search engine on participant time and workflow when seeking answers to clinical queries at the point of care. The secondary end points of our study were to assess whether effective question answering could lead to a better quality of work life and improve confidence in decision-making for the end user.

This study aimed (1) to gain user feedback for a novel context-specific proof-of-concept search engine and (2) to measure the time required to retrieve clinically relevant information with a novel engine compared to other search engines in the setting of hospital emergency care.

Methods

Study Design

We conducted a prospective direct pre- and postobservational pilot study in teams caring for medical emergency admissions examining the impact of access to Medwise.ai, a clinical query answering platform, versus information retrieval from guidelines saved on the hospital intranet using time-motion study methodology [3].

Study Setting

The study took place at the Ysbyty Gwynedd, Bangor. The Ysbyty Gwynedd is a district general hospital in North Wales with 550 beds covering all major specialties, including a dialysis unit and a 13-bedded critical care unit. Clinicians from 2 units in the hospital were recruited; the acute medical unit has 23 beds and is complemented by a same day emergency care unit for low-risk admissions.

Both units receive direct admissions from primary care and work closely with the emergency department (ED) to manage acute medical presentations. It is important to note that the traditional boundaries between emergency care and acute medicine have become less distinct in recent years due to increasing hospital overcrowding in the United Kingdom. While historically, patients were first seen in the ED by emergency physicians and then referred to the Acute Medical Unit for care by acute and general physicians, many acute and general medical teams now also work within the ED to manage the flow of acute medical patients.

During the day shift (8 AM to 8:30 PM), the on-call team consists of 4 doctors in training (a newly qualified foundation year 1 doctor, 2 core medical trainees with 1‐4 years of experience, and 1 medical registrar with 4 or more years of clinical experience and membership of the Royal College of Physicians) and an on-call consultant with a full specialist qualification. The sole task of the on-call team is the care of emergency admissions and emergencies of inpatients. Patients seen by doctors in training are subsequently reviewed by a consultant as part of the posttake ward round.

Intervention

Medwise.ai is a proof-of-concept search engine combining well-established information retrieval techniques with textual question answering and trained models that determine the best answer within a document. The search platform is available to the participating clinicians over a web app accessible via mobile web browser on both Android and iOS devices. Research staff trained the onboard clinicians and helped to install the Medwise.ai search on the participating clinicians' mobile devices ahead of observations.

Content development for Medwise.ai was informed by informal interviews (Multimedia Appendix 1) with staff and a single focus group, consisting of 14 participants, to identify relevant local guidelines and standard operating procedure documents to be included in the Medwise.ai platform, thus ensuring user buy-in and optimal functionality of the Medwise.ai platform. The local documents shared with Medwise.ai were in PDF or Microsoft Word format. It is important to note that searches using Medwise.ai were limited by the content available in the local repository; if the content did not exist within this repository, the AI would not be able to provide answers to queries. It should be noted that the underlying AI model's technical performance was not the focus of this study.

Research Procedure

A dedicated and trained member of the research team shadowed trainee doctors for 20 complete working days: observation days included 10 days without and 10 days with access to Medwise.ai. The 2 groups of participants were not matched for the hospital intranet versus AI-supported search engine using groups.

Doctors reviewed new admissions and conducted unscheduled reviews of previously admitted patients: the doctors were aware that they were observed, and written consent was taken prior to commencing the direct observation. The observed activity of doctors was entered into a work diary. For practical reasons, all observations were undertaken during office hours.

Sampling and Recruitment

The sample size was chosen pragmatically based on experience with previous time-motion studies [4]. Doctors working as part of the acute medical take throughout Ysbyty Gwynedd were recruited. To be included, participants had to be willing and able to give informed consent for participation in the study, be in possession of a smartphone to access web-based content, be permanent or locum staff, and be qualified as a doctor of any

XSL•FO
RenderX

grade. Participants had the right to withdraw from the study at any time. Participants received vouchers with a value of US $50 for participation in the study. Doctors who were only part of the clinical team for 4 hours or less were excluded from recruitment.

Candidates were approached by the principal investigator (CPS) or the dedicated research team member. Notification of the study was undertaken through existing WhatsApp groups and a Junior Doctors' Forum. Inclusion and exclusion criteria were identified during the initial screening.

Assessments

Participants' gender, specialty, grade, duration of work in their current post, and duration of clinical practice were recorded. Participants were shadowed for the duration of a shift but for a minimum of 6 hours per study day. Shadowing did not include times when doctors were consulting or examining patients.

Task duration for clinical queries was measured in minutes. Tasks were classified using a standardized list. Observed searches for information were classified by clinical content (diagnostic category, diagnostic pathway, treatment protocols, prognostication, scoring systems, normal values, and accessed data source). Confidence in decision-making was assessed at the end of each working day using a validated scale for self-assessment [5]. Confidence in decision-making and user satisfaction were measured using the Likert scale (How satisfied are you with the Medwise.ai product? On a Likert scale of 1-10) and the Net Promoter Score (NPS) [6] (How likely are you to recommend Medwise.ai to a friend or colleague?). NPS is a validated score that is used to measure user satisfaction. It is widely used in industry as a benchmark for comparing products within different industries. A positive NPS is considered as good, a score above 20 as favorable, and above 50 as excellent. Participants were also asked for suggestions for improvements of the platform.

Data Management and Statistics

Statistical analysis was primarily descriptive including key data items related to the frequency of clinical searches, the subject of such searches, the time taken for the searches, and graded feedback in relation to the app. Data were recorded on paper case report forms and subsequently transcribed into a Microsoft Excel database for further analysis. Data were anonymized through the removal of name, case report number, and age.

The comparison was made between the events during 10 observed shifts without the app and 10 observed shifts with the app. Diary data and data from questionnaires were analyzed using SPSS (IBM Corp). Kernel density plots were used to visualize the differences in observed task duration; Welch *t* test (2-tailed) was used to determine whether the apparent differences in distribution were likely to have occurred due to random chance or were due to a real difference in the mean task duration.

Given the sample size and the unknown baseline distribution of the variables in question, statistical comparison data from this study might inform subsequent power calculations for related research or a definitive trial.

Ethical Considerations

National Health Service (NHS) Research Ethics Review was not required as the study involved staff only as participants. The study was approved through Health and Care Research Wales and the Health Research Authority. No adverse events were anticipated. However, monitoring was in place in line with the sponsor safety reporting standard operating procedure. Data from focus group meetings were anonymized and deidentified. Written informed consent was given by all members of staff who were observed during the assessments. Participants were reimbursed with £50 (US $61.92) retail vouchers according to local regulations.

Results

Demographics and Baseline Characteristics

All assessments were undertaken between November 17, 2022, and February 8, 2023 (hospital intranet period) and July 6 and 27, 2023 (AI-supported search engine period).

The hospital intranet and AI-supported search engine groups both comprised 10 doctors each with 5 and 7 female doctors, respectively. The hospital intranet group had a median clinical experience of 23 months. Of these, 6.1 months were in their current specialty. The AI-supported search engine group had a median of 54 months of experience. Of these, 9.6 months were in their current specialty. No statistically significant differences between mean time in current specialty were observed between the groups (t=−1.01; 95% CI −11.4679 to 4.0179; P>.05).

Focus Group

The focus group was held on April 5, 2023, after the completion of the hospital intranet period. In total, 14 clinicians participated including foundation year 1 and 2 doctors, registrars, physician associates, and 3 consultant physicians. Participants were asked a list of prepared questions in relation to search habits, preferred information sources, and views about search engines (Multimedia Appendix 1).

Contributors discussed sources of information that they use for their clinical practice and recommended examples of web resources from other NHS organizations. Pros and cons of alternative models of information provision including share point sites were debated. Clinicians commented on the difficulties to find the right locally authorized resources in locations of the internet and intranet. On the other hand, they valued that Medwise.ai guidance is relevant for the organization in which the search is undertaken. Participants discussed their experience of using resources from other UK and international organizations if there is no identifiable local guidance for a topic. Doctors were concerned that Google searches might identify sources that contain misinformation. The challenge of identifying the right information in real time was described as crucial during the time pressures of acute and emergency care.

Descriptives

Overall, 67 searches were observed in the hospital intranet and 39 searches in the AI-supported search engine period. The mean number of clinical searches performed per shift was 6.7 (SD 3.43) in the hospital intranet group and 3.9 (SD 3.00) in the

XSL•FO
RenderX

AI-supported search engine group. Mobile phones were used for 40 of 67 (60%) searches in the hospital intranet group and 34 of 39 (85%) in the AI-supported search engine group. In the hospital intranet group, 11 searches were via Google. Other sources included apps, the British National Formulary, National Institute for Health and Care Excellence guidelines, and local guidelines. In the AI-supported search engine group, no search required Google.

Searches covered a large number of topics (Multimedia Appendices 2 and 3). The most commonly searched topics in both groups were medication-related (n=18 and n=10, respectively).

Participants in the AI-supported search engine group commented on the impact of Medwise.ai on workflow using free-text responses (Textbox 1).

Textbox 1. Free-text feedback provided by users collected at the end of observation.

"With more training [training of the AI system] this could be very helpful in speeding up carrying out jobs. Bit unfortunate study was carried out in final weeks of rotation, when doctors are the most comfortable with systems and need to search things very rarely" [Participant 1].

"Can see it has the potential to increase speed of doing jobs, thus means we see patients quicker" [Participant 2].

"May expediate time available for patient care if brings up more precise answers. When I know where to find the answers it is easier to go straight to the source" [Participant 3].

"As you get your answers under one second from hospital protocol, it saves time and helps you see more patients. also can provide more time for patient care rather than looking for protocol and guidelines in each app, it will be all in one. Highly recommended" [Participant 4].

"Probably not much difference. Perhaps better used in GP. I think it would be more useful in GP. In hospital it is okay but not majorly useful. I think over all it would be a good idea. But I think I search for stuff more in GP. I think overall ED staff would use it most in hospital" [Participant 5].

"Easy availability of resources saves up time and causes less disruption to the flow of clerking. Easy access to BCU/local guidelines is very helpful as some of these are difficult to find on the intranet. Great idea; would be really helpful" [Participant 6].

"Will reduce search time for searching guidelines resulting in more time for patient care. And all in one place makes it easy to find [information]. I am surprised that I have not needed to access my phone/ the app as yet today as this is unusual. Would usually look at betsinet multiple times during a shift for electrolyte levels. I anticipate that this app would be useful" [Participant 7].

"Makes it easier to search for relevant guidelines/policies/pathways etc. Easy and quick to use. User friendly. Appears to work well for searches during work time. Quickly found everything I searched for" [Participant 8].

"Will help find relevant guidelines much more quickly, which will make work more efficient" [Participant 9].

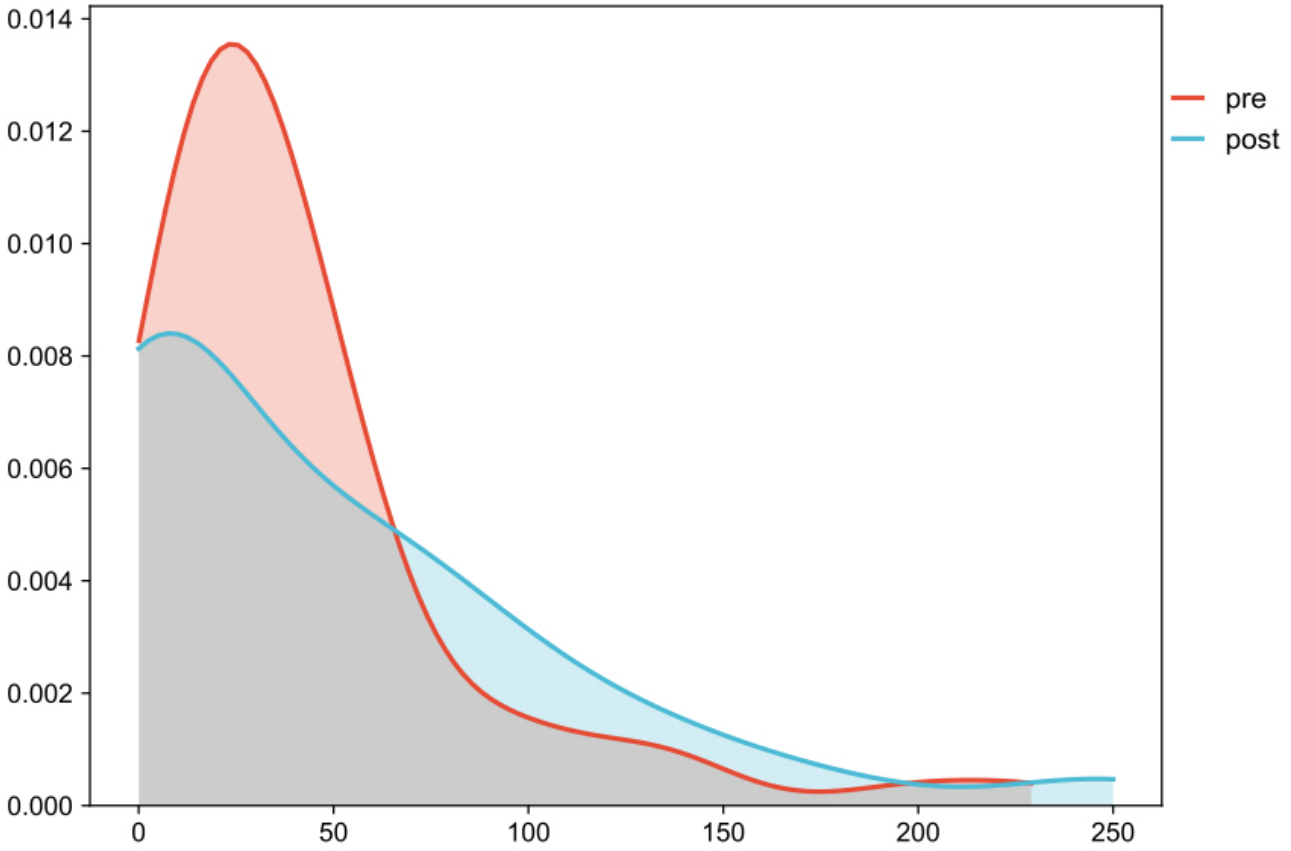
"Directed me to mdcalc which I already use as an app. No impact on above criteria (workflow, number of patients seen and time for each patient). Would be more useful to have easy access to trust guidelines such as electrolyte imbalances and chest pain etc." [Participant 10].

Comparison of Groups

Comparison of the 2 groups revealed that searches in the hospital intranet group had shorter search times than in the AI-supported search engine group (41.4 vs 88.1 seconds, respectively; Welch t=4.06; 95% CI 20.30-59.09; P<0.05). The differences in the observed task duration are presented as Kernel density plots in Figure 1. There were no statistically significant differences between the groups with regard to user satisfaction (Welch t=0.75; 95% CI −1.2607 to 2.6607; P>.05) and likelihood of a search result solving the query (χ^2=2.2; 95% CI −3.24% to 27.45%; P>.05). Participants of the AI-supported search engine group were likely to recommend Medwise.ai to a friend or colleague as reflected in their NPS of 20 (40% promoters and 20% detractors).

XSL•FO
RenderX

Figure 1. Kernel plot of the duration of searches pre (hospital intranet) and post (artificial intelligence–supported search engine) phase.



Overall, the group using Medwise.ai spent longer searching for information, and their searches were not more likely to be successful compared to the usual information retrieval methods.

Discussion

What We Have Shown

To the best of our knowledge, this is the first study to evaluate the impact of an AI-based information retrieval system on clinical workflow in acute hospital care. The design and implementation of a clinical information search engine integrating intranet- and internet-based information resources using machine learning was feasible and well-received by participants.

Searches in this small sample were not faster, and there was no higher likelihood of a successful search result. Despite this, doctors believe Medwise.ai has the potential to improve efficiency and workflow in the future and hence highly recommended the digital tool to their colleagues.

Strength and Weaknesses

One of the strengths of our study is that it uses objective measures of efficiency including search times and search outcomes rather than relying on self-reported questionnaire data from users. This mitigates recall bias. Additionally, the free-text feedback from users provided useful insights in that it shows that despite Medwise.ai's limitations, the users were optimistic about its potential in the future and expressed views in favor of its implementation in clinical workflow, as evident in quotes in Textbox 1.

It is worth noting that no specific training in evidence searching was provided to the participants beyond what they would have received as part of their usual medical education. This lack of additional training could be seen as both a strength and a limitation of our study. On one hand, it reflects the real-world scenario where clinicians often rely on their existing search skills when using new tools. This approach allows for a more authentic assessment of the AI-supported search engine's usability and effectiveness in a clinical setting. On the other hand, it raises questions about whether targeted training in evidence searching or in using the new platform might have improved search efficiency and outcomes. The ideal search engine should require minimal introduction or training, but understanding its functionality could potentially enhance search yields. Future studies might consider comparing the performance of users with and without specific search training to better understand the impact of such preparation on the effectiveness of AI-supported search tools in clinical practice.

The initially unfamiliar user interface could have been a significant contributor to the longer search time. Additionally, the introduction of the AI-supported search engine occurred toward the end of the rotation when doctors were most confident about the processes and management of patients. Participants had spent an average of 8 months in their current specialty, throughout which they would have regularly used traditional information retrieval sources.

Our observations and user feedback highlighted several areas for potential improvement in future iterations of the AI-supported search engine. First, we recognized that the platform was initially developed for desktop use, but mobile

XSL•FO
RenderX

access emerged as the preferred method in the acute care setting. In response, the platform's developers are working on a mobile-first version, including dedicated iOS and Android apps, as well as access via a chatbot interface on WhatsApp. This adaptation aims to improve user experience and reduce search times. Additionally, we noted that the more experienced intervention group might have been less accustomed to new digital formats, potentially impacting their interaction with the novel interface. To address this, future iterations could incorporate more intuitive design elements and provide brief, targeted training to familiarize users with the system's capabilities.

In our pilot, Medwise.ai was neither more nor less efficient, successful, or satisfying to users than traditional information retrieval sources. Searches using Medwise.ai were limited by the content available: if content did not exist within the local repository, then the AI will not be able to provide answers to queries. Finally, the local documents that were shared with Medwise.ai were in PDF or Microsoft Word format. Further development to transform these documents into more machine-readable formats with seamless integration with the search could further improve the performance of the Medwise.ai search engine. Like all AI-based tools, Medwise.ai's performance is contingent upon the quality and quantity of data used to train the tool. As the bank of searches builds up over time, Medwise.ai would be able to offer easier and faster access to reliable local guidelines.

The effectiveness of the search engine is dependent on the underlying AI models' performance. Although not a focus of this study, it is an area for future attention, particularly with the growing use of large language models.

What Others Have Shown

The NHS Long Term Workforce Plan [7] emphasizes the need to leverage the power of AI to improve efficiency and workflow. This case study is an example of a technological innovation to improve work efficiency with the end goal of improving patient care. It contributes to the literature focusing on evaluating the potential of AI-based technologies to improve service provision in the NHS [8-10].

Clinical Implications

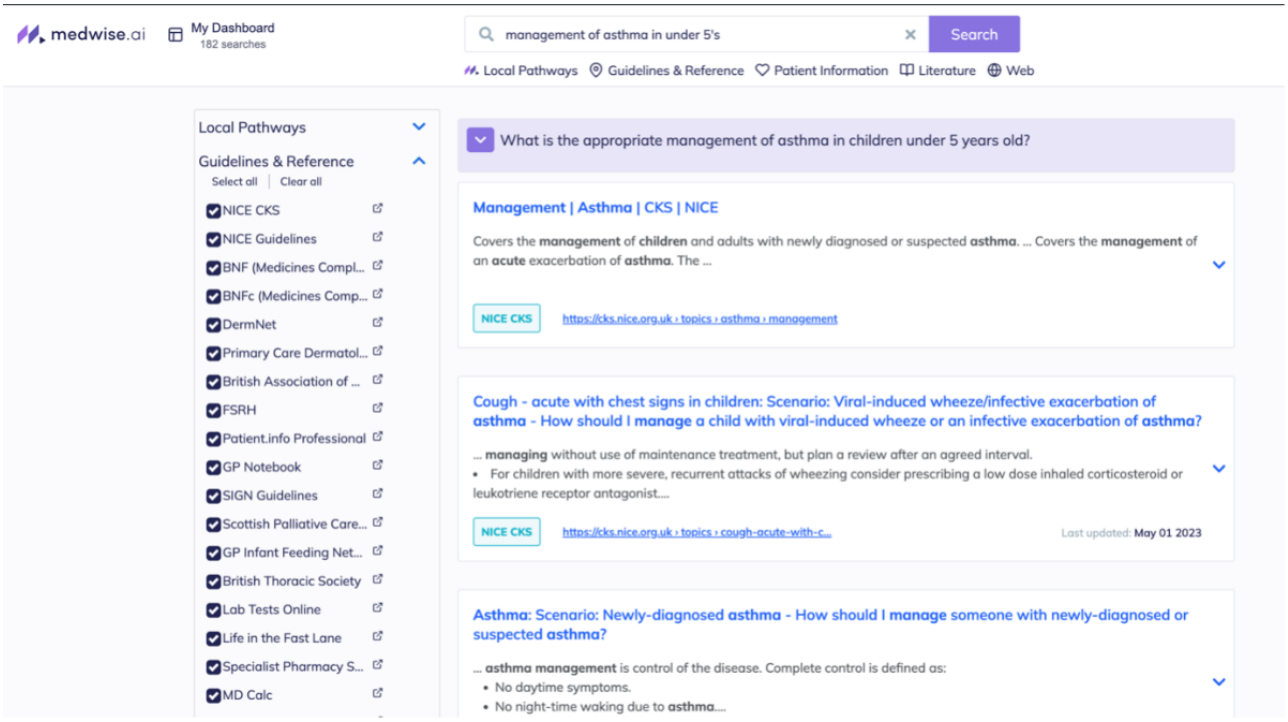
The search engine was implemented successfully during this trial. With the clear hierarchy of information sources, Medwise.ai might represent a source that has higher utility for clinicians at the bedside than public search engines such as Google or Safari. While it has been hypothesized that AI could aid to bring clinical guidelines to routine consultations [11], clinical apps remain scarce [12]. Moreover, evidence suggesting a significant reliance on colleagues and internet websites for information retrieval during shifts makes a strong case for the need for easy-to-access evidence-based local guidelines [13].

Research Recommendations

Given the small sample size and the unknown baseline distribution of the variables in question, statistical comparison is expected to be of limited value. A larger study with randomization of doctors at the beginning of rotational posts might provide a more robust assessment. The authors believe that less experienced clinicians (foundation year 1 trainees or advanced nurse practitioners in their first year of practice) would rely heavily on digital information sources and would be equally unfamiliar with other information retrieval methods and might hence be the most suitable group to evaluate effects on decision-making, safety, and possible clinical impact. More recent developments in AI such as large language models have shown the ability to encode clinical knowledge and answer medical questions [14]. Integrating these into the Medwise.ai engine might improve accuracy and speed. In the proof-of-concept platform, a full-text, query-augmented, retrieval engine configured and tuned for medical searches was used (Figure 2). No large language models were used. Further research is needed to evaluate how the accuracy and performance of the proof-of-concept search engine could be improved by incorporating large language models into the engine.

XSL•FO
RenderX

Figure 2. An image of the platform's search reply.



Conclusions

Implementation of Medwise.ai was feasible. In this small pilot study of an AI-supported search engine, we did not demonstrate increased workflow efficiency, search success, or satisfaction when compared with searches using the traditional hospital intranet. However, doctors in training believed the solution has the potential to do so in the future and hence recommend its implementation in clinical workflow. With added source content and integration of the Medwise.ai search with the next generation of AI large language models, it is likely that more benefits will be realized.

Acknowledgments

Marie Latham-Jones and Lynne Grundy were instrumental in establishing the connection between Betsi Cadwaladr University Health Board and Medwise.ai. Laura Longshaw was the driving force behind the progress of this grant. Without her encyclopedic knowledge of research frameworks, forensic attention to detail, unmeasurable patience, and immense commitment, the authors would not have succeeded. Without Gareth Jones' contribution to the recruitment of participants in the artificial intelligence–supported search engine group, the study would not have been completed on time. A big thank you! The study was funded through an Innovate UK grant administered through Medwise with Betsi Cadwaladr University Health Board as a coapplicant (Innovate UK grant application 10023176).

Conflicts of Interest

HLKT, LRUP, and MP are employees of Medwise AI Limited. The authors have no other conflict of interest.

Multimedia Appendix 1
Questions asked in the focus group.
[DOCX File, 16 KB - humanfactors_v12i1e52358_app1.docx]

Multimedia Appendix 2
Clustered column chart showing the distribution of search categories in the hospital intranet period (control) and artificial intelligence–supported search engine period (intervention). "General calc" comprised of calculating medication doses and converting units of measurements. "Interpreting invx" meant finding information about the significance of specific findings on investigations including plain films, electrocardiographs, and blood results. "Medication" meant medication-related information, that is, dose, route of administration, and interactions. "Scoring systems" meant users calculating prognostic scores including Well score. "Staff contact" meant looking up bleep numbers for other members of the health care workforce. "Term" meant looking up specific definitions. "Treatment protocol" meant treatment pathways for conditions and presenting complaints.
[JPEG File, 15 KB - humanfactors_v12i1e52358_app2.jpeg]

XSL•FO
RenderX

Multimedia Appendix 3
Word cloud of search terms.
[JPEG File, 132 KB - humanfactors_v12i1e52358_app3.jpeg]

References

1. Del Fiol G, Workman TE, Gorman PN. Clinical questions raised by clinicians at the point of care: a systematic review. JAMA Intern Med 2014 May;174(5):710-718. [doi: 10.1001/jamainternmed.2014.368] [Medline: 24663331]
2. Penders R, Mallet M. Survive On Call—a QI project to improve access to hospital clinical guidelines. BMJ Qual Improv Rep 2016;5(1):u210274.w4363. [doi: 10.1136/bmjquality.u210274.w4363] [Medline: 27239310]
3. Lopetegui M, Yen PY, Lai A, Jeffries J, Embi P, Payne P. Time motion studies in healthcare: what are we talking about? J Biomed Inform 2014 Jun;49:292-299. [doi: 10.1016/j.jbi.2014.02.017] [Medline: 24607863]
4. Sabin J, Khan W, Subbe CP, et al. "The time it takes…" How doctors spend their time admitting a patient during the acute medical take. Clin Med (Lond) 2016 Aug;16(4):320-324. [doi: 10.7861/clinmedicine.16-4-320] [Medline: 27481373]
5. Nagendran M, Chen Y. Real-time confidence of clinical decision making: a systematic review. Futur Healthc J 2017;17(2):82-90. [doi: 10.7861/futurehosp.6-2s-s82]
6. What is net promoter? NICE. URL: https://www.netpromoter.com/know/ [accessed 2022-06-03]
7. NHS Long Term Workforce Plan. NHS England. URL: https://www.england.nhs.uk/publication/nhs-long-term-workforce-plan/ [accessed 2025-01-03]
8. Akkus Z, Cai J, Boonrod A, et al. A survey of deep-learning applications in ultrasound: artificial intelligence-powered ultrasound for improving clinical workflow. J Am Coll Radiol 2019 Sep;16(9 Pt B):1318-1328. [doi: 10.1016/j.jacr.2019.06.004] [Medline: 31492410]
9. Bizzo BC, Almeida RR, Michalski MH, Alkasab TK. Artificial intelligence and clinical decision support for radiologists and referring providers. J Am Coll Radiol 2019 Sep;16(9 Pt B):1351-1356. [doi: 10.1016/j.jacr.2019.06.010] [Medline: 31492414]
10. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. N Med 2019 Jan;25(1):44-56. [doi: 10.1038/s41591-018-0300-7] [Medline: 30617339]
11. Manning CL. Artificial intelligence could bring relevant guidelines into every consultation. BMJ 2019 Aug 7;366:l4788. [doi: 10.1136/bmj.l4788] [Medline: 31391184]
12. Hamed E, Sharif A, Eid A, Alfehaidi A, Alberry M. Advancing artificial intelligence for clinical knowledge retrieval: a case study using ChatGPT-4 and link retrieval plug-in to analyze diabetic ketoacidosis guidelines. Cureus 2023 Jul;15(7):e41916. [doi: 10.7759/cureus.41916] [Medline: 37457604]
13. Daei A, Soleymani MR, Ashrafi-Rizi H, Zargham-Boroujeni A, Kelishadi R. Clinical information seeking behavior of physicians: a systematic review. Int J Med Inform 2020 Jul;139:104144. [doi: 10.1016/j.ijmedinf.2020.104144] [Medline: 32334400]
14. Singhal K, Azizi S, Tu T, et al. Large language models encode clinical knowledge. Nat New Biol 2023 Aug;620(7972):172-180. [doi: 10.1038/s41586-023-06291-2] [Medline: 37438534]

Abbreviations

AI: artificial intelligence
ED: emergency department
NHS: National Health Service
NPS: Net Promoter Score

Edited by A Kushniruk; submitted 03.09.23; peer-reviewed by C Lokker, F Alam; revised version received 14.09.24; accepted 01.12.24; published 25.02.25.

Please cite as:
Ejaz H, Tsui HLK, Patel M, Ulloa Paredes LR, Knights E, Aftab SB, Subbe CP
Comparison of a Novel Machine Learning–Based Clinical Query Platform With Traditional Guideline Searches for Hospital Emergencies: Prospective Pilot Study of User Experience and Time Efficiency
JMIR Hum Factors 2025;12:e52358
URL: https://humanfactors.jmir.org/2025/1/e52358
doi:10.2196/52358

© Hamza Ejaz, Hon Lung Keith Tsui, Mehul Patel, Luis Rafael Ulloa Paredes, Ellen Knights, Shah Bakht Aftab, Christian Peter Subbe. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 25.2.2025. This is an open-access article

XSL•FO
RenderX

distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Usability and User Experience Testing of a Co-Designed Electronic Patient-Reported Outcomes App ("MyPal for Adults") for Palliative Cancer Care: Mixed Methods Study

Panos Bonotis[1,2], DipEng; Pantelis Angelidis[2], PhD; Pantelis Natsiavas[1], MSc, PhD

[1]Institute of Applied Biosciences, Centre for Research and Technology Hellas, 6th km Charilaou-Thermi Road, Thessaloniki, Greece

[2]Department of Electrical & Computer Engineering, University of Western Macedonia, Kozani, Greece

Corresponding Author:
Panos Bonotis, DipEng
Institute of Applied Biosciences, Centre for Research and Technology Hellas, 6th km Charilaou-Thermi Road, Thessaloniki, Greece

Abstract

Background: Mobile health technologies have the potential to enhance the management, communication, and overall quality of life of patients with cancer. The MyPal project, using a participatory design approach, aims to provide palliative care support for patients with cancer through an electronic patient-reported outcomes eHealth platform.

Objective: This study aims to evaluate the usability and user experience of "MyPal for adults," a co-designed palliative care mobile app designed to support adults with cancer.

Methods: Representative users participated in a 4-step usability study using a "think-aloud" protocol, complemented with feature satisfaction, difficulty perceived, and design impression surveys along with a short, structured interview. Participants were also asked to provide quantitative feedback via the postuse System Usability Scale, User Experience Questionnaire, and Post-Study System Usability Questionnaire. The data were analyzed along the lines of the International Organization for Standardization (ISO) 9241‐210 framework.

Results: All participants found the intervention content useful, and they reported satisfactory usability, with a mean Post-Study System Usability Questionnaire score of 2.458 (SD 1.08) and a System Usability Scale score of 68.9. All aspects of the User Experience Questionnaire (attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty) surpassed usability quality benchmarks. The qualitative analysis identified 43 usability issues, primarily related to effectiveness and efficiency as defined in ISO 9241‐210.

Conclusions: In this study, we conducted a usability evaluation of the "MyPal for adults" app, a digital tool designed to enhance the palliative care experience. This approach identifies real-world usability issues, enabling iterative improvements in the eHealth platform's design. Our findings reveal a user-friendly interface and positive patient experiences. This study emphasizes the need to enhance mobile health platform usability, offering insights to improve digital palliative care.

Trial Registration: ClinicalTrials.gov NCT04370457; https://www.clinicaltrials.gov/study/NCT04370457

(JMIR Hum Factors 2025;12:e57342) doi:10.2196/57342

KEYWORDS

mHealth; app; digital intervention; think-aloud test; palliative care; cancer care; patient-reported outcomes; usability evaluation; mobile health; mobile phone

Introduction

Background

Hematological malignancies (HMs) account for about 6.5% of global cancer cases, primarily affecting older individuals around 70 years old [1]. Patients with HMs often experience distressing symptoms, which can lead to a lower quality of life (QoL) [2]. Consequently, palliative care has emerged as a vital health care practice to enhance the QoL of individuals facing serious, life-limiting illnesses, regardless of age or disease stage [3].

Effective palliative care relies on active patient-health care team communication, where electronic patient-reported outcome (ePRO) systems offer substantial promise [4,5]. These systems enable patients to report their health status, including personal evaluations, symptoms, psycho-emotional state, and adverse drug reactions via defined questionnaires. However, despite their potential to enhance patient care, the widespread integration of ePRO systems in routine health care and research remains limited [6].

XSL•FO
RenderX

Although smart device adoption in health care began increasing even before COVID-19 [7], enabling more direct dissemination of health information, complex interfaces often lead to app abandonment, especially among older users. The Google and Apple app stores offer around one million health-related apps, with approximately 300,000 in the mobile health (mHealth) category [8]. Despite this, mHealth app downloads are declining [9,10], raising concerns about their usability and acceptance, influenced by factors like privacy and usability [11].

To address these challenges, co-design approaches have gained interest [12]. Co-design involves collaborative efforts among researchers, designers, and end users who actively contribute to knowledge development, idea generation, and concept refinement [13]. This user-centered approach fosters a deeper understanding of user perspectives [14] and can optimize the development and usability of palliative care ePRO apps.

MyPal Project

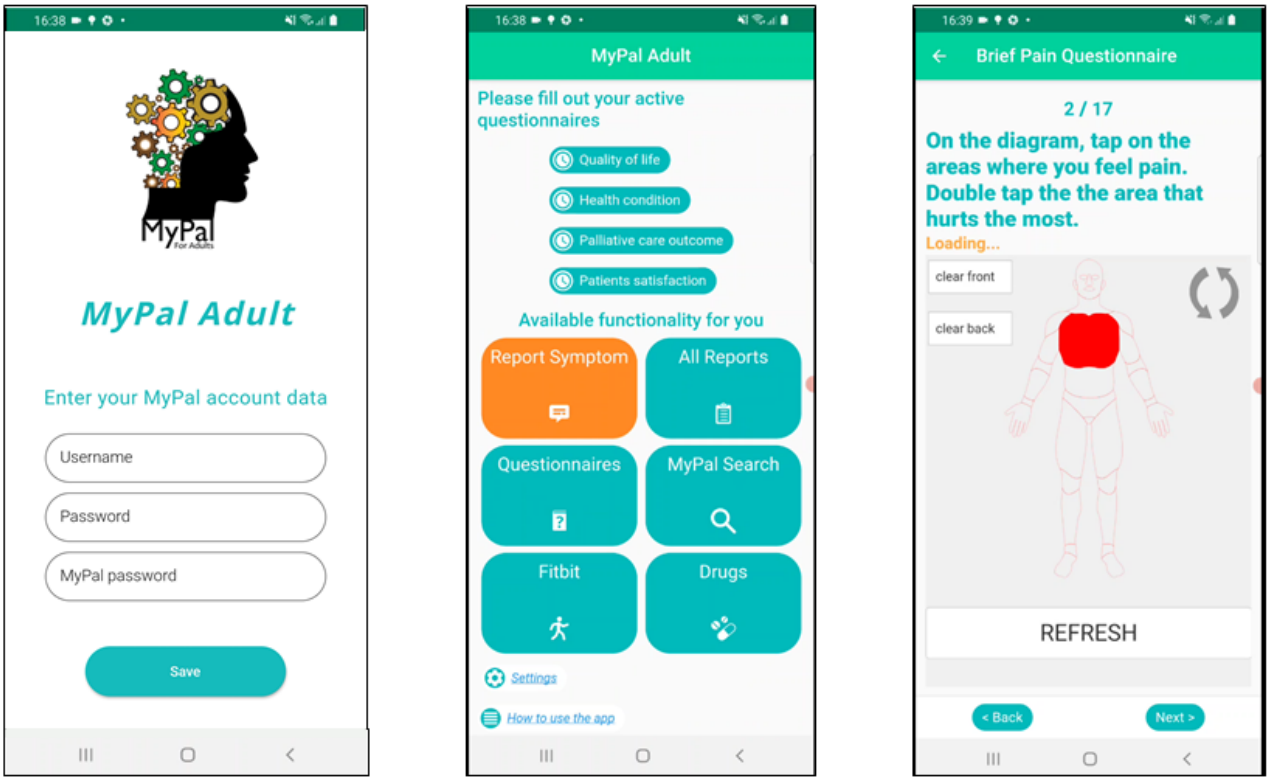
MyPal [15], a collaborative H2020 research project funded by the European Commission, uses eHealth technologies to support adults and children with cancer through advanced patient-reported outcome systems. Its primary aim is to develop and clinically assess 2 new ePRO-based palliative care interventions for adults and children, along with their carers, and health care professionals (HCPs) to enhance their QoL [16], improve effective communication, and management of the disease. MyPal uses a user-centered, iterative methodology involving end users from the project's inception to design various platform components. The design of the MyPal platform and its features were aligned with the needs of patients receiving palliative care for hematologic malignancies, as identified through research conducted during the early stages of the MyPal project. These include symptom management tools (eg, pain and fatigue tracking and spontaneous symptom reporting), psychosocial support features such as customizable motivational messages and emotional well-being assessments via ePROs, and tools for personalized care planning allowing patients to set individual goals and preferences. In addition, the app facilitates advanced care planning discussions by providing educational resources that address both treatment and QoL considerations.

The MyPal platform for adult patients consists of a mobile app for the patient, a web app for the HCP, and tools for the administrator. MyPal also supports young patients (children and adolescents) and their careers in the CHILD study [17] with a serious game for patients, a mobile app for the carer, and a web app for HCPs. These interventions and their associated technologies are to be tested in a randomized clinical trial conducted across 5 countries (Greece, Germany, Italy, Sweden, and the Czech Republic).

This usability study focuses on the "MyPal for adults" app [18] which has an end goal to facilitate (1) better self-management of the HM disease symptoms, and (2) timely assessment of the reported symptoms or QoL deteriorations by treating HCPs. Technical components are demonstrated in an openly available video [19]. The MyPal adult platform includes a mobile app (Figure 1) that supports periodic and spontaneous reporting of physical and psycho-emotional symptoms, using standardized ePROs for periodic reporting and custom electronic forms for spontaneous reporting. These reports, along with the standardized ePRO data, are directly accessible to HCPs via their dedicated platform. HCPs can view reports, track the status of questionnaires for each patient, and be alerted to any symptoms that meet predefined thresholds, ensuring timely clinical intervention. The app also offers features such as medication management, where the patient can set reminders for dosage intake, an HCP-tailored disease, and health information search engine, personalized motivational messages that are sent to patients based on their needs collected via screener questionnaires at specific time points, and tracking of lifestyle parameters like physical activity and sleep quality. Additionally, "MyPal for adults" provides secondary supportive features that are optional for use, including an app tutorial, tech support chat, and contact information for local MyPal liaisons. Overall, the app is provided in 5 languages: Greek, Italian, Swedish, Czech, and English (only for demonstration).

XSL•FO
RenderX

Figure 1. The MyPal mobile app for adult patients: login page, home screen, and example question from the electronic patient-reported outcome module. The interface tested was in Greek; the version shown here is in English to facilitate clear communication of the app's design and functionality.



Objective

This study aims to evaluate the usability and user experience (UX) of a co-designed, midfidelity version of the "MyPal for adults" mobile app using a 4-step framework, focusing on assessing content, functionality, and barriers experienced by end users.

In that context, the research questions addressed are: Is "MyPal for adults" usable, comprehensible, satisfactory, and acceptable to end users? If not, what are the usability and satisfaction issues in terms of number, type, and severity as evaluated by patients?

Methods

Explorative Usability and UX Assessment

We chose an exploratory approach to assess usability and UX through a 4-step process, including a coaching "think-aloud" protocol [20], structured interviews, and follow-up usability and UX measuring questionnaires: the Post-Study Systems Usability Questionnaire (PSSUQ) [21], the System Usability Scale (SUS) [22] and the User Experience Questionnaire (UEQ) [23] known for their established effectiveness in the literature. This process simulated a patient's journey in the MyPal Adult intervention study, as described in its protocol [2] (Figure 2). Given the COVID-19 pandemic restrictions and the Greek National Health System's advisories against hospital visits, we conducted all assessments remotely using digital tools.

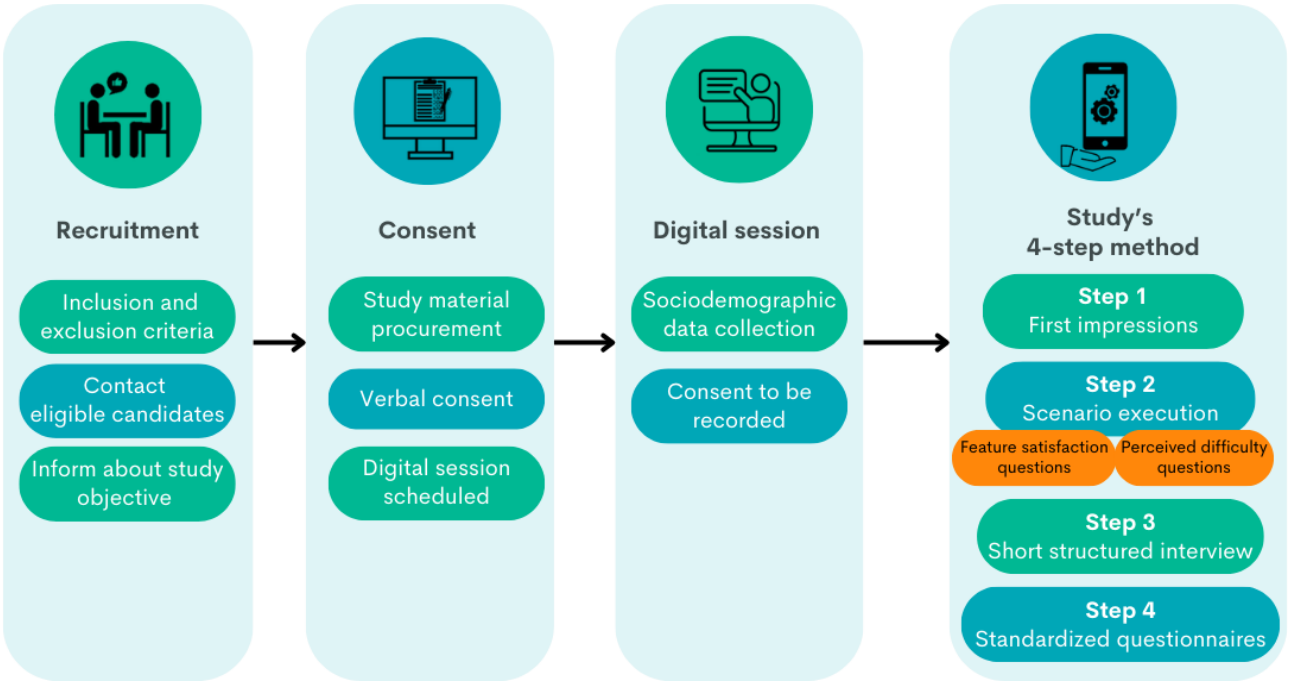
Usability, as per the ISO (International Organization for Standardization) usability framework (ISO 9241 - 210), is "the extent to which a product can be used by designated users to achieve specific goals with efficiency, effectiveness, and satisfaction, in a specific context use" [24]. The criteria of this framework are effectiveness, efficiency, satisfaction, and context of use. Typically, the concept of usability also covers comprehensibility, learnability, functionality (operability), attractiveness, appropriateness recognizability, accessibility, user error protection, and interface aesthetics [25].

As described in ISO 9241 - 210, effectiveness is the accuracy and completeness with which users achieve specified goals, such as symptom reporting. Efficiency refers to the resources used in relation to the results achieved, like the time needed to report a symptom. Satisfaction is defined as the "extent to which the user's physical, cognitive, and emotional responses that result from the use of a system, product, or service meet the user's needs and expectations" [24].

UX is defined as "user's perceptions and responses that result from the use or anticipated use of a system, product, or service." [26]. Good UX not only contributes to higher work motivation and performance but can also affect the well-being of users [27]. Achieving acceptable UX or usability requires systematic evaluation and iterative design adjustments as the app matures [28].

XSL•FO
RenderX

Figure 2. Participant journey in the usability study.



Participant Recruitment

Participants for this study were recruited through nonprobability methods, a mix of purposive sampling to target relevant participants, and convenience sampling to address recruitment challenges during the COVID-19 pandemic. They were recruited based on predefined inclusion criteria in the project's randomized clinical trial protocol [2]. Indicatively, adults (18 years or older) diagnosed with hematologic cancer, who were scheduled to receive or had previously been exposed to any line of treatment, and were users of an internet-connected device (smartphone or tablet). Exclusion criteria included patients requiring immediate referral for specialized palliative care or with a life expectancy of less than 3 months. To this end, given our institution's close collaboration with ELLOK (Greek Federation of Cancer Patients), members were contacted via telephone and email, facilitated by MyPal's clinical liaisons, ensuring the participation of members who are patients with hematologic cancer. Finally, since ELLOK had previously procured volunteers to consult on and test earlier prototypes, we added an exclusion criterion stating that no participant should have had any prior experience interfacing with the midfidelity app in focus. This exclusion criterion was added to ensure that all participants would be "naïve" users, that is, users who have no training or experience with the app, as would be the case for a patient enrolling in the MyPal study.

To address the recruitment difficulties posed by the pandemic and in consultation with our clinical liaisons, the inclusion criteria were expanded to include individuals with other chronic diseases sharing care needs similar to hematologic cancer (eg, autoimmune skin conditions and chronic pulmonary diseases) and recent hospital visits in the last 6 months were also planned to be included as a countermeasure.

Data Collection

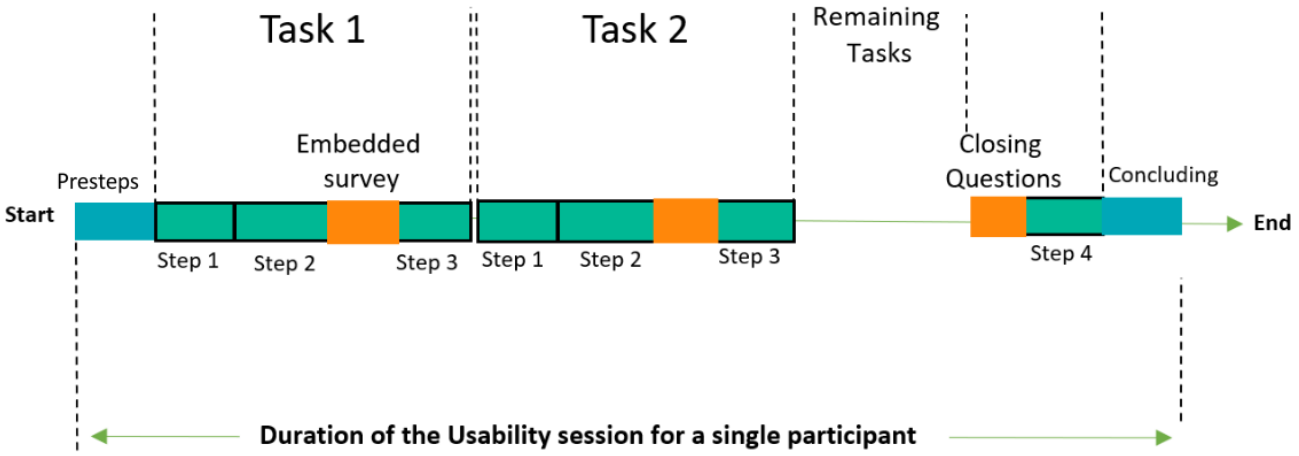
Global Process

In alignment with the MyPal Adult intervention study protocol, interested and eligible candidates were informed about this usability study objective and the MyPal project goal, its technologies, and their functionalities, and provided with a detailed informational sheet and contact details for any inquiries. The consent process was also conducted remotely using digital tools due to the COVID-19 pandemic restrictions. After confirming their participation in this usability study, they were emailed with every material a patient joining the MyPal Adult intervention study would receive, specifically the link to download the app, a user manual, and a quick-start guide, offering a summary of the project's goal and the platform capabilities. They were also asked to confirm the appointment for the usability session. Verbal consent was obtained once more by the moderator of the usability study prior to the session initiation.

Each participant completed a single 55- to 70-minute teleconference session, following a structured task-based use scenario (Figure 3). The moderator used a PC with a webcam and the Zoom (Zoom Communications, Inc) platform to record facial expressions and facilitate the sharing of the mobile device screen. Participants were asked to install the Zoom app on their smartphones and laptops prior to the session to enable screen sharing from their smartphones and record facial and vocal interactions via their laptops. For the app testing, participants used their own Android devices; however, the research team had a backup Android device available if needed.

XSL•FO
RenderX

Figure 3. Step sequence in a single think-aloud session.



Preparatory Steps

At the beginning of each session, participants provided sociodemographic information, including age and sex. They also answered 2 multiple-choice questions about their smartphone skills, based on their comfort level and experience with using smartphone apps, including frequency of use, ability to navigate apps, and familiarity with mobile technology in general, and 3 questions on smartphone app use (Multimedia Appendix 1). Following the introduction, the usability study's context, task completion logic (described in the task "scenario" in Multimedia Appendix 2), and the "think-aloud" process were explained in detail. An example task was also presented to familiarize participants with the "think-aloud" method.

Step 1: First Impressions

During the task "scenario" execution, participants navigated to various functionalities' screens. The moderator inquired about the screens' general eye-catching attributes and the users' initial impressions. For each screen, 2 questions were posed: (question 1.1) "What is your initial impression of this screen?" and (question 1.2) "How do you rate the design and information organization?" To gain more detailed insights, a third question was added for each screen (question 1.3): "What actions do you think you can perform on this screen?"

Step 2: Task "Scenario" With an Embedded Survey for Satisfaction and Perceived Difficulty

The second step involved executing tasks within the "scenario," where participants were encouraged to think aloud to identify usability issues, aligning with the "think-aloud" protocol. The "scenario" included in total of 16 tasks (Multimedia Appendix 2). Some of the tasks included multiple-choice questions in the 5-Likert and 7-Likert scale [29], to gather quantitative data on perceived difficulty and satisfaction with specific functionalities. While the satisfaction questions were inquired for every feature in the app, (mainly focusing on modules and elements regarded as important early on in the design ideations), the perceived difficulty questions were more targeted. Difficulty-perceived questions were only inquired during tasks that had performed poorly based on the results of prior assessment procedures (validation, rehearsal testing) within the MyPal project. In accordance with the "think-aloud" protocol, moderators worked with participants during task execution, meaning they would intervene to prompt participants to verbalize their thoughts, or "think-aloud." When a participant struggled with a task, the moderator noted it down and provided guidance. If participants repeatedly encountered issues with a specific action leading to task failure, moderators inquired about the problem's nature to capture participant perceptions and experiences.

Participants' verbalizations about their preferences or dislikes regarding the PROs content were not considered, as it was not the study's purpose. Nonetheless, participants were encouraged to voice their thoughts while going through the content to uncover potential user interface (UI) inconsistencies.

Step 3: Short Structured Interview

After completing the final task in the "scenario," participants were given the opportunity to add anything to their responses regarding feature satisfaction from step 2 and to address any follow-up questions related to the usability evaluation and overall UX. Additionally, 3 extra questions were posed as part of the structured interview following the last task of the "scenario:" (question 3.1) "How do you perceive the secondary complementary supportive features? Would you use them?" (question 3.2) "What would be your first action with this app?" and (question 3.3) "Do you feel the app expects something from you?"

Step 4: Standardized Questionnaires

Before concluding each session, quantitative feedback was gathered from participants using a set of well-established and validated questionnaires. These included the SUS, a 10-item questionnaire rated on a 5-point Likert scale ranging from Strongly Disagree to Strongly Agree, which assesses the overall usability of a system [22]. The UEQ measures 6 distinct dimensions of UX—attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty—using a 7-point semantic differential scale [23]. Finally, the PSSUQ, which is rated on a 7-point scale from Strongly Disagree to Strongly Agree, evaluates user satisfaction with system usability, covering areas such as task ease, efficiency, and interface quality [21]. Participants were then given the option to submit either only the PSSUQ or all 3 questionnaires.

XSL•FO
RenderX

Data Analysis

Steps 1 and 3 ("first impression" and "short, structured interviews") underwent inductive thematic analysis [30], involving transcription and translation of all participant responses into English, along with the calculation of standardized questionnaire scores. In step 2 (task "scenarios"), a deductive analysis aligned with the ISO framework categories was performed [30].

Ethical Considerations

The study received ethical approval from the Bioethical Committee of the Center for Research and Technology Hellas (ETH.COM-61). The study involved human participants and adhered to institutional and ethical guidelines for research involving human participants. Informed consent was obtained from all participants after they were provided with clear information regarding the purpose of the study, procedures involved, data handling, and their right to withdraw at any time without penalty. No waiver of consent was applied. Privacy and confidentiality were maintained throughout the study. All data collected were fully anonymized before analysis, and no personally identifiable information was stored or used in reporting. Any potential identifiers in screenshots or visual material were removed or blurred. Participants did not receive any form of remuneration or compensation for their participation in the study.

Results

Participant Demographics

A total of 9 participants took part in the usability and UX testing. They were recruited from 3 regions in Greece—Thessaloniki, Karditsa, and Heraklion—for feasibility and proximity reasons. The study group included 4 individuals with chronic lymphocytic leukemia (CLL), 2 senior citizens, and 3 adult representative users afflicted by a range of chronic ailments. The gender distribution was balanced, with 4 (44%) women and 5 (56%) men. Participants' ages ranged from 27 to 68 years old, with a mean age of 49.1 (SD 13.6) years, and a median of 17 (IQR 42-59) years. Regarding smartphone skills, participants self-reported varying levels of proficiency: 1 (11%) participant reported poor skills, 2 (22%) participants reported low skills, 2 (22%) participants had medium skills, 1 (11%) participant had high skills, and 3 (33%) participants considered their skills to be perfect. In terms of prior experience with similar apps, 2 (22%) participants had used health-related apps before, while the majority of participants (7/9, 78%) had no previous experience (Table 1).

Table . Participants' sociodemographic data.

Characteristics	Value (N=9)
Gender, n (%)	
Female	4 (44)
Male	5 (56)
Age (years), mean (SD)	49.1 (13.6)
Smartphone skills, n (%)	
Poor skills	1 (11)
Low skills	2 (22)
Medium skills	2 (22)
High skills	1 (11)
Perfect skills	3 (33)
Prior experience with similar apps, n (%)	
Yes	2 (22)
No	7 (78)
Participant's diagnosis, n (%)	
CLL[a]	4 (44)
Chronic arthritis	1 (11)
COPD[b]	3 (33)
HS[c] stage 3	1 (11)

[a]CLL: chronic lymphocytic leukemia.

[b]COPD: chronic obstructive pulmonary disease.

[c]HS: hidradenitis suppurativa.

XSL•FO
RenderX

Participants' Feedback

Step 1: First Impressions

Overview

This section presents the results from open-ended questions gathered during the first impressions collection step, translated into English. Each participant is referenced by an assigned ID (eg, P1 and P2). The analysis of these initial impressions is based on themes that emerged during the inductive thematic analysis.

Question 1.1: "What Is Your Initial Impression of This Screen?"

Every functionality and screen received a positive first impression. No one expressed any major dislikes or was annoyed by something particular in the design. The only negative comment was about the app's icon:

> *I don't particularly like that the icon's portraying a human head that is black like it is burned. Also, the gears appearing like they are inside the head makes me cringe a little bit.* [P4]

Question 1.2: "How Do You Rate the Design and Information Organization?"

As question 1.2 significantly overlapped with question 1.1, participants had already developed an impression of each screen's design by the time question 1.1 was posed. All participants neutrally accepted every screen in each functionality, with minor feedback primarily related to button placement and text entry field colors.

When analyzing participant responses, 2 distinct user groups emerged: young users (25-35 years of age) and senior users (45-70 years of age). While younger participants desired a more modern app design, they found the app's current design very user-friendly.

> *It would be very good if there was an additional UI with a more modern design, like with a side menu etc but I guess this would be more helpful with older patients as well. I don't really mind it as long as it proves easy-to-use and useful. Which I guess it will ... at least it seems that way.* [P8]

The color palette of the app and the simplicity of the design were very much appreciated by the more senior users.

> *I really like the colors, they are quite soothing, and give me the "health care" vibe. Everything seems quite simple and straightforward.* [P2]

Question 1.3: "What Actions Do You Think You Can Perform on This Screen?"

The main purpose of this question was to determine if the context of use for each functionality was evident from the screen's appearance and whether the action-probing screens were distinguishable from the informing and data-viewing screens. Almost all participants correctly identified the main purpose of each screen, except for one participant (P9), who was a senior patient with limited technology experience.

> *I understand that the purpose of this screen is to help me manage my disease. I guess it will ask me somehow about my medication. Can I speak to it? I don't understand how to do this or what to do.* [P9]

Step 2: Task "Scenario" With an Embedded Survey on Satisfaction and Perceived Difficulty

We identified 43 problems in the task "scenarios," which were classified according to the categories defined as part of the ISO 9241‐210 framework. The task can be found in Multimedia Appendix 2. Most problems (10/43, 23%) occurred during questionnaire submission at task 3. Other tasks with notable problems were drug entry (5/43, 12%) and reminder setting in the drug module (8/43, 19%) at tasks 11 and 12, respectively.

Table 2 presents the problems mapped to the main categories defined by the ISO usability framework, an example for each, and the transcript or description of the observation. It also presents the task where the problems occurred and the type of the problem. Some tasks could not be completed because of programming errors or bugs. When problems were caused by programming errors, they were classified as N/A (not applicable) in each ISO framework category and the task was considered failed. In those cases, a screenshot was displayed of the broken feature to enable the continuation of the study.

About 30% (13/43) were classified as "effectiveness" and took place mainly at the "add a drug" task, where users were asked to log their medication. A total of 23% (10/43) were classified as "efficiency" and occurred mainly at the registration process during the baseline filling-out questionnaire mainly affected by the repetition of text at each step of the registration process. In the MyPal Adult intervention study, prior to the "MyPal for adults" app's regular use, a baseline data collection is taking place immediately after logging into the app for the first time, as part of the app's registration process. The registration process was a relatively new feature, that was not thoroughly tested prior to this study, and problems were expected. It is presented in a step-by-step manner guiding the user to complete all baseline questionnaires and configure app initial settings. In addition, a notable efficiency issue was related to task 16, and the time needed to locate the emergency contact number of the local liaison team in the app. A smaller portion (9/43, 21%) related to "satisfaction" and was mainly noticed in the medication reminder setting task, which was perceived as overly complex by most participants, hindering their satisfaction with it. Two problems were classified as context of use. These problems were identified by a single participant, and both could be explained by the fact that this participant was a senior user, not experienced with mobile apps. The problems coded as the context of use included the following: (1) the participant tried to use the MyPal patient-tailored disease-related information search engine to "ask a question" (expecting a chat function) to the clinicians and (2) the size of some UI elements, that is, the module buttons, text displayed. Finally, 21% (9/43) of the problems were classified as N/A and were related to technical issues, such as software bugs or errors, that were outside the scope of usability-related classifications (effectiveness, efficiency, etc). These technical issues arose from system updates occurring during the usability testing and could not be

XSL•FO
RenderX

classified according to the ISO usability framework. Most of these problems occurred mainly in the Fitbit module when attempting to view the sleep data, setting drug reminders, and accessing the patient-tailored disease-related information search engine.

In Figures 4 and 5, we present the results of the feature satisfaction and perceived difficulty questionnaires which complemented the usability task "scenario." The question formulation can be also found in Multimedia Appendix 2.

As shown in the figures, most of the functionalities of the app were rated as satisfactory and easy to use. However, the medication reminder feature stood out as overly complex for some participants, leaving a portion either unsatisfied or indifferent. Another noteworthy observation concerns the questionnaire module. While most participants found it generally satisfactory, they struggled to locate it quickly, suggesting the need for a more intuitive or descriptive heading. Regarding the ePROs and questionnaires overall (including personalization surveys and symptom reporting), some dissatisfaction and difficulty were noted, largely due to efficiency issues. These stemmed from repetitive text at each question and high question volume, which impacted the UX, and even had some participants thinking it was a technical bug with the system.

Table . Problems with the app, classified according to the International Organization for Standardization framework.

ISO[a] categories	Problems (n=43), n (%)	"Scenario" section and task number (example)	Verbalization or observation (example)	Problem detected (example)
Effectiveness	13 (30)	App registration process (task 2)	"So, I can see here the clinics communication information, so I guess I tap on next"	Setting up the notification timing choice was not visually clear and as a result the participant did not notice it and moved on.
Efficiency	10 (23)	Custom personalization surveys and ePROs[b] impression (task 3)	"Why are the same texts repeated? Wouldn't it be better to have them all grouped so the texts wouldn't have to be repeated?"	There are a lot of seemingly repeated questions in a questionnaire, where only a small phrase would change, and all the other wordings would stay the same.
Satisfaction	9 (21)	Login and initial settings (task 1)	"Is this normal, or is it a bug? In any case it should not be like that"	Password needs to be protected. When the passwords are entered, they could be hiding with asterisks.
Context of use	2 (5)	Patient tailored disease-related information search engine (task 14)	Participant is entering a question directed to his clinicians	The participant did not understand that this is a search engine.
N/A[c]	9 (21)	Fitbit Module (task 9)	"It doesn't work. I am pressing the button but I cannot see my sleep data"	There is a technical issue with a specific button, the participant cannot watch his sleep data.

[a]ISO: International Organization for Standardization.

[b]ePRO: electronic patient-reported outcome.

[c]N/A: not applicable.

XSL•FO
RenderX

Figure 4. Feature satisfaction. ePRO: electronic patient-reported outcome.

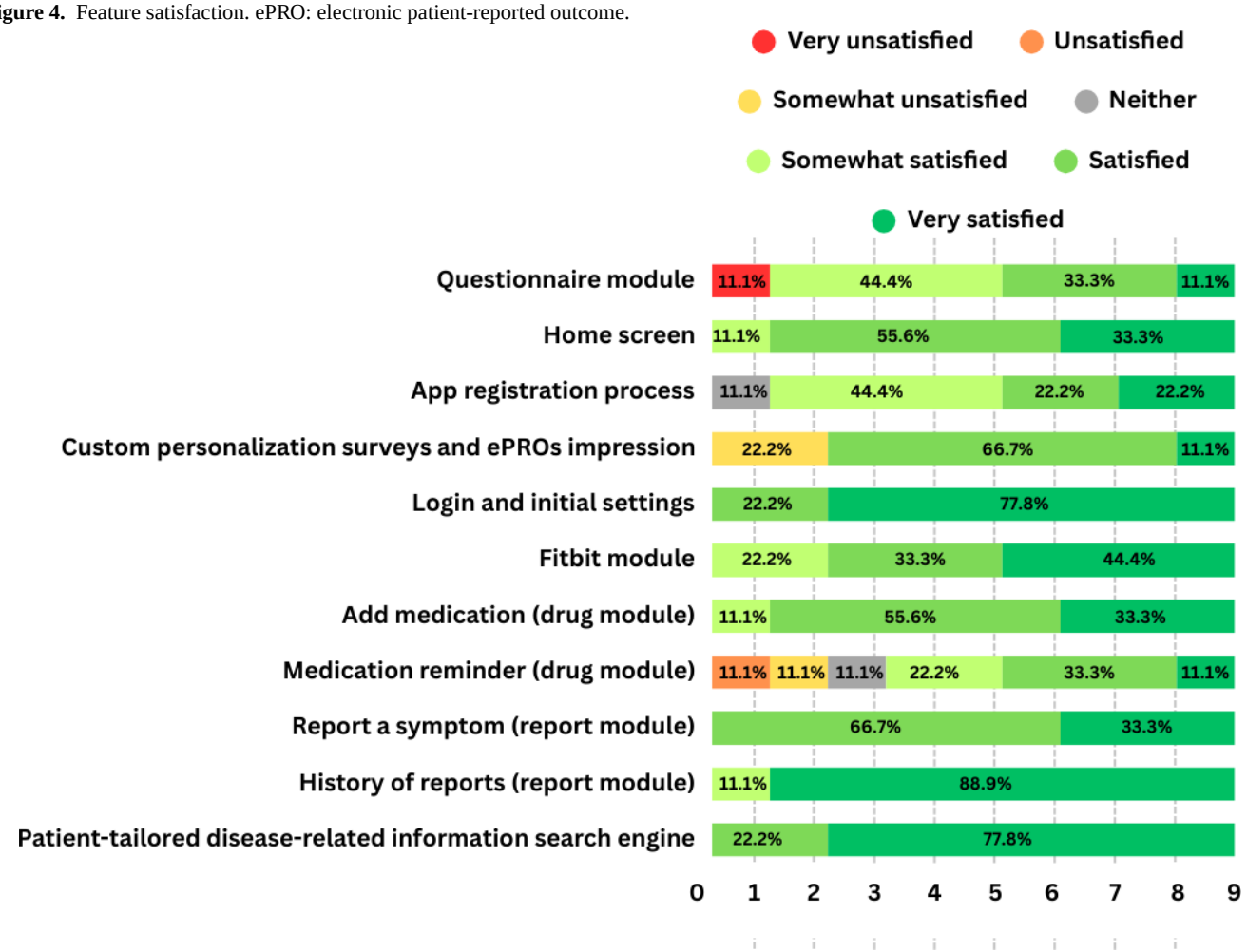
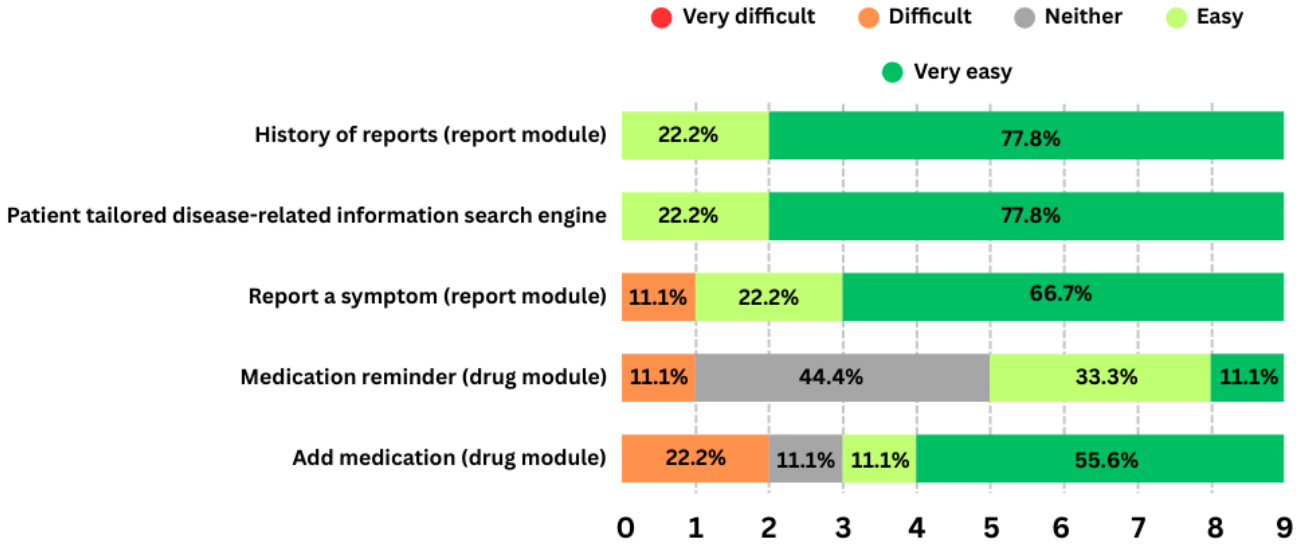


Figure 5. Task difficulty.



Step 3: Short Structured Interview

This section presents the results from the open-ended questions, translated into English, collected during the structured interview process that followed the task-based "scenario."

Question 3.1: "How Do You Perceive the Secondary Complementary Supportive Features? Would You Use Them?"

All participants expressed their willingness to use every complementary supportive feature. The tech support chat was very appreciated, as well as the app's manual.

XSL•FO
RenderX

> *I really like the fact that the manual of the app is in the home screen and so easily accessed.* [P3]
>
> *I really appreciate the chatting service for technical support. I feel like I could do this.* [P7]

Question 3.2: "What Would Be Your First Action With This App?"

This question aimed to determine if participants would prioritize pending questionnaires. However, participants' responses were inconclusive regarding this objective because all users, being new to the app, expressed a desire to explore and become familiar with it instead of immediately focusing on the active pending questionnaires.

> *I would like to familiarize myself with the app at first. I would just browse through everything and start playing with it, I guess.* [P8]
>
> *I would most likely first enter my medications and set reminders for them. Then I would probably report something retrospectively.* [P2]

Only 2 users expressed the desire to first navigate to the questionnaire module or the questionnaire pending list so they can verify they have completed everything is expected from them:

> *I guess first step is to familiarize myself with the app and make sure that I have reported everything I must since I can see in the pending list that there are a few questionnaires available for me.* [P4]
>
> *Well, I am a gadget geek so I would want to explore the whole app. But first I would navigate to the questionnaire module to make sure everything is set.* [P1]

Question 3.3: "Do You Feel the App Expects Something From You?"

This was a follow-up to question 2, ensuring that pending questionnaires were clear to all users. All participants except one did not grasp the main purpose of the pending questionnaire list on the home screen. The participant who understood its purpose admitted it was not obvious and required guidance from the moderator.

> *Yes, of course it is expecting me to answer these pending questionnaires displayed in this list. But I have to say if you had not asked me before I would not notice it right away. I would eventually but not right away.* [P5]

Step 4: Standardized Questionnaires

Following every session, the participants were asked to submit a follow-up questionnaire that included the postuse SUS, the UEQ, and the PSSUQ. All participants submitted every questionnaire. The mean scores of each questionnaire are presented in Table 3.

XSL•FO
RenderX

Table . Poststudy standardized questionnaire scores.

Standardized questionnaires	Value
PSSUQ[a,b]	
SYSUSE[c] score, mean (SD)	2.14 (0.74)
INFOQUAL[d] score, mean (SD)	2.70 (1.09)
INTERQUAL[e] score, mean (SD)	2.77 (1.48)
Overall PSSUQ score, mean (SD)	2.46 (1.08)
SUS[f,g] score	68.9
UEQ[h,i] means, mean (SD)	
Attractiveness	1.667 (0.574)
Perspicuity	1.444 (0.728)
Efficiency	1.500 (0.671)
Dependability	1.361 (0.6)
Stimulation	1.528 (0.52)
Novelty	1.306 (0.825)

[a]PSSUQ: Post-Study System Usability Questionnaire.

[b]Scale: 1=strongly agree or best score to 7=strongly disagree or worst score.

[c]SYSUSE: System Usefulness.

[d]INFOQUAL: Information Quality.

[e]INTERQUAL: Interface Quality.

[f]SUS: System Usability Scale.

[g]Scale: 0=worst usability to 100=best usability.

[h]UEQ: User Experience Questionnaire.

[i]Scale: –3=worst to +3=best.

Discussion

Principal Results and Findings

This study assesses the usability and UX of the "MyPal for adults" app in the early phases of the MyPal project. Results from the usability test yielded a marginally positive SUS score of 68.9, as per Brooke [22]. UX results were favorable, as evidenced by UEQ means and user feedback. PSSUQ also showed a positive outcome (mean 2.46, SD 1.08), with higher ratings for usefulness compared to information quality and UI. While all quantitative data indicated an acceptable score, a closer examination of participant responses revealed the primary issues with the app. User feedback emphasized the importance of tailoring specific interface elements, such as wording changes (eg, "Questionnaires" to "My Questionnaires"). Additionally, technical bugs in 2 features and repetitive text in ePRO screens reduced satisfaction and utility. Furthermore, certain functions, particularly the Medication module, were seen as overly complex, requiring refinement to enhance ease of use. Despite these challenges, users expressed a desire to engage with the app and appreciated its potential once they became familiar with its features.

Findings from step 1 confirm that all participants found the presented information comprehensible, with a clear organization of information. However, participants made numerous proposals, suggesting a potential need for a more personalized interface that accounts for age-related preferences. Older users appreciated the app's simplicity, while younger users preferred a more modern and visually engaging design. The feedback primarily focused on button placement, header wording, and contextual text. Given that these differences between age groups were consistently observed throughout the study, a dual-tiered UI could provide an effective solution. Such an approach would allow users to choose between a more visually appealing, modern interface and a minimalist, straightforward design better suited for older adults or those less familiar with technology. This tailored UI could help mitigate issues related to horizontal comprehension and ensure that every module's screen has its purpose understood by all participants.

Findings from step 2 revealed more major issues in the drug reminders and emergency contact detail modules compared to the minor issues found with the app's core functionality, the ePRO submission, which mainly related to wording and information displayed refinements. Participants found the medication reminder function confusing and overly complex and indicated they were unlikely to use it. Interestingly enough, during the "think-aloud" part of the study, most participants also dismissed this as an issue by pointing out that they already have a system and keep a medication log with reminders in place, making this feature possibly redundant. Locating emergency contact information was challenging for all participants, which in an emergency situation, could prove vital.

XSL•FO
RenderX

Participants suggested a more obvious approach, where the information can be found at the top of the home screen, which arguably could hinder usability by limiting the available screen to essential app functionalities and should be further explored. Additionally, 2 modules faced technical problems: (1) the search engine module and (2) the activity data display in the Fitbit module. The search engine module experienced unexpected bugs because the study was conducted on the development server, while the Fitbit module had difficulties connecting to the Fitbit server during the project's early development phase.

In general, based on the questionnaire results, the participant's attitude, and our overall impression and experience of the usability study, participants enjoyed using the app, and in many cases, they complemented their feedback with suggestions for UI-related details and expressed concerns about module titles or text. For instance, "Questionnaires" could be improved to "My Questionnaires" to clarify its purpose. An interesting realization was the importance of translating paper-based PROs into digital "ePROs" successfully. Unlike paper-based PROs presented on a single sheet, ePROs in the "MyPal for adults" app had individual screens for each question, leading to perceived information redundancy and concerns about falsely perceived technical issues. Participants also raised concerns about HCPs' ability to respond promptly to MyPal reports due to their busy schedules, which could affect the UX of patients, aligning with previous research on chronic pain monitoring [31].

Findings from step 3 indicate that participants were eager to explore and familiarize themselves with the app before engaging with specific tasks, such as completing pending questionnaires, an encouraging sign of positive anticipation for using the system. Additionally, many expressed a desire to explore the app, set up medication reminders, and retrospectively report symptoms. The accessibility of the app's manual was highly appreciated, especially by more senior participants, who also were positively surprised, and even relieved, when they realized the app had its own tech support functionality. Further investigation revealed that senior users anticipated complex background computations, and the availability of tech support reassured them. One of the app's core functionalities, facilitating ePRO submissions, was supported by in-app reminders and notifications, as well as a list of pending ePROs prominently displayed on the home screen to draw users' attention upon opening the app. While some participants quickly identified the pending questionnaire list, others required additional guidance, underscoring the need for clearer design cues or onboarding processes to ensure critical tasks were not missed.

Finally, the importance of iteration in the development cycle must be highlighted since all participants had problems with the presentation of the ePROs, and almost all participants but one missed the pending list of questionnaires, present on the home screen of the app. Although this could in part be attributed to the actual question which was formulated in a way to not betray its purpose, the fact that it was missed completely by almost every participant (all but one) suggests that a redesign of the app following the principles of UX is needed.

Challenges and Limitations

Study limitations include programming errors that interrupted tasks and potentially led to the loss of usability information. These errors endured due to concurrent bug fixes conducted alongside the study. The scheduling challenges related to usability testing caused by COVID-19, compounded by the Greek National Health System's discouragement of hospital visits, resulted in significant delays. Consequently, participants faced programming errors that impeded their ability to complete tasks without issue.

Beyond these limitations, the main challenges were related to the study participants. The purposive sampling method used in this study, while suitable for targeting specific user groups, does introduce inherent selection bias. Participants were chosen based on their medical conditions and alignment with clinical trial inclusion criteria. This limits the generalizability of our findings to the broader population but remains relevant to the specific user group we studied. Furthermore, the nature of usability testing and the absence of blinding may have influenced participant feedback, as users were aware they were interacting with the MyPal app. We aimed to mitigate this by simulating real-world use cases via a use scenario, where users would receive a thorough briefing during onboarding, conducted by a recruiter physician to explain the functionalities of the app. We also encouraged participants to provide candid and constructive feedback throughout the process.

Additionally, the participants had a very diverse level of smartphone skills which along with the age variation, posed a limitation. Age is a determining factor in the use of mobile app products, especially considering neurodegenerative conditions that affect older individuals may significantly impact dexterity, and thus navigation. While the study did not focus on participants with such conditions, the age range of 27 to 68 years presents its own challenges. However, the median of 17 (IQR 42-59) years suggests that the majority of participants were within a relatively concentrated age group, minimizing the impact of an extensive age gap.

Finally, participant recruitment was arduous due to the constraints imposed by the COVID-19 pandemic. Initially targeting users with hematologic malignancies and more specifically CLL or myelodysplastic syndrome, recruitment difficulties led to a shift toward enlisting senior patients coping with other chronic diseases that necessitate periodic hospital visits. While the original objective was to secure 10 - 15 participants, the study ultimately included 9 participants. Faulkner [32] has demonstrated that groups of 10 users can identify an average of 95% of usability problems, mitigating concerns regarding sample size. It is worth noting that the study's findings possess limited generalizability but can be partially applied to self-management mHealth systems using ePROs within the context of palliative cancer care, with a focus on a dedicated mobile app tailored for patients with cancer.

Comparison With Prior Work and Future Directions

Our usability evaluation of the "MyPal for adults" app demonstrates its potential to improve patient-provider communication, supporting the broader scope of palliative care

XSL•FO
RenderX

for patients managing chronic illnesses. Comparing our findings with prior research in the field of eHealth apps for palliative care, several key insights emerge. First, this study aligns with previous research that emphasizes the potential of eHealth apps to improve communication between patients and HCPs in palliative care settings [33]. Through the MyPal app, participants in this study reported improved channels for information exchange, empowering them to actively participate in their care decisions. This echoes findings from related studies [33,34], which have also highlighted how eHealth apps can bridge the communication gap in palliative care, providing a convenient and accessible platform for patients to engage with their health care providers.

While our usability evaluation yielded positive outcomes, it is important to note that, similar to much of the prior work in this field [33], this study primarily involved testing, feasibility assessments, and acceptability studies on relatively small patient populations. A comprehensive evaluation that can impact patient outcomes, including symptom management, QoL, and overall satisfaction with care, necessitates the execution of rigorous clinical trials [35]. This phase of assessment represents the final critical stage in establishing the app's effectiveness and suitability for integration into palliative care practices, as is the plan with the "MyPal for adults" app.

Future research should aim to broaden the use of the MyPal app to include a more diverse range of patients in palliative care. Additionally, the integration of the app with existing health care systems is crucial to enhance its functionality, providing HCPs with real-time patient data and enabling better care coordination. Finally, ongoing development efforts should prioritize a user-centric design approach, incorporating input and feedback from both patients and health care providers to maintain the app's intuitiveness, accessibility, and alignment with evolving user needs.

Conclusions

This usability study assessed the "MyPal for adults" app with a diverse user group, including individuals with chronic diseases, and patients with CLL. Participants generally had positive initial impressions of the app's design and organization, though some preferred a more modern interface. Although certain modules posed usability challenges, the ePRO modules were well-received. However, enhancing learnability, personalization, and addressing technical issues in some tasks is essential. This study emphasizes the importance of iterative, user-centered development to improve the app's usability and UX. Notably, older users favored simplicity, while younger users sought a contemporary design, indicating room for age-specific improvements. Given the app's role in palliative cancer care, effective ePRO implementation and thoughtful translation of paper-based PROs into digital formats are critical. Ensuring timely health care provider responses to patient reports is vital. These findings offer valuable insights for refining the MyPal app and may inform other self-management mHealth systems in palliative cancer care.

Acknowledgments

This work was partly funded by the MyPal project which has received funding from the European Union's Horizon 2020 research and innovation program under grant 825872 [15]. The authors express their sincere gratitude to ELLOK (Greek Federation of Cancer Patients) for their invaluable assistance in granting access to members for study participation. Special thanks go to Dr Christina Karamanidou for her significant role in participant recruitment and for establishing a communication pathway with ELLOK.

Authors' Contributions

PB led the study design, authored the primary text, and conducted the research. PN contributed to the study's conceptualization and provided critical manuscript revisions. PA offered manuscript edits and supported PB throughout the study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Questions asked during the preparatory steps.
[DOCX File, 19 KB - humanfactors_v12i1e57342_app1.docx]

Multimedia Appendix 2
Usage scenario and task list.
[DOCX File, 21 KB - humanfactors_v12i1e57342_app2.docx]

References

1. de Moraes Hungria VT, Chiattone C, Pavlovsky M, et al. Epidemiology of hematologic malignancies in real-world settings: findings from the hemato-oncology latin america observational registry study. J Glob Oncol 2019 Nov;5(5):1-19. [doi: 10.1200/JGO.19.00025] [Medline: 31774711]

XSL•FO
RenderX

2. Scarfò L, Karamanidou C, Doubek M, et al. MyPal ADULT study protocol: a randomised clinical trial of the MyPal ePRO-based early palliative care system in ADULT patients with haematological malignancies. BMJ Open 2021 Nov 2;11(11):e050256. [doi: 10.1136/bmjopen-2021-050256] [Medline: 34728446]
3. What is palliative care? Get Palliative Care. URL: https://getpalliativecare.org/whatis/ [accessed 2022-08-17]
4. Consolo L, Castellini G, Cilluffo S, Basile I, Lusignani M. Electronic patient-reported outcomes (e-PROMs) in palliative cancer care: a scoping review. J Patient Rep Outcomes 2022 Sep 23;6(1):102. [doi: 10.1186/s41687-022-00509-z] [Medline: 36138279]
5. Kiderlen TR, Schnack A, de Wit M. Essential barriers and considerations for the implementation of electronic patient-reported outcome (ePRO) measures in oncological practice: contextualizing the results of a feasibility study with existing literature. J Public Health (Berl) 2023 Dec;31(12):2071-2088. [doi: 10.1007/s10389-022-01767-3]
6. Karamanidou C, Natsiavas P, Koumakis L, et al. Electronic patient-reported outcome-based interventions for palliative cancer care: a systematic and mapping review. JCO Clin Cancer Inform 2020 Jul;4(4):647-656. [doi: 10.1200/CCI.20.00015] [Medline: 32697604]
7. Vaghefi I, Tulu B. The continued use of mobile health apps: insights from a longitudinal study. JMIR Mhealth Uhealth 2019 Aug 29;7(8):e12983. [doi: 10.2196/12983] [Medline: 31469081]
8. Baxter C, Carroll JA, Keogh B, Vandelanotte C. Assessment of mobile health apps using built-in smartphone sensors for diagnosis and treatment: systematic survey of apps listed in international curated health app libraries. JMIR Mhealth Uhealth 2020 Feb 3;8(2):e16741. [doi: 10.2196/16741] [Medline: 32012102]
9. Byambasuren O, Sanders S, Beller E, Glasziou P. Prescribable mHealth apps identified from an overview of systematic reviews. NPJ Digital Med 2018;1(1):12. [doi: 10.1038/s41746-018-0021-9] [Medline: 31304297]
10. Tremblay M, Latulippe K, Guay M, et al. Usability of a co-designed eHealth prototype for caregivers: combination study of three frameworks. JMIR Hum Factors 2021 Aug 18;8(3):e26532. [doi: 10.2196/26532] [Medline: 34406123]
11. Benjumea J, Ropero J, Rivera-Romero O, Dorronzoro-Zubiete E, Carrasco A. Privacy assessment in mobile health apps: scoping review. JMIR Mhealth Uhealth 2020 Jul 2;8(7):e18868. [doi: 10.2196/18868] [Medline: 32459640]
12. Noorbergen TJ, Adam MTP, Teubner T, Collins CE. Using co-design in mobile health system development: a qualitative study with experts in co-design and mobile health system development. JMIR Mhealth Uhealth 2021 Nov 10;9(11):e27896. [doi: 10.2196/27896] [Medline: 34757323]
13. Sanders EBN, Stappers PJ. Co-creation and the new landscapes of design. CoDesign 2008 Mar;4(1):5-18. [doi: 10.1080/15710880701875068]
14. Castro AR, Arnaert A, Moffatt K, Kildea J, Bitzas V, Tsimicalis A. Developing an mHealth application to coordinate nurse-provided respite care services for families coping with palliative-stage cancer: protocol for a user-centered design study. JMIR Res Protoc 2021 Dec 13;10(12):e34652. [doi: 10.2196/34652] [Medline: 34898464]
15. MyPal: fostering palliative care of adults and children with cancer through advanced patient-reported outcome systems. MyPal. URL: https://mypal-project.eu/ [accessed 2024-07-19]
16. Maramis C, Karamanidou C, Schera F, et al. Using electronic patient reported outcomes to foster palliative cancer care: the MyPal approach. Presented at: 2019 IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE); Oct 28-30, 2019; Athens, Greece. [doi: 10.1109/BIBE.2019.00079]
17. Hoffmann S, Schraut R, Kröll T, et al. AquaScouts: ePROs implemented as a serious game for children with cancer to support palliative care. Front Digital Health 2021;3(December):730948. [doi: 10.3389/fdgth.2021.730948] [Medline: 34957461]
18. Koumakis L, Schera F, Parker H, et al. Fostering palliative care through digital intervention: a platform for adult patients with hematologic malignancies. Front Digital Health 2021;3(December):730722. [doi: 10.3389/fdgth.2021.730722] [Medline: 34977857]
19. MyPal platform—fostering palliative care for cancer patients via epros. YouTube. URL: https://youtu.be/K32nGL2R7sk?si=xvwB_u-bZ0_5-q5b [accessed 2022-08-22]
20. Tirkkonen-Condit S. Think-aloud protocols. In: Brown K, editor. Encyclopedia of Language & Linguistics: Elsevier; 2006:678-686. [doi: 10.1016/B0-08-044854-2/00479-X]
21. Lewis JR. Psychometric evaluation of the Post-Study System Usability Questionnaire: the PSSUQ. Proc Hum Factors Ergon Soc Annu Meet 1992 Oct;36(16):1259-1260. [doi: 10.1177/154193129203601617]
22. Brooke J. SUS—a quick and dirty usability scale. In: Usability Evaluation in Industry: Taylor & Francis; 1995:189-194. [doi: 10.1201/9781498710411-35]
23. Laugwitz B, Held T, Schrepp M. Construction and evaluation of a user experience questionnaire. Presented at: USAB '08: Proceedings of the 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society on HCI and Usability for Education and Work; Nov 20, 2008; Austria. [doi: 10.1007/978-3-540-89350-9_6]
24. ISO 9241-11:2018(en), ergonomics of human-system interaction— part 11: usability: definitions and concepts. ISO. URL: https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en [accessed 2021-02-14]

XSL•FO
RenderX

25. ISO/IEC 25010:2011—systems and software engineering—systems and software quality requirements and evaluation (square)—system and software quality models. ISO. 2011. URL: https://www.iso.org/standard/35733.html [accessed 2025-04-17]
26. ISO 9241-210:2019(en), ergonomics of human-system interaction—part 210: human-centred design for interactive systems. ISO. URL: https://www.iso.org/obp/ui/#iso:std:iso:9241:-210:ed-2:v1:en [accessed 2021-03-18]
27. Kashfi P, Feldt R, Nilsson A. Integrating UX principles and practices into software development organizations: a case study of influencing events. J Syst Software 2019 Aug;154:37-58. [doi: 10.1016/j.jss.2019.03.066]
28. Saksono RD, Sulistyorini D, Sagita SR, Sadita L. Usability evaluation and interface design improvement for the Maxim application with user-centered design approach. J Sistem Inf (J Inf Sys) 2024 Apr;20(1):23-40. [doi: 10.21609/jsi.v20i1.1377]
29. Robinson J. Likert scale. In: Michalos AC, editor. Encyclopedia of Quality of Life and Well-Being Research: Springer Netherlands; 2014:3620-3621. [doi: 10.1007/978-94-007-0753-5_1654]
30. Miles MB, Huberman AM, Saldana J. Qualitative Data Analysis: A Methods Sourcebook, 4th edition: SAGE; 2014. [doi: 10.1080/10572252.2015.975966]
31. Varsi C, Ledel Solem IK, Eide H, et al. Health care providers' experiences of pain management and attitudes towards digitally supported self-management interventions for chronic pain: a qualitative study. BMC Health Serv Res 2021 Mar 25;21(1):275. [doi: 10.1186/s12913-021-06278-7] [Medline: 33766028]
32. Faulkner L. Beyond the five-user assumption: benefits of increased sample sizes in usability testing. Behav Res Methods Instrum Comput 2003 Aug;35(3):379-383. [doi: 10.3758/bf03195514] [Medline: 14587545]
33. Widberg C, Wiklund B, Klarare A. Patients' experiences of eHealth in palliative care: an integrative review. BMC Palliat Care 2020 Oct 14;19(1):158. [doi: 10.1186/s12904-020-00667-1] [Medline: 33054746]
34. Benze G, Nauck F, Alt-Epping B, et al. PROutine: a feasibility study assessing surveillance of electronic patient reported outcomes and adherence via smartphone app in advanced cancer. Ann Palliat Med 2019 Apr;8(2):104-111. [doi: 10.21037/apm.2017.07.05] [Medline: 29156896]
35. Qin M, Chen B, Sun S, Liu X. Effect of mobile phone app-based interventions on quality of life and psychological symptoms among adult cancer survivors: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2022 Dec 19;24(12):e39799. [doi: 10.2196/39799] [Medline: 36534460]

Abbreviations

CLL: chronic lymphocytic leukemia
ELLOK: Greek Federation of Cancer Patients
ePRO: electronic patient-reported outcome
HCP: health care professional
HM: hematological malignancy
ISO: International Organization for Standardization
mHealth: mobile health
PSSUQ: Post-Study System Usability Questionnaire
QoL: quality of life
SUS: System Usability Scale
UEQ: User Experience Questionnaire
UI: user interface
UX: user experience

Edited by A Kushniruk, Y Senathirajah; submitted 13.02.24; peer-reviewed by C Baxter, S Markham, S Ashraf, S Yordanov; revised version received 06.11.24; accepted 15.01.25; published 24.04.25.

Please cite as:
Bonotis P, Angelidis P, Natsiavas P
Usability and User Experience Testing of a Co-Designed Electronic Patient-Reported Outcomes App ("MyPal for Adults") for Palliative Cancer Care: Mixed Methods Study
JMIR Hum Factors 2025;12:e57342
URL: https://humanfactors.jmir.org/2025/1/e57342
doi:10.2196/57342

© Panos Bonotis, Pantelis Angelidis, Pantelis Natsiavas. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 24.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction

XSL•FO
RenderX

in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Satisfaction and Usability of a Commercially Available Medication Adherence App (Medisafe) Among Medically Underserved Patients With Chronic Illnesses: Survey Study

Christa Hartch[1,2*], PhD, RN, FNP-C; Mary S Dietrich[2,3*], MS, PhD; B Jeanette Lancaster[4*], PhD, RN, FAAN; Shelagh A Mulvaney[2,5*], PhD; Deonni P Stolldorf[2*], PhD, RN

[1]School of Nursing and Health Sciences, Manhattanville University, Purchase, NY, United States

[2]School of Nursing, Vanderbilt University, Nashville, TN, United States

[3]Department of Biostatistics, Vanderbilt University Medical Center, Vanderbilt University, Nashville, TN, United States

[4]Sadie Heath Cabiness Professor and Dean Emerita, School of Nursing, University of Virginia, Charlottesville, VA, United States

[5]Department of Biomedical Informatics, Vanderbilt University Medical Center, Vanderbilt University, Nashville, TN, United States

[*]all authors contributed equally

Corresponding Author:
Christa Hartch, PhD, RN, FNP-C
School of Nursing and Health Sciences
Manhattanville University
2900 Purchase Street
Purchase, NY, 10577
United States
Phone: 1 914 323 5365
Email: christa.hartch@mville.edu

Abstract

Background: Research supports the use of mobile phone apps to promote medication adherence, but the use of and satisfaction with these apps among medically underserved patients with chronic illnesses remain unclear.

Objective: This study reports on the overall use of and satisfaction with a medication adherence app (Medisafe) in a medically underserved population.

Methods: Medically underserved adults who received care for one or more chronic illnesses at a federally qualified health center (FQHC) were randomized to an intervention group in a larger randomized controlled trial and used the app for 1 month (n=30), after which they completed a web-based survey. Objective data on app usage were provided as secondary data by the app company.

Results: The participants were very satisfied with the app, with all participants (30/30, 100%) somewhat or strongly agreeing that they would recommend the app to family and friends. Participants strongly agreed (28/30, 93%) that the reminders helped them remember to take their medications at the correct time each day, and they (28/30, 93%) found the app easy to use. Additional features accessed by some included educational features and the adherence report. Participants noted the helpfulness of having a medication list on their phones, and some used it during medication reconciliation at doctor visits. Use of the Medfriend feature, which alerts a social support person if a medication is missed, was low (n=2), but those who used it were very positive about the feature.

Conclusions: A commercially available medication adherence app was found to be useful by participants, and they were satisfied with the app and the additional features provided. The use of medication adherence mobile phone apps has the potential to positively influence chronic disease management in a medically underserved population on a large scale.

Trial Registration: ClinicalTrials.gov NCT05098743; https://clinicaltrials.gov/study/NCT05098743

(JMIR Hum Factors 2025;12:e63653) doi:10.2196/63653

KEYWORDS

medication adherence; mHealth; mobile phone; app; medically underserved; chronic disease; satisfaction; usage; health disparities

XSL•FO
RenderX

Introduction

Background

Medication adherence is vital for those with chronic illnesses who require long-term medication therapies to maintain optimal health. For example, medication adherence and persistence with high blood pressure medications are known to significantly decrease the risk of both cardiac events and stroke [1,2]. In those with type 2 diabetes, medication adherence with hypoglycemics reduces microvascular complications [3]. Unfortunately, the burden of chronic diseases is increasing, with an estimated 60% of adults in the United States having 1 chronic disease and 40% having 2 or more chronic diseases [4]. The growth of chronic disease burden coupled with a lack of medication adherence is associated with increased health care expenditures due to increased demands on health care resources [5,6] and poor health outcomes such as worsening disease status and even death [5]. The economic impact of low medication adherence is estimated to cost the US health care system between US $100 billion and $290 billion annually [5-8].

Medication adherence is therefore particularly important in medically underserved populations who seek care at federally qualified health centers (FQHCs). These centers serve communities and populations with a demonstrable unmet need for health services [9]. These centers are reporting growth in more complex patient populations because their patients have higher rates of chronic conditions and social risk factors associated with poorer health outcomes [10]. Additionally, lower rates of medication adherence are seen in lower socioeconomic populations [11,12] and those with multiple chronic conditions [13]. The reasons for this are influenced by social determinants of health and the material and social conditions in which people live [14]. Adverse social determinants of health are associated with lower medication adherence [15].

Mobile health (mHealth) interventions, defined as the use of mobile wireless technologies for public health [16], have been cited as a potential way to reduce health disparities among chronically ill and medically underserved populations [17,18]. However, despite the promise of these technologies, researchers indicate that mHealth interventions remain understudied in medically underserved populations [17,18]. This is true of medication adherence apps, which can support patients in adhering to their medications through reminders, medication educational information, adherence data, and social support. Studies have shown mixed results for the interest in mobile phone interventions in vulnerable populations [18,19]. Furthermore, research testing commercially available apps to manage chronic disease in a racially and ethnically diverse sample found that the usability of the tested apps in this population was suboptimal [20]. Understanding and gathering detailed data from diverse perspectives regarding the user experience of medication adherence apps will provide important information that is needed to support wider implementation.

A recent meta-analysis of the effectiveness of mobile apps on medication adherence in adults with chronic illnesses found that medication adherence mobile apps, which are designed to be used across a range of multiple chronic health conditions, remain underexplored [21]. This meta-analysis reported that in general, patients have a high acceptance of medication apps, but none of the studies analyzed included medically underserved populations [21]. Eight studies have demonstrated increased medication adherence with the use of medication adherence apps [22-29], but only 3 of these were conducted in low-income medically underserved populations [27-29]. Two of those studies in underserved populations were focused on hypertension [27,28], and the other included hypertension and type 2 diabetes and was a post hoc analysis of a digital health offering using a cluster-randomized design [29]. Only 1 of these studies, conducted in an urban low-income population with hypertension, obtained satisfaction information on the intervention [28]. Satisfaction with the app was high, and most participants felt they would use the app or a similar program in the future. Participants agreed that the app made it easier to keep track of their medications and that having a medication list on their phone made it easier to take care of themselves. More detailed feedback from the participants or information on which features of the app were used was not gathered [28].

A high-quality, free, commercially available smartphone medication adherence app called Medisafe supports patients in adhering to their medication regimen across disease states [30]. It uses a variety of advanced features, such as daily reminders, which can be snoozed, rescheduled, and marked as taken or missed; medication educational information in the form of medication cards and videos; an interaction checker; customizable refill reminders; adherence reports; and the ability to designate a social support person to be notified if a medication is skipped [23]. A randomized controlled trial (RCT) mixed methods evaluation in patients with coronary artery disease examined the efficacy [23] and the utility, acceptability, and engagement [31] of the Medisafe app. This study was conducted in a large urban tertiary hospital in Sydney, Australia and did not focus on a medically underserved population. In addition to improving self-reported medication adherence, overall utility was rated positively, with participants indicating that having their medication list on their phone and receiving timed reminders were useful. Most participants engaged with the app and its features; found the app acceptable, convenient, and easy to use; planned to continue using the app; and would recommend it to a family member or friend [31].

A qualitative study explored the potential benefits and barriers of using a mobile medication app in a medically underserved population in the United States [18]. The researchers found that patients were willing to try smartphone apps but expressed concerns about affordability, the technology being too complicated, not keeping phones with them all the time, and not being able to use all the features [18]. That study exposed a knowledge gap regarding the perceptions and user experiences of medically underserved patients with chronic illnesses who use free commercially available medication adherence apps.

Purpose

To address the knowledge gaps, a larger RCT investigating the efficacy of the Medisafe app (reported elsewhere) [32] was performed for evaluating the overall use and satisfaction of patients with a variety of chronic illnesses in a medically

XSL•FO
RenderX

underserved population in an FQHC in the United States. The efficacy portion of the RCT found significant improvements in both medication adherence (Cohen *d*=0.52; *P*=.01) and medication self-efficacy (Cohen *d*=0.43; *P*=.04) for participants assigned to use the app compared to the usual care group [32]. As part of this RCT, participants assigned to the intervention arm provided feedback and usage data regarding their experience using the Medisafe app [32]. This manuscript presents the summaries of the perceptions of patients enrolled in the intervention arm of the RCT regarding the usefulness of and satisfaction with the app features after 1 month of use. Given the improvements observed in medication adherence and self-efficacy, understanding patients perceived usefulness and satisfaction with the app is important to address potential barriers for uptake and use in larger medically underserved patient populations who often receive care for chronic illnesses in FQHCs.

Methods

Setting and Recruitment

Participants were recruited from November 2021 through June 2022 from an outpatient adult medicine department in an FQHC in the northeastern United States. The inclusion criteria for the RCT study were as follows: (1) adults aged 18 years or older, (2) having the ability to speak and understand English, (3) personally owning and using an Android smartphone (version 5.0 or above and at least 88 MB of phone space) or iOS smartphone (version 13 or later and at least 165 MB of phone space), and (4) taking at least one medication for a chronic condition based on the computerized medical record at the health center. Patients were excluded if they: (1) were already using a medication reminder app or other electronic reminder system such as phone alarms, (2) owned a smartphone not capable of downloading the app, (3) had a diagnosis of severe dementia or serious mental illness, or (4) were otherwise unable to use a mobile phone or the medication reminder software either physically or cognitively. For this study, only those participants who were randomized to the intervention group and used the Medisafe app were invited to participate in the survey.

Recruitment involved an informational flyer, a referral form from clinicians at the health center, and in-person recruitment. The flyer and referral forms were available to clinicians, staff, and patients in the FQHC offices and at the reception desk. The form contained study information, the contact information of the principal investigator (PI), and a place for patients interested in participation to provide their contact information. The form also contained a section for health care providers (HCPs) to refer potential patients and a section for their signature to verify that the patient's medications listed in the electronic health record were correct and current. The PI (CH) conducted in-person recruitment at the FQHC on multiple days per week and worked with clinic staff to identify potentially eligible patients. Although a convenience sample was used (ie, patients visiting the clinic on any given day), the risk of selection bias was reduced by using the aforementioned 2-prong approach to identify eligible patients for recruitment, inviting all patients meeting the eligibility criteria to participate, and using random assignment to either the intervention or control group. The 2 groups were not statistically different [32]. The PI approached eligible patients at the end of the health center visit to inform them of the study. Once the PI confirmed participant eligibility and obtained informed consent, participants were randomized to either the intervention or control group. Additional details of study procedures for the RCT have been previously published [32].

Statistical Analysis

Based on a preliminary efficacy study for the RCT [33], a total sample of 60 participants was estimated to enable the detection of differences between the groups with Cohen *d* effect values of 0.6-0.7 (80% power; α=.05) for the quantitative study variables [32]. As 30 participants were randomized to the app intervention group, their usage and satisfaction data are presented in this manuscript. Descriptive statistics and frequency distributions were used to describe the sample and determine if the data were normally distributed. Qualitative participant responses were transcribed into an Excel spreadsheet, and the content was coded and summarized as themes by the researcher (CH) and PhD faculty advisor (DPS).

Ethical Considerations

This research was approved by the Vanderbilt University Institutional Review Board (IRB #211409) and is registered with ClinicalTrials.gov (NCT05098743). All participants received a copy of the consent form. Based on participant preference, informed consent was completed as either an IRB-approved e-consent form or a hard copy. The consent form contained a privacy and confidentiality protection description ensuring that all study data are deidentified. Participants received a US $25 gift card after completion of the baseline survey and a US $35 gift card after completion of the follow-up survey.

Medisafe App Intervention

The Medisafe app is a Health Insurance Portability and Accountability Act (HIPAA) compliant medication adherence app that is available at no cost in the iTunes and Google app stores. In previous studies, Medisafe was ranked highly among medication reminder apps [30,34]. The Medisafe app provides interactive and customizable daily timed reminders to reinforce medication taking at a set time every day through a push notification, equivalent to an alarm or text message. The reminders can be snoozed, rescheduled, or marked as taken or skipped, and they are repeated a total of 3 times in 10-minute intervals if the participant does not mark the medication dose. Additional features include educational information in the form of a medication database that includes written and video content [30]. The written content is in the form of a medication card, which Medisafe terms a leaflet, and it reviews what the medicine is used for, medication interactions, what to do if the user misses a dose, what the user should watch for, possible side effects, how it should be used, and where to store the medication. Some medications also have video content, consisting of a brief clip of an HCP reviewing the most important considerations when taking the medication, which can be viewed on tapping the information icon. There is also an interactions tab that lists possible interactions with the medications or food/alcohol.

XSL•FO
RenderX

Lastly, there is an interaction checker where participants can check for interactions between their medications. The app also has a Medfriend feature, which allows participants to designate a family member or friend as their support person. The Medfriend feature will alert the designated Medfriend who can provide peer support and additional reminders through text messages, emails, or a telephone call if the patient misses a dose. The language mode of the app can be switched, if desired, to multiple foreign languages, including Spanish.

The PI helped consented participants set up the app using a copy of the patient's medication list extracted from the electronic health record. The PI also reviewed how participants could access and edit their medications; access medication educational content; and indicate when a medication was taken, skipped, or rescheduled. The PI reviewed with participants additional app features such as Medfriend, medication interaction checking, adherence reports, and refill reminders. Participants were also shown how to access the help and support section in the app. Following the app set up, the PI provided previously developed educational materials as a take-home resource. These materials, specific to either an iPhone or Android smartphone, included a laminated "quick tips" card with short instructions on the reviewed features and how to access them. Additional detailed instructions on how to use the app were printed in a question-and-answer format and distributed to participants.

Data Collection and Study Procedures

All study data were collected using observation, a REDCap web-based survey, and secondary data provided by the app company and were obtained using a data-sharing agreement between institutions. REDCap is a secure web-based software platform designed to support data capture for research studies [35,36]. Following consent, all participants completed the baseline study survey. Those randomized to the intervention group also completed a survey at study end to obtain feedback on the app, including usability and satisfaction.

Observation

While setting up the app for the intervention group, the PI completed a study-specific observational behavioral checklist. The purpose of the checklist was to inform the researchers if participants had difficulty setting up the app and how long it took them to do so. The checklist included documenting whether the participant had difficulty visualizing the app and had difficulty with dexterity while setting up the app, and mentioning the number of times the participant's input of medications needed to be corrected. The length of time in minutes from starting the download of the app to completing app set up and reviewing the app was also documented.

Survey

After 1 month, based on preference, participants completed the follow-up survey online, by phone, or in-person at the health center. Participants who did not complete the follow-up survey within 10 days of the 1-month follow-up date received 2 reminders via phone, email, or text message.

Measures

The end-of-study survey included 11 questions that assessed satisfaction and usability using a 5-point Likert scale with responses ranging from "strongly disagree" to "strongly agree." The survey questions were developed and pilot tested before use [33]. Seven of the first 11 questions were developed by Santo et al [31] and were used with permission in this study, while the remaining 4 were developed by the researchers. Six additional questions asked about the use of additional features such as the educational information, Medfriend feature, interaction checker, adherence report, refill reminder, and additional morning reminder of the Medisafe app. These questions asked participants whether they used a given feature, and if they did, whether they found the feature useful. There were open-text response options available to elicit qualitative data from the app participants such as how a feature helped them manage their medications and what they found most useful about the feature. The remainder of the survey included 6 general use questions previously developed and pilot tested, 4 of which were "Yes/No" questions (eg, did you use the refill reminder and did you have technical issues with the smartphone app?). The remaining 2 questions assessed how often medication reminders were received each day and which language the participant used.

Secondary Data

Deidentified usage data were obtained from the Medisafe company at study completion. Medisafe provided the PI with objective user interactions with the Medisafe app, such as whether educational information in the form of a leaflet was accessed by participants and whether the Medfriend feature was used.

Analysis

Descriptive statistics and frequency distributions were used to describe the sample and determine if the data were normally distributed. Open-ended survey question responses were imported into an Excel (Microsoft Corp) spreadsheet, and the content was coded and summarized as themes by the researcher (CH) and the PhD faculty advisor (DPS). This approach was taken given the short, free-text, and limited responses received. Since the qualitative data came from the open-ended survey responses, data collection was based on sample size rather than data saturation.

Results

Participant Characteristics

Complete details are included in Table 1, and information can also be found in the app efficacy manuscript [32]. A flowchart of study participants in the main RCT can be found in Figure 1. The median age of the 30 participants using the app was 53.5 years (IQR 37-76 years). Most participants in the intervention group were non-White (23/29, 79%). Races/ethnicities were as follows: Asian (5/29, 17%), Black or African American (10/29, 35%), Hispanic/Latino (4/29, 14%), Native American or Alaska Native (1/29, 3%), and other (3/29, 10%). More than 75% of the participants had government insurance (25/30, 83%), and a small number of participants were uninsured (2/30, 7%). Brief

XSL•FO
RenderX

health literacy scores were generally high (median 12.0 of a possible score of 15, IQR 5-15). Slightly more than half of the participants (16/29, 55%) reported that it was either very or somewhat difficult to pay their monthly bills. The most common chronic illness was hypertension (22/30, 73%), followed by hyperlipidemia (19/30, 63%) and type 2 diabetes (14/30, 47%). Most participants (25/30, 83.3%) had 2 or more chronic illnesses.

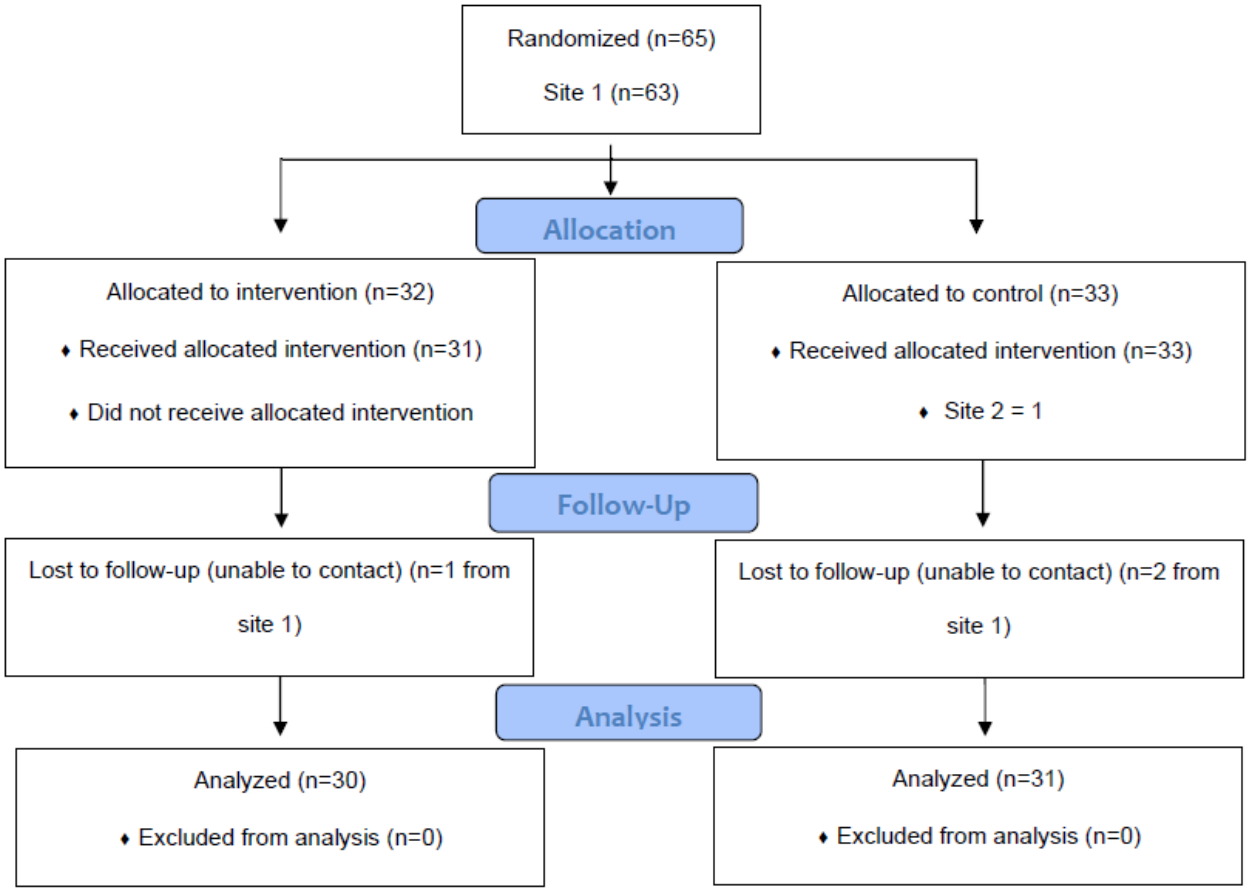
Table 1. Sociodemographic characteristics of the participants using the app.

Characteristic	Value, n (%)
Race/ethnicity (N=29)	
Asian	5 (17)
Black or African American	10 (35)
Hispanic/Latino	4 (14)
Native American or Alaska Native	1 (3)
White	6 (21)
Other	3 (10)
Marital status (N=30)	
Married/partnered	15 (50)
Single/never married	15 (50)
Employment status (N=30)	
Employed	15 (50)
Unemployed	12 (40)
Retired	3 (10)
Education (N=30)	
Some high school or less	8 (27)
High school graduate	4 (13)
College credit, no degree	8 (27)
Trade/vocational training	4 (13)
Associate’s degree	2 (7)
Bachelor’s degree or higher	4 (13)
Difficulty paying bills (N=29)	
Very difficult	6 (21)
Somewhat difficult	10 (35)
Not very difficult	8 (28)
Not at all difficult	5 (17)
Type of health insurance (N=30)	
Uninsured (sliding scale)	2 (7)
Government insurance	25 (83)
Private insurance	3 (10)
Current chronic illness (N=30)	
Hypertension	22 (73)
Type 2 diabetes	14 (47)
Hyperlipidemia	19 (63)
Asthma	5 (17)
Other[a]	11 (37)

[a]Includes depression (n=1), type 1 diabetes (n=1), chronic obstructive pulmonary disease (n=1), heart disease (n=1), cirrhosis (n=1), anxiety (n=1), gout (n=1), rheumatoid arthritis (n=1), thyroid disorder (n=1), hypothyroidism (n=2), gastroesophageal reflux disease (n=1), arthritis (chronic pain) (n=1), and fibromyalgia (n=1).

XSL•FO
RenderX

Figure 1. Flowchart of study participants in the larger randomized controlled trial of the efficacy of the app intervention.



Behavioral Observations While Setting Up the App

During app set up, some participants (4/30, 13%) expressed difficulty visualizing the app owing to the unavailability of their eyeglasses, which they stated were either in their car or left at home. No participants had difficulty with dexterity while setting up the app, and the median time it took from starting the app download to completing the set up and reviewing the app was 15 minutes (IQR 10.0-25.0; minimum 10, maximum 30 minutes). The majority of participants (21/30, 70%) did not need to be corrected when they entered the medications. However, 4 (13%) were corrected by the researcher once, 3 (10%) were corrected twice, and 2 (7%) were corrected thrice. Patients were corrected when they spelled the medication name incorrectly, chose the incorrect medication dose, or set an incorrect time for the reminder.

Satisfaction and Utility

Summaries of the participants' reports of satisfaction are presented in Table 2.

Most participants (27/30, 90%) strongly agreed that they liked the app design, while most (25/30, 83%) strongly agreed that it was useful to have their medication list on their smartphone. Some participants (2/30, 7%) mentioned the usefulness of the app when seeing other HCPs to indicate the medications they were taking during medication reconciliation. Furthermore, a large proportion of participants (28/30, 93%) strongly agreed that the reminders helped them to remember to take their medications at the correct time each day. The majority of participants strongly agreed that the app was easy to use (27/30, 90%) and convenient (28/30, 93%) and that they would continue using the app (26/30, 87%). A small number of participants (2/30, 7%) somewhat agreed that they would continue using the app. It is important to note that some participants (2/30, 7%) strongly disagreed that they would continue using the app, because they found the reminders annoying. All the participants (30/30, 100%) somewhat agreed or strongly agreed that they would recommend the app to family and friends.

XSL•FO
RenderX

Table 2. Satisfaction with the app (N=30).

Satisfaction information	Value (N=30), n (%)
Liked the app design	
Neutral	1 (3)
Somewhat agree	2 (7)
Strongly agree	27 (90)
It is easy to tap the correct icon with my finger	
Neutral	1 (3)
Somewhat agree	1 (3)
Strongly agree	28 (93)
I am able to see all the options in the app	
Neutral	3 (10)
Somewhat agree	1 (3)
Strongly agree	26 (87)
It is useful to have a medication list on the smartphone	
Somewhat disagree	1 (3)
Neutral	1 (3)
Somewhat agree	3 (10)
Strongly agree	25 (83)
Reminders helped me remember to take my medications at the correct time each day	
Neutral	1 (3)
Somewhat agree	1 (3)
Strongly agree	28 (93)
Found it easy to use the app	
Somewhat disagree	1 (3)
Neutral	1 (3)
Somewhat agree	1 (3)
Strongly agree	27 (90)
Found it easy to set up reminders in the app	
Neutral	3 (10)
Somewhat agree	2 (7)
Strongly agree	25 (83)
Found it convenient to have the app	
Strongly disagree	1 (3)
Somewhat disagree	1 (3)
Strongly agree	28 (93)
Found it useful to snooze the reminder	
Strongly disagree	1 (3)
Somewhat disagree	1 (3)
Neutral	10 (33)
Somewhat agree	3 (10)
Strongly agree	15 (50)
Will continue using the app	
Strongly disagree	2 (7)

XSL•FO
RenderX

Satisfaction information	Value (N=30), n (%)
Somewhat disagree	1 (3)
Neutral	1 (3)
Strongly agree	26 (87)
Will recommend the app to family and friends	
Somewhat agree	2 (7)
Strongly agree	28 (93)

The 2 participants (10%) who were dissatisfied with the app described the reasons why. One said:

> *It is annoying when you get a reminder and you are in the middle of doing your work. This is not for everybody. I work on the computer and on my cell phone and it is very distracting to receive the reminder in the middle of working on something. It might be better for someone who doesn't have as much going on. I find it very distracting.*

The other participant had technical difficulties but blamed it on the type of phone they had:

> *The bug thing with the notification alarms was a problem. I have an android - a cheap phone. My phones get destroyed because of the type of work I do.*

Use of Educational Information

Almost half of the participants (12/30, 40%), self-reported accessing educational information. To the contrary, objective usage data from Medisafe indicated that only 3 participants (10%) accessed the educational content, which was defined as cards termed “leaflets” or videos. According to Medisafe, this was done for a total of 14 medications. Additionally, Medisafe reported that only 1 participant accessed 4 different videos and 1 leaflet, 1 participant accessed 1 leaflet, and 1 participant accessed 6 different videos and 2 leaflets. Although not all participants actually accessed the information, those who reported accessing the educational information (12/12, 100%) found the information useful. When asked about how they used the educational information in the app, the participants reported learning about the side effects of the medications (6/12, 50%), reported that it was helpful for general knowledge (4/12, 33%), and mentioned using it to learn more about medication and food interactions (2/12, 17%).

Medfriend Feature

Based on usage data from Medisafe, only 1 participant (3%) used the Medfriend feature. That participant was very positive about the feature and reported that her husband would call her to say, “Are you taking your medicines?” She stated:

> *It gets him involved. It makes him recognize that I need support and I need to take the medicine. It makes me know he loves me.*

Another participant self-reported using the Medfriend feature, but there was no indication of use in the Medisafe data. The participant reported that when her husband was notified, he would send a text about her forgetting her medications and she would remember to take them. Those who did not use the Medfriend option were asked, “who might that person be for you?” Among those who responded (18/30, 60%), the top 3 most common responses were their sibling (4/18, 22%), their child (4/18, 22%), and their husband or wife (4/18, 22%).

Interaction Checker

All participants who used the interaction checker (5/30, 17%) agreed that it was useful. One participant reported that 2 of the medications she had been taking together should be taken separately and stated, “It was a lifesaver!” The other 4 participants expressed an appreciation for being able to have access to this type of information. The use of the interaction checker was distinct from the educational content and could not be verified in the Medisafe data as Medisafe does not register or track the use of the interaction checker.

Adherence Reports and Reminders

Participants (8/30, 27%) who checked their adherence report agreed it was useful. The adherence report provided them with a history of their daily missed and taken medications as well as a weekly adherence percentage based on what they reported when marking medication reminders in the app. Participants reported experiencing positive reinforcement for adhering to their medications through the adherence report, mentioned the affirmation they received when they saw a high percentage of adherence, and reported appreciating the positive reinforcement as useful. Some participants (2/30, 7%) mentioned that it incentivized them to reach higher levels of adherence.

Slightly less than half of the participants (13/30, 43%) received reminders to take their medications 2 times a day, while around one-third (9/30, 30%) received reminders 3 or more times a day. Reminders were generally well received:

> *I like the reminder. The shaking of the pill bottle helps me. Sometimes I will wake up at night and remember hearing the shaking pill bottle that day and I will get out of bed and check if I took my pills that day. I might be cooking with the grandkids and the first reminder goes off. I might ignore it but with the second reminder I might put the bottles on the counter so I can remember.*

Three participants (15%) mentioned the app’s helpfulness, particularly for those who have multiple chronic illnesses and take multiple medications.

> *This app was a lifesaver. I take a lot of different medication so sometimes I forget whether I took the*

XSL•FO
RenderX

> *medication or not. I can check the app to see if I took it or not.*

One of those 3 participants commented as follows:

> *It is perfect for people who have multiple illnesses and take a lot of different medications. I take eight different medications a day.*

Most of the participants (27/30, 90%) did not use the refill reminder. In their comments, a number of participants said they received automatic refill reminders from their pharmacy and therefore did not need this feature of the app. All participants (30/30, 100%) used the app in English, and the majority (24/30, 80%) did not make any changes, such as changing the time of a reminder, removing or adding a medication, or changing the medication dose in the app. Many participants (13/30, 43%) said they would use the app to manage someone else's medications.

Technical Difficulties Using the App

Some participants (4/30, 13%) mentioned they had technical difficulties. Of those who reported difficulties, 3 (75%) needed to allow notifications from the app to hear the reminders. Another user mentioned that they had to tap the "take all" icon a number of times before it registered and suggested that it should be made bigger or be more centrally placed. Participants gave additional feedback about the app when asked (20/30, 67%). In this section, participants (7/20, 35%) specifically mentioned liking the reminder.

> *It was really nice to hear that shaking sound. It was fun.*

Some participants (2/20, 10%) reported that the snooze function was particularly helpful when they were away from home.

> *The snooze option is helpful to use when I am out and don't have my medications. When I come back home it reminds me so I remember to take it.*

Social Support

The 2 participants (7%) who self-reported using the Medfriend feature were very enthusiastic.

> *It's a great app and I love it. My husband is on it for his meds and I am his Medfriend. I am also going to get my mother hooked up on it.*

The other participant shared her thoughts about the feature, highlighting her increased feelings of self-efficacy and social support.

> *This app is about being a team player. You are able to help me and I am able to help you. I can now say "I did it" "I can do this." This is a good app. I can't see anyone who is interested in their health not using this app. Since being introduced to this app I know that it is there for me.*

Discussion

Principal Findings

As part of an RCT using the Medisafe medication adherence app in a medically underserved population with a variety of chronic illnesses, behavioral observations on app use and satisfaction and usage data were gathered from participants in the intervention arm of the study and the Medisafe company. The quantitative RCT results (reported elsewhere) [32] found significant improvements in both medication adherence and medication self-efficacy for participants who used the app. The portion of the RCT presented in this manuscript, which collected behavioral observations and satisfaction and usage data from the intervention arm, identified that participants were satisfied with the app and found it useful. Even though the use of the additional features was generally low, those who used them found them useful. Most participants did not need help setting up the app. An important strength of this RCT is that it explored patients' perceptions of the usefulness of the app and their satisfaction with the app and therefore fills an important knowledge gap. This was done by collecting both quantitative and qualitative data through open-response questions, which gave voice to the perspectives of a low-income racially and ethnically diverse sample of adults with chronic illnesses receiving care in an FQHC [17,18]. As FQHCs are reporting a growth in the rates of treating complex chronic conditions [10] and there are lower rates of medication adherence among populations with lower socioeconomic status and those with multiple conditions [37], the implementation of tools to enhance medication adherence is imperative. Understanding the user experience with the Medisafe mobile app demonstrated that wider-scale use of the Medisafe app is feasible in a low-income population with multiple chronic illnesses. Systematic literature reviews have pointed out a gap in implementation studies of mobile app interventions in this population [38]. This study addressed an important knowledge gap by demonstrating that the use of a commercially available free medication adherence app is a viable option for medically underserved adult patients with chronic diseases.

Prior research found that medically underserved patients expressed reluctance about paying for a medication adherence app [18]. While not directly addressed in this study, some participants anecdotally asked before enrolling in the study if they would need to pay for the app, and when told it was free, they expressed interest in participating. This underscores the importance of not having patients incur additional costs for the technology and was a strength of this intervention.

Most participants were able to set up the app with minimal assistance, with a median duration time of 15 minutes during the behavioral observation. It is important to note that 30% of the participants needed to be corrected 1-3 times when setting up the app, thus pointing to the importance of helping some patients set up the app initially and checking that the medications are entered accurately. This highlights a difficulty with individuals setting up the app. Although not implemented in this study, another option is to import medications from other databases, such as Apple Health, or a pharmacy directly. This may shorten the time it takes to set up the app. Once the app is set up properly, in addition to HCPs assisting patients, the help and support page and the company contact could serve as a resource for patients.

Survey data indicated that satisfaction with the app was high, with most patients strongly agreeing it was easy to use. All

XSL•FO
RenderX

intervention group participants (30/30, 100%) strongly agreed that they would recommend the app to family or friends. This was higher than the proportion in the study by Santo et al [31], which used Medisafe and found that 78.6% of patients with coronary heart disease would recommend it [31], and the study by Anglada-Martinez et al [24], which used a similar app and found that 71.4% of patients receiving treatment for hypertension, dyslipidemia, heart failure, or HIV would recommend it [24]. Both these studies were, however, conducted outside of the United States. It may be that Americans are more familiar with app technologies and feel more comfortable recommending apps to others.

Qualitative research conducted in a medically underserved population with chronic illness regarding the use of medication adherence apps found that technical issues and complexity were concerns when setting up and using these apps [18]. One study involving a medication adherence app similar to Medisafe indicated that 50% of participants reported problems receiving reminders [24]. The Medisafe app used in this study is a commercially available app with high-quality assessment ratings [30,39], and in this study, technical issues were rare. The most common issue was not receiving the reminders until the participant allowed notifications from the app in their phone settings. No participants expressed that the app was technically complex, which was previously cited as a concern in this population but was not an issue in this study [18]. One participant had ongoing technical issues receiving reminders. These findings suggest that the Medisafe app can be implemented in this population from a technology standpoint, and participants did not find it difficult to use.

Similar to other studies involving the Medisafe app, feedback results point to receiving timed medication reminders as the most used aspect of the app [31]. Furthermore, feedback regarding the app aligns with the findings of other studies linking medication reminders with medication adherence [22,25,28,29]. Participants found the snooze function of the reminder helpful when they were not home to take their medications and used this function as a reminder to take their medications when they got home. The snooze function therefore was an important component of the app when participants experienced disruptions in their routines, such as being away from their medications. The findings also align with a previous study in a medically underserved population where participants indicated that disruptions in their daily routines negatively affected their medication adherence [18]. This study demonstrated that patients used the reminder feature when available, and the majority of patients found it helpful in improving medication adherence. The 1 participant who was bothered by the reminders used his phone for work and found the reminders distracting if he was using his phone for work purposes. The reminders predominantly targeted the phenomenon of forgetting, which has been found by a study to be the most likely cause of reported nonadherence in low-income uninsured patients with multiple chronic illnesses [13].

Research has shown that both patient knowledge of medications and their satisfaction with the information provided about their medications can improve medication adherence [40]. There was a discrepancy between the data reported by Medisafe and the number of participants who self-reported accessing educational information. Although the reason for this discrepancy is not clear, several possibilities exist. First, participants could have overstated the use of educational features. Another potential reason might be related to the specific data Medisafe defines as educational data. Medisafe does not collect data on the use of the interaction tracker or the “For You” tab at the bottom of the app and only collects data if a participant clicks on the educational leaflet and opens it up. In contrast, participants might have perceived content under the “For You” tab and drug interaction materials or videos as educational materials, resulting in a discrepancy between patient self-reported data and Medisafe data regarding accessing educational content. The feedback received demonstrated that participants who reported accessing the educational information (less than half) were very positive about doing so. The educational information was found to be useful for learning about side effects and food and medication interactions. Because individuals have different preferences for the amount of medication information they receive and the way that information is delivered [41], the modularity of the Medisafe app is useful to facilitate patient education in a practical and less burdensome way. The information is available at the patient’s fingertips and can be accessed as frequently as needed to learn what they want at their convenience. The educational app feature is also advantageous to HCPs as it alleviates some of the burden and time commitment associated with educating patients about their medications.

Social support has been found to have a positive effect on medication adherence [42-44]. Studies deploying digital technologies in the form of web-based online communities to provide social support have generally demonstrated that they can support people emotionally, socially, practically, and politically [45]. However, using technology to provide social support has not been studied in the context of a commercially available medication adherence mobile app. This study addressed this gap by studying social support via a commercially available app in the context of medication adherence. The Medisafe app offers social support in the form of Medfriend, and this is the first known study to incorporate this feature as part of the study intervention. Some studies have pointed out that online social support networks for those with specific chronic illnesses lessened the burden on relationships with family and friends, who are referred to as “offline” support persons [45,46]. However, despite asking and offering to demonstrate how to set up the Medfriend feature, usage of the Medfriend feature was very low. This study did not gather information on why participants chose not to set up the Medfriend feature, and this is a limitation of the study. It may have been because this feature was seen as too burdensome by the patients or their support people, most of whom were identified as family members. Patients might have avoided using Medfriend due to confidentiality concerns associated with this feature, which entails giving access to the user’s medication list, as many participants (13/30, 43%) were willing to use the app to manage someone else’s medications but chose not to share their own medication information. Another challenge regarding the Medfriend feature was that there was a discrepancy between Medisafe data and self-reported data on the use of the Medfriend feature. Despite the aforementioned concerns, the 2

XSL•FO
RenderX

participants who reported using the Medfriend feature were satisfied with it as they perceived that the app fostered social support. To further explore the social support feature of the app, research on dyads who use the app to manage the medications of family members might shed new light on the phenomenon of incorporating social support into mobile apps. By studying a subset of the population, including patients and their caregivers, the social support feature may be used more frequently. If the confidentiality of medication lists proves to be a barrier, a feature that dissociates specific medications from the reminder might address that concern. Support persons could receive a general text that their online Medfriend has not taken their medications without sharing details on the specific medications.

Participants reported that having a list of medications on their phone was beneficial, which was also noted in a population of patients with coronary heart disease who used the Medisafe app in Australia [31]. When managing chronic illnesses, patients are often referred outside of the FQHC setting or require hospitalization to receive care. Some participants mentioned using the phone medication list for medication reconciliation when seeing other HCPs. This finding is in contrast to that of another study of patients presenting to an emergency department setting, which found that emergency department patients rarely used their mobile phones to share their medication list during medication reconciliation [47]. Medication reconciliation can be facilitated through the adoption of these technologies. HCPs in both primary care and tertiary care settings should suggest and support patients with implementing researched medication adherence mobile apps. The sample of this study included many patients with multiple chronic conditions. These participants appreciated the ability of the app to work across multiple chronic illnesses and its helpfulness when taking multiple medications. This finding underscores the importance of advocating for the use of medication adherence apps like Medisafe, which can work across a range of illnesses and medications and can be easily adjusted when medications change over the disease trajectory. Additionally, participants who used the adherence report felt that it provided positive reinforcement and was an incentive to reach higher levels of adherence. Similar to what has been reported in other studies, a majority of participants reported not using the refill reminder because they already received text alerts from their pharmacy, which they found helpful [18]. When HCPs select apps for patients to enhance their medication adherence, careful attention to app features and evaluation of existing research findings, such as the findings of this study, are important.

Limitations

This study had several limitations. First, the study duration of 1 month does not provide insights regarding long-term patient satisfaction and continued use of app features, which are important aspects of chronic disease management. Though this study found high satisfaction and usability of the app during the first month of use, future studies should evaluate the role of time in app usage and satisfaction. Medisafe data and self-reported data showed that the uptake of educational information and the Medfriend social support feature was low. There was an unresolved discrepancy in the number of participants who reported accessing educational information and the actual usage identified from the Medisafe data. The discrepancy might be because Medisafe data only captured if the leaflet was accessed. Patients may have perceived accessing educational information as clicking the interaction button or clicking the “For You” tab at the bottom of the app, which Medisafe data did not capture. This can be clarified in the future by a more detailed definition of what constitutes the educational features of the app. Incorporating interviews to clarify subsequent survey results would strengthen future research studies. Another limitation of the study is that we did not gather participants’ inputs about why they chose not to use the Medfriend feature. Therefore, this study cannot speak about the potential benefits of this feature. Finally, although the app can be used in several languages and many patients who seek care at FQHCs speak a primary language other than English [48], the researchers were not able to incorporate multiple languages into the study protocol.

Future Research

FQHCs and primary care settings working with adults who are chronically ill should consider medication adherence mobile phone apps as acceptable and practical tools to support medication adherence. Future studies could include a larger sample, consider the use of the available provider portal, and consider the experiences of both providers and patients. Cost analysis could be performed, and hospitalization rates and long-term usage and health outcomes over time could be studied. This study was for a 30-day period, but a study with a longer duration is necessary to see if the use of the app is sustained over time. In this study, only 2 participants reported barriers to using the app, and a larger long-term study could further explore barriers to sustained use and strategies for maintaining engagement in this population. Future research should use mixed methods to provide insights into app modification, the nature of barriers to use, and how app features, such as the Medfriend feature, could more easily be implemented among patients who might benefit the most from such features. As uptake of the additional features of the app, such as educational information and the Medfriend option, was low in this study, future research using larger datasets could explore what types of patients chose to use specific features and why they did. We purposefully did not require certain features to be used because we wanted to organically discover which features were most often used, if any.

Studying the usability of the app and its associated effects in ethnic populations in various languages is an important area of future research as community health centers serve a large number of patients with limited English proficiency [48].

Conclusion

This study demonstrated that the medication adherence app is a useful, convenient, and feasible intervention in an FQHC setting. The various features of this app positively influenced the medication-taking behaviors of adults with one or more chronic illnesses. Participants were satisfied with the app and the features they chose to use. Reminders were viewed as helpful by the majority of participants. The medication list feature was particularly useful for patients who had multiple chronic

XSL•FO
RenderX

conditions and saw multiple providers, and some used it to facilitate medication reconciliation. The findings of this study have important clinical implications, as clinicians can recommend the use of medication adherence apps as tools to provide support in adhering to medication regimens and as additional tools to use during medication reconciliation.

Acknowledgments

We would like to extend our gratitude to the leadership and staff of the health center and the participants who made this work possible. Support for this study was provided by the Vanderbilt Institute for Clinical and Translational Research (VICTR) #VR56076, Student Achievement Research Award (SARA), and VICTR #VR60349. The research study was also supported by CTSA award UL1TR002243 from the National Center for Advancing Translational Sciences. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Center for Advancing Translational Sciences of the National Institutes of Health.

Data Availability

Limited deidentified data are available upon request from the corresponding author.

Conflicts of Interest

None declared.

References

1. Corrao G, Parodi A, Nicotra F, Zambon A, Merlino L, Cesana G, et al. Better compliance to antihypertensive medications reduces cardiovascular risk. J Hypertens 2011 Mar;29(3):610-618. [doi: 10.1097/HJH.0b013e328342ca97] [Medline: 21157368]
2. Vrijens B, Antoniou S, Burnier M, de la Sierra A, Volpe M. Current situation of medication adherence in hypertension. Front Pharmacol 2017 Mar 01;8:100 [FREE Full text] [doi: 10.3389/fphar.2017.00100] [Medline: 28298894]
3. Yu AP, Yu YF, Nichol MB. Estimating the effect of medication adherence on health outcomes among patients with type 2 diabetes--an application of marginal structural models. Value Health 2010 Dec;13(8):1038-1045 [FREE Full text] [doi: 10.1111/j.1524-4733.2010.00787.x] [Medline: 20946182]
4. Chronic Diseases in America. National Center for Chronic Disease Prevention and Health Promotion. URL: https://www.cdc.gov/chronic-disease/about/index.html [accessed 2024-12-26]
5. Osterberg L, Blaschke T. Adherence to medication. N Engl J Med 2005 Aug 04;353(5):487-497. [doi: 10.1056/nejmra050100]
6. Viswanathan M, Golin CE, Jones CD, Ashok M, Blalock SJ, Wines RC, et al. Interventions to improve adherence to self-administered medications for chronic diseases in the United States: a systematic review. Ann Intern Med 2012 Dec 04;157(11):785-795 [FREE Full text] [doi: 10.7326/0003-4819-157-11-201212040-00538] [Medline: 22964778]
7. Thinking outside the pillbox: A system-wide approach to improving patient medication adherence for chronic disease. New England Healthcare Institute. 2009. URL: https://www.akleg.gov/basis/get_documents.asp?session=28&docid=3370 [accessed 2024-12-26]
8. Peterson A, Takiya L, Finley R. Meta-analysis of trials of interventions to improve medication adherence. Am J Health Syst Pharm 2003 Apr 01;60(7):657-665. [doi: 10.1093/ajhp/60.7.657] [Medline: 12701547]
9. Medically Underserved Areas and Population (MUA/P). Health Resources and Services Administration. URL: https://bhw.hrsa.gov/workforce-shortage-areas/shortage-designation#mups [accessed 2024-12-26]
10. Health Centers are Providing Care to Growing Numbers of Patients with Complex Needs. National Association of Community Health Centers. URL: https://opus-nc-public.digitellcdn.com/uploads/nachc/articles/808aaa175320895896c418d5fa27d6f6.pdf [accessed 2024-12-26]
11. Osborn CY, Mayberry LS, Wagner JA, Welch GW. Stressors may compromise medication adherence among adults with diabetes and low socioeconomic status. West J Nurs Res 2014 Oct 25;36(9):1091-1110 [FREE Full text] [doi: 10.1177/0193945914524639] [Medline: 24569697]
12. Bidwal M, Lor K, Yu J, Ip E. Evaluation of asthma medication adherence rates and strategies to improve adherence in the underserved population at a Federally Qualified Health Center. Res Social Adm Pharm 2017 Jul;13(4):759-766. [doi: 10.1016/j.sapharm.2016.07.007] [Medline: 27595427]
13. Fernandez-Lazaro CI, Adams DP, Fernandez-Lazaro D, Garcia-González J, Caballero-Garcia A, Miron-Canelo JA. Medication adherence and barriers among low-income, uninsured patients with multiple chronic conditions. Res Social Adm Pharm 2019 Jun;15(6):744-753. [doi: 10.1016/j.sapharm.2018.09.006] [Medline: 30241872]
14. Reutter L, Kushner KE. 'Health equity through action on the social determinants of health': taking up the challenge in nursing. Nurs Inq 2010 Sep 16;17(3):269-280. [doi: 10.1111/j.1440-1800.2010.00500.x] [Medline: 20712665]

XSL•FO
RenderX

15. Wilder ME, Kulie P, Jensen C, Levett P, Blanchard J, Dominguez LW, et al. The impact of social determinants of health on medication adherence: a systematic review and meta-analysis. J Gen Intern Med 2021 May 29;36(5):1359-1370. [doi: 10.1007/s11606-020-06447-0] [Medline: 33515188]
16. mHealth: use of appropriate digital technologies for public health: report by the Director-General. World Health Organization. 2017. URL: https://iris.who.int/handle/10665/274134 [accessed 2024-12-26]
17. Peek ME. Can mHealth interventions reduce health disparities among vulnerable populations? Divers Equal Health Care 2017;14(2):44-45. [doi: 10.21767/2049-5471.100091]
18. Hincapie AL, Gupta V, Brown SA, Metzger AH. Exploring perceived barriers to medication adherence and the use of mobile technology in underserved patients with chronic conditions. J Pharm Pract 2019 Apr 06;32(2):147-153. [doi: 10.1177/0897190017744953] [Medline: 29212410]
19. Stockman M, Modzelewski K, Steenkamp D. Mobile health and technology usage by patients in the diabetes, nutrition, and weight management clinic at an urban academic medical center. Diabetes Technol Ther 2019 Jul;21(7):400-405. [doi: 10.1089/dia.2018.0369] [Medline: 31045447]
20. Sarkar U, Gourley GI, Lyles CR, Tieu L, Clarity C, Newmark L, et al. Usability of commercially available mobile applications for diverse patients. J Gen Intern Med 2016 Dec 14;31(12):1417-1426 [FREE Full text] [doi: 10.1007/s11606-016-3771-6] [Medline: 27418347]
21. Peng Y, Wang H, Fang Q, Xie L, Shu L, Sun W, et al. Effectiveness of mobile applications on medication adherence in adults with chronic diseases: a systematic review and meta-analysis. J Manag Care Spec Pharm 2020 Apr;26(4):550-561 [FREE Full text] [doi: 10.18553/jmcp.2020.26.4.550] [Medline: 32223596]
22. Morawski K, Ghazinouri R, Krumme A, Lauffenburger JC, Lu Z, Durfee E, et al. Association of a smartphone application with medication adherence and blood pressure control: the MedISAFE-BP randomized clinical trial. JAMA Intern Med 2018 Jun 01;178(6):802-809 [FREE Full text] [doi: 10.1001/jamainternmed.2018.0447] [Medline: 29710289]
23. Santo K, Singleton A, Rogers K, Thiagalingam A, Chalmers J, Chow CK, et al. Medication reminder applications to improve adherence in coronary heart disease: a randomised clinical trial. Heart 2019 Feb 27;105(4):323-329. [doi: 10.1136/heartjnl-2018-313479] [Medline: 30150326]
24. Anglada-Martínez H, Martin-Conde M, Rovira-Illamola M, Sotoca-Momblona JM, Sequeira E, Aragunde V, et al. Feasibility and preliminary outcomes of a web and smartphone-based medication self-management platform for chronically ill patients. J Med Syst 2016 Apr 12;40(4):99. [doi: 10.1007/s10916-016-0456-y] [Medline: 26872781]
25. Labovitz DL, Shafner L, Reyes Gil M, Virmani D, Hanina A. Using artificial intelligence to reduce the risk of nonadherence in patients on anticoagulation therapy. Stroke 2017 May;48(5):1416-1419. [doi: 10.1161/strokeaha.116.016281]
26. Svendsen M, Andersen F, Andersen K, Pottegård A, Johannessen H, Möller S, et al. A smartphone application supporting patients with psoriasis improves adherence to topical treatment: a randomized controlled trial. Br J Dermatol 2018 Nov 05;179(5):1062-1071. [doi: 10.1111/bjd.16667] [Medline: 29654699]
27. Chandler J, Sox L, Kellam K, Feder L, Nemeth L, Treiber F. Impact of a culturally tailored mHealth medication regimen self-management program upon blood pressure among hypertensive Hispanic adults. Int J Environ Res Public Health 2019 Apr 06;16(7):1226 [FREE Full text] [doi: 10.3390/ijerph16071226] [Medline: 30959858]
28. Patel S, Jacobus-Kantor L, Marshall L, Ritchie C, Kaplinski M, Khurana PS, et al. Mobilizing your medications: an automated medication reminder application for mobile phones and hypertension medication adherence in a high-risk urban population. J Diabetes Sci Technol 2013 May 01;7(3):630-639 [FREE Full text] [doi: 10.1177/193229681300700307] [Medline: 23759395]
29. Moorhead P, Zavala A, Kim Y, Virdi NS. Efficacy and safety of a medication dose reminder feature in a digital health offering with the use of sensor-enabled medicines. J Am Pharm Assoc (2003) 2017 Mar;57(2):155-161.e1. [doi: 10.1016/j.japh.2016.12.067] [Medline: 28159505]
30. Santo K, Richtering SS, Chalmers J, Thiagalingam A, Chow CK, Redfern J. Mobile phone apps to improve medication adherence: a systematic stepwise process to identify high-quality apps. JMIR Mhealth Uhealth 2016 Dec 02;4(4):e132 [FREE Full text] [doi: 10.2196/mhealth.6742] [Medline: 27913373]
31. Santo K, Singleton A, Chow CK, Redfern J. Evaluating reach, acceptability, utility, and engagement with an app-based intervention to improve medication adherence in patients with coronary heart disease in the MedApp-CHD study: a mixed-methods evaluation. Med Sci (Basel) 2019 Jun 04;7(6):68 [FREE Full text] [doi: 10.3390/medsci7060068] [Medline: 31167489]
32. Hartch CE, Dietrich MS, Lancaster BJ, Stolldorf DP, Mulvaney SA. Effects of a medication adherence app among medically underserved adults with chronic illness: a randomized controlled trial. J Behav Med 2024 Jun 21;47(3):389-404 [FREE Full text] [doi: 10.1007/s10865-023-00446-2] [Medline: 38127174]
33. Hartch CE, Dietrich MS, Stolldorf DP. Effect of a medication adherence mobile phone app on medically underserved patients with chronic illness: preliminary efficacy study. JMIR Form Res 2023 Dec 11;7:e50579 [FREE Full text] [doi: 10.2196/50579] [Medline: 38079192]
34. Morawski K, Ghazinouri R, Krumme A, McDonough J, Durfee E, Oley L, et al. Rationale and design of the medication adherence improvement support app for engagement-blood pressure (MedISAFE-BP) trial. Am Heart J 2017 Apr;186:40-47. [doi: 10.1016/j.ahj.2016.11.007] [Medline: 28454831]

XSL•FO
RenderX

35. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
36. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: Building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
37. Rolnick SJ, Pawloski PA, Hedblom BD, Asche SE, Bruzek RJ. Patient characteristics associated with medication adherence. Clin Med Res 2013 Jun 11;11(2):54-65 [FREE Full text] [doi: 10.3121/cmr.2013.1113] [Medline: 23580788]
38. Anderson-Lewis C, Darville G, Mercado RE, Howell S, Di Maggio S. mHealth technology use and implications in historically underserved and minority populations in the united states: systematic literature review. JMIR Mhealth Uhealth 2018 Jun 18;6(6):e128 [FREE Full text] [doi: 10.2196/mhealth.8383] [Medline: 29914860]
39. Park JYE, Li J, Howren A, Tsao NW, De Vera M. Mobile phone apps targeting medication adherence: quality assessment and content analysis of user reviews. JMIR Mhealth Uhealth 2019 Jan 31;7(1):e11919 [FREE Full text] [doi: 10.2196/11919] [Medline: 30702435]
40. Alhewiti A. Adherence to long-term therapies and beliefs about medications. Int J Family Med 2014 Feb 13;2014:479596-479598 [FREE Full text] [doi: 10.1155/2014/479596] [Medline: 24688792]
41. Horne R, Hankins M, Jenkins R. The Satisfaction with Information about Medicines Scale (SIMS): a new measurement tool for audit and research. Qual Health Care 2001 Sep 01;10(3):135-140 [FREE Full text] [doi: 10.1136/qhc.0100135] [Medline: 11533420]
42. DiMatteo MR. Social support and patient adherence to medical treatment: a meta-analysis. Health Psychol 2004 Mar;23(2):207-218. [doi: 10.1037/0278-6133.23.2.207] [Medline: 15008666]
43. Molloy GJ, Perkins-Porras L, Bhattacharyya MR, Strike PC, Steptoe A. Practical support predicts medication adherence and attendance at cardiac rehabilitation following acute coronary syndrome. J Psychosom Res 2008 Dec;65(6):581-586. [doi: 10.1016/j.jpsychores.2008.07.002] [Medline: 19027448]
44. Scheurer D, Choudhry N, Swanton KA, Matlin O, Shrank W. Association between different types of social support and medication adherence. Am J Manag Care 2012 Dec 01;18(12):e461-e467 [FREE Full text] [Medline: 23286676]
45. Kingod N, Cleal B, Wahlberg A, Husted GR. Online peer-to-peer communities in the daily lives of people with chronic illness: a qualitative systematic review. Qual Health Res 2017 Jan 12;27(1):89-99. [doi: 10.1177/1049732316680203] [Medline: 27956659]
46. van Uden-Kraan CF, Drossaert CHC, Taal E, Shaw BR, Seydel ER, van de Laar MAFJ. Empowering processes and outcomes of participation in online support groups for patients with breast cancer, arthritis, or fibromyalgia. Qual Health Res 2008 Mar 01;18(3):405-417. [doi: 10.1177/1049732307313429] [Medline: 18235163]
47. Stolldorf DP, Jones AB, Miller KF, Paz HH, Mumma BE, Danesh VC, et al. Medication discussions with patients with cardiovascular disease in the emergency department: an opportunity for emergency nurses to engage patients to support medication reconciliation. J Emerg Nurs 2023 Mar;49(2):275-286 [FREE Full text] [doi: 10.1016/j.jen.2022.12.002] [Medline: 36623969]
48. Serving Patients with Limited English Proficiency: Results of a Community Health Center Survey. National Association of Community Health Centers. 2013 Jul 23. URL: https://healthlaw.org/resource/serving-patients-with-limited-english-proficiency-results-of-a-community-healt/ [accessed 2024-12-26]

Abbreviations

FQHC: federally qualified health center
HCP: health care provider
PI: principal investigator
RCT: randomized controlled trial

Edited by A Kushniruk; submitted 25.06.24; peer-reviewed by P Cook, C Chiaranai; comments to author 11.09.24; revised version received 17.10.24; accepted 22.12.24; published 07.01.25.

Please cite as:
Hartch C, Dietrich MS, Lancaster BJ, Mulvaney SA, Stolldorf DP
Satisfaction and Usability of a Commercially Available Medication Adherence App (Medisafe) Among Medically Underserved Patients With Chronic Illnesses: Survey Study
JMIR Hum Factors 2025;12:e63653
URL: https://humanfactors.jmir.org/2025/1/e63653
doi:10.2196/63653
PMID:

XSL•FO
RenderX

©Christa Hartch, Mary S Dietrich, B Jeanette Lancaster, Shelagh A Mulvaney, Deonni P Stolldorf. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 07.01.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluating the Usability, Acceptability, User Experience, and Design of an Interactive Responsive Platform to Improve Perinatal Nurses' Stigmatizing Attitudes Toward Substance Use in Pregnancy: Mixed Methods Study

Michael Rubyan[1,2*], MPH, PhD; Yana Gouseinov[3*], BS; Mikayla Morgan[3*], BSN, RN, CCRN; Deborah Rubyan[4*], MPA; Divya Jahagirdar[1*], BDS, MHI; David Choberka[5*], PhD; Carol J Boyd[6,7*], PhD, MSN, RN; Clayton Shuman[2,3,6*], PhD, MSN, RN

[1]Department of Health Management and Policy, School of Public Health, University of Michigan, Ann Arbor, MI, United States

[2]Institute for Healthcare Policy and Innovation, University of Michigan, Ann Arbor, MI, United States

[3]Department of Systems, Populations, and Leadership, School of Nursing, University of Michigan, Ann Arbor, MI, United States

[4]Honey Locust Health, West Bloomfield, MI, United States

[5]Museum of Art, University of Michigan, Ann Arbor, MI, United States

[6]Center for the Study of Drugs, Alcohol, Smoking and Health, University of Michigan, Ann Arbor, MI, United States

[7]Women's and Gender Studies, University of Michigan, Ann Arbor, MI, United States

[*]all authors contributed equally

Corresponding Author:
Michael Rubyan, MPH, PhD
Department of Health Management and Policy
School of Public Health
University of Michigan
1415 Washington Heights
M3517 SPH2
Ann Arbor, MI, 48109
United States
Phone: 1 734 936 6331
Email: mrubyan@umich.edu

Abstract

Background: Perinatal nurses are increasingly encountering patients who have engaged in perinatal substance use (PSU). Despite growing evidence demonstrating the need to reduce nurses' stigmatizing attitudes toward PSU, limited interventions are available to target these attitudes and support behavior change, especially those reflecting the overwhelming evidence that education alone is insufficient to change practice behavior. Arts-based interventions are associated with increasing nursing empathy, changing patient attitudes, improving reflective practice, and decreasing stigma. We adapted ArtSpective for PSU—a previously evaluated, in-person, arts-based intervention to reduce stigma toward PSU among perinatal nurses—into an interactive, digital, and responsive platform that facilitates intervention delivery asynchronously.

Objective: This study aimed to evaluate the usability, acceptability, and feasibility of the interactive, responsive platform version of ArtSpective for PSU. Our goal was to elicit the strengths and weaknesses of the responsive platform by evaluating the user experience to identify strategies to overcome them.

Methods: This study used a mixed methods approach to explore the platform's usability, user experience, and acceptability as an intervention to address stigma and implicit bias related to PSU. Theatre testing was used to qualitatively assess usability and acceptability perspectives with nurses and experts; a modified version of the previously validated 8-item Abbreviated Acceptability Rating Profile was used for quantitative assessment. Quantitative data for acceptability and satisfaction were analyzed using descriptive statistics. All qualitative data were analyzed iteratively using an inductive framework analysis approach.

Results: Overall, 21 nurses and 4 experts in stigma, implicit bias, and instructional design completed theatre-testing sessions. The mean duration of interviews was 31.92 (SD 11.32) minutes for nurses and 40.73 (SD 8.57) minutes for experts. All participants indicated that they found the digital adaptation of the intervention to be highly acceptable, with mean acceptability items ranging

XSL•FO
RenderX

from 5.0 (SD 1.0) to 5.5 (SD 0.6) on a 1-6 agreement scale. Nurses reported high satisfaction with the platform, with mean satisfaction items ranging from 5.14 (SD 0.56) to 5.29 (SD 0.63) on a 1-6 agreement scale. In total, 1797 interview segments were coded from the theatre-testing sessions with 4 major themes: appearance, navigation, characterization, and overall platform, and 16 subthemes were identified. Consistent with the quantitative findings, the results were positive overall, with participants expressing high satisfaction related to the platform's appearance, the ease with which they could navigate the various modules, engagement, clarity of the presentation, and feasibility of being completed asynchronously.

Conclusions: Developing and evaluating the usability of a digital adaptation of ArtSpective for PSU resulted in strong support for the usability, acceptability, and satisfaction of the program. It also provided insight into key aspects related to acceptability and usability that should be considered when designing a digital adaptation of an arts-based intervention for health care providers.

(JMIR Hum Factors 2025;12:e67685) doi:10.2196/67685

KEYWORDS

digital platform; nurse training; usability; user experience; user interface; mixed methods; theatre testing; stigma; perinatal substance use; art intervention; acceptability; perinatal; substance use; pregnancy; perinatal nurse; feasibility; interactive responsive platform

Introduction

Background

Stigmatizing attitudes among nurses who treat patients with perinatal substance use (PSU) [1-7] is a major barrier to the implementation of evidence-based practices for maternally provided care (ie, breastfeeding and skin-to-skin contact) that improve infant outcomes; decrease symptom severity; and reduce pharmacologic treatment, length of stay, and costs [8-18]. Despite growing evidence that demonstrates the need to reduce nurses' stigmatizing attitudes toward PSU [17-20], limited interventions are available that target these attitudes and support behavior change among nurses. Interventions that have been studied use didactics to increase awareness about the need to reduce stigma and increase compassion [20]. However, these interventions do not reflect the overwhelming evidence that education alone is not sufficient to change practice behavior [21,22] and that stigmatized attitudes may prevent the application of education to individual practice [23]. Evidence-based interventions targeting stigmatizing attitudinal change are necessary to empower nurses to effectively reduce biases and judgments toward PSU and care for these patients.

One of the key weaknesses of traditional interventions in addressing stigma is that they are generally structured around passive learning and focus on knowledge deficits rather than intervening on stigma itself [24]. In contrast, increasing research demonstrates how arts-based interventions improve clinical skills and communication between health care providers, especially for stigmatized attitudes [25]. Arts-based pedagogy is associated with increasing nursing empathy, changing stigmatized attitudes toward patients, and improving reflective practice [26]. Photography and digital storytelling have been demonstrated to decrease stigma and increase provider willingness to help patients affected by substance use disorder [27]. In addition, creative storytelling has been effectively used to improve clinician attitudes toward persons with dementia, a highly stigmatized population [28].

ArtSpective for PSU: An Arts-Based Intervention

Overview

Based on existing evidence of efficacy for arts-based interventions to improve clinician attitudes toward stigmatized patient populations [29,30], we developed ArtSpective for PSU to facilitate perspective taking and change attitudes toward PSU among nurses to facilitate their engagement with mothers and promote evidence-based practice that is not impeded by biases and stigma. ArtSpective for PSU was initially designed as a synchronous intervention that is delivered in one 60-minute session to a small group of nurses through live videoconference [24,31-33]. The intervention is grounded in art pedagogy previously used by art museum educators and facilitates a perspective-taking exercise using curated fine-art photographs and guided narrative storytelling. Participants are shown images and instructed to select two to interact with for the intervention. Selected images use a documentary photographic style characterized by ambiguity in subject matter, which enables participants to write two short creative narratives from two different perspectives (one perspective per image): (1) the perspective of a subject in the photograph and (2) the outside perspective of the nurse. Perspective taking is described as taking the point of view of another person within the other's context. This is grounded in evidence that this type of contextualization is associated with taking actions and exerting behaviors to alleviate distress and improve outcomes [34]. Participants spend time discussing their narratives with partners during the synchronous session after writing a narrative about the photographs they reviewed. They also engage in facilitated dialogue related to the perspective-taking attitudes elicited during their participation. These conversations enabled participants to compare their responses to the photographs and notice when their interpretations displayed biased attitudes.

While participants previously reported high demand, satisfaction, and efficacy with the synchronous version of ArtSpective for PSU [24], they identified key barriers to scaling the intervention across hospitals and health systems, given its synchronous and in-person modality. These barriers are consistent with several implementation barriers that prevent engagement with arts-based interventions, including fluctuation

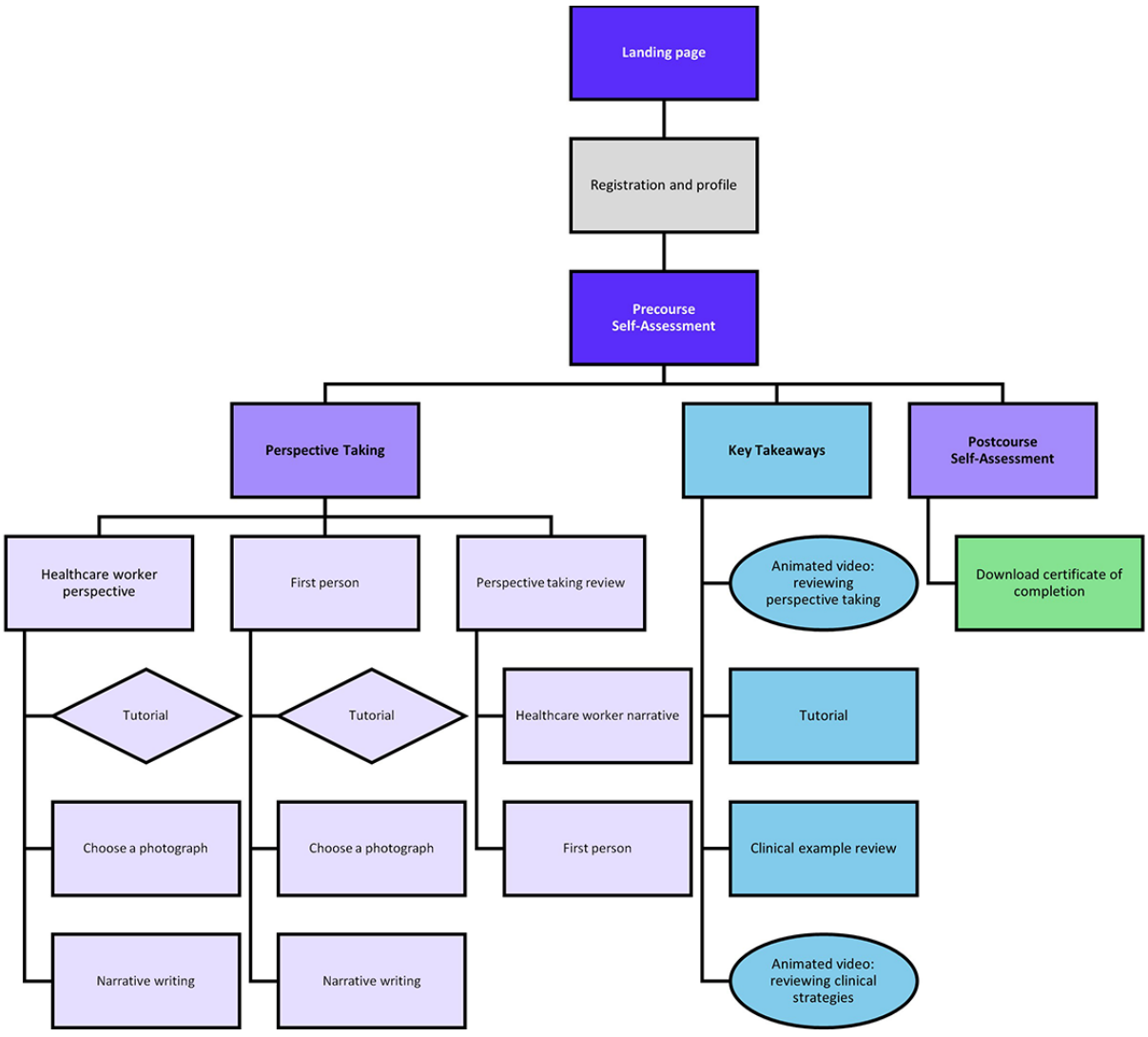
XSL•FO
RenderX

in participation, skepticism of efficacy, and obstacles related to delivery, especially in person [35]. Given these barriers, we adapted ArtSpective for PSU into an interactive, responsive platform that facilitates the delivery of the intervention asynchronously. Responsive platforms and digital delivery of interventions have been demonstrated to improve uptake and overcome implementation barriers related to scalability, especially for interventions designed for nurses [36-42].

Design and Development of Digital Intervention

A multidisciplinary team, including experts in nursing, labor and delivery, health informatics, health communication, and implementation science, developed the interactive responsive platform. Before developing the platform, we conducted 2 focus groups related to the synchronous version of the intervention (n=11) to identify specific requirements for the interactive, responsive platform. Participants noted the need for the adapted intervention to (1) take less time to complete, (2) allow for self-pacing, and (3) be delivered asynchronously [24]. Partnering with a public health communications firm, the synchronous intervention components were adapted into an interactive, responsive platform optimized for desktop, smartphone, or tablet interface. The platform incorporated electronic art exhibits and facilitation materials consisting of a 30-minute, web-based session that nurses could complete asynchronously (Figure 1). The platform design is grounded in the user experience design methodology, which is anchored in principles related to understanding users, user participation throughout the process of development, user-centered evaluation, iteration, a reflection of the entire user experience, and a multidisciplinary perspective to design and development [43,44]. The overall design is guided by Material 3, an iteration of Google Material Design, an open-source adaptable system of guidelines, components, and tools that support best practices of user interface design [45]. The platform also uses structured headings, alternative text descriptions, bold fonts, high-contrast color combinations, and captions to ensure accessibility across all users.

Figure 1. ArtSpective for perinatal substance use (PSU) platform structure.



Before the study's launch, the research team and health communications firm developed an interactive prototype of the digital adaptation of ArtSpective for PSU in Figma (Adobe). A panel of experts with extensive experience in perinatal nursing, PSU, art pedagogy, implementation science, and user interface and experience design examined all prototype elements in an iterative process as part of a heuristic evaluation. This process consisted of identifying platform functionalities individually

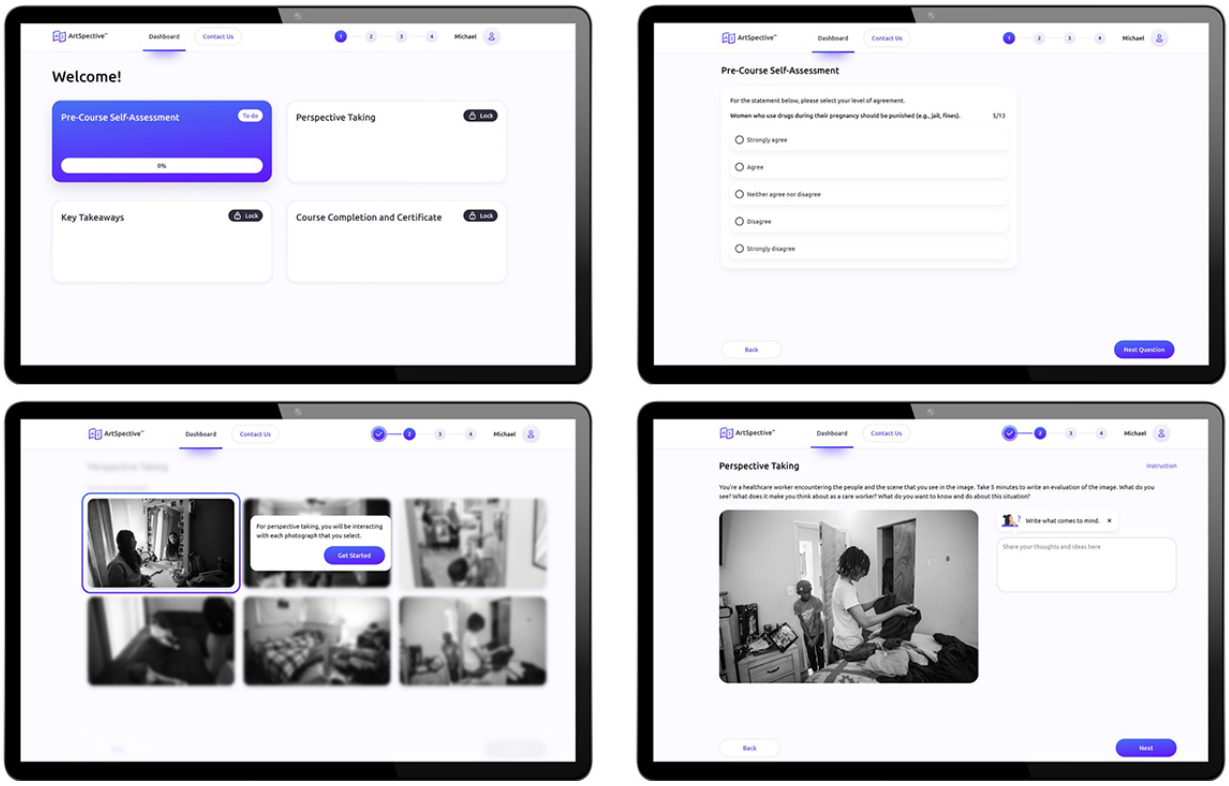
XSL•FO
RenderX

based on the list of heuristics for usable interfaces [46,47]. The health communications firm iteratively refined the prototype as user interface issues were identified.

Platform Flow

At the platform launch, users are asked to register for an account or log into their existing account. Following account creation, users complete a series of demographic questions as part of creating their profile. Once a profile has been completed, the user lands on a dashboard with the four core modules that comprise the intervention: (1) *Pre-Course Self-Assessment*, (2) *Perspective Taking*, (3) *Key Takeaways*, and (4) *Course Completion and Certificate*. Modules are displayed as symmetrical cards with large rounding on each card. Small, rounded badges are included on each module card to indicate if it has a status of "To Do," "Locked," or "Completed." Navigation is contained at the top of the interface, along with a module progress bar that uses numbers and colors to indicate progression through the platform (Figure 2).

Figure 2. Sample of the dashboard, along with the Pre-Course Self-Assessment and Perspective Taking modules, in the adapted intervention.

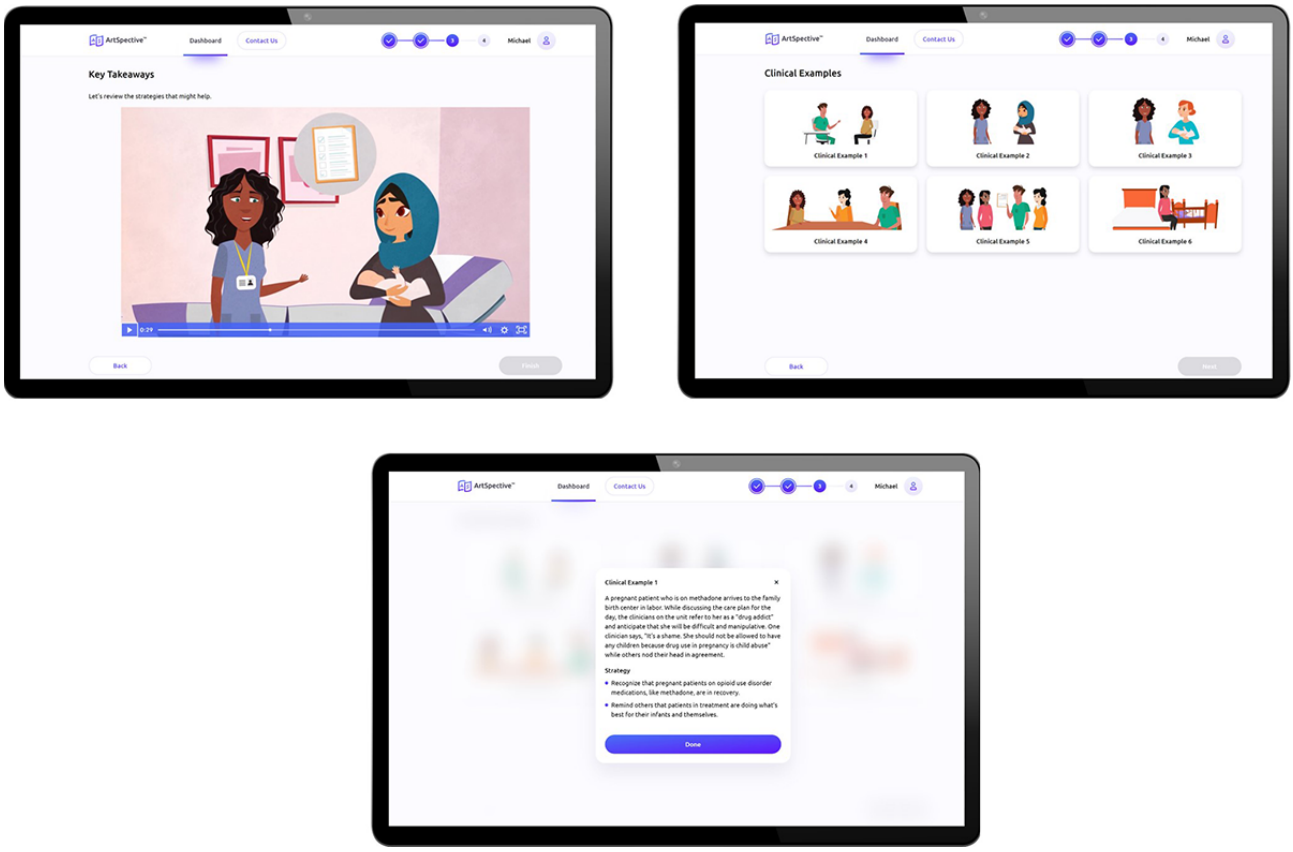


In the Perspective Taking module (Figure 2), participants are presented with 6 rounded symmetrical cards containing photographs to select from for the exercise. Before choosing a photograph, the user is prompted to complete a brief guided tutorial and then prompted to imagine that they are a health care worker encountering the scene they see in the image and directed to type a message into a textbox consisting of an evaluation of the image from their health care worker perspective. Following this, the participant is presented with the same 6 photographs (the first photograph they chose has a reduced opacity to indicate that it can no longer be selected). The participant is prompted to type a message into a different textbox that evaluates the second image from their personal perspective. After entering their messages, the participant sees each photograph they chose and can read the messages they typed without editing them.

In the Key Takeaways module (Figure 3), participants watch a 1-minute animated video depicting the perspective flipping exercise through the actions of animated characters, exploring how to limit assumptions in patient care through the lens of how the participant's health-care-worker perspective may have been different from their personal perspective in the exercise. After watching the video, participants review 6 rounded symmetrical cards that contain clinical examples and strategies illustrated with the characters from the animated video and watch a 90-second video articulating the core clinical examples and techniques that users reviewed in the previous phase to reinforce the content. On average, participants spend 10 minutes engaging with this module.

XSL•FO
RenderX

Figure 3. Sample of the video displayed in the Key Takeaways module and the Clinical Examples page presented in the adapted intervention.



Before evaluating the effectiveness of a digital adaptation, it is critical to examine its use and sustainability [48,49], including usability, acceptability, and feasibility with end users. Usability is characterized by how well specific users achieve a predetermined goal with efficiency and satisfaction in a particular context [50]. Acceptability relates to how end users think and feel about a digital intervention, while feasibility relates to how well it can be deployed with end users [51-53].

Purpose

This study aimed to evaluate the usability, acceptability, and feasibility of the interactive, responsive platform version of ArtSpective for PSU—designed to deliver an adapted version of the in-person, arts-based, perspective-taking intervention to improve nurses' stigmatizing attitudes toward PSU—following a mixed methods approach. Our goal was to elicit the strengths and weaknesses of the responsive platform by evaluating the user experience to identify strategies that overcome the weaknesses.

Methods

Study Design

The evaluation of the platform's usability, acceptability, and feasibility followed a mixed methods approach, combining target user experience evaluation facilitated through Zoom (Zoom Communications) videoconference and a validated measure of acceptability.

Procedure

Study Population and Recruitment

A purposive sample of nurses was invited to participate through an email sent to all nurses meeting eligibility criteria employed by a large, urban academic medical center in Chicago with high rates of birthing individuals with opioid use disorder and infants with neonatal opioid withdrawal syndrome. This was followed by snowball sampling in both community hospitals and academic health centers in Michigan and Chicago. Inclusion criteria consisted of (1) being credentialed as a registered nurse, (2) actively practicing in a perinatal unit (labor, postpartum, and neonatal intensive care unit), and (3) being at least 18 years old. In addition, a convenience sample of experts in instructional design, stigma, and implicit bias was recruited to participate. The target sample size was 24 nurses and 4 experts.

Methodology

We used theatre testing to qualitatively explore nurses' and experts' usability and acceptability perspectives. Theatre testing is a methodology that is integrated into the Assessment Decision Administration Production Topical Experts-Integration Training Testing (ADAPT-ITT) model and designed to facilitate the adaptation of evidence-based interventions [54]. It is a type of pretesting methodology that has been previously used to test messaging, such as television or print advertisements, with a key target demographic [55]. Participants who reflect the intended audience of the intervention are asked to respond to a demonstration of an intervention and answer questions designed to elicit their reactions to it [54]. A key strength of this methodology is its ability to rapidly gather feedback about a

XSL•FO
RenderX

prototype that reflects the intervention as end users would experience it. We modified the 8-item Abbreviated Acceptability Rating Profile (AARP) [56], a previously validated and widely used acceptability instrument, for quantitative acceptability assessment. The abbreviated nature of the instrument was designed to overcome barriers related to time-intensiveness. Modifications to the AARP for this study included specific references to perinatal substance. We complemented the instrument with a separate stand-alone set of 5 items related to satisfaction that were used previously when conducting concept testing of the synchronous intervention [24].

Theatre testing for qualitative evaluation provides a deeper understanding of users' perspectives on interface and platform design and individual user experience. Simultaneously, the quantitative data derived from the validated scale and our supplemental satisfaction questions provide precise measurements of acceptability and feasibility among nurses and experts. The goal was to explore both the usability and user experience of the platform and its acceptability and feasibility as an intervention to address stigma and implicit bias related to the treatment of PSU.

Data Collection

A total of 2 research team members (MR and CS) designed the theatre-testing session interview guide, which was piloted for presentation and clarity of information and workflow (Multimedia Appendix 1). MR (male) has a PhD and MPH. He is a clinical assistant professor with experience and training in implementation science and provider behavior change. CS (male) has a PhD, MSN, and RN. He is an assistant professor with training and experience in implementation science and care for vulnerable populations, including individuals with PSU and postpartum women. This study is reported according to the COREQ (Consolidated Criteria for Reporting Qualitative Research) guidelines (Multimedia Appendix 2) [57]. Participants were invited by email to individual sessions facilitated through Zoom videoconference and were informed that the researchers leading the study were developing a new platform to promote a behavior change intervention. Sessions lasted between 30 and 45 minutes between December 2022 and October 2023. Each session was conducted by one of two analysts on the research team (YG) who have extensive expertise in qualitative interviewing and have clinical experience caring for patients with substance use disorder. The analysts did not share their assumptions or motivations for facilitating the study with participants. The study procedure was explained to participants after they had logged into the Zoom session, and participants were informed that the interview would be video recorded and transcribed. Interviewers used the screen-sharing feature in Zoom to display the interactive prototype and clicked through the screens and prompts. To understand participant reactions to the platform, the concurrent think-aloud method was used by asking participants to verbalize what they noticed or did not understand [58]. No training or contextualization of the prototype were offered before the interview began to avoid biases related to preparation. Participants were given minimal assistance in understanding what was displayed, and interviewers only interjected with contextual information to encourage them to continue reflecting on what they saw. Following the demonstration, the interviewer debriefed each participant, using a semistructured interview guide that assessed overall user experience, design of the content in the platform, and ease of navigation, and performed a brief retrospective analysis of issues that participants noted during the demonstration [59]. Participants also completed a questionnaire using Qualtrics XM hosted at the University of Michigan that included questions about demographics, the AARP, and stand-alone satisfaction questions.

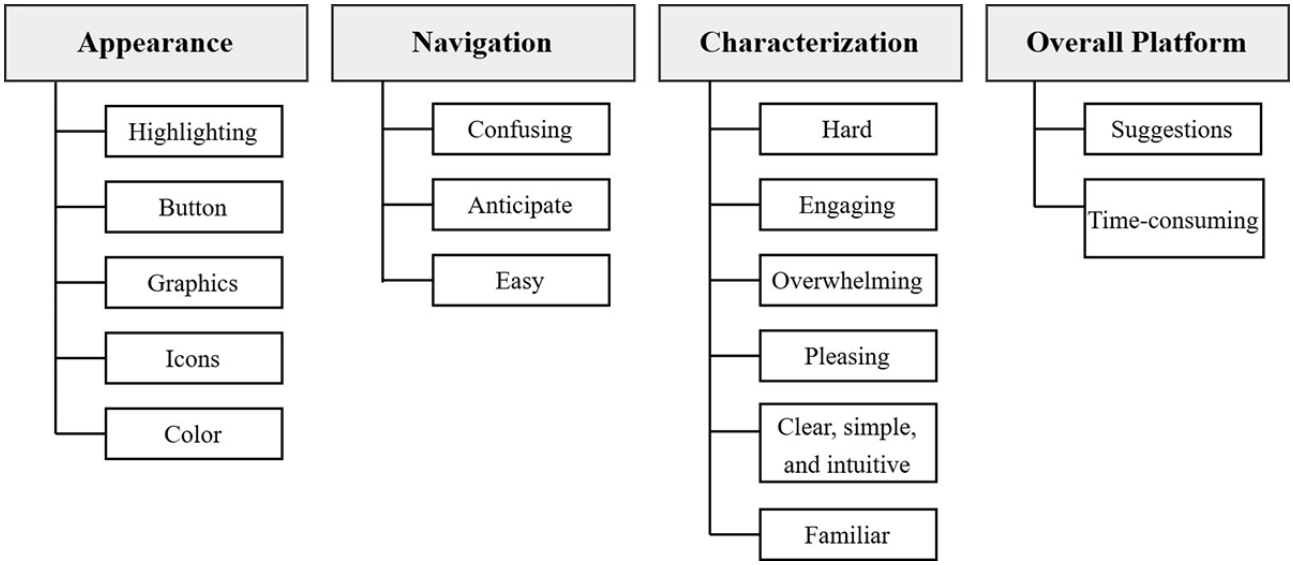
Nonparticipants did not attend the sessions, and no relationship with participants was established before they engaged with the study. No repeated interviews were conducted, and no field notes were taken. Sessions were recorded through Zoom Recordings, pseudonymized, and transcribed verbatim. Transcripts were not returned to participants for their review. Data were transcribed by Rev. Data were analyzed in MAXQDA (version 22.8.0; VERBI GmbH), a computer-assisted qualitative data analysis software. The research team met after 21 interviews with nurses to discuss data saturation and information power. Since no new themes emerged in later interviews, the research team concluded that information power had been reached, and no additional interviews were required. In addition, following the 4 interviews with experts, the research team met separately to discuss information power and concluded that since all 4 experts had provided similar feedback, information power had been reached.

Data Analysis

All data, including pilot sessions, were included in the final data analysis. Pilot sessions were included, as no changes were made to the interview guide or protocol after the pilot sessions were conducted. We conducted separate quantitative and qualitative data analyses. All quantitative data regarding acceptability and satisfaction were checked for completeness, and analyses were performed using descriptive statistics (mean, SD, frequency, and percentage) to describe demographic characteristics, responses to the modified AARP, and satisfaction items. All qualitative data were analyzed iteratively, using an inductive framework analysis approach [60]. Three multidisciplinary research team members with backgrounds in implementation science, perinatal nursing, and qualitative research methods (MR, YG, and MM) read the transcripts to inductively identify initial sets of codes. The team then met to discuss codes and map them to a coding schema, resulting in a codebook based on overarching themes. In addition, 2 research team members (YG and MM) independently coded segments in each transcript, meeting regularly with the entire research team to discuss discrepancies and refine or modify the codebook as necessary. This resulted in an iteratively developed, more detailed codebook. All themes were organized into major themes and subthemes that reflected quotes from the interview transcripts (Figure 4). Participants did not provide feedback on the findings.

XSL•FO
RenderX

Figure 4. Overview of overall themes and subthemes.



Ethical Considerations

The study was deemed exempt by the University of Michigan institutional review board (HUM00223937). Informed consent was obtained from all participants. Data have been deidentified. Participants received a US $30 cash card for their time.

Results

Overview

A total of 21 nurses and 4 experts completed theatre-testing sessions. No participants refused to participate. The mean duration of interviews was 31.92 (SD 11.32) minutes for nurses and 40.73 (SD 8.57) minutes for experts. Table 1 shows the demographic data of the participants. The majority of participants identified as female (22/25, 88%), and nearly half (12/25, 48%) of the participants were between 18 and 34 years old. Nearly half (10/21, 48%) of nurses practiced in the neonatal intensive care unit. Years of experience among nurses were equally divided across the three categories (<5, 5-10, and >10 years).

XSL•FO
RenderX

Table 1. Demographic characteristics of study participants (N=25).

Demographic characteristics	Value, n (%)
Sex	
Male	2 (8)
Female	22 (88)
Other	1 (4)
Race	
White	16 (64)
African American	3 (12)
Asian	4 (16)
Other	2 (8)
Ethnicity	
Hispanic	2 (8)
Not Hispanic	23 (92)
Age category (years)	
18-34	12 (48)
>35	13 (52)
Unit type (nurses only; n=21)	
NICU[a]	10 (48)
Labor and delivery	6 (28)
Postpartum	5 (24)
Years in practice (nurses only; n=21)	
<5	7 (33)
5-10	7 (33)
>10	7 (33)

[a]NICU: neonatal intensive care unit.

Quantitative Measures

All participants indicated that they found the digital adaptation of the intervention to be highly acceptable, with mean acceptability items ranging from 5.0 (SD 1.0) to 5.5 (SD 0.6) on a 1-6 agreement scale (Table 2). Nurses reported high satisfaction with the platform, with mean satisfaction items ranging from 5.14 (SD 0.56) to 5.29 (SD 0.63) on the same agreement scale (Table 3). Of note, after completing the intervention, participants recognized the severity of clinician stigma toward PSU; they recognized the intervention is needed to better understand and demonstrate compassion for mothers and infants affected by PSU.

XSL•FO
RenderX

Table 2. Acceptability of the intervention among all participants using the Abbreviated Acceptability Rating Profile (N=25).

Item	Score[a], mean (SD)
This is an acceptable program for addressing clinician stigma towards perinatal substance use.	5.20 (0.57)
This program should be effective in reducing clinician stigma towards perinatal substance use.	5.04 (0.79)
Clinician stigma toward perinatal substance use is severe enough to justify this program.	5.54 (0.64)
I would recommend this program to others.	5.00 (1.00)
This program would not have bad side effects on the clinician participant.	5.50 (0.65)
I liked this program.	5.00 (1.00)
This program is a good way to handle clinician stigma toward perinatal substance use.	5.08 (0.74)
Overall, the program would help clinicians.	5.24 (0.59)

[a]Items were rated on a 1- to 6-point agreement scale, with higher scores indicating greater levels of acceptability.

Table 3. Nurses' satisfaction with the intervention (n=21).

Item	Score[a], mean (SD)
The overall program was beneficial to me.	5.14 (0.56)
After completing the program, I have a better understanding of compassion for mothers and infants affected by perinatal substance use.	5.29 (0.63)
After completing the program, I have a better understanding of maternal agency related to caring for mothers and infants affected by perinatal substance use.	5.14 (0.64)
Because of this program, I will be more compassionate to mothers and infants affected by perinatal substance use.	5.25 (0.54)
This program was unique, unlike other training about stigma toward perinatal substance use.	5.19 (0.66)

[a]Items were rated on a 1- to 6-point agreement scale, with higher scores indicating greater satisfaction.

Qualitative Measures

Interviews and Themes

A total of 1797 interview segments were coded from the theatre testing sessions. A total of 4 major themes and 16 subthemes were identified. Major themes included appearance, navigation, characterization, and overall platform. A summary of themes is contained in Figure 4.

Appearance

The first major theme was related to the appearance of the platform and different aspects of the visual user experience and design. We identified five subthemes of appearance: (1) highlighting, (2) buttons, (3) graphics, (4) icons, and (5) color. Participants reflected on how highlighting was used in the navigation bar to indicate progression, module completion, and photograph selection during the Perspective Taking module. They observed that highlighted buttons helped them with navigation. They also noted that highlighting in the navigation bar provided a clear indication of their progression and appreciated the use of visual elements to denote locked or unlocked modules, noting the strength of the visual design of the platform. Participants frequently compared it with other training modules where highlighting as a visual cue was not present, which led to a lack of motivation by the user to complete a different training.

> *…the next button is literally highlighted when you're next and it literally walks you right through it and I think that's super important for courses like this because sometimes it's not easy, people are going to give up and just get right through it and not even complete it as they're supposed to.* [Participant 107]

Participants also appreciated that highlighting related to locked photos was used during the Perspective Taking module to help them know where to go next.

Participants noted the use of buttons throughout the platform and how they served as visual cues for navigation. They noticed that illuminating the buttons and giving them specific colors provided the signal to move to the next step.

> *I liked the visual reminders and cues, having the little padlock sign that's not where you are now. Good use of color and contrast for the buttons to say next or completed.* [Participant 106]

The "Next" button also encouraged them not to give up before completing the course. They also mentioned that the modal window containing a blue button with a check mark helped them to know that the course had concluded and that they could download a certificate of completion. Overall, the buttons served as visual cues for the participants, and the colors made them easy to understand.

XSL•FO
RenderX

Participants observed the strengths of how the text was formatted and sized, the arrangement of content, the use of animated characters, and photography. While most participants noted specific strengths related to monochromatic colors for all graphical elements, some participants pointed out that this seemed like too much of one color, making it more difficult to find specific elements.

> *I would maybe emphasize the woman as a different color. It's like the numbers are that color, the Continue button is that color and the woman is that color. I feel like it's a little overload on that color.* [Participant 119]

One suggestion they made was that the characters in the Key Takeaways module could be easier to understand if they were in full color. In addition, participants commented that the colorization of the graphics helped them understand what was coming next.

Most participants found the typography and the design of iconography and color made it easier to navigate overall. However, some noted that the overwhelming use of blue icons, buttons, and graphics made it difficult to see where to click next. Participants also indicated that icons in the menu choices helped to prompt them on what they needed to complete. They commented that the checkmarks (once completed) that appeared next to each clinical example in the Key Takeaways module helped clarify what needed to be reviewed before they could continue.

> *The same screen as before, just the first one has a checkbox now, so I would assume to continue through until they all get a checkbox. And currently the finish is grayed out, so I assume I still have something to do on the screen before I can move on.* [Participant 111]

Participants also noted that check mark icons in the menu choices helped guide them on what was completed and what still needed to be reviewed. The participants generally found the icons helpful and appealing, especially when combined with bolding and bulleted list formatting.

> *I think definitely the usability of this was a lot more simpler. I feel like other courses that I've taken, it's a little bit outdated and it's a little bit harder to navigate. I like big fonts and all this stuff. It's just nice. It's like bubbles and push, click, next, get started. That was really nice.* [Participant 103]

Finally, some participants observed that using the large checkmark icon at the end of the training was unclear, noting that they were unsure whether to save the certificate of completion or view it. However, they also mentioned that since the only choice was to click on the page, it led them to download the certificate, which clarified the user experience.

Overall, participants commented that the lavender color was a calming element in the platform. They also found the white background and blue colors appealing. Participants mentioned that the color scheme helped them to navigate, pointing out that the grayed-out objects that appeared not to be clickable helped to indicate how the platform progressed. They also noted their satisfaction with how the colored buttons guided participants to click through the different parts of each module.

> *I think the content is pretty calming, giving the lavender colors and just the very light design in general. There isn't too much going on, on the page, so it's not very distractive, and I could see where my tasks are.* [Participant 102]

Although some participants stated that there was too much blue and purple, others felt that a similar color was helpful for consistent navigation. In general, participants remarked that the consistent blue color for all the buttons made it clear where to click next and helped them know how to navigate.

Navigation

The second major theme was related to the user experience surrounding the platform's navigation and how users were able to interpret navigational elements. We identified three subthemes of navigation: (1) confusing, (2) anticipate, and (3) easy.

Overall, participants remarked that the platform was not confusing. However, some participants commented that tutorial features were overly prescriptive, which confused them about whether they were still in the tutorial or were supposed to begin their reflection in the Perspective Taking module.

> *I will say there were times I wasn't exactly sure if I had completed something, or I guess I should say some things were somewhat repetitive, like when I click the picture and then it pops up again and then it goes away, but then I have to click it again. That was kind of the only thing.* [Participant 117]

Participants reported that they knew what to do after each step and that the platform's visual cues helped them anticipate the next step. Overall, they observed that the step-by-step process, including check marks to indicate what was completed, the "Next" buttons that were illuminated when they were activated, and the grayed-out imagery to indicate what could not be selected, helped to guide them on what they needed to do next.

> *Think it was pretty easy to navigate, because once you complete something, it shows you that you've completed it. What else you need to do, there's buttons that says you need to do that. And once you complete a task, the next button or the finish button illuminates, then I feel that's your clue that you could move on.* [Participant 101]

Participants described the visual cues as being very helpful in making navigation easy. They also observed that the signup and log-in features were simple and easy to navigate.

> *…literally for each section, it would explain what the section is, and then in order to continue, you would have to either fill something out, or click the highlighted button.* [Participant 102]

Characterization

The third major theme was related to how participants characterized the platform across engagement, clarity, familiarity, and difficulty. We identified six subthemes of

XSL•FO
RenderX

characterization: (1) hard; (2) engaging; (3) overwhelming; (4) pleasing; (5) clear, simple, and intuitive; and (6) familiar.

There were very few instances of participants indicating difficulty, except for the requirement that they would need to respond to questions about substance use when they might not remember a case related to substance use in their caseload. Participants also noted that if they had personal experience with substance use, it would be more difficult to actively participate.

> *I also don't have any personal or family relationships with people with substance use. So I think in that sense, I don't... I wouldn't have as much hesitation or maybe nervousness in going through this module versus somebody who does have a personal experience. It might be a little challenging, perhaps.* [Participant 105]

Participants noted that one of the most engaging aspects of the platform was knowing in advance how many scenarios there would be during the Perspective Taking module and not being able to skip around in the platform.

> *I really liked the format of this. I've done quite a few modules in the last few months, and I wish they all looked like this because I think sometimes the ones that are just you're going through an automatic voiced PowerPoint and then you have a knowledge check every so often can just really drag you down. Anything that allows for more thoughtful learning I think is good. Yeah, I liked it a lot.* [Participant 108]

They were also very receptive to the active learning format that allowed them to practice a skill instead of reviewing didactic material. In addition, they found the activities on the platform incredibly engaging since they related to a topic they were already familiar with.

> *I think a lot of them are rather mundane and bleak and you just feel like, oh, I have to do another module. But this one was just intriguing. I think a lot of that was because of the uniqueness of it. So it didn't make me feel like, oh, geez, another module I have to do. It kept my interest.* [Participant 121]

A common observation among participants was that the monochromatic color scheme sometimes felt overwhelming to process. However, they reflected that this was modulated by the different photographs used. When presented with the Perspective Taking module, some participants also noted that the task of writing a narrative in 5 minutes felt like a significant task that may not be achievable. Finally, some participants shared that the number of "Next" and "Continue" buttons during the tutorials seemed overwhelming.

> *In the tutorial, a lot set up the same. I do feel like this one, there was a lot of "Hit the 'Next' button," and then "Hit the 'Continue' button," whereas if you just hit the "Next" button, could you not... Could you do without some of the "Next," and then "Next," again, button. It seemed like you had to hit the "Next" button at that end a lot. And I was like, I feel like that's redundant. If you hit the "Next" button, you should just be done.* [Participant 112]

There was high satisfaction with the platform layout, presentation clarity, and information flow, and the participants noted that the use of material design principles caused the platform to be visually pleasant.

> *I think the content is pretty calming, giving the lavender colors and just the very light design in general. There isn't too much going on the page, so it's not very distractive, and I could see where my tasks are.* [Participant 101]

Participants also noted that instructions and guidance were unambiguous and made progressing through the different modules easy. They were especially satisfied with how the design allowed them to see the difference between the writing activities in the Perspective Taking module and the didactic material in the Key Takeaways module. There was consensus on the dashboard's clarity about how to progress through the platform, and participants found the signup and sign-in flow to be very intuitive and accessible.

> *There's four distinct sections. They kind of go in an order. Once you finish one, it highlights that you're supposed to move on to the next. So there's not really any way to get kind of lost.* [Participant 105]

Participants noted that the Key Takeaways module structure that used cards, modal windows, and short video segments was clearly organized and easy to follow.

Finally, participants consistently reported that the format was very familiar and similar to other courses they had taken.

> *It reminds me of kind of the courses that we have to take in the hospital when we have to renew or look at our continuing education. It has kind of that similar feel.* [Participant 115]

Overall Platform

The fourth major theme was related to participant observations about the entire platform experience, with a particular focus on their suggestions for improvements and how they would characterize the time commitment required to complete the platform activities. We identified two subthemes: (1) suggestions and (2) time-consuming.

Participants suggested that during sign-up and profile creation, the platform should not request a profile picture, as this is too personal for an educational module connected to their workplace.

> *The profile picture thing, unless it was like pre-selected profile pictures, but I think it would just be kind of funny also if you were just doing an education module for them to ask for a profile picture.* [Participant 116]

In the Pre-Course Self-Assessment module, participants noted that it would be helpful if they were given more notice that they would be prompted to share their personal perspective as part of the training, which may be something not all nurses want to do. During the Perspective Taking module, participant suggestions included reducing the length of the tutorial to avoid redundancy, making the visuals more explicit about what role

XSL•FO
RenderX

they were assuming in each of the narrative writing steps, retaining instructions from one screen to the next to help recall what to do, and adding more arrows and icons to help with users who are less experienced with web-based platforms. In addition, participants suggested some improvements to the exercise that was being facilitated asynchronously in this module. This included a desire to learn other participants' perspectives and eliminate the review of their narratives after they wrote them, since they could not be edited. In the Key Takeaways module, participants suggested that the clinical examples and strategies could be separated more clearly to deliver more independent information. They also thought that these could be more consolidated to help participants remember them.

> *It might be nice if you were like, here, read this clinical example and then click a button and you say, actually, here's the strategies that we recommend for this situation, instead of just having it all on one page. If you separate those, it might be more clear to people this situation actually did not go well. Here's what we want to promote.* [Participant 106]

Finally, some participants suggested that the platform be more explicit about available subtitles for videos so that users who do not have headphones could still review the video content. In the Course Completion and Certificate module, participants suggested that the platform include an instrument that assesses competency related to the learning before providing a certificate, as well as a presentation of statistics and background data about maternal substance use, so that participants could carry more issue-related content with them after participating.

Very few participants noted that the platform was too time-intensive after completing the application session during the interview. Those participants who did observe this aspect said that when using the platform, they might need to take a break after the perspective-taking exercise before proceeding to the Key Takeaways module, since the perspective-taking exercise required writing, and the Key Takeaways module required listening and reading. Even though the total amount of time for the whole session was approximately 30-45 minutes (interview and application session together), they thought that taking a break between the two modules might be helpful.

> *...I see a few key takeaways, but by this point I'd probably take a two-minute break or five-minute break only because that perspective taking took a lot of my mental capacity, so I'll probably just rest and, I don't know, refresh my mind for a couple minutes. Then, I'll go ahead and press the key takeaway.* [Participant 103]

Discussion

Principal Findings

In this study, we evaluated the initial usability, acceptability, and feasibility of the interactive, responsive platform version of ArtSpective for PSU—an adapted version of an arts-based synchronous intervention designed to improve nurses' stigmatizing attitudes toward PSU—following a mixed methods approach. We engaged 21 perinatal nurses and 4 experts in instructional design, stigma, and implicit bias to participate in theatre-testing sessions where they qualitatively assessed the usability of the platform and quantitatively assessed acceptability. This mixed methods approach has been extensively used to evaluate usability [61-64] and is consistent with previous work that uses theatre testing to gather feedback on an intervention from a specific group of users [65-67].

The results from the theatre testing sessions were positive overall, with participants expressing high satisfaction with the appearance of the platform, the ease with which they were able to navigate the various modules, their engagement in the platform, the clarity of the presentation, and its feasibility to be completed asynchronously by perinatal nurses. These findings were consistent with the quantitative acceptability measures and complementary satisfaction items. Using a mixed methods approach provided a more holistic assessment of the platform's different usability, acceptability, and feasibility aspects among perinatal nurses and experts.

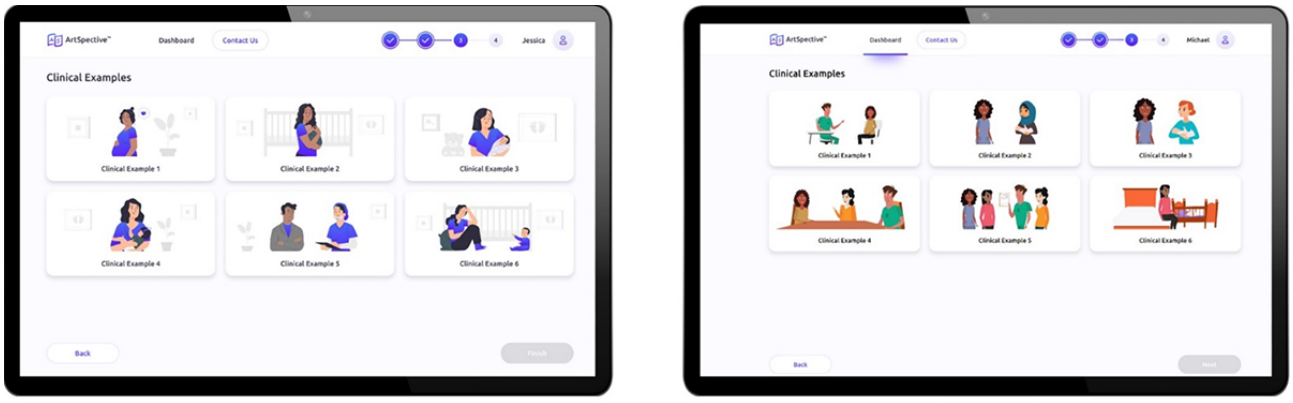
Furthermore, while quantitative and qualitative analyses indicated high acceptability and satisfaction, the theatre-testing sessions allowed for more granular data on platform usability. The use of theatre testing as a modality for conducting this type of evaluation also provided a more accessible method for practicing nurses to participate as it did not require them to allocate more significant time to engage in task-oriented usability testing, overcoming barriers related to engaging health care providers in this type of study [68-70] and engaging far more than the 5 participants that are minimally adequate to identifying more basic usability challenges [71]. Furthermore, an important implication of using theatre testing in this way is that we could collect granular usability data without investing fully in a developed platform. This provided for a more cost-effective development process that allowed developers to incorporate target user feedback before the final development of the platform.

In designing an interactive, responsive platform for perinatal nurses, we observed that while having a monochromatic color scheme consistent with material design principles may be helpful, it might also make it more difficult for nurses to discern differences. As a result, one of the updates made to the platform before development included changing the color scheme of illustrated characters in the Key Takeaways module from monochromatic to full color (Figure 5). In addition, while evidence related to effective asynchronous module instructional design recommends robust tutorials to guide participants [72,73], we observed that too many instructions could lead to more confusion and frustration. Therefore, future tutorial integration into interactive, responsive platforms for health care providers should be highly targeted and aligned with only the most essential information. This could be integrated through tooltips or info buttons that allow users to select which aspects of the platform they need more help clarifying. At the same time, our findings demonstrate that while an asynchronous platform may facilitate engagement in an intervention among health care providers, it must facilitate rapid content completion without sacrificing meaningful comprehension and overall effectiveness. Our findings indicate that achieving this goal requires a dynamic and modern design that also balances graphical and verbal cues

XSL•FO
RenderX

with a similar structure to other training platforms that allows users to recognize that the material is consistent with different pieces of training and, therefore, gain trust with users so that they are willing to engage in completing the intervention with a high degree of fidelity to each step.

Figure 5. Platform screens before and after the completion of theatre-testing sessions: (A) the original Clinical Examples screen with monochromatic characters and (B) the updated screen with full-color characters aligned with the animated videos.



A B

Comparison With Previous Work

Our findings related to acceptability and satisfaction were consistent with those related to our previous work, demonstrating satisfaction with the intervention [24]. This supports our conclusion that the platform addressed some of the dissatisfaction points identified in the previous study of the synchronous intervention, which illuminated the need for an asynchronous version of the intervention that would not compromise satisfaction overall among nurses.

Limitations

Our study is not without limitations. First, while we achieved information power after recruiting 21 perinatal nurses, 28% (7/25) of participants identified as either African American or Asian, and 88% (22/25) identified as female. While these demographics are somewhat representative of the perinatal nursing field, our results related to the usability of an interactive, responsive platform outside of the perinatal specialty might be more generalizable with a different group of participants. Furthermore, all participants were from two states in the Midwest, which may not be generalizable to providers in other regions of the United States. Second, although a larger sample of experts may have offered additional feedback, the sample included in this study was conveniently selected, represented multiple disciplines with varied expertise, and the research team concluded that information power was reached after receiving very similar feedback from all 4 experts being interviewed. Third, while theatre testing was an effective approach to gathering input from target users, we could not incorporate task-specific analyses into the sessions, as the theatre-testing approach is oriented around the presentation of the prototype. This enabled our study to include more nurses, but a task-oriented analysis, where we could compute a severity score for different issues encountered, may have allowed us to further augment our findings [74-76]. Fourth, while the validity and reliability of the AARP scale have been previously demonstrated [56], we did not have a sample large enough to evaluate the reliability of the AARP or the individual satisfaction items in our study. However, these items were administered primarily to pragmatically assess acceptability and satisfaction overall to inform future usability testing of the developed platform. Finally, since we were presenting a prototype of the intervention, not all features were fully implemented. For example, users reviewed a desktop version of the platform even though it was designed responsively. The decision to only review the desktop version was motivated by study feasibility. A prototype that switches between desktop and mobile formats may have confused participants as they went through the various modules. One alternative would be to review a separate mobile prototype. However, that would introduce feasibility challenges related to how much time perinatal nurses could allocate to participate in the study. To overcome these limitations, future studies might stratify participants into two groups where one group engages in theatre testing with a desktop-oriented prototype and the second group reviews a mobile prototype.

Conclusions

Developing and evaluating the usability of a digital adaptation of ArtSpective for PSU provided insight into key aspects related to usability, acceptability, and feasibility that should be considered when creating a digital adaptation of an arts-based intervention for health care providers. Overall, theatre-testing sessions identified key strengths and weaknesses of the prototype and provided opportunities to refine the platform's design before engaging in more costly development. Nurses and experts found the platform acceptable and were highly satisfied with its design and features. Future studies could evaluate the efficacy of the fully developed platform among practicing perinatal nurses and inform further refinements of the platform to ensure that health care providers in other specialties can engage with the intervention asynchronously.

XSL•FO
RenderX

Acknowledgments

This study was supported by the National Institute on Drug Abuse of the National Institutes of Health (1R21DA055067). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

The platform evaluated in this study was developed by Honey Locust Health. MR is a partial owner of Honey Locust Health, the firm that developed the digital adaptation of ArtSpective for perinatal substance use. All research performed in this study by MR was in his capacity as a University of Michigan investigator in adherence with a Conflict of Interest Management Plan established by the University of Michigan Office of Research Conflict of Interest Review Committee. CS, MR, DC, and CB are registered as inventors of ArtSpective for perinatal substance use. CS is the owner of ArtSpective LLC. All research performed in this study by CS was in his capacity as a University of Michigan investigator in adherence with a Conflict of Interest Management Plan established by the University of Michigan Office of Research Conflict of Interest Review Committee.

Multimedia Appendix 1
Semi-Structured Interview Guide.
[DOCX File , 16 KB - humanfactors_v12i1e67685_app1.docx]

Multimedia Appendix 2
COREQ (COnsolidated criteria for REporting Qualitative research) Checklist.
[PDF File (Adobe PDF File), 171 KB - humanfactors_v12i1e67685_app2.pdf]

References

1. Hirai AH, Ko JY, Owens PL, Stocks C, Patrick SW. Neonatal abstinence syndrome and maternal opioid-related diagnoses in the US, 2010-2017. JAMA 2021;325(2):146-155 [FREE Full text] [doi: 10.1001/jama.2020.24991] [Medline: 33433576]
2. Winkelman TNA, Villapiano N, Kozhimannil KB, Davis MM, Patrick SW. Incidence and costs of neonatal abstinence syndrome among infants with medicaid: 2004-2014. Pediatrics 2018;141(4):e20173520 [FREE Full text] [doi: 10.1542/peds.2017-3520] [Medline: 29572288]
3. Patrick SW, Davis MM, Lehmann CU, Cooper WO. Increasing incidence and geographic distribution of neonatal abstinence syndrome: United States 2009 to 2012. J Perinatol 2015;35(8):650-655 [FREE Full text] [doi: 10.1038/jp.2015.36] [Medline: 25927272]
4. Desai RJ, Hernandez-Diaz S, Bateman BT, Huybrechts KF. Increase in prescription opioid use during pregnancy among Medicaid-enrolled women. Obstet Gynecol 2014;123(5):997-1002 [FREE Full text] [doi: 10.1097/AOG.0000000000000208] [Medline: 24785852]
5. Bateman BT, Hernandez-Diaz S, Rathmell JP, Seeger JD, Doherty M, Fischer MA, et al. Patterns of opioid utilization in pregnancy in a large cohort of commercial insurance beneficiaries in the United States. Anesthesiology 2014;120(5):1216-1224 [FREE Full text] [doi: 10.1097/ALN.0000000000000172] [Medline: 24525628]
6. Martin CE, Longinaker N, Terplan M. Recent trends in treatment admissions for prescription opioid abuse during pregnancy. J Subst Abuse Treat 2015;48(1):37-42 [FREE Full text] [doi: 10.1016/j.jsat.2014.07.007] [Medline: 25151440]
7. Yazdy MM, Desai RJ, Brogly SB. Prescription opioids in pregnancy and birth outcomes: a review of the literature. J Pediatr Genet 2015;4(2):56-70 [FREE Full text] [doi: 10.1055/s-0035-1556740] [Medline: 26998394]
8. Patrick SW, Schumacher RE, Horbar JD, Buus-Frank ME, Edwards EM, Morrow KA, et al. Improving care for neonatal abstinence syndrome. Pediatrics. 2016. URL: http://publications.aap.org/pediatrics/article-pdf/137/5/e20153835/1060768/peds_20153835.pdf [accessed 2025-04-01]
9. Walsh MC, Crowley M, Wexelblatt S, Ford S, Kuhnell P, Kaplan HC, et al. Ohio Perinatal Quality Collaborative improves care of neonatal narcotic abstinence syndrome. Pediatrics 2018;141(4):e20170900. [doi: 10.1542/peds.2017-0900] [Medline: 29514974]
10. Grossman MR, Berkwitt AK, Osborn RR, Xu Y, Esserman DA, Shapiro ED, et al. An initiative to improve the quality of care of infants with neonatal abstinence syndrome. Pediatrics 2017 Jun;139(6) [FREE Full text] [doi: 10.1542/peds.2016-3360] [Medline: 28562267]
11. Asti L, Magers JS, Keels E, Wispe J, McClead RE. A quality improvement project to reduce length of stay for neonatal abstinence syndrome. Pediatrics 2015;135(6):e1494-e1500. [doi: 10.1542/peds.2014-1269] [Medline: 25941308]
12. Holmes AV, Atwood EC, Whalen B, Beliveau J, Jarvis JD, Matulis JC, et al. Rooming-In to treat neonatal abstinence syndrome: improved family-centered care at lower cost. Pediatrics 2016;137(6):e20152929. [doi: 10.1542/peds.2015-2929] [Medline: 27194629]
13. McKnight S, Coo H, Davies G, Holmes B, Newman A, Newton L, et al. Rooming-in for infants at risk of neonatal abstinence syndrome. Am J Perinatol 2016;33(5):495-501. [doi: 10.1055/s-0035-1566295] [Medline: 26588259]

XSL•FO
RenderX

14. Holmes AP, Schmidlin HN, Kurzum EN. Breastfeeding considerations for mothers of infants with neonatal abstinence syndrome. Pharmacotherapy 2017;37(7):861-869. [doi: 10.1002/phar.1944] [Medline: 28488805]
15. Pritham UA. Breastfeeding promotion for management of neonatal abstinence syndrome. J Obstet Gynecol Neonatal Nurs 2013;42(5):517-526. [doi: 10.1111/1552-6909.12242] [Medline: 24020477]
16. Abdel-Latif ME, Pinner J, Clews S, Cooke F, Lui K, Oei J. Effects of breast milk on the severity and outcome of neonatal abstinence syndrome among infants of drug-dependent mothers. Pediatrics 2006;117(6):e1163-e1169. [doi: 10.1542/peds.2005-1561] [Medline: 16740817]
17. Atwood EC, Sollender G, Hsu E, Arsnow C, Flanagan V, Celenza J, et al. A qualitative study of family experience with hospitalization for neonatal abstinence syndrome. Hosp Pediatr 2016;6(10):626-632. [doi: 10.1542/hpeds.2016-0024] [Medline: 27647475]
18. Cleveland LM, Bonugli R. Experiences of mothers of infants with neonatal abstinence syndrome in the neonatal intensive care unit. J Obstet Gynecol Neonatal Nurs 2014;43(3):318-329. [doi: 10.1111/1552-6909.12306] [Medline: 24754258]
19. Short VL, Cambareri K, Gannon M, Alexander K, Abatemarco DJ. A pilot study to assess breastfeeding knowledge, attitudes, and perceptions of individuals who work in perinatal opioid use disorder treatment settings. Breastfeed Med 2019;14(5):307-312. [doi: 10.1089/bfm.2018.0257] [Medline: 30888210]
20. Tobin KB. Changing neonatal nurses' perceptions of caring for infants experiencing neonatal abstinence syndrome and their mothers: an evidenced-based practice opportunity. Adv Neonatal Care 2018;18(2):128-135. [doi: 10.1097/ANC.0000000000000476] [Medline: 29595550]
21. Antunez AG, Rowe KA, Bain PA, Doherty GM, Dossett LA. Behavioral interventions on surgeons' preoperative decision-making: A scoping review and critical analysis. J Surg Res 2024;295:547-558. [doi: 10.1016/j.jss.2023.11.037] [Medline: 38086255]
22. Thapa D, Liu T, Chair SY. Multifaceted interventions are likely to be more effective to increase adherence to the ventilator care bundle: A systematic review of strategies to improve care bundle compliance. Intensive Crit Care Nurs 2023;74:103310. [doi: 10.1016/j.iccn.2022.103310] [Medline: 36154789]
23. Nyblade L, Stockton MA, Giger K, Bond V, Ekstrand ML, Lean RM, et al. Stigma in health facilities: why it matters and how we can change it. BMC Med 2019;17(1):25 [FREE Full text] [doi: 10.1186/s12916-019-1256-2] [Medline: 30764806]
24. Shuman CJ, Choberka D, Morgan ME, Boyd CJ. ArtSpective™ for prinatal substance use: feasibility of an arts-based intervention addressing stigma. J Addict Nurs 2024;35(2):43-50. [doi: 10.1097/JAN.0000000000000577] [Medline: 38829993]
25. Fancourt D, Finn S. What is the evidence on the role of the arts in improving health and well-being? a scoping review. World Health Organization. 2019. URL: https://www.who.int/europe/publications/i/item/9789289054553 [accessed 2025-04-03]
26. Rieger KL, Chernomas WM, McMillan DE, Morin FL, Demczuk L. Effectiveness and experience of arts-based pedagogy among undergraduate nursing students: a mixed methods systematic review. JBI Database System Rev Implement Rep 2016;14(11):139-239. [doi: 10.11124/JBISRIR-2016-003188] [Medline: 27941518]
27. Flanagan EH, Buck T, Gamble A, Hunter C, Sewell I, Davidson L. "Recovery Speaks": A photovoice intervention to reduce stigma among primary care providers. Psychiatr Serv 2016;67(5):566-569. [doi: 10.1176/appi.ps.201500049] [Medline: 26766754]
28. George DR, Stuckey HL, Whitehead MM. An arts-based intervention at a nursing home to improve medical students' attitudes toward persons with dementia. Acad Med 2013;88(6):837-842. [doi: 10.1097/ACM.0b013e31828fa773] [Medline: 23619065]
29. Elder NC, Tobias B, Lucero-Criswell A, Goldenhar L. The art of observation: impact of a family medicine and art museum partnership on student education. Fam Med 2006;38(6):393-398. [Medline: 16741837]
30. Georges JM. An emancipatory theory of compassion for nursing. ANS Adv Nurs Sci 2013;36(1):2-9. [doi: 10.1097/ANS.0b013e31828077d2] [Medline: 23370495]
31. Shuman CJ, Isaman DJ, Morgan ME, Kukora SK, Rubyan M, Veliz PT, et al. Validity and reliability of the modified attitudes about drug use in pregnancy scale. J Obstet Gynecol Neonatal Nurs 2022;51(3):324-335. [doi: 10.1016/j.jogn.2022.02.006] [Medline: 35341716]
32. Shuman CJ, Wilson R, VanAntwerp K, Morgan M, Weber A. Elucidating the context for implementing nonpharmacologic care for neonatal opioid withdrawal syndrome: a qualitative study of perinatal nurses. BMC Pediatr 2021;21(1):489 [FREE Full text] [doi: 10.1186/s12887-021-02955-y] [Medline: 34736443]
33. Shuman CJ, Weber A, VanAntwerp K, Wilson R. Engaging mothers to implement nonpharmacological care for infants with neonatal abstinence syndrome: perceptions of perinatal and pediatric nurses. Adv Neonatal Care 2020;20(6):464-472 [FREE Full text] [doi: 10.1097/ANC.0000000000000812] [Medline: 33009157]
34. Underwood B, Moore B. Perspective-taking and altruism. Psychological Bulletin 1982;91(1):143-173. [doi: 10.1037//0033-2909.91.1.143]
35. Gaiha SM, Salisbury TT, Usmani S, Koschorke M, Raman U, Petticrew M. Effectiveness of arts interventions to reduce mental-health-related stigma among youth: a systematic review and meta-analysis. BMC Psychiatry 2021;21(1):364 [FREE Full text] [doi: 10.1186/s12888-021-03350-8] [Medline: 34294067]

XSL•FO
RenderX

36. Singh L, Kanstrup M, Depa K, Falk A, Lindström V, Dahl O, et al. Digitalizing a brief intervention to reduce intrusive memories of psychological trauma for health care staff working during COVID-19: exploratory pilot study with nurses. JMIR Form Res 2021;5(5):e27473 [FREE Full text] [doi: 10.2196/27473] [Medline: 33886490]
37. Fiol-DeRoque MA, Serrano-Ripoll MJ, Jiménez R, Zamanillo-Campos R, Yáñez-Juan AM, Bennasar-Veny M, et al. A mobile phone-based intervention to reduce mental health problems in health care workers during the COVID-19 pandemic (PsyCovidApp): randomized controlled trial. JMIR Mhealth Uhealth 2021;9(5):e27039 [FREE Full text] [doi: 10.2196/27039] [Medline: 33909587]
38. Albott CS, Wozniak JR, McGlinch BP, Wall MH, Gold BS, Vinogradov S. Battle buddies: rapid deployment of a psychological resilience intervention for health care workers during the COVID-19 pandemic. Anesth Analg 2020;131(1):43-54 [FREE Full text] [doi: 10.1213/ANE.0000000000004912] [Medline: 32345861]
39. Andrews B, Klein B, Corboy D, McLaren S, Watson S. Therapist training in video chat technology for use in an adaptive digital mental health intervention: challenges, facilitators and implications for training models. Couns and Psychother Res 2023;23(3):818-829. [doi: 10.1002/capr.12667]
40. Jeon E, Park HE. Nursing intervention using smartphone technologies; a systematic review and meta-analysis. Stud Health Technol Inform 2015;210:321-325. [Medline: 25991158]
41. Haakma T, Tieben R, Sleven B, Buise M, van Mol M. Experiences of nurses with an innovative digital diary intervention in the intensive care unit: A qualitative exploration. Intensive Crit Care Nurs 2022;70:103197 [FREE Full text] [doi: 10.1016/j.iccn.2022.103197] [Medline: 35090798]
42. Yu J, Parsons GS, Lancastle D, Tonkin ET, Ganesh S. "Walking in Their Shoes": the effects of an immersive digital story intervention on empathy in nursing students. Nurs Open 2021;8(5):2813-2823 [FREE Full text] [doi: 10.1002/nop2.860] [Medline: 33743185]
43. Harte R, Glynn L, Rodríguez-Molinero A, Baker PM, Scharf T, Quinlan LR, et al. A human-centered design methodology to enhance the usability, human factors, and user experience of connected health systems: A three-phase methodology. JMIR Hum Factors 2017;4(1):e8 [FREE Full text] [doi: 10.2196/humanfactors.5443] [Medline: 28302594]
44. Nesbitt K, Beleigoli A, Du H, Tirimacco R, Clark R. User experience (UX) design as a co-design methodology: lessons learned during the development of a web-based portal for cardiac rehabilitation. Eur J Cardiovasc Nurs 2022;21(2):178-183. [doi: 10.1093/eurjcn/zvab127] [Medline: 35030261]
45. Material design. Google. 2024. URL: https://m3.material.io/ [accessed 2025-04-01]
46. Nielsen J, Molich R. Heuristic evaluation of user interfaces. : Association for Computing Machinery; 1990 Presented at: CHI '90: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; April 1-5, 1990; Seattle, WA p. 249-256. [doi: 10.1145/97243.97281]
47. Lenzner T, Kaczmirek L, Lenzner A. Cognitive burden of survey questions and response times: a psycholinguistic experiment. Applied Cognitive Psychology 2009;24(7):1003-1020. [doi: 10.1002/acp.1602]
48. Kanuri N, Arora P, Talluru S, Colaco B, Dutta R, Rawat A, et al. Examining the initial usability, acceptability and feasibility of a digital mental health intervention for college students in India. Int J Psychol 2020;55(4):657-673 [FREE Full text] [doi: 10.1002/ijop.12640] [Medline: 31867730]
49. Drotar D, Lemanek K. Steps toward a clinically relevant science of interventions in pediatric settings: introduction to the special issue. J Pediatr Psychol 2001;26(7):385-394. [doi: 10.1093/jpepsy/26.7.385] [Medline: 11553693]
50. Ergonomics of human-system interaction - Part 11: usability: definitions and concepts (ISO 9241-11). International Organization for Standardization. 2023. URL: https://www.iso.org/standard/63500.html [accessed 2025-04-01]
51. Nadal C, Sas C, Doherty G. Technology acceptance in mobile health: scoping review of definitions, models, and measurement. J Med Internet Res 2020;22(7):e17256 [FREE Full text] [doi: 10.2196/17256] [Medline: 32628122]
52. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res 2017;17(1):88 [FREE Full text] [doi: 10.1186/s12913-017-2031-8] [Medline: 28126032]
53. Perski O, Short CE. Acceptability of digital health interventions: embracing the complexity. Transl Behav Med 2021;11(7):1473-1480 [FREE Full text] [doi: 10.1093/tbm/ibab048] [Medline: 33963864]
54. Wingood GM, DiClemente RJ. The ADAPT-ITT model: a novel method of adapting evidence-based HIV Interventions. J Acquir Immune Defic Syndr 2008;47 Suppl 1:S40-S46. [doi: 10.1097/QAI.0b013e3181605df1] [Medline: 18301133]
55. National Institutes of Health National Cancer Institute. Making health communication programs work. U.S. Department of Health & Human Services. 2004. URL: https://www.cancer.gov/publications/health-communication/pink-book.pdf [accessed 2025-04-01]
56. Tarnowski KJ, Simonian SJ. Assessing treatment acceptance: the abbreviated acceptability rating profile. J Behav Ther Exp Psychiatry 1992;23(2):101-106. [doi: 10.1016/0005-7916(92)90007-6] [Medline: 1460097]
57. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
58. Jaspers MWM, Steen T, van den Bos C, Geenen M. The think aloud method: a guide to user interface design. Int J Med Inform 2004;73(11-12):781-795. [doi: 10.1016/j.ijmedinf.2004.08.003] [Medline: 15491929]

XSL•FO
RenderX

59. McDonald S, Zhao T, Edwards HM. Dual verbal elicitation: the complementary use of concurrent and retrospective reporting within a usability test. International Journal of Human-Computer Interaction 2013;29(10):647-660. [doi: 10.1080/10447318.2012.758529]
60. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013;13:117 [FREE Full text] [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]
61. Otis M, Zhu J, Mustafa-Kutana SN, Bernier AV, Ma Shum J, Soros Dupre AA, et al. Testing usability and feasibility of a mobile educator tool for pediatric diabetes self-management: mixed methods pilot study. JMIR Form Res 2020;4(5):e16262 [FREE Full text] [doi: 10.2196/16262] [Medline: 32356773]
62. Khairat S, Coleman C, Newlin T, Rand V, Ottmar P, Bice T, et al. A mixed-methods evaluation framework for electronic health records usability studies. J Biomed Inform 2019;94:103175 [FREE Full text] [doi: 10.1016/j.jbi.2019.103175] [Medline: 30981897]
63. Ossebaard HC, Seydel ER, van Gemert-Pijnen L. Online usability and patients with long-term conditions: a mixed-methods approach. Int J Med Inform 2012;81(6):374-387. [doi: 10.1016/j.ijmedinf.2011.12.010] [Medline: 22261086]
64. Miller J. The design cycle and a mixed methods approach for improving usability: a case study. Journal of Web Librarianship 2019;13(3):203-229. [doi: 10.1080/19322909.2019.1600451]
65. MacEntee K, Lee EOJ, Oudshoorn A, Abramovich A, Kiptui R, Ayuku D, et al. Using scenario videos with theatre testing method to adapt a peer navigation model to improve street-connected youth's access to HIV care in Kenya and Canada. Front Public Health 2022;10:975117 [FREE Full text] [doi: 10.3389/fpubh.2022.975117] [Medline: 36408034]
66. Davis T, DiClemente RJ, Prietula M. Using ADAPT-ITT to modify a telephone-based HIV prevention intervention for SMS delivery: formative study. JMIR Form Res 2020;4(10):e22485 [FREE Full text] [doi: 10.2196/22485] [Medline: 32831178]
67. MacEntee K, Lee EOJ, Oudshoorn A, Abramovich A, Kiptui R, Ayuku D, et al. Using scenario videos with theatre testing method to adapt a peer navigation model to improve street-connected youth's access to HIV care in Kenya and Canada. Front Public Health 2022;10:975117 [FREE Full text] [doi: 10.3389/fpubh.2022.975117] [Medline: 36408034]
68. Wray J, Gates B. Problems of recruiting participants for nursing research: a case study. NT Research 1996;1(5):366-373. [doi: 10.1177/136140969600100509]
69. Chlan L, Guttormson J, Tracy M, Bremer K. Strategies for overcoming site and recruitment challenges in research studies based in intensive care units. Am J Crit Care 2009;18(5):410-417 [FREE Full text] [doi: 10.4037/ajcc2009400] [Medline: 19723861]
70. Staggers N, Elias BL, Makar E, Alexander GL. The imperative of solving nurses' usability problems with health information technology. J Nurs Adm 2018;48(4):191-196. [doi: 10.1097/NNA.0000000000000598] [Medline: 29570144]
71. Nielsen J. Why you only need to test with 5 users. 2000. URL: https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/ [accessed 2025-04-01]
72. Chen F, Lui AM, Martinelli SM. A systematic review of the effectiveness of flipped classrooms in medical education. Med Educ 2017;51(6):585-597. [doi: 10.1111/medu.13272] [Medline: 28488303]
73. Galway LP, Corbett KK, Takaro TK, Tairyan K, Frank E. A novel integration of online and flipped classroom instructional models in public health higher education. BMC Med Educ 2014;14:181 [FREE Full text] [doi: 10.1186/1472-6920-14-181] [Medline: 25169853]
74. Bond RR, Finlay DD, Nugent CD, Moore G, Guldenring D. A usability evaluation of medical software at an expert conference setting. Comput Methods Programs Biomed 2014;113(1):383-395. [doi: 10.1016/j.cmpb.2013.10.006] [Medline: 24210868]
75. Kennedy B, Kerns E, Chan Y, Chaparro B, Fouquet S. Safeuristics! Do heuristic evaluation violation severity ratings correlate with patient safety severity ratings for a native electronic health record mobile application? Appl Clin Inform 2019;10(2):210-218 [FREE Full text] [doi: 10.1055/s-0039-1681073] [Medline: 30919397]
76. Cho H, Yen PY, Dowding D, Merrill JA, Schnall R. A multi-level usability evaluation of mobile health applications: a case study. J Biomed Inform 2018;86:79-89 [FREE Full text] [doi: 10.1016/j.jbi.2018.08.012] [Medline: 30145317]

Abbreviations

AARP: Abbreviated Acceptability Rating Profile
ADAPT-ITT: Assessment Decision Administration Production Topical Experts-Integration Training Testing
COREQ: Consolidated Criteria for Reporting Qualitative Research
PSU: perinatal substance use

XSL•FO
RenderX

Edited by A Kushniruk, E Borycki; submitted 18.10.24; peer-reviewed by R Marshall, Q Xu; comments to author 05.01.25; revised version received 01.03.25; accepted 27.03.25; published 08.05.25.

Please cite as:
Rubyan M, Gouseinov Y, Morgan M, Rubyan D, Jahagirdar D, Choberka D, Boyd CJ, Shuman C
Evaluating the Usability, Acceptability, User Experience, and Design of an Interactive Responsive Platform to Improve Perinatal Nurses' Stigmatizing Attitudes Toward Substance Use in Pregnancy: Mixed Methods Study
JMIR Hum Factors 2025;12:e67685
URL: https://humanfactors.jmir.org/2025/1/e67685
doi:10.2196/67685
PMID:40146964

©Michael Rubyan, Yana Gouseinov, Mikayla Morgan, Deborah Rubyan, Divya Jahagirdar, David Choberka, Carol J Boyd, Clayton Shuman. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 08.05.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Effects of a Mobile Storytelling App (Huiyou) on Social Participation Among People With Mild Cognitive Impairment: Pilot Randomized Controlled Trial

Di Zhu[1], PhD; Abdullah Al Mahmud[1*], PhD; Wei Liu[2*], PhD

[1]Centre for Design Innovation, Swinburne University of Technology, Melbourne, Australia

[2]School of Design, Southern University of Science and Technology, Shenzhen, China

[*]these authors contributed equally

Corresponding Author:
Abdullah Al Mahmud, PhD
Centre for Design Innovation
Swinburne University of Technology
John St
Melbourne, 3122
Australia
Phone: 61 03 9214 3830
Email: aalmahmud@swin.edu.au

Abstract

Background: As the prevalence of mild cognitive impairment (MCI) among older adults increases, so does the need to enhance social participation and cognitive functions through innovative interventions. Digital storytelling in group settings holds potential not only to foster social connections but also to integrate with traditional in-person activities, leveraging both for greater impact.

Objective: This study aimed to evaluate the feasibility of the Huiyou app in supporting group-based storytelling activities, aiming to enhance social participation for people with MCI. We focused on the app's ability to improve storytelling goal attainment, social connectedness, self-efficacy, and subjective happiness, comparing these outcomes between the experimental and control groups.

Methods: We randomly assigned 20 participants with MCI to either an intervention group or a control group, engaging them in the use of the Huiyou digital storytelling app over 4 weekly sessions of 45 minutes each. We measured outcomes through the Assessment of Life Habits questionnaire (77 items), particularly outdoor activities and interpersonal relationships; the Social Connectedness Scale–Revised (20 items); the General Self-Efficacy Scale (10 items), focusing on coping self-efficacy; and the Subjective Well-Being Scale (SWBS; 20 items), with a special emphasis on self-acceptance.

Results: The sample had an average age of 69.7 (SD 3.21) years, with no significant (P=.23) baseline differences between groups in age, sex, or educational background. Cognitive function, assessed via the Montreal Cognitive Assessment–Chinese questionnaire, also showed no significant differences at baseline (P=.20). Specifically, significant enhancements in the outdoor activity (mean value difference 0.171, SD 0.353; Cohen d=1.046; P=.03) and interpersonal adaptation experience subscales of the SWBS (mean value difference 0.167, SD 0.247; Cohen d=1.290; P=.01) were noted. Notably, storytelling performance improved markedly, evidenced by increases in story sharing duration and complexity. Although overall improvements in Assessment of Life Habits (P=.14), Social Connectedness Scale–Revised (P=.59), and Subjective Well-Being Scale (P=.26) scores were not statistically significant, the large effect sizes observed suggest potential benefits of the Huiyou app that might be obscured by the study's small sample size.

Conclusions: This study indicates that the Huiyou mobile storytelling app is feasible to enhance social participation and specific aspects of social functioning such as interpersonal adaptation for people with MCI. Despite the lack of significant changes in overall scores for key scales, observed effect sizes highlight a positive trend that merits further investigation. These results advocate for the continuation of digital intervention development to improve quality of life and social integration for individuals with MCI.

(JMIR Hum Factors 2025;12:e70177) doi:10.2196/70177

XSL•FO
RenderX

KEYWORDS

digital storytelling; technology-based intervention; social participation; people with mild cognitive impairment; digital health; mobile apps

Introduction

Background

People with mild cognitive impairment (MCI) face significant challenges in daily life due to cognitive deficits and emotional changes that reduce their ability to function socially [1]. Cognitive decline is closely linked to decreased social activities, with behavioral symptoms and social-cognitive deficits often leading to poor maintenance of social networks and difficulties using everyday technology [2,3]. Memory loss further undermines feelings of security and autonomy, often resulting in frustration and anxiety [4]. Despite these challenges, maintaining high-level social participation is crucial for people with MCI as it has been shown to prevent or slow down cognitive decline, improve quality of life, and reduce social isolation and loneliness [5-8]. Despite this, people with MCI often experience reduced social connections and activities, which are fundamental to social participation [9,10]. Conversely, limited social engagement can lead to further cognitive decline and increased isolation [11,12]. However, multiple barriers, including the subtle nature of MCI-related social decline, often impede participation in social activities. In response, various interventions have been developed to support social participation for people with MCI. Technology-based interventions, supported by modern technologies such as cameras, internet communication, and speech-to-text tools, have shown promise in enhancing social participation for people with MCI. These interventions include cognitive rehabilitation programs, educational tools, telecommunication systems, social robots, and self-management systems, all designed to improve social engagement and quality of life [13-15]. While some studies suggest that these interventions have outcomes comparable to those of traditional treatments, such as in the quality of social interactions [16], they offer the advantage of expanding care from hospitals to homes, reducing therapists' workload and enabling remote treatment [17]. Recent studies have demonstrated that patients engaging in digital technology–based serious game interventions experience notable improvements across multiple domains, including global cognitive function, executive function, attention, mood (notably reductions in depression), and activities of daily living, compared to those undergoing traditional rehabilitation or receiving standard medical care [18]. Given this context, digital storytelling interventions, which combine the benefits of technology with the therapeutic power of narrative, hold particular potential to further enhance social participation and cognitive engagement among individuals with MCI.

Digital storytelling is a powerful tool for enhancing cognitive engagement and fostering social connections, particularly among older adults with MCI. Engaging in storytelling requires individuals to recall past experiences, organize their thoughts, and articulate them coherently [19], which stimulates essential cognitive processes such as memory recall, attention, and language skills. For those with MCI, this activity offers a structured yet creative outlet for mental engagement, helping maintain and even improve cognitive functions that are often challenged by the condition [20]. In addition, storytelling fosters social connections by creating opportunities for individuals to share their personal narratives with others whether in person or through digital platforms [21]. Recent developments such as time travel tours have made digital storytelling more accessible to older adults by enabling them to create and share location-based narratives through a conversational web interface without requiring advanced technical skills [22]. This exchange of stories not only builds empathy, understanding, and a sense of community but also combats the social isolation that frequently accompanies cognitive decline [23]. When older adults share their stories in a group setting, they reinforce their own memories while contributing to the collective memory of the group, further strengthening social bonds [24]. As an intervention, digital storytelling uniquely addresses the cognitive and social challenges faced by people with MCI by combining the therapeutic benefits of storytelling with the accessibility and engagement potential of digital tools. This dual focus on cognitive engagement and social connection makes digital storytelling a valuable approach for improving the overall well-being of older adults with MCI, providing them with a meaningful way to connect with others and exercise their cognitive abilities.

Prior Work

Technology-based storytelling interventions involve recalling previous actions, events, and feelings using physical prompts (ie, photographs, household items, music, and audio recordings). Such prompts are now commonly stored and presented digitally [25]. However, there is some evidence, albeit not conclusive, suggesting that storytelling interventions or reminiscence therapy for dementia yield no significant effect on quality of life immediately following the intervention period when compared to a control group (no treatment; see the study by Wood [26]). In reviews of storytelling interventions and apps aimed at enhancing social participation among people with MCI [19,27], we identified 4 key themes: capture memory, saving memory, memory retrieval, and sharing. These themes represent the core aspects of the interventions, focusing on how individuals can record, store, recall, and share their personal experiences. These elements collectively contribute to the effectiveness of technology-based storytelling interventions.

Various digital storytelling applications are designed to assist older adults, particularly those with MCI or dementia, in enhancing their mood and increasing social engagement [28]. The main features for generating materials across these applications include the ability to upload personalized content such as photos, videos, and music [29]. Some applications offer additional capabilities such as selecting auto-play intervals [30], adding tags [31], or creating dynamic timelines [29]. For memory retrieval, features include auto-playing photos and videos, viewing materials on timelines, and using virtual reality to project realistic images [32]. Some applications support a

XSL•FO
RenderX

random display of materials to trigger reminiscence [33]. Story sharing capabilities vary, with some applications enabling sharing through smartphones, tablets, or dedicated platforms [34,35]. These features often include options to share edited storybooks, personal videos, or narrated slideshows with family, friends, or broader networks [25,30]. Additional features frequently involve creating structured stories, using prewritten scripts, integrating multimedia content, and supporting administrative activities such as tagging and organizing materials [36].

The platforms for these applications are primarily smartphones and tablets [32], with some using laptops, computers, or virtual reality devices [37]. The impact on social participation includes improving mood, increasing social interactions, enhancing subjective well-being, and fostering closer ties with friends and family [30]. However, specific outcomes vary, and not all applications provide detailed results on social participation. Overall, these digital storytelling applications demonstrate significant potential for enhancing social engagement and emotional well-being among older adults with cognitive impairments.

Digital storytelling interventions offer a promising avenue for enhancing social participation among people with MCI, but they also face significant barriers and limitations. One major issue is that most interventions target individual relationships rather than broader community-based connections [8,38], which can limit the social impact and scope of interaction. This focus on personal relationships might support social connections within families but does not necessarily encourage engagement with the wider community or strangers, which is essential for comprehensive social participation. Another barrier involves the difficulty that people with MCI often encounter with commercial digital storytelling applications [39]. These applications frequently require a level of digital literacy that can be challenging for this group even after training sessions. As a result, many interventions rely on caregivers or trained volunteers to help develop and share stories, which may not be sustainable or widely accessible [29,40]. This indicates a need for user-friendly applications designed to accommodate the specific needs and capabilities of people with MCI, reducing the learning curve and associated resource demands. There is a research gap in developing customized software that caters specifically to the usability needs of people with MCI [41,42]. The addition of group elements to digital storytelling can enhance engagement and foster a sense of community [43], yet current digital storytelling tools focus on individual experiences, not collaborative storytelling [30,44]. Privacy concerns also emerge when sharing stories with broader audiences [43,45]. While sharing personal stories online can create new connections and engagement, it also poses risks of exposing sensitive information or triggering unfavorable reactions from a wider audience. This necessitates careful consideration of privacy and safeguarding measures when designing digital storytelling interventions. Moreover, most studies use the basic versions of the interventions or applications, but few studies have been conducted as randomized controlled trials to measure the efficacy of the interventions or applications [19]. Recent international research underscores the potential of digital health interventions in supporting individuals with MCI. A systematic review and meta-analysis by Di Lorito et al [46] found that such interventions can produce positive effects on cognitive abilities among people with MCI and dementia. Furthermore, Karakose et al [47] conducted a comprehensive bibliometric and science mapping analysis, revealing the evolving landscape of digital addiction research, which is pertinent to understanding user engagement with digital health tools. These findings highlight the importance of integrating international perspectives to inform the development and implementation of digital storytelling applications such as Huiyou.

The Huiyou App

To bridge these research gaps, we co-designed an app, 会友 (Huiyou, derived from the pinyin pronunciation, which means “meeting new friends” in Chinese), with therapists and people with MCI [48]. It is inspired by a classic quote from Confucius in the *Analects*: “A gentleman seeks friendship through literature and reinforces goodness through friendship.” This statement emphasizes the idea that individuals cultivate friendships through literary exchange and support virtue through companionship. The name reflects a positive vision of fostering social connections through literature, friendship, and benevolence. Users can leverage technological means through the software to expand their social circles, facilitating deeper communication and connections with others. The name embodies the social nature of the software and its goal of promoting friendship. Huiyou effectively stores memories from daily life and encourages people with MCI to reminisce and discuss favorable memories with new friends. Huiyou has 2 main features: supporting people with MCI to conduct self-reflection daily (preparing materials using cues) on certain topics and facilitating group memory retrieval (presenting a story and promoting a discussion with group members). In addition, Huiyou uniquely integrates structured outdoor activities with digital storytelling to reinforce physical engagement alongside cognitive and social interaction. The 2 innovative aspects of Huiyou are embedded memory retrieval for capturing daily life and its ability to collect recent, valuable memories. The intervention design also deliberately combines self-generated memory prompts with shared group discussions to maximize both personal reflection and collective participation. Another feature is combined self-reflection and group reflection to enhance social interaction during the discussion in addition to sharing and participating in social activities outside the home. A previous publication provides more details about the design of Huiyou [48].

Study Objectives

This study contributes theoretically by advancing the understanding of digital storytelling to enhance social participation and emotional well-being for people with MCI. Empirically, it delivers one of the first pilot evaluations of a co-designed, culturally adapted mobile storytelling app for community-dwelling older adults with MCI in China. The integration of storytelling with structured outdoor engagement provides a practical model for accessible, technology-supported interventions in low-resource settings. This experimental study aimed to investigate the feasibility of the Huiyou digital

XSL•FO
RenderX

storytelling intervention for people with MCI. We hypothesized that (1) the Huiyou intervention would significantly enhance social participation among individuals with MCI, as evidenced by improved quantitative measures of social participation outcomes; and (2) the intervention would facilitate the achievement of predefined social participation goals, with quantitative assessments confirming goal attainment.

Methods

Ethical Considerations

Ethics Review Approvals or Exemptions

Ethics approval for this study was granted by the Swinburne University of Technology Human Research Ethics Committee (approval number 20226525-11105; September 26, 2022). The study was conducted in accordance with the principles of the Declaration of Helsinki. The study protocol was also published in advance [49].

Informed Consent

Informed consent was obtained from all participants prior to their involvement in the study. For participants with mild cognitive impairment (MCI), consent was obtained with the support of a formal or informal caregiver. Participants were also asked to articulate their reasons for participation to confirm understanding and willingness. Trained research assistants provided step-by-step guidance throughout each session, addressing confusion or distress in real time. Participants were continuously monitored for signs of cognitive overload or discomfort, with immediate support provided as needed.

Privacy and Confidentiality

All data were anonymized using participant ID codes, and no personally identifiable information was collected or stored. Data were securely stored on encrypted servers with access restricted to authorized research personnel. No identifiable images or personal data of participants are presented in this manuscript.

Compensation Details

No financial or material compensation was provided to participants. Participation was entirely voluntary, and support was offered to ensure a safe and respectful experience.

Study Design and Procedure

We conducted a 4-week experimental study with a waitlist control group to assess the impact of the Huiyou app on people with MCI (see Multimedia Appendix 1 for the CONSORT checklist). The randomization sequence for participant allocation was computer generated and managed through a Microsoft Excel spreadsheet. As this study involved active participation in the experimental activities, it was designed as a single-blind experiment. The experiment was conducted in 2 batches, with the second batch serving as a supplemental recruitment due to some participants' limited availability.

This study commenced with the installation and learning of the Huiyou software, along with pretesting, on December 18, 2023. The intervention period spanned the second to the fifth week, with posttesting conducted at the end of the fifth week and some participants opting for testing in the sixth week. The intervention sessions took place at the Jing Shi Psychology Centre Meeting Room in Tiantongyuan, Tongzhou District, Beijing.

After each intervention session, the theme for the following week was announced, with the first week's theme disclosed after the pretest. A WeChat group was established to facilitate the convenient communication of activity information and scheduling adjustments. Participants had the flexibility to upload the necessary materials for their stories through the Huiyou app at any time, including photos and videos.

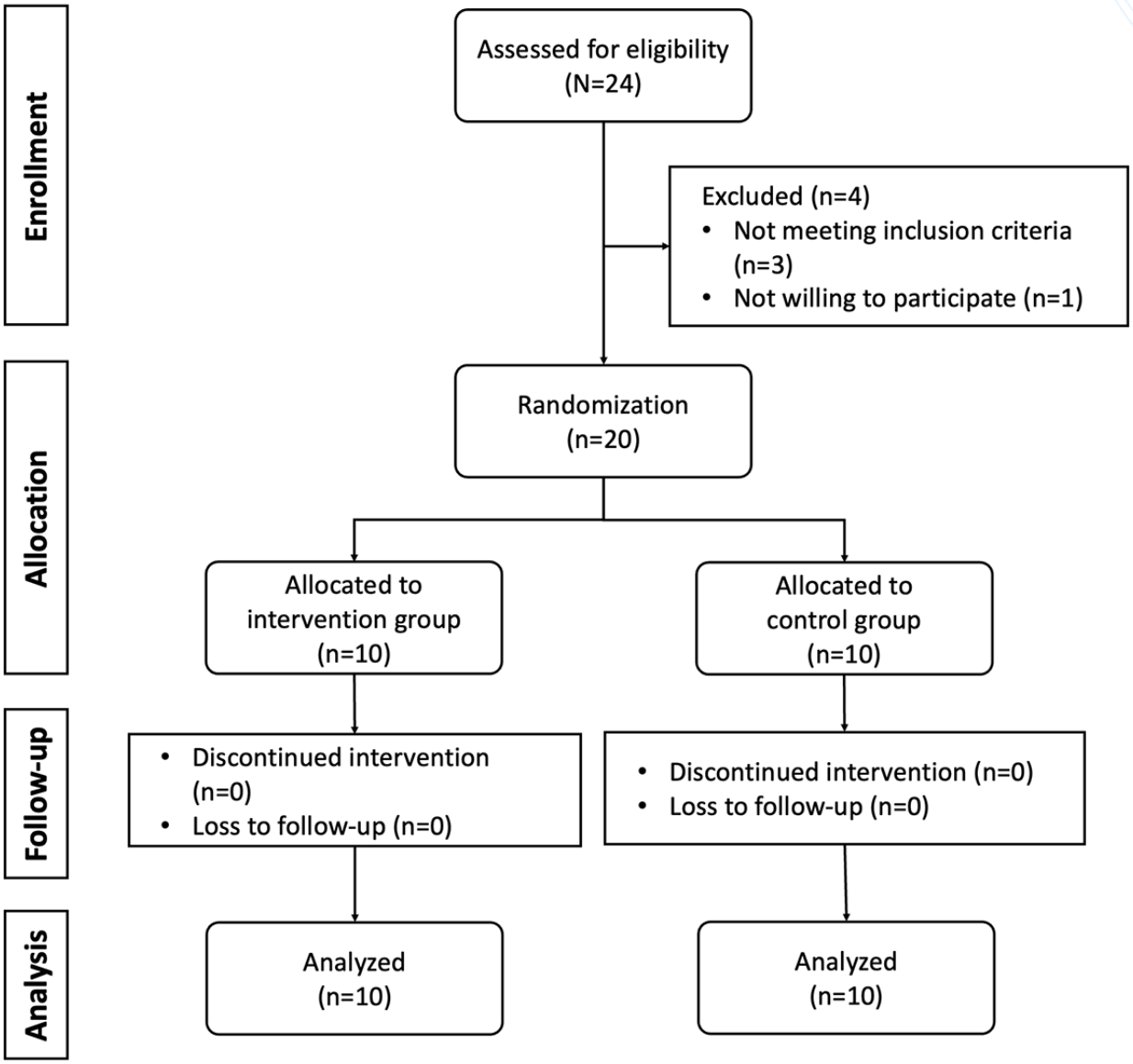
Recruitment

Recruitment for this study commenced in September 2023 and concluded in December 2023. We initiated participant selection by prescreening people with MCI from previous studies. Our collaboration with the Jing Shi Psychology Centre was pivotal in reaching out to potential participants. The center provided these individuals with detailed information about the study, including providing informed consent forms and an overview of the activities involved. A total of 20 participants enrolled in the study, with 10 (50%) randomized to the experimental group and 10 (50%) to the control group (Figure 1). The communication between the participants and staff in the partner institution before the formal research was aimed at gauging their interest in participating in our research. In addition, the center assisted in the recruitment process by conducting screenings of people with MCI using the Montreal Cognitive Assessment–Chinese (MoCA-C) questionnaire [50]. This meticulous and collaborative approach ensured that we recruited participants who met the study criteria and were well informed and motivated enough to contribute to our research on improving the lives of people with MCI. Participants were asked to complete the scales within a week.

The inclusion criteria for recruiting people with MCI were being clinically diagnosed with MCI using the MoCA-C screening tool, residing in community-dwelling complexes in Beijing, being aged ≥65 years, and having no significant visual or hearing impairments. In addition, participants were required to have adequate reading skills, especially for digital interfaces. Exclusion criteria were in place for individuals who had neurological conditions such as stroke or brain injury as these could affect daily functioning and impact study outcomes.

XSL•FO
RenderX

Figure 1. Flowchart for selecting the study participants.



Intervention: Huiyou Group

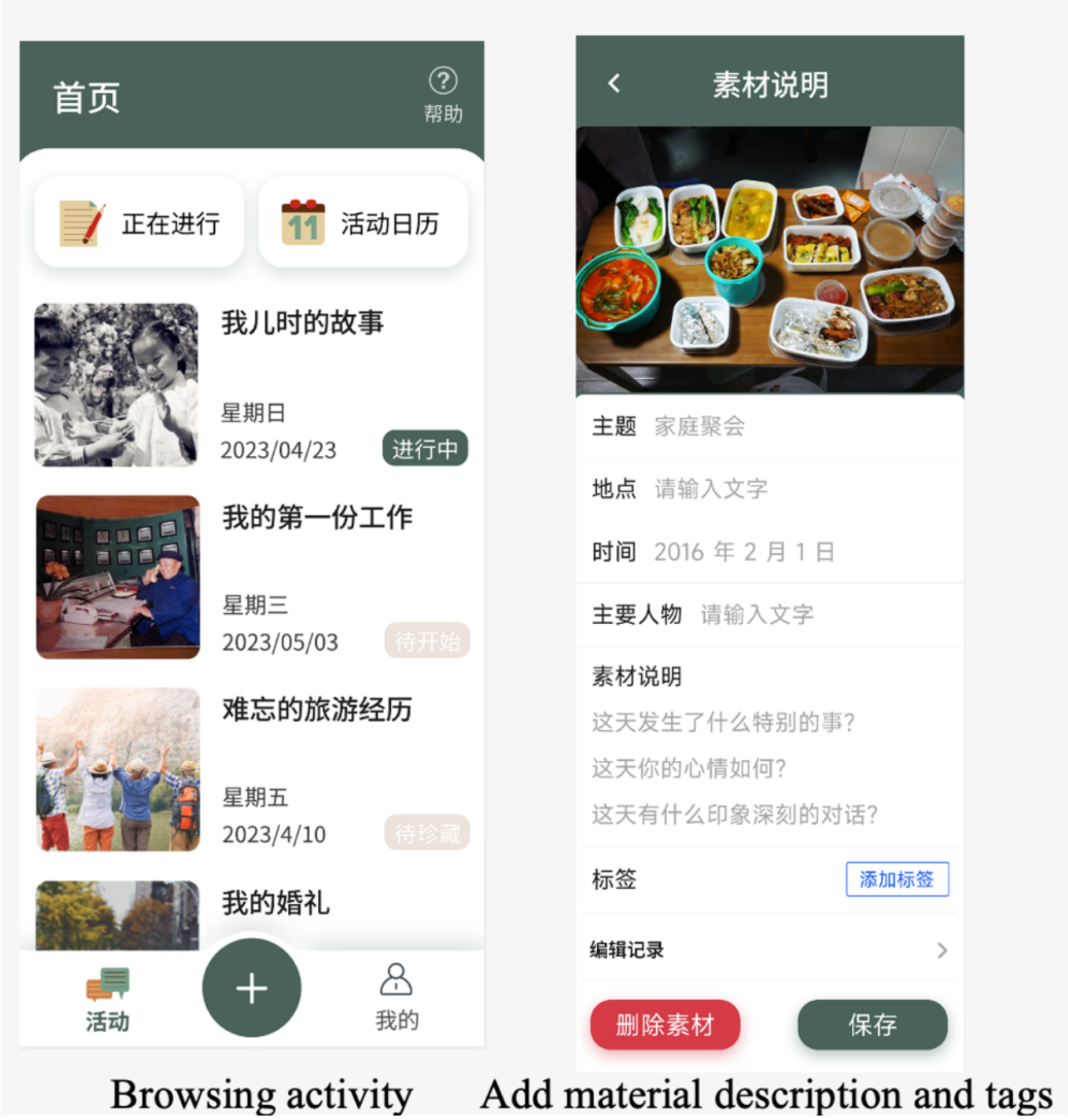
Huiyou is an Android-based mobile app designed to support digital storytelling co-designed with therapists and people with MCI [48]. Participants engage with weekly activity themes, where they are prompted to prepare and upload photos or videos relevant to 1 of 3 selected themes. Users can also edit keywords and categorize them for better organization. Once an activity begins, participants are grouped into small teams of 3 to 4 members for a storytelling session. A tablet serves as the interface for the storyteller, displaying prepared keywords to aid in their narrative, whereas a computer screen acts as the display terminal for the audience. Currently, owing to technical constraints, the display technology is managed through Keynote (Apple Inc), with researchers predownloading and processing materials to ensure smooth playback. Participants are given approximately 30 minutes before each activity to prepare their materials. The sessions are held weekly, with the timing determined by each group through discussions before each activity. Goal assessments are conducted at the end of the second and fourth weeks, with detailed records kept of goal completion, notes, and the duration of each story shared.

Figures 2 and 3 illustrate various functionalities of the Huiyou mobile app designed to enhance user interaction and memory recording. The workflow of Huiyou is as follows. Participants sign up for topics of interest, and we assist them in scheduling a sharing time. Huiyou assigns 3 missions, providing several guiding questions to help participants think about and prepare their materials in advance. During preparation, participants can type in memos and tips. During the sharing session, the pictures are displayed to group members, and participants can refer to their memos. This process ultimately creates a shared storytelling memory. The co-design and usability study of Huiyou provides a more detailed description of the app [48]. The first screenshot in Figure 2 shows the Browsing Activity page, offering a user-friendly interface for users to navigate through different activities seamlessly. The second screenshot in Figure 2 highlights the functionality for users to add material descriptions and tags, enabling them to organize and categorize their entries for better accessibility and management. The first

XSL•FO
RenderX

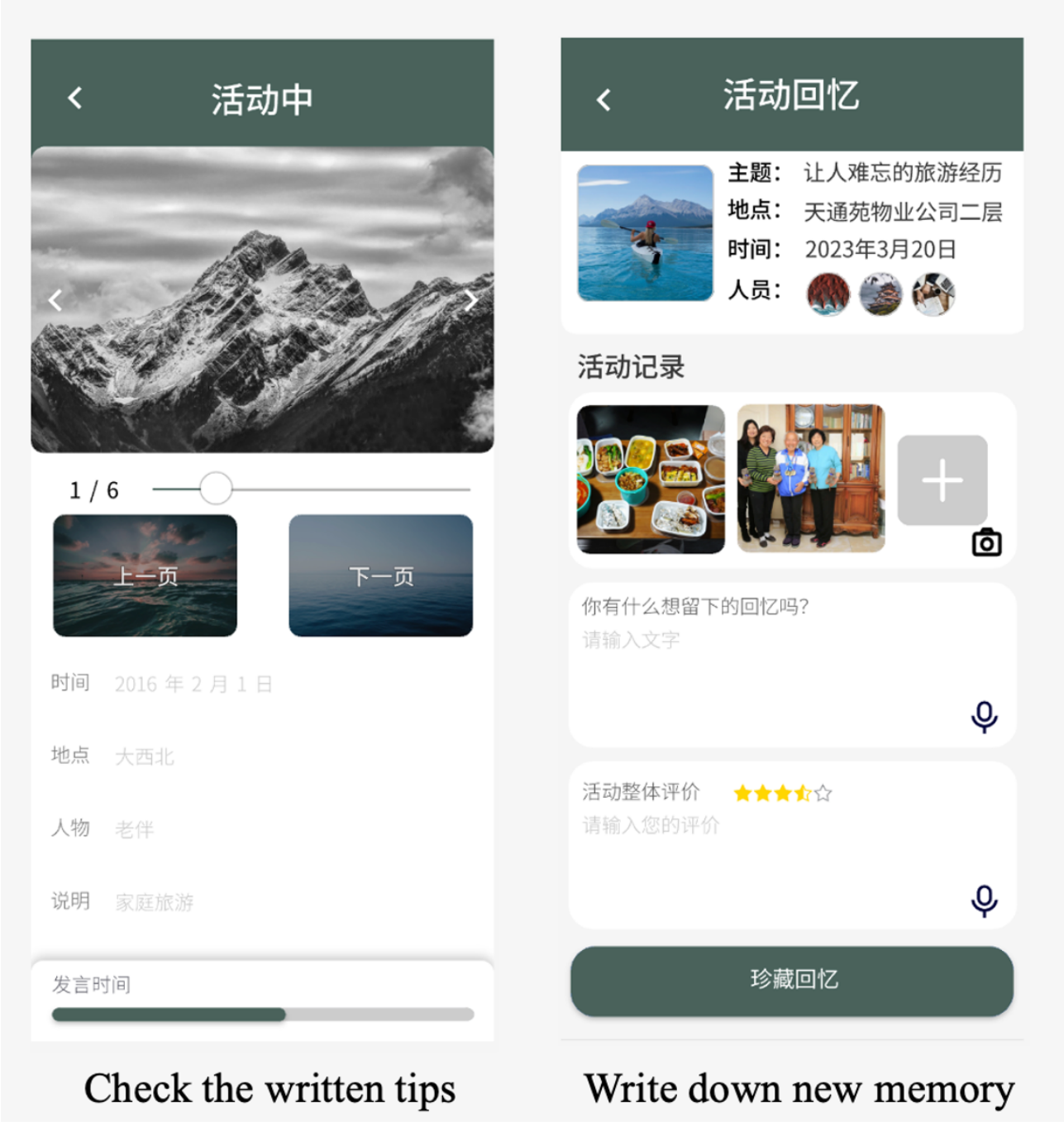
image in Figure 3 shows the Check the Written Tips section, where users can access practical advice on using the app effectively. Finally, the Write Down New Memory feature, which provides a structured format for users to document new memories with prompts and text fields to guide detailed recollections, is shown in the second screenshot of Figure 3.

Figure 2. Screenshots of Huiyou—memory collection.



XSL•FO
RenderX

Figure 3. Screenshots of Huiyou—story sharing.



Waitlist Control Group

The participants in the waitlist control group were involved in all evaluations at the beginning (T0) and end (T1) of the study. After completing the final evaluation (T1), they were offered the opportunity to learn about the Huiyou program, which they could potentially use in future research. This arrangement ensured that all participants received equal information about the program's features and potential benefits, although they did not use the program during the study period [49].

Measures

Assessments were completed at community locations in Beijing, China, at baseline (T0) and 4 weeks (T1). Baseline participant characteristics included age, sex, educational level, and MoCA-C score. This paper summarizes the assessments for social participation (Assessment of Life Habits [LIFE-H]; 77 items), social connectedness (Social Connectedness Scale–Revised [SCS-R]; 20 items), self-efficacy (General Self-Efficacy Scale [GSES]; 10 items), and subjective sense of happiness (Subjective Well-Being Scale [SWBS]; 20 items), as Table 1 shows. Further details of the study design can be found in the protocol paper [49].

XSL•FO
RenderX

Table 1. Study assessments and outcome measures, which were conducted on all participants at baseline and after 4 weeks.

Assessment	Outcome
Social participation	• The LIFE-H[a] (77 items): indoor activity, outdoor activity, interpersonal relationships, community life, and amusement and recreation
Goal attainment of storytelling	• Duration of story sharing • Number of key events in the story • Number of photos collected • Number of notes made on the photos
Social connectedness	• SCS-R[b] (20 items)
Self-efficacy	• GSES[c] (10 items): action self-efficacy and coping self-efficacy
Subjective sense of happiness	• SWBS[d] (20 items): interpersonal adaptation experience, mental health experience, family atmosphere experience, psychological balance experience, physical health experience, target value experience, social confidence experience, growth and progress experience, contented and abundant experience, and self-acceptance experience

[a]LIFE-H: Assessment of Life Habits.

[b]SCS-R: Social Connectedness Scale–Revised.

[c]GSES: General Self-Efficacy Scale.

[d]SWBS: Subjective Well-Being Scale.

All the scales used have validated Chinese versions, and we used these versions in this study. Social participation was evaluated using the Chinese version of the LIFE-H, a validated and reliable self-report questionnaire comprising 77 items [51]. This instrument assesses various aspects of daily living, including indoor activities, outdoor activities, interpersonal relationships, community life, and amusement and recreation. Goal attainment in storytelling was measured by recording several key metrics: the duration of story sharing, the number of key events included in each story, the number of photos collected, and the number of annotations made on these photos. Participants set baseline targets on a 10-point scale before commencing the activity and revisited these targets during the second and fourth weeks of the intervention to track progress and achievement [52]. Social connectedness was assessed using the SCS-R, a 20-item, well-established self-report questionnaire known for its validity and reliability in measuring individuals' sense of belonging and connection to others [53]. Self-efficacy levels were measured using the GSES, a validated 10-item questionnaire that evaluates both action self-efficacy and coping self-efficacy, reflecting the participants' confidence in managing and executing tasks as well as handling challenging situations effectively [54,55]. The subjective sense of happiness was evaluated using the SWBS, another reliable and valid self-report instrument consisting of 20 items. This scale captures a wide range of experiences related to well-being, including interpersonal adaptation, mental health, family atmosphere, psychological balance, physical health, goal achievement, social confidence, personal growth, contentment, and self-acceptance [56]. Interpersonal adaptation is the process of adjusting one's behavior and communication style to better fit in with others in social interaction [53].

In addition to the structured quantitative assessments, qualitative observations were conducted during the field-testing sessions to capture participants' engagement and social interaction patterns. Researchers took brief field notes after each session, documenting participants' verbal and nonverbal responses, interactions, and levels of engagement. These observations were intended to supplement the quantitative findings by offering contextual insights into the participants' experiences with the intervention [57]. A full systematic analysis of these qualitative data was outside the scope of this paper and has been reported separately in a companion article focusing specifically on the qualitative findings. In this paper, qualitative observations are referred to selectively to support the interpretation of key quantitative outcomes.

Statistical Analysis

Quantitative data analysis was conducted in the JASP software [58]. The analysis focused on key metrics related to the digital storytelling intervention: duration of story sharing, number of key events in the story, number of photos collected, and number of notes made on the photos. Participants submitted subjective reports at 3 junctures: baseline, the 2-week follow-up, and the 4-week follow-up. For each metric, we followed the Bangor Goal-Setting Interview Manual [52] to calculate the mean and variance of and changes in the outcomes reported by participants in addition to objective measures such as the actual duration of story sharing, the total photos collected, and the tally of notes made. To standardize the scoring across various metrics, a normalization method was applied: the maximum number of photos (n=9) was set as a perfect score of 10, with each participant's count proportionately scaled. Similarly, for the duration of story sharing and the number of notes, the highest recorded value was established as the benchmark for a full score of 10, with other values adjusted accordingly. Regarding the number of key events in the story, following transcription, our research team conducted a quantitative assessment to enumerate significant narrative elements within each shared story. Finally, to attain a comprehensive insight into the participants'

XSL•FO
RenderX

experiences and the intervention's impact, thematic coding was applied to analyze the data from the concluding user interviews.

To evaluate the intervention's impact, we calculated a standardized change rate for each outcome measure:



This calculation standardizes each participant's change relative to their baseline, providing a more sensitive measure of group differences. Independent *t* tests were conducted in JASP to compare these change rates between the intervention and control groups, with *P* values and effect sizes with 95% CIs reported. An α level of .05 was set for all 2-tailed tests. No interim analyses or stopping rules were applied, as the trial involved a small sample and posed minimal risk to participants. Participants and facilitators were not blinded due to the nature of the intervention. However, outcome assessors and data analysts were blinded by using coded group labels and de-identified data.

All randomized participants (N=20) were included in the final analysis and analyzed in the groups to which they were originally assigned, following an intention-to-treat approach. No participants were lost to follow-up. Minor missing data identified during data checking were subsequently completed through verification with source records, and no imputation was required. No subgroup or sensitivity analyses were prespecified or conducted, given the exploratory nature of the study. Harms were described narratively, as no adverse events were reported.

Harms Assessment

Given the non-invasive and low-risk nature of the digital storytelling intervention, no formal adverse event monitoring was implemented. Harms were defined broadly as any negative physical, emotional, or psychological responses associated with participation in the intervention. These were assessed non-systematically through participant self-reports and informal observations made by facilitators during and after each session. No adverse events or participant-reported harms were observed or reported throughout the trial.

Results

Sample Characteristics

A total of 20 participants were recruited and randomly assigned to either the intervention group or the waitlist control group (n=10, 50% per group). Demographic and cognitive assessments showed no significant differences between the groups, supporting the randomization protocol (Table 2). Participants were randomly assigned to the intervention or control group in a 1:1 ratio using simple randomization. The allocation sequence was generated by an independent researcher using Excel's RAND function and implemented via sequentially numbered, opaque, sealed envelopes. The personnel responsible for enrolling participants and assigning interventions did not have access to the allocation sequence, ensuring allocation concealment. Specifically, the difference in age was not statistically significant (*P*=.23), nor were differences in sex distribution (*P*=.29), educational attainment (*P*=.41), or MoCA-Ca scores (*P*=.20). Participants averaged 69.7 (SD 3.21) years of age, with no significant difference between groups (mean 68.80, SD 2.61 in the intervention group vs mean 70.60, SD 3.49 in the waitlist control group; *P*=.23). Most participants (16/20, 80%) were female, with sex distribution consistent across groups (*P*=.29). Educational levels and cognitive function (assessed using the MoCA-C) did not significantly differ between groups (*P*>.05 in all cases).

Table 2. Sample characteristics overall and by treatment condition (N=20).

Characteristic	Total	Intervention (n=10)	Waitlist control (n=10)	*P* value
Age (y), mean (SD; range)	69.7 (3.21; 65-76)	68.80 (2.61; 65-76)	70.60 (3.49; 65-74)	.23
Sex, n (%)				.29
Male	4 (20)	1 (10)	3 (30)	
Female	16 (80)	9 (90)	7 (70)	
Educational level, n (%)				.41
University	4 (20)	2 (20)	2 (20)	
High school	7 (35)	5 (50)	2 (20)	
Middle school	9 (45)	3 (30)	6 (60)	
MoCA-C[a] score, mean (SD)	23.25 (1.51)	23.7 (1.19)	22.8 (1.66)	.20

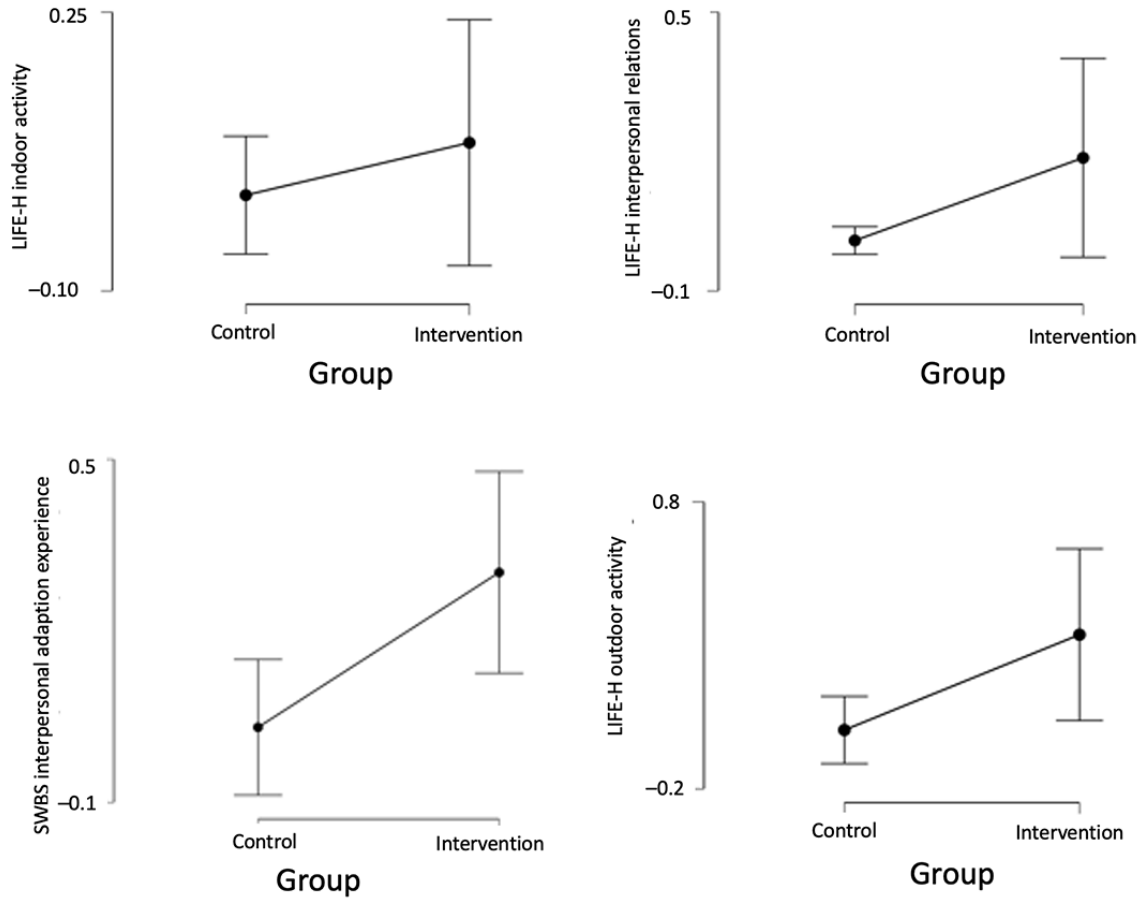
[a]MoCA-C: Montreal Cognitive Assessment–Chinese version.

Intervention Impact Analysis

According to statistical comparisons of change rates across key outcome measures (Multimedia Appendix 2), significant improvements were observed in outdoor activity on the LIFE-H (t_{18}=2.339; Cohen *d*=1.046; *P*=.03) and in interpersonal adaptation experience on the SWBS (t_{18}=2.884; Cohen *d*=1.290; *P*=.01), indicating that the intervention effectively enhanced participants' engagement in outdoor activities and interpersonal adaptation. In other measures, such as indoor activity on the LIFE-H (*P*=.40) and the GSES (*P*=.98), slight but nonsignificant improvements were noted. Figure 4 illustrates the change rates across outcome measures, with clear improvements for outdoor activity and interpersonal adaptation, whereas other measures showed overlapping trends, consistent with the nonsignificant *t* test results.

XSL•FO
RenderX

Figure 4. Selected outcome measure change rates—intervention versus control group. LIFE-H: Assessment of Life Habits; SWBS: Subjective Well-Being Scale.



Effect of Huiyou

To assess specific domains of social participation and subjective well-being, we used the LIFE-H, SCS-R, GSES, and SWBS. Pre- and postintervention *t* tests were conducted to compare group differences (Multimedia Appendix 3). Most measures showed no significant baseline differences between the experimental and control groups, confirming initial equivalence. Specifically, baseline comparisons showed no significant group differences in general LIFE-H (*P*=.14), indoor activity (*P*=.40), outdoor activity (*P*=.03), interpersonal relationships (*P*=.08), community life (*P*=.81), SCS-R (*P*=.59), GSES total (*P*=.98), action self-efficacy (*P*=.39), coping self-efficacy (*P*=.14), SWBS total (*P*=.26), interpersonal adaptation (*P*=.01), mental health (*P*=.42), family atmosphere (*P*=.41), psychological balance (*P*=.22), physical health (*P*=.21), target value (*P*=.21), social confidence (*P*=.46), and contented and abundant experience (*P*=.70). However, a significant baseline difference was observed in the self-acceptance experience on the SWBS, where a baseline difference (*P*=.007) was noted and considered in subsequent analyses.

Table 3 presents descriptive statistics for improvement rates across study measures. For LIFE-H categories, average improvement rates ranged from 0.045 (SD 0.156; general) to 0.171 (SD 0.353; outdoor activities), reflecting varied enhancements in daily living habits. The SCS-R showed a mean improvement of 0.098 (SD 0.192). Mental health experience on the SWBS demonstrated the largest variability (mean 0.448, SD 1.12; range −0.222 to 5), indicating individual differences in postintervention outcomes. Other subjective well-being domains, including interpersonal adaptation (mean 0.167, SD 0.247; *P*=.01) and psychological balance (mean 0.313, SD 0.514; *P*=.22), also showed positive trends; however, these results were not statistically significant.

Multimedia Appendix 2 shows the effect sizes (Cohen *d*) of improvement rates across various outcome measures, visually highlighting the magnitude of changes observed between the intervention and control groups. The red dashed line at 0.80 indicates the threshold for a large effect. As shown, interpersonal adaptation experience and outdoor activity demonstrated large effect sizes (Cohen *d*=1.29 and 1.05, respectively), suggesting substantial improvements attributable to the intervention. Several other outcomes, including overall LIFE-H, SCS-R, and SWBS scores, showed medium to small effects, whereas a few measures exhibited minimal or negative effects. This visual summary supports the interpretation that, while the intervention had varied impacts, it significantly enhanced certain aspects of social participation and well-being in people with MCI.

The remaining variables, including different facets of life habits (LIFE-H), general self-efficacy (GSES), social connectedness (SCS-R), and various dimensions of subjective well-being

XSL•FO
RenderX

(SWBS), did not show statistically significant changes. Small to moderate effect sizes were observed, suggesting that, while certain domains such as outdoor activity and interpersonal adaptation showed meaningful improvements, other areas such as community life, self-efficacy, and overall subjective well-being exhibited minimal to no change.

Table 3. Descriptive statistics and improvement rates in study measures.

	Improvement rate, mean (SD; range)	*t* test (*df*)	*P* value	Cohen *d*
LIFE-H[a] general	0.045 (0.156; −0.273 to 0.597)	1.566 (18)	.14	0.701
Indoor activity	0.053 (0.168; −0.181 to 0.653)	0.87 (18)	.40	0.389
Outdoor activity	0.171 (0.353; −0.2 to 1.414)	2.339 (18)	.03	1.046
Interpersonal relationships	0.097 (0.227; −0.02 to 0.923)	1.864 (18)	.08	0.834
Community life	0.06 (0.444; −0.56 to 1.778)	0.242 (18)	.81	0.108
SCS-R[b]	0.098 (0.192; −0.174 to 0.533)	−0.548 (18)	.59	−0.245
GSES[c]	−0.006 (0.174; −0.275 to 0.417)	−0.027 (18)	.98	−0.012
Action self-efficacy	0.012 (0.188; −0.278 to 0.357)	−0.882 (18)	.39	−0.394
Coping self-efficacy	−0.015 (0.227; −0.4 to 0.5)	1.562 (18)	.14	0.698
SWBS[d]	0.109 (0.189; −0.103 to 0.603)	1.16 (18)	.26	0.519
Interpersonal adaptation experience	0.167 (0.247; −0.125 to 0.667)	2.884 (18)	.01	1.29
Mental health experience	0.448 (1.12; −0.222 to 5)	0.83 (18)	.42	0.371
Family atmosphere experience	0.05 (0.185; −0.125 to 0.714)	0.845 (18)	.41	0.378
Psychological balance experience	0.313 (0.514; −0.2 to 2)	1.259 (18)	.22	0.563
Physical health experience	0.444 (0.94; −0.2 to 4)	−1.307 (18)	.21	−0.585
Target value experience	0.096 (0.247; −0.125 to 0.714)	1.317 (18)	.21	0.589
Social confidence experience	0.059 (0.256; −0.4 to 1)	0.762 (18)	.46	0.341
Contented and abundant experience	0.01 (0.238; −0.583 to 0.571)	−0.391 (18)	.70	−0.175
Self-acceptance experience	0.1 (0.247; −0.2 to 0.714)	−0.283 (18)	.78	−0.127

[a]LIFE-H: Assessment of Life Habits.

[b]SCS-R: Social Connectedness Scale–Revised.

[c]GSES: General Self-Efficacy Scale.

[d]SWBS: Subjective Well-Being Scale.

Goal Attainment of Digital Storytelling

Table 4 presents the goal attainment ratings at baseline and the 2- and 4-week follow-ups, alongside statistical comparisons across various measures related to digital storytelling activities among participants. The measures include the duration of story sharing, the number of key events in the story, the number of photos collected, and the number of notes made on the photos, each assessed through participant and objective ratings. At baseline, the duration of story sharing was rated at an average of 4.8 (SD 1.23), which increased to 9.2 (SD 0.79) at the 4-week follow-up. Similarly, the number of key events showed an increase from a baseline rating of 5.5 (SD 1.35) to 9.5 (SD 0.528), and the number of photos collected rose from an initial 5.2 (SD 1.39) to 9.7 (SD 0.48). Finally, the number of notes made on the photos increased from 5.2 (SD 0.92) to 9.7 (SD 0.48). These results demonstrate a significant improvement in all measured aspects of digital storytelling activities, reflecting enhanced participant engagement, storytelling complexity, and interaction with digital content over the course of the study.

Table 4 presents the mean attainment and changes in story sharing duration, number of key events, number of photos collected, and notes on the photos at baseline and the 2- and 4-week follow-ups among 50% (10/20) of the participants in the intervention group. The table illustrates a clear improvement in participants' performance from the initial assessment to subsequent follow-ups. Substantial improvements were observed across all digital storytelling performance metrics, including longer storytelling durations, greater number of key events, more photos collected, and more notes recorded.

XSL•FO
RenderX

Table 4. Goal attainment ratings at baseline and the 2- and 4-week follow-ups and statistical comparisons.

Measure	Participant rating (1-10), mean (SD)			Researcher rating (1-10), mean (SD)
	Baseline (T0)	2 wk	4 wk (T1)	
Duration of story sharing	4.8 (1.23)	8.7 (1.34)	9.2 (0.79)	6.8 (1.93)
Number of key events in the story	5.5 (1.35)	8.4 (1.43)	9.5 (0.53)	5.7 (1.06)
Number of photos collected	5.2 (1.39)	8.6 (1.26)	9.7 (0.48)	7.3 (1.49)
Number of notes on the photos	5.2 (0.92)	8.3 (1.77)	9.7 (0.48)	7.3 (1.49)

In addition to the quantitative findings, qualitative observations during the field-testing phase highlighted noticeable changes in participants' engagement and interaction patterns. Initially, several participants appeared hesitant to share their stories; however, by the third and fourth sessions, many became more active, showing greater willingness to discuss personal memories and respond to others' narratives. Emotional expressions such as laughter, nostalgia, and mutual encouragement were frequently observed during group discussions. For example, one participant shared the following—"We chatted away, made new friends. Without this activity, I wouldn't have met everyone, you know" (group 1; P1)—reflecting the social bonds that developed through the sessions. Another participant remarked the following—"I am very interested in this activity as it is a completely new experience for me. I have never done anything like it before" (group 2; P1)—highlighting the novelty and engagement stimulated by the Huiyou app.

Most participants also demonstrated increasing independence in navigating Huiyou, requiring less assistance over time. One participant expressed the following:

> *I managed to do everything by myself and didn't ask for help from anyone.* [Group 1; P2]

Furthermore, the intervention appeared to strengthen interpersonal relationships beyond the sessions, with a participant noting the following:

> *I've had more interactions with my husband now; we reminisce about the old times together.* [Group 1; P3]

These observations provide supplementary evidence of improved social connectedness, storytelling engagement, and digital competency among participants with MCI.

No subgroup or sensitivity analyses were prespecified or conducted in this study. Given the small sample size and the exploratory nature of the trial, the analysis was limited to evaluating primary and secondary outcomes at the group level.

No harms or unintended events were reported in either the intervention or waitlist control group throughout the study period. Participants did not report any emotional distress, discomfort, or adverse reactions related to the digital storytelling sessions or study procedures. Facilitators also did not observe any concerning behaviors or incidents during intervention delivery.

Discussion

Principal Findings

The principal findings of this pilot randomized controlled trial investigating the effects of the mobile storytelling app Huiyou on social participation and storytelling ability among people with MCI are compelling. This study demonstrated that the experimental group, which engaged with the Huiyou app, exhibited a significant improvement in outdoor activities compared to the control group. This was reflected by a Cohen *d* value of 1.046, indicating a large effect size, suggesting that the intervention significantly increased the frequency and possibly the quality of outdoor social interactions among participants. In addition, there was a significant enhancement in interpersonal adaptation experiences, with a Cohen *d* of 1.290, also indicating a substantial effect. However, although there were strong effects observed for the LIFE-H total score, SCS-R total score, SWBS total score, interpersonal relationships, psychological balance experience, physical health experience, and target value experience, these improvements did not reach statistical significance. This may be due to the small sample size.

Despite the lack of statistically significant changes across all measures, the large effect sizes observed suggest meaningful improvements that warrant further investigation. These findings underscore the potential value of the Huiyou app in enhancing social participation among individuals with MCI but also suggest that larger-scale studies are necessary to fully assess the app's impact. The intervention effectively increased the duration of story sharing, the number of key events detailed in the stories, the volume of photos collected, and the number of notes made on these photos. These improvements indicate an increase not only in participant engagement but also in the complexity and richness of the narratives shared by the participants. This multifaceted enhancement highlights the efficacy of the intervention in enhancing participants' storytelling skills and their overall experience with the digital platform. It reflects a broader impact on their engagement and the quality of their interactions. This improvement suggests that the intervention successfully boosted participants' ability to adapt to social interactions, enhancing their social integration and reducing feelings of isolation. These outcomes underline the potential of tailored digital storytelling interventions to enrich social engagement and overall quality of life for people with MCI.

XSL•FO
RenderX

Comparison With Prior Work

Social Participation and Storytelling Engagement

In our discussion regarding the Huiyou intervention, we noted a significant increase in outdoor activities after the intervention, echoing the findings of Ryu and Heo [59]. They noted a positive correlation between life satisfaction and physical activities, suggesting the various benefits of outdoor engagement. The significant improvement in outdoor activities observed in the experimental group of this study compared to the control group highlights that the intervention led to increased participation in these activities by older individuals. The strong effect size (Cohen *d*=1.046) indicates a substantial impact of the intervention, demonstrating its effectiveness in enhancing outdoor activity among participants. This notable increase not only supports the efficacy of our intervention but also suggests that engaging in outdoor activities could be particularly beneficial for improving the overall well-being and social engagement of people with MCI. This finding is crucial for developing future interventions aimed at increasing physical activity and social interactions in this population. Our structured weekly meet-ups as part of the intervention encouraged active participation in the external environment, as highlighted in the study by Ryu and Heo [59]. Their research expands on this connection by linking outdoor engagement not only to physical and psychological well-being but also to increased life satisfaction and health perception. This indicates that the intervention's impact on outdoor activities extends beyond immediate physical benefits, potentially influencing overall life satisfaction and health perception among participants, thus underlining the vital role of structured outdoor activities in improving quality of life for people with MCI. Furthermore, the holistic advantages of outdoor activities, bolstered by structured interventions, encompass physical health, optimism, and cultural engagement. This comprehensive approach to wellness emphasizes the importance of environmental factors in promoting outdoor activities, alongside the findings of Ryu and Heo [59] and Chang [60]. These studies highlight the multifaceted value of outdoor engagement in promoting health and well-being among people with MCI and stress the importance of both structured interventions and supportive environments in enhancing outdoor activity participation. The prototype developed in this study played a significant role in enhancing outdoor activities, particularly reflected in the notable differences in LIFE-H outdoor activity scores after the intervention. This aligns with existing research indicating the significant benefits of structured outdoor interventions, particularly those conducted in natural settings, for enhancing the physical and mental health of older adults. These studies highlight that such interventions can lead to improved physical activity, reduced stress, and better overall well-being [61]. These findings suggest that similar programs focusing on outdoor activities could be valuable in various contexts to promote a healthier lifestyle and greater social engagement. Specifically, rehabilitation programs that incorporate outdoor mobility have shown success in increasing activity levels, outdoor mobility, and endurance for older adults recovering from illness or injury. This suggests that incorporating structured outdoor activities into rehabilitation or wellness programs can have a notable impact on improving both physical endurance and quality of life for people with MCI [62]. Coming out of the room is a preliminary step toward engaging in outdoor activities. It is a crucial aspect of encouraging physical activity and social interaction, especially for people with MCI. By fostering a motivation to leave the room, interventions can lay the groundwork for enhanced outdoor activity, leading to improved health outcomes and overall well-being. Given this understanding, it is important to design interventions that address the barriers that may prevent individuals from coming out of their rooms. This might involve creating structured programs that are appealing and easily accessible, offering social incentives, and providing support to overcome potential anxiety or hesitation. Such efforts can increase the likelihood of engaging in outdoor activities, which in turn can lead to better physical health and more social connection.

Our findings on the Huiyou app underscore its substantial positive effects on interpersonal adaptation among participants with MCI. The significant improvement in interpersonal adaptation experience, highlighted by activities such as generating and sharing material, aligns with existing research on psychological interventions for MCI or dementia [63]. The intervention notably increased participants' scores compared to those in the control group, which suggests that the digital storytelling approach not only enhances interpersonal skills but also supports people with MCI in adapting more effectively to social interactions. The strong effect size (Cohen *d*=1.290) emphasizes the impact of the intervention, illustrating its potential to significantly improve social adaptability in individuals with MCI. These findings advocate for the integration of storytelling elements into interventions as a powerful mechanism to enhance social capabilities and address some cognitive impairments associated with social challenges. Interventions that focus on building social connections and promoting social participation have been shown to improve interpersonal adaptation and reduce feelings of isolation [64]. Technology-based interventions such as digital applications that support reminiscence activities can further enhance memory recall and strengthen social interactions [20]. This aligns with our findings that the "generate material" phase, involving the searching and sharing of photos, facilitated increased interactions with friends and family, thereby improving social bonds.

Furthermore, Bentley et al [37] highlighted the importance of connecting with places to strengthen social contact, promoting a unified system with friends and family to share places and stories. This concept dovetails with our findings, wherein participants engaged in deeper reminiscence with their spouses to recall more detailed memories. This process not only nurtured more positive interactions but also effectively reduced negative exchanges, highlighting the value of digital storytelling interventions in cultivating meaningful interpersonal relationships. The convergence of our results illustrates the multifaceted impact of digital storytelling interventions on enhancing social interaction and communication. Through facilitating the creation and sharing of content, such interventions promote active social participation and

XSL•FO
RenderX

reminiscence, contributing to improved interpersonal relationships among participants and their social circles.

Moreover, the study's intricate assessment of digital storytelling activities revealed considerable improvements in engagement, complexity of storytelling, and interaction with digital content, marking a significant stride in enhancing participants' quality of life. Digital storytelling interventions offer several benefits for older adults, including promoting mental health, building meaningful community connections, achieving digital literacy, mitigating negative ageism, and enhancing intellectual ability [65]. In addition, we found that these interventions can support self-expression and language expression, providing a platform for older adults to share their stories, which encourages active communication and engagement with others. By engaging in digital storytelling interventions, older adults can develop new skills, build social networks, and gain a sense of accomplishment, contributing to improved overall well-being. Given these observations, incorporating digital storytelling into rehabilitation and engagement programs could effectively improve participants' quality of life through enhanced interaction and complexity in storytelling while also strengthening social engagement and psychological well-being. The Huiyou app notably enhanced participants' storytelling engagement, as evidenced by increased narrative complexity and more active interaction with digital content. The structured sessions and targeted prompts encouraged deeper reflection and participation in storytelling activities. Furthermore, the process of generating and sharing stories fostered richer social interactions within participants' networks, highlighting a positive potential toward improved social connectedness. Complementary structured outdoor activities also contributed to significant improvements in outdoor activity engagement, supporting both physical and psychological well-being. Overall, the findings suggest that structured interventions integrating digital storytelling, social participation, and outdoor activities can meaningfully enhance quality of life and social well-being for people with MCI. Although a detailed qualitative analysis is reported separately, key observations during the intervention reinforce the quantitative results. The progressive increase in storytelling engagement, emotional sharing, and group interaction suggests that Huiyou successfully fostered a supportive social environment for people with MCI. These behavior changes align with previous research emphasizing the role of peer interaction and shared memory activities in promoting social participation. Furthermore, the gradual improvement in participants' digital literacy and confidence with the app underscores the potential for appropriately designed technology to support autonomy and reduce technology-related anxiety among older adults with cognitive impairment.

Self-Efficacy and Measurement Challenges

Our study examined the lack of significant changes in self-efficacy among older adults engaging in physical and social activities. It shed light on several critical factors that merit consideration. First, the intervention's duration was a potential factor influencing these outcomes. Existing literature supports the notion that short-term interventions may not be sufficiently impactful to alter self-efficacy levels significantly, particularly in older adults who may have long-held beliefs about their capabilities [66]. Self-efficacy develops and is influenced over time through mastery experiences, vicarious experiences, verbal persuasion, and interpretations of physiological and emotional states [67]. Consequently, the relatively brief span of our intervention might not have provided ample opportunities for these processes to effectuate a meaningful change in self-efficacy levels. Second, the challenge of detecting significant changes in self-efficacy is compounded by the high baseline self-efficacy scores observed among participants and the possible insensitivity of the measurement tools used. This observation aligns with findings of other studies that highlight the difficulty in measuring subtle shifts in self-efficacy among populations with already high levels of this construct [66,68]. The necessity for more refined measurement tools is emphasized, suggesting a need for instruments that can discern minor yet impactful changes in self-efficacy. Such tools would enable a more accurate assessment of intervention effects, particularly in contexts in which initial self-efficacy levels are high. Third, the stress associated with participation in activities represents another critical consideration. Our findings indicate that stress may undermine self-efficacy beliefs, resonating with previous research indicating that stress and anxiety can negatively impact one's confidence in performing future activities [69,70]. This means that incorporating strategies within interventions is crucial to managing stress and enhancing psychological resilience among participants. Thus, tailoring interventions to account for initial psychological states and using stress management techniques could be pivotal in supporting and possibly improving self-efficacy beliefs.

Given the cognitive and linguistic difficulties among people with MCI [71-73], measurement tools such as the SCS-R should be simplified for accurate evaluation. The linguistic complexity of some measurement tools, particularly the double-negative phrasing found in items of the SCS-R, may have posed cognitive challenges for participants with MCI. While this limitation was acknowledged, future research should pilot simplified or adapted versions of such instruments that are validated for populations with cognitive impairment to ensure accuracy and comprehension. Given these considerations, there is a pressing need for more straightforward or positively framed items in future studies, especially when assessing populations with cognitive impairments such as MCI. Simplifying the language used in measurement tools not only accommodates the cognitive profiles of these populations but also enhances the reliability and accuracy of the data collected.

While not all measures showed significant changes, certain areas such as physical health and self-acceptance exhibited positive trends toward improvement. This observation suggests that prototype-based interventions could contribute to a broader impact on these outcomes. Studies have shown that outdoor physical activity interventions in green and blue spaces can improve health outcomes, including well-being, mood, and physical performance [74]. Although our study indicated that structured activities play a pivotal role in enhancing outdoor activity engagement, the scope of these findings suggests the need for more extensive studies to validate these trends and determine the potential for these interventions to improve health outcomes across a wider range of indicators [75].

XSL•FO
RenderX

The results from the digital storytelling intervention suggest that further exploration in this field could lead to improved strategies for enhancing social participation and quality of life for people with MCI. The success of these interventions underscores the need to develop innovative approaches that effectively address the unique challenges faced by people with MCI.

Limitations

Primarily, the pilot nature of this study, with 10 participants in both the intervention and control groups, serves as a foundational step, limiting the generalizability of our findings. The modest sample size also restricts our ability to detect significant differences across some outcome measures, underscoring the need for expanded research to validate these initial results in a future study. This study was conducted as a pilot trial to explore the feasibility and preliminary effects of a digital storytelling intervention for people with MCI. Although the study received ethical approval from the relevant institutional review board, it was not prospectively registered. At the time of initiation, trial registration was not required for minimal-risk behavioral interventions under institutional policy. We acknowledge the absence of prospective registration as a limitation and recommend that future confirmatory trials undergo prospective trial registration in accordance with best practices. Given the pilot nature of this study (N=20), the statistical power was limited, and the findings should be interpreted with caution. However, the promising trends observed in this preliminary investigation justify the design and implementation of a larger, fully powered randomized controlled trial to validate the intervention's effectiveness and enhance generalizability to a broader population of people with MCI. Although randomization was used, a notable baseline difference in self-acceptance scores (SWBS) was observed between groups. This may have introduced bias into postintervention comparisons related to this variable. Future studies should consider stratified randomization or statistical adjustment to better control for such baseline differences. Concerning the sex distribution, our study observed a lower participation rate among men. Rather than reflecting a bias in recruitment or willingness to participate, this trend aligns with broader observations that, in China, men tend to engage less in collective activities compared to women. This aspect provides a cultural context to our findings and suggests avenues for further exploration into sex dynamics in social participation. A crucial point of consideration is the nature of the participants recruited through institutions, who tended to be more active or had previously engaged with psychological services. This selection bias suggests that the intervention's preliminary findings may not fully represent the broader population of older individuals with MCI, particularly those who are less active. Recognizing this, there is a pronounced need for future studies to delve into the intervention's effects on a more diverse and possibly less active older adult cohort, ensuring a comprehensive understanding of the intervention's benefits across varying levels of initial activity and engagement. Addressing feasibility concerns, the potential commuting barriers highlighted the importance of local community engagement in facilitating participation. Organizing activities within participants' communities could enhance the intervention's accessibility and reduce logistical barriers to participation. We could also use an online platform for digital storytelling to support participants who live far away from the group members.

Moreover, the scalability of the Huiyou intervention requires careful consideration. Technical barriers such as digital literacy gaps among older adults, limited access to compatible devices, and challenges in maintaining user engagement over time could hinder broader implementation. Contextual challenges, particularly in low- and middle-income settings, include resource constraints and insufficient community infrastructure to support ongoing group activities. These considerations are crucial for ensuring that digital storytelling interventions such as Huiyou can be effectively scaled to benefit wider populations of people with MCI.

Conclusions

The Huiyou intervention has shown promising effects on the well-being of people with MCI, significantly enhancing outdoor activities and interpersonal adaptation experiences. Notable improvements were observed in participants' engagement with digital storytelling, contributing to more complex narrative constructions. Consistent with existing research, these outdoor activities were found to contribute to both physical and psychological health, potentially affecting overall life satisfaction. Enhanced social connections were facilitated through digital content sharing, boosting interpersonal relationships. However, improvements in self-efficacy were limited, suggesting challenges in enhancing this area through short-term interventions, particularly given the cognitive limitations associated with MCI. These findings underline the importance of developing suitable measurement tools that can capture nuanced changes in this population.

Although most outcome measures did not reach statistical significance, the large effect sizes observed in outdoor activity and interpersonal adaptation outcomes suggest promising intervention potential. Given the limited sample size, these findings should be interpreted as preliminary, highlighting the need for further research using larger randomized controlled trials to substantiate and extend these early results. This study underscores the potential for digital interventions to improve quality of life for people with MCI and suggests a need for accessible and user-friendly digital platforms tailored to their specific needs. Future studies should expand on these findings by examining the intervention's effectiveness across a more diverse and larger population to fully leverage digital tools in supporting individuals with MCI.

XSL•FO
RenderX

Acknowledgments

The authors would like to extend their thanks to all the staff, people with mild cognitive impairment, carers, and therapists who contributed to this project. They also sincerely thank student researchers Mingyang Zhang, Lingjie Wu, and Lingjun Liu for their assistance in organizing the activities and collecting the data. In addition, they are grateful to Yuan Fang for helping double-check the data analysis and ensuring its quality. No funding was received for this research.

Data Availability

Due to ethical considerations and data protection policies regarding vulnerable populations, including older adults with mild cognitive impairment (MCI), the deidentified participant data, data dictionary, and statistical code are not available for external sharing. Access is strictly limited to the research team as approved by the institutional ethics committee, and will not be granted to other researchers or institutions.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT checklist.
[DOCX File , 33 KB - humanfactors_v12i1e70177_app1.docx]

Multimedia Appendix 2
Independent-sample t test for change rates in key variables.
[DOCX File , 18 KB - humanfactors_v12i1e70177_app2.docx]

Multimedia Appendix 3
Baseline comparison of social participation.
[DOCX File , 16 KB - humanfactors_v12i1e70177_app3.docx]

References

1. Green AF, Rebok G, Lyketsos CG. Influence of social network characteristics on cognition and functional status with aging. Int J Geriatr Psychiatry 2008 Sep;23(9):972-978 [FREE Full text] [doi: 10.1002/gps.2023] [Medline: 18449952]
2. Dillon C, Serrano CM, Castro D, Leguizamón PP, Heisecke SL, Taragano FE. Behavioral symptoms related to cognitive impairment. Neuropsychiatr Dis Treat 2013;9:1443-1455 [FREE Full text] [doi: 10.2147/NDT.S47133] [Medline: 24092982]
3. Small BJ, Dixon RA, McArdle JJ, Grimm KJ. Do changes in lifestyle engagement moderate cognitive decline in normal aging? Evidence from the Victoria Longitudinal Study. Neuropsychology 2012 Mar;26(2):144-155 [FREE Full text] [doi: 10.1037/a0026579] [Medline: 22149165]
4. Steeman E, de Casterlé BD, Godderis J, Grypdonck M. Living with early-stage dementia: a review of qualitative studies. J Adv Nurs 2006 Jun 30;54(6):722-738. [doi: 10.1111/j.1365-2648.2006.03874.x] [Medline: 16796664]
5. Fallahpour M, Borell L, Luborsky M, Nygård L. Leisure-activity participation to prevent later-life cognitive decline: a systematic review. Scand J Occup Ther 2016 Nov 20;23(3):162-197 [FREE Full text] [doi: 10.3109/11038128.2015.1102320] [Medline: 26586025]
6. Kuiper JS, Zuidersma M, Oude Voshaar RC, Zuidema SU, van den Heuvel ER, Stolk RP, et al. Social relationships and risk of dementia: a systematic review and meta-analysis of longitudinal cohort studies. Ageing Res Rev 2015 Jul;22:39-57. [doi: 10.1016/j.arr.2015.04.006] [Medline: 25956016]
7. Smith L, Shin JI, López Sánchez GF, Oh H, Kostev K, Jacob L, et al. Social participation and mild cognitive impairment in low- and middle-income countries. Prev Med 2022 Nov;164:107230 [FREE Full text] [doi: 10.1016/j.ypmed.2022.107230] [Medline: 36057392]
8. Al Mahmud A, Long KM, Harrington KD, Casey K, Bhar S, Curran S, et al. Developing a digital psychoeducational tool to reduce loneliness in older adults: a design case study. Int J Hum Comput Interact 2021 Aug 04;38(6):499-528. [doi: 10.1080/10447318.2021.1949854]
9. Kelly ME, Duff H, Kelly S, McHugh Power JE, Brennan S, Lawlor BA, et al. The impact of social activities, social networks, social support and social relationships on the cognitive functioning of healthy older adults: a systematic review. Syst Rev 2017 Dec 19;6(1):259 [FREE Full text] [doi: 10.1186/s13643-017-0632-2] [Medline: 29258596]
10. Zaine I, Frohlich DM, Rodrigues KR, Cunha BC, Orlando AF, Scalco LF, et al. Promoting social connection and deepening relations among older adults: design and qualitative evaluation of media parcels. J Med Internet Res 2019 Oct 03;21(10):e14112 [FREE Full text] [doi: 10.2196/14112] [Medline: 31584001]
11. Carmel S, Tur-Sinai A. Cognitive decline among European retirees: impact of early retirement, nation-related and personal characteristics. Ageing Soc 2021 Mar 10;42(10):2343-2369. [doi: 10.1017/S0144686X21000064]

XSL•FO
RenderX

12. Parsons-Suhl K, Johnson ME, McCann JJ, Solberg S. Losing one's memory in early Alzheimer's disease. Qual Health Res 2008 Jan 01;18(1):31-42. [doi: 10.1177/1049732307308987] [Medline: 18174533]
13. Beentjes KM, Neal DP, Kerkhof YJ, Broeder C, Moeridjan ZD, Ettema TP, et al. Impact of the FindMyApps program on people with mild cognitive impairment or dementia and their caregivers; an exploratory pilot randomised controlled trial. Disabil Rehabil Assist Technol 2023 Apr 27;18(3):253-265 [FREE Full text] [doi: 10.1080/17483107.2020.1842918] [Medline: 33245000]
14. Park JH, Park JH. Does cognition-specific computer training have better clinical outcomes than non-specific computer training? A single-blind, randomized controlled trial. Clin Rehabil 2018 Feb 20;32(2):213-222. [doi: 10.1177/0269215517719951] [Medline: 28726492]
15. Long KM, Casey K, Bhar S, Al Mahmud A, Curran S, Hunter K, et al. Understanding perspectives of older adults on the role of technology in the wider context of their social relationships. Ageing Soc 2022 Sep 19;44(6):1453-1476. [doi: 10.1017/S0144686x2200085x]
16. McCarron HR, Zmora R, Gaugler JE. A web-based mobile app with a smartwatch to support social engagement in persons with memory loss: pilot randomized controlled trial. JMIR Aging 2019 Jun 18;2(1):e13378 [FREE Full text] [doi: 10.2196/13378] [Medline: 31518270]
17. Zin NA, Elaklouk AM. Design science paradigm in the development of serious game for cognitive rehabilitation. Int J Adv Sci Eng Inf Technol 2017;7:118-124. [doi: 10.18517/IJASEIT.7.1.1171]
18. Gao Y, Liu N. Effects of digital technology-based serious games interventions for older adults with mild cognitive impairment: a meta-analysis of randomised controlled trials. Age Ageing 2025 Mar 28;54(4):afaf080. [doi: 10.1093/ageing/afaf080] [Medline: 40192627]
19. Rios Rincon AM, Miguel Cruz A, Daum C, Neubauer N, Comeau A, Liu L. Digital storytelling in older adults with typical aging, and with mild cognitive impairment or dementia: a systematic literature review. J Appl Gerontol 2022 Mar 19;41(3):867-880 [FREE Full text] [doi: 10.1177/07334648211015456] [Medline: 34009053]
20. Hashim A, Rias RM, Kamaruzaman MF. The use of personalized digital memory book as a reminiscence therapy for Alzheimer's disease (AD) patients. In: Proceedings of the 3rd International Visual Informatics Conference on Advances in Visual Informatics. 2013 Presented at: IVIC '13; November 13-15, 2013; Selangor, Malaysia p. 508-515 URL: https://link.springer.com/chapter/10.1007/978-3-319-02958-0_46 [doi: 10.1007/978-3-319-02958-0_46]
21. Zheng M, Zhang X, Xiao H. Effects of a WeChat-based life review program for patients with digestive system cancer: 3-arm parallel randomized controlled trial. J Med Internet Res 2022 Aug 25;24(8):e36000 [FREE Full text] [doi: 10.2196/36000] [Medline: 36006665]
22. Sapkota A, Conte S, Munteanu C. Time travel tours: an accessible social VR storytelling tool for older adults. In: Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems. 2025 Presented at: CHI EA '25; April 26-May 1, 2025; Yokohama, Japan p. 1-5 URL: https://dl.acm.org/doi/10.1145/3706599.3721196 [doi: 10.1145/3706599.3721196]
23. Xu L, Fields NL, Highfill MC, Troutman BA. Remembering the past with today's technology: a scoping review of reminiscence-based digital storytelling with older adults. Behav Sci (Basel) 2023 Dec 04;13(12):998 [FREE Full text] [doi: 10.3390/bs13120998] [Medline: 38131854]
24. Kim S, Chee KH, Gerhart O. Generativity in creative storytelling: evidence from a dementia care community. Innov Aging 2020;4(2):igaa002 [FREE Full text] [doi: 10.1093/geroni/igaa002] [Medline: 32258435]
25. Woods B, Subramaniam P. Digital life storybooks for people with dementia living in care homes: an evaluation. Clin Interv Aging 2016;11:1263-1276 [FREE Full text] [doi: 10.2147/CIA.S111097] [Medline: 27698556]
26. Wood D. Commentary: contribution of scaffolding to learning and teaching: interdisciplinary perspectives. Int J Educ Res 2018;90:248-251. [doi: 10.1016/j.ijer.2018.03.005]
27. Zhu D, Al Mahmud A, Liu W. Design requirements for a digital storytelling application for people with mild cognitive impairment (MCI). Digit Health 2024 Sep 19;10:20552076241282237 [FREE Full text] [doi: 10.1177/20552076241282237] [Medline: 39314819]
28. Li C, Lin X, Kang K, Hu J, Hengeveld B, Hummels C, et al. Interactive gallery: enhance social interaction for elders by story sharing. In: Proceedings of the 2017 International Workshop on Advances in Digital Cultural Heritage. 2017 Presented at: ADCH '17; June 28, 2017; Madeira, Portugal p. 116 URL: https://link.springer.com/chapter/10.1007/978-3-319-75789-6_8
29. Elfrink TR, Ullrich C, Kunz M, Zuidema SU, Westerhof GJ. The online life story book: a randomized controlled trial on the effects of a digital reminiscence intervention for people with (very) mild dementia and their informal caregivers. PLoS One 2021 Sep 15;16(9):e0256251 [FREE Full text] [doi: 10.1371/journal.pone.0256251] [Medline: 34525105]
30. Moon S, Park K. The effect of digital reminiscence therapy on people with dementia: a pilot randomized controlled trial. BMC Geriatr 2020 May 06;20(1):166 [FREE Full text] [doi: 10.1186/s12877-020-01563-2] [Medline: 32375661]
31. Potts C, Bond R, Ryan A, Mulvenna M, McCauley C, Laird E, et al. Ecological momentary assessment within a digital health intervention for reminiscence in persons with dementia and caregivers: user engagement study. JMIR Mhealth Uhealth 2020 Jul 06;8(7):e17120 [FREE Full text] [doi: 10.2196/17120] [Medline: 32420890]

XSL•FO
RenderX

32. Tominari M, Uozumi R, Becker C, Kinoshita A. Reminiscence therapy using virtual reality technology affects cognitive function and subjective well-being in older adults with dementia. Cogent Psychol 2021 Sep 01;8(1):1968991. [doi: 10.1080/23311908.2021.1968991]
33. Ngiam NH, Yee WQ, Teo N, Yow KS, Soundararajan A, Lim JX, et al. Building digital literacy in older adults of low socioeconomic status in Singapore (project wire up): nonrandomized controlled trial. J Med Internet Res 2022 Dec 02;24(12):e40341 [FREE Full text] [doi: 10.2196/40341] [Medline: 36459398]
34. Sarne-Fleischmann V, Tractinsky N, Dwolatzky T, Rief I. Personalized reminiscence therapy for patients with Alzheimer's disease using a computerized system. In: Proceedings of the 4th International Conference on PErvasive Technologies Related to Assistive Environments. 2011 Presented at: PETRA '11; May 25-27, 2011; Heraklion Crete, Greece p. 1-4 URL: https://dl.acm.org/doi/10.1145/2141622.2141679 [doi: 10.1145/2141622.2141679]
35. Derbring S, Barbos Nordström M, Svenningsson JA, Ekström A, Ingebrand E, Samuelsson C, et al. Effects of a digital reminiscing intervention on people with dementia and their care-givers and relatives. Ageing Soc 2021 Oct 27;43(9):1983-2000. [doi: 10.1017/s0144686x21001446]
36. Filoteo JV, Cox EM, Split M, Gross M, Culjat M, Keene D. Evaluation of ReminX as a behavioral intervention for mild to moderate dementia. Annu Int Conf IEEE Eng Med Biol Soc 2018 Jul;2018:3314-3317. [doi: 10.1109/EMBC.2018.8513049] [Medline: 30441098]
37. Bentley FR, Basapur S, Chowdhury SK. Promoting intergenerational communication through location-based asynchronous video communication. In: Proceedings of the 13th international conference on Ubiquitous computing. 2011 Presented at: UbiComp '11; September 17-21, 2011; Beijing, China p. 31-40 URL: https://dl.acm.org/doi/10.1145/2030112.2030117 [doi: 10.1145/2030112.2030117]
38. Zhu D, Al Mahmud A, Liu W. Social connections and participation among people with mild cognitive impairment: barriers and recommendations. Front Psychiatry 2023 Jul 5;14:1188887 [FREE Full text] [doi: 10.3389/fpsyt.2023.1188887] [Medline: 37476544]
39. Qi S, Sun Y, Yin P, Zhang H, Wang Z. Mobile phone use and cognitive impairment among elderly Chinese: a national cross-sectional survey study. Int J Environ Res Public Health 2021 May 26;18(11):5695 [FREE Full text] [doi: 10.3390/ijerph18115695] [Medline: 34073323]
40. Chonody J, Wang D. Connecting older adults to the community through multimedia: an intergenerational reminiscence program. Activ Adapt Aging 2013 Jan;37(1):79-93. [doi: 10.1080/01924788.2012.760140]
41. McCauley CO, Bond RB, Ryan A, Mulvenna MD, Laird L, Gibson A, et al. Evaluating user engagement with a reminiscence app using cross-comparative analysis of user event logs and qualitative data. Cyberpsychol Behav Soc Netw 2019 Aug;22(8):543-551. [doi: 10.1089/cyber.2019.0076] [Medline: 31403855]
42. Zhu D, Al Mahmud A, Liu W, Wang D. Digital storytelling for people with cognitive impairment using available mobile apps: systematic search in app stores and content analysis. JMIR Aging 2024 Oct 24;7:e64525 [FREE Full text] [doi: 10.2196/64525] [Medline: 39446478]
43. Njeru JW, Patten CA, Hanza MM, Brockman TA, Ridgeway JL, Weis JA, et al. Stories for change: development of a diabetes digital storytelling intervention for refugees and immigrants to minnesota using qualitative methods. BMC Public Health 2015 Dec 29;15(1):1311 [FREE Full text] [doi: 10.1186/s12889-015-2628-y] [Medline: 26715465]
44. Mulvenna M, Gibson A, McCauley C, Ryan A, Bond RR, Laird L, et al. Behavioural usage analysis of a reminiscing app for people living with dementia and their carers. In: Proceedings of the 2017 European Conference on Cognitive Ergonomics. 2017 Presented at: ECCE '17; September 19-22, 2017; Umeå, Sweden p. 35-38 URL: https://dl.acm.org/doi/10.1145/3121283.3121289 [doi: 10.1145/3121283.3121289]
45. Sehrawat S, Jones C, Orlando J, Bowers T, Rubins A. Digital storytelling: a tool for social connectedness. Gerontechnology 2017 Apr 28;16(1):56-61. [doi: 10.4017/GT.2017.16.1.006.00]
46. Di Lorito C, Bosco A, Rai H, Craven M, McNally D, Todd C, et al. A systematic literature review and meta-analysis on digital health interventions for people living with dementia and Mild Cognitive Impairment. Int J Geriatr Psychiatry 2022 Jun 19;37(6):5730 [FREE Full text] [doi: 10.1002/gps.5730] [Medline: 35588315]
47. Karakose T, Tülübaş T, Papadakis S. Revealing the intellectual structure and evolution of digital addiction research: an integrated bibliometric and science mapping approach. Int J Environ Res Public Health 2022 Nov 12;19(22):14883 [FREE Full text] [doi: 10.3390/ijerph192214883] [Medline: 36429603]
48. Zhu D, Al Mahmud A, Liu W. Digital storytelling intervention for enhancing the social participation of people with mild cognitive impairment: co-design and usability study. JMIR Aging 2024 Jan 17;7:e54138 [FREE Full text] [doi: 10.2196/54138] [Medline: 38231541]
49. Zhu D, Al Mahmud A, Liu W. Digital storytelling intervention to enhance social connections and participation for people with mild cognitive impairment: a research protocol. Front Psychiatry 2023 Oct 11;14:1217323 [FREE Full text] [doi: 10.3389/fpsyt.2023.1217323] [Medline: 37886113]
50. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Apr 30;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]

XSL•FO
RenderX

51. Noreau L, Fougeyrollas P, Vincent C. The LIFE-H: assessment of the quality of social participation. Technol Disabil 2002 Sep 29;14(3):113-118. [doi: 10.3233/tad-2002-14306]
52. Clare L, Kudlicka A, Oyebode JR, Jones RW, Bayer A, Leroi I, et al. Individual goal-oriented cognitive rehabilitation to improve everyday functioning for people with early-stage dementia: a multicentre randomised controlled trial (the GREAT trial). Int J Geriatr Psychiatry 2019 May;34(5):709-721 [FREE Full text] [doi: 10.1002/gps.5076] [Medline: 30724405]
53. Lee RM, Draper M, Lee S. Social connectedness, dysfunctional interpersonal behaviors, and psychological distress: testing a mediator model. J Couns Psychol 2001;48(3):310-318. [doi: 10.1037//0022-0167.48.3.310]
54. Schwarzer R, Bäßler J, Kwiatek P, Schröder K, Zhang JX. The assessment of optimistic self‐beliefs: comparison of the German, Spanish, and Chinese versions of the general self‐efficacy scale. Applied Psychology 2008 Jan 22;46(1):69-88. [doi: 10.1111/j.1464-0597.1997.tb01096.x]
55. Zeng G, Fung S, Li J, Hussain N, Yu P. Evaluating the psychometric properties and factor structure of the general self-efficacy scale in China. Curr Psychol 2020 Jul 11;41(6):3970-3980. [doi: 10.1007/S12144-020-00924-9]
56. Fang C, Ma H, Bao C, Wang Z, Li G, Sun S, et al. Urban–rural human settlements in China: objective evaluation and subjective well-being. Humanit Soc Sci Commun 2022 Oct 28;9(1):3. [doi: 10.1057/S41599-022-01417-9]
57. Lewis S. Qualitative inquiry and research design: choosing among five approaches. Health Promot Pract 2015 Apr 02;16(4):473-475. [doi: 10.1177/1524839915580941]
58. A fresh way to do statistics. JASP. URL: https://jasp-stats.org/ [accessed 2025-05-29]
59. Ryu J, Heo J. Relationships between leisure activity types and well-being in older adults. Leis Stud 2017 Sep 22;37(3):331-342. [doi: 10.1080/02614367.2017.1370007]
60. Chang PJ. Effects of the built and social features of urban greenways on the outdoor activity of older adults. Landsc Urban Plann 2020 Dec;204:103929. [doi: 10.1016/j.landurbplan.2020.103929]
61. Geohagen O, Hamer L, Lowton A, Guerra S, Milton-Cole R, Ellery P, et al. The effectiveness of rehabilitation interventions including outdoor mobility on older adults' physical activity, endurance, outdoor mobility and falls-related self-efficacy: systematic review and meta-analysis. Age Ageing 2022 Jun 01;51(6):afac120 [FREE Full text] [doi: 10.1093/ageing/afac120] [Medline: 35737601]
62. Singh B, Olds T, Curtis R, Dumuid D, Virgara R, Watson A, et al. Effectiveness of physical activity interventions for improving depression, anxiety and distress: an overview of systematic reviews. Br J Sports Med 2023 Sep 16;57(18):1203-1209 [FREE Full text] [doi: 10.1136/bjsports-2022-106195] [Medline: 36796860]
63. Yu K, Wild K, Dowling NM, Kaye JA, Silbert LC, Dodge HH. Emotional characteristics of socially isolated older adults with MCI using tablet administered NIH toolbox: I-CONECT study. Alzheimers Dement (Amst) 2022 Nov 11;14(1):e12372 [FREE Full text] [doi: 10.1002/dad2.12372] [Medline: 36381558]
64. Dickens AP, Richards SH, Greaves CJ, Campbell JL. Interventions targeting social isolation in older people: a systematic review. BMC Public Health 2011 Aug 15;11(1):647 [FREE Full text] [doi: 10.1186/1471-2458-11-647] [Medline: 21843337]
65. Chang H, Do Y, Ahn J. Digital storytelling as an intervention for older adults: a scoping review. Int J Environ Res Public Health 2023 Jan 11;20(2):1344 [FREE Full text] [doi: 10.3390/ijerph20021344] [Medline: 36674100]
66. McAuley E, Elavsky S, Motl RW, Konopack JF, Hu L, Marquez DX. Physical activity, self-efficacy, and self-esteem: longitudinal relationships in older adults. J Gerontol B Psychol Sci Soc Sci 2005 Sep 01;60(5):P268-P275. [doi: 10.1093/geronb/60.5.p268] [Medline: 16131621]
67. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977 Mar;84(2):191-215. [doi: 10.1037//0033-295x.84.2.191] [Medline: 847061]
68. Resnick B, Palmer MH, Jenkins LS, Spellbring AM. Path analysis of efficacy expectations and exercise behaviour in older adults. J Adv Nurs 2000 Jun 25;31(6):1309-1315. [doi: 10.1046/j.1365-2648.2000.01463.x] [Medline: 10849141]
69. Luszczynska A, Gutiérrez‐Doña B, Schwarzer R. General self‐efficacy in various domains of human functioning: evidence from five countries. Int J Psychol 2005 Apr;40(2):80-89. [doi: 10.1080/00207590444000041]
70. Schwarzer R. Self-Efficacy: Thought Control Of Action. New York, NY: Taylor & Francis; 2014.
71. Petersen RC. Mild cognitive impairment. N Engl J Med 2011 Jun 09;364(23):2227-2234. [doi: 10.1056/nejmcp0910237]
72. Taler V, Phillips NA. Language performance in Alzheimer's disease and mild cognitive impairment: a comparative review. J Clin Exp Neuropsychol 2008 Jul 05;30(5):501-556. [doi: 10.1080/13803390701550128] [Medline: 18569251]
73. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011 May;7(3):270-279 [FREE Full text] [doi: 10.1016/j.jalz.2011.03.008] [Medline: 21514249]
74. Galle SA, Deijen JB, Milders MV, De Greef MH, Scherder EJ, van Duijn CM, et al. The effects of a moderate physical activity intervention on physical fitness and cognition in healthy elderly with low levels of physical activity: a randomized controlled trial. Alzheimers Res Ther 2023 Jan 11;15(1):12 [FREE Full text] [doi: 10.1186/s13195-022-01123-3] [Medline: 36631905]
75. Bhatia D, Salbach NM, Akinrolie O, Alsbury-Nealy K, Dos Santos RB, Eftekhar P, et al. Outdoor community ambulation interventions to improve physical and mental health in older adults: a systematic review and meta-analysis. J Aging Phys Act 2022 Dec 01;30(6):1061-1074 [FREE Full text] [doi: 10.1123/japa.2021-0151] [Medline: 35418512]

XSL•FO
RenderX

Abbreviations

GSES: General Self-Efficacy Scale
LIFE-H: Assessment of Life Habits
MCI: mild cognitive impairment
MoCA-C: Montreal Cognitive Assessment–Chinese version
SCS-R: Social Connectedness Scale–Revised
SWBS: Subjective Well-Being Scale

Edited by A Kushniruk; submitted 17.12.24; peer-reviewed by E Opoku-Agyemang, S Papadakis; comments to author 20.04.25; revised version received 02.05.25; accepted 12.05.25; published 18.06.25.

Please cite as:
Zhu D, Al Mahmud A, Liu W
Effects of a Mobile Storytelling App (Huiyou) on Social Participation Among People With Mild Cognitive Impairment: Pilot Randomized Controlled Trial
JMIR Hum Factors 2025;12:e70177
URL: https://humanfactors.jmir.org/2025/1/e70177
doi:10.2196/70177
PMID:

©Di Zhu, Abdullah Al Mahmud, Wei Liu. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 18.06.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Understanding the Behavioral Determinants of First Responder App Adoption by Integrating Perspectives From the Unified Theory of Acceptance and Use of Technology and Health Belief Model: Cross-Sectional Survey

Cas von Winckelmann[1], MSc; Robyn Vanherle[1,2], PhD; Lara Schreurs[1], PhD; Olivier Hoogmartens[3], MSc; Heidi Salaets[4], PhD; Jan De Spiegeleer[5], PhD; Marc Sabbe[6], PhD; Kathleen Beullens[1], PhD

[1]Media Psychology Lab, Department of Communication Science, KU Leuven, Leuven, Belgium

[2]Research Foundation - Flanders, Brussels, Belgium

[3]Leuven Institute for Healthcare Policy, Department of Public Health and Primary Care, KU Leuven, Leuven, Belgium

[4]Interpreting Studies, Faculty of Arts Antwerp Campuses, KU Leuven, Antwerp, Belgium

[5]Statistics and Risks, Department of Mathematics, KU Leuven, Leuven, Belgium

[6]Emergency Medicine, Department of Public Health and Primary Care, KU Leuven, Leuven, Belgium

Corresponding Author:
Cas von Winckelmann, MSc
Media Psychology Lab
Department of Communication Science
KU Leuven
Parkstraat 45 (Box 3603)
Leuven, 3000
Belgium
Phone: 32 16 37 70 92
Email: cas.vonwinckelmann@kuleuven.be

Abstract

Background: Out-of-hospital cardiac arrests (OHCAs) are a leading cause of death worldwide, yet first responder apps can significantly improve outcomes by mobilizing citizens to perform cardiopulmonary resuscitation before professional help arrives. Despite their importance, limited research has examined the psychological and behavioral factors that influence individuals' willingness to adopt these apps.

Objective: Given that first responder app use involves elements of both technology adoption and preventive health behavior, it is essential to examine this behavior from multiple theoretical perspectives. Building on the unified theory of acceptance and use of technology (UTAUT) and health belief model (HBM), this study therefore developed an integrative framework to explain which behavioral determinants and demographic and health-related factors drive an individual's willingness to install a first responder app for OHCA.

Methods: We conducted a web-based cross-sectional survey (N=3660; mean age 49.95, SD 16.75 years; n=1909, 52.2% women) in June 2024 among Belgian adults. Behavioral determinants (UTAUT and HBM constructs), demographic (eg, age), and health-related (eg, cardiopulmonary resuscitation training experience) variables were measured using (multi-item) scales. Willingness to install the app served as the outcome variable. We developed a structural equation model using the *Lavaan* package in R and specified regression paths, on the one hand, between the behavioral determinants and willingness to install the app, and on the other hand, between the demographic and health-related factors and the behavioral determinants. Additionally, we conducted multiple group analyses to examine the moderating role of demographic and health-related factors on the relationships between the behavioral determinants and the willingness to install the app.

Results: Our results revealed that 2 UTAUT variables (ie, facilitating conditions: β=.07; *P*=.003 and social influence: β=.16; *P*<.001) and 3 HBM variables (ie, perceived susceptibility: β=.06; *P*=.003, perceived barriers: β=–.29; *P*<.001, and perceived benefits: β=.38; *P*<.001) were associated with willingness to install a first responder app for OHCA. Additionally, most demographic and health-related factors were indirectly related to willingness via behavioral determinants, with age being the sole moderator.

XSL•FO
RenderX

Specifically, a negative association between perceived severity and willingness was only observed among older adults. In addition, the positive relationship between perceived benefits and willingness was stronger for older adults compared to younger ones.

Conclusions: Overall, the results of this study have both theoretical and practical implications. Theoretically, this study finds its relevance in extending the UTAUT and HBM to altruistic mobile health apps and advancing our understanding of technology adoption in health contexts. Practically, the study's findings could inform real-life health campaigns aimed at enhancing citizen participation in first responder systems.

(JMIR Hum Factors 2025;12:e69934) doi:10.2196/69934

KEYWORDS

mHealth; mobile health; out-of-hospital cardiac arrest; first responder app; UTAUT; unified theory of acceptance and use of technology; health belief model; health technology adoption; cardiopulmonary resuscitation

Introduction

Background

With the widespread use of mobile phones, the potential of technology to improve health outcomes has significantly increased [1]. Many people, for example, use their mobile devices to install apps for medical and public health practices, better known as mobile health (mHealth) [2]. Most of these apps enable users to monitor their health by, for instance, tracking their steps or creating personalized diet plans [3,4]. Additionally, besides promoting self-tracking, some mHealth apps can also be used to enhance the health of others and benefit society overall. Examples of such altruistic mHealth apps include contact tracing apps (CTAs), used during the recent COVID-19 pandemic [5], and first responder apps.

First responder apps alert and mobilize trained citizens to provide first aid in emergency situations, such as experiencing an out-of-hospital cardiac arrest (OHCA) [6]. Experiencing an OHCA is a leading cause of death worldwide [7], and in order to tackle this public health problem, early recognition of an OHCA and responding quickly to the situation are crucial. The use of a first responder app could thus be critical in this case, as citizens can be alerted to provide cardiopulmonary resuscitation (CPR) before the arrival of professional medical help, thereby ensuring a higher chance of survival [6,8]. However, despite the significant value of these apps, little is still known about their integration within societies for 2 reasons. First, these apps offer health benefits for others but not directly for users themselves, thereby leaving a gap in our understanding of users' motivations behind adopting these altruistic apps. Second, first responder apps for OHCA differ substantially from other mHealth apps, as they are not designed for frequent use. Once installed, users are only required to engage with the app when they receive a notification requesting assistance for someone experiencing a cardiac arrest. As such, given this infrequent use, it is crucial to understand individuals' willingness to install first responder apps for OHCA in order to promote the adoption of such apps in society.

Therefore, this study aims to contribute to the literature by defining which behavioral and individual factors play a role in citizens' willingness to install a first responder app for OHCA. Given that first responder app use involves elements of both technology adoption and preventive health behavior, it is essential to examine this behavior from multiple theoretical perspectives, including the unified theory of acceptance and use of technology (UTAUT) [9] and the health belief model (HBM) [10]. Based on these theories, we will develop an integrated framework to explain the behavioral determinants that drive individuals' willingness to install altruistic mHealth apps, that is, first responder apps for OHCA. Additionally, demographic and health variables will be examined to provide a comprehensive understanding of factors driving adoption willingness.

Technology Perspective: UTAUT

To address the technology-related aspect of first responder app adoption, the study builds on the UTAUT model [9], a commonly used theoretical model for predicting the adoption of new technologies. The model consists of 4 key behavioral determinants of technology adoption: performance expectancy (ie, perceived effectiveness of the technology), effort expectancy (ie, the degree of ease associated with using the technology), social influence (ie, descriptive and injunctive norms), and facilitating conditions (ie, resources available that make it easier to adopt the new technology, such as internet access and owning a smartphone).

Given that this model integrates elements from various models and theories related to technology acceptance, it has been used frequently in the past to explain the adoption of apps that require frequent and consistent interaction, such as mobile banking apps [11], health monitoring apps [12,13], or travel apps [14]. Moreover, besides these frequently used apps, UTAUT has also been successful in explaining the adoption of CTAs [5,15-18]. Studies, for example, found that users were more inclined to install CTAs when they perceived the app as effective [5,17,18] and easy to use [18]. Additionally, social influence [16] and facilitating conditions, like smartphone ownership [19], were crucial predictors of CTA adoption. The study by Walrave et al [19], for example, showed that not owning a smartphone was one of the main reasons for not using a CTA.

These studies on CTAs, in turn, offer a relevant comparison for our study, as these apps, similar to first responder apps for OHCA, benefit public health and involve passive use (ie, only interacting when receiving a notification). Based on their insights, it is thus reasonable to assume that the factors outlined in the UTAUT model could also serve as effective predictors for the willingness to install a first responder app for OHCA—an app that is characterized by its altruistic nature, focus on immediate action, and infrequent use. In line with the prototype

XSL•FO
RenderX

willingness model [20], we focus on willingness to install a first responder app rather than behavioral intention or actual use, since no such app currently exists in Belgium, thus making those measures less feasible at this stage. Building on the UTAUT model, this study therefore extends the literature by forming the following hypotheses:

- Hypothesis 1—Performance expectancy is positively associated with the willingness to install a first responder app for OHCA.
- Hypothesis 2—Effort expectancy is negatively associated with the willingness to install a first responder app for OHCA.
- Hypothesis 3—Facilitating conditions are positively associated with the willingness to install a first responder app for OHCA.
- Hypothesis 4—Social influence is positively associated with the willingness to install a first responder app for OHCA.

Health Perspective: HBM

When installing a first responder app, individuals are not merely installing new software but are committing to potentially performing life-saving actions, that is, providing CPR to someone experiencing a cardiac arrest. Relying solely on a technological perspective, such as the UTAUT model, would therefore be insufficient to fully understand an individual's willingness to install a first responder app for OHCA. Logically, this decision requires consideration of the health-related implications and responsibilities associated with the app's use. Therefore, in addition to examining first responder app adoption through the lens of technology acceptance, it is essential to apply a health behavior change model to better understand the factors influencing the adoption of a first responder app for OHCA.

To do so, we will integrate the HBM [21], which is a widely recognized framework for understanding health-related behavior. The HBM suggests that individuals' decisions to take health actions are influenced by 5 different factors: perceived susceptibility (ie, belief in the risk of experiencing a health problem), perceived severity (ie, belief in the seriousness of the health problem), self-efficacy (ie, confidence in performing the health behavior), perceived benefits (ie, factors that facilitate the health behavior), and perceived barriers (ie, factors that hinder the health behavior). In this context, a perceived benefit might be personal relevance [22], while a potential barrier could include privacy concerns [23]. Self-efficacy was later added as the fifth variable [21].

The model is often used to explain health behaviors aimed at protecting or improving one's own health, such as condom use [24], or using mobile apps for managing chronic diseases [25]. However, limited research has explored the effectiveness of the HBM in predicting altruistic health behavior like providing CPR. Unlike self-beneficial behaviors, installing a first responder app, and thus committing oneself to potentially performing CPR, is a health behavior that only benefits another individual (ie, person experiencing a cardiac arrest). Several studies have examined the application of the HBM in the context of CTA adoption, which can also be considered an altruistic health behavior, as its primary goal is to minimize the spread of COVID-19. These studies indicate that self-efficacy, perceived barriers, and perceived benefits are the most significant predictors of adoption [18,26]. For instance, participants were more inclined to install the app when they felt confident in using it, recognized its benefits, and perceived fewer barriers to its adoption. While perceived severity and susceptibility have not consistently predicted adoption in studies on CTAs, they have shown positive associations with preventive health behaviors, such as mask-wearing [27].

Despite these insights related to CTAs, research on first responder apps and the determinants outlined in the HBM is limited. Still, building on findings from previous studies, the HBM likely also applies to the adoption of first responder apps for OHCA. For instance, as these apps can prevent severe outcomes, such as death or brain injury [28], individuals who perceive OHCA as a severe health risk may be more willing to install the app to be able to prevent such serious consequences. In addition, perceived susceptibility may influence willingness by heightening awareness of the app's necessity in addressing a widespread health risk in society. Furthermore, because of the similar altruistic and preventive nature of CTAs, we can expect that self-efficacy and perceived benefits may positively influence the willingness to install a first responder app, while perceived barriers may negatively influence it.

Building on these insights, this study therefore aims to provide an extensive overview of how the variables of the HBM can explain the willingness to install a first responder app for OHCA. Based on current literature on mHealth apps and preventive health behavior, we propose the following hypotheses:

- Hypothesis 5—Self-efficacy is positively associated with the willingness to install a first responder app for OHCA.
- Hypothesis 6—Perceived susceptibility is positively associated with the willingness to install a first responder app for OHCA.
- Hypothesis 7—Perceived severity is positively associated with the willingness to install a first responder app for OHCA.
- Hypothesis 8—Perceived barriers are negatively associated with the willingness to install a first responder app for OHCA.
- Hypothesis 9—Perceived benefits are positively associated with the willingness to install a first responder app for OHCA.

Influential Factors

Apart from the previously mentioned direct determinants of first responder app adoption, the HBM also identified a number of indirect factors affecting health behavior, including demographic factors and prior contact with the health problem [10]. Other health behavior change models such as the theory of planned behavior [29] and social cognitive theory [30] also emphasize the importance of demographic (eg, age, gender, and socioeconomic status [SES]) and social context variables in shaping individuals' beliefs, attitudes, and social norms. Moreover, these theoretical propositions have also been confirmed in empirical research as, for example, older people

XSL•FO
RenderX

tend to perceive health issues, such as COVID-19, as more severe [31].

Along with demographic variables, practicing a medical profession may also play a significant role in installing a first responder app for OHCA. Individuals working in the medical sector could, for example, perceive greater benefits from using such an app, as it can improve patient outcomes through bystander CPR [6,8]. Consequently, their perceived barriers to installing the app and expectations regarding its effectiveness may also be shaped by their professional experience and familiarity with emergency medical care.

In addition, as installing a first responder app for OHCA is a health-related behavior, the general health status of an individual and other relevant health variables should also be taken into account. First, given that CPR training significantly enhances the willingness to provide CPR [32,33], we could assume that trained individuals will be more likely to install a first responder app for OHCA, potentially via various behavioral determinants such as increased self-efficacy and perceived benefits. Second, the HBM highlights the relevance of prior contact with a health problem as a factor that can increase perceived susceptibility and severity, thereby influencing behavior [10]. In this study, having a cardiovascular disease and, therefore, an increased change of experiencing a cardiac arrest can be a relevant factor. Finally, indirect contact with health issues, such as having a family member or close acquaintance with a cardiovascular disease, could also affect individuals' perceptions and decisions to adopt health-related behaviors like installing a first responder app for OHCA.

Since relationships between these influential factors (ie, demographic and health variables) and behavioral determinants of first responder app adoption have not been tested before, we pose the following research question (RQ):

- RQ1—Do demographic and health variables influence the behavioral determinants of willingness to install a first responder app for OHCA (performance expectancy, effort expectancy, social influence, facilitating conditions, self-efficacy, perceived severity, perceived susceptibility, perceived barriers, and perceived benefits)?

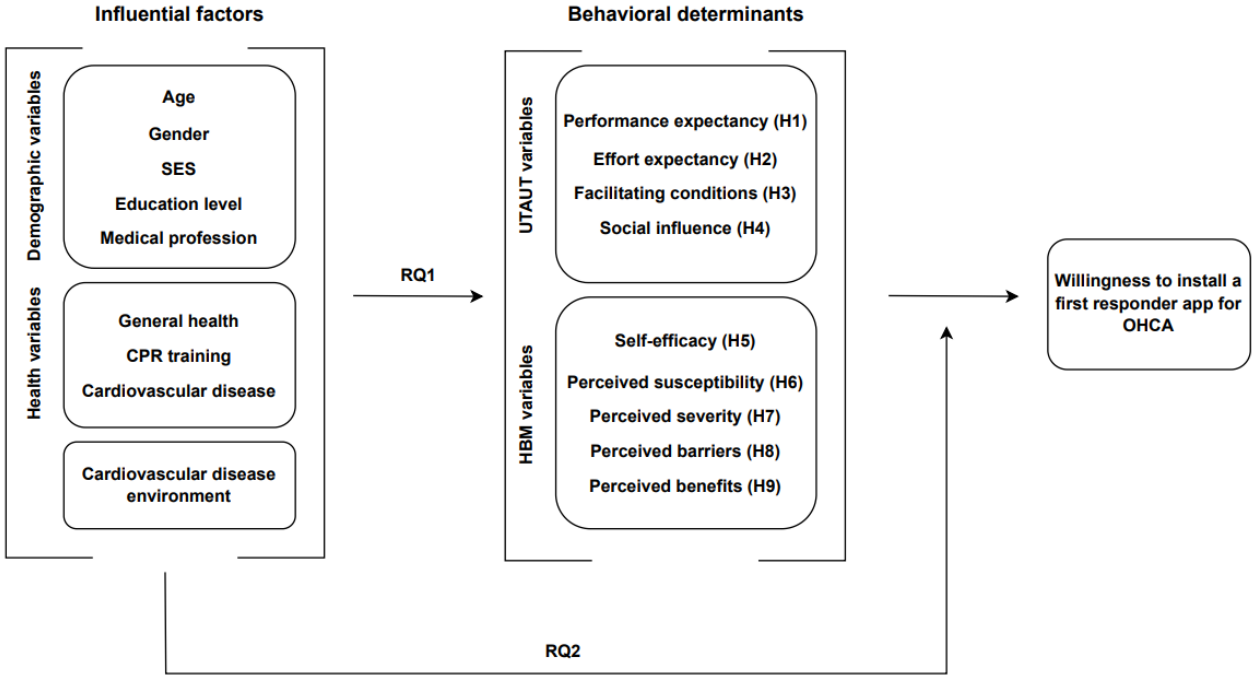
In addition to examining the variables mentioned earlier as predictors, it may be valuable to explore their potential moderating role. In the UTAUT model, gender and age already moderate the relationships between the key determinants and technology adoption [9,34]. For example, the relationship between performance expectancy and the intention to adopt a technology was stronger for men and young people, but the opposite was found for the other determinants [9,35]. Additionally, empirical research indicates that an individual's education level moderates relationships between certain UTAUT variables (ie, performance expectancy and social influence) and individuals' willingness to adopt new software [36]. This suggests that similar effects could be found for other demographic and health-related variables in the context of first responder OHCA apps. As there are again a limited amount of studies that investigated these moderations, we will examine the following RQ:

- RQ2—Do demographic and health variables moderate the relationships between the behavioral determinants (performance expectancy, effort expectancy, social influence, facilitating conditions, self-efficacy, perceived severity, perceived susceptibility, perceived barriers, and perceived benefits) and individuals' willingness to install a first responder app for OHCA?

Overall, to fully understand the willingness to install a first responder app for OHCA, this study will evaluate a self-developed integrative framework that incorporates determinants from the UTAUT and HBM, along with additional influential factors that are particularly relevant to this context (Figure 1).

XSL•FO
RenderX

Figure 1. Conceptual framework of understanding and predicting willingness to install a first responder app for OHCA. Arrows indicate hypothesized associations based on theory and do not imply a causal relationship due to the cross-sectional design. CPR: cardiopulmonary resuscitation; H: hypothesis; HBM: health belief model; OHCA: out-of-hospital cardiac arrest; RQ: research question; SES: socioeconomic status; UTAUT: unified theory of acceptance and use of technology.



Methods

Study Design, Setting, and Participants

To test our hypotheses and address our RQs, we conducted a cross-sectional survey in June 2024 among a representative sample of Belgian adults. Participants were recruited out of a large national panel managed by a research agency, which ensured representativeness based on key sociodemographic characteristics, including gender, age, education level, and region of residence. Using quota sampling, the agency, with help from the researchers, actively monitored sample composition and adjusted invitations during data collection (eg, targeting underrepresented age groups). Eligible participants were individuals aged 18 years or older residing in Belgium. Respondents were asked to fill in a web-based Qualtrics questionnaire in Dutch or French, given the bilingual nature of the country.

Ethical Considerations

The study received ethics approval from the Social and Societal Ethics Committee of the KU Leuven (G-2024-7825). At the beginning of the survey, participants were informed about the study's objectives, procedures, their right to opt out at any time, and the confidentiality of their data, and were required to provide informed consent before continuing. Given the web-based procedure, the involved researchers were available via email to answer questions and solve practical and logistic issues. As compensation, participants received digital points from the research agency that recruited them, which could be redeemed for cash or donated to charity. To ensure privacy and confidentiality, all data were collected anonymously, with no personal identifiable information recorded, and were used solely for research purposes. In addition, the data were kept on a secure server.

Study Size

Of the initial 3908 respondents, only those who completed more than the sociodemographic section of the questionnaire were included (n=3752). After removing 92 ineligible cases (underage, duplicate responses, or lack of consent), the final sample was 3660 participants, with a mean age of 49.95 (SD 16.75) years; 52.2% (n=1909) were women. The sample included 2205 Dutch-speaking and 1455 French-speaking respondents. Education levels varied, with 1.6% (n=58) having no formal education, 3.9% (n=144) primary, 36% (n=1319) secondary, 28.9% (n=1058) bachelor, 26.5% (n=971) master, and 3% (n=110) PhD. SES was assessed using the MacArthur Scale of Subjective Social Status [37]: 4% (n=149) low SES (scores 1-3), 70.3% (n=2576) medium SES (scores 4-7), and 25.5% (n=935) high SES (scores 8-10). Health-related data showed that 10.1% (n=369) worked as medical professionals, 58.8% (n=2153) had CPR training, 15.4% (n=564) had cardiovascular disease, and 30.4% (n=1113) knew someone with cardiovascular disease.

Assessments and Data Sources

Demographic Variables

Respondents' gender (1=man, 2=woman, and 3=other), year of birth, education level (1=no degree, 2=primary school degree, 3=secondary school degree, 4=bachelor degree, 5=master degree, and 6=PhD), and subjective SES were questioned [37]. Respondents were also asked if they practice a medical profession (1=yes and 0=no).

XSL•FO
RenderX

Health Variables

We asked respondents whether they had ever followed a CPR training (1=yes and 0=no) and whether they (1=yes and 0=no) or people in their surroundings (1=yes and 0=no) experienced a cardiovascular disease. Additionally, general health was measured with an adapted version of the Linear Analogue Scale [38]. Respondents had to indicate their general health condition on a scale from 1=worst to 100=best. On average, participants scored 71.11 (SD 18.05), suggesting that the overall health of our sample was good.

Willingness to Install a First Responder App for OHCA

To measure participants' willingness to install a first responder app for OHCA, participants first received an explanation of what is meant with a first responder app. They were told that a first responder app for OHCA is a mobile app that alerts and mobilizes volunteer citizens to provide first aid to someone experiencing a cardiac arrest. Additionally, we explained the different features of the app (eg, location-based technology), the goal of the app, and that such an app is currently being set up in Belgium. Next, we asked respondents if they ever heard of a first responder app (n=458, 13.5% heard of a first responder app) and to indicate on a 5-point Likert scale (1=totally disagree to 5=totally agree) to what extent the following statement applied to them: "If a first responder app for OHCA is available in Belgium, I am willing to install it." The mean score of 3.64 (SD 1.03) indicates a moderately high willingness to install the app.

Behavioral Determinants

UTAUT Determinants

Performance Expectancy

Based on the definition given by Venkatesh et al [9], performance expectancy was measured by a self-developed scale that asked respondents to indicate on a 5-point Likert scale (1=totally disagree to 5=totally agree) to what extent the following 2 statements applied to them: "This first responder app is an effective app for saving individuals experiencing a cardiac arrest" and "This first responder app would ensure that fewer people die from a cardiac arrest." In these statements, this refers to the particular first responder app for OHCA in Belgium that is being set up. We found a high correlation between the 2 items (r=0.75; P<.001) and merged them together in 1 variable. On average, participants scored 3.90 (SD 0.78), indicating a strong positive perception of the app's effectiveness.

Effort Expectancy

This was assessed using a 5-point Likert scale (1=totally disagree to 5=totally agree) item adapted from van der Waal et al [18], asking participants to rate the statement, "I think it will cost me a lot of time and energy to install this first responder app." The average score was 2.06 (SD 1.05), indicating that most participants found the app easy to install.

Facilitating Conditions

Facilitating conditions were measured with a 5-point Likert scale (1=totally disagree to 5=totally agree) item based on van der Waal et al [18], asking participants to rate the statement, "I have a smartphone with internet access, allowing me to install this first responder app." The mean score of 4.04 (SD 1.08) indicates high smartphone ownership and internet access among respondents.

Social Influence

Social influence was surveyed by measuring both injunctive and descriptive norms. We developed 2 items based on the definitions of Rimal and Real [39]. Injunctive norms were measured by asking respondents to indicate on a 5-point Likert scale (from 1=totally disagree to 5=totally agree) to what extent the following statement applied to them: "Most people in my environment would have a positive attitude toward this first responder app." Descriptive norms were measured by: "Most people in my environment would install this first responder app, if this app would be available in Belgium," and were evaluated on a 5-point Likert scale. Both norms were combined in social influences, as they were correlated (r=0.65; P<.001). With an average score of 3.28 (SD 0.78), the results indicate that social influences were moderately strong, meaning that the participants believed a fair number of people in their environment would install the first responder app and have a positive attitude toward it.

HBM Determinants

Self-Efficacy

A self-developed scale based on the definition of Bandura [30] asked respondents to indicate on a 5-point Likert scale (from 1=totally disagree to 5=totally agree) to what extent the following 2 statements applied to them: "It seems difficult to install this first responder app" and "I would be able to install this first responder app if I wanted to." The first item was recoded for analysis and then merged with the second item into 1 construct, as there was a significant correlation between the 2 items (r=0.45; P<.001). Participants scored high on self-efficacy (mean 3.92, SD 0.87), suggesting that they feel confident in their ability to successfully install a first responder app.

Perceived Susceptibility of an OHCA

Perceived susceptibility was assessed on both personal and societal levels using self-developed items based on El-Toukhy [40]. Respondents rated their own risk (personal perceived susceptibility) and the average person's risk (societal perceived susceptibility) of experiencing a cardiac arrest on a 5-point Likert scale (1=very low to 5=very high). Personal and societal susceptibility were correlated (r=0.56; P<.001) and combined into one measure with a mean of 2.90 (SD 0.64), indicating a moderate perceived susceptibility of an OHCA.

Perceived Severity of an OHCA

To measure respondents' perceived severity of an OHCA on a personal and societal level, a self-developed scale adapted from El-Toukhy [40] asked them to indicate on a 5-point Likert scale (from 1=totally disagree to 5=totally agree) to what extent the following 2 statements applied to them: "I believe that if I experience a cardiac arrest, it could have serious consequences for me" and "I believe that if someone experiences a cardiac arrest, it can have serious consequences for that person." The items, personal and societal perceived severity, respectively,

XSL•FO
RenderX

correlated with each other (r=0.71; P<.001), and a mean score was calculated by averaging the 2 items. On average, participants perceived an OHCA as very severe (mean 4.45, SD 0.74).

Perceived Barriers

Based on technology and health behavior literature [19,21,41], we created 14 different items that reflected various barriers that respondents could experience while installing a first responder app (Multimedia Appendix 1). Examples of these barriers included phone-related issues, digital difficulties, privacy violations, and credibility of the app. Respondents had to indicate on a 5-point Likert scale (from 1=totally disagree to 5=totally agree) to what extent the statements applied to them (eg, "I would not install this first responder app because I am afraid my privacy will not be guaranteed."). A manifest variable was created by summing the scores of all 14 items (range 14-70). Consequently, a low score on this variable indicates that a participant experiences few barriers, while a high score suggests that a participant encounters many barriers when it comes to installing a first responder app for OHCA. On average, participants scored 29.75 (SD 10.70), indicating that they generally did not feel substantially hindered in their ability to install the app.

Perceived Benefits

For perceived benefits, we developed 5 items that reflected various benefits respondents could perceive about installing a first responder app for OHCA (Multimedia Appendix 1), based on previous studies on altruistic health apps and providing CPR [22,32,42]. We measured benefits like personal relevance and altruistic satisfaction. Respondents had to indicate on a 5-point Likert scale (from 1=totally disagree to 5=totally agree) to what extent the statements applied to them (eg, "I would install this first responder app because it is my duty to help people."). A manifest variable was formed by adding the scores from all 5 items, resulting in a possible range of 5 to 25. A lower score on this variable reflects fewer perceived benefits, while a higher score indicates more benefits related to installing a first responder app. The average participant score was 18.17 (SD 3.47), suggesting that most respondents saw the app as advantageous.

Data Analysis

After data cleaning, reliability indices (Multimedia Appendix 2), descriptive statistics (Multimedia Appendix 3), and zero-order correlations (Multimedia Appendix 3) were calculated. Then, we developed a structural equation model (SEM) using the *Lavaan* package in R (R Foundation for Statistical Computing) to test the hypotheses and address the RQs.

All variables with 2 items were entered as latent variables, while single-item variables were entered as manifest variables. Subsequently, we specified the regression paths, which were drawn between the behavioral determinants (ie, performance expectancy [hypothesis 1], effort expectancy [hypothesis 2], facilitating conditions [hypothesis 3], social influence [hypothesis 4], self-efficacy [hypothesis 5], perceived susceptibility [hypothesis 6], perceived severity [hypothesis 7], perceived barriers [hypothesis 8], and perceived benefits [hypothesis 9]; independent variables) and the willingness to install a first responder app for OHCA (ie, dependent variable) to answer our 9 hypotheses. Additionally, to examine RQ1, paths were specified between the behavioral determinants as dependent variables and the influential factors (ie, demographic and health-related variables) as independent variables.

To explore RQ2 and examine the moderating role of the influential factors on the relationships between the behavioral determinants and the willingness to install a first responder app for OHCA, multiple group analyses were conducted. We created groups based on the median for age, SES, general health, and education. For age, we compared younger and older participants; for SES, we compared low and high SES groups; for health, we compared poor and good health; and for education, we compared lower and higher educational attainment. Gender was recoded as a dichotomous variable, as the group of 11 participants who selected "other" was too small for meaningful analysis; these responses were therefore treated as missing values. The variables for medical profession, CPR training, cardiovascular disease, and cardiovascular disease in one's environment were also dichotomous, meaning that participants were already divided into 2 distinct groups for each variable. The constrained model (ie, the model in which the paths of the behavioral determinants to willingness were specified to be the same across groups) was then compared with an unconstrained model (ie, the model in which the paths of the behavioral determinants to willingness were allowed to differ across groups), and a chi-square difference test was conducted to assess whether the model fit differed significantly between the groups.

The following indices were used to assess the model fit: chi-square/degrees of freedom, Comparative Fit Index, Tucker-Lewis Index, root mean square error of approximation (RMSEA), 90% CI for RMSEA, and standardized root mean square residual. The overall hypothesized model indicated an acceptable fit (χ^2_{95}=1230.8; P<.001; RMSEA=0.060; 90% CI 0.057-0.063; Comparative Fit Index=0.950; Tucker-Lewis Index=0.874; standardized root mean square residual=0.030). Although the chi-square (df) statistic and corresponding P value are reported, they were not used as the primary criterion for model acceptance or rejection, given that reliance on chi-square (df) alone can be overly restrictive in many research contexts. The study was preregistered on the Open Science Framework after data collection but before data analysis. The preregistration included the hypotheses and a detailed analysis plan (including model specifications, exclusion criteria, and handling of missing data). Additionally, an anonymized and cleaned version of the dataset and R code are available on Open Science Framework to enhance transparency and reproducibility of the analyses.

Results

Relations Between Behavioral Determinants and Willingness to Install a First Responder App for OHCA

Parameter estimates can be found in Table 1. First, regarding hypothesis 1 on the association between performance expectancy and willingness to install a first responder app for OHCA, we

XSL•FO
RenderX

found no significant relationship between the 2 variables. Hypothesis 2 was also rejected, as we did not find a significant association between effort expectancy and willingness to install a first responder app for OHCA. For hypothesis 3, we found that more facilitating conditions were related to a higher willingness to install a first responder app for OHCA, which consequently confirms the hypothesis. In addition, social influence positively predicted willingness to install a first responder app for OHCA, thus confirming hypothesis 4. However, there were no significant associations between willingness to install a first responder app for OHCA, self-efficacy (hypothesis 5), and perceived severity (hypothesis 7), thus refuting hypothesis 5 and hypothesis 7. Perceived susceptibility on the other hand was positively related with the willingness to install a first responder app for OHCA, confirming hypothesis 6. Finally, the strongest associations with willingness to install a first responder app for OHCA occurred for perceived barriers (hypothesis 8) and perceived benefits (hypothesis 9), confirming hypothesis 8 and hypothesis 9.

Table 1. Parameter estimates of the structural equation model: relations between behavioral determinants and willingness to install a first responder app (FRA) for out-of-hospital cardiac arrest[a].

Dependent variable and predictor	Estimate	SE	Standard estimate	*P* value
Intention FRA				
Performance expectancy	0.036	0.031	0.026	.24
Effort expectancy	–0.044	0.049	–0.045	.36
Facilitating conditions	0.062[b]	0.021	0.065[b]	.003
Social influences	0.228[c]	0.029	0.157[c]	<.001
Self-efficacy	–0.118	0.102	–0.097	.24
Perceived susceptibility	0.126[b]	0.042	0.061[b]	.003
Perceived severity	–0.033	0.024	–0.023	.17
Perceived barriers	–0.028[c]	0.003	–0.292[c]	<.001
Perceived benefits	0.111[c]	0.005	0.375[c]	<.001
Age	–0.006[c]	0.001	–0.090[c]	<.001
Gender[d]	0.022	0.027	0.011	.41
SES[e]	–0.017	0.011	–0.024	.12
Education level	–0.013	0.015	–0.012	.42
Medical profession[f]	0.055	0.047	0.016	.24
General health	0.002	0.001	0.028	.07
Cardiovascular disease[f]	0.029	0.040	0.010	.47
CPR[g] training[f]	–0.045	0.030	–0.021	.13
Cardiovascular disease in environment[f]	0.000	0.030	0.000	.99

[a]R^2 value is 0.446.

[b]P<.01.

[c]P<.001.

[d]Gender is coded: 1=woman, 2=man, and 3=other.

[e]SES: socioeconomic status.

[f]Medical profession, cardiovascular disease, CPR training, and cardiovascular disease in environment are coded: 1=yes and 0=no.

[g]CPR: cardiopulmonary resuscitation.

Relations Between the Influential Factors and the Behavioral Determinants

To answer RQ1, we examined how demographic and health variables related to the behavioral determinants of willingness to install a first responder app for OHCA (see Tables 2 and 3 for all parameter estimates).

XSL•FO
RenderX

Table 2. Parameter estimates of the structural equation model: relations between demographic variables and the behavioral determinants[a].

Dependent variable and predictor	Estimate	SE	Standard estimate	*P* value	R^2
Performance expectancy					0.015
Age	0.004[b]	0.001	0.093[b]	<.001	
Gender	–0.028	0.028	–0.019	.32	
SES[c]	0.007	0.011	0.014	.52	
Education level	–0.013	0.015	–0.018	.39	
Medical profession	–0.079	0.047	–0.032	.09	
Effort expectancy					0.095
Age	0.013[b]	0.001	0.203[b]	<.001	
Gender	–0.082	0.035	–0.040	.01	
SES	–0.034	0.014	–0.048	.01	
Education level	–0.099[b]	0.019	–0.095[b]	<.001	
Medical profession	0.366[b]	0.059	0.106[b]	<.001	
Facilitating conditions					0.073
Age	–0.012[b]	0.001	–0.190[b]	<.001	
Gender	0.026	0.036	0.012	.48	
SES	0.056[b]	0.015	0.076[b]	<.001	
Education level	0.082[b]	0.020	0.076[b]	<.001	
Medical profession	–0.167[d]	0.062	–0.047[d]	.007	
Social influences					0.015
Age	–0.003[b]	0.001	–0.076[b]	<.001	
Gender	0.007	0.027	0.005	.80	
SES	0.034[d]	0.011	0.070[d]	.002	
Education level	–0.015	0.015	–0.022	.30	
Medical profession	0.078	0.047	0.033	.09	
Self-efficacy					0.191
Age	–0.017[b]	0.001	–0.335[b]	<.001	
Gender	0.085[d]	0.032	0.051[d]	.009	
SES	0.028	0.013	0.049	.03	
Education level	0.110[b]	0.018	0.131[b]	<.001	
Medical profession	–0.358[b]	0.055	–0.129[b]	<.001	
Perceived susceptibility					0.428
Age	0.018[b]	0.001	0.594[b]	<.001	
Gender	0.018	0.014	0.018	.22	
SES	–0.015[d]	0.006	–0.045[d]	.008	
Education level	–0.057[b]	0.008	–0.115[b]	<.001	
Medical profession	0.123[b]	0.025	0.075[b]	<.001	
Perceived severity					0.068
Age	0.002	0.001	0.045	.02	

XSL•FO
RenderX

Dependent variable and predictor	Estimate	SE	Standard estimate	*P* value	R^2
Gender	–0.013	0.026	–0.009	.61	
SES	0.007	0.010	0.014	.51	
Education level	0.135[b]	0.014	0.195[b]	<.001	
Medical profession	–0.248[b]	0.044	–0.108[b]	<.001	
Perceived barriers					0.053
Age	–0.008	0.012	–0.012	.49	
Gender	–1.018[d]	0.360	–0.049[d]	.005	
SES	–0.377	0.146	–0.052	.01	
Education level	–0.726[b]	0.199	–0.068[b]	<.001	
Medical profession	2.882[b]	0.613	0.082[b]	<.001	
Perceived benefits					0.090
Age	0.003	0.004	0.014	.43	
Gender	0.158	0.115	0.023	.17	
SES	0.007	0.046	0.003	.89	
Education level	–0.115	0.063	–0.033	.07	
Medical profession	0.786[b]	0.195	0.068[b]	<.001	

[a]Gender is coded: 1=woman, 2=man, and 3=other. Medical profession is coded: 1=yes and 0=no.

[b]*P*<.001.

[c]SES: socioeconomic status.

[d]*P*<.01.

XSL•FO
RenderX

Table 3. Parameter estimates of the structural equation model: relations between health variables and the behavioral determinants[a].

Dependent variable and predictor	Estimate	SE	Standard estimate	*P* value	R^2
Performance expectancy					0.015
General health	0.003[b]	0.001	0.067[b]	.002	
Cardiovascular disease	–0.013	0.041	–0.006	.75	
CPR[c] training	0.087[b]	0.029	0.058[b]	.003	
Cardiovascular disease in environment	0.004	0.031	0.002	.90	
Effort expectancy					0.095
General health	–0.002	0.001	–0.043	.02	
Cardiovascular disease	0.098	0.052	0.034	.06	
CPR training	–0.211[d]	0.036	–0.099[d]	<.001	
Cardiovascular disease in environment	0.094	0.039	0.041	.02	
Facilitating conditions					0.073
General health	0.003[b]	0.001	0.053[b]	.005	
Cardiovascular disease	–0.018	0.054	–0.006	.73	
CPR training	0.154[d]	0.038	0.070[d]	<.001	
Cardiovascular disease in environment	0.025	0.041	0.011	.54	
Social influences					0.015
General health	0.002	0.001	0.040	.07	
Cardiovascular disease	0.013	0.041	0.007	.75	
CPR training	–0.072	0.029	–0.050	.01	
Cardiovascular disease in environment	0.000	0.031	0.000	.99	
Self-efficacy					0.191
General health	0.003[b]	0.001	0.058[b]	.009	
Cardiovascular disease	–0.071	0.049	–0.031	.14	
CPR training	0.180[d]	0.034	0.105[d]	<.001	
Cardiovascular disease in environment	–0.053	0.036	–0.029	.15	
Perceived susceptibility					0.428
General health	–0.003[d]	0.000	–0.094[d]	<.001	
Cardiovascular disease	0.038	0.022	0.027	.08	
CPR training	–0.031	0.015	–0.030	.04	
Cardiovascular disease in environment	0.056[b]	0.016	0.051[b]	.001	
Perceived severity					0.068
General health	0.001	0.001	0.038	.07	
Cardiovascular disease	0.052	0.038	0.027	.17	
CPR training	0.179[d]	0.027	0.126[d]	<.001	
Cardiovascular disease in environment	0.029	0.029	0.019	.31	
Perceived barriers					0.053
General health	–0.061[d]	0.011	–0.103[d]	<.001	
Cardiovascular disease	1.159	0.540	0.039	.03	
CPR training	–2.325[d]	0.380	–0.107[d]	<.001	

XSL•FO
RenderX

Dependent variable and predictor	Estimate	SE	Standard estimate	*P* value	R^2
Cardiovascular disease in environment	0.676	0.405	0.029	.09	
Perceived benefits					0.090
General health	0.014[d]	0.004	0.075[d]	<.001	
Cardiovascular disease	–0.062	0.172	–0.007	.72	
CPR training	1.818[d]	0.121	0.257[d]	<.001	
Cardiovascular disease in environment	0.530[d]	0.129	0.070[d]	<.001	

[a]Cardiovascular disease, CPR training, and cardiovascular disease in environment are coded: 1=yes and 0=no.

[b]*P*<.01.

[c]CPR: cardiopulmonary resuscitation.

[d]*P*<.001.

Demographic Variables

The results showed that practicing a medical profession was positively related to effort expectancy, perceived susceptibility, perceived barriers, and benefits but negatively with facilitating conditions, self-efficacy, and perceived severity (Table 2). Age was positively associated with performance expectancy, effort expectancy, and perceived susceptibility but negatively with facilitating conditions, social influences, and self-efficacy. Furthermore, education level correlated with 6 of the 9 behavioral determinants: negatively with effort expectancy, perceived susceptibility, and perceived barriers and positively with facilitating conditions, self-efficacy, and perceived severity. Gender was only positively related to self-efficacy and negatively related to perceived barriers. This means that men had higher levels of self-efficacy and women perceived more barriers to install a first responder app for OHCA. Finally, we found positive relationships between SES and facilitating conditions and social influences, whereas SES was negatively correlated with perceived susceptibility (Table 2).

Health Variables

General health proved to be an important variable, as the higher participants rated their health, the higher they scored on performance expectancy, facilitating conditions, self-efficacy, and perceived benefits, but the lower they scored on perceived susceptibility and barriers. Moreover, having done a CPR training was also positively related to 5 behavioral determinants (ie, performance expectancy, facilitating conditions, self-efficacy, perceived severity, and perceived benefits) but negatively with effort expectancy and perceived barriers. Surprisingly, we found no significant associations between any of the behavioral determinants and experiencing a cardiovascular disease oneself. Knowing someone in your close environment who is experiencing a cardiovascular disease, on the other hand, increased perceived susceptibility and perceived benefits of a first responder app for OHCA (Table 3).

We also explored the indirect relationships between the influential factors and willingness to install the app through the behavioral determinants. Only a limited number of these relationships were significant. Multimedia Appendix 4 reports the estimates.

The Moderating Role of the Influential Factors

Regarding RQ2, multiple group analyses were performed to examine potential moderating effects of the demographic and health variables (Table 4). Age was the only significant moderator in the SEM, with the chi-square difference test indicating a significant difference between the fits of the constrained and unconstrained model ($\Delta\chi^2_9$=23.690; *P*=.005). The post hoc tests revealed that age differences were only present for 2 relations within the model (Table 5). First, perceived severity was differently related to willingness to install a first responder app for OHCA. For young adults, there was no significant relation (β=.028; *P*=.49), and for older adults, there was a negative relationship (β=–.072; *P*=.03). Thus, for older people perceiving a cardiac arrest as severe related negatively to willingness to install the app. In addition, the association between perceived benefits and willingness to install a first responder app for OHCA was significantly stronger for older people (β=.123; *P*<.001) compared to younger people (β=.092; *P*<.001) .

XSL•FO
RenderX

Table 4. Multiple group analyses.

Moderator	$\Delta\chi^2$ (Δdf=9)	*P* value
Age	23.690[a]	.005
Gender	10.195	.34
SES[b]	7.478	.59
Education level	6.443	.70
Medical profession	11.397	.25
Health	8.091	.53
Cardiovascular disease	8.704	.47
CPR[c] training	9.901	.36
Cardiovascular disease in environment	4.890	.84

[a]*P*<.01.

[b]SES: socioeconomic status.

[c]CPR: cardiopulmonary resuscitation.

Table 5. Post hoc chi-square difference tests for age.

	$\Delta\chi^2$ (Δdf=1)	*P* value	β young	β old
Performance expectancy→intention to install FRA[a]	2.467	.17	—[b]	—
Effort expectancy→intention to install FRA	0.128	.72	—	—
Facilitating conditions→intention to install FRA	0.022	.88	—	—
Social influences→intention to install FRA	3.013	.08	—	—
Self-efficacy→intention to install FRA	0.885	.35	—	—
Perceived susceptibility→intention to install FRA	0.053	.82	—	—
Perceived severity→intention to install FRA	3.920[c]	.048	.028	–.072
Perceived barriers→intention to install FRA	0.161	.69	—	—
Perceived benefits→intention to install FRA	7.820[d]	.005	.092	.123

[a]FRA: first responder app.

[b]Not applicable.

[c]*P*<.05.

[d]*P*<.01.

Discussion

Overview

OHCAs are one of the leading causes of death worldwide, impacting approximately 350,000 people annually in Europe [7]. With survival rates ranging from 5% to 25% [7], first responder apps for OHCAs offer a vital solution to increase survival rates by mobilizing volunteer citizens to provide CPR in emergency situations. However, despite the emergence of these apps, little is still known about their adoption processes. Building on a comprehensive framework that integrates the UTAUT model, the HBM, and other contextual variables, this study therefore contributes to the literature by examining which factors influence citizens' willingness to install a first responder app for OHCA. While this framework suggests meaningful relationships between the variables studied, it is important to note that these are correlational and do not imply causality due to the cross-sectional nature of the study.

Main Findings Related to Hypotheses

This study offers partial evidence for the use of UTAUT determinants in understanding the adoption of first responder apps for OHCA, expanding our insights into how the model applies to altruistic mHealth apps. Specifically, only facilitating conditions (hypothesis 3) and social influence (hypothesis 4) were positively related to willingness to install a first responder app for OHCA, whereas performance and effort expectancy were not. First, aligning with facilitating conditions in studies on CTAs [18,19], people with internet and smartphone access were more willing to install a first responder app for OHCA, which is unsurprising given that these conditions are essential for installing and using a first responder app [19]. In addition, individuals who perceived more positive descriptive and injunctive norms were more willing to install a first responder

XSL•FO
RenderX

app for OHCA. This aligns with a systematic review on CTAs showing that social influence is one of the most significant predictors of novel app adoption [16]. Moreover, research shows that social influence is a crucial predictor of the intention to perform bystander CPR [43], which may also help clarify the relationship observed in this study.

Performance (hypothesis 1) and effort expectancy (hypothesis 2) were not significantly associated with willingness to install a first responder app for OHCA. While previous studies on CTAs and other health apps have reported mixed results for effort expectancy [5,18,44], performance expectancy has been classified as a strong predictor [13,15,17,18]. These insignificant findings in this study could be attributed to the fact that the first responder app has not yet been implemented in Belgium, thus making it hard for individuals to estimate the app's effectiveness (performance expectancy) and ease of use (effort expectancy) based solely on the brief description provided.

Regarding the HBM determinants, we identified 3 significant relationships (ie, perceived susceptibility, barriers, and benefits) with willingness to install a first responder app for OHCA. First, even though mixed findings occurred for perceived susceptibility in mHealth app literature [18,25,45], this variable was a significant predictor for the willingness to install a first responder app (hypothesis 6). This indicates that individuals who perceive themselves to be at a higher risk of experiencing a cardiac arrest are more inclined to adopt the app. Consequently, their personal risk assessment likely influences their perception of the app's importance in society. In addition, perceived barriers (hypothesis 8) and benefits (hypothesis 9) were significantly associated with willingness to install a first responder app for OHCA, with barriers negatively and benefits positively relating to willingness. In other words, the more barriers participants perceived, the less likely they were to install the app, while those who saw more benefits were more inclined to install. These findings align with previous research on health app adoption and bystander CPR, where barriers such as usability issues or privacy concerns [19,21,22,41], and perceived benefits like personal relevance and altruistic satisfaction [21,31,42], have been shown to be key predictors. Interestingly, perceived barriers and benefits emerged as the strongest predictors, highlighting their critical role in users' willingness to install a first responder app for OHCA.

Finally, there was no significant relationship between self-efficacy (hypothesis 5) or perceived severity (hypothesis 7) and willingness to install a first responder app for OHCA. The absence of a link with self-efficacy is surprising, as it contradicts studies on CTAs [18,26] and other mHealth technologies [46,47], where self-efficacy often emerged as a strong predictor, sometimes even the strongest. One possible explanation could lie in the difference between technology-related and health-related self-efficacy. In health research, self-efficacy often involves confidence in performing complex behaviors, like CPR. As such, it may not be self-efficacy related to feeling able to install the app, which can be fairly simple, but rather self-efficacy related to feeling capable of providing CPR if alerted, which may predict individuals' willingness to install the app. This study only measured technology-related self-efficacy, which might explain the absence of a significant link in the findings. Future studies should explicitly differentiate between technology-related and health-related self-efficacy to clarify their respective roles in shaping willingness to adopt first responder apps. Additionally, mixed results regarding perceived severity have been reported in the literature [18,48], suggesting variability in its influence across different health behaviors. In the context of first responder apps for OHCA, the lack of a significant relationship might be explained by the psychological and physical distance associated with such emergencies. Although citizens recognize that OHCA is a severe, life-threatening condition, they may perceive it as happening outside of their immediate, familiar environment and therefore do not feel a personal sense of responsibility. This feeling of detachment might lessen their willingness to install a first responder app.

Main Findings Related to RQs

We formulated 2 RQs to explore influential factors that may affect the behavioral determinants of willingness to install a first responder app for OHCA and to examine the potential moderating effects of these factors on the relationships between the determinants and willingness. As expected, age proved to be a significant variable for 2 main reasons. First, it was associated with 6 of the 9 behavioral determinants. Second, it was the only moderating factor in the model. More specifically, there was an age difference in the relation between perceived severity and willingness: no significant link was found for adults younger than 50 years of age, while a negative association was observed for those older than 50 years. Older adults who viewed cardiac arrest as severe were thus less willing to install a first responder app for OHCA. This finding is surprising; yet, it may suggest that older adults, who often have a more skeptical attitude toward technology [49], may associate severe medical emergencies with the need for professional intervention rather than assistance from lay responders. This perception could reduce their confidence in the app's usefulness or appropriateness. Alternatively, they may feel less personally capable or comfortable engaging with the app, especially in high-stakes situations. While these interpretations remain speculative, they highlight the need for further research into how age-related attitudes toward technology and emergency care shape the willingness to install health apps. Additionally, there was an age difference in the relationship between perceived benefits and willingness, with a stronger positive relation identified among older adults. The stronger positive relationship may indicate that older people, unlike digital natives, require clearer and more substantial advantages before they are willing to adopt new technologies. This aligns with Rogers' diffusion of innovations theory [50], which suggests that individuals assess the relative advantages and disadvantages of an innovation before deciding to adopt it. Moreover, it is in line with the UTAUT model that says that all relationships except for performance expectancy are stronger for older people [9,35].

In accordance with the HBM [10], also other demographic variables such as gender, SES, and education level were related to several of the behavioral determinants, and in that way indirectly to willingness to install a first responder app for OHCA. For example, younger people and men experienced more self-efficacy, which aligns with existing research on

XSL•FO
RenderX

technology-related self-efficacy [51,52], and people with a lower SES perceived themselves and their environment as more susceptible to a cardiac arrest. This relationship is logical, as general health showed a strong correlation with SES.

Health variables, although often not included within behavior change models [10,29], seemed to play a meaningful role. First of all, medical professionals perceived greater benefits of a first responder app for OHCA than those outside the medical field, which aligns with expectations, given the app's potential to improve patient outcomes [6]. However, they also reported more barriers, lower self-efficacy, higher effort expectancy, and lower scores on facilitating conditions and perceived severity. This pattern of results, where medical professionals score lower on multiple variables associated with willingness to install the app, may be attributed to several factors. First, they may be less technology-savvy or more resistant to health care digitalization, a trend supported by previous research [53]. Second, their familiarity with the consequences of a cardiac arrest could reduce the perceived severity, as the nature of their work may lead them to view a cardiac arrest as a more routine, manageable event. Third, they may have reservations about the general public's ability to provide effective CPR, leading to greater skepticism toward the app's effectiveness in nonprofessional hands. This skepticism may be unfounded, as individuals trained in CPR scored higher across 7 of the 9 behavioral determinants in this study, highlighting the potential benefits of promoting CPR training among the general public in first responder app adoption.

Furthermore, the general health status of individuals affected the majority of the behavioral determinants, but surprisingly, this was not the case for having cardiovascular disease. This is in contrast with the HBM, which states that prior experience or direct contact with a health issue typically influences perceptions and attitudes toward related health behaviors [10]. Indeed, one would assume that individuals with a cardiovascular disease, given their increased susceptibility to a cardiac arrest, would perceive themselves as more at risk and potentially recognize greater benefits in a first responder app for OHCA. Although we did find individual correlations between having a cardiovascular disease and several determinants (eg, positively related to effort expectancy, perceived susceptibility, and perceived barriers), their significance diminished when tested within the full model. Given the cross-sectional nature of our study, making it impossible to make claims about directionality, longitudinal research is needed to further clarify these correlations. On the contrary, knowing people in your environment who have a cardiovascular disease heightened individuals' awareness of personal risk and ensured more perceived benefits of a first responder app for OHCA.

Limitations

This study has several limitations that should be considered when interpreting the findings. First, as mentioned, the use of a cross-sectional survey design limits the ability to establish causal relationships. The relationships observed in the study should be interpreted as correlational, rather than causal. For example, while our model posits that social influence increases individuals' willingness to install a first responder app, it is also plausible that individuals who are already willing to install such an app subsequently develop stronger descriptive and injunctive norm perceptions. To further explore the directional effects and how all constructs within the model influence each other over time, future research should consider using longitudinal or experimental methodologies.

Second, many behavioral determinants were assessed with only 1 or 2 items, rather than more complex, multidimensional scales. This approach was intentional to minimize participant burden and was supported by literature in applied survey research and health psychology, which suggests that single-item measures can be as valid and reliable as their multifaceted counterparts [54]. However, we did not validate the single-item measures in our study, and future research is, therefore, recommended to validate these measures and compare them with multi-item scales to assess whether they perform equally well. This is particularly important, given that some single-item measures (eg, effort expectancy) targeted only specific facets of the constructs rather than capturing them in their entirety.

Third, while the study benefited from a large sample, the use of a web-based survey panel may have introduced self-selection biases. Individuals who join such panels and choose to participate in a CPR-related survey may be more technology-savvy and health-conscious than the general population. This could have led to an over- or underestimation of certain beliefs or attitudes toward the app. For instance, participants may have reported a higher willingness to install the app, as technology-savvy individuals may be more comfortable downloading and using health apps compared to the general public. Future studies should, therefore, consider using alternative or mixed method recruitment strategies, such as random digit dialing or community-based sampling, to validate the results in more diverse and representative populations.

Finally, as there is no existing first responder app for OHCA in Belgium, the study could not measure behavioral intention or actual app use, which is typically measured in studies applying models like UTAUT or the HBM. As a result, we were unable to determine whether participants' stated willingness would translate into real-world behavior. Additionally, the use of self-reported willingness to install the app may have been subjected to social desirability bias. For example, participants may have overstated their willingness to install the app to appear socially responsible or helpful, especially given the focus on emergency health situations. This intention-behavior gap is well documented in health behavior research and highlights an important limitation of the study. Future research should replicate this work in countries where such apps are already in use to examine whether the identified determinants also predict actual behaviors, such as installation rates, retention over time, and the frequency of CPR interventions triggered by the app.

Implications and Future Research

The findings of this study have both theoretical and practical implications. Theoretically, it finds its relevance in extending the UTAUT and HBM to altruistic mHealth apps that require active involvement, which has not been done before. Moreover, this study contributes to the field by developing a comprehensive

XSL•FO
RenderX

framework that also integrates relevant demographic and health variables, offering a more nuanced understanding of the factors influencing the first responder app for OHCA adoption. Practically, this study was conducted in Belgium, which is a country that currently lacks an established first responder system. As such, the knowledge provided by the study provides practical guidance regarding how to convince citizens in Belgium to participate in the first responder system that health care professionals are currently deploying. Building on this, future research could further enhance these efforts by designing and testing health messages in an experimental setting based on the key determinants of willingness to install a first responder app that we identified in this study. In that way, subsequent studies can refine communication strategies for real-life campaigns. Additionally, it would be valuable to replicate the proposed model in countries where first responder apps for OHCA are already available. Such studies could examine actual adoption behaviors and compare the model's applicability across different contexts.

This study provides an important first step in mapping the behavioral, demographic, and health-related factors of the Belgian population that influence the willingness to install a first responder app for OHCA. By identifying key behavioral determinants, such as social influence and perceived barriers or benefits, and other sociodemographic and health-related factors, we offer actionable insights that governments can use to shape implementation strategies. For example, campaigns could emphasize the app's societal benefits and social norms around helping others while also addressing privacy or usability concerns. Additionally, app developers may draw on these insights to design a more accessible and user-friendly first responder app for OHCA. Our findings also point to the need for tailored messaging, as one-size-fits-all communication strategies may not be effective for the entire Belgian population. For instance, older adults may respond differently to a message focused on a certain behavioral determinant than younger adults, suggesting that future research should explore how best to adapt outreach efforts to diverse groups. Future efforts to roll out such technology should be grounded in this type of population-informed research, ensuring that interventions are aligned with what people value and what might hold them back.

Acknowledgments

Generative artificial intelligence (ChatGPT) was used to improve the language and clarity of certain sections, as all authors are nonnative English speakers. However, it was not used to generate original content, ideas, or code. This work was supported by grants from Interne Fondsen KU Leuven (Internal Funds KU Leuven; CITSC/24/012). The authors thankfully acknowledge their support.

Data Availability

The datasets generated or analyzed during this study are available in the Open Science Framework repository (osf.io/6k7b5).

Authors' Contributions

The conceptualization of the study was carried out by KB, LS, RV, HS, JDS, MS, and OH. Methodological planning, including the survey design and analytical strategy, was performed by CvW, KB, LS, and RV. CvW was primarily responsible for data curation and conducted the formal analysis with support from LS and RV. CvW also wrote the initial draft of the manuscript. Subsequent reviewing and editing of the manuscript were performed by CvW, RV, LS, and KB. Supervision was provided throughout the research process by KB, LS, RV, HS, JDS, MS, and OH. Funding for the study was acquired by KB, LS, RV, HS, JDS, MS, and OH. All authors reviewed and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Items of the behavioral determinants.
[DOCX File , 15 KB - humanfactors_v12i1e69934_app1.docx]

Multimedia Appendix 2
Reliability of multi-item constructs.
[DOCX File , 15 KB - humanfactors_v12i1e69934_app2.docx]

Multimedia Appendix 3
Statistics and zero-order correlations.
[DOCX File , 24 KB - humanfactors_v12i1e69934_app3.docx]

XSL•FO
RenderX

Multimedia Appendix 4
Indirect relations.
[DOCX File , 23 KB - humanfactors_v12i1e69934_app4.docx]

References

1. mHealth: use of appropriate digital technologies for public health: report by the director-general. World Health Assembly 71. 2018. URL: https://iris.who.int/handle/10665/276430 [accessed 2024-06-21]
2. From innovation to implementation: eHealth in the WHO European region. World Health Organization. Regional Office for Europe. 2016. URL: https://iris.who.int/handle/10665/326317 [accessed 2024-06-21]
3. Hsu A, Hsu YF. Comprehensive analysis of dieting apps: effectiveness, design, and frequency usage. 2023 Presented at: IEEE 47th Annual Computers, Software, and Applications Conference (COMPSAC); June 26-30, 2023; Torino, Italy p. 549-557. [doi: 10.1109/compsac57700.2023.00079]
4. Middelweerd A, Mollee JS, van der Wal CN, Brug J, Te Velde SJ. Apps to promote physical activity among adults: a review and content analysis. Int J Behav Nutr Phys Act 2014;11(1):97 [FREE Full text] [doi: 10.1186/s12966-014-0097-9] [Medline: 25059981]
5. Walrave M, Waeterloos C, Ponnet K. Ready or not for contact tracing? Investigating the adoption intention of COVID-19 contact-tracing technology using an extended unified theory of acceptance and use of technology model. Cyberpsychol Behav Soc Netw 2021;24(6):377-383. [doi: 10.1089/cyber.2020.0483] [Medline: 33017171]
6. Scquizzato T, Belloni O, Semeraro F, Greif R, Metelmann C, Landoni G, et al. Dispatching citizens as first responders to out-of-hospital cardiac arrests: a systematic review and meta-analysis. Eur J Emerg Med 2022;29(3):163-172. [doi: 10.1097/MEJ.0000000000000915] [Medline: 35283448]
7. Gräsner JT, Wnent J, Herlitz J, Perkins GD, Lefering R, Tjelmeland I, et al. Survival after out-of-hospital cardiac arrest in Europe—results of the EuReCa TWO study. Resuscitation 2020;148:218-226. [doi: 10.1016/j.resuscitation.2019.12.042] [Medline: 32027980]
8. Derkenne C, Jost D, Roquet F, Dardel P, Kedzierewicz R, Mignon A, et al. Mobile smartphone technology is associated with out-of-hospital cardiac arrest survival improvement: the first year "Greater Paris Fire Brigade" experience. Acad Emerg Med 2020;27(10):951-962 [FREE Full text] [doi: 10.1111/acem.13987] [Medline: 32445436]
9. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425-478. [doi: 10.2307/30036540]
10. Rosenstock IM. The health belief model and preventive health behavior. Health Educ Monogr 1974;2(4):354-386. [doi: 10.1177/109019817400200405]
11. Saprikis V, Avlogiaris G, Katarachia A. A comparative study of users versus non-users' behavioral intention towards M-banking apps' adoption. Information 2022;13(1):30. [doi: 10.3390/info13010030]
12. Mitchell KM, Holtz BE, McCarroll AM. Assessing college students' perceptions of and intentions to use a mobile app for mental health. Telemed J E Health 2022;28(4):566-574. [doi: 10.1089/tmj.2021.0106] [Medline: 34166099]
13. Wei J, Vinnikova A, Lu L, Xu J. Understanding and predicting the adoption of fitness mobile apps: evidence from China. Health Commun 2021;36(8):950-961. [doi: 10.1080/10410236.2020.1724637] [Medline: 32041437]
14. Gupta A, Dogra N, George B. What determines tourist adoption of smartphone apps? J Hosp Tour Technol 2018;9(1):50-64. [doi: 10.1108/jhtt-02-2017-0013]
15. Chopdar PK. Adoption of Covid-19 contact tracing app by extending UTAUT theory: perceived disease threat as moderator. Health Policy Technol 2022;11(3):100651 [FREE Full text] [doi: 10.1016/j.hlpt.2022.100651] [Medline: 35855013]
16. Sujarwoto S, Maharani A. Facilitators and barriers to the adoption of mHealth apps for COVID-19 contact tracing: a systematic review of the literature. Front Public Health 2023;11:1222600 [FREE Full text] [doi: 10.3389/fpubh.2023.1222600] [Medline: 38145061]
17. Tomczyk S, Barth S, Schmidt S, Muehlan H. Utilizing health behavior change and technology acceptance models to predict the adoption of COVID-19 contact tracing apps: cross-sectional survey study. J Med Internet Res 2021;23(5):e25447 [FREE Full text] [doi: 10.2196/25447] [Medline: 33882016]
18. van der Waal NE, de Wit J, Bol N, Ebbers W, Hooft L, Metting E, et al. Predictors of contact tracing app adoption: integrating the UTAUT, HBM and contextual factors. Technol Soc 2022;71:102101. [doi: 10.1016/j.techsoc.2022.102101]
19. Walrave M, Waeterloos C, Ponnet K. Reasons for nonuse, discontinuation of use, and acceptance of additional functionalities of a COVID-19 contact tracing app: cross-sectional survey study. JMIR Public Health Surveill 2022;8(1):e22113 [FREE Full text] [doi: 10.2196/22113] [Medline: 34794117]
20. Gibbons FX, Gerrard M. Predicting young adults' health risk behavior. J Pers Soc Psychol 1995;69(3):505-517. [doi: 10.1037//0022-3514.69.3.505] [Medline: 7562392]
21. Rosenstock IM, Strecher VJ, Becker MH. Social learning theory and the health belief model. Health Educ Q 1988;15(2):175-183. [doi: 10.1177/109019818801500203] [Medline: 3378902]
22. Smoll NR, Walker J, Khandaker G. The barriers and enablers to downloading the COVIDSafe app—a topic modelling analysis. Aust N Z J Public Health 2021;45(4):344-347 [FREE Full text] [doi: 10.1111/1753-6405.13119] [Medline: 33970555]

XSL•FO
RenderX

23. Ayres-Pereira V, Pirrone A, Korbmacher M, Tjostheim I, Böhm G. The privacy and control paradoxes in the context of smartphone apps. Front Comput Sci 2022 Sep 9;4:1-17. [doi: 10.3389/fcomp.2022.986138]
24. Calaguas NP. Predictors of condom use among gay and bisexual men in the Philippines. Int J Sex Health 2020;32(3):188-198. [doi: 10.1080/19317611.2020.1770392]
25. Alpar P, Driebe T. Patients' attitudes toward apps for management of a chronic disease. In: Ahlemann F, Schütte R, Stieglitz S, editors. Innovation Through Information Systems. Berlin: Springer International Publishing; 2021:22-37.
26. Walrave M, Waeterloos C, Ponnet K. Adoption of a contact tracing app for containing COVID-19: a health belief model approach. JMIR Public Health Surveill 2020;6(3):e20572 [FREE Full text] [doi: 10.2196/20572] [Medline: 32755882]
27. Subedi S, Leal Filho W, Adedeji A. An assessment of the health belief model (HBM) properties as predictors of COVID-19 preventive behaviour. J Public Health (Berl) 2023;33(6):1329-1339. [doi: 10.1007/s10389-023-02109-7]
28. Saranya R, Kathirvel S. Chapter 1—principles and approaches in public health practice. In: Kathirvel S, Singh S, Chockalingam A, editors. Principles and Application of Evidence-Based Public Health Practice. Cambridge, MA: Academic Press; 2024:3-21.
29. Ajzen I. The theory of planned behavior. Organ Behav Hum Decis Process 1991;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
30. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Adv Behav Res Ther 1978;1(4):139-161. [doi: 10.1016/0146-6402(78)90002-4]
31. Bechard LE, Bergelt M, Neudorf B, DeSouza TC, Middleton LE. Using the health belief model to understand age differences in perceptions and responses to the COVID-19 pandemic. Front Psychol 2021;12:609893 [FREE Full text] [doi: 10.3389/fpsyg.2021.609893] [Medline: 33935865]
32. Daud A, Nawi A, Aizuddin A, Yahya M. Factors and barriers on cardiopulmonary resuscitation and automated external defibrillator willingness to use among the community: a 2016-2021 systematic review and data synthesis. Glob Heart 2023;18(1):46 [FREE Full text] [doi: 10.5334/gh.1255] [Medline: 37649652]
33. Dobbie F, Uny I, Eadie D, Duncan E, Stead M, Bauld L, et al. Barriers to bystander CPR in deprived communities: findings from a qualitative study. PLoS One 2020;15(6):e0233675 [FREE Full text] [doi: 10.1371/journal.pone.0233675] [Medline: 32520938]
34. Nunes A, Limpo T, Castro SL. Acceptance of mobile health applications: examining key determinants and moderators. Front Psychol 2019;10:2791 [FREE Full text] [doi: 10.3389/fpsyg.2019.02791] [Medline: 31920836]
35. Venkatesh V, Thong JYL, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Q 2012;36(1):157-178. [doi: 10.2307/41410412]
36. Gilda V, Putra RR. Moderating role of accounting knowledge and education level on UTAUT2 model in MSMEs. Owner 2024;8(1):769-783. [doi: 10.33395/owner.v8i1.1797]
37. Adler NE, Epel ES, Castellazzo G, Ickovics JR. Relationship of subjective and objective social status with psychological and physiological functioning: preliminary data in healthy White women. Health Psychol 2000;19(6):586-592. [doi: 10.1037//0278-6133.19.6.586] [Medline: 11129362]
38. Moons P, Van Deyk K, De Bleser L, Marquet K, Raes E, De Geest S, et al. Quality of life and health status in adults with congenital heart disease: a direct comparison with healthy counterparts. Eur J Cardiovasc Prev Rehabil 2006;13(3):407-413. [doi: 10.1097/00149831-200606000-00017]
39. Rimal RN, Real K. How behaviors are influenced by perceived norms: a test of the theory of normative social behavior. Commun Res 2005;32(3):389-414. [doi: 10.1177/0093650205275385]
40. El-Toukhy S. Parsing susceptibility and severity dimensions of health risk perceptions. J Health Commun 2015;20(5):499-511. [doi: 10.1080/10810730.2014.989342] [Medline: 25749555]
41. Garrett PM, White JP, Lewandowsky S, Kashima Y, Perfors A, Little DR, et al. The acceptability and uptake of smartphone tracking for COVID-19 in Australia. PLoS One 2021;16(1):e0244827 [FREE Full text] [doi: 10.1371/journal.pone.0244827] [Medline: 33481841]
42. Sun M, Waters CM, Zhu A. Public willingness, attitudes and related factors toward cardiopulmonary resuscitation: a grounded theory study. Public Health Nurs 2024;41(2):233-244. [doi: 10.1111/phn.13271] [Medline: 38111292]
43. Magid KH, Ranney ML, Risica PM. Using the theory of planned behavior to understand intentions to perform bystander CPR among college students. J Am Coll Health 2021;69(1):47-52. [doi: 10.1080/07448481.2019.1651729] [Medline: 31483211]
44. Holtz BE, Kanthawala S, Martin K, Nelson V, Parrott S. Young adults' adoption and use of mental health apps: efficient, effective, but no replacement for in-person care. J Am Coll Health 2025;73(2):602-610. [doi: 10.1080/07448481.2023.2227727] [Medline: 37399569]
45. Sutarno M. The interaction of perceived susceptibility with predictors of mHealth technology usage. Am J Health Behav 2022;46(6):716-728. [doi: 10.5993/AJHB.46.6.14] [Medline: 36721284]
46. Balapour A, Reychav I, Sabherwal R, Azuri J. Mobile technology identity and self-efficacy: implications for the adoption of clinically supported mobile health apps. Int J Inf Manag 2019;49:58-68. [doi: 10.1016/j.ijinfomgt.2019.03.005]

XSL•FO
RenderX

47. Zhang X, Han X, Dang Y, Meng F, Guo X, Lin J. User acceptance of mobile health services from users' perspectives: the role of self-efficacy and response-efficacy in technology acceptance. Inform Health Soc Care 2017;42(2):194-206. [doi: 10.1080/17538157.2016.1200053] [Medline: 27564428]
48. Wang X, Lee C, Jiang J, Zhu X. Factors influencing the aged in the use of mobile healthcare applications: an empirical study in China. Healthcare (Basel) 2023;11(3):396 [FREE Full text] [doi: 10.3390/healthcare11030396] [Medline: 36766970]
49. Digitale trends in Vlaanderen. imec. 2023. URL: https://www.imec.be/sites/default/files/2024-03/imec%20digimeter%202023%20Rapport.pdf [accessed 2024-10-30]
50. Rogers EM. Diffusion of Innovations. 3rd Edition. New York, NY: Free Press; 1983.
51. Aslam I, Arzeen N, Arzeen S, Muhammad H. Effects of gender and age on the level of students' ICT self-efficacy, self-directed learning, E-learning readiness, and engagement. Clin Couns Psychol Rev 2022;3(1):59-77. [doi: 10.32350/ccpr.31.05]
52. Gröstenberger E, Selinger A. The effect of gender and age on computer self-efficacy, computer anxiety and perceived enjoyment among Austrian secondary school teachers. MAP Educ Humanit 2023;4(1):1-9. [doi: 10.53880/2744-2373.2023.4.1]
53. Iyanna S, Kaur P, Ractham P, Talwar S, Najmul Islam A. Digital transformation of healthcare sector. What is impeding adoption and continued usage of technology-driven innovations by end-users? J Bus Res 2022;153:150-161. [doi: 10.1016/j.jbusres.2022.08.007]
54. Allen MS, Iliescu D, Greiff S. Single item measures in psychological science. Eur J Psychol Assess 2022;38(1):1-5. [doi: 10.1027/1015-5759/a000699]

Abbreviations

CPR: cardiopulmonary resuscitation
CTA: contact tracing app
HBM: health belief model
mHealth: mobile health
OHCA: out-of-hospital cardiac arrest
RMSEA: root mean square error of approximation
RQ: research question
SEM: structural equation model
SES: socioeconomic status
UTAUT: unified theory of acceptance and use of technology

Edited by A Kushniruk; submitted 11.12.24; peer-reviewed by A Ort, E Engel; comments to author 12.03.25; revised version received 09.05.25; accepted 10.07.25; published 09.09.25.

Please cite as:
von Winckelmann C, Vanherle R, Schreurs L, Hoogmartens O, Salaets H, De Spiegeleer J, Sabbe M, Beullens K
Understanding the Behavioral Determinants of First Responder App Adoption by Integrating Perspectives From the Unified Theory of Acceptance and Use of Technology and Health Belief Model: Cross-Sectional Survey
JMIR Hum Factors 2025;12:e69934
URL: https://humanfactors.jmir.org/2025/1/e69934
doi:10.2196/69934
PMID:

©Cas von Winckelmann, Robyn Vanherle, Lara Schreurs, Olivier Hoogmartens, Heidi Salaets, Jan De Spiegeleer, Marc Sabbe, Kathleen Beullens. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 09.09.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Participants' Perspectives on the iCareBreast Mobile-Based Perioperative Care Program for Women Undergoing Breast Cancer Surgery: Qualitative Process Evaluation

Yan Pang[1], MSc; Honggu He[1,2], BScMed, MSc, PhD; Minna Pikkarainen[3,4], PhD; Swee-Ho Lim[5], MBBS, MMed(Surgery), FRCSEd

[1]Alice Lee Centre for Nursing Studies, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore

[2]National University Health System, Singapore, Singapore

[3]Digitalization of Healthcare Services, Oslomet, Oslo Metropolitan University, Oslo, Norway

[4]Connected Health, Martti Ahtisaari Institute, University of Oulu, Oulu, Finland

[5]KK Women's and Children's Hospital, Singapore, Singapore

Corresponding Author:
Honggu He, BScMed, MSc, PhD
Alice Lee Centre for Nursing Studies
Yong Loo Lin School of Medicine
National University of Singapore
Level 5, Centre for Translational Medicine, Block MD6, 14 Medical Drive
Singapore, 117599
Singapore
Phone: 65 65167448
Fax: 65 67767135
Email: nurhhg@nus.edu.sg

Abstract

Background: Breast cancer treatment, particularly during the perioperative period, is often accompanied by significant psychological distress, including anxiety and uncertainty. Mobile health (mHealth) interventions have emerged as promising tools to provide timely psychosocial support through convenient, flexible, and personalized platforms. While research has explored the use of mHealth in breast cancer prevention, care management, and survivorship, few studies have examined patients' experiences with mobile interventions during the perioperative phase of breast cancer treatment.

Objective: This study aimed to explore the experiences of patients with breast cancer using iCareBreast, a mobile app designed to provide perioperative guidance and psychosocial support.

Methods: A qualitative approach was used to explore participant experiences. A total of 13 English- or Chinese-speaking participants from the intervention group of a clinical study were recruited via purposive sampling between April 2021 and February 2022. Semistructured individual phone interviews were conducted, audio-recorded, and transcribed verbatim. Thematic analysis was performed to identify key patterns of experience, focusing on usability, emotional impact, perceived value, and areas for future improvement.

Results: Overall, 4 main themes and 11 subthemes emerged from this study: (1) navigating the app with confidence and comfort, (2) making sense of treatment through relevant and evolving content, (3) finding emotional anchors in a time of uncertainty, and (4) advocating for broader use and continued motivation. Participants found the app user-friendly and appreciated its structure and locally relevant content, which helped reduce anxiety and enhance surgical preparedness. Features such as deep breathing exercises, motivational quotes, survivor stories, mindfulness practices, and peer support links offered emotional comfort and a sense of companionship. Participants strongly advocated for more personalized and adaptive content aligned with their treatment type and recovery progress. They also emphasized the value of interactive elements, such as video demonstrations and accessing messaging functions, to support sustained engagement. Many expressed the need for extended support throughout the adjuvant treatment phases, including chemotherapy and radiotherapy.

Conclusions: The iCareBreast app was perceived as a supportive tool during the perioperative period, helping patients navigate both informational and emotional challenges. However, the findings underscore the importance of extending content across the treatment continuum and enhancing personalization and interactivity. mHealth interventions should be responsive to patients'

XSL•FO
RenderX

evolving needs and integrated into clinical care pathways to provide timely, comprehensive, tailored, and ongoing support for women with breast cancer.

(JMIR Hum Factors 2025;12:e71686) doi:10.2196/71686

KEYWORDS

breast cancer; mobile health; psychosocial intervention; participant experience; qualitative study

Introduction

Background

Breast cancer is one of the most prevalent cancers among women worldwide and remains the leading cause of cancer diagnosis and mortality in Singapore [1,2]. Despite advancements in treatment, patients frequently experience heightened anxiety, depression, and uncertainty as they navigate surgical procedures and recovery [3-5]. Psychological distress affects approximately 50% of patients with breast cancer and is closely linked to factors such as short time since diagnosis, being in the preoperative phase, and concerns about body image [6]. Preoperative depressive symptoms are particularly concerning, with rates as high as 39% among patients undergoing mastectomy or lumpectomy, particularly in Asian countries [7]. These statistics highlight the significant psychological challenges faced by women during the perioperative period.

In addition to emotional distress, many patients report unmet information needs during the active treatment phase. During treatment preparation, patients often seek detailed explanations about the treatment plan and procedures, whereas their focus shifts to managing side effects and understanding long-term outcomes during treatment [8]. Supportive and goal-oriented strategies before surgery, including psychosocial interventions, have been shown to effectively reduce stress and improve the quality of life for patients with breast cancer [9-11]. These findings underscore the critical importance of comprehensive psychosocial support in addressing psychological distress and unmet information needs, ultimately improving outcomes during the perioperative period.

Traditional face-to-face psychosocial support mechanisms may not always be sufficient and convenient to meet the needs of patients with breast cancer, particularly during the brief and emotionally intense period from diagnosis to surgery. Mobile health (mHealth) interventions have emerged as innovative solutions that offer timely psychosocial support through accessible digital platforms. These interventions leverage flexibility, convenience, and cost-effectiveness to deliver personalized care. Research highlights mHealth's value across breast cancer care, including prevention, early detection, care management, and survivorship. Through interactive features such as pictures, audios, videos, and tracking diaries, these platforms provide educational resources, monitor health, promote lifestyle changes, facilitate emotional support, and enable communication with health care professionals (HCPs) [12]. Despite their potential, psychosocial interventions specifically targeting patients with breast cancer during the perioperative period are limited, particularly those delivered through mHealth or eHealth platforms. Furthermore, there is a limited understanding of patients' preferences for such solutions [13]. To address this gap, we developed an innovative e-supportive care mobile app, iCareBreast, designed for patients with breast cancer undergoing surgery. On the basis of the self-efficacy theory proposed by Bandura [14], the app aims to enhance patients' self-efficacy in perioperative care while addressing their specific needs during this critical period.

Objectives

Understanding participant experiences with these technologies is essential for uncovering the mechanisms behind their impact, optimizing their design to better meet patients' needs, and guiding future implementation [13,15]. This process evaluation focused on exploring the experiences of patients with breast cancer who used the iCareBreast mobile app, designed to deliver care support during their surgical journey. By gathering qualitative insights into each component of the app, the evaluation aimed to identify user experience, focusing on usability, strengths, weaknesses, and suggestions for future improvements. These insights are critical for refining the app and ensuring that it effectively addresses patients' needs, ultimately enhancing their perioperative care experience.

Methods

Study Design and Participants

This descriptive qualitative study is a follow-up to a randomized controlled trial (ClinicalTrials.gov ID NCT04172350) [16] that evaluated the effectiveness of iCareBreast, a mobile-based perioperative care program, on various health outcomes in improving perioperative self-efficacy, anxiety, depression, fatigue, health-related quality of life, and perioperative care satisfaction among women undergoing breast surgery in a public hospital in Singapore. The randomized controlled trial enrolled 123 women with early-stage breast cancer, randomized to either the iCareBreast intervention group (n=62) or the control group (n=61). While no statistically significant differences were found between the groups for most health outcomes, participants in the iCareBreast group reported higher perioperative care satisfaction.

All participants received standard multidisciplinary preoperative care, including consultations with a breast surgeon. The number and duration of consultations varied depending on case complexity and individual patient concerns. Patients were typically referred to a breast care nurse (BCN), who served as a nurse navigator providing education, emotional support, and care coordination. A short hospital stay (1-2 days) following breast-conserving surgery or mastectomy was typically required to ensure adequate pain control, postoperative monitoring, and wound care education. Postoperative follow-up appointments were scheduled with the breast surgeon. Medical social workers

XSL•FO
RenderX

were available, and referrals were made based on clinical judgment and patient-specific needs. In addition to standard care, participants in the intervention group used the iCareBreast app daily at home.

iCareBreast e-Support Care Program

The iCareBreast app was an interactive tool designed to guide and support patients through their perioperative journey. Tailored to each patient's surgery date, the app provided daily guidance and information over a 29-day period (14 days before surgery, the day of surgery, and 14 days after surgery). Aiming to enhance care management and patient engagement for both surgery preparation and recovery, iCareBreast included 4 main functions: educational content, perioperative care guidance, psychological support, and social support.

The educational components covered presurgery and postsurgery topics relevant to breast cancer, such as anatomy, treatment options, and anesthesia. The perioperative care included preoperative preparation, day-of-surgery guidance, and postoperative care guidance on wound and drain management, pain control, infection monitoring, physiotherapy, and discharge planning. The psychosocial support component provided daily motivational quotes, mindfulness exercises, and stories from breast cancer survivors, along with a summary of breast cancer support networks available in Singapore. In addition, participants could ask questions through the app's messaging system, with responses provided by HCPs and researchers within 24 hours.

The iCareBreast app was downloaded from the Google Play Store (for Android) or the Apple App Store (for iOS), available in participants' preferred language—either English or Mandarin. A unique activation code was used for login, ensuring confidentiality and preventing cross-group contamination. Further details have been published elsewhere [16].

Recruitment

Purposive sampling was used to select participants from the intervention group, categorized by age groups (21-49, 50-64, and ≥65 years), to ensure representation across different levels of technology competency. Participants in the intervention group were purposely invited to the interview phase at the same time they were enrolled in the main study. Interviews were conducted 2 weeks after surgery (immediately following the intervention) until data saturation was reached, which occurred with the 10th participant. Three additional patients with breast cancer were interviewed to confirm that no new information emerged, resulting in a total of 13 participants enrolled. All participants were informed that participation in the interview was optional and would not impact their involvement in the main study. They were also informed that the interviews would take approximately 10 to 20 minutes and would be audio-recorded.

Data Collection

All semistructured interviews were conducted via phone calls during the COVID-19 period (from April 2021 to February 2022). Interviews lasted an average of 20 minutes, with duration ranging from 12 to 32 minutes. An interview guide was developed to facilitate discussions about participants' experiences using the app, focusing on usability, strengths, weaknesses, and suggestions for future improvements.

Approximately 2 weeks after surgery, when participants had just completed their app use, a female research assistant who had no prior contact with participants trained in qualitative interviewing conducted individual phone interviews at mutually convenient times for both the interviewer and participants. The interview guide (Multimedia Appendix 1) was developed by an academic expert specializing in eHealth interventions, with another academic expert providing feedback on content clarity. A total of 10 interviews were conducted in English, and the remaining 3 in Mandarin. The Mandarin interviews were transcribed verbatim, then back translated into English by the same bilingual research assistant and verified for accuracy by another bilingual team member. Field notes were also taken to capture nonverbal cues (such as crying or laughing) to supplement the transcripts.

Data Analysis

Data were analyzed following the 6 phases of thematic analysis described by Braun and Clarke [17]. This approach enabled us to explore and interpret patterns of meaning across the dataset, with an emphasis on understanding participants' experiences and perspectives rather than counting specific responses. All interviews were transcribed verbatim by the same interviewer (research assistant) and cross-checked for accuracy by another team member. Data were analyzed concurrently with collection using inductive thematic analysis [17]. Two independent coders immersed themselves in the transcripts, repeatedly reading them. Codes were developed inductively from the data, and themes were shaped based on the depth and recurrence of meaning both within and across interviews. Initial coding was conducted independently by both coders line-by-line in the raw Microsoft Word documents, after which the codes were transferred to XMind Zen software to develop themes and subthemes related to participant experience [17,18]. A thematic map illustrating the connections between themes and subthemes was created after a third researcher reviewed and harmonized the findings [17]. Study rigor and reliability were maintained through prolonged engagement, peer debriefing, and by taking an audit trail and field notes [19]. Study finding was reported according to the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist (Multimedia Appendix 2) [20].

Ethical Considerations

Ethics approval was obtained from the SingHealth Centralised Institutional Review Board (reference number 2019/2632). In accordance with the principles of the Helsinki Declaration, the participants were fully informed about the study, given ample time for consideration, and assured of their voluntary participation and right to withdraw at any time. All participants provided informed consent, and their identities were anonymized. They were then randomly assigned to either the intervention or control group. No compensation was provided to the participants.

XSL•FO
RenderX

Results

Participant Demographics

A total of 13 participants were interviewed for this study, all of whom were women with breast cancer undergoing breast surgery. The mean age was 58.5 (SD 11.8) years, with an age range of 35 to 73 years. Overall, 31% (4/13) of the participants were aged 21 to 49 years, 31% (4/13) were aged 50 to 64 years, and the remaining 38% (5/13) were aged ≥65 years. The average hospital stay was 2.85 days. Of the 13 participants, most were Chinese (n=11, 85%), with 1 (8%) Malay and 1 (8%) from another ethnicity. All participants identified as religious believers. Just more than half of the participants (11/13, 85%) received a polytechnic-level education or less; 46% (6/13) were currently employed, and 46% (6/13) had an income of more than SGD 3000 (approximately US $2243). Overall, 62% (8/13) of the participants underwent mastectomy, including 3 who received immediate reconstruction. Furthermore, 31% (4/13) of the participants underwent lumpectomy, with one receiving immediate postlumpectomy reconstruction. Approximately half of the participants (6/13, 46%) received adjuvant treatment (chemotherapy or radiotherapy) after surgery during the study period. Other sociodemographic and clinical characteristics of the participants are provided in Table 1.

XSL•FO
RenderX

Table 1. Participants' sociodemographic and clinical characteristics (N=13).

Characteristics	Values
Age (y), mean (SD; range)	58.5 (11.8; 35-73)
Age group (y), n (%)	
21-49	4 (31)
50-64	4 (31)
≥65	5 (38)
Hospital stay length (days), mean (range)	2.85 (1-9)[a]
Marital status, n (%)	
Married	8 (62)
Single	4 (31)
Divorced or separated	1 (8)
Ethnicity, n (%)	
Chinese	11 (85)
Malay	1 (8)
Others	1 (8)
Religion, n (%)	
Buddhism	7 (54)
Christianity	5 (38)
Islam	1 (8)
Education, n (%)	
No formal education	1 (8)
Secondary school	7 (54)
Polytechnic or college	3 (23)
University	2 (15)
Employment, n (%)	
Employed	6 (46)
Unemployed	3 (23)
Retired	3 (23)
Others	1 (8)
Monthly household income per capita (SGD $), n (%)	
≤1000	4 (31)
1000-3000	3 (23)
3000-5000	3 (23)
>5000	3 (23)
Type of surgery, n (%)	
Mastectomy	5 (38)
Mastectomy with reconstruction	3 (23)
Lumpectomy	4 (31)
Lumpectomy with reconstruction	1 (8)
Chemotherapy, n (%)	
Yes	6 (46)
No	7 (54)
Radiotherapy, n (%)	

XSL•FO
RenderX

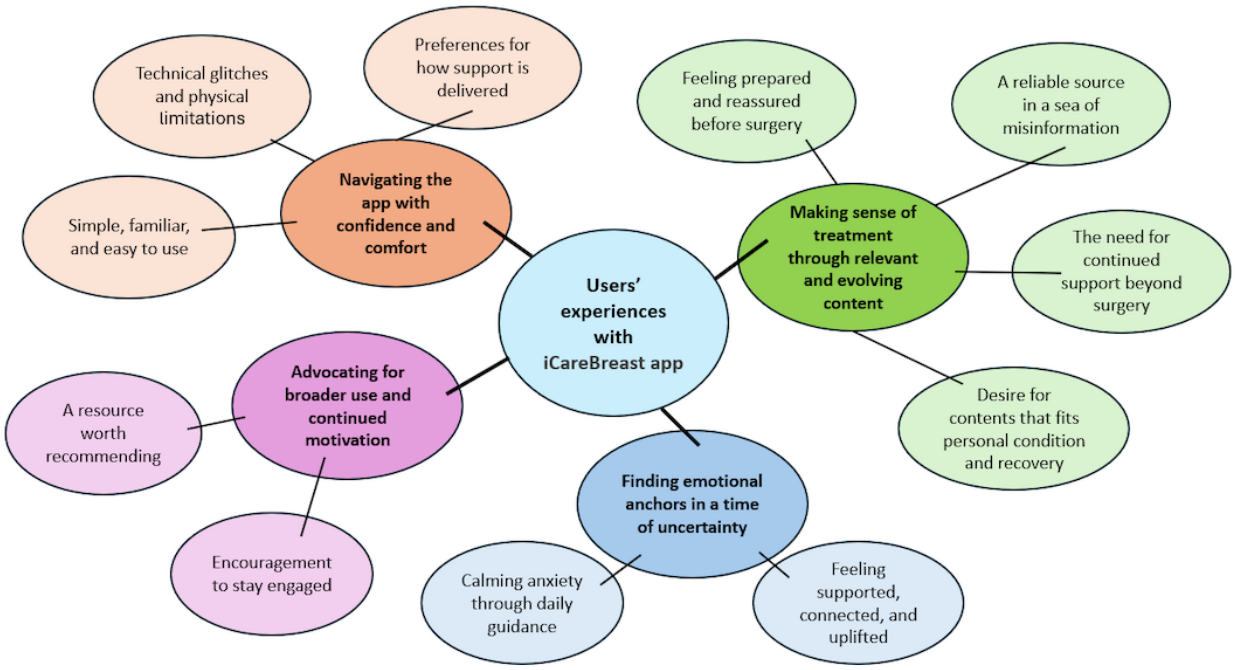
Characteristics	Values
Yes	5 (38)
No	8 (62)

[a]Extended coordination of postdischarge care due to no immediate home support.

Overview

Overall, 4 main themes and 11 subthemes emerged from the thematic analysis: navigating the app with confidence and comfort, making sense of treatment through relevant and evolving content, finding emotional anchors in a time of uncertainty, and advocating for broader use and continued motivation. A summary of the themes and subthemes is provided in Figure 1.

Figure 1. User experience thematic map for the iCareBreast app, showing 4 main themes and 11 subthemes.



Theme 1: Navigating the App With Confidence and Comfort

Subtheme 1.1: Simple, Familiar, and Easy to Use

Participants generally found the app easy to use and navigate, with a clean layout that suited users across ages, including older adults. While the content was released daily, the app's flexibility allowed them to preview upcoming content and catch up on any information from previous days that they may have missed. In addition, participants appreciated the daily app reminders, which helped engage them in completing their daily tasks consistently:

> *Simple, convenient. Also, there was no need to enter this, enter that, switch here and there.* [P5]

> *Yeah I appreciate the ah availability of the information before the actual day has passed.* [P1]

Subtheme 1.2: Technical Glitches and Physical Limitations

While most had a smooth experience of using the app, several users encountered login problems or faced issues such as the app hanging. However, these issues were quickly resolved by our technical support team in Finland, who developed the app and provided technical assistance throughout the study period. A few participants also noted that they were physically too unwell to use the app on the day of surgery and during the initial days after surgery, as they needed to remain in bed and focus on recovery:

> *That day there are some login technical issues. But then your colleague is setting it up. But then after that, this is no more issue.* [P12]

Subtheme 1.3: Preferences for How Support Is Delivered

Some participants missed out on using the in-app messaging feature due to its location or unfamiliarity, suggesting that making it more visually prominent would help them easily ask questions and receive support without needing to contact hospital staff. Others, particularly older users, preferred direct phone calls for urgent concerns. They noted that this feature is more appropriate for nonurgent issues. Many felt that shorter video-based instructions, particularly for physiotherapy or muscle relaxation exercises, would be easier to follow and maintain interest than written or static images:

> *I forgot to use, because it (messaging feature) was not part of the main interface, I think the design of that part was build it within the setting.* [P4]

> *I don't know if it itself as a reminder to the patients. But my therapy is simple so I will remember after a*

XSL•FO
RenderX

> *few times you see, that I don't need to see every, but if I keep seeing every day I will like I feel like boring...a video format will be good, because we can follow like an instructor that we can follow.* [P12]

Theme 2: Making Sense of Treatment Through Relevant and Evolving Content

Subtheme 2.1: Feeling Prepared and Reassured Before Surgery

Participants found the surgical guidance within the app clear and helpful, particularly as it walked newly diagnosed patients with breast cancer through the process step by step. Knowing what to expect for breast cancer and surgical procedures reduced uncertainty and helped them mentally prepared for surgery. The app's detailed knowledge on postoperative expectations and postsurgery care such as pain management, wound care, and lymphedema prevention was particularly valuable:

> *I feel self-prepared (for surgery). Uh, I know what's going to happen. what will happen. How am I going to accept it and take care of it.* [P9]

Subtheme 2.2: A Reliable Source in a Sea of Misinformation

Many participants felt that the content provided in the app was trustworthy and Singapore-relevant source of information. Many appreciated not having to seek additional resources, such as Google searches, which they found overwhelming or sometimes confusing. The volume and complexity of information shared during clinical visits often left them feeling overwhelmed and more likely to forget important details. Many participants appreciated that they could refer back to the app whenever needed, finding it a helpful resource to reinforce what they had learned during clinic visits:

> *Yeah, in fact this software is specifically for the patient at the hospital in Singapore, I think the content of the app is may more rigorous than my own online searching.* [P1]

> *So someone like me I need to refer to the app to refresh my memory. I'll go back to the apps again and see the exercise (physiotherapy) and all those posted on the app lah.* [P11]

Subtheme 2.3: The Need for Continued Support Beyond Surgery

Several users felt the iCareBreast content ended too soon, particularly as new symptoms or emotional challenges often arose later. They expressed a strong desire to extend the app's content to include the subsequent adjuvant treatment phase, with a particular emphasis on chemotherapy. They suggested adding information on different types of chemotherapy, side effect management, and ongoing recovery support to better prepare patients for this stage. Some participants also expressed interest in educational resources on radiotherapy, similar to the surgery-related guidance, covering topics such as types of radiation, key precautions, and postradiation care:

> *For me, the first 14 days were fine; my wound healed well, no pain, nothing unusual. But after the 14 days (end of app using), the pain started, and there was some discharge. Additional information could be helpful.* [P7]

> *Because like for me, the treatment after the surgery I think at least I have to go 9 more months of treatment so I think it would be helpful if we expand using the application after the surgery.* [P8]

Subtheme 2.4: Desire for Contents That Fits Personal Condition and Recovery

Some participants felt that, while the app provided general information, it lacked in-depth detail information based on their surgery type, treatment plan, or recovery stages. Participants felt that the app should adapt to their progress and treatment stages, with features such as adjusting physiotherapy exercises as they recover to provide more personalized support. They also hope for more guidance on nutrition, exercise, and long-term self-care as they progress in recovery:

> *Some people need chemotherapy and so on, so I think it would be great if it was added, it would be continuous for the patient. Because of this disease its treatment is relatively long…patient will just choose to see the content that relevant will be a long-term help for them.* [P1]

> *Additionally, this kind of information for faster recovery, like diet, I believe will be helpful, because I was trying to search if there's any in the app, I couldn't find much.* [P3]

Theme 3: Finding Emotional Anchors in a Time of Uncertainty

Subtheme 3.1: Calming Anxiety Through Daily Guidance

Many participants reported that the app's day-to-day guidance and recovery tips gave them a clearer sense of what to expect, helping to reduce anxiety and alleviate doubts. The amount of daily information in the app also made them feel less overwhelmed. After practicing the provided deep breathing exercises or mindfulness practices, participants felt less pain and pressure, more relaxed, and experienced improved sleep quality:

> *The daily updates on what would happen each day, like meeting the anesthetist and what they would do. It helped me know what to expect, so I didn't feel panicked.* [P7]

Subtheme 3.2: Feeling Supported, Connected, and Uplifted

The app served as a supportive emotional companion, helping participants feel cared for and less alone. They experienced a sense of connection, encouragement, hope, and optimism when reading stories from breast cancer survivors, realizing that many others had been through similar experiences. However, participants suggested adding more diverse breast cancer survivors' stories to strengthen this emotional support aspect. They also appreciated the peer group resources provided, which allowed them to seek additional support and participate in more activities. In addition, participants reported feeling more focused, hopeful, and motivated to think positively and appreciate the present, inspired by the app's motivational quotes and mindfulness practices. They suggested that adding more

XSL•FO
RenderX

mindfulness practice videos would further enhance these benefits:

> *Yeah, I think it can introduce two people (breast cancer survivors' story) a week, that make you feel there are a lot of people who are recovering... A kind of more encouragement.* [P2]

> *The app helped me a lot. One day, when I was feeling down, especially on Chinese New Year's Day, I was very negative. Then, I saw content in the app that talked about people in Africa who don't have food and children in other places who suffer from diseases. It reminded me to be grateful, and it completely changed my outlook. I stopped being resentful and started communicating better with my family.* [P13]

Theme 4: Advocating for Broader Use and Continued Motivation

Subtheme 4.1: A Resource Worth Recommending

Nearly all participants recommend the app to others in similar situations. They saw it as a valuable support tool that made a difference in both practical and emotional aspects of their recovery:

> *Yes, if others need something like this for the long term, I think I would recommend them to use.* [P10]

Subtheme 4.2: Encouragement to Stay Engaged

Participants expressed a desire for the app to provide ongoing motivational content to help them feel supported. Continued emotional support and inspiration were seen as key to encouraging long-term use of the app throughout their recovery:

> *We have this disease, diagnosis has been extremely scared, it is difficult to sleep. Have this thing (APP) is also a kind of encouragement, good. Isn't it? If the app keeps giving me encouragement, confidence, of course I will continue using. Otherwise, if recurrence happen and the app content would be forgotten for a long time.* [P6]

Discussion

Principal Findings

This qualitative study explored how women with breast cancer experienced and engaged with a mobile perioperative support app during the breast cancer perioperative period. Through thematic analysis of interviews with 13 users, 4 key themes were identified, highlighting how the app supported users in navigating their surgery treatment journey, making sense of medical information, managing emotional challenges, and suggesting its potential beyond surgery.

The findings indicate that participants generally felt confident using the iCareBreast app due to its user-friendly interface and smooth navigation across both iOS and Android platforms. The app delivered daily content tailored to each user's surgery date, which helped manage the information flow during an already emotionally overwhelming period. Participants valued this gradual delivery, noting it prevented the overload often experienced during clinic visits or conversations with others. The app's flexibility was another key strength. Users appreciated being able to catch up on missed information if they enrolled late, and some users found it helpful to preview upcoming content if desired. Daily app reminders were seen as helpful prompts that supported consistent engagement, particularly by older users who also appreciated the app's straightforward design. These features underscore the importance of designing digital tools that accommodate varying levels of digital literacy. A few users reported occasional app hang-ups or encountered login difficulties, particularly during scheduled app updates timed according to Finland's time zone, where the iCareBreast app was developed and maintained, which inadvertently affected users in Singapore. These technical challenges were quickly addressed, ensuring minimal disruption to the participants.

Research consistently highlights that patients with breast cancer experience distress and anxiety during the perioperative phase due to uncertainty about what to expect, lack of health-related information, and concerns about unexpected surgical outcomes [21]. These emotional states can exacerbate feelings of being overwhelmed by the overwhelming volume of information provided during hospital visits [22,23]. Our study findings supported this observation. Many participants reported relying on the iCareBreast app to recall guidance from physicians or physiotherapists. This reliance underscores a communication gap that contributes to heightened distress early in the perioperative journey [21]. Although preoperative education and surgery-related information have been shown to reduce anxiety, improve knowledge, and enhance satisfaction, particularly when delivered in person [23], limited consultation time and the fast-paced nature of clinical care often restrict such opportunities. In this context, digital tools such as mobile apps can play a valuable role in supplementing patient education. Participants in our study noted that the iCareBreast mobile app helped in addressing their doubts. They appreciated its trustworthy, locally relevant content, which they described as more accessible and less overwhelming than online searches, which were often confusing or inconsistent. For many users, the app served as a reassuring and structured source of support during the perioperative phase.

Participants described the app as a source of emotional comfort and companionship during a time of uncertainty. Features such as deep breathing exercises, mindfulness practices, daily motivational quotes, breast cancer survivor stories, and links to peer support groups were seen as helpful in reducing feelings of isolation, enhancing focus, and increasing hope. These findings align with the self-efficacy theory proposed by Bandura [14] and the social support components, which guided the conceptual framework of the app's content. The integration of practical knowledge with psychosocial support appeared to mitigate uncertainty-related anxiety and contributed to a more reassuring perioperative experience [21]. However, several participants expressed a desire for a greater variety of mindfulness videos and a wider selection of survivor stories. They felt that expanding these elements would offer additional encouragement, strengthen their sense of hope, and help them stay focused on the present. At the time of the study, the app included 3 survivor stories and 5 short mindfulness practices, which participants found limited. Enhancing this content may

XSL•FO
RenderX

deepen the emotional support provided and increase the app's impact on psychological well-being during the perioperative period [24].

At the same time, our findings point to areas for future development. Participants expressed a strong desire for the app's content and motivational support to extend beyond the perioperative period, encompassing the adjuvant treatment phase, including chemotherapy and radiotherapy. They also highlighted the value of content addressing long-term needs, such as practical advice on nutrition and strategies for maintaining overall health. Many participants noted that their informational and emotional needs evolved significantly approximately 2 weeks after surgery, when discussions about further treatments typically began during follow-up consultations. These findings underscore the importance of providing continuous education and psychosocial support across the entire cancer treatment trajectory [25]. Expanding the app to support active treatment phases could enhance patient engagement, improve preparedness, and help reduce the anxiety associated with uncertainty. To achieve this, the content should align closely with the patient's treatment progress and individual care plans. In addition, HCPs play a critical role in ensuring the ongoing and timely delivery of support that reflects each patient's evolving needs.

In addition to extended support, participants expressed the need for more detailed and personalized information tailored to their specific type of surgery, treatment plans, and recovery process. They noted that personalized physiotherapy exercise and in-depth explanations of treatment options would enhance the app's relevance to their individual experiences. This need for personalized information and care has also been emphasized in previous studies, which report that the complexity of breast cancer treatments, including various combinations of surgery, chemotherapy, radiotherapy, and hormonal therapies, creates a significant demand for individualized support [26]. To address these needs, future app iterations could incorporate features such as an initial intake assessment to identify individual treatment types, algorithms designed to tailor recommendations based on individual treatment plans, and interactive tools for progress monitoring [27-29]. These enhancements would enable the app to remain adaptable and relevant, providing patients with timely and personalized support as patients navigate the complexities of their breast cancer treatment.

In addition, participants expressed a preference for an interactive and engaging format. While the current app includes a messaging feature that allows participants to post nonurgent concerns related to breast cancer surgery, it was underused. Several participants mentioned that this feature was hidden within the settings, making them either unaware of its availability or prone to forgetting to use it. Others acknowledged its usefulness as an alternative means of communicating with HCPs but noted that its use was limited to nonurgent matters, and some preferred speaking with HCPs directly through a phone call. Importantly, many participants favored video content, particularly for deep breathing and physiotherapy exercises, because they found visual demonstrations easier to follow and more confidence building than text or image-based instructions. These findings highlight the need for mHealth interventions to prioritize interactive features to support long-term patient engagement. Future research delivering interventions via mHealth platforms should explore the integration of additional interactive components, such as enhanced video-based content, user-friendly messaging systems, and real-time feedback mechanisms. By aligning with patient preferences, such features have the potential to engage users more actively with digital health tools [30]. Aligning app design with patient preferences and use behaviors could increase engagement, enhance satisfaction with care, and contribute to improved health outcomes over time [13,27].

Recommendations for Future Clinical Practices and Research

This study highlights the significant emotional burden faced by patients with breast cancer in the perioperative period. Many participants experienced heightened anxiety and distress due to the uncertainty of surgery and its potential outcomes. During this time, patients often felt overwhelmed by the volume and complexity of information provided by HCPs. The challenges are compounded by the short window between surgical decision-making and the procedure itself-often just a few days, depending on hospital scheduling and clinical urgency. This limited time frame can make it difficult for patients to process their diagnosis, seek clarification, or access psychosocial support, particularly as preoperative consultations are typically brief and infrequent. In light of these findings, HCPs, particularly BCNs, play a proactive role in regularly assessing and addressing patients' informational, emotional, and psychosocial needs, particularly during the preoperative period. Early introduction of supportive resources, including digital tools, can help patients better prepare for surgery and reduce feelings of uncertainty. After surgery, as patients often transition quickly into adjuvant treatment, ongoing support remains crucial. BCNs play a vital role in maintaining ongoing communication and providing emotional support to help patients cope with the evolving demands of treatment and recovery.

Future research should focus on qualitative studies to explore patients' experiences and support needs throughout the continuum of care, from the waiting period for diagnosis, the period of diagnosis to surgery, and the transition into postsurgical and adjuvant treatment. Such studies could provide valuable insights into how mHealth interventions can be optimized to benefit more patients by extending their use period and tailoring support to address patients' evolving needs at each stage of their treatment journey. Further development of mHealth solutions should prioritize personalization, such as tailoring content based on individual treatment plans and recovery stages, to improve relevance and user engagement. Adding interactive elements such as real-time messaging and video demonstrations can make the app more engaging and easier to use. By designing features based on patients' preferences, future versions of the app could play a stronger role in supporting better health outcomes and making oncological care feel more personalized and responsive.

Study Strengths and Limitations

A key strength of this study was its use of mHealth technology to support patients with breast cancer during the perioperative

XSL•FO
RenderX

period, a phase often marked by high anxiety and limited access to psychosocial support. The iCareBreast app provided timely, accessible, and structured guidance, which was particularly valuable during the COVID-19 pandemic when traditional face-to-face support was not feasible. This study is one of the few that specifically explores user experiences with mHealth, specifically during the perioperative phase of breast cancer care—a phase with limited research focus and unmet patient needs. It offers valuable practical recommendations for future psychosocial or educational interventions using mHealth and provides insights into enhancing patient engagement with mHealth solutions in clinical practice.

However, this study has several limitations. First, participants were restricted to those who were technologically proficient and willing to provide feedback, potentially limiting the generalizability of the findings to the broader population of patients with breast cancer. In addition, the perspectives primarily reflect those of older adults, with a mean participant age of approximately 58.5 years, which may not fully capture the age-specific perspectives of younger patients with breast cancer. Furthermore, as with other qualitative studies, reliance on self-reported data may introduce recall bias.

Conclusions

In conclusion, this study provides valuable insights into participants' experiences using a mobile perioperative support app and highlights the evolving nature of their informational, emotional, and psychosocial needs throughout the early treatment journey. Participants found the app to be a helpful and accessible resource that enhanced their preparedness for surgery, reduced anxiety, and offered emotional reassurance. However, their feedback also revealed a strong need for extended, personalized, and interactive support beyond the immediate postoperative period. These findings underscore the importance of designing patient-centered mHealth interventions that align with the dynamic realities of cancer care. Future iterations should prioritize content continuity, personalization, and engagement strategies to ensure long-term engagement across the treatment continuum. Such enhancements hold promise for improving patients' experience, emotional well-being, and long-term outcomes in breast cancer care.

Acknowledgments

The authors would like to thank all the participants for their valuable contributions to this study. The authors would also like to thank the administrative staff and nursing team at KK Women's and Children's Hospital, Singapore, for their logistical support. The authors are grateful to Buddy Healthcare Ltd Oy for their role in developing the iCareBreast app and providing technical support throughout the study. This research work was funded by the Business Finland (grant 203/31/2018).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

YP was primarily responsible for data collection and drafting the manuscript. HH contributed to the study conceptualization, supervised the study as the main supervisor, provided administrative and material support, guided the intervention development, and participated in both drafting and critically reviewing the manuscript. MP contributed to the study as a cosupervisor, providing guidance on intervention development and supporting manuscript preparation and funding acquisition. SHL served as cosupervisor, overseeing the study, obtaining ethics approval, contributing to the intervention development, and assisting with manuscript preparation. All authors have contributed significantly and agree with the content of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide for participants in the intervention group.
[DOCX File , 18 KB - humanfactors_v12i1e71686_app1.docx]

Multimedia Appendix 2
COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist.
[PDF File (Adobe PDF File), 373 KB - humanfactors_v12i1e71686_app2.pdf]

References

1. Breast cancer. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/breast-cancer [accessed 2024-11-23]
2. Singapore cancer registry annual report 2022. National Registry of Diseases Office. 2024. URL: https://nrdo.gov.sg/docs/librariesprovider3/default-document-library/scr-ar-2022_web-report.pdf?sfvrsn=a9eb8c10_1 [accessed 2024-11-23]

XSL•FO
RenderX

3. Brady L. My journey with triple negative breast cancer. Plast Surg Nurs 2015;35(3):137-144. [doi: 10.1097/PSN.0000000000000107] [Medline: 26313679]
4. Wagner JF, Lüdders D, Hoellen F, Rody A, Banz-Jansen C. Treatment anxiety in breast cancer patients. Arch Gynecol Obstet 2019 May 22;299(5):1365-1371. [doi: 10.1007/s00404-018-05038-z] [Medline: 30671699]
5. Kyranou M, Puntillo K, Dunn LB, Aouizerat BE, Paul SM, Cooper BA, et al. Predictors of initial levels and trajectories of anxiety in women before and for 6 months after breast cancer surgery. Cancer Nurs 2014;37(6):406-417 [FREE Full text] [doi: 10.1097/NCC.0000000000000131] [Medline: 24633334]
6. Tang WZ, Mangantig E, Iskandar YH, Cheng SL, Yusuf A, Jia K. Prevalence and associated factors of psychological distress among patients with breast cancer: a systematic review and meta-analysis. BMJ Open 2024 Sep 26;14(9):e077067 [FREE Full text] [doi: 10.1136/bmjopen-2023-077067] [Medline: 39327054]
7. Leo CS, Cheng LJ, Lam XR, He H. Global prevalence and factors associated with preoperative depression in women undergoing breast surgery: a meta-analysis and meta-regression. Breast Cancer Res Treat 2024 Jun;205(3):425-438. [doi: 10.1007/s10549-023-07216-3] [Medline: 38492162]
8. Lu H, Xie J, Gerido LH, Cheng Y, Chen Y, Sun L. Information needs of breast cancer patients: theory-generating meta-synthesis. J Med Internet Res 2020 Jul 28;22(7):e17907 [FREE Full text] [doi: 10.2196/17907] [Medline: 32720899]
9. Taurisano P, Abbatantuono C, Verri V, Pepe I, Stucci LS, Taurino A, et al. Pre-surgery supportive and goal-oriented strategies are associated with lower post-surgery perceived distress in women diagnosed with breast cancer. BMC Psychol 2022 Jan 03;10(1):2 [FREE Full text] [doi: 10.1186/s40359-021-00714-3] [Medline: 34980291]
10. Jassim GA, Doherty S, Whitford DL, Khashan AS. Psychological interventions for women with non-metastatic breast cancer. Cochrane Database Syst Rev 2023 Jan 11;1(1):CD008729 [FREE Full text] [doi: 10.1002/14651858.CD008729.pub3] [Medline: 36628983]
11. Matthews H, Grunfeld EA, Turner A. The efficacy of interventions to improve psychosocial outcomes following surgical treatment for breast cancer: a systematic review and meta-analysis. Psychooncology 2017 May 02;26(5):593-607. [doi: 10.1002/pon.4199] [Medline: 27333194]
12. Jongerius C, Russo S, Mazzocco K, Pravettoni G. Research-tested mobile apps for breast cancer care: systematic review. JMIR Mhealth Uhealth 2019 Feb 11;7(2):e10930 [FREE Full text] [doi: 10.2196/10930] [Medline: 30741644]
13. Vaffis S, Whaley S, Axon DR, Hall-Lipsy E, Hincapie A, Slack M, et al. Features of cancer mHealth apps and evidence for patient preferences: scoping literature review. JMIR Cancer 2023 Apr 28;9:e37330 [FREE Full text] [doi: 10.2196/37330] [Medline: 37115587]
14. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977 Mar;84(2):191-215. [doi: 10.1037//0033-295x.84.2.191] [Medline: 847061]
15. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. BMJ 2015 Mar 19;350:h1258 [FREE Full text] [doi: 10.1136/bmj.h1258] [Medline: 25791983]
16. Pang Y, He H, Ng RP, Lee NK, Htein MM, Zhao XX, et al. Effectiveness of an innovative mobile-based perioperative care program for women undergoing breast cancer surgery (iCareBreast): randomized controlled trial. J Med Internet Res 2025 Apr 21;27:e71684 [FREE Full text] [doi: 10.2196/71684] [Medline: 40258266]
17. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
18. Boyatzis RE. Transforming Qualitative Information: Thematic Analysis and Code Development. Thousand Oaks, CA: SAGE Publications; 1998.
19. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis: striving to meet the trustworthiness criteria. Int J Qual Methods 2017 Oct 02;16(1):1609406917733847. [doi: 10.1177/1609406917733847]
20. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec 16;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
21. Grocott B, Reynolds K, Logan G, Hebbard P, El-Gabalawy R. Breast cancer patient experiences of perioperative distress and anxiety: a qualitative study. Eur J Oncol Nurs 2023 Apr;63:102299. [doi: 10.1016/j.ejon.2023.102299] [Medline: 36893576]
22. Schmidt A, Kowalski C, Pfaff H, Wesselmann S, Wirtz M, Ernstmann N. The influence of health literacy on information needs among women newly diagnosed with breast cancer, with special reference to employment status. J Health Commun 2015 Jun 18;20(10):1177-1184. [doi: 10.1080/10810730.2015.1018626] [Medline: 26086892]
23. Waller A, Forshaw K, Bryant J, Carey M, Boyes A, Sanson-Fisher R. Preparatory education for cancer patients undergoing surgery: a systematic review of volume and quality of research output over time. Patient Educ Couns 2015 May 23:S0738-3991(15)00229-3. [doi: 10.1016/j.pec.2015.05.008] [Medline: 26072423]
24. Haller H, Winkler MM, Klose P, Dobos G, Kümmel S, Cramer H. Mindfulness-based interventions for women with breast cancer: an updated systematic review and meta-analysis. Acta Oncologica 2017 Jul 07;56(12):1665-1676. [doi: 10.1080/0284186x.2017.1342862]

XSL•FO
RenderX

25. Smit A, Coetzee BJ, Roomaney R, Bradshaw M, Swartz L. Women's stories of living with breast cancer: a systematic review and meta-synthesis of qualitative evidence. Soc Sci Med 2019 Feb;222:231-245. [doi: 10.1016/j.socscimed.2019.01.020] [Medline: 30665063]
26. Cardoso F, Kyriakides S, Ohno S, Penault-Llorca F, Poortmans P, Rubio IT, et al. Early breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2019 Aug 01;30(8):1194-1220 [FREE Full text] [doi: 10.1093/annonc/mdz173] [Medline: 31161190]
27. Tuominen L, Leino-Kilpi H, Poraharju J, Cabutto D, Carrion C, Lehtiö L, et al. Interactive digital tools to support empowerment of people with cancer: a systematic literature review. Support Care Cancer 2024 May 31;32(6):396 [FREE Full text] [doi: 10.1007/s00520-024-08545-9] [Medline: 38816629]
28. Kirsch EP, Kunte SA, Wu KA, Kaplan S, Hwang ES, Plichta JK, et al. Digital health platforms for breast cancer care: a scoping review. J Clin Med 2024 Mar 27;13(7):1937 [FREE Full text] [doi: 10.3390/jcm13071937] [Medline: 38610702]
29. Holtdirk F, Zindler T, Mehnert A, Bültmann O, Weiss M, Mayer J, et al. Digital health applications to support patients with breast cancer: effects of two tailored, dialogue-based programs on quality of life. Eur J Cancer 2022 Nov;175:S56-S57. [doi: 10.1016/s0959-8049(22)01501-5]
30. Madanian S, Nakarada-Kordic I, Reay S, Chetty T. Patients' perspectives on digital health tools. PEC Innov 2023 Dec;2:100171 [FREE Full text] [doi: 10.1016/j.pecinn.2023.100171] [Medline: 37384154]

Abbreviations

BCN: breast care nurse
COREQ: Consolidated Criteria for Reporting Qualitative Research
HCP: health care professional
mHealth: mobile health

Edited by A Kushniruk; submitted 24.01.25; peer-reviewed by M Knobf, J Erdrich; comments to author 03.07.25; revised version received 07.08.25; accepted 12.08.25; published 10.09.25.

Please cite as:
Pang Y, He H, Pikkarainen M, Lim SH
Participants' Perspectives on the iCareBreast Mobile-Based Perioperative Care Program for Women Undergoing Breast Cancer Surgery: Qualitative Process Evaluation
JMIR Hum Factors 2025;12:e71686
URL: https://humanfactors.jmir.org/2025/1/e71686
doi:10.2196/71686
PMID:40928835

©Yan Pang, Honggu He, Minna Pikkarainen, Swee-Ho Lim. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 10.09.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Patient's Perceptions of a Centralized Virtual Ward for Remote Patient Monitoring in Primary Care: Qualitative Study

Alex Jaranka[1*], MD; Gunnar H Nilsson[1*], MD, PhD; Terese Stenfors[2*], Prof, PhD; Maria Hägglund[3*], Prof, PhD; Panagiotis Papachristou[1*], MD, PhD; Marina Taloyan[1*], Docent, PhD

[1]Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Alfred Nobels Allé 23, Huddinge, Sweden

[2]Department of Learning, Informatics, Management and Ethics, Karolinska Institutet, Stockholm, Sweden

[3]Department of Women's and Children's Health; Participatory eHealth and Health Data Research Group, Uppsala University, Uppsala, Sweden

[*]all authors contributed equally

Corresponding Author:
Alex Jaranka, MD
Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Alfred Nobels Allé 23, Huddinge, Sweden

Abstract

Background: Remote patient monitoring (RPM) has the potential to reduce in-clinic visits and promote proactive and preventive care for patients with chronic diseases in primary care. However, a decentralized approach to RPM in a primary health care (PHC) setting has not met stakeholders' expectations regarding scalability. This study introduces a centralized virtual ward (CVW)–led RPM, utilizing a multidisciplinary team approach to monitor patients with chronic diseases by clinicians who do not belong to the patients' PHC center.

Objective: This study aimed to gain a better understanding of patients' perceptions of CVW-led RPM for managing chronic diseases in a PHC setting.

Methods: In-depth interviews were conducted with 22 patients with chronic diseases enrolled at a PHC center in Stockholm, Sweden. The RPM project ran between October 2018 and April 2019 and included a total of 395 patients. Interviews followed a semistructured interview guide and were analyzed using qualitative content analysis.

Results: Primary care patients with chronic diseases expressed that their contact with the CVW felt impersonal but at the same time secure and accessible. They noted a lack of coordination and communication between the clinicians of the CVW and their PHC providers. Captured data resulted in 1 overarching theme "Sense of security and accessibility, but impersonal and uncoordinated" based on 5 categories: sense of security, care and self-care, accessibility, quality of care, and communication.

Conclusions: Our findings suggest that by addressing patients' needs for new organizational routines for patient-caregiver communication, RPM via centralized virtual wards can better realize the potential of this technology.

(JMIR Hum Factors 2025;12:e78780) doi:10.2196/78780

KEYWORDS

chronic disease; patients perception; primary care; qualitative study; remote patient monitoring; remote sensing technology; telemedicine

Introduction

Background

The future of health care needs a shift from reactive to proactive and preventive care to reduce the rising global burden of diseases [1]. The World Health Organization defines primary health care (PHC) as "a key process in the health system that supports first-contact, accessible, continued, comprehensive, and coordinated patient-focused care" [2]. A fundamental philosophy of PHC has been to deliver continuity of care to achieve positive patient experiences, greater patient satisfaction, increased treatment adherence, and improved patient outcomes [3].

Globally, PHC is facing numerous challenges, including rising costs, increased public demand for convenience, and technological advances driving personalized health care [4]. New models of health care delivery, including centralized coordination among providers, have been suggested to improve care quality and reduce hospitalizations [5]. Novel digital solutions may enable new forms of interventions and activities that address these challenges [6].

Remote patient monitoring (RPM) is a technology that allows health care providers to monitor, evaluate, and manage patient-reported health data through a remote interface. The collected data are transmitted to the health care provider for clinical review, care management, and patient education [7,8]. RPM enables clinical interventions performed by medical staff

XSL•FO
RenderX

who monitor health data, communicate with the patient, and make decisions as a part of the treatment [9].

A plethora of studies indicate that patients with chronic diseases perceive RPM as improving their access to care [10], sense of security [11,12], engagement [13], and disease-specific knowledge [14]. RPM also supports better clinical assessment [15,16], promotes better relationship-based continuity [17], and improves self-management [18]. However, recent studies have identified a lack of competence regarding the use of RPM in clinical care [19]. Also, given the resource constraints and time pressures in PHC, it can be challenging to introduce new tasks such as continuous monitoring of health data.

Virtual wards were first introduced in PHC in 2005, targeting individuals with a very high risk of future emergency hospitalization [20]. The intention was to use multidisciplinary teams and the daily routines of a hospital ward to provide health care in the community [21,22]. A community virtual ward project targeting older adults with complex health care and social care needs resulted in a reduction of emergency department visits and hospital admissions [23]. A centralized health information strategy in the PHC sector has fostered data-driven decision-making, improving the transparency of workflow and preventative care [24].

Combining the virtual ward concept with RPM could potentially offer a more feasible approach to implementing RPM in PHC. The prediction of prior work around CVW and RPM highlights the importance of care coordination and leverages the full range of health care skills within the system to promote patient-centered care [25]. Consequently, it encourages greater patient autonomy and self-management and provides a sense of reassurance [26]. The introduction of CVW models for monitoring patients with COVID-19 during the pandemic may have transformed patients' experiences of care, making the measurement of patient experience increasingly vital for building a strong, person-centered health care system [27].

Problem Statement

Despite growing interest among stakeholders, the implementation of RPM in PHC remains limited [28]. Patients support sharing their health data but are concerned about data privacy, security, and management [29]. Clinicians see the benefit of having access to more reported health data but are concerned about increased workload due to the volume of transferred health data [30]. Clinicians also worry that the modernization of health care might lead to a decline in continuity of care [31,32] and reduced patient safety [33]. A major concern among PHC professionals regarding the use of RPM is increased workload due to the transfer of a large amount of data and an increased number of requests from caretakers [34-36]. Health care managers have seen telemonitoring as a means for cost reduction. However, due to increased workload and the cost of technology, managers believed that service centralization would reduce the level of clinician involvement in frontline, resulting in cost savings [32].

A key distinction of the centralized virtual ward (CVW)–led RPM model compared to the traditional in-clinic RPM is that the CVW multidisciplinary care team solely focuses on digital patient interactions based on patient vitals transferred via the RPM platform. The CVW-led RPM team can therefore continuously monitor incoming patient data and interact with the patients without interruptions related to in-clinic workload. This allows the CVW-led RPM model to conduct the monitoring of chronic patients from several clinics, which could therefore address patient and clinician concerns and make the implementation of RPM in PHC more feasible according to health managers' consideration.

The novelty of this study is the examination of the patients' perception of CVW-led RPM in a PHC setting. To our knowledge, there are no previous publications on this topic.

Objective

The objective of this study was to explore patients' perceptions of CVW-led RPM for managing their chronic diseases in a PHC setting.

Methods

Study Overview

In this qualitative study, semistructured in-depth interviews were conducted with participants of a digital health intervention, in a Swedish PHC setting. This study is reported in accordance with the Consolidated Criteria for Reporting Qualitative Research checklist [37].

Centralized Virtual Ward

A private Swedish health care provider, with multiple PHC practices across the country, collaborated with researchers from Karolinska Institutet to establish a CVW-led RPM project at a PHC center in Stockholm. This project was initiated to evaluate the effects of a centralized model of RPM compared to the health care providers' previous experience of piloting an on-site RPM project, where patients sent their vitals to and communicated with their ordinary clinicians at the PHC.

To establish the CVW-led RPM project, an RPM platform provider was contracted by the health care provider. A specialist doctor in family medicine, 2 specialist district nurses, and an information technology assistant worked part-time at the CVW. All clinical staff at the collaborating PHC were informed about the project. In addition, 2 doctors and 3 nurses from the PHC received training and were given access to the RPM platform for monitoring and communicating with patients during the entire project period.

After the project was launched, the CVW team monitored patient data to detect early signs of deterioration. Elevated vital signs resulted in asynchronous patient contact initiated by the CVW via the platform. Based on the transmitted patient data, medical staff at the CVW were able to personalize treatment plans and make medication adjustments digitally. When signs of deterioration indicated that the participants needed a physical examination, the staff at the patients' PHC center were alerted by the CVW.

The CVW clinicians received training in digital communication with patients but did not conduct any physical examinations

XSL•FO
RenderX

during the intervention. Further details of the intervention are available in previously published papers [38,39].

Remote Patient Monitoring Project

The PHC staff created a list of eligible participants for the project enlisted at the PHC by using a database extraction tool (MedRave) [40], connected to its electronic health records. A flyer with written information about the intervention was mailed to the candidates. The flyer contained an email address and instructions on how to express interest in participating in the project. The information about the project was posted at the PHC center. PHC staff gave information to eligible candidates about the intervention during physical visits. CVW staff phoned the candidates who showed interest in participating in the intervention and gave further information about the project and assessed eligibility.

The inclusion criteria to the CVW-led RPM project were age above 18 years, being fluent in the Swedish language, having access to a smart device with the iOS operating system and internet access, and having one or more of the following chronic conditions: hypertension, chronic heart failure, type 2 diabetes, chronic obstructive pulmonary disease, anxiety, and mild depression. Pregnant women and patients with other psychiatric conditions were excluded. In total, 395 participants were enrolled during the first 3 months of the project. At the time of enrollment, 153 participants gave their written consent to be contacted about further research. The project started in October 2018 and ended after 6 months at the end of March 2019.

All the participants were scheduled for introductory group information sessions at the PHC (6-10 participants per session). The CVW enrolled the participants into the platform and decided on a monitoring plan for each participant based on their chronic disease(s). An information technology assistant from the RPM platform provider helped the participants to download an iOS app into their smartphone device and paired the app to the selected noninvasive sensor(s) via Bluetooth. The participants received verbal and written instructions on how to share their vitals (blood pressure, weight, sleep rhythm, step count), track their vitals, fill out forms, receive standardized information, and communicate asynchronously by chat or synchronously via prebooked video appointments with the CVW.

The participants were able to transmit health data or respond to questionnaires without restrictions in time or place. Vitals and answers to questionnaires were transferred via the participants' apps to a secure cloud server.

Recruitment of Study Participants

All 153 RPM project participants who consented to be contacted were invited by the last author (MT) via email (including 1 reminder) to receive information about this qualitative study. A total of 21 (13 female) patients responded to the invitation and accepted to be interviewed. The number of participants at each process stage for the recruitment to the in-depth interviews is presented in Table 1.

Table . Process stage for in-depth interviews.

Process stage	Number of participants
In the RPM[a] intervention	395
Agreed to be contacted for qualitative study information	153
Agreed to be interviewed	22
Completed the interview	22

[a]RPM: remote patient monitoring.

Data Collection

Due to the COVID-19 pandemic, the in-depth interviews were conducted digitally via the Zoom communication platform [41] and audio-recorded after the participants' verbal consent. Only the interviewer and the participant were present during the individual interview.

Interview Guide

A semistructured interview guide was designed to attain an in-depth understanding of participants' perception about the interaction with the CVW team. A pilot interview with 1 person not included in this study was performed by the first author (AJ, PhD student, M.D, specialist in family medicine, male). The final interview guide was approved by all the authors (Multimedia Appendix 1).

Between March and May 2021, 22 in-depth interviews were conducted. Each interview ranged between 45 and 90 minutes. An external research assistant (female) with previous experience in qualitative methods, without any affiliation to the participants, conducted the interviews after initial training in the semistructured interview guide and with the guidance of the last author of this study (MT, associate professor in family medicine; female).

Data Analysis

All interviews were recorded and transcribed verbatim. Data analysis was conducted in Swedish. The principles of qualitative content analysis were conducted according to Graneheim and Lundman [42]. A basic inductive content analysis process was applied [43]. The first and last authors selected meaningful phrases and sentences independently. The second step was to condense them to meaning units, and the third step was to set open codes on the manifest level to avoid interpretation and abstraction at the coding stage. The condensing of meaning units into codes was revised during physical meetings in discussions between the first and second authors, and after agreement, a new clustering of codes was conducted if needed. This ensured a common understanding and a reduction in the number of codes before aggregating these into subcategories

XSL•FO
RenderX

and categories. The strength of this method is that during the categorizations, the authors are going back to the original statements several times. After the coding and categorization described above, all the authors discussed the results and developed the overall theme (Textbox 1).

Textbox 1. Examples of meaning units, condensed meaning units, and codes.

Quote

"I take my blood pressure a number of times a week, then I get someone who reviews these values and thinks whether it's good or bad simply. So far, all the values have been satisfactory all the time, so it's that you feel calm, and you know that someone is checking it."

Meaning unit

I take my blood pressure and feel reassured that someone is checking and thinking whether it is good or bad.

Condensed meaning unit

I feel reassured.

Code

Someone is keeping track and is watching.

Ethical Considerations

This study was approved by the Regional Ethical Review board in Stockholm, with 2 amendments (Dnr 2018-625-31/5; 2018/1717 - 32; 2020 - 01890). An email with study information and a consent form was sent to the participants prior to the interview. All the participants gave written informed consent before they were scheduled for an in-depth interview, and oral informed consent was recorded verbally at the start of the interview. To ensure participant anonymity, the participants' identification was replaced by a unique identification number. This number was used to match the participants and their transcripts and to report participant quotations in this study. The participants did not receive any compensation for being interviewed.

The first author (AJ) was clinically active at the pilot PHC 1 year prior to the launch of the project. Thus, he did not participate in the data collection to avoid bias.

Results

Demographics

The 22 interviewed participants had a mean age of 61.2 years (SD 11.2; range 40-78 years of age). Most (n=13, 59%) participants were women. They had predominantly 1 medical condition (n=15, 68%), where 19 (86%) of these participants had hypertension. No participant was diagnosed with more than 3 medical conditions. In comparison to the CVW-led RPM project, the 385 participants had a mean age of 58 years, with 62.3% (n=240) being female participants. The participant characteristics of this study are listed in Table 2.

XSL•FO
RenderX

Table . Participant characteristics (n=22) with 1 or more chronic conditions.

Characteristics	Participants, n (%)
Age (y)	
40 - 49	4 (18)
50 - 59	5 (23)
60 - 69	8 (36)
70 - 79	5 (23)
Sex	
Female	13 (59)
Male	9 (41)
Medical condition	
Hypertension	19 (86)
Diabetes	6 (27)
Mental disorder	3 (14)
Congestive heart failure	1 (5)
Chronic obstructive pulmonary disease	1 (5)
Number of medical conditions	
One	15 (68)
Two	5 (23)
Three	2 (9)

Overall Theme

The following theme was constructed: “Sense of security and accessibility, but impersonal and uncoordinated” based on 5 categories: sense of security, care and self-care, accessibility, quality of care, and communication (Textbox 2).

Textbox 2. Subcategories, categories, and themes from content analysis. CVW: centralized virtual ward; PHC: primary health care.

Theme: Sense of security and accessibility, but impersonal and uncoordinated.

- Sense of security
  - Self-control
  - Care relationship between CVW and the patient
- Care and self-care
  - Care utilization at the PHC
  - Health awareness
- Accessibility
  - Care not dependent on space and time
- Quality of care
  - Basis for making health care decisions
  - Proactive care
- Communication
  - Participant communication with the CVW team
  - Lack of coordination between CVW team and other caregivers at the physical center

XSL•FO
RenderX

Sense of Security

A higher number of transferred vitals and correspondence with the CVW at signs of deterioration strengthened the patients' sense of security. The patients experienced that the CVW reacted promptly to deviating vital signs. Seeing that their vitals were stable induced a sense of security, which reduced their worries about a potential major medical event. Prompt actions taken by the CVW at signs of deterioration built further trust. The participants expressed that the CVW cared about them. A recurring argument for sense of security was the interaction with a trusted health care professional who gave clinical feedback:

> *Then you know that there is someone who cares, there is someone who sees my values every day which my regular doctor does not do. Then I feel safer, then there is a reason why I do this every morning.* [Participant 9]

The participants regarded their collected vitals as evidence when they interacted with health care professionals. They used this evidence to present their case when having a physical contact at their PHC. The clinicians' access to collected vitals and medical records gave a historical perspective, which gave the patients an elevated sense of security as illustrated:

> *What is positive is that it feels like there is a history all the time that you can keep track of, what it looks like for me… it feels like there is a history to see in a completely different way, which gives security…* [Participant 12]

RPM with interactions by the CVW sheds light on the importance of feeling safe as expressed by the following participant:

> *Now you start to realize that it is incredibly important that you feel safe, important for your health and the feeling of… well if something happens, yes, then I know.* [Participant 13]

Care and Self-Care

Continuous reporting of vitals gave the participants insights into their underlying condition and the symptom related to their chronic diseases. Discussing deviating vitals with the CVW enhanced their understanding about the causes of deterioration. Health awareness resulted in patients taking steps toward active lifestyle changes. Motivation for tracking vitals increased as it gave better insights about their condition, as 1 participant described:

> *… say that I had gone to the health centre every other month or quarterly and taken my blood pressure there, then I had not thought about the times between each visit that you may have felt bad and that it may have been because the blood pressure had been too high. Now it was very easy when you feel a little strange, this does not feel good, well then, I can take the blood pressure and then I have better understanding of how I feel right then and how it feels to have this high blood pressure.* [Participant 20]

Feedback on the collected vitals and being monitored also motivated some participants to take further steps toward active lifestyle changes, as expressed by the following participant:

> *Yes, but it is with this that I have started to move a lot more, it has really motivated me to do not just walking but also other physical activity like with swimming and cycling and things like that, there have really been changes there.* [Participant 3]

Insights about the correlation between lifestyle choices and the stability of vitals were generated in the dialogue between the participants and the CVW. These insights motivated the participants to change harmful lifestyle behaviors. The participants sensed that the CVW tried to improve their health outcomes by working both reactively and preventively. One participant described an aha feeling:

> *Just that feeling that it was nice that someone has an ambition to make it possible for me as a patient to improve. A little aha-feeling… It felt like someone cared, like "now we'll do good before it gets too bad," so that was positive.* [Participant 1]

Accessibility

The participants expressed that they refrained from contacting the PHC for what they perceive as bagatelle symptoms because they want other patients with greater medical needs to have access to care. This mindset might lead to deteriorating symptoms that might result in unnecessary complications. The RPM platform reduced thresholds for asking questions about their chronic condition, which enhanced the participants' sense of accessibility to the CVW compared to the medical staff at the PHC center as expressed:

> *I contact healthcare when I need to. If I had a question that was important to me, but that I wouldn't call the regular health centre about.…When I can write that easily, maybe I'll reach out to them more often if I have questions or such…. Usually, I don't do that; I go to the doctor when I am really unwell…* [Participant 16]

The participants were able to travel with their noninvasive sensors and contact the CVW to follow up on their vitals. This opened a new dimension of accessibility to health care for some participants who had been restricted from traveling:

> *A benefit is that you are not tied to sitting at home; even when I travel, I have the blood pressure monitor with me, and I can take those readings, which can also be used for monitoring.* [Participant 6]

Quality of Care

The difference between the number of data points collected by traditional health care versus RPM via the CVW raised new perspectives. The participants reflected that the collection of additional data points could result in better decision-making by the health care staff:

> *The staff I have contact with have significantly better information available once I contact them.* [Participant 9]

XSL•FO
RenderX

A higher frequency of data transfer by sharing of vital measurements leads to a better basis for medical decision-making for the CVW according to patients' experiences. They found that continuous control of vital signs allows both patients and clinicians to detect signs of deterioration. This allows clinicians to adjust treatment plans more stringently to stabilize patients' vitals. Medical adjustment made based on only 1 measurement is not as reliable as those made based on a higher frequency of measurements according to the following participants:

> *…it became more quantitative data of course because it felt like the case became more, it became a better basis in the continued care…* [Participant 20]

> *To go to the healthcare centre and do a measurement at some point in the year does not actually say much more than what it is right then. I think it is good to have some continuous measurements, partly it is a security to get information if there is something that seems strange.* [Participant 9]

To be reminded of reporting values was expressed as feeling that "someone taps on the shoulders." This proactiveness of the CVW was more appreciated than the traditional way of contacts from the health care center.

> *…if I did not leave my daily vital measures they contacted me. The healthcare centre does not do this in the same way, but it is this part that I appreciate, that there is someone who taps you on the shoulder and says that now is the time to leave some vital measures.* [Participant 19]

Communication

This category includes the communication between the CVW, the participants, and the PHC staff. A communication plan was established for participant referrals from the CVW to the PHC. The CVW was being seen as a quite separate entity, and some participants mentioned that it was not important where the CVW was located physically. The ease of use of the app, including the CVW response through the asynchronous channel, was expressed as faster, easier, and simpler compared with the communication channels with the PHC center:

> *…if they are in the healthcare centre or if they are sitting elsewhere, I have no idea and I do not care so much either. It does not matter to me.* [Participant 6]

> *Much simpler. Much faster response, much easier…Yes, I have much more communication with the healthcare now than I had before…* [Participant 4]

The participants identified that the clinicians at the PHC did not have access to their collected RPM data. In addition, the general practitioners were not interested in getting the information regarding the collected vitals. The lack of communication and collaboration between the PHC personnel and the CVW was mentioned by some participants:

> *I can't quite understand if this team collaborates, or to what extent they cooperate with the healthcare centre, or if they are updating them on the values that I submit.* [Participant 19]

> *But I don't know if that's the case; it doesn't seem like it when I've wanted to show my values to the doctors. There was a female doctor, and she said, 'That's not for us.'* [Participant 5]

The asynchronous communication, in combination with the standardized automatic replies and notifications sent to the participants, was described as impersonal:

> *So, it feels safe although it is quite impersonal, I would say.* [Participant 22]

> *Yes, it's not the same staff that I have contact with anymore; it could be random people reaching out who I don't really know who they are, so it becomes more anonymous that way…It becomes impersonal that way.* [Participant 11]

The participants noticed that there was a lack of coordination between the CVW and the PHC center. Also, some participants wished that their efforts of vital submissions would result in better transparency and utilization of health data between the patients, the CVW, and the PHC center. There was a sense of disappointment that the PHC staff did not have access to their health data through the platform and that they felt rejected when they tried to show their data to the health care personnel at the PHC center. One participant expressed lack of coordination as follows:

> *I lacked some natural spin-offs such as my general practitioner or others that I met with would be informed and kept track of this. But they seemed unaware that I was connected to this system.* [Participant 1]

Discussion

Principal Findings

The unique contribution of this study lies in exploring patients' perspective of CVW-led RPM in a PHC setting, bringing a new dimension into the field of RPM in comparison to on-site RPM by the PHC staff. A principal finding of this study was that the participants did not mind that the CVW staff did not belong to their ordinary PHC. The qualitative data highlight the participants' appreciation for engaged, accessible, and well-prepared CVW staff rather than an attachment to their PHC clinician. The participants experienced similar benefits of RPM as the traditional model of RPM, such as enhanced sense of security [12], improved accessibility [18], positive effects on self-care [44], better quality of care [45], and the need for improvement in communication between the patients and the staff [17,46]. In addition, this study underscores the challenge of data transfer and communication between the CVW staff and the PHC staff. The main theme of this study represents the main result: *Sense of security and accessibility, but impersonal and uncoordinated*.

Interpretation of Findings

The participants valued the CVW-led RPM in terms of staff medical competence, availability, and prompt feedback in

XSL•FO
RenderX

comparison to their regular clinical contact. This indicates that patients' trust can be earned through virtual interactions, even with unfamiliar health care staff. In this case, the CVW was able to earn the participants' trust by following a set of protocols to monitor incoming data to detect the signs of deterioration and reacted promptly by reaching out to the individual.

Similarly, Walker et al [17] found that RPM participants reported an increased sense of safety when clinicians responded to clinical signs of deterioration. A sense of safety is defined as an internal sense—an intrinsic state based on faith and trust in oneself and others [47]. The capacity for the RPM system to operate safely lies in the possibility of understanding a participant's symptoms and health behaviors. A sense of safety also relates to participants' ability to process health data and awareness that the health care professionals continuously monitored their data [12]. Although a false sense of security was not mentioned by the participants, one could argue that there is a potential risk for a false sense of security as an extended outcome of organizational and workflow challenges that arise during the introduction of new ways of working, as described in the literature [48] .

The automated response in the patient application gave a sense that someone was constantly monitoring their vitals. But the dismissal of RPM data by some individual clinicians at the PHC upon physical visits may have contributed to the participants' experience of uncoordinated care. However, the communication gaps need further investigation.

A higher number of interactions with the CVW with a faster response rate compared to their usual care at the PHC prompted a sense of accessibility to care. Access to asynchronous communication and data visualization by the RPM application gave the participants a new perspective regarding the accessibility and continuity of care. Previous studies have described that improved accessibility increases the chance of getting medical attention [49], resulting in the avoidance of hospitalization [50-52]. Hall et al [10] also describe that patients' access to technology can aid in the development of effective health behavior interventions. The ability to share vitals from the convenience of their residence at any time of the day resulted in the sense that they had better accessibility to the CVW than to their PHC center. By eliminating confinements of time and space, their interactions with the CVW felt more adequate since they were able to ask specific questions about their underlying health conditions. The participants sensed that the answers provided by the CVW were based on facts from the collected health data. This interpretation is consistent with prior findings that patient trust increases when health data are accurate and complete [53]. Improved accessibility also reduced the threshold for help-seeking behavior from the participants. In turn, the effects of continuous interactions between the participants and the CVW in combination with the prompt response rate from the CVW enhanced their sense of continuity of care. The combination of accessibility to care and sense of continuity of care became the driving factors for their sense of security.

Previous studies found that future care models of chronic disease management will shift toward self-care facilitated by digital technology, social support, and clinical expertise [54]. Self-care is based on long-term adaptations of lifestyle behaviors with adherence to preventive or therapeutic regimens with the potential of improving health outcomes [55]. The cornerstone in self-care is collaborative care between the patient and the health care provider and individual education in disease management [56,57]. Several studies indicate that RPM empowers patients by enhancing their motivation to adhere to the treatment plan, understanding of their chronic condition, and ability to follow their vitals [17,58]. The participants of this study described enhanced knowledge about their chronic condition, awareness and understanding of deviating vitals, and adherence to long-term lifestyle changes. The collection of vitals from the comfort of their residence enhanced the sense of measurement accuracy. In line with previous findings [59], we found that the collection of vitals is a major key for allowing patients to actively share their personal health data. By observing the longitudinal trends of their vitals, the participants gained a better understanding of the effects of their lifestyle choices, resulting in deeper insights into their chronic condition. As described by Quinn et al [60], patient-gathered health data allow the patients to better understand their disease process, monitor their symptoms, and feel connected to their health care provider. These insights, with their ability to discuss their observations with the CVW, evolved in revising their health goals, leading to self-knowledge that was used for better self-care.

Stavropoulou et al [61] describe the quality of care as "holistic care, addressing all patient needs with competency and aiming for the best patient outcomes". The participants expressed confidence in the CVW's ability to detect deterioration signs based on health data trends. A common notion among the participants was that a higher frequency of shared vitals resulted in a better basis for medical decision-making, resulting in higher quality of care in comparison to annual checks made at the PHC center. The participants anticipated a reduced risk of deterioration because the CVW reacted to deviated vitals and applied preventive measures for restoring the participants' health to achieve better health outcomes. However, 1 area of consideration regarding longitudinal collection of health data at a high frequency is the ethical aspects concerning boundaries of confidentiality and the clinician's access to vast amounts of data [62]. The participants in this study did not express concerns regarding compliance to data regulatory frameworks for protecting their data privacy.

The inclusion criteria requiring participants to have access to iOS smart devices and sufficient digital literacy have excluded less tech-savvy individuals or those from lower socioeconomic backgrounds. This selection bias potentially limits the generalizability of the findings to broader patient populations, consistent with prior research highlighting digital divides in remote health monitoring uptake [29].

This study identifies that clear organizational routines and communication strategies between patients, CVWs, and their PHC center are crucial for gaining patient confidence in a CVW-led RPM setting. Several participants expressed a sense of uncoordinated health care due to a lack of work routines for sharing information between the CVW and the PHC center. While RPM facilitates continuous data collection and clinical oversight, our findings highlight persistent communication gaps

XSL•FO
RenderX

between the CVW and patients' primary care providers, which may hinder integrated care delivery. Addressing these collaborative barriers is critical to realizing the full potential of CVW-led RPM. Setting clear goal-oriented communication strategy between caregivers and patients can increase the success rate for establishing RPM [63,64]. This finding might also be relevant to setting up a CVW-led RPM within PHC settings. Our participants appreciated the ability to communicate with the CVW without being confined to the opening hours of the PHC center. Although they appreciated the better accessibility to the CVW, they argued that the asynchronous way of communication felt impersonal. This impersonal sentiment can be due to the transactional nature of digital care with fewer nonverbal cues and limited opportunities for nuanced empathy that can hinder traditional elements of small talk, creating a sense of anonymity and the decreased opportunities for emotional and social support in comparison to standard clinical visits. Despite this limitation, enhanced patient agency, transparency, and collaborative goal-setting, all supported by CWV-led RPM, can foster a sense of partnership and engagement, thereby mitigating feelings of isolation.

Previous studies describe that an increased number of communications require new work routines for the health care personnel [58] and that there is a continuous need for improvement in the communication between the patients and the staff [17,46]. Our findings suggest that clear communication strategies and work routines could reduce the perception of impersonal interactions from the participants' perspective.

Despite advancements in RPM technology after the COVID-19 pandemic, there are still barriers for scalability of RPM in PHC settings [19], making the findings of this study relevant to understanding these challenges and potentials of CVW-led RPM.

Strengths and Limitations

To support trustworthiness [42], we described each phase of our study methodology from data collection, analysis of data, and interpretation. The credibility [42] of the results has been validated by internal considerations regarding the reliability and accuracy in every process of this study. To improve the reliability of the study data, clear coding rules were applied so that each researcher could independently interpret the data. The trustworthiness and credibility of the study were also strengthened by providing detailed contextual information, conducting member checks, and performing a pilot study before data collection.

A primary limitation of this study was the lack of patient involvement in the design process. The limitation to data collection correlates to the recruitment of participants to the RPM project where patients who lacked access to an iOS smart device were excluded. This reduced the potential number of candidates with a lower socioeconomic background. An additional limitation of this study is the lack of information about the participants' digital literacy and experience of using iOS smart devices. However, all the participants included in this study had access to iOS smart devices, and they received instructions for how to use the RPM app prior to getting connected to the CVW for sharing their vitals. To ensure dependability [65] on data collection, the participants were required to be fluent in the Swedish language. In addition, a semistructured interview guide was used to enhance dependability. This study did not explore the importance of how the underlying disease impacted the patient's perception of the CVW-led RPM. Such aspects need to be further investigated in further research.

Transferability [65] was assured by describing the study setting, methods, and time frames for the collection of data. Despite the fact that the pool of the participants may not represent all the participants of this RPM project, we recognize that there is a strength in the large sample size contributing to data collection. Since some participants had left the RPM project 18 months prior to being interviewed, recall bias may have been introduced. In addition, the duration of participation in the project varied among the participants, which might have affected their view and understanding of the organizational structure and the strengths and limitations of the project.

To obtain confirmability [65], the research team audited the results and discussed the analysis repeatedly. Research reflexivity was maintained through continuous dialogues within the research team regarding personal biases that might affect data interpretation. The final strength of the study material is the inclusion of various points of view from participants as well as the research team, which is required in qualitative studies.

Conclusion

Patient perceptions of a CVW-led RPM of chronic diseases in PHC could be condensed to a "Sense of security and accessibility, but impersonal and uncoordinated," with 5 underlying categories: sense of security, care and self-care, accessibility, quality of care, and communication. This study reveals the following learnings from a patient perspective regarding CVW-led RPM: sense of security due to the health care staffs' prompt reaction on signs of health deterioration, improved motivation for care and self-care, sense of accessibility to health care staff, and enhanced quality of care.

Our findings suggest that by addressing patients' needs for new organizational routines for patient-caregiver communication, RPM via centralized virtual wards can better realize the potential of this technology.

Future research is needed to explore the health care personnel's perceptions regarding their interactions between the CVW and PHC.

Acknowledgments

We would like to thank all the participants of this study who shared their experiences of the project. Without them, this study would not have been possible. We would also like to thank the research assistant Emma Hedberg Rundgren for conducting the

XSL•FO
RenderX

interviews, Prof. Ulrica Schwartz for her assistance with the ethics approval, and Dr. Elahe Vahabnezad for sharing outside perspectives and proofreading the manuscript.

Funding

Funding for the RPM project was provided by Vinnova and Region Stockholm.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Semistructured interview guide (translated from Swedish).
[PDF File, 139 KB - humanfactors_v12i1e78780_app1.pdf]

References

1. It's time to walk the talk: WHO independent high-level commission on noncommunicable diseases—final report. : World Health Organization; 2019 URL: https://www.who.int/publications/i/item/9789241517003 [accessed 2025-11-22]
2. Operational framework for primary health care: transforming vision into action. World Health Organization. 2020 URL: https://www.who.int/publications/i/item/9789240017832 [accessed 2025-11-27]
3. Jackson C, Ball L. Continuity of care: vital, but how do we measure and promote it? Aust J Gen Pract 2018 Oct;47(10):662-664. [doi: 10.31128/AJGP-05-18-4568] [Medline: 31195766]
4. Ellner AL, Phillips RS. The coming primary care revolution. J Gen Intern Med 2017 Apr;32(4):380-386. [doi: 10.1007/s11606-016-3944-3] [Medline: 28243869]
5. Realising the potential of primary health care. : OECD Publishing; 2020 URL: https://www.oecd.org/en/publications/realising-the-potential-of-primary-health-care_a92adee4-en/full-report.html [accessed 2025-12-27]
6. Davis MM, Freeman M, Kaye J, Vuckovic N, Buckley DI. A systematic review of clinician and staff views on the acceptability of incorporating remote monitoring technology into primary care. Telemed J E Health 2014 May;20(5):428-438. [doi: 10.1089/tmj.2013.0166] [Medline: 24731239]
7. Sarasohn-Kahn J. The connected patient: charting the vital signs of remote health monitoring. : California HealthCare Foundation; 2011 URL: https://www.chcf.org/wp-content/uploads/2017/12/PDF-TheConnectedPatient.pdf [accessed 2025-11-22]
8. El-Rashidy N, El-Sappagh S, Islam SMR, M El-Bakry H, Abdelrazek S. Mobile health in remote patient monitoring for chronic diseases: principles, trends, and challenges. Diagnostics (Basel) 2021 Mar 29;11(4):607. [doi: 10.3390/diagnostics11040607] [Medline: 33805471]
9. Marcolino MS, Oliveira JAQ, D'Agostino M, Ribeiro AL, Alkmim MBM, Novillo-Ortiz D. The impact of mHealth interventions: systematic review of systematic reviews. JMIR Mhealth Uhealth 2018 Jan 17;6(1):e23. [doi: 10.2196/mhealth.8873] [Medline: 29343463]
10. Hall AK, Dodd V, Harris A, McArthur K, Dacso C, Colton LM. Heart failure patients' perceptions and use of technology to manage disease symptoms. Telemed J E Health 2014 Apr;20(4):324-331. [doi: 10.1089/tmj.2013.0146] [Medline: 24483939]
11. Huniche L, Dinesen B, Nielsen C, Grann O, Toft E. Patients' use of self-monitored readings for managing everyday life with COPD: a qualitative study. Telemed J E Health 2013 May;19(5):396-402. [doi: 10.1089/tmj.2012.0135] [Medline: 23531094]
12. Ekstedt M, Nordheim ES, Hellström A, Strandberg S, Hagerman H. Patient safety and sense of security when telemonitoring chronic conditions at home: the views of patients and healthcare professionals—a qualitative study. BMC Health Serv Res 2023 Jun 6;23(1):581. [doi: 10.1186/s12913-023-09428-1] [Medline: 37340472]
13. Hanley J, Ure J, Pagliari C, Sheikh A, McKinstry B. Experiences of patients and professionals participating in the HITS home blood pressure telemonitoring trial: a qualitative study. BMJ Open 2013 May 28;3(5):e002671. [doi: 10.1136/bmjopen-2013-002671] [Medline: 23793649]
14. Gorst SL, Coates E, Armitage C. "It's sort of a lifeline": chronic obstructive pulmonary disease patients' experiences of home telehealth. Health Psychol 2016 Jan;35(1):60-68. [doi: 10.1037/hea0000246] [Medline: 26148189]
15. Hanley J, Fairbrother P, McCloughan L, et al. Qualitative study of telemonitoring of blood glucose and blood pressure in type 2 diabetes. BMJ Open 2015 Dec 23;5(12):e008896. [doi: 10.1136/bmjopen-2015-008896] [Medline: 26700275]
16. Rahimpour M, Lovell NH, Celler BG, McCormick J. Patients' perceptions of a home telecare system. Int J Med Inform 2008 Jul;77(7):486-498. [doi: 10.1016/j.ijmedinf.2007.10.006] [Medline: 18023610]
17. Walker RC, Tong A, Howard K, Palmer SC. Patient expectations and experiences of remote monitoring for chronic diseases: systematic review and thematic synthesis of qualitative studies. Int J Med Inform 2019 Apr;124:78-85. [doi: 10.1016/j.ijmedinf.2019.01.013] [Medline: 30784430]

XSL•FO
RenderX

18. Lundell S, Modig M, Holmner Å, Wadell K. Perceptions of home telemonitoring use among patients with chronic obstructive pulmonary disease: qualitative study. JMIR Mhealth Uhealth 2020 Jun 3;8(6):e16343. [doi: 10.2196/16343] [Medline: 32490844]
19. Hilty DM, Armstrong CM, Edwards-Stewart A, Gentry MT, Luxton DD, Krupinski EA. Sensor, wearable, and remote patient monitoring competencies for clinical care and training: scoping review. J Technol Behav Sci 2021;6(2):252-277. [doi: 10.1007/s41347-020-00190-3] [Medline: 33501372]
20. Lewis G, Wright L, Vaithianathan R. Multidisciplinary case management for patients at high risk of hospitalization: comparison of virtual ward models in the United kingdom, United States, and Canada. Popul Health Manag 2012 Oct;15(5):315-321. [doi: 10.1089/pop.2011.0086] [Medline: 22788975]
21. Case study: virtual wards at Croydon Primary Care Trust. : Croydon Primary Care Trust; 2006 URL: https://archive.kingsfund.org.uk/downloads/vd66vz88c?locale=en [accessed 2025-11-22]
22. Lewis G, Vaithianathan R, Wright L, et al. Integrating care for high-risk patients in England using the virtual ward model: lessons in the process of care integration from three case sites. Int J Integr Care 2013;13:e046. [doi: 10.5334/ijic.1150] [Medline: 24250284]
23. Lewis C, Moore Z, Doyle F, Martin A, Patton D, Nugent LE. A community virtual ward model to support older persons with complex health care and social care needs. Clin Interv Aging 2017;12:985-993. [doi: 10.2147/CIA.S130876] [Medline: 28721026]
24. Kandy MC, Abdulmajeed J, Gohel CN, Gibb JM, Al-Kuwari MG. Transforming primary healthcare services with centralized health intelligence: a case study from Qatar. Qatar J. Public Health 2024 Feb 2;2023(2). [doi: 10.5339/qjph.2023.8]
25. Vallée A, Arutkin M. The transformative power of virtual hospitals for revolutionising healthcare delivery. Public Health Rev 2024;45:1606371. [doi: 10.3389/phrs.2024.1606371] [Medline: 38962359]
26. Chaiyachati KH, Shea JA, Ward M, et al. Patient and clinician perspectives of a remote monitoring program for COVID-19 and lessons for future programs. BMC Health Serv Res 2023 Jun 27;23(1):698. [doi: 10.1186/s12913-023-09684-1] [Medline: 37370059]
27. Bajgain B, Rabi S, Ahmed S, Kiryanova V, Fairie P, Santana MJ. Patient-reported experiences and outcomes of virtual care during COVID-19: a systematic review. J Patient Rep Outcomes 2023 Dec 1;7(1):126. [doi: 10.1186/s41687-023-00659-8] [Medline: 38038800]
28. Hailu R, Sousa J, Tang M, Mehrotra A, Uscher-Pines L. Challenges and facilitators in implementing remote patient monitoring programs in primary care. J Gen Intern Med 2024 Oct;39(13):2471-2477. [doi: 10.1007/s11606-023-08557-x] [Medline: 38653884]
29. Baines R, Stevens S, Austin D, et al. Patient and public willingness to share personal health data for third-party or secondary uses: systematic review. J Med Internet Res 2024 Mar 5;26:e50421. [doi: 10.2196/50421] [Medline: 38441944]
30. Serrano LP, Maita KC, Avila FR, et al. Benefits and challenges of remote patient monitoring as perceived by health care practitioners: a systematic review. Perm J 2023 Dec 15;27(4):100-111. [doi: 10.7812/TPP/23.022] [Medline: 37735970]
31. Tammes P, Morris RW, Murphy M, Salisbury C. Is continuity of primary care declining in England? Practice-level longitudinal study from 2012 to 2017. Br J Gen Pract 2021 Jun;71(707):e432-e440. [doi: 10.3399/BJGP.2020.0935] [Medline: 33947666]
32. Fairbrother P, Pinnock H, Hanley J, et al. Continuity, but at what cost? The impact of telemonitoring COPD on continuities of care: a qualitative study. Prim Care Respir J 2012 Sep;21(3):322-328. [doi: 10.4104/pcrj.2012.00068] [Medline: 22875143]
33. Yuan CT, Dy SM, Yuanhong Lai A, et al. Challenges and strategies for patient safety in primary care: a qualitative study. Am J Med Qual 2022;37(5):379-387. [doi: 10.1097/JMQ.0000000000000054] [Medline: 35404306]
34. Appleton R, Williams J, Vera San Juan N, et al. Implementation, adoption, and perceptions of telemental health during the COVID-19 pandemic: systematic review. J Med Internet Res 2021 Dec 9;23(12):e31746. [doi: 10.2196/31746] [Medline: 34709179]
35. Ure J, Pinnock H, Hanley J, et al. Piloting tele-monitoring in COPD: a mixed methods exploration of issues in design and implementation. Prim Care Respir J 2012 Mar;21(1):57-64. [doi: 10.4104/pcrj.2011.00065] [Medline: 21785816]
36. Oudbier SJ, Souget-Ruff SP, Chen BSJ, Ziesemer KA, Meij HJ, Smets EMA. Implementation barriers and facilitators of remote monitoring, remote consultation and digital care platforms through the eyes of healthcare professionals: a review of reviews. BMJ Open 2024 Jun 10;14(6):e075833. [doi: 10.1136/bmjopen-2023-075833] [Medline: 38858155]
37. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
38. Wannheden C, Stenfors T, Stenling A, von Thiele Schwarz U. Satisfied or frustrated? A qualitative analysis of need satisfying and need frustrating experiences of engaging with digital health technology in chronic care. Front Public Health 2020;8:623773. [doi: 10.3389/fpubh.2020.623773] [Medline: 33585391]
39. Wannheden C, Roczniewska M, Hasson H, Karlgren K, von Thiele Schwarz U. Better self-care through co-care? A latent profile analysis of primary care patients' experiences of e-health-supported chronic care management. Front Public Health 2022;10:960383. [doi: 10.3389/fpubh.2022.960383] [Medline: 36211687]
40. Background. Medrave Software. 2017. URL: https://www.medrave.com/about [accessed 2025-11-25]

XSL•FO
RenderX

41. Corentin MB, Gibb S. Zoom: a spatial data visualization tool (version 2.0.6). 2014. URL: https://cbarbu.r-universe.dev/zoom#:~:text=Barbu%2C%20CM.%20(2014).%20zoom%3A%20A,Corresponding%20BibTeX%20entry%3A%20Access%20date:%202025-11-25 [accessed 2025-11-27]
42. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today 2004 Feb;24(2):105-112. [doi: 10.1016/j.nedt.2003.10.001] [Medline: 14769454]
43. Kyngäs H. Inductive content analysis. In: The Application of Content Analysis in Nursing Science Research: Springer; 2019:13-21. [doi: 10.1007/978-3-030-30199-6_2]
44. Andersen TO, Langstrup H, Lomborg S. Experiences with wearable activity data during self-care by chronic heart patients: qualitative study. J Med Internet Res 2020 Jul 20;22(7):e15873. [doi: 10.2196/15873] [Medline: 32706663]
45. Lu JF, Chi MJ, Chen CM. Advocacy of home telehealth care among consumers with chronic conditions. J Clin Nurs 2014 Mar;23(5-6):811-819. [doi: 10.1111/jocn.12156] [Medline: 23796027]
46. McCann L, Lewis L, Oduntan O, et al. Patients' and clinicians' experiences using a real-time remote monitoring system for chemotherapy symptom management (ASyMS): qualitative study. J Med Internet Res 2024 Dec 3;26:e53834. [doi: 10.2196/53834] [Medline: 39626227]
47. Petersson P, Blomqvist K. Sense of security—searching for its meaning by using stories: a Participatory Action Research study in health and social care in Sweden. Int J Older People Nurs 2011 Mar;6(1):25-32. [doi: 10.1111/j.1748-3743.2010.00211.x] [Medline: 21303461]
48. Tan SY, Sumner J, Wang Y, Wenjun Yip A. A systematic review of the impacts of remote patient monitoring (RPM) interventions on safety, adherence, quality-of-life and cost-related outcomes. NPJ Digit Med 2024 Jul 18;7(1):192. [doi: 10.1038/s41746-024-01182-w] [Medline: 39025937]
49. Berk ML, Schur CL, Cantor J. Ability to obtain health care: recent estimates from the Robert Wood Johnson Foundation National Access to Care Survey. Health Aff (Millwood) 1995;14(3):139-146. [doi: 10.1377/hlthaff.14.3.139] [Medline: 7498887]
50. Bindman AB, Grumbach K, Osmond D, et al. Preventable hospitalizations and access to health care. JAMA 1995 Jul 26;274(4):305-311. [doi: 10.1001/jama.1995.03530040033037] [Medline: 7609259]
51. Rosano A, Loha CA, Falvo R, et al. The relationship between avoidable hospitalization and accessibility to primary care: a systematic review. Eur J Public Health 2013 Jun;23(3):356-360. [doi: 10.1093/eurpub/cks053] [Medline: 22645236]
52. Lurie N, Ward NB, Shapiro MF, Brook RH. Termination from Medi-Cal—does it affect health? N Engl J Med 1984 Aug 16;311(7):480-484. [doi: 10.1056/nejm198408163110735] [Medline: 6379458]
53. Nundy S, Lu CYE, Hogan P, Mishra A, Peek ME. Using patient-generated health data from mobile technologies for diabetes self-management support: provider perspectives from an academic medical center. J Diabetes Sci Technol 2014 Jan;8(1):74-82. [doi: 10.1177/1932296813511727] [Medline: 24876541]
54. Kerari A, Bahari G, Alharbi K, Alenazi L. The effectiveness of the chronic disease self-management program in improving patients' self-efficacy and health-related behaviors: a quasi-experimental study. Healthcare (Basel) 2024 Apr 3;12(7):778. [doi: 10.3390/healthcare12070778] [Medline: 38610201]
55. Allegrante JP. Advancing the science of behavioral self-management of chronic disease: the arc of a research trajectory. Health Educ Behav 2018 Feb;45(1):6-13. [doi: 10.1177/1090198117749010] [Medline: 29258348]
56. Allegrante JP, Wells MT, Peterson JC. Interventions to support behavioral self-management of chronic diseases. Annu Rev Public Health 2019 Apr 1;40:127-146. [doi: 10.1146/annurev-publhealth-040218-044008] [Medline: 30601717]
57. Nelson EC, Meyer G, Bohmer R. Self-care: the new principal care. J Ambul Care Manage 2014;37(3):219-225. [doi: 10.1097/JAC.0000000000000036] [Medline: 24887523]
58. Johnsson N, Strandberg S, Tuvesson H, Fagerström C, Ekstedt M, Lindberg C. Delineating and clarifying the concept of self-care monitoring: a concept analysis. Int J Qual Stud Health Well-being 2023 Dec;18(1):2241231. [doi: 10.1080/17482631.2023.2241231] [Medline: 37506372]
59. Deering MJ, Siminerio E, Weinstein S. Issue brief: patient-generated health data and health IT. : National Learning Consortium; 2013 URL: https://www.healthit.gov/sites/default/files/patient_generated_data_factsheet.pdf [accessed 2025-11-22]
60. Quinn CC, Chard S, Roth EG, Eckert JK, Russman KM, Cross RK. The Telemedicine for Patients With Inflammatory Bowel Disease (TELE-IBD) clinical trial: qualitative assessment of participants' perceptions. J Med Internet Res 2019 Jun 3;21(6):e14165. [doi: 10.2196/14165] [Medline: 31162128]
61. Stavropoulou A, Rovithis M, Kelesi M, et al. What quality of care means? Exploring clinical nurses' perceptions on the concept of quality care: a qualitative study. Clin Pract 2022 Jun 30;12(4):468-481. [doi: 10.3390/clinpract12040051] [Medline: 35892437]
62. Ignatowicz A, Slowther AM, Elder P, et al. Ethical implications of digital communication for the patient-clinician relationship: analysis of interviews with clinicians and young adults with long term conditions (the LYNC study). BMC Med Ethics 2018 Feb 23;19(1):11. [doi: 10.1186/s12910-018-0250-0] [Medline: 29475437]
63. Sinnott C, Mc Hugh S, Browne J, Bradley C. GPs' perspectives on the management of patients with multimorbidity: systematic review and synthesis of qualitative research. BMJ Open 2013 Sep 13;3(9):e003610. [doi: 10.1136/bmjopen-2013-003610] [Medline: 24038011]

XSL•FO
RenderX

64. Freilich J, Nilsson GH, Ekstedt M, Flink M. “Standing on common ground”—a qualitative study of self-management support for patients with multimorbidity in primary health care. BMC Fam Pract 2020 Nov 17;21(1):233. [doi: 10.1186/s12875-020-01290-y] [Medline: 33203401]
65. Stahl NA, King J. Expanding approaches for research: understanding and using trustworthiness in qualitative research. Journal of Developmental Education 2020;44:26-28 [FREE Full text]

Abbreviations

CVW: centralized virtual ward
PHC: primary health care
RPM: remote patient monitoring

Edited by A Choudhury; submitted 09.06.25; peer-reviewed by HHL Chiu, Z Jonassen; revised version received 04.11.25; accepted 04.11.25; published 01.12.25.

Please cite as:
Jaranka A, Nilsson GH, Stenfors T, Hägglund M, Papachristou P, Taloyan M
Patient’s Perceptions of a Centralized Virtual Ward for Remote Patient Monitoring in Primary Care: Qualitative Study
JMIR Hum Factors 2025;12:e78780
URL: https://humanfactors.jmir.org/2025/1/e78780
doi:10.2196/78780

© Alex Jaranka, Gunnar H Nilsson, Terese Stenfors, Maria Hägglund, Panagiotis Papachristou, Marina Taloyan. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 1.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Family Caregivers' Perspectives on the Potential of Drone-Based Medication Delivery in Palliative Home Care: Qualitative Focus Group Study

Franziska Fink[1], Dr Rer Nat; Sabrina Claudia Otto[2], Mmel; Martin Grünthal[3], StEx; Anne Lehmann[1], MSc; Patrick Jahn[1], Prof Dr

[1]Translation Region for Digitalized Healthcare, Department of Internal Medicine, Faculty of Medicine, Martin Luther University Halle-Wittenberg, Magdeburger Straße 12, Halle, Germany

[2]Psychiatric Center Nordbaden, Clinic for Forensic Psychiatry and Psychotherapy, Heidelberger Str. 1, Wiesloch, Germany

[3]Pharmacy at the Bauhaus, Gropiusallee 50, Dessau-Roßlau, Germany

Corresponding Author:
Franziska Fink, Dr Rer Nat
Translation Region for Digitalized Healthcare, Department of Internal Medicine, Faculty of Medicine, Martin Luther University Halle-Wittenberg, Magdeburger Straße 12, Halle, Germany

Abstract

Background: Palliative care supports individuals with incurable, life-threatening illnesses, focusing on symptom management and quality of life. Access to timely care, including essential medications, is often limited, particularly in rural areas, leading to gaps in home-based care. Digital health technologies, including drone-based delivery systems, have the potential to address such logistical challenges. For these technologies to be effective, they must be adapted to the specific needs of patients and caregivers, which often differ from general health care contexts, especially in remote areas.

Objective: This study aimed to identify user needs and requirements for a drone-based medication delivery system designed to supplement traditional courier services in palliative care. The study also sought to explore practical considerations for integrating drones into home-based care workflows, with the goal of informing both technical design and clinical applicability.

Methods: We conducted 1 focus group (FG) and 4 semistructured individual interviews with family caregivers involved in home-based palliative care. Participants were recruited via local palliative care services in rural regions. The discussions focused on experiences with palliative care, medication logistics, needs, expectations, and concerns regarding drone-assisted delivery. Interviews and the FG were audio-recorded, transcribed, and analyzed using structured qualitative content analysis.

Results: A total of 10 caregivers participated (mean age 68.6, SD 12.3 years) in this study. Six participants took part in the FG and 4 participants were part of individually conducted interviews. Caregivers frequently reported long travel distances to obtain medications, sometimes leaving patients unattended and noting medication shortages, especially in the afternoons and on weekends. Participants highlighted the potential of drones to supplement existing courier services during periods of high demand. At the same time, caregivers expressed concerns about limited technical skills, particularly regarding mobile apps for ordering deliveries, and emphasized the need for simple, user-friendly systems.

Conclusions: These findings reveal logistical gaps in palliative care medication supply and design requirements for drone-based delivery systems. There is a clear demand for faster, more reliable access to essential medication, especially during evenings and weekends. Whether drone-assisted delivery can reduce hospitalizations and support patients' wishes to remain at home warrants further research. Specialized outpatient palliative care teams are central to eHealth adoption, providing technical guidance, emotional support, and confidence in using new technologies. Their involvement in communication and training is critical. Strengthening digital health literacy requires individual training and structural adjustments to meet the needs of older caregivers and those with limited digital affinity. Concerns about losing personal interaction remain significant; digital tools should complement, not replace, face-to-face care. Timely medication access could reduce caregiver stress and improve quality of life by supporting symptom control at home. To foster acceptance, caregivers need hands-on experience. Future studies should test drone delivery in real-world settings, addressing technological, psychological, and social factors.

(JMIR Hum Factors 2025;12:e80320) doi:10.2196/80320

KEYWORDS

palliative care; family caregivers; drone-based medication delivery; eHealth; technology acceptance; specialized outpatient palliative care; SOPC

XSL•FO
RenderX

Introduction

Over the past decades, most developed countries have seen declining mortality rates [1,2]. This increase in life expectancy is accompanied by a rise in chronic and incurable diseases, especially in old age, such as cardiovascular conditions, cancer, and diabetes. The end of life is often marked by enduring and heavy symptom burdens, prompting the development of specialized palliative care to support patients and families in leading a self-determined life and maintaining quality of life [3,4].

Most palliative patients prefer to remain in their familiar home environment. While 76% wish to spend their final days at home, only about 20% actually do, with nearly half dying unplanned in hospital settings [5]. The care setting significantly impacts quality of life; patients at home generally experience better outcomes than those in institutional settings [6].

As the number of terminally ill people increases, home care becomes more reliant on family caregivers. The level of support required varies depending on the illness [7]. This puts a tremendous amount of pressure on family caregivers, as they are often unable to cope with the stressful situation and may feel helpless in the face of the situation [3,8]. In this regard, it was indicated that the pattern of enduring in the social, psychological, physical, and spiritual domains at certain time points (diagnosis, return to home, recurrence, and terminal stage) is similar between patients and their caregivers, which is mainly influenced by the nature of the disease and the availability of professional care [3,9,10]. This implies that the needs of patients and caregivers should be addressed in parallel, as these key time points are identical for patients and caregivers [3,9].

Palliative care aims primarily at minimizing pain to preserve patients' dignity and life quality, which also positively affects caregivers' well-being [11,12]. Fast and flexible symptom management is crucial and often requires rapid medication adjustments by caregivers or specialized outpatient palliative care (SOPC) teams [13]. The inability to respond quickly and adjust necessary medication can jeopardize home care and lead to unwanted hospital admissions and a deterioration in the patient's and caregivers' quality of life. In other words, it is crucial to ensure rapid and direct access to essential medicines that improve quality of life, especially in rural areas.

In this context, drones may offer value by significantly improving delivery response times in health services [14]. Drones are increasingly applied in health care and medicine, particularly in areas such as public health, disaster response, telemedicine, and medical logistics [15]. In the context of medical transport, drones are used to deliver medications (eg, vaccines, pharmaceuticals), blood products, organs, defibrillators, and other essential medical equipment [15-20]. These medical drones offer substantial advantages for health care, including faster emergency response times, improved access to medical services in remote or hard-to-reach regions, and enhanced clinical outcomes, such as increased survival rates [14,21-23]. Compared with helicopters, drones offer a more cost-effective solution for transporting medical supplies. Their utility is especially pronounced in regions with challenging terrains such as mountains, deserts, or forests, in areas with limited road access, over long distances, or in locations impacted by large-scale natural disasters [18,19,24]. Overall, drones hold significant promise for improving health care delivery in such contexts [18]. Despite their considerable potential in health care, the adoption and effectiveness of drones are influenced by technological, organizational, and environmental factors (TOE barriers) [14,25,26]. In Germany and many other countries, structural and regulatory conditions, such as undefined legal frameworks and operational guidelines, currently limit the use of drones to research projects and prototype testing [14,27]. Most existing projects on medical drones have been conducted in Africa, where regulatory conditions are comparatively more flexible than in Europe or North America [27,28]. However, findings from African contexts cannot be directly generalized to European settings [29]. Among the factors affecting drone usage, usability, and design consistently emerge as the most critical. In addition, users' limited skills, negative perceptions of the technology, and the relative immaturity of drone systems pose significant barriers to broader adoption [14,25,26].

Despite increasing interest in drone-based medication delivery, robust evidence on its impact in real-world health care logistics remains limited. Most existing studies have focused on single scenarios [20,21,23,30-32] or relied on simulations [33,34], and thus have not been conducted in real-world settings. To date, 1 study covered the entire development process, from the initial app prototypes to the complete ordering and delivery workflow using an ordering app and drone over densely populated areas in Germany [35]. In particular, to our knowledge, no study has investigated drone-based medication delivery in the real-world context of palliative care, a setting characterized by complex symptom management, urgent medication needs, and a high reliance on family caregivers. Consequently, it remains unclear whether drones could meaningfully improve medication logistics, support care networks, or enhance symptom management and quality of life in outpatient or inpatient palliative care. To address these gaps, the present study is embedded within the PalliDrone project, which aims to co-create and evaluate a drone-based medication delivery system for palliative care [36,37]. The PalliDrone project comprises multiple intervention components (T0–T3), encompassing medical and nursing strategies. It will assess the usefulness, effectiveness, and benefits of drone delivery compared to conventional courier services, treating it as a complex health intervention [38]. In this context, the present study focuses on identifying user needs, potential benefits, and practical requirements for integration into existing palliative care processes. Specifically, the research objectives are:

1. To explore the needs, expectations, and concerns of family caregivers regarding drone-based medication delivery in palliative care.
2. To assess how drone delivery could alleviate burdens on caregivers and support timely access to essential medications.
3. To identify practical and organizational factors that may facilitate or hinder the implementation of drone-based medication logistics in real-world palliative care settings.

XSL•FO
RenderX

In this study, we aim to explore the practical requirements, opportunities, and challenges associated with drone-based medication delivery in palliative care for family caregivers. Specifically, the study seeks to inform both the technical readiness of a user-centered drone delivery system and its potential integration into clinical workflows. By assessing user needs, expectations, and concerns, the study provides insights that are relevant for the development of the drone system itself as well as for its safe, effective, and feasible use in routine palliative care practice. This dual focus ensures that the intervention addresses both technological feasibility and clinical applicability.

Methods

Overview

This study adheres to the COREQ (Consolidated Criteria for Reporting Qualitative Research) to enhance transparency, credibility, and completeness in qualitative reporting. A completed COREQ checklist is provided in Checklist 1. The COREQ checklist was used as a guiding framework during the preparation and reporting stages. Before data collection, the checklist was instrumental in informing the study design, ensuring that all pertinent aspects—including the composition and reflexivity of the research team, participant selection strategy, and detailed data collection procedures—were meticulously planned and documented in advance. During the process of manuscript development, the COREQ items were systematically reviewed to verify that each relevant element was transparently reported in the Methods and Results sections. This process ensured completeness and facilitated reproducibility.

Setting and Sampling

The focus group (FG) and interviews included family caregivers selected through purposive sampling. The sampling was used, with participants being purposively recruited to ensure a diverse range of experiences relevant to the study's objective. The selection of participants was conducted by the SOPC teams of the "Anhaltische Hospiz- und Palliativgesellschaft" in Dessau-Roßlau, Germany, and the "Leipziger Palliativgesellschaft" in Torgau, Germany. Participants were approached by telephone or via email. The SOPC teams identified potential participants who met the inclusion criterion of having provided care for a palliative patient within the past 6 months in rural regions. Within this sample, variation in relationship to the patient (spouse and other relative) and gender was sought to capture a range of perspectives. Although the data collection process did not entail the application of formal stratification criteria, such as age, caregiving burden, or technology literacy, the SOPC teams were encouraged to extend invitations to individuals with a range of backgrounds, thereby fostering heterogeneity in the ensuing discussions. Group effects were minimized, and no strict exclusion criteria were applied, reflecting the open, cocreative nature of the development process. There was no relationship to participants before the study.

Materials

A German-language FG and interview guide was developed through a systematic, multistage process informed by qualitative research principles rooted in a constructivist paradigm by AL and FF [30,39]. This method was chosen due to existing theoretical assumptions, with the aim of exploring a wide range of perspectives rather than reaching consensus. The guide focused on eliciting argumentative patterns, that is, the reasoning and justification underlying participants' views, rather than collecting biographical narratives [40]. The stepwise procedure (collecting, checking, sorting, and subsuming) ensured systematic coverage of relevant themes while maintaining openness to emergent issues [39]. First, potentially relevant questions were collected without concern for their final usability (collecting). Next, the questions were evaluated against qualitative guideline criteria (checking). Then, the content was sorted into thematic areas (sorting). Finally, these were subsumed into question blocks, each beginning with an open introductory question. The sequence of these blocks was also determined (subsuming). Critical review and feedback were provided by PJ, MG, and members of the PalliDrone project team (peer group). Feedback aimed to assess the clarity and relevance of items. No fixed instructions were given, and the guide was refined after each feedback round.

In terms of content, the guide development was informed by established models of technology acceptance, particularly the technology acceptance model (TAM 1 - 3) [41-43]. TAM proposes that technology adoption is primarily driven by perceived usefulness and perceived ease of use, complemented in later extensions by factors such as social, cognitive, and psychological factors. These factors collectively shape the intention to use. The Technology Usage Inventory expands upon the traditional technology acceptance factors of the TAM, such as perceived usability, usefulness, immersion, and accessibility (technology-specific), by incorporating key psychological constructs, including interest, curiosity, technology-related anxiety, and skepticism [44,45]. The guide built on these theoretical foundations and on instruments from previous research on cocreative development of drone-assisted delivery and mobile app [30]. Thus, the final guide includes four categories: (1) processes and problem situations in palliative care (is-state), (2) specification of the process flow and communication (usability), (3) possibilities for realizing technical integration (usefulness), and (4) concerns regarding the implementation and integration (concerns). The guide began with a narrative stimulus (description) to initiate the conversation, followed by open-ended thematic question blocks. These were built on previous responses (immanent) and introduced new themes (exmanent). Transitions between blocks were deliberately soft to encourage storytelling. Evaluative questions were asked at the end. The complete FG and interview guide is available in Multimedia Appendix 1.

Data Collection

All FG and interview discussions were conducted in German. The FG was conducted at "Leipziger Palliativgesellschaft" in Torgau, Germany. With the help of the "Anhalt-Hospiz Dessau", interviews were conducted at the homes of caregivers. Two

XSL•FO
RenderX

moderators (M) guided both formats: M1 (AL, female, research associate, experienced in mixed methods) actively led the sessions using an inductive questioning approach (starting with immanent before exmanent questions), while M2 (FF, female, postdoc, experienced in FG and interviews) remained more passive and managed the input setting. M1 moderated the discussion, asked questions, and encouraged participants to elaborate, drawing out a range of views. M2 documented interactions, atmosphere, and key statements through an ad-hoc protocol, which was later made available for participants to verify the accuracy of their contributions. Before the data collection, M1 received interview training from M2. Both moderators were familiar with the subject matter and guided their actions using the principle of "methodological understanding of others" [46].

At the beginning of each session, moderators introduced themselves and explained the procedure. A detailed overview of drone-based medication delivery was presented, followed by an open discussion about possible applications in palliative care. Participants reflected on urgency, expected relief for caregivers, and situations where traditional courier delivery would suffice.

Sessions were audio-recorded. Data validation was ensured through member checking. Participants received summaries of their respective FG discussions or interviews and were invited to provide feedback or corrections. This process helped to confirm the accuracy and credibility of the data while respecting participants' time and availability.

Participants were informed in writing and verbally about the recording and their right to withdraw consent at any time.

The FG lasted about 90 minutes; interviews lasted around 30 minutes, depending on the situation. In interviews, 1 member of the SOPC team was present, and in the FG, 2 members were present. This was done deliberately to provide emotional support for participants, given that recruitment targeted caregivers within 6 months of bereavement, which could evoke distress.

Data Analyses

All collected statements during the FG and interviews were repeated and discussed with participants. All participants confirmed for correctness. The participants' responses were transcribed and coded by 1 researcher in German (AL). Transcripts were not returned to participants. A simple transcription system was used, that is, only manifest content was analyzed [47,48]. The data analysis was conducted according to the content analysis method from Elo and Kyngäs [47] and was coded using a deductive approach moving from general to specific [49]. All data were coded with the MAXQDA 2022 program (VERBI Software GmbH). The unit of analysis (event) was defined as words and as one or several sentences that contain one theme of the guideline-based questions [50]. Thus, aspects that would fit the matrix were chosen from the data. The categories were formed along the guideline-based question complexes [30,44] (see Table 1).

To validate the concept, a second researcher (FF), familiar with the analyses, independently coded 20% of the transcripts to calculate the interrater reliability. In case of discrepancies, discussions were held between the 2 researchers until a consensus was reached. Cohen Kappa was computed for all variables [51]. The interrater reliability was $\kappa=0.71$ ($P<.001$) with a substantial agreement [51,52].

For reporting purposes, illustrative quotes and examples were translated into English. Translations were iteratively cross-checked by FF and AL of the research team to ensure that the original meaning, tone, and context were preserved.

XSL•FO
RenderX

Table . Structured analysis matrix showing the categorization of caregiver statements collected during focus groups and interviews. The study explored perspectives on drone-based medication delivery among family caregivers of recently deceased palliative patients (within 6 months) in rural areas in Germany, in 2024.

No.	Category	Definition	Example
1	Is-state	Describes the initial situation and current state of problems and work procedures in palliative care as well as the competence in handling medical apps and drones	—[a]
1.1	Work procedure	Describes the current work flow in palliative care	“We always plan for a week and adjust it every day, depending on what we need.”
1.2	Problems	Describes current problems in palliative care that justify the usefulness of the technology	“So we have to speed things up so that we can do all this.”
1.3	Competence	Describes the current level of knowledge and skills in handling medical apps and drones	“Well, I have no experience. Sometimes you see one flying and it’s a bit easy.”
2	Usability	Describes the necessities or conditions for the comprehensibility, application, and usability of the technology	—
2.1	Realization options	Describes different implementation options for the drone, app, and the general process	“It would really be more interesting for rural areas, where there is a bit more space and people have their yard.”
2.2	Communication	Describes the necessities or conditions for the communication of the technology	"I could imagine that. But then it has to be quick. Because when I call, I’m on a house call and a normal patient, I want to prescribe an antibiotic. It’s the weekend, Saturday if you like, I need it on duty today if you like and then I call. Then it’s really a matter of a minute and I don’t stay in this house call forever. And then I think it has to happen within five minutes.”
2.3	Drone process	Describes the necessities or conditions/characteristics of the process of technology	“If you are logged in, you will see all patients. When I log in, I also see all patients.”
2.4	Ordering	Describes the necessities or conditions for the ordering of medication	“But that’s basically just triggering the prescription. I could also call. I think in this case it’s quite practical via this app. Triggering the prescription.”
2.5	Delivery	Describes the necessities or conditions for the transfer of medication	“For example: It should be dropped off right here on the meadow, or somewhere up ahead. And when the person of SOPC team, triggers the order, they can decide: Delivery here or to the customer. And the caregiver can only decide: Delivery to my home.”
3	Usefulness	Describes the advantages, the benefits of the technology	"We’ve had it once or twice now, we had a ward round with the doctor. She wanted to inject a medication, but we didn’t have it there. It would be good to say, I need this now. It’ll be there in 10 min.”

XSL•FO
RenderX

No.	Category	Definition	Example
4	Concerns	Describes the dangers, the risk, the complications, the disadvantages, the difficulties that the technology brings with it	"I think there are problems with the fact that the patients we look after can't do that at all. [...] So I say, if you're there as a nurse yourself and you have the app on your duty phone, it may all work. You can use it, but the patients we look after, I'd say 70, 80 percent of them, can't do that. And they're not mobile either, they're immobile, they can't even go out the door. I think that's actually the big problem."

[a]Not available.

Ethical Considerations

The PalliDrone project was reviewed and approved by the Ethics Committee of Martin-Luther-University Halle-Wittenberg (case number 2024‐038, date of approval April 18, 2024) for the conduct of the PalliDrone study. All participants gave their informed written consent before the FG and semistructured interviews and were informed about the procedure, the general aim, and the reasons for the discussion or interview. Participants were informed of their right to withdraw at any time without any consequences. For participants who were approached for qualitative interviews following recent bereavement, special care was taken to ensure they had the opportunity to ask questions and receive adequate explanations before consenting. All data were anonymized before analysis. All personal identifiers (eg, names and locations) were removed during transcription. Transcripts and coded data were stored on secure, password-protected servers accessible only to the research team. Participants did not receive financial compensation for their participation but were offered reimbursement of travel expenses where applicable. No identifiable images or other identifying information about participants are presented in this manuscript or in the supplementary materials.

Results

Characteristics of Caregivers

Between June and July 2024, we collected data from 10 participants. Six participants took part in the FG, and 4 participants were part of individually conducted interviews. No one dropped out during the discussions or interviews. All caregivers (n=10; mean age 68.6, SD 12.3) lacked knowledge and skills in using drones, and only 30% (3/10) were familiar with using medication apps. Of the 10 participants, 30% (n=3) care for their relatives while they are still working. Additional findings are reported in Table 2.

Table . Demographic characteristics of study participants, including age, gender, technology competence, and education (N=10). Participants were family caregivers of recently deceased palliative patients (within 6 months) in rural areas in Germany.

Characteristic	Value
Age, mean (SD)	68.60 (12.29)
Age range (years)	49‐84
Gender, n	
Female	8
Male	2
Drone competence, n	0
Medication app competence, n	3
Professional active, n	3
Secondary school, n	6
High school, n	3
Profession degree, n	5
University degree, n	5

XSL•FO
RenderX

Qualitative Content Analyses

Overview

Is-State

With 80 - 81 utterances, talking about their daily lives and problems was the most reported topic during the interviews and FG. However, all participants reported that they worked well with the SOPC.

> *[…] I've already called at night when he gets his pain attacks again, […] it's wonderful, they [SOPC] were always available.*

Some of the participants reported that they had little or no technical experience (competence), did not own a tablet or smartphone, and therefore could not imagine ordering medication via app and drone. Technical difficulties or poor network coverage in rural areas are barriers to the use of drone-based medication delivery (problems). In particular, those who reported no problems with medication delivery said they did not see the point of having medication delivered by drone. While some of the caregivers stated that the medication delivery with the courier service has been good so far, most of the participants reported problems with long errands for medication and transportation tickets. One of the biggest difficulties is leaving their relatives alone for too long:

> *You come home and he's lying down again because he tried to get up and he couldn't. But those are situations where you're really pushing yourself to the limit. I could see [the drone] really taking the pressure off.*

They mentioned that home delivery may be a relief in any case.

Another major problem mentioned was delivery bottlenecks during evenings and weekends. One woman described the problem this way:

> *I also took over the final care of my uncle. He had lung cancer. And he was taken out of the hospital without medication. He was screaming in pain. I had an aunt who was completely overwhelmed. […] I got in the car and drove to the next town. Friday night. Pharmacies closed. No pharmacy available. […] We just didn't have the medicine. […] in the morning we had the same problem. It's a small town, but the pharmacy had been closed for three months because the owner was sick. So we had to go back to the next bigger town. Same problem. Who should I send? My aunt couldn't have made it. […] I was really lucky that the doctor came by chance, his family doctor, and they knew each other privately. She said I'll stay here until you get back and I drove myself. But that's where the problems start. If I had a drone, a lot of things would have been easier for me.*

In addition, many caregivers have little family support, are often older adults, are no longer able to drive, or have to travel long distances, making it difficult to provide care and obtain medications. The lack of time available to general practitioners was also mentioned as a problem.

Another mentioned problem was the e-prescription. A key issue with e-prescriptions is delayed availability; pharmacies often cannot access all medications immediately after the chip card is inserted, as updates may take additional time.

Usability

Participants emphasized that the delivery process should be clear and as simple as possible. Since caregivers often need more than just medication—such as transport tickets—they suggested drones could also deliver important documents. Generally, they found it difficult to imagine ordering medication via an app and drone but appreciated the idea of centralizing all information: "because then you only have one point of contact." Most participants could only envision drone delivery being useful in rural areas.

In terms of communication, participants stressed the importance of direct contact with SOPC services, especially for consultations:

> *And there are also side effects from these medications. That's the case when you have palliative care and questions quickly arise.*

They would be more willing to use an app if it enabled phone communication with general practitioners, SOPC teams, and pharmacies. A chat function was dismissed as too impersonal and lacking advisory value. One participant with a caregiver who spoke limited German suggested the app should support multiple languages.

Some participants preferred that the SOPC team initiate medication orders through the app, with individual orders placed only if necessary. They also valued having a choice between courier and drone delivery.

Regarding logistics, caregivers suggested that "the nursing service should be involved" and that medication be delivered directly to the SOPC service. While most found it hard to visualize the drone process, "find it hard to imagine how such a drone would work," they could imagine choosing half-hour delivery time slots and having "an alternative" between courier or drone.

Usefulness

Although usefulness was in sum the category with the lowest statements, it was the category with the most agreement that drones would be useful in "emergency situation where they're panicking or something," in rural areas, "over the weekend," or in situations where the relatives have to be left alone:

> *My husband was very ill at the time, he always said when I had to go away […]: Don't leave me alone. There was no one there […]. It's all such a problem […] so with the drone, that's a good thing.*

Concerns

Participants encounter a range of challenges when delivering medications via drone. These include adverse weather conditions, inadequate network infrastructure, dead spots in remote areas, malfunctioning technology, and users who are unable to operate the technology. In addition, data protection concerns and escalating costs for users further compound the

XSL•FO
RenderX

challenges. Notably, the primary concern among participants pertains to technical issues with the drone, such as its ability to reach its destination and its reliability. There is also a fear that "the drone may crash."

The participants expressed concerns that the introduction of drone-based medication delivery systems might result in a diminution of human interaction with pharmaceutical services:

> *I can well imagine that if someone is alone at home and then only has a telephone, they no longer have any relationship with other people. That's why I think it's very important for me, […] that there is [...] personal contact, with the doctor or, as is the case with palliative care, that human closeness is not lost.*

This is accompanied by "the concern that if you take advantage of this, you might get one less home visit."

Another concern raised was the logistics of medication delivery by drone if "people […] live in a new building, three floors up, are single, old, over 80 and even older, then the drone comes and drops the medicine in front of the door, downstairs, who's going to fetch it?"

This concern was further compounded by the possibility that the delivery would be made to a lower floor, necessitating the retrieval of the medicine by someone else, like children or pets. The participants expressed uncertainty regarding the procedural mechanisms involved in the administration of medication delivered by drone. Specifically, there was a lack of clarity concerning the necessity of signature authentication in such circumstances, given the assumption that the delivery would be formalized through the usage of a signature.

Discussion

Principal Findings

This study aimed to explore the practical requirements, opportunities, and challenges associated with drone-based medication delivery in palliative care for family caregivers by addressing their needs, relief, and requirements. FG and interview results revealed that family caregivers face challenges in assessing the usefulness of drone delivery, largely due to unfamiliarity with such technologies and generally low eHealth literacy [53]. Caregivers highlighted frequent delays in medication access, especially during evenings and weekends, which often led to emergency interventions and increased stress. They expressed a need for reliable and fast delivery solutions that operate beyond regular hours, while emphasizing the importance of maintaining human contact and trusted communication channels, particularly with SOPC teams.

Needs, Expectations, and Concerns of Family Caregivers Regarding Drone-Based Medication Delivery in Palliative Care

This study revealed that family caregivers face persistent challenges in accessing medications in a timely manner, particularly during evenings and weekends. These delivery gaps often result in emergency interventions or unplanned hospital admissions, increasing stress and contradicting patients' wishes to remain at home [5,6]. Such problems are especially acute during early stages of palliative care, before SOPC teams are established [3,8-10]. This highlights the need for a reliable, fast delivery system that operates even during off-hours, where conventional courier services may fail. Studies in rural medicine and emergency supply delivery have demonstrated drones' ability to overcome logistical and geographic barriers and reduce delivery times, especially in areas with limited infrastructure [15-24,32]. For example, the usage of drones for the delivery of defibrillators has led to a notable enhancement in accessibility to essential lifesaving equipment [21-23,32]. Consistent with the findings of our study, these investigations underscore the value of drones as a useful addition to existing logistics, rather than a replacement. They demonstrate the potential of drones to facilitate access to critical medical supplies through faster delivery [54]. However, there is a paucity of studies that have used drones within the framework of a living lab design to substantiate their effectiveness and practicality within a particular context. Moreover, palliative care differs from emergency contexts in its ongoing, sensitive care relationships where continuity and human interaction are paramount.

However, the family caregivers also highlighted a limited familiarity with drones and generally low eHealth literacy, making it difficult to assess the usefulness of the technology. The technological gap makes it difficult to assess what caregivers truly need in this context. This gap necessitates empowering both individuals through skill development and adapting social and technical infrastructures accordingly, because digital health literacy is an interplay between user competence, system readiness, and their interaction [55]. At the same time, the likelihood of requiring palliative care increases with age, raising concerns about a mismatch between technical competence and care needs. Although the most common palliative care services are symptom management (88%) and psychological support (81%), which often require face-to-face contact, eHealth can complement traditional support to some extent [56]. However, concerns regarding a potential loss of trusted human contact, particularly with SOPC teams, privacy risks, and the possibility that digital ordering systems could exclude those with limited technical competence were faced in this context. These concerns align with previous studies reporting that palliative care users value face-to-face relationships and clear communication as part of care delivery [53,57]. This indicates that the loss of human connection is a key concern, mirroring previous research [31,53,58]. Given disparities in access to palliative care globally, alternative delivery models that help patients and families communicate needs are essential [59]. This is particularly important for family caregivers, who frequently report feelings of helplessness in managing complex care situations [8,60].

Drone-Assisted Delivery as a Relief for Family Caregivers

Despite these concerns, caregivers acknowledged the potential benefits of drones, especially for logistical relief and emergency access. Participants described how delivery delays often required them to travel long distances to pharmacies, leaving patients unattended. These situations were particularly stressful for older adult caregivers, many of whom had limited mobility or transport options [61,62]. A reliable drone delivery system was

XSL•FO
RenderX

perceived as potentially reducing such logistical burdens by shortening delivery times and enabling rapid response during urgent situations, especially in rural or poorly connected areas [15-24,32]. Although most participants had no previous experience with drones, they acknowledged that faster access to medication could improve symptom control and help maintain patient comfort, which in turn might reduce caregivers' stress levels. Previous studies have demonstrated strong links between patient well-being and caregiver quality of life [3,7,10,62,63]. Managing pain effectively is a priority in palliative care, and fears related to pain, particularly fear of patients dying in pain, can exacerbate caregiver distress [63]. Faster access to medications via drones could reduce these fears and feelings of helplessness [21-23,32]. However, these anticipated benefits remain hypothetical and require further empirical testing in real-world palliative care contexts.

Practical and Organizational Factors for Implementation of Drone-Assisted Delivery

Several conditions emerged as critical for successful implementation. Integration into established SOPC workflows was considered essential to maintain trusted communication channels and ensure that drones complemented rather than replaced conventional courier services [53,57]. Strengthening collaboration with SOPC teams is crucial, as caregivers view these professionals as trusted partners for both emotional and technical support. Participants emphasized that the ordering process should be as clear and simple as possible, ideally with the SOPC team initiating medication orders and individual orders placed only when necessary. They valued having a single point of contact for all medication-related information (eg, within the app) and suggested that drones could also transport essential documents, such as transport tickets. Direct communication options, particularly via telephone, with SOPC teams, general practitioners, and pharmacies were seen as indispensable, while chat functions were dismissed as impersonal and lacking advisory value. Direct communication was also the most important aspect of the cocreative process for developing a delivery app for drone-assisted systems [31]. However, in the present study, participants were, on average, 18 years older (mean age 68.6,SD 12.3) and preferred not to communicate with the SOPC via app or chat. Moreover, multilingual app interfaces were recommended to accommodate caregivers with limited German proficiency.

In terms of logistics, caregivers expressed a preference for having the option to choose between courier and drone delivery, with flexible time slots (eg, 30-minute windows) for urgent deliveries. Involving nursing services and enabling direct delivery to SOPC facilities were seen as beneficial. While many participants found it difficult to imagine the operational aspects of drones, they recognized their potential for rural areas with limited infrastructure.

Environmental and technical challenges, such as weather conditions, potential mechanical failure, and flight range limitations, were viewed as possible constraints. Data protection, noise disturbance, and regulatory and liability issues were additional concerns that would need to be resolved before deployment [26,31]. Moreover, unequal access to necessary digital tools could exacerbate disparities among caregivers and patients from different socio-economic or geographic backgrounds [31]. These findings highlight that technological readiness alone is insufficient; social, infrastructural, and regulatory readiness must also be addressed before implementation [25,55].

Study Limitations

Although the sample reflects typical age and gender distributions, it is not representative due to the small number of participants. The recruitment process, conducted by SOPC staff, may have introduced a positive selection bias. Caregivers with an established trust relationship with the SOPC teams may have been more inclined to participate and potentially more open to innovative care concepts, including drone delivery. This could have led to an overrepresentation of participants who were generally supportive of new interventions. We also note that such participants may have been more motivated to engage in research activities, which could have influenced the generally positive reception of the concept. While drones in medical delivery were largely unknown to all participants, reducing the likelihood that only technology enthusiasts were included, clear rejections of drone use were rare. However, limited knowledge restricted the depth of the discussion. Nevertheless, data saturation can be assumed, as previous studies have also shown low awareness of drones in health care, likely due to the novelty of the technology and lack of discourse [31]. Despite this, the discussions revealed a clear need for drone delivery, particularly through the detailed accounts of caregivers' challenges. This supports the relevance of the topic, even with limited previous knowledge. Another limitation concerns the presence of SOPC team members during data collection, one in the interviews and 2 in the focus group. Their involvement was intended to provide emotional support for participants, who were within 6 months of bereavement. However, the SPOC involvement may have introduced social desirability bias, as SOPC staff were also involved in recruitment. Participants might have felt inclined to provide answers they believed aligned with SOPC expectations. To minimize this risk, SOPC staff did not actively contribute to the discussions, and all facilitation was conducted by the research team. However, the possibility that their presence influenced participants' responses cannot be fully excluded. All FGs were conducted in German; to preserve meaning, translations were cross-checked. Minor linguistic shifts were adjusted through a process of iterative translation.

Clinical Implications

This study reveals several clinical implications. First, integrating eHealth can help ensure timely medication supply during critical care gaps, such as evenings and weekends. Drone-based delivery could reduce unnecessary hospitalizations and support patients' preference to remain at home. Second, SOPC teams play a central role in facilitating eHealth use. Caregivers view them as essential for technical guidance, emotional support, and confidence in using new technologies. Their involvement in training and communication is therefore crucial. Third, strengthening digital health literacy should be a structural task. The study shows that a lack of eHealth literacy is a key obstacle to accepting and effectively using drone-based health care

XSL•FO
RenderX

solutions. Therefore, empowerment measures should focus not only on individual training but also on structural adjustments that better adapt technological offerings to the abilities and needs of target groups, such as older relatives and people with little affinity for technology. Fourth, technology must not replace human connection. Concerns about losing personal interaction in palliative care remain significant. Digital tools should complement in-person care and include communication features that maintain links with SOPC teams. Fifth, improving the quality of life through faster access to medication may ease pain and reduce caregiver stress. Drone delivery may help close critical supply gaps, a potential that future research should explore further. Ultimately, to foster acceptance, caregivers need hands-on experience with drone-based services to understand their practical value.

Conclusions

This study highlights both opportunities and challenges in implementing drone-based medication delivery in palliative care. Caregivers perceived drones as a promising solution for timely access to essential medication, particularly during critical care gaps such as evenings and weekends. It remains open for further research whether the integration of eHealth solutions like drone-assisted delivery could help reduce unnecessary hospitalizations to support patients' preference to remain at home. Successful adoption; however, requires more than technical feasibility. Low digital health literacy, especially among older caregivers, remains a key barrier. Effective eHealth must align with users' abilities, offering simple interfaces and integration into trusted systems like SOPC teams, who play a central role in providing technical guidance, emotional support, and confidence in new technologies. Their involvement in both communication and training processes is therefore critical for successful adoption. Strengthening digital health literacy should be addressed not only through individual training but also via structural adjustments that better adapt technology to the needs of older caregivers and those with limited affinity for digital tools.

Improving timely medication access could also reduce caregiver stress and improve patients' quality of life by supporting symptom control at home, which remains open for further research. Ultimately, to foster acceptance, caregivers need hands-on experience with drone-based services to understand their practical value. Future research should focus on real-world testing in living lab environments, evaluating feasibility, effectiveness, and integration into existing clinical workflows while ensuring that technological, psychological, and social needs are equally addressed.

Acknowledgments

We would like to thank the PalliDrone team for the exchange and discussion of the focus group guide and for evaluating the results of focus groups. A special thank you goes out to the "Leipziger Palliativgesellschaft" in Torgau, Germany, and the "Anhalt-Hospiz Dessau" for their help with participant recruitment. Generative artificial intelligence (DeepL Write) was used solely to assist in reviewing and improving the correctness of some English-language expressions in the manuscript. All content, interpretation of data, and conclusions were fully authored and verified by the study team. We acknowledge the financial support of the Open Access Publication Fund of Martin-Luther-University Halle-Wittenberg. This research was funded by the Federal Ministry of Education and Research Germany through the "TDG—Translational Region for Digitalized Healthcare" project 03WIR3128A. PJ received this funding. The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Authors' Contributions

Conceptualization: FF, PJ, and MG
Data Curation: SCO
Formal analysis: FF and AL
Funding acquisition: FF and PF
Investigation: FF, AL, and MG
Methodology: FF
Project administration: FF
Supervision: PJ
Visualization: FF and SCO;
Writing – original draft: FF
Writing – review & editing: PJ

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Interview guide.
[DOCX File, 39 KB - humanfactors_v12i1e80320_app1.docx]

Checklist 1
COREQ checklist.
[PDF File, 118 KB - humanfactors_v12i1e80320_app2.pdf]

References

1. Stahmeyer JT, Geyer S, Epping J, Tetzlaff J, Eberhard S. Healthcare expenditures and the role of demografic change: an analysis of statutory health insurance data. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2018 Apr;61(4):432-441. [doi: 10.1007/s00103-018-2713-3] [Medline: 29487976]
2. Vallin J, Meslé F. Convergences and divergences in mortality: a new approach of health transition. Demogr Res 2004;S2:11-44. [doi: 10.4054/DemRes.2004.S2.2]
3. Ann-Yi S, Bruera E. Psychological aspects of care in cancer patients in the last weeks/days of life. Cancer Res Treat 2022 Jul;54(3):651-660. [doi: 10.4143/crt.2022.116] [Medline: 35790196]
4. Carter GL, Lewin TJ, Gianacas L, Clover K, Adams C. Caregiver satisfaction with out-patient oncology services: utility of the FAMCARE instrument and development of the FAMCARE-6. Support Care Cancer 2011 Apr;19(4):565-572. [doi: 10.1007/s00520-010-0858-1] [Medline: 20349317]
5. Zich K, Sydow H. Palliativversorgung: sterbeort krankenhaus - regionale unterschiede und einflussfaktoren [Report in German]. : Bertelsmann Stiftung; 2015.
6. Lerzynski GA, Allan A, Murray SA. Die Bewertung der palliativmedizinischen Patientenversorgung mithilfe der Palliative Care Outcome Scale (POS) in verschiedenen Versorgungsformen [Article in German]. Palliativmedizin 2004 Mar;5(1):19-27. [doi: 10.1055/s-2003-814874]
7. Martín JM, Olano-Lizarraga M, Saracíbar-Razquin M. The experience of family caregivers caring for a terminal patient at home: a research review. Int J Nurs Stud 2016 Dec;64:1-12. [doi: 10.1016/j.ijnurstu.2016.09.010] [Medline: 27657662]
8. Bijnsdorp FM, Onwuteaka-Philipsen BD, Boot CRL, van der Beek AJ, Pasman HRW. Caregiver's burden at the end of life of their loved one: insights from a longitudinal qualitative study among working family caregivers. BMC Palliat Care 2022 Aug 10;21(1):142. [doi: 10.1186/s12904-022-01031-1] [Medline: 35945558]
9. Alam S, Hannon B, Zimmermann C. Palliative care for family caregivers. J Clin Oncol 2020 Mar 20;38(9):926-936. [doi: 10.1200/JCO.19.00018] [Medline: 32023152]
10. Murray SA, Kendall M, Boyd K, Grant L, Highet G, Sheikh A. Archetypal trajectories of social, psychological, and spiritual wellbeing and distress in family care givers of patients with lung cancer: secondary analysis of serial qualitative interviews. BMJ 2010 Jun 9;340:c2581. [doi: 10.1136/bmj.c2581] [Medline: 20538635]
11. Johnston EE, Rosenberg AR. Palliative care in adolescents and young adults with cancer. J Clin Oncol 2024 Feb 20;42(6):755-763. [doi: 10.1200/JCO.23.00709] [Medline: 37862672]
12. Steins M, Eschbach C, Villalobos M, Thomas M. Schmerztherapie in der palliativmedizin. Palliativmedizin 2018 Jan;19(1):47-56. [doi: 10.1055/s-0043-122667]
13. Leitlinienprogramm onkologie: palliativmedizin für patienten mit einer nicht heilbaren krebserkrankung [Book in German]. In: Langversion 2.2: Deutsche Krebsgesellschaft, Deutsche Krebshilfe, and AWMF; 2020.
14. Hiebert B, Nouvet E, Jeyabalan V, Donelle L. The application of drones in healthcare and health-related services in North America: a scoping review. Drones 2020;4(3):30. [doi: 10.3390/drones4030030]
15. Rosser JC, Vignesh V, Terwilliger BA, Parker BC. Surgical and medical applications of drones: a comprehensive review. JSLS 2018;22(3):e2018.00018. [doi: 10.4293/JSLS.2018.00018] [Medline: 30356360]
16. Balasingam M. Drones in medicine-the rise of the machines. Int J Clin Pract 2017 Sep;71(9). [doi: 10.1111/ijcp.12989] [Medline: 28851081]
17. Bhattacharya S, Hossain MM, Hoedebecke K, Bacorro M, Gökdemir Ö, Singh A. Leveraging unmanned aerial vehicle technology to improve public health practice: prospects and barriers. Indian J Community Med 2020;45(4):396-398. [doi: 10.4103/ijcm.IJCM_402_19] [Medline: 33623188]
18. Laksham KB. Unmanned aerial vehicle (drones) in public health: a SWOT analysis. J Family Med Prim Care 2019 Feb;8(2):342-346. [doi: 10.4103/jfmpc.jfmpc_413_18] [Medline: 30984635]
19. Poljak M, Šterbenc A. Use of drones in clinical microbiology and infectious diseases: current status, challenges and barriers. Clin Microbiol Infect 2020 Apr;26(4):425-430. [doi: 10.1016/j.cmi.2019.09.014]
20. Scalea JR, Pucciarella T, Talaie T, et al. Successful implementation of unmanned aircraft use for delivery of a human organ for transplantation. Ann Surg 2021 Sep 1;274(3):e282-e288. [doi: 10.1097/SLA.0000000000003630] [Medline: 31663974]
21. Claesson A, Fredman D, Svensson L, et al. Unmanned aerial vehicles (drones) in out-of-hospital-cardiac-arrest. Scand J Trauma Resusc Emerg Med 2016 Oct 12;24(1):124. [doi: 10.1186/s13049-016-0313-5] [Medline: 27729058]

XSL•FO
RenderX

22. Rosamond WD, Johnson AM, Bogle BM, et al. Drone delivery of an automated external defibrillator. N Engl J Med 2020 Sep 17;383(12):1186-1188. [doi: 10.1056/NEJMc1915956] [Medline: 32937053]
23. Sanfridsson J, Sparrevik J, Hollenberg J, et al. Drone delivery of an automated external defibrillator - a mixed method simulation study of bystander experience. Scand J Trauma Resusc Emerg Med 2019 Apr 8;27(1):40. [doi: 10.1186/s13049-019-0622-6] [Medline: 30961651]
24. Amukele TK, Hernandez J, Snozek CLH, et al. Drone transport of chemistry and hematology samples over long distances. Am J Clin Pathol 2017 Nov 2;148(5):427-435. [doi: 10.1093/ajcp/aqx090] [Medline: 29016811]
25. Rejeb A, Rejeb K, Simske S, Treiblmaier H. Humanitarian drones: a review and research agenda. Internet of Things 2021 Dec;16:100434. [doi: 10.1016/j.iot.2021.100434]
26. Rejeb A, Rejeb K, Simske SJ, Treiblmaier H. Drones for supply chain management and logistics: a review and research agenda. Int J Logist Res Appl 2023 Jun 3;26(6):708-731. [doi: 10.1080/13675567.2021.1981273]
27. Wang N, Christen M, Hunt M. Ethical considerations associated with “Humanitarian Drones”: a scoping literature review. Sci Eng Ethics 2021 Aug 3;27(4):51. [doi: 10.1007/s11948-021-00327-4] [Medline: 34342721]
28. Wang N. “Killing two birds with one stone”? a case study of development use of drones. Presented at: 2020 IEEE International Symposium on Technology and Society (ISTAS); Nov 12-15, 2020; Tempe, AZ, USA p. 2020. [doi: 10.1109/ISTAS50296.2020.9462187]
29. Stephan F, Reinsperger N, Grünthal M, Paulicke D, Jahn P. Human drone interaction in delivery of medical supplies: a scoping review of experimental studies. PLoS One 2022;17(4):e0267664. [doi: 10.1371/journal.pone.0267664] [Medline: 35482656]
30. Fink F, Kalter I, Steindorff JV, Helmbold HK, Paulicke D, Jahn P. Identifying factors of user acceptance of a drone-based medication delivery: user-centered design approach. JMIR Hum Factors 2024 Apr 30;11:e51587. [doi: 10.2196/51587] [Medline: 38687589]
31. Fink F, Paulicke D, Grünthal M, Jahn P. “Of course, drones delivering urgent medicines are necessary. But I would not use them until…” Insights from a qualitative study on users’ needs and requirements regarding the use of medical drones. PLoS One 2023;18(5):e0285393. [doi: 10.1371/journal.pone.0285393] [Medline: 37155650]
32. Zègre-Hemsey JK, Grewe ME, Johnson AM, et al. Delivery of automated external defibrillators via drones in simulated cardiac arrest: users’ experiences and the human-drone interaction. Resuscitation 2020 Dec;157:83-88. [doi: 10.1016/j.resuscitation.2020.10.006] [Medline: 33080371]
33. Mateen FJ, Leung KHB, Vogel AC, Cissé AF, Chan TCY. A drone delivery network for antiepileptic drugs: a framework and modelling case study in a low-income country. Trans R Soc Trop Med Hyg 2020 Apr 8;114(4):308-314. [doi: 10.1093/trstmh/trz131] [Medline: 31943110]
34. Pulver A, Wei R, Mann C. Locating AED enabled medical drones to enhance cardiac arrest response times. Prehosp Emerg Care 2016;20(3):378-389. [doi: 10.3109/10903127.2015.1115932] [Medline: 26852822]
35. Lehmann A, Kalter I, Jahn P, Fink F. A user journey: development of drone-based medication delivery—meeting developers and co-developers’ expectations. Designs 2025;9(2):27. [doi: 10.3390/designs9020027]
36. Farao J, Malila B, Conrad N, Mutsvangwa T, Rangaka MX, Douglas TS. A user-centred design framework for mHealth. PLoS One 2020;15(8):e0237910. [doi: 10.1371/journal.pone.0237910] [Medline: 32813711]
37. Konstantinidis EI, Petsani D, Bamidis PD. Teaching university students co-creation and living lab methodologies through experiential learning activities and preparing them for RRI. Health Informatics J 2021;27(1):1460458221991204. [doi: 10.1177/1460458221991204] [Medline: 33535859]
38. Skivington K, Matthews L, Simpson SA, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 2021 Sep 30;374:n2061. [doi: 10.1136/bmj.n2061] [Medline: 34593508]
39. Helfferich C. Die Qualität Qualitativer Daten [Book in German]: VS Verlag für Sozialwissenschaften; 2011.
40. Schulz M, Mack B, Renn O. Fokusgruppen in Der Empirischen Sozialwissenschaft [Book in German]: VS Verlag für Sozialwissenschaften; 2012.
41. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
42. Venkatesh V, Bala H. Technology acceptance model 3 and a research agenda on interventions. DECISION Sciences 2008 May;39(2):273-315 [FREE Full text] [doi: 10.1111/j.1540-5915.2008.00192.x]
43. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manage Sci 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
44. Kothgassner OD, Felnhofer A, Hauk N, Kastenhofer E, Gomm J, Kryspin-Exner I. TUI: Technology Usage Inventory 2012. URL: https://www.ffg.at/sites/default/files/allgemeine_downloads/thematische%20programme/programmdokumente/tui_manual.pdf?utm_source=chatgpt.com [accessed 2025-10-24]
45. Peters D, Calvo RA, Ryan RM. Designing for motivation, engagement and wellbeing in digital experience. Front Psychol 2018;9:797. [doi: 10.3389/fpsyg.2018.00797] [Medline: 29892246]
46. Przyborski A, Wohlrab-Sahr M. Qualitative Sozialforschung [Book in German]: De Gruyter; 2021.
47. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]

XSL•FO
RenderX

48. Gläser J, Laudel G. Experteninterviews Und Qualitative Inhaltsanalyse: Als Instrumente Rekonstruierender Untersuchungen [Book in German], 4th edition: VS Verl.: Wiesbaden; 2010.
49. Denney R. Review of Content Analysis in Communication Research, by B. Berelson. Audio Visual Communication Review 1954;2(1):64-67 [FREE Full text]
50. Winsler A, Fernyhough C, Mcclaren E, Way E. Private Speech Coding Manual: George Mason University; 2005.
51. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 1960 Apr;20(1):37-46. [doi: 10.1177/001316446002000104]
52. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174. [doi: 10.2307/2529310] [Medline: 843571]
53. Steindal SA, Nes AAG, Godskesen TE, et al. Advantages and challenges of using telehealth for home-based palliative care: systematic mixed studies review. J Med Internet Res 2023 Mar 13;25:e43684. [doi: 10.2196/43684] [Medline: 36912876]
54. Amukele T. Current state of drones in healthcare: challenges and opportunities. J Appl Lab Med 2019 Sep;4(2):296-298. [doi: 10.1373/jalm.2019.030106] [Medline: 31639681]
55. Bachmann A. Digitale gesundheitskompetenz: eine kurze übersicht zur aktuellen literatur: schweiz [Book in German]. : Bundesamt für Gesundheit; 2019.
56. Phongtankuel V, Meador L, Adelman RD, et al. Multicomponent palliative care interventions in advanced chronic diseases: a systematic review. Am J Hosp Palliat Care 2018 Jan;35(1):173-183. [doi: 10.1177/1049909116674669] [Medline: 28273750]
57. Holmen H, Riiser K, Winger A. Home-based pediatric palliative care and electronic health: systematic mixed methods review. J Med Internet Res 2020 Feb 28;22(2):e16248. [doi: 10.2196/16248] [Medline: 32130127]
58. Phongtankuel V, Adelman RD, Reid MC. Mobile health technology and home hospice care: promise and pitfalls. Prog Palliat Care 2018;26(3):137-141. [doi: 10.1080/09699260.2018.1467109] [Medline: 30505077]
59. Widberg C, Wiklund B, Klarare A. Patients' experiences of eHealth in palliative care: an integrative review. BMC Palliat Care 2020 Oct 14;19(1):158. [doi: 10.1186/s12904-020-00667-1] [Medline: 33054746]
60. Eißfeldt H, Vogelpohl V, Stolz M, et al. The acceptance of civil drones in Germany. CEAS Aeronaut J 2020 Sep;11(3):665-676. [doi: 10.1007/s13272-020-00447-w]
61. Tanco K, Prado B, Qian Y, et al. A comparison of caregiver burden of patients with advanced cancer in different palliative cancer care settings. J Palliat Med 2021 Dec;24(12):1766-1775. [doi: 10.1089/jpm.2021.0026] [Medline: 33926226]
62. Wadhwa D, Burman D, Swami N, Rodin G, Lo C, Zimmermann C. Quality of life and mental health in caregivers of outpatients with advanced cancer. Psychooncology 2013 Feb;22(2):403-410. [doi: 10.1002/pon.2104] [Medline: 22135229]
63. Walbaum C, Philipp R, Bokemeyer C, et al. Death anxiety in patients with advanced cancer and their family caregivers. J Pain Symptom Manage 2024 Dec;68(6):622-631. [doi: 10.1016/j.jpainsymman.2024.08.027] [Medline: 39182590]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
FG: focus group
SOPC: specialized outpatient palliative care
TAM: technology acceptance model

Edited by A Kushniruk; submitted 08.07.25; peer-reviewed by E Oluokun, P Onukak; revised version received 15.08.25; accepted 20.09.25; published 13.11.25.

Please cite as:
Fink F, Otto SC, Grünthal M, Lehmann A, Jahn P
Family Caregivers' Perspectives on the Potential of Drone-Based Medication Delivery in Palliative Home Care: Qualitative Focus Group Study
JMIR Hum Factors 2025;12:e80320
URL: https://humanfactors.jmir.org/2025/1/e80320
doi:10.2196/80320

© Franziska Fink, Sabrina Claudia Otto, Martin Grünthal, Anne Lehmann, Patrick Jahn. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 13.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Capturing Movement Behaviors in Latinas: Feasibility, Validity, and Acceptability Study of an Ecological Momentary Assessment Protocol

Jaclyn P Maher[1*], PhD; Peyton A Greco[1*], MS; Eugenia Camacho Fernandez[2], MPH; Brynn L Hudgins[1], PhD; Sandra E Echeverria[2], MPH, PhD

[1]Department of Kinesiology, University of North Carolina at Greensboro, 1408 Walker Ave, Greensboro, NC, United States

[2]Department of Public Health Education, University of North Carolina at Greensboro, Greensboro, NC, United States

[*]these authors contributed equally

Corresponding Author:
Jaclyn P Maher, PhD
Department of Kinesiology, University of North Carolina at Greensboro, 1408 Walker Ave, Greensboro, NC, United States

Abstract

Background: Latinas are one of the largest and fastest-growing female ethnic groups in the United States and have high levels of physical inactivity and sedentary behavior (SB), contributing to a disproportionate burden of chronic health conditions. An ecological momentary assessment (EMA) involves the use of smartphone-based data collected in real time to assess health behaviors and outcomes.

Objective: We examined the feasibility, validity, and acceptability of an EMA protocol assessing physical activity (PA) and SB in Latina adults.

Methods: For 7 days, 67 Latinas (average age 39 years, SD = 13.6; n=37, 55.2% earning less than US $50,000/year; n=53, 79.1% foreign-born; and n=49, 73.1% of Mexican or Mexican American origin) completed a signal-contingent EMA protocol with 3 prompts per day and wore an ActiGraph GT3X accelerometer to measure levels of PA and SB. EMA prompts inquired about current behavior, feelings, beliefs, social conditions, and contexts.

Results: Latinas completed 69.7% (892/1279) of EMA prompts. They were more likely to respond to EMA prompts when engaged in more SB (odds ratio [OR] 1.04, 95% CI 1.01-1.06) and less light-intensity PA (OR 0.97, 95% CI 0.94-0.99) in the 30 minutes around the prompt. Accelerometer data validated self-reported occasions of PA and SB via EMA. The majority of participants (>70%) were satisfied with the protocol and expressed interest in participating in future studies.

Conclusions: EMA is a feasible, valid, and acceptable methodology for capturing movement behaviors among Latinas, which can provide insights into the antecedents and consequences of these behaviors in their daily lives.

(JMIR Hum Factors 2025;12:e75855) doi:10.2196/75855

KEYWORDS

experience sampling; ambulatory assessment; Hispanic; physical activity; exercise; sitting; accelerometer; women

Introduction

Latina or Hispanic adult women (hereafter, “Latinas”) are one of the largest, fastest-growing female ethnic groups in the United States and engage in lower levels of physical activity (PA) compared to Latino men and non-Latina women [1,2]. Recent estimates suggest that one-third of Latinas in the United States report no moderate-to-vigorous-intensity physical activity (MVPA; outside of work), the highest of any racial or ethnic group [3]. Conversely, Latinas tend to engage in more light-intensity physical activity (LPA) and less sedentary behavior (SB) compared to their White and Black counterparts [4,5]. The unique movement behavior profiles among Latinas likely result from unique barriers and facilitators that Latinas encounter in their daily lives. Yet, our understanding of Latinas’ engagement in movement behaviors (ie, MVPA, LPA, and SB) and determinants of those behaviors in naturalistic settings is limited, hampering efforts to effectively intervene [6]. Therefore, it is not surprising that Latinas in the United States experience disproportionately high rates of chronic conditions closely linked to low levels of MVPA such as diabetes, hypertension, obesity, and certain cancers.

The ecological momentary assessment (EMA) has increased in popularity over the last 15 years in PA research as a methodological approach to better capture health behaviors and their determinants as they occur in daily life [7]. Specifically, EMA is a real-time data capture methodology designed to intensively and repeatedly assess participants’ behaviors,

XSL•FO
RenderX

experiences, feelings, and contexts surrounding a phenomenon of interest [8]. Such an approach can improve ecological validity and reduce recall errors [9], ultimately improving our understanding of when, where, with whom, and why movement behaviors occur.

Despite the potential advantages of this methodological approach, no studies have explicitly evaluated the feasibility, validity, and acceptability of EMA to capture movement behaviors in Latinas. Latinas are likely to face unique barriers to participation in research involving methodologies such as EMA given documented issues related to lack of trust in research, limited rapport with research staff, and cultural insensitivity of research to the unique values, traditions, and beliefs of the Latino/a community [10,11]. Other issues associated with broader research participation among Latinas include concerns about immigration or undocumented status, lack of access to Spanish-speaking staff or interpreters, and lack of childcare [10,11].

Therefore, there is a need to establish the feasibility, validity, and acceptability of a smartphone-based EMA study to assess movement behaviors in Latinas. In this study, we examined the feasibility, validity, and acceptability of a 7-day EMA protocol assessing PA and SB. To determine feasibility, we examined rates of compliance with elements of the study protocol (ie, EMA, accelerometer wear time). EMA studies across a variety of designs, samples, and research fields have reported an average compliance with EMA protocols of 79% [12]. However, given that EMA studies focused on movement behaviors in other samples of minoritized women have reported lower compliance [13] balanced with the need to capture representative and ecologically valid data on participants' daily experiences, we deemed a compliance rate of 70% to be acceptable. We also examined measurement reactivity to determine whether accelerometer-derived behavioral data were affected or changed because of answering the EMA prompts. With respect to validity, we investigated the degree to which EMA-reported behaviors corresponded with the amount of MVPA, LPA, and SB captured by the accelerometer over that same time. To our knowledge, we are not aware of any studies that have investigated compliance with, potential reactivity to, or validity of EMA reports with regard to LPA; however, given the unique movement behavior profiles of Latinas with high amounts of LPA and accumulating evidence regarding the health benefits of LPA [3-5,14], this is an important question to investigate. Finally, to establish acceptability, we used end-of-study feedback on specific elements of the study. This formative research is crucial for enhancing our understanding of the factors influencing movement behaviors and for developing personally tailored and context-sensitive movement-related interventions for Latinas.

Methods

Participants and Recruitment

Latinas living in Guilford County, North Carolina, and surrounding counties were recruited for participation in this study. Recruitment was done primarily through community centers serving predominantly Latina/o populations. Through conversations with community partners, effective methods for recruitment were identified. Print materials were placed at community locations. Further, announcements were made and recruitment tables with staff were present at specific community events (eg, English as a second language classes, soccer games, and family services). Community partners also communicated information about the study to organization patrons. Women were eligible if they were aged ≥18 years, self-identified as Latina or Hispanic, and had no limitations that would prevent engaging in PA. If women did not speak and read English or Spanish, were pregnant, or were unwilling to use their personal smartphone for the study, they were deemed ineligible to participate. We hired and trained bilingual, bicultural research staff to recruit, enroll, and check in with participants over the course of the study.

Procedures

Following eligibility screening, Latinas attended the introductory training session (day 1) at one of three community locations. At the training session, a research staff member provided a broad overview of the study (eg, procedures, risks or benefits, and compensation). After the overview, individuals were asked to review the consent form (written in English) and sign to indicate their willingness to participate in the study. Research staff indicated that they could read the consent form to participants, if desired. Although we did not collect specific records of whether participants read the consent themselves or had the staff member read it to them, based on anecdotal staff reports, most participants read the consent form themselves. Research staff were available to answer any questions individuals had. Once participants consented, they completed a baseline questionnaire to provide basic demographic and health information. Participants were also familiarized with an accelerometer, which was to be worn on the hip during all waking hours unless they were showering, bathing, or swimming. Research staff also provided more details on the EMA study procedures during the training session and assisted participants in downloading the smartphone app, LifeData, onto their personal smartphones. The app delivered the signal-contingent EMA protocol. The EMA protocol delivered a questionnaire at a random time within three 2-hour windows (ie, morning: 6:30-8:30 AM; afternoon: 12-2 PM; evening: 6-8 PM). The EMA surveys took 2 - 3 minutes to complete, and participants could receive between 29 and 35 items at each prompt depending on participant responses at the time of the prompt (eg, if they were at work and if they were alone). Screenshots of all EMA items can be found in Figure S1 in Multimedia Appendix 1. When reviewing the EMA items during the training sessions, as well as at the beginning of each EMA prompt, participants were presented with definitions of key constructs to help orient participants to what was meant by specific terms that appeared in the questions. For instance, participants were provided a definition of physical activity that aligned with MVPA (eg, moderate to hard physical effort and breathing harder than normal). The EMA protocol was available in English and Spanish and continued for 7 consecutive days.

After the delivery of an EMA prompt, participants had 30 minutes to answer the EMA. After 30 minutes, if the EMA prompt was still unanswered, the questionnaire became

XSL•FO
RenderX

inaccessible and was counted as missing. Research staff monitored compliance in real time and contacted all participants 2‐3 days into the EMA protocol to check in and address any issues related to questionnaire delivery, compliance, etc.

Following the 7-day study period, participants returned the study equipment at a community location and were compensated for their time. Community locations were also compensated for granting the research team space to meet with participants. Finally, participants were asked to complete a brief questionnaire to gather feedback about their experience in the study.

Measures

EMA-Reported Physical Activity and Sedentary Behavior

At each EMA prompt, participants were asked "What were you doing just before the phone went off?" and were given the option to select all responses that were applicable (eg, watching television or movies; using a computer, tablet, or phone; physical activity or exercise; running errands; and household chores [15-17]; see Multimedia Appendix 1 for a complete list of response options). Given our interest to determine the validity between self-reported PA and device-measured PA, occasions in which participants selected "physical activity/exercise" at the time of the prompt were classified as currently engaged in PA at that occasion. Responses were coded so that occasions in which PA was not reported were the reference (coded as 0) and occasions in which PA was reported were the comparison (coded as 1). Aside from occasions where participants reported engaging in "physical activity/exercise," they were asked if they were sitting while doing their current activity. We classified participants as engaging in SB at the time of the prompt if they reported sitting while doing their current activity. Responses were coded so that occasions where sitting was not reported were the reference (coded as 0) and occasions where sitting was reported were the comparison (coded as 1). Regardless of whether participants indicated they were currently engaged in PA or SB, they were not asked to report the duration of their current PA or SB.

Device-Based Physical Activity and Sedentary Behavior

The ActiGraph GT3X-BT accelerometer measured participants' PA and SB. Data from this device were downloaded in 30-second epochs, and movement behaviors were operationalized as minutes of MVPA (ie, ≥2020 counts/minute), LPA (ie, 100‐2019 counts/minute), and SB (<100 counts/minute) [18,19]. Time stamps from the device and EMA questionnaires were used to pair the 2 data sources and create windows of accelerometer data before, after, or around the EMA prompt. Movement behavior data accumulated in the 15 minutes before, 15 minutes after, or 30-minute window around (ie, ±15 minutes) the EMA prompt was used in analyses. Nonwear time was determined based on consecutive minutes of 0 activity counts [19]. Time intervals in which valid wear time amounted to at least half of the interval were considered valid and included in the analysis (eg, ≥7.5 minutes of wear in the 15 minutes before the prompt).

Time

We created variables representing the time of day and day of the week EMA prompts were delivered. EMA prompts delivered between 6:30 and 8:30 AM were classified as morning prompts, 12 and 2 PM as afternoon prompts, and 6 and 8 PM as evening prompts. Prompts delivered from Monday to Friday were classified as weekday prompts, and prompts delivered on Saturday and Sunday were classified as weekend prompts. Finally, an EMA prompt number variable (ranging between 1 and 21, as this is the maximum number of prompts that could be delivered as part of the study) was created to account for time in study.

Demographic, Family, Work, Social, and Structural Determinants

During the introductory training session, participants completed a baseline questionnaire that assessed demographic and health information including age, nativity, and self-reported health. Although not relevant to the aims of this manuscript, at each EMA prompt, participants reported on family, work, and social conditions and potential stressors they were experiencing. Latinas' geographic location at the time of the prompt was also recorded at each EMA prompt.

Study Feedback

At the end of the study, participants provided their perceptions of study procedures, their overall experience, and their willingness to complete smartphone-based or accelerometer-related study procedures in future studies. Items and response options are provided in the Results section.

Data Analysis

Our analytic approach was informed by previous work investigating the feasibility and validity of EMA in PA-related research in other populations [15-17]. As an intensive longitudinal study, we fit a series of multilevel models to account for the nested structure of our data (ie, observations nested within people) to investigate the aims previously defined [20,21]. Descriptive statistics and frequencies were calculated to describe compliance (ie, total answered prompts completed divided by total delivered prompts).

To determine if engagement in movement behaviors influenced the likelihood of completing an EMA prompt, we fit a series of multilevel logistic regression models. More specifically, these models examined if engaging in movement behaviors (ie, MVPA, LPA, and SB in the 30-minute window [±15 minutes] around the EMA prompt) as well as temporal (ie, time of day, day of week, and time in study) or demographic factors (ie, age, preferred language, self-reported health, nativity status, and employment status) influenced whether or not an EMA prompt was completed. EMA prompts were dummy coded (ie, prompts completed were coded as 1; prompts not completed were coded as 0). Prompts were considered not completed if a participant did not open the EMA questionnaire within 30 minutes of delivery or the prompt was answered but the participant did not answer all questions included in the EMA questionnaire (ie, did not provide complete data at the EMA prompt).

XSL•FO
RenderX

The next set of models examined whether the time spent engaging in PA (or SB) in the 15-minute window before the EMA prompt differed from the time spent engaging in PA (or SB) in the 15-minute window after the EMA prompt. This is relevant as it provides insights into whether the EMA prompt resulted in any reactivity (or changes in PA or SB as a result of being prompted about their movement behaviors) [22,23]. We created a dummy-coded variable representing the 15 minutes prior ("window"=0) and after the EMA prompt ("window"=1) to examine its association with MVPA, LPA, and SB (one movement behavior per model).

Our final set of models examined criterion validity. We fit multilevel linear regression models to determine associations between EMA-reported PA and SB and accelerometer-derived PA and SB, respectively. We used EMA-reported PA (or SB) at the time of the prompt as the main predictor to determine associations with accelerometer-derived MVPA, LPA, or SB (one movement behavior per model). For all models in which MVPA was the outcome, model results were consistent regardless of using raw or transformed MVPA variables, so we opted to use raw MVPA values to aid in the interpretability of findings.

Based on a priori selection of covariates and prior studies, we adjusted all multilevel models for time and day of the prompt, time in study (ie, EMA prompt number), age (grand mean centered), employment status (dummy coded), language of questionnaire (dummy coded), nativity (dummy coded), and self-reported health (eg, [3,5,6,12,15,23-25]). Previous simulation studies indicate that having at least 50 units at the high order level (in our case, participants at level 2) with between 5 and 30 observations per unit (in our case, we averaged 10 valid observations per participant) is sufficient to generate accurate unstandardized coefficient and standard error estimates [26]. Quantitative analyses were conducted in Stata v. 18 and SAS v9.4.

Lastly, to establish acceptability, we evaluated reports of satisfaction with, convenience of, and disruptions due to study participation. In addition to Likert items assessing acceptability, participants were given the opportunity to provide additional feedback via an open-ended question. Only one-third (23/67, 34.3%) of the participants responded to the open-ended question, and responses were brief, with an average length of 10 words. Therefore, we did not conduct a formal qualitative analysis but highlighted quotes to represent common responses across respondents.

Ethical Considerations

All study procedures were approved by the University of North Carolina at Greensboro's institutional review board (IRB FY22-483). Written informed consent was obtained from all individual participants included in the study. Several measures were implemented to ensure participant privacy and confidentiality, including meeting with participants in private or secluded areas at community locations, using ID numbers to link participant data, and storing electronic or paper records in secure locations. Participants could earn up to US $50 in gift cards: US $20 for attending the introductory session and US $30 for completing the EMA and accelerometer protocol (not prorated based on compliance).

Results

Participant Characteristics and Completion of EMA Prompts and Accelerometer Wear

A total of 67 Latinas completed the study. Table 1 displays the demographic characteristics of study participants. Participants were, on average, 39 years of age (SD = 13.6). About half of the sample met physical activity guidelines and indicated they were from households earning less than US $50,000 per year. Approximately three-quarters (53/67, 79.1%) of the sample were foreign-born. Similarly, about three-quarters (49/67, 73.1%) of the sample were identified as Mexican or Mexican American.

XSL•FO
RenderX

Table . Demographic characteristics of study participants.

Participant characteristics	Values
Age (years) (n=67), n (%)	
18 - 44	44 (65.7)
45 - 64	21 (31.3)
≥65	2 (3.0)
Education (n=67), n (%)	
Some high school	18 (26.9)
High school or equivalent	23 (34.3)
Some college	12 (17.9)
College degree or postsecondary	12 (17.9)
Missing	2 (3.0)
Annual household income (US $) (n=67), n (%)	
<10,000	3 (4.5)
10,000 - 24,999	11 (16.4)
25,000 - 49,999	23 (34.3)
≥50,000	20 (29.9)
Do not know or not sure	10 (14.9)
Marriage status (n=67), n (%)	
Married or cohabitating	48 (71.6)
Single	12 (17.9)
Other	7 (10.5)
Ethnic group (n=67), n (%)	
Mexican or Mexican American	49 (73.1)
Central American	13 (19.4)
Other	5 (7.5)
Language spoken at home (n=67), n (%)	
Only Spanish	33 (49.2)
Mostly Spanish	15 (22.4)
Mostly English	2 (3.0)
Spanish and English equally	17 (25.4)
Self-reported health (n=67), n (%)	
Excellent	8 (11.9)
Very good	13 (19.4)
Good	27 (40.3)
Fair	19 (28.4)
Currently employed (n=67), n (%)	
Yes	41 (61.2)
Country of birth (n=67), n (%)	
United States	14 (20.9)
Mexico	40 (59.7)
Other	13 (19.4)
Physical activity (n=67), n (%)	

XSL•FO
RenderX

Participant characteristics	Values
Meeting physical activity guidelines (based on average daily minutes of MVPA)[a]	37 (55.2)
Occasions engaging in physical activity or exercise (n=907), n (%)	
EMA-based[b] report	35 (3.9)
Occasions engaging in sedentary behavior (n=901), n (%)	
EMA-based report	457 (50.7)
Moderate to vigorous physical activity (minutes), mean (SD)	
Device-based assessment (15 minutes prior to EMA prompt)	0.44 (1.42)
Light physical activity (minutes), mean (SD)	
Device-based assessment (15 minutes prior to EMA prompt)	5.01 (3.72)
Sedentary behavior (minutes), mean (SD)	
Device-based assessment (15 minutes prior to EMA prompt)	9.32 (4.21)

[a]MVPA: moderate to vigorous intensity physical activity.

[b]EMA: ecological momentary assessment.

There were 1279 EMA prompts delivered across all participants. Of the EMA prompts delivered, participants answered 927 (72.5%) EMA prompts and completed 892 (69.7%, range: 5% - 100%) EMA prompts. On approximately two-thirds of occasions, participants had valid actigraphy data in the 15 minutes before (n=842, 65.8%) or 15 minutes after (n=860, 67.2%) the EMA prompt. Frequencies and descriptive statistics for EMA-reported and device-based movement behaviors are displayed in Table 1.

Factors Influencing Compliance With Study Procedures

Table 2 displays associations between demographic, temporal, and behavioral factors and compliance. There was no difference in odds of participants completing the EMA prompt (vs not) depending on the amount of MVPA engaged in in the 30-minute window around the EMA prompt (P>.05). However, Latinas were more likely to complete an EMA prompt if they engaged in more SB (odds ratio [OR] 1.04, 95% CI 1.01-1.06) and less LPA (OR 0.97, 95% CI 0.94-0.99) in the 30-minute window around the EMA prompt.

XSL•FO
RenderX

Table . Odds of completing an EMA[a] prompt based on participant characteristics and current behaviors and temporal context.[b]

	Model 1: Multilevel logistic regression with MVPA[c] as key movement behavior, OR[d] (95% CI)	Model 2: Multilevel logistic regression with LPA[e] as key movement behavior, OR (95% CI)	Model 3: Multilevel logistic regression with SB[f] as key movement behavior, OR (95% CI)
Fixed effects			
Intercept	0.66 (0.17 - 2.54)	0.84 (0.21 - 3.35)	0.32 (0.08 - 1.36)
Movement behavior (±15 minutes)	0.95 (0.90 - 1.01)	0.97[g] (0.94 - 0.99)	1.04[g] (1.01 - 1.06)
Afternoon	1.83[g] (1.08 - 3.10)	1.85[g] (1.09 - 3.14)	1.74[g] (1.02 - 2.95)
Evening	2.51[g] (1.47 - 4.29)	2.36[g] (1.38 - 4.04)	2.22[g] (1.29 - 3.81)
Weekend day	0.66[g] (0.43 - 0.99)	0.67 (0.44 - 1.02)	0.66[g] (0.44 - 0.99)
Notification number	1.08[g] (1.04 - 1.12)	1.08[g] (1.04 - 1.12)	1.08[g] (1.04 - 1.12)
Age	1.01 (0.97 - 1.04)	1.01 (0.98 - 1.04)	1.01 (0.98 - 1.04)
EMA in Spanish	1.04 (0.38 - 2.90)	1.07 (0.38 - 3.00)	1.09 (0.39 - 3.05)
Currently working	1.19 (0.57 - 2.48)	1.21 (0.57 - 2.54)	1.22 (0.59 - 2.56)
US born	1.26 (0.38 - 4.13)	1.23 (0.37 - 4.09)	1.25 (0.38 - 4.08)
Self-reported health	1.25 (0.85 - 1.85)	1.27 (0.85 - 1.88)	1.26 (0.85 - 1.86)
Random effects			
Intercept	1.41 (0.78 - 2.56)	1.45 (0.81 - 2.63)	1.47 (0.77 - 2.56)

[a]EMA: ecological momentary assessment.

[b]All models were based on 67 participants with 850 observations.

[c]MVPA: moderate to vigorous intensity physical activity.

[d]OR: odds ratio.

[e]LPA: light intensity physical activity.

[f]SB: sedentary behavior.

[g]$P<.05$.

Latinas were more likely to complete an EMA prompt in the afternoon, evening, and weekend compared to mornings and weekdays (except for the model with LPA predicting compliance). Participants were also more likely to complete an EMA prompt as participants received more notifications in the study (ie, time went on in the study). There was no difference in odds of completing an EMA prompt by between-person demographic factors (eg, age).

Potential Reactivity to EMA Prompts

Latinas engaged in significantly more SB in the 15 minutes following the EMA prompt (*B*=0.49, SE 0.21, *P*=.02; mean 9.56, SD 4.14 minutes) compared to the 15 minutes before the prompt (mean 9.07, SD 4.34 minutes). However, there was no difference in MVPA (*B*=−0.04, SE 0.07; *P*=.54) or LPA (*B*=−0.3, SE 0.19; *P*=.11) in the 15 minutes before the EMA prompt compared to the 15 minutes after the EMA prompt (data not shown in table).

Criterion Validity of EMA Responses

Table 3 displays results from models investigating the extent to which there was agreement between device-based movement behaviors and EMA reports of PA and SB. Latinas who self-reported engaging in PA at the time of the prompt averaged 3.6 more minutes of device-based MVPA (*B*=3.60, 95% CI 2.75 - 4.45) and 2.32 more minutes of LPA (*B*=2.32, 95% CI 0.17-4.61) relative to occasions when Latinas did not report being physically active at the time of the prompt. Further, device-based assessments of SB corresponded with self-reported behavior via EMA. At the time of the prompt, Latinas averaged 4.11 more minutes of device-based SB in the 30-minute window (ie, ±15 minutes) on occasions when they reported sitting at the time of the prompt relative to occasions when they did not report sitting at the time of the prompt (*B*=4.11, 95% CI 3.10-5.11).

XSL•FO
RenderX

Table . Predicting minutes of movement behaviors in the 30-minute window (±15 minutes) around the EMA[a] prompt to determine correspondence between EMA-reported and device-based behaviors.[b]

	Multilevel linear regression with MVPA[c] as outcome, unstandardized coefficient (SE)	Multilevel linear regression with LPA[d] as outcome, unstandardized coefficient (SE)	Multilevel linear regression with SB[e] as outcome, unstandardized coefficient (SE)
Fixed effects			
Intercept	0.14 (0.42)	8.89[f] (1.34)	17.09[f] (1.47)
EMA-reported PA[g]	3.60[f] (0.43)	2.32[f] (1.17)	—[h]
EMA-reported SB	—[h]	—[h]	4.10[f] (0.51)
Afternoon	–0.01 (0.28)	–0.86 (0.77)	1.84[f] (0.82)
Evening	0.09 (0.28)	–1.31 (0.76)	2.14[f] (0.82)
Weekend day	–0.25 (0.20)	–0.39 (0.52)	0.71 (0.57)
Notification number	0.01 (0.01)	0.02 (0.04)	–0.05 (0.05)
Age	–0.01 (0.01)	0.07[f] (0.03)	–0.07[f] (0.03)
EMA in Spanish	0.10 (0.27)	0.91 (0.92)	–1.29 (0.99)
Currently working	0.12 (0.19)	0.97 (0.67)	–0.57 (0.70)
US born	–0.01 (0.31)	–1.15 (1.05)	0.40 (1.13)
Self-reported health	0.14 (0.10)	0.25 (0.36)	–0.15 (0.38)
Random effects			
Intercept	0.03 (0.08)	2.46 (1.09)	2.77 (1.24)
Residual	4.90 (0.28)	34.67 (2.01)	40.35 (2.34)

[a]EMA: ecological momentary assessment.

[b]All models were based on 66 participants with 657 observations for models predicting MVPA and LPA, and 656 observations for the model predicting SB.

[c]MVPA: moderate to vigorous intensity physical activity.

[d]LPA: light intensity physical activity.

[e]SB: sedentary behavior.

[f]$P<.05$.

[g]PA: physical activity.

[h]Not applicable.

Adjusting for covariates may overfit models. Another set of unadjusted multilevel models was tested to determine correspondence between EMA-reported and device-based behaviors. Results did not differ substantially from the adjusted results presented in Table 3 regarding agreement between EMA-reported and device-based behaviors. Output from unadjusted multilevel models is in Multimedia Appendix 2.

Acceptability

As indicated in Table 4, most participants (48/66, 72.7%) were satisfied with the study. More than half of the participants indicated that the length of the questionnaires was reasonable (ie, 37/66, 56.1% strongly disagreed or disagreed that the EMA questionnaires took too much time to complete). There were mixed responses regarding the convenience of the timing of EMA prompts (ie, 21/66, 31.8% agreed or strongly agreed prompts were delivered at a convenient time; 15/66, 22.7% disagreed or strongly disagreed). One participant noted in open-ended comments about the study: "Sometimes questionnaires were too early, or I was busy with my kids, and I couldn't answer." Another participant indicated a preference for more time before EMA prompts became inaccessible, with one participant noting that they would have liked to "get more than 30 minutes to answer [the EMA prompt]." Most participants indicated that the preferred number of EMA prompts per day was 3 (52/66, 78.8%) and study days was 7 (53/66, 80.3%). Though in open-ended responses, some participants reported the desire for EMA questionnaires to take less time to complete. Nearly half of the participants indicated that the accelerometer was not a nuisance to wear (29/66, 43.9%) or were neutral about wearing it (18/66, 27.3%), while 28.8% (19/66) indicated it was a nuisance to wear. More than four-fifths (56/66, 84.8%) of the sample indicated that they were satisfied with the compensation provided for completing the study.

XSL•FO
RenderX

Table . Participant feedback on study procedures (n=66[a]).

Item and response options	Values, n (%)
It took too much time to complete the questionnaires sent over the telephone.	
Strongly disagree	14 (21.2)
Disagree	23 (34.8)
Neutral	20 (30.3)
Agree	7 (10.6)
Strongly agree	1 (1.5)
Missing	1 (1.5)
The hours when the EMA[b] prompts were sent over the telephone were convenient.	
Strongly disagree	6 (9.1)
Disagree	9 (13.6)
Neutral	30 (45.5)
Agree	18 (27.3)
Strongly agree	3 (4.5)
It was a nuisance wearing the activity monitor for 7 days.	
Strongly disagree	10 (15.2)
Disagree	19 (28.8)
Neutral	18 (27.3)
Agree	14 (21.2)
Strongly agree	5 (7.6)
This study demanded a lot of time.	
Strongly disagree	18 (27.3)
Disagree	26 (39.4)
Neutral	16 (24.2)
Agree	6 (9.1)
Strongly agree	0 (0)
I was satisfied with the compensation for this study.	
Strongly disagree	3 (4.5)
Disagree	0 (0)
Neutral	6 (9.1)
Agree	33 (50.0)
Strongly agree	23 (34.8)
Missing	1 (1.5)
Overall, what was your level of satisfaction with the study?	
Very unsatisfied	3 (4.5)
Unsatisfied	0 (0)
Neutral	14 (21.2)
Satisfied	33 (50.0)
Very satisfied	15 (22.7)
Missing	1 (1.5)
What is the maximum number of questionnaires you would be willing to complete on your mobile phone in a single day?	

XSL•FO
RenderX

Item and response options	Values, n (%)
3 prompts	52 (78.8)
4 - 5 prompts	8 (12.1)
6 - 8 prompts	6 (9.1)
What is the maximum number of consecutive days you would be willing to complete these study procedures?	
7 days	53 (80.3)
8 - 11 days	7 (10.6)
12+ days	6 (9.1)

[a]One participant did not provide study feedback.

[b]EMA: ecological momentary assessment.

Discussion

EMA can provide novel insights into the prediction and modeling of movement-related behaviors as it captures phenomena of interest in real-time and real-world environments. However, such a methodology is only useful and informative if participants are willing to comply with study procedures and the data collected through this method are valid. To our knowledge, this study is the first to investigate Latinas' willingness, compliance, and accuracy in reports of behaviors in an EMA protocol designed to assess movement-related behaviors. We found that among a sample of emerging to midlife Latinas of Mexican or Mexican American origin and born outside of the United States, these women were willing to engage in and comply with the EMA research, demonstrated strong agreement in reporting engagement in PA and SB when compared to device-based measures, and, overall, had positive perceptions of the study procedures. This study provides encouraging findings for future EMA research with Latinas to better understand and intervene on movement-related behaviors.

Latinas demonstrated acceptable levels of compliance with the EMA protocol, answering approximately 72.5% (927/1279) and completing 69.7% (892/1279) of EMA prompts in this study. Across a variety of research designs, samples, and disciplines, the average compliance with an EMA protocol is approximately 79% [12]. Compared to EMA studies of movement behaviors exclusively among samples of women, compliance rates documented in this study were similar [27,28] or better than previous research [13]. Some of the PA EMA studies among women have implemented strategies to entice EMA compliance by offering additional compensation based on the percentage of EMA prompts answered [27,29]. However, Wrzus and Neubauer [12] found that compliance did not differ if EMA studies used incentives dependent on answering minimum percentage of prompts compared to incentives without compliance reinforcement. It appears the main compliance issue women in our study expressed was the convenience of EMA prompts (eg, timing). Further design considerations regarding EMA prompts among Latinas may be warranted.

Rates of activity monitor nonwear documented in this study were surprising, considering that participants generally reported that it was not a nuisance to wear the activity monitor. Commercial wearable devices may be more advantageous than research-grade devices as commercial devices tend to be water-resistant, comfortable to wear, and stylish, all of which can promote continuous wear with limited reasons for device removal. However, researchers should weigh the pros and cons of using commercially available devices for research purposes [30].

Our findings suggest that the likelihood of completing an EMA prompt increases significantly if Latinas are engaged in higher levels of SB and less LPA around the time of the prompt. Previous EMA studies have similarly documented that rates of compliance are improved if a participant is engaged in high levels of SB around the EMA prompt and that engagement in MVPA around the EMA was unrelated to compliance [16,17]. Other PA EMA studies have documented compliance differences based on concurrent engagement in MVPA (particularly driven by vigorous intensity PA) [15,31], but we did not observe such differences in compliance. This may be due to the relatively low levels of MVPA in our study sample. We did, however, find that engaging in more LPA around the prompt decreased the likelihood of completing an EMA prompt. Engaging in higher levels of LPA may reflect a general busyness of daily life among women, which may explain poorer compliance at those times.

One concern with EMA protocols is that they can potentially induce measurement reactivity (ie, changes in a construct because of intensively assessing it) [22,23]. Our finding that Latinas sat for significantly more time immediately following the EMA prompt as compared to immediately before the EMA prompt aligns with previous findings [15]. Of note, estimates in our study suggest that the difference in SB before and after the EMA prompt was approximately 30 seconds. This volume of change in SB may not have practical implications for Latinas' health. Further, neither MVPA nor LPA significantly changed among Latinas because of completing the EMA prompt, aligning with previous work [15,23]. In general, findings from this study are promising as they suggest that completing an EMA prompt does not induce meaningful reactivity in PA or SB behaviors among Latinas.

Regarding the validity of EMA to assess PA and SB, on occasions when Latinas reported engaging in PA or SB via EMA, actigraphy-derived measures of those behaviors were greater than when those behaviors were not reported via EMA. Importantly, actigraph-derived MVPA increased by nearly 4

XSL•FO
RenderX

minutes when Latinas reported engaging in PA at the EMA prompt, but LPA only increased by about 2 minutes. This suggests that Latinas accurately differentiated between intensities of PA as they were instructed during the introductory training session and reminded at the beginning of each EMA prompt. Given the relatively low prevalence of MVPA in the ±15-minute window around the EMA prompt across all observations (mean 0.96, SD 2.87 minutes), the increase of nearly 4 minutes of device-based PA when PA was reported at the prompt indicates that participant reports were consistent with the behavioral engagement. The agreement between EMA-reported and device-based movement behaviors is consistent with findings from studies involving diverse samples of emerging, midlife, and older adults. These studies have shown increases of 2-4 minutes in device-based PA on occasions when PA was reported through EMA prompts, compared to occasions when it was not [15-17,32]. Ultimately, these findings indicate that Latinas accurately report their current movement behaviors through EMA.

Finally, the overall positive perceptions of the EMA study indicate that this type of smartphone-based study design is acceptable to assess movement behaviors in a sample of Latinas that have immigrated to the United States. Based on participants' feedback, further refining ways to minimize study demands and improve the convenience of study procedures would be advantageous. Additional strategies such as customizing to participants' sleep and wake time schedules or allowing more time to complete the EMA prompt could address issues of convenience. Lengthening the available window in which Latinas can respond to an EMA prompt may help limit potential burden. However, when deciding whether to lengthen the availability of the EMA questionnaire, researchers should weigh the potential benefits and costs of such a decision considering the phenomenon of interest they are interested in capturing. Finally, considerations around the number of items and time to complete at each EMA questionnaire may be necessary. Some EMA studies have reduced the number of items in EMA prompts, while also providing variety in the constructs assessed, by randomizing items so that only a portion of constructs are assessed at each EMA prompt [15,33]. Though such design considerations may only be applicable in studies with a larger number of EMA prompts within days or more days of assessment to ensure that all constructs are adequately captured to reflect the changing contexts of daily life.

Strengths of this study include the rigorous interrogation of factors that may influence potential feasibility and validity of a movement behavior EMA protocol among a socially vulnerable population of Latinas. However, study limitations should be noted. Although simulation studies indicate that in intensive longitudinal study samples of more than 50 participants with assessments within participants can generate reliable fixed and random estimates as well as standard errors [26], our sample size was relatively small (N=67), which may limit the generalizability of findings. Additionally, we obtained a larger percentage of Latinas meeting PA guidelines compared to national averages. Further, most of our sample identified as Mexican or Mexican American and was born outside of the United States. There may be important differences in terms of feasibility, validity, and acceptability depending on Hispanic origin and nativity status. Relatedly, we did not take height and weight measurements to determine participants' Body Mass Index (BMI). Some studies have indicated differences in EMA compliance by BMI [15]. Examining differences in compliance by BMI is an important direction for future EMA research among Latinas. With respect to our study design, because a random, signal-contingent prompting schedule was used, it is likely that not all bouts of PA and SB were captured. Further, EMA items did not assess the duration or intensity of PA or SB.

The findings from our study indicate that EMA methods are a feasible, valid, and acceptable approach to use with Latinas to understand PA and SB. Although some strategies may need to be implemented in future work to enhance compliance with smartphone and accelerometer-based study procedures, our study provides timely, new evidence that EMA can be effectively used among Latinas to capture movement behaviors. Latinas represent a vulnerable and understudied population in PA research. Using EMA in such a group can shed new light on the antecedents and consequences of movement behaviors, leading to new intervention targets delivered in the context of Latinas' daily lives. The findings from this study represent an essential first step in that process.

Acknowledgments

This research was supported by a Faculty Research Grant, Office of Research and Engagement, University of North Carolina at Greensboro. The authors would also like to thank Ashley Banegas, MPH, for her assistance with recruitment and data collection.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

All authors made substantial contributions to the conception or design of the work; the acquisition, analysis, or interpretation of data for the work; in drafting the work or reviewing it critically for important intellectual content; and provided final approval of the version submitted.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Ecological momentary assessment (EMA) screenshots.
[DOCX File, 956 KB - humanfactors_v12i1e75855_app1.docx]

Multimedia Appendix 2
Predicting minutes of movement behaviors in the 30-minute window (±15 minutes) around the EMA prompt unadjusted for covariates to determine correspondence between EMA-reported and device-based behaviors. EMA: ecological momentary assessment.
[DOCX File, 19 KB - humanfactors_v12i1e75855_app2.docx]

References

1. Arredondo EM, Sotres-Alvarez D, Stoutenberg M, et al. Physical activity levels in U.S. Latino/Hispanic adults. Am J Prev Med 2016 Apr;50(4):500-508. [doi: 10.1016/j.amepre.2015.08.029]
2. Neighbors CJ, Marquez DX, Marcus BH. Leisure-time physical activity disparities among Hispanic subgroups in the United States. Am J Public Health 2008 Aug;98(8):1460-1464. [doi: 10.2105/AJPH.2006.096982] [Medline: 18048795]
3. Adult physical inactivity outside of work. US Centers for Disease Control and Prevention. 2025 Jan 31. URL: https://www.cdc.gov/physical-activity/php/data/inactivity-maps.html [accessed 2025-10-15]
4. Chang YJ, Bellettiere J, Godbole S, et al. Total sitting time and sitting pattern in postmenopausal women differ by Hispanic ethnicity and are associated with cardiometabolic risk biomarkers. J Am Heart Assoc 2020 Feb 18;9(4):e013403. [doi: 10.1161/JAHA.119.013403] [Medline: 32063113]
5. Hawkins MS, Storti KL, Richardson CR, et al. Objectively measured physical activity of USA adults by sex, age, and racial/ethnic groups: a cross-sectional study. Int J Behav Nutr Phys Act 2009 Jun 3;6(1):31. [doi: 10.1186/1479-5868-6-31] [Medline: 19493347]
6. Echeverría SE. Improving Latino health through equity-centered physical activity research. Am J Public Health 2024 Jul;114(S6):S436-S438. [doi: 10.2105/AJPH.2024.307764] [Medline: 39083733]
7. Dunton GF. Ecological momentary assessment in physical activity research. Exerc Sport Sci Rev 2017 Jan;45(1):48-54. [doi: 10.1249/JES.0000000000000092] [Medline: 27741022]
8. Stone AA, Shiffman S. Ecological momentary assessment (EMA) in behavioral medicine. Ann Behav Med 1994 Sep 1;16(3):199-202. [doi: 10.1093/abm/16.3.199]
9. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
10. Calderón JL, Baker RS, Fabrega H, et al. An ethno-medical perspective on research participation: a qualitative pilot study. MedGenMed 2006 Apr 25;8(2):23. [Medline: 16926762]
11. George S, Duran N, Norris K. A systematic review of barriers and facilitators to minority research participation among African Americans, Latinos, Asian Americans, and Pacific Islanders. Am J Public Health 2014 Feb;104(2):e16-e31. [doi: 10.2105/AJPH.2013.301706] [Medline: 24328648]
12. Wrzus C, Neubauer AB. Ecological momentary assessment: a meta-analysis on designs, samples, and compliance across research fields. Assessment 2023 Apr;30(3):825-846. [doi: 10.1177/10731911211067538] [Medline: 35016567]
13. Campbell KL, Wang Y, Kuhn AP, Black MM, Hager ER. An ecological momentary assessment study of physical activity behaviors among mothers of toddlers from low-income households. BMC Womens Health 2021 Mar 22;21(1):120. [doi: 10.1186/s12905-021-01243-2] [Medline: 33752659]
14. Amagasa S, Machida M, Fukushima N, et al. Is objectively measured light-intensity physical activity associated with health outcomes after adjustment for moderate-to-vigorous physical activity in adults? A systematic review. Int J Behav Nutr Phys Act 2018 Jul 9;15(1):65. [doi: 10.1186/s12966-018-0695-z] [Medline: 29986718]
15. Dunton GF, Liao Y, Kawabata K, Intille S. Momentary assessment of adults' physical activity and sedentary behavior: feasibility and validity. Front Psychol 2012;3:260. [doi: 10.3389/fpsyg.2012.00260] [Medline: 22866046]
16. Maher JP, Rebar AL, Dunton GF. Ecological momentary assessment is a feasible and valid methodological tool to measure older adults' physical activity and sedentary behavior. Front Psychol 2018;9:1485. [doi: 10.3389/fpsyg.2018.01485] [Medline: 30158891]
17. Maher JP, Sappenfield K, Scheer H, Zecca C, Hevel DJ, Kennedy-Malone L. Feasibility and validity of assessing low-income, African American older adults' physical activity and sedentary behavior through ecological momentary assessment. J Meas Phys Behav 2021;4(4):343-352. [doi: 10.1123/jmpb.2021-0024]
18. Matthews CE, Chen KY, Freedson PS, et al. Amount of time spent in sedentary behaviors in the United States, 2003-2004. Am J Epidemiol 2008 Apr 1;167(7):875-881. [doi: 10.1093/aje/kwm390] [Medline: 18303006]
19. Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc 2008 Jan;40(1):181-188. [doi: 10.1249/mss.0b013e31815a51b3] [Medline: 18091006]
20. Bolger N, Davis A, Rafaeli E. Diary methods: capturing life as it is lived. Annu Rev Psychol 2003;54(1):579-616. [doi: 10.1146/annurev.psych.54.101601.145030] [Medline: 12499517]

XSL•FO
RenderX

21. Raudenbush SW, Bryk AS. Hierarchical Linear Models: Applications and Data Analysis Methods: SAGE Publications; 2002.
22. French DP, Sutton S. Reactivity of measurement in health psychology: how much of a problem is it? What can be done about it? Br J Health Psychol 2010 Sep;15(Pt 3):453-468. [doi: 10.1348/135910710X492341] [Medline: 20205982]
23. Maher JP, Arigo D, Baga K, et al. Measurement reactivity in ecological momentary assessment studies of movement-related behaviors. J Meas Phys Behav 2024 Apr 15;7(1). [doi: 10.1123/jmpb.2023-0035]
24. Tate AD, Fertig AR, de Brito JN, et al. Momentary factors and study characteristics associated with participant burden and protocol adherence: ecological momentary assessment. JMIR Form Res 2024 Apr 24;8:e49512. [doi: 10.2196/49512] [Medline: 38656787]
25. Wang S, Yang CH, Brown D, Cheng A, Kwan MYW. Participant compliance with ecological momentary assessment in movement behavior research among adolescents and emerging adults: systematic review. JMIR Mhealth Uhealth 2025 Feb 11;13(1):e52887. [doi: 10.2196/52887] [Medline: 39933165]
26. Maas CJM, Hox JJ. Sufficient sample sizes for multilevel modeling. Methodology (Gott) 2005 Jan;1(3):86-92. [doi: 10.1027/1614-2241.1.3.86]
27. Arigo D, Hevel DJ, Bittel K, Maher JP. Within-person examination of the exercise intention-behavior gap among women in midlife with elevated cardiovascular disease risk. Psychol Sport Exerc 2022 May;60:102138. [doi: 10.1016/j.psychsport.2022.102138] [Medline: 35531355]
28. Elavsky S, Kishida M, Mogle JA. Concurrent and lagged relations between momentary affect and sedentary behavior in middle-aged women. Menopause 2016 Aug;23(8):919-923. [doi: 10.1097/GME.0000000000000645] [Medline: 27326822]
29. Do B, Wang SD, Naya CH, Dunton GF, Mason TB. Momentary and personal characteristics predicting maternal fruit and vegetable preparation for children using ecological momentary assessment. Eat Behav 2021 Apr;41:101492. [doi: 10.1016/j.eatbeh.2021.101492] [Medline: 33677396]
30. Balbim GM, Marques IG, Marquez DX, et al. Using Fitbit as an mHealth intervention tool to promote physical activity: potential challenges and solutions. JMIR Mhealth Uhealth 2021 Mar 1;9(3):e25289. [doi: 10.2196/25289] [Medline: 33646135]
31. Ponnada A, Thapa-Chhetry B, Manjourides J, Intille S. Measuring criterion validity of microinteraction ecological momentary assessment (micro-EMA): exploratory pilot study with physical activity measurement. JMIR Mhealth Uhealth 2021 Mar 10;9(3):e23391. [doi: 10.2196/23391] [Medline: 33688843]
32. Knell G, Gabriel KP, Businelle MS, Shuval K, Wetter DW, Kendzor DE. Ecological momentary assessment of physical activity: validation study. J Med Internet Res 2017 Jul 18;19(7):e253. [doi: 10.2196/jmir.7602] [Medline: 28720556]
33. Maher JP, Hevel DJ, Bittel KM, Hudgins BL, Labban JD, Kennedy-Malone L. Studying movement-related behavioral maintenance and adoption in real time: protocol for an intensive ecological momentary assessment study among older adults. JMIR Res Protoc 2023 Jul 28;12:e47320. [doi: 10.2196/47320] [Medline: 37505805]

Abbreviations

EMA: ecological momentary assessment
LPA: light-intensity physical activity
MVPA: moderate-to-vigorous-intensity physical activity
OR: odds ratio
PA: physical activity
SB: sedentary behavior

Edited by A Malik; submitted 11.04.25; peer-reviewed by S Harris, SH Park; revised version received 04.09.25; accepted 24.09.25; published 05.11.25.

Please cite as:
Maher JP, Greco PA, Fernandez EC, Hudgins BL, Echeverria SE
Capturing Movement Behaviors in Latinas: Feasibility, Validity, and Acceptability Study of an Ecological Momentary Assessment Protocol
JMIR Hum Factors 2025;12:e75855
URL: https://humanfactors.jmir.org/2025/1/e75855
doi:10.2196/75855

© Jaclyn P Maher, Peyton A Greco, Eugenia Camacho Fernandez, Brynn L Hudgins, Sandra E Echeverria. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 5.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use,

XSL•FO
RenderX

distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Perceptions of User-Generated Content as a Source of Health Messages in Smoking Cessation Mobile Interventions: Focus Group Study

Michael Wakeman[1,2], BS; Lydia Tesfaye[1,3], BS; Tim Gregory[4], BA; Erin Leahy[4], BA; Gunnar Baskin[4], BS; Greg Gruse[4], BS; Brandon Kendrick[4], MA; Sherine El-Toukhy[1], MA, PhD

[1]Division of Intramural Research, National Institute on Minority Health and Health Disparities, 11545 Rockville Pike, Rockville, MD, United States

[2]Pritzker School of Medicine, University of Chicago, Chicago, IL, United States

[3]School of Medicine, Emory University, Atlanta, GA, United States

[4]ICF NEXT (United States), Reston, VA, United States

Corresponding Author:
Sherine El-Toukhy, MA, PhD
Division of Intramural Research, National Institute on Minority Health and Health Disparities, 11545 Rockville Pike, Rockville, MD, United States

Abstract

Background: Health messages are integral to smoking cessation interventions. Common approaches to health message development include expert-crafted messages and audience-generated messages, which produce messages that can be monotonic, didactic, and limited in number. We introduce an alternative approach to health message development that relies on user-generated content available on open-content platforms as a source of health messages.

Objective: We examined the acceptability of user-generated content curated from Twitter (subsequently rebranded X) as a source of health support messages in a newly developed smoking cessation mobile intervention called Quit Journey and the optimal timing and frequency with which health messages can be deployed to support app users in real time.

Methods: A total of 12 semistructured focus groups were held with 38 young adults with low socioeconomic status who smoked cigarettes, wanted to quit, and were aged 18 to 29 years. Focus groups were held virtually on GoTo Meeting, audio recorded, and transcribed verbatim. Deductive thematic analysis was used, with themes based on 5 constructs from the second unified theory of acceptance and use of technology (ie, effort expectancy, facilitating conditions, hedonic motivation, performance expectancy, and social influence) and negative, neutral, and positive sentiment.

Results: Participants perceived user-generated content positively (56/108, 51.9% of the quotes) and focused on their perceived usefulness (37/108, 34.3% of the quotes). User-generated content was perceived as authentic, nonrepetitive support from people with similar real-life experiences. Negative or sarcastic user-generated content elicited negative reactions from participants. Participants preferred receiving 3 or fewer daily messages, ideally before cravings. Suggestions focused on the need to screen user-generated content before its inclusion in the app library and allow app users to customize message frequency and timing.

Conclusions: User-generated content was deemed an acceptable source of health messages. This content can improve the efficacy and effectiveness of smoking cessation interventions by increasing their pool of unique messages that may be better received and more persuasive than expert-curated content. User-generated content can be used to curate health messages for all medical conditions and behaviors with relevant publicly available online content for integration in behavioral interventions given its high volume, brevity, and narrative-like nature. Future research is needed to investigate the effects of user-generated content on health behaviors and identify the theoretical mechanisms for these effects.

(JMIR Hum Factors 2025;12:e76804) doi:10.2196/76804

KEYWORDS

focus group discussions; low socioeconomic status; smoking cessation apps; young adults; user-generated content; social media; Twitter; X

Introduction

Health messaging is integral to smoking cessation efforts, including behavioral interventions [1,2]. Effective health messages are understandable; informative; engaging; and, most importantly, capable of inducing behavior change [3]. Despite the importance of health messages to behavioral interventions, the process of health message development is “commonly

XSL•FO
RenderX

hidden within a black box" [4]. The most common is a top-down approach to message development, where interventions include expert-crafted health messages that embody clinical or expert guidelines, determinants of behavior change as outlined in theoretical frameworks (eg, health belief model, theory of planned behavior, and elaboration likelihood model), message tactics (eg, gain framed and narratives), and approaches (eg, targeting and tailoring) [3]. Alternatively, target populations have been involved in developing health messages in a bottom-up approach, although this approach is less frequently used [5]. We propose a different approach, where health messages are curated from online open-content sources for use in smoking cessation interventions.

There are limitations to current approaches to health message development. Expert-constructed messages can be perceived as dry, didactic, or unrelatable and can be met with persuasion resistance from audiences [6,7]. Target population–constructed messages can be difficult to generate, especially at scale. Both approaches to message development generate a finite number of messages that are recycled over the course of an intervention. Message repetitiveness can result in message fatigue, which leads to unfavorable message attitudes and reduced message effectiveness, credibility, and behavioral intentions [8,9]. This is particularly important given the need for a substantial pool of candidate messages to cover the extended treatment times of smoking cessation interventions, which usually last 4 to 6 weeks (and often require multiple quit attempts), and the need for personalized message content and intensity under adaptive interventions [10,11].

Web 2.0 and social media platforms have an endless reservoir of user-generated health content owing to their active users, including 2 billion for Instagram and 600 million for Twitter (subsequently rebranded X) [12]. User-generated content is defined as "content [that] is created or produced by the general public rather than by paid professionals and primarily distributed on the internet" [13]. There is evidence suggesting that people turn to online platforms, including social media, for health purposes. For instance, health discussion forums, social media platforms, and crowdsourcing platforms have user-generated smoking cessation content [14,15]. In 2018, a total of 70.14% of Americans used digital means to access health information, whereas 14.02% shared health information via social media platforms [16]. However, research on Web 2.0 and social media platforms has been limited to examining tobacco advertising and audience-targeting strategies and their associations with tobacco use behaviors [17-19]; analyzing opinions and attitudes regarding tobacco products and the viability of predicting user behavior based on online media posts [20-23]; and, finally, engaging with and disseminating antitobacco messaging and cessation interventions [14,24-26].

User-generated content remains a practically untapped resource for persuasive health messaging. There is evidence suggesting that users trust content created by those going through similar experiences [27]. User-generated online content is also akin to unscripted peer-to-peer communication, which is typically conversational and can include narrative elements or storytelling. These content characteristics can induce behavior change through various mechanisms, such as reducing counterarguments and facilitating role-modeling of the behavior [28-31]. For example, peer messages have been associated with greater engagement with tobacco interventions compared to expert messages [32], whereas emotional and personal testimonials have been associated with greater quitting among individuals with low and middle socioeconomic status (SES) who smoke than among those with high SES [33].

In this study, we explored reactions to the potential of user-generated content as a source of support messages in a smoking cessation intervention [34] called Quit Journey that targets individuals with low SES who smoke cigarettes. The app will feature an on-demand message library and just-in-time support messages to be sent to users when they are at risk of a lapse or relapse. These messages will be drawn from Twitter, a microblogging platform. As target audiences' reactions to health messages are critical determinants of their effectiveness [3,35], we gauged the acceptance of the use of Twitter messages for cessation support in Quit Journey among young adults with low SES who smoked and their opinions on the frequency and timing of the just-in-time support messages. This study was part of the preparation phase of the multiphase optimization strategy for the development and evaluation of Quit Journey [36,37].

Methods

Ethical Considerations

The National Institutes of Health institutional review board exempted this study on October 11, 2019, under category 2, research that only includes interactions involving educational tests, survey procedures, interview procedures, or observation of public behavior (title 45 of the Code of Federal Regulations, part 46.10(d)(2)), and category 3, research involving benign behavioral interventions (title 45 of the Code of Federal Regulations, part 46.10(d)(3)). ICF International's institutional review board exempted this study on November 19, 2019, under category 2 and approved an amendment on February 26, 2020. All participants verbally consented to take part in the study. Participants received US $150 as compensation. Participants were assigned numeric identification numbers that were used during the focus group discussions and in the transcripts.

Participants and Recruitment

We partnered with UserWorks, Inc, to recruit a convenience sample of 38 young adults with low SES who smoked. Individuals with low SES are a priority for smoking cessation efforts, with high prevalence of cigarette smoking among adults with low incomes (18.3%) and those without a high school diploma or who passed the General Educational Development test (20.1% and 30.7%, respectively) [38].

Eligibility criteria included being aged 18 to 29 years; having a low SES, as evidenced by being neither a 4-year college graduate nor a college enrollee [39-41]; being a current smoker who smoked at least 100 cigarettes in their lifetime and reported smoking every day or some days; willingness to quit within 6 months; not currently using any smoking cessation aids or noncigarette combustible tobacco products; being a smartphone owner; and speaking English. Recruitment took place between January 2020 and April 2020. The sample size was deemed

XSL•FO
RenderX

sufficient to achieve data saturation, and none of the participants withdrew from the study [42,43]. The 32-item COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist can be found in Table S1 in Multimedia Appendix 1 [44].

Procedures

We conducted 12 virtual focus groups lasting approximately 1.5 hours each on GoTo Meeting using a topic guide informed by the second unified theory of acceptance and use of technology (Multimedia Appendix 1) [45,46]. In this paper, we focus solely on Quit Journey's message library and just-in-time support feature, which are under development. The moderator described the concept behind the on-demand message library and just-in-time support feature in 8 focus groups and presented mock app pages with placeholder messages and 5 sample tweets in 4 focus groups. For example, one sample tweet read the following:

> *...after seeing a smokers teeth now im glad I haven't touched a cig in months!!*

All sample tweets can be found in Note S1 in Multimedia Appendix 1. In total, 31.6% (12/38) of the participants were involved in 2 focus groups, one in which the moderator presented the concept behind utilizing user-generated content in Quit Journey and one in which the moderator presented app mock pages with placeholder user-generated content. Our work on perceptions of smoking cessation apps and of Quit Journey and its individual features has been published elsewhere [47-51; Wakeman M, unpublished data, April 2025].

We conducted dry runs to ensure that the sessions ended on time. TG, a male user experience strategist, moderated the discussions. EL, a female strategic communications and marketing project director, was a backup moderator and recorded notes. TG, EL, and SE-T were the only nonparticipants present during the focus group discussions, and none had a previous relationship with the participants. Participants were informed that the moderators were unaffiliated with the research group that commissioned the study. The focus groups were audio recorded and auto-transcribed by GoTo Meeting. Three members of our staff verified these transcripts against the audio files. Transcripts and findings were not returned to participants for feedback.

Analysis

We adopted a deductive thematic approach to data analysis [52]. After an initial review of the transcripts, we developed codes and corresponding themes based on 5 constructs from the second unified theory of acceptance and use of technology: effort expectancy, facilitating conditions, hedonic motivation, performance expectancy, and social influence [45,46]. Furthermore, we coded the transcripts for negative, neutral, and positive sentiment; design concepts; suggestions for improvement; and Quit Journey app features [49-51]. Briefly, effort and performance expectancies were defined as "perceived ease or effortfulness" and "perceived usefulness or helpfulness" of mobile apps and their features, respectively. Facilitating conditions were defined as "factors that can aid or impede [their] uptake or use," whereas social influence referred to "perceived importance of significant others' recommendations and approval" in considering use of mobile apps and their features. Hedonic motivation was defined as "perceived fun, pleasure, or enjoyment (or lack thereof) associated with [their] use." Negative and positive sentiment captured statements that indicated "a sense of disapproval, criticism, or skepticism" or "a sense of approval, praise, or certainty" about mobile apps, respectively. Neutral sentiment captured remarks that "did not have positive or negative tone, contained equal number of positive- and negative-toned remarks, or were conditional in nature." The design concept code captured statements related to the visual depiction and aesthetics of mobile apps or their features. Finally, the suggestion code captured statements that were "concerned with improvements, modifications, or adjustments" to mobile apps or their features, whereas the app feature code captured the Quit Journey app feature being discussed (eg, carbon monoxide tracking; Wakeman M, unpublished data, April 2025) [47-51]. Quotes that did not fit under a technology acceptance construct but still conveyed a particular sentiment were coded as "not applicable."

We introduced new codes if the content in the transcripts did not fit the predefined ones. In coding the transcripts, we applied the predominant theme in each quote even if it was not mentioned explicitly (ie, in response to a moderator's question). When one theme was contingent on another, we coded the quote for the underpinning theme. We followed a multicoding approach where multiple domains (eg, technology acceptance and sentiment) could be applied to a single quote but only 1 code could be selected from each domain (eg, negative, neutral, or positive).

MW and LT independently coded the focus group transcripts. Using the ATLAS.ti qualitative software (version 8; ATLAS.ti Scientific Software Development GmbH), we calculated 2 intercoder agreement measures: the Krippendorff c-alpha and Krippendorff cu-alpha. The first, a measure of separating relevant and irrelevant content (ie, whether coders identified texts of similar locations and lengths), was 0.82 [53]. The second, a metric related to semantic domain reliability (ie, whether coders coded for the presence or absence of groups of related themes), was 0.66 for technology acceptance, 0.70 for sentiment, and 0.94 for app features [53]. MW, LT, and SE-T discussed discrepancies and resolved them.

Results

Overview

Sample characteristics are shown in Table 1, and detailed participant characteristics can be found in Table S2 in Multimedia Appendix 1.

XSL•FO
RenderX

Table . Participant characteristics (N=38) [47-51].

Characteristics	Participants, n (%)
Sex	
Female	20 (52.6)
Male	18 (47.4)
Race and/or ethnicity	
American Indian or Alaska Native	1 (2.6)
Asian, Native Hawaiian, or Pacific Islander	3 (7.9)
Black or African American	11 (28.9)
Hispanic or Latino	6 (15.8)
White	16 (42.1)
Mixed	1 (2.6)
Highest level of education	
Lower than high school	3 (7.9)
High school graduate	10 (26.3)
High school equivalent	3 (7.9)
Some college, no degree	18 (47.4)
2-year associate degree	4 (10.5)
Smoking frequency	
Every day	30 (78.9)
Some days	8 (21.1)
Quit time frame	
7 days	11 (28.9)
30 days	22 (57.9)
6 months	5 (13.2)
Smartphone operating system	
Android	21 (55.3)
iOS	17 (44.7)

User-generated content was an acceptable source of smoking cessation support messages. Of 108 extracted quotes, roughly half (n=56, 51.9%) reflected a positive sentiment (Table 2). Discussions largely focused on performance expectancy (n=37, 34.3%), followed by effort expectancy (n=7, 6.5%), hedonic motivation (n=7, 6.5%), and facilitating conditions (n=1, 0.9%). All quotes appear verbatim in Tables S3 and S4 in Multimedia Appendix 1.

Table . Distribution of the number of quotes by technology acceptance and sentiment toward the user-generated message library. Column totals add up to 100% within each theme, whereas overall row totals add up to 100% across a semantic domain (N=108).

Technology acceptance domain	Sentiment, n/N (%)			Total, n/N (%)
	Negative	Neutral	Positive	
Effort expectancy	1/7 (14.3)	0 (0)	6/7 (85.7)	7/108 (6.5)
Facilitating conditions	0 (0)	1/1 (100)	0 (0)	1/108 (0.9)
Hedonic motivation	0 (0)	3/7 (42.9)	4/7 (57.1)	7/108 (6.5)
Performance expectancy	9/37 (24.3)	3/37 (8.1)	25/37 (67.6)	37/108 (34.3)
Social influence	0 (0)	0 (0)	0 (0)	0 (0)
Not applicable	19/56 (33.9)	16/56 (28.6)	21/56 (37.5)	56/108 (51.9)
Total	29/108 (26.9)	23/108 (21.3)	56/108 (51.9)	108/108/ (100)

XSL•FO
RenderX

Performance Expectancy

Most quotes on the usefulness of user-generated content for smoking cessation (25/37, 67.6%) overlapped with positive sentiment. Reasons for its perceived usefulness included providing nonrepetitive, authentic content from people "going through the same thing as you" (P35). User-generated content was also seen as a validation of the intended audiences' emotions and experiences while imparting a sense of community and social support:

> *I know I get annoyed when...something that's...giving you motivational quotes and it'll just start like repeating them because they only have a few.* [P11]

> *Yeah, I feel like it'd be very motivating to have that sense of community [from the tweets] and to see it working in real-time and kind of have that push to be inspired to keep going, if you see actual people getting benefits from it.* [P10]

Participants emphasized the benefits of knowing the experiences of others, which could improve their self-efficacy and motivation to quit:

> *I think [the tweets] would be super helpful for people.... The community who have ended up quitting smoking, I think it would be super helpful to see what they have to say and what they think about it after the fact. So, yeah, I like that a lot.* [P08]

> *I think it's nice to have a personal touch and [the tweet] gives the user a reminder that it's possible to achieve the outcome you're wanting and...there's proof from other people that have been able to overcome the...addiction that we're all trying to kick.... There are key points that I can see in some of these [example tweets] that I'm thinking to myself are applicable to me and concerns that I have. So, it's just another reminder and incentive and...positive reinforcement that I think is a great idea.* [P14]

Quotes with a negative sentiment (9/37, 24.3%) captured skepticism about the relevance of others' success stories or their motivating impact on individual participants' quit journeys:

> *I just don't think that...reading what other people are doing...is like helpful to me and...if I'm quitting smoking, I don't care if someone else...is a "proud owner of a tobacco free body," right. Like, there's lots of people who don't smoke. But...that's not my journey.* [P11]

Negative sentiment also reflected participants' reactions to sample tweets that were perceived as sarcastic or ironic in tone or those that referenced the adverse health effects of smoking:

> *I don't like the first [example tweet], because I feel like that would just make me feel bad, especially if like my teeth are already yellow from smoking.... If you're...struggling and then someone else is, like, "Oh, yay." Like, you know, "I'm doing so well and blah blah blah," and...that can make you feel kind of...bad. And then...if you feel that you might just start smoking more because you're like, what's the point, I'm a failure.* [P11]

Other Technology Acceptance Themes

Participants perceived Quit Journey's tweet-based message library as both easy and fun to use. Quotes reflecting effort expectancy, or the perceived ease of using the message library, and hedonic motivation, or the perceived enjoyment of using the message library, were generally positive (6/7, 85.7% and 4/7, 57.1%, respectively):

> *I think it will be fun [to include tweets].* [P16]

Discussions rarely focused on conditions that could facilitate or impede the use of a Twitter-based message library or just-in-time support messages. For example, one participant had no privacy concerns, noting that they would be comfortable with the app collecting their ratings of messages to personalize message content as the app learned their preferences.

Suggestions

An additional 105 quotes focused on participants' suggestions related to message content, timing, and frequency. First, to maintain the authenticity of user-generated content (ie, ensuring that app users know that the messages were authored by real people), participants suggested that messages be attributed to the poster's name, Twitter handle, or online platform. Others suggested that the app's library maintain the visual presentation of Twitter messages to signal the messages' authenticity. Finally, participants suggested that the tweets should be unedited or unaltered, including emojis and any spelling or grammatical errors unless they jeopardized the meaning of the message. If edited, it should be minimal to aid comprehension:

> *If you're going to have [tweets], yeah, it's better to know that it's from a real person than like a computer-generated thing.* [P11]

> *Just mention the name of the platform the messages are from but exclude the name of the people who posted. Keep them anonymous.* [P06]

> *We use social media a lot, we do a lot of misspelling, we miss punctuations, and all of that stuff. So...if that is gonna be coming from an actual person...you want to show...not just the good, but the ugly too [in the tweets], you know not everyone can spell.* [P16]

Participants emphasized that sarcastic, ironic, or negative tweets should not be included in the app's message library:

> *I would just be cautious with what [Twitter] messages you use.... But I like the idea, like, getting a notification. Maybe...every day or twice a day or something that has something kind of inspirational to help you. But...I think it's...really a matter of choosing ones...that are all positive.* [P05]

Regarding push messages, participants suggested tailoring the messages based on age, smoking behavior, and ratings of previous messages, among other factors. Furthermore, participants emphasized the importance of being able to customize the timing and frequency of notifications. For example, although feedback was mixed, most participants preferred to receive support messages before a craving occurred.

XSL•FO
RenderX

Participants suggested that they would be receptive to 1 to 10 messages per day, with 3 messages per day being the most common:

> *I guess [I would prefer to receive messages] before, or during the craving, because if it comes afterwards, then you probably already smoked one.* [P10]

> *So, in my opinion, it would be that if you're making a really good progress, the less [tweets] that you may need, or it can be the opposite...if you're doing really good, more [tweets].... I don't know. I think maybe start off three a day—morning, noon, night.* [P16]

> *I feel like [the tweets are] really personal and probably even could like learn me in a way where if I'm having...a bad day or...there's certain times where I feel like I need a little more support then.... Just to kind of like throw it at me and then if I have to change it then I will change it, you know?* [P02]

Discussion

Principal Findings

This study showed that user-generated content was an acceptable source of support messages in Quit Journey, a smoking cessation intervention targeting individuals with low SES who smoke. The perceived usefulness of user-generated content reflected core characteristics of effective health messaging, including their brevity, relatedness to target populations, and narrative-like nature, while avoiding pitfalls common to expert-designed messages, such as monotony, repetitiveness, and didacticism [3,6]. Suggestions to improve user-generated content for use in Quit Journey's message library and just-in-time support feature included refraining from overly negative or sarcastically toned messages and tailoring the messages to users' characteristics, time-varying needs, and message preferences [3]. Cumulatively, these results suggest that we can leverage user-generated content to build Quit Journey's message library. More broadly, user-generated content marks a departure from the limited number of top-down, expert-designed messages commonly used in health interventions, which can be subject to persuasion resistance and message fatigue [8,9]. Moreover, the use of user-generated content exponentially increases the pool of candidate messages, which allows for the tailoring of message content, frequency, and timing to users' needs. This is particularly true for technology-based smoking cessation interventions, where the capture of time-varying psychological, physiological, and contextual factors is increasingly feasible [54,55]. Our proposed approach of relying on user-generated content is applicable to practically all health conditions and behaviors with relevant publicly available online content, including addictive and lifestyle behaviors [56,57].

Comparison to Prior Work

User-generated content includes elements of effective health messages, including its brevity and narrative-like nature, which can improve message persuasiveness and effectiveness in promoting behavior change [30,31]. Indeed, our participants attributed the usefulness of user-generated content to its realism and role-modeling from others who have attempted to quit, successfully or otherwise. For example, one participant noted their aversion to "computer-generated" messages (P11), whereas another noted how these messages could improve self-efficacy, saying that "if someone else can do it, maybe...you can do it too" (P21). These positive perceptions may also reflect involvement with message creators and trust in the health information provided as personal stories by those perceived to be similar to them and with firsthand experiences [27,29,58]. Evidence shows that peer messages and emotional and narrative content have been used to effectively treat nicotine dependence and are associated with greater engagement with tobacco cessation interventions [32,59]. Notably, tweets are similar in length to SMS text messages, which have proven successful as smoking cessation interventions [60]. While user-generated content has been shown to influence nonhealth decision-making (eg, purchasing decisions) [61,62], future research should examine its persuasive effect on health behaviors, identify mechanisms of its effect (eg, self-efficacy and perceived norms), and compare its effectiveness to other content generation approaches (eg, expert-crafted messages and artificial intelligence conversational agents or chatbots [63]).

Open-source content negates message fatigue associated with message repetition [8,9]. A typical smoking cessation intervention is 6 to 8 weeks long and includes a fixed number of support messages (eg, the National Cancer Institute's SMS text messaging intervention SmokefreeTXT) [64]. An intervention recipient is bound to receive the same messages during the intervention once they have exhausted the number of messages dedicated to a particular purpose (eg, craving support). Furthermore, individuals who smoke often attempt to quit several times before they are successful [65]. This means that they will be exposed to the same messages if they reset their quit day or re-enroll in an intervention. Additionally, tailoring messages to individuals' characteristics and time-varying factors as they undergo an intervention quickly becomes a challenge [66]. This could be one of the reasons why many smoking cessation interventions use only basic tailoring strategies (eg, addressing an intervention recipient by their name) [67]. A large pool of candidate smoking-related messages curated from Twitter or other platforms can facilitate tailoring of intervention message content.

Our results show that the appeal of user-generated content as a source of cessation support messages was dependent on message content and attributes. Participants generally preferred positive messages over those perceived as sarcastic or ironic, consistent with previous research [68]. Furthermore, perceptions were mixed on messages that focused on the adverse health effects of smoking. These discussions reflect debates on the use of fear appeals (ie, messages that include a threat) and gain- versus loss-framed messages (ie, messages that emphasize benefits of compliance with message recommendations or losses associated with noncompliance, respectively) in health interventions [69,70]. There is evidence suggesting that fear appeals can be effective in promoting smoking cessation [71-73], including among individuals with low SES who smoke [2]. It is noteworthy that the persuasiveness of a message depends on various factors, such as the behavior in question (eg, one time vs repeated and prevention vs detection), audience

XSL•FO
RenderX

characteristics (eg, readiness for change), and other message elements (eg, inclusion of self-efficacy statements). For example, evidence shows that gain-framed messages can be slightly more effective in promoting some preventive behaviors, including smoking cessation [74,75]. Additionally, there are unintended and ethical consequences of fear appeals and loss-framed messages and questions about their long-term effects on behavior [76]. Taken together, feedback from our participants and existing literature suggests a need to identify the underlying theoretical concepts conveyed in user-generated messages and investigate the effect of different types of user-generated messages on behavioral outcomes at various stages of a smoking cessation intervention.

Participants' suggestions centered on message attribution and content and the frequency and timing of push messages. They emphasized the appeal of attributing the messages in Quit Journey to their actual authors. This is consistent with research showing that message source can increase trust and operate as a cue to persuasion and attitude change [27,77]. Relatedly, participants noted the need for the messages to be authentic and human-authored, suggesting that the tweets retain their conversational tones and even grammatical errors. Finally, participants emphasized the need to tailor message content to personal and dynamic factors [66,78,79]. For instance, participants suggested several tailoring factors, including cravings, mood, and stage of quitting. Furthermore, participants suggested that, on average, they would be open to receiving up to 3 daily push notifications but preferred to have the ability to customize the frequency and timing of these messages. While these requests are common [80], they defeat the purpose of just-in-time support [10]. A compromise could be to allow users some control over the number of daily messages they want to receive and giving them options to customize other aspects of the intervention (eg, app aesthetics).

Strengths and Limitations

To our knowledge, our approach to curating publicly available user-generated content in a smoking cessation mobile intervention (and in health interventions generally) is the first of its kind [22,54]. In this study, we demonstrated initial acceptance of this approach among the target population of our smoking cessation intervention, Quit Journey. Study strengths include a diverse group of participants, with no more than 50% being from one race and/or ethnicity. Limitations include the reliance on educational level as a sole indicator for SES [39-41]. The focus groups were held virtually due to the COVID-19 pandemic, which may have impacted engagement in discussions. Additionally, the pandemic had mixed effects on smoking behaviors [81], which consequently could have affected participants' perceptions of our smoking cessation mobile app and its components, including the proposed user-generated message library to support smoking cessation. Participants were only shown 5 example tweets, which were not representative of the wide range of smoking-related tweets that could have elicited different reactions from participants. Furthermore, because the message library and just-in-time support features of Quit Journey were still under development, participants were presented with mock pages with placeholder messages. Participants were not representative of all individuals who smoke, such as older individuals, some of whom may not be as receptive to social media content as the young adults in our study [82]. Finally, due to scheduling conflicts, focus groups had unequal numbers of participants.

Conclusions

This study provided evidence of the acceptability of user-generated content as a source of support messages in smoking cessation interventions. Implementing this new approach to developing health messages requires additional research to select and evaluate user-generated content with the target audience before these messages are integrated into a behavioral intervention. Although we focused on smoking cessation, our proposed approach to health message development can be applied to other health conditions and risky behaviors to accommodate increasing needs for content-diverse, relatable, and short messages that can support individual users and their dynamic needs during behavior change.

Acknowledgments

ICF International's services for this study were performed under the support of the National Institute on Minority Health and Health Disparities through contract 75N91019F00132 with the National Cancer Institute. The National Cancer Institute had no input on the development of Quit Journey or its features or on study conceptualization or data analysis and interpretation. The efforts of MW, LT, and SE-T have been supported by the Division of Intramural Research of the National Institute on Minority Health and Health Disparities. Charmaine Chan, Kristyn Kamke, and Zahra Ansari reviewed the transcripts against the audio files. Rahaf Zaza extracted quotes in an exploratory phase of the analysis.

Funding

This research is supported (in part) by the Intramural Research Program of the National Institutes of Health (NIH; ZIAMD000011). The contributions of the NIH authors are considered works of the United States Government. The findings and conclusions presented in this paper are those of the authors and do not necessarily reflect the views of the NIH or the US Department of Health and Human Services.

Data Availability

All data generated or analyzed during this study are included in this published paper and its supplementary information files.

XSL•FO
RenderX

Authors' Contributions

Conceptualization: SE-T
Data curation: TG, EL, and SE-T
Formal analysis: MW and LT
Funding acquisition: SE-T
Methodology: SE-T
Project administration: BK and SE-T
Software: GB and GG
Writing—original draft: MW
Writing—review and editing: MW, LT, TG, EL, GB, GG, BK, and SE-T

Conflicts of Interest

GB, TG, GG, EL, and BK are employed by ICF NEXT. All other authors declare no other conflicts of interest.

Multimedia Appendix 1
Themes and illustrative quotes of participants' perceptions of performance for a Twitter/X message-based library and illustrative quotes of participants' suggestions for improving a Twitter/X message-based library.
[DOCX File, 68 KB - humanfactors_v12i1e76804_app1.docx]

References

1. US Department of Health and Human Services. Smoking cessation: a report of the surgeon general. : US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease, Prevention and Health Promotion, Office on Smoking and Health; 2020 URL: https://www.hhs.gov/sites/default/files/2020-cessation-sgr-full-report.pdf [accessed 2025-12-4]
2. Durkin S, Brennan E, Wakefield M. Mass media campaigns to promote smoking cessation among adults: an integrative review. Tob Control 2012 Mar;21(2):127-138. [doi: 10.1136/tobaccocontrol-2011-050345] [Medline: 22345235]
3. National Cancer Institute. Making health communication programs work. : US Department of Health and Human Services, Public Health Service, National Institutes of Health, National Cancer Institute; 2002 URL: https://www.iccp-portal.org/sites/default/files/resources/pink-book.pdf [accessed 2025-12-4]
4. Maar MA, Yeates K, Toth Z, et al. Unpacking the black box: a formative research approach to the development of theory-driven, evidence-based, and culturally safe text messages in mobile health interventions. JMIR Mhealth Uhealth 2016 Jan 22;4(1):e10. [doi: 10.2196/mhealth.4994] [Medline: 26800712]
5. Quinn-Scoggins HD, Murray RL, Quaife SL, et al. Co-development of an evidence-based personalised smoking cessation intervention for use in a lung cancer screening context. BMC Pulm Med 2022 Dec 15;22(1):478. [doi: 10.1186/s12890-022-02263-w] [Medline: 36522781]
6. Slater MD, Buller DB, Waters E, Archibeque M, LeBlanc M. A test of conversational and testimonial messages versus didactic presentations of nutrition information. J Nutr Educ Behav 2003;35(5):255-259. [doi: 10.1016/s1499-4046(06)60056-0] [Medline: 14521825]
7. Dudley MZ, Squires GK, Petroske TM, Dawson S, Brewer J. The use of narrative in science and health communication: a scoping review. Patient Educ Couns 2023 Jul;112:107752. [doi: 10.1016/j.pec.2023.107752] [Medline: 37068426]
8. So J. Counterproductive effects of overfamiliar antitobacco messages on smoking cessation intentions via message fatigue and resistance to persuasion. Psychol Addict Behav 2022 Dec;36(8):931-941. [doi: 10.1037/adb0000776] [Medline: 34914408]
9. Reinhard MA, Schindler S, Raabe V, Stahlberg D, Messner M. Less is sometimes more: How repetition of an antismoking advertisement affects attitudes toward smoking and source credibility. Soc Influ 2014 Apr 3;9(2):116-132. [doi: 10.1080/15534510.2013.790839]
10. Carpenter SM, Menictas M, Nahum-Shani I, Wetter DW, Murphy SA. Developments in mobile health just-in-time adaptive interventions for addiction science. Curr Addict Rep 2020 Sep;7(3):280-290. [doi: 10.1007/s40429-020-00322-y] [Medline: 33747711]
11. Naughton F. Delivering "just-in-time" smoking cessation support via mobile phones: current knowledge and future directions. Nicotine Tob Res 2017 Mar;19(3):379-383. [doi: 10.1093/ntr/ntw143] [Medline: 27235703]
12. Most popular social networks worldwide as of February 2025, by number of monthly active users. Statista. URL: https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/ [accessed 2025-12-04]
13. Daugherty T, Eastin MS, Bright L. Exploring consumer motivations for creating user-generated content. J Interact Advert 2008 Mar;8(2):16-25. [doi: 10.1080/15252019.2008.10722139]
14. Prochaska JJ, Pechmann C, Kim R, Leonhardt JM. Twitter=quitter? An analysis of Twitter quit smoking social networks. Tob Control 2012 Jul;21(4):447-449. [doi: 10.1136/tc.2010.042507] [Medline: 21730101]

XSL•FO
RenderX

15. Struik L, Khan S, Assoiants A, Sharma RH. Assessment of social support and quitting smoking in an online community forum: study involving content analysis. JMIR Form Res 2022 Jan 13;6(1):e34429. [doi: 10.2196/34429] [Medline: 35023834]
16. Ratcliff CL, Krakow M, Greenberg-Worisek A, Hesse BW. Digital Health Engagement in the US Population: Insights From the 2018 Health Information National Trends Survey. Am J Public Health 2021 Jul;111(7):1348-1351. [doi: 10.2105/AJPH.2021.306282] [Medline: 34014759]
17. Donaldson SI, Dormanesh A, Perez C, Majmundar A, Allem JP. Association between exposure to tobacco content on social media and tobacco use: a systematic review and meta-analysis. JAMA Pediatr 2022 Sep 1;176(9):878-885. [doi: 10.1001/jamapediatrics.2022.2223] [Medline: 35816331]
18. Kresovich A, Norris AH, Carter CC, Kim Y, Kostygina G, Emery SL. Deciphering influence on social media: a comparative analysis of influential account detection metrics in the context of tobacco promotion. Soc Media Soc 2024 Jan;10(1). [doi: 10.1177/20563051231224268] [Medline: 39575100]
19. Jo CL, Kornfield R, Kim Y, Emery S, Ribisl KM. Price-related promotions for tobacco products on Twitter. Tob Control 2016 Jul;25(4):476-479. [doi: 10.1136/tobaccocontrol-2015-052260] [Medline: 26130449]
20. Tamersoy A, De Choudhury M, Chau DH. Characterizing smoking and drinking abstinence from social media. HT ACM Conf Hypertext Soc Media 2015 Sep;2015:139-148. [doi: 10.1145/2700171.2791247] [Medline: 26640831]
21. Cole-Lewis H, Pugatch J, Sanders A, et al. Social listening: a content analysis of e-cigarette discussions on Twitter. J Med Internet Res 2015 Oct 27;17(10):e243. [doi: 10.2196/jmir.4969] [Medline: 26508089]
22. Fu R, Kundu A, Mitsakakis N, et al. Machine learning applications in tobacco research: a scoping review. Tob Control 2023 Jan;32(1):99-109. [doi: 10.1136/tobaccocontrol-2020-056438] [Medline: 34452986]
23. Rose SW, Jo CL, Binns S, Buenger M, Emery S, Ribisl KM. Perceptions of menthol cigarettes among Twitter users: content and sentiment analysis. J Med Internet Res 2017 Feb 27;19(2):e56. [doi: 10.2196/jmir.5694] [Medline: 28242592]
24. Reuter K, Wilson ML, Moran M, et al. General audience engagement with antismoking public health messages across multiple social media sites: comparative analysis. JMIR Public Health Surveill 2021 Feb 19;7(2):e24429. [doi: 10.2196/24429] [Medline: 33605890]
25. Rocheleau M, Sadasivam RS, Baquis K, et al. An observational study of social and emotional support in smoking cessation Twitter accounts: content analysis of tweets. J Med Internet Res 2015 Jan 14;17(1):e18. [doi: 10.2196/jmir.3768] [Medline: 25589009]
26. Naslund JA, Kim SJ, Aschbrenner KA, et al. Systematic review of social media interventions for smoking cessation. Addict Behav 2017 Oct;73:81-93. [doi: 10.1016/j.addbeh.2017.05.002] [Medline: 28499259]
27. Kim Y. Trust in health information websites: a systematic literature review on the antecedents of trust. Health Informatics J 2016 Jun;22(2):355-369. [doi: 10.1177/1460458214559432] [Medline: 25518944]
28. Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory: Pearson; 1986.
29. Green MC. Narratives and cancer communication. J Commun 2006 Aug 1;56(suppl_1):S163-S183. [doi: 10.1111/j.1460-2466.2006.00288.x]
30. Kreuter MW, Green MC, Cappella JN, et al. Narrative communication in cancer prevention and control: a framework to guide research and application. Ann Behav Med 2007 Jun;33(3):221-235. [doi: 10.1007/BF02879904] [Medline: 17600449]
31. de Graaf A, Sanders J, Hoeken H. Characteristics of narrative interventions and health effects: a review of the content, form, and context of narratives in health-related narrative persuasion research. Rev Commun Res 2016;4:88-131. [doi: 10.12840/issn.2255-4165.2016.04.01.011]
32. Coley HL, Sadasivam RS, Williams JH, et al. Crowdsourced peer- versus expert-written smoking-cessation messages. Am J Prev Med 2013 Nov;45(5):543-550. [doi: 10.1016/j.amepre.2013.07.004] [Medline: 24139766]
33. Durkin SJ, Biener L, Wakefield MA. Effects of different types of antismoking ads on reducing disparities in smoking cessation among socioeconomic subgroups. Am J Public Health 2009 Dec;99(12):2217-2223. [doi: 10.2105/AJPH.2009.161638] [Medline: 19833980]
34. Chou WYS, Prestin A, Lyons C, Wen KY. Web 2.0 for health promotion: reviewing the current evidence. Am J Public Health 2013 Jan;103(1):e9-18. [doi: 10.2105/AJPH.2012.301071] [Medline: 23153164]
35. Davis KC, Nonnemaker J, Duke J, Farrelly MC. Perceived effectiveness of cessation advertisements: the importance of audience reactions and practical implications for media campaign planning. Health Commun 2013;28(5):461-472. [doi: 10.1080/10410236.2012.696535] [Medline: 22812702]
36. Collins LM, Baker TB, Mermelstein RJ, et al. The multiphase optimization strategy for engineering effective tobacco use interventions. Ann Behav Med 2011 Apr;41(2):208-226. [doi: 10.1007/s12160-010-9253-x] [Medline: 21132416]
37. Landoll RR, Vargas SE, Samardzic KB, Clark MF, Guastaferro K. The preparation phase in the multiphase optimization strategy (MOST): a systematic review and introduction of a reporting checklist. Transl Behav Med 2022 Feb 16;12(2):291-303. [doi: 10.1093/tbm/ibab146] [Medline: 34850214]
38. Cornelius ME, Loretan CG, Jamal A, et al. Tobacco product use among adults - United States, 2021. MMWR Morb Mortal Wkly Rep 2023 May 5;72(18):475-483. [doi: 10.15585/mmwr.mm7218a1] [Medline: 37141154]
39. Galobardes B, Shaw M, Lawlor DA, Lynch JW, Davey Smith G. Indicators of socioeconomic position (part 1). J Epidemiol Community Health 2006 Jan;60(1):7-12. [doi: 10.1136/jech.2004.023531] [Medline: 16361448]

XSL•FO
RenderX

40. Shavers VL. Measurement of socioeconomic status in health disparities research. J Natl Med Assoc 2007 Sep;99(9):1013-1023. [Medline: 17913111]
41. Adler NE, Newman K. Socioeconomic disparities in health: pathways and policies. Health Aff (Millwood) 2002;21(2):60-76. [doi: 10.1377/hlthaff.21.2.60] [Medline: 11900187]
42. Hennink M, Kaiser BN. Sample sizes for saturation in qualitative research: a systematic review of empirical tests. Soc Sci Med 2022 Jan;292:114523. [doi: 10.1016/j.socscimed.2021.114523] [Medline: 34785096]
43. Saunders B, Sim J, Kingstone T, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018 Jul;52(4):1893-1907. [doi: 10.1007/s11135-017-0574-8]
44. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
45. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003 Sep 1;27(3):425-478. [doi: 10.2307/30036540]
46. Venkatesh V, Thong JYL, Xu X. Consumer acceptance and use of information technology: extending the Unified Theory of Acceptance and Use of Technology. MIS Q 2012 Mar 1;36(1):157-178. [doi: 10.2307/41410412]
47. Wakeman M, Tesfaye L, Gregory T, Leahy E, Kendrick B, El-Toukhy S. Perceptions of the use of mobile technologies for smoking cessation: focus group study with individuals of low socioeconomic status who smoke. JMIR Form Res 2024 Oct 11;8:e58221. [doi: 10.2196/58221] [Medline: 39392684]
48. Tesfaye L, Wakeman M, Baskin G, et al. A feature-based qualitative assessment of smoking cessation mobile applications. PLOS Digit Health 2024 Nov;3(11):e0000658. [doi: 10.1371/journal.pdig.0000658] [Medline: 39571041]
49. Wakeman M, Tesfaye L, Baskin G, et al. Perceptions toward an attentional bias modification mobile game among individuals with low socioeconomic status who smoke: qualitative study. JMIR Serious Games 2025 May 9;13:e59515. [doi: 10.2196/59515] [Medline: 40344199]
50. Tesfaye L, Wakeman M, Baskin G, et al. Acceptability of a carbon monoxide monitoring feature of a smoking cessation mobile application: a virtual focus group study. BMJ Public Health 2025 [FREE Full text] [doi: 10.1136/bmjph-2024-002354]
51. Tesfaye L, Wakeman M, Gregory T, et al. Acceptability of a smart lighter for tracking cigarette smoking: A focus group study. Digit HEALTH 2025;11:20552076251323998. [doi: 10.1177/20552076251323998] [Medline: 40151637]
52. Braun V, Clarke V. Thematic analysis. In: Cooper H, Coutanche MN, McMullen LM, editors. APA Handbook of Research Methods in Psychology: Foundations, Planning, Measures, and Psychometrics/Research Designs: Quantitative, Qualitative, Neuropsychological, and Biological/Data Analysis and Research Publication: American Psychological Association; 2012:57-71. [doi: 10.1037/13620-004]
53. Friese S. ATLAS.ti 8 Windows User Manual. : ATLAS.ti Scientific Software Development GmbH; 2020 URL: https://atlasti.com/manuals-and-documents [accessed 2025-12-04]
54. Vilardaga R, Casellas-Pujol E, McClernon JF, Garrison KA. Mobile applications for the treatment of tobacco use and dependence. Curr Addict Rep 2019 Jun;6(2):86-97. [doi: 10.1007/s40429-019-00248-0] [Medline: 32010548]
55. Imtiaz MH, Ramos-Garcia RI, Wattal S, Tiffany S, Sazonov E. Wearable sensors for monitoring of cigarette smoking in free-living: a systematic review. Sensors (Basel) 2019 Oct 28;19(21):4678. [doi: 10.3390/s19214678] [Medline: 31661856]
56. Cavazos-Rehg PA, Krauss MJ, Sowles SJ, Bierut LJ. "Hey everyone, I'm drunk." an evaluation of drinking-related Twitter chatter. J Stud Alcohol Drugs 2015 Jul;76(4):635-643. [doi: 10.15288/jsad.2015.76.635] [Medline: 26098041]
57. Cavazos-Rehg PA, Krauss M, Fisher SL, Salyer P, Grucza RA, Bierut LJ. Twitter chatter about marijuana. J Adolesc Health 2015 Feb;56(2):139-145. [doi: 10.1016/j.jadohealth.2014.10.270] [Medline: 25620299]
58. Moyer-Gusé E. Toward a theory of entertainment persuasion: explaining the persuasive effects of entertainment-education messages. Commun Theory 2008 Aug;18(3):407-425. [doi: 10.1111/j.1468-2885.2008.00328.x]
59. Durkin SJ, Wakefield MA, Spittal MJ. Which types of televised anti-tobacco campaigns prompt more quitline calls from disadvantaged groups? Health Educ Res 2011 Dec;26(6):998-1009. [doi: 10.1093/her/cyr048] [Medline: 21715652]
60. Whittaker R, McRobbie H, Bullen C, Rodgers A, Gu Y, Dobson R. Mobile phone text messaging and app-based interventions for smoking cessation. Cochrane Database Syst Rev 2019 Oct 22;10(10):CD006611. [doi: 10.1002/14651858.CD006611.pub5] [Medline: 31638271]
61. Chen T, Samaranayake P, Cen X, Qi M, Lan YC. The impact of online reviews on consumers' purchasing decisions: evidence from an eye-tracking study. Front Psychol 2022;13:865702. [doi: 10.3389/fpsyg.2022.865702] [Medline: 35756238]
62. Bae S, Lee T. Product type and consumers' perception of online consumer reviews. Electron Markets 2011 Dec;21(4):255-266. [doi: 10.1007/s12525-011-0072-0]
63. Whittaker R, Dobson R, Garner K. Chatbots for smoking cessation: scoping review. J Med Internet Res 2022 Sep 26;24(9):e35556. [doi: 10.2196/35556] [Medline: 36095295]
64. Prutzman YM, Wiseman KP, Grady MA, et al. Using digital technologies to reach tobacco users who want to quit: evidence from the National Cancer Institute's Smokefree.gov Initiative. Am J Prev Med 2021 Mar;60(3 Suppl 2):S172-S184. [doi: 10.1016/j.amepre.2020.08.008] [Medline: 33663705]
65. Chaiton M, Diemert L, Cohen JE, et al. Estimating the number of quit attempts it takes to quit smoking successfully in a longitudinal cohort of smokers. BMJ Open 2016 Jun;6(6):e011045. [doi: 10.1136/bmjopen-2016-011045]

XSL•FO
RenderX

66. Koslovsky MD, Hébert ET, Swartz MD, et al. The time-varying relations between risk factors and smoking before and after a quit attempt. Nicotine Tob Res 2018 Sep 4;20(10):1231-1236. [doi: 10.1093/ntr/ntx225] [Medline: 29059413]
67. Hoeppner BB, Hoeppner SS, Seaboyer L, et al. How smart are smartphone apps for smoking cessation? A content analysis. Nicotine Tob Res 2016 May;18(5):1025-1031. [doi: 10.1093/ntr/ntv117] [Medline: 26045249]
68. Ybarra ML, Prescott TL, Holtrop JS. Steps in tailoring a text messaging–based smoking cessation program for young adults. J Health Commun 2014 Dec;19(12):1393-1407. [doi: 10.1080/10810730.2014.901441] [Medline: 24766267]
69. Tannenbaum MB, Hepler J, Zimmerman RS, et al. Appealing to fear: a meta-analysis of fear appeal effectiveness and theories. Psychol Bull 2015 Nov;141(6):1178-1204. [doi: 10.1037/a0039729] [Medline: 26501228]
70. Witte K, Allen M. A meta-analysis of fear appeals: implications for effective public health campaigns. Health Educ Behav 2000 Oct;27(5):591-615. [doi: 10.1177/109019810002700506] [Medline: 11009129]
71. Emery SL, Szczypka G, Abril EP, Kim Y, Vera L. Are you scared yet?: evaluating fear appeal messages in tweets about the tips campaign. J Commun 2014 Apr;64:278-295. [doi: 10.1111/jcom.12083] [Medline: 25429162]
72. McAfee T, Davis KC, Alexander RL, Pechacek TF, Bunnell R. Effect of the first federally funded US antismoking national media campaign. Lancet 2013 Dec 14;382(9909):2003-2011. [doi: 10.1016/S0140-6736(13)61686-4] [Medline: 24029166]
73. Skubisz C, Miller GA, Hinsberg L, Kaur S, Miller GA. Tips from former smokers: a content analysis of persuasive message features. Int Q Community Health Educ 2016 Oct;37(1):13-20. [doi: 10.1177/0272684X16685253] [Medline: 30238856]
74. O'Keefe DJ, Jensen JD. The relative persuasiveness of gain-framed and loss-framed messages for encouraging disease prevention behaviors: a meta-analytic review. J Health Commun 2007;12(7):623-644. [doi: 10.1080/10810730701615198] [Medline: 17934940]
75. Gallagher KM, Updegraff JA. Health message framing effects on attitudes, intentions, and behavior: a meta-analytic review. Ann Behav Med 2012 Feb;43(1):101-116. [doi: 10.1007/s12160-011-9308-7] [Medline: 21993844]
76. Myrick J. The Role of Emotions in Preventative Health Communication, 1st edition: Lexington Books; 2015. [doi: 10.5771/9780739191484]
77. Petty RE, Cacioppo JT. The elaboration likelihood model of persuasion. In: Berkowitz L, editor. Advances in Experimental Social Psychology: Academic Press; 1986:123-205. [doi: 10.1016/S0065-2601(08)60214-2]
78. Perski O, Kwasnicka D, Kale D, et al. Within-person associations between psychological and contextual factors and lapse incidence in smokers attempting to quit: a systematic review and meta-analysis of ecological momentary assessment studies. Addiction 2023 Jul;118(7):1216-1231. [doi: 10.1111/add.16173] [Medline: 36807443]
79. Perski O, Li K, Pontikos N, et al. Classification of lapses in smokers attempting to stop: a supervised machine learning approach using data from a popular smoking cessation smartphone app. Nicotine Tob Res 2023 Jun 9;25(7):1330-1339. [doi: 10.1093/ntr/ntad051] [Medline: 36971111]
80. Jakob R, Harperink S, Rudolf AM, et al. Factors influencing adherence to mHealth apps for prevention or management of noncommunicable diseases: systematic review. J Med Internet Res 2022 May 25;24(5):e35371. [doi: 10.2196/35371] [Medline: 35612886]
81. Bakaloudi DR, Evripidou K, Siargkas A, Breda J, Chourdakis M. Impact of COVID-19 lockdown on smoking and vaping: systematic review and meta-analysis. Public Health 2023 May;218:160-172. [doi: 10.1016/j.puhe.2023.02.007] [Medline: 37043948]
82. Prensky M. Digital natives, digital immigrants part 1. On Horizon 2001 Sep;9(5):1-6. [doi: 10.1108/10748120110424816]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
SES: socioeconomic status

Edited by A Kushniruk; submitted 30.04.25; peer-reviewed by HS Bernardino, S Baig; revised version received 31.08.25; accepted 16.10.25; published 17.12.25.

Please cite as:
Wakeman M, Tesfaye L, Gregory T, Leahy E, Baskin G, Gruse G, Kendrick B, El-Toukhy S
Perceptions of User-Generated Content as a Source of Health Messages in Smoking Cessation Mobile Interventions: Focus Group Study
JMIR Hum Factors 2025;12:e76804
URL: https://humanfactors.jmir.org/2025/1/e76804
doi:10.2196/76804

© Michael Wakeman, Lydia Tesfaye, Tim Gregory, Erin Leahy, Gunnar Baskin, Greg Gruse, Brandon Kendrick, Sherine El-Toukhy. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 17.12.2025. This is an open-access

XSL•FO
RenderX

article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Improving Mental Health Referral Systems in Rural Australia: Co-Design Study With Health Professionals and Consumers

Kate Bartel[1,2*], PhD; Asini Malinka Siriwardene[1*], BSc; Paul Worley[1,2], MBBS, PhD; Kim Pearce[2], BSc; Niranjan Bidargaddi[3], PhD; Bronwin Patrickson[3], PhD; Simon Moody[1], BN; Sharon Wingard[1]; Martin Jones[4,5], MSc, DN; Brian McKenny[6], BMBS; Darryl Cameron[7†], MPH; Sharon Lawn[2,8], MSW, PhD; Amy E Mendham[1,2], PhD

[1]Riverland Academy of Clinical Excellence, Riverland Mallee Coorong Local Health Network, 10 Maddern Street, Berri, Australia

[2]College of Medicine and Public Health, Rural and Remote Health, Flinders University, Adelaide, Australia

[3]College of Medicine and Public Health, Flinders Health & Medical Research Institute, Digital Health Research Lab, Flinders University, Adelaide, Australia

[4]School of Nursing and Midwifery, Edith Cowan University, Bunbury, Australia

[5]Department of Rural Health, University of South Australia, Whyalla, Australia

[6]Barossa Hills Fleurieu Local Health Network, Angaston, Australia

[7]Moorundi Aboriginal Community Controlled Health Service Inc, Murray Bridge, Australia

[8]Lived Experience Australia, Adelaide, Australia

[†]deceased

[*]these authors contributed equally

Corresponding Author:
Amy E Mendham, PhD
Riverland Academy of Clinical Excellence, Riverland Mallee Coorong Local Health Network, 10 Maddern Street, Berri, Australia

Abstract

Background: In rural Australia, geographical isolation, limited resources, and complex health care navigation create significant barriers to mental health care access. Mental health care professionals and organizations often work in segregation, exacerbating existing barriers. Digital technology provides an opportunity to improve communication between providers and streamline workflows while supporting a diverse range of consumers.

Objective: This co-design study aimed to identify rural community needs and explore digital solutions to enhance mental health service delivery pathways.

Methods: Using a design-thinking methodology, we conducted focus groups and workshops with 17 participants (7 consumers and caregivers and 10 health care professionals) from a rural region to understand mental health service needs, systemic challenges, and design potential digital solutions. Thematic analysis followed a grounded theory approach, involving systematic coding and theme development through an iterative consensus process.

Results: Access to mental health care emerged as the central theme. Rural community participants reported strong community connections but faced challenges, including limited technological innovation and substantial travel burdens. Health care professionals highlighted critical systemic pressures: underresourcing, overwhelmed clinicians with extensive waitlists, and complex referral processes. Both groups identified overlapping barriers in service limitations and system navigation. During the design phase, we developed personas capturing consumer and health care professional experiences and conceptualized an integrated digital solution comprising a health care professional dashboard and a consumer-facing app with caregiver access to enhance service coordination.

Conclusions: The study demonstrated strong stakeholder support for implementing an integrated digital solution to enhance rural mental health service delivery. Further research is required to build upon the solution prior to testing, optimizing, and scaling.

(JMIR Hum Factors 2025;12:e73460) doi:10.2196/73460

KEYWORDS

co-design; digital solution; focus group; referral; mental health; rural health care

XSL•FO
RenderX

Introduction

Mental health burden has significant economic and social impacts [1], particularly in rural areas [2,3], which typically have fewer trained professionals, limited services, and less service innovations compared to metropolitan areas [3-6]. In Australia, 7 million rural residents face substantial barriers to mental health care access. These obstacles include fewer resources, limitations in technology, lack of knowledge of services, distance to services, and a deficiency in culturally sensitive practices [7], with complex and often ineffective treatment pathways [3].

General practitioners (GPs), as the primary contact for mental health support [8,9], face considerable challenges when connecting patients with appropriate mental health care, primarily due to time constraints in determining suitable referrals and limited availability of services in rural communities [10,11]. Adding to the complexity, service availability, fee structure, and communication methods vary between providers and change frequently, reflecting poor integration between services. This exacerbates the problem caused by inefficiencies and complexities in the health care system, decreasing the efficiency of care provision and inadvertently resulting in potentially less effective care [2].

The Australian government has provided the Initial Assessment and Referral Decision Support Tool (IAR-DST), a clinical decision support tool to assist with mental health care level determination. While IAR-DST determines care need, clinicians require additional support to rapidly identify suitable local providers and optimize referral processes [12]. Without such support, rural GPs may struggle to initiate the right referral on their first attempt where limited service availability already constrains their options. This may lead to rework, delayed treatment, and patient disengagement [13]. Moreover, many rural Australian people forgo seeking support altogether due to the difficulties in accessing GPs, so the importance of getting the referral processes right on the first attempt is even more imperative [10].

Digital referral platforms could augment existing assessment tools by automating provider matching and streamlining referral workflows [14]. Such systems show promise in reducing administrative load while supporting culturally and linguistically diverse consumers, for example, through better matching of services and more informed options [15]. However, their utility in rural Australian settings remains unexplored. Awareness of the needs and context of the community, what resources are available, and how their systems operate is invaluable [5,8]. Thus, to explore whether this could be a useful strategy, the community in which it would be implemented must be involved in the co-design process.

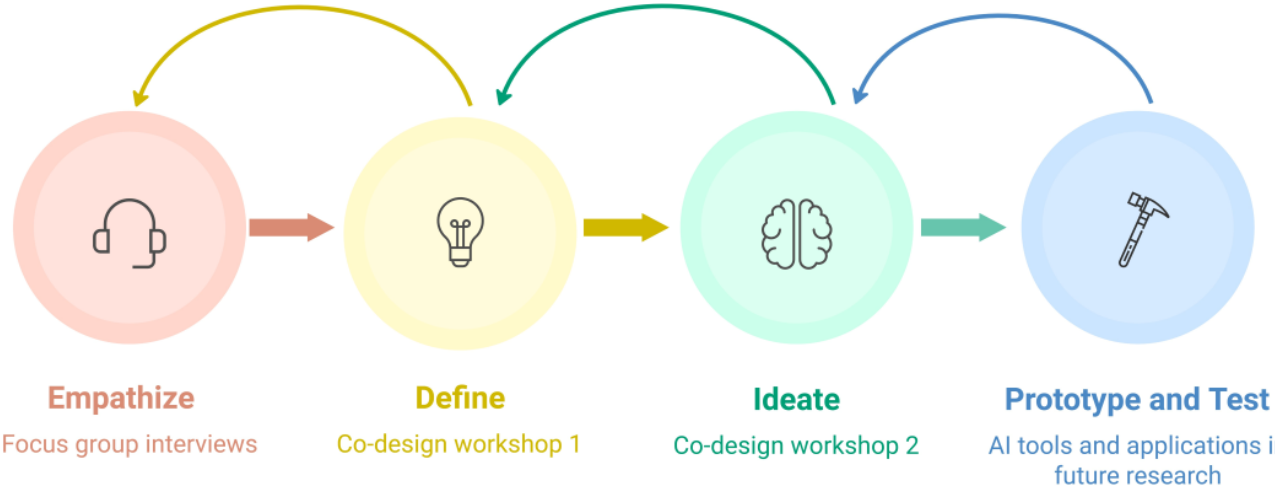
The primary aim of this formative study was to explore digital solutions for mental health care referral pathways in a rural community using participatory and human-centered design-thinking methods. Specifically, this study explored if the rural community wants a digital solution for mental health referral pathways, how it would be useful, and who would benefit.

Methods

Study Design

Participatory and human-centered design-thinking methodologies informed the structure and approach for all participant activities (Figure 1). Design-thinking methodology is an iterative, user-centric process that aims to create a solution that is desirable, feasible, and viable [16] to end users, stakeholders, and the supporting infrastructure. Accordingly, consumers' and health care professionals' lived experiences, preferences, and needs are prioritized to develop a technological solution to improve referral efficiencies in rural communities.

Figure 1. Participatory and human-centered design-thinking methodologies (adapted from NNGroup: Design thinking 101). AI: artificial intelligence.



Design-Thinking Methodology

Design-thinking phases were reflected in the study design:

1. Empathize: Semistructured focus groups with consumers and health care professionals to develop a deep understanding of the rural community needs.

XSL•FO
RenderX

2. Define: A workshop to start the co-design process with combined consumer and health care professional input to prioritize the pain points in the mental health care system.
3. Ideate: A second workshop with consumers and health care professionals to brainstorm possible solutions.

Setting

The study took place in the Riverland, South Australia, which is a small, rural region with a high level of rurality (Modified Monash Model 5), which entails considerable geographical remoteness and smaller population sizes. The Riverland area represents a typical low-resourced and low socioeconomic rural context with a Socio-Economic Indexes for Areas score of 996.8, indicating a level of socioeconomic conditions that falls below the national average.

Participants

Consumers were recruited through researcher and clinician networks and included those with lived experience of mental health issues and formal or informal caregivers of people who experienced mental health issues. To be eligible to participate, consumers needed to be 18 years of age or older and have interacted with a health service regarding their mental health, or the mental health of a person they cared for, within the past 24 months. Health care professionals were recruited using purposeful sampling. Those who worked in the Riverland region in South Australia and had experience working with mental health consumers in the previous 24 months were invited to participate. The services approached included all public health sites, private mental health services, nongovernment organizations (NGO), charities, and Aboriginal services. Prospective participants were contacted via email or telephone and informed of the nature and purpose of the study, prior to consenting.

Ethical Considerations

This project was approved by the Department of Health and Wellbeing Ethics Committee (2023/HRE00227) and the Aboriginal Health and Research Ethics Committee (04-23-1104). All participants provided written consent prior to their inclusion in the study and were remunerated for their time according to South Australian Health Department guidelines. Participants' data were deidentified and stored on a secure electronic server.

Data Collection

Data collection occurred from June to September 2024.

Empathize: Focus Groups

All focus groups followed a semistructured format as provided in Multimedia Appendix 1. Consumers (n=7 - 10) participated in two 90-minute, face-to-face focus group discussions that were facilitated by experienced coinvestigators (KP and SL), including a lived experience representative (SL). The first session explored past experiences with mental health care and expectations of mental health management. The second focus group discussion explored current digital engagement behaviors and willingness to explore digital solutions. Health care professionals (n=7 - 10) participated in two 60-minute focus group discussions conducted on the web (via Microsoft Teams) with the same facilitators. The first session explored the coordination of care and navigation across the mental health system. The second focus group discussion centered on experiences with technology as a solution and willingness to explore further a digital solution.

Following the completion of the focus group interviews, all participants were invited to participate in two 3-hour face-to-face design workshops. The workshop was facilitated by 2 external design-thinking experts. A consumer representative from a consumer and caregiver mental ill-health advocacy group (SL) and members of the research team with a background in psychology or focus group research assisted (KB, AEM, and AMS). All face-to-face focus groups and workshops were held in a local hospital meeting room.

Define: Design Workshop 1

Participants were asked to choose from consumer scenarios presented (as outlined in the Results section). For each scenario, consumers and health care professionals worked together in groups (n=4 - 6) to map the key interactions that happen along the consumer mental health "journey." Following the mapping process, these interactions and timelines were discussed with the entire group to fill in any gaps and highlight barriers in the system (pain points).

Ideate: Design Workshop 2

Key journey moments across the scenarios in workshop 1 were prioritized for conversation in workshop 2. Using these "key moments," the room was split into 2 groups (1 group of consumers and 1 group of health care professionals). Each group discussed the pain points in the referral pathways and identified possible technological solutions. Following these small group discussions, the core pain points and technological solutions were discussed with the entire group. Following on from earlier conversations, participants were provided with a scope to focus on interactions between using mobile phone apps and health care professional–facing dashboards.

Data Analysis

Focus groups were recorded, transcribed, and deidentified. Field notes were written by AMS and KB. Data were thematically coded using qualitative analysis software NVivo (version 14; Lumivero), guided by a grounded theory approach. The thematic coding process was managed by AMS and KB, in discussion with AEM, BP, and SL. Open coding was performed by AMS and checked for accuracy by KB, with any discrepancies discussed and resolved through consensus. Axial and selective coding were subsequently conducted by AEM, AMS, BP, KB, and SL [17]. Saturation of data was discussed (ie, no new themes were being generated by the data). Themes and subthemes were finalized by AEM, AMS, BP, KB, and SL. This study followed the COREQ (Consolidated Criteria for Reporting Qualitative Research) reporting guidelines, ensuring transparency in reporting researcher reflexivity, study design, and data analysis (Checklist 1) [18].

The iterative process of design-thinking methodology was implemented throughout the study, with data collected and analyzed in each phase to advise the next phases of data collection. Participants were provided with a verbal summary

XSL•FO
RenderX

of previous session content, at the commencement of the session, and given the chance to comment or provide correction. Preliminary analyses of the focus groups were discussed with the research team, and external user-experience designers advised on the planning of co-design workshop 1. Participant outputs were digitally collated, and user personas were developed. User personas are a fictional representation of real-life users who share common characteristics, behaviors, goals, and pain points [19,20]. This effectively narrows down the intended users of the platforms and creates a more targeted approach, reflecting the "define" stage of the design-thinking process. Moreover, it creates a more detailed representation of the consumers to inform effective and targeted design decisions while maintaining the anonymity of the real study participants and reflecting their experiences accurately. Insights from co-design workshop 2 were similarly digitized and analyzed to identify key themes, which were then grouped to uncover opportunities. These opportunities were mapped back to the user personas to ensure that the solutions created serve the user.

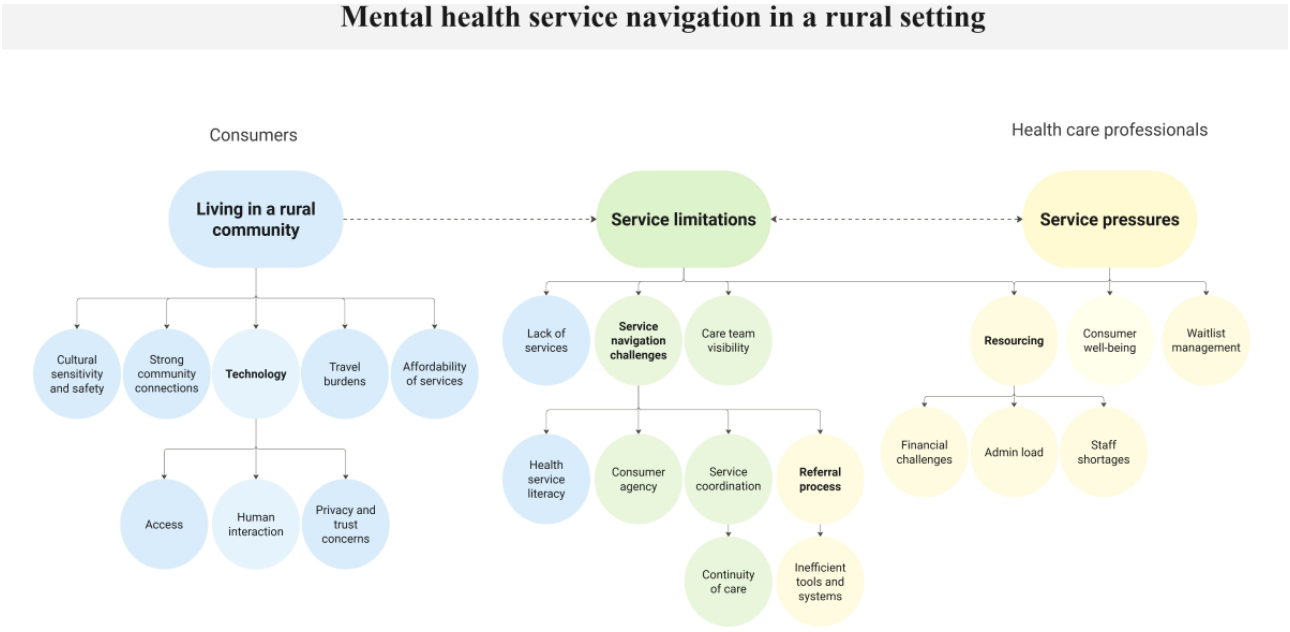
Results

Overview

A total of 7 consumers and 10 health care professionals participated in focus groups and design workshops. In total, 10 consumers were contacted; 2 were unavailable, 1 did not attend, providing no notice, and 7 participated in at least 1 focus group or workshop. A total of 15 health professionals or organizations were contacted. In total, 3 professionals were unavailable, 2 did not return contact, and 10 attended a minimum of 1 focus group or workshop. Health care professionals included a psychiatrist, GPs (n=2), a clinical psychologist in private practice, NGOs and nonprofit managers (n=2), mental health nurses (n=2), and an Aboriginal nurse unit manager. Of the consumers, 5 had experienced a mental health issue, and 2 cared for someone with a mental health issue in the previous 24 months. One of the 2 Aboriginal consumers was an Aboriginal Elder.

The unifying theme identified across consumer and health care professional focus groups was "access to mental health care in a rural community." A coding tree displays an overview of the main themes and subthemes identified from the focus group discussion (Figure 2), with example quotes provided in Table 1.

Figure 2. Coding tree of the main themes and subthemes identified from the focus group discussions.



XSL•FO
RenderX

Table . Themes and subthemes from focus groups.

Themes and subthemes	Quotes
Theme 1: living in a rural community (consumer identified)	
1.1. Cultural sensitivity and safety	• "I heard a whisper, they're getting a morgue up here. I don't know how true this is. But they want to know how to make it culturally appropriate. So they're getting someone from Adelaide to come up here to tell us" [Consumer 2].
1.2. Strong community connections	• "We are just one really, really big family" [Consumer 4—in relation to community group members in the Riverland].
1.3. Travel burdens	• "Sometimes it's transport because that's terrible up here" [Consumer 1—assisting people with mental health support].
1.4. Affordability of services	• "Clients don't have money to pay to see a GP" [Health care professional 1].
1.5 Technology	
1.5.1. Access	• "But I'm happy to use my phone. Get my SMSs when I've got appointments, they text me say your appointment" [Consumer 4]. • "A lot of people who use our services don't have smart phones. Or if they do, they have very limited capabilities and often don't have access to internet" [Health care professional 2]. • "I don't know how to use a computer. But I know how to use my phone" [Consumer 4].
1.5.2. Human interaction	• "I'd rather stand in a line at a grocery store rather than using the checkout because I like face-to-face contact" [Consumer 3—regarding worries about using technology]. • "The least human interaction I can have sometimes, the better. If my social battery is full, empty actually, I'll opt for that" [Consumer 5 regarding a fuel app that allows users to pay from their phone at the bowser].
1.5.3. Privacy and trust concerns	• "Could this not just be—instead of, 'How's your Wednesday been? Let's write this. What are your good thoughts, bad thoughts, gratitudes. Something to work on tomorrow.' Instead of it being so invasive, could it not be a basic K10 every three weeks? Or just a check-in every so often to see how you're travelling, instead of being an every single day thing" [Consumer 5].
Theme 2: service limitations (consumer and health care professional identified)	
2.1. Lack of services	• "Up here, if you need to get referred onto any specialist service, like a psychiatrist or a psychologist. They're just not available and you just can't make an appointment with one" [Consumer 1]. • "Specialist services not being available in Riverland" [Health care professional 6—in relation to the biggest pressure point they face in the Riverland with mental health services].
2.2. Care team visibility	• "It would be helpful to know who is part of this client's care team" [Health care professional 1].
2.3. Service navigation challenges	
2.3.1. Health service literacy	• "I'm still not aware of all the services here or what's available" [Consumer 3].
2.3.2. Consumer agency	• "The person's got to make a choice. It has to be their choice" [Consumer 1]. • "Some sort of system that would involve the client, so that information that was being shared was accessible by the client and was maybe even directed by the client" [Health care professional 1].

XSL•FO
RenderX

Themes and subthemes	Quotes
2.3.3. Service coordination	• "It would be really helpful in terms of me then, if I could somehow use that system to communicate with their GP or other people in their team" [Health care professional 5].
2.3.3.1. Continuity of care	• "Not getting your own GP when you want to go and see the doctor" [Consumer 4].
2.3.4. Referral process	• "When you look at how many services are in town, I don't know what they all are. I don't know who's referred people to whom" [Consumer 1]. • "Yes and if they require an actual official referral" [Consumer 5 in response to Consumer 1].
2.3.4.1. Inefficient tools and systems	• "Depending on the GP, I find it varying levels of useful" [Health care professional 5—addressing referrals they receive from general practitioners]. • "The other thing is that when the clinician who's assessing the patient at the intake, if that assessment is just limited to a tool or is too superficial or is mostly looking at past records. Especially for risk assessment, then usually they get it wrong" [Health care professional 6].
Theme 3: service pressures (health care professional identified)	
3.1. Consumer well-being	• "A lot of clients are completely disengaged with mental health services, despite them being significantly unwell" [Health care professional 2]. • "I think it would be beneficial for lots of family and carers to be involved and have access to the apps, because they're obviously usually the first to know when people are deteriorating. It gives them resources too, how to navigate the step model of care for their person" [Health care professional 1].
3.2. Waitlist management	• "I find waitlist management really hard, because it's impossible for me to predict how long someone's going to need services for" [Health care professional 5].
3.3. Resourcing	• "This hospital's constantly bed blocked now. I work closely with [psychiatrist] and the on-call psychiatrist to determine whether community treatment would be more beneficial" [Health care professional 3].
3.3.1. Financial challenges	• "Or finances which is obviously what [Health care professional 6] has sort of alluded to as well" [Health care professional 4—in response to the biggest mental health service challenges faced].
3.3.2. Admin load	• "It's another administrative job that nobody wants to do" [Health care professional 1].
3.3.3. Staff shortages	• "Recruitment and retention of staff is a big one" [Health care professional 4—in response to the biggest mental health service challenges faced].

Empathize: Focus Groups

Overview

Consumer insights were primarily shaped by the experience of living in a rural community. Such experiences were underpinned by the effects of geographical isolation and lower population densities, often resulting in service limitations. These limitations may, in turn, stem from broader service pressures faced by health care professionals such as staff shortages and a lack of resources, contributing to the unique challenges faced by rural communities in accessing mental health care.

Living in a Rural Community

In spite of the challenges faced, consumers living in rural settings had strong community connections, describing themselves as "one really, really big family" (Consumer 4), feeling a sense of pride in residing in the country. Aboriginal participants echoed these sentiments by having support networks to care for those with mental health issues. Regardless of strong community connections, trust in health care systems was varied, with Aboriginal participants less likely to trust in health care systems. As an Aboriginal Elder observed: "That's where it's

XSL•FO
RenderX

different for Aboriginal people. The health system was where they stole kids" (Consumer 2).

Historically, Aboriginal people in Australia have faced systemic discrimination [21], leading to distrust in the health care system, underscoring the importance of cultural safety and sensitivity when developing the technological solution. Distrust in health care systems may also extend to distrust in technological products issued by these systems, as corroborated by consumers, regardless of Aboriginality. Lacking cultural and rural appropriateness, privacy, and data mishandling concerns may contribute to this varied trust.

When asked regarding the general experience of using digital technology tools, consumers raised concerns about accessibility, as some were unable to afford smartphones, phone credit, or had limited digital literacy.

Consumers reinforced that rural mental health services were harder to access due to affordability or geographic location, often requiring numerous visits to metropolitan Adelaide for specialist services. Of the services that were available in the Riverland, most were said to have long wait times, or consumers were not aware of their existence.

Service Limitations and Potential Solutions

Health care professionals and consumers voiced that available mental health services were hard to find and difficult to navigate for consumers in rural areas. Despite being a lifelong resident in the Riverland, a consumer was "still not aware of all the services here or what's available" (Consumer 1), and another reinforced that it would be helpful to "create some kind of a roadmap to the services" (Consumer 5). This lack of health service literacy may be a by-product of ineffective communication and coordination between mental health services, GPs, and hospital services. Participants identified that effectively increasing interservice communication can lead to a better continuity of care, improving mental health outcomes for patients.

Care team visibility and effective communication may assist in avoiding duplication of services, as "there's a lot of overservicing that happens in [the mental healthcare] space" (Health care professional 1), thus reducing workloads and service limitations. Consumers suggested that this may empower and encourage them to actively participate in their care while reducing the risk of confusion and misunderstanding. Consumers felt strongly regarding patient choice, having the necessary information to make an informed decision regarding their care. In this respect, the role of the caregiver was also emphasized by both consumers and health care professionals, as caregivers can advocate for patients when they cannot do so themselves. Caregivers are also "usually the first to know when [patients] are deteriorating" (Health care professional 1). Health care professionals respectively advocated for the consumers, valuing the importance of care "directed by the client" (Health care professional 1) and sharing information with the consumers.

Service Pressures and Potential Solutions

The predominant challenges and frustrations for health care professionals were categorized under the central theme of service pressures. Services were impacted by limited resource availability, which included a lack of finances for infrastructure and increased administrative loads due to staff shortages and attrition.

Most mental health services in the Riverland were characterized by prolonged waiting periods for patients, as providers were overburdened. Methods for the management of waitlists varied between health care providers. Participants suggested that if consumers were able to inform the health care service of their symptoms improving or worsening, health care professionals would be able to reprioritize consumers and better manage waitlists.

While both consumers and health care professionals agreed that displaying waitlists in the technology tool would be quite helpful, services would struggle to update waitlists regularly, as it would be the "first thing that slips when practices get busy" (Health care professional 1).

Health care professionals unanimously agreed that referral processes between services were underoptimized and "siloed" (Consumer 1), significantly impacting service delivery. Each service used varied processes and systems for referral and interservice communication, and most relied on inefficient and archaic tools such as fax machines: "It can actually delay consumer care because the fax might not be working, or got rejected" (Health care professional 3).

All health care professionals except 1 were unfamiliar with the IAR-DST, preferring practical, more familiar tools that allow flexibility in clinical judgment. Clinician discrepancies in tool use were often not visible to other health care professionals, impacting the efficiency of the referral process. In a similar vein, out-of-date waitlists and service directories place pressure on GPs to efficiently refer patients. For developing the technological solution, health care professionals prioritized consumer advocacy and recognized the importance of addressing consumer pain points in improving service provision.

Define: Design Workshop 1

Overview

Workshop 1 was conducted with mixed groups of health care professionals and consumers. Based on the preliminary insights from focus groups in the empathize stage, the following consumer scenarios were developed.

1. An older person living alone in an isolated rural location with limited family support, moderate mental health challenges, and no internet access.
2. A young adult living and working in town with severe mental health challenges requiring frequent journeys between metro services and rural community support services.
3. A single parent with mild mental health challenges living rurally and accessing GP and NGO supports.
4. A young person attending a rural high school, accessing services and supports following an acute mental health admission at the local hospital.
5. A farmer living on a rural property, experiencing isolation and hesitant to seek support for his mental health challenges.

XSL•FO
RenderX

Participants chose scenarios 1, 2, and 3 for the journey mapping exercise. Each step of the consumer journey was mapped with the current channel that the consumer engages in, the value of this interaction, the software capabilities to replace or augment, the value that could be added from this augmentation, and the outcome for the consumer.

Scenario 1

An older person living alone in an isolated rural location with limited family support, moderate mental health challenges, and no internet access.

The role of the caregiver was particularly stressed in this scenario due to the consumer's limited access to technology. All participants agreed that the caregivers should be given access to the technological solution at the discretion of the consumer. The caregiver will thus be able to facilitate consumer care through the digital tool.

Community support groups, or another trusted person, were proposed to facilitate mental health care support for consumers when a caregiver was unavailable. This was particularly pertinent for some Aboriginal consumers and those of a lower socioeconomic status, who may lack smartphones or internet access. Conversely, SMS text messaging through basic mobile phones appeared to be more readily available and accepted across the population. By incorporating an SMS text messaging feature into a clinician platform, accessibility may be enhanced for individuals with limited technological literacy or low socioeconomic backgrounds.

Scenario 2

A young adult living and working in town with severe mental health challenges requiring frequent journeys between metro services and rural community support services.

Current channels for patient-mental health service interactions were identified to be initial care visits (GP, hospital admissions, and community mental health groups), crisis services, specialized services, telehealth appointments, potential police contact, and ongoing support groups. Participants valued positive relational interactions in these channels, specifically feeling safe, respected, and avoiding behavioral stereotyping from decision makers such as law enforcement and doctors, as it can lead to inappropriate care.

Consumers and health care professionals reiterated that navigating the health care system with available services was overwhelming and confusing. Moreover, health care professionals acknowledged that patient handovers were inadequate, particularly between metropolitan and rural services. To address this, real-time information sharing, such as discharge summaries between services in the dashboard, was suggested.

Trust and privacy issues were identified as potential concerns for the clinician dashboard and consumer app. Both consumers and health care professionals emphasized the importance of data encryption to protect the transfer of sensitive patient data between services. Including the consumer in the care team ecosystem was suggested to entrust the consumer with handling of their own information, giving the ability to share or unshare data with selected providers.

Scenario 3

A single parent with mild mental health challenges living rurally and accessing GP and NGO supports.

Participants reaffirmed that navigating the available services in the Riverland was overwhelming, advocating for a centralized directory of services in the digital tool with customization specific to the user and their mental health requirements. The potential user in this scenario was presumed to have hectic schedules, leading to missed appointments or medications. To combat this, participants proposed the software to embed a medication and appointment management system that consolidates and integrates the user's existing apps to prevent software fatigue.

User Personas

Informed by the insights derived from the 3 scenarios, the following user personas were developed (Figures 3 and 4): "Anna," a health care professional persona, and "Peter," a consumer persona.

Health care professional persona "Anna" was structured around a GP, as this is often the initial point of contact for mental health care in Australia [22]. One of the biggest challenges in mental health care quoted by a clinician (Health care professional 6) was the lack of specialist services in the Riverland. Health care provider bottlenecks and frustrations presented previously are consolidated under pain points: inefficient referral processes, lack of communication between services, underresourcing, and difficulties in maintaining patient well-being. The desired outcome for Anna is to ultimately improve patient care by coordinating with other services effectively. The attribute scale displays Anna as proficient in technology, care coordination as poor, service navigation confidence as moderate, and time management as poor. Technological channels for providing care entail electronic health record systems and general practice software, telehealth services, and patient communication through mobile devices and email.

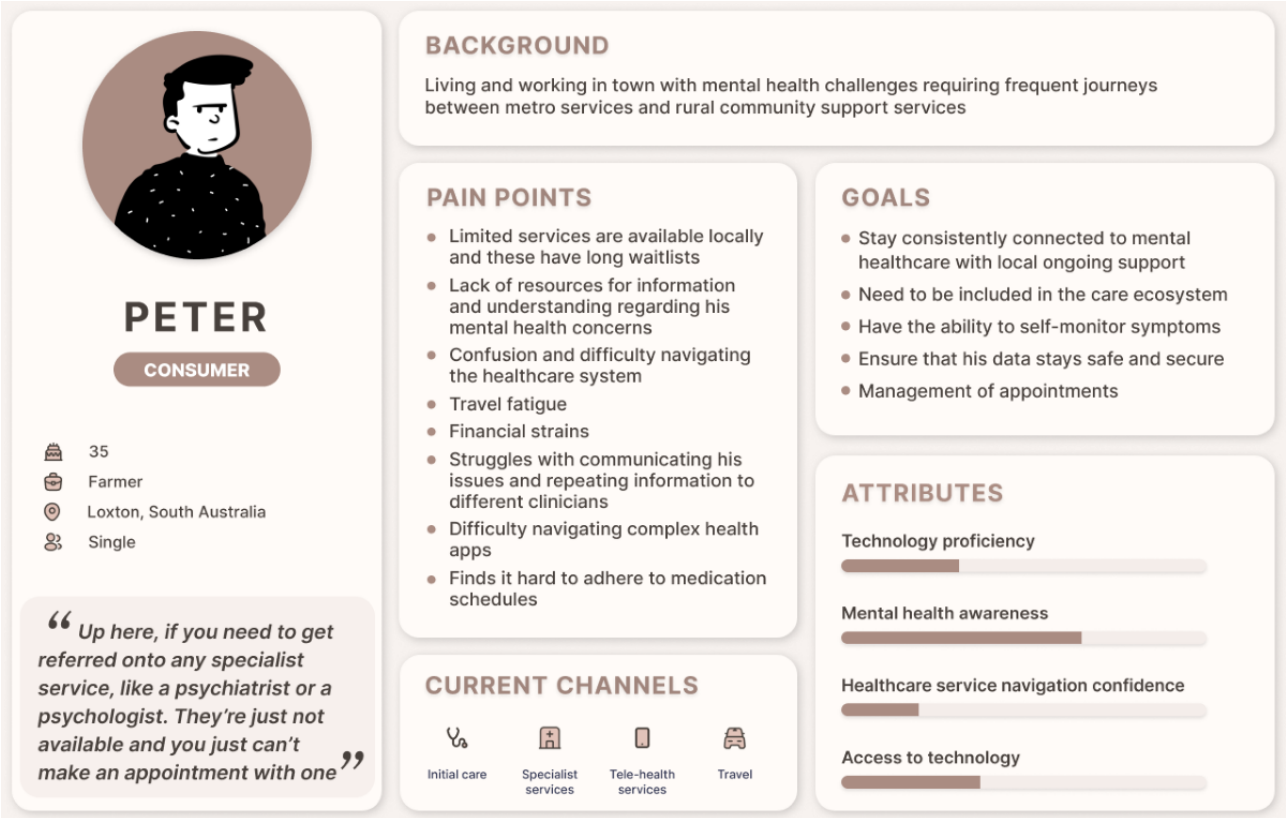
Consumer persona "Peter" was framed around scenario 2. Similar to health care professionals, access and availability of specialist services were quoted to be a pressure point by a consumer (Consumer 1). An amalgamation of consumer issues raised in focus groups and scenarios and pain points underscores rural access issues such as service access, long waitlists, and service navigation. Peter envisions more consumer-driven care and better support throughout his mental health care journey. The attribute scale presents technology proficiency as low, mental health awareness as moderate to high, health service navigation as poor, and access to technology as moderate. Peter's current channels for accessing care include a GP for primary care, traveling for specialist services, and telehealth services for potential ongoing care.

XSL•FO
RenderX

Figure 3. Health care professional persona “Anna”.



Figure 4. Consumer persona “Peter”.



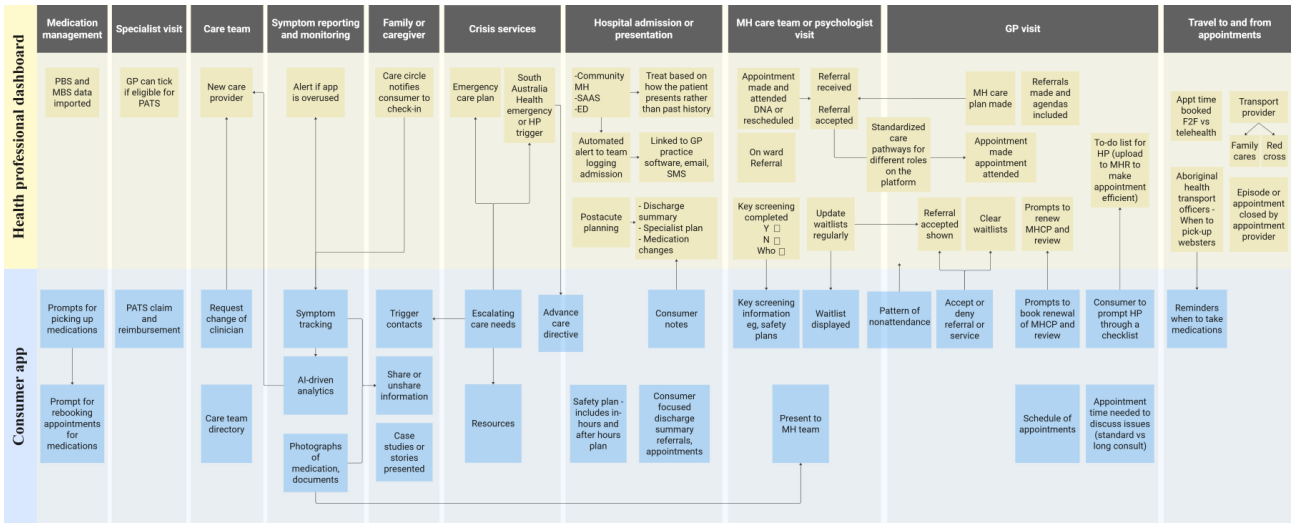
Ideate: Design Workshop 2

The content of the personas was used to guide participant discussions in workshop 2 and ensure that consumer and health care professional goals were kept central, and ideated solutions were addressing respective pain points.

XSL•FO
RenderX

Co-design workshop 2 iteratively revised bottlenecks in prior phases and generated possible solutions through an interaction map between the proposed clinician dashboard and consumer app (Figure 5). An overarching theme of "communication" was identified across the clinician and consumer journey. It was considered essential for health care professionals to be able to easily communicate with each other, and for the consumer's journey to be easily visible, within the constraints of privacy and confidentiality. Health care professionals articulated the desire for consumer-specific information (eg, referrals made to other agencies) as well as generic organization information (eg, requirements for valid referrals), relevant to the consumer they are seeing, to be made clearly accessible. Likewise, communication was a key aspect of consumer agency, both for consumers being aware of their options and being able to communicate required information to their health provider, even if that was prompted via an app and discussed in a session. Participants also provided personal reflections of the study and its aim, ultimately affirming their support for a consumer-facing app and a health care professional–facing dashboard.

Figure 5. Interaction map outlining communication pathways within and between a health professional dashboard and consumer app. AI: artificial intelligence; DNA: did not attend; ED: emergency department; F2F: face to face; GP: general practitioner; HP: health care professional; MBS: Medicare benefits schedule; MH: mental health; MHCP: Mental Health Care Plan; MHR: My Heath Record; PATS: patient assistance transport scheme; PBS: pharmaceutical benefits scheme; SAAS: South Australia Ambulance Service.



Prototype and Test

Low-fidelity prototypes were developed by the research team based on ideated solutions and re-evaluated using the personas. These displayed key interactions between the consumer app and the health care professional dashboard. The main features will be applied and tested in the future include but are not limited to the connectivity between platforms used by different providers, self-assessment and patient assessment tools that are required for referral and monitoring processes, medication management, social network building, transport, and telehealth options.

Discussion

Principal Findings

The core purpose of this study was to initiate co-designing a digital platform with consumers and health care professionals, which may enhance referral processes and service navigation in rural mental health care. Three core themes emerged, which include (1) consumers identifying that "living in a rural community" has both positive aspects (strong community bonds) and negative challenges (access to technology and innovation); (2) health care professionals highlighting "service pressures" such as waitlists, consumer well-being, and resourcing; and finally, (3) both consumers and health care professionals identifying "service limitations," such as navigation challenges and issues around an inefficient referral process. The human-centered technological solution that was explored in the workshops highlights the need for "communication" to integrate the consumer, caregiver, and providers, marking a critical step toward optimizing rural mental health care.

In this study, consumers highlighted both challenges and positive aspects of living in a rural community. Struggles included limited access to mental health services, often requiring travel to metropolitan locations for necessary care. Despite these challenges, rural consumers were often mentioned to be supported by dedicated caregivers and benefited from strong community bonds, especially among the Aboriginal population. These findings are consistent with prior studies [7] that identified barriers such as limited resources and distance to services while also highlighting the importance of person-centered and collaborative care. Thus, a key implication for future iterations of the prototype is to integrate community-based support systems to reduce strain caused by remoteness. Participants in this study further mentioned the importance of integrating other health services and community groups into the platform. While this is important for delivering a holistic service, it is outside the scope of this study and requires further investigation of limitations such as data sharing barriers. Future research could explore opportunities to integrate health services, providing additional support to consumers.

Rural health care professionals in this study faced pressures of underresourcing and access challenges. Similar to previous research [23], this resulted in mental health care systems that

XSL•FO
RenderX

were overly complex, impacting health care professionals in coordinating care and consumers in navigating the required care. Our study showed that both providers and consumers deemed the current referral processes to be inefficient due to a lack of care coordination and limited communication between services, exacerbated by existing rural service limitations [7]. Participants in this study supported the view that a dashboard for health care professionals could address these referral inefficiencies by integrating referral tracking, communication tools, and standardizing or displaying referral processes for various services. Participants also agreed that a consumer-facing app complements the dashboard by offering consumers the agency to input information for their care team and displaying real-time service updates. An app could also allow for a consumer-nominated, trusted community member to access particular information, for example, appointment times. Thus, transport to appointments could be facilitated if the consumer is unable to take themselves. Although the digital platform may not directly resolve resource constraints, it is likely to reduce administrative loads for health care professionals, enhance care coordination, and provide a more consumer-centered approach to mental health care by improving the accessibility of services for consumers [24]. This could also afford a flow-on effect, whereby practitioner burnout is decreased and retention is improved. In particular, in rural locations where GPs take on a high proportion of mental health cases, easy, appropriate referral could reduce compassion fatigue and burnout [8].

Prior studies have explored the use of digital platforms to improve service navigation [15] and potentially improve communication among providers [25,26]. However, these platforms did not integrate well with existing infrastructure used by clinicians; thus, they resisted uptake [27]. Furthermore, many of these solutions were designed for urban settings and do not resolve the challenges faced by rural consumers and providers. Rural Australia, compared to metropolitan areas, has fewer public mental health systems [28] but a higher prevalence of private organizations that often operate in isolation. The health care professional dashboard could enable communication with these private organizations and public systems, facilitating care coordination. Thus, this study distinguished itself by its focus on a rural community. This uniquely positioned the study to co-design a platform to improve rural mental health referral pathways. Due to the location, there were fewer services to coordinate than a metropolitan region; yet, the population could benefit substantially from a platform that decreases the sometimes wastefully siloed service provision and enhances communication between service providers, consumers, and their care teams.

Strengths and Limitations

A key strength of this study lies in its participatory approach, engaging both consumers and health care professionals in the design process, ensuring that the digital platforms developed are grounded in real-world experiences [29] and tailored to the unique challenges of rural mental health care. By using design-thinking methodology, the prototyped solutions are human-centric, contextually relevant, and functional [16]. The percentage of Aboriginal participants employed in this study was representative of the community demographics; however, further in-depth Aboriginal input is needed prior to implementing the solution. This study was also limited by its total number of participants. While this is reflective of a pilot study, it may also limit the generalizability of the findings to broader contexts; thus, future studies with a greater number of participants are necessary to further prototype the provider platform and consumer-facing app in continued consultation with end users. This includes testing the feasibility and acceptability and determining which cohorts would benefit most. Furthermore, the solution should be tested in other rural Australian regions to assess scalability.

Conclusions

The proposed human-centered technological solution aims to integrate the consumer, the caregiver, and the provider, marking a critical step toward optimizing rural mental health care. The co-design process empowered consumers to actively engage in their mental health care journey while suggesting options to streamline system inefficiencies for overburdened clinicians. Ultimately, this platform may not only improve patient well-being and satisfaction but also boost the morale and retention of rural mental health professionals—a critical factor in addressing the current workforce shortage. Importantly, designing a solution in collaboration with rural communities may help to ensure a trusted, sustainable, and scalable implementation [29] in the future.

Acknowledgments

The authors acknowledge the contributions of DC, who passed away before the publication of this manuscript. The study was produced with the financial and other support of the Health Translation SA (HTSA) Medical Research Future Fund Catalyst Grant Scheme, led by HTSA, and supported by The Hospital Research Foundation Group.

Authors' Contributions

KB and AEM were involved in the conceptualization, methodology, data collection, data analysis, and writing the manuscript. AMS was involved in data collection, data analysis, creation of figures, and writing the manuscript. PW, NB, SM, SW, MJ, BM, and DC were involved in the conceptualization and methodology. KP was involved in the conceptualization, methodology, and data collection. BP and SL were involved in the conceptualization, methodology, data collection, and data analysis. All authors reviewed and edited the final manuscript.

XSL•FO
RenderX

Conflicts of Interest

NB is a shareholder and director of CareMappr Pty Ltd. NB is also a software developer who has developed a platform for use in mental health care. NB was not involved in recruitment, data collection, or data analysis.

Multimedia Appendix 1
Focus group session guide.
[DOCX File, 50 KB - humanfactors_v12i1e73460_app1.docx]

Checklist 1
COREQ checklist.
[PDF File, 446 KB - humanfactors_v12i1e73460_app2.pdf]

References

1. Arias D, Saxena S, Verguet S. Quantifying the global burden of mental disorders and their economic value. EClinicalMedicine 2022 Dec;54:101675. [doi: 10.1016/j.eclinm.2022.101675] [Medline: 36193171]
2. Calder R, Dunkin R, Rochford C, Nichols T. Australian Health Services: too complex to navigate. A review of the national reviews of Australia's health service arrangements. : Australian Health Policy Collaboration; 2019.
3. Morales DA, Barksdale CL, Beckel-Mitchener AC. A call to action to address rural mental health disparities. J Clin Transl Sci 2020 May 4;4(5):463-467. [doi: 10.1017/cts.2020.42] [Medline: 33244437]
4. Andrilla CHA, Patterson DG, Garberson LA, Coulthard C, Larson EH. Geographic variation in the supply of selected behavioral health providers. Am J Prev Med 2018 Jun;54(6 Suppl 3):S199-S207. [doi: 10.1016/j.amepre.2018.01.004] [Medline: 29779543]
5. Rathod S, Pinninti N, Irfan M, et al. Mental health service provision in low- and middle-income countries. Health Serv Insights 2017;10:1178632917694350. [doi: 10.1177/1178632917694350] [Medline: 28469456]
6. Stephenson E. Mental disorders and access to mental health care. Statistics Canada. 2023. URL: https://www150.statcan.gc.ca/n1/pub/75-006-x/2023001/article/00011-eng.htm [accessed 2025-09-04]
7. Kavanagh BE, Corney KB, Beks H, Williams LJ, Quirk SE, Versace VL. A scoping review of the barriers and facilitators to accessing and utilising mental health services across regional, rural, and remote Australia. BMC Health Serv Res 2023 Oct 4;23(1):1060. [doi: 10.1186/s12913-023-10034-4] [Medline: 37794469]
8. Rattray M, Fuss B, Shelby-James T, et al. Exploring practice staffs' perspectives regarding professional roles, confidence, and motivations in delivering mental health care within rural general practice: a qualitative exploration. BMC Prim Care 2024 Nov 13;25(1):395. [doi: 10.1186/s12875-024-02648-2] [Medline: 39538205]
9. Moroz N, Moroz I, D'Angelo MS. Mental health services in Canada: barriers and cost-effective solutions to increase access. Healthc Manage Forum 2020 Nov;33(6):282-287. [doi: 10.1177/0840470420933911]
10. Farrer LM, Walker J, Harrison C, Banfield M. Primary care access for mental illness in Australia: patterns of access to general practice from 2006 to 2016. PLoS One 2018;13(6):e0198400. [doi: 10.1371/journal.pone.0198400] [Medline: 29856836]
11. van Spijker BA, Salinas-Perez JA, Mendoza J, et al. Service availability and capacity in rural mental health in Australia: analysing gaps using an Integrated Mental Health Atlas. Aust N Z J Psychiatry 2019 Oct;53(10):1000-1012. [doi: 10.1177/0004867419857809] [Medline: 31250654]
12. Rosenberg S, Iorfino F, Hickie I. Redesigning primary mental health care for enhanced access, equity and quality. : The University of Sydney; 2023 URL: https://www.croakey.org/wp-content/uploads/2023/08/redesigning-primary-mental-health-care.pdf [accessed 2025-09-04]
13. Bellairs-Walsh I, Byrne SJ, Bendall S, et al. Working with young people at risk of suicidal behaviour and self-harm: a qualitative study of Australian general practitioners' perspectives. Int J Environ Res Public Health 2021 Dec 8;18(24):12926. [doi: 10.3390/ijerph182412926] [Medline: 34948536]
14. Santa K, Dixon C, Ganga RN, et al. Facilitating access to mental health services: a stakeholder-driven improvement of the children and young people (CYP) as one referral platform. Int J Environ Res Public Health 2024 Jun 16;21(6):784. [doi: 10.3390/ijerph21060784] [Medline: 38929030]
15. Kevat A, Manohar J, Bate N, Harris K. Online referral and immediate appointment selection system empowers families and improves access to public community paediatric clinics. J Paediatr Child Health 2019 Apr;55(4):454-458. [doi: 10.1111/jpc.14228] [Medline: 30238684]
16. Doorley S, Holcomb S, Klebahn P, Segovia K, Utley J. Design Thinking Bootleg. Stanford University. 2018. URL: https://dschool.stanford.edu/tools/design-thinking-bootleg [accessed 2025-09-04]
17. Ezzy D. Qualitative Analysis: Routledge; 2002.
18. Tong A, Sainsbury P, Craig J. Consolidated Criteria for Reporting Qualitative Research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]

XSL•FO
RenderX

19. Miaskiewicz T, Kozar KA. Personas and user-centered design: how can personas benefit product design processes? Design Studies 2011 Sep;32(5):417-430. [doi: 10.1016/j.destud.2011.03.003]
20. Pruitt J, Adlin T. The Persona Lifecycle: Keeping People in Mind Throughout Product Design: Morgan Kaufmann; 2006. [doi: 10.1145/1167867.1164070]
21. Dismantling systemic racism—perspectives from First Nations peoples. Australian Human Rights Commission. 2024. URL: https://humanrights.gov.au/about/news/media-releases/dismantling-systemic-racism-perspectives-first-nations-peoples [accessed 2025-08-09]
22. General practice: health of the nation 2022. The Royal Australian College of General Practitioners. 2022. URL: https://www.racgp.org.au/getmedia/80c8bdc9-8886-4055-8a8d-ea793b088e5a/Health-of-the-Nation.pdf [accessed 2025-09-04]
23. Evans J, Horn K, Cowan D, Brunero S. Development of a clinical pathway for screening and integrated care of eating disorders in a rural substance use treatment setting. Int J Ment Health Nurs 2020 Oct;29(5):878-887. [doi: 10.1111/inm.12722] [Medline: 32233124]
24. Hews-Girard J, Bright K, Barker M, et al. Mental health provider and youth service users' perspectives regarding implementation of a digital mental health platform for youth: a survey study. Digit Health 2024;10:20552076241289179. [doi: 10.1177/20552076241289179] [Medline: 39421310]
25. Davenport TA, LaMonica HM, Whittle L, et al. Validation of the InnoWell platform: protocol for a clinical trial. JMIR Res Protoc 2019 May 31;8(5):e13955. [doi: 10.2196/13955] [Medline: 31152524]
26. LaMonica HM, Iorfino F, Lee GY, et al. Informing the future of integrated digital and clinical mental health care: synthesis of the outcomes from project synergy. JMIR Ment Health 2022 Mar 9;9(3):e33060. [doi: 10.2196/33060] [Medline: 34974414]
27. O'Shea P, Bates S, Katz I, Gendera S. Project synergy (phase 2) implementation evaluation: final report part A (summary). : UNSW Social Policy Research Centre; 2023 URL: https://ds8526jcpbygs.cloudfront.net/wp-content/uploads/2023/06/InnoWellPartA_Summary55.pdf [accessed 2025-09-04]
28. Rural and remote health. Australian Institute of Health and Welfare. 2024. URL: https://www.aihw.gov.au/reports/rural-remote-australians/rural-and-remote-health [accessed 2025-09-04]
29. Russell K, Rosenbaum S, Varela S, Stanton R, Barnett F. Fostering community engagement, participation and empowerment for mental health of adults living in rural communities: a systematic review. Rural Remote Health 2023 Mar;23(1):7438. [doi: 10.22605/RRH7438] [Medline: 36966523]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
GP: general practitioner
IAR-DST: Initial Assessment and Referral Decision Support Tool
NGO: nongovernment organization

Edited by A Kushniruk; submitted 05.03.25; peer-reviewed by M Drummond, P McEnhill; revised version received 04.06.25; accepted 09.06.25; published 18.09.25.

Please cite as:
Bartel K, Siriwardene AM, Worley P, Pearce K, Bidargaddi N, Patrickson B, Moody S, Wingard S, Jones M, McKenny B, Cameron D, Lawn S, Mendham AE
Improving Mental Health Referral Systems in Rural Australia: Co-Design Study With Health Professionals and Consumers
JMIR Hum Factors 2025;12:e73460
URL: https://humanfactors.jmir.org/2025/1/e73460
doi:10.2196/73460

© Kate Bartel, Asini Malinka Siriwardene, Paul Worley, Kim Pearce, Niranjan Bidargaddi, Bronwin Patrickson, Simon Moody, Sharon Wingard, Martin Jones, Brian McKenny, Darryl Cameron, Sharon Lawn, Amy E Mendham. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 18.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploratory Research With a Health Consumer Group on Social Robot Use Among Older Adults: Qualitative Study

James Sadler, MRes; Aila Khan, PhD; Omar Mubin, PhD; Michael Lwin, PhD

School of Business, Western Sydney University, 169 Macquarie Street, Parramatta, Australia

Corresponding Author:
James Sadler, MRes
School of Business, Western Sydney University, 169 Macquarie Street, Parramatta, Australia

Abstract

Background: There is an increased focus on involving members of the public in health research. These types of groups, such as "health consumer groups," bring different expertise to inform the design of a research study. There is a growing general concern about older adults' acceptance and use of technologies. This becomes critical when it involves health care services.

Objective: To understand the use of social robots among older adults, it is prudent to gauge stakeholders' perspectives on optimal research design. In line with the philosophy of the "triple helix model," researchers sought the expertise and guidance of a health consumer group.

Methods: Researchers recruited an expert health consumer group for this study. This included 5 participants from an 8-member panel. Semistructured interviews were conducted. Each interviewee was introduced to visual stimuli of assistive technologies, older adults, and social robots. Subsequently, they were asked for their perspectives on what they viewed and to provide guidance on how to best design upcoming research on these phenomena.

Results: Key themes were derived from the interview transcripts with the health consumer group members. Findings include panel members' advice and guidance on explaining the research aims to technology-averse older adults, approaching data collection from this demographic, and, finally, their perceptions of the appearance of social robots.

Conclusions: The advice and guidance of this expert health consumer, in tandem with researchers and industry partners, substantially aid in advancing research efforts toward social robot use among technology-averse older adults in Australia. This research provides vital information, including how best to approach data collection about social robots from this demographic.

(JMIR Hum Factors 2025;12:e70462) doi:10.2196/70462

KEYWORDS

older adults; assistive technology; social robots; health care stakeholders; health consumer group; triple helix model of innovation; quality of life

Introduction

Health Consumer Groups

A health consumer group, also called the consumer council, is a platform that brings together community members to provide health care–related consumer views, opinions, and experiences to help guide research projects and to disseminate research findings among the community [1]. A "consumer" is a receiver or a potential receiver of health care services. However, other terms are also used to describe health care recipients, laypersons, users, service users, patients, and clients [2].

There is an increased focus on involving differing perspectives from the public domain in developing health-specific interventions [3]. This involvement of public representatives enables a broad range of differing ideas, perspectives, and resolutions to be brought forward [1]. Such public involvement is advocated for governing boards [4], social care [5], and consumer research [6]. It is recognized that the expertise of members of the public is unique and is different from possessing professional skills [7]. Consequently, this enables various stakeholders in the health care sphere to have their perspectives incorporated, enabling health care professionals to design and implement effective strategies [1]. While the use of "expert opinion" in health research is common [4], there is now a more significant push to increase the involvement of patients and the public in health research [8]. Public participation in health research refers to research conducted with or by the public rather than to, about, or for them [8]. Consumer and community participation in health care is well-recognized [9]. However, the case for public involvement in health research has, over the years, become more widespread [9].

Health consumer groups in the Australian health care networks are commonplace. Ordinarily, they comprise ex-health care professionals, retired professionals who worked adjacent to the health care sector, and stakeholders with a wealth of experience with the health care sector throughout their lives [1].

XSL•FO
RenderX

Subsequently, their perspectives are vital, as they offer an outside view of issues occurring within the health care sector, which may need to be seen or recognized by professionals [1].

Rationale for Using Health Consumer Groups in Research

Social robot research among older adults in an Australian context is still emerging within the literature. Australia's older adults are particularly hesitant and resistant to newer forms of technology, which pose expected barriers to the acceptance and use of social robots by Australia's older adults in a clinical setting [10,11]. Enabling outside perspectives, such as those of members of a health consumer group, allows researchers to access valuable information regarding potential causes of these barriers. Theoretical, political, and ethical reasons are given in favor of public involvement with research. Theoretically, according to World Health Organization guidelines, there is a difference between disease and illness. A disease is a "physiological and clinical abnormality" [12]. On the other hand, an illness is a persistent sickness and is a "more subjective concept related to personal experience of a disease" [12]. Research studies require the expertise of professionals and input from "experts" who may have witnessed and experienced living through a medical condition.

The political argument for involving the public members in health research is based on the premise that, as citizens and taxpayers, individuals who use health care services have the right to influence research funded through public monies [13]. Finally, the traditional model of a passive patient and an active doctor is no longer valid. Similar to the ethical obligation of health to involve patients in service delivery [2], there is a growing call for community input in research conducted within those communities.

Roles Played by Members of Health Consumer Groups

Members of health consumer groups play a critical role in ensuring clinically relevant research is being undertaken. Health group members are essential in identifying priority areas for research. They can identify research gaps from a layperson's perspective [13]. First, consumer participation helps determine the type of information communities need [13]. Second, they can help with the design and development of innovations. Next, they can help recruit relevant participants for research studies [8]. Finally, members of a health consumer group can ensure that the information produced is in a format and language that is easily comprehensible [13].

Limitations of Health Consumer Groups

There are challenges in using health consumer groups for exploratory research. First, as a preselected group, the perspectives of the consumer group members may not be representative of opinions held by the entire demographic or community [14]. Second, the size of the consumer group plays a crucial role in the depth of the discussion conducted in qualitative research [14]. If a consumer group is too large and the resources of researchers are spread too thin, the engagement can become disorganized, difficult to manage, and lead to inadequate data collection [15]. If the consumer group is too small, the scope of information could be further limited [15]. Finally, there is a chance of misunderstanding some of the technical information and jargon [15]. Not all health consumer group members are health care professionals. This may create knowledge deficits or misunderstandings and can lead to misconceptions about the research, interventions, policies, or practices discussed within the interview processes [14].

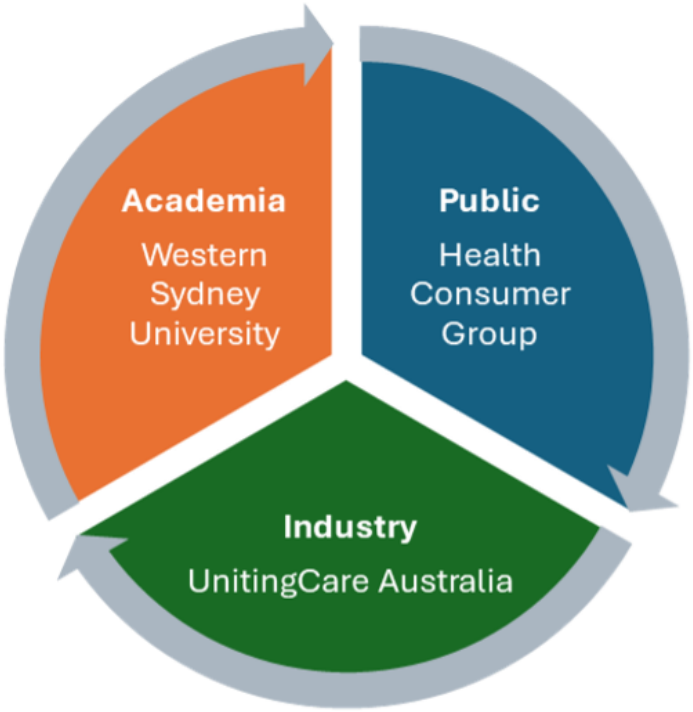
Triple Helix Model

The triple helix model (Figure 1) inspired the use of a health consumer group for this study. First proposed by Etzkowitz and Leydesdorff [16], the "triple helix model" (Figure 1) initially referred to the interactions among higher education, government, and industry to encourage economic, social, and cultural development in a region. Each component of the helix model plays a commensurate role in fostering economic, social, and cultural development [17]. For example, higher education focuses on research and the dissemination of innovation, industry on the production of goods and services, and the government or the public on serving the interests of people, including legislative and regulatory mechanisms that encourage and foster the diffusion of innovation into a population [17].

The triple helix model has mainly been used to improve the innovation process [16]. However, this study uses an "end-user enriched triple helix" approach [18], which considers the end-users' perspective. Innovations do not occur spontaneously; in fact, understanding the user's perceptions and interacting with them is critical to the success of a technological innovation [19]. The lack of input from users may lead to missed opportunities. The use of the triple helix model is based on the premise that involving the end users and adopting a participatory approach would lead to a more effective innovation development process. As older adults are the logical end users [18] of such technologies as social robots in the aged care sector, the triple helix model provides an essential apparatus to include their perspectives in the research process while incorporating academic and industry interests [17].

XSL•FO
RenderX

Figure 1. The triple helix model (as applied in this study).



Social Robots and Older Adults

Assistive technology use among older adults is commonplace. Assistive technologies are ordinarily used to improve the quality of life of older adults by overcoming a physical or cognitive deficit [20]. Newer forms of assistive technology, such as social robots, have been deployed among older adults in a range of contexts with demonstrable effectiveness (Figure 2A). Social robots are machines equipped with a degree of artificial intelligence and human-like mechanics that can mimic or replicate a human interaction with a user [20]. These interactions include companionship, gameplay, touch, eye movement, handshaking, hugging, greetings, verbal interactions, and recognizing cues (Figure 2B). In the aged care sphere, where the number of older adults is increasing, and the number of care staff is declining, social robots can provide valuable supplements to care practices, social interactions, connectedness, and subsequent improvements in the quality of life for older adults (Figure 2B) [21,22].

Primarily, research on social robots among older adults has focused on mimicking social interactions and providing companionship (Figure 2B) [23,24]. As demonstrated by Khosla et al [25], interactions between social robots and older adults in care were associated with significant improvements in observed states of well-being among residents. This is reinforced by Ge and Schleimer [26], whose analysis of older adults' use and acceptance of social robots demonstrated a marked improvement in general well-being. The benefits of social robots to older adults in care transcend social interactions and companionship. Consistent with Fasola and Mataric [27] and Kyong et al [28], deploying a socially assistive robot to assist in exercise routines has been shown to provide positive motivation and outcomes for older adults in care settings. Socially assistive robots can also motivate older adults to engage in physical activities (Figure 2A). According to Macis et al [29], the deployment of a socially assistive robot into a residential aged care facility motivated residents to try and maintain a daily stretching and tai chi routine. Pharmacological disbursement is another area where socially assistive robots have been deployed effectively. According to Smarr et al [30], reminders to take medications and management systems for pharmacology by socially assistive robots in care facilities have produced positive outcomes and a greater level of compliance.

Studies show that the use of social robots can improve the quality of life; however, acceptance of the technology is a key barrier in Australia [31]. To better explore and understand this phenomenon, researchers contend that it is important to seek broader contextual information from differing players in the health care sphere, such as assessing the acceptance of the technology by a health consumer group. This would aid in overcoming these barriers in Australia. In addition, Australia has an aging population, and it is expected that the proportion of Australians aged 65 and older will represent 21% - 23% of the population in 2066. As mentioned above, with a major shortage of health care workers in Australia, the social robots in aged care present an opportunity to fulfill this gap. Thus, understanding the adoption of social robots in aged care in Australia has become a national priority.

XSL•FO
RenderX

Figure 2. (A) Pepper leading group physio. (B) Mario along with a patient with Alzheimer disease (credit: Getty Images).



Research Objectives

Social robot research among older adults in an Australian context is still in its infancy. Research using a health consumer group in designing an academic study with older adults and social robots is limited. Before testing social robots as an intervention among older adults in a clinical aged care setting, it is prudent to gauge the perceptions of an expert health consumer group. Subsequently, the objective of this research is to apply the triple helix model of innovation to seek guidance from a health consumer group to aid research efforts in the adoption of social robots among older adults.

Methods

Participant Recruitment

UnitingCare Australia recommended the research team contact the Sydney Partnership for Health, Education, Research and Enterprise for advice on this study. An element of the Sydney Partnership for Health, Education, Research and Enterprise is the Age and Aging Clinical Academic Group (AAA CAG), which acts as an expert health consumer panel related to aged care [1]. The work of the AAA CAG is to guide and inform research, policy, practice, and support assistive technology use to further healthy aging and independent living for Australia's aged population [1]. The AAA CAG consists of 8 members, many of whom have worked within, adjacent to, or have extensive experience in the Australian health care sector. People who worked in AAA and CAG formed the population of the study. To reduce bias, each person was invited to participate in the study. Therefore, the participants were selected randomly from the population.

Participant Information

Overall, 5 members of AAA CAG were interviewed, of whom 3 were deselected due to lack of interest, age, and time constraints. Of the 5 interviewed members, 3 were males and 2 were females between the ages of 66 and 73 years. Overall, 2 were from non–English-speaking backgrounds (non-ESB) and 3 were from English-speaking backgrounds (ESB; Table 1.

Table . Participant demographic information.

Participant number	Age (years)	Sex	Language background
P1	73	Male	Non–English speaking
P2	66	Male	English speaking
P3	73	Female	English speaking
P4	72	Female	Non–English speaking
P5	66	Male	English speaking

Participant Interviews

Semistructured interviews were conducted with each member via Zoom (Zoom Communications, Inc.). Participants were asked background questions from a preselected pool about their perceptions and experiences with assistive technologies. Visual stimuli, including images and short video clips about social robots and older adults, were used to give the participants an insight into this novel technology. Examples included images of a social robot leading a seated physiotherapy routine to older residents in a care facility. Similarly, a BBC YouTube clip of an independently living older man interacting with PEPPER, a social robot, was also used in the interview process. This was designed to mirror past studies that have also used videos and visual stimuli to demonstrate the capability of robots [32,33]. Finally, researchers asked participants for their opinions, feelings, and perceptions regarding the use of social robots in a research project with older adults. Each interview lasted approximately 45 minutes, and audio recordings were taken using Zoom. Professional transcriptions of each recording were

XSL•FO
RenderX

generated using Notta AI software (Notta Inc.), and each participant was deidentified and provided with a participant number (P1-P5). Manual content analysis was conducted on each transcription. Transcripts were read and reread manually to ensure maximum retention.

Data Analysis

The initial content analysis was inductive, following the guidelines of Braun and Clarke [34,35]. This involved familiarization with the data by reading and rereading the interview transcripts. During this process, the researchers identified and reviewed initial codes, which were later grouped into themes. The themes were subsequently confirmed through meetings with the research team and deductively grouped using the protocols developed by the Australian Clinical Trials Alliance. Three key themes emerged across the 5 interviews with members of the health consumer panel.

Ethical Considerations

Ethics approval was sought from the Western Sydney Human Research Ethics Committee, which approved application H15666 to seek information from the AAA CAG. As the study involved an expert consumer panel, who is involved with aged care in an ancillary capacity, its ethics classification was a low-negligible risk. The AAA CAG meets monthly to discuss various research projects, technologies, interventions, policies, and practices currently within the health care sector. As part of the recruitment process, the authors of this paper were asked to present their research study via a Zoom meeting to all AAA CAG panel members. The chair of AAA CAG permitted the research team to contact each interested member for interviews by providing each member's email address. Interviews, rather than a focus group discussion, were preferred. Interviews ensured that all members were free to voice their opinions without hesitation. Each participant met the inclusion criteria: a member of the AAA CAG, above 65 years of age, willingness and availability to participate in the study, and some experience with aged care consumer products. After the preselection phase, researchers distributed participant information sheets, consent forms, and other informational materials to members who indicated interest in participating. Informed consent was obtained prior to conducting interviews. All participants were deidentified and any identifiable information was obscured within this research and held by researchers only. Gift cards in the amount of $50 were distributed to each participant for their participation in this project.

Results

Overview

This section outlines the main themes that emerged as the researchers analyzed the transcripts. In line with previous research on engaging consumers and communities [36], we focused on identifying topics that provided guidelines for researchers on how to use social robots with older adults. Participants' quotes are presented with the following information: participant identification number, gender, age, and language background (ESB or non-ESB).

Communicating the Research Aim to Tech-Averse Participants

Researchers may find communicating the interview's research aims to the participants easy and straightforward. However, it can be a challenge with some population segments. Older adults could be skeptical about new technologies, including social robots. Not surprisingly, some may hold opposing views even about standard assistive technologies.

> … *my mother-in-law - an elderly lady-had to be convinced to get a walking stick and get it fitted properly as well. She said she didn't need it, but she does.* [P3, Female, 73 years, ESB]

Hesitation to use tools or equipment by older adults appears to be shared across different backgrounds and demographics:

> … *it's nothing to do with education. It's to do with a mindset [of this age group]*. [P4, Female, 72 years, non-ESB]

One research participant was particularly vocal about the need to explain the research aims to older adults:

> … *the biggest thing that I can see is to be able to communicate, talk to them [participants]*. [P1, Male, 73 years, non-ESB]

The same participant offered advice, stating:

> *Initially, they might [resist], but when they can see the benefits of it, I don't think there will be any resistance [to the research project]*.

Panel members of this consumer health group reiterated that undertaking human-robot interaction research is meaningful if older adults are convinced about its benefits. As mentioned by a non–English-speaking panel member:

> *Initially, they might [resist], but when they can see the benefits of it, I don't think there will be any resistance...* [P1, Male, 73 years, non-ESB.]

A suggestion on how to communicate the research aims of a social robot project to older adults highlighted the role of family members:

> *If there was a family member or carer involved and they also were accepting and recommending [involvement in the research project], I think yes... that would certainly get through the first barrier.* [P5, Male, 66 years, ESB]

Data Collection Methods With Older Adults

The consumer health panel members advised researchers on collecting data from older adults. The participants' overarching opinions were "direct" questions and "face-to-face" interviews. The research team also received advice on involving "health professionals" and "family members" during the data collection phase. One panel member (P1, male, 73 years, ESB) warned the research team against using videos to communicate the idea of a social robot. Despite older adults' generally known aversion to technology, this English-speaking male participant strongly advocated for a direct interaction of the robot with older research participants.

XSL•FO
RenderX

Appearance of the Social Robot

The research team also sought advice on the robot's appearance. Two members of the consumer panel had previously been exposed to humanoids:

> *...In Paris, in 2017, ... this little white machine came up and was talking to us in English.* [P3, Female, 73 years, ESB]

Another participant revealed the introduction of a baby seal, Paro, at his previous workplace. These panel members did not find the social robot "scary." Such participants quickly commented that this may not be the case with the average older adult who would interact with the humanoid for the first time:

> *Look..., it goes back to that idea of dystopian...it's sort of big brotherish.* [P5, Male, 66 years, ESB]

Another panel member (P4, female, 72 years, non-ESB) stated that while the robot's image was "acceptable"… "the eyes could be more friendly." This participant also hesitated about the robot's look:

> *The expression,..., the expression, yeah...* [P4, Female, 72 years, non-ESB]

Interestingly, another female participant (P3, 73 years, ESB) appreciated the "feminine form" of the robot and that it was "not confronting with the height and size."

Discussion

Principal Findings

In this section, we highlight 3 key themes identified from the data analysis. We also highlight the managerial implications of the results for each theme. The primary aim of this paper was to seek advice and guidelines from members of a health consumer group regarding human-robot interaction research with older adults. Such advice becomes even more critical when the consensus is that a participant group is vulnerable [37]. There is a general concern about older adults' vulnerability to technology, especially regarding their security and privacy [38].

Key Theme 1: Communicating the Research Aim to Tech-Averse Participants

The first theme that emerged from our analysis was the importance of explaining the aims of a research project to an anxious participant group. Previous researchers [39] have already acknowledged that conducting research with older adults presents challenges; however, the most identified challenges center around obtaining consent and ensuring privacy during interviews. The discussion with the health consumer group revealed that presenting the research topic to tech-averse participants, such as older adults, could hinder gaining cooperation for interviews and discussions [40]. The results demonstrate that demographic variables have little role to play in older adults' attitudes toward technology. Researchers working on tech-related topics should involve family members, peers, or other familiar clinicians to ensure smooth participation in the research project. Furthermore, identifying an opinion leader within the group and encouraging them to use technology will have a positive impact on the group's adoption of the technology. This snowball effect will be more powerful than trying to convince everyone to adopt the technology. In addition, to encourage tech-averse participants to adopt the technology, consistent messaging about its benefits is required. For example, "the robot can help with my exercise routine" and "the robot exercise routine is entertaining and fun." Highlighting the health and entertainment benefits is more likely to reduce resistance to technology. Furthermore, it is also important to change the mindset; for example, a 30-minute robot tai chi exercise as part of their daily routine could help change their mindset. Robots delivering these exercises daily can help older individuals change their daily behavior and change their mindset toward regular exercise.

Key Theme 2: Data Collection Methods With Older Adults

Another commonly discussed topic with the health consumer group focused on data collection methods. Despite increasing evidence of older adults' comfort in using the internet [41], our panel of experts felt that the best way to collect data was face-to-face, allowing older adults to interact directly with a social robot. While it may not be feasible to "house" the robot for extended periods at a facility [42], the findings suggest short-term interactions in a safe environment would be an ideal way to introduce older adults to humanoids. For example, short 30-minute robot-driven exercises daily would work better than 2 - 3 hours of activity. The findings also indicate that effective communication between the robot and the individual is critical in the adoption of the robot. Older individuals would find it difficult to communicate with the robot if it were at the house for only 1 - 2 days. Thus, having the robot for several weeks or months in the house would help older individuals learn to communicate with the robot more effectively. This would also help change the "mindset" toward being more active and exercising.

Key Theme 3: Appearance of the Social Robot

Finally, the health consumer group provided mixed feedback on the robot's appearance and features. Recent studies [43] have found that older adults label the robot as "friendly." However, it could be that some of these studies used smaller types of robots in their projects. Other researchers [44] have reported older adults' dislike for a machine that tries to look like a human—a fake human. Our results demonstrate the need for further research in this area. It is necessary to consider users' needs by investigating them in greater detail. At this stage, the results show conflicting arguments from the health consumer group. For example, some preferred robots with eyes that were more friendly. However, they did not like the expressions of the robot. Other participants highlighted that the feminine features of the robot made it look less confrontational. While these findings are limited to a few participants, it is clear that the personalization and customization of robot features are necessary to enhance adoption. It was interesting to find that the adoption of the robot depends on the individual's preferences for the robot's features. However, academic research in this area is limited, and future studies should investigate the personalization and the personification of robots.

XSL•FO
RenderX

Limitations and Future Studies

Several essential caveats deserve comment. First, this research study was conducted with an 8-member health consumer group, resulting in a small sample size for interviews. Next, our study considered only one type of social robot—Pepper. It is quite possible that a different robot type may have resulted in different opinions by the participants. Finally, it is important to consider the presence of bias in participant responses due to their values and attitudes toward robots. The industry partner (UnitingCare Research) recommended future research to focus on the health consumer group. This would provide a standard unit of analysis in future studies, and it provides researchers with a standard research process to compare against. One of the key learnings from the study is that when you are recruiting participants for the health consumer group, it is important to vet the participants. Future studies should develop a vetting process for the health consumer group. This would help future studies have a standardized selection process and help researchers compare the results more accurately. Furthermore, the study was limited to the qualitative methodology, and future studies should validate the findings using empirical modeling.

As this research is the foundation for future studies, including social robots and older adults, the results highlight how best to approach studying this phenomenon. Future studies should consider these recommendations to effectively design social robot interventions among older adults in the Australian aged care context. However, the findings are limited to the Australian aged care context, and future researchers should explore different cultural backgrounds. For example, in the Middle East, religion is at the core of their livelihood, and therefore, the way individuals interact with the robot may differ substantially from that of older Australian adults. The Japanese are more likely to adopt the robot; thus, future researchers may compare whether culture impacts the adoption process. Comparisons between these contexts are important to understand the factors driving robot adoption. These findings could be used to advise future studies in aged care.

Conclusions

Health consumer groups are vital health care stakeholders, whose experiences, perspectives, concerns, and recommendations form the basis for designing robust interventions to improve the quality of life for specific demographics through products and services. After being pretested among an expert health care consumer panel by the authors, the findings of this study should provide confidence to another element of the triple helix model: industry. However, as mentioned in the limitations of this study, social robot research among older adults in Australia is still emerging. This study represents nascent exploratory research that requires further testing. The authors, in conjunction with industry partners, recommend field testing of social robots among older adults to assess their acceptance and use, as they are the intended end users of the technology.

Authors' Contributions

Conceptualization: JS and AK
Validation: JS and AK
Writing - original draft preparation: JS
Visualization: AK
Supervision: AK, OM, and ML

Conflicts of Interest

None declared.

References

1. Age & ageing. Sphere. 2024 Jun 6. URL: https://thesphere.com.au/clinical-themes/age-ageing/ [accessed 2025-08-25]
2. Boote J, Telford R, Cooper C. Consumer involvement in health research: a review and research agenda. Health Policy 2002 Aug;61(2):213-236. [doi: 10.1016/s0168-8510(01)00214-7] [Medline: 12088893]
3. Baxter S, Clowes M, Muir D, Baird W, Broadway-Parkinson A, Bennett C. Supporting public involvement in interview and other panels: a systematic review. Health Expect 2017 Oct;20(5):807-817. [doi: 10.1111/hex.12491] [Medline: 27535876]
4. Florin D, Dixon J. Public involvement in health care. BMJ 2004 Jan 17;328(7432):159-161. [doi: 10.1136/bmj.328.7432.159] [Medline: 14726350]
5. INVOLVE strategy 2012-2015. : National Institute for Health Research; 2024 Apr 22 URL: https://core.ac.uk/reader/186751140 [accessed 2025-09-12]
6. Roberts P, Wild K, Washington K, Mountford C, Capewell J, Priest H. Inclusion of lay people in the pre-registration selection process. Nurs Stand 2010;24(48):42-47. [doi: 10.7748/ns2010.08.24.48.42.c7930] [Medline: 20806610]
7. Mosconi P, Satolli R, Colombo C, Villani W. Does a consumer training work? A follow-up survey of the PARTECIPASALUTE training programs. Health Res Policy Syst 2012 Sep 1;10:27. [doi: 10.1186/1478-4505-10-27] [Medline: 22938140]
8. Russell J, Fudge N, Greenhalgh T. The impact of public involvement in health research: what are we measuring? Why are we measuring it? Should we stop measuring it? Res Involv Engagem 2020;6:63. [doi: 10.1186/s40900-020-00239-w] [Medline: 33133636]

XSL•FO
RenderX

9. Liabo K, Boddy K, Bortoli S, et al. Public involvement in health research: what does “good” look like in practice? Res Involv Engagem 2020;6:11. [doi: 10.1186/s40900-020-0183-x] [Medline: 32266085]
10. Lancioni GE, Singh NN. Assistive Technologies for People with Diverse Abilities, 1st edition: Springer; 2014.
11. Wu YH, Wrobel J, Cornuet M, Kerhervé H, Damnée S, Rigaud AS. Acceptance of an assistive robot in older adults: a mixed-method study of human-robot interaction over a 1-month period in the Living Lab setting. Clin Interv Aging 2014;9:801-811. [doi: 10.2147/CIA.S56435] [Medline: 24855349]
12. Australia’s health. Australian Institute of Health and Welfare. 2014 Jun 9. URL: https://www.aihw.gov.au/reports-data/australias-health [accessed 2025-09-12]
13. Shea B, Santesso N, Qualman A, et al. Consumer-driven health care: building partnerships in research. Health Expect 2005 Dec;8(4):352-359. [doi: 10.1111/j.1369-7625.2005.00347.x] [Medline: 16266423]
14. Tausch AP, Menold N. Methodological aspects of focus groups in health research. Glob Qual Nurs Res 2016 Jan 1;3:233339361663046. [doi: 10.1177/2333393616630466]
15. Wiles LK, Kay D, Luker JA, et al. Consumer engagement in health care policy, research and services: a systematic review and meta-analysis of methods and effects. PLOS ONE 2022;17(1):e0261808. [doi: 10.1371/journal.pone.0261808] [Medline: 35085276]
16. Etzkowitz H, Leydesdorff L. The triple helix -- university-industry-government relations: a laboratory for knowledge based economic development. EASST Review 1995;14(1):14-19 [FREE Full text]
17. Cai Y, Lattu A. Triple helix or quadruple helix: which model of innovation to choose for empirical studies? Minerva (Arlingt Va) 2022 Jun;60(2):257-280. [doi: 10.1007/s11024-021-09453-6]
18. Bunders JFG, Broerse JEW, Zweekhorst MBM. The triple helix enriched with the user perspective: a view from Bangladesh. J Technol Transf 1999 Aug;24(2-3):235-246. [doi: 10.1023/A:1007811607384]
19. Buchanan RA. The atmospheric railway of I.K. Brunel. Soc Stud Sci 1992 May;22(2):231-243. [doi: 10.1177/030631292022002003]
20. Shin HR, Um SR, Yoon HJ, et al. Comprehensive senior technology acceptance model of daily living assistive technology for older adults with frailty: cross-sectional study. J Med Internet Res 2023 Apr 10;25:e41935. [doi: 10.2196/41935] [Medline: 37036760]
21. Hegel F, Muhl C, Wrede B, Hielscher-Fastabend M, Sagerer G. Understanding social robots. Presented at: 2009 Second International Conferences on Advances in Computer Human Interactions; Jul 1-7, 2009; Cancun, Mexico. [doi: 10.1109/ACHI.2009.51]
22. Kang HS, Koh IS, Makimoto K, Yamakawa M. Nurses’ perception towards care robots and their work experience with socially assistive technology during COVID-19: a qualitative study. Geriatr Nurs (Lond) 2023 Mar;50:234-239. [doi: 10.1016/j.gerinurse.2023.01.025]
23. Pirhonen J, Tiilikainen E, Pekkarinen S, Lemivaara M, Melkas H. Can robots tackle late-life loneliness? Scanning of future opportunities and challenges in assisted living facilities. Futures 2020 Dec;124:1-12. [doi: 10.1016/j.futures.2020.102640] [Medline: 33041358]
24. Sun R, Liu Y. Mortality of the oldest old in China. J Aging Health 2006 Feb;18(1):37-55. [doi: 10.1177/0898264305281103] [Medline: 16470963]
25. Khosla R, Nguyen K, Chu MT. Human robot engagement and acceptability in residential aged care. International Journal of Human–Computer Interaction 2017 Jun 3;33(6):510-522. [doi: 10.1080/10447318.2016.1275435]
26. Ge GL, Schleimer SC. Robotic technologies and well-being for older adults living at home. JSM 2023 Feb 28;37(3):340-350. [doi: 10.1108/JSM-03-2022-0076]
27. Fasola J, Mataric M. A socially assistive robot exercise coach for the elderly. Journal of Human-Robot Interaction 2013;2(2):2. [doi: 10.5898/JHRI.2.2.Fasola]
28. Freedman S, Mataric MJ, Cunningham MJ, Lopez B. A hands-off physical therapy assistance robot for cardiac patients. Presented at: 9th International Conference on Rehabilitation Robotics, 2005 ICORR 2005; Jun 28 to Jul 1, 2005; Chicago, IL, USA. [doi: 10.1109/ICORR.2005.1501114]
29. Macis D, Perilli S, Gena C. Employing socially assistive robots in elderly care. Presented at: UMAP ’22; Jul 4-7, 2022; Barcelona, Spain. [doi: 10.1145/3511047.3537687]
30. Smarr CA, Prakash A, Beer JM, Mitzner TL, Kemp CC, Rogers WA. Older adults’ preferences for and acceptance of robot assistance for everyday living tasks. Proc Hum Factors Ergon Soc Annu Meet 2012 Sep;56(1):153-157. [doi: 10.1177/1071181312561009] [Medline: 25284971]
31. Huppertz C, Forbrig TA, Lengert-Brzozowski S, Gräske J. Associations between older adults’ loneliness and acceptance of socially assistive robots: a cross-sectional study. J Gerontol Nurs 2023 Apr;49(4):21-26. [doi: 10.3928/00989134-20230309-04] [Medline: 36989477]
32. Fortunati L, Manganelli AM, Höflich J, Ferrin G. How the social robot sophia is mediated by a YouTube video. New Media & Society 2024 Jun;26(6):3390-3409. [doi: 10.1177/14614448221103176]
33. Saplacan D, Schulz T, Torresen J, Pajalic Z. Health professionals’ views on the use of social robots with vulnerable users: a scenario-based qualitative study using story dialogue method. Presented at: 2023 32nd IEEE International Conference

XSL•FO
RenderX

on Robot and Human Interactive Communication (RO-MAN); Aug 28-31, 2023; Busan, Republic of Korea. [doi: 10.1109/RO-MAN57019.2023.10309644]
34. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
35. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qualitative Research in Sport, Exercise and Health 2019 Aug 8;11(4):589-597. [doi: 10.1080/2159676X.2019.1628806]
36. Todd AL, Nutbeam D. Involving consumers in health research: what do consumers say? Public Health Res Pract 2018 Jun 14;28(2):2. [doi: 10.17061/phrp2821813] [Medline: 29925089]
37. Fordham M, Lovekamp WE, Thomas DSK, et al. Understanding social vulnerability. In: Thomas DSK, Philips BD, Lovekamp WE, et al, editors. Social Vulnerability to Disasters, 2nd edition: CRC Press; 2013:1-29 URL: https://www.academia.edu/8600704/Understanding_Social_Vulnerability_in_Social_Vulnerability_to_Disasters_2nd_ed_ [accessed 2025-09-12]
38. Murthy S, Bhat K, Das S, Kumar N. Individually vulnerable, collectively safe: the security and privacy practices of households with older adults. In: Proceedings of the ACM on Human-Computer Interaction: Association for Computing Machinery; 2021, Vol. 5:1-24. [doi: 10.1145/3449212]
39. Hall S, Longhurst S, Higginson IJ. Challenges to conducting research with older people living in nursing homes. BMC Geriatr 2009 Aug 24;9:1-8. [doi: 10.1186/1471-2318-9-38] [Medline: 19703277]
40. Hu HM. A qualitative study on why older adults may be reluctant to participate in learning activities. Journal of Adult and Continuing Education 2024 May;30(1):39-55. [doi: 10.1177/14779714231191354]
41. Brondani MA, MacEntee MI, O'Connor D. Email as a data collection tool when interviewing older adults. Int J Qual Methods 2011 Sep;10(3):221-230. [doi: 10.1177/160940691101000303]
42. Ostrowski AK, Breazeal C, Park HW. Mixed-method long-term robot usage: older adults' lived experience of social robots. Presented at: 2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI); Mar 7-10, 2022; Sapporo, Japan. [doi: 10.1109/HRI53351.2022.9889488]
43. Bayles MA, Lee J, Kadylak T, Rogers WA. Understanding older adults' initial perceptions of robot appearance and function: implications for acceptance. Gerontechnology 2023 Aug 31;22(2):1-7. [doi: 10.4017/gt.2023.22.2.bay.08]
44. Wu YH, Fassert C, Rigaud AS. Designing robots for the elderly: appearance issue and beyond. Arch Gerontol Geriatr 2012;54(1):121-126. [doi: 10.1016/j.archger.2011.02.003] [Medline: 21349593]

Abbreviations

CAG: Age and Aging Clinical Academic Group
ESB: English-speaking background
non-ESB: non–English-speaking background

Edited by A Kushniruk; submitted 22.12.24; peer-reviewed by CK Tan, F Shojaei; revised version received 29.04.25; accepted 03.06.25; published 24.09.25.

Please cite as:
Sadler J, Khan A, Mubin O, Lwin M
Exploratory Research With a Health Consumer Group on Social Robot Use Among Older Adults: Qualitative Study
JMIR Hum Factors 2025;12:e70462
URL: https://humanfactors.jmir.org/2025/1/e70462
doi:10.2196/70462

© James Sadler, Aila Khan, Omar Mubin, Michael Lwin. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 24.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Testing the Acceptability and Feasibility of a Gender-Informed Smoking Cessation mHealth App for Women: Mixed Methods Approach

Osnat C Melamed[1,2,3,4], MSc, MD; Kamna Mehra[3], DNB, MSc, MBBS; Allison Gayapersad[3], PhD; Roshni Panda[3], PhD; Nadia Minian[1,2,3,4,5], PhD; Laurie Zawertailo[2,3,4,5], PhD; Leslie Buckley[6,7], MD, MPH; Marta Maslej[7,8], PhD; Lorraine Greaves[9,10], PhD; Andreea C Brabete[9], PhD; Jonathan Rose[11], PhD; Matt Ratto[12,13], PhD; Peter Selby[1,2,3,7,14], MBBS, MHSc

[1]INTREPID Lab, Nicotine Dependence Service, Centre for Addiction and Mental Health, 1025 Queen Street West, Toronto, ON, Canada
[10]School of Population and Public Health, Faculty of Medicine, University of British Columbia, Vancouver, Canada
[11]Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, Toronto, Canada
[12]Faculty of Information, University of Toronto, Toronto, Canada
[13]Schwartz Reisman Institute for Technology and Society, University of Toronto, Toronto, Canada
[14]Dalla Lana School of Public Health, University of Toronto, Toronto, Canada
[2]Department of Family and Community Medicine, University of Toronto, Toronto, ON, Canada
[3]Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, ON, Canada
[4]Institute of Medical Sciences, University of Toronto, Toronto, ON, Canada
[5]Department of Pharmacology and Toxicology, University of Toronto, Toronto, ON, Canada
[6]Addictions Division, Centre for Addiction and Mental Health, Toronto, ON, Canada
[7]Department of Psychiatry, University of Toronto, Toronto, ON, Canada
[8]Krembil Centre for Neuroinformatics, Centre for Addiction and Mental Health, Toronto, ON, Canada
[9]Centre of Excellence for Women's Health, Vancouver, BC, Canada

Corresponding Author:
Osnat C Melamed, MSc, MD
INTREPID Lab, Nicotine Dependence Service, Centre for Addiction and Mental Health, 1025 Queen Street West, Toronto, ON, Canada

Abstract

Background: Cigarette smoking is a leading cause of preventable morbidity and mortality worldwide. Women who smoke face greater health risks than men, including higher rates of cardiovascular disease and more pronounced declines in lung function. Despite this, women experience lower success rates with conventional smoking cessation treatments, due in part to unique sex- and gender-related factors influencing smoking behavior and barriers to quitting. Digital health tools, such as mobile health apps, offer a promising avenue for delivering accessible, tailored smoking cessation support to women.

Objective: This study evaluated the acceptability and feasibility of the "My Change Plan–Women" (MCP-W) app, a gender-specific smoking cessation mobile health intervention co-designed with women who smoke, clinicians, and researchers, to address women's unique needs in smoking cessation.

Methods: We conducted a single-group, prospective, sequential mixed methods study with 30 women who smoke in Ontario, Canada. Participants used the MCP-W app for 28 days. Acceptability was defined as ≥50% of participants endorsing "agree" or "strongly agree" to the statement "using the app is likely to help me make changes to my smoking habits." Feasibility was defined as ≥50% of participants using the app for 7 or more days during the trial period. Quantitative data on acceptability, smoking behavior, and motivation to quit were collected at baseline and follow-up via REDCap (Research Electronic Data Capture) surveys. App usage metrics were captured through Google Analytics. Semistructured interviews explored participants' experiences using the app and were thematically analyzed using the theoretical framework of acceptability.

Results: At follow-up, 37% (11/30; 95% CI 21%-56%) of participants rated the MCP-W app as acceptable, falling below the predefined threshold (≥50%) and indicating that the intervention "needs further work." Feasibility criteria were met, with 60% (18/30) of participants using the app for 7 or more days. Notably, acceptability was higher among those who used the app for more than 14 days (7/11, 64%) compared with those with lower usage (4/19, 21%). Average daily cigarette consumption decreased from 16.4 to 14.6 cigarettes, and the number of participants reporting at least 1 smoke-free day in the previous week increased from 7% (2/27) to 22% (6/27). Qualitative findings revealed that women with higher motivation to quit found the app more

XSL•FO
RenderX

helpful, particularly its behavior change tools (eg, cigarette tracking and identifying triggers) and gender-specific content. However, women facing stress, mental health challenges, or low readiness found it harder to engage. Participants suggested enhancements including customizable reminders, more interactive content, and live or artificial intelligence–based emotional support.

Conclusions: The MCP-W app is a feasible intervention for delivering gender-specific smoking cessation support. However, its acceptability was limited to a third of users with high levels of motivation. Improvements to interactivity and support features may enhance its relevance and uptake among women with complex barriers to quitting.

International Registered Report Identifier (IRRID): RR2-10.2196/60677

(JMIR Hum Factors 2025;12:e71683) doi:10.2196/71683

KEYWORDS

women's health; gender; smoking cessation; mHealth; digital health; co-development; feasibility; mobile phone; mobile health

Introduction

Cigarette smoking is one of the leading risk factors of morbidity and mortality globally contributing to the development of cancers, lung diseases, and heart diseases [1]. While the prevalence of smoking among women in some Western countries may be reducing overall, the rate of decline is slowing, and the contribution of smoking to deaths among women is increasing [2,3]. Smoking causes more harm in women compared with men. Relative to men, women who smoke have higher rates of cardiovascular events and experience greater decline in lung function as a result of smoking [4]. When trying to quit smoking to protect their health, women experience less success than men when offered evidence-based smoking cessation treatment. The disparity can be partly attributed to sex- and gender-related factors affecting motives for smoking and barriers to quitting, which are rarely addressed in treatment. This highlights the need for smoking cessation programs to take into account the unique needs of women and incorporate them into treatment.

Various sex- and gender-related factors affect smoking and quitting behaviors among women. Cosgrove et al [5] found differences between the effect of nicotine on the dopaminergic pathway in the brains of men and women, where men showed a greater dopamine response to smoking than women. This suggests that men smoke for the pleasurable effects of nicotine, whereas smoking has additional drivers in women. Others theorized that stressful life experiences in women become a conditioned cue for smoking (ie, sociopharmacology) [6]. This theory is backed by evidence of high smoking rates among women facing material hardship, domestic violence, mental illness, and lone parenting. Biological factors, as well as diverse social and psychological factors (ie, sex- and gender-related factors), contribute to a greater difficulty in smoking cessation among women relative to men [4,7-11]. Previous work suggests a number of unique challenges experienced by women in quitting, including using smoking to cope with stress or manage emotions and control weight by suppressing their appetite [12]. Behavioral interventions that focus explicitly on the unique experiences of women trying to quit could improve smoking cessation success [11].

Mobile health (mHealth) technology, such as smartphone apps, can deliver smoking cessation treatment to wider populations in a more cost-effective manner [12-16]. For greater effectiveness, mobile app content can be tailored to address the unique smoking cessation needs of women. Indeed, 4 apps dedicated to helping women reduce or quit smoking have been identified in a systematic review of the literature [13]; 2 of them focus on women who are pregnant [17,18], 1 on modifying multiple health behaviors among women [19], and 1 for Chilean women using which sends motivational messages and strategies to women to reduce or quit smoking [20]. This trend is encouraging, but there is room for further work focused on a wider, more general population of women who smoke.

The "My Change Plan–Women" (MCP-W) app is one such additional mHealth intervention. The original version of the app, My Change Plan (MCP), was developed to support the smoking cessation efforts of patients enrolled in the Smoking Treatment for Ontario Patients (STOP) program and was a digital evolution of a paper-based booklet, also called "My Change Plan" [21,22]. The MCP app provides evidence-based behavioral change techniques to support tobacco reduction and cessation that can be helpful for both women and men in developing their quit plan. The quit plan includes goal setting, identifying smoking triggers, and developing coping strategies which are grounded in theory [23-25]. It allows users to track their cigarettes and cravings and evaluate financial savings from reducing or quitting smoking. It also provides evidence-based information in the app's library about smoking cessation medications, local supports for smoking cessation (eg, smoker's helpline), and the impact of physical activity, nutrition, stress, and sleep on smoking and smoking cessation. The MCP app has been available for iPhone (Apple Inc) and Android (Google) smartphones, in their respective app stores, for over 5 years. The MCP-W app retains MCP gender-neutral content and complements it with women-specific content that focuses on women's unique motives for smoking and the barriers to quitting they experience. The women-specific content is delivered through 4 animation videos embedded in the app, and testimonials from women who smoked are included in the app's library. The animated videos focus on women-specific motives for smoking, challenges faced by women when trying to quit, and strategies that have helped women overcome them. The testimonials include quotes from women on topics such as "desire to quit," "quit strategies," "quit struggles," and "encouragement."

The women-specific version of this app, the MCP-W app, was co-designed in collaboration with women who smoke (ie, patient partners) and a multidisciplinary team of researchers and

XSL•FO
RenderX

clinicians with expertise in tobacco addiction, implementation science, sex and gender science, women's health, public health, and app development [26]. The MCP-W app was based on motivational interviewing. This is a counseling method aimed to promote behavioral change in a compassionate and nonjudgmental manner [27]. It respects women's autonomy to make changes to their smoking at their own pace, when they feel ready, and does not pressure them into quitting by using scare or shame tactics.

The MCP-W app was developed between July 2022 and August 2023, and a testable version was finalized in January 2024. The aim of this study is to test whether women who smoke find the MCP-W app to be an acceptable tool for the delivery of gender-informed smoking cessation support, and whether the MCP-W app can feasibly deliver this content.

Methods

Study Design and Setting

We conducted a single-group, prospective, sequential mixed methods study to test the acceptability and feasibility of the MCP-W app at the Center for Addiction and Mental Health (CAMH) INTREPID Lab in Toronto, Ontario, Canada. A comprehensive protocol of this study was published previously [26].

Participant Recruitment and Study Procedures

Purposive sampling was conducted between February and May 2024 to recruit 30 women participants who smoked cigarettes to test the MCP-W app. The study recruitment flyer was shared through monthly presentations to STOP clinicians, who provide smoking cessation treatment in primary care clinics across Ontario, Canada. In addition, recruitment flyers were provided to health care practitioners at the Nicotine Dependence Clinic at CAMH and posted on the CAMH research boards across campus. Furthermore, past treatment-seeking individuals (STOP program participants), who agreed to participate in future research studies, were contacted about the study.

Women who expressed interest in participation were screened for eligibility over the telephone. We included women between 25 and 65 years of age, who smoked at least 5 tobacco cigarettes per day, owned a smartphone, had continuous internet access, and had experience using smartphone apps. Those who did not feel comfortable using smartphone apps, or were pregnant, breastfeeding, or trying to become pregnant in the next 3 months were ineligible to participate. Individuals with severe general or mental health conditions that could interfere with study participation were also excluded. This criterion was assessed during the eligibility screening and applied to those unable to complete study procedures due to a physical or mental illness (eg, current hospitalization). Eligible participants provided informed consent over WebEx (Cisco Systems), a secure online platform to conduct videoconference meetings, after a discussion of the study procedures, risks, benefits, and rights as a study participant. During this WebEx meeting, participants also completed the baseline survey (Multimedia Appendices 1 and 2) through a unique REDCap (Research Electronic Data Capture) survey link sent by email, where they provided information about their sociodemographic characteristics, acceptability of the app, smoking behaviors, smoke-free days in the past week, and their motivation to quit smoking. They downloaded the MCP-W app (either an Android or an iPhone version) and were provided a brief demonstration of the various app features.

Participants were granted access to the app for a period of 28 days. After the 28-day trial completion, participants were sent an automatic email through REDCap with a follow-up survey. The follow-up survey asked participants about the acceptability of the app, smoking behaviors, smoke-free days in the past week, motivation to quit smoking, the usability of the app using the System Usability Scale (SUS) [28], and how likely they were to recommend the app to others (Multimedia Appendices 2 and 3). In addition, participants were invited to schedule the follow-up semistructured WebEx interview about their experiences using the app. Interviews were conducted by a research staff member who did not have any previous relationship with the participants. The interview guide consisted of questions about the app based on the 7 constructs of the theoretical framework of acceptability (TFA; ie, affective attitude, burden, ethicality, intervention coherence, opportunity costs, perceived effectiveness, and self-efficacy) [29], as well as additional questions about barriers to quitting smoking, the app's helpfulness in quitting smoking, and strategies to make the app helpful for women from diverse backgrounds (Multimedia Appendix 4). The interview lasted for an average of 34 (SD 11, range 18 - 62) minutes. The interviews were automatically transcribed by the WebEx Assistant Software. Research staff members reviewed, cleaned, and verified the transcripts for accuracy.

Data Analysis

Primary Outcomes

The primary outcomes for the study included acceptability and feasibility of the app. Acceptability of the app was measured by asking the question "using the app is likely to help me make changes to my smoking habits" at baseline and follow-up surveys, which could be answered on a 5-point Likert scale from "strongly agree" to "strongly disagree." The proportion of participants endorsing "agree" or "strongly agree" was calculated. The app would be considered "acceptable" if the majority (≥50%) of participants rate it as such. Similarly, if 26% to 50% reported the app as acceptable, then it would be deemed "needs further work"; otherwise, it would be deemed "unacceptable" if the proportion was 25% or lower. Nonresponse values were imputed to "do not agree" since low study participation most likely indicated low acceptability of the app.

Feasibility of the app was assessed using Google Analytics data. Participants were assigned unique log-in IDs and the number of times they accessed the app and clicked on the videos was collected using Google Analytics. The app would be considered "feasible" if the proportion of participants who accessed the app for 7 days or more out of 28 days was 50% or more, "needs further work" if 26% to 49% , and "not feasible" if 25% or less. Since the Google Analytics data was available for all participants, there was no missing data on feasibility measures. The binomial 95% CIs were calculated to assess precision of

XSL•FO
RenderX

estimates, but the decision about acceptability and feasibility rested on point estimates.

Secondary and Tertiary Outcomes

The secondary outcome measures consisted of participants' motivation to quit smoking. Changes in mean scores of motivation from baseline to follow-up were calculated for each of the 3 validated subscales: "importance," "commitment," and "ability" [30]. Participants rated the following statements: (1) it is important for me to change my smoking habits, (2) I am trying to change my smoking habits, and (3) I could change my smoking habits, using a Likert scale where 1=definitely not, 2=probably not, 3=maybe, 4=probably, and 5=definitely.

The tertiary outcomes consisted of smoking behavior metrics including change in the number of cigarettes smoked daily and the number of smoke-free days in the last 7 days. Quantitative data was analyzed using Stata v.16 (StataCorp). Based on the observed distribution, we used McNemar tests for binary outcomes, exact Wilcoxon signed-rank tests for paired ordinal data, and negative binomial regression for cigarettes per day. Demographic data were analyzed descriptively.

Qualitative Data Analysis

A deductive analysis approach was used to analyze the qualitative data [31]. Using Microsoft Excel, the interview data elicited from the 26 participants were organized into categories based on the 7 TFA constructs and 2 additional questions asked in the interview guide. KM read and reread all transcripts to become familiar with and code the data. AG and OM reviewed the transcripts and the extracted and coded data to ensure accuracy and consistency. KM, OM, and AG met weekly to discuss and refine the codes. Coded data were structured into themes and relevant quotes were selected to support themes and subthemes.

Ethical Considerations

The study was approved by the CAMH Research Ethics Board (#122/2023). Participants provided informed consent via WebEx (Cisco Systems). Participants were provided an honorarium of CAD $30 (a currency exchange rate of US $1=CAD $1.33796 was applicable at the time of data collection) for completing the baseline survey and an additional CAD $30 for completing the posttrial interview. No identifying information about the participants has been included in this manuscript.

Results

Characteristics of Study Participants

A total of 383 women were contacted about the study (369 from the STOP program and 16 from other sources), out of which 55 women expressed interest and were screened for eligibility. Of the total, 49 women were found to be eligible, out of which 32 women provided informed consent and completed the baseline survey. Furthermore, 30 women were able to download the app, and 2 women could not do so due to technical issues, 27 completed the follow-up survey, and 26 completed the interview (3 participants were lost to follow-up). The baseline demographic data of the 30 participants are described in Table 1.

XSL•FO
RenderX

Table . Characteristics of 30 study participants.

Demographic characteristics	Value
Age (y), mean (SD)	49.5 (9.4)
Gender, n (%)	
Woman	30 (100)
Marital status[a], n (%)	
Never legally married	10 (33)
Legally married (and not separated)	10 (33)
Separated or divorced	4 (13)
Living with a partner but not married (common law)	5 (17)
Highest level of education completed, n (%)	
Some high school	1 (3)
High-school diploma	5 (17)
Some college	5 (17)
College diploma	13 (43)
University degree	6 (20)
Total household income last year before taxes (CAD $[b]), n (%)	
40,000 or less	8 (27)
40,000 to 80,000	12 (40)
More than 80,000	10 (33)
Racial background, n (%)	
Métis	2 (7)
Latin American (Hispanic or Latin American descent)	1 (3)
Middle Eastern (eg, Arab, Persian, and West Asian descent [Afghan, Egyptian, Iranian, Kurdish, Lebanese, Turkish, etc])	1 (3)
South Asian (eg, Bangladeshi, Indian, Indo-Caribbean, Pakistani, and Sri Lankan descent)	1 (3)
White (eg, European descent)	25 (83)
Overall health, n (%)	
Excellent	0 (0)
Very good	5 (17)
Good	15 (50)
Fair	9 (30)
Poor	1 (3)
Chronic health conditions, n (%)	
High blood pressure	7 (23)
High cholesterol	8 (27)
Heart disease	2 (7)
Diabetes	6 (20)
Chronic bronchitis, emphysema, or COPD[c]	3 (10)
Rheumatoid arthritis	3 (10)
Chronic pain	11 (37)
Cancer	1 (3)
Depression	11 (37)

XSL•FO
RenderX

Demographic characteristics	Value
Anxiety	13 (43)
Schizophrenia	1 (3)
Substance use disorder (other than tobacco or caffeine)	2 (7)
Problem gambling	1 (3)
Cigarettes smoked per day, mean (SD)	16.4 (7.5)

[a]One participant selected "Don't know/prefer not to answer."

[b]Data collected in Canadian dollars. A currency exchange rate of US $1=CAD $1.33796 was applicable at the time of data collection.

[c]COPD: chronic obstructive pulmonary disease.

Primary Outcome Measures

Acceptability of the MCP-W App

Following the 28-day trial period, participants responded to the following question: "Please indicate how much you agree or disagree with the following statement: Using the app is likely to help me make changes to my smoking habits." In total, 11 out of 30 (37%, 95% CI 21%-56%) participants endorsed "agree or strongly agree." This result did not reach our predefined threshold for acceptability (≥50%), but landed in the "needs further work" range (26% - 49%). We assessed the acceptability of the intervention before and following the 28-day trial, in which the proportion of participants endorsing "agree or strongly agree" decreased from 43% (13/30) before using the app to 37% (11/30) at follow-up; however, this decrease was not statistically significant (Table 2).

Table . Primary, secondary, and tertiary outcome measures.

Outcome measures	Baseline	Follow-up	*P* value
Primary outcome			
Acceptability[a], n (%)			.09
Agree or strongly agree	13 (43)	11 (37)	
Neutral, disagree, or strongly disagree	14 (47)	16 (53)	
Secondary outcome			
Motivation[b], mean (SD)			
Importance	4.7 (0.6)	4.7 (0.6)	>.99
Commitment	4.6 (0.7)	4.5 (0.7)	.22
Ability	4.1 (0.9)	3.7 (1.0)	.40
Tertiary outcomes			
Cigarettes smoked per day[b], mean (SD)	16.4 (7.5)	14.6 (7.0)	.01
1 or more smoke-free days in the last 7 days[b], n (%)	2 (7)	6 (22)	.01

[a]Acceptability was measured on a 5-point Likert scale using the question: "Using the app is likely to help me make changes to my smoking habits."

[b]Data from 27 participants who completed the follow up survey.

Feasibility of the MCP-W App

Based on the Google Analytics data, 18 (60%, 95% CI 41%-77%) participants used the app for 7 or more days, which fulfilled the 50% feasibility criteria. Furthermore, 14 (47%) participants used the app for 1 - 7 days, 5 (17%) used it for 8 - 14 days, 3 (10%) for 15 - 21 days, and 8 (27%) for 22 - 28 days. Based on the number of days the MCP-W app was used, the proportion of participants who found the app acceptable was 14% (2/14; 1 - 7 days), 40% (2/5; 8 - 14 days), 67% (1/3; 15 - 22 days), and 63% (3/8; 22 - 28 days; Multimedia Appendix 5). There was a statistically significant correlation between the number of days the app was used and agreement with the acceptability of the app (P<.001).

Secondary and Tertiary Outcome Measures

Motivation to Quit

Participants had a high level of motivation with respect to "importance" of quitting smoking (mean score of 4.7, SD 0.6), and it did not significantly change from baseline to follow-up. There were no statistically significant changes in the "commitment" and "ability" measures (Table 2).

XSL•FO
RenderX

Smoking Behaviors

The number of cigarettes smoked per day significantly reduced from a mean of 16.4 (SD 7.5) cigarettes at baseline to a mean of 14.6 (SD 7.5) cigarettes at follow-up. The number of participants endorsing 1 or more smoke-free days in the last 7 days significantly increased from 2 (7%) at baseline to 6 (22%) at follow-up (Table 2). Based on the number of days the MCP-W app was used, the proportion of participants who reported a smoke-free day in the last 7 days at follow-up was 0% (1 - 7 days), 20% (8 - 14 days), 100% (15 - 22 days), and 25% (22 - 28 days). Furthermore, 1 participant reported quitting smoking at follow-up.

System Usability

The mean total score of the SUS was 70.2 (SD 12). The mean score of the response for the question "How likely are you to recommend this app to others?" was 6 (SD 2.0; on a scale of 0 to 10).

Qualitative Inquiry Into the Acceptability and Feasibility of the MCP-W App

Table 3 summarizes key themes and insights from qualitative inquiry with MCP-W participants.

Table . Summary of key themes and insights from qualitative results.

Theme	Subthemes or key insights	Supporting quotes
Motivation and readiness to quit	• App more useful for those highly motivated or ready to quit • Lower readiness linked to lower engagement	• "I was pretty motivated... that gave me a definitive start date." [ID 27] • "I'm not there yet... the app helped …but I think you have to be a 100% ready." [ID 01]
Barriers to quitting smoking	• Stress, mental health, daily hardships • Habitual use and nicotine addiction • Social contexts and emotional triggers	• "The more stressed I am, the more I smoke..." [ID 19] • "It's really hard to quit... the brain always makes excuses to get it" [ID 28]
Helpful app features	• Cigarette tracking increased awareness • Trigger identification and coping strategies • Resource library, visual feedback (savings, graphs, progress) • Gender-specific content resonated with users	• "...just reminding me to get outside and…do something other than just …smoking..." [ID 11] • "Everything they were talking about (in the videos) is how I felt..." [ID 18]
Features in need of improvement	• Burden of frequent tracking/logging • Usability issues (eg, journal entries, notifications) • Desire for more interactivity and real-life testimonials	• "It just seemed like the app wanted me to go into it every time..." [ID 07] • "I would like to get more like a free open option to write stuff down" [ID 19]
Desire for additional support	• Suggestions for live support or artificial intelligence–based chat features • App more effective when paired with counseling or pharmacotherapy	• "A live connection... someone like ...a sponsor." [ID 14] • "Maybe they should be starting that (medication) already before they use the app, so they already have a jump start, so to speak" [ID 29]

Motivation and Readiness to Quit Smoking

We found that participants' motivation and readiness to quit smoking influenced the acceptability and feasibility of the MCP-W app. More than half (n=16) of all participants interviewed mentioned that they were motivated to quit smoking. The app was seen as a helpful tool particularly for those highly motivated to quit. Some doubted the app's utility for women with lower levels of motivation or readiness.

> *Well, right now for my journey, I'm at a point where I know - I know I've got to quit smoking for my health, for my family's health, but mainly for mine just because I'm getting older I smoked for so long. Um, here's my cough…* [ID 03]

> *Well, I was pretty motivated when you first contacted me, so I think that's what helped [quit] was with the trial starting that that gave me a definitive start date or an end date, I should say.* [ID 27]

> *And I'm so tired of smoking, I don't want to smoke anymore. But I'm not there yet. I don't think, cause, I mean, I could analyze my smoking and I could, you know, the app helped me with that and, and figure it out and the triggers and… But, I think you have to be a 100% ready.* [ID 01]

Participants who mentioned that they were motivated and ready to quit smoking appeared to find the MCP-W app feasible, which they demonstrated by engaging with the MCP-W app longer than participants who mentioned they were not as motivated or ready. For example, one of the following

XSL•FO
RenderX

participants who was ready to quit used the MCP-W app for 29 days compared with the other participant who was not ready and only used the MCP-W app for 14 days.

> *But I'm determined right now with this app… It like with the full 29 days I've been using the app so far…* [ID 31]

> *I feel like, I have no motivation or willpower or something, and I'm just not making real changes here. So this is just kind of like, maybe it'll help other women, but not me.* [ID 16]

Some participants felt that the app is of most use to women who are ready to quit. Since they did not feel ready, it was less relevant to their needs.

> *If I was further into this stage of quitting smoking or wanting to quit smoking, it would definitely - it would help me. I just - I'm not there yet, so it was easy for me to not-not use the app, to not look at it. Right?* [ID 03]

> *I think that if somebody is ready, that it would be a valuable tool. It would be like an accountability tool… So I think I would be a good candidate for that if and when I was ready, I was just not ready…* [ID 32]

Participants cited numerous barriers to quitting smoking, some related to gender and others to the nature of the tobacco addiction including daily stressors (eg, employment, housing, finance, and relationships; n=16), the addictive nature of nicotine (eg, lifelong habit and hard to give up; n=14), boredom or nonstructured daily routine (n=4), social smoking (eg, with partners, friends, and work colleagues or managers; n=3), and mental health challenges (eg, depression; n=3).

> *The more stressed I am, the more I smoke, so, stress is a big trigger. I think when you're looking at, like, certainly me being a single mom, everything's on me to do all the time. So, I think quitting smoking would be a lot easier if I wasn't so stressed out all the time.* [ID 19]

> *I think that I would not have to love nicotine. Or I think that I would not have ever had to start smoking… But personally, I just don't think, once a person starts and once their brain like becomes addicted to nicotine, it's, it's really hard to quit cause the brain always makes excuses to get it.* [ID 28]

Some participants mentioned that the tailored content in the app was helpful to overcome gender-specific barriers to quitting.

> *I was able to cut out a few cigarettes when I'm walking to the bus stop. I made a conscious effort, because I thought, oh, that is what I really could skip. Um, most of my problem is that it is habit. I think the reasons why we smoke and when we're smoking, when you see it [in the MCP-W app], it really hits home. It's like, wow.* [ID 02]

> *I think since I've started using this [MCP-W app], I have reduced. So proud of that, um. It's just a matter of reducing to nothing.* [ID 10]

> *So, it actually made me think about that in retrospect of where I've been and what I've been doing and even just reminding me to get outside and go and do something other than just sitting here smoking. Just seeing that sheer number and volume day by day, helped me cut back. By the end of the month, it helped me cut back probably 10 to 11 cigarettes per day.* [ID 11]

MCP-W App Features That Were Helpful

Participants provided additional insights about the acceptability and feasibility of the MCP-W app. Factors that facilitated the acceptability and feasibility of the MCP-W app included the available tools and resources, for example, the feature of tracking the number of cigarettes they smoked. Participants (n=15) mention this feature helped them become more aware of the number of cigarettes they were smoking and held them accountable for changing their behavior (n=6). When tracking their cigarettes, they also became more aware of their smoking triggers such as emotions, places, and people they were with (n=10).

> *I like the idea that when I did use it, that it would ask me my, my attitude or who I was smoking with, which then if I used it properly, then it would trigger to when I am actually smoking. And then it would make me more aware of where I need to maybe change my surroundings or the urge to smoke.* [ID 04]

Participants mentioned that identifying their triggers and recording their coping strategies was helpful (n=15).

> *I did like the survey where you could put your motivations for quitting, how many you smoked a day, what to do, what are your triggers, how you could alleviate those stresses… when you think about it beforehand, so that when you're in the moment, you're able to go back to that app, look at it and say, "Okay, I said I would do this."* [ID 23]

Participants thought that the library feature provided helpful information about resources on quitting smoking, medications, such as nicotine replacement therapy, managing weight when quitting, and so on (n=14).

> *The library section had quite a bit of information as well with different smoking cessation medications and different supports and things like that that are very helpful.* [ID 25]

Other features found helpful by participants included visualizing their smoking through progress graphs (n=7), receiving notifications about reduction in number of cigarettes (n=6), and feedback about money saved (n=4).

> *I liked the notifications. So, when it said to me, like, you've smoked, you know, so many cigarettes less this morning than you did yesterday, that was kind of like a pat in the back moment for me.* [ID 01]

A number of participants provided feedback on the gender-specific content embedded in the app. In total, 9 participants mentioned the topics depicted in the videos and testimonials tailored to women who smoke resonated with them

XSL•FO
RenderX

and 5 mentioned that the videos motivated them or provided them information they did not know.

> *You know, I was reading the testimonials of people and it's, it's real people and you, you know, you can relate to them and then it just gives you more motivation to want to quit smoking.* [ID 04]

> *I think every video hit home, because everything they were talking about is how I felt or how I feel, right? So it kind of, that reminded me of like my own personal journey as well. So there was a lot of things in those videos that was just totally me.* [ID 18]

MCP-W App Features That Needed Improvement

Using the app requires a time commitment which was perceived as a burden to some. Half the participants (n=13) mentioned they found it difficult to find time or forgot to track every cigarette they smoked, and several (n=12) mentioned they needed reminders to log their cigarettes, with some specifying the need for customizable reminders 2 or 3 times a day. Indeed, 5 participants mentioned that women with children would have difficulty logging their cigarettes multiple times a day every day. Furthermore, 4 participants also reported logging factors associated with every cigarette was too tedious.

> *It just seemed like the app wanted me to go into it every time I had a cigarette and I just don't have that kind of time or resources to do that, especially while I'm at work - and like if you've got a young mom or anything like that, there's no way they could do that. Like, it'd be better if we could go in a couple of times a day and just pinpoint on a timeline when we had cigarettes.* [ID 07]

Some participants described functionality and design issues with key features of the MCP-W app that may have affected the usability. For example, some participants mentioned the journal did not allow independent entries (n=7), clicking on notifications did not open the app (n=2), or logging previous cigarettes did not allow logging the factors associated with them (n=4).

> *My journal part… I couldn't enter anything… I think it would be really beneficial if we can write things like 'Today really sucks and I'm really struggling with quitting smoking"… I would like to get more like a free open option to write stuff down.* [ID 19]

> *Sometimes I noticed on my phone, I'd get a notification pop up. And then I'd click on it, but it never said anything. So, I don't know if there were some problems with that.* [ID 10]

A few participants (n=5) wished for the MCP-W app to be more engaging in order to maintain participants' interest over time. Participants suggested that the MCP-W app should include interactive activities (ie, gamification) or dynamic content to keep interest levels high (n=8).

> *Maybe that's a good time to have something integrated into the app, like a game to play…* [ID 19]

> *Because there was not a true engagement. At the end of the day an app, it passes the excitement like a movie or like a show, like anything. It passes the time. So if there's no more like new things or like more engagement you will lose just interest naturally, I think…* [ID 22]

In total, 15 participants acknowledged that the app contains both gender-specific and gender-neutral content. Furthermore, 2 suggested that the library could be improved by including more gender-related information such as hormonal influences on smoking. They also preferred human actors in the videos as opposed to the animated characters (n=9), with additional success stories and strategies to overcome challenges to quit (n=6).

> *I think having a section to deal with hormones would be very beneficial like… things have changed drastically for me there, going through menopause… So, I think having that information involved, so you can take a closer look at yourself, because hormones trigger everything in your life, really.* [ID 11]

> *Oh, maybe, like real people (in the videos) would be nice and also maybe people who have been successful in what worked the best for them.* [ID 28]

A few participants (n=6) mentioned that they would like to have a feature that allows the user to connect to a support person or artificial intelligence (AI)–based chatbot for more personalized support.

> *A live connection where you can phone somebody and speak to them live. And not for psychological, like, not a psychiatrist or psychologist, but just someone, like, what do they call-a sponsor?* [ID 14]

> *I think in terms of quitting and smoking, you need more live support, that's one. But like I mentioned before, I think if it had like a chat open or linked to other participants that shared the experience, like, "Okay, how are you doing this day."* [ID 22]

Overall, participants (n=5) mentioned that the app can be a helpful tool, particularly when it is combined with other types of support such as behavioral counseling or smoking cessation medication, rather than a stand-alone solution.

> *Well, it did [help] on the few days that I had, for whatever reason, fewer smokes than other days. I think it did tell me like something like "You didn't have a cigarette this morning"… I think I'm stuck in some personal issues right now that that make it hard. I think you have to already be committed to doing something seriously and not have a lot of extraneous problems that are gonna defeat you before you even get started. Maybe somebody should already be starting patches or have a bunch of the other, you know, whether it's that medication, you know, pills or whatever they are, or other things. Maybe they should be starting that already before they use the app, so they already have a jump start, so to speak. Otherwise, just jumping into the app, it's really hard to just change everything without more, more of a push.* [ID 29]

XSL•FO
RenderX

In summary, participants' motivation and readiness to quit smoking significantly influenced the acceptability and feasibility of the MCP-W app. Women motivated to quit found the app helpful for tracking smoking behaviors, identifying triggers, and providing gender-specific resources. However, individuals with lower readiness to quit engaged less with the app. Participants recommended improvements, such as providing additional success stories of women, enhanced interactivity (eg, gamification), and personal coaching options. While the app could be beneficial, especially when paired with other cessation tools, it was not perceived as a stand-alone solution, particularly for individuals who face significant barriers to quitting.

Discussion

Principal Findings

In this study, we aimed to test the acceptability and feasibility of a gender-specific app for smoking cessation among women who smoke using a mixed method approach. Our study found that the MCP-W app was acceptable to slightly more than a third of women, with 37% (11/30) endorsing it as a helpful tool that can support their smoking cessation needs. As such, it did not meet the a priori acceptability criteria (>50%). Nonetheless, the app met feasibility criteria, with 60% (18/30) of participants using it for 7 or more days. The finding that the MCP-W app did not meet the predefined threshold for acceptability (only 11/30, 37% of participants rated the app as likely to help with smoking cessation) has important implications for future scalability and deployment. This result suggests that without additional support or enhancements, the MCP-W app in its current form may not be suitable as a stand-alone tool for a broad population of women who smoke. Instead, the app may be most useful for a subset of women who are already highly motivated to quit. This emphasizes the need for further development work to improve engagement and relevance, particularly for women exhibiting ambivalence about quitting. Future iterations of the app may benefit from incorporating more dynamic features (eg, personalized reminders, gamification, or peer support options) and tailoring content based on users' readiness to change.

The proportion of participants needed to consider the app acceptable (50%) was not met; however, the proportion of participants who used the app more frequently (more than 14 d) found it acceptable (7/11, 64%), as compared with the proportion of participants (4/19, 21%) who found it acceptable but used the app less frequently (14 d or less). The significant correlation between acceptability and feasibility of an app could be bidirectional, representing the increased acceptability of the app resulting in more use or conversely, as users engage with the app for longer, they find it more helpful. Previous research has found a similar correlation between acceptability and feasibility, even though they have been recognized as separate constructs [32,33]. In addition, the proportion of participants who mentioned they would find the MCP-W app acceptable at baseline was 43% (13/30), demonstrating their low pre-existing beliefs about the app's ability to help them make changes to their smoking. Research has shown that pre-existing beliefs about effectiveness can affect engagement with digital health interventions, for example, perceiving technology as distracting is linked with lower use metrics [34]. There may be a need for more education and awareness about the potential helpfulness of digital health interventions among women who are trying to reduce or quit smoking.

MCP-W use was associated with a positive change in participant smoking behaviors. The average number of daily cigarettes reduced from 16.4 to 14.6, and more women reported that they attempted to quit after using the MCP-W compared with baseline (6 vs 2). The latter was more pronounced in women who used the app more versus less than 14 days. This is an encouraging finding, although it should be interpreted with caution given that our study did not have a control group. This finding is consistent with other reports from digital interventions for smoking cessation [13,35] that speak to the potential effectiveness of digital health tools to support individuals in quitting. Our sample included women with a high level of motivation for quitting, which did not change significantly across the study timeline. Women rated the importance of quitting smoking as 4.7/5 at both baseline and follow-up. However, their perceived "ability" to quit smoking, measured by the statement "I could change my smoking habits (reduce or quit)" decreased in a nonsignificant way from 4.1/5 to 3.7/6. Since many women reported tracking their cigarettes and reducing during the trial, it is possible that these actions crystalized the challenges they face in their day-to-day lives, thus rating their ability to quit as lower.

Women's attitudes toward quitting were closely linked with whether they rated the MCP-W app as acceptable and the relative intensity of their use of the app (feasibility). Despite the majority of women reporting high motivation for quitting for common reasons, such as improving one's health, saving money, and for their families or children, many felt they were not ready to make this change. Lack of readiness was centered around life stressors such as material hardship, interpersonal conflict, lone parenting, mental illness, and medical problems. Dealing with the above was prioritized, and quitting smoking was left to be dealt with in the future, when life circumstances would change. This notion is consistent with the previous literature on women and smoking [7]. This explains, in part, why some women did not find the app acceptable. The app cannot directly change difficult life circumstances but is aimed at strengthening women's resilience in light of them. It is possible that combining the use of the app with support from cessation counselors could provide women with clarity regarding the individual barriers they face and guide them to develop helpful strategies to overcome those barriers [36].

The MCP-W app contained evidence-based gender-neutral behavior change techniques meant to facilitate change via goal setting, tracking daily cigarettes, and identifying triggers for smoking [23]. Most women found these techniques helpful. This is consistent with previous research on effective features that are recommended for use in smoking cessation apps [13]. However, in our study, women who were not ready to quit tended to put less effort into these behavioral tasks (ie, logging smoked cigarettes and identifying smoking triggers), compared with women who were ready to quit. Some women also commented on the efforts needed to enter information into the

XSL•FO
RenderX

app (ie, goal setting and tracking), and felt it was a big task, particularly from women with very little free time, like single mothers. It is possible that newer technologies will be able to make these tasks less tedious and easier to perform. For example, voice to app features, where you do not need to enter information manually to your device but use a voice command to do so (eg, "MCP-W, record I just smoked one cigarette in my car as I am driving back from work"). Voice-activated app functions are widely used today [37] and should be explored in the future as a method to reduce the burden of tracking a target behavior.

The MCP-W app contained gender-specific components including 4 animation videos and testimonials from women who quit smoking. The majority of women appreciated the tailored content and felt they could relate to the challenges expressed by other women. They sense that others sharing your struggles and understanding you is considered a key tenet of peer-supported activities in addiction recovery [38]. In addition, seeing other women succeed in quitting smoking may also improve one's self-efficacy of quitting as suggested by the social cognitive theory [39]. It could also help reduce the stigma and shame that is associated with smoking that is known to undermine cessation treatment seeking in some women [40]. It is possible that this will provide women with future gains of increased readiness to ask health care providers and family members for cessation support instead of smoking in secret or denying their smoking [41].

Overall, several participants who described themselves as motivated and ready to quit reported using the behavioral tools in the app with greater engagement. This helped them identify triggers for smoking and come up with alternatives to smoking in these situations. Other participants wished for live support to address challenges, such as chronic medical conditions or mental illness, that come in the way of quitting. This support is most likely to be achieved through interactive counseling by a human or agent. Nonetheless, our results suggest that the MCP-W could be helpful to a subset of women who are willing to use the app in line with their strong commitment to quitting. These findings also indicate that future research should consider stratifying or targeting participants based on motivation and readiness to quit, as well as testing the MCP-W app as part of a blended intervention model (eg, combined with counseling or pharmacotherapy). For example, for women who struggle to engage with the app due to life challenges and low motivation for quitting, a live counseling alternative could be doubly helpful. First, it provides ongoing encouragement and emotional support, and second, it solves problems like how to overcome specific barriers in one's life, such as having a smoking male partner. Alternately, some women felt that the support could be delivered using a virtual agent such as an AI-based chatbot counselor. Given the advances in the use of AI technologies in health care and the early success of smoking cessation chatbots [35], this is an avenue for scalable low-cost support that should be tested in the future.

Strengths and Limitations

The strengths of the study include the co-development of gender-specific content for a smoking cessation app by collaborating with patient partners and subject matter experts [42]. In addition, the use of mixed methods ensured that we captured in-depth contextual information about the quantitative findings of the app's acceptability and feasibility using qualitative data from the interviews [43]. However, most of the participants in this study were treatment-seeking individuals that reported continuing to smoke at the end of a treatment course with the STOP program. This represents a selection bias toward individuals with treatment resistance, which could limit generalizability to women who smoke in the general population [44]. In addition, the study sample size was small and nonparametric tests had to be conducted to explore the acceptability and feasibility. We also acknowledge that the study sample primarily consisted of middle-aged women with higher levels of education, which may limit the generalizability of our findings to younger individuals or those experiencing socioeconomic disadvantage. However, this was an exploratory study, and the attrition rate of participants was low, which allowed us to conduct appropriate statistical analyses.

Conclusions

The MCP-W app could serve as a valuable tool to support women in their journey to quit smoking. By delivering gender-specific smoking cessation content through a pre-existing gender-neutral platform, the app addresses unique needs. The acceptability and feasibility of the app are closely linked to the motivation and readiness of the woman attempting to quit. Improvements such as incorporating interactive features and offering additional personalized support could enhance its relevance. Feedback from participants highlighted the app's most helpful features, providing valuable insights. These findings underscore not only the current value of the MCP-W app but also its potential to evolve through continuous improvement, leading to greater effectiveness in women-focused smoking cessation interventions.

Acknowledgments

The authors wish to acknowledge the Center for Addiction and Mental Health (CAMH) Information Management Group for programming the My Change Plan–Women app, those contributing to the development of the animation videos—Mathangee Lingam, Darrelle Carpentier, Sophia Alcasabas, Adina Coriou—and patient partners Natasha Jaczek, Jung No, Lesley Quair, Mirabelle Infante, and Susan Haney involved in the development of the app content. This study was funded by the 2022 CAMH Womenmind Seed Grant competition.

XSL•FO
RenderX

Data Availability

The datasets generated or analyzed during this study are not publicly available due to Research Ethics Board–related privacy restrictions but are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare that they have no conflicting interests with respect to this manuscript; however, some authors have general disclosures to report. OCM receives salary support from the Department of Family and Community Medicine at the University of Toronto, and the INTREPID lab at CAMH. OCM reports receipt of funding from Canadian Institutes of Health Research. NM reports receipt of funding from Canadian Institutes of Health Research, Canadian Cancer Society, and from the Discovery Fund of CAMH. PS reports receipt of funding from Canadian Institutes of Health Research, Canadian Cancer Society, Pfizer Inc, Pfizer Canada, and Ontario Lung Association. PS also reports that through an open-tender process, Johnson & Johnson, Novartis, and Pfizer Inc are vendors of record for having provided smoking cessation pharmacotherapy for research studies at free or discounted rates. PS holds the Vice-Chair, Research and Giblon Professor in Family Medicine Research, a University Named Professorship at the University of Toronto.

Multimedia Appendix 1
Baseline demographic survey instrument.
[DOCX File, 20 KB - humanfactors_v12i1e71683_app1.docx]

Multimedia Appendix 2
Surveys for baseline and follow-up.
[DOCX File, 15 KB - humanfactors_v12i1e71683_app2.docx]

Multimedia Appendix 3
System Usability Scale survey instrument for the follow-up survey.
[DOCX File, 17 KB - humanfactors_v12i1e71683_app3.docx]

Multimedia Appendix 4
Interview guide.
[DOCX File, 21 KB - humanfactors_v12i1e71683_app4.docx]

Multimedia Appendix 5
Acceptability of the app based on number of app use days (acceptability is defined by participants who “agreed” or “strongly agreed” with the following statement: “Using the app is likely to help me make changes to my smoking habits”).
[PNG File, 13 KB - humanfactors_v12i1e71683_app5.png]

References

1. Kenfield SA, Stampfer MJ, Rosner BA, Colditz GA. Smoking and smoking cessation in relation to mortality in women. JAMA 2008 May 7;299(17):2037-2047. [doi: 10.1001/jama.299.17.2037] [Medline: 18460664]
2. Thun M, Peto R, Boreham J, Lopez AD. Stages of the cigarette epidemic on entering its second century. Tob Control 2012 Mar;21(2):96-101. [doi: 10.1136/tobaccocontrol-2011-050294] [Medline: 22345230]
3. Jafari A, Rajabi A, Gholian-Aval M, Peyman N, Mahdizadeh M, Tehrani H. National, regional, and global prevalence of cigarette smoking among women/females in the general population: a systematic review and meta-analysis. Environ Health Prev Med 2021 Jan 8;26(1):5. [doi: 10.1186/s12199-020-00924-y] [Medline: 33419408]
4. Allen AM, Oncken C, Hatsukami D. Women and smoking: the effect of gender on the epidemiology, health effects, and cessation of smoking. Curr Addict Rep 2014 Mar;1(1):53-60. [doi: 10.1007/s40429-013-0003-6] [Medline: 27213132]
5. Cosgrove KP, Wang S, Kim SJ, et al. Sex differences in the brain’s dopamine signature of cigarette smoking. J Neurosci 2014 Dec 10;34(50):16851-16855. [doi: 10.1523/JNEUROSCI.3661-14.2014] [Medline: 25505336]
6. Leventhal AM. The sociopharmacology of tobacco addiction: implications for understanding health disparities. Nicotine Tob Res 2016 Feb;18(2):110-121. [doi: 10.1093/ntr/ntv084] [Medline: 25890832]
7. Greaves L. The meanings of smoking to women and their implications for cessation. Int J Environ Res Public Health 2015 Jan 27;12(2):1449-1465. [doi: 10.3390/ijerph120201449] [Medline: 25633033]
8. Elicia Nademin M, Napolitano MA, Xanthopoulos MS, Fava JL, Richardson E, Marcus B. Smoking cessation in college-aged women: a qualitative analysis of factors important to this population. Addict Res Theory 2010 Dec;18(6):649-666. [doi: 10.3109/16066351003660601]

XSL•FO
RenderX

9. Voci S, Veldhuizen S, Tien S, et al. A comparison of quit outcomes for men and women in a smoking cessation program offering personalized nicotine replacement therapy and counseling in primary care clinics. Nicotine Tob Res 2021 Aug 29;23(10):1673-1681. [doi: 10.1093/ntr/ntab082] [Medline: 33912963]
10. Smith PH, Weinberger AH, Zhang J, Emme E, Mazure CM, McKee SA. Sex differences in smoking cessation pharmacotherapy comparative efficacy: a network meta-analysis. NICTOB 2017;19(3):ntw144. [doi: 10.1093/ntr/ntw144]
11. Smith PH, Bessette AJ, Weinberger AH, Sheffer CE, McKee SA. Sex/gender differences in smoking cessation: a review. Prev Med 2016 Nov;92:135-140. [doi: 10.1016/j.ypmed.2016.07.013] [Medline: 27471021]
12. Torchalla I, Okoli CTC, Bottorff JL, Qu A, Poole N, Greaves L. Smoking cessation programs targeted to women: a systematic review. Women Health 2012;52(1):32-54. [doi: 10.1080/03630242.2011.637611] [Medline: 22324357]
13. Vilardaga R, Casellas-Pujol E, McClernon JF, Garrison KA. Mobile applications for the treatment of tobacco use and dependence. Curr Addict Rep 2019 Jun;6(2):86-97. [doi: 10.1007/s40429-019-00248-0] [Medline: 32010548]
14. Bhattacharya A, Vilardaga R, Kientz JA, Munson SA. Lessons from practice: designing tools to facilitate individualized support for quitting smoking. ACM Trans Comput Hum Interact 2017;2017(3057-70):3057-3070. [doi: 10.1145/3025453.3025725] [Medline: 29123362]
15. Whittaker R, McRobbie H, Bullen C, Rodgers A, Gu Y. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2016 Apr 10;4(4):CD006611. [doi: 10.1002/14651858.CD006611.pub4] [Medline: 27060875]
16. Bert F, Giacometti M, Gualano MR, Siliquini R. Smartphones and health promotion: a review of the evidence. J Med Syst 2014 Jan;38(1):24346929. [doi: 10.1007/s10916-013-9995-7] [Medline: 24346929]
17. Tombor I, Shahab L, Brown J, Crane D, Michie S, West R. Development of SmokeFree Baby: a smoking cessation smartphone app for pregnant smokers. Transl Behav Med 2016 Dec;6(4):533-545. [doi: 10.1007/s13142-016-0438-0] [Medline: 27699682]
18. Dotson JAW, Pineda R, Cylkowski H, Amiri S. Development and evaluation of an iPad application to promote knowledge of tobacco use and cessation by pregnant women. Nurs Womens Health 2017;21(3):174-185. [doi: 10.1016/j.nwh.2017.04.005] [Medline: 28599739]
19. Gordon JS, Armin J, D Hingle M, et al. Development and evaluation of the See Me Smoke-Free multi-behavioral mHealth app for women smokers. Transl Behav Med 2017 Jun;7(2):172-184. [doi: 10.1007/s13142-017-0463-7] [Medline: 28155107]
20. Martinez-Gutierrez J, Domínguez A, López C, et al. “Appagalo” a customized mobile health intervention (mHealth) for smoking cessation in women: a randomized controlled trial. Tob Use Insights 2023;16:1179173X231152316. [doi: 10.1177/1179173X231152316] [Medline: 36844174]
21. University of Toronto. My Change Plan smoking cessation app. Apple App Store. URL: https://apps.apple.com/ca/app/mychangeplan/id1405709545 [accessed 2025-09-04]
22. My Change Plan. Intrepid Lab CAMH. URL: https://intrepidlab.ca/en/teach/Documents/My%20Change%20Plan%20Edition%208.pdf [accessed 2025-09-04]
23. Michie S, Richardson M, Johnston M, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
24. Clinical Practice Guideline Treating Tobacco Use and Dependence 2008 Update Panel, Liaisons, and Staff. A clinical practice guideline for treating tobacco use and dependence: 2008 update. A U.S. Public Health Service report. Am J Prev Med 2008 Aug;35(2):158-176. [doi: 10.1016/j.amepre.2008.04.009] [Medline: 18617085]
25. Hartmann-Boyce J, Livingstone-Banks J, Ordóñez-Mena JM, et al. Behavioural interventions for smoking cessation: an overview and network meta-analysis. Cochrane Database Syst Rev 2021 Jan 4;1(1):CD013229. [doi: 10.1002/14651858.CD013229.pub2] [Medline: 33411338]
26. Melamed OC, Mehra K, Panda R, et al. A gender-informed smoking cessation app for women: protocol for an acceptability and feasibility study. JMIR Res Protoc 2024 Dec 10;13:e60677. [doi: 10.2196/60677] [Medline: 39433391]
27. Miller WR, Rollnick S. Motivational Interviewing: Helping People Change: Guilford Press; 2012.
28. Hyzy M, Bond R, Mulvenna M, et al. System Usability Scale benchmarking for digital health apps: meta-analysis. JMIR Mhealth Uhealth 2022 Aug 18;10(8):e37290. [doi: 10.2196/37290] [Medline: 35980732]
29. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res 2017 Jan 26;17(1):88. [doi: 10.1186/s12913-017-2031-8] [Medline: 28126032]
30. Miller WR, Johnson WR. A natural language screening measure for motivation to change. Addict Behav 2008 Sep;33(9):1177-1182. [doi: 10.1016/j.addbeh.2008.04.018] [Medline: 18558466]
31. Azungah T. Qualitative research: deductive and inductive approaches to data analysis. QRJ 2018 Nov 15;18(4):383-400. [doi: 10.1108/QRJ-D-18-00035]
32. Weiner BJ, Lewis CC, Stanick C, et al. Psychometric assessment of three newly developed implementation outcome measures. Implement Sci 2017 Aug 29;12(1):108. [doi: 10.1186/s13012-017-0635-3] [Medline: 28851459]
33. Cho E, Lyon AR, Tugendrajch SK, Marriott BR, Hawley KM. Assessing provider perceptions of training: initial evaluation of the acceptability, feasibility, and appropriateness scale. Implement Res Pract 2022;3(26334895221086269):26334895221086269. [doi: 10.1177/26334895221086269] [Medline: 37091090]

XSL•FO
RenderX

34. Borghouts J, Eikey E, Mark G, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. J Med Internet Res 2021 Mar 24;23(3):e24387. [doi: 10.2196/24387] [Medline: 33759801]
35. He L, Balaji D, Wiers RW, Antheunis ML, Krahmer E. Effectiveness and acceptability of conversational agents for smoking cessation: a systematic review and meta-analysis. Nicotine Tob Res 2023 Jun 9;25(7):1241-1250. [doi: 10.1093/ntr/ntac281] [Medline: 36507916]
36. Urquhart C, Jasiura F, Poole N, Nathoo T, Greaves L. Liberation!: Helping Women Quit Smoking: A Brief Tobacco Intervention Guide: British Columbia Centre of Excellence for Women's Health; 2012.
37. Chung AE, Griffin AC, Selezneva D, Gotz D. Health and fitness apps for hands-free voice-activated assistants: content analysis. JMIR Mhealth Uhealth 2018 Sep 24;6(9):e174. [doi: 10.2196/mhealth.9705] [Medline: 30249581]
38. Tracy K, Wallace SP. Benefits of peer support groups in the treatment of addiction. Subst Abuse Rehabil 2016;7(143-54):143-154. [doi: 10.2147/SAR.S81535] [Medline: 27729825]
39. Bandura A. Social cognitive theory: an agentic perspective. Asian J of Social Psycho 1999 Apr;2(1):21-41 [FREE Full text] [doi: 10.1111/1467-839X.00024]
40. David JC, Fonte D, Sutter-Dallay AL, et al. The stigma of smoking among women: a systematic review. Soc Sci Med 2024 Jan;340(116491):116491. [doi: 10.1016/j.socscimed.2023.116491] [Medline: 38096599]
41. Hwang SW, Lee HJ, Kim CH, et al. The effect of hidden female smoking on the relationship between smoking and cardiovascular disease. Cardiol J 2021;28(5):716-727. [doi: 10.5603/CJ.a2020.0001] [Medline: 31960946]
42. Silvola S, Restelli U, Bonfanti M, Croce D. Co-design as enabling factor for patient-centred healthcare: a bibliometric literature review. Clinicoecon Outcomes Res 2023;15(333-47):333-347. [doi: 10.2147/CEOR.S403243] [Medline: 37220481]
43. Wasti SP, Simkhada P, van Teijlingen ER, Sathian B, Banerjee I. The growing importance of mixed-methods research in health. Nepal J Epidemiol 2022 Mar;12(1):1175-1178. [doi: 10.3126/nje.v12i1.43633] [Medline: 35528457]
44. Selby P, Voci S, Zawertailo L, Baliunas D, Dragonetti R, Hussain S. Public health impact of a novel smoking cessation outreach program in Ontario, Canada. BMC Public Health 2018 Sep 14;18(1):1117. [doi: 10.1186/s12889-018-6012-6] [Medline: 30217187]

Abbreviations

AI: artificial intelligence
CAMH: Center for Addiction and Mental Health
MCP: My Change Plan
MCP-W: My Change Plan–Women
mhealth: mobile health
REDCap: Research Electronic Data Capture
STOP: Smoking Treatment for Ontario Patients
SUS: System Usability Scale
TFA: Theoretical Framework of Acceptability

Edited by H Witteman; submitted 28.01.25; peer-reviewed by C Martinez-Vispo, D Verran; revised version received 30.05.25; accepted 30.05.25; published 25.09.25.

Please cite as:
Melamed OC, Mehra K, Gayapersad A, Panda R, Minian N, Zawertailo L, Buckley L, Maslej M, Greaves L, Brabete AC, Rose J, Ratto M, Selby P
Testing the Acceptability and Feasibility of a Gender-Informed Smoking Cessation mHealth App for Women: Mixed Methods Approach
JMIR Hum Factors 2025;12:e71683
URL: https://humanfactors.jmir.org/2025/1/e71683
doi:10.2196/71683

© Osnat C Melamed, Kamna Mehra, Allison Gayapersad, Roshni Panda, Nadia Minian, Laurie Zawertailo, Leslie Buckley, Marta Maslej, Lorraine Greaves, Andreea C Brabete, Jonathan Rose, Matt Ratto, Peter Selby. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 25.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Design Guidelines for a Game-Based Physical Rehabilitation System: Focus Group Study

Ahmad M S Elaklouk[1*], PhD; Ratna Zuarni Ramli[2*], PhD; Samah M M Alakklouk[3], MSc; Norsafinar Rahim[3], PhD

[1]School of Computing and Informatics, Universiti Teknologi Brunei, Jalan Tungku Link Gadong, Bandar Seri Begawan, Brunei Darussalam

[2]Universiti Teknologi MARA Cawangan Negeri Sembilan Kampus Seremban, Seremban, Malaysia

[3]Centre for Instructional Technology and Multimedia, Universiti Sains Malaysia, Penang, Malaysia

[*]these authors contributed equally

Corresponding Author:
Ahmad M S Elaklouk, PhD
School of Computing and Informatics, Universiti Teknologi Brunei, Jalan Tungku Link Gadong, Bandar Seri Begawan, Brunei Darussalam

Abstract

Background: Incorporating serious games and advancements in information and communication technologies into physical rehabilitation can substantially enhance the process, provide unique benefits, and improve its effectiveness and efficiency. While recent literature highlights various game-based interventions for physical rehabilitation, there is a lack of comprehensive guidance on how to design and develop systems that effectively address the actual needs of therapists, practitioners, and individuals with physical disabilities.

Objective: The objective of this study was to explore the intentions, needs, and desires of therapists and other practitioners, as well as to examine the factors and determinants influencing the effectiveness and efficacy of game-based physical rehabilitation, since therapists and other health care practitioners play crucial roles in both patient recovery and the establishment of an effective game-based therapy.

Methods: A design science approach was adopted to achieve this research objective. A focus group of 27 participants was conducted to gather feedback, identify user needs, and understand the requirements for game-based physical rehabilitation. The participants first tested commercially available games and then evaluated mock-ups of the proposed game prototypes.

Results: This study provides essential design insights and guidelines for designers and researchers, focusing on the practical needs and requirements of game-based physical rehabilitation systems.

Conclusions: As proof of concept, these guidelines will be used in the next phase of our research, which involves designing and developing a game-based physical rehabilitation system.

(JMIR Hum Factors 2025;12:e67336) doi:10.2196/67336

KEYWORDS

serious games; physical rehabilitation; game design; brain damage; kinect; game; gamification; therapy; therapist; practitioner; game-based; physical; rehabilitation; virtual reality; VR; movement; motor

Introduction

The application of technology in rehabilitation is experiencing significant and rapid expansion [1]. The integration of serious games and technological advancements enhances the rehabilitation process, creating new opportunities to address stakeholders' needs and expectations by (1) enhancing the quality and efficacy of rehabilitation, (2) offering alternatives that sustain patient motivation by mitigating the repetitive nature of traditional rehabilitation, (3) providing tailored therapy levels, and (4) overcoming limitations in resources and facilities associated with conventional rehabilitation methods [2-13]. For example, individuals who have experienced a stroke are motivated to use technologies and games to facilitate their functional recovery [14]. Furthermore, the study by Cruz et al [15] revealed that commercial gaming consoles and off-the-shelf games are more affordable and offer numerous options and variations, which could facilitate the implementation of virtual reality games in clinical practice. However, the usability of virtual reality games is a crucial factor to consider in the design of these interventions. The study by Daoud et al [16] introduced a game-based rehabilitation system that includes 3 exercises and a computerized assessment method. This system evaluates the accuracy of right arm movements during gameplay by analyzing tracking data collected from a Microsoft Kinect sensor. The study by Jayasree-Krishnan et al [17] discussed several barriers to the seamless integration of technological solutions into clinical care, including issues with accessibility

XSL•FO
RenderX

to quality rehabilitation, adaptability to individual patient differences, compliance, and engagement with the rehabilitation process.

In recent years, the implementation of diverse technologies has strengthened telerehabilitation, enhancing rehabilitation processes and ensuring patients receive the necessary services [11,18,19]. Integrating physiotherapy with technologies such as virtual reality and video games enhances the rehabilitation of poststroke patients compared to physiotherapy alone. This approach offers an immersive and interactive platform that boosts patient engagement and motivation during rehabilitation, while also providing a safe environment for practicing and improving physical abilities [2,20,21]. Evidence from a range of studies highlights its potential to improve outcomes for individuals with various physical impairments [2]. Takei et al [22] assessed the safety, feasibility, and acceptability of using the Nintendo Switch video game Ring Fit Adventure for upper and lower limb strength training in older adults with musculoskeletal conditions, alongside conventional physiotherapy over 6 sessions. The authors concluded that this approach is a safe and feasible adjunct to a rehabilitation program, as long as it is supervised, particularly in the initial stages. In the study by Wittmann et al [23], the game "Meteors" was introduced, which involves a virtual robotic arm that replicates the player's arm movements to catch meteors falling onto a planet. In the study by Wittmann et al [24], the game "Slingshot" was used to enhance arm coordination and improve the precision of aiming and extending arm movements. This game focuses on exercising the flexion and extension of the elbow. In this game, scoring is based on performance, and the difficulty level can be dynamically adjusted to maintain the patient's motivation and commitment. In the study by Goršič et al [25], 4 video games were developed for arm rehabilitation. One game involved competitive play, where the patient competed against another individual, such as a friend, relative, or therapist. In addition, there were 2 cooperative games, where the patient and another player worked together to play against the computer, and 1 single-player game where the patient played alone against the computer. The study revealed that competitive games significantly enhanced functional recovery and improved patients' quality of life more than conventional rehabilitation exercises.

In the study by Ma et al [26], a full-body rehabilitation video game was developed using Kinect V2 as an input device. The game accommodates both sitting and standing positions, supporting various movements such as gross motor actions (steps, jumps, and squats) and fine motor gestures (handshakes, palm rotations, and hand opening and closing). This system tracks the player in 3D space and registers real-time data, demonstrating positive outcomes in motor function and the performance of basic daily activities among chronic poststroke patients. The study by [19] argued that successful games in the industry, known for their immersive qualities, incorporate 8 key features: game control technology, display technology, rewards, social elements, avatars, game difficulty, music or sound, and graphical realism. They proposed that future research could explore different game genres, aspects like game narrative and competition, improvements in user interface, and the integration of immersive technologies such as virtual and augmented reality.

The exploration of therapeutic alternatives, combined with new technologies, enables professionals to use a variety of tools for patient rehabilitation and treatment. However, it is crucial to comprehend the functionalities, potential applications, and limitations of these tools [27]. Future research should aim for higher methodological quality, larger sample sizes, and well-defined rehabilitation programs to mitigate inconsistencies in evidence within this field [3,28].

Video games designed for therapeutic use have distinct requirements compared to regular video games. There are differences in user characteristics, where the patient serves as the primary player and the therapist as a secondary user, in contrast to players of regular video games. These users also have different objectives, and the context varies [27]. Considering the patient as the sole end user for any rehabilitation intervention is an aberration and diverges from standard practice, as therapists are primarily responsible for motivating, guiding, and assessing the patient. The study by Krishnan et al [14] revealed that, although individuals who have experienced a stroke were amenable to integrating exergames into their rehabilitation regimen, they noted that these games could not replace the essential role of therapist supervision. Integrating new technologies into clinical practice is complex, as it requires addressing the needs of both clients and clinicians. By considering these factors, technology developers can improve the chances that clinicians will embrace and adopt innovative technologies [29].

The study by Hamilton et al [30] highlighted 2 key factors for optimizing patient engagement: adequate support and a perceived benefit from the technology. It concluded that patients are more likely to engage with technology during rehabilitation when it is personalized by a therapist to fit their specific needs. Therapists and other health care practitioners play crucial roles in both patient recovery and the establishment of an effective rehabilitation system [31,32].

Since rehabilitation aims to restore patient independence and enhance the quality of life, achieving optimal outcomes requires a cohesive multidisciplinary team comprising doctors, psychologists, occupational therapists, nurses, social workers, and physiotherapists collaborating effectively [27]. Despite numerous studies on technology acceptance, there is still a lack of understanding regarding the factors that influence health care professionals' adoption of technological innovations [27,33]. The study by Jung et al [34] argued that although there has been speculation about the roles of therapists in game-based therapy and suggestions regarding varying levels of patient engagement, no comprehensive studies have been conducted to investigate these matters in depth. Understanding how games influence and support the different roles of therapists remains an area for future study [34]. The study by Wu et al [35] argued that, although some serious games have been developed for physical therapy, little work has been conducted through a participatory design approach. The authors highlight the importance of involving diverse stakeholders in the design process to create more effective and user-centric serious games for rehabilitation.

XSL•FO
RenderX

Therefore, effective implementation and use require collaboration from all parties involved in the process, including physicians, nurses, social workers, and other allied health professionals [27]. Thus, the objective of this study is to explore the intentions, needs, and desires of therapists and other practitioners, as well as to examine the factors and determinants influencing the effectiveness and efficacy of game-based physical rehabilitation.

Methods

Ethical Considerations

This study was conducted in accordance with ethical standards for research involving human participants. Ethics approval was obtained from the Ethics and Research Committee, Scientific Research Department, University College of Science and Technology, Gaza, Palestine.

All participants were provided with comprehensive information regarding the study's purpose, procedures, and ethical considerations, enabling them to make an informed decision about their involvement. Informed consent was obtained from all participants, who were explicitly informed of their right to decline or withdraw from the study at any point without any repercussions. Participants' personal information and responses were treated with strict confidentiality. Data were anonymized and stored securely, ensuring that individual identities could not be disclosed in any reports or publications.

Study Design Overview

This study used a design science research methodology, encompassing 2 iterations. The first iteration investigated the therapeutic game requirements and proposed guidelines to guide developers and practitioners in the design of therapeutic games; this iteration will be documented in this paper. The second iteration will use the findings from the first iteration to construct the main artifact, which is a rehabilitation gaming system for physical rehabilitation. The earlier work outlines a more detailed description of the design science research methodology [36,37]. Therefore, in the first iteration, the limited expertise of therapists in using game-based physical rehabilitation for acquired brain injuries significantly prolonged the process of determining the appropriate research strategy. Hence, before any actual game development, conducting a focus group with therapists and related practitioners is a great starting point as it helps generate interesting discussions between physical therapists, occupational therapists, service providers, special educators, psychologists, game designers, software engineers, other practitioners, and the researchers. This process is crucial to understanding how rehabilitation exercises can be translated into a suitable game by identifying the required movements, constraints, difficulty levels, and outputs to design game interventions that align with these exercises. To achieve this objective, the first purpose of the focus group is to evaluate the low-fidelity game prototypes (paper mock-ups) developed and assess how well these prototypes meet the requirements of the rehabilitation exercises. These low-fidelity game prototypes were developed through collaborative team brainstorming sessions and an in-depth review of the literature [38-57]. The second purpose of the focus group is to evaluate the commercial games. The study by Pedraza-Hueso et al [58] determined the feasibility of using the Xbox Kinect gaming console (Kinect Adventure game series) in the rehabilitation of a single patient. This study highlights a usability gap, noting that further research is needed to assess the usability of the Xbox Kinect gaming system in clinical settings and across different patient populations. Whole-body motion capture VR allows a unique opportunity for individuals to experience a heightened sense of realism during task-specific therapeutic activities. However, clinicians need to be able to match a game's components to an individual's functional deficits [59]. Even though none of these games were designed for clinical purposes. The study by Rizzo and Kim [60] stated that "Designers of rehabilitation tasks can benefit from examining the formulas that commercial game developers use." For example, the study by Burke et al [61], discussed that scoring mechanisms as well as mechanisms that maintain player engagement are commonly used in commercial games and can be used similarly in a rehabilitation setting. Therefore, testing commercial games such as the Kinect Adventure game series and Motionsports Adrenaline with therapists can also provide valuable insights into design, usability, and the various mechanisms used to motivate and maintain patient engagement. It can also help identify the missing components that make these games unsuitable for physical rehabilitation. Based on the focus group results, we can then move forward to create high-fidelity prototypes.

Focus Group

The purpose of the focus group was to gather feedback from participants, explore their needs and expectations for a serious game, and gain insights into the requirements for game-based physical rehabilitation. The focus group (Figures 1-3) was held at a designated presentation venue in a hotel in Palestine and was divided into 2 sessions. The first session, lasting approximately 3.5 hours, involved the evaluation of selected commercial games. It began with an introductory briefing by the principal investigator, who outlined the study's objectives. Research assistants initially facilitated the gameplay, helping participants become familiar with the interface. As the session progressed, some participants became comfortable navigating the game menus independently. Following the gameplay, an open discussion was conducted to gather participants' feedback on their experiences and expectations. A scheduled break followed the conclusion of this session.

XSL•FO
RenderX

Figure 1. Focus group session briefing by the researcher.

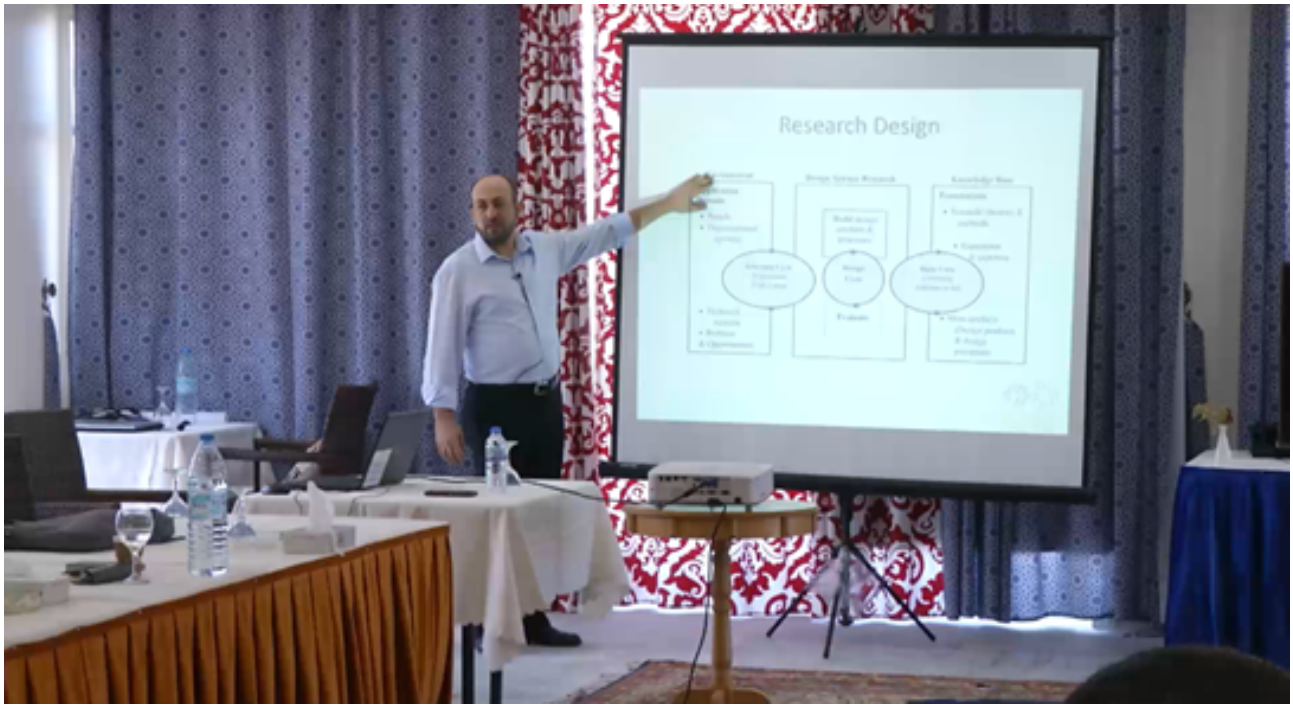


Figure 2. Focus group session. The participants are playing the Kinect Adventures game.



XSL•FO
RenderX

Figure 3. Focus group session. The participants are playing the Kinect Motionsports Adrenaline game.



The second session, held after the break and lasting approximately 4 hours, focused on evaluating the game prototype mock-ups developed by the research team. The principal investigator presented each mock-up sequentially using a projector. Participants were then provided with structured questionnaires, which allowed them to share detailed feedback and offer suggestions for improvement. The completed questionnaires were collected by the research assistants for further analysis. Both sessions were audio- and video-recorded to support detailed qualitative data analysis.

Participants

The focus group comprised 27 participants, including 4 occupational therapists, 10 physiotherapists, 2 game designers, 3 software engineers, 3 special educators, 3 psychologists, and 2 researchers. Table 1 shows the list of focus group participants, encompassing all participant categories.

Table . List of participants.

Role	Participant ID
Occupational therapists	P1-P4
Physiotherapists	P5-P14
Game designers	P15-P16
Software engineers	P17-P19
Special educators	P20-P22
Psychologists	P23-P25
Researchers	P26-P27

Intervention Using Commercial Games

Motion sensing and tracking technologies, such as the high-performance 3D depth-sensing camera Astra Pro Plus, can be used in diverse applications, including robotics, serious games, virtual reality, and augmented reality. Motion-sensing devices, such as the Kinect, fall under the category of gesture-tracking devices, using a user interface known as a natural user interface. Consequently, participants moved around the play area, raised and lowered their arms, and adjusted their body positions to complete the tasks. A total of 2 commercially available games were selected: Kinect Adventures and Motionsports Adrenaline. Kinect Adventures was selected for its entertaining nature and its variety of mini-games, including Space Pop, 20,000 Leaks, Reflex Ridge, River Rush, and Rally Ball. These mini-games offer diverse scenarios and require different body movements. For instance, in 20,000 Leaks, players use their hands and feet to plug holes that appear when fish hit the tank glass. In Rally Ball, players hit a virtual ball toward targets, while in Reflex Ridge, they avoid obstacles while collecting coins. The second game chosen is Motionsports Adrenaline, which offers a range of sports experiences. It includes six distinct mini-games—Wingsuit, Kitesurfing, Kayaking, Mountain Biking, Rock Climbing, and Extreme Skiing—designed to simulate high-stakes stunts that most people would not attempt in real life. Unlike Microsoft's Kinect Sports series, it stands out with a focus on realistic visuals, distinguishing it from many other Kinect games.

XSL•FO
RenderX

Intervention Using Low-Fidelity Paper Mock-Ups

Rabbit Hunting

This game prototype displays several holes, with a rabbit hiding inside one of them. Once the game begins, the rabbit pops out of one of the 5 holes and stays in a hole for a short period of time (depending on the patient's ability), as shown in Figure S1 in Multimedia Appendix 1. The player must hit the appearing rabbit with their hand to score a point. Using motion sensing devices, such as Kinect, the arm movement of the patient is mapped on the corresponding limb of the avatar (a model of a human represented on the screen). Feedback, such as a buzzing sound, an updated score counter, and a change in the rabbit's color, is displayed to the patient on the screen. The game prototype aims to score as many points as possible within a specified time limit, which can be adjusted by the therapist according to the patient's ability. The game becomes more challenging by increasing its speed or introducing a cat that the player must avoid. The therapeutic goals are to improve shoulder, elbow, and wrist movement; enhance multifunctional ability; develop eye-hand coordination; and strengthen muscle power.

Picking the Balls

The game prototype comprises three levels, each designed with increasing difficulty. The patient's arm movements are tracked using motion-sensing devices, such as Kinect, and mirrored by a human avatar displayed on the screen. In the first level (see Figure S2 in Multimedia Appendix 1), the player simply places balls into a basket, with no requirement regarding color or order. Upon successful completion, the player advances to the second level (see Figure S3 in Multimedia Appendix 1), where colored balls are introduced. In this stage, the system prompts the player to pick a ball matching the color shown on the screen. In the final level (see Figure S4 in Multimedia Appendix 1), balls are marked with random numbers, and the player must place them into the basket in either ascending or descending numerical order. In the advanced levels, the patient will be asked which hand to use to pick up the balls. The therapeutic goal is to improve cognitive skills, endurance, eye-hand coordination, awareness, and the range of motion (ROM) of the arm.

Organizing Eggs

This prototype shows an egg appearing in the middle of the screen. Using motion sensing devices, such as Kinect, the avatar (the virtual character portraying the player on the screen) mimics the movements that the player makes. The player has to pick and put each egg in the specified place in the basket as shown in Figure S5 in Multimedia Appendix 1. The player gets a point for every egg put in the right place. The therapeutic goal is to improve motor functioning, balance, strengthen the muscle power in the whole upper limb, and movement of reach, grasp, and release.

Shuffling Shelves

This game prototype intends to work on the patient's cognitive and physical capabilities, as the patient has to grab or pick the book from the upper shelf and put it on the lower one (see Figure S6 in Multimedia Appendix 1), depending on both the book and the shelf's colors. The arm movement of the patient is mapped on the corresponding limb of the avatar (a model of a human represented on the screen) using the motion-sensing device recognition feature. The game begins with a series of colored books appearing on the upper shelf, above the divided and colored lower shelves. After the patient grabs the book and puts it on the suitable shelf, he gets a point. The therapeutic goals include improving endurance and balance, enhancing trunk control, developing right-left discrimination, refining the patient's coordination skills, and promoting arm movements such as extension and shoulder abduction.

Hold the Boat

This prototype focuses on the arm's strength and stability. Using the motion sensing device recognition feature, the avatar (the virtual character portraying the player on the screen) mimics the movements that the player makes. In this game prototype, the patient has to hold the boat and keep moving it to the goal (the end of the channel). Visual cues appear on the screen to guide the player, as shown in Figure S7 in Multimedia Appendix 1. The patient must not only reach the game objects (the boat) but also accompany their movements toward the final destination. This will promote the patient's stability and persistence. The therapeutic goals are to improve grasping, hand awareness, stability, and shoulder ROM.

Reaping the Oranges

This game prototype includes 2 modes: single-player and multiplayer. In the single mode, the patient plays alone, collecting as many oranges as possible from the tree during a certain period of time. Using the motion sensing device recognition feature, the avatar (the virtual character portraying the player on the screen) mimics the movements that the player makes. The player collects the oranges by simply moving their hands to touch the orange on the tree, then placing it in the baskets displayed on the screen, as shown in Figure S8 in Multimedia Appendix 1. The game ends when the time limit is reached. The patient's score is displayed, which is the number of oranges collected during the game. In the multiplayer mode, the goal of the game does not change, but there are two options: competition or collaboration. In the competition, both participants (either 2 patients or a patient and a therapist) must collect as many oranges as possible. The winner will be the one who reaps the most oranges. In the collaboration, both participants must collect as many oranges as possible and place them in shared baskets. The game ends when the time limit is reached or when they reach the target number of oranges, which can be set and controlled by the therapist. In the next game levels, the patients can be asked to use either their right or left arm to place the oranges in the right or left basket, depending on their condition. The therapeutic goal is to strengthen the upper limb, improve eye-hand coordination, range of motion for the arm, "Right-Left" discrimination, weight shifting, abduction, and adduction of the upper limbs of patients, and social interaction. In the next two game prototypes, the objective remained the same as described above, with the difference that the game objects the player must touch would be moving rather than static. This adds complexity, promoting stability and persistence.

XSL•FO
RenderX

Ice Balls

In this game prototype, the patient has to move his hands (stretching his arms at different lengths) to catch ice ball drops from the top of the screen as shown in Figure S9 in Multimedia Appendix 1. The game begins with a small number of ice balls falling slowly from the top of the screen. The arm movement of the patient is mapped on the corresponding limb of the avatar (a model of a human represented on the screen) using the motion-sensing device recognition feature. The more ice balls a player catches, the higher the score is. In the next levels, the number and speed of dropping ice balls increased, and in more difficult places to reach. The therapeutic goal is to improve upper limb movement and coordination, motor planning, timing of reaching, and the supination and pronation movement of the forearm.

Blowing the Bubbles

In this game prototype, the player must move their hand (or both hands). Using the motion sensing device recognition feature, the avatar (the virtual character portraying the player on the screen) mimics the movements that the player makes. The game starts with a girl on the side of the screen generating bubbles, and the player starts collecting points by blowing those bubbles with his or her hand as shown in Figure S10 in Multimedia Appendix 1. The game gets harder by creating more bubbles and making them faster and harder to blow. The therapeutic goal is to improve supination and pronation movements of the forearm, as well as shoulder and elbow muscle power and elbow ROM.

Matching Shapes

This game prototype focuses on cognitive and movement skills. The game starts as several shapes appear randomly on the screen (see Figure S11 in Multimedia Appendix 1). The arm movement of the patient is mapped on the corresponding limb of the avatar (a model of a human represented on the screen) using the motion sensing device recognition feature. The patient must use both arms to select matching shapes or identical symbols. The earlier levels have fewer and simpler shapes, while the advanced levels present more complex shapes. The player gets a point for correctly matching shapes. The therapeutic goal is to improve bilateral hand functioning, movement for the whole upper limb, coordination, and depth perception.

Colored Boxes

This game prototype focuses on cognitive and movement rehabilitation. The arm movement of the patient is mapped on the corresponding limb of the avatar (a model of a human represented on the screen) using the motion-sensing device recognition feature. The patient uses both upper limbs to transport boxes from one side to the other. On the left of the screen, the colored boxes (red, yellow, and blue) gradually depart from one of the 3 levels (the moving belts). The player stands in the middle of the screen using his upper extremities to make a bridge connecting both sides so that the appropriate colored box moves towards the same truck color as shown in Figure S12 in Multimedia Appendix 1. As the game levels advance, the speed of the colored boxes increases. The therapeutic goal is to improve bilateral hand functioning, endurance, coordination of movement in the upper limb, and hand control.

Fallen Apples

In this prototype, the patient raises their hands and moves them left and right, causing the apples to drop, as shown in Figure S13 in Multimedia Appendix 1. Using the motion sensing device recognition feature, the avatar (the virtual character portraying the player on the screen) mimics the movements that the player makes. As the game begins, the trees appear with apples on them, and the player starts moving their hands left and right. A correct movement causes the apples to drop, earning the patient a point. As the levels increase, the difficulty also rises, requiring the player to move their arms to more challenging angles. The therapeutic goals are to strengthen the upper limb, maintain balance, enhance motor planning, and refine movement timing.

Boat Driver

This game prototype relies on the player's ability to use both hands simultaneously. Using the motion sensing device recognition feature, the avatar (the virtual character portraying the player on the screen) mimics the movements that the player makes. The game starts with the boat in a river, as shown in Figure S14 in Multimedia Appendix 1. The player tries to lead the boat to the goal by moving their hands to steer the boat's rudder, aiming to reach a specific point within a given time. Visual cues appear on the screen to direct the player's movement. The game gets more challenging by adding obstacles that the player must avoid before reaching the final destination. The game prototype mechanics focus on improving patients' speed, endurance, eye-hand coordination, rigidity, ROM, and bilateral movement.

Flying for Gold

In this prototype, the player emulates the movement of a flying bird by raising their arms and moving them up and down toward the locations where the golden coins appear. Using the motion sensing device recognition feature, the virtual flying bird on the screen mimics the movements that the player makes. The game starts with the bird appearing on the screen, as shown in Figure S15 in Multimedia Appendix 1, followed by golden coins that appear randomly. The player must collect the coins using the movement mentioned above; the more coins they collect, the higher their score will be. The therapeutic goals are to improve coordination, shoulder movement, and endurance.

The Rowing Boat

The prototype is designed to enable the patient to use both arms simultaneously. Using the motion sensing device recognition feature, the avatar (the virtual character portraying the player on the screen) mimics the movements that the player makes. The game starts with the 2-paddled boat in a river (Figure S16 in Multimedia Appendix 1), and the player starts moving his arms (bilateral movement) to reach a specific point at a given time. Visual cues will appear on the screen to guide the player. The faster the player moves the paddles, the quicker they reach the finish line. In the advanced levels, obstacles are introduced, and the player must navigate around them to reach the destination. The prototype mechanics (targeted exercises) focus

XSL•FO
RenderX

on improving patients' speed, ROM, bilateral movement, body awareness, and reducing rigidity.

The Steam/Air Pump Game

The game begins with the appearance of a pump and a steering wheel. The objective is for the player to rotate the steering wheel or valve multiple times to release steam, ignite a fire, and cook food within a specified time limit, as illustrated in Figure S17 in Multimedia Appendix 1. Points are awarded when the player turns the steering wheel in the correct direction. Visual cues are displayed on the screen to guide the player's movements. The game mechanics are designed to enhance movement speed, reduce joint rigidity, improve ROM, and promote bilateral coordination. Figure S18 in Multimedia Appendix 1 shows an alternative scenario within the Air Pump game prototype.

Results

This section presents the opinions and statements of expert participants regarding the commercial games and game prototypes (mock-ups) that were presented during the focus group. It also includes observations made by the researchers. Therefore, several design implications and recommendations are highlighted.

The participants were excited to test the Kinect-based commercial games, exploring their features and expressing strong enthusiasm for their potential applications. However, the therapist (P4) pointed out that the Kinect Adventures and Motionsports Adrenaline games may not be entirely suitable for patients and would require significant modifications to settings such as speed and difficulty. He added that the extent of a patient's impairment increases the necessity for customizable difficulty settings. Participants (P20 and P22) who tested the adventure game observed a lack of clear initial in-game guidance, which required the facilitator to provide instructions on how to play. Therapists (P1, P2, P6, P11, and P13) agreed that the mock-up prototypes were effective due to their simplicity and focus on specific movements, whereas the commercial games lacked this focus. It is important to note that these games are typically created for healthy individuals and often provide negative feedback upon losing, which may be unsuitable for patients. However, depending on the patient's rehabilitation stage and their motor and cognitive abilities, these games can be a fun alternative to conventional rigorous training, particularly for maintaining a training regimen at home after clinical sessions. Participants (P3, P7, and P14) mentioned that in certain cases, the primary goal is to get patients moving and to create a positive experience with motion. Although correct movement is important, it is not always the primary focus. The commercial games have the potential to motivate patients, serving as a foundation for more advanced exercises. However, to be effective, they need adjustable difficulty levels, minimal distracting elements to maintain focus, simplicity, easy navigation, and clear feedback.

Participants (P4, P6, P10, and P12) emphasized that for a game-based rehabilitation system to be effective, therapists' roles should closely align with their responsibilities in traditional therapy. Once the therapist identifies the therapeutic goals, the game should enable them to assign activities and adjust the exercises to suit the individual impairment characteristics of each patient. In addition, it should enable therapists to track patients' progress and monitor the quality of their exercise movements.

A participant (P19) highlighted that rehabilitation game systems cannot operate independently (as stand-alone applications). They need to be part of a comprehensive framework that involves patients, therapists, clinicians, hospitals, and other relevant stakeholders, due to the multidisciplinary nature of the rehabilitation process. Participants (P5 and P11) believed that the use of rehabilitation games could help reduce therapists' workloads by automating exercise movements and providing monitoring and evaluation of patients' performance. Participants (P17 and P19) suggested expanding the game system to support multiple stations, enabling therapists to work with several patients simultaneously. After each session, patient data would be stored in a database for easy access, review, and performance analysis.

Participants noted that in conventional rehabilitation, therapists used various methods to deliver instructions, including verbal, gestural, and physical approaches. For example, therapists might say, "Move your hand up," and often combine verbal cues with demonstrations of the movements. When patients had difficulty understanding the instructions through verbal or gestural means, therapists provided physical assistance, guiding the patients through the movements until they understood the instructions. Therefore, the game-based system should use multiple methods to communicate instructions and explain game-related rules to patients, particularly those with cognitive impairments. This includes offering audio and visual instructions through text, images, and animations on the screen. In addition, it is important to provide feedback and multiple visual stimuli during gameplay to help direct and maintain patients' focus on the required exercises throughout the therapy sessions. In addition, a participant (P17) mentioned that for the game system to be a viable complement to in-hospital rehabilitation, it should provide clear and sufficient instructions for home-based exercises, allowing patients to participate independently. Another participant (19) suggested including video recording functionality to enable therapists to record themselves while playing the game and demonstrate the correct movements for patients to follow at home. Also, participants (P2) proposed that, following a patient's discharge from the hospital or rehabilitation center, it would be beneficial if the game-based system could record the patient while they practice at home. This would allow therapists to review and evaluate these recordings to monitor the patient's adherence and ensure the exercises are being performed correctly.

A participant (P15) proposed using an avatar to mirror patients' movements, as illustrated in the mock-up prototypes, enabling them to observe their actions in real time. To enhance instructional clarity, a second avatar could be included to demonstrate the correct exercise movements.

When asked about the importance of patients seeing themselves and their actions while exercising, as depicted in most of the presented game prototypes, most therapists appreciated the

XSL•FO
RenderX

visual representation on the screen. The psychologists (P24 and P25) noted that, for youths and children, using a fun character to demonstrate movements on the screen could be beneficial, as many prefer not to see themselves in real images due to potential negative self-perception. A psychologist (P23) added that role-playing in games is impactful because players develop emotional connections with their avatars. As avatars evolve through different stages, they play a key role in sustaining motivation. Important motivational factors include providing diverse avatar options to boost feelings of autonomy, which in turn fosters positive emotions and connections through the avatar's development and advancement (P14). In the Reaping the Oranges mock-up game, a participant (P4) recommended updating the avatar to a farmer. This change would let players take on the role of a farmer, harvesting oranges from a tree to fill a basket. He also suggested adding different game scenarios to involve a variety of physical movements. For example, instead of placing the basket beside the player, the avatar could carry a small basket on their head. In this setup, the player would need to reach up to pick oranges from the tree and place them into the basket on their head. A participant (P9) proposed incorporating different game scenarios for subsequent or advanced levels. For example, oranges could appear on a tall tree that the player's avatar cannot reach by hand. In this case, the avatar could use a small basket, either on its head or held in its hands, to catch the falling oranges before they hit the ground. The avatar would move left or right depending on where the orange will fall, considering the patient's conditions and limitations (P9). He also suggested that with each new game level, the system could change the shape and size of the tree and add animated objects to the environment to increase visual variety. Participants (P7, P11, and P13) suggested that the game system should allow them to choose from a list of available exercises. This way, each patient can be assigned different exercises based on what the therapist wants to evaluate.

A participant (P23) pointed out that consideration should be taken carefully when choosing game themes, as some themes could trigger negative reactions in patients. For example, a car racing game might be uncomfortable for someone who has been involved in a car accident. To address these concerns, it's essential to provide a variety of themes to suit different needs and sensitivities. A participant (P15) proposed that farm life could be an excellent theme for game-based physical rehabilitation, and many others (P1, P2, P11, P13, P15, P16, P20, and P22) agreed with this suggestion. This theme could appeal to a broad range of patients, including elderly individuals who may have connections to rural environments from their youth, as well as younger adults and children. The peaceful and relaxing nature of farm life can instill a sense of purpose in patients, potentially improving their commitment to the rehabilitation process. Furthermore, the theme of farm life aligns with the slow pace of seasonal changes and the growth cycles of crops and animals, reflecting the long-term nature of rehabilitation. This theme emphasizes the need for patients to commit to intensive exercise over an extended period, much like the ongoing care required in farming. For example, various farming activities can be adapted into game exercises targeting different movements, such as shearing sheep, herding flocks, milking cows, sowing seeds, planting trees, weeding, fishing, mending fences, cutting grass, picking fruit, harvesting wheat, chopping wood, feeding chickens, selling corn, irrigating, collecting honey, gathering eggs, watering plants, driving a tractor, trimming hedges, spraying pesticides, raking leaves, and scaring birds. A participant (P3) proposed enhancing the game environment by adding a farmer's house and including more exercises related to activities of daily living (ADLs). Similarly, a participant (P13) emphasized the importance of designing the game system to support the practice of ADL-related movements and tasks. To be effective, game scenarios should replicate real-life activities such as washing dishes, brushing teeth, combing hair, grasping a spoon, holding a cup, shaving, and moving objects. These scenarios should be presented through virtual representations performed by the in-game avatar, incorporating relevant joint movements—such as shoulder abduction and adduction, elbow flexion and extension, and forearm pronation and supination—to interact meaningfully with virtual objects.

A participant (P9) suggested providing a pool of games, each targeting different exercise movements. Another participant (P21) noted that many of the mock-up game prototypes, including The Steam/Air Pump Game, The Rowing Boat, Boat Driver, Colored Boxes, Matching Shapes, Reaping the Oranges, Shuffling Shelves, Organizing Eggs, Rabbit Hunting, and Fallen Apples, could be adapted to fit within a farm life theme under a unified game system. This would provide a variety of game exercises centered around the farm life concept.

Participants (P20 and P23) revealed that the multiplayer aspect of the Reaping the Oranges game, whether through competition, collaboration, or connecting with other patients face-to-face or online, could enhance engagement and enjoyment for many patients. However, it is important to recognize that such interactions may cause stress for some individuals. Therefore, while the game has the potential to enhance the overall experience, it is crucial to consider the diverse preferences and needs of different patients.

A participant (P10) suggested that, given recent technological advancements, it is possible to enable immersive experiences through head-mounted display devices, which are now lightweight and affordable. These devices provide immersive 3D game experiences that can help sustain patient motivation and adherence.

Therapists and other practitioners have been asked about how to transform a physical rehabilitation procedure into a game. It is found that various devices are employed to facilitate therapeutic processes. For instance, during physiotherapy sessions, patients employ various auxiliary devices to facilitate task execution, such as the Shoulder Pulley T Type and the Shoulder Wheel, as depicted in Figure S19 in Multimedia Appendix 1. The shoulder wheel is employed to regain the range of motion in the shoulder joints. Although these devices support the execution of therapeutic movements, they often lack motivational aspects to engage patients. Participants (P3, P10, and P20) noted that integrating gamification with these instrumented devices could fill this gap by turning rehabilitation procedures into interactive and motivating experiences. This

XSL•FO
RenderX

approach represents a valuable opportunity for enhancing therapeutic outcomes and deserves further investigation.

Participants (P22, P24, and P14) argued that adapting existing devices used in physiotherapy sessions and transforming them into game controllers could shift patients' focus from the monotonous repetitions of their exercises to engaging with the game itself. Consequently, actions performed with physical objects in the real world are mapped and reflected in the virtual environment.

The researchers suggested the following game scenarios to demonstrate how these devices can be used: (1) by placing a screen in front of the patients, the Shoulder Wheel can be used to explore a virtual city or world, allowing them to engage with the game and experience tailored stimuli and positive reinforcement. Actions performed by the player using physical objects in the real world are mapped and reflected in the virtual environment. (2) The Shoulder Wheel can also be used for a variety of therapeutic exercises. When operated with both hands, its movements can be integrated into a racing car simulator. Alternatively, when used with one hand, it can be adapted into a firefighter pump game, allowing patients to simulate controlling a water hose to help extinguish a fire. (3) The Shoulder Pulley T Type can be used to simulate the flapping of a bird or the soaring of an aircraft through the skies.

Games based on this setup can be designed to balance patient enjoyment (entertainment value) with an emphasis on the quality of exercise movements (therapeutic value), thereby enhancing both engagement and the effectiveness of the rehabilitation process.

Most participants emphasized that a game-based physical rehabilitation system should offer flexible configuration. It should accommodate diverse cognitive abilities and gameplay preferences by incorporating essential features such as therapy-specific physical movements, adjustable game mechanics, and customizable visual aesthetics.

A participant (P2), who has experience using games in rehabilitation, expressed dissatisfaction with the preprogrammed difficulty levels in games such as adventure games. The participant emphasized the need for more flexible control over game parameters, allowing therapists to adjust the difficulty to suit the individual needs and progress of each patient.

A participant (P16) suggested that developers could take advantage of existing commercial games to create diverse game scenarios. He added that a game like Kinect Adventures: 20,000 Leaks could be modified by simplifying its design, enabling therapists to control aspects such as difficulty level, speed, and range of motion, and to tailor patients' interactions with virtual game objects.

A participant (P14) argued that the game prototypes presented in the Intervention Using Low-Fidelity Paper Mock-Ups section are designed to permit only predefined exercise movements to interact with or play a specific game. For example, in the Blowing the Bubbles game, patients use reaching movements to blow the bubbles. In the Boat Driver game, patients perform circular movements to control the boat. A game system that enables therapists to customize and align different physical movements with game actions (such as using circular motions to maneuver a jet fighter) can enhance the gaming experience for patients and support their rehabilitation goals more effectively.

Most participants believed that a game system with reconfigurable visual elements, tailored to patients' preferences and needs, can enhance their engagement and participation in game-based therapy.

Participants (P4, P7, P22, and P24) noted that patients may react differently to the visual design of rehabilitation games. Individuals with cognitive impairments often struggle to process highly realistic graphics, which can lead to disengagement. In such cases, simplified 2D or cartoon-like visuals may improve comprehension and promote participation in therapy. In addition, older patients or those with negative attitudes toward gaming may view these games as childish and lacking therapeutic value. To address this concern, a participant (P15) suggested that game designs could be adjusted to reflect real-world therapeutic activities, thereby reinforcing their role as serious tools for rehabilitation. Therefore, a game-based system should be able to adjust its visual settings and reconfigure its appearance to match a patient's preferences and needs, as this is crucial for enhancing engagement and participation in game-based therapy. This highlights the need for future research to develop tools that enable therapists to adjust and tailor games to individual patients' preferences and needs.

On the other hand, the controllers of the rehabilitation game system should be designed to enable patients to practice various exercise movements while intentionally suppressing their compensatory behaviors. A participant (P11) suggested that the controller could be configured by the therapist to recognize only the horizontal movements of the patient's hand, or to consider trunk movements as input, thereby reducing trunk-related compensatory actions. Likewise, for stroke patients who unintentionally involve their shoulder or elbow, the game system can allow the therapist to address this by using the compensating joint as an input. Furthermore, the system should be adjustable and flexible, allowing the therapist to choose which hand or side of the patient (left or right) will be used for game interaction.

A participant (P5) suggested implementing a dedicated user interface for therapists to enable efficient control of the game system. This interface would allow therapists to configure game settings. For instance, if a patient begins to show compensatory behaviors during gameplay, the therapist can adjust the game settings to reduce the range of motion or slow down the game pace. This allows the patient to perform movements without severe compensatory behaviors. When patients experience fatigue or pain, therapists can make the games less challenging or pause them. In addition, therapists can have full control over the games to maintain a balance between therapeutic and engagement values.

Another participant (P6) argued that in conventional rehabilitation sessions, therapists must physically interact with patients to correct inappropriate movements. This often requires therapists to stand or sit next to or behind their patients, meaning they need to continuously reposition themselves throughout the

XSL•FO
RenderX

therapy sessions. Thus, a dedicated and portable user interface is essential, allowing therapists to control game-based therapy more efficiently using handheld devices like tablets or mobile phones. This interface would facilitate better interaction with the games and improve the moderation of therapy sessions.

Consequently, based on the earlier discussion, it is evident that there is a need for game systems with high configurability. Such systems would allow therapists to customize therapy programs to better meet individual patient needs. These systems should offer the flexibility required for therapists to tailor therapy programs to each patient's specific needs. However, the increased complexity of game system operations places additional demands on therapists and requires comprehensive training to effectively manage and moderate game-assisted therapy sessions.

Therapists and other practitioners emphasized the importance of proper training to effectively utilize rehabilitation games as a therapeutic tool. For instance, a participant (P1) remarked, "I am motivated to use games with my patients, but I don't know what to do." Therapists explained that improper use of rehabilitation games can result in patient pain or injury, making it challenging for therapists who are new to or lack experience with game-based therapy. Participants (P22 and P25) noted that using rehabilitation games without previous training can make it difficult to determine how to tailor the therapy to meet patients' needs effectively. With proper training, therapists would gain a clearer understanding of their roles, leading to more efficient use of the games and an improvement in the quality of game-based therapy sessions. Therefore, thorough training on the game system is crucial for therapists to let them grasp the system's details and learn the most effective ways to use it. Consequently, it is essential to thoroughly assess and prioritize the usability of the game system for therapists to ensure it is both effective and user-friendly. This underscores the necessity for future research to examine the usability and effectiveness of interactive mechanisms that allow therapists to configure and customize game-based interventions.

On the other hand, therapists revealed that the data collected from the game-based intervention can provide various insights. While the scores at the end of a game session are commonly used to measure performance, additional metrics can be gathered to offer insights into patients' motor and cognitive abilities. These include the following: (1) score: higher scores may indicate better performance in response to challenging tasks; (2) time: longer session durations suggest a greater number of repetitions of the prescribed exercises; (3) level: advancing to higher game levels is associated with increased challenges; (4) distance: Assessing the distance covered by the affected limbs helps determine the level of muscle engagement and effort exerted during the exercise; (5) changes in motion direction: a higher frequency of changes in movement direction might indicate inaccuracies in execution or that the movements were performed without a clear purpose; (6) initial movement range/amplitude: this metric is crucial for evaluating patients' progress and understanding their development over time; and (7) number of sessions: tracking the number of sessions played and any interruptions can help therapists determine whether patients felt uncomfortable with specific games or if they were disengaged and did not complete the prescribed exercises. In addition, one therapist mentioned that electromyography (EMG) sensors could be used to monitor muscle activity during exercises. By recording EMG readings, therapists can assess movement patterns, measure muscle fatigue, identify disorders, and evaluate the effectiveness of exercise interventions.

Movement correction measure: A key factor in the effectiveness of game-based rehabilitation systems is the accurate assessment of patient movements during gameplay sessions. However, according to a participant (P13), the measures taken to prevent incorrect execution of exercises in motion-sensing technology–based games are limited, as patients at risk of injury are typically not permitted to use motion-tracking systems. It remains the therapist's responsibility to evaluate and manage these situations. Another participant (P5) highlighted that some of the major benefits of using this technology include engaging patients, making exercise enjoyable, providing opportunities for social interaction, and offering feedback on exercises. The experts were less worried about the games enforcing or ensuring correct movement and addressing compensatory actions, as they would personally guide and decide when and how patients should use motion games.

According to the discussed metrics, these metrics are expected to provide valuable insights for therapists to evaluate a patient's recovery progress. For instance, low scores and extended playtimes may indicate that patients are struggling to complete the required tasks, which can result in frustration and a negative perception of the game. In addition, covering longer distances with the affected limb and making fewer changes in motion direction suggests that patients are performing movements with greater amplitude, indicating they are in a more advanced stage of recovery. Conversely, if patients cover short distances with frequent changes in motion direction, it indicates that they either struggle to perform wide gestures or exhibit inaccuracies in their movements. Thus, while the patient enjoys the game, these parameters can be collected and stored for later analysis. This allows therapists to monitor and assess the patient's progress and development over time, providing insights into the quality and effectiveness of the game exercises performed. Another therapist suggested developing an online game-based system accessible to both patients and therapists. Therapists would log in to the system to view their patients, and once a specific patient is selected, the therapist could access their records, review collected data, and customize each game's settings as needed to better suit the patient's needs.

Discussion

Principal Findings

The observations, feedback, and suggestions from the focus group were analyzed, leading to the development of several guidelines to inform the design and development of an effective game-based physical rehabilitation system. To increase compliance and ensure patients perform their exercises correctly, the game must maintain patient engagement and provide support during exercises. This can be achieved through appropriate feedback. Instant and clear feedback is crucial for providing a meaningful gameplay experience. It helps patients understand

XSL•FO
RenderX

the outcomes of their actions, reflects their performance in exercises, and offers various in-game achievements. Feedback should include visual, audible, and haptic elements to enhance the user's interpretation of the gameplay and maintain their engagement. In addition, a game-based system needs to facilitate the practice of movements and tasks related to ADLs.

The social aspect of the game system is crucial, enabling patients to play with family members, friends, or other patients. Highlighting this feature enhances the overall experience. In addition, incorporating a tracking mechanism for the quality of the patient's movements in the game system is advantageous. This will effectively support and facilitate the rehabilitation process, encouraging patients to perform the exercises accurately. In addition, a potential game-based system should be capable of tracking patients' progress. This feature could benefit patients by providing positive reinforcement and encouraging compliance and also assist physiotherapists by enabling them to monitor and assess patient development. In addition, simple art styles and graphics, as seen in the mock-up prototypes, can motivate players and make the games more approachable, a feature that participants seemed to appreciate. This cartoon-like, simplistic art style has benefits such as lower processing power requirements and encouraging a more playful state of mind compared to realistic graphics. However, different demographics might respond differently to this style, as many people today are accustomed to games with nearly photo-realistic graphics, such as Call of Duty. People generally tend to prefer familiar styles. On the other hand, since rehabilitation programs typically involve multiple exercises, the game system should be designed to incorporate them. Therefore, the game system should be expandable to include a variety of exercises, supporting the entire rehabilitation process.

Furthermore, the game system should offer significant flexibility and adaptability to suit each patient's specific degree of injury. It is essential for a potential game-based rehabilitation system to allow therapists to customize the game exercises. Since individuals have different needs based on their injury and stage of rehabilitation, customization ensures that the game can better accommodate these unique patient requirements and preferences. Given that the rehabilitation process often extends over a long period, the focus of the training changes throughout this time. Therefore, the game system should include mechanisms for adapting game interventions to accommodate the different stages and focuses of rehabilitation. On the other hand, ensure that patients can concentrate on their exercises rather than dealing with technical details. The findings highlight the need for a rehabilitation game to be simple to install and start. It should be easy to begin without lengthy or complex setup processes. For home use, the game system should prioritize easy installation, fast start-up, and intuitive usability. Table 2 presents a summary and categorization of the design guidelines for game-based physical rehabilitation systems.

XSL•FO
RenderX

Table . Design guidelines for the game-based physical rehabilitation system.

Category	Guidelines
Gameplay design and usability	• Support simplified gameplay mechanics to enhance usability and accommodate patients with varying cognitive and physical abilities. • Integrate multimodal instructional methods, including audio, visual, text, and animation, to support understanding of game mechanics and therapeutic tasks. • Include accessibility features such as simplified controls, multiple visual modes, easy-to-navigate menus, and customizable interface styles to support diverse user needs, allowing the system to cater to different user preferences, cognitive loads, and therapeutic requirements. • Support game-based exercises tailored to the cognitive and physical capacities of patients. • Provide instant feedback and multiple visual stimuli during gameplay to guide and sustain patients' focus on the required exercises. • Use positive reinforcement and reward mechanisms to motivate patients and enhance engagement during therapy. • Support home-based exercise modules that are self-explanatory and include clear, structured instructions to enable independent participation. • Embed clear and intuitive tutorials, supported by visual and auditory cues, as well as explicit in-game guidance, to encourage independent use and reduce the need for external support. • Provide optional multiplayer modes to promote social interaction and engagement when appropriate for the therapeutic context.
Game themes and scenarios	• Provide a variety of themes and scenarios to suit different needs and sensitivities. • Avoid game themes that may trigger distress. Themes should be carefully selected to prevent negative reactions (eg, racing games for car accident survivors) and instead incorporate neutral, calming settings (eg, farm life with seasonal progression) that align with rehabilitation goals. • Incorporate activities of daily living (ADLs), such as fruit picking, brushing teeth, and combing hair, to enhance task relevance and promote functional recovery. • Use stylized, age-appropriate avatars rather than live camera images of patients to enhance comfort and engagement during gameplay. • Offer diverse avatar choices and thematic avatars matching game contexts. • Incorporate avatar customization and progression systems and offer upgrade options to enhance autonomy and sustain patient motivation throughout therapy. • Implement a dual-avatar system, where one avatar mirrors the patient's movements and another, representing the instructor or therapist, demonstrates the correct form to facilitate self-correction.

XSL•FO
RenderX

Category	Guidelines
Therapeutic focus and therapist control	• Implement therapist interfaces that allow dynamic adjustment of exercise parameters such as speed, difficulty, and range of motion during rehabilitation sessions. • Enable therapists to adjust difficulty, customize visuals, and adapt game mechanics to accommodate patient limitations and cognitive impairments. • Allow therapists to target specific motor functions by selecting exercises from a game-based exercise list and assigning them based on individual patient impairments and needs. • Provide a portable, handheld (tablet or phone) interface for therapists to seamlessly control sessions, including real-time form alerts, emergency stop, pause, and override functions to ensure patient safety, correct movement, and efficient session management. • Include tools that allow therapists to translate and map physical rehabilitation movements into game mechanics, ensuring clinical relevance and effectiveness. • Enable the therapist to define movement constraints: the system should be adjustable and flexible, allowing the therapist to choose which hand or side of the patient (left or right) will be used for game interaction. • Provide video recording functionality for therapists to demonstrate exercises remotely. • Enable multiuser support, centralized patient databases, and access for multiple stakeholders or practitioners.
Movement quality, monitoring, and data management	• Incorporate features to track performance metrics, including scores, time, directional changes, movement precision, range of motion, limb engagement, muscle activation, and compensatory movements. • Implement AI-assisted compensatory movement detection with real-time alerts to prevent patients from persisting through pain or incorrect form. • Store data for longitudinal analysis to inform therapy adjustments. • Integrate motion tracking with electromyography (EMG) data to enable comprehensive monitoring of muscle activity during exercises. • Include postsession analytics and progress monitoring dashboards. • Develop comprehensive dashboards for therapists that display and visualize patient progress, compensation behaviors, adherence levels, and movement quality trends over time.
Technology and device integration	• Support the integration of virtual reality or augmented reality and diverse interaction technologies to enhance the effectiveness and engagement of rehabilitation experiences. • Support the integration of EMG biofeedback to track muscle activation, fatigue thresholds, intervention efficacy, and exercise effectiveness. • Gamify existing rehabilitation devices, such as shoulder wheels and pulleys, by transforming them into game controllers that map real-world movements to virtual environments. This allows patients to navigate and interact with the game, thereby enhancing the effectiveness of rehabilitation. • Implement an online platform that supports secure data access and remote configuration.

Limitations

Therapists and health care professionals were actively involved in this phase of the study due to their critical role in implementing effective game-based rehabilitation. Although the inclusion of brain injury survivors (including those with traumatic injuries and stroke) is equally significant, it was beyond the scope of this paper. Their perspectives and participation will be incorporated into future phases of the research and addressed in subsequent publications.

Conclusions

Incorporating serious games and advances in information and communication technologies into physical rehabilitation can enhance the rehabilitation process, offer benefits beyond those of conventional methods, and improve both the effectiveness and efficiency of rehabilitation. The aim is to (1) enhance the effectiveness of game-based therapy for patients who struggle to engage with or benefit from it, and (2) assist therapists by offering a game-based rehabilitation system tailored to their needs. To investigate the factors, determinants, requirements, and actual needs for potential game-based rehabilitation, a focus

XSL•FO
RenderX

group with therapists and other relevant health care professionals is organized. This decision aims to evaluate the ideas with real rehabilitation therapists and practitioners. To effectively discuss the ideas with therapists and other related professionals before starting development, available commercial games are tested, and various paper game mock-ups are evaluated. This approach facilitated productive discussions, validated the direction, and is expected to optimize our development timeline. The focus group discussions yielded several guidelines that will serve as the foundation for the next phase of our research. This phase will involve the design, development, and implementation of a game-based physical rehabilitation system aligned with these guidelines.

Acknowledgments

The authors gratefully acknowledge the focus group participants for their time and for sharing their valuable perspectives. This research was funded by Universiti Teknologi Brunei (UTB). UTB/GSR/2/2023 (19). We thank UTB for the financial support of this research.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Illustration of gameplay.
[PDF File, 1489 KB - humanfactors_v12i1e67336_app1.pdf]

References

1. Riaz S, Shakil Ur Rehman S, Hassan D, Hafeez SJS. Gamified exercise with Kinect: can Kinect-based virtual reality training improve physical performance and quality of life in postmenopausal women with osteopenia? A randomized controlled trial. Sensors (Basel) 2024 Jun 1;24(11):3577. [doi: 10.3390/s24113577] [Medline: 38894367]
2. Bhise S, Rathi M, Dabadghav R, Atre J. Use of virtual reality in physical rehabilitation: A narrative review. Curr Med Res Pract 2024;14(3):122-127. [doi: 10.4103/cmrp.cmrp_225_23]
3. Eckert M, Domingo Soria B, Terroso Gil N. Finding effective adjustment levels for upper limb exergames: focus group study with children with physical disabilities. JMIR Serious Games 2023 Jan 13;11:e36110. [doi: 10.2196/36110] [Medline: 36637882]
4. Brooks AL. Computer graphics, video games, and gamification impacting (re) habilitation, healthcare, and inclusive well-being. In: Encyclopedia of Computer Graphics and Game: Springer; 2024:426-438.
5. Chan-Víquez D, Fernández-Huertas H, Chacón-Vargas F, et al. Family expectations and demand for home-based videogaming therapy in children with cerebral palsy in Costa Rica: a mixed methods study. Disabil Rehabil 2025 Feb;47(4):1012-1023. [doi: 10.1080/09638288.2024.2362952] [Medline: 38946018]
6. Tăut D, Pintea S, Roovers J, Mañanas MA, Băban A. Play seriously: Effectiveness of serious games and their features in motor rehabilitation. A meta-analysis. NRE 2017;41(1):105-118. [doi: 10.3233/NRE-171462]
7. Alfieri FM, da Silva Dias C, de Oliveira NC, Battistella LR. Gamification in musculoskeletal rehabilitation. Curr Rev Musculoskelet Med 2022 Dec;15(6):629-636. [doi: 10.1007/s12178-022-09797-w] [Medline: 36301514]
8. Wang L, Chen JL, Wong AMK, Liang KC, Tseng KC. Game-based virtual reality system for upper limb rehabilitation after stroke in a clinical environment: systematic review and meta-analysis. Games Health J 2022 Oct;11(5):277-297. [doi: 10.1089/g4h.2022.0086] [Medline: 36252097]
9. Tuah NM, Ahmedy F, Gani A, Yong L. A survey on gamification for health rehabilitation care: applications, opportunities, and open challenges. Information 2021;12(2):91. [doi: 10.3390/info12020091]
10. Li Y, Liu Y, Leung AY, Montayre J. Technology-assisted motor-cognitive training among older adults: rapid systematic review of randomized controlled trials. JMIR Serious Games 2025 Jun 3;13:e67250. [doi: 10.2196/67250] [Medline: 40460432]
11. Dou J, Wang J, Gao X, et al. Effectiveness of telemedicine interventions on motor and nonmotor outcomes in Parkinson disease: systematic review and network meta-analysis. J Med Internet Res 2025 Jun 3;27:e71169. [doi: 10.2196/71169] [Medline: 40460428]
12. Mayer TA, Koska D, Harsch AK, Maiwald C. Effectiveness of self-training with a web-based digital health application versus physiotherapy in the treatment of disorders of the patella: randomized controlled trial. J Med Internet Res 2025 May 5;27:e66463. [doi: 10.2196/66463] [Medline: 40323642]
13. Sheehy L, Taillon-Hobson A, Sveistrup H, et al. Home-based nonimmersive virtual reality training after discharge from inpatient or outpatient stroke rehabilitation: parallel feasibility randomized controlled trial. JMIR Rehabil Assist Technol 2025 Mar 28;12:e64729. [doi: 10.2196/64729] [Medline: 40153779]

XSL•FO
RenderX

14. Krishnan S, Mandala MA, Wolf SL, Howard A, Kesar T. Perceptions of stroke survivors regarding factors affecting adoption of technology and exergames for rehabilitation. PM R 2023 Nov;15(11):1403-1410. [doi: 10.1002/pmrj.12963]
15. Cruz L, Santos Mendes FA, Neri SGR, Carregaro RL. Principles and practical uses of virtual reality games as a physical therapy strategy. In: Everyday Technologies in Healthcare: CRC Press; 2019:235-257.
16. Daoud MI, Alhusseini A, Ali MZ, Alazrai RJS. A game-based rehabilitation system for upper-limb cerebral palsy: a feasibility study. Sensors (Basel) 2020 Apr 24;20(8):2416. [doi: 10.3390/s20082416] [Medline: 32344557]
17. Jayasree-Krishnan V, Ghosh S, Palumbo A, Kapila V, Raghavan P. Developing a framework for designing and deploying technology-assisted rehabilitation after stroke: a qualitative study. Am J Phys Med Rehabil 2021 Aug 1;100(8):774-779. [doi: 10.1097/PHM.0000000000001634] [Medline: 33141773]
18. Alarcón-Aldana AC, Callejas-Cuervo M, Bo APL. Upper limb physical rehabilitation using serious videogames and motion capture systems: a systematic review. Sensors (Basel) 2020 Oct 22;20(21):5989. [doi: 10.3390/s20215989] [Medline: 33105845]
19. Umar N, Yuyun Y, Gani H. Generating game immersion features for immersive game selection. ILKOM Jurnal Ilmiah ;14(3):264-274. [doi: 10.33096/ilkom.v14i3.1224.264-274]
20. Takei K, Morita S, Watanabe Y. Acceptability of physical therapy combined with nintendo ring fit adventure exergame for geriatric hospitalized patients. Games Health J 2024 Feb;13(1):33-39. [doi: 10.1089/g4h.2023.0009] [Medline: 37566481]
21. Peláez-Vélez FJ, Eckert M, Gacto-Sánchez M, Martínez-Carrasco Á. Use of virtual reality and videogames in the physiotherapy treatment of stroke patients: a pilot randomized controlled trial. Int J Environ Res Public Health 2023 Mar 8;20(6):4747. [doi: 10.3390/ijerph20064747] [Medline: 36981652]
22. Takei K, Morita S, Watanabe Y, Suganami M, Inao MJD, Technology RA. Safety, feasibility, and acceptability of physiotherapy combined with strength training using active video games for older patients with musculoskeletal conditions. Disabil Rehabil Assist Technol 2024 Apr;19(3):641-647. [doi: 10.1080/17483107.2022.2115562] [Medline: 36001528]
23. Wittmann F, Lambercy O, Gassert RJS. Magnetometer-based drift correction during rest inIMU arm motion tracking. Sensors (Basel) 2019 Mar 15;19(6):1312. [doi: 10.3390/s19061312] [Medline: 30884745]
24. Wittmann F, Held JP, Lambercy O, et al. Self-directed arm therapy at home after stroke with a sensor-based virtual reality training system. J Neuroeng Rehabil 2016 Aug 11;13(1):75. [doi: 10.1186/s12984-016-0182-1] [Medline: 27515583]
25. Goršič M, Cikajlo I, Novak D. Competitive and cooperative arm rehabilitation games played by a patient and unimpaired person: effects on motivation and exercise intensity. J Neuroeng Rehabil 2017 Mar 23;14(1):23. [doi: 10.1186/s12984-017-0231-4] [Medline: 28330504]
26. Ma M, Proffitt R, Skubic M. Validation of a Kinect V2 based rehabilitation game. PLoS ONE 2018;13(8):e0202338. [doi: 10.1371/journal.pone.0202338] [Medline: 30142631]
27. Junior IPG, Garcia FM, Silva K, Ferreira FX, Silva JRD, Chagas WMD. A review of the use of new technologies in physical therapy rehabilitation: possibilities and challenges with Xbox and kinect. IJAUDIT 2018;4(1):1. [doi: 10.1504/IJAUDIT.2018.100190]
28. Burin-Chu S, Baillet H, Leconte P, Lejeune L, Thouvarecq R, Benguigui N. Effectiveness of virtual reality interventions of the upper limb in children and young adults with cerebral palsy: A systematic review with meta-analysis. Clin Rehabil 2024 Jan;38(1):15-33. [doi: 10.1177/02692155231187858] [Medline: 37499213]
29. Tatla SK, Shirzad N, Lohse KR, et al. Therapists' perceptions of social media and video game technologies in upper limb rehabilitation. JMIR Serious Games 2015 Mar 10;3(1):e2. [doi: 10.2196/games.3401] [Medline: 25759148]
30. Hamilton C, McCluskey A, Hassett L, Killington M, Lovarini M. Patient and therapist experiences of using affordable feedback-based technology in rehabilitation: a qualitative study nested in a randomized controlled trial. Clin Rehabil 2018 Sep;32(9):1258-1270. [doi: 10.1177/0269215518771820] [Medline: 29696990]
31. Elaklouk AM, Mat Zin NA, Shapii A. Investigating therapists' intention to use serious games for acquired brain injury cognitive rehabilitation. J King Saud Univ Comput Inf Sci 2015 Apr;27(2):160-169. [doi: 10.1016/j.jksuci.2014.03.019]
32. Huang Z, McCoy D, Cooper R, Crytzer TM, Chi Y, Ding D. Wearable-enhanced mHealth intervention to promote physical activity in manual wheelchair users: single-group pre-post feasibility study. JMIR Rehabil Assist Technol 2025 Jun 5;12:e70063. [doi: 10.2196/70063] [Medline: 40472238]
33. Olimb Hillkirk A, Skavberg Roaldsen K, Johnsen HM. Physiotherapists' user acceptance of a lower limb robotic exoskeleton in specialized rehabilitation: qualitative exploratory study. JMIR Rehabil Assist Technol 2025 Apr 16;12:e68233. [doi: 10.2196/68233] [Medline: 40238235]
34. Jung HT, Park T, MAhyar N, Park S, Ryu T, Kim Y, et al. Rehabilitation games in real-world clinical settings: practices, challenges, and opportunities. ACM Trans Comput Hum Interact 2020;27(6):1-43. [doi: 10.1145/3418197]
35. Wu KL, Wang YH, Hsu YC, Shu YC, Chu CH, Lin CA. Developing a motion sensor-based game to support frozen shoulder rehabilitation in older adults through a participatory design approach. Games Health J 2024 Oct;13(5):335-345. [doi: 10.1089/g4h.2023.0097] [Medline: 38934148]
36. Mat Zin NA, Elaklouk AM. Design science paradigm in the development of serious game for cognitive rehabilitation. Int J Adv Sci Eng Inf Technol 2017;7(1):118. [doi: 10.18517/ijaseit.7.1.1171]

XSL•FO
RenderX

37. Elaklouk AMS, Zuarni Ramli R, Edris I, Padilla-Valdez N, Jumaily AA. A framework for tele-rehabilitation gaming system. Presented at: 2023 6th International Conference on Applied Computational Intelligence in Information Systems (ACIIS); Oct 23-25, 2023; Bandar Seri Begawan, Brunei Darussalam. [doi: 10.1109/ACIIS59385.2023.10367329]
38. Burke J, Morrow P, McNeill M, McDonough S, Charles D, editors. Vision based games for upper-limb stroke rehabilitation. Presented at: 2008 International Machine Vision and Image Processing Conference; Sep 3-5, 2008; Portrush, Ireland. [doi: 10.1109/IMVIP.2008.16]
39. Nixon ME, Howard AM, editors. Applying gaming principles to virtual environments for upper extremity therapy games. Presented at: 2013 IEEE International Conference on Systems, Man, and Cybernetics; Oct 13-16, 2013; Manchester, UK. [doi: 10.1109/SMC.2013.585]
40. Powell V, Powell W. Therapy-led design of home-based virtual rehabilitation. In: Powell V, Powell W, editors. Presented at: 2015 IEEE 1st Workshop on Everyday Virtual Reality (WEVR); Mar 23, 2015; Arles, Camargue, Provence, France. [doi: 10.1109/WEVR.2015.7151688]
41. Pirovano M, Lanzi PL, Mainetti R, Borghese NA, editors. The design of a comprehensive game engine for rehabilitation. Presented at: 2013 IEEE International Games Innovation Conference (IGIC); Sep 23-25, 2013; Vancouver, BC, Canada. [doi: 10.1109/IGIC.2013.6659160]
42. Abd Latif MH, Md.Yusof H, Sidek SN, Shikhraji MS, Safie MH. A gaming-based system for stroke patients physical rehabilitation. In: Abd Latif MH, Md.Yusof H, Sidek SN, Shikhraji MS, Safie MH, editors. Presented at: 2014 IEEE Conference on Biomedical Engineering and Sciences (IECBES); Kuala Lumpur, Malaysia. [doi: 10.1109/IECBES.2014.7047594]
43. d'Ornellas MC, Cargnin DJ, Prado ALC, editors. Thoroughly approach to upper limb rehabilitation using serious games for intensive group physical therapy or individual biofeedback training. Presented at: 2014 Brazilian Symposium on Computer Games and Digital Entertainment; 2014: IEEE; Nov 12-14, 2014; Porto Alegre, Brazil. [doi: 10.1109/SBGAMES.2014.22]
44. Muijzer F. Development of an Automated Exercise Detection and Evaluation System Using the Kinect Depth Camera: University of Twente; 2014.
45. Jaume-i-Capó A, Martínez-Bueso P, Moyà-Alcover B, Varona J. Interactive rehabilitation system for improvement of balance therapies in people with cerebral palsy. IEEE Trans Neural Syst Rehabil Eng 2014 Mar;22(2):419-427. [doi: 10.1109/TNSRE.2013.2279155] [Medline: 24122561]
46. Cary F, Postolache O, PSJIJoSS G, Systems I. Kinect based system and serious game motivating approach for physiotherapy assessment and remote session monitoring. Int J Smart Sens Intell Syst 2020;7(5):1-6. [doi: 10.21307/ijssis-2019-131]
47. Madeira RN, Costa L, Postolache O. PhysioMate - pervasive physical rehabilitation based on NUI and gamification. In: Postolache O, editor. Presented at: 2014 International Conference and Exposition on Electrical and Power Engineering (EPE); Oct 16-18, 2014; Iasi, Romania. [doi: 10.1109/ICEPE.2014.6969982]
48. Macedo ACM. Serious game for motion disorders rehabilitation of parkinson's disease patients. : Instituto Superior Técnico de Lisboa; 2014.
49. Paraskevopoulos I. The development and applications of serious games in the public services: defence and health. : Brunel University London; 2014.
50. Kallmann M, Camporesi C, Han J, editors. VR-assisted physical rehabilitation: adapting to the needs of therapists and patients. 2013 Presented at: Virtual Realities: International Dagstuhl Seminar; Jun 9-14, 2013; Dagstuhl Castle, Germany. [doi: 10.1007/978-3-319-17043-5_9]
51. Baranyi R, Reisecker F, Lederer N, Gobber M, Grechenig T. WristDroid - a serious game to support and motivate patients throughout their wrist rehabilitation. In: Reisecker F, Lederer N, Gobber M, Grechenig T, editors. Presented at: 2014 IEEE Conference on Biomedical Engineering and Sciences (IECBES); Dec 8-10, 2014; Kuala Lumpur, Malaysia. [doi: 10.1109/IECBES.2014.7047617]
52. Martins T, Carvalho V, Soares F. A serious game for rehabilitation of neurological disabilities: premilinary study. In: Martins T, Carvalho V, Soares F, editors. Presented at: 2015 IEEE 4th Portuguese Meeting on Bioengineering (ENBENG); Feb 26-28, 2015; Porto, Portugal. [doi: 10.1109/ENBENG.2015.7088843]
53. Baldominos A, Saez Y, Pozo CD. An approach to physical rehabilitation using state-of-the-art virtual reality and motion tracking technologies. Procedia Comput Sci 2015;64:10-16. [doi: 10.1016/j.procs.2015.08.457]
54. Jaume-I-Capó A, Samcovic A. Interactive multimedia system using serious game for users with motor disabilities. Telfor 2015;7(2):97-102. [doi: 10.5937/telfor1502097J]
55. Pachoulakis I, Papadopoulos N, Spanaki C. Parkinson's disease patient rehabilitation using gaming platforms: lessons learnt. IJBES 2015 Oct 31;2(4):1-12 [FREE Full text] [doi: 10.5121/ijbes.2015.2401]
56. Pedraza-Hueso M, Martín-Calzón S, Díaz-Pernas FJ, Martínez-Zarzuela M. Rehabilitation using Kinect-based games and virtual reality. Procedia Comput Sci 2015;75:161-168. [doi: 10.1016/j.procs.2015.12.233]
57. Pirovano M. The design of exergaming systems for autonomous rehabilitation [article in italian]. Politecnico. 2015. URL: https://www.politesi.polimi.it/handle/10589/108826 [accessed 2025-09-20]
58. Paavola JM, Oliver KE, Ustinova KI. Use of X-box Kinect gaming console for rehabilitation of an individual with traumatic brain injury: a case report. J Nov Physiother 2013;03(1):1-6. [doi: 10.4172/2165-7025.1000129]

XSL•FO
RenderX

59. Chanpimol S, Seamon B, Hernandez H, Harris-Love M, Blackman MR. Using Xbox kinect motion capture technology to improve clinical rehabilitation outcomes for balance and cardiovascular health in an individual with chronic TBI. Arch Physiother 2017;7:1-11. [doi: 10.1186/s40945-017-0033-9] [Medline: 28824816]
60. Rizzo A“, Kim GJ. A SWOT analysis of the field of virtual reality rehabilitation and therapy. Presence: Teleoperators & Virtual Environments 2005 Apr;14(2):119-146. [doi: 10.1162/1054746053967094]
61. Burke JW, McNeill M, Charles D, Morrow P, Crosbie J, McDonough S, editors. Serious games for upper limb rehabilitation following stroke. Presented at: 2009 Conference in Games and Virtual Worlds for Serious Applications; Mar 23-24, 2009; Coventry, United Kingdom.

Abbreviations

ADL: activity of daily living
EMG: electromyography
ROM: range of motion

Edited by A Kushniruk; submitted 08.10.24; peer-reviewed by E Vallefuoco, F Noveletto, NA Mat Zin; revised version received 18.06.25; accepted 10.07.25; published 26.09.25.

Please cite as:
S Elaklouk AM, Ramli RZ, M Alakklouk SM, Rahim N
Design Guidelines for a Game-Based Physical Rehabilitation System: Focus Group Study
JMIR Hum Factors 2025;12:e67336
URL: https://humanfactors.jmir.org/2025/1/e67336
doi:10.2196/67336

© Ahmad M S Elaklouk, Ratna Zuarni Ramli, Samah M M Alakklouk, Norsafinar Rahim. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 26.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Clinician Needs and Requirements for a Decision Aid Navigator: Qualitative Study

Brad Morse[1], MA, PhD; Carrie Reale[2], RN-BC, MSN; An T Nguyen[3], OTD, OTR/L, BCG; Erin Latella[1], BA; Hannah Bauguess[1], MA; Shilo Anders[2], PHD; Pamela Roberts[3], PHD; Spencer L SooHoo[3], MS, PHD; Robert El-Kareh[4], MPH, MD; Andrey Soares[1], PHD; Lisa Schilling[1], MSPH, MD

[1]Division of General Internal Medicine, University of Colorado Anschutz Medical Campus, 1890 North Revere Court, Aurora, CO, United States

[2]Vanderbilt University Medical Center, Nashville, TN, United States

[3]Cedars-Sinai Medical Center, Los Angeles, CA, United States

[4]University of California, San Diego, San Diego, United States

Corresponding Author:
Brad Morse, MA, PhD
Division of General Internal Medicine, University of Colorado Anschutz Medical Campus, 1890 North Revere Court, Aurora, CO, United States

Abstract

Background: Decision aids (DAs) are important tools that support shared decision-making (SDM) between clinicians and patients, enabling patients to be more informed and engaged in decisions regarding their care. The use of DAs can increase patient knowledge, reduce decisional regret, and engage the clinician and patient in meaningful dialog. Despite proven effectiveness in enhancing patient-centered care, a gap remains in clinician use of DAs. Known clinician barriers to using DAs include (1) time constraints, (2) concerns about the match between patient need and available DAs, (3) forcing users to leave the electronic health record (EHR) to access DAs, and (4) the burden of manually entering data into the DA.

Objective: This qualitative study identified the needs and requirements of clinicians to inform the design of a clinician-facing, EHR-integrated, Substitutable Medical Applications, Reusable Technologies (SMART; SMART Health IT) on Fast Healthcare Interoperability Resources (FHIR) (HL7) app, the Decision Aid Navigator (DEAN; University of Colorado Anschutz Medical Campus). The Navigator identifies and surfaces DAs that are relevant to a patient's health care conditions (eg, atrial fibrillation), current care (eg, not on anticoagulation), and demographics (eg, check the youngest age for the stroke prevention), and facilitates documentation of SDM discussions and decisions.

Methods: We conducted 13 semistructured interviews with clinicians who were recruited from 4 academic medical centers. Interviews included a demonstration of an initial, mid-fidelity, DEAN app prototype that was designed to address DA use and barriers described in the literature. The interviews focused on clinician context and use of the prototype, affordances and barriers to using the system, and clinician needs and requirements of the system. We used qualitative content analysis to code and reduce the data, using a consensus-making approach, and identify emerging themes.

Results: We identified 3 overarching themes: (1) streamlined functionality may simplify workflow and decrease the burden of DA use and SDM; (2) clinicians need appropriate competencies to effectively use the Navigator and relevant DAs; and (3) trust that the Navigator suggests prevetted DAs. Unanimously, clinicians shared that the DEAN Navigator should be integrated into the EHR. To accomplish this clear priority, clinicians stated that they needed the requisite competencies to successfully use the tool within their workflow and build trust with the tool itself.

Conclusions: Better tools to support and harness the benefits of SDM are needed. Overcoming the barriers of using DAs is paramount. Tools designed and developed to support DA use must be integrated into the EHR efficiently to create an opportunity for uptake of the technology by busy clinicians. If tools like DEAN can streamline the cumbersome process of documenting the use of DAs, more clinicians may potentially use DAs with their patients, given the right context and appropriate DA.

(JMIR Hum Factors 2025;12:e69756) doi:10.2196/69756

KEYWORDS

qualitative research; focus groups; user-centered design; shared decision-making; decision aid

XSL•FO
RenderX

Introduction

Background

Shared decision-making (SDM) and the use of decision aids (DA) tools improve patient-centered outcomes, such as satisfaction with understanding the risks and benefits of health care decisions and less decisional regret [1-3]. SDM is a deliberative process and is defined as: "an approach where clinicians and patients share the best available evidence when faced with the task of making decisions, and where patients are supported to consider options, to achieve informed preferences" [4]. Some believe the SDM process is best accomplished with the use of DAs [5], which are tools specifically designed to provide easily understandable information to facilitate clinician and patient communication and understanding of information to support informed decision-making, as compared to usual care [5].

DAs often elicit a patient's preferences and values and present evidence-based comparisons that are easily understandable to support clinicians and patients in their discussions and decision-making. A 2024 Cochrane review of 209 studies provided strong evidence that patient decision aids, DAs that target health care decisions made by patients, resulted in increased knowledge, more accurate risk perception, congruency between values and care choices, decreased decisional conflict and regret for patients, patient satisfaction with their decision, and improved communication between clinicians and patients [5,6]. In some scenarios, patients were less likely to choose major invasive surgery over conservative care and were more likely to start needed medications [6]. DAs come in many forms, including static text, videos [7], interactive websites [8], and even telenovelas [9], to support various learning preferences, cultures, and attention spans.

It is essential that patients and their care teams can work together to achieve patients' health goals, patient-centered outcomes, and effective, high-value health care systems. As a primary process of involving patients directly in their care, SDM is explicitly supported both by the National Academy of Health (NAM) (formerly the Institute of Medicine) [10,11] and the Affordable Care Act [12,13]. For instance, Medicare includes requirements for SDM as a condition of reimbursement for implantable cardiac defibrillator [14,15], lung cancer screening [16], and the left atrial appendage closure (WATCHMAN) device placement [15].

Barriers Facing Clinician DA Uptake

Similar to other clinical decision support (CDS) tools, DAs face many implementation barriers, such as increased workload and knowledge demands [17], and organizational factors (eg, culture and priorities) [18]. The result is the slow adoption of DAs and SDM approaches. Stacey et al found that the main barriers for uptake of DAs were (1) lack of funding to support implementation of DAs in clinical practice, including training in using DAs, (2) outdated DAs, and (3) clinician reluctance to use DAs due to increased workload, quality concerns, health record integration issues, manual data entry, and documentation burden [19,20].

Clinician DA use to support decision-making by patients is directly influenced by how cumbersome the process is to access and use the DA within existing clinical workflows [21]. Furthermore, clinicians desire a flexible approach to decision-making with patients [22]. These existing gaps in the literature suggest a need to more fully understand the complex clinical phenomenon and subjective experiences of clinicians to be better able to leverage Health Information Technology (HIT) to support expanded use of DAs.

The 5 Rights of CDS

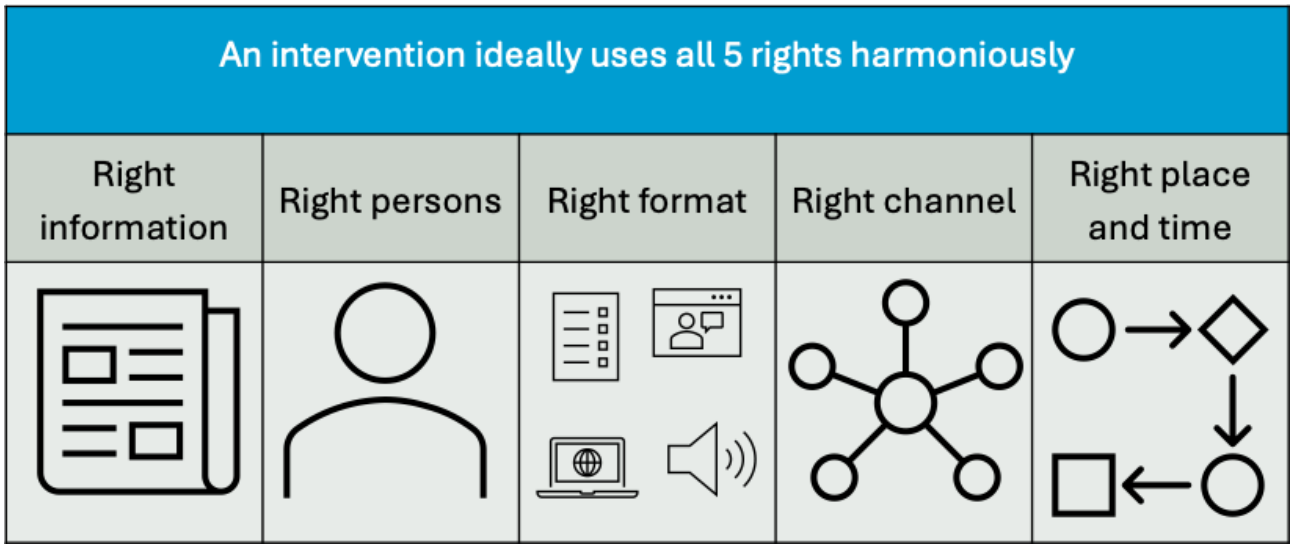
Osheroff's "The 5 Rights of CDS," shown in Figure 1, outline a framework for effective CDS implementation [23,24]. We applied the "5 Rights of CDS" framework to SDM support, specifically the facilitation of DA use in the SDM process. SDM and DA implementation face many of the same implementation barriers as CDS interventions. The "5 Rights of CDS" address many of the known barriers to SDM and the use of DAs. DAs provide the right information (eg, the relevant, current, and unbiased DA) to the right person (eg, clinician or patient) at the right time (eg, before the decision has been made), in the right format (eg, best format for the clinician or patient given the topic, such as written text or video and the patient's primary language), and in the right channel (eg, within the electronic health record [EHR] patient portal).

Despite the proven benefits of DAs and SDM [6], informatics systems to address the barriers that health systems, clinicians, and patients face when attempting to use DA are lacking. The 5 Rights of CDS and the use of DA are closely aligned with the aim to improve patient care by ensuring that the right information reaches the patient or clinician at the right time. Aligning the principles of the 5 Rights and SDM can improve the quality of care and support patient engagement to make important health decisions.

Based on the documented advantages of using DAs, the 5 Rights of CDS, and the research still needed in making DAs accessible in existing workflows, the objective of this study was to understand clinician needs and requirements for a HIT solution that surfaces relevant DAs. In our proposed solution, DAs are launched in the SMART (SMART Health IT) on FHIR (HL7) [25] Shared Decision Aid Navigator System (DEAN; University of Colorado Anschutz Medical Campus) navigator app and are surfaced in the customized DEAN EHR tab. DEAN was designed and developed as an interoperable SMART on FHIR [25] app and therefore is highly applicable and generalizable.

XSL•FO
RenderX

Figure 1. The 5 rights of clinical decision support (adapted from [23]) .



Methods

Overview

We (BM) conducted semistructured interviews (SSI), while other study staff were present to take notes (EL) and assist with facilitation (AS and LS), with clinicians identified at 4 US academic medical centers across the United States using purposive sampling [26,27]. Partners at the academic medical centers identified potential interview participants. We approached individuals via email to gauge interest in participating. Eligibility requirements for the SSI included clinicians with previous DA use experience and at least 20% clinical effort. Interviews continued until we reached thematic saturation, defined as not hearing any substantive new information from research participants.

Interview Guide and Data Collection Interaction

After a review of the literature to understand the already reported barriers and potential facilitators to DA use by clinicians, and alignment with the "5 Rights," we (LS, AS, and BM) developed an SSI guide to elicit insight from potential future users of the Navigator and ensure the same core questions were asked of all the participants while allowing flexibility to ask probing questions on other topics that arose during the discussion (Multimedia Appendix 1).

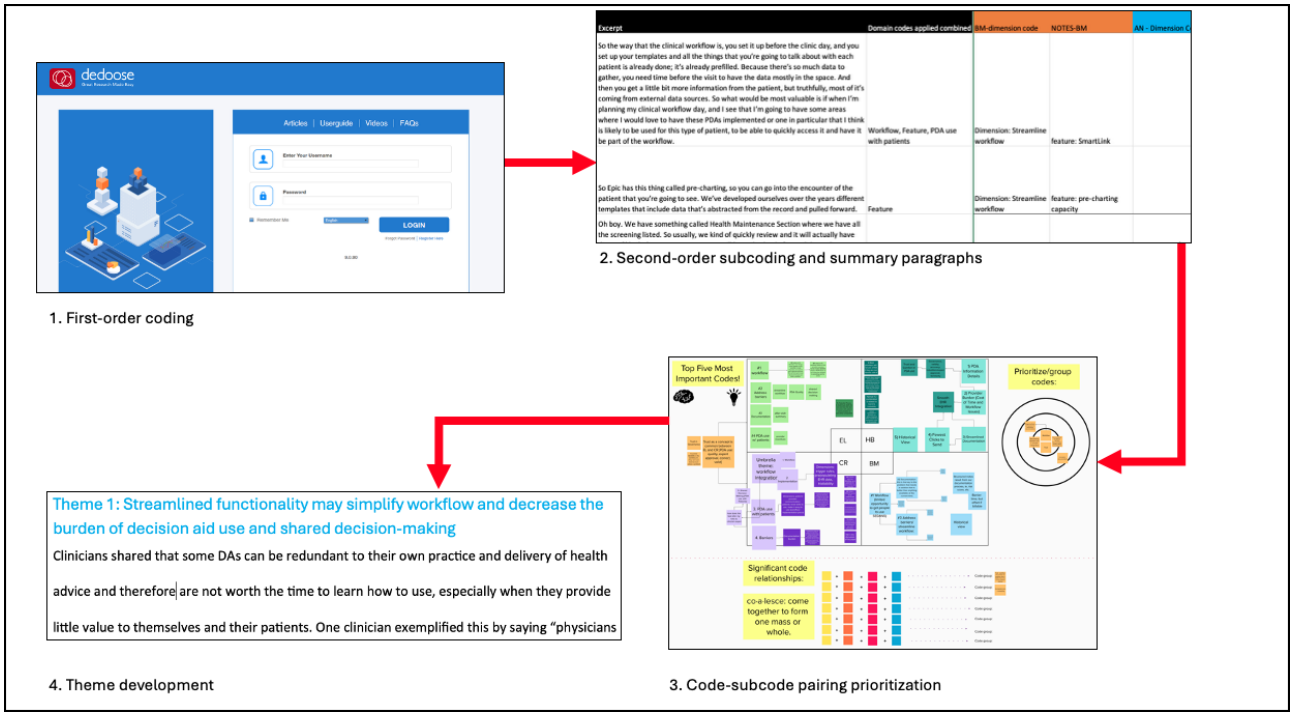
The SSI began with an elicitation of the participants' perspectives and experiences using DAs and having SDM discussions. This served to kick start the interview and provided the interviewer with the necessary personal context to conduct the remainder of the interview. As the purpose of the SSI was to obtain clinician input on our initial solution to the known barriers to DA use and SDM, we demonstrated a clickable, interactive prototype of DEAN (referred to as the Navigator henceforth) to demonstrate our proposed solution and collect insights from clinicians on how to further refine the prototype to support clinician use of DAs. The demonstration reviewed the known barriers and proposed solutions and facilitated input on the prototype features and functions from a user's perspective. Each interview lasted 1 hour and was conducted on the Zoom (Zoom Video Communications Inc) [28] video conferencing platform. Sessions were recorded and professionally transcribed verbatim for analysis.

Content Analysis Overview

We then conducted a 4-step qualitative coding process to analyze the interview transcripts using the Rapid and Rigorous Qualitative Data Analysis (RaDaR) technique [29-31]. This process included (1) creation and application of first-order codes, (2) application of second-order subcodes for a more nuanced and detailed coding scheme and a reduction of the code-subcode pairings to individual summary paragraphs [29,32], (3) identification and prioritization of relationships in the data, and (4) articulation of salient relationships into qualitative themes describing clinician needs and requirements of the Navigator (Figure 2).

XSL•FO
RenderX

Figure 2. Content analysis overview.



Qualitative Analysis Step 1: First-Order Coding

To accomplish data reduction and thematic development, a codebook was developed that contained the initial a priori codes, ie, first-order codes [33]. We then applied the codes to the interview transcripts to identify expected phenomenon (eg, barriers to the use of DAs) [32] using Dedoose (SocioCultural Research Consultants LLC), a qualitative analysis software [34]. If emergent codes (ie, new codes) were necessary during the coding process, the team defined the emergent code and discussed adding them to the codebook on a consensus basis (agreement from all 5 team members needed for the emergent code to be added). Coding was completed by the core qualitative team (AN, CR, SA, and BM). Each transcript was coded to consensus by 2 team members [33,35,36], a process in which discrepancies are discussed until consensus is reached [37], increasing the rigor of the coding [33,35,37].

Qualitative Analysis Step 2: Second-Order Subcoding and Summary Paragraphs

Once all the transcripts received codes, the Dedoose excerpts were entered into an Excel spreadsheet. Coders then applied more nuanced second-order subcodes. Subcodes were applied independently to Excel excerpts by at least 2 qualitative team members, and consensus was reached (AN, CR, EL, HB, SA, and BM) [36]. Excerpts that received the same code and subcode combination were grouped together and then reduced through synthesis to a single summary paragraph for each pairing.

Qualitative Analysis Steps 3: Code-Subcode Pairing Prioritization

Team members (CR, EL, HB, and BM) then reviewed all the code-subcode pairing summary paragraphs to identify the most relevant pairings to the needs and requirements of clinicians to support DA use. The saliency of the code-subcode pairing was determined by the pairing's information power [38], that is, the richness and relevance of the data in light of the study's objective and questions. Relationships were identified between code-subcode pairings that could be used to articulate the needs and requirements of clinicians for Navigator use.

Qualitative Analysis Steps 4: Theme Development

Mural [39] was used to visualize relevant code-subcode pairings and to interactively coproduce themes (BM, CR, EL, and HB) that characterized and described clinician needs and requirements. We reviewed and finalized the drafted themes ensuring that they accurately reflected our data. Themes were presented to the DEAN team (LS, AS, EL, PR, SS, RE, SA, and CR) for verification of relevance to our research question.

Ethical Considerations

The study was approved by the Colorado Multiple Institution Review Board (COMIRB #22 - 1061) and adhered to all relevant ethical guidelines, including informed consent procedures. We attest to maintaining privacy and confidentiality of research subjects' data and/or identity. The project received a waiver of informed consent. We distributed a postcard consent form outlining the study to all participants prior to data collection. Participants were informed that they could stop their participation at any time. Data were anonymized. Clinicians were compensated $100 for participating in the SSIs.

Results

Overview

Twenty clinicians were contacted to participate in the SSIs, of which 13 participated in them, representing 4 academic medical centers, while 7 did not participate (Table 1). One of the 7 clinicians did not respond, 3 were too busy, 3 did not meet the criteria (1 had no experience with DAs and 2 did not meet the minimum clinical full-time equivalent, FTE requirement of

XSL•FO
RenderX

20%). Interviews were approximately 60 minutes long. The data collected provided data saturation for understanding the needs and requirements of clinicians interacting with the Navigator.

We identified 3 primary themes. Theme 1 describes opportunities identified by the clinicians that may support the uptake of the Navigator. Themes 2 and 3 describe needs and requirements, training and trust, which clinicians reported as necessary to use a technology like the Navigator that surfaces relevant DAs.

Table . Semistructured interview participant demographics (n=13).

Demographic variables	Participants (N=13), n (%)
Sex identity	
Female	3 (23)
Male	8 (62)
Not reported	2 (15)
Age (years)	
26 - 32	1 (8)
33 - 40	5 (38)
41 - 49	4 (31)
50+	3 (23)
Race	
Asian	2 (15)
White	10 (77)
Not reported	1 (8)
Specialty	
Cardiology	1 (8)
Emergency Medicine	2 (15)
Family Practice	1 (8)
Geriatrics	1 (8)
Internal Medicine	5 (38)
Pediatrics	2 (15)
OB/GYN-Oncology	1 (8)
Clinical FTE[a] (remainder of time spent on informatics)	
20% - 50%	10 (77)
51% - 100%	3 (23)

[a]FTE: full-time equivalent.

Theme 1: Streamlined Functionality May Simplify Workflow and Decrease the Burden of DA Use and SDM

Clinicians shared that some DAs can be redundant to their own practice and delivery of health advice and therefore are not worth the time to learn how to use, especially when they provide little value to themselves and their patients. One clinician exemplified this by saying, “Physicians are already overburdened. Having to bring up another tool, the Navigator, and explain another decision [DA] to the patient takes more time.” Clinicians also shared that while they believed DAs can be valuable, they are not as significant an issue to tackle with their patients as there are more urgent problems the patient is experiencing.

According to the clinicians we interviewed, to combat the burdensome barriers and enable more clinicians to use DAs, the Navigator should address the following requirements: (1) integration into the EHR, (2) data collection burden reduction, and (3) documentation burden reduction. The DA delivery system that centralizes patient EHR data, populates required DA data, allows DAs to open within the EHR, and facilitates documentation of the SDM discussion in the EHR, are components that clinicians cited as desirable.

Too many clicks will be detrimental to the adoption of the DEAN technology. This concern was expressed by a clinician who asked, “If these tools exist, are they going to be within the existing workflow division or are they going to add extra clicks?” Another clinician warned, “All the third-party vendors try to make it seem like it’s seamless. It never is. And busy

XSL•FO
RenderX

clinicians don't fall for the seamless talk," expressing the importance of immediately and consistently demonstrating EHR integration in the existing workflow. Several clinicians recognized that automating tasks (eg, documentation) may increase time in the visit for patient engagement and enhanced patient-clinician relationships, sharing that "The clinician should spend time catching up with the patient, asking how they're doing and having that human aspect to it."

To overcome common barriers, it is not only important that the Navigator is integrated into the EHR, but that the Navigator documentation process be integrated within the EHR note-writing function to make it as easy as possible. Automated note population of DA use with patients was strongly preferred. Some physicians stated they document their use of the DA, including discussing risks with the patients and what they and the patient decided together. "The report, that would then populate into the note, which I think would be brilliant and is a perfect way to include it in the template." However, some find it difficult to balance thorough notes that are simple for patients to understand. In addition, some physicians find it challenging to either write a note with the patient during the visit (which can be very time-intensive) or to remember enough about the visit to write about it later. Templates, or even having the patient come in already engaged with the DA by entering required data, make the documentation process easier for clinicians.

Theme 2: Clinicians Need Appropriate Competencies to Effectively Use the Navigator and Relevant DAs

The use of a DA delivery system, like the Navigator, to support SDM involves interaction with complex technology and health care information. Clinicians communicated that training would be important before using the Navigator. They also felt they needed the knowledge and skills necessary to use the DAs. For clinicians, this includes the training required to access appropriate DAs in the Navigator and use the Navigator within the fast-paced and time-pressured clinical workflows, as well as possessing sufficient knowledge and experience with the clinical content of a DA. Currently, clinicians report being overburdened and learning how to use the Navigator and various DAs to implement the DSM processes into their workflow would be time-consuming. Many clinicians shared that they see many patients in a day and learning a new tool like the Navigator would only add to their busy schedules.

Clinicians identified the need to determine whether a DA is well designed and evidence-based. Alternatively, clinicians would rely on others (eg, a health system governing group or a medical organization) to endorse the DA. Clinicians described the substantial effort expended to adopt a new DA and stay current on the underlying evidence base. Evidence can change rapidly or be based on studies that are unfamiliar to the clinician. "I have very deep knowledge of the CDS tools that I use and understand how well they were validated and who did the studies and how high quality the studies were…we do journal clubs to talk about these tools…There is a depth of understanding and trust that happens before we apply them in clinical practice." Another clinician added, "Some of the other barriers that come up are comfort with the data. Like, you must know the material well yourself to be able to use some of these…and it changes all the time, so you must really be up to date."

Clinicians indicated that competency with the Navigator and specific DAs could be gained through trial-and-error adaptive learning. The opportunity to learn through practice was seen as a strategy to gain the experience and knowledge required to comfortably integrate DAs into clinical workflows. "I think trialability, (eg, using the tool before having to use it with patients) is another way to get over the mental model of adoption." Trialability will add work to the busy clinician's schedule, however, "And if you just gave me a new tool (DA and the Navigator) and say, hey, this is good for head injury patients, I would say, all right, I'll read about it after shift, and I may apply to the future." Relatedly, some clinicians highlighted the Navigator's ability to increase their competence and knowledge, stating, "there's a lot of times when not only am I having trouble explaining a complicated decision to a patient, but there's also actually information in these tools [the relevant DA] that I don't know."

Theme 3: Clinicians Must Trust the Navigator to Implement Use in Their Standard Workflow

A common response to the tool demonstration was that trust in the Navigator is a crucial prerequisite for the successful adoption of DAs. Clinicians are often reluctant to adopt technologies, such as the Navigator, when their benefits are uncertain [40]. Two types of trust were identified by clinicians.

Trust That the Navigator Is Aligned With Clinic Workflow

One clinician recognized the importance that streamlining the busy and burdensome clinical workflow may have on clinician's assessment of the Navigator. Trusting that the Navigator works smoothly within the busy clinical workflow is important because an indirect patient assessment of both the clinician and Navigator may result in, "…we're dealing with time and trust" of the patient. Violation of this expectation could erode clinician trust in the Navigator. Another clinician recognized that the Navigator might streamline the use of DAs, but gaining personal experience with the tool is an essential component for building trust in the system, "…I think after you try these three or four times, then the mental model barrier goes away. You're like, 'Oh, this is kind of helpful. It's not usurping me.'"

Trust That the Navigator Suggests Prevetted DAs They Can Use

Finally, selecting a quality DA and becoming familiar enough with its content to use it effectively during a visit was a concern for clinicians. Physicians feel that they need to establish familiarity with and trust in a specific DA in advance of using the DA during a patient encounter. One clinician shared, "If there were a lot of these [DAs] and I wasn't personally familiar with them, deeply familiar, then I would have difficulty…trying to make sure that I'm comfortable that this information applies to their particular case." Clinicians felt that while information embedded in the Navigator about a DA, such as evidence base or journal citations, could support clinicians' comfort with using a DA, clinicians would still have to become familiar with the DA to trust it enough to put it into practice. However, clinicians

XSL•FO
RenderX

stated that if DAs were prevetted by subject matter experts or committees at the local institution, uptake would be easier because the institution had previously endorsed veracity and use of the DA.

Discussion

Principal Findings

Our findings in Theme 1 reveal that clinicians have concerns about their ability to use DA and have SDM discussions during routine clinical care due to visit length time constraints, which are further impacted by technology that does not fit within their workflow and the need to address the most pressing and medically serious issues during visits. Workflow-concordant technology is essential for adoption and uptake, and clinicians stressed the importance of seamless EHR integration, prevetted and suggested DAs, DAs that do not require duplicate data entry, and the facilitation of the documentation of SDM discussions. While every concern matters, and clinicians and patients should work together to prioritize the visit agenda, clinicians are trained to recognize signs of urgent or life-threatening conditions that must take precedence. For instance, addressing symptoms of a possible heart attack would understandably take precedence over discussing sleep difficulties, regardless of the existence of the perfect insomnia management DA. Clinicians also expressed concerns about redundancy and a lack of value to DAs due to a clinician's expert knowledge of a topic or a simplistic decision. Many clinician recommendations and patient decisions do not require use of a DA, such as the recommendation and decision to take antibiotics for confirmed bacterial pneumonia. This contrasts with the high-stakes decision to have an implantable cardiac defibrillator, which is highly dependent on a patient's desire to lengthen their life and the distress caused by expected and realized electronic shocks. It is an obvious and understandable stance that DAs should only be used when they provide added value to complex SDM discussions. Although research suggests that clinicians recognize that they use paternalistic methods more often than they would prefer and that they have more limited discussions than they feel are optimal, mainly focusing on treatment options but not comparative pros and cons, and patient values and preferences. Many patients would like to be more engaged in their health care decisions [41,42].

Our findings formulated in Theme 2 indicate that clinicians may need to acquire new skills and knowledge to use a tool such as the Navigator and to effectively integrate DA use into existing health care workflows. In a study by Abbass et al [43], who found that lack of appropriate training was a key barrier to technology use, our research participants also commented on the need for both Navigator training and time to become familiar with a DA's content before using it with patients. A qualitative analysis of expert interviews in the context of artificial intelligence [44] described core competencies that clinicians need to effectively use these types of advanced technological tools within their clinical workflow [44]. The competencies identified include basic knowledge about the purpose and application of the tools; how to evaluate the quality, accuracy, and contextual appropriateness of the tools; how to adapt to changes in roles and workflows resulting from implementation; and the need to participate in continuing education related to use.

Similar competencies will be required for DA use in SDM, and with a tool like the Navigator, implementation planning and system training curriculum should account for the additional learning needed to successfully operationalize DAs in real-world settings. When clinicians participate in education opportunities on DAs, they are more likely to share a DA with a patient and engage in formal SDM [45]. With an increased skill set, it may follow that patient satisfaction with DAs and the SDM process would also improve. However, due to the high potential for ongoing changes to DA content, as new evidence and care standards emerge, continuing education will also be required. Continuing education should focus on SDM and DAs as a teachable skill [46].

As explained in Theme 3, while the benefits of DAs have been documented, the lag in widespread "trust" will persist until we can address multiple barriers. When designing and developing HIT that promotes trust, Jones et al [47] recommend designers clearly explain how the technology works, involve end-users in the development of the HIT, provide training and support for targeted end-users, allow for customization by end-users so that they can tailor the way they use the technology, and develop feedback mechanisms so that designers and developers can address system issues [47]. In our interviews, clinicians highlighted issues consistent with the recommendations made by Jones et al [47]. They wanted to trust that the Navigator system and the process of its implementation would work in their existing workflow. Therefore, before clinicians can trust the Navigator, which is mandatory, they need to be assured that the Navigator meets their needs and requirements as outlined in Themes 1 and 2.

The goal of health care technology is to provide greater convenience and clinical task efficiency, yet Abbas et al[43] cite key findings that barriers exist at technological, individual, and institutional levels that prevent widespread trust and adoption of health care technology. Specifically, these findings highlight compatibility, reliability, usability, training, and governance as barriers designers need to overcome to build clinician trust in new technologies [43]. Concerns were shared in the interviews that the Navigator would be vulnerable to these same trust issues and would need to be addressed before uptake.

The results of this study informed the future design, evaluation, and implementation plans of our proposed DEAN system. The rich qualitative data confirmed many of the functional requirements our team had previously identified based on known barriers to DA use and SDM. It also allowed further specification of some functional requirements and advanced the prototype to be used in further user-centered design sessions. For example, we sought to ensure the Navigator interface would integrate with the EHR in a manner consistent with EHR use in clinical care, opening within the EHR and not within an external browser. We also confirmed the need to create a system to support institutional curation of vetted DAs. In response to clinician input, we designed a system that surfaced Suggested DAs based on patients' current problems and medications, in a

XSL•FO
RenderX

manner that was noninterruptive, as clinicians expressed that DA use would be based on agenda prioritization and the need for a formal SDM discussion and the added value of the DA to the discussion. We also planned to streamline the filing of the SDM discussion note to the EHR through a variety of commonly used pathways. Our hope is that data-informed improvements in DA use workflows through the DEAN system, including the Navigator, will support greater uptake of DAs and enhance SDM conversations for clinicians and patients.

Furthermore, we believe the Navigator could lead to more trust on the part of the clinician. For example, the Navigator is designed to have a tile in the interface for each DA. Each DA tile will have an information button that provides readily available information on topics such as the references for the evidence on which the DA was developed and the DA creators. The information button also provides information to indicate the governing person or committee at the institution responsible for vetting and releasing the DA within the Navigator. Finally, we believe that training in Navigator use and time for clinicians to become familiar with the DAs will increase clinician trust of the tool.

Limitations

Some participants conflated the idea of risk calculators with DAs, but the facilitators were able to clarify the difference and provide relevant examples. We only conducted interviews with academic clinicians and community clinicians; practices may have different opinions. If anything, we suspect their time constraints and need to focus on revenue-generating activities may be even greater for community physicians. This could lead to varying stances, such as a clinician never making the time to use a DA and participate in a thorough SDM discussion, or a clinician who welcomes the timesaving and workflow-aligned functionality that the Navigator offers.

Conclusions

Future directions include the further development of the DEAN system, including the Navigator, and other backend systems to support DA curation and documentation of key facts such as DA governing entities, DA authors, evidence citations, evidence in the form of brief written summaries, and DA version date, including when the DA was last updated. Our hope is that the process of including clinicians early in the process of the DEAN system design leads to a tool that is workflow aligned and addresses the major barriers associated with DA use and SDM discussions. This should result in a system that has greater uptake and adoption, ultimately leading to better patient engagement and satisfaction with their medical decisions.

Acknowledgments

The authors would like to thank all the participants that assisted with the user-centered design discovery phase. Research reported in this publication was supported by the Agency for Healthcare Research and Quality (AHRQ) under award number 5R18HS028578-03. The content is solely the responsibility of the authors and does not necessarily reflect the official views of the NIH or AHRQ.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

LS, AS, and BM, conceptualized the study. BM, EL, AS, and LS collected the data. BM, EL, ATN, CR, SA, and HB analyzed the data. BM, EL, CR, and HB contributed to the first draft of the paper, along with contributions from all the authors. LS, AS, and BM performed supervision of the project and finalized the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
DEAN: Semistructured interview guide for Navigator requirements.
[DOCX File, 52 KB - humanfactors_v12i1e69756_app1.docx]

References

1. Frank L, Basch E, Selby JV, Patient-Centered Outcomes Research Institute. The PCORI perspective on patient-centered outcomes research. JAMA 2014 Oct 15;312(15):1513-1514. [doi: 10.1001/jama.2014.11100] [Medline: 25167382]
2. van Tol-Geerdink JJ, Leer JWH, Wijburg CJ, et al. Does a decision aid for prostate cancer affect different aspects of decisional regret, assessed with new regret scales? A randomized, controlled trial. Health Expect 2016 Apr;19(2):459-470. [doi: 10.1111/hex.12369] [Medline: 25940277]

XSL•FO
RenderX

3. Ehrbar V, Germeyer A, Nawroth F, et al. Long-term effectiveness of an online decision aid for female cancer patients regarding fertility preservation: knowledge, attitude, and decisional regret. Acta Obstet Gynecol Scand 2021 Jun;100(6):1132-1139. [doi: 10.1111/aogs.14108] [Medline: 33538329]
4. Elwyn G, Laitner S, Coulter A, Walker E, Watson P, Thomson R. Implementing shared decision making in the NHS. BMJ 2010 Oct 14;341:c5146. [doi: 10.1136/bmj.c5146] [Medline: 20947577]
5. Stacey D, Lewis KB, Smith M, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2024 Jan 29;1(1):CD001431. [doi: 10.1002/14651858.CD001431.pub6] [Medline: 38284415]
6. Stacey D, Légaré F, Lewis K, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2017 Apr 12;4(4):CD001431. [doi: 10.1002/14651858.CD001431.pub5] [Medline: 28402085]
7. Volandes AE, Paasche-Orlow MK, Barry MJ, et al. Video decision support tool for advance care planning in dementia: randomised controlled trial. BMJ 2009 May 28;338:b2159. [doi: 10.1136/bmj.b2159] [Medline: 19477893]
8. Col NF, Ngo L, Fortin JM, Goldberg RJ, O'Connor AM. Can computerized decision support help patients make complex treatment decisions? A randomized controlled trial of an individualized menopause decision aid. Med Decis Making 2007;27(5):585-598. [doi: 10.1177/0272989X07306781] [Medline: 17873260]
9. Jibaja-Weiss ML, Volk RJ, Granchi TS, et al. Entertainment education for breast cancer surgery decisions: a randomized trial among patients with low health literacy. Patient Educ Couns 2011 Jul;84(1):41-48. [doi: 10.1016/j.pec.2010.06.009] [Medline: 20609546]
10. Alston C, Elwyn G, Fowler F, et al. Shared decision-making strategies for best care: patient decision aids. NAM Perspectives 2014;4(9). [doi: 10.31478/201409f]
11. Briere R. Institute of Medicine: Crossing the Quality Chasm: National Academies Press; 2001. [Medline: 25057539]
12. The Patient Protection and Affordable Care Act Section 933.: Authenticated US Government Information (GPO); 2010 Mar 23 URL: https://www.congress.gov/111/plaws/publ148/PLAW-111publ148.pdf [accessed 2025-07-15]
13. Patient preference information. The Food and Drug Administration. 2015. URL: http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM446680.pdf
14. Centers for Medicare and Medicaid Services. Decision memo for implantable cardioverter defibrillators. 2018. report no.: CAG-00157R4. 2018 URL: https://www.cms.gov/medicare-coverage-database/details/nca-decision-memo.aspx?NCAId=288 [accessed 2019-02-02]
15. National coverage determination (NCD) for percutaneous left atrial appendage closure (LAAC). Centers for Medicare & Medicaid Services. URL: https://www.cms.gov/medicare-coverage-database/details/ncd-details.aspx?NCDId=367 [accessed 2021-08-17]
16. Decision memo for screening for lung cancer with low dose computed tomography (LDCT)(CAG-00439N). : Centers for Medicare & Medicaid Services; 2015.
17. Meunier PY, Raynaud C, Guimaraes E, Gueyffier F, Letrilliart L. Barriers and facilitators to the use of clinical decision support systems in primary care: a mixed-methods systematic review. Ann Fam Med 2023;21(1):57-69. [doi: 10.1370/afm.2908] [Medline: 36690490]
18. O'Donnell S, Cranney A, Jacobsen MJ, Graham ID, O'Connor AM, Tugwell P. Understanding and overcoming the barriers of implementing patient decision aids in clinical practice*. Evaluation Clinical Practice 2006 Apr;12(2):174-181. [doi: 10.1111/j.1365-2753.2006.00613.x]
19. Stacey D, Suwalska V, Boland L, Lewis KB, Presseau J, Thomson R. Are patient decision aids used in clinical practice after rigorous evaluation? A survey of trial authors. Med Decis Making 2019 Oct;39(7):805-815. [doi: 10.1177/0272989X19868193] [Medline: 31423911]
20. Ankolekar A, Dekker A, Fijten R, Berlanga A. The benefits and challenges of using patient decision aids to support shared decision making in health care. JCO Clin Cancer Inform 2018 Dec;2(2):1-10. [doi: 10.1200/CCI.18.00013] [Medline: 30652607]
21. Ng CJ, Lee PY, Lee PY, Abdullah KL. Health innovations in patient decision support: bridging the gaps and challenges. Australas Med J 2013;6(2):95-99 [FREE Full text] [doi: 10.4066/AMJ.2013.1655] [Medline: 23483776]
22. Yu CH, Ke C, Jovicic A, et al. Beyond pros and cons - developing a patient decision aid to cultivate dialog to build relationships: insights from a qualitative study and decision aid development. BMC Med Inform Decis Mak 2019 Sep 18;19(1):186. [doi: 10.1186/s12911-019-0898-5] [Medline: 31533828]
23. Osheroff JA, Teich J, Levick D, et al. Improving Outcomes with Clinical Decision Support: An Implementer's Guide: HIMSS Publishing; 2012. URL: https://www.taylorfrancis.com/books/mono/10.4324/9780367806125/improving-outcomes-clinical-decision-support-jerome-osheroff-jonathan-teich-donald-levick-luis-saldana-ferdinand-velasco-dean-sittig-kendall-rogers-robert-jenders
24. Peek ME, Wilson SC, Gorawara-Bhat R, Odoms-Young A, Quinn MT, Chin MH. Barriers and facilitators to shared decision-making among African-Americans with diabetes. J Gen Intern Med 2009 Oct;24(10):1135-1139. [doi: 10.1007/s11606-009-1047-0] [Medline: 19578818]
25. Mandel JC, Kreda DA, Mandl KD, Kohane IS, Ramoni RB. SMART on FHIR: a standards-based, interoperable apps platform for electronic health records. J Am Med Inform Assoc 2016 Sep;23(5):899-908. [doi: 10.1093/jamia/ocv189] [Medline: 26911829]

XSL•FO
RenderX

26. Campbell S, Greenwood M, Prior S, et al. Purposive sampling: complex or simple? Research case examples. J Res Nurs 2020 Dec;25(8):652-661. [doi: 10.1177/1744987120927206] [Medline: 34394687]
27. Etikan I, Musa SA, Alkassim RS. Comparison of Convenience Sampling and Purposive Sampling. Am J Theor Appl Stat 2016;5(1):1. [doi: 10.11648/j.ajtas.20160501.11]
28. Video conferencing, cloud phone, webinars, chat, virtual events. Zoom. URL: https://zoom.us/ [accessed 2021-07-09]
29. Watkins DC. Rapid and rigorous qualitative data analysis: The "RADaR" technique for applied research. Int J Qual Methods 2017;16(1):1609406917712131.
30. Watkins DC, Abelson JM, Jefferson SO. "Their depression is something different... it would have to be": findings from a qualitative study of black women's perceptions of depression in black men. Am J Mens Health 2013 Jul;7(4 Suppl):45S-57S. [doi: 10.1177/1557988313493697] [Medline: 23784520]
31. Watkins DC, Allen JO, Goodwill JR, Noel B. Strengths and weaknesses of the Young Black Men, Masculinities, and Mental Health (YBMen) Facebook project. Am J Orthopsychiatry 2017;87(4):392-401. [doi: 10.1037/ort0000229] [Medline: 27977287]
32. Kleinheksel AJ, Rockich-Winston N, Tawfik H, Wyatt TR. Demystifying content analysis. Am J Pharm Educ 2020 Jan;84(1):7113. [doi: 10.5688/ajpe7113] [Medline: 32292185]
33. Bradley EH, Curry LA, Devers KJ. Qualitative data analysis for health services research: developing taxonomy, themes, and theory. Health Serv Res 2007 Aug;42(4):1758-1772. [doi: 10.1111/j.1475-6773.2006.00684.x] [Medline: 17286625]
34. Great research made easy. Dedoose. URL: https://dedoose.com/ [accessed 2023-04-20]
35. Alok S, Raveendran J, Shaheen M. Conflict management strategies used by Indian software companies: a summative content analysis. IUP Journal of Soft Skills 2014;8(4):47-61 [FREE Full text]
36. Barr PJ, Haslett W, Dannenberg MD, et al. An audio personal health library of clinic visit recordings for patients and their caregivers (HealthPAL): user-centered design approach. J Med Internet Res 2021 Oct 22;23(10):e25512. [doi: 10.2196/25512] [Medline: 34677131]
37. Richards KAR, Hemphill MA, Richards KAR, Hemphill MA. A practical guide to collaborative qualitative data analysis. J Teach Phys Educ 2018 Apr;37(2):225-231. [doi: 10.1123/jtpe.2017-0084]
38. Malterud K, Siersma V, Guassora AD. Information Power: Sample Content and Size in Qualitative Studies Qualitative Research in Psychology: Expanding Perspectives in Methodology and Design, 2nd edition: American Psychological Association; 2021:67-81. [doi: 10.1037/0000252-004]
39. Unlock productivity & empower transformation. MURAL. URL: https://www.mural.co/ [accessed 2025-07-09]
40. Wang T, Tan JYB, Liu XL, Zhao I. Barriers and enablers to implementing clinical practice guidelines in primary care: an overview of systematic reviews. BMJ Open 2023 Jan 6;13(1):e062158. [doi: 10.1136/bmjopen-2022-062158] [Medline: 36609329]
41. Driever EM, Tolhuizen IM, Duvivier RJ, Stiggelbout AM, Brand PLP. Why do medical residents prefer paternalistic decision making? An interview study. BMC Med Educ 2022 Mar 8;22(1):155. [doi: 10.1186/s12909-022-03203-2] [Medline: 35260146]
42. Wexler R, Gerstein BS, Brackett C, et al. Decision aids in the United States: the patient response. Int J Pers Cent Med ;5(3):105-111. [doi: 10.5750/ijpcm.v5i3.517]
43. Abbas RM, Carroll N, Richardson I, Beecham S. Trust factors in healthcare technology: a healthcare professional perspective. : University of Limerick; 2018. URL: https://researchrepository.ul.ie/articles/conference_contribution/Trust_factors_in_healthcare_technology_a_healthcare_professional_perspective/19851862
44. Russell RG, Lovett Novak L, Patel M, et al. Competencies for the use of artificial intelligence–based tools by health care professionals. Acad Med 2023 Mar 1;98(3):348-356. [doi: 10.1097/ACM.0000000000004963] [Medline: 36731054]
45. Légaré F, Adekpedjou R, Stacey D, et al. Interventions for increasing the use of shared decision making by healthcare professionals. Cochrane Database Syst Rev 2018 Jul 19;7(7):CD006732 [FREE Full text] [doi: 10.1002/14651858.CD006732.pub4] [Medline: 30025154]
46. Thériault G, Bell NR, Grad R, Singh H, Szafran O. Teaching shared decision making: an essential competency. Can Fam Physician 2019 Jul;65(7):514-516. [Medline: 31300437]
47. Jones C, Thornton J, Wyatt JC. Enhancing trust in clinical decision support systems: a framework for developers. BMJ Health Care Inform 2021 Jun;28(1):e100247. [doi: 10.1136/bmjhci-2020-100247] [Medline: 34088721]

Abbreviations

CDS: clinical decision support
DA: decision aid
DEAN: Decision Aid Navigator
EHR: electronic health record
FHIR: Fast Healthcare Interoperability Resources
HIT: Health Information Technology
NAM: National Academy of Health

XSL•FO
RenderX

PDA: patient decision aid
RaDaR: Rigorous and Accelerated Data Reduction
SDM: shared decision-making
SSI: semistructured interview

Edited by A Kushniruk; submitted 06.12.24; peer-reviewed by D Ubbink, FH Leung, R Appiah-Kubi; revised version received 14.05.25; accepted 27.05.25; published 21.08.25.

Please cite as:
Morse B, Reale C, Nguyen AT, Latella E, Bauguess H, Anders S, Roberts P, SooHoo SL, El-Kareh R, Soares A, Schilling L
Clinician Needs and Requirements for a Decision Aid Navigator: Qualitative Study
JMIR Hum Factors 2025;12:e69756
URL: https://humanfactors.jmir.org/2025/1/e69756
doi:10.2196/69756

© Brad Morse, Carrie Reale, An T Nguyen, Erin Latella, Hannah Bauguess, Shilo Anders, Pamela Roberts, Spencer L SooHoo, Robert El-Kareh, Andrey Soares, Lisa Schilling. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 21.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Symptom Management Preference and Persona Development for Mobile Health Design Targeting Chinese Older Adult Patients With Breast Cancer: Descriptive Qualitative Study

Danyu Li[1], PhD; Yanlin Zhu[1], BSc; Changrong Yuan[1], PhD; Qingmei Huang[1], PhD; Yang Yang[2,3], MSc; Zhaohui Geng[4], PhD; Xiaotong Yuan[5], BSc; Fulei Wu[1], PhD

[1]School of Nursing, Fudan University, 305 Fenglin Road, Shanghai, China

[2]Department of Nursing, Fudan University Shanghai Cancer Center, Shanghai, China

[3]Department of Oncology, Shanghai Medical College, Fudan University, Shanghai, China

[4]School of Nursing, Shanghai University of Traditional Chinese Medicine, Shanghai, China

[5]International Medical Services, Peking Union Medical College Hospital, Beijing, China

Corresponding Author:
Fulei Wu, PhD
School of Nursing, Fudan University, 305 Fenglin Road, Shanghai, China

Abstract

Background: Mobile health (mHealth) for breast cancer care can greatly benefit patients' symptom management. Although research supports the effectiveness of mHealth, older adult patients with breast cancer often face difficulties using it, hindering them from accessing effective symptom management possibilities. Understanding the preference for mHealth among this population is crucial for providing insights into effective mHealth design.

Objective: This study aimed to better understand the symptom management preference using mHealth for Chinese older adult patients with breast cancer and use the approach of personas to inform the mHealth design.

Methods: This was a descriptive qualitative study. In total, 17 patients with breast cancer aged 60 years and older were recruited from tertiary hospitals in Shanghai, China, using purposive sampling. Data were collected through one-on-one interviews. Content analysis was used to identify the factors that influence participants' symptom management preference using mHealth. The categories of influencing factors of preference informed the persona template and guided the development of the persona.

Results: We identified 3 major categories affecting participants' preference for mHealth, including social interaction patterns, mHealth literacy, and symptoms. The following five personas were developed: (1) Positive Manager, (2) Dependent Parent, (3) Management Isolationist, (4) Image Manager, and (5) Clinician Dependent. We provide insights into how these personas can be used when designing and implementing mHealth for symptom management support.

Conclusions: Key factors influencing symptom management preference using mHealth among Chinese older adult patients with breast cancer and personas developed based on that can foster a better understanding of this population and initiate future mHealth design and implementation.

(JMIR Hum Factors 2025;12:e71448) doi:10.2196/71448

KEYWORDS

qualitative; mHealth; preference; symptom management; older adults; personas

Introduction

Breast cancer is among the main illnesses that threaten the lives and health of older adult women in China, with a double incidence peak at 45 - 55 and 70 - 74 years of age [1]. The mean age for breast cancer diagnosis in China is 45 - 55 years, which is considerably younger than that in Western countries, where the average age is between 48 and 50 years. However, due to the current unrestrained population aging, a substantial number of older adult patients with breast cancer may emerge. Although only 16.6% of patients with breast cancer were aged 65 years or older in 2018 in China, the incidence rate is estimated to increase to 27% by 2030 [2], leading to growing concerns regarding the management of older adult patients with breast cancer. According to previous studies, older adult patients with breast cancer experience higher rates of comorbidities, poorer performance status, limited social support, and difficulty with transportation, thus leading to different treatment approaches than those provided for their younger counterparts or undertreatment [3]. Older adult patients with breast cancer may face a complicated recovery process, insufficient information access, and a lack of emotional support [4,5].

XSL•FO
RenderX

Fortunately, mobile health (mHealth), which refers to medical and public health practices supported by mobile devices [6], is a promising tool that supports symptom management by providing unprecedented access to specialist clinical diagnostics and treatment advice [7]. Currently, mHealth is providing substantial improvements in human lifestyle behaviors and chronic condition management, facilitating the popularity and accessibility of health interventions [8,9]. Regarding disease management for patients with breast cancer, mHealth has demonstrated great potential for improving lifestyles, managing symptoms, and providing information and emotional support [10-14].

Integrating mHealth solutions for older adults can be significantly beneficial to this population, but the process faces various challenges [15]. For example, many mHealth tools may not be designed considering age-related dilemmas, such as vision or dexterity limitations, thus causing accessibility and usability problems [16]. Furthermore, the older adults from lower socioeconomic backgrounds may face affordability barriers to smartphone access and internet connectivity [17,18]. Moreover, a lack of tailored mHealth designed specifically for their needs can decrease the older adults' willingness to use mHealth, preventing them from adapting to technology and receiving the associated benefits [19,20]. Thus, the successful usage and adoption of mHealth are fairly poor in aging patients, accounting for a large proportion of the customer base for innovative health care devices [21].

Considering the importance of understanding the unique needs of older adult patients with breast cancer, we adopt personas to guide analyses of the needs of this population. The concept of persona was first introduced by Alan Cooper in 1999 [22]. It refers to virtual figures that share specific needs and unique personalities and are categorized based on key information about a target group [23]. Personas are detailed personas that embody the characteristics, behaviors, motivations, goals, and pain points of the identified user segments, thereby enabling more precise interventions [24,25]. Personas serve as the main character in a narrative, a scenario-based approach to intervention design that iteratively generates concepts (in the envision phase), provides feedback to enhance design coherence and appropriateness (in the refine phase), and provides a powerful communication tool to help developers understand the design rationale and prioritize features based on user needs. Considering persona was continuously adopted in different health contexts [20], the age-related challenges in the aging population facing mHealth, in part, can be compensated for by designing senior-friendly mHealth devices using persona.

The specific needs related to mHealth for improving symptom management among Chinese older adult patients with breast cancer have received little focus, and the insufficient use of personas requires further exploration. We typically focus on patients undergoing chemotherapy in this study because of the long duration and cumulative nature of this treatment phase, which is characterized by consistent and increasingly severe side effects such as fatigue, nausea, and immunosuppression. Understanding these patients' unique challenges is essential for developing effective mHealth solutions to improve their symptom management and overall quality of life. Therefore, this study aims to (1) understand the symptom management behavior and preference for mHealth among this population using qualitative data, (2) develop personas representing patients with different symptom management behavior and preference for mHealth, and (3) provide guidance for integrating personas into the design of mHealth.

Methods

Sampling

This study was conducted from April 2022 to February 2023 at the Fudan University Shanghai Cancer Center and Huadong Hospital in Shanghai, China. The purposive sampling method was adopted to enroll participants. The inclusion criteria were as follows: (1) diagnosed with primary breast cancer, (2) aged 60 years and older, (3) undergoing chemotherapy after breast cancer surgery or within 2 weeks of chemotherapy ending, (4) expected survival time of more than 6 months, and (5) know about their own disease status and can express their thoughts and options. The exclusion criteria were as follows: (1) patients currently receiving radiation therapy and (2) patients with vital organ dysfunction. The sample size was guided by the principle of data saturation. Interviews were discontinued after a minimum of 10 sessions, once no new themes emerged across 3 successive interviews [26].

Data Collection

Eligible participants were identified by nurse leaders in the breast oncology departments based on predefined inclusion and exclusion criteria. After briefly introducing the study aims and obtaining verbal agreement from potential participants to be contacted, the nurse leaders referred the interested individuals to the research team.

Before each interview, participants received a detailed explanation of the study from the research team and signed written informed consent forms. General participant characteristics were then recorded, including age, education, chemotherapy stage, primary carer, and whether living with a primary carer.

Subsequently, one of the authors (YZ) and the corresponding author (FW) conducted the face-to-face interviews. YZ is a nursing undergraduate student who took classes in qualitative study, and FL is a lecturer and experienced qualitative researcher with a research focus on breast cancer and imparted qualitative research methodology in the university. Before the beginning of data collection, YZ was trained by FW, covering the principles of qualitative research, techniques of conducting semi-structured interviews (building rapport, active listening, and probing), and mock interview exercises based on the interview guide. Each interview was audio-recorded and lasted 30 - 50 minutes. During data collection, the interviewers summarize and confirm their interpretation of what a participant said during data collection as a member checking to secure the trustworthiness.

Data were collected through semi-structured interviews using open-ended questions. The study aimed to gather qualitative data about participants' behavior and patterns for symptom management and preference for using mHealth; therefore, the

XSL•FO
RenderX

questions were developed carefully around this topic. The original questions were refined after trial interviews with 3 potential patients, and the final interview outlines are presented in Textbox 1. The data from trial interviews were not included in the final data analysis.

Textbox 1. Interview outline around patients' symptom management behavior and preference for using mHealth to manage symptoms.

1. Since starting chemotherapy, how do you feel?
2. What physical, mental, and social interaction changes have you noticed?
3. What are the most dominating symptoms you feel?
4. How do you manage your symptoms? What are the barriers and facilitators of your symptom management experience?
5. Did you get any help to manage your symptoms? Regarding help and resources from your spouse, children, relatives, friends, and health care providers?
6. How do you think about your symptom management skills, attitudes, and experience?
7. How do you think about managing your symptoms using mHealth (eg, knowledge distribution apps, virtual health consultation websites, symptom assessment apps)? Are there any concerns that might discourage you from using mHealth?
8. What features would you find most helpful in mHealth for symptom management?

Ethical Considerations

This study was reviewed and approved by the Institutional Review Board of Fudan University School of Nursing (IRB no. TYSQ2021-03-05). All participants provided informed consent prior to data collection.

Data Analysis

The audio-recorded interviews were transcribed within 24 hours after interview completion. All transcripts were managed and coded using NVivo 12. Data were analyzed using qualitative content analysis, and transcripts were coded inductively [27]. By using inductive coding, the data drove the analytical process, ensuring a comprehensive exploration of patterns and themes based on the interview transcripts. This approach allows the researcher to derive insights and meaning directly from the data rather than imposing preconceived coding categories. YZ and DL analyzed the data independently, and the main steps are as follows: (1) immersing in the data by line-by-line reading of the transcript; (2) selecting the unit of analysis; (3) making sense of the data as a whole; (4) open coding; and (5) group and categorization. After open coding, the preliminary codebook was developed and was iteratively refined through constant comparison and discussion between 2 researchers. Codes were then clustered into subcategories and higher-level categories. Intercoder reliability was addressed through regular meetings to compare coding results and resolve discrepancies via discussion until consensus was reached. Investigator triangulation was implemented by involving 2 researchers in data coding and theme development. Discrepancies were resolved through discussion with the other 2 researchers (CY and FW, with 25 y and 10 y of qualitative research, respectively) to reach consensus. Investigator triangulation was used to ensure a comprehensive and trustworthy exploration of subcategories and categories grounded in the interview transcripts.

Persona Segmentation

After the qualitative data analysis, we identified the categories and subcategories of symptom management behavior and preference for using mHealth to manage symptoms, which could be used as the persona segmentation template. Patients with similar features of categories and subcategories were grouped. Based on these groupings, a persona was created to represent the shared patterns within each subgroup [28]. This ensured that each persona reflected not just individual traits but collective tendencies observed across participants. These personas were sent back to participants for validation to ensure an accurate representation of their characteristics and experiences. Final groupings were refined through team discussions and reviewed by qualitative experts to ensure consistency.

Persona Creation

Subsequently, we designed the persona skeleton, a layout describing the key information that the persona includes [28]. After finalizing the details of each persona, we used visual icons to highlight their key features and enhance vividness. To further humanize the personas, we assigned them names and incorporated representative quotes to reflect each target group's personality and context [25]. The visual elements were hand-drawn by the research team to maintain authenticity and engagement.

Results

Participants

A total of 17 participants were approached and agreed to participate in the interviews. Participant characteristics are outlined in Table 1. The participants had a mean age of 69.6 (range 61 - 80) years. All the participants were female and had undergone radical mastectomy. The most common primary caregivers were their adult children.

XSL•FO
RenderX

Table . Sociodemographic characteristics of the participants.

No.	Age	Education	Chemotherapy stage	Primary carer	Live with primary carer
N1	72	Primary school	2nd-4th cycle	Children	Yes
N2	73	Illiteracy	5th-8th cycle	Children	Yes
N3	66	High school	2nd-4th cycle	Spouse	Yes
N4	68	Primary school	2nd-4th cycle	Children	Yes
N5	79	Illiteracy	5th-8th cycle	Children	Yes
N6	81	Illiteracy	Chemo-ended	Children	Yes
N7	78	Illiteracy	Chemo-ended	Children	Yes
N8	63	Middle school	5th-8th cycle	Spouse	Yes
N9	61	High school	2nd-4th cycle	Children	No
N10	65	High school	2nd-4th cycle	Children	No
N11	80	Primary school	5th-8th cycle	Children	Yes
N12	66	Middle school	5th-8th cycle	Patient self	NA
N13	65	High school	2nd~4th cycle	Patient self	NA
N14	71	Primary school	Chemo-ended	Children	Yes
N15	64	College	2nd-4th cycle	Children	No
N16	69	High school	Chemo-ended	Children	Yes
N17	63	Middle school	5th-8th cycle	Spouse	Yes

Symptom Management Behaviors and Preference for Using mHealth to Manage Symptom

We identified 3 categories, 7 subcategories, and corresponding descriptions (Table 2). The results describe the symptom management behaviors and preference for using mHealth to manage symptoms among Chinese older adult breast cancer patients.

Table . Symptom management behavior and preference for using mHealth to manage symptoms.

Categories	Subcategories	Description
Social interaction patterns	Living conditions	Living with children or spouse who provide 24/7 care or living alone without a specific caregiver
	Social dynamics	Regular interacting with families or friends or completely being isolated and staying at home
mHealth literacy	mHealth access	Having or not having mobile devices (phone or computer) and Internet connection
	mHealth usage	Knowing or not knowing how to use mHealth devices or with amount of experience
	mHealth interest	Being willing or unwilling to interact with others via mobile devices
Symptoms	Symptom diversity	Experiencing many or only a few symptoms simultaneously
	Symptom severity	Experiencing tremendous or bearable symptom burden

Personas

We classified our findings into 5 personas, with differences in social interaction patterns, mHealth literacy, and symptoms. We typologized the personas as follows: Positive Manager, Dependent Parent, Management Isolationist, Image Manager, and Clinician-Dependent (Table 3).

XSL•FO
RenderX

Table . Persona overview.

Features	Positive Manager	Dependent Parent	Management Isolationist	Image Manager	Clinician-Dependent
					
Age	65	66	81	60	70
Education	High school	Primary school	Illiteracy	College	Primary school
Living conditions	Lives with spouse and children providing care	Lives with children who provide 24/7 care	Lives with spouse	Lives with spouse	Lives with children
Social dynamics	Regular interaction with family and friends	Regularly interacts with family members	Isolated, stays mostly at home	Seeks interactions with patient support groups	Regular interaction with clinicians
mHealth access	Has mobile devices and Internet access	Limited access, often relies on children for tech support	No mobile devices or Internet	Has access	Has mobile devices but prefers clinician support
mHealth usage	Actively uses mHealth, searches for professional health information	Uses mHealth only with children's assistance	Refuses to use mHealth, relies solely on children	Uses mHealth to communicate with peer groups but may need help	Relies heavily on clinicians, limited personal mHealth use
mHealth interest	Interested in web-based clinician interactions and reliable information	Prefers family involvement in mHealth, little personal interest	No interest in using mHealth directly	Prefers engaging with online patient communities for social and emotional support	Trusts only clinician-guided mHealth solutions
Symptom diversity	Experiences multiple symptoms (eg, gastrointestinal issues)	Deals with a mix of symptoms, mainly related to chemotherapy	Experiences chronic pain and discomfort	Focuses on managing self-esteem and emotional symptoms	Multiple symptoms, particularly concerned with treatment side effects
Symptom severity	Manages ssymptoms well but occasionally need support	Struggles with symptom severity	Bears with severe symptoms with willpower	Faces moderate symptom severity but is more focused on appearance	Could bear with symptoms with close clinician guidance
Specific need	mHealth to support chemotherapy-related symptom management, online clinician interaction	mHealth for monitoring by children, simple interface for basic needs	mHealth for children to track appointments and care schedules	mHealth for emotional support, patient-patient communication	mHealth with reliable, clinician-vetted information and consultation options
Design point	Design should focus on providing tips for chemotherapy-related symptom management and ensuring access to reliable professional information, with easy access to online clinician consultations.	Create a simple, user-friendly interface that allows children to monitor patient health and symptom management, ensuring notifications for appointments and medical updates.	Design tools that allow children to remotely manage patient schedules, appointments, and care, with minimal interaction from the patient. Focus on simplifying updates and reminders.	Design should emphasize patient-to-patient interaction for emotional support, with features to help rebuild self-esteem and manage appearance-related concerns.	Provide a trusted, clinician-validated platform with regular updates, symptom management tips, and easy access to clinician consultations.

Persona of Positive Manager

Quotes:*I think I can battle with breast cancer, and I will eventually win. Long days are there (Textbox 2).*

XSL•FO
RenderX

Textbox 2. Persona of positive manager.

Characteristics:

- Social interaction patterns: *I live with my caring spouse and supportive children, who play a crucial role in my daily life and health journey. I am deeply grateful for their presence and support, and I cherish every moment with them. Family and friends are my pillars of support. Regular gatherings and conversations keep me mentally engaged and emotionally fulfilled.*
- mHealth literacy: *Having access to mobile devices and the internet has been a game-changer. It empowers me to stay informed about my health and treatment options, which gives me a sense of control. I actively use mHealth apps to educate myself on managing chemotherapy-related symptoms. The ability to connect with healthcare professionals online reassures me and provides personalized advice. I completed high school and have always valued learning. I find comfort in online interactions with healthcare providers. Their expertise and reassurance help me navigate the challenges of treatment with greater confidence.*
- Symptom diversity and severity: *Dealing with various symptoms, such as digestive issues and occasional fatigue, can be challenging. However, learning how to manage them effectively has been empowering. While I generally manage my symptoms well, there are moments when I need extra support. Knowing I can reach out to my family or healthcare team for guidance makes a big difference.*

Specific preference: Specifically, I require tailored symptom tracking to monitor chemotherapy-related symptoms accurately. Personalized guidance is essential, offering tips on managing issues like gastrointestinal discomfort and fatigue. Regular updates and alerts keep me informed of symptom changes and treatment updates.

Persona of Dependent Parent

Quotes:*I don’t know what to do and have lost hope for the future. Please tell anything to my children (Textbox 3).*

Textbox 3. Persona of dependent parent.

Characteristics:

- Social interaction patterns: I live with my children, who provide around-the-clock care. We have a close-knit family dynamic, and I rely heavily on them for both emotional support and practical assistance. While I interact regularly with my family members, most of my social interactions revolve around family gatherings and caregiving responsibilities.
- mHealth literacy: I have limited experience with technology and rely on my children to navigate mHealth tools. They assist me in using healthcare apps for basic tasks like appointment scheduling and medication reminders. Understanding complex medical information online is challenging for me, so I prefer simplified explanations and direct guidance from my family.
- Symptom diversity and severity: I manage a mix of symptoms related to chemotherapy, including fatigue and occasional digestive issues. These symptoms can be overwhelming at times, impacting my daily activities and emotional well-being. While I strive to cope independently, I often rely on my children’s support to manage symptom flare-ups and seek medical advice when needed.

Specific preference: I require mHealth tools that simplify symptom monitoring and provide clear instructions for managing chemotherapy side effects. Specifically, I need a user-friendly app that allows my children to track my health status, appointments, and medication schedules effortlessly. Timely alerts and reminders are essential to ensure I adhere to treatment plans and receive prompt medical attention for any symptom changes. Access to reliable information and simplified health advice through mHealth would greatly enhance my ability to manage symptoms and maintain my overall well-being.

Persona of Management Isolationist

Quotes:*I am so severely ill. I don’t want to say anything (Textbox 4).*

Textbox 4. Persona of management isolationist.

Characteristics:

- Social interaction patterns: I prefer solitude and spend most of my time at home, away from social interactions. Living alone with minimal external contact, I find comfort in solitude and independence, handling my health challenges mostly on my own.
- mHealth literacy: I don’t use mobile devices or the internet regularly, which limits my access to mHealth tools. I rely solely on face-to-face interactions with healthcare providers or family members for medical information and support.
- Symptom diversity and severity: Managing chronic pain and discomfort is a daily struggle for me. These symptoms significantly impact my mobility and quality of life, requiring ongoing management strategies to cope independently.

Specific preference: I require healthcare solutions that respect my preference for isolation while providing essential medical support. An ideal mHealth tool would allow minimal digital interaction but provide critical health updates and remote consultations with healthcare professionals as needed. This approach would enable me to maintain my independence while ensuring timely medical guidance and support.

XSL•FO
RenderX

Persona of Image Manager

Quotes:*I haven't retired. I have to go back to work with a normal body shape (Textbox 5).*

Textbox 5. Persona of image manager.

Characteristics:

- Social interaction patterns: I actively seek interaction with online patient support groups and communities. These platforms provide me with valuable emotional support and information-sharing opportunities, helping me navigate the challenges of chemotherapy with peers who understand my experiences.
- mHealth literacy: I use mHealth apps to connect with patient communities and access reliable health information. I am comfortable navigating online resources and value the empowerment that comes from educating myself about my health and treatment options.
- Symptom diversity and severity: While managing chemotherapy symptoms, I focus on maintaining my appearance and emotional well-being. I deal with moderate symptom severity, such as fatigue and appearance-related concerns, which impact my self-esteem.

Specific preference: I need mHealth tools that facilitate patient-to-patient interactions and provide emotional support. Specifically, I seek platforms that foster community engagement and offer resources for managing appearance-related symptoms. Access to reliable information and peer support online is crucial for enhancing my resilience and emotional health during treatment.

Persona of Clinician-Dependent

Quotes:

I only rely on health care providers, and I want to contact doctors when I need (Textbox 6)

Textbox 6. Persona of clinician-dependent.

Characteristics:

- Social interaction patterns: I maintain regular interactions with healthcare providers, relying heavily on their expertise and guidance in managing my health. I prioritize direct communication with doctors and healthcare professionals for medical advice and treatment decisions.
- mHealth literacy: While I have access to mobile devices, I primarily use them to facilitate communication with healthcare providers. I prefer direct contact with doctors for accurate medical information rather than relying on digital health tools or online resources.
- Symptom diversity and severity: Managing multiple symptoms related to chemotherapy is a significant concern for me. I depend on clinicians to monitor my symptoms closely and provide tailored treatment plans that address my specific health needs.

Specific preference:

My specific need is a reliable mHealth platform that facilitates direct and timely communication with healthcare providers. I prefer tools that offer secure messaging or virtual consultations with doctors, ensuring I can reach out for medical advice promptly when needed. This direct access to clinicians is crucial for me in managing chemotherapy-related symptoms effectively and receiving personalized care.

Discussion

Overview

This study depicted the differences in symptom management behaviors and preference for using mHealth to manage symptoms among the older adult breast cancer patients in China and, based on the differences, built 5 typical personas: Positive Manager, Dependent Parent, Management Isolationist, Image Manager, and Clinician-Dependent. This study outlines the complete process of persona development, including qualitative data collection, persona segmentation, and creation, building upon previous reviews of persona technologies [28]. The strict adherence to the consensus process [28] and early user involvement [29] suggests that our findings may serve as a valuable reference for future persona development and user-centered design.

The qualitative interviews we conducted focused on our target and special population—older adult breast cancer patients. The combination of cancer and aging renders this population particularly vulnerable, as cancer impacts their physical health while aging contributes to increased comorbidities and a decline in overall resilience [30]. Despite previous studies that have primarily focused on the treatment and survival of this population, few have closely examined their self-management behaviors or actively engaged them in the information age to ensure they fully benefit from modern technologies [31]. Notably, we found only 1 previous paper that recognized the need for health technology in chronic disease management [21], and no existing literature specifically addresses the preferences of older adult breast cancer patients. Therefore, our study identifies 3 key categories—interaction patterns, mHealth literacy, and symptom diversity and severity—that revealed the self-management behavior patterns and preference of mHealth for symptom management of older adult breast cancer patients. This research addresses a critical gap in the literature, highlighting the importance of understanding these factors to enhance mHealth solutions for this vulnerable population.

Following this, we developed 5 personas based on the categories and subcategories identified. This approach adheres to standard persona construction practices by using data-driven classification [28]. Moreover, the personas expand traditional persona content to include social dynamics, mHealth access, mHealth usage, mHealth interest, symptom diversity, symptom severity, specific needs, and design points to enhance richness in understanding older adult breast cancer patients and their context [32].

XSL•FO
RenderX

According to previous studies, the enriched insights gained provide valuable guidance for functional requirements, design, and implementation strategies tailored specifically for this target population [21]. The structure and content of our personas, which incorporate the key features of symptom management, can serve as templates for developing personas for other health consumer target populations [21].

We believe that the personas developed can effectively highlight both common and distinct traits among users. The shared characteristics of the 5 personas should serve as the primary focus for core functionalities, while the unique traits will inform the development of supplementary and optional features. From our research findings, we have initially identified 2 core functionalities: symptom reporting and educational materials.

The symptom self-reporting feature will enable users to log their symptoms daily, facilitating real-time monitoring and proactive management of potential side effects, which is crucial for breast cancer management for older adult patients [33]. For instance, for the Positive Manager, visualizing longitudinal symptom assessments could enhance their understanding of health trends over time. In the case of Dependent Parents, the mHealth product could facilitate a dyad approach where patients and their children both engage with the system, allowing symptom reports to be sent directly to their children via push notifications.

Meanwhile, the educational materials will provide tailored information to enhance users' understanding of their condition and treatment, empowering them to make informed decisions regarding their health. For the Clinician-Dependent, it is crucial to integrate a voice icon and profile icon, enabling medical advice to be easily accessible and conveyed by a trusted physician. Regarding the educational features of the mHealth product, Positive Manager and Imanage Manager are inclined to learn and manage their health and could benefit from self-directed reading materials and a patient support network, while Clinician Dependent, who prefers individualized attention from health care providers, could be offered one-on-one video appointments. Given that the target population consists of older adult patients who often experience memory and comprehension difficulties, the development team must carefully manage the amount of information presented to avoid cognitive overload. A modular approach can effectively cater to users, providing essential knowledge for those who prefer not to delve deeper while allowing for broader exploration through optional modules. Active information seekers, such as the Positive Manager and Image Manager, can benefit from this modular design, which ensures continuous content updates and relevance.

As for the role of personas in the design process, they serve as educational tools for design and usability teams. They provide insights that convey project progress to sponsors and offer a "word picture" that familiarizes team members with target user groups. Personas structure user typification, enhancing communication within development teams. As noted by Hill and Barteck [34], once established, personas become the primary medium for design discussions, aiding in resolving group conflicts by referencing user perspectives. For instance, team members might ask, "What would the Clinician-Dependent think?" while designing the interface and algorithm. Additionally, based on interaction design principles like the double-diamond design, the iterative process often involves usability testing [35]. Personas can help predict who the primary and secondary users will be. For breast cancer patients, spouses and children are typically considered secondary users, as they can facilitate adoption and usage and are often willing to assist primary users in receiving better care. Therefore, detailed personas can serve as a foundation for selecting and screening participants for interactive prototyping and usability testing, directly influencing the solution's design and development [21]. For instance, when considering Dependent Parents, it is crucial to involve their children as secondary users of the mHealth tool, as their experiences and insights are equally valuable.

During the implementation stage, personas serve as essential tools to guide the deployment of mHealth solutions. For the Positive Manager, promotional efforts such as posters and advertisements can effectively reach this group. For the Image Manager, word-of-mouth within social circles is a key avenue for adoption. The Clinician Dependent persona can be introduced to mHealth solutions through doctors or nurses during appointments, making it crucial to integrate mHealth into the clinical workflow. For the Dependent Parent, the focus can be on caregivers who accompany them to appointments, using this relationship as a key promotional touchpoint. Additionally, for long-term adherence, especially with Dependent Parents, their children should be considered the primary channel to improve compliance and ensure ongoing engagement with the mHealth tool.

According to previous studies, personas do not need to be recreated from scratch for every application [36]. Our 5 specified personas—The Positive Manager, The Dependent Parent, The Management Isolationist, The Image Manager, and The Clinician Dependent—can be reused by both researchers and practitioners. With China's rapidly aging population and increasing prevalence of cancer, innovative approaches to patient management are urgently needed [37]. These personas offer valuable insights that can address the current gap in China's user-centered design practice, particularly in mHealth solutions tailored to older adult breast cancer patients. Their reusability makes them adaptable for various health care applications, enhancing design efficiency and consistency across future projects.

Limitations

Our study presents some limitations that should be acknowledged. First, we relied on interviews for data collection, without using focus groups. While one-to-one interviews provided in-depth insights into individual experiences, focus groups could have uncovered collective perspectives and group dynamics, particularly valuable when studying older adult populations. Second, the personas we developed are based on qualitative data, which, while rich in detail, are inherently subjective. The personas were constructed by researchers based on interpretations of the qualitative data, which may introduce some bias. Quantitative approaches to persona development, such as surveys or statistical analysis, could offer more

XSL•FO
RenderX

objective, testable profiles that may provide additional validity and generalizability to the findings [38].

Conclusions

This study identified symptom management behaviors and preferences for managing symptoms using mHealth among older adult Chinese breast cancer patients. Further, we developed 5 personas that reflect the diverse preferences of this population. These personas play a crucial role across multiple stages of mHealth development: helping to determine core functionalities during the planning phase, guiding design and development by aligning product features with user needs, and supporting implementation by facilitating adoption and engagement strategies. Together, these insights contribute to more effective, user-centered mHealth solutions.

Acknowledgments

This study was supported by the National Natural Science Foundation of China [grant number 72104055]. This fund played a role in data collection.

Authors' Contributions

DL contributed to writing – original draft preparation and visualization; YZ involved in investigation and formal analysis; CY involved in conceptualization and data curation; writing – review & editing; QH involved in methodology; YY contributed to resources; ZG contributed to methodology; XY contributed to resources; FW contributed to conceptualization; funding acquisition, project administration and formal analysis.

Conflicts of Interest

None declared.

References

1. Wild CP, Weiderpass E, Stewart BW. World Cancer Report 2020: Cancer Research for Cancer Prevention: International Agency for Research on Cancer; 2020. URL: http://publications.iarc.fr/586 [accessed 2025-08-19]
2. Fan L, Strasser-Weippl K, Li JJ, et al. Breast cancer in China. Lancet Oncol 2014 Jun;15(7):e279-e289. [doi: 10.1016/S1470-2045(13)70567-9] [Medline: 24872111]
3. Tesarova P. Breast cancer in the elderly—should it be treated differently? Reports of Practical Oncology & Radiotherapy 2013 Jan;18(1):26-33. [doi: 10.1016/j.rpor.2012.05.005]
4. Wong RT, Cafferky BM, Alejandro JP. Chronic disease and elder mistreatment: a meta-analysis. Int J Geriatr Psychiatry 2022 Jan;37(1). [doi: 10.1002/gps.5640] [Medline: 34633703]
5. Vermunt N, Harmsen M, Westert GP, Olde Rikkert MGM, Faber MJ. Collaborative goal setting with elderly patients with chronic disease or multimorbidity: a systematic review. BMC Geriatr 2017 Jul 31;17(1):167. [doi: 10.1186/s12877-017-0534-0B2] [Medline: 28760149]
6. Ryu S. Book Review: mHealth: new horizons for health through mobile technologies: based on the findings of the Second Global Survey on eHealth (Global Observatory for eHealth Series, Volume 3). Healthc Inform Res 2012;18(3):231. [doi: 10.4258/hir.2012.18.3.231B4] [Medline: 22509475]
7. Rowland SP, Fitzgerald JE, Holme T, Powell J, McGregor A. What is the clinical value of mHealth for patients? NPJ Digit Med 2020;3(1):4. [doi: 10.1038/s41746-019-0206-xB1] [Medline: 31970289]
8. Aguiar M, Trujillo M, Chaves D, Álvarez R, Epelde G. mHealth apps using behavior change techniques to self-report data: systematic review. JMIR Mhealth Uhealth 2022 Sep 9;10(9):e33247. [doi: 10.2196/33247B2] [Medline: 36083606]
9. Haldane V, Koh JJK, Srivastava A. User preferences and persona design for an mHealth intervention to support adherence to cardiovascular disease medication in Singapore: a multi-method study. JMIR Mhealth Uhealth 2019 May 28;7(5):e10465. [doi: 10.2196/10465B2] [Medline: 31140445]
10. Harder H, Holroyd P, Burkinshaw L. A user-centred approach to developing bWell, a mobile app for arm and shoulder exercises after breast cancer treatment. J Cancer Surviv 2017 Dec;11(6):732-742. [doi: 10.1007/s11764-017-0630-3B2] [Medline: 28741202]
11. Lozano-Lozano M, Melguizo-Rodríguez L, Fernández-Lao C. Association between the use of a mobile health strategy app and biological changes in breast cancer survivors: prospective pre-post study. J Med Internet Res 2019 Aug 14;21(8):e15062. [doi: 10.2196/15062B2] [Medline: 31414667]
12. Min YH, Lee JW, Shin YW. Daily collection of self-reporting sleep disturbance data via a smartphone app in breast cancer patients receiving chemotherapy: a feasibility study. J Med Internet Res 2014 May 23;16(5):e135. [doi: 10.2196/jmir.3421B2] [Medline: 24860070]
13. Gehrke A, Lee SS, Hilton K. Development of the cancer survivor profile-breast cancer (CSPro-BC) app: patient and nurse perspectives on a new navigation tool. J Cancer Surviv 2018 Jun;12(3):291-305. [doi: 10.1007/s11764-017-0668-2] [Medline: 29524014]

XSL•FO
RenderX

14. Saevarsdottir SR, Gudmundsdottir SL. Mobile apps and quality of life in patients with breast cancer and survivors: systematic literature review. J Med Internet Res 2023 Jul 26;25:e42852. [doi: 10.2196/42852B2] [Medline: 37494111]
15. Forman DE, LaFond K, Panch T, Allsup K, Manning K, Sattelmair J. Utility and efficacy of a smartphone application to enhance the learning and behavior goals of traditional cardiac rehabilitation: a feasibility study. J Cardiopulm Rehabil Prev 2014;34(5):327-334. [doi: 10.1097/HCR.0000000000000058B3] [Medline: 24866355]
16. Steinman L, Heang H, van Pelt M. Facilitators and barriers to chronic disease self-management and mobile health interventions for people living with diabetes and hypertension in Cambodia: qualitative study. JMIR Mhealth Uhealth 2020 Apr 24;8(4):e13536. [doi: 10.2196/13536B2] [Medline: 32329737]
17. Palas JU, Sorwar G, Hoque MR, Sivabalan A. Factors influencing the elderly's adoption of mHealth: an empirical study using extended UTAUT2 model. BMC Med Inform Decis Mak 2022 Dec;22(1):191. [doi: 10.1186/s12911-022-01917-3] [Medline: 35871682]
18. Chen YRR, Schulz PJ. The effect of information communication technology interventions on reducing social isolation in the elderly: a systematic review. J Med Internet Res 2016 Jan 28;18(1):e18. [doi: 10.2196/jmir.4596] [Medline: 26822073]
19. Alam MZ, Khanam L. Comparison of the young aged and elderly female users' adoption of mHealth services. Health Care Women Int 2022;43(10-11):1259-1283. [doi: 10.1080/07399332.2022.2039149] [Medline: 35195507]
20. Kampmeijer R, Pavlova M, Tambor M, Golinowska S, Groot W. The use of e-health and m-health tools in health promotion and primary prevention among older adults: a systematic literature review. BMC Health Serv Res 2016 Sep 5;16 Suppl 5(Suppl 5):290. [doi: 10.1186/s12913-016-1522-3] [Medline: 27608677]
21. LeRouge C, Ma J, Sneha S, Tolle K. User profiles and personas in the design and development of consumer health technologies. Int J Med Inform 2013 Nov;82(11):e251-e268. [doi: 10.1016/j.ijmedinf.2011.03.006] [Medline: 21481635]
22. Cooper A. The Inmates Are Running the Asylum: Springer; 1999. URL: https://ptgmedia.pearsoncmg.com/images/9780672326141/samplepages/0672326140.pdf [accessed 2025-08-19]
23. Liu Y, Encelle B, Sehaba K. A user-centered approach to design a mobile application for chronic pain management. MMC_C 2020 Dec 31;81(1-4):24-29 [FREE Full text] [doi: 10.18280/mmc_c.811-405]
24. Bartels SL, Taygar AS, Johnsson SI. Using personas in the development of eHealth interventions for chronic pain: a scoping review and narrative synthesis. Internet Interv 2023 Apr;32:100619. [doi: 10.1016/j.invent.2023.100619B2] [Medline: 37273935]
25. Ledel Solem IK, Varsi C, Eide H. A user-centered approach to an evidence-based electronic health pain management intervention for people with chronic pain: design and development of EPIO. J Med Internet Res 2020 Jan 21;22(1):e15889. [doi: 10.2196/15889] [Medline: 31961331]
26. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res 2016 Nov;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
27. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
28. Alsaadi B, Alahmadi D. The use of persona towards human-centered design in health field: review of types and technologies. In: Alsaadi B, Alahmadi D, editors. Presented at: 2021 International Conference on e-Health and Bioengineering (EHB); Iasi, Romania. [doi: 10.1109/EHB52898.2021.9657744]
29. Eisma R, Dickinson A, Goodman J, Syme A, Tiwari L, Newell AF. Early user involvement in the development of information technology-related products for older people. Universal Access in the Information Society 2004 Jun 1;3(2):131-140. [doi: 10.1007/s10209-004-0092-z]
30. Mohile SG, Xian Y, Dale W, et al. Association of a cancer diagnosis with vulnerability and frailty in older medicare beneficiaries. JNCI Journal of the National Cancer Institute 2009 Sep 2;101(17):1206-1215. [doi: 10.1093/jnci/djp239]
31. Litvak DA, Arora R. Treatment of elderly breast cancer patients in a community hospital setting. Arch Surg 2006 Oct;141(10):985-990. [doi: 10.1001/archsurg.141.10.985] [Medline: 17043276]
32. Kwon JY, Moynihan M, Lau F. Seeing the person before the numbers: personas for understanding patients' life stories when using patient-reported outcome measures in practice settings. Int J Med Inform 2023 Apr;172:105016. [doi: 10.1016/j.ijmedinf.2023.105016] [Medline: 36758303]
33. Fusco D, Allocca E, Villani ER, Franza L, Laudisio A, Colloca G. An update in breast cancer management for elderly patients. Transl Cancer Res 2018 Apr;7(S3):S319-S328. [doi: 10.21037/tcr.2018.03.21]
34. Hill V, Bartek V. Telling the user's story. 2007 Mar 30 Presented at: Proceedings of the 2007 Symposium on Computer Human Interaction for the Management of Information Technology p. 6. [doi: 10.1145/1234772.1234794]
35. Gustafsson D. Analysing the double diamond design process through research & implementation. 2019 URL: https://aaltodoc.aalto.fi/items/3a6b3fa8-14a4-4b39-809f-cf31ab66c17c [accessed 2025-08-19]
36. Pruitt J, Adlin T. The Persona Lifecycle: Keeping People in Mind Throughout Product Design: Elsevier; 2010.
37. Ju W, Zheng R, Zhang S. Cancer statistics in Chinese older people, 2022: current burden, time trends, and comparisons with the US, Japan, and the Republic of Korea. Sci China Life Sci 2023 May;66(5):1079-1091. [doi: 10.1007/s11427-022-2218-x] [Medline: 36543994]
38. Mulder S, Yaar Z. The User Is Always Right: A Practical Guide to Creating and Using Personas for the Web: New Riders 2006.

XSL•FO
RenderX

Abbreviations

mHealth: mobile health

Edited by A Kushniruk; submitted 19.01.25; peer-reviewed by J Chen, O Abasilim; revised version received 15.06.25; accepted 10.07.25; published 26.08.25.

Please cite as:
Li D, Zhu Y, Yuan C, Huang Q, Yang Y, Geng Z, Yuan X, Wu F
Symptom Management Preference and Persona Development for Mobile Health Design Targeting Chinese Older Adult Patients With Breast Cancer: Descriptive Qualitative Study
JMIR Hum Factors 2025;12:e71448
URL: https://humanfactors.jmir.org/2025/1/e71448
doi:10.2196/71448

© Danyu Li, Yanlin Zhu, Changrong Yuan, Qingmei Huang, Yang Yang, Zhaohui Geng, Xiaotong Yuan, Fulei Wu. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 26.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring Clinician Experiences With a Digital Platform Supporting Orthopedic Care That Integrates Patient-Generated Health Data: Qualitative Study of Early Users

Dmitri S Katz[1], PhD; Daniel Gooch[1], PhD; Linda Price[1], PhD; Irum Rauf[1], PhD; Oliver Pearce[1,2], MBBS; Blaine Price[1], MSc

[1]School of Computing and Communications, The Open University, Walton Hall, Milton Keynes, United Kingdom
[2]Trauma and Orthopaedics, Milton Keynes University Hospital, Milton Keynes, United Kingdom

Corresponding Author:
Daniel Gooch, PhD
School of Computing and Communications, The Open University, Walton Hall, Milton Keynes, United Kingdom

Abstract

Background: Digital care platforms that integrate patient-generated health data (PGHD) alongside education and communication tools have been recognized as potential instruments in transforming health care from clinician-centered to a more patient-centered approach. This transformation is driven by the potential of PGHD to provide deeper insights into patients' conditions, facilitate personalized care, improve patient quality of life, reduce inefficiencies in data collection, and empower patients. Yet, actual implementation within clinical settings is still at early stages; therefore, impacts on clinical care remain limited.

Objective: This study sought to explore the benefits, challenges, and opportunities of integrating PGHD into orthopedic care by analyzing the reflections of early adopter surgeons and physiotherapists who have used a digital care management platform.

Methods: This qualitative study used thematic analysis of interviews conducted with surgeons and physiotherapists (n=9) from a clinical unit that was among the first to trial "mymobility," an industry-produced software platform (Zimmer Biomet). The participants were recruited using snowball sampling, and interviews were conducted from June to July 2022. The interviews focused on work practices, use of digital tools, experiences with PGHD, and experiences with the mymobility software. Thematic analysis was conducted using NVivo software (QSR International Pty Ltd), focusing on identifying key themes and insights.

Results: The study identified several benefits of integrating PGHD into orthopedic care, including improved patient education, enhanced communication and assessment, and increased patient motivation and adherence. However, several challenges were also noted, such as increased clinician workload, questionable data utility, lack of patient centricity, and inability to tailor software to clinical contexts. Suggested opportunities included improving dashboard design, personalizing physiotherapy, and using collected data for improving clinical care.

Conclusions: The integration of PGHD into orthopedic care shows promise, largely in areas suggested by the literature. However, significant challenges remain. Future research should focus on addressing solvable challenges, such as improving software user interface design and functionality, while embracing the possibility that some challenges lack clear solutions and will likely require careful balancing of design tensions. The findings highlight the need for ongoing development and refinement of PGHD-inclusive systems to better support clinical practice and patient outcomes.

(JMIR Hum Factors 2025;12:e65216) doi:10.2196/65216

KEYWORDS

patient-generated health data; PGHD; orthopedic care; digital health; remote monitoring; patient-centered care; clinician workload; patient engagement; health informatics; orthopedics; quality of life; physiotherapists; surgeons

Introduction

Overview

There are frequent calls within academic literature and popular media to transform health care systems from the traditional "clinician-centered" model to a more "patient-centered" approach, an approach characterized by shared clinical decision-making and increased patient autonomy [1] while also helping to reduce costs through coordinated and accessible personalized services [2]. An important and often asserted mechanism for this transition is increasing the use of patient-generated health data (PGHD) within clinical care. PGHD are information collected by individuals (and informal caregivers) on aspects such as their treatment records, symptoms, biometric data, and lifestyle preferences [3]. The use of PGHD has been promoted for various reasons: providing deeper insight into a patient's condition [4], facilitating

XSL•FO
RenderX

personalization of care [5], improving patient quality of life [6], reducing inefficiencies and insufficiencies in data collected solely at clinical visits [7], reducing retrospective recall bias [8], and empowering patients by helping them to understand their progress [9].

Using PGHD has become increasingly feasible, as technologies for self-tracking of health and wellness data have become affordable and culturally acceptable to consumers [10]. The ease of data collection for patients and ubiquity of these devices are informally influencing clinical interactions, as patients increasingly present their self-collected data during clinical visits [5], and recent studies suggest that clinicians are increasingly open to using PGHD at least as supplementary data within clinical workflows [11] [4].

Despite the potential benefits, there remain extensive known barriers to implementing PGHD within clinical care. For example, West et al [5] identified the nature of the data (structure, completeness, reliability, lack of context, and relevance), selective disclosure, clinician suspicions of underlying patient psychiatric disorders, insufficient clinician time, insufficient clinician expertise, clinician overload, interoperability, and negative impacts to workflow. Other notable barriers include clinician reluctance to assume responsibility for motivating patients to collect data [4], clinician concerns that reliance on PGHD could exacerbate health disparities [6], negative impacts on patient-clinician relationships [12], and resistance to change within clinical culture [7]. Given such diverse challenges, it is perhaps unsurprising that formal integration of PGHD within clinical practice remains limited [13]. Therefore, there remains a lack of knowledge on diverse PGHD-related topics such as how to use PGHD to improve health care outcomes [3], impact on patient-clinician communication within surgical domains [12], best practices to incorporate into clinical workflows [14], and impacts in supporting clinical practice [11].

This study sought to understand work practices, acceptance of digital tools, and attitudes toward PGHD as part of a broader care management platform to gain insights that could guide the design of digital systems that incorporate PGHD alongside other patient support functionalities in orthopedic care. To this end, we conducted interviews with surgeons and physiotherapists who were members of an “early adopter” orthopedic unit that had (1) been given the opportunity to use an early release of industry-produced patient journey software that offered various tools for interacting with patients and their data and (2) participated over several years in the development and deployment of a novel device to enable orthopedic patients to self-collect pain data [15]. Focusing on this context foregrounds real-world exposure and engagement with PGHD, with the reported experiences helping to fill current gaps in knowledge in the clinical use of PGHD, guiding future research on the use of these data in a clinical context.

Background

Total knee arthroplasty (TKA) and total hip arthroplasty (THA) are common elective surgeries for older adults, where an original joint damaged by arthritis or trauma is replaced with metal and plastic implants. Due to high cumulative expenses resulting from this increasingly common procedure, TKA or THA operation is an important target for optimizing care [16]. To this end, fast-track programs that seek to reduce patient time in the hospital have become standard, successfully reducing costly hospital stays without impairing outcomes or increasing readmissions [17]. However, despite the successes of this approach, there remain opportunities for improvement that could potentially be supported with increased use of PGHD. For example, patients who have had a TKA or THA operation frequently have risk factors that can be related to readmission (eg, obesity, cardiovascular disease, diabetes, and lifestyle factors) [16]. Yet, current clinical systems often fail to consider what occurs outside the boundaries of the health care system [18]. This lack of vital information about patients’ comorbidities and lifestyle can hinder timely decisions about surgical scheduling and personalizing patient care, and there is also an urgent need to increase proactive care to enhance patient satisfaction, promote long-term recovery, and reduce unnecessary hospital stays. Given the importance of events occurring outside of the clinical setting for improving outcomes, the importance of engaging patients in collaborative care, and the potential for nonclinical data to inform policies that could improve outcomes, TKA or THA appears as a promising domain for increased PGHD integration.

PGHD and THA or TKA

The adoption and affordability of wearable devices have increased the feasibility of integrating PGHD into THA or TKA care, and there have been numerous studies exploring potential benefits and limitations. Toogood et al [19] used ankle-worn Fitbit step counters to track daily steps in patients who have had their THA operation. They found this method feasible and found that by differentiating activity levels based on factors like age, BMI, surgical approach, and discharge destination, it might be possible to identify slower recovery cases that could benefit from interventions. However, other studies cautioned that while tracking was technically feasible, the resulting data had practical limitations. For example, Crizer et al [20] studied using step counts to monitor preoperative and postoperative recovery but found a weak correlation with patient-reported outcomes. The authors suggested that combining objective measures like step counts with patient-reported outcomes would provide a more comprehensive assessment. Similarly, Bendich et al [21] examined data from wearable sensors worn by patients both before and after total joint surgery. They reported that alterations in sensor data 6 weeks after surgery correlated closely with patients’ reported improvements in well-being; however, the raw sensor data did not directly correspond with patients’ self-reported feelings, suggesting that to effectively predict postsurgery outcomes using sensors, it is essential to analyze how the data evolve over time rather than solely focusing on absolute values. These papers highlight that it is not sufficient to aggregate and report data to clinicians, but rather that meaningful interpretation and context are essential. We found numerous papers that applied tracking devices to support patient monitoring outside the clinical environment. However, we found little research on the reflections of orthopedic clinicians on the use of PGHD and more specifically their experiences with PGHD-inclusive clinical software. In addition, relatively few

XSL•FO
RenderX

studies on PGHD integration have examined the use of patient-generated data from the clinician's perspective rather than the patient's perspective [5], which is important, as clinician engagement has previously been recognized as a significant challenge for integrating patient self-reporting systems into clinical practice [7]. This paper aims to fill these gaps through exploring clinicians' real-world experiences with the benefits and challenges of integrating PGHD into TKA or THA workflows and through analysis identifying opportunities to improve these systems.

Methods

Ethical Considerations

Ethics clearance was granted through the Open University Human Research Ethics board (HREC/3776/Mehta). Preconditions for participation were professional capability in either of the 2 selected clinical roles, experience with TKA or THA, and spoken English. All participants were employed in these roles in the United Kingdom. Participants were assured anonymity. Prior to the interview, participants received an email confirmation of the time and date, as well as receiving a copy of the questions and a consent form that was returned by email. Participation was voluntary, and all participants were informed of their right to withdraw at any time. The interviews lasted from 30 to 60 minutes and were conducted by the same member of the research team (LP). Participants were offered either face-to-face or digital interviews to accommodate their clinical schedules. The interviewees' responses were anonymized according to the ethics committee procedure. Each institution's data protection policy was adhered to, ensuring that data were confidentially stored. All transcripts were transcribed and stored anonymously, and the data were stored on a secure system only accessible to the research team. Interviewees received approximately US $34 honorarium for taking part in the interviews.

Study Design

This study was designed as a formative qualitative investigation to surface early insights from clinicians engaging with PGHD-inclusive digital tools. It was not intended to quantify impact or compare outcomes but rather to inform future system design and identify areas warranting further evaluation through quantitative or mixed methods research.

We chose to interview 2 essential and complementary clinical roles in the TKA or THA domain: orthopedic surgeons who are central to the procedure [22] and physiotherapists whose work has been shown to improve postsurgical recovery [23]. These 2 clinical roles are directly involved in interpreting and responding to PGHD within the mymobility platform (Zimmer Biomet). Surgeons reviewed data to monitor recovery progress and inform follow-up care, while physiotherapists engaged with exercise adherence and symptom reports to adjust rehabilitation plans. Their combined perspectives were essential for understanding how PGHD are integrated into orthopedic workflows. Participants were selected using purposive sampling to capture a range of clinical and implementation perspectives from early adopters of the platform. Recruitment occurred during the COVID-19 pandemic, which introduced additional constraints such as staff unavailability, clinical backlogs, and limited access to nonurgent personnel. These conditions restricted the expansion of the sample. Our team was composed of human-computer interaction professionals (an interdisciplinary field that studies how people interact with digital technologies, with the goal of designing systems that are usable, effective, and aligned with users' needs) and clinicians, a pairing recommended by Cheng et al [24]. Our team also included a change management specialist with experience in digital system implementation. This team member supported interview design and interpretation by contributing expertise on organizational readiness, clinician engagement, and adoption patterns as well as conducting interviews. This team member's perspective helped to frame questions around perceived barriers and enablers in relation to system-wide change processes. Our research was also guided by a sociotechnical systems design (STSD) approach, which seeks to balance human, social, organizational, and technical requirements. This perspective recognizes that digital systems—particularly those introduced into complex clinical environments—often fail when they overlook the interdependent and often hidden dynamics of the social and organizational contexts in which they are deployed [25]. The STSD perspective was used as an interpretive lens for understanding clinician experiences and guiding the interpretation of clinician reflections in relation to their work systems, roles, and digital interactions. It was not applied as a formal evaluative framework, and no inferential claims are made. Given the qualitative nature and small sample size of this study, STSD served to sensitize analysis to the dynamic interplay between human, technical, and organizational factors in digital orthopedic care. Our research group has been collaborating with clinicians within this group since 2016 on the use of tracking devices to assess compliance [26] as well as research and design of tangible devices for patient pain logging [15,27,28]. In addition, members of this unit are participating in early testing of "mymobility," an industry-produced orthopedic care software platform that integrates diverse PGHD from Apple Watch and iPhone into clinical dashboards. These interviews were conducted from June to July 2022, a period when, due to the pandemic, many typically in-person services had been replaced with app-based communication, which by necessity gave clinicians additional exposure to using health-relevant data collected by patients. Drawing on our ongoing relationship with the orthopedic center at Milton Keynes University Hospital, we used snowball sampling to recruit 4 orthopedic surgeons (1 female and 3 male) and 5 physiotherapists (4 female and 1 male; Table 1). All had previous experience with TKA or THA procedures, with experience ranging from 1.5 to 25 years. Clinicians had varying levels of exposure working with PGHD at Milton Keynes University Hospital. While we did not track or analyze the number of patients whose PGHD were reviewed by each participant, all interviewees were directly involved in the early use of the digital platform in clinical care. Participants routinely engaged with patients contributing PGHD as part of their roles, with varying degrees of exposure depending on their clinical responsibilities. The goal of this study was to explore clinicians' perspectives on PGHD integration, not to quantify its use across

XSL•FO
RenderX

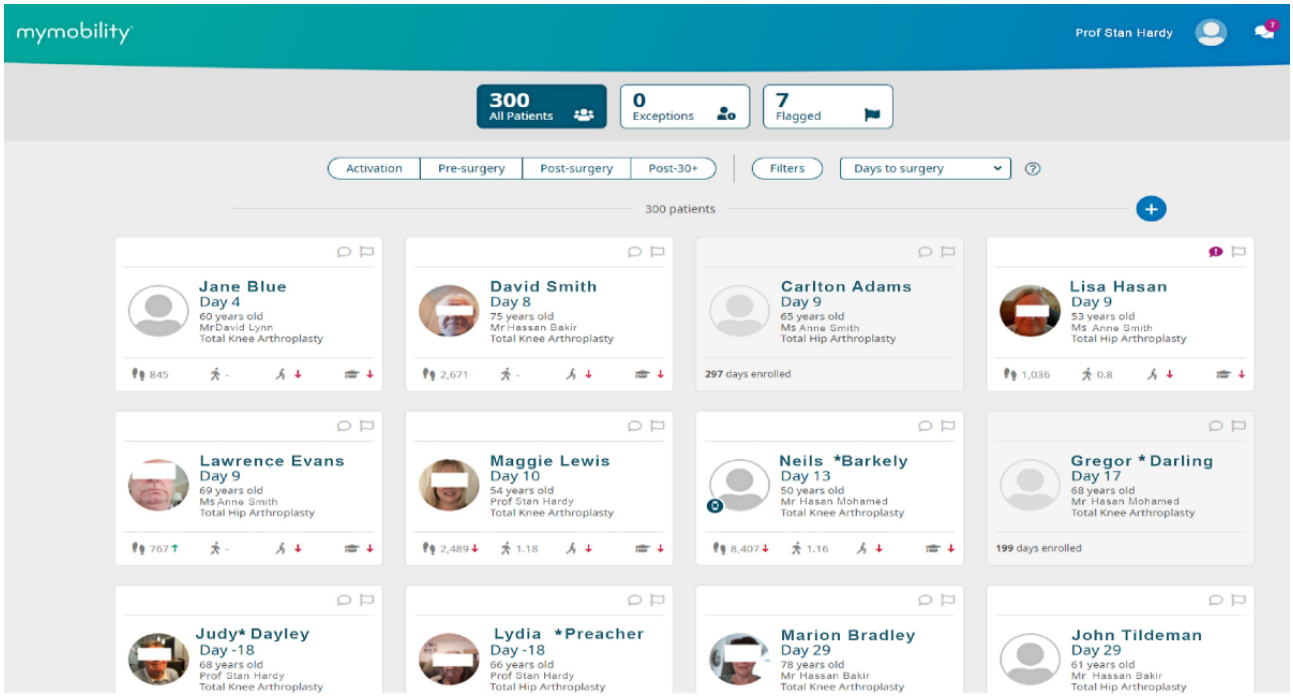
patient populations. In the reporting on interviews, S1-4 will be used for surgeons and P1-5 for physiotherapists.

Table . Demographics of participants.

Participant	Age (years)	Sex	Ethnicity	Role	Experience (years)
Surgeon 1	56	Male	Asian	Associated specialist	25
Surgeon 2	46	Male	Asian	Consultant orthopedic surgeon	25
Surgeon 3	45	Male	Asian	Trauma and orthopedic consultant, hip and knee	25
Surgeon 4	53	Female	White	Consultant orthopedic and trauma surgeon	15
Physiotherapist 1	25	Female	White	Rotational physiotherapist	1.5
Physiotherapist 2	56	Female	White	Senior team lead physio in trauma and orthopedics	32
Physiotherapist 3	46	Female	White	Physiotherapist on trauma and orthopedics	12
Physiotherapist 4	29	Male	White	Musculoskeletal and orthopedic outreach physiotherapist	4
Physiotherapist 5	32	Female	White	Specialist musculoskeletal physiotherapist	3

Questions were based on our literature review and previous experience researching the requirements and design of digital systems. The interview questions were piloted and iterated with an advising senior orthopedic surgeon and a physiotherapist. The questions for the semistructured interviews were structured within four sections: (1) current work practices, (2) use of digital tools, (3) experiences with PGHD, and (4) experiences with the mymobility software platform. For the last set of questions, participants were provided with anonymized screenshots of PGHD from the mymobility app for their reference (Figures 1 and 2).

Figure 1. Patient overview.



XSL•FO
RenderX

Figure 2. Patient progress.

XSL•FO
RenderX

mymobility Software

mymobility [29] is a digital care management platform tailored for orthopedic patients, supporting a variety of procedures through care plans for knee, hip, shoulder, and spine surgeries. This platform seeks to facilitate communication and data sharing between patients and health care providers throughout the surgical process. Key features include:

- Patient engagement and compliance: The app seeks to foster patients' engagement in their surgical journey by providing educational materials, procedural information, and personalized exercise plans.
- Data monitoring and insights: Through automated remote monitoring with Apple Watch and iPhone and patient-reported outcomes, the app tracks patient progress, allowing patients' and clinicians' dashboards to view collected data.
- Care team management: Through a unified platform for digital care management and telemedicine, the app seeks to support communication and coordination among patients and care team members.

Data Analysis

The anonymized transcribed interviews were analyzed using reflexive thematic analysis (RTA), a qualitative approach that emphasizes the researcher's active role in constructing meaning and identifying patterns across data [30,31]. RTA was selected for its suitability in capturing subjective clinician perspectives without imposing a fixed coding frame, aligning with our study's exploratory and practice-based context. We adopted a realist stance—seeking to report participants' experiences and meanings—while using an inductive, data-driven approach without applying a pre-existing theory or model. The analysis followed Braun and Clarke's [30,31] 6-phase process: familiarization, coding, theme development, theme review, and final definition and naming.

Although the sample was limited (n=9), this reflects the formative nature of the study and recruitment constraints within a single National Health Service site during ongoing pandemic restrictions. While no formal assessment of data saturation was conducted, thematic redundancy across participants suggested that core experiential patterns were robustly captured. Consistent with Braun and Clarke's [30,31] critique of saturation as a criterion borrowed from grounded theory, we recognized that RTA does not seek conceptual closure but rather embraces interpretative plurality and the coconstruction of meaning between researchers and data.

Analysis was conducted with NVivo software (version 12; QSR International Pty Ltd). The interviews were audio-recorded and transcribed. The answers were then collected in subcategories using open coding. Finally, similar open codes were grouped together and discussed in frequent subsidiary meetings with team members. While not all software functions within the mymobility platform are exclusively PGHD-related (education and messaging), and as the software integrates PGHD as a core functionality that can influence these interactions, and as our approach was not sufficiently fine-grained to distinguish between the impact and interactions of each of the features, we report on clinician reflections on these aspects as well.

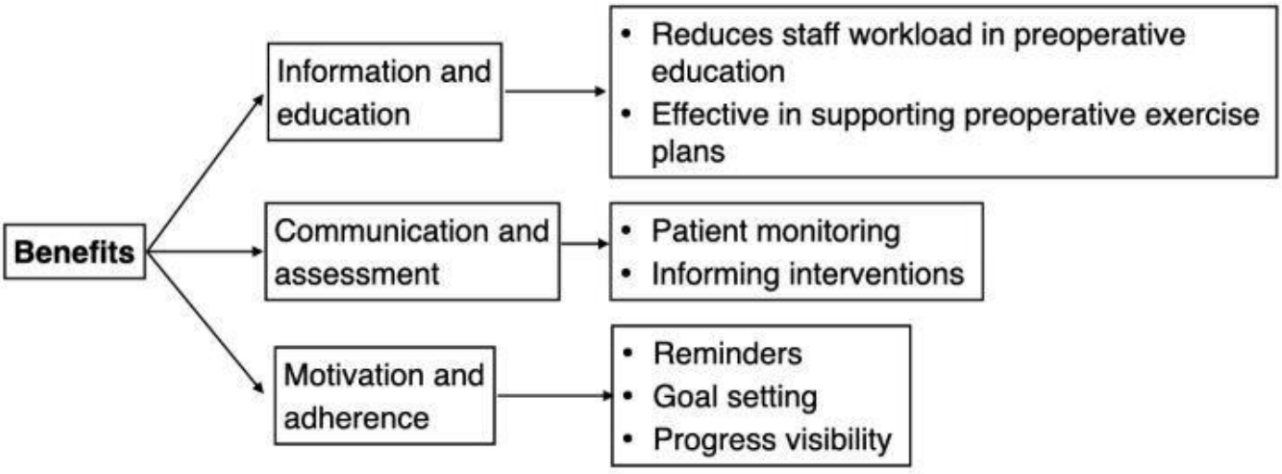
Results

Benefits

Overview

The mymobility app demonstrated benefits in patient information and education, communication and assessment, and patient motivation and adherence (Figure 3). Participants noted that the app reduced staff workload by providing preoperative information and exercise plans, ensuring that patients were well-prepared for surgery. Remote monitoring capabilities offered insights into patient activities and adherence, aiding clinical decisions and interventions. Additionally, features like reminders and goal setting improved patient adherence to clinical protocols.

Figure 3. Benefits.



Information or Education

Participants viewed the app as a valuable tool for reducing staff workload related to preoperative education. Several noted that it helped patients access key information and prepare for surgery independently. One physiotherapist explained that patients "get all the information ... we normally spend a lot of time going through with them," giving them "an idea of what they're getting

XSL•FO
RenderX

themselves into" (Physiotherapist 3). Another remarked that patients find the preoperative education "very beneficial" (Physiotherapist 5). The app also supported home-based exercise routines, with participants noting that the protocol-based rehabilitation content was "absolutely fine prior to the operation" (Physiotherapist 4) and that it allowed patients to better "understand the exercises" (Surgeon 1).

Communication and Assessment

Clinicians consistently valued the app's capacity to monitor patients remotely, particularly in terms of activity tracking, exercise adherence, and recovery progression. For example, one physiotherapist highlighted how access to pre-operative step counts offered a useful baseline for recovery expectations: "How active they are pre-op will influence their speed of recovery after op" (Physiotherapist 1). Another clinician appreciated the insight into adherence and difficulty with exercises, noting that the app enabled them to "see what exercises they're doing ... and how well they're adhering to it" (Physiotherapist 4).

Several participants described the platform as offering an unusually comprehensive overview. One surgeon described it as "brilliant ... it's even doing your gait analysis ... the data we're capturing, it's unbelievable" (Surgeon 3). These data were also used to assess postoperative recovery patterns, helping clinicians gauge improvements or asymmetries over time: "It's quite beneficial to give us a bit more of a guide as to how much load they're putting through their operated limb" (Physiotherapist 5).

Beyond passive monitoring, participants emphasized the potential for the system to inform real-time interventions. One surgeon noted that the data made it easier to detect when patients were deviating from expected progress and intervene earlier if needed (Surgeon 1), while a physiotherapist described how being able to see prior prescribed exercises allowed for better continuity of care and responsive treatment adjustments (Physiotherapist 3). Another surgeon discussed how longitudinal data helped prioritize surgical scheduling for patients with deteriorating function: "We go through the app ... and we are able to assess that they need to be prioritized" (Surgeon 1).

Patient Motivation or Adherence

Participants valued the app's features that supported patient motivation and adherence to rehabilitation protocols. Automated reminders were seen as particularly useful, helping patients follow exercises more consistently: "They seem to adhere a lot better to it when the phone pings at them and tells them to do it" (Physiotherapist 4). Similarly, goal-setting features were appreciated for creating clear expectations and providing structure: "Setting the goals ... is a very good, nice incentive for the patient" (Surgeon 1).

Clinicians also highlighted how progress tracking helped patients maintain perspective during recovery. One physiotherapist noted that visualizing change over time enabled more constructive conversations: "It's really nice to be able to say, 'But look how you were here' ... you actually forget quite how bad you were at the start" (Physiotherapist 3).

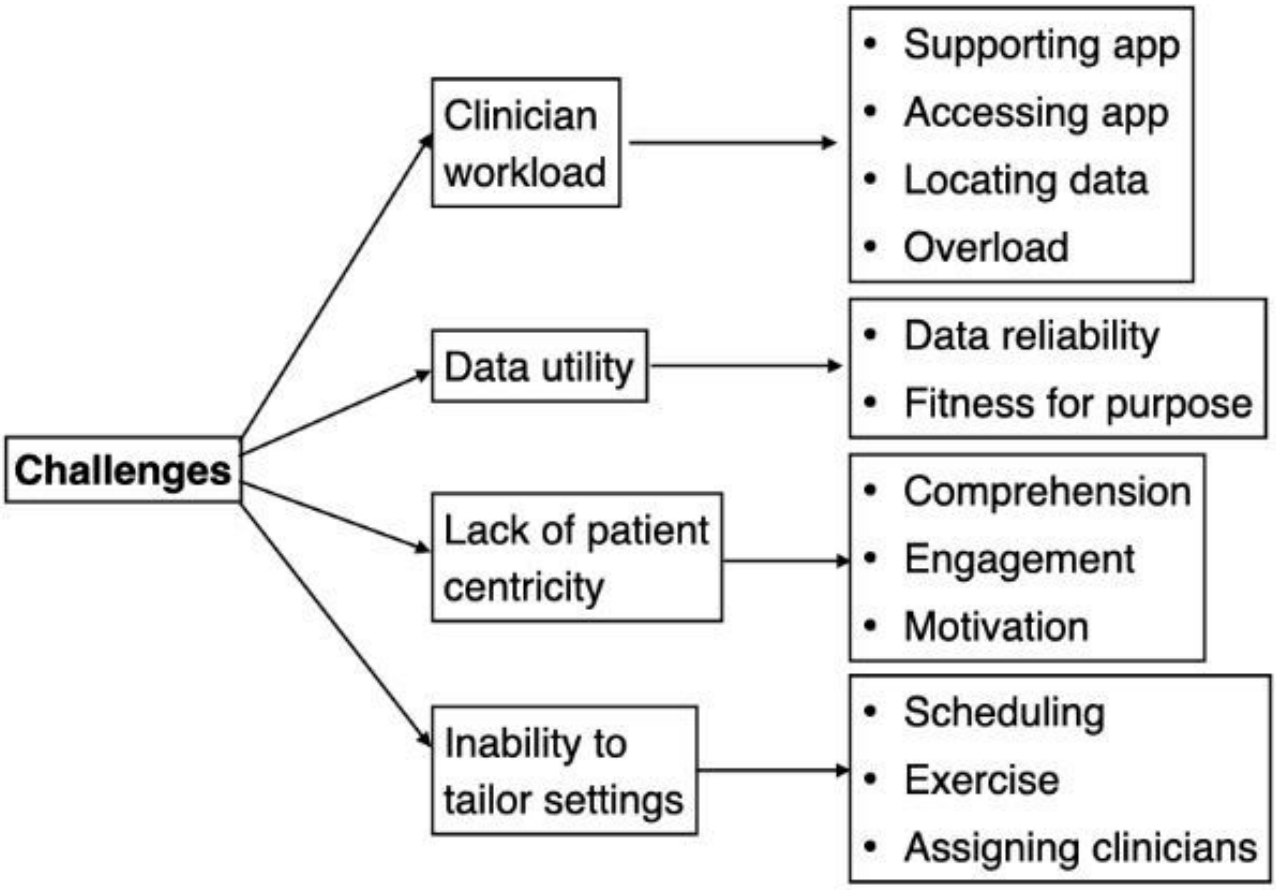
Challenges or Weaknesses

Overview

Despite the benefits noted in the previous section, there were also many noted shortcomings, which draw attention to the ongoing challenges in designing and implementing PGHD-inclusive software. In the following sections, we discuss ways in which the software was reported to often increase clinician workload, data utility was often questionable, was ill-suited for patient needs, and the software lacked the flexibility to be tailored to clinical workflows (Figure 4).

XSL•FO
RenderX

Figure 4. Reported challenges.



Clinician Workload or Time

Clinicians reported that the app introduced additional workload and usability challenges. For example, managing patient queries about technical issues, such as step-count tracking, was time-consuming (Physiotherapist 4). Accessing the platform also required logging into multiple systems, which several participants described as cumbersome or unrealistic within the constraints of outpatient care: “It’s too difficult to log on to lots of different things” (Physiotherapist 5).

Beyond accessibility, participants expressed frustration with the fragmentation and volume of data, which often lacked clinical clarity. One surgeon noted, “Too much information and not quite sure what the information it is giving me is ... I don’t even know what double support percentage means” (Surgeon 2). Others echoed this sentiment, describing the system as cognitively overwhelming and poorly integrated into existing workflows. As one clinician put it, “It’s a lot of data” (Physiotherapist 1).

Utility

Beyond issues of access and cognitive load, clinicians expressed doubt about the reliability and clinical utility of patient-generated data. Some were concerned that behavioral data (eg, step counts) lacked the necessary context. As one physiotherapist explained, “If someone’s doing 3000 steps but limping badly and in significant pain, that’s not particularly useful for us” (Physiotherapist 4). Others questioned the accuracy of passive data collection, noting that reduced activity might reflect patients not carrying their phones, rather than a true change in function.

Several participants also highlighted the insufficiency of quantitative metrics alone, emphasizing the continued need for direct clinical assessment: “You have to look at the patient, you can’t just look at the data” (Physiotherapist 2). For many, the data offered limited actionable value, and some felt it had minimal influence on therapeutic decision-making. As one participant concluded, “It doesn’t change our practice in any way, and it almost makes it longer” (Physiotherapist 4).

Lack of Patient Centricity

Clinicians emphasized that the app’s design did not fully account for the cognitive, motivational, and contextual limitations faced by many patients. Some described difficulties with patients who struggled to understand what was being asked of them, especially those with cognitive impairments or who were frail and living in care settings. As one clinician put it, “A lot of the patients we see here are also quite frail, quite elderly ... I think it may just be one too many things for them to manage” (Surgeon 4). Another noted that misunderstandings about the purpose or expectations of PGHD were common: “One of the biggest challenges is everyone’s understanding of exactly what’s being asked of them and how they measure that accurately” (Physiotherapist 5).

In addition to these accessibility issues, clinicians also reflected on the unintended psychological effects of goal-based metrics and countdowns. While some patients found feedback motivating, others reacted negatively to perceived regressions

XSL•FO
RenderX

or delays. “If they see they’re slipping back,” one physiotherapist explained, “then it’s that defeat ... ‘What’s the point?’” (Physiotherapist 3).

Inability to Tailor to Patient Journey

Clinicians expressed frustration with the app’s limited flexibility to accommodate real-world variations in patient pathways and care team structures. One key issue was the fixed countdown timer, which could become inaccurate after surgery and block access to relevant content: “Her app is now telling her it’s 52 days to her surgery ... but she’s just had it” (Physiotherapist 3). This lack of adaptability also extended to exercise programming. Physiotherapists reported that postoperative exercise options were too generic to meet the diverse needs of patients and could not be tailored to individual recovery trajectories. As one noted, “There’s only two programs ... some exercises are too easy for certain patients, some are too difficult ... there’s no way of changing them” (Physiotherapist 5). In some cases, clinicians advised patients to ignore or delete the app in favor of more appropriate alternatives.

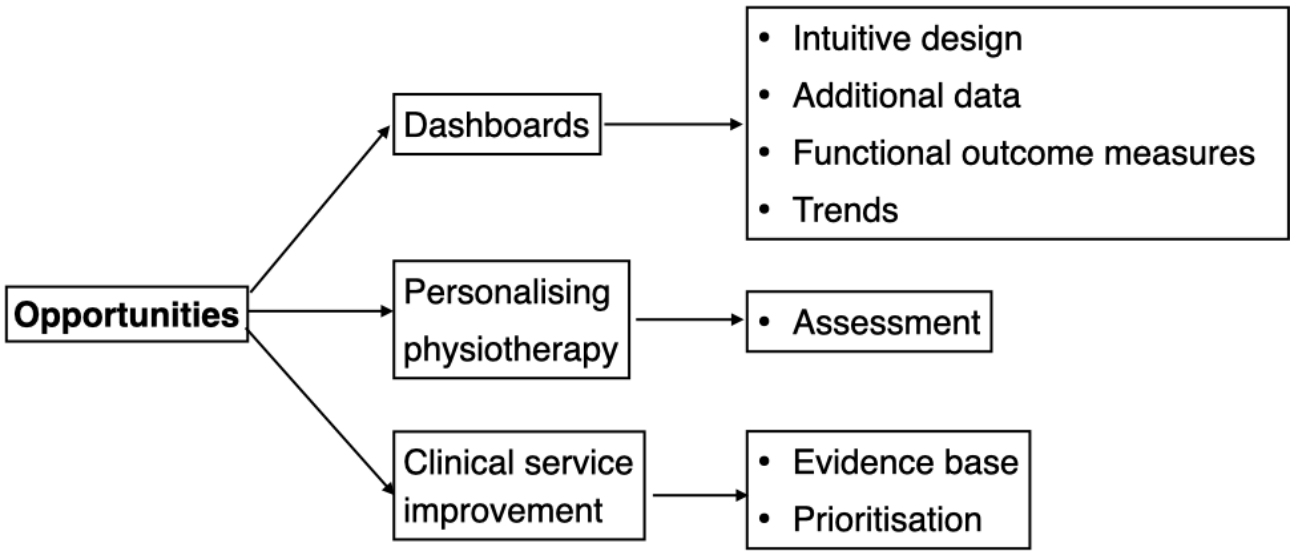
In addition to content rigidity, participants also pointed to deficiencies in role-based functionality. They reported being unable to assign specific clinicians to patients, which led to communication breakdowns and unclear responsibilities. “There’s no way of us assigning ourselves to that patient on the app,” one physiotherapist explained, “so they may still get pinged with exercises we’re not using” (Physiotherapist 5). Surgeons likewise described limited visibility into patient activity between referral and surgery, undermining continuity of care (Surgeon 4).

Opportunities for Improvement

Overview

As well as the benefits and shortcomings noted in the previous section, our analysis suggests several opportunities for enhancing future systems in dashboard design, personalization of physiotherapy, and clinical service improvement. Participants suggested making dashboards more intuitive and informative, with real-time notifications and interest in adding additional data such as heart rate and pain levels. They also emphasized the need for trend identification in patient data to better track progress over time. Personalizing physiotherapy was highlighted as crucial, with tailored exercises based on individual weaknesses and goals, especially after the operation. Finally, the potential of PGHD to improve clinical care and surgical prioritization was recognized, suggesting that more detailed and personalized patient data might enhance treatment outcomes and resource management (Figure 5).

Figure 5. Opportunities.



Dashboard Design

Clinicians proposed several enhancements to dashboard design and data presentation. Many expressed a desire for a more intuitive overview of patients with real-time updates and alerts for significant changes. One surgeon asked, “Why can’t I have [notifications] for my patients?” and described wanting a sidebar dashboard to monitor patient status more easily (Surgeon 2). Participants also suggested visual elements like color coding to draw attention to significant changes in patient mobility (Surgeon 3).

Requests for additional data included heart rate and pain levels, especially during the early recovery phase. “Pain is usually the really big indicator ... in that first six-week period,” noted one physiotherapist (Physiotherapist 4). Others emphasized the importance of context—knowing not just pain scores, but how they relate to exercises, walking aid use, and movement patterns (Physiotherapist 5).

There was strong support for capturing functional outcome measures that reflect daily living. Participants emphasized activities like dressing, standing from a chair, driving, or using public transport as better indicators of recovery than static scores. As one clinician put it, “It’s all about their ability to get on with daily life” (Physiotherapist 3).

Finally, clinicians stressed the value of visualizing patient trends over time. One surgeon explained that even if absolute scores remain low, improvements in mobility or independence may still represent substantial recovery: “Trend is more important than absolute” (Surgeon 4). Another participant noted that

XSL•FO
RenderX

having longitudinal data could significantly improve follow-up consultations after long gaps (Physiotherapist 3).

Personalizing Physiotherapy

Clinicians highlighted the need to personalize physiotherapy plans based on individual patient characteristics, especially postoperatively. “I give different exercises to most of my patients depending on where their weakness is, where their limitation in movement is, and what their goals are,” explained one physiotherapist (Physiotherapist 5). Others emphasized that while protocol-based rehabilitation may suffice preoperatively, postoperative recovery is highly variable. “After the operation, so much depends on what changes ... everyone’s different,” noted Physiotherapist 4.

Clinical Service Improvement

Participants saw potential for PGHD to improve clinical services by informing population-level trends, surgical practices, and prioritization processes. One surgeon noted that, with sufficient data, PGHD could help identify common rehabilitation challenges or communication barriers and enable more targeted support: “We can look into what works for most of the patients ... and target those patients specifically” (Surgeon 1). Others reflected on how such data could prompt critical self-assessment of surgical decisions—for example, identifying stiffness patterns that might relate to prosthetic fit (Surgeon 4).

PGHD were also seen as a tool for refining surgical prioritization. “If their mobility is decreasing rapidly, that’s very important ... so we can try to get them in and have their operation quicker” (Surgeon 1). The dynamic use of deterioration data to rerank patients on waiting lists was described as a promising but underused opportunity.

Discussion

Principal Findings

In this section, we discuss the benefits, challenges, and opportunities raised through our analysis. First, we examine the software’s potential to streamline patient education, improve communication, and enhance motivation. We then explore the hurdles of increased clinician workload, clinician concerns about the reliability or credibility of patient-submitted data, and patient engagement variability. We finally discuss emergent tensions and propose further research to help resolve ongoing challenges.

Benefits

The interviews suggested that the mymobility software with PGHD integration could offer meaningful benefits in the orthopedic care setting. Structured education plans and essential information were appreciated for reducing staff workload and helping patients to prepare for operations. The communication and data sharing functions were useful for monitoring patients and, in some cases, helping to inform interventions. The use of reminders, goal setting, and the ability for patients to view progress over time were noted as showing promise for motivating patients to follow pre- and postoperative care plans. Perhaps most promising for asserting the value of PGHD integration was Surgeon 1 noting that there were specific cases where the shared data had provided evidence for prioritizing care. However, it is important to note that app-based education and preoperative physiotherapy support, some of the more successful noted aspects of the app, do not necessarily require PGHD integration. It remains for further research to determine if PGHD could help with further personalization, for example, with support for comorbidities [32], and if in these cases, benefits would outweigh operational challenges. Physiotherapist 4 noted that the great variance in postoperative outcomes requires in-depth knowledge of the patient’s abilities and progress, which could potentially be supported by greater PGHD integration, though the interviewees expressed concerns that this was more reliably and with less effort accomplished through in-person contact (Physiotherapist 4 and Physiotherapist 2). It is noteworthy and promising that many of the potential benefits of PGHD cited in the literature review were to at least some degree reflected in these interviews and in some cases delivered by the software. As per Cohen et al [4], PGHD did help to provide insights; as per West et al [5], it did help to facilitate care personalization; and as per Caldeira et al [9], the collected data showed promise in motivating and engaging patients. While these benefits might not be universally reported by all clinicians, the responses suggest that this software is making strides in delivering the promised benefits of PGHD integration, and this study provides some empirical evidence that it is viable to integrate PGHD within orthopedic clinical practice.

Challenges

Despite the noted benefits, clinicians reported many barriers to integrating PGHD within clinical care, often echoing the known barriers cited in our review. Particularly problematic were the many ways the app could increase clinician workload, from requiring time to support patients in using the app (Physiotherapist 4), interacting with the platform (Surgeon 2), and locating and making use of data (Physiotherapist 5). Potential benefits of the platform were also complicated by the clinicians’ continued reluctance to trust the validity of PGHD (Physiotherapist 4) and challenges in applying data to decision-making (Physiotherapist 2). Furthermore, clinicians expressed concerns that the need for patients to participate with this approach was unsuitable for many of their patients, echoing the study by Jim et al [6], which noted that the use of PGHD in treatment could increase care disparities. The interviews also highlighted the challenges in tailoring software to reflect individual needs, such as the selection of postoperative exercises or the assigning of clinicians. While the noted interface design shortcomings are likely correctable, the inability to overcome the larger barriers cited within the review is concerning, given the resources likely invested in this software’s design and development. In addition, defining the requirements for the many configurations needed to adequately address personalization of the patient journey likely remains a stubborn problem [33]. This is especially vital, as clinicians noted how this inability to properly reflect the individual patient journey can easily lead to software rejection. Similarly, this leads to another broader question as to how many features should be included within PGHD systems, given how a single deficiency area such as improper notifications can lead to product abandonment.

XSL•FO
RenderX

Opportunities

Clinicians suggested many realizable and relatively standard suggestions for improving product usability, such as making interface terms easier to understand, expanded use of color coding, and indications of how patient progress deviates from norms or expectations. There were also calls for increasing convenience through new features such as notifications of important events and streamlining sign-ins, as well as potentially integrating additional data sources like heart rate, range of movement, pain levels, gait patterns, and increased integration of functional outcome measures to support a more holistic understanding of patients' needs. Such suggestions appear relatively straightforward and reflect established design practices and features within other products. However, other suggestions, such as automated exercise adaptation to individual patient need, might pose greater challenges, given that the clinicians emphasized the importance of human interaction in making these adjustments. While the use of PGHD for clinical service improvement appears especially promising, it might require standardization and trust in PGHD, which might still be culturally lacking at this time.

Design Tensions for Further Research

The software demonstrated capacity to deliver value to clinical care, and many of the noted design deficiencies and suggested improvements are likely amenable to improvement through established iterative human-centered or co-design techniques. Other deficiencies relate to broader societal concerns, which are outside of the ability of software systems to address. In this section, we identify key areas for further research that consist of trade-offs or tensions that might imply conflicting requirements, which cannot be "solved" but rather require design decisions that recognize that trade-offs are inherent. While not necessarily novel, their persistence in this software after significant industry investment suggests that software design alone is insufficient to drive PGHD into the mainstream.

Leveraging Bigger Data Versus Burdening Clinicians

Participants were impressed with the range of available data presented within the mymobility app and even suggested adding additional data. However, they also noted that expansive data could be overwhelming, could increase workload, and that they were often unsure of how to apply such data to their practice. Research question: How can PGHD dashboards be designed to decrease clinician workload? How can the system deliver the right data at the right time to the right clinical role?

Standardization Versus Personalization in Clinical Practice

While standardized protocols and data collection methods are necessary for consistency and reliability, clinicians emphasized the importance of personalized care, particularly in postoperative physiotherapy, where individual patient needs vary significantly. Research question: How can PGHD systems balance the need for standardized data collection and protocols with the flexibility required to personalize patient care effectively?

Comprehensive Versus Targeted Software

Study participants related challenges resulting from accessing information and functionalities from multiple systems, which could support the need for a single comprehensive approach. Yet, participants also noted how one inappropriate feature could lead to system rejection. Research question: What is the appropriate balance in the number of features to include within PGHD-inclusive orthopedic software? How should this be determined?

Adaptive Versus Static Design Approaches

It is recognized that health care is complex with outcomes emerging from multiple interacting systems needing frequent adaption to shifting contexts [34]. This creates a tension between the validation needed to gain regulatory approval for clinical software systems and the dynamic requirements within clinical contexts. Research question: How can clinical software systems be designed that can be validated yet are sufficiently adaptable to variable workflows, preferences, and evolving needs?

Limitations

This study has several limitations. The sample size was limited due to real-world constraints during the COVID-19 pandemic, including staffing shortages and scheduling disruptions. Although we sought diverse perspectives, these logistical challenges prevented broader recruitment and may have introduced selection bias. Additionally, recruitment was limited to a single orthopedic center, which may constrain the transferability of findings. This was due to institutional access constraints and pandemic-related limitations. Future research should explore multicenter participation to capture a broader range of experiences with PGHD integration. Recruitment methods using snowball sampling and existing professional relationships may also introduce selection bias, and the reliance on self-reported data could affect accuracy due to recall bias or socially desirable responses. Additionally, the varying levels of technological familiarity among clinicians and the unique circumstances of the COVID-19 pandemic, which increased reliance on digital tools, may have influenced the results. Future research should address these limitations by including a larger, more diverse sample across multiple health care settings and ideally using random sampling methods to better capture the evolving nature of PGHD integration in clinical practice. Furthermore, as the study was limited to clinicians working with patients who have had TKA or THA operations, the findings may not be transferable to other orthopedic contexts such as trauma, spine, or sports medicine. Future research should investigate the applicability and value of PGHD integration across a wider range of orthopedic conditions.

Conclusions

This study sought to close gaps in knowledge about the real-world application of PGHD integrative software within orthopedic practice. To this end, surgeons and physiotherapists were interviewed to learn more about their experiences and suggestions for improvement. The findings suggest that the mymobility software delivered some meaningful benefits. Structured education plans and the providing of essential information were praised for reducing staff workload and

XSL•FO
RenderX

helping patients prepare for operations. The communication and data-sharing functions showed promise for monitoring patients and informing interventions. Additionally, reminders, goal setting, and progress-tracking features showed the capacity to motivate patients to adhere to pre- and postoperative care plans. Notably, the integration of PGHD helped prioritize care in specific cases, indicating its potential in enhancing clinical decision-making. While some benefits, such as app-based education and preoperative physiotherapy support, do not necessarily require PGHD integration, the software still aligns with literature-reviewed benefits, providing insights, facilitating personalized care, and motivating patients. While there remained a clear need for further interface and interaction design, the software appears to be making strides in delivering on the promised benefits of PGHD within a clinical context.

However, many of the known barriers to PGHD integration remain unresolved. Clinicians reported challenges such as the app increasing their workload by requiring time to support patients in using the app, interacting with the platform, and locating and making use of data. Potential benefits were further hampered by clinicians' reluctance to trust the validity of PGHD and challenges in applying data to decision-making. Additionally, the patient participation required for a PGHD approach posed significant challenges, especially for those who are older or have cognitive impairments. Although the mymobility platform has already been scaled commercially, this study shows that availability does not guarantee seamless clinical integration. Participants described persistent barriers—such as limited time to engage with PGHD, doubts about its clinical relevance, and challenges supporting diverse patient populations—that complicated effective use in practice. These findings suggest that even at scale, PGHD platforms require ongoing refinement of interface design, clinician workflows, and support strategies to ensure meaningful and sustainable adoption.

While many of the perceived benefits and barriers align with prior findings from similar digital health interventions, this study adds value by contextualizing those dynamics within the day-to-day realities of orthopedic care. The role-specific insights from surgeons and physiotherapists offer a practical understanding of how PGHD are actually interpreted, adapted, and applied within clinical workflows.

Future research should explore how to address trade-offs or tensions that may not be fully "solvable" but instead require thoughtful design decisions. Research and development efforts should aim to navigate these complexities in ways that improve usability, utility, and adoption. Further studies should also build on these findings by using quantitative or mixed methods approaches to evaluate the measurable impact of PGHD on clinical workflows, patient engagement, and health outcomes.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: BP (lead), DSK (equal), DG, LP, OP, IR
Data curation: LP (lead), DSK (supporting)
Formal analysis: DSK (lead), LP (equal)
Investigation: DSK (lead), LP (equal), BP, DG, OP, IR (supporting)
Methodology: DSK (lead), BP, DG, LP, OP, IR (supporting)
Writing—original draft: DSK (lead), BP, DG, LP, OP, IR (supporting)
Writing—review and editing: DSK (lead), BP, DG, LP, OP, IR (supporting)

Conflicts of Interest

OP is the UK Lead for the Zimmer Biomet Rapid Recovery Program; he received a travel grant to present his work in 2022. All other authors have no conflicts of interest.

References

1. Grover S, Fitzpatrick A, Azim FT, et al. Defining and implementing patient-centered care: an umbrella review. Patient Educ Couns 2022 Jul;105(7):1679-1688. [doi: 10.1016/j.pec.2021.11.004] [Medline: 34848112]
2. Dickson N. Confident your patients get the best care? You might be surprised how care can improve. Physician's Weekly. 2024. URL: https://www.physiciansweekly.com/confident-your-patients-get-the-best-care-you-might-be-surprised-how-care-can-improve/ [accessed 2024-06-18]
3. Omoloja A, Vundavalli S. Patient generated health data: benefits and challenges. Curr Probl Pediatr Adolesc Health Care 2021 Nov;51(11):101103. [doi: 10.1016/j.cppeds.2021.101103] [Medline: 34799255]
4. Cohen DJ, Keller SR, Hayes GR, Dorr DA, Ash JS, Sittig DF. Integrating patient-generated health data into clinical care settings or clinical decision-making: lessons learned from project HealthDesign. JMIR Hum Factors 2016 Oct 19;3(2):e26. [doi: 10.2196/humanfactors.5919] [Medline: 27760726]
5. West P, Van Kleek M, Giordano R, Weal MJ, Shadbolt N. Common barriers to the use of patient-generated data across clinical settings. 2018 Apr 21 Presented at: CHI '18; Apr 21-26, 2018; Montreal QC Canada p. 1-13. [doi: 10.1145/3173574.3174058]

XSL•FO
RenderX

6. Jim HSL, Hoogland AI, Brownstein NC, et al. Innovations in research and clinical care using patient-generated health data. CA Cancer J Clin 2020 May;70(3):182-199. [doi: 10.3322/caac.21608] [Medline: 32311776]
7. Basch E, Abernethy AP. Supporting clinical practice decisions with real-time patient-reported outcomes. J Clin Oncol 2011 Mar 10;29(8):954-956. [doi: 10.1200/JCO.2010.33.2668] [Medline: 21282536]
8. Yoo DW, De Choudhury M. Designing dashboard for campus stakeholders to support college student mental health. 2019 May 20 Presented at: PervasiveHealth'19; May 20-23, 2019; Trento Italy p. 61-70 URL: https://dl.acm.org/doi/proceedings/10.1145/3329189 [doi: 10.1145/3329189.3329200]
9. Caldeira C, Costa Figueiredo M, Dodakian L, de Souza CRB, Cramer SC, Chen Y. Towards supporting data-driven practices in stroke telerehabilitation technology. Proc ACM Hum-Comput Interact 2021 Apr 13;5(CSCW1):1-33. [doi: 10.1145/3449099]
10. Lai AM, Hsueh PYS, Choi YK, Austin RR. Present and future trends in consumer health informatics and patient-generated health data. Yearb Med Inform 2017 Aug;26(1):152-159. [doi: 10.15265/IY-2017-016] [Medline: 29063559]
11. Guardado S, Karampela M, Isomursu M, Grundstrom C. Use of patient-generated health data from consumer-grade devices by health care professionals in the clinic: systematic review. J Med Internet Res 2024 May 31;26:e49320. [doi: 10.2196/49320] [Medline: 38820580]
12. Lordon RJ, Mikles SP, Kneale L, et al. How patient-generated health data and patient-reported outcomes affect patient-clinician relationships: a systematic review. Health Informatics J 2020 Dec;26(4):2689-2706. [doi: 10.1177/1460458220928184] [Medline: 32567460]
13. Giordanengo A, Årsand E, Woldaregay AZ, et al. Design and prestudy assessment of a dashboard for presenting self-collected health data of patients with diabetes to clinicians: iterative approach and qualitative case study. JMIR Diabetes 2019 Jul 9;4(3):e14002. [doi: 10.2196/14002] [Medline: 31290396]
14. Tiase VL, Hull W, McFarland MM, et al. Patient-generated health data and electronic health record integration: a scoping review. JAMIA Open 2021 Feb 15;3(4):619-627. [doi: 10.1093/jamiaopen/ooaa052]
15. Price BA, Kelly R, Mehta V, McCormick C, Ahmed H, Pearce O. Feel my pain: design and evaluation of painpad, a tangible device for supporting inpatient self-logging of pain. Presented at: CHI Conference on Human Factors in Computing Systems; Apr 21-26, 2018; Montreal, QC, Canada p. 169. [doi: 10.1145/3173574.3173743]
16. Yu S, Garvin KL, Healy WL, Pellegrini VDJ, Iorio R. Preventing hospital readmissions and limiting the complications associated with total joint arthroplasty. J Am Acad Orthop Surg 2015 Nov;23(11):e60-e71. [doi: 10.5435/JAAOS-D-15-00044] [Medline: 26498587]
17. Petersen PB, Jørgensen CC, Kehlet H, Lundbeck Foundation Centre for Fast-track Hip and Knee Replacement Collaborative Group. Fast-track hip and knee arthroplasty in older adults-a prospective cohort of 1,427 procedures in patients ≥85 years. Age Ageing 2020 Apr 27;49(3):425-431. [doi: 10.1093/ageing/afz176] [Medline: 31868901]
18. Batalden M, Batalden P, Margolis P, et al. Coproduction of healthcare service. BMJ Qual Saf 2016 Jul;25(7):509-517. [doi: 10.1136/bmjqs-2015-004315] [Medline: 26376674]
19. Toogood PA, Abdel MP, Spear JA, Cook SM, Cook DJ, Taunton MJ. The monitoring of activity at home after total hip arthroplasty. Bone Joint J 2016 Nov;98-B(11):1450-1454. [doi: 10.1302/0301-620X.98B11.BJJ-2016-0194.R1]
20. Crizer MP, Kazarian GS, Fleischman AN, Lonner JH, Maltenfort MG, Chen AF. Stepping toward objective outcomes: a prospective analysis of step count after total joint arthroplasty. J Arthroplasty 2017 Sep;32(9S):S162-S165. [doi: 10.1016/j.arth.2017.02.058] [Medline: 28343831]
21. Bendich I, Chung C, Hwang K, et al. Changes in prospectively collected longitudinal patient-generated health data are associated with short-term patient-reported outcomes after total joint arthroplasty: a pilot study. Arthroplast Today 2019 Mar;5(1):61-63. [doi: 10.1016/j.artd.2019.01.005] [Medline: 31020024]
22. Walters JL, Mackintosh S, Sheppard L. The journey to total hip or knee replacement. Aust Health Rev 2012 May;36(2):130-135. [doi: 10.1071/AH11050] [Medline: 22624631]
23. Coulter CL, Scarvell JM, Neeman TM, Smith PN. Physiotherapist-directed rehabilitation exercises in the outpatient or home setting improve strength, gait speed and cadence after elective total hip replacement: a systematic review. J Physiother 2013 Dec;59(4):219-226. [doi: 10.1016/S1836-9553(13)70198-X] [Medline: 24287215]
24. Cheng KG, Hayes GR, Hirano SH, Nagel MS, Baker D. Challenges of integrating patient-centered data into clinical workflow for care of high-risk infants. Pers Ubiquit Comput 2015 Jan;19(1):45-57. [doi: 10.1007/s00779-014-0807-y]
25. Baxter G, Sommerville I. Socio-technical systems: from design methods to systems engineering. Interact Comput 2011 Jan;23(1):4-17. [doi: 10.1016/j.intcom.2010.07.003]
26. Kelly R, Jones S, Price B, Katz D, McCormick C, Pearce O. Measuring daily compliance with physical activity tracking in ambulatory surgery patients: comparative analysis of five compliance criteria. JMIR Mhealth Uhealth 2021 Jan 26;9(1):e22846. [doi: 10.2196/22846] [Medline: 33496677]
27. Price L, Rauf I, Gooch D, Katz D, Pearce O, Price B. Children's perspectives on pain-logging: insights from a co-design approach. 2024 Jul Presented at: DIS '24; Jul 1-5, 2024; Copenhagen Denmark p. 1306-1318 URL: https://dl.acm.org/doi/proceedings/10.1145/3643834 [doi: 10.1145/3643834.3661597]
28. Pearce O, Gooch D, Price L, Price B. The challenges of technology adoption in the NHS: lessons learnt from deploying PROMs collection devices. Bulletin 2023 Nov;105(8):376-378. [doi: 10.1308/rcsbull.2023.131]

XSL•FO
RenderX

29. mymobility. Zimmer Biomet. URL: https://www.zimmerbiomet.com/en/products-and-solutions/zb-edge/mymobility.html#01-overview [accessed 2024-05-12]
30. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
31. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport Exerc Health 2019 Aug 8;11(4):589-597. [doi: 10.1080/2159676X.2019.1628806]
32. Jansson M, Vuorinen AL, Harjumaa M, et al. The digital patient journey solution for patients undergoing elective hip and knee arthroplasty: protocol for a pragmatic randomized controlled trial. J Adv Nurs 2020 Jun;76(6):1436-1448. [doi: 10.1111/jan.14343] [Medline: 32133684]
33. Tunis R, Fleischmann KR, Smith ADR. Designing for personalization in personal informatics: barriers and pragmatic approaches from the perspectives of designers, developers, and product managers. 2024 Jul Presented at: Designing Interactive Systems Conference; Jul 1-5, 2024; Copenhagen, Denmark p. 584-596. [doi: 10.1145/3643834.3661622]
34. Greenhalgh T, Papoutsi C. Studying complexity in health services research: desperately seeking an overdue paradigm shift. BMC Med 2018 Jun 20;16(1):95. [doi: 10.1186/s12916-018-1089-4] [Medline: 29921272]

Abbreviations

PGHD: patient-generated health data
RTA: reflexive thematic analysis
STSD: sociotechnical systems design
THA: total hip arthroplasty
TKA: total knee arthroplasty

Edited by A Kushniruk; submitted 08.08.24; peer-reviewed by H Gandhi, Z Chen; revised version received 30.05.25; accepted 09.07.25; published 29.08.25.

Please cite as:
Katz DS, Gooch D, Price L, Rauf I, Pearce O, Price B
Exploring Clinician Experiences With a Digital Platform Supporting Orthopedic Care That Integrates Patient-Generated Health Data: Qualitative Study of Early Users
JMIR Hum Factors 2025;12:e65216
URL: https://humanfactors.jmir.org/2025/1/e65216
doi:10.2196/65216

© Dmitri S Katz, Daniel Gooch, Linda Price, Irum Rauf, Oliver Pearce, Blaine Price. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 29.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

An Informatics-Based, Payer-Led, Low-Intensity Multichannel Educational Campaign Designed to Decrease Postdischarge Utilization for Medicare Advantage Members: Retrospective Evaluation

Danica Fernandes[1], MBS; Elise Kokonas[2], BA; Jai Bansal[1], MA; Ken Hayashima[1], MS; Brian Hurley[1], MS; Annabel Ryu[1], MS; Snehal Mhatre[1], MS; Mohammed Ghori[1], MS; Kelly Jean Craig[3], PhD; Amanda L Zaleski[3], PhD; Lily Vogel[1], MBA; Alena Baquet-Simpson[4], MD; Daniel Reif[1], BS

[1]Analytics & Behavior Change, Aetna, CVS Health, New York, NY, United States

[2]Marketing, Aetna, CVS Health, New York, NY, United States

[3]Clinical Evidence Development, Medical Affairs, CVS Health, Corporate Office, Wellesley, MA, United States

[4]Medical Affairs, Aetna, CVS Health, Wellesley, MA, United States

Corresponding Author:
Kelly Jean Craig, PhD
Clinical Evidence Development, Medical Affairs, CVS Health, Corporate Office, Wellesley, MA, United States

Abstract

Background: Readmission avoidance initiatives have been a priority for the Centers for Medicare & Medicaid Services for over a decade; however, interventions are often high-intensity, costly, and resource-intensive, and therefore, rarely scalable or sustainable. Large national payers are in a unique position to leverage data to identify members in real-time who are at high risk of readmission to prioritize the scaled delivery of tailored behavior change techniques to provide an educational intervention to modify health behaviors.

Objective: This study aims to examine the impact of an informatics-driven, multichannel educational messaging campaign implemented to decrease 30- and 90-day acute inpatient readmissions and emergency department (ED) visits among Medicare Advantage members of a large national payer.

Methods: A quality improvement initiative was designed and implemented to provide an evidence-based outreach campaign using human-centered design and behavior change principles to deliver multiple intervention functions, including timely, contextual, and relevant delivery of education, enablement, and persuasion, to reinforce health-promoting behaviors related to planned or unplanned inpatient admissions. Outcomes, including 30- and 90-day acute inpatient readmissions and ED visits, were retrospectively evaluated from Medicare Advantage members enrolled in a large national health plan residing across the United States between May 2020 and July 2022. Leveraging utilization management data, rules-based logic identified members (N=368,393) with a planned acute inpatient procedure (ie, preadmission) or discharged from an acute hospital stay (ie, postdischarge) within 15 days. Members were sequentially assigned to a standard (N=141,223) or an enhanced (N=227,470) messaging group, whereby the standard group received usual outreach and the enhanced group received an educational intervention via a messaging campaign deployed through multiple low-intensity communication channels (eg, text message, email, direct mail) in addition to standard outreach.

Results: Members who received enhanced outreach had fewer relative 30-day acute inpatient readmissions (−4.1%, 95% CI −5.5% to −2.7%; *P*<.001) and ED visits (−3.4%, 95% CI −5.0% to −1.7%; *P*<.001) compared with members receiving standard outreach. Similarly, these findings persisted for relative 90-day outcomes such that members receiving enhanced outreach experienced fewer acute inpatient readmissions (−5.4%, 95% CI −6.5% to −4.3%; *P*<.001) and ED visits (−3.8%, 95% CI −5.0% to −2.5%; *P*<.001) compared with members receiving standard outreach messaging.

Conclusions: Behavior change techniques deployed via educational interventions as low-intensity multi-channel outreach is an effective strategy to reduce avoidable 30- and 90-day inpatient readmissions and ED visits in recently discharged Medicare Advantage members (primarily >65 years).

(JMIR Hum Factors 2025;12:e63841) doi:10.2196/63841

XSL•FO
RenderX

KEYWORDS

clinical informatics; digital health; health behavior; hospital readmission; personalized education; population health; human-centered design; design; human centered; educational; medicare; behavior change; outreach campaign; readmission; inpatient; messaging campaign; messaging; utilize; utilization

Introduction

Acute inpatient readmissions pose a significant challenge to the US health care system, affecting patient outcomes, health care costs, and overall quality of care. Despite increasing health system focus on readmission avoidance initiatives, hospital readmissions remain a major burden for health care organizations, payers, and patients. On average, there are ≈3.8 million readmissions per year, contributing to US $452.4 billion in US health care costs, of which Medicare members account for ≈60% [1]. As such, the Centers for Medicare & Medicaid Services introduced the Hospital Readmission Reduction Program (HRRP) in 2012 with the primary goal of addressing excessive hospital readmissions and improving the quality of health care delivery [2]. Key activities of the HRRP include identifying high-risk patients, comprehensive discharge planning, patient education, medication reconciliation, timely follow-up appointments, care coordination with primary care providers, and proactive outreach to patients postdischarge to monitor their health status and address any concerns. Under the HRRP, the Centers for Medicare & Medicaid Services imposes financial penalties on hospitals with higher-than-expected 30-day readmission rates for certain conditions, including, but not limited, to acute myocardial infarction, heart failure, and pneumonia [2].

Now, 10+ years since its inception, there has been a dearth of literature focused on reducing hospital readmissions. Of the existing studies, a majority focus on high-intensity interventions that are costly and resource-intensive, such as home visits and clinician-led telephonic outreach to deliver high-touch member education, discharge planning, care coordination, and transitions of care [3-6]. While such interventions have shown great promise, they are generally not universally sustainable for large-scale implementation. In addition, many readmission avoidance programs primarily focus on specific high-risk groups, leaving a significant portion of the patient population underserved. Thus, there is an unmet need to identify practical and cost-effective strategies that can be operationalized and delivered at scale.

Large national payers hold a unique position in the health care landscape owing to continuous data ingestion of claims and clinical informatics data. Leveraging this rich and diverse data foundation, payers are well-poised to identify members in real-time who are at high risk of readmission to prioritize the frequency and intensity of behavior change techniques and interventions. Combined with clinical expertise and other foundational capabilities (ie, multi-channel tools, interoperability, and plan benefit design), these data-informed insights have great potential to enable the delivery of low-cost interventions to modify members' health behaviors associated with a planned inpatient procedure or following discharge from an acute hospital stay. These tailored intervention functions complement and extend existing initiatives across the health care ecosystem, enhance the patient and member care experience, and effectively optimize the use of health care resources on a population level.

As such, the purpose of this study was to explore the impact of a payer-led quality improvement initiative that used behavior change techniques to deliver low-intensity interventions, primarily an education-based campaign with persuasion and enablement, as an innovative approach to address the persistent challenge of hospital readmissions and related health care use outcomes. Specifically, it was hypothesized that Medicare Advantage members receiving enhanced outreach using a multichannel educational campaign would exhibit lower relative 30- and 90-day acute inpatient readmissions and emergency department (ED) visits compared with members receiving standard of care messaging (ie, usual outreach).

Methods

Study Overview

This study represents a retrospective evaluation of a quality improvement intervention. Briefly, a personalized, evidence-based, informatics-enabled campaign framework was designed and implemented with the overall goal to reduce avoidable hospital readmissions in the Medicare Advantage (primarily >65 years) population. Eligible members were sequentially assigned to a standard or enhanced messaging intervention consisting of evidence-based, multichannel lay education. Administrative claims data were deidentified, aggregated, and analyzed to compare relative changes in 30- and 90-day acute inpatient readmissions and ED visits in a group of Medicare Advantage members that received the enhanced messaging campaign compared with those who received the standard messaging campaign.

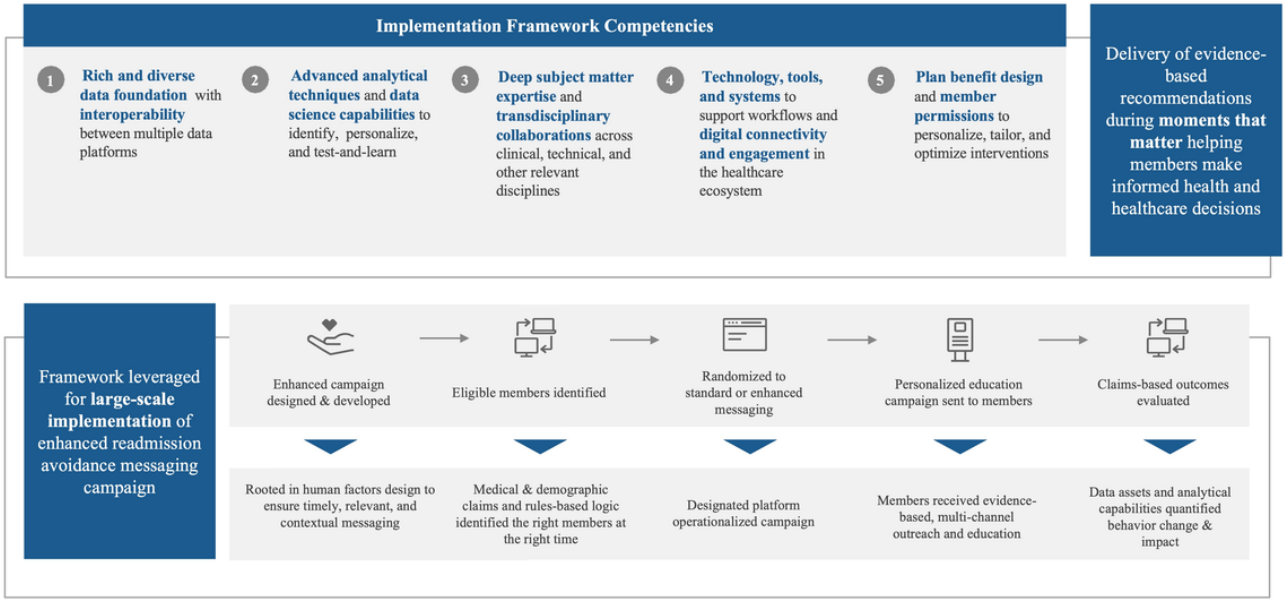
Overview of the Campaign Framework

A quality improvement initiative was designed and implemented to provide an evidence-based outreach campaign using behavior change techniques to deliver multiple intervention functions, including education, enablement, and persuasion, to modify health behaviors related to planned or unplanned inpatient admissions. An education-based outreach intervention was designed with the intention of supporting Medicare Advantage members in taking their "next best action" to better health through analytics-informed behavioral changes to modify health behavior before and after hospital admission. The campaign framework was operationalized across six key components (Figure 1): (1) a rich and diverse data foundation; (2) interoperability between multiple internal and external data platforms; (3) application of analytical techniques and data science capabilities, (4) a designated platform to serve as the centralized technology to operationalize campaigns; (5) deep subject matter expertise, and (6) connectivity and engagement

XSL•FO
RenderX

in the health care ecosystem to deliver actionable recommendations for members, providers, or both.

Figure 1. A high-level overview of the campaign framework. Key foundational competencies directly enable the identification, delivery, and evaluation of an informatics-based, readmission avoidance educational messaging campaign for Medicare members of a large national payer. Potential members are identified through rule-based logic that is highly dependent on a rich and diverse data foundation. Model sets are derived from a data warehouse that continuously ingests data from numerous internal platforms with interoperability. Advanced analytics identifies and routes eligible members to a dedicated platform to operationalize the delivery of evidence-based, multichannel outreach and education campaigns (standard or enhanced). Data assets and analytical capabilities facilitate the retrospective evaluation of behavior change and downstream impacts on health care use patterns.



Member Identification

The campaign was deployed to Medicare Advantage members enrolled in a large national health plan with an acute inpatient index event and discharged to a home setting between May 2020 and July 2022. One index admission was associated with each member's study eligibility, which occurs once per 180-day window. Members were excluded from study evaluation if they were discharged to a nonhome setting (eg, skilled nursing facilities, long-term care); enrolled in Dual-Eligible Special Needs Plans; had nonimpactable conditions (ie, trauma, pregnancy); or were admitted for maternity-related or behavioral health events.

Standard Versus Enhanced Messaging Group Allocation

Existing technological capabilities were leveraged to operationalize campaigns. Specifically, a marketing technology application (previously developed in-house in 2018) served as the centralized tool to enable identification, risk stratification, randomization, channel delivery, engagement, and reporting. Briefly, the platform architecture integrates multiple internal, third-party, and other external data, systems, and services. Defined modules enable data input, processing, and campaign deployment. Precise member targeting and agile data science and marketing pods drive rapid and iterative test-and-learn experimentation approaches that support high-volume test cells to evaluate the impact of campaign-related interventions.

Leveraging these standardized processes, members were sequentially assigned to standard or enhanced messaging groups following response adaptive-randomization procedures with dynamic adjustment for covariate imbalance [7]. Briefly, the standard intervention consisted of usual outreach messaging (ie, single-channel, single-time point evidence-based educational messaging) to members with conditions that were high-risk for readmission while the enhanced intervention used member- and condition-informed personalization to augment the mode, frequency, timing and content of educational messaging in an expanded population (Table 1).

XSL•FO
RenderX

Table . Comparison of standard versus enhanced messaging intervention components.

Intervention component	Standard intervention	Enhanced intervention
Population	• High-risk or select conditions	• All eligible members with an acute inpatient index event
Data sources	• Medical and administrative claims	• Daily list of acute inpatient stays • Precertification vendor data • Provider information • Plan benefit structure
Personalization logic	• Member-centric content (see messaging content below)	• Machine learning to optimize member's channel preferences • Just-in-time messaging • Tailored content (see messaging content below)
Channel delivery	• Channels informed by member permissions • Variable; often single channel	• Multichannel (eg, email, direct mail, SMS, and IVR[a]) approach • Provider (PCP)[b] fax notifications
Timing		
Preadmission	• None	• Within 15 d prior to admission
Postdischarge	• High-risk or select conditions	• 3 - 5 d after discharge to home
Recovery	• High-risk or select conditions	• 7 - 30 d after previous message
Frequency	• One time	• Multiple touchpoints
Messaging content	• Clinically relevant (ie, condition-specific) • Informed by evidence-based guidelines	• Behavior change technique groupings (eg, goals and planning, shaping knowledge), and intervention functions (eg, education, persuasion, and enablement) • Multi-component communication campaign (knowledge sharing, reminders, tracker tool, reinforcement of provider follow-up, complications awareness, etc) • Rooted in human-centered design • Condition-specific considerations • Reinforcement of additional plan-enabled resources and support (eg, transportation)

[a]IVR: interactive voice response.

[b]PCP: primary care provider.

Standard Intervention Description

The standard group received the usual outreach messaging. In general, usual member-facing communication leverages a combination of demographic information, health history, and behavioral economics to create personalized, evidence-based awareness and education campaigns. Personalized messaging campaigns are deployed through multichannel modes of communication delivery, including: digital (eg, SMS, multimedia message service, email, interactive voice response [IVR], 1:1 personalized websites or mobile applications); traditional (eg, direct mail, telephonic); or in-person (eg, provider-led, retail brick-and-mortar store) channels.

General principles of member-facing campaigns are to (1) serve as a trusted source of evidence-based, credibly sourced health education; (2) provide strategies to overcome barriers to enable decision-making; and (3) deliver personalized, relevant, and timely information that facilitates informed health care decisions. Education-based campaigns are grounded in well-established, evidence-based guidelines and best-practice recommendations for population health (eg, United States Preventive Services Task Force, Centers for Disease Control and Prevention, and American Heart Association). All educational content reinforces or references publicly available health education intended for lay audiences and approved by an internal panel of medical directors with subject matter expertise.

Primary outcomes of member-facing communications vary depending on the use case but are designed to favorably impact clinical outcomes, appropriate health care utilization, medical health care spend, quality metrics, or member satisfaction. For example, many messaging campaigns are designed to favorably augment outcomes related to primary (eg, vaccinations), secondary (eg, screenings), tertiary (eg, disease management),

XSL•FO
RenderX

and quaternary (eg, prevent overmedicalization) care activities. Note that standard outreach communications may differ (ie, frequency, intensity, and type) based on member permissions, plan benefit design, gaps in care, and member risk/need. In the event of potentially redundant outreach initiatives, standard campaigns may be suppressed to facilitate test-and-learn outreach campaigns that are most likely to optimize member outcomes. Specific to this use case, examples of interventions that members may receive include payer-led readmission avoidance program support; postdischarge, telephonic-based, nurse care manager outreach; pharmacist-based medication reconciliation; home health services; remote monitoring; and social support services.

Enhanced Intervention Description

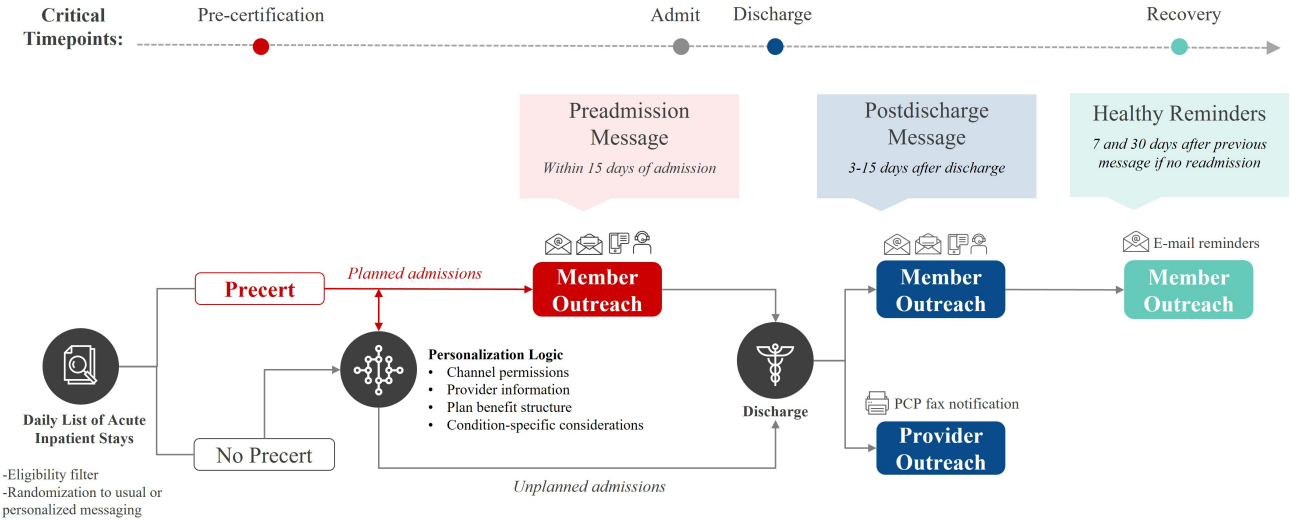
The enhanced group received usual (ie, standard) outreach messaging in addition to personalized readmissions-focused educational outreach using a low-intensity multichannel approach. The enhanced campaign was designed and implemented by a transdisciplinary team composed of experts across data science, clinical, economic, behavioral, marketing, and public health disciplines. Human factors principles were integrated into an agile, rapid-cycle testing framework [8,9] to ensure usability, accessibility, and real-world applicability of readmission reduction interventions for the identified user and context; Medicare members with a recent or planned hospitalization. User requirements for the enhanced intervention were informed by best practices for readmission prevention [2], including recommended transitional care activities (ie, timely postdischarge provider follow-up) and self-monitoring and self-management activities (ie, medication adherence, symptom tracking, and postacute recovery). Human-centered design principles were then applied to enhance local adaptability based on anticipated needs, barriers, and facilitators that were informed by the existing literature [3,9-16] and in-house medical directors with subject matter expertise. These specifications informed the development of the final campaign that was pilot-tested against design requirements prior to large-scale implementation. Continuous feedback loops, including call transcripts, surveys, and evaluations, enabled iterative refinements to messaging, workflows, and intervention delivery, ensuring alignment with user needs and real-world constraints.

The overall goal of the campaign was to support, empower, and encourage members to engage in clinical actions aligned with guideline-directed care. Personalized, timely outreach at critical touchpoints in a member's journey (ie, preadmission and postdischarge) ensured that messaging was contextually relevant and most likely to translate to goal-concordant behavior change related to planned or unplanned hospital stay. Messaging was grounded in behavioral economics of decision-making, specifically the "foot-in-the-door" technique [17], which facilitates engagement by first encouraging small, manageable tasks (eg, filling out a preadmission checklist or a postdischarge recovery tracker), which often translates to participation with a larger ask, such as avoiding unnecessary health care utilization [17]. Additionally, behavior change technique groupings (eg, goals and planning, shaping knowledge) [18], and intervention functions (eg, education, persuasion, and enablement) [19] were incorporated to reinforce health-promoting behaviors.

The campaign (Figure 2) identified members with a planned acute inpatient procedure (ie, preadmission) using the precertification process or those who were discharged from an acute hospital stay (ie, postdischarge), both within a 15-day event window time span. Just-in-time lay educational messaging was strategically timed to reinforce key recovery actions at critical touchpoints in the member's journey. The outreach strategy used all available communication channels based on member preferences, ensuring inclusivity across digital and nondigital mediums. Preadmission outreach provided timely education on essential preparation steps, including instructions to (1) prepare the home environment, (2) schedule provider follow-up visits and manage prescription fills, and (3) stay on track with recovery plans. Postdischarge outreach supported recovery during the critical 30-day period by delivering tools and information. All members received a recovery tracker calendar with designated fields to record self-reported signs, symptoms, and notes, along with stickers to reinforce daily use. The recovery tracker was designed for portability, enabling members to easily bring to follow-up appointments to facilitate patient-provider communication. Messaging emphasized the importance of provider follow-up, medication adherence, and symptom recognition for complications that would warrant further care. In addition, contact information was provided to facilitate inbound telephonic outreach to assist with potential questions or needs.

XSL•FO
RenderX

Figure 2. Readmissions campaign member flow. Overall campaign logic flow and cadence containing both preadmission (where applicable) and postdischarge messaging and channels. Potential channels for initial member outreach were informed by member permissions and included: e-mail, direct mail, SMS, and interactive voice response (IVR). Provider fax notifications were also sent to all members with a provider on file. PCP: primary care provider.



Examples of enhanced member-facing preadmission and postdischarge messaging can be referenced in Multimedia Appendices 1 and 2, respectively. Human-centered design ensured messaging was accessible and engaging [11]. For example, plain language and large fonts optimized readability and usability, addressing common barriers to comprehension. To accommodate varying levels of digital literacy, all members received direct mail (at minimum), with reinforcement through digital channels (eg, SMS, email, IVR) based on member preferences, ensuring engagement was not limited by access or comfort with technology. Operational efficiencies were implemented to improve the timeliness of direct mail delivery. Creatives were preprinted and mailed daily to ensure arrival within 3 days of identification. Personalization was further enhanced to improve relevance and usefulness using rules-based logic informed by provider relationships, medical and social needs, and plan benefit structure. For example, for members with a provider of record, fax notifications were incorporated to reinforce patient-provider communication and care coordination. Where plan benefits allowed, postdischarge messaging included transportation options and additional support for medical and social needs.

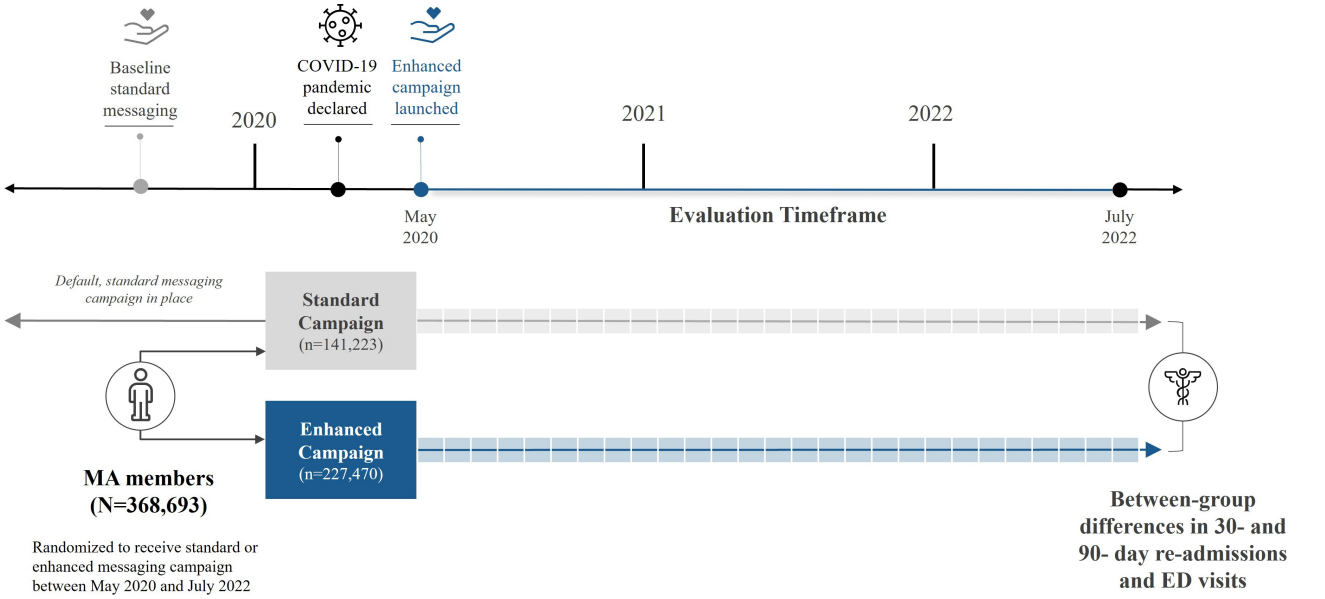
Campaign Evaluation

Data Sources and Study Setting

Retrospective demographic and medical claims data were deidentified, aggregated, and analyzed to determine the impact of the enhanced versus standard educational campaign on primary outcomes for the evaluation timeframe (ie, May 2020 to July 2022). The study design overview is provided in Figure 3. Claims data included ED diagnoses, procedures, sites of care, and provider information. Additional data fields included aggregations of the above information in the forms of medical cases, episode treatment groups, and chronic condition flags. Demographic information collected during health insurance enrollment included self-reported sex, age, plan benefit design, location, and census tract statistics.

XSL•FO
RenderX

Figure 3. Study design overview. The campaign experimentation framework was leveraged to conduct a retrospective analysis of 30- and 90-day outcomes (ie, hospital readmissions and emergency department [ED] visits) between members who received the enhanced messaging campaign compared with a cohort of members who received the standard campaign.



Inclusion and Exclusion Criteria

All participants were enrollees of a Medicare Advantage health plan provided by a large national payer. Inclusion criteria included: (1) receipt of ≥1 standard or enhanced messaging campaign intervention throughout the study evaluation period. Potential participants were excluded if they: (1) did not meet inclusion criteria or (2) were members of a commercial, dual coverage, or indemnity insurance plan.

Ethical Considerations

The Sterling Institutional Review Board reviewed and approved the study (#10132) as an exempt study under 45 CFR 46.104(d)(4). Informed consent was not obtained, as an exemption determination was provided. In addition, a waiver of Health Insurance Portability and Accountability Act authorization for the use and disclosure of aggregated, deidentified member data was obtained. No compensation was provided.

Statistical Analyses

Two-tailed unpaired *t* tests examined between-group differences in baseline characteristics to ensure adaptive randomization techniques effectively achieved balance. Change in acute inpatient readmission rate was calculated as the difference in the average number of readmissions per 100 members (ie, readmission rate) between the standard and enhanced outreach groups. The relative reduction in inpatient readmissions was calculated as the between-group difference in readmission rates divided by the standard outreach group's readmission rate and expressed as a percentage ±95% CI. Change in ED visits was calculated as the difference in the average number of ED visits per 100 members between the standard and enhanced outreach groups. The relative decrease in ED visits was calculated as the between-group difference in ED visits divided by the standard outreach group's ED visit rate and expressed as a percentage ±95% CI.

Secondary analyses explored differences in 30- and 90-day acute inpatient readmissions and ED visits by age brackets (ie,<65 yr, 65 - 69 yr, 70 - 74 yr, 75 - 79 yr, 80 - 84 yr, and >80 yr), geographic region (ie, Midwest, Northeast, Southeast, West), geographic area (ie, urban, suburban, rural), as well as common comorbidities (eg, dementia, diabetes mellitus, low back pain, and obesity).

Covariate balance was confirmed by assessing standardized mean differences for all variables reported. Chi-square tests were used to test the statistical significance of the outreach impact on primary outcomes. Two-tailed unpaired *t* tests were used to explore between-group differences, with statistical significance defined as $P<.05$.

Results

Baseline Characteristics

Table 2 details the baseline characteristics of the standard (N=141,223) and enhanced (N=227,470) outreach groups. On average, the total included study population (N=368,693) was comprised of older adult (mean age:75.3 yr) males (n=181,083 [49.1%]) and females (n=187,610 [50.9%]); a majority of which residing in rural geographies (n=189,594 [51.4%]) with top 3 comorbidities being heart conditions (n=153,353 [41.6%]), diabetes mellitus type 1 or 2 (n=115,432 [31.3%]), and low back pain (n=114,954 [31.2%]). There were no between-group differences in baseline demographic characteristics, risk factors, or comorbidities (all $P>.12$). Of note, mean age was marginally higher in the enhanced (75.4 yr) versus standard (75.3 yr) group ($P<.001$), though this is not considered clinically significant.

XSL•FO
RenderX

Table . Baseline demographic and clinical characteristics among the total study population and comparison by group.

Characteristic	Total population (N=368,693)	Standard outreach (n=141,223)	Enhanced outreach (n=227,470)	SMD[a]	*P* value
Sex, n (%)				0.005	.12
Male	181,083 (49.1)	69,623 (49.1)	111,460 (49.0)		
Female	187,610 (50.9)	71,600 (50.7)	116,010 (51.0)		
Age (years), mean (SD)	75.3 (9.9)	75.4 (9.9)	75.3 (10.0)	–0.010	<.001
Age at admission (years), n (%)					
<65	37,889 (10.3)	14,687 (10.4)	23,202 (10.2)	–0.005	.91
65 - 69	56,269 (15.3)	21,466 (15.2)	34,803 (15.3)	0.002	.89
70 - 74	76,578 (20.8)	28,809 (20.4)	47,769 (21.0)	0.012	.89
75 - 79	75,245 (20.4)	28,386 (20.1)	46,859 (20.6)	0.012	.88
80 - 84	58,677 (15.9)	22,737 (16.1)	35,940 (15.8)	–0.006	.88
>85	64,035 (17.5)	25,138 (17.9)	38,897 (17.2)	–0.016	.87
US region, n (%)					
Midwest	102,952 (27.9)	39,260 (27.8)	69,518 (28.0)	0.003	.85
Northeast	124,446 (33.8)	48,016 (34.0)	83,413 (33.6)	–0.007	.84
Southeast	92,401 (25.1)	35,306 (25.0)	62,369 (25.1)	0.001	.86
West	48,894 (13.2)	18,641 (13.1)	32,705 (13.2)	0.004	.90
Geographic area, n (%)					
Urban	90,950 (24.7)	35,447 (25.1)	55,503 (24.4)	–0.015	.85
Suburban	88,149 (23.9)	33,329 (23.6)	54,820 (24.1)	0.010	.87
Rural	189,594 (51.4)	72,447 (51.2)	117,147 (51.5)	0.005	.80
Comorbidities, n (%)					
Behavioral health[b]	96,488 (26.2)	36,436 (25.8)	60,052 (26.4)	0.013	.86
Cancer	89,059 (24.2)	33,329 (23.6)	55,730 (24.5)	0.021	.87
COVID-19	12,308 (3.3)	4802 (3.4)	7507 (3.3)	0.000	.95
Dementia	21,298 (5.8)	8332 (5.9)	12,966 (5.7)	–0.008	.93
Diabetes mellitus	115,432 (31.3)	43,779 (31.0)	71,653 (31.5)	0.011	.85
Heart conditions[c]	153,353 (41.6)	58,043 (41.1)	95,310 (41.9)	0.017	.83
Hepatitis	5530 (1.5)	2118 (1.5)	3412 (1.5)	0.004	.97
HIV	1106 (0.3)	424 (0.3)	682 (0.3)	0.006	.99
Low back pain	114,954 (31.2)	43,073 (30.5)	71,881 (31.6)	0.024	.85
Multiple sclerosis	1843 (0.5)	706 (0.5)	1137 (0.5)	–0.004	.99
Obesity	82,336 (22.3)	30,928 (21.9)	51,408 (22.6)	0.017	.88
Parkinson disease	5162 (1.4)	1977 (1.4)	3185 (1.4)	0.003	.97
Rheumatoid arthritis	10,920 (3.0)	4095 (2.9)	6824 (3.0)	0.006	.95
Renal conditions[d]	71,385 (19.4)	27,256 (19.3)	44,129 (19.4)	0.006	.88
SUD[e]	20,506 (5.6)	7767 (5.5)	12,738 (5.6)	0.005	.94

[a]SMD: standardized mean difference.

[b]Anxiety, bipolar disorder, depression, psychoses, eating disorders, and postpartum disorders.

[c]Heart failure, congenital heart disease, and ischemic heart disease.

[d]End-stage renal disease and chronic renal failure.

XSL•FO
RenderX

[e]SUD: substance use disorder.

Thirty-Day Inpatient Readmissions and ED Visits

Members in the enhanced outreach group experienced a 4.1% (95% CI 2.7% - 5.5%) relative reduction in acute inpatient readmissions compared with the standard outreach group (*P*<.001). In addition, the enhanced outreach group experienced a 3.4% (95% CI 1.7% - 5.0%) relative reduction in unplanned ED visits compared with the standard outreach group (*P*<.001).

Secondary analyses (Table 3) revealed that members between the ages of 70 and 74 years experienced the greatest relative decrease (5.8%) in acute inpatient readmissions, followed by members who were 65 - 69 yr (5.3%), and 75 - 79 yr (4.5%) (all *P*<.05). Members who reside in the Northeast (8.0%) and Midwest (3.4%) experienced a greater relative reduction in acute inpatient readmissions (both *P*<.05) compared with those in the Southeast or West, as did those in suburban (7.6%) compared with members in urban (3.8%) and rural areas (both *P*<.05). Certain comorbidities also demonstrated a relationship with decreased frequency of 30-day readmissions (Table 2), including behavioral health, cancer, diabetes mellitus type 1 or 2, and heart conditions (all *P*<.05).

XSL•FO
RenderX

Table . Relative change in 30-day acute inpatient readmissions and emergency department (ED) visits by demographic and clinical characteristics.

Characteristic	Change in 30-day inpatient readmissions (%), (95% CI)	*P* value	Change in 30-day ED visits (%), (95% CI)	*P* value
Age at admission (years)				
<65	−2.3 (−5.7 to 1.2)	.38	−1.1 (–4.8 to 2.6)	.57
65 - 69	−5.3 (−8.5 to −2.1)	.02	−1.9 (–5.5 to 1.7)	.36
70 - 74	−5.8 (−8.6 to −2.9)	.005	−5.7 (–9.1 to −2.4)	.001
75 - 79	−4.5 (−7.2 to −1.8)	.02	−6.7 (–10.0 to −3.4)	<.001
80 - 84	−4.1 (−7.1 to −1.1)	.06	1.9 (−1.5 to 5.2)	.36
>85	−1.4 (−4.1 to 1.3)	.51	−4.0 (−7.3 to −0.7)	.03
Region				
Midwest	−3.4 (−5.7 to −1.1)	.04	−2.8 (−5.3 to −0.2)	.04
Northeast	−8.0 (−10.2 to −5.8)	<.001	−5.7 (−8.3 to −3.2)	<.001
Southeast	−0.1 (−2.4 to 2.2)	.95	−1.8 (−4.6 to 1.1)	.28
West	−2.4 (–5.7 to 0.8)	.34	−2.0 (−5.7 to 1.6)	.34
Geographic area				
Urban	−3.8 (−6.1 to −1.4)	.03	–3.8 (–6.9 to –0.7)	.02
Suburban	−7.6 (−10.2 to −5.1)	<.001	–6.8 (–9.8 to –3.7)	<.001
Rural	−2.3 (−4.0 to −0.6)	.07	–1.9 (–3.7 to –0.1)	.07
Comorbidities				
Behavioral health[a]	−3.4 (−5.6 to −1.1)	.049	–2.7 (–5.2 to –0.1)	.05
Cancer	−4.2 (−6.5 to −1.9)	.02	−5.4 (−8.3 to −2.5)	<.001
COVID-19	−0.3 (−6.0 to 5.5)	.95	−8.4 (−15.5 to −1.3)	.008
Dementia	−0.7 (−5.2 to 3.8)	.85	−5.8 (−11.5 to −0.2)	.04
Diabetes mellitus	−4.4 (−6.4 to −2.4)	.003	–4.6 (–7.0 to –2.2)	<.001
Heart conditions [b]	–3.1 (–4.8 to –1.4)	.02	−2.4 (−4.5 to −0.4)	.04
Hepatitis	−5.8 (−14.1 to 2.4)	.31	−21.6 (−34.8 to −8.3)	<.001
HIV	−5.5 (−26.0 to 15.0)	.70	−30.1 (−60.3 to 0.1)	.005
Low back pain	−3.0 (−5.2 to −0.9)	.06	−5.8 (−8.3 to −3.3)	<.001
Multiple sclerosis	−2.8 (−19.8 to 14.2)	.83	1.8 (−17.3 to 20.9)	.86
Obesity	−2.7 (−5.2 to −0.3)	.15	−2.1 (−4.9 to 0.8)	.21
Parkinson disease	−5.7 (−15.3 to 3.9)	.42	−16.5 (−29.3 to −3.7)	.001
Rheumatoid arthritis	−7.2 (−14.0 to −0.3)	.14	−10.0 (−18.1 to −1.9)	.02
Renal conditions[c]	−1.7 (−4.0 to 0.6)	.33	0.9 (−2.0 to 3.7)	.59
SUD[d]	−3.8 (−8.2 to 0.7)	.25	1.1 (−4.0 to 6.2)	.67

[a]Anxiety, bipolar disorder, depression, psychoses, eating disorders, and postpartum disorders.

[b]Heart failure, congenital heart disease, and ischemic heart disease.

[c]End-stage renal disease and chronic renal failure.

[d]SUD: substance use disorder.

Ninty-Day Inpatient Readmissions and ED Visits

These trends (ie, 30-day) persisted at 90-days postdischarge such that the enhanced outreach group experienced a 5.4% (95% CI 4.3% - 6.5%) and 3.8% (95% CI 2.5% - 5.0%) relative reduction in acute inpatient readmissions and ED visits, respectively, compared with the standard outreach group (both *P*<.001).

Secondary analyses revealed between-group differences across age, geographic region, geographic area, and comorbidities (Table 3). Similar to the 30-day results, members between the ages of 70 and 74 years experienced the greatest relative

XSL•FO
RenderX

reduction in 90-day acute inpatient readmissions on the order of 8.5% (vs 2.2% - 5.9% for all other age groups). Members with behavioral health, cancer, diabetes mellitus type 1 or 2, heart conditions, low back pain, obesity, Parkinson disease, and renal-related conditions receiving enhanced messaging demonstrated greater reductions in acute inpatient readmissions at 90 days postdischarge (all $P<.05$) compared to members receiving standard messaging (Table 4).

Table . Relative change in 90-day acute inpatient readmissions and emergency department (ED) visits by demographic and clinical characteristics.

Characteristic	Change in 90-day inpatient readmissions (%), (95%CI)	*P* value	Change in 90-day ED visits (%), (95% CI)	*P* value
Age at admission (years)				
<65	–5.1 (–7.7 to –2.4)	<.001	–2.3 (–5.4 to 0.9)	.015
65 - 69	–3.9 (–6.4 to –1.5)	.007	–1.4 (–4.2 to 1.5)	.27
70 - 74	–8.5 (–10.7 to –6.3)	<.001	–8.9 (–11.5 to –6.3)	<.001
75 - 79	–5.6 (–7.7 to –3.5)	<.001	–5.1 (–7.6 to –2.7)	<.001
80 - 84	–5.9 (–8.2 to –3.7)	<.001	–0.7 (–3.2 to 1.8)	.56
>85	–2.2 (–4.2 to –0.2)	.09	–1.8 (–4.1 to 0.6)	.11
Region				
Midwest	–4.7 (–6.5 to –3.0)	<.001	–2.8 (–4.8 to –0.9)	<.001
Northeast	–8.5 (–10.1 to –6.9)	<.001	–5.3 (–7.2 to –3.3)	<.001
Southeast	–3.0 (–4.8 to –1.2)	.006	–3.3 (–5.5 to –1.0)	<.001
West	–2.6 (–5.1 to –0.1)	.10	–3.6 (–6.6 to –0.6)	.005
Geographic area				
Urban	–4.9 (–6.7 to –3.2)	<.001	–4.2 (–6.6 to –1.8)	<.001
Suburban	–9.9 (–11.9 to –8.0)	<.001	–6.1 (–8.5 to –3.7)	<.001
Rural	–3.2 (–4.4 to –1.9)	<.001	–2.7 (–4.1 to –1.3)	<.001
Comorbidities				
Behavior health[a]	–4.8 (–6.6 to –3.1)	<.001	–3.7 (–5.6 to –1.7)	<.001
Cancer	–4.4 (–6.2 to –2.7)	<.001	–4.4 (–6.5 to –2.3)	<.001
COVID-19	–0.5 (–4.6 to 3.7)	.85	–4.3 (–9.4 to 0.8)	.01
Dementia	–0.3 (–3.5 to 3.0)	.89	–3.5 (–7.6 to 0.6)	.02
Diabetes mellitus	–4.9 (–6.4 to –3.4)	<.001	–5.1 (–6.9 to –3.3)	<.001
Heart conditions[b]	–3.9 (–5.2 to –2.7)	<.001	–4.4 (–6.0 to –2.8)	<.001
Hepatitis	–3.7 (–9.7 to 2.3)	.24	–10.8 (–20.1 to –1.5)	<.001
HIV	–6.5 (–22.1 to 9.2)	.45	–31.7 (–54.6 to –8.8)	<.001
Low back pain	–6.4 (–8.0 to –4.7)	<.001	–6.3 (–8.2 to –4.4)	<.001
Multiple sclerosis	7.4 (–3.9 to 18.7)	.33	–0.6 (–14.8 to 13.5)	.92
Obesity	–4.5 (–6.4 to –2.7)	<.001	–5.3 (–7.5 to –3.0)	<.001
Parkinson disease	–8.5 (–15.8 to –1.3)	.04	–14.8 (–24.4 to –5.1)	<.001
Rheumatoid arthritis	–4.5 (–9.4 to 0.5)	.13	–5.1 (–11.2 to 0.9)	.04
Renal conditions[c]	–3.1 (–4.8 to –1.4)	.002	–2.6 (–4.8 to –0.4)	.004
SUD[d]	–2.9 (–6.2 to 0.3)	.11	–0.5 (–4.7 to 3.6)	.67

[a]Anxiety, bipolar disorder, depression, psychoses, eating disorders, and postpartum disorders.

[b]Heart failure, congenital heart disease, and ischemic heart disease.

[c]End-stage renal disease and chronic renal failure.

[d]SUD: substance use disorder.

XSL•FO
RenderX

Discussion

Principal Findings

This study sought to explore the impact of a payer-led, quality improvement initiative that was designed and implemented to deliver a low-intensity, multichannel educational outreach campaign to modify health behaviors with the goal of decreasing unplanned and avoidable hospital readmissions and ED visits at 30- and 90-days. Consistent with our hypothesis, members receiving the enhanced outreach campaign intervention experienced greater relative reductions in both outcomes compared with the standard outreach group at both 30- and 90-days postdischarge. Put into context, a 4% decrease in 30-day readmissions would translate to approximately US $600,000 in avoidable health care costs for every 1000 members or ≈US $136.5M for the included study sample of 227,470 members (assuming a conservative estimate of US $15,000 per readmission) [1].

The effectiveness of low-intensity, postdischarge follow-up has been examined in previous studies; however, a majority of research has focused on provider-led settings with direct clinician engagement and escalation mechanisms. Bressman et al [20] found that an automated, bi-directional, text-based follow-up intervention from their primary care physician following an index admission translated to 41% lower (adjusted odds ratio) 30-day acute inpatient readmissions or ED visits among 374 patients recently hospitalized (compared with a control, no messaging cohort). Similarly, Patel et al [21] were the first to evaluate the impact of provider-facing secure text messaging on patient outcomes among ≈6400 patients on inpatient services. Patients whose providers were engaged via text messaging to communicate patient information and to help facilitate medical decision-making demonstrated a significant decrease in length of stay on the order of ≈1 day. The larger effect size observed in these studies is likely attributed to differences in setting and intervention design. For example, their interventions were conducted within a single health care system, enabling direct provider engagement and follow-up, whereas this study implemented educational messaging at scale across a nationally distributed MA population. Additionally, the smaller sample size may have influenced effect size estimates by reducing the heterogeneity of outcomes.

This study expands upon the existing literature by evaluating the impact of a payer-led readmission avoidance campaign at scale. Despite the modest relative reductions observed in our study, the large-scale implementation across ≈368k members underscores the potential population-level impact of informatics-driven outreach. The success of the low-intensity multichannel educational outreach campaign is likely multifactorial. Access to extensive claims and clinical informatics data played a critical role in identifying high-risk patients in real-time, enabling the timely deployment of intervention efforts to reduce readmissions. Integrating these data into the outreach campaign enabled precision and effectiveness in reaching the target population. Unlike resource-intensive high-intensity interventions, the low-intensity campaign allowed for engagement with a larger number of members, ensuring that a broader population had equitable access to the necessary support and education to effectively manage their health at a critical moment that matters in their health care journey.

This multichannel outreach campaign embraced a holistic approach to care that was rooted in behavior change. Behavior change technique groupings (eg, goals and planning, shaping knowledge) [18], and intervention functions (eg, education, persuasion, and enablement) [19] were applied to modify health behaviors with the goal of reducing hospital readmissions. Numerous components that combined various self-management techniques, including knowledge acquisition, independent health monitoring, medication adherence, and lifestyle changes, comprehensively addressed member needs in a tailored manner. The multichannel aspect of outreach was designed to reinforce educational messaging that was personalized, contextual, and relevant. Additionally, targeted messaging and outreach efforts were deployed to both members and their providers on record, whenever possible. Specifically, campaign messaging encouraged members to proactively schedule postdischarge follow-up visits and communicate potential complications that would warrant additional clinical support. Simultaneously, payer-led provider alerts regarding their patients' recent hospitalization equipped providers with pertinent information, enabling them to deliver more informed and timely interventions. Overall, this dyadic approach (ie, member-provider) aimed to bridge the communication gap and foster continuity of care by reinforcing the relationship between members and their existing care team.

Last, the low-intensity delivery of communications allowed for broader inclusion of Medicare-insured members across different regions of the United States compared with higher-intensity outreach, which is often cost and labor-intensive. Previous studies examining readmission reduction strategies have found that higher-intensity outreach combining multiple methods is more effective than only using one method, although these methods (eg, home visits, telephone calls, case management) are resultantly more costly compared with the outreach channels deployed in this study [22]. To the best of our knowledge, the study represents the first to investigate the effectiveness of combining multiple low-intensity outreach channels for this use case.

Limitations

There are several limitations to this study. First, the study population is limited to Medicare Advantage members of one large national insurer in the United States with geographic concentrations in the Northeast and Midwest regions. A majority of members in this study self-report primary residence in rural areas (n=189,594 [51%]); notably higher than the average US population (14%) [23], which may limit generalizability to the broader US population. It is possible that differences in access to broadband internet and digital engagement tools may have influenced the effectiveness of electronic messaging, particularly among rural populations. However, the inclusion of multiple outreach channels, including direct mail and provider notifications, helped mitigate these potential disparities. It is difficult to systematically assess differences in access to

XSL•FO
RenderX

technology by geographical location, as multichannel outreach was informed by member channel permissions. In addition, this study was not designed to examine differences in digital technology by geographic location. However, contemporary analysis (Feb 2025) suggests minimal differences in access and preference for digital outreach methods among Medicare Advantage members residing across rural, urban, and suburban locations, with 68% - 69% of members permitting SMS, 75% - 78% permitting email, and 92% - 93% permitting IVR contact across all regions. Interestingly, email open rates (2024 data) indicate slightly lower engagement among rural members (60.5%, $P<.0001$) compared with suburban (64.3%) and urban (62.8%), which may signal potential differences in digital engagement and warrant additional exploration. Nevertheless, this study presents outcomes for a population often underrepresented in the literature and advances our understanding of potential disparities in hospital readmissions.

Second, this study infers readmission and ED visit reductions to be driven by the campaign interventions and does not consider additional (ie, not plan sponsored) interventions. Third, the time horizon may not fully capture the impact of the COVID-19 pandemic on members' behaviors, which continues to evolve. Base rates are unable to be reported due to commercial data use agreements; however, these findings from this large-scale intervention provide valuable insights for designing and implementing future readmission reduction strategies across diverse health care settings. Finally, the possibility always exists that unmeasured confounding variables and potential selection bias may influence outcomes.

Despite the recognized limitations, this study possesses several key strengths. This study represents a retrospective evaluation of a rigorously designed and pragmatic quality improvement intervention. The included study population consisted of a large sample size of 368,693 Medicare Advantage-insured, geographically diverse, older adults across the nation. By focusing on Medicare Advantage beneficiaries, the study addresses a critical population with unique health care needs, contributing to the development of effective strategies to reduce readmissions and improve member and patient outcomes in the population most likely to have gaps in care. The use of claims data enriches the results, providing real-world evidence of outcomes. The integration of additional member data, including plan benefit structure and overlapping programs, enhances the study's precision and effectiveness in minimizing confounding bias. Finally, the campaign's grounded approach in evidence-based practice and health behavior theoretical constructs reinforces the robustness of the intervention and its potential applicability to the generalized population.

Future Research

There are several opportunities for additional exploration. Immediate future research will aim to explore the incremental effect of additional channels. Interestingly, post hoc analysis within the enhanced outreach group revealed that members who received more channels experienced fewer ED admissions than those receiving fewer channels. Specifically, members who received 4 outreach channels had fewer 30-day acute inpatient readmissions and ED visits than members receiving campaigns via one outreach channel, indicating an additive effect of incremental channels ($P<.001$). While interesting, we caution that these results are preliminary and warrant additional evaluation in a controlled, hypothesis-driven study.

In addition, investigating the impact of low-intensity channels based on specific sub-populations, diseases or conditions, and health care settings can provide targeted insights for tailored interventions. As previously mentioned, incorporating more granular data on channel engagement by geographic location could potentially refine multichannel strategies to optimize outreach effectiveness, particularly in rural populations. Future explorations should explore whether reliance on postcharge digital messaging could unintentionally widen the "digital divide" or disparities in access to care. Additionally, exploring best practices in human-centered design that increase member satisfaction, engagement, and resultant clinical and economic outcomes will further optimize and elucidate campaign effectiveness.

Conclusions

Payer-led, personalized educational messaging using multiple low-intensity channels of delivery, including digital health communications such as text message, email, and IVR, can effectively reduce inpatient readmissions and ED visits within 30- and 90-days of discharge among the Medicare Advantage member population.

Acknowledgments

The authors wish to thank Laure Salomon, Josh Weiner, Ali Keshavarz, and Rahul Kak for making this study possible; Shahin Taghikhani and Michael Kolor for their data science support; Szu-Min Yu, Maegan North, Nick Mcburney, Adam Sharaf, Kate Cao, Alexis Cason, Isabella (Ang) Li, Kratika Agrawal, Laurence Chalude, and Carly Levin for their marketing and data engineering support; Dawn Merris, Kirsten Sanderson-Cornelius, Dr Michael Raymond, and Dr Christine Lawless for their business and clinical support; and Dr Dorothea Verbrugge and Mohammad Emami for their publication expertise and guidance. This study was funded by CVS Health Corporation.

Data Availability

The data sets generated or analyzed during this study are not publicly available due to commercial data use agreements.

XSL•FO
RenderX

Conflicts of Interest

All authors of this publication are currently or previously employed by CVS Health Corporation. KH and BH own stock in the company, and EK owns equity in the company. DR, DF, ALZ, KJTC, and ABS own both stock and equity in the company. LV and JB were formerly employed by CVS Health Corporation and received stock, equity, or both.

Multimedia Appendix 1
Screenshot displaying enhanced preadmission messaging sent to members via direct mail channel prior to an upcoming hospital stay. Preadmission outreach content focused on priority actions to take in advance of a hospital stay, with emphasis on preparation for the eventual return home (eg, nutritious meal planning, removal of fall risk hazards), scheduling follow-up visits prior to admission, and the importance of medication filling and adherence).
[PNG File, 2608 KB - humanfactors_v12i1e63841_app1.png]

Multimedia Appendix 2
Screenshot displaying enhanced postdischarge messaging sent to members via direct mail following hospital discharge to home. Postdischarge outreach content focused on a member's individual recovery journey by providing a "recovery tracker" calendar for personalized tracking of pain levels, upcoming appointments, and any other key notes to share with their care team at their next follow-up visit. The content reinforced the importance of going to follow-up visits, filling prescriptions and taking them as directed, as well as common warning signs suggestive of a potential complication that would warrant care.
[PNG File, 1218 KB - humanfactors_v12i1e63841_app2.png]

References

1. Weiss AJ, Jiang HJ. Overview of clinical conditions with frequent and costly hospital readmissions by payer, 2018. In: Healthcare Cost and Utilization Project (HCUP) Statistical Briefs: Agency for Healthcare Research and Quality (US); 2021. URL: https://pubmed.ncbi.nlm.nih.gov/34460186/ [Medline: 34460186]
2. Hospital readmission reduction program. : Centers for Medicare & Medicaid Services; 2022 Sep 9 URL: https://www.cms.gov/medicare/payment/prospective-payment-systems/acute-inpatient-pps/hospital-readmissions-reduction-program-hrrp [accessed 2025-05-12]
3. Jack BW, Chetty VK, Anthony D, et al. A reengineered hospital discharge program to decrease rehospitalization: A randomized trial. Ann Intern Med 2009 Feb 3;150(3):178-187. [doi: 10.7326/0003-4819-150-3-200902030-00007] [Medline: 19189907]
4. Thomsen K, Fournaise A, Matzen LE, Andersen-Ranberg K, Ryg J. Does geriatric follow-up visits reduce hospital readmission among older patients discharged to temporary care at a skilled nursing facility: a before-and-after cohort study. BMJ Open 2021 Aug 13;11(8):e046698. [doi: 10.1136/bmjopen-2020-046698] [Medline: 34389564]
5. Flynn J, Marino M, Fields S. Reducing hospital readmissions through primary care practice transformation: this study found that a "culture of continuity" using processes that strengthen outpatient-inpatient caregiver communication improves patient care. J Fam Pract 2014;63(2).
6. Kripalani S, Chen G, Ciampa P, et al. A transition care coordinator model reduces hospital readmissions and costs. Contemp Clin Trials 2019 Jun;81:55-61. [doi: 10.1016/j.cct.2019.04.014] [Medline: 31029692]
7. Ning J, Huang X. Response-adaptive randomization for clinical trials with adjustment for covariate imbalance. Stat Med 2010 Jul 30;29(17):1761-1768. [doi: 10.1002/sim.3978] [Medline: 20658546]
8. Ergonomics of human-system interaction—part 210: human-centred design for interactive systems. International Organization for Standardization. 2019. URL: https://www.iso.org/obp/ui/#iso:std:iso:9241:-210:ed-2:v1:e [accessed 2025-05-05]
9. Harte R, Glynn L, Rodríguez-Molinero A, et al. A human-centered design methodology to enhance the usability, human factors, and user experience of connected health systems: A three-phase methodology. JMIR Hum Factors 2017 Mar 16;4(1):e8. [doi: 10.2196/humanfactors.5443] [Medline: 28302594]
10. Hernandez AF, Greiner MA, Fonarow GC, et al. Relationship between early physician follow-up and 30-day readmission among medicare beneficiaries hospitalized for heart failure. JAMA 2010 May 5;303(17):1716-1722. [doi: 10.1001/jama.2010.533] [Medline: 20442387]
11. Hume K, Tomsik E. Enhancing patient education and medication reconciliation strategies to reduce readmission rates. Hosp Pharm 2014 Feb;49(2):112-114. [doi: 10.1310/hpj4902-112] [Medline: 24623862]
12. Johnson A, Guirguis E, Grace Y. Preventing medication errors in transitions of care: a patient case approach. J Am Pharm Assoc (2003) 2015;55(2):e264-e274. [doi: 10.1331/JAPhA.2015.15509] [Medline: 25749270]
13. Peikes D, Chen A, Schore J, Brown R. Effects of care coordination on hospitalization, quality of care, and health care expenditures among medicare beneficiaries: 15 randomized trials. JAMA 2009 Feb 11;301(6):603-618. [doi: 10.1001/jama.2009.126] [Medline: 19211468]
14. Rask KJ, Hodge J, Kluge L. Impact of contextual factors on interventions to reduce acute care transfers II implementation and hospital readmission rates. J Am Med Dir Assoc 2017 Nov 1;18(11):991. [doi: 10.1016/j.jamda.2017.08.002] [Medline: 28967602]

XSL•FO
RenderX

15. Fu BQ, Zhong CC, Wong CH, et al. Barriers and facilitators to implementing interventions for reducing avoidable hospital readmission: systematic review of qualitative studies. Int J Health Policy Manag 2023;12:7089. [doi: 10.34172/ijhpm.2023.7089] [Medline: 37579466]
16. Gardella JE, Cardwell TB, Nnadi M. Improving medication safety with accurate preadmission medication lists and postdischarge education. Jt Comm J Qual Patient Saf 2012 Oct;38(10):452-458. [doi: 10.1016/s1553-7250(12)38060-4] [Medline: 23130391]
17. Freedman JL, Fraser SC. Compliance without pressure: the foot-in-the-door technique. J Pers Soc Psychol 1966 Aug;4(2):195-202. [doi: 10.1037/h0023552] [Medline: 5969145]
18. Michie S, Wood CE, Johnston M, Abraham C, Francis JJ, Hardeman W. Behaviour change techniques: the development and evaluation of a taxonomic method for reporting and describing behaviour change interventions (a suite of five studies involving consensus methods, randomised controlled trials and analysis of qualitative data). Health Technol Assess 2015 Nov;19(99):1-188. [doi: 10.3310/hta19990] [Medline: 26616119]
19. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6(42):42. [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
20. Bressman E, Long JA, Honig K, et al. Evaluation of an automated text message-based program to reduce use of acute health care resources after hospital discharge. JAMA Netw Open 2022 Oct 3;5(10):e2238293. [doi: 10.1001/jamanetworkopen.2022.38293] [Medline: 36287564]
21. Patel MS, Patel N, Small DS, et al. Change in length of stay and readmissions among hospitalized medical patients after inpatient medicine service adoption of mobile secure text messaging. J Gen Intern Med 2016 Aug;31(8):863-870. [doi: 10.1007/s11606-016-3673-7] [Medline: 27016064]
22. Fønss Rasmussen L, Grode LB, Lange J, Barat I, Gregersen M. Impact of transitional care interventions on hospital readmissions in older medical patients: a systematic review. BMJ Open 2021 Jan 8;11(1):e040057. [doi: 10.1136/bmjopen-2020-040057] [Medline: 33419903]
23. Parker K. Demographic and economic trends in urban, suburban and rural communities. : Pew Research Center; 2018 URL: https://www.pewresearch.org/social-trends/2018/05/22/demographic-and-economic-trends-in-urban-suburban-and-rural-communities/ [accessed 2025-05-12]

Abbreviations

ED: emergency department
HRRP: Hospital Readmission Reduction Program
IVR: interactive voice response

Edited by A Kushniruk; submitted 10.07.24; peer-reviewed by E Borycki, MA Jacobs; revised version received 24.02.25; accepted 30.03.25; published 27.05.25.

Please cite as:
Fernandes D, Kokonas E, Bansal J, Hayashima K, Hurley B, Ryu A, Mhatre S, Ghori M, Craig KJ, Zaleski AL, Vogel L, Baquet-Simpson A, Reif D
An Informatics-Based, Payer-Led, Low-Intensity Multichannel Educational Campaign Designed to Decrease Postdischarge Utilization for Medicare Advantage Members: Retrospective Evaluation
JMIR Hum Factors 2025;12:e63841
URL: https://humanfactors.jmir.org/2025/1/e63841
doi:10.2196/63841

© Danica Fernandes, Elise Kokonas, Jai Bansal, Ken Hayashima, Brian Hurley, Annabel Ryu, Snehal Mhatre, Mohammed Ghori, Kelly Jean Craig, Amanda L Zaleski, Lily Vogel, Alena Baquet-Simpson, Daniel Reif. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 27.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring Youth Perspectives on Digital Mental Health Platforms: Qualitative Descriptive Study

Sarah Daniel[1]; Lauren Volcko[1], BMSc; Emilie Bassi[1], MSc; Julia Hews-Girard[2,3,4,5], RN, PhD; Katherine Bright[3,6,7,8,9,10], RN, MN, PhD; Marianne Barker[1], PhD; Lia Norman[1], BA; Karina Pintson[1]; Geneca Henry[1], BA; Sumaya Soufi[1], BS; Chukwudumbiri Efrem Omorotionmwan[1], BS; Melanie Fersovitch[11], BN; Leanne Stamp[11], B.Ed; Karen Moskovic[11], BN; David W Johnson[2,12], MD; Gina Dimitropoulos[1,2,4,5], MSW, PhD

[1]Faculty of Social Work, University of Calgary, MacKimmie Tower, 2500 University Dr NW, Calgary, AB, Canada

[10]Recovery Alberta, Women's Mental Health Clinic, Edmonton, AB, Canada

[11]Alberta Health Services, Edmonton, AB, Canada

[12]Department of Pediatrics, Cumming School of Medicine, Calgary, AB, Canada

[2]Alberta Children's Hospital Research Institute, Calgary, AB, Canada

[3]Faculty of Nursing, University of Calgary, Calgary, AB, Canada

[4]Mathison Centre for Mental Health Research and Education, Calgary, AB, Canada

[5]Hotchkiss Brain Institute, Calgary, AB, Canada

[6]School of Nursing and Midwifery, Mount Royal University, Calgary, AB, Canada

[7]School of Nursing, Thompson Rivers University, Kamloops, BC, Canada

[8]Faculty of Rehabilitation Medicine, University of Alberta, Edmonton, AB, Canada

[9]Heroes in Mind, Advocacy and Research Consortium, Edmonton, AB, Canada

Corresponding Author:
Gina Dimitropoulos, MSW, PhD
Faculty of Social Work, University of Calgary, MacKimmie Tower, 2500 University Dr NW, Calgary, AB, Canada

Abstract

Background: The increasing prevalence of mental health disorders among youth underscores the need for accessible and effective interventions. Digital mental health (dMH) platforms like Innowell offer promising solutions by increasing access to mental health care for young people. Innowell is a web-based platform that supports youth mental health by providing personalized measurement-based care in collaboration with a youth's health care providers. However, understanding youth perspectives on these platforms is crucial for ensuring successful implementation and sustained engagement.

Objective: This study aimed to explore youth perspectives on the implementation of the Innowell platform, identifying key factors influencing uptake, engagement, and long-term retention.

Methods: A qualitative descriptive approach was used to examine youth perspectives. Data were collected through 9 focus groups and 1 interview, involving 39 participants aged 15‐24 years from urban (23/39, 59%) and rural (16/39, 41%) communities in Alberta, Canada. Participants were recruited through mental health clinics and community organizations. Thematic analysis was conducted on the transcripts to identify factors that support or hinder engagement with the platform.

Results: Participants emphasized the importance of privacy, security, and personalization in building trust in the platform, with 72% (28/39) reporting that clear communication about data protection would increase their likelihood of use. Progress tracking features, such as symptom trend visualizations and diaries, were identified by 65% (25/39) of participants as critical for sustaining engagement. Ease of use was highlighted, with 58% (23/39) preferring mobile app functionality over web-based interfaces. Dynamic content and personalized notifications were suggested as strategies to maintain long-term use, with 64% (25/39) of participants valuing customizable reminders to encourage daily interactions. Rural participants (16/39, 41%) noted the need for offline functionality due to inconsistent internet access. In addition, participants recommended features such as crisis support, professional communication channels, and access to local mental health resources.

Conclusions: Youth-centered design is essential for enhancing the usability and engagement of dMH platforms like Innowell. Key features prioritized by participants included privacy, security, progress tracking, and personalization. Dynamic and user-friendly interfaces, along with the ability to customize notifications and access professional support, were critical for fostering long-term engagement. Insights from this study provide actionable recommendations for optimizing dMH platforms to meet the mental health needs of young people, particularly in diverse urban and rural settings. Future research should explore implementation strategies tailored to specific user demographics to enhance the scalability and impact of dMH interventions.

XSL•FO
RenderX

(JMIR Hum Factors 2025;12:e69907) doi:10.2196/69907

KEYWORDS

digital mental health; youth mental health; youth perspectives; engagement and retention; trust and privacy; personalization

Introduction

The number of youth and young adults affected by mental health disorders has increased in the last 10 years [1,2]. Adolescence and early adulthood, ages 12 - 24 years [3], is a peak period of vulnerability for mental ill-health, with 62.5% of mental health disorders emerging before the age of 25 [4]. Mental health conditions of concern include depression, anxiety, eating disorders, bipolar disorder, and substance abuse disorders among others [1,2,4]. Globally, mental health disorders are the leading cause of disability in youth and are the sixth leading cause of disability adjusted life years [5]. The number of youths underdiagnosed or undertreated remains a significant global issue due to barriers to accessing youth mental health services [6,7]. This involves social stigma, limited awareness of available mental health resources [7], and concerns about the therapeutic relationship, including confidentiality and privacy [8,9].

Digital mental health (dMH) has the potential to address the treatment gap of those who do not receive services and increase accessibility of mental health care to youth and young adults [10,11]. The dMH includes both mobile apps and web-based platforms [12] that can augment in-person and traditional mental health services. Youth tend to be more familiar with the digital world and use the internet at progressively younger ages [13]. Given that social interactions and educational commitments are increasingly taking place through digital modes, youth are becoming more inclined to seek mental health support through the internet [14-16]. Research suggests that stigmatizing beliefs surrounding mental illness drive many youth to self-initiate their mental health care by accessing the internet for educational information or assistance regarding mental health concerns, because it is perceived to be more confidential and private than face-to-face services [17]. Of youth aged 13 - 17, approximately 22.2% use internet-based services to seek information regarding behavioral issues [17]. Therefore, there is a growing need for youth-oriented dMH platforms to address some of the challenges of traditional face-to-face mental health care [18,19].

Youth and young adult perspectives on dMH platforms are critical to facilitate retention and uptake of these platforms [20]. Including youth perspectives in dMH research ensures that platforms are accessible, relevant and align with their unique needs and experiences [20,21]. A qualitative study conducted by McKenna et al [22] found that youth valued measurement-based care when embedded in a digital platform and used in their sessions with their therapists because it helped them track their health information between appointments. Generally, dMH platforms have been noted by youth to increase accessibility to mental health care [23], however, the information presented on these platforms must be easy to understand [23]. Youth also noted that dMH may aid self-determination and promote anonymity when asking for mental health support [23].

Despite the benefits of dMH for youth, knowledge regarding the reasons youth use and continue to engage with these platforms is limited [24]. To date, research suggests that reasons for lack of engagement in dMH platforms include limited in-person elements, for example, face-to-face support from a parent, peer or professional [13], privacy concerns [25], technical difficulties [25], and inadequate personalization of the platform [25,26].

A 2019 systematic review and meta-analysis assessing the effectiveness of dMH platforms for youth with anxiety and depression revealed that out of the 34 identified interventions reported, none involved young people in the design process [27]. In contrast, a 2020 review identified 30 platforms that did involve youth design [28]; however, the impact of this has not yet been fully explored as few are widely available for use [29]. One example of a codesigned, personalized, and youth-centered dMH platform that is available in Australia and Canada is Innowell [10,30-32]. The perspectives of youth on the facilitators and challenges in using dMH to access mental health resources and as an adjunct to mental health interventions are scarce.

This study seeks to ascertain the perspectives of youth and young adults regarding the potential barriers and facilitators of using dMH. Understanding what youth expect from a dMH platform can inform recommendations for implementation. To that end, this project asked the following research question: what influences youth engagement with and sustained retention of digital mental health platforms?

Methods

Ethical Considerations

This study is part of a larger multimethod research project aimed at investigating the implementation of a measurement-based care dMH platform for Albertan youth and young adults and was approved by the University of Calgary's Research Ethics Board (REB20-1137) [30,31]. Informed consent was obtained from all participants, and participants under 18 years were required to complete a decision-making capacity assessment to not require parental consent. Participants received instruction to avoid disclosing identifying information, and during transcription any direct identifiers (names, cities, and workplaces) were removed to protect participants' privacy. Focus group participants were compensated for their time at a rate of CAD $25 (US $18.07) per hour.

Recruitment

Recruitment occurred from November 2020 to June 2022. Participants were recruited using self-selected sampling, where participating communities and Innowell testing sites displayed and shared study recruitment posters. Project Leads coordinated with leaders in partnering communities to display the recruitment poster in mental health clinics and community

XSL•FO
RenderX

spaces. In some cases, clinicians discussed the opportunity to participate in the study with their clients aged 15 - 24 years. Interested youth then emailed the study coordinator to express their interest in participating. Inclusion criteria for participants required being between the ages of 15 to 24 years, proficient in written and spoken English, residing in Alberta, and finally, able to use web-based technologies on a computer, tablet, or smartphone. Interested participants were emailed the date and time of the focus group, along with a link to complete the informed consent form and a brief demographics survey to assess participant characteristics. In accordance with the research ethics protocol, a decision-making capacity assessment was administered to participants under the age of 18 years. A perfect score of 4 out of 4 was required to be determined a mature minor, for which parental consent would not be required. If a participant scored lower than 4/4, the missed items were reviewed and retested with a limit of 3 total attempts. Participants without a perfect score after 3 trials were determined to not be mature minors and were ineligible to participate without parental consent. Once the completed consent form was received, the participants were sent a link to join the focus group via Zoom Video Conferencing Software.

Data Collection

Focus groups were conducted as the primary source of data collection. One interview was conducted at the request of the participant. A semistructured interview guide was created to focus on general thoughts and opinions on the Innowell platform and what would support or hinder youth from using it in their community. All focus groups and interviews were facilitated virtually via Zoom Video Conferencing software, audio recorded, and transcribed verbatim by a professional transcriptionist [30,31]. Each focus group lasted up to 90 minutes in length and included 2 to 8 participants. The focus groups and interview included one facilitator and 1 youth research partner note-taker or chat moderator. Participants were first introduced to Innowell through a short video and demonstration of the platform before the focus group discussion. In total, 9 focus groups and one interview were conducted with 39 participants (n=39). Participants represented both urban and rural communities across the province of Alberta, Canada. Sample characteristics are in Table 1.

Table . Demographic characteristics of the participants (N=39) were analyzed.[ab]

		Values
Sex, n (%)		
	Female	25 (64.1)
	Male	7 (17.9)
	Missing	7 (17.9)
Sexual orientation, n (%)		
	Heterosexual	25 (64.1)
	LGBTQIA2S+	10 (25.6)
	Missing	4 (10.2)
Age, n (%)		
	15 - 17 years	21 (53.8)
	18 - 24 years	16 (41.0)
	Missing	1 (2.5)
Ethnicity[c], n		
	Other North American origins	7
	European origins	21
	Asian origins	5
	Other	6
Student or employment status, n (%)		
	Employed	22 (56.4)
	Student	29 (74.3)

[a]"Other" includes sociodemographic information for fewer than five participants to protect anonymity.

[b]Missing includes "prefer not to answer" and "missing" responses. These are combined for clarity and confidentiality.

[c]Values for ethnicity do not equal 100% as participants were allowed to select multiple responses.

XSL•FO
RenderX

Data Analysis

This study used a qualitative descriptive methodology to thoroughly capture the thoughts, perspectives, and experiences of participants who have experiences in health care settings [33,34]. This allowed for the exploration of youth perceptions of challenges and enablers of dMH implementation that are grounded in their lived-experiences of navigating mental health within their communities.

The qualitative analysis of the focus group and interview data was conducted using a 6-stage thematic analysis process: familiarization with the data, generating codes, constructing themes, revising themes, defining themes, and producing the report [30,31,35]. This process followed an iterative codebook approach, allowing for the development of consistency with the generation of codes across focus groups [35]. Three young adult researchers familiarized themselves with the transcripts. They then independently created preliminary codes with support and oversight from the qualitative lead. After coding, they exchanged transcripts and collaborated to resolve any discrepancies, ensuring the validity of the codes through consensus. The coders and the research team then developed themes from the codes. The team engaged in group discussions about the themes, actively documenting their insights to maintain a rigorous and transparent process. Following these 6 steps enhanced the accuracy and credibility of the analysis [30,31]. The sample of 9 focus groups and 1 interview was deemed sufficient based on the concept of "information power," which suggests a qualitative sample's quality relies on its richness, relevance, and depth rather than the number of participants. This approach ensures ample information was gathered to draw meaningful conclusions from the content of the data [30,31,36].

Research Team Positionality

Our team comprises qualitative leads with extensive experience in both research and frontline clinical support, bringing a wealth of knowledge about the complexities of health care settings and a deep understanding of how to effectively support youth. The study team also included youth and young adult coresearchers throughout the entirety of the study, including recruitment, data collection and analysis, and writing the research findings. Youth coresearchers contributed to the research process and offered guidance and knowledge in engaging youth participants. The study team also focused on the power dynamics inherent in conducting focus groups with youth. Recognizing these dynamics, the study team prioritized creating a safe and supportive environment by including a youth coresearcher in every session. The study team also actively encouraged youth to choose their level of engagement, including whether to keep their videos on or off or use the chat, ensuring that their comfort and agency were prioritized in the research process. Through these strategies, we aimed to foster a more equitable dialogue and enhance the validity of our findings.

The Innowell Platform

The dMH platform, Innowell, is an Australian web-based tool youth use in conjunction with their mental health providers [37-39]. Specifics regarding the Innowell platform are described elsewhere [30-32,38,39]. Previous work from the research group has explored the implementation of this platform in school settings, mental health providers' perceptions of using and engaging dMH with youth, and the creation process of an implementation protocol for using dMH in stratified care.

Results

Demographics Summary

The study included participants from various communities across Alberta (withheld to protect anonymity). Participants were aged 15 - 17 years (n=21) and 18 - 24 years (n=16), with one missing response for age. Regarding sex, 7 participants identified as male, 25 as female, and 6 did not report. Sexual orientation was diverse, with most participants identifying as straight (n=25) and 10 as Lesbian, Gay, Bisexual, Queer or Questioning, Intersex, Asexual, 2-Spirited, and other orientations (LGBQIA2S+), combined to protect anonymity. Ethnic backgrounds included European origins (n=21), Other North American (n=7), Asian origins (n=5), six as other or unsure. Regarding employment or student status, 22 were employed and 29 were students. For a full overview of the demographics, see Table 1.

Theme 1: Building Trust—A Digital Safe Space

Confidentiality, Privacy and Data Security

Participants in the study agreed that youth valued and trusted a dMH platform that would grant them the "comfort of …confidentiality and being up front and honest," (FG2). By improving confidentiality and privacy, dMH enables a safe space for youth disclosure of mental health challenges and requests for support. A highly secure and confidential dMH platform would enable youth to access mental health support by reducing the stigma associated with seeking in-person mental health care. One participant stated:

> *There's always that stigma around mental health …[by] having it so that you don't have to be right in-person with the doctor is just a lot easier for people to maintain confidentiality and to be more expressive.* [FG6]

Participants across all focus groups emphasized a deciding factor in their use of dMH would be the strength of data privacy and security features determining how their information is used and stored. As illustrated by one participant, "You have to remind the people that are using this (dMH) that you're not collecting their data" (FG9). Another participant indicated, "I would want to know who can view the information, where the information is going etc," (FG4). To build trust with youth, participants strongly advocated that dMH must clearly state how user data is used and protected. Factors that contributed to building trust in dMH security measures included the appearance and professionalism of the platform. Participants identified that commercial advertisements decreased the platform's appeal, with one stating:

> *I would be hesitant to download that app knowing that I could still be targeted with ads or that information could be leaked to third parties and trying to sell me something with it.* [FG9]

XSL•FO
RenderX

The dMH was seen as having the potential to build safe spaces for youth disclosure if "you have that comfort of some sort of confidentiality and being upfront and honest" (FG2) when providers are onboarding youth to use dMH in their care.

Keeping "Control Over the Situation": Access to Information and Involvement of Others

Participants frequently reported that youth must have jurisdiction over who accesses their personal information on dMH, including providers and parents or guardians. Many participants advocated for reassurance that their information was "protected, nobody else is going to see [their] information [and] it's not going to be shared with anybody outside of [their] clinician and [themselves] or [their] support person" (FG6).

To facilitate this level of privacy, some participants identified "certain healthcare providers that [they] would not want involved in [their] mental health care," and suggested that youth are "able to pick the adults that they were going to be disclosing information to" on dMH platforms (FG1). All participants noted that youth must have control over the level of parent or guardian involvement in dMH. In fact, some participants who reported having a conflictual relationship with their parent or guardian were less inclined to use dMH if family members had any involvement in their care. Overall, youth wanted "control over the situation" to defend their confidentiality and privacy: "...if someone is being added they know they can say, 'Hey, I don't want this person here'" (FG6).

Theme 2: Platform Features to Maximize Youth Engagement

Onboarding or Getting Started With dMH

The participants agreed that youth must be onboarded to dMH to increase digital literacy and trust and generate enthusiasm for dMH's utility in their care. Once receptive to the idea of dMH, participants expressed a preference to receive collaborative onboarding by a provider or dMH expert. One participant imagined receiving a "thorough" walk-through in which they receive a "step-by-step explanation of what you would do [led by] the professionals behind the app" (FG4). Other participants emphasized the need to understand the purpose and the features of dMH to increase the likelihood of engaging with dMH applications. Participants suggested that the onboarding process to dMH be educational, informing youth of "the why behind using it [as] people lose interest when there isn't much detail explained" (FG4).

Progress Tracking

Participants frequently reported the value of tracking their mental health within dMH, including visualizing trends through graphs. One participant noted:

> *I know for myself, I'm really interested in data and trends about myself and my health and being able to broaden that to my mental health is kind of a cool way to introduce it.* [FG9]

Relatedly, a participant noted that seeing progress through dMH tracking may be motivating when a change in their mental health is not perceptible on a day-to-day basis to them and others in their lives.

> *It is about getting better, but it's not always going to be about seeing this dramatic increase and having this weight lifted off your shoulders. I think being able to document things like, "What happened to me today?" Or "how's this week been?" Or "how am I feeling or dealing with whatever's going on?"* [FG6]

A platform feature allowing youth to store and detail personal information, including symptoms and important goals, can serve as an invaluable tool for tracking progress . One participant expressed, "Maybe there could be areas where you could store things and keep things, maybe make a diary page or something" (FG2). This feature would enable users to document their thoughts, feelings, and experiences over time, facilitating self-reflection and self-awareness. Overall, participants noted that they would use dMH regularly if it served to track and enable reflection on their mental health progress and fluctuations in symptoms.

User Experience

Participants throughout the focus groups valued user-friendly design of dMH, including ease of navigation and accessibility. One participant described a facilitator to consistent dMH use as being if "the dashboard itself is easy to understand and navigate" (FG9). Participants expressed a clear preference for an app, or mobile-based, interface compared to a web-based one: "I think it should be an app in the App store because if that is an app in the App Store, it would be accessible to everyone" (FG1). Participants indicated it would increase the chances of discovering dMH, as "youth "looking for meditation [applications] or…games [might] come across it" (FG1). In addition, participants saw potential in the dMH interface facilitating access to mental health education and resources. One participant remarked:

> *It could give you helpful information to make your mental health better and to give you more activities to help you and places to access phone numbers or places to contact if you need help. That would be amazing.* [FG1]

Other participants suggested that dMH could direct youth to local mental health resources and crisis hotlines.

The ability to customize the platform's features and resources to suit a youth's unique needs was also desirable to participants. A participant suggested users be granted the ability to "edit when notifications are sent, when I can talk to someone, who I can talk to and more would be great so it won't be stressful and I can run on my own time"(FG2). A personalized dMH could cater to a wider "variety of people and it can connect [them] with the right resources" (FG4). Another suggested utility of dMH included direct communication with a professional, who "could reach out to you through the app. You could maybe have a personal chat area" (FG2). Other participants expressed interest in dMH platforms with specific features supporting youth in crisis.

> *I think having emergency functions be very accessible would be important as well. Something like a 'chat*

XSL•FO
RenderX

> *now' function that's right on the home screen or close to it so that if someone is in crisis, they can access the help they need.*

This highlights the critical need for in-app support for urgent situations, accessible and emergency functions, complementing the platform's diverse set of resources, including educational materials, links to available mental health services and crisis hotlines.

Theme 3: Sustaining Engagement—Strategies for Optimum User Retention

Barriers to Consistency: The Challenge of Integration

Across all focus groups, participants identified barriers to integrating dMH into their lives. One participant noted, "For me personally, using apps in the past, after a certain point you just kind of get bored of it" (FG6). Another difficulty was prioritizing dMH in the face of other pressing demands and daily routines: "I think that I've had mental health apps in the past, but it's just remembering to do it and having it as a priority it was really tough…." (FG7). Participants identified competing online priorities such as social media as a barrier to using dMH: "Why would I use a mental health app when I can go on Instagram instead. There's nothing enticing me to use the app" (FG6), and "Social media gets in the way of using mental health apps sometimes" (FG3). Overall, these quotes underscore the necessity for dMH platforms to be engaging, easy to integrate into daily routines, and compelling enough to compete with other online activities to ensure sustained user retention.

Suggestions to Reduce Distractions and Maintain Attention

An additional strategy to encourage motivation and continued use of dMH mentioned within the focus groups was the integration of reminder notifications. One participant stated:

> *...if I'm getting notified all the time to maybe update or maybe there's like a quote of the day every day that it'll send me… I think that would be very useful, and something very appealing to the eye.* [FG2]

Another participant provided an example of how notifications and prompts promoted reflection and participation with the platform.

> *I have the app "Calm" on my phone … you can get a notification at the end of the day asking, "How are you feeling at the end of the day?" Sometimes, a lot of the time I just swipe up and sometimes I actually take the time to think about how I'm actually feeling. It's nice for the most part.* [FG3]

However, other participants did not find consistent prompts necessary:

> *Some apps I know I have deleted before if I have set the notifications on and they are just constantly sending me notifications to the point where it gets annoying.* [FG2]

Incorporating timely and personalized notifications could, therefore, play a role in keeping young users engaged and motivated to consistently use dMH platforms, as suggested by one participant.

> *If you check your phone, you see the app which reminds you to do your daily check in or maybe a reminder to do it later in the day around lunchtime or afternoon when you accomplish things or have things that affect your mood. I think that would be very important and useful for people.* [FG2]

The ability for dMH platforms to incorporate additional elements to increase engagement was valued by participants, reflected in a broader sentiment that applications lacking new content tend to be forgotten: "If an app never has anything new on it or different things to try, I find I forget about the app after a while" (FG2). Likewise, one participant noted the power of rewards to attract attention to the dMH platform.

> *People are very motivated by rewards and attention as well. Because like social media—even though you're not getting a specific reward, you're getting attention from others, you're getting communication and it kind of feeds you … so, I think if either some kind of award or like "Congrats you've made it this far! Here's like a little boost of encouragement," or something like that. People like that—that positive reinforcement.* [FG6]

Therefore, incorporating timely, personalized notifications can significantly enhance the appeal and retention of dMH platforms, ensuring they remain a valuable and consistently used resource for young users.

Discussion

Principal Findings

This study aimed to explore youth perspectives on optimizing the implementation of dMH platforms. Three themes were identified from the inductively coded data. The first theme emphasizes the importance of trust in dMH, protecting youth information and creating a safe digital space. The second theme is the identification of specific platform features to enhance the engagement of youth on the platform. The third theme describes the need for dMH platforms to maintain ongoing engagement of youth by removing obstacles to consistent use of the dMH platform and highlighting strategies to reduce user fatigue and boredom.

Theme 1: Building Trust—A Digital Safe Space

Our findings suggest that transparency in a dMH's privacy and confidentiality protections is essential to build youth trust. Other research supports confidentiality's importance for youth across both in-person and digital mental health care settings [8]. Further supporting our results, youths' knowledge of a platform's privacy and security features was identified as a critical factor in increasing engagement with dMH [40]. Despite confidentiality measures contributing to youths' trust in a dMH platform, a study identified this as insufficient for long-term, sustainable engagement from youth [18].

Intertwined with their emphasis on confidentiality, participants valued the power and autonomy to restrict who can see their

XSL•FO
RenderX

mental health information on the dMH platform. Youth most expressed concern about parents and guardians being involved in their dMH care. The findings from our study underscore the critical importance of maintaining youth autonomy and control within dMH platforms, including the ability to selectively involve providers, parents, or guardians. Consistent with our findings, in a preimplementation survey of youth interest in a dMH platform, parental consent was identified as a deterrent for youth participation in dMH [41]. Contrasting to our findings, a study investigating youth attitudes toward parental involvement in dMH reported that youth favored parent access to the dMH patient portal, predicting it would increase accountability and engagement, and enable provider follow-ups [42]. The difference in results may stem from the setting: their study focused on clinical care, where parental involvement is more common, whereas our study was based in community settings, potentially encouraging a greater emphasis on independence and privacy. Our findings suggest that to maximize engagement and effectiveness, dMH platforms should allow users to decide who has access to their data, including the flexibility to invite parents or guardians when it aligns with the youth's preferences and needs.

In line with our finding that that youth value security and privacy measures on dMH platforms, dMH platforms should have robust data security measures to reinforce confidentiality promises. Supporting our findings, a review article identified confidentiality as a "key ethical concern" for users of all ages on dMH platforms [43]. The emotional and psychological effects of data breaches on victims are long-reaching and can include a loss of confidence in the related technology, acute symptoms of posttraumatic stress disorder, and even learned helplessness [44]. Our participants wanted to be informed about dMH's security and privacy measures to build trust in the platform. More investigation is necessary, but dMH platforms should maintain user trust by clearly communicating the security features present on the platform.

Theme 2: Platform Features to Maximize Youth Engagement

Our findings suggest youth prioritize their autonomy when choosing to use dMH platforms but are more inclined to engage if recommended by a trusted individual, including a provider or peer. However, some youth expressed a general resistance to any recommendations about dMH, valuing their own decisions about whether to use dMH. These contrasting findings suggest that eliciting youths' consideration about using dMH must be explored to minimize the risk of rejection and resistance to dMH. Inquiries about dMH from trusted individuals may be followed by youth considering dMH.

Once committed to using a dMH platform, our findings suggest that a positive onboarding experience can facilitate positive user engagement. Youth participants in this study recommend that educational information about dMH and a collaborative approach to engaging youth is critical for increasing the continuous use of dMH. Similar to our findings, a 2023 scoping review of therapeutic dMH platforms for schizophrenia identified thorough initial training including, educational sessions and a hands-on support team as a predictor of user engagement levels [45]. To increase trust in reluctant youth, we recommend that dMH platforms facilitate a virtual orientation process that grants youth the opportunity to practice navigating through the platform and educates them about its intent, and security and privacy features.

Our findings indicated progress tracking features such as graphs and diaries could bolster user motivation and thus lead to higher user engagement. It has been found that increased insight into a user's own personal health, along with empowerment, knowledge, and curiosity may increase engagement with the platform [25]. A research article postulates a psychological mechanism for how digital technologies use progress tracking to increase user motivation and engagement [46]. Another paper described the opportunities and challenges of passive and active progress tracking, arguing that the latter is more conducive to habit formation than a passive, automated one [47]. This suggests that active progress tracking, such as the diary feature suggested by participants , can be more effective in fostering user engagement than passive strategies such as data visualization. Some participants identified the potential for progress tracking to be discouraging for youth whose progress is nonlinear or in a crisis. However, although not contributing to user motivation, progress tracking features could bring youth's attention to unconscious trends of improvement, decline, or stagnation in their mental health.

Our findings suggest an intuitive and accessible user interface is conducive to increased engagement with dMH platforms. Other literature supports our findings, including a qualitative study on the dMH platform Artemis; youth testers preferred a "simple, modern, and consistent" user design and "calming" color schemes instead of identifying technical issues as problematic and disruptive to user engagement [48]. User interface design, content quality, and personalization significantly influence youth engagement [49]. Visually engaging and user-friendly interfaces are crucial in encouraging youth to engage with digital health tools. Our findings indicated that youth preferred an application-based delivery over a web-browser-based dMH platform.

In our findings, participants emphasized dMH, including crisis resources such as a "chat now" feature. In a review of dMH applications' features, a stark 1 out of 31 CBT applications targeting depression were identified that included explicit crisis support features, defined as prominent, identifiable, and intended for use during mental health crises. Such features were not interactive, including a detailed crisis plan or prominent links to suicide prevention hotlines or support centres. This limited presence of crisis support in existing dMH applications aligns with the youth's call for such features on dMH platforms [47].

Theme 3: Sustaining Engagement—Strategies for Optimum User Retention

Participants in our findings highlighted several factors that hindered their ability to consistently use dMH platforms, including the challenges of maintaining engagement, integrating usage into daily routines, and competing with other digital distractions. This decline in engagement over time reflects a common issue with adherence to digital health interventions [50-52]. To that end, designing platforms that offer ongoing

XSL•FO
RenderX

value, such as personalized content, may help foster long-term use. Further, nonpersonalized interfaces can result in frustration and disengagement [13]. A systematic review by Opie et al [53] found that personalized feedback primarily delivered through a dMH internal messaging system from providers to youth was linked to a positive increase in user outcomes.

Participants also identified competition with social media as a barrier to using dMH, as youth often use social media to seek and maintain relationships for educational purposes and entertainment [54]. The immediate gratification and diversity of applications offered by social media can divert attention from health-focused applications. However, by incorporating these aspects, platforms promoting social connectedness enabled by social networks through peer support have been shown to increase engagement [25,53,55]. In contrast, research showed that integrating social media may affect the credibility of the dMH platform [53]. To compete, dMH platforms must offer a similarly compelling user experience. The increased face-to-face interactions facilitated by proper and consistent dMH use by providers can similarly engage youth users. Participants emphasized push notifications and dynamic visual and interactive content would optimize youth retention and engagement (Table 2). Consistent with existing research, participants emphasized the need for customizable, well-timed, relevant notifications that fit naturally into their daily routines [26,56]. In addition, the use of rewards or positive reinforcement to motivate individuals was mentioned within focus groups, which is consistent with previous research on persuasive design techniques [52,57]. However, collecting data for personalized notifications may impact users' privacy without youths' informed consent. Research also suggests that the development of notification fatigue, where users feel overwhelmed by constant alerts, can lead to disengagement or deletion of the platform [58]. Similarly, while not youth-specific, notifications promoted the engagement of users who consistently use the platforms rather than inactive users [59].

Our findings can be situated within a Self-Determination Theory framework, which suggests that human behavior is motivated by 3 innate psychological needs: autonomy, competence, and relatedness [60]. Participants emphasized privacy controls and personalization, which align with Self-Determination Theory. Progress-tracking features and user-friendly design can increase youth's sense of competence. The need for relatedness can be linked to the crisis support and provider communication features suggested by youth for dMH. The satisfaction of these psychological needs may facilitate engagement and long-term retention for youth on dMH.

Practical Recommendations for dMH Implementation

Throughout focus groups, youth participants provided recommendations on how to implement and maintain dMH platforms, including building trust by using these platforms, features that promoted engagement and ways to maintain youth attention and engagement throughout dMH adoption. The themes outlined in this paper offered the ability to provide practical recommendations for implementing dMH platforms in clinical settings for youth and young adults dealing with mental health concerns. Table 2 details the recommendations based on our findings. It is important to note that dMH implementation relies on an individual's access to an electronic device and the internet, which may be feasible in low-resource settings.

XSL•FO
RenderX

Table . Recommendations are based on participant focus groups.

Theme and recommendations		Description
Theme 1: Building Trust		
	Ensuring privacy, control, and security	• Prioritize transparent data usage policies and emphasize the platform's commitment to data security to foster trust among youth. • Empower youth by offering customizable privacy settings that allow them to control who accesses their mental health information. Platforms should provide flexibility for users to selectively involve providers, parents, or guardians, ensuring youth needs are met. • Build trust by clearly communicating privacy policies and data security protocols before onboarding. Highlight confidentiality without overemphasizing security features to avoid triggering concerns about privacy invasions. • Strengthen data security measures to protect youth from the emotional impacts of potential data breaches and foster confidence in the platform's safety.
Theme 2: Platform Features for Engagement		
	Streamlined, collaborative onboarding processes	• Implement a comprehensive, step-by-step onboarding process hosted on the digital mental health (dMH) platform and engaging their provider. The onboarding process should be educational, explaining the purpose and key features of the platform to sustain user interest.
	Progress tracking and a user-friendly interface	• Offer customizable tracking tools like graphs and diaries to help youth monitor their mental health and increase autonomy. Active tracking can foster engagement and support self-reflection. • Platforms should emphasize that progress tracking is not just about linear improvement but also about understanding personal growth, even in times of crisis. • Develop simple and intuitive platforms with customizable features. Allow youth to personalize notifications, communication settings, and resources to suit their needs and preferences. • Mobile-based applications are preferred over a web platform for ease of access.
	Crisis support and professional communication	• Include easily accessible crisis resources. Youth mentioned the importance of timely crisis interventions. Facilitating communication with professionals via the platform can provide an added layer of support, making the platform a go-to resource in times of need.
Theme 3: Sustaining Engagement		

XSL•FO
RenderX

Theme and recommendations	Description
Guided dMH[a] platforms	• Ongoing human guidance and support can improve adherence by providing regular check-ins or nudges as well as personalized feedback from a professional to maintain motivation. • Knowing that professional support is available can help youth build confidence when navigating dMH to promote sustained use. • Regular interaction with a professional or trusted individual can promote a sense of accountability, leading to higher engagement rates as opposed to unguided approaches.

[a]dMH: digital mental health.

Strengths and Limitations

The project's strengths included involving youth from different geographic regions in rural and urban communities where Wi-Fi varied in accessibility. We recruited a diverse sample of youth with respect to geography, sector of involvement, gender, and age. Another important strength is that the focus groups were facilitated by a researcher, youth researcher partner, and implementation lead. A summary of the focus group was shared with participants within 24 hours to receive additional reflections and confirm the accuracy of what was captured from the focus groups.

There are several limitations worth noting. The focus groups were conducted during the COVID-19 pandemic, which may have influenced youths' reflections about using online applications and e-tools via dMH. The increased usage of digital platforms during the pandemic period may have made participants more receptive to dMH compared to prepandemic conditions. The pandemic may have also heightened youth's mental health concerns, influencing their expectations and needs from dMH. This could affect the generalizability of our findings in postpandemic settings. However, the research team probed broadly about the youths' perceptions of dMH before the pandemic to obtain reflections that were not dependent on the context at the time. While the youth had a brief explanation and introduction to the platform before each focus group, all were conducted before youth were onboarded or using the dMH; consequently, the feedback about dMH was general rather than specific to a particular platform like Innowell. In addition, only Albertan youth participated in the focus groups, affecting the generalizability beyond our context.

Conclusions

In this paper, youth perspectives on facilitators of the implementation of dMH perspectives were explored. In total, 3 major themes emerged through thematic analysis. The first, "Building Trust," includes strong confidentiality and robust data security measures as crucial for youth trust in a dMH platform. Second, "Platform Features to Maximize Engagement" includes a thorough onboarding process, a simple and intuitive user interface, and progress tracking to engage with youth users. The third, "Sustaining Engagement," centers on strategies like notifications and reward systems to increase the likelihood that users remain users of dMH platforms. The findings informed recommendations for developing dMH to increase youth uptake of dMH platforms, increase engagement with dMH platforms, and ensure that youth stay committed to dMH platforms.

Acknowledgments

The authors gratefully acknowledge the financial and project support provided by the Alberta Children's Hospital Foundation, Alberta Health, the Partnership for Research and Innovation in the Health System (PRIHS), and Canadian Institutes of Health Research through the Operating Grant: SPOR iCT Rewarding Success.

Authors' Contributions

Recruitment was carried out by MF, LS, and KM. GD and DJ designed the overall study, while GD developed the methods and supervised the preparation and writing of the paper. Data collection was performed by EB, KP, and LN. Coding was conducted by SD, SS, and CEO. The preparation, analysis, and writing of the manuscript were carried out by SD, LV, and EB. KB, JH-G, GH, and MB supported the analysis. All authors carefully reviewed the results and approved the final version of the manuscript.

Conflicts of Interest

None declared.

References

XSL•FO
RenderX

1. Keyes KM, Gary D, O'Malley PM, Hamilton A, Schulenberg J. Recent increases in depressive symptoms among US adolescents: trends from 1991 to 2018. Soc Psychiatry Psychiatr Epidemiol 2019 Aug;54(8):987-996. [doi: 10.1007/s00127-019-01697-8] [Medline: 30929042]
2. Wiens K, Bhattarai A, Pedram P, et al. A growing need for youth mental health services in Canada: examining trends in youth mental health from 2011 to 2018. Epidemiol Psychiatr Sci 2020 Apr 17;29:e115. [doi: 10.1017/S2045796020000281] [Medline: 32299531]
3. Bonnie RJ, Stroud C, Breiner H, et al. Young adults in the 21st century. In: Investing in the Health and Well-Being of Young Adults: National Academies Press; 2015. [Medline: 25855847]
4. Solmi M, Radua J, Olivola M, et al. Age at onset of mental disorders worldwide: large-scale meta-analysis of 192 epidemiological studies. Mol Psychiatry 2022 Jan;27(1):281-295. [doi: 10.1038/s41380-021-01161-7] [Medline: 34079068]
5. Erskine HE, Moffitt TE, Copeland WE, et al. A heavy burden on young minds: the global burden of mental and substance use disorders in children and youth. Psychol Med 2015 May;45(7):1551-1563. [doi: 10.1017/S0033291714002888] [Medline: 25534496]
6. Kourgiantakis T, Markoulakis R, Lee E, et al. Access to mental health and addiction services for youth and their families in Ontario: perspectives of parents, youth, and service providers. Int J Ment Health Syst 2023 Mar 14;17(1):4. [doi: 10.1186/s13033-023-00572-z] [Medline: 36918893]
7. Radez J, Reardon T, Creswell C, Orchard F, Waite P. Adolescents' perceived barriers and facilitators to seeking and accessing professional help for anxiety and depressive disorders: a qualitative interview study. Eur Child Adolesc Psychiatry 2022 Jun;31(6):891-907. [doi: 10.1007/s00787-020-01707-0] [Medline: 33502596]
8. Kafka JX, Kothgassner OD, Felnhofer A. A matter of trust: confidentiality in therapeutic relationships during psychological and medical treatment in children and adolescents with mental disorders. J Clin Med 2024 Mar 18;13(6):1752. [doi: 10.3390/jcm13061752] [Medline: 38541975]
9. Martínez-Hernáez A, DiGiacomo SM, Carceller-Maicas N, Correa-Urquiza M, Martorell-Poveda MA. Non-professional-help-seeking among young people with depression: a qualitative study. BMC Psychiatry 2014 Apr 28;14(1):124. [doi: 10.1186/1471-244X-14-124] [Medline: 24774644]
10. Davenport TA, Cheng VWS, Iorfino F, et al. Flip the clinic: a digital health approach to youth mental health service delivery during the COVID-19 pandemic and beyond. JMIR Ment Health 2020 Dec 15;7(12):e24578. [doi: 10.2196/24578] [Medline: 33206051]
11. Kruzan KP, Fitzsimmons-Craft EE, Dobias M, Schleider JL, Pratap A. Developing, deploying, and evaluating digital mental health interventions in spaces of online help- and information-seeking. Procedia Comput Sci 2022;206:6-22. [doi: 10.1016/j.procs.2022.09.081] [Medline: 37063642]
12. Wies B, Landers C, Ienca M. Digital mental health for young people: a scoping review of ethical promises and challenges. Front Digit Health 2021;3:697072. [doi: 10.3389/fdgth.2021.697072] [Medline: 34713173]
13. Lehtimaki S, Martic J, Wahl B, Foster KT, Schwalbe N. Evidence on digital mental health interventions for adolescents and young people: systematic overview. JMIR Ment Health 2021 Apr 29;8(4):e25847. [doi: 10.2196/25847] [Medline: 33913817]
14. Burns JM, Davenport TA, Durkin LA, Luscombe GM, Hickie IB. The internet as a setting for mental health service utilisation by young people. Med J Aust 2010 Jun 7;192(S11):S22-S26. [doi: 10.5694/j.1326-5377.2010.tb03688.x] [Medline: 20528703]
15. Pretorius C, Chambers D, Cowan B, Coyle D. Young people seeking help online for mental health: cross-sectional survey study. JMIR Ment Health 2019 Aug 26;6(8):e13524. [doi: 10.2196/13524] [Medline: 31452519]
16. Sweeney GM, Donovan CL, March S, Forbes Y. Logging into therapy: adolescent perceptions of online therapies for mental health problems. Internet Interv 2019 Mar;15:93-99. [doi: 10.1016/j.invent.2016.12.001] [Medline: 30792959]
17. Rickwood DJ, Mazzer KR, Telford NR. Social influences on seeking help from mental health services, in-person and online, during adolescence and young adulthood. BMC Psychiatry 2015 Mar 7;15:40. [doi: 10.1186/s12888-015-0429-6] [Medline: 25886609]
18. Dallinger VC, Krishnamoorthy G, du Plessis C, et al. Utilisation of digital applications for personal recovery amongst youth with mental health concerns. Int J Environ Res Public Health 2022 Dec 14;19(24):16818. [doi: 10.3390/ijerph192416818] [Medline: 36554700]
19. Lattie EG, Stiles-Shields C, Graham AK. An overview of and recommendations for more accessible digital mental health services. Nat Rev Psychol 2022 Feb;1(2):87-100. [doi: 10.1038/s44159-021-00003-1] [Medline: 38515434]
20. Stiles-Shields C, Ramos G, Ortega A, Psihogios AM. Increasing digital mental health reach and uptake via youth partnerships. Npj Ment Health Res 2023;2(1):1-4. [doi: 10.1038/s44184-023-00030-1] [Medline: 37483392]
21. Bell IH, Thompson A, Valentine L, Adams S, Alvarez-Jimenez M, Nicholas J. Ownership, use of, and interest in digital mental health technologies among clinicians and young people across a spectrum of clinical care needs: cross-sectional survey. JMIR Ment Health 2022 May 11;9(5):e30716. [doi: 10.2196/30716] [Medline: 35544295]
22. McKenna S, Hutcheon A, Gorban C, Song Y, Scott E, Hickie I. Youth perspectives on "highly personalised and measurement-based care": qualitative co-design of education materials. Health Expect 2024 Aug;27(4):e14137. [doi: 10.1111/hex.14137] [Medline: 38976368]

XSL•FO
RenderX

23. Høgsdal H, Kyrrestad H, Rye M, Kaiser S. Exploring adolescents' attitudes toward mental health apps: concurrent mixed methods study. JMIR Form Res 2024 Jan 15;8(1):e50222. [doi: 10.2196/50222] [Medline: 38224474]
24. Clarke AM, Kuosmanen T, Barry MM. A systematic review of online youth mental health promotion and prevention interventions. J Youth Adolesc 2015 Jan;44(1):90-113. [doi: 10.1007/s10964-014-0165-0] [Medline: 25115460]
25. Borghouts J, Eikey E, Mark G, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. J Med Internet Res 2021;23(3):e24387. [doi: 10.2196/24387]
26. Ludlow K, Russell JK, Ryan B, et al. Co-designing a digital mental health platform, "Momentum", with young people aged 7-17: a qualitative study. Digit Health 2023;9:20552076231216410. [doi: 10.1177/20552076231216410] [Medline: 38033517]
27. Grist R, Croker A, Denne M, Stallard P. Technology delivered interventions for depression and anxiety in children and adolescents: a systematic review and Meta-analysis. Clin Child Fam Psychol Rev 2019 Jun;22(2):147-171. [doi: 10.1007/s10567-018-0271-8] [Medline: 30229343]
28. Bevan Jones R, Stallard P, Agha SS, et al. Practitioner review: Co-design of digital mental health technologies with children and young people. J Child Psychol Psychiatry 2020 Aug;61(8):928-940. [doi: 10.1111/jcpp.13258] [Medline: 32572961]
29. Agapie E, Chang K, Patrachari S, Neary M, Schueller SM. Understanding mental health apps for youth: focus group study with Latinx youth. JMIR Form Res 2022 Oct 18;6(10):e40726. [doi: 10.2196/40726] [Medline: 36256835]
30. Bassi EM, Bright KS, Norman LG, et al. Perceptions of mental health providers of the barriers and facilitators of using and engaging youth in digital mental-health-enabled measurement based care. Digit Health 2024;10:20552076241253093. [doi: 10.1177/20552076241253093] [Medline: 38726214]
31. Dimitropoulos G, Bassi EM, Bright KS, et al. Implementation of an electronic mental health platform for youth and young adults in a school context across Alberta, Canada: thematic analysis of the perspectives of stakeholders. JMIR Ment Health 2024 Jan 17;11:e49099. [doi: 10.2196/49099] [Medline: 38231558]
32. Iorfino F, Cross SP, Davenport T, et al. A digital platform designed for youth mental health services to deliver personalized and measurement-based care. Front Psychiatry 2019;10:595. [doi: 10.3389/fpsyt.2019.00595] [Medline: 31507465]
33. Bradshaw C, Atkinson S, Doody O. Employing a qualitative description approach in health care research. Glob Qual Nurs Res 2017;4:2333393617742282. [doi: 10.1177/2333393617742282] [Medline: 29204457]
34. Colorafi KJ, Evans B. Qualitative descriptive methods in health science research. HERD 2016 Jul;9(4):16-25. [doi: 10.1177/1937586715614171] [Medline: 26791375]
35. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qualitative Research in Sport, Exercise and Health 2019 Aug 8;11(4):589-597. [doi: 10.1080/2159676X.2019.1628806]
36. Malterud K, Siersma V, Guassora AD. Information power: sample content and size in qualitative studies. In: Qualitative Research in Psychology: Expanding Perspectives in Methodology and Design, Vol. 2021:67-81. [doi: 10.1037/0000252-004]
37. Iorfino F, Scott EM, Carpenter JS, et al. Clinical stage transitions in persons aged 12 to 25 years presenting to early intervention mental health services with anxiety, mood, and psychotic disorders. JAMA Psychiatry 2019 Nov 1;76(11):1167-1175. [doi: 10.1001/jamapsychiatry.2019.2360] [Medline: 31461129]
38. McKenna S, Piper S, Capon W, et al. The polarization of clinician and service staff perspectives after the use of health information technology in youth mental health services: implementation and evaluation study. JMIR Hum Factors 2023 Jul 25;10:e42993. [doi: 10.2196/42993] [Medline: 37490321]
39. Ospina-Pinillos L, Davenport T, Iorfino F, et al. Using new and innovative technologies to assess clinical stage in early intervention youth mental health services: evaluation study. J Med Internet Res 2018 Sep 10;20(9):e259. [doi: 10.2196/jmir.9966] [Medline: 30201602]
40. Cheng VWS, Piper SE, Ottavio A, Davenport TA, Hickie IB. Recommendations for designing health information technologies for mental health drawn from self-determination theory and co-design with culturally diverse populations: template analysis. J Med Internet Res 2021 Feb 10;23(2):e23502. [doi: 10.2196/23502] [Medline: 33565985]
41. Cavazos-Rehg P, Min C, Fitzsimmons-Craft EE, et al. Parental consent: a potential barrier for underage teens' participation in an mHealth mental health intervention. Internet Interv 2020 Sep;21:100328. [doi: 10.1016/j.invent.2020.100328] [Medline: 32528858]
42. Hoopes AJ, Brandzel SD, Luce C, et al. What do adolescents and their parents need from mental health integration in primary care? A qualitative exploration of design insights. J Pediatr Health Care 2022;36(6):570-581. [doi: 10.1016/j.pedhc.2022.06.006] [Medline: 35953380]
43. Lustgarten SD, Garrison YL, Sinnard MT, Flynn AW. Digital privacy in mental healthcare: current issues and recommendations for technology use. Curr Opin Psychol 2020 Dec;36:25-31. [doi: 10.1016/j.copsyc.2020.03.012] [Medline: 32361651]
44. Inkster B, Knibbs C, Bada M. Cybersecurity: a critical priority for digital mental health. Front Digit Health 2023;5:1242264. [doi: 10.3389/fdgth.2023.1242264] [Medline: 37781452]
45. Simões de Almeida R, Marques A. User engagement in mobile apps for people with schizophrenia: a scoping review. Front Digit Health 2022;4:1023592. [doi: 10.3389/fdgth.2022.1023592] [Medline: 36703941]
46. Peters D, Calvo RA, Ryan RM. Designing for motivation, engagement and wellbeing in digital experience. Front Psychol 2018;9:797. [doi: 10.3389/fpsyg.2018.00797] [Medline: 29892246]

XSL•FO
RenderX

47. Stawarz K, Preist C, Tallon D, Wiles N, Coyle D. User experience of cognitive behavioral therapy apps for depression: an analysis of app functionality and user reviews. J Med Internet Res 2018 Jun 6;20(6):e10120. [doi: 10.2196/10120] [Medline: 29875087]
48. Burn AM, Ford TJ, Stochl J, Jones PB, Perez J, Anderson JK. Developing a web-based app to assess mental health difficulties in secondary school pupils: qualitative user-centered design study. JMIR Form Res 2022 Jan 10;6(1):e30565. [doi: 10.2196/30565] [Medline: 35006079]
49. Borghouts J, Pretorius C, Ayobi A, Abdullah S, Eikey EV. Editorial: factors influencing user engagement with digital mental health interventions. Front Digit Health 2023;5:1197301. [doi: 10.3389/fdgth.2023.1197301] [Medline: 37143934]
50. Achilles MR, Anderson M, Li SH, Subotic-Kerry M, Parker B, O'Dea B. Adherence to e-mental health among youth: considerations for intervention development and research design. Digit Health 2020;6:2055207620926064. [doi: 10.1177/2055207620926064] [Medline: 32547775]
51. Lillevoll KR, Vangberg HCB, Griffiths KM, Waterloo K, Eisemann MR. Uptake and adherence of a self-directed internet-based mental health intervention with tailored e-mail reminders in senior high schools in Norway. BMC Psychiatry 2014 Jan 21;14:14. [doi: 10.1186/1471-244X-14-14] [Medline: 24443820]
52. Liverpool S, Mota CP, Sales CMD, et al. Engaging children and young people in digital mental health interventions: systematic review of modes of delivery, facilitators, and barriers. J Med Internet Res 2020 Jun 23;22(6):e16317. [doi: 10.2196/16317] [Medline: 32442160]
53. Opie JE, Vuong A, Welsh ET, Esler TB, Khan UR, Khalil H. Outcomes of best-practice guided digital mental health interventions for youth and young adults with emerging symptoms: part II. A systematic review of user experience outcomes. Clin Child Fam Psychol Rev 2024 Jun;27(2):476-508. [doi: 10.1007/s10567-024-00468-5] [Medline: 38634939]
54. Kircaburun K, Alhabash S, Tosuntaş Ş, Griffiths MD. Uses and gratifications of problematic social media use among university students: a simultaneous examination of the Big Five of personality traits, social media platforms, and social media use motives. Int J Ment Health Addiction 2020 Jun;18(3):525-547. [doi: 10.1007/s11469-018-9940-6]
55. Ravaccia GG, Johnson SL, Morgan N, Lereya ST, Edbrooke-Childs J. Experiences of using the digital support tool MeeToo: mixed methods study. JMIR Pediatr Parent 2022 Oct 20;5(4):e37424. [doi: 10.2196/37424] [Medline: 36264619]
56. Grant C, Widnall E, Cross L, Simonoff E, Downs J. Informing the development of an E-platform for monitoring wellbeing in schools: involving young people in a co-design process. Res Involv Engagem 2020;6(1):51. [doi: 10.1186/s40900-020-00219-0] [Medline: 32908677]
57. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen J. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152. [doi: 10.2196/jmir.2104] [Medline: 23151820]
58. Smith AC, Fowler LA, Graham AK, et al. Digital overload among college students: implications for mental health app use. Soc Sci (Basel) 2021;10(8):279. [doi: 10.3390/socsci10080279] [Medline: 34305199]
59. Boucher EM, Raiker JS. Engagement and retention in digital mental health interventions: a narrative review. BMC Digit Health 2024;2(1):52. [doi: 10.1186/s44247-024-00105-9]
60. Deci EL, Ryan RM. The "what" and "why" of goal pursuits: human needs and the self-determination of behavior. Psychol Inq 2000 Oct;11(4):227-268. [doi: 10.1207/S15327965PLI1104_01]

Abbreviations

dMH: digital mental health

Edited by A Kushniruk; submitted 22.12.24; peer-reviewed by M Hirschi, O Ogorry; revised version received 16.03.25; accepted 22.03.25; published 29.05.25.

Please cite as:
Daniel S, Volcko L, Bassi E, Hews-Girard J, Bright K, Barker M, Norman L, Pintson K, Henry G, Soufi S, Omorotionmwan CE, Fersovitch M, Stamp L, Moskovic K, Johnson DW, Dimitropoulos G
Exploring Youth Perspectives on Digital Mental Health Platforms: Qualitative Descriptive Study
JMIR Hum Factors 2025;12:e69907
URL: https://humanfactors.jmir.org/2025/1/e69907
doi:10.2196/69907

© Sarah Daniel, Lauren Volcko, Emilie Bassi, Julia Hews-Girard, Katherine Bright, Marianne Barker, Lia Norman, Karina Pintson, Geneca Henry, Sumaya Soufi, Chukwudumbiri Efrem Omorotionmwan, Melanie Fersovitch, Leanne Stamp, Karen Moskovic, David W Johnson, Gina Dimitropoulos. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 29.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License

XSL•FO
RenderX

(https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Preuse Acceptance of a Family-Centered, Need-Based, and Interprofessional Perinatal Care Mobile Health Intervention: Exploratory Study

Kristina Killinger[1], MSc, Dr med; Verena Seyfried[1], MSc; Katharina Brusniak[2], Dr med; Markus Wallwiener[3], Prof Dr med; Michael Abou-Dakn[4], Prof Dr med; Dorothea Scholle[4], Dr med; Stephanie Wallwiener[5], Prof Dr med

[1]Department of Gynecology and Obstetrics, University Hospital Heidelberg, Im Neuenheimer Feld 440, Heidelberg, Germany

[2]Department of Anesthesiology and Intensive Care Unit, Florence-Nightingale-Hospital, Duesseldorf, Germany

[3]Department of Gynecology, University Hospital Halle, Halle, Germany

[4]Department of Obstetrics and Gynecology, St. Joseph's Hospital Berlin, Berlin, Germany

[5]Department of Obstetrics and Perinatal Medicine, University Hospital Halle, Halle, Germany

Corresponding Author:
Kristina Killinger, MSc, Dr med
Department of Gynecology and Obstetrics, University Hospital Heidelberg, Im Neuenheimer Feld 440, Heidelberg, Germany

Abstract

Background: The perinatal period is one of the most vulnerable times a woman experiences. Multidimensional, interprofessional, and personalized support is needed to improve outcomes in women's and children's health while strengthening partner relationships at the same time. Although a vast amount of support services already exist in Germany for psychosocial counseling during the perinatal period, groups who are especially at risk do not take advantage of them.

Objective: Family eNav is an app-based intervention developed by experts in the field of medical and psychosocial support to help young parents navigate through primary and secondary care services in Germany according to their needs. It also empowers patient and parenting perspectives through self-education and symptom monitoring for different settings, for example, mental health and preterm birth. While the intervention will be evaluated in a multicenter, randomized, controlled trial, the focus here lies on the conception of the app, demand among patients, and preuse acceptance.

Methods: During the conception phase, we conducted an explorative study with prospective users and experts in the perinatal psychosocial field to understand the need and preuse acceptance of the intervention. We interviewed 20 participants with a semistructured guide, analyzing their responses using systematic text condensation. Additionally, we conducted a short survey on general questions concerning digitalization within the health care system among the participants.

Results: We established two main themes: (1) access and barriers to health care and psychosocial services and (2) high preuse acceptance of app-based intervention. Health care and psychosocial providers indicated that there is a high demand for their services, which cannot always be met immediately, and at the same time, they are doubtful of reaching those individuals most in need. Prospective users and health and social care providers alike showed great interest in the perinatal navigator and suggested a variety of needs and content requirements to be included. Regionality, availability, and individualized content were underlined as success factors for high user acceptance. Barriers consisted of data protection concerns, as well as denial of their own needs.

Conclusions: Our findings show great acceptance for an app-based intervention on the part of both prospective users and service providers. Feedback on requirements and content, as well as possible barriers, was taken into consideration while developing the app.

(JMIR Hum Factors 2025;12:e66658) doi:10.2196/66658

KEYWORDS

perinatal; mobile app; need-based care; patient empowerment; preuse acceptance; family; digital health; smartphone; eHealth; preterm birth; woman; Germany; psychosocial support; parenting; survey; telehealth; telemedicine

Introduction

Pregnancy and early parenthood represent a turning point in life. Sufficient multidimensional support during pregnancy may be beneficial for mother-child relationships [1,2], whereas a lack of support can lead to manifold psychosocial risk factors [3]. Exposure to psychosocial risk factors, such as maternal depression, stress, and low social support, negatively affects pregnancy and child development [4,5]. Maternal and children's health and social inequality are tightly linked [6-9]. Therefore,

XSL•FO
RenderX

improving maternal psychosocial risk factors during this phase represents an important public health aim.

In 2015, a cross-sectional survey in Germany of over 7500 families with small children showed that around 40% of families have three and more stressors, consisting of biographic, perinatal, and psychosocial measures, cumulatively, whereby at least 25% of participants reported that the parenting role was associated with stress [10]. A vast amount of support services already exist in Germany. The ministry for families, seniors, women, and children even offers web-based services through a family portal [11]. For families in troublesome situations, early preventative measures are outlined, from diminishing exposure to violence to promoting a positive environment for development, which are all compiled—and subsidized—within the initiative called "Frühe Hilfen" (translation: Early Help) [12]. Frühe Hilfen conducted several studies on how socioeconomic factors are related to obstacles to using early childhood prevention services. They found that families from lower socioeconomic backgrounds knew less about assistance programs in general and were less likely to actively make use of them. Programs such as home visits or family midwives, who provide support during the first year after childbirth, were more often used by families with lower socioeconomic backgrounds because they were offered to them on the basis of need, and parents did not have to actively look for them [13,14].

Women show a high tendency to search for pregnancy-related information on the web and via smartphone apps [15,16]. The number of health apps for pregnant women that are easy to access and can introduce new ways to perinatal care delivery, especially in low socioeconomic settings, is growing [15,17]. The quality and effectiveness of mobile apps vary tremendously, though, which can lead to mistrust and discontinuance of the app [18,19]. Nonetheless, multiple randomized controlled trials (RCTs) showed health-promoting effects of using mobile health (mHealth) apps, for example, in weight management, gestational diabetes management, and maternal mental health [20,21]. Introducing a digital tool that has been developed with a user- and provider-centered approach opens a realistic opportunity to find evidence-based information and also lowers the threshold for gaining access to this information by digitally interlinking families in need with their regional service providers.

Our research project, Family eNav, is a multicenter health service program funded through the innovation funds of the Joint National Committee (German: Gemeinsamer Bundesausschuss). It was our objective to develop an app-based perinatal guide to better coordinate need-based primary and secondary prevention care and empower patient and parenting perspectives through self-education and symptom monitoring. Intensifying the link between health and social system support through an app-based navigator may ameliorate care for pregnant women and young families, especially those with psychosocial risk factors, and hence may lead to higher quality of life, relieve stress, and strengthen partner and parent-child relationships.

In the conception phase, we tried to develop the content basis for the guide by evaluating the demand for and acceptance of such an app-based intervention. The aim of this study was therefore to gain a deeper understanding of young families' needs and reasons for using services to establish the demand for an app-based perinatal guide. At the same time, we tried to shed light on the provider's perspective on how to successfully integrate an app-based perinatal navigator to determine factors of success and possible barriers.

Methods

Study Design

The study was conducted at the University Hospital of Heidelberg with pregnant women and young mothers, as well as health care experts [22]. Study participants were recruited between November 2021 and March 2022. We used purposive sampling to include all perspectives necessary for the conception of an app-based perinatal guidance program. Overall, we recruited 9 prospective users, among whom 5 users were pregnant and 4 users had infants, and 13 experts from all areas of perinatal services (Table 1). Eligibility criteria for users were older than 18 years of age and fluency in German. To be eligible for participation, service providers had to be involved in the care of pregnant women or young families. We conducted interviews until information saturation was reached.

XSL•FO
RenderX

Table . Study participant information.

	Age (years)	Highest level of education	Duration of interview (min)
Patient			
Pregnant in week 32 (interview 1)	34	University degree	18
Pregnant in week 28 (interview 2)	32	University degree	33
Pregnant in week 20 (interview 3)	30	Secondary level	21
Mother of 6-week-old infant (interview 4)	29	University degree	19
Mother of 4-month-old infant (interview 5)	31	Secondary level	29
Mother of 8-week-old infant (interview 6)	29	Secondary level	63
Mother of 3-month-old twins (interview 7)	30	University degree	38
Pregnant in week 11 (interview 8)	33	Secondary level	29
Pregnant in week 12 (interview 9)	31	Vocational training	17
Expert			
Course teacher (interview 10)	34	University degree	21
Psychologist or psychotherapist (interview 11)	65	University degree	30
Physiotherapist (interview 12)	27	Vocational training	29
Midwife (interview 13)	61	Vocational training	28
Coordinator at Frühe Hilfen (interview 14)	48	University degree	38
Pediatrician (interview 15)	62	Medical examination	36
Professor of medical psychology (interview 16)	46	Habilitation	24
Preterm development aid worker (interview 17)	41	University degree	45
Frühe Hilfen (interview 18)	56	University degree	43
Frühe Hilfen (interview 19)	37	University degree	29
Head of Gynecology Department (interview 20)	64	Medical degree	59

Semistructured interviews with a mean length of 32.45 (SD 9.54) minutes were held by 3 trained investigators following a previously established interview guide with primarily open-ended questions. At the end of the semistructured interviews, we added a general survey of the participants' perspective on digital health with primarily closed-ended questions, using a scaling system (Multimedia Appendix 1). We performed systematic text condensation on the responses to the questions, which is a descriptive method concentrating on people's experiences rather than the underlying meaning [23]. Through meaning units, we established a coding scheme to form predominant themes in accordance with our research questions. Researcher triangulation was used to increase reliability. Specific quotes were translated into English to underline the main perspectives. The answers to the digital health survey were assessed through descriptive statistics.

The study aimed to explore the perceived status quo of perinatal services and investigate whether there was a demand for a digital navigator, and also determine preuse acceptance of the planned mHealth intervention study [24]. Second, we wanted to give our prospective users and collaborators a voice within the conception phase of the app. Data were analyzed descriptively.

Ethical Considerations

The study was approved by the ethical committee of the University of Heidelberg (S-344/2022). Informed consent was gathered prior to the interviews. Participants gave oral and written permission to record the interviews and perform data

XSL•FO
RenderX

analyzation. They were later transcribed verbatim. The data were anonymized. Participants did not receive any compensation.

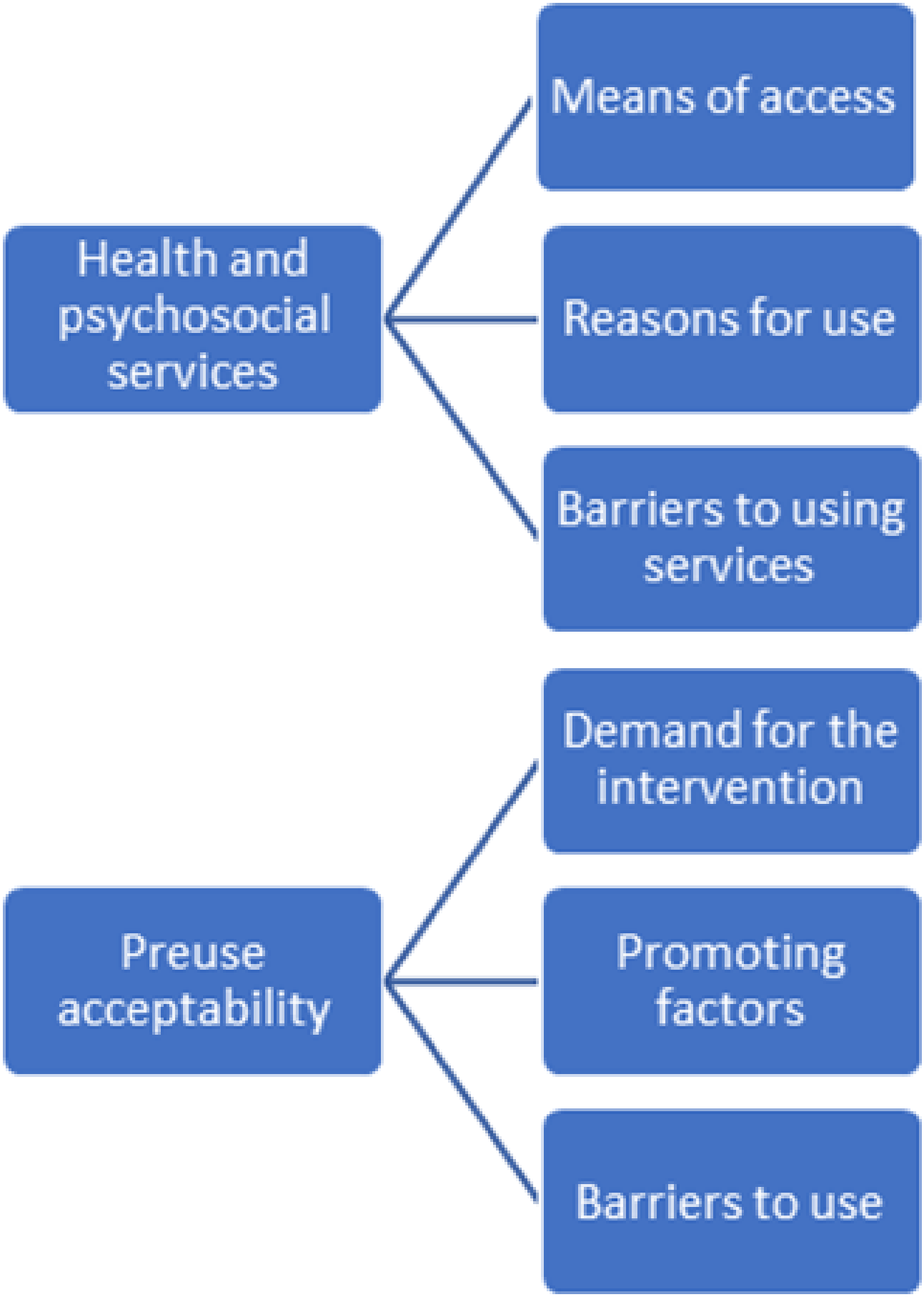
Results

Overview

Through our analysis, we established the following main themes: health care and psychosocial services, and preuse acceptance of an app-based intervention (Figure 1). We listed the most important meaning units for preuse acceptance by comparing the views of our prospective users and experts.

XSL•FO
RenderX

Figure 1. Main themes and categories retrieved from 20 interviews during the conception period of an app-based intervention study, called Family eNav.



XSL•FO
RenderX

Health Care and Psychosocial Services: Access and Barriers

Means of Access to Psychosocial and Health Care Services

Internet-based methods, such as search engines, social media, and local home pages to access perinatal services, made up at least 50% of the responses from users and experts. Most experts said they were contacted via email. Conventional methods, such as the telephone or the formal referral system, made up the rest. Numerous experts additionally stressed the importance of word-of-mouth recommendations.

> *I mainly google things or ask colleagues who already were pregnant for recommendations.* [Interview 2]

Reasons for Using and Not Using Services

The willingness to take advantage of health care services during and after the pregnancy was high in our study group. Almost all participants at least tried to take part in activities such as antenatal classes, yoga for pregnant women, baby swimming classes, or courses for family bonding. Reasons for not taking advantage of services were a lack of interest and time, or low perceived value of the service itself, or low perceived qualification of the personnel, lack of knowledge, and higher barriers due to COVID-19 restrictions on services. Some minimized the use of certain services because of disadvantages due to cost and organizational load. Sometimes, a certain level of distrust was shown when services did not show transparently what they offered.

> *Because I think this is what happens to all parents, they just don't have enough time and if they had more time, maybe they would care more about things like this, but especially if they have to search things for a long time, then this is why it fails.* [Interview 7]

Our experts described their situation as being in high demand and even having to work with waiting lists for their services. According to them, the reason is a shortage of personnel and not saturation of demand. They even raised concerns about their outreach to families in need. Most experts see ignorance of their services as the most important reason that they are not used. They stated that barriers to asking for help are naturally high and that the process of finding help is time-consuming; therefore, many families cannot overcome such a barrier. Additionally, they suggested that families in need are often in denial about their own condition or even consider social and health services as a threat that needs to be controlled. Psychological conditions such as lethargy due to depression constitute another barrier to looking for and taking advantage of services. All experts recognized that a digital tool to navigate families would be helpful in lowering access barriers.

> *In our experience, I would say, very paradoxically, the smaller the problem, the higher the parents' willingness to ask for help and vice versa.* [Interview 14]

> *Well, I think the process of looking for help is complicated for postpartum depression, for example, because until you find someone to take you in, there is just not enough capacity in small cities. I think it would help to have some kind of shortcut to a contact person.* [Interview 10]

High Preuse Acceptance of App-Based Intervention

Overview

> *You see, hardly anybody reaches for a book or magazine or something similar while waiting for something. Usually, you have your phone with you.* [Interview 6]

This quote shows the importance of mobile devices in our patients' everyday lives. During our interviews, we saw great acceptance for an app-based intervention on the part of both prospective users and the service providers. As suitable devices, smartphones and handhelds were most often named. Service providers found a computer format more convenient for their daily work.

> *As a provider, the computer would be the most appropriate, but I think for most of the users the mobile phone is most appropriate.* [Interview 16]

We created a mind-map showing the most relevant points concerning the acceptance of our navigator, both from a user and provider standpoint (Multimedia Appendix 2).

High Demand for App-Based Navigator

All our prospective users saw the benefit of an app-based navigator to lower the threshold. They especially mentioned how they missed a sense of direction when looking for specific services. Especially, primipara considered their lack of experience as a source of uncertainty throughout their pregnancy. When those people in their surroundings have had little or have had bad experiences, they see the guidance through an app as a source of trust.

> *Because especially when you are expecting your first child you know zero of what to expect. There is this big unknown territory.* [Interview 6]

Not only do prospective users see the wide range of offers and the inherent lack of transparency here as hindering, but experts, too, emphasized the importance of having clear structures and orientation within health care services to ameliorate pre- and postnatal care.

> *And then we decide within the situation which help systems are needed. And I think, this is like a jungle for women but also for us midwives.* [Interview 13]

Another main improvement this study's participants found in using app-based navigators is time-efficient self-education, which gives them a feeling of security and strengthens their partnerships when sharing new information within the app. Most of our prospective users emphasized that it is important to have one app that combines all aspects of pregnancy, the postpartum period, and early childhood, instead of having a multitude of books and apps for only certain subjects.

> *For example, we bought the book Oh I am growing where all development steps are explained because I did not want to have an app, but now the book is actually just sitting on the bookshelf.* [Interview 7]

XSL•FO
RenderX

Service providers also had a very positive perception of the improvements a future digital navigator could bring. They considered an app where all health-related information could be collected as an improvement in efficiency, making it easier to treat their patients or adequately help their clients. Some of them could envision the app as part of an early warning system, making interventions available sooner to those who need them. Concerning the target group, they thought an app might have a better reach due to the low threshold for gaining information on services and due to the privacy aspect of self-assessing one's need through psychological check-ups and symptom tracking tools. They stated that a digital navigator might be able to lift the burden from some overcrowded postnatal wards and go hand in hand with hospitals' psychosocial support concepts. Like our prospective users, they believe that the app can help give a sense of security and orientation throughout this challenging time.

> *And I would recommend it because it will convey knowledge to parents, mothers, so as to give a little bit of assurance and orientation in this new phase and an important way to network. In the broadest sense of I know where to turn to when I need it and this will indirectly positively influence the children because mothers will feel more confident in their role as mothers or fathers in their role as fathers.* [Interview 14]

Promoting Factors for an App-Based Navigator

Design-wise, our interviewees found a clear-cut structure with a user-friendly interface, as well as easy-to-understand instructions and messages, to be most agreeable. Both experts and prospective users emphasized that anything that would lower the threshold will enhance user engagement.

> *(…)the more I feel at ease, because through colors, font, structure or something else it might leave me with a certain feeling, the more I would be willing to spend time on the app.* [Interview 6]

Both experts and prospective users expressed the need to have a filter function on the displayed services. Filters should include distance and availability of services to promote an efficient search. Prospective users would also look for ratings or comparison tools for services. They would like to have a direct link to the services and information on whether certain offers fall within their insurance benefits. To contact the services, most users would like to have a direct link to the home page or their email address, or telephone number, and some would even prefer to have a direct chat option.

Prospective users listed the following as important content for the self-education part of the navigator: financial aspects such as how to apply for parental benefits; information on the development of the fetus with individual updates according to the gestational age; information on prenatal care, advice on symptoms during the pregnancy, symptom- and weight-tracking tools; calendar function for all doctors' and health care services' appointments and deadlines; checklists for daily questions; information on delivery options and health care providers; support for the postpartum period; and list of postnatal courses, childcare services, information on breastfeeding, activities for babies, and nutritional advice pre- and postpartum.

In addition to the aforementioned content, experts included: relief opportunities for mothers such as family midwives, motherhood nurses, or simple food delivery options; help on finding a midwife and information about their importance; emergency telephone number when dealing with negative feelings after a traumatizing birth or due to postnatal depression; lists of medical doctors and psychologists; lists and information on where to give birth; organizational check lists pre- and postpartum; nutritional and sport advice during pregnancy and nursing period; advice on sleep rhythm and signaling of babies, as well as development milestones; and networking possibilities such as self-help groups.

When asked whether the women wanted to have their partner included in the app, most found it to be helpful. Experts also distinguished the importance of including partners.

> *I think this might be even more important than for myself. As a mother, one is a lot closer to the baby, and therefore, it is easier to find out what is wrong with it. Especially when the father is home alone then he has something like a guidebook. It could be a good back-up to look up things.* [Interview 1]

The last promoting factor that both sides brought up in nearly every interview is the trustworthiness of the information and the contacts that are displayed, and how this might be the key factor for women to continue using and working with the app.

> *The maximum of neutrality and the maximum of transparency and the maximum of the maximum of credibility. This is the basic prerequisite, and this needs to be conveyed.* [Interview 20]

Barriers to App Use

Possible reasons for not using such an app were dominated by data protection concerns, and in some form, a fear of becoming dependent on the smartphone. This was connected to time-consuming installation processes or questionnaires, which could lead to user dissatisfaction or even to withdrawal from the intervention. It was also mentioned that if the information was not kept up-to-date, users would stop using the app soon after realizing this.

> *I think first and foremost, data protection reasons. I think there are many who might be a little timid in this case. As to maybe their data might be given to third parties or maybe they must disclose too much.* [Interview 8]

Some participants stated that people might not be well informed or generally did not acknowledge their own need for help. Here, experts found lower education status, language barriers, and psychological conditions to be the cause.

> *They think they already know everything and will manage solely on their gut instinct. Otherwise, I do not see a reason because applications are omnipresent and only rarely do people shy away from using them.* [Interview 10]

XSL•FO
RenderX

Others thought that users might miss the personal contact when dealing with their problems through an app. Experts had doubts in particular about video-based therapy options, which they had already had to incorporate in their daily work due to COVID-19 restrictions.

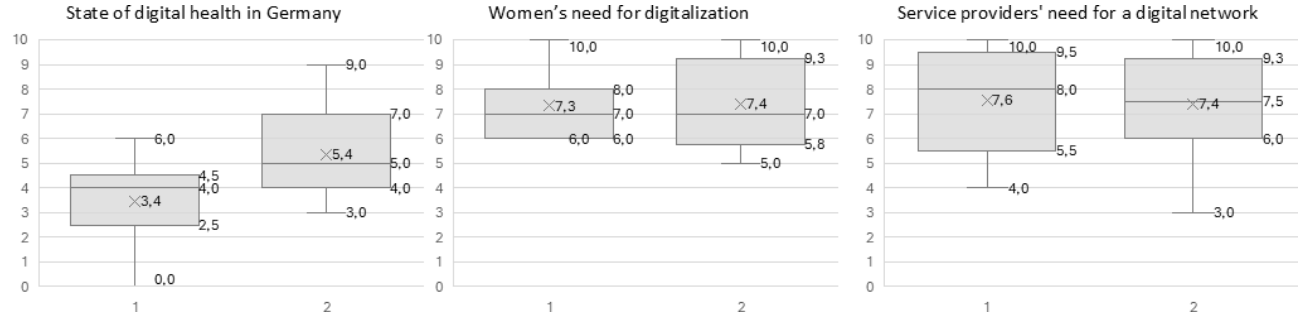
> *Maybe that is just because of anonymity, the nonverbal, the missing conversation.* [Interview 13]

In general, some assessed that the market for app-based interventions might already be saturated. Prospective users said they would not use an app if it were not free of charge. At the same time, none of the prospective users wanted to have advertisements within the app.

Digital Health Survey

At the end of the interview, we asked our participants some general questions about their opinions concerning digital health in Germany (Multimedia Appendix 1). They indicated that the state of digital health in Germany is at least mediocre. However, experts still had a higher opinion than our prospective users, where one even gave a zero, which was outside the given range of 1 - 10. When asked about the need for more digital interlinkage between women and services, both sides estimated that the need is fairly high. Here, opinions did not differ between experts and users (Figure 2). We asked whether not only interlinking providers and users but also providers with providers might be beneficial. All the respondents thought that creating a digital network for service providers would improve access for women. A total of 13 out of 19 respondents thought it would also ameliorate the quality of services, while 6 respondents thought it would not have any effect on the quality. When we asked our service providers specifically, 82% indicated it would enhance their efficiency (Multimedia Appendix 3).

Figure 2. Boxplots of results from the digital health survey; possible answers ranged from 1 to 10 (1 meaning the least and 10 meaning the highest); 1=perspective users, 2=experts; marked within the plot are mean (X), median, and IQRs.



Discussion

Improving Access to Psychosocial Services in Germany

This study showed that most women already access information about perinatal services online and have a high willingness to take advantage of general perinatal services, in particular. Barriers consist mainly of a lack of knowledge, time and affordability constraints, and a certain degree of distrust in the qualification and effectiveness of the services. On the other hand, most experts stated that the services they offered are in high demand, and both pregnant women and young mothers reflected this in their interviews. This outcome can be argued in light of the sociodemographic background of our study participants and the region of Heidelberg, where the annual household average income is approximately €24,400 (approximately US $27,469), which is among the highest in Germany [25]. Our multicenter approach to the RCT, involving centers from low-income areas such as Berlin Tempelhof and more rural areas such as Jena, will provide a more representative study population. The next question we asked is whether the demand is high relative to what is being offered and the capacity of services, or whether services are saturated for other reasons, such as underfunding and understaffing. This fact was explicitly mentioned by some of our experts. While stating that their resources are limited, they also saw an urgent need to reach more families in need, suggesting that their services needed improved access points. This finding compares well with similar studies in Germany showing that there is, in fact, a need to improve access to perinatal services for hard-to-reach subgroups [14,26].

There is a significant amount of evidence pointing in the direction of mHealth for improving access to psychosocial treatments [27]. Furthermore, multiple studies showed that mHealth interventions had positive effects on a large range of psychosocial measurements, including self-management of health, acceptance of pregnancy, anxiety, depressive symptoms, perceived stress, mental well-being, coping, and self-efficiency. It also showed that mHealth interventions positively influenced social support from partners or health care providers. The greatest effect was observed in highly vulnerable pregnant women, leading to the conclusion that a perinatal navigator can effectively help vulnerable groups that we are missing right now [28-31]. In prior studies, interlinking women with service providers or enhancing self-help through mHealth interventions seemed to be the key concept. We took this a step further by asking whether it might be beneficial not only to interlink providers and users but also to connect providers with providers. All the respondents thought that access could be facilitated for women by creating a digital network for service providers. Most of the respondents thought it would also improve the quality of services. When we asked our service providers specifically, 82% indicated it would enhance their efficacy. Creating interdisciplinary networks within perinatal care is an ongoing aim for health care providers and has been shown to be beneficial, for example, in mental health settings [32,33]. So far, interprofessional networks have been based on personal meetings and consultations, and to our knowledge, a telemedical

XSL•FO
RenderX

network has not yet been implemented within perinatal care. We submit that regional interdisciplinary networks can be created and implemented via a mobile navigator and will improve access to and uptake of perinatal services.

High Preuse Acceptance Promising for RCT Outcomes

All of our study participants gave positive feedback and showed high preuse acceptance. Most pregnant and postpartum women have already used apps during their pregnancy. Promoting factors for app use from the user's perspective were time-efficient self-education concerning evidence-based information, positive feedback mechanisms through symptom tracking, direct and easy contact opportunities to regional perinatal and psychosocial services, and networking possibilities. Providers estimated a destigmatization of risk factors, the low threshold for information, and for access to specific services, as well as the chance to better monitor high-risk pregnancy, as success factors for a perinatal app-based navigator. On the subject of perinatal and postpartum depression, a study from Varma et al [34] showed comparable results on acceptance and utility. Especially during the COVID-19 pandemic, we saw an increased interest in the subject of telemedical monitoring. Several studies showed that virtual visits and communication through telemedicine are highly accepted on both the patient and provider sides of health care [35-37]. Barriers to app use expressed by future users were mainly unawareness, data protection concerns, complexity of app use, hidden costs, or in-app advertisements, as well as the already high number of apps on the market. Providers added language barriers, low education profile of users, and mismatch of app design, denial of necessity, and psychological lethargy, as well as a general digital skepticism, to this list. Similar barriers to telehealth uptake, including level of education, telehealth literacy, and unawareness, were found in a multinational review by Kruse et al [38]. What differs from other studies is our family-centered approach in a multidisciplinary context, allowing multiple entry points, as well as a personalized solution for women and young families navigating through this phase.

Limitations

There are several limitations to this study. The first and perhaps greatest limitation concerns our study participants' socioeconomic status and level of education, which is only partly representative of the foreseen cohort of women with psychosocial risk factors from hard-to-reach socioeconomic niches. Furthermore, all participants were from the same geographic area, which might lead to results not being generalizable nationally. Nonetheless, tailoring the app to the needs of prospective users and health care providers should help the intervention to be more successful. Testing the mHealth intervention in a multicenter setting will help with representative data. Another limitation is the small sample size, which is characteristic of qualitative research methodology. To validate the data, we used researcher triangulation. Interviews were conducted by two investigators, and data interpretation was done in a team of three. By introducing a small survey with closed-ended questions, we were able to triangulate common notions on digital health.

Conclusions

This study demonstrated high acceptance of an app-based perinatal navigator both from the users' perspective and the psychosocial and health care providers' points of view. It also showed that creating better networks between users and providers, but also between providers and providers, will help improve access and quality of care. Feedback on requirements and content, as well as possible barriers, informed our concept and helped in creating all parts of the self-education guide. All information was curated by a team of health care professionals from different specialties, including midwives, gynecologists, psychologists, and pediatricians. The results from this study underscore the necessity for a multicenter RCT on the perinatal navigator.

Acknowledgments

This study is funded as part of a multicenter randomized controlled trial, called Family eNav (01VSF20002), through the innovation funds of the Joint National Committee (German: Gemeinsamer Bundesausschuss). For the publication fee, we acknowledge financial support by the University of Heidelberg.

Authors' Contributions

KK transcribed, curated, analyzed, and interpreted data for the work. KK is responsible for drafting the paper and agrees to be accountable for all aspects of the work. VS and KB drafted the interview guide and are responsible for recruitment and interviewing the study participants. VS and KB were part of the data curation and interpretation of the work. MW, MA-D, and DS helped conceptualize the study design, as well as reviewed the process and final draft, and gave important contextual information. SW, as principal investigator of Family eNav, gave the idea to the project, as well as conceptual assistance for this exploratory study. SW reviewed the data and gave important input for the interpretation of the data. All coauthors examined the manuscript, made corrections, and approved it for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.

XSL•FO
RenderX

[DOCX File, 17 KB - humanfactors_v12i1e66658_app1.docx]

Multimedia Appendix 2
Mind map of expert and user perspectives.
[PNG File, 259 KB - humanfactors_v12i1e66658_app2.png]

Multimedia Appendix 3
Digital health survey results.
[DOCX File, 17 KB - humanfactors_v12i1e66658_app3.docx]

References

1. Kinsey CB, Baptiste-Roberts K, Zhu J, Kjerulff KH. Birth-related, psychosocial, and emotional correlates of positive maternal-infant bonding in a cohort of first-time mothers. Midwifery 2014 May;30(5):e188-e194. [doi: 10.1016/j.midw.2014.02.006] [Medline: 24650812]
2. Yarcheski A, Mahon NE, Yarcheski TJ, Hanks MM, Cannella BL. A meta-analytic study of predictors of maternal-fetal attachment. Int J Nurs Stud 2009 May;46(5):708-715. [doi: 10.1016/j.ijnurstu.2008.10.013] [Medline: 19081091]
3. Lancaster CA, Gold KJ, Flynn HA, Yoo H, Marcus SM, Davis MM. Risk factors for depressive symptoms during pregnancy: a systematic review. Am J Obstet Gynecol 2010 Jan;202(1):5-14. [doi: 10.1016/j.ajog.2009.09.007] [Medline: 20096252]
4. Walker SP, Wachs TD, Grantham-McGregor S, et al. Inequality in early childhood: risk and protective factors for early child development. Lancet 2011 Oct 8;378(9799):1325-1338. [doi: 10.1016/S0140-6736(11)60555-2] [Medline: 21944375]
5. Sabates R. The impact of multiple risk factors on young children's cognitive and behavioural development. Children Soc 2015;29:95-108. [doi: 10.1111/chso.12024]
6. Lampert T, Kuntz B. Effects of poverty for health and health behavior of children and adolescents : results from KiGGS Wave 2. Bundesgesundheitsbl Gesundheitsforsch Gesundheitssch 2019 Oct;62(10):1263-1274. [doi: 10.1007/s00103-019-03009-6] [Medline: 31529186]
7. Pinker V, Brandstetter S, Tischer C, et al. Determinants of maternal health four weeks after delivery: cross-sectional findings from the KUNO-kids health study. BMC Public Health 2021 Sep 15;21(1):1676. [doi: 10.1186/s12889-021-11667-y] [Medline: 34525999]
8. Dennis EF, Webb DA, Lorch SA, Mathew L, Bloch JR, Culhane JF. Subjective social status and maternal health in a low income urban population. Matern Child Health J 2012 May;16(4):834-843. [doi: 10.1007/s10995-011-0791-z]
9. CSDH. Closing the gap in a generation: health equity through action on the social determinants of health: final report of the commission on social determinants of health. : World Health Organization; 2008 URL: https://iris.who.int/handle/10665/43943 [accessed 2025-05-13]
10. Lorenz S, Ulrich SM, Sann A, Liel C. Self-reported psychosocial stress in parents with small children. Dtsch Arztebl Int 2020 Oct 16;117(42):709-716. [doi: 10.3238/arztebl.2020.0709] [Medline: 33559586]
11. Familie F, Senioren F, Jugend B. [Online calculators and interactive applications]. BMFSFJ. URL: https://www.bmfsfj.de/bmfsfj/service/online-rechner [accessed 2024-08-08]
12. Online course “new to early intervention” [Web page in German]. Nationales Zentrum Frühe Hilfen. URL: https://www.fruehehilfen.de/ [accessed 2024-08-08]
13. Neumann A, Renner I. Barriers to accessing early intervention services [Article in German]. Bundesgesundheitsbl 2016 Oct;59(10):1281-1291. [doi: 10.1007/s00103-016-2424-6]
14. Eickhorst A, Schreier A, Brand C, et al. Knowledge and use of different support programs in the context of early prevention in relation to family-related psychosocial burden [Article in German]. Bundesgesundheitsbl 2016 Oct;59(10):1271-1280. [doi: 10.1007/s00103-016-2422-8]
15. Wallwiener S, Müller M, Doster A, et al. Pregnancy eHealth and mHealth: user proportions and characteristics of pregnant women using web-based information sources—a cross-sectional study. Arch Gynecol Obstet 2016 Nov;294(5):937-944. [doi: 10.1007/s00404-016-4093-y]
16. Bert F, Gualano MR, Brusaferro S, et al. Pregnancy e-health: a multicenter Italian cross-sectional study on internet use and decision-making among pregnant women. J Epidemiol Community Health 2013 Dec 1;67(12):1013-1018. [doi: 10.1136/jech-2013-202584] [Medline: 24072743]
17. Guerra-Reyes L, Christie VM, Prabhakar A, Harris AL, Siek KA. Postpartum health information seeking using mobile phones: experiences of low-income mothers. Matern Child Health J 2016 Nov;20(Suppl 1):13-21. [doi: 10.1007/s10995-016-2185-8] [Medline: 27639571]
18. Scott KM, Gome GA, Richards D, Caldwell PHY. How trustworthy are apps for maternal and child health? Health Technol 2015 Apr;4(4):329-336. [doi: 10.1007/s12553-015-0099-x]
19. Krebs P, Duncan DT. Health app use among US mobile phone owners: a national survey. JMIR Mhealth Uhealth 2015 Nov 4;3(4):e101. [doi: 10.2196/mhealth.4924] [Medline: 26537656]

XSL•FO
RenderX

20. Chan KL, Chen M. Effects of social media and mobile health apps on pregnancy care: meta-analysis. JMIR Mhealth Uhealth 2019 Jan 30;7(1):e11836. [doi: 10.2196/11836] [Medline: 30698533]
21. Overdijkink SB, Velu AV, Rosman AN, van Beukering MD, Kok M, Steegers-Theunissen RP. The usability and effectiveness of mobile health technology-based lifestyle and medical intervention apps supporting health care during pregnancy: systematic review. JMIR Mhealth Uhealth 2018 Apr 24;6(4):e109. [doi: 10.2196/mhealth.8834] [Medline: 29691216]
22. Bogner A, Littig B, Menz W. Interviewing Experts: Springer; 2009.
23. Malterud K. Systematic text condensation: a strategy for qualitative analysis. Scand J Public Health 2012 Dec;40(8):795-805. [doi: 10.1177/1403494812465030] [Medline: 23221918]
24. Nadal C, Sas C, Doherty G. Technology acceptance in mobile health: scoping review of definitions, models, and measurement. J Med Internet Res 2020 Jul 6;22(7):e17256. [doi: 10.2196/17256] [Medline: 32628122]
25. Disposable income of private households in the metropolitan regions in Germany in 2019 [Web page in German]. Statista. 2022. URL: https://de.statista.com/statistik/daten/studie/789220/umfrage/verfuegbares-einkommen-privater-haushalte-in-den-metropolregionen-in-deutschland/ [accessed 2024-09-12]
26. Walz H, Bohn B, Sander J, et al. Access to difficult-to-reach population subgroups: a family midwife based home visiting service for implementing nutrition-related preventive activities—a mixed methods explorative study. AIMS Public Health 2015;2(3):516-536. [doi: 10.3934/publichealth.2015.3.516] [Medline: 29546123]
27. Parikh SV. Improving access to psychosocial treatments—integrating patient, provider, and systems approaches. Can J Psychiatry 2015 Jun;60(6):242-244. [doi: 10.1177/070674371506000602] [Medline: 26175321]
28. Sakamoto JL, Carandang RR, Kharel M, et al. Effects of mHealth on the psychosocial health of pregnant women and mothers: a systematic review. BMJ Open 2022 Feb 15;12(2):e056807. [doi: 10.1136/bmjopen-2021-056807] [Medline: 35168981]
29. Bayrampour H, Trieu J, Tharmaratnam T. Effectiveness of eHealth interventions to reduce perinatal anxiety: a systematic review and meta-analysis. J Clin Psychiatry 2019 Jan 22;80(1):18r12386. [doi: 10.4088/JCP.18r12386] [Medline: 30688418]
30. Khademioore S, Ebrahimi E, Khosravi A, Movahedi S. The effect of an mHealth application based on continuous support and education on fear of childbirth, self-efficacy, and birth mode in primiparous women: a randomized controlled trial. PLoS ONE 2023;18(11):e0293815. [doi: 10.1371/journal.pone.0293815] [Medline: 37910495]
31. Hassdenteufel K, Müller M, Abele H, et al. Using an electronic mindfulness-based intervention (eMBI) to improve maternal mental health during pregnancy: results from a randomized controlled trial. Psychiatry Res 2023 Dec;330:115599. [doi: 10.1016/j.psychres.2023.115599] [Medline: 37988816]
32. Asarnow JR, Rozenman M, Wiblin J, Zeltzer L. Integrated medical-behavioral care compared with usual primary care for child and adolescent behavioral health. JAMA Pediatr 2015 Oct 1;169(10):929. [doi: 10.1001/jamapediatrics.2015.1141]
33. Hense H, Mathiebe J, Helfer S, et al. Evaluation of the telemedical health care network "SAFE BIRTH" for pregnant women at risk, premature and sick newborns and their families: study protocol of a cluster-randomized controlled stepped-wedge trial. BMC Health Serv Res 2024 Feb 14;24(1):200. [doi: 10.1186/s12913-024-10667-z] [Medline: 38355579]
34. Varma DS, Mualem M, Goodin A, et al. Acceptability of an mHealth app for monitoring perinatal and postpartum mental health: qualitative study with women and providers. JMIR Form Res 2023 Jun 7;7:e44500. [doi: 10.2196/44500] [Medline: 37285185]
35. Harrison TN, Sacks DA, Parry C, Macias M, Ling Grant DS, Lawrence JM. Acceptability of virtual prenatal visits for women with gestational diabetes. Womens Health Issues 2017 May;27(3):351-355. [doi: 10.1016/j.whi.2016.12.009]
36. Foster KE, Casola AR, Uzumcu Z, Wodoslawsky S, Kelly C. Outpatient maternity care and telemedicine use perceptions in the COVID-19 pandemic: a 2020 CERA survey. Women Health 2022;62(5):402-411. [doi: 10.1080/03630242.2022.2072051] [Medline: 35531759]
37. Hofmann G, Hampanda K, Harrison MS, Fasano M, Nacht A, Yeoman M. Virtual prenatal and postpartum care acceptability among maternity care providers. Matern Child Health J 2022 Jul;26(7):1401-1408. [doi: 10.1007/s10995-022-03412-7] [Medline: 35292887]
38. Kruse CS, Karem P, Shifflett K, Vegi L, Ravi K, Brooks M. Evaluating barriers to adopting telemedicine worldwide: a systematic review. J Telemed Telecare 2018 Jan;24(1):4-12. [doi: 10.1177/1357633X16674087] [Medline: 29320966]

Abbreviations

mHealth: mobile health
RCT: randomized controlled trial

XSL•FO
RenderX

Edited by A Kushniruk; submitted 22.09.24; peer-reviewed by A Herzeg, M Gonzalez; revised version received 18.03.25; accepted 23.03.25; published 12.06.25.

Please cite as:
Killinger K, Seyfried V, Brusniak K, Wallwiener M, Abou-Dakn M, Scholle D, Wallwiener S
Preuse Acceptance of a Family-Centered, Need-Based, and Interprofessional Perinatal Care Mobile Health Intervention: Exploratory Study
JMIR Hum Factors 2025;12:e66658
URL: https://humanfactors.jmir.org/2025/1/e66658
doi:10.2196/66658

© Kristina Killinger, Verena Seyfried, Katharina Brusniak, Markus Wallwiener, Michael Abou-Dakn, Dorothea Scholle, Stephanie Wallwiener. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 12.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring Clinicians' and Patients' Acceptance and Utilization of a Digital Solution to Support Individualized Care in Diabetes Specialist Outpatient Care (DigiDiaS): Qualitative Study

Maria Aadland Mollestad[1], MSc; Annesofie Lunde Jensen[2,3*], PhD; Heidi Holmen[1,4*], PhD; Tone Singstad[5], MSc; Eirik Årsand[6], PhD; Jacob Andreas Winther[5], PhD; Astrid Torbjørnsen[1], PhD

[1]Department of Nursing and Health Promotion, Faculty of Health Sciences, OsloMet—Oslo Metropolitan University, Postboks 4, St Olavs plass, Oslo, Norway

[2]SDCA—Steno Diabetes Center Aarhus, Aarhus University Hospital, Aarhus, Denmark

[3]Department of Clinical Medicine, Aarhus University, Aarhus, Denmark

[4]The Intervention Centre, Oslo University Hospital, Oslo, Norway

[5]Division of Medicine, Akershus University Hospital, Lillestrøm, Norway

[6]Department of Computer Science, Faculty of Science and Technology, UiT The Arctic University of Norway, Tromsø, Norway

[*]these authors contributed equally

Corresponding Author:
Maria Aadland Mollestad, MSc
Department of Nursing and Health Promotion, Faculty of Health Sciences, OsloMet—Oslo Metropolitan University, Postboks 4, St Olavs plass, Oslo, Norway

Abstract

Background: With the increasing prevalence of type 1 diabetes alongside limited health care resources, the need for more sustainable health care services is apparent. Central to ensuring the standard of diabetes care while simultaneously optimizing resource utilization is improved patient-clinician communication and the provision of individualized care. Digital outpatient solutions incorporating patient-reported outcome measures (PROMs) have been introduced in diabetes outpatient care over recent years; however, features and delivery methods differ, and existing studies on their use and perceived clinical value are limited. Furthermore, clinicians' acceptance has been suggested as a key factor in the sustainability of digital solutions. Thus, to support the implementation of digital outpatient solutions perceived as valuable by clinicians and patients, we need more knowledge about how they are accepted and utilized in clinical practice.

Objective: This study investigates how clinicians and patients with type 1 diabetes accept and utilize a digital outpatient solution to support individualized care in the context of full-scale implementation at a diabetes specialist outpatient clinic. Furthermore, we aim to explore the synchronous interaction between patients and clinicians when they are allowed to prepare through the filling and reviewing of asynchronous PROMs before consultations.

Methods: This qualitative study uses interpretive description as a methodological approach. The digital outpatient solution features various components, including PROM questionnaires, asynchronous chat, remote consultations, e-learning, and information distribution. Data were collected through semistructured interviews with 10 clinicians and 20 patients with type 1 diabetes and observations of consultations. The data from the patient and clinician interviews (267 A4 pages) were analyzed separately before being jointly analyzed in the context of the findings from the observations (40 A4 pages).

Results: Our analysis resulted in the following three main themes that describe the interplay between clinicians' and patients' acceptance, utilization, and perceived clinical value of a digital outpatient solution: (1) clinicians' acceptance of the digital outpatient solution influences patients' acceptance, (2) variations in the use of different features influence the extent of individualized care, and (3) clinicians' and patients' utilization influences perceived care efficiency and quality. Those who demonstrated higher acceptance and more extensive utilization reported that the solution was more valuable in enhancing individualized care efficiency and quality.

Conclusions: This study highlights the interplay between clinicians' and patients' acceptance, utilization, and perceived clinical value of a digital outpatient solution in diabetes specialist outpatient care. Our findings suggest that when clinicians and patients understand why and how digital solutions are used, such solutions can enhance care efficiency and quality, contributing to sustainable health care. Future research should aim to gain an in-depth understanding of clinicians' and patients' acceptance, as well as the effectiveness of change management strategies when implementing digital outpatient solutions in diabetes specialist outpatient care.

XSL•FO
RenderX

(JMIR Hum Factors 2025;12:e70301) doi:10.2196/70301

KEYWORDS

telemedicine; eHealth; mHealth; patient-reported outcome measures; outpatient care; patient-centered care; diabetes mellitus; type 1 diabetes; interpretive description

Introduction

Background

As of 2022, about 8.75 million individuals worldwide were living with type 1 diabetes [1], a number predicted to increase to between 13.5 and 17.4 million by 2040 [2]. The rising global prevalence of type 1 diabetes, combined with limited health care resources, poses significant challenges to health care systems worldwide in their effort to provide high-quality care [3]. This escalating pressure calls for more sustainable health care services to ensure high diabetes care standards while simultaneously optimizing resource utilization. Thus, facilitating efficient communication between patients and their health care providers, emphasizing patients' individual needs, and providing timely support will be increasingly important in the immediate future. Emerging digital solutions designed to promote patient communication and evaluate health status to individualize care offer opportunities and potential strategies to navigate this successfully [3,4].

An important aspect of diabetes care involves educating and supporting patients in developing sufficient self-management skills [5]. Diabetes self-management, which includes maintaining blood glucose levels within the target range, adopting a healthy lifestyle, and managing potential psychological issues, can help minimize the risk of complications and improve quality of life [6,7]. However, patients' required self-management support level varies significantly, underlining the need for person-centered and individualized care [5]. Patient involvement becomes increasingly important, and digital solutions can be tailored to enable patients to participate actively in their treatment and self-management. One way of involving patients is through digital patient-reported outcome measures (PROMs), which entail collecting health information directly from patients without interpretation by a health care professional [8]. Relevant PROMs in type 1 diabetes may encompass patients' perceived glycemic control, self-management activities, and diabetes distress. PROMs can provide a comprehensive understanding of patients' needs and provide valuable information for clinicians when delivering person-centered and individualized care [9,10]. Simultaneously, PROMs can foster a self-reflection process within patients, leading to increased engagement in their treatment and self-management [9].

In recent years, digital outpatient solutions with PROMs have been developed and examined in diabetes specialist outpatient care [11-13]. However, according to a scoping review, these solutions' features and delivery methods vary [4]. Some digital PROM-based outpatient solutions that aim to individualize care seem to focus on the psychological aspects of living with type 1 diabetes [14,15], while others also include other factors, such as glucose monitoring, self-management activities, and lifestyle [13,16]. The delivery methods range from patients completing the PROMs at a touch screen at the clinic to utilizing web-based solutions [4]. Furthermore, the literature on how the solutions are used is limited, mainly comprising small pilot and feasibility studies [4], underscoring the need for further research.

DigiDiaS—Individualizing Diabetes Care Through a Digital Solution

In November 2021, a multicomponent digital outpatient solution was implemented at the diabetes specialist outpatient clinic at Akershus University Hospital in eastern Norway through the DigiDiaS project [17]. The solution mainly relies on PROMs entailing glucose monitoring, diabetes distress, and self-management activities. However, it also incorporates additional features aimed at individualizing care and supporting communication between clinicians and patients with type 1 diabetes, such as asynchronous chat and video consultations. For patients, use of the solution is voluntary and requires their consent upon app download. As of April 2024, 1354 out of approximately 1800 patients with type 1 diabetes had been included in the digital outpatient solution. All clinicians at the outpatient clinic utilize the digital outpatient solution, comprising diabetes specialist nurses and physicians working with type 1 diabetes patients.

For digital solutions aiming to individualize care to be implemented successfully, they must be accepted and utilized by both clinicians and patients, and clinicians' acceptance has been suggested as a key factor in new digital solutions' sustainability [18,19]. Acceptability, a core concept in digital health [20], can be defined as a multifaceted construct reflecting the extent to which people delivering or receiving a health care intervention consider it appropriate [19]. This acceptability can, in turn, influence utilization in the sense of user engagement and intervention effectiveness [20]. While previous research states that patients in diabetes outpatient care generally seem to welcome the use of digital PROMs [14,16,21,22] and the opportunity to self-report concerns [23,24], clinicians' acceptance and perceptions of such measures' usefulness vary [11,12,16]. Thus, we need more knowledge about how digital solutions are being accepted and utilized by clinicians and patients in diabetes specialist outpatient care [4]. Furthermore, no prior research has explored the synchronous interaction between clinicians and patients after asynchronous preparation using PROMs in a digital solution [4]. Consequently, the clinical value of using digital PROM-based solutions in diabetes care has yet to be established in terms of enhancing care quality and sustainable utilization of health care resources.

Objective

This study investigates how clinicians and patients with type 1 diabetes accept and utilize a digital solution to support individualized care in the context of full-scale implementation at a diabetes specialist outpatient clinic. Furthermore, we aim to explore the synchronous interaction between patients and

XSL•FO
RenderX

clinicians when they are allowed to prepare through the filling and reviewing of asynchronous PROMs before consultations.

Methods

Design

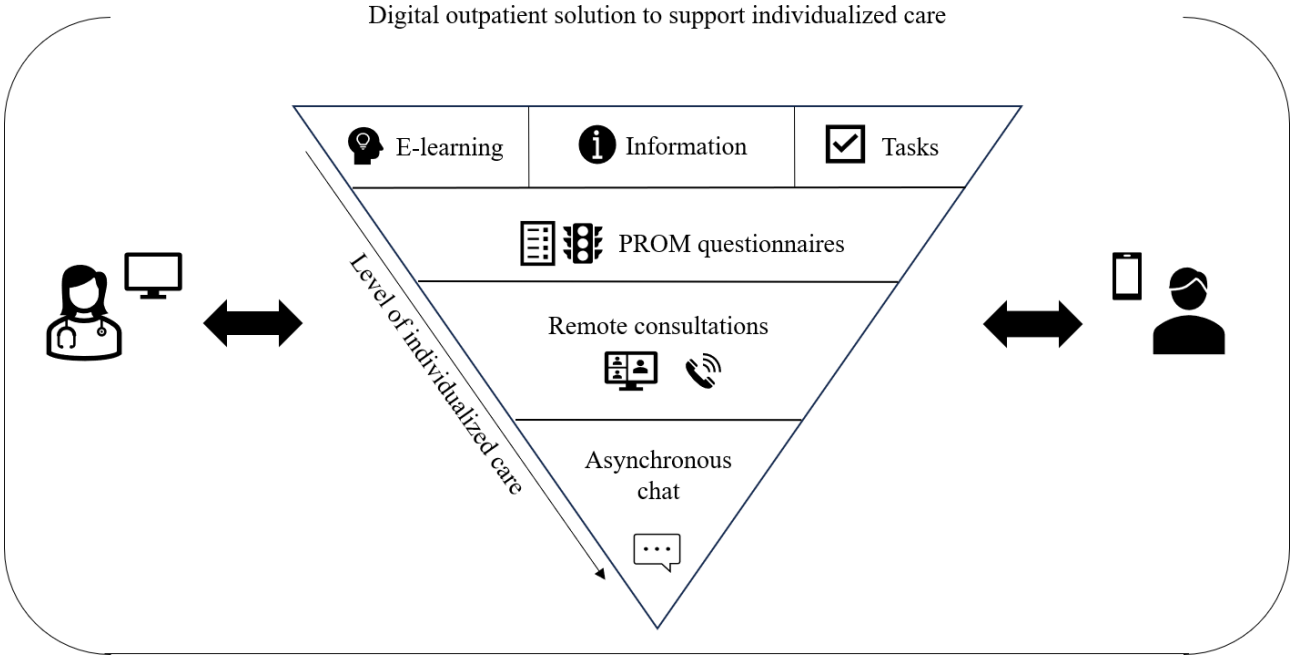
This qualitative study uses interpretive description as a methodological approach, comprising a framework developed to conduct research arising from clinical practice and aiming to improve practice [25]. The interpretive description process is inductive and pragmatic, providing flexibility to adapt the research process to the specific research questions [25]. Our study is embedded within a larger multimethod prospective observational study investigating digital PROMs in diabetes specialist outpatient care [17].

The Multicomponent Digital Outpatient Solution

The previously detailed solution [17] was facilitated via Dignio Connected Care, offering web-based software used by clinicians and an app for patients [26], along with JOIN Norsk Helsenett, which provides a platform for video consultations [27]. Clinicians received training on the use of the digital outpatient solution, with ongoing support provided during and after the implementation process. Clinicians use the web-based Dignio platform to review completed PROM questionnaires from patients before a scheduled consultation, assign tasks, provide e-learning, and distribute digital information to individual patients and groups. During the outpatient clinic's operating hours on weekdays, diabetes specialist nurses asynchronously respond to chat messages from patients, consulting a physician if necessary. On the patients' side, the solution's app enables them to complete assigned PROM questionnaires before consultations, receive information and e-learning, cancel consultations, reschedule or change upcoming consultations at the outpatient clinic to phone- or video consultations, and chat asynchronously with their clinicians as needed. The chat feature is mainly used by patients to report and ask questions about nonacute medical issues (self-initiated patient-reported outcomes) and submit practical queries. Should immediate assistance be required, patients are instructed to either call the clinic directly or dial the national emergency number.

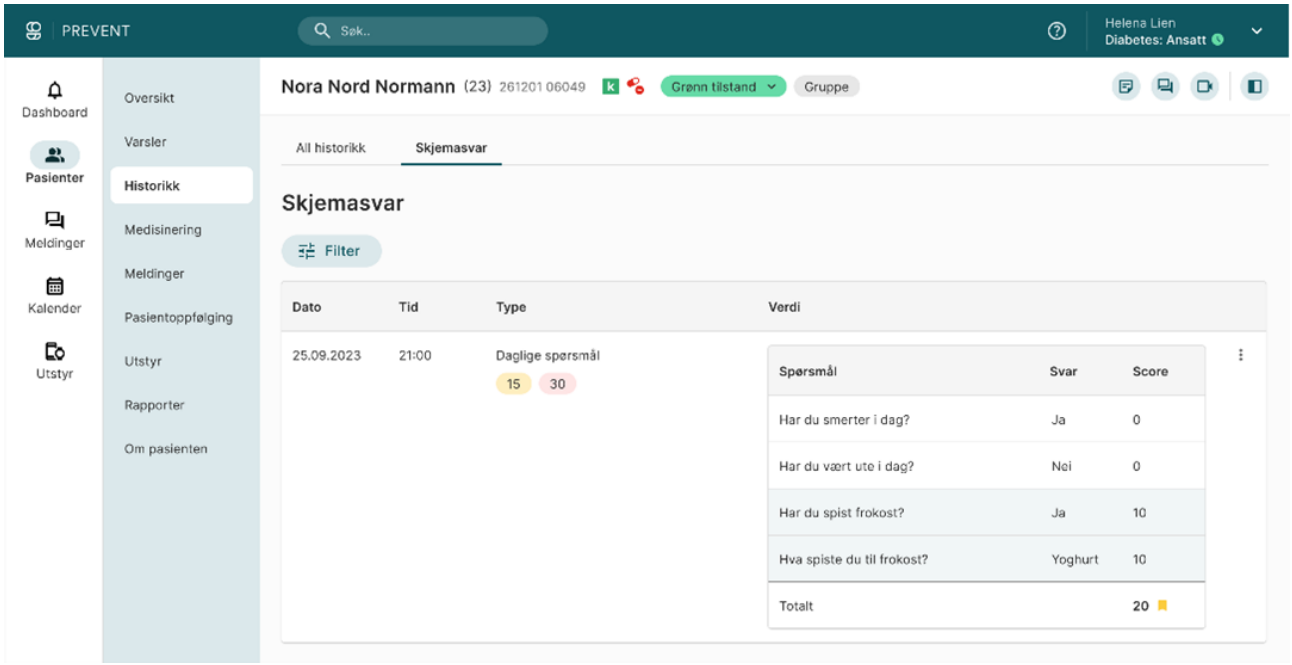
The PROM questionnaire sent to patients before consultations is a modified and translated version of the DiabetesFlex Care PROM questionnaire originating from Denmark [13]. The questionnaire is multidimensional and is triaged based on preset thresholds using a traffic light system, which indicates patient responses' severity [13]. Norwegian adjustments to the Danish version were made based on a user involvement study conducted at Akershus University Hospital, involving patients, diabetes specialist nurses, endocrinologists, management, and researchers [23]. In addition to PROMs, the questionnaire contains open text boxes for subjective responses. Figure 1 illustrates the digital outpatient solution's different features. Figure 2 shows a screenshot of the web-based digital platform used by clinicians, while Figure 3 displays screenshots of the app used by patients.

Figure 1. The digital outpatient solution aims to individualize care through improved patient-clinician communication. Each feature can be used based on an individual patient's need for support. The triangular presentation illustrates the different levels of individualized care that each feature promotes, from uniform to highly individualized features. For example, while e-learning is a uniform feature received by those patients in need, the asynchronous chat is highly individualized and tailored to address a specific question at hand. PROM: patient-reported outcome measure.



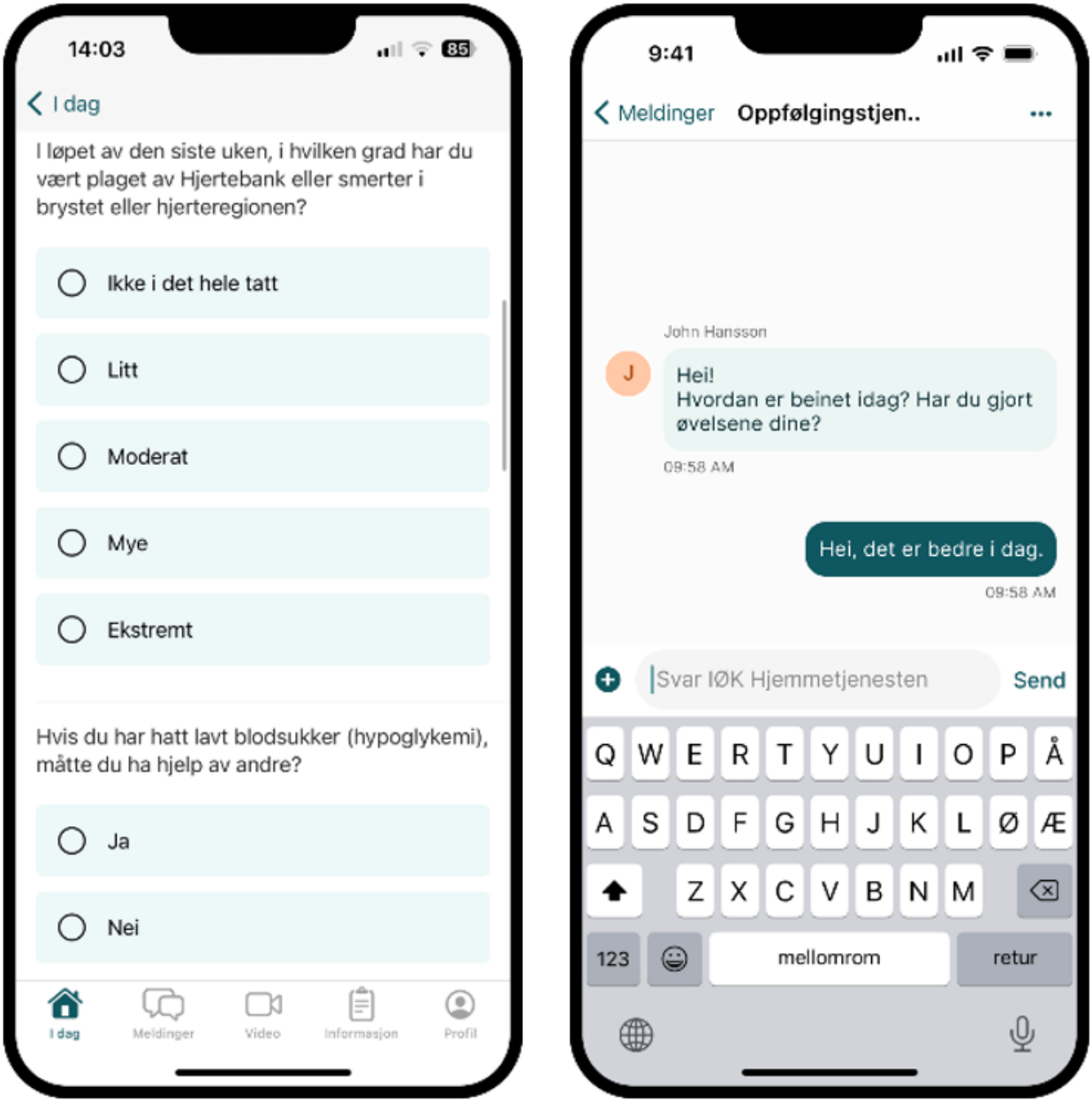
XSL•FO
RenderX

Figure 2. Screenshot from the dashboard used by clinicians. The screenshot is from a demo user and is intended as an example to illustrate the design.



XSL•FO
RenderX

Figure 3. Screenshots from the app used by patients. The screenshot on the left illustrates how the patient-reported outcome measure questionnaire appears in the app, while the screenshot on the right shows an example of a chat conversation. These screenshots are from a demo user and are intended to illustrate the design.



Participant Recruitment

Clinicians working with patients with type 1 diabetes at the outpatient clinic were recruited through convenience sampling [25], and the clinic manager invited them to participate. They were recruited based on availability, as we wanted to interfere as little as possible with clinical practice. All 10 clinicians invited to participate in the study, comprising 6 diabetes specialist nurses and 4 physicians, consented to participate. This represented approximately half the physicians and all the diabetes specialist nurses at the outpatient clinic, excluding the clinic manager and 1 newly hired nurse. Patients invited were adults with type 1 diabetes who had consented to use the digital outpatient solution and had completed the PROM questionnaire before their next consultation. Patients were recruited through purposive sampling [25] to ensure variation in gender, age, ethnicity, and frequency of digital interactions. Information about the study was sent to patients via the digital outpatient solution before their next upcoming consultation. At the start of the consultation, their clinician asked whether they wanted to participate. Of the 23 patients invited to participate, 20 consented. Written consent was obtained from all study participants.

Data Collection

The first author collected data through observations and semistructured interviews between August 2023 and April 2024. An observation guide (Multimedia Appendix 1) and interview guides (Multimedia Appendices 2 and 3) were developed based on informal observations that the first author conducted at the outpatient clinic between February and March 2023, as well as aspects of the Method for Assessment of Telemedicine

XSL•FO
RenderX

framework [28]. The guides were reviewed in a reference group comprising a user representative, researchers, clinicians, and management from the outpatient clinic. The observation guide facilitated insight into how the digital outpatient solution was used to individualize care, while the interviews sought to capture both clinicians' and patients' perspectives on their utilization of possibilities provided by the digital outpatient solution.

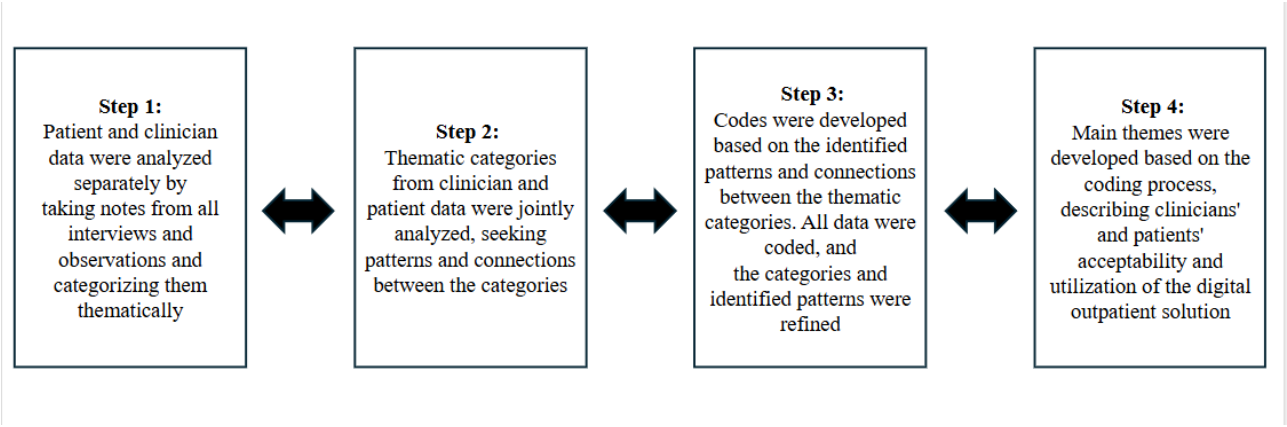
Scheduled consultations involving all participating patients and 6 out of 10 participating clinicians were observed. The observed consultations took place at the clinic (17), over the phone (2), or via video (1). Observations of remote consultations were conducted on-site, in the same room with the clinician at the outpatient clinic. Observation notes (40 A4 pages) were written down manually in the observation guide, digitally recorded, and stored. The notes in the observation guide contained descriptions of the type of consultation, topics discussed, context, use of PROM responses during the consultations, and analytical thoughts.

Individual semistructured interviews with clinicians were conducted before the observations, whereas patient interviews were conducted within 2 weeks after the observed consultations. The interviews were conducted at the clinic, over the phone, or via video based on participants' preferences regardless of whether they were clinicians or patients. During the interviews, participants were asked questions about their attitudes toward the use of health technology in general, experiences with the implementation of the digital outpatient solution and using different features, the solution's perceived usefulness, and potential improvements. Clinician interviews lasted between 28 and 52 minutes (median 37 minutes), and patient interviews lasted for 20-47 minutes (median 28 minutes). All interviews were transcribed verbatim (267 A4 pages).

Data Analysis

Data analysis was conducted in line with interpretive description [25]. We concurrently collected data and conducted a constant comparative analysis. This made the process dynamic by moving back and forth between the different steps of the analysis, as illustrated in Figure 4.

Figure 4. Illustration of the dynamic process of the constant comparative analysis.



The analysis aimed to inductively explore how clinicians and patients utilize a digital outpatient solution to individualize care. Through our interpretation process, we revealed a strong link between acceptability and utilization. In the final step of the analysis, 3 main themes were developed based on the coding process that describe clinicians' and patients' acceptance and utilization of the digital outpatient solution. The theoretical framework for acceptability was used to determine that the findings pertain to acceptance and utilization [19]. The first author conducted the analysis, with ALJ, HH, and AT contributing to analytical discussions during each step. These themes were discussed further with the remaining authors. NVivo 14 software (Lumivero) was used to manage and code the data [29], supplemented by visual exploration techniques to identify connections and patterns in the data. The visual exploration process involved positioning the thematic categories in relation to each other to identify patterns in clinicians' and patients' acceptance and utilization of the digital outpatient solution.

Credibility

To ensure the credibility and epistemological integrity of our study, we have provided a comprehensive description of the study design and analysis process [25]. We aimed to establish representative credibility through prolonged engagement with the outpatient clinic prior to data collection. Furthermore, we enhanced the credibility of our findings by triangulating data, collecting information from multiple sources [30]. By thoroughly describing our results, we aim to demonstrate our analytic logic and interpretive authority to ensure trustworthiness [25].

Ethical Considerations

The study was assessed by the Norwegian Agency for Shared Services in Education and Research (reference no. 456954) and the data protection officer at Akershus University Hospital (22/09709 - 13). The project did not require approval by the regional committees for medical and health research ethics (reference no. 407539).

We used a service for sensitive data (TSD), developed at the University of Oslo, which was designed for storing and processing sensitive data in compliance with the Norwegian Personal Data Act and Health Research Act [31]. Written consent was obtained either on paper and stored securely, or digitally via Nettskjema [32], requiring a governmental ID portal

XSL•FO
RenderX

login to access the information, and stored in TSD. All interviews were audiorecorded using the Nettskjema-diktafon app [33], a solution that immediately transfers all audio files encrypted for storage in TSD. MAM transcribed the interviews using the f4transkript program [34]. AT reviewed all recordings and transcripts throughout the analytical process, informing the interpretations. The privacy and confidentiality of research participants' data and identities were maintained.

Results

Participants' Characteristics

The patients' characteristics, including their use of the digital outpatient solution, are shown in Table 1, and the clinicians' characteristics, including their diabetes care experience, are shown in Table 2.

Table . Characteristics of patients and their use of diabetes technology and the digital outpatient solution (N=20).

Characteristics and usage	Values
Gender (female), n (%)	10 (50)
Age (years), median (minimum-maximum)	36 (18 - 65)
Diabetes duration (years), median (minimum-maximum)	21 (1 - 50)
Glucose monitoring, n (%)	
Continuous glucose monitor	19 (95)
Glucometer	1 (5)
Treatment, n (%)	
Insulin pump	12 (60)
Insulin pen	8 (40)
Number of completed PROM[a] questionnaires, median (minimum-maximum)	1 (1 - 7)
Number of chat conversations, median (minimum-maximum)	2 (0 - 11)
Has used remote consultations, n (%)	12 (60)

[a]PROM: patient-reported outcome measure.

Table . Characteristics of clinicians and their diabetes care experience (N=10).

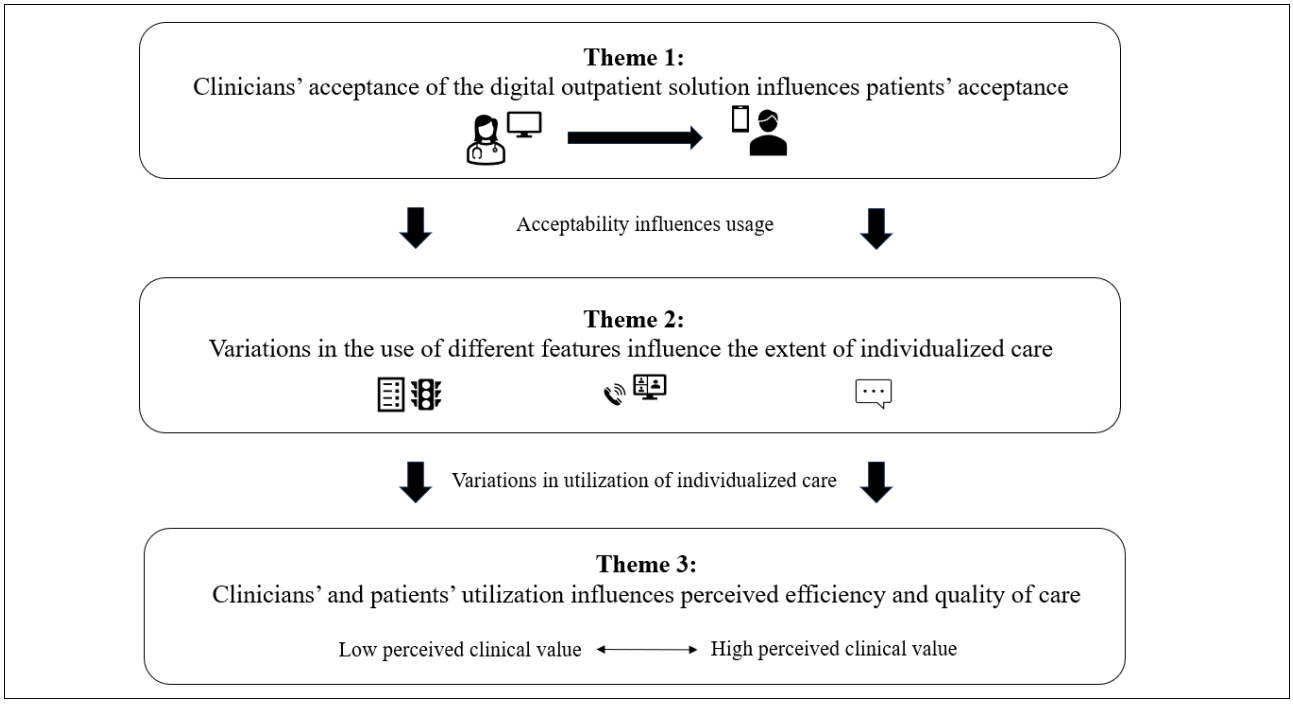
Characteristics and diabetes care experience	Values
Gender (female), n (%)	9 (90)
Diabetes specialist nurse, n (%)	6 (60)
Physician, n (%)	4 (40)
Nurses' diabetes care experience (years), minimum-maximum	5 - 25
Physicians' diabetes care experience (years), minimum-maximum	5 - 15

Overview of the Identified Patterns

Our analysis identified three main themes that comprise a pattern describing the complex dynamics of how clinicians and patients accept and use the digital outpatient solution to support communication and individualize care: (1) clinicians' acceptance of the digital outpatient solution influences patients' acceptance, (2) variations in the use of different features influence the extent of individualized care, and (3) clinicians' and patients' utilization influences perceived care efficiency and quality. Figure 5 provides an overview of how the 3 themes are connected. Notably, the participants did not emphasize the possibility of assigning and receiving tasks, information, and e-learning, so it was not a focus of the analysis.

XSL•FO
RenderX

Figure 5. Illustration of how clinicians and patients accept and use the digital outpatient solution to support individualized care.



Clinicians' Acceptance of the Digital Outpatient Solution Influences Patients' Acceptance

This theme describes how a variation was observed in how clinicians and patients accepted the digital outpatient solution, with an interplay between the two. Specifically, clinicians' acceptance seemed to shape patients' understanding of the solution's use and, subsequently, their acceptance.

Clinicians' understanding of the digital outpatient solution's potential and their confidence in using its different features varied. They also expressed variations in their personal commitment toward mastering the digital outpatient solution's features and adjusting their existing workflows. Some clinicians felt confident using the PROM questionnaires, chat, and remote consultations, viewing them as positive means to support individualized care:

> *I really like it. Both the digital outpatient solution and I absolutely love video consultations because they shorten the time spent during consultations and we stay more on topic. And with the apps we use, we have access to much more data than before.* [Clinician 1]

Conversely, other clinicians expressed uncertainty about using some features of the solution. While most felt comfortable using the chat and remote consultations, some were uncertain about how to use the PROM questionnaires before and during consultations, questioning whether the benefits outweighed the additional preparation time required of them and their patients:

> *I feel it is a very comprehensive and somewhat time-consuming questionnaire for patients to fill out. I'm a bit worried that patients might find it to be too much work.* [Clinician 10]

Furthermore, clinicians' own understanding and acceptance of the solution seemed to affect how they introduced the solution to patients. Clinicians who indicated a solid understanding of the solution's possibilities to individualize care expressed that they spent time explaining the various features to patients, during which they emphasized how it could enhance patients' self-management awareness and improve communication with their clinicians. Patients who felt well informed about the solution's potential also indicated higher acceptance levels:

> *As we talked, he realized that this was his tool, and he could expect that I had read his responses before consultations. So, it suited him very well. He also understood that it was a way to make him more aware of his disease… Of course, it creates an expectation that I, as a clinician, should keep myself updated on their responses, but with him, I also experienced a positive expectation, that now we are talking, we can clear big and small problems out of the way early.* [Clinician 3]

However, clinicians expressing less familiarity with the solution's possibilities often focused on the chat feature as a means for patients to communicate with the clinic between consultations. Some reported that they did not mention the use of PROM questionnaires at all when presenting the solution to patients. This led to uncertainty among patients who received these questionnaires before consultations without adequate information about their purpose:

> *It was a bit like, what is this, I thought…and it seemed that you had the option to decline to fill out the questionnaire. I believe there was an option to not fill out the questionnaire without losing the appointment. But I thought, OK, I'll just fill out the form, it's probably a questionnaire they have. But I thought the information was poor.* [Patient 14]

Simply receiving the questionnaire without any clear instructions about its intended use was insufficient to gain patients'

XSL•FO
RenderX

acceptance. Clinicians' and patients' acceptance of the digital outpatient solution influenced further to what extent they used the different features to improve communication and individualized care.

Variations in the Use of Different Features Influence the Extent of Individualized Care

This theme describes how clinicians and patients used PROM questionnaires, the option for remote consultation, and the chat feature to individualize care. It outlines the variations among participants' use, as well as the interplay between clinicians' and patients' usage patterns.

PROM Questionnaires Can Facilitate Individualized Consultations Through Mapping Needs and Reflection

The most pronounced differences in use of the solution were related to use of patients' PROM responses and how these responses influenced consultations. While some clinicians admitted to forgetting to review the patients' PROM responses before a consultation occasionally, most clinicians made it a routine part of their consultation preparation. Some clinicians used PROM responses to tailor consultations and set agendas, ensuring that the patient's needs were emphasized. When PROM responses were used actively, patients reported feeling more engaged in their treatment, with increased opportunities to discuss their concerns and priorities. This also included discussing positive PROM responses, as this could serve as a motivational factor in patients' self-management process:

> *I have previously felt that these consultations are a bit "so-so". We talked about blood glucose monitoring and maybe not so much more. So, you don't really know, and maybe the clinician doesn't know either, which direction the conversation will take. So, the fact that I have already thought about what I want to discuss, and they are aware of my thoughts and feelings, has added great value to my consultations.* [Patient 8]

When patients reported challenging issues through the questionnaire, clinicians could address these concerns specifically, improving their communication with their patients. However, some clinicians used the PROM responses less often during the consultations, which led their patients to question the value and purpose of the information they provided through the questionnaire:

> *What then is the value of the questionnaire? It was not mentioned and none of the elements I reported were brought up... so I question the value of it. Questionnaires for the sake of questionnaires?* [Patient 6]

If clinicians did not address patients' PROM responses during the consultations explicitly, we observed that most patients did not bring it up. However, patients noted that this lack of use made it more challenging to engage actively in their consultations. Still, most patients used the PROM questionnaire as a tool for reflection to some extent to inform and guide their self-management process and prepare for consultations.

Consultations Tailored to Needs and Personal Preferences

With the digital outpatient solution's implementation, most patients experienced a notable shift toward more individualized consultations. In addition to the use of PROM questionnaires, which affected the consultations' content, the shift was facilitated by patients being asked whether they wanted to meet remotely with their clinicians by either phone or video. Both clinicians and patients agreed that certain matters, such as physical checks or discussing difficult or complex issues such as mental health challenges, were best addressed in person. However, many found remote consultations more convenient for simpler requirements, such as support when starting to use a new insulin pump:

> *I have used it in the follow-up when patients have started using an insulin pump. Initially, a physical meeting is required to get started with the pump, but the subsequent follow-up is conducted digitally. This is possible because patients have uploaded their pump data. This allows us to see all the necessary information while they are at home, enabling us to make necessary adjustments.* [Clinician 1]

However, both clinicians and patients expressed personal preferences regarding using remote consultations. While some preferred communicating remotely and used it regularly, others had a strong preference for in-person consultations:

> *I have always chosen to decline that option because I prefer to speak with people face-to-face. I have used video meetings a lot for work and other settings, and that's not a problem, but I feel much more comfortable face-to-face.* [Patient 10]

Patients regularly chose between physical attendance at the clinic and video or phone consultations based on needs or personal preferences.

Chat Messages Facilitate Individualized Care When Needed

Most patients used the chat feature, which all the diabetes specialist nurses used consistently, facilitating targeted care within a short time frame. Chat was highlighted as the most frequently used feature of the digital outpatient solution. For clinicians, chat presented an opportunity to maintain contact with patients and even reduce consultation frequency if adequate follow-up could be conducted via chat:

> *We reach more patients when they have the opportunity to send a message. We dare, in a way, to schedule their next consultation in half a year because they can write to us in the meantime if they need to.* [Clinician 1]

When confronted with complex patient inquiries via chat that warranted comprehensive written responses or extensive dialogue with the patient, clinicians found transitioning to immediate remote consultations useful in providing the necessary support. These chat-initiated remote consultations, serving as an extension of the chat interaction, enriched the overall communication process:

XSL•FO
RenderX

> *When the patients write messages, we see how extensive the issue is, whether it's a short question, or if it's a question that requires having a conversation around what the patient is wondering about. So, I often ask if we can have a conversation instead of answering a dissertation...And there is a lot of good feedback on that. I also think we can resolve issues quickly so that they don't have to wait until their next scheduled consultation to get answers. So, I believe that we are promoting health more in the way we are working now.* [Clinician 3]

The patients found it easier to contact the clinic through chat. Moreover, patients reported that chat supported their self-management, as they used it to address urgent concerns immediately instead of waiting for their next consultation or calling the clinic. Many patients viewed phone calls as a barrier, and they experienced a lower threshold for writing their questions in chat. Patients and clinicians used the chat feature for a range of queries, from addressing practical questions about upcoming consultations to discussing treatment-related matters, such as insulin adjustment. Chat also provided a sense of security for patients, knowing that help was just a message away:

> *It certainly helps me to have the reassurance that I can simply send a message instead of having to call and wait for a response and all that, so it's very convenient. Let us say I'm at work and wondering about something, I can just send a message, and I know I'll get a response soon. Especially when I don't have the opportunity to be on the phone, it's a great comfort knowing there's someone there to help.* [Patient 4]

Although some patients had used chat only once or twice, or not at all, they reported a sense of security from knowing that they had the option to send a chat message when necessary. This underscores the chat feature's latent value, demonstrating its significance even when not used actively.

Clinicians' and Patients' Utilization Influences Perceived Care Efficiency and Quality

Both clinicians and patients acknowledged the digital outpatient solution's potential clinical value in enhancing efficiency by streamlining communication and improving care quality through individualized follow-up based on patients' needs. However, the extent of the perceived clinical value varied among clinicians and patients and was reflected in how they utilized different features of the solution. Clinicians' and patients' utilization, that is, their practical and effective use of the digital outpatient solution, appeared to influence their experiences regarding the solution's benefits, with those demonstrating extensive utilization typically perceiving the solution as more valuable. Furthermore, this theme describes clinicians' and patients' perspectives on how the different features can contribute to enhanced care efficiency and quality.

The most prominent differences in clinicians' and patients' perceived clinical value were related to the PROM questionnaires. When utilized, both clinicians and patients found that the questionnaires helped focus on patient-identified key areas for managing their life with type 1 diabetes, simultaneously making consultations more efficient by prioritizing topics. From the patients' perspective, the PROM questionnaires were also viewed as the clinic's acknowledgment of the complex nature of living with type 1 diabetes. This comprehensive focus on all aspects of living with diabetes was believed to enhance care quality by enabling the patients to address challenging issues more comfortably:

> *It's easier to sit and write it on the phone than to bring it up in a consultation because the questionnaire is a lot about your thoughts and feelings, and it is not always easy to talk about… and when you are in a consultation, they often ask you how you are doing and then you usually just say that it's going well even though it's not.* [Patient 11]

However, patients stressed that the extent to which clinicians utilized their responses was crucial in perceiving the questionnaire as valuable from the patients' perspective. Although not all patients experienced remote consultations, most recognized their clinical value in increasing efficiency. Some patients also expressed a wish to try remote consultations in the future, particularly if circumstances prevented them from meeting clinicians in person at the outpatient clinic. Both patients and clinicians appreciated remote consultation's potential to increase efficiency and flexibility. Patients clearly benefited from a reduction in travel time, while some clinicians valued the possibility of working remotely. Both parties also noted that remote consultations tended to be more concise, without the usual small talk evident during in-person consultations. However, this led to a few clinicians expressing concerns about potentially missing important information, with some worrying that care quality provided remotely was not as high as in person. Most patients did not share the same concern, as they would choose clinic consultations for issues they felt were better discussed in person.

Both clinicians and patients emphasized the chat feature's role in elevating care efficiency and quality. Clinicians found that chat reduced the volume of calls to the clinic, allowing them to respond at their convenience, rather than being interrupted by calls during other tasks. Although responses from the clinic could take a few hours or even a couple of days, patients perceived the overall process as more efficient and flexible than calling. Furthermore, chat enabled clinicians to give more thought-out advice, differing from the immediate responses provided over the phone and enhancing the perceived quality of the responses provided to patients:

> *After nearly two years of using the solution, we appreciate receiving written questions, as it allows us to respond at our convenience. This also allows us to thoroughly consider our response and consult with a colleague before replying if necessary. I think this is a good approach.* [Clinician 4]

Patients appreciated the chat feature's value in offering an accessible way to receive reassuring responses to their questions, thereby easing everyday management of their condition. Both clinicians and patients viewed the chat feature, largely driven by patient-initiated conversations, as a key element of the digital

XSL•FO
RenderX

outpatient solution for individualizing care to meet patients' needs.

Discussion

Principal Findings

This qualitative study provides insight into how clinicians and patients accept and utilize a digital outpatient solution in a diabetes outpatient clinic differently, highlighting an interplay between their usage patterns. Specifically, clinicians' acceptance seems to shape patients' understanding and acceptance of the solution's resources, subsequently influencing clinicians' and patients' use. Furthermore, this seems to influence clinicians' and patients' perceived clinical value of the solution's different features. Those demonstrating higher acceptance and more extensive utilization perceived the solution as more valuable, recognizing its potential to enhance care efficiency and quality.

To our knowledge, no extant literature has highlighted the interplay between clinicians' and patients' acceptance, utilization, and perceived clinical value of digital outpatient solutions during full-scale implementation in diabetes specialist outpatient care. The literature mostly encompasses pilot and feasibility studies, with an emphasis either on the patients' or the clinicians' perspective [4]. Although the literature in this field is limited, studies have indicated that patients generally embrace PROM-based digital outpatient solutions [16,22], whereas clinicians often express ambivalence and challenges related to time, resources, and care quality [11,12,16]. Our findings particularly align with how clinicians express ambivalence about using digital outpatient solutions, revealing variations in satisfaction and use of the solution. Moreover, our study goes further than extant studies in exploring the complex interplay between clinicians and patients, specifically by demonstrating that clinicians' acceptance influences patients' acceptance and their collective utilization of the solution. This interplay underscores the important role of clinicians' understanding and acceptance of the digital outpatient solution for successful implementation [19].

The importance of clinicians' acceptance has been emphasized in literature examining telehealth services in different contexts [18]. Regarding this specific multicomponent digital solution, our findings suggest that both clinicians and patients perceived combining different features to improve the efficiency and quality of individualized care as clinically valuable and appropriate for implementation in clinical practice [19]. However, variations in acceptability across individuals seemingly affect clinicians' and patients' perceived clinical value of each feature, influencing their utilization in the sense of user engagement and intervention effectiveness [20]. For instance, most participants used the chat feature and found it valuable for individualizing care, aligning with extant literature arguing that a chat feature can strengthen clinician-patient relationships in type 1 diabetes care [35]. Remote consultations were used based on individual preferences and needs and were valued particularly for their role in enhancing communication efficiency, a benefit also reported in earlier literature [36]. However, we observed wide variations in use and perceived clinical value of PROM questionnaires. Extant literature has argued that PROMs' usefulness in clinical practice depends on patients reporting their concerns, clinicians' proficiency in utilizing information, and both patients and clinicians perceiving them as useful [37]. This suggests that the effectiveness of using PROMs to individualize care depends on clinicians' and patients' ambition and capacity to leverage opportunities that PROMs offer. Without active engagement from both clinicians and patients, PROM questionnaires can be perceived as a burden rather than an asset, adding to workloads without enhancing perceived clinical value.

Implications for Clinical Practice

Overview

To ensure successful implementation, it is crucial to address both human factors and organizational change [38]. We argue that an in-depth understanding of human factors, such as users' acceptance, is essential when implementing a digital outpatient solution in diabetes care. Based on our findings, we emphasize the importance of clinic management in addressing the challenges and implications that clinicians and patients encounter during implementation, as these factors can significantly impact the solution's sustainable utilization.

Importance of User Acceptance and the Clinicians' Role in Implementation

Based on our findings, which underscore the important role of clinicians' acceptance in the sustainable utilization of the digital outpatient solution, we argue that clinicians should be a prioritized group during implementation. When clinicians accept and understand how the digital outpatient solution can be utilized, they are in a better position to provide the necessary information for patients to start using it. Extant literature has recognized clinicians' role in implementing digital outpatient solutions successfully [18] and has suggested further exploration of workflow changes that clinicians encounter when implementing digital solutions in diabetes care [4]. Furthermore, clinicians' role has also been highlighted in the context of other diagnostic groups, such as epilepsy, in which it has been argued that PROM-based follow-up represents substantially changed outpatient care and that implementation success is highly dependent on clinicians' daily management of the solution [38].

The Need for Organizational Involvement and Strategy

Ensuring clinicians' acceptance and competence to provide the necessary information and education to patients about how to utilize digital outpatient solutions seems to require organizational strategies that emphasize education, training, and ongoing support throughout and after implementation. As part of the implementation of this digital outpatient solution, all clinicians underwent training and were offered ongoing support as needed. Nevertheless, the persistence of some clinicians' hesitance to change implies that more comprehensive organizational involvement may be necessary for optimal implementation. Strategies that address clinicians' individual support needs might help foster more universal acceptance among clinicians.

Furthermore, our findings indicate that the extent of organizational involvement required for implementation to be

XSL•FO
RenderX

successful can differ, depending on the effort required from clinicians and patients to adapt to the new practices. In this study, the implementation of the chat feature and offering patients the choice of remote consultations were viewed as particularly valuable. However, the implementation of PROM questionnaires required considerable changes in clinicians' workflow and their care delivery methods, while concurrently requiring increased patient involvement. Extant literature has highlighted the need for clinicians to adapt their care methods when implementing PROMs in diabetes care [12] and that the efficiency of using PROMs in clinical practice seems to increase in line with clinicians' and patients' understanding of why and how they are used [37]. Consequently, we suggest that more structured education and follow-up from clinic management are needed on the use of PROM questionnaires, compared with other, more easy-to-use features of the digital outpatient solution, such as the chat feature and remote consultations.

Change management, a systematic approach to managing the human elements of change [39], could inspire strategies to ensure utilization when implementing a digital outpatient solution. This approach involves a set of processes designed to guide an organization through a transition [39] and has been applied in other contexts implementing digital services [40], as well as other settings in health care services [41]. We propose that using a change management approach adapted to the specific context could facilitate future implementations of digital outpatient solutions in diabetes specialist outpatient care. However, implementing change notably often requires more time and resources than anticipated, particularly under pressure [42]. Given the increasing patient population in diabetes care [2], as well as the constraints of limited health care resources, potential implementation challenges should be considered before introducing new digital outpatient solutions in diabetes specialist outpatient care. Thoughtful consideration of potential implementation challenges and proactive planning for their resolution could enhance the likelihood of successful implementation.

Limitations

Despite providing valuable insights, this study has some limitations. The data, collected from 1 single diabetes outpatient clinic and specific to 1 particular digital outpatient solution, may not present patterns transferable to other settings or digital solutions. However, we did find that patients and clinicians were reporting similar experiences in other studies, although the content and delivery methods in their digital outpatient solutions differed [11,12,16,22]. These studies were also conducted in Scandinavian countries, providing a patient population and health care setting comparable with our study.

Another potential limitation is that some of the clinicians interviewed in this study also participated in the user involvement study previously conducted at Akershus University Hospital [23], which aimed to develop this digital outpatient solution and the care trajectory. The clinicians' involvement in the solution's development process may have influenced their attitudes and understanding of its use. This implies that not all clinicians may have had the same foundation or knowledge to utilize the solution, potentially contributing to observed variations in acceptance, utilization, and perceived clinical value. Still, this variation in clinicians contributes to bringing different perspectives with the digital solution into the study, and the findings emphasize the importance of understanding why and how such solutions are used for their effective utilization.

All patients recruited for this study had used the solution to some extent, having completed at least 1 PROM questionnaire. We chose to focus on this group to gain insight into the utilization of the different features. However, we acknowledge a limitation in that we have not considered patients who have accepted using the digital outpatient solution but have not done so for various reasons. Understanding the barriers that these patients face in using the solution could provide valuable insight, potentially helping to facilitate a smoother transition to digital follow-up for patients currently not using the solution.

Finally, as this is a qualitative study, we are gaining deep insight into how the solution is utilized in clinical practice. Although our sample size concerning the clinicians could have been larger, it still contained 96 A4 pages of transcribed data. This shows a high degree of information power [43]. However, using a mixed methods design or further investigating patients' and clinicians' acceptability and utilization quantitatively could deepen our understanding and enhance the validity and reliability of our findings.

Conclusions

This study explored the interplay between clinicians' and patients' acceptance, utilization, and perceived clinical value of a digital outpatient solution in diabetes specialist outpatient care. We identified variations in the use of the solution's features, suggesting that certain features might require greater knowledge and efforts from clinicians and patients to be utilized efficiently, particularly the use of PROM questionnaires. Our findings imply that when clinicians and patients understand why and how such solutions are used, they can enhance care efficiency and quality, contributing to sustainable health care services. We argue that an in-depth understanding of clinicians' acceptance is essential when implementing a digital outpatient solution. Furthermore, using a change management approach, adapted to the specific context, could facilitate future implementations to ensure sustainable utilization. Future research should look deeper into clinicians' and patients' acceptance, as well as use of change management strategies when implementing digital outpatient solutions in diabetes specialist outpatient care.

Acknowledgments

The authors want to thank the members of their reference group for their valuable insight into the study's development, as well as all the clinicians and patients who consented to participate. This study has received funding from Akershus University Hospital

XSL•FO
RenderX

and Oslo Metropolitan University. Sikt AI chat, a GPT technology that ensures privacy and data security [44], was used for language improvements during manuscript preparation.

Data Availability

The datasets generated in this study are available only in Norwegian, and as they are part of a larger and ongoing study, data are currently not publicly available. Please contact the corresponding author for further information. The Norwegian version of the DiabetesFlex questionnaire can be made available upon request to TS. The English version 1 is published elsewhere [13].

Authors' Contributions

All authors contributed to the study design. TS and MAM recruited the participants, and MAM conducted data collection and transcribed all the material. AT reviewed the transcripts and MAM conducted the data analysis, with ALJ, HH, and AT contributing to analytical discussions. TS and JAW ensured that the content was relevant for clinical practice. EÅ reviewed the content in light of his knowledge of digital solutions in diabetes care. MAM drafted the manuscript, and all authors reviewed and edited the content. All authors take responsibility for the final manuscript's content.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Observation guide.
[DOCX File, 25 KB - humanfactors_v12i1e70301_app1.docx]

Multimedia Appendix 2
Interview guide—clinicians.
[DOCX File, 32 KB - humanfactors_v12i1e70301_app2.docx]

Multimedia Appendix 3
Interview guide—patients.
[DOCX File, 36 KB - humanfactors_v12i1e70301_app3.docx]

References

1. Graham D Ogle FW, Gregory GA, Maniam J. Type 1 diabetes estimates in children and adults. : International Diabetes Federation; 2022 URL: https://diabetesatlas.org/media/uploads/sites/3/2025/03/IDF-T1D-Index-Report.pdf [accessed 2025-06-15]
2. Gregory GA, Robinson TIG, Linklater SE, et al. Global incidence, prevalence, and mortality of type 1 diabetes in 2021 with projection to 2040: a modelling study. Lancet Diabetes Endocrinol 2022 Oct;10(10):741-760. [doi: 10.1016/S2213-8587(22)00218-2]
3. Herman WH. The global burden of diabetes: an overview. In: Dagogo-Jack S, editor. Diabetes Mellitus in Developing Countries and Underserved Communities: Cham: Springer; 2017:1-23. [doi: 10.1007/978-3-319-41559-8_1]
4. Torbjørnsen A, Jensen AL, Singstad T, Weldingh NM, Holmen H. Patient-reported outcome measures in diabetes outpatient care: a scoping review. BMJ Open Diabetes Res Care 2023 Nov;11(6):e003628. [doi: 10.1136/bmjdrc-2023-003628] [Medline: 37963648]
5. Holt RIG, DeVries JH, Hess-Fischl A, et al. The management of type 1 diabetes in adults. A consensus report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). Diabetologia 2021 Dec;64(12):2609-2652. [doi: 10.1007/s00125-021-05568-3] [Medline: 34590174]
6. Reddy M, Rilstone S, Cooper P, Oliver NS. Type 1 diabetes in adults: supporting self management. BMJ 2016 Mar 10;352:i998. [doi: 10.1136/bmj.i998] [Medline: 26965473]
7. Jones A, Vallis M, Pouwer F. If it does not significantly change HbA1c levels why should we waste time on it? A plea for the prioritization of psychological well-being in people with diabetes. Diabet Med 2015 Feb;32(2):155-163. [doi: 10.1111/dme.12620] [Medline: 25354315]
8. Weldring T, Smith SMS. Patient-reported outcomes (PROs) and patient-reported outcome measures (PROMs). Health Serv Insights 2013;6(61-8):61-68. [doi: 10.4137/HSI.S11093] [Medline: 25114561]
9. Greenhalgh J, Gooding K, Gibbons E, et al. How do patient reported outcome measures (PROMs) support clinician-patient communication and patient care? A realist synthesis. J Patient Rep Outcomes 2018 Dec;2:42. [doi: 10.1186/s41687-018-0061-6] [Medline: 30294712]
10. Santana MJ, Haverman L, Absolom K, et al. Training clinicians in how to use patient-reported outcome measures in routine clinical practice. Qual Life Res 2015 Jul;24(7):1707-1718. [doi: 10.1007/s11136-014-0903-5] [Medline: 25589231]

XSL•FO
RenderX

11. Haugstvedt A, Hernar I, Graue M, et al. Nurses' and physicians' experiences with diabetes consultations and the use of dialogue tools in the DiaPROM pilot trial: a qualitative study. Diabet Med 2021 Jun;38(6):e14419. [doi: 10.1111/dme.14419] [Medline: 33035378]
12. Nielsen BK, Terkildsen MD, Jensen AL, Pedersen MZ, Hollesen MR, Lomborg K. Endocrinologists' use of patient-reported outcome measures in the care of people with diabetes: a qualitative study. Diabetes Epidemiol Manag 2023 Oct;12:100159. [doi: 10.1016/j.deman.2023.100159]
13. Laurberg T, Schougaard LMV, Hjollund NHI, Lomborg KE, Hansen TK, Jensen AL. Randomized controlled study to evaluate the impact of flexible patient-controlled visits in people with type 1 diabetes: the DiabetesFlex Trial. Diabet Med 2022 May;39(5):e14791. [doi: 10.1111/dme.14791] [Medline: 35028992]
14. Hernar I, Graue M, Richards D, et al. Electronic capturing of patient-reported outcome measures on a touchscreen computer in clinical diabetes practice (the DiaPROM trial): a feasibility study. Pilot Feasibility Stud 2019;5:29. [doi: 10.1186/s40814-019-0419-4] [Medline: 30820340]
15. Bachmeier CAE, Waugh C, Vitanza M, et al. Diabetes care: addressing psychosocial well-being in young adults with a newly developed assessment tool. Intern Med J 2020 Jan;50(1):70-76. [doi: 10.1111/imj.14355] [Medline: 31081194]
16. Skovlund SE, Troelsen LH, Noergaard LM, Pietraszek A, Jakobsen PE, Ejskjaer N. Feasibility and acceptability of a digital patient-reported outcome tool in routine outpatient diabetes care: mixed methods formative pilot study. JMIR Form Res 2021 Nov 3;5(11):e28329. [doi: 10.2196/28329] [Medline: 34730545]
17. Torbjørnsen A, Spildo I, Mollestad MA, et al. Investigating digital patient-reported outcome measures in patient-centered diabetes specialist outpatient care (DigiDiaS): protocol for a multimethod prospective observational study. JMIR Res Protoc 2024 Mar 5;13:e52766. [doi: 10.2196/52766] [Medline: 38441955]
18. Wade VA, Eliott JA, Hiller JE. Clinician acceptance is the key factor for sustainable telehealth services. Qual Health Res 2014 May;24(5):682-694. [doi: 10.1177/1049732314528809] [Medline: 24685708]
19. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res 2017 Jan 26;17(1):88. [doi: 10.1186/s12913-017-2031-8] [Medline: 28126032]
20. Perski O, Short CE. Acceptability of digital health interventions: embracing the complexity. Transl Behav Med 2021 Jul 29;11(7):1473-1480. [doi: 10.1093/tbm/ibab048] [Medline: 33963864]
21. Hernar I, Graue M, Strandberg RB, et al. Young adults with type 1 diabetes and their experiences with diabetes follow-up and participation in the DiaPROM pilot trial: a qualitative study. Diabet Med 2021 Jun;38(6):e14535. [doi: 10.1111/dme.14535] [Medline: 33547702]
22. Jensen AL, Schougaard LMV, Laurberg T, Hansen TK, Lomborg K. Flexible patient-reported outcome-based telehealth follow-up for type 1 diabetes: a qualitative study. Scand J Caring Sci 2023 Sep;37(3):662-676. [doi: 10.1111/scs.13154] [Medline: 36775917]
23. Holmen H, Singstad T, Ribu L, Jensen AL, Weldingh NM, Torbjørnsen A. Adapting a patient-reported outcome measure to digital outpatient specialist health care services for type 1 diabetes: user involvement study. JMIR Hum Factors 2022 Nov 15;9(4):e38678. [doi: 10.2196/38678] [Medline: 36378513]
24. Eeg-Olofsson K, Johansson UB, Linder E, Leksell J. Patients' and health care professionals' perceptions of the potential of using the digital diabetes questionnaire to prepare for diabetes care meetings: qualitative focus group interview study. J Med Internet Res 2020 Aug 19;22(8):e17504. [doi: 10.2196/17504] [Medline: 32812884]
25. Thorne S. Interpretive Description: Qualitative Research for Applied Practice: Routledge; 2016.
26. Solution. Dignio. URL: https://dignio.com/solution/ [accessed 2025-06-15]
27. Helsenett N. About Norsk helsenett—connecting Norwgian health services. Norsk helsenett. URL: https://www.nhn.no/om-oss/about-us-connecting-norwegian-health-services [accessed 2025-06-06]
28. Kidholm K, Ekeland AG, Jensen LK, et al. A model for assessment of telemedicine applications: MAST. Int J Technol Assess Health Care 2012 Jan;28(1):44-51. [doi: 10.1017/S0266462311000638] [Medline: 22617736]
29. Allsop DB, Chelladurai JM, Kimball ER, Marks LD, Hendricks JJ. Qualitative methods with Nvivo software: a practical guide for analyzing qualitative data. Psych 2022;4(2):142-159. [doi: 10.3390/psych4020013]
30. Farmer T, Robinson K, Elliott SJ, Eyles J. Developing and implementing a triangulation protocol for qualitative health research. Qual Health Res 2006 Mar;16(3):377-394. [doi: 10.1177/1049732305285708] [Medline: 16449687]
31. About TSD. University of Oslo. URL: https://www.uio.no/english/services/it/research/sensitive-data/about/index.html [accessed 2025-06-06]
32. Short introduction to nettskjema. University of Oslo. 2018. URL: https://www.uio.no/english/services/it/adm-services/nettskjema/about-nettskjema.html [accessed 2025-06-06]
33. Nettskjema-diktafon. University of Oslo. 2019. URL: https://www.uio.no/english/services/it/adm-services/nettskjema/help/nettskjema-dictaphone.html [accessed 2025-06-06]
34. F4transcript. University of Oslo. 2024. URL: https://www.sv.uio.no/psi/english/research/groups/quali-lab/Qualitative%20Research%20Portal/interview-and-transcribing/f4transcript [accessed 2025-06-06]

XSL•FO
RenderX

35. Miele F, Clementi S, Gennaro R, et al. Text messaging and type 1 diabetes management: qualitative study exploring interactions among patients and health care professionals. JMIR Diabetes 2019 May 10;4(2):e11343. [doi: 10.2196/11343] [Medline: 31094332]
36. Bertuzzi F, Stefani I, Rivolta B, et al. Teleconsultation in type 1 diabetes mellitus (TELEDIABE). Acta Diabetol 2018 Feb;55(2):185-192. [doi: 10.1007/s00592-017-1084-9] [Medline: 29209814]
37. Carfora L, Foley CM, Hagi-Diakou P, et al. Patients' experiences and perspectives of patient-reported outcome measures in clinical care: a systematic review and qualitative meta-synthesis. PLoS One 2022;17(4):e0267030. [doi: 10.1371/journal.pone.0267030] [Medline: 35446885]
38. Mejdahl CT, Schougaard LMV, Hjollund NH, Riiskjær E, Thorne S, Lomborg K. PRO-based follow-up as a means of self-management support - an interpretive description of the patient perspective. J Patient Rep Outcomes 2017;2(1):38. [doi: 10.1186/s41687-018-0067-0] [Medline: 30238083]
39. Hiatt JM, Creasey TJ. Change Management: The People Side of Change: Prosci; 2003.
40. Kho J, Gillespie N, Martin-Khan M. A systematic scoping review of change management practices used for telemedicine service implementations. BMC Health Serv Res 2020 Sep 1;20(1):815. [doi: 10.1186/s12913-020-05657-w] [Medline: 32873295]
41. Harrison R, Fischer S, Walpola RL, et al. Where do models for change management, improvement and implementation meet? A systematic review of the applications of change management models in healthcare. J Healthc Leadersh 2021;13:85-108. [doi: 10.2147/JHL.S289176] [Medline: 33737854]
42. Chreim S, Williams BEB, Janz L, Dastmalchian A. Change agency in a primary health care context: the case of distributed leadership. Health Care Manage Rev 2010;35(2):187-199. [doi: 10.1097/HMR.0b013e3181c8b1f8] [Medline: 20234224]
43. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res 2016 Nov;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
44. Sikt AI chat. Sikt. URL: https://sikt.no/en/tjenester/sikt-ai-chat [accessed 2025-06-15]

Abbreviations

PROM: patient-reported outcome measure

Edited by A Kushniruk; submitted 19.12.24; peer-reviewed by KL Mauco, M Grady; revised version received 29.04.25; accepted 12.05.25; published 08.07.25.

Please cite as:
Mollestad MA, Jensen AL, Holmen H, Singstad T, Årsand E, Winther JA, Torbjørnsen A
Exploring Clinicians' and Patients' Acceptance and Utilization of a Digital Solution to Support Individualized Care in Diabetes Specialist Outpatient Care (DigiDiaS): Qualitative Study
JMIR Hum Factors 2025;12:e70301
URL: https://humanfactors.jmir.org/2025/1/e70301
doi:10.2196/70301

© Maria Aadland Mollestad, Annesofie Lunde Jensen, Heidi Holmen, Tone Singstad, Eirik Årsand, Jacob Andreas Winther, Astrid Torbjørnsen. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 8.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Adapting a Text Messaging Intervention to Improve Diabetes Medication Adherence in a Spanish-Speaking Population: Qualitative Study

Jacqueline Seiglie[1,2], MD; Seth Tobolsky[2], MD; Ricaurte Crespo Trevino[3], MD; Lluvia Cordova[2], BS; Sara Cromer[1,2], MD; A Enrique Caballero[4], MD; Margarita Alegria[5,6], PhD; J Jaime Miranda[7], MD, PhD; Deborah Wexler[1,2*], MD, MSc; Lindsay Mayberry[8*], PhD

[1]Department of Medicine, Harvard Medical School, Boston, MA, United States

[2]Diabetes Unit, Massachusetts General Hospital, 50 Staniford St, Suite 340, Boston, MA, United States

[3]Mass General Brigham Salem Hospital, Salem, MA, United States

[4]Division of Endocrinology, Diabetes and Hypertension, Brigham and Women's Hospital, Boston, MA, United States

[5]Disparities Research Unit, Department of Medicine, Massachusetts General Hospital, Boston, MA, United States

[6]Departments of Medicine & Psychiatry, Harvard Medical School, Boston, MA, United States

[7]Sydney School of Public Health, Faculty of Medicine and Health, University of Sydney, Sydney, Camperdown, Australia

[8]Vanderbilt University Medical Center, Nashville, TN, United States

[*]these authors contributed equally

Corresponding Author:
Jacqueline Seiglie, MD
Department of Medicine, Harvard Medical School, Boston, MA, United States

Abstract

Background: Latino adults with type 2 diabetes (T2D) have higher rates of diabetes medication nonadherence than non-Hispanic White adults. REACH (Rapid Encouragement/Education And Communications for Health) is a text message platform based on the information-motivation-behavioral skills model that addresses barriers to adherence and was shown to improve adherence and glycated hemoglobin (HbA1c) levels, but it is only available in English.

Objective: This study aimed to report the multiphase, stakeholder-driven adaptation of the REACH barriers to diabetes medication adherence content to a Latino population (REACH-Español).

Methods: This was a qualitative study using focus groups. We identified potentially eligible patients (≥18 y old, Latino ethnicity, Spanish-language preference, and T2D diagnosis) using a Mass General Brigham Hospital query. Eligible patients were invited to participate in a focus group conducted in Spanish between April 13 and November 9, 2023. A total of 5 focus groups were conducted. Focus groups 1‐3 centered on ranking 40 barriers to diabetes medication adherence (derived from REACH and the extant literature), whereas focus groups 4‐5 centered on translation and cultural modifications of the original SMS text message content associated with each of the REACH barriers. Barriers were mapped onto information-motivation-behavioral constructs. We used descriptive statistics to summarize participant characteristics. Focus groups were audio-recorded, professionally transcribed, and analyzed with thematic content analysis using NVivo (Lumivero).

Results: In total, 22 participants attended the focus groups. The mean (SD) age was 63.2 (11) years, 55% (n=10/22) were female, and the mean HbA1c level was 8.5%. All participants were born in Latin America or the Caribbean and spoke Spanish as their preferred language, and 54.5% (12/22) had completed middle-school education or less. Among the top 10 ranked barriers, 50% (n=5) corresponded to information, 20% (n=2) to social motivation, 20% (n=2) to behavioral skills, and 10% (n=1) to personal motivation. Personal motivation barriers (medication burden and fear of side effects) and behavioral skills (forgetting to take medication) emerged as important themes in the focus groups.

Conclusions: A stakeholder-driven approach to intervention adaptation identified and prioritized relevant barriers to diabetes medication adherence among Latino adults with T2D and facilitated the adaptation of the REACH platform to a Spanish-speaking population.

(JMIR Hum Factors 2025;12:e66668) doi:10.2196/66668

XSL•FO
RenderX

KEYWORDS

diabetes; diabetes mellitus; DM; type 2 diabetes; T2DM; type 1 diabetes; endocrinology; medication adherence; adherence; Latino adults; stakeholder engagement; text messaging; messaging; intervention; multiphase; nonadherence; glycated hemoglobin; HbA1c; Latino; stakeholder; medication; Spanish-speaking; population

Introduction

Suboptimal medication adherence, generally defined as taking less than 80% of prescribed medications [1], is reported by nearly 50% of adults with type 2 diabetes (T2D) [2,3] and is associated with poor glycemic management [1,2,4] and higher risk of all-cause mortality [5]. Rates of suboptimal diabetes medication adherence are higher in racial and ethnic minority populations [6,7], particularly among individuals with lower English language proficiency [8]. An estimated 60% of Latino adults with limited English language proficiency report nonadherence to diabetes medications, compared with 52% of Latino adults with English proficiency and 38% of non-Hispanic White individuals [8]. Unique barriers to diabetes medication adherence have been reported by Latino adults with T2D [8-18], including negative perceptions regarding diabetes medication use (particularly insulin) [9,12,18], misunderstanding of the ongoing need for glucose-lowering pharmacotherapy once glycated hemoglobin (HbA1c) has improved [17], and challenges with regimen complexity even with the use of a Spanish-language interpreter [8,19]. While interventions that address patient-reported barriers to diabetes medication adherence can improve glycemic management [20,21], interventions that address barriers to diabetes medication adherence of relevance to Latino adults with T2D are lacking.

Among numerous behavioral theories that explain medication adherence behavior [22], the information-motivation-behavioral skills (IMB) model is an important theory-based model for conceptualizing and addressing domains associated with diabetes medication adherence [23,24]. The IMB model is built on three key components that influence behavior change, that are (1) information (ie, accurate knowledge about the disease and medications), (2) motivation (ie, beliefs and attitudes plus social support and social norms), and (3) behavioral skills (ie, objective and perceived abilities) [23,25]. When applied to diabetes medication adherence behavior, the IMB components allow for the identification of barriers to medication adherence in each domain [25]. Several interventions that have incorporated this model to address self-reported barriers to medication adherence have shown improvements in diabetes medication adherence and glycemic management among adults with T2D, including among Latino adults [21,23]. However, patient-centered tools for Latino adults with T2D based on the IMB model that are scalable and that do not rely on in-person educational sessions, which can be resource-intensive, do not yet exist.

REACH (Rapid Encouragement/Education And Communications for Health) is an innovative text message tool based on the IMB model that delivers text message content tailored to individual patient-reported barriers to diabetes medication adherence [21,26]. REACH delivers daily diabetes self-management 1-way texts, daily 2-way texts that ask about medication adherence, and weekly texts that provide adherence feedback and encouragement based on responses to 2-way texts. REACH was shown to improve diabetes medication adherence (by an average of one-third to one-half day per week over 12 months), reduce barriers to diabetes medication adherence (including cost-related barriers), improve dietary behavior and physical activity, and significantly reduce HbA1c at 6 months compared with usual care (0.74% reduction among patients with baseline HbA1c≥8.5%) [21]. However, despite its efficacy, scalability, and high participant engagement [27], REACH has not been adapted to other contexts. REACH could serve as an important clinical tool to address barriers to diabetes medication adherence among Latino adults with T2D. In this study, we used a multistage, stakeholder-driven, iterative method of language adaptation to adapt the original REACH barrier content to a Latino, Spanish-speaking population with T2D, focusing on identifying barriers to diabetes medication adherence of relevance to this population.

Methods

Study Design, Setting, and Participant Eligibility

We conducted a qualitative study by carrying out focus groups at the Massachusetts General Hospital (MGH) Chelsea HealthCare Center, one of the MGH Community Health Care Centers that serve a patient population in the greater Boston area that is majority Latino or Spanish-speaking [28]. Eligible participants had a T2D diagnosis (based on diagnostic codes, confirmed with electronic health record review), were ≥18 years old, were taking at least 1 glucose-lowering medication, and reported Spanish as their preferred written and spoken language. We excluded patients with auditory or communication difficulties.

Participant Recruitment

We searched the MGH Research Patient Data Registry to identify individuals who received care at MGH within the previous 2 years and who were ≥18 years old, had a diagnosis of T2D, and had registered Spanish as their preferred language. A total of 3709 patients met this criteria, of whom 500 lived in the city of Chelsea. We mailed out opt-out letters to 350 of these participants (selected randomly), followed by telephone calls to screen for the eligibility criteria described above. Among 30 eligible patients willing to participate in the study, 22 unique participants attended one of the focus groups 1 - 5. A total of 10 participants who attended focus groups 1 - 3 also attended either focus group 4 or 5. The reason for participation in multiple focus groups was to engage patients in the iterative process of content adaptation, which took place in two steps: (1) identification of barriers to diabetes medication adherence of relevance to the study population (focus groups 1 - 3) and (2) adaptation of SMS text message content addressing the relevant barriers. Enrollment occurred between March and November of 2023. Given that the original REACH content was created with input from communication experts who edited the text messages to be readable and understandable (ie, written at the

XSL•FO
RenderX

sixth-grade reading level or below) [26], the adaptation of REACH barrier content was also guided by an emphasis on reading appropriateness for the study population. Thus, for eligible patients interested in participating in the study, we conducted a health literacy screening to tailor the focus group activities to the literacy level of the study population. We used the 1-item Brief Health Literacy Screening Tool (BRIEF) questionnaire ("How confident are you filling out medical forms by yourself?"), which has been validated for Spanish-speaking populations as a quick assessment of health literacy [29].

Ethical Considerations

Participants received US $50 for focus group attendance. The Mass General Brigham Institutional Review Board approved all aspects of the study (protocol 2022P002678). At the beginning of each focus group, participants received an IRB-approved information sheet, which outlined the objectives of the study, procedures, benefits and risks, confidentiality, as well as an emphasis on voluntary participation. All participants agreed to participate in the focus groups.

Measures

At the start of each focus group, participants were asked to fill out a sociodemographic and diabetes history questionnaire. The data that were collected through this questionnaire were age, sex, ethnicity, country of birth, preferred language, educational attainment, household income, health insurance, and diabetes history (age at diagnosis, diabetes medications, and complications).

REACH Barriers to Diabetes Medication Adherence

The development of the REACH barrier content has been described in studies by Nelson et al [25,26]. The REACH study team conducted a thorough review of published studies describing patient-reported barriers to adherence to glucose-lowering medications. The literature review yielded 68 barriers to taking oral glucose-lowering medications and 7 barriers to taking insulin; similar barriers were collapsed, resulting in 36 adherence barriers [25]. A 2-step interactive procedure was subsequently used to rank these barriers among adults with T2D. Each of the 36 barriers was mapped to the IMB constructs, and content experts drew on identified studies to develop SMS text messages addressing each barrier [26]. For instance, the barrier "My daily medicine routine is too complicated to keep track of" generated 21 SMS text messages to address this concern, such as "to help keep things simple, take your medications at the same time you're doing something else, like brushing your teeth or eating meals."

Adaptation Methodology

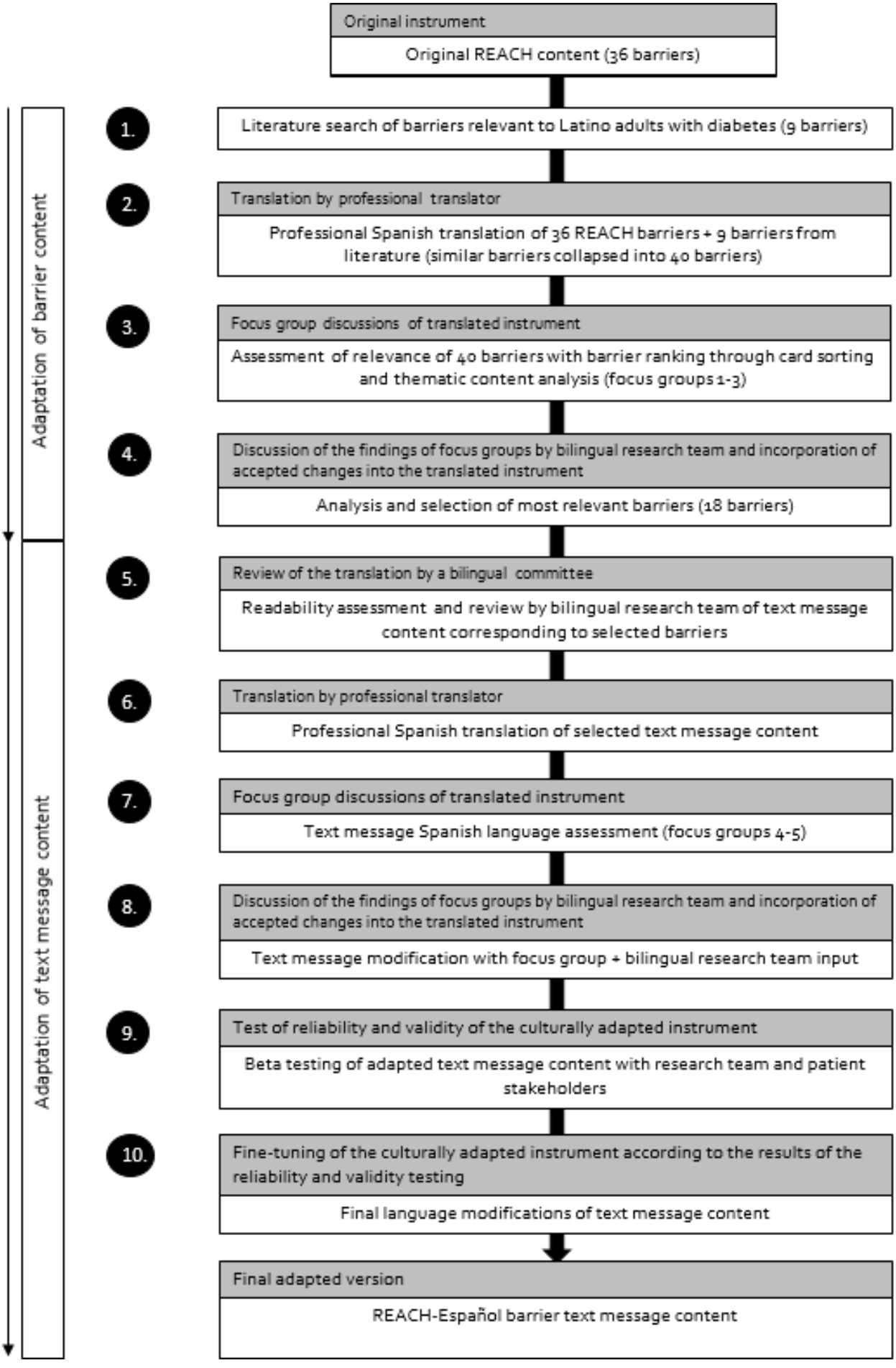
Description

To ensure that the REACH-Español barrier content was culturally and linguistically appropriate for the target population, we were guided by the multistage, iterative method for cross-cultural adaptation proposed by Matías-Carrelo et al [30] This multistage method relies on the linguistic review of the translated instrument by incorporating input from multiple stakeholders, including both a multinational bilingual committee composed of bilingual (English-Spanish) researchers familiar with the instrument and stakeholders from the populations being studied. This adaptation methodology aims to maintain the content, semantic, and technical equivalence. It is superior to the translation and back-translation method, which may not yield cultural equivalency of the content being translated, nor capture key sociocultural factors [30]. In this study, two groups of bilingual stakeholders reviewed the SMS text message content being translated (1) a bilingual research team and (2) focus group participants. Figure 1 summarizes the 10-step adaptation process used in this study (the Matías-Carrelo et al [30] framework is shown in gray and the corresponding adaptations in white).

XSL•FO
RenderX

Figure 1. Adaptation and development of REACH-Español text message content on barriers to diabetes medication adherence (gray boxes outline original adaptation framework proposed by Matias-Carrelo et al [30]; corresponding modifications are shown in the white boxes). REACH: Rapid Encouragement/Education And Communications for Health.

XSL•FO
RenderX



XSL•FO
RenderX

Literature Review of Barriers and Translation

We conducted a systematic review of published studies describing patient-reported barriers to diabetes medication adherence in Spanish-speaking adults with T2D. We identified an additional 9 barriers relevant to Spanish-speaking adults that were not included in the original REACH barrier list (Multimedia Appendix 1). We collapsed similar barriers, resulting in 40 barriers to diabetes medication adherence with potential relevance to Spanish-speaking adults with T2D (Multimedia Appendix 2).

Barrier Translation

The 40 barriers' names were professionally translated into Spanish (Multimedia Appendix 1).

Barrier Assessment

To assess the relevance (and irrelevance) of the 40 barriers to diabetes medication adherence among Latino adults with T2D, we conducted 3 focus groups. Our screening for health literacy at recruitment showed that 33% of participants interested in attending one of the focus groups had low health literacy ("not at all" confident in filling out medical forms by yourself). Therefore, we tailored the focus group activities to a lower health literacy level by focusing on nonwritten activities and using visual props and cues to guide the discussion (Multimedia Appendix 3). We used a 2-step interactive process adapted from the original REACH barrier development process [25]. In the first activity, we encouraged an open discussion by asking participants to think about "things that make taking diabetes medications difficult" and to share these with the group; we wrote each item on a whiteboard as the discussion progressed. In the second activity, we asked participants to rank each of the 40 barriers to diabetes medication adherence referenced earlier, according to personal relevance. To make the activity easy to follow, we used a "traffic light" model of 3 colored envelopes and asked participants to place each written barrier name (eg, "I worry that taking diabetes medicines for a long time will be bad for me") into a folder, based on its relevance to them (Multimedia Appendix 3): red (never), yellow (sometimes), or green (always). Focus groups were audio-recorded and professionally transcribed in Spanish.

Barrier Analysis and Selection

Upon completion of the barrier ranking activity, barriers that were placed in the red "never" envelope received a score of 1, those placed in the yellow "sometimes" folder received a 2, and barriers placed in the green "always" folder received a 3. We then calculated the proportion of participants that reported each barrier as sometimes or always (Multimedia Appendix 4). We retained 23 barriers that were reported as relevant ≥33% of participants (mean proportion for all barriers was 32%) plus 1 barrier (regimen complexity) that emerged as an important theme in the focus groups. We then combined barriers with similar content (19 barriers) and removed 1 barrier ("I think brand name medicine works better than generic medicine") because of some recent recalls of generic glucose-lowering medications which may have affected perceptions at the time [31], resulting in 18 final barriers for REACH-Español.

Review of Existing Barrier Text Message Content

To adapt the text message content corresponding to each barrier in a Spanish-speaking population, we selected the most relevant barriers that were identified in focus groups 1 - 3. We reviewed the original REACH text message English content corresponding to each of the selected barriers and conducted a readability check of each text message using the Flesch-Kincaid readability grade level test, ensuring that the language used was grade 6 or lower.

Text Message Content Translation

The text message content corresponding to each of the 18 barriers was professionally translated into Spanish.

Text Message Spanish Language Assessment

To ensure that the Spanish text message content was culturally and linguistically appropriate for the target population, we invited the same group of participants that attended focus groups 1 - 3 to attend 1 of 2 additional focus groups (focus groups 4 - 5). To solicit participant feedback in the focus groups, we projected each text message on a Microsoft PowerPoint presentation and gave each participant a green ("Yes) and a red ("No") paddle (Multimedia Appendix 5). We asked participants to approve or disapprove of messaging according to ease of understanding by raising the paddle when each text message was projected and read aloud. We discussed messages that received "No" votes.

Text Message Content Modification

Following the completion of focus groups 4 - 5, we excluded text messages not well-received by participants (not relevant, tone did not resonate). Then, our bilingual research team modified specific language based on participant feedback, with a focus on selecting language that was more broadly understood by Spanish-speaking participants from diverse geographic and cultural backgrounds with distinct dialects. Focus groups were audio-recorded and professionally transcribed in Spanish.

Beta Testing

The revised text message content was sent to our digital health partner, MEMOTEXT, which will deliver the text messages to participants in the pilot study of REACH-Español. To ensure that the text message content delivery worked as intended and to subject the content to one more round of language modification, we conducted beta testing of the text message content. Beta testing entailed having the research team members as well as 3 patient stakeholders (participants who attended focus groups 4 or 5) receive and interact with the REACH-Español text message content over the course of 4 weeks. All members of the beta testing process provided input on a weekly basis on either technical or linguistic aspects that required modification. Patient stakeholders provided their impression of the program via an exit interview.

Fine-Tuning

Throughout the beta testing process, we identified several technical issues, mostly pertaining to the sent time of the text messages (ie, too close to bedtime, welcome text message sent after content message). MEMOTEXT revised the platform throughout the beta testing process based on user feedback.

XSL•FO
RenderX

After all user input was incorporated, a final version of the REACH-Español text message content was created.

Analysis

We used descriptive statistics (means, SDs, and proportions) to summarize participant characteristics. For the qualitative analysis, an "open-coding" approach (inductive) was used first to identify themes in the transcripts (in Spanish) from focus groups 1 - 3. Themes were then categorized as barriers and facilitators to diabetes medication adherence. Second, we used deductive coding to map the codes into the IMB construct. Coding was conducted by native Spanish-speaking research team members. Analyses were conducted using Microsoft Excel and NVivo 14 (Lumivero).

Results

Focus Group Participants

A total of 22 participants attended focus groups 1 - 5. The first 3 groups were attended by 18 participants; 10 of these participants also attended either focus group 4 or 5. Table 1 summarizes the demographic and clinical characteristics of the study participants. The mean age was 63.2 (SD 11) years, 55% (10/22) were female, 59% (13/22) were categorized as having high health literacy (based on the 1-item BRIEF questionnaire), although 55% (12/22) reported an educational attainment of elementary or middle school only. Over two-thirds (16/22 73%) of participants were born outside of the United States, largely in Central America. Participants had a mean (SD) diabetes duration of 20.3 (12) years. In addition, 59% (13/22) of participants reported insulin use.

Table . Characteristics of the focus group population.

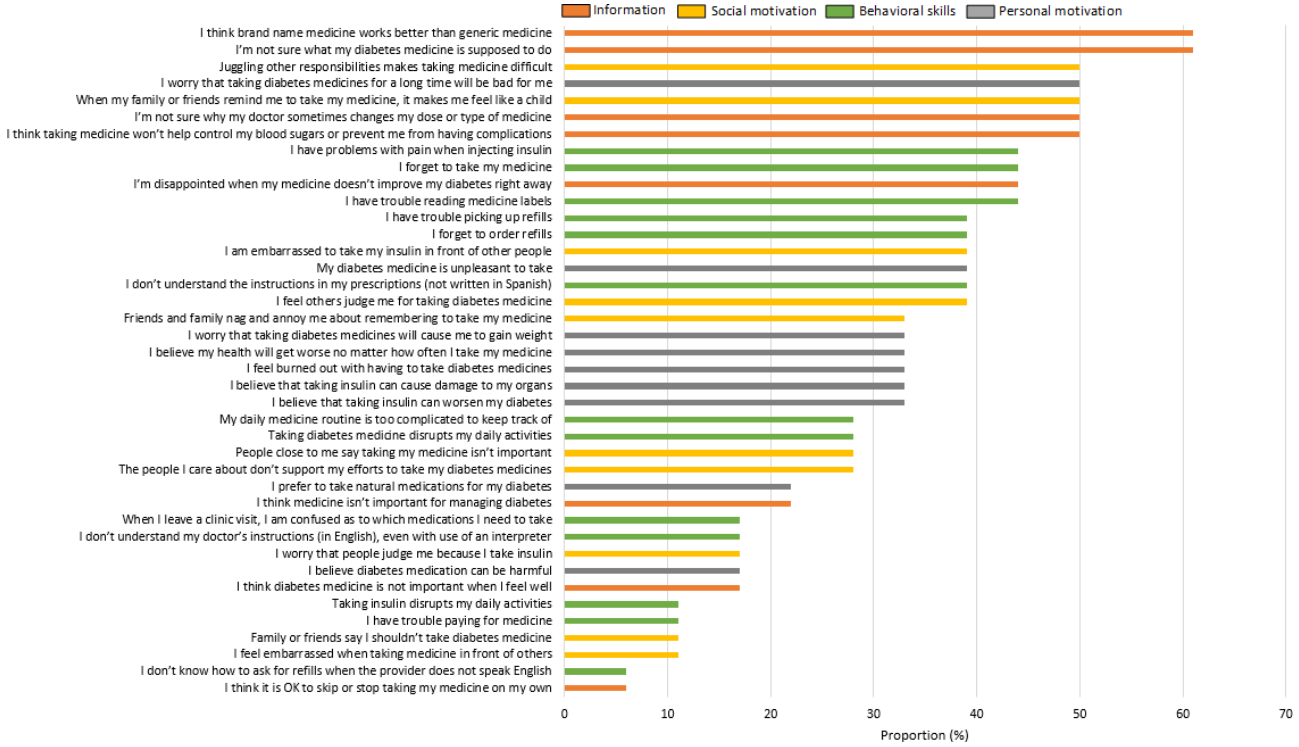
Characteristic	Overall (n=22)
Age (y), mean (SD)	63.2 (11)
Sex (female), n (%)	10 (55)
HbA_{1c}, mean (SD)	8.5 (1)
HbA_{1c}<7%, n (%)	2 (9)
HbA_{1c}>7%, n (%)	20 (91)
Age at diabetes diagnosis (y), mean (SD)	44.2 (13)
Diabetes duration (y), mean (SD)	20.3 (12)
Number of diabetes medications, mean (SD)	2.9 (1)
Insulin use, n (%)	13 (59)
Health literacy quintile, n (%)	
1 (lowest)	3 (14)
2	1 (4)
3	2 (9)
4	3 (14)
5 (highest)	13 (59)
Educational attainment, n (%)	
Elementary or middle school	12 (55)
Some secondary schooling	4 (18)
High school or GED	4 (18)
Some university	2 (9)
Place of birth, n (%)	
Colombia	1 (5)
Dominican Republic	1 (5)
El Salvador	5 (23)
Guatemala	5 (23)
Honduras	3 (14)
Peru	1 (5)
Puerto Rico	6 (27)

XSL•FO
RenderX

Barrier Selection

Multimedia Appendix 2 shows the list of 40 barriers to diabetes medication adherence that were assessed for relevance in focus groups 1‐3. Barriers are shown in order of higher to lower proportion of participants reporting each barrier as relevant sometimes or always. Figure 2 shows the list of 40 barriers, categorized into the IMB model and ranked according to the proportion of participants who reported the barrier as relevant sometimes or always. Among the top 10 barriers, 50% (n=5) corresponded to the information category, 20% (n=2) to the social motivation category, 20% (n=2) to behavioral skills, and 10% (n=1) to personal motivation.

Figure 2. Proportion of participants who reported each diabetes medication adherence barrier as relevant sometimes or always (barriers were shortened to improve readability, please refer to Multimedia Appendix 4 for full barrier wording).



Barriers to Diabetes Medication Adherence Based on Thematic Analysis

Table 2 outlines the key themes that emerged in focus groups 1‐3 based on participant quotes, grouped into barriers to diabetes medication adherence, and mapped into the IMB model. In total, 5 themes emerged in response to the question "What are some difficulties that come to mind when taking your diabetes medications?" In order of relevance (higher or lower number of times mentioned), most participants agreed that medication burden was one of the most important barriers to diabetes medication adherence, followed by fear of side effects or of new medication, difficulty or confusion with medication changes, forgetting to take medications, and frustration with diagnosis or lack or improvement.

XSL•FO
RenderX

Table . Diabetes medication adherence barriers reported by focus group participants (what are some difficulties that you encounter when taking your diabetes medications?).

Barriers to diabetes medication adherence: What are some difficulties that you encounter when taking your diabetes medications?			
Themes	Participant Quotes	IMB[a] category	Existing REACH[b] barrier
Medication burden			
Too many medications and exhaustion with taking medications daily	"It's not difficult to take medications, it's just that we are disgusted by having to take so many medications." (Female, 62 y old, El Salvador) "I am exhausted with having to take medications every day." (Female, 45 y old, Guatemala) "Sometimes I don't want it….there is a rejection, a rejection [of the medication]." (Female, 45 y old, El Salvador)	Personal Motivation	I feel burned out with having to take diabetes medicines
Interferes with daily activities	"Yes, [it interferes] with daily chores…if the next dose is close enough, then I skip it and inject the following dose."(Male, 63 y old, El Salvador). "Well, sometimes it's very difficult to wake up, knowing that you won't be drinking a cup of coffee with bread, but something else, to take the diabetes pills." (Female, 74 y old, Puerto Rico)	Social Motivation	Taking diabetes medicine disrupts my daily activities/ Taking insulin disrupts my daily activities
Regimen complexity	"It's very difficult because I take ten medications in the morning and eight at night." (Male, 63 y old, El Salvador). "In my case, I have to take two pills in the morning and two at night…and the daytime ones, I can't get used to taking them. I take all four at night." (Male, 63 y old, El Salvador)	Behavioral skills	My daily medicine routine is too complicated to keep track of
Juggling other responsibilities	"I have difficulty taking the medications because I start working at 7AM, so if I am late for work I don't take any medications until my break time, around 10am…and then the other medications I take at 9PM, when I come back from work." (Female, 45 y old, Guatemala)	Social Motivation	Juggling other responsibilities makes medicine difficult
Fear of side effects or of new medications			
Low blood sugars	"And there are times when it [blood sugar] drops suddenly and it gives me anxiety to find something to eat because I feel like as he [other participant] mentioned, tremulous, it's horrible." (Male, 56 y old, Guatemala)	Personal Motivation	N/A[c]
Side effects	"Perhaps I will not feel it now but I will feel it [side effects] over time." (Male, 65 y old, Guatemala)	Personal Motivation	I worry that taking diabetes medicines for a long time will be bad for me
New medications	"You are taking a risk, if you take the medication, it alleviates something but it will damage something else." (Female, 63 y old, Puerto Rico)	Information	N/A

XSL•FO
RenderX

Barriers to diabetes medication adherence: What are some difficulties that you encounter when taking your diabetes medications?			
Themes	Participant Quotes	IMB[a] category	Existing REACH[b] barrier
Organ damage	"Because you know that there are medications that damage the kidneys and then the doctor tells you, I'm sorry, it was the medication [that caused the kidney damage], so you live in fear." (Female, 71 y old, Puerto Rico)	Personal Motivation	I believe diabetes medication can be harmful.
Difficulty or confusion with medication changes			
Medication changes	"I had a problem with the insulin because, I mean, my sugar was very high so the doctor said 'we are going to do an experiment,' they did an experiment with me that they put me until a certain level and that's where it [the dose] remained." (Male, 65 y old, Guatemala) "Because they have me like a guinea pig, like playing with me, making experiments…sincerely don't know what to do." (Male, 56 y old, Guatemala)	Information	I'm not sure why my doctor sometimes changes my dose or type of medicine
Medication instructions	"Í would leave [the clinic] and would write down the [medication] changes but then I would lose the paper…so I wouldn't know if I had to put 35 [units] in the morning or 38 in the afternoon…and since I had a visit every 3 mo, I would suffer very much." (Male, 77 y old, Guatemala)	Information	When I leave a clinic visit, I am confused as to which medications I need to take
Forget to take medications			
	"In reality, it's not difficult to take the medications…the problem is that I forget to take the medications because of my age." (Male, 77 y old, Guatemala) "Yes, I have one difficulty in taking my medications. That is, I am very forgetful." (Male, 54 y old, El Salvador)	Behavioral skills	I forget to take my medicine
Frustration with diagnosis or lack of improvement			
Treatment doesn't work	"At times I feel very frustrated because so many medications that you are taking and the sugar does not go down….that puts me depressed." (Female, 62 y old, El Salvador)	Information	I'm disappointed when my medicine doesn't improve my diabetes right away
Difficulty accepting diagnosis	"I have diabetes since the last two years. I come from a family with diabetes. I got divorced and in the process, the diabetes appeared. So, I started taking the metformin because I didn't want to accept it. It was something very hard and I have difficulty accepting it." (Female, 45 y old, Guatemala)	Personal motivation	N/A

XSL•FO
RenderX

Barriers to diabetes medication adherence: What are some difficulties that you encounter when taking your diabetes medications?			
Themes	Participant Quotes	IMB[a] category	Existing REACH[b] barrier
Labile blood sugars	"The sugar went down but now I have a problem that the sugar went up, so there's no balance, how to control this, sincerely it is very, very complicated for me and I don't know what to do." (Male, 56 y old, Guatemala)	Information	N/A

[a]IMB: information-motivation-behavioral skills model.

[b]REACH: Rapid Encouragement/Education And Communications for Health.

[c]N/A:not applicable

Facilitators to Diabetes Medication Adherence Based on Thematic Analysis

Multimedia Appendix 6 shows the 5 themes that emerged in response to the question, "What motivates you to take your diabetes medications." In order of relevance (higher to lower), fear as motivation was an important facilitator, followed by belief in treatment benefit, personal responsibility, and will to take medications, family support and medication organization, and getting used to the medication.

Language Modification of Text Message Content

Focus groups 4 - 5, which centered on eliciting feedback on the translated text message language content, generated rich discussion. A total of 111 text messages (related to barriers 1 - 8) were reviewed in focus group 4 and 100 text messages (related to barriers 9 - 18) were reviewed in focus group 5. Participants were generally enthusiastic about the text message content and provided feedback on wording that seemed confusing, too complex, or less likely to be understood by broad Spanish-speaking audiences. For instance, the word "comprimidos" was changed to "pastillas" and the preferred word for refills was "reposiciones" instead of "prescripciones." In total, 12 messages were excluded since patients did not like the tone, felt that they were difficult to understand, or felt that they were not relevant. Some examples of these excluded text messages included "If taking your diabetes medications is unpleasant, try to push through. Remember that not taking them may make you feel worse" and "What advice would you give your friends who forget to refill their medications? Use those tips for yourself too!"

Beta Testing Patient Stakeholder Input

In the final step of REACH-Español content adaptation, 3 patient participants who had attended one of the focus groups were invited to participate in the beta testing process over the course of 4 weeks. At the end of beta testing, the following patient stakeholder input was provided:

> *The texts are very important because when you are old, you know things from experience, but you forget things that are important. Like that the test strips need to be in a hermetically sealed box. The text messages are very helpful with remembering to take the medications. [Male, 77 years old, Guatemala]*

> *The texts are helping to remind me to take my medications. When I look and I see the message [any of the messages], I run to take my medication. The texts are very clear. It's really helping me. [Female, 45 years old, Guatemala]*

> *The content is helpful. For example, learning about what food to eat to raise a low blood sugar. It's helping to remind me to take my medications. I think this program is best for older people who have the time and attention to follow the advice. [Male, 63 years old, El Salvador]*

Discussion

Principal Findings and Comparison With Previous Works

In this study, we outline the multistage process used to adapt the REACH text message content [26], which addresses patient-reported barriers to diabetes medication adherence, to a Latino population. The adaptation process centered on the theory-based IMB skills model for medication adherence and engagement of patient stakeholders to identify barriers to diabetes medication adherence with specific relevance to Latino adults with T2D (with Spanish-speaking preference). We found that among the top 10 most commonly reported barriers to diabetes medication adherence among Latino adults with diabetes, 50% (n=5) corresponded to information gaps, 20% (n=2) to social motivation, 20% (n=2) to behavioral skills, and 10% (n=1) to personal motivation. In contrast, our thematic analysis showed that personal motivation barriers (medication burden and fear of side effects) were the highest concern for focus group participants, followed by behavioral skills (forgetting to take medications), and information barriers (difficulty or confusion with medication changes and frustration with diagnosis or lack of improvement). These findings suggest that using the IMB model as a framework to identify barriers to diabetes medication adherence is useful in this population and that different qualitative approaches (open discussion vs ranking exercise) can provide complementary data to inform the adaptation process of tailored interventions.

By engaging focus group participants in an interactive barrier ranking exercise conducted through card sorting, we were also able to select highly relevant barriers to diabetes medication adherence and exclude those with lower salience from the study.

XSL•FO
RenderX

Through this exercise, we found that the priority of some barriers differed in this population compared with the original REACH intervention. For instance, "I'm not sure what my diabetes medicine is supposed to do" was reported as relevant by 61% (11/18) of participants in this study, whereas this barrier was relevant only for 16% (38/237) of participants in the REACH development study [25]. Similarly, "I am not sure why my doctor sometimes changes my dose or type of medicine" was relevant to 50% (9/18) of participants in this study, whereas only 15% (36/237) of participants in the REACH development study reported that this barrier was relevant. These findings underscore the importance of stakeholder-driven adaptation approaches given that core intervention components (in this case, the list of barriers), differed across populations.

We also found that the iterative process of translating and modifying the original intervention content with stakeholder input (bilingual research team, focus group participants) enriched the adaptation process by incorporating language diversity and nuance. By eliciting input from participants through interactive focus group activities, we were able to capture both the thematic and linguistic or cultural relevance of the content being translated. For instance, we excluded certain text messages that did not resonate with the audience. We revised the language with modifications considered more broadly understood across Spanish-speaking Latin American and Caribbean populations. Choosing language wording with a broader reach was particularly important in this study given the heterogeneity of the Spanish language across Latin America and the Caribbean. Our observations are consistent with previous studies that have highlighted the importance of engaging bilingual stakeholders (including patients) throughout the language adaptation process [30,32,33].

Although barriers to diabetes medication adherence have been previously reported in the Latino population [8,10,11,17,19,34], most studies have not used a theoretical framework to identify such barriers. The IMB model has been empirically validated as a framework to identify barriers to diabetes medication adherence [23], but only 1 intervention based on the IMB model has documented its use in Latin American adults with diabetes [35]. Our study supports the use of the IMB framework to better ascertain barriers to diabetes medication adherence among Latino adults, which can be helpful in the development of interventions to address these barriers. Our qualitative analysis also highlights the importance of treatment complexity in this patient population and underscores the need for interventions that can identify and reduce treatment burden and regimen complexity. Although these challenges to diabetes self-management have been well described [36-39], scalable interventions that can reduce the treatment burden are lacking, particularly in the Latino population.

Digital health technology can improve diabetes medication adherence behavior [21,40] and HbA1c [21,40,41] and is recommended by the American Diabetes Association to support diabetes self-care [42]. However, among the 90 (approximately) digital health apps currently available for diabetes self-management, only one-third are available in Spanish, of which 94% (90/96) have a readability level above that recommended for patient education material [43]. Some digital medication adherence support platforms are currently under study in Latino and Spanish-speaking populations [44,45]. However, none of these platforms provide educational content tailored to patient-reported barriers to diabetes medication adherence. The REACH-Español platform adapted in this study offers this key innovation, in addition to health literacy–informed content, which could facilitate broader engagement of prospective participants and long-term medication adherence.

Our study has the key strength of having adapted an intervention with iterative input from stakeholders who could be representative of the population most likely to use the intervention. This user-centered approach is recommended in diabetes digital health research [46,47] to ensure that end user perspectives are represented in the development of interventions and content is tailored to user needs. However, our study also has several limitations. First, the sample size was relatively small, although we achieved thematic saturation for barrier selection after conducting 3 focus groups. Second, we recruited participants from a single health center and did not have representation from every Latin American and Caribbean country, which could limit the generalizability of our findings. In addition, although the barriers to diabetes medication adherence selected in this study were based on relevance to the focus group participants, the order of relevance may differ in other Latin American populations living elsewhere in the United States or with a different insulin usage profile. For instance, "I have trouble paying for my medicine" was not a highly ranked barrier but this may be an important barrier in other areas of the country with lower insurance coverage. Thus, should the REACH-Español content be adapted to other contexts, using the IMB model to identify population-specific barriers could be helpful. Finally, the focus of this study was to adapt the barrier-specific text message content. While text message content pertaining to diet and physical activity was directly translated from English to Spanish, the content did not undergo the stakeholder-driven adaptation process described here, which will be the subject of a future study.

Conclusions

In this study, we report the multiphase, stakeholder-driven adaptation of the REACH digital health platform text message content, which addresses barriers to diabetes medication adherence, to a Latino population with T2D. We found key barriers to diabetes medication adherence that were relevant to this population, largely related to information gaps. These findings can contribute to the fields of patient-important outcomes and treatment burden to guide respectful and meaningful patient engagement in the management of T2D. Furthermore, engaging a diverse group of bilingual stakeholders (bilingual research team, focus group participants) early and throughout the iterative adaptation process was integral to ensuring content and linguistic relevance of the text messages in the REACH-Español platform. The feasibility, acceptability, usability, and preliminary efficacy of REACH-Español is under study in a 6-month pilot randomized controlled trial, which could broaden our understanding of the role of tailored, digital health interventions on improving diabetes self-care among Latino adults with T2D.

XSL•FO
RenderX

Acknowledgments

The authors thank the MGH Chelsea HealthCare Center and the MGH Diabetes Research Center, particularly Mary Larkin, Melissa Calverley, and Arantxa Medina for their contributions to this study and for their assistance in organizing the focus groups.

JS was supported by grant K23DK135798 from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) and by the MGH Physician-Scientist Development Award. SC was supported by American Diabetes Association grant # 7 - 21-JDFM-005.

Authors' Contributions

JS, DW, and LM co-conceived the study. JS, ST, and RCT conducted the focus groups and collected the data. JS conducted the analysis. JS, DW, and LM wrote the first draft of the manuscript, and all authors provided critical inputs on multiple iterations. All authors have approved the final version.

Conflicts of Interest

DW reports serving on a data monitoring committee for Novo Nordisk. SC's spouse works for Depuy-Synthes; ad hoc consulting for Alexion Pharmaceuticals, Patients Square Capital, and Wolters-Kluwer. All other authors declare no competing interests.

Multimedia Appendix 1
List of REACH barriers and additional barriers from literature search of potential relevance to Spanish-speaking adults with type 2 diabetes (table shows the barriers to diabetes medication adherence used in the original REACH intervention, mapped into the Information Motivation Behavioral skills model; additional barriers with described relevance to Spanish-speaking adults with type 2 diabetes are shown in gray).
[DOCX File, 14 KB - humanfactors_v12i1e66668_app1.docx]

Multimedia Appendix 2
Condensed list of barriers used in focus groups 1-3 (table shows the barriers to diabetes medication adherence used in the adaptation of REACH to REACH-Español. The 9 barriers that were identified in the literature with potential relevance for Latino adults with type 2 diabetes were added to the list and combined with similar REACH barriers [gray]).
[DOCX File, 16 KB - humanfactors_v12i1e66668_app2.docx]

Multimedia Appendix 3
Barrier ranking exercise. Top: "Traffic light" visual tool used to illustrate the ranking categories and corresponding folders used in the ranking exercise. Bottom: Barriers to diabetes medication adherence cards used in the ranking exercise.
[DOCX File, 773 KB - humanfactors_v12i1e66668_app3.docx]

Multimedia Appendix 4
Barriers to diabetes medication adherence assessed in focus groups conducted with Latino adults with type 2 diabetes and corresponding proportion of participants who reported that the barrier was relevant sometimes or always.
[DOCX File, 16 KB - humanfactors_v12i1e66668_app4.docx]

Multimedia Appendix 5
Focus group activity to elicit participant input regarding text message content and language. Yes (Si) and No banners were used by participants to prompt discussion on text message content and language.
[DOCX File, 205 KB - humanfactors_v12i1e66668_app5.docx]

Multimedia Appendix 6
Facilitators to diabetes medication adherence based on thematic content analysis.
[DOCX File, 13 KB - humanfactors_v12i1e66668_app6.docx]

References

1. Egede LE, Gebregziabher M, Echols C, Lynch CP. Longitudinal effects of medication nonadherence on glycemic control. Ann Pharmacother 2014 May;48(5):562-570. [doi: 10.1177/1060028014526362] [Medline: 24586059]
2. Osborn CY, Gonzalez JS. Measuring insulin adherence among adults with type 2 diabetes. J Behav Med 2016 Aug;39(4):633-641. [doi: 10.1007/s10865-016-9741-y] [Medline: 27062271]

XSL•FO
RenderX

3. Piragine E, Petri D, Martelli A, Calderone V, Lucenteforte E. Adherence to oral antidiabetic drugs in patients with type 2 diabetes: systematic review and meta-analysis. J Clin Med 2023 Mar 2;12(5):1981. [doi: 10.3390/jcm12051981] [Medline: 36902770]
4. Aiken LH, Sloane DM, Bruyneel L, et al. Nurse staffing and education and hospital mortality in nine European countries: a retrospective observational study. Lancet 2014 May 24;383(9931):1824-1830. [doi: 10.1016/S0140-6736(13)62631-8] [Medline: 24581683]
5. Ho PM, Rumsfeld JS, Masoudi FA, et al. Effect of medication nonadherence on hospitalization and mortality among patients with diabetes mellitus. Arch Intern Med 2006 Sep 25;166(17):1836-1841. [doi: 10.1001/archinte.166.17.1836] [Medline: 17000939]
6. Osborn CY, Cavanaugh K, Wallston KA, et al. Health literacy explains racial disparities in diabetes medication adherence. J Health Commun 2011;16 Suppl 3(Suppl 3):268-278. [doi: 10.1080/10810730.2011.604388] [Medline: 21951257]
7. Yang Y, Thumula V, Pace PF, Banahan BF, Wilkin NE, Lobb WB. Predictors of medication nonadherence among patients with diabetes in Medicare Part D programs: a retrospective cohort study. Clin Ther 2009 Oct;31(10):2178-2188. [doi: 10.1016/j.clinthera.2009.10.002] [Medline: 19922889]
8. Fernández A, Quan J, Moffet H, Parker MM, Schillinger D, Karter AJ. Adherence to newly prescribed diabetes medications among insured Latino and White patients with diabetes. JAMA Intern Med 2017 Mar 1;177(3):371-379. [doi: 10.1001/jamainternmed.2016.8653] [Medline: 28114642]
9. Baghikar S, Benitez A, Fernandez Piñeros P, Gao Y, Baig AA. Factors impacting adherence to diabetes medication among urban, low income Mexican-Americans with diabetes. J Immigr Minor Health 2019 Dec;21(6):1334-1341. [doi: 10.1007/s10903-019-00867-9] [Medline: 30798408]
10. Bailey GR, Barner JC, Weems JK, et al. Assessing barriers to medication adherence in underserved patients with diabetes in Texas. Diabetes Educ 2012;38(2):271-279. [doi: 10.1177/0145721711436134] [Medline: 22316643]
11. Banuelos Mota A, Feliz Sala EE, Perdomo JM, et al. Assessing barriers to medication adherence among Latinos with diabetes: a cross-sectional study. J GEN INTERN MED 2020 Feb;35(2):603-605. [doi: 10.1007/s11606-019-05041-3] [Medline: 31161564]
12. Billimek J, August KJ. Costs and beliefs: understanding individual- and neighborhood-level correlates of medication nonadherence among Mexican Americans with type 2 diabetes. Health Psychol 2014 Dec;33(12):1602-1605. [doi: 10.1037/hea0000020] [Medline: 24295022]
13. Colby JA, Wang F, Chhabra J, Pérez-Escamilla R. Predictors of medication adherence in an urban Latino community with healthcare disparities. J Immigr Minor Health 2012 Aug;14(4):589-595. [doi: 10.1007/s10903-011-9545-7] [Medline: 22069023]
14. Hu J, Amirehsani K, Wallace DC, Letvak S. Perceptions of barriers in managing diabetes: perspectives of Hispanic immigrant patients and family members. Diabetes Educ 2013;39(4):494-503. [doi: 10.1177/0145721713486200] [Medline: 23640301]
15. Morkos M, Tahsin B, Fogelfeld L. Factors associated with diabetes control in predominately African American and Hispanic population with newly diagnosed type 2 diabetes. J Racial Ethn Health Disparities 2021 Apr;8(2):332-338. [doi: 10.1007/s40615-020-00785-9] [Medline: 32557276]
16. Andreina Millan RR. Myths and misconceptions about insulin therapy among Latinos/ Hispanics with diabetes: a fresh look at an old problem. J Diabetes Metab 2015;06(1):1-5. [doi: 10.4172/2155-6156.1000482]
17. Rustveld LO, Pavlik VN, Jibaja-Weiss ML, Kline KN, Gossey JT, Volk RJ. Adherence to diabetes self-care behaviors in English- and Spanish-speaking Hispanic men. Patient Prefer Adherence 2009 Nov 3;3:123-130. [doi: 10.2147/ppa.s5383] [Medline: 19936154]
18. Smith-Miller CA, Berry DC, Miller CT. Diabetes affects everything: type 2 diabetes self-management among Spanish-speaking hispanic immigrants. Res Nurs Health 2017 Dec;40(6):541-554. [doi: 10.1002/nur.21817] [Medline: 28877552]
19. Wilson E, Chen AHM, Grumbach K, Wang F, Fernandez A. Effects of limited English proficiency and physician language on health care comprehension. J Gen Intern Med 2005 Sep;20(9):800-806. [doi: 10.1111/j.1525-1497.2005.0174.x] [Medline: 16117746]
20. Mayberry LS, Mulvaney SA, Johnson KB, Osborn CY. The messaging for diabetes intervention reduced barriers to medication adherence among low-income, diverse adults with type 2. J Diabetes Sci Technol 2017 Jan;11(1):92-99. [doi: 10.1177/1932296816668374] [Medline: 27595710]
21. Nelson LA, Greevy RA, Spieker A, et al. Effects of a tailored text messaging intervention among diverse adults with type 2 diabetes: evidence from the 15-month REACH randomized controlled trial. Diabetes Care 2021 Jan;44(1):26-34. [doi: 10.2337/dc20-0961] [Medline: 33154039]
22. Holmes EAF, Hughes DA, Morrison VL. Predicting adherence to medications using health psychology theories: a systematic review of 20 years of empirical research. Value Health 2014 Dec;17(8):863-876. [doi: 10.1016/j.jval.2014.08.2671] [Medline: 25498782]
23. Mayberry LS, Osborn CY. Empirical validation of the information-motivation-behavioral skills model of diabetes medication adherence: a framework for intervention. Diabetes Care 2014;37(5):1246-1253. [doi: 10.2337/dc13-1828] [Medline: 24598245]

XSL•FO
RenderX

24. Osborn CY, Egede LE. Validation of an information-motivation-behavioral skills model of diabetes self-care (IMB-DSC). Patient Educ Couns 2010 Apr;79(1):49-54. [doi: 10.1016/j.pec.2009.07.016] [Medline: 19699601]
25. Nelson LA, Wallston KA, Kripalani S, LeStourgeon LM, Williamson SE, Mayberry LS. Assessing barriers to diabetes medication adherence using the Information-Motivation-Behavioral skills model. Diabetes Res Clin Pract 2018 Aug;142:374-384. [doi: 10.1016/j.diabres.2018.05.046] [Medline: 29879495]
26. Nelson LA, Mayberry LS, Wallston K, Kripalani S, Bergner EM, Osborn CY. Development and usability of REACH: a tailored theory-based text messaging intervention for disadvantaged adults with type 2 diabetes. JMIR Hum Factors 2016 Sep 8;3(2):e23. [doi: 10.2196/humanfactors.6029] [Medline: 27609738]
27. Nelson LA, Spieker A, Greevy R, LeStourgeon LM, Wallston KA, Mayberry LS. User engagement among diverse adults in a 12-month text message-delivered diabetes support intervention: results from a randomized controlled trial. JMIR Mhealth Uhealth 2020 Jul 21;8(7):e17534. [doi: 10.2196/17534] [Medline: 32706738]
28. U.S. Census Bureau. QuickFacts: chelsea city, massachusetts. 2024. URL: https://www.census.gov/quickfacts/fact/table/chelseacitymassachusetts/PST045223 [accessed 2025-04-21]
29. Singh R, Coyne LS, Wallace LS. Brief screening items to identify spanish-speaking adults with limited health literacy and numeracy skills. BMC Health Serv Res 2015 Sep 14;15:374. [doi: 10.1186/s12913-015-1046-2] [Medline: 26370120]
30. Matías-Carrelo LE, Chávez LM, Negrón G, Canino G, Aguilar-Gaxiola S, Hoppe S. The Spanish translation and cultural adaptation of five mental health outcome measures. Cult Med Psychiatry 2003 Sep;27(3):291-313. [doi: 10.1023/a:1025399115023] [Medline: 14510096]
31. Teasdale B, Hudspeth A, Kucera K, et al. Safety vs price in the generic drug market: metformin. Am J Manag Care 2024 Apr;30(4):161-168. [doi: 10.37765/ajmc.2024.89450] [Medline: 38603530]
32. Cortés DE, Gerena M, Canino G, et al. Translation and cultural adaptation of a mental health outcome measure: the BASIS-R(c). Cult Med Psychiatry 2007 Mar;31(1):25-49. [doi: 10.1007/s11013-006-9043-x] [Medline: 17219078]
33. Ramos Z, Alegría M. Cultural adaptation and health literacy refinement of a brief depression intervention for Latinos in a low-resource setting. Cultur Divers Ethnic Minor Psychol 2014 Apr;20(2):293-301. [doi: 10.1037/a0035021] [Medline: 24588455]
34. Hofer R, Choi H, Mase R, Fagerlin A, Spencer M, Heisler M. Mediators and moderators of improvement in medication adherence: secondary analysis of a community health worker-led diabetes medication self-management support program. Health Educ Behav Off Publ Soc Public Health Educ 2017:285-296. [doi: 10.1177/1090198116656331]
35. Osborn CY, Amico KR, Cruz N, et al. A brief culturally tailored intervention for Puerto Ricans with type 2 diabetes. Health Educ Behav 2010 Dec;37(6):849-862. [doi: 10.1177/1090198110366004] [Medline: 21076128]
36. Eton DT, Ridgeway JL, Egginton JS, et al. Finalizing a measurement framework for the burden of treatment in complex patients with chronic conditions. Patient Relat Outcome Meas 2015;6:117-126. [doi: 10.2147/PROM.S78955] [Medline: 25848328]
37. Haider S, El Kawkgi O, Clark J, et al. Beyond hemoglobin A1c: a videographic analysis of conversations about quality of life and treatment burden during clinical encounters for diabetes care. Endocrine 2021 Sep;73(3):573-579. [doi: 10.1007/s12020-021-02757-3] [Medline: 34189680]
38. Sorkin DH, Billimek J, August KJ, et al. Mental health symptoms and patient-reported diabetes symptom burden: implications for medication regimen changes. Fam Pract 2015 Jun;32(3):317-322. [doi: 10.1093/fampra/cmv014] [Medline: 25846216]
39. Spencer-Bonilla G, Serrano V, Gao C, et al. Patient work and treatment burden in type 2 diabetes: a mixed-methods study. Mayo Clin Proc Innov Qual Outcomes 2021 Apr;5(2):359-367. [doi: 10.1016/j.mayocpiqo.2021.01.006] [Medline: 33997635]
40. Shrivastava TP, Goswami S, Gupta R, Goyal RK. Mobile app interventions to improve medication adherence among type 2 diabetes mellitus patients: a systematic review of clinical trials. J Diabetes Sci Technol 2023 Mar;17(2):458-466. [doi: 10.1177/19322968211060060] [Medline: 34861793]
41. Chamany S, Walker EA, Schechter CB, et al. Telephone intervention to improve diabetes control: a randomized trial in the New York City A1c Registry. Am J Prev Med 2015 Dec;49(6):832-841. [doi: 10.1016/j.amepre.2015.04.016] [Medline: 26232903]
42. American Diabetes Association Professional Practice Committee. 7. Diabetes technology: standards of care in diabetes-2024. Diabetes Care 2024 Jan 1;47(Suppl 1):S126-S144. [doi: 10.2337/dc24-S007] [Medline: 38078575]
43. Rodriguez JA, Singh K. The Spanish availability and readability of diabetes apps. J Diabetes Sci Technol 2018 May;12(3):719-724. [doi: 10.1177/1932296817749610] [Medline: 29291639]
44. Fortmann AL, Savin KL, Clark TL, Philis-Tsimikas A, Gallo LC. Innovative diabetes interventions in the U.S. Hispanic population. Diabetes Spectr 2019 Nov;32(4):295-301. [doi: 10.2337/ds19-0006] [Medline: 31798285]
45. Philis-Tsimikas A, Fortmann AL, Godino JG, et al. Dulce Digital-Me: protocol for a randomized controlled trial of an adaptive mHealth intervention for underserved Hispanics with diabetes. Trials 2022 Jan 28;23(1):80. [doi: 10.1186/s13063-021-05899-x] [Medline: 35090520]
46. Martinez W, Threatt AL, Rosenbloom ST, Wallston KA, Hickson GB, Elasy TA. A patient-facing diabetes dashboard embedded in a patient web portal: design sprint and usability testing. JMIR Hum Factors 2018 Sep 24;5(3):e26. [doi: 10.2196/humanfactors.9569] [Medline: 30249579]

XSL•FO
RenderX

47. Nelson LA, Threatt AL, William Martinez SWA, Mayberry LS. Chapter 4 - agile science: what and how in digital diabetes research. In: Klonoff DC, Kerr D, Mulvaney SA, editors. Diabetes Digital Health: Elsevier; 2020:51-63. [doi: 10.1016/B978-0-12-817485-2.00004-3]

Abbreviations

BRIEF: Brief Health Literacy Screening Tool
HBA_{1c}: glycated hemoglobin
IMB: information-motivation-behavioral
MGH: Massachusetts General Hospital
REACH: Rapid Encouragement/Education And Communications for Health
T2D: type 2 diabetes

Edited by J Waterson; submitted 04.10.24; peer-reviewed by AN Ali, S Ranjbaran; revised version received 06.03.25; accepted 12.03.25; published 01.05.25.

Please cite as:
Seiglie J, Tobolsky S, Crespo Trevino R, Cordova L, Cromer S, Caballero AE, Alegria M, Miranda JJ, Wexler D, Mayberry L
Adapting a Text Messaging Intervention to Improve Diabetes Medication Adherence in a Spanish-Speaking Population: Qualitative Study
JMIR Hum Factors 2025;12:e66668
URL: https://humanfactors.jmir.org/2025/1/e66668
doi:10.2196/66668

© Jacqueline Seiglie, Seth Tobolsky, Ricaurte Crespo Trevino, Lluvia Cordova, Sara Cromer, A Enrique Caballero, Margarita Alegria, J Jaime Miranda, Deborah Wexler, Lindsay Mayberry. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 1.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Perception and Evaluation of a Knowledge Transfer Concept in a Digital Health Application for Patients With Heart Failure: Mixed Methods Study

Madeleine Flaucher[1], MSc; Sabrina Berzins[1], MSc; Katharina M Jaeger[1], MSc; Michael Nissen[1], MSc; Jana Rolny[2]; Patricia Trißler[2]; Sebastian Eckl[2]; Bjoern M Eskofier[1,3], Prof Dr; Heike Leutheuser[1], Dr. Ing.

[1]Machine Learning and Data Analytics Lab, Department Artificial Intelligence in Biomedical Engineering, Friedrich-Alexander-Universität Erlangen-Nürnberg, Carl-Thiersch-Str. 2b, Erlangen, Germany

[2]ProCarement GmbH, Forchheim, Germany

[3]Institute of AI for Health, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany

Corresponding Author:
Madeleine Flaucher, MSc
Machine Learning and Data Analytics Lab, Department Artificial Intelligence in Biomedical Engineering, Friedrich-Alexander-Universität Erlangen-Nürnberg, Carl-Thiersch-Str. 2b, Erlangen, Germany

Abstract

Background: Digital health education can enhance the quality of life of patients with heart failure by providing accessible and tailored information, which is essential for effective self-care and self-management.

Objective: This work aims to develop a mobile health knowledge transfer concept for heart failure in a user-centered design process grounded in theoretical frameworks. This approach centers on enhancing the usability, patient engagement, and meaningfulness of mobile health education in the context of heart failure.

Methods: A user-centered design process was employed. First, semistructured stakeholder interviews were conducted with patients (n=9) and medical experts (n=5). The results were used to develop a health knowledge transfer concept for a mobile health app for heart failure. This concept was implemented as a digital prototype based on an existing German mobile health app for patients with heart failure. We used this prototype to evaluate our concept with patients with heart failure in a study composed of user testing and semistructured patient interviews (n=7).

Results: Stakeholder interviews identified five themes relevant to mobile health education: individualization, content relevance, media diversity, motivation strategies, and trust-building mechanisms. The evaluation of our prototype showed that patients value the adaptation of content to individual interests and prior knowledge. Digital rewards such as badges and push notifications can increase motivation and engagement but should be used with care to avoid overload, irrelevance, and repetition.

Conclusions: Our findings emphasize the importance of tailoring mobile health education to the specific needs and preferences of patients with heart failure. At the same time, they also highlight the careful implementation of motivation strategies to promote user engagement effectively. These implications offer guidance for developing more impactful interventions to improve health outcomes for this population.

(JMIR Hum Factors 2025;12:e56798) doi:10.2196/56798

KEYWORDS

health literacy; digital Literacy; user-centered design; digital health app; heart failure; mixed methods study; user centered deign; usability; patient engagement; mHealth app; development

Introduction

In 2019, heart failure affected more than 56 million people worldwide [1]. The diagnosis of heart failure entails severe health consequences for patients and significant financial resources for public health systems [2]. In light of an aging population, heart failure is becoming increasingly important in society and requires new approaches for prevention, management, and therapy to mitigate health impact [3,4].

Heart failure requires a high level of active patient involvement [5]. Self-monitoring of physiological parameters, medication intake, nutrition, and physical activity are crucial factors in preventing the hospitalizations of patients [6,7]. In this context, health literacy plays a major role since high levels of health literacy in patients with heart failure promote patient empowerment, self-care practices, and medication adherence [8-10]. Health literacy can be described as the capacity of a person to access and understand health-related information, place it in context, and use it responsibly to promote and

XSL•FO
RenderX

maintain good health [11]. With a focus on heart failure and other cardiovascular diseases, The American Heart Association concludes that health literacy is crucial for effective treatments for cardiovascular diseases and acknowledges information technology as a potential path to improving health literacy [12]. In the long-term, low health literacy is associated with an increased risk for hospitalizations and death in patients with heart failure [13].

To effectively deliver information to patients and achieve learning effects, mobile health technology offers great potential. This applies to populations with low health literacy in particular [14-16]. As smartphone access has become widespread across various demographic groups, including those with lower socioeconomic status, they provide an accessible platform for delivering tailored health interventions to underserved populations [17,18]. Enhancing health literacy requires attention not only to disease-specific content but also to digital health literacy, empowering users to navigate health-related inquiries online effectively [14,19,20]. However, several barriers have been identified, including the access to apps, the readability of content as well as the usability of digital health services [14,21]. In the context of heart failure, evidence about the effects of mobile health for promoting health literacy is still limited. Allida et al identified no increase in heart failure knowledge through digital educational content and unclear evidence on self-efficacy, self-care, and health-related quality of life [22]. In contrast, several recent studies demonstrated positive effects on quality of life, hospitalization rates, and self-management abilities [23-25]. However, despite the potential benefits, these approaches lack the incorporation of theoretical frameworks for engaging, appealing, and effective knowledge transfer. Furthermore, patients with heart failure are hardly involved in the development process of mobile health interventions.

Thus, the aim of this work was to develop a concept for a sustainable and engaging knowledge transfer with a specific focus on the needs and preferences of patients with heart failure. Theoretical frameworks for mobile health knowledge transfer and the involvement of end-users throughout the design process build a strong foundation for an engaging, user-friendly, and effective mobile health intervention. A digital prototype served as a tangible representation of our concept, which we evaluated with patients with heart failure.

Methods

Study Design Overview

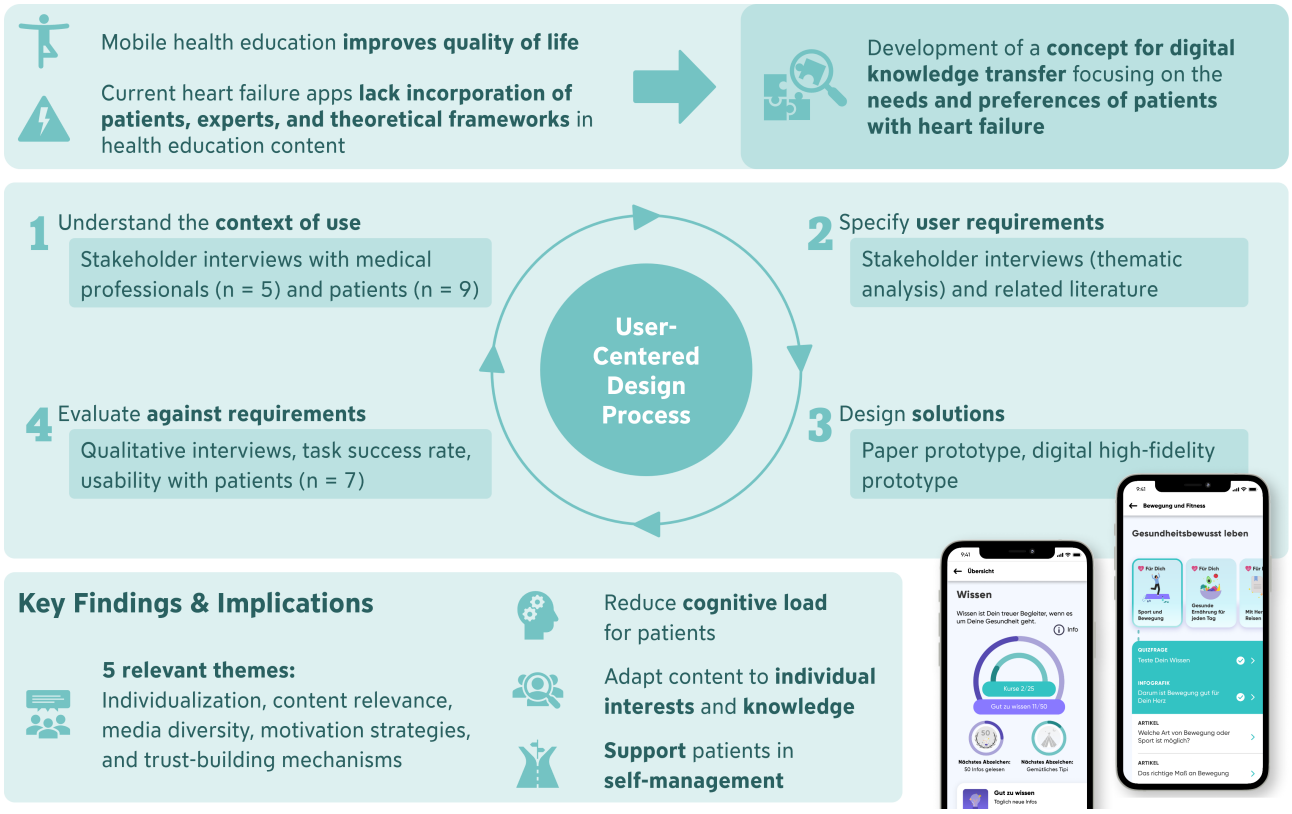
The foundation of this work was the user-centered design process, a common method in interaction design [26]. It emphasizes the involvement of the potential end-users throughout the developmental phases to ascertain that the resulting concepts align with their needs, preferences, and experiences of the target group. This iterative approach often incorporates techniques such as user interviews, focus groups, prototyping, and usability testing to refine the design based on continuous feedback. The graphical abstract (Figure 1) provides an overview of the methodology used.

First, we used current literature and semistructured qualitative interviews to understand the context of use and derive requirements. Based on the findings, possible design solutions were drafted as a paper prototype. Thereafter, a digital high-fidelity prototype was implemented. This digital prototype was then evaluated in a study focusing on the identified requirements of the target group.

The exemplary app we used for the prototype design is the ProHerz app (ProCarement Gmbh, Germany), a medical-grade mobile app for heart failure. This app is intended to support patients in managing their health by tracking vital signs and medication intake. Additionally, the company provides a ‘CareCenter,’ where, depending on the product used, medical experts regularly monitor the patients’ health parameters and stay in personal contact via messages and phone calls.

XSL•FO
RenderX

Figure 1. Graphical abstract illustrating the integration of the user-centered design process in the development of a health knowledge transfer concept for patients with heart failure.



Stakeholder Interviews

Recruitment

Qualitative semistructured interviews were conducted with patients (n=9) and medical experts (n=5). This approach is a flexible data collection method that combines predefined questions with the freedom to explore emerging topics to further enhance the understanding of participants' experiences and perspectives [27]. All patients were recruited from the ProHerz user base. Thus, all were patients with heart failure and previous experience with the ProHerz app. We chose this group because of their experience using a mobile health app for self-management. To allow for meaningful comparisons of feedback, our recruitment strategy aimed to include participants with varying levels of experience using the ProHerz app, aligning with best practices in end-user and usability testing. Medical experts were recruited through ProCarement, where they interact with the ProHerz users in their daily work. All experts were nurses with long-term experience in heart failure care. Their close exchange gives them a broad perspective on patients' characteristics, needs, and skills.

Procedure

The interview guidelines were based on our research aims and current literature. Both stakeholder interviews covered similar aspects (Figure 2). The guidelines aimed to identify patients' specific needs and requirements regarding content and features within a mobile health knowledge transfer.

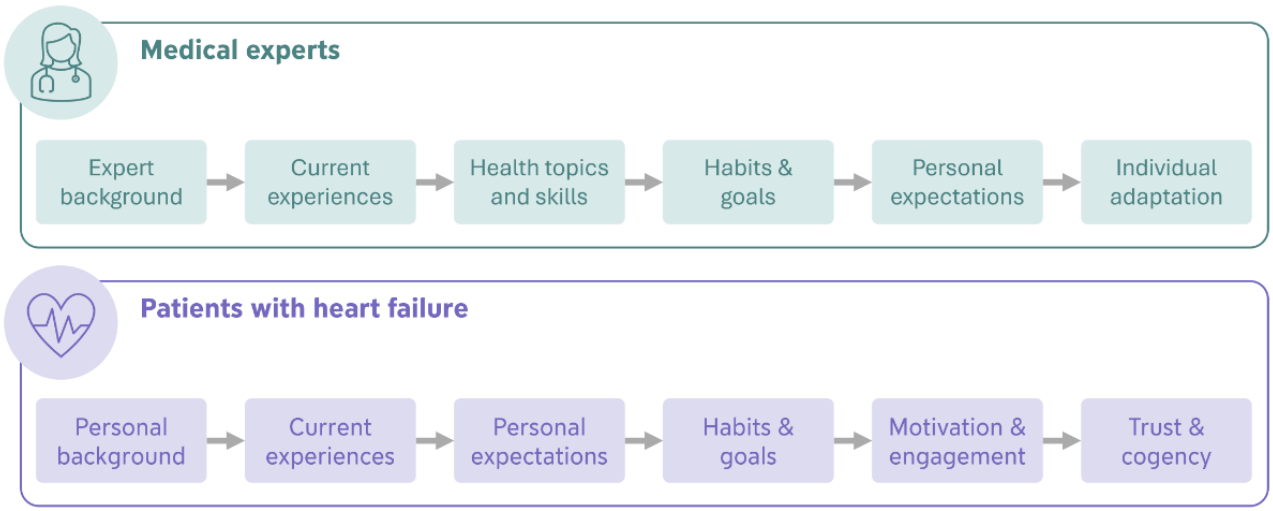
The interviews with the medical experts targeted explicitly at their perspective on essential knowledge and skill requirements that should be covered in a mobile health knowledge transfer concept. Besides, specific features, content presentation, and adaptation strategies for individual patient needs were discussed.

In the patient interviews, we inquired about experiences with specific features, their current interactions with the ProHerz app, and their expectations regarding knowledge transfer within a mobile health app designed for heart failure. Furthermore, the interviews included discussions on patients' health-related routines and goals and identifying requirements for trust-building and cogency. Additionally to the interview, patients were asked to complete the European Health Literacy Survey Questionnaire (EU-HLS-Q16) to assess their health literacy level [25,26]. Through this tool, participants rate general tasks related to health literacy based on their level of difficulty.

All interviews with medical experts were conducted via Microsoft Teams. Patient interviews were conducted in person.

XSL•FO
RenderX

Figure 2. The topics included in the interview with both stakeholder groups.



Prototype Development

The development of the knowledge transfer concept was based on the findings from the interviews and related literature.

We incorporated the framework of Riegel et al to structure the topics [28]. The content categories in the prototype are oriented toward the three self-care processes for heart failure: maintenance, symptom perception, and management. Information in the prototype was structured based on the principles for effective e-learning by Clark and Mayer and the recommendations for designing health-literate mobile apps by Broderick et al [29,30]. This included, for example, dividing information into manageable units, using different types of media, and writing in a conversational style.

Additionally, we considered possible cognitive limitations of users in the context of heart failure [22,31-34]. According to the Cognitive Load Theory, a reduction of cognitive load is able to improve learning outcomes. The theory describes three types of cognitive loads in the context of learning: intrinsic, extraneous, and germane cognitive load. We aimed at reducing the extraneous cognitive load by simplifying navigation, avoiding unnecessary elements, and ensuring a clear, uncluttered interface design. The intrinsic load was managed by sequencing information into smaller units aligned with the complexity of topics. The germane load was supported by including interactive elements such as quizzes or visual aids, encouraging active engagement.

Personalizing health education to address individual needs has been shown to improve health outcomes in areas such as physical activity, nutrition, and adherence to screening recommendations [35]. In the context of health education for heart failure, users face unique challenges through varying levels of health literacy and cognitive impairments [13]. As such, a personalized approach that takes relevant individual factors into account can meet diverse needs of patients, ensuring information is both accessible and relevant [22]. For this reason, we decided to incorporate the possibility of adapting the contents based on a short survey participants answered in the beginning.

The derived themes of the interviews were brought together with findings from the health literacy assessment that additionally revealed the needs of patients. We drafted a workflow for the mobile knowledge transfer and implemented this as a low-fidelity paper prototype.

For the first usability test, we embedded the paper prototype in Useberry (Useberry User Testing Technologies IKE, Greece). Useberry is a tool to test the usability of the paper-based prototype digitally. We aimed to find usability issues as early as possible in our design. For this initial testing, we invited healthy participants through personal contact. At this stage, the priority was gathering quick insights from users who were familiar with mobile apps to identify basic usability issues before moving on to the larger-scale user testing with patients with heart failure. This reduced the burden and required time and effort of the actual patients. Test users (n=5) were asked to perform different navigation-related tasks within the prototype. Outcome measures were the task completion rate and the recorded clicks.

After this initial usability testing, the paper prototype was then transferred into a digital high-fidelity prototype using Figma (Figma GmbH, Germany). This prototype included an example questionnaire for individualizing the information content based on the date of diagnosis, implanted devices (pacemaker and defibrillator), smoking habits, topics of interest, the preferred media formats and the New York Heart Association (NYHA) functional class (Figure 3). The NYHA classification is a commonly used system for categorizing the severity of heart failure based on physical activity limitations caused by cardiac symptoms. It ranges from Class I (no limitations) to Class IV (highest limitations) [36]. After the questionnaire, the basic landing page of the app was presented with an additional link to the health knowledge section. The main part of our prototype was the knowledge section with an overview and link to the available courses and the short information messages. Additionally, badges for learning-related achievements can be found there. Finally, an example course was implemented on the topic of sports and exercise (Figure 4).

XSL•FO
RenderX

Figure 3. Questionnaire for individual adaptation (left), home screen of the app with the link to the knowledge section (middle), and the main page of the knowledge section (right) of the digital prototype.

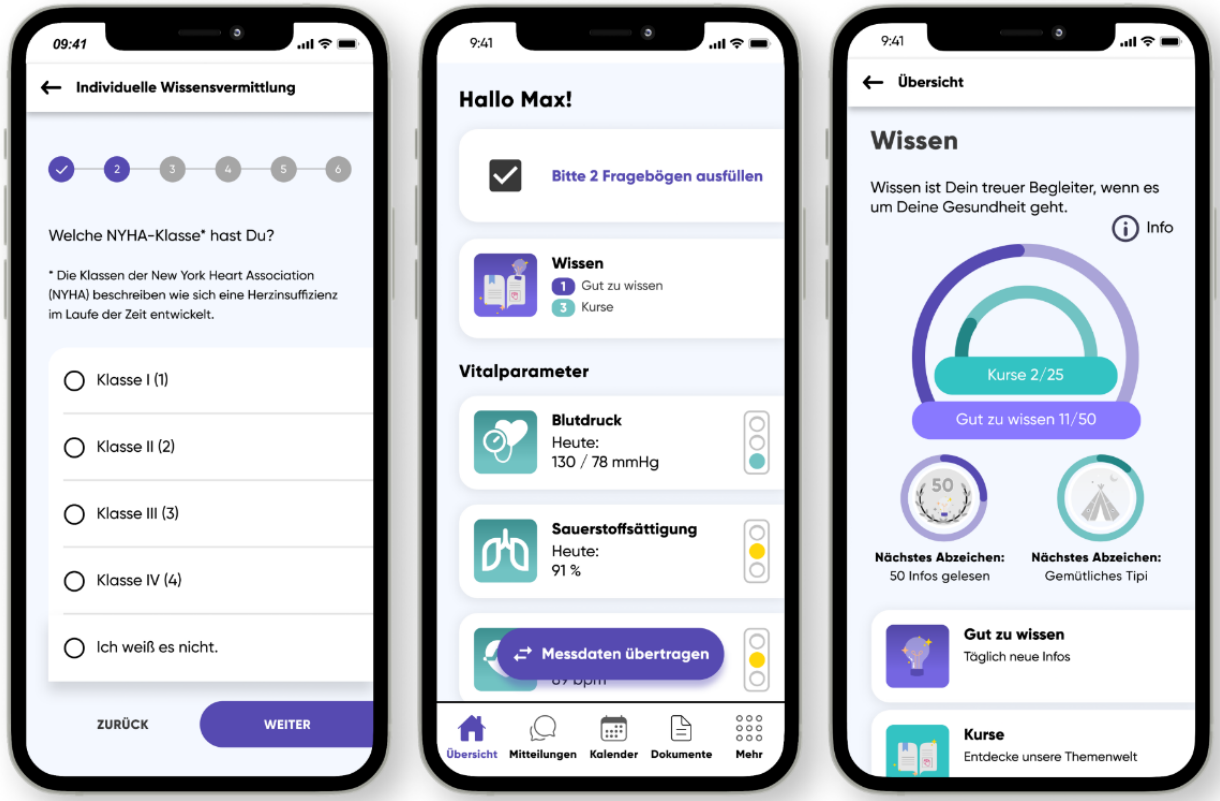
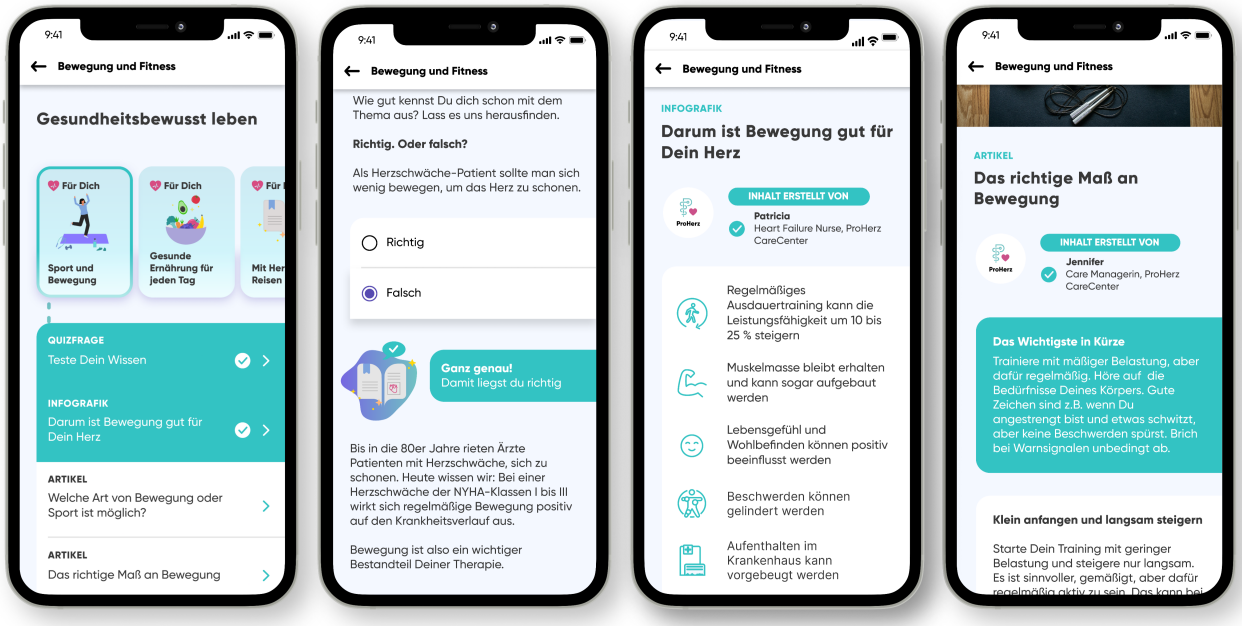


Figure 4. Structure of one exemplary course with (1) the overview of the content, (2) a quiz question, (3) an infographic presenting key facts, and (4) an article starting with information about the author of the text and a summary of the most important aspects of the article.



Concept Evaluation

Recruitment

For evaluating the knowledge transfer concept, we recruited patients with heart failure (n=7), who were already experienced in using a mobile health app. Therefore, all participants were recruited from the user base of ProCarement. Four participants already took part in the previous stakeholder interviews conducted as part of this study.

XSL•FO
RenderX

Procedure

After being instructed about the study procedure, participants were asked to complete 11 tasks with the digital prototype. Due to feasibility reasons, patients accessed the prototype on a computer interface. This approach allowed participants to interact with the app more comfortably during the evaluation, while also enabling researchers to observe user interactions in detail and collect qualitative feedback efficiently. Tasks included small, simple tasks and more complex, time-consuming tasks. An overview of all tasks is listed in Table 1. Afterwards, semistructured interviews were conducted to obtain the participants' overall impression and to identify features and characteristics that have the potential to improve engagement and increase health literacy.

Table . Tasks given to participants in the evaluation of the digital prototype of the health knowledge transfer concept.

No.	Description
1	Navigate to the questionnaire for individualization
2	Complete the questionnaire for individualization
3	Navigate to the knowledge section
4	Navigate to informative notifications, read the first unread notification
5	Navigate to the educational courses
6	Navigate to the course category 'symptoms'
7	Navigate to completed courses
8	Select a course within a course category
9	Navigate to recommended courses and select a recommended course
10	Complete one educational course
11	Navigate to the overview of badges

Analyses

All qualitative data obtained during interviews were analyzed using thematic analysis based on the approach of Braun and Clarke [37]. This method involves identifying, analyzing, and reporting patterns (themes) within the data, allowing for a rich and detailed interpretation of the participants' experiences and perspectives. The process includes familiarizing oneself with the data, generating initial codes, searching for themes, reviewing themes, and defining and naming them to ensure a comprehensive understanding of the qualitative insights.

Ethical Considerations

All presented substudies were approved by the institutional review board of the Friedrich-Alexander-Universität Erlangen-Nürnberg (approval no. 22 - 233-S), and we obtained written informed consent from all participants. The collected data were pseudonymized after collection to ensure participant confidentiality and data protection.

Results

Stakeholder Interviews

Demographics

The stakeholder interviews with patients and medical experts were conducted in August 2022. We recruited 9 patients with a mean (SD) age of 67 (7) years. Patients received their diagnosis between 15 and 3 years prior to this study. All patients were using the ProHerz app for at least 4 months (Table 2). Seven identified as women, and 2 identified as men. Additionally, 5 medical experts were recruited (Table 3) with work experience between 6 and 11 years.

Table . Characteristics of patients with heart failure included in the stakeholder interviews.

Interviewee ID	Age (y)	Identified as	Years since heart failure diagnosis	Months of app experience with ProHerz
P1	60	female	5	9
P2	58	female	3	5
P3	75	female	3	15
P4	78	female	5	15
P5	63	female	4	16
P6	75	female	9	4
P7	69	female	15	5
P8	60	male	3	4
P9	68	male	9	6

XSL•FO
RenderX

Table . Characteristics of medical experts included in the stakeholder interviews.

Interviewee ID	Identified as	Expert experience	Years of work experience with patients with heart failure
Expert 1	female	Heart failure nurse, licensed practical nurse	11
Expert 2	male	Licensed practical nurse	6
Expert 3	female	Licensed practical nurse	6
Expert 4	female	Bachelor of Science in Nursing, specialist in anesthesia and intensive care	9
Expert 5	male	State-qualified nurse	8

Thematic Analysis

We derived five themes relevant to a knowledge transfer concept from the stakeholder interviews with patients with heart failure and medical experts.

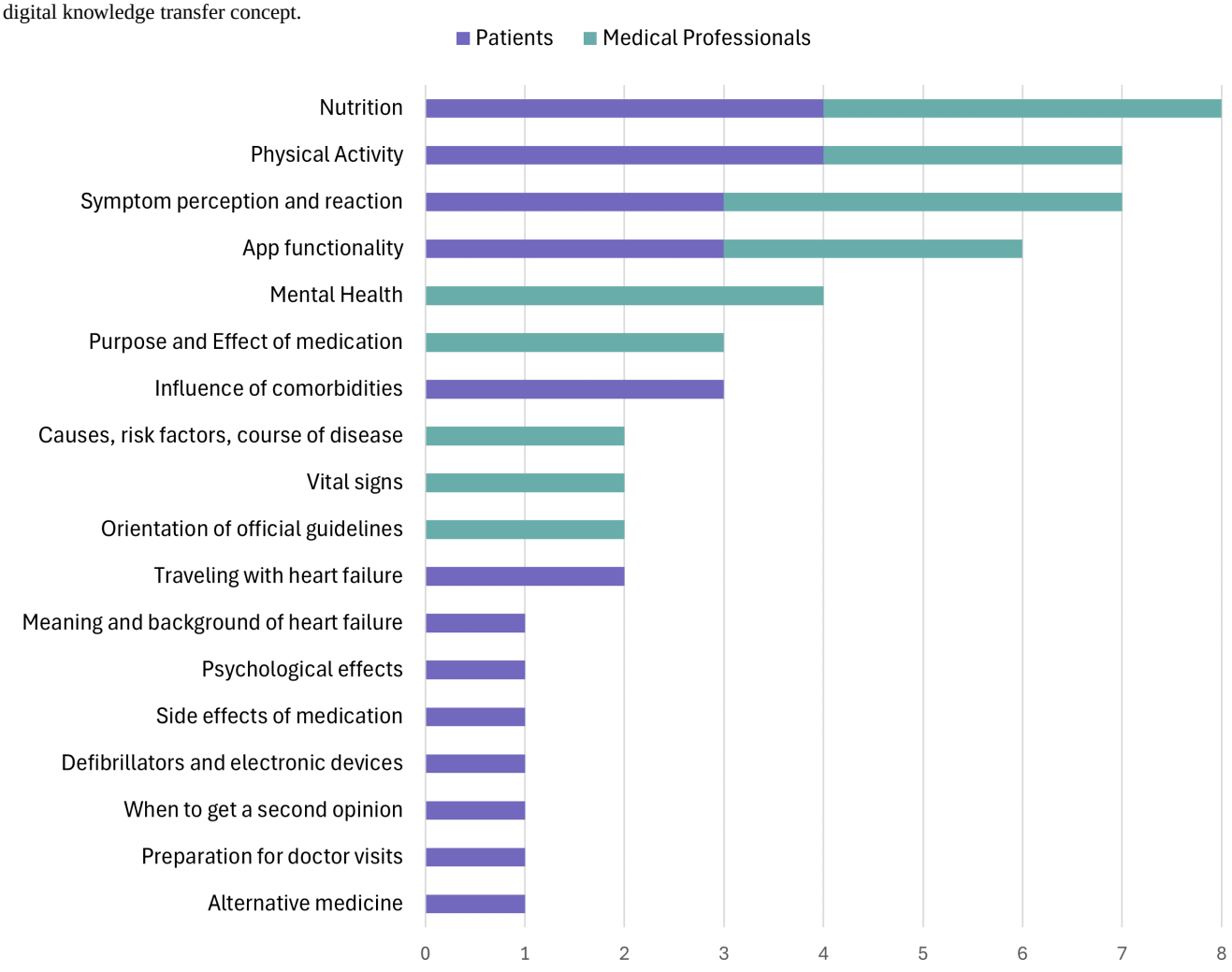
Individual Adaptation of Health Information

The interviews revealed a positive perspective towards the adaptation of content and push notifications to the different levels of knowledge of the users (4/5 experts, 2/9 patients). Our results suggest individualization based on different information such as the NYHA class, walking distance, app usage duration, physical condition age, prior knowledge, or personal preferences that could be determined for example through a quiz within the app.

Relevant Health Topics for Patients With Heart Failure

Several relevant health topics were identified within the stakeholder interviews (Figure 5). Health topics most requested by both patients and medical experts included nutrition (4/5 experts, 4/9 patients), physical activity (3/5 medical experts, 4/9 patients), symptom perception and reaction (4/5 medical experts, 3/9 patients), and app functionality (3/5 medical experts, 3/9 patients).

Figure 5. Health topics mentioned in the stakeholder interviews with patients and medical experts and that were considered relevant to be included in a digital knowledge transfer concept.



XSL•FO
RenderX

Diversity of Media Formats

Most of the interviewed patients (7/9 patients) preferred information in text format. For example, 1 participant stated, "you can read it two or three times if you say it's something that's difficult to absorb, maybe a text would be better" (P5, translated). Additionally, videos were perceived as useful by all medical experts but only desired by 2 patients. Related to the presentation of the information, patients mentioned using fewer technical terms and more simple language (2/9 patients) and shortly highlighting the most important aspects (2/9 patients). This can support easy comprehension of contents, which was also mentioned as a factor for the trustworthiness of health information (3/9 patients).

Promotion of Motivation, Engagement, and Joy

Medical experts (4/5 experts) suggested a reward system to increase engagement with the information contents. However, it should be ensured that this reward system is suitable for the specific target group of users (2/5). Three patients also mentioned the wish to receive positive feedback for progress within the app. Personal interest in the content was also mentioned (3/9) to have a positive influence on engagement.

Credibility and Trust

From the interviews with patients, we identified several factors related to the credibility of information. Positively perceived was the specification of authors (3/9 patients), especially if the authors are physicians (2/9 patients) or authors with an academic background (2/9 patients). In addition, the provider of the mHealth app should have a good reputation (3/9 patients). Factors restraining trust and credibility could be advertising (1/9 patients), authors without medical background (1/9 patients), or irrelevant information (1/9 patients).

Health Literacy

Health literacy levels based on the EU-HLS-Q16 were between 10 and 16 with an average (SD) score of 13.6 (2.1). Participants mostly had difficulties deciding on preventive actions based on media information (5/9 patients) and determining whether information about health risks in media is trustworthy (4/9 patients).

Evaluation of the Digital Prototype

Demographics

We evaluated the proposed digital prototype of the knowledge transfer concept with 7 patients with heart failure (Table 4). The mean (SD) age was 65 (6) years. Patients received their diagnosis between 2 and 24 years prior to this study. All patients were using the ProHerz app for at least 5 months. Four participants identified as women, and 3 as men. This group was less diverse in terms of ProHerz experience, which was a result of recruiting participants based on availability and willingness to participate.

Table . Characteristics of patients included in the evaluation study of the knowledge transfer concept.

Interviewee ID	Age (y)	Gender	Years since heart failure diagnosis	Months of app experience
P1	60	female	5	9
P2	58	female	3	5
P3	75	female	3	15
P5	63	female	4	16
P10	59	male	9	15
P11	71	male	2	15
P12	69	male	24	10

Usability

The mean task completion rate was 84%; 5 of the 11 tasks given during the study were completed by all participants. The lowest completion rate occurred with task 4, which was completed by 4 of 7 participants (Table 5).

XSL•FO
RenderX

Table . Task completion rate for all tasks performed by the participants within the usability testing.

No.	Description	Completion rate (%)
1	Navigate to the questionnaire for individualization	100
2	Complete the questionnaire for individualization	100
3	Navigate to the knowledge section	100
4	Navigate to informative notifications, read the first unread notification	57
5	Navigate to the educational courses	71
6	Navigate to the course category ‘symptoms’	71
7	Navigate to completed courses	71
8	Select a course within a course category	71
9	Navigate to recommended courses and select a recommended course	86
10	Complete one educational course	100
11	Navigate to the overview of badges	100

Interview Results

The thematic analysis of the semistructured interviews with patients within the evaluation of our proposed concept resulted in 3 main themes.

Knowledge Transfer

In the evaluation of the digital prototype, the majority of participants (4/7 patients) classified the prospective range of information as valuable. The organization and structure of contents received positive feedback from most participants (5/7 patients). Specifically, the clear arrangement of information (3/7 patients) and the definition of specific categories (2/7 patients) were highlighted. Patients also perceived the individual adaptation as positive, because personally irrelevant content would be hidden (4/7 patients).

To further improve the concept, participants indicated the necessity to update contents in response to potentially evolving needs over time due to lifestyle or health condition changes (2/7 patients).

Motivation and Engagement

A majority (5/7 patients) expressed an intention to utilize the course content in the future. Notably, reminders about feasible health-promoting measures were identified as a motivational factor by 3 out of 7 patients, with a specific emphasis from 2 participants on messages related to physical activity.

Concerns were raised about the potential negative impact of irrelevant, repetitive, or overly frequent notifications (4/7 patients). Conversely, aligning information with individual interests was identified as a key factor in promoting joy and sustaining engagement among users.

Patients expressed joy while answering the quiz questions, which may have contributed to a more engaging experience (2/7 patients). There was a suggestion to place quiz questions at the end of a course to assess the comprehension of the content effectively.

Participants generally responded positively to medical staff monitoring of progress, considering it a supportive element in their engagement with the course. However, 1 participant viewed social comparison negatively. Furthermore, intangible rewards, such as badges, were not universally well-received, with 1 out of 7 participants expressing a negative opinion, stating that it is solely virtual and not a real reward (P11).

Ease of Use

Out of 7 patients, 3 patients explicitly mentioned the app as easy to use. Other participants did not comment on this aspect, neither positively nor negatively. Concerns regarding the ease of use included potential initial overload when first using the app and potential individual difficulties with the technology and navigation.

Discussion

Principal Findings and Implications

In this study, we developed a concept for mobile health knowledge transfer for patients with heart failure. We grounded this concept on the requirements of patients with heart failure and current theoretical frameworks. From the patient and caregiver perspectives, we derived five central themes for a health knowledge transfer concept.

Our findings contribute to the understanding of user preferences and content organization in digital health interventions targeted at users with heart failure. The integration of diverse methodologies, including qualitative interviews, usability testing, and participant feedback, fortifies the study’s comprehensiveness, providing insights into user engagement, health literacy promotion, and the practical implications for digital health interventions. The results of this study represent an important step toward the development of a mobile health knowledge transfer tool for heart failure patients. By identifying relevant features and characteristics, this work provides a foundation for further refinement and testing, ultimately aiming to support health literacy and improve patient outcomes. It

XSL•FO
RenderX

becomes evident that a concept should focus on individualization, topic relevance, media diversity, motivation strategies, and trust-building mechanisms to optimize its impact on health literacy and engagement of patients with heart failure. Based on the findings, identified in the stakeholder interviews, we developed a concept for knowledge transfer in a digital health app for patients with heart failure.

We evaluated our concept with a digital prototype and identified features and characteristics that are perceived as supportive in increasing health literacy, that can promote motivation, engagement, and joy, and that are able to convince patients of a mobile health knowledge transfer. Most prominent was the wish for content that is individually adapted and balanced regarding the topics covered as well as media formats used to present those topics.

The results of the prototype evaluation presented both encouraging feedback and areas for refinement. Participants highlighted the app's structured design and ease of navigation, emphasizing that these features facilitated a straightforward user experience and enhanced the clarity of the presented information. The quiz elements, in particular, were frequently mentioned as engaging and motivating, showcasing the potential of interactive components to sustain user interest and reinforce learning. These positive responses suggest that the app effectively addressed some of its primary goals, including improving knowledge transfer and supporting patient engagement.

Conversely, the feedback also identified challenges that should be addressed in future iterations. Participants expressed concerns of getting overwhelmed, potentially hindering comprehension. This may reflect concerns about managing excessive information or complex interactions within the app, particularly in the context of their health condition. For users with chronic conditions like heart failure, the introduction of too much content or functionality in a single interface could exacerbate feelings of being overwhelmed and lead to disengagement or frustration [38]. Furthermore, badges and achievements were mentioned to be not motivating despite research showing the positive effects of such gamification elements [39,40]. Older users or those less familiar with gamification might find such features irrelevant or even patronizing if they do not see clear value in them. Other gamification elements such as progress bars, interactive storylines, or more useful rewards may be more effective in motivating users. Therefore, future iterations could explore the inclusion of a simple setup and customization menu to enable or disable features such as gamification and other user-preferred functionalities.

Accessibility is a critical consideration in designing digital health systems to ensure they meet the needs of diverse users. While none of the participants in our study raised concerns or reported difficulties with accessibility, indicating that the system was well-received by the target population, it is important to account for potential challenges in broader user groups. Features such as larger fonts, high-contrast visuals, simplified navigation, and optional audio support can further enhance usability, particularly for individuals with age-related changes in vision, hearing, or varying levels of digital literacy.

The high rates of completion for most tasks suggest that the interface and navigation design are intuitive, enabling users to locate and interact with the desired functions efficiently. This aligns with the positive feedback received during the interviews conducted within the evaluation. However, with a few tasks having a lower completion rate, some aspects of the design need further refinement in future iterations. Factors such as differences in technology familiarity, cognitive load, or previous experience with the ProHerz app could have contributed to these discrepancies. Interactive tutorials to onboard users and a further adaptation of content may be able to better account for different requirements, levels of experience, or cognitive abilities. In particular, task 4 with a completion rate of only 57% was the most problematic task for participants. The participants were not able to locate the button to navigate to the informative notifications. The reason for this is most likely the design of the progress bar on the same page (Figure 3, right) that users expected to be the button to the next page. Ensuring more consistent visual cues for interactive elements could further reduce the potential for misinterpretation and enhance the overall user experience.

Comparison With Prior Work

The need for individualization is in line with the findings of Giordan et al [41], calling for further customization especially related to the educational level and digital literacy of users. Our results additionally highlight the need for customization regarding personal interests and lifestyle and the importance of regular adaptation to changing circumstances in the life and knowledge of users. Future applications require new strategies to individually adapt educational content that can change over time. This can include regularly soliciting user feedback through surveys, using quizzes to assess understanding, or adaptive learning algorithms based on user behavior within the mobile app [42].

The topics covered in our knowledge transfer concept are in line with the recommendations of the European Society of Cardiology. They included information about heart failure, symptom monitoring, self-management, medication, fluid intake, implantable devices, physical activity, nutrition, smoking, sleep mental health, and traveling [43]. The information was presented using different media formats, which relates to current literature emphasizing the benefits of this approach. A variety of formats can support users with different health literacy levels and can deepen the learning effects [29,43,44].

Strengths and Weaknesses

Several limitations apply to this study. Firstly, the participant pool featured a restricted number of individuals, all German-speaking and residents in Germany. Moreover, participants were well-acquainted with the ProHerz app, potentially influencing their perspectives and feedback, especially regarding the ease of use. Given the advanced health literacy levels of the participants, our findings may not comprehensively represent the needs of individuals with lower health literacy. The initial usability testing was conducted with a convenience sample due to limited patient availability; therefore, results from this group need to be interpreted with caution. During the evaluation phase, 4 participants already

XSL•FO
RenderX

were part of the stakeholder interviews in the first phase of this study due to limited participant availability. Lastly, all participants accessed the app on a laptop rather than a mobile phone, introducing a potential discrepancy, as some users reported difficulties with computer usage and the overall interaction experience may be different. These limitations underscore the necessity for cautious interpretation and highlight avenues for future research refinement.

Our study combines several strengths: we were the first to conceptualize a mobile health knowledge transfer within a user-centered design approach. By prioritizing patient needs and requirements, and involving healthcare professionals to ensure relevance and meaningfulness, the study sets a precedent in fostering impactful and patient-centric digital health interventions. The incorporation of theoretical frameworks for mobile health education enriches the conceptual foundation, contributing to the overall robustness of the study and its potential impact on the enhancement of patient-centered care.

Conclusions

We were the first to develop a mobile health knowledge transfer for heart failure grounded on theoretical frameworks within a user-centered design process.

By addressing individual health literacy levels, cognitive challenges, and personal preferences, this knowledge transfer concept can help patients better understand and manage their condition. The active involvement of healthcare professionals ensures that the resulting intervention is meaningful and aligned with relevant healthcare aspects. The findings also emphasize a discrepancy between what healthcare professionals perceive as important for patients and what patients themselves prioritize, particularly when it comes to identifying relevant healthcare topics. Our results are paving the way to more personalized and effective mobile health education for individuals managing heart failure. The concept can be used as a foundation for other digital platforms and mobile health apps to provide tailored educational content and supporting clinicians in delivering consistent, patient-specific information in home settings.

Future research opportunities should focus on longitudinal studies to assess the sustained impact of this knowledge transfer concept on health literacy, self-care behavior, and patient outcomes. In this context, validated usability scales alongside user interviews should be incorporated during iterative development processes, while also addressing potential biases in design and testing to ensure more robust and generalizable findings. Additionally, possible strategies for better individualization of health information should be further explored. To better understand the impact of personalization, future studies could incorporate a comparison of user experiences before and after customizing the app, providing valuable insights into how individualization influences usability and user satisfaction. This presents an opportunity to investigate adaptive learning algorithms that dynamically tailor content based on user progress and preferences.

Acknowledgments

This project is supported by the Bavarian State Ministry of Health and Care, project grant number PBN-MGP-2010-0004-DigiOnko. The authors would like to acknowledge the financial support from Deutsche Forschungsgemeinschaft and Friedrich–Alexander Universität Erlangen-Nürnberg within the funding program Open Access Publication Funding.

Authors' Contributions

MF conceived the study, developed the study design, analyzed data, assisted in interviews, and wrote the manuscript draft. SB and JR assisted in developing the study design, interview conduction, and data analysis. KMJ, MN, and HL assisted in writing the manuscript. PT, SE, BME, and HL initiated and supervised the project. All authors reviewed the final manuscript.

Conflicts of Interest

JR, SE, and PT are employees of ProCarement GmbH. All other authors declare no conflicts of interest.

References

1. Yan T, Zhu S, Yin X, et al. Burden, trends, and inequalities of heart failure globally, 1990 to 2019: a secondary analysis based on the global burden of disease 2019 study. J Am Heart Assoc 2023 Mar 21;12(6):e027852. [doi: 10.1161/JAHA.122.027852] [Medline: 36892088]
2. Heidenreich PA, Fonarow GC, Opsha Y, et al. Economic issues in heart failure in the United States. J Card Fail 2022 Mar;28(3):453-466. [doi: 10.1016/j.cardfail.2021.12.017] [Medline: 35085762]
3. Roth GA, Forouzanfar MH, Moran AE, et al. Demographic and epidemiologic drivers of global cardiovascular mortality. N Engl J Med 2015 Apr 2;372(14):1333-1341. [doi: 10.1056/NEJMoa1406656] [Medline: 25830423]
4. Guasti L, Dilaveris P, Mamas MA, et al. Digital health in older adults for the prevention and management of cardiovascular diseases and frailty. A clinical consensus statement from the ESC Council for Cardiology Practice/Taskforce on Geriatric Cardiology, the ESC Digital Health Committee and the ESC Working Group on e-Cardiology. ESC Heart Fail 2022 Oct;9(5):2808-2822. [doi: 10.1002/ehf2.14022] [Medline: 35818770]
5. Dunn P, Margaritis V, Anderson C. Understanding health literacy skills in patients with cardiovascular disease and diabetes. TQR 2017 Jan 2. [doi: 10.46743/2160-3715/2017.2548]

XSL•FO
RenderX

6. Michalsen A, König G, Thimme W. Preventable causative factors leading to hospital admission with decompensated heart failure. Heart 1998 Nov;80(5):437-441. [doi: 10.1136/hrt.80.5.437] [Medline: 9930040]
7. Wen CP, Wai JPM, Tsai MK, et al. Minimum amount of physical activity for reduced mortality and extended life expectancy: a prospective cohort study. The Lancet 2011 Oct;378(9798):1244-1253. [doi: 10.1016/S0140-6736(11)60749-6]
8. Evangelista LS, Rasmusson KD, Laramee AS, et al. Health literacy and the patient with heart failure--implications for patient care and research: A consensus statement of the Heart Failure Society of America. J Card Fail 2010 Jan;16(1):9-16. [doi: 10.1016/j.cardfail.2009.10.026] [Medline: 20123313]
9. Matsuoka S, Tsuchihashi-Makaya M, Kayane T, et al. Health literacy is independently associated with self-care behavior in patients with heart failure. Patient Educ Couns 2016 Jun;99(6):1026-1032. [doi: 10.1016/j.pec.2016.01.003] [Medline: 26830514]
10. Wu JR, Reilly CM, Holland J, Higgins M, Clark PC, Dunbar SB. Relationship of health literacy of heart failure patients and their family members on heart failure knowledge and self-care. J Fam Nurs 2017 Feb;23(1):116-137. [doi: 10.1177/1074840716684808] [Medline: 28795936]
11. Sørensen K, Van den Broucke S, Fullam J, et al. Health literacy and public health: a systematic review and integration of definitions and models. BMC Public Health 2012 Jan 25;12(1):80. [doi: 10.1186/1471-2458-12-80] [Medline: 22276600]
12. Magnani JW, Mujahid MS, Aronow HD, et al. Health literacy and cardiovascular disease: fundamental relevance to primary and secondary prevention: a scientific statement from the American Heart Association. Circulation 2018 Jul 10;138(2):e48-e74. [doi: 10.1161/CIR.0000000000000579] [Medline: 29866648]
13. Fabbri M, Murad MH, Wennberg AM, et al. Health Literacy and outcomes among patients with heart failure: a systematic review and meta-analysis. JACC Heart Fail 2020 Jun;8(6):451-460. [doi: 10.1016/j.jchf.2019.11.007] [Medline: 32466837]
14. Kim H, Xie B. Health literacy in the eHealth era: A systematic review of the literature. Patient Educ Couns 2017 Jun;100(6):1073-1082. [doi: 10.1016/j.pec.2017.01.015] [Medline: 28174067]
15. Conard S. Best practices in digital health literacy. Int J Cardiol 2019 Oct 1;292:277-279. [doi: 10.1016/j.ijcard.2019.05.070] [Medline: 31230937]
16. Zippel-Schultz B, Palant A, Helms TM. Digitale Lösungen für die versorgung von herzinsuffizienzpatienten. In: Pfannstiel MA, Holl F, Swoboda WJ, editors. mHealth-Anwendungen Für Chronisch Kranke Wiesbaden [Book in German]: Springer Fachmedien Wiesbaden; 2020:135-155. [doi: 10.1007/978-3-658-29133-4_8]
17. Beneito-Montagut R, Rosales A, Fernández-Ardèvol M. Emerging digital inequalities: a comparative study of older adults' smartphone use. Social Media + Society 2022 Oct;8(4):20563051221138756. [doi: 10.1177/20563051221138756]
18. Mackert M, Champlin SE, Holton A, Muñoz II, Damásio MJ. eHealth and Health Literacy: a research methodology review. J Comput-Mediat Comm 2014 Apr;19(3):516-528. [doi: 10.1111/jcc4.12044]
19. Mein E, Fuentes B, Soto Más F, Muro A. Incorporating digital health literacy into adult ESL education on the US-Mexico border. Rhetor Prof Commun Glob 2012 Dec;3(1):162-174. [Medline: 23730533]
20. Blunck D, Kastner L, Nissen M, Winkler J. The effectiveness of patient training in inflammatory bowel disease knowledge via Instagram: randomized controlled trial. J Med Internet Res 2022 Oct 19;24(10):e36767. [doi: 10.2196/36767] [Medline: 36260385]
21. Klimis H, Thakkar J, Chow CK. Breaking barriers: mobile health interventions for cardiovascular disease. Can J Cardiol 2018 Jul;34(7):905-913. [doi: 10.1016/j.cjca.2018.02.012] [Medline: 29887218]
22. Allida S, Du H, Xu X, et al. mHealth education interventions in heart failure. Cochrane Database Syst Rev 2020 Jul 2;7(7):CD011845. [doi: 10.1002/14651858.CD011845.pub2] [Medline: 32613635]
23. Bakogiannis C, Tsarouchas A, Mouselimis D, et al. A patient-oriented app (ThessHF) to improve self-care quality in heart failure: from evidence-based design to pilot study. JMIR Mhealth Uhealth 2021 Apr 13;9(4):e24271. [doi: 10.2196/24271] [Medline: 33847599]
24. Ahn JA. Heart failure-smart life ii: Effects of a mobile-based education and self-management program for patients with heart failure and their family caregivers. J Am Coll Cardiol 2022 Mar;79(9):2027. [doi: 10.1016/S0735-1097(22)03018-2]
25. Heiney SP, Donevant SB, Arp Adams S, Parker PD, Chen H, Levkoff S. A smartphone app for self-management of heart failure in older African Americans: feasibility and usability study. JMIR Aging 2020 Apr 3;3(1):e17142. [doi: 10.2196/17142] [Medline: 32242822]
26. Richter M, Flückiger M. User-Centred Engineering: Springer Berlin Heidelberg; 2014. [doi: 10.1007/978-3-662-43989-0]
27. Adeoye‐Olatunde OA, Olenik NL. Research and scholarly methods: semi‐structured interviews. J Am Coll Clin Pharm 2021 Oct;4(10):1358-1367 [FREE Full text] [doi: 10.1002/jac5.1441]
28. Riegel B, Dickson VV, Faulkner KM. The situation-specific theory of heart failure self-care. J Cardiovasc Nurs 2016;31(3):226-235. [doi: 10.1097/JCN.0000000000000244] [Medline: 25774844]
29. Clark RC, Mayer RE. E-Learning and the Science of Instruction: Proven Guidelines for Consumers and Designers of Multimedia Learning, 4th edition: Hoboken: Wiley; 2016.
30. Broderick J, Devine T, Lemerise AJ, et al. Designing health literate mobile apps. NAM Perspectives 2014 Jan 28;4(1) [FREE Full text] [doi: 10.31478/201401a]
31. van den Hurk K, Reijmer YD, van den Berg E, et al. Heart failure and cognitive function in the general population: the Hoorn Study. Eur J Heart Fail 2011 Dec;13(12):1362-1369. [doi: 10.1093/eurjhf/hfr138]

XSL•FO
RenderX

32. Currie K, Rideout A, Lindsay G, Harkness K. The association between mild cognitive impairment and self-care in adults with chronic heart failure: a systematic review and narrative synthesis. J Cardiovasc Nurs 2015;30(5):382-393. [doi: 10.1097/JCN.0000000000000173] [Medline: 24988321]
33. Vogels RLC, Scheltens P, Schroeder‐Tanka JM, Weinstein HC. Cognitive impairment in heart failure: a systematic review of the literature. European J of Heart Fail 2007 May;9(5):440-449. [doi: 10.1016/j.ejheart.2006.11.001]
34. Wilson EAH, Wolf MS. Working memory and the design of health materials: a cognitive factors perspective. Patient Educ Couns 2009 Mar;74(3):318-322. [doi: 10.1016/j.pec.2008.11.005] [Medline: 19121915]
35. Cohn WF, Lyman J, Broshek DK, et al. Tailored educational approaches for consumer health: a model to address health promotion in an era of personalized medicine. Am J Health Promot 2018 Jan;32(1):188-197. [doi: 10.1177/0890117116671082] [Medline: 29214832]
36. Bennett JA, Riegel B, Bittner V, Nichols J. Validity and reliability of the NYHA classes for measuring research outcomes in patients with cardiac disease. Heart Lung 2002 Jul;31(4):262-270. [doi: 10.1067/mhl.2002.124554]
37. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
38. Issom DZ, Henriksen A, Woldaregay AZ, Rochat J, Lovis C, Hartvigsen G. Factors influencing motivation and engagement in mobile health among patients with sickle cell disease in low-prevalence, high-income countries: qualitative exploration of patient requirements. JMIR Hum Factors 2020 Mar 24;7(1):e14599. [doi: 10.2196/14599] [Medline: 32207692]
39. Shakoor A, Mohansingh C, van der Boon RMA, Brugts JJ, Schaap J. Gamification and its potential for better engagement in the management of heart failure or quality of care registries: a viewpoint. Curr Heart Fail Rep 2025 Dec;22(1):4. [doi: 10.1007/s11897-024-00692-8]
40. Berglund A, Jaarsma T, Berglund E, Strömberg A, Klompstra L. Understanding and assessing gamification in digital healthcare interventions for patients with cardiovascular disease. Eur J Cardiovasc Nurs 2022 Aug 29;21(6):630-638. [doi: 10.1093/eurjcn/zvac048] [Medline: 35709297]
41. Bezerra Giordan L, Tong HL, Atherton JJ, et al. The use of mobile apps for heart failure self-management: systematic review of experimental and qualitative studies. JMIR Cardio 2022 Mar 31;6(1):e33839. [doi: 10.2196/33839] [Medline: 35357311]
42. Tong HL, Quiroz JC, Kocaballi AB, et al. Personalized mobile technologies for lifestyle behavior change: A systematic review, meta-analysis, and meta-regression. Prev Med 2021 Jul;148:106532. [doi: 10.1016/j.ypmed.2021.106532] [Medline: 33774008]
43. McDonagh TA, Metra M, Adamo M, et al. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. Eur Heart J 2021 Sep 21;42(36):3599-3726. [doi: 10.1093/eurheartj/ehab368] [Medline: 34447992]
44. Smith B, Magnani JW. New technologies, new disparities: The intersection of electronic health and digital health literacy. Int J Cardiol 2019 Oct 1;292:280-282. [doi: 10.1016/j.ijcard.2019.05.066] [Medline: 31171391]

Abbreviations

NYHA: New York Heart Association

Edited by A Kushniruk; submitted 05.04.24; peer-reviewed by R Marshall, W Luan; revised version received 24.01.25; accepted 28.01.25; published 31.03.25.

Please cite as:
Flaucher M, Berzins S, Jaeger KM, Nissen M, Rolny J, Trißler P, Eckl S, Eskofier BM, Leutheuser H
Perception and Evaluation of a Knowledge Transfer Concept in a Digital Health Application for Patients With Heart Failure: Mixed Methods Study
JMIR Hum Factors 2025;12:e56798
URL: https://humanfactors.jmir.org/2025/1/e56798
doi:10.2196/56798

© Madeleine Flaucher, Sabrina Berzins, Katharina M Jaeger, Michael Nissen, Jana Rolny, Patricia Trißler, Sebastian Eckl, Bjoern M Eskofier, Heike Leutheuser. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 31.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Utilization of Electronic Consultations (eConsults) to Address Emerging Questions Related to Long COVID-19 in Ontario, Canada: Mixed Methods Analysis

Jatinderpreet Singh[1,2], MD, MSc, MASc; Michael Quon[3,4,5], MD; Danica Goulet[6,7], MSc; Erin Keely[5,7,8,9,10], MD; Clare Liddy[2,6,7,9,10], MD, MSc

[1]Department of Family Medicine, Queen's University, Kingston, ON, Canada
[10]Ontario Health, Toronto, ON, Canada
[2]Department of Family Medicine, University of Ottawa, Ottawa, ON, Canada
[3]Department of Medicine (General Internal Medicine), The Ottawa Hospital, Ottawa, ON, Canada
[4]Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada
[5]Department of Medicine, University of Ottawa, Ottawa, ON, Canada
[6]C T Lamont Primary Health Care Research Centre, Bruyère Health Research Institute, Ottawa, ON, Canada
[7](Ontario) eConsult Centre of Excellence, The Ottawa Hospital, Ottawa, ON, Canada
[8]Division of Endocrinology and Metabolism, The Ottawa Hospital, Ottawa, ON, Canada
[9]Ottawa Hospital Research Institute, Ottawa, ON, Canada

Corresponding Author:
Clare Liddy, MD, MSc
Department of Family Medicine, University of Ottawa, Ottawa, ON, Canada

Abstract

Background: Long COVID is an often debilitating condition affecting millions of people. Its diverse clinical presentations make effective diagnosis and management at the primary care level difficult, while specialist services for long COVID face extensive wait times. An electronic consultation (eConsult) program in Ontario developed a long COVID specialist group to allow primary care providers (PCPs) prompt access to specialist advice for patients with long COVID.

Objective: This study aims to assess patterns of service use, response times, impact, and clinical content of eConsult cases submitted to an eConsult long COVID specialist group in Ontario.

Methods: This study is a mixed methods analysis of eConsults submitted by PCPs to the long COVID specialist group of 2 eConsult services (Champlain eConsult BASE and Ontario eConsult) between June 1, 2021, and July 31, 2022. Data sources included the use data collected automatically by the services, responses to a mandatory closeout survey, and the content of PCP questions and specialist responses (Champlain eConsult BASE service only). Clinical questions or responses were analyzed using 2 validated taxonomies. Descriptive statistics were used for survey responses and use data.

Results: A total of 40 PCPs submitted 47 eConsults through Champlain eConsult BASE and 197 PCPs submitted 228 cases through Ontario eConsult. The median specialist response time was 0.6 (IQR 0.19-2.36; mean 1.7, SD 2.29) days. The 5 most common symptoms of long COVID were fatigue (14/47, 30%), dyspnea (7/47, 15%), cough (6/47, 13%), altered sense of smell (ie, anosmia and parosmia; 6/47, 13%), and cognitive changes (6/47, 13%). The five main question categories asked by PCPs were: (1) management of chronic symptoms of COVID-19, (2) need for additional work-up or follow-up testing, (3) community resources to support or manage patients with long COVID, (4) diagnostic clarification, and (5) guidance regarding COVID-19 vaccination.

Conclusions: The long COVID groups provided rapid access to a multispecialty service that facilitated the avoidance of unnecessary face-to-face referrals. An assessment of eConsults highlighted 5 common question types, providing insight into potential gaps in knowledge among PCPs that could help guide medical education and policy.

(JMIR Hum Factors 2025;12:e58582) doi:10.2196/58582

KEYWORDS

COVID-19; long COVID; eConsult; consultation; Canada; mixed methods analysis; diagnosis; primary care; electronic consultation; COVID specialist; specialist; patient; assessment; COVID-19 vaccination; vaccination; symptom; medical education; web-based consultation; teleconsultation; web-based consultations

XSL•FO
RenderX

Introduction

The COVID-19 pandemic has had an unprecedented impact on health care systems and economies worldwide. As of November 2023, there have been over 700 million reported cases of COVID-19, resulting in over 6.5 million deaths [1]. As the pandemic began spreading, the potential for long-term morbidity from COVID-19 became apparent. The World Health Organization reports that 10% - 20% of those infected with COVID-19 report incomplete recovery months beyond the acute illness—a condition commonly referred to as long COVID or post-COVID-19 condition [2].

A recent Statistics Canada survey reported that 14.8% of respondents with a previous positive test or suspected infection for COVID-19 experienced ongoing symptoms at least 3 months after infection [3]. Potential symptoms of long COVID are varied, and people often experience symptoms across multiple organ systems.

Primary care is ideally positioned to diagnose and provide management for patients with long COVID. Unfortunately, given its diverse clinical presentations and limited empirical evidence guiding management, many primary care providers (PCPs) lack the specialized knowledge to effectively identify and manage long COVID. In certain jurisdictions, long COVID specialty clinics have been established to support primary care. However, demand and wait times for these clinics make timely access a challenge. There is a need for innovative tools to support PCPs in identifying and delivering care for those with long COVID.

One solution to these challenges is electronic consultation (eConsult), which allows PCPs to ask clinical questions regarding their patients to specialists using a secure web-based platform. In response to the pandemic, an eConsult service operating in Ontario, Canada, launched a long COVID specialist group, allowing PCPs to connect with specialists across multiple disciplines with expertise in long COVID to help with the diagnosis and management of their patients. Previous studies assessing this eConsult service in Ontario have demonstrated improved access to specialty care, reduced need for face-to-face specialist visits, cost savings, and high physician satisfaction [4-8].

In this study, we examined eConsult cases submitted to a long COVID specialist group in Ontario to assess patterns of service use, response times, the impact on the need for formal face-to-face referrals, and the content of clinical questions being asked.

Methods

Study Design

This study is a mixed methods analysis of eConsults submitted by PCPs to the long COVID specialist group between June 1, 2021, and July 31, 2022.

Ethical Considerations

This study received ethics approval from the Ottawa Health Science Network Research Ethics Board (REB protocol 2009848 - 01). The Ottawa Health Science Network Research Ethics Board waived informed consent for the study due to its retrospective, cross-sectional nature, so individualized patient consent was not obtained. All nonanonymized data were stored on secure servers that are accessible only by approved users who signed a privacy and confidentiality form.

Participants

Participants in the study were PCPs who were registered to 1 of the 2 eConsult services and who submitted at least 1 eConsult to the long COVID group within the study timeframe.

eConsult Services

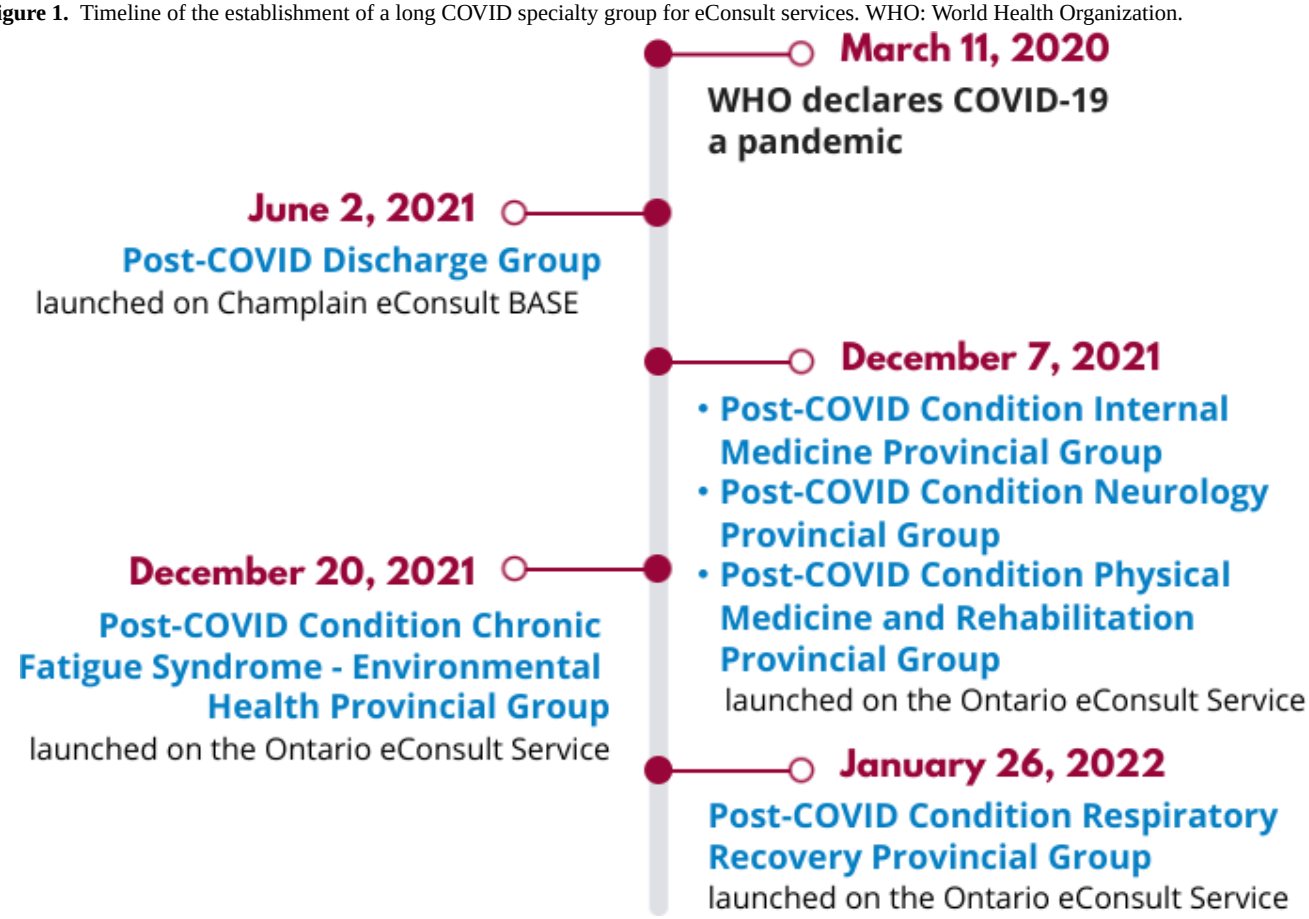
Two multispecialty eConsult services currently operate in Ontario under the umbrella of Ontario eConsult. The first service, the Champlain eConsult BASE (Building Access to Specialists through eConsultation) service, was initially piloted in 2010 across the Champlain region of eastern Ontario, which includes Ottawa and its surrounding communities and has a population of around 1.3 million people. The second service, the Ontario eConsult Service, made eConsult available to all PCPs across Ontario in 2018. Both services are funded through the Ontario Ministry of Health. On June 2, 2021, a long COVID specialty group was established on the Champlain eConsult BASE service. Later in the year, a similar service was made available on the Ontario eConsult Service, providing PCPs with access to five different long COVID specialty subgroups that were rolled out between December 2021 to January 2022: (1) Post-COVID Condition Internal Medicine Provincial Group, (2) Post-COVID Condition Neurology Provincial Group, (3) Post-COVID Condition Phys Med/Rehab Provincial Group, (4) Post-COVID Condition Chronic Fatigue Syndrome—Environ Health Provincial Group, and (5) Post-COVID Condition Respiratory Recovery Provincial Group.

Measures considered when recruiting specialists to eConsult BASE-managed groups include the specialist's experience, level of interest, and ability to maintain an adequate case volume, as well as maintaining equity across and within regions or organizations. The services leveraged existing eConsult specialists and recommendations from peers to identify eligible specialists.

In total, 19 specialists were available to respond to long COVID questions during the study period (5 on the Champlain BASE and 14 on the Ontario eConsult service) from various specialty backgrounds, including internal medicine, infectious disease, neurology, physiatry and respirology (Figure 1).

XSL•FO
RenderX

Figure 1. Timeline of the establishment of a long COVID specialty group for eConsult services. WHO: World Health Organization.



PCPs initiate an eConsult by logging onto a secure web portal and selecting the specific specialty service from a dropdown menu. The PCP fills out a structured electronic form and submits the case along with any pertinent attachments (eg, laboratory or imaging data). The case is assigned to a specialist from the chosen group, who responds within 7 days (the median specialist response time is typically 1 day [5]). When a case is closed, the PCP completes a mandatory close-out questionnaire that gathers information about the case's outcome, whether a referral was originally contemplated and whether it was ultimately avoided as a result of the specialist's advice, and the case's educational value and capacity to serve as material for continuing medical education by using a 5-point Likert scale (Multimedia Appendix 1).

Data Collection and Analysis

We computed descriptive statistics to examine the overall use of the long COVID specialist group from June 1, 2021, to July 31, 2022. Data from both the Ontario eConsult and Champlain eConsult BASE services were analyzed to determine the total number of cases submitted to the long COVID specialty group. Time stamps for each eConsult were assessed to examine the response interval for specialist responses, and the total self-reported time billed was used to determine the amount of time each specialist spent responding to the eConsult.

Following the established methodology, case transcripts of PCP questions from eConsults submitted through the Champlain eConsult BASE between June 1, 2021, and July 31, 2022, were retrospectively reviewed. Case-level data were accessible for Champlain eConsult BASE but not for Ontario eConsult. Two investigators (MQ and JS) reviewed cases retrospectively using 2 validated taxonomies: the International Classification for Primary Care, version 2 and Ely et al's taxonomy of generic clinical questions. The International Classification for Primary Care, version 2 was used to classify clinical topics and patient-presenting symptoms. The taxonomy of generic clinical question was used for question classification. In many cases, a single eConsult case had multiple questions, and patients often presented with multiple symptoms; therefore, each case was not restricted to a single clinical question or presenting complaint or symptom.

We performed a descriptive analysis of all eConsult close-out questionnaires completed by the PCPs. Specifically, we sought to determine the PCP's course of action after the eConsult (eg, face-to-face referral avoided and face-to-face referral still needed) and to gain insight into the nature of the advice given to the PCP (eg, new advice given and advice confirmed course of action initially contemplated).

Results

Overview

A total of 40 PCPs submitted 47 eConsults through Champlain eConsult BASE between June 1, 2021, and July 31, 2022, and 197 PCPs submitted 228 cases through Ontario eConsult between December 1, 2021, and July 31, 2022. These cases comprised 0.2% (n=23,067) of the total number of closed cases submitted to Champlain eConsult BASE and 0.5% (n=45,494)

XSL•FO
RenderX

of all Ontario eConsult closed cases during their respective time frames.

The median time for a specialist to respond to the eConsult was 0.6 (IQR 0.19 - 2.36; mean 1.7, SD 2.29) and the average time the specialist billed in responding to each eConsult was 27.5 (SD 18.53; median 20, IQR 15-45) minutes.

Clinical Content of eConsults

The retrospective taxonomy analysis assessed the clinical questions PCPs asked in the 47 cases submitted through Champlain eConsult BASE. Patients had a median age of 43 (IQR 6-65) and presented 24 unique symptoms across cases (Table 1).

Table . Percentage of patients with reported symptom or finding.

Symptom or finding	Patients (n=47), n (%)
Fatigue	14 (30)
Dyspnea	7 (15)
Altered smell (anosmia, parasomnia)	6 (13)
Cough	6 (13)
Cognitive changes	6 (13)
Chest pain	5 (11)
Palpitations	3 (6)
Altered taste (dysgeusia, ageusia)	3 (6)
Congestion	3 (6)
Alopecia	2 (4)
Myalgia	2 (4)
Rash	2 (4)
Paresthesia	2 (4)
Headache	2 (4)
Tremor	2 (4)
Insomnia	2 (4)
Fever	1 (2)
Eye pain	1 (2)
Tinnitus	1 (2)
Arthralgia	1 (2)
Back pain	1 (2)
Abdominal pain	1 (2)
Edema	1 (2)
Asthenia	1 (2)

The 5 most common symptoms were fatigue (14/47, 30%), dyspnea (7/47, 15%), cough (6/47, 13%), altered sense of smell (ie, anosmia and parosmia; 6/47, 13%), and cognitive changes (6/47, 13%). The five main question categories asked by PCPs were: (1) management of chronic symptoms of COVID-19, (2) need for additional work-up or follow-up testing, (3) community resources to support or manage patients with long COVID, (4) diagnostic clarification, and (5) guidance regarding COVID-19 vaccination (Table 2; sample questions in Multimedia Appendix 2). The most common topics pertained to guidance on the management of chronic symptoms of COVID-19 (22/47 cases, 47%), the need for additional work-up or follow-up testing (20/47 cases, 43%), suggestions for community resources to help support or manage patients with chronic COVID-19 (9/47 cases, 19%), and diagnostic clarification (7/47 cases, 15%).

XSL•FO
RenderX

Table . Content of clinical questions asked of post-COVlD-19 condition specialists by electronic consultation (eConsult).

Content topic	eConsult cases (n=47), n (%)
Management of chronic symptoms of COVID-19	22 (47)
Need for additional work-up or follow-up testing	20 (43)
Community resources to support or manage patients	9 (19)
Diagnostic clarification	7 (15)
Guidance related to COVID-19 vaccination	7 (15)
Necessary to refer to specialist	6 (13)
Interpretation of testing	4 (9)
Other	3 (6)

Outcome of eConsults

PCPs completed the post-eConsult questionnaire for 268 of the 275 cases submitted on both the Champlain eConsult BASE and Ontario eConsult services during the study timeframe. Through eConsult, PCPs were able to confirm a course of action that they had initially contemplated in 35% (95/268) of cases and received advice on a new or additional course of action in 59% (157/268) of cases (Table 3). In 38% (102/268) of cases, PCPs had initially contemplated a face-to-face referral but found it unnecessary after the eConsult, while an actual face-to-face referral was needed in only 24% (64/268) of cases (Table 4).

Table . Impact of electronic consultation (eConsult) on PCP[a] course of action.

PCP survey answer	Responses (n=268), n (%)
I got good advice for a new or additional course of action	157 (59)
I was able to confirm a course of action that I originally had in mind	95 (35)
I got good advice for a new or additional course of action that I am not able to implement[b]	5 (2)
I did not find the advice very useful[c]	5 (2)
Other	6 (2)

[a]PCP: primary care provider.

[b]Response option available only for the Champlain eConsult BASE Service.

[c]Response option available only for the Ontario eConsult Service.

Table . Impact of electronic consultation (eConsult) on need for face-to-face referral.

PCP[a] survey answer	Responses (n=268), n (%)
Referral was originally contemplated but now avoided at this stage	102 (38)
Referral was not originally contemplated and is still not needed	88 (33)
Referral was originally contemplated and is still needed	57 (21)
Referral was not originally contemplated, but eConsult process resulted in a referral being initiated	7 (3)
There was no particular benefit to using eConsult in this case[b]	2 (1)
Other	12 (4)

[a]PCP: primary care provider.

[b]Response option available only for the Ontario eConsult Service.

Discussion

Principal Results

This study demonstrates the considerable potential of an eConsult service to support PCPs navigating the complexity and diversity of long COVID symptoms. By facilitating quick, low-barrier access to specialist advice, eConsult empowered PCPs to provide the best possible care for patients with long COVID while avoiding long wait times for specialist referrals. PCPs were able to get prompt specialist advice from a specialist with a median response time of just over 12 hours. The main question types identified in the study related to (1) the management of chronic symptoms of COVID-19, (2) the need for additional work-up or follow-up testing, (3) community resources to support or manage patients with long COVID, (4)

XSL•FO
RenderX

diagnostic clarification, and (5) guidance regarding COVID-19 vaccination. A face-to-face referral was not needed after the majority of eConsult cases, so patients could be given an expert-guided treatment plan without waiting to see a specialist. To our knowledge, this is the first study to examine an innovative electronic solution to help support PCPs in managing long COVID.

Limitations

This study has several limitations. First, the taxonomy analysis used data from a single health region in Ontario, which may impact the generalizability of our findings. As mentioned above, we only had access to case-level data from the Champlain eConsult BASE service and not the Ontario eConsult service. The Champlain region is a culturally and linguistically diverse region and our findings were externally validated by specialist team members who answered eConsults from across Ontario and confirm the nature of the questions identified in this study are representative of common questions asked across the province. Furthermore, this study only examined eConsults submitted to the long COVID group and it is possible that relevant COVID-19 cases were submitted to other specialty groups such as infectious disease. Finally, neither service collects detailed demographic information about the PCPs using the service, so we were unable to report any of this information in our results section.

Comparison With Prior Work

With studies and data continuing to emerge about long COVID, there remain significant challenges in diagnosing and managing patients with this condition (O'Hare et al [9]). A systematic review conducted by Macpherson et al [10] found patients commonly expressed concerns about the lack of knowledge about long COVID, and often found they received conflicting or inconsistent advice from providers. There is currently no effective validated treatment for long COVID [11]. This challenge with management was evident in our findings, as approximately half of the questions PCPs posed were related to advice on management, and one-fifth inquired about community resources to support with management. Furthermore, 59% of PCPs surveyed reported receiving new advice that changed their management plans.

Recognizing the burden of long COVID, many jurisdictions worldwide have implemented multispecialty long COVID clinics and rehabilitation centers. Previous studies have shown that these rehabilitation programs can have a positive impact on various patient outcomes [12,13], such as increased quality of life, reduced fatigue, improved functional status, and better mental health [13]. Unfortunately, the demand for such clinics is high, with many reporting long wait times. Digital strategies such as eConsult provide a means for quicker access to support patients and initiate management. They also have the potential to reduce demand by allowing the PCP to provide guided care; approximately 40% of PCPs surveyed highlighted that a face-to-face referral was not needed as a result of the advice they received through eConsult. This may lessen the backlog of seeing long COVID specialists by decreasing the number of unnecessary in-person referrals. Additionally, we were able to quickly adapt the eConsult service to implement long COVID groups, whereas setting up specific specialty clinics requires greater coordination, time, and cost.

One of the other challenges in identifying and managing long COVID relates to its diversity of presentation; the Public Health Agency of Canada has reported over 100 different symptoms associated with long COVID. In the 47 eConsults examined in this study, 24 unique symptoms were reported, with the 5 most common being fatigue, dyspnea, cough, altered sense of smell (ie, anosmia and parosmia), and cognitive changes. These findings are in line with other international studies. For example, a UK-based study conducted by Carfi et al [14] also noted fatigue as the most common long COVID symptom (51%), followed by dyspnea (35%), arthralgia (25%), and concentration difficulties (25%). This diversity in presentations can lead to challenges with diagnosis, which is particularly important as the symptoms of long COVID overlap with other serious and possibly life-threatening complications associated with COVID-19, such as pulmonary embolism, myocarditis, and organizing pneumonia. Several publications have reported delays in the diagnosis of life-threatening conditions such as pulmonary embolism as they were mistaken for symptoms of COVID-19 (Yousefazai and Bhimaraj [15]; Melazzini et al [16]). In total, 15% (7/47) of PCP questions related to diagnostic clarification, while 43% (20/47) of eConsults asked about the need for additional testing to rule out other conditions. This highlights the importance of getting timely advice to confirm the diagnosis and avoid delays in the diagnosis of possibly life-threatening conditions. The diversity of presentations also demonstrates the need for a multidisciplinary approach to providing support for patients with long COVID. To address this need, Ontario eConsult established 5 unique subgroups that PCPs can select (ie, internal medicine, neurology, physical medicine or rehabilitation, chronic fatigue syndrome, and respirology) involving providers from various specialties to better support PCPs and their patients in getting helpful and appropriate advice in a timely manner.

Conclusions

The long COVID groups available through the Champlain eConsult BASE and Ontario eConsult services provided rapid access to a multispecialty service that facilitated the avoidance of unnecessary face-to-face referrals. Long COVID is a multisystemic condition that is often debilitating, with a significant impact on one's quality of life and mental health. Given the lack of knowledge around long COVID, limited timely access to specialized long COVID clinics, and the possibility of delayed diagnosis of life-threatening conditions associated with COVID-19 (eg, pulmonary embolism), there is an urgent need for innovative digital solutions such as eConsult to better support PCPs to provide patients with timely access to specialty advice. An assessment of eConsults highlighted 5 common question types, providing insight into potential gaps in knowledge among PCPs that could help guide medical education and policy.

XSL•FO
RenderX

Acknowledgments

This research received funding from Ontario Health.

Conflicts of Interest

CL and EK are the cofounders of the eConsultBASE service but have no commercial interest in the service. EK is the executive director of the Ontario eConsult Centre of Excellence and receives salary support from Ontario Health. EK completes occasional electronic consultations (eConsults) as a specialist through the service for which she is reimbursed. CL is the director of evaluation for the Ontario eConsult Centre and receives salary support from Ontario Health.

Multimedia Appendix 1
Mandatory close-out survey.
[DOCX File, 13 KB - humanfactors_v12i1e58582_app1.docx]

Multimedia Appendix 2
Sample questions asked of post-COVlD-19 condition specialists.
[DOCX File, 19 KB - humanfactors_v12i1e58582_app2.docx]

References

1. COVID-19 dashboard. World Health Organization. URL: https://covid19.who.int/ [accessed 2025-02-20]
2. Coronavirus disease (COVID-19): post COVID-19 condition. World Health Organization. 2023. URL: https://www.who.int/news-room/questions-and-answers/item/coronavirus-disease-(covid-19)-post-covid-19-condition [accessed 2025-02-20]
3. Long-term symptoms in canadian adults who tested positive for COVID-19 or suspected an infection, january 2020 to august 2022. Statistics Canada. 2022. URL: https://www150.statcan.gc.ca/n1/daily-quotidien/221017/dq221017b-eng.htm [accessed 2025-02-20]
4. Singh J, Garber GE, Keely E, Guglani S, Liddy C. Evaluation of an electronic consultation service for COVID-19 care. Ann Fam Med 2022;20(3):220-226. [doi: 10.1370/afm.2807] [Medline: 35606132]
5. Liddy C, Mitchell R, Guglani S, et al. The provincial spread and scale of the Ontario eConsult service: evaluation of the first 2 years. Ann Fam Med 2022;20(3):262-265. [doi: 10.1370/afm.2812] [Medline: 35606133]
6. Liddy C, Drosinis P, Armstrong CD, McKellips F, Afkham A, Keely E. Improving access to allied health professionals through the Champlain BASE. BMJ Open 2016;6(6):e010920. [doi: 10.1136/bmjopen-2015-010920]
7. Keely E, Williams R, Epstein G, Afkham A, Liddy C. Specialist perspectives on Ontario provincial electronic consultation services. Telemed e-Health 2019 Jan;25(1):3-10. [doi: 10.1089/tmj.2018.0012]
8. Liddy C, Sethuram C, Mihan A, Moroz I, Keely E. Primary care providers' perspectives on the Ontario eConsult program. Telemed J E Health 2021;27(9):1039-1045. [doi: 10.1089/tmj.2020.0338]
9. O'Hare AM, Vig EK, Iwashyna TJ, et al. Complexity and challenges of the clinical diagnosis and management of long COVID. JAMA Netw Open 2022 Nov 1;5(11):e2240332. [doi: 10.1001/jamanetworkopen.2022.40332] [Medline: 36326761]
10. Macpherson K, Cooper K, Harbour J, Mahal D, Miller C, Nairn M. Experiences of living with long COVID and of accessing healthcare services: a qualitative systematic review. BMJ Open 2022 Jan 11;12(1):e050979. [doi: 10.1136/bmjopen-2021-050979] [Medline: 35017239]
11. Davis HE, McCorkell L, Vogel JM, Topol EJ. Long COVID: major findings, mechanisms and recommendations. Nat Rev Microbiol 2023 Mar;21(3):133-146. [doi: 10.1038/s41579-022-00846-2] [Medline: 36639608]
12. Smith JL, Deighton K, Innes AQ, et al. Improved clinical outcomes in response to a 12-week blended digital and community-based long-COVID-19 rehabilitation programme. Front Med (Lausanne) 2023;10:1149922. [doi: 10.3389/fmed.2023.1149922] [Medline: 37293307]
13. Jimeno‐Almazán A, Franco‐López F, Buendía‐Romero Á, et al. Rehabilitation for post‐COVID‐19 condition through a supervised exercise intervention: a randomized controlled trial. Scand Med Sci Sports 2022 Dec;32(12):1791-1801. [doi: 10.1111/sms.14240] [Medline: 36111386]
14. Carfì A, Bernabei R, Landi F, Gemelli Against COVID-19 Post-Acute Care Study Group. Persistent symptoms in patients after acute COVID-19. JAMA 2020 Aug 11;324(6):603-605. [doi: 10.1001/jama.2020.12603] [Medline: 32644129]
15. Yousefzai R, Bhimaraj A. Misdiagnosis in the COVID-19 Era: when zebras are everywhere, don't forget the horses. JACC Case Rep 2020 Aug;2(10):1614-1619. [doi: 10.1016/j.jaccas.2020.04.018] [Medline: 32342051]
16. Melazzini F, Reduzzi M, Quaglini S, Fumoso F, Lenti MV, Di Sabatino A. Diagnostic delay of pulmonary embolism in COVID-19 patients. Front Med (Lausanne) 2021;8:637375. [doi: 10.3389/fmed.2021.637375] [Medline: 33996852]

Abbreviations

eConsult: electronic consultation
PCP : primary care provider

XSL•FO
RenderX

Edited by A Kushniruk; submitted 21.03.24; peer-reviewed by K Peeters; revised version received 20.12.24; accepted 15.01.25; published 28.02.25.

Please cite as:
Singh J, Quon M, Goulet D, Keely E, Liddy C
The Utilization of Electronic Consultations (eConsults) to Address Emerging Questions Related to Long COVID-19 in Ontario, Canada: Mixed Methods Analysis
JMIR Hum Factors 2025;12:e58582
URL: https://humanfactors.jmir.org/2025/1/e58582
doi:10.2196/58582

© Jatinderpreet Singh, Michael Quon, Danica Goulet, Erin Keely, Clare Liddy. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 28.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Understanding Appropriation of Digital Self-Monitoring Tools in Mental Health Care: Qualitative Analysis

Lena de Thurah[1], MSc; Glenn Kiekens[1,2,3], PhD; Jeroen Weermeijer[1], PhD; Lotte Uyttebroek[1], MSc; Martien Wampers[1], PhD; Rafaël Bonnier[1], MSc; Inez Myin-Germeys[1], PhD

[1]Department of Neurosciences, Center for Contextual Psychiatry, KU Leuven, Herestraat 49 ON5B, bus 1029, Leuven, Belgium

[2]Research Unit of Clinical Psychology, Faculty of Psychology and Educational Sciences, KU Leuven, Leuven, Belgium

[3]Department of Medical and Clinical Psychology, Tilburg University, Tilburg, Netherlands

Corresponding Author:
Lena de Thurah, MSc
Department of Neurosciences, Center for Contextual Psychiatry, KU Leuven, Herestraat 49 ON5B, bus 1029, Leuven, Belgium

Abstract

Background: Digital self-monitoring tools, such as the experience sampling method (ESM), enable individuals to collect detailed information about their mental health and daily life context and may help guide and support person-centered mental health care. However, similar to many digital interventions, the ESM struggles to move from research to clinical integration. To guide the implementation of self-monitoring tools in mental health care, it is important to understand why and how clinicians and clients adopted, adapted, and incorporated these tools in practice.

Objective: Therefore, this study examined how clinicians and clients within a psychiatric center appropriated an ESM-based self-monitoring tool within their therapy.

Methods: Twelve clinicians and 24 clients participated in the piloting of the ESM tool, IMPROVE. After utilizing the tool, 7 clinicians and 11 clients took part in semistructured interviews. A thematic framework analysis was performed focusing on participants' prior knowledge and expectations, actual use in practice, and potential future use of ESM tools.

Results: Many participants experienced that the ESM tool provided useful information about clients' mental health, especially when clinicians and clients engaged in collaborative data interpretation. However, clinicians experienced several mismatches between system usability and their technical competencies, and many clients found it difficult to comply with the self-assessments. Importantly, most participants wanted to use digital self-monitoring tools in the future.

Conclusions: Clinicians' and clients' choice to adopt and integrate self-monitoring tools in their practice seems to depend upon the perceived balance between the added benefits and the effort required to achieve them. Enhancing user support or redesigning ESM tools to reduce workload and data burden could help overcome implementation barriers. Future research should involve end users in the development of ESM self-monitoring tools for mental health care and further investigate the perspectives of nonadopters.

(JMIR Hum Factors 2025;12:e60096) doi:10.2196/60096

KEYWORDS

digital self-monitoring; technology appropriation; experience sampling method; mental health care; mental health; self-monitoring; digital health; adoption; implementation; thematic; usability; interview; experience; attitude; opinion; perception; perspective; acceptance

Introduction

To improve access to and facilitate person-centered mental health care, digital technologies are increasingly being deployed to collect and share health-related data, deliver care, and support individuals in managing their health [1,2]. The experience sampling method (ESM; also termed ecological momentary assessment) is a structured diary technique that enables individuals to collect detailed information about their mental health in their daily lives [3]. Using smartphone apps, the ESM prompts individuals to complete brief self-assessments multiple times daily, assessing their emotional states, behaviors, and psychological symptoms. By allowing individuals to collect this information and share it with their clinicians, the ESM can ensure that therapeutic decisions are aligned with clients' everyday experiences and needs. For example, by providing insights into which activities an individual commonly engages in, and how they respond to different daily life situations and stressors, clients and clinicians can make shared decisions about the focus of treatment [4].

Despite its potential, the ESM, similar to many other digital mental health interventions [5,6], struggles to move from

XSL•FO
RenderX

piloting and trialing to actual clinical implementation. To address this research-to-practice gap, scholars have made calls to adopt user-centered design strategies for developing digital mental health tools [7]. User-centered design emphasizes the importance of conducting in-depth analyses of users' goals, needs, and context of use to inform technology design [8]. Nonetheless, most evaluations of digital health technologies focus primarily on quantifiable, technical aspects such as system performance, usability scores, and cost benefits [9,10]. However, an often overlooked but crucial step in implementing the ESM and self-monitoring tools is understanding how and why people integrate these technologies into their daily practices [11]. For example, what are the goals people hope technologies will help them achieve, what difficulties do individuals encounter when using new technologies, and what strategies do they apply to overcome technology use obstacles?

Studying the situated use of the ESM can provide essential knowledge into how users appropriate and make sense of these tools [12,13]. Technology appropriation can be examined in three main stages [13]. The first stage involves users' prior knowledge and expectations, which determine their initial willingness to adopt the technology. Expectations about potential benefits (eg, simplifying tasks) and the effort required (eg, time spent on training) are particularly crucial in shaping the initial interest [14,15]. In the second phase, users explore and evaluate the technology's capabilities, while becoming familiar with its functions. They adapt their practices to integrate the technology (eg, maintaining an internet connection), and adapt the technology to suit their needs (eg, disabling certain features or altering settings) [13]. They will test the technology's applicability in different situations and determine what features are most useful to them. Finally, if users find that the technology has added value and helps them achieve their goals and tasks, they might develop new routines and workflows that allow them to integrate and use the technology in their daily practice, thus moving into the third stage of persistent use [13].

To date, the appropriation of ESM tools in health care has largely been overlooked. While one study has examined how individuals with back pain appropriated ESM tools for pain management [16], there is a lack of research in the field of mental health care. As a consequence, we have a limited understanding of how individuals with mental illnesses and their clinicians interact with and make sense of digital ESM tools, and what might lead them to adopt or abandon these technologies. However, understanding this is vital for facilitating successful implementation [15,17]. To address this knowledge gap, we undertook a pilot implementation study, using the ESM tool "IMPROVE," a clinical prototype tool informed by research examining clinicians' and clients' design preferences [18,19]. IMPROVE was used as a part of therapy following a 3-step intervention, in which clinicians and their clients could (1) personalize the tool to clients' specific problems and situations, (2) self-monitor clients' mental health and daily activities via an app, and (3) review the collected data in summarized graphs in an online dashboard. To inform the further development and implementation of clinical ESM tools, this paper aimed to evaluate participants' (1) prior knowledge and expectations, (2) actual use in practice, and (3) potential future integration of ESM-based self-monitoring tools. Addressing these gaps in the literature will provide valuable insights into why and how clients and clinicians in mental health care choose to adopt or abandon digital self-monitoring tools in therapy.

Methods

Ethical Considerations

All participants provided written informed consent, and all procedures were approved by the medical ethics committee of KU Leuven (S64244). Participants were given a study ID to ensure anonymity, and interviews were pseudonymized during the transcription by removing personally identifiable information and replacing them with pseudo codes. Finally, participants were allowed to freely use the IMPROVE tool after completion of the study but were not given additional compensation.

Study Design and Recruitment

We conducted a pilot study within the University Psychiatric Center KU Leuven in Belgium in which we involved clients and clinicians in testing and evaluating IMPROVE. The research team was unable to access the clinic directly, due to COVID-19 restrictions. Therefore, invitations to participate were emailed to all psychiatrists and psychologists affiliated with the psychiatric center (N=142). The study aimed to enroll 12 clinicians, and initial invitations were followed by up to 2 reminder emails. For those interested, online informational sessions with the research team were organized. Participating clinicians were asked to recruit at least 1 client from their practice. By allowing clinicians to select clients, the referred sample provided a practical representation of the clients with whom clinicians were likely to use the tool within a real-life context.

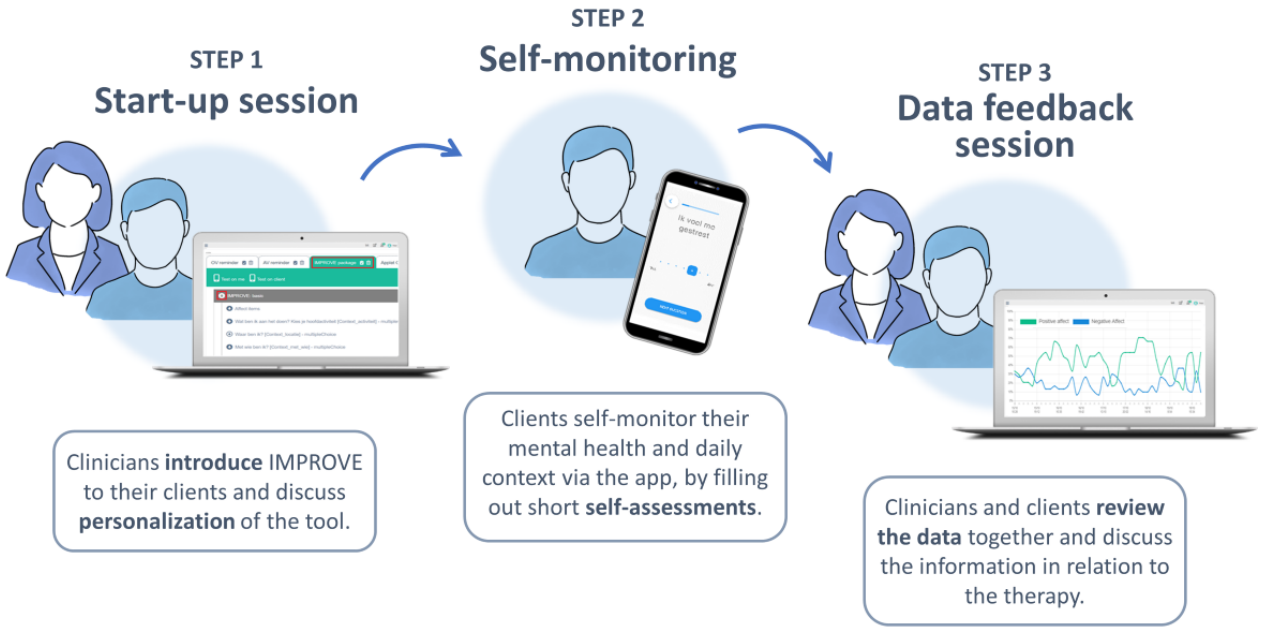
The inclusion criteria were kept broad to reflect the diverse reality of clinical practice. To be eligible, clinicians needed to be certified mental health professionals and proficient in Dutch, which was the language the intervention was designed in. Clients had to be 18 years or older and also proficient in Dutch. To replicate real-world implementation, clients were asked to use their own smartphones. As a result, only clients owning a smartphone with at least 3G coverage were eligible to participate.

IMPROVE was run on the digital platform m-Path [20]. To enable clinicians to use IMPROVE, they received a manual with instructions on using the tool in their practice (Multimedia Appendix 1: IMPROVE training manual) and were invited to join an online training session. Clinicians and their clients were then requested to complete the 3-step intervention (Figure 1). In the first step, clinicians introduced IMPROVE to their clients and assisted them in downloading the app on their smartphones. They also discussed personalization of the tool, such as adding symptom-specific questions or adjusting the notification schedule. In the second step, clients self-monitored their mental health for 6 consecutive days. They received 10 semirandom notifications daily, prompting them to complete brief self-assessments on mood and context. Additionally, they received a morning notification to assess their sleep and an evening notification to evaluate the day. In the third step,

XSL•FO
RenderX

clinicians and clients reviewed the collected data together, exploring factors such as clients' activities, mood variability, and mood-context interactions.

Figure 1. Overview of the steps included in the IMPROVE intervention.



Data Collection and Analysis

Due to COVID-19 measures, data were collected remotely via online surveys in RedCap (Research Electronic Data Capture, Vanderbilt University) [21] and through Skype interviews (Version MSO; Skype Technologies). In some cases, poor connections affected the quality of the interview recordings; however, the overall data collection was not hindered by remote methods. Demographic information was collected at study enrollment and participants who completed the intervention were invited to participate in a semistructured interview (Multimedia Appendix 2: Interview guides). Interviews were audio recorded and transcribed verbatim. We performed a thematic analysis [22] in which we first conducted top-down content-coding of the transcripts using a thematic framework that identified interview material related to clinicians' and clients': (1) previous knowledge and expectations, (2) actual use in practice, and (3) potential future integration of the IMPROVE tool (Table 1). Hereafter, we performed additional inductive in vivo coding and used pattern and focused coding to create labels for the in vivo codes and compiled them into additional subthemes [23]. LdT undertook all primary and secondary coding, which was then revised by the coauthors and discussed in multiple peer debriefing sessions [24]. After each session, the coding categorizations and labels were modified and refined based on feedback from coauthors. To explore the potential impact of the different themes on the discontinued use of IMPROVE, additional information was extracted from informal participant contact records kept by the research team.

Table . Thematic framework for top-down coding of interviews.

Theme	Description
Prior knowledge and expectations	Participants talk about any prior experiences or knowledge they have regarding the ESM[a] or self-monitoring tools (analog or digital). Participants talk about their motivations and expectations related to using the ESM or self-monitoring tools and participating in the study.
Actual use in practice	Participants talk about how they used the IMPROVE tool and how they judged these experiences. Mentioning eg, what purpose they used it for, when, and how often they used it.
Potential future integration	Participants talk about their interest in using the IMPROVE tool (or similar tools) in the future. Participants talk about potential changes that they find would be relevant for future use and integration of the IMPROVE tool.

[a]ESM: experience sampling method.

XSL•FO
RenderX

Results

Participant Characteristics

Nineteen clinicians expressed interest in participating in the study, of whom 12 were enrolled. While a systematic record of the number of clients approached by clinicians was not maintained, at least 29 clients were invited, with 24 agreeing to participate. Of the enrolled participants, 8/12 clinicians (67%) and 17/24 clients (71%) completed the intervention; among these, 7/8 clinicians (88%) and 11/17 clients (65%) subsequently participated in an interview. A demographic summary of interview participants can be consulted in Multimedia Appendix 3: Demographic summary.

Thematic Analysis

A complete overview of the themes and subthemes of the analysis and how frequently they were mentioned by participants can be found in Multimedia Appendix 4: Overview of themes. Tabulated in vivo code summaries of participants' individual experiences can be found in Multimedia Appendix 5: Individually summarized experiences. Additionally, illustrative summaries are displayed in text boxes.

Previous Experience and Expectations

Individuals' experiences and expectations are important indicators of their initial willingness to adopt technologies. While only 1 clinician had experience using digital self-monitoring tools, most clinicians and some clients had experience using analog self-monitoring or diary techniques. Some clinicians also had experience with different forms of digital tools within their practice (eg, online training platforms, video consultations, and virtual reality). Similarly, several clients had experience using mental health apps, eg for practicing breathing techniques and cognitive-behavioral therapy exercises. Despite their limited experience, clinicians generally expected that IMPROVE would offer benefits over analog registration methods, allowing them to more easily collect and summarize information about their clients. However, several clinicians also anticipated that there would be challenges and limitations for using the tool. The most commonly mentioned was that low digital literacy could be a barrier for both clients and clinicians (Clinician 0100 Textbox 1). Clinicians' motivation for testing the tool was primarily driven by curiosity and the conviction that they need to master digital tools, as these are becoming increasingly dominant in the health care sector. Interestingly, most clients were primarily motivated by a wish to help and contribute to research but also expressed a curiosity toward what digital mental health tools might offer (Client 0501 Textbox 2). Some clients hoped that IMPROVE would allow them and their clinicians to get a better understanding of their mental health problems.

XSL•FO
RenderX

Textbox 1. "Might use digital self-monitoring tools in the future"– selected participants' experiences.

CLINICIAN 0100:

Prior knowledge and expectations

- I had no experience with digital self-monitoring, but I am currently testing VR technology in therapy. I have previously worked with analog self-registration in therapy.
- I expected it would be technically challenging to work with the tool.
- I want to try using new methods in therapy and I was eager to try the tool.

Actual use in practice

- It was difficult to find things in the dashboard and I needed help from the research team to use the tool. After a while using the platform got easier, but it required a lot of time to get started.
- It was difficult to draw conclusions about my clients' data, but I asked my client for clarifications when I was not able to interpret their data.

Potential future integration

- The tool gave me more information about what happens in clients' lives.
- I might use the tool again if it is made easier to use.

CLIENT 0503:

Prior knowledge and expectations

- I had no experience with self-monitoring apps.
- I hoped to gain more insight into my mental health.

Actual use in practice

- Assessment frequency was high, but okay for 1 week.
- I think I responded to almost all notifications, but I sometimes missed notifications because I forgot my phone. Also, when I was not feeling well I didn't respond to the notifications.
- Identifying, labeling, and scoring emotions on a scale are difficult.
- My therapist had difficulties operating the dashboard, so we didn't go into detail with the data.

Potential future integration

- I might use the tool again, but it is tiring to do for a long time.
- I prefer using the tool as a part of therapy, but I would like to have access to the data myself.

CLIENT 0802:

Prior knowledge and expectations

- I have no experience with health apps, but I use analog methods to keep track of my mental health.
- Digital tools allow you to do more than analog tools, so I thought it would be interesting to try.

Actual use in practice

- The number of notifications was okay, but I sometimes missed notifications because I was working or sleeping.
- Using the tool during the therapy session made it easier to recollect things that had happened.

Potential future integration

- I might consider using the tool again.

XSL•FO
RenderX

Textbox 2. "Not likely to use digital self-monitoring tools in the future"– selected participants' experiences.

CLIENT 0501:

Prior knowledge and expectations

- I had no experience with self-monitoring apps.
- I was hesitant to participate because I do not like self-assessments.
- I thought it was worth trying and I wanted to contribute to improving mental health care.

Actual use in practice

- It was difficult always to remember to have my phone with me.
- Notifications were sometimes disturbing and the assessment frequency was too high.
- It was important to me that I did not miss notifications; therefore, fear of missing notifications would sometimes stress me.
- Self-reflection was sometimes difficult and confronting.
- It was interesting to look at my data.

Potential future integration

- I don't find it likely that I will use the tool again.

Actual Use in Practice

Participants' willingness to test the IMPROVE tool was influenced by their initial expectations that digital self-monitoring would provide them with benefits and advantages. Below we describe how clinicians and clients used the tool in practice and how they evaluated its capabilities and contextual fit.

Step 1: Start-Up Session

Clinicians generally found that the time investment required to learn to use IMPROVE was too much considering the time they had available. Experiencing time constraints was also one of the most common reasons reported by clinicians who did not complete the intervention. Personalizing clients' questionnaires in particular asked for extra time and effort from clinicians. Despite emphasizing that the personalization of the IMPROVE questionnaires made the tool more relevant, clinicians made limited use of the available personalization options. Most clients also expressed that personalization of the tool is desirable but that clinicians did not discuss this with them.

Furthermore, while most clinicians were convinced that using the IMPROVE tool would become easier with practice, many clinicians indicated that the complexity of the tool was too high. In particular, clinicians found that navigating the dashboard was not intuitive and that it was difficult to find things and set up the clients' questionnaires. Thus, many experienced that they did not have adequate competencies to make full use of the tool; "I constantly felt I was doing something wrong," 1 clinician explained. However, at the same time, clinicians made limited use of the training manual and expressed a wish for more in-person support. This was also reflected by several clinicians contacting the research team for additional support after the initial training (Clinician 0100,Textbox 1). Interestingly, some clinicians resolved their need for support by organizing training sessions with colleagues.

Step 2: Self-Monitoring

The majority of clients reported that they made an effort to comply with the notifications, but many found that the frequency of the self-assessments (10 notifications per day) was too high. A few clients indicated that they would consciously skip assessments because they did not consider it important to respond to all notifications. Conversely, some clients also mentioned that missing notifications would make them feel guilty and annoyed, or that the thought of missing assessments made them nervous and stressed (Client 0501, Textbox 2). Although clients reported that completing the assessments did not take them long, many experienced difficulties responding to the notifications within the fixed 15-minute response window. Several reasons for this were voiced, of which being occupied with work and other daily activities was the most frequent (Client 0802, Textbox 1; Client 0601, Textbox 3). Having an irregular day rhythm or sleep schedule that did not match the notification schedule, needing to carry one's phone and ensure an internet connection, and feeling unwell or tired were also factors that made complying difficult. Importantly, clients who dropped out of the study also reported a lack of time, difficulties complying with the assessments, notifications stress, and feeling unwell as reasons for not completing the intervention.

XSL•FO
RenderX

Textbox 3. "Likely to use digital self-monitoring tools in the future"– selected participants' experiences.

CLINICIAN 0600:

Prior knowledge and expectations

- I have no experience with digital self-monitoring, but I have used analog self-monitoring techniques in therapy.
- I think digital tools are easier to use than analog self-monitoring methods. Digital tools are used more and more, and we need to offer clients tools that can help them achieve their goals.

Actual use in practice

- It takes time and effort to learn to use the tool.
- I didn't personalize my client questionnaire, because I wanted to start with the basics.
- I monitored my clients' responses and contacted them if I could see they weren't responding.
- There was a lot of data; I started with simpler visualizations and gradually added things. If I didn't know the client beforehand, I would not be able to make sense of the data.

Potential future integration

- The tool provided more details and an overview and allowed us to go more in-depth with problems that we already knew were there. We used the tool to identify what was important for the client.
- I will certainly consider using this tool again with my clients.

CLIENT 0601:

Prior knowledge and expectations

- I had no experience with health apps.
- I had no specific expectations and I'm generally skeptical about health apps, but I wanted to help research.

Actual use in practice

- I responded to as many notifications as I could, but I missed a lot because I was busy. I didn't consider it a problem that I missed some notifications. I also had to remember to bring my phone and connect it to 4G.
- It is sometimes difficult to assess whether your mood changed and how much.
- I was amazed to see what came out of the data. My therapist was able to do a lot with the data, despite my low compliance.

Potential future integration

- The tool provided useful information and allowed me to get to know myself better.
- I would be interested in using the tool again, but I would like to have access to the data myself.

CLINICIAN 1200:

Prior knowledge and expectations

- I used digital self-monitoring tools before in therapy.
- Digital self-monitoring tools are easier to use and give better insight into clients' lives than other methods. However, self-monitoring can be difficult for some clients.

Actual use in practice

- Setting up the tool took a lot of time.
- It was interesting to look at the data, but the many visualizations were somewhat overwhelming.

Potential future integration

- The tool quickly provides you with an overview of how your client is doing.
- I would like to use the tool again, but it should be made less burdensome for clients. It would be interesting to monitor clients for a longer period to evaluate their progress and the effect of treatments.

CLIENT 1201:

Prior knowledge and expectations

- I had no experience with mental health apps.

XSL•FO
RenderX

- I expected the tool would make it easier for my therapist to understand my problems.
- I think there is a need for more mHealth in mental health care and that this tool might help me and others.

Actual use in practice

- There are too many notifications and they are sometimes disturbing. People should be allowed to snooze notifications if they are busy.
- I did my best to respond to the notifications. On good days, it is okay to complete the assessments, but on bad days, it is difficult.
- My therapist identified moments when I was feeling bad and tried to understand these.

Potential future integration

- The tool can make you aware of what you need to work on, and it makes it easier to monitor the effects of your treatment.
- I would like to keep using the tool to check how I am doing.

Several clients found the practice of labeling and rating their emotions difficult. Some indicated that they were unsure whether they completed the self-assessments correctly and expressed a need for more guidance from their clinicians on how to respond to the assessments. Several clients also expressed a wish for more open questions that would allow them to describe their experiences in more detail, as they did not find that the default questions were sufficiently able to capture their experiences. Interestingly, while clinicians were not instructed to do so, some monitored their clients' responses during the self-monitoring week as a form of remote monitoring of clients (Clinician 0600, Textbox 3).

Step 3: Data Feedback Session

To get the most out of the data feedback session, most clinicians reviewed their clients' data in preparation for the session. Many felt that preparation was necessary for them to understand the data and be able to discuss it with clients. Many clinicians found it challenging to interpret and navigate the graphs, which they often perceived as overwhelming. This generally led clinicians to be selective in the data they would review and discuss with clients. All clinicians reviewed the data with their clients and some also encouraged their clients to give their interpretation of the data or to provide further clarification (Clinician 0100, Textbox 1*)*. Similarly, clients indicated that the support of a clinician was crucial for them to make sense of their data. While most clients reported that they could understand the graphs, they also indicated relying on their clinician to help them interpret the data. These findings indicate that, despite challenges, most clinicians could provide useful feedback to their clients based on the self-monitoring data.

Potential Future Integration

Despite encountering challenges, both clinicians and clients described that the IMPROVE tool provided them with useful insights. Next, we will explore how these experiences influenced participants' willingness to integrate and use digital self-monitoring tools in their future practice. Clinicians and clients were generally open to using digital self-monitoring tools again. Participants who expressed the greatest interest in continued use more often experienced an added value of using IMPROVE (see Textbox 3). This included having access to more detailed information about clients' mental health, which helped create an overview and clarity, and identify focus points to discuss in therapy. Clinicians who had initial positive expectations and believed digital tools were important in supporting mental health care were also more willing to use digital self-monitoring again. The clinicians who were more hesitant reported more difficulties using IMPROVE and expressed a greater need for support (see Textbox 1).

Clients who expressed a low interest in using digital self-monitoring tools in the future also reported negative reactivity in the form of stress and increased negative emotions during self-monitoring (see Textbox 2). Interestingly, all these clients had high compliance during the self-monitoring. However, clients overall reported that the assessment frequency should be lowered to make future self-monitoring less burdensome. Moreover, some clients and clinicians expressed an interest in using the tool for extended periods to monitor progress and the effects of therapy. Finally, clients and clinicians also expressed a desire to adapt the self-assessment further and monitor factors more specific to the individual client.

Discussion

Principal Findings and Implications

To guide the successful implementation of digital self-monitoring tools in mental health care, it is pivotal to understand how people react, respond, and adapt to these technologies. We examined how clinicians and clients in a psychiatric center appropriated the ESM-based self-monitoring tool IMPROVE within their therapeutic practices. In line with existing theories, our results indicate that clinicians' and clients' willingness to adopt and integrate digital self-monitoring tools into their practice depend on the perceived balance between added benefits and the effort necessary to accomplish these [17]. The main benefits identified in our study included better information about clients' mental health, while system usability and assessment burden demanded extra effort from participants. Below we discuss the potential implications of this for the implementation of the ESM tools in clinical practice.

While some struggled more than others, all interviewed participants managed to reach a basic level of use of the IMPROVE tool within the intervention period. Furthermore, we observed a great interest and curiosity towards digital mental health tools, especially among clinicians. Many clinicians experienced that using IMPROVE added value and generated more clarity and focus in the therapy, by providing more detailed

XSL•FO
RenderX

information about the clients' mental health. Several clients also experienced these benefits. Our findings suggest that a collaborative effort between clients and clinicians is key to maximizing the experienced benefits of the ESM tools. While many clinicians struggled to construct a frame for the interpretation of clients' data, several turned to their clients to help make sense of the data. Clients also stressed the importance of having a clinician support them in interpreting their data and doubted whether they would be able to do this themselves. Other studies similarly highlighted the value of patients providing clinicians with additional information about patient-generated health data to unveil the subjective meanings of the data [25]. This indicates that collaborative data interpretation is an essential component in the clinical application of the ESM tools, which boosts the perceived usefulness of the tool. Therefore this component should be emphasized and strengthened in future implementation initiatives.

Most participants in our study expressed willingness to use the ESM or similar tools in the future, although many emphasized the necessity of making changes and improvements to the tool. This aligns with previous findings that clinicians and clients are generally interested in the ESM-based self-monitoring [18,19,26]. However, the need for further adaption was reflected by participants encountering several challenges in using and integrating IMPROVE into their practice. Many clinicians experienced difficulties navigating the tool's functionalities, which made it difficult for them to start using the tool and apply more advanced features, such as personalizing clients' questionnaires. Clinicians, therefore, expressed a need for more in-person support. Unfortunately, these types of usability issues are common in digital health tools [27]. A study examining usability problems of eHealth applications found that system navigation, interface design, and lack of built-in guidance and support accounted for 69% of users' usability issues [27]. Furthermore, consistent with other research [28,29], our study indicated that technology literacy and attitude can influence willingness to use digital tools. More anthropological field research on mental health care workers' existing technology habits and literacy might help inform the system design of the ESM tools and understand users' context and capabilities. Similarly, engaging end-users in co-design processes could help tackle usability issues that could lead users to abandon the tools. Finally, our results show that adequate support for clinicians and clients is needed to facilitate the implementation of the ESM tools in mental health care. Especially in the early adoption phases, users should be supported in familiarizing themselves with the tools and integrating them into their work routines. Potential solutions could include more built-in guidance functions in the tool or establishing additional structures (eg, service centers) that can provide direct user support.

Furthermore, our findings emphasize that self-monitoring demands a lot of clients, and can be difficult and burdensome for people with mental health problems. While clients expressed a wish to comply with the self-assessments, they often found it practically challenging. In line with other self-monitoring studies [30], we found that competing activities (eg, working) and technical issues (eg, device or internet access) were the most common reasons for missing assessments. To comply with the assessments, several clients had to change their phone habits, eg, ensuring that they always had their phone with them, and had notifications and internet connection on. As it is known that existing habits are important predictors of technology acceptance [17], this need to change habits can become a threat for sustained clinical implementation of ESM tools. Furthermore, some clients reported negative reactivity to self-monitoring, which was associated with less interest in using the tool again. Other studies also found that while self-monitoring can be motivating and helpful, it can also become a stressful activity that clients feel obliged to comply with [31]. This highlights the need to investigate how self-monitoring tools can be made less burdensome for users while still producing valuable information. Potential solutions could include the integration of passive monitoring [32,33] or adaptive assessment schemes that allow for periods with lower and higher assessment intensity [34,35].

Limitations

The findings of the study should be interpreted considering the following limitations. One limitation is that the study relied on referral sampling, which might have influenced the representativeness and diversity of the sample. Furthermore, the study was conducted in a large psychiatric center of a university in Belgium. Clients and clinicians within different health care settings or with different demographic backgrounds might have different experiences of using digital self-monitoring tools. Another limitation is the potential overrepresentation of people with a positive attitude towards digital mental health tools. Several clinicians for example recruited clients who they expected were interested in using digital tools and had the necessary technical literacy. Furthermore, none of the participants who dropped out of the intervention participated in the interviews. However, using contact records data, we established that some of the challenges experienced by participants who were interviewed were also reported as reasons for dropout by participants who did not complete the intervention. Future research should aim to better understand the views of nonadoptors, identify the main reasons for nonadoption, and work towards tackling these barriers in the design and implementation of ESM self-monitoring tools.

Conclusion

ESM-based self-monitoring can provide clinicians and clients with useful information about how clients' daily life activities correspond to fluctuations in their mental health. These benefits seem to increase when clinicians and clients engage in collaborative interpretation and sense-making of the clients' data. Therefore, ESM tools have a clear potential to support a person-centered approach to mental health care rooted in clients' daily life experiences. However, our study highlights that the effort required by clinicians and clients to integrate ESM tools into daily practice remains substantial due to system usability challenges and the burden of repeated assessments. Addressing these issues in future ESM tool developments will be crucial for successful implementation. One potential solution to tackle these barriers is enhancing user support. Another is modifying key features of ESM tools to better suit users and their contexts, such as reducing the frequency of assessment and data information load. Engaging end-users directly in this process

XSL•FO
RenderX

through user-centered design approaches may ensure a better fit between the tools, their context of use, and user goals and ultimately facilitate better adoption.

Acknowledgments

The authors would like to thank all clinicians and clients for participating in the study and contributing to this work. We also thank Merijn Mestdagh and Stijn Verdonck for programming the necessary software functionalities on the m-Path software platform. In addition, the authors would like to thank Silke Apers, Daphne Tuyaerts, and Tessa Biesemans for their contribution to data collection and Ana Teixeira and Daphne Tuyaerts for their help in designing the study. Finally, the authors would like to thank the Leuven IMMERSE patient advisory board members for providing feedback on the interpretation of the study results. This project received funding from the European Union's Horizon 2020 research and innovation program (777084; DynaMORE project and 945263; IMMERSE). GK is supported by the Research Foundation–Flanders (12ZZM21N and 1204924N).

Data Availability

To avoid potential identification of study participants, the complete interview transcripts will not be made publicly available.

Authors' Contributions

LT, GK, JW, MW, and IMG were involved in designing the study and setting up the data collection. Data analysis was undertaken in a collaboration between LT and LU. The first manuscript draft was written by LT with the support of GK. All authors contributed to the critical review of the manuscript and approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
IMPROVE training manual.
[DOC File, 4491 KB - humanfactors_v12i1e60096_app1.doc]

Multimedia Appendix 2
Interview guides.
[DOC File, 35 KB - humanfactors_v12i1e60096_app2.doc]

Multimedia Appendix 3
Demographic summary.
[DOC File, 25 KB - humanfactors_v12i1e60096_app3.doc]

Multimedia Appendix 4
Overview of themes.
[DOC File, 30 KB - humanfactors_v12i1e60096_app4.doc]

Multimedia Appendix 5
Individually summarized experiences.
[DOC File, 40 KB - humanfactors_v12i1e60096_app5.doc]

References

1. Lupton D. The digitally engaged patient: Self-monitoring and self-care in the digital health era. Soc Theory Health 2013 Aug;11(3):256-270. [doi: 10.1057/sth.2013.10]
2. World Health Organization. World Mental Health Report: Transforming Mental Health for All 2022.
3. Myin-Germeys I, Kasanova Z, Vaessen T, et al. Experience sampling methodology in mental health research: new insights and technical developments. World Psychiatry 2018 Jun;17(2):123-132. [doi: 10.1002/wps.20513] [Medline: 29856567]
4. van Os J, Verhagen S, Marsman A, et al. The experience sampling method as an mHealth tool to support self-monitoring, self-insight, and personalized health care in clinical practice. Depress Anxiety 2017 Jun;34(6):481-493. [doi: 10.1002/da.22647] [Medline: 28544391]
5. Peiris D, Miranda JJ, Mohr DC. Going beyond killer apps: building a better mHealth evidence base. BMJ Glob Health 2018;3(1):e000676. [doi: 10.1136/bmjgh-2017-000676] [Medline: 29527353]

XSL•FO
RenderX

6. Torous J, Myrick K, Aguilera A. The need for a new generation of digital mental health tools to support more accessible, effective and equitable care. World Psychiatry 2023 Feb;22(1):1-2. [doi: 10.1002/wps.21058] [Medline: 36640397]
7. Mohr DC, Weingardt KR, Reddy M, Schueller SM. Three problems with current digital mental health research... and three things we can do about them. Psychiatr Serv 2017 May 1;68(5):427-429. [doi: 10.1176/appi.ps.201600541] [Medline: 28412890]
8. Gulliksen J, Göransson B, Boivie I, Persson J, Blomkvist S, Cajander Å. Key principles for user-centred systems design. In: Seffah A, Gulliksen J, Desmarais MC, editors. Human-Centered Software Engineering — Integrating Usability in the Software Development Lifecycle: Springer Netherlands; 2005:17-36. [doi: 10.1007/1-4020-4113-6_2]
9. Weermeijer JDM, Wampers M, de Thurah L, et al. Usability of the experience sampling method in specialized mental health care: pilot evaluation study. JMIR Form Res 2023 Nov 21;7:e48821. [doi: 10.2196/48821] [Medline: 37988137]
10. Simons CJP, Drukker M, Evers S, et al. Economic evaluation of an experience sampling method intervention in depression compared with treatment as usual using data from a randomized controlled trial. BMC Psychiatry 2017 Dec 29;17(1):415. [doi: 10.1186/s12888-017-1577-7] [Medline: 29284448]
11. Fidock J, Carroll J, Rynne A. Evaluating information systems: an appropriation perspective. In: Hart DN, Gregor SD, editors. Information Systems Foundations: The Role of Design Science: ANU Press; 2010:121-142 URL: http://www.jstor.org/stable/j.ctt24h3z6.9 [doi: 10.22459/ISF.12.2010.06]
12. Bal M, Benders J, Vermeerbergen L. "Bringing the covert into the open": a case study on technology appropriation and continuous improvement. Int J Environ Res Public Health 2022 May 23;19(10):6333. [doi: 10.3390/ijerph19106333] [Medline: 35627869]
13. Carroll J. Completing design in use: closing the appropriation cycle [abstract]. Presented at: 12th ECIS; 2004; Turku, Finland URL: http://aisel.aisnet.org/ecis2004/44 [accessed 2025-02-18]
14. Chiu TML, Eysenbach G. Stages of use: consideration, initiation, utilization, and outcomes of an internet-mediated intervention. BMC Med Inform Decis Mak 2010 Nov 23;10:73. [doi: 10.1186/1472-6947-10-73] [Medline: 21092275]
15. Fidock JJ, Carroll J. The Model of Technology Appropriation: A Lens for Understanding Systems Integration in A Defence Context: ACIS; 2006.
16. Kaur E, Delir Haghighi P, Cicuttini FM, Urquhart DM. Smartphone-based ecological momentary assessment for collecting pain and function data for those with low back pain. Sensors (Basel) 2022 Sep 19;22(18):7095. [doi: 10.3390/s22187095] [Medline: 36146442]
17. Venkatesh V, Thong JYL, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Q 2012;36(1):157. [doi: 10.2307/41410412]
18. Bos FM, Snippe E, Bruggeman R, Wichers M, van der Krieke L. Insights of patients and clinicians on the promise of the experience sampling method for psychiatric care. Psychiatr Serv 2019 Nov 1;70(11):983-991. [doi: 10.1176/appi.ps.201900050] [Medline: 31434558]
19. Weermeijer J, Kiekens G, Wampers M, Kuppens P, Myin-Germeys I. Practitioner perspectives on the use of the experience sampling software in counseling and clinical psychology. Behav Inf Technol 2024 Feb 17;43(3):540-550. [doi: 10.1080/0144929X.2023.2178235]
20. Mestdagh M, Verdonck S, Piot M, et al. m-Path: an easy-to-use and highly tailorable platform for ecological momentary assessment and intervention in behavioral research and clinical practice. Front Digit Health 2023;5:1182175. [doi: 10.3389/fdgth.2023.1182175] [Medline: 37920867]
21. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: Building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208. [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
22. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
23. Saldana J. The Coding Manual for Qualitative Researchers: SAGE Publications; 2012.
24. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis: striving to meet the trustworthiness criteria. Int J Qual Methods 2017;16(1):1609406917733847. [doi: 10.1177/1609406917733847]
25. West P, Van Kleek M, Giordano R, Weal MJ, Shadbolt N. Common barriers to the use of patient-generated data across clinical settings. 2018 Apr 21 Presented at: CHI '18; Apr 21-26, 2018; Montreal QC Canada p. 1-13 URL: https://dl.acm.org/doi/proceedings/10.1145/3173574 [doi: 10.1145/3173574.3174058]
26. de Thurah L, Kiekens G, Sips R, Teixeira A, Kasanova Z, Myin-Germeys I. Using experience sampling methods to support clinical management of psychosis: the perspective of people with lived experience. Psychiatry Res 2023 Jun;324:115207. [doi: 10.1016/j.psychres.2023.115207] [Medline: 37087818]
27. Broekhuis M, van Velsen L, Peute L, Halim M, Hermens H. Conceptualizing usability for the eHealth context: content analysis of usability problems of eHealth applications. JMIR Form Res 2021 Jul 27;5(7):e18198. [doi: 10.2196/18198] [Medline: 34313594]
28. Kabakus AK, Bahcekapili E, Ayaz A. The effect of digital literacy on technology acceptance: an evaluation on administrative staff in higher education. Journal of Information Science 2023. [doi: 10.1177/01655515231160028]
29. Jang M, Aavakare M, Nikou S, Kim S. The impact of literacy on intention to use digital technology for learning: a comparative study of Korea and Finland. Telecomm Policy 2021 Aug;45(7):102154. [doi: 10.1016/j.telpol.2021.102154]

XSL•FO
RenderX

30. Ziesemer K, König LM, Boushey CJ, et al. Occurrence of and reasons for "missing events" in mobile dietary assessments: results from three event-based ecological momentary assessment studies. JMIR Mhealth Uhealth 2020 Oct 14;8(10):e15430. [doi: 10.2196/15430] [Medline: 33052123]
31. Wienroth M, Lund Holm Thomsen L, Høstgaard AM. Health technology identities and self. Patients' appropriation of an assistive device for self-management of chronic illness. Sociol Health Illn 2020 Jun;42(5):1077-1094. [doi: 10.1111/1467-9566.13079] [Medline: 32157709]
32. Medich M, Cannedy SL, Hoffmann LC, et al. Clinician and patient perspectives on the use of passive mobile monitoring and self-tracking for patients with serious mental illness: user-centered approach. JMIR Hum Factors 2023 Oct 24;10:e46909. [doi: 10.2196/46909] [Medline: 37874639]
33. Saunders KEA, Bilderbeck AC, Panchal P, Atkinson LZ, Geddes JR, Goodwin GM. Experiences of remote mood and activity monitoring in bipolar disorder: a qualitative study. Eur Psychiatry 2017 Mar;41:115-121. [doi: 10.1016/j.eurpsy.2016.11.005] [Medline: 28135594]
34. Bucci S, Gumley A, Walters J, et al. CONNECT: using electronic devices (e.g. smartphones, smartwatches) to predict relapse of psychosis - research protocol. ISRCTN Registry 2024. [doi: 10.1186/ISRCTN24746936]
35. King ZD, Yu H, Vaessen T, Myin-Germeys I, Sano A. Investigating receptivity and affect using machine learning: ecological momentary assessment and wearable sensing study. JMIR Mhealth Uhealth 2024 Feb 7;12:e46347. [doi: 10.2196/46347] [Medline: 38324358]

Abbreviations

ESM: experience sampling method

Edited by A Kushniruk; submitted 03.05.24; peer-reviewed by K Mouloudj, M Galani; revised version received 06.12.24; accepted 23.12.24; published 03.03.25.

Please cite as:
de Thurah L, Kiekens G, Weermeijer J, Uyttebroek L, Wampers M, Bonnier R, Myin-Germeys I
Understanding Appropriation of Digital Self-Monitoring Tools in Mental Health Care: Qualitative Analysis
JMIR Hum Factors 2025;12:e60096
URL: https://humanfactors.jmir.org/2025/1/e60096
doi:10.2196/60096

© Lena de Thurah, Glenn Kiekens, Jeroen Weermeijer, Lotte Uyttebroek, Martien Wampers, Rafaël Bonnier, Inez Myin-Germeys. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 3.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Development and Testing of the Kids Hurt App, a Web-Based, Pain Self-Report App for First Nations Youths: Mixed Methods Study

Karlee Francis[1,2*], BSc; Julie Francis[3], RN; Margot Latimer[4,5*], RN, PhD; Hayley Gould[6], BA; Shante Blackmore[2,7*], BSc, MD; Emily MacLeod[5*], MA

[1]Eskasoni Health Centre, Eskasoni, Nova Scotia, Canada
[2]Dalhousie University, Halifax, Nova Scotia, Canada
[3]Tajikeim k Mi'kmaw Health & Wellness, Membertou, Nova Scotia, Canada
[4]School of Nursing, Dalhousie University, 6299 South Street, Halifax, Nova Scotia, Canada
[5]IWK Health, Halifax, Nova Scotia, Canada
[6]Nova Scotia Health, Sydney, Nova Scotia, Canada
[7]Nova Scotia Health, Halifax, Nova Scotia, Canada
[*]these authors contributed equally

Corresponding Author:
Margot Latimer, RN, PhD
School of Nursing, Dalhousie University, 6299 South Street, Halifax, Nova Scotia, Canada

Abstract

Background: First Nations children and youths may have unique ways to convey their health needs that have not been recognized by health providers. This may contribute to the disparity between high rates of mental health and physical pain and low rates of treatment for the conditions they experience. Evidence suggests that a colonial history has resulted in poor experiences with the health care system, lack of trust with health providers, and miscommunication between clinicians and patients. Contemporary ways, using both Indigenous and Western knowledge, are needed to bridge the gap in communicating pain.

Objective: The aim of this qualitative study was to test the usability and clinical feasibility of the Kids Hurt App with First Nations youths and clinicians working with youths.

Methods: Using a Two-Eyed Seeing approach, the Kids Hurt App was developed using concepts from validated mood and pain assessment apps combined with community-based research that gathered First Nations youths and clinicians perspectives on quality, intensity, and location of pain and hurt. The Kids Hurt App contains 16 screens accessible on any web-based device.

Results: In total, 3 rounds of low-fidelity testing (n=19), 2 rounds of high-fidelity testing (n=20), and 2 rounds of clinical feasibility testing (n=10) were conducted with First Nations youths (10 - 19 years) to determine the relevance, validity, and usability of the Kids Hurt App. High-fidelity testing was also conducted with 15 clinicians after completing the high-fidelity youth sessions. Youths had constructive suggestions that were used to improve the app in subsequent rounds of version testing. There was one main discrepancy between youths and clinicians related to preference for how best to visually convey pain. The youth's preference was maintained in the app.

Conclusions: All youths in all rounds of testing indicated that they would use the Kids Hurt App if it was available to them in a health care setting, with most clinicians noting that the app would be useful in practice.

(JMIR Hum Factors 2025;12:e48370) doi:10.2196/48370

KEYWORDS

app; eHealth; pain; Indigenous; First Nations; children; youths; mobile phone

Introduction

Overview

Indigenous peoples in Canada are comprised of 3 distinct groups: First Nations, Inuit, and Métis. They reside across Canada, from large cities to small isolated communities [1]. In comparison to non-Indigenous peoples, Indigenous peoples are the fastest-growing cohort in Canada with a significantly younger population [1]. Indigenous children younger than 14 years of age represent 25.4% of the total Indigenous population, while the same age group in the non-Indigenous population only compromises 16% [1]. Repeated reports of racism in health care, such as those in the In Plain Sight report describing Joyce

XSL•FO
RenderX

Echaquan's fatal health care experience in Quebec [2], combined with research findings [3,4] suggest that health care providers need to understand the structural and historical factors that contribute to the current racial disparities faced by Indigenous peoples [5]. A legacy of colonialism and present-day racism is a determinant of health that is unique to Indigenous peoples so efforts need to be made to create better understanding, appropriate assessment, and improved care for health conditions such as pain [6].

While 1 in every 5 Canadian people experiences chronic pain, this number is significantly higher in Indigenous communities, and most types of pain are more prevalent in Indigenous peoples [7]. Indigenous children are particularly vulnerable to experiencing pain. In one study examining pain diagnoses, a large sample of First Nations children (n=2631) who were compared to an age-, sex-, and location-matched sample of non–First Nations children, demonstrated that the First Nations cohort experienced significantly more physical pain–related diagnoses for 10 of the 13 pain indicators studied [7]. Further to this, while ear- and throat-related diagnoses were the 2 most diagnosed pain conditions for both groups, the First Nations cohort was significantly less likely than the non–First Nations cohort to visit a specialist for these same conditions, demonstrating inequities in diagnosis and treatment [7]. Additionally, these data showed that children (0 - 9 years) who had a physical pain diagnosis were more likely to have a mental health diagnosis in adolescence, a finding only evident in the non–First Nation cohort who were seemingly able to access mental health care [7].

Cultural variances in pain expression have also been found between Indigenous and non-Indigenous children and youths [8,9]. Latimer et al [8] concluded that Mi'kmaq children were stoic and hid their pain. Children reported that they would not cry when enduring pain and preferred instead to be "brave" or "tough it out" [8]. These findings led researchers to conclude that health clinicians may not have ways to accurately assess or document the level of pain a child or youth is experiencing. Similar findings of stoicism were reported in Indigenous peoples in Australia, with participants often being described as quiet about their pain and not reporting pain due to higher pain tolerance, fear of Western medicine, or intercultural communication difficulties [10]. Clinicians are trained to use numerical or face pain scales to assess pain; however, only minimal research has been done to determine if these self-report pain scales are culturally appropriate for use with Indigenous children, and indeed to our knowledge, none in First Nations children [11,12]. Community members reported that numerical pain rating assessment scales are confusing, and it is difficult to attach hurt or pain to one measure such as a face or number [9]. The evidence suggests that there is a clinical practice gap in assisting Indigenous youths to convey their pain to health providers. This study endeavored to find a culturally appropriate mechanism to support youths in sharing their pain and to address reports of a lack of clinician's ability to recognize pain, which creates a barrier to appropriate health care [13]. In response to these needs, the Kids Hurt App [14] was cocreated with Indigenous and clinical partners. It was created with the notion that Indigenous peoples have many effective ways of communicating knowledge that include visual images and storytelling. This led the developers to explore how both Indigenous and Western ways could be used to develop a better assessment tool that puts the technology in the hands of the youths. The aim of this qualitative study was to test the usability and feasibility of the Kids Hurt App with First Nations youths and clinicians working with youths.

Background

Two-Eyed Seeing

The project used a Two-Eyed Seeing (TES) approach, a term coined by Elders Murdena and Albert Marshall, which recognizes the benefit of seeing Indigenous knowledge and experience from one eye while also seeing strengths from another eye's perspective that is different from one's own [15]. The Western perspective is most typically how health care is currently delivered. When both eyes are used, the ultimate benefit is achieved by effectively enhancing the health of Indigenous peoples through the practical sharing of knowledge [16].

Digital Health Interventions for Indigenous Children and Youths

Although there is a growing body of literature on digital health and app-based interventions to manage pain in children and youths, research on the development and usability of app-based interventions for Indigenous peoples is limited. In a review of digital health solutions for Indigenous mental well-being across Canada, the United States, Australia, and New Zealand, Hensel et al [17] indicated that co-designed app-based interventions had been shown to reduce anxiety and depression symptoms and contribute to well-being. The Aboriginal and Islander Mental Health Initiative for Youth App, or AIMhi-Y, was developed in consultation with Indigenous youths from Australia and the Torres Strait Islands [18]. The youths identified apps as a potential way to mitigate barriers to accessing help and highlighted the need for a strength-based approach [18,19]. A randomized controlled trial found that iBobbly, a suicide prevention app developed in consultation with Indigenous youths living in Australia, had only a minimal, nonsignificant impact on psychological well-being outcomes, although qualitative data indicated the app was helpful and that construct validity issues may have been the reason why there was not a significant impact [20]. To date, there is no research on the development or usability of apps for Indigenous children and youths in Canada [21].

Development of the Kids Hurt App

The Kids Hurt App was developed by the Aboriginal Children's Hurt and Healing Initiative for children aged 10 - 19 years. Funding for the app development and testing was received from the Canadian Institutes of Health Research (FRN# 162455 & SCA-145102), Indigenous Health Nursing Research Chair, Dalhousie University (FRN# 167603), and IWK Health Foundation. The Aboriginal Children's Hurt and Healing Initiative is a broad partnership, consisting of Indigenous community leaders, clinicians, elders, youths, researchers from universities, and health delivery systems. In keeping with Ownership, Control, Access and Possession (OCAP) principles

XSL•FO
RenderX

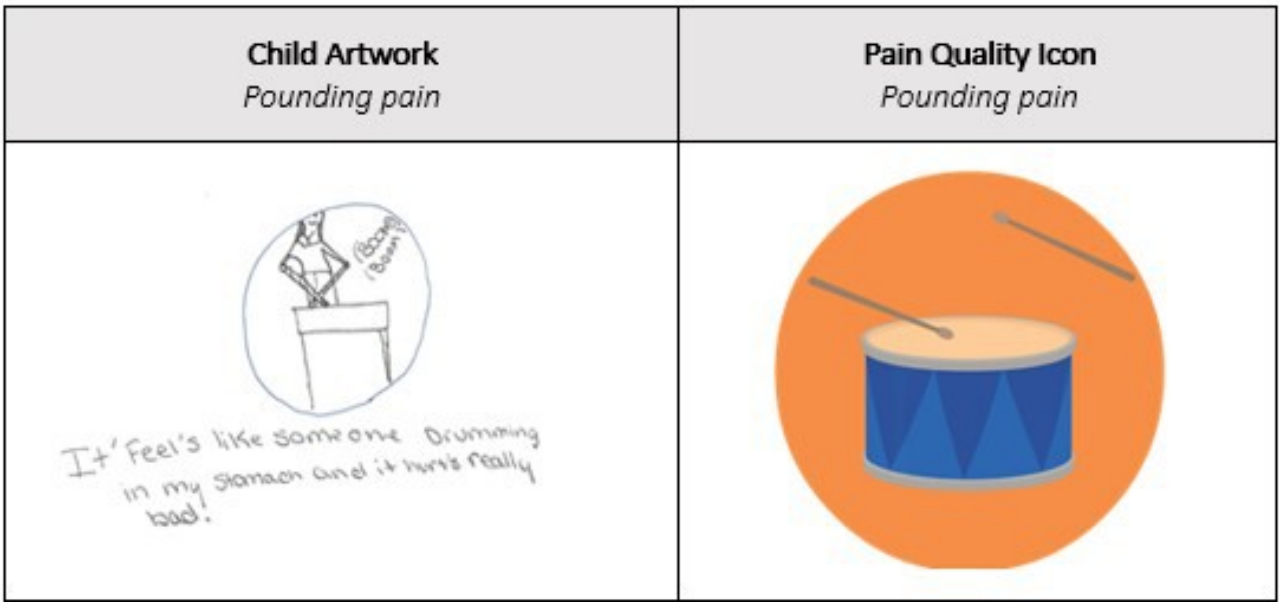
[22], the app is owned by communities involved in this research study. OCAP is a registered trademark of the First Nations Information Governance Centre and describes the principles of ownership, control, access, and possession to ensure that First Nations have control over data processes including ownership and how their information is used. The app was developed to facilitate pain communication between clinicians and Indigenous youths in a culturally safe manner. It is not meant to replace the history-taking step of a health care interaction but rather serves to augment the quality of communication exchange by giving youths an opportunity to share their pain story in a comfortable, culturally safe way.

The app was developed using TES to ensure both Indigenous and Western concepts and content were included. In total, 16 "screens" or interfaces were developed. Multimedia Appendix 1 provides an example of the final 16 screens developed. Previous versions of these screens were used throughout various stages of testing.

The first app screen asks users their gender and age (see Screen 1 in Multimedia Appendix 1), with the subsequent screen providing them with the option to create an avatar or choose a predesigned one (see Screen 2-3 in Multimedia Appendix 1). First Nations youths were involved in the development of the avatar designs through initial content validity testing to ensure that visual representation was present. Once an avatar is chosen, it is prompted to answer a series of responses to physical and emotional pain prompts. The inclusion of both physical and emotional pain prompts provides an example of how TES [15] was used throughout the development of the app. Earlier design ideas for the app had only physical pain prompts; however, community feedback suggested that incorporating both physical and emotional pain prompts would better represent the Indigenous holistic view of health, which involves "a healthy balance of 4 elements or aspects of wellness: physical, emotional, mental, and spiritual" [23]. While this version of the Kids Hurt App does not incorporate all 4 wellness aspects, emotional and physical pain were chosen because they were the most reported types of pain in previous research [7]. Future app iterations are expected to include spiritual and mental dimensions. There is also evidence that physical and emotional pain are interconnected, and the app presents an opportunity to highlight this for both youths and clinicians.

Additional incorporation of TES [15] involved the use of the International Association for the Study of Pain assessment tool guidelines [24] in the development, ensuring three key assessment features were embedded into the app: (1) pain location (see Screen 4-6 in Multimedia Appendix 1), (2) pain quality (see Screen 7 in Multimedia Appendix 1, and (3) pain intensity (see Screen 8 in Multimedia Appendix 1). Pain quality icons were conceptualized based on previous art-based research with Indigenous youths [8,9] as well as the Iconic Pain Assessment Tool [25] and the Pain Quality Icon (QuILT) [26] as illustrated in Figure 1.

Figure 1. Pain quality icon development.



Pain intensity was initially captured using 2 scales: a 10-point "face scale," mimicking the Faces Pain Scale—Revised (FPS-R) [27], and a newly developed "Jar of Hurt," building on the concept of the "Pieces of Hurt" tool [28] and respecting the nation's language and word for "pain" which is "hurt" [8]. A jar was used as a common, everyday item that would provide the imagery of filling something up with pain or hurt. In addition, details of *where* users were when they got hurt (ie, school) and *what* they were doing (ie, playing) were also included with options based on common places of injury noted in the First Nations Regional Health Survey [29].The inclusion of the *where* and *what* aspects of the app allowed the incorporation of storytelling (see Screen 9-10 in Multimedia Appendix 1), an approach with significant cultural ties for Indigenous populations [30].

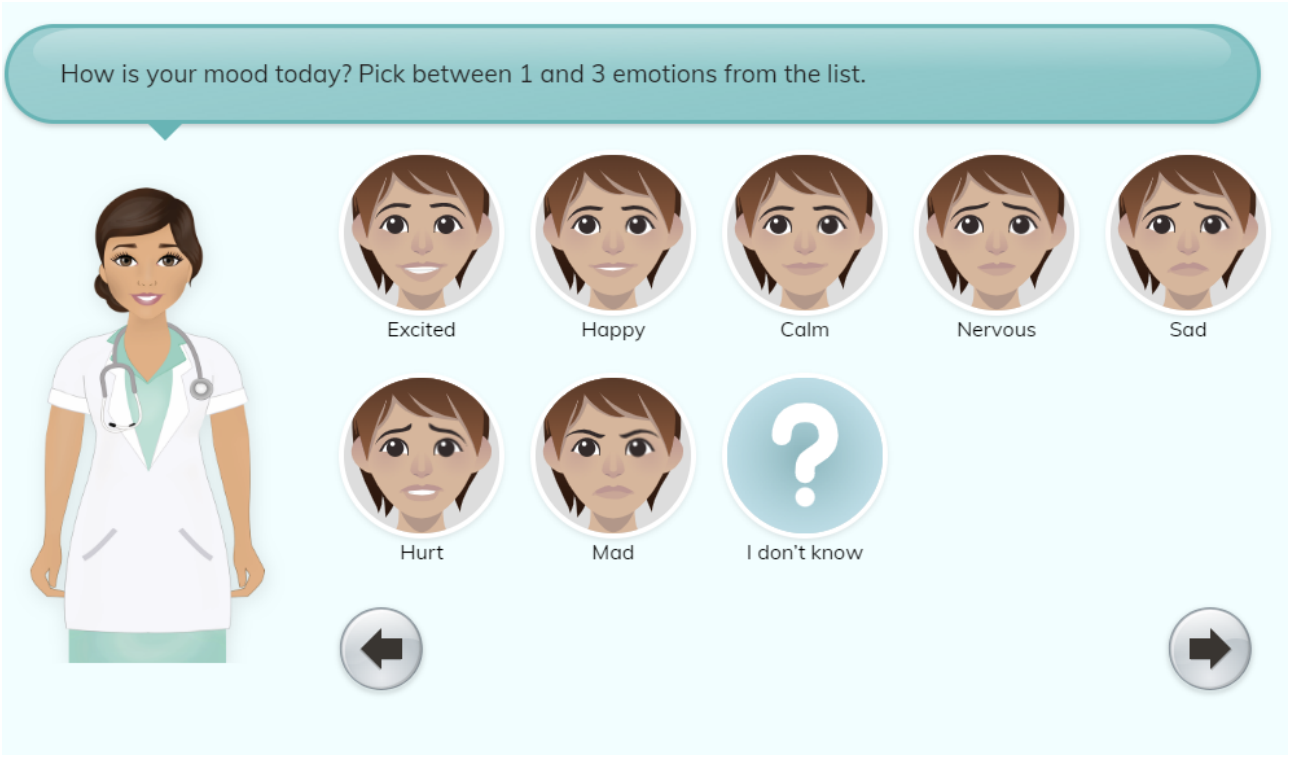
After completing questions regarding the user's physical pain, users are then prompted to provide information on their emotional pain through a series of similar screens (see Screen 11-14 in Multimedia Appendix 1). Emotional pain icons were initially chosen based on 22 common emotions identified from

XSL•FO
RenderX

20 validated mood apps (Multimedia Appendix 2). Through the initial content validity testing with First Nations youths, these 22 emotions were narrowed down to 8, which were identified as most likely to be used to describe their emotional mood. The resulting 8 emotions, both "negative" and "positive," provided a range of options and avoided bias in responses (Figure 2). Emotional pain questions focused on users' current emotional state (happy, sad, and anxious), intensity of those emotions (using the 10-point Jar of Hurt), and any physical symptoms that may be associated with their emotional pain (ie, chest tightness).

Upon answering both the physical and emotional pain questions, information is presented on a final storyboard screen (see Screen 15 in Multimedia Appendix 1).

Figure 2. Emotional pain icons.



Methods

Study Design

Once an initial version of the app was finalized three types of testing took place: (1) low-fidelity testing, (2) high-fidelity testing, and (3) clinical feasibility testing. This testing or development strategy mirrored the process published by Stinson et al [31].

Ethical Considerations

Prior to project initiation, the study was approved by both an Indigenous and tertiary pediatric hospital research ethics board (Mi'kmaw Ethics Watch & IWK Health Research Ethics Board; #1017831). Informed consent was obtained from all participants or their parents or guardians prior to taking part in study procedures. Youth participants received a CAD $10 (US $6.98) gift card after completing the testing. Clinician participants were not compensated. Data were deidentified immediately following data collection sessions.

Participants

First Nation youth participants were recruited through social media platforms and community organizations or contacts to complete low-fidelity, high-fidelity, and clinical feasibility testing. Clinicians recruited from a First Nation Health Centre and a pediatric health center unit specializing in emotional or physical pain–related services (rheumatology, pain clinic, etc) also completed high-fidelity testing. Eligibility criteria for youth participants included (1) self-identifying as Indigenous, (2) residing in a First Nation community, and (3) being between 10 and 19 years of age.

Data Collection

Overview

The study took place over a 5-year period between April 2017 and March 2022. Significant delays were experienced due to the COVID-19 pandemic. All youth data collection sessions were completed within the youth's Indigenous community to ensure that local support resources were available if a youth felt distressed. A trained mental health clinician was available as a resource for each session, and smudging was offered for youths who took part in the clinical feasibility sessions.

Low-Fidelity Testing

Low-fidelity testing is a qualitative usability approach where youths were able to view a paper-based version of the app. Youths completed a demographic survey prior to reviewing the app screens. In total, 3 rounds of low-fidelity testing took place. During each round, youths were shown paper-based screenshots of the Kids Hurt App and asked what they liked and disliked about the design through a series of semistructured questions.

XSL•FO
RenderX

Sessions were audio-recorded and transcribed for clarity and accuracy. After each round of testing, feedback was reviewed and integrated into the app. Screenshots of the updated version of the app were then shown to the next round of participating youths until no further changes were suggested based on the methods of Stinson et al [31].

High-Fidelity Testing

High-fidelity testing is similar to low-fidelity testing; however, participants were able to review a fully functioning version of the app rather than paper-based screenshots. Youths completed a demographic survey prior to using the app, and all sessions were audio-recorded and transcribed. Two rounds of high-fidelity testing were completed with participating youths to determine functionality and acceptability and to allow for any suggested revisions to be shared. The app was viewed on an iPad, and the youths worked through the app in the presence of the research assistant. At the start of each session, youths were asked to complete the app by themselves and to talk out loud about any likes, dislikes, or difficulties. The length of time it took for youths to complete the app from start to finish was recorded. Then, they were asked to go through the app for the second time with the research assistant who asked them semistructured questions. Once finished, they were asked open-ended questions on their overall thoughts regarding likes, dislikes, ease of use, and whether they would use the app in a health care setting if it was available. Feedback from each round was incorporated until no further changes were suggested (2 rounds). Interviews with the youths were audio-recorded in each of the 2 rounds to maintain the integrity and accuracy of the findings.

After the youths completed the high-fidelity testing, clinicians were then invited to take part in high-fidelity testing. Clinicians were invited to use the app and then answer a series of questions either through a digital survey or during a focus group. Focus group sessions were transcribed to ensure data accuracy. Questions assessed the clarity and importance of the various app components and also allowed space for participants to provide suggestions on what they would like to change or add for each screen. Demographic questions were also asked to determine participant's profession and types of experience.

Clinical Feasibility Testing

Clinical feasibility testing was completed with youths to help determine if the app works well, if it is easy to use, and if youths found it relevant for use in their care. Youths were asked to work through the app with the research assistant present and then answer a series of questions in a digital survey format. Questions ranged from demographic questions to those regarding the app's usefulness as well as functionality. Once an initial 5 youths had completed their review of the app, changes were incorporated, and an additional 5 youths were invited to take part. After 2 rounds of clinical feasibility testing with very positive results and minimal changes, it was determined that a third round was not needed.

Data Analysis

Demographic survey responses from all rounds of testing were analyzed using the SPSS Statistics (version 24; IBM Corp) database system to provide an overall breakdown of age and gender. Field notes and transcribed interviews were reviewed to determine any changes needed for subsequent rounds of testing. Revisions associated with more large-scale app functionality that were unable to be incorporated between testing rounds, such as incorporating a feature to track pain over time, were added to a database for future recommendations.

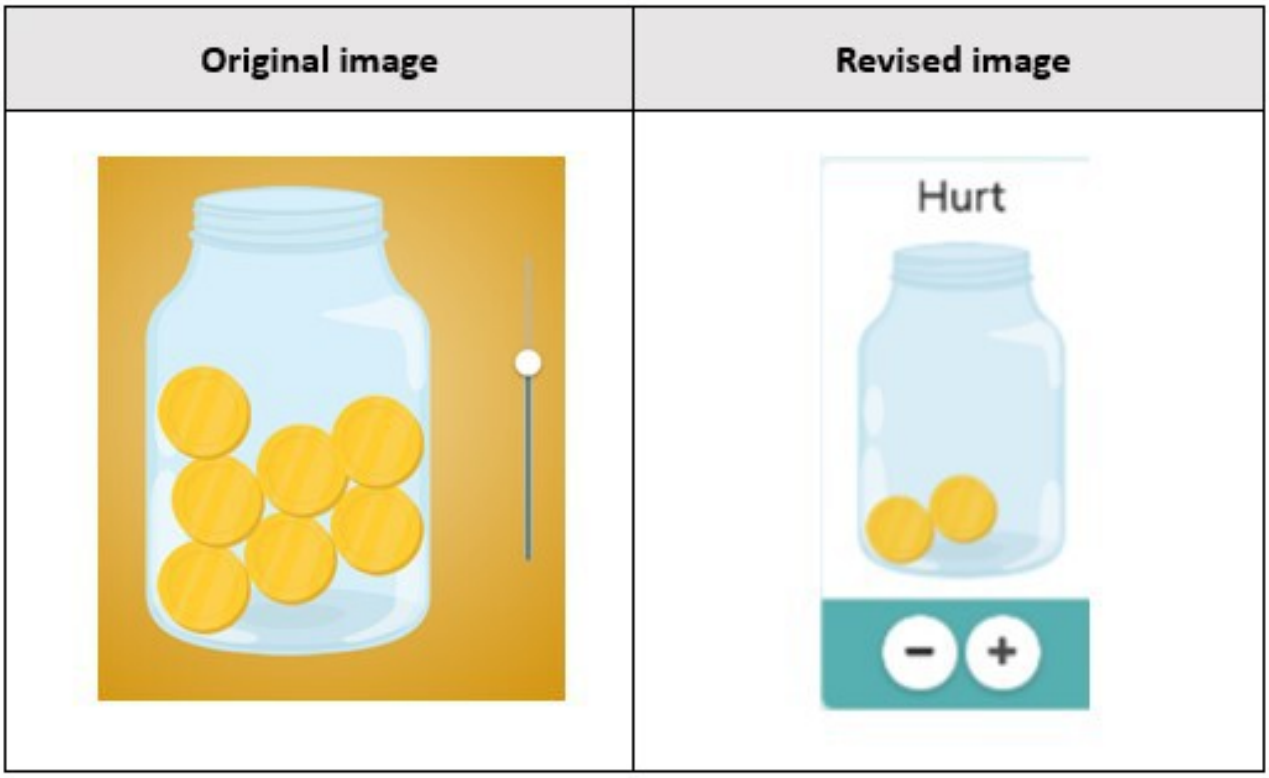
Results

Low-Fidelity Testing

A total of 19 youths participated in 3 rounds of low-fidelity testing (round 1: 7 youths, round 2: 6 youths, and round 3: 6 youths). In total, 63% (n=12) of the participants were identified as female. The average age for all participants was 13 (SD 2.46) years. In the first round of low-fidelity testing, 1 participant suggested that the Jar of Hurt scale should use a plus or minus button to increase or decrease the amount of hurt rather than the original "slider" option (Figure 3). In subsequent rounds, youths were shown both the slider and the plus or minus option and asked their preference. The majority of youths (n=11) preferred the plus or minus option. The directions for describing how to use the Jar of Hurt were also identified during low-fidelity testing as needing clarification. A total of 6 of the 19 youths identified this issue. "I don't know, it's a little confusing. It sounds complicated" (Female, 14 years). The language was simplified, and it was agreed that it would be further reviewed in the high-fidelity rounds.

XSL•FO
RenderX

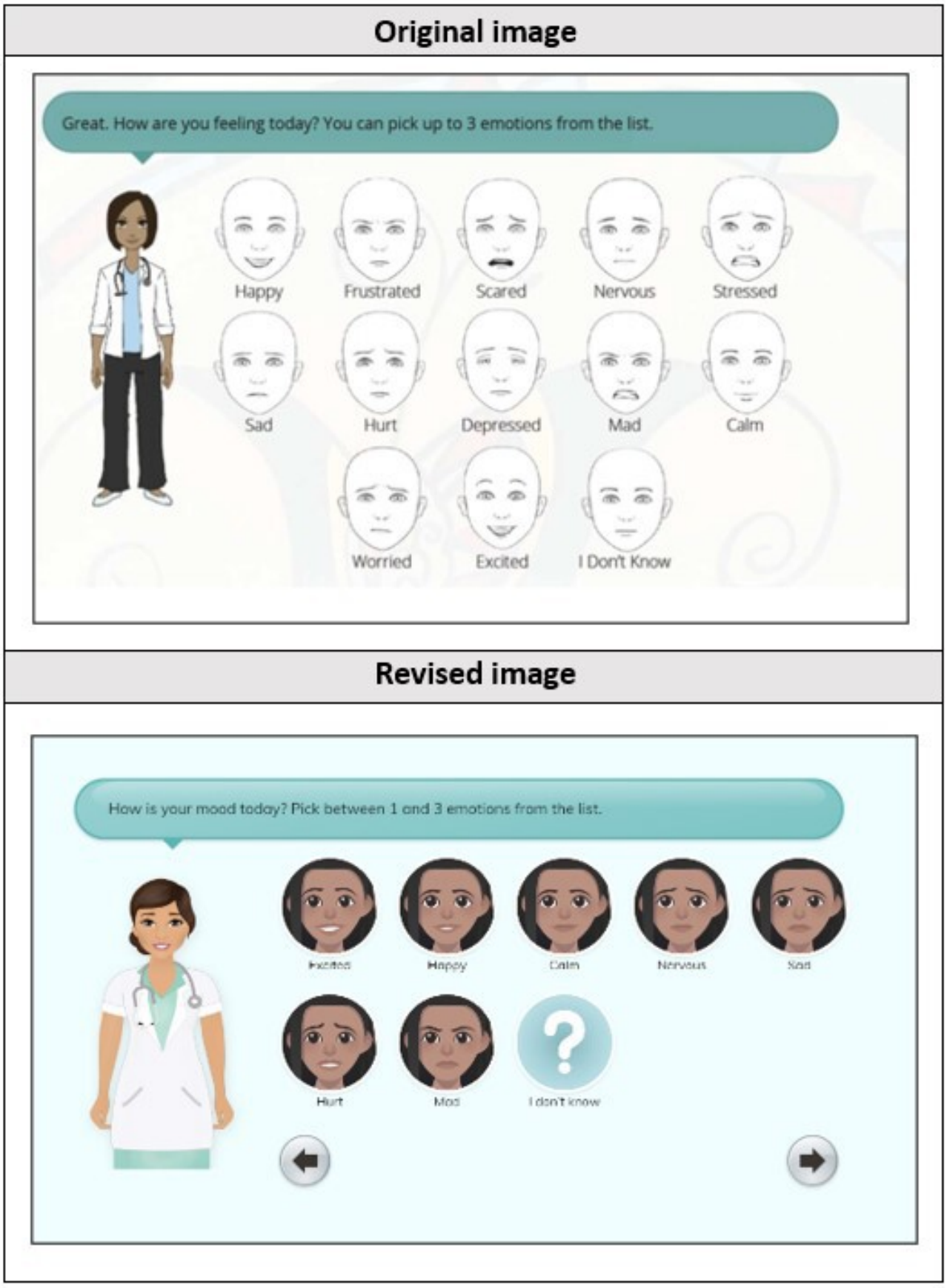
Figure 3. Jar of Hurt—plus or minus button addition.



The final noted change from low-fidelity testing was an overall dislike of the way the emotion face icons looked. One youth described the faces as “creepy.” To address this, the facial icons were redesigned to look like the user’s avatar rather than a generic face as illustrated in Figure 4.

XSL•FO
RenderX

Figure 4. Emotion faces—revised. Original image depicts 10 emotions rather than the finalized 8 emotions, as this was an initial mock-up screenshot for testing purposes only.



XSL•FO
RenderX

High-Fidelity Testing

A total of 20 youths participated in the 2 rounds of high-fidelity testing (round 1: 10 youths and round 2: 10 youths). In total, 70% (n=14) of participants were female, and the average age for youth participants was 16 (SD 2.46) years. The average time it took youths to complete the app was 3.61 minutes (216.6 seconds; SD 48). All youth participants identified enjoying the app and agreed it would be helpful to them.

> *It was pretty helpful. I know when you go to your doctor, you don't really want to open up. Sometimes you're nervous. At least the app helps explain how you feel.* [Female, 17 years]

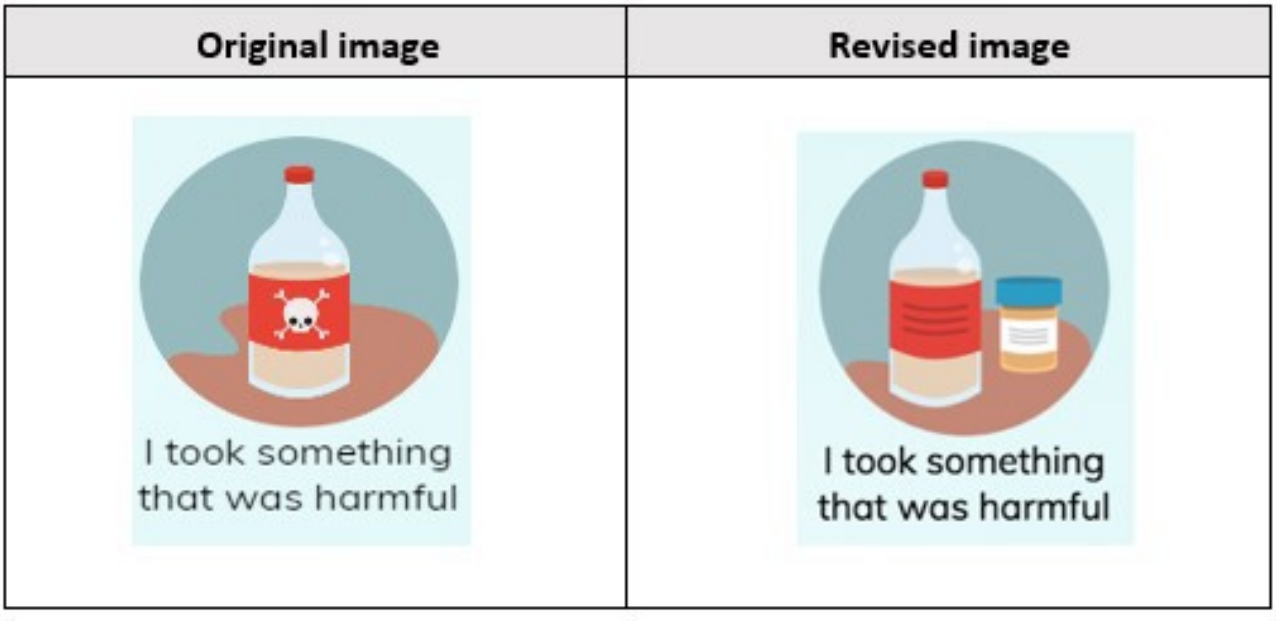
Confusion remained regarding the use of the Jar of Hurt with 2 of 10 youths in the first round of testing describing it as confusing. As a result, examples of an empty jar and a full jar were added beside the plus or minus buttons to provide a visual of how to increase or decrease the amount of pain or hurt (see Screen 7 and 8 in Multimedia Appendix 1). After this change was implemented, no confusion was noted in the remaining high-fidelity testing round.

Participants also suggested revisions to the "I took something harmful" icon in the "how you were hurt" section. Initially, this icon included a skull and crossbones to symbolize poison; however, youths suggested that something more representative of drugs or alcohol should be included.

> *I think this would be something chemical, based on the picture and skull thing. But I think something that can be added is maybe drugs or alcohol. Because the person may not always feel comfortable saying it but hitting it on the button without thinking and if it's right in front of you, they may be honest.* [Female, 18 years]

This icon was revised to better reflect youth participant's perception of which icons are most relevant to them (Figure 5).

Figure 5. "Harmful" icon changes.



A final revision resulting from the youth high-fidelity sessions was determining a preference for which scale to use to represent the quantity of pain. Youths were shown a scale with facial images similar to those in the FPS-R and superimposed in the avatar [27] (Figure 6) and the Jar of Hurt scale (Figure 7) and asked which they felt would best allow them to quantify their pain. The majority (n=12, 60%) preferred the Jar of Hurt scale. One youth shared their reason why:

> *If you had to choose one option, I would say the jar scale would be a lot better. Because maybe some people, they know how to hide it. Their pain. So, I'd say the jar scale is a lot better.* [Male, 16 years]

XSL•FO
RenderX

Figure 6. Faces Pain Scale—Revised expressions avatar adapted faces scale.

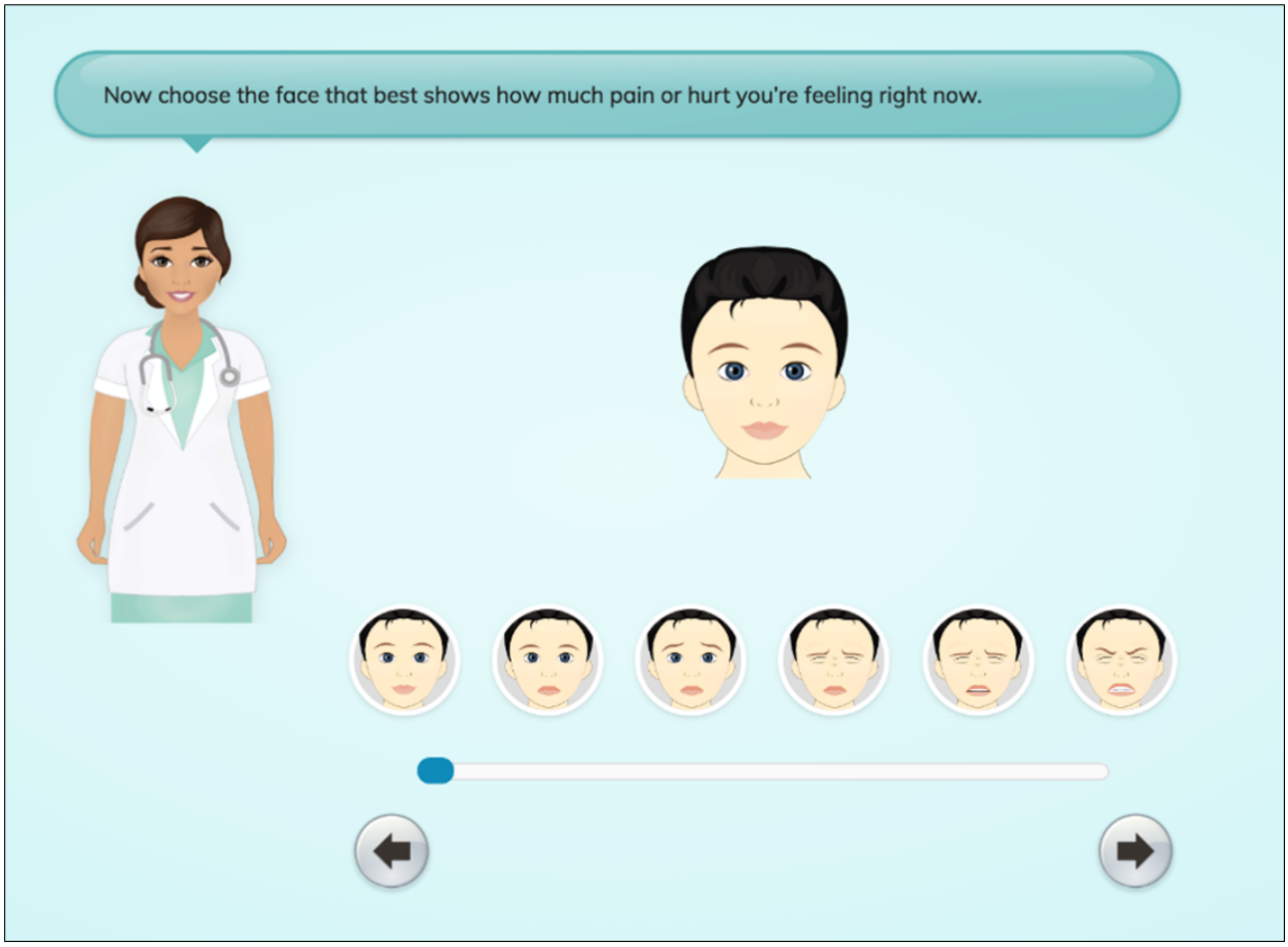
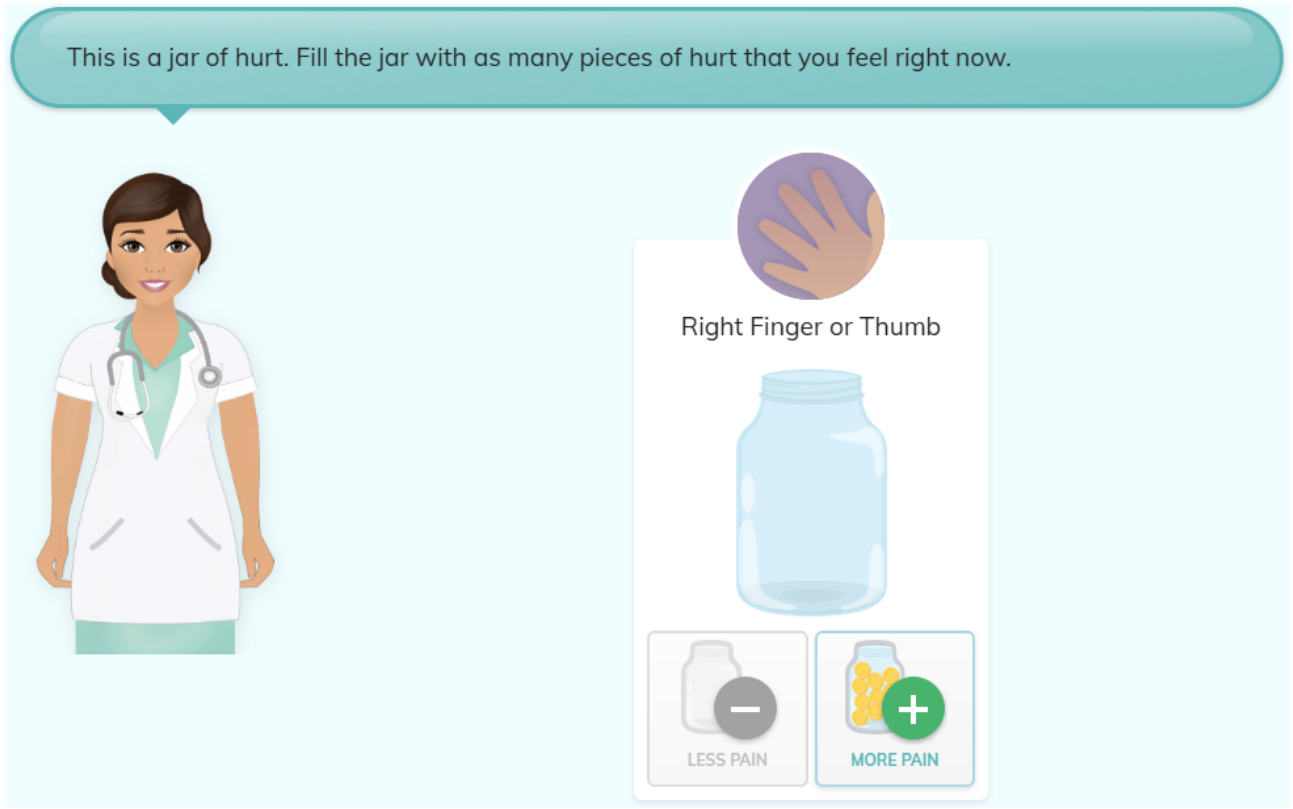


Figure 7. Jar of Hurt scale.



XSL•FO
RenderX

In addition to youths, a total of 15 clinicians participated in high-fidelity testing. In total, 10 clinicians completed this through a digital survey, and 5 participants took part in a focus group session. Most survey respondents were nurses (n=7, 70%), and focus group participants included physicians, physiotherapists, nurses, and psychologists.

In total, 14 of 15 clinicians found the app to be moderately or extremely useful. Like youths, clinicians were asked to identify their preference between the Jar of Hurt or the FPS-R adapted scale (Figures 6 and 7) through the survey questions. A total of 7 of the 9 clinician respondents identified a preference for the FPS-R adapted scale.

On 2 different screens, clinicians suggested adding an "other" option. For the quality of pain screen, 4 of 5 clinicians from the focus group felt that an "other" option should be included to allow users to add their own quality of pain descriptors. Similarly, for the emotional pain screen, all clinician participants suggested adding an "other" option to allow users the flexibility of adding any emotion they felt, even if not in the presented list. These additions were made in subsequent versions of the app.

Clinical Feasibility Testing

A total of 10 youths completed clinical feasibility testing (round 1: 5 and round 2: 5). Results from the clinical feasibility testing were very positive, with 80% (n=8) of respondents identifying that they found the app easy to use and 90% (n=9) saying they would use the app in a hospital or health care setting. No technical errors were found. Participants suggested that the Jar of Hurt method was more geared toward younger youths but also noted that they were happy to use the scale either way. Suggestions for change were mainly larger design changes identified as "future changes" including having a text or email function to send results to themselves or their care providers, having the ability to track changes in their pain or hurt over time, and including translation in their Indigenous language.

Discussion

Principal Findings

In this study of clinicians and First Nations youths, 3 types of testing occurred in the development of the Kids Hurt App: low-fidelity, high-fidelity, and clinical feasibility. For the 2 groups of participants, each subsequent round of testing reduced the app modifications. Overall changes were minor, but important observations were thought to increase the usefulness of the app for users.

Recent literature has identified that app-based interventions may remove barriers to accessing care for youths [18,32,33]. Apps are accessible and may mitigate confidential concerns and concerns about stigma. Smartphone-based apps may be particularly useful, as authors from one study found that smartphones are a mainstream technology and are an advantage within the health care system [34]. While existing research suggests that Western pain assessment tools are not appropriate for Indigenous peoples [10], this study demonstrates how an app can be developed collaboratively and engage youths to share their pain respecting both Indigenous and Western knowledge systems. All youths stated that they would use the Kids Hurt App if it was available in the health care setting. Youths reaffirmed that they often found it difficult to articulate their feelings and felt that the app could help them communicate their needs. This app allows them to share their pain and help facilitate enhanced communication.

One notable finding in this study was the process of determining which pain scale to use, the "Jar of Hurt" scale or the "face scale." In the initial development stages of the app, the "Jar of Hurt" concept was chosen based on previous research regarding the "Pieces of Hurt" tool [28]. While this research was completed with younger children than our target population, the concept of using objects rather than faces was appealing based on previous research indicating that Indigenous children are often stoic in their expression of pain and may hide their pain rather than displaying it outwardly through facial expressions [8]. Through high-fidelity testing, 60% (n=12) of youths preferred the "Jar of Hurt" scale over the "face scale," and while it was noted in clinical feasibility testing that the scale was more geared toward younger youths, participants identified being happy to use the tool, regardless of that. Clinician preference was strongly noted for the face scale likely because of its similarity to the FPS-R scale, a well-known pain measurement tool [27]. However, given that the app was developed to provide youths with a valid way to convey their pain and hurt and to develop an alternative to Western pain assessment tools, the "Jar of Hurt" was maintained as the preference for conveying pain intensity. This finding highlights the importance of taking the time to validate patient indicators of pain consistent with different age groups and populations.

Improved Health Care Experience

During a typical patient visit, a clinician obtains the patient's history, current pain experience, and future needs in a short amount of time [35]. This can be challenging if a patient finds it difficult to articulate their pain and hurt for a range of reasons, such as stoicism or power dynamic, further causing delay in timely and appropriate care [34,36]. In this study, participants identified the app as "easy to use" and were able to complete the app screens in an average of 3.61 minutes (216.6 seconds; SD 48). The youths' suggestions for app improvement, such as the Jar of Hurt being a better way to capture pain instead of a face that may hide pain, provide some indication that the Kids Hurt App may facilitate better communication, and consequently health care experience, by allowing Indigenous youths to convey their pain through an app. Other comments such as the youths who indicated pushing a button might be easier than talking could imply that the app offers a nonthreatening and safe way for youths to share their stories. Research has shown that youths have high technological literacy and are comfortable using eHealth apps [37,38]. In a generation where technology is fluent, the Kids Hurt App can give youths a voice in the health care system, in a medium they are comfortable with, to personally advocate for their physical and emotional pain care.

Limitations

At the time of testing, the app was not yet available in the youth participant nation's language. The current version of the app [14] is fully translated into Mi'kmaw. Communication between

XSL•FO
RenderX

the provider and patient should always be clear and effective, and having health information accessible in one's primary language is important [39]. Thus, we recommend that the app be available in the nation's primary language in any further testing. An additional limitation is that testing was only completed with 2 First Nation communities so results cannot be assumed to be accurate for all Indigenous nations.

Conclusions

Culturally safe care acknowledges the attitudes, knowledge, and behaviors of another's culture in a respectful manner. If a provider encounters an Indigenous patient, it would be important to understand the patient's unique historical legacies, ways of communicating, and any potential intergenerational trauma that may influence pain report [40]. Implications of creating and using an app that has been codeveloped by the end users (First Nations youths) increase the usefulness of the app and provide a space for the provider and patient relationship to be optimized for best care outcomes. The short app completion time is an indication of the minimal effort it takes to offer a culturally appropriate alternative to pain self-report and may reduce miscommunication and repeat visits for the same condition. Future research will include testing the app with non-Indigenous youths as well as in various health settings such as mental health, tertiary, primary care, and emergency departments. Further development of the app related to youth's self-report of effective strategies that reduce pain and hurt is planned. The Kids Hurt App is a culturally safe and evidence-based communication tool that may allow First Nations youths to comfortably share their pain and hurt stories. Each nation may be different in its communication and will need to review the app for usability and feasibility in its own communities.

Acknowledgments

The authors would like to thank the First Nations youths and the clinicians for taking the time to support the development and testing of the Kids Hurt App.

Conflicts of Interest

None declared.

Multimedia Appendix 1
App screens.
[PDF File, 1154 KB - humanfactors_v12i1e48370_app1.pdf]

Multimedia Appendix 2
Mood app inventory.
[PDF File, 229 KB - humanfactors_v12i1e48370_app2.pdf]

References

1. Indigenous population continues to grow and is much younger than the non-Indigenous population, although the pace of growth has slowed. Statistics Canada. 2022. URL: https://www150.statcan.gc.ca/n1/daily-quotidien/220921/dq220921a-eng.htm [accessed 2025-01-24]
2. Turpel-Lafond ME. In Plain Sight: addressing Indigenous-specific racism and discrimination in BC health care. 2020. URL: https://engage.gov.bc.ca/app/uploads/sites/613/2020/11/In-Plain-Sight-Full-Report-2020.pdf [accessed 2025-01-24]
3. Hamed S, Bradby H, Ahlberg BM, Thapar-Björkert S. Racism in healthcare: a scoping review. BMC Public Health 2022 May 16;22(1):988. [doi: 10.1186/s12889-022-13122-y] [Medline: 35578322]
4. Mahabir DF, O'Campo P, Lofters A, Shankardass K, Salmon C, Muntaner C. Experiences of everyday racism in Toronto's health care system: a concept mapping study. Int J Equity Health 2021 Mar 10;20(1):74. [doi: 10.1186/s12939-021-01410-9] [Medline: 33691682]
5. Goodman A, Fleming K, Markwick N, et al. "They treated me like crap and I know it was because I was Native": the healthcare experiences of Aboriginal peoples living in Vancouver's inner city. Soc Sci Med 2017 Apr;178:87-94. [doi: 10.1016/j.socscimed.2017.01.053] [Medline: 28214449]
6. McNally M, Martin D. First Nations, Inuit and Métis health: considerations for Canadian health leaders in the wake of the Truth and Reconciliation Commission of Canada report. Healthc Manage Forum 2017 Mar;30(2):117-122. [doi: 10.1177/0840470416680445] [Medline: 28929885]
7. Latimer M, Rudderham S, Lethbridge L, et al. Occurrence of and referral to specialists for pain-related diagnoses in First Nations and non-First Nations children and youth. CMAJ 2018 Dec 10;190(49):E1434-E1440. [doi: 10.1503/cmaj.180198] [Medline: 30530610]
8. Latimer M, Simandl D, Finley A, et al. Understanding the impact of the pain experience on Aboriginal children's wellbeing: viewing through a Two-Eyed Seeing lens. First Peoples Child Fam Rev 2014;9(1):22-37. [doi: 10.7202/1071791ar]
9. Latimer M, Young S, Dell C, Finley GA. Aboriginal children and physical pain: what do we know? IJIH 2012;9(1):7-14. [doi: 10.18357/ijih91201212389]

XSL•FO
RenderX

10. Strong J, Nielsen M, Williams M, Huggins J, Sussex R. Quiet about pain: experiences of Aboriginal people in two rural communities. Aust J Rural Health 2015 Jun;23(3):181-184. [doi: 10.1111/ajr.12185] [Medline: 25945587]
11. Jimenez N, Garroutte E, Kundu A, Morales L, Buchwald D. A review of the experience, epidemiology, and management of pain among American Indian. Alaska Native Aboriginal Can peoples J Pain 2011;12(5):511-522. [doi: 10.1016/j.jpain.2010.12.002] [Medline: 21330217]
12. Ellis JA, Ootoova A, Blouin R, et al. Establishing the psychometric properties and preferences for the Northern Pain Scale. Int J Circumpolar Health 2011 Jun;70(3):274-285. [doi: 10.3402/ijch.v70i3.17823] [Medline: 21703128]
13. Empathy, dignity, and respect: creating cultural safety for Aboriginal people in urban health care. Health Council of Canada. 2012. URL: https://learningcircle.ubc.ca/files/2014/05/Empathy-dignity-and-respect-Creating-cultural-safety-for-Aboriginal-people-in-urban-health-care.pdf [accessed 2025-01-24]
14. Kids Hurt App. ACHH Initiative. 2024. URL: https://www.kidshurtapp.com [accessed 2025-01-24]
15. Marshall A, Marshall M, Bartlett C. Two-Eyed Seeing in medicine. In: Greenwood M, Leeuw S, Lindsay N, Reading C, editors. Determinants of Indigenous Peoples' Health in Canada: Beyond the Social: Canadian Scholars Press; 2015:16-24.
16. Hovey RB, Delormier T, McComber AM, Lévesque L, Martin D. Enhancing Indigenous health promotion research through Two-Eyed Seeing: a hermeneutic relational process. Qual Health Res 2017 Jul;27(9):1278-1287. [doi: 10.1177/1049732317697948] [Medline: 28682710]
17. Hensel JM, Ellard K, Koltek M, Wilson G, Sareen J. Digital health solutions for Indigenous mental well-being. Curr Psychiatry Rep 2019 Jul 1;21(8):68. [doi: 10.1007/s11920-019-1056-6] [Medline: 31263971]
18. Povey J, Sweet M, Nagel T, et al. Drafting the Aboriginal and Islander Mental Health Initiative for Youth (AIMhi-Y) app: results of a formative mixed methods study. Internet Interv 2020 Sep;21:100318. [doi: 10.1016/j.invent.2020.100318] [Medline: 32477884]
19. Dingwall KM, Puszka S, Sweet M, Nagel T. "Like drawing into sand": acceptability, feasibility, and appropriateness of a new e‐mental health resource for service providers working with Aboriginal and Torres Strait Islander people. Aust Psychol 2015 Feb 1;50(1):60-69. [doi: 10.1111/ap.12100]
20. Tighe J, Shand F, McKay K, Mcalister TJ, Mackinnon A, Christensen H. Usage and acceptability of the iBobbly app: pilot trial for suicide prevention in Aboriginal and Torres Strait Islander youth. JMIR Ment Health 2020 Dec 1;7(12):e14296. [doi: 10.2196/14296] [Medline: 33258782]
21. Noronha N, Smith SJ, Martin Hill D, et al. The use of mobile applications to support Indigenous youth wellbeing in Canada. IJCAR 2021;8(1):124-134. [doi: 10.7202/1077724ar]
22. The First Nations principles of OCAP. First Nations Information Governance Centre. 2012. URL: https://fnigc.ca/ocap-training/ [accessed 2025-01-24]
23. An overview of Aboriginal health in Canada. National Collaborating Centre for Aboriginal Health. 2013. URL: https://www.ccnsa-nccah.ca/docs/context/FS-OverviewAbororiginalHealth-EN.pdf [accessed 2025-01-24]
24. Terminology. International Association for the Study of Pain. 2018. URL: https://www.iasp-pain.org/resources/terminology/ [accessed 2025-1-24]
25. McMahon E, Wilson-Pauwels L, Henry JL, Jenkinson J, Sutherland B, Brierley M. The Iconic Pain Assessment Tool: facilitating the translation of pain sensations and improving patient-physician dialogue. J Biocommun 2008;34(2):20-24 [FREE Full text]
26. Lalloo C, Kumbhare D, Stinson JN, Henry JL. Pain-QuILT: clinical feasibility of a web-based visual pain assessment tool in adults with chronic pain. J Med Internet Res 2014 May 12;16(5):e127. [doi: 10.2196/jmir.3292] [Medline: 24819478]
27. Hicks CL, von Baeyer CL, Spafford PA, van Korlaar I, Goodenough B. The Faces Pain Scale-Revised: toward a common metric in pediatric pain measurement. Pain 2001 Aug;93(2):173-183. [doi: 10.1016/S0304-3959(01)00314-1] [Medline: 11427329]
28. Hester NK. The preoperational child's reaction to immunization. Nurs Res 1979;28(4):250-255. [Medline: 255924]
29. First Nations Regional Health Survey (RHS) 2008/10: national report on adults, youth and children living in First Nations. First Nations Information Governance Centre. 2012. URL: https://fnigc.ca/wp-content/uploads/2020/09/5eedd1ce8f5784a69126edda537dccfc_first_nations_regional_health_survey_rhs_2008-10_-_national_report_adult_2.pdf [accessed 2025-01-24]
30. Chan AS. Chapter 11: storytelling, culture, and indigenous methodology. In: Bainbridge A, Formenti L, West L, editors. Discourses, Dialogue and Diversity in Biographical Research 2021:170-185. [doi: 10.1163/9789004465916_012]
31. Stinson JN, Jibb LA, Nguyen C, et al. Development and testing of a multidimensional iPhone pain assessment application for adolescents with cancer. J Med Internet Res 2013 Mar 8;15(3):e51. [doi: 10.2196/jmir.2350] [Medline: 23475457]
32. Seko Y, Kidd S, Wiljer D, McKenzie K. Youth mental health interventions via mobile phones: a scoping review. Cyberpsychol Behav Soc Netw 2014 Sep;17(9):591-602. [doi: 10.1089/cyber.2014.0078] [Medline: 25007383]
33. Uhlhaas P, Torous J. Digital tools for youth mental health. NPJ Digit Med 2019;2:104. [doi: 10.1038/s41746-019-0181-2] [Medline: 31646184]
34. Suso-Ribera C, Mesas Á, Medel J, et al. Improving pain treatment with a smartphone app: study protocol for a randomized controlled trial. Trials 2018 Feb 27;19(1):145. [doi: 10.1186/s13063-018-2539-1] [Medline: 29482614]

XSL•FO
RenderX

35. Tai-Seale M, McGuire TG, Zhang W. Time allocation in primary care office visits. Health Serv Res 2007 Oct;42(5):1871-1894. [doi: 10.1111/j.1475-6773.2006.00689.x] [Medline: 17850524]
36. Pathak EB, Wieten SE, Wheldon CW. Stoic beliefs and health: development and preliminary validation of the Pathak-Wieten Stoicism Ideology Scale. BMJ Open 2017 Nov 14;7(11):e015137. [doi: 10.1136/bmjopen-2016-015137] [Medline: 29138193]
37. Thackeray R, Hunter M. Empowering youth: use of technology in advocacy to affect social change. J Comput Mediat Commun 2010;15(4):575-591. [doi: 10.1111/j.1083-6101.2009.01503.x]
38. Pretorius C, Coyle D. Young people's use of digital tools to support their mental health during Covid-19 restrictions. Front Digit Health 2021;3:763876. [doi: 10.3389/fdgth.2021.763876] [Medline: 34927133]
39. Molina RL, Kasper J. The power of language-concordant care: a call to action for medical schools. BMC Med Educ 2019 Nov 6;19(1):378. [doi: 10.1186/s12909-019-1807-4] [Medline: 31690300]
40. Dell EM, Firestone M, Smylie J, Vaillancourt S. Cultural safety and providing care to Aboriginal patients in the emergency department. CJEM 2016 Jul;18(4):301-305. [doi: 10.1017/cem.2015.100] [Medline: 26565079]

Abbreviations

FPS-R: Faces Pain Scale—Revised
OCAP: Ownership, Control, Access and Possession
TES: Two-Eyed Seeing

Edited by A Kushniruk, EM Borycki; submitted 20.04.23; peer-reviewed by D Dunsmuir, N Mungoli; revised version received 11.10.24; accepted 27.11.24; published 03.03.25.

Please cite as:
Francis K, Francis J, Latimer M, Gould H, Blackmore S, MacLeod E
Development and Testing of the Kids Hurt App, a Web-Based, Pain Self-Report App for First Nations Youths: Mixed Methods Study
JMIR Hum Factors 2025;12:e48370
URL: https://humanfactors.jmir.org/2025/1/e48370
doi:10.2196/48370

© Karlee Francis, Julie Francis, Margot Latimer, Hayley Gould, Shante Blackmore, Emily MacLeod. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 3.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Understanding Experiences of Telehealth in Palliative Care: Photo Interview Study

Mahima Kalla[1], BE, PhD; Teresa O'Brien[1], BMedSci, PhD; Olivia Metcalf[1], PhD; Rashina Hoda[2], BSc, PhD; Xiao Chen[3], PhD; Andy Li[2], BCS, BEngEE (Hons); Catriona Parker[4], BAppSc, MAppSc, PhD; Michael Edward Franco[2,5,6], MBBS; Sam Georgy[7], BSc, BEng, MBA, MSc; Kit Huckvale[1], MBChB, BSc, MSc, PhD; Christopher Bain[8], MBBS, MInfoTech, PhD; Peter Poon[9,10], MBBS, PGDipPM, CHIA

[1]Centre for Digital Transformation of Health, Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne, Melbourne Connect, 700 Swanston St, Carlton, Australia

[10]Department of Medicine, School of Clinical Sciences at Monash Health, Faculty of Medicine,Nursing & Health Sciences, Monash University, Melbourne, Australia

[2]Faculty of Information Technology, Monash University, Melbourne, Australia

[3]School of Information and Physical Sciences, College of Engineering, Science and Environment, University of Newcastle, Callaghan, Australia

[4]Transfusion Research Unit, Faculty of Medicine, Nursing & Health Sciences, Monash University, Clayton, Australia

[5]Faculty of Medicine, Nursing and Health Sciences, Monash University, Clayton, Australia

[6]St Vincent's Health, Melbourne, Australia

[7]Healthdirect Australia, Sydney, Australia

[8]Alliance for Digital Health At Monash, Faculty of Information Technology, Monash University, Clayton, Australia

[9]Supportive and Palliative Care Unit, Monash Health, Melbourne, Australia

Corresponding Author:
Mahima Kalla, BE, PhD
Centre for Digital Transformation of Health, Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne, Melbourne Connect, 700 Swanston St, Carlton, Australia

Abstract

Background: It is widely accepted that the COVID-19 pandemic has accelerated the era of online health care delivery, including within community palliative care. This study was part of a larger project involving a collaboration between universities, health care services, government agencies, and software developers that sought to enhance an existing telehealth (video call) platform with additional features to improve both patient and health care professional (HCP) experience in a palliative care context.

Objective: The aim of this study was to understand palliative care patients' and HCPs' experiences of telehealth delivery in a palliative care context in Victoria, Australia. For the purposes of this study, telehealth included consultations by both video and telephone calls. By better understanding users' experiences and perceptions of telehealth, we hoped to determine users' preferences for new telehealth enhancement features.

Methods: A total of 6 health care professionals and 6 patients were recruited from a major tertiary hospital network's palliative care unit in Victoria, Australia. Participants were asked to generate 3 - 5 photographs depicting their telehealth experiences. These photographs were used as visual aids to prompt discussion during subsequent one-on-one interviews. Intertextual analysis was conducted to identify key themes.

Results: A total of 3 overarching themes emerged: comfort (or lack thereof) afforded by telehealth, connection considerations in telehealth, and care quality impacts of telehealth. Patients (n=6) described telehealth as supporting their physical and psychological comfort and maintaining connection with HCPs, yet there were specific situations where it failed to meet their needs or impacted care quality and delayed treatment. HCPs (n=6) recognized the benefit of telehealth for patients but reported several limitations of telehealth, in particular due to lack of physical examination opportunities. Participants indicated that 2 types of connection were imperative for effective telehealth delivery: technical connection (eg, good internet connectivity or clear phone line) and interpersonal connection (ie, good rapport and therapeutic alliance between the HCPs and patients). Often technical connection issues impeded the development of interpersonal connection between the HCPs and patients in telehealth.

Conclusions: The findings presented in this study combined with other co-design activities, which are outside the scope of this paper, indicated the potential value of a telehealth enhancement feature that generates patient-facing clinical consultation summaries. Our team has developed a video telehealth enhancement feature (or "add-on"), which will enable clinicians to distill key actionable advice and self-management guidance discussed during teleconsultations for a take-home summary document for patients. The add-on's prototype has also been subjected to an initial simulation study, which will be reported in a future publication.

XSL•FO
RenderX

(JMIR Hum Factors 2025;12:e53913) doi:10.2196/53913

KEYWORDS

consultation summary; digital scribe; qualitative research; telehealth; digital health; photo-elicitation; palliative care; photo interview; qualitative research; photographs; intertextual analysis

Introduction

In Australia, chronic illness continues to be a major challenge in health care [1] with more than three-quarters (78.6%) of Australians impacted by at least one chronic condition [2]. Many of these chronic conditions are life-limiting [3], and thus patients may benefit from palliative care [4]. The 2018 National Palliative Care Strategy recommends the approach of using palliative care as a "means to improve the quality of life of patients and their families affected with life-threatening illnesses through the prevention and relief of suffering by means of early identification, assessment, and treatment of pain and other problems" [5]. Studies have shown that palliative care improves quality of life by improving independence, providing symptom relief and pain control, and supporting the physical, emotional, and spiritual needs of patients, their carers, and loved ones [6-9]. Furthermore, palliative care is linked with a significant decrease in health care use and prolongment of survival [10].

Specialist palliative care in Australia is delivered using different models of care depending on the geographic location and available resources [11]. Many of these models of care were transformed during the COVID-19 pandemic to promote physical distancing [12,13]. One such approach for the delivery of specialist palliative care in Australia was a move to telehealth to limit hospital attendance and in-home visits [12]. The implementation of telehealth in palliative care in Australia demonstrated promise beyond the pandemic, with the potential to improve communication, facilitating more equitable access to palliative care (especially for those in rural and regional areas), avoiding lengthy travel to hospitals or clinics (particularly when patients are very unwell), and potentially assisting with timely symptom relief [14]. However, for continued sustainability and success of telehealth models, telehealth needs to adapt and evolve based on feedback from patients and health care professionals (HCPs) [15].

The importance of user involvement in the implementation of palliative care services has been recognized and prioritized by researchers to improve productivity, quality, and relevance of care [16]. The aim of this study was to understand palliative care patients' and HCPs' experiences of telehealth delivery in a palliative care context in Victoria, Australia. For the purposes of this study, telehealth included consultations by both video and telephone calls. This study was part of a larger project involving a collaboration between universities, health care services, government agencies, and software developers, which sought to enhance an existing telehealth (video call) platform with additional features to improve both patient and HCP experience. By better understanding users' experiences and perceptions of telehealth, we hoped to determine users' preferences for other telehealth enhancement features.

Methods

Participants

Palliative care professionals and patients were recruited from the Monash Health Supportive and Palliative Care Unit at a tertiary hospital network in Victoria, Australia. HCPs, including palliative care medical and nursing staff, and interpreters involved in the Unit's use of telehealth were invited to participate by email. The Unit has video telehealth palliative care clinics for oncology patients. In addition, during COVID-19 outbreak, the Unit provided a telehealth outreach service when in-person home visits were not possible.

Patients aged 18 years and older, with conversational English proficiency, currently using telehealth to access palliative care were invited through email invitations and postcards. A total of 6 patients and 6 HCPs participated in the study. All participants lived in Victoria and spoke English as their primary language. The patients' ages ranged between 54 and 61 years old. In addition, 4 of the 6 patient participants identified as female and 2 identified as male. The HCPs' ages ranged between 33 and 62 years old, with 5 identifying as female and 1 as male. The duration in their health care roles ranged from 6 months to 8 years. Their roles were 'doctor' (n=3), 'interpreter' (n=2), and 'nursing clinical coordinator' (n=1). Our 6 patient participants have been deidentified as P009, P012, P013, P014, P015, and P016, and HCP participants as HCP001, HCP003, HCP 004, HCP005, HCP010, and HCP011.

Data Collection

The study used a qualitative photo interviewing technique [17,18]. Participants were requested to generate 3‐5 photographs based on certain prompts and share them with the research team. HCPs were asked to click photographs to depict their everyday work life in the context of telehealth models of care, what telehealth represents in their work, general experiences, and opportunities for improving current telehealth models. Similarly, patients were asked to generate photographs to illustrate their experience of receiving health care through telehealth, what telehealth represents for them as patients and their everyday lives, and opportunities for improving telehealth models of care. Multimedia Appendix 1 and Multimedia Appendix 2 shows the specific instructions provided to both HCPs and patients.

Participants emailed their photographs to a member of the research team. Subsequently, telephone interviews, lasting between 30 and 45 minutes, were conducted with each HCP and patient participant. During the interviews, participants were asked to elaborate on what each of the photos meant for them and how the photographs related to their experiences of telehealth. During the interviews, use of participant-generated photographs helped trigger participants' memories and invoke their deeper consciousness, values, attitudes, and personal

XSL•FO
RenderX

meanings of telehealth [19-21]. The visual aids also helped the researcher and participant reach a shared understanding of the topic [22]. All the interviews were audio-recorded and transcribed verbatim by a professional transcribing company. The interview transcripts were deidentified, and photos blurred to protect participants' identities. The anonymized transcripts were imported into Microsoft Word and the photos inserted at the point where they were discussed during the interviews, allowing for an intertextual analysis in which the participants' photos and words could be viewed together [23].

Data Analysis

A multistep data analysis process was conducted, involving intertextual analysis wherein participants' words and photos were viewed together and analyzed using a 3-step process involving (1) preview, (2) review, and (3) cross-photo comparison [23]. In the first "preview" stage, the first author reviewed all the photos and identified recurring motifs. These motifs were recorded in a PowerPoint slide deck and shared with the research team at a meeting for feedback and further interpretation. Subsequently, a more in-depth "review" was conducted wherein one researcher coded all the interview transcripts with relevant themes and subthemes in more granular detail. The emergent codebook was shared and workshopped with the research team to further distill the emerging patterns.

The final stage in the data analysis process was "cross-photo comparison." During this phase, the research team discussed the similarities and differences within and across participant groups (eg, within patient participants' or within HCPs' accounts, or across patients' and HCPs' accounts). Often the patients' and HCPs' accounts provided contrasting experiences of the same facet of telehealth use (referred as "counter-examples"). For example, where patients valued the convenience of telehealth and being able to liaise with their doctors from anywhere, HCPs found patients' use of telehealth in certain places (eg, public spaces) disruptive and uncongenial to the therapeutic process. To organize the emergent themes, a storyboard was developed by the first author containing illustrative quotes and photographs from both patients and HCPs. This story board has been provided in Multimedia Appendix 3 . The storyboard was presented to the research team and subsequently, manuscript development commenced.

Ethical Considerations

Ethics approval for the study was obtained from the Monash Health Human Research Ethics Committee (project identification number 79681). All participants were provided plain language summaries of the research project, and what their participation would entail. All participants provided informed written consent. They also had the option to stop the interview at any point. Only approved members of the research team had access to the identifiable participant information. All interview transcripts were deidentified before the commencement of the data analysis process. Photos generated by participants showing any humans were also blurred to preserve participants' privacy.

Results

Overview

A total of 3 themes were identified: comfort (or lack thereof) afforded by telehealth, connection considerations in telehealth, and care quality impacts of telehealth. It should be noted that while these themes are presented as distinct categories, they are not necessarily mutually exclusive. In addition, in some cases the themes included counter-examples, as mentioned above.

Comfort (Or Lack Thereof) Afforded by Telehealth

For some participants, telehealth afforded them greater physical and psychological comfort. However, for some others, telehealth hampered their physical and psychological comfort. Therefore, this theme elucidates how telehealth use impacted patients' and HCPs' physical and psychological comfort.

Physical Comfort

Palliative care patients often experience physical discomfort, including pain and fatigue, which in turn impacts their ability to visit clinics for in-person consultations. Patient participants advised that they had arranged their homes in ways that minimized their discomfort. Telehealth enabled them to remain in this environment, unlike in-person appointments, which lead to physical exertion due to travel and wait times, with this issue compounded for rural patients.

Patient participant P015 captured this theme when discussing a photo (Figure 1) of the room where they spent most of their time. They reported being "in pain 24/7" and explained how telehealth helped minimize their discomfort:

> *...You can actually see how close my bed; my chair is to my office desk. And so, I just use the walking frame, I get to the chair very easily. I sit in my chair, I'm listening to a bit of music, or I'm just watching TV…I'll sit in the wait in the chair, waiting for the doctor to come in.* [P015]

XSL•FO
RenderX

Figure 1. Patient home office for telehealth consultations (P015).



By contrast, some HCPs associated telehealth with physical discomfort. HCPs working in the palliative care outreach service during the pandemic, as opposed to the Unit's dedicated telehealth clinics, reported that private rooms were not always available. Therefore, telehealth appointments were often conducted in shared workspaces, which could be noisy and distracting, impacting their ability to effectively conduct quality consultations with patients. For example, HCP001 included a photo of the shared office space (Figure 2) where they often conducted telehealth consultations and stated:

> *It gets very loud because we're [all] talking… So then anybody else talking on the phone, someone else has to talk up to talk over the top of that, and then anybody having a conversation with each other in the office has to talk up.* [HCP001]

XSL•FO
RenderX

Figure 2. Palliative care service shared office space (HCP001).



Psychological Comfort

Alongside physical comfort, telehealth provided patients with psychological comfort by reducing the stress associated with in-person appointments. Participants reported a variety of psychological stressors associated with in-person appointments, which were mitigated by using telehealth, for example, travel and wait times, risk of exposure to COVID-19 in hospitals, forgetting key questions for their doctors, and financial stress, particularly for rural patients who may need to make overnight accommodation arrangements for brief appointments. P013 provided a photo of a piggybank (Figure 3) and stated:

> *…going to appointments, car wear and tear, fuel for my daughter, parking fees, petrol... [With] telehealth, [I] don't have to spend that money… [the money saved could be spent on] all the bits and bobs that they keep chucking at me to take.* [P013]

XSL•FO
RenderX

Figure 3. Piggy bank (P013).



The psychological comfort associated with the freedom to stay in one's chosen location and receive health care from the comfort of their own homes is also illustrated in P014's photo (Figure 4) and quote as follows:

My bush family are my reason for living out here … This is my family. Because of telehealth, I get to stay home and enjoy them. You know, I don't lose a day or two being on the road away from them. I know it's only animals but for some people, you know, animals are our family. [P014]

XSL•FO
RenderX

Figure 4. Life in 'the bush' (P014).



For some outreach service HCPs, however, the sudden large-scale transition to telehealth in the aftermath of the COVID-19 outbreak, led to some psychological discomfort at their perceived inability to provide their usual standard of care. Given their professional training had primarily been in face-to-face care, they felt ill-equipped to efficaciously transition to online care. HCP005 provided a photo of themselves holding their stethoscope up to the camera (Figure 5), demonstrating the limitation of physical examination in telehealth consultations. HCP005 said that the return to more in-person consultations was "almost a relief" post–COVID-19 pandemic. They further elaborated:

> *…often to be thorough in your assessment, as a doctor, you've been trained [to] do your history, examination and do investigations. I guess telehealth takes away … what we're trained to do.* [HCP005]

XSL•FO
RenderX

Figure 5. Health care professional (HCP) uncomfortable about limitations of telehealth (HCP005).



Connection Considerations in Telehealth

Participants indicated that 2 types of connection were imperative for effective telehealth delivery: technical connection (eg, good internet connectivity or clear phone line) and interpersonal connection (ie, good rapport and therapeutic alliance between HCPs and patients). Often technical connection issues impeded the development of interpersonal connection between the HCPs

XSL•FO
RenderX

and patients in telehealth. This theme illustrates how technical connection and interpersonal connection impacted the participants' telehealth experiences.

Technical Connection

The technical aspects of connection included the skills required to use the technology, internet and mobile phone connectivity, and specific hardware and software concerns. HCP participants noted that their patients had first been introduced to the telehealth model of care during the pandemic with varying levels of technological proficiency. Patients with limited technological proficiency needed to build their confidence and skills over time, with support from HCPs, support staff, and family members. HCPs described situations where no one had checked whether or not a patient's computer had required microphone and video capabilities. Internet connectivity and bandwidth issues were also reported by both patients and HCPs. HCPs, in particular, due to their volume of telehealth calls, reported a lot of time lost to troubleshooting connectivity issues, which impacted the time available for clinical conversations. Often participants had to transition to telephone calls. For example, P012 shared an image of their landline home phone (Figure 6), describing it as follows:

> *It is the most reliable tool in my house. It never drops out. Whereas the Internet … the internet on my computer can be slow or drop out or whatever at a telehealth appointment. And my mobile phone drops out, it's unreliable as well. But this phone is not unreliable. So if I lose a connection with any of the doctors, I know I can ring them on that phone … and it's not going to drop out. That's really important … So the landline is a good back up source for me for telehealth. I always have it sitting there in case something happens.* [P012]

XSL•FO
RenderX

Figure 6. Landline home phone as back-up (P012).

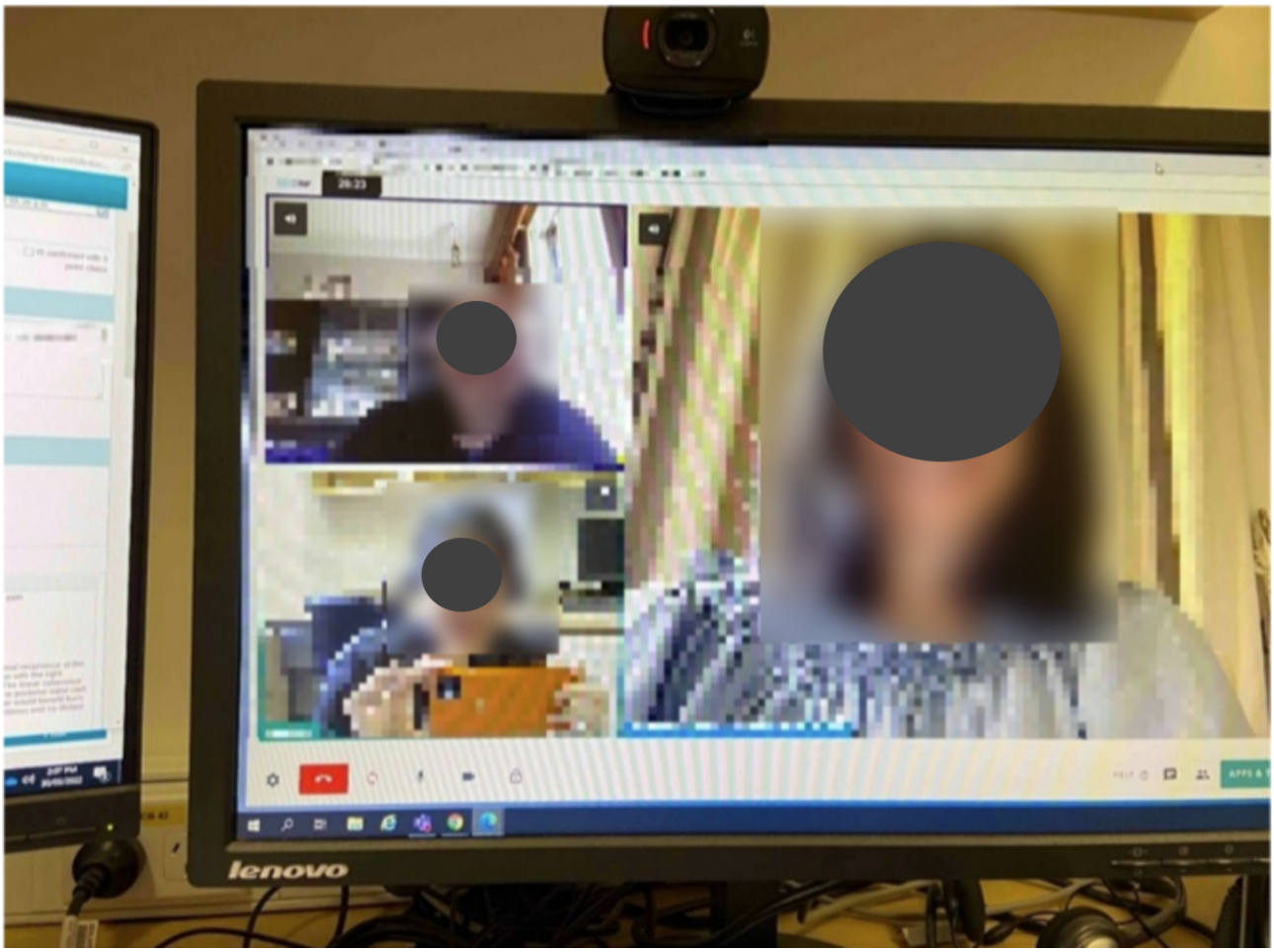


Nevertheless, participants also reported that in the absence of connectivity issues, and with the appropriate skill level to use, and availability of the right technology, telehealth can be very useful and seamless. For example, HCP003 illustrated a good experience of using telehealth involving both a patient and their carer in Figure 7, and recounted:

> *… essentially showing how we can make it work when everyone is capable of using telehealth. On the right … that's actually the patient's son … But it just shows how it can work. So … having patient caregiver in telehealth, it is all very possible. The video worked well the audio worked well ... there was no barrier there.* [HCP003]

XSL•FO
RenderX

Figure 7. Positive experience of telehealth (HCP003).



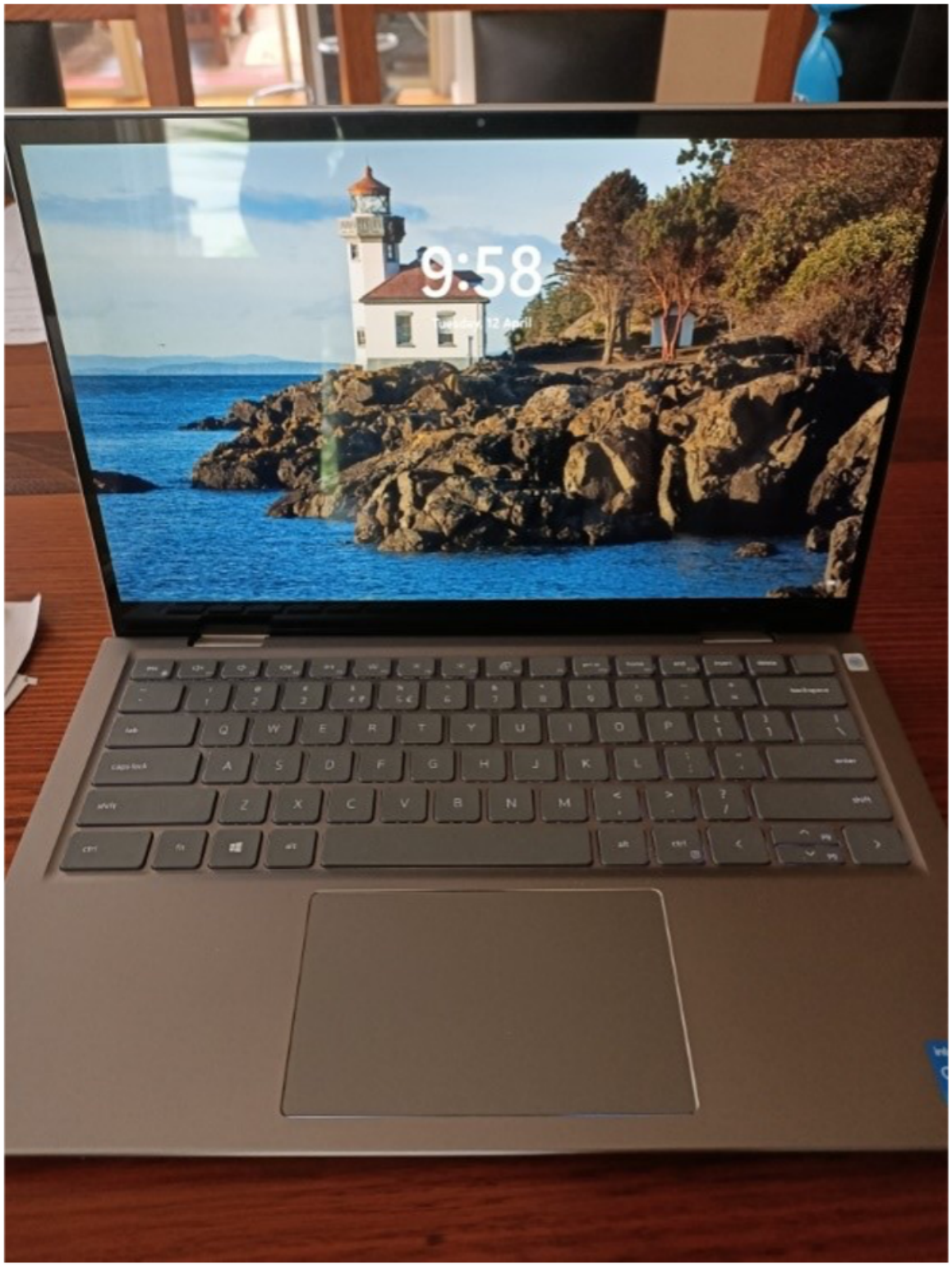
Interpersonal Connection

Participants expressed a general preference for video consultations over telephone consultations because of more opportunities for eye contact, and reading of facial expressions or other nonverbal cues. The notion of power dynamics between HCPs and patients also emerged, with some participants stating that conducting the consultations over telehealth made patients feel more relaxed and open, as opposed to feeling nervous in front of their HCP in person. P012 provided an image of their laptop (Figure 8), which they used for video consultations and stated:

> *I've noticed for me that I open up and talk more ... But when I'm in his rooms, I'm intimidated when I'm in person with him, face to face. I'm more conscious of myself and self-conscious. Whereas over the computer I'm not. I'm really confident and I just pour everything out.* [P012]

XSL•FO
RenderX

Figure 8. Patient's laptop used for video consultations makes them feel more relaxed (P012).



The type of technology used for video consultations also played a role in the quality of the clinical conversations. As previously mentioned, participants sometimes had to rely on backup options such as mobile phones or landlines for consultations due to technological failures, which also impacted the quality of their telehealth consultations. HCP011 shared a photo of an outreach service consultation that they had joined using their mobile phone, because the laptop and desktop were being used by other HCPs (Figure 9), and reflected:

> *The pictures are very small. So it just does not give you a realistic feeling of seeing a patient.* [HCP011]

XSL•FO
RenderX

Figure 9. Video consultation conducted on a mobile phone (HCP011).

XSL•FO
RenderX

In addition, participants generally agreed that periodic in-person consultations interspersed between their regular telehealth appointments were necessary to generate rapport between patients and HCPs. P015 had video telehealth oncology appointments every 6 - 8 weeks for over a year, despite which they said, "I've never met my oncologist," by which they meant that they had never met their oncologist in person. Even P012, who felt they were able to be more open through telehealth said:

> *the doctors should touch base with you specifically in person after a period of at least six months…. seeing them person to person sort of reconnects the relationship.* [P012]

In considering the benefits of in-person care, HCP010 said:

> *…we still want that relationship with the patient, we want to look them in the face and an important aspect of interpreting is body language. So even if you see them on the screen, [it] is not the same.* [HCP010]

HCP005 explained that some patients, "feel a bit suspicious or not completely trusting when you are not there [in person]."

Separately, HCP004 and HCP005 both also said that breaking bad news was more difficult through telehealth because of their inability to offer comfort through nonverbal communication like touch.

Care Quality Impacts of Telehealth

This final theme elucidates participants' perceptions about how telehealth impacts the quality of health care for patients.

Patient participants expressed mixed feelings regarding their satisfaction with the telehealth model of health care. While some participants perceived that the care they received through telehealth was satisfactory, others expressed concerns around issues such as lack of physical examination, impact on treatment timeliness, and communication.

Patients who described telehealth care positively viewed the crucial elements of care as unaffected. P014 explained that being able to share documents and show the HCP medications meant that:

> *[telehealth] doesn't stop the communication process in anyway…as [a] patient I'm not missing anything.* [P014]

Participant P014 was also able to monitor their blood pressure and oxygen saturation at home and share this information with their HCP. HCP004 saw telehealth as working well when there was another complementary team, for example, "hospital-in-the-home," which offered in-person care and provided clinical information to the palliative care patients. They also described the benefits of using telehealth to determine when an in-person visit was necessary:

> *it works best where there's opportunity for, or being able to identify that you can't rely on telehealth alone.* [HCP004]

Participants who believed that the lack of an in-person examination had compromised their care were more negative about telehealth. For example, P013 shared a photo of their arm in their lymphoedema sleeve (Figure 10) to describe the way telehealth had limited their care:

> *I was seeing her [my HCP] pretty much monthly on telehealth. As my arm would flare up, there was still not much she could do because we weren't allowed in … she finally got face to face with me, measured me up for the sleeve, so that was six months without proper diagnosis.* [P013]

Figure 10. Lymphoedema sleeve (P013).



XSL•FO
RenderX

This participant perceived that their symptoms had not been taken as seriously or treated as effectively because the HCPs involved had not been able to examine them.

P016 used a photo of their skin (Figure 11) to discuss the implications of limited physical examination opportunities in telehealth:

> *I was trying to explain to the doctor what the problem was … he actually couldn't see me, what I was trying to explain he couldn't see … And we didn't really get to the bottom of what it was.* [P016]

Figure 11. Patient finding it difficult to show the health care professional (HCP) her skin redness concern (P016).



For this participant, in-person care was important not only in this specific instance, but also because:

> *the doctor would pick up from the last time he saw you or your oncologist who last saw you [on] sort of things that you don't notice in yourself … I just feel as I'm getting more unwell, that I need to have more face-to-face appointments than telehealth appointments, but they tend to push you towards telehealth.* [P016]

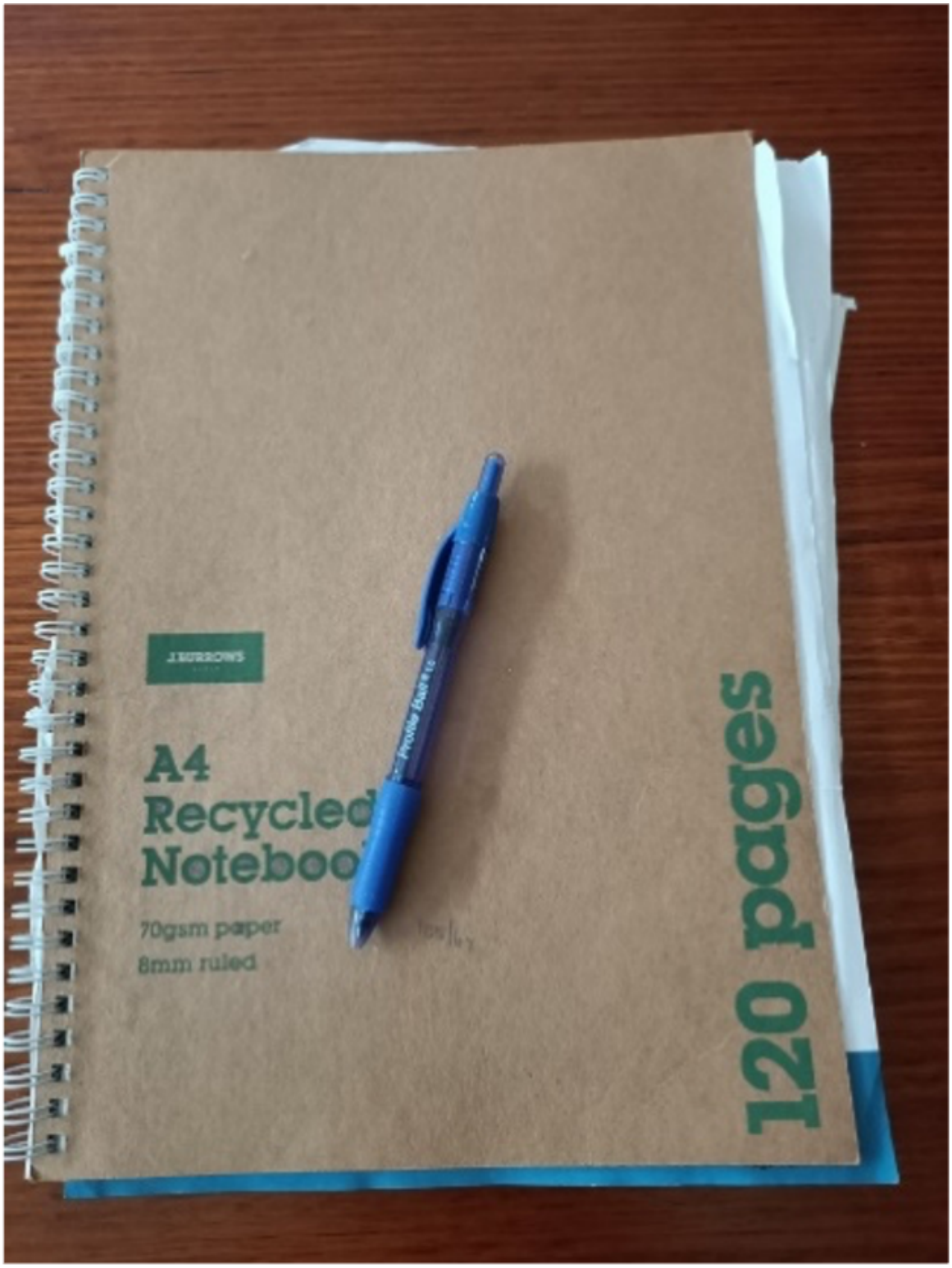
HCPs also stated that telehealth sometimes resulted in less thorough assessments. HCP001 explained that this may be the result of poor technical connection, whereas HCP005 described it as a problem when the assessment required a third party to facilitate the telehealth consultation, such as staff in an aged care facility, whose other duties meant they did not have the requisite time to devote to the consultation.

On the other hand, a positive aspect of the telehealth model of care for patients was their ability to feel more in control of their self-care and self-management during telehealth consultations. For example, patient P012 took a photo of their notebook (Figure 12) to illustrate that they felt more comfortable bringing prewritten notes for discussion and scribing notes during a telehealth consultation than they would in person:

> *Notebook [is] very important ... Leading up to all my appointments I have, I make notes. And then when my appointment comes up, I cross stuff off … that's really important. For me, it's my notes. That way, I know I'm not going to forget anything. And whatever they say to me that's relevant or important, something I had to do or whatever, I'll write it down … I probably wouldn't do that at the doctor's … But online I can just do it on the side while I'm talking to them.* [P012]

XSL•FO
RenderX

Figure 12. Patient feels more in control of their ability to take notes before and during telehealth consultations (P012).



Similarly, another patient P009 recounted a similar experience of feeling more comfortable about their ability to capture important information before and during a telehealth consultation by having family members present with them from the comfort of their homes:

> *Part of the brain tumors is that I can't remember stuff afterwards. So I have to have someone there who's like my virtual notepad … they can answer things that I may have forgotten especially to do with medication because it seems to be the main one or [one of the] side effects … So it works well as in, they've [husband and/or daughter] got an area where they can be comfortable … I have a pad and pen here. But of course I have sometimes husband and/or daughter not only remembering stuff like that's said but going back and reminding me of stuff I need to tell the oncologist that I may have forgotten about.* [P009]

XSL•FO
RenderX

Discussion

Principal Findings

Our intertextual analysis yielded 3 key themes: comfort (or lack thereof) afforded by telehealth, connection considerations in telehealth, and care quality impacts of telehealth. Participants' relationship to telehealth was nuanced. Generally, participants acknowledged the myriad of benefits that telehealth brings, such as greater psychological and physical comfort for patients, especially those based in rural areas. However, there was also a general sense that telehealth should be combined with in-person care for optimal patient-HCP therapeutic alliance and better patient outcomes. It should be noted that this study's data collection commenced during the pandemic, amid nation-wide efforts to rapidly scale telehealth models of care across Australia. Eastman et al [15] sought to understand patients' and HCPs' experiences of telehealth in community palliative care. Their findings were congruent with ours, suggesting that telehealth is being increasingly used in palliative care and generally patients and HCPs are supportive of using this model of health care, combined with regular in-person touchpoints.

A salient finding from our study was the notion of greater power symmetry between patients and HCPs in telehealth compared with in-person consultations. Patients alluded to feeling more empowered, in control, and psychologically comfortable during teleconsultations. This translated into them feeling more confident about "pouring out" information in front of their HCPs, than they would face-to-face. Furthermore, the safety of a device's screen afforded patients the space to take notes during consultations. Having teleconsultations from the comfort of their homes where family members could also be present where needed, meant that patients felt more at ease. In some cases, family members also served the important function of remembering what had been discussed during consultations, which enhanced patients' quality of care.

Our team conducted additional co-design activities with consumers which are outside the scope of this article. In those activities too, participants had reported the importance of taking notes before, during, and after their telehealth appointments. Participants described forgetting important details from their appointments and relying on their family to attend consultations to remember what had happened during the appointment.

Literature suggests that 40% - 80% of medical information and recommendations provided by HCPs during clinical consultations is forgotten by patients almost immediately. The poor information recall may have downstream impacts for treatment adherence [24]. Issues of memory recall may be exacerbated in palliative care patients who experience complex illness symptomatology. There is some evidence to suggest that patient-facing clinical consultation recordings can be helpful for patients in general. A scoping review conducted by Tsulukidze et al [25] in 2014 suggested that patients place a high value on audio recordings of clinical consultations and benefit from subsequently listening to the consultation recordings.

The participants of our study reported in this paper and subsequent co-design activities indicated that having a summary of the consultation at the end of telehealth session would mean that this information is kept secure and in one place for patients. Memory and cognitive deficits are often experienced by palliative care patients. Therefore, a mechanism to record salient information discussed during consultations would be particularly useful for this patient population.

The literature on patient-facing consultation summaries generated during telehealth is currently limited. A non-systematic literature scan in Google Scholar on this subject yielded no relevant papers. The closest relevant literature appears to be about artificial intelligence (AI) enabled digital scribes or documentation systems, which can help automate clinical documentation tasks ordinarily conducted by humans [26-28]. However, the literature on patient-facing clinical summaries in telehealth (whether generated manually by HCPs or through AI-enabled algorithms) is not yet developed.

The paucity of literature on this subject combined with the participants' unanimous views emergent in our investigations, emphasized the need for innovating in the area of patient-facing clinical consultation summaries. Since the conduct of the study presented in this paper, our team of software developers has created a telehealth-enhancement add-on feature which seeks to generate patient-facing summaries of clinical consultations during video telehealth sessions. The prototype was developed in an agile manner, involving think-aloud testing activities with end users. Subsequently, the minimum viable product (MVP) was subjected to a simulation study to generate early-stage evidence about end-users' perceptions of the add-on's potential clinical usefulness. The findings of our think-aloud and simulation studies will be reported in future papers. However, based on participants' accounts presented in this paper, we expect that the provision of a patient-facing document that summarizes the telehealth consultation and distills key actionable clinical advice for patient self-management, might enhance patient experience and self-management.

Limitations

The limitations of this project include the small number of participants interviewed. The small sample size was partly due to challenges in recruiting palliative care patients. Patients' complex symptomatology and in some cases, limited life expectancy, impacts their ability to participate in research studies. In addition, as the data collection was conducted during the pandemic, this posed additional challenges for HCP recruitment. Nevertheless, our multimodal data collection approach involving the use of photographs and interviews combined with dual perspectives of patients and HCPs meant that we had a rich dataset to generate an in-depth understanding of users' experiences of telehealth.

In this instance, we did not collect information about, or pose eligibility restrictions regarding participants' ethnic and racial background, socioeconomic status, and level of experience with telehealth. This decision was driven by the small number of palliative care participant population available to participate in the study. Future telehealth research may include

XSL•FO
RenderX

contextualisation of emergent findings through the lens of one or more of these socio-demographic characteristics.

Conclusion

This study explored patients' and HCPs' experiences of telehealth in a palliative care context. A total of 3 themes emerged: comfort (or lack thereof) afforded by telehealth, connection considerations in telehealth, and care quality impacts of telehealth. The findings presented in this article combined with other co-design activities, which are outside the scope of this paper, indicated the potential value of a telehealth enhancement add-on feature that generates patient-facing clinical consultation summaries. Our team has developed a video telehealth enhancement feature, which will enable clinicians to distill key actionable advice and self-management guidance discussed during teleconsultations in a take-home summary document for patients. The add-on prototype has been subjected to a preliminary simulation study which will be reported in a future publication.

Acknowledgments

We thank Ms Roxanne Kritharidis who conducted all of the interviews, cleaned transcripts, and prepared the data for analysis. We thank Dr Jennifer Bibb who obtained ethics approvals, supported data collection, and participated in data analysis and writing of the final co-design report on which this work is based. We thank Dr Elton Lobo who contributed to analysis discussions and started early paper drafts. Thank you to Dr Kate Robins-Browne who led data analysis and drafted first versions of this paper. We also thank Prof Victoria Palmer who designed the photo interviewing study as part of the co-design stage for the early parts of this project, oversaw data collection and analysis, and contributed to the early draft of the paper. This research is supported by the Digital Health Cooperative Research Centre Limited (DHCRC), Monash University, Healthdirect Australia, Monash Health, University of Melbourne, and Victorian Department of Health. DHCRC is funded under the Commonwealth Government Cooperative Research Centres (CRC) Program. Finally, we thank all the participants of this study who generously contributed their time towards this research.

Data Availability

The datasets generated during and analyzed during this study are not publicly available due to the sensitive nature of the data and vulnerable nature of the population.

Generative artificial intelligence was not used for any part of the study reported in this paper nor for manuscript development.

Authors' Contributions

MK contributed to formal analysis, writing–original draft, writing–review and editing, and project administration. TOB contributed to formal analysis, writing–original draft, writing–review and editing, and project administration. CP and MEF contributed to writing–review and editing. CB, RH, and PP contributed to writing–review and editing and funding acquisition. XC, AL, and SG contributed to writing–review and editing and software. OM and KH contributed to writing–review and editing and project administration.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Photo instructions—health care professionals (HCPs).
[PDF File, 246 KB - humanfactors_v12i1e53913_app1.pdf]

Multimedia Appendix 2
Photo instructions—patients.
[PDF File, 202 KB - humanfactors_v12i1e53913_app2.pdf]

Multimedia Appendix 3
The 3 Cs storyboard.
[PDF File, 1357 KB - humanfactors_v12i1e53913_app3.pdf]

References

1. Harris MF, Harris-Roxas B, Knight AW. Care of patients with chronic disease: achievements in Australia over the past decade. Med J Aust 2018 Jul 16;209(2):55-57. [doi: 10.5694/mja18.00333] [Medline: 29996749]

XSL•FO
RenderX

2. Health conditions prevalence - key findings on selected long-term health conditions and prevalence in australia. Australian Bureau of Statistics. URL: https://www.abs.gov.au/statistics/health/health-conditions-and-risks/health-conditions-prevalence/latest-release [accessed 2022-10-19]
3. Deaths in Australia. Australian Institute of Health and Welfare. URL: https://www.aihw.gov.au/reports/life-expectancy-death/deaths-in-australia/contents/leading-causes-of-death [accessed 2022-10-19]
4. Earp M, Cai P, Fong A, et al. Hospital-based acute care in the last 30 days of life among patients with chronic disease that received early, late or no specialist palliative care: a retrospective cohort study of eight chronic disease groups. BMJ Open 2021 Mar 24;11(3):e044196. [doi: 10.1136/bmjopen-2020-044196] [Medline: 33762238]
5. National Palliative Care Strategy 2018. Australian Commission on Safety and Quality in Health Care. URL: https://www.safetyandquality.gov.au/sites/default/files/2019-06/national_palliative_care_strategy_2018.pdf [accessed 2022-10-19]
6. Wajid M, Rajkumar E, Romate J, George AJ, Lakshmi R, Simha S. Why is hospice care important? An exploration of its benefits for patients with terminal cancer. BMC Palliat Care 2021 May 17;20(1):1-8. [doi: 10.1186/s12904-021-00757-8] [Medline: 34001076]
7. Otani H, Morita T, Igarashi N, Shima Y, Miyashita M. A nationwide survey of bereaved family members' perception of the place patients spent their final days: is the inpatient hospice like or unlike a home? Why? Palliat Med Rep 2020;1(1):174-178. [doi: 10.1089/pmr.2020.0058] [Medline: 34223474]
8. Currow DC, Agar MR, Phillips JL. Role of hospice care at the end of life for people with cancer. J Clin Oncol 2020 Mar 20;38(9):937-943. [doi: 10.1200/JCO.18.02235] [Medline: 32023154]
9. Shrank WH, Russell K, Emanuel EJ. Hospice carve-in-aligning benefits with patient and family needs. JAMA 2020 Jul 7;324(1):35-36. [doi: 10.1001/jama.2020.8459] [Medline: 32496529]
10. de Avila V, Paik JM, de Avila L, et al. Hospice care utilisation among elderly patients who died with hepatocellular carcinoma in the United States. JHEP Rep 2021 Apr;3(2):100236. [doi: 10.1016/j.jhepr.2021.100236] [Medline: 33748729]
11. Mitchell GK. Palliative care in Australia. Ochsner J 2011;11(4):334-337. [Medline: 22190884]
12. Pasanen L, Le Gautier R, Wong A, et al. Telehealth in outpatient delivery of palliative care: a qualitative study of patient and physician views. Pall Supp Care 2023 Dec;21(6):980-987. [doi: 10.1017/S1478951522000670]
13. Fitzgerald DA, Nunn K, Isaacs D. Consequences of physical distancing emanating from the COVID-19 pandemic: an Australian perspective. Paediatr Respir Rev 2020 Sep;35:25-30. [doi: 10.1016/j.prrv.2020.06.005] [Medline: 32690355]
14. Jiang B, Bills M, Poon P. Integrated telehealth-assisted home-based specialist palliative care in rural Australia: a feasibility study. J Telemed Telecare 2023 Jan;29(1):50-57. [doi: 10.1177/1357633X20966466] [Medline: 33079611]
15. Eastman P, Dowd A, White J, Carter J, Ely M. Telehealth: rapid adoption in community palliative care due to COVID-19: patient and professional evaluation. BMJ Support Palliat Care 2024 May;14(e1):e438-e442. [doi: 10.1136/bmjspcare-2021-002987]
16. Daveson BA, de Wolf-Linder S, Witt J, et al. Results of a transparent expert consultation on patient and public involvement in palliative care research. Palliat Med 2015 Dec;29(10):939-949. [doi: 10.1177/0269216315584875] [Medline: 25931336]
17. Burton A, Hughes M, Dempsey RC. Quality of life research: a case for combining photo-elicitation with interpretative phenomenological analysis. Qual Res Psychol 2017 Oct 2;14(4):375-393. [doi: 10.1080/14780887.2017.1322650]
18. Bugos E, Frasso R, FitzGerald E, True G, Adachi-Mejia AM, Cannuscio C. Practical guidance and ethical considerations for studies using photo-elicitation interviews. Prev Chronic Dis 2014 Oct 30;11:E189. [doi: 10.5888/pcd11.140216] [Medline: 25357257]
19. Glaw X, Inder K, Kable A, Hazelton M. Visual methodologies in qualitative research: autophotography and photo elicitation applied to mental health research. Int J Qual Methods 2017;16(1). [doi: 10.1177/1609406917748215]
20. Todd A, Holmes H, Pearson S, et al. "I don't think I'd be frightened if the statins went": a phenomenological qualitative study exploring medicines use in palliative care patients, carers and healthcare professionals. BMC Palliat Care 2016 Jan 29;15(1):1-7. [doi: 10.1186/s12904-016-0086-7] [Medline: 26822776]
21. Cederholm EA. The use of photo‐elicitation in tourism research – framing the backpacker experience. Scandinavian Journal of Hospitality and Tourism 2004 Jan;4(3):225-241. [doi: 10.1080/15022250410003870]
22. Pyyry N, Hilander M, Tani S. Photography and photo elicitation as visual methods. In: Creative Methods for Human Geographers: SAGE Publications; 2021:75-86. [doi: 10.4135/9781529739152.n6]
23. Palmer VJ, Furler J. A room with a view: a metaphor analysis of Vietnamese women's representations of living with depression using photo elicitation. Visual Studies 2018 Jul 3;33(3):251-263. [doi: 10.1080/1472586X.2018.1525309]
24. Kessels RPC. Patients' memory for medical information. J R Soc Med 2003 May;96(5):219-222. [doi: 10.1177/014107680309600504] [Medline: 12724430]
25. Tsulukidze M, Durand MA, Barr PJ, Mead T, Elwyn G. Providing recording of clinical consultation to patients - a highly valued but underutilized intervention: a scoping review. Patient Educ Couns 2014 Jun;95(3):297-304. [doi: 10.1016/j.pec.2014.02.007] [Medline: 24630697]
26. Coiera E, Kocaballi B, Halamka J, Laranjo L. The digital scribe. NPJ Digit Med 2018;1(1):58. [doi: 10.1038/s41746-018-0066-9] [Medline: 31304337]
27. Ghatnekar S, Faletsky A, Nambudiri VE. Digital scribe utility and barriers to implementation in clinical practice: a scoping review. Health Technol (Berl) 2021;11(4):803-809. [doi: 10.1007/s12553-021-00568-0] [Medline: 34094806]

XSL•FO
RenderX

28. van Buchem MM, Boosman H, Bauer MP, Kant IMJ, Cammel SA, Steyerberg EW. The digital scribe in clinical practice: a scoping review and research agenda. NPJ Digit Med 2021 Mar 26;4(1):57. [doi: 10.1038/s41746-021-00432-5] [Medline: 33772070]

Abbreviations

AI: artificial intelligence
HCP: health care professional
MVP: minimum viable product

Edited by A Kushniruk; submitted 13.11.23; peer-reviewed by G Carolus, L Rotteau, P Koppel; revised version received 30.08.24; accepted 14.10.24; published 11.02.25.

Please cite as:
Kalla M, O'Brien T, Metcalf O, Hoda R, Chen X, Li A, Parker C, Franco ME, Georgy S, Huckvale K, Bain C, Poon P
Understanding Experiences of Telehealth in Palliative Care: Photo Interview Study
JMIR Hum Factors 2025;12:e53913
URL: https://humanfactors.jmir.org/2025/1/e53913
doi:10.2196/53913

© Mahima Kalla, Teresa O'Brien, Olivia Metcalf, Rashina Hoda, Xiao Chen, Andy Li, Catriona Parker, Michael Edward Franco, Sam Georgy, Kit Huckvale, Christopher Bain, Peter Poon. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 11.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Effects of Digital eHealth Versus Onsite 2-Day Group-Based Education in 255 Patients With Irritable Bowel Syndrome: Cohort Study

Birgitte Berentsen[1,2], MSc, PhD; Camilla Thuen[1], MSc; Eline Margrete Randulff Hillestad[1,2], MSc; Elisabeth Kjelsvik Steinsvik[2], MD, PhD; Trygve Hausken[1], MD, PhD; Jan Gunnar Hatlebakk[1,2], MD, PhD

[1]Department of Clinical Medicine, University of Bergen, Bergen, Norway

[2]Department of Medicine, National Center for Functional Gastrointestinal Disorders, Haukeland University Hospital, Bergen, Norway

Corresponding Author:
Birgitte Berentsen, MSc, PhD
Department of Clinical Medicine, University of Bergen, Bergen, Norway

Abstract

Background: Irritable bowel syndrome (IBS) has a high worldwide prevalence and there are few effective treatment options. Patient education can influence patient behavior that subsequently may lead to changes in attitudes and skills necessary for maintenance or improvement in management of symptom severity and quality of life. However, as postdiagnostic patient education can be resource demanding, assessment of digital approaches and verification of their effectiveness is warranted.

Objective: This cohort study aimed to investigate the effects of a digital web-based multidisciplinary eHealth program on the domains of symptom severity (Irritable Bowel Syndrome Symptom Severity Scale [IBS-SSS]), quality of life (irritable bowel syndrome quality of life [IBS-QOL]), anxiety and depression (Hospital Anxiety and Depression Scale), and a measure of general client satisfaction (client satisfaction questionnaire), compared with an onsite multidisciplinary 2-day group-based education program ("IBS-school"), in 2 cohorts of 255 patients with IBS.

Methods: Patients diagnosed with IBS, aged 15-70 years, were enrolled after referral to the Section of Gastroenterology at Haukeland University Hospital, Norway. In total, 132 patients were recruited to the eHealth program and 123 to the IBS-school group for comparison. Data were self-reported and collected digitally at enrollment and after 3 months, between 2017 and 2019. Furthermore, 71 attending the eHealth program and 49 attending the IBS-school completed the questionnaires at 3 months. Intervention response was defined as a reduction of ≥50 points on the IBS-SSS.

Results: Patients attending the eHealth program reported a significant reduction in IBS symptom severity 3 months after treatment (n=71), compared with patients attending the IBS-school (n=50). Overall, patients categorized as intervention responders in both programs showed a significant reduction in symptom severity at 3 months. Here, 41% (29/71) of patients attending the eHealth program reported a mean IBS-SSS reduction of 103 (SD 72.0) points (P<.001). In addition, these patients reported reduced anxiety (P>.001) and depression (P=.002) and enhanced quality of life (P=.03), especially the degrees of dysphoria, body image, food avoidance, health worry, interference with activity, relations, and social relations. Patients responding to the IBS-school intervention (18/50, 36%) reported a mean IBS-SSS reduction of 119 (SD 86.2) points (P<.001), and reduced depression scores (P=.046), but no difference in overall quality of life. Both groups reported the respective interventions as "good" quality health care programs, scoring them 23.5 (SD 4)—the eHealth program 23.5 (SD 4), and the IBS-school 24.2 (SD 4)—on the client satisfaction questionnaire.

Conclusions: We conclude that the digital multidisciplinary eHealth program has a significant effect on IBS symptom severity in a portion of patients; it is useful as a tool in disease self-management and does not result in worse symptom scores than an onsite multidisciplinary 2-day group-based education program after 3 months. We believe these results indicate that a digital eHealth approach is preferable to an onsite multidisciplinary 2-day group-based education program covering the same topics.

(JMIR Hum Factors 2025;12:e43618) doi:10.2196/43618

KEYWORDS

irritable bowel syndrome; IBS; eHealth; internet-guided; patient education; self-management; self-reported; patient behavior; quality of life; QOL; anxiety; depression; gastrointestinal; physiotherapist; kinesiology; cognitive behavioural therapy; CBT; Hospital Anxiety and Depression Scale; HADS; client satisfaction questionnaire; CSQ; Mann-Whitney U test; nonparametric; Wilcoxon test; neurogastroenterology

XSL•FO
RenderX

Introduction

Irritable bowel syndrome (IBS) is a chronic disorder manifested by recurrent abdominal pain and alterations in stool form or frequency [1,2]. The condition affects between 4% and 9.2% of the global population [3-5] and it is highly heterogenous. IBS' unclear etiology involves multifactorial disturbances of the bidirectional communication between the gut and the brain, including visceral hypersensitivity, low-grade inflammatory responses, intestinal motility disturbances, alterations of central nervous system processing, and alterations in gut microbiota composition [6]. However, no clear biomarker or therapeutic target for IBS has been identified. The condition lacks both a cure and medication that gives sufficient symptom relief, a fact that highlights the necessity of integrating nonpharmacological approaches including patient education in patient care [7]. Patients with IBS require personalized treatment for successful symptom relief. Approaches may include physical therapy, cognitive behavioral therapy (CBT), hypnotherapy, mindfulness and exposure therapy, and comprehensive dietary guidance by registered dietitians such as the low FODMAP (fermentable oligosaccharides, disaccharides, monosaccharides, and polyols) diet. However, access to these treatment options is often limited due to lack of trained professionals, travel distances, and cost. Thus, more accessible treatment options are warranted. Web-based interventions, patient education, and self-management has been demonstrated to be effective in patients suffering from chronic diseases, including IBS [8-10]. A systematic review from 2017 concluded with mixed results regarding the effectiveness of web-based mindfulness-based interventions compared with active control treatment conditions such as CBT. However, the study showed that treatment targeting symptoms of IBS had the largest effect size improvements [11]. A longitudinal qualitative study from 2020 showed that both telephone-based CBT and web-based CBT for IBS were positively received and had lasting positive impacts on participants' understanding of IBS symptoms, quality of life, and IBS-related behaviors [12]. A recent Japanese randomized controlled trial has shown that a multidisciplinary eHealth self-management program can reduce the severity of IBS-symptoms and improved the quality of life [13].

Indeed, limited health care resources, national priority guidelines, and sparse treatment options may restrict any long-term follow-up that patients with IBS may request from a secondary or tertiary care institution. Since 2012, Haukeland University Hospital has offered patients with IBS an onsite multidisciplinary 2-day group-based education program, a so-called "IBS-school." The multidisciplinary approach is designed to provide patients with evidence-based information and practical skills that may lead to improved IBS management, reduced IBS symptoms, and enhanced quality of life. Based on our clinical experience with group education, we have developed a novel internet-guided multidisciplinary self-care management program for patients with IBS, from now on referred to as the "eHealth program." The content in the eHealth program is based on topics covered in the IBS-school, but the patients also have access to digital clinical support during the program. Herein, we hypothesized that the eHealth program would have an equally good effect on symptom severity, quality of life, and patient satisfaction, compared with a cohort of patients attending IBS-school. Data from neither program have been published before. In this study, we aimed to investigate the effects of a digital web-based multidisciplinary eHealth program on the domains of symptom severity (Irritable Bowel Syndrome Symptom Severity Scale [IBS-SSS]), quality of life (irritable bowel syndrome quality of life [IBS-QOL]), anxiety and depression (Hospital Anxiety and Depression Scale [HADS]), and a measure of general client satisfaction (client satisfaction questionnaire [CSQ-8]), compared with an onsite multidisciplinary 2-day group-based education program, IBS-school, in 2 cohorts of 255 patients with IBS.

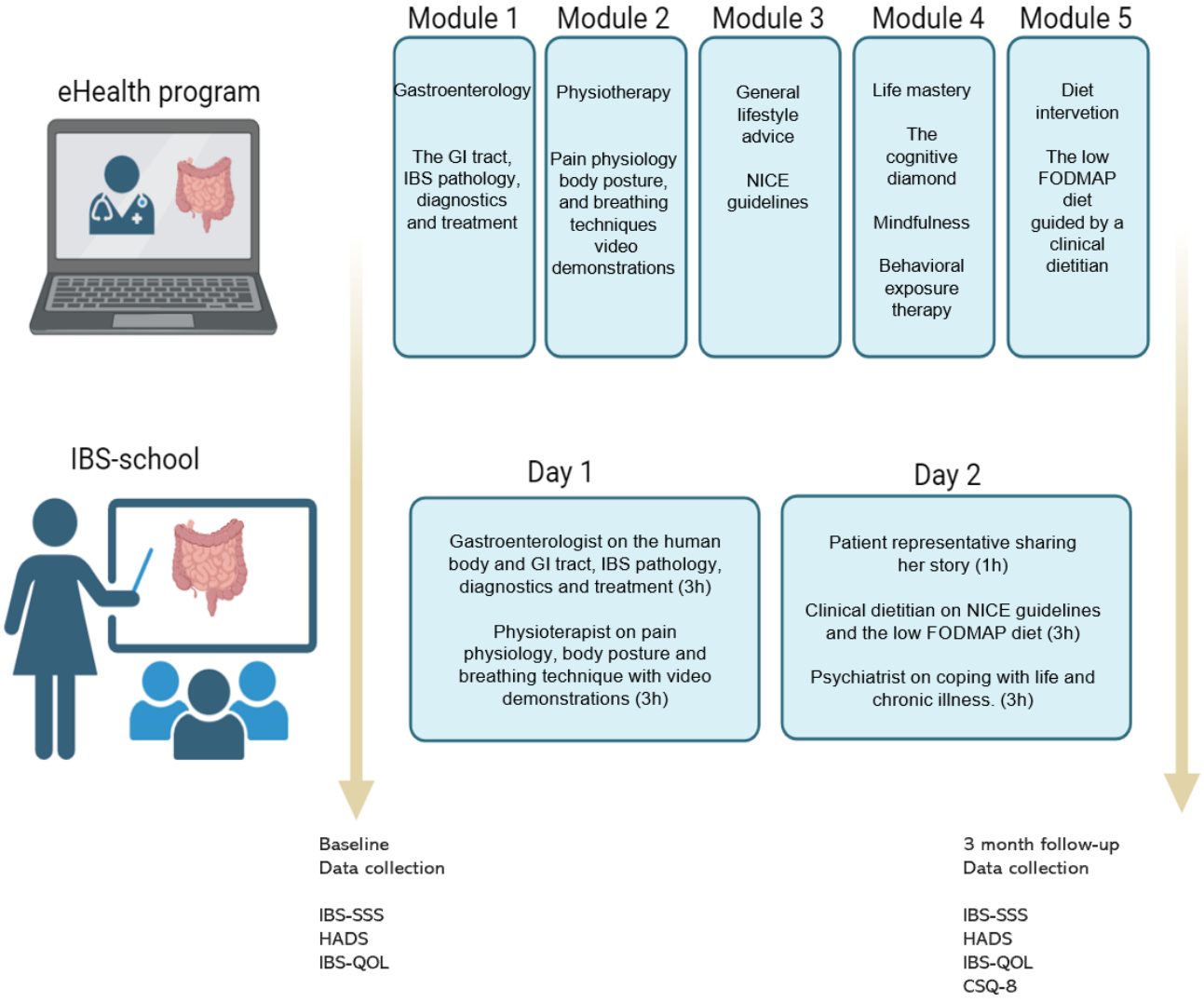
Methods

Patient Sample, Randomization, and Treatments

In this study**,** 255 patients between 15 and 70 years were recruited after being accepted for patient education at Haukeland University Hospital in Norway between 2017 and 2019. Random patients on the waiting list to the onsite multidisciplinary 2-day group-based education program, IBS-school, were contacted by phone by a study nurse and offered to attend the novel multidisciplinary digital 5-module eHealth program, from here on referred to as the "eHealth program." Patients attending the IBS-school were recruited on site when attending the 2-day group-based education program. Patients received oral and written information before giving written consent and sent it to the hospital by post. Furthermore, 123 patients that attended the IBS-school were used for comparison (N=255, 1:1). Here, patients were recruited by a study nurse on site. They received an oral and written information about the study before signing consent (Figure 1). Inclusion criteria included (1) being referred to receiving patient education on IBS by a gastroenterologist or general practitioner after the diagnosis of IBS had been determined (*International Classification of Primary Care, Second Editon* [*ICPC-2*] code D93; *International Statistical Classification of Diseases, Tenth Revision* [*ICD-10*] code K58), (2) being between 18 and 70 years, and (3) understanding written and oral Norwegian. There were no specific exclusion criteria.

XSL•FO
RenderX

Figure 1. Intervention overview. CSQ-8: client satisfaction questionnaire; FODMAP: fermentable oligosaccharides, disaccharides, monosaccharides, and polyols; GI: gastrointestinal; HADS: Hospital Anxiety and Depression Scale; IBS: irritable bowel syndrome; IBS-QOL: irritable bowel syndrome quality of life; IBS-SSS: Irritable Bowel Syndrome Symptom Severity Scale; NICE: National Institute for Health and Care Excellence.



Ethical Considerations

Eligible patients gave informed written consent and participants were given the option to withdraw from the study at any time point without a specific reason. The data were deidentified and stored on a secure hospital server, and analysis was performed on anonymous data. None of the participants were compensated. The study was approved by the Regional Ethical Committee of Western Norway (REC-2016/1098).

Interventions

The Irritable Bowel Syndrome–School

Patients attended an onsite multidisciplinary 2-day group-based education program, IBS-school, at Haukeland University Hospital. The number of participants varied between 15 and 40 each month. The program involved lectures and question and answer sessions by 4 health care professionals. Refer to Figure 1 for the intervention outline. On day 1, a gastroenterologist talked about the human body and the gastrointestinal system, what IBS is, what causes it, diagnostics, and treatment options (3 hours). Second, a physiotherapist was giving a lecture on body posture and breathing techniques including demonstrations, introducing pain physiology, and the function of the human nervous system (3 hours). On day 2, a patient representative shared her personal experience with IBS (1 hour), and a clinical dietitian gave a lecture on National Institute for Health and Care Excellence (NICE) guidelines [14] that summarizes the most recent recommendations on IBS in adults in primary care, and the low FODMAP diet [15] (3 hours). Finally, a psychiatrist gave a lecture on coping with life including health worry, social relations, tiring thoughts and feelings, adaptations, and symptoms. A nurse with specialization in gastroenterology hosted the group education program to create a comfortable environment for exchange of personal experiences and participate in group assignments.

The eHealth Program

Patients who were enrolled in the digital eHealth program had access to a comprehensive, multidisciplinary web-based program that consisted of 5 modules (refer to Figure 1 for the outline). The modules consisted of instructive texts, videos, animations, and images over 150 web-pages on a digital treatment platform by CheckWare AS [16]. In addition, patients carried out “home assignments” based on principles of CBT, a protocol by Ljótsson et al [17] at the Swedish Karolinska Institute, followed by an

XSL•FO
RenderX

optional low FODMAP diet intervention guided by a clinical dietitian. The eHealth program allowed a secure login (eg, bank-ID) and digital communication between patient and a clinical dietitian throughout the process, at the patient's own request and need. The patients were expected to finish the program over a period of 3 - 12 weeks, at his or her own pace, all dependent on their motivation and work capacity. Module contents have been described further in this study. Module 1 describes the body and the gastrointestinal system. In this first module, the patient gets an introduction to what IBS is, how it is diagnosed, and what causes it. The patient learns about the function of our digestive system and how it is regulated, and how it can be disturbed in people with IBS. Module 2 describes the posture and breathing techniques. This module focuses on the connection between IBS and musculoskeletal disorders. Many people with IBS may have a "hyperactive" nervous system that can cause pain and physical maladjustments. In this module, a physiotherapist introduces the patient to pain physiology and how the nervous system works. There is a practical section with useful exercises for people with IBS. The module will give the patient an understanding of how long-term pain occurs, why it often persists, and how to influence it. Module 3 consists of diet and lifestyle advices. There is no miracle cure for IBS, but many people experience improvement by following some general advice. In this module, we look at lifestyle advice that has been shown by research to improve symptoms in people with IBS. It is recommended to try this before eliminating other foods or following strict diets. Module 4 focuses on coping with life. In this module, the patient learns techniques from cognitive therapy including "the cognitive diamond," mindfulness, and practice systematic exposure exercises has previously shown beneficial effects for patients with IBS. The protocol has been described in a study by Ljótsson et al [17]. Module 5 describes the dietary intervention with a low FODMAP diet. This is a dietary treatment that provides symptom relief in approximately 70% of patients with IBS [18]. In this module, the patient will be introduced to the low FODMAP diet in both theory and practice. The patient had access to digital guidance by a clinical dietitian during the entire course of the study.

Questionnaires

Overview

Patients completed questionnaires related to IBS symptom severity (IBS-SSS) [2], quality of life, (IBS-QOL) [19], and anxiety and depression (HADS) [20] upon enrollment at baseline and after 3 months.

The IBS-SSS is considered the gold standard measure of IBS symptoms and contains 5 questions that measure the frequency of abdominal pain, the severity of abdominal distention, dissatisfaction with bowel habits, and interference with quality of life, scored in the range of 0 - 500. A higher score indicating worse condition, scores <175 represent mild IBS symptoms, 175 - 300 represents moderate severity, scores >300 represent severe IBS [2].

The IBS-QOL is a condition-specific measure for assessing health-related quality of life in IBS. It consists of 34 items, each with a 5-point response scale in a range of 0 - 100, where the higher score indicates a better IBS specific quality of life. There are 8 subscale scores for the IBS-QOL: dysphoria, interference with activity, body image, health worry, food avoidance, social reaction, sexual, and relationships [21].

HADS is a scale of 14 items designed to measure anxiety and depression, 7 items for each subscale (ie, anxiety and depression). The total score is the sum of the 14 items, and for each subscale the score is the sum of the respective 7 items that range from 0 to 21 [20].

CSQ-8 [22] was completed at 3 months. The questionnaire is an 8-item measure with a total score ranging from 8 to 32, with the higher number indicating greater satisfaction with treatment.

Primary Outcome Measure and Definition of Intervention Responders

In the field of IBS research, a change of 50 points in IBS-SSS score has shown to reliably indicate improvement in IBS symptom severity [2]. Our primary outcome measure was a >50-point reduction in IBS-SSS at 3 months, compared with baseline. Patients reporting a ≥50-point reduction in IBS-SSS at 3 months were categorized as "responders" to the intervention. Patients reporting <50-points on IBS-SSS were categorized as "nonresponders" to the intervention.

Statistical Analysis

Statistical analyses were performed using SPSS Statistics (version 26, IBM) for Microsoft Windows. Baseline patient characteristics and questionnaires were first presented according to intervention—eHealth group and IBS-school. At 3 months after intervention, patients were categorized as responders or nonresponders. For comparison between groups, unpaired *t* test were performed for parametric data and Mann-Whitney *U* for nonparametric data. For comparisons before and after interventions, paired *t* tests were performed for parametric data and Wilcoxon signed rank test for nonparametric data. A *P* value <.05 was considered statistically significant. The intervention responder or nonresponder analysis to the eHealth program or IBS-school was carried out performing a paired *t* test on data from patients who filled out the 3-month questionnaires (n=71 and n=50, respectively). A χ^2 test of independence was used to assess differences between intervention response.

Results

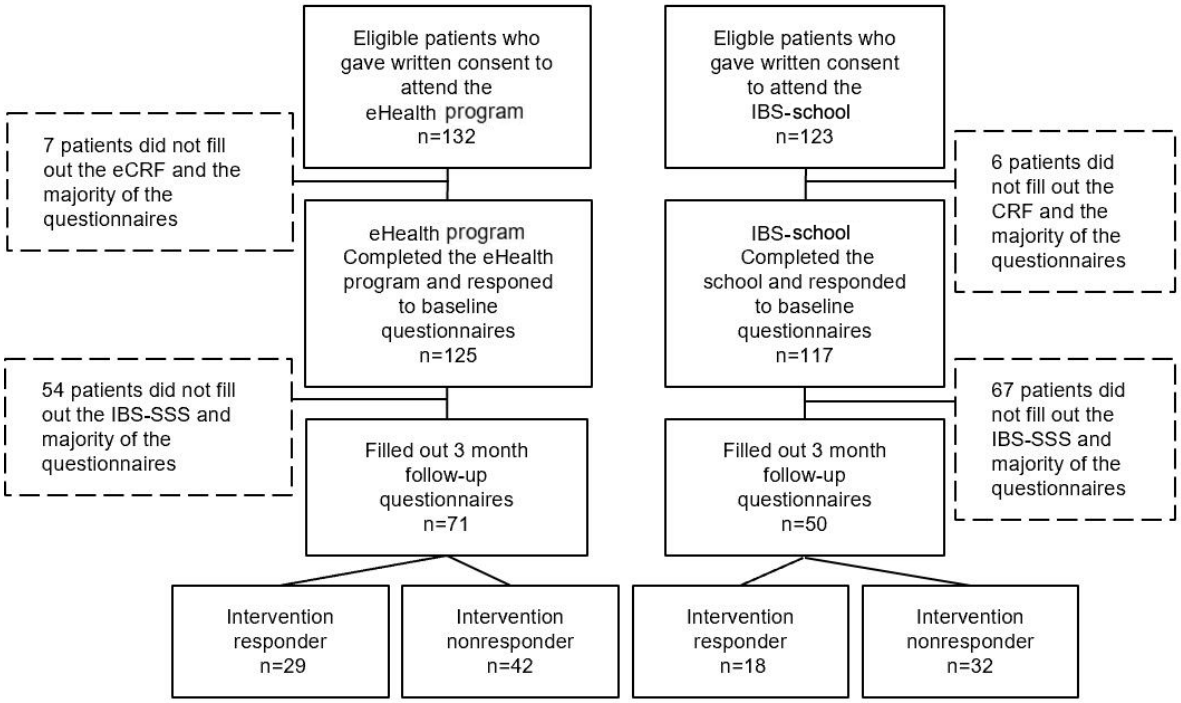
Patients and Characteristics

A flowchart of participating patients is illustrated in Figure 2. Of the 132 eligible patients who gave written consent to participate in the eHealth program group of the study, 7 did not fill out the minimum requirement including the electronic case report form and the IBS-SSS questionnaire. Of the 125 patients who completed the program, 54 patients did not respond to the IBS-SSS questionnaire at 3 months. Hence, the follow-up data from patients attending the eHealth program reduced to n=71 at 3 months. Of the 123 eligible patients who gave written consent to participate in the IBS-school group of the study, 5 did not fill out the minimum requirement including the electronic case report form and the IBS-SSS questionnaire. Thus, of the 118 patients who completed the program, 69 patients did not

XSL•FO
RenderX

respond to the IBS-SSS questionnaire at 3 months. Hence, the follow-up data from patients attending the eHealth program reduced to n=49 at 3 months.

Figure 2. Flow chart of participating patients with IBS. Patients reporting a ≥50-point reduction in the IBS-SSS at 3 months were categorized as "responders to the intervention". Patients reporting <50-points on the IBS-SSS were categorized as "nonresponders to the intervention". CRF: case report form; eCRF: electronic case report form; IBS: irritable bowel syndrome; IBS-SSS: Irritable Bowel Syndrome Symptom Severity Scale.



Participants were predominantly female (190/235, 81%) with a mean age of 38.3 (SD 12.4) years (Characteristics are summarized in Table 1). An unpaired *t* test showed that there was no difference in age between groups. A Mann-Whitney *U* test was performed to evaluate whether sex differed between groups, reveling no significant difference between patients attending the eHealth program and patients attending the IBS-school (Table S1 in Multimedia Appendix 1, U=6407.0, z=−1.394, P=.16). Furthermore, participants at the IBS-school and eHealth program displayed similar baseline characteristics upon enrollment, including moderate to severe IBS symptom severity.

There was no significant difference between most relevant features at baseline, except for anxiety and depression, as shown in Table 1. Here, patients attending the eHealth program scored an average of 11 (SD 5.2) which corresponds to mild to moderate anxiety, while patients attending the IBS-school scored 9.4 (SD 5.1), corresponding to a mild level of anxiety (2-tailed t_{232}=2.37, P=.02). Thus, both patient groups displayed an anxiety score typical of a clinical case of anxiety (HADS-anxiety ≥8 is considered a clinical case [20]), and patients attending the eHealth program presented a significantly higher baseline score than patients attending the IBS-school. Furthermore, patients attending the eHealth program reported significantly higher depression scores than patients attending the IBS-school (mean 7, SD 4.1 vs mean 5.9, SD 4, respectively; t_{232}=2.01, P=.045). However, both groups displayed an average depression score corresponding to nonclinical severity of depression (HADS-depression ≤8 is considered noncase [20]) which gives the difference little relevance on group level.

XSL•FO
RenderX

Table . Baseline patient characteristics and questionnaire data of 242 patients with irritable bowel syndrome (IBS) at enrollment to the eHealth program or IBS-school interventions. Unpaired *t* test unless otherwise stated; the reduced number from total number of participants indicate missing data.

	eHealth program		IBS-school		Significance	
	n	Mean (SD)	n	Mean (SD)	*t* test (*df*)	*P* value
Participant age	125	38.3 (12.4)	117	38.4 (13.4)	−0.06 (234)	.95
Gender						
Men	27	—[a]	18	—	—	—
Women	92	—	98	—	—	.16[b]
IBS symptom severity (IBS-SSS[c])	125	282.7 (82.5)	117	268.5 (87.9)	1.29 (240)	.20
Anxiety (HADS-A[d])	124	11 (5.2)	112	9.4 (5.1)	2.37 (232)	.02[e]
Depression (HADS-D[f])	124	7 (4.1)	112	5.9 (4)	2.01 (232)	.045[e]
IBS-QOL [g]	123		117			
Overall score		50.4 (21.9)		51.8 (21.0)	−0.48 (238)	.63
Body image		41.0 (21.3)		44.6 (22.9)	−1.27 (238)	.21
Dysphoria		47.0 (23.5)		50.0 (24.0)	−0.97 (238)	.33
Food avoidance		29.0 (22.2)		28.3 (21.9)	0.25 (238)	.80
Health worry		52.5 (22.9)		57.0 (23.0)	−1.46 (238)	.15
Interference with activity		45.0 (21.9)		44.2 (21.3)	0.32 (238)	.90
Relationships		56.7 (23.2)		57.6 (22.6)	−0.27 (238)	.79
Social relations		51.6 (22.1)		54.3 (22.6)	−0.93 (238)	.35
Sexual activity		57.8 (29.6)		57.8 (32.9)	−0.05 (238)	.96

[a]Not applicable.

[b]Mann-Whitney *U* test (Mean rank, *U*, *z* in Table S1 in Multimedia Appendix 1).

[c]IBS-SSS: Irritable Bowel Syndrome Symptom Severity Scale.

[d]HADS-A: Hospital Anxiety and Depression Scale – Anxiety.

[e]Significant at the $P<.05$ level.

[f]HADS-D: Hospital Anxiety and Depression Scale – Depression.

[g]IBS-QOL: irritable bowel syndrome quality of life.

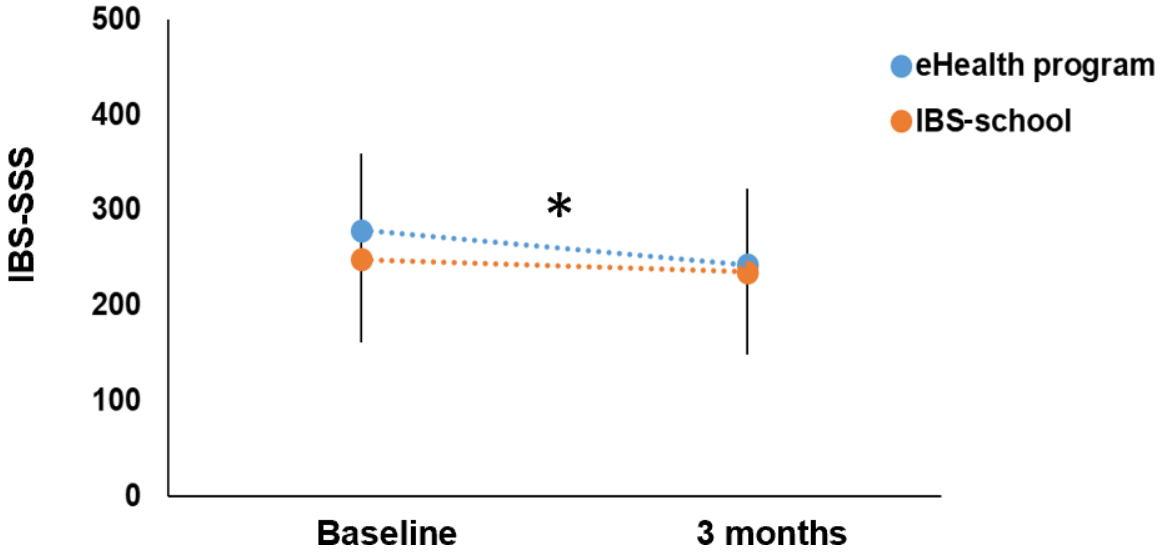
Symptom Relief and Enhanced Quality of Life

Overview

A paired *t* test revealed that patients attending the eHealth program reported a significant reduction in IBS symptom severity at 3 months, shown in Figure 3 (mean 279.3, SD 80 vs mean 242.7, SD 79.3; $t_{70}=4.91$, $P<.001$). Comparably, patients attending the IBS-school did not report a significant reduction in symptom severity (mean 248.6, SD 86.6 vs mean 235.4, SD 98.9; $t_{49}=0.85$, $P=.40$). Refer to Table S2 of Multimedia Appendix 1 for data. However, none of the groups achieved a clinically meaningful improvement of ≥50 points on IBS-SSS at 3 months. Patients in the eHealth program-group reported a 37-point reduction, whereas patients attending IBS-school reported a 13-point reduction in IBS-SSS total score (Table S2 of Multimedia Appendix 1). Thus, for further analysis we categorized participants as a “responder” or “non-responder” to the intervention, where the sample response threshold for responders was set as ≥50-point decrease in total IBS-SSS score, a threshold that has been demonstrated to correlate with improvements in clinical symptoms [2].

XSL•FO
RenderX

Figure 3. Overall change in IBS-SSS over time. On average, patients who completed the 3-month questionnaires at the IBS-school reported a 13 point reduction in IBS-SSS (n= 50), whereas eHealth program participants reported a 37-point reduction (n=71). A decrease of 50 points is defined as a clinical response to the intervention, categorizing the patient as a "responder to the intervention." The IBS-SSS eHealth program score at baseline was 279 (SD 80) and 243 (SD 79) at 3 months, *P*<.001; the IBS-SSS IBS-school score at baseline was 248 (SD 87) and 235 (SD 99) at 3 months (*P*=.40). Paired *t* test results are summarized in Table S2 in Multimedia Appendix 1. **P*<.001, indicating a significant difference in IBS-SSS scores before and after eHealth program intervention. Blue dot represents the eHealth program; orange dot represents the IBS-school. Paired *t* test was performed. Error bars present SD. Lower n indicate missing data. The reduced number from the total number of participants indicate missing data. IBS: irritable bowel syndrome. IBS-SSS: Irritable Bowel Syndrome Symptom Severity Scale.



Responders

Of the 71 patients, 29 (41%) attending the eHealth program responded to the intervention with a reduction in IBS-SSS score of ≥50 points, compared with 36% (18/50) who attended the IBS-school after 3 months (Figure 3). The results in this section are summarized in Tables 2 and 3. Patients categorized as responders to the eHealth program reported a significant mean reduction in IBS-SSS score of 103 (SD 72.0) points with a moderate effect size of 0.56 (t_{28}=7.62, Cohen *d*=1.33, *P*<.001). In addition, eHealth program–responding patients reported a significant reduction in anxiety, changing from a clinical case to a noncase classification (mean −4.7, SD 6.2; t_{28}=6.89, Cohen *d*=0.85, *P*<.001). Here, the effect size of 0.39 indicated a moderate effect. On average, these patients also reported an increase in overall quality of life, summarized in Tables 2 and 3 (mean 5.80, SD 19.1; t_{28}=−2.28, Cohen *d*=−0.30, *P*=.03). Here, features such as dysphoria (mean 18, SD 23.3; t_{28}=−2.1, Cohen *d*=−0.85, *P*<.001) and body image (mean 14.8, SD 18.6; t_{28}=−3.9, Cohen *d*=−0.72, *P*<.001) increased significantly with a moderate effect size. Features such as food avoidance (mean 6.3, SD 22.5; t_{28}=−2.10, Cohen *d*=−0.13, *P*=.045), health worry (mean 10.8, SD 23.5; t_{28}=−3.44, Cohen *d*=−0.43, *P*=.002), interference with activity (mean 13.4, SD 18.8; t_{28}=−4.03, Cohen *d*=−0.63, *P*<.001), relations (mean 9.5, SD 22.9; t_{28}=−2.72, Cohen *d*=−0.37, *P*=.01), and social relations (mean 10.7, SD 19.4; t_{28}=−3.62, Cohen *d*=−0.56, *P*=.001), all improved significantly, compared with baseline. The changes in sexual activity were not significantly different after 3 months (mean 5.4, SD 29.8; t_{28}=−1.52, Cohen *d*=−0.19, *P*=.14).

XSL•FO
RenderX

Table . Changes in symptom severity, anxiety and depression, and quality of life 3 months after the eHealth program in patients with irritable bowel syndrome (IBS) who attended the eHealth program. Lower n indicate missing data; the reduced number from the total number of participants indicate missing data.

	Responder						Nonresponder					
	n	mean (SD)	*t* test (*df*)	Cohen *d*	Effect size *r*	*P* value	n	mean (SD)	*t* test (*df*)	Cohen *d*	Effect size *r*	*P* value
IBS symptom severity (IBS-SSS[a])	29	−102.8 (72.0)	7.62 (28)	1.33	0.56	<.001[b]	42	9.2 (70.4)	−1.73 (41)	−0.13	−0.06	.09
Anxiety (HADS-A[c])	29	−4.7 (6.2)	6.80 (28)	0.85	0.39	<.001[b]	39	−2.4 (4.8)	3.65 (38)	0.52	0.25	.001[b]
Depression (HADS-D[d])	29	−3.0 (4.6)	3.47 (28)	0.76	0.36	<.001[b]	39	−1.3 (3.8)	1.89 (38)	0.34	0.17	.67
IBS-QOL[e], overall score	29	5.80 (19.1)	−2.28 (28)	−0.30	−0.15	.03[b]	37	2.5 (24.8)	−0.79 (36)	−0.10	−0.05	.44
Body image		14.8 (18.6)	−3.90 (28)	−0.72	−0.34	.001[b]		3.8 (25.4)	−1.24 (36)	−0.16	−0.08	.15
Dysphoria		18.0 (23.3)	−2.10 (28)	−0.85	−0.39	<.001[b]		2.8 (24.3)	−1.20 (36)	−0.11	−0.05	.24
Food avoidance		6.3 (22.5)	−2.10 (28)	−0.26	−0.13	.045[b]		4.8 (21.7)	−1.94 (36)	−0.22	−0.11	.06
Health worry		10.8 (23.5)	−3.44 (28)	−0.43	−0.21	.002[b]		1.8 (23.5)	−0.60 (36)	−0.08	−0.04	.55
Interference with activity		13.4 (18.8)	−4.03 (28)	−0.63	−0.30	<.001[b]		2.5 (23.65)	−1.11 (36)	−0.11	−0.05	.28
Relationships		9.5 (22.9)	−2.72 (28)	−0.37	−0.18	.01[b]		4.6 (21)	−1.81 (36)	−0.20	−0.10	.08
Social relations		10.7 (19.4)	−3.62 (28)	−0.56	−0.27	.001[b]		10.7 (24.6)	−0.53 (36)	−0.46	−0.23	.60
Sexual activity		5.4 (29.8)	−1.52 (28)	−0.19	−0.10	0.14		5.4 (30.13)	−0.56 (36)	−0.19	−0.09	.58

[a]IBS-SSS: Irritable Bowel Syndrome Symptom Severity Scale.

[b]Significant at the $P<.05$ level.

[c]HADS-A: Hospital Anxiety and Depression Scale – Anxiety.

[d]HADS-D: Hospital Anxiety and Depression Scale – Depression.

[e]IBS-QOL: irritable bowel syndrome quality of life.

XSL•FO
RenderX

Table . Changes in symptom severity, anxiety and depression, and quality of life in patients with irritable bowel syndrome (IBS) 3 months after the IBS-school. Lower n indicate missing data; the reduced number from the total number of participants indicate missing data.

	Responder						Nonresponder					
	n	mean (SD)	*t* test (*df*)	Cohen *d*	Effect size *r*	*P* value	n	mean (SD)	*t* test (*df*)	Cohen *d*	Effect size *r*	*P* value
IBS symptom severity (IBS-SSS)[a]	18	−119.3 (86.2)	5.45 (17)	1.33	0.56	<.001[b]	32	46.5 (77.8)	−4.08 (31)	−0.57	−0.27	<.001[b]
Anxiety (HADS-A[c])	16	−2.5 (6.8)	1.97 (15)	0.43	0.21	.07	29	−1.9 (5.3)	2.02 (28)	0.38	0.19	.05
Depression (HADS-D[d])	16	−2.3 (4.4)	2.17 (15)	0.55	0.27	.046[b]	29	−0.5 (3.9)	0.56 (28)	0.12	0.06	.58
IBS-QOL[e], overall score	8	2.4 (27.9)	−0.49 (7)	−0.08	−0.04	.64	24	−11.4 (23.8)	3.16 (23)	5.52	0.25	.004[b]
Body image		1.68 (26.5)	−0.57 (7)	−0.05	−0.02	.59		−4.0 (22.4)	1.38 (23)	0.18	0.09	.18
Dysphoria		17 (27.2)	−6.76 (7)	−0.62	−0.30	<.001[b]		−3.3 (21.7)	1.11 (23)	0.15	0.08	.28
Food avoidance		5.3 (28)	−0.62 (7)	−0.17	−0.09	.56		0.8 (21.6)	−0.13 (23)	−0.03	−0.02	.85
Health worry		13.7 (15.3)	−2.76 (7)	−0.67	−0.32	.03[b]		0 (22.4)	−0.01 (23)	0	0	.20
Interference with activity		9.3 (28.1)	−3.14 (7)	−0.30	−0.15	.02[b]		−0.5 (19.1)	0.19 (23)	0.02	0.01	.85
Relationships		21.2 (22.6)	−3.50 (7)	−0.87	−0.40	.01[b]		18.7 (17)	−2.11 (23)	−0.86	−0.39	.046[b]
Social relations		7.8 (21.4)	−1.23 (7)	−0.27	−0.13	.26		6.4 (24.2)	1.84 (23)	0.26	0.13	.08
Sexual activity		5.4 (29.8)	1.05 (7)	−0.19	−0.10	.33		2.8 (31.6)	0.81 (23)	0.08	0.04	.43

[a]IBS-SSS: Irritable Bowel Syndrome Symptom Severity Scale.

[b]Significant at the $P<.05$ level.

[c]HADS-A: Hospital Anxiety and Depression Scale – Anxiety.

[d]HADS-D: Hospital Anxiety and Depression Scale – Depression.

[e]IBS-QOL: irritable bowel syndrome quality of life.

Patients who responded to the IBS-school intervention reported an average reduction in IBS symptom severity score of 119 (SD 86.2) points (n=18; t_{17}=5.54), Cohen d=1.33, P<.001). In addition, patients reported a reduction in depression scores (mean −2.3, SD 4.4; t_{15}=2.17, Cohen d, P=.046). However, these patients did not report significant improvements in anxiety or overall quality of life (mean 2.5, SD 6.8; t_{15}=1.97, Cohen d=0.43, P=.07 and mean 2.4, SD 27.9; t_7=−0.57, Cohen d=−0.08, P=.64, respectively). Furthermore, the number of responding patients were too low for further in-depth statistical analysis of quality of life (n=8). Data for these analyses are presented in Table S3 of Multimedia Appendix 1.

Nonresponders

Patients who were nonresponders to the eHealth program reported no improvement in symptom severity or quality of life (mean 9.2, SD 70.4; t_{41}=−1.73, Cohen d=−0.13, P=.09) and a significant decrease in anxiety scores (mean −2.4, SD 4.8, t_{38}=3.65, Cohen d=0.52, P=.001), compared with baseline. However, both changes were of small effect (r=−0.27 and 0.25, respectively). Results are summarized in Tables 2 and 3. There were no significant changes in depression scores or overall quality of life or the respective sub-categories (P>.05). Nonresponding patients attending IBS-school reported significantly enhanced symptom scores (mean 46.5, SD 77.8; t_{31}=−4.08, Cohen d=−0.57, P<.001). However, it is worth noting that these enhanced symptoms scores were not above the 50-point threshold for clinical relevance. Furthermore, these

XSL•FO
RenderX

patients reported a significantly reduced overall quality of life (mean −11.4, SD 23.8; t_{23}=3.16, Cohen d=5.52, P=.004) and a significant increase in the domain of relationships (mean 18.7, SD 17; t_{23}=−2.11, Cohen d=−0.86, P=.046). Baseline characteristics for these analyses are presented in Table S3 of Multimedia Appendix 1.

Association Between Type of Intervention and Outcome

A χ^2 test of independence analysis was used to test whether the type of intervention was independent from the intervention outcome. Results showed that the proportion of patients who responded to the intervention in both groups (29/71 vs 18/50) were the same. Hence, there was no significant association between the type of intervention and response to the intervention (Pearson χ^2_1=0.2, Φ=0.041, P=.65). Thus, we conclude that there is no difference in treatment response between the eHealth program and IBS-school in affecting symptom severity scores. Data from this analysis is reported in Table S4 of Multimedia Appendix 1.

Patient Satisfaction

Patient satisfaction was investigated in both groups 3 months after enrollment (CSQ-8). Patients who attended the IBS-school reported a mean score of 24.2 (SD 3.7), compared with patients attending the eHealth program with a mean score of 23.5 (SD 4.0), which are both equivalent to “good” health care offers [22]. An unpaired t test revealed no significant difference between group scores (t_{115}=−1.032, P=.31; Table 4).

Table . Difference in client satisfaction scores (CSQ-8) between patients attending the eHealth program and irritable bowel syndrome (IBS)–school. A χ^2 test of independence was conducted. The reduced number from the total number of participants indicate missing data.

		n	mean (SD)	t test (df)	Cohen d	Effect size r	P value
Overall				−1.032 (115)	−0.193	−0.01	.31
	eHealth program	68	23.46 (4.0)				
	IBS-school	49	24.20 (3.67)				

Discussion

Principal Findings

In this study, we have shown that the novel digital multidisciplinary eHealth program has a significant reducing effect on IBS symptom severity and is useful as a tool in disease self-management. In total, 41% (29/71) of participants reported significant and clinically relevant symptom relief. Furthermore, we show that the eHealth program is safe, as patients not responding to the intervention reported unchanged symptoms and quality of life at 3 months. In addition, eHealth intervention–responding patients reported significant benefits on multiple domains of IBS-related quality of life such as body image, food avoidance, health worry, interference with activity, relations, and social relations. Levels of anxiety were significantly reduced, and levels of dysphoria were improved.

Comparably, 36% (18/50) of participants reported a clinically significant effect to the onsite multidisciplinary 2-day group-based education program, the so-called “IBS-school.” Thus, our results indicate that the digital multidisciplinary eHealth program may be equally effective to the IBS-school, which is often a standard treatment offer to newly diagnosed patients. Furthermore, patients responding to the IBS-school intervention did not report any significant improvements in quality of life or in anxiety, but a small not clinically meaningful decrease in depression scores. The number of IBS-school responders were too low for more in-depth statistical analysis on the domains of quality of life.

A χ^2 test of independence showed no difference between intervention outcome in the 2 groups; hence the eHealth program did not have a better intervention response than the IBS-school on the measures of symptom severity. However, the eHealth program had a significant effect on other aspects of IBS symptomatology, including anxiety and quality of life, whereas IBS-school did not.

Patients rated both the eHealth program and the IBS-school as good health care offers on measures of patient satisfaction of health care quality, scoring them 23.5 and 24.2 out of a maximum of 32 points, respectively.

We believe these results indicate that a digital eHealth approach, designed to provide patients with evidence-based information and practical skills, is preferable to an onsite multidisciplinary 2-day group-based education program covering the same topics.

Comparison With Previous Work

Initially at baseline in both groups, the greatest impairment in quality of life was observed for the subscale of food avoidance followed by body image, inactivity, and dysphoria. This order of impairment is similar to findings by Drossmann et al [21] in an international study from 2009 (n=1966). However, our findings show lower scores on all subscales except dysphoria and health worries, which are in the same magnitude. In comparison with a newer study on quality of life by Kopczyńska et al [23] (n=87), our baseline results show much lower scores on both overall and all subscales of quality of life. A recent Vietnamese study showed that patient education, lifestyle, and dietary intervention, administered by clinical pharmacists, improved IBS related quality of life compared with standard medical therapy over 8 weeks [10]. These study patients also showed a much higher baseline and postintervention quality of life-scores compared with our study. We speculate that our study participants represent a group with more severe IBS because they are referred to a tertiary health care institution that due to capacity issues has to prioritize those patients who need it the most. However, the low baseline scores on food avoidance in our study indicate that both our interventions with a key focus on diet was the right call. This core problem was targeted

XSL•FO
RenderX

because food is a known important trigger of IBS-symptoms [24]. On the domain of food avoidance, we observed an improvement in eHealth program-responders, compared with baseline (Tables 2 and 3). No such improvement was observed in patients attending the IBS-school. We speculate that this is due to the differences in comprehensiveness and durability. The eHealth program is designed to engage the patient interactively to learn how to make appropriate changes in diet. It offers a thorough guidance with videos and instructions on how to follow the low FODMAP diet, including the exclusion and reintroduction of foods. In addition, the patient has the option of digital question and answer sessions with a clinical dietitian throughout the program. The IBS-school is also designed for patients to learn how to self-help, but the program is much shorter with its 2 days of physical attendance. In the light of these results, we deduce that the eHealth program is not inferior to the established IBS-school as a health care offer. In fact, we show that an eHealth intervention program can provide the patient with self-help tools that can lead to reduced gastrointestinal symptoms and enhance quality of life for patients with IBS. This aligns with a recent randomized controlled study by Tayama et al [13] (n=40) showing that a multidisciplinary eHealth self-management program leads to an increased intake of FODMAP-groups and subsequently a more extensive diet. Severe food avoidance and dietary restriction is previously reported in 13% (829/955) of IBS-patients and this subgroup of IBS-patients reported more severe IBS-symptoms, reduced quality of life, and reduced intake of nutrients [24]. Thus, providing patients with evidence-based information, practical tools, and support by a clinical dietitian is important in clinical care of patients with IBS.

eHealth programs in the form of apps, internet-guided programs, or telehealth has recently accelerated as useful tools in clinical medicine. In an American study on satisfaction during COVID-19, most patients with IBS reported high satisfaction rates and ease of use with telehealth [25]. Here the authors reported on multiple benefits including the patient having to take less time off from work and improved access to the care team. Their most commonly reported challenges with telehealth included feeling impersonal and being unable to address all of their issues or concerns. A majority felt that telehealth was as good as or better than face-to-face visits and would use telehealth for future care. Only approximately 10% (130/1311) of the patients remained dissatisfied. In 2020, a Polish study showed that an educational program combined with elements of behavioral therapy, individualized for patients with IBS, is an important part of therapy [26]. In addition, as a part of a dietetic-led gastroenterology service in primary care, feasibility, acceptability, and cost-efficiency of using webinars to deliver first-line patient education for patients with IBS, has been shown to be successful [27]. A meta-analysis of chronic gastrointestinal illness interventions (19 studies conducted in 8 countries, n=3193) showed that eHealth gastrointestinal interventions improved patients' quality of life, psychological distress, medication adherence, and illness-related knowledge [28]. The meta-analysis also showed that eHealth gastrointestinal interventions significantly reduced the number of patient visits to the hospital. Taken together with our results, these findings support eHealth interventions holding decent promise in improving outcomes for patients with IBS.

Comparably on patient satisfaction, a randomized controlled trial by Lackner et al [29] showed that patients who received 4 gastroenterologist-led patient education sessions over 2 weeks reported 26.6 in patient satisfaction. Here, 43.5% (63/145) of patients reported improvement in addition to a higher satisfaction score than in our study. However, both these and our results reflect that subjective symptom relief may not be required for the patient to experience the treatment as useful. This is also highlighted by our patients attending the IBS-school who did not experience significant enhancement in quality of life nor reduction in symptom severity, but still reported a higher satisfaction than patients who objectively benefited more from the eHealth program.

Limitations

All 255 recruited patients completed the programs. However, there was a very high percentage that did not submit the 3-month questionnaires. In total, 41% (54/132) percent of patients attending the eHealth program, and 54% (67/123) of the patients who attended the IBS-school did not submit the 3-month follow-up questionnaires on IBS symptoms severity (Figure 1). Unfortunately, this occurred despite multiple efforts and reminders (phone calls, emails, and SMS text messages) from the research team encouraging the participants to respond. First, this high percentage of missing data are significant and may have affected the results of the study even though the rates appear similar across the groups. However, high dropout rates are a common issue in eHealth studies and known as "the law of attrition" [30]. This warrants the need for a larger study where an intention-to-treat analysis can be carried out without diminishing the power of the study. Second, eHealth literacy is defined by Norman and Skinner [31] as "the ability to seek, find, understand, and appraise health information from electronic sources and apply the knowledge gained to addressing or solving a health problem". The individual patient's level of health literacy was not mapped during the study and variations in these abilities may have affected our results. Third, the patient population represents all subtypes of IBS and there are no analyses focusing on the differences in response between patients with predominant diarrhea, constipation, or a mix of the 2. Comorbidities or other additional diagnoses are common in IBS [32] but have not been excluded in this study and may have affected the results. Fourth, the use of drugs during the study period have not been reported. Hence, many drugs have side effects such as nausea, vomiting, diarrhea, constipation, flatulence, which are symptoms that the patient may confused with IBS symptoms. Fifth, a high placebo effect, which can be up to 40%, is a known challenge in clinical studies on IBS [20]. Although this study has not been designed with a control group, the placebo effect may have affected our results in either group and have not been adjusted for. However, we may speculate that the phenomenon has affected patients in both groups equally and importantly, the placebo effect may recede after 12 weeks, which was our end point [33]. Sixth, as both programs are broad and cover a variety of information, advice, and treatment including the low FODMAP diet and principles of CBT, another limitation of this study is the unknown specifics patients were

XSL•FO
RenderX

responding to in either program. Indeed, there are no measures on cognition verifying the direct effects of the principles of CBT or exposure therapy. This needs to be further investigated in a prospective study, and we acknowledge that an in-depth explanation for our observed benefits after attending the eHealth program remain to be clarified. Thus, these aspects are objectives in our currently ongoing randomized controlled trial [34]. In the light of limited primary and secondary health care resources, it will be useful to develop prediction tools to identify which patients may achieve improvement in both symptom severity and domains of quality of life. For these stratification analyses to be clinically meaningful, the number of participants need to be higher than reported in this study and performed in and randomized controlled trial.

Conclusions

We conclude that the digital multidisciplinary eHealth program has a significant effect on IBS symptom severity in a portion of patients, and is useful as a tool in disease self-management. In addition, it does not result in worse symptom scores than an onsite multidisciplinary 2-day group-based education program after 3 months. We believe these results indicate that a digital eHealth approach, that include benefits such as 3 months unlimited access to quality assured information and treatment with documented effect, is preferable to an onsite multidisciplinary 2-day group-based education program covering the same topics.

Acknowledgments

The authors wish to thank all patients who participated at the irritable bowel syndrom (IBS)–school and the eHealth program. In addition, we wish to thank Haukeland University Hospital's “Learning and Mastering Center” and their wonderful employees, the patient educators psychiatrist Tor Jacob Moe, physiotherapist Eirik Østhammer, psychosomatic psychiatrist Jørn Bødtker, patient representative Cecilie Ågotnes, IBS-school host nurse Per Førde Refsnes, clinical dietitians Cecilie Larsen and CT, CheckWare AS, and eHealth program developer Jan Gunnar Hesthammer, Section for eHealth, Haukeland University Hospital. We also wish to thank the Department of Medicine and Section for Gastroenterology at Haukeland University Hospital, Tine Nordgreen and Research Centre for Digital Mental Health Services, Centre for Research-Based Innovation, Haukeland University Hospital, and Bjrann Ljottson, Karolinska Institutet, and Pelle Lindfors, Sahlgrenska, Sweden. This work was supported by Helse Vest under grants HV1148 and R-10923.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables.
[XLSX File, 47 KB - humanfactors_v12i1e43618_app1.xlsx]

References

1. Ford AC, Sperber AD, Corsetti M, Camilleri M. Irritable bowel syndrome. Lancet 2020 Nov 21;396(10263):1675-1688. [doi: 10.1016/S0140-6736(20)31548-8] [Medline: 33049223]
2. Francis CY, Morris J, Whorwell PJ. The irritable bowel severity scoring system: a simple method of monitoring irritable bowel syndrome and its progress. Aliment Pharmacol Ther 1997 Apr;11(2):395-402. [doi: 10.1046/j.1365-2036.1997.142318000.x] [Medline: 9146781]
3. Cassar GE, Youssef GJ, Knowles S, Moulding R, Austin DW. Health-related quality of life in irritable bowel syndrome: a systematic review and meta-analysis. Gastroenterol Nurs 2020;43(3):E102-E122. [doi: 10.1097/SGA.0000000000000530] [Medline: 32487960]
4. Oka P, Parr H, Barberio B, Black CJ, Savarino EV, Ford AC. Global prevalence of irritable bowel syndrome according to Rome III or IV criteria: a systematic review and meta-analysis. Lancet Gastroenterol Hepatol 2020 Oct;5(10):908-917. [doi: 10.1016/S2468-1253(20)30217-X] [Medline: 32702295]
5. Sperber AD, Bangdiwala SI, Drossman DA, et al. Worldwide prevalence and burden of functional gastrointestinal disorders, results of Rome Foundation Global Study. Gastroenterology 2021 Jan;160(1):99-114. [doi: 10.1053/j.gastro.2020.04.014] [Medline: 32294476]
6. Hillestad EMR, van der Meeren A, Nagaraja BH, et al. Gut bless you: the microbiota-gut-brain axis in irritable bowel syndrome. World J Gastroenterol 2022 Jan 28;28(4):412-431. [doi: 10.3748/wjg.v28.i4.412] [Medline: 35125827]
7. Camilleri M. Diagnosis and treatment of irritable bowel syndrome: a review. JAMA 2021 Mar 2;325(9):865-877. [doi: 10.1001/jama.2020.22532] [Medline: 33651094]
8. Ankersen DV, Weimers P, Bennedsen M, et al. Long-term effects of a web-based low-FODMAP diet versus probiotic treatment for irritable bowel syndrome, including shotgun analyses of microbiota: randomized, double-crossover clinical trial. J Med Internet Res 2021 Dec 14;23(12):e30291. [doi: 10.2196/30291] [Medline: 34904950]

XSL•FO
RenderX

9. Lindfors P, Axelsson E, Engstrand K, et al. Online education is non-inferior to group education for irritable bowel syndrome: a randomized trial and patient preference trial. Clin Gastroenterol Hepatol 2021 Apr;19(4):743-751. [doi: 10.1016/j.cgh.2020.04.005] [Medline: 32289541]
10. Thong VD, Ngoc Phuc N, Quynh BTH. Effectiveness of educational intervention carried out by clinical pharmacists for the quality of life of patients with irritable bowel syndrome: a randomized controlled trial. JGH Open 2021 Feb;5(2):242-248. [doi: 10.1002/jgh3.12477] [Medline: 33553662]
11. Toivonen KI, Zernicke K, Carlson LE. Web-based mindfulness interventions for people with physical health conditions: systematic review. J Med Internet Res 2017 Aug 31;19(8):e303. [doi: 10.2196/jmir.7487] [Medline: 28860106]
12. Hughes S, Sibelli A, Everitt HA, et al. Patients' experiences of telephone-based and web-based cognitive behavioral therapy for irritable bowel syndrome: longitudinal qualitative study. J Med Internet Res 2020 Nov 20;22(11):e18691. [doi: 10.2196/18691] [Medline: 33216002]
13. Tayama J, Hamaguchi T, Koizumi K, et al. Efficacy of an eHealth self-management program in reducing irritable bowel syndrome symptom severity: a randomized controlled trial. Sci Rep 2024 Jan 3;14(1):4. [doi: 10.1038/s41598-023-50293-z] [Medline: 38172498]
14. Hookway C, Buckner S, Crosland P, Longson D. Irritable bowel syndrome in adults in primary care: summary of updated NICE guidance. BMJ 2015 Feb 25;350:h701. [doi: 10.1136/bmj.h701] [Medline: 25716701]
15. Ireton-Jones C. The low FODMAP diet: fundamental therapy in the management of irritable bowel syndrome. Curr Opin Clin Nutr Metab Care 2017 Sep;20(5):414-419. [doi: 10.1097/MCO.0000000000000398] [Medline: 28617710]
16. CheckWare. URL: https://www.checkware.com [accessed 2025-01-14]
17. Ljótsson B, Hesser H, Andersson E, et al. Provoking symptoms to relieve symptoms: a randomized controlled dismantling study of exposure therapy in irritable bowel syndrome. Behav Res Ther 2014 Apr;55:27-39. [doi: 10.1016/j.brat.2014.01.007] [Medline: 24584055]
18. Whelan K, Martin LD, Staudacher HM, Lomer MCE. The low FODMAP diet in the management of irritable bowel syndrome: an evidence-based review of FODMAP restriction, reintroduction and personalisation in clinical practice. J Hum Nutr Diet 2018 Apr;31(2):239-255. [doi: 10.1111/jhn.12530] [Medline: 29336079]
19. Patrick DL, Drossman DA, Frederick IO, DiCesare J, Puder KL. Quality of life in persons with irritable bowel syndrome: development and validation of a new measure. Dig Dis Sci 1998 Feb;43(2):400-411. [doi: 10.1023/a:1018831127942] [Medline: 9512138]
20. Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. Acta Psychiatr Scand 1983 Jun;67(6):361-370. [doi: 10.1111/j.1600-0447.1983.tb09716.x] [Medline: 6880820]
21. Drossman DA, Morris CB, Schneck S, et al. International survey of patients with IBS: symptom features and their severity, health status, treatments, and risk taking to achieve clinical benefit. J Clin Gastroenterol 2009 Jul;43(6):541-550. [doi: 10.1097/MCG.0b013e318189a7f9] [Medline: 19384249]
22. Larsen DL, Attkisson CC, Hargreaves WA, Nguyen TD. Assessment of client/patient satisfaction: development of a general scale. Eval Program Plann 1979;2(3):197-207. [doi: 10.1016/0149-7189(79)90094-6] [Medline: 10245370]
23. Kopczyńska M, Mokros Ł, Pietras T, Małecka-Panas E. Quality of life and depression in patients with irritable bowel syndrome. Prz Gastroenterol 2018;13(2):102-108. [doi: 10.5114/pg.2018.75819] [Medline: 30002768]
24. Melchior C, Algera J, Colomier E, Törnblom H, Simrén M, Störsrud S. Food avoidance and restriction in irritable bowel syndrome: relevance for symptoms, quality of life and nutrient intake. Clin Gastroenterol Hepatol 2022 Jun;20(6):1290-1298. [doi: 10.1016/j.cgh.2021.07.004] [Medline: 34229035]
25. Yu C, Tovar S, Shi J, et al. Satisfaction with telehealth among patients with irritable bowel syndrome during COVID-19: cross-sectional survey findings. Clin Transl Gastroenterol 2022 Sep 1;13(9):e00515. [doi: 10.14309/ctg.0000000000000515] [Medline: 35981244]
26. Sierżantowicz R, Lewko J, Jurkowska G. The impact of an individual educational program on the quality of life and severity of symptoms of patients with irritable bowel syndrome. Int J Environ Res Public Health 2020 Jun 13;17(12):4230. [doi: 10.3390/ijerph17124230] [Medline: 32545800]
27. Williams M, Barclay Y, Harper L, Marchant C, Seamark L, Hickson M. Feasibility, acceptability and cost efficiency of using webinars to deliver first-line patient education for people with irritable bowel syndrome as part of a dietetic-led gastroenterology service in primary care. J Hum Nutr Diet 2020 Dec;33(6):758-766. [doi: 10.1111/jhn.12799] [Medline: 32816367]
28. Rohde JA, Barker JO, Noar SM. Impact of eHealth technologies on patient outcomes: a meta-analysis of chronic gastrointestinal illness interventions. Transl Behav Med 2021 Feb 11;11(1):1-10. [doi: 10.1093/tbm/ibz166] [Medline: 31731292]
29. Lackner JM, Jaccard J, Keefer L, et al. Improvement in gastrointestinal symptoms after cognitive behavior therapy for refractory irritable bowel syndrome. Gastroenterology 2018 Jul;155(1):47-57. [doi: 10.1053/j.gastro.2018.03.063] [Medline: 29702118]
30. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11. [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]

XSL•FO
RenderX

31. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. J Med Internet Res 2006 Jun 16;8(2):e9. [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]
32. Shiha MG, Aziz I. Review article: physical and psychological comorbidities associated with irritable bowel syndrome. Aliment Pharmacol Ther 2021 Dec;54 Suppl 1:S12-S23. [doi: 10.1111/apt.16589] [Medline: 34927759]
33. Spiller RC. Problems and challenges in the design of irritable bowel syndrome clinical trials: experience from published trials. Am J Med 1999 Nov 8;107(5A):91S-97S. [doi: 10.1016/s0002-9343(99)00086-8] [Medline: 10588179]
34. Digital treatment of irritable bowel syndrome (IBS) (DIGIBS). National Library of Medicine. 2023. URL: https://clinicaltrials.gov/study/NCT06117865?cond=Irritable%20Bowel%20Syndrome&term=digibs&rank=1 [accessed 2025-01-14]

Abbreviations

CBT: cognitive behavioral therapy
CSQ-8: client satisfaction questionnaire
FODMAP: fermentable oligosaccharides, disaccharides, monosaccharides, and polyols
HADS: Hospital Anxiety and Depression Score
IBS: irritable bowel syndrome
IBS-QOL: irritable bowel syndrome quality of life
IBS-SSS: Irritable Bowel Syndrome Symptom Severity Scale
***ICD-10*:** *International Statistical Classification of Diseases, Tenth Revision*
***ICPC-2*:** *International Classification of Primary Care, Second Edition*

Edited by A Kushniruk; submitted 26.10.22; peer-reviewed by A Al-Asadi, I Mesters, S Hughes; revised version received 31.08.24; accepted 14.10.24; published 03.02.25.

Please cite as:
Berentsen B, Thuen C, Hillestad EMR, Steinsvik EK, Hausken T, Hatlebakk JG
The Effects of Digital eHealth Versus Onsite 2-Day Group-Based Education in 255 Patients With Irritable Bowel Syndrome: Cohort Study
JMIR Hum Factors 2025;12:e43618
URL: https://humanfactors.jmir.org/2025/1/e43618
doi:10.2196/43618

© Birgitte Berentsen, Camilla Thuen, Eline Margrete Randulff Hillestad, Elisabeth Kjelsvik Steinsvik, Trygve Hausken, Jan Gunnar Hatlebakk. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 3.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Unveiling Sociocultural Barriers to Breast Cancer Awareness Among the South Asian Population: Case Study of Bangladesh and West Bengal, India

Fahmida Hamid[1*], PhD; Tania Roy[1*], PhD

New College of Florida, Sarasota, FL, United States

[*]all authors contributed equally

Corresponding Author:
Fahmida Hamid, PhD
New College of Florida
5800 Bay Shore Rd
Sarasota, FL, 34243
United States
Phone: 1 (941) 487 5000
Fax: 1 (941) 487 5000
Email: fhamid@ncf.edu

Abstract

Background: Bangladesh and West Bengal, India, are 2 densely populated South Asian neighboring regions with many socioeconomic and cultural similarities. In dealing with breast cancer (BC)–related issues, statistics show that people from these regions are having similar problems and fates. According to the Global Cancer Statistics 2020 and 2012 reports, for BC (particularly female BC), the age-standardized incidence rate is approximately 22 to 25 per 100,000 people, and the age-standardized mortality rate is approximately 11 to 13 per 100,000 for these areas. In Bangladesh, approximately 90% of patients are at stages III or IV, compared with 60% in India. For the broader South Asian population, this figure is 16%, while it is 11% in the United States and the United Kingdom. These statistics highlight the need for an urgent investigation into the reasons behind these regions' late diagnoses and treatment.

Objective: Early detection is essential for managing BC and reducing its impact on individuals. However, raising awareness in diverse societies is challenging due to differing cultural norms and socioeconomic conditions. We aimed to interview residents to identify barriers to BC awareness in specific regions.

Methods: We conducted semistructured interviews with 17 participants from West Bengal and Bangladesh through Zoom (Zoom Video Communications). These were later transcribed and translated into English for qualitative data analysis. All our participants were older than 18 years, primarily identified as female, and most were married.

Results: We have identified 20 significant barriers to effective BC care across 5 levels—individual, family, local society, health care system, and country or region. Key obstacles include neglect of early symptoms, reluctance to communicate, societal stigma, financial fears, uncertainty about treatment costs, inadequate mental health support, and lack of comprehensive health insurance. To address these issues, we recommend context-specific solutions such as integrating BC education into middle and high-school curricula, providing updates through media channels like talk shows and podcasts, promoting family health budgeting, enhancing communication at cultural events and religious gatherings, offering installment payment plans from health care providers, encouraging regular self-examination, and organizing statewide awareness campaigns. In addition, social media can be a powerful tool for raising mass awareness while respecting cultural and socioeconomic norms.

Conclusions: Fighting BC or any fatal disease is challenging and requires support from various dimensions. However, studies show that raising mass awareness is crucial for the early detection of BC. By adopting a sensitive and well-informed approach, we aim to improve the early detection of BC and help reduce its impact on South Asian communities.

(JMIR Hum Factors 2025;12:e53969) doi:10.2196/53969

KEYWORDS

Bangladesh; West Bengal; India; Asia; breast cancer; awareness; early detection; screening; sociocultural barriers; health knowledge; cultural; stigma; social media; socioeconomic

XSL•FO
RenderX

Introduction

According to the GLOBOCAN (Global Cancer Observatory) 2020 report on the estimates of cancer incidence and mortality among the people of 185 countries (produced by the International Agency for Research on Cancer), female breast cancer (BC) has been identified as the most commonly diagnosed cancer (11.7%) and is the fifth leading cause of cancer mortality worldwide [1]. The same report suggests that among women, BC accounts for 1 in 4 cancer cases and for 1 in 6 cancer deaths, ranking first for incidence in most of the countries (159 of 185 countries). Through numerous studies, researchers found that early detection can significantly reduce fatal outcomes, and awareness is one of the keys to early detection [2]. Spreading awareness is a challenge for different societies for various reasons, and researchers have been working on identifying those for a long time. In this work, we assess the barriers to mass awareness (and thus far early detection) of BC among people in 2 very densely populated but culturally similar and geographically neighboring areas—Bangladesh and West Bengal, India.

Bangladesh (area: 57,321 square miles) and West Bengal, India (area: 34,267 square miles) are 2 territories in South Asia. Bangladesh's official language is Bengali, and the mass people primarily use the Bengali language for daily conversation. In total, 8% (91,276,115/1,210,569,573) of the Indian population speaks Bengali and mainly resides east of the country, with a concentration in West Bengal [3]. The population density of Bangladesh is 1328.68 people per sq km (collected from the web), and in West Bengal, India is 1028 people per sq km (according to their 2021 census). The deep sociocultural linkages between the Bengali linguistic community in eastern India and Bangladesh remain stable even after India's partition and independence in 1947. The unemployment rate in Bangladesh is 5.23%, and in West Bengal, it is 5.2% (as of June 2022). A total of 70.54% (64,385,546/91,276,115) of people from West Bengal follow Hinduism, and 27.01% (24,654,825/91,276,115) follow Islam. In Bangladesh, 91.04% (150,360,405/169,828,921) are Muslims, and 7.95% (13,130,109/169,828,921) are Hindus. The life expectancy in West Bengal, India, is 71.2 years [4], and in Bangladesh, it is 73.57 years [5]. The approximate sex ratio (as of 2022) in Bangladesh is 102.12 males per 100 females, whereas in West Bengal, it is 105 males for every 100 females. The average literacy in Bangladesh is 74.91% [6], and in West Bengal is 76.26% [6]. The commonality in the history and sociocultural environment of eastern India and Bangladesh has created a pseudo-homogeneous society where language is one of the fundamental elements in this connectivity. Hence, our work focuses on understanding attitudes, perceptions, and behaviors around BC-related issues in these 2 densely populated areas.

Unfortunately, there are no available sources of comprehensive epidemiological, pathological, and outcomes data for patients with BC [7,8] for the chosen regions; so, we are using the closest information from the GLOBOCAN 2020 report [1] to estimate the incidence and mortality rate of BC in this part of the world (Table 1). The factors contributing to the rising incidence and mortality of BC in low-income countries such as Bangladesh and West Bengal, India, are multifaceted. It is important to acknowledge that the solutions required for addressing these challenges differ significantly from those implemented in high-income countries [9]. Therefore, adopting models or policies directly from high- or middle-income countries, which have demonstrated success in tackling BC, may not be applicable in this context.

Table 1. Age-standardized incidence rate and age-standardized mortality rate of breast cancer in India and Bangladesh (GLOBOCAN 2020 report) and some relevant sources like the World Health Organization.

Scope	Age-standardized incidence rate in 100,000	Age-standardized mortality rate in 100,000
World	47.8 [1]	13.6 [1]
South Central Asia (Bangladesh and India are geographically part of the area)	26.2 [1]	13.1 [1]
West Bengal, India	25.2 [8]	13.62 [10]
Bangladesh	21.4 [8]	10.69 [11]

Despite being classified as least developed countries, Bangladesh and West Bengal, India, are home to several highly equipped and specialized cancer hospitals, some for BC only; for example, the National Institute of Cancer Research and Hospital (Dhaka, Bangladesh) or Tata Medical Center (Kolkata, West Bengal, India). These institutions, staffed with highly skilled and experienced surgeons, oncologists, and gynecologists, offer mammograms, CT (computed tomography) scans, and other modern screening and medical facilities. However, advanced medical facilities are scarce in rural and suburban areas, and the capacity of existing institutions often falls short of meeting the demand. In addition, there are also significant uncertainties regarding the overall cost of treatment, the duration of care, mental health support, and gaps in preventative measures. Given this context, our goal is to identify key gaps in existing approaches to raising awareness about BC and addressing the barriers that prevent people from seeking preventative care and timely treatment. By targeting these blind spots, we aim to make meaningful changes that promote early detection and improve access to necessary medical services.

In this work, to understand the barriers to BC awareness, we conducted an open-ended interview with the residents from both regions about BC awareness (to assess their current state of knowledge, attitude and readiness, awareness, and so on) After carefully analyzing our interview data, we highlighted some common obstacles in both regions or societies, by grouping

XSL•FO
RenderX

them into different societal levels such as individual, family, local society, health care system, and state or country. In the upcoming sections, we discuss the study design (Methods), the highlighted barriers with some root causes, and suggested solutions (Results); then, we suggest how technology and social media can play a crucial role in resolving some of those effectively (Discussion).

Methods

Study Design

The main objective was to evaluate the general knowledge and awareness of the participants regarding BC, identify sociocultural factors that could hinder mass awareness campaigns, and examine health care disparities and systemic issues. We conducted individual interviews with the participants using a semi structured questionnaire that included a variety of specific questions as well as 1 open-ended question (Table 2). The participants' responses included personal experiences and anecdotal stories.

Table 2. Our questionnaire.

Category	Questions
Awareness	• People of which age range do you think are at high risk of getting breast cancer? • What are the most common symptoms of Breast Cancer? • What are the common misconceptions about Breast Cancer? • Are there any media-related programs that have provided information regarding Breast Cancer or Breast Cancer diagnosis or progression of the disease? Example: Television talk shows, YouTube, WhatsApp, Facebook groups, or movie or television shows • Are you aware that breast cancer can affect anybody regardless of their sex and gender classification? If not, what gender and sex demographic do you think have the possibility of getting Breast Cancer?
Self-diagnosis	• Do you know and voluntarily do the Breast test yourself at home?
Post-diagnosis support	• What were the support mechanisms (nonpharma) provided to you? For example, mental health or NGO support, nutrition, or self-care. • Was the financial impact of the proposed treatment plan discussed with the doctor?
Open-ended	• Is there anything you would like to share about your journey or experience with Breast Cancer?

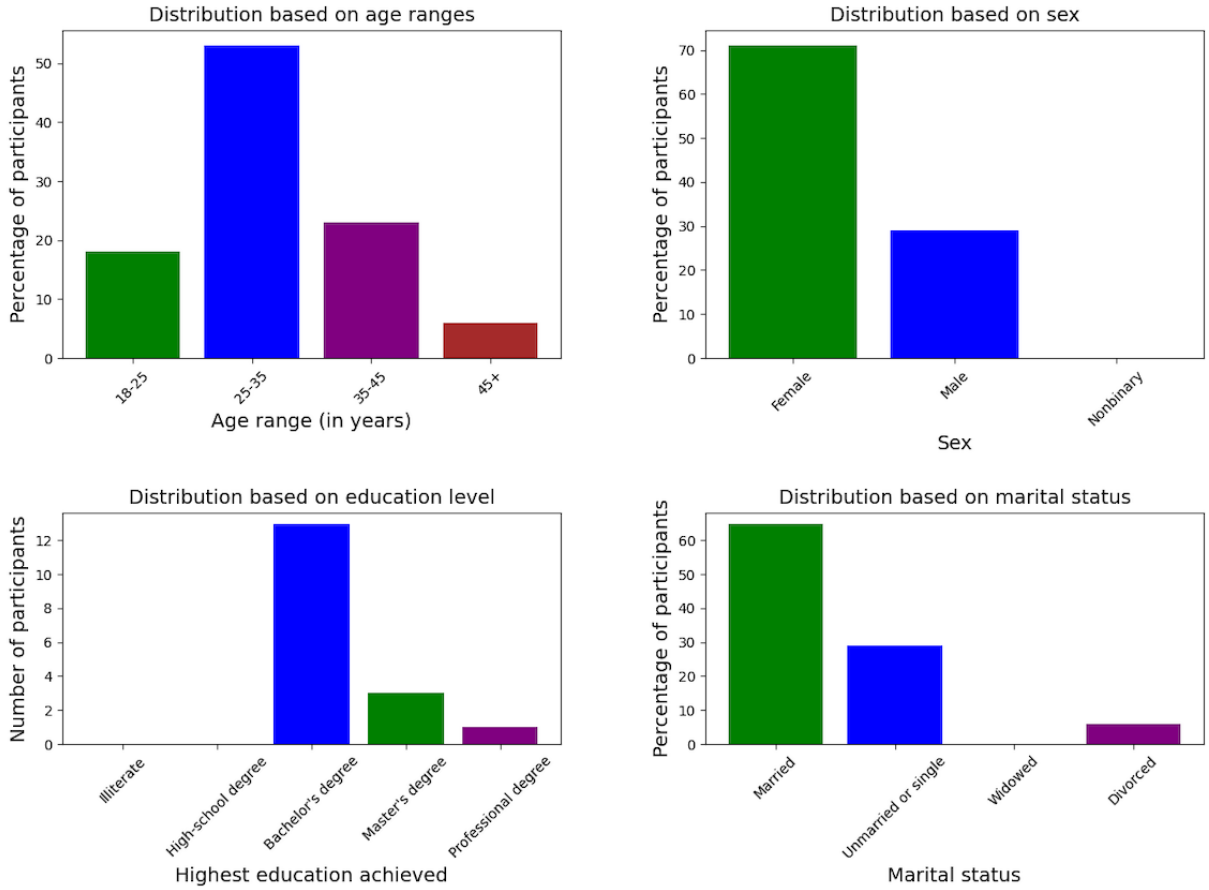
[a]NGO: nongovernmental organization.

Participant Selection and Recruitment

We recruited participants through snowball sampling [12], using mutual connections and social media platforms such as Facebook (Meta). Semistructured interviews were conducted with 17 participants from West Bengal and Bangladesh. All our participants were above the age of 18 years, and recruitment was focused on participants who resided (at the time of study) in the 2 regions mentioned above. Our participants primarily identified as female, with an average age of 25-35 years, and most were married (Figure 1). The participants selected the language during the interview; most spoke Bengali. The researchers conducted the study through Zoom (Zoom Video Communications), and the audio recordings were transcribed and translated into English for data analysis.

XSL•FO
RenderX

Figure 1. Distribution of participants.



Demographics

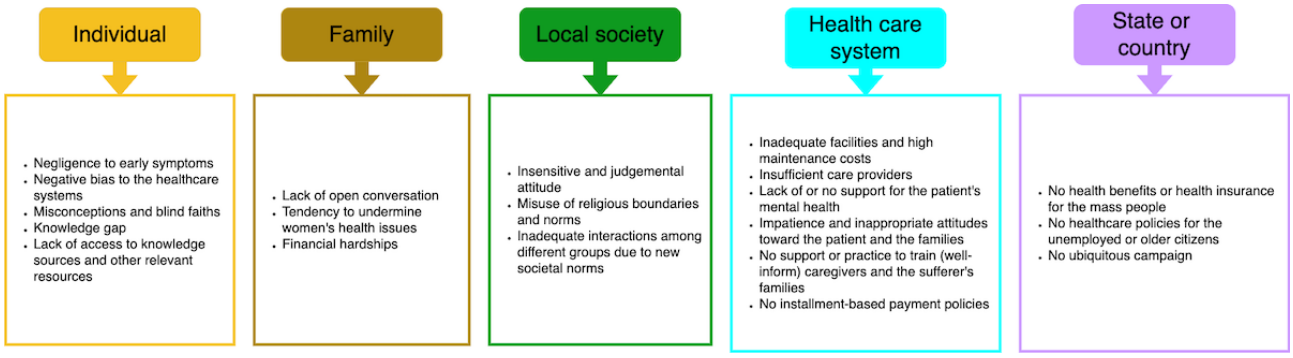
We aimed to recruit a representative sample of participants from diverse educational and financial backgrounds from the region, as socioeconomic status and education play a huge role in the participants' access to health care in the region. Also, a cultural stereotype related to educational qualifications is widely propagated, relating educational accomplishments to open-mindedness, access to health care, and overall "success" in life [13]. Our participants self-declared their financial status as middle- and high-economic classes and were all in professions such as business, architecture, social development work, or education. Most of our participants had completed their undergraduate degrees (n=13), few had graduate degrees (n=3), and some were currently working toward a professional degree (n=1). The socioeconomic status of our participants is important in the larger context of interpreting the results and will be discussed further in the upcoming section.

Data Collection and Analysis

We transcribed the interview audio recordings, which were between 15 and 45 minutes long, and translated into English. Our qualitative data analysis follows a tiered approach and an inductive methodology [14,15] to generate themes. In tier 1, the initial coder examined the transcript identifying recurring themes, clustering them based on similarity, and using the descriptive coding methodology to label them. This clustering formed the foundation for subsequent analysis. Moving to tier 2, a second coder independently reviewed the clustered themes and descriptive terms associated with them, providing validation by either corroborating existing clusters or proposing new ones as necessary. Finally, tier 3 involved collaborative discussions between both coders to assess the relevance and significance of the clustered themes and descriptive terms associated with them. Through iterative deliberations and consensus-building, the final classification of themes was determined (Figure 2). As our interviews were semistructured (where the participants had the freedom to take the narrative and provide more information beyond a fixed set of questions), this method of generating themes bottom up and assigning descriptive terms to it allowed us the flexibility to show nuance and highlight subthemes. Both coders were familiar with the language spoken by the participants (Bengali) to parse through the linguistic, cultural, and social significance of the data generated during the interview sessions.

XSL•FO
RenderX

Figure 2. Outcome of our qualitative data analysis: major barriers to breast cancer awareness in the region.



The hierarchical representation of our themes and subsequent categories (Figure 2) is inspired by the presentation of the barrier tree [16] developed for a mobile-based BC monitoring study for the people in rural Bangladesh around 2012. Notably, several factors (such as feeling shy, fear, lack of familiarity, undermining women problems, religious belief, shared beliefs and practices, communication problems, scarcity of doctors, inconsistent patient data, long-term monitoring, and poverty) identified then are still highly relevant and came up during our analysis. The distinctions between both works are listed below.

First, most of our participants are city dwellers, but they shared stories and experiences from different parts (city, suburb, or rural) of the country or state during their interviews. Haque et al [16] studied in rural setups about 1.2 decades ago.

Second, we present the barriers by clustering them into 5 socioeconomic themes (individual, family, local society, health care system, and country or state). Haque et al [16] presented their findings into categories like identification and disclosure, achieving treatment, continuation of treatment, environmental issues, and user issues.

Third, while many research articles in the domain of BC primarily adopt a medical science perspective, such as the study by Haque et al [16], where the team visited a clinic or hospital to gather data and generate a barrier tree, our approach diverges by examining BC through a societal framework. By doing so, we aim to offer a comprehensive analysis that encompasses both the medical and societal impacts of BC. This approach allows us to explore the same problem from different lenses, shedding light on the multifaceted challenges and implications of BC beyond the purely medical aspect.

Finally, as a novel addition to the field, we delve into the mental health relevancy of patients with cancer and their caregivers, a crucial aspect that was not addressed in the other work.

We discuss the outcome of the qualitative analysis in the next section.

Ethical Considerations

The institutional review board of New College of Florida approved the study (Protocol #22-037), and we strictly followed all protocols to ensure data privacy and integrity, demonstrating our commitment to ethical standards. For the study, participation was voluntary; participants did not have to participate and could stop at any time. There were no penalties or loss of benefits or opportunities if the participant did not participate or decided to stop once started. Their decision to participate or not to participate did not affect their job status, employment record, employee evaluations, or advancement opportunities. The participants did not receive any benefit from the research team. There was no cost to participate. This research was considered minimal risk. Minimal risk means that the study risks are the same as those in daily life.

Results

The results of the data analysis are discussed in 2 phases; first, barrier identification and second, awareness raising.

Barrier Identification

The majority of our participants were aware of BC's symptoms, possible treatment mechanisms, costs, survival rate, and so on. During the interviews, they were able to list most of the significant symptoms of BC, like deformed breasts or nipples, lumps in the breast and armpit, discharge of pus or liquids, irritation, flaky breast skin, irregular breast pains, overall fatigue and weight loss, and so on. All but 3 participants knew of at least 1 person who has or had BC—either in their own or extended family, in the neighborhood, or in a friend's family—indicating how predominant this disease has become (Tables 3 and 4). A total of 14 participants "had heard of" self–breast-examinations, 13 participants "knew how to" perform self–breast-examinations, and 7 participants "had performed" self–breast-examinations (Figure 3). A greater than 50% decrease in the number of participants who had conducted self–breast examination is noteworthy.

XSL•FO
RenderX

Table 3. Number of responses to the question: "How many patients with breast cancer do you know of?"

Percentage of participants	Number of patients or survivors with BC[a]
18	0
35	1
35	2
6	3
6	4

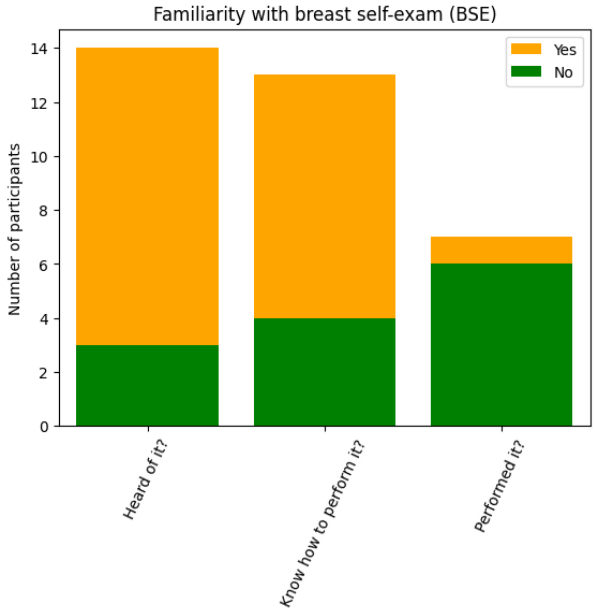
[a]BC: breast cancer.

Table 4. Patients with breast cancer in immediate family.

Percentage of participants	Number of patients or survivors with BC[a] in immediate family
59	0
24	1
17	2

[a]BC: breast cancer.

Figure 3. Familiarity test for "breast self-examination.".



The following sections share the barriers we highlighted from our data. Since the barriers often have cause-and-effect relationships and circular causality, we will not present them according to the identified societal levels like in Figure 2; instead, we would discuss them under the same section or title if they mostly appeared together during the interviews.

Tendency to Neglect Early Symptoms and Knowledge Gap

Our survey participants agreed that although BC is well-known, many people are unaware of the symptoms and often miss early signs. Specifically, when trying to identify symptoms in themselves, they either don't know what to look for or tend to disregard early signs of discomfort. For example, 1 participant stated:

> *One day, my mom had severe stomach pain. We took her to the doctor, who asked us to complete a USG of her abdomen as the doctor suspected that my mom had a tumor in her stomach. While doing the USG for the tumor, the doctor detected that she had gotten breast cancer, not a tumor…… it was already a late detection. It's at a deadly stage already. The treatment would have been much less painful if she had been diagnosed earlier.* [b007]

A similar story also stated the reluctance to pay heed to early symptoms or disclose them:

> *Though she (a neighbor) suspected it, she did not disclose it initially. When the pain became unbearable, then only she went to the doctor. She went to the doctors in the middle stage (not the early stage).* [b011]

XSL•FO
RenderX

> *One of my aunts (a family friend) had been diagnosed with breast cancer. However, it had already spread to the throat and jaw when she was diagnosed. So, the doctors could not remove/cure it, and she passed away.* [b008]

These responses reflect the results Amin et al [17] found in a survey that the research team conducted in Bangladesh in 2020; out of the 9 symptoms of BC, approximately 50% (N=500) of participants were unaware of 5 or more symptoms.

Misconceptions

Although none of our participants were medical professionals, our participant pool consisted of individuals who had all completed at least an undergraduate degree, were socioeconomically affluent, and were self-professed, open-minded individuals. Due to their educational backgrounds, we assumed that the number of misconceptions about BC would be lower than other groups. However, our participants had several deep-rooted misconceptions about BC, which are listed below.

First, it can only afflict females. Surprisingly, 10 out of 17 participants (59%) believed BC could only afflict females. The rest of the participants agreed BC occurred, irrespective of sex.

Second, it is more likely to afflict married women and those who have recently become mothers. Several participants stated that in married women, especially after giving birth to a child or during the breastfeeding period, it is highly likely to get inflicted with BC.

Third, it is more likely to afflict older women. Several participants responded that BC only affects older women; when probed further, they believed that the age range between 25 and 45 was not particularly susceptible to BC. Reports have suggested that Indian women in their early thirties till fifties are at considerable risk of developing BC, and the incidence risk increases till its peak when they reach 50-64 years of age. Furthermore, 1 in 28 Indian women are likely to develop BC during their lifetime [18-21]. It is more (1 in 22) for urban women than for the rural group (1 in 60). Amin et al [17] report similar statistics for Bangladesh (22.5 per 100,000 women and mean age being 41.8 years and 56% of the diagnoses were for reproductive-age women).

At the end of our interviews, we informed our participants that (1) BC is statistically more dominant among females but can afflict anyone, irrespective of gender and (2) there are various types (metastatic, inflammatory, and so on) of BCs that afflict people from various age ranges, so directly linking it to a particular age, marital status, or maternity may not be the best idea.

These factual misconceptions and knowledge gaps about BC are major factors related to the diagnosis of BC at an early stage. Our participant responses reinforced the need to disseminate accurate information to raise awareness.

Social Stigma and Lack of Open Conversation

First, beauty and breasts—judgmental attitudes. In addition to being an organ, breasts play a significant role in defining femininity and women's role in South Asian cultures. The fear of the disease attacks the fundamentals of female identity as a woman's breast conjures her sexuality and her capacity to nurture [22]. In patriarchal societies, such as Bengalis in West Bengal and Bangladesh, irrespective of educational or financial gains and social liberties through the years, women are perceived to have the roles of mothers, wives, or homemakers, whereas men are mostly seen as primary breadwinners of the family [23].

According to a report by a Pakistani Non-Governmental Organization, Aurat Foundation and USAID (United States Agency for International Development; 2016), in South Asian cultures, feminine traits include emotional, nurturing, passive, good homemakers, fair-skinned, without body or facial hair, long-haired, with a narrow waist, and sexually submissive. In contrast, masculine traits include being assertive, primary breadwinners, risk-taking, rational, and brave [24]. The ultimate danger of this notion to a Bengali (or maybe many people from other parts of the world) is that they feel one becomes "lesser feminine" if one has any breast health-related issue. On top of that, many people assume that BC's ultimate aftereffect is mastectomy, if not death, which adds a new level of stigma [25,26].

> *When I was very young, I learned that one of my friend's mom had breast cancer. At that time, my parents were very conservative; they would not open up in front of other family members/us. I learned later that doctors had to remove her breast. But the good thing is, she recovered totally, and it's been 20 years. She has an everyday life. I believe she received good treatment. Every treatment doesn't go wrong!* [b009]

> *Of the few breast-cancer and other cancer survivors I know of, I say it is common for breast cancer patients to feel humiliated or more mentally upset than other cancer patients. …… I don't notice any initiative in our society in this matter; neither is it commonly seen to provide some counseling for the patients for overcoming/fighting cancer.* [b005]

For several reasons, anyone with BC is reluctant to share their experiences. In many cases, patients with BC suffer from depression [17] due to a lack of a support system to debunk the body image and societal myths. The image of an ideal woman propagated in the media (television or social media) reinforces the narrative that is sometimes unhealthy to a person's body image [27]. Particularly the lack of influential media personalities sharing their stories of struggle and recovery or the absence of characters that represent the struggles of patients or survivors of cancer in television shows or media portrayals impacts the perception of the journey and struggles.

Our participants echoed these sentiments in their statements:

> *There is a taboo around breast cancer, regardless of class and status. People do not talk about it. We learn about it later, like once someone has recovered or died of breast cancer. Mostly, no one shares their journey with breast cancer.* [b014]

> *She (a breast cancer survivor) used to talk to me and hang out with our family. However, she and I never*

XSL•FO
RenderX

> *talked about it. Neither I asked, nor she opened up. If it is not very necessary, we will not share. That is the mindset.* [b012]

The term cancer comes with a connotation of impending doom and death; however, the significantly higher chances of patients surviving cancer because of early detection strategies and awareness are not emphasized at all [28]. When it comes to the 5-year overall survival, a study reported it to be 95% for stage I patients, 92% for stage II, 70% for stage III, and only 21% for stage IV patients [29]. It is important to highlight these statistics widely to encourage early detection and awareness. Survival rates of patients with BC are lower compared with Western countries, primarily attributed to early age onset, late-stage detection and delayed initiation of definitive management, and inadequate or fragmented treatment [28].

Second, cultural sensitivities. In addition to not expressing discomfort, awareness campaigns in India and Bangladesh must be wary of the strict cultural, religious, and moral censorship, which does not allow the posting of revealing images of breasts. People, in general, are brought up in an environment where they do not feel comfortable discussing sexual health or women's health concerns, including BC. The term "breast" is perceived through the lens of hypersexualization and hence comes under deep scrutiny from moral and religious sensitivities. As researchers embark on developing awareness campaigns, it becomes crucial to exercise cultural sensitivity, especially considering the conservative nature of the audience in these regions. Being mindful of cultural nuances is essential to ensure that the true essence of the campaign is not overshadowed by moral policing or societal restrictions. It is important to point out that women's health or sexual health and the term breast are not taboo; we cannot undo decades of cultural teaching and biases in a day. Hence respecting the norms and designing awareness campaigns around them would be a more sensitive approach.

> *She (my paternal aunt, who was a teacher) had lung infections and troubled breathing due to breast cancer, but the doctor tracked it back to breast cancer after doing several medical tests; she was never open about the symptoms (or discomforts with her breasts) upfront to the doctor.* [b001]

> *People are not comfortable talking about it. Since I am the son, my mom also never shared about her symptoms or sufferings.* [b007]

Third, women's health care is a necessity, not a luxury. Irrespective of South Asian patriarchal societies or more progressive Western societies, a woman's role in society, economy, and family cannot be ignored. The roles of caregiver, parent, and wife for a rural woman with added burdens of maintaining appearance and body image in the case of an urban woman all contribute to the issues that the contributions women make to the world are "nonproductive" or "nonessential." This perception sometimes from women themselves (in some cases, the family members) leads to the health concerns of women being ignored or underprioritized [23]. This includes a lack of attention to proper nutrition, timely intervention, care after diagnosis, and adequate mental health support. This practice or habit stood out as one of the major reasons why female patients often get a late detection of health-related problems.

> *We wait and hope that the discomforts will disappear automatically after a while – so we tend to linger. I have noticed this among my friends and family (especially females), including myself; we tend not to go to doctors so easily. I am not sure if it's a cultural thing! But it is less about financial conditions and more about awareness.* [b014]

Certainly, that mindset is changing gradually. Nonetheless, there is a need for increased awareness in this particular area:

> *A woman in the family will likely delay her treatment as she thinks it's the least important issue and is probably putting financial pressure on the family. But sometimes, they forget that delayed treatments may invite worse consequences.* [b003]

The Disconnect Between Health Care Models and People

Although none of the existing health care models worldwide are perfect, in the context of a deadly disease like BC and health-related supports, we noted the flaws, inconsistencies, and lack of patient-centered policies in the regions.

First, lack of a yearly health examination policy. Unlike in the United States or other parts of the world, Bengalis usually do not go for yearly health examinations, or it is not a common practice among doctors to prescribe yearly health checkups. The lack of adequate resources, such as available medical professionals, provider-centric consultations, and a symptom-based health care model, is partially responsible for this issue [30]. The region requires immediate advocacy for a robust primary health care system and the use of existing community care efforts to enhance awareness.

Second, out-of-pocket medical expenses. Health examinations are expensive, and it is unusual for the region to support health insurance employers [31]. There is no system of national health insurance in Bangladesh [31]. People whose employers cover health benefits also must pay to the facilities out of pocket first and then later ask for a reimbursement. This cost barrier leads to patients not seeking treatment early on.

Third, lack of integrated service and long-term plans. Those on tight budgets usually have to prioritize the price over reliability and integrated service available to few modern (and more expensive) health care facilities. Patients and their family members are often not fully aware of the overall costs (long-term plans alongside short-term plans) and risk factors so that they can make informed decisions.

> *Talking to the patient and their family about the situation and possibilities is crucial. Only then can we make an informed decision. But if the doctor dictates the options, the patient's family has to follow the order, no matter how expensive or if they have to lose their property in arranging the financial support.* [b009]

Fourth, open conversation between patients and the doctors. Due to various reasons, including high volume of patients,

XSL•FO
RenderX

doctors seem to spend much less time with the patients, which has been a consistent story we heard from our participants.

> *The more renowned the doctors are, the less interactive they are, or the less compassionate they are. There is much less scope and the chance of openly conversing with them.* [b014]

Fifth, insensitive and judgmental commenting. Medical professionals lack adequate training when addressing patients and their concerns. A dearth of empathy compounded by the enormous patient loads leads to negative interactions with the health care providers, and patients or family members do not feel comfortable asking essential questions. An insensitive comment may cause severe damage to the patient's (or their caregiver's) confidence and motivation. Effective doctor-patient communication is determined by the doctor's bedside manner, which patients judge as a major indicator of their doctor's general competence [32,33]. Doctor-patient communication skills are based on verbal and nonverbal communication, where 22% is transferred through voice tone, 55% through visual cues, and 7% verbally [34]. Lack of training to demonstrate effective behavior leads to scarring patient outcomes, one such is listed below:

> *I remember an oncologist's comment to his patient (lung cancer survivor, 65+ years of age) regarding the side effects of Chemo/radiotherapy: you don't have your natural hair anymore! So why worry about that?* [b003]

Sixth, the doctor's gender. BC disproportionately afflicts the female population, and since it concerns an intimate body part, it is common among the patients and their family members to withhold the treatment until they find a female specialist; in short, they prioritize the doctors' gender over professional expertise [32]. In a cross-sectional study conducted with 1257 female patients in Udaipur, India (August 2014), by Nagarajappa et al [32] found an overwhelming 85% (1069/1257) of the respondents said yes to the question: "I would always prefer being treated by a doctor of my gender." In some other similar studies conducted in 2019, researchers found that the most important attributes of maternal health care facility choice for Bangladeshi women were consistent access to a female doctor, the availability of branded drugs, respectful provider attitudes, and a continuum of maternal health care [35,36].

> *My aunt and uncle delayed the treatment as they were hesitant about choosing a doctor; they were looking for a female doctor, but there were not many female breast cancer specialists in Bangladesh during that period. My uncle (my mother's cousin) had hesitation. Had she been diagnosed earlier, she would have had a better chance of survival. She was a survivor, but her entire family suffered greatly due to the delay. My uncle later regretted his decision. I remember him expressing that to my mother (his sister).* [b003]

Although 61.1% (198/324) of medical students in India identify as female, a study by Bajpai et al [37] in 2020 found that men led 67.3% (218/324) of oncology teams in India. These numbers are similar to the numbers in the United States, where out of 667 female respondents, 442 identified as academic oncologists versus nonacademic ones [38]. This gender disbalance among physicians leads to patients not seeking timely medical care.

Seventh, lack of trust. Many people in the region believe that:

> *…an institute (hospital and caregiving center) that is capitalist in nature cannot have people's best interest at heart.* [b016]

This may play a significant role in people's hesitancy to seek health advice unless necessary.

Mental Health Support Infrastructure

All participants expressed concerns about lacking mental health support systems. Most reported that:

- family members and caregivers are never made aware of the patient's mental health, its relevance, and how to handle them [26].
- "I didn't know and never thought mental health could be impacted directly by these (or any) diseases." [b007]
- there is no infrastructure to support the mental health of patients suffering from terminal or deadly diseases.
- "Mental health is considered taboo, more significant of a taboo than breast cancer. People don't usually talk about it. Doctors also don't do a good job in this regard. Providing professional mental health support is not common in India/West Bengal to battle cancer." [b016]
- "Since cancer is associated with a lot of uncertainty, fear, and stress, the moment patients get diagnosed with cancer, they must get counseling. Mental support from proper channels can help patients gather the strength and courage to defeat it." [b010]

Role of Media

Media has always played a significant role in spreading awareness among the general population on various issues. We wanted to learn how our participants felt about the role of media around them in spreading BC awareness. A survey conducted by YouGov Global Media in India identified that in 2022, 59% of Indians under the age of 24 years depended on websites or web apps as their primary media source, whereas people above the age of 45 years depended on television (live or nonlive broadcast) as a primary media source [39]. In Bangladesh, USAID conducted a similar survey and found that 46.6% of respondents mentioned television, and then 33.1% said the internet was a primary source of information media [40]. Similarly, in India, 60% of the population in the age group of 45 years and above depend on television for news, whereas younger audiences tend to lean toward updates from social media instead of watching traditional television news (which broadcasts news at specific hours).

Social media and traditional media (television and newspapers) can substantially spread awareness among people. We wanted to understand whether the media's contribution to raising awareness was limited to a date or time (eg, Breast Cancer Awareness Month) or whether there is any consistent campaigning.

Our participants have reported seeing numerous posts, symbols (such as the pink ribbon or black screen), and talk shows on specific days, weeks, or months of the year (such as during

XSL•FO
RenderX

Breast Cancer Awareness Day). However, these efforts are transitory. They come and go; many posts are left unattended, unread, and unfollowed, and many programs and talk shows are left unwatched. The lack of consistent programming on any platform that regularly sends out relevant messages throughout the year leads to reduced engagement and interest.

> *I have seen many programs aired on the national television of Bangladesh, but I believe that is not sufficient…. I haven't noticed any program regularly on this topic. If we talk about social media, I don't see enough posts relevant to this topic.* [b004]

> *People share (even if that's at a minimal level, especially on special occasions like Breast Cancer Month). But others may not have time to read it carefully due to their busy schedules. I sometimes see such posts but don't have time to read them thoroughly. I feel until it hits someone, they are likely to ignore (or blindside) it.* [b003]

Most BC-related posts (eg, Figure 4) on media platforms include either a pink bow [41,42] or wearing the color pink at campaign events [43,44]. To urban audiences' mammograms are widely promoted by health care professionals or organizations [41,45].

Figure 4. Media coverage, rallies, and campaigns.



The absence of relevant information about the disease, prognosis, and symptoms is stark and leaves the audience wondering how to perceive this awareness campaign. Often, momentum is lost, and true engagement is not achieved, leading to lackluster awareness numbers.

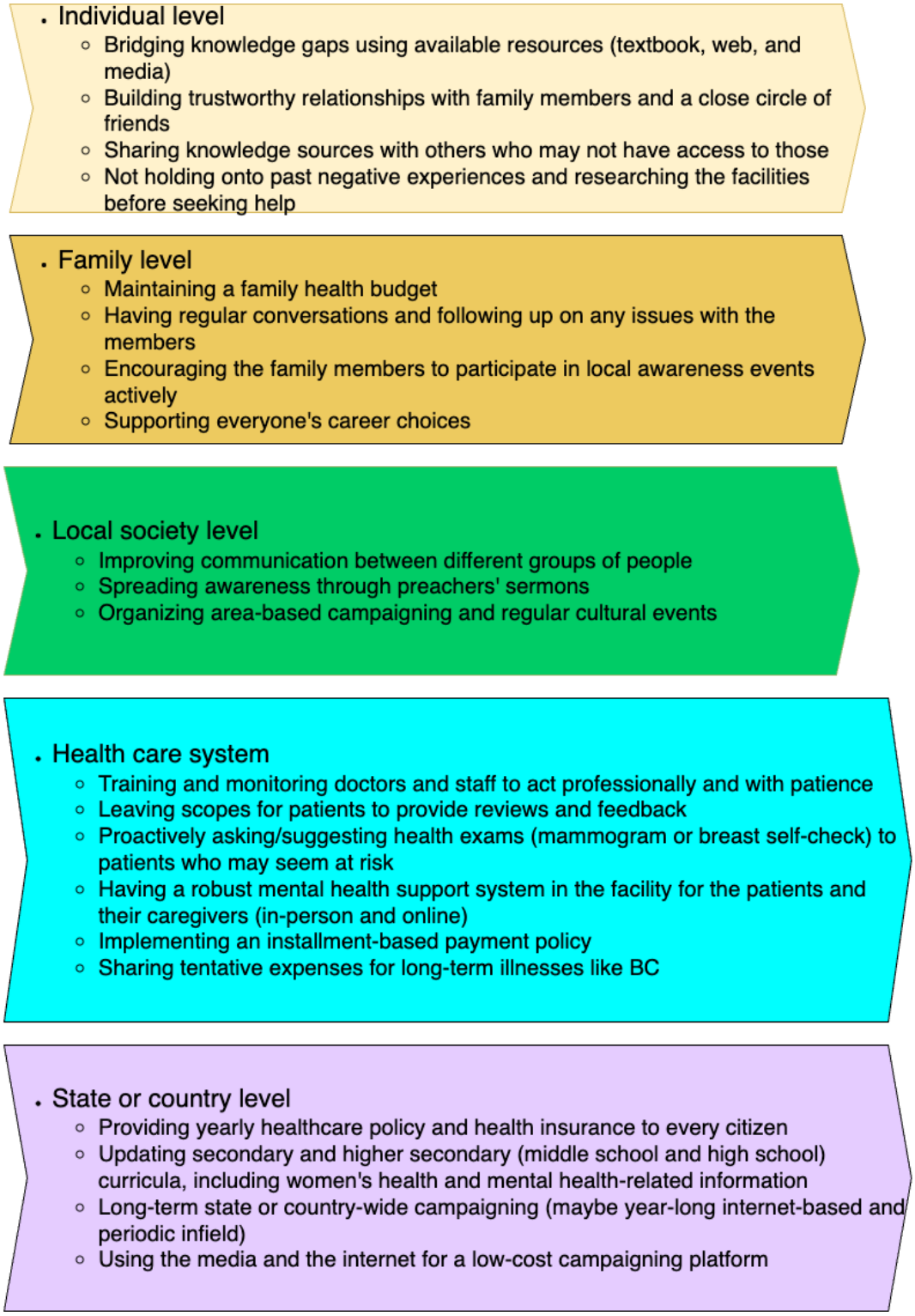
> *We see advertisements relating to breast cancer, sometimes on the television or elsewhere, or billboards, but then it is very short and crisp. And it just says that we should be aware of breast cancer, but then what to be aware of it is never mentioned anywhere.* [b017]

Awareness Raising

Many of the findings are codependent and mutually recursive. For example, lack of trust in health care systems in the area and financial hardships implicitly push people to neglect early symptoms of deadly diseases; lack of open conversation and social stigma block them from reaching out for mental health support, to name a few. Our discussions with the participants highlighted several means to help spread mass awareness in the regions. In this section, we present the summary of our suggestions to achieve mass awareness through Figure 5 and briefly discuss them right after.

XSL•FO
RenderX

Figure 5. Our recommendation to help raise awareness. BC: breast cancer.



Designing a Yearly Health Care Routine for Everyone

> *Society must believe that early diagnosis can help increase survival and reduce suffering. We need to build the practice of going through a yearly checkup.* [b008]

Given the population, unemployment, and other issues, achieving mass success in building this practice will require support and cooperation from international organizations (such as the World Health Organization) and local governments.

Spreading the Voice of Survivors

We often hear words from specialists (doctors and scientists) on mass media (eg, Shastho Kotha [46]). Aside from their knowledgeable notes, the programs will draw attention to the mass scale if BC survivors' interviews are included regularly. This will also reduce fear, hesitancy, and social stigma. When

XSL•FO
RenderX

people see interviews of BC survivors belonging to the same community (age, gender, social status, and so on), it will help everyone to open up about similar issues.

> *I believe cancer changes our views and ways of life. We must break the practice of neglecting minor symptoms and not seeking medical advice. We need to improve trust and healthy communication between doctors and patients. Roughly 50% of the population is female. So, whichever cancer is more probabilistic to which population, they need to take it seriously. Financial hurdles should not be the first concern. It will be beneficial if there is a platform where people can talk about it and not treat it as something that cannot be discussed.* [b016]

Sharing the experiences of BC survivors as they navigate treatment plans, manage financial burdens, and embark on their journey toward recovery can play a crucial role in dispelling fear, reducing anxiety, and challenging the taboos surrounding BC. Narratives that illuminate the resilience of these survivors, highlight the support of their families, and shed light on the obstacles they overcome contribute to humanizing this daunting illness.

Using the Grassroots Health Workers

> *NGOs like Surjer Hashi Clinic, Shobuj Chata Community Clinic, etc., wildly succeeded in child immunization and birth control in Bangladesh. So, if cancer awareness (like breast and cervical cancer) is taken as an agenda by such organizations, I am hopeful it will succeed.* [b003]

If the health workers spread a simple flyer about BC with some visuals and depictions in Bangla during their weekly visits, it will help raise awareness in rural areas quickly.

> *Bangladesh has successfully rolled out the idea of condoms and contraception, though we are considered a very conservative society. Major religions of our land directly oppose the idea of birth control. Yet the success came not because of the paid advertisements or programs in media but because of sending medical advice to the grassroots level. So, I believe we can succeed in growing awareness about breast cancer if the Government takes such programs into the hand.* [b005]

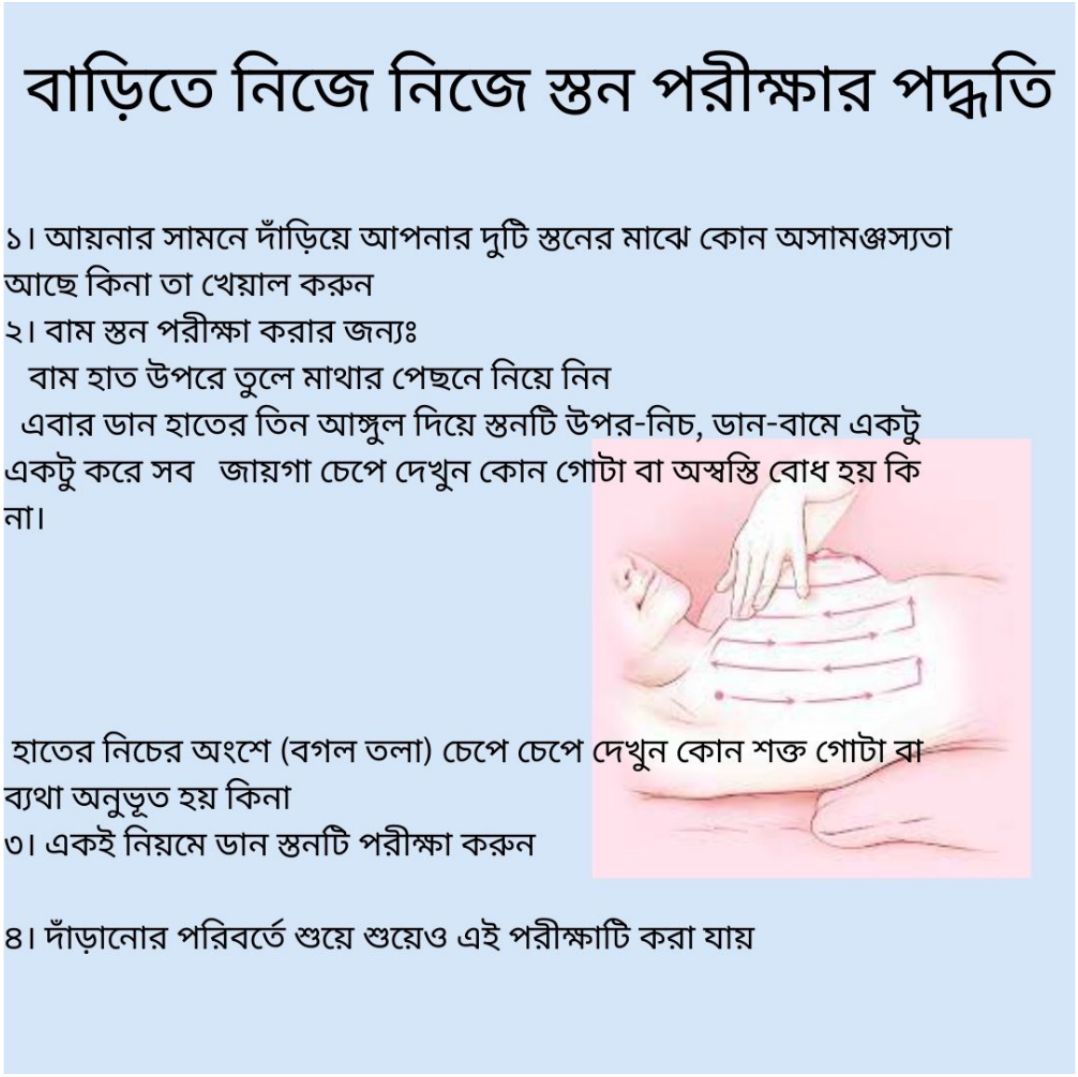
Establishing a Collaborative Effort on Social Media

Bangladesh's internet penetration rate stood at 31.5% (52.58 million/167.1 million) of the total population at the start of 2022 [47], and India's 47% (658 million/1.40 billion) of the population has internet access [48]. India and Bangladesh have both shown steady growth in mobile cell phone service, too. Instant messaging apps such as WhatsApp (Meta; 487 million users in India and 40 million users in Bangladesh) [49,50]. We strongly believe these platforms can be used to spread accurate awareness messages to users, particularly if government health organizations partner.

We suggest regularly preparing WhatsApp message broadcasts or mass forwards to groups, patients, or individuals from credible organizations such as local government branches (eg, zilla panchayat or district offices), medical associations (eg, Bangladesh Medical Association or Medical Council of India), or nonprofit organizations (eg, Sabuj Sathi). Sending these messages individually to users also avoids making anyone uncomfortable with the content, keeping in mind the social and religious sentiments or barriers in these regions. WhatsApp's ubiquitous presence in the digital communication ecosystem of both countries removes additional app downloads or installations. We propose using WhatsApp forwards (illustrated in Figure 6) as a cost-effective, easy-to-understand, and time-efficient method for dissemination. However, it is essential that these messages are reviewed by an expert before distribution to ensure accuracy and reliability.

XSL•FO
RenderX

Figure 6. Proposed self-breast examination instruction flyer.



Inclusive Texts in Secondary and Higher Secondary (Middle and High School) Curricula

Many people in both regions finish their education till middle- or high-school level. Some inclusive articles (written with appropriate languages and cultural values in mind) regarding women's health and the terminal and most predominant diseases in society will help spread awareness. The students and their families will gradually become aware; thus, many of the causes blocking the attention can be addressed thus far. Because textbooks can reach families where social media or entertainment media cannot.

> *I grew up in a village. I can talk to you from that context.… It isn't easy to spread awareness in the villages. For example, I grew up in a Muslim family where watching TV was not allowed, as it is anticipated that kids will be misguided by watching cartoons or inappropriate programs. So, getting this knowledge through TV was/is challenging for many kids like me.* [b006]

Per our research, no story, article, or novel about BC is included in the secondary and higher-secondary curricula of Bangladesh and West Bengal, India.

Educating the Caregivers and the Family Members

Besides building infrastructure for mental health support, doctors and other associates must educate the family members and caregivers about the patient's mental health and how to deal with it. We heard stories suggesting how a family can motivate a survivor to fight BC:

> *Someone very close to me has breast cancer. Initially, when it was diagnosed, the doctor prescribed her some medicine but barely communicated with her about her mental health. So, she melted down. After battling breast cancer for two years, she started self-motivating and regained her mental strength. She is a mother of two young kids, so that pushed her to keep going and not give up. She involved herself in activities like Yoga. Due to regaining mental strength, she started handling the treatment better (I am not suggesting that the cancer has disappeared, but she started driving it better).* [b002]

> *She (a relative) went into depression after the mastectomy. But since she has kids, she motivated herself. Her family also supported her very strongly. This support has helped her come out of the trauma. She is still following up with her breast cancer-related*

XSL•FO
RenderX

treatment. It is four years that she has been receiving treatments. [b011]

Discussion

In this research study, we aimed to understand the attitudes and perceptions of South Asian people, specifically people from Bangladesh and West Bengal, India, regarding BC. Our goal was to identify common barriers to raising mass awareness about BC in this population. In the results section, we discussed the significant barriers to BC awareness, which included a lack of knowledge and open conversation, negligence of early symptoms, negative bias toward health care systems, judgmental attitudes, and insensitive comments toward patients and their families, and financial hardship with limited options for financial support. We also provided suggestions to effectively raise mass awareness, including encouraging open conversations among family members and different groups, discussing relevant topics in social events and places of religious practices, integrating the topic into school curriculums, and disseminating accurate information through social and traditional media.

We present, in this work, the summaries of our findings in Figure 2 and Figure 5, where we clustered the critical points into 5 themes based on different societal levels—individual, family, local society, health care system, and state or country. Our findings highlighted the need for long-term planning and cooperation between people and the government, collaboration between national and international organizations, the use of media and education systems, and the importance of modern technologies and mass media in raising BC awareness.

Considering our dataset size (N=17) and the demographics of the participants, particularly their socioeconomic status and educational background, it is important to acknowledge that our findings are not fully representative of the entire population. Specifically, our dataset may not adequately capture the experiences of individuals living in rural areas, those below the poverty line, or those in the wealthiest brackets. These socioeconomic statuses certainly have disparate health care experiences and struggles. Our participant pool does offer valuable insights into the perspectives of the upper-middle-class segment of society and gives a general overview of barriers that should be further investigated.

For future work, it will be essential to extend this study to include participants from a wider range of economic and educational backgrounds. This broader approach will enable us to evaluate the generalizability of our findings and recommendations across diverse demographic groups and to uncover any critical issues that may have been overlooked in the current analysis. Such an expansion will contribute to a more comprehensive understanding of the nuances of this important issue and increase the applicability of our conclusions to a broader audience.

Conflicts of Interest

None declared.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021;71(3):209-249 [FREE Full text] [doi: 10.3322/caac.21660] [Medline: 33538338]
2. Breast Cancer.: World Health Organization; 2024. URL: https://www.who.int/news-room/fact-sheets/detail/breast-cancer [accessed 2024-11-27]
3. Home | Government of India --- censusindia.gov.in? Government of India. URL: https://censusindia.gov.in/census.website/ [accessed 2023-07-26]
4. Press Information Bureau. Average Life Expectancy --- pib.gov.in,? Government of India. URL: https://pib.gov.in/PressReleseDetail.aspx?PRID=1606209 [accessed 2023-10-08]
5. Bangladesh Life Expectancy 1950-2023. URL: https://www.macrotrends.net/countries/BGD/bangladesh/life-expectancy [accessed 2023-07-31]
6. West Bengal Population| Sex Ratio| Literacy rate. 2023. URL: https://www.census2011.co.in/census/state/west+bengal.html [accessed 2023-10-19]
7. Leong SPL, Shen ZZ, Liu TJ, Agarwal G, Tajima T, Paik NS, et al. Is breast cancer the same disease in Asian and Western countries? World J Surg 2010;34(10):2308-2324 [FREE Full text] [doi: 10.1007/s00268-010-0683-1] [Medline: 20607258]
8. Hossain MS, Ferdous S, Karim-Kos HE. Breast cancer in South Asia: a Bangladeshi perspective. Cancer Epidemiol 2014;38(5):465-470. [doi: 10.1016/j.canep.2014.08.004] [Medline: 25182670]
9. Islam MR. Urbanization in Bangladesh: The prevalence of breast cancer brings unique challenges. 2021. URL: https://ascopost.com/issues/october-25-2021/urbanization-in-bangladesh-the-prevalence-of-breast-cancer-brings-unique-challenges/ [accessed 2023-10-08]
10. INDIA: BREAST CANCER. www.worldlifeexpectancy.com. 2024. URL: https://www.worldlifeexpectancy.com/india-breast-cancer [accessed 2024-11-27]
11. BANGLADESH: BREAST CANCER. worldlifeexpectancy.com. 2020. URL: https://www.worldlifeexpectancy.com/bangladesh-breast-cancer [accessed 2024-11-27]

XSL•FO
RenderX

12. Naderifar M, Goli H, Ghaljaie F. Snowball sampling: a purposeful method of sampling in qualitative research. Strides Dev Med Educ 2017;14(3). [doi: 10.5812/sdme.67670]
13. Jenks AC. From "lists of traits" to "open-mindedness": emerging issues in cultural competence education. Cult Med Psychiatry 2011;35(2):209-235. [doi: 10.1007/s11013-011-9212-4] [Medline: 21560030]
14. Bingham AJ, Witkowsky P. Deductive and inductive approaches to qualitative data analysis. In: Analyzing and Interpreting Qualitative Data: After the Interview. Washington DC: Sage; 2021:133-146.
15. Mattimoe R, Hayden MT, Murphy B, Ballantine J. Approaches to analysis of qualitative research data: a reflection on the manual and technological approaches. Accounting, Finance & Governance Review 2021;27(1):54-69 [FREE Full text]
16. Haque M, Kawsar M, Adibuzzaman S, Ahamed S, Love R, Dowla R, et al. Findings of e-ESAS: a mobile based symptom monitoring system for breast cancer patients in rural Bangladesh. 2012 Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; 2012 May 5-10; Austin Texas USA. [doi: 10.1145/2207676.2208532]
17. Amin MN, Uddin MG, Uddin MN, Rahaman MZ, Siddiqui SA, Hossain MS, et al. A hospital based survey to evaluate knowledge, awareness and perceived barriers regarding breast cancer screening among females in Bangladesh. Heliyon 2020;6(4):e03753 [FREE Full text] [doi: 10.1016/j.heliyon.2020.e03753] [Medline: 32382673]
18. Statistics of Breast Cancer in India. URL: https://cytecare.com/blog/breast-cancer/statistics-of-breast-cancer/ [accessed 2023-07-26]
19. Malvia S, Bagadi SA, Dubey US, Saxena S. Epidemiology of breast cancer in Indian women. Asia Pac J Clin Oncol 2017;13(4):289-295. [doi: 10.1111/ajco.12661] [Medline: 28181405]
20. Breast Cancer Risk in American Women. URL: https://www.cancer.gov/types/breast/risk-fact-sheet [accessed 2023-07-28]
21. Breast Cancer Epidemiology and Risk Factors: Epidemiology, Reproductive Factors, Endogenous Sex Hormones and Risk of Breast Cancer. URL: https://emedicine.medscape.com/article/1697353-overview [accessed 2023-07-26]
22. Dodge DG, Jarowenko AM. Diagnosed with breast cancer: stigmatized or member of an empowered sisterhood? In: Stigma and Prejudice: Touchstones in Understanding Diversity in Healthcare. Cham: Springer International Publishing; 2016:145-166.
23. Mansoor T, Abid S. Negotiating femininity, motherhood and beauty: experiences of Pakistani women breast cancer patients. Asian Journal of Women's Studies 2020;26(4):485-502. [doi: 10.1080/12259276.2020.1859076]
24. Agha N, Tarar MG. Battling breast cancer: women’s narratives of struggle, family support and survival from rural Sindh, Pakistan. Pak. J. Women Stud 2019;26(1):21-40 [FREE Full text]
25. McClelland SI. Gender and sexual labor near the end of life: advanced breast cancer and femininity norms. Women's Reprod Health 2017;4(1):29-45. [doi: 10.1080/23293691.2017.1276367]
26. Daniel S, Venkateswaran C, Hutchinson A, Johnson MJ. 'I don't talk about my distress to others; I feel that I have to suffer my problems...' voices of indian women with breast cancer: a qualitative interview study. Support Care Cancer 2021;29(5):2591-2600 [FREE Full text] [doi: 10.1007/s00520-020-05756-8] [Medline: 32955655]
27. Nagar I, Virk R. The struggle between the real and ideal: impact of acute media exposure on body image of young Indian women. Sage Open 2017;7(1):215824401769132. [doi: 10.1177/2158244017691327]
28. Mehrotra R, Yadav K. Breast cancer in India: present scenario and the challenges ahead. World J Clin Oncol 2022;13(3):209-218 [FREE Full text] [doi: 10.5306/wjco.v13.i3.209] [Medline: 35433294]
29. Maurya AP, Brahmachari S. Current status of breast cancer management in India. Indian J Surg 2020;83(6):316-321. [doi: 10.1007/s12262-020-02388-4]
30. van Weel C, Kassai R, Qidwai W, Kumar R, Bala K, Prasad Gupta P, et al. Primary healthcare policy implementation in South Asia. BMJ Glob Health 2016;1(2):e000057 [FREE Full text] [doi: 10.1136/bmjgh-2016-000057] [Medline: 28588938]
31. Story HL, Love RR, Salim R, Roberto AJ, Krieger JL, Ginsburg OM. Improving outcomes from breast cancer in a low-income country: lessons from Bangladesh. Int J Breast Cancer 2012;2012:423562 [FREE Full text] [doi: 10.1155/2012/423562] [Medline: 22295245]
32. Nagarajappa R, Bapat S, Asawa K, Daryani H, Gupta VV, Chaturvedi P, et al. Perception of sexual ramification among women toward male doctors in today's Indian society. Transl Behav Med 2018;8(5):771-775. [doi: 10.1093/tbm/ibx027] [Medline: 29385565]
33. Hall JA, Roter DL, Rand CS. Communication of affect between patient and physician. J Health Soc Behav 1981;22(1):18-30. [doi: 10.2307/2136365]
34. van den Brink-Muinen A, Verhaak PFM, Bensing JM, Bahrs O, Deveugele M, Gask L, et al. Communication in general practice: differences between European countries. Fam Pract 2003;20(4):478-485. [doi: 10.1093/fampra/cmg426] [Medline: 12876125]
35. Mahumud RA, Alamgir NI, Hossain MT, Baruwa E, Sultana M, Gow J, et al. Women's preferences for maternal healthcare services in Bangladesh: evidence from a discrete choice experiment. J Clin Med 2019;8(2):132 [FREE Full text] [doi: 10.3390/jcm8020132] [Medline: 30678044]
36. Begum SA, Mahmud T, Rahman T, Zannat J, Khatun F, Nahar K, et al. Knowledge, attitude and practice of Bangladeshi women towards breast cancer: a cross sectional study. Mymensingh Med J 2019;28(1):96-104. [Medline: 30755557]
37. Bajpai J, Mailankody S, Nair R, Surappa ST, Gupta S, Prabhash K, et al. Gender climate in Indian oncology: national survey report. ESMO Open 2020;5(2):e000671 [FREE Full text] [doi: 10.1136/esmoopen-2020-000671] [Medline: 32354723]

XSL•FO
RenderX

38. Oncologists in Mumbai Address Elephant in Room: Gender Gap Among Cancer Specialists. URL: https://timesofindia.indiatimes.com/city/mumbai/oncologists-address-elephant-in-room-gender-gap-among-cancer-specialists/articleshow/86761995.cms [accessed 2023-07-26]
39. Bureau M. Digital media dominates consumers' choice in India but growth in consumption appears to be slowing: Report. 2022. URL: https://www.medianews4u.com/digital-media-dominates-consumers-choice-in-india-but-growth-in-consumption-appears-to-be-slowing-report/ [accessed 2023-07-27]
40. Final Report, Assessment of the Media Sector in Bangladesh. USAID. 2022. URL: https://pdf.usaid.gov/pdf_docs/PA00Z9T6.pdf [accessed 2024-11-27]
41. Check-up camp for breast cancer in Lucknow. Lucknow News. 2023. URL: https://timesofindia.indiatimes.com/city/lucknow/check-up-camp-for-breast-cancer-in-lucknow/articleshow/98055384.cms [accessed 2023-07-27]
42. Breast Cancer Awareness Campaign for Early Detection of Cancer. 2022. URL: https://financialsamachar.com/breast-cancer-awareness-campaign-for-early-detection-of-cancer/ [accessed 2023-07-27]
43. Fresh Tissue observes 'Pink Day' on the occasion of breast cancer awareness month. 2022. URL: https://www.tbsnews.net/economy/corporates/fresh-tissue-observes-pink-day-occasion-breast-cancer-awareness-month-516138 [accessed 2023-07-28]
44. Shifting Tide: Breast Cancer is Number 1 Killer of Women. 2026. URL: https://timesofindia.indiatimes.com/home/science/shifting-tide-breast-cancer-is-number-1-killer-of-women/articleshow/51107576.cms [accessed 2023-07-27]
45. Ruby General Hospital YouTube page. Breast Cancer Awareness From Chandan Sen, Television Actor. 2021. URL: https://www.youtube.com/watch?v=OKfG0i2ZLM0 [accessed 2023-07-27]
46. BanglaVision PROGRAM YouTube page. Shastho Kotha| Ep 27| Tahamina Mukta| BanglaVision Program 2019. 2019. URL: https://www.youtube.com/watch?v=twxY5NVrto0 [accessed 2023-10-19]
47. Kemp S. Digital 2022: Bangladesh. Data Reportal Global Digital Insights. URL: https://datareportal.com/reports/digital-2022-bangladesh [accessed 2023-07-31]
48. Kemp S. Digital 2022: India. Data Reportal Global Digital Insights. URL: https://datareportal.com/reports/digital-2022-india [accessed 2023-07-31]
49. Number of WhatsApp users in selected countries worldwide as of June 2021. Statista. URL: https://www.statista.com/statistics/289778/countries-with-the-most-facebook-users/ [accessed 2023-07-31]
50. Solutions O. Ridesharing App OBHAI On WhatsApp - A First in Bangladesh. URL: https://www.obhai.com/blog/ride-with-obhai-through-whatsapp.html [accessed 2023-07-31]

Abbreviations

BC: breast cancer
CT: computed tomography
GLOBOCAN: Global Cancer Observatory
USAID: United States Agency for International Development

Edited by A Kushniruk, Y Gong; submitted 25.10.23; peer-reviewed by S Chaudhry, J Berian, DW Han; comments to author 04.03.24; revised version received 31.05.24; accepted 08.11.24; published 10.01.25.

Please cite as:
Hamid F, Roy T
Unveiling Sociocultural Barriers to Breast Cancer Awareness Among the South Asian Population: Case Study of Bangladesh and West Bengal, India
JMIR Hum Factors 2025;12:e53969
URL: https://humanfactors.jmir.org/2025/1/e53969
doi:10.2196/53969
PMID:

©Fahmida Hamid, Tania Roy. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 10.01.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Significance of a Cerebrovascular Accident Outcome Prediction Model for Patients, Family Members, and Health Care Professionals: Qualitative Evaluation Study

Corinne G Allaart[1,2], MSc; Sanne van Houwelingen[1,2], MSc; Pieter HE Hilkens[2], MD; Aart van Halteren[1,3], PhD; Douwe H Biesma[4], MD, PhD; Lea Dijksman[2], PhD; Paul B van der Nat[2,5,6], PhD

[1]Department of Computer Science, Faculty of Science, Vrije Universiteit Amsterdam, Amsterdam, Netherlands
[2]Department of Value Improvement, St. Antonius Hospital, Nieuwegein, Netherlands
[3]Philips Research, Eindhoven, Netherlands
[4]Leiden Unversity Medical Center, Leiden, Netherlands
[5]IQ Healthcare, Radboud University Medical Center, Nijmegen, Netherlands
[6]Santeon, Utrecht, Netherlands

Corresponding Author:
Corinne G Allaart, MSc
Department of Computer Science
Faculty of Science
Vrije Universiteit Amsterdam
Boelelaan 1111
Amsterdam, 1081 HV
Netherlands
Phone: 31 20 5989898
Email: c.g.allaart@vu.nl

Abstract

Background: Patients with cerebrovascular accident (CVA) should be involved in setting their rehabilitation goals. A personalized prediction of CVA outcomes would allow care professionals to better inform patients and informal caregivers. Several accurate prediction models have been created, but acceptance and proper implementation of the models are prerequisites for model adoption.

Objective: This study aimed to assess the added value of a prediction model for long-term outcomes of rehabilitation after CVA and evaluate how it can best be displayed, implemented, and integrated into the care process.

Methods: We designed a mock-up version, including visualizations, based on our recently developed prediction model. We conducted focus groups with CVA patients and informal caregivers, and separate focus groups with health care professionals (HCPs). Their opinions on the current information management and the model were analyzed using a thematic analysis approach. Lastly, a Measurement Instrument for Determinants of Innovations (MIDI) questionnaire was used to collect insights into the prediction model and visualizations with HCPs.

Results: The analysis of 6 focus groups, with 9 patients, 4 informal caregivers, and 8 HCPs, resulted in 10 themes in 3 categories: evaluation of the current care process (information absorption, expectations of rehabilitation, prediction of outcomes, and decision aid), content of the prediction model (reliability, relevance, and influence on the care process), and accessibility of the model (ease of understanding, model type preference, and moment of use). We extracted recommendations for the prediction model and visualizations. The results of the questionnaire survey (9 responses, 56% response rate) underscored the themes of the focus groups.

Conclusions: There is a need for the use of a prediction model to assess CVA outcomes, as indicated by the general approval of participants in both the focus groups and the questionnaire survey. We recommend that the prediction model be geared toward HCPs, as they can provide the context necessary for patients and informal caregivers. Good reliability and relevance of the prediction model will be essential for its wide adoption.

(JMIR Hum Factors 2025;12:e56521) doi:10.2196/56521

XSL•FO
RenderX

KEYWORDS

cerebrovascular accident; machine learning; artificial intelligence; visualization; focus groups; questionnaire; informal caregivers; health care professionals

Introduction

Cerebrovascular accident (CVA), the collective name for ischemic stroke and intracerebral hemorrhagic stroke, affects 15 million people each year globally [1]. A CVA can have a range of symptoms that vary from patient to patient. Unilateral paralysis or weakness, numbness or loss of vision, speech difficulties and other cognitive dysfunctions, ataxia, diplopia, and nonorthostatic dizziness are some of these symptoms [2]. After the acute phase of treatment, patients often undergo a long rehabilitation process, although the length, intensity, and outcome can vary greatly between patients. In the Netherlands, 65% of CVA patients undergo rehabilitation at home after their hospitalization. The remaining 35% of CVA patients visit a rehabilitation center or geriatric rehabilitation center. Among patients in this group, 25% will not return to their homes and will go to nursing homes [3].

To provide the best care, health care professionals (HCPs) and patients should decide together which path to take in the care process [4]: shared decision-making. Efficient and practical goals can be determined when patients are involved in decision-making, leading to a more appropriate and suitable discharge location. Thus, CVA patients should be involved in decision-making regarding where rehabilitation will take place and setting rehabilitation goals [5]. A considerable number of CVA patients prefer active involvement in decision-making and consider themselves capable of doing so [6]. An important part of supporting active involvement in the decision-making process is to offer patients insights into expected outcomes.

Recovery is very heterogeneous for CVA patients in terms of speed, process, and outcomes, and is often considered too complex for HCPs to predict. As such, little information about expected medium- and long-term outcomes (3 months to >1 year) is provided to patients. Over the last few years, using artificial intelligence (AI) methods, several outcome prediction models have been developed. Some models on certain aspects of functional status have been developed, such as rehabilitation of motoric upper limb skills [7] or independent walking [8]. Others have reported testing at rehabilitation admission [9] or the use of wearable sensors [10] to provide useful data to predict postrehabilitation CVA outcomes. Most of these models use machine learning algorithms on structured clinical data, with varying accuracy [11,12]. Deep learning has offered the opportunity to combine more data and more different data types, such as imaging or free text, into a prediction model [13] and as such increase the accuracy of predictions. Several studies have investigated how deep learning can lead to better CVA outcome predictions [13,14] to reach clinically acceptable reliability. One of our previous studies [15] showed that combining structured clinical data and perfusion computed tomography (CT) scans can lead to an increase in performance.

While having accurate and reliable predictions is important, the explainability and clinical relevance of AI methods are essential for their adoption in medical practice [16,17], and many prediction model implementations currently fall short in these aspects [18]. Using a functional tool based on AI in a real-world setting would require both HCPs and patients to understand and trust the outcome prediction and see its relevance. Understanding a predicted outcome can be challenging for healthy people, but even more for CVA patients. These patients can have impaired cognition and often have a higher age [2]. Thus, information should be provided in an easy and clear manner for low cognitive effort. For example, when numerical data are presented, it is recommended to use simple visual forms such as flow or bar charts and circle diagrams [19].

The discharge interview usually addresses the duration and location of the rehabilitation. There are typically multiple possibilities after discharge, such as going home without or with home therapy, outpatient rehabilitation, geriatric rehabilitation, and clinical rehabilitation treatment [3]. A personalized prediction of CVA outcomes and the length of the rehabilitation process would allow HCPs to better inform patients and informal caregivers. This could form a basis for shared decision-making, as a more informed conversation about the preferences of the patient can be held. This research aims to assess the added value of a prediction model for long-term outcomes of rehabilitation after CVA, to investigate how its results can be best displayed, and to determine the best way to implement and integrate it into the care process.

Methods

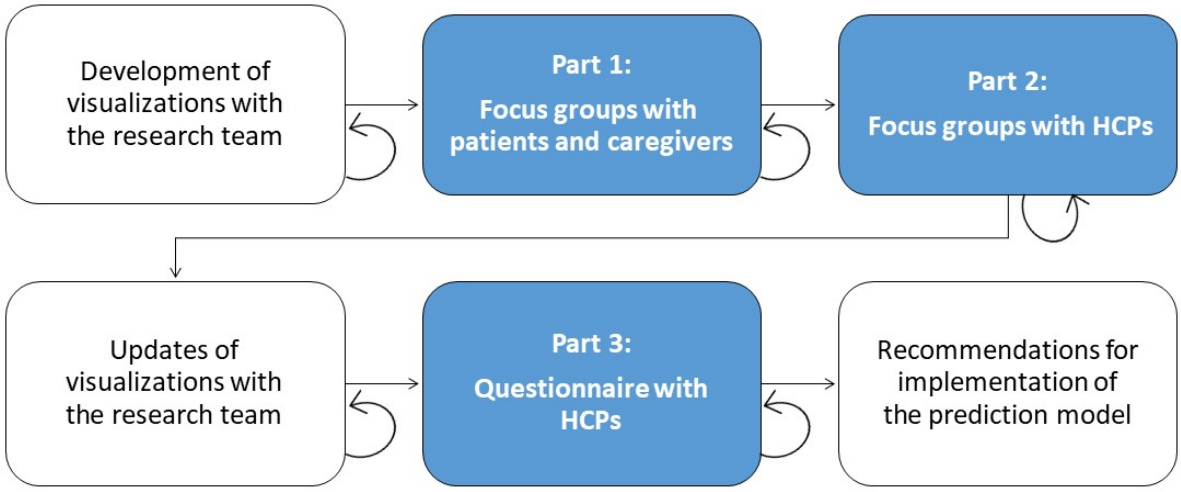
Design

A qualitative study was designed with the goal of evaluating the added value of a prediction model for stroke outcomes in CVA care. More specifically, this study aims to investigate how to best display the outcomes of the model, and integrate and implement them in the care process in an appropriate manner. We based our study on a prediction model previously designed in our hospital [15]. Before the qualitative evaluation, the research team developed a mock-up of a display of the prediction model, together with several different visualizations of the model outcome and interpretability. The research team included a neurologist and an expert on medical AI. The mock-up design was iterative, allowing for an update after each part of the study.

The study consisted of 3 parts: focus group discussions with CVA patients and their informal caregivers, focus group discussions with HCPs, and a follow-up questionnaire to HCPs, as shown in Figure 1. In both sets of focus groups, the opinions of patients and HCPs on the current information management and their needs with regard to the prediction model were explored and analyzed using a thematic analysis approach. The questionnaire was used to evaluate the finalized prediction model and visualizations with the HCPs.

XSL•FO
RenderX

Figure 1. Schematic overview of the different parts of the study. The circular arrow refers to the iterative aspect of each step. HCP: health care professional.



Focus Groups

Study Population

Patients who were admitted to St. Antonius Hospital between January 2020 and December 2021 for a CVA were selected using convenience sampling. Patients who had a CVA in the previous 3 months or who did not speak Dutch were excluded. We selected participants with a wide variation in age, gender, rehabilitation type (medical specialist, geriatric, home, or outpatient), and education level. An overview of the characteristics of the selected participants can be found in Table 1. Patients were approached by phone, and information letters were sent by email or postal address. Patients could bring their informal caregivers to the focus groups if this was of added value due to impaired cognition or higher age. The research team thought that an informal caregiver could assist in interpreting the model and could discuss patient preferences. Another advantage of including informal caregivers is that they are often involved in the entire hospitalization process of patients, and their opinions can therefore be valuable. Participants signed up by contacting the researcher and were asked to sign an informed consent form before the focus group started.

Table 1. Cerebrovascular accident participants in focus groups.

Participant	Age range (years)	Gender	Level of education	Rehabilitation
P1	71-75	Female	Vocational education	Geriatric rehabilitation center
P2	41-45	Male	Vocational education	Rehabilitation at home
P3	41-45	Female	Vocational education	Rehabilitation center
P4	71-75	Male	Applied university	Rehabilitation at home
P5	81-85	Female	High school	Rehabilitation at home
P6	66-70	Female	Applied university	Rehabilitation at home
P7	41-45	Male	High school	Rehabilitation at home
P8	76-80	Female	Vocational education	Geriatric rehabilitation center
P9	61-65	Female	Vocational education	Rehabilitation center

For HCPs, participants were recruited from a pool of HCPs involved in the discharge or rehabilitation process in the hospital. This included neurologists, rehabilitation doctors, junior doctors, nurses, and nurse specialists. All participants were approached via email or in person. The research team aimed to have 3 participants in each focus group, specifically a medical specialist, a general doctor, and a nurse.

Data Collection

The focus groups were led by 2 researchers (CGA and SvH). SvH moderated the focus groups with patients, and CGA moderated the focus groups with HCPs. We opted for the separation of patients and HCPs in different focus groups, such that the participants in both groups would be able to speak more freely about their experiences. We started with 3 focus groups for patients and their informal caregivers and 3 focus groups for HCPs, and we assessed whether saturation of data was reached or more focus groups were necessary. Each focus group had different participants.

The focus groups consisted of 2 phases. The first phase of the focus group was intended to explore the needs and expectations of the use of a prediction model during the discharge interview. This started with a reflection on the discharge interview and rehabilitation process, what expectations were created, what shared decisions were made, how the process went, and other

XSL•FO
RenderX

questions about topics and components that were addressed in the prediction model. The goal was also to determine how a prediction model could contribute to better expectations and improve the discharge interview.

In the second phase of the focus groups, different options for visualizations of the prediction model were discussed. We have elaborated on their design in the section on visualizations. Before the second phase of the focus group started, there was a short explanation of the visualizations. After this, the participants were asked to discuss several topics. First, whether the visualizations were clearly explained. Second, they were asked about relevance and what they think about the visualizations being used. During the focus groups with HCPs, this also included the predictors included in the model. Finally, they were asked to discuss which of the visualizations they preferred and why.

Focus Group Guide

Focus group guides were used, consisting of open and follow-up questions for the 2 phases of the focus groups (Multimedia Appendix 1 for patients and Multimedia Appendix 2 for HCPs). The focus group guides consisted of a few open questions, aimed at starting a discussion on information provision to patients in the current care process and the suitability of the prediction model and its visualizations. The focus groups followed an iterative design, which meant that after each focus group, the questions in the interview guide were adjusted if certain topics needed to be added or changed.

Data Analysis

Analysis of the focus groups was performed by researchers CGA and SvH in a thematic analysis approach using Atlas.ti [20]. CGA and SvH are early stage researchers with educational training and some research experience in focus group research. The thematic analysis was performed as follows. First, the researchers transcribed and anonymized the interviews. Then, both researchers separately open coded the transcripts. The data were coded axially by clustering codes by meaning, and the clusters were assigned themes. When all 6 transcripts were axially coded, the data were selectively coded to determine the relevance and consistency of the themes [21]. Finally, the themes were defined by noting the characteristics of the themes for consistent use. After analysis, the researchers held discussions about the data until a consensus was reached. We aimed for research triangulation by having 2 researchers independently analyze the data and iterate until agreement.

Questionnaire Survey

Study Population

The study population of the questionnaire consisted of HCPs. Similar to the focus groups, the participants were recruited from the pool of HCPs involved in the discharge or rehabilitation process in the hospital. However, only HCPs who directly worked with the prediction mode, such as junior doctors and neurologists, were asked to fill out the questionnaire. They were contacted by email or phone, or in person.

Data Collection

The last part of this study consisted of a questionnaire for a final analysis of the visualizations by HCPs. A simulated visualization (updated after the focus groups; see section visualizations) was shown to the HCPs for different scenarios for possible situations of patients in a discharge interview. They were asked to answer a modified Measurement Instrument for Determinants of Innovations (MIDI) questionnaire [22], which is a tool for describing the helpfulness and possibilities of medical technological devices. We selected 2 relevant sections out of 4 in the questionnaire regarding the innovation itself and the end users. The other 2 sections regarding work environment and sociopolitical environment were out of scope and more suitable for a later implementation phase. The questionnaire was adapted to fit the setting of the prediction model and was distributed via RedCap [23]. The questions were divided into 5 sections: general opinions, advantages and disadvantages, effects on patients, effects on colleagues, and opinions on implementation. The questionnaire consisted of 21 questions on a Likert scale and 1 open question.

Data Analysis

No statistical analyses were performed on the questionnaire, as the MIDI questionnaire is not a validated questionnaire and the sample size is small. Instead, the questions were meant to provide a formalized evaluation of the prediction model and its visualizations. Both the general score of the prediction model and the agreement have been reported.

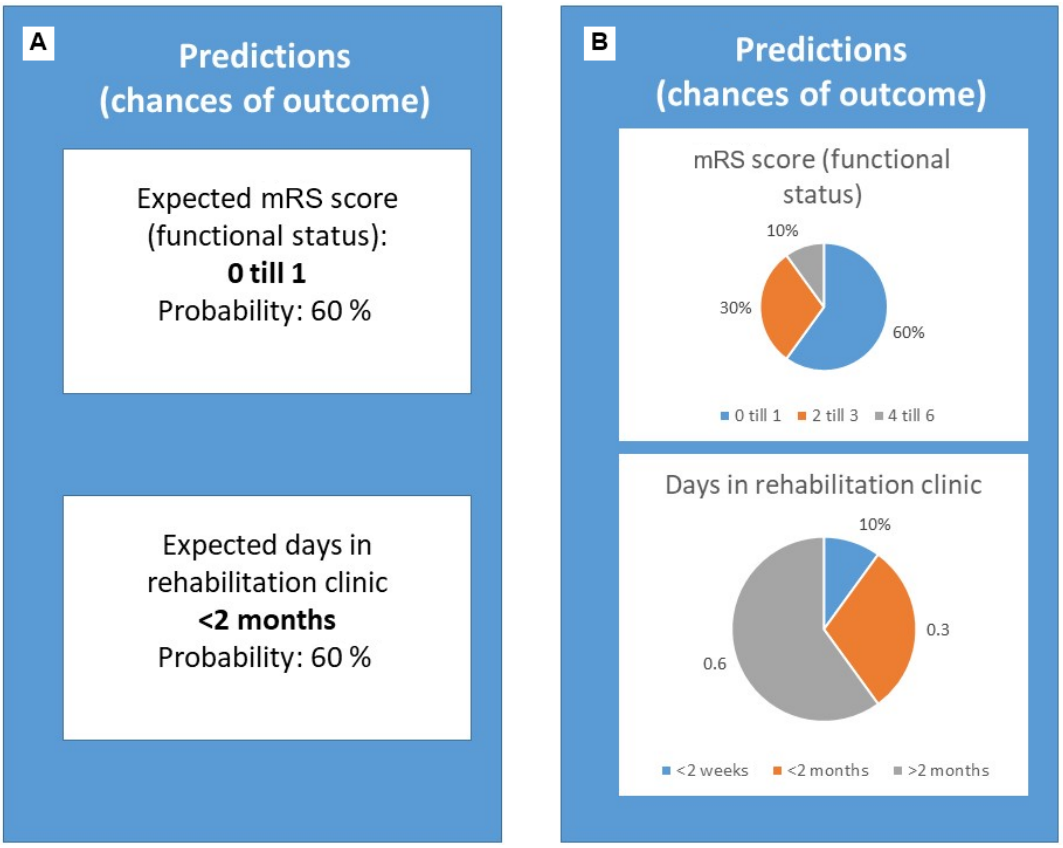
Prediction Model and Visualizations

The model visualizations were based on the model described in a previous article [24]. This model aims to predict a 2-fold outcome: functional outcome (modified Rankin scale [mRS] score) after 3 months and length of stay in a rehabilitation center. Two feature sets were used as predictors: hospital data and data from geriatric rehabilitation clinics. The hospital data were based on registry data from the Dutch Acute Stroke Audit [25], such as medical history, treatment data, and admission data, and the other data included intake data, admission data, and discharge data.

The design of the visualizations followed an iterative setup with the research team. Before the focus groups, the research team created a mock-up, with outcomes and visualizations for 2 fictional patients, as a functional model with real patient data was not yet available. Based on examples and previous literature about explainable AI in health care [26,27], a choice of 2 different visualization styles was made. One with a more simplified textual approach having clear and limited information, and one with a more detailed visualized approach having more information in the form of diagrams. We chose circle and bar charts for this style, as they are considered the most “approachable style” of visualization of numbers [19]. The visualizations of the outcomes in both styles are shown in Figure 2.

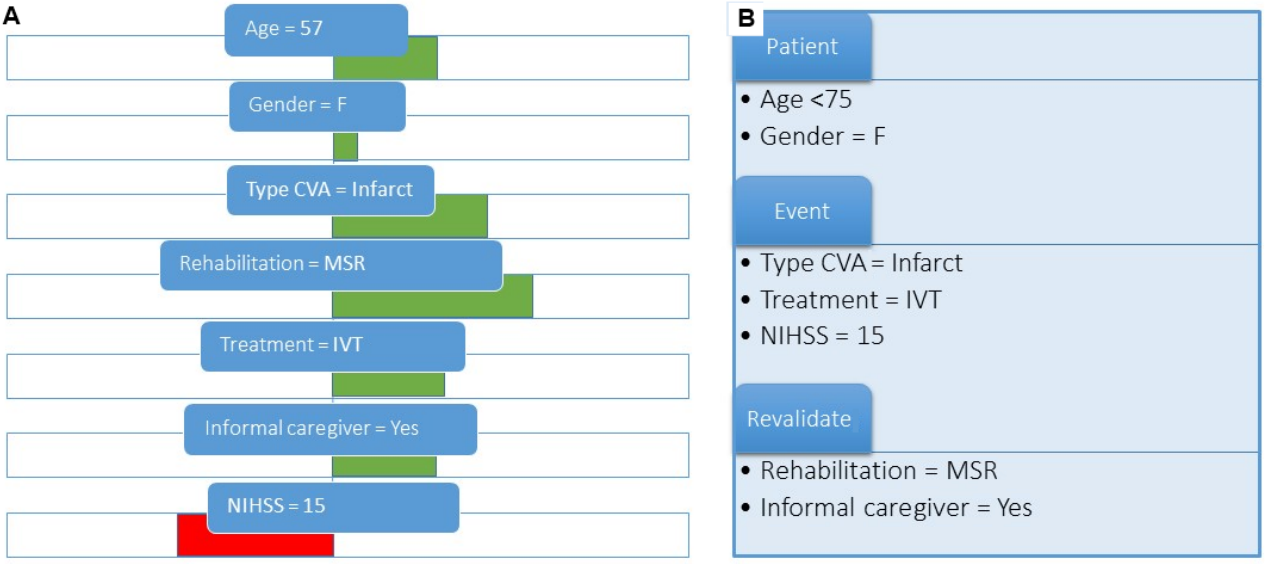
XSL•FO
RenderX

Figure 2. Different visualizations of outcomes (translated from Dutch). (A) Textual limited information; (B) Visual detailed information. mRS: modified Rankin scale.



Next to outcome visualizations, we also included explanations of the predicted outcomes by the models, which have been shown to be important for the adoption of AI in medicine [14]. We opted to show the model explainability through feature importance, where for the predicted outcome of the model for each patient, the contribution of each feature is shown. This can also be described as local interpretability, which has been further explained previously [19,23]. For example, it can show that the age of that specific patient contributed 20% to the decision of their personal outcome. This can be both positive or negative feature importance, indicating whether it made the predicted outcome more or less likely. The detailed visualization (Figure 3A) shows this in red and green bars. For the simplified version (Figure 3B), only the most important features for the outcome were mentioned (not the extent of their contribution).

Figure 3. Different visualizations of influencing factors (translated from Dutch). (A) Visual detailed information. The color and length of the bars indicate the size of the effect on the outcome. (B) Textual limited information. CVA: cerebrovascular accident; IVT: intravenous thrombolysis; MSR: medical specialistic rehabilitation; NIHSS: National Institutes of Health Stroke Scale.



XSL•FO
RenderX

Visualizations were developed for 2 fictional patients in 2 different visualization styles, resulting in 4 examples. The complete translated visualizations can be found in Multimedia Appendix 3. Based on the feedback from the focus groups, 1 visualization style was selected, and the visualizations were modified in accordance with the focus group feedback. These final visualizations were used for the questionnaire.

Ethical Considerations

The Medical Ethics committee of Utrecht (MEC-U) declared that this study is not subject to the Dutch Medical Research Involving Human Subjects Act (WMO) and issued a non-WMO statement under number W22.007, and subsequently, the board of directors of St. Antonius Hospital issued a statement of no objection (permission to perform) under number Z22.009. All participants provided informed consent.

Results

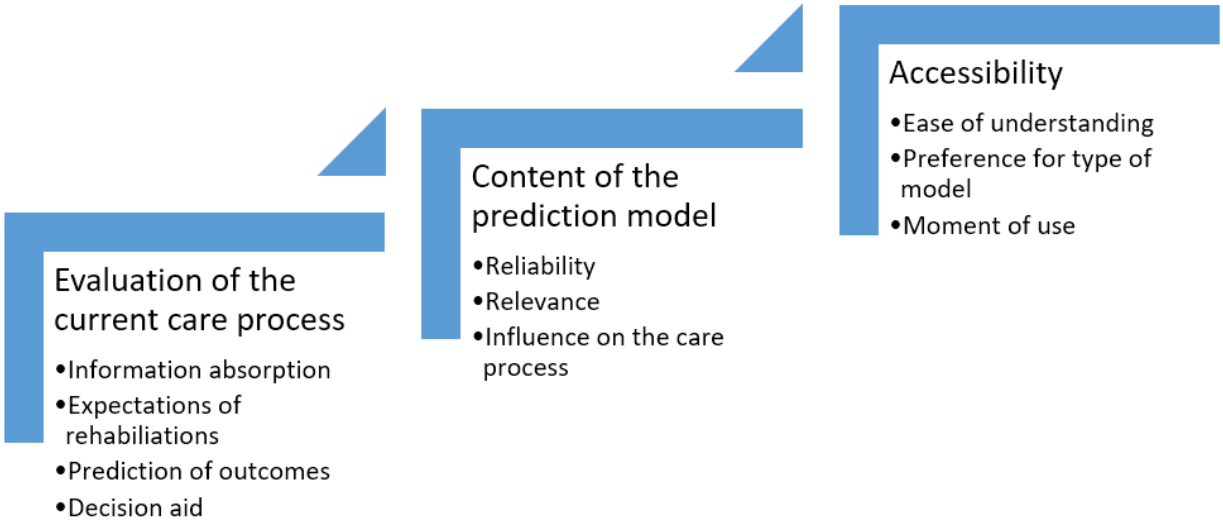
Focus Groups

The focus groups took place in April and May 2022 and lasted for 60-90 minutes with patients and caregivers and for about an hour with HCPs. We conducted 3 focus groups of each type (total 9 patients, 4 informal caregivers, and 8 HCPs). After these sessions, further focus groups were not deemed necessary, as saturation was reached.

A few minor changes were made to the interview guide after the first focus group with patients. We put less focus on the discharge interview and more on the questions about the care process in general for patients and the family interview for HCPs. For the focus groups with patients and informal caregivers, the transcripts were sent to these participants for a member check. None of the participants had any remarks on the transcripts.

The results of the focus groups have been organized by theme in prediction model development in 3 sections: evaluation of the current care process, content of the prediction model, and accessibility of the model (Figure 4).

Figure 4. The 10 themes in the focus groups divided into 3 categories.



Evaluation of the Current Care Process

Information Absorption

Some of the patients said that they either could not remember or did not understand the information given at the discharge interview. This could be due to the symptoms of their CVA. For example, a patient developed aphasia because of the CVA. As a result, she did not know what questions to ask during the discharge interview and did not manage to do so. However, most patients said this was due to information overload caused by discharge taking place shortly after the CVA, and the impressions and emotions during admission. On the other hand, most of the informal caregivers said that they received the information provided during the discharge interview and helped the patients with the information when necessary.

The patients and informal caregivers who could remember said that they were positive about the information provision, calling it clear, open, and informative, with a lot of room for questions. HCPs mentioned that they often found it hard to estimate whether patients and caregivers truly understood the given information, saying that patients always say they understand, even if they do not. They also mentioned language barriers as a common issue. Moreover, they were not sure whether the extra resources given, such as websites or information flyers, were read and understood by patients, as they did not discuss these with them.

Expectations of Rehabilitation

Most patients and informal caregivers had little idea of what to expect from rehabilitation and expressed that they felt thrown in at the deep end. One patient described her lack of expectations as follows:

> *I just let everything wash over me. That was all I could do.*

Only 1 patient said he knew what he could expect in his personal situation, even if daily tasks proved more difficult. Informal

XSL•FO
RenderX

caregivers expressed wishes for more information from the hospital about rehabilitation, with 1 informal caregiver expressing that he would have wanted a rehabilitation plan from the hospital:

> *[...] my mom is not the first one with a CVA, so there is a certain expectation and based on that […] protocols are made. It's my first time, and then there's nothing […] that's hard. Something has to be done about that. That's what I missed.*

Eventually, most patients were happy with the rehabilitation, except for a few bad experiences with HCPs. HCPs' experiences matched with the views of the patients and caregivers, and they stated that they do not have a good overview of rehabilitation and therefore cannot give patients an expectation of what will happen during rehabilitation. Some HCPs believed that this information should be given in rehabilitation clinics. They also mentioned that there could be a disparity in the expectations between patients and informal caregivers. In those situations, patients often felt like they were ready to be discharged home, while caregivers would still be very worried and push for inpatient rehabilitation.

Predictions of Outcomes

Almost all HCPs mentioned the difficulty of making individual predictions. They explained that this was due to patients, even similar patients, having varying recovery paths and because they did not have oversight over recovery during rehabilitation. Only the 2 most experienced neurologists felt more comfortable giving general predictions using age and the severity of the event. One of the junior doctors mentioned that they did not feel like they had enough knowledge and data to make such predictions.

> *I also don't know the exact scientific evidence from the top of my head […] for someone with a hemiparesis after 3 months.*

These statements were in agreement with the experiences of the patients. In the perception of the patients, if the physician did predict outcomes, it was not very precise. For example, they mentioned that the physician predicted that they would recover gradually in the first 3 months. Moreover, they said that no data in the form of numbers or percentages were used.

We asked all participants what they considered the most important outcomes to have a prediction on. Three outcomes were mainly mentioned: functional status in the medium or long term (multiple months or years), time of rehabilitation, and the chance of a new CVA. A new event was something that many patients were very scared about. They wanted to know how high the chance of recurrence was and wanted reassurance that it was as low as possible. HCPs mentioned it was not necessary to have a prediction of recurrence, as medications will always be given to minimize those chances.

Decision Aid

Many patients and informal caregivers felt like the physician took their opinions into account. According to the participants, it varied for each patient whether they had a choice in discharge location. Most participants said they did not have a choice in this decision. However, most participants felt that they could not make a better decision than a physician could, and they accepted the physician's opinion and professionalism. According to the participants, physicians have more knowledge about a CVA than patients and therefore know what is better for them.

> *You have to imagine; it was all new to us. So […] how can your opinion be different from what they recommend to you? It's almost impossible because you're in a situation like that so you can't judge it.*

Physicians mentioned shared decision-making. There is often not really any discussion on where the patient should be discharged.

Content of the Prediction Model

Reliability

Many patients were initially skeptical about a model that would predict outcomes. This was based on the physician's statement that it is not possible to make personal predictions. Several patients and informal caregivers said they did not believe that a physician could make accurate predictions and that consequences could only be seen after rehabilitation. Furthermore, some caregivers also found the prediction too absolute and programmed.

> *If it's such a well-founded prediction but we asked the neurologist about the whole thing, and he couldn't really give us any answers about predictions […]. Won't we get the wrong expectations from an overview like this?*

HCPs were also skeptical of the reliability of the model. They similarly argued that they see a lot of variation between patients and wondered if it is possible to make accurate predictions. In general, they believed that all included predictors would be usable. However, the HCPs questioned the reliability of the model due to the absence of certain important predictors, such as pre-existent functioning (Barthel score), detailed comorbidity, and the character of the patient. They acknowledged that the last aspect would be hard to measure. Moreover, if it is just a tool for information, the doctors argued that it is not that bad if the prediction is not 100% accurate**.**

> *[perfect performance] would be preferable, but you can never predict that […] maybe the important part is how to translate model performance to the patient*

One of the informal caregivers felt that there should be a disclaimer with the model because the outcome could never be 100% right.

Relevance

While most HCPs were generally enthusiastic about the proposed model, another point of skepticism was regarding relevance, as they were hesitant about how to explain the predicted outcomes to patients, especially regarding percentages. Patients confirmed these concerns. One of the patients said he did not like percentages because one can never say what 90%, for example, truly quantifies. Patients agreed that it is better to have a conversation and explain the information instead of simply showing the model. HCPs mentioned that while a general model for outcome prediction would be good, a more useful option would be if there were more different outcomes. For

XSL•FO
RenderX

example, where it would be more useful if we could split it into several types of complaints or symptoms (like aphasia or paralysis).

There was disagreement among patients regarding whether they wanted predictions at all. Some stated that they would be curious to know their predictions. Another participant stated that the information he received from his physician was sufficient and that he did not need a more specific prediction. Some of the caregivers said they would find it convenient to use a prediction model during a discharge conversation and that a prediction could supplement the information given by HCPs. Most HCPs were slightly worried about the risk of giving too much information, as it is currently already too much. However, some physicians also thought that a prediction could help clarify the situation, making the other information more digestible.

Influence on the Care Process

There were several points discussed concerning the possible influence of a prediction model on the care process. There was agreement among HCPs that it would be good for the patient as guidance if it is understandable for the patient. Junior doctors especially thought it would be good for HCPs as well for guidance in conversations with patients, reaching a better ground to make prognoses and overcoming unconscious biases about patients.

While most patients and informal caregivers expressed the positive aspects of guidance similar to HCPs, some patients worried that if the prediction is negative, it could be scary or it could demotivate the rehabilitation process. The patients stated that a lot of motivation is necessary for the best recovery.

> *And if I were [example patient] and 82 [years old] I don't know if I would be up for knowing […] that bad outcome*

However, most patients and informal caregivers did not want HCPs to hold back negative results from them. Moreover, doctors mentioned they would not be worried about this, as they are used to delivering bad news.

The main importance for HCPs was a way to inform patients and to offer clearer information for expectation management in the further rehabilitation process. As such, using this prediction tool would not have a direct result on the treatment or outcome of patients. There was the question of why this would be useful, and it was expressed that the goal of the tool should be clear; otherwise, it might not be used.

Accessibility

Ease of Understanding

The HCPs in general found the prediction model and the visualizations quite clear, and most thought that with explanation the model should be understandable for patients. However, several patients and informal caregivers found it hard to understand. Even after another explanation, they struggled to fully comprehend the meaning. One patient mentioned this might be due to one of the complications of his CVA being a lack of concentration and thus not having enough concentration to read. They struggled with the medical terminology, even after explanation. One patient said that it might be convenient for a physician but it is too hard for a patient. Most patients and caregivers understood the visualizations if they looked at them for a little longer but concurred that an explanation is needed from someone with a medical background or a more detailed but simple explanation is needed in additional text.

Preference for a Model Style

There were differences among patients regarding the preferences for the style of visualizations. Most of the participants preferred the prediction model with more graphics and less text because it was more visual and easier to understand. One participant stated that when a patient has symptoms like trouble with reading or concentration, it is easier to see the graphs.

> *I also think that for people who have […] more symptoms, who understand it a […] less easily […] the pictures are a little bit more informative.*

A few patients said the bar charts for feature importance did not have their preferences, as they contained too much information and the features could not be changed.

All HCPs preferred the more detailed visualization, and they argued that they contained more information and made the prediction more "alive." Interestingly, they thought patients would not have the same preferences. The HCPs also argued for integration into the electronic health record (EHR), as this would help with ease of use and adoption in the care process. Some HCPs were afraid that due to the high workload and the model not having a direct impact on the care process, the model might not be used if it is not easily accessible.

Moment of Use

Whether to use the model during the discharge interview or during the family interview, which happens a few days into admission with informal caregivers, was a point of discussion among HCPs. It could be beneficial for both. The discharge interview can be very rushed, and during the family interview, there might be more time to go over the predictions. However, this would be earlier in the process, which means that less information might be included in the predictions. Moreover, they agreed that having some form of access afterward would be good. Different options were mentioned: at the rehabilitation center, with the general practitioner, or at the 3-month follow-up.

The patients generally agreed with this view. One patient stated that it would be good to see the prediction when you feel like it, but it would have been too much during the discharge interview.

> *It is useful but […] too extensive [for] a normal person […] does not understand a thing. [...] Then it is explained by someone with a medical background […] who can perhaps explain it better. […] Maybe after a few months go back and look at it again […]. What are the expectations now, what are the expectations later.*

Questionnaire Survey

Based on the results of the focus groups, we extracted a set of recommendations. These recommendations were either immediately included in the visualizations of the prediction

XSL•FO
RenderX

model or served as general recommendations for further development or for the context of the implementation in current care. The recommendations are summarized in Table 2.

Table 2. Recommendations from the focus groups.

Category	Theme	Recommendation	Incorporation
Evaluation of current care	Information absorption	Recognize that the information might not be understood at the discharge interview	Implementation
Evaluation of current care	Expectations of rehabilitation	Create a better overview of rehabilitation for patients, informal caregivers, and health care professionals	Context of care
Evaluation of current care	Prediction of outcomes	Highlight that the model meets the needs of patients, informal caregivers, and health care professionals regarding providing a prediction of the outcome	Immediately
Evaluation of current care	Decision aid	Make sure the patient's opinion is heard in the decision process, for example, through shared decision-making	Context of care
Accessibility	Ease of understanding	The target audience of visualizations should be health care professionals and not patients	Immediately
Accessibility	Moment of use	Do not limit the use of prediction to the discharge interview	Implementation
Accessibility	Model preference	Visualize outcome and feature importance (visualization version b)	Immediately
Content of the model	Reliability	Include the general predictive performance of the model, for example, according to the margin of error	Future research
Content of the model	Relevance	mRS[a] score in 3 categories, including the possible functional score, measured during hospital stay	Future research
Content of the model	Influence on the care process	Clarify the goal of the model (informing patients) to improve adherence	Immediately

[a]mRS: modified Rankin scale.

The visualizations were adapted according to the recommendations. We selected the more “visualized” versions (see Figures 2B and 3A). Moreover, we focused on HCPs as our target audience and therefore removed explanations of well-known abbreviations and concepts, such as the National Institutes of Health Stroke Scale (NIHSS) score. We also specified that the moment of use could be the discharge or family interview and that the goal of the prediction model was to inform patients. Moreover, we included more detailed information on feature importance, based on our underlying model. The adapted visualizations can be found in Multimedia Appendix 3.

The questionnaire was sent to 16 relevant HCPs in St. Antonius Hospital, and we received 9 responses. The answers to the closed questions are summarized in Table 3. For each question, we have reported scores. Moreover, we have presented the average score and SD for each section. Negatively phrased questions (questions 4, 8, and 9) were scored in reverse.

XSL•FO
RenderX

Table 3. Questionnaire results reported on a Likert scale (0-5).

Question[a]	Score, mean (SD)
General statements	
1. The prediction model matches well with my usual way of working.	4.0 (0.5)
2. The prediction model is based on factually correct information.	3.7 (0.5)
3. The prediction model offers all the information and materials necessary to work with it properly.	3.6 (0.5)
4. The prediction model is too complicated for me to use.	4.1 (0.8)
5. I expect that the effects of the use of the prediction model will be clearly visible.	2.9 (0.4)
6. I think the prediction model is suitable for my patients.	3.8 (1.0)
Advantages and disadvantages	
7. The use of the prediction model gives me the opportunity to inform my patients better.	3.8 (1.0)
8. The use of the prediction model during discharge or family interviews will cost me more time than usual.	3.2 (0.5)
9. I find it too complicated to use the prediction model during discharge or family interviews.	3.9 (0.8)
Prediction model and patients	
11. I expect that with the prediction model, my patients will be better informed about the expected functional status after 3 months and the duration of rehabilitation.	3.4 (0.5)
12. I think it is important to use the prediction model to better inform my patients about the expected outcomes: the functional status after 3 months, and the duration of rehabilitation.	4.0 (0.5)
13. I consider it part of my job to use the prediction model.	3.6 (0.7)
14. Patients will generally be satisfied if I use this prediction model.	3.8 (0.4)
15. Patients will generally be cooperative if I use this prediction model.	3.7 (0.5)
Prediction model and colleagues	
16. I can count on sufficient help from my colleagues when necessary while using the prediction model.	3.5 (0.5)
17. How many doctors who conduct discharge interviews will actually use this prediction model?	3.8 (0.7)
18. To what extent will (your fellow) neurologists expect you to use the prediction model?	3.2 (0.4)
19. When it comes to working with the prediction model, how much do you care about the opinion of (your fellow) neurologists?	3.7 (0.5)
Application of the prediction model	
20. If you wanted to, do you think you would be able to apply the prediction model during the discharge or family discussions?	3.9 (0.3)
21. I have sufficient knowledge to be able to use the prediction model.	3.8 (0.8)
22. To what extent are you aware of the content of the prediction model?	3.1 (1.1)

[a]Question 10 was an open question (What are the other advantages or disadvantages of this prediction model?) and is not included in the table.

We considered statements with a score of ≥3.5 and an SD of <1.0 to be sufficient to positive, with agreement among HCPs. Most statements were received as such, and we considered the general feedback to be moderately positive. Moreover, in all categories of statements, we noted that most statements were at least moderately positive with agreement, highlighting that there is no topic where the prediction model consistently underperforms. The high scores for "*The prediction model matches well with my usual way of working*" (score 4.0), "*I think it is important to use the prediction model to better inform my patients about the expected outcomes: the functional status after 3 months, and the duration of rehabilitation*" (score 4.0), and "*If you wanted to, do you think you would be able to apply the prediction model during the discharge or family discussions?*" (score 3.9) indicate a high willingness to implement the prediction model.

For some questions/statements, the HCPs were not positive on average (score of <3.5) or were not in agreement (SD ≥1.0). The statement "*I expect that with the prediction model, my patients will be better informed about the expected functional status after 3 months and the duration of rehabilitation*" had a score of 3.4. This was just under our set threshold. Interestingly, the statement "*I think it is important to use the prediction model to better inform my patients about the expected outcomes: the functional status after 3 months, and the duration of rehabilitation*" had a score of 4.0. It shows that the HCPs, even when not sure about the positive outcome of use, still found it important to use the prediction model. The same concern was highlighted in 2 statements that had high disagreement among the HCPs: "*I think the prediction model is suitable for my patients*" (score 3.8, SD 1.0) and "*The use of the prediction model gives me the opportunity to inform my patients better*" (score 3.8, SD 1.0). This shows consistency with the focus

XSL•FO
RenderX

groups indicating that there seems to be variety in the extent to which HCPs are positive about the effects of the prediction model.

In the other statements that were not positive on average or had high disagreement, we also noted similar findings that came up in the focus groups as well. The statement "*Iexpect that the effects of the use of the prediction model will be clearly visible*" had the lowest score of 2.9. This is not completely unexpected as the model's primary function is to inform and not provide decision support. Moreover, the issue of high workload was reflected in the statement "*The use of the prediction model during discharge or family interviews will cost me more time than usual*" (score 3.2). Lastly, the statement "*To what extent are you aware of the content of the prediction model?*" had both a low score and high disagreement (score 3.1, SD 1.1), indicating that more detailed explanation and training are necessary to understand the model.

The statement "*To what extent will (your fellow) neurologists expect you to use the prediction model?*" had a low score of 3.2. This shows that many HCPs would consider the use of such a prediction model to be more elective. Next to the scaled questions, we also asked 1 open question "*what are the other advantages or disadvantages of this prediction model?*" A big point of contention was whether the prediction model would be accurate enough, as certain factors like recovery during the time in the hospital and mental resilience were not included. A positive aspect mentioned a few times was that it could also be used to preinform the rehabilitation clinics for helping with planning.

Discussion

Principal Findings

This study shows the added value of a prediction model for long-term outcomes of rehabilitation after stroke in CVA care under certain conditions. The need for a prediction model was consistently shown in focus groups and a questionnaire survey, as evidenced by the general approval of participants. Several recommendations can be made, most importantly on the following 5 conditions: primary target, timing of discussing the predictions, different focuses of HCPs and patients, goal of the prediction model, and visualization of the model.

First, the primary target audience of the prediction model should be HCPs. While the information provided is of interest to both patients and informal caregivers, not all patients can fully understand the outcomes of the prediction model. This resembles the findings of other studies that investigated different digital health tools for stroke patients, which reported complaints of difficulty [28] or disinterest among some patients [29]. Our study specifically found that patients prefer HCPs to provide information and guide them through the visualizations of the model.

Second, the timing for using predictions should also be considered. As patients mentioned, it could be more useful when targeted toward patients at different points in the care process. Here, it is interesting to see if the model can be simplified, as was requested by interested patients, and how that should be offered. Another study on digital health tools for patients with brain injuries [30] highlighted the importance of user testing among these patients. This would be a good way to ensure that the simplified model fits the patient's needs and understanding.

Third, for HCPs, the reliability and relevance of the model are key. HCPs focus largely on the reliability and relevance of the model. It is essential to not only provide good reliability and relevance but also clearly communicate the aspects to HCPs. Reliability is essential to validate the performance of the model. Therefore, changes in outcome reporting were considered to be more in line with the current care [31]: divide the mRS score into 3 categories and include predictive factors. By including the general predictive performance of the model (eg, according to the margin of error), the reliability can be better communicated to HCPs. These 2 points of reliability and relevance are very much in agreement with the guidelines for the implementation of AI in health care [32,33]. While the performance and generalizability of the model were not the focus of this study, they should be properly validated before model implementation. We found that patients focused more on the impact on rehabilitation, the emotional aspects, and understandability.

Fourth, HCPs and patients should understand the goal of the prediction model. This understanding can improve adherence and the willingness of HCPs to cooperate, which are essential for the adoption of the model in practice [32]. This matched our results, as this was one of the main points of the HCPs in both focus groups and the questionnaire survey. We aimed to clarify the goal of the prediction model after the focus groups. However, the questionnaires showed that the goal of the prediction model was still unclear to some HCPs, even after a more detailed explanation in the introduction of the questionnaire. This could be mitigated by properly informing HCPs, for example, by providing in-person training or a detailed demonstration of the model. It is also essential to highlight the goal of the explainability aspect of the model. The visualizations showed to what extent certain factors contributed to the model's decision, but that did not mean changing those factors would necessarily lead to a different outcome for the patient. For example, while the model might assign high importance to a certain treatment in the prediction of negative outcomes, it does not mean that not performing this treatment would lead to a better outcome.

Finally, all participant groups showed a preference for a more visual model. This matches the general consensus in visualization science, which argues for presenting numerical data in simple visualizations, such as bar or pie charts [15]. Our results showed that this is even more relevant for patients who have experienced a stroke. However, the visualizations that we used relied on red and green bars to highlight the importance of features, which can be an issue for colorblind people [34]. While none of our participants reported issues related to colorblindness, it should be adapted to a more inclusive visualization.

Strengths and Limitations

One of the main strengths of this study was the involvement of different stakeholder groups, including patients, informal

XSL•FO
RenderX

caregivers, and HCPs, with different roles. Moreover, the double setup with separate focus groups and questionnaires allowed for a check of agreement between the results of the different methods. One limitation to consider is our use of convenience sampling, which could have skewed the results. People with very severe stroke or with trouble reading or understanding were not represented. For HCPs, we mainly included people who were already open to research and new developments. The interviews with patients and informal caregivers were performed several months after the CVA, so their actual responses at the time of discharge might have been different. Moreover, we used mock-ups for the visualizations and not "real" predictions based on actual patient data. Although the visualizations were based on our existing models and data as much as possible, this was not a personal prediction. This made the visualizations more difficult to understand than expected, and the focus groups might have been more efficient with a functional model based on the actual patient data.

For the questionnaires, we had a relatively small sample size, and therefore, we did not perform any statistical analyses. As there was high agreement among the HCPs and the majority (9/16) of invited HCPs responded, it is likely they are a representative sample of the HCP group in the hospital, but we cannot know this for sure. Moreover, this group did not include all HCPs involved in stroke care. It is possible that physical therapists or general practitioners would have a different outlook on the prediction models, as they might be more involved in the rehabilitation of patients or involved at a different point in rehabilitation care. Lastly, this study was limited to a single hospital in the Netherlands. Differences between hospitals and especially between medical practices among countries can impact the generalizability of our results. Combining these limitations, we would recommend a quantitative follow-up study based on our questionnaire. Preferably, this should be a multicenter and international study to confirm the generalizability of our results outside the confines of our hospital.

Implications and Future Directions

From this study, we can extract lessons to be considered when implementing AI in medicine in general. An important aspect is that the need for information provision is highly variable between groups as well as individuals. The results of the prediction model can be highly personal, and similarly, the needs of a patient and the setting and specific illness of the patient can be personal. It is essential to clearly define the purpose of the tool, as there are not only different levels of understanding but also different priorities. For example, patients do not always appreciate detailed information, especially if the information is not actionable but only informative. It is also essential to consider a model that can have multiple purposes. A model, such as ours, could also be used for advanced resource planning in rehabilitation clinics by estimating the future demand on the clinics and evaluating the effectiveness of the rehabilitation.

An important point is to investigate how to include patients who struggle with technology due to language, socioeconomic, or medical issues. For example, this study excluded non-Dutch speakers, but HCPs mentioned that language issues could further complicate explaining the expectations of recovery. This difficulty in communication should not be increased by a new technological tool or prediction model. Relying more on technology can lead to more inequality in care [35], especially for vulnerable patients, such as stroke patients [32]. One effort that can prevent such issues is the development of a comprehensive framework for AI implementation in health care and the creation of a roadmap for AI implementation. Such frameworks are being developed [31,33], but none of them are exhaustive, and having resources that detail considerations for AI application and evaluation are still necessary to develop AI with meaningful medical applications [33].

Conclusions

The findings of this study show that participants are generally positive toward a prediction model for long-term outcomes of rehabilitation after stroke in CVA care under certain conditions, with a general preference for a more visualized prediction model. The prediction model should be geared toward HCPs, as they can provide the context necessary for patients and their informal caregivers. For HCPs, good reliability and relevance of the prediction model are essential for its proper integration. We recommend a quantitative follow-up study to confirm these results in multicenter and international settings.

Acknowledgments

This study is part of the project Enabling Personalized Interventions (EPI) with project number 628.011.028 of the research program Commit2Data-Data2Person, which is partly financed by the Dutch Research Council (NWO).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Guide for focus groups with patients (translated from Dutch).
[DOCX File , 16 KB - humanfactors_v12i1e56521_app1.docx]

Multimedia Appendix 2
Guide for focus groups with health care professionals (translated from Dutch).

XSL•FO
RenderX

[DOCX File , 15 KB - humanfactors_v12i1e56521_app2.docx]

Multimedia Appendix 3
Finalized visualizations (translated from Dutch).
[PNG File , 537 KB - humanfactors_v12i1e56521_app3.png]

References

1. Pan B, Jin X, Jun L, Qiu S, Zheng Q, Pan M. The relationship between smoking and stroke: A meta-analysis. Medicine (Baltimore) 2019 Mar;98(12):e14872 [FREE Full text] [doi: 10.1097/MD.0000000000014872] [Medline: 30896633]
2. Hankey GJ. Stroke. Lancet 2017 Feb 11;389(10069):641-654. [doi: 10.1016/S0140-6736(16)30962-X] [Medline: 27637676]
3. Geerars M, Wondergem R, Pisters MF. Decision-making on referral to primary care physiotherapy after inpatient stroke rehabilitation. J Stroke Cerebrovasc Dis 2021 May;30(5):105667 [FREE Full text] [doi: 10.1016/j.jstrokecerebrovasdis.2021.105667] [Medline: 33631474]
4. Armstrong MJ. Shared decision-making in stroke: an evolving approach to improved patient care. Stroke Vasc Neurol 2017 Jun;2(2):84-87 [FREE Full text] [doi: 10.1136/svn-2017-000081] [Medline: 28959495]
5. Krishnan S, Hay CC, Pappadis MR, Deutsch A, Reistetter TA. Stroke survivors' perspectives on post-acute rehabilitation options, goals, satisfaction, and transition to home. J Neurol Phys Ther 2019 Jul;43(3):160-167 [FREE Full text] [doi: 10.1097/NPT.0000000000000281] [Medline: 31205230]
6. Prick JCM, Zonjee VJ, van Schaik SM, Dahmen R, Garvelink MM, Brouwers PJAM, Santeon VBHC stroke group. Experiences with information provision and preferences for decision making of patients with acute stroke. Patient Educ Couns 2022 May;105(5):1123-1129 [FREE Full text] [doi: 10.1016/j.pec.2021.08.015] [Medline: 34462248]
7. Smith M, Ackerley S, Barber P, Byblow W, Stinear C. PREP2 algorithm predictions are correct at 2 years poststroke for most patients. Neurorehabil Neural Repair 2019 Aug;33(8):635-642. [doi: 10.1177/1545968319860481] [Medline: 31268414]
8. Smith M, Barber AP, Scrivener BJ, Stinear CM. The TWIST tool predicts when patients will recover independent walking after stroke: an observational study. Neurorehabil Neural Repair 2022 Jul;36(7):461-471. [doi: 10.1177/15459683221085287] [Medline: 35586876]
9. Harari Y, O'Brien MK, Lieber RL, Jayaraman A. Inpatient stroke rehabilitation: prediction of clinical outcomes using a machine-learning approach. J Neuroeng Rehabil 2020 Jun 10;17(1):71 [FREE Full text] [doi: 10.1186/s12984-020-00704-3] [Medline: 32522242]
10. O'Brien M, Lanotte F, Khazanchi R, Shin S, Lieber R, Ghaffari R, et al. Early prediction of poststroke rehabilitation outcomes using wearable sensors. Phys Ther 2024 Feb 01;104(2):pzad183 [FREE Full text] [doi: 10.1093/ptj/pzad183] [Medline: 38169444]
11. Bonkhoff AK, Grefkes C. Precision medicine in stroke: towards personalized outcome predictions using artificial intelligence. Brain 2022 Apr 18;145(2):457-475 [FREE Full text] [doi: 10.1093/brain/awab439] [Medline: 34918041]
12. Wang W, Kiik M, Peek N, Curcin V, Marshall IJ, Rudd AG, et al. A systematic review of machine learning models for predicting outcomes of stroke with structured data. PLoS One 2020;15(6):e0234722 [FREE Full text] [doi: 10.1371/journal.pone.0234722] [Medline: 32530947]
13. Campagnini S, Arienti C, Patrini M, Liuzzi P, Mannini A, Carrozza MC. Machine learning methods for functional recovery prediction and prognosis in post-stroke rehabilitation: a systematic review. J Neuroeng Rehabil 2022 Jun 03;19(1):54 [FREE Full text] [doi: 10.1186/s12984-022-01032-4] [Medline: 35659246]
14. Herzog L, Kook L, Hamann J, Globas C, Heldner MR, Seiffge D, et al. Deep learning versus neurologists: functional outcome prediction in LVO stroke patients undergoing mechanical thrombectomy. Stroke 2023 Jul;54(7):1761-1769. [doi: 10.1161/STROKEAHA.123.042496] [Medline: 37313740]
15. Borsos B, Allaart CG, van Halteren A. Predicting stroke outcome: A case for multimodal deep learning methods with tabular and CT Perfusion data. Artif Intell Med 2024 Jan;147:102719 [FREE Full text] [doi: 10.1016/j.artmed.2023.102719] [Medline: 38184355]
16. Combi C, Amico B, Bellazzi R, Holzinger A, Moore JH, Zitnik M, et al. A manifesto on explainability for artificial intelligence in medicine. Artif Intell Med 2022 Nov;133:102423 [FREE Full text] [doi: 10.1016/j.artmed.2022.102423] [Medline: 36328669]
17. Amann J, Blasimme A, Vayena E, Frey D, Madai VI, Precise4Q consortium. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. BMC Med Inform Decis Mak 2020 Nov 30;20(1):310. [doi: 10.1186/s12911-020-01332-6] [Medline: 33256715]
18. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, DECIDE-AI expert group. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nat Med 2022 May;28(5):924-933. [doi: 10.1038/s41591-022-01772-9] [Medline: 35585198]
19. Trevena LJ, Bonner C, Okan Y, Peters E, Gaissmaier W, Han PKJ, et al. Current challenges when using numbers in patient decision aids: advanced concepts. Med Decis Making 2021 Oct;41(7):834-847 [FREE Full text] [doi: 10.1177/0272989X21996342] [Medline: 33660535]

XSL•FO
RenderX

20. ATLAS.ti Scientific Software Development GmbH. Atlas.ti. URL: https://atlasti.com/ [accessed 2024-12-18]
21. Green J, Thorogood N. Qualitative Methods for Health Research. Thousand Oaks, CA: SAGE Publications; 2018.
22. Fleuren MAH, Paulussen TGWM, Van Dommelen P, Van Buuren S. Towards a measurement instrument for determinants of innovations. Int J Qual Health Care 2014 Oct;26(5):501-510 [FREE Full text] [doi: 10.1093/intqhc/mzu060] [Medline: 24951511]
23. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: Building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
24. Kassem J, Allaart C, Amiri S, Kebede M, Müller T, Turner R, et al. Building a Digital Health Twin for Personalized Intervention: The EPI Project. In: Commit2Data. Open Access Series in Informatics (OASIcs), Volume 124. Wadern, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik; 2024.
25. Kuhrij LS, Wouters MW, van den Berg-Vos RM, de Leeuw F, Nederkoorn PJ. The Dutch Acute Stroke Audit: Benchmarking acute stroke care in the Netherlands. Eur Stroke J 2018 Dec;3(4):361-368 [FREE Full text] [doi: 10.1177/2396987318787695] [Medline: 31236484]
26. Duell J. A Comparative Approach to Explainable Artificial Intelligence Methods in Application to High-Dimensional Electronic Health Records: Examining the Usability of XAI. arXiv. 2021. URL: https://arxiv.org/abs/2103.04951 [accessed 2024-12-18]
27. Ahmad M, Eckert C, Teredesai A. Interpretable Machine Learning in Healthcare. In: BCB '18: Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics. 2018 Presented at: ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics; August 29, 2018 to September 1, 2018; Washington, DC, USA. [doi: 10.1145/3233547.3233667]
28. Epalte K, Tomsone S, Vētra A, Bērziņa G. Patient experience using digital therapy "Vigo" for stroke patient recovery: a qualitative descriptive study. Disabil Rehabil Assist Technol 2023 Feb;18(2):175-184. [doi: 10.1080/17483107.2020.1839794] [Medline: 33155507]
29. Bally ELS, Cheng D, van Grieken A, Ferri Sanz M, Zanutto O, Carroll A, et al. Patients' perspectives regarding digital health technology to support self-management and improve integrated stroke care: qualitative interview study. J Med Internet Res 2023 Apr 04;25:e42556 [FREE Full text] [doi: 10.2196/42556] [Medline: 37014677]
30. Miao M, Morrow R, Salomon A, Mcculloch B, Evain J, Wright MR, et al. Digital health implementation strategies coproduced with adults with acquired brain injury, their close others, and clinicians: mixed methods study with collaborative autoethnography and network analysis. J Med Internet Res 2023 Sep 19;25:e46396 [FREE Full text] [doi: 10.2196/46396] [Medline: 37725413]
31. van de Sande D, Van Genderen ME, Smit JM, Huiskens J, Visser JJ, Veen RER, et al. Developing, implementing and governing artificial intelligence in medicine: a step-by-step approach to prevent an artificial intelligence winter. BMJ Health Care Inform 2022 Feb;29(1):e100495 [FREE Full text] [doi: 10.1136/bmjhci-2021-100495] [Medline: 35185012]
32. Verma A, Towfighi A, Brown A, Abhat A, Casillas A. Moving towards equity with digital health innovations for stroke care. Stroke 2022 Mar;53(3):689-697 [FREE Full text] [doi: 10.1161/STROKEAHA.121.035307] [Medline: 35124973]
33. de Hond AAH, Leeuwenberg AM, Hooft L, Kant IMJ, Nijman SWJ, van Os HJA, et al. Guidelines and quality criteria for artificial intelligence-based prediction models in healthcare: a scoping review. NPJ Digit Med 2022 Jan 10;5(1):2 [FREE Full text] [doi: 10.1038/s41746-021-00549-7] [Medline: 35013569]
34. Katsnelson A. Colour me better: fixing figures for colour blindness. Nature 2021 Oct;598(7879):224-225. [doi: 10.1038/d41586-021-02696-z] [Medline: 34608306]
35. Yao R, Zhang W, Evans R, Cao G, Rui T, Shen L. Inequities in health care services caused by the adoption of digital health technologies: scoping review. J Med Internet Res 2022 Mar 21;24(3):e34144 [FREE Full text] [doi: 10.2196/34144] [Medline: 35311682]

Abbreviations

AI: artificial intelligence
CVA: cerebrovascular accident
HCP: health care professional
MIDI: Measurement Instrument for Determinants of Innovations
mRS: modified Rankin scale

XSL•FO
RenderX

Edited by A Choudhury; submitted 22.01.24; peer-reviewed by A Hassan, T Abd El-Hafeez, R Gore, F Lanotte; comments to author 16.09.24; revised version received 29.10.24; accepted 20.11.24; published 22.01.25.

Please cite as:
Allaart CG, van Houwelingen S, Hilkens PHE, van Halteren A, Biesma DH, Dijksman L, van der Nat PB
The Significance of a Cerebrovascular Accident Outcome Prediction Model for Patients, Family Members, and Health Care Professionals: Qualitative Evaluation Study
JMIR Hum Factors 2025;12:e56521
URL: https://humanfactors.jmir.org/2025/1/e56521
doi:10.2196/56521
PMID:

©Corinne G Allaart, Sanne van Houwelingen, Pieter HE Hilkens, Aart van Halteren, Douwe H Biesma, Lea Dijksman, Paul B van der Nat. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 22.01.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Exploring the Users' Perspective of the Nationwide Self-Exclusion Service for Gambling Disorder, "Spelpaus": Qualitative Interview Study

Johanna Tjernberg[1], RN; Sara Helgesson[1], RN; Anders Håkansson[1], PhD; Helena Hansson[1], PhD

Department of Clinical Sciences Lund, Faculty of Medicine, Lund University, Lund, Sweden

Corresponding Author:
Anders Håkansson, PhD
Department of Clinical Sciences Lund
Faculty of Medicine
Lund University
Psychiatry, Baravägen 1
Lund, 221 00
Sweden
Phone: 46 703135677
Email: anders_c.hakansson@med.lu.se

Abstract

Background: Problem gambling and gambling disorder cause severe social, psychiatric, and financial consequences, and voluntary self-exclusion is a common harm reduction tool used by individuals with gambling problems.

Objective: The aim of this study was to explore users' experience of a novel nationwide, multioperator gambling self-exclusion service, "Spelpaus," in Sweden and to inform stakeholders and policy makers in order to improve harm reduction tools against gambling problems.

Methods: Semistructured interviews were conducted with 15 individuals who reported self-perceived gambling problems and who had experience of having used the self-exclusion service Spelpaus in Sweden. Interviews were transcribed and analyzed through qualitative content analysis.

Results: We identified 3 categories and 8 subcategories. The categories were (1) reasons for the decision to self-exclude, (2) positive experiences, and (3) suggestions for improvement. The subcategories identified a number of reasons for self-exclusion, such as financial reasons and family reasons, and positive experiences described as a relief from gambling; in addition, important suggestions for improvement were cited, such as a more gradual return to gambling post–self-exclusion, better ways to address loopholes in the system, and transfer from self-exclusion to treatment.

Conclusions: Voluntary self-exclusion from gambling, using a nationwide multioperator service, remains an appreciated harm-reducing tool. However, transfer from self-exclusion to treatment should be facilitated by policy making, and loopholes allowing for breaching of the self-exclusion need to be counteracted.

(JMIR Hum Factors 2025;12:e66045) doi:10.2196/66045

KEYWORDS

gambling disorder; gambling addiction; behavioral addiction; harm reduction; self-exclusion; voluntary self-exclusion; Spelpaus; lived experience; human factors; usability; qualitative study

Introduction

Background

Gambling disorder is a behavioral addictive disorder known to cause severe financial, social, and mental health consequences to affected individuals [1-3]. Among the harm-reducing or harm-preventive strategies used against gambling-related harm, voluntary self-exclusion is a well-established responsible gambling measure aiming to prevent and reduce the damage that can be caused by gambling [4].

Through self-exclusion, gamblers can voluntarily suspend themselves from gambling. However, 1 well-known limitation of self-exclusion services may be the possibility to gamble with another operator than the one the gambler has been self-excluded

XSL•FO
RenderX

from [5]. Gamblers are therefore usually required to self-exclude from multiple different operators. Consequently, in recent years, to further prevent gamblers from breaching their self-exclusion, it has become more common to offer gamblers the possibility to self-exclude from multiple operators at the same time.

In January 2019, the legislation for the Swedish gambling market changed, and monopoly was replaced by a license-regulated market [6]. This regulation meant that gambling operators, including overseas operators, needed to apply for a Swedish gambling license to operate on the Swedish market. In connection to the new regulations, the nationwide self-exclusion service "Spelpaus" was launched. Gambling operators with a Swedish license must adhere to this self-exclusion service, along with a series of other measures for responsible gambling. Spelpaus offers the possibility to self-exclude from all gambling operators with a Swedish license for 1, 3, 6, or 12 months. Self-exclusion cannot be canceled prematurely, and if the 12-month option is chosen, the self-exclusion continues beyond the 12 months, unless the individual actively chooses to cancel it after that period. Self-exclusion via Spelpaus also prohibits gambling operators from sending direct advertisements to self-excluded customers.

Spelpaus is in many ways a unique harm-reducing tool since it is nationwide, covers all licensed gambling operators (both online and land based), and can be accessed separately from the gambling operators. Thus, it theoretically represents a novel harm reduction strategy over and above other responsible gambling measures. Indeed, theoretically, this system may have a larger coverage and fewer loopholes (ie, fewer practical possibilities to turn into other gambling modalities and gamble despite self-exclusion) compared to a self-exclusion system involving 1 or few gambling operators. Around 100 gambling companies have a Swedish gambling license, with most of them operating exclusively online.

The prevalence of gambling problems in the Swedish population is estimated to be approximately 0.5% on a diagnostic level and a total of 1%-2% with at least moderate-risk gambling according to the Problem Gambling Severity Index [7,8]. Formal treatment seeking, or treatment provision, is low; less than 1000 people annually receive a gambling disorder diagnosis in the whole of Sweden, and even when taking other care providers into considerations, it can be assumed that only a small minority of people with a gambling problem on a diagnostic level receive formal treatment [9]. In contrast, low formal treatment seeking is often described in gambling disorder; however, a more informal but still active help-seeking behavior appears to be common, and voluntary self-exclusion from gambling can be part of this [10].

In Sweden, around 110,000 inhabitants are enrolled in Spelpaus [11], a number that has steadily increased since the launch in 2019 [12]. Thus, we believe that an increasing number of people with problematic gambling are attempting to intervene against their problem through self-exclusion, although we must consider that people without gambling problems also choose to self-exclude.

The steady increase in registrations in Spelpaus indicates that it is perceived as a helpful tool for many people. However, there are also limitations in the self-exclusion service, making it possible to breach the self-exclusion. Such a limitation is that Spelpaus only covers operators with a Swedish license, therefore making it possible for gamblers to gamble with overseas operators outside the Swedish license gambling market. In a 2020 web survey among gamblers, 38% of gamblers stated that they had gambled with unlicensed operators, while being self-excluded through Spelpaus [13], and in a later survey in 2022, the corresponding figure was 49% [14]. In another survey, conducted by the Swedish Gambling Authority, 25% gamblers stated that they had gambled with unlicensed operators during their self-exclusion [15]. In addition, continued gambling despite self-exclusion remains part of the clinical picture in clinical gambling disorder treatment in this setting [16].

Previous international experience of responsible gambling tools has demonstrated not only the feasibility of self-exclusion services and their use by individuals with gambling problems [4,17,18] but also that users experience limitations of these services [10,18]. Altogether, previous research calls for more in-depth examinations of user experience with self-exclusion services and suggestions for improvement of their effectiveness. In particular, this is relevant for the present type of self-exclusion service, which theoretically has been designed to overcome well-known limitations of older self-excluson systems; this novel Swedish service (1) is government based and independent of gambling operators, such that an individual does not need to enter a gambling site in order to self-exclude; (2) covers all licensed operators in the country and therefore prevents the possibility to change between licensed operators during an ongoing self-exclusion period; and (3) involves both land-based and online-based operators.

Thus, altogether, the relatively recent introduction of a unique, large-scale, multioperator self-exclusion service in Sweden has attracted a large number of users, but it is hindered by the risk of breaching one's self-exclusion to an extent that is clearly of relevance in the treatment setting. Therefore, it is of great value to study which effects of the Spelpaus service are perceived as favorable and challenging by its users and to also study this in more detail than in previous quantitative survey studies.

Aim

Based on the aforementioned research gap, the aim of this study was to enhance comprehension of the user experience of the Spelpaus service within the gambling community and to acquire a more profound insight into the reasons for use of this harm reduction tool and its advantages and challenges. Thereby, the overarching goal of the study was to inform policy makers about potential drawbacks and the potential for improvement of the harm reduction tool.

Methods

Study Design

This research was exploratory in nature. A semistructured interview guide was developed to ensure the exploration of diverse perspectives and experiences related to Spelpaus. Notably, this study represents the first attempt to comprehensively examine user experiences of Spelpaus, thus

XSL•FO
RenderX

justifying the use of qualitative content analysis for its methodological flexibility. The protocol of this study has been published previously [19].

An initial interview guide was developed based on the study's aim and research questions. A test interview with a peer support worker who had experience with self-exclusion services was conducted, leading to refinements of some questions to be more open ended. Further details of the interview questions are provided in the study protocol [19].

Context

The study was conducted in Sweden, a country with approximately 10.4 million inhabitants, where gambling is prevalent, with around 56% of individuals aged from 16 to 84 years having engaged in gambling activities in the past year. Moreover, 18% gamble at least once per month. The prevalence of gambling is higher among men aged 45-84 years, and online gambling has increased markedly in popularity in Sweden over the past decade [20]. Problem gambling in the present setting is predominated by online casino gambling and sports betting, with online casino gambling representing a large majority of patients seeking treatment at a gambling disorder facility in Sweden [16].

Sweden has a license gambling market since 2019, where online and land-based gambling operators are allowed to operate, provided they adhere to a number of responsible gambling regulations, one of which is adherence to the nationwide Spelpaus service. The legal gambling age in Sweden is 18 years [6]. After being a more traditional, land-based gambling market, but with a large predominance of a government-owned gambling monopoly for many years [21], the Swedish gambling market was gradually increasingly affected by overseas online operators, which represented a large proportion of visible gambling advertising despite their unregulated status [22]. This led to the decision to liberalize the gambling market into a licensed, but controlled, market from 2019. Given the new license-based gambling market since 2019, and thereby the relative recency of a unique self-exclusion service, the present setting is of interest to assess the qualities and challenges of such a new harm reduction instrument.

Potential stakeholders and users of the study's findings may include policy makers, public health officials, mental health professionals, gambling addiction counselors, researchers in the field of addiction studies, advocacy groups, and individuals and families affected by problem gambling. The insights garnered from this research could inform the development of targeted interventions, public health campaigns, regulatory measures, and support services aimed at addressing and mitigating the adverse effects of gambling behavior within the Swedish population.

Recruitment Process

Criterion sampling was used in the recruitment process. Participants were selected based on 2 critera: (1) being 18 years old or older during the time of the interview and (2) having previous or ongoing experience of using Spelpaus (any experience was enough, and it was not a requirement that the self-exclusion be in close temporal association with the interview). Recruitment was conducted online through social media advertisements managed by Trialy, a company specializing in research recruitment. Potential participants received information about the study and registered their interest via the Trialy website, completing a short online survey regarding their age and experience with Spelpaus. Eligible individuals were contacted by the first author (JT) and provided with additional study information. They were then sent written information about the study and about participants' rights and asked to sign and return (by mail) an informed consent form. When written consent was obtained, an interview was scheduled. Individuals were recruited during spring-summer 2023.

Sample

In total, 13 men and 2 women, aged between 31 and 62 years, participated in the study, representing various regions of Sweden. All participants reported either a gambling problem or identified themselves as being at risk, with prior or current experience using the Spelpaus self-exclusion service. This gender distribution aligned with existing research findings on gambling demographics in Sweden, which indicate a higher prevalence of gambling-related issues among men [16]. The gambling types involved in each person's gambling pattern were reported. Although the study did not systematically record whether gambling occurred (and had occurred) in online-based or land-based settings, the distribution of gambling types reported by the participants was comparable with the gambling types typically seen in Swedish problem gambling treatment settings, where online casino and sports betting is predominant [16].

Interviews

All interviews were digital (both video and audio) and were carried out by the first author (JT) of the paper, who has experience in qualitative interviewing and psychiatric research, especially in psychiatric research, ensuring a sensitive and nuanced data collection approach. The interviews were audiorecorded but not videorecorded and lasted between 23 and 67 minutes. During the interview process, the authors continuously reviewed and assessed the interviews to discern recurring patterns. Following the completion of 12 interviews, the emergence of new categories began to decline. In the final 2 interviews, only categories closely aligned with those identified previously surfaced. Consequently, the authors made the decision to conclude recruitment for the study. All interviews were then transcribed verbatim by the first and second authors (JT and SH).

As compensation for their time and efforts in the study, all participants received 2 cinema tickets.

Researcher Characteristics and Reflexivity

Our research team included professionals in psychiatry and social work, with substantial experience in qualitative research and interviewing. Throughout the study, we reflected on our backgrounds and preconceptions, particularly regarding gambling disorders and harm reduction strategies, and their potential influence on data interpretation. This reflexivity was integrated into our coding and analysis discussions to include

XSL•FO
RenderX

multiple perspectives and thereby to expand our own perspectives.

To ensure trustworthiness, triangulation was used throughout the analysis process. Multiple researchers independently coded the data, followed by discussions to reach consensus, minimizing the influence of individual biases and ensuring comprehensive data representation.

Data Analysis

This study used qualitative content analysis as the primary method for data analysis, chosen for its flexibility in systematically describing textual data, aligning well with the exploratory nature of this research. The analysis focused on manifest content, which involves identifying clear, descriptive categories that summarize data without deeper interpretive meanings. As noted by Graneheim and Lundman [23], categories represent the descriptive level of content and express the manifest content of the text, consistent with this study's goals. We deliberately chose not to identify themes, aiming instead to capture the descriptive aspects of participants' experiences, a recommended approach when focusing on manifest content rather than latent content [24,25]. Although themes can provide interpretative layers, we chose not to use them in this analysis, maintaining a descriptive focus to capture clear, observable patterns in participants' experiences with Spelpaus. This approach reflects the strength of qualitative content analysis in emphasizing the descriptive aspect without requiring thematic depth when the aim is to focus on manifest content [23].

The qualitative content analysis followed Graneheim and Lundman's [23] approach, focusing on identifying categories that represented the manifest content of participants' responses. Key analysis steps included repeated readings of the transcripts by 3 authors (JT, SH, and HH), identifying meaning units, condensing those units, and labeling them with codes, which were then grouped into categories based on similarities. The transcribed interviews were read through repeatedly and independently, and then the material was discussed by all authors. The content of the interviews related to the aim of the study was divided into meaning units, which were then condensed and labeled with codes. Similarities and differences in the codes were identified before they were divided into different categories and subcategories (Table 1). All steps of the analysis process were carried out separately and then discussed within the group to reach consensus. Various meetings within the group were held during the analysis process to discuss the different findings before a consensus about the results was reached.

Table 1. Examples of the data analysis process.

Meaning unit	Condensed meaning units	Code	Category	Subcategory
"If you should have a gambling stop like this, it should really apply to everything, because the brain will always find a way otherwise. This is the good thing—that it is a stop. You can't make a phone call, you can't just change your mind, because now it is a stop. I think that is very important." [Male participant, 31 years old]	The gambling stop should really apply to everything. Because the brain will always find a way otherwise. The good thing is that it is a stop, you cannot just change your mind. I think that is important.	Satisfaction with Spelpaus	Positive experiences	Prevention of future gambling problems
"Well, the downside is that it gives a little bit of a false sense of security. At least when it comes to certain people...as if you gamble on other gambling sites...it doesn't help anything." [Female participant, 35 years old]	The downside is that it gives a little bit of a false sense of security. When it comes to certain people who gamble on other sites, it does not help anything.	Suggestion for improvement: to also stop overseas gambling sites	Suggestions for improvement	Addressing loopholes

Ethical Considerations

The study was conducted according to the principles of the Declaration of Helsinki and was approved by the Swedish Ethical Review Authority (approval number 2022/06933-01, amendment 2023/01684-02). All participants received oral and written information about the study and about participants' rights, prior to signing a consent form. Written informed consent was obtained from all participants.

Results

Participant Details

A detailed description of the 15 participants is presented in Table 2.

XSL•FO
RenderX

Table 2. Background information about the participants (N=15).

Characteristics	Participants, n (%)
Social relationships	
In a relationship	8 (53)
Have children	9 (60)
Gambling form	
Casino	3 (20)
Sports betting	1 (7)
Horse race betting	0
Lottery	1 (7)
Casino + sports betting	2 (13)
Casino + poker	3 (20)
Casino + horse race betting	1 (7)
>2 different gambling forms	4 (27)
Enrolled in Spelpaus at the time of the interview	
Yes	10 (67)
No	3 (20)
N/A[a]	2 (13)
In active gambling at the time of the interview	
Yes	3 (20)
No	11 (73)
N/A	1 (7)
Number of times using Spelpaus	
1	4 (27)
≥2	11 (73)
Chosen time interval of self-exclusion	
12 months	7 (47)
1, 3, or 6 months	3 (20)
Both of the above	5 (33)
Overseas gambling during Spelpaus	
Yes	6 (40)
No	9 (60)

[a]Not applicable.

Categories and Subcategories

As shown in Table 3, we identified 3 categories: motivations and reasons for self-exclusion (including triggers, such as financial strain and concern for close relationships), perceived positive experiences (eg, reduced gambling urges and improved quality of life), and challenges and suggestions for improvement (notably, breaches through unlicensed gambling sites and lack of integrated support).

XSL•FO
RenderX

Table 3. Categories and subcategories.

Category	Subcategories
Reasons for the decision to self-exclude	• Precipitating factors • Self-awareness of gambling habits • Concern for close ones
Positive experiences	• Prevention of future gambling problems • Improved quality of life
Suggestions for improvement	• Integrated support systems • Addressing loopholes • Practical enhancements

Reasons for the Decision to Self-Exclude

The motivations underlying the choice to use the self-exclusion service were multifaceted, often encompassing a combination of factors.

Precipitating Factors

Numerous participants reported specific events that precipitated heightened gambling activity, ultimately leading to the decision to self-exclude. Common triggers included winning a large amount of money and quickly gambling it away, traumatic family incidents, interpersonal conflicts, financial struggles, and the onset of the COVID-19 pandemic. For instance, 1 (7%) participant articulated:

> *I was facing unemployment without insurance, which exacerbated my gambling tendencies. Although I secured employment thereafter, the meager income did not deter my gambling. The resurgence of COVID-19 in September restarted my gambling habits, leading to a downward spiral.* [Male participant, 44 years old]

Additionally, some participants opted for self-exclusion to prevent the risk of others accessing their gambling accounts as part of gambling-related criminal behavior. An example of this was the following interviewee who had a gambling problem but who also had a family member with a severe gambling problem:

> *...At the same time, this has helped us a lot every time she has wished to gamble on our accounts.* [Female participant, 35 years old]

Another reason to self-exclude was because the individual wished to facilitate debt settlement, as self-exclusion was a prerequisite for debt relief applications.

> *I applied for a debt settlement, but since I did not take it [self-exclusion] for a year, I did not get it…I would like to make another attempt for a debt settlement application, but then I probably have to take a 12-month period [of self-exclusion] as well.*

Self-Awareness of Gambling Habits

A prevalent reason for self-exclusion was an enhanced awareness of one's gambling behavior. Participants acknowledged excessive gambling tendencies, albeit with varying perspectives on whether they had developed a full-fledged gambling problem or were teetering on the brink of one. Fueled by these insights, individuals opted for Spelpaus to forestall further progression into problem gambling or to address existing issues. As 1 (7%) participant reflected:

> *I have had a realization, particularly in recent months...I recognize that I need to address this behavior before it escalates.* [Female participant, 56 years old]

This introspective process often unfolded gradually, sometimes facilitated by professional guidance or peer support.

Concern for Close Ones

Many participants described the adverse impact of their gambling habits on loved ones, with 10 (67%) individuals mentioning that their loved ones had been negatively impacted to a great extent. Financial strain, deceit, and compromised relationships were common themes. Some participants lamented fractured relationships resulting from their gambling, while others acknowledged the toll gambling took on family bonds and quality time spent together. One participant expressed:

> *It is not just about the money lost—it is the betrayal, the lies...and the time squandered gambling instead of being with my family that truly stings.* [Male participant, 39 years old]

Although self-exclusion was often an independent decision, familial influence or shared deliberation with loved ones also played a role. The following 2 (13%) citations are examples of this:

> *The relationship with my parents is starting to get a little bit better; for many years, it was pretty bad. They have helped me out with money a lot of times. Every time I call them, they think I am going to ask them for money again, so it [gambling] has hurt our relationship a lot.* [Male participant, 44 years old]

> *If you fail in a game or something, you get kind of aggressive, depressed, and angry. You also get very affected emotionally. If you bet on 7 games and 1 of them fails, you get upset, angry, or sad. It has affected my close ones the most.* [Male participant, 40 years old]

Positive Experiences

Participants identified several benefits associated with using the Spelpaus self-exclusion service.

XSL•FO
RenderX

Prevention of Future Gambling Problems

Self-exclusion was perceived as a preemptive measure to safeguard against relapse into gambling behaviors. Even individuals who had not experienced gambling urges for an extended period recognized the potential for future vulnerability. By enacting self-exclusion, participants closed the door to impulsive gambling, thus mitigating the risk of relapse. Several participants credited the Spelpaus service for preventing immediate urges to gamble. A couple of the participants who began gambling again described how they after the end of the self-exclusion period managed to gamble in a more controlled way than before.

> *After the self-exclusion, it was much more under control. I gambled a lot less, bet smaller amounts…if I felt I need to limit it now or I have to stop now. When you feel it is starting to be too much or take too much time, you feel a little self-control. Then I take a free weekend or several days free from betting, and after that, I bet a little, and then it feels like everything is under control.* [Male participant, 40 years old]

Another participant described the partial effect on gambling, primarily on impulse-driven uncontrolled gambling:

> *A part of me does not really want to say goodbye and never gamble again, but I want it to be at a reasonable level. It is at ATG [Swedish horse-betting company]. I am staying there. I think I gamble for 96 (Swedish) kronor a week now, but it was still that break (self-exclusion) that forced me to deal with it somehow.* [Male participant, 55 years old]

Improved Quality of Life

Many participants reported enhanced well-being during the self-exclusion period. Freed from the grip of gambling, individuals experienced a newfound sense of calm, happiness, and control over their lives. Moreover, self-exclusion provided an opportunity for personal growth and self-improvement, allowing participants to redirect their focus toward constructive pursuits, such as physical exercise and social interactions. One individual reflected:

> *You can relax because you have given away the control to something else…once I have chosen to self-exclude, I cannot do anything about it…and I get so mentally relaxed, I do not have to think about it at all anymore.* [Male participant, 32 years old]

Financial stability and restored trust within familial relationships were also cited as positive outcomes, as 1 (7%) respondent expressed:

> *My family can tell that I feel better when I do not gamble. I am happier, more positive, and I do not call them and ask for help. Of course, they notice a big difference when I do not gamble.* [Male participant, 31 years old]

In addition, the self-exclusion period provided an opportunity for reflection within the motivational process:

> *When I had my Spelpaus, I became quite aware of the problem, so I had a lot of thoughts about the need to do something about this problem.* [Male participant, 40 years old]

Suggestions for Improvement

Participants identified several areas for enhancement within the Spelpaus self-exclusion service.

Integrated Support Systems

Acknowledging self-exclusion as 1 component of the multifaceted approach to gambling cessation, participants emphasized the need for additional support mechanisms. One participant expressed:

> *My fear is that when the self-exclusion disappears, I have not come so far that I will not get stuck in gambling again…I think it is important that you try to solve your problems when you are self-excluded.* [Male participant, 43 years old]

Peer support groups, counseling services, and psychotherapy were deemed vital complements to self-exclusion. However, participants expressed reluctance to seek out support independently, underscoring the importance of proactive outreach and integrated support systems from within Spelpaus. Some of the participants who considered themselves to be more at risk of developing a gambling problem expressed that Spelpaus alone was enough for them.

> *If you take a year, and you have the opportunity to end it, then there needs to be some kind of dialogue before, and some kind of amount limit.* [Male participant, 55 years old]

> *I think it should be combined with a guaranteed CBT [cognitive behavioural therapy] or diary where you write what you feel, what you can do and what you think you should be able to do. Then of course, in the best of worlds, group therapy is perhaps even better than being digital.* [Male participant, 55, years old]

Addressing Loopholes

Participants raised concerns regarding the efficacy of Spelpaus in preventing access to overseas gambling sites. The absence of coverage for nondomestic operators posed a significant loophole, enabling some individuals to circumvent self-exclusion measures. Suggestions for improvement included enhanced international cooperation and stricter regulations on gambling advertisements.

Practical Enhancements

Participants proposed various technical and logistical improvements to Spelpaus, including more flexible and expanded self-exclusion options, longer time intervals, and a more deliberative process for ending self-exclusion periods. Mainly, participants asked for time intervals longer than 12 months. Several participants saw it as problematic to be able to gamble immediately after the end of a self-exclusion period. Additionally, participants advocated for heightened demands on gambling operators to deter relapse and streamline the self-exclusion process.

> *The self-exclusion service should be able to be developed so you cannot access overseas sites…when*

XSL•FO
RenderX

> *you select it, it should be covering all sites within the EU [European Union] or something like that…that you have a self-exclusion service for all of Europe when you choose it.* [Male participant, 62 years old]
>
> *The best thing would have been if there was a self-exclusion site where you are self-excluded from all gambling…that you have some kind of cooperation with other countries and everything and that it will be a stop everywhere.* [Female participant, 35 years old]

Several respondents stated that the possibility of returning to gambling after the self-exclusion period should still be hindered or slowed down in order to prevent impulse-related relapses even after the predetermined self-exclusion period is over. Some examples of this was an "embargo" period during which gambling still could not be reopened. Some participants found themselves waiting to gamble immediately after the self-exclusion period ended, so they suggested there be some kind of deposit limit for all clients coming back from self-exclusion, a mandatory telephone call or other message to individuals for whom self-exckusion is ending, prohibition of direct advertising to individuals immediately after a self-exclusion period, or mechanisms making the risk of gambling during alcohol influence smaller (eg, through prohibition of gambling during the night).

One respondent described the need for an embargo period:

> *After 12 months, you should not be able to cancel it on the same day; it should be 1, 2, 3 months notice…otherwise, when 12 months have passed, you can just gamble without any restrictions.* [Male participant, 32 years old]

Other respondents described the need for other possible measures to apply in the policy of self-exclusion:

> *After the self-exclusion ended, the advertising started coming again. It becomes like a reminder. So one could simply get going and gamble again!* [Male participant, 31 years old]
>
> *There should be some time to think it over when a client has asked to come back to gambling again. Some time should elapse, and then the person should be asked again. After some time, you should be asked about whether you really want to gamble again. There are people who get a kick, but after some time, that feeling migh pass again.* [Male participant, 44 years old]

Discussion

Principal Findings

This study aimed to deepen the understanding of the user experience of a nationwide, multioperator self-exclusion service, beyond what is known from recent quantitative research studies with web panel members and clinical data from patients in treatment for a gambling disorder. The potential challenges of a multioperator self-exclusion service have been demonstrated in surveys, showing that primarily among individuals with highly intense gambling practices, self-exclusion is popular but often breached by its users [13,14]. However, advantages of being self-excluded, and a deeper understanding of how this is related to the challenges of the method, largely remain to be understood.

Here, this qualitative study resulted in the description of 3 overarching categories that provide a deeper understanding of the problem: the main reasons for the decision to self-exclude, the advantages perceived by the users of this method, and the users' suggestions for improvements.

The decision to self-exclude from gambling platforms is often prompted by a sudden realization of escalating gambling behavior and a subsequent desire to regain control. Although the development of a disordered gambling pattern or financial difficulties were reported to have contributed to the decision to self-exclude, it was also reported, less expectedly, that self-exclusion may also be triggered by a wish to prevent another person from gambling on one's account. This may be a less expected reason for self-exclusion and will merit further investigation, especially in relation to the role of gambling in criminal behavior. Gambling or conducting financial transactions using somebody else's identity may represent one of the features of a severe gambling problem and also constitutes a criminal act [26].

Importantly, the role of the gambler's concerned significant others emerged as one of the reasons for self-exclusion. This highlights the importance of further addressing the interplay between individuals with gambling problems and their close ones and may also facilitate treatment processes if families of the affected individual are actively involved. This finding is consistent with previous research indicating that concerned significant others can favor treatment seeking and a favorable treatment outcome in their near ones with gambling problems [27,28]. In addition, the active participation of partners in the treatment of patients with a gambling disorder can be beneficial, both to the outcome of the affected patients and to the well-being of patients and their partners [29]. Thus, altogether, concerned significant others appear to play an important role through different phases of help seeking in individuals with problem gambling, even in the earlier harm-reducing interventions that self-exclusion is meant to represent.

Furthermore, the data showed that one reason for self-exclusion may be the intention to demonstrate one's willingness to quit gambling when formally applying for a public service where this is either required or believed to have an effect favoring the application process. This type of instrumental decision to self-exclude in order to obtain specific services or to fulfill expectations of external decision makers rarely has been reported in this context before. Whether this happens as part of an individual motivational process or entirely based on the requirements from others remains to be studied.

When external factors are cited as a reason to self-exclude, whether it be the concern of family members or the requirements from a treatment setting or authority, it can be argued that the choice to self-exclude does not represent the same decisive step in a motivational process as it does when an individual self-excludes based on their own decision to stop gambling in order to prevent or alleviate symptoms of addiction. Thus, the

XSL•FO
RenderX

possibly different levels of personal commitment associated with self-excluding will be relevant to include in the theoretical framework of self-exclusion in the future.

Overall, many favorable effects from self-exclusion were cited, despite previous experience where research has highlighted some major limitations of this service. Previous studies have been conducted using quite different methodologies, relying on larger quantitative data sets from online surveys [13,14] and clinical documentation [16], and where the findings were primarily the high rates of breaching one's self-exclusion within this service. This paper, importantly, adds to the literature regarding the present type of multioperator, nationwide self-exclusion service; in previous publications from this setting, it has been highlighted that the rate of breaching of one's self-exclusion is high in people with a gambling disorder [16] or with intensive online gambling patterns [14]. Here, several respondents described a large favorable impact on life after having self-excluded. Therefore, the description of more favorable experiences from this study provides a valuable experience, which contributes to the overall picture of this model, and strongly deepens the understanding of how this service fairs in affected individuals, over and above previous analyses conducted with quantitative methodology.

Although self-exclusion offered a respite from gambling-related stressors, several suggestions for improvements of the system emerged from the interviews. Participants emphasized the importance of holistic support systems and regulatory measures to augment its efficacy as a harm reduction tool. Suggestions for improvement partly followed what could be expected from previous web survey studies on gamblers; as many individuals who self-exclude also struggle with the risk of relapsing on overseas and nonlicensed gambling sites, users desired further legislative efforts making it possible to also self-exclude from other services, including those that are registered outside the present jurisdiction. Gambling in unlicensed, overseas online casinos has been reported to be the most common means of breaching one's self-exclusion period, as reported in the online surveys previously conducted here in this setting [13,14]. Thus, one major implication of this is the political process of how to exclude unlicensed gambling operators from the gambling market or to prevent financial transactions to gambling companies operating abroad. The legal framework and technical boundaries of this go beyond the scope of this study, but these findings highlight the need for gambling regulations to stretch beyond the controlling of the land-based gambling opportunity within each geographical jurisdiction.

One important and relatively novel suggestion was the linking between a mere harm-reducing self-exclusion service and actual therapeutic efforts against the addictive disorder. Importantly, this would move the purpose of the self-exclusion service from being an anonymous electronic support tool to actually providing a conduit to treatment. Treatment seeking in gambling disorder is known to be low, with several barriers reported [18]. Importantly, it has been reported that despite perceived barriers against treatment seeking, individuals with gambling problems may use more informal ways to seek help and not necessarily formal enrolment into psychotherapeutic treatment. For example, this may involve access to online advice or other online tools to facilitate a behavioral change [10]. In that aspect, any degree of referral from self-exclusion to different degrees of online support, advice, or treatment may be a way forward to improve the outcome of a person with gambling problems, over and above the effect of self-exclusion per se.

Such a system of transferring self-excluding clients to addiction treatment has not, to the best of our knowledge, been described in previous literature, and this may suggest a new area of research. For example, in such a system, it would have to be considered whether there is a potential barrier against enrolling in a self-exclusion service if it may, under some conditions, trigger a personal contact from a gambling operator, from a public authority, or from a treatment provider. However, it has been demonstrated that unsolicited motivational telephone calls to people displaying hazardous gambling habits may be a way to lower wagering and increase harm-reducing measures [30,31] and that acceptability of such motivational efforts appears to be high [32]. Thus, if self-exclusion from gambling would trigger an outreach effort aiming to improve the individual's gambling, it could be argued that this would not require a formal therapist-guided treatment but possibly may involve a brief, personalized intervention but with the capacity to refer individuals to further treatment when needed. Such individualized, personalized normative feedback has been suggested to be effective in other treatment or harm reduction interventions in problem gambling [33,34] and can initiate or enhance an internal motivational process in affected individuals, over and above the effects from the sole self-exclusion.

Likewise, another expansion of the self-exclusion tool, as suggested by this study, would be to slow down the procedure of opening up gambling opportunities after the self-exclusion period is terminated. This would also be a novel strategy that would require further study, but potentially, given the loss of the control component in gambling disorder, a more gradual onset of gambling after the self-exclusion period may provide further possibility to consider one's choice to gamble and may reduce the harm potentially associated with a sudden onset of high-risk gambling. Likely, any tool that slows down access to online gambling may potentially have a harm-reducing and relapse-preventing effect [35].

Overall, including the considerations made by interviewees in this study and their suggestions for improvement, a self-exclusion service generally appears to be promising and possible to increase even further. However, the major obstacle to this method, the risk of breaching one's self-exclusion through the use of gambling sites outside of the system, may remain. Here, one also has to consider the fact that the choice to self-exclude may cause the gambling patterns upon a potential relapse to happen on an unlicensed market and thereby with a lower degree of control than inside a licensed market. Previous experience, however, has pointed out that in individuals with extremely intense gambling patterns, such gambling patterns may happen on many operators, and a limit set upon one's gambling in one setting may be followed by gambling on another site. Still, even when this happens, such as during a deposit limit imposed during the COVID-19 pandemic in Sweden, still many clients report such a limit to be favorable and to have decreased their gambling pattern, rather than the

XSL•FO
RenderX

opposite [36]. Thus, although a method such as self-exclusion cannot hinder all types of gambling, it may still have motivational effects that limit the money lost and limits some of the consequences even in a high-risk gambling behaviour.

Strengths and Limitations

This study has a number of potential limitations, but given the novelty of the research topic and the merits of an in-depth qualitative study design, it also has considerable strengths. Recruitment via social media advertisements and digital interviews can be considered as both a strength and a limitation. The recruitment method enabled recruitment without any geographical limitations, therefore enabling a wide sample of individuals with different backgrounds, gambling habits, and experiences. However, it is possible that advertisements in social media attract interviewees with slightly different views on and experiences of self-exclusion from gambling than help-seeking individuals or individuals currently attending some type of land-based or online-based gambling venue. In contrast, the capacity of a qualitative design to identify a range of different aspects of a behaviour may benefit from this type of broad recruitment.

Digital interviews may represent another limitation, as these cannot readily be compared with physical interviews. Small changes in tone and body language are easily overlooked when the interview is held online. Meanwhile, the use of digital interviews in this study substantially facilitated data collection, widened the possibility to participate from the whole country rather than only locally, and may have facilitated recruitment of affected individuals who may not readily disclose their personal suffering and show up in person in a mental health–and addicion-oriented research unit. Indeed, challenges, but also potential advantages, of online qualitative research interview have been discussed in the research literature in recent years [37], pointing to the fact that such online interviews can play an important role in qualitative research.

Another potential limitation is the time aspect; some of the participants had recent or even ongoing experiences of Spelpaus, while others had been using it further back in time. Although this may be seen as a limitation, their information about their Spelpaus-related experience still is of great value (ie, as their ongoing or previous experience).

Additionally, the absence of a diagnosis of gambling gisorder as an inclusion criterion means that assessments of participants' gambling problems relied solely on self-reported statements and estimations, which may not capture the full spectrum of gambling-related issues. This limitation underscores the possibility that experiences with Spelpaus may vary based on individuals' level of gambling problems.

However, this study addressed a novel research issue and has strengths in a number of aspects. This is, to the best of our knowledge, the first qualitative study centered around the users' experiences of the relatively novel and unique nationwide self-exclusion service. Consequently, it offers updated and in-depth insights into gamblers' experiences with Spelpaus, elucidating its advantages and challenges as a harm reduction tool.

Conclusion

This qualitative study aimed to deepen the knowledge about and understanding of a novel nationwide service for self-exclusion from gambling. This includes the experience of individuals who have a history of being self-excluded, including their reasons for self-exclusion and their experience of the system's merits and challenges. Based on the interviews, this study suggests that although breaching one's self-exclusion remains a challenge, users of this service also report major advantages from being self-excluded. Users suggest further improvements to the self-exclusion model, such as international collaboration in order to prevent overseas gambling operators offering gambling outside of the self-exclusion system and also structured links from self-exclusion to further help and support, possibly also involving formal treatment. The study provided new information about self-exclusion and about a broader ranges of reasons for self-excluding, over and above the reporting of severe gambling problems (eg, the role of self-exclusion in the relationship with concerned significant others or to prove willingness to quit gambling). Several favorable effects from self-excluding were reported, in addition to a need for clearer obstacles against going back to gambling once self-exclusion ends.

Conflicts of Interest

AH has research funding for other projects from AB Svenska Spel, which is the state-owned gambling operator of Sweden, and from Svenska Spel's independent research council, as well as from the research council of Systembolaget, the Swedish alcohol monopoly. None of these organizations had any role in, or influence on, this project. The project itself was carried out thanks to funding from the research council of Svenska Spel. The funder did not have any influence on the procedures of this research or on the interpretation and reporting of results.

References

1. Langham E, Thorne H, Browne M, Donaldson P, Rose J, Rockloff M. Understanding gambling related harm: a proposed definition, conceptual framework, and taxonomy of harms. BMC Public Health 2016 Jan 27;16(1):80 [FREE Full text] [doi: 10.1186/s12889-016-2747-0] [Medline: 26818137]
2. Potenza MN, Balodis IM, Derevensky J, Grant JE, Petry NM, Verdejo-Garcia A, et al. Gambling disorder. Nat Rev Dis Primers 2019 Jul 25;5(1):51. [doi: 10.1038/s41572-019-0099-7] [Medline: 31346179]

XSL•FO
RenderX

3. Wardle H, Reith G, Langham E, Rogers RD. Gambling and public health: we need policy action to prevent harm. BMJ 2019 May 08;365:l1807 [FREE Full text] [doi: 10.1136/bmj.l1807] [Medline: 31068335]
4. Matheson FI, Hamilton-Wright S, Kryszajtys DT, Wiese JL, Cadel L, Ziegler C, et al. The use of self-management strategies for problem gambling: a scoping review. BMC Public Health 2019 Apr 29;19(1):445 [FREE Full text] [doi: 10.1186/s12889-019-6755-8] [Medline: 31035978]
5. Kraus L, Loy JK, Bickl AM, Schwarzkopf L, Volberg RA, Rolando S, et al. Self-exclusion from gambling: a toothless tiger? Front Psychiatry 2022 Sep 23;13:992309 [FREE Full text] [doi: 10.3389/fpsyt.2022.992309] [Medline: 36213894]
6. Ministry of Finance. Gambling Act (2018:1138). Swedish Gambling Authority. URL: https://www.spelinspektionen.se/globalassets/dokument/engelsk/oversatt-spellagen/english-spellagen-sfs-2018_1138-uppdat-sfs-2024_255.pdf [accessed 2025-01-21]
7. Abbott M, Romild U, Volberg R. The prevalence, incidence, and gender and age‐specific incidence of problem gambling: results of the Swedish longitudinal gambling study (Swelogs). Addiction 2017 Nov 24;113(4):699-707. [doi: 10.1111/add.14083]
8. Gabellini E, Lucchini F, Gattoni ME. Prevalence of problem gambling: a meta-analysis of recent empirical research (2016-2022). J Gambl Stud 2022 Dec 31;39(3):1027-1057. [doi: 10.1007/s10899-022-10180-0]
9. Håkansson A, Karlsson A, Widinghoff C. Treatment seeking for gambling disorder in nationwide register data – observations around a major shift in legislation. Front Public Health 2024 Mar 19;12:1293887 [FREE Full text] [doi: 10.3389/fpubh.2024.1293887] [Medline: 38566789]
10. Rodda S, Dowling N, Lubman D. Gamblers seeking online help are active help-seekers: time to support autonomy and competence. Addict Behav 2018 Dec;87:272-275. [doi: 10.1016/j.addbeh.2018.06.001] [Medline: 29935737]
11. Exclude yourself from gambling!. Spelpaus. 2024. URL: https://www.spelpaus.se/en#:~:text=This%20is%20where%20you%20can,real%2Dmoney%20gambling%20in%20Sweden.&text=A%20self%2Dexclusion%20via%20Spelpaus,before%20the%20date%20you%20choose [accessed 2025-01-21]
12. Håkansson A, Widinghoff C, Berge J. Self-exclusion from gambling – a measure of COVID-19 impact on gambling in a highly online-based gambling market? Int J Environ Res Public Health 2021 Jul 09;18(14):7367 [FREE Full text] [doi: 10.3390/ijerph18147367] [Medline: 34299817]
13. Håkansson A, Widinghoff C. Gambling despite nationwide self-exclusion - survey in online gamblers in Sweden. Front Psychiatry 2020 Dec 2;11:599967 [FREE Full text] [doi: 10.3389/fpsyt.2020.599967] [Medline: 33343428]
14. Håkansson A, Komzia N. Self-exclusion and breaching of self-exclusion from gambling: a repeated survey study on the development of a nationwide self-exclusion service. Harm Reduct J 2023 Aug 08;20(1):107 [FREE Full text] [doi: 10.1186/s12954-023-00822-w] [Medline: 37553657]
15. Spelvanor online 2023 [Online gambling habits 2023]. Swedish Gambling Authority. URL: https://www.spelinspektionen.se/globalassets/dokument/statistik/enkatundersokning/spelvanor-online-rapport-juli-2023.pdf [accessed 2025-01-21]
16. Miles M, Rothschild J, Åkesson G, Håkansson A. Nationwide, multioperator self-exclusion and psychiatric comorbidity in patients with gambling disorder: a retrospective chart review study from a regional treatment unit. J Addict 2023;2023:5532259 [FREE Full text] [doi: 10.1155/2023/5532259] [Medline: 37808466]
17. LaPlante DA, Nelson SE, LaBrie RA, Shaffer HJ. Stability and progression of disordered gambling: lessons from longitudinal studies. Can J Psychiatry 2008 Jan 01;53(1):52-60. [doi: 10.1177/070674370805300108] [Medline: 18286872]
18. Suurvali H, Cordingley J, Hodgins DC, Cunningham J. Barriers to seeking help for gambling problems: a review of the empirical literature. J Gambl Stud 2009 Sep 24;25(3):407-424. [doi: 10.1007/s10899-009-9129-9] [Medline: 19551495]
19. Håkansson A, Tjernberg J, Hansson H. Effects and limitations of a unique, nationwide, self-exclusion service for gambling disorder and its self-perceived effects and harms in gamblers: protocol for a qualitative study. JMIR Res Protoc 2023 Nov 14;12:e47528 [FREE Full text] [doi: 10.2196/47528] [Medline: 37962917]
20. Public Health Agency of Sweden. Swelogs (The Swedish Longitudinal Gambling Study). Public Health Agency of Sweden: Sweden; 2023.
21. Binde P. Gambling in Sweden: the cultural and socio-political context. Addiction 2014 Feb;109(2):193-198. [doi: 10.1111/add.12103] [Medline: 23297702]
22. Håkansson A, Widinghoff C. Television gambling advertisements: extent and content of gambling advertisements with a focus on potential high-risk commercial messages. Addict Behav Rep 2019 Jun;9:100182 [FREE Full text] [doi: 10.1016/j.abrep.2019.100182] [Medline: 31193826]
23. Graneheim U, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today 2004 Feb;24(2):105-112. [doi: 10.1016/j.nedt.2003.10.001] [Medline: 14769454]
24. Jacobsen D. Förståelse, beskrivning och förklaring - Introduktion till samhällsvetenskaplig metod för hälsovård och socialt arbete. Lund: Studentlitteratur; 2012.
25. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000 Aug;23(4):334-340. [doi: 10.1002/1098-240x(200008)23:4<334::aid-nur9>3.0.co;2-g] [Medline: 10940958]
26. Binde P, Cisneros Örnberg J, Forsström D. Criminogenic problem gambling: a study of verdicts by Swedish courts. Int Gambl Stud 2021 Dec 08;22(3):344-364. [doi: 10.1080/14459795.2021.2002383]

XSL•FO
RenderX

27. Grant JE, Kim SW, Kuskowski M. Retrospective review of treatment retention in pathological gambling. Compr Psychiatry 2004 Mar;45(2):83-87. [doi: 10.1016/j.comppsych.2003.12.005] [Medline: 14999657]
28. Ingle PJ, Marotta J, McMillan G, Wisdom JP. Significant others and gambling treatment outcomes. J Gambl Stud 2008 Sep 5;24(3):381-392. [doi: 10.1007/s10899-008-9092-x] [Medline: 18320296]
29. Tremblay J, Dufour M, Bertrand K, Saint-Jacques M, Ferland F, Blanchette-Martin N, et al. Efficacy of a randomized controlled trial of integrative couple treatment for pathological gambling (ICT-PG): 10-month follow-up. J Consult Clin Psychol 2023 Apr;91(4):221-233. [doi: 10.1037/ccp0000765] [Medline: 36521131]
30. Jonsson J, Hodgins DC, Munck I, Carlbring P. Reaching out to big losers leads to sustained reductions in gambling over 1 year: a randomized controlled trial of brief motivational contact. Addiction 2020 Aug 13;115(8):1522-1531. [doi: 10.1111/add.14982] [Medline: 31977104]
31. Hakansson A, Franklin K, Dahlström M, Lyckberg A. Motivational telephone intervention to risk gamblers by a state-owned gambling operator in Sweden. Front Psychiatry 2024 Jan 30;15:1343733 [FREE Full text] [doi: 10.3389/fpsyt.2024.1343733] [Medline: 38352656]
32. Håkansson A, Franklin K, Dahlström M, Lyckberg A. Responsible gambling through a motivational telephone intervention to high-risk gamblers - an evaluation of user satisfaction and subjective intervention effects. Front Psychol 2022 Nov 30;13:917216 [FREE Full text] [doi: 10.3389/fpsyg.2022.917216] [Medline: 36532977]
33. Neighbors C, Rodriguez LM, Rinker DV, Gonzales RG, Agana M, Tackett JL, et al. Efficacy of personalized normative feedback as a brief intervention for college student gambling: a randomized controlled trial. J Consult Clin Psychol 2015 Jun;83(3):500-511 [FREE Full text] [doi: 10.1037/a0039125] [Medline: 26009785]
34. Marchica L, Derevensky JL. Examining personalized feedback interventions for gambling disorders: a systematic review. J Behav Addict 2016 Mar;5(1):1-10 [FREE Full text] [doi: 10.1556/2006.5.2016.006] [Medline: 28092190]
35. Newall PWS. Reduce the speed and ease of online gambling in order to prevent harm. Addiction 2023 Feb 24;118(2):204-205. [doi: 10.1111/add.16028] [Medline: 36002979]
36. Håkansson A, Sundvall A, Lyckberg A. Effects of a national preventive intervention against potential COVID-19-related gambling problems. A self-report survey in online gamblers. JMIR Form Res 2022 Mar 09;6(3):e33066 [FREE Full text] [doi: 10.2196/33066] [Medline: 34678751]
37. Keen S, Lomeli-Rodriguez M, Joffe H. From challenge to opportunity: virtual qualitative research during COVID-19 and beyond. Int J Qual Methods 2022 Jun 04;21:16094069221105075. [doi: 10.1177/16094069221105075] [Medline: 35692956]

Edited by A Kushniruk; submitted 02.09.24; peer-reviewed by M Bellringer, R Lantto; comments to author 14.10.24; revised version received 23.12.24; accepted 08.01.25; published 31.01.25.

Please cite as:
Tjernberg J, Helgesson S, Håkansson A, Hansson H
Exploring the Users' Perspective of the Nationwide Self-Exclusion Service for Gambling Disorder, "Spelpaus": Qualitative Interview Study
JMIR Hum Factors 2025;12:e66045
URL: https://humanfactors.jmir.org/2025/1/e66045
doi:10.2196/66045
PMID:39888656

©Johanna Tjernberg, Sara Helgesson, Anders Håkansson, Helena Hansson. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 31.01.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Barriers and Facilitators to User Engagement and Moderation for Web-Based Peer Support Among Young People: Qualitative Study Using the Behavior Change Wheel Framework

Ananya Ananya[1], MSc; Janina Tuuli[1], MSc; Rachel Perowne[1], MSc; Leslie Morrison Gutman[1], PhD

Department of Clinical, Educational and Health Psychology, Faculty of Brain Sciences, University College London, London, United Kingdom

Corresponding Author:
Leslie Morrison Gutman, PhD
Department of Clinical, Educational and Health Psychology
Faculty of Brain Sciences
University College London
1-19 Torrington Place
London, WC1E 7HB
United Kingdom
Phone: 44 8059 0939
Email: l.gutman@ucl.ac.uk

Abstract

Background: Peer support groups or web-based chats for young people offer anonymous peer support in judgment-free spaces, where users may share their thoughts and feelings with others who may have experienced similar situations. User engagement is crucial for effective web-based peer support; however, levels of engagement vary. While moderation of peer support groups can have a positive impact on the engagement of young people, effective moderation can be challenging to implement.

Objective: This study aimed to identify barriers and facilitators to user engagement with, and moderation of, web-based peer support groups among young people aged 16 to 25 years and to provide recommendations for enhancing this service.

Methods: Drawing upon the Theoretical Domains Framework (TDF) and the Behavior Change Wheel (BCW), this study conducted qualitative interviews and gathered open-ended questionnaires from service users and moderators of The Mix, the United Kingdom's leading web-based mental health platform providing peer support groups for young people. Semistructured interviews were conducted with 2 service users and 8 moderators, and open-ended questionnaires were completed by 7 service users. Themes were coded using the Capability, Opportunity, Motivation, and Behavior (COM-B) model and the TDF. The BCW tools were then used to identify relevant behavior change techniques to improve user engagement in, and moderation of, the service.

Results: Thematic analysis revealed a total of 20 inductive themes within 10 TDF domains—9 (45%) for engagement and 11 (55%) for moderation. Of these 20 themes, 3 (15%) were facilitators of engagement, 7 (35%) were facilitators of moderation, 4 (20%) were barriers to moderation, and 6 (30%) barriers to engagement. Results suggest that skills, knowledge, beliefs about consequences, intentions, emotions, and the social and physical environment are important factors influencing service users and moderators of group chats. In particular, supporting the improvement of memory, attention, and decision-making skills of those involved; adapting the physical environment to facilitate effective interactions; and reducing negative emotions are suggested to optimize the value and effectiveness of peer support groups for young people's mental health for both the service users and moderators of these services.

Conclusions: The study demonstrates the effectiveness of the BCW approach and the use of the TDF and COM-B model to understand the influences on behavior in a systematic manner, especially for mental health and well-being interventions. The findings can be applied to design structured interventions to change behaviors related to the engagement with, and moderation of, web-based peer support groups and, in turn, improve mental health outcomes for young people.

(JMIR Hum Factors 2025;12:e64097) doi:10.2196/64097

XSL•FO
RenderX

KEYWORDS

internet; moderation; engagement; youth; teenager; adolescent; peer support; web-based group; user engagement; support group; barrier; facilitator; Theoretical Domains Framework; Behavior Change Wheel; qualitative; interview; behavior change technique; thematic analysis

Introduction

Background

Young people aged 16 to 25 years are particularly vulnerable to mental health difficulties [1]. Recent evidence shows an increase in reported mental health problems for young people aged 16 to 25 years, particularly following the COVID-19 pandemic [2,3]. For example, in young people aged 17 to 19 years in England, rates of probable mental disorders rose from 17.7% in 2020 [4] to 23.3% in 2023 [5]. Nevertheless, health care systems and support often overlook the needs of young people [6]. In the United Kingdom, 75% of young people experience delays in accessing mental health support, leading to a worsening of their condition [7].

Web-based mental health communities are a viable option to bridge the mental health service gap for young people. They offer anonymous peer support in judgment-free spaces, where users may share their thoughts and feelings with others who may have experienced similar situations [8]. Web-based peer support can be asynchronous or synchronous, providing online support in a group format, with or without moderation [9,10]. Synchronous or real-time support, such as web-based chats, provides users with in-the-moment assistance without the delays that can occur within asynchronous services [11,12]. Notably, studies have shown the efficacy of web-based peer support platforms compared to in-person talk therapies [13]. They offer clinical effectiveness [14] and scalability for public health impact, enabling outreach to a larger population [13]. Therefore, there is a strong argument for the use of web-based peer support platforms for young people as they address key barriers and hold the potential to enhance mental health treatment by providing a safe, accessible, and effective means of support.

Despite numerous benefits, studies have also highlighted challenges to web-based peer mental health support that limit its effectiveness [15], including varying engagement rates [14]. Low rates of engagement can lead to negative outcomes, such as perceived exclusion and isolation [16], whereas high rates have been found to improve young people's mental health [17]. Effective moderation of web-based, user-led mental health services has been found to improve user engagement [18]. However, there are few studies examining user engagement in, and moderation of, synchronous web-based peer support groups, especially capturing the perspectives of both service users and moderators [11,19]. Service users, with their lived experience of giving and receiving support, can offer valuable insights and suggestions for improving these platforms. Moderators, due to their proximity to users and responsibility for supporting positive interactions, safety, and engagement [14], may have practical insights to enhance platform effectiveness [20]. A systematic examination of these behaviors from different viewpoints would enable the identification of tailored intervention strategies to improve moderation of web-based peer support services and enhance user engagement. Using the Behavior Change Wheel (BCW) [21], this qualitative research study investigates the barriers and facilitators to user engagement and moderation for web-based peer support groups among young people aged between 16 and 25 years and then proposes tailored strategies for optimization.

User Engagement in Web-Based Peer Support

Digital health research has defined user engagement as the extent and subjective experience (characterized by attention, interest, and affect) of use [22]. Studies measure engagement in terms of use time, log-ins, or module completion [23,24], with approximately 60% of studies measuring attendance alone to assess engagement [25]. However, although passive involvement or observational participation in web-based peer support groups has recognized benefits [26], considering only attendance falls short as it omits active participation and interaction, which fosters the sense of community and support unique to this service. Therefore, this study considers engagement in the broader sense, incorporating any observable written contribution, such as comments expressing thoughts and feelings or support for others. This inclusive approach captures subjective experience, aids barrier and facilitator identification, and can inform interventions for enhanced meaningful engagement.

Previous research highlights key barriers and facilitators to young people's engagement with digital mental health interventions and services [27,28]. Barriers include interventions that are perceived as unappealing or unhelpful, technical issues, privacy concerns, and young people lacking time or remembering to use the intervention. Facilitators include the personalization and flexibility offered by digital interventions, along with effective design, usability, opportunities to build connections, and the potential for a rewarding user experience [27,28].

However, there is a gap in research examining influences on young people's engagement in synchronous web-based peer support groups for mental health. Barriers and facilitators are likely to differ from those in other digital mental health services, as peer-to-peer support is viewed as less stigmatizing and more relatable [29]. In addition, synchronous interactions have a different pace compared to other formats [11]. Recent research focusing on peer-to-peer chats has centered on moderators' roles without including service user voice [30].

Moderation of Web-Based Peer Support

Moderation involves managing content, safeguarding, and providing support to users when needed [18]. It can be carried out by health professionals or trained volunteers [19]. Effective moderation has been found to have a positive effect on user engagement and has also been helpful in preventing "toxic" web-based settings [18]. Negative effects of web-based platforms such as trolling and stalking may also be prevented by effective moderation [16]. In web-based chat groups and

XSL•FO
RenderX

discussion forums, moderators ensure adherence to group guidelines to enhance user safety and prevent service users from becoming overly dependent on one another [31,32].

There has been limited research on the moderation of web-based mental health interventions [19]. Synchronous group settings pose distinct challenges as they require real-time monitoring of the chat and interaction between peers and moderators. In a recent study, Deng et al [30] explored moderation of peer support for a web-based mental health community. In this study, although moderators were qualified mental health professionals, they experienced challenges, such as dealing with emotionally triggering problems or hostile users. In a similar environment, Saha et al [33] highlighted challenges, such as a lack of concern for moderators' well-being and moderators' uncertainty about when to intervene. However, the services examined in these studies were not specifically designed for young people. Thus, there is a need to understand moderation and, importantly, how to enhance it, particularly in the context of real-time group chats for young people moderated by young volunteers, in a way that is structured and rooted in theory.

To address these research gaps, this study triangulates moderator and service user perspectives using a theoretically based behavior change approach, providing a comprehensive examination of user engagement and the moderation process in synchronous web-based mental health peer chats. The BCW provides a systematic and comprehensive framework to enhance understanding of behavior and support the identification of linked, relevant behavior change techniques (BCTs) to address the identified barriers and optimize group chat engagement [21].

BCW Framework

The BCW comprises interconnected tools to guide decision-making and facilitate the systematic development and evaluation of behavioral interventions [21]. At its center lies the capability, opportunity, motivation, and behavior (COM-B) model of behavior, which supports the identification of the barriers and facilitators to a behavior (Figure 1 [21]). Capability refers to an individual's physical as well as psychological capacity to perform a specific activity. Motivation is the cognitive processes that drive human behavior. Finally, opportunity refers to external factors that are outside the individual and may facilitate behavior.

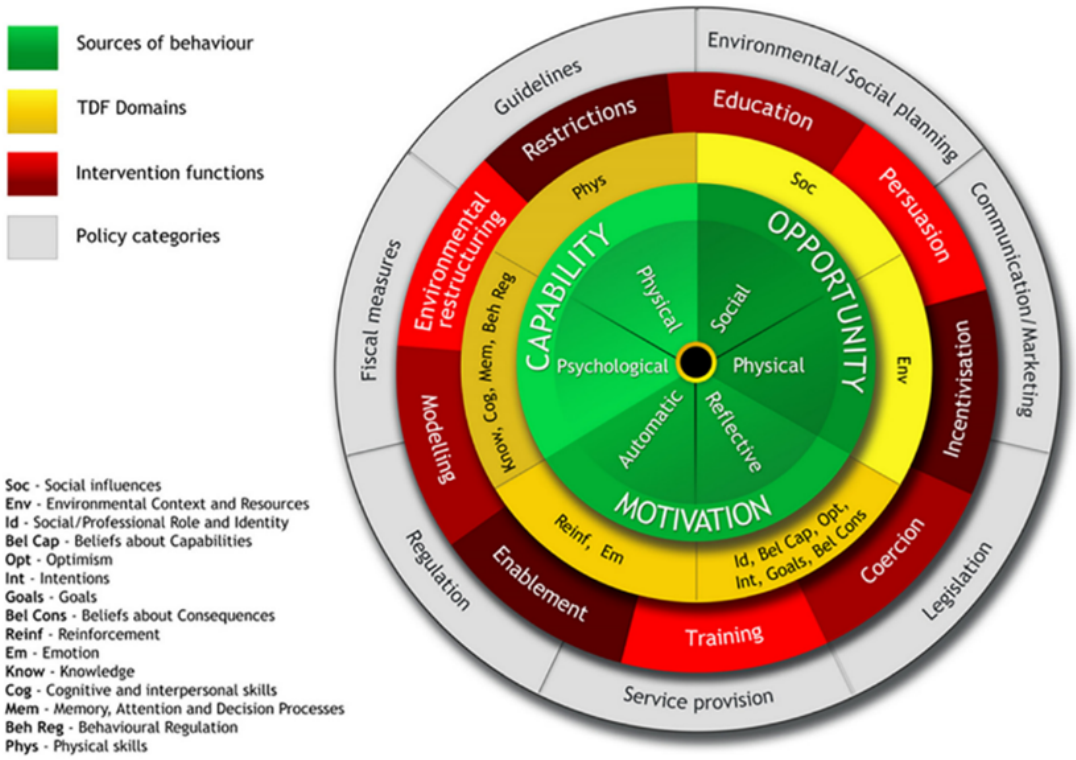
The components of COM-B can be further elaborated into 14 domains using the Theoretical Domains Framework (TDF), providing a more granular understanding of the influences on behavior [34,35]. In the next layer of the BCW are intervention functions, which are the broad categories of means by which an intervention can change behavior. The outer layer consists of policy categories to support long-term, system-wide implementation.

The Behavior Change Techniques Taxonomy (BCTT) comprises 93 observable and replicable BCTs. The BCTT is a reliable method for specifying, interpreting, and implementing BCTs, which are considered "active ingredients" to facilitate behavior change either alone or in combination [36]. Expert consensus allows for mapping of the COM-B and TDF domains to intervention functions and corresponding BCTs, guiding appropriate strategies for change [37].

The BCW approach has been successfully applied in previous research to understand the barriers and facilitators to the use and delivery of digital mental health platforms for young people, including webchat counseling [28,38] and moderation of self-harm content in web-based discussion boards [39].

XSL•FO
RenderX

Figure 1. The Behavior Change Wheel and Theoretical Domains Framework (TDF) (adapted from Michie et al [21]).



This Study

Given the lack of research in this area, this qualitative study systematically explores the barriers and facilitators to the user engagement and moderation of synchronous web-based peer group chats for young people and proposes evidence-based strategies for improvement. Data were gathered from open-ended questionnaires and in-depth, semistructured interviews with young service users and moderators of The Mix, a web-based mental health support service for young people aged <25 years. Using the BCW [21], the following research questions were addressed:

1. Using the COM-B and TDF, what are the barriers and facilitators to young people's engagement in web-based peer support groups?
2. Using the COM-B and TDF, what are the barriers and facilitators to moderation of web-based peer support groups?
3. Using the BCTT, what strategies can be used to improve user engagement in, and moderation of, web-based peer support groups for young people's mental health?

Methods

Participants

This study collected data from service users and moderators of The Mix, the United Kingdom's leading digital mental health support service for young people aged <25 years. The Mix offers free mental health support, including a helpline, phone and webchat counseling, crisis support, discussion boards, and moderated live group chats. These synchronous group chats are freely available for young people aged 13 to 25 years. Table 1 summarizes the peer-to-peer mental health group chats of The Mix analyzed in this study.

The Mix circulated a poster to moderators (total population of 52) and service users to promote participation. Interested participants contacted the researcher through email, and their involvement was entirely voluntary. As shown in Table 2, a total of 2 service users were interviewed, and a further 7 service users completed an open-ended questionnaire about their engagement. A total of 8 moderators (n=5, 63% staff and n=3, 38% volunteers) participated in interviews about moderation. Of these 8 moderators, 4 (50%) staff members and 1 (13%) volunteer also completed a separate interview about user engagement. As this was a qualitative study, it aimed to gather rich and detailed data from participants, similar to previous studies in a similar context [38,39], rather than to achieve a representative sample.

All moderators were women, while most of the service users were either women or did not self-identify their gender, with the exception of 1 man and 1 nonbinary service user. Moderators were aged between 19 and 30 years, and service users were aged between 18 and 24 years (Table 2). Moderators were involved with The Mix ranging from 6 months to as long as 8 years and worked for 1 to 2 hours per week.

Semistructured interviews, based on the TDF, were conducted via the Microsoft Teams platform and transcribed for analysis. The interviews lasted between 30 to 60 minutes each and were designed to gain in-depth service user and moderator insights into experiences of group chats regarding peer interactions, structure of the chats, and feelings about contributing. Questions included, for example, "How long do you spend on The Mix's chat per week?" and "How do other people in the group chat

XSL•FO
RenderX

make it easier or more difficult for you to share your thoughts, feelings and experiences or support others with theirs?" Questions for moderators only included the following: "How do other people influence how you contribute to the moderation of posts on the group chats?"

For user engagement only, participants were offered the alternative option to complete an open-ended Microsoft Forms questionnaire to combat any difficulties young people felt in expressing themselves in an interview and may have hindered participation [28]. The questionnaire included 5 open-ended questions, exploring what made it easier or harder for users to share their experiences, what influenced their decision to participate or not, and what changes they would make to the chat if given the opportunity.

Table 1. The Mix's peer-to-peer mental health chat groups.

Group chat type	Description	Participants, n	Structure	Participation	Privacy	Moderation
Support chat	Peer support for mental health struggles	Variable	Open discussions	Active participation	No sign-up required, no password protection	Moderated by trained young volunteers and staff
Support circle	More intimate setting, where everyone takes turns to obtain support while the rest of the group listens	≤5	Rotating support and sharing	Sign up to receive support or as a "listener"	Password protected, full session attendance	Moderated by trained young volunteers and staff

Table 2. Participant information.

Participant	Data	Age (y)	Gender
SU1[a]	Interview	24	Man
SU2	Interview	18	Nonbinary
SU3	Questionnaire	18-25	Did not self-identify
SU4	Questionnaire	18-25	Did not self-identify
SU5	Questionnaire	18-25	Woman
SU6	Questionnaire	18-25	Woman
SU7	Questionnaire	18-25	Did not self-identify
SU8	Questionnaire	18-25	Did not self-identify
SU9	Questionnaire	18-25	Did not self-identify
MS1[b]	Interview (engagement and moderation)	28	Woman
MS2	Interview (engagement and moderation)	30	Woman
MS3	Interview (engagement and moderation)	24	Woman
MV4[c]	Interview (moderation)	26	Woman
MV5	Interview (engagement and moderation)	24	Woman
MV6	Interview (moderation)	19	Woman
MS7	Interview (moderation)	23	Woman
MS8	Interview (moderation)	22	Woman

[a]SU: service user.

[c]MS: moderator staff.

[d]MV: moderator volunteer.

Data Analysis

Data were thematically analyzed using the 6-phase process developed by Braun and Clarke [40]. NVivo (Lumivero) and Excel (Microsoft) were used as analytical tools. The initial stage involved gaining familiarity with the transcribed interviews. A deductive coding approach was adopted initially, which involved reviewing items of data from the transcripts and questionnaires and organizing them into the TDF domains, providing a "start list" [41,42] of themes to be categorized as barriers or facilitators. The process was repeated using an inductive analysis to generate specific subthemes within each TDF theme. These subthemes were then reviewed, defined, and included in the codebook [43]. A codebook was developed to facilitate coding and was iteratively updated. As new data were added to TDF domains, facilitators and barriers were modified, expanded, or recategorized using the constant comparison method to identify patterns and variations within the dataset [44]. Triangulation

XSL•FO
RenderX

was completed with data from both service users and moderators to explore consistencies and contradictions in influences on young people's engagement with group chats. Reflexivity was prioritized to ensure openness and challenge in interpretations [45]. To ensure coding reliability and validity, 2 researchers independently cross-coded 1 transcript; a reliability check was conducted, and agreement was reached on any discrepancies.

Following data analyses, the coded TDF barriers were mapped to corresponding intervention types, using the BCW, and then specific BCTs were selected using the links specified in Cane et al [46] as well as the more recently developed Theory and Techniques Tool [47,48]. The selection of the relevant and appropriate intervention types and BCTs was informed by an appraisal of the affordability, practicability, effectiveness and cost-effectiveness, acceptability, side effects, and equity criteria [21]. This approach has been used in a multitude of studies [49,50] to evaluate intervention strategies. The selection of policy options was out of scope because the study was focused on a single organization rather than system level. Finally, the BCTs were operationalized, and intervention strategies were proposed based on a review of previous literature.

Ethical Considerations

Low-risk ethics approval was obtained from the University College London Ethics Committee (Z6364106/2023/03/149 social research, 25069/001). Participant information and a consent form were provided to interested parties to sign digitally and confirm consent. Participants were advised that participation was completely voluntary and they could withdraw at any time, up to 4 weeks after their data were collected. As reimbursement for the interviews, participants were given a £10 (US $12.45) voucher. Questionnaire respondents could opt into a lottery prize draw to win 1 of 2 £10 (US $12.45) vouchers by providing their email address. To protect participant privacy and confidentiality, data were fully anonymized.

Results

Overview

Table 3 sets out the themes identified as both barriers and facilitators to user engagement and moderation. These are set out according to both COM-B and TDF frameworks.

XSL•FO
RenderX

Table 3. The COM-B[a] and TDF[b] themes and subthemes for user engagement and moderation.

COM-B and TDF themes	User engagement subthemes	Moderation subthemes
Psychological capability		
Cognitive and interpersonal skills	• Difficulty expressing feelings and needs (barrier) • Skills to listen and validate (facilitator)	• Developing skills (facilitator)
Knowledge	—[c]	• Understanding the guidelines (facilitator)
Memory, attention, and decision-making	• Memory recall of guidelines (barrier)	• Responding quickly in complex situations (barrier)
Behavioral regulation	—	• Holding back on advice (barrier)
Reflective motivation		
Beliefs about consequences	• Judgment and confidentiality concerns (barrier)	• Impact of moderation (facilitator)
Social and professional role and identity	—	• Congruence between social and professional identity (facilitator)
Intentions	• Wanting to support others (facilitator)	• Intending to shape young people's perceptions (facilitator)
Automatic motivation		
Emotions	• Fear and anxiety (barrier)	• Distressing subject matter (barrier)
Physical opportunity		
Environment, context, and resources	• Overload and lack of structure (barrier)	• Organizational resources (facilitator) • Lack of visual cues (barrier)
Social opportunity		
Social influences	• Similarity and familiarity (facilitator) • Integration of new users and gaps in support (barrier)	• Support offered (facilitator)

[a]COM-B: Capability, Opportunity, Motivation, and Behavior.

[b]TDF: Theoretical Domains Framework.

[c]Not applicable.

Young People's Engagement

As shown in Table 3, a total of 9 barriers and facilitators were identified as subthemes across 7 domains of the TDF. Themes were mainly consistent across support chat and support circle responses, but a small number of differences arose between the 2, and these are indicated, where relevant, in the following subsections.

Cognitive and Interpersonal Skills (Psychological Capability)

Difficulty Expressing Feelings and Needs

Several service users and moderators described the difficulty users face in expressing their emotions and thoughts, which can prevent them from opening up in the chat. Subthemes are described below, together with illustrative quotes. Service user quotes are indicated by "SUX," volunteer moderator quotes are indicated by "MVX," and staff moderator quotes are indicated by "MSX," where "X" is the participant number, to identify each participant as per Table 2. One user commented the following:

> *[What makes it more difficult to open up is] not being able to express myself about feelings.* [SU8]

Moderators also noted that vague statements about one's emotional state without specific requests for support may be a barrier to others' engagement, as they provide insufficient guidance for other users to offer relevant support. For example, one user stated the following:

> *I think it depends on how someone phrases the question...if they're sort of saying, "And I'm feeling really low today" or "And I can't do this anymore," like some of those trickier comments.* [MV5]

Skills to Listen and Validate

Both moderators and users identified that having skills to effectively validate and listen to others facilitated peer support and engagement. One moderator noted the following:

XSL•FO
RenderX

> *The people that tend to offer the most peer support are those that...know how to validate someone's experiences.* [MS1]

Users described how they did this, with one user stating the following:

> *Everyone listened and waited their turn...I listened to others, and they encourage people waiting their turn to help others.* [SU6]

Personal experiences that related to others' struggles made it easier to connect and offer support, but 1 user stated the following:

> *There are some situations where I can't exactly relate...so I stay quiet.* [SU5]

Memory, Attention, and Decision Processes (Psychological Capability): Memory Recall of Guidelines

Difficulties in absorbing, retaining, and applying guidelines posed a significant barrier to a couple of users. Users indicated that they struggled to remember guidelines due to content volume, distractions, and cognitive limitations, impacting their ability to engage. Even long-term users found it challenging, with 1 individual expressing the following:

> *I can't actually remember what [the guidelines are]. They're very basic from what I do remember.* [SU1]

Mental health crisis situations further exacerbated guideline oversight, leading some users to unintentionally deviate from the chat's intended mode of engagement in the moment. Moderators noted that this issue was worsened by some users, particularly regular ones, who found the guidelines unengaging and tended to ignore them.

Beliefs About Consequences (Reflective Motivation): Judgment and Confidentiality Concerns

This was a barrier revolving around young people's apprehensions about being judged, misunderstood, or facing negative reactions when sharing personal thoughts and experiences, particularly in the presence of new members. One participant stated the following:

> *Some people feel like a new person may judge more or less or may be offended more or be offended less. So [withdrawing] very much sort of safeguarding themselves, protecting themselves, and protecting others.* [SU1]

This could also lead users "to not give too much information" (SU7). Similarly, a moderator highlighted concerns about users fearing judgment, with 1 moderator stating the following:

> *...judging that person for talking about that cause I think that's probably what they worry about quite a lot.* [MV5]

Both moderators and users mentioned concerns about confidentiality, giving examples of where their privacy was violated by other members, leading to negative consequences. As a result, they became hesitant to share. One moderator stated the following:

> *Some members also worry confidentiality will be broken so they don't open up in fear of the police being called.* [MS2]

Intentions (Reflective Motivation): Wanting to Support Others

Overall, the intention to create a supportive environment was a facilitator, which motivated users to engage to support others. One user expressed the following:

> *Most of the time, I try my best to offer support and be there for the person too. I just want people to be able to open up to me, and when they do, I try to be a good friend.* [SU2]

However, moderators were concerned that some users may prioritize self-support, limiting their involvement in supporting others and welcoming newcomers. One moderator stated the following:

> *Often times, the young people are so busy sharing their own experiences that they don't make enough space to support others who may be going through a similar thing.* [MS1]

Emotions (Automatic Motivation): Fear and Anxiety

A frequently mentioned barrier by both users and moderators related to the user's fear of being perceived as a burden by sharing their thoughts, feelings, and experiences. One user expressed the following:

> *As for not opening up so much that's more when I feel like I'm a burden and not wanted, but that's me not The Mix.* [SU4]

Users identified that "Worrying what people think" (SU7) in terms of oversharing or repeating themselves caused them to hold back or stop sharing altogether. Another user shared similar sentiments, noting that opening up more "doesn't really make me feel better, it just makes me anxious and guilty" (SU2). These feelings of self-doubt and anxiety hindered users from fully opening up and receiving the support they needed.

Environmental Context and Resources (Physical Opportunity): Overload and Lack of Structure

A barrier identified by almost all participants was the lack of structure and organization within the support chat, especially during busy times. This could result in confusion, overwhelm, and a sense of disorganization. Users sometimes struggled to fully engage and provide or receive support effectively because, as 1 user said, "When you have a lot of people, you do get to a point where you do have like five, ten different conversations going on, all different levels of importance so it does get a bit confusing" (SU1) and "It can be hard to type that fast" (SU6). One moderator added the following:

> *It can be hard for someone to post, and their message may get missed.* [MV5]

Another user highlighted the following:

> *Support Chat needs to be more organized...The way it currently is...stresses me out, I find myself leaving midway most of the time.* [SU9]

XSL•FO
RenderX

Social Influences (Social Opportunity)

Similarity and Familiarity

The presence of individuals who had experienced similar situations or who were known to each other was highlighted by moderators and users as helping foster a sense of belonging and encouraging engagement. One user noted the following:

> *It helps that there are people who know what you've been through, so you don't feel alone. It makes you feel more comfortable to open up.* [SU1]

It was seen to be "Easier to have people with similar problems and get ideas surrounding mental health teams etc." (SU3). Several long-term users also highlighted that regular participation by both users and moderators builds trust and familiarity, facilitating engagement. One user stated the following:

> *Because we've been with [the moderators] for some time, they understand our situations a little bit more so it makes it easier for us to open up to them and...we can trust them.* [SU5]

Moderators shared similar insights, noting that regular users "feel quite attached to the moderators, so they'll leave support circle and join support chat for a bit and then come back" (MV5).

Integration of New Users and Gaps in Support

A frequently mentioned barrier by several members and most moderators related to how some regular users tended to chat among themselves or with the moderators, making it difficult for new users to integrate into the group. This could lead to new users feeling like outsiders and being less likely to engage. One user stated the following:

> *When I first started coming to circle, it was very awkward for me because the other people in there had been doing it a lot longer and I felt like a bit of an outsider.* [SU5]

New users described not receiving responses to messages, which seemed to have a notable impact on them. One of them stated the following:

> *What prompts me to withdraw is when some of my messages get ignored. I know it's not on purpose but it kind of makes me feel unwanted.* [SU4]

This was a theme that was also recognized by multiple moderators as a deficiency in social support.

Moderation

As shown in Table 3, a total of 11 barriers and facilitators were identified as subthemes across 10 domains of the TDF. The following paragraphs describe the themes in more detail, along with selected quotes.

Cognitive and Interpersonal Skills (Psychological Capability): Developing Skills

Most moderators discussed how they developed their skills over time through a continuous learning process, which helped them become more effective and refine abilities such as active listening and conflict management:

> *But it's like a continuous learning process, so I wouldn't say they give you kind of training at the beginning of you know, how to be an active listener and how to give that specific type of support.* [MS1]

Observing and practicing moderating also helped develop the necessary skills for moderation. One moderator expressed the following:

> *I think a lot of it is just sort of the exposure and just sort of the repeated attendance of chats and you sort of build, you work out sort of what things you've said have gone down well.* [MS7]

Knowledge (Psychological Capability): Understanding the Guidelines

Knowing and understanding the guidelines and procedures laid out by The Mix was a recurrent enabling theme of moderation, mentioned by all moderators, facilitating the identification of what is and is not deemed inappropriate. One moderator expressed the following:

> *And then the handbook also has things around the tech side of things, so using the platform how to use the moderator functions as well, so we're able to mute people. We're able to freeze them and then remove them from the room as well.* [MS3]

The guidelines were also particularly helpful for directing individuals at risk through appropriate safeguarding procedures. One moderator stated the following:

> *"It's got things from and sort of managing young people that come in who are in crisis. So many kinds of questions to ask safely within the room and the kind of signpost we can give.* [MS2]

Memory, Attention, and Decision Processes (Psychological Capability): Responding Quickly in Complex Situations

The complexity of group chats acted as a barrier for most moderators, presenting challenges by their fast-paced nature and situations that moderators may not have come across before. Attentively responding to the chats becomes challenging. One moderator expressed the following:

> *But there's only so fast you can type, and only so many conversations you can kind of have going on at once that you can't talk to everybody.* [MS8]

In addition, dealing with novel or gray situations caused confusion about what is acceptable in the chat and what is not, making decision-making difficult. One moderator stated the following:

> *There's still some situations that catch me off guard. You know, some people saying things that I've not come across before, experiencing things that I've not come across before.* [MV5]

XSL•FO
RenderX

Behavioral Regulation (Psychological Capability): Holding Back on Advice

Having a natural inclination to offer advice made it difficult for some moderators to self-monitor their responses. Going against instincts and refraining from giving advice when young people shared their problems required self-monitoring of behavior. One moderator expressed the following:

> *It is like to go against that nature and leave it open to the room to support each other and not give advice.* [MS1]

As moderators found it difficult to hold back, effective moderation became challenging. One moderator stated the following:

> *We can't give medical advice, it's tough...we shouldn't really be saying (giving advice) to someone.* [MV5]

Belief About Consequences (Reflective Motivation): Impact of Moderation

Overall, the perceived outcome of moderation for young people, as understood by moderators, facilitated their moderation. Moderators had different beliefs on the impact that moderation had on young people. Some believed that moderation helped create a safe space. One moderator expressed the following:

> *I think young people find it really reassuring to have, like, moderators there because there's a lot of online support spaces, but they're not always moderated or kept safe.* [MS2]

Some moderators felt that young people could feel left out and unsupported. One moderator stated the following:

> *And sometimes the young people can feel like the moderators are not responding to them.* [MS3]

The vast coverage of the impact of moderation was also highlighted by moderators. One moderator expressed the following:

> *I think you do notice the impact because it covers everything OK, like mental health is the main domain. But it's also like everything like education, careers, drugs and alcohol. Just like everything that someone could be going through. So I think you do notice a wider impact.* [MV4]

Social or Professional Role and Identity (Reflective Motivation): Congruence Between Social and Professional Identity

The alignment of personal values and social identity with the professional role of a moderator facilitated moderation for all moderators. One moderator expressed the following:

> *I love volunteering and I love working in a church in the voluntary sector and working with charities. So and I like working with young people. A lot of my work has been with young people. So and yeah, I think it does. It does fit in that respect personally.* [MS3]

The role of a moderator aligned with the identity that moderators had built for themselves. One moderator stated the following:

> *I also work in mental health in my 9 to five job, so that really helps and that's why I wanted to start it in the first place because I started the chat at university and I was trying to broaden my experience working with people with mental health conditions.* [MS7]

Intentions (Reflective Motivation): Intending to Shape Young People's Perceptions

Some moderators expressed a deliberate intention to shape and manage young people's perceptions and expectations of the chat. One moderator expressed the following:

> *I've tried over the years to change the young peoples' perception of what the chat should be used for, I think that in the spirit of trying to make it more of a group conversation like a group chat.* [MS1]

This intention served as a strong facilitator for moderators. One moderator stated the following:

> *So I guess how we manage those situations, yeah, we can warn community members like, we encourage them to take a step back and let the moderators manage the situation. And yeah, we always let them know like they might be removed from the room if they don't like, listen to us.* [MS2]

Emotion (Automatic Motivation): Distressing Subject Matter

Most moderators identified that topics discussed in group chats, such as self-harm and suicidal thoughts, could take an emotional toll on moderators and cause stress and exhaustion. One moderator stated the following:

> *The topics that are being discussed are quite distressing. Examples are feeling very, very low and sometimes it becomes a bit of an echo chamber of you know, they're all talking about feeling like they want to end their life or feeling like self-harming.* [MS1]

These topics also gave way to discussions about the larger mental health ecosystem and evoked concern among the moderators for the young people. One moderator expressed the following:

> *Hard when you see them struggling. And so yeah, I think when that happens, you like worry quite a lot and it does kind of be on your mind for like a few days.* [MS2]

These distressing thoughts and related feelings inevitably acted as a barrier for moderators.

Environmental Context and Resources (Physical Opportunity)

Organizational Resources

The multitude of resources provided by The Mix was appreciated by all moderators. These encompassed guidelines, debrief forms, newsletters, and more, helping them moderate effectively. One moderator stated the following:

XSL•FO
RenderX

> *It really helps to have that handbook there to walk you through those guidelines and how to respond.* [MS3]

Moderation was facilitated through the ready availability of these resources, providing moderators with content that could be used in the chat. One moderator expressed the following:

> *And we have our Mod handbook which is really helpful in terms of giving us like example messages to pop in the group chat and to manage certain situations.* [MS7]

Lack of Visual Cues

Working in a web-based environment meant that moderators had to work without any visual cues, such as facial expressions. One moderator stated the following:

> *I think moderating is quite a unique task...especially when it's online.[MS2]*

This absence made it difficult for moderators to gauge the intensity of young people's emotions and thus required extra effort and attention, which could be a barrier for some moderators. One moderator expressed the following:

> *I guess you have to put a bit of extra effort into that because you don't have the visual cues of like, nodding and all that stuff.* [MS1]

Social Influences (Social Opportunity): Support Offered

Support offered during and after moderation (eg, support from supervisors and emails checking up on moderators) acted as an enabling influence. All respondents had an overwhelmingly positive response toward the support that they received from their supervisors and believed that this support helped them moderate. A respondent stated the following:

> *[The Mix] never makes you feel unappreciated. Like we get quite regular emails of just a reminder that you're all doing really good things...and I think yeah, when you have had a difficult chat and sort of things, they're feeling a bit hard, just things like that really make a difference.* [MS7]

In addition, volunteers reported that the presence of a supervisor in the chats made moderation easier. One moderator stated the following:

> *The moderator would flag it to the supervisor on shift and just say you know ohh it looks like these two are having a bit of a disagreement. What should we do?* [MS1]

One moderator expressed the following:

> *They are really understanding about wanting someone else to step in with a conversation that hits too close to home or needing 5 mins to yourself during the session. There is also a debrief form and if I say anything I felt unhappy with, they always chase it up to check on me or ask if I want more training in the area.* [MV6]

BCTs to Address Barriers

To optimize user engagement, Table 4 presents the BCTs identified to address the core barriers and enhance the facilitators, along with examples.

Table 5 shows the potential intervention types along with the corresponding BCTs to overcome the identified barriers to moderation.

XSL•FO
RenderX

Table 4. Barriers to engagement and strategies for change.

TDF[a] domain	Barrier	Intervention types	Potential BCTs[b]	Operationalization of BCTs
Cognitive and interpersonal skills	Difficulty expressing feelings and needs	Training	Instruction on how to perform a behavior (4.1[c]); behavioral practice/rehearsal (8.1)	Self-reflection worksheets: provide optional worksheets with guidance on identifying emotions and the support needed before users' participation in the group chat.
Memory, attention, and decision processes	Memory recall of guidelines	Enablement	Commitment (1.9)	Group agreement (in the support chat): collectively agree to abide by principles and behaviors aligned with the guidelines at the start of the support chat.
Memory, attention, and decision processes	Memory recall of guidelines	Education	Prompts/cues (7.1)	Reminders: implement periodic reminders or prompts about the chat guidelines to reinforce their importance and improve memory recall.
Beliefs about consequences	Judgment and confidentiality concerns	Education	Information about emotional consequences (5.6)	Share success stories: highlight success stories or testimonials from other young individuals who have benefited from engaging in peer-to-peer group chats.
Emotions	Fear and anxiety	Enablement	Reduce negative emotions (11.2)	Clarify burden misconceptions: provide information and resources from peers about the value of sharing and supporting each other, and share examples of challenges discussed by previous service-users in groups chats to address misconceptions that sharing is a burden and facilitate engagement.
Environmental context and resources	Overload and lack of structure	Environmental restructuring	Restructuring the physical environment (12.1)	Group similar topics and establish topic rotation: encourage moderators and members to brainstorm and group similar topics together either during the support chat (eg, via a poll website such as Slido) or before the chat (eg, via a poll on the discussion boards), allowing for more focused and meaningful discussions instead of fragmented conversations (alternative: create more themed chat sessions).
Environmental context and resources	Overload and lack of structure	Environmental restructuring	Restructuring the social environment (12.2)	Limit chat size: implement a cap on the number of members allowed in the support chat at 1 time to maintain a manageable and supportive group size.
Environmental context and resources	Overload and lack of structure	Environmental restructuring	Adding objects to the environment (12.5)	Emoticons: consider moving to a new chat software that enables members to offer immediate reactions to others' messages via emoticons rather than messages.
Social influences	Integration of new users and gaps in support	Enablement	Social support (unspecified; 3.1) and social support (practical; 3.2)	Buddy system: assign new users a designated "welcoming buddy" (peer) whose role is to encourage or practically facilitate interactions during their initial sessions; grouping or pairing system: pairing or grouping members in the chat as each other's dedicated "supporters" for the session to enhance peer support and encourage more engagement.

[a]TDF: Theoretical Domains Framework.

[b]BCT: behavior change technique.

[c]Behaviour Change Technique code numbers as per the Behaviour Change Technique Taxonomy provided in Michie et al [21].

XSL•FO
RenderX

Table 5. Barriers to moderation and strategies for change.

TDF[a] domain	Barrier	Intervention types	Potential BCTs[b]	Operationalization of BCTs
Memory, attention, and decision processes	Responding quickly to complex situations	Training	Behavioral practice/rehearsal (8.1[c])	Practice sessions: create practice moderation sessions that simulate the quick nature of group chats, which could help moderators to enhance their attention and decision-making processes.
Behavioral regulation	Holding back on advice	Training	Self-monitoring of behavior (2.3)	Diaries: ask moderators to make a note of situations and monitor when they feel instinctively inclined to offer advice to avoid it in the future.
Emotion	Distressing subject matter	Enablement	Reduce negative emotions (11.2)	Positive messaging: providing information about the positive impact of moderation through regular updates (eg, weekly or biweekly emails such as "Here's the impact that you helped deliver!") to increase moderators' positive emotions and reduce concerns about young people.
Environmental context and resources	Lack of visual cues	Environmental restructuring	Prompts/cues (7.1)	Phrase book: prompts or cues may be provided to compensate for the lack of visual cues. A booklet of phrases or "words to look out for" may be provided such that moderators may look out for these words to gauge if an individual is at risk, which may otherwise be missed.

[a]TDF: Theoretical Domains Framework.

[b]BCT: behavior change technique.

[c]Behaviour Change Technique code numbers as per the Behaviour Change Technique Taxonomy provided in Michie et al [21].

Discussion

Principal Findings

This study addresses a research gap by investigating the influences on engagement with, and moderation of, synchronous web-based peer support group chats to support young people's mental health. Using the BCW framework, thematic analysis revealed a total of 20 themes, 9 (45%) for engagement and 11 (55%) for moderation. Of the 20 themes, 3 (15%) were facilitators of engagement, 7 (35%) were facilitators of moderation, 4 (20%) were barriers to moderation, and 6 (30%) were barriers to engagement. The following discussion focuses on the COM-B and TDF themes that were prominent and common across both user engagement and moderation. The findings of this study are discussed in relation to previous research, and the potential BCTs to address the identified barriers to engagement and moderation are contextualized.

Facilitators Common to Moderation and Engagement

Cognitive and interpersonal skills and knowledge (psychological capability) enabled users to express their feelings and needs. Both moderators and users benefitted from having the skills to listen and respond to other users. Moderators developed these skills through practice and were supported by a strong knowledge and understanding of the chat guidelines. Similar themes have been found in other modes of web-based mental health support [39].

Intentions (reflective motivation) were also core themes for moderators and users—both were driven by their desire to help others and ensure that the chat was a valuable resource for users. This sentiment is echoed elsewhere in the literature on web-based support groups, where listeners aimed to create a "safe and warm" space for their clients [51]. Similarly, moderators in this study consciously resolved to provide a safe and nonjudgmental space for young people. This is a theme that can be seen elsewhere in the literature, relating to other web-based peer mental health communities [30].

Finally, for both engagement and moderation, social influences (social opportunity) was another facilitator, where other users, moderators, or supervisors were seen as understanding, supportive, and appreciative of others' needs. For users, this meant being able to share experiences in a safe environment, and moderators felt valued, appreciated, and encouraged in their moderation. Such support also promotes moderators' mental health. As has been previously suggested by Aldamman et al [52], perceived organizational support was positively related to mental well-being, reduced emotional exhaustion, and reduced stress among humanitarian volunteers. This aspect is also linked to the potential of moderators to contribute to the enhancement of the mental well-being of service users, as demonstrated by Perry et al [19].

Barriers Common to Moderation and Engagement

Group chats present a complex environment due to their dynamic and fast-paced nature, with multiple users. Users could find it difficult to remember and adhere to the guidelines, especially in pressurized situations (eg, crises). Although moderators were familiar with the guidelines, they found it challenging to make real-time decisions, particularly during busy chats when many users were interacting rapidly with one another. This inherent complexity can function as a barrier to effective moderation and engagement. With a continuous inflow of new messages in group chats, it becomes challenging to focus on messages and identify any specific theme, as noted by Li et al [53]. Moderators in this study also encountered difficulties in promptly and accurately assessing messages for potential at-risk situations, under time pressure, while simultaneously

XSL•FO
RenderX

ensuring that all the young people in the chat felt supported. The BCT of "Behavioral practice" can help equip moderators to handle novel situations quickly as they arise. This BCT has been successfully used to "positively influence" behavior in the context of user training [54] and in interventions supporting shared decision-making [55].

The challenges with rapid decision-making of how to respond partly resulted from the chat environment being unstructured and uncontrolled in terms of the volume and speed of message exchanges; this was a barrier for moderators and users. Implementing emoticons (BCT "adding objects to the environment") could offer an alternative and rapid way of engaging with messages, helping to express emotions while reducing message overload [56]. Previous studies have also highlighted participants of a Cognitive Behavioral Therapy-based peer support platform feeling overwhelmed and stressed due to message volume and unfamiliar dynamics [57]. This underscores the need for accessible and well-organized chat platforms. One promising BCT is "restructuring the physical environment." Similar topics could be grouped within the platform or support chat sessions could be themed to allow users to explore and self-organize into groups based on shared characteristics and pain points, as has been implemented elsewhere [58].

Another aspect of the environment that was highlighted in previous studies was the lack of visual clues to support decision-making. This was a particular barrier for moderators, impeding connection and relationship building [39,59]. The BCT "prompts/cues" has been previously suggested in a similar context to improve moderation [39]. These could be in the form of electronic prompts or suggested phrases that pop-up when a young person writes a phrase that may need to be flagged. Such prompts may support moderators to identify individuals at risk and compensate for the lack of visual cues. This BCT has been successfully used to prompt action by the user in various contexts [60] and can also be used to simultaneously address barriers related to memory difficulties.

Given the sensitive nature of the discussions within web-based peer support chats, it is perhaps unsurprising that emotions can constitute another barrier. The BCT of "reduce negative emotions" can help address this. For users, the fear and anxiety associated with disclosure can prevent them from sharing in the chats, hindering their access to the help they need. Previous interventions report participants finding it helpful to know that others were undergoing similar emotions, reducing feelings of isolation [61]. Therefore, providing information and resources from peers about the value of sharing and supporting each other, and sharing examples of challenges discussed in group chats by previous service users may facilitate engagement. This could further build a sense of a supportive and friendly community on the platform, which facilitates willingness to share feelings and difficulties in other contexts [57]. Dealing with distressing topics such as mental health issues, suicidal thoughts, and other issues that young people are dealing with often becomes emotionally exhausting for moderators. In this case, the BCT "reduce negative emotions" could involve providing moderators with information about the positive impact of their moderation, to counter any negative emotions. A scoping review highlighted that this BCT has frequently been used in the development of mental health interventions [62]. Resilience-building training may also be helpful in equipping moderators to deal with the "emotional cost" of content moderation [63].

Barriers Specific to User Engagement

The barrier of "difficulties expressing feelings and support needs" experienced by users aligns with a prior study on webchat counseling engagement among young people, which found users lacking self-expression skills [28]. "Constructive emotional sharing" is a skill that can be improved with practice [64]. Promising BCTs to address this barrier include "instruction on how to perform a behavior" and "behavioral practice/rehearsal." For example, self-reflection worksheets could encourage users to identify their feelings and practice written statements before participating in group chats [64]. Statements such as "I feel...when...because" could be shared, leading to more personalized responses [65]. These BCTs could prove particularly beneficial for newcomers or individuals less familiar with the dynamics of group chat interaction, also creating positive spillover to a further barrier "integration of new users and gaps in support." A barrier identified here aligns with a previous study, which found that receiving empathic comments initially has a significant cascading effect, motivating individuals to reciprocate and offer support to others [66]. To further encourage integration and support for new users, the BCT "social support" could be introduced by assigning a designated "welcoming buddy" whose role is to support new members during their initial sessions or pairing members in the chat as each other's "supporter" to enhance peer support and encourage more engagement. Prior research has shown that shared interests encouraged conversation between new pairs in a peer support intervention, who were strangers when they were initially paired [67].

Having judgment and confidentiality concerns constituted another barrier. Overcoming such concerns can contribute to more fluid communication in web-based peer support and are a core component of building trust [67]. The BCT "information about emotional consequences" may address this barrier through sharing success stories, statistics, or personal narratives. Within the recovery model, this increases a sense of acceptance, understanding, and authenticity, particularly for new service users who tend to experience these fears more than long-term users [68].

Barriers Specific to Moderation

Moderators found it difficult to hold back on offering advice. Rooted in the logic of care [69], moderation is contextualized and involves an empathetic dialogue or interaction between moderators and users [70]. In such an open-ended process, holding back, given the natural flow of an interaction, proved a challenge for moderators. "Self-monitoring of behavior" could be a potential BCT that may help moderators to keep their behavior in check and change it. This could be achieved by encouraging moderators to note when they instinctively offer advice or feel like they want to do so, allowing them to reflect and monitor their own behavior if such situations arise again. This technique has proved effective in many behavior change

XSL•FO
RenderX

interventions in different contexts [71,72], particularly for health behaviors [73,74].

Limitations

The findings of this study need to be considered in light of certain limitations. The sampling of this study was reliant on self-selection, which may have introduced bias and favored participation by more confident users and moderators, potentially influencing the barriers and facilitators identified [75]. Although the participants were assured that their responses would remain confidential and would not have any consequences with respect to their relationship with The Mix, they may nevertheless have given socially desirable responses, whether intentionally or subconsciously [76]. Furthermore, it is noteworthy that the study participants were predominately women, making the sample relatively homogeneous. However, the overall population of moderators and group chat users at The Mix is also mainly women (>70% of users and 95% of moderators). Finally, the study's findings are grounded in the data obtained from a single mental health group chat forum, which could limit the generalizability of the results to other chat contexts operating under different circumstances.

Conclusions

Group chats are an increasingly popular form of digital mental health intervention, and this study contributes to building the evidence base, which can help optimize them as a safe and timely form of mental health support for young people. It is particularly valuable as it examines synchronous group chats, which are characterized by in-the-moment empathic interactions and emotional connections. Through using the COM-B and TDF, the study found that skills and knowledge, beliefs about consequences and intentions, emotions, and the social and physical environment are important factors influencing both the users and moderators of group chats. In particular, supporting the improvement of memory, attention, and decision-making skills of those involved; adapting the physical environment to facilitate effective interactions; and reducing negative emotions are suggested to optimize the value and effectiveness of group chats for young people's mental health support for both the users and moderators of these services. The intervention types and BCTs proposed serve to emphasize the importance of training and support, particularly for moderators, in this important role. The study also further demonstrates the effectiveness of the BCW approach and the use of the TDF and COM-B to understand the influences on behavior in a systematic manner, especially for mental health and well-being interventions.

A natural progression of this work would be to implement and evaluate the interventions proposed in the study and gauge to what extent the suggested BCTs reduce the identified barriers. In addition, the fidelity of the BCTs could also be assessed to understand the nuances of intervention delivery [77] to facilitate contextualized tailoring of the intervention. In an environment where digital mental health interventions for young people continue to grow in significance, this study and future suggested studies can contribute toward ensuring that they are evidence based and consider the voices of young people themselves.

Conflicts of Interest

None declared.

References

1. Ogden T, Hagen KA. Adolescent Mental Health: Prevention and Intervention. London, UK: Routledge; 2018.
2. Kauhanen L, Wan Mohd Yunus WM, Lempinen L, Peltonen K, Gyllenberg D, Mishina K, et al. A systematic review of the mental health changes of children and young people before and during the COVID-19 pandemic. Eur Child Adolesc Psychiatry 2023 Jun 12;32(6):995-1013 [FREE Full text] [doi: 10.1007/s00787-022-02060-0] [Medline: 35962147]
3. Creswell C, Shum A, Pearcey S, Skripkauskaite S, Patalay P, Waite P. Young people's mental health during the COVID-19 pandemic. The Lancet Child Adolescent Health 2021 Aug;5(8):535-537. [doi: 10.1016/s2352-4642(21)00177-2]
4. Mental Health of Children and Young People in England 2022 - wave 3 follow up to the 2017 survey. National Health Service England. 2022 Nov 29. URL: https://digital.nhs.uk/data-and-information/publications/statistical/mental-health-of-children-and-young-people-in-england/2022-follow-up-to-the-2017-survey [accessed 2025-01-12]
5. Mental Health of Children and Young People in England, 2023 - wave 4 follow up to the 2017 survey. National Health Service England. 2023 Nov 21. URL: https://digital.nhs.uk/data-and-information/publications/statistical/mental-health-of-children-and-young-people-in-england/2023-wave-4-follow-up [accessed 2025-01-12]
6. McGorry PD, Mei C, Chanen A, Hodges C, Alvarez-Jimenez M, Killackey E. Designing and scaling up integrated youth mental health care. World Psychiatry 2022 Feb 11;21(1):61-76 [FREE Full text] [doi: 10.1002/wps.20938] [Medline: 35015367]
7. Facts and figures. Children & Young People's Mental Health Coalition. URL: https://cypmhc.org.uk/resources/facts-and-figures/ [accessed 2025-01-12]
8. Yang D, Kraut RE, Smith T, Mayfield E, Jurafsky D. Seekers, providers, welcomers, and storytellers: modeling social roles in online health communities. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 2019 Presented at: CHI '19; May 4-9, 2019; Glasgow, UK. [doi: 10.1145/3290605.3300574]
9. Ali K, Farrer L, Gulliver A, Griffiths KM. Online peer-to-peer support for young people with mental health problems: a systematic review. JMIR Ment Health 2015 May 19;2(2):e19 [FREE Full text] [doi: 10.2196/mental.4418] [Medline: 26543923]

XSL•FO
RenderX

10. Castelnuovo G, Gaggioli A, Mantovani F, Riva G. From psychotherapy to e-therapy: the integration of traditional techniques and new communication tools in clinical settings. Cyberpsychol Behav 2003 Aug;6(4):375-382. [doi: 10.1089/109493103322278754] [Medline: 14511449]
11. Dowling M, Rickwood D. Online counseling and therapy for mental health problems: a systematic review of individual synchronous interventions using chat. J Technol Hum Serv 2013 Jan;31(1):1-21 [FREE Full text] [doi: 10.1080/15228835.2012.728508]
12. Tibbs M, O'Reilly A, Dwan O'Reilly M, Fitzgerald A. Online synchronous chat counselling for young people aged 12-25: a mixed methods systematic review protocol. BMJ Open 2022 Apr 25;12(4):e061084 [FREE Full text] [doi: 10.1136/bmjopen-2022-061084] [Medline: 35470202]
13. Naslund JA, Deng D. Addressing mental health stigma in low-income and middle-income countries: a new frontier for digital mental health. Ethics Med Public Health 2021 Dec;19:100719 [FREE Full text] [doi: 10.1016/j.jemep.2021.100719] [Medline: 35083375]
14. Fortuna KL, Naslund JA, LaCroix JM, Bianco CL, Brooks JM, Zisman-Ilani Y, et al. Digital peer support mental health interventions for people with a lived experience of a serious mental illness: systematic review. JMIR Ment Health 2020 Apr 03;7(4):e16460 [FREE Full text] [doi: 10.2196/16460] [Medline: 32243256]
15. Griffiths KM, Reynolds J, Vassallo S. An online, moderated peer-to-peer support bulletin board for depression: user-perceived advantages and disadvantages. JMIR Ment Health 2015 Apr 24;2(2):e14 [FREE Full text] [doi: 10.2196/mental.4266] [Medline: 26543919]
16. Easton K, Diggle J, Ruethi-Davis M, Holmes M, Byron-Parker D, Nuttall J, et al. Qualitative exploration of the potential for adverse events when using an online peer support network for mental health: cross-sectional survey. JMIR Ment Health 2017 Oct 30;4(4):e49 [FREE Full text] [doi: 10.2196/mental.8168] [Medline: 29084710]
17. Merchant R, Goldin A, Manjanatha D, Harter C, Chandler J, Lipp A, et al. Opportunities to expand access to mental health services: a case for the role of online peer support communities. Psychiatr Q 2022 Jun 28;93(2):613-625 [FREE Full text] [doi: 10.1007/s11126-022-09974-7] [Medline: 35226251]
18. Wadden D, August T, Li Q, Althoff T. The effect of moderation on online mental health conversations. Proc Int AAAI Conf Web Soc Media 2021;15(1):751-763. [doi: 10.1609/icwsm.v15i1.18100]
19. Perry A, Pyle D, Lamont-Mills A, du Plessis C, du Preez J. Suicidal behaviours and moderator support in online health communities: a scoping review. BMJ Open 2021 Jun 30;11(6):e047905 [FREE Full text] [doi: 10.1136/bmjopen-2020-047905] [Medline: 34193497]
20. Kraut RE, Resnick P, Kiesler S. Building Successful Online Communities: Evidence-Based Social Design. Cambridge, MA: MIT Press; 2012.
21. Michie S, Atkins L, West R. The Behaviour Change Wheel: A Guide to Designing Interventions. Sutton, UK: Silverback Publishing; 2014.
22. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun 13;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
23. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011 Aug 05;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]
24. Sieverink F, Kelders SM, van Gemert-Pijnen JE. Clarifying the concept of adherence to eHealth technology: systematic review on when usage becomes adherence. J Med Internet Res 2017 Dec 06;19(12):e402 [FREE Full text] [doi: 10.2196/jmir.8578] [Medline: 29212630]
25. Lakind D, Bradley WJ, Patel A, Chorpita BF, Becker KD. A multidimensional examination of the measurement of treatment engagement: implications for children's mental health services and research. J Clin Child Adolesc Psychol 2022 Jul 16;51(4):453-468 [FREE Full text] [doi: 10.1080/15374416.2021.1941057] [Medline: 34269632]
26. James TL, Villacis Calderon ED, Bélanger F, Lowry PB. The mediating role of group dynamics in shaping received social support from active and passive use in online health communities. Inf Manag 2022 Apr;59(3):103606. [doi: 10.1016/j.im.2022.103606]
27. Liverpool S, Mota CP, Sales CM, Čuš A, Carletto S, Hancheva C, et al. Engaging children and young people in digital mental health interventions: systematic review of modes of delivery, facilitators, and barriers. J Med Internet Res 2020 Jun 23;22(6):e16317 [FREE Full text] [doi: 10.2196/16317] [Medline: 32442160]
28. Mawdsley G, Richiello M, Gutman LM. Barriers and facilitators of young people's engagement with webchat counselling: a qualitative analysis informed by the Behaviour Change Wheel. Couns Psychother Res 2022 Mar 23;22(3):725-735. [doi: 10.1002/capr.12528]
29. Onken SJ, Dumont JM, Ridgway P, Dornan DH, Ralph RO. Mental health recovery: what helps and what hinders? A national research project for the development of recovery facilitating system performance indicators. National Association of State Mental Health Program Director and National Technical Assistance Center for State Mental Health Planning. 2002 Oct. URL: https://www.nasmhpd.org/sites/default/files/MHSIPReport%281%29.pdf [accessed 2025-01-12]

XSL•FO
RenderX

30. Deng D, Rogers T, Naslund JA. The role of moderators in facilitating and encouraging peer-to-peer support in an online mental health community: a qualitative exploratory study. J Technol Behav Sci 2023 Feb 16;8(2):128-139 [FREE Full text] [doi: 10.1007/s41347-023-00302-9] [Medline: 36810998]
31. Prescott J, Rathbone AL, Brown G. Online peer to peer support: qualitative analysis of UK and US open mental health Facebook groups. Digit Health 2020 Dec 10;6:2055207620979209 [FREE Full text] [doi: 10.1177/2055207620979209] [Medline: 33354335]
32. Strand M, Eng LS, Gammon D. Combining online and offline peer support groups in community mental health care settings: a qualitative study of service users' experiences. Int J Ment Health Syst 2020 May 29;14(1):39 [FREE Full text] [doi: 10.1186/s13033-020-00370-x] [Medline: 32514303]
33. Saha K, Ernala SK, Dutta S, Sharma E, de Choudhury M. Understanding moderation in online mental health communities. In: Proceedings of the 12th International Conference on Social Computing and Social Media. Participation, User Experience, Consumer Experience, and Applications of Social Computing. 2020 Presented at: SCSM 2020; July 19-24, 2020; Copenhagen, Denmark. [doi: 10.1007/978-3-030-49576-3_7]
34. Michie S, Johnston M, Abraham C, Lawton R, Parker D, Walker A. Making psychological theory useful for implementing evidence based practice: a consensus approach. Qual Saf Health Care 2005 Feb 01;14(1):26-33 [FREE Full text] [doi: 10.1136/qshc.2004.011155] [Medline: 15692000]
35. Cane J, O'Connor D, Michie S. Validation of the theoretical domains framework for use in behaviour change and implementation research. Implement Sci 2012 Apr 24;7(1):37 [FREE Full text] [doi: 10.1186/1748-5908-7-37] [Medline: 22530986]
36. Michie S, Johnston M. Theories and techniques of behaviour change: developing a cumulative science of behaviour change. Health Psychol Rev 2012 Mar;6(1):1-6. [doi: 10.1080/17437199.2012.654964]
37. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6(1):42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
38. Richiello MG, Mawdsley G, Gutman LM. Using the Behaviour Change Wheel to identify barriers and enablers to the delivery of webchat counselling for young people. Couns Psychother Res 2021 May 04;22(1):130-139. [doi: 10.1002/capr.12410]
39. Perowne R, Gutman LM. Barriers and enablers to the moderation of self-harm content for a young person's online forum. J Ment Health 2024 Jun 16;33(3):357-365 [FREE Full text] [doi: 10.1080/09638237.2022.2069721] [Medline: 35574666]
40. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
41. Azungah T. Qualitative research: deductive and inductive approaches to data analysis. Qual Res J 2018 Oct 31;18(4):383-400. [doi: 10.1108/qrj-d-18-00035]
42. Bradley EH, Curry LA, Devers KJ. Qualitative data analysis for health services research: developing taxonomy, themes, and theory. Health Serv Res 2007 Aug 24;42(4):1758-1772 [FREE Full text] [doi: 10.1111/j.1475-6773.2006.00684.x] [Medline: 17286625]
43. Maguire M, Delahunt B. Doing a thematic analysis: a practical, step-by-step guide for learning and teaching scholars. All Ireland J High Educ 2017 Oct 31;9(3) [FREE Full text] [doi: 10.62707/aishej.v9i3.335]
44. Hallberg LR. The “core category” of grounded theory: making constant comparisons. Int J Qual Stud Health Well Being 2009 Jul 12;1(3):141-148. [doi: 10.1080/17482620600858399]
45. Berger R. Now I see it, now I don’t: researcher’s position and reflexivity in qualitative research. Qual Res 2013 Jan 03;15(2):219-234. [doi: 10.1177/1468794112468475]
46. Cane J, Richardson M, Johnston M, Ladha R, Michie S. From lists of behaviour change techniques (BCTs) to structured hierarchies: comparison of two methods of developing a hierarchy of BCTs. Br J Health Psychol 2015 Feb 12;20(1):130-150. [doi: 10.1111/bjhp.12102] [Medline: 24815766]
47. Johnston M, Carey RN, Connell Bohlen LE, Johnston DW, Rothman AJ, de Bruin M, et al. Development of an online tool for linking behavior change techniques and mechanisms of action based on triangulation of findings from literature synthesis and expert consensus. Transl Behav Med 2021 May 25;11(5):1049-1065 [FREE Full text] [doi: 10.1093/tbm/ibaa050] [Medline: 32749460]
48. Welcome to the Theory and Techniques Tool. Theory and Techniques Tool. URL: https://theoryandtechniquetool.humanbehaviourchange.org/ [accessed 2025-01-12]
49. Allison AL, Purkiss D, Lorencatto F, Miodownik M, Michie S. Improving compostable plastic disposal: an application of the Behaviour Change Wheel intervention development method. Front Sustain 2022 Aug 18;3 [FREE Full text] [doi: 10.3389/frsus.2022.968152]
50. Gutman LM, Perowne R, Younas F, O'Hanrachtaigh E. Making hybrid work for diverse staff in higher education: a behaviour change approach. High Educ Q 2023 Dec 15;78(3):784-806. [doi: 10.1111/hequ.12487]
51. Barak A. Emotional support and suicide prevention through the internet: a field project report. Comput Hum Behav 2007 Mar;23(2):971-984. [doi: 10.1016/j.chb.2005.08.001]

XSL•FO
RenderX

52. Aldamman K, Tamrakar T, Dinesen C, Wiedemann N, Murphy J, Hansen M, et al. Caring for the mental health of humanitarian volunteers in traumatic contexts: the importance of organisational support. Eur J Psychotraumatol 2019 Dec 03;10(1):1694811 [FREE Full text] [doi: 10.1080/20008198.2019.1694811] [Medline: 31839900]
53. Li H, Kraut RE, Zhu H. Technical features of asynchronous and synchronous community platforms and their effects on community cohesion: a comparative study of forum-based and chat-based online mental health communities. J Comput Mediat Commun 2021;26(6):403-421. [doi: 10.1093/jcmc/zmab016]
54. Wood CE, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. Applying the behaviour change technique (BCT) taxonomy v1: a study of coder training. Transl Behav Med 2015 Jun 19;5(2):134-148 [FREE Full text] [doi: 10.1007/s13142-014-0290-z] [Medline: 26029276]
55. Agbadjé TT, Elidor H, Perin MS, Adekpedjou R, Légaré F. Towards a taxonomy of behavior change techniques for promoting shared decision making. Implement Sci 2020 Aug 20;15(1):67 [FREE Full text] [doi: 10.1186/s13012-020-01015-w] [Medline: 32819410]
56. Aldunate N, Villena-González M, Rojas-Thomas F, López V, Bosman CA. Mood detection in ambiguous messages: the interaction between text and emoticons. Front Psychol 2018 Apr 4;9:423 [FREE Full text] [doi: 10.3389/fpsyg.2018.00423] [Medline: 29670554]
57. Browne NL, Carragher NO, O'Toole A, Pimm J, Ryder J, Thew GR. Evaluating user experiences of SHaRON: an online CBT-based peer support platform. The Cogn Behav Ther 2022 Apr 11;15:e18. [doi: 10.1017/s1754470x22000150]
58. O'Leary K, Bhattacharya A, Munson SA, Wobbrock JO, Pratt W. Design opportunities for mental health peer support technologies. In: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing. 2017 Presented at: CSCW '17; February 25-March 1, 2017; Portland, OR. [doi: 10.1145/2998181.2998349]
59. Lau PL, Jaladin RA, Abdullah HS. Understanding the two sides of online counseling and their ethical and legal ramifications. Procedia Soc Behav Sci 2013 Nov;103:1243-1251. [doi: 10.1016/j.sbspro.2013.10.453]
60. Morrissey EC, Corbett TK, Walsh JC, Molloy GJ. Behavior change techniques in apps for medication adherence: a content analysis. Am J Prev Med 2016 May;50(5):e143-e146. [doi: 10.1016/j.amepre.2015.09.034] [Medline: 26597504]
61. Borghouts J, Eikey E, Mark G, de Leon C, Schueller SM, Schneider M, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. J Med Internet Res 2021 Mar 24;23(3):e24387 [FREE Full text] [doi: 10.2196/24387] [Medline: 33759801]
62. Martinengo L, Jabir AI, Goh WWT, Lo NYW, Ho MHR, Kowatsch T, et al. Conversational agents in health care: scoping review of their behavior change techniques and underpinning theory. J Med Internet Res 2022 Oct 03;24(10):e39243 [FREE Full text] [doi: 10.2196/39243] [Medline: 36190749]
63. Steiger M, Bharucha TJ, Venkatagiri S, Riedl MJ, Lease M. The psychological well-being of content moderators: the emotional labor of commercial moderation and avenues for improving support. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. 2021 Presented at: CHI '21; May 8-13, 2021; Yokohama, Japan. [doi: 10.1145/3411764.3445092]
64. van Dyke K. Feeling stuck? How to express your feelings. Psych Central. 2021 Jun 17. URL: https://psychcentral.com/lib/feeling-stuck-how-to-express-your-feelings [accessed 2025-01-12]
65. Doherty G, Coyle D, Sharry J. Engagement with online mental health interventions: an exploratory clinical study of a treatment for depression. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2012 Presented at: CHI '12; May 5-10, 2012; Austin, TX. [doi: 10.1145/2207676.2208602]
66. Chen Y, Xu Y. Exploring the effect of social support and empathy on user engagement in online mental health communities. Int J Environ Res Public Health 2021 Jun 26;18(13):6855 [FREE Full text] [doi: 10.3390/ijerph18136855] [Medline: 34206719]
67. Andalibi N, Flood MK. Considerations in designing digital peer support for mental health: interview study among users of a digital support system (Buddy Project). JMIR Ment Health 2021 Jan 04;8(1):e21819 [FREE Full text] [doi: 10.2196/21819] [Medline: 33393909]
68. Peck CE, Lim MH, Purkiss M, Foley F, Hopkins L, Thomas N. Development of a lived experience-based digital resource for a digitally-assisted peer support program for young people experiencing psychosis. Front Psychiatry 2020 Jul 2;11:635 [FREE Full text] [doi: 10.3389/fpsyt.2020.00635] [Medline: 32714223]
69. Mol A. The Logic of Care: Health and the Problem of Patient Choice. Milton Park, UK: Taylor & Francis; 2008.
70. Ruckenstein M, Turunen LL. Re-humanizing the platform: content moderators and the logic of care. New Media Soc 2019 Sep 19;22(6):1026-1042. [doi: 10.1177/1461444819875990]
71. Keller J, Kwasnicka D, Wilhelm LO, Lorbeer N, Pauly T, Domke A, et al. Hand washing and related cognitions following a brief behavior change intervention during the COVID-19 pandemic: a pre-post analysis. Int J Behav Med 2022 Oct 29;29(5):575-586 [FREE Full text] [doi: 10.1007/s12529-021-10042-w] [Medline: 34843096]
72. Berry R, Kassavou A, Sutton S. Does self-monitoring diet and physical activity behaviors using digital technology support adults with obesity or overweight to lose weight? A systematic literature review with meta-analysis. Obes Rev 2021 Oct 30;22(10):e13306. [doi: 10.1111/obr.13306] [Medline: 34192411]
73. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol 2009 Nov;28(6):690-701. [doi: 10.1037/a0016136] [Medline: 19916637]

XSL•FO
RenderX

74. Garnett CV, Crane D, Brown J, Kaner EF, Beyer FR, Muirhead CR, et al. Behavior change techniques used in digital behavior change interventions to reduce excessive alcohol consumption: a meta-regression. Ann Behav Med 2018 May 18;52(6):530-543 [FREE Full text] [doi: 10.1093/abm/kax029] [Medline: 29788261]
75. Gautier PA, Klaauw BV. Selection in a field experiment with voluntary participation. J Appl Econometri 2012 Jan 18;27(1):63-84. [doi: 10.1002/jae.1184]
76. Bispo Júnior JP. Social desirability bias in qualitative health research. Rev Saude Publica 2022 Nov 18;56:101 [FREE Full text] [doi: 10.11606/s1518-8787.2022056004164] [Medline: 36515303]
77. Toomey E, Hardeman W, Hankonen N, Byrne M, McSharry J, Matvienko-Sikar K, et al. Focusing on fidelity: narrative review and recommendations for improving intervention fidelity within trials of health behaviour change interventions. Health Psychol Behav Med 2020 Mar 12;8(1):132-151 [FREE Full text] [doi: 10.1080/21642850.2020.1738935] [Medline: 34040865]

Abbreviations

BCT: behavior change technique
BCTT: Behavior Change Techniques Taxonomy
BCW: Behavior Change Wheel
COM-B: Capability, Opportunity, Motivation, and Behavior
TDF: Theoretical Domains Framework

Edited by A Kushniruk; submitted 08.07.24; peer-reviewed by D Levine, L Valentine; comments to author 24.11.24; revised version received 05.12.24; accepted 31.12.24; published 05.02.25.

Please cite as:
Ananya A, Tuuli J, Perowne R, Gutman LM
Barriers and Facilitators to User Engagement and Moderation for Web-Based Peer Support Among Young People: Qualitative Study Using the Behavior Change Wheel Framework
JMIR Hum Factors 2025;12:e64097
URL: https://humanfactors.jmir.org/2025/1/e64097
doi:10.2196/64097
PMID:

©Ananya Ananya, Janina Tuuli, Rachel Perowne, Leslie Morrison Gutman. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 05.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Safety of Digital Mental Health Interventions: Findings and Recommendations From a Qualitative Study Exploring Users' Experiences, Concerns, and Suggestions

Rayan Taher[1], MSc; Daniel Stahl[2], PhD; Sukhi Shergill[1,3], MD, PhD; Jenny Yiend[1], PhD

[1]Department of Psychosis Studies, King's College London, London, United Kingdom

[2]Department of Biostatistics and Health Informatics, King's College London, London, United Kingdom

[3]Kent and Medway Medical School, Canterbury, United Kingdom

Corresponding Author:
Jenny Yiend, PhD
Department of Psychosis Studies
King's College London
Institute of Psychiatry, Psychology & Neuroscience (IoPPN)
16 De Crespigny Park
London, SE5 8AB
United Kingdom
Phone: 44 020 7848 0002
Email: jenny.yiend@kcl.ac.uk

Abstract

Background: The literature around the safety of digital mental health interventions (DMHIs) is growing. However, the user/patient perspective is still absent from it. Understanding the user/patient perspective can ensure that professionals address issues that are significant to users/patients and help direct future research in the field.

Objective: This qualitative study aims to explore DMHI users' experiences, views, concerns, and suggestions regarding the safety of DMHIs.

Methods: We included individuals aged 18 years old or older, having experience in using a DMHI, and can speak and understand English without the need for a translator. Fifteen individual interviews were conducted. Deductive thematic analysis was used to analyze the data.

Results: The analysis of the interview transcripts yielded 3 main themes: Nonresponse: A Concern, a Risk, and How Users Mitigate It, Symptom Deterioration and Its Management, and Concerns Around Data Privacy and How to Mitigate Them.

Conclusions: The results of this study led to 7 recommendations on how the safety of DMHIs can be improved: provide "easy access" versions of key information, use "approved by..." badges, anticipate and support deterioration, provide real-time feedback, acknowledge the lack of personalization, responsibly manage access, and provide genuine crisis support. These recommendations arose from users' experiences and suggestions. If implemented, these recommendations can improve the safety of DMHIs and enhance users' experience.

(JMIR Hum Factors 2025;12:e62974) doi:10.2196/62974

KEYWORDS

digital mental health; safety; user perspective; patient perspective; qualitative; risks; risk mitigation; deterioration; nonresponse; data safety

Introduction

Digital mental health interventions (DMHIs) are mental health interventions that are delivered through digital platforms such as mobile apps, websites, or virtual reality [1]. Some of the added benefits of mental health interventions that are delivered digitally are improved accessibility, scalability, convenience, and the potential for anonymous engagement [2]. To realize these benefits, users need to trust that these interventions are effective and safe [2]. The evidence shows that DMHIs can be as effective as traditional face-to-face therapies, especially for common mental health disorders such as depression and anxiety

XSL•FO
RenderX

[3,4]. However, the safety of DMHIs is still an evolving field [1,2,5,6].

DMHI's users face similar risks to those in face-to-face therapy, such as deterioration in symptoms, novel symptoms (experiencing new mental health symptoms during treatment), and nonresponse [1]. Deterioration of symptoms, observed in approximately 3%-10% of psychotherapy cases [7,8], signifies a phenomenon where patients' conditions worsen during therapy. Deterioration is the most common side effect of mental health therapies (face to face and digital) [1]. There is a debate in the literature about whether deterioration is a normal and integral part of therapy or an unnecessary side effect [9,10]. A recent experts' consensus study concluded that short-term deterioration that occurs during therapy is part of therapy and should not be considered a safety concern; however, deterioration still needs to be monitored to ensure that it is not chronic, severe, and does not lead to an adverse event such as the patient dropping out of treatment [11]. Nonresponse occurs when the therapy is not effective in relieving the target symptoms [12]. It is considered a negative outcome as it hinders access to more effective treatments, spontaneous remission, and may prolong or even increase distress [12]. Additionally, the digital nature of DMHIs introduces additional risks to mental health therapies such as technical issues and privacy concerns [5,13].

The expanding body of literature addressing the safety of DMHIs is noteworthy [1,5,6]. While considerable attention has been devoted to exploring the safety of DMHIs (how it is assessed, analyzed, and reported), a notable gap exists in the qualitative understanding of individual perspectives on the topic. Existing qualitative studies in this field have either focused on the viewpoints of health professionals and medical students [14,15] or have sought user opinions on specific digital innovations usually as part of a wider program of development work. A few studies have sought service users' views on digital interventions more generally [16-18], but no work to our knowledge has investigated user/patient perspectives specifically about the safety of these technologies [19,20]. Understanding the user/patient perspective on safety can help direct future research in the field to ensure professionals focus on issues significant to users/patients; identifying users' concerns can help professionals address these issues, leading to higher rates of adherence and engagement. For that reason, this qualitative study aims to explore DMHIs users' experiences, views, concerns, and suggestions regarding the safety of DMHIs.

Methods

Design and Aim

This qualitative study utilized individual interviews to explore users' experiences, views, concerns, and suggestions about the safety of DMHIs. The Consolidated Criteria for Reporting Qualitative Research (COREQ) was used to report the results of this study [21].

Recruitment

The study included individuals aged 18 years or older, with previous experience using a DMHI, and speaking and understanding English without the need for a translator. The DMHI needed to be a mental health intervention that was provided via a tech-based medium (eg, app, website, virtual reality) and targeted a specific mental health condition. Participants were recruited using nonpurposive sampling by posting advertisements on authors' own professional social media platforms, such as X (formerly Twitter; X Corp.) and LinkedIn (Microsoft Corporation), an online participant recruiting platform (MQ; MQ Mental Health Research), the university's (King's College London) research volunteering circular email, and through DMHI trials whose participants consented for their details to be shared for future research. A total of 54 potential participants reached out; 15 (28%) were eligible and participated in the study, 7 (13%) did not complete the online eligibility form, and 32 (59%) were ineligible. Eligibility was determined through screening questions asked of potential participants to gather further details about the specific intervention used. The main reasons for ineligibility were that the intervention used was telehealth (eg, face-to-face therapy conducted via video call) or did not target a specific mental health condition (eg, mindfulness apps).

Participants

A total of 15 participants were recruited to participate in this study. Researchers initially estimated a sample size of 6-16 participants based on pragmatic recommendations from the literature, suggesting that 6-16 interviews provide sufficient information power [22]. Then the final sample size (15 participants) was determined based on the richness of the data and their ability to sufficiently answer the research question [23]. Of the 15 participants, 12 (80%) were females and 3 (20%) were males. All participants lived in the United Kingdom at the time of the interviews. Participants had an average age of 30 years (SD 6.43 years; range 19-42 years). On average, participants had used a DMHI for 8 months (SD 10.51 months; range 1-36 months). The DMHIs used by participants in this study were Beating the Blues, Calm Harm, FREED-M, Happify, Molehill Mountain, Moodkit, Silver Cloud, Sleepio, STOP app, Woebot, Youper, and an online intervention for Bulimia. See Table 1 for more information on the interventions used by participants in this study. Some participants used the DMHI for both anxiety and depression. Other than that, no other participants used the DMHI for more than 1 target symptom or condition.

XSL•FO
RenderX

Table 1. The DMHIs[a] used by participants in this study (N=15).

DMHI	Values, n (%)	Duration (months) of use, mean (range)
DMHI's target symptom/condition		
Depression and anxiety	5 (33)	3.4 (1-6)
Depression	2 (13)	19.25 (2.5-36)
Self-harm	2 (13)	12.75 (1.5-24)
Paranoia	1 (7)	3 (N/A[b])
Insomnia	1 (7)	1.5 (N/A)
Postnatal depression	1 (7)	24 (N/A)
Bulimia	1 (7)	1.5 (N/A)
Anxiety in autism	1 (7)	3 (N/A)
Eating disorders	1 (7)	3 (N/A)
DMHI's format		
Web-based	7 (47)	2.5 (1-6)
App-based	6 (40)	11.5 (1.5-36)
Artificial intelligence chatbot	2 (13)	15 (6-24)
Therapist involvement		
Self-administered (users independently used the DMHI without any support)	12 (80)	9.33 (1.5-36)
Hybrid (users independently used the DMHI while receiving regular support)	3 (20)	1.66 (1-2)
How participants found the DMHI		
Health care professional	7 (47)	N/A
Social media	2 (13)	N/A
App store	2 (13)	N/A
DMHI research	2 (13)	N/A
Work (via human resources)	1 (7)	N/A
University website	1 (7)	N/A

[a]DMHI: digital mental health intervention.

[b]N/A: not applicable.

Materials

Individual interviews with participants were conducted and recorded online via Microsoft Teams (Microsoft Corporation). Interviews were semistructured. See Multimedia Appendix 1 for the topic guide. The interviewer (RT) used prompts to facilitate and guide the discussion. RT has experience conducting individual interviews for research purposes.

Procedure

Participants viewing the study advertisement were asked to email the researcher if they were interested. The researcher replied to introduce them to the study, share the participant information sheet (PIS), and request that they complete an online (Qualtrics) questionnaire to check eligibility. The same researcher contacted eligible participants to check that they had read the PIS, answered any questions that they had, and asked if they were interested in participating in the study. Those expressing a desire to participate were emailed an online form that includes a few questions (details below) and an e-consent form (using Qualtrics) to sign and a link to book a 1-hour slot for the interview. The online form asked for demographic details such as gender and age, details about the intervention (name and intended purpose), and how long they used the intervention for. All interviews were audio recorded. At the end of the interview, participants were sent a thank you email and a £20 (US $25) voucher as compensation for their time. Interviews lasted on average 40 minutes. Recordings were automatically transcribed by Microsoft Teams and were verified for accuracy by RT. Once transcription was complete all recordings were deleted.

Analysis

Thematic Analysis Process

The 15 transcripts were uploaded onto NVivo (Lumivero, LLC) for analysis [24]. We used deductive thematic analysis to analyze the data. Thematic analysis is a process used to identify

XSL•FO
RenderX

patterns or themes within qualitative data to answer or explore a research question [25,26]. The analysis was conducted collaboratively by 2 researchers (RT and JY), who followed Braun and Clarke's [25] step-by-step guide to conduct a thematic analysis by familiarizing themselves with the 15 transcripts and coding the data. They then organized the codes based on relatedness, reviewing them, and defining and naming them as subthemes and themes [25]. The thematic analysis acknowledges "the researcher's reflective and thoughtful engagement with their data, and their reflexive and thoughtful engagement with the analytic process is essential" [25]. It recognizes the potential benefits of using multiple coders, such as achieving richer interpretations, however, it does not view this as a requirement [26]. Researchers are discouraged from attempting to provide accounts of "accurate" or "reliable" coding or pursuing consensus among multiple coders [26].

Given the research question and acknowledging that participants' experiences and perspectives on safety may vary in this study, the analysis was used to reflect the range of experiences of participants and highlight how these might differ, rather than attempting to merge these experiences into a single, unified interpretation [25]. Once the results of the study were ready, they were shared with all 15 participants to review and ensure that they were representative of their experiences. Three participants responded and said that they agreed with the results and did not offer any additional insights or suggest any alterations.

Researcher Reflexivity

A critical realist epistemology was adopted for this study, where the researchers aimed to explore participants' subjective experiences, acknowledge them as "real," and recognize the researchers' inability to fully access that reality [25]. The researchers were aware of their reflexivity [27]; at the time of this study, they all worked on a separate clinical trial that aimed to assess the efficacy and safety of a specific DMHI. The first author (RT) was completing her PhD on the safety of DMHIs. This study was 1 of 4 separate pieces of work for the PhD (other work comprising a systematic review, a methodology paper, and an experts' consensus study). This study was not directly related to or in any way part of the clinical trial that researchers were working on. As a team, the researchers were invested in learning how users/patients experience risks, react to them, and what risks matter to them in order to contribute to the field and improve their approach to safety.

Ethical Approval

Ethical Clearance was provided for this study by the King's College London (reference number LRS/DP-22/23-35403).

Results

Overview

The analysis of the data using deductive thematic analysis led to 3 major themes:

- Nonresponse: A Concern, a Risk, and How Users Mitigate It
- Symptom Deterioration and Its Management
- Concerns Around Data Privacy and How to Mitigate Them

Theme 1: Nonresponse: A Concern, a Risk, and How Users Mitigate It

Assessing the Effectiveness of DMHIs

Under this theme, participants spoke about their concerns regarding the DMHI being ineffective, experiencing ineffectiveness/nonresponse as a risk, and the methods they used to assess whether a DMHI was safe and effective.

Concerns Around Nonresponse

Users of DMHIs were concerned about the potential ineffectiveness of these interventions. Will these interventions be able to help them? Are these interventions evidence-based? One important area of concern is illustrated as follows:

> *Umm, I was concerned with like how helpful it would actually be, being that it is an online thing and like I'm not actually talking to a person you know...I was even actually concerned when I started the first session, whether the program would have lasting effects on actually helping me or supporting my mental health.* [Participant 6, used a DMHI for depression for 2.5 months]

It is likely that participants were doubtful about their interventions' effectiveness because they were struggling and in such emotional pain that they were unsure how a technology with no human could alleviate their pain and improve their mood.

Nonresponse as a Risk

Participants also spoke about the risk of the DMHI being unhelpful and ineffective, and how that at times led to further frustration, deterioration, and self-blame, for example:

> *I found it ineffective, if I'm honest. I think it was the set of six or eight weeks...I remember that the second week I burst into tears. It just felt so pointless. I can't remember what set me off, but it just felt so pointless.* [Participant 9 used a DMHI for an eating disorder for 1.5 months]

Another participant explained how the ineffectiveness of the DMHI led to a deterioration in her symptoms and feelings of isolation and self-blame. She said:

> *So, it's sort of added to that frustration when it was making the situation worse...So now, I felt like it would lead me to do sort of negative coping strategies...I'd be like get angry and irritable with people, or I'd go and overeat....it definitely sort of furthered the thoughts that there was like no one to help me...It made low moments even worse...It almost triggered sort of thoughts of like, oh, something's wrong with me. Why can't the program help me?* [Participant 14 used a DMHI for post-natal depression for 24 months]

Users' Method for Assessing Safety and Effectiveness

Users used 2 main methods to assess whether they thought a DMHI was safe and effective: (1) social proofing, which refers

XSL•FO
RenderX

to the tendency to follow the behavior of others as a guide for one's own actions [28]; and (2) assessing the contribution of experts or a trusted body. Some participants opened up about finding it difficult to assess the safety of a DMHI, and not knowing how to do that.

Participants wanted to know that professionals were involved in the development of the DMHI and that scientific research has been conducted to assess it, saying:

> *I suppose I would want to know how it was developed and in partnership with mental health practitioners...so I think with credibility, I guess things like whether they have worked with the university or with kind of recognized academics and done any kind of scientific research and rather than just user testing.* [Participant 11 used a DMHI for insomnia for 1.5 months]

In this participant's case in particular, knowing that the intervention was evidence-based was very important because she had struggled with insomnia for more than a decade and had tried many things (face-to-face therapy and medication) that were not helpful for her. Users also relied on other users' experiences. They checked reviews, ratings, and social media groups to find out more about the DMHI. The following quote is an example of this:

> *Uh, I have this habit of looking up these things online, so I would look up reviews of the app online. Maybe even check out any details about how helpful it has been. If people have had good experiences, bad experiences, they've had negative experiences, what have they been about and how I could avoid them. I would maybe even like look at Facebook groups.* [Participant 6 used a DMHI for depression for 2.5 months]

Some participants were honest about not knowing how to check the safety of the DMHI, and needing the support of professionals to be able to do so, they said:

> *I mean, unless it was like referred by my GP or kind of, you know, promoted through kind of official channels like the NHS website or something, I don't know how I would even check that an app has all the right checks, and you know safeguarding approvals or whatever.* [Participant 13 used a DMHI for depression and anxiety for 2 months]

Other participants suggested that the DMHI needs to assess users' suitability, saying:

> *I guess before someone's able to access the app kind of going through, I don't know, some sort of risk assessment on like who would find it useful.* [Participant 7 used a DMHI for paranoia for 3 months]

Another participant thought that this could be achieved by the DMHI clearly stating its intended use and target population:

> *I think the app would need to be really explicit about the limitations and sort of say up front like this is not for severe mental health issues or this is for maintenance.* [Participant 14 used a DMHI for post-natal depression for 24 months]

Theme 2: Symptom Deterioration and Its Management

Addressing Symptom Deterioration in DMHIs

Under this theme, participants spoke about experiencing symptom deterioration. Participants also gave their feedback on one of the methods used to support them when experiencing or struggling to cope (referral to other services and crisis support) and their suggestions on how deterioration can be managed.

Deterioration of Symptoms

Participants talked about how using a DMHI and dealing with their mental health struggles at times led to a deterioration in their mental health symptoms because it made them think about things that were upsetting. One participant who was struggling with sleep said:

> *I did find thinking more closely about my trouble with sleep did initially make me more anxious about sleep and made it harder to sleep. So, like the kind of tracking and then realizing that actually that was a really bad night...sometimes makes it harder to sleep the next night, by bringing it to the forefront.* [Participant 11 used a DMHI for insomnia for 1.5 months]

It is important to note that this participant found her intervention effective in helping her manage her insomnia and improve her sleep quality. This aligns with recent findings from an expert consensus study, which concluded that symptom deterioration is not a safety concern of DMHIs but rather a normal part of therapy [11].

In some cases, the inflexibility of the predetermined content in the DMHI and its inability to cater specifically to each user's needs (ie, lack of personalization) meant that the DMHI was unable to relate to users' emotional state and could lead to a deterioration in symptoms, for example:

> *So sometimes the AI (chatbot) like would give me suggestions that didn't really fit my situation. So, I'm like, you know what? Forget it. I'm not even going to do it, and I would feel worse afterwards because I wanted to express it. And then I'm just sitting here typing things, and it's not helping. It wasn't really built to recognize that CBT isn't effective for certain situations.* [Participant 14 used a DMHI for post-natal depression for 24 months]

Managing Deterioration

As deterioration is the most common negative effect of DMHIs [1], the authors asked participants how they think DMHIs could support them through it. Some participants said that normalizing deterioration would be very helpful. Participant 1 explained how that could be done in a hybrid model:

> *I think one way to support...is to have a video call therapy session. With an agent for example to make me understand that these things are normal. So, at that I would be reassured that I'm getting back to*

XSL•FO
RenderX

> *normal.* [Participant 1 used a DMHI for depression for 36 months]

Another participant explained how that could be done in a nonhybrid model:

> *Uh, maybe you know if the app had a mood tracker. The algorithm could check if you're feeling low right after therapy. It might just send them a message "Hey, if you're feeling down, you just check, you just had therapy. This could be normal" and that would be like, uh, fair enough.* [Participant 5 used a DMHI for depression and anxiety for 1 month]

Another participant suggested using regular check-ins to detect deterioration and provide users with support accordingly, saying:

> *Umm, I think the biggest thing is regular check-ins. That's quite an important thing about how they're finding it and what particularly is so difficult, maybe even not slowing it down, but having a bit of flexibility around kind of OK, you found this section of the app pretty difficult.* [Participant 10 used a DMHI for self-harm for 1.5 months]

Signposting to Other Services for Support

Most participants (11/15, 73%) were provided with emergency numbers, within the DMHI, to call in case they felt that they could not cope and needed further support. Some participants shared that they found this support helpful:

> *They gave me the contact numbers of like mind and Samaritans in case I needed urgent help. I'm using an online service. If I did need help, I could contact these services, which is actually really helpful because once or twice when I really felt like I was troubled at night, this did come in handy.* [Participant 6 used a DMHI for depression for 2.5 months]

However, other participants felt that the signposting to crisis support within the DMHI was tokenistic, sharing:

> *It's hard to feel like it's a genuine thing. It feels almost like a boilerplate that they put in every conversation. It doesn't really feel like there's thought going into it like "Ohh I recognize that the program can't help you with this. This would be better for a psychiatrist.* [Participant 14 used a DMHI for post-natal depression for 24 months]

Further analysis of these data highlighted that the participants who had a positive experience with the DMHI (ie, found it helpful) experienced the referral to crisis support information positively. By contrast, participants who found the DMHI unhelpful and were frustrated with it found the crisis support information unhelpful and ingenuine. Thus, users' relationship with the DMHI and their feelings toward it informed how they felt about being referred to other services. Users who found the DMHI helpful were likely to view the information about other services as a helpful bonus, whereas those who found their DMHI unhelpful were likely to doubt its genuine concern for them and thus viewed such referrals as a mere box-ticking exercise.

Theme 3: Concerns Around Data Privacy and How to Mitigate Them

Participant Perspectives on Data Privacy

Under this theme, participants spoke about their concerns around data privacy and their suggestions on how to help ease these concerns.

Concerns Around Confidentiality and Data Privacy

Participants were concerned about their data. Was the DMHI confidential? If not, who is it sharing their data with? Participants talked about their concern that the data might be shared with their health care team without their consent:

> *I had concerns about the information that I was putting in, and the kind of data that might have been collected. It's kind of that worry that what you're writing isn't actually confidential or you know that it could go back to someone else. I think one of the biggest worries for me was that what I was inputting in the app might have been given to my psychologist or somehow, you know, connected with the NHS or something like that.* [Participant 10 used a DMHI for self-harm for 1.5 months]

It is worth noting that, given the sensitivity of participant 10's struggle with self-harm, it is understandable why they were particularly concerned about their data being shared—even with their health care provider.

A User-Friendly Data Policy

Participants expressed their frustration with the vagueness and complexity of how DMHIs present their data protection policy and had suggestions on how that can be improved. Some thought that DMHIs should make key data protection information available in a simpler and more readily accessible format:

> *Maybe just share more information...like make it clear what you do to protect users' data. I don't want to have to go through, you know, all of your Terms and conditions, privacy policies and things like that to find out what it is. I mean, nobody's actually going to do that. I would never actually do it, so it would be helpful if it was just clearly mentioned somewhere.* [Participant 6 used a DMHI for depression for 2.5 months]

Another participant said that all they wanted from a DMHI is to be honest about what data they are storing and why:

> *They could kind of like emphasize maybe that your data isn't stored under an identifiable name that leads to you. Or say we are storing the data, but it is confidential and we're doing it to help more people. That's all...just being open about it.* [Participant 8 used a DMHI for an eating disorder for 3 months]

It is worth noting that the data from this study did not show any association between duration of use and participants' experiences. Participants had mixed responses to the DMHI when they used it for a short period (1.5 months, eg, participants 10 and 11), and those who used the DMHI for a long period (24 months; eg, participant 14) did not necessarily have a good

XSL•FO
RenderX

experience. However, Table 1 does show that the DMHIs that were targeting low mood–related areas were used for the longest period, such as depression (on average 19.25 months), self-harm (on average 12.75 months), and postnatal depression (on average 24 months). It is also noted that self-administered interventions were used for longer periods compared with hybrid interventions (on average 9.33 months vs 1.66 months). This is expected given that self-administered interventions require fewer resources and clinician time.

Discussion

Advancing Knowledge on DMHI Safety

This research was successful in exploring and understanding users' experiences, views, concerns, and suggestions regarding the safety of DMHIs. Until now, such findings have been absent from the literature. These findings contribute to advancing the field of digital mental health safety by providing valuable evidence of the viewpoints and experiences of its target population.

Principal Findings

Overview

The main findings of this study are presented in Table 2 (also see Figure 1) using user-friendly language and in the form of recommendations.

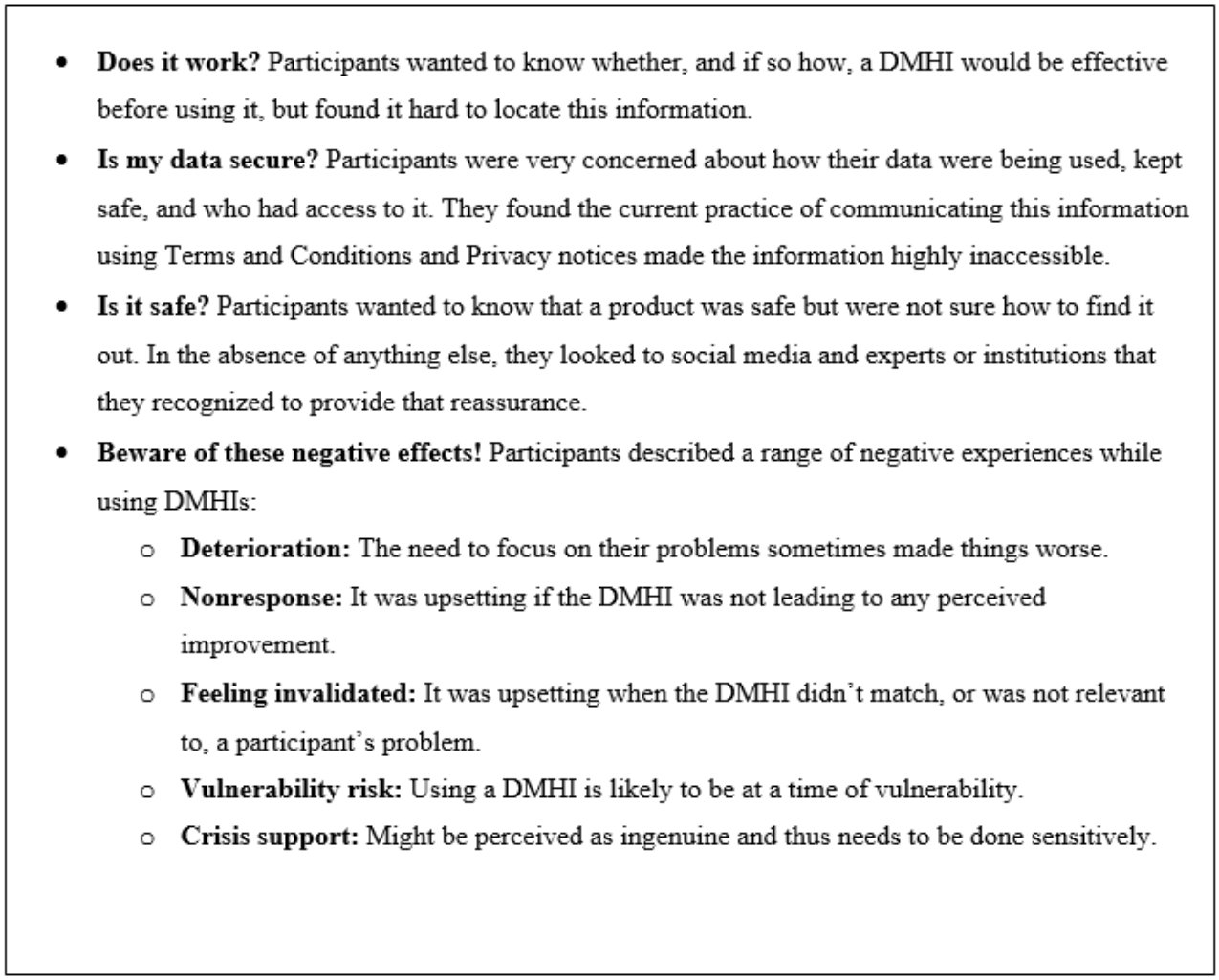
Table 2. User-informed recommendations to improve DMHI[a] safety.

Recommendations	Description
1. Provide 'easy access' versions of key information	Ensure that DMHI product manufacture and approval include a requirement to provide readily accessible, easy-to-read lay summaries of key information. At a minimum, these should cover (1) evidence of effectiveness; (2) data usage, security measures, and access (ie, who can access the data); and (3) potential negative effects.
2. Use 'Approved by...' badges	Introduce a sectorwide, widely recognized, branded badge to provide top-level reassurance of the quality and safety of any DMHIs bearing that badge.
3. Acknowledge the lack of personalization	DMHIs should flag to users their inability to be fully personalized and adaptable to an individual user's needs to mitigate feelings of invalidation and disappointment.
4. Anticipate and support deterioration	Before using a product, users should be alerted to possible mood or symptom deterioration, given normalizing information, and signposted to relevant support to help mitigate these effects should they occur.
5. Provide real-time feedback	DMHIs should internally track users' progress and provide feedback on whether they are benefiting as expected. Where there is no individual benefit, despite the appropriate use of the product, users should be automatically advised to seek alternative support.
6. Provide genuine crisis support	Content should include an acknowledgment of the DMHIs' limitations and a summary of each crisis service's support. It is recommended to consult the target population on wording to ensure genuine concern for users is communicated.
7. Responsibly manage access	DMHIs should incorporate an assessment of suitability focusing on risk levels and the appropriateness of the intervention for users' specific mental health conditions and severity. This could be done with a simple set of initial built-in questions which output a recommendation to use, or not use, the product based on the user's response.

[a]DMHI: digital mental health intervention.

XSL•FO
RenderX

Figure 1. Key findings (in lay language). DMHI: digital mental health intervention.



Users' Concerns Around the Safety of DMHIs

In this study, users expressed 2 primary concerns regarding the safety of DMHIs: (1) whether a product would be effective and (2) whether their data would be secure and confidential. Evidence around the effectiveness of DMHIs is usually disseminated in academic peer-reviewed articles. This poses an accessibility challenge for the typical users, which is only slightly tempered by recent initiatives toward making open-access publications the norm in academia. It has been documented that uncertainty around the effectiveness of a DMHI is a key barrier to its use [15]. To address users' concerns about effectiveness, it is crucial to translate scientific findings into lay language and publish them in user-friendly formats to improve accessibility (recommendation 1). This dovetails well with the increasing requirements placed upon academics to evidence the impact of their work.

Regarding users' second major concern, data safety and confidentiality, participants in this study suggested that DMHIs provide their users with clear and concise information on how their data are used, stored, and who has access to it. Recognizing that users often overlook traditional lengthy privacy policies and terms and conditions [29], DMHIs could supplement these by providing users with a layperson's summary of how their data are being used, kept safe, and who has access to them (recommendation 1). Clear and transparent communication about data and privacy would help build trust, a key component of the therapeutic relationship [30]. The evidence to date suggests that the digital therapeutic alliance is both relevant and important in DMHIs [19,20].

Assessing a DMHI's Safety From Users' Perspective

It is important to understand how users assess the safety of DMHIs, as this will inform professionals about where users look for safety information and so where and how best to provide it. Participants in this study described methods for assessing the safety of a DMHI that reflected social proof, a concept first attributed to Robert Cialdini [31,32]. Social proof refers to situations where people use opinions and information from others similar to themselves to influence personal choices, decisions, and behaviors, especially if uncertain [31,32]. In the present context, this involved reading online product reviews on websites and mobile app stores. Other participants expressed that they did not know how to assess the safety of a DMHI. In a different qualitative study, medical students shared a similar experience, expressing difficulty in identifying which DMHIs were evidence-based [15]. These students also criticized the lack of guidance available for users on how to find evidence-based DMHIs [15].

XSL•FO
RenderX

There are already regulatory bodies equipped and responsible for assessing the efficacy and safety of DMHIs and making recommendations for use. In the United Kingdom, this includes NICE (The National Institute for Health and Care Excellence) and the MHRA (Medicines and Healthcare Products Regulatory Agency). The NHS (National Health Service) previously had a health app store called the "NHS Health Apps Library," but it was decommissioned in 2021 due to the increasing complexity of maintaining the library and ensuring the safety and effectiveness of the listed apps. An alternative to this has emerged through the Health App Library provided by the Organization for the Review of Care and Health Apps (ORCHA), in collaboration with health providers such as NHS Trusts. This library offers a list of mobile apps that have been reviewed by ORCHA for effectiveness and safety. One effective way to demarcate a product's safety and efficacy status would be for regulators to introduce a branded, recognizable stamp or badge to be displayed by products achieving prespecified minimum safety and efficacy requirements (recommendation 2). What those requirements should be, however, would entail significant additional research to achieve a consensus across industry, regulatory bodies, academia, developers, and users. Nevertheless, this approach would capitalize on users' existing tendency to seek safety information via the product's mobile app page or website and is therefore likely to ultimately be an effective means of disseminating key information on which end users can base their decisions. A notable advancement in this direction is Google's revision of their app store's health policy, which mandates that starting from May 31, 2024, all health apps posted on their app store must prove compliance with relevant laws and regulations (privacy policy, ethics approval, and certification when required) [33].

Risks Experienced by DMHIs' Users and Their Suggestions on How to Mitigate Them

Participants in this study spoke about 3 risks they had experienced as a result of using a DMHI: feeling invalidated by the DMHI, deterioration in their symptoms, and nonresponse. Users noted that, unlike a human therapist, the DMHIs' inability to be fully responsive, personalized, and adaptable to each user's needs left them feeling invalidated and unheard. Such experiences undermine the therapeutic relationship. There is evidence that personalization in a DMHI fosters therapeutic alliance [20], and thus the lack of it is likely to undermine this alliance. The weaker the digital therapeutic relationship, the more this is likely to undermine the effectiveness of the DMHI [30]. In a qualitative study involving Australian psychologists and their experiences with DMHIs, the psychologists expressed the view that DMHIs are inferior to face-to-face therapy due to their limited capacity for personalization [14]. In another qualitative study, medical students made the same comparison between DMHIs' and health professionals' ability to provide personalized therapy [15]. This is a limitation of current technologies, which might change with the future advances of artificial intelligence. However, for now, it is important for DMHIs to acknowledge and communicate this limitation to users to mitigate feelings of invalidation (recommendation 3).

Participants suggested that DMHIs can support users experiencing deterioration by informing them about the possibility, normalizing it, and providing pathways to relevant support when it occurs (recommendation 4). These suggestions align with those made by digital mental health professionals in a recent consensus statement [11]. Other studies have suggested implementing an automated process within DMHIs to monitor and flag when participants' symptoms deteriorate beyond a predefined threshold [5]. Although that threshold would be subject to individual conditions and clinical opinion within any specific context, a useful starting point would be to adopt the clinical "rule of thumb" that considers a 20% change in symptoms as a meaningful variation [34].

"Nonresponse" is a documented potential side effect of DMHIs [1,12]. It was interesting to see how nonresponse from users' perspective was almost a compound negative effect that led to a cascade of unwanted effects including feelings of frustration, hopelessness, deterioration, isolation, and self-blame. One way to address this would be for DMHIs to track users' clinical outcomes, identify those experiencing nonresponse, and provide them with targeted pathways to further support as a way of mitigating these possible adverse consequences (recommendation 5).

The Risk Mitigation Methods Experienced by DMHIs' Users

When discussing how DMHIs mitigate risks and safeguard users, users highlighted the importance of signposting to other sources of support. The majority of users (11/15) were provided with details of emergency numbers and other mental health services; however, not all users found these helpful. Some users felt that this was a checkbox exercise that DMHIs needed to complete and that the crisis support provided was ingenuine. Users of DMHIs may feel this way because of their experiences with mental health services, and because most mental health services/interventions tend to provide crisis support information. Signposting to a different service needs to be done delicately to ensure that the user feels cared for. For that, DMHIs need to be careful about how they present crisis support details. Including an acknowledgment of the DMHI's limitations, a concern for the user, and a description of the support that each crisis service provides might help make users feel that the intention to provide support is more genuine (recommendation 6). Soliciting user input on how such information is phrased and presented would also be of benefit (recommendation 6).

Only 1 user among our sample of 15 was informed of the potential side effects of the DMHI that they were using. A recent systematic review on the safety of DMHIs found that only one-third of the interventions informed their users of their adverse events or possible side effects [1]. The even lower level of side effect awareness in our study might be due to the commonplace practice of embedding side effect information within inaccessible or often unread documentation (eg, terms and conditions or instructions for use) [6]. Indeed, as reported above, our participants told us that they found these documents particularly impenetrable, suggesting some may have missed out on important side effect information. To ensure user safety, it is important to improve the visibility and accessibility of side effect information by adopting new methods of communication (recommendation 1). This might include a digital equivalent of

XSL•FO
RenderX

listing possible side effects on medication labels. Existing regulations already require a digital product label to be displayed within the product itself and this label includes "Cautions" and "Warnings." It would be a relatively simple matter to add a section "Possible Side Effects" as a further requirement.

However, simply adding information is unlikely, on its own, to meet users' needs as identified by our study. In addition (and as already discussed earlier), serious consideration should be given to adopting the practice of requiring "easy read" or "lay summary" versions of key information which is provided alongside full and formal versions. This is now standard practice in domains such as academia, governmental, and other public sector organizations. Our study suggests that, as a minimum, this should apply to information on data security and side effects.

Finally, to minimize this risk, our sample of participants/users recommended that DMHIs should evaluate each user's suitability, with a focus on assessing risk levels and determining the appropriateness of the intervention for their specific mental health condition and severity. These assessments can be included as a standard procedure before onboarding a user onto a DMHI, similar to how patients are screened before they receive face-to-face therapy (recommendation 7).

Limitations

There are a few limitations to this study. The participants in this study were mostly female (12/15, 80%). The DMHIs used by participants were mostly self-administered (12/15, 80%), and thus results might be biased by their experiences. Additionally, recruiting for a study to explore users' perspectives on the safety of DMHIs might have attracted individuals who have experienced such issues. It is important to be aware of how the sample of participants in this study could have shaped the results. This is expected in qualitative studies, which aim to explore and understand the experiences and opinions of a sample of the population [35].

Conclusions

The results of this study led to 7 user-informed recommendations on how the safety of DMHIs can be improved. These recommendations arose from users' experiences and suggestions. The key findings (Figure 1) and recommendations of this paper could improve the safety of DMHIs, enhance users' experience, address some of their concerns, and foster a more trusting therapeutic relationship between the user and the DMHI.

Acknowledgments

We are grateful to all our participants for taking the time to participate in our study and share with us their valuable experiences and opinions. We also would like to express our gratitude to the National Institute for Health and Care Research (NIHR) Biomedical Research Centre hosted at South London and Maudsley NHS Foundation Trust in partnership with King's College London. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, the Department of Health and Social Care, the ESRC, or King's College London.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Topic guide.
[DOCX File , 24 KB - humanfactors_v12i1e62974_app1.docx]

References

1. Taher R, Hsu C, Hampshire C, Fialho C, Heaysman C, Stahl D, et al. The safety of digital mental health interventions: systematic review and recommendations. JMIR Ment Health 2023 Oct 09;10:e47433 [FREE Full text] [doi: 10.2196/47433] [Medline: 37812471]
2. Brown P, Prest B, Miles P, Rossi V. The development of National Safety and Quality Digital Mental Health Standards. Australas Psychiatry 2022 Apr 27;30(2):154-157. [doi: 10.1177/10398562211042361] [Medline: 34569321]
3. van Orden ML, Kraaijeveld JC, Spijker AT, Silven AV, Bonten TN, Chavannes NH, et al. Preliminary effects of a digital mental health intervention for depression and anxiety. Clinical eHealth 2022 Dec;5:44-51. [doi: 10.1016/j.ceh.2022.06.002]
4. Kambeitz-Ilankovic L, Rzayeva U, Völkel L, Wenzel J, Weiske J, Jessen F, et al. A systematic review of digital and face-to-face cognitive behavioral therapy for depression. NPJ Digit Med 2022 Sep 15;5(1):144-144 [FREE Full text] [doi: 10.1038/s41746-022-00677-8] [Medline: 36109583]
5. Gómez Bergin AD, Valentine AZ, Rennick-Egglestone S, Slade M, Hollis C, Hall CL. Identifying and Categorizing Adverse Events in Trials of Digital Mental Health Interventions: Narrative Scoping Review of Trials in the International Standard Randomized Controlled Trial Number Registry. JMIR Ment Health 2023 Feb 22;10:e42501. [doi: 10.2196/42501]
6. Martinez-Martin N, Kreitmair K. Ethical issues for direct-to-consumer digital psychotherapy apps: addressing accountability, data protection, and consent. JMIR Ment Health 2018 Apr 23;5(2):e32 [FREE Full text] [doi: 10.2196/mental.9423] [Medline: 29685865]

XSL•FO
RenderX

7. Berk M, Parker G. The elephant on the couch: side-effects of psychotherapy. Aust N Z J Psychiatry 2009 Sep 01;43(9):787-794. [doi: 10.1080/00048670903107559] [Medline: 19670051]
8. Batterham PJ, Christensen H, Calear AL, Werner-Seidler A, Kazan D. Rates and predictors of deterioration in a trial of internet-delivered cognitive behavioral therapy for reducing suicidal thoughts. Archives of Suicide Research 2020 Nov 20;26(2):937-947. [doi: 10.1080/13811118.2020.1848671] [Medline: 33215554]
9. Linden M. How to define, find and classify side effects in psychotherapy: from unwanted events to adverse treatment reactions. Clin Psychol Psychother 2013 Jan 18;20(4):286-296. [doi: 10.1002/cpp.1765] [Medline: 22253218]
10. Terry NP, Gunter TD. Regulating mobile mental health apps. Behav Sci Law 2018 Mar 16;36(2):136-144 [FREE Full text] [doi: 10.1002/bsl.2339] [Medline: 29659069]
11. Taher R, Bhanushali P, Allan S, Alvarez-Jimenez M, Bolton H, Dennison L, et al. Bridging the gap from medical to psychological safety assessment: consensus study in a digital mental health context. BJPsych Open 2024 Jun 03;10(4):e126-e126 [FREE Full text] [doi: 10.1192/bjo.2024.713] [Medline: 38828683]
12. Rozental A, Andersson G, Boettcher J, Ebert DD, Cuijpers P, Knaevelsrud C, et al. Consensus statement on defining and measuring negative effects of internet interventions. Internet Interventions 2014 Mar;1(1):12-19. [doi: 10.1016/j.invent.2014.02.001]
13. Bradstreet S, Allan S, Gumley A. Adverse event monitoring in mHealth for psychosis interventions provides an important opportunity for learning. J Ment Health 2019 Oct 26;28(5):461-466 [FREE Full text] [doi: 10.1080/09638237.2019.1630727] [Medline: 31240970]
14. Scott S, Knott V, Finlay-Jones AL, Mancini VO. Australian psychologists experiences with digital mental health: a qualitative investigation. J Technol Behav Sci 2022 Aug 16;8(4):1-11 [FREE Full text] [doi: 10.1007/s41347-022-00271-5] [Medline: 35991293]
15. Dederichs M, Weber J, Pischke CR, Angerer P, Apolinário-Hagen J. Exploring medical students' views on digital mental health interventions: a qualitative study. Internet Interv 2021 Sep;25:100398 [FREE Full text] [doi: 10.1016/j.invent.2021.100398] [Medline: 34026567]
16. Allan S, Beedie S, McLeod HJ, Farhall J, Gleeson J, Bradstreet S, et al. Using EMPOWER in daily life: a qualitative investigation of implementation experiences. BMC Psychiatry 2023 Aug 17;23(1):597-597 [FREE Full text] [doi: 10.1186/s12888-023-05096-x] [Medline: 37592231]
17. Arnold C, Williams A, Thomas N. Engaging with a web-based psychosocial intervention for psychosis: qualitative study of user experiences. JMIR Ment Health 2020 Jun 19;7(6):e16730 [FREE Full text] [doi: 10.2196/16730] [Medline: 32558659]
18. Berry N, Lobban F, Bucci S. A qualitative exploration of service user views about using digital health interventions for self-management in severe mental health problems. BMC Psychiatry 2019 Jan 21;19(1):35-13 [FREE Full text] [doi: 10.1186/s12888-018-1979-1] [Medline: 30665384]
19. Tong F, Lederman R, D'Alfonso S, Berry K, Bucci S. Conceptualizing the digital therapeutic alliance in the context of fully automated mental health apps: a thematic analysis. Clin Psychol Psychother 2023 Apr 11;30(5):998-1012. [doi: 10.1002/cpp.2851] [Medline: 37042076]
20. Tong F, Lederman R, D'Alfonso S, Berry K, Bucci S. Digital therapeutic alliance with fully automated mental health smartphone apps: a narrative review. Front Psychiatry 2022 Jun 22;13:819623-819612 [FREE Full text] [doi: 10.3389/fpsyt.2022.819623] [Medline: 35815030]
21. Braun V, Clarke V. Successful Qualitative Research A Practical Guide for Beginners. London, UK: SAGE Publications Ltd; Mar 01, 2013.
22. Moltu C, Stefansen J, Svisdahl M, Veseth M. Negotiating the coresearcher mandate - service users' experiences of doing collaborative research on mental health. Disabil Rehabil 2012;34(19):1608-1616. [doi: 10.3109/09638288.2012.656792] [Medline: 22489612]
23. Braun V, Clarke V. To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. Qualitative Research in Sport, Exercise and Health 2019 Dec 26;13(2):201-216. [doi: 10.1080/2159676x.2019.1704846]
24. NVivo (version12). QSR International Pty Ltd. URL: https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home [accessed 2025-01-31]
25. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
26. Byrne D. A worked example of Braun and Clarke’s approach to reflexive thematic analysis. Qual Quant 2021 Jun 26;56(3):1391-1412. [doi: 10.1007/s11135-021-01182-y]
27. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qualitative Research in Sport, Exercise and Health 2019 Jun 13;11(4):589-597. [doi: 10.1080/2159676x.2019.1628806]
28. Scott W, Barden J. Social Proof. London, UK: Routledge; Apr 22, 2022:36-64.
29. Wykes T, Lipshitz J, Schueller SM. Towards the design of ethical standards related to digital mental health and all its applications. Curr Treat Options Psych 2019 Jul 5;6(3):232-242. [doi: 10.1007/s40501-019-00180-0]

XSL•FO
RenderX

30. Boucher EM, Harake NR, Ward HE, Stoeckl SE, Vargas J, Minkel J, et al. Artificially intelligent chatbots in digital mental health interventions: a review. Expert Rev Med Devices 2021 Dec 31;18(sup1):37-49 [FREE Full text] [doi: 10.1080/17434440.2021.2013200] [Medline: 34872429]
31. Cialdini RB. Influence, New and Expanded: The Psychology of Persuasion. New York, NY: HarperCollins; Oct 27, 2013.
32. Cialdini R. Influence: Science and Practice. London, UK: Pearson Publishing; 2009.
33. Google LLC. Health content and services. Google LLC. URL: https://support.google.com/googleplay/android-developer/answer/13996823?sjid=11315650021394300636-EU [accessed 2025-01-31]
34. Hobbs C, Lewis G, Dowrick C, Kounali D, Peters TJ, Lewis G. Comparison between self-administered depression questionnaires and patients' own views of changes in their mood: a prospective cohort study in primary care. Psychol Med 2020 Jan 20;51(5):853-860. [doi: 10.1017/s0033291719003878]
35. Leung L. Validity, reliability, and generalizability in qualitative research. J Family Med Prim Care 2015;4(3):324-327 [FREE Full text] [doi: 10.4103/2249-4863.161306] [Medline: 26288766]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
DMHI: digital mental health intervention
MHRA: Medicines and Healthcare Products Regulatory Agency
NHS: National Health Service
NICE: The National Institute for Health and Care Excellence
ORCHA: Organization for the Review of Care and Health Apps
PIS: participant information sheet

Edited by A Kushniruk; submitted 06.06.24; peer-reviewed by I Wilson, L McCann; comments to author 14.10.24; revised version received 07.11.24; accepted 24.11.24; published 07.02.25.

Please cite as:
Taher R, Stahl D, Shergill S, Yiend J
The Safety of Digital Mental Health Interventions: Findings and Recommendations From a Qualitative Study Exploring Users' Experiences, Concerns, and Suggestions
JMIR Hum Factors 2025;12:e62974
URL: https://humanfactors.jmir.org/2025/1/e62974
doi:10.2196/62974
PMID:

©Rayan Taher, Daniel Stahl, Sukhi Shergill, Jenny Yiend. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 07.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Provider Perspectives on the Use of Mental Health Apps, and the BritePath App in Particular, With Adolescents at Risk for Suicidal Behavior: Qualitative Study

Frances Lynch[1*], PhD, MSPH; Julie Cavese[1,2*], PsyD, MA; Lucy Fulton[1*], AA; Nancy Vuckovic[3*], PhD; David Brent[4*], MD, MS

[1]Kaiser Permanente Center for Health Research, Portland, OR, United States

[2]Cavese Counseling, Portland, OR, United States

[3]Vuckovic Consulting, Vuckovic Consulting, Portland, OR, United States

[4]Department of Psychiatry, University of Pittsburgh Medical Center, Pittsburgh, PA, United States

[*]all authors contributed equally

Corresponding Author:
Frances Lynch, PhD, MSPH
Kaiser Permanente Center for Health Research
3800 N Interstate Ave
Portland, OR, 97227
United States
Phone: 1 503 335 2400
Fax: 1 503 335 2428
Email: frances.lynch@kpchr.org

Abstract

Background: Many youth with significant mental health concerns face limited access to mental health services. Digital programs, such as mobile apps designed to address mental health issues, have the potential to expand access to strategies for managing these conditions. However, few mental health apps are specifically designed for youth experiencing severe concerns, such as suicidal ideation. BritePath is a new app developed to enhance communication and interaction between providers and youth at risk for suicidal behavior.

Objective: This study aims to explore health care providers' opinions and concerns regarding the use of mental health apps for youth at significant risk of suicidal behavior.

Methods: We conducted individual semistructured interviews with 17 providers across 7 states. Interviews were conducted via video, recorded, and transcribed. Codes were developed using a team-based approach, with discrepancies resolved through team discussions.

Results: Most providers were aware of mental health apps in general and expressed interest in trying the BritePath app with patients experiencing depression, suicidality, or both. Analyses identified 4 key themes related to mental health apps: (1) almost all providers viewed mental health apps as an adjunct to, rather than a replacement for, psychotherapy visits; (2) most providers were concerned about the cost of apps and youth access to them; (3) providers noted the challenge of maintaining patient engagement with apps over time; and (4) providers were concerned about patient privacy, in terms of both data shared with app developers and data privacy within families. Analyses of providers' opinions specifically about the BritePath app identified 4 additional themes: (1) providers believed that access to safety plans within BritePath could be beneficial for youth at risk for suicidal behavior; (2) providers reported that BritePath's interactive features could enhance communication between providers and youth; (3) providers appreciated BritePath's flexibility and the ability for both youth and providers to tailor its content to individual needs; and (4) providers expressed concerns about integrating BritePath into clinical workflows within health systems.

Conclusions: The use of mental health apps is expanding, yet there is limited understanding of how to effectively integrate these tools into mental health treatment. Providers are increasingly referring patients to mental health apps, and most expressed interest in trying the BritePath app for patients with depression, suicidality, or both. However, providers also identified several concerns, particularly regarding privacy and safety.

XSL•FO
RenderX

(JMIR Hum Factors 2025;12:e64867) doi:10.2196/64867

KEYWORDS

depression; adolescent; suicidality; safety plan; mental health; apps; suicide

Introduction

Background

Adolescent suicidal behavior, suicide ideation, and depression are major public health problems that have increased significantly over the past 20 years [1-3] and have been exacerbated by the COVID-19 pandemic and its sequelae [4-8]. For example, between 2019 and 2021, the number of female high school students reporting that they had seriously considered attempting suicide increased by 6% [3]. These increases in suicidal ideation, behavior, and depression have disproportionately affected youth from historically minoritized racial and ethnic groups [2-4]. Once youth are identified as having a significant risk of suicidal behavior, they are typically treated in emergency departments or urgent care settings or hospitalized for stabilization before beginning outpatient treatment [9]. However, many youth remain at high risk of suicidal behavior even after acute treatment or stabilization in intensive settings. Managing youth with high levels of suicidal risk requires close communication between patients and clinical providers, such as psychiatrists or mental health therapists (referred to collectively as providers hereafter). Evidence-based care for these youth includes regular monitoring of depression symptoms, suicidal ideation, and suicidal behavior [9-11]. However, most health systems fall short of maintaining this level of close communication due to barriers such as lack of time and resources, as well as difficulties in staying in touch with at-risk youth after they leave intensive settings or acute treatment.

To address these concerns, researchers are developing new mobile apps aimed at improving both the efficiency of communication with youth at risk for suicidal behavior and shared decision-making between providers and adolescent patients. These tools are increasingly promoted to address mental health concerns in general and depression in particular [12-16]. Empirically based apps for depression range from cognitive behavioral self-help programs [15,16] to screening [16,17] or mood-tracking apps [18]. Recent research indicates that these tools are generally acceptable to providers [13,17]. However, there are several limitations to their use in clinical practice. Much of the evidence comes from surveys asking providers about the acceptability of apps [19] or studies that have not focused on specific concerns about particular apps [13]. Few studies have examined the use of apps within the context of mental health treatment [15]; instead, most have focused on patients' use of apps independent of providers [13]. To date, no prior research has examined the acceptability of apps designed to support shared decision-making between providers and youth at high risk for suicidal behavior. While some experts have raised potential ethical and safety concerns about using digital interventions with high-risk populations [19,20], improving connections to treatment after nonfatal suicidal behavior has the potential to enhance the treatment trajectory and long-term outcomes for these youth.

Study Goal

The aims of this study were to characterize providers' (1) opinions on the barriers and benefits of using apps in mental health care in general, (2) perspectives on the use of apps with adolescents with depression or suicidal thoughts and young adult patients, (3) interest in and willingness to use a recently developed app (BritePath) intended for use with adolescents with suicidal thoughts or young adult patients, and (4) barriers to implementing the BritePath app in routine clinical practice.

Methods

Implementation of BritePath: Benefits and Barriers

This qualitative study was conducted as part of the Center for Enhancing Treatment and Utilization for Depression and Emergent Suicidality (ETUDES Center), an ALACRITY Center funded by the National Institute of Mental Health (NIMH P50MH115838), aimed at helping providers better support youth experiencing severe depression, suicidality, or both. This study explores issues related to the use of mental health apps in general with this population and examines potential benefits and barriers to implementing an already developed app, BritePath, in health systems that have not participated in the research studies developing BritePath and are not affiliated with the ETUDES Center. To date, all information regarding the acceptability of the BritePath app has been gathered within the institutions that developed it.

The BritePath App

The BritePath app was developed to improve communication between providers and youth with depression or suicidal thoughts. It consists of 3 integrated components: (1) *Guide2BRITE*, an electronic guide for mental health providers that includes step-by-step instructions for onboarding a patient's safety plan into the app, as well as guidance on discussing key treatment components such as emotion regulation and distress tolerance skills; (2) the *BRITE app*, a personalized and interactive safety plan and self-monitoring tool for youth; and (3) the clinician dashboard, *BRITEBoard*, which allows providers to track youth's app use, distress levels, and treatment progress while facilitating communication and collaboration among mental health and primary care providers.

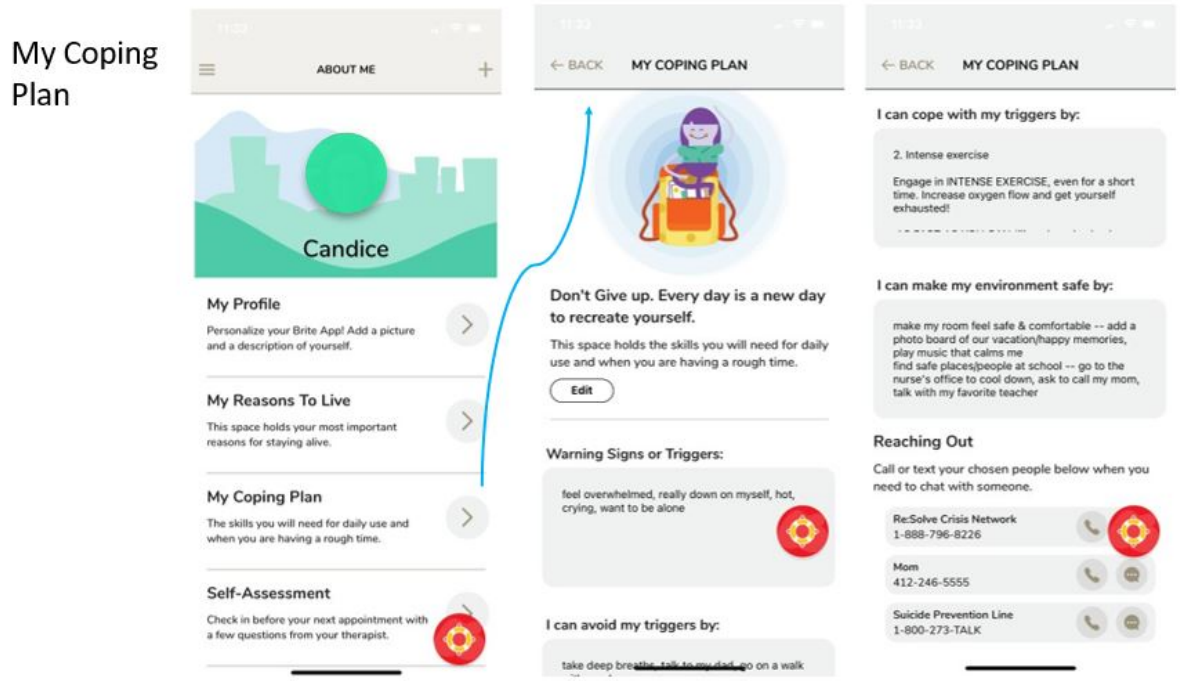
BritePath promotes self-monitoring and self-management through personalized strategies to avoid or cope with triggers for suicidal urges [21]. It is based on BRITE, a patient-facing safety planning app developed by researchers in psychiatry and psychology at the University of Pittsburgh and the University of Texas Southwestern Medical Center [21]. To our knowledge, BritePath is the first smartphone-based safety planning app to be tested in clinical trials involving adolescents with severe depression or suicidal thoughts. BritePath provides providers

XSL•FO
RenderX

with tools to onboard their adolescent and young adult patients to the app. This includes assisting youth in developing and personalizing a safety plan, incorporating distress tolerance and emotion regulation strategies, and monitoring symptomatic progress for at-risk youth in their care. Figure 1 provides a screenshot from the BritePath app for a mock patient.

BritePath was first tested for clinical effectiveness in a pilot randomized trial that evaluated BritePath in combination with a related intervention. The study found that the combined intervention reduced the rate of suicide attempts by 50% in the 6 months following hospital discharge [21]. This pilot study also suggested that more frequent use of BritePath increased reasons for living. However, the small sample size did not allow for comprehensive testing of BritePath alone. A more recent, larger randomized trial found that BritePath is as effective as usual care for youth hospitalized for suicidality. Additionally, youth who received BritePath were less likely to have a subsequent suicide attempt and had a longer time until a repeat attempt [22].

Figure 1. Example of the safety plan in the BritePath app.



Sampling and Recruitment

We identified a convenience sample of mental health and primary care providers who have cared for patients with mental health concerns. Providers were recruited through health systems participating in the Mental Health Research Network [23] and via snowball sampling [24,25], in which interviewed providers recommended additional potential participants. No specific level of digital experience or skills was required for participation in the study.

Ethical Approval

All study procedures were approved by the Institutional Review Board of the Kaiser Permanente Center for Health Research (KPCHR) at Kaiser Permanente Northwest (approval ID: MOD20050381-003), where the study was conducted. Interviews were conducted via KPCHR Teams, and digital recordings and transcripts were securely stored within the primary organization's firewall-protected file service.

Data Collection

Semistructured interviews were conducted remotely via Microsoft Teams. The interview guide (Multimedia Appendix 1) was developed by our medical anthropologist expert (NV) in coordination with the principal investigator (FL). Additionally, a qualitative investigator from the larger BritePath team reviewed the guide for content. The interviewer (FL) used a semistructured interview guide (Multimedia Appendix 1) to ensure consistency across interviews. The framework that guided the design of our qualitative interview guide focused on usability and concept testing. We examined whether providers from different backgrounds and training would consider using BritePath. The included questions aligned with several frameworks for adopting new technology, such as the Diffusion of Innovation theory [26,27]. The interviews assessed providers' opinions on using apps with patients and their impressions of the benefits and barriers of mental health apps in general. Participants were then shown a series of visuals (screenshots) depicting what a user would see while interacting with the BritePath app. These prompts ensured that participants had a clear and consistent understanding of the app's features and user interface. During this portion of the interview, participants were asked about their initial impressions of the BritePath app, as well as the barriers and potential benefits of using it in their practice. They were also given the opportunity to suggest features they would like to see in future app developments. Each interview lasted 30-40 minutes and was audio recorded and transcribed by an independent transcriptionist. Transcripts were entered into Atlas.ti [28], a qualitative analysis software program

XSL•FO
RenderX

used for coding and managing data. Atlas.ti was then used to generate reports for the analysis of interview transcripts.

Analysis

Thematic analysis, a method for analyzing patterns of meaning in a data set, was used to interpret the data by identifying themes and their interconnections [29-32]. Following established thematic analysis procedures [30], members of the research team (JC, LF, and NV) first familiarized themselves with the data and, together with FL, developed an initial code list consisting of a priori codes reflecting the interview questions. JC and LF coded the first interview using this initial code list and identified additional codes while reviewing the transcripts. The full team met to review coding consistency between reviewers and to evaluate the need for new, emergent codes. Modifications to the code list were made based on discussion and consensus. This process was repeated, with 2 team members coding 2 additional transcripts each. The full team met to discuss differences between coders for these additional transcripts and found minimal discrepancies in the codes applied. Upon confirming sufficient consistency in code usage and identifying no additional emerging codes, a master code list was finalized and used to code the remaining transcripts. After coding all transcripts, the full study team individually reviewed reports of coded interview segments to identify themes. The final themes were then derived through discussion and consensus. This formal, team-based analysis procedure was used to enhance the credibility and trustworthiness of the findings.

Results

Providers' Perspectives on Mental Health Apps and the BritePath App for Youth at Risk of Suicidal Behavior

We conducted semistructured interviews with a convenience sample of 17 providers from 7 states in the United States (Table 1). Our goal was to understand providers' opinions on the use of apps for youth with mental health concerns in general and the use of the BritePath app for youth at risk of suicidal behavior. First, we asked providers about their experiences using mental health apps with patients. Next, we gathered their opinions on a specific app, BritePath, designed to improve patient engagement and communication with providers for youth at risk of suicidal behavior. The interview guide is available in Multimedia Appendix 1.

Table 1. Provider demographics (N=17).

Characteristic	Values, n (%)
Age (years) group	
25-34	4 (23)
35-44	6 (35)
45 and over	7 (41)
Sex at birth	
Female	15 (88)
Male	2 (22)
Provider type	
Physician	2 (22)
Therapist/social worker	8 (47)
Psychologist	7 (41)
Type of setting	
Large health system	10 (22)
Smaller group practice	4 (47)
Individual practice	3 (41)

Providers' Opinions About the Use of Apps for Mental Health Concerns in General

Overview

All providers reported familiarity with mental health apps, and most stated that they had referred at least one patient to a mental health–focused app—for example, to help patients learn a skill such as relaxation. However, none of the providers reported regularly using any mental health app during treatment visits or in an interactive manner with patients. Each theme is discussed in more detail below.

Theme 1: Mental Health Apps Have the Potential to Be Valuable as an Adjunct to Therapy Visits

Providers reported that they felt mental health apps could add value to the therapeutic relationship as an adjunct or complement to face-to-face or video-based psychotherapy sessions. However, most did not believe that mental health apps could serve as a substitute for psychotherapy visits. Several providers noted that these apps can offer access to exercises for learning new skills or encourage the use of tools discussed in therapy, making them a useful complement to treatment sessions.

XSL•FO
RenderX

> *I think they can be really great adjuncts if you have the time to utilize those as a clinician. I am always interested for my patients of course for them doing things in the many hours outside of therapy, so I see you from 1:00 to 2:00 and then there are all these other hours that in theory you're doing the things that we were practicing in that one hour. If there are apps that can promote, engage, and promote the use of those strategies, whether it's thought records or sleep diaries or mood meters, things like that. And then they can bring that back in and we can look at it together and that facilitates the conversation, I think that would be great.* [P5]

> *And this way as a clinician I am teaching a skill, but I want it to be reinforced by this tool that mainly my client really loves but eventually is a way to train the brain to think a new and different way.* [P4]

Providers also noted that this approach to learning skills may be particularly useful for teenagers, who frequently engage with their phones.

> *I think they're a great compliment to therapy. I think they're really useful as a tool. I am not a huge fan of the app-based therapies you know. Like text therapy and stuff like that, I think it's not a good fit for everybody of course. But I think the apps themselves as far as tools and teaching and reinforcing skills I think are really helpful. And I would imagine for teenagers, I am assuming most of them are on their phones quite a bit. So that might be kind of a really easy transition or way to incorporate into something they're already doing. I think it would be cool especially for teens if there was a way to have some sort of feedback or back and forth with their therapist so they could actually share what they're doing, that would be neat. I think there's a lot of potential there for younger people.* [P14]

Although providers generally expressed support for mental health apps, most also raised several concerns. They described issues from both the provider and patient perspectives.

Theme 2: Cost/Access to Tool

The cost of purchasing an app was noted as a potential barrier. Providers suggested that they would only consider recommending apps with relatively low fees or a 1-time payment rather than ongoing costs.

> *Obviously free is best. Especially for teens, even young adults that are trying to make ends meet. If not then maybe something that's a one-time fee. If it was a subscription that would be harder for that population. I think that's something that older adults could possibly swing, but unless the parents are paying for it, it doesn't seem like that would be a good match.* [P13]

> *On the patient side yeah I think it would depend on how expensive the app would be. If they've got to pay $50/month that certainly could be prohibitive for folks.* [P5]

In addition, providers expressed concern that some youth, particularly those from low-income backgrounds, might lose phone service or have limited minutes, which could hinder access to the app and reduce engagement with its tools. They also noted that youth might have only intermittent access to mental health apps for various reasons, such as an inability to consistently pay for phone service or depleting their data or minutes. This loss of access could be frustrating for youth trying to use mental health apps and could interfere with their engagement with the app's content.

> *Most of our patients do have smart phones so that's usually not a barrier. But sometimes getting their cell phones shut off has happened before. So whether or not they'll have continuous access I think has been a barrier in the past.* [P8]

Theme 3: Engagement

Many providers expressed concern that a common barrier to using mental health apps is a lack of patient engagement. Several participants shared experiences of patients downloading apps but never using them beyond the initial session. Others noted that apps often lose their appeal quickly, making sustained engagement a challenge.

> *Yeah, and I think the literature is pretty clear that if you send patients off to use some sort of behavioral health app, they'll use it once or twice. They'll check it out and then use drops of precipitously.* [P1]

> *I think patients often download and forget about them which is probably the biggest one is the engagement with some of these apps is low. After a short amount of time they might forget about it and then they don't really come back to it. So in the context of therapy that was less of an issue because I would be there and be encouraging that use. So that's one of the barriers if people were just using it out in the real world.* [P9]

Providers described challenges in getting youth to engage with an app at all. Several noted concerns that apps must be visually appealing and up-to-date to attract teens, who are often highly tech-savvy. Based on their experience, if an app was not engaging enough, usage would decline, similar to what they had observed with other patient-focused apps.

> *I think we all know that teens are very sophisticated with apps and tech so I think having it be interesting and up to date is I think if they're going to use it, it's got to be flashy in some way.* [P3]

Theme 4: Patient Privacy

Most providers expressed concern about patient privacy. Some noted that their patients had raised concerns about the privacy of data collected by apps, which could make them hesitant to use mental health apps. Providers also emphasized that health systems would need to ensure patient privacy before integrating an app into treatment.

> *I know so many patients just in general are very concerned about their privacy. So I think that would be at the top of my list is just knowing that I could have confidence in saying to this person this app has*

XSL•FO
RenderX

> *been vetted in these different ways. This is the security that it offers. Maybe it's even password protected when you sign in so that way if your mom picks up the phone and she can't get into that particular app. I feel like that's a deeply personal thing. And I would want to make sure that I am protecting my patient's confidentiality and I can speak to how this app does that.* [P2]

> *I think always a concern is privacy and making sure the data is secure and HIPAA compliant. Especially working in health systems we're always concerned about technology and the privacy of patient information.* [P9]

Some providers also discussed concerns about patient privacy at home. Before recommending an app, they wanted to be confident in their understanding of how patient data were being protected.

> *On the other hand, if their parents are checking their phone and they don't feel like they have privacy they might not be able to use it in the same way that they might have otherwise had they had more privacy.* [P8]

Providers' Opinions About the Use of the BritePath App

Overview

All of the themes providers discussed regarding mental health apps in general were also reported in relation to the BritePath app. For instance, providers expressed privacy concerns both broadly and specifically in the context of BritePath. After hearing a description of BritePath, all but one provider expressed interest in trying this type of app for patients with depression, suicidality, or both. In addition to the general themes identified for mental health apps, several additional themes emerged specific to BritePath.

Theme 1: May Make Visits More Efficient

Providers described several ways in which an app such as BritePath could make therapy visits more time-efficient. Some noted that BritePath could help patients gain insight into the relationship between their mood and the activities they engage in between visits.

> *I think such a common thing with younger teens is being able to reflect back on how their week went or how their time went. And often times having something that is kind of tracking in the moment gives so much more actual information of when there were any spikes in anxiety or in distress in some way to help be like oh I saw this happened on this day, what happened then. To help anchor them into being able to talk about some things. So I can see that being a real benefit.* [P15]

> *I think it's so important for people to be able to see visually what the distribution is of the moods throughout the week. And also to remember that they're not always in that low point. So having some check ins throughout and maybe not only on their really bad days, right, but check ins every day would be helpful probably so they could be able to track down only the downs but maybe the ups or the little bit better.* [P14]

BritePath was also seen as a tool for providers to gather information on patient moods and their use of techniques between in-person sessions. Providers noted that access to this information via the app could make sessions more focused and productive by reducing the need to collect it during visits and allowing for better session preparation. Additionally, real-time patient-reported data on the app were perceived as more complete and valid than recalled information shared during a visit.

> *If you had the time to review what someone did or didn't do since you last saw them before they even walked in the door...I've saved at least 10 minutes of the precious 45 that we have together. I can launch straight into hey your mood meter was great over the last week. Or I see you check in on your safety plan on Tuesday, tell me a little bit about that? I don't have to say so how did the last week go? Oh Tuesday was bad. Tell me more about Tuesday and then wait for them to bring up the fact that they had to use their safety plan on Tuesday. We can just get straight to it. And it helps...not only the time saver, but hopefully not having to deal with retrospective recall.* [P5]

> *I think too if we have a session with a patient and they're just in that moment trying to recall their week is not as impactful as if you can actually go back and see what they're feeling in any particular time and probably would really help with recall bias.* [P8]

Providers also discussed how the BritePath app could facilitate communication with youth and enhance their engagement in the treatment process.

> *I think...leveraging digital health tool...I think another place this can be helpful is if kids engage in it, it gives you more to talk about in the session. You know doing sessions with teenagers a lot of times it's pulling teeth to get them to say anything. At least this would provide them with one, another way to keep them engaged and two, more fodder for talking to someone that may not be interested in filling 45 minutes' worth of talking.* [P5]

Theme 2: Provides Access to Safety Plans

Most providers noted that a key benefit of BritePath was giving patients convenient and consistent access to their safety plans through the app. They believed that because most youth carry their phones, they would be more likely to have their safety plan readily available in a crisis—unlike a paper version, which could be easily misplaced.

> *I think the nice thing about it, is it's all in one place. It's easily accessible on our phone. If they're in distress at any given moment they can pull it out and it's all right there. That makes it much more likely that they're going to use their skills in the moment, versus us giving them a piece of paper or emailing*

XSL•FO
RenderX

> *them or whatever else technology that would be used in the past. So I think that's a huge advantage.* [P13]

> *So patients can have a safety plan in the EHR but if they can't access it. We print it out and who knows where they left it. This gives them something, most people are pretty good about not leaving without their phone. So, it's a way that's there, it's accessible.* [P1]

However, providers also had questions and concerns about the functionality of storing the safety plan on a youth's phone. Some wondered whether the safety plan could be easily integrated into the patient's electronic health record. Another concern was whether providers would be notified when a patient accessed or used their safety plan.

> *Would the provider be alerted if there was a change? For example, if someone were to use the program or use part of their safety plan would that then ping me as a provider? Or would I just need to go and check on the patient status every so often to know those sorts of things? Another thing I might be curious about is how it might communicate with my EHR, if there was a way to import the information or download, even the safety plan for example so that way I could then upload it into the patient's chart. I think that would be important for me too.* [P2]

Theme 3: Offers an Additional Way to Communicate With Patients

Providers noted that BritePath could serve as an alternative communication tool for patients. Some recounted experiences with teens who struggled to express themselves and suggested that BritePath could offer these youth another way to connect with their providers. One provider also noted that this feature could enhance engagement in treatment for some patients.

> *Obviously, someone who is busy, if they're actually busy or just having issues with talking with a clinician face to face, this is a way to avoid that stress of direct interpersonal interaction. So that's a positive. And adolescents have all sorts of reasons why they sometimes don't want to engage with care. This is a way for them to get some care even if they're not feeling like interacting with their clinician that week. That I see as a benefit. Asynchronous so it can be on their time rather than someone else's time.* [P12]

> *If you think about teenagers, they are much more comfortable with technology and sometimes that's a great place to start for therapy or treatment with them because they might not feel comfortable talking to a professional.* [P9]

Several providers valued BritePath's 2-way communication feature, noting that it could facilitate shared decision-making and provide real-time updates to the treatment plan.

> *I think there's a real value in that shared decision-making component that can go into an app like the one you're describing. And it also becomes an important way of sort of communicating.* [P1]

> *I think that two-way communication is what I think could be the most useful if it's going between the patient and the therapist and you're able to interact back and forth. So in terms of modifying treatment plan or knowing what is resonating with the person.* [P3]

Theme 4: Flexibility and Personalization of the BritePath App

A number of providers expressed enthusiasm for BritePath's flexibility and personalization features. Several noted that allowing youth to customize the app with photos, videos, motivational reminders, or prompts for coping strategies could help sustain engagement, increase usage frequency, and accommodate diverse patient needs.

> *I think that's definitely a bonus. Sometimes the out of the box app works just fine. But the more you can tailor it or personalize it to someone's specific needs, the more likely they are to use it, so I think that that could definitely be useful!* [P9]

> *I love that there can be general tools but also some really individualized personal things in there in too. I just love the idea of having names and photos of people, loved ones and pets and things like that.* [P14]

One provider highlighted BritePath's potential to empower youth by allowing them to personalize their treatment and safety plan.

> *The place a teen usually goes to for soothing support regulation is the phone. But it often feels like the phone is just riddled with pitfalls of things that could make things worse. Maybe don't go on Instagram and compare your life to other people. So the fact that there's something right next to it potentially on their phone for them to dive into instead of something that could be potentially more disregulating or harmful I think is fabulous. The more opportunity I think there is to customize what's happening in that app the more I'd be inclined to really empower my client to tap into a wise place, tap into wiser parts of themselves, have those places and parts load what's in the app. So then they have access to it in those harder moments. I could see that being a particularly empowering thing.* [P11]

The collaborative functionality of the app was also seen as a benefit, as it enables providers to modify the app's content in collaboration with patients over time as their needs change.

> *I think that two-way communication is what I think could be the most useful if it's going between the patient and the therapist and you're able to interact back and forth. So in terms of modifying treatment plan or knowing what is resonating with the person.* [P3]

Theme 5: Concerns About Integration of BritePath Into Clinical Workflow

Providers expressed additional concerns specific to BritePath, beyond those related to mental health apps in general. One

XSL•FO
RenderX

concern was its impact on clinical workflow. For instance, providers questioned how well the app would integrate with existing electronic medical record systems. They also expressed concerns about its ease of use for busy clinicians.

> *Another thing I might be curious about is how it might communicate with my EHR, if there was a way to import the information or download, even the safety plan for example so that way I could then upload it into the patient's chart. I think that would be important for me too.* [P2]

> *Do I have to have a certain log in? Do I see a list of patients that I have that are using it? How cumbersome is it to get into that? And what's the patient's understanding and expectations for me, either monitoring that data or accessing that data or is there even some sort of portal to do that? Or is it more that the patient brings in their phone with their smart phone app and sort of shows me what they've been up to?* [P1]

In some cases, these concerns amplified existing worries about mental health apps in general, such as what might happen if a patient experiences a crisis. Additionally, several providers raised concerns about managing patient expectations regarding 24/7 access to their provider and ensuring that patients understood how to handle crises when providers were unavailable.

> *The only thing that has come up before... as far as cell phone and texting technology is somebody trying to text you in crisis while you're off work. You're not checking, so if there's something that presents in the app that is alarming or suggests that this client is in danger of harming themselves and you haven't checked the app in whatever amount of time, that would be I guess the concern for me too.* [P13]

> *I would want to make sure that whatever the app is that it stays within the boundaries of how I am available to the client or not. And I'd want it to be clear, it's tricky because if I can get a report about where they're at conceivably I could get a report when they're having a moment of increased distress and does that require me to then reach out to them too? So, I'd want it set up to be like hey either this app only releases this data to me X amount of times, so if you put in that you're having a hard time I am not going to see it.* [P11]

Discussion

Principal Findings

The focus of this qualitative study was to gain a deeper understanding of provider opinions on the use of mental health apps in general and the BritePath app specifically for treating teen and young adult patients with depression, suicidality, or both. Providers in this study had no prior exposure to the BritePath app. Most were aware of mental health apps in general and expressed interest in trying the BritePath app with patients with depression, suicidality, or both. Nearly all providers viewed mental health apps as a complement to other mental health treatments rather than a replacement for psychotherapy visits.

Mental Health Apps in General

Most providers were supportive of mental health apps, recognizing their potential to provide access to information and skills that could improve youth mental health. Key benefits included their availability at any time—particularly when providers were not accessible, such as late at night—and their ability to reinforce skills such as stress reduction between sessions. However, providers differed in their level of comfort with these tools. Some frequently recommended resources such as online cognitive behavioral therapy programs or anxiety management apps, while others were less enthusiastic. Most providers noted that mental health apps might be particularly useful for younger individuals, as they frequently use phones and other digital devices. These findings align with existing research on provider and patient perspectives, which indicate broad support for digital mental health tools [13,17,33-35]. However, most providers emphasized that apps alone were not sufficient for youth with significant mental health concerns. Many explicitly stated that these tools should complement, rather than replace, psychotherapy.

All providers expressed concerns about patient privacy. They discussed potential risks related to data collection by apps, including unauthorized access by individuals other than the youth (eg, parents) and uncertainties about how app providers might use collected data. While privacy concerns are common with app usage in general [13], providers emphasized that data collected by mental health apps are particularly sensitive and require high standards of privacy and protection. Several providers noted that app usage has become so commonplace that privacy policies are often overlooked, which may be especially problematic for youth. One provider specifically highlighted the difficulty in determining how commercial apps safeguard data privacy, which they found concerning.

Many providers also raised concerns about maintaining youth engagement with apps. Difficulty in sustaining engagement with digital tools has been noted in prior studies [36,37]. Several providers pointed out that they currently have little to no insight into how engaged their patients are with mental health apps. This uncertainty tempered some providers' optimism about the potential of these tools as a strong aid in treatment.

Providers also identified several logistical barriers to the effective use of mental health apps. The cost was a common concern, particularly the potential inaccessibility of apps that require ongoing payments (eg, monthly fees). Additionally, several providers noted that youth often experience inconsistent access to phones due to service suspensions, late payments, or lack of internet access. While these concerns have been raised in studies on app use in general [13], they may be especially critical for youth with mental health concerns, as intermittent access could disrupt their ability to use these tools as a coping mechanism.

BritePath App

Each of the potential benefits and concerns providers expressed about mental health apps in general were also raised in relation

XSL•FO
RenderX

to BritePath, including data privacy, patient engagement, and cost-related concerns. In some cases, providers expressed stronger opinions—both positive and negative—specifically regarding BritePath. For example, concerns about patient privacy were heightened due to the app's inclusion of sensitive information on symptoms and suicidality. By contrast, providers strongly supported BritePath's potential to enhance psychotherapy visits by facilitating more in-depth collaboration between providers and patients on skill development.

Providers raised several additional themes regarding BritePath. Many highlighted the value of the app's personalization features, noting that these could enhance patient engagement. Most providers expressed enthusiasm for BritePath's ability to facilitate communication between sessions and help providers prepare more effectively for upcoming appointments.

Providers also raised concerns about ensuring that safety issues were thoroughly addressed if BritePath were to be used in practice. While they appreciated the idea of having a safety plan integrated into the app, they emphasized the need for built-in safeguards, such as pop-up messages directing patients to emergency services when necessary. Additionally, providers wanted to ensure that patients understood that communication through the app would not guarantee a 24/7 response from providers.

Practical Implications for Research and Clinical Practice

Providers suggested several ideas for future research on apps addressing suicidality in youth. Integrating ambient phone data, such as activity tracking, could enable more timely and accurate recording of safety plan activities, reducing reliance on self-reporting. Additionally, visually linking mood and activity data, along with incorporating pop-up mood assessments, could enhance patient awareness of the relationship between mood and behavior. Increasing personalization in mood tracking and allowing patients to annotate their mood data may further improve engagement with the app.

Although providers were enthusiastic about the potential benefits of a customizable, readily accessible safety plan on a patient's phone, many also raised concerns about ensuring the security of this information. From a clinical perspective, safeguarding safety plans on the device would be essential for successful implementation within health systems. Additionally, embedding a clear and transparent patient consent process into the app—one that explains who will have access to their data and why—would be crucial in helping patients understand the risks associated with using a tool such as BritePath. Furthermore, ensuring that BritePath is compatible with different phone operating systems and can integrate with various electronic health record systems would be key to facilitating widespread adoption in clinical practice.

Strengths and Limitations

This study provides new insights into providers' perspectives on the use of mental health apps for youth at risk of suicidal behavior. We acknowledge the potential for sampling bias. To address this, we aimed to recruit a diverse range of providers from different health systems affiliated with the NIMH Mental Health Research Network [23]. While this sample may not be fully representative of providers across all health systems, participants were drawn from health systems in 7 US states. Although the sample is not fully representative of all providers in the United States, it includes a variety of geographic areas, organizational arrangements, and provider types, ensuring a diverse range of perspectives. In designing our qualitative interview guide, we focused on usability and concept testing rather than adhering to a specific technology adoption theory, such as the Technology Acceptance Model [26] or the Diffusion of Innovation Model [27]. However, our interview questions covered many of the key domains included in these models. We invited providers who delivered mental health services within these health systems to participate, without requiring a specific level of experience with apps or other digital tools in their practice, as we aimed to capture a broad range of experience levels. It is possible that providers with a greater interest in app use were more likely to participate.

One goal of this study was to explore providers' opinions about the BritePath app, which was developed for use with youth at risk for suicide. Specifically, we sought feedback from providers with no prior clinical experience or knowledge of BritePath to assess its potential for future implementation in health systems. Our focus was on understanding the broader applicability of mental health apps, with a particular emphasis on BritePath. While we collected limited demographic data on providers, it is possible that unmeasured factors influenced their perspectives on the use of mental health apps. For example, although respondents represented a range of age groups, most were younger providers, who may have been more comfortable with apps in general than older providers. Despite these limitations, this study offers new insights into providers' perspectives on mental health apps, including the BritePath app, and highlights key concerns regarding their use with young patients at risk for suicide.

Conclusions

The use of mental health apps is expanding, yet more research is needed to determine how they can be most effectively integrated into mental health treatment. Most providers expressed interest in using the BritePath app for patients with depression, suicidality, or both; however, concerns about privacy and safety remain.

Acknowledgments

This study was supported by the National Institute of Mental Health: The Center for Enhancing Triage and Utilization for Depression and Emergent Suicidality in Pediatric Primary Care (grant P50MH115838). The development and testing of BritePath have been funded by the National Institute of Mental Health (grants R34 MH100451 and P50MH115838), the American Foundation

XSL•FO
RenderX

for Suicide Prevention (AFSP), and a clinical use version institutional review board protocol (STUDY20050047) for collecting usability data. The authors thank Ms. Jill Pope for editorial assistance, Dr. Megan Hamm for feedback on the interview guide, Ms. Brandie George-Milford for editorial support, and Dr. David Brent for leading the P50 grant.

Conflicts of Interest

The authors have an SBIR proposal with Ksana Health to commercialize the BRITE app.

Multimedia Appendix 1
Qualitative interview guide.
[DOCX File , 3849 KB - humanfactors_v12i1e64867_app1.docx]

References

1. Van Meter AR, Knowles EA, Mintz EH. Systematic review and meta-analysis: international prevalence of suicidal ideation and attempt in youth. J Am Acad Child Adolesc Psychiatry 2023 Sep;62(9):973-986. [doi: 10.1016/j.jaac.2022.07.867] [Medline: 36563876]
2. Xiao Y, Cerel J, Mann JJ. Temporal trends in suicidal ideation and attempts among US adolescents by sex and race/ethnicity, 1991-2019. JAMA Netw Open 2021 Jun 01;4(6):e2113513 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.13513] [Medline: 34125218]
3. Madigan S, Korczak DJ, Vaillancourt T, Racine N, Hopkins WG, Pador P, et al. Comparison of paediatric emergency department visits for attempted suicide, self-harm, and suicidal ideation before and during the COVID-19 pandemic: a systematic review and meta-analysis. Lancet Psychiatry 2023 May;10(5):342-351 [FREE Full text] [doi: 10.1016/S2215-0366(23)00036-6] [Medline: 36907199]
4. Brewer AG, Doss W, Sheehan KM, Davis MM, Feinglass JM. Trends in suicidal ideation-related emergency department visits for youth in Illinois: 2016-2021. Pediatrics 2022 Dec 01;150(6):39-48. [doi: 10.1542/peds.2022-056793] [Medline: 36373281]
5. Bridge JA, Ruch DA, Sheftall AH, Hahm HC, O'Keefe VM, Fontanella CA, et al. Youth suicide during the first year of the COVID-19 pandemic. Pediatrics 2023 Mar 01;151(3):1-15 [FREE Full text] [doi: 10.1542/peds.2022-058375] [Medline: 36789551]
6. Lantos JD, Yeh H, Raza F, Connelly M, Goggin K, Sullivant SA. Suicide risk in adolescents during the COVID-19 pandemic. Pediatrics 2022 Feb 01;149(2):2-7. [doi: 10.1542/peds.2021-053486] [Medline: 34977942]
7. McCoy K, Kohlbeck S. Intersectionality in pandemic youth suicide attempt trends. Suicide Life Threat Behav 2022 Oct;52(5):983-993 [FREE Full text] [doi: 10.1111/sltb.12895] [Medline: 35735265]
8. Durante JC, Lau M. Adolescents, suicide, and the COVID-19 pandemic. Pediatr Ann 2022 Apr;51(4):e144-e149. [doi: 10.3928/19382359-20220317-02] [Medline: 35417315]
9. Brent DA, Goldstein TR, Benton TD. Bridging gaps in follow-up appointments after hospitalization and youth suicide: mental health care disparities matter. JAMA Netw Open 2020 Aug 03;3(8):1-3 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.13100] [Medline: 32780119]
10. Fontanella CA, Warner LA, Steelesmith DL, Brock G, Bridge JA, Campo JV. Association of timely outpatient mental health services for youths after psychiatric hospitalization with risk of death by suicide. JAMA Netw Open 2020 Aug 03;3(8):e2012887-e2012812 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.12887] [Medline: 32780122]
11. Doupnik SK, Rudd B, Schmutte T, Worsley D, Bowden CF, McCarthy E, et al. Association of suicide prevention interventions with subsequent suicide attempts, linkage to follow-up care, and depression symptoms for acute care settings. JAMA Psychiatry 2020 Oct 01;77(10):1021-1030. [doi: 10.1001/jamapsychiatry.2020.1586]
12. Blueprint for youth suicide prevention: strategies for clinical settings. American Academy of Pediatrics. 2021. URL: https://www.aap.org/suicideprevention [accessed 2023-05-31]
13. Van AM, Turna J, Khalesi Z, Pullia K, Patterson B. There is an app for that! The current state of mobile applications (apps) for DSM-5 obsessive-compulsive disorder, posttraumatic stress disorder, anxiety and mood disorders. Depress Anxiety 2017 Jun;34(6):526-539. [doi: 10.1002/da.22657] [Medline: 28569409]
14. Kao C, Liebovitz DM. Consumer mobile health apps: current state, barriers, and future directions. PM R 2017 May;9(5S):S106-S115. [doi: 10.1016/j.pmrj.2017.02.018] [Medline: 28527495]
15. Chan S, Li L, Torous J, Gratzer D, Yellowlees PM. Review of use of asynchronous technologies incorporated in mental health care. Curr Psychiatry Rep 2018 Aug 28;20(10):1-11. [doi: 10.1007/s11920-018-0954-3] [Medline: 30155593]
16. Martinengo L, Stona A, Griva K, Dazzan P, Pariante CM, von Wangenheim F, et al. Self-guided cognitive behavioral therapy apps for depression: systematic assessment of features, functionality, and congruence with evidence. J Med Internet Res 2021 Jul 30;23(7):1-17 [FREE Full text] [doi: 10.2196/27619] [Medline: 34328431]
17. Radovic A, Anderson N, Hamm M, George-Milford B, Fascetti C, Engster S, et al. Perceived facilitators of and barriers to implementation of a decision support tool for adolescent depression and suicidality screening: focus group and interview study. JMIR Ment Health 2021 Sep 15;8(9):1-16 [FREE Full text] [doi: 10.2196/26035] [Medline: 34524090]

XSL•FO
RenderX

18. Caldeira C, Chen Y, Chan L, Pham V, Chen Y, Zheng K. Mobile apps for mood tracking: an analysis of features and user reviews. AMIA Annu Symp Proc 2017;2017:495-504 [FREE Full text] [Medline: 29854114]
19. Marshall JM, Dunstan DA, Bartik W. Clinical or gimmickal: the use and effectiveness of mobile mental health apps for treating anxiety and depression. Aust N Z J Psychiatry 2020 Jan;54(1):20-28. [doi: 10.1177/0004867419876700] [Medline: 31552747]
20. Larsen M, Nicholas J, Christensen H. A systematic assessment of smartphone tools for suicide prevention. PLoS One 2016 Apr 13;11(4):e0152285-e0152214 [FREE Full text] [doi: 10.1371/journal.pone.0152285] [Medline: 27073900]
21. Kennard BD, Goldstein T, Foxwell AA, McMakin DL, Wolfe K, Biernesser C, et al. As safe as possible (ASAP): a brief app-supported inpatient intervention to prevent postdischarge suicidal behavior in hospitalized, suicidal adolescents. Am J Psychiatry 2018 Sep 01;175(9):864-872 [FREE Full text] [doi: 10.1176/appi.ajp.2018.17101151] [Medline: 30021457]
22. Goldstein TR, Kennard BD, Porta G, Miller AO, Aguilar K, Bigley K, et al. Bridging gaps in care following hospitalization for suicidal adolescents: as safe as possible (ASAP) and BRITE app. J Am Acad Child Adolesc Psychiatry 2024 Jul 18:1-13 [FREE Full text] [doi: 10.1016/j.jaac.2024.06.008] [Medline: 39032815]
23. Mental Health Research Network. URL: https://mhresearchnetwork.org/accessed [accessed 2023-06-01]
24. Goodman LA. Snowball sampling. Ann Math Statist 1961 Mar;32(1):148-170. [doi: 10.1214/aoms/1177705148]
25. Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. Adm Policy Ment Health 2015 Sep;42(5):533-544 [FREE Full text] [doi: 10.1007/s10488-013-0528-y] [Medline: 24193818]
26. Jacob C, Sanchez-Vazquez A, Ivory C. Understanding clinicians' adoption of mobile health tools: a qualitative review of the most used frameworks. JMIR Mhealth Uhealth 2020 Jul 06;8(7):1-20 [FREE Full text] [doi: 10.2196/18072] [Medline: 32442132]
27. Warty RR, Smith V, Salih M, Fox D, Mcarthur SL, Mol BW, et al. Barriers to the diffusion of medical technologies within healthcare: a systematic review. IEEE Access 2021;9:139043-139058. [doi: 10.1109/access.2021.3118554]
28. Atlas.ti. 2020 Jan 31. URL: https://atlasti.com/ [accessed 2020-01-31]
29. Davis MM, Currey JM, Howk S, DeSordi MR, Boise L, Fagnan LJ, et al. A qualitative study of rural primary care clinician views on remote monitoring technologies. J Rural Health 2014;30(1):69-78 [FREE Full text] [doi: 10.1111/jrh.12027] [Medline: 24383486]
30. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
31. Joffe H. Thematic Analysis. In: Harper D, Thompson AR, editors. In Qualitative Research Methods in Mental HealthPsychotherapy: A Guide for StudentsPractitioners. Harper DThompson A. editors. Chichester: Wiley-Blackwell 2012, pp. Edinburgh, Scotland: Wiley-Blackwell; Jun 28, 2011:209-223.
32. Nicholas H, McDowall A. When work keeps us apart: a thematic analysis of the experience of business travellers. Community, Work & Family 2012 Aug;15(3):335-355. [doi: 10.1080/13668803.2012.668346]
33. Melia R, Monahan L, Duggan J, Bogue J, O'Sullivan M, Young K, et al. Exploring the experiences of mental health professionals engaged in the adoption of mobile health technology in Irish mental health services. BMC Psychiatry 2021 Aug 19;21(1):1-11 [FREE Full text] [doi: 10.1186/s12888-021-03426-5] [Medline: 34412601]
34. Silfee V, Williams K, Leber B, Kogan J, Nikolajski C, Szigethy E, et al. Health care provider perspectives on the use of a digital behavioral health app to support patients: qualitative study. JMIR Form Res 2021 Sep 28;5(9):1-12 [FREE Full text] [doi: 10.2196/28538] [Medline: 34529583]
35. Anderson K, Burford O, Emmerton L. Mobile health apps to facilitate self-care: a qualitative study of user experiences. PLoS One 2016 May 23;11(5):1-21 [FREE Full text] [doi: 10.1371/journal.pone.0156164] [Medline: 27214203]
36. Amagai S, Pila S, Kaat AJ, Nowinski CJ, Gershon RC. Challenges in participant engagement and retention using mobile health apps: literature review. J Med Internet Res 2022 Apr 26;24(4):1-12 [FREE Full text] [doi: 10.2196/35120] [Medline: 35471414]
37. Williams JE, Pykett J. Mental health monitoring apps for depression and anxiety in children and young people: a scoping review and critical ecological analysis. Soc Sci Med 2022 Feb 12;297:1-15 [FREE Full text] [doi: 10.1016/j.socscimed.2022.114802] [Medline: 35192989]

Abbreviations

ETUDES Center: Center for Enhancing Treatment and Utilization for Depression and Emergent Suicidality
KPCHR: Kaiser Permanente Center for Health Research
NIMH: National Institute of Mental Health

XSL•FO
RenderX

Edited by H Witteman; submitted 12.08.24; peer-reviewed by R Melia, C Jacob; comments to author 04.11.24; revised version received 10.01.25; accepted 13.01.25; published 26.02.25.

Please cite as:
Lynch F, Cavese J, Fulton L, Vuckovic N, Brent D
Provider Perspectives on the Use of Mental Health Apps, and the BritePath App in Particular, With Adolescents at Risk for Suicidal Behavior: Qualitative Study
JMIR Hum Factors 2025;12:e64867
URL: https://humanfactors.jmir.org/2025/1/e64867
doi:10.2196/64867
PMID:

©Frances Lynch, Julie Cavese, Lucy Fulton, Nancy Vuckovic, David Brent. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 26.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Legitimacy as Social Infrastructure: A Critical Interpretive Synthesis of the Literature on Legitimacy in Health and Technology

Sydney Howe[1], MSc; Carin Uyl-de Groot[1], PhD; Rik Wehrens[1], PhD

Erasmus School of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands

Corresponding Author:
Sydney Howe, MSc
Erasmus School of Health Policy and Management
Erasmus University Rotterdam
Burgemeester Oudlaan 50
Rotterdam, 3062 PA
The Netherlands
Phone: 31 107547119
Email: howe@eshpm.eur.nl

Abstract

Background: As technology is integrated into health care delivery, research on adoption and acceptance of health technologies leaves large gaps in practice and provides limited explanation of how and why certain technologies are adopted and others are not. In these discussions, the concept of legitimacy is omnipresent but often implicit and underdeveloped. There is no agreement about what legitimacy is or how it works across social science disciplines, despite a prolific volume of the literature centering legitimacy.

Objective: This study aims to explore the meaning of legitimacy in health and technology as conceptualized in the distinctive disciplines of organization and management studies, science and technology studies, and medical anthropology and sociology, including how legitimacy is produced and used. This allows us to critically combine insights across disciplines and generate new theory.

Methods: We conducted a critical interpretive synthesis literature review. Searches were conducted iteratively and were guided by preset eligibility criteria determined through thematic analysis, beginning with the selection of disciplines, followed by journals, and finally articles. We selected disciplines and journals in organization and management studies, science and technology studies, and medical anthropology and sociology using results from the Scopus and Web of Science databases and disciplinary expert–curated journal lists, focusing on the depth of legitimacy conceptualization. We selected 30 journals, yielding 796 abstracts.

Results: A total of 97 articles were included. The synthesis of the literature allowed us to produce a novel conceptualization of legitimacy as a form of social infrastructure, approaching legitimacy as a binding fabric of relationships, narratives, and materialities. We argue that the notion of legitimacy as social infrastructure is a flexible and adaptable framework for working with legitimacy both theoretically and practically.

Conclusions: The legitimacy as social infrastructure framework can aid both academics and decision makers by providing more coherent and holistic explanations for how and why new technologies are adopted or not in health care practice. For academics, our framework makes legitimacy and technology adoption empirically approachable from an ethnographic perspective; for decision makers, legitimacy as social infrastructure allows for a practical, action-oriented focus that can be assessed iteratively at any stage of the technology development and implementation process.

(JMIR Hum Factors 2025;12:e48955) doi:10.2196/48955

KEYWORDS

legitimacy; health technology; infrastructure; literature review; technology adoption; health care governance; technology acceptance; health care delivery; social infrastructure; critical interpretive synthesis

XSL•FO
RenderX

Introduction

Background

Technological innovations and health care delivery have become increasingly intertwined in recent decades [1,2]. With each promising innovation, new questions are raised about how to integrate the technology into health domains and the social implications of doing so [2-8]. This problem has given rise to evaluation processes such as health technology assessment [9] and a range of government institutions, such as the National Institute for Health and Care Excellence in the United Kingdom, to help make decisions about the appropriate use of technologies in various health care contexts [10].

Questions of legitimacy are implicit in debates about embedding technologies in health care, and these implicit questions about legitimacy permeate the literature about health and technology. While there is limited conceptual work on legitimacy in health and technology specifically, within the small body of work that exists, legitimacy is explored through a wide range of theoretical lenses and conceptualizations. However, there is no consensus about what legitimacy is or how it works nor is there much conversation among different branches of the literature. Across social science disciplines and medicine, legitimacy is affiliated with many concepts, including "acceptability" [11], "reasonableness" [12], and "transparency" [3,4]; these terms are sometimes conflated with legitimacy itself. Definitions of legitimacy range just as widely, especially among different disciplines: in organization and management studies (OMS), legitimacy is often defined as "normative appropriateness" [13,14], although science and technology studies (STS) and medical anthropology and sociology (MAS) include definitions such as "the acceptability of claims to authority" [15,16] and "accountability for reasonableness" [10] when the term is defined. Some disciplines, such as OMS, follow a strong tradition of explicit theorization of legitimacy, while others, such as medicine, usually do not explicitly problematize or define the word "legitimacy," although legitimacy and similar concepts may play a substantial role in discussions in the field. Within the same discipline, some authors use legitimacy interchangeably with other concepts [6], while others focus on a single subcategory of legitimacy developed from a narrow line of theoretical thought [17]. Given the breadth and depth of the literature on legitimacy in health and technology, it is clearly an important topic. A cross-disciplinary conceptualization of legitimacy that can be applied to questions about embedding technologies in health care would allow for a deeper understanding of these processes.

We chose to focus on legitimacy for several reasons. Legitimacy specifically is widely used but poorly defined within multiple fields studying health and technology and points toward social and organizational processes. Even though legitimacy is clearly important in studies of health and technology across many disciplines, it is abstract and there is very little cohesion both within and across social science disciplines around how it is produced and how it is used.

To understand the role of legitimacy in relation to the embedding and governance of technology in health care, we fleshed out the different ideas and underlying assumptions authors bring to the OMS, STS, and MAS literature. This paper explores three research questions: (1) What does legitimacy mean in the context of health and technology? (2) How is legitimacy produced? (3) How is legitimacy used to explain the challenges of embedding technologies in health care? On the basis of a synthesis of the literature, we developed a novel conceptualization of legitimacy as social infrastructure. This allowed us to focus on the aspects of legitimacy related to norms, materialities, semiotics, and specific relationships among stakeholders and larger systemic contexts.

Existing frameworks guiding policy intentions around technology adoption in health care often do not reflect the realities of everyday practice nor do they focus on the social processes of embedding technologies in health care [18]. Current models and frameworks for embedding technologies (in health care), such as the Unified Theory of Acceptance and Use of Technology [19] and the Technology Acceptance Model (TAM) [20,21], provide useful information about the judgments and perceptions of various actors in relation to a specific technology. While these models provide useful insights into actors' *intention* to adopt, as well as the ethical considerations they may perceive and institutional norms they may face, they do not often explain how and why certain technologies are or are not adopted or accepted *in practice* among particular groups and thus can provide only limited governance support [22]. Most technology adoption and acceptance models deal primarily with behavioral intention [20,21]. These models attempt to explain individual behavior intentions (as a proxy for actual behavior) through an analysis of a variety of factors, which can include beliefs, attitudes, perceived control, and perceived norms. This leaves substantial gaps, including the lack of attention to external factors impacting the embedding process [20] and any difference between behavioral intention and practice. Focusing on the impact of legitimacy within technology embedding processes can help fill these gaps by allowing for more comprehensive and nuanced conversations about technology adoption and acceptance.

A Cross-Disciplinary Legitimacy Conceptualization

The goal of this review is to produce a theoretical synthesis of how the legitimacy of health and technology is conceptualized. We aim to generate theoretical insights that will allow for a cross-disciplinary approach to legitimacy. To do this, we used the critical interpretive synthesis (CIS) review method to produce a cross-disciplinary understanding [23] of how legitimacy is produced and maintained without privileging one discipline's theoretical traditions over another's [24].

We used OMS, STS, and MAS as a disciplinary framework to understand the different strands of reasoning within legitimacy studies. OMS includes the most clearly developed theorization of legitimacy as a concept and several of the best-known review articles [14,25]. Legitimacy studies is a well-defined subarea of OMS research that explores legitimacy in the context of management, entrepreneurship, and institutional endeavors. OMS primarily focuses on how legitimacy works at the level of organizations and institutions, rather than among individual people or through technologies.

XSL•FO
RenderX

STS explores the social and cultural implications of science, technology, and technology development [26,27]. STS tends to focus on controversies, the unforeseen implications of technology development and adoption, and "black box" technologies [28] that inhibit full understanding. This contributes to a strong disciplinary interest in legitimacy, particularly of technology and medicine. STS centers interactions with technologies and technical change in legitimacy discussions, rather than separating the motivations of human actors from technologies and practices.

Although MAS developed from different traditions, they have partially converged in the study of medical cultures and interactions from a qualitative perspective [29], which is also the focus of legitimacy studies in MAS. MAS generally takes a skeptical view of taken-for-granted institutions. Many studies explore the impact of this taken-for-grantedness on people who are harmed by or left out of these institutions and systems. This leads to a natural questioning of legitimacy and deeper exploration of baseline assumptions, particularly in a health care context. Together, these disciplines explore health and health technologies with less focus on institutions and more focus on humans and connections among them. We hope that greater synthesis of different conceptualizations of legitimacy could provide a more holistic perspective about how and why technologies are or are not embedded in health care that integrates ideas from outside current frameworks based on adoption, acceptance, and implementation.

This review does not focus on ethics, although some articles that deal with ethics are included in this review. Within ethics, there is general consensus that legitimacy in health and technology is produced through ethical judgments and values-based reasoning [30-33]. While there is more cohesion in ethics than in other disciplines that deal with legitimacy, the scope of the literature is narrower and generally limited to moral or ethical legitimacy [3,34]. Legitimacy is often defined in ethical terms across several disciplines, including management, STS, and health policy. Articles with an ethical focus on health care technology adoption usually prioritize normative and moral dimensions of legitimacy, which may or may not be related to other aspects that contribute to overall legitimacy, such as utility or social acceptability [35]. While moral legitimacy is an important dimension of the subject, it has also already been studied and reviewed extensively and is far more clearly defined within the literature than other dimensions of legitimacy. Therefore, we chose to focus on other aspects of legitimacy to fill gaps in the existing social science literature. In addition, our review does not center legal approaches. Legal literature defines legitimacy according to principles of a participatory democracy [36]. Because the focus of this review is on social processes rather than democratic principles, we have chosen to center other means of legitimacy production and thus other branches of the literature, while acknowledging that democratic principles can play an important role [15,37].

Methods

Overview

In a research landscape that increasingly encourages and necessitates interdisciplinary collaboration, a CIS [24] allows us to develop a richer conceptual frame that can bring these different disciplines into conversation with one another [23,38]. The CIS method specifically addresses the limitations of conventional systematic reviews in cases where the objective is to generate a critical analysis of a body of the literature [24]. It draws on a tradition of qualitative inquiry: rather than summarizing the literature using preidentified key concepts, CIS uses an inductive, iterative process to interpret data, develop key concepts, and synthesize concepts with data through the process of the review itself [39]. Therefore, CIS produces a critical interpretation of the data reported in literature, which allows questioning of taken-for-granted assumptions [24].

We found CIS to be a particularly appropriate review style for an exploration of legitimacy in health and technology because, unlike traditional review styles, CIS is adept at dealing with "a large, amorphous, and complex body of literature" [24]. CIS synthesizes findings in a way that is both conceptually valuable to researchers and useful in informing policy. Furthermore, the CIS review style can manage the abstract nature of legitimacy across 3 disciplines without reducing the complexity of existing literature.

CIS is methodological and rigorously structured. However, it is also labor intensive compared to other review styles and deprioritizes transparency and full replicability available through systematic reviews. As in other kinds of interpretive research, different researchers using the same materials may have come to different conclusions, and full transparency of analysis cannot be expected at the same level as quantitative studies. We chose the CIS method in service of greater interpretive flexibility and conceptual depth and because it allowed us to account for the diversity of qualitative research without privileging a small range of studies that met specific and, here, less-relevant methodological standards.

Despite many empirical papers that address some aspects of legitimacy in relation to health and technology, there is a much more limited number of articles dealing with legitimacy in the domains of health and technology at a conceptual level. Therefore, we adopted broad definitions of both technology and health, including as many articles with strong legitimacy theorization as possible. "Health" in this paper refers to "human health and well-being"; "technology" includes "any devices, machinery, or equipment developed for the application of scientific knowledge that may impact human health." We acknowledge that there are potentially many other ways to operationalize these concepts that could be relevant to this review and health care in general [40]. Given the broad, conceptual focus of the review, as well as the wide range of distinctive technologies considered, we introduced 2 measures that allowed us to provide a clear focus. First, we zoomed in on conceptualizations of legitimacy; it was outside the scope to examine the specific empirical contexts and details of these conceptualizations (ie, the use of specific technologies within

XSL•FO
RenderX

specific health practices). Second, looking at specific disciplines proved helpful to demarcate boundaries around our topic. We used academic disciplines as the basis of our theoretical categories to give structure to a broad and abstract review, following in a strong disciplinary synthesis tradition established by other CIS researchers [23,38].

Our review was conducted in 4 cycles [38]. In the first cycle, we conducted an exploratory search in relevant journals with many articles on legitimacy in the context of health and technology. Using this search and subsequent iterations, we refined inclusion criteria (for instance, a threshold for legitimacy theorization) to determine a list of disciplines, journals within each discipline, and articles within these journals that would be a part of our review (see Multimedia Appendix 1 for details). In the second cycle, we read and coded included articles directly from the text, then grouped codes into thematic areas. Third, we analyzed legitimacy in health and technology within each discipline through interpretive disciplinary summaries. Finally, we developed a synthesizing argument across all 3 disciplines, based on these interpretive summaries.

Search Strategy

Journals within each discipline were selected for the review manually. Given the interdisciplinary nature of this review, we sought to include relevant journals within each discipline from a wide range to ensure diversity of perspectives.

The first author ran an initial search in Google Scholar and the Web of Science (WoS) database using the terms “legitimacy,” “health,” and “technology.” However, it became clear that regardless of how specific the search criteria became, most records returned through a database search would not include deep conceptualizations of the term “legitimacy.” This is because this term is commonly used across many disciplines as an unproblematic descriptor of technologies, legal processes, narrative choices, medical decisions, policies, and institutions. Therefore, we switched to a more iterative, manual strategy for discipline and subsequently journal selection (Textbox 1). The first author searched through a few of the major journals (based on impact factor or prior knowledge of the review team) in depth to understand whether studies addressing questions of legitimacy in a specific discipline generally included deeper conceptualization. This method allowed us to narrow down the disciplines to be included (see Multimedia Appendix 1 for details).

XSL•FO
RenderX

Textbox 1. Selected journals.

Organization and management studies

- Journal of Business Venturing
- Journal of Management Studies
- Organization Studies
- Organization Science
- Administrative Science Quarterly
- Health Care Management Review
- Strategic Entrepreneurship Journal
- Health Care Analysis
- Strategic Organization
- European Journal of Public Health

Science and technology studies

- Science as Culture
- Science, Technology and Human Values
- Social Studies of Science
- Social Epistemology
- Journal of Responsible Innovation
- AI & Society
- Big Data and Society
- Risk Analysis
- Information Communication & Society
- Research Policy

Medical anthropology and sociology

- Medical Anthropology
- Medical Anthropology Quarterly
- Anthropology & Medicine
- Qualitative Health Research
- Critical Public Health

  (interdisciplinary journal: some articles included within science and technology studies)
- Sociology of Health & Illness
- Health: An Interdisciplinary Journal for the Social Study of Health, Illness and Medicine
- Social Science and Medicine

  (interdisciplinary journal: some articles included within science and technology studies)
- Sociology-The Journal of the British Sociological Association
- Sociological Theory

Upon the selection of disciplines, journals were first included based on impact factor: the top 10 journals in each discipline according to the WoS formed the base of the list. Then these journals were screened for topical relevance, and lower-ranked journals were excluded when the topic was not relevant. Finally, journals were added from library guides created by universities that are well-known in the field, especially if no definitive rankings were available from the WoS (STS) or there was too much disciplinary overlap in the rankings (MAS). This resulted in a list of approximately 20 journals per discipline. We then screened all journals for legitimacy, health, and technology results to produce the final list, prioritizing journals that contained significant theorization and conceptual work around legitimacy (see Multimedia Appendix 1 for details).

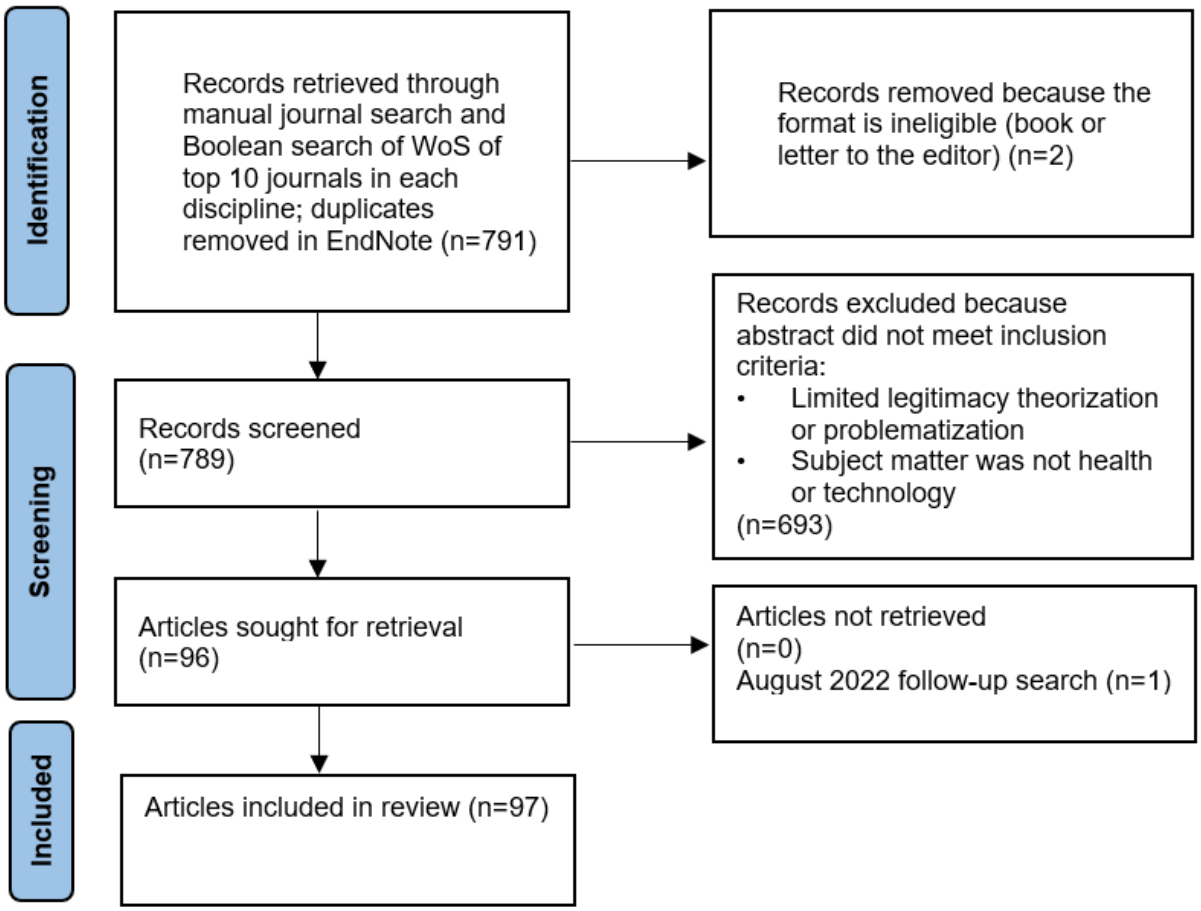
XSL•FO
RenderX

Selection Process

Overview

A total of 791 abstracts were selected for initial review in April 2021 (Figure 1). Articles were sorted based on inclusion criteria: an article was included if there was evidence of explicit theorization or problematization of legitimacy in the abstract and the subject of the article was health or technology or both. A final search for more recent articles was run in August 2022, resulting in 1 additional inclusion. Multimedia Appendix 1 gives more information about the inclusion criteria.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of inclusions. WoS: Web of Science.



Triangulation

The first author (SH) read all 791 abstracts and categorized them according to the inclusion criteria as "include," "exclude," or "review." The last author (RW reviewed 79/791, 10% of the total articles' abstracts at random (every 10th article when arranged in alphabetical order by author name). The resulting inclusion list closely matched SH's selection. In the 7 disagreements, the 2 reviewers came to a joint decision about each contested article and edited the inclusion criteria. SH then reviewed the abstracts again. With no disagreements following this protocol, the literature list of 96 articles was finalized in March 2021, with an addition of 1 article in August 2022, bringing the total to 97, including, by discipline, 48 for STS, 29 for OMS, and 20 for MAS [4-8,10,12,13,15,16,37,41-126].

Synthesis Methods

The goal of our initial analysis was to develop "synthetizing constructs" for each discipline, which "build on the explanations and interpretations of the constituent studies and are simultaneously consistent with the original results while extending beyond them" [24]. To accomplish this, SH first coded themes across a random selection of 30 articles spread equally across the 3 disciplines in Atlas.ti (ATLAS.ti Scientific Software Development GmbH). Inductive coding began with words taken from the article text. We made an exception for deductive codes relating to assumptions made as part of our "critical and reflexive approach" to this review [24]; for instance, coding whether articles discussed legitimacy as a process, perception, or asset based on our background understanding of existing legitimacy theorization across multiple disciplines. SH grouped codes by theme within each discipline. This thematic analysis was performed with the help of tables of code co-occurrences, generated in Atlas.ti, which showed overlap of unique codes within each document; this allowed us to understand when codes could be combined into a theme. We also used network visualizations, which helped us understand how different aspects of legitimacy may relate to one another on a broader scale (see Multimedia Appendix 2 for an example of one of these visualizations). This process allowed reviewer SH to maintain a differentiated perspective on each discipline and acted as a memory aid for the reviewer given the volume of the literature. SH then wrote summaries of the approach to

XSL•FO
RenderX

legitimacy within each discipline, which were discussed with RW. SH coded the remaining articles using Excel (Microsoft Corporation) while editing the disciplinary summaries as necessary. Multimedia Appendix 1 provides an example.

SH developed a synthesizing argument that could connect constructs of legitimacy across all 3 disciplines. SH tested the fit of different conceptualizations with the disciplinary summaries and then explored the development of these metaphors with RW. After SH had read 90% (87/97) of the articles, no further edits or additions to the disciplinary summaries or synthesizing argument were necessary, likely indicating data saturation.

Results

Organization and Management Studies

What Does Legitimacy Mean?

The basis of legitimacy epistemology in OMS can be traced to the highly influential paper by Suchman [14]. His definition of legitimacy still forms the basis of most subsequent theorizing in this field: "Legitimacy is a generalized perception or assumption that the actions of an entity are desirable, proper, or appropriate within some socially constructed system of norms, values, beliefs, and definitions" [14].

This definition builds from and expands on Weber's [127] understanding of legitimate authority, among other conceptualizations [41], focusing on the application of Weberian legitimacy concepts such as "tradition, affectual attitudes, value, and legality" [127] within a "socially constructed system of norms, values, beliefs, and definitions" [14]. Bitektine and Haack [128] further expanded on the Weberian connection to Suchman [14] with a multilevel legitimacy theory that differentiates among legitimacy judgments at the individual and collective levels [42]. Because of strong theoretical cohesion in this discipline, authors draw generously from these theoretical traditions, and most legitimacy subtypes are variations of normative legitimacy. Normative legitimacy deals with the norms, systems, and underlying assumptions that result in an institution, technology, or system being considered legitimate, and is generated through systems and social processes that may not be made explicit until there is a legitimacy crisis [41].

This definition is notable for its understanding of legitimacy as a perception of actions based on socially constructed norms and implicit conceptualization of legitimacy as a taken-for-granted state. However, perhaps more important is what is left out: because legitimacy is "generalized," it is not dependent on materialities or physical space, nor does the definition of legitimacy change specific to relationships among institutions or individuals. Our critical interpretive analysis shows that these kinds of assumptions seem to help determine the lines of reasoning for most OMS legitimacy scholars included in this review.

Our analysis shows that OMS legitimacy studies constitute a cohesive body of literature in which theoretical contributions are made incrementally. For instance, most authors outline exactly how they are adding to the existing literature (usually based on the definition by Suchman [14]) in a specific way. For instance, articles often divide legitimacy into ever more granular categories such as "network legitimacy" [43] and "actional legitimacy" [44] or focus on a particular aspect of the definition by Suchman [14] to explore in more depth, such as "normative appropriateness" [129]. Therefore it is easy to trace the lineage of theory. However, because every article adheres to the same framework, there is less diversity of thought. In light of this, a division could be made between technology literature and health care literature. Technology literature, usually with an entrepreneurship bent, tends to focus on audience perception and innovation framing through discourse, by teaching entrepreneurs how to appeal through discourse to different kinds of funding audiences [45]. For example, entrepreneurs seeking crowd-based funding are advised to make "contribution claims" emphasizing identity and added value to the funding community, rather than institutional ties or the significance of potential scientific or social advancements that may legitimate the entrepreneur to other audiences [45]. Health care literature focuses mainly on institutional legitimacy, for instance, by examining the negotiations involved in academic health center mergers [46] or the role of physician executives in institutional contexts [47]. For example, Kitchener [46] emphasizes that even as new sources of legitimacy for an academic health center became necessary in the wake of both a merger and a changing environment, the aim was to create an "aura" of legitimacy within a new institutional logic; in other words, even though the source of legitimacy changed, it was still defined through institutions.

How Is Legitimacy Produced?

On the basis of the literature reviewed for this study, conceptualizations of legitimacy production in OMS fall into 3 categories. Legitimacy-as-process, or legitimation, is "an interactive process of social construction" [25], often with clearly delineated steps toward generating legitimacy [48,130]. Articles using this conceptualization tend to track change over time [13,49,50], for instance, by examining the legitimation journey of a single innovation over decades [50]. Legitimacy as perception or "a social judgment, an evaluation, a socio-cognitive construction" [25] is used most frequently in relation to audiences or the question "legitimate to whom?" usually when examining innovations within larger organizational contexts [45,51,52]. Finally, a legitimacy-as-property conceptualization, in which legitimacy is viewed as an asset to be gained, increased, or lost [25], emphasizes how an entity gains legitimacy rather than differences between legitimate or illegitimate entities; this conceptualization appears to play a smaller role in more recent work. Legitimacy as a process and legitimacy as a perception are often used simultaneously within the same paper; there is an acknowledgment that both perspectives can offer valuable insights. For instance, McKnight and Zietsma [53] describe the process through which new ventures can achieve "optimal distinctiveness" (a perception) through framing and collaboration strategies (a process). Similarly, Garud et al [54] discuss entrepreneurial storytelling as a legitimation process that relies on audience perceptions, thereby promoting an understanding of audience perceptions

XSL•FO
RenderX

as part of analyses of different narrative techniques, which are framed as processes of legitimation [53].

How Is Legitimacy Used?

Within the literature reviewed, we noted that legitimacy is often assumed in OMS to be necessary for resource acquisition and is mostly relevant in an institutional setting. Because resource acquisition is the *raison d'être* of business in capitalism, legitimacy is considered essential. All literature included in our review aims to explain legitimation journeys or to advise institutions and entrepreneurs on how to move toward legitimacy. The stated goal is often to advance legitimacy theory within OMS, rather than explore empirical evidence beyond specific case studies.

Our analysis shows that nearly all legitimacy problems in OMS are framed as problems of categorization: either an institution or innovation has been miscategorized by its intended audience, or it has not yet been categorized [51]. Because of this, there are several articles that deal with the paradox of legitimacy for disruptive innovations: the innovation must conform to a category to be seen as legitimate, but it will not be seen as necessary unless it also distinguishes itself within the category [45,53,54]. Articles often describe how entrepreneurs can strike this balance or analyze how an older innovation navigated the paradox. For instance, Ruef and Scott [41] produced a multidimensional model of legitimacy designed to help hospitals use knowledge of legitimacy norms and categorization to "attract managerial legitimacy." Other authors provide articles that, while academic, can also be interpreted as how-to guides for entrepreneurs attempting to formulate narratives for different audiences. Articles in this category posit that entrepreneurs must fit the narrative to the correct category of both the legitimacy problem [54] and the audience's perspective [45].

Articles discussing specific innovations often explore problems of categorization through "institutional logics" [45,46,55] or the forms normative legitimacy may take within a particular institution. These articles acknowledge that different institutions have different norms, and therefore, entrepreneurs and innovators must adapt their storytelling to "fit" with the institution. Our analysis of the OMS literature showed that institutions that are taken for granted have no narrative flexibility. This means that any changes made not only require buy-in from professionals but also immediate practical implementation of changes (such as shifting patient education responsibilities from physicians to nurses) to avoid becoming mired in default institutional processes [56]. However, an entrepreneurial venture wrestling with liability of newness [53] has nearly endless opportunities for narrative reframing and limited risk of stagnation in nonexistent institutional processes [45].

Summary

In short, OMS seems to heavily theorize granular definitions of legitimacy, derived mostly from normative legitimacy. Most articles describe how organizations can move from illegitimate to legitimate status. Legitimacy must be negotiated at multiple levels and can be conceptualized as a process, perception, and asset. On the basis of our initial analysis, legitimacy in OMS rests on taken-for-grantedness and categorization according to institutional logics; it is hampered by liability of newness [51] and difficulties fitting into an institutional logic or other category.

We now move on from the critical interpretation phase of CIS methodology and toward synthesis. We develop synthetic constructs to metaphorically unite major themes within each discipline. While the OMS line of reasoning is somewhat homogeneous, our interpretive analysis shows that there is acknowledgment that different institutions use different systems of thought, or institutional logics, to judge and negotiate legitimacy. Stakeholders must navigate these different systems by fitting themselves into clear categories. However, the system by which legitimacy is judged and negotiated is rarely fully navigable to any one stakeholder.

In this way, navigating legitimacy in OMS is like using a highway system. The roads themselves (in this case, pathways to legitimacy) are clear and concrete but can only be used by particular categories of vehicles in particular locations, in the same way that, for instance, entrepreneurs must use particular discourses (ie, roads) to legitimate their invention to certain audiences (locations) [45,50,54]. The metaphor of the highway system allows us to capture OMS' conceptualization of legitimacy as a rule-bound system of connections, driven by institutions. The highway system metaphor shows us that the OMS perspective on legitimacy is unusual for how fully theorized and comprehensive it is. However, in the same way that mapping a highway system alone does not describe everything that happens on the road, this metaphor also helps us articulate gaps in OMS theorization. OMS focuses on pathways to legitimacy, institutional involvement, and categorization; however, it can be difficult to explore the moments of legitimacy breakdown, unexpected contradictions within legitimacy journeys, and the influence of materialities on legitimacy within the concrete and bounded pathways that make up legitimacy conceptualization in OMS.

Science and Technology Studies

What Does Legitimacy Mean?

Within the legitimacy literature reviewed, nearly all STS articles dealt with health care or science practices on the edge of scientific validity or public acceptance, usually by focusing on controversial technologies. Authors seemed to pay significant attention to the audience for these controversies. Innovations explored within STS also include changes in health care and scientific systems or discourses, for instance, the process of creating new institutions within an academic environment [57].

In this way, it seems that legitimacy in STS can only be discussed easily when it is in question. This may be linked to STS's longstanding focus on controversies as strategic locations of research and the STS tendency to "stay with the trouble" [131]. Taken-for-grantedness may be considered such an integral part of legitimacy that it raises red flags in a discipline that is highly sensitive to unintended consequences.

However, in the STS literature reviewed, legitimacy itself is rarely the subject of an article. Our interpretive analysis shows that legitimacy is still undertheorized in STS, in that many

XSL•FO
RenderX

articles do not reference other legitimacy literature or define the term. When legitimacy is explored in depth, authors tend to deal with it as a type of knowledge coproduced at the site of controversy between human and nonhuman actors. For instance, 2 articles examine the demarcation of boundaries in discourse about future-facing technologies; in these articles, narratives about nanotechnology and forensic science both impact and are impacted by developments in the technology [16,58]. For instance, Selin [16] demonstrates that even if a single vision of the future of nanotechnology is accepted by scientists over all others, competing visions may still have an impact on the direction of technological development via other avenues, such as governance and funding. We interpret this as evidence that legitimacy is coproduced between the narrative around new scientific disciplines and concrete innovations within those disciplines.

A consequence of the focus on disputed boundaries that we observed in the STS literature seems to be that STS scholars treat legitimacy as mutable. Boundary work [59] articles conceptualize legitimacy as a process that happens through negotiation within relationships among people and institutions. Boundary work deals with how people and institutions negotiate changing roles, limits, and responsibilities, often in relation to knowledge practices [59]. Within our review, most authors seem to use boundary work only as an explanatory framework for discursive strategies, but some articles include writing and practices as instances of boundary work [60]. For instance, Granja and Machado [58] describe how forensic scientists clearly demarcate their expertise and the limits of their role in the justice system through both discursive strategies and practical performances of accountability such as fastidious record keeping as a strategy to maintain their legitimacy. While perception can play a role in this process, we found no STS boundary work articles within the literature reviewed that consider legitimacy to be merely an asset. We interpreted this as possible evidence of the STS disinclination to consider any relationship or object to be fixed and a general disciplinary focus on processes of knowledge production.

How Is Legitimacy Produced?

We found that processes of legitimacy coproduction of new technologies in health care are described in relationships among a variety of human and nonhuman actors [28]. STS seems to resist explanations that pin legitimacy's presence or absence on a single explanation or strategy. Instead, by examining the influence of nonhuman actors, STS encourages readers to embrace complexity in legitimacy relationships. For example, Barker [61] discusses how a drug marketing campaign legitimated a contested illness; McLevey [62] details how legitimacy goals for policy recommendations change how think tanks make knowledge; and Greco [63] uses the concept of "scientific ideology" to explore how the placebo effect threatens the legitimacy of scientific rationality in biomedicine. In these examples, the influence of unexpected nonhuman actors shapes legitimacy in a porous way; this influence adds complexity to legitimacy relationships. These examples thus turn the "expected" legitimacy relationship upside down. Barker shows that a treatment can make an illness legitimate (rather than an illness becoming legitimate and then treatment being developed). McLevey [62] demonstrates that knowledge may be made to legitimize policy recommendations (rather than knowledge first being created and then informing policy decisions). Finally, Greco [63] posits that a biomedical effect can threaten the legitimacy of biomedicine (rather than biomedicine creating biomedical effects that reinforce its legitimacy).

How Is Legitimacy Used?

Through our critical interpretive analysis, legitimacy in STS appears to be yet another "two-faced Janus" [28], in that it is coproduced by both taken-for-granted processes and unexpected assemblages, which may appear contradictory. For instance, Brown and Michael [64] note that the institutional mechanisms for validating scientific discovery play a part in legitimizing medical use of pig organs. However, the authors also find that these taken-for-granted institutional processes are not enough to legitimate transspecies transplantation. Rather, porcine donor organs are legitimated in their interpretation through an unexpected assemblage of institutional processes and scientific and cultural judgments about pigs. While anatomical similarities between humans' and pigs' organs exist from a scientific perspective, human-pig differences legitimate porcine donor organs (over, for instance, monkey organs) from a cultural perspective [64]. These apparent contradictions do not threaten the legitimacy of porcine donor organs; instead, the authors interpret these contradictions as necessary to establish legitimacy for a controversial innovation in medicine.

While most articles in the reviewed STS literature use the concept of legitimacy to explore complexity in relationships among technologies, ideas, people, and institutions, a small subset of the literature consists of discourse analyses focusing on language as the only vehicle of legitimation [4,62]. In these articles, there is no room for other means of legitimation, such as nondiscursive practices, language-free materialities, or physical space. In addition, policy-oriented articles about discourse-based legitimation strategies tend to view these strategies as mutually exclusive: "legitimacy can only be achieved if complexity is reduced via language" [65]. This appears to be a radical departure from the assembled view of legitimation present in most STS articles, in that legitimacy is produced by specific arguments designed to convince one party that another is legitimate, rather than *through* the complexity of negotiation among people, technologies, ideas, and institutions. However, both STS subgroups use legitimacy as a relational concept, whether relationships are implied through arguments or complex negotiations.

Summary

Our analysis demonstrates that STS approaches legitimacy through controversy and contestation. Because of its empirical orientation and wide variety, the way legitimacy is approached in STS cannot be viewed as theoretically cohesive. We found that STS explores legitimacy as an artifact of knowledge coproduction, centering nonhuman actors and complex legitimacy processes through discourse, practices, or both. In one conceptualization, legitimacy is viewed as a constant and complex negotiation among actors and their contexts. However, discourse analyses often view the complexity of these relationships as a problem for legitimacy rather than an integral

XSL•FO
RenderX

aspect of the concept. Despite this divergence, nearly all STS literature reviewed seems to assume that legitimacy is relational and includes nonhuman actors, contexts, and practices of knowledge production.

Moving toward our synthetic constructs, we posit that in STS, the coproduction of legitimacy through both taken-for-granted processes and unexpected assemblages functions such as the mutable infrastructure of a computer operating system. The hardware of the computer does not change (much), even though the overall capacity of the computer relies on constant updates of existing software. New operating systems build on what came before while adding something new to stay relevant. Similarly, the legitimacy of technologies in health care is viewed predominantly as being built on underlying infrastructures, such as evidence-based medicine (EBM) and laboratory protocols. However, it is also considered to be shaped by unexpected influences, such as the coexistence of traditional medicines with EBM standards [66] and popular conceptualizations of new technologies that do not align with the use of those technologies by experts [58]. Furthermore, legitimization of a technology may have unexpected consequences, such as a medicine legitimizing an illness [61]. Legitimacy, according to STS scholars, must be constantly negotiated among nonhuman actors through an iterative process that is most visible at the boundaries. Through the metaphor of computer operating systems, STS emphasizes a dualistic perspective of legitimacy. Legitimacy in STS depends on the mutability and unexpected creativity of practices, actors, and context (ie, software); simultaneously, this "software" of legitimacy must be built on top of preexisting systems that are much more rarely changed in substantial ways (ie, hardware). Without these preexisting systems, the knowledge production practices delineated by STS scholars have nowhere to operate.

Medical Anthropology and Sociology

What Is the Meaning of Legitimacy?

On the basis of our critical interpretation of the MAS literature under review, we found that MAS tends to center relationships among people within larger systems and physical spaces, rather than the nature of the systems or institutions themselves. Similar to STS, the MAS literature reviewed usually discusses controversies rather than strategies of legitimation.

Because few articles use concrete definitions of legitimacy or conceptualizations that reference any other body of literature, even if they discuss legitimacy issues in depth using empirical evidence, we found legitimacy to be undertheorized within MAS as a discipline. Several articles use terms such as "medical legitimacy" [6] or "narratives of legitimation" without definitions [67]. Some define legitimacy only in relation to another concept [12,68,69]. More than any other discipline, MAS seems to veer toward legitimacy definitions that are limited by subject and narrow [59] in scope: for instance, "being legitimately on sick leave requires both a certified illness and the inability to perform one's work as a result of that illness" [70]. In our interpretation, this appears to allow for more empirical grounding but prevents conversation with other literature about legitimacy as a larger concept. Within the MAS literature reviewed, legitimacy is always relative, in that there is no universal legitimacy that works across all contexts.

However, there is still some cohesion visible among a subset of articles in the discipline. Several authors frame legitimacy as a form of "symbolic capital" in the Bourdieuvian sense [69,71,132] that can be used for resource acquisition. These articles and others also use boundary work [59] to describe the legitimacy of professional roles as a negotiation of boundaries mostly through discourse [72]. In combination with the disciplinary tendency to focus on the disenfranchised, these framings paint legitimacy as a power issue [37].

The literature reviewed seems to prioritize nuanced accounts of both social processes in contested situations and the spaces and materialities that shape the contestation of a technology. For instance, in an exploration of medical quackery in South Africa, Hornberger [68] shows that even when using medically useless technologies, providers of these technologies can construct legitimacy for their patients because of contextual elements, such as lower education levels, local beliefs, and problems accessing proven medical treatments. Spaces, materialities, and relativity appear to be especially important in articles about disconnects between policy and practice in professional settings [73]. For example, Perrotta and Geampana [7] explore how in vitro fertilization (IVF) professionals navigate the disconnect between the standards of EBM and the enthusiasm of patients for the latest technologies. In another account of IVF, Barnreuther [84] notes that because the research took place in a physician private home in India, as opposed to an institutional setting in a Western country, the resulting technical success was deemed illegitimate and went unrecognized in both medicine and research for decades. We find that these kinds of nuanced accounts present a deeper and more empirically grounded understanding of the meaning of legitimacy in context, despite a lack of theorization.

How Is Legitimacy Produced?

Our critical interpretation of the MAS literature emphasizes the intertwined nature of practices, discourses, social processes, and materialities in the production of legitimacy. Authors often seem to assume these are impossible to separate from each other without losing meaning. For instance, in a study of the use of restraint in a hospital psychiatry ward, McKeown et al [74] examine how boundaries of space (ie, office or patient area), identity (ie, health care professional or patient), and diagnosis work together to legitimize or delegitimize the use of force in context. In this way, legitimacy is defined, not merely influenced, by the environment in which it is generated [75].

When legitimacy is conceptualized as a power problem, categorization seems to solidify the standing of already-legitimate actors. Meanwhile, power appears to become less accessible for the illegitimate when professional boundaries are unclear [69,76-78] or impermeable to outsiders [8,79,80]. Common MAS topics include contested illnesses and attempts of patients to receive treatment [70,80,81] and role negotiations of nonmedical practitioners within the medical establishment [69,79,82-84]. In both cases, power imbalance is usually framed as the primary obstacle to legitimacy, which is necessary for a more egalitarian redistribution of power. Although some authors

XSL•FO
RenderX

argue it is possible to have some power without legitimacy [72,82,84], the physical location of institutions, people, and technologies still defines flows of resources, manifesting in, for instance, differential treatment of patients at a rural hospital that would be unlikely at an urban hospital due to differences in hospital conditions rather than patient deservingness [85]. On the basis of these examples, our interpretation is that MAS and OMS have some overlap in this area in that both disciplines seem to approach legitimacy as an issue of access to resources.

How Is Legitimacy Used?

MAS literature reviewed often examines how legitimacy can flow through the taken-for-granted artifacts of professionalism. These artifacts include prescriptions [78] and complicated-looking fake medical machines [68]. Because the object itself is imbued with legitimacy through specific qualities in the professional context (eg, an appearance of being "official" or "high tech"), artifacts in these articles seem to reinforce the legitimacy of the pharmacist writing the prescription or the practitioner using the machine. Material artifacts may also include larger systems such as old IT infrastructures or preexisting contracts that impede desired change within institutions [73], implicitly reinforcing the legitimacy of institutional systems. While these materialities impact potential changes in legitimacy, they are usually not the subject of debate in the MAS literature. Therefore, it seems that the legitimacy of the prescription-writing pharmacist may be contested, but the prescription itself usually is not [78].

Summary

Our critical interpretation of the MAS literature is that legitimacy in this discipline is a relative and context-dependent concept. Legitimacy is often understood to be intertwined with power relations, but authors disagree about the nature of this connection and its implications. Finally, MAS articles center sites of controversy in which the subject has already been deemed illegitimate by focusing on space, materialities, and relationships. MAS authors appear to use these lenses to question "default" narratives of legitimacy in favor of exploring how and why a technology, system, group, or experience came to be viewed as illegitimate.

Here we develop our final disciplinary synthetic construct. Our interpretive analysis found that in MAS, legitimacy is physical, material, and always relative. Therefore, we employ the metaphor of a building's plumbing: each part of the system, from the pipes to the faucets, performs a different function, which defines both their physical form and location. No two buildings can have identical plumbing systems, because no two buildings, even if constructed similarly, are situated in exactly the same location—thus, they must connect to the local environment in different ways. In much the same way, the materialities of health and technology, and how these materialities interact with spaces and actors, are considered to define the legitimacy of health and technology in MAS. A "quack" technology reads as legitimate instead of "fake" because of both its physical form and the spatial context it inhabits [68]; an innovation such as IVF in the "wrong" space may never be recognized as innovation at all [84]. In both cases, if we understand legitimacy to be an infrastructure that is dependent on both space and materialities, we can better connect the different underlying systems and elements that comprise legitimacy for health technologies. A sink connected to a sewer line requires multiple specific parts, arranged in a particular way, running through a particular location, to function. In the cases noted earlier, materials, the spaces in which they are developed, and the spaces in which they are used must be arranged in particular ways for technologies to be considered legitimate, whether the audience is the international scientific establishment [84] or residents of a rural South African community [68]. The metaphor of a plumbing system emphasizes the categorical importance of context and physical materials in the construction of legitimacy: in MAS, legitimacy cannot exist without these. Therefore, this metaphor also helps us understand why MAS generally avoids comprehensive definitions or theorization of legitimacy; other than a few basic building blocks, systems of legitimacy for MAS are essentially viewed as nontransferrable between contexts.

Synthesizing Argument: Legitimacy as Social Infrastructure

Legitimacy is omnipresent and important within all 3 disciplines. However, it is often undertheorized and the subject of many implicit assumptions across MAS and STS literature. In OMS, conversely, increasingly detailed theorization results in the proliferation of "subtypes" of legitimacy with limited empirical support (see Table 1 for more details). Our synthesizing argument allows us to bring these different disciplinary traditions into conversation with each other through a new conceptual framework for legitimacy.

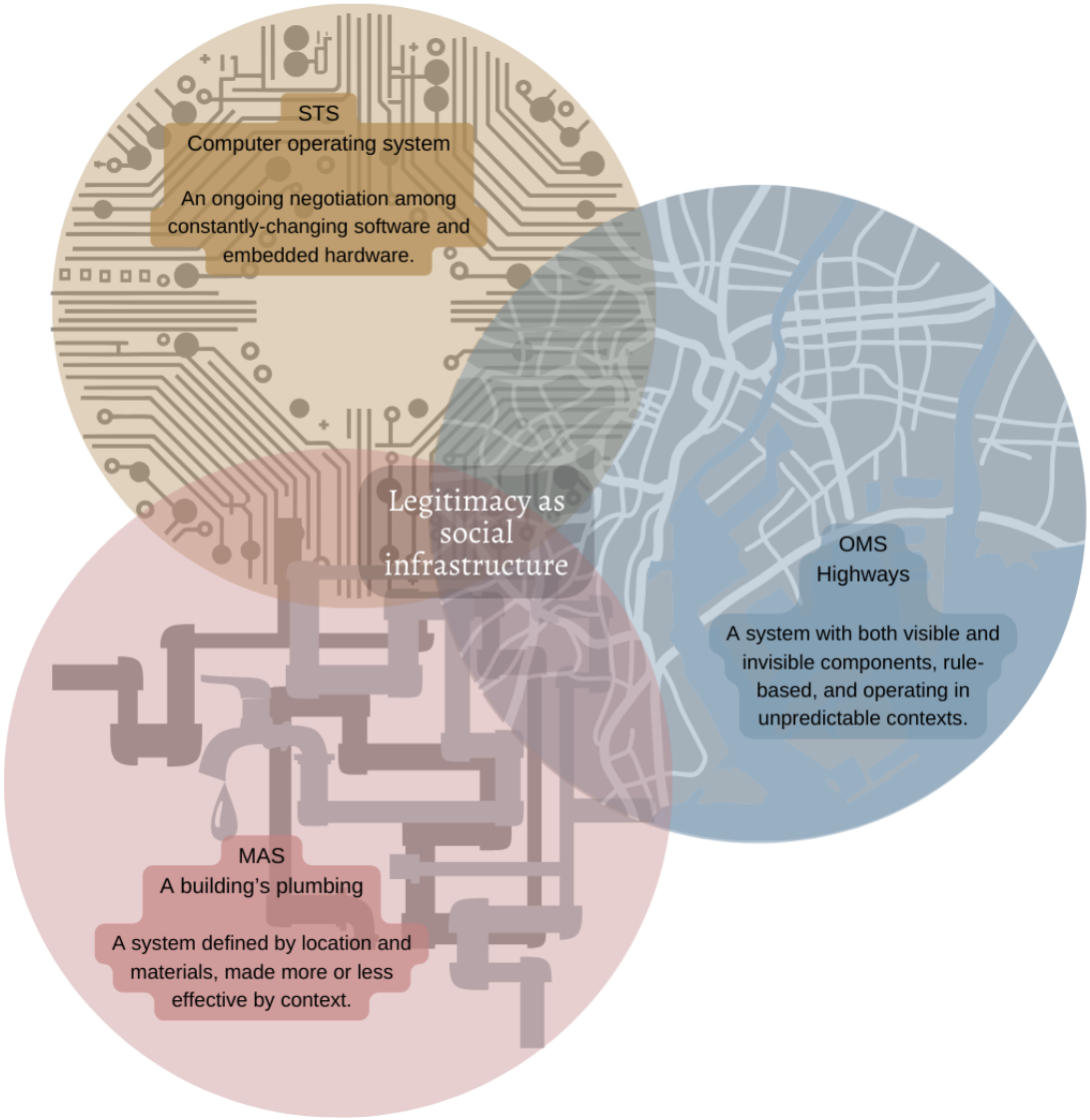
Using the 3 disciplinary metaphors we developed, we propose that legitimacy can be conceptualized as a social infrastructure, based on the criteria for an infrastructure by Star and Ruhleder [133]. If we combine the metaphors of a highway system (OMS), computer operating system (STS), and plumbing system (MAS), we begin to see the building blocks of an infrastructure that can support an entire city (Figure 2). Each metaphor shows a particular element of city infrastructure (ie, understanding of legitimacy) clearly. However, similar to a layered map of a city's interconnecting infrastructures, we gain a fuller picture of how legitimacy works in health and technology when we layer the metaphors to view their intersections and divergences. Knowing how computer systems monitor traffic and manage stoplights helps us better understand both computer systems and highways; similarly, an STS understanding of how unexpected assemblages may work with taken-for-granted processes can help us better understand how these unexpected assemblages may impact theory about taken-for-granted processes in practice within OMS. Therefore, such a layered map allows for a richer conceptualization in which there is room for different approaches to legitimacy without denying the validity of any single conceptualization.

XSL•FO
RenderX

Table 1. Disciplinary characteristics: legitimacy in organization and management studies (OMS), science and technology studies (STS), and medical anthropology and sociology (MAS).

	OMS	STS	MAS
Theoretical foundations	Major focus on theory, granular definitions of legitimacy common; normative legitimacy most prevalent	Some theoretical exploration but limited theoretical cohesion	Limited theoretical exploration, reliance on granular definitions based on empirical evidence
Conceptualized as	A process (of negotiation); perception; or asset, and sometimes as several of these at once.	A product of knowledge coproduction; a negotiation among human and nonhuman actors, contexts, and epistemic cultures	A relative and context-dependent concept intertwined with power relationships
Purpose of literature	Describe how organizations can move from illegitimate to legitimate status	Investigate instances of controversy and contestation	Explore how and why a subject came to be viewed as illegitimate
Most important related concepts	Taken-for-grantedness; categorization and standardization	Relationships, taken-for-grantedness, and unexpected assemblages	Materialities, relationships, and space

Figure 2. Disciplinary metaphors. MAS: medical anthropology and sociology; OMS: organization and management studies; STS: science and technology studies.



Similar to a physical infrastructure, legitimacy can be understood as a binding fabric that makes possible certain types of relationships among people, institutions, and technologies. In our conceptualization, legitimacy is an inherently relational concept: it can only exist between two or more entities, rather than within the perceptions of a single person or institution. The notion of social infrastructure makes space for legitimacy to exist outside of human judgments, without devaluing the importance of perception in the production of legitimacy. Just as infrastructures consist of both materialities and abstract systems (eg, roads are made of concrete, but their use is facilitated by traffic laws), social infrastructure melds semiotic,

XSL•FO
RenderX

material, and normative aspects of legitimacy and explores how they might work together. Infrastructure thus allows for the integration of disparate and seemingly contradictory aspects of making legitimacy. While Star [134] has emphasized that infrastructure is inherently a "relational property," we use the word "social" to focus on how legitimacy supports specific types of relationships among people, institutions, and technologies; in particular, how these relationships are shaped by implicit and intangible norms and values [135]. For instance, understanding the nuances and norms of pharmacists' relationships with both physicians and health care institutions reveals why a simple policy change allowing pharmacists to prescribe medication legally is not enough to legitimize the practice [78]. This moves us away from conversations about legitimacy as a conduit for resources and toward more clarity about how legitimacy acts as a binding fabric within social relationships.

Descriptions of legitimacy across all 3 disciplines fit the criteria for an infrastructure by Star and Ruhleder [133]. For instance, legitimacy is usually only visible "upon breakdown" [133]. Contestation and controversy are the only sites of legitimacy exploration in MAS and STS; in OMS, legitimacy is discussed in terms of desired change in legitimacy status (ie, moving from illegitimacy to legitimacy), and contestation is captured explicitly through concepts such as "liability of newness" [53,54]. In addition, maintenance of this social infrastructure happens through continued adherence to legitimizing protocols and norms within relationships; deviations can result in a crisis of legitimacy. In this sense, legitimacy is characterized by a strong fit with predefined categories and "standards" [133] in many articles [45,53,65,85], even when legitimacy itself is not defined. In our conceptualization of legitimacy, norms and values make up the "installed base" into which legitimacy is embedded [133] and are responsible for legitimacy's taken-for-granted nature.

Legitimacy also acts similar to infrastructure, in that it can be used to invisibly support the work of organizations, technologies, or people [55,86,133], often by allowing for acquisition of resources or power [69,87]. In fact, values and norms are often so deeply entrenched and invisible that even the infrastructure supporting them (ie, legitimacy) becomes taken-for-granted and invisible itself. This combination can make it extremely difficult for any contradictory experiences or knowledge to find expression within a society's taken-for-granted, unquestioned truths [85,88,89,132]. Therefore, it is understandable why the moments of crisis when legitimacy is "visible upon breakdown" usually occur when norms and values clash with each other. For example, in the exploration of IVF by Perotta and Geampana [7], new reproductive technologies are a site of legitimacy crisis because they are heavily desired by patients but cannot adhere to standards of EBM. In this situation, physicians must navigate among the medical value of providing patients with the care they seek, the medical value of providing evidence-based care, and the medical norm of adhering to preset medical standards that may not have kept up with available care options.

Finally, the origins of legitimacy parallel those of any other infrastructure. Legitimacy is built on existing social systems: it "wrestles with the inertia of the installed base" [133] and borrows both strengths and weaknesses from that base, resulting in legitimation by once-functional mechanisms that are now relics. For instance, Sheard et al [73] show how desired changes in practices in a hospital setting are stalled by existing social, technical, and bureaucratic systems and entrenched social practices, despite a lack of effectiveness of these entrenched systems and a desire for change among the people acting within them. Overcoming this systemic inertia requires the integration of new practices and systems into the installed base, regardless of the current utility of that base; attempts to completely overhaul entrenched institutional practices or technologies without attention to the installed base rarely succeed.

In all 3 metaphors, infrastructure serves as a system through which different elements that contribute to legitimacy can negotiate with one another. A social infrastructure conceptualization reflects the nature and parameters of these negotiations. Thus this lens allows us to see where different perspectives on legitimacy could provide a more holistic picture of how legitimacy works in practice, encouraging a collaborative approach to conflicting views on legitimacy.

We hope this conceptualization can provide a useful lens to better understand how legitimacy is produced and how it works among institutions, technologies, and people. We aim to move the conversation toward a more holistic perspective on embedding technologies in health care that goes beyond models based primarily on judgments and perceptions.

Discussion

Principal Findings

Our conceptualization of legitimacy as social infrastructure allows for deeper analysis and understanding of the meaning, production, and uses of legitimacy when embedding technologies in health care. Our review contributes to the OMS literature by bringing in outside perspectives. Cross-fertilization among the 3 disciplines reigns in the proliferation of legitimacy theory in OMS while providing a better connection with empirical research. STS and MAS literature benefit from grounding empirical exploration within a more cohesive theoretical framework. Our review also helps contextualize studies of legitimacy that are more limited in scope; for instance, while discourse analyses of legitimacy are useful, this review demonstrates that discourse comprises just one aspect of legitimacy production [67,90,91]. By allowing for many definitions of legitimacy, the notion of legitimacy as social infrastructure is adaptable enough to produce meaning across disciplines while providing ways to explore relational, material, and semiotic aspects. By providing a framework through which to explore many conceptualizations and understandings of legitimacy, we hope LSI can facilitate fruitful cross-disciplinary collaborations of legitimacy researchers. Specifically, LSI allows different conceptualizations to coexist without competing with one another, providing a framework in which researchers can build on the work of other disciplines without having to disregard the traditions of their own discipline. In addition, this framework provides value to acceptance, adoption, and implementation researchers and other stakeholders interested

XSL•FO
RenderX

in better explaining how technologies are embedded in health care. Through an interdisciplinary understanding of legitimacy and a focus on relationships and moments of breakdown, LSI provides a mechanism to study technology embedding both contextually and over time. In this way, LSI produces a more comprehensive understanding of how the interconnected social, medical, environmental, and institutional elements surrounding a technology may interact in ways that impact legitimacy for a given technology, and thus, whether a technology will become embedded in health care.

Embedding Beyond Current Frameworks

Legitimacy as social infrastructure provides an alternative to current models describing how technologies are embedded in health care. We can use this conceptualization to account for gaps between policy intentions and practices when embedding technologies. Social infrastructure produces a more comprehensive understanding of the issues underpinning technology adoption than frameworks that hinge on trust [136,137], transparency [4,137], and acceptability [19], especially when those frameworks are based on behavioral intention [20,21,138]. Articles included in this review have shown that the complexities of embedding technologies in health care often clash with the best-intentioned policy choices because these frameworks generally do not account for the full scope of legitimizing processes within health care delivery systems. This is particularly evident in articles that focus on contested illnesses [80,89,92] and professional boundary work negotiations in light of new policies or collaborations in health care [47,56,77,78,93,94]. For example, Weiss and Sutton [78] demonstrate that even though policy changes explicitly allow pharmacists to prescribe medication, actually doing so in practice involves a more complex negotiation of roles and hierarchies to legitimize pharmacist prescribing activities. The new prescribing policy, based on notions of trust and acceptability, is not enough. For patients whose contested illnesses do not meet EBM criteria [80], negotiations and individual relationships are essential to access any kind of medical care [95]. Here, EBM criteria designed to increase transparency and trust in medicine prevent patients with real symptoms from accessing necessary care.

The many technology acceptance and adoption frameworks available currently have difficulty explaining adequately how and why some technologies are embedded in health care but others are not; in other words, “the literature does not specify the conditions for full use to be achieved” [138]. Similar to legitimacy, these terms are not clearly defined across the literature [138]: “acceptance” and “acceptability” generally are defined in terms of perceptions of potential users [138,139] and separate from both “adoption” and “appropriation,” which deal with the processes and practices of using the technology [138]. Other models aim to identify factors contributing to acceptance of new technologies. TAM is the most well-known example; it has been adapted and added to by numerous authors since its inception [20,21,138,140]. However, similar to most adoption and acceptance frameworks, TAM is based on behavioral intention, which has limited utility in predicting actual behavior [141,142] and does not generally attempt to integrate other systemic or relational aspects of technology acceptance. Within this review, Reay et al [56] demonstrated that gaining discursive buy-in, the focus of behavioral intention models of technology adoption, “is not enough” to legitimize new practices in a clinical setting without action and interpersonal, structural support.

Other frameworks have attempted to integrate sociocultural aspects of technology acceptance, specifically in low- and middle-income countries, although acceptance still generally focuses on individual users rather than systemic or relational factors [143]. More unusually, the nonadoption, abandonment, scale-up, spread, and sustainability framework by Greenhalgh et al [22], which is explicitly directed at health technologies, attempts to fill gaps left by TAM and other models by focusing on nonadoption and abandonment of technology instead of acceptance. This framework incorporates norms and values to some extent but does not focus specifically on legitimacy and the impact of norms and values on technology development.

The legitimacy as social infrastructure framework (LSI framework) is different from these models: rather than starting with adoption and acceptance of technologies, our conceptualization focuses on the relationships and negotiations that result in adoption and acceptance or not. If we take legitimacy as a necessary element in technology adoption and acceptance, conceptualizing legitimacy as social infrastructure allows us to understand how adoption, acceptance, and implementation frameworks may work together to explain the process of embedding technologies in health care. For instance, rather than conducting acceptance, adoption, and implementation studies separately in a health technology development project, researchers could use the conceptualization of legitimacy as social infrastructure as a framework to iteratively explore the binding fabric of relationships and legitimacy narratives that make up the landscape in which a new technology is being developed and provide advice to decision makers based on these findings. This focus would also allow project stakeholders to contextualize the results of a traditional acceptability or adoption study to better support the end goal of successfully embedding a technology in health care. Doing so would move conversation beyond simple adoption and acceptance and toward a more holistic perspective about this embedding process based instead on an analysis of legitimacy.

Focus on Relationships and Narratives

Momentum for embedding technologies in health care often fades due to unforeseen complexities [144] when moving technologies from laboratories to real-world environments. Particularly for mobile health technologies in which social and relational elements play a heightened role, the LSI framework could prove extremely useful in both anticipating and navigating complexity in the real world by describing changing relationships and narratives. Legitimacy as social infrastructure could provide a better blueprint that takes contextual factors into account earlier and avoids some of the acceptance and adoption bottlenecks that develop in linear models.

The struggles and unintended consequences of embedding technology in practice have been well-documented in STS and OMS [58,61,93,96]. The LSI framework conceptualizes how these documented struggles impact legitimacy and, thus,

XSL•FO
RenderX

technology adoption and acceptance. STS understands new technologies (including those designed for health care) to be both directional and open-ended [23]. Technologies are embodied with certain scripts and ideas, produced both through the development process and purposely in the minds of developers and other stakeholders, impacting how they can be used and when. However, the scripts managing this directionality never completely determine how a technology is used in practice [23,145]. Integrating technologies into real-world environments necessarily involves tinkering [22,23,144] by users, developers, and other stakeholders, often contradicting a technology's directionality entirely. For instance, as Perrotta and Geampana [7] demonstrate, new medical technology may be used in ways that contradict its original EBM-derived script, given a particular set of circumstances generated by patients, economics, and the pace of technology development. What happens next is a negotiation that determines how real-world context, practices, and technology scripts interact to generate a legitimacy relationship. Conceptualizing legitimacy as social infrastructure allows us to both understand the nature of these interactions and how they may or may not lead to technology adoption and acceptance in a health care setting. For instance, Sheard et al [73] discuss the complex systems and legitimacy relationships that impact the ability of care providers in the United Kingdom to act on patient feedback; this feedback is solicited in the first place to increase trust in and acceptability of the medical system. If we view patient feedback frameworks as a form of health care technology, legitimacy as social infrastructure helps us better understand and anticipate the relative importance of backgrounded relational and structural systems [73] that determined whether patient feedback was incorporated successfully in the health care institutional systems under study. An infrastructural understanding reveals how the 3 issues identified (ie, normative legitimacy, structural legitimacy, and organizational readiness [14,73]) work together to support or deter change in a complex organizational setting. A binding fabric of relationships ties all of these issues together and clarifies why the quality of certain relationships is more important than others. This focus also adds to understanding of the phenomenon of "institutional entrepreneurs." Relationships likely play a role in how institutional entrepreneurs avoid being "confined by the status quo," allowing them to change organizations from within despite legitimacy challenges for others [73].

Future Research and Applications

Given the utility of this conceptualization in bringing together 3 very different disciplinary traditions around legitimacy, future research could expand on this work to incorporate perspectives from disciplines, such as ethics and law, that we could not include here. In addition, it is possible that legitimacy is usually not explicitly conceptualized in medicine and psychology due to an implicit discipline-wide understanding of the values that underpin the concept. In other words, norms and legitimacy in medicine and psychology may be so closely intertwined with the goal of providing care that legitimacy may not be contested or questioned from these disciplinary perspectives. These kinds of taken-for-granted assumptions make it important to explore the origins, production, and meanings of legitimacy more thoroughly in the psychology and medical fields. Our conceptualization could allow for such investigation in fields in which the language describing legitimacy is either absent or perceived to be self-evident.

Potential applications of legitimacy as social infrastructure in practice have a wide range thanks to the flexibility and adaptability of the conceptualization. For example, to manage risks appropriately, regulators and others must have knowledge and understanding of the subject of regulation, in this case, health care technologies. This includes not only technical knowledge, but epistemological understanding of what is seen to constitute valid forms of knowledge, what is being left out, and "system knowledge of the context regulation is operating in" [146]. By approaching the embedding of technologies in health care as a result of knowledge coproduction, the notion of legitimacy as social infrastructure can provide deeper understandings of the systems and context of health technologies. It can then help regulators differentiate valid from invalid forms of knowledge in context. Regulators also experience legitimacy concerns themselves, in that regulation must be legitimate to function [9,147]. An understanding of legitimacy as social infrastructure could provide regulators with clarity about the practices, decisions, and structures that may impact their own legitimacy. For instance, knowing more about how their relationships with various stakeholders contribute to their legitimacy could help regulators take stakeholders' viewpoints into account at the most impactful moments and more appropriately narrate their decisions for different audience needs.

Limitations

Given the broad focus of our topic "legitimacy in health and technology," we included a heterogenous collection of empirical topics within each discipline. While this allowed for a broad discussion of legitimacy in health and technology, we did not produce one overarching definition of legitimacy; rather, we developed a conceptualization that can be used as a framework with the diversity of definitions already available. Similarly, the disciplinary boundaries are somewhat "fuzzy" [24]: OMS is a subdiscipline of sociology, STS overlaps with both MAS and OMS, and MAS is a combination of 2 disciplines with similar approaches. We found these disciplinary constructs useful because, despite their overlap, they highlight substantial differences among social science approaches to legitimacy.

In addition to the legal and ethical boundaries previously described, in this review, we did not engage with several other topics related to the legitimacy of technologies in health care. While the word "legitimacy" is used frequently in medicine and health economics, very few articles in these fields define the concept or explore it in any depth. Rather, the vast majority use the word uncritically within a surface explanation of the importance of a problem, either by claiming without evidence that legitimacy is an asset held by organizations, technologies, or people [148-150] or by mentioning that the legitimacy of a technology is under threat but without fully discussing why or how [151,152]. When legitimacy is defined, it is often reduced to a series of related concepts, such as "expertise, reciprocity, and trustworthiness," without discussing how these interact to

XSL•FO
RenderX

create legitimacy [153], although there are a few exceptions [9]. Due to the general lack of critical exploration, medicine and health economics were excluded as disciplines. This is an area that would benefit from further research using the LSI framework.

For similar reasons, we focused our search strategy within journals known to provide deep conceptualization of legitimacy (see Multimedia Appendix 1 for further details). Given the abstract and conceptual nature of our subject, a database search proved inadequate for finding all relevant articles; however, searching only within specific journals, despite rigorous selection, may have biased study findings. We feel the benefits of closely adhering to CIS protocol in this manner outweighed this issue due to our study's interpretive goals and the abstract subject matter.

Conclusions

The notion of legitimacy as social infrastructure provides a deeper understanding of how and why technologies are or are not embedded in health care. Social infrastructure focuses on the aspects of legitimacy related to norms, semiotics, materialities, and specific relationships among stakeholders and institutional contexts without excluding or privileging one form of knowledge production over the other. Through this conceptualization, the production and use of legitimacy can be studied ethnographically by following the social infrastructural networks and relationships that constitute and maintain legitimacy. By moving beyond judgment- and perception-based frameworks, legitimacy as social infrastructure approaches the problem of embedding technologies in health care through the lens of knowledge coproduction. The LSI framework enables empirical data collection and analysis that transcends disciplinary boundaries without negating the discrete lines of theoretical thought of any one discipline.

Given the interdisciplinary nature of health care policy and practice and the wide range of stakeholders involved, clearer and deeper communication around legitimacy could have far-reaching impacts. This could include new frameworks explaining technology adoption and acceptance, greater consensus about what makes a health care technology or policy decision legitimate, and strategic direction for decision makers working with disruptive health care technologies within entrenched systems.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Method details: selection of journals, disciplines, and articles.
[PDF File (Adobe PDF File), 234 KB - humanfactors_v12i1e48955_app1.pdf]

Multimedia Appendix 2
Methods: disciplinary network visualization.
[PDF File (Adobe PDF File), 529 KB - humanfactors_v12i1e48955_app2.pdf]

References

1. Ancker JS, Kern LM, Abramson E, Kaushal R. The Triangle Model for evaluating the effect of health information technology on healthcare quality and safety. J Am Med Inform Assoc 2012;19(1):61-65 [FREE Full text] [doi: 10.1136/amiajnl-2011-000385] [Medline: 21857023]
2. McDougall A, Goldszmidt M, Kinsella EA, Smith S, Lingard L. Collaboration and entanglement: an actor-network theory analysis of team-based intraprofessional care for patients with advanced heart failure. Soc Sci Med 2016 Sep;164:108-117 [FREE Full text] [doi: 10.1016/j.socscimed.2016.07.010] [Medline: 27490299]
3. Mittelstadt BD, Allo P, Taddeo M, Wachter S, Floridi L. The ethics of algorithms: mapping the debate. Big Data Soc 2016 Dec 01;3(2). [doi: 10.1177/2053951716679679]
4. de Fine Licht K, de Fine Licht J. Artificial intelligence, transparency, and public decision-making: why explanations are key when trying to produce perceived legitimacy. AI Soc 2020 Mar 19;35(4):917-926. [doi: 10.1007/s00146-020-00960-w]
5. Lehoux P, Williams-Jones B, Grimard D, Proulx S. Technologies of the self in public health: insights from public deliberations on cognitive and behavioural enhancement. Crit Public Health 2017 Mar 15;27(3):373-383. [doi: 10.1080/09581596.2017.1300637]
6. Fishman JR, Flatt MA, Settersten RA. Bioidentical hormones, menopausal women, and the lure of the "natural" in U.S. anti-aging medicine. Soc Sci Med 2015 May;132:79-87 [FREE Full text] [doi: 10.1016/j.socscimed.2015.02.027] [Medline: 25795991]
7. Perrotta M, Geampana A. The trouble with IVF and randomised control trials: professional legitimation narratives on time-lapse imaging and evidence-informed care. Soc Sci Med 2020 Aug;258:113115 [FREE Full text] [doi: 10.1016/j.socscimed.2020.113115] [Medline: 32593117]

XSL•FO
RenderX

8. Petersen A, Tanner C, Munsie M. Between hope and evidence: how community advisors demarcate the boundary between legitimate and illegitimate stem cell treatments. Health (London) 2015 Mar 02;19(2):188-206. [doi: 10.1177/1363459314555240] [Medline: 25367895]
9. Cleemput I, Franken M, Koopmanschap M, le Polain M. European drug reimbursement systems' legitimacy: five-country comparison and policy tool. Int J Technol Assess Health Care 2012 Sep 17;28(4):358-366. [doi: 10.1017/s0266462312000529]
10. Russell J, Greenhalgh T. Affordability as a discursive accomplishment in a changing National Health Service. Soc Sci Med 2012 Dec;75(12):2463-2471. [doi: 10.1016/j.socscimed.2012.09.026] [Medline: 23103349]
11. Ward JK, Cafiero F, Fretigny R, Colgrove J, Seror V. France's citizen consultation on vaccination and the challenges of participatory democracy in health. Soc Sci Med 2019 Jan;220:73-80 [FREE Full text] [doi: 10.1016/j.socscimed.2018.10.032] [Medline: 30408684]
12. Wamsiedel M. Reasonableness: legitimate reasons for illegitimate presentations at the ED. Sociol Health Illn 2018 Nov 10;40(8):1347-1360. [doi: 10.1111/1467-9566.12776] [Medline: 29989187]
13. Haack P, Pfarrer MD, Scherer AG. Legitimacy - as - feeling: how affect leads to vertical legitimacy spillovers in transnational governance. J Manag Stud 2013 Dec 23;51(4):634-666. [doi: 10.1111/joms.12071]
14. Suchman MC. Managing legitimacy: strategic and institutional approaches. Acad Manag Rev 1995 Jun;20(3):571-610 [FREE Full text] [doi: 10.2307/258788]
15. Lövbrand E, Pielke RJ, Beck S. A democracy paradox in studies of science and technology. Sci Technol Hum Values 2010 Aug 26;36(4):474-496. [doi: 10.1177/0162243910366154]
16. Selin C. Expectations and the emergence of nanotechnology. Sci Technol Hum Values 2007 Mar 01;32(2):196-220. [doi: 10.1177/0162243906296918]
17. Fredriksson M, Tritter JQ. Disentangling patient and public involvement in healthcare decisions: why the difference matters. Sociol Health Illn 2017 Jan 11;39(1):95-111. [doi: 10.1111/1467-9566.12483] [Medline: 27862007]
18. Wehrens R, Sihag V, Sülz S, van Elten H, van Raaij E, de Bont A, et al. Understanding the uptake of big data in health care: protocol for a multinational mixed-methods study. JMIR Res Protoc 2020 Oct 22;9(10):e16779 [FREE Full text] [doi: 10.2196/16779] [Medline: 33090113]
19. Dwivedi YK, Shareef MA, Simintiras AC, Lal B, Weerakkody V. A generalised adoption model for services: a cross-country comparison of mobile health (m-health). Gov Inf Q 2016 Jan;33(1):174-187. [doi: 10.1016/j.giq.2015.06.003]
20. Yarbrough AK, Smith TB. Technology acceptance among physicians: a new take on TAM. Med Care Res Rev 2007 Dec 05;64(6):650-672. [doi: 10.1177/1077558707305942] [Medline: 17717378]
21. Holden RJ, Karsh BT. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010 Feb;43(1):159-172 [FREE Full text] [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]
22. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
23. Carboni C, Wehrens R, van der Veen R, de Bont A. Conceptualizing the digitalization of healthcare work: a metaphor-based Critical Interpretive Synthesis. Soc Sci Med 2022 Jan;292:114572 [FREE Full text] [doi: 10.1016/j.socscimed.2021.114572] [Medline: 34839086]
24. Dixon-Woods M, Cavers D, Agarwal S, Annandale E, Arthur A, Harvey J, et al. Conducting a critical interpretive synthesis of the literature on access to healthcare by vulnerable groups. BMC Med Res Methodol 2006 Jul 26;6:35 [FREE Full text] [doi: 10.1186/1471-2288-6-35] [Medline: 16872487]
25. Suddaby R, Bitektine A, Haack P. Legitimacy. Acad Manag Ann 2017 Jan;11(1):451-478. [doi: 10.5465/annals.2015.0101]
26. Miller CA, Smith-Doerr L, Fouché R, Felt U. The Handbook of Science and Technology Studies. Cambridge, MA: MIT Press; 2017.
27. Woolgar S. The turn to technology in social studies of science. Sci Technol Hum Values 1991 Jan 01;16(1):20-50. [doi: 10.1177/016224399101600102]
28. Latour B. Science in Action: How to Follow Scientists and Engineers Through Society. Cambridge, MA: Harvard University Press; 1987.
29. Conrad P. Parallel play in medical anthropology and medical sociology. Am Soc 1997 Dec;28(4):90-100. [doi: 10.1007/s12108-997-1021-4]
30. Dawson A, Verweij M. Public health and legitimacy: or why there is still a place for substantive work in ethics. Public Health Ethics 2014 Jul 14;7(2):95-97. [doi: 10.1093/phe/phu018]
31. Tomlinson M, Chopra M, Hoosain N, Rudan I. A review of selected research priority setting processes at national level in low and middle income countries: towards fair and legitimate priority setting. Health Res Policy Syst 2011 May 15;9(1):19 [FREE Full text] [doi: 10.1186/1478-4505-9-19] [Medline: 21575144]
32. Boada JP, Maestre BR, Genís CT. The ethical issues of social assistive robotics: a critical literature review. Technol Soc 2021 Nov;67:101726. [doi: 10.1016/j.techsoc.2021.101726]
33. Crawford L. Moral legitimacy: the struggle of homeopathy in the NHS. Bioethics 2016 Feb 25;30(2):85-95 [FREE Full text] [doi: 10.1111/bioe.12227] [Medline: 26806451]

XSL•FO
RenderX

34. Morley J, Machado CC, Burr C, Cowls J, Joshi I, Taddeo M, et al. The ethics of AI in health care: a mapping review. Soc Sci Med 2020 Sep;260:113172. [doi: 10.1016/j.socscimed.2020.113172] [Medline: 32702587]
35. Baeroe K, Baltussen R. Legitimate healthcare limit setting in a real-world setting: integrating accountability for reasonableness and multi-criteria decision analysis. Public Health Ethics 2014 Mar 14;7(2):98-111. [doi: 10.1093/phe/phu006]
36. Schmidt VA. Democracy and legitimacy in the European Union revisited: input, output and 'throughput'. Political Stud 2012 Aug 06;61(1):2-22. [doi: 10.1111/j.1467-9248.2012.00962.x]
37. Martin GP. Representativeness, legitimacy and power in public involvement in health-service management. Soc Sci Med 2008 Dec;67(11):1757-1765. [doi: 10.1016/j.socscimed.2008.09.024] [Medline: 18922611]
38. Borst RA, Wehrens R, Bal R. Sustaining knowledge translation practices: a critical interpretive synthesis. Int J Health Policy Manag 2022 Dec 19;11(12):2793-2804 [FREE Full text] [doi: 10.34172/ijhpm.2022.6424] [Medline: 35279039]
39. Timmermans S, Tavory I. Theory construction in qualitative research: from grounded theory to abductive analysis. Soc Theory 2012 Sep 10;30(3):167-186. [doi: 10.1177/0735275112457914]
40. Berntsen GK, Gammon D, Steinsbekk A, Salamonsen A, Foss N, Ruland C, et al. How do we deal with multiple goals for care within an individual patient trajectory? A document content analysis of health service research papers on goals for care. BMJ Open 2015 Dec 10;5(12):e009403 [FREE Full text] [doi: 10.1136/bmjopen-2015-009403] [Medline: 26656243]
41. Ruef M, Scott WR. A multidimensional model of organizational legitimacy: hospital survival in changing institutional environments. Admin Sci Q 1998 Dec;43(4):877-904. [doi: 10.2307/2393619]
42. Haack P, Schilke O, Zucker L. Legitimacy revisited: disentangling propriety, validity, and consensus. J Manag Stud 2020 Aug 15;58(3):749-781. [doi: 10.1111/joms.12615]
43. Provan KG, Lamb G, Doyle M. Building legitimacy and the early growth of health networks for the uninsured. Health Care Manage Rev 2004;29(2):117-128. [doi: 10.1097/00004010-200404000-00005] [Medline: 15192984]
44. Wilmoth S. Foundation trusts and the problem of legitimacy. Health Care Anal 2004 Jun;12(2):157-169. [doi: 10.1023/b:hcan.0000041188.79809.05] [Medline: 15487816]
45. Fisher G, Kuratko DF, Bloodgood JM, Hornsby JS. Legitimate to whom? The challenge of audience diversity and new venture legitimacy. J Bus Ventur 2017 Jan;32(1):52-71. [doi: 10.1016/j.jbusvent.2016.10.005]
46. Kitchener M. Mobilizing the logic of managerialism in professional fields: the case of academic health centre mergers. Organ Stud 2002 May 01;23(3):391-420. [doi: 10.1177/0170840602233004]
47. Hoff TJ. The paradox of legitimacy: physician executives and the practice of medicine. Health Care Manage Rev 1999;24(4):54-64. [doi: 10.1097/00004010-199910000-00007] [Medline: 10572789]
48. Girschik V. Managing legitimacy in business‐driven social change: the role of relational work. J Manag Stud 2020 Jan 07;57(4):775-804. [doi: 10.1111/joms.12544]
49. Patala S, Korpivaara I, Jalkala A, Kuitunen A, Soppe B. Legitimacy under institutional change: how incumbents appropriate clean rhetoric for dirty technologies. Organ Stud 2017 Dec 17;40(3):395-419. [doi: 10.1177/0170840617736938]
50. Cattani G, Ferriani S, Lanza A. Deconstructing the outsider puzzle: the legitimation journey of novelty. Organ Stud 2017 Dec;28(6):965-992. [doi: 10.1287/orsc.2017.1161]
51. Alexy O, George G. Category divergence, straddling, and currency: open innovation and the legitimation of illegitimate categories. J Manag Stud 2012 Dec 18;50(2):173-203. [doi: 10.1111/joms.12000]
52. Golant BD, Sillince JA. The constitution of organizational legitimacy: a narrative perspective. Organ Stud 2007 Aug 01;28(8):1149-1167. [doi: 10.1177/0170840607075671]
53. McKnight B, Zietsma C. Finding the threshold: a configurational approach to optimal distinctiveness. J Bus Ventur 2018 Jul;33(4):493-512. [doi: 10.1016/j.jbusvent.2018.03.004]
54. Garud R, Schildt HA, Lant TK. Entrepreneurial storytelling, future expectations, and the paradox of legitimacy. Organ Sci 2014 Oct;25(5):1479-1492. [doi: 10.1287/orsc.2014.0915]
55. Patriotta G, Gond JP, Schultz F. Maintaining legitimacy: controversies, orders of worth, and public justifications. J Manag Stud 2011 Mar 17;48(8):1804-1836. [doi: 10.1111/j.1467-6486.2010.00990.x]
56. Reay T, Goodrick E, Casebeer A, Hinings CR. Legitimizing new practices in primary health care. Health Care Manage Rev 2013;38(1):9-19. [doi: 10.1097/HMR.0b013e31824501b6] [Medline: 22269782]
57. O'Kane C, Mangematin V, Geoghegan W, Fitzgerald C. University technology transfer offices: the search for identity to build legitimacy. Res Policy 2015 Mar;44(2):421-437. [doi: 10.1016/j.respol.2014.08.003]
58. Granja R, Machado H. Forensic DNA phenotyping and its politics of legitimation and contestation: views of forensic geneticists in Europe. Soc Stud Sci 2023 Dec 30;53(6):850-868 [FREE Full text] [doi: 10.1177/0306312720945033] [Medline: 32729409]
59. Gieryn TF. Boundary-work and the demarcation of science from non-science: strains and interests in professional ideologies of scientists. Am Sociol Rev 1983 Dec;48(6):781-795 [FREE Full text]
60. Jacobson B. Constructing legitimacy in geoengineering discourse: the politics of representation in science policy literature. Sci Culture 2018 Apr 25;27(3):322-348. [doi: 10.1080/09505431.2018.1465910]
61. Barker KK. Listening to Lyrica: contested illnesses and pharmaceutical determinism. Soc Sci Med 2011 Sep;73(6):833-842. [doi: 10.1016/j.socscimed.2011.05.055] [Medline: 21840638]

XSL•FO
RenderX

62. McLevey J. Understanding policy research in liminal spaces: think tank responses to diverging principles of legitimacy. Soc Stud Sci 2015 Apr 30;45(2):270-293. [doi: 10.1177/0306312715575054] [Medline: 26477208]
63. Greco M. The ambivalence of error: "scientific ideology" in the history of the life sciences and psychosomatic medicine. Soc Sci Med 2004 Feb;58(4):687-696. [doi: 10.1016/s0277-9536(03)00220-x] [Medline: 14672585]
64. Brown N, Michael M. Switching between science and culture in transpecies transplantation. Sci Technol Hum Values 2001 Jan 01;26(1):3-22. [doi: 10.1177/016224390102600101] [Medline: 15362238]
65. Flink T, Kaldewey D. The new production of legitimacy: STI policy discourses beyond the contract metaphor. Res Policy 2018 Feb;47(1):14-22. [doi: 10.1016/j.respol.2017.09.008]
66. Cloatre E. Law and biomedicine and the making of 'genuine' traditional medicines in global health. Crit Public Health 2019 Jun 04;29(4):424-434 [FREE Full text] [doi: 10.1080/09581596.2019.1594696] [Medline: 31391708]
67. Foley L, Faircloth CA. Medicine as discursive resource: legitimation in the work narratives of midwives. Sociol Health Illn 2003 Mar;25(2):165-184 [FREE Full text] [doi: 10.1111/1467-9566.00330] [Medline: 14498938]
68. Hornberger J. Who is the fake one now? Questions of quackery, worldliness and legitimacy. Crit Public Health 2019 Jun 04;29(4):484-493. [doi: 10.1080/09581596.2019.1602719]
69. MacGregor C, Petersen A, Munsie M. From the margins to mainstream: how providers of autologous 'stem cell treatments' legitimise their practice in Australia. Health (London) 2021 Jan 13;25(1):51-68. [doi: 10.1177/1363459319846927] [Medline: 31081381]
70. Flinkfeldt M. 'Filling one's days': managing sick leave legitimacy in an online forum. Sociol Health Illn 2011 Jul 11;33(5):761-776 [FREE Full text] [doi: 10.1111/j.1467-9566.2011.01330.x] [Medline: 21561459]
71. Pushkar P, Tomkow L. Clinician-led evidence-based activism: a critical analysis. Crit Public Health 2020 Nov 27;31(2):235-244. [doi: 10.1080/09581596.2020.1841112]
72. Jones M, Pietilä I. Alignments and differentiations: people with illness experiences seeking legitimate positions as health service developers and producers. Health (London) 2020 May 16;24(3):223-240. [doi: 10.1177/1363459318800154] [Medline: 30222007]
73. Sheard L, Marsh C, O'Hara J, Armitage G, Wright J, Lawton R. The patient feedback response framework - understanding why UK hospital staff find it difficult to make improvements based on patient feedback: a qualitative study. Soc Sci Med 2017 Apr;178:19-27 [FREE Full text] [doi: 10.1016/j.socscimed.2017.02.005] [Medline: 28189820]
74. McKeown M, Thomson G, Scholes A, Jones F, Downe S, Price O, et al. Restraint minimisation in mental health care: legitimate or illegitimate force? An ethnographic study. Sociol Health Illn 2020 Mar 27;42(3):449-464 [FREE Full text] [doi: 10.1111/1467-9566.13015] [Medline: 31657030]
75. Busfield J. The moral economy of the pharmaceutical industry: legitimising prices. Health (London) 2021 May 17;25(3):271-287. [doi: 10.1177/1363459319879474] [Medline: 31623464]
76. Sanders T, Harrison S. Professional legitimacy claims in the multidisciplinary workplace: the case of heart failure care. Sociol Health Illn 2008 Mar 15;30(2):289-308 [FREE Full text] [doi: 10.1111/j.1467-9566.2007.01052.x] [Medline: 18290937]
77. Lees S, Palmer J, Procureur F, Blanchet K. Contested legitimacy for anthropologists involved in medical humanitarian action: experiences from the 2014-2016 West Africa Ebola epidemic. Anthropol Med 2020 Jun 04;27(2):125-143 [FREE Full text] [doi: 10.1080/13648470.2020.1742576] [Medline: 32363909]
78. Weiss MC, Sutton J. The changing nature of prescribing: pharmacists as prescribers and challenges to medical dominance. Sociol Health Illn 2009 Apr;31(3):406-421 [FREE Full text] [doi: 10.1111/j.1467-9566.2008.01142.x] [Medline: 19055585]
79. Mizrachi N. Epistemology and legitimacy in the production of anorexia nervosa in the journal Psychosomatic Medicine 1939–1979. Sociol Health Illn 2002 Jun 05;24(4):462-490. [doi: 10.1111/1467-9566.00304]
80. Zavestoski S, Brown P, McCormick S, Mayer B, D'Ottavi M, Lucove JC. Patient activism and the struggle for diagnosis: Gulf War illnesses and other medically unexplained physical symptoms in the US. Soc Sci Med 2004 Jan;58(1):161-175. [doi: 10.1016/s0277-9536(03)00157-6] [Medline: 14572929]
81. Mik-Meyer N, Obling AR. The negotiation of the sick role: general practitioners' classification of patients with medically unexplained symptoms. Sociol Health Illn 2012 Sep 05;34(7):1025-1038 [FREE Full text] [doi: 10.1111/j.1467-9566.2011.01448.x] [Medline: 22384857]
82. Nathan S, Stephenson N, Braithwaite J. Sidestepping questions of legitimacy: how community representatives manoeuvre to effect change in a health service. Health (London) 2014 Jan 30;18(1):23-40. [doi: 10.1177/1363459312473617] [Medline: 23364312]
83. Tovey P. Contingent legitimacy: U.K. alternative practitioners and inter-sectoral acceptance. Soc Sci Med 1997 Oct;45(7):1129-1133. [doi: 10.1016/s0277-9536(97)00010-5] [Medline: 9257404]
84. Bärnreuther S. Innovations 'out of place': controversies over IVF beginnings in India between 1978 and 2005. Med Anthropol 2016 Oct 05;35(1):73-89. [doi: 10.1080/01459740.2015.1094066] [Medline: 26436991]
85. Andersen HM. "Villagers": differential treatment in a Ghanaian hospital. Soc Sci Med 2004 Nov;59(10):2003-2012. [doi: 10.1016/j.socscimed.2004.03.005] [Medline: 15351468]
86. Latimer J. Commanding materials: (re)legitimating authority in the context of multi-disciplinary work. Sociology 2004 Oct 01;38(4):757-775. [doi: 10.1177/0038038504045863]

XSL•FO
RenderX

87. Kwak K, Yoon H. Unpacking transnational industry legitimacy dynamics, windows of opportunity, and latecomers' catch-up in complex product systems. Res Policy 2020 May;49(4):103954. [doi: 10.1016/j.respol.2020.103954]
88. Caiata-Zufferey M. Genetically at-risk status and individual agency. A qualitative study on asymptomatic women living with genetic risk of breast/ovarian cancer. Soc Sci Med 2015 May;132:141-148. [doi: 10.1016/j.socscimed.2015.03.037] [Medline: 25813728]
89. Paxman CG. "Everyone thinks I am just lazy": legitimacy narratives of Americans suffering from fibromyalgia. Health (London) 2021 Jan 20;25(1):121-137. [doi: 10.1177/1363459319857457] [Medline: 31216871]
90. Sapir A. Mythologizing the story of a scientific invention: constructing the legitimacy of research commercialization. Organ Stud 2019 Jan 28;41(6):799-820. [doi: 10.1177/0170840618814575]
91. Vaara E, Monin P. A recursive perspective on discursive legitimation and organizational action in mergers and acquisitions. Organ Sci 2010 Feb;21(1):3-22. [doi: 10.1287/orsc.1080.0394]
92. Clark S, Botterill LC. Contesting facts about wind farms in Australia and the legitimacy of adverse health effects. Health (London) 2018 Jul 01;22(4):337-355. [doi: 10.1177/1363459317693407] [Medline: 28401817]
93. Edwards KT. Methods of legitimation: how ethics committees decide which reasons count in public policy decision-making. Soc Sci Med 2014 Jul;113:34-41. [doi: 10.1016/j.socscimed.2014.04.043] [Medline: 24833251]
94. Vargas R. How health navigators legitimize the Affordable Care Act to the uninsured poor. Soc Sci Med 2016 Sep;165:263-270. [doi: 10.1016/j.socscimed.2016.01.012] [Medline: 26806315]
95. Haldar M, Engebretsen E, Album D. Legitimating the illegitimate: how doctors manage their knowledge of the prestige of diseases. Health (London) 2016 Nov;20(6):559-577. [doi: 10.1177/1363459315596798] [Medline: 26245482]
96. Kannan-Narasimhan R. Organizational ingenuity in nascent innovations: gaining resources and legitimacy through unconventional actions. Organ Stud 2014 Mar 05;35(4):483-509. [doi: 10.1177/0170840613517596]
97. Carr ES, Obertino-Norwood H. Legitimizing evidence: The trans-institutional life of evidence-based practice. Soc Sci Med 2022 Oct;310:115130. [doi: 10.1016/j.socscimed.2022.115130] [Medline: 36070634]
98. Albert M, Paradis E, Kuper A. Interdisciplinary promises versus practices in medicine: the decoupled experiences of social sciences and humanities scholars. Soc Sci Med 2015 Feb;126:17-25. [doi: 10.1016/j.socscimed.2014.12.004] [Medline: 25500163]
99. Alexiou K, Wiggins J. Measuring individual legitimacy perceptions: scale development and validation. Strateg Organ 2018 May 09;17(4):470-496. [doi: 10.1177/1476127018772862]
100. Andersen RS, Tørring ML, Vedsted P. Global health care-seeking discourses facing local clinical realities: exploring the case of cancer. Med Anthropol Q 2015 Jun 30;29(2):237-255. [doi: 10.1111/maq.12148] [Medline: 25355457]
101. Bartlett A, Lewis J, Reyes-Galindo L, Stephens N. The locus of legitimate interpretation in Big Data sciences: lessons for computational social science from -omic biology and high-energy physics. Big Data Soc 2018 May 08;5(1). [doi: 10.1177/2053951718768831]
102. Bensimon CM, Smith MJ, Pisartchik D, Sahni S, Upshur RE. The duty to care in an influenza pandemic: a qualitative study of Canadian public perspectives. Soc Sci Med 2012 Dec;75(12):2425-2430 [FREE Full text] [doi: 10.1016/j.socscimed.2012.09.021] [Medline: 23089615]
103. Carnabuci G, Operti E, Kovács B. The categorical imperative and structural reproduction: dynamics of technological entry in the semiconductor industry. Organ Sci 2015 Dec;26(6):1734-1751. [doi: 10.1287/orsc.2015.1011]
104. Dodworth K, Stewart E. Legitimating complementary therapies in the NHS: campaigning, care and epistemic labour. Health (London) 2022 Mar 07;26(2):244-262 [FREE Full text] [doi: 10.1177/1363459320931916] [Medline: 32508138]
105. Goodrick E, Reay T. Florence Nightingale endures: legitimizing a new professional role identity. J Manag Stud 2009 Nov 11;47(1):55-84. [doi: 10.1111/j.1467-6486.2009.00860.x]
106. Goodwin D, Pope C, Mort M, Smith A. Access, boundaries and their effects: legitimate participation in anaesthesia. Sociol Health Illn 2005 Sep 11;27(6):855-871 [FREE Full text] [doi: 10.1111/j.1467-9566.2005.00477.x] [Medline: 16283902]
107. Hillman A. 'Why must I wait?' The performance of legitimacy in a hospital emergency department. Sociol Health Illn 2014 May 20;36(4):485-499 [FREE Full text] [doi: 10.1111/1467-9566.12072] [Medline: 24053721]
108. Hirschkorn KA. Exclusive versus everyday forms of professional knowledge: legitimacy claims in conventional and alternative medicine. Sociol Health Illn 2006 Jul 03;28(5):533-557 [FREE Full text] [doi: 10.1111/j.1467-9566.2006.00506.x] [Medline: 16910946]
109. Kim PH, Croidieu G, Lippmann S. Responding from that vantage point: field position and discursive strategies of legitimation in the U.S. wireless telegraphy field. Organ Stud 2016 Jul 09;37(10):1417-1450. [doi: 10.1177/0170840616634132]
110. Kushida S, Kawashima M, Abe T. Why this clinic now? A context-sensitive aspect of accounting for visits. Soc Sci Med 2020 Nov;265:113278. [doi: 10.1016/j.socscimed.2020.113278] [Medline: 32947182]
111. Lê G. Trading legitimacy: everyday corruption and its consequences for medical regulation in southern Vietnam. Med Anthropol Q 2013 Sep 08;27(3):453-470. [doi: 10.1111/maq.12052] [Medline: 24248995]
112. Lehoux P, Daudelin G, Hivon M, Miller FA, Denis JL. How do values shape technology design? An exploration of what makes the pursuit of health and wealth legitimate in academic spin-offs. Sociol Health Illn 2014 Jun 21;36(5):738-755. [doi: 10.1111/1467-9566.12097] [Medline: 24444436]

XSL•FO
RenderX

113. Lewis M. De-legitimising complementary medicine: framings of the Friends of Science in Medicine-CAM debate in Australian media reports. Sociol Health Illn 2019 Jun 21;41(5):831-851. [doi: 10.1111/1467-9566.12865] [Medline: 30793332]
114. Lischka JA. Explicit terror prevention versus vague civil liberty: how the UK broadcasting news (de)legitimatise online mass surveillance since Edward Snowden's revelations. Inf Commun Soc 2016 Jul 18;20(5):665-682. [doi: 10.1080/1369118x.2016.1211721]
115. Lockett A, Currie G, Waring J, Finn R, Martin G. The role of institutional entrepreneurs in reforming healthcare. Soc Sci Med 2012 Feb;74(3):356-363. [doi: 10.1016/j.socscimed.2011.02.031] [Medline: 21440968]
116. Manigart S. The founding rate of venture capital firms in three European countries (1970–1990). J Bus Ventur 1994 Nov;9(6):525-541. [doi: 10.1016/0883-9026(94)90020-5]
117. Markard J, Wirth S, Truffer B. Institutional dynamics and technology legitimacy – a framework and a case study on biogas technology. Res Policy 2016 Feb;45(1):330-344. [doi: 10.1016/j.respol.2015.10.009]
118. McKendrick DG, Carroll GR. On the genesis of organizational forms: evidence from the market for disk arrays. Organ Sci 2001 Dec;12(6):661-682. [doi: 10.1287/orsc.12.6.661.10083]
119. Ozcan YA, Shukla RK, Tyler LH. Organizational performance in the community mental health care system: the need fulfillment perspective. Organ Sci 1997 Apr;8(2):176-191. [doi: 10.1287/orsc.8.2.176]
120. Petkova AP, Rindova VP, Gupta AK. No news is bad news: sensegiving activities, media attention, and venture capital funding of new technology organizations. Organ Sci 2013 Jun;24(3):865-888. [doi: 10.1287/orsc.1120.0759]
121. Ridge D, Kokanovic R, Broom A, Kirkpatrick S, Anderson C, Tanner C. "My dirty little habit": patient constructions of antidepressant use and the 'crisis' of legitimacy. Soc Sci Med 2015 Dec;146:53-61. [doi: 10.1016/j.socscimed.2015.10.012] [Medline: 26498732]
122. Stevenson FA, Britten N, Barry CA, Bardley CP, Barber N. Perceptions of legitimacy: the influence on medicine taking and prescribing. Health 2002;6(1):85-104. [doi: 10.1177/1363459302006001446]
123. Swartz A. Legacy, legitimacy, and possibility: an exploration of community health worker experience across the generations in Khayelitsha, South Africa. Med Anthropol Q 2013 Jun 26;27(2):139-154. [doi: 10.1111/maq.12020] [Medline: 23804283]
124. Uibu M. The emergence of new medical pluralism: the case study of Estonian medical doctor and spiritual teacher Luule Viilma. Anthropol Med 2021 Dec 28;28(4):445-460. [doi: 10.1080/13648470.2020.1785843] [Medline: 32854514]
125. van Dijk S, Berends H, Jelinek M, Romme AG, Weggeman M. Micro-institutional affordances and strategies of radical innovation. Organ Stud 2011 Nov 03;32(11):1485-1513. [doi: 10.1177/0170840611421253]
126. Wienroth M. Value beyond scientific validity: let's RULE (Reliability, Utility, LEgitimacy). J Responsible Innov 2020 Oct 30;7(sup1):92-103. [doi: 10.1080/23299460.2020.1835152]
127. Weber M. Economy and Society: A New Translation. Cambridge, MA: Harvard University Press; 2019.
128. Bitektine A, Haack P. The "macro" and the "micro" of legitimacy: toward a multilevel theory of the legitimacy process. Acad Manag Rev 2015 Jan;40(1):49-75. [doi: 10.5465/amr.2013.0318]
129. Haack P, Sieweke J. The legitimacy of inequality: integrating the perspectives of system justification and social judgment. J Manag Stud 2017 Nov 29;55(3):486-516. [doi: 10.1111/joms.12323]
130. Johnson C, Dowd TJ, Ridgeway CL. Legitimacy as a social process. Annu Rev Sociol 2006 Aug 01;32(1):53-78. [doi: 10.1146/annurev.soc.32.061604.123101]
131. Haraway D. When species meet: staying with the trouble. Environ Plan D 2010 Jan 01;28(1):53-55. [doi: 10.1068/d2706wsh]
132. Bourdieu P, Nice R. Outline of a Theory of Practice. Cambridge, UK: Cambridge University Press; 1977.
133. Star SL, Ruhleder K. Steps toward an ecology of infrastructure: design and access for large information spaces. Inf Syst Res 1996 Mar;7(1):111-134. [doi: 10.1287/isre.7.1.111]
134. Star SL. The ethnography of infrastructure. Am Behav Sci 1999 Nov 01;43(3):377-391. [doi: 10.1177/00027649921955326]
135. Monat J, Gannon T. What is systems thinking? A review of selected literature plus recommendations. Am J Syst Sci 2015 Jul;4(1):11-26. [doi: 10.5923/j.ajss.20150401.02]
136. Fox G, Connolly R. Mobile health technology adoption across generations: narrowing the digital divide. Inf Syst J 2018 Jan 29;28(6):995-1019. [doi: 10.1111/isj.12179]
137. Reddy S, Allan S, Coghlan S, Cooper P. A governance model for the application of AI in health care. J Am Med Inform Assoc 2020 Mar 01;27(3):491-497 [FREE Full text] [doi: 10.1093/jamia/ocz192] [Medline: 31682262]
138. Nadal C, Sas C, Doherty G. Technology acceptance in mobile health: scoping review of definitions, models, and measurement. J Med Internet Res 2020 Jul 06;22(7):e17256 [FREE Full text] [doi: 10.2196/17256] [Medline: 32628122]
139. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 2011 Mar 19;38(2):65-76 [FREE Full text] [doi: 10.1007/s10488-010-0319-7] [Medline: 20957426]
140. Davis FD. A technology acceptance model for empirically testing new end-user information systems : theory and results. Massachusetts Institute of Technology. 1986. URL: https://dspace.mit.edu/handle/1721.1/15192 [accessed 2023-10-02]
141. Manski CF. The use of intentions data to predict behavior: a best-case analysis. J Am Stat Assoc 1990 Dec;85(412):934-940. [doi: 10.2307/2289590]

XSL•FO
RenderX

142. Conner M, Norman P. Understanding the intention-behavior gap: the role of intention strength. Front Psychol 2022 Aug 4;13:923464 [FREE Full text] [doi: 10.3389/fpsyg.2022.923464] [Medline: 35992469]
143. Peiris D, Praveen D, Johnson C, Mogulluru K. Use of mHealth systems and tools for non-communicable diseases in low- and middle-income countries: a systematic review. J Cardiovasc Transl Res 2014 Nov 11;7(8):677-691. [doi: 10.1007/s12265-014-9581-5] [Medline: 25209729]
144. Wehrens R, Bal R. Health programs struggling with complexity: a case study of the Dutch 'PreCare' project. Soc Sci Med 2012 Jul;75(2):274-282. [doi: 10.1016/j.socscimed.2012.02.052] [Medline: 22571889]
145. Verran H. Number as an inventive frontier in knowing and working Australia's water resources. Anthropol Theory 2010 Jun 07;10(1-2):171-178. [doi: 10.1177/1463499610365383]
146. Black J, Murray AD. Regulating AI and machine learning: setting the regulatory agenda. Eur J Law Technol 2019;10(3).
147. Pace J, Pearson SA, Lipworth W. Improving the legitimacy of medicines funding decisions: a critical literature review. Ther Innov Regul Sci 2015 May 30;49(3):364-368. [doi: 10.1177/2168479015579519] [Medline: 30222405]
148. Brown B, Cheraghi-Sohi S, Jaki T, Su TL, Buchan I, Sperrin M. Understanding clinical prediction models as 'innovations': a mixed methods study in UK family practice. BMC Med Inform Decis Mak 2016 Aug 09;16:106 [FREE Full text] [doi: 10.1186/s12911-016-0343-y] [Medline: 27506547]
149. Benniche S. Using blockchain technology to recycle cancer drugs. Lancet Oncol 2019 Jun;20(6):e300. [doi: 10.1016/s1470-2045(19)30291-8]
150. Mandeville KL, Barker R, Packham A, Sowerby C, Yarrow K, Patrick H. Financial interests of patient organisations contributing to technology assessment at England's National Institute for Health and Care Excellence: policy review. BMJ 2019 Jan 16;364:k5300 [FREE Full text] [doi: 10.1136/bmj.k5300] [Medline: 30651227]
151. Pocock SJ, Stone GW. The primary outcome is positive - is that good enough? N Engl J Med 2016 Sep 08;375(10):971-979 [FREE Full text] [doi: 10.1056/NEJMra1601511] [Medline: 27602669]
152. Myers SP, Xue CC, Cohen MM, Phelps KL, Lewith GT. The legitimacy of academic complementary medicine. Med J Aust 2012 Jul 16;197(2):69-70. [doi: 10.5694/mja12.10491] [Medline: 22794038]
153. Penedo FJ, Oswald LB, Kronenfeld JP, Garcia SF, Cella D, Yanez B. The increasing value of eHealth in the delivery of patient-centred cancer care. Lancet Oncol 2020 May;21(5):e240-e251. [doi: 10.1016/s1470-2045(20)30021-8]

Abbreviations

CIS: critical interpretive synthesis
EBM: evidence-based medicine
IVF: in vitro fertilization
MAS: medical anthropology and sociology
OMS: organization and management studies
STS: science and technology studies
TAM: Technology Acceptance Model
WoS: Web of Science

Edited by A Kushniruk; submitted 19.05.23; peer-reviewed by M Barbosa, A Geampana, G Berntsen; comments to author 11.12.23; revised version received 24.01.24; accepted 01.12.24; published 05.03.25.

Please cite as:
Howe S, Uyl-de Groot C, Wehrens R
Legitimacy as Social Infrastructure: A Critical Interpretive Synthesis of the Literature on Legitimacy in Health and Technology
JMIR Hum Factors 2025;12:e48955
URL: https://humanfactors.jmir.org/2025/1/e48955
doi:10.2196/48955
PMID:40053717

©Sydney Howe, Carin Uyl-de Groot, Rik Wehrens. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 05.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Benefits and Barriers to mHealth in Hypertension Care: Qualitative Study With German Health Care Professionals

Susann May[1,2], MPH; Felix Muehlensiepen[1,2], Dr rer med; Eileen Wengemuth[1,2], Dr rer nat; Frances Seifert[1,2], MSc; Martin Heinze[3], Prof Dr Med; Dunja Bruch[2], MSc; Sebastian Spethmann[4,5,6], Prof Dr Med

[1]Brandenburg Medical School Theodor Fontane, Center for Health Services Research, Faculty of Health Sciences, Rüdersdorf, Germany

[2]Brandenburg Medical School Theodor Fontane, Department of Cardiovascular Surgery, Heart Center Brandenburg, Faculty of Health Sciences, Bernau, Germany

[3]Brandenburg Medical School, Immanuel Hospital Rüdersdorf, University Clinic for Psychiatry and Psychotherapy, Rüdersdorf, Germany

[4]Deutsches Herzzentrum der Charité, Department of Cardiology, Angiology and Intensive Care Medicine, Berlin, Germany

[5]corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Charité – Universitätsmedizin Berlin, Berlin, Germany

[6]German Center for Cardiovascular Research (DZHK), Partner Site Berlin, Berlin, Germany

Corresponding Author:
Susann May, MPH
Brandenburg Medical School Theodor Fontane
Center for Health Services Research, Faculty of Health Sciences
Seebad 82/83
Rüdersdorf, 15562
Germany
Phone: 49 3391 39 145 91
Email: susann.may@mhb-fontane.de

Abstract

Background: Digital health technologies, particularly mobile health (mHealth) apps and wearable devices, have emerged as crucial assets in the battle against hypertension. By enabling lifestyle modifications, facilitating home blood pressure monitoring, and promoting treatment adherence, these technologies have significantly enhanced hypertension treatment.

Objective: This study aims to explore the perspectives of health care professionals (HCPs) regarding the perceived benefits and barriers associated with the integration of mHealth apps into routine hypertension care. Additionally, strategies for overcoming these barriers will be identified.

Methods: Through qualitative analysis via semistructured interviews, general practitioners (n=10), cardiologists (n=14), and nurses (n=3) were purposefully selected between October 2022 and March 2023. Verbatim transcripts were analyzed using qualitative content analysis.

Results: The results unveiled 3 overarching themes highlighting the benefits of mHealth apps in hypertension care from the perspective of HCPs. First, these technologies possess the potential to enhance patient safety by facilitating continuous monitoring and early detection of abnormalities. Second, they can empower patients, fostering autonomy in managing their health conditions, thereby promoting active participation in their care. Lastly, mHealth apps may provide valuable support to medical care by offering real-time data that aids in decision-making and treatment adjustments. Despite these benefits, the study identified several barriers hindering the seamless integration of mHealth apps into hypertension care. Challenges predominantly revolved around data management, communication contexts, daily routines, and system handling. HCPs underscored the necessity for structural and procedural modifications in their daily practices to effectively address these challenges.

Conclusions: In conclusion, the effective usage of digital tools such as mHealth apps necessitates overcoming various obstacles. This entails meeting the information needs of both HCPs and patients, tackling interoperability issues to ensure seamless data exchange between different systems, clarifying uncertainties surrounding reimbursement policies, and establishing the specific clinical benefits of these technologies. Active engagement of users throughout the design and implementation phases is crucial for ensuring the usability and acceptance of mHealth apps. Moreover, enhancing knowledge accessibility through the provision of easily understandable information about mHealth apps is essential for eliminating barriers and fostering their widespread adoption in hypertension care.

Trial Registration: German Clinical Trials Register DRKS00029761; https://drks.de/search/de/trial/DRKS00029761

XSL•FO
RenderX

International Registered Report Identifier (IRRID): RR2-10.3389/fcvm.2022.1089968

(JMIR Hum Factors 2025;12:e52544) doi:10.2196/52544

KEYWORDS

hypertension; mHealth apps; digital health; physicians; nurses; HCP; qualitative interviews; health care professional; cardiologists; mHealth; Germany; general practitioners; blood pressure monitoring; qualitative study; qualitative content analysis

Introduction

Hypertension is a significant health problem affecting millions of people worldwide [1]. In Germany, too, hypertension is one of the most common diseases affecting a large proportion of the population. Prevention and effective treatment of hypertension are therefore crucial to prevent associated serious complications such as ischemic heart disease, strokes, and renal disease, and to improve the quality of life of those affected [2-4].

In recent years, the development of digital health technologies has experienced a considerable upswing, opening up promising perspectives for the prevention and management of hypertension [5]. Mobile health (mHealth) apps, wearables, and other digital tools enable patients to monitor their blood pressure levels, track lifestyle changes, and set health goals. These technological advances also offer health care professionals (HCPs) new opportunities to improve patient care and promote healthier lifestyles [6]. mHealth apps in particular have been shown to improve the treatment of hypertension [7].

In Germany, the following professional groups are primarily involved in the treatment of people with hypertension. General practitioners play a central role in the diagnosis, treatment, and long-term care of patients with hypertension. They are often the first point of contact for patients with hypertension and coordinate further care. Cardiologists specialize in the diagnosis and treatment of heart disease, including hypertension. They can treat complex cases of hypertension and offer specialized diagnosis and treatment options. Internists can also play a role in the care of patients with hypertension, especially if there are comorbidities or complex medical problems. Nursing professionals (nurses, nursing assistants) play a crucial role in the care of patients with hypertension. They help implement blood pressure control measures, educate patients, and help monitor symptoms and side effects [8]. In October 2020, physicians in Germany were allowed to prescribe digital therapeutics (DiGA [Digitale Gesundheitsanwendungen]) in standard care for the first time [9]. DiGAs can be prescribed by physicians for a 3-month use period and is fully reimbursed by the statutory insurance companies. To obtain approval, mHealth providers must demonstrate the safety, functionality, quality, data security, and a fundamental benefit in a clinical study to obtain DiGA status. Patients receive a written prescription for a DiGA based on respective indications (eg, depression) and have to send it to their insurance company.. At the time this study was conducted, there was no DiGA that is specific for the prevention of hypertension. In perspective, however, it is expected that mHealth apps for the prevention of hypertension will obtain DiGA status and will be made available to patients. In their recent study, Dahlhausen et al [10] pointed out the significant role of HCPs, specifically physicians, in promoting mHealth use among patients. In order to understand the implementation of mHealth into hypertension care, it is necessary to understand the perspectives of physicians, nurses, and medical assistants and their perception of benefits and challenges, as well as their information needs, in order to best integrate mHealth into hypertension care.

Nevertheless, a certain frustration is noticeable among physicians in Germany. Not least because there have been repeated technical problems in the past with the connection to the German telematics infrastructure (TI). The TI comprises a range of technical components that are relevant for operation in medical practices. These include the practice management system (Praxis Verwaltungssystem), which supports the organization and documentation of practice tasks, and the connector, a piece of hardware that enables access to the TI and communicates with the eHealth card terminal and the Praxis Verwaltungssystem via a secure network. The eHealth card terminal is used to use the electronic health card and to check practice ID cards. Other components include mobile card terminals, practice or institution ID cards, a virtual private network access service, and the electronic health professional card, which is required for various applications such as the electronic doctor's letter and prescriptions [11]. However, half of the physicians have to struggle with technical problems at least once a week, compared to 36% in 2020. Overall, frustration with the digitization process has increased [12].

While there is a wealth of research focusing on the patient perspective regarding digital health tools [13-15], there remains a lack of comprehensive evidence on how physicians perceive and integrate these technologies into their clinical practice. Despite the promising potential of digital health tools, the potential obstacles and barriers to widespread adoption of these technologies in the clinical setting have not been thoroughly explored. A deeper understanding of the structural and individual factors that influence the implementation and adoption of digital prevention approaches is imperative. Integrating digital prevention approaches into the health care system requires a clear delineation of the role and positioning of these technologies within the broader health care landscape. A thorough investigation of how digital health tools can be effectively integrated into the existing health care infrastructure and how they are perceived by the various stakeholders in the health care sector is crucial.

Addressing these research gaps can contribute significantly to a more nuanced understanding of how digital health and specifically mHealth apps can be optimally used for hypertension prevention. It can also identify potential barriers and pave the way for strategies to improve clinical adoption and seamless integration of these technologies. In this study, we explore the following research questions: Which benefits

XSL•FO
RenderX

do HCPs see in mHealth apps for prevention? What are possible structural and individual barriers to the use of mHealth apps? How can the challenges be met?

Methods

Study Design

This paper reports on our exploration of the HCPs' perspectives regarding mHealth apps in hypertension prevention. We conducted in-depth interviews with general practitioners, cardiologists, and nurses. This study is part of the DiPaH project [16], which examines structural and individual factors in different stakeholders that influence the use of digital preventive measures in patients with arterial hypertension in Germany.

Ethical Considerations

This study was approved by the Ethics Committee of the Brandenburg Medical School Theodor Fontane (E-02-20220620). All experimental protocols were approved by a named institutional or licensing committee. All methods were carried out in accordance with relevant guidelines and regulations. Participants were informed verbally and in writing about the purpose, the procedure, the significance of the study, and the benefits and risks that may be associated with it and had the opportunity to ask questions. They were also informed that they had the right to withdraw their consent to participate in the study at any time, either verbally or in writing, without giving reasons. They were also informed that personal data would be recorded and stored, whereby the data would be pseudonymized, but that no data would be published that could be used to identify the person. For security reasons, the data received from the participants were always stored in a password-protected folder on a secure desktop computer. Written consent was obtained after participants were given the opportunity to ask questions. Patients were not involved in designing this study. Participants were offered €75.00 (approximately US $77.89) as an incentive for their participation in the study.

Recruitment

Participants were selected using purposive sampling [17]. We included HCPs who were currently actively involved in the treatment of people with hypertension, such as general practitioners, cardiologists, and nurses. The participants were selected according to defined categories, aiming to generate a heterogeneous sample with the broadest variety of different perspectives [18]. The following categories were included: region of practice (rural vs urban), main area of practice (outpatient vs inpatient), and gender (male vs female). A further inclusion criterion was interest and willingness in participating in an interview. Participants were recruited both by systematic invitation and by snowball sampling. The following sources were used for recruitment (Textbox 1).

Textbox 1. Sources used for recruitment.

- Social media accounts of the university (Brandenburg Medical School Theodor Fontane), such as Facebook, Instagram, and Twitter
- Association of Statutory Health Insurance Physicians Brandenburg (KVBB): newsletter
- General practitioners' association Brandenburg: information study participation at events
- Personal contacts among colleagues (snowball sampling)
- Interested persons contacted the study team by telephone or e-mail, and it was checked in advance whether they could be included before an interview was conducted

Data Collection

A preliminary semi-structured interview guide was drafted by a multiprofessional team (SM, FM, DB, and SSp). The semistructured interview guide consisted of open-ended questions that explored how the participants are attuned to digital applications and mHealth apps, what benefits and barriers they see in digital technologies for prevention, what information needs they have, and how these can be met. The focus was on hypertension apps with the following functions: educating about hypertension, monitoring blood pressure, and promoting adherence, although no specific apps were offered or recommended to participants. Sample interview questions included: Which mHealth apps in the context of hypertension treatment are used in your practice, or what experiences have you had with it? What benefits do you see in the use of mHealth apps? What barriers do you see in the use of mHealth apps? How well informed are you about mHealth apps, and where are knowledge deficits? In addition, sociodemographic data were collected, including profession, gender, age, number of inhabitants of the place of practice, duration of professional activity, and setting. To ensure clarity and relevance of the questions, we did a pilot test of the interview guide with 5 eligible participants recruited from clinics and outpatient physicians in the study's catchment area (refer to Multimedia Appendix 1).

All interviews were conducted via telephone during October 2022 to March 2023. The interviews were recorded and transcribed verbatim in accordance with data protection guidelines. Data collection continued until no substantially new findings emerged and saturation of content was reached. Saturation of content is defined as code saturation, when no additional issues are identified, and meaning saturation, when no further dimensions, nuances, or insights of issues can be found [19]. We chose to conduct 27 interviews, as this number has proven to be adequate in previous studies to reach theoretical saturation in the exploration of digital approaches [20,21]. Field notes (a short summary of the interview) were taken following each interview to ensure better comprehensibility of the interview situation. The field notes themselves were not included in the analysis.

XSL•FO
RenderX

Data Analysis

Qualitative analysis of the interviews was performed iteratively by the study team (SM and FM) based on Kuckartz's structured qualitative content analysis [22] using MAXQDA Analytics Pro 2022, Release (version 22.1.0; Verbi GmbH). After transcription of the audio material, the analysis began with a screening of the interview texts, whereupon the interviews were coded. Relevant text passages from the interview material were coded according to a deductive-inductive procedure. Main categories and subcategories were formed inductively from the codes. Until a common understanding of all the emerging categories was achieved, consensus discussions were held continuously in the research group. Other categories were developed deductively based on the research questions and merged into the coding tree. By then, the process of gathering data had already concluded. Two researchers (SM and FM) separately applied the established category system to analyze the complete data set, ensuring that the process could be traced and replicated. The data was analyzed in German. To present the findings, significant excerpts from the transcriptions were chosen as representative quotes. These quotes were translated into English by native speakers and incorporated into the manuscript. The manuscript has been compiled in accordance with the COREQ (Consolidated Criteria for Reporting Qualitative Research; refer to Multimedia Appendix 2) [23].

Results

In total, 27 interviews were conducted and analyzed until theoretical saturation was reached. The mean duration of the interviews was 42 (range 33-56) minutes. The mean age of the participants (N=27) was 50 (range 35-74) years. Most participants were female (14 female/13 male). A total of 14 cardiologists, 3 nurses, and 10 general practitioners participated. In total, 9 persons worked in the inpatient setting and 17 persons worked in the outpatient setting. One person worked in both settings. Four of the HCPs had already had experience with digital applications in daily practice or were also working with them. Detailed characteristics of study participants are shown in Table 1.

Table 1. Detailed characteristics of study participants.

Characteristics	Participants (N=27), n (%)
Professional group	
General practitioner	10 (37)
Cardiologist	14 (52)
Nurse	3 (11)
Duration of professional activity	
>10	9 (33)
10-20	8 (30)
21-30	7 (26)
31-40	3 (11)
Gender	
Female	14 (52)
Male	13 (48)
Age (years)	
>40	2 (7)
40-50	11 (41)
51-60	11 (41)
<60	3 (11)
Number of inhabitants	
>10,000	2 (7)
10-100,000	11 (41)
100,001-1 million	4 (15)
<1 million	10 (37)
Setting[a]	
Inpatient	10 (37)
Outpatient	18 (67)

[a]One participant works in both the inpatient and outpatient settings.

XSL•FO
RenderX

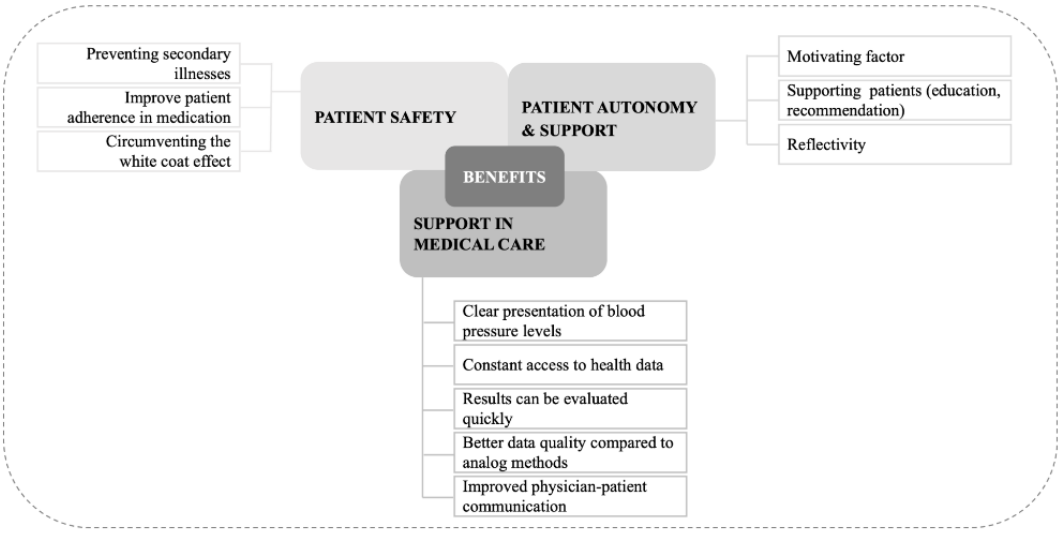
Themes

The analysis explored 7 key themes, categorized into benefits and barriers. Themes identified as benefits included (1) patient safety, (2) patient autonomy and support, and (3) support in medical care. Themes identified as barriers were organized by the following context: (4) handling of data, (5) in the context of communication, (6) in daily routines, and (7) in dealing with the systems.

Benefits

Although few of the HCPs use or have used mHealth apps in their daily practice, they were able to mention benefits they expected to gain from using mHealth apps. In the context of benefits, 3 overarching themes could be identified: potential to increase patient safety, patient autonomy, and support in medical care (Figure 1 and Multimedia Appendix 3).

Figure 1. Coding tree of the benefits regarding mHealth apps in hypertension prevention. mHealth: mobile health.



A different number of subcategories could be assigned to the 3 main categories. The following 3 subcategories were assigned to the “Patient safety” category: preventing secondary illnesses, improving patient adherence in medication, and circumventing the white coat effect. The following 3 subcategories were assigned to the category “Patient autonomy and support”: motivating factor, supporting patients, and reflectivity. The following 5 subcategories were categorized under the category “Support in medical care”: clear presentation of blood pressure levels, constant access to health data. Results can be evaluated quickly, better data quality compared to analog methods, and improved physician-patient communication.

Patient Safety

Respondents see the potential of mHealth apps primarily in the prevention of secondary diseases. In addition, participants see the potential of mHealth apps to improve patient adherence to treatment in terms of medication adherence, which in turn has an impact on patient safety.

The respondents attribute a better representation to the blood pressure levels measured at home and subsequently documented. The interviewees hope that through the use of mHealth apps, patients will consistently document their blood pressure and thus always have an overview of their blood pressure levels. This is because, in contrast to the blood pressure levels measured in the practice, they seem to correspond more closely to the reality of life. The “white coat effect” is thus circumvented.

Patient Autonomy and Support

An mHealth app can offer a variety of features that serve as a motivating factor for patients from the perspective of health care providers. Gamified elements, progress tracking, and reward systems can provide incentives for developing and maintaining health-promoting behaviors, strengthening patients' self-motivation, and fostering engagement in their health.

From the HCP’s perspective, an mHealth app can be a valuable addition to medical care by providing support to patients in several ways. First, the mHealth app can provide personalized advice based on the patient's individual needs and health goals. This can include dietary recommendations, lifestyle changes, or stress management tips. Second, mHealth apps can act as reminders, reminding patients to take their medication, which could improve adherence to therapy.

An mHealth app can also include reflectivity and help patients comprehend information about their blood pressure to their behaviors, which can support them reduce or avoid unfavorable behaviors.

Support in Medical Care

By using mHealth apps, the HCPs assume that blood pressure levels can be presented in a clear and understandable way. mHealth apps allow patients to access their levels on their smartphones or other digital devices. Patients can track the development of their blood pressure over time and identify possible trends or abnormalities so that they can seek medical help timely if necessary.

Patients have constant access to their current health data. This enables real-time tracking of their health status, even if they are not undergoing immediate medical treatment.

XSL•FO
RenderX

mHealth apps provide quick results that can be evaluated immediately. Physicians can thus make timely decisions about the treatment and care of their patients.

HCPs attribute better data quality to mHealth apps compared to analog methods.

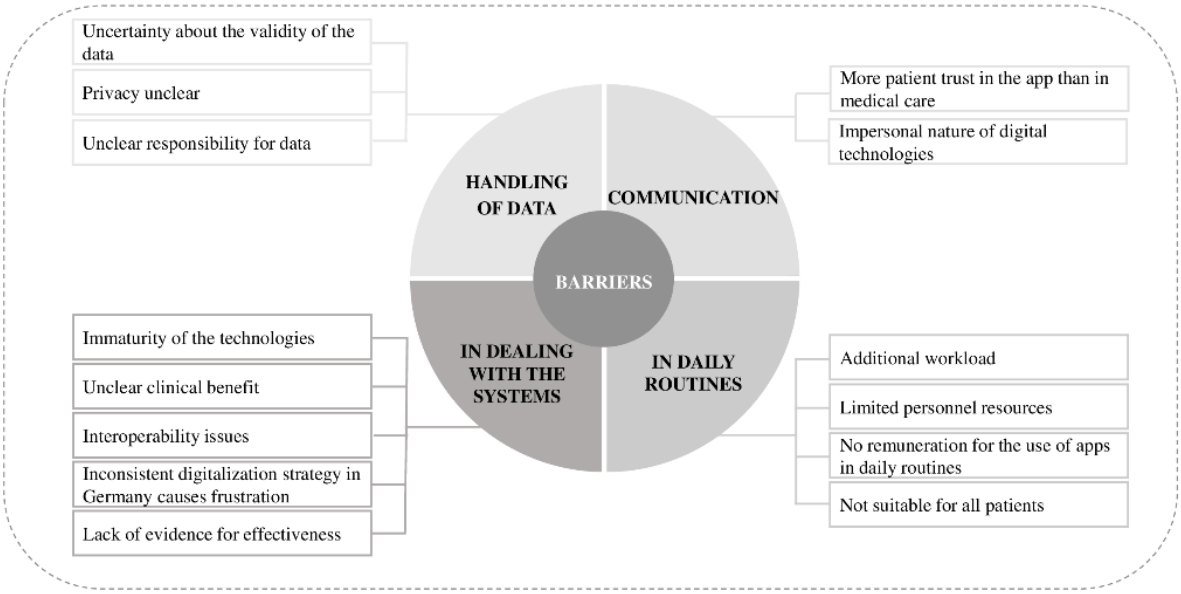
Using mHealth apps, patients can send the data they collect in their daily environment directly to the medical facility before visiting the physician. This allows the physician to check the blood pressure levels in advance and prepare specifically for the conversation with the patient. The time saved helps to make the physician-patient conversation more efficient and targeted so that relevant questions and concerns of the patient can be better addressed.

The benefits initially included constant access to health data, which was particularly mentioned by participants in the outpatient sector. This group also appreciated the motivating factor that the apps could provide to support them in their health routines. Participants in large cities also preferred the app's clear display of their blood pressure readings, which helped them to better understand their health data. The time saved in doctor-patient communication was also mentioned as a benefit by the outpatient participants.

Barriers

We identified barriers in 4 different aspects of the HCPs work life: handling of data, in the context of communication, in daily routines, and in dealing with the systems (Figure 2 and Multimedia Appendix 4). A different number of subcategories could be assigned to the 4 main categories. The following 3 subcategories were assigned to the category “Handling of data”: uncertainty about the validity of the data, privacy unclear, and unclear responsibility for data. The following two subcategories were assigned to the category “in the context of communication”: more patient trust in the app than in medical care and the impersonal nature of digital approaches. The following 4 categories were assigned to the category “in daily routines”: additional workload, limited personnel resources, no remuneration for the use of apps in daily routines, and not suitable for all patients. The following 5 subcategories were assigned to the category “in dealing with the systems”: immaturity of the technologies, unclear clinical benefit, interoperability issues, and inconsistent digitalization strategy in Germany, which causes frustration and lack of evidence for effectiveness.

Figure 2. Coding tree of the barriers regarding mHealth apps in hypertension prevention. mHealth: mobile health.



Handling of Data

One of the main barriers to the use of apps for blood pressure monitoring is uncertainty about the validity of the data collected. Concerns about the accuracy of data documented by patients can undermine HCPs’ confidence in the data and willingness to use such tools.

Another important concern related to the use of mHealth apps in the treatment of hypertension is privacy. Fear of data breaches or unauthorized access to patient information may limit the use of such tools.

Unclear responsibility for data presents another hurdle. Physicians and nurses feel uncertain about who is responsible for managing and interpreting the data they collect. The wealth of data generated by mHealth apps can be overwhelming, and the lack of clear structures and policies for appropriate use of the data can be challenging.

In the Context of Communication

The HCPs expect that a potential barrier to communication when using mHealth apps is the tendency of some patients to trust the mHealth app more than their medical care. This misconception could lead to patients relying solely on the mHealth app for solutions or seeing its advice as better than that of their physician. This could hinder communication with HCPs and lead to a lack of collaboration between patients and physicians.

XSL•FO
RenderX

Another barrier perceived by HCPs is the impersonal nature of mHealth apps. In contrast to direct interaction, the use of mHealth apps could lead to a lower intensity of exchange between patients and HCPs. In addition, mHealth apps may be perceived as less empathetic and less tailored to the individual needs of patients.

In Daily Routines

The HCPs apprehend that introducing mHealth apps in medical practice can initially lead to an additional workload for them. The introduction of new technologies requires training, the adaptation of workflows, and integration into existing IT systems, which can initially lead to time and resource constraints.

In many health care settings, resources are lacking in order to engage HCPs in the active use of mHealth apps in practice. Adoption and integration of new technologies requires time, training, and technical support, which can be challenging in an already strained health care system. Limited capacity impedes readiness to use mHealth apps for hypertension care.

At present, the use of mHealth apps is not sufficiently reimbursed in everyday health care in Germany. HCPs do not receive additional compensation or incentives for the use and deployment of digital tools in the care of hypertensive patients. This lack of compensation may cause some HCPs to be hesitant to use digital tools, as it may entail additional tasks without financial compensation.

Although mHealth apps offer many benefits, they may not be suitable for all patients. Older people or those with limited digital capabilities may have difficulty using digital applications. In addition, some patients, especially those with hypertension, may overmeasure their blood pressure, which can negatively impact their blood pressure levels. According to HCPs, digital tools for monitoring blood pressure may be less appropriate for such patients.

In Dealing With the Systems

A major barrier to the use of mHealth apps in hypertension care is the immaturity of the technologies. HCPs suspect that mHealth apps do not work reliably, have bugs, or do not have all the necessary features to meet the demands of medical practice. As a result, they may be reluctant to recommend digital technologies for fear that the systems will not exist at a later date.

Another barrier to the use of mHealth apps in hypertension care is the unclear actual utility and clinical effectiveness of such tools from the perspective of HCPs.

The diversity of mHealth apps and their interfaces is another barrier for HCPs. The incompatibility or lack of interoperability between systems can hinder smooth data exchange and integration.

The insufficient as well as inconsistent digitalization strategy in Germany is a source of frustration for HCPs. There is often a lack of a clear, long-term vision and of a structured approach to drive the integration of digital technologies in health care. The development and implementation of digital health solutions is often fragmented and inconsistent, leading to confusion, ineffective solutions, and a return to analog methods. These negative experiences are a significant barrier to the adoption of digital tools.

The actual effectiveness and long-term impact of mHealth apps in treating hypertension are poorly researched or demonstrated from the perspective of HCPs. The lack of evidence on the clinical effectiveness of such tools raises doubts and influences their willingness to use them.

Specific Aspects in Various Subgroups

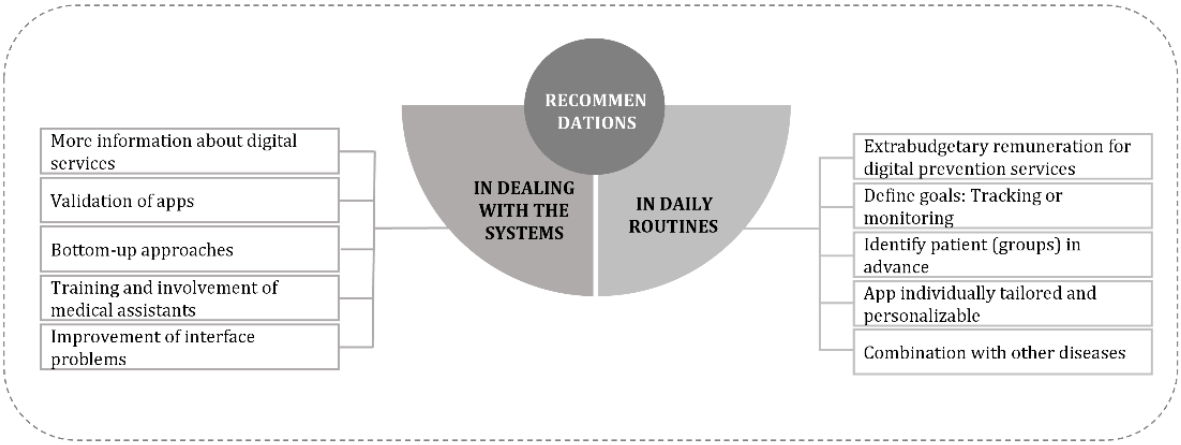
Participants from the outpatient sector in particular pointed out interoperability problems that could hinder the smooth exchange of data. In addition, the unclear clinical benefits of the apps were not recognized by everyone, especially not by the participants from the outpatient sector. It was also noted from the outpatient sector that not all apps are suitable for every patient. Concerns about data protection were also expressed by participants in the outpatient sector, who were worried about the security of their personal health data. Finally, the lack of reimbursement for the use of health apps from the outpatient sector was cited as a further barrier. Women expressed concerns about the validity of the data collected by the apps.

Recommendations for Successful Integration of mHealth Apps Into the Health Care Landscape From HCPs' Perspective

In the context of recommendations, 2 themes could be identified: in daily routines and in dealing with the systems (Figure 3 and Multimedia Appendix 5). A different number of subcategories could be assigned to the 2 main categories. The following 5 categories were assigned to the category "in daily routines": extrabudgetary remuneration for digital prevention services, defining goals: tracking or monitoring, identifying patient (groups) in advance, app individually tailored and personalizable, and combination with other diseases. The following 5 subcategories were assigned to the category "in dealing with the systems": more information about digital services, validation of apps, bottom-up approaches, training and involvement of medical assistants, and improvement of interface problems.

XSL•FO
RenderX

Figure 3. Coding tree of the recommendations for successful integration of mHealth apps into the health care landscape. mHealth: mobile health.



In Daily Routines

One of the fundamental challenges in integrating hypertension mHealth apps into routine medical practice is that the time spent using and counseling patients is not adequately compensated. To address this problem and increase HCPs' motivation to use digital prevention services, extrabudgetary reimbursement for the implementation and support of these apps in HCPs' practices should be considered. According to the participants, such remuneration would ensure that HCPs are appropriately compensated and thus create an incentive to increasingly integrate mHealth apps into their daily work.

It is crucial to define clear goals for the use of mHealth apps in hypertension prevention. HCPs should work with patients to determine which parameters or data should be collected by the mHealth app and which specific health goals are being pursued. The HCPs assume that a distinction can be made between tracking and monitoring approaches. While tracking helps patients to collect data on their blood pressure, lifestyle, and other relevant parameters themselves, monitoring enables the physician to regularly monitor and analyze the collected data. The clear definition of goals facilitates the use of the mHealth apps and ensures that the collected data can be used meaningfully and effectively.

The HCPs mentioned that not all patients benefit equally from mHealth apps. It is important to identify in advance which patients or patient groups could benefit most from a hypertension prevention mHealth app. Some patients may already be well informed and motivated to use digital health technologies on their own, while others may need more support and guidance. By identifying those patients who could best benefit from mHealth app use, clinicians can target their mHealth services more effectively and efficiently.

Another important improvement is that hypertension apps should be customized and personalizable. Each patient is unique, and the mHealth app should therefore be customizable to each individual's specific needs, health goals, and preferences. Personalization options could include selection of relevant health goals, medication reminders, or customization of the app interface to the user's personal preferences. By tailoring the app to each patient's individual needs, the user experience is likely to be enhanced and the likelihood of long-term use increased.

Hypertension often occurs in combination with other conditions, such as diabetes or obesity. To improve the care of multimorbid patients, mHealth apps for hypertension could be combined with other health applications to provide comprehensive and integrated care.

In Dealing With the Systems

It is important that HCPs have access to comprehensive and reliable information about the functionalities, efficacy, safety, and privacy policies of the various hypertension apps. A transparent and evidence-based presentation of this information can help to address concerns about the quality and reliability of the apps and stimulate physicians' interest in using these digital tools.

Careful validation of mHealth apps is essential to ensure that hypertension management apps actually deliver the promised benefits and provide medically robust data. HCPs need reliable information about the scientific evidence and clinical validity of the apps in order to effectively integrate them into their clinical practice.

To improve the adoption and integration of apps for hypertension, it is crucial to promote bottom-up approaches. This means including the user perspective, including physicians and medical staff, in the development and implementation of mHealth apps from the very beginning. Involving users in the development process makes it possible to adapt mHealth apps to the specific needs and requirements of physicians and ensure high usability.

Successful integration of hypertension apps requires not only training of physicians but also the involvement of other HCPs, especially medical assistants. Medical assistants play an important role in helping physicians use mHealth apps and interact with patients. Extensive training of medical assistants on how the mHealth apps work and are useful, as well as how to effectively use the data collected, can help ensure that the integration of the apps is seamless in the practice workflow.

Interoperability and compatibility between different mHealth apps and existing practice IT systems are critical. Often,

XSL•FO
RenderX

physicians and medical assistants face the challenge that the interfaces between the mHealth apps and existing IT systems are not sufficiently optimized, which can lead to technical difficulties and data inconsistencies. Targeted improvement of the interface issues enables physicians to work smoothly with the mHealth apps and seamlessly integrate the captured data into their practice workflows.

Discussion

Principal Results

This qualitative study is the first to explore the factors that enable or hinder the use of mHealth apps for the prevention of hypertension from the perspective of general practitioners, cardiologists, and nurses in Germany. Until now, mHealth apps for the prevention of hypertension are only fragmentarily used in the care of hypertensive patients. From the perspective of the caregivers, the main potentials of mHealth apps lie in increasing patient safety, patient autonomy, and support in medical care. However, the presented results point to major challenges in data handling, patient-HCP communication, integration into standard care, and interoperability. Measures such as extra-budgetary remuneration for digital preventive services, setting targets through tracking or monitoring, identifying patient groups in advance, adapting and personalizing mHealth apps, and combining them with other diseases could, in the view of the HCPs, help to overcome the barriers. In dealing with the systems, the HCPs mention aspects such as more information about digital services, validation of mHealth apps, bottom-up approaches, training and involvement of medical assistants, and improvement of interface problems.

Comparison With Prior Work

Our results are in line with earlier research in mHealth and beyond: physicians and nurses recognize the added values and positive potential of mHealth apps in supporting patients, fostering patient autonomy, and improving medication adherence [24,25]. The attributed benefits are supported by evidence: for instance, Li et al [26] report improvements in self-management behavior and medication adherence in adults using mHealth. Furthermore, self-monitoring of hypertension-related behaviors via smartphone apps combined with tailored advice has a modest but potentially clinically significant effect on blood pressure reduction [27].

Despite the advantages of mHealth, as of yet the barriers hinder their integration into the medical routine. One of the main concerns is the handling of data. Physicians frequently doubt the validity of data collected by patients and have concerns regarding data privacy and security. This can impede trust in the data and willingness to use such mHealth apps.

The mode of communication between physicians and patients can also be influenced by the use of mHealth apps. Some patients may tend to place more trust in the instructions provided by the mHealth app than in the recommendations of their medical caregivers. This could hinder communication with HCPs and impede collaboration between patients and physicians. However, studies have already demonstrated that there are contrary findings to the presumed concerns: findings of a systematic, narrative review show a generally positive influence of mHealth apps on physician-patient communication and relationships, which at times was correlated with better health outcomes [28]. And perhaps the use of apps can make visits more efficient, because blood pressure levels can be transmitted to the practice in advance, and conversations can be focused on essential aspects.

In the interviews, it is stated that mHealth consumes more resources and deteriorates patient communication. In rheumatology, digital health and especially mHealth are considered resource-saving. Digitization creates more time for patient interaction, which enhances the quality of care. If questions have already been answered beforehand, there may be extra time in the conversation with the patient to focus on essential aspects, which can positively impact the perceived quality of care [29]. Furthermore, it is emphasized in rheumatology care that this would allow more time for patient consultations. This is a contradiction. However, rheumatology is more digitized compared to hypertension care because of the greater shortage of personnel and disease burden.

Furthermore, apps can initially mean additional workload for HCPs. The introduction of new technologies requires training, adaptation of workflows, and integration into existing IT systems, which can lead to time and resource constraints. In addition, due to limited resources, many medical facilities may not be adequately prepared to actively integrate digital technologies into their practice. To be able to use digital technologies in their daily routine, physicians also need the appropriate digital skills [30]. In a recent measurement of physicians' professional digital health literacy, they scored 53.1 out of a possible 100 points [31]. For nurses, this figure is only slightly higher at 54.5 points. The most difficult task for physicians and nurses is “helping patients assess the trustworthiness of the digital health information they find,” followed by “helping patients find the digital health information that is relevant to them.” In the future, physicians and nurses will need support especially in this area so that the digital transformation in health care can be driven forward.

A lack of integration of digital tools into the reimbursement structure of the health care system in Germany is also a barrier to the use of such applications. Currently, physicians only receive remuneration for prescribing DiGAs, which does not translate to monitoring apps that have not obtained the DiGA status. There are ongoing discussions about which remuneration model is suitable for mHealth. Labinsky et al [32] reported that only 13% of patients that were prescribed a DiGA completed the whole program. This may also be due to a lack of incentives for physicians to ask if the apps are being used. Considering the low adherence with DiGA, performance-based compensation appears to be appropriate.

Implications

Several recommendations for the integration of mHealth apps into medical practice can be derived from the results. It is important to provide HCPs with low-threshold access to comprehensive and reliable information about available mHealth apps to address their concerns about quality and privacy. Validation of mHealth apps for effectiveness and clinical

XSL•FO
RenderX

validity is critical to convince physicians that these apps provide real value [33]. Bottom-up approaches or participatory study designs that include the user perspective, including physicians and medical assistants in the development process, can help ensure that apps are better tailored to the needs and requirements of medical practice and have a higher usability.

Involving HCPs, especially medical assistants, is also important as they play a key role in helping physicians use mHealth apps and interact with patients. Extensive training of medical staff on how the apps work and their benefits can help ensure that the integration of the apps is seamless in the practice workflow.

Our results underline the importance of staggered use of the mHealth app for effective blood pressure control in patients with hypertension. A high frequency of use at the beginning of treatment helps both patients and health care providers to quickly identify and implement individualized strategies to lower blood pressure. Once the target values have been stably achieved, the frequency of use can be gradually reduced so that the mHealth app can serve as a monitoring tool without unnecessarily increasing the time required by medical staff. This flexible handling could allow the mHealth app to be optimally integrated into everyday clinical practice while minimizing the burden on health care providers in terms of data management.

In addition, it is important to improve the interface issues between digital health apps and existing IT systems. Interoperability and compatibility between different digital apps and medical practice software systems are critical to ensure smooth integration and use of the apps. Targeted improvement of the interfaces can enable physicians to work smoothly with the apps and seamlessly integrate the collected data into their practice workflows.

The use of digital care services could be beneficial if the goal of use were addressed in advance in physician-patient communication. For example, use in the sense of self-tracking would not necessarily require medical intervention. Whereas the use in the sense of monitoring is associated with a control function, and physicians are actively involved in the process. Further studies are also needed to analyze the benefits and effectiveness of digital care services.

Based on the current findings and recommendations made by the HCPs, various stakeholders (eg, policy makers, health insurance companies, physicians' associations, associations of the scientific medical societies, patient organizations, health care researchers, or app developers) in different fields of the health care system should take action to push the integration of mHealth app in daily practice (Textbox 2).

Textbox 2. Recommendations for the integration of mobile health (mHealth) apps into medical practice.

Structural conditions

- Extrabudgetary remuneration for digital prevention services
- Improvement of interface problems
- Training and involvement of medical assistants
- More information about digital offers

App development

- Bottom-up approaches and user experience design
- App individually tailored and personalizable
- Combination with other diseases
- Validation of mHealth apps

Use in patient care

- Assessment of patients' eligibility to use mHealth
- Identify patient (groups) in advance
- Define goals: tracking or monitoring

Strength and Limitations

Our study is the first to investigate the benefits and barriers of using mHealth apps in preventing hypertension in Germany. We drew closely on the reality of care and the perspectives of the physicians and nurses. The qualitative interviews allowed for obtaining an in-depth understanding of the experiences with or without apps in preventing hypertension and expected advantages and disadvantages. However, there are certain limitations to our study. The results might be subject to a selection bias because perhaps people who were more interested in the topic took part in the study. Due to the sampling strategy, we may not have been able to reach everyone and did not record important aspects. A generalization of the results is therefore impossible. In addition, patients were not interviewed at this time.

The authors are aware that this general exploration can only be a first step given the breadth and diversity of the topic "hypertension mHealth apps." Further studies must be conducted in the context of specific health care situations to build upon these findings. These limitations directly translate into opportunities for further research: We are planning to extend

XSL•FO
RenderX

the study to a larger sample in Germany [16] and to validate the qualitative results with a questionnaire survey.

Further research should also focus on how to identify patients who may or may not be eligible for mHealth apps in hypertension care; how HCPs can be effectively informed about mHealth and digital health in general, as well as what reimbursement strategy will reinforce the effective use of mHealth in the care of chronic conditions.

Conclusions

HCPs recognize the benefits of mHealth apps for patient safety or to support disease management. However, at present, best practices for implementing mHealth apps in preventing hypertension do not exist. In order to be able to use digital tools such as mHealth apps efficiently, the barriers, such as information needs, interoperability issues, uncertain remuneration, and vague clinical effectiveness, need to be overcome. Above all, users need to be involved in design and implementation processes. Furthermore, low knowledge needs to be eliminated by offering access to low-threshold information on mHealth apps.

Acknowledgments

The authors would like to thank all interview partners and participants for their valuable contributions and time. We would like to express our special thanks to the Association of Statutory Health Insurance Physicians Brandenburg (KVBB) and the General Practitioners' Association Brandenburg for their support during the recruitment process. This work was supported by the Innovation Committee at the Federal Joint Committee (Innovation Fund, grant number 01VSF18014).

Data Availability

All data generated or analyzed during this study are included in this published article (and its supplementary information files). For further questions regarding the reuse of data, please contact the corresponding author (susann.may@mhb-fontane.de).

Authors' Contributions

SM made substantial contributions to the design of the work, the data acquisition and analysis, and the interpretation of the data; drafted the manuscript and had full access to all the data in the study; and takes responsibility for the integrity of the data and the accuracy of the data analysis. FM made substantial contributions to the design of the work, the data acquisition and analysis, and the interpretation of the data; drafted the manuscript and had full access to all the data in the study; and takes responsibility for the integrity of the data and the accuracy of the data analysis. EW made substantial contributions to the design of the work and the data acquisition and analysis and substantively revised the manuscript. FS made substantial contributions to the data analysis and substantively revised the manuscript. MH made substantial contributions to the data acquisition and analysis and substantively revised the manuscript. DB made substantial contributions to the design of the work, obtaining funding, and data acquisition and analysis, and substantively revised the manuscript. SSp made substantial contributions to the design of the work, obtaining funding, and data acquisition and analysis, and substantively revised the manuscript. All authors have approved the submitted manuscript. In addition, the authors have agreed both to be personally accountable for the author's own contributions and to ensure that questions related to the accuracy or integrity of any part of the work, even ones in which the author was not personally involved, are appropriately investigated, resolved, and the resolution documented in the literature. SM and FM contributed equally to this work (shared first authorship). DB and SSp contributed equally to this work (shared last authorship).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Excerpt of the interview guide.
[DOCX File , 23 KB - humanfactors_v12i1e52544_app1.docx]

Multimedia Appendix 2
COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist.
[PDF File (Adobe PDF File), 56 KB - humanfactors_v12i1e52544_app2.pdf]

Multimedia Appendix 3
Coding tree with anchor quotes: benefits regarding mHealth apps in hypertension prevention. mHealth: mobile health.
[DOCX File , 15 KB - humanfactors_v12i1e52544_app3.docx]

Multimedia Appendix 4

XSL•FO
RenderX

Coding tree with anchor quotes: barriers regarding mHealth apps in hypertension prevention. mHealth: mobile health.
[DOCX File , 16 KB - humanfactors_v12i1e52544_app4.docx]

Multimedia Appendix 5
Coding tree with anchor quotes: recommendations for successful integration of mHealth apps into the health care landscape. mHealth: mobile health.
[DOCX File , 15 KB - humanfactors_v12i1e52544_app5.docx]

References

1. Forouzanfar MH, Liu P, Roth GA, Ng M, Biryukov S, Marczak L, et al. Global burden of hypertension and systolic blood pressure of at least 110 to 115 mm Hg, 1990-2015. J Am Med Assoc 2017;317(2):165-182 [FREE Full text] [doi: 10.1001/jama.2016.19043] [Medline: 28097354]
2. Neuhauser H, Kuhnert R, Born S. 12-Month prevalence of hypertension in Germany. J Health Monit 2017;2(1):51-57 [FREE Full text] [doi: 10.17886/RKI-GBE-2017-016] [Medline: 37151301]
3. Zhou B, Perel P, Mensah GA, Ezzati M. Global epidemiology, health burden and effective interventions for elevated blood pressure and hypertension. Nat Rev Cardiol 2021;18(11):785-802 [FREE Full text] [doi: 10.1038/s41569-021-00559-8] [Medline: 34050340]
4. Olsen MH, Angell SY, Asma S, Boutouyrie P, Burger D, Chirinos JA, et al. A call to action and a lifecourse strategy to address the global burden of raised blood pressure on current and future generations: the Lancet Commission on Hypertension. Lancet 2016;388(10060):2665-2712. [doi: 10.1016/S0140-6736(16)31134-5] [Medline: 27671667]
5. Parati G, Goncalves A, Soergel D, Bruno RM, Caiani EG, Gerdts E, et al. New perspectives for hypertension management: progress in methodological and technological developments. Eur J Prev Cardiol 2023;30(1):48-60. [doi: 10.1093/eurjpc/zwac203] [Medline: 36073370]
6. Kuwabara A, Su S, Krauss J. Utilizing digital health technologies for patient education in lifestyle medicine. Am J Lifestyle Med 2020;14(2):137-142 [FREE Full text] [doi: 10.1177/1559827619892547] [Medline: 32231478]
7. Mikulski BS, Bellei EA, Biduski D, De Marchi ACB. Mobile health applications and medication adherence of patients with hypertension: a systematic review and meta-analysis. Am J Prev Med 2022;62(4):626-634. [doi: 10.1016/j.amepre.2021.11.003] [Medline: 34963562]
8. Nationale VersorgungsLeitlinie Hypertonie (2023). URL: https://register.awmf.org/assets/guidelines/nvl-009l_S3_Hypertonie_2023-06.pdf [accessed 2024-04-04]
9. Bundesministerium für Gesundheit. Bundesgesundheitsminister Jens Spahn im Bundestag zum Digitale-Versorgung-Gesetz (DVG). URL: https://www.bundesgesundheitsministerium.de/presse/reden/dvg-23-lesung.html#c16594 [accessed 2020-12-15]
10. Dahlhausen F, Zinner M, Bieske L, Ehlers JP, Boehme P, Fehring L. There's an app for that, but nobody's using it: insights on improving patient access and adherence to digital therapeutics in Germany. Digit Health 2022;8:20552076221104672 [FREE Full text] [doi: 10.1177/20552076221104672] [Medline: 35811758]
11. Telematikinfrastruktur: alle Infos zur TI und ihren Anwendungen (2024). URL: https://abrechnungsstelle.com/telematikinfrastruktur/ [accessed 2024-04-04]
12. Müller T, Padmanabhan P, Richter L, Silberzahn T. e-Health Monitor 2022. URL: https://www.mckinsey.com/industries/life-sciences/our-insights/german-e-health-offerings-expand-but-adoption-remains-uneven#/ [accessed 2023-05-05]
13. Madanian S, Nakarada-Kordic I, Reay S, Chetty T. Patients' perspectives on digital health tools. PEC Innov 2023;2:100171 [FREE Full text] [doi: 10.1016/j.pecinn.2023.100171] [Medline: 37384154]
14. Morrissey EC, Casey M, Glynn LG, Walsh JC, Molloy GJ. Smartphone apps for improving medication adherence in hypertension: patients' perspectives. Patient Prefer Adherence 2018;12:813-822 [FREE Full text] [doi: 10.2147/PPA.S145647] [Medline: 29785096]
15. McBride CM, Morrissey EC, Molloy GJ. Patients' experiences of using smartphone apps to support self-management and improve medication adherence in hypertension: qualitative study. JMIR Mhealth Uhealth 2020;8(10):e17470 [FREE Full text] [doi: 10.2196/17470] [Medline: 33112251]
16. Bruch D, Muehlensiepen F, May S, Wengemuth E, Johannsen O, Reber KC, et al. Digital preventive measures for arterial hypertension (DiPaH) - a mixed-methods study protocol for health services research. Front Cardiovasc Med 2022;9:1089968 [FREE Full text] [doi: 10.3389/fcvm.2022.1089968] [Medline: 36704456]
17. Etikan I, Musa SA, Alkassim RS. Comparison of convenience sampling and purposive sampling. Am J Theor Appl Stat 2016;5(1):1-4 [FREE Full text]
18. Robinson RS. Purposive sampling. In: Michalos AC, editor. Encyclopedia of Quality of Life and Well-Being Research. Netherlands: Springer Dordrecht; 2014.
19. Hennink MM, Kaiser BN, Marconi VC. Code saturation versus meaning saturation: how many interviews are enough? Qual Health Res 2017;27(4):591-608 [FREE Full text] [doi: 10.1177/1049732316665344] [Medline: 27670770]

XSL•FO
RenderX

20. Stenzel R, Hadaschik K, May S, Grahammer M, Labinsky H, Welcker M, et al. Digitally-supported patient-centered asynchronous outpatient follow-up in rheumatoid arthritis - an explorative qualitative study. BMC Health Serv Res 2022;22(1):1297 [FREE Full text] [doi: 10.1186/s12913-022-08619-6] [Medline: 36307779]
21. Muehlensiepen F, May S, Hadaschik K, Vuillerme N, Heinze M, Grahammer M, et al. Digitally supported shared decision-making and treat-to-target in rheumatology: a qualitative study embedded in a multicenter randomized controlled trial. Rheumatol Int 2023;43(4):695-703 [FREE Full text] [doi: 10.1007/s00296-022-05224-y] [Medline: 36229671]
22. Kuckartz U. Qualitative inhaltsanalyse. 4th ed. Weinheim: Beltz Juventa; 2018.
23. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
24. Wangler J, Jansky M. Gesundheits-Apps in der hausarztbasierten versorgung ? Empirische befunde zur perspektive von allgemeinmedizinern und patienten. In: Pfannstiel MA, Holl F, Swoboda WJ, editors. mHealth-Anwendungen für chronisch Kranke. Wiesbaden: Springer Gabler; 2020.
25. Morrissey EC, Glynn LG, Casey M, Walsh JC, Molloy GJ. New self-management technologies for the treatment of hypertension: general practitioners' perspectives. Fam Pract 2018;35(3):318-322. [doi: 10.1093/fampra/cmx100] [Medline: 29088438]
26. Li R, Liang N, Bu F, Hesketh T. The effectiveness of self-management of hypertension in adults using mobile health: systematic review and meta-analysis. JMIR Mhealth Uhealth 2020;8(3):e17776 [FREE Full text] [doi: 10.2196/17776] [Medline: 32217503]
27. Kassavou A, Wang M, Mirzaei V, Shpendi S, Hasan R. The association between smartphone app-based self-monitoring of hypertension-related behaviors and reductions in high blood pressure: systematic review and meta-analysis. JMIR Mhealth Uhealth 2022;10(7):e34767 [FREE Full text] [doi: 10.2196/34767] [Medline: 35819830]
28. Qudah B, Luetsch K. The influence of mobile health applications on patient - healthcare provider relationships: a systematic, narrative review. Patient Educ Couns 2019;102(6):1080-1089. [doi: 10.1016/j.pec.2019.01.021] [Medline: 30745178]
29. Mühlensiepen F, Kurkowski S, Krusche M, Mucke J, Prill R, Heinze M, et al. Digital health transition in rheumatology: a qualitative study. Int J Environ Res Public Health 2021;18(5):2636 [FREE Full text] [doi: 10.3390/ijerph18052636] [Medline: 33807952]
30. Muehlensiepen F, Petit P, Knitza J, Welcker M, Vuillerme N. Factors associated with telemedicine use among German general practitioners and rheumatologists: secondary analysis of data from a nationwide survey. J Med Internet Res 2022;24(11):e40304 [FREE Full text] [doi: 10.2196/40304] [Medline: 36449333]
31. Schaeffer D, Griese L, Haarmann A. [Professional health literacy of general practitioners - results of the HLS-PROF]. Gesundheitswesen 2024 [FREE Full text] [doi: 10.1055/a-2350-6377] [Medline: 39191374]
32. Labinsky H, Gupta L, Raimondo MG, Schett G, Knitza J. Real-world usage of digital health applications (DiGA) in rheumatology: results from a German patient survey. Rheumatol Int 2023;43(4):713-719 [FREE Full text] [doi: 10.1007/s00296-022-05261-7] [Medline: 36543961]
33. Ventola CL. Mobile devices and apps for health care professionals: uses and benefits. P T 2014;39(5):356-364 [FREE Full text] [Medline: 24883008]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
DiGA: Digitale Gesundheitsanwendungen
HCP: health care professional
mHealth: mobile health
TI: telematics infrastructure

Edited by A Kushniruk; submitted 07.09.23; peer-reviewed by K Lakshminarayan, RJ Katz, YR Park, G Alnooh; comments to author 07.03.24; revised version received 25.04.24; accepted 31.12.24; published 10.03.25.

Please cite as:
May S, Muehlensiepen F, Wengemuth E, Seifert F, Heinze M, Bruch D, Spethmann S
Benefits and Barriers to mHealth in Hypertension Care: Qualitative Study With German Health Care Professionals
JMIR Hum Factors 2025;12:e52544
URL: https://humanfactors.jmir.org/2025/1/e52544
doi:10.2196/52544
PMID:

XSL•FO
RenderX

©Susann May, Felix Muehlensiepen, Eileen Wengemuth, Frances Seifert, Martin Heinze, Dunja Bruch, Sebastian Spethmann. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 10.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Perceived Risks, Mitigation Strategies, and Modifiability of Telehealth in Rural and Remote Emergency Departments: Qualitative Exploration Study

Christina Tsou[1*], BN, MPH, JD; Justin Yeung[2*], BMedSci, MBBS; Melanie Goode[2*], BSc; Josephine Mcdonnell[2*], BSc, MBChB; Aled Williams[2*], MBChB, MPH+TM; Stephen Colin Andrew[2*], BSc; Jenny Tetlow[2*], BEc; Andrew Jamieson[2*], MBBCh, MOccMed; Delia Hendrie[3*], PhD; Christopher Reid[3*], BA, MSci, PhD; Sandra Thompson[4*], BSc, MBBS, MPH, PhD

[1]WA Country Health Service, Curtin University, Perth, Australia

[2]WA Country Health Service, Perth, Australia

[3]Curtin University, Perth, Australia

[4]Western Australian Centre for Rural Health, University of Western Australia, Perth, Australia

[*]all authors contributed equally

Corresponding Author:
Christina Tsou, BN, MPH, JD
WA Country Health Service
Curtin University
Kent Street
Perth, 6102
Australia
Phone: 61 862661800
Email: tsouc01@gmail.com

Abstract

Background: Telehealth is a recognized and rapidly evolving domain in the delivery of emergency medicine. Research suggests a positive impact of telehealth in patients presenting for emergency care; however, the regional challenges of acute telemedicine delivery have not been studied. The WA Country Health Service (WACHS) established the Emergency Telehealth Service (ETS) in 2012 to provide telehealth and other technology-enabled services to regional Western Australian hospitals and clinics. The WACHS ETS supports 87 rural and remote WACHS-operated hospitals as well as 10 non-WACHS health clinics via high-definition audio-visual equipment installed in the resuscitation bay of the emergency department (ED) at each site. This 12-year practical application of emergency telemedicine offers a unique opportunity to explore the experiences and perceptions of clinicians delivering virtual care to rural and remote communities.

Objective: This study explores the perceptions of ETS clinicians regarding acceptability, appropriateness, and clinical decision-making when delivering emergency telemedicine in rural and remote settings.

Methods: This qualitative study used semistructured interviews to explore the perspectives of ETS clinicians regarding the factors influencing their clinical decision-making. It explored how ETS clinicians determine and modify clinical risks associated with using audio-visual equipment to deliver care. Emerging themes were compared with the concepts arising from the interim guidance of the Medical Board of Australia, and both the Australian and New Zealand, and American Colleges of Emergency Medicine.

Results: Overall, 16 doctors, 4 clinical nurse coordinators, and a nurse educator from WACHS ETS provided their experiences and perspectives. Accurate clinical decisions, especially regarding patient disposition, were crucial to virtual care. Timeliness and accuracy were enhanced through a mutual learning model grounded in the local context. Mitigation strategies such as improvisation and flexible technology use compensated for technological barriers. Nonmodifiable risk factors included patients' presenting complaints, clinical urgency of presentation, ED capability, clinician scope of practice, and, if a transfer was required, the distance between the ED of original presentation and the hospital of definitive care.

Conclusions: Telehealth can enhance clinical decision-making in rural and remote EDs, and ETS clinicians can prioritize patient safety through a lens incorporating both local hospital capabilities and community contexts. Even for the most experienced

XSL•FO
RenderX

clinicians, telehealth was not comparable to face-to-face communication in all circumstances. The impact of the ETS on the scope of the regional emergency medicine practice and on the building of clinical skills warrants further study in relation to its overall effectiveness and cost-effectiveness in rural and remote EDs. These findings identify areas for further qualitative research while providing a rich contextual background for rigorous quantitative analysis of the effectiveness of the ETS.

(JMIR Hum Factors 2025;12:e58851) doi:10.2196/58851

KEYWORDS

emergency telemedicine; implementation effectiveness; clinical effectiveness; risk aversion; risk mitigation; rural and remote; emergency departments

Introduction

Background

Telehealth has rapidly emerged as a means of providing health advice and care to people at sites where health providers at the hospitals of initial presentation (presenting hospitals) may lack capability or require assistance with the diagnosis and treatment of presenting clinical conditions. In recognition of the rapidly evolving domains of telehealth in the delivery of emergency medicine (EM) in the United States [1] and Australia and Aotearoa (New Zealand) [2], the Australasian College of Emergency Medicine (ACEM) [2], American College of Emergency Physicians [1], and Medical Board of Australia [3] deliberated on interim principles, considerations, and minimum standards to guide emergency telemedicine ("the interim guiding principles from the Colleges of Emergency Medicine"). Telehealth in the emergency department (ED) aspires to improve the quality of health care for patients in remote and other settings by increasing access to specialists and optimizing health service utilization and disposition coordination, including timely access to definitive care [2].

Guiding Principles on Emergency Telemedicine Implementation

These interim principles provide minimum standards for telehealth practice in the ED. As the Colleges of Emergency Medicine have identified, the use of telehealth in EM is in the formative stage and currently lacks a strong evidence base for clinical practice in the context of emergency care.

The ACEM and Australian New Zealand Fellow of ACEM (FACEM) Telemedicine Community of Practice (ANZFTCOP), superseded by a formal ACEM Emergency Telehealth Network, raised the importance of local context and the unplanned and exigent nature of emergency care [3]. The ANZFTCOP distinguished the nature of emergency medical practice from the broader telehealth discussion as it involves the provision of acute care without an ongoing clinical relationship either in-person or virtually [3]. The Community of Practice also pointed to circumstances where it may not be possible to access emergency physician expertise without recourse to telemedicine [3]. This is common for many rural and remote Australian EDs. The ANZFTCOP suggested that the decision regarding the appropriateness of conducting a telemedicine session instead of an in-person consultation should be based on individual clinical circumstances and expert clinician judgment [3]. Table 1 provides an amalgamated summary of the key guiding principles from these position statements and guidance documents.

XSL•FO
RenderX

Table 1. Amalgamated summary of the interim guidance on emergency telemedicine implementation.

Key themes	What is emergency telehealth and what should it do?	What is not emergency telehealth and what should it not do?
Locally grounded	• Should be an important complement for locally and regionally provided comprehensive health care. • Should be part of regional models of emergency care delivery validated by all relevant stakeholders across the public health system. • Should have regard for the local context and be grounded in the relationship with local clinicians including but not limited to knowledge of the local population, service availability, and patient pathways. • Should have appropriate escalation procedures in place to activate emergency services.	• Should not replace local care. • Should not be considered as a substitute for face-to-face consultations. • Should not involve low-value care that would not have otherwise been provided. • Should not be a stand-alone solution to address health workforce capacity and maldistribution.
Of similar quality and standard to ordinary care	• Assessments and consultations: (1) should be thorough within the limitations of telehealth and virtual consultations, (2) should replicate the components of an emergency medicine consultation, and (3) should be cognizant of the peculiarities or distinctive requirements of specific conditions. • Should implement quality assurance programs to monitor clinical performance, patient outcomes, and integrated ongoing care.	—[a]
Timely care	• Should ensure that definitive care is not delayed.	• Should not create additional barriers to accessing emergency medical care.
Appropriateness of care	• Should ensure that the patient's presenting problem is suitable to be assessed and managed remotely.	• Not appropriate for all clinical consultations.

[a]Not applicable.

Study Setting

The WA Country Health Service (WACHS) is one of the largest country health services in the world in terms of geographical coverage (2.55 million square kilometers, 96% of the land mass of Western Australia), with a population of 531,934 (18% of the Western Australian population; population density of 0.21 per square kilometer). It is organized into 7 distinct and diverse regions with 118 health facilities, accounting for 41% of the total number of emergency presentations across Western Australia (over 40,000 presentations per year) [4]. Most hospitals outside the large regional centers are staffed mainly by nursing staff supported by a resident general practitioner (GP) credentialed to provide medical services at the hospital. There are variable numbers of consultant specialists providing services at the larger regional resource centers. The majority of WACHS hospitals do not have on-site access to emergency medicine physicians (FACEMs).

The WACHS established the Emergency Telehealth Service (ETS) in 2012 to provide telehealth and other technology-enabled services to regional Western Australian hospitals and clinics. The ETS is the earliest and most established service stream of the WACHS Command Centre, which supports 87 WACHS hospitals and 10 additional non-WACHS facilities through high-definition videoconferencing services [5]. The Command Centre is a 24/7 virtual clinical and operational hub including services for general medical inpatients, mental health, midwifery, after-hours palliative care, and transfer coordination. It supports clinicians in regional hospitals and nursing posts by providing ready access to specialist clinicians who use technology, videoconferencing, and real-time data to assist in delivering quality patient care [6].

In addition to over 38,000 clinical consultations per year, the ETS supports rural and remote clinicians to maintain skills and knowledge through an education program and opportunities for professional development [5]. The consulting clinicians are located across multiple geographical areas (Perth, across regional Western Australia, other Australian jurisdictions, and internationally). Most medical practitioners are FACEMs complemented by EM-credentialed GPs and emergency nurse practitioners. The central hub of the Command Centre is based in Perth, and referrals are coordinated by a team of clinical nurses and clinical nurse consultants.

Referrals to the ETS vary across hospital types but are mainly from small hospitals (81.5%), nursing posts (9.7%), and integrated district health services (8.2%), and the ETS is least used by regional resource centers (0.7%) [4,7]. The proportions of ED presentations involving the ETS also vary across hospital types, with nursing posts and small hospitals involving the ETS in 21.3% and 23% of all ED presentations, respectively. Between July 2018 and April 2023, the uptake for nursing posts ranged from 2.8% for nonurgent presentations to 62.1% for emergency presentations that must be attended within 10 minutes. For small hospitals, the uptake ranged from 5.3% for nonurgent presentations to 47.9% for presentations requiring resuscitation. The more remotely located EDs appear to have greater reliance on the ETS especially for the management of more urgent cases [4,7].

Evidence on Telehealth Implementation in Rural and Remote EDs

Studies have identified factors that negatively influence the experiences of regional clinicians in rural and remote EM practices. These include [8,9] the lack or absence of medical backup or resources, the geographical and social isolation of

XSL•FO
RenderX

rural communities, poor health system coordination, the loneliness experienced by regional clinicians, and the challenges of health care provision in geographically isolated locations [10-12].

The challenges these conditions pose are community specific [8], and studies found that rural and remote EM physicians drew on the strength of professional relationships to improvise, solve problems, and create systems of support to meet local needs in order to survive and thrive in the middle of rural and remote challenges [8,9].

Published studies on the use of telehealth in EM suggested positive impacts on patient care, including benefits such as cost reduction, improved quality of care, reduced mortality rate, reduced patient treatment time, and reduced time between first contact and treatment [13]. Moreover, in the rural and remote context, studies have reported reduced patient transfers from rural facilities to major centers [10-12,14-16] and improved capability of rural centers [10,17,18]. However, studies have not discussed the complexities of rural and remote EM practices in depth or taken them into account in their analyses.

The use of telehealth in rural and remote EDs to help alleviate these regional EM practice challenges has not been studied. Whether or how emergency telemedicine in rural and remote settings improves patient health outcomes and the factors influencing the clinical decisions of emergency telehealth clinicians remain uncertain. There is a need for quality evidence to better understand these factors to help identify the gap in the effectiveness of emergency telehealth [12,14,15,17,19-27].

Studies have, however, identified challenges in the implementation of virtual care, and emergency telehealth is one of the practice areas considered [14,22,23]. These include technical difficulties [11,12,23,25-27], factors impeding the process of care [10,14,16,19-21,24-26], and factors impacting patient privacy, confidentiality, and data security [24-26].

It has been proposed that to understand the factors influencing job satisfaction, the complex environments around rural and remote EM practices must be considered [8]. However, the impact of environmental factors on clinical decision-making in rural ED practice has not been considered.

The clinicians' perceived clinical risks have been identified as a challenge in rural and remote EM practices [8,9]; however, how individuals perceive the risks in rural and remote emergency telehealth practices and the impacts on their clinical practices have not been explored previously. In rural and remote EDs, it is possible to assume that when the perceived clinical risk is high, the approach to diagnosis and treatment may change, leading to increased rates of transfer. The distinctions between these perceived clinical risks when delivering face-to-face versus virtual EM services warrant investigation.

To the best of our knowledge, specific links between the challenges of clinical reasoning and decision-making in virtual emergency care and the rural and remote context have not been identified. However, these factors are critical for understanding the effectiveness of implementing telehealth in rural and remote EDs.

Study Aims

From the perspective of WACHS clinicians delivering the ETS, this study explored the perceptions of emergency telehealth clinicians regarding the acceptability, appropriateness, and determinants in clinical decision-making when delivering emergency telehealth in rural and remote settings. This study aims to understand the following: (1) the factors influencing clinical decisions during an emergency telehealth consultation; (2) the impact of involving emergency telehealth clinicians in rural and remote ED consultations on the health workforce and the safety of clinical service delivery; and (3) how the clinical risks associated with emergency telehealth consultation can be modified, including whether improving the status of factors that are amenable to change can positively influence the safe and effective delivery of the ETS.

Methods

Overview

This qualitative study used semistructured interviews to explore the perspectives of ETS clinicians regarding the factors influencing their clinical decision-making. It explored how ETS clinicians determine and modify associated clinical risks. Emerging themes were compared with the concepts arising from the interim guidance of the Medical Board of Australia and both the Australian and New Zealand, and American Colleges of Emergency Medicine. The WACHS ETS was used to examine contextual factors influencing the quality and perceived effectiveness of emergency telehealth in rural and remote areas.

Sampling Techniques and Study Participants

A combination of convenience, snowball, and purposive sampling was used to obtain a diverse perspective on the ETS model of care. Convenience sampling was required given the nature of ED work and participant availability to recruit FACEMs, GPs, ETS nurse coordinators, and clinical nurses. A purposive sampling of WACHS ETS physicians who were interstates or overseas occurred toward the end of the participant recruitment.

On December 17, 2020, an invitation email was sent to all ETS doctors and nurses by one of the FACEM members of the research team. This was followed by individual follow-up emails on January 22, 2021, to ETS doctors, as well as emails and in-person follow-ups by one of the ETS clinical nurse consultants in the research team.

Interview Guide Development

The interview protocol (guide) was developed in collaboration with the clinical researchers at the Command Centre. This interview guide was then distributed to the full research team for review.

After piloting the interview protocol with a select leadership group, the protocol was updated to include probes on whether an electronic stethoscope would help provide confidence in clinical decision-making, as well as adding questions related to the clinical governance, service legitimacy, and legality of the ETS practice. Minor changes were also made to the wording and ordering of the content areas and probes within each content

XSL•FO
RenderX

area. The final version of the interview protocol is presented in Multimedia Appendix 1.

Simultaneous Data Collection and Analysis

Semistructured in-depth interviews explored the experiences and opinions of the clinicians involved in ETS implementation and the factors considered in clinical decision-making.

Participants were given the option to participate in an interview via MS Teams (Microsoft Corp) or phone, from a place and at a time of their choosing. Interviews were audio recorded, automatically transcribed, and repeatedly played back in the open-coding phase of the analysis to validate the accuracy of the transcription. Due to the nature of emergency work, the recruitment and conduct of the interview occurred opportunistically to catch moments when the clinicians were available. This meant that interviews sometimes occurred outside rostered work times, with some conversations interrupted by work-related demands.

Data analysis proceeded simultaneously with data collection and issues arising from information [28]. The early phase of the analysis included open and axial coding [29], identifying the emerging themes with open coding from clinician interviews. The emerging themes were then compared with the guiding principles identified from the interim guidance documents of the Colleges of Emergency Medicine and Medical Board of Australia. The key guiding principles were introduced as axial codes to place the open codes within the context of EM practice. In the final phase of the analysis, risk factors considered by clinicians in their clinical decision-making during an emergency telehealth consultation were extracted and classified into technology related, presenting hospital/location related, or EM related. The modifiability of each of the factors was assigned by the lead author and confirmed by co-authors based on author experience, the rural and remote context of Western Australia, the public health service landscape, and the resources available to the WACHS. The factual findings reported in this paper were discussed with and confirmed by the WACHS Command Centre.

Ethics Approval

This research obtained ethics approval from the WACHS Human Research Ethics Committee (approval number: RGS0000003076) and reciprocal ethics approval from the Curtin University Human Research Ethics Committee (approval number: HRE2019-0740-09). Site access was governed by the WACHS research governance policy and procedures. A participant information and consent form was attached to individual email invitations to the participants to review, sign, and return before the scheduled interview. For participants who attended the interview but were unable to sign the consent form, verbal consent was obtained before the commencement of the interview. Interview recordings were transcribed verbatim by the first author. The transcripts and recordings were allocated a participant code and stored in a password-protected electronic folder that is separated from the folder where the participant register is stored.

Results

Interviews and Participants

Most interviews were conducted via MS Teams (17/21, 81%), and the remaining were conducted by phone (4/21, 19%). Participants were in their private homes, a dedicated home office space, or Western Australian health facilities.

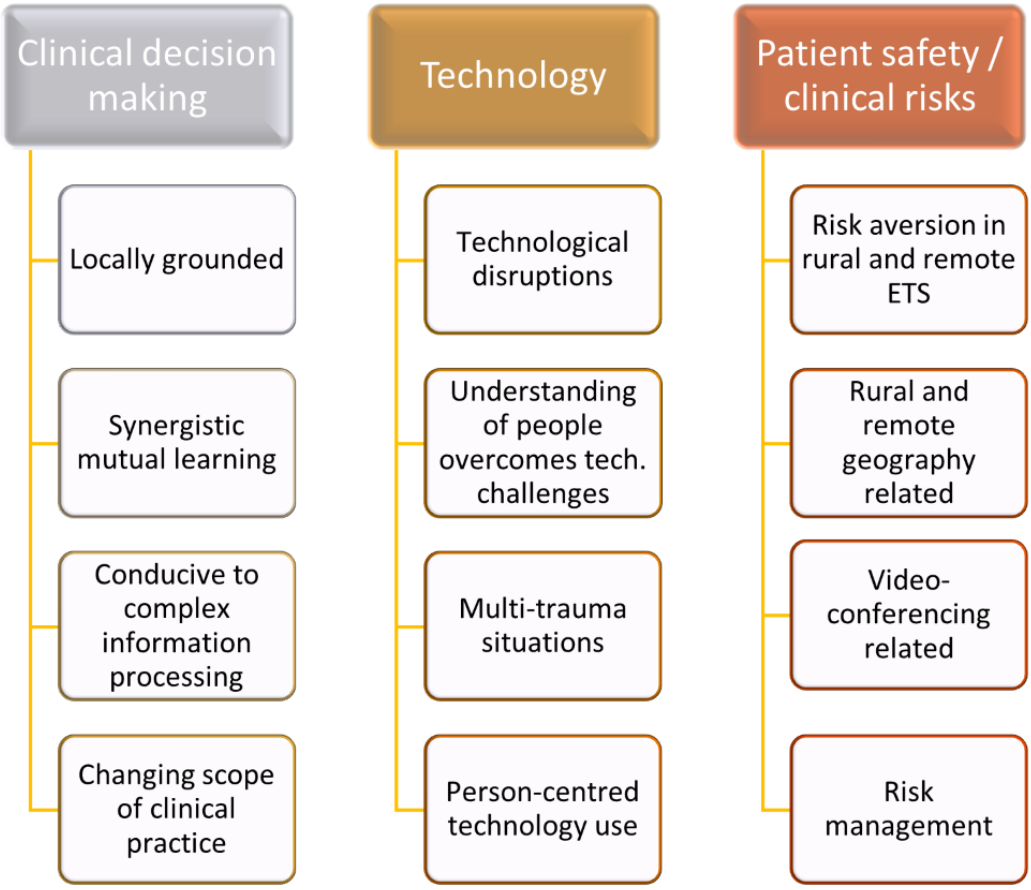
Sixteen WACHS ETS doctors and 5 ETS nurses participated in the interviews. There were 9 FACEMs, 5 GPs, 2 ED registrars, 4 clinical nurse coordinators, and 1 nurse educator. All participating ETS nurses were based at the Command Centre in Perth, while 5 of the ETS medical practitioner participants were based in Perth, 6 were based in regional Western Australia, 4 were overseas, and 1 was interstate. Two senior rural GPs provided their perspectives from both the receiving and provisioning ends of the ETS (as GPs working in regional locations). Around a quarter of ETS doctors and 16% of ETS nurses in the Command Centre participated in the interviews. All ETS nurses, except 1, were recruited by the same clinical researcher through face-to-face invitations, while 83% (15/18) of the ETS doctors who participated in the interviews responded to the initial invitation to participate. The purposive recruitment yielded 2 additional ETS participants to add weight to the perspectives of overseas or interstate doctors. Multimedia Appendix 2 lists the details of the participants, including home base, professional designation, gender, ETS role, and perspectives they contributed to this study.

Domains That Emerged From the Simultaneous Data Collection and Analysis

The major thematic domains identified in this analysis were clinical decision-making, technology, and patient safety or clinical risk. The themes and subthemes have been outlined below. Figure 1 provides a representation of the domains and themes.

XSL•FO
RenderX

Figure 1. Thematic domains and subthemes that emerged in this study. ETS: Emergency Telehealth Service; Tech.: technology.



Clinical Decision-Making: Benefits of Involving the ETS to Support Clinical Decisions

Participants identified 3 key benefits of ETS involvement in rural and remote clinical services. Hospital clinicians reported that having the ETS improved the willingness and confidence of regional clinicians to practice in remote locations. The ETS enabled synergistic mutual learning through the collaborative patient management of site-based and ETS clinicians. Further, ETS FACEMs experienced the benefit of ETS practice over site-based practice as it allowed them to be more cognitively focused when processing complex information. Participants touched on the changing scope of clinical practice as part of their discourse on clinical decision support.

Locally Grounded Clinical Decisions

Participants described timely and appropriate disposition (admission, discharge, or transfer) of patients as the most significant clinical decision for rural and remote EDs. Disposition of ED patients, especially from less-resourced remote EDs, has significant patient safety and clinical risk implications. Clinical decisions pertaining to patients presenting to small and remote EDs differed from those of better-resourced urban or regional center EDs as site-related factors needed to be considered.

All participants commented that multiple interacting factors were at play when deciding to transfer a patient for ongoing management. Table 2 lists the factors they considered in their clinical reasoning, and only 1 was related to technology. The predominant consideration was the resources available locally within the presenting hospital and in the local community. The next consideration was the distance to a destination site of definitive treatment if a patient's condition deteriorated. The final consideration was the timing of transfer and its appropriateness according to an individual patient's health condition.

XSL•FO
RenderX

Table 2. Thought processes in clinical decisions including patient disposition considerations via rural and remote emergency telehealth.

Thought processes of decision-making clinicians	Technology related	Hospital/location related	Patient related	Emergency medicine related
Normal diagnostic decision-making: what is wrong with the patient?	No	No	No	Yes
Can the patient be adequately assessed for all relevant possibilities at the site? What [differential diagnosis] am I potentially missing and is that putting the patient at risk?	Yes	Yes	No	No
Is this patient safe to go home medically and socially?	No	Yes	Yes	No
Is there a local GP[a] in town to care for the patient the next day?	No	Yes	No	No
Am I happy for the patient to wait till the next day for follow-up in a community?	No	Yes	Yes	Yes
Considering the patient's location, the weather conditions, and the local resources available, is the patient safe to stay at this site and/or be admitted locally for observation?	No	Yes	No	No
Does the patient need to be retrieved earlier [compared to the standard of a tertiary center or regional center ED[b]]?	No	Yes	No	No
Where and how do I transfer them? Is this patient safe to go somewhere with private transport or road ambulance, or do they require aeromedical transfer?	No	Yes	Yes	No
What does the patient want? (some patients do not want to leave their community)	No	No	Yes	No

[a]GP: general practitioner.

[b]ED: emergency department.

Synergistic Mutual Learning With Joint Clinical Decision-Making

ETS physicians were reliant on the local knowledge of hospital clinicians at the site of patient presentation to facilitate and consider community, social, and resource factors in clinical decisions for regional patients. Rural-based hospital clinicians reported benefits from the assistance of ETS clinicians in making clinical decisions, especially their support in connecting to appropriate and available resources.

Regional-based participants reported that the involvement of ETS colleagues supported them in answering the questions listed in Table 2, with increased speed and confidence. An ETS physician who also worked at a site based on a WACHS ED stated:

> *It makes certain that we make the best decisions for our patients. ETS does make the decision-making process much easier in the region. I've got somebody else online who can call me if I'm looking after somebody who's critically ill. They can help organize transfers and ICU admission, they can help me with drug doses, so I can…crack on and do what I need to do. And I know that I'm working in a safe environment.* [ETS Doctor #3]

Participants also stated that they acquired a different skill set working in a supportive role to clinicians in remote facilities, including ways of communicating instructions to presenting hospital clinicians (mainly nursing staff), gathering medical history, prescribing medications, and efficiently analyzing and synthesizing presenting complaints.

For ETS clinicians who were FACEMs mainly based in metropolitan hospitals, participation in ETS consultations provided a new and unfiltered experience in the rural and remote clinical practice. Many reported that this changed the way they practiced in their substantive role on the floor of tertiary EDs:

> *I think [the learning has been] in both directions, one, it makes me realize how resource rich we are in a tertiary hospital. Two, it makes me realize that the nurses and other staff out there [in regional WA] are often just as competent, just as knowledgeable, and just as skillful and often braver than we are in tertiary hospitals. It certainly helps me stay grounded when I'm accepting referrals. [Tertiary hospital] has a big catchment, we receive a lot of these people from peripheral hospitals and nurse run clinics. And one that certainly helped me know where all these places are, when I'm on the receiving end of the phone…I'm able to explain to other staff exactly how resource limited, a lot of these clinics are. And there is a reason they haven't had blood tests and haven't had imaging. And in fact, there's no doctor there. So they have just been seen by nurses must come to be seen by us, we don't have to be difficult about accepting them. I think that's been quite useful.* [ETS Doctor #43]

Conducive to Complex Clinical Information Processing

Participants reported that their ability to control and prioritize their consultations in the virtual environment assisted their efficiency in processing complex information and enabled more objective assessment than working in person on the ED floor, given their distance from the patient. Participants reported the

XSL•FO
RenderX

benefits of providing virtual care compared to working on an ED floor:

> *If you're on an ED floor you are interrupted constantly because you're physically there…one of the most frustrating things is you're constantly distracted, and you can't finish a task from A to B, I think that adds hugely to burnout. …But you are focused here [at ETS], …we can only see one camera…one patient at a time. …And the chaos is more manageable… you can choose to not be interrupted. So I feel that it's very satisfying, because you're so cognitively focused. And I think that can potentially lead to better quality care… because you can just make better decisions.* [ETS Doctor #6]

With the ability to control the environment they practice in, ETS FACEMs found the virtual environment conducive to them being more cognitively focused.

Changed Scope of Clinical Practice

The impacts on the scope of practice of regionally based clinicians, especially nurses, were dichotomously reported. Most participants reported an increasing scope of site-based clinical practice with ETS doctors guiding local nurses to perform procedures that they otherwise would not have the confidence to do on their own. A contrary perspective was the engagement of ETS clinicians in treatment and procedures or clinical decisions, which local clinicians would have performed independently through a local process in place prior to ETS implementation. For example, regarding giving a tetanus booster, 1 participant commented:

> *…we'll get consults of patients who need a tetanus booster, they've had a small wound a few days ago, …I don't know what they would have done before, I think they would have done it and had a process in place to do it.* [ETS Doctor #2]

The changing scope of practice was noted by the participants. This reflected policy changes to increase medical practitioner involvement in ED consultations and was an emerging theme that warrants further in-depth exploration.

Technology

The second thematic domain was technology, which included technological (digital) disruption and mitigation strategies in emergency telehealth consultations.

Participants were of the view that the WACHS ETS had already made most of the technological and process improvements and considered that any further changes on the technological front would have incremental benefits in effectiveness for improving patient health outcomes. One FACEM participant stated:

> *A lot of the system problems are the fact that we have a hundred places that are tiny and it's hard to get [for example] a pathologist, or experienced nurses for long stays. A lot of it is system problems that are difficult to modify.* [ETS Doctor #43]

> *We probably made the easy gains and the incremental changes you can make to improve things from here are big and costly, and difficult. For example, attracting and keeping people in remote communities and providing enough staffing to let GPs take holidays and do professional development and be replaced by someone, rather than leaving the community to rely on telehealth because there is no one else to cover.* [ETS Doctor #43]

Participants described the ways in which they modify how they use the telehealth technology or adjust the process of care to improve the quality of ETS consultations.

Technological (Digital) Disruptions

A commonly expressed view among ETS physician participants was that during videoconferencing, one is expected to lose nonvisual cues and some verbal cues due to technological (digital) disruptions. Clinicians reflecting on the videoconference consultation experience suggested that the amount and quality of information a clinician could gather through video came down to their level of understanding of people as individuals. They referred to the nuances of patients' expressions observable during videoconferencing and being able to relate to the regional context of the patients. Table 3 provides a summary of the circumstances around the technological barriers during an ETS consultation and the mitigation strategies participants used to overcome these technological disruptions.

XSL•FO
RenderX

Table 3. Technology-related issues of Emergency Telehealth Service technology and mitigation strategies.

ETS[a] technology-related issues	Technology barrier			Mitigation strategy		
	Barrier	Circumstances	Participants	Strategy	Incompatibility	Participants
Positioning of the monitor where the image of the ETS physician appears (ie, at the foot of the ED[b] stretcher)	• Sending the wrong message that the doctor is looking down on the patient. • Abrupt appearance does not allow ETS to ease into the scene as in the case of an in-person ED floor.	When people on the floor are not adequately prepared and people are unaware why the ETS clinician is appearing on the screen, there can be disruptions in the actions already taking place in the ETS bay.	ETSD[c] #1, 4, 6, 7, 10, and 13; ETSN[d] #3	• Prepare patients before the VC[e] link is up; explain the process before and after VC consultation, including the steps leading up to transfer when interhospital transfer is indicated • ETS physician is identified as not from Perth and thus does not appear as "someone from Perth talking down on us"	• In a critical encounter, there is no time and space to prepare patients and people in the room. • Only available to regional-based ETS clinicians.	ETSN #3; ETSD #7
Talking into an open space; broadcasting of sound from the VC setup into the ETS bay	• Impersonal, with privacy and confidentiality potentially compromised; concurrent private conversations can occur from adjacent ED bays, which can create distracting noise during the consultation process.	Especially when multiple parties are involved in the conversation	ETSD #1, 4, 7, 10, and 13; ETSN #3	• Mute VC sound and use a telephone directly with site-based clinicians or patients for audio to ensure private conversations. • Transition to telephone only after an initial meet and greet with VC. • Clinicians on the floor are encouraged to wear headphones to enable switching voice channels between different target audiences.	• If children are involved, communicate with parents or carers	ETSN #3 and 4
The ETS clinician's visual field of the ETS bay is restricted by the camera's visual field	• ETS clinicians are not aware of others in the room.	Consulting before realizing a family member or another patient is in the room; knowing who is in the room and their relationship with the patient can change the dialogue and alter the instructions given regarding the next steps.	ETSD #1 and 6	• Use a telephone before screen appearance, find out who is present, and clarify the purpose of involvement. • Self-introduction, role of ETS consultation, and a name shown at the base of the screen.	—[f]	ETSD #2, 4, 8, 9, 12, 13, 15, and 43
Zoom in or out function adjusted locally without the ETS clinician's awareness	• Unable to zoom in to see what needs to be seen.	Significant (disastrous) impact in critical situations.	ETSD #5	• Having a reliable camera that the ETS doctor can control.	—	ETSD #4 and 14

[a]ETS: Emergency Telehealth Service.

[b]ED: emergency department.

[c]ETSD: ETS doctor.

[d]ETSN: ETS nurse.

[e]VC: videoconferencing.

[f]Not applicable.

The videoconferencing equipment setup and the restriction of the camera's visual field created technological barriers affecting the quality of ETS consultation. Physical distance was the perceived distance between either side of the high-definition audio-visual equipment installed in the resuscitation bay of the ED at each site.

XSL•FO
RenderX

Participants experienced constraints of the camera setup made consultations impersonal and visual cues easily missed or distorted, and they shared how they adapted in their ETS practice to ensure effective visualization. One doctor said the technical issues and logistics of commencing a consultation with sudden appearance on a screen and sound broadcasting as key differences to working on the floor:

> *…, you got to spend time to dial in, and, then suddenly just appear on the screen in front of the patient… sometimes they're a bit shocked because you're big and loud on the screen… especially if they're drunk, confused that can exacerbate that [shock]. And sometimes I don't know what the volume is, sometimes I'm absolutely booming. And then everybody can hear me in the ED… they got to find the mute button, unmute me and decrease the volume…* [ETS Doctor #6]

A few participants suggested that adding a reliable remote-control pan or zoom would considerably improve the quality of the ETS physical assessment:

> *…great history and reading the cues if someone is lying very still, are they squirming around, are they making reasonable eye contact, are they coherent when they talk, all of that sort of stuff... you know looking for facial droop, obvious distal neurological deficit can be done really well [with reliable and high resolution camera].* [ETS Doctor #14]

Understanding of People Can Overcome Technological Challenges

Through reflective practice, participants redefined and modified their practice over the ETS to mitigate the shortcomings of the videoconference equipment setup. There were varied responses to technological (digital) disruptions in interpersonal communications. ETS doctors gave examples of how they used visual and imperfect auditory cues to assist with formulating their working diagnosis when direct eye contact was not possible via videoconferencing. Visual cues, such as closing hands and fidgeting, were used to the clinician's advantage to help with diagnosis. Humor and a nonjudgmental approach were helpful in getting patients to reveal their alcohol issues and to work collaboratively with patients over videoconferencing to reach the best outcome for them. These human approaches made a notable difference in the virtual physical assessment process.

Multitrauma Situations With One Videoconference Bay

In a multitrauma situation when only 1 doctor was available on site, teamwork was required. For example, when the site-based doctor was busy with hands-on procedures in collaboration with an ETS FACEM, a local nurse stepped up to perform the team leader tasks under the guidance of ETS clinical nurses. When there were multiple casualties, all casualties would need to rotate to be in front of the cameras (typically 2 perpendicular pan/tilt/zoom cameras), and by using headphone channel switching, ETS clinicians could converse with each of the site clinicians individually. When there was a requirement to "shuffle" beds in and out of the camera visual field, adequate explanation ahead of the action was noted as an indispensable component of patient-clinician communication.

Person-Centered Technology Use

Facing variations in the quality of the internet connection, fixed positioning of cameras, and the broadcasting of sound into an open space of the ETS bay at regional EDs, ETS clinicians developed approaches to work with and via the presenting hospital clinicians to improve the quality of videoconference consultation while also protecting the privacy and confidentiality of patients. Overcoming technological limitations required users to troubleshoot and adapt to circumstances, requiring flexibility to ensure the effectiveness of virtual care. This finding pertains more to the exigent nature of EM and less to the rural and remote context of ETS delivery.

Patient Safety and Clinical Risk Considerations

The final thematic domain was related to the perceived risks of virtual care identified by the study participants (Table 4). The central aim was to ensure patient safety when using telehealth. If the ETS team perceived that patient safety was compromised by remaining at the local facility (particularly if there was no on-site medical practitioner), they would strongly consider transferring patients to an appropriate hospital for ongoing management. ETS clinicians who had higher levels of perceived clinical risk had a lower threshold to transferring their patients. Participants were confident that with experience (both at the site and virtually), some transfers could be avoided. The modifiability of risks as assessed in this study identified areas for potential improvement and the factors to consider in the effectiveness dialogue.

XSL•FO
RenderX

Table 4. Factors that make emergency telemedicine clinicians risk averse in rural and remote areas in Western Australia.

Clinical risks and sources of risks in the Emergency Telehealth Service practice	Technology related	Hospital/location related	Emergency medicine related	Is it modifiable?
Insufficient or lack of local resources				
Nature of ED[a] cases: sudden deterioration not always predictable	No	No	Yes	Not modifiable
Distance/time required to transfer to a location with more resources against the risk of deterioration	No	Yes	No	Not modifiable
No short-stay unit: cannot keep the patient after hours	No	Yes	No	Potentially
Local workforce capacity and capability				
Clinicians on the ED floor are unable to perform the required physical assessment	No	Yes	No	Limited
Procedural (IV[b] access and treatment procedures)	No	Yes	No	Potentially
No doctor or insufficient nurses onsite to admit the patient	No	Yes	No	Potentially
No GP[c] in town to follow-up on the next day	No	Yes	No	Limited
Unable to secure a clear diagnosis				
Unable to access required testing or imaging	No	Yes	No	Not modifiable
Unable to fully examine: chances of missing a clinical cue are greater	Yes	No	No	Limited
Not 100% confident with the clinical picture despite full complement of clinical information	No	No	Yes	Not modifiable
Variable internet quality: disrupts physical assessment	Yes	No	No	Limited
Camera is not always reliable	Yes	No	No	Modifiable
Individual factors				
Physician personality and approaches to risk	No	No	Yes	Not modifiable
Physicians are less comfortable with the reliability of their own physical assessment over VC[d] compared to face-to-face evaluation	Yes	No	No	Potentially
Patient preference (to stay closer to home)	No	No	Yes	Not modifiable
Patients and family members are risk averse to telehealth consultation	Yes	No	No	Potentially

[a]ED: emergency department.

[b]IV: intravenous.

[c]GP: general practitioner.

[d]VC: videoconferencing.

Concept of Risk Aversion in the WACHS ETS

Risk aversion impacts transfer rates from the site of presentation. The level of risk ETS clinicians were willing to take in keeping patients locally depended on 4 broad categories risk considerations. The first and second considerations of risks were related to the presenting hospital capacity and capability (ie, the health service resources and workforce). Risk aversion was associated with increasing remoteness of the presenting hospital. The third category of risk was unanimously reported by all the ETS physician participants. They reported that they were not able to examine the patient adequately or be completely confident with the clinical information relayed by site-based clinicians. This concern reflected how well ETS clinicians felt they could apply strategies to mitigate technological (digital) disruptions. The final category reflected individual clinician factors, such as their personality and their own tendency to be risk averse, and was present in EM generally, irrespective of whether it was place based or virtual. Factors related to the nature of EM and patients presenting to EDs would not have changed this risk even if managing clinicians were able to be "teleported" to the scene in person, and these aspects have not been discussed further in this paper.

Risk Aversion Associated With the Hospital of Presentation and the Rural and Remote Geography

Of 17 risks, 8 emerged from the ETS risk-aversion discussion related to the hospital of presentation and the rural and remote geography. These were less readily modifiable than the videoconference-related risks. The least modifiable risk was related to the inability to access the required tests or imaging, and this was reported by multiple clinicians. If point-of-care testing and the pathological approach available locally were not adequate to reach a clear diagnosis, the patient would be

XSL•FO
RenderX

transferred for further investigation. The extent to which this was considered an issue varied, with 1 participant (ETS doctor #3) having a view that point-of-care tests, routine pathology, and x-ray imaging available locally were generally satisfactory and sufficient for many of the presentations, and this view differed from that of many other participants.

Another factor associated with the presenting health facility was related to the regional workforce capacity and capability to perform the required physical assessment or procedure, or to keep patients locally for observation onsite or in the community. This mainly applied to sites where there were no medical practitioners. These factors had limited modifiability at the time of the consultation but were potentially adaptable. When the ETS clinician was unable to work with the clinician onsite to overcome the limitations of physical assessment over videoconferencing, the workforce capacity at the presenting hospital impacted the ability to secure a clear diagnosis and manage the patient adequately. Workforce capability was also related to the inability to perform required procedures ranging from establishing vascular access, inserting an intraosseous needle, and simple suturing to performing more complicated procedures, which could have avoided transfer. Generally, this is related to nursing-only facilities staffed by clinicians with limited scope of practice. Even when a diagnosis was made and the required procedures are performed, a patient might still be transferred if there was no doctor or insufficient nursing capacity to keep the patient on site for further observation, including if there was no GP in town the next day to follow-up in the community.

The final factor adding to the risk aversion of ETS clinicians was related to resources available at the hospital of presentation. The key issue raised here was the absence of a short stay–type inpatient unit to keep patients in the hospital for overnight observation. The distance and time required to transfer patients to a location with higher levels of capacity and capability were considered against the risk of deterioration of the patient's condition and the assessment of clinical risk. These factors were not modifiable and subject to the treating clinician's judgment of the patient's health condition.

Videoconference-Related Risk Aversion

Of the 4 categories of ETS clinical risk, 2 included factors that were related to the inability to secure a clear diagnosis and were related to videoconferencing. Many participants discussed the need to transfer patients because they were not confident with the clinical assessment via video. These included their perception of missing significant clinical signs, due to being unable to directly palpate or examine the patient, which impeded reaching a clear diagnosis. For example, 1 FACEM participant commented as follows:

> *…spending a lot more time thinking about exactly what I'm going to do and why… you don't have any of the nice reassurance normal investigations you would normally have. You are putting yourself more at risk by doing telemedicine because the chances of missing something are greater because you don't have access to all the little things that would make you much more comfortable.* [ETS Doctor #14]

ETS physician participants also reported issues with internet connectivity and speed, and many made comments on issues associated with the reliability of the camera affecting their ability to formulate a clear diagnosis. As discussed above, issues related to the camera were associated with the way it was physically positioned in the allocated treatment bay; operational limitations with pan, tilt, or zoom; and the camera technology itself. The internet speed and connectivity were more challenging as they were impacted by the internet provision where both the patients and ETS clinicians were physically based (nationally and internationally).

Finally, user acceptance of videoconference consultation differed among participants. Not seeing patients face-to-face made some clinicians more cautious and risk averse, with treating clinicians also reporting that some patients or family members were risk averse to telehealth consultations due to a lack of experience or understanding of this alternative modality of clinical service delivery. Several ETS clinicians reported that they spent more time second-guessing their observations and decisions, suspecting that they are less reliable over videoconferencing compared to a face-to-face meeting.

Risk Management for Patient Safety Instead of Risk Aversion

Complementing the higher level of potential and perceived risk associated with videoconference consultations, participants shared a range of perspectives on the perception of clinical risk:

> *…am I risk averse? I am probably less. It's not my personality. My personality is not risk averse. And I think most emergency physicians are not risk averse. That's not the nature of our business. There's risk everywhere we are risk managers.* [ETS Doctor #2]

Despite observing the potential risks associated with videoconference consultations, participants believed that those who were overly risk averse would not participate in telehealth, and it was a matter of finding a balance between the perceived benefits of ETS and the expected risks associated. A pragmatic view was that ETS physicians were no more cautious on camera compared to being on the ED floor; however, due to incomplete data points to inform an evidence-based clinical decision, interhospital transfer may be suggested for the required investigations. In terms of confidence in their decisions, some participants reported that they felt comfortable with their clinical decision-making as FACEM training helped balance relatively scant information in undifferentiated patients, and this was no different from the decision they would make in a face-to-face meeting:

> *Well, that's my job. It's not so much about being confident. Well, you've got to be confident in your decision because otherwise you can't afford to be in the [emergency medicine] business. I mean, you know, if you if you're worried about the patient, then you do something to alleviate that worry… Occasionally, there are times when it is difficult… for example, if a patient has a communication problem in either they are deaf, or they can't speak properly, and so my interaction is sufficiently degraded but otherwise, for*

XSL•FO
RenderX

> *[most] times, I don't think it does make a particular difference to be honest. The difference between me not being there and being there is far more evident in the skills that are brought face-to-face. And those are nearly all procedurals.* [ETS Doctor #5]

Discussion

Overview

The experiences and perspectives of WACHS ETS clinicians contributed to an in-depth understanding of the factors influencing clinical decision-making during an emergency telehealth consultation in rural and remote settings. The study progressed the understanding of the mutual learning opportunities presented during the joint decision-making process, the perceived clinical risks of emergency telehealth practice in the rural and remote context, and the technological barriers imposed on clinical decisions and the modifiability of these perceived risks.

Principal Results: Meeting the Principles of the ETS Practice

This research described the mutually beneficial effects of ETS consultation recognized by the participants in their clinical practice. These were in line with the principles and interim standards stipulated by the Colleges of Emergency Medicine that telehealth should complement and be part of the regional model of emergency care delivery validated by all stakeholders [1-3]. This finding also aligns with previous research, which highlights how rural and remote ED clinicians rely on the support and expertise of their colleagues to navigate and succeed in the unique challenges of rural and remote settings [8-10,17,18]. It further explored how this dynamic operated when colleagues provided support through a virtual presence. This research provides additional insights into how collegiality during emergency telehealth consultations supports various aspects of practice. These include enhancing confidence in rural and remote clinical practices, advancing the practice of virtual EM, and incorporating the rural and remote context into collaborative clinical decision-making involving stakeholders on both sides of the technology.

Clinical Decision-Making

The participants described clinical decision-making mediated by a complement of ordinary technologies. Considerable regard was given to the context of the hospital of initial ED presentation, the community where the patient is from, and the regional Western Australia health service delivery landscape. ETS clinicians were reliant on the local knowledge of presenting hospital clinicians, the patients, and their family members to incorporate these local contexts into collaborative decision-making. The capacity building was mutually beneficial for regionally based clinicians and ETS clinicians. Their complementary skills and synergies in what they were able to contribute to patient care became apparent during ETS consultations. In this paper, we have referred to this phenomenon as “synergistically mutual learning.”

Many participants described their experience of improved efficiency in processing complex clinical information through better focus. This reflected the time-sensitive nature of the ED, especially when the hospital of definitive treatment was hours away, and the organized chaos in multitrauma situations. This finding represents a novel contribution from the perspective of emergency physicians conducting virtual ED consultations. The quality of clinical reasoning was reported to be superior in ETS practice where there was greater reliance on synthesizing medical history, reviewing clinical records, and carrying out focused processing of multiple sources of information. In situations where there was a heavy reliance on physical examinations and complex laboratory or imaging investigations, clinicians were cognizant of the peculiarities and distinctive requirements of the specific conditions and replicated, as much as possible, the components of an EM consultation. This was in line with the guiding principles stipulated by the Colleges of Emergency Medicine to deliver a service of similar quality and standards as ordinary care [1,2].

The perspectives of ETS clinicians highlight a critical aspect of improvement in the overall efficiency of emergency care delivery not previously reported in studies of the efficiency of emergency telemedicine within receiving hospitals [10-12,14-16].

Clinical Use of Technology

At the forefront of technological change is how humans personalize technologies. The way technologies are used is a significant part of the innovation in the EM practice in rural and remote settings. This study extended past findings of technological difficulties impeding care processes [10,14,16,19-21,24-26], patient privacy, confidentiality, and data security [24-26] by explaining how emergency telehealth practitioners introduced human factors to the innovative use of emergency telehealth through their understanding of people and tailoring the use of the same technology to the needs of patients on the other side of the camera as described in Table 3. To the best of our knowledge, this is the first description of how human users innovate to close the distance across technological barriers in an emergency telehealth scenario.

Modifiability of Patient Safety and Clinical Risks Identified

Risk factors related to ETS technology were modified through upgraded camera resolution, fixed camera positioning, increased user experience, improvisation, and consumer preparation. The quality of physical examinations was generally considered unlikely to match the quality of face-to-face examinations, especially when examinations of hard-to-visualize areas of the body or procedural interventions were indicated. However, emergency telemedicine using videoconference facilities has been considered to have reached a plateau and might be sufficient to cover a great majority of presentations. Additional investment in other examination equipment will need to be carefully evaluated as the benefits of further equipment upgrades are likely to be incremental against the costs of purchasing the equipment.

To meet the standard of replicating components of an EM consultation stipulated in the interim guidance [1,2], the need for an appropriately skilled and competent local workforce

XSL•FO
RenderX

cannot be overlooked. Workforce capabilities were potentially modifiable through education, including pre- and postvocational training opportunities; however, modifiability was offset by the ongoing issue of the turnover of clinical staff in rural and remote hospitals. This said, participants provided examples of how inter- and intraregional movement of clinical staff helped to retain the necessary skill sets within Western Australia. The extent to which clinical skills acquired through ETS involvement in ED consultations benefited the overall capability of the regional clinical workforce in WA fell outside the scope of this paper; however, this is an area of significant importance and relevance to the effectiveness of the ETS.

The pivotal but nonmodifiable consideration in clinical decision-making was the physical distance from definitive treatment locations and presenting hospitals. Local hospital resources, capacity and capability were beyond the scope of this study, but it was acknowledged that these had limited modifiability, at least in the short term.

This paper did not explore the impact of the ETS on the quality and standard of care in larger regional EDs, and the dominant view that EM specialists are risk managers is of significance. Any manifested risk aversion is likely related to patient safety considerations given that rural and remote contextual factors are nonmodifiable or have limited modifiability. Despite the dominant view that the ETS added value to presenting hospital clinical decisions, there were also concerns that pre-existing processes and capacity had been disrupted by the care processes now available through the ETS. The change in the scope of practice in presenting hospitals, particularly regarding nursing staff, because of ETS involvement is a subject for further detailed exploration to reach an optimal balance.

Novel Contributions of This Study

Prior work has paid considerable attention to advocating for telehealth policy reform [30] and identifying the barriers and enablers of telehealth implementation in rural and remote EDs. This is the first study to systematically detail the complex interrelationships between technological and human interfaces. The barriers to quality and safety of emergency telehealth consultations and mitigation strategies used by the WACHS ETS clinicians when collaborating with rural and remote EDs were detailed. These perceived clinical risks and risk mitigation strategies added new insights into the contextual factors influencing the effective implementation of rural and remote ETS.

Findings around technological challenges were consistent with the findings of previous studies that pointed to variations in usability, depending on the setting and situation. Unlike previous studies that compared the approach of different types or combinations of telehealth technologies (eg, videoconferencing, store and forward, and telemonitoring) [21], this study embraced the use of available technology as a norm in rural and remote ETS practices in Western Australia. The focus on the strategies used in WACHS ETS delivery to overcome technological barriers in this study makes a novel contribution to advancing our knowledge on the modifiability of the risks associated with rural and remote virtual care and helps to inform future quantitative analysis and interpretation of findings on the effectiveness of rural and remote ETS.

Limitations and Future Research

The findings reported in this paper are subject to several limitations. Clinical nurse coordinators are the first contact points for accessing the ETS, and this triaging (referred to as ETS prioritization to distinguish from the Australasian Triage Scoring system) plays a pivotal role in case review order and selection. The ETS prioritization process ensured that the presenting hospital and community context were incorporated into timely and appropriate care from the first point of ETS contact. This paper focused on regional context considerations influencing clinical decisions around the disposition of ED patients during a virtual consultation and did not examine regional context factors already incorporated during the process of ETS prioritization. Future reports on ETS prioritization would be an essential addition to the body of knowledge on emergency telehealth implementation in the rural and remote context.

Telehealth, especially emergency telehealth, is a relatively new area in clinical practice with evolving clinical governance standards. Medical indemnity issues were discussed during the interviews but did not emerge as a dominant theme in the data analysis. WACHS ETS physicians are particularly focused on patient safety. However, based on available data, it is difficult to conclude the extent to which medicolegal considerations are at play in the response to the risks associated with the ETS practice.

The financial viability or cost-effectiveness of virtual care is best explored quantitatively. It is worth noting that participants were value conscious in evaluating disposition options. However, they were not able to comment on the financial viability of the service as it was not within the role of most of these clinician participants. The implementation factors identified inevitably influence the effectiveness of telehealth in rural and remote EDs but are insufficient to determine the effectiveness of the ETS. This study represents a step toward determining the effectiveness of ETS delivery and is an area for future research.

The recruitment of participants may have been subject to social desirability bias, and some nonresponders may have had an alternative opinion about the service, which they felt did not align with the health service's official position. The presenting health condition is one of the central tenets in this research, although the data collection did not focus on specific health conditions. Our intention in this paper was to capture the overarching themes of clinical decision-making and how they affect the safe and effective delivery of rural and remote ETS to improve patient health outcomes. There were no comparisons between the treatments of urgent and nonurgent presentations because formative analysis showed the tendency of using the ETS for more urgent cases. This focus was due to the higher proportion of FACEMs among the study participants, who were more likely to encounter urgent cases.

Conclusion

Telehealth has the potential to improve the processes of clinical decision-making in rural and remote EDs. The participants, all

XSL•FO
RenderX

of whom participated in delivering the ETS, perceived improved accuracy, confidence, and speed of clinical decisions. The reach and effectiveness of the ETS in addressing the identified health needs remain to be tested.

Patient safety was central to all clinical decisions, which were grounded in the capability of presenting hospitals and the local community context. The risks of delay in definitive care for those patients requiring transfer for ongoing management associated with the geographical distance to the hospital for definitive treatment and the presenting hospital capability were not amenable to change. The importance of social and geographical contexts for appropriate telehealth delivery was highlighted in this study.

Technological barriers to the safe and effective delivery of the ETS were amenable to improvisation and improvements to infrastructure, but overall, participants were aware that they could not reach the same reliability as that in face-to-face consultation for all presenting conditions. Although the ETS was an acceptable and feasible option for ED consultation, the maintenance and sustainability of the ETS as part of routine practice and policy are dependent on presenting conditions.

Investment into rural and remote telehealth has the potential to support regionally based clinicians both directly in patient care and indirectly by promoting their (health workforce) longevity in regional areas through the ongoing development of the virtual care network. The effect of the ETS on the scope of the regional ED practice and the building of clinical skills within the region warrants further in-depth exploration including its potential implications for the overall effectiveness and cost-effectiveness of telehealth services in rural and remote EDs.

Acknowledgments

This research was initiated by the WA Country Health Services (WACHS) and Curtin University, and received financial support from both institutions as well as the Digital Health Collaborative Research Centre (DHCRC) in the form of PhD scholarships. The WACHS also contributed in-kind support in the form of clinician time, industry supervision, and administrative and research infrastructure support. This research received further in-kind contributions from the School of Population Health at Curtin University, WA Centre for Rural Health at the University of Western Australia, WACHS Command Centre, and WACHS Country Health Innovation Project Support Office in the form of academic supervision, industry mentorship, and knowledge exchange.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide for Emergency Telehealth Service clinicians.
[DOCX File , 18 KB - humanfactors_v12i1e58851_app1.docx]

Multimedia Appendix 2
The professional designations and perspectives of study participants.
[DOCX File , 17 KB - humanfactors_v12i1e58851_app2.docx]

References

1. Sikka N, Gross H, Joshi AU, Shaheen E, Baker MJ, Ash A, et al. Defining emergency telehealth. J Telemed Telecare 2019 Dec 11;27(8):527-530. [doi: 10.1177/1357633x19891653]
2. Interim Position Statement: Teleheatlh in Emergency Medicine (S843). Australasian College for Emergency Medicine. URL: https://www.mdanational.com.au/-/media/files/pdfs-linked-to-library-items/s843_interim_statement_telehealth.pdf [accessed 2025-01-30]
3. Public consultation - Guidelines: Telehealth consultations with patients. Medical Board of Australia. URL: https://www.medicalboard.gov.au/News/Past-Consultations/2023.aspx [accessed 2023-07-24]
4. WA Country Health Service. Emergency Department Data Collection. Perth, Western Australia: WA Country Health Service; 2025.
5. Emergency Telehealth Service. Government of Western Australia. URL: https://www.wacountry.health.wa.gov.au/Our-patients/Telehealth/Emergency-Telehealth-Service [accessed 2025-01-30]
6. Command Centre. Government of Western Australia. URL: https://www.wacountry.health.wa.gov.au/Our-services/Command-Centre [accessed 2025-01-30]
7. WA Country Health Service. Emergency Telehealth Collection. Perth, Western Australia: WA Country Health Service; 2023.
8. Kolhatkar A, Keesey A, Bluman B, Lynn B, Wilkinson T. Understanding how emergency medicine physicians survive and thrive in rural practice: a theoretical model. Rural and Remote Health 2017;17(4):1-14. [doi: 10.22605/rrh4285]

XSL•FO
RenderX

9. Setrinen Hansen N, Mikkelsen S, Bruun H, From E, Milling L. Physicians' experiences working in emergency medicine in a rural area in Northern Sweden: a qualitative study. Rural and Remote Health 2021 Dec;21(4):1-11. [doi: 10.22605/rrh6672]
10. du Toit M, Malau-Aduli B, Vangaveti V, Sabesan S, Ray RA. Use of telehealth in the management of non-critical emergencies in rural or remote emergency departments: A systematic review. J Telemed Telecare 2017 Oct 05;25(1):3-16. [doi: 10.1177/1357633x17734239]
11. Brainard JS, Ford JA, Steel N, Jones AP. A systematic review of health service interventions to reduce use of unplanned health care in rural areas. J Eval Clin Pract 2016 Apr 28;22(2):145-155. [doi: 10.1111/jep.12470] [Medline: 26507368]
12. Culmer N, Smith T, Stager C, Meyer H, Quick S, Grimm K. Evaluation of the triple aim of medicine in prehospital telemedicine: A systematic literature review. J Telemed Telecare 2019 Jun 25;26(10):571-580. [doi: 10.1177/1357633x19853461]
13. Sharifi Kia A, Rafizadeh M, Shahmoradi L. Telemedicine in the emergency department: an overview of systematic reviews. Z Gesundh Wiss 2023;31:1193-1207 [FREE Full text] [doi: 10.1007/s10389-021-01684-x] [Medline: 35103232]
14. Kelton DK, Szulewski A, Howes D. Real-time video telemedicine applications in the emergency department: a scoping review of literature. Canadian Journal of Emergency Medicine 2018 Nov 22;20(6):920-928. [doi: 10.1017/cem.2017.382] [Medline: 28829008]
15. Nadar M, Jouvet P, Tucci M, Toledano B, Sicotte C. Impact of synchronous telemedicine models on clinical outcomes in pediatric acute care settings: a systematic review. Pediatric Critical Care Medicine 2018 Dec;19(12):e662-e671. [doi: 10.1097/PCC.0000000000001733] [Medline: 30234678]
16. Winburn AS, Brixey JJ, Langabeer J, Champagne-Langabeer T. A systematic review of prehospital telehealth utilization. Journal of Telemedicine and Telecare 2017 Jun 22;24(7):473-481. [doi: 10.1177/1357633x17713140]
17. Lazarus G, Permana AP, Nugroho SW, Audrey J, Wijaya DN, Widyahening IS. Telestroke strategies to enhance acute stroke management in rural settings: A systematic review and meta-analysis. Brain and Behaviour 2020 Oct 18;10(10):e01787 [FREE Full text] [doi: 10.1002/brb3.1787] [Medline: 32812380]
18. Ward MM, Jaana M, Natafgi N. Systematic review of telemedicine applications in emergency rooms. International Journal of Medical Informatics 2015 Sep;84(9):601-616. [doi: 10.1016/j.ijmedinf.2015.05.009] [Medline: 26072326]
19. Boggan JC, Shoup JP, Whited JD, Van Voorhees E, Gordon AM, Rushton S, et al. Effectiveness of acute care remote triage systems: a systematic review. J Gen Intern Med 2020 Jul 02;35(7):2136-2145 [FREE Full text] [doi: 10.1007/s11606-019-05585-4] [Medline: 31898116]
20. Eder PA, Reime B, Wurmb T, Kippnich U, Shammas L, Rashid A. Prehospital telemedical emergency management of severely injured trauma patients. Methods of Information in Medicine 2018 Nov;57(5-06):231-242. [doi: 10.1055/s-0039-1681089] [Medline: 30875702]
21. Rogers H, Madathil KC, Agnisarman S, Narasimha S, Ashok A, Nair A, et al. A systematic review of the implementation challenges of telemedicine systems in ambulances. Telemedicine Journal and E-Health 2017 Sep;23(9):707-717. [doi: 10.1089/tmj.2016.0248] [Medline: 28294704]
22. Salmoiraghi A, Hussain S. A systematic review of the use of telepsychiatry in acute settings. Journal of Psychiatric Practice 2015 Sep;21(5):389-393. [doi: 10.1097/PRA.0000000000000103] [Medline: 26348806]
23. Kimmel HJ, Brice YN, Trikalinos TA, Sarkar IN, Ranney ML. Real-time emergency department electronic notifications regarding high-risk patients: a systematic review. Telemedicine Journal and E Health 2019 Jul;25(7):604-618 [FREE Full text] [doi: 10.1089/tmj.2018.0117] [Medline: 30129886]
24. Marsh-Feiley G, Eadie L, Wilson P. Telesonography in emergency medicine: A systematic review. PLoS One 2018 May 3;13(5):e0194840 [FREE Full text] [doi: 10.1371/journal.pone.0194840] [Medline: 29723198]
25. Gattu R, Teshome G, Lichenstein R. Telemedicine applications for the pediatric emergency medicine. Pediatric Emergency Care 2016;32(2):123-130. [doi: 10.1097/pec.0000000000000712]
26. Pak D, Pak K. Effectiveness of telemedicine in acute / emergency care settings versus face to face patient care: a systematic literature review. Journal of the International Society for Telemedicine and eHealth 2016;3:e14-1-e14-13 [FREE Full text]
27. Is Prehospital Emergency Telemedicine Implementation Feasible In Non - Traditional EMS Settings: A Systematic Literature Review. The University of Arizona. 2017. URL: https://repository.arizona.edu/handle/10150/623626 [accessed 2025-01-30]
28. Corbin J, Strauss A. Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. Thousand Oaks, CA: SAGE Publications, Inc; 2014.
29. Strauss A, Corbin JM. Grounded Theory in Practice. Thousand Oaks, CA: SAGE Publications, Inc; 1999.
30. Rangachari P, Mushiana SS, Herbert K. A scoping review of applications of the Consolidated Framework for Implementation Research (CFIR) to telehealth service implementation initiatives. BMC Health Serv Res 2022 Nov 30;22(1):1450 [FREE Full text] [doi: 10.1186/s12913-022-08871-w] [Medline: 36447279]

Abbreviations

ACEM: Australasian College of Emergency Medicine
ANZFTCOP: Australian New Zealand FACEM Telemedicine Community of Practice

XSL•FO
RenderX

ED: emergency department
EM: emergency medicine
ETS: Emergency Telehealth Service
FACEM: Fellow of Australasian College of Emergency Medicine
GP: general practitioner
WACHS: WA Country Health Service

Edited by A Kushniruk, D Kaufman; submitted 26.03.24; peer-reviewed by J Job, GA Kingston, N Harvey; comments to author 07.07.24; revised version received 01.12.24; accepted 23.01.25; published 15.04.25.

Please cite as:
Tsou C, Yeung J, Goode M, Mcdonnell J, Williams A, Andrew SC, Tetlow J, Jamieson A, Hendrie D, Reid C, Thompson S
Perceived Risks, Mitigation Strategies, and Modifiability of Telehealth in Rural and Remote Emergency Departments: Qualitative Exploration Study
JMIR Hum Factors 2025;12:e58851
URL: https://humanfactors.jmir.org/2025/1/e58851
doi:10.2196/58851
PMID:40232817

©Christina Tsou, Justin Yeung, Melanie Goode, Josephine Mcdonnell, Aled Williams, Stephen Colin Andrew, Jenny Tetlow, Andrew Jamieson, Delia Hendrie, Christopher Reid, Sandra Thompson. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 15.04.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Experiences of Wheelchair Users With Spinal Cord Injury With Self-Tracking and Commercial Self-Tracking Technology ("In Our World, Calories Are Very Important"): Qualitative Interview Study

Vasiliki Mylonopoulou[1], PhD; Katerina Cerna[2], PhD; Alexandra Weilenmann[1], Prof Dr; Mattias Rost[1], PhD; Tobias Holmlund[3], PhD

[1]Applied IT, University of Gothenburg, Gothenburg, Sweden

[2]School of Information Technology, Halmstad University, Halmstad, Sweden

[3]Neurobiology, Care Sciences and Society, Karolinska Institute, Stockholm, Sweden

Corresponding Author:
Vasiliki Mylonopoulou, PhD
Applied IT
University of Gothenburg
Forskningsgången 6
Gothenburg, 417 56
Sweden
Phone: 46 0702517497
Email: vasiliki.mylonopoulou@gu.se

Abstract

Background: Commercial wearable and mobile wellness apps and devices have become increasingly affordable and ubiquitous. One of their aims is to assist the individual wearing them in adopting a healthier lifestyle through tracking and visualizing their data. Some of these devices and apps have a wheelchair mode that indicates that they are designed for different types of bodies (eg, wheelchair users with spinal cord injury [SCI]). However, research focuses mainly on designing and developing new condition-specific self-tracking technology, whereas the experiences of wheelchair users with SCI using self-tracking technology remain underexplored.

Objective: The objectives of this study were to (1) provide a comprehensive overview of the literature in the field of self-tracking technology and wheelchair users (as a basis for the study), (2) present the self-tracking needs of wheelchair users with SCI, and (3) present their experiences and use of commercial self-tracking technology.

Methods: We conducted semistructured interviews with wheelchair users with SCI to understand their experiences with self-tracking and self-tracking technologies, their self-tracking needs, and how they changed before and after the injury. The interviews were thematically analyzed using an inductive approach.

Results: Our findings comprised three themes: (1) being a wheelchair user with SCI, (2) reasons for self-tracking, and (3) experiences with self-tracking technologies and tools. The last theme comprised 3 subthemes: self-tracking technology use, trust in self-tracking technology, and calorie tracking.

Conclusions: In the Discussion section, we present how our findings relate to the literature and discuss the lack of trust in commercial self-tracking technologies regarding calorie tracking, as well as the role of wheelchair users with SCI in the design of commercial self-tracking technology.

(JMIR Hum Factors 2025;12:e65207) doi:10.2196/65207

KEYWORDS

wheelchair; spinal cord injury; tracking; self-tracking; wellness technology; calories; health inequalities; inclusive design in mobile health; design; lifestyle app; artificial intelligence; AI

XSL•FO
RenderX

Introduction

Background

Every year, up to 500,000 people experience spinal cord injury (SCI) and, as a result, become wheelchair users [1]. This population must adjust to their new body, learn to use a wheelchair, and sustain an active lifestyle. An active lifestyle is vital for their independent living and quality of life. If they are not active, their shoulders can be injured because they are overused in their daily life without being trained for specific activities (eg, getting in and out of the wheelchair). Another risk is that they can gain weight, which requires them to change to a wider wheelchair, thus rendering more places inaccessible.

As commercial self-tracking technology (eg, smartwatches) becomes increasingly common, wheelchair users can access and benefit from its advancements. Some of this technology (eg, Apple Watch and Runkeeper) has already implemented accessibility settings for wheelchair users. Up to now, research on self-tracking technology has focused on (1) chair-ables, artifacts placed on or in the wheelchair to promote a healthier lifestyle [2-4]; and (2) studies on wheelchair users' experiences with self-tracking technology that was provided or introduced to them [5-7]. However, the everyday lived experience of self-tracking and related technology (ie, self-tracking technology) of wheelchair users with SCI is underresearched. This can result in the exclusion and underrepresentation of this specific group in the research and development of commercial wearables. Excluding part of the population as "special users" is an ethical and social issue [8,9]. In addition, more commercial technologies are anticipated to become accessible to a broader audience once the European Accessibility Act (EAA) is applied. This act focuses on digital products and services, underlining the importance of mainstream technology being accessible to people with disabilities and older adults [10].

The objective of this study is to build on previous research on self-tracking and self-tracking technology among wheelchair users with SCI [5-7,11] to include lived experience of self-tracking and self-tracking technology. Through an interview study, we aimed to answer the following question: How do wheelchair users with SCI experience self-tracking and commercial self-tracking technology in their everyday lives?

This section first provides an overview of SCI and the impact of the injury on people's lives. This is followed by a short overview of existing research on wheelchair users and self-tracking technology for health and well-being. Finally, we position our research in the context of contemporary society and technological research focusing on the role of wheelchair users with SCI.

SCI Overview

Worldwide, approximately 90 million people live with SCI, many of whom are either underage or part of the working population [1,12]. Each year, between 250,000 and 500,000 people sustain an injury to their spinal cord [1]. The average lifetime costs of an SCI starting at the age of 25 years range from €0.45 million to €2.1 million (US $0.49 million-$2.3 million), which exceeds those related to dementia, multiple sclerosis, and cerebral palsy [1]. Poor quality of life for people with SCI, their families, and caregivers—in combination with the socioeconomic burdens caused by the injuries—make it imperative to find solutions for this condition.

A motor-complete SCI leads to complete muscle mass loss in the legs and, thus, life as a wheelchair user. For a walking person, the muscles in the legs are the main consumers of energy both in an active state (eg, when the person exercises or moves) and in a rest state (eg, sleeping or sitting) [13,14]. In addition, an injury above thoracic level 6 provokes an altered (physiological) autonomic response during both rest and activity [15]. This results in an altered metabolic response, leading to reduced energy burn in both states [13-15]. Body weight balance is partly explained as caloric balance based on calorie intake and output [16-18]. After an SCI, the total energy burn (ie, calorie burn) is much lower than before. Generally, after the initial acute phase, people with SCI lose weight due to the loss of muscle mass and the healing process. During rehabilitation, most people get used to their light body weight and narrow wheelchair, facilitating an active and independent daily life. Throughout the rehabilitation phase, body weight tends to increase (approximately ≥2 kg during the first year) [19]. This relates to the lack of compensatory caloric intake due to the lower metabolic demands (total daily energy burn), which leads to energy imbalance with higher calorie intake than energy burn [19]. Increased body weight can have many consequences, such as higher stress on shoulder muscles due to extra weight being carried each time the person transfers to and from their wheelchair [20]. Moreover, it becomes more difficult to manage daily activities such as self-care. The resulting shoulder pain and reduced independence are often devastating [20]. Thus, calories are extremely important for people with SCI. Maintaining an appropriate calorie balance becomes much more achievable when energy burn and calorie intake can be monitored.

Wheelchair Users With SCI and Self-Tracking Technology for Health and Well-Being

Overview

Self-tracking is a voluntary, reflexive practice in which people collect data about themselves to enhance self-awareness and understanding by observing, noting, and interpreting various aspects of their lives [21]. Self-tracking technology is any technology that supports people in self-tracking practice. This technology can include apps—where the user registers information about their health and well-being—smartwatches or other wearable technology (wearables), or sensors embedded in assistive technology such as a wheelchair (chair-ables). Self-tracking technology is not assistive technology as it does not aim to improve the functional ability of people with disabilities or enable and enhance their participation and inclusion in different domains of life [22]. However, self-tracking technologies can offer inclusivity through adjusting functionality to include people with disabilities.

In this study, we focused on the experiences of people with SCI with mainstream technology as we aimed to contribute to an inclusive design. However, we recognize the importance of

XSL•FO
RenderX

understanding and learning from assistive technology for our target demographic. This understanding could direct inclusive design research and practice as it provides insights into adjusting or expanding the mainstream technology to be more inclusive.

Assistive Technology

Chair-ables have been researched in terms of tracking specific behaviors and self-care–related injuries. Let us take the case of pressure releases, which are vital for wheelchair users in avoiding pressure ulcers potentially contributing to premature death [23]. A chair-able can support wheelchair users with pressure release exercises. Sensors can be placed on the wheelchair to measure in-seat movement, classify the activity from the sensors into weight shifts, and inform the user with their data as well as sending push notifications upon achieving goals related to pressure releases [2]. In 2021, Ahad et al [4] developed and evaluated the reliability of an algorithm and a chair-able that could support people on wheelchairs in doing their pressure release exercises. The year after, design considerations for supporting electric wheelchair users with SCI on their pressure releases were published [24]. The authors [24] implemented context-aware and unobtrusive reminders delivered by the wheelchair for incomplete or incorrect pressure release exercises. The interaction with the wheelchair in this context [24] was preferable to an interaction with an app. Similarly, but in the context of self-tracking apps, Amann et al [25,26] used participatory design to design, develop, and evaluate an evidence-based app that uses a smart camera to prevent pressure injuries.

As for self-care, Büyüktür et al [3] presented requirements for semiautomated tracking to support people with SCI in their self-care by interviewing patients and health care professionals. Their main findings were that the patients' routines change from the clinic to the home environment and each patient decides to focus on what they value the most (eg, sleeping instead of waking up to change position during their sleep). Each patient and their informal caregivers try different routines in the home environment until they find one through which they can track the behavior and with which they can comply. Thus, Büyüktür et al [3] recommended a highly tailored and flexible system that can follow the patient's routines as they change and give them feedback. In general, a systematic literature review [11] showed that there is research on self-management of SCI in the home environment and clinic as well as the promotion of physical activity and a healthy lifestyle (eg, wellness and fitness). For example, one of the articles reviewed [27] presented that an intervention that tracked the activity of wheelchair users with SCI and motivated them through just-in-time messaging had the potential to improve physical activity for wheelchair users with SCI. Similarly, an interview study [28] explored the barriers that people with SCI face to staying physically active: (1) lack of tailored physical activity forums with wheelchair users with SCI and (2) lack of personalized fitness-tracking technology tailored to wheelchair-based activities.

Mo et al [29] explored the information needs of wheelchair users, and they identified a lack of tailored, accurate, and affordable tracking for wheelchair users both commercially and academically. Even though they argued for inclusive technology, the start of filling this gap could be in the assistive technology literature. Li et al [30] introduced WheelPoser, a system that tracks and identifies wheelchair users' on-the-move activity. They argued that they could do that by using 4 small sensors. Their dataset, code, and models are available for everyone to use, which allows mainstream technology to become more inclusive by using the same system as an add-on to their trackers or only the dataset to train their own systems. Huang et al [31], recognizing the lack of data on wheelchair users for training artificial intelligence algorithms on pose estimation, developed and evaluated a dataset to fill that gap. Their dataset is publicly available as well. Another article published in a disability forum explored the relationship between the successful completion of activity guidelines for people with SCI and the fitness or health status of the person with SCI [32]. This paper referred to activity guidelines [33,34] for people with SCI and an understanding of how mainstream technology can tailor its features (eg, daily goals) to a specific SCI group.

Inclusive Technology

Research underlines the need for inclusive and accurate commercial technology for wheelchair users through studies [35] and literature reviews [36,37]. The participants in the study by Li et al [35] showed interest in tracking their activity, but they did not trust the accuracy of contemporary commercial technology, understanding that those using powered wheelchairs could track some data on their activity that were relevant to them, such as distance. A literature review explored how academic publications in the personal informatics field address the information needs of wheelchair users [36]; it consolidated recommendations to tackle the design challenges that researchers and practitioners may face when attempting to design for including wheelchair users' activity tracking and information needs. In 2019, Moon et al [37] conducted a literature review of studies on the inclusivity of contemporary digital technology. They advocated for inclusive and universal design to become an integral part of commercial digital technology's development process and consolidated methods that can support the process.

Another systematic literature review [11] published in 2023 examined the use of mobile health in supporting people with SCI to maintain or improve their health. It showed that research lacked results regarding the role of commercial self-tracking technology for health and wellness to support people with SCI. The authors [11] attributed this gap in research to the fact that commercial self-tracking devices are inaccurate in their calculations for wheelchair users. However, in 2021, a study [38] evaluated the accuracy of the Apple Watch Series 1, showing that the watch was accurate for high-frequency strokes on a wheelchair (eg, a treadmill) but not for low-frequency strokes (eg, walking). Even though the research was published recently, the results of this research may not correspond to the current version of the Apple Watch (Series 7).

Related to commercial trackers and their use by of wheelchair users with SCI or their preferences, there is research limited to commercial technology given to them for a specific amount of time. For example, Ungurean and Vatavu [7] published a study in 2022 exploring the needs of wheelchair users related to activity trackers regarding different self-tracking technologies

XSL•FO
RenderX

(eg, rings, watches, armbands, and chair-ables). The researchers [7] interviewed participants after letting them watch videos of 2 different self-tracking technologies (eg, ring and armband) to understand their perceptions, preferences, and willingness to use wearables. Some of their participants were already wearable technology users, and the researchers mentioned that this could have supported them in answering the interview questions [7]. However, the study's results focused on the videos that the participants watched.

Malu and Findlater [5] based their research on the work by Carrington et al [6]. Carrington et al [6] tracked variables of wheelchair basketball players and presented the variables to the athletes to understand their needs and potential data use. In addition, Malu and Findlater [5] included wheelchair users who were not necessarily athletes. Their participants were interviewed, participated in design workshops, and tried 2 commercial technologies. Specifically, they used 3 fitness trackers for 30 minutes before they compared them and participated in a design session for designing a fitness tracker. Finally, both studies agreed that the values that people in wheelchairs may want to see are energy burned (calories), distance, GPS data, and pulse. The participants in the study by Malu and Findlater [5] added that calorie intake is equally important to energy burned. Finally, Malu and Findlater [5], unlike Carrington et al [6], showed that commercial self-tracking devices using GPS and comparing mobility in steps between different dates were acceptable and of value to their participants.

In 2019, Helle and Rosenbeck Gøegb [39] conducted a focus group interview study with 7 wheelchair users who were members of a basketball team. The study aimed to explore wheelchair users' experiences with current self-tracking devices and future requirements. Their key results showed that their participants only found the distance and time parameters useful and perceived calorie consumption, pulse, steps, and training intensity as inaccurate [39]. Their study participants commented that "some activity trackers provide reminders that assume that the user can walk" [39]. The authors concluded that accurate activity trackers designed for wheelchair users are needed. It is not obvious from this publication whether the participants were daily users of the technology or whether the technology was given to them for study purposes, and we do not know whether they were professional athletes.

To summarize, research on self-tracking and self-tracking technology and its use by wheelchair users with SCI is primarily focused on either building technology for them or gathering requirements and their opinions on contemporary commercial self-tracking technology, with which they may have limited experience [5-7,11,25,39]. Our research enriches this body of knowledge [5-7,11,25,39] with lived experience of people who track or tracked their activity or use or used commercial self-tracking technology. This study also increases the body of knowledge of everyday lived experiences with self-tracking devices and self-tracking for an often underrepresented community in technological research—wheelchair users with SCI [40].

More Reasons for Inclusion in the Design of Commercial Technology

People with disabilities are often considered users of "special" products designed explicitly for them instead of active participants in the mainstream market [41]. "Special" targeted products would have a smaller share in the market, and they may be more expensive and, therefore, inaccessible for many people with disabilities as people with disabilities—at least in the European Union (EU) [42]—tend to be at a higher risk of financial challenges. According to Eskyte [41], dividing the market into "average" and "vulnerable" users is an obstacle to incorporating accessibility requirements into general consumer product development. In 2025, all digital services and products in the EU market need to be accessible according to the EAA [10], which will urge practitioners to design digital products and services with a diverse audience in mind.

Publications in technological fields support that building things for people with disabilities often means building better things for everyone [9], and a systematic literature review [11] shows that research on technology for people with SCI tends toward user-centered design. Even though user-centered design focuses on the user by collecting information about them, building empathy—often by simulating their health conditions [43]—and potentially involving them in key stages of the design and development process, it does not seem to be enough. Bennett and Rosner [44] argued for a more participatory way of designing; instead of designing for the person who is different from the designers (eg, wheelchair users), we should design *with* them as part of the design team. Their research indicates that, by trying to simulate their life experiences, we do not treat them as equals but as something different from us, different from the norm [44]. While the solution to the limitations of current empathy-building methods seems to be the participation of the specific user group in the design process, participation does not guarantee technology acceptance [26] and may be hard to achieve [40]. For example, in the case of SCI, Kabir et al [40] support that, due to the impact of the condition on the user, for example, speaking or fatigue issues may have been excluded from the research because the collection methods of the designers and researchers are inaccessible. Kabir et al [40] describe that the methods need to be adjusted based on each participant's abilities, something that some publication forums and reviewers may consider a methodological weakness of a study.

This paper does not argue about replacing the involvement of people with SCI in research or user research. However, it gives an idea of how current people with SCI experience commercial self-tracking technology and describes some of their needs. Even though researchers and practitioners will need to involve people with SCI when designing accessible products, this study can act as a starting point to understand some people's experiences with SCI, their self-tracking habits, and their use of self-tracking technology.

Methods

We will first present the research context and then how the data collection and analysis were conducted.

XSL•FO
RenderX

Context of the Study

This research was conducted as part of a long-lasting cooperation among researchers in the human-computer interaction field; a researcher in physiotherapy focusing on wheelchair users with SCI; and a center for SCI in Gothenburg, Sweden. This cooperation aims to explore the use of affordable commercial wearables and build an application that uses the wearables' data to accurately calculate the collected variables specifically for wheelchair users with SCI. The project is based on research exploring the energy outtake for wheelchair users with SCI on specific exercises [21-24]. Within the scope of this project, we conducted interviews to understand *how wheelchair users with SCI experience self-tracking and the use of commercial self-tracking technology in their everyday lives*.

Ethical Considerations

At the start of the cooperation, the planned research was reviewed and evaluated for its ethical integrity following the process of the Department of Applied IT at the University of Gothenburg, Sweden. This qualitative study interviewed people with SCI on their tracking habits and use of commercial self-tracking technology. This type of research does not fall under any of the categories described by the Riksdag (the supreme decision-making body of the Kingdom of Sweden) as research requiring ethical clearance as no sensitive personal data or other personal data were collected or published [45]. The study did not have access to or collect interviewees' personal information prior to, during, or after interviews. The participants were informed about the study through an informed consent form and could ask questions before taking part. The information sheet also included details on how they could withdraw. The data of the participants were handled according to the Swedish implementation of the European General Data Protection Regulation. Participants received no compensation.

Recruitment, Data Collection, and Data Analysis

We conducted 9 semistructured interviews [46,47] with wheelchair users with SCI. Knowing from the literature [40] that we may face challenges in recruiting participants for interviews, we used nonprobability convenience sampling using the network of the SCI center in Gothenburg, Sweden, and personal connections. Eligible for participation were people aged >18 years who had SCI for more than a year and who used a wheelchair to move. If participants did not fulfill all the aforementioned criteria, they were excluded. The aim was to have a deep understanding of the experiences of the people who contacted us rather than having a certain quantity of data.

Due to the geographical distribution of the participants and to accommodate the interviews at any time the participants wanted (even if it was on short notice), we conducted the interviews and recorded them over Zoom video call (Zoom Video Communications). In that way, the participants could also show us how they tracked themselves if needed. The interview guide had three groups of questions: (1) introductory and background questions, aiming to get to know the interviewees and their habits concerning daily life and exercise activity, including mobility; (2) questions about experiences of self-tracking and self-tracking technologies before and after their injury, aiming to understand the relationship of the interviewee with self-tracking and self-tracking technologies before and after their injury; and (3) closure of the interview questions, aiming to allow the interviewees to add anything they wanted us to know and understand their interest in the results of the study. We did not collect any actual measurements of their physical activity, nor did we try to evaluate their physical activity based on any tool as the focus of this study was to understand their experiences with self-tracking and self-tracking technologies.

The average duration of the interviews was 45 (SD 12.9) minutes, with the shortest being 30 minutes long and the longest being 69 minutes long (390 min; 6.5 h of video recordings in total). All participants were Swedish; 89% (8/9) of the interviews were conducted in English, and 11% (1/9) were conducted in Swedish. The participants interviewed in English could communicate clearly and fluently in English. The interviews were transcribed, and the first author analyzed them through inductive thematic analysis [48] following the steps by Braun and Clarke [49]. The first author familiarized themselves with the data and found codes and initial themes (steps 1-3). All authors reviewed, defined, and named the themes (steps 4-5). In our qualitative and explorative study, we focused on providing a rich description of our participants' experiences with self-tracking and self-tracking technology rather than the frequency with which they talked about each theme. The quotes used were modified to remove repetitions and for grammatical correctness. The quotes taken from the Swedish-language interview were translated into English.

Methodological Limitations

To invite the participants, we used the connections mentioned previously. In the invitation, we informed participants that it would be an online interview; thus, participants who had difficulty speaking due to their injury or other reasons may have been excluded [40]. We expected a small sample who would have the time and energy to participate. To counteract this, we invited people with SCI to participate regardless of whether they were current or past users of self-tracking technology as long as they had tracked something related to their injury or physical activity at some point. We focused on analyzing and obtaining a deeper understanding of the experiences of the 9 people rather than on collecting more data.

Results

Overview

A total of 22% (2/9) of the participants were female, reflecting the gender ratio of SCI at 20% female [50]. In total, 22% (2/9) of the participants had never used self-tracking technology, but they had all tracked some variables. Apart from participant P5, no one was currently an athlete. All the participants (9/9, 100%) were well past their first year of living with SCI, characterized as phase 3 or the chronic phase [40]. No participant had any serious issues with breathing or speaking during the interview regardless of their level of injury. More details about each participant's sex, age, years with SCI, and use of self-tracking technology, as well as the code used to refer to them in the following sections, can be found in Table 1.

XSL•FO
RenderX

Table 1. Demographics of the participants, including their use of self-tracking technology.

Participant code	Sex	Age group (y)	Years with SCI[a]	Use of self-tracking technology
P1	Male	50s	35	Yes
P2	Female	40s	31	Yes
P3	Male	30s	5	Yes
P4	Male	≥65	29	No
P5	Male	30s	12	Yes
P6	Male	40s	13	Yes
P7	Female	≥65	40	No
P8	Male	60s	6	Yes
P9	Male	50s	—[b]	Yes

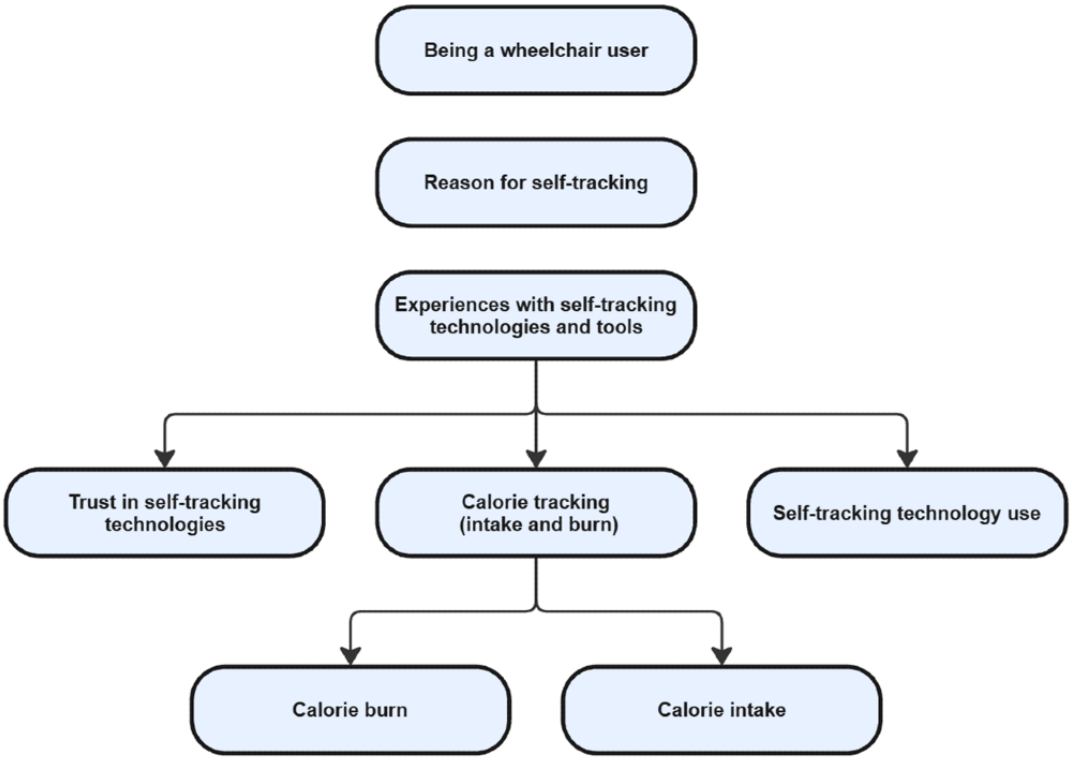
[a]SCI: spinal cord injury.

[b]Not applicable.

In total, 78% (7/9) of the participants had extensively experienced self-tracking technology; however, this high level of engagement with technology should be considered within the context of the study. People living in Sweden use IT daily for mundane things (eg, mobile payment is common; 92% of the population) [51]. In addition, Sweden is among the 7 countries in the EU with the lowest gap in employment between people with and without disabilities, which may have allowed our participants to have enough income [52] to purchase self-tracking technologies.

The thematic analysis resulted in 3 themes, one of which had 3 subthemes. The first theme, *being a wheelchair user with SCI*, included codes relevant to the participants' descriptions of their activity levels, themselves, and their lives. The second theme, *reasons for self-tracking*, included codes related to the participants' self-tracking habits. The third theme, *experiences with self-tracking technologies and tools*, included subthemes related to self-tracking technology use, calorie tracking (intake and burn), and trust in self-tracking technologies. Figure 1 provides a more illustrative understanding of how the themes are structured. The figure is designed to be read from top to bottom and left to right. A tabular form of the figure can be found in Multimedia Appendix 1.

Figure 1. The final themes that resulted from the thematic analysis in a diagrammatic representation. Multimedia Appendix 1 shows the representation of the themes in a tabular form with a summary of the main findings.



XSL•FO
RenderX

Being a Wheelchair User With SCI

This presentation of the participants aims to increase our empathy and understanding of how they experience their lives as wheelchair users and as former walking persons. Wheelchair users with SCI experience life first as walking people and then as wheelchair users. This means that they must adjust to their new body and develop new skills and often see themselves and their life divided into 2 periods—before and after the injury. For example, P8, who had a physical job and changed to a desk job after becoming a wheelchair user, mentioned that "before the injury, I did not track [my activity]. You could see I was physically in a good place. I was very masculine. After the injury, it is another story...you need to train, of course, a lot of rehabilitation." The various SCIs can influence people's bodies and lives differently, widening the diversity of their bodies.

Our participants dealt with the change in their skills in five ways: (1) adjusting their previous knowledge to the new situation, (2) giving up old habits, (3) being challenged to keep old habits, (4) reflecting on the behavior of their body, and (5) acquiring new skills. For example, regarding item 1, P3 adjusted his previous knowledge to the new situation in the following way:

> *I knew what because of my experience before the injury. I knew how to train, [but] I had to adapt to the situation I got into. I could not do everything in the same way, but I got some tips from the rehabilitation clinic, and I tested them myself.*

Others had advanced skills as non–wheelchair users that they lost when they became wheelchair users. For example, regarding item 2, P9 preferred to give up old habits "since I know that I have been playing badminton at a certain [high] level, it is frustrating to find myself being bad at it in the wheelchair. So maybe I would rather play tennis or basketball," whereas others saw their new life as a challenge to keep doing what they used to do. For example, regarding item 3, P1, who used to ski before his injury, felt that sit skiing would not be a challenge. When he tried it and realized how difficult it was for him, he was challenged to keep skiing:

> *I tried [sit ski], then it was so super hard.... So somewhere, it triggered me that I need to practice to get better at it.*

Regarding item 4, P9 reflected that his body did not react the same while running. As now he was not using his legs but his arms to move the wheelchair, it was harder for him to raise his pulse and feel the effects of exercise before his arms got tired. All of them had to learn to use a wheelchair, and the learning experience differed from person to person (item 5). For example, P4, one of the main organizers of a wheelchair community, mentioned that, when he started using a wheelchair, there was not much support but, in his community, they support others with basic wheelchair skills:

> *[W]e [practice] with balancing the wheelchair [on the back wheels] without dropping it back. [It] is very useful and hard skill.... It is very dangerous [as] you can hit your head.*

Everyday activities take more time based on the level of the injury and how it has influenced the body. For example, P7 mentioned that "I have to take care of myself and then work...it takes a lot of time," so she did not have much time to dedicate to exercising. P6, who worked at a rehabilitation center where people stayed for several weeks, said the following:

> *[R]ehabilitation is a lot about...you come here for 4, 6, 8, 10 weeks and you lay a foundation, but to get to be autonomous/independent. You need to continue when you get home because that's where you really need to put in the effort. And I tried to be clear with patients that this is the beginning. When you come home, that's where the struggle or the real-life starts.*

P6 underlined that, when one is in the rehabilitation clinic, they are free from daily responsibilities that may increase when they return to their daily life and take time from the exercise levels they need to sustain. Finally, P4 and P8 pointed out that, in addition to SCI, one develops more chronic conditions due to aging (eg, P8 kept track of his diet due to type 2 diabetes, and P4 tracked his blood pressure and related medication because he had a myocardial infarction).

To summarize, wheelchair users with SCI need to learn to know and adjust to their new bodies—bodies that are diverse by nature combined with the type of injury. When they learn basic wheelchair skills, they must deal with everyday life, which of course takes time for everyone, but for them, it can take time from their exercise. Exercise is a vital part of the high level of activity they need to sustain to be independent and healthy.

Reasons for Self-Tracking

Habitually, inpatients in rehabilitation clinics are told to document their activities or have them tracked during any type of rehabilitation or training related to their injury. However, few of our participants described this as tracking. For example, P9 mentioned the following:

> *If one is in rehab, one should keep track so they can see that the exercise is working...*

P7 remembered tracking during her injury-related training:

> *[Y]ou get a mentor that you can check up to after a couple of weeks, then you must make a new measure and see how you are doing. That was very useful.*

P8 reflected on the lack of technology and feedback during rehabilitation:

> *I think I get the idea [of tracking] from having to fill out all those forms when I was in the rehab center. Every week I have to write on paper, my weight, how much I lost when I went to the toilet, how much I ate on a particular day...and then the doctor concluded...[what I should change]...with the technology we have today just putting that on paper was not doing anything to me...I was just waiting for the outcome from the doctor.*

Our participants had different motivations for tracking. A total of 44% (4/9) of them (P1, P2, P5, and P9) mentioned that they had started tracking as they were practicing a sport. They tracked because their coach had told them to, or their coach did the

XSL•FO
RenderX

tracking for them. However, all of them kept tracking their behavior even after stopping exercising. P2 tracked variables related to swimming (eg, swimming style, time, and meters) as she was training for the Paralympics; nowadays, she tracks the activities she does through her Apple Watch as "it's so simple...the watch does it for me" and also uses it as a motivation to move—"it kind of reminds me to move" and "it keeps track, and it tells me you did a good [job] today." P1 mentioned his reaction the first time he was asked to track variables related to skiing (eg, weather conditions, exercise type, and timing) and how he came to keep tracking his activities:

> *Well, it was the trainer who said, you need to keep track on what you do. Why I've already done it. That was my answer. But then I started to do it. And then I realized that let's say after one season, I looked in the diary [and thought] oh, I need a lot of those things and [those things] didn't pay off. Maybe I should do less of those things since they don't pay off so much and maybe I should increase the things that pay off.*

P9 continued reflecting on his motivation for tracking his activities using his Apple Watch:

> *[I]t is not so much that I want to be in a national team or anything, it is more important that I can stay healthy and be a part of family activities and continue to do that. That is the most important thing to me.*

P3 mentioned keeping track of his training, especially when making changes to his training activities:

> *When you gonna train intensively for a month or something, I think it's interesting to keep track of your training. If you have a goal or race as well, which I don't right now.*

He reflected the following:

> *Yeah, I think most parts of tracking is for motivation for me at least, and then also, it's just interesting to see some data.*

P2 and P3 mentioned that they used tracking to increase motivation, as did P6:

> *[I]t's mostly motivation to see that to give myself a pat on the shoulder.*

Some participants (P1, P5, P6, and P9) tracked their activity before their injury as they were practicing different sports or participating in competitions. On the other hand, P4, P7, and P8 were not interested in tracking or were motivated to track only after their injury. P4 tracked his activity only as part of a game his colleague created—a point system that considered activities one did (eg, walking) and measurements of weight, waist, and chest. The person with the most points won. The point system was adjusted for P4 as he was in a wheelchair. P4 mentioned this game:

> *It was like, a challenge, you know...we challenge ourselves and even checked together "now I'm up to 2000 points" or whatever.*

P7, who had a higher level of injury, mentioned that she was never interested in tracking but that she kept track of her weight and the time it took her to move between places in her house:

> *[I]f you go to the bathroom, I got to make it in 10 minutes, [to] compete with myself to do with it as quickly as possible.... The only thing I track is my weight. My heart rate is no use because I can't get [much variation]*

P8 mentioned that he started feeling the need to track after his injury; before, he could simply see that he was muscular.

Summarizing, our participants performed a lot of tracking during their rehabilitation period to support the health care professionals following them, but the tracking was done on paper. Those training before their injury had a habit of tracking different variables and found value in tracking, but not everyone felt the need to track. In any case, the injury changed the way in which our participants tracked, either from not needing to track to needing to track or by changing the reasons for tracking and the variables they tracked.

Experiences With Self-Tracking Technologies and Tools

In this section, we will present the experiences of our participants with self-tracking technologies and tools through three subthemes: (1) use of self-tracking technologies, (2) tracking calories, and (3) trust in self-tracking technologies.

Self-Tracking Technology Use

The participants had experience using the Apple Watch, Runkeeper, Polar Beat, Pulse Polar Watch, and Nike+. While these technologies automate tracking and keep score of data, the participants also used manual means of tracking data about themselves. They mentioned using Microsoft Excel, a stopwatch, Lifesum, and paper as examples of manual means of self-tracking. One participant (P5) also mentioned that they "track in my mind." However, the most recurrent technologies in the responses were Apple Watch and Runkeeper.

Before their injury, only P6 and P9 used technology (Nike+ Running App) to track their activity. P1 kept a logbook with ski-related variables (he continued after the injury until he bought his Apple Watch), and P5's coach tracked his activity and gave him a schedule or advice accordingly. After their injury, participants P3, P5, P6, and P9 started using the Runkeeper app, which includes a wheelchair user mode. P3 and P5 used it to measure their speed and the distance they ran, and P6 and P9 used it to train for marathons or similar events. However, P9 mentioned that he had given up on Runkeeper because it did not work well with his Apple Watch:

> *Nike, that's actually what I used when [training for the] world run because it's, it's pre-installed in my Apple watch. So, when I got the Apple Watch, the first thing I did was download RunKeeper on my Apple watch. But I noticed that it wasn't attached to it with all its [features]. But there was also already a running up called Nike.*

P1, P2, and P9 were using Apple Watch to track their activities. Apple Watch includes an option for wheelchair users. In the

XSL•FO
RenderX

default mode, the Apple Watch tracks many activities (eg, biking, swimming, and skiing). However, if one sets it to the wheelchair user mode, only 4 activities become available: wheeling in- or outdoors and walking or running speed. P2 preferred to track her activities separately and used the default mode:

> *So most of the time I use the activities that are in the watch. I'm aware that they [calories] are not adapted for me and my muscle mass. That is not the important thing to me. For me, the important thing is that kind of keeps track and tells me that—okay, you did a good job.*

For P2, it was more important to track her activities and obtain positive feedback than to have access to accurate measurements of some of her activities.

The Apple Watch also displays 3 concentric rings: the inner blue ring shows how many times in the day one has moved (ie, changed position), the middle green circle shows how many minutes of brisk activity one has had, and the outer red circle indicates the active calories burned throughout the day. When P2 referred to the positive feedback she received, she meant how full these rings were. P1 and P9—who used the wheelchair mode on their watches—also looked at those rings for a quick daily overview. P9 mentioned the following:

> *[I] look at it but I don't pay too much attention to it.... I can see that I have done my full day's work and that's kind of the analysis I do, not so much else.*

Some of our participants used other apps and devices. For example, P8 had tried the Samsung Health app, but the measurements provided were not relevant to him. P1, one of the Apple Watch users, used a digital food diary (Lifesum) compatible with his watch to track his diet. P5, who was using Runkeeper, also used Polar Beat when training to measure his heart rate; however, he also considered the limitations of the technology. Specifically, he said the following:

> *I don't know if it's correct, but they mean that if you're a tetraplegic [injury level], your heart rate can't go over 120 Mm. And you can push as hard as you can. And you can't get over 120 But on the Polar-Beat that I [wear around the arm], I have been up to 136, I don't trust the one you put around [the chest] like this.*

This section shows that our participants were familiar with self-tracking tools. Many used or tried digital self-tracking technology after their injury, and some had done it even before. They tracked variables that could influence their sports performance or support them in understanding their body (eg, pulse and diet). They also tracked in less quantifiable ways, for example, what activities they did (eg, skiing or swimming) or the completion of goals set on their Apple Watch, which provided motivation. However, all participants mentioned calories and their measurement and consumption. Some expressed the need for a "lifestyle app" supporting them in healthy living as a whole—both in activity and diet tracking.

Tracking the Calories

Overview

When our participants talked about their experiences with self-tracking technology, the discussion always turned to calorie tracking or similar variables (such as weight and diet). Keeping a healthy lifestyle and not gaining weight is an issue for wheelchair users, and the impact of gaining weight on their life quality is tremendous. P9 was the most descriptive on how gaining weight could influence his life quality:

> *I just want one thing; I don't want to have to change my wheelchair into a wider one. If I get fatter, for example, it would be harder to make all these movements that I have to do in and out of the car every day, in and out of my bed every day in and out of the shower every day, or maybe every other day. Also, if because I'm getting fatter, if I have to get a wider wheelchair, I wouldn't get into certain doors, for example, in hotel rooms because my wheels are wider apart...I don't want that to happen and I'm ready to try any mechanical device to help me from not having to do that.*

This quote illustrates the potentially more serious consequences of gaining weight for wheelchair users and the need to track the calorie burn and, if possible, intake. P5 mentioned that "in our world, calories are very important," and P1 commented on weight gain and aging:

> *[W]eight is a problem all the time, especially now when I am getting a bit older, and it is easier to gain weight than losing it.*

Even P7, who seemed uninterested in tracking her activities and did not use any self-tracking technology, mentioned that "tracking it's not so important for me" but "I track weight because it is important for us." When our participants discussed tracking their calories, they mostly talked about calorie burn (P1, P2, P3, P5, P6, P8, and P9). Only 44% (4/9) of the participants (P1, P3, P4, and P5) mentioned calorie intake or diet, and only 11% (1/9; P7) mentioned weight (which can be perceived as a balance between calorie intake and burn).

Calorie Burn

As wheelchair users with SCI have less muscle mass than a walking person, they consume calories differently. P1 mentioned the following:

> *It is really hard for a paraplegic [injury level] to burn calories. It takes so much more time to burn the same amount as you [a full-muscle ability person]. If you're out jogging for one hour, I must wheel maybe for three.*

Therefore, to burn enough calories, they need to spend more time exercising.

Few commercial self-tracking technologies have their features adjusted to wheelchair users. Runkeeper and Apple Watch were 2 technologies mentioned by our participants that consider wheelchair users in some of their features. Our participants did not trust the self-tracking technology they used regarding calories and, instead, used other variables as a proxy for calorie

XSL•FO
RenderX

burn. Consequently, they managed their calories by staying active or using self-tracking technology as a motivation to stay active. P6 used self-tracking technology (FitNotes and Runkeeper) to keep a high level of activity in addition to motivating himself:

> *[One session] wasn't good enough, next week, I'm doing two sessions. And I'm taking the time [for an extra session] by putting something else out of my schedule. Or if, last week, I trained five times. I tell myself I did something good. Continue!*

P5, who used the Runkeeper app, underlined that he checked the distance and speed provided by the app but not the calories despite mentioning that calories were very important to him. He also discarded the calories that his wearable showed (Polar Beat, which measures heartbeat and calories), saying that "I don't care about the calories because I know...the calories are all wrong." P3 also mentioned heartbeat as an indicator for understanding that his training had an impact:

> *[H]eartbeat is always good because if your heartbeat races, you're doing something.*

He also reflected on a noncommercial smartwatch he had tried as part of another study that measured steps, finding steps not to be a good indicator or proxy for understanding his activity levels:

> *I do not know if I am moving a lot, I have more steps on my watch, but I do not know how it works and I do not know what it reduces it and what it does not.*

Similarly, P8 reflected that the visualization of steps in commercial self-tracking technology as an activity proxy was irrelevant to him:

> *Samsung health (smartphone) counts my steps which is not connected with the reality...Like FitBit.*

Apple Watch is one of the self-tracking technologies that has adjusted some of its features for wheelchair users (eg, wheeling in- or outdoors and walking or running pace); however, only P1 trusted and followed their calorie burn. In contrast, P9 pointed out the following:

> *I have my Apple Watch on when I do [strength exercises]. I guess it's not smart...it tracks some of the movements I make as a push in my wheelchair. It was the same with that would that pole pooling...it doesn't track it as something else than kind of push with the wheelchair.*

Thus, the watch counted his calories based on the strokes he did to move his wheelchair, and he argued that, regardless of what exercise he did, the watch would always measure only wheeling in- or outdoors and walking or running pace but not the actual exercise. Therefore, the calories could not be estimated based on a different activity but only on the strokes he did, so he did not pay attention to them but used the watch mainly for motivation. P2, another Apple Watch user, said the following:

> *[A]ll other activities are in there, but they are not adapted for people with no full muscle ability. With that said, I had kind of decided that it doesn't really matter if I spent 300 calories or 500 calories.*

She added that she used the watch to motivate herself to move. Thus, she preferred to use the Apple Watch in the default mode rather than in the wheelchair mode, favoring the ability to track different activities over a more accurate calorie estimation.

Some participants also used nontechnological means to measure their calories manually. For example, P1, who was part of a medical research project on SCI, became aware of the connection between his energy consumption and his pulse and took action based on that:

> *[B]efore being part of the project about energy consumption for a disabled person, I had no way to determine how much energy I consume during the day but after that, I made charts because I have a really strong correlation between pulse and how much energy I consume.*

To summarize, a person with SCI consumes calories at a slower pace than a full-muscle, able-bodied person and needs to be more active to burn the same amount of calories. Most commercial technology currently in the market does not account for this difference, which leads to apps incorporating the wheelchair symbol without adjusting the calorie burn for this type of user or apps calculating the calorie burn correctly but only for limited activities. Thus, our participants estimated their calorie burn based on how many exercise sessions they did or their heart rate increase. Self-tracking technology was used mainly for motivation, and steps were not perceived as a good proxy for activity tracking or calorie burn.

Calorie Intake

Weight is impacted not only by calorie expenditure but also by calorie intake. For example, P4 said that "Exercises are good but if you don't stop eating too much, you will never lose weight." Just like P3 adjusted his exercise knowledge to the new situation, P5 needed to adjust what he knew about diet. Before the injury, P5 had a personal trainer who planned his weekly exercise and diet, for example, to gain muscle. However, P5 mentioned that he found that he could no longer follow what he had learned from his trainer after the injury:

> *When I ended up in a wheelchair, I was used to having the [diet and exercise] paper in the back of my head. But every other person in a wheelchair said you can't eat as much as you did before. So today I'm eating small portions but more often. I don't do breakfast, lunch, dinner.*

P5 had to adapt his weight management after becoming a wheelchair user by avoiding the traditional food distribution throughout the day.

Even though participants were interested in counting their calorie intake, only P1 used a dietary app, called Lifesum. He mentioned that the intake of calories was also important and that it would be interesting to have a daily calorie goal and take note of the type of diet one follows, if applicable. P1 said the following:

XSL•FO
RenderX

> *An app that could measure what I eat, and I can put in parameters like, [amount of] calories [that] are ok for the day, and any kind of diet I follow.*

Another participant, P7, felt that she had control over her diet owing to a course organized for people with SCI that she took after her injury:

> *[In this course] you start tracking everything, they check your weight and height, how you train and exercise, how you feel—everything. And then you get a plan and learn a lot [about] bad calories, and you know exactly how much you are going to eat and how to change...you get a mentor that you can check in with after a couple of weeks, then you [must] make a new measure or everything and see how you are doing. That was very useful.*

Here, instead of using a tracking tool, the user learned how to manage their calorie intake in an educational setting, when episodic measurements were made, and then made sense of them with the support of their mentor.

The participants expressed that calorie intake was as important as calorie burn. They described that they had to adjust their eating habits after the injury to fit their new bodies. Even though only P1 used self-tracking technology to track their diet, other participants mentioned that they had to undergo training or go through an adjustment period to learn how much they could eat by tracking their diet.

Trust in Self-Tracking Technology

P1, P2, P3, P5, P6, and P9 used self-tracking technology to track their activities but did not trust this technology, especially regarding the calories. P1 trusted his Apple Watch because "Apple bought a survey conducted in the US that was about paraplegics, working [out], and how much calories [they burn] with which workload. And they put that algorithm in the watch. For example, let's say a paraplegic person has a 30% slower or lower metabolism compared to able-bodied, that is in consideration in this app." He compared his watch to Runkeeper:

> *I'm out wheeling with 100 pulses; I get depressed because it's like 210 calories. But if you go to RunKeeper, and you have been jogging for one hour, you have 700 calories. So, if you use that app, you will be so fooled and tricked by it. Because it is hard for a paraplegic to burn calories.*

P1 made a comment on commercial apps using the wheelchair symbol deceptively:

> *[Y]ou have apps on the market RunKeeper, or whatever, you download it because it has a wheelchair symbol [thinking] Oh, yeah, this is a cool company, they even bother to [consider] wheelchair [users]. Yeah, but people don't understand. That's a chart with calories from an able-bodied person. They just replace the running person symbol with the wheelchair symbol.*

P2 was dissatisfied that, as a wheelchair user, she could only track 4 activities on the Apple Watch. Therefore, she preferred to use the watch without the wheelchair user setting to keep track of the different activities rather than accurately tracking the calories—as the quote in the previous theme shows. Despite this, P2 stated that they were an "Apple family" and that she trusted Apple because "it's people I know that are interested in technology that buy all the gear and try all the gear out that's Apple Watch is good I'm buying kind of trusted them Yeah. And my husband recommended and said that it would be worth it [now] that I had already started to add more exercise into my life, and I thought that could be a good fun thing to try." In this case, what made the technology more trustworthy was not the actual technological features but the social context. On the other hand, P9 did not trust his Apple Watch to accurately calculate the calories because it tracked everything as strokes on a wheelchair regardless of the type of activity—as his quote in the previous theme showed. However, he mentioned that he trusted his Apple Watch more than an app on his phone as he could not see how an app could track accurately:

> *[T]here was a wheelchair feature that I can put [in the RunKeeper], instead of running or cycling, I put it on the chair. But for example, it didn't track my pushes, because it can't track my pushes, it seems it's only on the phone. So, I think it helps that it's on my Apple watch on my hand to be able to keep track of [the pushes]. So, if I only have the phone in my, pouch that I have on my sitting pillow, or I have it on my arm higher, it doesn't track my movement in the same way.*

P9 expressed his distrust of Runkeeper because it did not provide the user with an explanation of how the app measures the calories of a wheelchair user.

Similarly, P3 and P5 mentioned that they did not trust the technology they used (Runkeeper and Polar Beat) regarding the calories it presented them with. Specifically, P5 mentioned the following:

> *Now it's like, if I go for a walk, it shows how far and the pace and even it [RunKeeper] shows the calories. But I do not think the calories are exact, so I don't care so much about the calories.*

He explained that, apart from calories, "I trust just the RunKeeper because it's like a map, a GPS. You have walked three kilometers. You got the pace for each kilometer." This quote illustrates that, despite P5 distrusting some features of the app (calories), he trusted others that were more easily understandable. He added the following:

> *So for now, I'm pretty happy with the Polar-Beat it is just the calories are not...and it's not reliable.*

Although the user considered himself satisfied with the Polar Beat use, he chose not to rely on the calorie measurement.

Summarizing, trust in self-tracking technology (or the lack of it) is influenced by the perception of the participants regarding how the technology works and how the development company accounts for different bodies related to calorie burn.

XSL•FO
RenderX

Discussion

Principal Findings

This section recaps our contribution followed by a short discussion on trust and transparency related to self-tracking technology. Then follows a section with the implications of commercial self-tracking technology, which concludes with directions for future research.

To address the question of how wheelchair users with SCI experience self-tracking and commercial self-tracking technology in their everyday lives, we interviewed 9 wheelchair users with SCI who had experience with self-tracking or self-tracking technology. The last theme, *experiences with self-tracking technology and tools* in Figure 1 divided into 3 subthemes, shows that current self-tracking technology fails to accommodate wheelchair users' needs concerning calorie management in 2 ways. First, the current commercial technology calorie calculation is based on a body with full muscle ability and, therefore, is inaccurate for people with SCI (subtheme: *calorie tracking*), and second, the users do not trust the calories in the self-tracking technology (subtheme: *trust in self-tracking technology*). Nevertheless, the wheelchair users we studied continue to use self-tracking technology to motivate them to stay active and to monitor variables such as heartbeat, speed, and activity schedule as proxies for their energy consumption (subtheme: *self-tracking technology use*).

This research has limitations due to the small number of interviews and the positioning of the research in a technologically advanced country (Sweden), which also has 1 of the 7 lowest gaps in unemployment between people with and without disabilities in the EU. If the study had been conducted in a different context with, for example, lower technological literacy and use or lower employability of people with disabilities, the results may have differed, particularly on the number of people who have used self-tracking technology. Finally, all of our participants (9/9, 100%) were White. If our participants had had a darker skin tone, the results may have differed, particularly in relation to the *trust in self-tracking technology* theme as self-tracking technology is even less accurate on darker skin [53]. Nevertheless, this study adds to the body of knowledge by validating previous studies on the different variables that wheelchair users want to see. It expands the current body of knowledge by focusing on long-lived experiences of self-tracking and contemporary commercial self-tracking technology, especially on calorie management—the most important factor influencing the life quality of wheelchair users with SCI.

Our research contributes to the scientific community as follows. It extends the work by Malu and Findlater [5] on fitness trackers on wheelchairs by interviewing wheelchair users with SCI who had lived, daily experience with tracking their activity or self-tracking technologies. In line with the work by Malu and Findlater [5], our findings in the *calorie tracking* subtheme indicate that one of the most important values to be measured is calories (intake and burn), followed by distances, GPS data, and pulse, as indicated by the *self-tracking technology use* subtheme. However, our findings in the *calorie burn* subtheme indicate that steps are not a good variable to measure for wheelchair users, which contradicts the study by Malu and Findlater [5] but is in line with the work by Carrington et al [6] and Helle and Rosenbeck Gøegb [39].

Our research, specifically the *self-tracking technology use* subtheme, is in line with the work by Helle Rosenbeck Gøegb[39] regarding the tracking of distance and time. In addition, the *trust in self-tracking technologies* subtheme is partially in line with the results of Helle and Rosenbeck Gøegb [39] showing a lack of trust in technology related to some variables (calorie burn and training intensity). Our participants, similarly to those in the study by Helle and Rosenbeck Gøegb [39], trusted the self-tracking technology regarding their speed measurements and were interested in these measurements; however, neither the participants in the aforementioned study nor our participants trusted calorie burn tracking (see the *trust in self-tracking technology* subtheme). Some of our participants showed a lack of trust in training intensity as well (similarly to those in the study by Helle and Rosenbeck Gøegb [39]); for example, smartwatches counted every activity as a stroke on a chair without distinguishing it from upper-body strength training (see the *trust in self-tracking technology* subtheme). Finally, our findings in the *reasons for self-tracking* theme suggest that being able to track different activities and calorie burn and keep track of exercise schedules, as well as seeing activity goals being reached, was motivational for our participants to keep the active lifestyle they needed, which is in line with the findings of past research [27,28,54].

Trust and Transparency

Participants' trust in self-tracking technologies was heavily influenced by their understanding and perception of how these devices worked. For example, Apple Watch users often trusted the device more than phone apps, believing that the watch's ability to track arm movements made it inherently more accurate than a phone carried in a pocket. However, this trust was conditional and limited; participants expressed confidence in the calorie estimates only for the 4 wheelchair-specific activities provided by the device. Regarding other activities in which calorie expenditure was approximated using these predefined categories, they expressed skepticism about accuracy. This led to varying strategies—some participants attempted to fit all their activities into 1 of the 4 wheelchair-specific categories to achieve more reliable calorie estimates, whereas others discarded calorie tracking entirely and used the watch in nonwheelchair mode to access a broader range of features.

Previous work on trust in self-tracking technology has revealed that users often have a *mental model* of how apps or devices work and calculate data [55]. Trust in self-tracking technologies, particularly in the context of wheelchair users with SCI, hinges significantly on the transparency of how data are collected, processed, and presented. Participants in our study exhibited a nuanced understanding of their devices, often forming mental models to rationalize the data's validity. However, the perceived "black box" nature of calorie calculations—whether on the Apple Watch or Runkeeper—fostered skepticism. This reinforces previous findings that transparency regarding algorithms and data processing is critical for trust [56]. On the

XSL•FO
RenderX

basis of this, we suggest that designers prioritize user education by including detailed explanations of their algorithms, especially when adapting generic metrics such as calorie burn for niche user groups, such as wheelchair users. This aligns with broader calls [57,58] for designing technologies that foster informed trust.

Furthermore, trust in technology is not isolated from social factors [59]. Apple Watch users' trust was partially rooted in social recommendations and previous brand experiences. This suggests that trust may extend beyond technical accuracy to include the brand's perceived reliability and commitment to inclusion. Investigating the interplay between social networks, brand trust, and technology adoption among wheelchair users could provide further insights into how collective validation influences individual trust decisions.

Distrust in calorie tracking was not merely technical but also emotional, linked to fears of inaccuracy impacting critical life decisions (eg, managing weight to maintain mobility). This illustrates that trust is not just a cognitive process but a deeply affective one. Incorporating participatory design with wheelchair users at every stage of the design and development process—testing, iteration, and after launch—can mitigate such emotional concerns. By co-designing solutions, developers can address both the technical and psychological dimensions of trust.

We observed how distrust often led participants to bypass calorie metrics altogether, choosing instead to focus on other variables or manual methods. This self-empowerment reflects resilience but also signals a failure of the technology to deliver on its promise. Therefore, tools must enhance user agency by aligning design goals with real-world constraints and expectations.

Finally, we observed a mismatch between representation and actual design. Runkeeper users appreciated the app's reliance on GPS and speed data, which they found straightforward and comprehensible. However, despite this, they distrusted the app's calorie calculations even when using its wheelchair mode and described feeling misled by its implied inclusivity. The use of wheelchair symbols created an initial perception of inclusivity but, ultimately, resulted in feelings of deception among participants due to inaccurate calorie calculations. This reflects a broader issue in which symbols and icons signaling accessibility and inclusiveness are in fact merely superficial. Misrepresentations such as these can diminish trust but also risk alienating the very communities that the technologies claim to support. Implications such as these are discussed further in the next section.

Implications for Commercial Self-Tracking Technology for Wheelchair Users With SCI

The most important need and valuable variable for our participants (subtheme: calorie tracking), and of the participants of previous research with wheelchair users, was the accuracy of calorie consumption [6,39,60]. Previous research [6,39,60] and our participants indicated that commercial fitness and wellness technology that uses the wheelchair symbol does not necessarily calculate accurately the calories burned by wheelchair users (subtheme: *trust in self-tracking technologies*). Research [6,39,60] also indicates that specific technology for wheelchair users needs to be developed, personalized, and tailored to their needs, which underlines the importance of tailoring mainstream self-tracking technology to the user group that it intends to include [61]. Furthermore, there is extensive research on wheelchair user–specific technology for self-tracking health and well-being [27,28,54], as well as exploring current technology with wheelchair users to inspire designs of wheelchair user–specific fitness and wellness technology [5]. Even though a big part of the aforementioned literature focuses on assistive technology or technology specifically for wheelchair users, mainstream self-tracking technology that aims to be inclusive toward wheelchair users can take advantage of what assistive technology research can offer (eg, databases and artificial intelligence models) [30,31], as well as what the medical field offers, which already has some physical activity guidelines for wheelchair users [14,32,34,62].

On the other hand, the production of new wheelchair-specific self-tracking technology (which is not assistive technology) divides the market into wheelchair users and non–wheelchair users. Recent research [41] indicates that, by excluding specific populations from the development of commercial products, the incorporation of accessibility requirements into general consumer products comes to a halt. In addition, special technology could cost more as it targets a specific population. Regarding wheelchair users, special technology may not be a financially viable solution as they have a higher risk of financial issues [42]. Working with affordable materials to create new technology for a specific population can also lead to them not using it because it could be stigmatizing [63]. From 2025 onward, digital services and products in the EU market should be accessible according to the EAA [10,49]. Our findings show that the self-tracking technology currently in the market is not prepared to be accessible to wheelchair users with SCI. This technology is not trusted when it comes to calorie calculation. Our participants described how they used proxy variables to estimate their calorie consumption and activity level. They expressed the need for a self-tracking technology that accurately calculates their calorie intake and burn during a variety of activities. Many of the issues with self-tracking technology that uses the wheelchair symbol can be solved by involving wheelchair users in the design and development phase. This involvement will make the technology more relevant to the people it aims to include (ie, wheelchair users) and open the market to broader audiences [64-66].

On the basis of the literature presented in this paper and our findings, we see a need for future research and development regarding inclusive self-tracking technology for wheelchair users with SCI supported and inspired by assistive technology literature.

First, there is a need to explore ways to make affordable technology already used by wheelchair users closer to their needs and make it more accurate regarding calorie burn (ie, track a variety of activities and calorie intake). Our study can be seen as a starting point for this exploration of commercial fitness self-tracking technology and its use by wheelchair users with SCI, complementing previous literature on similar subjects [5-7,11].

XSL•FO
RenderX

Second, we argue that there is a need to confront and acknowledge the limitations of the methods used to conduct similar research and alternate them for the research and practice to be more inclusive. According to Kabir et al [40], data collection and design methods may be inaccessible to some populations, leading researchers to exclude them as they perceive them as "difficult to reach and research." Instead, they suggest having a palette of methods and adjusting them based on the participants' abilities [40]. Similarly, Moon et al [37] provide a comprehensive toolbox for more inclusive user research when designing with wheelchair users in mind. In addition, Li et al [30] and Huang et al [31] provide freely available concrete tools (ie, databases and code) for developing an inclusive self-tracking technology to be used by anyone wanting to make self-tracking technology more inclusive.

Third, regardless of all the support that this paper and the literature provide, the inclusion of wheelchair users as an integral part of the design process of the self-tracking technology targeting this demographic should be seen as mandatory for the product to be relevant to the new audience (ie, wheelchair users) [65,66].

Future Research

Our next steps are a series of participatory workshops in which we—together with people with SCI and the support of the SCI center we partnered with—identify further tracking needs and co-design a first version of an app that could support them in learning about their new body in different stages of their postinjury phase. In the workshops, we plan to include the physiotherapist who cooperates with us to include the health care perspective without compromising the power balance. His participation will also create an environment in which the participants can learn more about SCI and how the health care system perceives SCI and them as patients. To create a link for the communication between health care professionals and people with SCI, we will use the connections of our partner at Karolinska Institute to investigate the needs of health care professionals (ie, which of the data tracked by people with SCI are relevant to the health care professionals and in what form to fit their practice when consulting with their patients).

Acknowledgments

The authors thank all the people who dedicated their time to participate in the interviews. Without their contribution and their rich experiences, this study would have been impossible. The authors thank their collaborators from the spinal cord injury center that supported this research and helped them find participants for this study.

Authors' Contributions

VM led the research and the writing process of the paper. VM designed and conducted the interviews and the initial thematic analysis with support from KC. AW and MR participated in the later stages of the thematic analysis and contributed to the overall writing of the paper. TH contributed to the writing of the paper and was responsible for the medical accuracy of the paper regarding spinal cord injury.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The final themes resulting from the thematic analysis in a tabular form.
[PDF File (Adobe PDF File), 28 KB - humanfactors_v12i1e65207_app1.pdf]

References

1. International perspectives on spinal cord injury. World Health Organization. URL: https://www.who.int/publications/i/item/international-perspectives-on-spinal-cord-injury [accessed 2024-04-29]
2. Sprigle S, Sonenblum S, O'Brien JJ. Transitioning a research tool into a consumer product: the wheelchair in-seat activity tracker. In: Proceedings of the 2019 Conference on Design of Medical Devices. 2019 Presented at: DMD '19; April 15-18, 2019; Minneapolis, MN p. 1-5. [doi: 10.1115/DMD2019-3211]
3. Büyüktür AG, Ackerman MS, Newman MW, Hung PY. Design considerations for semi-automated tracking self-care plans in spinal cord injury. In: Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare. 2017 Presented at: PervasiveHealth '17; May 23-26, 2017; Barcelona, Spain p. 183-192 URL: https://dl.acm.org/doi/10.1145/3154862.3154870 [doi: 10.1145/3154862.3154870]
4. Ahad N, Sonenblum SE, Davenport MA, Sprigle S. Validating a wheelchair in-seat activity tracker. Assist Technol 2022 Sep 03;34(5):588-598 [FREE Full text] [doi: 10.1080/10400435.2021.1893237] [Medline: 33617402]
5. Malu M, Findlater L. Toward accessible health and fitness tracking for people with mobility impairments. In: Proceedings of the 10th EAI International Conference on Pervasive Computing Technologies for Healthcare. 2016 Presented at: PervasiveHealth '16; May 16-19, 2016; Cancun, Mexico p. 170-177 URL: https://dl.acm.org/doi/10.5555/3021319.3021344 [doi: 10.4108/eai.16-5-2016.2263329]

XSL•FO
RenderX

6. Carrington P, Chang K, Mentis H, Hurst A. "But, I don't take steps": examining the inaccessibility of fitness trackers for wheelchair athletes. In: Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility. 2015 Presented at: ASSETS '15; October 26-28, 2015; Lisbon, Portugal p. 193-201 URL: https://dl.acm.org/doi/10.1145/2700648.2809845 [doi: 10.1145/2700648.2809845]
7. Ungurean OC, Vatavu RD. “I gave up wearing rings:” insights on the perceptions and preferences of wheelchair users for interactions with wearables. IEEE Pervasive Comput 2022 Jul 1;21(3):92-101. [doi: 10.1109/MPRV.2022.3155952]
8. Sloane M. Inequality is the name of the game: thoughts on the emerging field of technology, ethics and social justice. In: Proceedings of the Weizenbaum Conference 2019 on Challenges of Digital Inequality - Digital Education, Digital Work, Digital Life. 2019 Presented at: EFTESJ '19; May 16-17, 2019; Berlin, Germany p. 1-9 URL: https://tinyurl.com/2b2r7sjj [doi: 10.34669/wi.cp/2.9]
9. Shew A. Ableism, technoableism, and future AI. IEEE Technol Soc Mag 2020 Mar;39(1):40-85. [doi: 10.1109/mts.2020.2967492]
10. European accessibility act. European Parliament. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32019L0882 [accessed 2023-03-10]
11. Bernard RM, Seijas V, Davis M, Volkova A, Diviani N, Lüscher J, et al. Mobile health self-management support for spinal cord injury: systematic literature review. JMIR Mhealth Uhealth 2023 Apr 26;11:e42679 [FREE Full text] [doi: 10.2196/42679] [Medline: 37099372]
12. Spinal locomotor circuits: organization and repair after injury. European Commission. 2011. URL: https://cordis.europa.eu/project/id/201144/es [accessed 2024-04-29]
13. Holmlund T, Ekblom-Bak E, Franzén E, Hultling C, Wikmar LN, Wahman K. Energy expenditure in people with motor-complete paraplegia. Spinal Cord 2017 Aug 4;55(8):774-781. [doi: 10.1038/sc.2017.27] [Medline: 28374809]
14. Holmlund T, Ekblom-Bak E, Franzén E, Hultling C, Wahman K. Energy expenditure after spinal cord injury in people with motor-complete tetraplegia or motor-complete paraplegia. Spinal Cord 2018 Mar 14;56(3):274-283. [doi: 10.1038/s41393-017-0024-4] [Medline: 29238094]
15. Krassioukov A. Autonomic function following cervical spinal cord injury. Respir Physiol Neurobiol 2009 Nov 30;169(2):157-164. [doi: 10.1016/j.resp.2009.08.003] [Medline: 19682607]
16. Piaggi P, Thearle MS, Bogardus C, Krakoff J. Lower energy expenditure predicts long-term increases in weight and fat mass. J Clin Endocrinol Metab 2013 Apr;98(4):E703-E707 [FREE Full text] [doi: 10.1210/jc.2012-3529] [Medline: 23418317]
17. Piaggi P, Vinales KL, Basolo A, Santini F, Krakoff J. Energy expenditure in the etiology of human obesity: spendthrift and thrifty metabolic phenotypes and energy-sensing mechanisms. J Endocrinol Invest 2018 Jan 24;41(1):83-89 [FREE Full text] [doi: 10.1007/s40618-017-0732-9] [Medline: 28741280]
18. Piaggi P. Metabolic determinants of weight gain in humans. Obesity (Silver Spring) 2019 May 23;27(5):691-699 [FREE Full text] [doi: 10.1002/oby.22456] [Medline: 31012296]
19. Chen Y, Henson S, Jackson AB, Richards JS. Obesity intervention in persons with spinal cord injury. Spinal Cord 2006 Feb 16;44(2):82-91. [doi: 10.1038/sj.sc.3101818] [Medline: 16103891]
20. Chen Y, Cao Y, Allen V, Richards JS. Weight matters: physical and psychosocial well being of persons with spinal cord injury in relation to body mass index. Arch Phys Med Rehabil 2011 Mar;92(3):391-398. [doi: 10.1016/j.apmr.2010.06.030] [Medline: 21238950]
21. Lupton D. The diverse domains of quantified selves: self-tracking modes and dataveillance. Econ Soc 2016 Apr 15;45(1):101-122. [doi: 10.1080/03085147.2016.1143726]
22. Global report on assistive technology. World Health Organization. URL: http://apps.who.int/bookorders [accessed 2025-01-27]
23. Redelings MD, Lee NE, Sorvillo F. Pressure ulcers: more lethal than we thought? Adv Skin Wound Care 2005 Sep;18(7):367-372. [doi: 10.1097/00129334-200509000-00010] [Medline: 16160463]
24. Motahar T, Ghosh I, Wiese J. Identifying factors that inhibit self-care behavior among individuals with severe spinal cord injury. In: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. 2022 Presented at: CHI '22; April 29-May 5, 2022; New Orleans, LA p. 1-16 URL: https://dl.acm.org/doi/10.1145/3491102.3517658 [doi: 10.1145/3491102.3517658]
25. Amann J, Fiordelli M, Brach M, Bertschy S, Scheel-Sailer A, Rubinelli S. Co-designing a self-management app prototype to support people with spinal cord injury in the prevention of pressure injuries: mixed methods study. JMIR Mhealth Uhealth 2020 Jul 09;8(7):e18018 [FREE Full text] [doi: 10.2196/18018] [Medline: 32673241]
26. Amann J, Fiordelli M, Scheel-Sailer A, Brach M, Rubinelli S. Opportunities and challenges of a self-management app to support people with spinal cord injury in the prevention of pressure injuries: qualitative study. JMIR Mhealth Uhealth 2020 Dec 09;8(12):e22452 [FREE Full text] [doi: 10.2196/22452] [Medline: 33295876]
27. Hiremath SV, Amiri AM, Thapa-Chhetry B, Snethen G, Schmidt-Read M, Ramos-Lamboy M, et al. Mobile health-based physical activity intervention for individuals with spinal cord injury in the community: a pilot study. PLoS One 2019 Oct 15;14(10):e0223762 [FREE Full text] [doi: 10.1371/journal.pone.0223762] [Medline: 31613909]

XSL•FO
RenderX

28. Lakshminarayanan R, Canori A, Ponnada A, Nunn M, Read MS, Hiremath SV, et al. Exploring opportunities to improve physical activity in individuals with spinal cord injury using context-aware messaging. Proc ACM Hum Comput Interact 2022 Nov 11;6(CSCW2):1-27. [doi: 10.1145/3555628]
29. Mo W, Singh A, Holloway C. Exploring information needs for tracking to support using wheelchairs in everyday life. In: Extended Abstracts of the CHI Conference on Human Factors in Computing Systems. 2024 Presented at: CHI EA '24; May 11-16, 2024; Honolulu, HI p. 1-7 URL: https://dl.acm.org/doi/10.1145/3613905.3650933 [doi: 10.1145/3613905.3650933]
30. Li Y, Mollyn V, Yuan K, Carrington P. WheelPoser: sparse-IMU based body pose estimation for wheelchair users. In: Proceedings of the 26th International ACM SIGACCESS Conference on Computers and Accessibility. 2024 Presented at: ASSETS '24; October 27-30, 2024; St. John's, NL p. 1-17 URL: https://dl.acm.org/doi/10.1145/3663548.3675638 [doi: 10.1145/3663548.3675638]
31. Huang W, Ghahremani S, Pei S, Zhang Y. WheelPose: data synthesis techniques to improve pose estimation performance on wheelchair users. In: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems. 2024 Presented at: CHI '24; May 11-16, 2024; Honolulu, HI p. 1-25 URL: https://dl.acm.org/doi/10.1145/3613904.3642555 [doi: 10.1145/3613904.3642555]
32. Hoevenaars D, Holla JF, Postma K, van der Woude LH, Janssen TW, de Groot S. Associations between meeting exercise guidelines, physical fitness, and health in people with spinal cord injury. Disabil Rehabil 2023 Mar 20;45(6):1030-1037 [FREE Full text] [doi: 10.1080/09638288.2022.2048910] [Medline: 35306932]
33. Tweedy SM, Beckman EM, Geraghty TJ, Theisen D, Perret C, Harvey LA, et al. Exercise and sports science Australia (ESSA) position statement on exercise and spinal cord injury. J Sci Med Sport 2017 Feb;20(2):108-115. [doi: 10.1016/j.jsams.2016.02.001] [Medline: 27185457]
34. Martin Ginis KA, van der Scheer JW, Latimer-Cheung AE, Barrow A, Bourne C, Carruthers P, et al. Evidence-based scientific exercise guidelines for adults with spinal cord injury: an update and a new guideline. Spinal Cord 2018 Apr;56(4):308-321. [doi: 10.1038/s41393-017-0017-3] [Medline: 29070812]
35. Li Y, Li F, Carrington P. Breaking the “inescapable” cycle of pain: supporting wheelchair users’ upper extremity health awareness and management with tracking technologies. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. 2023 Presented at: CHI '23; April 23-28, 2023; Hamburg, Germany p. 1-17 URL: https://dl.acm.org/doi/10.1145/3544548.3580660 [doi: 10.1145/3544548.3580660]
36. Motahar T, Wiese J. A review of personal informatics research for people with motor disabilities. Proc ACM Interact Mob Wearable Ubiquitous Technol 2022 Jul 07;6(2):1-31. [doi: 10.1145/3534614]
37. Moon NW, Baker PM, Goughnour K. Designing wearable technologies for users with disabilities: accessibility, usability, and connectivity factors. J Rehabil Assist Technol Eng 2019 Aug 13;6:2055668319862137 [FREE Full text] [doi: 10.1177/2055668319862137] [Medline: 35186318]
38. Glasheen E, Domingo A, Kressler J. Accuracy of Apple Watch fitness tracker for wheelchair use varies according to movement frequency and task. Ann Phys Rehabil Med 2021 Jan;64(1):101382 [FREE Full text] [doi: 10.1016/j.rehab.2020.03.007] [Medline: 32335302]
39. Helle T, Rosenbeck Gøegb K. Wheelchair users’ experiences with and need of activity trackers. In: Proceedings of the 2019 International Conference on Advancement of Assistive Technology in Europe. 2019 Jul 17 Presented at: ESERA '19; August 26-30, 2019; Bologna, Italy p. 1 URL: https://www.ucviden.dk/en/publications/wheelchair-users-experiences-with-and-need-of-activity-trackers-3 [doi: 10.3233/tad-190001]
40. Kabir KS, Alsaleem A, Wiese J. The impact of spinal cord injury on participation in human-centered research. In: Proceedings of the 2021 ACM Designing Interactive Systems Conference. 2021 Presented at: DIS '21; June 28-July 2, 2021; Virtual Event p. 1902-1914. [doi: 10.1145/3461778.3462122]
41. Eskytė I. Disabled people’s vulnerability in the European single market: the case of consumer information. J Consum Policy 2019 Jul 19;42(4):521-543. [doi: 10.1007/S10603-019-09422-3]
42. Disability statistics - financial situation - statistics explained. Eurostat. URL: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Disability_statistics_-_financial_situation#Struggling_to_make_ends_meet [accessed 2023-01-20]
43. Kouprie M, Visser FS. A framework for empathy in design: stepping into and out of the user's life. J Eng Des 2009 Oct;20(5):437-448. [doi: 10.1080/09544820902875033]
44. Bennett CL, Rosner DK. The promise of empathy: design, disability, and knowing the "other". In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 2019 Presented at: CHI '19; May 4-9, 2019; Glasgow, UK p. 1-13 URL: https://dl.acm.org/doi/10.1145/3290605.3300528 [doi: 10.1145/3290605.3300528]
45. Lag (2003:460) om etikprövning av forskning som avser människor. Sveriges Riksdag. URL: https://www.riksdagen.se/sv/dokument-och-lagar/dokument/svensk-forfattningssamling/lag-2003460-om-etikprovning-av-forskning-som_sfs-2003-460/ [accessed 2024-04-29]
46. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today 2004 Feb;24(2):105-112. [doi: 10.1016/j.nedt.2003.10.001] [Medline: 14769454]
47. Patton MQ, Welch JK. Qualitative evaluation and research methods. Mod Lang J 1992;76(4):543. [doi: 10.2307/330063]
48. Williams M, Moser T. The art of coding and thematic exploration in qualitative research. Int J Manag Rev 2019;15(1):45-55 [FREE Full text]

XSL•FO
RenderX

49. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
50. Raguindin PF, Muka T, Glisic M. Sex and gender gap in spinal cord injury research: focus on cardiometabolic diseases. A mini review. Maturitas 2021 May;147:14-18 [FREE Full text] [doi: 10.1016/j.maturitas.2021.03.004] [Medline: 33832642]
51. E-tjänster. Svenskarna och internet. URL: https://svenskarnaochinternet.se/rapporter/svenskarna-och-internet-2024/e-tjanster/#betaltjanster-och-swish [accessed 2025-02-07]
52. Majority of persons with disabilities locked out of quality employment. European Disability Forum. URL: https://www.edf-feph.org/majority-of-persons-with-disabilities-locked-out-of-quality-employment/ [accessed 2025-02-07]
53. Colvonen PJ, DeYoung PN, Bosompra NO, Owens RL. Limiting racial disparities and bias for wearable devices in health science research. Sleep 2020 Oct 13;43(10):zsaa159 [FREE Full text] [doi: 10.1093/sleep/zsaa159] [Medline: 32893865]
54. Mason L, Dickinson P, Gerling K, De Angeli A. Design goals for playful technology to support physical activity among wheelchair users. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 2019 Presented at: CHI '19; May 4-9, 2019; Glasgow, UK p. 1-12 URL: https://dl.acm.org/doi/10.1145/3290605.3300262 [doi: 10.1145/3290605.3300262]
55. Zhang Y, Trace CB. The quality of health and wellness self‐tracking data: a consumer perspective. J Assoc Inf Sci Technol 2021 Oct 14;73(6):879-891. [doi: 10.1002/ASI.24591]
56. Kizilcec RF. How much information?: effects of transparency on trust in an algorithmic interface. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. 2016 Presented at: CHI '16; May 7-12, 2016; San Jose, CA p. 2390-2395 URL: https://dl.acm.org/doi/10.1145/2858036.2858402 [doi: 10.1145/2858036.2858402]
57. Felzmann H, Fosch-Villaronga E, Lutz C, Tamò-Larrieux A. Towards transparency by design for artificial intelligence. Sci Eng Ethics 2020 Nov 16;26(6):3333-3361. [doi: 10.1007/s11948-020-00276-4]
58. Vorm ES, Combs DJ. Integrating transparency, trust, and acceptance: the intelligent systems technology acceptance model (ISTAM). Int J Hum Comput Interact 2022 May 20;38(18-20):1828-1845. [doi: 10.1080/10447318.2022.2070107]
59. Riegelsberger J, Sasse MA, McCarthy JD. The mechanics of trust: a framework for research and design. Int J Hum Comput Interact 2005 Mar;62(3):381-422. [doi: 10.1016/J.IJHCS.2005.01.001]
60. Malu M, Chundury P, Findlater L. Exploring accessible Smartwatch interactions for people with upper body motor impairments. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI '18; April 21-26, 2018; Montreal, QC p. 1-12 URL: https://dl.acm.org/doi/10.1145/3173574.3174062 [doi: 10.1145/3173574.3174062]
61. Warms C. Physical activity measurement in persons with chronic and disabling conditions: methods, strategies, and issues. Fam Community Health 2006;29(1 Suppl):78S-88S [FREE Full text] [doi: 10.1097/00003727-200601001-00012] [Medline: 16344640]
62. Holmlund T, Ekblom-Bak E, Franzén E, Hultling C, Wahman K. Intensity of physical activity as a percentage of peak oxygen uptake, heart rate and Borg RPE in motor-complete para- and tetraplegia. PLoS One 2019 Dec 3;14(12):e0222542 [FREE Full text] [doi: 10.1371/journal.pone.0222542] [Medline: 31794548]
63. Carrington P, Hurst A, Kane SK. Wearables and chairables: inclusive design of mobile input and output techniques for power wheelchair users. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2014 Presented at: CHI '14; April 26-May 1, 2014; Toronto, ON p. 3103-3112 URL: https://dl.acm.org/doi/10.1145/2556288.2557237 [doi: 10.1145/2556288.2557237]
64. Waller S, Bradley M, Hosking I, Clarkson PJ. Making the case for inclusive design. Appl Ergon 2015 Jan;46 Pt B:297-303. [doi: 10.1016/j.apergo.2013.03.012] [Medline: 23538129]
65. Bødker S, Kyng M. Participatory design that matters—facing the big issues. ACM Trans Comput Hum Interact 2018 Feb 13;25(1):1-31. [doi: 10.1145/3152421]
66. Wilkinson CR, De Angeli A. Applying user centred and participatory design approaches to commercial product development. Des Stud 2014 Nov;35(6):614-631. [doi: 10.1016/j.destud.2014.06.001]

Abbreviations

EAA: European Accessibility Act
EU: European Union
SCI: spinal cord injury

XSL•FO
RenderX

Edited by M Waite; submitted 12.08.24; peer-reviewed by A Panahi, A Castan; comments to author 22.10.24; revised version received 07.02.25; accepted 07.03.25; published 15.04.25.

Please cite as:
Mylonopoulou V, Cerna K, Weilenmann A, Rost M, Holmlund T
Experiences of Wheelchair Users With Spinal Cord Injury With Self-Tracking and Commercial Self-Tracking Technology ("In Our World, Calories Are Very Important"): Qualitative Interview Study
JMIR Hum Factors 2025;12:e65207
URL: https://humanfactors.jmir.org/2025/1/e65207
doi:10.2196/65207
PMID:

©Vasiliki Mylonopoulou, Katerina Cerna, Alexandra Weilenmann, Mattias Rost, Tobias Holmlund. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 15.04.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Blended Mobile-Based Interventions With Integrated Virtual Reality Exposure Therapy for Anxiety Disorders: Thematic Analysis of Patient Perspectives

Jari Planert[1*], MSc; Anne-Sophie Hildebrand[1*], MSc; Alla Machulska[1], PhD; Kati Roesmann[2], Prof Dr; Marie Neubert[1], PhD; Sebastian Pilgramm[3], Prof Dr; Juliane Pilgramm[1], PhD; Tim Klucken[1], Prof Dr

[1]Department of Psychology, Clinical Psychology and Psychotherapy, University of Siegen, Siegen, Germany

[2]Institute of Psychology, Clinical Psychology and Psychotherapy in Childhood and Adolescence, University of Osnabrück, Osnabrück, Germany

[3]Department of Psychology, IU International University of Applied Sciences, Erfurt, Germany

[*]these authors contributed equally

Corresponding Author:
Jari Planert, MSc
Department of Psychology
Clinical Psychology and Psychotherapy
University of Siegen
Obergraben 21
Siegen, 57072
Germany
Phone: 49 271 740 ext 3887
Email: jari.planert@uni-siegen.de

Abstract

Background: Guided mobile-based interventions may mitigate symptoms of anxiety disorders such as panic disorder, agoraphobia, or social anxiety disorder. With exposure therapy being efficacious in traditional treatments for these disorders, recent advancements have introduced 360° videos to deliver virtual reality exposure therapy (VRET) within mobile-based interventions.

Objective: Despite ongoing trials evaluating the treatment's efficacy, research examining patient perceptions of this innovative approach is still scarce. Therefore, this study aimed to explore patient opinions on specific treatment aspects of mobile-based interventions using mobile VRET and psychotherapeutic guidance for anxiety disorders.

Methods: A total of 11 patients diagnosed with panic disorder, agoraphobia, or social anxiety disorder who had previously taken part in the experimental conditions of 2 randomized controlled trials for a mobile intervention including mobile VRET participated in cross-sectional, retrospective interviews. Using a semistructured interview format, patients were asked to reflect on their treatment experiences; personal changes; helpful and hindering aspects; their motivation levels; and their encounters with the mobile-based intervention, manualized treatment sessions, and the mobile VRET.

Results: Thematic analysis led to the formation of 14 themes in four superordinate categories: (1) perceived treatment outcomes, (2) aspects of the mobile intervention, (3) experiences with mobile VRET, and (4) contextual considerations. Patients offered their insights into factors contributing to treatment success or failure, delineated perceived treatment outcomes, and highlighted favorable aspects of the treatment while pointing out shortcomings and suggesting potential enhancements. Most strikingly, while using a blended app-based intervention, patients highlighted the role of psychotherapeutic guidance as a central contributing factor to their symptom improvement.

Conclusions: The findings of the thematic analysis and its diverse patient perspectives hold the potential to guide future research to improve mobile-based treatment options for anxiety disorders. Insights from these patient experiences can contribute to refining mobile-based interventions and optimizing the integration of VRET in accordance with patients' preferences, needs, and expectations.

(JMIR Hum Factors 2025;12:e60957) doi:10.2196/60957

KEYWORDS

virtual reality exposure therapy; anxiety disorders; internet- and mobile-based interventions; blended therapy; eHealth

XSL•FO
RenderX

Introduction

Background

Anxiety disorders such as social anxiety disorder (SAD), panic disorder (PD), or agoraphobia (AG) are a pervasive public health issue with 12-month prevalence rates between 0.9% and 1.2%, causing severe impairment for those affected [1-3]. SAD, PD, and AG are characterized by intense fear reactions toward specific situations or bodily sensations that are normally perceived as harmless or nonthreatening [4]. In SAD, typical anxiety-provoking situations include social interactions (eg, being at the center of attention and meeting new people) as the fear is linked to negative evaluation, criticism, or embarrassment and rejection [5]. In PD, panic attacks are triggered by an overinterpretation of occasional bodily arousal (eg, elevated heartbeat) as life-threatening (eg, having a myocardial infarction) [6]. The catastrophic misinterpretation of such sensations can then lead individuals to be afraid of losing control or experience fear of death. In contrast, AG entails paniclike fear in situations in which escape is difficult (shopping queues, the subway, or car rides) as patients often describe their main concern revolving around losing control, having a medical emergency, being trapped, or not receiving help [7]. Affected individuals often describe anticipatory anxiety as especially troublesome, which may explain why all 3 disorders are often associated with avoidance and safety behavior [8,9].

Cognitive behavioral therapy (CBT) has been found highly effective in treating PD and AG. [10]. Typical methods used in CBT for these disorders are psychoeducation, behavioral experiments, and exposure therapy [11]. However, despite compelling evidence supporting the treatment's efficacy [12], most affected individuals fail to receive treatment (ie, according to Stein et al [13], less than a quarter of patients are admitted for treatment). One reason that keeps individuals in need from receiving treatment is the increasingly exhausted supply capabilities in mental health care, leading to prolonged waiting periods and large pretreatment attrition rates [14]. This is further aggravated by the fact that treatment recommendations such as exposure therapy often remain unfollowed due to time constraints, high costs, and resource demands [15,16].

Exposure therapy, one of the most effective treatments for anxiety disorders, involves the systematic confrontation with feared stimuli or situations to reduce avoidance behavior and facilitate fear extinction. There are several underlying mechanisms that may explain the effect of exposure therapy. First, early theories emphasize that repeated confrontation with anxiety-inducing stimuli promotes habituation [17], leading to a reduction in fear responses and supporting a learning process whereby patients experience that anxiety can diminish without avoidance. Second, confrontation also addresses general therapeutic factors such as problem activation and problem-solving, which have been repeatedly shown to be key mechanisms in psychotherapy [18]. Current theories suggest that the central mechanism in exposure therapy is inhibitory learning, which is induced by expectancy violations. In this framework, to optimize exposure success, it is key to challenge patients' predictions about feared outcomes [19,20]. Despite the documented efficacy and repeated validation of exposure therapy, it remains underused in clinical practice. A contributing factor is the underdiagnosis of anxiety disorders, which reduces access to appropriate treatment [21,22]. However, even when correctly diagnosed, the implementation of exposure therapy is not guaranteed. Reasons for this vary, including logistical challenges, a lack of resources, concerns about patient distress, or insufficient training [16]. In addition, patients might be reluctant to confront feared stimuli or discontinue treatment, which poses additional considerable barriers.

A recent approach to bridge the shortage of treatment capabilities in mental health care is using internet- and mobile-based interventions (IMIs) [23], which include smartphone apps that are tailored for specific psychological disorders such as SAD, PD, and AG. Most of these apps incorporate basic CBT interventions and have been shown to be effective in reducing symptoms, for example, through (a combination of) psychoeducation, cognitive restructuring, mindfulness, and relaxation techniques [24-26]. IMIs can follow a stand-alone approach or may also include varying degrees of psychotherapeutic guidance. Regarding the latter, guided treatment programs have been shown to be superior to stand-alone approaches in reducing clinical symptoms, as indicated by larger acceptability, engagement, and adherence [23,26]. These factors are crucial as they have been linked to improved treatment outcomes and sustained recovery [26]. Possible advantages of IMIs include flexible use at almost any time and place and relieving the psychotherapists' time burden [27]. While the use of IMIs could constitute an option to bridge scarce treatment capabilities, little is known about the potential of IMIs with exposure components to overcome these boundaries and provide less resource-intensive and more accessible exposure therapy as the first-line treatment option for anxiety disorders [28]. Therefore, the question arises of how exposure therapy can be integrated within IMIs.

One way to reduce barriers in the delivery of exposure therapy is the transfer of anxiety-eliciting situations to a virtual reality (VR) environment. VR exposure therapy (VRET) allows for the creation of specific scenarios that may be challenging or impractical to replicate in real life (in vivo) and that are tailored to elicit fear responses [29-31]. These simulations can be adjusted according to each patient's unique needs (ie, massive or gradual approach) and specific concerns while therapists maintain full control during such sessions [32,33]. For instance, therapists can pause, repeat, or pick scenarios based on the patients' progress, feedback, or observed reactions, whereas no unforeseeable events can appear as in in vivo therapy. Even if a certain level of risk is necessary for successful exposure therapy, which cannot be guaranteed by the artificial nature of VRET, the so-called immersion (ie, the feeling of realism created in VR atmospheres) helps uphold the patients' apprehension of imminent danger. This, in turn, can activate central fears and allow for their correction. Moreover, aversive risks such as physical harm can mostly be ruled out in VRET, rendering it to a certain degree safer than in vivo exposure therapy. So far, VRET has been used in the treatment of multiple anxiety disorders, such as SAD or AG, to reduce symptoms and improve quality of life successfully [34,35]. For instance,

XSL•FO
RenderX

patients with SAD have experienced significant reductions in fear of public speaking or social interactions [36-39]. Patients with AG have had less anxiety and shown less avoidance when traveling or enduring crowded places [40]. The quantity and duration of VRET sessions varies in the literature, typically lasting between 30 and 60 minutes and being conducted 8 to 12 times. Compared to exposure in vivo, increased acceptance and lower refusal rates have been reported in the past, making VRET a low-cost treatment option with a high success rate [30,41].

VRET has been used in mainly 2 ways in the existing literature. The first involves a computer-based system, allowing patients to engage with avatars, affect their surroundings, or receive immediate feedback during exposure [42]. Advantages of this approach include a high degree of immersion, the ability to tailor scenarios to individual needs, and the option to simulate dynamic (eg, social) interactions. The second approach uses 360° video systems, which present prerecorded scenarios without the possibility of affecting these through interaction. Notable advantages of this method are low costs, effortless scalability, and integrability with other mental health approaches such as IMIs as it requires considerably less technical equipment than interactive VR systems [43]. However, theoretical models of SAD [5] suggest that approaches with higher degrees of interaction might be more beneficial as fear is linked to interpersonal feedback or interaction [44]. Interestingly, both approaches have been found efficacious in reducing symptoms of SAD, PD, and AG [45]. However, research that directly compares interactive VRET systems and VRET via 360° video systems for the treatment of the aforementioned disorders is still scarce.

Taken together, the use of an IMI might reduce waiting times, whereas VRET might help implement exposure without high costs and resources [30]. To benefit from and provide additional advantages from the combined use, new approaches most recently have tried to integrate mobile VRET into IMIs [46-48]. Instead of using expensive VR equipment, patients use a head-mounted construction in which they can insert their smartphone as a screen and display 360° videos. By doing so, it has been demonstrated that treating acrophobia via self-guided VRET is possible [46,49]. Currently, randomized controlled trials (RCTs) are ongoing to investigate the efficacy of improving symptoms of SAD, PD, and AG using this innovative treatment approach [47,48,50].

To date, compared to a robust body of quantitative research, there are considerably less qualitative studies that have examined VRET. Research investigating the patient perspective has found that immersion and engagement are critical factors for its success, with patients highly valuing the ability to confront fears in a controlled and virtual environment [51,52]. However, perceived challenges are limited interactivity in prerecorded scenarios and technical difficulties, which may decrease the immersive experience and, therefore, the perceived effectiveness of the intervention [51,52]. Regarding the use of IMIs, qualitative research on patients' opinions has highlighted high accessibility and flexibility as major strengths while raising concerns about the lack of relational and emotional support inherent to the oftentimes self-guided nature of IMIs [53,54]. Although the cited literature provides valuable insights into the perspective of patients using VRET and IMIs separately, qualitative research regarding the integration of mobile VRET within guided IMIs is scarce. This research gap is addressed in this study.

Objectives

This study aimed to investigate patients' perspectives on integrating mobile VRET into guided IMIs for SAD, PD, and AG. Specifically, the goal was to identify key aspects that patients perceive as relevant to improvement or deterioration, including factors associated with the treatment's limitations, benefits, and outcomes [55,56]. This study also attempted to explore how patients navigate through the different treatment components, what their subjective reasons were for experiencing change or not, and their views on the combination of app-based treatment with psychotherapeutic guidance. On this basis, ideas can be formed on how to improve further similar approaches by facilitating patient adherence.

Methods

Participants

Patients were recruited from 2 RCTs that used a blended treatment approach comprising an IMI with in-app mobile VRET and appointments with a licensed psychotherapist. The work by Planert et al [47] and Hildebrand et al [48] provides further details concerning the trials.

Inclusion criteria were (1) fulfilling the *International Classification of Diseases, 10th Revision*, diagnostic criteria for one of the following psychological disorders: *agoraphobia, unspecified* (F40.00), *agoraphobia with panic disorder* (F40.01), *agoraphobia without panic disorder* (F40.02), *social phobia* (F40.1), or *panic disorder* (F41.0); (2) being aged ≥18 years; and (3) having completed the treatment in one of the RCTs [47,48]. Exclusion criteria for participation in the trials included a history of cardiovascular conditions (eg, stroke or myocardial infarction), respiratory disorders (eg, asthma or chronic obstructive pulmonary disease), epilepsy, pregnancy, impaired vision, schizophrenia, severe depression, acute suicidality, or substance use disorders. None of the patients received outpatient psychotherapy during the treatment period apart from the therapy provided as part of the study. Only participants from the treatment group in either of the 2 trials were invited upon completion of the RCT. A total of 11 patients (n=5, 45% patients with SAD; n=6, 55% patients with PD or AG) agreed to partake in the interviews.

Mobile-Based Intervention

The IMI in this blended treatment was an app for mobile phones incorporating 3 distinct treatment programs specifically developed for SAD, PD, and AG (Invirto [57]), of which patients completed 1. The app is based on a CBT treatment and consists of structured modules targeting different treatment aspects: goal setting, motivation, psychoeducation, tension regulation, cognitive techniques, exposure therapy, behavioral experiments, emotion regulation, exposure, reflection, and relapse prevention. For example, tension regulation was provided via breathing exercises, incorporating mindfulness-based techniques to help

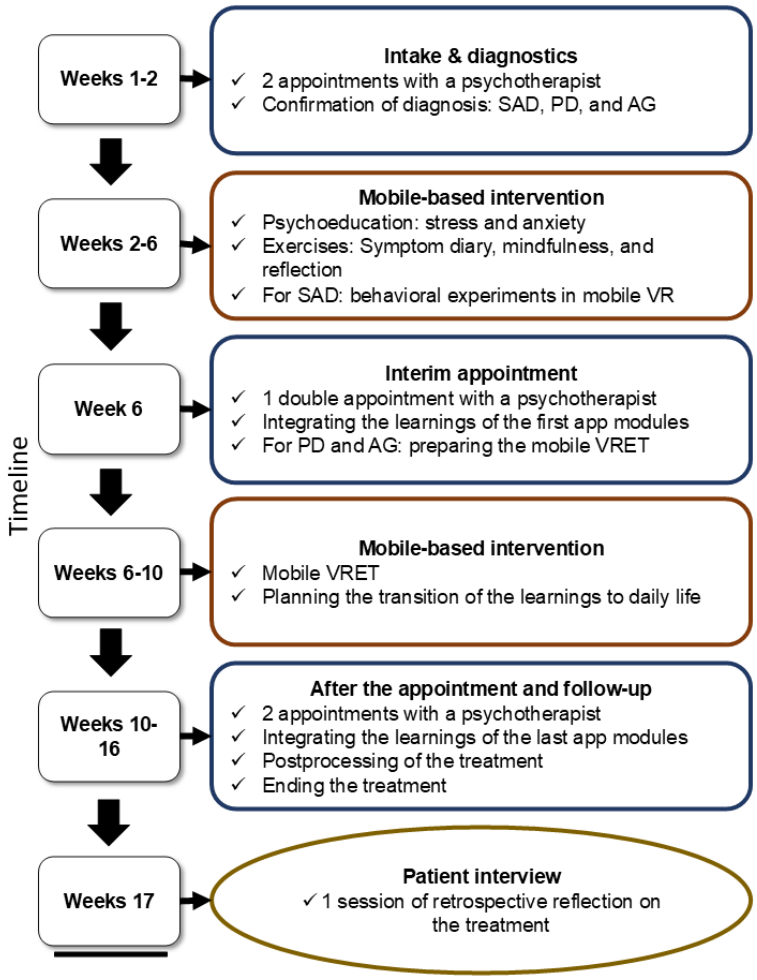
XSL•FO
RenderX

patients manage physiological arousal. Unlike classic behavioral therapy methods, which rely heavily on face-to-face interactions, the app enabled patients to work autonomously through the modules. While the psychotherapists provided some guidance during scheduled appointments, the intervention allowed patients to progress within the structured program at their own pace. Modules incorporated a mix of prerecorded video sequences of psychotherapists, interactive experiments, self-assessments, exercises, and a wide range of potentially fear-evoking situations in a controlled VR environment. Using 360° videos, patients were able to choose scenarios such as giving a public speech, undergoing a job interview, or being in a crowded space or public transport that were thought to provoke or aimed at provoking symptoms of their specific anxiety disorder.

Procedure

Throughout the blended treatment, 5 psychotherapeutic appointments were conducted (for details, see the work by Planert et al [47] and Hildebrand et al [48]). As the mobile-based app offers 3 different treatment programs depending on the anxiety disorder, a joint recruitment process took place. Figure 1 provides an overview of the treatment procedure and study schedule.

Figure 1. Procedure of the treatment process and study schedule. On the left, the boxes indicate the corresponding week of the trial. Blue boxes indicate psychotherapeutic sessions, whereas the red boxes indicate periods in which the patients worked independently with the treatment app. The yellow box marks the time point at which the patient interviews took place. AG: agoraphobia; PD: panic disorder; SAD: social anxiety disorder; VR: virtual reality; VRET: virtual reality exposure therapy.



At study entry, patients received 2 diagnostic appointments in which they were assessed for eligibility. During these appointments, the diagnosis was checked, and an anamnesis took place. Those diagnosed with PD, AG, or SAD received access to the app (ie, 1 of the 3 treatment programs), which was available for 90 days after prescription. They were instructed to complete the first modules (self-guided phase I), including psychoeducation on the corresponding disorder and exercises (eg, identifying and visualizing triggers of fear and exercising diaphragmatic breathing for PD to reduce hyperventilation and promote relaxation [58]). To check on the completion of the self-guided phase I, patients were called to check up on their progress and whether they were ready for the next appointment. The self-guided phase I was set to 6 weeks; however, progressing to the next appointment was only possible after completing all modules from phase I.

After completing the self-guided phase I, there was a third appointment to evaluate what patients had learned in the last modules and prepare for the next phase. As the treatment programs slightly differed with respect to the specific diagnosis, the content of the third appointment differed as well. For patients with SAD, the third appointment included evaluating the previous psychoeducation, exercises, and the behavioral

XSL•FO
RenderX

experiments in mobile VR. In addition, the patients were prepared for the transfer of those behavioral experiments into everyday life. For patients with PD or AG, the evaluation included the previous psychoeducation and exercises without VR. The patients were prepared for behavioral exercises (eg, running upstairs) and exposure to mobile VR (eg, sitting in a car or train) via 360° videos without the possibility of interacting (eg, by provoking reactions from objects or persons nearby). Patients received additional guidance to apply previously learned breathing techniques and integrate these during exposure tasks, ensuring that patients were able to handle physiological symptoms of panic that would arise. After the third appointment, all patients had 4 weeks to complete the remaining modules (self-guided phase II). This included exposure in vivo, emotion regulation, and relapse prevention in all treatment programs. For patients with PD and AG, this also included exposure in VR in multiple scenarios (eg, a car or plane ride or a supermarket). Afterward, patients and psychotherapists evaluated the complete intervention in a fourth appointment. After 6 weeks, patients had a follow-up appointment to assess long-term changes in symptoms and were asked to participate in a half-standardized interview. The obtained data were retrospective insights from patients who completed the whole treatment and were able to extensively use the app in their daily routine.

Half-Standardized Interview Procedure

All interviews were held in German and took, on average, 30 minutes and 40 seconds (SD 6 min). The interviews were conducted by author JP, a clinical psychologist. After describing the interview's purpose and how data protection was ensured, each interview began with an open-ended question, providing interviewees with an initial opportunity to freely discuss the aspects of the treatment they found most notable. This was done by asking the following: "How did you experience the treatment you were taking part in?" Afterward, the interviewer followed a semistructured manual based on the Client Change Interview [59]. The Client Change Interview establishes aspects that are potentially associated with a client's change in psychotherapy, such as helpful and hindering aspects, personal or psychological changes they made, or attributions for these changes. In this study, the interview manual was expanded through specific questions about the app included in the study, along with the mobile VRET, the study design of the RCT, and the expected role of the specific treatment in mental health care. The interview was semistructured, and for every point that the interviewee brought up, the interviewer was instructed to ask follow-up questions to ensure an in-depth understanding. All the interviews were semantically transcribed by student assistants. The interview procedure was conducted under the supervision of researchers AM and TK to ensure consistency and adherence to ethical guidelines.

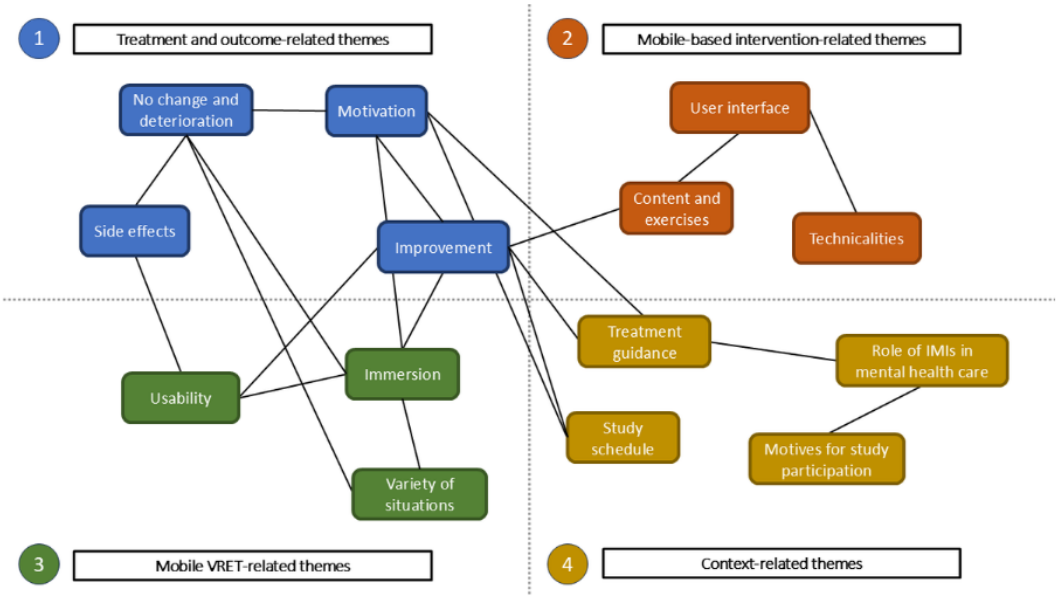
Data Analysis

To analyze the interview data, a thematic content analysis was conducted according to Braun and Clarke [55,56]. The analysis was carried out by authors JP (first coder) and ASH (second coder), both clinical psychologists, who independently from one another reviewed the data. An inductive approach was chosen to generate codes directly from the transcripts of the interviews.

The qualitative data analysis was conducted using MAXQDA 2022 (version 22.7.0; VERBI GmbH) [60]. As a first step, both coders intuitively went over a part of the interviews in a summarizing manner. This was done to establish codes representing the content of the interviews. At this point, codes were not structured. Afterward, the established codes were reviewed by both coders collaboratively to ensure accuracy and check on errors or parts and messages that the other coder might have missed. The generated codes were then interpreted and grouped into overarching categories (eg, VR-related content and IMI-related content) by 1 coder. The second coder then checked the categories and the assigned codes. Afterward, the codes in 1 overarching category were grouped into multiple themes by the second coder and checked by the first coder. Both coders then reflected on and interpreted the categories and themes with their corresponding codes and depicted them in a thematic map (Figure 2) to visually illustrate the findings.

XSL•FO
RenderX

Figure 2. Thematic map of the established themes within their respective categories. Themes are divided by category. Links between the themes indicate thematic relatedness depending on the statements of the interviewed patients. IMI: internet- and mobile-based intervention; VRET: virtual reality exposure therapy.



Ethical Considerations

The interview procedure after the follow-up appointment was approved by the Research Ethics Committee of the University of Siegen through an amendment to the corresponding ethics statement (ER_48_2021, amendment AZ_56/2023) All participants provided informed consent before participation, which could be withdrawn at any time without giving any reason. All collected data were pseudonymized to protect participants' confidentiality. No monetary compensation was issued.

Results

Overview

As the interviews were conducted in German, extracted quotes were translated from German into English literally. The extracted quotes appear indented and in italics in the sections corresponding to each theme. Table 1 shows the demographic characteristics of the 11 interviewees.

Table 1. Demographic information of the study sample.

Patient code	Age range (y[a])	Gender	Primary diagnosis (*ICD-10*[b])	Symptom change[c]
P01	20-39	Woman	Social anxiety disorder	–8.5%
P02	20-39	Woman	Panic disorder	–50%
P03	20-39	Woman	Social anxiety disorder	–28.4%
P04	>40	Woman	Agoraphobia	–92.3%
P05	>40	Woman	Social anxiety disorder	–20.2%
P06	>40	Woman	Panic disorder	–73.2%
P07	>40	Woman	Panic disorder	–27.6%
P08	20-39	Woman	Agoraphobia	–70.1%
P09	>40	Man	Agoraphobia with panic disorder	–19.4%
P10	20-39	Woman	Social anxiety disorder	+10%
P11	20-39	Man	Social anxiety disorder	–3.8%

[a]Information on age is given as ranges to protect the anonymity of the interviewed patients.

[b] *ICD-10: International Classification of Diseases, 10th Revision.*

[c]Symptom change indications refer to either of 2 disorder-specific questionnaires: the Questionnaire for Social Anxiety and Social Competence Deficits or the Panic and Agoraphobia Scale.

The themes of the various interview segments were divided into 4 overarching categories: "treatment and outcome–related themes," "mobile-based intervention–related themes," "VRET-related themes," and "context-related themes" (Figure 2).

XSL•FO
RenderX

Category 1: Treatment and Outcome–Related Themes

The first category covered themes that were related to patient-perceived aspects of the blended treatment or its outcomes. In total, this category consisted of 4 themes: symptom improvement, no change and symptom deterioration, side effects, and motivation throughout treatment.

Theme 1: Symptom Improvement

Throughout the treatment, many of the interviewed patients experienced symptom improvement. Indications as to how and why their symptoms might have improved are covered by this category's first theme. While some patients reported a perceived symptom remission, others indicated that, while the treatment helped, they felt not fully recovered:

> *I was able to learn a lot and made big advancements. By now, I oftentimes go by car alone and I feel like I make progress every day, I also stopped thinking about my anxiety.* [P04; position 8]

> *Well, I still have anxiety. Partly, it got better.* [P03; position 18]

Various reasons were given for the perceived improvement. While some patients stressed the superior role of one component of the therapy (eg, therapeutic appointments, VRET, or psychoeducation), some pointed out that the comprehensive approach with a treatment app, mobile VRET, and psychotherapeutic appointments was helpful:

> *Yea, I'd say the therapist rather helped me, coming from an interpersonal understanding.* [P02; position 29]

> *It was an interplay. There was the app, in which I learned about anxiety and its causes. And then I had a person that accompanied or led me. The appointments helped me to keep going and push through.* [P06; position 50]

While sometimes the mobile VRET was highlighted as an especially helpful ingredient, others claimed that everything but the mobile VRET was beneficial:

> *And these experiments with the VR-headset. They really helped me, because in these situations, I really had anxiety, but I knew it was not real and it doesn't matter how I behave.* [P03; position 51]

> *I think the video took an hour, I was on the highway and I was in the passenger seat. But being the passenger was no problem for me. That's the same as if I was in somebody's car, that doesn't bother me either.* [P09; position 54]

In addition to the treatment components, symptom insight, learning to reflect on these symptoms, and receiving access to treatment after a long search were mentioned as crucial factors. Reflecting on their treatment progress, participants illuminated different protective factors that facilitated their symptom improvement, such as engaging modules, social support, hope, optimism, and motivation to use the app.

Theme 2: No Change and (Short-Term) Symptom Deterioration

Some patients reported retrospectively no change or even deterioration of their initial symptomatology. They brought up that the treatment was not able to address all aspects of their anxiety due to the manualization and the one-size-fits-all approach. In addition, they reported a lack of individualization and personalization:

> *It's not possible as it is a pretty general format. For example, symptoms I don't have are described or you are put in a state that does not cause anxiety. That is something I suspected to happen in the beginning.* [P11; position 108]

One patient brought up that, through psychoeducation, symptoms worsened by shifting the focus toward them:

> *I always watched the first few modules at night. This was usually the time when I got my panic attacks and listening to the modules was kind of a trigger for me. For example, it's about what happens in your brain and what symptoms it can have. After hearing about the symptoms, I started to notice them and that is how my panic was caused. Very unpleasant, nausea and shortness of breath. When the person in the app talked about it, it sometimes made it even worse for me.* [P02; position 48-49]

Theme 3: Side Effects

The theme of side effects covered codes addressing unwanted changes associated with the treatment but not related to the patients' anxiety disorders. Side effects included simulation sickness (eg, "I got a little headache" [P02; position 36]) and the surfacing of other problems (eg, trauma), which were not addressed by the app:

> *I think there still is something. For now, I'd use the app, but through the treatment, other areas arose that were problematic, which were kind of hidden before. In that sense, you could say that it had a negative effect.* [P10; position 44]

Theme 4: Motivation Throughout the Treatment

Motivation was oftentimes addressed by the interviewed patients, depicting it as a dynamic process throughout the treatment. Many mentioned it as a crucial factor to the success or failure of their treatment. Different factors were given as reasons for an increase or decrease in motivation. Patients described the motivation to reduce symptoms to be able to master feared situations and increase well-being:

> *I was totally motivated, I wanted it, no matter what. For me it was so important, as from day to day, my life completely shifted and I couldn't deal with it. Before, I was a happy and funny person and suddenly, I was afraid of everything.* [P07; position 64]

During the interviews, patients emphasized the necessity of being motivated for the, in large part, self-reliant treatment. Therefore, a lack of motivation was brought up as a potential risk factor for treatment success:

XSL•FO
RenderX

> *In the beginning [I was] very motivated, but at some point, that faded a little. And then I had a few weeks doing nothing. Then I realized I should do something and did a few exercises. But sometimes, I lack the motivation to sit down at home and engage [with the app]. I think it lacks the person that in some way "kicks your butt."* [P03; position 74]

However, a high initial motivation was no guarantee for long-term engagement with the treatment app. For example, a patient reported a reduction in motivation to engage with the app due to a quick initial symptom reduction and overall improvement. Eventually, this made it harder to find the required time for the treatment or resulted in difficulties facing their anxiety in the frame of the treatment and a lack of translation into their everyday life:

> *In the beginning, it was fairly easy [to work with the app] because I knew it had to happen. The weeks before that weren't really worth living. Once I got a grip on my anxiety, also because of the app and the therapy, it got harder, as it is quite time-consuming and I couldn't estimate how long the next module would take.* [P02; position 62]

Category 2: Mobile-Based Intervention–Related Themes

The second category covered 3 themes regarding the mobile-based app used in the intervention evaluated in this study: app content and exercises (theme 1), user interface (theme 2), and technicalities (theme 3). The displayed codes do not concern the mobile VRET, which is covered in another category (category 3: VRET-related themes).

Theme 1: Content and Exercises

The first theme covered patients' opinions concerning the content and exercises of the app. Patients viewed the modules covering psychoeducation, the in-app push notifications, and the option to repeat contents as especially helpful:

> *I was surprised, how well the content was specialized to my panic disorder. I guess there is different content available that is shown to you, depending on what you have.* [P02; position 62]

> *Also, it was nice that I received notifications.* [P04; position 33]

Patients had mixed opinions on the relaxation exercises. Some described them as "extremely helpful" (P08; position 16-17), whereas others perceived them as rather misleading and distressing:

> *These breathing or mindfulness exercises partly worsened my panic. Well, definitely the breathing exercises, as I tried it a lot in the beginning. As soon as I gave room to my panic and tried to breathe against it, it became worse and made it take longer.* [P02; position 34]

Some patients expressed a desire for in-app contact with other patients or their psychotherapist to share progress or solve issues with peers:

> *It could be that for social anxiety, this isn't realizable, but I could well imagine that it works like group therapy. So you could talk about it. "How was the situation for you?" "What was your take-home message?"* [P11; position 144]

Theme 2: User Interface

Overall, the app was perceived as well built regarding quality and user interface:

> *And I found [the app] qualitatively well-made. And it was good to listen to it and to read through it.* [P11; position 96]

The option to choose the in-app psychotherapist was brought up as an advantage. However, the importance of an option to speed up videos in the modules was highlighted. To further facilitate usability, a function to consume the content while being absent from the screen was desired:

> *Not being able to lock the screen, not being able to close the app. These are mechanisms that in some way make sense, but most of the time I used the app while cleaning, doing my laundry, or doing my homework. And then your cellphone is in another room and you have to click to proceed. That wasn't severe, but also not comfortable.* [P10; position 86]

Theme 3: Technicalities

The sign-up process was marked as a problem source, and a transparent guide on how to obtain access to the treatment app and a function to download content were deemed helpful. However, it was also pointed out that technical difficulties, if they occurred, were quickly solved by the app support:

> *One time a situation didn't load, so I contacted the support and they solved it fairly quickly, on the same day. It worked directly.* [P11; position 88]

Category 3: VRET-Related Themes

The third category consisted of 3 themes covering opinions that were related to the mobile VRET modules of the treatment app. Participants discussed their experiences in the VRET, including the degree of immersion, the usability, and the variety of situations displayed on the treatment app.

Theme 1: Immersion

The first theme covered the perceived degree of immersion. Regarding this theme, several patients emphasized the feeling of immersion as helpful for their improvement and the ability of the app to create immersion and elicit fear:

> *And then I had anxiety, even though it was not real. It was quite surprising that I was able to do that. It was quite interesting and for some part even fun. Overall, I think it was mostly interesting to get to know more about [the anxiety].* [P03; position 34]

> *I think that is especially because I am conscious about being in a virtual situation. Other aspects, such as noises from the outside felt different. It is so advanced that you can completely exchange certain things,*

XSL•FO
RenderX

> *which somehow felt very real, without the real consequences, however.* [P11; position 79]

On the other hand, other patients found the prerecorded situations unrealistic, which could have negatively affected the degree of immersion. In that sense, it was addressed as a shortcoming that the mobile VRET consisted of 360° videos without possible interaction. This was perceived as especially problematic by patients with SAD as the prerecorded audience did not react to them:

> *You wear the VR-headset and you are sat in front of an audience, which is obviously filmed and the VR doesn't really matter, because no matter what I do, they will always react the same. And that is not what causes my fear, it rather is the direct contact with others.* [P01; position 26]

> *So sometimes, when I misspoke, I had the feeling that it usually would've been a situation that would have caused a reaction. That took a bit of the immersion away. But I cannot estimate how much.* [P11; position 82]

In addition, some patients with PD or AG described a lack of anxiety when confronted with situations in mobile VR:

> *What I didn't like, was the virtual [reality], yes, that didn't match. I have to be honest, I sat on an airplane with weird people that did not look real and made repetitive movements. The airplane didn't even shake or do anything that would have caused something in me. Or being alone in the basement. You would imagine this being creepy, yes? But nothing happened, not even a mouse hushed by or anything. I tried a lot, but that wasn't for me. Or a shopping queue, nothing happened. Unrealistic. But for the rest, the app was great.* [P07; position 28]

Theme 2: Usability

The second theme was the usability of the VRET. While some patients did like facing feared situations at home, others felt discomfort when using mobile VRET:

> *My parents were home for the whole day so I could only train in situations when no one was at home. At some time in the evening, when I was sure that no one could hear me.* [P05; position 66]

Some patients described problems with the compatibility of the VR app and their smartphone despite their device meeting the manufacturer's specified minimum technical requirements. For example, they mentioned lags in the 360° videos:

> *It should have worked with my phone, but it lagged so much, for five seconds it worked and then it stood still for another five.* [P09; position 52]

Regarding the VR headset, patients had problems with adequate fit and the glasses fogging up on the inside:

> *The VR headset did not really fit well and I got a bit of a headache because I wasn't able to correctly adjust the resolution.* [P02; position 36]

Another patient brought up that the weight of the head-mounted construction with the smartphone inside made it difficult to wear the equipment for an extended period:

> *I found the headset very heavy. At some point, I had to put my elbows on my knees and by that point, it did not feel like not being at home on the couch.* [P10; position 107-108]

Finally, it was suggested that, as the app was only available for 90 days, the VR equipment should be only distributed temporarily as, when the license ends, the patient has no further use for the equipment that is tailored to the treatment app.

Theme 3: Variety of VR Content

A third theme covered the situations available in VR within the treatment app. Patients described the situations presented on the app as aligning well with fear-evoking situations in real life. They also liked the variety of agoraphobic situations they were able to experience without having to face these situations in real life:

> *So first of all, being able to identify the anxiety and then being able to directly test it. It was nice to directly go into the specific situations. So, thinking back, I would have never gone into the situation [in reality]. It was practical, I did not need to go somewhere to go through the fear.* [P06; position 100]

However, the situations were criticized for the lack of eventfulness. Patients did expect options to further increase fear and adjust situations:

> *I thought that the situations would increase in intensity. I thought that eventually, I would stand on a huge stage and give a speech. And I thought that something VR-related would happen with the people. Not being in a prerecorded situation, but something like a chat.* [P11; position 100]

While the app did implement situations for SAD (eg, giving a talk in front of an audience) and AG (eg, a bus ride), one patient criticized the lack of available situations for eliciting bodily arousal PD–related symptoms:

> *In the end, some modules did not fit my panic disorder anymore but fitted more to other anxiety disorders.... But I think that also was because I wasn't able to pick a location that would cause my panic. For me, it is rather a personal, bodily thing. I tried one exercise on the airplane but after two minutes it was boring as it did not elicit panic.* [P02; position 34-36 and position 62]

Category 4: Context-Related Themes

The fourth category comprised themes that addressed the framework in which the patients received the treatment.

Theme 1: Significance of Treatment Guidance

The treatment consisted of an app with mandatory psychotherapeutic sessions for the preparation and postprocessing of certain modules. Many of the patients addressed this guidance in their interviews. For most, this was a crucial ingredient for symptom improvement or a protective

XSL•FO
RenderX

factor against deterioration. In addition, patients who did not particularly like the treatment app or the mobile VRET components still regarded the psychotherapeutic sessions as helpful:

> *The therapist definitely helped. Just being able to talk and also for my understanding. So, the empathy and to accept it as normal and not be upset about it.* [P02; position 27-28]

Despite the guidance, it was criticized that patients could not share the completed modules with their psychotherapist. However, the combination of the treatment app and the psychotherapeutic sessions was brought up as a well-rounded mix by most patients. One of the reasons given was that the appointments facilitated the integration of the content learned on the app:

> *After I held the speech [in VR] I had an appointment and the therapist was able to improve my anxiety a little, which I would have never achieved on my own. That's why it was very good that I still had the appointments. Completely on my own, it wouldn't have been possible.* [P03; position 38]

Some patients addressed the desire for more appointments accompanying the app-based treatment:

> *I would've appreciated it if the time between the appointments was shorter and that we would've met more often and spoken more.* [P07; position 82]

Theme 2: Study Schedule

Patients attended psychotherapeutic appointments at prearranged intervals as part of the RCTs. Most patients indicated not feeling like "being part of a research scenario" (P01; position 54) but being glad to contribute to research findings. However, to enable the manualized appointments in the experimental condition, the patients were asked to engage with the treatment app between the scheduled sessions. While this was perceived as helpful for staying motivated by most patients, some patients perceived it as stressful. The schedule was linked to treatment outcomes and limited the flexibility of the treatment:

> *First I felt relieved because usually, you have to wait a long time for a therapy spot. I was happy about any help, after all. It went quickly and I felt relieved, but also obligated to do this and that module between the appointments.... A downside was that time-wise, you are independent, but also in this study combination. If you have this time frame in which you have to finish certain modules before coming to the next appointment, it is a downside for me, as it limits my flexibility.* [P02; position 13-14 and 56-57]

Theme 3: Motives for Study Participation

In their interviews, patients expressed their motives for participating in this study. For example, one patient was motivated to improve the supply situation for women in mental health care:

> *It is super important to generate new findings, therefore I find [participating in research] always positive. Especially, I think for many diseases, women are underrepresented, and when I can contribute to improving this and technology.* [P02; position 69-70]

Another motive was receiving access to therapy as this is often hindered by long waiting times. Patients hoped to be prioritized by contributing to research:

> *I find the first step very hard and also exhausting when you are looking for a therapist. And when through [the study], you can quickly talk to someone, I consider this helpful, as you are rewarded for your first step and you don't need to call another 20 [psychotherapists] and receive 19 rejections.* [P01; position 62]

Theme 4: Role of Mobile-Based Interventions in Mental Health Care

The fourth theme consisted of different opinions regarding the role of app treatment in mental health care. As such, IMIs were considered as an add-on to conventional therapy (eg, as an alternative to traditional homework or to bridge the waiting period before beginning psychotherapy):

> *Sometimes I think [waiting times] take way too long. The anxiety or the problems get more severe the longer it takes until they finally receive treatment. With this, you can at least work a bit on your own.* [P07; position 198]

In addition, IMIs were considered as having fewer barriers. This way, they could be a well-suited first step for patients searching for help:

> *I think many people are afraid to attend therapy in the first place or to completely trust another person. So I think it is easier to start the treatment with an app. For example, writing down things in the app without the therapist noticing. So I can imagine that this could lessen the barrier to starting a treatment.* [P11; position 67]

However, while IMIs were perceived as suitable for complementing therapy or bridging waiting times, some patients also described limitations in their applicability. They thought that IMIs might only be suited for certain aspects of treatment or patients. For example, some people might prefer books over using their smartphones to inform themselves about their disorders. P05 mentioned that she perceived the app treatment as helpful as she already knew what her problem was:

> *If you have problems in many areas of your life and the app asks you to name your problems and then you type down an eternally long text, in that case, a real conversation is more practical. However, if you know your problem and just want to work on it.... For me, that was very helpful and I think that could also be helpful for others.* [P05; position 82]

Patients consistently made clear that, from their experience, IMIs could be a complementary tool but should not replace psychotherapy:

XSL•FO
RenderX

> *Difficult. I think that depends on the person. But I cannot imagine [IMIs replacing therapy]. Not in the current state.* [P03; position 96]

Discussion

Principal Findings

Overview

The goal of this study was to systematically analyze patients' perspectives concerning the integration of mobile VRET into guided IMIs for treating anxiety disorders. This was done to inform how patients perceive this innovative treatment approach and where chances and challenges might lie. This way, in 4 overarching categories, 14 themes were identified: symptom improvement, no change and deterioration, side effects, motivation, content and structure, user interface, technicalities, immersion, usability, variety of situations, treatment guidance, study schedule, motives for study participation, and role of IMIs in mental health care.

Perspectives on Treatment Outcomes

The first category of themes captured patients' experiences with symptom alleviation throughout the treatment. Most striking about the themes in this category was that patients articulated a variety of reasons for their change or no change in symptoms and linked these to different aspects of the treatment. Improvement was explained by oftentimes highlighting the helpful interactions with the guiding psychotherapist but also by pointing out especially helpful treatment components such as psychoeducation, which was easy to follow as well as informational. Mobile VRET was also frequently brought up as allowing for the controlled encounter of feared and previously avoided situations [30]. Thus, fear responses can be elicited in a controlled virtual setting, enabling a confrontation of fears in line with the principles of in vivo exposure therapy, leveraging gradual exposure and habituation while maintaining the control of a simulated environment. By reducing logistical barriers, mobile VRET highlights the potential it holds in mental health care [35], which is further supported by meta-analytic findings yielding no inferiority of VRET compared to in vivo exposure therapy concerning primary treatment outcomes [61-63]. No change was explained by the treatment components not meeting the individual symptomatology, showing that the superficial modality of an app-based treatment might not fit every patient. The results give rise to the argument that treatment apps alone may not provide the depth of therapeutic interaction needed for comprehensive mental health care. Regarding the theme of side effects, how to deal with adverse events related to the treatment needs to be reflected on [23], as it happened to one patient, who learned through the treatment that there were other issues they needed to deal with for which there was no room in the manualized short-term treatment. Without the frame of the study context, this occurrence could have gone unnoticed, and the patient would most likely not have received further treatment.

Bringing together these findings with the results of previous quantitative studies, it can be stated that there are very few but promising studies on the efficacy of mobile VRET in IMIs. These emphasize clinical improvements and provide initial evidence of treatment benefits, particularly for components such as mobile VRET [46,64]. While additional quantitative studies have yet to profoundly support the treatment's efficacy, the current qualitative studies can expand the findings by showing that the patient-perceived treatment outcomes are often tied to specific treatment components, such as psychoeducation, VRET, or tension regulation. These components have also been found efficacious in quantitative research [42,65], whereas other qualitative studies have determined motivation, symptom insight, and social support as crucial factors in treatments incorporating IMIs [52,53]. In comparison, the qualitative approach used in this study highlighted that the personally perceived commitment, motivation, and outcome expectancy were intertwined with personal impairments (anticipatory anxiety limiting daily routine and reinforcing the circle of avoidance and increase in fear over time) and treatment goals (building confidence for certain scenarios and gaining tools to manage physiological symptoms), which cannot be taken into account in many quantitative approaches.

Patients pointed out motivation as a central construct when addressing treatment outcomes. While motivation has also been found to be a key factor in traditional psychotherapy [66], it appears to play an even bigger role when the patient is self-reliant in working through most of the treatment content [67]. In particular, motivation was described as a dynamic process that is highest in the beginning, when the psychological strain is highest, and decreases throughout the treatment. After its decrease, it is then associated with no change or even deterioration, resembling previous findings on motivation in traditional psychotherapy [68]. For this treatment approach, this finding highlights the need for additional motivators as the treatment progresses to protect the patient from demotivation and losing benefits from the treatment.

Perspectives on the Treatment App

The second category of themes covered patient opinions regarding the treatment app. Most strikingly, psychoeducation was brought up as an overwhelmingly positive treatment component that can be delivered well via an app, which is in line with previous research on patient perspectives on digital treatment modalities [69]. In addition, a high degree of individualization (eg, choosing the voice of the in-app psychotherapist) and an adjustable user interface were stressed as preferable. An aspect that faced repeated criticism was the in-app relaxation. While mindfulness interventions have been found efficacious in mental health care [70], this sample encompassed patients with anxiety disorders. It might be that, for such disorders, in particular regarding PD, relaxation is better delivered via psychotherapeutic interaction than through an app considering the possibility of relaxation-induced panic attacks [71]. It is also known that confrontation with situations and contexts that have previously been avoided contributes to an initial increase in fear or discomfort, which is positively associated with treatment success [72]. However, in the framework of self-guided or blended treatment, patients might be more inclined to discontinue treatment as they perceive the initial negative emotional states as threatening, thereby reinforcing avoidance behavior and, ultimately, intensifying the risk of unwanted events. For patients, it was also important to

XSL•FO
RenderX

have quick access to the treatment app and be able to use it without barriers. One of these barriers was long informational sections, which patients could not speed up or listen to while being active (eg, walking or cleaning) as the app would require the user to click to move along. These findings are comparable to those of other qualitative studies on opinions on app treatment as, in general, a desire for a high degree of personalization stood out in this study as well [69]. However, while implementing the patients' suggestions might lead to more treatment acceptance, whether these suggestions could raise other issues has to be carefully considered for each suggestion on its own. For example, allowing for the locking of the screen or opening of other apps while using the treatment app could lead to less focus on the treatment and, eventually, a decline in motivation or less treatment success.

Perspectives on Mobile VRET

In the third category, themes were listed that were related to the mobile VRET of the treatment app. The most discussed aspect was the degree to which the head-mounted construction with the patient's smartphone serving as a display was able to create an immersive experience for VRET [43]. While it was welcomed by some patients, who described being pleased with the high immersion, others criticized a lack of interaction given the fact that the mobile VRET consisted of prerecorded 360° videos. As previously outlined, different systems are used in VRET that are either equipped with dynamic interaction options or, as in this study, based on 360° videos that do not offer the chance to affect the scenarios or characters in them. Patients addressed this by pointing out a lack of interpersonal interaction during VRET, which made the experience less immersive. This aligns with the clinical presentation of SAD, where interpersonal interaction plays a central role [5,44], and suggests that the degree of the immersive interaction required might vary depending on the disorder. Nonetheless, many participants across all 3 anxiety disorders found the 360° videos personally beneficial for treating their anxiety, appreciating the practical benefits of accessibility and ease of use, supporting their engagement during the exposure tasks. It is important to note that these evaluations were conducted without directly comparing 360° videos to interactive VRET systems, which could be addressed in future research. In this study, the hardware was identified as a problem source in the theme of usability. These findings suggest that mobile VRET still needs time to overcome these particular challenges but, overall, can be useful in treating SAD, PD, and AG.

Although numerous studies have drawn out the possibilities of VRET in mental health care [35], it is yet to be seen how the intriguing treatment approach of using mobile VRET can be further improved to be well received by patients. One way could be to integrate desired elements such as the possibility to choose from a large pool of fear-evoking situations with optional adjustments, such as the crowdedness of certain places or the pace of car, bus, or tram rides, as brought up by the interviewed patients.

Perspectives on the Study's Context

Themes in the fourth category, which were context related, showed that, for the patients, the psychotherapeutic supervision was very important, the study schedule may have had an influence on symptom change, and most patients could well imagine using (guided) treatment apps but could not imagine them as a substitute for psychotherapy. Overall, the interviewed patients highly valued the psychotherapeutic sessions as part of the guided treatment approach. Despite the app being the main ingredient of their treatment, the psychotherapeutic guidance was mentioned in all interviews and was always brought up positively or with the request to have more psychotherapeutic sessions. This finding is in line with those of previous literature that found guided IMIs to be superior to unguided ones [26], indicating that alliance as a mechanism of change also holds up for digitally delivered interventions [73]. While the general modality of using IMIs is partly aimed at reducing psychotherapeutic involvement [23], how far this reduction is taken should be carefully considered for each treatment approach. In this study, many interviewed patients distinguished the psychotherapeutic appointment as the most remarkable factor in their treatment, whereas conceptually, one could argue that it was only intended as a supporting element, with the main focus laying on the treatment app with mobile VRET [47,48].

The thematic connection between the theme of study schedule and the theme of symptom improvement originated in patients giving aspects of the study schedule as reasons for their improvement. For example, it was mentioned that the psychotherapeutic appointments and calls from the study team served as deadlines and reminders to finish the modules in the treatment app, which then led to a more holistic understanding of etiological and maintaining factors underlying their anxiety disorder. Here, it is important to note that these checkups would not have occurred in a real clinical setting. Despite making sure that the RCTs from which the interviewed patients were recruited accurately reflected real-world clinical settings, it is essential to acknowledge that the patients still engaged in a structured intervention within a controlled environment [47,48]. Therefore, these minimal but necessary alterations could have influenced the quantitative results. Beyond trials investigating the treatment efficacy, it is important to also further identify characteristics in patients (eg, demographic and clinical variables, personality traits, and coping mechanisms), as well as process (eg, fear reduction during exposure) and context variables that may predict the treatment outcomes of these interventions and develop clear indications as to who should gain access to these treatment options [72,74,75]. In addition, for this reason, future interventions could increasingly incorporate deadlines and checkups to increase the patients' commitment to engage with the otherwise self-reliant treatment contents.

While treatment acceptance is a crucial factor in mostly self-delivered treatments, in their interviews, patients also reflected on the role of IMIs in the mental health care system. Due to the scarcity of treatment capabilities, they could imagine IMIs as a way to bridge waiting times. However, they could not imagine the tested treatment substituting psychotherapy as a whole, which is comparable to other qualitative research on patient attitudes toward IMIs [54]. This finding is also in line with those of research on the perspective of psychotherapists

XSL•FO
RenderX

on IMIs, who welcomed the digital treatment option as an aid for certain aspects but also could not imagine it as a substitute for psychotherapy [76].

Relating this study's results to the current situation of outpatient psychotherapy, several advantages and challenges of the investigated treatment modality can be highlighted. As accessibility and a low threshold for engaging with IMIs were appreciated, they might be advantageous in bridging waiting times or reducing stigma or hesitancy associated with seeking psychotherapeutic treatment [26]. This might be especially beneficial in countries with long waiting times, such as Germany [77]. Another advantage was that the treatment functioned as digital homework, contributing to problem activation outside the treatment, which has been associated with treatment success in the past [18]. Patients also listed the cost and scalability of the treatment as potential advantages over outpatient psychotherapy. However, oftentimes, the importance of psychotherapeutic guidance was emphasized, which is why a challenge might lie in the self-guided nature of mobile VRET. Namely, the procedure partly opposes current treatment guidelines [10], indicating that initial exposure attempts are best psychotherapeutically accompanied to prevent adverse effects or misapplication (ie, premature termination). In addition, a lack of individualized support remains a challenge, which was criticized by patients in direct comparison to regular outpatient psychotherapy.

Limitations and Future Research

While this study was capable of drawing a multifaceted picture of patients' opinions, some limitations need to be addressed when interpreting the findings. First, the findings of this study are derived from a blended IMI with integrated mobile VRET, indicating that the observed benefits or challenges may not directly translate to other treatment modalities. For instance, experiences with unguided IMIs or stationary VRET may differ from those with blended IMIs with mobile VRET. In addition, while the reliance on a qualitative method of data analysis to interpret the obtained interview transcripts offers rich insights, it also inherently requires caution when interpreting causality. Such an approach captures patient perspectives but does not allow for claims regarding the treatment's efficacy, which would require further exploration through controlled, comparative, and quantitative research designs. Future trials of a confirmatory nature are needed to further investigate the exploratory findings of this study.

Another point worth noting is that the interviews were conducted as part of a larger RCT. From the patients' perspective, it was interpreted that the study schedule, despite being incorporated into an outpatient clinical setting, might be linked to the obtained results. Differences between the study and treatment as usual include, for example, a high degree of standardization, required comparability of therapeutic settings, frequency and duration of therapy, and a reliance on treatment manuals. Therefore, for future research, it could be beneficial to conduct studies in a more naturalistic setting. This could include investigating treatment adherence and feasibility in less structured clinical settings such as standard outpatient care. Another relevant option for future research could be investigating challenges to the treatment's accessibility in contexts with different resources and infrastructure (eg, rural vs urban regions) as IMIs and mobile VRET in particular are theorized to possibly address disparities in the supply of mental health care by reducing geographical barriers and providing scalable solutions in underserved areas.

Conclusions

This study explored patients' perspectives on IMIs embedding mobile VRET for SAD, PD, and AG. As such, various components such as the psychoeducational content were listed as beneficial in the intervention, and motivation was continuously perceived as a crucial factor for treatment engagement. The mobile VRET was welcomed as an innovative add-on to overcome anxiety but, for some patients, lacked immersion and personal fit due to the reliance on prerecorded 360° videos without interaction. Finally, the psychotherapeutic guidance of the intervention was appreciated as especially beneficial and stood out as a central aspect in most of the interviews as it helped the patients integrate their learning from the IMI and served as a deadline to engage with the IMI before the next treatment session. This study has shown that, in the further development of IMIs and other associated interventions such as VRET, it is important not only to empirically test the efficacy of such procedures or their acceptance among psychotherapists but also not to ignore the perspectives of patients and those affected as these are highly valuable to further improve existing procedures as well as ensure their applicability. For future research, it is suggested to explore patients' perspectives on other treatment modalities or investigate the findings of this study using confirmatory research designs.

Acknowledgments

The authors thank the patients who participated in this study and the subsequent interview procedure. JP and ASH are funded by the German Research Foundation, Research Training Group 2493 (DFG project number: 398510439). No monetary (or any other type of) remuneration has been received by the developing team of the treatment app used in this trial.

Conflicts of Interest

None declared.

References

XSL•FO
RenderX

1. Barrera TL, Norton PJ. Quality of life impairment in generalized anxiety disorder, social phobia, and panic disorder. J Anxiety Disord 2009 Dec;23(8):1086-1090 [FREE Full text] [doi: 10.1016/j.janxdis.2009.07.011] [Medline: 19640675]
2. Stein DJ, Scott KM, de Jonge P, Kessler RC. Epidemiology of anxiety disorders: from surveys to nosology and back. Dialogues Clin Neurosci 2017 Jun;19(2):127-136 [FREE Full text] [doi: 10.31887/DCNS.2017.19.2/dstein] [Medline: 28867937]
3. Alonso J, Angermeyer M, Bernert S, Bruffaerts R, Brugha TS, Bryson H, et al. Prevalence of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. Acta Psychiatr Scand Suppl 2004(420):21-27. [doi: 10.1111/j.1600-0047.2004.00327.x] [Medline: 15128384]
4. Dilling H, Mombour W, Schmidt MH, World Health Organization. Internationale Klassifikation Psychischer Sto¨rungen: ICD-10, Kapitel V (F), Klinisch-Diagnostische Leitlinien. Geneva, Switzerland: World Health Organization; 1991.
5. Lecrubier Y, Wittchen HU, Faravelli C, Bobes J, Patel A, Knapp M. A European perspective on social anxiety disorder. Eur Psychiatry 2000 Feb;15(1):5-16. [doi: 10.1016/s0924-9338(00)00216-9] [Medline: 10713797]
6. Angst J. Panic disorder: history and epidemiology. Eur Psychiatry 2020 Apr 16;13(S2):51s-55s. [doi: 10.1016/s0924-9338(98)80014-x]
7. Wittchen HU, Gloster AT, Beesdo-Baum K, Fava GA, Craske MG. Agoraphobia: a review of the diagnostic classificatory position and criteria. Depress Anxiety 2010 Feb;27(2):113-133. [doi: 10.1002/da.20646] [Medline: 20143426]
8. Kashdan TB, Goodman FR, Machell KA, Kleiman EM, Monfort SS, Ciarrochi J, et al. A contextual approach to experiential avoidance and social anxiety: evidence from an experimental interaction and daily interactions of people with social anxiety disorder. Emotion 2014 Aug;14(4):769-781 [FREE Full text] [doi: 10.1037/a0035935] [Medline: 24749634]
9. Pfaltz MC, Michael T, Meyer AH, Wilhelm FH. Reexperiencing symptoms, dissociation, and avoidance behaviors in daily life of patients with PTSD and patients with panic disorder with agoraphobia. J Trauma Stress 2013 Aug;26(4):443-450. [doi: 10.1002/jts.21822] [Medline: 23893375]
10. Bandelow B, Aden I, Alpers GW, Benecke A, Benecke C, Deckert J, et al. Deutsche S3-leitlinie behandlung von angststörungen: version 2. AWMF. 2021. URL: https://register.awmf.org/assets/guidelines/051-028k_S3_Behandlung-von-Angststoerungen_2021-06.pdf [accessed 2025-04-12]
11. Bandelow B, Michaelis S, Wedekind D. Treatment of anxiety disorders. Dialogues Clin Neurosci 2022 Apr 01;19(2):93-107. [doi: 10.31887/dcns.2017.19.2/bbandelow]
12. Heimberg RG. Cognitive-behavioral therapy for social anxiety disorder: current status and future directions. Biol Psychiatry 2002 Jan 01;51(1):101-108. [doi: 10.1016/s0006-3223(01)01183-0] [Medline: 11801235]
13. Stein DJ, Lim CC, Roest AM, de Jonge P, Aguilar-Gaxiola S, Al-Hamzawi A, et al. The cross-national epidemiology of social anxiety disorder: data from the world mental health survey initiative. BMC Med 2017 Jul 31;15(1):143 [FREE Full text] [doi: 10.1186/s12916-017-0889-2] [Medline: 28756776]
14. Issakidis C, Andrews G. Pretreatment attrition and dropout in an outpatient clinic for anxiety disorders. Acta Psychiatr Scand 2004 Jun;109(6):426-433. [doi: 10.1111/j.1600-0047.2004.00264.x] [Medline: 15117287]
15. Freiheit SR, Vye C, Swan R, Cady M. Cognitive-behavioral therapy for anxiety: is dissemination working? Behav Ther 2004;27(2):25-32 [FREE Full text]
16. Pittig A, Hoyer J. Exposition aus sicht niedergelassener verhaltenstherapeutinnen und verhaltenstherapeuten. Zeitschrift für klinische Psychologie Psychopathologie 2018 Apr 10;46(4):223-235. [doi: 10.1026/1616-3443/a000441]
17. Foa EB, Kozak MJ. Emotional processing of fear: exposure to corrective information. Psychol Bull 1986;99(1):20-35. [doi: 10.1037//0033-2909.99.1.20]
18. Gassmann D, Grawe K. General change mechanisms: the relation between problem activation and resource activation in successful and unsuccessful therapeutic interactions. Clin Psychol Psychother 2006 Feb 02;13(1):1-11. [doi: 10.1002/cpp.442]
19. Pittig A, Pittig R. Individualisierte exposition bei angststörungen. Psychotherapie 2024 Jan 08;69:67-82. [doi: 10.1007/s00278-023-00698-y]
20. Craske MG, Hermans D, Vervliet B. State-of-the-art and future directions for extinction as a translational model for fear and anxiety. Philos Trans R Soc Lond B Biol Sci 2018 Mar 19;373(1742):20170025 [FREE Full text] [doi: 10.1098/rstb.2017.0025] [Medline: 29352025]
21. Fernández A, Haro JM, Martinez-Alonso M, Demyttenaere K, Brugha TS, Autonell J, et al. Treatment adequacy for anxiety and depressive disorders in six European countries. Br J Psychiatry 2007 Feb;190:172-173 [FREE Full text] [doi: 10.1192/bjp.bp.106.023507] [Medline: 17267936]
22. Horenstein A, Heimberg RG. Anxiety disorders and healthcare utilization: a systematic review. Clin Psychol Rev 2020 Nov;81:101894. [doi: 10.1016/j.cpr.2020.101894] [Medline: 32818687]
23. Ebert DD, Van Daele T, Nordgreen T, Karekla M, Compare A, Zarbo C, et al. Internet- and mobile-based psychological interventions: applications, efficacy, and potential for improving mental health. Eur Psychol 2018 May 14;23(2):167-187. [doi: 10.1027/1016-9040/a000318]
24. Machulska A, Eiler TJ, Kleinke K, Grünewald A, Brück R, Jahn K, et al. Approach bias retraining through virtual reality in smokers willing to quit smoking: a randomized-controlled study. Behav Res Ther 2021 Jun;141:103858. [doi: 10.1016/j.brat.2021.103858] [Medline: 33862407]

XSL•FO
RenderX

25. Weisel KK, Zarski AC, Berger T, Krieger T, Schaub MP, Moser CT, et al. Efficacy and cost-effectiveness of guided and unguided internet- and mobile-based indicated transdiagnostic prevention of depression and anxiety (ICare Prevent): a three-armed randomized controlled trial in four European countries. Internet Interv 2018 Apr 15;16:52-64 [FREE Full text] [doi: 10.1016/j.invent.2018.04.002] [Medline: 30775265]
26. Domhardt M, Geßlein H, von Rezori RE, Baumeister H. Internet- and mobile-based interventions for anxiety disorders: a meta-analytic review of intervention components. Depress Anxiety 2019 Mar;36(3):213-224. [doi: 10.1002/da.22860] [Medline: 30450811]
27. Ebert DD, Cuijpers P, Muñoz RF, Baumeister H. Prevention of mental health disorders using internet- and mobile-based interventions: a narrative review and recommendations for future research. Front Psychiatry 2017 Aug 10;8:116 [FREE Full text] [doi: 10.3389/fpsyt.2017.00116] [Medline: 28848454]
28. Deacon BJ, Farrell NR. Therapist barriers to the dissemination of exposure therapy. In: Storch E, McKay D, editors. Handbook of Treating Variants and Complications in Anxiety Disorders. New York, NY: Springer; Feb 28, 2013.
29. Wechsler TF, Kümpers F, Mühlberger A. Inferiority or even superiority of virtual reality exposure therapy in phobias?-a systematic review and quantitative meta-analysis on randomized controlled trials specifically comparing the efficacy of virtual reality exposure to gold standard in vivo exposure in agoraphobia, specific phobia, and social phobia. Front Psychol 2019 Sep 10;10:1758 [FREE Full text] [doi: 10.3389/fpsyg.2019.01758] [Medline: 31551840]
30. Bush J. Viability of virtual reality exposure therapy as a treatment alternative. Comput Hum Behav 2008 May;24(3):1032-1040. [doi: 10.1016/j.chb.2007.03.006]
31. Machulska A, Roesmann K, Eiler TJ, Grünewald A, Brück R, Klucken T. Der einsatz von virtueller realität in der psychotherapeutischen Praxis: aktueller forschungsstand, chancen, risiken und herausforderungen. Psychother Forum 2021 Dec 03;25:169-176. [doi: 10.1007/s00729-021-00185-2]
32. Powers MB, Emmelkamp PM. Virtual reality exposure therapy for anxiety disorders: a meta-analysis. J Anxiety Disord 2008;22(3):561-569. [doi: 10.1016/j.janxdis.2007.04.006] [Medline: 17544252]
33. Anderson PL, Price M, Edwards SM, Obasaju MA, Schmertz SK, Zimand E, et al. Virtual reality exposure therapy for social anxiety disorder: a randomized controlled trial. J Consult Clin Psychol 2013 Oct;81(5):751-760. [doi: 10.1037/a0033559] [Medline: 23796315]
34. Geraets CN, Veling W, Witlox M, Staring AB, Matthijssen SJ, Cath D. Virtual reality-based cognitive behavioural therapy for patients with generalized social anxiety disorder: a pilot study. Behav Cogn Psychother 2019 Mar 27;47(6):745-750. [doi: 10.1017/s1352465819000225]
35. Emmelkamp PM, Meyerbröker K. Virtual reality therapy in mental health. Annu Rev Clin Psychol 2021 May 07;17:495-519. [doi: 10.1146/annurev-clinpsy-081219-115923] [Medline: 33606946]
36. Horigome T, Kurokawa S, Sawada K, Kudo S, Shiga K, Mimura M, et al. Response to Morina et al.'s criticisms of Horigome et al.'s recent report on the efficacy of virtual reality exposure therapy for social anxiety disorder. Psychol Med 2021 Oct 19;53(5):2179-2180. [doi: 10.1017/s0033291721004086]
37. Meyerbröker K, Morina N. The use of virtual reality in assessment and treatment of anxiety and related disorders. Clin Psychol Psychother 2021 May;28(3):466-476 [FREE Full text] [doi: 10.1002/cpp.2623] [Medline: 34097318]
38. Morina N, Kampmann I, Emmelkamp P, Barbui C, Hoppen TH. Meta-analysis of virtual reality exposure therapy for social anxiety disorder. Psychol Med 2021 May 18;53(5):2176-2178. [doi: 10.1017/s0033291721001690]
39. Horigome T, Kurokawa S, Sawada K, Kudo S, Shiga K, Mimura M, et al. Virtual reality exposure therapy for social anxiety disorder: a systematic review and meta-analysis. Psychol Med 2020 Nov;50(15):2487-2497. [doi: 10.1017/S0033291720003785] [Medline: 33070784]
40. Botella C, García‐Palacios A, Villa H, Baños RM, Quero S, Alcañiz M, et al. Virtual reality exposure in the treatment of panic disorder and agoraphobia: a controlled study. Clin Psychol Psychother 2007 May 25;14(3):164-175. [doi: 10.1002/cpp.524]
41. Garcia-Palacios A, Botella C, Hoffman H, Fabregat S. Comparing acceptance and refusal rates of virtual reality exposure vs. in vivo exposure by patients with specific phobias. Cyberpsychol Behav 2007 Oct;10(5):722-724. [doi: 10.1089/cpb.2007.9962] [Medline: 17927544]
42. Krijn M, Emmelkamp PM, Olafsson RP, Biemond R. Virtual reality exposure therapy of anxiety disorders: a review. Clin Psychol Rev 2004 Jul;24(3):259-281. [doi: 10.1016/j.cpr.2004.04.001] [Medline: 15245832]
43. Amin A, Gromala D, Tong X, Shaw C. Immersion in cardboard VR compared to a traditional head-mounted display. In: Proceedings of the 8th International Conference on Virtual, Augmented and Mixed Reality. 2016 Presented at: VAMR 2016; July 17-22, 2016; Toronto, Canada. [doi: 10.1007/978-3-319-39907-2_25]
44. Morina N, Brinkman W, Hartanto D, Kampmann IL, Emmelkamp PM. Social interactions in virtual reality exposure therapy: a proof-of-concept pilot study. Technol Health Care 2015 Sep 01;23(5):581-589. [doi: 10.3233/thc-151014]
45. Ionescu A, Van Daele T, Rizzo A, Blair C, Best P. 360° videos for immersive mental health interventions: a systematic review. J Technol Behav Sci 2021 Sep 29;6:631-651. [doi: 10.1007/s41347-021-00221-7]
46. Donker T, Cornelisz I, van Klaveren C, van Straten A, Carlbring P, Cuijpers P, et al. Effectiveness of self-guided app-based virtual reality cognitive behavior therapy for acrophobia: a randomized clinical trial. JAMA Psychiatry 2019 Jul 01;76(7):682-690 [FREE Full text] [doi: 10.1001/jamapsychiatry.2019.0219] [Medline: 30892564]

XSL•FO
RenderX

47. Planert J, Machulska A, Hildebrand AS, Roesmann K, Otto E, Klucken T. Self-guided digital treatment with virtual reality for panic disorder and agoraphobia: a study protocol for a randomized controlled trial. Trials 2022 May 21;23(1):426 [FREE Full text] [doi: 10.1186/s13063-022-06366-x] [Medline: 35597959]
48. Hildebrand AS, Roesmann K, Planert J, Machulska A, Otto E, Klucken T. Self-guided virtual reality therapy for social anxiety disorder: a study protocol for a randomized controlled trial. Trials 2022 May 12;23(1):395 [FREE Full text] [doi: 10.1186/s13063-022-06320-x] [Medline: 35549980]
49. Campos D, Bretón-López J, Botella C, Mira A, Castilla D, Mor S, et al. Efficacy of an internet-based exposure treatment for flying phobia (NO-FEAR Airlines) with and without therapist guidance: a randomized controlled trial. BMC Psychiatry 2019 Mar 06;19(1):86 [FREE Full text] [doi: 10.1186/s12888-019-2060-4] [Medline: 30841930]
50. Schultz J, Baumeister A, Schmotz S, Moritz S, Jelinek L. Efficacy of an internet-based intervention with self-applied exposure therapy in virtual reality for people with panic disorder: study protocol for a randomized controlled trial. Trials 2023 Aug 12;24(1):521 [FREE Full text] [doi: 10.1186/s13063-023-07536-1] [Medline: 37573377]
51. Mayer G, Gronewold N, Polte K, Hummel S, Barniske J, Korbel JJ, et al. Experiences of patients and therapists testing a virtual reality exposure app for symptoms of claustrophobia: mixed methods study. JMIR Ment Health 2022 Dec 05;9(12):e40056. [doi: 10.2196/40056] [Medline: 36469413]
52. Lindner P, Rozental A, Jurell A, Reuterskiöld L, Andersson G, Hamilton W, et al. Experiences of gamified and automated virtual reality exposure therapy for spider phobia: qualitative study. JMIR Serious Games 2020 Apr 29;8(2):e17807 [FREE Full text] [doi: 10.2196/17807] [Medline: 32347803]
53. Asklund I, Samuelsson E, Hamberg K, Umefjord G, Sjöström M. User experience of an app-based treatment for stress urinary incontinence: qualitative interview study. J Med Internet Res 2019 Mar 14;21(3):e11296 [FREE Full text] [doi: 10.2196/11296] [Medline: 30869644]
54. Krakowczyk JB, Truijens F, Teufel M, Lalgi T, Heinen J, Schug C, et al. Evaluation of the e-mental health intervention make it training from patients' perspectives: qualitative analysis within the reduct trial. JMIR Cancer 2024 Apr 09;10:e53117 [FREE Full text] [doi: 10.2196/53117] [Medline: 38592764]
55. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
56. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport Exerc Health 2019 Jun 13;11(4):589-597. [doi: 10.1080/2159676X.2019.1628806]
57. Invirto homepage. Invirto. URL: https://invirto.de/ [accessed 2025-04-08]
58. Hopper SI, Murray SL, Ferrara LR, Singleton JK. Effectiveness of diaphragmatic breathing on physiological and psychological stress in adults: a quantitative systematic review protocol. JBI Database System Rev Implement Rep 2018 Jun;16(6):1367-1372. [doi: 10.11124/JBISRIR-2017-003477] [Medline: 29894405]
59. Elliott R, Slatick E, Urman M. Qualitative change process research on psychotherapy: alternative strategies. Psychol Beitr 2001;43(3):69-111 [FREE Full text]
60. Rädiker S, Kuckartz U. Analyse Qualitativer Daten Mit MAXQDA. Heidelberg, Germany: Springer VS; 2019.
61. van Loenen I, Scholten W, Muntingh A, Smit J, Batelaan N. The effectiveness of virtual reality exposure-based cognitive behavioral therapy for severe anxiety disorders, obsessive-compulsive disorder, and posttraumatic stress disorder: meta-analysis. J Med Internet Res 2022 Feb 10;24(2):e26736 [FREE Full text] [doi: 10.2196/26736] [Medline: 35142632]
62. Shin B, Oh J, Kim BH, Kim HE, Kim H, Kim S, et al. Effectiveness of self-guided virtual reality-based cognitive behavioral therapy for panic disorder: randomized controlled trial. JMIR Ment Health 2021 Nov 22;8(11):e30590. [doi: 10.2196/30590] [Medline: 34813486]
63. Meyerbroeker K, Morina N, Kerkhof GA, Emmelkamp PM. Virtual reality exposure therapy does not provide any additional value in agoraphobic patients: a randomized controlled trial. Psychother Psychosom 2013;82(3):170-176. [doi: 10.1159/000342715] [Medline: 23548832]
64. Adema B, Thams F, Angern J. Wirksamkeit von invirto – die therapie gegen angst. Invirto. 2022. URL: https://web.invirto.de/documents/studie/weitere-informationen-studienergebnisse-invirto.pdf [accessed 2025-04-13]
65. Donker T, Griffiths KM, Cuijpers P, Christensen H. Psychoeducation for depression, anxiety and psychological distress: a meta-analysis. BMC Med 2009 Dec 16;7:79 [FREE Full text] [doi: 10.1186/1741-7015-7-79] [Medline: 20015347]
66. Keithly LJ, Samples SJ, Strupp HH. Patient motivation as a predictor of process and outcome in psychotherapy. Psychother Psychosom 1980;33(1-2):87-97. [doi: 10.1159/000287417] [Medline: 7375631]
67. Weightman M. Digital psychotherapy as an effective and timely treatment option for depression and anxiety disorders: implications for rural and remote practice. J Int Med Res 2020 Jun;48(6):300060520928686 [FREE Full text] [doi: 10.1177/0300060520928686] [Medline: 32527170]
68. Ryan RM, Deci EL. A self-determination theory approach to psychotherapy: the motivational basis for effective change. Can Psychol 2008;49(3):186-193. [doi: 10.1037/a0012753]
69. Vo V, Auroy L, Sarradon-Eck A. Patients' perceptions of mHealth apps: meta-ethnographic review of qualitative studies. JMIR Mhealth Uhealth 2019 Jul 10;7(7):e13817 [FREE Full text] [doi: 10.2196/13817] [Medline: 31293246]
70. Davis DM, Hayes JA. What are the benefits of mindfulness? A practice review of psychotherapy-related research. Psychotherapy (Chic) 2011 Jun;48(2):198-208. [doi: 10.1037/a0022062] [Medline: 21639664]

XSL•FO
RenderX

71. Ley R. Panic attacks during relaxation and relaxation-induced anxiety: a hyperventilation interpretation. J Behav Ther Exp Psychiatry 1988 Dec;19(4):253-259. [doi: 10.1016/0005-7916(88)90054-7] [Medline: 3148637]
72. Roesmann K, Leehr EJ, Böhnlein J, Gathmann B, Herrmann MJ, Junghöfer M, et al. Mechanisms of action underlying virtual reality exposure treatment in spider phobia: pivotal role of within-session fear reduction. J Anxiety Disord 2023 Dec;100:102790. [doi: 10.1016/j.janxdis.2023.102790] [Medline: 37879242]
73. Theurer C, Wilz G, Lechner-Meichsner F. Clients' and therapists' experiences of five general change mechanisms during an internet-based cognitive behavioral intervention for family caregivers. J Clin Psychol 2021 Dec;77(12):2798-2816. [doi: 10.1002/jclp.23253] [Medline: 34599844]
74. Böhnlein J, Altegoer L, Muck NK, Roesmann K, Redlich R, Dannlowski U, et al. Factors influencing the success of exposure therapy for specific phobia: a systematic review. Neurosci Biobehav Rev 2020 Jan;108:796-820. [doi: 10.1016/j.neubiorev.2019.12.009] [Medline: 31830494]
75. Leehr EJ, Roesmann K, Böhnlein J, Dannlowski U, Gathmann B, Herrmann MJ, et al. Clinical predictors of treatment response towards exposure therapy in virtuo in spider phobia: a machine learning and external cross-validation approach. J Anxiety Disord 2021 Oct;83:102448. [doi: 10.1016/j.janxdis.2021.102448] [Medline: 34298236]
76. Hildebrand AS, Planert J, Machulska A, Margraf LM, Roesmann K, Klucken T. Exploring psychotherapists' attitudes on internet- and mobile-based interventions in Germany: thematic analysis. JMIR Form Res 2024 Nov 07;8:e51832 [FREE Full text] [doi: 10.2196/51832] [Medline: 39510514]
77. Rabe-Menssen C, Ruh M, Dazer A. Die versorgungssituation seit der reform der psychotherapie-richtlinie 2017. Deutsche PsychotherapeutenVereinigung. 2019. URL: https://www.dptv.de/fileadmin/Redaktion/Bilder_und_Dokumente/Wissensdatenbank_oeffentlich/DPtV_Hintergrund/DPtV-Hintergrund-2019-1.pdf [accessed 2025-04-13]

Abbreviations

AG: agoraphobia
CBT: cognitive behavioral therapy
IMI: internet- and mobile-based intervention
PD: panic disorder
RCT: randomized controlled trial
SAD: social anxiety disorder
VR: virtual reality
VRET: virtual reality exposure therapy

Edited by A Kushniruk; submitted 27.05.24; peer-reviewed by M Franzen, K Schmidt; comments to author 24.11.24; revised version received 18.01.25; accepted 29.01.25; published 24.04.25.

Please cite as:
Planert J, Hildebrand AS, Machulska A, Roesmann K, Neubert M, Pilgramm S, Pilgramm J, Klucken T
Blended Mobile-Based Interventions With Integrated Virtual Reality Exposure Therapy for Anxiety Disorders: Thematic Analysis of Patient Perspectives
JMIR Hum Factors 2025;12:e60957
URL: https://humanfactors.jmir.org/2025/1/e60957
doi:10.2196/60957
PMID:

©Jari Planert, Anne-Sophie Hildebrand, Alla Machulska, Kati Roesmann, Marie Neubert, Sebastian Pilgramm, Juliane Pilgramm, Tim Klucken. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 24.04.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Using Gesture and Speech to Control Surgical Lighting Systems: Mixed Methods Study

Nima Zargham[1], Dr Ing; Anke V Reinschluessel[2], Dr Ing; Andre Mühlenbrock[1]; Thomas Muender[1], Dr Ing; Timur Cetin[3]; Verena Nicole Uslar[3], Dr rer nat; Dirk Weyhe[4], Prof Dr Med; Rainer Malaka[1], Prof Dr; Tanja Döring[1], Prof Dr

[1]Faculty of Mathematics and Computer Science, University of Bremen, Bremen, Germany

[2]Department of Computer and Information Science, University of Konstanz, Konstanz, Germany

[3]Department of Human Medicine, University of Oldenburg, Oldenburg, Germany

[4]Pius Hospital Oldenburg, Oldenburg, Germany

Corresponding Author:
Nima Zargham, Dr Ing
Faculty of Mathematics and Computer Science
University of Bremen
Bibliothekstr. 5
Bremen, 28359
Germany
Phone: 49 42121864401
Email: zargham@uni-bremen.de

Abstract

Background: Surgical lighting systems (SLSs) provide optimal lighting conditions for operating room personnel. Current systems are mainly adjusted by hand; surgeons either accommodate the light themselves or communicate their requirements to an assistant to ensure optimal surgical conditions. This poses challenges to maintaining sterility, proper accessibility, and illumination and can lead to potential collision problems. Furthermore, the personnel operating the light may not have deep medical knowledge or equipment expertise.

Objective: This paper introduces a touch-free interaction concept for controlling an SLS using speech and gestures.

Methods: We used an iterative, user-centered design approach with participatory design sessions. This process involved conducting a literature review, several observations of actual surgical sites, and engaging stakeholders through interviews and focus groups. In addition, we carried out 2 user studies: one in a virtual reality setup and another in a living laboratory environment.

Results: Our findings indicate that our interaction concept is a viable alternative for controlling an SLS. Despite some technical limitations, surgical experts found the system intuitive and useful, recognizing the significant potential for touch-free lighting adjustments in the operating room. The combination of speech and gesture modalities was seen as helpful and even necessary, with some interactions better suited to one modality over the other. Offering both modalities for each interaction provided greater flexibility.

Conclusions: Our findings suggest that our proposed touch-free interaction concept can enhance surgical conditions and has the potential to replace traditional adjustment.

(JMIR Hum Factors 2025;12:e70628) doi:10.2196/70628

KEYWORDS

speech interaction; gesture recognition; operating theater; surgical lighting systems; smart lighting; artificial intelligence

Introduction

Background

The illumination of the operating room (OR) is critical to ensuring optimal surgical conditions. Surgical lighting systems (SLSs) are designed to provide optimal lighting for the surgical staff in the OR. They provide settings for adjusting the intensity of the light, its color temperature, and the size of the illuminated field, which are mainly controlled with control panels situated on the lamp's body. Such adjustments to the SLS are necessary for various reasons, such as differentiating between only

XSL•FO
RenderX

marginally different-looking tissues [1]. Making modifications to the lighting conditions requires the surgeon or an assistant to operate the control panel with their hands, which is often located in an ergonomically unfavorable position, such as above the surgeon's head [2].

Generally, an OR is required to have two sets of distinctive lighting types: (1) general room lights (ambient lights) and (2) focused and precise procedural surgical lights to illuminate areas of incision (surgical lights) [3]. Optimal surgical illumination combines sufficient ambient light with the ability to apply focused light at specific positions and angles throughout different stages of an operation [4]. Current SLSs are primarily controlled and adjusted by hand, and surgeons are eager to operate the SLSs themselves while keeping sterility. It is essential to look for an ergonomic solution for the SLS adjustments to maintain the surgical staff's focus on their primary task and keep distractions to a minimum. Hence, controlling the SLS should be intuitive and straightforward to keep the mental load to a minimum and increase interaction safety [1]. Moreover, during an operation, the surgical staff must reposition the SLS several times to ensure optimal surgical illumination. Research shows that every 7.5 minutes, a change in the lighting position takes place [2]. Such position changes are also commonly performed by hand. However, performing such adjustments by hand can lead to collisions and other mechanical problems. In addition, this raises concerns regarding sterility and increases potential safety risks [5]. Even though it is required to sterilize all equipment the surgical staff comes in contact with during an operation, this still might not be done sufficiently due to the lack of information and attention of the OR personnel, use of the wrong sterilization product, or insufficient and ineffective work of sterilization [5]. During an operation, surgeons often need to interact with the OR equipment, such as radiological images and lights [6]. Due to sterility concerns, surgeons may be unable to interact with such equipment and depend on their assistants. Communicating this need with the personnel might be complex and lead to further errors if the assistant and the surgeon cannot communicate properly. Consequently, there is a possibility of interruptions in the workflow.

In this work, we propose the use of natural user interfaces (NUIs) [7,8] to provide surgeons with an intuitive interface for adjusting the SLS. NUIs have the potential to enable clinicians to have freehand control of the OR equipment while maintaining sterility. Data from several studies highlight the potential and advantages of using NUIs in ORs [1,6,9-13]. This paper presents an interaction concept that combines voice and gesture to control the SLS. We used a touch-free approach to protect sterility in the OR and ultimately increase patient safety. To develop this interaction concept, we used an iterative user-centered design approach with participatory design sessions. First, we conducted a literature review and had several observations of actual surgical sites to develop the initial list of interactions. Using this list, through co-design, we discussed individual interactions for both modalities of speech and hand gestures with 7 experts using interviews and developed an initial interaction concept. Subsequently, we conducted 2 focus groups to discuss the concept with the OR personnel. The revised interaction concept based on the focus groups' feedback was then used in a virtual reality (VR) prototype and evaluated by 6 experts. Considering the results of the VR evaluation, we developed a prototype in a living laboratory and conducted a laboratory study with 12 surgeons and surgical assistants. Our findings suggest that our interaction concept is a valid alternative for controlling the SLS. Even though certain technical limitations were experienced, surgical experts found the system highly useful and intuitive and saw great potential for touch-free approaches to lighting adjustments in the OR. This paper contributes the following: (1) the design and implementation of a touch-free interaction system for controlling the SLS informed by domain experts and (2) 2 user studies providing an exploratory evaluation of how this system can support surgical personnel. The implications of our work can support researchers and medical companies when adapting touch-free approaches for adjusting SLS. It further gives an understanding of the potential and challenges of such approaches. The broader insights of this research can also be valuable for the general use of NUIs in the OR to enhance surgical conditions and increase patient safety.

All figures used in the manuscript can be viewed in Multimedia Appendix 1.

Related Work

Many OR personnel believe that lighting is a major inconvenience [14]. Previous literature identified the following types of problems with SLS: *mechanical*, *collision*, *accessibility*, and *illumination* problems [2,15,16]. *Mechanical* problems include the requirement of excessive force to move the SLS, which can make a 1-hand light adjustment difficult and occasionally cause the SLS to get jammed. In such cases, surgeons usually ask for the help of other OR personnel. This can delay the completion of the adjustment and could lead to possible errors because the personnel responsible might not have the necessary medical knowledge or familiarity with the equipment [2]. *Collision* problems refer to the times when the SLS collides with other OR equipment or against the heads of the personnel [2]. *Accessibility* problems occur when surgeons have to stand up or move to perform an adjustment because the SLS is out of reach [2]. *Illumination* problems refer to the issues that the surgeons or the surgical staff experience regarding the SLS lighting conditions, such as wound illumination, refocusing the light beam, and increasing illumination levels [15].

In the following sections, we provide an overview of the previous literature on smart surgical light control and touch-free techniques for an OR, including voice and gesture interaction in the OR.

Smart Light Control in the OR

Appropriate lighting during surgery is critical for patient safety [17]. Researchers have looked into various methods to enhance the lighting conditions in the OR and the interaction with the SLS. Teuber et al [18] presented a method for autonomous positioning of surgical lamps during open surgeries using a depth camera and robotic arms. The authors tested their algorithm in a VR simulation and found that their method was robust and could ensure close-to-optimal lighting conditions in real-world surgeries. Burger [15] developed a technique for

XSL•FO
RenderX

autonomous, real-time adjustment of an SLS by identifying the wound based on heat sources to achieve complete automation of an SLS in the OR. The findings suggested that the combination of optical and thermal cameras with stereo image-processing techniques could be used to identify and track a heat source. Moreover, predefined scene lighting has been commonly examined and used for medical environments, such as angiography rooms [3]. This is accomplished by defining specific lighting features, such as the light position, intensity, angle, and color. Mühlenbrock et al [19,20] introduced a novel lighting system consisting of numerous small lighting modules permanently mounted on the ceiling. This system actively avoids casting shadows using depth sensors. Through simulations and a subsequently constructed prototype, which was evaluated with surgeons, it was demonstrated that in open abdominal surgeries, this system can illuminate a surgical wound more effectively than conventional surgical lights and was deemed more suitable for use in the operating theater by surgeons. In addition, point cloud recordings were used to investigate how the positioning and optimization of these light modules affect the achievable brightness at the surgical site [21]. Other studies have shown the potential of freehand control for operating the OR lights. Hartmann and Schlaefer [9] used an RGBD (red, green, and blue, plus depth) camera for gesture tracking to study gesture-based control of OR lights. Their results showed that gesture control could be quickly learned and effectively used to operate OR lights. Similarly, to offer the surgeons a sterile and ergonomically comfortable operation, Dietz et al [1] developed a freehand gesture control using a Kinect (Microsoft Corporation) system for the SLS. Participants reported high usability scores for the system and rated it ergonomically more comfortable than the conventional control panels and less likely to distract them from their tasks. In this work, we build on the lighting modules introduced by Mühlenbrock et al [20] in our living laboratory study, implementing a freehand control method to adjust the lighting conditions in the OR.

Touch-Free Techniques in the OR

An extensive body of literature has previously looked into touch-free methods for controlling OR equipment. As input modalities, gaze [22-24], gestures [12,22,25-33], or voice [11,34-36], for example, were researched. Hatscher et al [22] used gaze tracking and foot gestures on an interactive floor as input modalities for hands-free interaction to interact with medical image data. Their results show that selection is accomplished faster via gaze than with a foot-only approach, but gaze and foot easily interfere when used simultaneously. To minimize the dependency on OR personnel, Zaman et al [35] investigated voice- and foot-based interactions for the surgeon to interact with 2D images in VR. Their findings suggest that both modalities have acceptable usability. While the voice command system was perceived as more comfortable, foot-based interaction was more efficient.

The use of voice technology in health care is growing. Research has shown that voice-based commands could lower the cognitive and physical cost of accessing specific features or performing certain actions [37]. Feedback from experts revealed that surgeons have previously requested more voice-controlled features in ORs, including voice commands for adjusting lamp positions [14]. Often, in critical situations during an operation, the surgeon and the assistants must use both hands to ascertain the patient's health [38]. In this context, the advantage of voice interaction is that the surgeons can concentrate on the operation while simultaneously controlling other components using their voice [38,39]. The most crucial point is that the surgeon's hands are always unencumbered. Previous research has looked into the use of voice interaction in an OR [36,40-43]. El-Shallaly et al [40] evaluated a voice interface regarding the use of time and surgical staff during laparoscopic cholecystectomy. Their results showed that the voice interface significantly optimized the operating time and the use of surgical staff. In a comparison study, Alapetite et al [39] compared a voice interaction system with the traditional touchscreen and keyboard interface to update anesthesia records during crisis situations. Their results showed an accurate speech-based registration performance even during emergencies and time-critical scenarios. Punt et al [44] compared the efficiency, reliability, and user satisfaction of controlling an endoscope's zoom and light intensity using voice control, a touch panel, or manual control by an assistant. Most participants experienced voice control as the quickest interface, although in reality, it was slower than the other 2 conditions. Recent advances in natural language processing and artificial intelligence have extensively enhanced the efficacy of such systems, addressing some of the most prominent concerns with these systems, such as speech recognition. These advances provide new opportunities to explore the use of voice in various domains, including health care [45]. The emergence of large language models and generative artificial intelligence technologies such as ChatGPT (OpenAI) has ushered in notable improvements in this regard [46]. Another commonly proposed modality for hand-free interaction is gesture control. Gesture-based systems have been considered a novel application to adjust medical equipment [32,33]. Previous research has looked into navigating magnetic resonance imaging and computer tomography images using arm gestures during surgery [6,13,31]. Nestorov et al [30] designed a touch-free medical image control system using Leap Motion (Leap Motion, Inc) and Kinect. Their results showed an average acceptability rate for both systems. Nonetheless, surgeons and radiologists found Kinect to have better utility and be more beneficial for most clinicians. Overall, each of these modalities has demonstrated potential in ORs, improving the working conditions for surgeons. Researchers have also investigated the combination of different interaction modalities. Data from several studies suggest that incorporating voice and gesture inputs can be helpful in the OR [10,29,42]. To operate a system without the risk of equipment contamination, Ebert et al [10] developed a system prototype that allows for touch-free control of medical images using a depth camera and voice recognition software. The authors found that the touch-free system performs favorably and removes a potential vector for infection, protecting both patients and staff. Similarly, Hötker et al [42] demonstrated the feasibility and utility of using speech and gesture for reviewing images in interventional radiology to navigate through image data in a sterile environment. Their results showed high-command recognition rates and stability under different lighting conditions. In a study investigating the uses of voice control versus gestural control in the OR, Mentis et al [11]

XSL•FO
RenderX

designed a system that allows for both modalities at the user's choice.

Authors suggest that the benefits of each technology are circumstantial and recommend enabling functionalities to be achieved using both modalities to have a more flexible combination of their benefits as the context of the procedure requires [11]. Previous studies have highlighted that voice interaction is better fitted for trigger functions or functionality switches, whereas hand or body gestures are more suitable for continuous manipulation of parameters [11,38]. Although alternative methods for controlling the OR equipment have been largely investigated, to our knowledge, using speech and gesture to operate the SLS has yet to be evaluated in this context. Moreover, previous literature still lacks an understanding of the experts' expectations and perceptions of such systems. Building on the previous work, in this paper, we engage with OR personnel to design a touch-free interaction concept for controlling the SLS in an iterative user-centered approach.

Methods 1: Initial Concept

Overview

In this work, we sought to develop and evaluate a touch-free interaction concept for controlling an SLS. For this, we used an iterative, user-centered approach to address the needs and requirements of the OR personnel. We could not directly translate existing touch-free practices to controlling an SLS, as the use case in the OR is unique due to its constraints (eg, patient safety or sterility aspects) and requires extra attention. Therefore, we chose a human-centered approach to best integrate and fulfill the needs of the OR personnel. The iterative process allowed us to enhance the concept by reevaluating it several times. This approach also helped us identify possible weaknesses and issues at the early stages. Furthermore, it enabled us to fill the gap between our intended design and how the experts interpret it [47,48]. We first sat in on an actual surgery to observe how surgical personnel interact with the SLS; on the basis of the observations and related work, we curated a list of required SLS interactions, which were discussed with experts and further refined using a focus group.

Assessing OR Light Adjustments

To come up with our initial concept, we first reviewed the existing literature to explore previous work on contactless approaches for controlling surgical equipment [1,10,29,38,42]. A short summary of our review is mentioned in the Related Work section of this paper. Furthermore, to better understand how the OR personnel, especially the lead surgeon, interact with the current surgical lights, we attended an open liver surgery. The main goal was to identify what type of interactions occur with the lights, which personnel modify the SLS, how often it is adjusted, and how the adjustments are made throughout different phases of an operation. In total, 7 surgical personnel were involved in this surgery. The operation took around 81 minutes. For this surgery, 2 satellite lamps were used as focused lights to illuminate the wound. During this time, 21 light adjustments were made (6 major and 15 minor), that is, approximately 1 adjustment for every 4 minutes. Major light adjustments involve significant changes to overall illumination or focus to accommodate procedural needs, while minor adjustments are precise tweaks to optimize visibility without altering the general lighting setup. We observed that all the light adjustments were made using a single hand during the surgery. The reasons for light adjustments were mainly due to the lead surgeon or assistants changing poses, using other OR equipment, such as the ultrasound device, or changing the light's focus. Only around 35% of adjustments took the shortest route to get to the next position. The lead surgeon only adjusted the SLS 20% of the time. Three other assistant surgeons made the remaining adjustments. We observed that light adjustments led to interruptions and interference with the personnel's task at hand. For instance, the lead surgeon had to stop what they were doing to adjust the light and then continue with the work. Considering our literature research and observation of the surgery, we developed a list of possible interactions to perform with the SLS, which includes the *light switch*, *light beam position*, *size of the illuminated field*, *light intensity*, and *predefined lights*. On the basis of the recommendations from the previous research, we chose the 2 modalities of *speech* and *gesture* to use for our touchless interaction concept.

Expert Interviews

After developing a list of possible interactions with the light, we conducted 7 interview sessions with the OR personnel to better match our concept to their needs and discuss our modalities and individual interactions, with the aim of designing an initial interaction concept.

Procedure

All interview sessions were held remotely via video calls while the experimenter recorded verbal statements and observations. The participants were asked to give informed consent and fill in the demographics questionnaire before the session. Participants were briefed about the interview procedure at the beginning of each session. We explained our concept and research goals, including our intention to use voice and hand gestures to control the SLS. We first asked about the experts' impression of the concept when the interview started. After that, they provided feedback about each interaction and its positive or negative aspects. We then asked for their suggestions to enhance the current concept. For further qualitative analysis, we audio-recorded the sessions. The interview sessions took between 35 and 45 minutes. The interview recordings were summarized, then analyzed and coded based on domain summaries [49,50], where the themes are structured around a shared topic rather than shared meaning, with the goal of capturing the diversity of meaning in relation to a specific subject or area of focus [51]. The transcripts of the interviews were coded by a researcher using inductive coding [52,53], where a single quote could be assigned to multiple codes. In addition, we collected insightful and unique statements.

Participants

We recruited 7 senior physicians (1 female and 6 male) who were aged between 41 and 58 (mean 47, SD 6.16) years, all with at least 17 (mean 22, SD 5.12) years of medical experience and active roles in surgeries. Moreover, 4 were general and

XSL•FO
RenderX

visceral surgery specialists, 1 was a neurosurgeon, 1 was a urological surgeon, and 1 was an ear-nose-throat specialist.

Ethical Considerations

Ethics approval was received for every step of our evaluation from the University of Oldenburg ethics board. For each step, participants gave informed consent before participation. All data collected during the study were anonymized following the sessions to ensure participant confidentiality and privacy. No personally identifiable information was stored or retained. Participation in the study was entirely voluntary, and no compensation was provided.

Results 1

Overview

The experts generally liked the idea of using voice and gesture to control the SLS. They found it beneficial to have more control over the lights and avoid having additional personnel to make such changes. One expert said, “It is exhausting to ask an assistant to control the light. It takes extra effort” (P1). Another mentioned, “Everything you have to ask somebody else to do for you should have smart control” (P4). Furthermore, one expert pointed out that using such a concept can reduce unnecessary movement in the OR, for example, moving around for activities such as regulating or switching the lights. All participants agreed that multiple personnel should be able to adjust the SLS. This includes the anesthesiologist, the instrument nurse, and the assistant surgeons. One participant suggested that restricting access to certain personnel in special cases could be helpful. One participant suggested designing the interactions based on what is commonly used in ORs, such as similar sentences or phrases that are now used as an assistance request, “Easiest way is always to follow what you have done before” (P3). Two other experts recommended designing familiar gestures based on real-world interactions. For our experts, minor position changes were the most critical interactions often used during surgeries. Four participants mentioned that major position adjustments are less often performed in surgeries than fine adjustments, which matches our observations. Regarding the light intensity, 4 found it helpful to adjust the light intensity intraoperatively. On the other hand, 3 found it unnecessary to adjust the brightness of the SLS. In addition, 6 out of the 7 participants agreed that using predefined lights is beneficial. Participants found it helpful to record specific light characteristics to be able to reload them again afterward. One expert said, “It is an absolutely useful feature to save specific light positions and go back to it later” (P2). Another participant mentioned, “It would be beneficial if we could define certain positions based on the positions of the organs and just load them” (P3). One expert judged predefined lighting conditions to be an unnecessary feature. Furthermore, 3 participants suggested changing the light color to be helpful during the surgery. One surgeon highlighted that using the color green during laparoscopic surgery is quite beneficial, and it would be helpful to make this change with the new concept. Participants also pointed out possible concerns regarding each modality. Although all participants gave highly positive feedback on the concept, some were uncertain about the efficiency of the technology. Five participants were concerned about the noise level in the OR and whether voice input could be used in such an environment. Participants were worried about the recognition of the gestures and speech commands. Two mentioned that the unrecognized commands could frustrate the personnel and add more stress. Four participants were worried about misrecognition and false positives where the OR personnel give input, but the system does something otherwise. Moreover, participants raised concerns about having multiple hands near the wound, which might be problematic for gesture recognition. In addition, 6 out of the 7 participants suggested using a wake word for the voice input to minimize false recognition by the system. Four experts emphasized the importance of choosing a suitable wake word, suggesting using words or phrases not commonly used in the OR. A similar initiation step was also recommended for the gesture interaction to avoid false recognition. Moreover, 2 participants pointed out the importance of considering lighting scenarios that different surgical disciplines can use. Experts mentioned that designing smart lighting systems may not suit every type of surgery. For instance, the positions are typically predefined in spinal surgeries, and there is little variation. Moreover, in some instances, the surgeons use different lighting sources, such as headlights.

Initial Concept

Overview

On the basis of the discussions and findings from the expert interviews, previous literature, and technical considerations, we developed an initial interaction concept (Table 1).

Table 1. Interactions in the initial concept with their respective modality and action to perform the interaction.

Interaction	Modality	Action
System activation	Speech	“Activate”
Light switch	Speech	“Surgical lights on”
Light beam position (minor)	Gesture	Move hand for positioning
Light beam position (major)	Speech	“Position three”
Light intensity	Speech	“Intensity 75 percent”
Size of the illuminated field	Gesture	Distance between the two hands
Predefined positions	Speech	“Standard mode”
Previous position	Speech	“Previous position”

XSL•FO
RenderX

System Activation

To interact with the system, the system first needs to be activated using a wake word. This activation is done via speech using the command "activate." After using the wake word, there is a 5-second window to interact with the system. The system returns to standby mode when an action is performed, awaiting the next wake word. If no action happens, the system returns to standby mode after 5 seconds.

Light Switch

Switching the lights on and off is done using speech. Users can turn the SLS on by saying, "surgical lights on!" and, similarly, switch it off by saying, "surgical lights off!"

Light Position

The major position changes are done using speech interaction. The possibility is to either use numbers for each major position (eg, "position two") or specific names (eg, "position stomach"). The surgical team can add custom positions with their desired names for that position. Fine (ie, minor) adjustments are performed with a combination of speech and gestures. After the wake word, users must use the command "edit position!" The system will then follow the personnel's hand. After finding the right spot, the position can be set by saying "save position!" The position is saved afterward, and the hands are no longer tracked as the system returns to standby mode. The personnel can always command the system to return to the previous position by saying, "previous position."

Light Intensity

The intensity of light is also adjusted by speech. Users can use percentages to modify the intensity, for example, "intensity 75 percent."

Predefined Settings and Modes

Different predefined illumination modes can be set where specific lighting conditions are saved for multiuse purposes. For instance, in dark mode, the lights are off, and the personnel can better focus on the information from the monitors (eg, the ultrasound device). The modes are set using speech, where users can give commands such as "standard mode," "dark mode," or "laparoscopic mode" (which turns the light green).

Field of Illumination

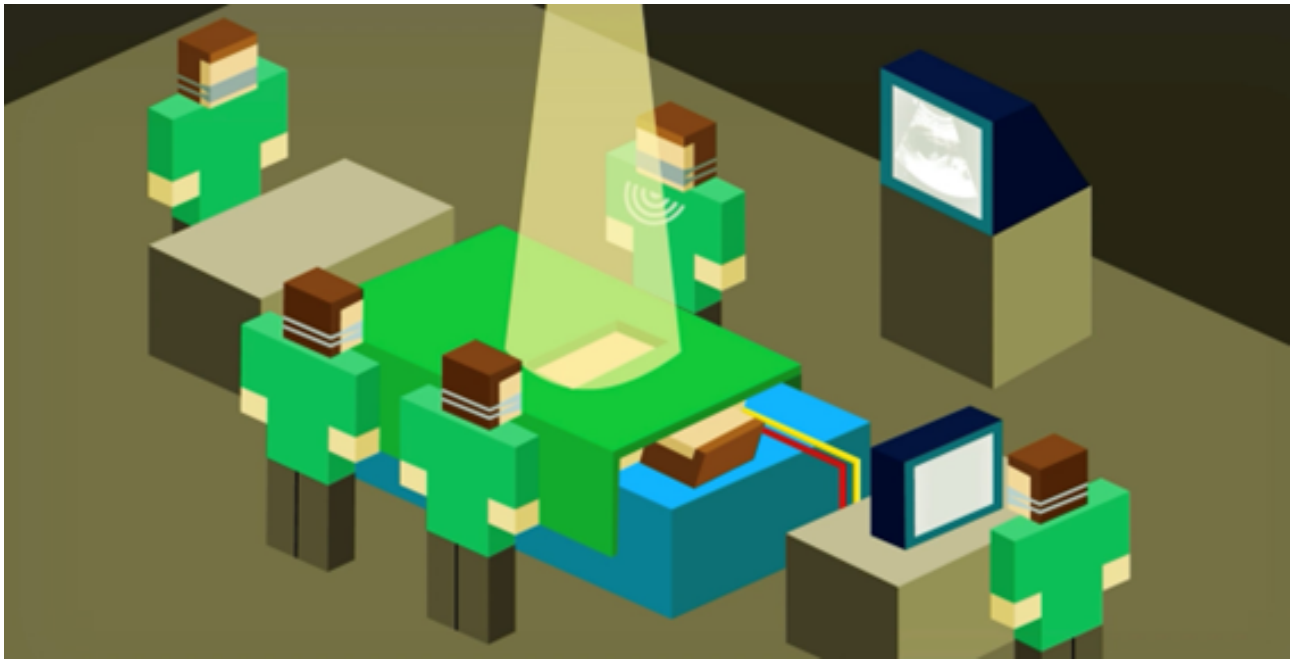
To adjust the size of the illuminated field, we chose a combination of hand gestures and speech. After system activation, users must use the command "edit angle." The system would then track the surgeon's hands and calculate the distance between them. The more distant they are, the wider the angle, and vice versa. After finding the right size, the position can be set by saying "save." The position is saved afterward, and the system returns to standby mode.

Methods 2: Focus Groups

Overview

We conducted 2 focus groups with the surgical personnel to evaluate and discuss our initial interaction concept. We used an approach similar to scenario-based design methods [54] and vignette experiments [55]. We developed animated videos demonstrating the interactions based on the concept (Figure 1). In these videos, we visualized how each interaction could be done in the OR using the previously defined and described input modalities. This allowed us to investigate the interaction and technologies without worrying about technological concerns. The main goal for the focus group sessions was to design the input format for each interaction, including individual voice commands and different hand gestures. Furthermore, we wanted to discuss the appropriate modality for individual interaction.

Figure 1. A frame from the concept animation that was used in the focus group sessions to help visualize the interactions for the experts.



XSL•FO
RenderX

Procedure

We conducted 2 focus groups, inviting different target groups. Both sessions were conducted in a hybrid format. The first focus group (FG1) consisted of 6 surgeons (2 female surgeons and 4 male surgeons), all present in a hospital meeting room, with 2 researchers assisting them in person (Figure 2). Two other researchers joined in via video call. The second focus group (FG2) was conducted with 4 surgical assistants (one female surgical assistant and 3 male surgical assistants). Similar to the first session, all participants and 3 researchers were present in a hospital meeting room, and 1 researcher joined via video call. We grouped surgical assistants separately from surgeons to allow for a comfortable reflection without the potential influence of the hierarchical structure within surgical teams. Both sessions were video recorded for further analysis. Participants gave informed consent and filled in a demographics questionnaire at the beginning of each session, which was followed by a brief overview of the session's procedure. Our interaction method was then presented along with the chosen modalities as well as the primary research goals. Next, individual interactions and possible input forms for each were discussed. The participants then acted on these suggestions in Wizard-of-Oz style, where they would make hand gestures or say the speech commands and the researchers would modify the lamp to simulate the interactions so that the participants could experience their suggested interactions better. Afterward, we demonstrated our animated concept video and collected the participants' feedback. We discussed the animated videos after the initial discussion about the interactions to avoid biasing participants with our chosen interaction set and modalities. In the end, we left the floor open for further comments or suggestions from the experts. The first session took approximately 90 minutes (FG1), and the second took approximately 80 minutes (FG2). On top of the video recordings in each session, 2 researchers noted verbal statements and their own observations. The researchers reviewed, analyzed, and summarized the session recordings of both focus groups.

Figure 2. A surgeon acting out the interactions in the first focus group session.



Results 2

Overview

The focus group results generally mirrored those obtained from interviewing experts. The focus groups led to further insights regarding participants' concerns with the system and how such systems should be designed. Participants again pointed out that the interactions should be familiar to the personnel: "It would be nice to have a similar approach like the current way of moving the lamps" (P2 in FG1). All participants agreed that multiple personnel should be able to access the system. One participant mentioned, "When both surgeon's hands are occupied, an assistant should be able to control the light" (P5 in FG2). In both focus groups, experts agreed that the interactions should be kept as simple and short as possible. Participants believed it would be valuable if the new interaction concept could help shorten the operating time.

Appropriate Modality

Three participants agreed that speech would be more practical than gestures overall, as they believed gestures could be more distracting and lead to collisions. Participants agreed that speech would be more suitable for switching the lights on or off. All

XSL•FO
RenderX

recommended that hand gestures should be used to perform fine adjustments. Regarding brightness and field of illumination, there were mixed feelings. Two participants believed speech would be more appropriate, while others suggested using gestures. The feedback about adjusting the size of the illuminated field varied among participants. Some believed this was best performed by speech using some sort of a scale, such as percentages, while others thought it would be more accurate using hand gestures. Two participants highlighted that the hand gestures should all be made using 1 hand. They believed that occupying both hands to adjust the SLS could delay the operation. The participants of both focus groups mainly agreed that position changes should be done with hand gestures, as they seem more precise. It was difficult for the participants to imagine speech commands that could accurately and quickly adjust the position of the light's focus.

The general agreement in the sessions regarding the voice commands was that they needed to be as short as possible. Experts believed that more extended commands would be harder to remember. More importantly, brief commands would save time, a critical aspect of surgeries. Using phrases not commonly used in the OR was recommended to avoid false recognition. There was uncertainty about the language used for the speech commands. Three participants believed the commands should be given in the local language (German), as it is easier to learn, remember, and correctly pronounce them. Others suggested using a language different from the local one, such as English, because it is not commonly spoken in the OR. It was recommended that the system should be flexible and able to handle multiple phrases per action, meaning multiple commands should be able to perform the same interaction.

Individual Interactions

We received multiple suggestions for a wake word, including "sun," "sunlight," "OP light," and "light." Among these, "light" was favored by most as it is short and easy to remember. Moreover, it was recommended that an activation step be used for hand gestures to avoid false positives. One participant suggested holding a gesture for a specific amount of time before giving the actual gesture. One participant recommended simulating pressing a light switch with hand gestures to switch the lights on and off. Another suggested opening and closing a fist to switch the lights. Participants mainly agreed that "lights on" or "lights off" are appropriate commands for speech regarding the light switch. For positioning the light beam, 2 participants tried to grab the beam of light and move it to a different position. Others in the group also found this form of interaction natural. Participants had difficulty imagining an appropriate speech command for positioning. One participant suggested using distance and a direction (eg, "20 centimeters left"), but others found this complex and time-consuming. Participants mainly agreed that having the possibility to go back to the previous position could be beneficial during operations. Regarding light intensity, one of the participants recommended using percentages. Another suggested using numbers from 1 to 10. For gesture interaction, a participant suggested moving a fist up and down to adjust the light intensity, similar to dragging a slider. Similar to the light intensity, numeric commands were recommended for speech input to adjust the illumination field. For gestures, 1 participant suggested using both hands in an L-shaped form (fully opened thumb and index finger, while other fingers are closed) and moving them closer or further away to make adjustments. The group also favored this, although it did not follow the one-hand-in-use recommendation. Predefined light settings were highly favored by the participants. All participants agreed that such a feature could be valuable for the OR and could help save time. One participant said, "We can use a specific setup and call it 'scenario 1', and then just say 'scenario 1' if we want to use this setting" (P3 in FG1).

Feedback on Animated Videos

Participants commented on the initial interaction concept based on the animated video. They generally liked the interactions and found them intuitive. Experts found some speech commands long and highlighted that they must be quick and short. Participants found the wake word sufficient but suggested using their recommended phrases instead. They found it appropriate to switch the lights with speech. Similarly, participants believed that modifying the light's position with hand gestures is intuitive and would be challenging with speech. Nevertheless, they found it helpful to use speech for predefined positions, including the previous position. It was mentioned that hand gestures used for the concept are easy to remember. Two participants recommended adjusting the system with both modalities to give the personnel more options and to allow them to choose the best based on circumstances.

Concerns

Similar to the interviews, experts had concerns regarding the reliability of the technology. They were skeptical about the accuracy of recognition of both speech and gesture. One expert mentioned, "It is critical to lower the stress level in the OR. At the moment, with this technology, I feel a bit stressed." (P1 in FG1). Participants were worried about the use of speech in the OR due to the high noise level. They repeatedly noted that surgeons may be unable to adjust the lamps as accurately as the current lights. Furthermore, 2 participants were concerned about the additional cognitive load on the surgical personnel caused by complicated interactions. They were worried that the surgical personnel might not remember the commands quickly enough in a stressful situation. Three participants suggested having a backup control for such situations. Regarding speech, 1 participant was skeptical about recognizing commands with different dialects.

Adjustments to the Interaction List

Considering the insights of the interviews and focus groups, we refined the interaction concept. Regarding speech commands, we shortened some of the commands and changed the wake word based on the expert recommendations to "Hey, light!" Because "light" was the favorite suggestion during focus groups, we adjusted it based on the technical recommendation by adding "hey" in the beginning to enhance the recognition and prevent false activation of the system. We also modified the predefined lights into saving and loading positions, where users could save specific lighting conditions and load them later on. To activate the gesture system, in line with recommendations from Dietz et al [1], the user could hold the gesture they will use to edit the

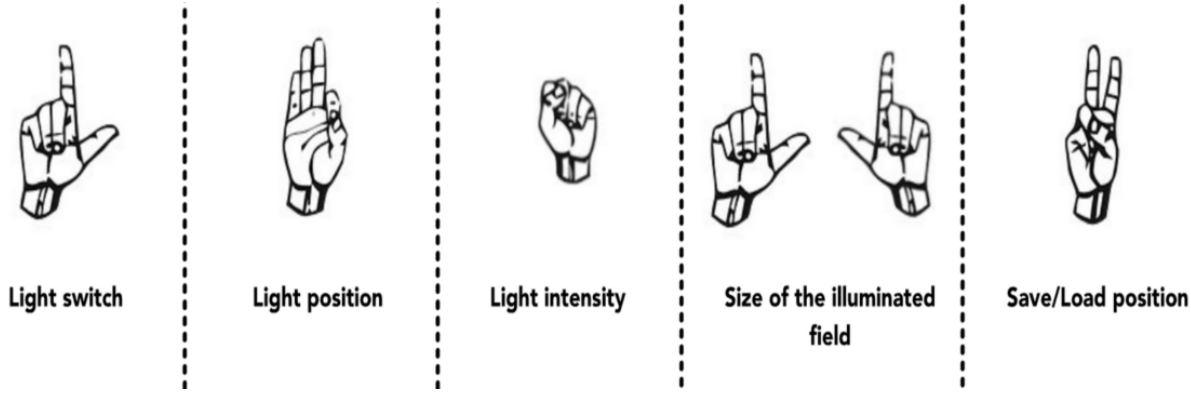
XSL•FO
RenderX

light settings for a certain period to perform an action. In this way, the probability of false positives will also be lowered, as it is less likely that surgical personnel will hold a specific gesture for a specific period. For instance, if the user wants to adjust the position of the light beam, they must hold the gesture for positioning the light first for a specific period. Once the system has been activated, the user is notified with auditory feedback that the gesture has been recognized, allowing the command (while maintaining the gesture) to be executed. We empirically set this time to 1.5 seconds after our initial testing sessions. The gesture system is deactivated when the command is executed until the next activation. Unlike Dietz et al [1], the user must only hold 1 gesture to activate and execute the command. In their approach, the activation gesture differed from the interaction gesture. To switch the lights on or off with hand gestures, users must hold an "L-shaped" hand gesture (Table 2) and move their hands in front, resembling the act of switching the lights on or off in the real world. Users must hold an "OK" gesture to position the light beam. The light would then follow the tip of their thumb and stop tracking once the hand is opened (Figure 3).

Table 2. Interactions in the virtual reality prototype for speech.

Interaction	Speech
System activation	"Hey, light!"
Light switch	"Lights on/off"
Light position	"20 centimeters left"
Light intensity	"Intensity 75"
Size of the illuminated field	"Angle 15"
Save or load positions	"Save/load position two"

Figure 3. Interactions in the virtual reality prototype for gestures.



This gesture was chosen as it would resemble grabbing the beam of light and moving it around. To modify light intensity, we implemented the "fist" gesture. Users have to move their fists up or down to increase or decrease the light intensity. We chose the "peace" gesture to save and load the positions. To save, the user has to move the hand forward and backward for loading. To choose a specific number to save or load, the hand has to be moved left to choose slot number 1 or right to choose slot number 2. For instance, if a user wants to save a specific position as position 2, they first have to move the hand forward and then to the right. This hand gesture was inspired by changing gears in a car. To change the size of the illuminated field, both hands have to be held as "L-shaped." The closer the hands would get, the smaller the size, and vice versa. The complete list of interactions can be found in Table 2.

Methods 3: VR User Study

Overview

After adjusting the interaction concept, we implemented it in a VR prototype. In this prototype, one could adjust the SLS in a virtual OR using gesture and speech interaction. Both modalities were supported for every individual interaction. We designed the interactions based on the experts' recommendations, technical considerations, and suggestions from the aforementioned literature. After developing the VR prototype, we consulted 2 surgeons and discussed the interactions that had been implemented with them before conducting the study. On the basis of their feedback, we made final adjustments to the prototype. We then tested the prototype with 2 novice users to identify and resolve further issues. For this study, we did not implement the functionality to return to the previous position to reduce study time and prioritize more essential interactions, as this was considered a minor feature.

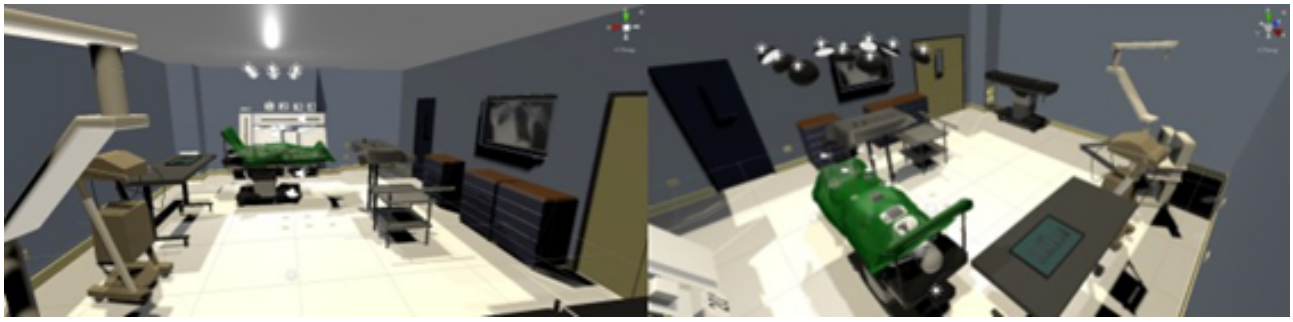
Implementation

The prototype was implemented in Unity 3D [56]. We used Valve Index [57] VR glasses to conduct the study. A virtual OR was designed to simulate a natural environment (Figure 4). We used the Picovoice Speech-to-Text library [58] for speech recognition and the Valve Index's 2 built-in cameras for gesture recognition using the hand-tracking SDK by HTC Vive [59].

XSL•FO
RenderX

A neural network classified the detected hands after matching features had been computed.

Figure 4. The virtual operating room used for the virtual reality study.

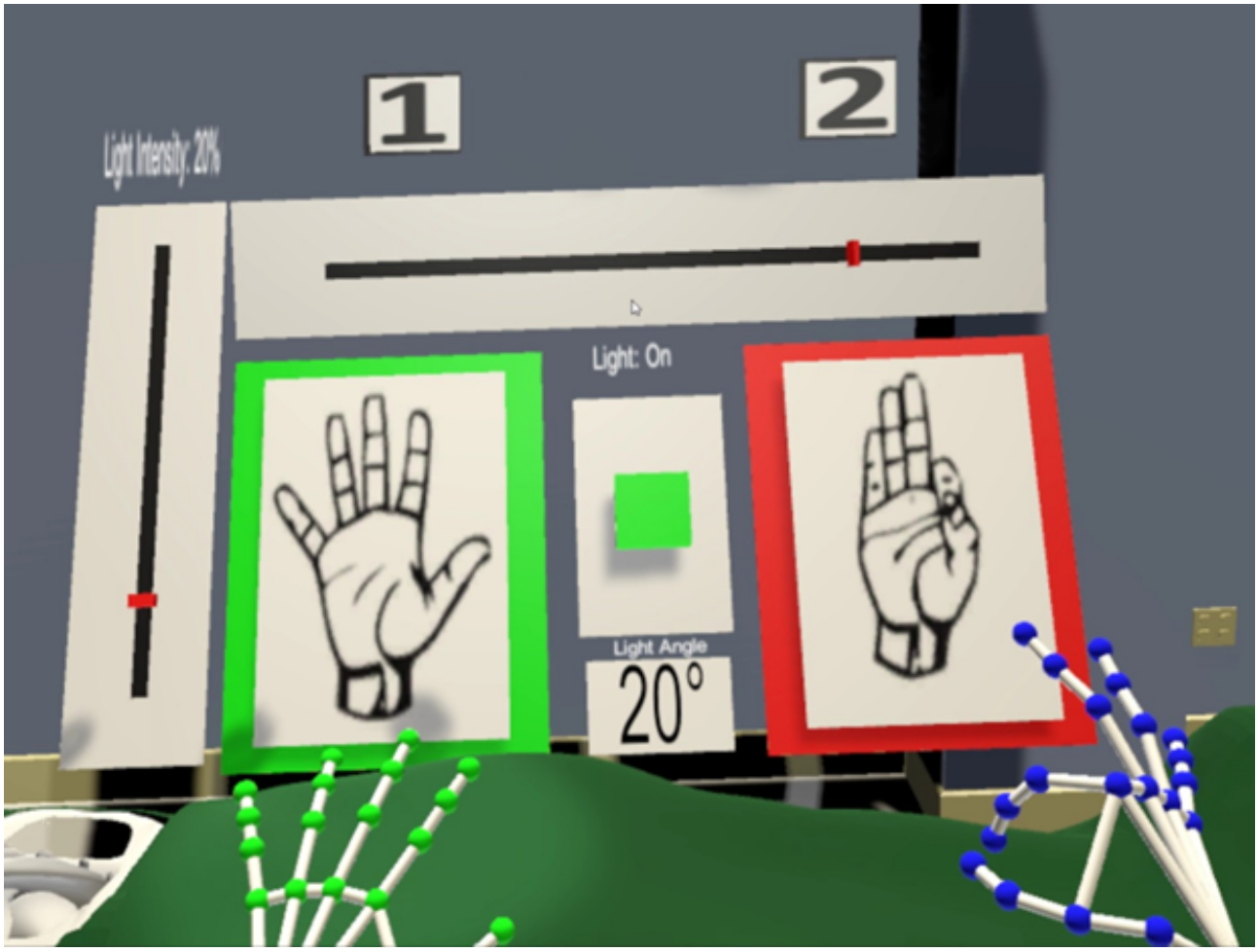


Feedback System

Both auditory and visual feedback were developed for the prototype based on recommendations from previous research [1]. Brief auditory feedback was played after the system was activated and commands were executed. The audio clips were short and subtle to not annoy or overload users with auditory feedback.

The visual feedback system was aimed at providing supplementary information regarding the hand gestures so the participants could understand the interactions better. This feedback system would show participants their current recognized hand gestures for both left and right hands using 2 separate boxes (Figure 5). The border of the boxes representing hand gestures could be either red or green. The border would turn green when a gesture was held for 1.5 seconds. This meant that activation occurred, and the designated interaction could be done. The border is red if activation has not yet taken place. A green button in the center represents the light switch. The button would be pressed when users hold the designated gesture for the light switch (“L-shaped”) and move forward. Below this button, a number shows the size of the illuminated field. For this, we used degrees rather than the diameter or the radius of the light, as the diameter could change with repositioning on uneven surfaces. Furthermore, a slider on the left side showed the degree of light intensity. Another slider at the top showed the save and load slots.

Figure 5. Visual feedback implemented in the virtual reality prototype.



XSL•FO
RenderX

Participants

Six experts took part in our study. The information about the participants' demographics can be found in Table 3. The participants included a chief physician, 2 senior physicians, and 3 assistant physicians. All were from the general and visceral surgery discipline. All surgeons regularly take part in operations (at least 2 times a week), besides one of the assistant surgeons, who takes part in operations once every 2 weeks. Only 2 surgeons had previous experience with VR, and 1 surgeon had experience with voice-based systems.

Table 3. Demographic information of the participants in the virtual reality study.

Participants	Sex	Age group (y)	Experience (y)
1	Male	56-60	26-30
2	Male	31-35	6-10
3	Male	31-35	1-5
4	Female	36-40	11-15
5	Female	36-40	6-10
6	Female	31-35	1-5

Procedure

In the beginning, the participants were informed about the procedure and risks of the study. After giving informed consent, the experts were given a detailed explanation of how the gesture and speech systems work and how the feedback system functions. They then spent 15 to 20 minutes familiarizing themselves with the VR environment. Furthermore, during this time, the participants were also asked to complete all the interactions at least once using both modalities to familiarize themselves with both speech and gesture systems. Subsequently, the experimenters explained the tasks for the study. Each participant had to perform the task in 3 rounds: once only using speech; once only using gesture; and in the last iteration, participants were free to choose the modality for each interaction. For the first 2 iterations, we changed the order of the modality for each participant to avoid order effects. In each iteration, the participants had to start by switching on the lights. They then had to position the light and adjust the illumination field, so it was positioned between the 2 drawn white rings (Figure 6). Then, they adjusted the light intensity by making it as bright as possible while not so bright that the gray tones could no longer be distinguished (Figure 7). Finally, they had to save the light settings and turn off the lights. After each iteration, participants answered a series of interview questions. The questions addressed aspects such as learnability, memorability, familiarization, audio and visual feedback systems, suggestions for improvement, and opinions on individual commands. Each session took approximately 40 to 50 minutes.

Figure 6. Examples of setting the illumination field and position of the light (left: wrong position; middle: correct setting; and right: field of illumination is too large).



XSL•FO
RenderX

Figure 7. Examples of light intensity settings (left: too dark; middle: correct setting; and right: too bright).



Results 3

Overview

Our results showed that participants found the concept compelling and helpful. All experts stated that they could imagine using the system regularly in the OR as a new system or as an add-on to the conventional lights. We witnessed that the experts often prioritized one modality over the other. Three participants suggested that gestures were generally more intuitive and easier to use, while others believed that speech was more convenient and faster to learn. One reason mentioned for this prioritization was that it was easier to master 1 than the 2 modalities. Two participants pointed out that the touch-free aspect of the interactions was a great advantage and very beneficial in the OR. They mentioned that such a system helps keep sterility and avoids asking other staff members to perform specific actions. On the other hand, 2 experts believed that the system is still not ready to fully replace a conventional SLS due to error rates and the time it takes to execute the commands, as it would be a potential risk in time-critical or emergency situations such as stopping heavy bleeding. They pointed out that the precision needs to be extremely high and the interactions even quicker.

Feedback Systems

All participants agreed that the audio feedback was helpful. Four experts mentioned that such feedback is necessary, especially for triggering the commands. Two participants noted that the auditory feedback could be turned off after a training phase for visible interactions such as positioning the light. However, 2 others mentioned that they could not imagine the system working well without auditory feedback. Participants found the activation tone of the speech system more helpful than the confirmation tone because the successful execution of the command is mainly confirmed by the visible changes in the light (with the exception of the save and load command). Regarding the visual feedback interface, 5 participants found it to be helpful, especially in the beginning. Four experts mentioned that they could use the system without visual feedback after a short training. However, 1 participant found the visual feedback disruptive and suggested only using auditory feedback.

Interactions

All but one participant found the gesture system intuitive and easy to learn. All agreed that with a short familiarization phase, one could understand and learn to interact with the system. However, 2 participants indicated that our familiarization phase was too short for the gesture interactions, and they recommended prolonging that to 30 to 60 minutes. Two participants found the execution of gestures to be ergonomic and quick. Three participants highlighted that knowing the commands from similar activities helped them to learn the interactions faster and the execution of the commands easier. Five experts believed that gesture interaction was more appropriate for positioning the light, as it was more precise. One participant found the load and save gesture command rather tricky and unintuitive. Another participant noted that it was unclear how to stop the hand tracking after performing the command. One participant found the gesture recognition not precise enough. The error rate for this user was higher than that of all other participants. Similar to the gesture system, all participants perceived speech commands as intuitive, fitting, and easy to remember. Two participants added that they particularly liked that the commands were short, which can be advantageous when used in the OR. One participant highlighted the high accuracy of the speech system. Two participants specifically praised the wake word “Hey, light!” and found it appropriate and fitting. One participant remarked that the trigger functions or functionality switches [11,38] are very suitable with voice commands because performing them requires little cognitive and physical effort (eg, light switch). A major criticism that was pointed out by 3 participants was about setting continuous parameters with speech, such as positioning the light beam. They found such interactions overly complex and confusing. One expert noted that estimating the numbers as measurements needed for the voice command is difficult. All participants wore medical masks during the study. We did not observe any negative impacts caused by wearing the mask. One of the participants believed that using the German language would be more convenient for the speech system. Others found English to be appropriate and sufficient. One participant experienced difficulties with the speech commands due to their accent.

Modality Preferences

Participants’ reasons for choosing a modality over another were based on the number of commands needed for execution, the

XSL•FO
RenderX

speed of the interaction, the cognitive and physical effort, and the precision of the commands. All but one participant used a mixture of both speech and gestures to go through the tasks in the last iteration. Table 4 shows the chosen modality of all participants for individual interactions in the third iteration. Among the 5 subtasks were 2 trigger functions and 3 continuous parameter settings. We observed that parameters were preferred to be set by gesture, as the dynamic feedback would help to interact more precisely. For such interactions, multiple speech commands were often needed to set the parameters accordingly. In the subsequent interview, we asked the participants why they chose a particular modality for each command in the last iteration. During these interviews, we found that, although participants might have preferred one modality over the other, 2 participants chose the less favored one to give it another chance. One participant who experienced problems using the speech system due to unclear English pronunciation performed all the tasks in the last iteration with the gesture. One participant indicated that they mainly worked with speech, as the familiarization phase was too short for them, and they forgot how to carry out the command with the other modality. All participants stated they could use the system entirely without further help after a short training period. Two participants stated that having both modalities simultaneously during surgeries is beneficial. One said that they would be able to use voice commands when their hands are busy or gesture when the noise level in the room is very loud. No participant expressed a negative characteristic of using both modalities simultaneously. Participants particularly emphasized the advantage of dynamic feedback and fine motor skills when setting continuous parameters with gestures, as well as speed and the low cognitive and physical effort in trigger functions when using speech.

Table 4. The selected modalities for each interaction in the third iteration of the virtual reality study.

Interaction type	Speech	Gesture
Light switch	5	1
Position	1	5
Intensity	3	3
Field of illumination	0	6
Save or load	4	2

Methods 4: Living Laboratory Study

Overview

The interaction concept was implemented in a prototype to be tested in a living laboratory. Similar to the VR study, participants could adjust the SLS in a living laboratory using gestures and speech, where every interaction could be made with both modalities. In this study, we used smart lighting modules introduced by Mühlenbrock et al [20] to test our interaction concept. This novel lighting system is specifically designed for the OR, which uses an array of small lighting modules mounted on the ceiling [20].

Smart Ceiling Lights

Each module is equipped with 2 motors for directional adjustment. The modules are arranged in a 7×8 grid with distances of 0.36×0.35 m and are controlled by a central computer that adjusts their intensity and orientation (Figure 8). To minimize shadows, the system uses geometry recognition of objects and personnel within the surgical area via 3 Kinect (version 2) [60] sensors, which are registered with each other and with the room. However, the automatic shadow prevention feature was disabled during the study. The lighting system provides an application programming interface that accepts commands for the absolute target position of the light in 3D space, relative adjustments to this position, and changes in lighting intensity. It then adjusts the motors and intensity of the lights accordingly. While individual modules are capable of providing up to 50 klx, compliance with medical standards limits the system to a maximum of 160 klx. The paper by Mühlenbrock et al [20] provides further details on the ceiling lights.

XSL•FO
RenderX

Figure 8. Smart ceiling lights used for the living laboratory study, consisting of an array of small lighting modules.



Interaction Concept Integration

We integrated our interaction software with the lighting system application programming interface to be able to interact with the lights using our touch-free concept. The prototype was implemented in Python (Python Software Foundation). We used Kinect Azure [61] for tracking hand gestures and used the Picovoice Speech-to-Text library [58] for speech recognition. A neural network classified the detected hands after matching features had been computed.

Due to the limitations regarding the motors in individual light modules, we were not able to change the field of illumination adequately. Therefore, we excluded this interaction from this study. The rest of the interactions were identical to the ones in the VR study (Table 2).

Moreover, in a realistic surgical environment, multiple hands may appear within the camera's field of view, complicating accurate hand recognition. To address this, we used the Kinect RGB camera and trained the system to track hands wearing a specific color of gloves—in this case, blue. This way, we could ensure that the system only tracks the designated surgeon's hands to adjust lights. Testing confirmed that only hands with blue gloves were detected. A notable advantage of this method is its ability to restrict access to authorized users based on glove color. However, a limitation arises when gloves become bloodstained, as red-tinted areas may hinder recognition. To mitigate this issue, the algorithm could be adapted to reassign red pixels to a skin-like color space to ensure consistent recognition.

Participants

For this study, we recruited 12 experts. Other than 1 medical student, all other participants were surgeons or assistant surgeons, and they regularly took part in operations (at least 2 times a week). The information about the demographics of the participants can be found in Table 5. All participants were from the general and visceral surgery disciplines.

Table 5. Demographic information of the participants in the living laboratory study.

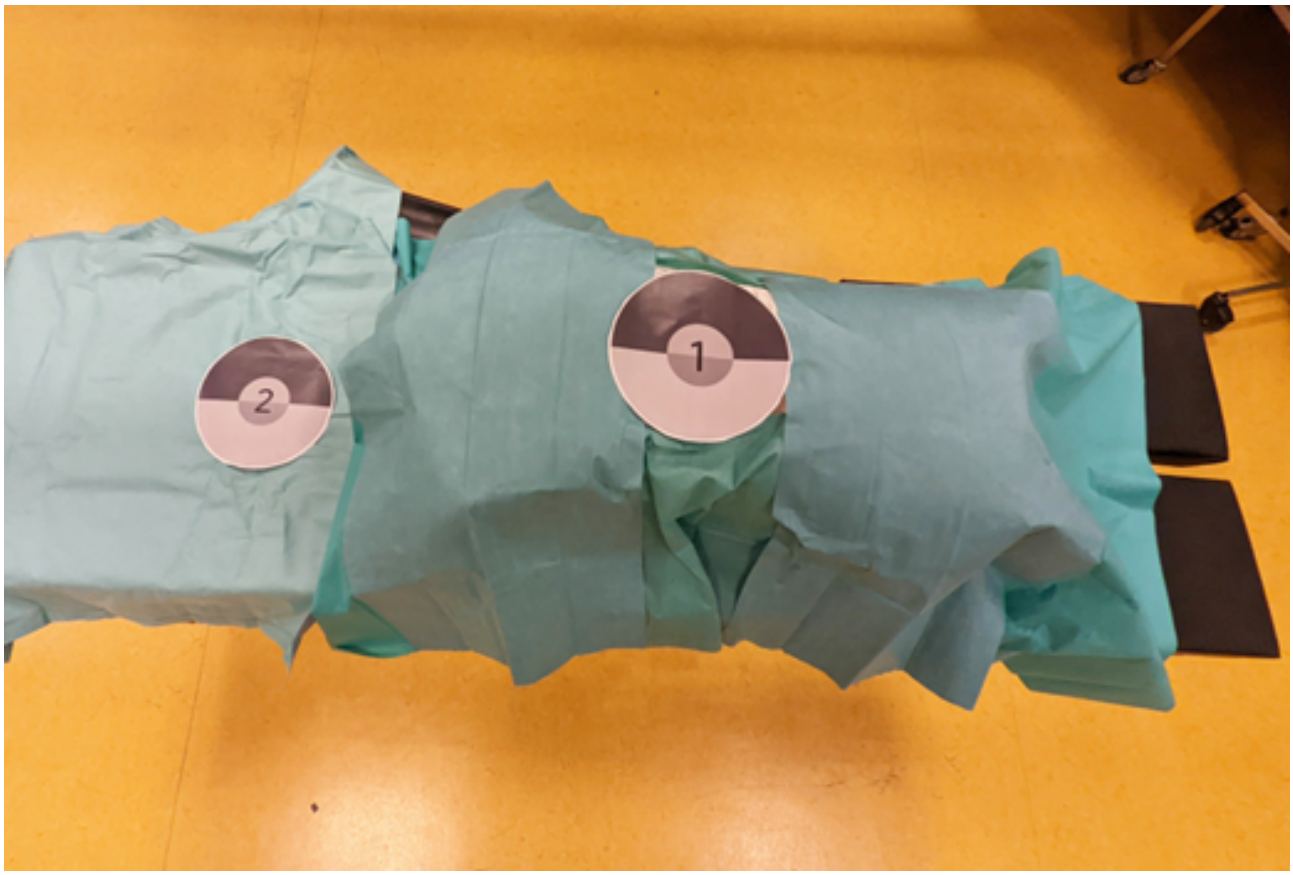
Participants	Sex	Age group (y)	Experience (y)
P1	Female	21-25	1-5
P2	Male	41-45	16-20
P3	Male	46-50	21-25
P4	Female	26-30	1-5
P5	Male	31-35	6-10
P6	Male	26-30	1-5
P7	Male	31-35	1-5
P8	Male	26-30	1-5
P9	Female	41-45	11-15
P10	Male	36-40	11-15
P11	Male	31-35	6-10
P12	Female	26-30	1-5

XSL•FO
RenderX

Procedure

After a short welcoming conversation, participants were briefed about the study procedure and gave informed consent. All participants were given a detailed explanation of how the interaction with the system works and how to use both modalities of gesture and speech. Participants then spent around 10 minutes familiarizing themselves with the system by performing all the supported interactions at least once using both modalities. Afterward, the experimenter explained the tasks to perform for the study. Similar to our VR experiment, each participant had to perform the task in 3 rounds. However, this time they could choose which modality to use for individual interactions in every iteration. For each iteration, we used 2 circular position marks, numbered mark 1 and mark 2 (Figure 9). Participants had to position the light on each ring, save that position with the mark's associated number, and switch between the 2 positions by loading the previously defined positions. The light intensity had to be set differently for each mark. One had to have a higher intensity and one, a lower intensity. Adjusting the size of the illuminated field was not considered in this study due to technical limitations with the physical lights in the living laboratory. The participants chose the order of the tasks to perform, and the experimenters suggested no strict order. At the end of each iteration, participants took a short break and continued with the next one. Once the last iteration was through, a semistructured interview was conducted where experts answered a series of questions about the overall system, learnability, memorability, familiarization, audio feedback, suggestions for improvement, and opinions on individual commands. Sessions took approximately 30 (mean 31.91, SD 4.85) minutes.

Figure 9. Setup of the user study in the living laboratory.



Results 4

Interaction Ratings

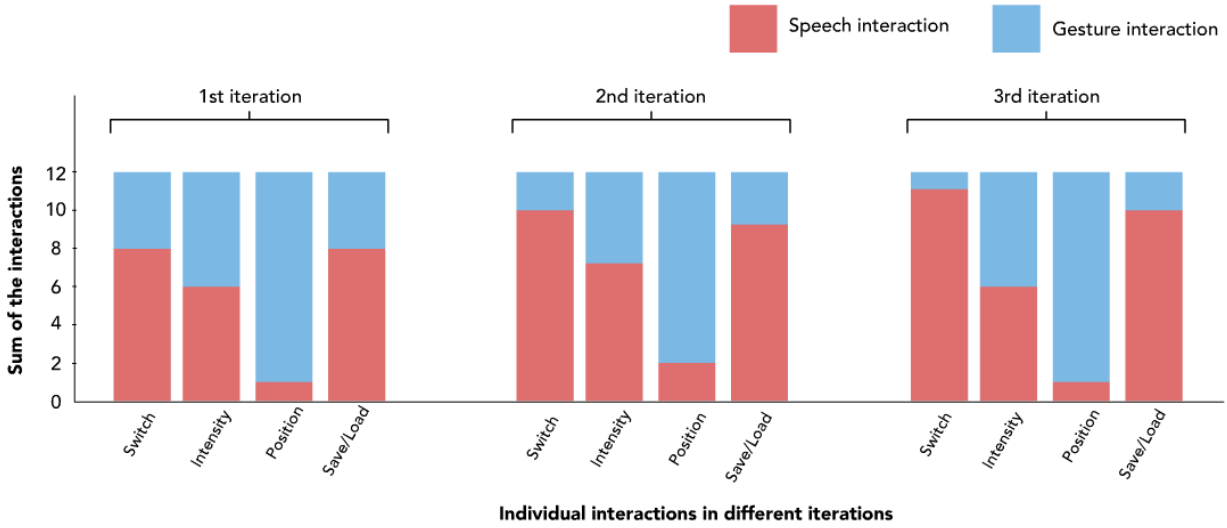
Our results indicated that most participants (11/12, 92%) found both speech and gesture interactions appropriate for their intended actions. Most participants (9/12, 75%) envisioned using such a system in the OR regularly. All participants agreed that, with a brief training period, they could use the system without needing ongoing technical support. While 2 participants suggested that the training might take up to a week, the remaining participants believed it could be completed in under an hour. They described the interactions as "fast and easy" (P3). We witnessed that the experts' performance improved over time. We saw that they worked easier with the system in the second and third iterations compared to their initial attempts. This trend suggests a positive learning effect with increased exposure to the system. Initially, during the first 2 iterations, participants were open to using both speech and gesture modalities for various interactions. However, by the final iteration, clear preferences for specific modalities had emerged depending on the task (Figure 10). For certain interactions, distinct modality preferences became apparent. Participants found speech more suitable for trigger functions, such as switching lights or saving and loading lighting conditions. Conversely, gestures were preferred for continuous parameters, such as positioning the light. Opinions on controlling light intensity were divided: half of the participants found gesture

XSL•FO
RenderX

interactions more intuitive, while the other half preferred using speech. All participants considered audio feedback as highly crucial for their interactions. A few participants (3/12, 25%) suggested using distinct sounds for different actions to help differentiate them more clearly. Participants found audio feedback particularly beneficial for specific tasks, such as positioning, as it provided essential information about the system's current state and alerted them when an action was not recognized.

Despite general satisfaction with individual interactions, participants provided several recommendations for adjustments and suggested new interaction features. Two participants recommended using shorter speech commands for adjusting light intensity. They proposed a simplified system where intensity is set on a scale of 1 to 10 rather than using percentage values. One participant (P5) suggested adding a feature to disable gesture tracking via speech commands to prevent misrecognition and unintended hand tracking. Participants also recommended the ability to revert to a previous light configuration. There were suggestions to combine modalities for specific interactions. For instance, 1 expert proposed a mixed approach for the save and load function: initiating the action with a gesture and then using speech to assign it to a specific slot (eg, "Save two"). However, 4 participants found the gesture-based save or load interaction unintuitive and overly complex. Two participants noted that this interaction was cognitively demanding because it required memorizing which position corresponded to each number. One participant suggested using specific names for save slots (eg, "save stomach") instead of numerical labels. In addition, 5 participants explicitly mentioned that using speech commands for positioning was inappropriate. They found it required excessive effort and approximation, making it less effective compared to the gesture.

Figure 10. Bar chart showing the experts' selected modality for each interaction on all 3 iterations.



Experimenter Observations

We observed that the field of view of the depth camera used to track hand movements positioned above the surgical bed was limited. Experts noted the desire to move the surgical light into zones that fell outside this camera's range. This limitation could be addressed by integrating multiple cameras, which would expand the tracking area and potentially enhance gesture recognition accuracy. Despite the system's capabilities, we still encountered several recognition errors for both speech and gesture modalities. To ensure reliability, these systems require better training and increased accuracy, especially in a high-stakes environment like an OR where patient safety is critical. The current system's precision needs to be enhanced to meet the strict demands of surgical settings. We found that participants' proficiency with the system improved over time with increased use. This suggests that future implementations should include longer and more comprehensive training sessions to help users become adept at operating the system efficiently. Four participants reported that the gesture-based positioning of the surgical light was slow. This issue stemmed from the limitations of the motors used in the light modules, which caused a noticeable delay in moving the light. Addressing these mechanical limitations could improve the system's responsiveness and enhance the user experience.

Discussion

Principal Findings

In this work, we examined a hands-free solution for controlling an SLS in the OR. In an iterative design process, we developed a system that uses speech and gestures as interaction modalities to control the SLS. We observed that surgeons and practitioners found such a system highly valuable and practical. They believed that this system could certainly support sterility and give them more direct control over the SLS.

Our results highlighted that the combination of modalities was a quite helpful approach, supporting the previous literature [10,29,42]. We witnessed certain interactions (specifically, triggers) that were much better suited to be performed by speech and others better suited with gesture (continuous parameters), in line with previous work [11,38]. In line with the recommendation from Mentis et al [11], our participants also suggested having the opportunity to be able to use both modalities for each interaction to have higher flexibility. We

XSL•FO
RenderX

witnessed that experts often preferred one modality over the other, largely based on perceived cognitive load and familiarity with the modality. Participants' opinions differed on which modality they found less cognitively demanding; some considered speech interaction easier due to its directness and hands-free nature, while others preferred gestures for visual immediacy. These findings highlight the potential of multimodal systems to enhance interaction effectiveness. The clear observation that certain tasks are better suited to one modality over another highlights the advantages of leveraging the strengths of both speech- and gesture-based interactions. Although mastering both interaction methods might initially pose a cognitive challenge for users, we believe this can be mitigated through brief technical training and the design of intuitive interactions and significantly reduce cognitive demands over time. Moreover, multimodal systems also serve as a critical redundancy layer—ensuring functionality even if one modality fails. In addition, providing users with the flexibility to choose their preferred modality fosters greater satisfaction and increases the likelihood of broader adoption.

The list of functions we planned for the adjustment of the lights was consistently found to be practical and useful by participants. Notably, participants highly favored predefined light settings, aligning with findings from previous research [3], as these settings simplify operations and save time for the personnel. The individual interactions were mainly found to be intuitive and usable. However, certain interactions that were not matched well with their modality were criticized, for example, saving and loading light settings with gesture. As this is considered a trigger function, speech seems to be the more intuitive modality to use. Participants found the freehand control aspect of using speech quite helpful. This meant the surgeons could use both hands at all times and focus on their primary tasks, as pointed out by previous work [38,39]. Furthermore, it is common that during surgeries, multiple hands are actively working on the patient, making it challenging to identify the correct hand for gesture control. Another concern with gestures is that they can cause collisions, a problem that does not occur with speech. On the other hand, speech is less appropriate for fine adjustments and more exact changes, as seen in our evaluations. In addition, a loud OR environment and problems with speech recognition, such as the accuracy for nonnative speakers or people with accents or certain dialects, could also be raised. These issues further emphasize the benefits of a multimodal system, where challenges with one modality can be mitigated by leveraging the strengths of the other.

A key advantage of our proposed system was the enhanced agency it could offer to the main surgeon. This was particularly appreciated when the surgeons' hands were occupied, as they could still control the lights, providing them with higher control and efficiency. Experts highlighted the benefit of eliminating the need to rely on additional personnel for adjustments, potentially reducing errors and saving time [2].

Throughout our examination, we observed that the biggest concern raised by the experts was the technological limitations, such as accuracy and functionality problems with the systems. In different stages of our evaluation, participants faced certain challenges with inaccuracies where the developed system did not behave as the experts intended it to do. The experts repeatedly highlighted that in such a sensitive context where patients' safety could be impacted, there is no room for error. Hence, the system ideally has to function error-free with perfect efficacy. As previous work highlighted, events that require unnecessary actions from the surgeons could hinder their performance [2]. It comes as no surprise that systems deployed in sensitive environments, such as health care, must prioritize reliability and error mitigation. Experts' emphasis on error-free operation highlights the unique stakes in surgical contexts, where minor inaccuracies could risk patient safety. In such high-stakes applications, designers must aim for the highest technical accuracy and consider appropriate backup methods and error recovery means to safeguard against unanticipated failures. Several technical concerns were highlighted, including the high noise levels in the OR and worries about their potential impact on speech command accuracy, as well as the challenges of achieving the high precision required for fine adjustments. Addressing these challenges is essential, and future research should focus on developing solutions that specifically mitigate these issues to ensure reliable system performance. Moreover, we observed that few participants did not trust the technology to work seamlessly due to old experiences or unfamiliarity with the technology (eg, "speech will not work well!"). Our participants highlighted that the lack of trust in the technology could raise the stress levels of the surgeons. Skepticism stemming from users' past experiences or unfamiliarity with the technology underscores the importance of fostering trust through transparency, education, and hands-on training with novel systems. Bridging the gap between perceived and actual system reliability through effective communication and usability testing is essential for promoting user trust and ensuring successful integration. A gradual adoption in low-risk contexts (such as VR simulation) can build confidence before full integration in critical tasks.

Limitations and Future Work

Our results offer valuable insights into using a touch-free interaction approach for controlling the SLS in the OR. Nevertheless, our work has certain limitations that require acknowledgment.

In our iterative evaluation, we mainly relied on qualitative approaches to assess our concept and interaction system. This provided us with in-depth insights into user perceptions, behaviors, and subjective experiences, enabling us to refine and improve the system based on detailed feedback and observations. However, for future work, incorporating quantitative approaches will be crucial to measuring usability, cognitive load, and user experience more objectively. By using metrics such as task completion times, error rates, physiological measurements, and standardized questionnaires, we can obtain a comprehensive understanding of the system's performance and its impact on users, ensuring a more robust and reliable evaluation. Moreover, our designed system was still a prototype, and participants encountered technical challenges at various stages. Future work should also focus on better training the systems to improve recognition rates, providing a more reliable foundation for such interaction methods.

XSL•FO
RenderX

Our studies primarily consisted of laboratory experiments and interviews. As an initial approach, this selection of methods is reasonable, providing controlled conditions to refine the system and gather preliminary insights. Nevertheless, testing the system in more complex environments and realistic situations—such as in noisy environments, with multiple hand-in scenarios, and across various surgical disciplines—would be beneficial. This would help us observe and address potential challenges and limitations, ensuring the system's robustness, adaptability, and effectiveness in real-world applications.

Finally, in our experiments, we observed that longer training sessions with the system could have led to better performance from the experts. Therefore, we recommend extended familiarization phases to allow users to become more acquainted with the system, enhancing their proficiency and effectiveness. Future work could also explore the optimal duration and structure of these training sessions.

Conclusions

This work sought to develop an intelligent interaction concept for controlling an SLS. We designed a touch-free interaction concept using speech and hand gestures to control the SLS. In an iterative process, using a user-centered design approach, we used a 5-step design process including a literature review and OR visits, interviews with surgeons and assistant surgeons, focus groups with the OR personnel, a VR evaluation, and a living laboratory study. Our findings suggest that experts find our interaction concept highly useful and beneficial for the OR. Despite facing technical challenges, surgeons saw great potential for such a touch-free approach and could imagine using it regularly in the OR. Moving forward, further refinement and validation of such a multimodal system will be essential to fully realize its potential in real-world surgical settings.

Acknowledgments

This work was supported by the German Federal Ministry of Education and Research (BMBF) project Smart-OT. The authors sincerely thank all their project partners for their support and contribution to this research.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary figures and illustrations. This appendix includes all figures referenced throughout the paper.
[ZIP File (Zip Archive), 8492 KB - humanfactors_v12i1e70628_app1.zip]

References

1. Armin D, Stephan S, Andreas P, Klaus F, Eduard R. Contactless surgery light control based on 3d gesture recognition. In: Proceedings of the 2nd Global Conference on Articial Intelligence. 2016 Sep Presented at: GCAI '16; September 29-October 2, 2016; Berlin, Germany p. 138-146 URL: https://easychair.org/publications/paper/L1 [doi: 10.29007/zmz9]
2. Knulst AJ, Mooijweer R, Jansen FW, Stassen LP, Dankelman J. Indicating shortcomings in surgical lighting systems. Minim Invasive Ther Allied Technol 2011 Sep;20(5):267-275. [doi: 10.3109/13645706.2010.534169] [Medline: 21082904]
3. Rostenberg B, Barach PR. Design of cardiovascular operating rooms for tomorrow's technology and clinical practice — part 2. Prog Pediatr Cardiol 2012 Jan;33(1):57-65. [doi: 10.1016/j.ppedcard.2011.12.010]
4. Curlin J, Herman CK. Current state of surgical lighting. Surg J (N Y) 2020 Apr 19;6(2):e87-e97 [FREE Full text] [doi: 10.1055/s-0040-1710529] [Medline: 32577527]
5. Ay F, Gencturk N. Disinfection and sterilization related situations for patient safety in operation rooms. Int J Caring Sci 2018;11(1):607 [FREE Full text]
6. Reinschluessel AV, Teuber J, Herrlich M, Bissel J, van Eikeren M, Ganser J, et al. Virtual reality for user-centered design and evaluation of touch-free interaction techniques for navigating medical images in the operating room. In: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems. 2017 Presented at: CHI EA '17; May 6-11, 2017; Denver, CO p. 2001-2009 URL: https://dl.acm.org/doi/10.1145/3027063.3053173 [doi: 10.1145/3027063.3053173]
7. Volkmar G, Muender T, Wenig D, Malaka R. Evaluation of natural user interfaces in the creative industries. In: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems. 2020 Presented at: CHI EA '20; April 25-30, 2020; Honolulu, HI p. 1-8 URL: https://dl.acm.org/doi/abs/10.1145/3334480.3375201#:~:text=Our%20results%20indicate%20that%20NUIs,to%20achieve%20the%20intended%20outcome [doi: 10.1145/3334480.3375201]
8. Malaka R, Döring T, Fröhlich T, Muender T, Volkmar G, Wenig D, et al. Using natural user interfaces for previsualization. EAI Endorsed Trans Creat Technol 2021 Mar 16;8(26):169030. [doi: 10.4108/eai.16-3-2021.169030]
9. Hartmann F, Schlaefer A. Feasibility of touch-less control of operating room lights. Int J Comput Assist Radiol Surg 2013 Mar 19;8(2):259-268. [doi: 10.1007/s11548-012-0778-2] [Medline: 22806717]

XSL•FO
RenderX

10. Ebert LC, Hatch G, Ampanozi G, Thali MJ, Ross S. You can't touch this: touch-free navigation through radiological images. Surg Innov 2012 Sep 06;19(3):301-307. [doi: 10.1177/1553350611425508] [Medline: 22064490]
11. Mentis HM, O'Hara K, Gonzalez G, Sellen A, Corish R, Criminisi A, et al. Voice or gesture in the operating room. In: Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems. 2015 Presented at: CHI EA '15; April 18-23, 2015; Seoul, Republic of Korea p. 773-780 URL: https://dl.acm.org/doi/10.1145/2702613.2702963 [doi: 10.1145/2702613.2702963]
12. Mewes A, Saalfeld P, Riabikin O, Skalej M, Hansen C. A gesture-controlled projection display for CT-guided interventions. Int J Comput Assist Radiol Surg 2016 Jan 10;11(1):157-164. [doi: 10.1007/s11548-015-1215-0] [Medline: 25958060]
13. Hatscher B, Mewes A, Pannicke E, Kägebein U, Wacker F, Hansen C, et al. Touchless scanner control to support MRI-guided interventions. Int J Comput Assist Radiol Surg 2020 Mar 13;15(3):545-553. [doi: 10.1007/s11548-019-02058-1] [Medline: 31520326]
14. Patkin M. What surgeons want in operating rooms. Minim Invasive Ther Allied Technol 2003 Nov 10;12(6):256-262. [doi: 10.1080/13645700310021133] [Medline: 16754107]
15. Burger JM. Automated surgical light positioning system using thermal-imaging and optical techniques. Stellenbosch University. URL: https://scholar.sun.ac.za/items/a8c26721-325a-495c-acae-96aa2d484cee [accessed 2024-04-29]
16. Conrad C, Konuk Y, Werner PD, Cao CG, Warshaw AL, Rattner DW, et al. A quality improvement study on avoidable stressors and countermeasures affecting surgical motor performance and learning. Ann Surg 2012 Jun;255(6):1190-1194 [FREE Full text] [doi: 10.1097/SLA.0b013e318250b332] [Medline: 22584632]
17. Hemphälä H, Osterhaus W, Larsson P, Borell J, Nylén P. Towards better lighting recommendations for open surgery. Light Res Tech 2020 Feb 25;52(7):856-882. [doi: 10.1177/1477153520903355]
18. Teuber J, Weller R, Kikinis R, Oldhafer KJ, Lipp MJ, Zachmann G. Autonomous surgical lamps. Universität Bremen. URL: https://cgvr.cs.uni-bremen.de/research/asula/ [accessed 2024-04-29]
19. Mühlenbrock A, Weller R, Zachmann G. Optimizing the illumination of a surgical site in new autonomous module-based surgical lighting systems. In: Proceedings of 2022 International Conference on Medical Imaging and Computer-Aided Diagnosis. 2022 Presented at: MICAD '22; November 20-21, 2022; Leicester, UK p. 293-303 URL: https://link.springer.com/chapter/10.1007/978-981-16-6775-6_24 [doi: 10.1007/978-981-16-6775-6_24]
20. Mühlenbrock A, Huscher H, Uslar VN, Cetin T, Weller R, Weyhe D, et al. A novel, autonomous, module-based surgical lighting system. ACM Trans Comput Healthcare 2025 Jan 08;6(1):1-29. [doi: 10.1145/3696670]
21. Andre M, Timur C, Dirk W, Gabriel Z. Optimizing the arrangement of fixed light modules in new autonomous surgical lighting systems. In: Proceedings of the 2022 Conference on Image-Guided Procedures, Robotic Interventions, and Modeling. 2022 Presented at: IGPRIM '22; February 22-23, 2022; San Diego, CA p. 2022 URL: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/12034/2611058/Optimizing-the-arrangement-of-fixed-light-modules-in-new-autonomous/10.1117/12.2611058.short [doi: 10.1117/12.2611058]
22. Hatscher B, Luz M, Nacke LE, Elkmann N, Müller V, Hansen C. GazeTap: towards hands-free interaction in the operating room. In: Proceedings of the 19th ACM International Conference on Multimodal Interaction. 2017 Presented at: ICMI '17; November 13-17, 2017; Glasgow, UK p. 243-251 URL: https://dl.acm.org/doi/10.1145/3136755.3136759 [doi: 10.1145/3136755.3136759]
23. Zhu H, Rohling RN, Salcudean SE. Hand-eye coordination-based implicit re-calibration method for gaze tracking on ultrasound machines: a statistical approach. Int J Comput Assist Radiol Surg 2020 May 22;15(5):837-845. [doi: 10.1007/s11548-020-02143-w] [Medline: 32323208]
24. Black D, Unger M, Fischer N, Kikinis R, Hahn H, Neumuth T, et al. Auditory display as feedback for a novel eye-tracking system for sterile operating room interaction. Int J Comput Assist Radiol Surg 2018 Jan 27;13(1):37-45 [FREE Full text] [doi: 10.1007/s11548-017-1677-3] [Medline: 29079993]
25. Hettig J, Saalfeld P, Luz M, Becker M, Skalej M, Hansen C. Comparison of gesture and conventional interaction techniques for interventional neuroradiology. Int J Comput Assist Radiol Surg 2017 Sep 24;12(9):1643-1653. [doi: 10.1007/s11548-017-1523-7] [Medline: 28120179]
26. Lee D, Choi J, Park JI, Kim M. Practical surgical interface using a RGB-D camera. In: Proceedings of the 18th IEEE International Symposium on Consumer Electronics. 2014 Presented at: ISCE '14; June 22-25, 2014; Jeju, South Korea p. 1-2 URL: https://ieeexplore.ieee.org/document/6884492 [doi: 10.1109/isce.2014.6884492]
27. Wachs JP, Stern HI, Edan Y, Gillam M, Handler J, Feied C, et al. A gesture-based tool for sterile browsing of radiology images. J Am Med Inform Assoc 2008 Feb 28;15(3):321-323. [doi: 10.1197/jamia.m241]
28. Wachs J, Stern H, Edan Y, Gillam M, Feied C, Smithd M, et al. Real-time hand gesture interface for browsing medical images. Int J Intel Comput Med Sci Image Process 2008 Jan;2(1):15-25. [doi: 10.1080/1931308x.2008.10644149]
29. O'Hara K, Gonzalez G, Sellen A, Penney G, Varnavas A, Mentis H, et al. Touchless interaction in surgery. Commun ACM 2014 Jan;57(1):70-77. [doi: 10.1145/2541883.2541899]
30. Nestorov N, Hughes P, Healy N, Sheehy N, OHare N. Application of natural user interface devices for touch-free control of radiological images during surgery. In: Proceedings of the 29th International Symposium on Computer-Based Medical Systems. 2016 Presented at: CBMS '16; June 20-24, 2016; Dublin, Ireland p. 229-234 URL: https://ieeexplore.ieee.org/document/7545992 [doi: 10.1109/cbms.2016.20]

XSL•FO
RenderX

31. Yusoff YA, Basori AH, Mohamed F. Interactive hand and arm gesture control for 2D medical image and 3D volumetric medical visualization. Procedia Soc Behav Sci 2013 Nov;97:723-729. [doi: 10.1016/j.sbspro.2013.10.293]
32. Grange S, Fong T, Baur C. M/oris: a medical/operating room interaction system. In: Proceedings of the 6th international conference on Multimodal interfaces. 2004 Presented at: ICMI '04; October 13-15, 2004; Borough, PA p. 159-166 URL: https://dl.acm.org/doi/10.1145/1027933.1027962 [doi: 10.1145/1027933.1027962]
33. Hansen L, Siebert M, Diesel J, Heinrich MP. Fusing information from multiple 2D depth cameras for 3D human pose estimation in the operating room. Int J Comput Assist Radiol Surg 2019 Nov 6;14(11):1871-1879. [doi: 10.1007/s11548-019-02044-7] [Medline: 31388959]
34. Gibby JT, Swenson SA, Cvetko S, Rao R, Javan R. Head-mounted display augmented reality to guide pedicle screw placement utilizing computed tomography. Int J Comput Assist Radiol Surg 2019 Mar 22;14(3):525-535. [doi: 10.1007/s11548-018-1814-7] [Medline: 29934792]
35. Zaman A, Roy A, Fatema K, Farin NJ, Doring T, Malaka R. Explore voice and foot-based interaction techniques to navigate 2D radiological images in the virtual reality operation theatre. In: Proceedings of the 22nd International Conference on Computer and Information Technology. 2019 Presented at: ICCIT '19; December 18-20, 2019; Dhaka, Bangladesh p. 1-7 URL: https://ieeexplore.ieee.org/document/9038175 [doi: 10.1109/iccit48885.2019.9038175]
36. Grinshpoon A, Sadri S, Loeb GJ, Elvezio C, Feiner SK. Hands-free interaction for augmented reality in vascular interventions. In: Proceedings of the 2018 IEEE Conference on Virtual Reality and 3D User Interfaces. 2018 Presented at: VR '18; March 18-22, 2018; Reutlingen, Germany p. 751-752 URL: https://ieeexplore.ieee.org/document/8446259 [doi: 10.1109/vr.2018.8446259]
37. Kim YS, Dontcheva M, Adar E, Hullman J. Vocal shortcuts for creative experts. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 2019 Presented at: CHI '19; May 4-9, 2019; Scotland, UK p. 1-4 URL: https://tinyurl.com/3abretrk [doi: 10.1145/3290605.3300562]
38. Mewes A, Hensen B, Wacker F, Hansen C. Touchless interaction with software in interventional radiology and surgery: a systematic literature review. Int J Comput Assist Radiol Surg 2017 Feb 19;12(2):291-305. [doi: 10.1007/s11548-016-1480-6] [Medline: 27647327]
39. Alapetite A. Speech recognition for the anaesthesia record during crisis scenarios. Int J Med Inform 2008 Jul;77(7):448-460. [doi: 10.1016/j.ijmedinf.2007.08.007] [Medline: 17904900]
40. El-Shallaly GE, Mohammed B, Muhtaseb MS, Hamouda AH, Nassar AH. Voice recognition interfaces (VRI) optimize the utilization of theatre staff and time during laparoscopic cholecystectomy. Minim Invasive Ther Allied Technol 2005 Jul 10;14(6):369-371. [doi: 10.1080/13645700500381685] [Medline: 16754183]
41. Perrakis A, Hohenberger W, Horbach T. Integrated operation systems and voice recognition in minimally invasive surgery: comparison of two systems. Surg Endosc 2013 Feb 28;27(2):575-579. [doi: 10.1007/s00464-012-2488-9] [Medline: 22926891]
42. Hötker AM, Pitton MB, Mildenberger P, Düber C. Speech and motion control for interventional radiology: requirements and feasibility. Int J Comput Assist Radiol Surg 2013 Nov 13;8(6):997-1002. [doi: 10.1007/s11548-013-0841-7] [Medline: 23580026]
43. Nathan CA, Chakradeo V, Malhotra K, D'Agostino H, Patwardhan R. The voice-controlled robotic assist scope holder AESOP for the endoscopic approach to the sella. Skull Base 2006 Aug 17;16(3):123-131 [FREE Full text] [doi: 10.1055/s-2006-939679] [Medline: 17268585]
44. Punt MM, Stefels CN, Grimbergen CA, Dankelman J. Evaluation of voice control, touch panel control and assistant control during steering of an endoscope. Minim Invasive Ther Allied Technol 2005 Jul 10;14(3):181-187. [doi: 10.1080/13645700510033967] [Medline: 16754161]
45. Zargham N, Fetni ML, Spillner L, Muender T, Malaka R. "I know what you mean": context-aware recognition to enhance speech-based games. In: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems. 2024 Presented at: CHI '24; May 11-16, 2024; Honolulu, HI p. 1-18 URL: https://dl.acm.org/doi/10.1145/3613904.3642426 [doi: 10.1145/3613904.3642426]
46. Bubeck S, Chandrasekaran V, Eldan R, Gehrke J, Horvitz E, Kamar E, et al. Sparks of artificial general intelligence: early experiments with GPT-4. arXiv Preprint posted online March 22, 2023 [FREE Full text]
47. Xu A, Rao H, Dow SP, Bailey BP. A classroom study of using crowd feedback in the iterative design process. In: Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing. 2015 Presented at: CSCW '15; March 14-18, 2015; Vancouver, BC p. 1637-1648 URL: https://dl.acm.org/doi/10.1145/2675133.2675140 [doi: 10.1145/2675133.2675140]
48. Elkins J. Art Critiques: A Guide. Washington, DC: New Academia Publishing; 2012.
49. Connelly LM, Peltzer JN. Underdeveloped themes in qualitative research: relationship with interviews and analysis. Clin Nurse Spec 2016;30(1):52-57. [doi: 10.1097/NUR.0000000000000173] [Medline: 26626748]
50. Braun V, Clarke V, Hayfield N, Terry G. Thematic Analysis: A Practical Guide. Thousand Oaks, CA: Sage Publications; 2019.
51. Morgan H. Understanding thematic analysis and the debates involving its use. Qual Rep 2022 Oct 03;27(10):2079-2091 [FREE Full text] [doi: 10.46743/2160-3715/2022.5912]

XSL•FO
RenderX

52. Thomas DR. A general inductive approach for analyzing qualitative evaluation data. Am J Eval 2006 Jun 01;27(2):237-246. [doi: 10.1177/1098214005283748]
53. Chandra Y, Shang L. Inductive coding. In: Chandra Y, Shang L, editors. Qualitative Research Using R: A Systematic Approach. Cham, Switzerland: Springer; 2019:91-106.
54. Carrol JM. Five reasons for scenario-based design. In: Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences. 1999 Presented at: HICSS '99; January 5-8, 1999; Maui, HI p. 30-51 URL: https://ieeexplore.ieee.org/document/772890 [doi: 10.1109/hicss.1999.772890]
55. Aguinis H, Bradley KJ. Best practice recommendations for designing and implementing experimental vignette methodology studies. Organ Res Methods 2014 Aug 25;17(4):351-371. [doi: 10.1177/1094428114547952]
56. Go create. Unity Technologies. URL: https://unity.com/ [accessed 2025-02-01]
57. Valve Index. Valve Corporation. URL: https://www.valvesoftware.com/en/index [accessed 2025-02-01]
58. Picovoice: voice ai platform. Picovoice. URL: https://picovoice.ai/ [accessed 2025-02-01]
59. Htc Vive hand-tracking SDK. HTC Vive. URL: https://www.vive.com/us/support/focusvision/category_howto/hand-tracking.html [accessed 2024-11-26]
60. Microsoft Kinect v2. Microsoft. URL: https://www.microsoft.com/en-us/download/details.aspx?id=44561 [accessed 2025-02-01]
61. Microsoft Azure Kinect. Microsoft. URL: https://azure.microsoft.com/en-in/products/kinect-dk [accessed 2025-02-01]

Abbreviations

NUI: natural user interface
OR: operating room
SLS: surgical lighting system
VR: virtual reality

Edited by A Kushniruk; submitted 27.12.24; peer-reviewed by J Nosta, FZ Chelali; comments to author 14.02.25; revised version received 18.02.25; accepted 23.03.25; published 12.05.25.

Please cite as:
Zargham N, Reinschluessel AV, Mühlenbrock A, Muender T, Cetin T, Uslar VN, Weyhe D, Malaka R, Döring T
Using Gesture and Speech to Control Surgical Lighting Systems: Mixed Methods Study
JMIR Hum Factors 2025;12:e70628
URL: https://humanfactors.jmir.org/2025/1/e70628
doi:10.2196/70628
PMID:

©Nima Zargham, Anke V Reinschluessel, Andre Mühlenbrock, Thomas Muender, Timur Cetin, Verena Nicole Uslar, Dirk Weyhe, Rainer Malaka, Tanja Döring. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 12.05.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Experiences With a Multicomponent Digital Behavioral Pain Management Intervention for Adults With Sickle Cell Disease: Qualitative Analysis of the CaRISMA Trial

Olubusola B Oluwole[1], MD, MS; Lakeya S McGill[1], PhD; Kadeem Gayle[1], DMH; Julia A O'Brien[2], RN, PhD; Kaleab Z Abebe[1,3], PhD; Charles R Jonassaint[1], MHS, PhD; Megan Hamm[1,3], PhD

[1]Department of Medicine, University of Pittsburgh, Pittsburgh, PA, United States

[2]Department of Acute and Tertiary Care, School of Nursing, University of Pittsburgh, Pittsburgh, PA, United States

[3]Center for Biostatistics & Qualitative Methodology, Pittsburgh, PA, United States

Corresponding Author:
Olubusola B Oluwole, MD, MS
Department of Medicine
University of Pittsburgh
Medical Arts Building, 4th floor
3708 Fifth Avenue
Pittsburgh, PA, 15260
United States
Phone: 1 412 647 6412
Email: Oluwoleob2@upmc.edu

Abstract

Background: Chronic pain is prevalent among adults with sickle cell disease (SCD) and can be worsened by psychosocial factors such as depression and inadequate social support. Effective behavioral interventions (eg, cognitive behavioral therapy [CBT]) exist for chronic pain in various populations; however, few have been developed to address chronic pain in SCD. Several barriers have restricted the development and dissemination of CBT pain interventions in SCD, such as limited accessibility and time constraints. Digital interventions provide accessible and cost-effective pain management tools, offering self-management strategies, real-time monitoring, and personalized treatment options. Yet, there are limited data regarding patients' experiences with such interventions within the SCD population. The Cognitive Behavioral Therapy and Real-Time Pain Management Intervention for Sickle Cell Via Mobile Applications (CaRISMA) trial evaluated the effectiveness of a digital CBT intervention compared with a digital educational intervention for pain management in SCD. Evaluating participants' experiences can guide refinement of digital pain interventions in SCD.

Objective: This study aimed to gain a deeper understanding of the lived experiences of participants in the CaRISMA trial and to determine how to better adapt this intervention to the SCD population. The study examined individuals' overall experience with the trial and their perspectives of the trial components: a health coach, a chatbot-delivered digital CBT program, and an electronic pain diary.

Methods: Respondents were randomly selected to participate in semistructured interviews at (1) baseline, (2) the end of the intervention period at 3 months, and (3) the postintervention time point at 6 months or beyond. Interviews were audiotaped, transcribed verbatim, and analyzed using conventional content analysis.

Results: A total of 48 participants (women: 33/48, 69%) completed the interviews, with 24 and 19 completing midpoint and postintervention interviews, respectively. Participants generally had a positive experience in the trial. Many found value in learning about the connection between pain and mental health, considering it an important aspect of their well-being. The health coach played a key role in offering personalized support and guidance. Although the chatbot reinforced pain management strategies, its usefulness and engagement varied based on participants' prior knowledge of SCD. The pain diary helped increase self-awareness of pain patterns but was perceived as tedious and irrelevant by those without current pain episodes.

Conclusions: This qualitative substudy of the CaRISMA trial showed that participants valued the personalized support of the health coach, education about the connection between stress and pain, and the self-reflection fostered by the pain diary. These findings highlight the potential of digital, patient-centered approaches to address the multifaceted needs of SCD care. For digital

XSL•FO
RenderX

interventions, the inclusion of personalized support with ongoing communication appears to be a critical component that can influence treatment adherence and effectiveness.

Trial Registration: ClinicalTrials.gov NCT04419168; https://clinicaltrials.gov/study/NCT04419168

International Registered Report Identifier (IRRID): RR2-10.2196/29014

(JMIR Hum Factors 2025;12:e73719) doi:10.2196/73719

KEYWORDS

behavioral; digital intervention; sickle cell disease; chronic pain; pain

Introduction

Sickle Cell Disease

Sickle cell disease (SCD) is a genetic disorder characterized by recurrent acute episodes of severe pain, commonly referred to as "pain crises," and chronic pain that can persist between these episodes. The unpredictable nature of SCD-related pain, combined with the physical and emotional burden of living with a chronic illness, leads to high rates of depression, stress, and reduced social support [1,2]. Additionally, stress and depression can further exacerbate pain, creating a cycle in which physical and emotional distress reinforce each other. Historically, pain interventions in SCD have focused on pharmacologic treatments, yet comprehensive biopsychosocial approaches that address both the physical and emotional aspects of SCD pain management remain limited. The 2019 American Society of Hematology guidelines for the management of SCD recommend a comprehensive approach to pain management, including pharmacologic treatments alongside behavioral health interventions that address psychological and social contributors to pain [1]. Despite these recommendations, access to integrated biopsychosocial care remains limited, leaving many individuals with unmet behavioral health and psychosocial coping support needs [3].

Cognitive Behavioral Therapy for Chronic Pain

Cognitive behavioral therapy (CBT) is a first-line, evidence-based treatment for chronic pain endorsed by the American College of Physicians and the United States Centers for Disease Control and Prevention [4,5]. CBT effectively improves pain outcomes in several chronic conditions, including SCD [6-8], by addressing psychosocial contributors to pain such as maladaptive thought patterns, promoting behavioral activation, and equipping individuals with coping strategies. CBT has also been linked to improved emotional regulation, self-efficacy, and health-related quality of life in patients with chronic pain.

Digital Interventions for SCD

Despite its effectiveness, access to CBT has been limited by barriers such as lack of transportation, costs, provider shortages, and scheduling conflicts. Digital behavioral interventions offer a promising avenue to expand access to behavioral health support, especially for individuals with chronic illnesses like SCD who face barriers to in-person care; however, the use of digital interventions in SCD has been limited. Digital CBT programs have demonstrated effectiveness for improving pain-related outcomes, coping skills, and mental health across chronic pain populations. Furthermore, when compared with non-CBT mobile apps, CBT-based apps demonstrated greater efficacy for behavior change than non-CBT apps, likely due to structured use of CBT techniques, engagement in cognitive reframing tasks, and empowerment of users [9]. However, there is limited evidence on the effectiveness and user experiences with digital behavioral health programs specifically designed for individuals with SCD [6,10-13], a population with unique medical, cultural, and psychosocial needs. Individuals with SCD report mixed experiences with available digital health interventions: Some found these tools helpful for learning coping strategies and tracking symptoms, while others noted limitations such as insufficient personalization or a lack of cultural relevance to the unique challenges of SCD [14]. In the context of SCD, where pain is deeply intertwined with stress, mood, and self-efficacy, digital CBT offers a targeted and scalable tool for helping individuals manage pain both psychologically and behaviorally, but more high-quality clinical trials of digital CBT interventions are needed.

Initial Development of a Digital CBT Intervention: Beginning With Addressing the Challenges in SCD Clinical Trials

Clinical trials in SCD face significant recruitment and retention challenges. Insufficient patient participation is a major impediment to the advancement of SCD research, with many studies terminating due to low enrollment [15,16]. A systematic review of 174 SCD clinical trials found that 14% were terminated specifically for failure to enroll participants, with only 35% of completed studies resulting in published manuscripts [17]. These challenges can be particularly pronounced for behavioral interventions, which require consistent engagement and are time-intensive for the participants. Our early research aimed to simultaneously address enrollment challenges and develop an accessible intervention by prioritizing people with lived experience. In collaboration with the SCD community, our group developed a digital CBT app adapted for adults with SCD. To explore the feasibility and preliminary impact of digital behavioral health tools in SCD, our group previously conducted a pilot randomized controlled trial with adolescents and young adults evaluating an earlier version of our CBT app compared with a CBT app for the general population [18]. The study demonstrated improvements in pain, self-efficacy, and depressive symptoms, highlighting the potential for digital interventions to address the biopsychosocial needs of this population. However, about 40% of participants did not engage with the digital interventions, revealing significant barriers to sustained participation.

XSL•FO
RenderX

Prior Work Refinement and Expansion of the Digital CBT Intervention: Further Addressing Engagement Challenges

To better understand the determinants of digital intervention engagement, we previously conducted a qualitative study using focus groups of adolescents and adults with SCD [18]. Participants identified 5 key engagement factors, including the importance of peer connection, personalized content and coaching, preferences regarding the coach's background (eg, a coach with lived SCD experience), the value of journaling and symptom tracking, and flexible engagement options [13]. These insights, along with data from our pilot randomized controlled trial, helped inform the design of the Cognitive Behavioral Therapy and Real-Time Pain Management Intervention for Sickle Cell via Mobile Applications (CaRISMA) trial. We randomized participants to 1 of 2 health coach-supported interventions: digital CBT versus digital patient education [19]. Participants also completed pain diaries during the study period. All components were designed to align with the preferences and barriers identified in our prior work, with the goal of creating a flexible, patient-centered digital platform for SCD pain management. Peer support delivered by health coaches was a core component of the intervention, as research suggests that peer-delivered models can build trust, improve engagement, and address barriers to mental health care in historically underserved populations [20-24]. In the primary outcomes study, we found that both the CBT and education interventions led to an improvement in pain interference; however, there was no significant difference between the arms [19].

This study focuses on participants' experiences in the CaRISMA trial and with the various components of the digital CBT program. Most of the previously published studies on digital interventions focus on single-component interventions [25]. Multicomponent, app-based interventions are less common but emerging [26]. The CaRISMA trial incorporated a multicomponent, culturally tailored platform that included peer-supported digital CBT, symptom tracking, and personalized chatbot engagement. Although there was no difference between groups, we found that digital CBT leads to improved pain-related outcomes [19]. Therefore, a critical next step is to better understand participants' experiences with the intervention components to in order to refine future approaches and enhance the likelihood of successful implementation in real-world settings.

Objective of This Study

This qualitative study aimed to gain a deeper understanding of participants' experiences with the CaRISMA trial, focusing on their perceptions of the health coach, chatbot, and pain diary components. By identifying key factors influencing engagement and satisfaction, we sought to better understand barriers to trial enrollment and retention in the SCD population, evaluate the potential for digital behavioral interventions to address this population's unique needs, and inform the implementation of culturally appropriate, patient-centered digital health interventions for individuals living with SCD.

Methods

Study Design and Setting

The protocol for CaRISMA has been previously published [27]. The CaRISMA trial enrolled adults with SCD who reported chronic pain (ie, pain at least 4 days a week over the past 3 months or longer). Participants of the CaRISMA trial were recruited from 7 comprehensive sickle cell centers and from 4 community-based organizations, either virtually or in person. Participants were randomized 1:1 to either the digital CBT or Education arm. Both arms had access to a health coach and a chat system via a chatbot. All participants were asked to complete daily pain assessments via an electronic pain diary app.

For this qualitative study, we aimed to recruit a diverse group of trial participants, stratifying our sample by depression severity assessed using the 9-item Patient Health Questionnaire (PHQ-9; PHQ-9≥10 vs PHQ-9<10) and trial arm (CBT vs Education). Trial participants were randomly selected for 1-to-1 semistructured interviews at baseline. We sought 24 interview participants from each study arm (CBT or Education), and within each of those 2 sets of 24 participants, we sought 12 participants with high depression severity (PHQ-9≥10) and 12 participants with low depression severity (PHQ-9<10). If a randomly selected participant declined to participate in the baseline interview, another participant was randomly selected until we had achieved a total of 48 interviews. This target was chosen based on a strong likelihood of achieving thematic saturation, the point at which additional interviews are unlikely to provide new insights [28,29]. Participants who completed baseline interviews were invited to participate in follow-up interviews at midtrial (3 months, end of intervention) and post-trial (between 6 months and 12 months) time points.

Intervention Components

Digital CBT

The digital CBT arm included a 12-week, tailored program designed for adults with SCD. It taught participants how to identify negative thoughts and emotions, use cognitive skills and problem-solving techniques, and apply effective coping strategies. The program focused on skill development through practice, homework assignments, and regular check-ins with a health coach.

Digital Education

The digital Education arm included comprehensive information about SCD and pain management through an interactive digital platform. Participants learned about chronic pain, lifestyle modifications, and SCD fundamentals while reinforcing their knowledge through interactive assessments. The program encouraged active learning by enabling users to discuss concepts with their health coach.

Health Coach

Health coaches—many of whom had personal connections as SCD patients, caregivers, or advocates—offered weekly support through phone calls or text messages, depending on participant preferences. Their primary role was to provide emotional support

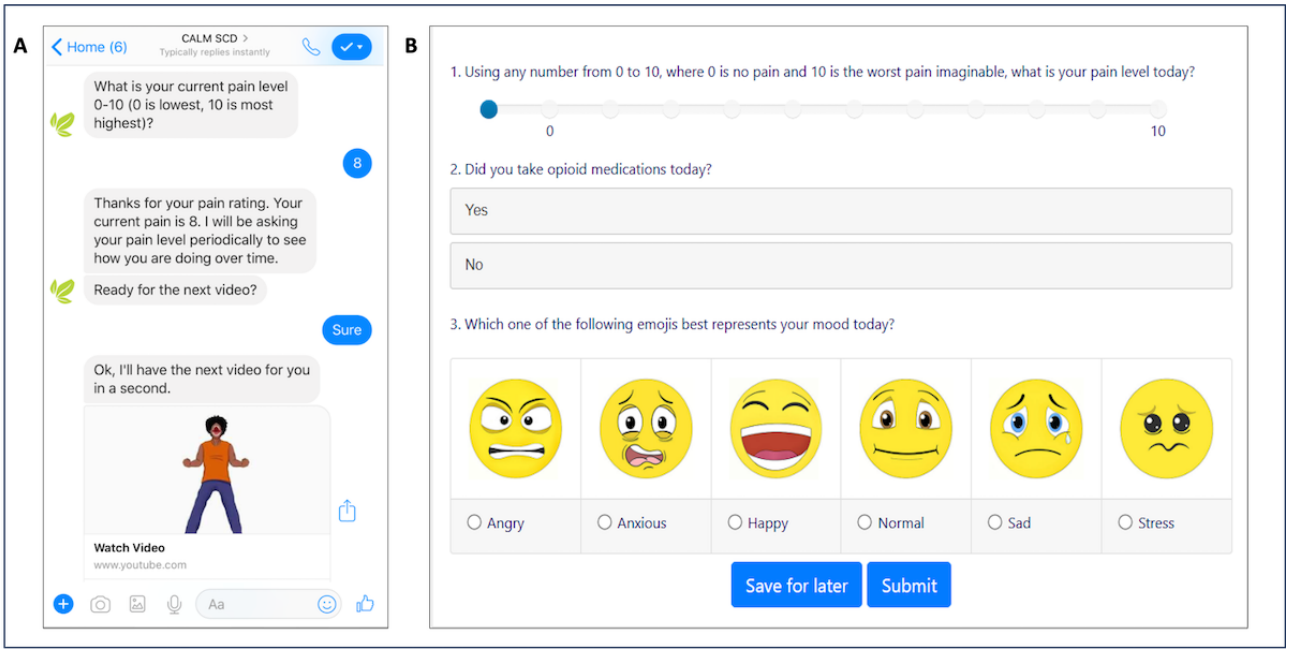
XSL•FO
RenderX

and encourage active engagement with the digital platform. Each coach underwent comprehensive training and attended weekly group supervision sessions facilitated by a mental health professional and research lead.

Chatbot

The CaRISMA interventions were delivered by a scripted chatbot (Figure 1), which was developed in collaboration with patients, families, and community partners. The chatbot was accessed through the Facebook Messenger app and offered tailored responses, educational materials, and motivational content in various formats including videos, GIFs, and images, all adapted to participants' inputs. It also provided round-the-clock conversational support, adapting responses to each participant's unique circumstances and needs. Users were able to monitor their progress through the program, revisit materials they had completed, and view their accomplishments via progress metrics and achievement rewards.

Figure 1. Cognitive Behavioral Therapy and Real-Time Pain Management Intervention for Sickle Cell via Mobile Applications (CaRISMA) components including (A) chatbot interaction and (B) pain diary tracking.



Pain Diary

Patients were asked to track pain levels daily through a mobile web app, as illustrated in Figure 1. They received text reminders during 2-week assessment windows at the start of the study and at 3 months, 6 months, and 12 months. Between these formal tracking periods, they were encouraged to keep logging their pain scores in the app.

Qualitative Analysis

Analytic Approach

We used a qualitative descriptive approach for data collection and analysis, aiming to directly capture and summarize participants' experiences, thoughts, and feelings related to the interview topics. This approach is frequently used in health and medical research to provide clear and practical insights [28,30]. Here, we sought a holistic understanding of SCD patients' prior experiences with pain, depression, and treatment for either or both conditions.

Our interview guides for baseline, midpoint, and final interviews (Multimedia Appendix 1) were designed to explore participants' experiences throughout the trial. Baseline interviews focused on participants' previous experiences with SCD pain management and their initial impressions of the study. Midpoint interviews gathered feedback on their experiences partway through the trial and any changes they had noticed. The final time point interviews sought information on durable effects and knowledge gained from their trial experiences.

Data Collection and Coding Procedures

Data collection and analysis were conducted by skilled qualitative researchers at the Center for Biostatistics and Qualitative Methodology's Qualitative Core, under the supervision of a qualitative methodologist. Once a randomly selected CaRISMA participant agreed to be interviewed, the interview was scheduled at a time that was convenient for them. To accommodate CaRISMA participants who were located across multiple geographic locations, interviews were conducted telephonically or via teleconferencing software according to the interviewee's preference. Interviews were audio-recorded and transcribed verbatim, with identifying details redacted. Although we hoped to have one consistent interviewer across the time points, the duration of the study and the logistics resulted in interviews being conducted by 3 interviewers, all of whom received the same training.

Concurrent with data collection, interviewers completed a summary template after each interview to document key topics discussed and capture emerging themes. These summaries allowed the team to assess thematic saturation, which was achieved. A trained qualitative analyst inductively developed an initial codebook based on the content of the interviews. This draft codebook, complete with detailed code definitions, was

XSL•FO
RenderX

circulated to the study team to ensure that it reflected relevant topical and theoretical themes important to the study team, as well as the content of the interviews. The codebook was slightly revised to incorporate new codes specific to themes emerging in the midpoint and final interviews. Once codebooks were finalized, qualitative coding was completed with the assistance of MAXQDA 2022 software. For each data set, 10 transcripts were co-coded and then reviewed by 2 independent, trained qualitative coders from the Center for Biostatistics and Qualitative Methodology to ensure quality and consistency in coding. Any coding discrepancies identified were adjudicated by the coders to full agreement. The primary coder completed the coding of the remaining transcripts following the standards established in the codebook and refined through the adjudication process. This coding served as the basis of a conventional content analysis [18], in which the primary coder prepared summaries describing participants' responses related to each interview guide domain at each time point. These summaries were reviewed by the qualitative methodologist and other study team members to identify the most relevant findings for reporting. Additional manuscripts are planned to present findings related to pain management and depression. This manuscript focuses specifically on participants' experiences with and perceptions of the CaRISMA trial.

Ethical Considerations

This study was approved by the University of Pittsburgh Institutional Review Board (STUDY20110346). Informed consent for this study was included in the original consent process for participation in the CaRISMA trial. No compensation was provided for this portion of the study. Participants were compensated for their broader participation in the CaRISMA trial. All study data were stored on secure University of Pittsburgh servers, accessible only to authorized members of the study team. Specifically, access was limited to members of the Qualitative Core who conducted the analysis. Audio files and interview transcripts were labeled with unique study identifiers and did not include participant names or other direct identifiers. All transcripts were de-identified, with personally identifying information redacted during the transcription process

Results

Participant Characteristics

A total of 48 participants completed baseline interviews. Of those 48 participants, 24 completed a midpoint interview, and 19 completed a final time point interview following the conclusion of their interactions with CaRISMA. At baseline, the mean age of participants was 36 (SD 9.7) years, and the majority were women (33/48, 69%). Most participants were recruited through clinic visits (38/48, 79%). About two-thirds (33/48, 69%) had attended some college or higher education, while 31% (15/48) had some high school education or a high school diploma or equivalent. One-third of participants (16/48, 33%) were employed, and about one-half of participants (25/48, 52%) reported being on disability. The demographic and clinical characteristics of participants at the baseline, midpoint, and final interview time points are summarized in (Table 1).

Here, we focus on participants' experiences in the trial, focusing on the unique aspects of the CaRISMA program, including the health coach, chatbot, and pain diary.

XSL•FO
RenderX

Table 1. Baseline demographic and clinical characteristics of study participants across time points.

Characteristic	Baseline (N=48)	Midpoint (n=24)	Final (n=19)
Age (years), mean (SD)	36.1 (9.7)	35.7 (10.8)	34.2 (10.8)
Sex, n (%)			
Female	33 (69)	8 (33)	6 (32)
Male	15 (31)	16 (67)	13 (68)
Race, n (%)			
Black/African American	45 (94)	23 (96)	19 (100)
Multiple/other/White	3 (6)	1 (4)	0 (0)
Education, n (%)			
Some high school	3 (6)	1 (4)	1 (5)
High school or equivalent	12 (25)	5 (21)	6 (32)
Some college	23 (48)	13 (54)	8 (42)
Completed college	6 (13)	3 (13)	3 (16)
Graduate studies	4 (8)	2 (8)	1 (5)
Employment status, n (%)			
Employed	16 (33)	7 (29)	8 (42)
Unemployed	7 (15)	4 (17)	2 (11)
On disability	25 (52)	13 (54)	9 (47)
PROMIS[a] pain interference score, mean (SD)	62.2 (7.2)	60.6 (6.9)	60.4 (7.8)
Average pain intensity, mean (SD)	4.0 (2.5)	3.7 (2.3)	3.5 (2.5)
Pain episode frequency, mean (SD)	47.6 (14.3)	48.1 (13.2)	46.5 (13.7)
ASCQ-Me[b] measures, mean (SD)			
Social functioning impact	46.9 (8.2)	47.5 (8.0)	47.8 (9.8)
Emotional impact	48.0 (10.0)	49.3 (9.3)	50.2 (11.0)
Number of pain diary entries (during the 2-week reminder period), mean (SD)	9 (4)	10 (4)	8 (4)
Depressive symptoms (PHQ-9[c]), mean (SD)	10 (5)	9 (4)	10 (5)
Depressive symptom severity, n (%)			
Minimal (0-4)	5 (12)	4 (19)	2 (13)
Mild (5-9)	17 (40)	8 (38)	7 (44)
Moderate (10-14)	8 (19)	5 (24)	3 (19)
Moderately severe (15-19)	12 (28)	4 (19)	4 (25)
Severe (20-27)	1 (2)	0 (0)	0 (0)
Brief PCS[d], mean (SD)	9 (4)	9 (4)	8 (4)
GAD-7[e], mean (SD)	9 (5)	7 (4)	8 (5)
SCSES[f], mean (SD)	29 (6)	29 (5)	29 (6)

[a]PROMIS: Patient-Reported Outcomes Measurement Information System.

[b]ASCQ-Me: Adult Sickle Cell Quality of Life Measurement Information System.

[c]PHQ-9: 9-item Patient Health Questionnaire.

[d]PCS: Pain Catastrophizing Scale.

[e]GAD-7: 7-item Generalized Anxiety Disorder.

[f]SCSES: Sickle Cell Self-Efficacy Scale.

XSL•FO
RenderX

Overall Trial Experience

Most participants reported a positive overall experience with the CaRISMA study, emphasizing that it provided a "safe space" to learn and express themselves without fear of judgment. One participant noted:

> *I enjoy doing it. I really enjoyed—like I said, I learned a lot. And I like the fact that I can say how I feel in a safe, safe space. [...] the thing I think I appreciated the most about the study was, no matter what I say, good, bad, or indifferent, it is appreciated, and I felt safe. I can say it exactly how I feel.* [P31]

Others appreciated the study's focus on the connection between chronic illness and mental health, noting that such an approach was rare but much needed. One participant noted:

> *I feel like the study went well. [...] The fact that you guys thought of doing something like this to think on mental health and just treating symptoms differently, I felt like was great for the community because a lot of us did not know anything about CBT.* [P68]

Additionally, several participants found the exploration of the link between pain and depression particularly valuable, highlighting the need for greater awareness of this relationship within the sickle cell community. One participant noted:

> *Yeah, it was good overall. I think that I was quite impressed that there's something being studied about, you know, like taking care of the relationship between chronic illness and mental health. I hadn't seen that before so that was a good one for me. Yeah, so it was good generally.* [P107]

Despite these positive aspects, participants also expressed a desire for more follow-up and reminders throughout the study. Many felt that regular check-ins and clearer communication could have improved their engagement and overall experience:

> *Yeah, I feel like that's, that's, that would help the study be so much more successful or everybody's feeling on the same page, just because — I mean, not saying that it's not successful now, but I feel like, you know, if you have people checking me in frequently and letting you know, hey, you know, you had this option, this option, this option, you maybe should try it out, see if this helps, whatever. It was just it was a lack of communication.* [P104]

Participants expressed mixed preferences regarding in-person versus social media–based engagement, with some valuing the sense of community and real-time connection that in-person meetings offer, while others appreciated the convenience and accessibility of digital platforms. One participant emphasized the benefits of in-person interactions, stating "It would be nice to be in-person, you know, we all get together now" (P110). Another noted the challenges of an in-person engagement:

> *Oh, I like in-person. Um, the only thing is, you know, distance, transportation, things like that, schedules of people that are working, you know. But in-person is always a good one, because you, for me, the most important thing about any person is a sense of community, um, and connection to someone else in real-time, right. But other than that, I think on, like, a walk through a study, because, you know, scheduling conflict, and time and transport and all those things might make in-person maybe a bit more difficult.* [P107]

Participants also had mixed opinions on the most effective social media platform for engagement, with some favoring Snapchat due to its artificial intelligence chatbot feature, which they believed could enhance interactions and provide a more conversational experience. Others preferred Instagram, citing its widespread use and strong messaging capabilities, suggesting that a more active presence with regular updates and resources could increase engagement. Another participant recommended a separate app for the study. If a separate app is not feasible, then Instagram is the next best option:

> *Yeah, I, I think, I think the app will be the best route. But if, if you couldn't, if, if it was added on other social media platform, I feel like Instagram will be a real, a real good one, because I think Instagram is probably more widely used, especially when it comes to, like, messaging. Um, so I, I think it will be, yeah, I think you'll get more of a reach on Instagram. Because I've been on CaRISMA Instagram page, and, I mean, it's, it's okay, but it don't, it don't really have [...] life on it. You know, that, that will be the only thing that they will have to do is just bring more life to the page.* [P68]

Others saw value in leveraging existing platforms to reach a broader audience. Overall, participants highlighted the importance of selecting a platform that aligns with user preferences while ensuring consistent engagement and accessibility.

Health Coach Interactions

Participants appreciated that the health coach had SCD *or* knew someone (eg, a family member) who has SCD. Participants thus felt understood by the health coach. For instance, one participant said, "I just like to be able to talk to my coach, because she's going through what I'm going through. So, it's, it's good to be able to talk to somebody" (P30). Another participant similarly said, "I loved it [the health coach]. It was like, you know, talking to somebody that I can, you know, relate to. Um, that was, that was a great thing. You know? Easy to communicate. You know" (P40).

Talking to someone who was knowledgeable proved to be very helpful. This led to enduring effects beyond the health coach interactions, with participants noting that they learned the importance of talking and reaching out to someone when they are upset or in pain. In at least one case, a participant sought counseling after completion of the CaRISMA program, as a result of learning how helpful it could be to talk to someone:

> *There is one thing that we did choose to implement, which was, you know, just kind of self-care and mental health check-ins, and that's something that on the side of after speaking with my CaRISMA coach about that, that's something that I decided to sign myself*

XSL•FO
RenderX

> *up for. So, I signed up for counseling through my church and we have little weekly check-ins; so that's been very helpful. [...] So yes, I, I would, I would think that, that was definitely was something that, that's helpful in the event that there is a depression or anxiety episode that rises.* [P53]

Participants also appreciated that the health coach was willing to be flexible when scheduling meetings, as the health coach understood the unpredictability of SCD. For instance, one participant said,

> *I feel like she's [...] very helpful. She's, uh, very, like, mindful of my time and my, my situation, like she understands that there are some days where I could just really just be ready to collab and have a long conversation. And then there are other days where I may not feel well and I'm not interested in really just talking on the phone, especially if I'm in the hospital or at home trying to relax. So, that's why I like; she comes in and texts and if, if we scheduled an appointment and that's not good day, she's very flexible. So, uh, I kinda like how everything is going with, with the coach that I have.* [P53]

Some participants lamented limited interactions with their health coach, wishing that they had more time to get to know them. For example, one participant said, "I kinda wish I had the, more time to talk to her more, so we could kinda get like a better understanding of each other" (P67). The participant added that, ideally, they would like to speak to the health coach "every week because with sickle cell stuff can change so quickly, so you never know what I could be doing this week or what could happen next week. So, I would say, like, every week but I know, like, schedules and work and school and different stuff puts a burden on me trying to check in with her every week. So, I mean, text messages is fine, but conversation is different than texting, so, that's what I mean" (P67).

Chatbot Use and Feedback

Most participants found the chatbot to be helpful, barring technical difficulties. Participants primarily enjoyed the chatbot if they gained new insights from it. Some participants said that the information they learned helped to reframe the way they think about their SCD. One participant found this to be an enlightening experience:

> *I actually learned something new. I didn't think I was going to learn anything. [...] So we know that the sickle cells are crescent shaped. [...] Well, I learned that different types of, you know, genotypes sometimes the sickle cells are shaped differently. Because I have [inaudible] mine are more shaped like crystals instead of crescents. So that was new. I was like what! And I love crystals [laugh]. You know, they kind of made me like the sickle cell a little bit. Like okay, girl, we're twins! [...] I really did not know that. I was like oh my God I've had sickle cell for 34 years and I did not know that.* [P82]

Some participants found the chatbot helpful for reinforcing pain management strategies, while others felt it provided limited new information for those already knowledgeable about their condition. For participants who gained new insights, the chatbot was a valuable learning tool.

Some participants also noted that the insight they gained challenged certain "preconceived notions" they held about SCD:

> *Um, yeah, there's been...um, information that I...probably didn't think about...before. On that. And, that you had, um — you know, it's just like you think this is logically, this is certain information you had in your head about stuff, and you find out that, oh, my god, that they makes sense, this, this — do you understand what I mean? [...] Okay. So, if I had a preconceived notion on a particular aspect of sickle cell; I don't know, if for any reason like…it could be like an old wives' tale. And then, on looking, going through that I get a different information and I'm like, oh, that makes sense; why did I ever think this is what it was? Yeah, that's what I mean.* [P69]

In contrast, participants who did not find the chatbot to be helpful reported that it did not teach them anything new. The participant explained that the chatbot was more useful for "people who [...] don't know much about their own disease":

> *To be honest, I think it was more unhelpful to me, because the questions that it asked are, are questions that I already know the answer to. Does that make sense? Like, um, some people don't know what kind of sickle cell they have, or they don't know their triggers or, or things like that. And my mom taught us growing up, like, to be able to speak for yourself when you went to the doctor, so I know what kind of sickle cell I have. I know — you know what I'm saying — my triggers, I know, um, my medications and the dosages and stuff like that. So, it would be helpful for some people who aren't really, I guess who don't know much about their own disease. Does that make sense? But I...I think I know my disease and my body pretty well to where that was, it wasn't helpful to me.* [P80]

Frequently, participants who did not use the chatbot suggested that it should take a more proactive role in encouraging participants to interact with it, rather than relying on users to engage on their own. For example, when asked what would improve their use of the chatbot, one participant asked for a reminder: "Train the bot to just kind of reach out and be, like, 'Hey, you haven't answered the question or asked a question in a couple days; is everything good or would you like to resume'-type thing" (P39). The need for reminders was particularly salient for anyone who did not regularly interact with Facebook Messenger outside of the study and might therefore not see it in the course of their day.

Use of the Pain Diary

One participant added that filling out the pain diary—and being able to show the pain diary to their health care providers—was very helpful:

> *And also, to see that part in the, in the study where they was asking, like, how you feel, like that was, that's something I wish was everywhere, honestly; I*

XSL•FO
RenderX

> *wish I had that everywhere. I wish you could, like, show the doctors how you feel and whatnot, because even if they did put it everywhere, I feel like if it was standardized, they probably would know how to deal with different pain levels, you know, and, and that would probably be able to help them structure, um, medication doses. Like I feel like that would be so helpful cuz they can see it, and it's not more of you trying to struggle and explain it to them; they already see it; they know exactly how this is affecting you. This is, this is what works for them, many other warriors on this level; let's give them the same — and I'm telling you, I feel like it would be so helpful. I feel like that needs to be implemented in every hospital for every, every person just — or not anybody. Um, but yeah, I, I, I think that was probably the, the biggest thing for me.* [P68]

Some participants found completing the pain diary daily redundant or tedious, particularly on pain-free days, leading them to develop their own strategies to manage this task. One noted:

> *To me personally because it would come every day. I would get the email for it every day. And at a certain point it felt redundant if I wasn't in pain for a few days or weeks at a time. Having to fill it out every day. So it got to the point where if I had pain I would say okay yeah I'll go fill it out now. Instead of every day putting no pain.* [P115]

Some participants appreciated completing the pain diary daily, even on pain-free days, as it allowed them to track their symptoms and emotions. It also helped them reframe their experience, reminding them that not all days are "bad" and that pain episodes eventually subside. For example, one participant stated:

> *I mean, I like the pain diary because then it shows me, like, you know, what I write. Pretty much write down everything on my own. To do it every day just to know okay I'm feeling bad this day but this day I was okay. So that kind of helps.* [P104]

Participants also appreciated having a record of their pain and pain medication usage. As one participant described it:

> *Not that I enjoy, but I do look forward to the daily pain diary; it just, it does make me stop and think about, you know, what pain I'm in and have I taken meds and do I, you know, what can I do to not decrease but to minimize the pain and, in turn, minimize, I guess, the amount of medication. So that is helpful.* [P58]

Thus, participants felt that the pain diary helped to increase their awareness of their pain and feelings, helping them make more informed decisions about their pain management strategies.

Representative quotes are summarized by intervention component in Table S1 in Multimedia Appendix 2.

Discussion

Principal Findings

Digital health interventions have emerged as a promising tool to enhance chronic pain management and coping. Based on feedback from people with lived experience, we designed the CaRISMA trial to equip adults with SCD with the tools, resources, and support systems necessary to improve long-term chronic pain outcomes. In the primary CaRISMA trial, we found that both digital interventions (CBT and Education) resulted in improvements in the primary pain outcome, pain interference, at 6 months [19]. Additionally, greater engagement with health coaches was associated with greater improvements in pain interference. The qualitative study reported here enhances our understanding of the first peer-supported digital CBT intervention for adults with SCD by examining participant experiences with its key components: the health coach, chatbot, and pain diary.

Overall Experience

Overall, participants reported a positive experience with the CaRISMA trial, appreciating the safe space provided for open discussions and the opportunity to learn about the connection between pain and mental health. Although some participants preferred face-to-face interactions, logistical barriers such as transportation and health limitations reinforced the value of flexible digital formats, consistent with findings in other chronic disease populations [31,32]. Regarding the delivery platform, some participants valued engagement via social media (eg, Facebook, Instagram), appreciating its convenience, accessibility, and ability to provide on-demand support. Participants also had varying opinions on which social media platform would be most effective, with some favoring Snapchat due to its chatbot feature, while others preferred Instagram for its broad user base and messaging capabilities. Given the evolving nature of social media, leveraging multiple digital platforms may enhance engagement by catering to diverse communication preferences. Alternatively, considering the shifting landscape of social media usage, consolidating resources into a single, dedicated app, separate from social media, could provide a centralized and streamlined user experience. These findings also highlight the importance of offering flexible, accessible digital health interventions that accommodate the unique needs, preferences, and logistical challenges faced by individuals living with chronic conditions like SCD.

Health Coach Experience

Participants strongly valued their health coach's lived experience with SCD, highlighting the need for culturally concordant and condition-specific health coaching. In other chronic disease populations, tailored peer support has contributed to enhanced trust, self-efficacy, satisfaction, engagement, and adherence by offering emotional support, shared experiences, and practical guidance [20,21,33]. In the CaRISMA trial, many participants appreciated that their health coach had SCD or personal connections to someone with SCD, which contributed to feeling understood and supported. Some participants also reported that interactions with their health coach motivated them to seek mental health counseling, demonstrating the intervention's

XSL•FO
RenderX

potential for long-lasting impact that possibly extends beyond pain management.

Participants found the health coach relationship so valuable that many expressed a desire for more frequent contact. We designed the CaRISMA program to consist of weekly check-ins; however, the unpredictability of pain episodes for both participants and health coaches interfered with check-ins. Our findings are aligned with those of previous studies demonstrating that more frequent, flexible check-ins improve engagement and adherence in chronic disease populations [34,35]. To optimize future interventions, health coach check-ins should be frequent and adaptable, incorporating scheduled and on-demand support. A multimodal approach, incorporating text messages, phone calls, and video check-ins, could help meet individual needs.

Chatbot Experience

The chatbot received mixed responses; some participants found it helpful for reinforcing pain management strategies, while others felt it provided limited new information. These differences may reflect variation in prior knowledge, learning preferences, higher health literacy, or the degree to which participants engaged with other components of the intervention. Digital interventions incorporating chatbots have been recognized as an effective way to deliver behavioral coaching, track symptoms, and strengthen patient engagement [36]. Although some participants had a positive or eye-opening experience with the feature, others suggested that the chatbot should be more proactive, initiating conversations and providing reminders rather than waiting for users to engage. Previous research highlights that adaptive, artificial intelligence–driven chatbots that tailor responses based on user engagement and knowledge levels improve overall effectiveness in digital health interventions [37]. This approach could be incorporated into future digital interventions for adults with SCD.

Pain Diary Experience

The pain diary helped some participants track pain trends, but others found it redundant and unnecessary on pain-free days. Daily pain diaries are widely used in both SCD and broader chronic pain populations to monitor pain patterns and inform treatment strategies. In previous studies, patient experiences with these diaries have varied, reflecting both benefits and challenges. For instance, in the Pain in Sickle Cell Epidemiology Study (PiSCES), 60% of participants submitted fewer than 5 months of the expected 6 months of pain diaries, indicating challenges in sustained daily reporting [38]. In the CaRISMA trial, some participants mentioned that tracking their pain helped them gain perspective on their condition, allowing them to recognize that not every day was a “bad” day. This is a sentiment echoed in prior research on digital self-monitoring tools for chronic pain [14,39]. However, others recommended the use of the pain diary only when participants are experiencing pain. Research has shown that user fatigue is a common barrier to adherence in digital health interventions, particularly when data entry is required daily without immediate benefits [14,40,41]. To enhance usability, customizable pain tracking options allowing participants to self-select tracking frequency might be helpful. Additionally, integrating mood tracking and other symptom assessments alongside the pain diary may provide a more holistic view of well-being.

Comparison With Prior Work and Future Directions

The CaRISMA trial offered a novel approach to pain management in SCD by integrating a culturally adapted, multicomponent digital intervention within a historically marginalized and medically underserved population. Our focus on a historically marginalized group with specific cultural needs contributes new evidence on tailoring digital behavioral interventions for high-need populations. Unlike previous digital health trials that often focus on single-component interventions, the CaRISMA trial evaluated a comprehensive model combining digital CBT, peer support via health coaches, symptom monitoring, and chatbot technology. This study also highlights the importance of tailoring digital behavioral health interventions to the unique needs of individuals with SCD. The findings suggest that a multicomponent approach can enhance engagement and effectiveness. Future trials should consider the following recommendations: (1) enhanced health coaching that involves frequent and flexible coach check-ins, including on-demand support, to improve participant engagement; (2) an adaptive and interactive chatbot that offers personalized education based on an individuals’ knowledge, needs, and preferences (eg, chatbot with adjusted tone and content based on patient-reported depression and pain severity or providing layered content depending on the user’s baseline familiarity with SCD education and management); and (3) flexible and customizable self-monitoring tools, such as pain diaries that allow users to adjust tracking frequency, skip entries on low-pain days, or annotate with additional symptoms and reflections, as well as view visual summaries of pain trends to support provider communication and self-awareness. Future research should also develop and test digital interventions that incorporate other therapeutic approaches, such as acceptance and commitment therapy [42,43] and mindfulness-based stress reduction [44,45]. Digital interventions may also offer a form of cognitive escapism, a concept referring to psychological relief and distraction from distressing symptoms. For instance, Anto et al [46] found that digital tools can help users disengage from pain and emotional burdens through structured, engaging content. Within the context of SCD, a digital CBT program may provide both therapeutic and escapist benefits, helping individuals better manage their symptoms while finding temporary relief from the psychological weight of chronic illness, which may be relevant to coping and engagement. These processes should be further investigated in adults with SCD.

Strengths and Limitations

This study has several strengths. First, to our knowledge, this is the first qualitative study to explore patient experiences with a peer-supported digital CBT intervention tailored for chronic pain in adults with SCD. Second, it fills a gap in understanding SCD patients’ perspectives on digital pain tools by examining multiple intervention components (health coach, chatbot, and mobile pain diary) and how each contributed to engagement and self-management. Third, by evaluating participants’ willingness and ability to use a mobile pain-tracking platform in a real-world setting, the study provides valuable insights for

XSL•FO
RenderX

the design and implementation of future digital health interventions. Finally, this study captured real-world barriers and facilitators to digital health engagement in a medically underserved, high-burden population with specific, actionable feedback regarding reminder frequency, chatbot utility, and diary fatigue that can inform digital health optimization.

However, the study also has limitations. First, this study does not account for participants' underlying mental health conditions, which may have influenced their responses to the interventions. Second, although participants were randomly selected for interviews, selection bias remains possible, as only 48 of the 115 individuals contacted agreed to participate. Third, attrition across interview time points may have introduced bias, as participants who remained engaged may differ systematically from those who dropped out. Although no clear patterns were observed among nonresponders, those who disengaged may differ in unmeasured ways from those who completed all interviews. This may reflect health issues, competing demands, limited perceived benefit, logistical barriers, or stronger motivation to remain involved. It is also possible that retention was skewed toward individuals with particularly strong feelings about the intervention or more positive experiences with the interventions who were more motivated to remain engaged. Although much of our feedback was positive, it is possible that more critical voices did not engage, although we have no data to support this speculation. Last, most participants had at least some college education, which may limit generalizability to individuals with lower educational attainment. As participants were also recruited from specialty centers and community-based organizations, findings may not reflect the broader SCD population

Conclusions

Overall, participants found the CaRISMA program to be a supportive space for learning, self-reflection, and symptom management. Digital interventions offer a scalable and accessible approach to addressing pain care disparities within the underserved SCD population. Regular, proactive communication emerged as essential for participant engagement, with inconsistent interaction likely contributing to dropout rates in this and similar studies. Future digital interventions should prioritize regular communication among participants, health coaches, and research teams to enhance retention and overall effectiveness. Additional research should focus on developing culturally appropriate, patient-centered, and adaptable interventions that address the diverse needs within the SCD population.

Acknowledgments

This study would not have been possible without participation and support of the individuals living with sickle cell disease who participated in the Cognitive Behavioral Therapy and Real-Time Pain Management Intervention for Sickle Cell via Mobile Applications (CaRISMA) trial. We are also grateful to Andrea Ball for her assistance with manuscript editing. This research reported in this publication was funded through a Patient-Centered Outcomes Research Institute (PCORI) Award (CER-2018C2-13320; principal investigators: CJ and KZA).

Data Availability

Due to the sensitive nature of the data, the data sets generated and analyzed during this study are not publicly available. However, de-identified numeric data are available from the corresponding author upon reasonable request.

Authors' Contributions

This qualitative study was conceptualized by MH and CJ. MH contributed to the methodology, investigation, data collection, and formal analysis of the study. OBO, MH, LM, and KG reviewed the transcripts and emergent themes to develop the thematic framework and manuscript outline. OBO and MH led the drafting of the original manuscript with contributions from LM. All authors were involved in writing, reviewing, and editing of the final version.

Conflicts of Interest

CJ has ownership in the company that develops the Painimation app, one of the exploratory outcome measures for the trial. The remaining authors have no conflicts of interest to declare.

Multimedia Appendix 1
Interview guides.
[DOCX File , 31 KB - humanfactors_v12i1e73719_app1.docx]

Multimedia Appendix 2
Representative Quotes by Intervention Component (Health Coach, Chatbot, and Pain Diary).
[DOCX File , 21 KB - humanfactors_v12i1e73719_app2.docx]

References

XSL•FO
RenderX

1. Adam S, Flahiff CM, Kamble S, Telen MJ, Reed SD, De Castro LM. Depression, quality of life, and medical resource utilization in sickle cell disease. Blood Adv 2017 Oct 24;1(23):1983-1992 [FREE Full text] [doi: 10.1182/bloodadvances.2017006940] [Medline: 29296845]
2. Jonassaint CR, Jones VL, Leong S, Frierson GM. A systematic review of the association between depression and health care utilization in children and adults with sickle cell disease. Br J Haematol 2016 Jul 15;174(1):136-147. [doi: 10.1111/bjh.14023] [Medline: 26991317]
3. Brandow A, Carroll CP, Creary S, Edwards-Elliott R, Glassberg J, Hurley RW, et al. American Society of Hematology 2020 guidelines for sickle cell disease: management of acute and chronic pain. Blood Adv 2020 Jun 23;4(12):2656-2701 [FREE Full text] [doi: 10.1182/bloodadvances.2020001851] [Medline: 32559294]
4. Dowell D, Ragan KR, Jones CM, Baldwin GT, Chou R. CDC clinical practice guideline for prescribing opioids for pain - United States, 2022. MMWR Recomm Rep 2022 Nov 04;71(3):1-95 [FREE Full text] [doi: 10.15585/mmwr.rr7103a1] [Medline: 36327391]
5. Qaseem A, Wilt TJ, McLean RM, Forciea MA. Noninvasive treatments for acute, subacute, and chronic low back pain: a clinical practice guideline from the American College of Physicians. Ann Intern Med 2017 Apr 04;166(7):514-530. [doi: 10.7326/m16-2367]
6. Sil S, Lai K, Lee JL, Gilleland Marchak J, Thompson B, Cohen L, et al. Preliminary evaluation of the clinical implementation of cognitive-behavioral therapy for chronic pain management in pediatric sickle cell disease. Complement Ther Med 2020 Mar;49:102348 [FREE Full text] [doi: 10.1016/j.ctim.2020.102348] [Medline: 32147059]
7. Thorn BE, Eyer JC, Van Dyke BP, Torres CA, Burns JW, Kim M, et al. Literacy-adapted cognitive behavioral therapy versus education for chronic pain at low-income clinics. Ann Intern Med 2018 Apr 03;168(7):471-480. [doi: 10.7326/m17-0972]
8. Williams ACDC, Fisher E, Hearn L, Eccleston C. Psychological therapies for the management of chronic pain (excluding headache) in adults. Cochrane Database Syst Rev 2020 Aug 12;8(8):CD007407 [FREE Full text] [doi: 10.1002/14651858.CD007407.pub4] [Medline: 32794606]
9. Tudor-Sfetea C, Rabee R, Najim M, Amin N, Chadha M, Jain M, et al. Evaluation of two mobile health apps in the context of smoking cessation: qualitative study of cognitive behavioral therapy (CBT) versus non-CBT-based digital solutions. JMIR Mhealth Uhealth 2018 Apr 18;6(4):e98 [FREE Full text] [doi: 10.2196/mhealth.9405] [Medline: 29669708]
10. Palermo TM, Lalloo C, Zhou C, Dampier C, Zempsky W, Badawy SM, et al. A cognitive-behavioral digital health intervention for sickle cell disease pain in adolescents: a randomized, controlled, multicenter trial. Pain 2024 Jan 01;165(1):164-176 [FREE Full text] [doi: 10.1097/j.pain.0000000000003009] [Medline: 37733479]
11. Palermo TM, Zempsky WT, Dampier CD, Lalloo C, Hundert AS, Murphy LK, et al. iCanCope with Sickle Cell Pain: design of a randomized controlled trial of a smartphone and web-based pain self-management program for youth with sickle cell disease. Contemp Clin Trials 2018 Nov;74:88-96 [FREE Full text] [doi: 10.1016/j.cct.2018.10.006] [Medline: 30316999]
12. Palermo T, Dudeney J, Santanelli JP, Carletti A, Zempsky WT. Feasibility and acceptability of internet-delivered cognitive behavioral therapy for chronic pain in adolescents with sickle cell disease and their parents. J Pediatr Hematol Oncol 2018 Mar;40(2):122-127 [FREE Full text] [doi: 10.1097/MPH.0000000000001018] [Medline: 29176462]
13. Nikolajski C, O'Brien J, Nardo E, Szigethy E, Jonassaint C. Tailoring a digital mental health program for patients with sickle cell disease: qualitative study. JMIR Ment Health 2023 Apr 06;10:e44216 [FREE Full text] [doi: 10.2196/44216] [Medline: 37023443]
14. Lalloo C, Nishat F, Zempsky W, Bakshi N, Badawy S, Ko YJ, et al. Characterizing user engagement with a digital intervention for pain self-management among youth with sickle cell disease and their caregivers: subanalysis of a randomized controlled trial. J Med Internet Res 2022 Aug 30;24(8):e40096 [FREE Full text] [doi: 10.2196/40096] [Medline: 36040789]
15. Peters-Lawrence MH, Bell MC, Hsu LL, Osunkwo I, Seaman P, Blackwood M, Sickle Cell Disease Clinical Research Network (SCDCRN). Clinical trial implementation and recruitment: lessons learned from the early closure of a randomized clinical trial. Contemp Clin Trials 2012 Mar;33(2):291-297 [FREE Full text] [doi: 10.1016/j.cct.2011.11.018] [Medline: 22155024]
16. Haywood C, Lanzkron S, Diener-West M, Haythornthwaite J, Strouse JJ, Bediako S, IMPORT Investigators. Attitudes toward clinical trials among patients with sickle cell disease. Clin Trials 2014 Jun 13;11(3):275-283 [FREE Full text] [doi: 10.1177/1740774513519876] [Medline: 24532686]
17. Lebensburger JD, Hilliard LM, Pair LE, Oster R, Howard TH, Cutter GR. Systematic review of interventional sickle cell trials registered in ClinicalTrials.gov. Clin Trials 2015 Dec 17;12(6):575-583 [FREE Full text] [doi: 10.1177/1740774515590811] [Medline: 26085544]
18. Parchuri E, Nardo E, O'Brien J, Crosby L, Porter J, Palermo T, et al. The effect of an adapted digital mental health intervention for sickle cell disease on engagement: a pilot randomized controlled trial. Res Sq 2023 Jun 28:2023 [FREE Full text] [doi: 10.21203/rs.3.rs-3073103/v1] [Medline: 37461733]
19. Jonassaint C, Lalama CM, Carroll CP, Badawy SM, Hamm ME, Stinson JN, et al. Digital cognitive behavioral therapy vs education for pain in adults with sickle cell disease. Blood Adv 2024 Dec 24;8(24):6257-6266 [FREE Full text] [doi: 10.1182/bloodadvances.2024013861] [Medline: 39374587]

XSL•FO
RenderX

20. Brodar KE, Carlisle V, Tang PY, Fisher EB. Identification and characterization of peer support for cancer prevention and care: a practice review. J Cancer Educ 2022 Jun 05;37(3):645-654 [FREE Full text] [doi: 10.1007/s13187-020-01861-8] [Medline: 32892278]
21. Thompson DM, Booth L, Moore D, Mathers J. Peer support for people with chronic conditions: a systematic review of reviews. BMC Health Serv Res 2022 Mar 31;22(1):427 [FREE Full text] [doi: 10.1186/s12913-022-07816-7] [Medline: 35361215]
22. Gaiser MG, Buche JL, Wayment CC, Schoebel V, Smith JE, Chapman SA, et al. A systematic review of the roles and contributions of peer providers in the behavioral health workforce. Am J Prev Med 2021 Oct;61(4):e203-e210. [doi: 10.1016/j.amepre.2021.03.025] [Medline: 34175172]
23. Puschner B, Repper J, Mahlke C, Nixdorf R, Basangwa D, Nakku J, et al. Using Peer Support in Developing Empowering Mental Health Services (UPSIDES): background, rationale and methodology. Ann Glob Health 2019 Apr 05;85(1):1 [FREE Full text] [doi: 10.5334/aogh.2435] [Medline: 30951270]
24. Sokol R, Fisher E. Peer support for the hardly reached: a systematic review. Am J Public Health 2016 Jul;106(7):e1-e8. [doi: 10.2105/ajph.2016.303180]
25. Gandy M, Pang ST, Scott AJ, Heriseanu AI, Bisby MA, Dudeney J, et al. Internet-delivered cognitive and behavioural based interventions for adults with chronic pain: a systematic review and meta-analysis of randomized controlled trials. Pain 2022 Oct 01;163(10):e1041-e1053. [doi: 10.1097/j.pain.0000000000002606] [Medline: 35121696]
26. Morcillo-Muñoz Y, Sánchez-Guarnido AJ, Calzón-Fernández S, Baena-Parejo I. Multimodal chronic pain therapy for adults via smartphone: randomized controlled clinical trial. J Med Internet Res 2022 May 11;24(5):e36114 [FREE Full text] [doi: 10.2196/36114] [Medline: 35373776]
27. Badawy SM, Abebe KZ, Reichman CA, Checo G, Hamm ME, Stinson J, et al. Comparing the effectiveness of education versus digital cognitive behavioral therapy for adults with sickle cell disease: protocol for the Cognitive Behavioral Therapy and Real-time Pain Management Intervention for Sickle Cell via Mobile Applications (CaRISMA) study. JMIR Res Protoc 2021 May 14;10(5):e29014 [FREE Full text] [doi: 10.2196/29014] [Medline: 33988517]
28. Kim H, Sefcik JS, Bradway C. Characteristics of qualitative descriptive studies: a systematic review. Res Nurs Health 2017 Feb 30;40(1):23-42 [FREE Full text] [doi: 10.1002/nur.21768] [Medline: 27686751]
29. Guest G, Bunce A, Johnson L. How many interviews are enough? Field Methods 2006 Feb 01;18(1):59-82. [doi: 10.1177/1525822x05279903]
30. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000 Aug;23(4):334-340. [doi: 10.1002/1098-240x(200008)23:4<334::aid-nur9>3.0.co;2-g] [Medline: 10940958]
31. Matsumoto K, Hamatani S, Shimizu E. Effectiveness of videoconference-delivered cognitive behavioral therapy for adults with psychiatric disorders: systematic and meta-analytic review. J Med Internet Res 2021 Dec 13;23(12):e31293 [FREE Full text] [doi: 10.2196/31293] [Medline: 34898445]
32. Fernandes LG, Devan H, Fioratti I, Kamper SJ, Williams CM, Saragiotto BT. At my own pace, space, and place: a systematic review of qualitative studies of enablers and barriers to telehealth interventions for people with chronic pain. Pain 2022 Feb 01;163(2):e165-e181. [doi: 10.1097/j.pain.0000000000002364] [Medline: 34433776]
33. Dennis C. Peer support within a health care context: a concept analysis. Int J Nurs Stud 2003 Mar;40(3):321-332. [doi: 10.1016/s0020-7489(02)00092-5] [Medline: 12605954]
34. Kim H, Tietsort C, Posteher K, Michaelides A, Toro-Ramos T. Enabling self-management of a chronic condition through patient-centered coaching: a case of an mHealth diabetes prevention program for older adults. Health Commun 2020 Dec 12;35(14):1791-1799. [doi: 10.1080/10410236.2019.1663583] [Medline: 31514532]
35. Bushey MA, Kroenke K, Weiner J, Porter B, Evans E, Baye F, et al. Telecare management of pain and mood symptoms: adherence, utility, and patient satisfaction. J Telemed Telecare 2019 Jun 20;26(10):619-626. [doi: 10.1177/1357633x19856156]
36. Aggarwal A, Tam CC, Wu D, Li X, Qiao S. Artificial intelligence-based chatbots for promoting health behavioral changes: systematic review. J Med Internet Res 2023 Feb 24;25:e40789 [FREE Full text] [doi: 10.2196/40789] [Medline: 36826990]
37. Li Y, Liang S, Zhu B, Liu X, Li J, Chen D, et al. Feasibility and effectiveness of artificial intelligence-driven conversational agents in healthcare interventions: a systematic review of randomized controlled trials. Int J Nurs Stud 2023 Jul;143:104494. [doi: 10.1016/j.ijnurstu.2023.104494] [Medline: 37146391]
38. Seftel AD. Re: a prospective diary study of stuttering priapism in adolescents and young men with sickle cell anemia: report of an international randomized control trial; the Priapism in Sickle Cell Study (PISCES study). Journal of Urology 2011 May;185(5):1837-1838. [doi: 10.1016/s0022-5347(11)60219-x]
39. Meyerowitz-Katz G, Ravi S, Arnolda L, Feng X, Maberly G, Astell-Burt T. Rates of attrition and dropout in app-based interventions for chronic disease: systematic review and meta-analysis. J Med Internet Res 2020 Sep 29;22(9):e20283 [FREE Full text] [doi: 10.2196/20283] [Medline: 32990635]
40. Goodday SM, Karlin E, Brooks A, Chapman C, Harry C, Lugo N, et al. Value of engagement in digital health technology research: evidence across 6 unique cohort studies. J Med Internet Res 2024 Sep 03;26:e57827 [FREE Full text] [doi: 10.2196/57827] [Medline: 39226552]

XSL•FO
RenderX

41. Reade S, Spencer K, Sergeant JC, Sperrin M, Schultz DM, Ainsworth J, et al. Cloudy with a chance of pain: engagement and subsequent attrition of daily data entry in a smartphone pilot study tracking weather, disease severity, and physical activity in patients with rheumatoid arthritis. JMIR Mhealth Uhealth 2017 Mar 24;5(3):e37 [FREE Full text] [doi: 10.2196/mhealth.6496] [Medline: 28341616]
42. Lai L, Liu Y, McCracken LM, Li Y, Ren Z. The efficacy of acceptance and commitment therapy for chronic pain: a three-level meta-analysis and a trial sequential analysis of randomized controlled trials. Behav Res Ther 2023 Jun;165:104308. [doi: 10.1016/j.brat.2023.104308] [Medline: 37043967]
43. Levin ME, Krafft J, Twohig MP. An overview of research on acceptance and commitment therapy. Psychiatr Clin North Am 2024 Jun;47(2):419-431. [doi: 10.1016/j.psc.2024.02.007] [Medline: 38724128]
44. Haller H, Breilmann P, Schröter M, Dobos G, Cramer H. A systematic review and meta-analysis of acceptance- and mindfulness-based interventions for DSM-5 anxiety disorders. Sci Rep 2021 Oct 14;11(1):20385 [FREE Full text] [doi: 10.1038/s41598-021-99882-w] [Medline: 34650179]
45. Burns JW, Jensen MP, Thorn B, Lillis TA, Carmody J, Newman AK, et al. Cognitive therapy, mindfulness-based stress reduction, and behavior therapy for the treatment of chronic pain: randomized controlled trial. Pain 2022 Feb 01;163(2):376-389. [doi: 10.1097/j.pain.0000000000002357] [Medline: 34074945]
46. Anto A, Basu A, Selim R, Foscht T, Eisingerich AB. Open-World Games' affordance of cognitive escapism, relaxation, and mental well-being among postgraduate students: mixed methods study. J Med Internet Res 2024 Dec 17;26:e63760 [FREE Full text] [doi: 10.2196/63760] [Medline: 39689301]

Abbreviations

CaRISMA: Cognitive Behavioral Therapy and Real-Time Pain Management Intervention for Sickle Cell via Mobile Applications
CBT: cognitive behavioral therapy
PHQ: Patient Health Questionnaire
PiSCES: Priapism in Sickle Cell Study
SCD: sickle cell disease

Edited by G Eysenbach, A Kushniruk; submitted 17.03.25; peer-reviewed by A Eisingerich, A Al-Asadi; comments to author 05.05.25; revised version received 29.06.25; accepted 14.07.25; published 05.08.25.

Please cite as:
Oluwole OB, McGill LS, Gayle K, O'Brien JA, Abebe KZ, Jonassaint CR, Hamm M
Experiences With a Multicomponent Digital Behavioral Pain Management Intervention for Adults With Sickle Cell Disease: Qualitative Analysis of the CaRISMA Trial
JMIR Hum Factors 2025;12:e73719
URL: https://humanfactors.jmir.org/2025/1/e73719
doi:10.2196/73719
PMID:

©Olubusola B Oluwole, Lakeya S McGill, Kadeem Gayle, Julia A O'Brien, Kaleab Z Abebe, Charles R Jonassaint, Megan Hamm. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 05.08.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Sociotechnical Approach to Bring-Your-Own-Device Security in Hospitals: Development and Pilot Testing of a Maturity Model Using Mixed Methods Action Research

Tafheem Ahmad Wani[1,2], PhD; Antonette Mendoza[2], PhD; Kathleen Gray[2], PhD

[1]School of Psychology and Public Health, Department of Public Health, La Trobe University, Melbourne, Australia

[2]Centre for Digital Transformation of Health, University of Melbourne, Melbourne, Australia

Corresponding Author:
Tafheem Ahmad Wani, PhD
School of Psychology and Public Health
Department of Public Health
La Trobe University
Level 5, Health Sciences 2, La Trobe University
Plenty Road, Bundoora, Victoria, Australia
Melbourne, 3086
Australia
Phone: 61 451906170
Email: t.wani@latrobe.edu.au

Abstract

Background: Bring your own device (BYOD) adoption in health care improves clinician productivity, but introduces cybersecurity risks due to weak security controls, human error, and policy circumvention. Existing security frameworks and models are technocentric, while overlooking sociotechnical factors such as clinician behavior, workflow integration, and organizational culture. This misalignment reduces their effectiveness in health care settings. In addition, hospitals vary in structure, resources, and BYOD use, necessitating a flexible yet structured approach to assess security maturity and prioritize improvements, which is lacking in existing models.

Objective: This study aims to develop and pilot a hospital BYOD security maturity model that integrates technical, policy, and human factors for a structured assessment and improvement of BYOD security in health care.

Methods: This study used mixed methods action research to design and pilot a hospital BYOD security maturity model. Surveys and interviews with IT managers and clinicians shaped the model, which was trialed at a public metropolitan hospital in Victoria, Australia. Participants completed a maturity assessment and joined a 90-minute co-design workshop that prioritized 6 key domains and proposed improvements. Descriptive statistics and thematic analysis guided refinements to improve clarity and usability.

Results: The model was initially developed with 22 domains across 3 key dimensions: technology, policy, and people, each structured across 5 maturity levels to support systematic progression in hospital BYOD security. On the basis of participant feedback during the refinement process, 2 training-related domains were merged, resulting in a final model with 21 domains. The technology dimension includes domains such as identity, access, and authentication management; device security; and clinical communication, ensuring technical controls align with hospital policies and workflows. The policy dimension focuses on governance, covering areas such as BYOD strategy, regulatory compliance, and incident response, to establish clear security guidelines and enforcement mechanisms. The people dimension addresses human factors, including security awareness training, stakeholder involvement, and security culture, fostering staff engagement and adherence to security protocols. A maturity assessment survey conducted at a public metropolitan hospital in Victoria, Australia, revealed an overall maturity level of 2.04. Key areas for improvement included identity and access management, clinical communication security, and governance transparency. A 90-minute co-design workshop identified challenges and proposed solutions for the top 6 priority domains. Recommendations included implementing single sign-on, defining a formal BYOD strategy, enhancing secure communication tools, and improving stakeholder engagement.

Conclusions: The model can serve as a valuable tool for hospitals and policy makers, offering actionable recommendations to strengthen BYOD security. The pilot implementation demonstrated its practical applicability, helping the hospital identify security

XSL•FO
RenderX

gaps and develop a road map for structured enhancements. Further validation across diverse health care settings will enhance its adaptability and long-term impact.

(JMIR Hum Factors 2025;12:e71912) doi:10.2196/71912

KEYWORDS

bring your own device; BYOD; security; privacy; mobile health; mHealth; mobile phone; maturity model; sociotechnical security; policy and governance; artificial intelligence; AI

Introduction

Background

Bring your own device (BYOD) is the practice of using personal devices, including laptops, smartphones, and tablets, for professional purposes. This practice has witnessed a significant increase in adoption across various sectors, especially within the health care domain [1]. A survey conducted in 2021 among hospital clinicians in Australia revealed that 87% of respondents engaged in BYOD for tasks related to their hospital responsibilities [2]. The ubiquitous and multifunctional nature of contemporary mobile devices attracts clinicians, who use BYOD for activities such as clinical documentation, accessing electronic medical records (EMRs), diagnostic decision support, and clinical communication [3-6]. For health care professionals, BYOD allows enhanced flexibility and convenience, thereby enhancing their mobility and productivity. Nevertheless, despite its benefits, BYOD is recognized as a contributor to increased cybersecurity vulnerabilities, which presents a significant concern for health care institutions [7,8]. A critical dimension of BYOD implementation is the insufficient security measures associated with clinicians' personal devices, rendering them more susceptible to security breaches, particularly due to human error and misuse. These elements are acknowledged as key contributors to data breaches within the health care sector [7,9,10]. In certain situations, clinicians may sometimes prioritize immediate patient care over adhering to IT security protocols, further exacerbating risks [11,12]. The risk escalates within a BYOD environment, wherein users exert greater control over hospital data residing on their personal devices, thus increasing the likelihood of security oversights. For instance, the proliferation of EMR use on mobile devices intensifies the difficulty of integrating security measures into routine workflows [13,14]. Moreover, adherence to health care privacy regulations further complicates the adoption of BYOD, as hospitals strive to reconcile security requirements with clinical operational efficiency [9].

Multiple academic studies have explored security issues associated with BYOD use, addressing challenges such as device theft, network vulnerabilities, malware, unencrypted data, and improper employee behavior [15-17]. However, research in the health care sector is comparatively limited. Several researchers emphasize the need for targeted academic investigations to offer a comprehensive and objective understanding of this field. They have highlighted that a one-size-fits-all approach to BYOD security is insufficient, and existing best practices and recommendations lack the nuanced approaches needed for specialized sectors such as health care, where security solutions must balance operational efficiency with strict regulatory requirements [18,19]. The intricate nature and multifaceted dynamics of technology implementation within health care calls for focused research on BYOD security in hospital environments. On the contrary, previous research has highlighted the limitations of existing BYOD security frameworks in health care, noting a predominant focus on technical solutions without adequately addressing the sociotechnical dimensions critical in clinical environments [15,16]. Existing frameworks emphasize technical controls such as access management and encryption (the process of converting data into a coded format to prevent unauthorized access, ensuring data security both in transit and at rest) but often overlook the importance of user behavior, organizational culture, and workflow integration [17,20]. This technocentric approach has been associated with increased vulnerability to cyberattacks due to neglected social aspects of security [21,22].

Recognizing these gaps, previous work involved developing a preliminary hospital BYOD security framework that integrated sociotechnical considerations, encompassing technical controls, policies, and human factors [10]. This framework outlined a 7-stage process—plan, identify, protect, detect, respond, recover, and assess and monitor—with recommendations tailored to the unique challenges of BYOD security management in hospitals. Subsequent empirical studies involving hospital IT managers and clinicians provided practical insights, further informing and refining the framework [2,23-25].

However, translating this comprehensive framework into actionable and practical sociotechnical security controls presents challenges. Hospitals vary significantly in their organizational structures, workflows, and available resources, making it impractical to implement all recommendations uniformly. There is a need for a flexible tool that allows hospitals to assess their current BYOD security posture and prioritize improvements in a manner that aligns with their specific contexts.

Maturity models have been used in information systems as instruments for assessment and systematic enhancement of organizational practices [26,27]. In information security, maturity models aid organizations in identifying cybersecurity gaps and developing tailored countermeasures that are feasible and aligned with organizational goals [28]. Nonetheless, existing maturity models in cybersecurity are predominantly technocentric and compliance-oriented, lacking in sociotechnical considerations and specific applicability to the health care sector [26,28-32]. A 2025 systematic review of cybersecurity maturity models identified several persistent gaps, including the absence of sector-specific frameworks and the limited integration of behavioral and organizational dimensions [33]. Similarly, other contemporary studies tend to focus on national cybersecurity capabilities, broad information security themes, or narrowly

XSL•FO
RenderX

defined applications, such as incident response, and hence do not address the full range of health care–specific challenges, nor do they provide comprehensive sociotechnical maturity frameworks [34-36]. The complexity of health care systems, characterized by intricate interactions between technology, clinical workflows, communication, and regulatory requirements, necessitates a maturity model that encompasses both technical and sociotechnical dimensions [37-39].

Objectives

Given the absence of a tailored maturity model addressing hospital BYOD security and its unique challenges, this study aims to bridge this gap by developing a comprehensive and holistic hospital BYOD security maturity model. Therefore, the aim of our research is to propose and present a comprehensive hospital BYOD security maturity model. Furthermore, we describe a pilot implementation that demonstrates the model's practical application to BYOD security in a specific hospital organizational context. In addition, the model presented aims to advance maturity model theory by offering a granular, domain-based structure that enables targeted progression mapping. This overcomes the abstraction of traditional models by operationalizing maturity in specific, interrelated hospital functions.

By bridging the gap between theoretical frameworks and practical implementation, this study aims to provide hospitals with a practical tool to enhance their BYOD security posture in a manner that is adaptable to their unique organizational contexts. The resulting maturity model is intended to facilitate a systematic approach to improving BYOD security, integrating both technical controls and sociotechnical factors, such as clinician behavior and organizational culture.

Methods

The method section comprises 2 subsections. First, we describe the method used to develop and refine the proposed hospital BYOD security maturity model. Then, we outline the method used to pilot its implementation.

Development and Application of the Proposed Hospital BYOD Security Maturity Model

Maturity Model Development

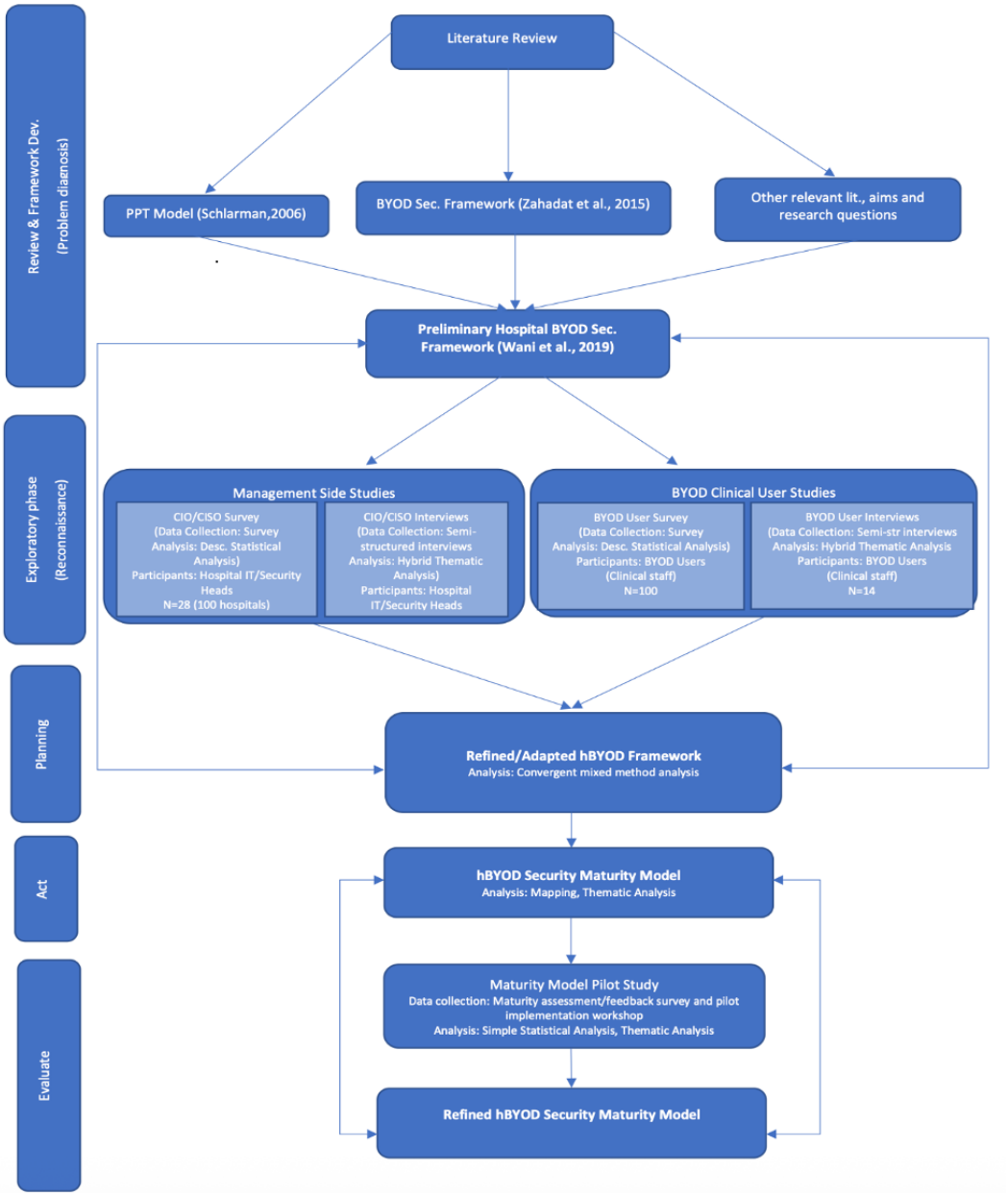
Before the maturity model, the authors developed a hospital BYOD security framework providing a comprehensive set of sociotechnical guidelines for BYOD security in hospitals across the BYOD life cycle. This framework was developed using a mixed methods action research (MMAR) approach, providing a robust foundation for the creation of the hospital BYOD (hBYOD) maturity model. The MMAR methodology enabled the integration of both qualitative and quantitative data (collected through empirical studies), ensuring that the framework addressed sociotechnical factors, including technical controls, policies, and human behaviors, to balance security with clinical productivity [40-42]. The quantitative studies involved surveys of IT management staff across hospitals to assess existing security practices and BYOD users, such as clinicians, to evaluate their security behaviors and compliance. The qualitative component included semistructured interviews with IT management staff to explore contextual factors influencing BYOD security decisions and with clinicians to understand how security measures impact their workflows and productivity. This comprehensive approach ensured that both the management and user perspectives were incorporated into the framework's development.

For the purposes of the maturity model, recommendations were mapped from the hospital BYOD security framework to the 5 maturity levels across all domains of the model. These descriptions and levels were triangulated among the 3 researchers to ensure consistency, accuracy, and alignment with the framework. By aligning the framework's recommendations with the maturity levels, elementary guidelines correspond to the lower levels, while more complex and advanced guidelines align with the higher levels, in accordance with the defined criteria for each level. This mapping ensures structured progression from basic to optimized practices within each domain.

Figure 1 illustrates the comprehensive process involved in developing the hBYOD maturity model, from initial problem analysis to the finalization of the model [10,19,43].

XSL•FO
RenderX

Figure 1. Hospital bring-your-own-device (hBYOD) security maturity model—Development process overview. BYOD: bring your own device; CIO: chief information officer; CISO: chief information security officer; PPT; people, policy, technology.



The proposed maturity model was initially developed with 22 domains across the 3 dimensions of people, process, and technology (PPT; 7 domains for technology, 7 for policy, and 8 for people), selected for their relevance to hospital BYOD security, following extensive research that included literature reviews and a series of empirical studies [2,23-25]. On the basis of participant feedback during the refinement process, 2 training-related domains were merged, resulting in a final model with 21 domains. Each domain represents a specific aspect of hospital BYOD security.

Each domain represents a specific aspect of hospital BYOD security. Incorporating the PPT dimensions ensured a holistic sociotechnical approach when assessing BYOD security maturity [43]. Every domain is defined across 5 maturity levels, ranging from level 1 to level 5, with unique definitions for each level in all 21 domains. These levels are aligned with prominent academic and industry-based maturity models [31,44-46]. An analysis of these models reveals the following definitions corresponding to each maturity level:

1. Nonexistent or very limited and elementary implementation, unreliable, completely reliant on manual effort, and major changes required.
2. Ad hoc or informal implementation, reactive, significant room for improvement, and mostly reliant on manual effort.
3. Well-defined but with gaps in management and enforcement, minimal automation, a high degree of manual effort, compliant, and moderate room for improvement.
4. Well-managed and enforced, structured practices, moderate level of automation, proactive, moderately agile, best practice, and low room for improvement.
5. Continually improving and highly agile and real-time, high levels of automation, complete integration, enterprise-wide practices, innovative, and up to date with little or no room for improvement.

XSL•FO
RenderX

The 5 maturity levels across all domains were derived from the hospital BYOD security framework (Multimedia Appendix 1), ensuring each level has a distinct and well-defined definition and characterization. Detailed mappings of the hospital BYOD security framework's recommendations to the maturity levels are provided in Multimedia Appendix 2.

Assessment Process Using the Maturity Model

The maturity model can be used to assess a hospital's BYOD security maturity through a straightforward process. For each of the 22 domains, a brief description of cybersecurity controls or practices is provided for each of the 5 maturity levels. Assessors rate their hospital from level 1 to level 5 in each domain, based on how closely their practices align with the provided descriptions. Assessors are expected to possess substantial knowledge of the described activities and have tangible evidence of the specified security controls.

When multiple assessors evaluate the same hospital, an average rating can be calculated for each domain to determine the overall maturity level. In addition, average ratings can be computed for each of the 3 PPT dimensions, as well as an overall average, providing cumulative ratings. This approach allows the model to identify specific areas for improvement and to represent the general capability of the hospital, based on individual domain maturity ratings and the cumulative mean maturity rating, respectively.

Prioritization and Roadmap for Improvement

It is important to note that areas for improvement can be prioritized or ranked based on the hospital's needs, future mission, capabilities, and feasibility considerations. Hospitals can plan a road map for enhancement by referring to the higher maturity levels in specific domains as guidance. For example, if a hospital's maturity rating in a particular domain is level 2, the descriptions of the progressive levels (levels 3, 4, and 5) can provide actionable insights on how stepwise improvements can be achieved. Depending on the hospital's capabilities and resources, it may aim for the level that is feasible to attain within a given time frame.

Pilot Implementation Study

This section outlines the methodology used for the pilot implementation study.

Ethical Considerations

Ethics approval was obtained from the hospital's ethics committee and the University of Melbourne, where the study was conducted (Ethics approval project no. 556/21). Informed consent was obtained from all participants through an information statement, and participation was voluntary, with an option to withdraw at any time. Survey responses were collected anonymously, and workshop data were pseudonymised and deidentified to ensure privacy and confidentiality. No patient data were collected. All data were securely stored on password-protected university systems. No compensation was offered to participants.

Study Site

The maturity assessment and improvement study were conducted at a major public tertiary metropolitan hospital in Victoria, Australia.

Study Participants and Recruitment

Overview

Two key groups of BYOD security stakeholders from the participating hospital were involved in this study. Participants were selected based on their involvement and understanding of BYOD security management, decision-making, or use within the hospital. The selection process was facilitated through nominations from the hospital's chief privacy officer to ensure the inclusion of individuals with relevant expertise and diverse perspectives. The criteria for participant selection included the following.

Technology Managers

IT professionals, cybersecurity managers, and policy makers responsible for BYOD decision-making, such as developing or implementing BYOD security strategies. They were chosen for their expertise in technical security controls, policy formulation, and compliance measures.

Clinical Stakeholders (Device Users)

Clinicians who understood BYOD use and clinical needs, such as those with IT and informatics roles. They were selected to provide insights into the impact of BYOD security controls on aspects such as clinical efficiency, usability, and compliance with hospital policies.

Nominated individuals were provided with a plain language statement and were included in the study upon giving informed consent. A total of 10 participants from the hospital took part in the study.

Study Design

The study was conducted in 2 stages.

Maturity Assessment and Model Feedback Survey (Before the Workshop)

A BYOD security maturity assessment and feedback survey was developed using Qualtrics (Qualtrics International Inc; Multimedia Appendix 3). Participants received a survey link and had 1 week to complete it independently before the workshop. Both technology managers and clinical stakeholders assessed BYOD security maturity by selecting the most appropriate level for each domain based on their knowledge and available tangible evidence. They also provided feedback on the accuracy, applicability, and clarity of the model, with an option to suggest improvements or indicate if they lacked sufficient knowledge to assess a domain.

Following the maturity assessment, participants were asked for their reflective comments on the structure of the model. This included (1) identifying any missing domains that should be included, (2) suggesting the removal of any domains deemed unnecessary, and (3) proposing reclassification of domains from one dimension to another.

XSL•FO
RenderX

At the end of the survey, participants were asked to prioritize and rank the domains within each PPT dimension based on their perceived importance for improving the BYOD security maturity of the hospital.

Maturity Improvement Discussion Workshop

Following the maturity assessment, a 90-minute Zoom (Zoom Communications, Inc) workshop was held with the same survey participants, including IT and clinical representatives providing management and user perspectives on BYOD use.

A participatory co-design approach was used during the workshop, enabling these stakeholder groups to collaboratively discuss challenges and propose recommendations related to the prioritized domains of BYOD security identified from the initial survey [47,48]. To maximize the effectiveness of the workshop, design and evaluation guidelines by Thoring et al [49] were followed.

The web-based workshop tool MURAL was used to document participant responses. Each group had 2 facilitators who managed time, addressed queries, and recorded key discussions. The full workshop session was recorded, transcribed, and deidentified.

Analysis

The maturity assessment data underwent descriptive statistical analysis to determine mean maturity ratings across the hospital's BYOD security domains. The top 6 priority domains (two from each of the technology, policy, and people dimensions) were identified based on participant responses.

For qualitative analysis, data from the maturity assessment, model improvement feedback, and workshop sessions were analyzed using thematic analysis [50]. A deductive approach was applied to assess how participant feedback aligned with predefined BYOD security maturity domains, followed by an inductive review to identify emerging challenges, contextual insights, and proposed solutions. This combined approach ensured that both structured evaluation and exploratory findings contributed to the refinement of the model and the formulation of consolidated, prioritized security recommendations for the hospital.

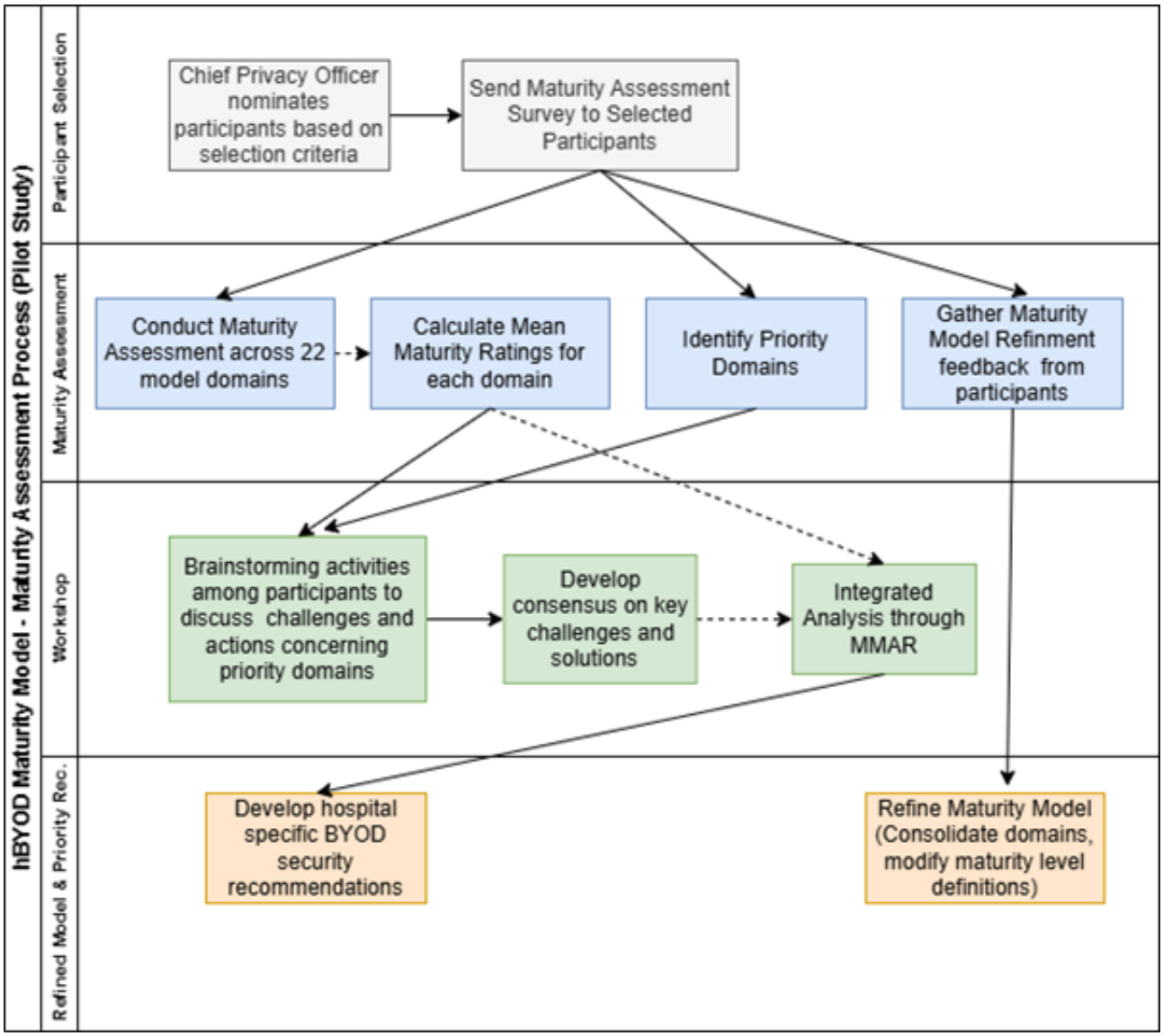
Participant Characteristics

A total of 10 participants from the model hospital, representing both participant groups—7 technology managers and 3 clinical stakeholder representatives—participated in the maturity assessment survey and the following discussion workshop.

Figure 2 provides an overview of the maturity assessment and model refinement process.

XSL•FO
RenderX

Figure 2. Maturity assessment and model refinement process overview. BYOD: bring your own device; hBYOD: hospital bring your own device; MMAR: mixed methods action research.



Results

This section is organized into three parts: (1) maturity model description, (2) feedback and refinements, and (3) hospital maturity assessment and workshop recommendations. The first part provides an overview of the maturity model domains, the second summarizes participant feedback and resulting revisions, and the third presents the findings from the maturity assessment survey and subsequent workshop conducted at the model hospital.

Maturity Model Description

Overview

The proposed maturity model was initially developed with 22 domains across 3 dimensions—technology (7 domains), policy (7 domains), and people (8 domains). Following participant feedback, 2 training-related domains under the people dimension were merged, resulting in a final model comprising 21 domains. This revision is detailed in the Feedback and Refinements subsection.

While this section provides only a brief overview of the 3 dimensions and their respective domains, the complete model, including detailed definitions for each of the 5 maturity levels per domain, along with clearly defined, measurable criteria, is presented in Multimedia Appendix 2. To aid interpretation, the operational indicators for each level have been bolded in the revised version of the appendix. These indicators are designed to support objective evaluation and guide structured improvement efforts.

Technology

Overview

The technology dimension demonstrates a clear progression from basic security practices to an advanced, integrated sociotechnical approach, where technology aligns with organizational policies and user needs. Each maturity level builds on the previous, addressing specific security and usability challenges relevant to BYOD use in hospitals. Table 1 presents a concise summary of the technology domains, offering a high-level overview of their structure and key focus areas.

XSL•FO
RenderX

Table 1. Technology dimension domain definition and descriptions (7 domains).

Technology dimension	Definition and description	Levels summary
Identity, access, and authentication management	Technologies and processes for verifying user identity and controlling access to hospital systems and data in a BYOD[a] environment	• Level 1: basic password-only authentication without additional protections for BYOD access. • Level 2: complex passwords with enforced policies, providing minimal improvement for personal device security. • Level 3: dual-factor authentication (eg, OTP or token) to enhance security for BYOD users. • Level 4: RBAC[b] tailored to BYOD device use, restricting access based on job roles. • Level 5: federated SSO[c] for seamless and secure access across multiple hospital systems from BYOD devices.
Storage and backup	Tools and controls for securely storing and backing up hospital data accessed or stored on BYOD devices, ensuring separation from personal data	• Level 1: no controls for securing or backing up data accessed via BYOD devices. • Level 2: minimal separation between personal and hospital data, with limited backup options for BYOD users. • Level 3: virtualizationd ensures BYOD users access data in secure, isolated environments without local storage. • Level 4: containerizatione fully separates hospital data on BYOD devices, with secure backup options. • Level 5: private cloud infrastructure with automated, encrypted backups for data accessed through BYOD devices.
Device security	Security measures to protect BYOD devices, including encryption[f], antivirus, and compliance monitoring	• Level 1: no restrictions or security requirements for BYOD devices. • Level 2: manual security controls, such as user-installed antivirus and basic settings. • Level 3: blacklist and white list capabilities to restrict unauthorized apps and ensure BYOD compliance. • Level 4: end point protectiong with real-time threat detection and device monitoring for BYOD devices. • Level 5: UEMh for centralized, automated security management of all BYOD devices.
Network security	Security protocols to protect hospital networks from threats introduced by BYOD use, including firewalls and segmentation	• Level 1: basic firewalls providing limited protection against BYOD-related threats. • Level 2: network segmentation separates hospital systems from BYOD devices to minimize risk. • Level 3: firewalls integrated with IDSi and IPSj) to identify and mitigate BYOD-originated threats. • Level 4: secure web gatewaysk filter internet traffic from BYOD devices, ensuring safe access. • Level 5: real-time threat detection using AIl and analytics to proactively secure hospital networks from BYOD threats.
Application security	Safeguards to secure applications in a BYOD environment, preventing unauthorized access and data breaches	• Level 1: no restrictions, allowing any applications on BYOD devices without security oversight. • Level 2: limited to basic applications with minimal security checks for BYOD use. • Level 3: vulnerable silos where hospital apps on BYOD gadgets lack integration or adequate oversight. • Level 4: approved white-listed applications with robust security policies for BYOD use. • Level 5: virtualized applications providing isolated and secure environments for BYOD users.
BYOD management and automation	Systems and tools to manage, monitor, and automate security measures for BYOD devices within the hospital	• Level 1: manual BYOD management with no centralized tools. • Level 2: ad hoc controls implemented reactively for BYOD issues. • Level 3: MDMm solutions enforce basic policies for BYOD compliance. • Level 4: MAMn tools control apps and permissions on BYOD gadgets. • Level 5: UEM provides comprehensive, automated security and operational management for BYOD gadgets.

XSL•FO
RenderX

Technology dimension	Definition and description	Levels summary
Clinical communication, photography, and file sharing	Secure platforms and protocols for communication, photography, file sharing, and data transfer using BYOD devices in clinical settings	• Level 1: undefined systems, with personal apps frequently used on BYOD devices for communication. • Level 2: generic communication platforms with limited security features for BYOD users. • Level 3: limited secure platforms partially integrated into hospital workflows for BYOD communication. • Level 4: fully integrated platforms with strong encryption and compliance features tailored for BYOD devices. • Level 5: dedicated systems specifically designed for secure clinical communication and file sharing on BYOD devices.

[a]BYOD: bring your own device.

[b]RBAC: role-based access control. A method of restricting system access based on an individual's job role to enforce security and limit exposure to sensitive data.

[c]SSO: single sign-on. An advanced form of SSO that allows users to access multiple apps and systems across different organizations or domains using a single authentication mechanism.

[d]A technology that allows hospital apps and data to be accessed in a controlled, virtual environment rather than being stored directly on BYOD devices, reducing security risks and enhancing centralized management.

[e]A security method that isolates hospital data and apps within a secure, controlled environment on a BYOD device, preventing unauthorized access to sensitive information while maintaining separation from personal apps.

[f]The process of converting data into a coded format to prevent unauthorized access, ensuring data security both in transit and at rest.

[g]End point protection: protection measures applied to end-user devices (eg, mobile phones, laptops) to prevent privacy or security breaches.

[h]UEM: unified end point management. A comprehensive security solution to managing and securing all devices (eg, desktops, mobiles, and tablets) from a single platform.

[i]IDS: intrusion detection system. Security tools are designed to detect, alert, and prevent potential cyberattacks by monitoring network traffic and system activities.

[j]IPS: intrusion prevention system.

[k]A security solution that monitors and filters internet traffic from devices to enforce hospital security policies, block malicious websites, prevent data leaks, and protect against web-based threats.

[l]AI: artificial intelligence.

[m]MDM: mobile device management. A security solution that allows organizations to control, monitor, and enforce policies on personal and organization-owned mobile devices.

[n]MAM: mobile application management. A security solution that focuses on controlling and securing hospital-approved apps on BYOD devices without managing the entire device, enabling secure access to clinical apps while maintaining user privacy.

The following sections provide a brief description of the 5 maturity levels across the 7 technology domains, illustrating the progression from minimal to advanced security capabilities and their application in BYOD contexts within hospitals.

Level 1: Minimal Security Controls

This foundational level exhibits basic security with limited controls. Authentication is password-based, without advanced access restrictions or separation between hospital and personal data on devices. Device security is minimal, allowing unrestricted personal device use, and network protection relies on basic firewall configurations. Application security is nonexistent, with employees using personal, unsecured applications. There are no centralized controls over BYOD use, and clinical communication occurs through personal apps, such as WhatsApp (Meta Platforms, Inc), posing significant data privacy and security risks.

Level 2: High Complexity and Minimal Automation

Security capabilities improve slightly, introducing single-factor authentication with complex passwords and limited device-level restrictions, mainly for high-risk devices. Network segmentation and guest network access are applied to separate personal and hospital data minimally, but data security largely depends on manual user management. BYOD management remains largely manual, and clinical communication continues through generic applications, offering little dedicated security.

Level 3: Basic Access Controls

At this level, more structured controls emerge, including basic identity and access management (IAM; a framework of policies and technologies ensuring that only authorized individuals can access specific hospital resources and systems) solutions with application-level virtualization (a technology that allows hospital applications and data to be accessed in a controlled, virtual environment rather than being stored directly on BYOD, reducing security risks and enhancing centralized management) for limited access control. Device security uses blacklist and white list capabilities to enforce some restrictions, while network security is strengthened with advanced firewalls and intrusion detection systems (security tools designed to detect, alert, and prevent potential cyberattacks by monitoring network traffic and system activities) and intrusion prevention systems. Although personal cloud options are permitted for backup, there is some degree of logical separation between personal and hospital data. Basic mobile device management (MDM; a security solution that allows organizations to control, monitor, and enforce policies on personal and organization-owned mobile

XSL•FO
RenderX

devices) tools are used for BYOD control and secure clinical communication platforms are available, although uptake remains minimal due to limited usability and integration.

Level 4: Role Based and Partially Integrated

This level introduces advanced IAM solutions with role-based access control, supporting more sophisticated, role-specific permissions. Data storage is managed through sanctioned cloud platforms with access restrictions based on user roles, enhancing data protection. End point protection (protection measures applied to end-user devices [eg, mobile phones and laptops] to prevent privacy or security breaches) incorporates network visualization and secure web gateways (a security solution that monitors and filters internet traffic from devices to enforce hospital security policies, block malicious websites, prevent data leaks, and protect against web-based threats), providing greater oversight of device security. Hospital-approved applications are made accessible through a secure app store, and mobile application management (MAM; a security solution that focuses on controlling and securing hospital-approved applications on BYOD devices without managing the entire device, enabling secure access to clinical apps while maintaining user privacy) tools enable more consistent BYOD management. Dedicated platforms for clinical communication offer basic integration with hospital systems, supporting more secure and streamlined workflows.

Level 5: Advanced, Integrated Security

The highest maturity level reflects an enterprise-level approach with fully integrated, automated security practices. IAM solutions use artificial intelligence for enhanced authentication management, including federated single sign-on (SSO; an advanced form of SSO that allows users to access multiple apps and systems across different organizations or domains using a single authentication mechanism) across all hospital applications, significantly improving user experience and security. Storage and backup solutions are handled via a private cloud with complete data isolation, ensuring data security. Comprehensive end point protection and real-time network monitoring provide robust defense against cybersecurity threats. Application security includes sophisticated isolation and restriction features, protecting sensitive data. Unified end point management (a comprehensive security solution to managing and securing all devices [eg, desktops, mobiles, and tablets] from a single platform) enables centralized, automated control over all BYOD devices, and dedicated clinical communication platforms integrate seamlessly with hospital systems, supporting efficient and secure communication workflows.

Policy

Overview

The policy dimension outlines a structured progression in BYOD policy development and governance within health care organizations. This dimension emphasizes the progression from a lack of structure in BYOD policy and security management to a comprehensive, proactive governance framework that aligns legal, technical, and operational elements of BYOD security within health care environments. Table 2 provides a brief overview of the policy domains.

XSL•FO
RenderX

Table 2. Policy dimension: domain definition and descriptions (7 domains).

Policy dimension	Definition and description	Levels summary
BYOD[a] strategy	The overarching plan for integrating BYOD into hospital workflows while ensuring security and compliance	• Level 1: no defined strategy for BYOD use. • Level 2: BYOD addressed informally as part of general IT strategy. • Level 3: a dedicated BYOD strategy focusing on security and compliance. • Level 4: comprehensive strategy covering clinical and nonclinical workflows related to BYOD use. • Level 5: proactively updated strategy aligned with evolving threats and hospital needs.
Policy components	Key elements of policies that define acceptable use, privacy, and security requirements for BYOD use	• Level 1: no specific BYOD policy components defined. • Level 2: basic acceptable use guidelines included in broader IT policies. • Level 3: dedicated BYOD policy components addressing security and privacy. • Level 4: comprehensive policies tailored for BYOD workflows and compliance requirements. • Level 5: regularly reviewed and updated policies integrating feedback and regulatory changes.
Regulatory compliance	Ensuring BYOD use adheres to health care regulatory standards and data security requirements	• Level 1: no understanding or monitoring of regulatory requirements for BYOD. • Level 2: minimal awareness of applicable regulations for BYOD use, with inconsistent adherence. • Level 3: regulatory requirements are clearly defined and communicated. • Level 4: compliance is routinely monitored and documented. • Level 5: automated systems ensure real-time adherence to regulatory standards for BYOD devices.
Policy enforcement	Mechanisms to monitor and ensure compliance with BYOD policies, including audits and corrective actions	• Level 1: no enforcement mechanisms for BYOD policies. • Level 2: limited BYOD monitoring with occasional audits. • Level 3: formal processes for BYOD monitoring and periodic reviews. • Level 4: balanced enforcement system integrating penalties and incentives. • Level 5: automated, real-time monitoring and enforcement of BYOD policies.
Incident response	Procedures for managing and resolving security incidents involving BYOD devices	• Level 1: no formal incident response processes for BYOD • Level 2: general IT incident management applies inconsistently to BYOD • Level 3: defined BYOD-specific incident response plans • Level 4: formalized processes with clear roles and rapid response capabilities • Level 5: automated, proactive incident response systems for BYOD threats
Lost device policy	Guidelines for handling lost or stolen BYOD devices, including data recovery and remote wiping	• Level 1: no procedures for lost or stolen BYOD devices • Level 2: staff follow ad hoc guidance for reporting and recovery • Level 3: defined policy for data wiping and device recovery • Level 4: selective data wiping to protect sensitive information • Level 5: real-time device tracking and automated data wiping capabilities
Accountability and governance	Defined structures and roles for overseeing and ensuring accountability for BYOD security	• Level 1: no defined roles or governance for BYOD security • Level 2: informal BYOD governance with limited oversight • Level 3: BYOD governance exists but is included in general IT management • Level 4: formal structures with clear roles and responsibilities for BYOD oversight • Level 5: agile BYOD governance processes integrated into organizational leadership and IT strategy

[a]BYOD: bring your own device.

The following subsections outline the progression of maturity levels within the policy dimension.

Level 1: Undefined BYOD Policy

At this foundational level, there is no formal BYOD strategy or dedicated policy for IT security concerning BYOD. Legal risks are not identified, and there are no established processes for enforcement, measurement, or compliance verification. Incident management is ad hoc and slow, with no defined roles or responsibilities for BYOD governance, making the organization highly vulnerable to security incidents.

Level 2: Basic, Informal Policy

BYOD policy is incorporated into the general IT strategy, with only loosely defined BYOD security aspects and references to legislation. Legal risks are vaguely defined, and compliance verification is minimal, with a general incident management process lacking specific standard operating procedures for BYOD incidents. Lost or stolen devices are not addressed, and governance is informal, lacking a clear structure and authority.

XSL•FO
RenderX

Level 3: Defined Policy With Basic Governance

A dedicated BYOD policy is established, offering high-level guidance. Legal risks are identified, but they remain complex and difficult for the staff to understand. There is basic enforcement, with a high-level compliance verification process. A BYOD incident management process is defined, but response times are slow, and devices are only wiped out when it is lost or stolen. Governance is somewhat integrated, with BYOD security part of the general IT and security governance framework.

Level 4: Comprehensive Policy With Formal Governance

At this level, a comprehensive BYOD strategy covers both clinical and nonclinical aspects, with well-defined legal risks that are regularly monitored. Compliance is measured through a balanced system, and a defined BYOD incident management process is in place, supported by a dedicated security team. The policy allows for selective data wiping on lost devices, and a formal BYOD governance structure is established to ensure accountability and clarity in roles.

Level 5: Advanced, Proactive Policy

The highest maturity level features a fully developed and continually updated BYOD strategy that is aligned with the latest threat, business, and legal landscapes. Compliance verification is automated with real-time measurement. The incident management process is highly automated and proactive, with swift responses and advanced capabilities. Governance is robust, ensuring that all aspects of BYOD policy are tightly managed within a well-defined structure, allowing for adaptability and scalability in response to evolving security needs.

People

Overview

The people dimension emphasizes the evolution from minimal awareness and isolated efforts to a fully integrated, proactive security culture. As maturity increases, the organization's workforce becomes more knowledgeable, engaged, and adept at managing BYOD security, aligned with the organization's overarching goals for patient safety and data protection in health care environments. Table 3 provides a brief overview of the people domains:

The following subsections outline the progression of maturity levels within the people dimension.

XSL•FO
RenderX

Table 3. People dimension: domain definitions and descriptions (8 domains).

People dimension	Definition and description	Levels summary
BYOD[a] security awareness and training	Programs to educate staff on the risks and security best practices specific to BYOD device use	• Level 1: no awareness programs or training for staff • Level 2: generic training offered occasionally, with minimal focus on BYOD security • Level 3: dedicated BYOD training programs addressing basic security risks • Level 4: comprehensive, role-specific training with practical scenarios for BYOD security • Level 5: regularly updated, workflow-integrated training emphasizing BYOD best practices
Training dissemination	Methods and channels for delivering BYOD-specific security training to staff effectively	• Level 1: no structured BYOD training dissemination • Level 2: periodic reminders or basic presentations • Level 3: online modules accessible to staff, covering BYOD security • Level 4: multichannel training, including workshops and hands-on sessions tailored to staff roles • Level 5: integrated dissemination through various media, including just-in-time training tools
Training importance	Emphasizing the role of training in maintaining security when using BYOD devices in the hospital environment	• Level 1: BYOD training not prioritized or mandatory • Level 2: training is optional and provided only to interested staff • Level 3: training viewed as important and made available to all staff • Level 4: training made mandatory with management oversight • Level 5: training incentivized and fully embedded in hospital practices, with performance linked to adherence
Management support	Leadership commitment and staff support to prioritizing and funding for BYOD security	• Level 1: no visible support from leadership for BYOD security initiatives • Level 2: limited acknowledgment of BYOD security risks • Level 3: moderate support, with funding allocated for basic BYOD security measures • Level 4: high-level support with strategic investments in BYOD security • Level 5: unrestricted commitment, with BYOD security as a priority in organizational strategy
Stakeholder involvement	Engagement of key stakeholders, including clinical staff and IT teams, in shaping BYOD security strategy, practices, and procedures	• Level 1: no involvement of stakeholders in BYOD decisions • Level 2: minimal IT department involvement without clinical input • Level 3: limited leadership and stakeholder input in shaping BYOD policies • Level 4: regular feedback and participation from clinical and IT teams • Level 5: continuous collaborative engagement of stakeholders in all BYOD-related strategies and practices
Security culture	Promoting a proactive organizational approach to securing hospital and personal data in a BYOD environment, fostering awareness and collaboration	• Level 1: no emphasis on security culture; reactive approach to incidents • Level 2: awareness limited to certain departments • Level 3: general staff awareness of security risks but minimal collaboration • Level 4: proactive initiatives to promote a security-conscious culture across the organization • Level 5: continuous improvement and organization-wide security commitment embedded into the culture
Usability and productivity	Balancing security measures for BYOD devices with ease of use to enhance clinical workflows and staff efficiency	• Level 1: BYOD security measures such as extensive restrictions and complex passwords, impede usability, leading to workarounds • Level 2: basic usability improvements, such as reduced sign-on, implemented without significant productivity gains • Level 3: improved usability through simplified processes such as basic sign-on • Level 4: workflow-optimized BYOD security measures, such as single sign-on and enterprise productivity BYOD apps • Level 5: fully integrated, seamless user experience enhancing both security and productivity through an optimized, secured, and unrestrictive BYOD environment
Expertise and skills improvement	Building and maintaining technical skills to manage and secure BYOD devices effectively	• Level 1: no focus on BYOD-related technical skills • Level 2: limited skill-building through ad hoc initiatives • Level 3: dedicated IT staff receive basic BYOD security training • Level 4: regular training and certifications provided for key staff • Level 5: advanced, ongoing skill development programs to maintain high levels of BYOD expertise

[a]BYOD: bring your own device.

Level 1: Limited Awareness and Reactive Approach

At this initial stage, there is no formal IT security training for staff regarding BYOD. Employees rely solely on personal knowledge and awareness, without any structured support or guidance from management on cybersecurity issues. There is a lack of stakeholder involvement in the BYOD strategy, leading

XSL•FO
RenderX

to a restrictive environment with low expertise and inadequate capacity to handle BYOD-related security issues.

Level 2: Basic Guidance With Minimal Organizational Support

This level introduces optional, generic IT security training, with basic guidelines and reminders sent to staff. Management provides limited support for cybersecurity initiatives, with only IT or executive groups involved in BYOD strategy development. The organization operates within a reactive security culture, where employees create workarounds due to limited resources. Individual staff members handle BYOD security issues independently, rather than through coordinated departmental efforts.

Level 3: Structured Training With Moderate Stakeholder Engagement

A dedicated BYOD security training program is in place, prioritizing cybersecurity with moderate management support. Clinical leaders play a minimal role in BYOD security decision-making, and there is general awareness among staff about the importance of patient data protection. However, resistance to change exists within some groups. Basic features, such as SSO, are introduced to simplify workflows, and dedicated IT personnel manage BYOD security.

Level 4: Integrated Security Culture With High Stakeholder Involvement

BYOD security training becomes a mandatory, strategic, and practical component of the organization's cybersecurity approach, with a high level of commitment and active management support. Clinical stakeholders, including BYOD users, are consulted in strategy development and changes, fostering a proactive security culture. The organization demonstrates strong capabilities to manage BYOD-related sociotechnical issues and leverages enterprise productivity features to enhance security and efficiency.

Level 5: Tailored Support and Proactive, Adaptive Security Culture

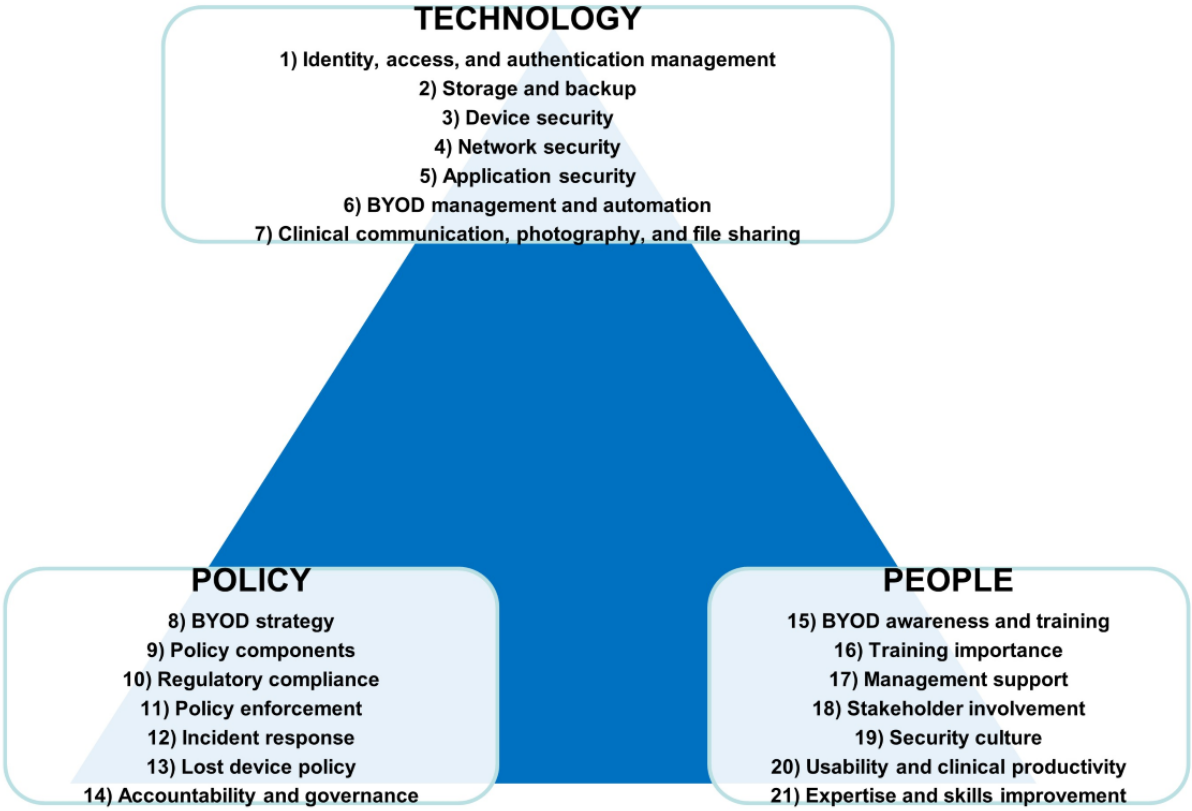
At the highest maturity level, BYOD security training is regularly updated and tailored to align with specific workflows, making cybersecurity a top priority. There is ongoing, unrestricted support from management, along with a collaborative process involving all stakeholders in BYOD strategy development. The organization promotes a proactive and continually improving security culture, where security concerns are autonomously managed through skilled improvement programs. This level represents a mature, integrated approach where the workforce is fully engaged and prepared to handle BYOD security challenges effectively.

Feedback and Refinements

Participant feedback led to both content and structural refinements of the maturity model. Content changes included clarifying technical terms, enhancing policy definitions, and tailoring guidance to BYOD use scenarios. Structurally, overlapping domains were consolidated for clarity and coherence. Detailed revisions, including before-and-after descriptions, are provided in Multimedia Appendix 4. Figure 3 illustrates the revised 21 domains across the 3 dimensions of the hBYOD security maturity model.

XSL•FO
RenderX

Figure 3. Hospital bring-your-own-device (hBYOD) security maturity model—dimensions and domains. BYOD: bring your own device.



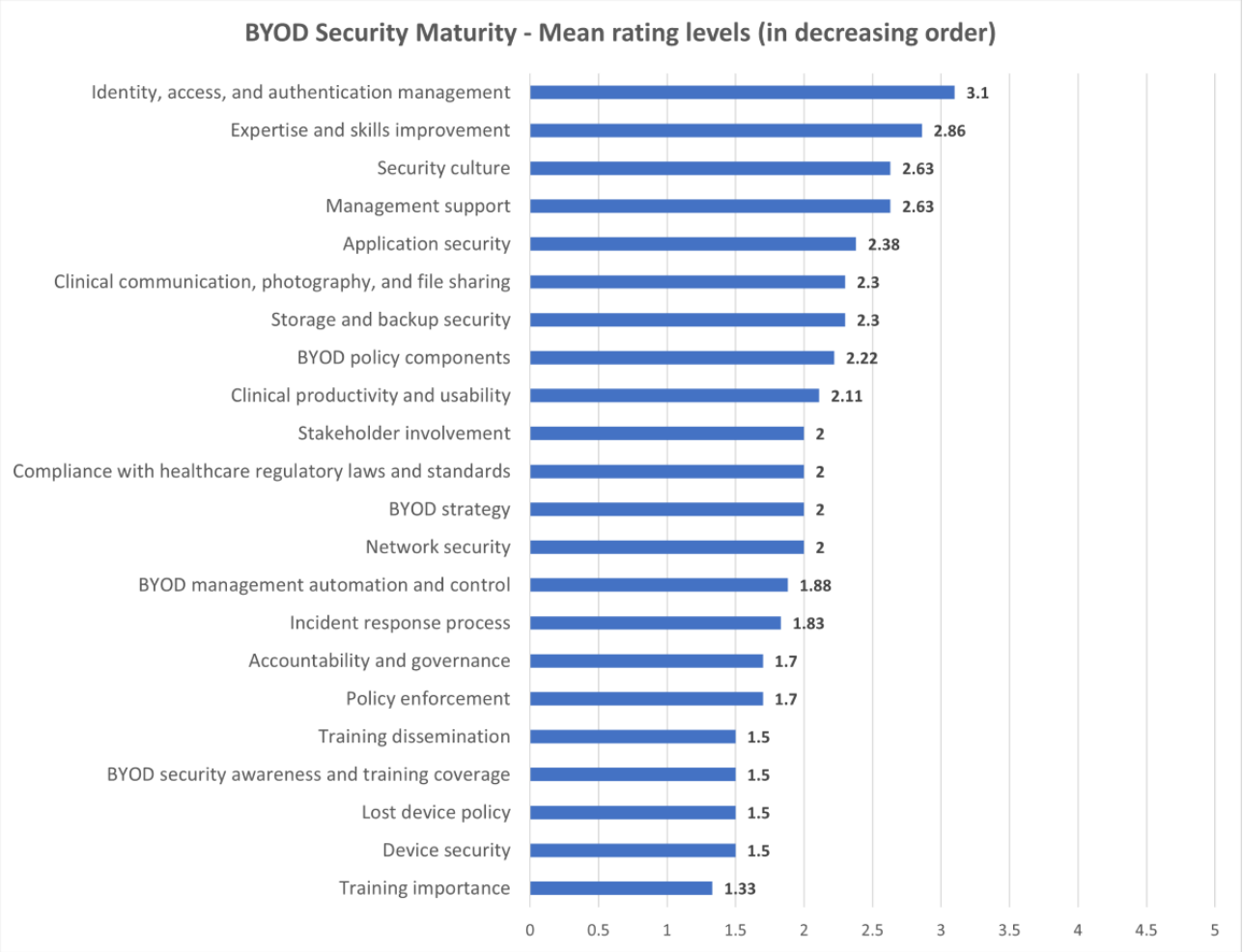
Pilot Study: BYOD Security Maturity Assessment and Recommendations for Model Hospital

Maturity Assessment Survey Results

The hospital BYOD maturity assessment evaluated security practices for the model hospital across the initially envisaged 22 domains, categorized into technology, policy, and people dimensions. The overall mean maturity level was 2.04, indicating foundational security practices. Technology scored an average of 2.21, policy scored an average of 1.85, and people scored an average of 2.07. Figure 4 provides a summary of the assessment results.

XSL•FO
RenderX

Figure 4. Bring-your-own-device (BYOD) security maturity assessment results for the model hospital.



Workshop Findings and Recommendations

Overview

In the maturity assessment survey, participants ranked domains to identify priority areas. Given time constraints, only the top 2 ranked domains from each PPT dimension were discussed in the workshop. The following summarizes key challenges and solutions mapped to the maturity model recommendations:

Priority Domain 1 (Technology): Identity, Access, and Authentication Management

Challenge 1: Cumbersome Authentication Requirements

Clinicians faced productivity challenges due to complex authentication processes, such as logging into multiple hospital applications, setting up 2-factor authentication, and managing regular BYOD updates. The varying levels of technology expertise among clinicians further exacerbated the issue, increasing dependency on IT support staff.

Recommendations

The first recommendation was BYOD workflow analysis. Tailor authentication processes to align with clinical workflows, minimizing disruption (usability and clinical productivity: level 4). Define access levels based on roles and analyze data flows to ensure security without impacting productivity (IAM: level 4). The second recommendation was advanced IAM solutions. Invest in role-based access control, multifactor authentication, and SSO services to streamline IAM across device types and locations, enhancing efficiency and reducing IT overhead (IAM: level 4).

Challenge 2: Managing BYOD Device Identities Across Multiple Health Services

Clinicians working across multiple health services encounter conflicts in BYOD requirements, making it impractical to implement device management solutions such as MDM.

Recommendations

The first recommendation was *hospital-owned devices or device segregation*. Encourage using separate BYOD devices for different health services to simplify management. Alternatively, provide hospital-owned devices to staff working at multiple health services after following an approval process. Allow full administrative control over devices for IT administrators (device security: level 3). The second recommendation was centralized identity management**.** Implement state-level centralized identity systems to provide a unified authentication experience across multiple health services, ensuring compatibility and reducing complexity (IAM: level 4; BYOD management automation and control: level 4).

XSL•FO
RenderX

Priority Domain 2 (Technology): Clinical Communication, Photography, and File Sharing

Challenge 1: Lack of a Sanctioned Platform for Clinical Communication

The absence of a formal strategy led to the use of multiple platforms for clinical communication, photography, and file sharing, with no consistency. Hospitals relied on tools such as Microsoft Teams, SharePoint, and OneDrive, while clinicians often used personal apps such as WhatsApp and Messenger for communication.

Recommendations

The first recommendation was strategy and policy development. Define and formalize a comprehensive strategy outlining approved platforms, processes for obtaining patient consent, and guidelines on information sharing and separation of personal and hospital data on BYOD devices (clinical communication, photography, and file sharing: level 3). The second recommendation was stakeholder engagement. Ensure buy-in from clinicians and other stakeholders to promote adherence to the strategy, minimizing policy circumvention (stakeholder engagement: level 4).

Challenge 2: Use of Unsecure Platforms

Unsecure platforms for communication and file-sharing posed risks, such as data leakage and unauthorized access, especially when patient data or photographs were stored on personal devices or cloud storage.

Recommendations

The first recommendation was to formalize generic platforms. Restrict the use of unsecure personal apps and formalize hospital-approved platforms such as Teams, SharePoint, and OneDrive, leveraging their encryption and secure storage capabilities to reduce risks (clinical communication, photography, and file sharing: level 2). The second recommendation was to invest in dedicated platforms. Opt for secure clinical communication and file-sharing platforms tailored for hospital use. These platforms should integrate with clinical systems, support electronic patient consent for photography and data sharing, and improve usability, thereby enhancing clinical workflows (clinical communication, photography, and file sharing: level 4).

Priority Domain 3: BYOD Strategy

Challenge 1: Lack of Strategic Direction

Workshop participants highlighted the absence of a clear strategic framework for BYOD security. Sociotechnical controls were implemented on an ad hoc basis, with critical aspects such as approved applications, device definitions, information exchange, and policy enforcement measures left undefined. Inconsistencies in related policies further compounded the issue, creating confusion among staff.

Recommendations

The first recommendation was to develop a comprehensive strategy. Establish a formal BYOD strategy addressing both clinical and nonclinical aspects, with a strong focus on patient and staff data privacy and security (BYOD strategy: level 3). The second recommendation was to consolidate existing controls. Streamline and formalize existing sociotechnical controls to eliminate inconsistencies and conflicts in related policies (BYOD strategy: level 3). The third recommendation was effective communication. Disseminate the strategy through various channels to ensure staff understand and comply with BYOD security guidelines, fostering clarity, direction, and user buy-in (stakeholder involvement: level 3).

Priority Domain 4: Accountability and Governance

Challenge 1: Undefined Roles and Responsibilities

Workshop participants highlighted a lack of clarity regarding BYOD-related roles and responsibilities, leading to confusion among staff. Clinicians were unsure about boundaries for BYOD use, such as when and where devices could be used (eg, within the hospital or after work hours) and for what purposes. In addition, there was insufficient guidance on actions required to ensure patient data privacy, responsible BYOD use, and compliance with legal and hospital policies.

Recommendations

The first recommendation was to define and communicate roles: Clearly outline BYOD-related roles and responsibilities through user agreements or employment contracts (accountability and governance: level 3). The second recommendation was a BYOD governance framework. Implement a governance structure that specifies reporting lines and promotes accountability (accountability and governance: level 4). The third recommendation was policy alignment. Ensure these expectations are integrated into the hospital's BYOD strategy to enhance compliance and improve user behavior (BYOD strategy: level 3).

Challenge 2: Lack of Transparency

Staff expressed concerns about the rationale behind certain security measures, such as MDM, which they perceived as intrusive to personal privacy. This lack of transparency created mistrust, with staff uncertain about the implications of security requirements such as dual-factor authentication and device restrictions.

Recommendations

The first recommendation was to enhance transparency. Clearly explain the rationale behind BYOD security measures, addressing staff concerns about personal privacy and productivity (security culture: level 3). The second recommendation was effective communication. Use the BYOD policy, user agreements, and training sessions to inform staff about privacy protections and hospital efforts to maintain productivity (BYOD awareness and training: level 3; stakeholder involvement: level 3). The third recommendation was to build trust. Foster confidence by demonstrating how personal data are safeguarded and ensuring staff understand the benefits of implemented security measures (accountability and governance: level 4; security culture: level 4).

XSL•FO
RenderX

Priority Domain 5: Stakeholder Involvement

Challenge 1: Minimal Input From Clinical Stakeholders

Both clinical and IT management groups identified the lack of clinician input in BYOD policy and strategy development as a key issue. This gap often led to policies misaligned with clinicians' workflow and productivity needs, resulting in limited buy-in. The compliance-driven focus of IT management, aimed at meeting state regulatory requirements, was seen as a contributing factor. Clinicians also raised concerns about meeting device requirements and the additional costs of using personal mobile data. On the IT side, engaging busy clinicians in strategy discussions was challenging, as cybersecurity often took a lower priority compared to their clinical responsibilities.

Recommendations

The first recommendation was to involve senior clinical stakeholders. Engage senior informatics officers, such as chief medical informatics officers or chief nursing informatics officers, in BYOD policy development and decision-making. These leaders can act as bridges between clinicians and IT management (stakeholder involvement: level 3). The second recommendation was to facilitate regular engagement. Establish regular forums, such as monthly meetings, to discuss BYOD security issues, raise awareness, and gather clinician feedback (stakeholder involvement: level 4). The third recommendation was to promote cybersecurity awareness. Use informatics officers to educate clinicians on the importance of cybersecurity, integrating it into their roles to encourage active participation in related activities (stakeholder involvement: level 5).

Priority Domain 6: Security Culture

Challenge 1: Convenience-First Environment

Participants noted that in a hospital setting, patient care often takes precedence, leading to a preference for convenience over data security. Clinicians frequently circumvent policies by using unsanctioned platforms, such as WhatsApp, for tasks such as sharing patient photos or clinical information.

Recommendations

The first recommendation was user-friendly policies and tools. Minimize workflow disruption by ensuring BYOD tools and processes are intuitive and easy to use (usability and clinical productivity: level 4; security culture: level 3). The second recommendation was comprehensive training programs. Enhance staff awareness of BYOD security risks through interactive and tailored training methods, such as phishing simulations and role-specific modules (BYOD security awareness and training: level 4). The third recommendation was strategic awareness campaigns. Educate clinicians on the impact of policy circumvention to foster compliance and reduce reliance on insecure platforms (security culture: level 4).

Challenge 2: Resistance to Change

Clinicians may resist BYOD security measures, citing fears of personal privacy intrusion or productivity losses. For example, concerns about device monitoring or objections to multifactor authentication and mandatory updates are common.

Recommendations

The first recommendation was to change management programs. Develop systematic approaches to address resistance, incorporating strategies to reduce disruption and alleviate concerns (security culture: level 4). The second recommendation was the hospital's commitment to cybersecurity. Ensure visible management support to emphasize the importance of cybersecurity (management support: level 4). The third recommendation was to engage change champions. Use influential clinicians to advocate for security measures, explain their rationale, and address concerns about privacy and workflow impacts within their departments (security culture: level 5).

The workshop structure and screenshots of brainstorming notes from the MURAL worksheets used by the clinical and IT management groups are provided in Multimedia Appendix 5.

Discussion

Principal Findings

This study introduces a sociotechnical maturity model that offers a holistic approach to addressing hospital BYOD security risks. By combining technical controls with organizational policies and human factors, the model fills existing gaps in the literature by integrating technical controls with organizational policies and human factors. It uniquely defines maturity levels across 21 specific domains, making it modular and providing detailed guidance throughout all areas of BYOD security. In addition, encompassing the key dimensions of PPT, it offers a comprehensive framework that spans the entire BYOD life cycle within hospitals. This aligns with the recommendations of several researchers who highlighted the importance of integrating sociotechnical elements into information security management [22,51-53]. Furthermore, a prominent feature of the maturity model is its accommodation of unique health care sociotechnical factors. Domains, such as clinical communication and collaboration, clinical productivity and usability, and compliance with health care regulatory laws and standards, better represent the complexity of cybersecurity in health care settings, especially concerning BYOD environments. This focus is supported by McLeod and Dolezel [54], who emphasized the critical role of regulatory compliance and user behavior in health care cybersecurity. The model also contributes theoretically by operationalizing abstract socio-organizational factors, such as security culture, into discrete, assessable maturity levels. This allows hospitals to systematically measure and improve human-centered dimensions of security previously overlooked in dominant frameworks.

This model also directly responds to several limitations identified in cybersecurity maturity literature. For instance, Büyüközkan and Güler [33] noted the lack of industry-specific models and insufficient emphasis on human and organizational factors. While their model is broad and literature based, it does not include detailed, structured maturity levels or any health care–specific focus. Bitzer et al [35] present a sociotechnical maturity model with tangible evaluation metrics, but it is limited to incident response and does not extend across the broader health care cybersecurity landscape. Similarly, Smyrlis et al [36] offer a health care cybersecurity risk assessment model

XSL•FO
RenderX

that is entirely technical and lacks both maturity level structuring and attention to human or organizational dimensions. AlDaajeh and Alrabaee [34] center their work on national-level cybersecurity capabilities rather than institutional or clinical implementation. In contrast, our model provides a modular, multidimensional, and health care–specific approach that can be practically applied within hospitals, offering clearly defined maturity levels across technology, policy, and people dimensions to guide structured, real-world improvement in BYOD security. In addition, the structured and modular design of the model facilitates alignment with internationally recognized cybersecurity frameworks, such as ISO/IEC 27001 and the National Institute of Standards and Technology Cybersecurity Framework [55]. By covering key areas, including governance, access control, incident response, and compliance monitoring, the model enables translation of these high-level standards into context-specific, actionable practices. It supports hospitals in operationalizing global best practices through enforceable policies, measurable progress indicators, and a commitment to continuous improvement, all while addressing the distinctive sociotechnical challenges posed by BYOD in health care environments.

The model was developed through MMAR, which synergistically integrated perspectives from both IT management and clinical users. MMAR facilitated the consolidation and critical analysis of findings from multiple empirical studies, enabling the refinement of the preliminary framework and the mapping of recommendations to formulate a comprehensive maturity model that ultimately provides a structured road map for enhancing BYOD security capabilities within hospitals. Being incremental and iterative in nature, MMAR can allow successive refinements to the maturity model based on evolving insights by incorporating emerging cybersecurity threats, advancements in technology, and evolving hospital security needs. As demonstrated in this study, stakeholder engagement through IT and clinical representatives played a crucial role in refining the maturity model. Therefore, methodologically, this study contributes to cybersecurity maturity model development by applying MMAR, enabling co-creation with end users. This iterative, participatory process grounds the model in real-world hospital contexts and contributes to theory by demonstrating the value of embedded stakeholder engagement in model formulation.

In terms of model updating, a cyclical evaluation process can be applied for future updates to the model through a structured step-by-step approach. First, periodic reassessments should be conducted through a regular review of industry best practices, emerging security frameworks, and hospital-wide BYOD security evaluations, including surveys and compliance audits, to identify gaps and risks. Second, real-world case studies can be integrated by analyzing security incidents, policy effectiveness, and user compliance trends, ensuring the model reflects actual BYOD security challenges in hospitals. Third, expert consultations with cybersecurity professionals, health care IT leaders, and policy makers can help incorporate regulatory updates, advancements in security technologies, and evolving best practices. Fourth, refining domain definitions and maturity levels based on the best practices review, stakeholder feedback, and empirical findings will ensure that the model accurately represents new threats, technologies, and governance strategies. Finally, validation and implementation should occur through structured feedback loops, pilot testing across diverse health care settings, and continuous monitoring, enabling iterative refinements that enhance the model's adaptability and long-term applicability in hospital environments. This stepwise approach, enabled by the incremental nature of MMAR, ensures that the BYOD security maturity model remains dynamic, evidence based, and responsive to the evolving cybersecurity landscape in health care.

The technology dimension covers various BYOD security technical controls and practices across domains such as IAM, storage and backup security, device security, network security, application security, and clinical communication, photography, and file sharing. The lower levels reflect basic practices, such as single-factor authentication systems, mixing of hospital and personal data on BYOD devices, use of basic firewalls, no restriction or control over BYOD use, and use of personal and unsecure apps for clinical communication and collaboration. Such practices may lead to a higher risk of BYOD-related data breaches, as they suggest a complete dependence of BYOD security on user action due to a lack of control over BYOD-related user activities. As we move to higher levels, advanced technical controls, such as enterprise IAM solutions, federated and unified SSO, private cloud for hospital data storage and backup, unified end point protection to manage BYOD devices, network segmentation and visualization, software-defined networking, and a dedicated secure platform for clinical communication integrated with hospital clinical systems, are stipulated. These features minimize the risk of hospital data leakage, restrict unsecure behavior, improve clinical workflow, and also enable better management and control over hospital BYOD use. Such practices align with the topmost or "optimized" level of prominent information security maturity models [28].

The policy dimension focuses on hospital BYOD security strategy, governance, and compliance factors. Lower levels specify policy practices such as a lack of a formal BYOD governance structure, ad hoc and inconsistent policies on BYOD security, a slow incident management process, no or unchecked policy enforcement, and poor compliance to relevant regulations or standards. These practices indicate a lack of direction and guidance on BYOD security, which can result in poor management of BYOD security incidents and greater chances of breach of relevant legislation or regulations. Higher or more advanced levels are characterized by a well-defined, formalized, and comprehensive BYOD strategy and related policies, which are regularly updated in view of changes to the threat, business, and legal landscape. Real-time, automated, and advanced compliance measurement and incident management processes are also key features of higher BYOD policy maturity levels. Therefore, hospitals at higher BYOD security maturity levels are better guided and equipped to prevent, detect, and respond to BYOD security incidents. Such a proactive approach, incorporating technology-enabled solutions for real-time compliance monitoring, aligns with prominent adaptive security management frameworks emphasizing the importance of

XSL•FO
RenderX

continuously enhancing security practices by leveraging real-time data to promptly detect security incidents, monitor compliance, and adapt policies as needed [56,57].

Finally, the people dimension addresses the "human" element of hospital BYOD security and therefore consists of domains such as BYOD security awareness and training, stakeholder involvement, security culture, and clinical productivity and usability. The lower maturity levels of the "people" dimension indicate a reactive security culture where there is no or low commitment from hospital IT management and clinicians alike to improve BYOD security practices. It is characterized by minimal stakeholder input on BYOD security management, a restrictive BYOD environment, a lack of BYOD security training, and minimal skills or low expertise to manage BYOD security issues. Such measures may lead to unsafe security behavior among hospital staff, circumvention of policies, and a tension between clinical and IT departments, ultimately increasing chances of BYOD security breaches. As far as higher maturity levels are concerned, they indicate a proactive security culture, where cybersecurity is seen as a top priority by management and staff alike. Therefore, hospitals at an advanced state of BYOD security handle cybersecurity affairs autonomously and may implement measures such as dedicated BYOD security training tailor-made to clinical specialty, ongoing management support, and a regular stakeholder consultative process for BYOD security strategy or policy development. Eventually, these practices can improve staff commitment and knowledge of BYOD security while also ensuring that BYOD use is made productive, in addition to being secure. A balanced approach that effectively balances productivity with security can lead to a positive security culture, thereby improving compliance and reducing risks [58-60].

The pilot implementation demonstrated how the hospital BYOD security framework can be applied to assess current practices and codevelop tailored recommendations. The maturity survey identified key gaps, while the workshop enabled stakeholder-driven discussions to guide progress from the "as-is" to the "to-be" state. This participatory approach proved valuable in addressing BYOD security challenges that are particularly important in complex and dynamic environments such as hospitals [61].

Change Management Considerations

While the maturity model provides a structured approach to hospital BYOD security, its implementation presents several challenges that hospitals must consider. Resource requirements vary depending on the hospital's current BYOD security posture, as advancing to higher maturity levels may require investment in MDM solutions, end point security tools, secure authentication systems, staff training programs, and dedicated cybersecurity personnel. Change management strategies are essential to ensure smooth adoption, particularly when introducing stricter BYOD policies, access controls, or security monitoring tools that may impact clinical workflows. Engaging stakeholders, including IT teams, clinicians, and hospital administrators, through structured communication, tailored training, and phased rollouts can minimize disruption and improve adherence. Potential resistance from stakeholders, especially clinical staff, must be addressed by demonstrating that security measures do not hinder usability or productivity. Establishing leadership buy-in and appointing clinical and IT champions can further support organizational alignment and reduce friction during implementation. Furthermore, hospitals should ensure that BYOD security policies are clinically feasible, providing secure yet user-friendly alternatives (eg, approved hospital communication apps instead of WhatsApp) to prevent the circumvention of policies through shadow IT. Integration with existing systems is another key consideration, as BYOD security measures must be compatible with EMRs, secure messaging platforms, hospital networks, and cloud-based collaboration tools without disrupting access to critical clinical information.

Significance

The maturity model serves as a valuable tool for assessing the current state of BYOD security management within a hospital, leading to a better understanding of strengths, weaknesses, and gaps in the existing security posture. It facilitates structured improvement toward a higher maturity level. Both internal and external technology or security auditors can use the model. Internal auditors, such as the hospital's IT or cybersecurity practitioners, can periodically assess BYOD security to set new improvement goals. Hospitals can also engage specialized auditors to carry out the assessment process.

Hospitals can follow a systematic approach to improve their BYOD security posture using the maturity model (Textbox 1).

In addition, government agencies, such as health departments, can use the maturity model as an audit tool to assess whether a hospital's BYOD security aligns with established standards.

XSL•FO
RenderX

Textbox 1. Improving bring your own device (BYOD) security posture using the hospital BYOD security maturity model.

1. Understanding the current maturity state: conduct a BYOD security assessment across the 21 domains. This step establishes the baseline or "as-is" state, crucial for understanding the current BYOD security position.
2. Identifying priority domains: determine critical areas or domains of BYOD security that require attention. Prioritization is essential because hospitals have finite resources and may not be able to improve all domains simultaneously. Improvement should be carried out step by step, addressing priority areas first. Involving all relevant stakeholders, especially technical and clinical personnel, ensures that user requirements are considered.
3. Setting improvement goals: develop achievable and measurable improvement goals based on the progressive maturity levels specified by the model. Identify which maturity level to aim for in each priority domain, considering hospitals' unique needs, constraints, and regulatory requirements. Improvement goals should be clearly defined using quantifiable benchmarks or key performance indicators, such as increasing SSO adoption rates, training completion percentages, or MDM enrollment levels, to track meaningful progress. Goals should be feasible, with a well-thought-out plan that gains consensus among all relevant stakeholders within the hospital.
4. Measuring progress: continuously monitor the progress of set goals against measurable, time-bound metrics set in stage 3 to systematically measure progress and ensure structured, data-driven maturity advancement. Verify that all actions required to advance to the next maturity level have been completed, providing concrete evidence of progress.
5. Revisiting maturity improvement: recognize that BYOD security management is an ongoing process requiring regular cycles of improvement. Establish a reasonable time frame to revisit and enhance controls and processes to counter evolving and complex BYOD security threats.

Limitations

The utility of the proposed hospital BYOD security maturity model has not been extensively tested across diverse hospital environments. Although significant effort was invested to enhance the model's accuracy by mapping maturity levels to recommendations derived from empirical studies and aligning them with industry standards, the model's practical effectiveness remains to be fully validated. Feedback from study participants was incorporated to refine the model further; however, additional validation and evaluation studies are required in various hospital contexts to improve its accuracy and efficacy. In addition, the rapidly changing nature of cybersecurity threats and technologies means that the model may require continuous updates to remain relevant. New types of security risks, emerging technologies, and changes in legal requirements could necessitate revisions to the model's domains and maturity levels. Another consideration is the approach to maturity assessment; while averaging assessors' ratings provides an overview, it can obscure key issues in the maturity assessment process. A structured consensus-building method, such as the Delphi approach used in the development of the Essentials of Cybersecurity in Healthcare Organizations framework [62], could enhance assessment reliability by integrating expert opinions systematically. Future research should focus on implementing such consensus-driven assessment methods alongside extensive validation studies in diverse hospital settings to refine the model, enhance its generalizability, and ensure it remains a practical tool for improving BYOD security in the ever-evolving landscape of health care cybersecurity. In addition, future validation studies may provide further insights into refining behavioral assessment methodologies, potentially incorporating more direct behavioral metrics where needed to better evaluate how security behaviors are enacted in practice.

Equal representation of both clinical and technical stakeholders was initially planned; however, due to limited availability, only 3 clinical representatives were able to participate in the workshop. Nonetheless, these clinicians had consulted with their broader teams beforehand and conveyed shared concerns and perspectives during the assessment. This imbalance in representation may have introduced a bias toward technical perspectives in the model's pilot results. Future research should aim for a more balanced stakeholder composition and a larger, more diverse sample to ensure broader applicability and validation.

Conclusions

This study developed a sociotechnical BYOD security maturity model tailored for hospitals, addressing gaps in existing cybersecurity models. By integrating technical controls, policies, and human factors, the model provides a holistic approach to managing BYOD security risks. The pilot implementation demonstrated how hospitals can use the model to create a structured, modular road map for enhancing security while balancing the unique sociotechnical complexities of health care.

Acknowledgments

The authors extend their sincere appreciation to the University of Melbourne for its support through the Melbourne Research PhD Scholarship. They are also deeply grateful to all study participants and the participating organization for their valuable contributions to the research. The authors further acknowledge and thank everyone who played a role in the successful completion of the study. Special thanks are extended to Adj Professor Frank Smolenaers for his ongoing support and guidance throughout the project.

XSL•FO
RenderX

Authors' Contributions

TAW was responsible for conceptualization, data curation, formal analysis, investigation, methodology, project administration, software, resources, visualization, project administration, and writing—original draft. AM was responsible for conceptualization, data curation, methodology, project administration, validation, resources, writing—review and editing, and supervision. KG was responsible for conceptualization, data curation, methodology, project administration, validation, resources, writing—review and editing, and supervision.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Hospital bring-your-own-device security framework recommendations and evidence.
[DOCX File , 81 KB - humanfactors_v12i1e71912_app1.docx]

Multimedia Appendix 2
Final hospital bring-your-own-device security maturity model description.
[DOCX File , 40 KB - humanfactors_v12i1e71912_app2.docx]

Multimedia Appendix 3
Maturity Assessment Survey.
[DOCX File , 23 KB - humanfactors_v12i1e71912_app3.docx]

Multimedia Appendix 4
Maturity model revisions.
[DOCX File , 19 KB - humanfactors_v12i1e71912_app4.docx]

Multimedia Appendix 5
Workshop structure and brainstorming notes.
[DOCX File , 1028 KB - humanfactors_v12i1e71912_app5.docx]

References

1. BYOD market size and share analysis - growth trends and forecasts (2025 - 2030). Mordor Intelligence. 2024. URL: https://www.mordorintelligence.com/industry-reports/byod-market [accessed 2024-07-15]
2. Wani TA, Mendoza A, Gray K, Smolenaers F. BYOD usage and security behaviour of hospital clinical staff: an Australian survey. Int J Med Inform 2022 Sep;165:104839. [doi: 10.1016/j.ijmedinf.2022.104839] [Medline: 35930927]
3. Armstrong KA, Semple JL, Coyte PC. Replacing ambulatory surgical follow-up visits with mobile app home monitoring: modeling cost-effective scenarios. J Med Internet Res 2014 Sep 22;16(9):e213 [FREE Full text] [doi: 10.2196/jmir.3528] [Medline: 25245774]
4. Moreau M, Paré G. Early clinical management of severe burn patients using telemedicine: a pilot study protocol. Pilot Feasibility Stud 2020 Jul 04;6(1):93 [FREE Full text] [doi: 10.1186/s40814-020-00637-7] [Medline: 32637150]
5. Nerminathan A, Harrison A, Phelps M, Alexander S, Scott KM. Doctors' use of mobile devices in the clinical setting: a mixed methods study. Intern Med J 2017 Mar;47(3):291-298. [doi: 10.1111/imj.13349] [Medline: 27925381]
6. Schooley B, San Nicolas-Rocca T, Burkhard R. Patient-provider communications in outpatient clinic settings: a clinic-based evaluation of mobile device and multimedia mediated communications for patient education. JMIR Mhealth Uhealth 2015 Jan 12;3(1):e2 [FREE Full text] [doi: 10.2196/mhealth.3732] [Medline: 25583145]
7. Schiff JL. The 4 biggest healthcare IT headaches. CIO. URL: https://www.cio.com/article/230060/the-4-biggest-healthcare-it-headaches.html [accessed 2018-03-01]
8. Williams J. Left to their own devices how healthcare organizations are tackling the BYOD trend. Biomed Instrum Technol 2014 Sep 02;48(5):327-339. [doi: 10.2345/0899-8205-48.5.327] [Medline: 25244198]
9. Sansurooh K, Williams P. BYOD in ehealth: herding cats and stable doors, or a catastrophe waiting to happen? In: Proceedings of the 3rd Australian eHealth Informatics and Security Conference. 2014 Presented at: AEISC '14; December 1-3, 2014; Perth Australia p. 3-23 URL: https://www.researchgate.net/publication/299367032_BYOD_in_ehealth_herding_cats_and_stable_doors_or_a_catastrophe_waiting_to_happen [doi: 10.4225/75/5798284331b46]
10. Wani TA, Mendoza A, Gray K. BYOD in hospitals-security issues and mitigation strategies. In: Proceedings of the 2019 Australasian Computer Science Week Multiconference. 2019 Presented at: ACSW '19; January 29-31, 2019; Sydney, Australia p. 1-10 URL: https://dl.acm.org/doi/10.1145/3290688.3290729 [doi: 10.1145/3290688.3290729]

XSL•FO
RenderX

11. Nunes P, Antunes M, Silva C. Evaluating cybersecurity attitudes and behaviors in Portuguese healthcare institutions. Procedia Comput Sci 2021;181:173-181. [doi: 10.1016/j.procs.2021.01.118]
12. Stobert E, Barrera D, Homier V, Kollek D. Understanding cybersecurity practices in emergency departments. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. 2020 Presented at: CHI '20; April 25-30, 2020; Honolulu, HI p. 1-8 URL: https://dl.acm.org/doi/10.1145/3313831.3376881 [doi: 10.1145/3313831.3376881]
13. Alanazi AT. Clinicians' perspectives on healthcare cybersecurity and cyber threats. Cureus 2023 Oct;15(10):e47026 [FREE Full text] [doi: 10.7759/cureus.47026] [Medline: 37965389]
14. Zhan Y, Ahmad SF, Irshad M, Al-Razgan M, Awwad EM, Ali YA, et al. Investigating the role of Cybersecurity's perceived threats in the adoption of health information systems. Heliyon 2024 Jan 15;10(1):e22947 [FREE Full text] [doi: 10.1016/j.heliyon.2023.e22947] [Medline: 38148811]
15. de-Marcos L, Cilleruelo C, Junquera-Sánchez J, Martínez-Herráiz JJ. A framework for BYOD continuous authentication: case study with soft-keyboard metrics for healthcare environment. In: Proceedings of the 3rd International Conference on Applied Informatics. 2020 Presented at: ICAI '20; October 29-31, 2020; Ota, Nigeria p. 347-358 URL: https://link.springer.com/chapter/10.1007/978-3-030-61702-8_24 [doi: 10.1007/978-3-030-61702-8_24]
16. Eke CI, Norman AA, Mulenga M. Machine learning approach for detecting and combating bring your own device (BYOD) security threats and attacks: a systematic mapping review. Artif Intell Rev 2023 Jan 17;56(8):8815-8858. [doi: 10.1007/S10462-022-10382-3]
17. Kanyi D, Ogao P. KANYI BYOD security framework: for secure access and use of mobile devices in a BYOD environment. Int J Netw Commun 2018;8(4):347-358.
18. Arregui D, Maynard S, Ahmad A. Mitigating BYOD information security risks. University of Melbourne. 2015. URL: http://minerva-access.unimelb.edu.au/handle/11343/56627 [accessed 2018-03-01]
19. Zahadat N, Blessner P, Blackburn T, Olson BA. BYOD security engineering: a framework and its analysis. Comput Secur 2015 Nov;55:81-99. [doi: 10.1016/j.cose.2015.06.011]
20. Siponen MT, Oinas-Kukkonen H. A review of information security issues and respective research contributions. Data Base Adv Inf Syst 2007 Feb 28;38(1):60-80. [doi: 10.1145/1216218.1216224]
21. Davis MC, Challenger R, Jayewardene DN, Clegg CW. Advancing socio-technical systems thinking: a call for bravery. Appl Ergon 2014 Mar;45(2):171-180 [FREE Full text] [doi: 10.1016/j.apergo.2013.02.009] [Medline: 23664481]
22. Malatji M, Von Solms S, Marnewick A. Socio-technical systems cybersecurity framework. Inf Comput Secur 2019 Jun 12;27(2):233-272. [doi: 10.1108/ics-03-2018-0031]
23. Wani TA, Mendoza A, Gray K. Hospital bring-your-own-device security challenges and solutions: systematic review of gray literature. JMIR Mhealth Uhealth 2020 Jun 18;8(6):e18175 [FREE Full text] [doi: 10.2196/18175] [Medline: 32554388]
24. Wani TA, Mendoza A, Gray K, Smolenaers F. Status of bring-your-own-device (BYOD) security practices in Australian hospitals – a national survey. Health Policy Technol 2022 Sep;11(3):100627. [doi: 10.1016/j.hlpt.2022.100627]
25. Wani TA, Mendoza A, Gray K. BYOD security practices in Australian hospitals – a qualitative study. In: Proceedings of the 6th International Conference on HCI for Cybersecurity, Privacy and Trust. 2024 Presented at: HCII '24; June 29-July 4, 2024; Washington, DC p. 138-158 URL: https://link.springer.com/chapter/10.1007/978-3-031-61379-1_10 [doi: 10.1007/978-3-031-61379-1_10]
26. Akinsanya OO, Papadaki M, Sun L. Current cybersecurity maturity models: how effective in healthcare cloud? In: Proceedings for the 27th International Conference on Case-Based Reasoning co-located with the 27th International Conference on Case-Based Reasoning. 2019 Presented at: ICCBR '19; September 8-12, 2019; Otzenhausen, Germany p. 211-222 URL: https://pearl.plymouth.ac.uk/secam-research/905/
27. Le NT, Hoang DB. Capability maturity model and metrics framework for cyber cloud security. Scalable Comput Pract Exp 2017 Nov 24;18(4):277-290. [doi: 10.12694/scpe.v18i4.1329]
28. Al-Matari OM, Helal IM, Mazen SA, Elhennawy S. Adopting security maturity model to the organizations' capability model. Egypt Inform J 2021 Jul;22(2):193-199. [doi: 10.1016/j.eij.2020.08.001]
29. ISO/IEC 27002:2022(en): information security, cybersecurity and privacy protection — information security controls. International Organization for Standardization. URL: https://www.iso.org/obp/ui/en/#iso:std:iso-iec:27002:ed-3:v2:en [accessed 2021-06-18]
30. Bowen P, Kissel R. Program Review for Information Security Management Assistance (PRISMA). National Institute of Standards and Technology. 2007. URL: https://csrc.nist.gov/publications/detail/nistir/7358/final [accessed 2025-05-29]
31. CMMI cybermaturity platform. ISACA. URL: https://www.isaca.org/enterprise/cmmi-cybermaturity-platform [accessed 2021-06-18]
32. Mohammed I, Bade AM. Cybersecurity capability maturity model for network system. Int J Dev Res 2019;09:22-27 [FREE Full text]
33. Büyüközkan G, Güler M. Cybersecurity maturity model: systematic literature review and a proposed model. Technol Forecast Soc Change 2025 Apr;213:123996. [doi: 10.1016/j.techfore.2025.123996]
34. AlDaajeh S, Alrabaee S. Strategic cybersecurity. Comput Secur 2024 Jun;141:103845. [doi: 10.1016/j.cose.2024.103845]
35. Bitzer M, Häckel B, Leuthe D, Ott J, Stahl B, Strobel J. Managing the inevitable – a maturity model to establish incident response management capabilities. Comput Secur 2023 Feb;125:103050. [doi: 10.1016/j.cose.2022.103050]

XSL•FO
RenderX

36. Smyrlis M, Floros E, Basdekis I, Prelipcean DB, Sotiropoulos A, Debar H, et al. RAMA: a risk assessment solution for healthcare organizations. Int J Inf Secur 2024 Mar 01;23(3):1821-1838. [doi: 10.1007/S10207-024-00820-4]
37. Carayon P, Schoofs Hundt A, Karsh B, Gurses AP, Alvarado CJ, Smith M, et al. Work system design for patient safety: the SEIPS model. Qual Saf Health Care 2006 Dec;15 Suppl 1(Suppl 1):i50-i58 [FREE Full text] [doi: 10.1136/qshc.2005.015842] [Medline: 17142610]
38. Harrison MI, Koppel R, Bar-Lev S. Unintended consequences of information technologies in health care--an interactive sociotechnical analysis. J Am Med Inform Assoc 2007;14(5):542-549 [FREE Full text] [doi: 10.1197/jamia.M2384] [Medline: 17600093]
39. Sittig DF, Singh H. A new socio-technical model for studying health information technology in complex adaptive healthcare systems. In: Patel VL, Kannampallil TG, Kaufman DR, editors. Cognitive Informatics for Biomedicine: Human Computer Interaction in Healthcare. Cham, Switzerland: Springer; 2015:59.
40. Creswell JW, Tashakkori A. Editorial: differing perspectives on mixed methods research. J Mix Methods Res 2007 Oct 01;1(4):303-308. [doi: 10.1177/1558689807306132]
41. Ivankova NV. Applying mixed methods in community-based participatory action research: a framework for engaging stakeholders with research as a means for promoting patient-centredness. J Res Nurs 2017 Apr 18;22(4):282-294. [doi: 10.1177/1744987117699655]
42. Ivankova N, Wingo N. Applying mixed methods in action research: methodological potentials and advantages. Am Behav Sci 2018 Apr 27;62(7):978-997. [doi: 10.1177/0002764218772673]
43. Schlarman S. The people, policy, technology (PPT) model: core elements of the security process. Inf Syst Secur 2006 Dec 21;10(5):1-6. [doi: 10.1201/1086/43315.10.5.20011101/31719.6]
44. Akinsanya OO, Papadaki M, Sun L. Towards a maturity model for health-care cloud security (M2HCS). Inf Comput Secur 2019 Dec 16;28(3):321-345. [doi: 10.1108/ics-05-2019-0060]
45. Essential eight maturity model. Australian Signals Directorate. URL: https://www.cyber.gov.au/acsc/view-all-content/publications/essential-eight-maturity-model [accessed 2025-05-29]
46. Mohamad Stambul MA, Razali R. An assessment model of information security implementation levels. In: Proceedings of the 2011 International Conference on Electrical Engineering and Informatics. 2011 Presented at: ICEEI '11; July 17-19, 2011; Bandung, Indonesia p. 1-6 URL: https://ieeexplore.ieee.org/document/6021561 [doi: 10.1109/iceei.2011.6021561]
47. Pilemalm S, Timpka T. Third generation participatory design in health informatics--making user participation applicable to large-scale information system projects. J Biomed Inform 2008 Apr;41(2):327-339 [FREE Full text] [doi: 10.1016/j.jbi.2007.09.004] [Medline: 17981514]
48. Zamenopoulos T, Alexiou K. Co-Design As Collaborative Research. Bristol, UK: University of Bristol/AHRC; 2018.
49. Thoring K, Mueller R, Badke-Schaub P. Workshops as a research method: guidelines for designing and evaluating artifacts through workshops. In: Proceedings of the 53rd Hawaii International Conference on System Sciences. 2020 Presented at: HICSS '20; January 7-10, 2020; Maui, HI p. 5036-5045 URL: https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/c78f65c6-8097-4099-a2a5-09b384c17e69/content [doi: 10.24251/hicss.2020.620]
50. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013 Sep 18;13:117 [FREE Full text] [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]
51. Sabbagh BA, Kowalski S. ST(CS)2 - featuring socio-technical cyber security warning systems. In: Proceedings of the 2012 International Conference on Cyber Security, Cyber Warfare and Digital Forensic. 2012 Presented at: CyberSec '12; June 26-28, 2012; Kuala Lumpur, Malaysia p. 312-316 URL: https://ieeexplore.ieee.org/document/6246110 [doi: 10.1109/cybersec.2012.6246110]
52. Siponen M, Willison R. Information security management standards: problems and solutions. Inf Manage 2009 Jun;46(5):267-270. [doi: 10.1016/j.im.2008.12.007]
53. Soomro ZA, Shah MH, Ahmed J. Information security management needs more holistic approach: a literature review. Int J Inf Manag 2016 Apr;36(2):215-225. [doi: 10.1016/j.ijinfomgt.2015.11.009]
54. McLeod A, Dolezel D. Cyber-analytics: modeling factors associated with healthcare data breaches. Decis Support Syst 2018 Apr;108:57-68. [doi: 10.1016/j.dss.2018.02.007]
55. Hajdarevic K, Allen P, Spremic M. Proactive security metrics for bring your own device (BYOD) in ISO 27001 supported environments. In: Proceedings of the 24th International Conference on Telecommunications Forum. 2016 Presented at: TELFOR '16; November 22-23, 2016; Belgrade, Serbia p. 4 URL: https://ieeexplore.ieee.org/document/7818717 [doi: 10.1109/telfor.2016.7818717]
56. Donalds C, Barclay C. Beyond technical measures: a value-focused thinking appraisal of strategic drivers in improving information security policy compliance. Eur J Inf Syst 2021 Sep 21;31(1):58-73. [doi: 10.1080/0960085x.2021.1978344]
57. Tøndel IA, Line MB, Jaatun MG. Information security incident management: current practice as reported in the literature. Comput Secur 2014 Sep;45:42-57. [doi: 10.1016/j.cose.2014.05.003]
58. Mikuletič S, Vrhovec S, Skela-Savič B, Žvanut B. Security and privacy oriented information security culture (ISC): explaining unauthorized access to healthcare data by nursing employees. Comput Secur 2024 Jan;136:103489. [doi: 10.1016/j.cose.2023.103489]

XSL•FO
RenderX

59. Parsons K, Calic D, Pattinson M, Butavicius M, McCormac A, Zwaans T. The Human Aspects of Information Security Questionnaire (HAIS-Q): two further validation studies. Comput Secur 2017 May;66:40-51. [doi: 10.1016/j.cose.2017.01.004]
60. Sohrabi Safa N, Von Solms R, Furnell S. Information security policy compliance model in organizations. Comput Secur 2016 Feb;56:70-82. [doi: 10.1016/j.cose.2015.10.006]
61. Pollini A, Callari TC, Tedeschi A, Ruscio D, Save L, Chiarugi F, et al. Leveraging human factors in cybersecurity: an integrated methodological approach. Cogn Technol Work 2022 Jun 11;24(2):371-390 [FREE Full text] [doi: 10.1007/s10111-021-00683-y] [Medline: 34149309]
62. O'Brien N, Grass E, Martin G, Durkin M, Darzi A, Ghafur S. Developing a globally applicable cybersecurity framework for healthcare: a Delphi consensus study. BMJ Innov 2020 Nov 23;7(1):199-207. [doi: 10.1136/bmjinnov-2020-000572]

Abbreviations

BYOD: bring your own device
EMR: electronic medical record
hBYOD: hospital bring your own device
IAM: identity and access management
MAM: mobile application management
MDM: mobile device management
MMAR: mixed methods action research
PPT: people, process, and technology
SSO: single sign-on

Edited by H Witteman; submitted 29.01.25; peer-reviewed by L Coventry, L Ma; comments to author 10.03.25; revised version received 25.03.25; accepted 12.05.25; published 13.08.25.

Please cite as:
Wani TA, Mendoza A, Gray K
A Sociotechnical Approach to Bring-Your-Own-Device Security in Hospitals: Development and Pilot Testing of a Maturity Model Using Mixed Methods Action Research
JMIR Hum Factors 2025;12:e71912
URL: https://humanfactors.jmir.org/2025/1/e71912
doi:10.2196/71912
PMID:40802372

©Tafheem Ahmad Wani, Antonette Mendoza, Kathleen Gray. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 13.08.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Diverse Attitudes and Experiences With Technology Use During the COVID-19 Pandemic Among Asian American and Pacific Islander Adults (the COMPASS Study): Survey Study

Linda G Park[1], MS, PhD; Serena Chi[2], BA; Myka Lay[3], BS; Nicole Phan[1], BS; Janice Y Tsoh[4,5,6], PhD; Oanh L Meyer[7], MAS, PhD; Bora Nam[1], MSN, PhD; Van Ta Park[1], MPH, PhD

[1]Department of Community Health Systems, University of California, San Francisco, School of Nursing, San Francisco, CA, United States

[2]School of Medicine, Yale University, New Haven, CT, United States

[3]Samuel Merritt University, Oakland, CA, United States

[4]Department of Psychiatry and Behavioral Sciences, University of California, San Francisco, School of Medicine, San Francisco, CA, United States

[5]Asian American Research Center on Health, University of California, San Francisco, San Francisco, CA, United States

[6]Multiethnic Health Equity Research Center, University of California, San Francisco, San Francisco, CA, United States

[7]Department of Neurology, University of California, Davis, School of Medicine, Sacramento, CA, United States

Corresponding Author:
Linda G Park, MS, PhD
Department of Community Health Systems
University of California, San Francisco
School of Nursing
490 Illinois Street, Floor 9
San Francisco, CA, 94158
United States
Phone: 1 415 502 6616
Email: linda.park@ucsf.edu

Abstract

Background: The COVID-19 pandemic forced the world to quarantine to slow the rate of transmission, causing communities to transition into virtual spaces. Asian American and Pacific Islander communities faced the additional challenge of discrimination that stemmed from racist and xenophobic rhetoric in the media. Limited data exist on technology use among Asian American and Pacific Islander adults during the height of the COVID-19 shelter-in-place period and its effect on their physical and mental health.

Objective: This study aims to examine Asian American and Pacific Islander adults' attitudes, perspectives, and experiences regarding their use of technology during the COVID-19 pandemic.

Methods: We collaborated with community partners and used social media to distribute the COVID-19 Effects on the Mental and Physical Health of Asian Americans and Pacific Islanders Survey Study, a nationwide multilingual survey available in English, Chinese, Korean, Samoan, and Vietnamese. The survey was administered from October 2020 to February 2021, and participants rated their level of agreement (1=*not at all* to 5=*extremely*) on 6 items assessing their attitudes toward technology use. Thematic analysis was conducted on responses to the open-ended question "Is there anything else you want to tell us about your use of technology during COVID-19?" The qualitative responses were reviewed, analyzed, coded, and organized into corresponding themes.

Results: The mean age of respondents was 45.9 (SD 16.3; range 18-98) years, with 5398 participants completing the quantitative survey and 1115 (20.66%) providing unique responses to the open-ended question. In the quantitative survey, 68% (3671/5398) of the respondents reported being comfortable using technology; the majority indicated that it helped them keep up with the news (4318/5398, 79.99%), maintain social connections (4102/5398, 75.99%), and provide care for others (2537/5398, 46.99%). However, responses were mixed regarding the usefulness of technology for health: 39.99% (2159/5398) agreed that it was helpful for mental health but disagreed regarding physical health. Four main themes emerged from the qualitative analysis: (1) technology was critical for functioning across many aspects of life and maintaining physical, mental, and emotional well-being; (2) technology was often the only means of interpersonal social connections; (3) overuse led to negative physical and mental health outcomes; and (4) technology use was associated with multiple challenges and barriers.

XSL•FO
RenderX

Conclusions: Our findings revealed diverse perspectives and experiences related to technology use by Asian American and Pacific Islander adults during the height of the COVID-19 pandemic. Dependence on technology may have exacerbated social inequities, particularly for those with lack of access to devices and Wi-Fi and limited English proficiency, affecting their ability to work, apply for jobs, and communicate virtually. Further qualitative research would be beneficial in amplifying the perspectives of Asian American and Pacific Islander adults to uncover concerns and address health disparities.

(JMIR Hum Factors 2025;12:e64999) doi:10.2196/64999

KEYWORDS

COVID-19; technology use; Asian American and Pacific Islander; mental health; physical health; social connections; virtual space; health disparities; thematic analysis; community partners; social media; multilingual survey; qualitative research; SARS-COV-2; coronavirus; infectious; pandemic; Asian American; Pacific Islander; mixed methods; COVID-19 Effects on the Mental and Physical Health of Asian Americans and Pacific Islanders Survey Study; COMPASS; adult; mobile health; mHealth

Introduction

Background

The impact of technology on society in the realm of health care access and delivery has been profound and multifaceted. The dependence on technology was magnified during the COVID-19 pandemic, underscoring the critical role that technology plays in nearly every form of communication. The pandemic caused widespread loss of life, a myriad of health consequences, and unprecedented physical and social isolation [1]. During this period, technology served as a vital tool for accessing health care, continuing remote work and education, and operating businesses. To protect older adults and others considered vulnerable, shelter-in-place policies were implemented to varying degrees across the world. As a result, many were forced to rely on virtual forms of communication, such as videoconferencing and online chat. In most cases, virtual communication became the only means of social interaction.

The pandemic also brought the digital divide to the forefront. Disparities related to COVID-19 and health care access were exacerbated for rural communities, racial and ethnic minority groups, individuals with low income, and other populations considered vulnerable. These groups were disproportionately burdened by the digital divide, compounding existing structural disadvantages [2]. These inequities may have been particularly acute for individuals with limited English proficiency in virtual workspaces and job interviews, where reliance on verbal communication can heighten communication challenges [3]. Research has also revealed that the widening of the digital divide intensified challenges, especially for rural residents [4]. However, to our knowledge, limited research exists on the use of digital technology among Asian American and Pacific Islander populations in the United States during the pandemic, particularly among older adults (aged ≥60 y) and adults with limited English proficiency. Asian American and Pacific Islander populations were especially vulnerable during the pandemic, as xenophobic attitudes became more prevalent due to racist statements linking China to COVID-19 [5]. As a result, many individuals from these communities experienced racism, discrimination, and hate crimes, which contributed to fear, loneliness, and social isolation [6,7].

A Pew Research Center survey of American adults conducted in April 2021 revealed that the internet served as a "lifeline" for many during the pandemic, but respondents also reported "struggles" associated with the increased use of and reliance on the internet [8]. Such surveys provide insights into general patterns of digital technology use, purposes, and associated challenges among the general American adult population. However, limited data exist on digital technology use patterns, perceived health-related support, and attitudes toward digital technology use among Asian American and Pacific Islander adults. Examining these gaps may assist in culturally and linguistically tailoring appropriate digital technology programs for these diverse populations. In addition, a better understanding of the impact of digital technology use is needed to inform clinical recommendations for physical and mental health.

Our prior work with the COVID-19 Effects on the Mental and Physical Health of Asian Americans and Pacific Islanders Survey Study (COMPASS) revealed that increased technology use during the pandemic was associated with poorer mental and physical health outcomes among participants [9]. Specifically, those who reported increasing their technology use by ≥7 hours per day experienced higher levels of depressive symptoms and worse physical health, regardless of age or ethnic background [9]. These findings suggest that even with high levels of digital access, increased technology use does not necessarily equate to positive health outcomes. Still, little is known about digital technology use among Asian American and Pacific Islander adults. In the COMPASS study, the majority of Asian American and Pacific Islander participants reported owning a smartphone (96.2%), home internet access (90%), a desktop or laptop computer (83.4%), a tablet (61.1%), cable television (43.5%), and a streaming device (45.7%).

Objectives

Building on previous findings, we aimed to address gaps in knowledge regarding technology use among Asian American and Pacific Islander populations by analyzing data from COMPASS, which examined mental and physical health implications associated with technology use during the height of the COVID-19 shelter-in-place period among Asian American and Pacific Islander adults in the United States. In this study, we quantitatively describe COMPASS participants' attitudes toward technology use and qualitatively explore their perspectives and experiences of technology use during the COVID-19 pandemic.

XSL•FO
RenderX

Methods

Recruitment

A total of 5411 individuals participated in the study from October 2020 to February 2021. The survey was offered on the COMPASS website [10] in English, simplified Chinese, traditional Chinese, Korean, Vietnamese, and Samoan or by telephone with assistance from COMPASS staff and community partners (Multimedia Appendix 1). The survey was also offered by telephone in Cantonese, Mandarin, Vietnamese, Korean, and Samoan by multilingual COMPASS staff. The survey was shared through personal and professional networks via social media, email, telephone, health promotion webinars, flyers, and ethnic media across the United States. The 26 COMPASS community advisory board members and 15 community partners that serve Asian American and Pacific Islander populations supported recruitment efforts by promoting the survey among their networks. In addition, recruitment leveraged the Collaborative Approach for Asian Americans and Pacific Islanders Research and Education (CARE) registry [11,12], contacting eligible participants interested in participating in health research via email and telephone. The goal of the CARE registry was to increase the participation of Asian American and Pacific Islander individuals in research related to aging, Alzheimer disease and related dementias, and caregiving.

Ethical Considerations

The COMPASS research project was approved by the institutional review board of the University of California San Francisco (UCSF; 20-31925). Informed consent was provided by participants before survey completion. Participants had the option to receive a US $10 gift card upon completing the survey.

Eligibility and Procedures

Eligible participants were required to be aged ≥18 years; reside in the United States; and identify as Asian, Asian American, Native Hawaiian, Pacific Islander, or any combination of these with other racial or ethnic identities. Participants were also required to be able to read in English, Chinese (simplified or traditional), Korean, Samoan, or Vietnamese. We used the World Health Organization's instrument adaptation process [13], including backward and forward translation, to guide the multilanguage survey's scale and material development. All translated instruments were reviewed by multiple research team members for accuracy and cultural relevance. Non-English survey responses were translated into English through online translation services before analysis.

Methodology, Data Collection, and Measures

We applied a cross-sectional study design that included COMPASS quantitative survey data and qualitative data from an open-ended question asking, "Is there anything else you want to tell us about your use of technology during COVID-19?" Free-text responses of any length were accepted. The data were stored in REDCap (Research Electronic Data Capture; Vanderbilt University) tools hosted at UCSF [14,15]. For sociodemographic information, participants were queried about their age, sex, ethnic or cultural background, marital status, country of birth, education, employment status, annual household income, and English proficiency.

The technology use survey was developed by the COMPASS research team to explore patterns of technology use during the COVID-19 pandemic. Participants were queried about the number of hours they used technology per day and whether their technology use increased during the COVID-19 crisis. Responses were categorized into five groups: (1) "did not increase," (2) "increased by 1 to 2 hours per day," (3) "increased by 3 to 4 hours per day," (4) "increased by 5 to 6 hours per day," and (5) "increased by ≥7 hours per day." In addition, participants were asked about their comfort level with using technology and their engagement in various technology-mediated activities during the COVID-19 pandemic. These activities included video chatting with friends or family, video meetings for work-related activities, telephone or video visits with health care providers, exercise or physical fitness activities, mental health activities, keeping in touch with friends and family through social media, leisure activities or hobbies, and accessing news. Moreover, participants were asked to evaluate the helpfulness of technology. The survey encompassed statements concerning the use of technology for keeping up with the news, maintaining social connections, caregiving, supporting mental health, and enhancing physical health. Respondents indicated their level of agreement with each statement on a 5-point Likert scale ranging from 1=*not at all* to 5=*extremely*.

Analysis

Descriptive statistics were used to report the sociodemographic characteristics and technology use patterns of the sample. Qualitative analysis was conducted using content analysis. Two trained independent reviewers initially organized the raw participant responses and analyzed the data to identify subthemes and categorize significant statements. Using an iterative process and a constant comparative method, the research team refined the coding scheme, finalized the themes, and identified patterns and relationships across the qualitative data.

Coding and organization of interview data were conducted using Microsoft Word and Microsoft Excel. A third coder independently cross-checked the transcripts to ensure the consistency of the interpretation and summarized the coded content within thematic categories. After this initial round of analysis, the coders collaboratively reviewed subthemes and worked together to consolidate them under broader thematic umbrellas. Through continued discussion, 5 major themes were identified. Codes were developed to support each theme, with refined subthemes providing additional structure. Participant responses were then organized according to these codes to illustrate and support the thematic findings.

The first coder (ML) is a board-certified registered nurse and research analyst at UCSF, with extensive experience in public health, clinical nursing, and Asian American and Pacific Islander health promotion initiatives. Their professional background spans diverse health care settings, with a focus on community health; health equity; and holistic, family-centered care. The second coder (NP) is a project policy analyst at UCSF with a strong background in qualitative research and experience

XSL•FO
RenderX

engaging with research participants in both English and languages across several Asian American and Pacific Islander communities. Their expertise includes community-based participatory research, coordinating outreach initiatives, and developing volunteer programs for Asian American and Pacific Islander students. A third team member (LGP), also experienced in qualitative methods, served as an adjudicator to resolve discrepancies in data interpretation and assist in finalizing the thematic structure. The diverse backgrounds of all three coders informed their analytical lens, particularly in identifying sociocultural influences and health disparities within participants' narratives. Their lived experiences and cultural identities contributed to a reflexive and culturally responsive approach to qualitative analysis.

Results

Participant Characteristics

Only individuals who provided responses to the open-ended question (n=1115) were included in the analysis. The mean age of the participants was 45.9 (SD 16.3; range 18-98) years. The majority identified as Vietnamese (399/1115, 35.78%) and ethnic Chinese (336/1115, 30.13%), followed by those who identified as Korean (161/1115, 14.44%) and Filipino (69/1115, 6.19%). Most of the participants completed the qualitative question in the survey in English (665/1115, 59.64%), identified as female (731/1115, 65.56%), were either married or living with a partner (723/1115, 64.84%), and were born in a foreign country (776/1115, 69.6%). In addition, a significant portion of participants reported limited English proficiency (322/1115, 28.88%). Table 1 provides an overview of the demographic characteristics of the participants.

XSL•FO
RenderX

Table 1. Descriptive characteristics of the participants (n=1115).

Characteristics	Participants
Age (y), mean (SD; range)	45.9 (16.3; 18-98)
Age groups (y), n (%)	
<30	185 (16.6)
30-39	150 (13.4)
40-49	198 (17.8)
50-59	227 (20.4)
≥60	355 (31.8)
Questionnaire asked for cultural group, n (%)	
Asian Indian	68 (6.1)
Ethnic Chinese[a]	336 (30.1)
Filipino	69 (6.2)
Hmong	98 (2.1)
Japanese	61 (5.5)
Korean	161 (14.4)
Native Hawaiian or Pacific Islander	15 (1.3)
Vietnamese	399 (35.8)
Other or mixed	83 (7.5)
Survey response language, n (%)	
English	665 (59.6)
Chinese	78 (7)
Korean	81 (7.3)
Vietnamese	288 (25.8)
Samoan	3 (0.3)
Sex, n (%)	
Female	731 (65.6)
Male	380 (34.1)
Other or declined to answer	40 (0.9)
Country of birth, n (%)	
United States	338 (30.3)
Outside the United States	776 (69.6)
Don’t know	1 (0.1)
Limited English proficiency, n (%)	
Yes	322 (28.9)
No	793 (71.1)
Marital status, n (%)	
Single	275 (24.7)
Married or living with partner	723 (64.8)
Separated, divorced, or widowed	113 (10.1)
Declined to answer	4 (0.4)
Employment status, n (%)	
Full time	441 (39.6)
Part time	185 (16.6)

XSL•FO
RenderX

Characteristics	Participants
Homemaker	108 (9.7)
Unemployed	134 (12)
Retired	207 (18.6)
Other or declined to answer	82 (7.4)
Education, n (%)	
High school or less	246 (22.1)
Some college or technical school	162 (14.5)
Bachelor's degree	379 (34)
Master's degree or higher	318 (28.5)
Declined to answer	10 (0.9)
Household income (US $), n (%)	
≤25,000	282 (25.3)
>25,000-75,000	337 (30.2)
>75,000-150,000	232 (20.8)
>150,000	77 (6.9)
Declined to state	97 (8.7)

[a]Ethnic Chinese includes individuals who identified as mainland Chinese, Hong Kongers, Taiwanese, or Huaren.

The participants' technology use characteristics are presented in Table 2. The majority (644/1115, 57.8%) reported using technology for 1 to 8 hours per day. A notable portion (803/1115, 72%) reported increased technology use due to the COVID-19 pandemic. More than half of the respondents (627/1115, 56.23%) said that they were very or extremely comfortable using technology. During the pandemic, the most common activity was video chatting with friends or family, with 902 (81%) of the 1115 participants using platforms such as Zoom, WhatsApp, or FaceTime for communication. In addition, 614 (55.07%) participants engaged in video meetings for work-related activities, while 575 (51.57%) used technology for telephone or video visits with health care providers. Exercise or physical fitness activities were performed by 420 (37.67%) of the 1115 participants, while 187 (16.77%) reported engaging in mental health activities through technology. Social media served as a means of keeping in touch with friends and family for 866 (77.67%) of the 1115 participants, and 517 (46.37%) used technology for leisure activities or hobbies. Furthermore, 850 (76.23%) participants accessed news through technology platforms.

XSL•FO
RenderX

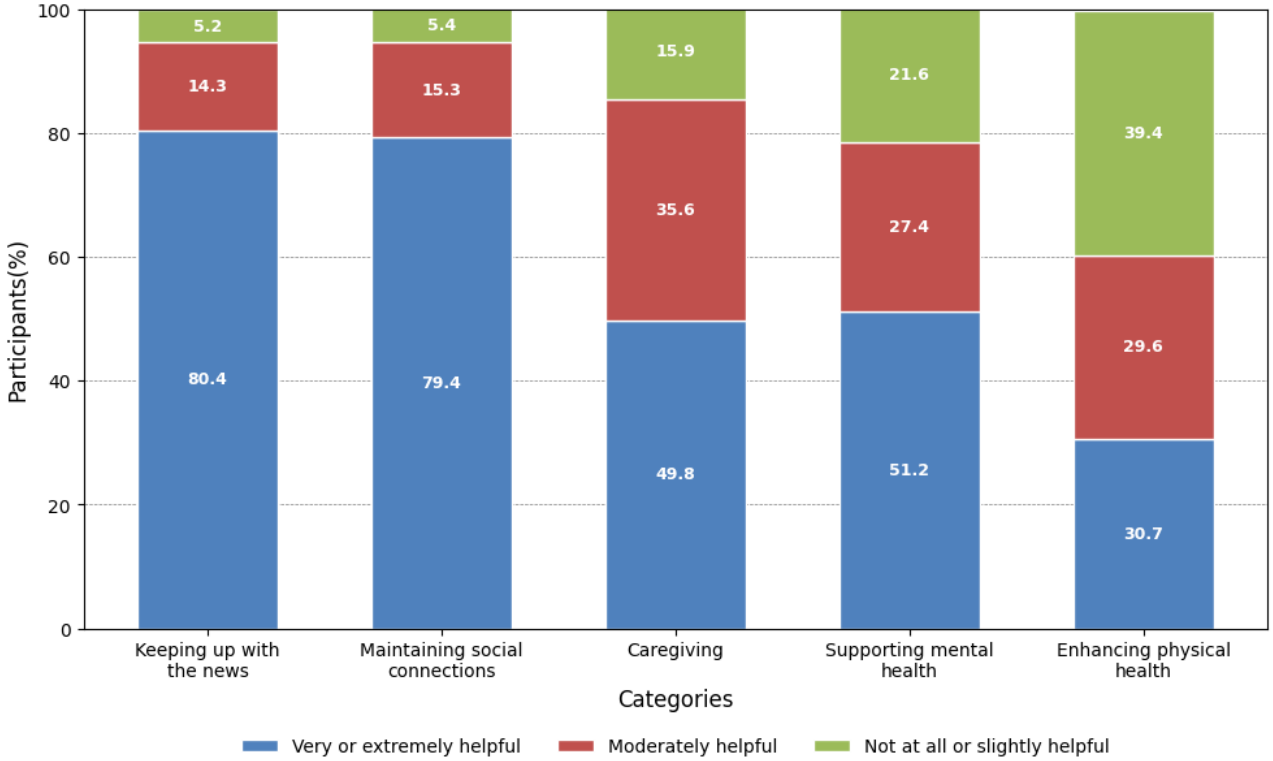
Table 2. Self-reported technology use of participants (n=1115).

Characteristics	Participants, n (%)
Total technology use (h/d)	
<1	66 (5.9)
1-4	321 (28.8)
5-8	323 (29)
9-12	238 (21.4)
13-16	120 (10.8)
≥17	46 (4.1)
Change in technology use due to the COVID-19 pandemic	
Decreased or stayed the same	292 (26.2)
Increased by 1-2 h/d	275 (24.7)
Increased by 3-4 h/d	311 (28)
Increased by 5-6 h/d	128 (11.5)
Increased by ≥7 h/d	89 (8)
Comfortable using technology	
Not at all	204 (18.4)
Slightly	91 (8.2)
Moderately	187 (16.9)
Very	301 (27.1)
Extremely	326 (29.4)
Types of technology used during the COVID-19 pandemic	
Video chatting with friends or family (eg, Zoom, WhatsApp, and FaceTime)	902 (80.9)
Video meetings for work-related activities	614 (55.1)
Telephone or video visit with health care provider	575 (51.6)
Exercise or activities for physical fitness	420 (37.7)
Mental health activities	187 (16.8)
Keeping in touch with friends and family through social media	866 (77.7)
Leisure activities or hobbies	517 (46.4)
Accessing news	850 (76.2)
None	20 (1.8)
Other (please specify)	20 (1.8)

The findings (Figure 1) indicate that the majority of participants (896/1115, 80.35%) found staying informed by accessing news through technology very or extremely helpful. In terms of keeping in touch with social connections, a significant majority (885/1115, 79.37%) perceived technology as very or extremely helpful, with only a small proportion (60/1115, 5.38%) reporting slightly or not at all helpful. Similarly, in the context of caregiving, a substantial portion of participants (938/1115, 84.12%) viewed technology as moderately or extremely helpful. However, perceptions regarding mental health support through technology varied, with a notable portion of participants (874/1115, 78.38%) finding it moderately or extremely helpful, while others (241/1115, 21.61%) indicated that technology was slightly or not helpful. Finally, participants' perceptions of technology's helpfulness for physical health varied, with a considerable proportion (565/1115, 50.67%) reporting technology being moderately or very helpful, while others (439/1115, 39.37%) expressed that technology was slightly or not helpful.

XSL•FO
RenderX

Figure 1. Survey responses to whether using technology was helpful in different aspects of life during the COVID-19 pandemic (n=1115).



Technology use among participants revealed a spectrum of use patterns (Table 2). Notably, 36.23% (404/1115) of the participants reported using technology for ≥9 hours per day. Specifically, 5.91% (66/1115) of the participants used technology for <1 hour per day, while 28.78% (321/1115) used it for 1 to 4 hours per day, 21.34% (238/1115) for 5 to 8 hours per day, 21.34% (238/1115) for 9 to 12 hours per day, 10.76% (120/1115) for 13 to 16 hours per day, and 4.12% (46/1115) for ≥17 hours per day.

Themes Related to Technology Use

Of the 5411 participants who completed the quantitative survey, 1115 (20.66%) provided responses to the open-ended question "Is there anything else you want to tell us about your use of technology during COVID-19?" A thematic analysis revealed four main themes related to digital technology use during the COVID-19 pandemic: (1) technology was critical for functioning across many aspects of life and maintaining physical, mental, and emotional well-being; (2) technology was often the only means of interpersonal social connections; (3) overuse led to negative physical and mental health outcomes; and (4) technology use was associated with multiple challenges and barriers.

Technology Was Critical for Functioning Across Many Aspects of Life and Maintaining Physical, Mental, and Emotional Well-Being

Overview

Technology use was found to be critical in maintaining daily activities related to skills and hobbies, work, and fitness during the pandemic. All these functions were reported to be vital for maintaining mental health, given the limited access to physical connection and ability to socialize outside of their homes. Participants reported feeling very isolated and believed that they would not have coped as well as they did without technology:

> *Without technology (internet, social networks, video streaming, online shopping), I'm not sure how I could survive this shelter-in-place.* [Ethnic Chinese participant, male, aged 44 y]

This reflects a common theme of isolation and disconnection experienced during the lockdown and shelter-in-place period. Many individuals were forced to physically distance from their community networks, resulting in increased reliance on technology to connect with others. In addition, many were unable to carry out routine errands, such as shopping, making delivery services their only option. One participant similarly emphasized her dependence on technology:

> *I would have gone absolutely insane without the internet during this quarantine.* [Filipino participant, female, aged 40 y]

During the lockdown, for many Asian American and Pacific Islander individuals, online access was the only means of connection to others. The internet was the safest way to follow public health guidelines, protect themselves and their loved ones from COVID-19 infection, and maintain their mental and emotional well-being. In terms of helpful aspects of technology, a participant described using social media to spread information and raise awareness on social issues relevant to Asian American and Pacific Islander individuals:

> *I used social media to spread information about anti-Asian discrimination and Black Lives Matter.* [Vietnamese participant, female, aged 49 y]

XSL•FO
RenderX

This statement reflects the broader context of the time, including the emotional unrest associated with the lockdown and the concurrent political movements and cultural propaganda directed toward the Asian American and Pacific Islander community. It highlights how these challenges were experienced by many individuals in this community and how they used social media as a tool for connection and raising awareness.

Skills and Hobbies

Many participants found technology to be beneficial for learning new skills and hobbies to pass time and for entertainment during the lockdown. These activities helped alleviate boredom and loneliness and created opportunities to explore new interests, for example, learning to cook, acquiring linguistic skills, and learning to play musical instruments. Such engagement was reported to help ease stress and remain occupied during the lockdown:

> *In my personal experience, the use of technology allowed me to learn and be exposed to people and places that may not have been possible. Also, I've become more knowledgeable with technology.* [Filipino participant, female, aged 62 y]

This quote illustrates how participants used technology to adapt and find new ways to connect with others as well as engage in new activities to promote positive habits and increase knowledge during the lockdown.

Another participant described how technology supported the continuation of daily activities and helped ease pandemic-related stress:

> *I have become better at the interplay between technology & non-technology- using technology to enhance aspects of my non-technology life like learning instruments or exercise.* [Chinese participant, female, aged 47 y]

This comment reinforces the theme of participants using technology to master new skills and hobbies to improve their quality of life during the shelter-in-place period.

Work

Learning new skills also included improving technology skills, such as adapting to online platforms for work and school. This included learning to navigate various tools and interfaces such as Zoom, Google Classroom, Google Workspace, Microsoft Office, and others:

> *I have had to learn to use a lot of technology tools for work within a very short period of time - Google classroom, Google slides, Google docs, Google drive, Ladibug document camera, Screencastify, Boom cards, interactive pdf, and of course, Zoom.* [Chinese (Huaren) participant, female, aged 51 y]

This quote reflects the transition to online forums and the use of new technological tools and how participants were required to quickly adapt to maintain productivity for work and school tasks during the lockdown.

Other benefits related to work included reduced commuting time and costs, which participants associated with increased efficiency and more time to focus on work-related tasks:

> *Good for my work I don't have to commute everyday no Bart [train]/bus fares don't have to wake up early to prepare going to work.* [Filipino participant, female, aged 68 y]

This supports the finding that new advances in technology and the incorporation of digital tools into work and school routines enabled participants to save time on commuting and increase productivity.

Fitness

Technology was reported to be especially helpful for monitoring and tracking fitness goals, particularly through devices such as Fitbit and Apple Watch:

> *I rely heavily on Fitbit to monitor and achieve physical fitness goals. This in turn has helped lower my blood pressure after these months of working out consistently.* [Korean participant, female, aged 66 y]

A Chinese participant (female, aged 60 y) shared that she was "doing virtual exercise classes daily," and another participant highlighted the role of activity tracking:

> *Apple Watch (and other activity trackers) have been helpful to keep me moving.* [Native Hawaiian participant, male, aged 25 y]

These quotes reflect the usefulness of technology for achieving consistent physical activity and monitoring daily progress for many participants. Thus, technology helped participants meet their fitness goals and was a tool to support positive health outcomes.

Technology Was Often the Only Means of Interpersonal Social Connections

Overview

Technology helped address participants' feelings of boredom and isolation by allowing access to streaming services, online shopping, and social media, as well as enabling them to feel connected to others. Staying connected through technology was also vital for work and health-related reasons, for example, when checking in on loved ones who were immunocompromised:

> *It's been very important in keeping me sane and connected to others.* [Vietnamese participant, female, aged 35 y]

This quote supports the finding that technology was essential in maintaining social interaction during the pandemic.

Connections for Social Interactions

During the COVID-19–related lockdown, many participants were unable to physically check in with friends and family and had to rely on alternative methods to maintain social connection. Technology use was reported to ease negative feelings by allowing quick and effective communication. Many participants with family members and friends residing out of the country used platforms such as Facebook Messenger, WeChat,

XSL•FO
RenderX

KakaoTalk, LINE, or WhatsApp for daily interactions and updates. Social media, Zoom, and Skype were also heavily used by participants to connect with loved ones and meet new people during this period of isolation. These tools were reported to be very useful for addressing communication challenges and easing worries of not being physically present:

> *Would have been much harder to cope without technology! Being able to video conference or text friends has helped maintain a sense of social connection while also being safe.* [Chinese participant, female, aged 51 y]

This quote further showcases technology's ability to bridge gaps in social connection, enabling participants to use tools to interact with members of their communities and support one another while maintaining safe distancing practices.

Technology was also reported to help participants meet new people and find online groups centered around similar interests and hobbies:

> *I have been involved in decisions about moving my community chorus to the virtual platform, a less than satisfactory replacement to singing together. But we have been able to reach out to singers from far and wide, something we would not have been able to do before!* [Chinese participant, female, aged 66 y]

This quote illustrates how technology enabled participants to meet others with common interests and hobbies, further expanding their communities and social interactions. Another participant shared a similar experience:

> *It's been an eye-opener to realize the capacity of the internet to connect people from all over the world through events, etc. I volunteer with our Girl Scout council and create virtual events for girls, who have attended from across the country!* [Japanese participant, female, aged 67 y]

This quote reflects the theme of technology facilitating the formation of new online communities, allowing participants to build relationships across geographic boundaries through common interests.

A participant expressed the importance of social connection with other Asian American and Pacific Islander individuals in the face of the racism and xenophobia that some experienced:

> *We Asians face a lot of discrimination. We can only communicate with friends online because it's risky to go out, and technology supports this.* [Chinese participant, female, aged 36 y]

This quote reflects participant concerns regarding the hate directed at Asian American and Pacific Islander communities during the lockdown and illustrates how technology enabled safe social interaction and communication.

Connections to Help Others

Technology was also reported to assist in the care of others, such as ordering groceries and essential supplies through online platforms:

> *It was helpful that I could order groceries for myself and my elderly parents on Instacart and also I am able to Zoom with them a few times a week so that we can "see" each other make sure everything is ok.* [Asian Indian participant, female, aged 53 y]

The availability of no-contact services provided a practical alternative to in-person errands, helping participants reduce their risk of exposure to COVID-19 infection while continuing to care for loved ones.

Connections for Work

Participants reported benefits associated with the shift to remote work. One noted advantage was the ability to expand business contacts across geographic regions and use platforms such as Zoom to widen professional networks and communicate with people in different time zones.

Connections for Health Reasons

Participants reported that technology use also facilitated access to wellness resources and communication with health care professionals. This included using technology to connect with meditation groups, community meetings, workshops, online physical activity classes, and telehealth services. Technology also enabled video calls with loved ones in hospital when in-person visits were not allowed:

> *Use of technology in the hospital provides an effective source of communication between the patient and his/her family members. I witnessed how helpful and meaningful to keep in touch with our patients' family by using smart phones or tablets, to see their loved one who is in the hospital in the screen really brought them a true happiness and brought positive effects to patients' wellbeing.* [Filipino participant, female, aged 52 y]

This comment illustrates how technology was used to care for others and protect individuals considered vulnerable by maintaining safe distancing. It also reflects concern regarding disconnecting from loved ones during the lockdown and how using technology allowed participants to continue to provide care and support to their communities.

Overuse Led to Negative Physical and Mental Health Outcomes

A common theme among participants was that increased technology use during the lockdown led to overuse, contributing to negative mental and physical health outcomes, including burnout. Participants described how increased reliance on technology use led to the feeling that they needed a break from technology and having no outlets away from technology use, feelings of increased reliance, and addiction to technology and social media.

Negative Mental Health

Participants also described needing mental health breaks due to feeling overwhelmed by constant COVID-19 updates from news outlets and social media. Some reported feeling addicted to social media and finding it difficult to step away due to loneliness or fear of "missing out," which resulted in increased anxiety and mental stress:

XSL•FO
RenderX

> *Technology, specifically social media has been incredibly stressful but addictive for me. I find myself much more anxious and unhappy when scrolling through other peoples profiles. But I feel this need to know otherwise I will miss out.* [Vietnamese participant, female, aged 24 y]

This quote reflects the broader theme of technology's impact on mental health during the pandemic, highlighting how constant access to social media and news updates made it difficult to disconnect for fear of missing essential information.

Technology also enabled participants to stay informed about the pandemic by providing access to news updates from outside of their immediate communities, including other cities, states, and countries. However, this constant exposure to negative news was also associated with increased anxiety and stress:

> *It's a double-edged sword—on one hand it has provided access to news and connection to family and friends, but it still feels isolating, lonely, and awkward to try and convey emotions and connections through the use of social media only.* [Asian Indian participant, female, aged 24 y]

This quote supports the finding that participants struggled to use technology effectively to maintain connection. Technology use helped participants stay connected with loved ones, but it also contributed at times to feelings of isolation and loneliness.

Discrimination Experienced by Asian American and Pacific Islander Individuals

Some participants reported negative experiences related to discrimination and racism during the pandemic, particularly on social media. Others expressed concerns about language barriers, xenophobia, and racism throughout this period:

> *I have been more sensitive to racially related news and posts on socially media.* [Chinese participant, female, aged 56 y]

> *[I] found that social media can be a bit more toxic during this period. Especially with all the negative news surrounding social justice, racism, the election, etc.* [Chinese participant, male, aged 28 y]

These quotes express the emotional and mental impact of frequent exposure to news and social media during a period marked by political unrest, social justice movements, and the pandemic, which resulted in participants feeling overwhelmed.

Burnout

Participants described feeling fatigued from constant technology use throughout the day but unable to disconnect due to the need to use technology for school and work:

> *I noticed that my technology usage has significantly increased during the COVID-19 pandemic. In doing so, I realized that I need to take screentime breaks so that I don't get exhausted from it.* [Filipino participant, female, aged 23 y]

Participants reported exhaustion from prolonged use of digital devices and limited opportunities to disconnect, given the lack of in-person alternatives for connection with others during the pandemic. This led to participants incorporating breaks from technology to maintain their mental health.

Some individuals reported being busier during the pandemic due to increased work hours and workload. A participant shared her experience as a frontline worker:

> *During the pandemic I have never been busier. Following my shifts at the hospital, I attend nightly community, task force, and organizing meetings learning about resources and responses to the COVID-19 pandemic that is disproportionately affecting our community of immigrant front-line workers with high levels of high-risk exposure. Technology has made it easier to connect local and national Filipino organizations, but Zoom fatigue coupled with clinical burnout is a real hazard.* [Filipino participant, female, aged 30 y]

This quote illustrates how frontline workers experienced long shifts and overwhelming workloads to protect their patients.

Participants also reported experiencing Zoom fatigue and feeling exhausted from daily online meetings. Some also reported struggling to separate work from home after transitioning to fully remote work, leaving little downtime to destress:

> *I was working from home 4 days per week before the pandemic, so I was just on my computer all day every day already. Now I'm on it more because I can't go out on weekends or in the evenings.* [Chinese participant, female, aged 35 y]

The combination of seemingly nonstop screen time from both work and personal activities contributed to mental and emotional exhaustion and resulted in feelings of burnout.

Negative Physical Health

Technology use during the pandemic was also found to be associated with negative physical health outcomes. Participants reported increased screen time and reduced physical activity, which contributed to symptoms such as fatigue, eyestrain, dry eyes, headaches, back and neck pain, wrist and hand pain, loss of sleep, dizziness, and weight gain:

> *Using electronic devices for long hours makes my shoulders, neck and back pains. However, I don't feel safe to see a doctor or physical therapist.* [Taiwanese participant, female, aged 37 y]

> *Too much screen time, more sitting for work related issues, common illnesses experienced, like headache and back pain.* [Asian Indian participant, female, aged 47 y]

These narratives reflect the physical toll of sustained technology use that compounded over time during the pandemic. Notably, the hesitation to seek care due to safety concerns highlights how the pandemic exacerbated both physical strain and barriers to accessing care for technology-related health issues.

Technology Use Was Associated With Multiple Challenges and Barriers

Many participants shared concerns regarding increased technology use, including language-related communication

XSL•FO
RenderX

challenges in virtual settings; increased costs; poor connectivity; and difficulty adapting to technology, particularly among older adults. These barriers may have been more pronounced for individuals who were underresourced or living in areas with limited internet access.

Challenges of Being a Nonnative English Speaker

Virtual work environments introduced new barriers for some Asian American and Pacific Islander individuals, as described by a participant:

> *As a nonnative English speaker/immigrant, I felt that I couldn't perform as well during a Zoom job interview as I became very self-conscious about my accent or mistakes that I make when I speak English. I think the Zoom interview environment has negatively impacted my chance of landing a job because it encourages interviewers to rely more on interviewee's speech as there are limited nonverbal cues that contributes to effective communication.* [Korean participant, male, aged 33 y]

This quote highlights the participant's perceived disadvantages when navigating a virtual job interview, with interviewers unable to meet interviewees and assess them as a whole. With first impressions limited to screen interactions, the complete dependence on technology was a barrier to connection and, in this participant's case, their success. Furthermore, relying primarily on speech disadvantaged the participant, a nonnative English speaker, heightening their fear of judgment and perceived risk of discrimination.

Increased Costs

Participants reported experiencing increased or additional expenses related to their increased technology use. These included upgrading internet service, paying higher data fees, and incurring increased electricity costs to support working or studying from home:

> *I had to upgrade my internet to be able to service people for work and during my internship for graduate school and it cost more to upgrade. Also, my electricity bill has gone up since I'm working and interning at home.* [Japanese participant, female, aged 41 y]

Increased costs posed additional challenges for individuals and families and was an added stressor, especially for those who were already facing financial strain.

Connectivity

Increased technology use at home also led to connectivity challenges related to school and work. Many participants reported that increasingly busy networks and limited bandwidth in certain areas led to decreased connectivity efficiency, which caused disruptions in work and school tasks. This often led to meeting disruptions, resulting in additional stress and loss of time to address connectivity issues:

> *With remote schooling and working remote, even in the suburbs, there are challenges with connectivity as everyone is online and using the internet for school and work. If the wireless networks and bandwidth are not strong, this often negatively impacts accessing internet for school and work.* [Filipino participant, female, aged 47 y]

This quote underscores that the usefulness of technology depends on reliable infrastructure. These challenges were especially pronounced for individuals in rural areas and for older adults who already faced barriers such as unfamiliarity with technology, financial constraints, and lack of support.

Technology Use Among Older Adults

Several participants reported difficulty navigating technology, with only 8.2% (91/1115) indicating that they felt slightly comfortable using technology and 18.4% (204/1115) reporting that they were not at all comfortable. The transition to increased technology use during the pandemic disproportionately affected older Asian American and Pacific Islander adults who struggled to adapt and often lacked the support necessary to learn how to operate technological tools. A participant expressed difficulty understanding how to use technology, stating that she would benefit from someone teaching her:

> *Please teach low tech adult/senior citizens like me how to use our iPhones, MacBook etc. We are dependent on our 7-12 year old grandchildren for tutorials (if we are lucky to have them nearby). Apple stores don't teach us anymore.* [Filipino participant, female, aged 77 y]

This quote reflects the generational gap in technology familiarity and the desire of older adults to improve their technology literacy and decrease their dependence on younger family members in this regard. It also highlights the frustration that some older adults felt due to limited opportunities for formal instruction and underscores the need for structured, age-inclusive technology education tailored to older adults.

Some participants expressed concern that their older adult loved ones did not know how to use technology to contact resources or reach out to family members for assistance during the pandemic:

> *I am grateful that I am comfortable and experienced in all aspects of technology. However, my mother (who does not live with me) is not, and it always worries me that she wouldn't be able to contact me if she has a pressing need, or that I am not around to help her immediately troubleshoot technological difficulties.* [Taiwanese participant, female, aged 31 y]

This participant's sentiment underscores the stress that families experienced when older adults lacked the skills to use technology independently. Limited technology access can create safety and communication barriers, especially during emergencies and periods of physical separation.

Another participant similarly reported worries regarding an older loved one:

> *Using technology is difficult with my elderly mom as she does not understand how to use it. As she is in an assisted living home, I cannot show her and the staff*

XSL•FO
RenderX

> *has limited time to help. For the hearing impaired or for those who are not familiar with screen interactions, technology is limited in its usefulness for meeting social needs.* [Chinese participant, female, aged 50 y]

This quote encapsulates the compounded challenges that older adults face when navigating technology, even in assisted living settings, and highlights how barriers such as impairments or unfamiliarity with digital interfaces can significantly hinder social connectivity and access to essential support.

Discussion

Principal Findings

This study describes the use of technology and its impact on the well-being and experiences of Asian American and Pacific Islander adults during the COVID-19 pandemic. While other studies have investigated the broad implications of technology on mental and physical health, the use of technology and its application to the culturally specific experiences of Asian American and Pacific Islander adults during the pandemic was largely unknown. The quantitative survey data revealed significant changes in lifestyle and communication behaviors that occurred during the pandemic. Nearly half of the respondents (528/1115, 47.5%) reported increasing their technology use by ≥3 hours per day during the pandemic, with up to 7.98% (89/1115) indicating an increase of ≥7 hours per day compared to before the pandemic. The 4 themes identified in this study highlight the mixed sentiment toward technology and social media use among Asian American and Pacific Islander adults, alongside notable challenges and barriers.

In the periods of social isolation during the pandemic, technology was a central force in facilitating individual well-being and fostering connections within the community [16]. Virtual platforms enhanced both the professional and hobby-related skills of participants, while connecting individuals with common interests. In addition to supporting mental and emotional health, technology benefited the physical health of participants through the use of fitness trackers and virtual interactions with health care professionals. Moreover, for the Asian American and Pacific Islander community, technology assumed a special significance by providing a sense of security amid rising anti-Asian discrimination. Virtual platforms and networks helped individuals stay informed about relevant social issues and incidents affecting their communities. Fearing for their safety, Asian American and Pacific Islander adults relied on technology, such as video chats and telephone calls, to maintain social connections without leaving home. Older adults were also better supported by family members who used mobile delivery services to order food and other essentials.

However, alongside the benefits of using technology, the detrimental effects of overuse on both mental and physical health were also revealed by this study. Participants reported concerns such as addiction to social media, burnout, and videoconferencing fatigue, which exacerbated feelings of loneliness and mental strain. For Asian American and Pacific Islander adults in particular, the increased exposure to negative news and discrimination on social media contributed to heightened awareness of discrimination, which in turn led to psychological distress. In addition to these negative impacts on mental health, technology overuse was associated with physical symptoms such as fatigue, eyestrain, back pain, headaches, and weight gain.

Our analysis of technology-related barriers for Asian American and Pacific Islander adults underscored how technology may exacerbate existing disparities within the community. Notably, technology may amplify inequities in employment opportunities, particularly evident in the challenges faced by nonnative English speakers during job interviews conducted over platforms such as Zoom. The shift to virtual interviews accentuated concerns about linguistic proficiency. Asian American and Pacific Islander participants expressed concern that interviewers may prioritize verbal communication over nonverbal cues, placing nonnative English speakers at a disadvantage. Moreover, increased reliance on technology also imposed financial burdens, potentially compounding economic disparities within the community. In addition, the growing global reliance on technology presented a significant barrier for older adults in the Asian American and Pacific Islander community who struggle to learn new technology. This challenge had a ripple effect on family members of these older adults, as they faced increased distress fearing an inability to contact older adult members in emergent situations or in times of social isolation.

While the themes of our study reveal a spectrum of positive and negative impacts of technology on Asian American and Pacific Islander adults, such mixed results regarding technology use are congruent with current literature sampling the general population. Our finding that technology was critical in maintaining individual mental and emotional well-being during the pandemic is supported by both studies concluding that the use of technology and digital tools was beneficial for young adults during the COVID-19 pandemic [17] and studies involving older adults [18,19]. In alignment with our findings, these studies reported reduced feelings of loneliness, better self-rated health, fewer chronic illnesses, higher subjective well-being, and fewer depressive symptoms. The use of technology as a facilitator for the feeling of community connection has also been well reported. A qualitative study involving older adults and their relatives highlighted how the use of technology-mediated communication by older adults during the pandemic afforded meaningful interconnections in their social lives and promoted cross-generational connections among family members [20].

Our finding that technology overuse can lead to detrimental effects on the mental and physical health of Asian American and Pacific Islander adults is also congruent with prior literature. Prior research has shown that moderate or severe depression levels were associated with greater time spent watching television and using computers (>6 h/d) for American adults [21]. In addition, our parent study of Asian American and Pacific Islander individuals found that those whose technology use during the pandemic increased by ≥7 hours per day had a 1.32-point (95% CI 1.00-1.64; *P* value for the trend <.001) higher Patient Health Questionnaire-4 score [9]. Negative impacts on physical health in the Asian American and Pacific Islander population are also congruent with our previous study

XSL•FO
RenderX

results, which detected significant changes in self-reported physical health scores with ≥3 hours per day of technology use [9], as well as independent studies reporting that technology use negatively affects physical activity and sleep routines, leading to headaches, neck pain, myopia, digital eye syndrome, and cardiovascular risk factors [22,23].

Finally, and perhaps most notably, our fourth theme, which discusses the multiple challenges and barriers to technology use for Asian American and Pacific Islander adults, contributes to the highly necessary investigation of racial disparities in an increasingly digitized world. Nonnative English speakers in this study described feeling disadvantaged in virtual workspaces, an experience recognized by a growing number of studies. In a case study of a multinational corporation after a shift to remote work triggered by the COVID-19 pandemic, researchers found that language-based discrimination in virtual spaces was primarily organizational and more subtle than in physical settings [3]. Virtual spaces also seemed to expedite the process of excluding migrant professionals, in part due to the emergence of separate meetings and parallel virtual channels for informal conversation [3]. Other barriers observed in our study, such as increased costs, connectivity challenges, and the lack of technological support for older adults, add to the growing body of evidence that technology may be exacerbating, rather than eliminating, socioeconomic and health disparities through the digital divide. This divide has significant clinical implications for the Asian American and Pacific Islander community.

This study illustrates ways in which Asian American and Pacific Islander adults found technology to be beneficial for mental and physical health. These results contradict our prior study results, which indicated that an increase in technology use during the pandemic correlated with negative mental health outcomes in the form of higher Patient Health Questionnaire-4 scores and lower physical health scores [6]. This discrepancy may be due to the variance in the framing of questions between the two studies. Our prior study (COMPASS) analyzed numeric responses from a structured set of questions, while this study analyzed responses to the open-ended question “Is there anything else you want to tell us about your use of technology during COVID-19?” This may have permitted respondents to consider alternative, positive applications of technology, such as its use as a tool to safeguard against an anti–Asian American and Pacific Islander social environment.

As revealed by our findings, while the COVID-19 pandemic accelerated the use of technology and telehealth, it also exacerbated existing challenges and barriers to accessing and using needed health care, evident in the growing digital divide and lack of digital literacy, as discussed in prior literature [24-26]. Asian American and Pacific Islander adults, particularly those with limited English proficiency, older adults, individuals with low income, and those living in rural areas or neighborhoods experiencing internet connectivity issues or lacking necessary digital literacy skills to effectively use telehealth platforms, face more challenges in adopting and using telehealth [27,28]. This leads to increased health disparities due to inequitable access to necessary health care. The barriers observed in our study, such as poor digital literacy, internet access issues, and lack of technology assistance, have been proven to be key drivers of the digital divide [29]. As such, identifying and resolving these barriers within the Asian American and Pacific Islander population is vital for the advancement of equitable health care and for appropriately addressing technology use as a social determinant of health. While many members of Asian American and Pacific Islander communities have become significantly more aware of the need to use such technologies, it is both crucial and timely to provide culturally and linguistically appropriate services, as well as training and resources to improve digital literacy, to address the identified disparities.

Consistent with previous research [30], our findings have revealed that dependence on technology, such as increased screen time, was associated with negative mental health. Conversely, avoidance of technology due to anxiety or lack of familiarity can prevent individuals from accessing telehealth services, potentially leading to delayed treatment or untreated health issues. Clinicians should be aware of these potential issues and integrate assessments of technology use and its impact on mental health into their care plans. Using a patient-centered approach to provide the needed support and resources to manage technology use in a helpful manner and offering alternative care options for patients who prefer not to use technology are important strategies in addressing these concerns.

The digital divide—highlighted during the COVID-19 pandemic as a disparity between the “haves” and “have-nots”—has been well documented [31]; however, more research is required to better understand the needs of diverse populations, not only by race and ethnicity, but also by English proficiency and age. This study contributes important knowledge in this area and may be helpful for developing social and educational programs to serve such diverse populations (eg, Asian American and Pacific Islander adults with low socioeconomic status and limited English proficiency).

Limitations

This study has several limitations. We used convenience sampling through community networks, social media, and the CARE registry, which may not fully capture the diversity of the broader Asian American and Pacific Islander population. There may be selection bias, with underrepresentation of individuals with limited internet access or those not engaged with community organizations. In addition, only 20.66% (1115/5398) of the respondents to the quantitative survey provided open-ended responses, which formed the basis for the qualitative themes presented in this study. In terms of generalizability, although the age range of participants was 18 to 98 years, the mean age was 45.9 (SD 16.3) years, which may reflect a generally more technology-savvy group. Furthermore, the majority of the respondents (665/1115, 59.64%) completed the open-ended question in English, likely reflecting a sample with a higher level of acculturation to the United States. Finally, we reported aggregated data across numerous Asian American and Pacific Islander groups, which may not represent the unique experiences of specific backgrounds within these communities.

XSL•FO
RenderX

Conclusions

This study presented the diverse attitudes and experiences related to technology use among Asian American and Pacific Islander adults during the COVID-19 pandemic. There were numerous benefits of using technology during a period of mandated physical and social isolation; however, there were many adverse experiences, including the detrimental effects of technology overuse on both mental and physical health. Important considerations for Asian American and Pacific Islander adults include awareness of how technology may exacerbate existing disparities, particularly among those with limited English proficiency. Further qualitative research would be beneficial in amplifying the perspectives of Asian American and Pacific Islander adults to uncover concerns and address health disparities.

Acknowledgments

The authors are grateful to all COVID-19 Effects on the Mental and Physical Health of Asian Americans and Pacific Islanders Survey Study (COMPASS) participants who completed the study. The authors would also like to acknowledge their community and academic partners in supporting the outreach and data collection. This study was conducted in collaboration with multiple academic institutions and community partners that serve diverse Asian American and Pacific Islander populations: the University of California San Francisco, University of California Davis, International Children Assistance Network, National Asian Pacific Center on Aging, Philip Jaisohn Memorial Foundation, Chinese Community Center, Hmong Community Center of Minnesota, Greater Boston Chinese Golden Age Center, United Sikhs, Memory and Aging Center at the University of California San Francisco, East Palo Alto Tennis and Tutoring, Yu-Ai Kai, India Community Center, and American Samoa Community Cancer Coalition. The authors sincerely thank Ms Hannah Jang for her administrative support on this manuscript. This research project is supported by the National Institute on Aging of the National Institutes of Health (R24AG063718 and R01AG083926). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Technology survey questions.
[DOCX File , 18 KB - humanfactors_v12i1e64999_app1.docx]

References

1. Hwang TJ, Rabheru K, Peisah C, Reichman W, Ikeda M. Loneliness and social isolation during the COVID-19 pandemic. Int Psychogeriatr 2020 Oct 26;32(10):1217-1220 [FREE Full text] [doi: 10.1017/S1041610220000988] [Medline: 32450943]
2. Eruchalu CN, Pichardo MS, Bharadwaj M, Rodriguez CB, Rodriguez JA, Bergmark RW, et al. The expanding digital divide: digital health access inequities during the COVID-19 pandemic in New York City. J Urban Health 2021 Apr;98(2):183-186 [FREE Full text] [doi: 10.1007/s11524-020-00508-9] [Medline: 33471281]
3. Back H, Piekkari R. Language-based discrimination in multilingual organizations: a comparative study of migrant professionals' experiences across physical and virtual spaces. J World Bus 2024 Apr;59(3):101518 [FREE Full text] [doi: 10.1016/j.jwb.2024.101518]
4. Lai J, Widmar NO. Revisiting the digital divide in the COVID-19 era. Appl Econ Perspect Policy 2021 Mar 12;43(1):458-464 [FREE Full text] [doi: 10.1002/aepp.13104] [Medline: 33230409]
5. Findling MG, Blendon RJ, Benson J, Koh HK. COVID-19 has driven racism and violence against Asian Americans: perspectives from 12 national polls. Health Aff Forefront 2022 Apr 12;4:1-10 [FREE Full text] [doi: 10.1377/forefront.20220411.655787]
6. Park VT, Tsoh JY, Dougan M, Nam B, Tzuang M, Park LG, et al. Racial bias beliefs related to COVID-19 among Asian Americans, native Hawaiians, and pacific islanders: findings from the COVID-19 effects on the Mental and Physical Health of Asian Americans and Pacific Islanders Survey Study (COMPASS). J Med Internet Res 2022 Aug 09;24(8):e38443 [FREE Full text] [doi: 10.2196/38443] [Medline: 35658091]
7. Ta Park VM, Dougan MM, Meyer OL, Nam B, Tzuang M, Park LG, et al. Discrimination experiences during COVID-19 among a national, multi-lingual, community-based sample of Asian Americans and pacific islanders: COMPASS findings. Int J Environ Res Public Health 2022 Jan 14;19(2):924 [FREE Full text] [doi: 10.3390/ijerph19020924] [Medline: 35055744]
8. Mcclain C, Vogels EA, Perrin A, Sechopoulos S, Rainie L. The internet and the pandemic. Pew Research Center. 2021 Sep 01. URL: https://www.pewresearch.org/internet/2021/09/01/the-internet-and-the-pandemic/ [accessed 2024-05-19]

XSL•FO
RenderX

9. Park LG, Meyer OL, Dougan MM, Golden B, Ta K, Nam B, et al. Social support and technology use and their association with mental and physical health during the COVID-19 pandemic among Asian Americans: the COMPASS cross-sectional study. JMIR Public Health Surveill 2023 Jan 23;9:e35748 [FREE Full text] [doi: 10.2196/35748] [Medline: 36395324]
10. Brief report. COMPASS. URL: https://compass.ucsf.edu/ [accessed 2025-05-29]
11. Ta Park VM, Meyer OL, Tsoh JY, Kanaya AM, Tzuang M, Nam B, et al. The collaborative approach for Asian Americans and pacific islanders research and education (CARE): a recruitment registry for Alzheimer's disease and related dementias, aging, and caregiver-related research. Alzheimers Dement 2023 Feb;19(2):433-443 [FREE Full text] [doi: 10.1002/alz.12667] [Medline: 35420258]
12. Collaborative approach for AAPI research and education (CARE). University of California San Francisco. URL: https://careregistry.ucsf.edu/home [accessed 2024-06-04]
13. Process of translation and adaptation of instruments. World Health Organization. 2018. URL: https://terrance.who.int/mediacentre/data/WHODAS/Guidelines/WHODAS%202.0%20Translation%20guidelines.pdf [accessed 2025-05-29]
14. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
15. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
16. Alghamdi NS, Alghamdi SM. The role of digital technology in curbing COVID-19. Int J Environ Res Public Health 2022 Jul 07;19(14):8287 [FREE Full text] [doi: 10.3390/ijerph19148287] [Medline: 35886139]
17. Pretorius C, Coyle D. Young people's use of digital tools to support their mental health during COVID-19 restrictions. Front Digit Health 2021;3:763876 [FREE Full text] [doi: 10.3389/fdgth.2021.763876] [Medline: 34927133]
18. Chopik WJ. The benefits of social technology use among older adults are mediated by reduced loneliness. Cyberpsychol Behav Soc Netw 2016 Sep;19(9):551-556 [FREE Full text] [doi: 10.1089/cyber.2016.0151] [Medline: 27541746]
19. Cotten SR, Ford G, Ford S, Hale TM. Internet use and depression among retired older adults in the United States: a longitudinal analysis. J Gerontol B Psychol Sci Soc Sci 2014 Sep;69(5):763-771. [doi: 10.1093/geronb/gbu018] [Medline: 24671896]
20. Badawy A, Solberg M, Obstfelder AU, Alnes RE. Together, at a distance: experiences with a novel technology for social contact among older people and their relatives in Norway during the COVID-19 pandemic. BMC Geriatr 2023 Apr 06;23(1):218 [FREE Full text] [doi: 10.1186/s12877-023-03869-3] [Medline: 37024822]
21. Madhav KC, Sherchand SP, Sherchan S. Association between screen time and depression among US adults. Prev Med Rep 2017 Dec;8:67-71 [FREE Full text] [doi: 10.1016/j.pmedr.2017.08.005] [Medline: 28879072]
22. Hilty DM, Armstrong CM, Smout SA, Crawford A, Maheu MM, Drude KP, et al. Findings and guidelines on provider technology, fatigue, and well-being: scoping review. J Med Internet Res 2022 May 25;24(5):e34451 [FREE Full text] [doi: 10.2196/34451] [Medline: 35612880]
23. Pandya A, Lodha P. Social connectedness, excessive screen time during COVID-19 and mental health: a review of current evidence. Front Hum Dyn 2021 Jul 22;3(50):60002-60017. [doi: 10.3389/fhumd.2021.684137]
24. Martins Van Jaarsveld G. The effects of COVID-19 among the elderly population: a case for closing the digital divide. Front Psychiatry 2020 Nov 12;11:577427 [FREE Full text] [doi: 10.3389/fpsyt.2020.577427] [Medline: 33304283]
25. Medero K, Merrill Jr K, Ross MQ. Modeling access across the digital divide for intersectional groups seeking web-based health information: national survey. J Med Internet Res 2022 Mar 15;24(3):e32678 [FREE Full text] [doi: 10.2196/32678] [Medline: 35289761]
26. Tai DB, Shah A, Doubeni CA, Sia IG, Wieland ML. The disproportionate impact of COVID-19 on racial and ethnic minorities in the United States. Clin Infect Dis 2021 Feb 16;72(4):703-706 [FREE Full text] [doi: 10.1093/cid/ciaa815] [Medline: 32562416]
27. Hsueh L, Huang J, Millman AK, Gopalan A, Parikh RK, Teran S, et al. Disparities in use of video telemedicine among patients with limited English proficiency during the COVID-19 pandemic. JAMA Netw Open 2021 Nov 01;4(11):e2133129 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.33129] [Medline: 34735015]
28. Rodriguez JA, Saadi A, Schwamm LH, Bates DW, Samal L. Disparities in telehealth use among California patients with limited English proficiency. Health Aff (Millwood) 2021 Mar 01;40(3):487-495. [doi: 10.1377/hlthaff.2020.00823] [Medline: 33646862]
29. Saeed SA, Masters RM. Disparities in health care and the digital divide. Curr Psychiatry Rep 2021 Jul 23;23(9):61 [FREE Full text] [doi: 10.1007/s11920-021-01274-4] [Medline: 34297202]
30. Zhang Y, Li Q, Hu W, Zhan N, Zou J, Wang J, et al. The relationships between screen time and mental health problems among Chinese adults. J Psychiatr Res 2022 Feb;146:279-285. [doi: 10.1016/j.jpsychires.2021.11.017] [Medline: 34785040]
31. Ramsetty A, Adams C. Impact of the digital divide in the age of COVID-19. J Am Med Inform Assoc 2020 Jul 01;27(7):1147-1148 [FREE Full text] [doi: 10.1093/jamia/ocaa078] [Medline: 32343813]

XSL•FO
RenderX

Abbreviations

CARE: Collaborative Approach for Asian Americans and Pacific Islanders Research and Education
COMPASS: COVID-19 Effects on the Mental and Physical Health of Asian Americans and Pacific Islanders Survey Study
UCSF: University of California San Francisco

Edited by A Kushniruk; submitted 01.08.24; peer-reviewed by K Wang, H Nguyen, A Valdez; comments to author 11.03.25; revised version received 24.05.25; accepted 10.07.25; published 10.09.25.

Please cite as:
Park LG, Chi S, Lay M, Phan N, Tsoh JY, Meyer OL, Nam B, Park VT
Diverse Attitudes and Experiences With Technology Use During the COVID-19 Pandemic Among Asian American and Pacific Islander Adults (the COMPASS Study): Survey Study
JMIR Hum Factors 2025;12:e64999
URL: https://humanfactors.jmir.org/2025/1/e64999
doi:10.2196/64999
PMID:40928837

©Linda G Park, Serena Chi, Myka Lay, Nicole Phan, Janice Y Tsoh, Oanh L Meyer, Bora Nam, Van Ta Park. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 10.09.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Barriers and Enablers to Young People's Posting, Responding, and Reading Behaviors on Mental Health Forums Using the Behavior Change Wheel: Qualitative Study

Zhen Lim[1], MSc; Si Yin Lim[1], MSc; Siqi Lu[1], MSc; Leslie Morrison Gutman[1], PhD

University College London, London, United Kingdom

Corresponding Author:
Leslie Morrison Gutman, PhD
University College London
1-19 Torrington Place
London, WC1E 7HB
United Kingdom
Phone: 44 020 7679 1897
Email: l.gutman@ucl.ac.uk

Abstract

Background: Web-based mental health forums have the potential to play a significant role in providing accessible support for young people, supplementing in-person services and contributing positively to their mental well-being. However, limited engagement often constrains their impact and effectiveness in supporting young people.

Objective: Using the Behavior Change Wheel framework, this qualitative study explores the barriers and facilitators to young people's engagement with web-based mental health forums, focusing on the behaviors of creating new posts, responding to posts, and reading posts. Behavior change techniques (BCTs) are identified to address these barriers.

Methods: Semistructured interviews were conducted with 13 young people aged 17-25 years who use UK-based youth mental health forums. Three participants self-identified as men, 8 as women, 1 as nonbinary, and 1 chose not to disclose their gender. Transcripts were coded using the Theoretical Domains Framework (TDF), followed by inductive theme generation. TDF barriers were then mapped to BCTs to suggest intervention strategies.

Results: Thematic analysis revealed ten inductive themes across 5 TDF domains. Of these, 3 were enablers, 2 were barriers, and 5 functioned as both enablers and barriers. The findings indicated that skills, beliefs about consequences, emotions, and the social and physical environment are key influences on young people's engagement with web-based mental health forums. Positive emotions experienced after using the forums enabled posting, responding, and reading behaviors. Enablers of more active participation (ie, posting and responding) included anonymity and positive interactions with other users. The presence of moderators acted as an enabler for all 3 behaviors by providing a safe environment, but also as a barrier to posting, as moderation could restrict the content of users' posts. Similarly, mobile access facilitated posting, responding, and reading, whereas layouts not optimized for mobile use acted as barriers to typing and reading on the go.

Conclusions: This study contributes to the existing knowledge base by examining the different ways in which young people engage with youth mental health forums. Different strategies may be prioritized and adopted depending on whether forum providers aim to increase more active forms of engagement (eg, posting and responding, which can be encouraged by fostering positive interactions with other users) or overall engagement (eg, establishing clear rules of engagement and optimizing web page content for mobile access can benefit all forms of engagement). These insights can help improve the delivery of youth mental health forums and foster a positive ecosystem of support for young people.

(JMIR Hum Factors 2025;12:e71549) doi:10.2196/71549

KEYWORDS

internet; online; adolescents; discussion board; evaluation; qualitative; survey

XSL•FO
RenderX

Introduction

Young People's Mental Health and the Role of Web-Based Forums

Globally, young people are experiencing a significant decline in mental health, with rising levels of reported mental health problems [1,2]. The COVID-19 pandemic, in particular, accelerated the deterioration of young people's mental health [3]. In the United Kingdom, 1 in 3 young people aged 18-24 report symptoms of anxiety, depression, or both—an increase since 2000, when fewer than 1 in 4 (24%) reported these problems [4]. However, a recent report identified significant gaps in early mental health support for young people, with a rising number reaching crisis point [5]. From 2022 to 2023, the Children's Commissioner's annual report on children's mental health [6] found that 28% of children and young people referred to mental health services were still waiting for support, with some waiting up to 2 years.

Viable stopgaps for young people include web-based channels such as social media or mental health forums, which are typically available 24/7 and can potentially provide accessible mental health support at scale. As digital natives, young people are adept at using the internet to seek support and are often favorable toward this approach, as it offers greater convenience and anonymity [7]. While social media may have a wider reach [8], mental health forums can be seen as a more focused and structured channel for young people to obtain mental health support [9]. Engaging in youth mental health forums can offer benefits such as emotional support, alleviation of feelings of isolation and stress, and reductions in symptoms of depression and self-harm [10,11]. However, potential drawbacks include increases in mental distress and self-harm tendencies, particularly when safeguards are insufficient in forums addressing sensitive or distressing topics [9]. Thus, it is important to continually explore ways to increase the effectiveness of these forums while reducing their potential harmful impacts. In addition, young people's low engagement in web-based mental health forums remains a challenge, potentially limiting their effectiveness [12]. Research shows that mental health forums often fail to achieve their intended purpose because of low engagement rates [13], and users are unlikely to continue participating in inactive forums due to the limited support available [14]. Sustaining user activity in web-based forums helps to build a pool of knowledge and ensure that the forum remains relevant, facilitates the creation of online connections, and provides necessary support [14,15]. Therefore, there is a need to better understand the factors influencing engagement to maximize potential benefits.

Research shows that young people engage in web-based mental health forums in various ways. Many studies dichotomize them into *posters*, who are active participants contributing most of the content, and *lurkers*, who are passive participants and read posts without contributing [16,17]. Posters may create initial posts that introduce topics for discussion, often focusing on their own concerns, questions, thoughts, and lived experiences [18]. They may also respond to others' posts, providing emotional or informational support to the user who shared the original comment [19]. Lurkers, on the other hand, read posts either to find specific information or to browse more generally [20]. Young people's engagement in web-based mental health forums therefore encompasses multiple behaviors, including posting original content, responding to others, and reading posts.

While these behaviors should be treated as distinct, as they have different motivating and hindering factors [21], existing studies often examine youth mental health forum use more generally [22,23] or focus only on posters and lurkers [15,16]. However, this limits our understanding of the factors influencing young people's engagement across a range of different behaviors. Using the Behavior Change Wheel (BCW) approach [21], this qualitative study explores the barriers and enablers to young people's engagement in mental health forums, focusing on posting, responding, and reading behaviors, and identifies strategies for improvement. Such an examination can reveal both unique and shared influences and support the development of targeted interventions, offering a holistic view of young people's engagement with youth mental health forums.

Research Background

Current research has highlighted the benefits of mental health forums for young people. Web-based forums foster a sense of community and friendship, providing spaces where young people can receive practical advice on a variety of topics [18]. They not only offer informal support but also signpost young people to formal support services when necessary [24,25]. Thus, they can serve as important conduits, facilitating access to relevant professional services and encouraging help-seeking behavior among young people who need it. Furthermore, unlike professional or in-person services, web-based forums connect young people around the clock and therefore serve as valuable supplements to other mental health services within the broader ecosystem [19].

The characteristics of web-based mental health forums also enable different types of engagement among young people. As these forums contain a wealth of information and messages that are typically archived for long periods, young people can access relevant content at any time [9]. Reading the repository of existing discussion threads therefore provides immediate, personally relevant, and actionable information [14], such as personal experiences on managing specific situations and techniques for improving one's own mental health [26,27]. Reading posts also allows young people to reflect on their personal difficulties through the lens of others' experiences, for instance, by reframing common challenges as manageable [28].

In contrast to passively reading content, young people who actively use web-based forums—by initiating and replying to posts—can benefit from the therapeutic nature of writing and from providing support to other users in need [28]. The veil of anonymity offered by most web-based forums reduces users' feelings of threat and inhibition when posting or responding [29,30], creating vast potential to connect young people who share similar life experiences and emotions but may not be in close geographical or physical proximity [9]. Consequently, young people may feel more motivated to post on web-based forums because they can voice their opinions with greater confidence and less fear of judgment [14]. Posting also

XSL•FO
RenderX

contributes to the exchange of emotional and informational resources within the online space and encourages further engagement from users [31]. "Reply reciprocity," whereby users tend to receive more replies to their posts if they respond to others, further underscores the importance of encouraging participation, as this positive feedback loop promotes sustained user activity [31].

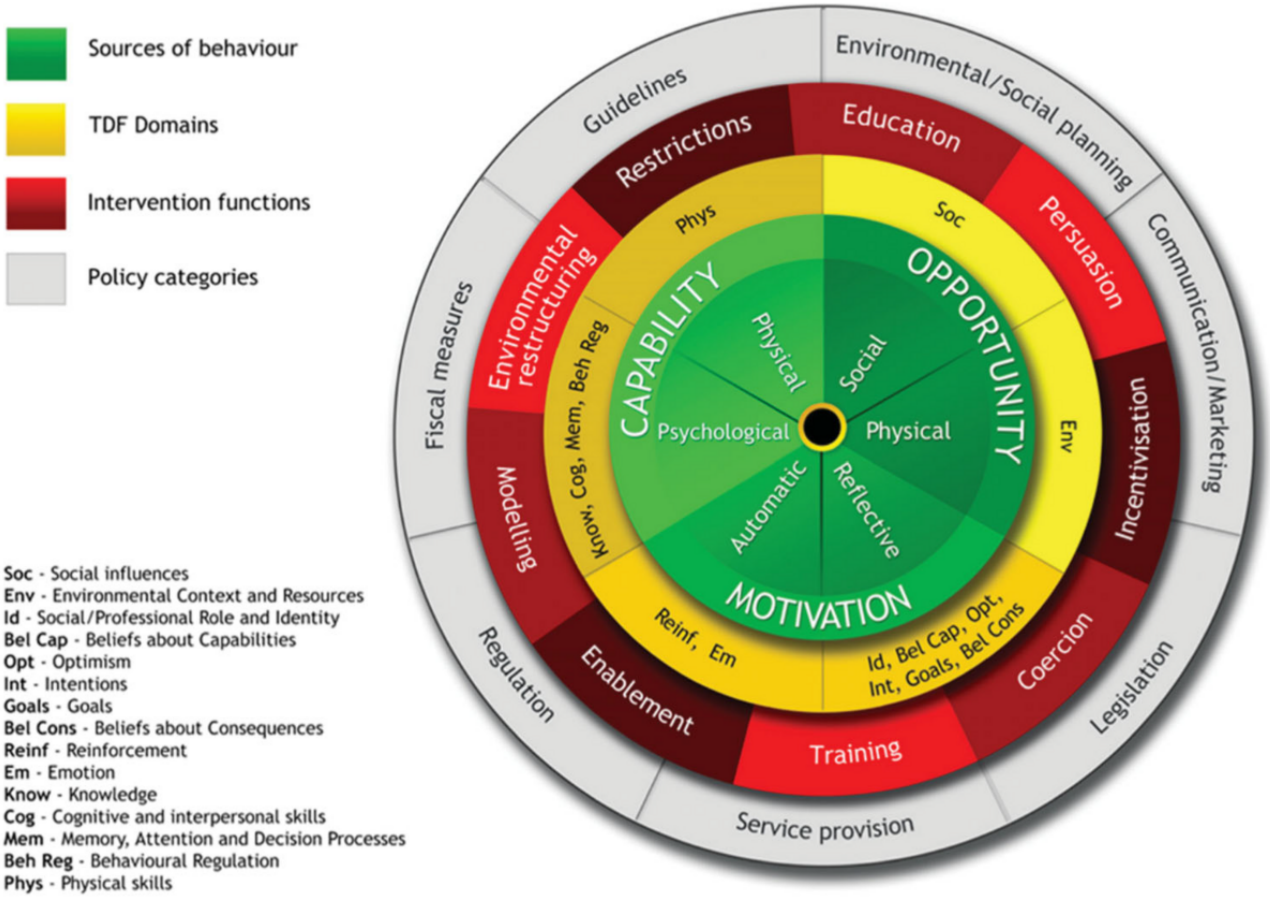
Several barriers to the use of web-based mental health forums have also been identified. Concerns about confidentiality and privacy can discourage posting on these forums, as can difficulties in accurately conveying emotions through writing [22]. In addition, reading about the challenges faced by other users may be triggering and induce negative moods [32]. Young people are also deterred from reading posts when forums lack high-quality resources or relevant topics [18,22]. Fears of causing harm to oneself (eg, worsening one's mental health condition) or to others (eg, making others feel worse) have likewise been identified as barriers to active participation [23]. Furthermore, some young people cite the lack of orientation for new users—making it harder to fit into a new online environment—as an additional barrier to active participation in mental health forums [32].

While existing studies have examined factors influencing engagement and proposed ways to enhance young people's experiences with mental health forums, such efforts are rarely guided by a comprehensive intervention framework like the BCW. Moreover, most studies on young people's use of web-based mental health forums take a one-dimensional view of engagement, without exploring the specific behaviors involved in interacting with these platforms. This limits researchers' and practitioners' ability to identify the distinct factors influencing different engagement behaviors and to develop effective, targeted solutions.

Behavior Change Wheel Framework

The BCW (Figure 1) is a comprehensive framework developed through the integration of 19 existing behavior change theories [21]. It offers a structured, evidence-based approach for diagnosing the barriers and enablers of behavior, which in turn informs the design and evaluation of interventions [33]. Central to the BCW is the COM-B model, which posits that capability (physical and psychological), opportunity (social and physical), and motivation (reflective and automatic) interact to produce behavior. The Theoretical Domains Framework (TDF) is often used alongside the COM-B model to provide greater granularity [34] and a more detailed understanding of different types of behavioral influences.

Figure 1. The Behavior Change Wheel and Theoretical Domains Framework (TDF).



Nine intervention types that enable behavior change form the layer surrounding the COM-B model, followed by an outer layer outlining 7 policy options for delivering these interventions [35]. The Theory and Techniques Tool [36] facilitates the mapping of TDF domains to behavior change techniques (BCTs) within the behavior change techniques taxonomy (BCTTv1). The taxonomy contains 93 evidence-based BCTs that are observable and replicable, facilitating behavior change [37].

XSL•FO
RenderX

This enables identified barriers to target behaviors to be linked to specific BCTs to achieve the desired change. The BCW approach has been successfully applied to enhance web-based mental health services for young people, including webchat counseling [38,39], helpline services [40], peer support groups [41], and self-harm content on a mental health forum [42].

Current Study

This study applies the BCW framework [20] to systematically explore the perceived barriers and facilitators of engagement in UK-based web-based youth mental health forums among young people aged 16-25 years. Through qualitative interviews, this study addresses the following research questions:

- Using the COM-B/TDF, what are the barriers and enablers to young people's engagement with youth mental health forums, specifically in relation to posting, responding, and reading?
- Using the BCTTv1, which BCTs can be applied to enhance young people's engagement with youth mental health forums, focusing on posting, responding, and reading?

This study can inform strategies to optimize the design and delivery of youth mental health forums, helping to better harness the potential of these web-based platforms in supporting young people's mental health and well-being.

Methods

Participants

The study targeted young people aged 16-25 years who had read, posted, or responded to posts in UK-based youth mental health forums. Participants were recruited via convenience sampling, with support from The Mix, powered by Mental Health Innovations—a UK-based charity that provides mental health support for young people, including crisis helplines, support groups, and web-based forums. The Mix helped publicize the study by sharing a recruitment poster and registration link via their web pages and social media channels. Additionally, the study was promoted on LinkedIn (LinkedIn Corporation) to broaden the pool of potential participants. Interested young people registered through an online form, which collected information such as age, gender, ethnicity, and frequency of youth mental health forum use.

Of the 32 young people who registered interest, 12 did not meet the eligibility criteria (eg, outside the target age range or not using youth mental health forums) and were excluded. Of the remaining 20 eligible participants, 6 did not provide consent after receiving the participant information sheet and consent form. Fourteen participants were interviewed; however, 1 interview was later excluded due to possible duplication.

The mean age of the 13 participants was 21 years, with a range of 17-25 years; 3 participants self-identified as male, 8 as female, 1 as nonbinary, and 1 did not wish to disclose their gender. Table 1 provides further details on participants' characteristics.

Table 1. Description of participants' characteristics.

Participant ID	Age (years)	Gender	Self-identified ethnicity	Experience with replying to posts on youth mental health forums	Mental health forums used
1	19	Woman	White British	Weekly	The Mix
2	25	Man	Indian	Weekly	The Mix
3	21	Woman	Iranian	Daily	The Mix
4	20	Man	Black African	Weekly	The Mix
5	23	Woman	Multiple ethnic groups	Never	The Mix
6	25	Woman	White British	Monthly or less	The Mix
7	18	Nonbinary	White British	Weekly	The Mix
8	23	Did not disclose	White British	Never	The Mix
9	23	Woman	Asian	Never	The Mix
10	23	Man	White British	Daily	The Mix and Kooth
11	24	Woman	Black African	Daily	Kooth
12	18	Woman	Black Caribbean	Monthly or less	The Mix
13	17	Woman	Multiple ethnic groups	Weekly	The Mix

Procedure

The researchers (ZL, SYL, and SL) conducted 1-on-1 semistructured interviews with participants from June to July 2024. These interviews were held online via Microsoft Teams. Participants could choose to respond using either audio conferencing or written chat, depending on their stated preference. All participants were asked the same set of questions based on an interview guide developed using COM-B and TDF domains, along with general questions that allowed for a broader

XSL•FO
RenderX

range of responses. The guide was developed by ZL, SYL, and SL in consultation with the senior author (LMG), and the questions were adapted from previous studies with a similar focus [38]. Sample questions were "What discourages you from responding to a post?" (general question), "What do you think are the benefits of using web-based mental health forums?" (reflective motivation: beliefs about consequences), and "Do others such as friends, family, or other users affect whether you use web-based mental health forums?" (social opportunity: social influences). The full interview schedule is provided in Multimedia Appendix 1.

Before recruitment began, the interview schedule was piloted with 4 young people from different demographic backgrounds (eg, gender, age range) to ensure that the questions were easily understood and that the interview could be completed within 30-45 minutes. Based on their feedback, the schedule was revised to use less technical language to enhance comprehension. The revised version also included additional prompts to encourage participants to provide more in-depth responses.

Publicity and participant recruitment began after the interview schedule was finalized. Interested individuals who met all eligibility criteria were provided with a consent form and a participant information sheet outlining the details of the study. After reviewing the information sheet and returning a signed consent form, participants were scheduled for 1-on-1 online interviews conducted by the research team. The interviews explored 3 types of behavioral engagement with web-based forums: reading forum content, creating new posts, and responding to existing posts. All interviews were recorded and transcribed. Participants received £25 (US $36) worth of vouchers as compensation for their participation.

Data Analysis

The behaviors of posting, responding to posts, and reading posts on youth mental health forums were coded separately by ZL, SYL, and SL. For each behavior, analysis followed a deductive approach guided by the Framework Analysis method, which included the stages of transcription, data familiarization, indexing, charting, and data interpretation [43]. Transcription and data familiarization involved repeatedly reviewing the interview content to develop a deeper understanding of the data. Indexing referred to applying an existing analytical framework to the dataset, while charting involved systematically organizing the data [44]. To support the indexing and charting process, a codebook was developed to define how each TDF domain applied to the specific engagement behaviors associated with youth mental health forum use (see Multimedia Appendix 2 for an example). The TDF served as a useful framework for coding and conducting a behavioral diagnosis to identify participants' perceived barriers and facilitators to forum use.

To enhance the reliability and consistency of coding, and to provide additional perspectives on the data, half of the total transcripts were second-coded by either ZL, SYL, or SL. Feedback from this process was used to refine the codebook applied throughout the deductive analysis phase. This was an iterative process in which the codebook was continually updated during transcript coding. The overall agreement rate exceeded 80% (122/146, 83.5%), and all coding discrepancies were discussed and resolved through mutual agreement.

After indexing and charting the data using the co-developed codebook, an inductive analysis approach was adopted to identify more specific and concrete themes. Candidate themes were examined to ensure internal consistency, and the data extracts within each theme were reviewed to confirm accurate representation, with further refinements made where necessary. Elements of Braun and Clarke's [45] reflexive thematic analysis were incorporated, emphasizing reflective and purposeful engagement with the data informed by the researchers' subjective experiences.

Following the deductive and inductive analyses, core inductive themes were shortlisted by ZL and LMG [46] based on their frequency and their importance and relevance to at least 2 of the 3 target behaviors. This criterion was used to identify key areas for improvement. Based on the identified TDF domains, associated BCTs with conclusive evidence-based links in the Theory and Techniques Tool were considered, while those with inconclusive or absent links were excluded. The final selection of BCTs and corresponding intervention strategies was informed by previous research on web-based forums and assessed for appropriateness and feasibility within this context by ZL, SYL, and SL. The selected strategies were further evaluated using the APEASE (Affordability, Practicability, Effectiveness, Acceptability, Side-effects, and Equity) criteria [21].

Ethical Information

This study was approved by the UCL Research Ethics Committee (reference number 16583/003; approval date: 27 July 2023). Informed consent was obtained, with researchers highlighting that participation was voluntary and that participants' sharing would be kept confidential.

Results

Overview

Ten core inductive themes were generated and mapped to the corresponding TDF domains (Table 2). These comprised 3 enablers, 2 barriers, and 5 that functioned as both enablers and barriers. Each theme is described in detail below, supported by participant quotes identified by their participant ID (Table 1). Themes specific to creating posts, responding to posts, and reading posts were also identified (see Multimedia Appendix 3).

XSL•FO
RenderX

Table 2. Description of core inductive themes.

Theme	Barrier/enabler	Description of theme	Relevant behaviors
Positive emotions	Enabler	Young people feel positive emotions (eg, relief, calmness) after using youth mental health forums.	Posting, responding, and reading
Beliefs about helpfulness of the reply	Barrier	Young people are discouraged from creating a new post or responding if they felt that the replies would not be helpful.	Posting and responding
Presence of moderators	Barrier/enabler	Presence of moderators on youth mental health forums makes the web space safer and reduces antisocial behavior, but may also limit the content of posts.	Posting, responding, and reading
Relevance of information	Barrier/enabler	Content that is not properly categorized or not personally relevant to young people discouraged reading and responding, but may encourage posting behavior.	Posting, responding, and reading
Accessing content on mobile phones	Barrier/enabler	The ability to access forum content on mobile phones facilitates usage, but when the layout is not optimized, it leads to difficulty in typing and reading on the go.	Posting, responding, and reading
Anonymity	Enabler	Anonymity reduces the sense of judgment from other users and encourages the sharing of thoughts.	Posting and responding
Time taken to obtain support	Barrier/enabler	Being able to quickly find the information needed can encourage young people to read posts, while the slow support and response prevent them from posting new content.	Posting and reading
Interactions with other users	Barrier/enabler	Positive experiences with other users facilitate posting and responding, while negative interactions discourage young people from doing so.	Posting and responding
Empathy	Enabler	Empathy helps young people relate to the original poster and provide more respectful and effective responses.	Responding and reading
Poor communication skills	Barrier	Poor communication skills hindered young people from creating and responding to posts in youth mental health forums. This barrier also had implications for reading, making it harder to understand posts.	Posting, responding, and reading

Automatic Motivation (TDF: Emotions)—Positive Emotions (Enabler)

Experiencing positive emotions was identified as an enabler for half of the young people using youth mental health forums across posting, responding, and reading behaviors. For those who actively engaged by posting or responding, these emotions often arose from the sense of connection with others. For example, P12 shared: “I feel relieved, I feel I'm able to interact (with) someone, I'm able to at least remedy a situation for someone, and I feel satisfied (with) myself.” P10 expressed similar sentiments, stating that they felt “a relief and calmness within that you've actually add(ed) value to someone’s life”. Even participants who only read forum posts reported experiencing positive emotions, often in the form of validation for their thoughts or behaviors. For example, P1 described: “it’s nice to read the posts and see that I’m not the only one feeling a certain way or experiencing a specific thing.”

Reflective Motivation (TDF: Beliefs About Consequences)—Beliefs About Helpfulness of the Reply (Barrier)

A barrier that discouraged a small number of young people from posting or responding was their belief about the helpfulness of replies. For example, P6 noted that responses are “not always helpful, depending on who responds.” P8 expressed a similar concern, highlighting that users may not “(get) the answers you wanted.” Specifically for the behavior of responding to posts, P1 reported sometimes withholding their response, explaining that they “might say the wrong thing to the person and end up making them feel worse or leave them feeling more confused.” Others expressed similar concerns, noting that they were cautious about responding to “avoid upsetting or triggering other users” [P1], or to ensure they did not “make things worse for anyone” [P6].

Physical Opportunity (TDF: Environmental Context and Resources)

Presence of Moderators (Barrier/Enabler)

For most young people, the presence of moderators in the forums could act as either a barrier or an enabler to posting, responding, and reading posts. Some participants noted that moderators helped make the online space safer, which in turn encouraged forum engagement. As P4 explained, “professional moderation on the online environment can ensure that the discussion boards remain (a) safe and respect(ful) space.” Participants also highlighted that moderators helped prevent exposure to content they preferred not to see, such as “seeing feeds that you think you don’t want to associate with or relate with, or even see” [P11], including “sexually explicit posts, political posts, terrorism, crimes” [P8]. Conversely, a small number of young

XSL•FO
RenderX

people felt that moderators could limit the content of posts and restrict self-expression. As P13 noted, "you're not able to adequately share what you're thinking about...the moderators put a lot of limitations that are making you not maybe express yourself the way you would have done if the moderator had no restrictions."

Relevance of Information (Barrier/Enabler)

A small group of young people noted that the relevance of information could act as either a barrier or an enabler for engagement, including posting, responding, and reading. For example, P6 explained, "there are a lot of subsections which...is needed but sometimes posts are in the wrong area or there are so many that things get lost" [P6]. Some young people also reported that the content they encountered on the forums was "not always relevant" and often focused on the "popular ones" [P3]. This perception that discussion board information might not be personally relevant or useful sometimes discouraged them from reading posts. As P3 explained, "whenever I read stuff about other people there, their situation isn't like mine." These participants therefore preferred creating posts on the forums rather than searching for information in existing posts, as posting could elicit responses that were more personally relevant and "felt more tailored to me and my current situation" [P6]. However, the inability to find relevant information also acted as a barrier, limiting young people's capacity to share knowledge and provide useful responses. P8 also noted that the presence of posts that were "repetitive or non-unique" discouraged her from responding and added that she "would respond to posts more (if) they were more varied."

Accessing Content on Mobile Phone (Barrier/Enabler)

For a small number of young people, accessing forum content on mobile phones could function as either an enabler or a barrier for posting, responding, and reading behaviors. One participant highlighted that the convenience of using her phone facilitated forum engagement: "the fact I can use it both on my phone and my laptop make me use it more because I can access it whenever I want" [P1]. However, others reported that the website's design and mobile experience were not optimized, which hindered their ability to use the forum on the go. As P5 elaborated, "If you're checking in via mobile phone, it can be a little awkward to navigate, some areas oversized or needing to be zoomed in upon, not always straightforward to type something in." Others shared similar challenges, with P9 stating that "using the website on cellphone is not convenient...It could be better if it launch apps," while P1 suggested that a mobile application could be helpful, allowing users to "access it easier" and "get notifications."

Anonymity (Enabler)

In the interviews, most young people highlighted anonymity as a key factor facilitating active engagement in forums, particularly for posting and responding. A secure, anonymous, and private environment was considered crucial for users to feel comfortable sharing their thoughts and concerns, thereby encouraging young people to create posts to seek help. For example, "I think the biggest contributor to me using it is the fact that it's anonymous so I don't have to worry about people I know seeing my posts" [P1] and "the sense of anonymity to most people so it's easier to open up about something and get the support" [P6]. When crafting responses, the sense of anonymity enabled young people to "freely talk about your situation or feelings without anyone judging you or questioning or invalidating you" [P1] and to "open up without the fear of people knowing who you are" [P11]. Anonymity facilitated engagement by allowing participants to share their thoughts with reduced concern about judgment from others.

Time Taken to Obtain Support (Barrier/Enabler)

For a small group of participants, the time required to obtain support could act as either an enabler or a barrier to posting and reading, depending on how quickly responses were received. Some participants preferred reading existing posts rather than creating new ones to seek information, as the immediacy of support available through reading was more appealing in the context of web-based forums. The wealth of information in the existing forum posts led some participants to view reading posts as a more desirable alternative to creating new posts. As P12 explained, "once I log in, I notice that most of the information that I needed to post, was already sometimes already there." This view was echoed by P8, who noted, "generally making posts means you can be waiting a couple of days for a response, whereas when you search and read information you get what you need immediately." The immediacy of support thus encouraged participants to read posts, while the potential for delayed responses discouraged them from posting.

Social Opportunity (TDF: Social Influences)—Interactions With Other Users (Barrier/Enabler)

For a small group of young people, interactions with other forum users specifically influenced active engagement (ie, posting and responding), acting as either an enabler or a barrier depending on the nature of these interactions. Participants who had positive experiences were more likely to increase their posting or responding behaviors, whereas those with negative experiences were less inclined to engage. For example, P6 described the role of reciprocity in deciding whether to reply, stating, "if someone has supported me before and they post I would be more likely to reply...It just feels like you are then helping someone who has helped you before." Conversely, negative experiences with other users discouraged young people from responding. As P11 explained, "usually if its someone that I know doesn't like me I will never reply to them...if there have been issues in the past with people it makes you want to avoid them." The impact of negative interactions extended beyond specific users to affect overall posting and responding behaviors. For example, "negative comments or replies by others" [P7] or instances where others "answer(ed) you rudely...really discourage(d)" [P13] deterred users from posting in general, whereas experiences of "peer support and empathy and understanding" [P4] encouraged engagement.

Psychological Capability (TDF: Skills)

Empathy (Enabler)

For most young people, empathy was identified as an important skill that helped them relate to the original poster when reading

XSL•FO
RenderX

posts and providing responses. Specifically, having empathy helped users "understand the situation...(and) how they're feeling" [P2], facilitating "emotional connection" and "driv(ing) you to offer support" [P4]. It also provided a foundation for young people to reply "respectfully and in an understanding way" [P6]. P11 elaborated that a lack of empathy could lead to negative outcomes: "they may tend to just reply without any emotions and for the person that has gone through this, it's hurtful and this discourages them from posting something that they feel like it's very personal or made them feel like that." P1 echoed these points, noting that empathy is necessary to "fully understand what that person is feeling before you begin to reply...otherwise the person may end up feeling more isolated or misunderstood."

Poor Communication Skills (Barrier)

Half of the participants reported that the poor communication skills of posters and responders discouraged them from engaging in posting and responding on youth mental health forums. For example, P3 stated, "I don't have great communication skills...that's why I don't always tend to write." P4 added that having good writing skills helped to "express your ideas and what you are going through." This barrier also affected the behavior of reading, representing a Physical Opportunity (TDF: Environmental Context and Resources) barrier, as young people were deterred from reading posts that were poorly communicated. This included posts that lacked clarity and coherence, such as those with incorrect grammar and punctuation, which made them more difficult for participants to read and understand. P13 explained, "if you're reading through a post that has grammatical mistakes...you can interpret it differently, but if it is clear, then you just get the information the way it is." In addition, some posts were described as "vague" [P9] or "confusing" [P10], with P9 adding that "some of the posts didn't make much sense...even after reading them."

BCTs

To address the identified barriers, the Theory and Techniques Tool was used to identify relevant BCTs. The most appropriate intervention components were then selected based on an APEASE assessment, as detailed in Table 3.

XSL•FO
RenderX

Table 3. Proposed intervention components and APEASE[a] evaluation.

Barrier (TDF[b] domain)	Suggested BCT[c]	Description	APEASE evaluation
Presence of moderators (Environmental Context and Resources)	• Adding objects to the environment • Restructuring the physical environment	• Clearly state rules and guidelines about content that is not permitted, and provide different interaction mechanisms for users	• Creating clear guidelines is an affordable, acceptable, and practicable way of setting the boundaries of what can be discussed in web-based forums. Having different interaction mechanisms can also be effective in providing the required social support for young people with different support needs [47].
Relevance of information (Environmental Context and Resources)	• Restructuring the physical environment	• Implement recommender engines (ie, tools which use machine learning to recommend relevant items to users) to improve the recommendation of relevant content	• While it may not be the most affordable option, it is the most effective, practicable, and acceptable in addressing stated barrier [48]. Thus, organizations are suggested to consider this option if resources permit.
Accessing content on a mobile phone (Environmental Context and Resources)	• Restructuring the physical environment	• Optimize website layout on mobile interface • Develop a mobile app	• An effective, practicable, and acceptable option to improve the readability and usability of forums on mobile devices. [49] • Effective and acceptable by young people, but may be less affordable than optimizing layout on the mobile interface [50]. Organizations may wish to consider this if resources permit.
Time taken to obtain support (Environmental Context and Resources)	• Restructuring the physical environment	• Leverage technology to facilitate quicker responses (eg, using automated triage systems; artificial intelligence as virtual assistants to help with replies) • Young people could be involved in developing the artificial intelligence tool to ensure it is acceptable, effective, and inclusive.	• Effective and acceptable, having the potential positive side effect of reducing the workload of moderators [51,52]. However, it is less affordable than having manual processes in place to prioritize replies, and organizations may wish to consider this if resources permit.
Interactions with other users (Social Influences)	• Social reward • Social support (unspecified)	• Moderators validate young people who create their first post or respond to posts; badges • Unmoderated forums to introduce moderators who can actively look out for negative interactions	• This is an affordable, practicable, and acceptable way to increase positive interactions on the forum, as such social interactions could improve mood and increase intrinsic motivation [53]. • For forums without moderators, this is an effective, practicable, and acceptable way to manage negative interactions between users [9,28].
Poor communication skills; self-understanding of emotions (Skills)	• Instruction on how to perform behavior	• Online self-administered training on ways to communicate effectively; self-paced learning tools/modules	• This option of using training to improve young people's communication skills is practicable [54]. An online self-administered training was assessed to be more effective than other methods of instruction, such as email (not suitable as it may be too wordy and hard for young people to absorb the information) or in-person training (more resource-intensive and difficult to accommodate available timeslots).
Beliefs about helpfulness of reply (Beliefs About Consequences)	• Information about social and environmental consequences	• Customized email outreach highlighting the positive consequences of posting and replying to forum posts	• This is an affordable and practicable option to convey succinct information. Using a personalized and persuasive tone in these types of communications has been shown to be effective in increasing participation in web-based communities [55].

[a]APEASE: Affordability, Practicability, Effectiveness, Acceptability, Side-effects, and Equity.

[b]TDF: Theoretical Domains Framework.

[c]BCT: behavior change technique.

XSL•FO
RenderX

Discussion

Principal Findings

This study addresses an existing gap by examining the factors influencing young people's engagement in youth mental health forums, including creating posts, responding to posts, and reading posts. Ten core inductive themes were identified across 5 TDF domains, including Skills, Beliefs About Consequences, Emotions, and the Social Influences and Environmental Context and Resources. Of these, 3 themes were enablers, 2 were barriers, and 5 functioned as both enablers and barriers. Some themes were relevant to all 3 behaviors, such as the presence of moderators, while others were specific to more active participation (ie, posting and responding), such as positive interactions with other users. While most themes were relevant to either active participation or overall engagement, some acted as a barrier for 1 behavior but an enabler for another, or were specific to responding and reading only. For example, the ability to quickly find the information needed encouraged reading posts, whereas slow support and responses acted as a barrier to posting new content. Empathic skills were important for both reading and responding. Most barriers were related to physical and social opportunity, suggesting that adjustments to the service structure could enhance young people's engagement. Below, the enablers of engagement are discussed first, followed by the barriers, along with the proposed contextualized BCTs to optimize engagement.

Enablers of Engagement With Youth Mental Health Forums

A range of capability, opportunity, and motivation factors influence young people's use of youth mental health forums. Regarding automatic motivation, this study found that both passive and active engagement with the forum can generate positive feelings, which serve as reinforcement to encourage similar behaviors in the future. This aligns with Smith-Merry et al's [28] findings, which reported that users of web-based mental health forums viewed responding as therapeutic and associated with positive emotions.

Regarding psychological capability, many young people in the study highlighted empathy as an important skill for responding to and reading others' posts. Empathy enables users to better understand others' situations and provide sensitive, thoughtful replies. Perowne and Gutman's [42] study found that empathy in communication was a key skill for moderators of self-harm content on young people's web-based forums, and this same skill is required by users to effectively understand and respond to others' posts.

Regarding physical opportunity, young people in this study reported that a platform allowing anonymity enabled them to express their opinions with less fear of judgment, thereby facilitating posting and responding behaviors. This finding aligns with research in other fields, such as education and physical health, which indicates that anonymity generally promotes self-disclosure and facilitates engagement with web-based forums [56-58].

Barriers to Engagement and Suggested BCTs

Presence of Moderators

Previous studies [9,59] have demonstrated that forum moderators play a crucial role in maintaining a safe online environment for participation. This was echoed in our study, where participants reported that the presence of moderators facilitated engagement by mitigating potential adverse effects of anonymity, such as bullying or disrespectful behavior resulting from disinhibition [14]. This study extends existing literature by highlighting a potential drawback of moderation: some young people may be reluctant to post for fear of negative consequences, such as being banned, if they freely express their thoughts or experiences. The BCTs of "adding objects to the environment" and "restructuring the physical environment" may help mitigate this issue. Specifically, forum providers can reduce ambiguity by offering clear rules or guidelines regarding content that should not be posted [60], giving young people greater certainty about what is acceptable on the forums. Additionally, alternative interaction mechanisms (eg, 1-on-1 chats, if available) can be offered for users to share personal anecdotes of a sensitive nature, ensuring they continue to feel supported for their mental health needs [47]. Providing such avenues allows young people to express their thoughts safely while reducing the likelihood of posting inappropriate content on web-based forums.

Relevance of Information

Consistent with past studies [29], another barrier to reading and responding to posts in youth mental health forums was the difficulty in finding relevant information. Most forums organize posts by recency (ie, the most recent posts appear at the top) rather than providing customized content tailored to users' interests. Furthermore, when forum content is not categorized intuitively or specifically enough (eg, a category is too generic and includes posts spanning a wide range of sub-topics), it hinders young people's ability to find personally relevant information. This, in turn, affects their ability to read or share knowledge and reduces their confidence in replying effectively. Interestingly, the inability to find relevant information may actually increase posting behavior, as users attempt to obtain content that is more personally relevant. However, when young people post solely because they cannot easily locate existing information, it can contribute to more repetitive content on the forums. Thus, it is important to optimize content relevance to reduce repetitive information and maintain the quality of interactions in the forums. To help young people find information that is relevant for reading and responding, forum providers could apply the BCT of "restructuring the physical environment," delivering more relevant or customized content based on users' previously indicated preferences. Implementing recommender engines based on participant preferences has been shown to enhance forum engagement, and this approach could be applied in the current context, as young people are more likely to read and respond to posts they find personally relevant [48].

XSL•FO
RenderX

Accessing Content on a Mobile Phone

Many young people value the convenience, immediacy, and ease of accessing services via their mobile phones [61]. Consequently, difficulties in using web-based forums on mobile devices—such as challenges with reading or typing—generally hindered posting, reading, and responding behaviors. By applying the BCT of "restructuring the physical environment," providers of youth mental health forums could optimize the layout of their websites for mobile interfaces. Additionally, forum providers might consider developing a mobile app to further facilitate ease of use on mobile devices. Converting websites into mobile apps offers several advantages, including the ability to provide instantaneous notifications (if users allow) and to sort content based on user preferences [50]. Providing more customized content through mobile apps could also help address the barrier of difficulty in finding relevant information on the forums. This, in turn, may lead to higher engagement and increased responding behavior.

Time Taken to Obtain Support

Efficiency is a key factor influencing user engagement in help-seeking and information-seeking via web-based forums [62]. Studies by Browne et al [63] and Lobban et al [12] found that forum users valued timely responses to their posts, as this helped them address their current difficulties. Similarly, users who read posts to seek help may be motivated by the desire for immediate support. To minimize the time required to obtain assistance on web-based forums, providers could apply the BCT of "restructuring the physical environment," leveraging technology to facilitate faster responses. For example, previous studies have highlighted the utility of automated triage systems to help moderators identify posts requiring immediate attention [51], as well as the use of artificial intelligence as virtual assistants to respond during off-peak hours [52]. These technological interventions can reduce reliance on manual post monitoring and increase the efficiency of responses.

Interactions With Other Users

Consistent with previous studies [14], some young people reported reservations about posting or responding due to prior negative interactions with other users. By contrast, others noted that they were more inclined to respond to users who had previously replied to them, supporting the concept of reply reciprocity [31,64]. This factor specifically affects active engagement (ie, posting and responding) and underscores the importance of fostering positive interactions on the forums to create a reinforcing feedback loop. To encourage greater posting and responding, forum providers could apply the "social reward" BCT, with moderators placing added emphasis on recognizing users who have made their first post or responded to others, thereby motivating continued engagement. Applying the BCT of "social support (unspecified)," it is important for unmoderated youth mental health forums to introduce moderators. These moderators should actively monitor the forums for negative interactions and work to diffuse and resolve conflicts efficiently [9,28]. By increasing positive interactions and reducing negative ones, overall active engagement with youth mental health forums can be facilitated.

Poor Communication Skills

Existing literature highlights communication skills as a key factor influencing forum engagement. Poor communication skills can act as a barrier, limiting young people's ability to post about their mental health issues or respond to others' posts. Consequently, young people had difficulty understanding poorly written posts, which hindered their reading. Previous research also indicates that longer posts can impede reading, as users are more likely to lose focus unintentionally [65]. Unclear or incoherent posts and replies can also affect users' ability to understand and act on the information provided [66,67]. To enhance communication skills, youth mental health forum providers could implement the "instruction on how to perform behavior" BCT. This might involve an online, self-administered training module that young people can access at their convenience. The training could include guidance on posting and replying effectively (eg, being clear and concise) and emphasize the importance of responding to other users' replies when they agree with the content. Although not in the same context, previous research has shown that online, self-administered social skills training can effectively reduce social fear and enhance social skills [54]. This highlights the potential of using online, self-administered training to enhance young people's skills and confidence in contributing positively to forums.

Beliefs About Helpfulness of Reply

A recent systematic review indicated that perceived effectiveness of professional help is a key factor influencing mental health help-seeking behavior. Adolescents' beliefs about the helpfulness of advice or interventions can significantly affect their likelihood of seeking support; for example, they may avoid seeking help if they have doubts about its effectiveness [68]. An empirical study of a web-based mental health community intervention also found that the quality and effectiveness of responses influenced young people's behavior in creating posts to seek help online [69]. To address the barrier of negative perceptions regarding the helpfulness of replies, forum providers could consider making the benefits of using web-based mental health forums more salient to users. The BCT "information about social and environmental consequences" could be used to emphasize the positive outcomes of posting or replying to forum posts. This could be paired with an active call to action, encouraging users to create posts or respond to existing ones. Implementation could include customized email outreach to new forum members or to existing members who have not yet posted or responded. The email could highlight the positive outcomes of posting or responding helpfully to other users, including validating those who have replied to the original post. It could also encourage recipients to complete the online, self-administered training mentioned above, helping to build their confidence in posting or replying. Directly reaching out to forum members with a persuasive message can effectively promote responding behavior. In similar contexts, studies have shown that actively inviting users to participate in discussions on web-based forums leads to increased engagement [55].

XSL•FO
RenderX

Limitations

The following limitations should be considered when interpreting the findings of this study. First, the use of convenience sampling may have introduced self-selection bias, meaning that individuals who consented to participate in the interviews could differ systematically from those who did not [70]. A similar study by Prescott et al [29] found that participants who volunteered for interviews were more likely to emphasize the advantages of web-based mental health forums rather than the disadvantages, possibly because individuals with more positive experiences were more inclined to share their opinions. To mitigate this, participants were assured of the anonymity of their responses and encouraged to provide honest feedback. Future research could also aim for a more representative sample, including a balanced mix of genders, as there may be gender differences in the use of online platforms for mental health support [8]. Additionally, most participants in this study used a single web-based mental health forum in the United Kingdom, which may limit the generalizability of the findings to other forums in different contexts. Future research could explore the barriers and enablers to posting, responding, and reading posts on other mental health forums targeting different age groups or populations. As the proposed intervention strategies were not evaluated in this study, future research could assess their effectiveness in optimizing young people's engagement with web-based mental health forums. Finally, as forums represent only a subset of the online mental health ecosystem, it may be valuable to explore how these strategies could be applied more broadly to other platforms, such as social media, to maximize their impact.

Conclusions

In today's digitally connected age, the internet serves as a valuable conduit for knowledge sharing and building connections. Youth mental health forums have the potential to provide accessible support, complement in-person services, and positively contribute to young people's mental well-being. However, a lack of engagement can limit the impact and effectiveness of these forums in supporting young people. This study contributes to the existing knowledge by examining the various ways in which young people engage with youth mental health forums. Different strategies may be prioritized depending on whether forum providers aim to increase active engagement (eg, encouraging posting and responding through positive interactions) or overall engagement (eg, implementing moderators and clear rules to support both active and passive participation).

Overall, these findings demonstrate the utility of the BCW approach in conducting behavioral diagnoses and developing behavior change interventions, contributing to the evidence supporting the application of behavior change frameworks in digital mental health interventions. This can help advance the emerging field of digital mental health intervention engagement, ultimately leading to higher-quality forum services that provide better support for young people.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview schedule.
[DOCX File , 17 KB - humanfactors_v12i1e71549_app1.docx]

Multimedia Appendix 2
Codebook.
[DOCX File , 19 KB - humanfactors_v12i1e71549_app2.docx]

Multimedia Appendix 3
Unique themes for each type of engagement behavior.
[DOCX File , 15 KB - humanfactors_v12i1e71549_app3.docx]

References

1. Ford K, Freund R. Young lives under pressure: protecting and promoting young people's mental health at a time of global crises. Young Lives. 2022 Nov. URL: https://www.younglives.org.uk/publications/young-lives-under-pressure-protecting-and-promoting-young-peoples-mental-health-time [accessed 2024-07-15]
2. Marchant N, Masiga J. How to tackle the mental health crisis facing young people. World Economic Forum. URL: https://www.weforum.org/stories/2024/07/mental-health-crisis-young-people/ [accessed 2024-06-30]
3. Montero-Marin J, Hinze V, Mansfield K, Slaghekke Y, Blakemore S, Byford S, MYRIAD Team. Young people's mental health changes, risk, and resilience during the COVID-19 pandemic. JAMA Netw Open 2023 Sep 05;6(9):e2335016. [doi: 10.1001/jamanetworkopen.2023.35016] [Medline: 37733343]
4. McCurdy C, Murphy L. We've only just begun. Resolution Foundation. URL: https://www.resolutionfoundation.org/publications/weve-only-just-begun/ [accessed 2024-07-14]

XSL•FO
RenderX

5. Crenna-Jennings W. Non-specialist mental health support for young people in England. Education Policy Institute. URL: https://epi.org.uk/publications-and-research/non-specialist-mental-health-support-for-young-people-in-england/ [accessed 2024-07-14]
6. Children's mental health services 2022-23. Children's Commissioner. 2024 Mar 14. URL: https://www.childrenscommissioner.gov.uk/resource/childrens-mental-health-services-2022-23/ [accessed 2024-06-23]
7. Evans S. The challenge and potential of the digital age. Transactional Analysis Journal 2017 Dec 28;44(2):153-166. [doi: 10.1177/0362153714545312]
8. Faverio M, Anderson M, Park E. Teens, social media and mental health. Pew Research Center. 2025 Apr 22. URL: http://www.pewresearch.org/internet/2025/04/22/teens-social-media-and-mental-health/ [accessed 2025-10-11]
9. Hanley T, Prescott J, Gomez KU. A systematic review exploring how young people use online forums for support around mental health issues. J Ment Health 2019 Oct 03;28(5):566-576. [doi: 10.1080/09638237.2019.1630725] [Medline: 31267813]
10. Barak A, Dolev-Cohen M. Does activity level in online support groups for distressed adolescents determine emotional relief. Couns and Psychother Res 2007 Feb 17;6(3):186-190. [doi: 10.1080/14733140600848203]
11. Horgan A, McCarthy G, Sweeney J. An evaluation of an online peer support forum for university students with depressive symptoms. Arch Psychiatr Nurs 2013 Apr;27(2):84-89 [FREE Full text] [doi: 10.1016/j.apnu.2012.12.005] [Medline: 23540518]
12. Lobban F, Akers N, Appelbe D, Chapman L, Collinge L, Dodd S, et al. Clinical effectiveness of a web-based peer-supported self-management intervention for relatives of people with psychosis or bipolar (REACT): online, observer-blind, randomised controlled superiority trial. BMC Psychiatry 2020 Apr 14;20(1):160 [FREE Full text] [doi: 10.1186/s12888-020-02545-9] [Medline: 32290827]
13. van Mierlo T. The 1% rule in four digital health social networks: an observational study. J Med Internet Res 2014 Feb 04;16(2):e33 [FREE Full text] [doi: 10.2196/jmir.2966] [Medline: 24496109]
14. Marshall P, Booth M, Coole M, Fothergill L, Glossop Z, Haines J, et al. Understanding the impacts of online mental health peer support forums: realist synthesis. JMIR Ment Health 2024 May 09;11:e55750 [FREE Full text] [doi: 10.2196/55750] [Medline: 38722680]
15. Preece J, Nonnecke B, Andrews D. The top five reasons for lurking: improving community experiences for everyone. Computers in Human Behavior 2004 Mar;20(2):201-223. [doi: 10.1016/j.chb.2003.10.015]
16. Leshed G. Posters, lurkers, and in between: a multidimensional model of online community participation patterns. Cornell University. 2005 Jan. URL: https://www.cs.cornell.edu/~gl/HCII2005.pdf [accessed 2024-06-06]
17. Himelboim I, Gleave E, Smith M. Discussion catalysts in online political discussions: content importers and conversation starters. J Comput Mediat Commun 2009;14(4):771-789. [doi: 10.1111/j.1083-6101.2009.01470.x]
18. Prescott J, Hanley T, Ujhelyi K. Peer communication in online mental health forums for young people: directional and nondirectional support. JMIR Ment Health 2017 Aug 02;4(3):e29 [FREE Full text] [doi: 10.2196/mental.6921] [Medline: 28768607]
19. Banwell E, Hanley T, De Ossorno Garcia S, Mindel C, Kayll T, Sefi A. The helpfulness of web-based mental health and well-being forums for providing peer support for young people: cross-sectional exploration. JMIR Form Res 2022 Sep 09;6(9):e36432 [FREE Full text] [doi: 10.2196/36432] [Medline: 36083629]
20. Koh J, Kim Y, Butler B, Bock G. Encouraging participation in virtual communities. Commun ACM 2007 Feb 01;50(2):68-73. [doi: 10.1145/1216016.1216023]
21. Michie S, Atkins L, West R. The Behaviour Change Wheel: A Guide to Designing Interventions. Sutton, UK: Silverback Publishing; 2014.
22. Chan JK, Farrer LM, Gulliver A, Bennett K, Griffiths KM. University students' views on the perceived benefits and drawbacks of seeking help for mental health problems on the internet: a qualitative study. JMIR Hum Factors 2016 Jan 19;3(1):e3 [FREE Full text] [doi: 10.2196/humanfactors.4765] [Medline: 27026140]
23. Breuer L, Barker C. Online support groups for depression. Sage Open 2015 Apr 08;5(2):2158244015574936. [doi: 10.1177/2158244015574936]
24. Bond CS, Ahmed OH. Can I help you? Information sharing in online discussion forums by people living with a long-term condition. J Innov Health Inform 2016 Nov 10;23(3):620-626 [FREE Full text] [doi: 10.14236/jhi.v23i3.853] [Medline: 28059696]
25. Malik S, Coulson NS. The therapeutic potential of the internet: exploring self-help processes in an internet forum for young people with inflammatory bowel disease. Gastroenterol Nurs 2011;34(6):439-448. [doi: 10.1097/SGA.0b013e318237a9ba] [Medline: 22129797]
26. Liao Siling J, Truss K, Philips L, Eastwood O, Bendall S. Young people's journeys of recovery from trauma: a qualitative study of narratives from internet forums. Psychol Trauma 2021 Mar;13(3):293-301. [doi: 10.1037/tra0000990] [Medline: 33661674]
27. Smit D, Vrijsen JN, Groeneweg B, Vellinga-Dings A, Peelen J, Spijker J. A newly developed online peer support community for depression (Depression Connect): qualitative study. J Med Internet Res 2021 Jul 12;23(7):e25917 [FREE Full text] [doi: 10.2196/25917] [Medline: 34255659]

XSL•FO
RenderX

28. Smith-Merry J, Goggin G, Campbell A, McKenzie K, Ridout B, Baylosis C. Social connection and online engagement: insights from interviews with users of a mental health online forum. JMIR Ment Health 2019 Mar 26;6(3):e11084 [FREE Full text] [doi: 10.2196/11084] [Medline: 30912760]
29. Prescott J, Hanley T, Ujhelyi Gomez K. Why do young people use online forums for mental health and emotional support? Benefits and challenges. British Journal of Guidance & Counselling 2019 May 20;47(3):317-327. [doi: 10.1080/03069885.2019.1619169]
30. Bargh JA, McKenna KYA. The internet and social life. Annu Rev Psychol 2004 Feb;55(1):573-590. [doi: 10.1146/annurev.psych.55.090902.141922] [Medline: 14744227]
31. Pan W, Shen C, Feng B. You get what you give: understanding reply reciprocity and social capital in online health support forums. J Health Commun 2017 Jan 27;22(1):45-52 [FREE Full text] [doi: 10.1080/10810730.2016.1250845] [Medline: 28027009]
32. Smithson J, Sharkey S, Hewis E, Jones RB, Emmens T, Ford T, et al. Membership and boundary maintenance on an online self-harm forum. Qual Health Res 2011 Nov 29;21(11):1567-1575. [doi: 10.1177/1049732311413784] [Medline: 21715606]
33. Michie S, West R. Behaviour change theory and evidence: a presentation to government. Health Psychology Review 2013 Mar;7(1):1-22. [doi: 10.1080/17437199.2011.649445]
34. Atkins L, Francis J, Islam R, O'Connor D, Patey A, Ivers N, et al. A guide to using the theoretical domains framework of behaviour change to investigate implementation problems. Implement Sci 2017 Jun 21;12(1):77 [FREE Full text] [doi: 10.1186/s13012-017-0605-9] [Medline: 28637486]
35. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
36. Theory and Technique Tool. 2024. URL: https://theoryandtechniquetool.humanbehaviourchange.org/ [accessed 2024-06-12]
37. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95 [FREE Full text] [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
38. Richiello MG, Mawdsley G, Gutman LM. Using the Behaviour Change Wheel to identify barriers and enablers to the delivery of webchat counselling for young people. Couns and Psychother Res 2021 May 04;22(1):130-139. [doi: 10.1002/capr.12410]
39. Mawdsley G, Richiello M, Gutman LM. Barriers and facilitators of young people's engagement with webchat counselling: a qualitative analysis informed by the Behaviour Change Wheel. Couns and Psychother Res 2022 Mar 23;22(3):725-735. [doi: 10.1002/capr.12528]
40. Campagnola M, Burlibasa A, Gutman LM. Barriers and enablers to the delivery of email communication for a helpline service for young people. PEC Innov 2022 Dec;1:100059 [FREE Full text] [doi: 10.1016/j.pecinn.2022.100059] [Medline: 37213768]
41. Ananya A, Tuuli J, Perowne R, Gutman LM. Barriers and facilitators to user engagement and moderation for web-based peer support among young people: qualitative study using the behavior change wheel framework. JMIR Hum Factors 2025 Feb 05;12:e64097 [FREE Full text] [doi: 10.2196/64097] [Medline: 39908551]
42. Perowne R, Gutman LM. Barriers and enablers to the moderation of self-harm content for a young person's online forum. J Ment Health 2024 Jun 16;33(3):357-365 [FREE Full text] [doi: 10.1080/09638237.2022.2069721] [Medline: 35574666]
43. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013 Sep 18;13:117 [FREE Full text] [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]
44. Goldsmith L. Using framework analysis in applied qualitative research. TQR 2021 Jun 20;26(6):2061-2076. [doi: 10.46743/2160-3715/2021.5011]
45. Braun V, Clarke V. One size fits all? What counts as quality practice in (reflexive) thematic analysis? Qualitative Research in Psychology 2020 Aug 12;18(3):328-352. [doi: 10.1080/14780887.2020.1769238]
46. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2008 Jul 21;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
47. Saha K, Ernala S, Dutta S, Sharma E, De CM. Understanding moderation in online mental health communities. 2020 Jul Presented at: 22nd HCI International Conference; July 19-24, 2020; Copenhagen, Denmark p. 87-107. [doi: https://doi.org/10.1007/978-3-030-49576-3_7]
48. Albatayneh N, Ghauth K, Chua F. Utilizing learners' negative ratings in semantic content-based recommender system for e-learning forum. Educ Technol Soc 2018 Jan:112-125 [FREE Full text]
49. Radilova M, Kamencay P, Matuska S, Benco M, Hudec R. Tool for optimizing webpages on a mobile phone. 2020 Jul Presented at: 2020 43rd International Conference on Telecommunications and Signal Processing (TSP); July 7-9, 2020; Milan, Italy. [doi: 10.1109/tsp49548.2020.9163449]
50. Jovenn C, Subaramaniam K, Jalil A. The development of a forum mobile application for students. 2019 Presented at: 2019 IEEE 9th International Conference on System Engineering and Technology (ICSET); October 7, 2019; Shah Alam, Malaysia p. 90-95. [doi: 10.1109/icsengt.2019.8906430]

XSL•FO
RenderX

51. Milne DN, McCabe KL, Calvo RA. Improving moderator responsiveness in online peer support through automated triage. J Med Internet Res 2019 Apr 26;21(4):e11410 [FREE Full text] [doi: 10.2196/11410] [Medline: 31025945]
52. Liu X, Pankiewicz M, Gupta T, Huang Z, Baker RS. A step towards adaptive online learning: exploring the role of GPT as virtual teaching assistants in online education. Open Science Framework. 2025 Feb. URL: https://doi.org/10.35542/osf.io/rw45b [accessed 2025-03-12]
53. Kraut R, Resnick P, Kiesler S. Building Successful Online Communities: Evidence-Based Social Design. Boston, MA: The MIT Press; 2012.
54. Lehenbauer M, Kothgassner OD, Kryspin-Exner I, Stetina BU. An online self-administered social skills training for young adults: results from a pilot study. Computers & Education 2013 Feb;61:217-224. [doi: 10.1016/j.compedu.2012.09.007]
55. Harper F, Frankowski D, Drenner S. Talk amongst yourselves: inviting users to participate in online conversations. 2007 Presented at: IUI '07: Proceedings of the 12th International Conference on Intelligent User Interfaces; January 28-31, 2007; Honolulu, HI p. 62-71. [doi: 10.1145/1216295.1216313]
56. Wadden D, August T, Li Q, Althoff T. The effect of moderation on online mental health conversations. 2021 May 22 Presented at: Proceedings of the International AAAI Conference on Web and Social Media; June 7-10, 2021; Virtual p. 751-763. [doi: https://doi.org/10.48550/arXiv.2005.09225]
57. Roberts LD, Rajah-Kanagasabai CJ. "I'd be so much more comfortable posting anonymously": identified versus anonymous participation in student discussion boards. AJET 2013 Nov 08;29(5):452. [doi: 10.14742/ajet.452]
58. Jones R, Sharkey S, Ford T, Emmens T, Hewis E, Smithson J, et al. Online discussion forums for young people who self-harm: user views. Psychiatrist 2018 Jan 02;35(10):364-368. [doi: 10.1192/pb.bp.110.033449]
59. Kendal S, Kirk S, Elvey R, Catchpole R, Pryjmachuk S. How a moderated online discussion forum facilitates support for young people with eating disorders. Health Expect 2017 Feb 03;20(1):98-111 [FREE Full text] [doi: 10.1111/hex.12439] [Medline: 26725547]
60. Mokkenstorm JK, Mérelle SYM, Smit JH, Beekman ATF, Kerkhof AJFM, Huisman A, et al. Exploration of benefits and potential harmful effects of an online forum for visitors to the suicide prevention platform in the Netherlands. Crisis 2020 May;41(3):205-213. [doi: 10.1027/0227-5910/a000627] [Medline: 31657643]
61. O'Dea B, Han J, Batterham PJ, Achilles MR, Calear AL, Werner-Seidler A, et al. A randomised controlled trial of a relationship-focussed mobile phone application for improving adolescents' mental health. J Child Psychol Psychiatry 2020 Aug;61(8):899-913 [FREE Full text] [doi: 10.1111/jcpp.13294] [Medline: 32683737]
62. Kruzan KP, Fitzsimmons-Craft EE, Dobias M, Schleider JL, Pratap A. Developing, deploying, and evaluating digital mental health interventions in spaces of online help- and information-seeking. Procedia Comput Sci 2022;206:6-22 [FREE Full text] [doi: 10.1016/j.procs.2022.09.081] [Medline: 37063642]
63. Browne NL, Carragher NO, O'Toole A, Pimm J, Ryder J, Thew GR. Evaluating user experiences of SHaRON: an online CBT-based peer support platform. tCBT 2022 Apr 11;15:e18. [doi: 10.1017/s1754470x22000150]
64. Chang PF, Whitlock J, Bazarova NN. "To respond or not to respond, that is the question": the decision-making process of providing social support to distressed posters on Facebook. Social Media + Society 2018 Feb 15;4(1):205630511875929. [doi: 10.1177/2056305118759290]
65. Forrin ND, Mills C, D'Mello SK, Risko EF, Smilek D, Seli P. TL;DR: longer sections of text increase rates of unintentional mind-wandering. The Journal of Experimental Education 2020 Jun 09;89(2):278-290. [doi: 10.1080/00220973.2020.1751578]
66. Berry N, Lobban F, Bucci S. A qualitative exploration of service user views about using digital health interventions for self-management in severe mental health problems. BMC Psychiatry 2019 Jan 21;19(1):35 [FREE Full text] [doi: 10.1186/s12888-018-1979-1] [Medline: 30665384]
67. Murnane EL, Cosley D, Chang P, Guha S, Frank E, Gay G, et al. Self-monitoring practices, attitudes, and needs of individuals with bipolar disorder: implications for the design of technologies to manage mental health. J Am Med Inform Assoc 2016 May;23(3):477-484. [doi: 10.1093/jamia/ocv165] [Medline: 26911822]
68. Radez J, Reardon T, Creswell C, Lawrence PJ, Evdoka-Burton G, Waite P. Why do children and adolescents (not) seek and access professional help for their mental health problems? A systematic review of quantitative and qualitative studies. Eur Child Adolesc Psychiatry 2021 Feb 21;30(2):183-211 [FREE Full text] [doi: 10.1007/s00787-019-01469-4] [Medline: 31965309]
69. Li J, Liu D, Wan C, Liang Z, Zhu T. Empirical study of factors that influence the perceived usefulness of online mental health community members. Psych J 2023 Apr;12(2):307-318. [doi: 10.1002/pchj.629] [Medline: 36726193]
70. Costigan CL, Cox MJ. Fathers' participation in family research: is there a self-selection bias? Journal of Family Psychology 2001;15(4):706-720. [doi: 10.1037/0893-3200.15.4.706]

Abbreviations

APEASE: Affordability, Practicability, Effectiveness, Acceptability, Side-effects, and Equity
BCT: behavior change technique
BCTT: behavior change techniques taxonomy
BCW: Behavior Change Wheel

XSL•FO
RenderX

COM-B: capability, opportunity, and motivation—behavior
TDF: Theoretical Domains Framework

Edited by H Witteman; submitted 21.01.25; peer-reviewed by L Prothero, A Zheluk; comments to author 09.10.25; revised version received 11.10.25; accepted 12.10.25; published 31.10.25.

Please cite as:
Lim Z, Lim SY, Lu S, Gutman LM
Barriers and Enablers to Young People's Posting, Responding, and Reading Behaviors on Mental Health Forums Using the Behavior Change Wheel: Qualitative Study
JMIR Hum Factors 2025;12:e71549
URL: https://humanfactors.jmir.org/2025/1/e71549
doi:10.2196/71549
PMID:

©Zhen Lim, Si Yin Lim, Siqi Lu, Leslie Morrison Gutman. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 31.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Cycling-Based Telerehabilitation: Acceptability and Feasibility Study

Sara Arlati[1], PhD; Vera Colombo[1], PhD; Marta Mondellini[1,2], MS; Roberta Nossa[3], PhD; Chiara Grasso[4], MS; Mauro Rossini[5], MS; Emilia Biffi[3], PhD; Alessia Fumagalli[4], MD; Eleonora Diella[3], MS; Eleonora Guanziroli[6], PhD; Nicole Sanna[7], MS; Emilia Ambrosini[7], Prof Dr; Marco Sacco[1], MS; Franco Molteni[6], MD

[1]Institute of Intelligent Industrial Technologies and Systems for Advanced Manufacturing, National Research Council of Italy, Via Previati 1/E, Lecco, Italy

[2]Department of Psychology, Università Cattolica del Sacro Cuore, Milano, Italy

[3]Scientific Institute, IRCCS Medea, Bosisio Parini, Italy

[4]Respiratory Unit, IRCCS INRCA (Italian National Research Centre On Aging), Casatenovo, Italy

[5]Valduce Hospital, Como, Italy

[6]Villa Beretta Rehabilitation Center, Costa Masnaga, Italy

[7]Department of Electronics Information and Bioengineering, Politecnico di Milano, Milano, Italy

Corresponding Author:
Sara Arlati, PhD
Institute of Intelligent Industrial Technologies and Systems for Advanced Manufacturing, National Research Council of Italy, Via Previati 1/E, Lecco, Italy

Abstract

Background: Telerehabilitation is a promising solution to provide continuity of care. Most existing telerehabilitation platforms focus on rehabilitating upper limbs, balance, and cognitive training, but exercises improving cardiovascular fitness are often neglected.

Objective: The objective of this study is to evaluate the acceptability and feasibility of a telerehabilitation intervention combining cognitive and aerobic exercises.

Methods: A virtual reality–based dual-task exercise exploiting a cycle ergometer was designed, developed, and integrated with a commercially available telerehabilitation platform. Patients with different conditions were enrolled and administered subjective questionnaires investigating attitudes toward technology, usability, technology acceptance, and subjective workload. Their therapists were interviewed, and adherence and performance data were analyzed.

Results: In total, 26 patients with neurological or post-COVID symptoms were included. Their attitude toward technology (range: 0‐5) did not change after the training period (pre: 3.44 [IQR 0.63]; post: 3.50 [IQR 0.48]); the platform was rated usable and acceptable. Frustration and physical and mental workload were present, especially among younger participants. The adherence was moderate, but individual differences were present (0.59 [IQR 0.54]). The therapists highlighted the potential of remote rehabilitation programs but also identified some limitations.

Conclusions: This study proved the feasibility and acceptability of a customized virtual reality–based telerehabilitation program allowing for the safe implementation of aerobic cycling-based dual-task training. The solution was judged meaningful for dehospitalized patients, although some environmental and technical barriers should be overcome to implement telerehabilitation more effectively.

(JMIR Hum Factors 2025;12:e71099) doi:10.2196/71099

KEYWORDS

telemedicine; virtual reality; exergaming; aerobic exercise; usability

Introduction

Rehabilitation comprises a series of interventions aimed at optimizing the functioning of an individual; it can play an essential role in regaining or maintaining autonomy during activities of daily living and reduce the personal, social, and economic impact of several acute or chronic conditions. On the contrary, without regular rehabilitation care, individuals with neurological, rheumatological, musculoskeletal, or cognitive disorders may experience a worsening of their condition, reduced quality of life, and emotional distress [1].

XSL•FO
RenderX

In such a context, a possible solution for the sustainable implementation of continuity of care is telerehabilitation, i.e., the remote delivery of rehabilitation services using information and communication technologies [2-5]. Telerehabilitation increases accessibility to care services in remote areas and developing countries [6] and enables patients to engage in higher doses of therapy or repetition, thus improving their performance at impairment, activity, and participation levels [7,8]. Existing studies have proven that telerehabilitation can improve patients' adherence and satisfaction [9], allowing clinical personnel to assist patients who cannot travel or have mobility issues and leading to improved clinical outcomes similar to those obtained with standard therapy [10].

The importance of such remote care-related services has been further emphasized during the COVID-19 pandemic, which hit Italy severely in February 2020 [11]. The emergency that arose significantly overloaded hospitals and led to halting or limiting several clinical services [12]. Such an interruption caused detrimental effects on many patients who were in need of rehabilitation, raising the need to develop efficient, safe, and long-term remote rehabilitative programs [13].

Currently, telerehabilitation can be administered via video and phone calls from a therapist or through a platform that provides the patients with videos, messages, emails, and links to educational materials. Also, it can be supported by virtual reality (VR), a digital technology that allows the patient to train in a computer-generated scenario [14]. VR could generally offer many advantages to rehabilitation interventions [15]: it allows for the simulation of ecological and controlled environments in which patients can train safely and promotes the transferring of the acquired abilities to real life [16,17]. Also, VR has been proven to enhance performance and contribute to keeping patients' motivation at a high level [18,19]. Thus, it can play a key role in promoting adherence, especially for patients with chronic conditions and needing (life-)long treatments.

When COVID arose, several telerehabilitation services were commercially available on the market [2,4]. However, most of them provided only cognitive exercises; a few focused on upper and lower body muscle strength and mobility [20] but did not address lower limb coordination and cardiovascular fitness [21], which—instead—could be very helpful for patients with post-COVID syndrome and chronic conditions in general [22-24].

This work was conceived to respond to the lack of rehabilitation services during the second wave of the COVID pandemic (September to December 2020). In particular, it aimed to design and develop a novel VR-based exercise supporting lower body and cardiovascular training to complement the existing exercises dedicated to upper limbs and cognitive functions in a commercial telerehabilitation platform.

Afterward, the feasibility and acceptability of a customized telerehabilitation program, possibly including such an aerobic exercise, were assessed in different categories of patients.

The remainder of this paper is organized as follows. The Methods section describes the qualitative study performed; the Equipment subsection, besides technical aspects, presents the design choices and the rationale for developing the novel VR-supported exercise. The following sections contain, respectively, the study outcomes, their discussion, and the conclusions we drew with some recommendations for future work.

Methods

Study Design

The study presented in this work was a pilot, interventional, and multicentric study conducted between July 2022 and March 2023.

Ethical Considerations

The study was approved by the Ethics Committees of Ospedale Valduce – Villa Beretta Rehabilitation Center (VB) (Prot. N. 150/2022; date of approval: February 24, 2022), IRCCS E. Medea (MEDEA) (Prot. N. 23/22 – CE; date of approval: March 31, 2022), IRCCS INRCA Casatenovo (INRCA) (Prot. N. 14/22; date of approval: May 26, 2022). All the procedures were performed under the Declaration of Helsinki. All adult participants gave their written informed consent. In the case of participants below legal age, informed written consent was obtained from their parents or legal guardians. No compensation for participation was given.

Equipment

Requirements and Design

As this work was conceived to give an immediate response during the pandemic, the following three requirements were set: (1) the system must accommodate different conditions, including post-COVID patients, and include different exercises for motor, aerobic, and cognitive training; (2) it must be safe for home use; and (3) it had to be available in a short time.

The first two requirements were not easy to satisfy, as, in general, program interventions focusing on balance and gait were delivered in person due to the potential risks of falls [25]. Experiences reported in the literature either occurred while sitting or lying down or were dedicated to orthopedic patients with no other co-morbidities [20]. Furthermore, in most cases, the focus was on increasing muscle strength and mobility rather than promoting coordination and improving cardiovascular fitness [21].

We used a cycle ergometer to satisfy safety and controllability requirements and the need to support an aerobic training program for lower limbs. In this way, in fact, aerobic training could be carried out safely while seated and holding the cycle ergometer handlebars [26,27]. Moreover, using an ergometer allows for workload adjustment and, thus, the possibility of modifying the difficulty of the proposed task according to each patient's characteristics [28]. Finally, cycling and walking share common features: both movements are periodical, require alternative joints' flexion/extension, and require the contraction of agonist/antagonist muscles in a coordinated pattern [29]. To ensure safety related to the accomplishment of physical activity, additional physiological sensors were added to the system (i.e., heart rate (HR) and oxygen saturation) to interrupt the exercise when it was too exhausting.

XSL•FO
RenderX

To answer the third requirement, two elements were considered. First, we decided to use commercially available equipment already certified by Medical CE. Second, we investigated the availability of telerehabilitation companies to integrate a cycling-based task. These requirements were met by the cycle ergometer from COSMED (Rome, Italy) and the Virtual Reality Rehabilitation System (VRRS) produced by Khymeia Group (Padova, Italy). Both systems are further described in the following section, along with the developed virtual environment.

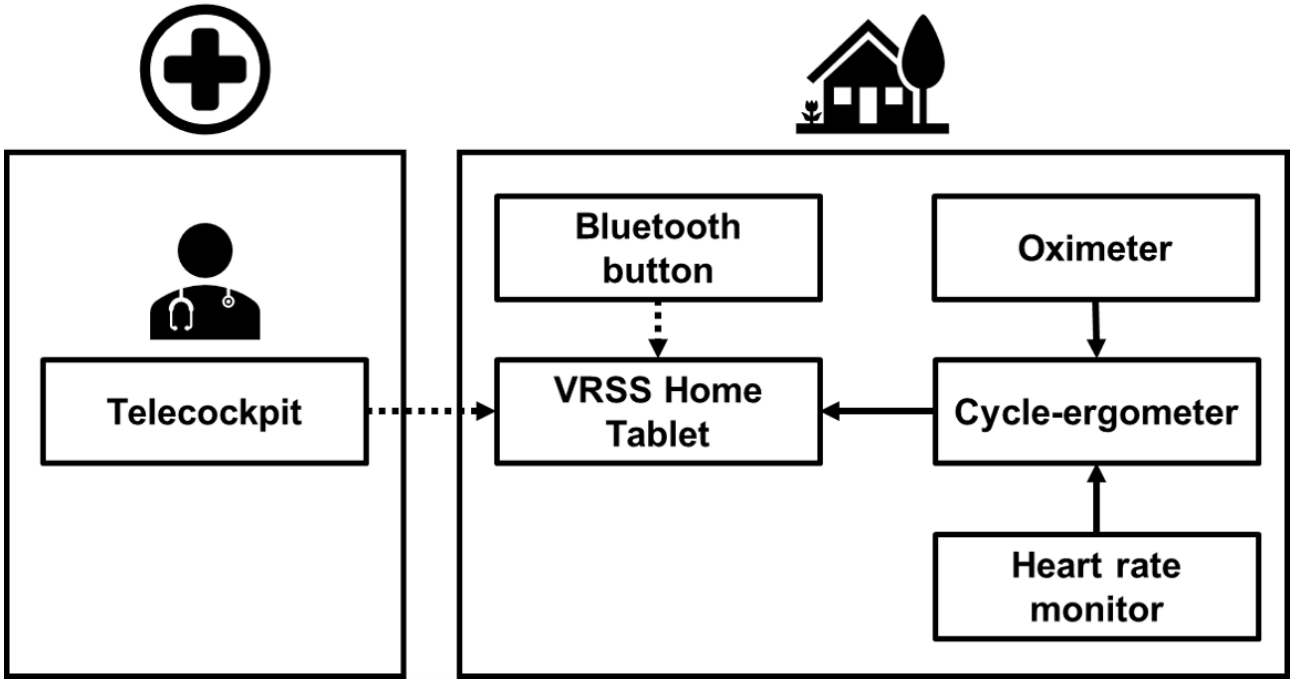
The Telerehabilitation System

The VRRS is a commercial telerehabilitation system composed of a tele-cockpit operated by the therapists in the clinics and a patient kit. The VRRS tele-cockpit allows for the creation of a daily rehabilitation plan and assessment of patients' compliance and performance. Through the tele-cockpit, the therapist can also video call the patients to assist her or him during the exercise execution. The patient kit comprises a Windows Surface laptop (VRRS Home Tablet), wearable sensors, and interaction devices (e.g., inertial measurement units). The VRRS system offers, per se, several modules dedicated to rehabilitating specific functions, e.g., cognitive, speech, neuromotor, and respiratory exercises [30].

For the sake of this study, the VRSS system was integrated with medical-certified cycle ergometers and a custom-built VR environment. In total, 3 different models of cycle ergometers were used: E100, E5, and E150 Pediatric. All were equipped with a chest band with a HR monitor; E5 models also integrated an oximeter to measure oxygen saturation (SpO_2).

The VR environment designed and developed with Unity (Unity Technologies) to support dual-task training was named ARTEDIA and made available, among other VRSS applications, on the Home Tablet. The technical connection details shown in Figure 1 are further described in [26].

Figure 1. A schema of the ARTEDIA application as integrated into the Virtual Reality Rehabilitation System (VRRS) from Khymeia [26].



The exercise occurs in a virtual park where the patient can travel along a predefined path; the forward velocity is synchronized according to the current ergometer's revolutions per minute (RPMs).

The cognitive exercise was based on the "go/no-go" paradigm, i.e., the participant has to respond when a "go" target appears; conversely, when a "no-go" target appears, participants should not react. During cycling, targets appear randomly at the edges of the path. The patient had to press the Bluetooth button mounted on the cycle-ergometer handlebar to kill the monsters ("go" targets), whereas animals ("no-go" targets) required no action (Figure 2). At each button press, the target disappeared, and visual feedback showed if the reaction was due or not. In case of a missed target, no feedback was provided.

XSL•FO
RenderX

Figure 2. A screenshot of the ARTEDIA application with a no-go target appearing along the route. On the right side, a patient doing the exercise is shown.



The therapist could customize the ARTEDIA exercise by selecting, for each session, the level of difficulty from 1 to 3 (targets frequency and highlight), the cycle ergometer workload (30-70 W), and the duration of the exercise (10, 15, or 20 minutes). Also, the therapist could define the maximum HR (range: 70 - 200 beats per minute) and minimum SpO_2 values (range: 85% - 100%). The exercise was automatically interrupted if the physiological measurement overcame these safety thresholds for more than 10 seconds. At the end of each session, correct answers, errors, missed targets, average RPMs, HR, and SpO_2 (if available) were saved.

Participants

The participants in this study were patients in diverse conditions but suitable to receive an unsupervised training program following their Individual Rehabilitation Project (IRP). Enrolled patients were adults with post-COVID syndrome (INRCA) and stroke (VB), children, adolescents, and young adults with neuromotor diseases (MEDEA).

Participants had to be older than 12. Participants above legal age (18 years) had to be able to provide informed written consent; for children and adolescents below 18, the consent was obtained by their parent or legal guardian. All patients were selected among the ones referring to the 3 involved clinical centers and were judged in need of an additional training period after discharge. All the participants had to live in Lombardy in a house with an available space of 1.5×1 meters (to accommodate the cycle ergometer easily), own a good internet connection, and be assisted by a caregiver who could support home-based training. Exclusion criteria were pain; tracheostomy; reduced trunk control; history of seizure; severe respiratory, renal, hepatic, or cardiological conditions; and the presence of motor, sensory, or cognitive deficits preventing the execution of the exercises. Patients taking part in other rehabilitative programs were also excluded.

Sample size estimation was made based on the *attitude toward technology* questionnaire developed by Huygelier et al. [31] (see paragraph Outcomes). Considering an effect size of 0.69 (as in the original work, i.e., [31]), α=0.05, and power =0.9, we obtained a sample size of 24 patients.

Study Protocol

Baseline

All patients were assessed for enrollment using the Trunk Control Test [32], the Visual Analog Scale for pain [33], and the Mini-Mental State Examination [34]. If considered eligible to participate in the study, each patient, depending on his or her condition, was administered a series of clinical measures to assess his or her baseline and define the IRP. These were the 6-minute walk test [35], the modified British Medical Research Council Questionnaire [36], the Quality of Upper Extremity Skills Test [37], the Gross Motor Function Measure [38], the Motricity Index [39], and the Berg Balance Scale [40].

Whenever possible, clinical scales were complemented with measures of cycling power to help the clinicians define ARTEDIA workload. Power measures were obtained using commercially available sensorized pedals (X-Power, SRM GmbH). This test saw 3 minutes of cycling on the cycle ergometer with a fixed workload set from 30 to 100 W. Data were analyzed to extract mean right and left power and the index of symmetry (*S*) [41]. The index of symmetry varies between 0 and 1, where 0 represents perfect symmetry, and 1 is the maximum power displacement (i.e., a leg has no or negative power).

Individual Rehabilitation Project

Depending on the baseline assessment, each patient was assigned to a complete telerehabilitation program, including lower limb aerobic and cognitive training with ARTEDIA, or an intervention foreseeing only Khymeia VRSS Home tablet exercises targeting cognitive and upper limb training.

The training was executed in unsupervised settings for all patients; most received the VRRS home tablet at their home. In total, 6 of them performed the rehabilitation program in a

XSL•FO
RenderX

room that one of the clinics (VB) set up to execute the training program. Such a room was near the clinic but only opened, closed, and sanitized by non-medical personnel; thus, patients were completely autonomous in following their IRP.

Each patient had access to the telerehabilitation system from 15 to 60 days; patients were instructed to train as defined by the therapist in their IPR and set in their VRSS Home Tablet profile.

Outcomes

Both subjective and objective outcomes were considered to achieve the objectives of the study.

The main outcome was attitude toward the technology; it was assessed before (T0) and after (T1) the intervention. This questionnaire, developed by Huygelier et al. [31], aimed at assessing the perceived ease of use, perceived usefulness, and anticipations about VR among older adults and was composed of 18 items to be assessed on a Likert scale from 1 (completely disagree) to 5 (completely agree).

Files stored on the VRSS home tablet were examined to extract objective data regarding patients' adherence. From these files, we extracted the total number of sessions, the total exercise time per day, and the total time spent using ARTEDIA; the following parameters were available for each ARTEDIA exercise: duration, mean RPMs, mean HR, mean SpO_2 (when present), number of correct answers, omissions, and errors.

As for subjective outcomes, the following dimensions were investigated at T1: perceived workload, usability, and technology acceptance. The questionnaires we used were the NASA Task Load Test (NASA-TLX) (considering raw scoring [42,43]), the System Usability Scale (SUS) [44], and the modified version of the Technology Acceptance Measure 3 (TAM3) [45] questionnaires, respectively.

NASA-TLX is a subjective method for measuring mental workload and includes 6 factors measured through a single question that has to be answered on a scale where 0 indicates no demand, and 20 maximum demand.

The TAM3 evaluates the acceptability of the system according to 13 dimensions. Responses to these questions were measured on a 5-point Likert scale, where 1 indicated strongly disagree and 5 strongly agree.

The SUS comprised 10 questions assessed on a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree). The scores are totaled and then multiplied by 2.5 to calculate the final score, which ranges from 0 to 100. Higher scores reflect superior usability.

At the end of the experimental campaign, all the therapists involved in the study in the 3 clinical centers (n=7) were interviewed to provide feedback about the system. The questions of the semi-structured interview are reported in the Supplementary Materials in Multimedia Appendix 1.

Statistical Analysis

Collected data are presented using medians and interquartile ranges; we report mean and standard deviations whenever needed for comparison with previous works. Pre- and post-intervention attitudes toward technology were compared using a paired Wilcoxon rank-sum test. Groups were compared using the Kruskal-Wallis test, and post hoc significance was investigated with Dunn's test. Correlations among subjective variables and between subjective and objective variables were calculated using Spearman's correlations. The reliability of TAM3 subscales was analyzed using Cronbach α. For all tests, the significance level was set at $P<.05$.

Results

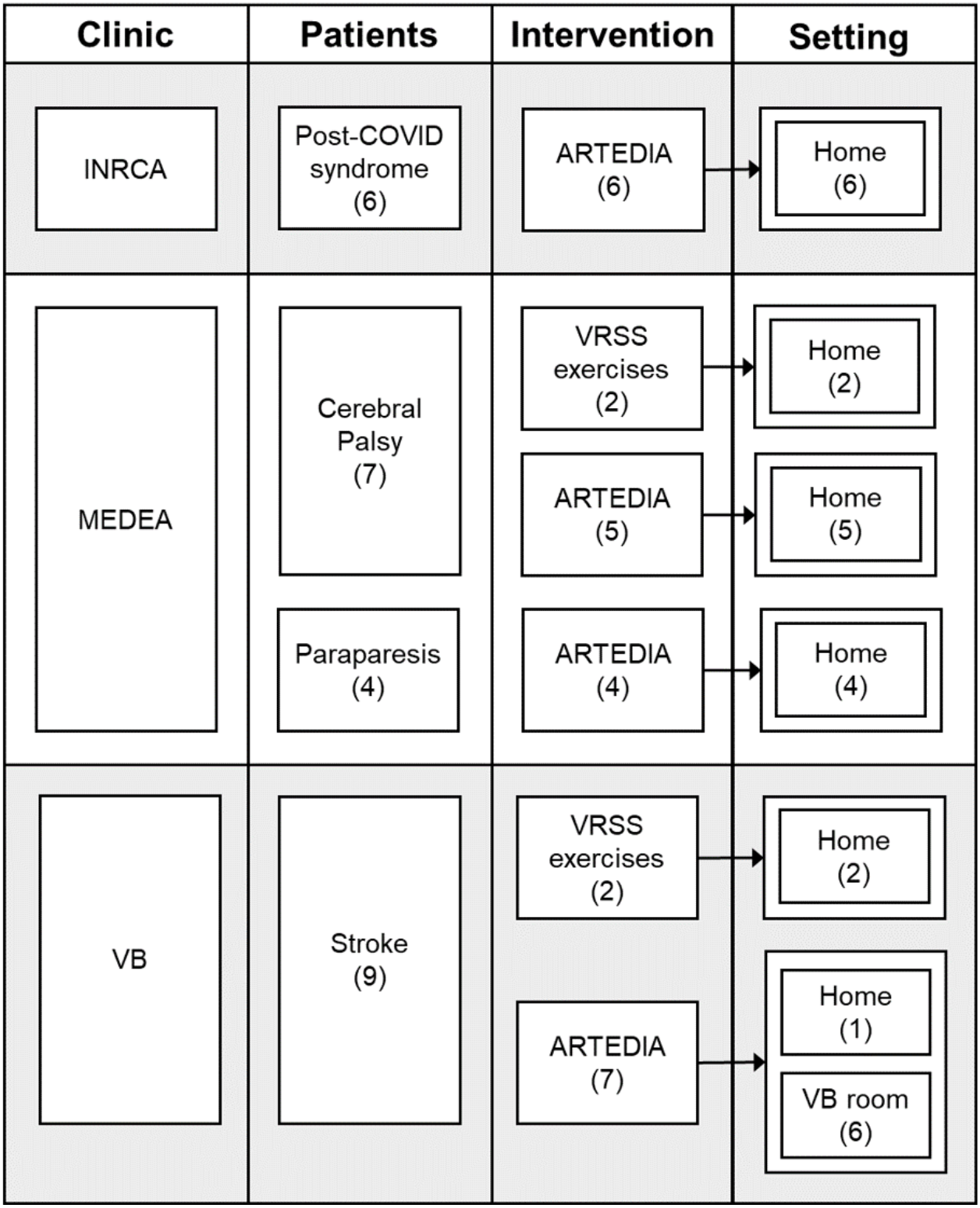
Participant Characteristics

In total, 26 patients were enrolled in the 3 clinical centers. All participants had the maximum score in the Trunk Control Test and no pain, except for one (MEDEA2), who had a score of 2 due to foot pain. The patients' final allocation and their health conditions are shown in Figure 3. Baseline data for each participant is reported in Multimedia Appendices 2-4.

There were 3 dropouts. MEDEA8 (paraparesis) left the study because of the occurrence of pain, whereas VB008 was not able to cycle safely. One of the participants (VB006) had to leave the study after having performed 4 sessions of training due to a call for a surgical procedure, but questionnaire data were collected. All dropouts had the ARTEDIA exercise in their IRP.

XSL•FO
RenderX

Figure 3. Final study allocation and the number of patients involved divided for any of the 3 clinics (INRCA, MEDEA, VB). In between brackets, the number of patients is reported. VRRS exercises indicate the exercises available on the Khymeia Home tablet; ARTEDIA stands for programs containing our newly developed exercise.



Attitude Toward Technology

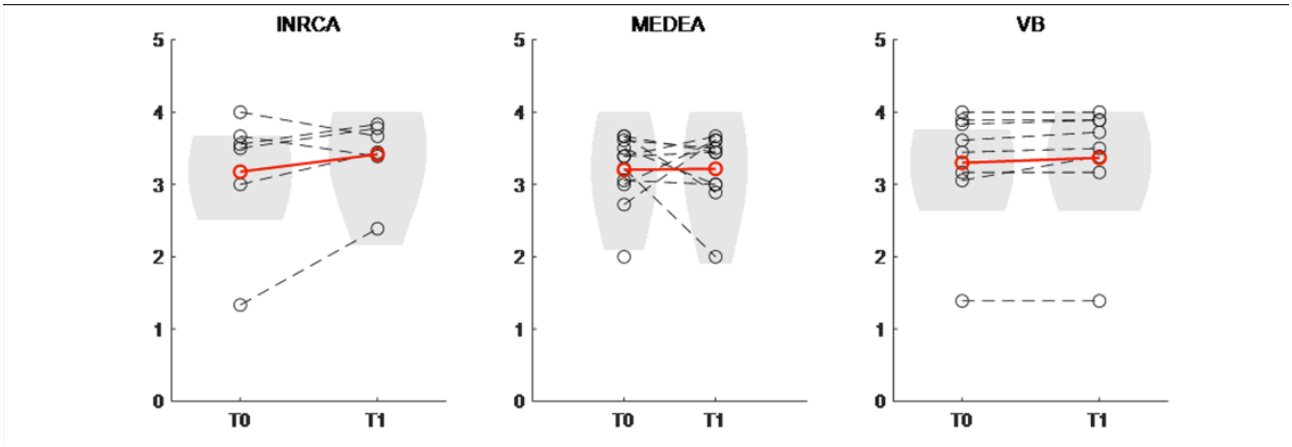
The median attitude toward technology considering all patients was 3.44 (IQR 0.63, mean: 3.23 (SD 0.71)) out of 5 at T0 and 3.5 (IQR 0.48, mean: 3.40 (SD 0.48)) at T1. Figure 4 shows the attitude toward technology in the 3 categories of patients.

The attitude toward technology was positive (above 3) at T0 for most patients; such a value was preserved at T1 in most cases, especially among adults. One patient in the post-COVID group rated their attitude at T1 over 1 point with respect to T0 (INRCA2), and 1 patient with cerebral palsy reached 0.9 (MEDEA2). One young patient (MEDEA7) showed a decrease greater than 1 point at T1. No significant differences were

XSL•FO
RenderX

recorded at the paired Wilcoxon test (P>.05), either considering the whole population of the patients enrolled by each clinic singularly.

Figure 4. Attitude toward technology (range: 0 - 5) in the 3 categories of patients as enrolled by the 3 clinics (INRCA, MEDEA, VB) measured pre- (**T0**) and post-intervention (**T1**). Each dashed gray line represents the observed scores of 1 participant, while the red solid line represents the group average. The gray area represents the density plots of the observed mean attitude scores.



Objective Data

Table 1 contains data regarding the IRP of each patient and his or her hours of actual exercise with the Home Kit, considering both ARTEDIA and VRSS Home Tablet exercises. The adherence, calculated according to the IRP each patient received, was 0.59 (IQR 0.54, min: 0.16, max: 1).

XSL•FO
RenderX

Table . Information about each patient Individual Rehabilitation Project and their adherence to the overall program the therapist prescribed. Data regarding upper limb and cognitive training are presented for the Virtual Reality Rehabilitation System (VRSS); more details are provided for the ARTEDIA exercise.

				ARTEDIA			
ID patient	Hours ARTEDIA	Hours VRSS exercises	Adherence	Total number of sessions	No. days of training	No. sessions/ day	Mean duration/ day (min)
INRCA1	4.79	0	36%	15	14	1.14	21.71
INRCA2	6.54	0	49%	22	18	1.11	20.45
INRCA3	7.92	0	59%	24	18	1.28	75.24
INRCA4	3.77	0	100%	15	11	1.18	18.06
INRCA5	12.60	0	94%	38	36	1.03	20.44
INRCA6	12.85	0	28%	51	25	2.00	30.42
MEDEA1	8.19	11.45	98%	33	17	2.00	29.91
MEDEA2	6.44	11.97	92%	30	17	1.65	20.91
MEDEA3	0.00	12.95	65%	NA[a]	NA	NA	NA
MEDEA4	8.12	13.62	100%	31	16	1.88	29.98
MEDEA5	7.77	10.70	92%	32	17	1.76	26.69
MEDEA6	1.76	2.97	24%	6	4	1.00	18.52
MEDEA7	3.27	6.51	49%	17	7	2.14	26.36
MEDEA9	0.00	14.88	74%	NA	NA	NA	NA
MEDEA10	2.63	3.42	30%	11	10	1.00	15.04
MEDEA11	2.19	2.05	21%	10	5	1.80	25.08
VB001	0	19.67	89%	NA	NA	NA	NA
VB002	0	8.13	37%	NA	NA	NA	NA
VB003[b]	4.54	0	76%	18	9	1.78	26.86
VB004	5.48	0	52%	22	13	1.77	26.30
VB005[b]	0.70	0	16%	4	1	2.00	21.03
VB006[b]	2.87	0	72%	16	10	1.50	15.64
VB007[b]	4.49	0	41%	22	17.00	1.24	15.23
VB009[b]	[c]—	—	—	—	—	—	—

[a]NA: not applicable.

[b]patients who did the training in the clinic-devoted room.

[c]—: missing data.

No differences were present among groups considering total hours of actual training and adherence ratio. Nonetheless, the total hours planned by the therapists were different (H(2)= 8.57, *P*=.013), with patients from MEDEA receiving an IRP foreseeing significantly more hours of exercise than both INRCA (*P*=.04) and VB (*P*=.03). Table 2 displays data regarding the physical and cognitive performance of each patient who performed ARTEDIA (i.e., session with cycle ergometer) at least once.

XSL•FO
RenderX

Table . ARTEDIA performance considering all sessions. The correctness ratio was calculated by dividing the number of correct guesses by the total number of stimuli (targets or distractors).

	Physical performance						Cognitive performance	
ID patient	Mean WL[a] (W)	Mean HR[b] threshold (bpm)	Mean HR (bpm)	Mean SpO_2[c] threshold (%)	Mean SpO_2 (%)	Mean RPM[d]	Level	Correctness ratio
INRCA1	40.67	128.67	51.84	90.00	97.60	63.18	2.87	0.96
INRCA2	42.73	130.32	87.77	90.32	88.87	64.21	2.91	0.91
INRCA3	30.83	130.00	113.73	90.00	94.41	78.01	3.00	0.98
INRCA4	70.00	132.94	98.17	NA[e]	NA	76.48	2.37	0.98
INRCA5	33.68	131.32	80.58	90.00	93.95	65.05	2.26	0.98
INRCA6	30.00	120.00	97.84	90.00	74.80	62.37	3.00	0.91
MEDEA1	54.85	176.97	150.72	NA	NA	68.63	3.00	0.99
MEDEA2	62.33	180.00	123.59	NA	NA	64.47	1.00	0.95
MEDEA4	30.32	180.00	136.49	NA	NA	63.66	2.94	0.95
MEDEA5	59.69	180.00	113.19	NA	NA	60.91	2.66	0.98
MEDEA6	30.00	180.00	115.90	NA	NA	39.69	1.00	0.80
MEDEA7	52.35	180.00	101.25	NA	NA	63.47	1.12	0.92
MEDEA10	30.00	180.00	133.05	NA	NA	61.03	1.91	0.48
MEDEA11	52.00	180.40	112.10	NA	NA	63.41	1.80	0.65
VB003	30.00	120.00	99.76	NA	NA	62.96	1.00	0.99
VB004	30.00	137.27	81.44	NA	NA	48.46	1.00	0.95
VB005	30.00	125.00	112.99	NA	NA	62.16	1.00	1.00
VB006	30.00	132.19	74.43	NA	NA	65.14	1.00	0.94
VB007	30.00	132.73	88.21	NA	NA	75.25	1.00	0.98
VB009	—	—	—	NA	NA	—	—	—

[a]WL: workload.

[b]HR: heart rate.

[c]SpO_2: oxygen saturation.

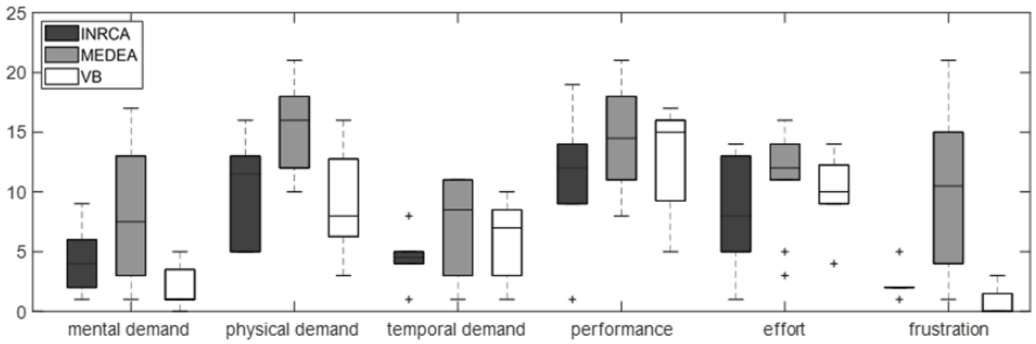
[d]RPM: rounds per minute.

[e]NA: not applicable.

Subjective Data

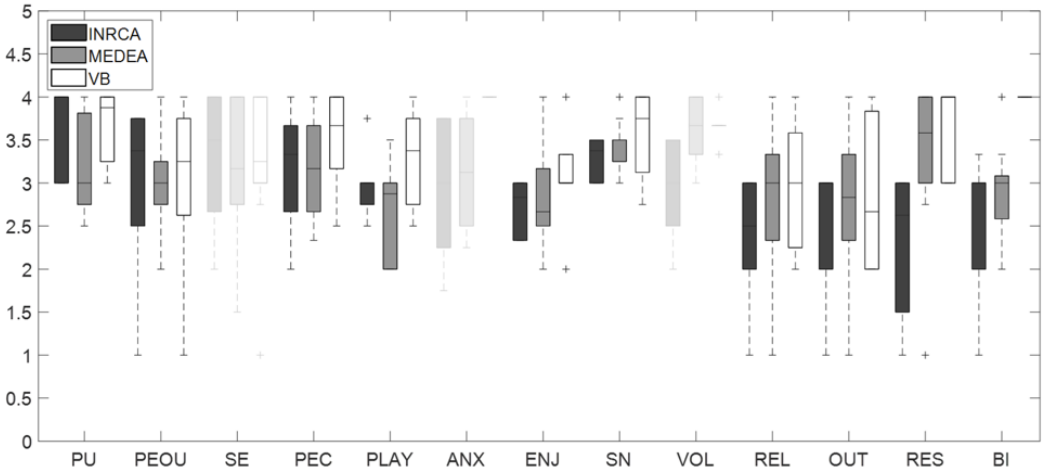
The workload measured at T1 is shown in Figure 5. Usability score was found to be in the “acceptable” and “excellent” ranges according to [46]; its median value was 80 (IQR 20, mean 75.73 (SD 17.45)). Figure 6 and Table 3 depict the correlations among TAM3 subscales and their reliability.

Figure 5. Perceived workload in the 3 populations of patients as enrolled by the 3 clinics (INRCA, MEDEA, VB). The reported items are the subscales measured by the NASA-Task Load IndeX (TLX).



XSL•FO
RenderX

Figure 6. TAM3 subscale values in the 3 populations of patients as enrolled by the 3 clinics (INRCA, MEDEA, VB). Non-reliable scales are grayed out. PU: perceived usefulness; PEOU: perceived ease of use; SE: self-efficacy; PEC: perception of external control; PLAY: playfulness; ANX: anxiety; ENJ: enjoyment; SN: subjective norm; VOL: voluntariness; REL: relevance; OUT: output quality; RES: result demonstrability; BI: behavioral intention.



XSL•FO
RenderX

Table . Correlations among TAM3 subscales and their reliability.

Variable	PU[a]	PEOU[b]	SE[c]	PEC[d]	PLAY[e]	ANX[f]	ENJ[g]	SN[h]	VOL[i]	REL[j]	OUT[k]	RES[l]	BI[m]	α
PU														
r_s		0.41	0.33	0.34	0.35	0.16	0.35	0.07	0.23	0.36	0.31	0.08	0.25	0.97
P value		.07	.15	.14	.13	.51	.13	.78	.32	.12	.18	.72	.30	
PEOU														
r_s			0.72	0.62	0.19	0.60	0.31	-0.24	0.34	0.22	0.36	0.51	0.47	0.87
P value			<.001	<.001	.42	<.001	.19	.31	.14	.34	.12	.02	.04	
SE														
r_s				0.69	0.28	0.48	0.29	-0.04	0.36	0.38	0.42	0.42	0.49	0.56
P value				<.001	.24	.03	.21	.85	.12	.10	.06	.07	.03	
PEC														
r_s				-	0.35	0.51	0.38	0.11	0.58	0.28	0.53	0.63	0.34	0.62
P value					.13	.02	.09	.65	.01	.24	.02	<.001	.14	
PLAY														
r_s						0.13	0.64	0.28	0.18	0.59	0.62	0.43	0.37	0.70
P value						.60	<.001	.24	.45	.01	<.001	.06	.11	
ANX														
r_s							0.45	0.18	0.67	0.31	0.22	0.64	0.75	0.51
P value							.05	.46	<.001	.18	.34	<.001	<.001	
ENJ														
r_s								0.10	0.49	0.65	0.70	0.50	0.50	0.86
P value								.67	.03	<.001	<.001	.02	.03	
SN														
r_s								-	0.25	0.34	0.06	0.18	0.54	0.64
P value									.29	.15	.81	.44	.01	
VOL														
r_s									-	0.32	0.43	0.62	0.45	0.25
P value										.17	.06	<.001	.05	
REL														
r_s										-	0.78	0.50	0.62	0.87
P value											<.001	.03	<.001	
OUT														
r_s											-	0.59	0.37	0.91
P value												.01	.11	

XSL•FO
RenderX

Variable	PU[a]	PEOU[b]	SE[c]	PEC[d]	PLAY[e]	ANX[f]	ENJ[g]	SN[h]	VOL[i]	REL[j]	OUT[k]	RES[l]	BI[m]	α
RES														
r_s												-	0.59	0.91
P value													.01	
BI														
r_s													-	0.95
P value														

[a]PU: perceived usefulness.

[b]PEOU: perceived ease of use.

[c]SE: self-efficacy.

[d]PEC: perception of external control.

[e]PLAY: playfulness.

[f]ANX: anxiety.

[g]ENJ: enjoyment.

[h]SN: subjective norm.

[i]VOL: voluntariness.

[j]REL: relevance.

[k]OUT: output quality.

[l]RES: result demonstrability.

[m]BI: behavioral intention.

In terms of subjective outcomes, group differences emerged in terms of perceived cognitive workload, in particular in the subscales of physical demand (H(2)=8.00, *P*=.018) and frustration (H(2)=11.36, *P*=.035); MEDEA reached higher levels than VB for both variables (*P*=0.03and *P*=.002, respectively). Significant differences among groups were also present in the case of technology acceptance, in particular for the subscales of result demonstrability (H(2)=13.13, *P*=.001) and behavioral intention (H(2)=11.76, *P*=.003). Post-hoc tests showed that VB results were more positive than both MEDEA and INRCA (*P*<.007 in all cases). No other significant group differences were present.

Mean attitude toward technology at T1 was negatively correlated with physical demand (r_s=−0.48, *P*=.018), effort (r_s=−0.41, *P*=.049), and frustration (r_s=−0.52, *P*=.01) as measured by the NASA-TLX. Mean attitude at T0 was significantly correlated with frustration only (r_s=−0.51, *P*=.01).

Mean attitude (T1) and technology acceptance subscale correlations were significant (*P*<.05) in the case of perceived usefulness (r_s=0.52), perceived ease of use (r_s=0.55), playfulness (r_s=0.50), relevance (r_s=0.44), and behavioral intention (r_s=0.50); the correlation of attitude with usability was significant (r_s=0.68, *P*<.001).

None of the subjective variables was correlated to total hours of training or adherence ratio.

Considering those who did ARTEDIA, the difference between the HR threshold and the mean HR measured throughout the sessions was correlated only with frustration (r_s=0.55, *P*=.017). The correctness ratio (i.e., the number of targets identified correctly divided by the total targets) correlated significantly and negatively with mental demand and frustration (r_s=−0.62, *P*=.007 and r_s=−0.53, *P*=.023, respectively). No other correlation was significant.

Therapists' Feedback

The results of the semi-structured interview are presented in Table 4.

XSL•FO
RenderX

Table . The therapists' feedback grouped according to themes and positive and negative comments. When the comments were related to one participant in particular, their ID is reported between brackets.

	Pros	Cons
Perceived usefulness	• Implementation of continuity of care • Home-based treatment favored compliance • Possibility of lower limb training • Patients' monitoring throughout time	• Exercises are generic and not well-targeted • No easy assessment of the performance quality (i.e., if compensation was used) • Impossible to establish the reason behind pain occurrence (MEDEA7)
Accessibility	• Accessibility for chronic patients not receiving other care • Accessibility for patients living in remote areas	• Patients must fit precise conditions (high cognitive functioning, no spasticity) • Patients must be available to schedule calls with the therapist during working hours (MEDEA11) • The safety thresholds were not suitable for one patient (INRCA4) • Internet connection was not always stable enough • Architectural barriers impeded the installation of the setup (stairs, lack of space)
Usability	• Good autonomous management in most cases • Scheduling daily exercises promoted adherence • Engaging scenario promoting motivation	• Technical issues • Difficulty in providing support in case of malfunctioning

As per technical issues, we have tracked the following occurrences in patients from MEDEA: the sensors were not responding properly or were sending wrong input data (n=1; MEDEA7); server updates did not allow the therapist to update the IRP and the patient to access the daily training (n=2; MEDEA4, MEDEA5); generic malfunctioning of the system (n=1, MEDEA6). For one post-COVID patient (INRCA4), the minimum threshold set for oxygen saturation was still too high to proceed with the training. In this case, the patient replaced the sensor with another that was not linked to the system and monitored themselves according to the therapist's indications. Insufficient power of the internet connection and architectural barriers were mentioned in all the interviews.

Discussion

Principal Findings

The results of this study indicated the feasibility and acceptability of a customized VR-based telerehabilitation program that allowed for the safe implementation of lower limb and aerobic training. This result aligns with the outcomes in previous works studying telerehabilitation interventions, either with or without VR support. Furthermore, conducting this study allowed the provision of rehabilitation services to patients who had limited access to the clinic due to geographical distance or had ended their hospitalization.

The study was performed by enrolling a heterogeneous sample of patients, comprising post-COVID patients and adults, adolescents, and children with neuromotor impairments. It allowed us to demonstrate the feasibility of our proposed customizable telerehabilitation program to address diverse issues and needs. Although heterogeneity may constitute a limitation, it can be welcomed when scientific (e.g., pragmatic) studies are aimed to demonstrate the effectiveness in a population as similar as possible to the target population who may benefit from the intervention [47]. Moreover, it could help design the future delivery of targeted care, accounting for patients' diverse conditions [48].

Regarding the attitude toward technology, contrary to what was found in [31], in which Huygelier et al. recorded an increase from 3.4 to 3.9, we did not find any improvement after the treatment. This outcome may be explained by considering that the original work used immersive VR, whereas our system used more familiar technologies, such as laptops and smartphones. Therefore, it is possible that our participants did not change their minds due to the lack of a "wow" effect [49]; also, their anticipation and perception toward the use of digital technology were probably correct at the baseline and confirmed at the end of the experience [50]. Nonetheless, the patients' relationship and approach to technology did not change, and no technical factors affected technology adoption and caused participants to leave the program. This was also confirmed by usability and acceptance outcomes, which were both satisfactory (all technology acceptance subscale scores were >2.7, up to 5), and the positive correlations, found at T1, between attitude and usability, and attitude and perceived ease of use, playfulness, and intentions to reuse the system.

On the other hand, as mentioned by the therapists in their interview, our sample was selected to fit the requirements of mostly autonomous system use (i.e., cognitive level and presence of a caregiver), thus introducing a possible bias in this sense. Indeed, literature shows that the general acceptance of mHealth applications still encounters patients' resistance in more than half of the cases, resulting in minimal engagement and no clinical improvements [51,52].

XSL•FO
RenderX

Our patients' adherence to the proposed program was moderate-to-good (59%), in line with other works proposing unsupervised physical activity programs [53,54] or rehabilitation activities for patients [55]. Other interventions indeed recorded higher adherence, but they also accounted for a stricter sample selection: e.g., a tablet-guided exercise program for older adults obtained 85% adherence, but only 36% of patients accepted to participate in the program and completed it [56].

This issue has already been discussed in the literature, especially when dealing with programs promoting physical exercise: participants in scientific studies are volunteers and, therefore, inherently motivated to exercise or willing to be involved. This means that the results of trials engaging patients in exercise are mostly not generalizable to the general population.

Previous work tried to identify the factors that may influence adherence; a review conducted in 2021 identified 14 key factors affecting adherence in rehabilitation programs [57]; among the others, there was the involvement of professionals from several disciplines, exploration of patients' barriers and facilitators, participants' education, integration in daily living and social support, which we did not investigate. Among older adults, factors promoting an active lifestyle were also identified as personal factors: younger age, education, history of regular exercise, less depression, fewer impairments, and chronic conditions could all positively influence adherence to training programs [58,59].

This dependence upon personal factors was possibly present also in our case, as individual differences, not detected by our collected data, were present. Only in a few cases, analyzing all the information together, was it possible to formulate a hypothesis for higher or lower adherence. For instance, MEDEA7 encountered some problems with the sensors and experienced footache, and the therapist was not able to investigate the cause of the pain remotely. This has impacted their attitude toward technology (showing a decrease of over 1 point from T0 to T1) and possibly their adherence (49%). In the case of INRCA6, MEDEA6, MEDEA 10, and VB002, it is plausible that their initial condition—worse than other peers'—negatively impacted adherence, as they had to make more effort to participate in the program.

These considerations are to be considered cautiously: they are hypotheses, and in most cases, both the adherence and the total training hours were not directly attributable to any of the variables investigated. Because of this, we encourage future work to focus more on personal and anamnestic characteristics we have not explored.

Another element that could contribute to assessing the acceptability of an intervention and influencing patients' adherence is workload, i.e., the cost incurred by an individual while achieving a certain performance on a task with specific demands [60]. The concept of workload per se is hard to interpret, as it is impossible for many tasks to define a "redline" above which the performance degrades [61].

Indeed, in our case, it is plausible to think that some mental effort would be required to perform the proposed exercises per se, even without any influence of the technological system. Considering the performance data related to ARTEDIA dual-task exercise, it emerged that lower cognitive results were indeed linked to higher cognitive demands.

Instead, both cognitive and physical performances were negatively correlated with frustration. In the NASA-TLX questionnaire, the frustration item foresees the investigation of "how insecure, discouraged, irritated, and annoyed" one feels while accomplishing a task, i.e., it examines the unsuccessful fulfillment of objectives after repetitive attempts due to obstacles and negative feedback [62]. Also in this case, it is impossible to discriminate precisely if the frustration was due to the task itself (either cycling or go/no-go task) or the system management. However, some considerations could be made by examining the significant differences among groups: patients from MEDEA reported more frustration and physical effort and were also the group reporting the higher occurrence of problems, such as VRRS updates that caused temporary unavailability of the system, sensor malfunctioning, and the impossibility of the therapist to assess the cause of pain remotely. The relationship between frustration and technical issues has emerged before [63,64].

Another consideration that could be made is related to the absolute dose of therapy: the group of younger participants from MEDEA was also the one receiving a more intense training plan. Although no correlations were identified (possibly because of the small sample), this outcome was in agreement with previous studies, which highlighted that lower-dose interventions could fit better the patients' schedules and, therefore, be more accepted [65,66]. This was highlighted by the case of MEDEA11, who was mostly unavailable for the video calls with the therapists and had a low adherence rate, possibly because of other duties of daily living (21%). Given this, the topic of the appropriate dose to promote adherence in telerehabilitation deserves to be further explored in future works.

On the contrary, VB patients were less subjected to problems, possibly because the system was regularly checked and sanitized (as it was shared and close to the clinic, with one exception only), and—as mentioned before—they trained significantly less.

Despite the reduced adherence, due possibly to the need to travel to reach the room close to the clinic, the hybrid solution has been appreciated and, therefore, deserves further analysis, especially in the case of non-intensive training, e.g, for maintenance of muscle fitness and promotion of an active lifestyle among chronic patients.

In such a setting, sharing a single piece of equipment saves personnel time, money, and equipment management duties (e.g., handling transfers, installations, and conformity checks at the end of the training). Plus, environmental barriers such as lack of space and adequate internet connection are excluded, and prompt support can be provided in the case of technical or clinical issues. On the other hand, the accessibility to the service remains available only for those patients who can travel and reach the facility.

XSL•FO
RenderX

Limitations

This work has some limitations. First of all, the sample was heterogeneous, and although this choice allowed testing our telerehabilitation program with diverse conditions, this limited the generalization to the populations of adults and children with neuromuscular or post-COVID-related impairments. Second, the adherence to the intervention was very different, and we could not determine which factors influenced it either positively or negatively. Finally, rehabilitation sessions were carried out in two different scenarios (home and clinic-devoted room), and study participants were part of different age groups; this may have influenced the patients' subjective experiences. Although the in-depth evaluation of differences among subgroups goes beyond the scope of this work, it would have been interesting to investigate the effects of other factors, such as age or setting. Unfortunately, this was not possible due to the small sample. Nonetheless, the preliminary assessment of the feasibility of the proposed intervention paves the way for future trials in which more participants per group should be enrolled.

Conclusions

This work presents a feasibility study of a novel telerehabilitation system allowing cycling-based dual-task training. The system was preliminarily assessed with a sample of patients with different conditions and demonstrated usable, meaningful, and acceptable. Moderate acceptance was found by therapists who recognized several strengths but also a number of issues. For the effective implementation of a complete telerehabilitation program, including lower limbs, some improvements are still required to overcome environmental and technical barriers.

In the future, larger and more homogeneous populations should be engaged to extend the obtained results and increase their generalizability. A higher number of participants will also allow us to delve deeper into subgroup differences, highlighting which features are considered more relevant by each patient or age group to enhance personalization even more. Particular attention should be paid to the therapy dose and the rehabilitation setting to find the correct balance between treatment effectiveness and suitability with patients' daily life in order to promote optimal compliance.

Acknowledgments

The authors would like to acknowledge Khymeia Group for their support in carrying out this project and the integration of the cycle ergometer within their VRRS system. Moreover, the authors would like to acknowledge the support of Dr. Tito Dinon, who designed and realized the 3D-printed supports.

The work was supported by Fondazione Cariplo (ARTEDIA grant). It was partially supported by the Italian Ministry of Health (Ricerca Corrente 2024/2025 to EB).

Data Availability

The data associated with the paper are not publicly available. Data is available, in anonymized form, from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Questionnaire for the therapists.
[DOCX File, 13 KB - humanfactors_v12i1e71099_app1.docx]

Multimedia Appendix 2
Characteristics of INRCA participants at the baseline; 6MWT: 6 minutes walking test in meters and velocity; mMRC: modified British Medical Research Council Questionnaire; S: index of symmetry; Pr/Pl: mean cycling power of right/left leg.
[DOCX File, 14 KB - humanfactors_v12i1e71099_app2.docx]

Multimedia Appendix 3
Characteristics of MEDEA participants at the baseline; *: drop-out; 6MWT: 6 minutes walking test in meters; QUEST: Quality of Upper Extremity Skills Test, only administered in participants with upper limb impairments; GMFM: Gross Motor Function Measure, administered only for children below 18 years old; S: index of symmetry; Pr/Pl: mean cycling power of right/left leg. NA: Not Applicable; /: missing value.
[DOCX File, 15 KB - humanfactors_v12i1e71099_app3.docx]

Multimedia Appendix 4

XSL•FO
RenderX

Characteristics of VB participants at the baseline; *: drop out; S: index of symmetry; Pr/Pl: mean cycling power of right/left leg. NA: Not Applicable; /: missing value.
[DOCX File, 14 KB - humanfactors_v12i1e71099_app4.docx]

References

1. Kocyigit BF, Assylbek MI, Yessirkepov M. Telerehabilitation: lessons from the COVID-19 pandemic and future perspectives. Rheumatol Int 2024 Apr;44(4):577-582. [doi: 10.1007/s00296-024-05537-0] [Medline: 38321330]
2. Brigo E, Rintala A, Kossi O, et al. Using telehealth to guarantee the continuity of rehabilitation during the COVID-19 pandemic: a systematic review. Int J Environ Res Public Health 2022 Aug 19;19(16):10325. [doi: 10.3390/ijerph191610325] [Medline: 36011959]
3. Seron P, Oliveros MJ, Gutierrez-Arias R, et al. Effectiveness of telerehabilitation in physical therapy: a rapid overview. Phys Ther 2021 Jun 1;101(6):pzab053. [doi: 10.1093/ptj/pzab053] [Medline: 33561280]
4. Suso-Martí L, La Touche R, Herranz-Gómez A, Angulo-Díaz-Parreño S, Paris-Alemany A, Cuenca-Martínez F. Effectiveness of telerehabilitation in physical therapist practice: an umbrella and mapping review with meta-meta-analysis. Phys Ther 2021 May 4;101(5):pzab075. [doi: 10.1093/ptj/pzab075] [Medline: 33611598]
5. Simmich J, Ross MH, Russell T. Real-time video telerehabilitation shows comparable satisfaction and similar or better attendance and adherence compared with in-person physiotherapy: a systematic review. J Physiother 2024 Jul;70(3):181-192. [doi: 10.1016/j.jphys.2024.06.001] [Medline: 38879432]
6. Paul L, Thomson K, Asibey SO, et al. Views of service users, their family or carers, and health care professionals on telerehabilitation for people with neurological conditions in Ghana: qualitative study. JMIR Mhealth Uhealth 2024 Mar 27;12:e49501. [doi: 10.2196/49501] [Medline: 38536218]
7. Alt Murphy M, Munoz-Novoa M, Heremans C, et al. European Stroke Organisation (ESO) guideline on motor rehabilitation. Eur Stroke J 2025 May 22:23969873251338142. [doi: 10.1177/23969873251338142] [Medline: 40401760]
8. Thanakamchokchai J, Khobkhun F, Phetsitong R, et al. Effectiveness of telerehabilitation on the International Classification of Functioning, Disability, and Health framework outcomes during the COVID-19 pandemic: a systematic review and meta-analysis of randomized controlled trials. Digit Health 2025;11:20552076251325993. [doi: 10.1177/20552076251325993] [Medline: 40162161]
9. Harkey LC, Jung SM, Newton ER, Patterson A. Patient satisfaction with telehealth in rural settings: a systematic review. Int J Telerehab 2020;12(2):53-64. [doi: 10.5195/ijt.2020.6303]
10. Timoteo EF, Silva DF, Oliveira TD, José A, Malaguti C. Real-time telerehabilitation for chronic respiratory disease and post-COVID-19: a systematic review and meta-analysis. J Telemed Telecare 2025 Aug;31(7):991-1004. [doi: 10.1177/1357633X241241572] [Medline: 38594927]
11. Bezzini D, Schiavetti I, Manacorda T, Franzone G, Battaglia MA. First wave of COVID-19 pandemic in Italy: data and evidence. Adv Exp Med Biol 2021;1353:91-113. [doi: 10.1007/978-3-030-85113-2_6] [Medline: 35137370]
12. Luca RD, Rifici C, Terranova A, et al. Healthcare worker burnout during the first COVID-19 lockdown in Italy: experiences from an intensive neurological rehabilitation unit. J Int Med Res 2023 Jul;51(7):3000605231182664. [doi: 10.1177/03000605231182664] [Medline: 37486238]
13. Rispoli V, Díaz Crescitelli ME, Cavallieri F, et al. Needs and perceptions of patients with dystonia during the COVID-19 pandemic: a qualitative framework analysis of survey responses from Italy. Front Neurol 2022;13:808433. [doi: 10.3389/fneur.2022.808433] [Medline: 35785354]
14. Rizzo A, Thomas Koenig S, Talbot TB. Clinical results using virtual reality. J Technol Hum Serv 2019 Jan 2;37(1):51-74. [doi: 10.1080/15228835.2019.1604292]
15. Hao J, Pu Y, Chen Z, Siu KC. Effects of virtual reality-based telerehabilitation for stroke patients: a systematic review and meta-analysis of randomized controlled trials. J Stroke Cerebrovasc Dis 2023 Mar;32(3):106960. [doi: 10.1016/j.jstrokecerebrovasdis.2022.106960] [Medline: 36586244]
16. Catania V, Rundo F, Panerai S, Ferri R. Virtual reality for the rehabilitation of acquired cognitive disorders: a narrative review. Bioengineering (Basel) 2024;11(1):35. [doi: 10.3390/bioengineering11010035]
17. Aderinto N, Olatunji G, Abdulbasit MO, Edun M, Aboderin G, Egbunu E. Exploring the efficacy of virtual reality-based rehabilitation in stroke: a narrative review of current evidence. Ann Med 2023 Dec 12;55(2). [doi: 10.1080/07853890.2023.2285907]
18. Huang W, Roscoe RD, Johnson‐Glenberg MC, Craig SD. Motivation, engagement, and performance across multiple virtual reality sessions and levels of immersion. Comput Assist Learn 2021 Jun;37(3):745-758 [FREE Full text] [doi: 10.1111/jcal.12520]
19. Howard MC. A meta-analysis and systematic literature review of virtual reality rehabilitation programs. Comput Human Behav 2017 May;70:317-327. [doi: 10.1016/j.chb.2017.01.013]
20. Jansson MM, Rantala A, Miettunen J, Puhto AP, Pikkarainen M. The effects and safety of telerehabilitation in patients with lower-limb joint replacement: A systematic review and narrative synthesis. J Telemed Telecare 2022 Feb;28(2):96-114. [doi: 10.1177/1357633X20917868] [Medline: 32316837]

XSL•FO
RenderX

21. Nikolaev VA, Nikolaev AA. Recent trends in telerehabilitation of stroke patients: a narrative review. NeuroRehabilitation 2022;51(1):1-22. [doi: 10.3233/NRE-210330] [Medline: 35527574]
22. Barbara C, Clavario P, De Marzo V, et al. Effects of exercise rehabilitation in patients with long coronavirus disease 2019. Eur J Prev Cardiol 2022 May 25;29(7):e258-e260. [doi: 10.1093/eurjpc/zwac019] [Medline: 35078233]
23. López LP, Coll-Andreu M, Torras-Garcia M, et al. Aerobic exercise and cognitive function in chronic severe traumatic brain injury survivors: a within-subject A-B-A intervention study. BMC Sports Sci Med Rehabil 2024 Sep 27;16(1):201. [doi: 10.1186/s13102-024-00993-4] [Medline: 39334396]
24. Wirth B, Ferreira TD, Mittelholzer M, Humphreys BK, Boutellier U. Respiratory muscle endurance training reduces chronic neck pain: a pilot study. J Back Musculoskelet Rehabil 2016 Nov 21;29(4):825-834. [doi: 10.3233/BMR-160695] [Medline: 27002668]
25. Park S, Tang A, Pollock C, Sakakibara BM. Telerehabilitation for lower extremity recovery poststroke: a systematic review and meta-analysis protocol. BMJ Open 2022 Mar 9;12(3):e055527. [doi: 10.1136/bmjopen-2021-055527] [Medline: 35264359]
26. Colombo V, Mondellini M, Tauro G, et al. Rehabilitation of post-COVID patients: a virtual reality home-based intervention including cardio-respiratory fitness training. Presented at: Extended Reality: First International Conference, XR Salento 2022; Jul 6, 2022; Lecce, Italy p. 3-17. [doi: 10.1007/978-3-031-15546-8_1]
27. Mrakic-Sposta S, Di Santo SG, Franchini F, et al. Effects of combined physical and cognitive virtual reality-based training on cognitive impairment and oxidative stress in MCI patients: a pilot study. Front Aging Neurosci 2018;10:282. [doi: 10.3389/fnagi.2018.00282] [Medline: 30327596]
28. Colombo V, Mondellini M, Gandolfo A, et al. Usability and acceptability of a virtual reality-based system for endurance training in elderly with chronic respiratory diseases. In: Virtual Reality and Augmented Reality 2019:87-96. [doi: 10.1007/978-3-030-31908-3_6]
29. Ambrosini E, Ferrante S, Pedrocchi A, Ferrigno G, Guanziroli E, Molteni F. A novel biofeedback cycling training to improve gait symmetry in stroke patients: a case series study. Presented at: 2011 IEEE International Conference on Rehabilitation Robotics; Jul 29-1, 2011; Zurich, Switzerland.
30. Fascio E, Vitale JA, Sirtori P, Peretti G, Banfi G, Mangiavini L. Early virtual-reality-based home rehabilitation after total hip arthroplasty: a randomized controlled trial. J Clin Med 2022 Mar 22;11(7):1766. [doi: 10.3390/jcm11071766] [Medline: 35407373]
31. Huygelier H, Schraepen B, van Ee R, Vanden Abeele V, Gillebert CR. Acceptance of immersive head-mounted virtual reality in older adults. Sci Rep 2019 Mar 14;9(1):4519. [doi: 10.1038/s41598-019-41200-6] [Medline: 30872760]
32. Collin C, Wade D. Assessing motor impairment after stroke: a pilot reliability study. J Neurol Neurosurg Psychiatry 1990 Jul;53(7):576-579. [doi: 10.1136/jnnp.53.7.576] [Medline: 2391521]
33. Carlsson AM. Assessment of chronic pain. I. Aspects of the reliability and validity of the visual analogue scale. Pain 1983 May;16(1):87-101. [doi: 10.1016/0304-3959(83)90088-X] [Medline: 6602967]
34. Magni E, Binetti G, Bianchetti A, Rozzini R, Trabucchi M. Mini-Mental State Examination: a normative study in Italian elderly population. Eur J Neurol 1996 May;3(3):198-202. [doi: 10.1111/j.1468-1331.1996.tb00423.x] [Medline: 21284770]
35. Enright PL. The six-minute walk test. Respir Care 2003 Aug;48(8):783-785. [Medline: 12890299]
36. Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. 2018 URL: https://goldcopd.org/wp-content/uploads/2017/11/GOLD-2018-v6.0-FINAL-revised-20-Nov_WMS.pdf [accessed 2025-08-19]
37. DeMatteo C, Law M, Russell D, Pollock N, Rosenbaum P, Walter S. The reliability and validity of the quality of upper extremity skills test. Phys & Occupational Therapy In Pediatrics 1993 Mar 1;13(2):1-18. [doi: 10.1300/J006v13n02_01]
38. Russell DJ, Rosenbaum PL, Cadman DT, Gowland C, Hardy S, Jarvis S. The gross motor function measure: a means to evaluate the effects of physical therapy. Dev Med Child Neurol 1989 Jun;31(3):341-352. [doi: 10.1111/j.1469-8749.1989.tb04003.x] [Medline: 2753238]
39. Demeurisse G, Demol O, Robaye E. Motor evaluation in vascular hemiplegia. Eur Neurol 1980;19(6):382-389. [doi: 10.1159/000115178] [Medline: 7439211]
40. Blum L, Korner-Bitensky N. Usefulness of the Berg Balance Scale in stroke rehabilitation: a systematic review. Phys Ther 2008 May;88(5):559-566. [doi: 10.2522/ptj.20070205] [Medline: 18292215]
41. Ambrosini E, Parati M, Peri E, et al. Changes in leg cycling muscle synergies after training augmented by functional electrical stimulation in subacute stroke survivors: a pilot study. J Neuroeng Rehabil 2020 Feb 27;17(1):35. [doi: 10.1186/s12984-020-00662-w] [Medline: 32106874]
42. Hart SG. NASA-task load index (NASA-TLX); 20 years later. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2006 Oct;50(9):904-908. [doi: 10.1177/154193120605000909]
43. Byers JC, Bittner AC, Hill SG. Traditional and raw task load index (TLX) correlations: are paired comparisons necessary? Adv Ind Ergon Saf 1989:481-485.
44. Brooke J. A “quick and dirty” usability scale. In: Usability Evaluation in Industry 1996:189-194.
45. Venkatesh V, Bala H. Technology acceptance model 3 and a research agenda on interventions. Decision Sciences 2008 May;39(2):273-315 [FREE Full text] [doi: 10.1111/j.1540-5915.2008.00192.x]

XSL•FO
RenderX

46. Bangor A, Kortum P, Miller P. Determining what individual SUS scores mean: Adding an adjective rating scale. Journal of usability studies 2009;4:114-123 [FREE Full text]
47. Giraudeau B, Caille A, Eldridge SM, Weijer C, Zwarenstein M, Taljaard M. Heterogeneity in pragmatic randomised trials: sources and management. BMC Med 2022 Oct 28;20(1):372. [doi: 10.1186/s12916-022-02569-w] [Medline: 36303153]
48. Cotton RJ, Segal Rick RL, Seamon BA, et al. Precision rehabilitation: optimizing function, adding value to health care. Arch Phys Med Rehabil 2022 Sep;103(9):1883-1884. [doi: 10.1016/j.apmr.2022.04.017] [Medline: 35690092]
49. Mondellini M, Arlati S, Gapeyeva H, et al. User experience during an immersive virtual reality-based cognitive task: a comparison between Estonian and Italian older adults with MCI. Sensors (Basel) 2022 Oct 27;22(21):8249. [doi: 10.3390/s22218249] [Medline: 36365947]
50. Busch PA, Hausvik GI, Ropstad OK, Pettersen D. Smartphone usage among older adults. Comput Human Behav 2021 Aug;121:106783. [doi: 10.1016/j.chb.2021.106783]
51. Jacob C, Sezgin E, Sanchez-Vazquez A, Ivory C. Sociotechnical factors affecting patients' adoption of mobile health tools: systematic literature review and narrative synthesis. JMIR Mhealth Uhealth 2022 May 5;10(5):e36284. [doi: 10.2196/36284] [Medline: 35318189]
52. Pfammatter AF, Mitsos A, Wang S, Hood SH, Spring B. Evaluating and improving recruitment and retention in an mHealth clinical trial: an example of iterating methods during a trial. Mhealth 2017;3:49. [doi: 10.21037/mhealth.2017.09.02] [Medline: 29184901]
53. Cox KL, Burke V, Gorely TJ, Beilin LJ, Puddey IB. Controlled comparison of retention and adherence in home- vs center-initiated exercise interventions in women ages 40–65 Years: The S.W.E.A.T. study (Sedentary Women Exercise Adherence Trial). Prev Med 2003 Jan;36(1):17-29. [doi: 10.1006/pmed.2002.1134] [Medline: 12473421]
54. Perri MG, Anton SD, Durning PE, et al. Adherence to exercise prescriptions: effects of prescribing moderate versus higher levels of intensity and frequency. Health Psychol 2002 Sep;21(5):452-458. [doi: 10.1037/0278-6133.21.5.452] [Medline: 12211512]
55. Kilbride C, Scott DJM, Butcher T, et al. Safety, feasibility, acceptability and preliminary effects of the Neurofenix platform for Rehabilitation via HOMe based gaming exercise for the upper-limb post Stroke (RHOMBUS): results of a feasibility intervention study. BMJ Open 2022 Feb 28;12(2):e052555. [doi: 10.1136/bmjopen-2021-052555] [Medline: 35228279]
56. Mosca IE, Salvadori E, Gerli F, et al. Analysis of feasibility, adherence, and appreciation of a newly developed tele-rehabilitation program for people with MCI and VCI. Front Neurol 2020;11:583368. [doi: 10.3389/fneur.2020.583368] [Medline: 33329326]
57. Collado-Mateo D, Lavín-Pérez AM, Peñacoba C, et al. Key factors associated with adherence to physical exercise in patients with chronic diseases and older adults: an umbrella review. Int J Environ Res Public Health 2021 Feb 19;18(4):1-24. [doi: 10.3390/ijerph18042023] [Medline: 33669679]
58. Neumann A, König HH, Hajek A. Determinants of patient use of telemental health services: representative cross-sectional survey from Germany. JMIR Ment Health 2025 Jun 13;12:e70925. [doi: 10.2196/70925] [Medline: 40513010]
59. Patel SB, Iqbal FM, Lam K, Acharya A, Ashrafian H, Darzi A. Characterizing Behaviors That Influence the Implementation of Digital-Based Interventions in Health Care: Systematic Review. J Med Internet Res 2025 Jun 12;27:e56711. [doi: 10.2196/56711] [Medline: 40505130]
60. Weidemann A, Russwinkel N. Investigation of frustration. 2019 Sep 8 Presented at: MuC'19; Sep 8-11, 2019; Hamburg Germany p. 819-824 URL: https://dl.acm.org/doi/proceedings/10.1145/3340764 [doi: 10.1145/3340764.3344915]
61. Grier RA. How high is high? A meta-analysis of NASA-TLX global workload scores. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2015 Sep;59(1):1727-1731. [doi: 10.1177/1541931215591373]
62. Papini MR, Ludvigson HW, Amsel A. Frustration theory: an analysis of dispositional learning and memory. Am J Psychol 1994;107(4):604. [doi: 10.2307/1423003]
63. Lange-Drenth L, Schulz H, Suck I, Bleich C. Barriers, facilitators, and requirements for a telerehabilitation aftercare program for patients after occupational injuries: semistructured interviews with key stakeholders. JMIR Form Res 2023;8:e51865. [doi: 10.2196/51865]
64. Chen Y, Chen Y, Zheng K, et al. A qualitative study on user acceptance of a home-based stroke telerehabilitation system. Top Stroke Rehabil 2020 Mar;27(2):81-92. [doi: 10.1080/10749357.2019.1683792] [Medline: 31682789]
65. Li I, Bui T, Phan HT, Llado A, King C, Scrivener K. App-based supplemental exercise in rehabilitation, adherence, and effect on outcomes: a randomized controlled trial. Clin Rehabil 2020 Aug;34(8):1083-1093. [doi: 10.1177/0269215520928119] [Medline: 32508183]
66. Reginster JY, Rabenda V, Neuprez A. Adherence, patient preference and dosing frequency: understanding the relationship. Bone 2006 Apr;38(4 Suppl 1):S2-S6. [doi: 10.1016/j.bone.2006.01.150] [Medline: 16520104]

Abbreviations

ANX: anxiety
BI: behavioral intention
ENJ: enjoyment

XSL•FO
RenderX

HR: heart rate
INRCA: IRCCS INRCA Casatenovo
IRP: Individual Rehabilitation Project
MEDEA: IRCCS E. Medea
NASA-TLX: Nasa Task Load Index
OUT: output quality
PEC: perception of external control
PEOU: perceived ease of use
PLAY: playfulness
PU: perceived usefulness
REL: relevance
RES: result demonstrability
RPM: revolutions per minute
SE: self-efficacy
SN: subjective norm
SUS: System Usability Scale
TAM3: Technology Acceptance Measure 3
VB: Villa Beretta Rehabilitation Center
VOL: voluntariness
VR: virtual reality
VRRS: Virtual Reality Rehabilitation System

Edited by A Kushniruk; submitted 10.01.25; peer-reviewed by E Peri, S Lopes; revised version received 16.06.25; accepted 26.06.25; published 10.09.25.

Please cite as:
Arlati S, Colombo V, Mondellini M, Nossa R, Grasso C, Rossini M, Biffi E, Fumagalli A, Diella E, Guanziroli E, Sanna N, Ambrosini E, Sacco M, Molteni F
Cycling-Based Telerehabilitation: Acceptability and Feasibility Study
JMIR Hum Factors 2025;12:e71099
URL: https://humanfactors.jmir.org/2025/1/e71099
doi:10.2196/71099

© Sara Arlati, Vera Colombo, Marta Mondellini, Roberta Nossa, Chiara Grasso, Mauro Rossini, Emilia Biffi, Alessia Fumagalli, Eleonora Diella, Eleonora Guanziroli, Nicole Sanna, Emilia Ambrosini, Marco Sacco, Franco Molteni. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 10.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Immersive Virtual Reality eHealth Intervention to Reduce Anxiety and Depression in Pregnant Women: Randomized Controlled Trial

Marta Jimenez-Barragan[1], BSc, MA, PhD; Amparo Del Pino Gutierrez[1], BSc, MA, PhD; Gloria Sauch Valmaña[2], BSc, MA, PhD; Olga Monistrol[3], RN, MSc, PhD; Carme Monge Marcet[4], BSc; Mar Pallarols Badia[4], MBBS, MA; Ignasi Garrido[5], BSc, MA, PhD; Anna Carmona Ruiz[6], MBBS, MA; Oriol Porta Roda[7], MBBS, MA, PhD; Cristina Esquinas[1*], BSc, MA, PhD; Gemma Falguera Puig[8*], BSc, MA, PhD

[1]Department of Public Health, Mental Health and Maternal and Child Health Nursing, Faculty of Nursing, Universitat de Barcelona, Barcelona, Spain

[2]Health Promotion in Rural Areas Research Group, Gerència Territorial de la Catalunya Central, Institut Català de la Salut, Institut Universitari per la Recerca a l'Atenció Primària de Salut Jordi Gol i Gurina (IDIAPJGol), Institut Català de la Salut, Barcelona, Spain

[3]Departament of Nursing and Physiotherapy, Campus Igualada, Universitat de Lleida, Lleida, Spain

[4]Department of Gynecology and Obstetrics, Assistential Foundation Mútua Terrassa, Mútua Terrassa, Terrassa, Spain

[5]Psychiatry and Mental Health Service. Eating Disorder Unit, Mútua Terrassa, Terrassa, Spain

[6]Department of Obstetrics and Gynaecology, Fundacio Assitencial Mollet, Mollet, Spain

[7]Department of Gynecology and Obstetrics, Mútua Terrassa, Terrassa, Spain

[8]Primary Care Group, Germans Trias i Pujol Research Institute (IGTP), Research Group Atenció a la Salut Sexual i Reproductiva (GRASSIR), Institut Universitari d'Investigació en Atenció Primària Jordi Gol (IDIAP Jordi Gol), Institut Català de la Salut, Barcelona, Spain

[*]these authors contributed equally

Corresponding Author:
Marta Jimenez-Barragan, BSc, MA, PhD
Department of Public Health, Mental Health and Maternal and Child Health Nursing
Faculty of Nursing
Universitat de Barcelona
Feixa Llarga s/n 08907 Hospitalet de Llobregat, Barcelona, Spain
Barcelona, 08907
Spain
Phone: 34 934024256
Email: marta.jim@gmail.com

Abstract

Background: Mental health during pregnancy is a critical factor influencing maternal and fetal outcomes. Anxiety and depression affect up to 30% of pregnant women, with significant consequences for maternal well-being and child development. Despite this, interventions during pregnancy remain limited, creating a need for innovative, accessible solutions.

Objective: This study aimed to evaluate the effectiveness of an immersive virtual reality (IVR) eHealth intervention in reducing anxiety and depression symptoms in women during pregnancy.

Methods: A 2-arm, randomized controlled trial was conducted across 5 primary care centers in Catalonia, Spain, between October 2021 and May 2024. The study included pregnant women (N=70) aged ≥18 years with moderate anxiety and depression symptoms (Edinburgh Postnatal Depression Scale [EPDS] scores: 9-12) at 12 to 14 weeks of gestation. They were randomly assigned (1:1) to an IVR intervention or standard care group. The intervention group engaged in daily 14-minute IVR mindfulness and relaxation sessions for 6 weeks. Mental health outcomes were assessed using the EPDS and State-Trait Anxiety Inventory.

Results: The intervention group demonstrated significant reductions in EPDS scores, with a mean decrease from 11.32 (SD 0.96) to 7.25 (SD 1.32; *P*<.001), compared to an increase in the control group from 11.32 (SD 0.94) to 16.23 (SD 1.25; *P*<.001). Similarly, State-Trait Anxiety Inventory scores improved markedly in the intervention group (mean decrease from 57.94, SD 5.23 to 35.03, SD 6.12; coefficient –30.47, 95% CI −45.23 to −15.72; *P*<.001), while the control group experienced a nonsignificant increase (from 66.10, SD 5.89 to 72.91, SD 6.34; *P*=.68). High adherence rates were observed, with 79% (26/33) of participants completing ≥30 sessions. Participant satisfaction was high, with 87% (29/33) reporting being "very satisfied" with the intervention.

XSL•FO
RenderX

Conclusions: The IVR eHealth intervention significantly reduced symptoms of anxiety and depression, demonstrating its potential as an accessible and effective tool for mental health support during pregnancy. High adherence and satisfaction levels underscore its feasibility and acceptability. Future research should explore the long-term effects and scalability of IVR interventions in diverse settings.

Trial Registration: ClinicalTrials.gov NCT05756205; https://clinicaltrials.gov/study/NCT05756205

International Registered Report Identifier (IRRID): RR2-10.1186/s12912-023-01440-4

(JMIR Hum Factors 2025;12:e71708) doi:10.2196/71708

KEYWORDS

virtual reality; eHealth; pregnancy; mental health; anxiety; depression; randomized controlled trial; antenatal care

Introduction

Background

Mental health during pregnancy is a crucial aspect of women's overall well-being, influencing not only the mother but also the development of the fetus and the subsequent outcomes for the child [1,2]. Up to 30% of women may experience mental health disorders during pregnancy, including anxiety and depression [3-5]. These disorders, often interrelated, have a significant impact not only on the mother's health but also on the neurological development of the fetus, potentially leading to complications, such as preterm birth, low birth weight, and long-term emotional difficulties [6]. These adverse effects not only affect the immediate postpartum period but can also influence the child's long-term development, with possible repercussions on behavior, cognition, and mental health [7].

Despite their importance, interventions to address mental health disorders during pregnancy have been limited, with many focusing on the postpartum period, leaving a critical need for attention during pregnancy itself [8]. This deficit in intervention during pregnancy is concerning, as early detection and appropriate treatment of these disorders are essential to ensuring optimal health outcomes for both the mother and the baby [9]. Inadequate treatment or lack of interventions may increase the risk of obstetric complications, such as preterm birth or low birth weight, as well as long-term emotional and behavioral problems in the child [10,11].

Low-intensity interventions, as described by the National Institute for Health and Care Excellence, are particularly important for women experiencing mild to moderate symptoms of anxiety and depression. The National Institute for Health and Care Excellence recommends these interventions as a first step before considering more intensive treatments, as they can provide effective support with fewer adverse effects and are accessible to a wide population [12]. In this context, digital health technologies, known as eHealth, offer a promising new avenue for improving access to care and supporting the mental health of pregnant women [13].

The use of eHealth technologies is growing exponentially worldwide, supported by guidelines from international organizations, such as the World Health Organization, which promote the use of these technologies to improve public health [14]. According to World Health Organization, eHealth technologies can play a key role in expanding access to mental health services [15]. Their application in perinatal mental health has shown promising results in preliminary studies, especially in reducing symptoms of anxiety and depression [16]. These technologies offer the possibility of providing low-intensity interventions, such as mindfulness, relaxation, and other stress management techniques, in a way that is accessible and convenient for women [17]. In addition, mobile apps allow for closer and more personalized monitoring of women's mental health, offering continuous support and enabling quicker interventions in the event of concerning symptoms [18]. eHealth interventions can be a flexible alternative for those women who, for various reasons, cannot easily access traditional mental health services, offering solutions tailored to their individual needs and daily routines [19].

In this context, immersive virtual reality (IVR) is emerging as an innovative tool in the health field. IVR allows users to immerse themselves in simulated environments in a fully immersive manner, offering therapeutic experiences that can be highly effective in managing anxiety, depression, and other mental disorders [20]. Several studies have demonstrated that IVR can help reduce stress and improve emotional well-being in patients, providing cognitive and emotional distraction that facilitates relaxation and symptom reduction [21]. Moreover, its ability to recreate controlled environments makes it useful for exposure therapy and other psychological interventions [22]. This technology, combined with other eHealth tools, can offer personalized and highly effective solutions for pregnant women experiencing mental health issues, opening new possibilities for their management and treatment [23].

Furthermore, the integration of eHealth and mobile health technologies is gaining traction as an effective means of delivering mental health interventions. These technologies provide a direct, low-cost, and engaging manner to deliver care, making mental health support more accessible [24,25]. The global increase in the use of these technologies has led to a growing body of research supporting their efficacy, particularly in treating mood disorders during pregnancy [26,27]. Recent studies have also underscored the importance of mobile health interventions specifically designed for pregnant women, highlighting their effectiveness in reducing anxiety and depression symptoms during pregnancy [5,28].

Objectives

The objective of this study was to evaluate the effectiveness of an IVR eHealth intervention in improving mental health outcomes during pregnancy, specifically by comparing symptoms of anxiety and depression between pregnant women

XSL•FO
RenderX

in the intervention and control groups, assessing satisfaction with antenatal care between the 2 groups, and monitoring adherence to the intervention by tracking the number of completed IVR sessions.

Methods

Study Design

This study was a 2-arm, prospective, randomized controlled trial (RCT) designed to assess the efficacy of a low-intensity eHealth intervention in improving mental well-being among pregnant women. The CONSORT (Consolidated Standards of Reporting Trials) guidelines for RCTs were followed. Participants were randomly assigned (1:1) to the intervention or control group using a computer-generated randomization sequence. Specifically, the Epidat (version 4.2; Epidemiology Service of the Dirección Xeral de Saúde Pública, Xunta de Galicia) software was used to generate the randomization list, ensuring allocation concealment and minimizing selection bias. To maintain the integrity of the randomization process, the allocation sequence was generated independently by the principal investigator (MJ-B) before recruitment. Randomization was stratified by recruitment center to control potential site-specific variations. Allocation was concealed from participants and researchers responsible for data collection to reduce potential biases. Participants and health care providers were masked to group allocation to the extent possible in an intervention of this nature.

The study included pregnant women aged ≥18 years who attended primary care centers affiliated with the sexual and reproductive health care (*Atenció a la Salut Sexual i Reproductiva* [ASSIR; Sexual and Reproductive Health Care in English]). It was conducted across 5 primary care centers (ASSIR) in Catalonia, Spain, between October 2021 and May 2024. Between 12 and 14 weeks of gestation, participants who scored between 9 and 12 on the Edinburgh Postnatal Depression Scale (EPDS)—validated Spanish version [29]—were recruited and randomly allocated to either the intervention group, which received the eHealth intervention, or the control group, which received standard care. Participants were also screened for mental health conditions using the State-Trait Anxiety Inventory (STAI)—validated Spanish version [30]. Those with a diagnosis of severe psychiatric disorder, ongoing treatment by mental health specialists, and a history of gender-based violence were excluded.

Procedure

Participants who scored between 9 and 12 on the EPDS, indicating symptoms of moderate anxiety and depression, were considered for inclusion in the study. Participants were randomly assigned (1:1) to the intervention or control condition. The principal investigator generated a randomization list using the Epidat (version 4.2) software, and all study participants were masked ensuring the randomization process. Women assigned to the control group were informed that their follow-up would proceed as usual, while those in the intervention group received additional information about the eHealth intervention, including the use of IVR.

Intervention

The intervention group received an eHealth-based mental health program, which included the use of IVR technology to deliver mindfulness and relaxation exercises designed to reduce anxiety and depression during pregnancy. The IVR intervention lasted 14 minutes per session and was performed daily for 6 weeks. These sessions were based on mindfulness techniques, conscious breathing, and progressive muscle relaxation, with the aim of providing an immersive therapeutic environment that would promote stress reduction and improve the emotional well-being of the participants. The content of the intervention was developed by Delaguila Games [31], a company specializing in virtual reality applied to health care (Figure 1). The sessions were designed to be self-guided and accessible from home (Figure 2), using an Oculus Go (Meta Reality Labs) virtual reality headset. The program contained 3 main modules that could be selected individually or together, adapting to the needs and preferences of each participant. To ensure adherence and effectiveness of the intervention, the total number of sessions completed by each participant was recorded.

XSL•FO
RenderX

Figure 1. Screenshots of the virtual reality experience.



Figure 2. A pregnant woman using the immersive virtual reality application at home (with written patient permission).



Control

Participants in the control group received the usual antenatal care provided by the public health system. This refers to the pregnancy follow-up protocol established by the Department of Health of the Generalitat de Catalunya, which outlines standard guidelines for prenatal care and monitoring [32].

Sample

Pregnant women aged ≥18 years who were followed at the primary care centers of sexual and reproductive health care (ASSIR) of Mútua Terrassa in Barcelona, Spain, were invited to participate in the study.

In a previous published interventional work [33] using a mindfulness program to evaluate changes in the EPDS score in pregnant women, the authors reported a mean difference of −2.56 (SE 0.72) between the 2 groups of study (intervention and control) after the program. From that, the SD was calculated as SE=SD/Root(n), so SD=SE × Root(n). In this case, as the n was different in both groups (51 and 45), we used the minimum of both (n=45), obtaining a joint SD of 4.83. Assuming a power of 80% to detect differences in contrast of the null hypothesis H_0: $\mu_1=\mu_2$ using a bilateral Student *t* test for 2 independent samples, a significance level of 5%, and assuming a mean difference between both groups of −3.5 (SD 4.83 units), it was necessary to include 31 women in the nonintervention group

XSL•FO
RenderX

and 31 women in the intervention group, totaling 66 patients in the study. Considering that the expected drop-out rate is 5%, it was necessary to recruit 70 women (35 in each group).

Measures

Primary Outcome: Mental Health Status

The mental health status of the participants was measured using the EPDS and the STAI. The EPDS is a self-report scale comprising 10 items, where participants choose 1 of 4 responses that best describes how they felt during the past week. The EPDS has been validated in Spanish and shows good sensitivity (79%) and specificity (95.5%), with a Cronbach α of 0.87 [29]. Scores range from 0 to 30, with scores ≥13 indicating a positive screen for depression. The STAI assesses 2 dimensions of anxiety: state anxiety (a temporary condition in response to a specific situation) and trait anxiety (a general tendency to perceive situations as threatening) [30]. Both scales have demonstrated robust psychometric properties in various populations.

Secondary Outcomes

Satisfaction With Pregnancy Care

Satisfaction with pregnancy care was measured using a validated self-completed questionnaire [34], consisting of 28 items that assess different aspects of care, including equipment, accessibility, organization of the consultation, and the competence of the staff. The scale uses a 5-point Likert system, and scores are categorized as satisfied (scores >3) or very satisfied (scores >4). The reliability of this tool has been demonstrated with a Cronbach α of 0.92 [35].

Adherence to the Intervention Protocol

Adherence to the intervention protocol was monitored by tracking the number of sessions completed by each participant in the intervention group. In addition, participants in this group were asked to complete a questionnaire evaluating the usability and accessibility of the eHealth technology used during the study [36]. The number of sessions and time spent on each session were recorded to assess adherence and participant engagement with the intervention.

Pulse Measurements

Pulses were recorded before and after the intervention.

Participants were encouraged to report any adverse events, discomfort, or unexpected side effects experienced during or after the IVR sessions. If the participant experienced dizziness, nausea, or an increase in anxiety symptoms, she was instructed to stop the session and rest. If symptoms persisted, participants could contact the research team via an email address specially designed to resolve any problems. Participants had the right to withdraw from the study at any time, without having to justify their decision, and without this affecting their regular prenatal follow-up. In addition, the study team, made up of midwives and mental health professionals, was available to provide additional psychological support if necessary, ensuring that participants felt safe and supported throughout the intervention.

Data Collection and Management

Data were collected at 2 time points: baseline (12-14 weeks of gestation) and 6 weeks after the start of the intervention. All data were securely stored and anonymized, with access restricted to authorized personnel only. Participants' personal information was protected according to the ethical standards of Mútua Terrassa.

Statistical Analysis

Qualitative variables were described using absolute frequencies and percentages. Quantitative variables were summarized using the mean, SD, and median. The Kolmogorov-Smirnov test was used to assess the normality of distributions.

Sociodemographic and clinical characteristics were compared across groups. For quantitative variables, Mann-Whitney *U* tests were conducted. The chi-square test (or Fisher exact test for frequencies <5) was used to compare categorical variables. The effectiveness of the intervention was evaluated using mixed-effects models to assess changes in mental health outcomes over time by group. The interaction between time and intervention, as well as potential confounding factors such as age, previous admissions, current planned pregnancy, and personal history of anxiety, were also considered in these analyses.

In addition, the postintervention scores of the EPDS and STAI-E were transformed into binary variables. For EPDS, a cutoff score of 9 was selected based on bibliographic criteria [29], while for STAI-E, the third quartile value (>86 points) was used [30]. Finally, several logistic regression models were conducted for the primary outcomes with the study group included as an independent variable. These models were also adjusted for the previously mentioned confounding factors, including preintervention EPDS and STAI-E scores. Hosmer-Lemeshow goodness-of-fit tests were performed to assess the overall fit of the models [37]. Receiver operating characteristic curves were also calculated for this approach.

Statistical significance was set at $P<.05$, and all analyses were performed using R version 4.2.1 (R Foundation for Statistical Computing). For the mixed-effects models, we used the *nlme* package with the *lme* function and the *stats* package with the *glm* function in the logistic models.

Ethical Considerations

The study protocol was approved by the Drug Research Ethics Committee of Mútua Terrassa (B1803). Informed consent was obtained from all participants before their inclusion in the study. The study was conducted in accordance with the Declaration of Helsinki and relevant European Union regulations on data protection. The informed consent process was conducted following strict ethical guidelines to ensure that all participants fully understood the nature of the study and its potential risks and benefits, and that their participation was voluntary. The steps taken for the informed consent process are provided in Textbox 1.

These measures were taken to guarantee that all participants provided truly informed and voluntary consent, minimizing any risk of coercion or undue influence.

XSL•FO
RenderX

Textbox 1. Informed consent process.

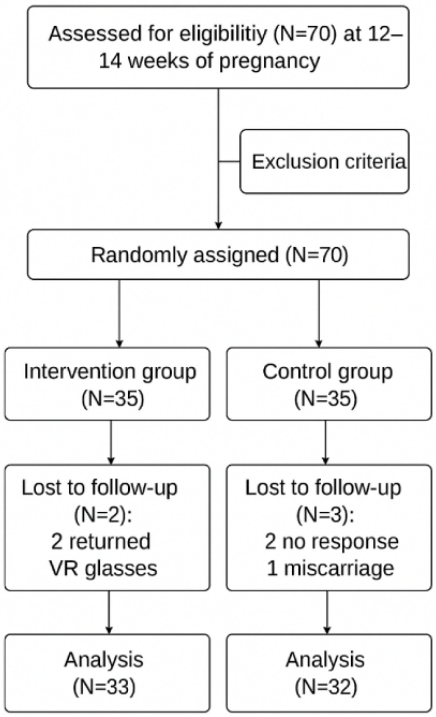
- A comprehensive explanation: before the enrollment, all participants received a detailed explanation of the study's objectives, procedures, potential benefits, and possible risks. This was done through both written and verbal communication by trained health care professionals.
- Time for reflection and questions: participants were given sufficient time to read the informed consent document and ask any questions. Researchers provided clarifications to ensure that participants had a clear understanding before signing the consent form.
- Assessment of understanding: to ensure comprehension, participants were asked to summarize key aspects of the study in their own words. If any misunderstandings were identified, additional explanations were provided.
- Voluntary participation and withdrawal rights: it was explicitly stated that participation was entirely voluntary, and that participants could withdraw at any time without any consequences for their health care or treatment.
- Safeguards against coercion: to prevent undue influence, recruitment was conducted by independent health care professionals who were not involved in participants' routine medical care. No financial or material incentives were provided to encourage participation.
- Ethics approval: the study protocol, including the informed consent procedures, was reviewed and approved by the Drug Research Ethics Committee of Mútua Terrassa (B1803), ensuring compliance with ethical and legal standards.

Results

Overview

The study included 70 participants with a mean age of 31.9 (SD 4.83) years, evenly distributed between the intervention group (n=35, 50%) and the control group (n=35, 50%; Figure 3). There were 2 losses to follow-up in the intervention group and 3 in the control group. All cases of loss to follow-up occurred in the period after the intervention.

Figure 3. CONSORT (Consolidated Standards of Reporting Trials) participant flowchart for the randomized controlled trial. VR: virtual reality.



Sociodemographic and Clinical Characteristics

Regarding sociodemographic and clinical characteristics, the intervention group had a higher mean age (in years) compared with the control group (33.77, SD 3.46 vs 30.03, SD 5.3; *P*=.003). Furthermore, a significantly higher proportion of participants in the intervention group reported having their own income (35/35, 100% vs 21/35, 60%; *P*<.001) and a university-level education (24/35, 69% vs 8/35, 23%; *P*<.001;Table 1).

XSL•FO
RenderX

Table 1. Comparison of sociodemographic and clinical characteristics between the control and intervention groups (N=70).

Characteristic	Total (N=70)	Control group (n=35)	Intervention group (n=35)	*P* value
Age (y), mean (SD)	31.9 (4.83)	30.03 (5.3)	33.77 (3.46)	.003[a]
Country of origin, n (%)				
Spain	50 (71)	17 (49)	33 (94)	<.001[b]
Morocco	10 (14)	10 (29)	0 (0)	—[c]
Other countries	10 (14)	8 (23)	2 (6)	—
Religious affiliation, n (%)				
Christian	21 (30)	9 (26)	12 (34)	.001[b]
Muslim	10 (14)	10 (29)	0 (0)	—
Atheist	36 (51)	15 (43)	21 (60)	—
Other	3 (4)	1 (3)	2 (6)	—
Educational level, n (%)				
University	32 (46)	8 (23)	24 (69)	<.001[b]
Secondary	30 (43)	19 (54)	11 (31)	—
Primary	6 (9)	6 (17)	0 (0)	—
Employment status , n (%)				
Specialized worker	37 (53)	11 (31)	26 (74)	<.001[b]
Seeking employment	12 (17)	7 (20)	5 (14)	—
Other occupations	21 (30)	17 (49)	4 (11)	—
Personal income, n (%)				
Yes	56 (80)	21 (60)	35 (100)	<.001[b]
No	14 (20)	14 (40)	0 (0)	—

[a]*P* value results comparing the study groups by Mann-Whitney *U* test are considered statistically significant when <.05.

[b]*P* value results comparing the study groups by chi-square tests are considered statistically significant when <.05.

[c]Not available.

Regarding prepregnancy mental health treatment history, participants who reported treatment history had received a variety of interventions, including individual or group psychotherapy, pharmacological treatment with antidepressants or anxiolytics, and follow-up by mental health professionals in primary or specialist health care. The length of treatment varied among participants, ranging from brief interventions <6 months to continued treatment >2 years. Disaggregated data about specific treatment adherence or the exact type of medication used by each participant were not available in this study.

Regarding obstetric characteristics, 86% (30/35) of pregnancies in the intervention group were planned, compared with 57% (20/35) in the control group (*P*=.02; Table 2). No significant differences were found between the groups regarding the history of cesarean sections, traumatic deliveries, or high-risk pregnancies (all *P*>.05).

XSL•FO
RenderX

Table 2. Clinical and obstetric characteristics.

Category and characteristic	Total (N=70), n (%)	Control group (n=35), n (%)	Intervention group (n=35), n (%)	*P* value[a]
Clinical				
Current smoker	10 (14)	7 (20)	3 (9)	.31
Mental health treatment before pregnancy	40 (57)	15 (43)	25 (71)	.03
Family history of psychiatric disorders	22 (31)	12 (34)	10 (29)	.80
Planned current pregnancy	50 (71)	20 (57)	30 (86)	.02
Current pregnancy achieved through assisted reproduction	7 (10)	1 (3)	6 (17)	.11
History of anxiety	47 (67)	18 (51)	29 (83)	.01
History of depression	18 (26)	8 (23)	10 (29)	.78
Obstetric[b]				
Previous miscarriage	18 (26)	7 (20)	11 (31)	.41
Previous cesarean section	4 (6)	2 (6)	2 (6)	.99
Previous traumatic childbirth experience	11 (16)	5 (14)	6 (17)	.99
High-risk current pregnancy	33 (47)	15 (43)	18 (51)	.63
Partner involvement in care	68 (97)	33 (94)	35 (100)	.50
Previous fetal loss	0 (0)	0 (0)	0 (0)	.99
Method of conception: IVF[c]	7 (10)	1 (3)	6 (17)	.11
Prepregnancy mental health diagnosis	20 (29)	8 (23)	12 (34)	.03

[a]*P* value results comparing the study groups by chi-square test are considered statistically significant when <.05.

[b]Gestational week of all participants at enrollment was 12 to 14 weeks.

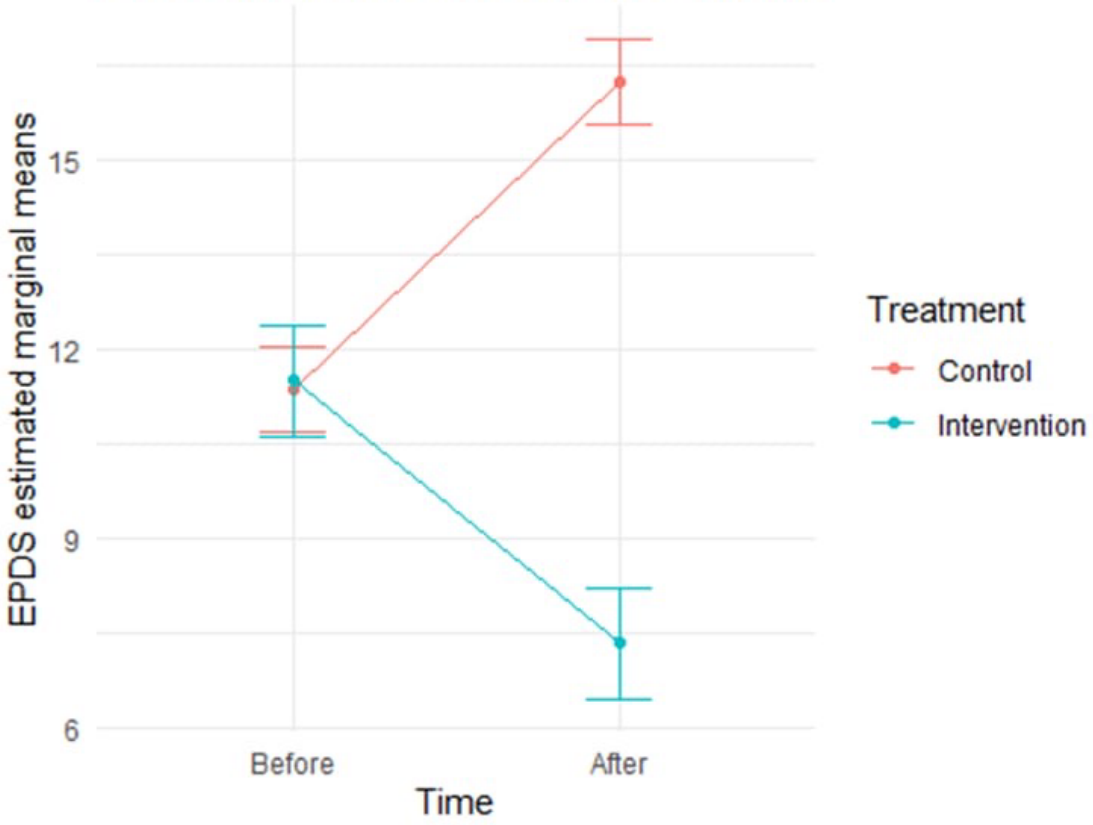
[c]IVF: in vitro fertilization.

Comparison of EPDS Scores Between Groups Before and After the Intervention

At baseline, the EPDS score was comparable between both groups (*P*=.56). Following the intervention, the mean EPDS score significantly decreased in the intervention group (from 11.32 to 7.25; *P*<.001), whereas it increased in the control group (from 11.32 to 16.23; *P*<.001). The adjusted mixed-effects model revealed a significant interaction between time and treatment (coefficient=−9.03, 95% CI −11.29 to −6.76; *P*<.001), indicating a substantial reduction in EPDS scores attributable to the intervention over time (Figure 4).

XSL•FO
RenderX

Figure 4. Interaction between time and treatment in the Edinburgh Postnatal Depression Scale (EPDS) score.



Furthermore, the adjusted logistic regression model demonstrated that participants in the intervention group had a significantly lower likelihood of exhibiting postintervention EPDS scores ≥9 points compared to the control group (odds ratio [OR] 0.03, 95% CI 0.01-0.8; *P*=.001; Table 3). In addition, a planned pregnancy was associated with a lower probability of achieving EPDS scores ≥9 points. In contrast, a higher preintervention EPDS score was associated with an increased likelihood of achieving EPDS scores ≥9 points (OR 3.32, 95% CI 1.52-9.58; *P*=.008). The model demonstrated good calibration, as evidenced by the Hosmer-Lemeshow goodness-of-fit test (*P*=.10) and the area under the receiver operating characteristic curve (area under curve=0.82).

Table 3. Multivariate logistic regression model for postintervention Edinburgh Postnatal Depression Scale (EPDS) score (>9 points).

Variable	Multivariate[a] odds ratio (95% CI)	*P* value[b]
Intervention group	0.03 (0-0.18)	.001[b]
Preintervention EPDS score	3.32 (1.52-9.58)	.008
Age (y)	0.99 (0.8-1.12)	.93
Personal income	2.2 (0.12-41.26)	.58
Current planned pregnancy	0.1 (0.01-0.65)	.03
Personal history of anxiety	4.48 (0.66-50.4)	.16

[a]Hosmer–Lemeshow goodness-of-fit test, *P*=.10. Area under the receiver operating characteristic curve=0.82 (95% CI 0.78-0.85).

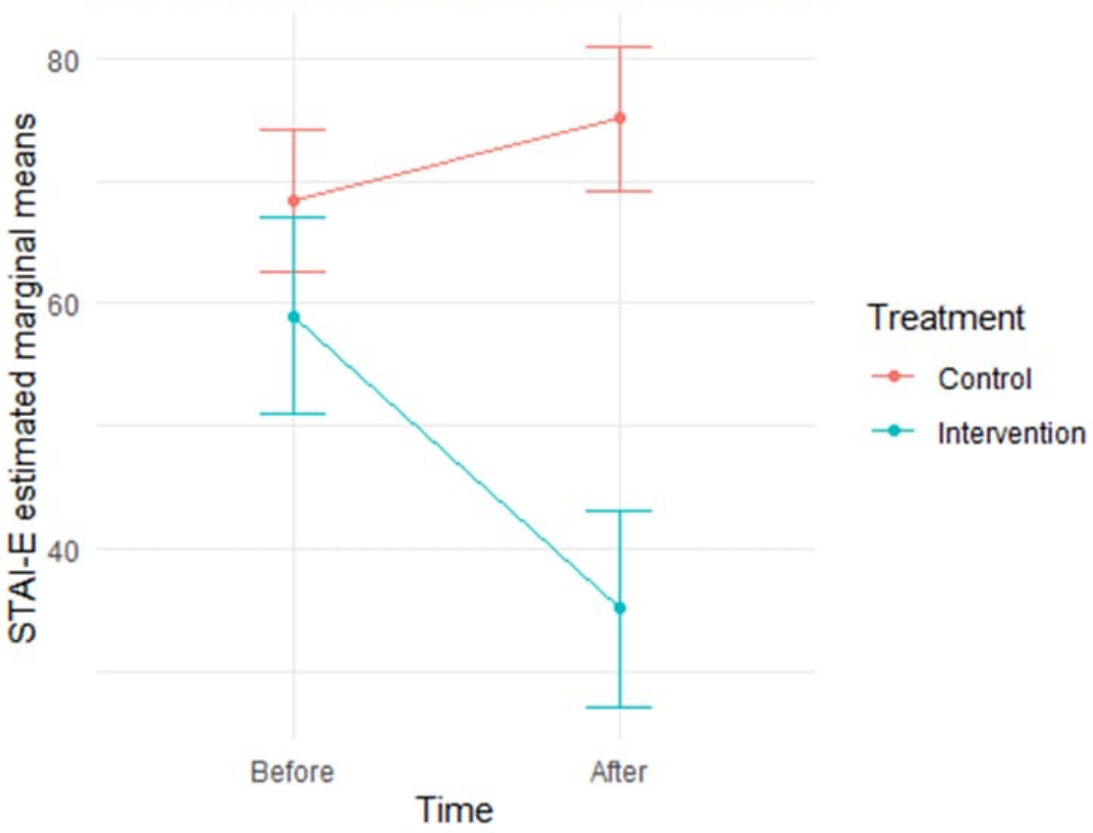
[b]*P* value is considered statistically significant when <.05.

Comparison of STAI-E Scores Before and After the Intervention

At baseline, the STAI-E score was similar in both groups (*P*=.12). Following the intervention, the intervention group experienced a significant reduction in mean scores (from 57.94 to 35.03; *P*<.001), whereas the control group showed an increase (from 66.10 to 72.91; *P*<.001). The adjusted mixed-effects model identified a significant interaction between time and treatment (coefficient=−30.47, 95% CI −45.23 to −15.72; *P*<.001; Figure 5), indicating a substantial improvement attributable to the intervention over time, similar to the effects observed with the EPDS. Conversely, the control group exhibited a nonsignificant change, with mean scores increasing slightly from 66.10 to 72.91 (*P*=.68).

XSL•FO
RenderX

Figure 5. Interaction between time and treatment in the State-Trait Anxiety Inventory (STAI)-E score.



Furthermore, the adjusted logistic regression model demonstrated that participants in the intervention group had a significantly lower likelihood of presenting postintervention STAI-E scores >86 points compared to the control group (OR 0.05, 95% CI 0.01-0.40; *P*=.01). In addition, a higher preintervention STAI-E score was associated with an increased probability of exceeding the threshold (OR 1.04, 95% CI 1.01-1.08; *P*=.02;Table 4).

Table 4. Multivariate logistic regression model for postintervention State-Trait Anxiety Inventory (STAI) scores (state; >86 points)

Variable	Multivariate[a] odds ratio (95% CI)	*P* value[b]
Intervention group	0.09 (0.00-0.73)	.05[b]
Preintervention STAI-E score	1.04 (1.01-1.08)	.02
Age (y)	0.83 (0.64-1.01)	.08
Personal income	0.40 (0.05-2.46)	.34
Current planned pregnancy	0.32 (0.05-1.75)	.20
Personal history of anxiety	0.95 (0.16-6.05)	.96

[a]Hosmer–Lemeshow goodness-of-fit test, *P*=.18. Area under the receiver operating characteristic curve=0.79 (95% CI 0.65-0.88).

[b]*P* value is considered statistically significant when <.05.

Pulse Measurements

Before intervention, the intervention group had a mean pulse of 85.91 (SD 5.47) beats per minute. After intervention, the mean pulse decreased to 76.94 (SD 6.62) beats per minute, indicating a significant improvement.

Intervention Adherence

Among the 33 participants in the intervention group, 32 (97%) provided valid data for adherence analysis. The average number of completed sessions was 37.84 (SD 9.83), with a median of 42 sessions. A total of 17 (51%) participants completed all 42 sessions, while 26 (79%) participants completed at least 30 sessions. Only 6 (18%) participants completed fewer than 30 sessions, and 1 (3%) participant had no adherence data available.

Participant Satisfaction

Participant satisfaction with the eHealth intervention was quantitatively assessed. Of the participants in the intervention group, 87% (29/35) reported being “very satisfied” with the intervention, while 9% (3/35) were “satisfied,” and 3% (1/35) expressed a neutral stance. No participants reported dissatisfaction.

XSL•FO
RenderX

In addition, 91% (32/35) of participants found the intervention "easy to use," and 86% (30/35) stated that it "helped them feel more supported during their pregnancy." These quantitative measures highlight the high level of acceptability and usability of the intervention.

No serious adverse events were reported. Only mild cases of dizziness (3/35, 9%) and temporary nausea (2/35, 6%) were recorded, all of which resolved spontaneously within a few minutes without requiring interruption of the intervention.

Discussion

Principal Findings

This study examined the efficacy of a low-intensity eHealth intervention supported by IVR technology in improving mental health outcomes during pregnancy. The results provide evidence that the intervention contributed to significant reductions in anxiety and depression scores among participants in the intervention group, compared to the control group. Specifically, participants in the intervention group experienced a decrease in EPDS scores from 11.32 to 7.25, while the control group showed an increase from 11.32 to 16.23. A mixed-effects model revealed a significant interaction between time and group (coefficient=−9.03; P<.001), emphasizing the intervention's effect on mitigating depressive symptoms. Similarly, for STAI-E, the intervention group's score decreased from 61.89 to 35.03, whereas the control group's score increased to 72.91. These findings underscore the intervention's capacity to address acute and persistent anxiety symptoms during pregnancy. The reduction in EPDS and STAI-E scores could be related to changes in cortisol levels, sleep quality, or social support [38].

The choice of IVR as an intervention tool is based on the growing scientific evidence supporting its efficacy in reducing anxiety and depression. Previous studies have shown that IVR can induce states of relaxation, facilitate emotional regulation, and reduce stress through immersion in controlled environments designed to promote psychological well-being [18,19]. Its ability to isolate the user from external stimuli and generate an immersive experience promotes greater adherence to, and engagement with, the intervention, which is particularly beneficial in pregnant populations. Furthermore, previous research has shown that IVR activates neurobiological mechanisms associated with the reduction of stress and anxiety [20], reinforcing its suitability as a low-risk, nonpharmacological strategy that is easy to implement in this context. Although the intervention was not specifically validated in women from various cultural contexts, it was designed to be accessible and applicable to a broad spectrum of users. The content of the intervention focused on mindfulness and relaxation exercises, techniques widely used and recommended in perinatal mental health care. To ensure that it was accepted and understood, a neutral and accessible language was used, without specific cultural references that could limit its applicability in different settings. In addition, priority was given to usability and ease of integration into the daily routine of the participants, which are the key aspects to maximize adherence to the intervention.

Our approach in this study is based on modern theories of emotional processing, which postulate that emotional responses are mediated by neural networks associated with memory and fear regulation [39]. IVR facilitates immersion in relaxing and controlled environments that induce a physiological relaxation response, promoting regulation of the autonomic nervous system and reducing activation of the amygdala, a key structure in the response to stress and anxiety [19]. Furthermore, the immersive experience in IVR promotes attentional focus and cognitive distraction, mechanisms that have been shown to be effective in reducing symptoms of anxiety and depression [18,20]. The intervention is based on mindfulness-based therapy, which has been widely validated in the treatment of anxiety and depression [16]. IVR enhances this practice by providing an immersive environment that facilitates the experience of presence and engagement with the activity, optimizing therapeutic benefits. Repeating sessions over 6 weeks favors the consolidation of emotional regulation patterns through neuroplasticity mechanisms, generating sustained changes in the perception of stress and anxiety [21]. Thus, the theoretical model that supports our intervention combines the physiological regulation of stress through modulation of the autonomic nervous system, cognitive distraction and attentional focus facilitated by immersion, and the progressive internalization of emotional coping strategies through the repeated practice of mindfulness in a safe and structured environment.

Although our study protocol did not include a specific assessment of sleep quality or anxiety sensitivity, some publications suggest that anxiety and depression during pregnancy are closely related to these factors. Previous studies have shown that mindfulness- and relaxation-based interventions, such as the one used in our study, can improve sleep quality and reduce stress levels [40,41]. Furthermore, it has been shown that anxiety sensitivity can act as a vulnerability factor for the development of anxiety and depressive symptoms, which reinforces the relevance of addressing these aspects in future research [42].

There are several mechanisms that may explain the observed benefits of IVR-based intervention on mental health. First, the cognitive and emotional distraction effect plays a key role in reducing symptoms of anxiety and depression. The immersive nature of virtual reality allows attention to be diverted from stressors, providing a controlled and immersive environment that promotes relaxation. This mechanism has been demonstrated in previous studies, in which exposure to virtual environments effectively reduced the perception of stress and pain in clinical contexts. Second, the induction of physiological relaxation may be a determining factor. The IVR intervention includes mindfulness and guided relaxation techniques, which have been associated with reduced autonomic activation, decreased heart rate, and better emotional regulation. These effects are consistent with our findings, in which a significant reduction in symptoms of anxiety and depression, as well as a decrease in heart rate, was observed in the intervention group. Third, the influence on neuroplasticity and emotional engagement could contribute to the positive effects of IVR. Immersive experiences can modulate neural circuits associated with stress and mood regulation, promoting adaptive cognitive

XSL•FO
RenderX

and emotional responses. Furthermore, the activation of reward-related brain areas during the virtual experience can improve mood and motivation, which would explain the high levels of adherence and satisfaction observed in our study. Finally, the increase in the perception of support and self-efficacy could mediate the effectiveness of the intervention. By providing a structured, engaging, and easily accessible tool for mental well-being, IVR allows users to take an active role in managing their symptoms. This sense of control has previously been linked to better outcomes in digital health care interventions. In any case, future studies should explore these mechanisms in greater depth by using neuroimaging techniques, physiological monitoring, and qualitative evaluations of the user's experience.

Regarding the generalizability of the findings, we acknowledge that the study was conducted in a specific region of Spain, which might limit its applicability to other populations. Nevertheless, to strengthen external validity, the following 2 strategies were implemented: (1) participants were recruited from 5 primary care centers (ASSIR), allowing for the inclusion of women with different sociodemographic profiles within the public health system; and (2) the virtual reality intervention used does not depend on specific cultural or contextual factors, facilitating its implementation in other prenatal care settings. Our findings are consistent with previous studies on eHealth and virtual reality for the management of anxiety and depression, suggesting that this technology may be effective in several pregnant populations.

With regard to the representativeness of the sample and the strategies to improve diversity, participants were selected in public primary care centers, ensuring equal access to the intervention and avoiding biases associated with the use of private services. Broad inclusion criteria were established, including women aged ≥18 years with moderate symptoms of anxiety and depression (EPDS scores: 9-12), which favored the heterogeneity of the sample. All participants had to speak and understand Spanish, ensuring understanding of the intervention and evaluation of the results without linguistic bias. Although differences were observed in the educational level of the participants, we performed adjusted analyses to control possible effects of these variables on the results.

In our study, no serious adverse events were reported. A pilot study in the United States analyzed the feasibility of IVR for distraction and pain management in women in labor and reported a high acceptance with no serious side effects [43]. While in the general population the use of IVR may be associated with effects such as dizziness, nausea, or disorientation—known as cybersickness—these effects appear to depend on the design of the intervention, the duration of exposure, and the individual characteristics of the users [44]. Although the evidence so far suggests that IVR is a safe and well-tolerated intervention in pregnant women, we encourage future research evaluating potential adverse effects in the long term and in more diverse populations.

Comparison With Previous Literature

Our findings align with previous literature emphasizing the potential of digital mental health interventions for perinatal populations. For example, Bell et al [45] demonstrated the effectiveness of virtual reality in managing anxiety and depression in diverse clinical settings. Similarly, studies such as Pallavicini et al [46] highlighted the ability of virtual reality games to elicit positive emotions and reduce state anxiety through immersive and interactive experiences. These findings suggest that IVR technology's immersive nature enhances user engagement and emotional regulation, supporting its use as a tool for mental health interventions during pregnancy.

In addition, Fuster-Casanovas et al [47] observed a significant rise in the adoption of eHealth tools for managing mental health in Catalonia, further highlighting the role of digital platforms in enhancing accessibility to psychological support. Their findings resonate with our study's results, as both suggest that digital tools, particularly IVR, can bridge gaps in care by offering scalable and engaging solutions for vulnerable populations. Moreover, Bell et al [48] reported that virtual reality interventions significantly reduced anxiety in adolescent patients in hospital settings when immersive, therapeutic applications were used, providing further evidence for the clinical utility of such technologies in diverse populations.

Satisfaction

Participant satisfaction with the IVR-based intervention was remarkably high, as evidenced by postintervention surveys. Participants reported that the immersive nature of the technology facilitated relaxation and emotional engagement, which contributed to the observed reductions in anxiety and depression symptoms. These findings are consistent with previous studies, such as Bell et al [48], which demonstrated that virtual reality interventions elicited positive emotional responses and were well-received by users. In addition, Pallavicini et al [46] noted that the interactive elements of virtual reality contributed to enhanced user experiences, making the intervention more enjoyable and effective. Further support comes from Al Kuwaiti et al [49], who highlighted that virtual reality's adaptability and immersive environments increase patient satisfaction in therapeutic applications. The positive feedback underscores the acceptability of IVR as a feasible mental health intervention, particularly in populations with limited access to traditional therapies.

Adherence

Adherence to the intervention protocol was high, with 97% (32/33) of participants providing valid data. Among them, 79% (26/33) completed at least 30 sessions. This is in line with findings from Fuster-Casanovas et al [47] who highlighted the ease of use and accessibility of eHealth tools as key factors in ensuring consistent engagement. High adherence rates were also observed in studies using similar virtual reality–based interventions, such as Bell et al [45], which attributed participant compliance to the engaging and user-friendly nature of the technology. In addition, findings by Garcia et al [50] suggested that virtual reality interventions designed with user-centric features significantly improved adherence by maintaining user motivation and interest. Furthermore, Indovina et al [51] demonstrated that adherence to virtual reality–based pain management programs was strongly associated with the level of interactivity and the ability to tailor content to user preferences. These findings suggest that the immersive and

XSL•FO
RenderX

interactive characteristics of IVR not only enhance satisfaction but also foster sustained participation, which is critical for achieving meaningful mental health outcomes.

We found that user-centered design and personalization of IVR is very important. As far as individual needs assessment is concerned, before implementing the IVR intervention, it is essential to conduct an initial assessment of each pregnant woman's needs, preferences, and expectations. This may include questionnaires about her anxiety and depression levels, as well as her preferences regarding types of exercises (eg, mindfulness, relaxation, and guided imagery) and virtual environments (eg, beaches, forests, and mountains) [23]. This assessment would allow for the creation of a personalized intervention plan that fits each woman's specific needs, thus increasing adherence to, and satisfaction with, the intervention [52]. As far as adaptation of virtual environments is concerned, they should be designed to be inclusive and culturally sensitive. For example, women from certain cultural backgrounds may feel more comfortable with environments that reflect their natural or cultural environment. This could involve adapting landscapes, sounds, and narratives to be culturally relevant [53]. Furthermore, environments should be adjustable in sensory intensity (eg, sounds, colors, and movement) to suit individual preferences and avoid overstimulation, especially in women who may be more sensitive during pregnancy [54]. Regarding flexibility of duration and frequency of the IVR session, they should be flexible to suit pregnant women's routines and energy levels. Some women may prefer shorter but more frequent sessions, while others may benefit from longer but less frequent sessions [55]. The ability to pause and resume sessions is also important, as pregnant women may need frequent breaks due to physical discomfort or fatigue [56]. The incorporation of real-time feedback systems can improve the personalization of the intervention. For example, heart rate sensors or short questionnaires during sessions can provide information about the woman's emotional state, allowing the IVR content to be adjusted based on her responses. This dynamic approach can help ensure that the intervention is always relevant and effective for each woman [56]. Finally, to ensure that the IVR intervention is truly user-centered, it is essential to involve pregnant women in the design and development process. This may include focus groups, pilot testing, and satisfaction surveys to gather direct feedback about the content and usability of the intervention [57]. Active user participation not only improves the acceptance of the intervention but also increases the likelihood that it will be effective and sustainable in the long term [58].

Strengths and Limitations

While our study contributes to the growing evidence base for IVR interventions, it is not without limitations. First, the relatively small sample size restricts the generalizability of the findings. Second, the reliance on self-reported adherence and satisfaction data introduces potential reporting biases. In addition, the absence of long-term follow-up limits our ability to evaluate the sustained impact of the intervention on postpartum mental health.

Another limitation to consider is related to the extrapolation of these results to women from disadvantaged socioeconomic backgrounds, or whose technological limitations could be restrictive, given that the availability of, and familiarity with, technologies such as IVR can differ significantly between socioeconomic groups. Also, the representativeness of the sample is limited because the participants generally had a high educational level and good access to health care and technological resources, which could bias the results toward greater effectiveness and adherence. Another potential limitation is related to possible biases derived from self-reporting of adherence to, and satisfaction with, the intervention, which may reflect a social desirability bias. Finally, it is worth noting that long-term follow-up was not conducted, which limits conclusions about the sustainability of the treatment's effect on postpartum mental health. Future research should therefore consider, including more diverse samples, including women from different socioeconomic and cultural backgrounds, as well as evaluating alternative methods of technological access that facilitate the equal participation of women in vulnerable situations. In addition, future studies should implement longer follow-ups to assess the sustainability of the observed impact on mental health and delve deeper into possible biases arising from self-reporting and the specific characteristics of the studied group.

Nonetheless, this study has several strengths. The use of IVR technology represents an innovative and accessible approach to perinatal mental health care. Rigorous adjustments for baseline differences strengthen the reliability of our findings, and the high adherence rates suggest feasibility for real-world implementation. Importantly, the results of this study provide a foundation for future research to build upon, particularly in exploring the long-term effects and cost-effectiveness of IVR interventions for pregnant women.

Clinical Implications and Future Directions

From a clinical perspective, the integration of IVR-based eHealth programs into prenatal care offers a promising avenue to address gaps in mental health support. These programs can serve as scalable and engaging solutions, particularly in contexts where traditional resources are limited. Future studies should explore how to customize IVR content to meet individual needs and preferences, potentially enhancing both engagement and outcomes. Fajnerova et al [59] emphasized the potential of multiuser virtual environments for facilitating group-based mental health interventions, which could be explored further in maternal care settings.

Our primary objective in this RCT was to assess the immediate efficacy of the IVR intervention on symptoms of anxiety and depression during pregnancy, using standardized instruments (EPDS and STAI). However, the evaluation of durability of these beneficial effects beyond the immediate intervention period is essential to confirm its clinical utility and inform maternal mental health policies. We therefore plan to conduct future longitudinal studies that include extended postpartum follow-up periods (eg, from 6 to 12 months after birth) to assess whether reductions in prenatal symptoms of anxiety and depression symptoms are maintained after delivery. These future studies will adopt longitudinal designs to analyze the persistence of improvements in maternal mental health outcomes, including

XSL•FO
RenderX

the potential prevention of postpartum depression (PPD) and anxiety. In addition, long-term secondary outcomes related to maternal and infant health will be studied, specifically assessing maternal psychological well-being, mother-child bonding, parenting stress, and infant development trajectory using validated tools. These assessments will help determine whether the benefits observed during pregnancy have significant long-term effects, contributing to improved mother-child interactions and improved infant development trajectories. In addition, to improve the robustness and generalizability of the findings, future studies will include larger and more diverse cohorts, multicenter trials, and economic analyses to assess the cost-effectiveness of IVR interventions and their long-term effects on postpartum mental health. We believe that this comprehensive approach will provide robust evidence about the clinical relevance, sustainability, and scalability of IVR interventions in perinatal care settings.

We are also aware that there are some limitations related to the implementation of IVR-based interventions and eHealth technologies in real-world clinical settings. Effective implementation will require rigorous compliance with data protection regulations (eg, General Data Protection Regulation), clear informed consent processes, and robust cybersecurity measures to protect confidentiality among users. Furthermore, it is critical to consider the question of equity. The sociodemographic characteristics observed in this study suggest a potential bias toward participants with higher educational levels and stable economic situations, which could limit the applicability of the intervention in socioeconomically disadvantaged populations. Addressing these equity concerns involves ensuring technological accessibility, affordability, and ease of use across diverse socioeconomic groups and geographic locations, while also considering differences in technological ability and internet connectivity. Financial constraints are also a barrier, particularly relevant in the context of national health care services, where resources are often limited and require careful budgeting. The implementation of these technologies may require considerable initial investments, as well as clear financial strategies to ensure long-term sustainability and cost-effectiveness. Finally, practical challenges, such as integration into existing health care services, obtaining acceptance by health care professionals, and devising sustainable financing models must also be considered. For all of these reasons, future research should focus on longitudinal analyses, assessing not only immediate efficacy but also sustained use and real-world feasibility, with the aim of developing scalable and inclusive solutions tailored to different clinical contexts and populations.

In relation to how to integrate IVR-based interventions into the health care system, we propose different key aspects as shown in Textbox 2.

Textbox 2. Proposed strategies to integrate immersive virtual reality (IVR)–based interventions into health care system.

Implementation within primary care and maternity services

- Midwives and health care professionals could incorporate IVR sessions into routine prenatal visits, offering guided relaxation and mindfulness exercises as a complementary mental health support tool.
- The intervention could be integrated into antenatal education programs, in which pregnant women receive structured sessions on stress management and emotional well-being using IVR technology.

Home-based digital health care support

- Pregnant women could be prescribed access to the IVR application for home use, with health care professionals monitoring adherence and outcomes through a digital platform.
- A mobile app could complement the IVR sessions by providing additional psychological support, reminders, and educational content tailored to individual needs.

Scalability within public health care system

- Collaboration with public health care providers could facilitate the incorporation of IVR technology into maternal mental health care programs, ensuring accessibility for a broad population.
- Future research should focus on cost-effectiveness analyses and feasibility studies to assess the sustainability of implementing IVR interventions on a larger scale.

Regarding to how to implement IVR in routine clinical practice, we consider the following recommendations for the integration of IVR into antenatal care protocols:

1. Early detection of anxiety and depression symptoms: through the implementation of IVR, routine screening for anxiety and depression symptoms could begin during antenatal visits, using validated tools, such as EPDS and STAI-E. Women presenting with moderate anxiety or depression symptoms (scores within the range established in our study) could be referred for the IVR intervention as part of a stepped care approach.
2. Training of professional health care staff, especially midwives and primary care nurses, in the use of IVR technology and in supervising sessions: this would include training in setting up the equipment, explaining relaxation and mindfulness exercises, and monitoring patients' progress. It would also be important to provide guidelines about how to address potential questions or concerns from patients regarding the use of technology.
3. Access to technology: primary care centers could have IVR equipment available at antenatal follow-up consultations. Alternatively, lending IVR devices to patients for use at

XSL•FO
RenderX

home could be considered if adequate follow-up is ensured. The intervention could be designed to be self-guided, with preprogrammed sessions that patients could complete independently, but with the possibility of contacting a health care professional in case of questions or need for additional support.

4. Integration into the individualized care plan, including the IVR intervention and tailoring to each patient's specific needs and the context of their pregnancy: this could include coordination with other mental health professionals, such as psychologists or psychiatrists, in cases where a more intensive approach is required. In addition, it could be combined with other low-cost interventions, such as prenatal exercise programs, mindfulness workshops, or support groups, to offer a holistic approach to mental well-being during pregnancy.
5. Continuous monitoring and evaluation: a continuous monitoring system should be established to evaluate the effectiveness of the intervention and adjust it as necessary. This could include collecting data about adherence, patient satisfaction, and changes in anxiety and depression symptoms over time. Health care professionals could use the data collected to identify patients who could benefit from additional or more intensive interventions.
6. Cultural and socioeconomic considerations: as accessibility and acceptance of technology may vary depending on the cultural and socioeconomic context, it is important to adapt the intervention to the specific needs of each population. This could include translating content into different languages or adapting IVR scenarios to make them culturally relevant.

We have also reflected on the potential long-term effects of the IVR intervention. In the effects on postpartum mental health, we should consider the following: first, anxiety and depression during pregnancy are known risk factors for the development of PPD [8]. Because our intervention demonstrated a significant reduction in anxiety and depression symptoms during pregnancy, it is plausible that it may also contribute to a decreased risk of PPD [60]. Previous studies have shown that interventions that improve mental health during pregnancy have a protective effect against PPD, suggesting that IVR could have a positive impact on women's emotional well-being after delivery [61]. Second, maternal mental health is closely related to the quality of the mother-child bond [62]. A reduction in anxiety and depression symptoms during pregnancy could facilitate a stronger and more positive bond between mother and infant, which in turn could have beneficial effects on the child's emotional and social development [6]. Future studies could assess the impact of the IVR intervention on the quality of the mother-child bond using validated scales, such as the Postpartum Bonding Questionnaire [63]. In addition, regarding the effects on child development, we can consider the following: first, maternal anxiety and depression during pregnancy have been associated with an increased risk of problems in the child's cognitive and emotional development, including learning difficulties, attention problems, and emotional disorders [64]. By reducing these symptoms during pregnancy, the IVR intervention could contribute to a healthier development of the child [16]. Longitudinal studies could evaluate the impact of the intervention on the cognitive and emotional development of children using tools such as the Bayley Scales of Infant and Toddler Development or the Child Behavior Checklist [65]. Second, some research suggests that maternal stress and anxiety during pregnancy may be associated with an increased risk of neurodevelopmental disorders, such as attention-deficit/hyperactivity disorder or autism spectrum disorder [66]. Although the relationship is not direct, it is possible that reducing stress and anxiety during pregnancy through the IVR intervention may have a protective effect in this regard [61]. Future studies could explore this hypothesis by long-term follow-up of children whose mothers participated in the intervention. Finally, regarding future research, we consider the following: first, to assess the long-term effects of the IVR intervention, longitudinal studies following women and their children for several years after delivery would be necessary. These studies could measure not only maternal mental health, but also children's cognitive, emotional, and social development [60]. Second, it is important to explore the underlying biological and psychological mechanisms that could explain how reducing anxiety and depression during pregnancy influences postpartum mental health and child development. For example, changes in cortisol levels or other biological markers associated with stress could be investigated [6]. Third, future research could explore the effect of combining IVR intervention with other postpartum interventions, such as parenting support programs or cognitive behavioral therapies, to maximize benefits on maternal mental health and child development [16].

We consider that feasibility, scalability, and costs are key aspects for the IVR interventions to be effectively implemented, especially in underserved populations where access to mental health care is limited. With regard to feasibility and scalability in real-world settings, we consider the following: first, the IVR intervention requires the use of virtual reality devices, which can vary in cost and complexity. In our study, we used midrange devices that are relatively affordable and easy to use. These devices could be purchased by primary care centers or maternal health clinics or even rented to reduce initial costs. IVR technology is becoming more accessible and device costs have decreased in recent years, making it easier to implement in resource-limited settings. Second, implementation of the IVR intervention would require basic training of health care personnel, such as midwives, nurses, or psychologists, in using the technology and supervising sessions. This training could be done in a short period and would not require advanced technical skills. Furthermore, the intervention could be designed to be self-guided, which would reduce the need for constant supervision by health care staff. Third, the IVR intervention is highly adaptable and could be tailored to meet the needs of various populations, including those with language or cultural barriers. For example, relaxation and mindfulness scenarios could be adapted to reflect culturally relevant environments, and content could be translated into different languages. Regarding cost and resource estimation, we consider the following:

1. Device cost: the cost of IVR devices can vary between €300 (US $327) and €600 (US $654) per unit, depending on the brand and features. For a primary care center with a

XSL•FO
RenderX

moderate workload, several devices could be purchased for rotational use.
2. Content development: the development of IVR content (eg, mindfulness exercises and relaxation) may involve an initial cost, but once developed, the content can be used repeatedly without significant additional expenses.
3. Maintenance: maintenance costs are relatively low, as modern IVR devices are long-lasting and require few upgrades.
4. Training: training of health care staff could be completed in a short period, such as a one- or two-day workshop, at an estimated cost of €500 (US $545) to €1000 (US $1090), depending on the number of participants. This training would not require advanced technical skills.
5. Implementation and long-term savings: the intervention could be designed to be self-guided, reducing the need for constant supervision by health care staff. In the long term, the IVR intervention could generate significant savings by reducing the need for more expensive interventions (eg, pharmacological or individual therapies), particularly for cases of mild to moderate anxiety and depression. Additionally, improved mental health in pregnant women could help reduce obstetric and neonatal complications related to stress and depression.
6. Adaptability: the IVR intervention is highly adaptable and could be tailored to meet the needs of various populations, including those with language or cultural barriers. For example, relaxation and mindfulness scenarios could be culturally adapted and translated into different languages.

All currency conversions from euros to US dollars were calculated using the average exchange rate at the time of the study (€1=US $1.09).

In addition, by improving the mental health of pregnant women, costs associated with obstetric and neonatal complications related to stress and depression could be reduced.

In relation to scalability in underserved populations, we consider the following: First, the IVR intervention could be particularly useful in rural or remote areas where access to mental health care professionals is limited. The devices could be transported between different health facilities or even used at home under remote supervision. In addition, the portable nature of IVR devices allows their use in communities with limited infrastructures if there is access to electricity and adequate space for sessions. Second, to facilitate implementation in underserved populations, collaborations could be established with nongovernmental organizations or public health agencies already working in these areas. These organizations could provide logistical and financial support for the acquisition and maintenance of the devices.

In this study, women with a history of diagnosed anxiety or depression were excluded because this clinical subgroup could present a different evolution compared to those with recent and moderate symptoms, which could affect the validity of the results. However, we consider that future studies could evaluate the effectiveness of this intervention in populations with a previous psychiatric history. We also consider that future studies could further explore the influence of pregnancy planning on the effectiveness of digital interventions for perinatal mental health.

It is worth remembering that IVR has been applied in other areas of mental health: first, IVR has been used for posttraumatic stress disorder, in which it has proved to be an effective treatment tool, particularly in the context of exposure therapy. The technology allows for the simulation of controlled and safe environments where patients can gradually confront the stimuli that trigger their symptoms, facilitating desensitization and emotional processing [67]. Studies have shown that IVR can be particularly useful in war veterans, accident survivors, or survivors of violence, in whom exposure to virtual environments can help reduce the symptoms of reexperiencing, avoidance, and hyperarousal associated with posttraumatic stress disorder [68]. Furthermore, IVR can be combined with evidence-based therapies, such as cognitive behavioral therapy, to improve treatment outcomes [69]. Second, IVR has also shown potential in the treatment of substance-use disorders, particularly in relapse prevention. Virtual environments can simulate high-risk situations (eg, bars or parties) where patients can practice coping strategies in a safe and controlled environment [70]. In addition, IVR can be used for aversion therapy, where patients are exposed to virtual stimuli associated with substance use, which can help reduce craving and dependence [71]. Recent studies have shown that IVR can be effective in treating alcohol, tobacco, and other drug addictions; improving treatment adherence; and reducing relapse rates [72]. Third, IVR has been widely used in the treatment of specific phobias, such as fear of flying, heights, or closed spaces. Gradual exposure to these stimuli in a virtual environment allows patients to confront their fears in a controlled manner, which reduces anxiety and improves functioning [73]. Furthermore, IVR can be useful in the treatment of generalized anxiety, by providing relaxing environments and mindfulness exercises that help reduce symptoms of excessive worry and muscle tension [46]. Fourth, IVR has also shown potential in supporting individuals with autism spectrum disorders, particularly in developing social skills and reducing anxiety in social situations. Virtual environments can simulate complex social interactions, allowing patients to practice and improve their communication skills in a safe environment [74]. Fifth, IVR has been used as a complementary tool in the management of chronic pain, which is often associated with mental health disorders, such as depression and anxiety. Virtual environments can distract patients from the pain and provide relaxation techniques that improve their emotional well-being [51].

Finally, we would like to mention the integration of artificial intelligence and machine learning into IVR interventions. We believe that these new tools have great potential to improve the personalization and effectiveness of IVR. Artificial intelligence–based algorithms could allow the individualization of IVR experiences based on, for example, emotional state or therapeutic response.

Conclusions

The findings of this study highlight the significant potential of IVR-based eHealth interventions in improving mental health outcomes during pregnancy. By reducing symptoms of anxiety

XSL•FO
RenderX

and depression, the intervention demonstrated its capacity to address critical mental health challenges faced by pregnant women, aligning with the growing emphasis on integrating digital health solutions into prenatal care. This intervention, grounded in immersive and interactive technology, not only achieved significant reductions in EPDS and STAI-E scores but also fostered high adherence and participant satisfaction. These results underscore the feasibility of implementing IVR-based programs as accessible, low-intensity, mental health interventions within broader health systems. Such programs hold promise for addressing disparities in access to care, particularly for vulnerable populations or those with limited availability of traditional mental health services.

Despite these, conclusions should be interpreted with caution given the relatively small sample size, the socioeconomic and educational homogeneity of the participants, and the possible presence of biases derived from self-reporting in the assessment of adherence and satisfaction. Furthermore, the lack of long-term follow-up limits the ability to determine the sustainability of the observed benefits. Therefore, although the findings support the potential efficacy of IVR-based interventions in specific contexts, additional studies with more diverse samples and longer follow-ups are required to confirm these results and explore their applicability in different socioeconomic and cultural contexts.

Future research should aim to address these limitations while exploring the scalability, cost-effectiveness, and cultural adaptability of IVR-based interventions. In addition, longitudinal studies are needed to evaluate the sustained impact of such interventions on postpartum mental health and their potential influence on maternal and child outcomes.

Acknowledgments

The authors would like to extend our gratitude to all the participants who generously contributed their time to this study. Special thanks to midwives in the primary care centers and professionals who supported the recruitment and facilitation of this research, including Centre d'Atenció Primària (CAP; Primary Care Center in English) Rambla, CAP Can Tries, CAP Rubí, CAP Can Mates, and CAP Can Parellada.

We also wish to thank Delaguila Games [31], the developers of the immersive virtual reality technology used in this study. Their innovation and technical expertise played a crucial role in creating a user-centered intervention tailored to the needs of pregnant individuals.

Finally, we are grateful to the research team, midwives, and health care professionals involved in the implementation and execution of the study.

This research was supported by funding from the Departament de Salut, Generalitat de Catalunya (PERIS Grant SLT006/17/249). This study was partially funded by the Official College of Nurses of Barcelona [75] as part of the Nurse Research Projects Grants (PR-642/2023). In addition, the study received the Mustela Scholarship for midwifery research projects awarded by the Mustela Foundation (FAME 2023).

Authors' Contributions

CE and GFP contributed equally to this work and should be considered senior authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT-EHEALTH checklist.
[PDF File (Adobe PDF File), 33 KB - humanfactors_v12i1e71708_app1.pdf]

References

1. Ajong AB, Kenfack B, Ali IM, Yakum MN, Ukaogo PO, Mangala FN, et al. Calcium supplementation in pregnancy: an analysis of potential determinants in an under-resourced setting. PLoS One 2023;18(10):e0292303 [FREE Full text] [doi: 10.1371/journal.pone.0292303] [Medline: 37796953]
2. Ajong AB, Yakum MN, Aljerf L, Ali IM, Mangala FN, Onydinma UP, et al. Association of hypertension in pregnancy with serum electrolyte disorders in late pregnancy among Cameroonian women. Sci Rep 2023 Nov 28;13(1):20940 [FREE Full text] [doi: 10.1038/s41598-023-47623-6] [Medline: 38017060]
3. Mental health in pregnancy. Royal College of Psychiatrists. 2018 Nov. URL: https://www.rcpsych.ac.uk/mental-health/treatments-and-wellbeing/mental-health-in-pregnancy [accessed 2025-04-14]
4. Bauer A, Knapp M, Parsonage M. Lifetime costs of perinatal anxiety and depression. J Affect Disord 2016 Mar 01;192:83-90 [FREE Full text] [doi: 10.1016/j.jad.2015.12.005] [Medline: 26707352]

XSL•FO
RenderX

5. Ghazanfarpour M, Bahrami F, Rashidi Fakari F, Ashrafinia F, Babakhanian M, Dordeh M, et al. Prevalence of anxiety and depression among pregnant women during the COVID-19 pandemic: a meta-analysis. J Psychosom Obstet Gynaecol 2022 Sep;43(3):315-326 [FREE Full text] [doi: 10.1080/0167482X.2021.1929162] [Medline: 34165032]
6. van den Bergh BR, Mulder EJ, Mennes M, Glover V. Antenatal maternal anxiety and stress and the neurobehavioural development of the fetus and child: links and possible mechanisms. A review. Neurosci Biobehav Rev 2005 Apr;29(2):237-258. [doi: 10.1016/j.neubiorev.2004.10.007] [Medline: 15811496]
7. Dunkel Schetter C, Tanner L. Anxiety, depression and stress in pregnancy: implications for mothers, children, research, and practice. Curr Opin Psychiatry 2012 Mar;25(2):141-148 [FREE Full text] [doi: 10.1097/YCO.0b013e3283503680] [Medline: 22262028]
8. Goodman JH, Chenausky KL, Freeman MP. Anxiety disorders during pregnancy: a systematic review. J Clin Psychiatry 2014 Oct;75(10):e1153-e1184. [doi: 10.4088/JCP.14r09035] [Medline: 25373126]
9. Grote NK, Bridge JA, Gavin AR, Melville JL, Iyengar S, Katon WJ. A meta-analysis of depression during pregnancy and the risk of preterm birth, low birth weight, and intrauterine growth restriction. Arch Gen Psychiatry 2010 Oct 04;67(10):1012-1024 [FREE Full text] [doi: 10.1001/archgenpsychiatry.2010.111] [Medline: 20921117]
10. Antenatal and postnatal mental health: clinical management and service guidance. National Institute for Health and Care Excellence. 2024. URL: https://www.nice.org.uk/guidance/cg192 [accessed 2025-04-14]
11. Recommendations on digital interventions for health system strengthening. World Health Organization. 2019 Jun 6. URL: https://www.who.int/publications/i/item/9789241550505 [accessed 2025-04-14]
12. Common mental health problems: identification and pathways to care. National Institute for Health and Care Excellence. 2011 May 25. URL: https://www.nice.org.uk/guidance/cg123 [accessed 2025-04-14]
13. mHealth: new horizons for health through mobile technologies. World Health Organization. 2011. URL: https://www.afro.who.int/publications/mhealth-new-horizons-health-through-mobile-technologie [accessed 2025-04-14]
14. Ryu S. Book Review: mHealth: new Horizons for Health through mobile technologies: based on the findings of the Second Global Survey on eHealth (Global Observatory for eHealth Series, Volume 3). Healthc Inform Res 2012;18(3):231. [doi: 10.4258/hir.2012.18.3.231]
15. Lee EW, Denison FC, Hor K, Reynolds RM. Web-based interventions for prevention and treatment of perinatal mood disorders: a systematic review. BMC Pregnancy Childbirth 2016 Feb 29;16(1):38 [FREE Full text] [doi: 10.1186/s12884-016-0831-1] [Medline: 26928898]
16. Shi Z, MacBeth A. The effectiveness of mindfulness-based interventions on maternal perinatal mental health outcomes: a systematic review. Mindfulness (N Y) 2017 Jan 19;8(4):823-847 [FREE Full text] [doi: 10.1007/s12671-016-0673-y] [Medline: 28757900]
17. Marc I, Toureche N, Ernst E, Hodnett ED, Blanchet C, Dodin S, et al. Mind-body interventions during pregnancy for preventing or treating women's anxiety. Cochrane Database Syst Rev 2011 Jul 06;2011(7):CD007559 [FREE Full text] [doi: 10.1002/14651858.CD007559.pub2] [Medline: 21735413]
18. Maples-Keller JL, Bunnell BE, Kim S, Rothbaum BO. The use of virtual reality technology in the treatment of anxiety and other psychiatric disorders. Harv Rev Psychiatry 2017;25(3):103-113. [doi: 10.1097/hrp.0000000000000138]
19. Bohil CJ, Alicea B, Biocca FA. Virtual reality in neuroscience research and therapy. Nat Rev Neurosci 2011 Nov 03;12(12):752-762. [doi: 10.1038/nrn3122] [Medline: 22048061]
20. Fodor LA, Cote CD, Cuijpers P, Szamoskozi S, David D, Cristea IA. The effectiveness of virtual reality based interventions for symptoms of anxiety and depression: a meta-analysis. Sci Rep 2018 Jul 09;8(1):10323 [FREE Full text] [doi: 10.1038/s41598-018-28113-6] [Medline: 29985400]
21. Donker T, Blankers M, Hedman E, Ljótsson B, Petrie K, Christensen H. Economic evaluations of internet interventions for mental health: a systematic review. Psychol Med 2015 Aug 03;45(16):3357-3376. [doi: 10.1017/s0033291715001427]
22. Anderson P, Jané-Llopis E. Mental health promotion and mental disorder prevention: a policy for Europe. Radboud University Nijmegen. 2005. URL: https://ec.europa.eu/health/archive/ph_projects/2002/promotion/fp_promotion_2002_a01_16_en.pdf [accessed 2025-04-14]
23. Mohr DC, Burns MN, Schueller SM, Clarke G, Klinkman M. Behavioral intervention technologies: evidence review and recommendations for future research in mental health. Gen Hosp Psychiatry 2013 Jul;35(4):332-338 [FREE Full text] [doi: 10.1016/j.genhosppsych.2013.03.008] [Medline: 23664503]
24. Deady M, Kay-Lambkin F, Teesson M, Mills K. Developing an integrated, internet-based self-help programme for young people with depression and alcohol use problems. Internet Interv 2014 Jul;1(3):118-131. [doi: 10.1016/j.invent.2014.06.004]
25. Vukovic V, Favaretti C, Ricciardi W, de Waure C. Health technology assessment evidence on e-health/m-health technologies: evaluating the transparency and thoroughness. Int J Technol Assess Health Care 2018 Feb 19;34(1):87-96. [doi: 10.1017/s0266462317004512]
26. Jimenez-Barragan M, Del Pino Gutierrez A, Garcia JC, Monistrol-Ruano O, Coll-Navarro E, Porta-Roda O, et al. Study protocol for improving mental health during pregnancy: a randomized controlled low-intensity m-health intervention by midwives at primary care centers. BMC Nurs 2023 Sep 07;22(1):309 [FREE Full text] [doi: 10.1186/s12912-023-01440-4] [Medline: 37674184]

XSL•FO
RenderX

27. O'Connor E, Rossom RC, Henninger M, Groom HC, Burda BU. Primary care screening for and treatment of depression in pregnant and postpartum women: evidence report and systematic review for the US preventive services task force. JAMA 2016 Jan 26;315(4):388-406. [doi: 10.1001/jama.2015.18948] [Medline: 26813212]
28. Dadi AF, Miller ER, Bisetegn TA, Mwanri L. Global burden of antenatal depression and its association with adverse birth outcomes: an umbrella review. BMC Public Health 2020 Feb 04;20(1):173 [FREE Full text] [doi: 10.1186/s12889-020-8293-9] [Medline: 32019560]
29. Garcia-Esteve L, Ascaso C, Ojuel J, Navarro P. Validation of the Edinburgh Postnatal Depression Scale (EPDS) in Spanish mothers. J Affect Disord 2003 Jun;75(1):71-76. [doi: 10.1016/s0165-0327(02)00020-4] [Medline: 12781353]
30. Spielberger CD, Gorsuch RL, Lushene RE. STAI. Cuestionario de ansiedad estado-rasgo. Hogrefe. URL: https://www.hogrefe-tea.com/public/catalogo/producto/STAI--CUESTIONARIO-DE-ANSIEDAD-ESTADO-RASGO [accessed 2025-04-14]
31. Virtual reality for health care. Delaguila Games. URL: https://delaguila.games/ [accessed 2025-04-13]
32. Protocol de seguiment de l'embaràs a Catalunya. Generalitat de Catalunya. 2018. URL: https://salutpublica.gencat.cat/web/.content/minisite/aspcat/promocio_salut/embaras_part_puerperi/protocol_seguiment_embaras/protocol-seguiment-embaras-2018.pdf [accessed 2025-01-17]
33. Pan WL, Gau ML, Lee TY, Jou HJ, Liu CY, Wen TK. Mindfulness-based programme on the psychological health of pregnant women. Women Birth 2019 Feb;32(1):e102-e109. [doi: 10.1016/j.wombi.2018.04.018] [Medline: 29752225]
34. Maderuelo JA, Haro AM, Pérez F, Cercas LJ, Valentín AB, Morán E. [Women's satisfaction with pregnancy follow-up. Differences among healthcare levels]. Gac Sanit 2006;20(1):31-39 [FREE Full text] [doi: 10.1157/13084125] [Medline: 16539991]
35. Maderuelo-Fernández J, Rodríguez-Martín B, Martín-Baranera M. Satisfacción de las mujeres con la atención prenatal: desarrollo y validación de un cuestionario. Gac Sanit 2006;20(6):453-459. [doi: 10.1157/13094880]
36. Tcha-Tokey K, Christmann O, Loup-Escande E, Richir S. Proposition and validation of a questionnaire to measure the user experience in immersive virtual environments. Int J Virtual Real 2016 Jan 01;16(1):33-48. [doi: 10.20870/IJVR.2016.16.1.2880]
37. Hosmer DWJ, Lemeshow S, Sturdivant RX. Applied Logistic Regression. Hoboken, NJ: John Wiley & Sons; 2013.
38. Světlák M, Malatincová T, Halámková J, Barešová Z, Lekárová M, Vigašová D, et al. The effectiveness of three mobile-based psychological interventions in reducing psychological distress and preventing stress-related changes in the psycho-neuro-endocrine-immune network in breast cancer survivors: study protocol for a randomised controlled trial. Internet Interv 2023 Apr;32:100628 [FREE Full text] [doi: 10.1016/j.invent.2023.100628] [Medline: 37273931]
39. Pessoa L. A network model of the emotional brain. Trends Cogn Sci 2017 May;21(5):357-371 [FREE Full text] [doi: 10.1016/j.tics.2017.03.002] [Medline: 28363681]
40. Chiesa A, Serretti A. Mindfulness-based stress reduction for stress management in healthy people: a review and meta-analysis. J Altern Complement Med 2009 May;15(5):593-600. [doi: 10.1089/acm.2008.0495] [Medline: 19432513]
41. Alvaro PK, Roberts RM, Harris JK. A systematic review assessing bidirectionality between sleep disturbances, anxiety, and depression. Sleep 2013 Jul 01;36(7):1059-1068 [FREE Full text] [doi: 10.5665/sleep.2810] [Medline: 23814343]
42. Taylor S. Anxiety Sensitivity: Theory, Research, and Treatment of the Fear of Anxiety. Milton Park, UK: Taylor & Francis; 2014.
43. Frey DP, Bauer ME, Bell CL, Low LK, Hassett AL, Cassidy RB, et al. Virtual reality analgesia in labor: the VRAIL pilot study-a preliminary randomized controlled trial suggesting benefit of immersive virtual reality analgesia in unmedicated laboring women. Anesth Analg 2019 Jun;128(6):e93-e96. [doi: 10.1213/ANE.0000000000003649] [Medline: 31094789]
44. Simón-Vicente L, Rodríguez-Cano S, Delgado-Benito V, Ausín-Villaverde V, Cubo Delgado E. Cybersickness. A systematic literature review of adverse effects related to virtual reality. Neurologia (Engl Ed) 2024 Oct;39(8):701-709 [FREE Full text] [doi: 10.1016/j.nrleng.2022.04.007] [Medline: 39396266]
45. Bell IH, Nicholas J, Alvarez-Jimenez M, Thompson A, Valmaggia L. Virtual reality as a clinical tool in mental health research and practice . Dialogues Clin Neurosci 2020 Jun;22(2):169-177 [FREE Full text] [doi: 10.31887/DCNS.2020.22.2/lvalmaggia] [Medline: 32699517]
46. Pallavicini F, Argenton L, Toniazzi N, Aceti L, Mantovani F. Virtual reality applications for stress management training in the military. Aerosp Med Hum Perform 2016 Dec 01;87(12):1021-1010. [doi: 10.3357/AMHP.4596.2016] [Medline: 28323588]
47. Fuster-Casanovas A, Miró Catalina Q, Vidal-Alaball J, Escalé-Besa A, Carrión C. eHealth in the management of depressive episodes in Catalonia's primary care from 2017 to 2022: retrospective observational study. JMIR Ment Health 2024 Jan 18;11:e52816 [FREE Full text] [doi: 10.2196/52816] [Medline: 38236631]
48. Bell IH, Pot-Kolder R, Rizzo A, Rus-Calafell M, Cardi V, Cella M, et al. Advances in the use of virtual reality to treat mental health conditions. Nat Rev Psychol 2024 Jul 10;3(8):552-567. [doi: 10.1038/S44159-024-00334-9]
49. Al Kuwaiti A, Nazer K, Al-Reedy A, Al-Shehri S, Al-Muhanna A, Subbarayalu AV, et al. A review of the role of artificial intelligence in healthcare. J Pers Med 2023 Jun 05;13(6):951 [FREE Full text] [doi: 10.3390/jpm13060951] [Medline: 37373940]

XSL•FO
RenderX

50. Garcia LM, Birckhead BJ, Krishnamurthy P, Sackman J, Mackey IG, Louis RG, et al. An 8-week self-administered at-home behavioral skills-based virtual reality program for chronic low back pain: double-blind, randomized, placebo-controlled trial conducted during COVID-19. J Med Internet Res 2021 Feb 22;23(2):e26292-e21152. [doi: 10.2196/26292] [Medline: 33484240]
51. Indovina P, Barone D, Gallo L, Chirico A, de Pietro G, Giordano A. Virtual reality as a distraction intervention to relieve pain and distress during medical procedures: a comprehensive literature review. Clin J Pain 2018 Sep;34(9):858-877. [doi: 10.1097/AJP.0000000000000599] [Medline: 29485536]
52. Kelders SM, Kok RN, Ossebaard HC, van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152. [doi: 10.2196/jmir.2104] [Medline: 23151820]
53. Kreuter MW, Lukwago SN, Bucholtz RD, Clark EM, Sanders-Thompson V. Achieving cultural appropriateness in health promotion programs: targeted and tailored approaches. Health Educ Behav 2003 Apr;30(2):133-146. [doi: 10.1177/1090198102251021] [Medline: 12693519]
54. Riva G, Baños RM, Botella C, Mantovani F, Gaggioli A. Transforming experience: the potential of augmented reality and virtual reality for enhancing personal and clinical change. Front Psychiatry 2016;7:164 [FREE Full text] [doi: 10.3389/fpsyt.2016.00164] [Medline: 27746747]
55. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011 Aug 05;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]
56. Schueller SM, Tomasino KN, Mohr DC. Integrating human support into behavioral intervention technologies: the efficiency model of support. Clin Psychol Sci Pract 2016 Nov 17;24(1):27-45. [doi: 10.1111/cpsp.12173]
57. Sanders EB, Stappers PJ. Co-creation and the new landscapes of design. CoDesign 2008 Mar;4(1):5-18. [doi: 10.1080/15710880701875068]
58. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. J Med Internet Res 2015 Jan 30;17(1):e30 [FREE Full text] [doi: 10.2196/jmir.4055] [Medline: 25639757]
59. Fajnerova I, Hejtmánek L, Sedlák M, Jablonská M, Francová A, Stopková P. The journey from nonimmersive to immersive multiuser applications in mental health care: systematic review. J Med Internet Res 2024 Nov 07;26(2):e60441-e60200 [FREE Full text] [doi: 10.2196/60441] [Medline: 39509153]
60. Dennis CL, Dowswell T. Psychosocial and psychological interventions for preventing postpartum depression. Cochrane Database Syst Rev 2013 Feb 28;2013(2):CD001134. [doi: 10.1002/14651858.CD001134.pub3] [Medline: 23450532]
61. Field T. Prenatal depression effects on early development: a review. Infant Behav Dev 2011 Feb;34(1):1-14. [doi: 10.1016/j.infbeh.2010.09.008] [Medline: 20970195]
62. Brockington IF, Fraser C, Wilson D. The Postpartum Bonding Questionnaire: a validation. Arch Womens Ment Health 2006 Sep;9(5):233-242. [doi: 10.1007/s00737-006-0132-1] [Medline: 16673041]
63. Talge NM, Neal C, Glover V, Early Stress, Translational Research and Prevention Science Network: Fetal and Neonatal Experience on Child and Adolescent Mental Health. Antenatal maternal stress and long-term effects on child neurodevelopment: how and why? J Child Psychol Psychiatry 2007;48(3-4):245-261 [FREE Full text] [doi: 10.1111/j.1469-7610.2006.01714.x] [Medline: 17355398]
64. Ronald A, Pennell CE, Whitehouse AJ. Prenatal maternal stress associated with ADHD and autistic traits in early childhood. Front Psychol 2010;1:223 [FREE Full text] [doi: 10.3389/fpsyg.2010.00223] [Medline: 21833278]
65. Bayley N. Bayley Scales of Infant and Toddler Development– Third Edition. San Antonio, TX: Harcourt Assessment; 2006.
66. Glover V. Prenatal stress and its effects on the fetus and the child: possible underlying biological mechanisms. Adv Neurobiol 2015;10:269-283. [doi: 10.1007/978-1-4939-1372-5_13] [Medline: 25287545]
67. Rothbaum BO, Hodges LF, Ready D, Graap K, Alarcon RD. Virtual reality exposure therapy for Vietnam veterans with posttraumatic stress disorder. J Clin Psychiatry 2001 Aug;62(8):617-622. [doi: 10.4088/jcp.v62n0808] [Medline: 11561934]
68. Difede J, Cukor J, Jayasinghe N, Patt I, Jedel S, Spielman L, et al. Virtual reality exposure therapy for the treatment of posttraumatic stress disorder following September 11, 2001. J Clin Psychiatry 2007 Nov;68(11):1639-1647. [Medline: 18052556]
69. Gonçalves R, Pedrozo AL, Coutinho ES, Figueira I, Ventura P. Efficacy of virtual reality exposure therapy in the treatment of PTSD: a systematic review. PLoS One 2012;7(12):e48469 [FREE Full text] [doi: 10.1371/journal.pone.0048469] [Medline: 23300515]
70. Bordnick PS, Carter BL, Traylor AC. What virtual reality research in addictions can tell us about the future of obesity assessment and treatment. J Diabetes Sci Technol 2011 Mar 01;5(2):265-271 [FREE Full text] [doi: 10.1177/193229681100500210] [Medline: 21527092]
71. Lee JH, Kwon H, Choi J, Yang BH. Cue-exposure therapy to decrease alcohol craving in virtual environment. Cyberpsychol Behav 2007 Oct;10(5):617-623. [doi: 10.1089/cpb.2007.9978] [Medline: 17927528]

XSL•FO
RenderX

72. Hone-Blanchet A, Wensing T, Fecteau S. The use of virtual reality in craving assessment and cue-exposure therapy in substance use disorders. Front Hum Neurosci 2014;8:844 [FREE Full text] [doi: 10.3389/fnhum.2014.00844] [Medline: 25368571]
73. Botella C, Fernández-Álvarez J, Guillén V, García-Palacios A, Baños R. Recent progress in virtual reality exposure therapy for phobias: a systematic review. Curr Psychiatry Rep 2017 Jul;19(7):42. [doi: 10.1007/s11920-017-0788-4] [Medline: 28540594]
74. Parsons S, Mitchell P. The potential of virtual reality in social skills training for people with autistic spectrum disorders. J Intellect Disabil Res 2002 Jun;46(Pt 5):430-443. [doi: 10.1046/j.1365-2788.2002.00425.x] [Medline: 12031025]
75. Official College of Nurses of Barcelona homepage. Official College of Nurses of Barcelona. URL: https://www.coib.cat/ca-es/inici.html [accessed 2025-04-13]

Abbreviations

ASSIR: Atenció a la Salut Sexual i Reproductiva (Sexual and Reproductive Health Care)
CONSORT: Consolidated Standards of Reporting Trials
EPDS: Edinburgh Postnatal Depression Scale
IVR: immersive virtual reality
OR: odds ratio
PPD: postpartum depression
RCT: randomized controlled trial
STAI: State-Trait Anxiety Inventory

Edited by A Kushniruk; submitted 24.01.25; peer-reviewed by L Aljerf, G Seguranyes; comments to author 26.02.25; revised version received 20.03.25; accepted 04.04.25; published 30.04.25.

Please cite as:
Jimenez-Barragan M, Del Pino Gutierrez A, Sauch Valmaña G, Monistrol O, Monge Marcet C, Pallarols Badia M, Garrido I, Carmona Ruiz A, Porta Roda O, Esquinas C, Falguera Puig G
Immersive Virtual Reality eHealth Intervention to Reduce Anxiety and Depression in Pregnant Women: Randomized Controlled Trial
JMIR Hum Factors 2025;12:e71708
URL: https://humanfactors.jmir.org/2025/1/e71708
doi:10.2196/71708
PMID:

©Marta Jimenez-Barragan, Amparo Del Pino Gutierrez, Gloria Sauch Valmaña, Olga Monistrol, Carme Monge Marcet, Mar Pallarols Badia, Ignasi Garrido, Anna Carmona Ruiz, Oriol Porta Roda, Cristina Esquinas, Gemma Falguera Puig. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 30.04.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Smartphone Apps and Wearables for Health Parameters in Young Adulthood: Cross-Sectional Study

Gaia Leuzzi[1,2], PhD; Mirko Job[1], PhD; Aldo Scafoglieri[2], PhD; Marco Testa[1], PhD

[1]Department of Neurosciences, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health, University of Genoa, Campus of Savona, Via Magliotto, 2, Savona, Italy

[2]Department of Physical Education and Rehabilitation, Experimental Anatomy Research Group (EXAN), Vrije Universiteit Brussel, Brussels, Belgium

Corresponding Author:
Marco Testa, PhD
Department of Neurosciences, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health, University of Genoa, Campus of Savona, Via Magliotto, 2, Savona, Italy

Abstract

Background: Fostering innovative and more effective interventions to support active aging strategies from youth is crucial to help this population adopt healthier lifestyles using technologies they are already familiar with. Mobile health (mHealth), particularly apps and wearables, represents a promising approach due to its versatility, ease of use, and ability to monitor multiple health variables simultaneously. Moreover, these devices offer opportunities for personalization and support in health behavior change, making them valuable tools for shaping healthy habits from a young age.

Objective: This study aims to (1) investigate whether young adults (18 - 26 years old) use apps or wearables to monitor or improve their health variables (ie, physical activity, diet, and mental health); (2) examine how they use them; (3) identify the most commonly used apps and wearables and the most frequently monitored health variables across these domains; and (4) evaluate the importance of different characteristics and functions of apps and wearables for health purposes.

Methods: This cross-sectional study used a public involvement framework to enhance the research quality and was conducted through an anonymous web survey disseminated across Italy over a 3-month period. The survey consisted of 5 sections: (1) demographics, (2) mobile apps and wearable devices for physical activity and sports, (3) mobile apps and wearable devices for diet, (4) mobile apps and wearable devices for mental health, and (5) preferences regarding mobile apps and wearable devices. Participants were eligible if they were young adults who reported using at least one app or wearable device to monitor at least one health variable (eg, steps, training, sleep, calorie intake). No additional eligibility criteria were applied.

Results: A total of 693 questionnaires were analyzed for aims 1 and 4, with the sample showing an equal gender distribution (females: 363/693, 52.4%). For aims 2 and 3, a total of 317 questionnaires were included. Participants using an app or wearable for physical activity accounted for 320 (46.2%), while 60 (8.7%) and 156 (22.5%) reported use for diet and mental health, respectively. Moreover, the frequency of use was predominantly on a daily basis, particularly for wearables. The app and wearable characteristics identified as most important were user-friendliness, free access to content, loading speed, and icon clarity.

Conclusions: Findings suggest that Italian young adults, particularly women, predominantly use wearables over apps to track health data, with both being checked on a daily basis. Physical activity is the most frequently monitored domain, likely due to its ease of tracking, while diet and mental health receive less attention. Overall, these tools are used more for monitoring than for actively improving health-related variables. The most valued characteristics identified by young adults include ease of use, free access to all content, and fast loading speed. These insights should guide the design and refinement of digital health interventions targeting this population.

(JMIR Hum Factors 2025;12:e64629) doi:10.2196/64629

KEYWORDS

exercise; mobile apps; diet; mental health; mHealth; apps; digital health; smartphones; wearables; parameters; young adulthood; adolescents; teenagers; young adult; cross-sectional study; public; involvement; interventions; active aging; strategies; healthy lifestyle; technology; physical activity

Introduction

Promoting healthy aging, or “active aging” [1], helps prevent excessive pressure on medical systems in the forthcoming years [2], while improving both individual and collective health [3]. This strategy combines theoretical and practical interventions to teach people, starting from youth, how to monitor and enhance health through a virtuous and sustainable lifestyle [4]. Reaching a large population in real time is crucial for the success

XSL•FO
RenderX

of this slow and consistent process. In the last decades, mobile health (mHealth) [5,6] interventions have provided general and personalized information to support decision-making and behavior change. In this context, mobile apps designed for smartphones are particularly popular among young people and adults [7], offering continuous data collection via built-in sensors and adaptability to different scenarios [8,9]. Given the rapid evolution of this technology, it is crucial to understand its use and effectiveness in creating sustainable and useful mHealth apps. To promote lasting behavioral changes for active aging, apps must specifically target young adults [10] and adapt to their needs. Educating young adults (18-26 years old) [11] about healthy lifestyles using familiar tools [12] encourages long-term adoption. Therefore, it is important to explore how apps and wearables for health purposes are used by this specific population to manage domains such as physical activity (PA), diet, and mental health. In particular, considering that health is defined as a state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity [13], these domains might play a role in health promotion for early active aging strategies.

The Health Belief Model (HBM) is a useful framework for understanding health-related behaviors [14]. It suggests that individuals are more likely to engage in health-promoting actions, such as using apps or wearables, if they perceive susceptibility to health issues, recognize their severity, and believe the benefits outweigh the barriers. Cues to action, such as app notifications, and self-efficacy in performing healthy actions further drive behavior change. By incorporating these elements, apps and wearables can motivate young adults to adopt healthy habits, supporting early active aging strategies.

In 2023, Italy had 5.3 million young adults, with about 1 million overweight or obese, 2.3 million not practicing any sports, and 2.4 million with chronic diseases [15]. Hence, it is fundamental not only to monitor but also to help this population effectively manage and improve its health. Moreover, in 2023, students enrolled in secondary high schools in Italy were reported to be 2.7 million, while 1.8 million were enrolled in universities [15].

Additionally, identifying the most and least appreciated features of these apps and wearables, along with their functions, is crucial for creating preventive and educational interventions [16] that assist young people in monitoring and improving their health [17-19], well-being, and quality of life through familiar technology tailored to their needs and expectations.

Therefore, this study aims to (1) determine whether young adults use mobile apps and wearable devices to monitor and improve their health variables (ie, PA, diet, and mental health); (2) understand how young adults use mobile apps and wearable devices to monitor and improve these health variables; (3) assess the most used mobile apps and wearables in terms of brands, as well as the most monitored variables via both apps and wearables; (4) evaluate the most important functions and characteristics of mobile apps and wearable for health variables based on young adults' opinion.

Methods

Study Design

This work consists of a cross-sectional study performed via an anonymous web survey distributed across the Italian territory. The survey was created in accordance with the International Handbook of Survey Methodology [20] and the Declaration of Helsinki [21], and was distributed via Microsoft Forms (Office 365 Suite), which ensures secure and ease of data collection, as well as anonymity in compliance with General Data Protection Regulation [22] policies. Before accessing the survey, each participant was presented with an informative note about the study and data processing.

Participation was completely voluntary, and completion of the survey could be interrupted at any time and for any reason, without the need for explanation; in such cases, no data were saved. The reporting of this manuscript follows the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines [23]. At each phase of the research, a habitual mobile app user was involved to ensure the "public involvement," as discussed below.

Ethical Considerations

Ethical approval for this study was granted by the University of Genoa's Ethical Committee for Research (CERA2023.24, approved on April 27, 2023). The study adhered to good clinical practice guidelines for the ethical conduct of research involving human participants. Participants were not compensated for their involvement, as clearly stated in the information sheet and informed consent form provided before their participation. All participants reviewed the study details and provided informed consent before taking part. Data were anonymized at the point of collection, as the use of Microsoft Fforms allowed participants to complete the survey without submitting personal information unless explicitly requested by the researchers. Data were collected and stored in accordance with current regulations, in a password-protected digital archive accessible only to the researchers directly involved in the study.

Web Survey

The web survey was created by a panel of 5 professionals from different fields, such as sports science and health, psychology, engineering, and physiotherapy, and a habitual mobile app user pursuing a master's degree in communication sciences. Existing questionnaires were reviewed to identify relevant questions, but none were deemed suitable for this specific research. Consequently, a preliminary draft of the survey was developed for evaluation and subsequent validation. The final web survey was divided into 5 sections: (1) demographics; (2) mobile apps and wearable devices for PA and sport; (3) mobile apps and wearable devices for food and diet; (4) mobile apps and wearable devices for mental health; and (5) preferences regarding mobile apps and wearable devices. Once imported into Microsoft Forms, the questions were organized with branching logic, allowing participants to skip sections on unused mobile apps or wearable devices, thereby enabling faster and more efficient completion. Following the demographics section, each subsequent section began with an initial yes-or-no question,

XSL•FO
RenderX

giving participants the option to either skip that section or proceed. The estimated total completion time was about 5 minutes. The full survey is available in Multimedia Appendix 1.

The web survey was assessed for face and content validity [24] by 10 potential participants who used mobile apps to monitor their PA, diet, or mental health. These individuals met the study's inclusion criteria and were asked to provide feedback on the survey before its adoption in the data collection process, to ensure data reliability and quality. Specifically, each participant had either a face-to-face meeting or a video call with the same professional, who explained the research aims, provided the informative note and informed consent documents, and, once consent was obtained, initiated the survey completion. Afterward, participants were asked questions about the survey's characteristics, and all notes and suggestions were recorded. The feedback received was subsequently discussed by the panel of experts to produce the final version of the web survey. Face and content validity assessments are crucial in survey creation to ensure that the main topic under investigation is correctly addressed and that the questions and answers are understandable to the target population. This ensures that the results can be accurately analyzed and compared with the research question. In particular, content validity is provided by the panel of experts in the field investigation, whereas face validity is assessed by potential participants. Additionally, to evaluate the internal consistency of the survey, a Cronbach α [25] analysis was performed. The results confirmed good internal consistency (Cronbach α=0.891), as values of 0.70 and above are generally considered acceptable. The scale ranges from 0 to 1, with 1 indicating the highest level of internal consistency. The higher the value obtained for the survey items, the better the internal consistency is considered to be. No other validation methods were applied to this survey.

Public Involvement

Our work employed the "study-focused" public-involvement framework [26], considering the point of view of the target population (ie, young adults aged 18 - 26 years using at least one app or wearable to monitor or improve PA, diet, or mental health) across all phases of the study [20]. The public representative was (1) a 23-year-old female student; (2) enrolled in the first year of a master's degree program in communication science; and (3) actively engaged with a mobile app for PA and possessing previous experience in using a mobile app for diet. Given the exploratory nature of this study, which aimed to investigate the prevalence of app and wearable use for health purposes among young adults as a whole, and not among specific subgroups (eg, by gender), a single representative, male or female, of the target population was deemed sufficient and was not expected to influence the results. The focus was on understanding general behaviors rather than subgroup-specific patterns, such as those based on gender.

The public representative volunteered to actively participate in formulating the research question and played a crucial role in creating, validating, and distributing the web survey. After the data collection period, she was also involved in the data analysis and the overall interpretation of the results, which was fundamental for better understanding the needs of the target population and maximizing the value of the collected data. Moreover, the public representative contributed to the writing and review of the final draft of this paper.

Participants

To answer the first and last aims (1 and 4) of the study, participants were considered eligible if they were young adults (ie, aged 18 - 26 years) and could read and understand the informative note and the web survey.

For aims 2 and 3, participants were eligible if they used at least one mobile app or wearable device for PA, diet, or mental health, and could read and understand the informative note and the web survey.

No other limitations were set for the population or for the type of app or wearable device used.

Setting and Modalities of Contact

The web survey was distributed via Microsoft Forms (Office 365 Suite) in July and from September to October 2023. Potential participants were reached through various modalities, including face-to-face, digital, and printed invitations (see Multimedia Appendix 2). Printed advertisements were placed in locations frequently visited by young adults, both education-related and not (eg, university campuses, local groups headquarters, libraries). Social media platforms (eg, Facebook, Instagram, WhatsApp), email, and high school and university mailing lists were also used to disseminate invitations to complete the survey. To maximize nationwide reach, all 73 universities in Italy were contacted. Of these, only 7 agreed to assist in distributing the survey. The list of universities was publicly available online, and each university's website was visited to obtain a contact email to explain the study rationale and request assistance with its promotion. The universities and high schools that agreed to collaborate were free to choose the dissemination method they considered most appropriate. To help balance the potential sample of participants, voluntary and social groups were also invited to share the study invitations. Concerning email, researchers used addresses that were publicly available in the "Contacts" section of relevant websites to reach potentially interested groups. Additionally, participants were encouraged to share the invitation with others who might be interested in participating. The recruitment process aimed to reach a diverse and broad audience, including individuals both engaged and not engaged in education, and from different regions of the country. Efforts were made to maintain a balance in region of origin, gender representation, and age distribution within the predefined age range. As the web survey was anonymous, and to prevent multiple responses, the option to submit more than 1 questionnaire was disabled directly in the platform settings. To further ensure data reliability, the dataset was checked for duplicates in Microsoft Excel after data collection. In cases of potential duplicates, variables such as age, gender, height, weight, and education were examined; no duplicates were found.

XSL•FO
RenderX

Statistical Analysis

Sociodemographic data were analyzed using descriptive statistics, with percentages reported where applicable, and continuous variables presented as mean (SD). Binomial logistic regression analyses were performed to assess the influence of gender and education level on the use of apps and wearables. The last section of the questionnaire was analyzed using different methods. The first method defined consensus as being reached for each statement when 2 conditions were met [27]: (1) at least 51% of participants responded with "quite important" or "very important," and (2) the median score exceeded 2.5. In our case, considering the IQRs, if the median was above 2.5, it could be assumed that at least 75% of the responses fell in the neutral or positive categories (ie, 2, 3, and 4). The second method involved the calculus of a ratio value between positive and negative answers collected. Moreover, a descriptive evaluation of results obtained from the Likert-type scale was adopted.

For evaluating the main variables monitored through mobile apps and wearables, the variables were divided into specific areas of interest (ie, PA, diet, and mental health) and aggregated into categories based on their similarities. PA-related variables were divided into "PA metrics" (ie, daily steps, distance covered, daily activity, training, number of weekly trainings, calorie consumption, elevation, and speed/step per kilometer) and "physiological parameters" (ie, heart rate, weight, oxygen saturation, sleep, stress, rest, body composition, menstrual cycle monitoring, and respiratory frequency). Diet parameters were grouped into "nutritional metrics" (ie, calorie consumption, macronutrient count, and water) and "body metrics" (ie, weight and body composition). Finally, mental health parameters were categorized as "sleep," "mental well-being" (ie, stress, mood, and emotions management), and "mindfulness and meditation practices" (ie, breathing, meditation, mindfulness). Additionally, the motivation behind the use of these mobile apps and wearables was presented, with results split between the 2 types of mHealth.

The last section of the questionnaire investigated mobile app and wearable preferences, and the results of each question were graphically reported. Participants were asked to indicate their level of agreement with each statement on a 5-point Likert-type scale. The possible answers were as follows: (0) "not important at all," (1) "unimportant," (2) "neutral," (3) "quite important," and (4) "very important". The results were then aggregated into 3 levels of agreement: negative (ie, 0 and 1), neutral (ie, 2), and positive (ie, 3 and 4). Moreover, an analysis of IQRs was performed, and medians and the distribution of responses were reported.

Bias

Young adults with different levels of education were contacted to reduce any possible selection bias. To better describe the target population, invitations were disseminated across the entire Italian territory without restrictions, with the aim of reaching participants with diverse characteristics such as age, gender, occupation, and education level. Other universities and several secondary high schools were also contacted to increase the number of responses. Detailed modalities of dissemination are reported in the "Setting and Modalities of Contact," and study limitations are discussed in the "Strengths and Limitations" section.

Results

A total of 821 responses were received, and after applying the inclusion and exclusion criteria, a final sample of 693 questionnaires was analyzed. Specifically, 42 questionnaires were excluded because participants did not consent to data treatment, 83 did not meet the study's age range (ie, 18 - 26), and 3 were excluded due to unusable data. Demographic data for the final sample are presented in Table 1.

XSL•FO
RenderX

Table . Demographic characteristics of the sample (n=693).

Characteristics	Values
Age (years), mean (SD)	21.2 (2.5)
Gender, n (%)	
Female	363 (52.4)
Male	329 (47.5)
Nonbinary	1 (0.1)
BMI (kg/m^2), mean (SD)	22.3 (3.03)
Last educational degree, n (%)	
Secondary school diploma	95 (13.7)
High school diploma	414 (59.7)
Higher education	184 (26.6)
Not using any mobile app or wearable, n (%)	317 (45.7)
Physical activity (n=144), n (%)	
1 app	106 (73.6)
2 or more apps	38 (26.4)
Diet (n=52), n (%)	
1 app	45 (86.5)
2 or more apps	7 (13.5)
Mental health (n=45), n (%)	
1 app	37 (82.2)
2 or more apps	8 (17.8)

Data regarding the use of mobile apps or wearables for PA, diet, or mental health variables are reported in Table 2 and include only those participants who reported using at least one app or wearable to monitor at least one health variable (n=376).

Table . Total number and percentage of participants using mobile apps or wearables for physical activity, diet, and mental health (n=693)..

Variables	Yes	No	Eliminated[a]
Physical activity, n (%)	320 (46.2)	338 (48.8)	35 (5.1)
App and wearable, n	56 (8.1)	—[b]	—
Only app, n	88 (12.7)	—	—
Only wearable, n	176 (25.4)	—	—
Diet, n (%)	60 (8.7)	622 (89.8)	11 (1.6)
App and wearable, n	15 (2.2)	—	—
Only app, n	37 (5.3)	—	—
Only wearable, n	8 (1.2)	—	—
Mental health, n (%)	156 (22.5)	524 (75.6)	13 (1.9)
App and wearable, n	19 (2.7)	—	—
Only app, n	26 (3.8)	—	—
Only wearable, n	111 (16.0)	—	—

[a]Data from each section were eliminated if participants provided conflicting answers within that section.

[b]Not available.

Table 3 ranks and details the most used mobile apps and wearables by brand and further divides them into specific spheres of interest (ie, PA, diet, and mental health).

XSL•FO
RenderX

Table . Ranking by brand of the most used mobile apps and wearables (n=376).

App	Physical activity (n=320)		Diet (n=60)		Mental health (n=156)	
	App (n=144)	Wearable (n=232)	App (n=52)	Wearable (n=23)	App (n=45)	Wearable (n=130)
1, n (%)	• Apple Health: 62 (43.0)	• Apple Watch: 81 (35)	• Yazio: 15 (28.8)	• Apple: 8 (34.8)	• Apple Health: 7 (15.5)	• Apple: 39 (30)
2, n (%)	• Samsung Health: 25 (17.4)	• Xiaomi: 37 (16)	• MyFitnessPal: 11 (21.1) • FatSecret: 11 (21.1)	• Xiaomi: 7 (30.4)	• Samsung Health: 6 (13.3) • Google Fit: 6 (13.3) · • FitBit: 6 (13.3)	• Amazfit: 35 (27)
3, n (%)	• WeWard: 18 (12.5)	• Amazfit: 25 (10.7) • Garmin: 25 (10.7)	• Samsung Health: 7 (13.5)	• Samsung: 4 (17.6)	• Calm: 5 (11.1)	• Samsung: 12 (9.2) • Garmin: 12 (9.2)
4, n (%)	• Google Fit: 14 (9.7)	• Samsung: 20 (8.6)	• Apple Health: 5 (9.61)	• Fitbit: 1 (4.3) • Garmin: 1 (4.3) • Huawei: 1 (4.3) • Redmi: 1 (4.3)	• Garmin Connect: 4 (8.9) • Zepp Life: 4 (8.9) • Headspace: 4 (8.9)	• Huawei: 10 (7.6)
5, n (%)	• Garmin Connect: 11 (7.6)	• Fitbit: 12 (5.2) • Huawei: 12 (5.2)	• Macros: 3 (5.8) · • Melarossa: 3 (5.8)	• N/A[a]	• Flo: 3 (6.7) • Serenity: 3 (6.7)	• Fitbit: 9 (6.8)
6, n (%)	• Zepp Life: 9 (6.2)	• Honor: 5 (2.2)	• Lifesum: 2 (3.8)	• N/A	• N/A	• Honor: 5 (3.8)
7, n (%)	• Komoot: 7 (4.9)	• Polar: 4 (1.7)	• Google Fit: 1 (1.9) · • Fitdays: 1 (1.9) • Lose it: 1 (1.9)	• N/A	• Daylio: 1 (2.2) · • Stoic: 1 (2.2) · • Mindshift: 1 (2.2)	• Google: 1 (0.8) • Polar: 1 (0.8) • Oura: 1 (0.8) · • Real me: 1 (0.8) • Liujo: 1 (0.8) • Suunto: 1 (0.8) • Whoop: 1 (0.8) • Xiaomi: 1 (0.8)
Others, n (%)	• A1[b]: 52 (36.1)	• A2[c]: 11 (4.7)	• N/A	• N/A	• N/A	• N/A

[a]N/A: not applicable.

[b]A1: SweatCoin, Strava, Fitbit, iSkiTracker, Pedometer, MyWellness, My workout plan, Huawei health, Hevy, Relive, Pokemon Go, Fitdays, Adidas Running, Nike training, GloryFit, Nike Run, etc.

[c]A2: Oppo, Google, Real me, Redmi, Suunto, and Liujjo.

Furthermore, the frequency of use of both mobile apps and wearables was evaluated and is presented in Table 4.

XSL•FO
RenderX

Table . Frequency of use of mobile apps and wearables.

Variables	Physical activity		Diet		Mental health	
	Mobile app	Wearable	Mobile app	Wearable	Mobile app	Wearable
Total users, n	144	232	52	23	45	130
Every day, n (%)	53 (36.8)	148 (63.8)	31 (59.6)	16 (69.6)	16 (35.6)	91 (70.0)
5 or 6 times a week, n (%)	20 (13.9)	24 (10.3)	10 (19.2)	3 (13.0)	4 (8.9)	12 (9.2)
3 or 4 times a week, n (%)	31 (21.5)	27 (11.6)	2 (3.8)	2 (8.7)	8 (17.8)	10 (7.7)
1 or 2 times a week, n (%)	20 (13.9)	22 (9.5)	4 (7.7)	1 (4.3)	9 (20.0)	8 (6.2)
Less than once a week, n (%)	20 (13.9)	11 (4.7)	5 (9.6)	1 (4.3)	8 (17.8)	9 (6.9)

Moreover, motivation for the use of apps and wearables, as well as the health variables monitored via these devices, is reported in Multimedia Appendix 3. The vast majority of the sample did not report using premium or paid content (310/320, 96.9%, participants monitoring PA; 58/60, 96.7%, monitoring diet; 152/156, 97.4%, monitoring mental health). Those who did use these additional functionalities primarily sought more detailed information on specific aspects such as sleep, rest periods, training, macro- and micronutrients, or personalized advice based on their data. Similarly, community functions were not used by most participants (289/320, 90.3%, monitoring PA; 54/60, 90%, monitoring diet; 142/156, 91%, monitoring mental health; see Multimedia Appendix 4)

Additionally, an analysis of the frequency distribution of answers based on gender and level of education was conducted. Binomial logistic regressions were also performed on these parameters to explore possible associations between the use of apps and wearables and gender or education level. Regarding gender, the percentage of females monitoring PA is higher than that of males (216/693, 31.2% vs 138/693, 19.9%), while nonusers are 191 out of 693 (27.6%) for males and 147 out of 693 (21.2%) for females. A slight difference was also found in diet monitoring, with 325 out of 693 (46.9%) females and 297 out of 693 (42.9%) males reporting no monitoring. For mental health, gender differences were observed, with 105 out of 693 (15.2%) females and 64 out of 693 (9.2%) males reporting monitoring. Logistic regressions examining gender influence on the monitoring of PA, diet, and mental health revealed a slight positive correlation for males in each sphere. Specifically, the models account for 2.4% for PA, 1.0% for diet, and 1.1% for mental health. Full results of these binomial logistic regressions are reported in Multimedia Appendix 5. Regarding education level, the frequency distribution for females and males differed across categories: (1) secondary school diploma (females: 23/693, 3.3%; males: 72/693, 10.4%), (2) high school diploma (females: 228/693, 32.9%; males: 185/693, 26.7%), (3) bachelor (females: 100/693, 14.4%; males: 4/693, 0.6%), and (4) master's degree (females: 12/693, 1.7%; males: 4/693, 0.6%). Logistic regression models explained 5.0% of the variance for PA, 0.5% for diet, and 2.4% for mental health. For PA, having a bachelor's or master's degree showed a slight negative influence on the use of apps or wearables, and the same pattern was observed for mental health, while for diet, only a master's degree has a slight negative influence on use. Full results of these analyses are reported in Multimedia Appendix 5. Finally, none of the differences related to gender or education level were statistically significant (see Multimedia Appendix 5).

For the fifth and final part of the questionnaire, all 693 participants were included in the analysis. Considering the assessment of the most important characteristics of mobile apps and wearables, 4 elements emerged as particularly important for young adults: (1) user-friendliness, (2) free access to all content, (3) loading speed, and (4) icon clarity. Figure 1 presents the aggregated responses for each question according to the level of attributed importance.

XSL•FO
RenderX

Figure 1. Mobile apps and wearables' functions preferences (n=693).

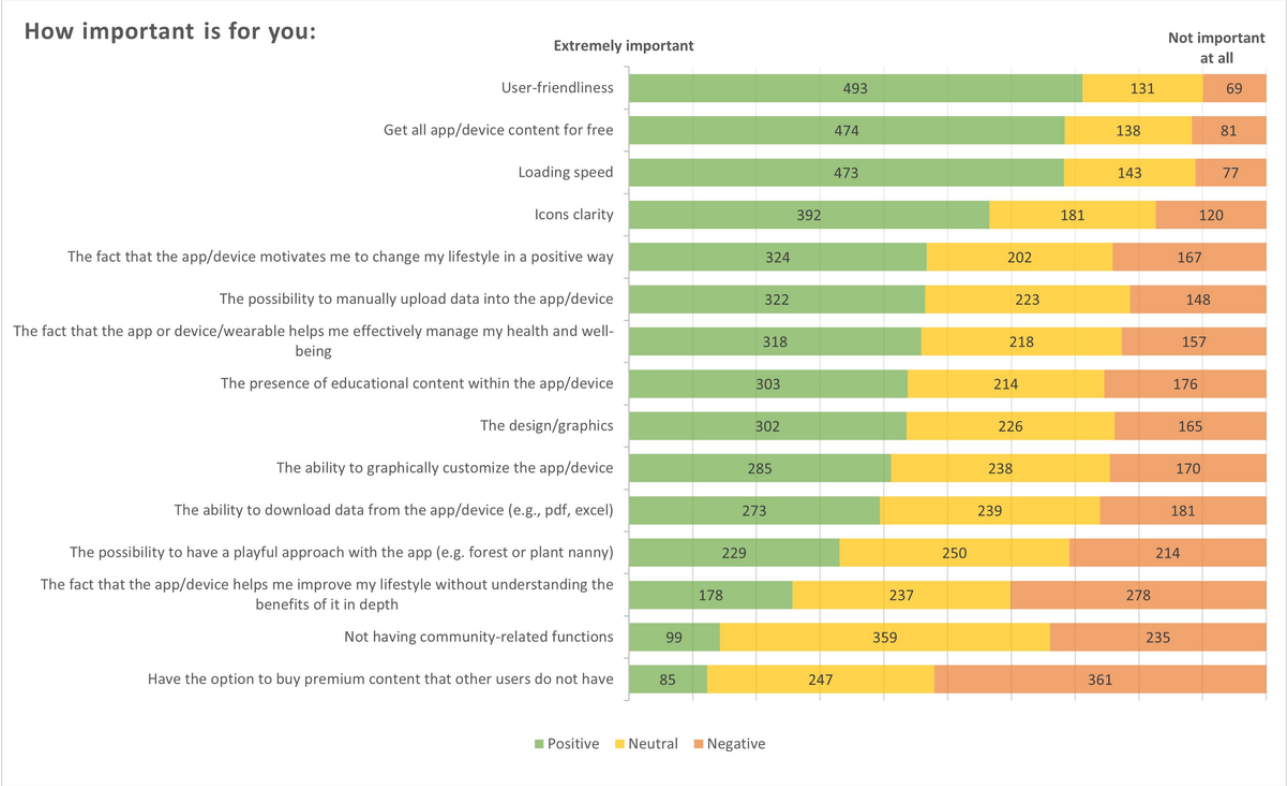
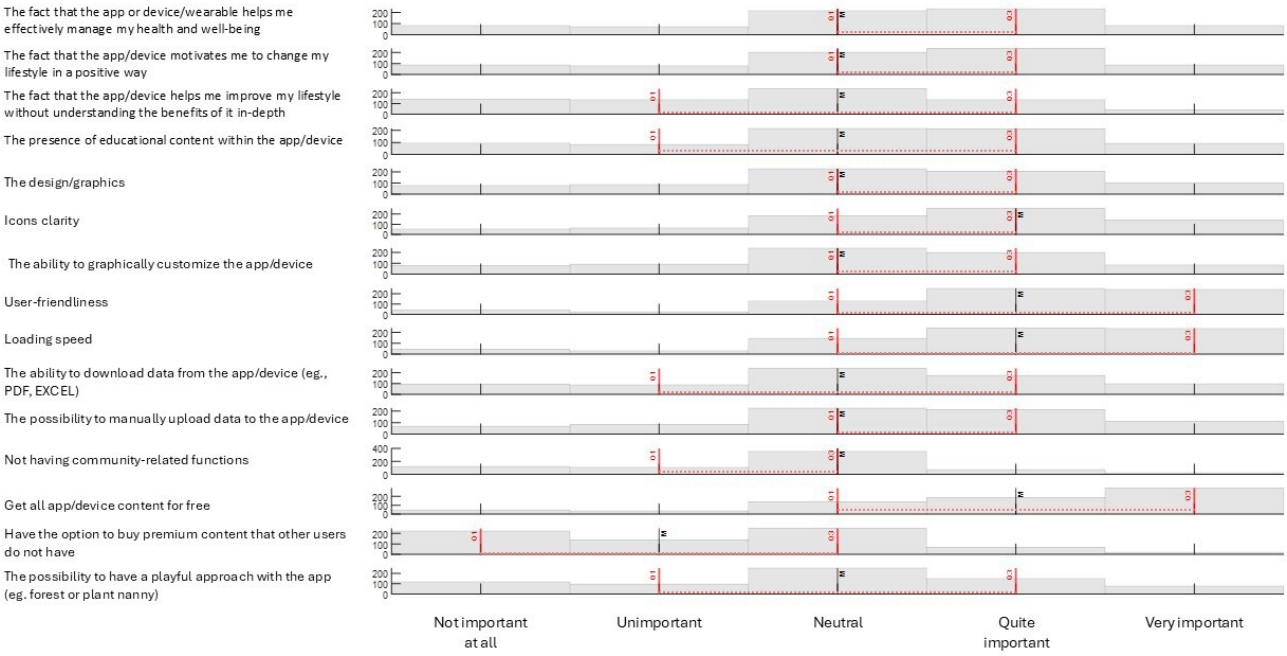


Figure 2 shows the distribution of responses, including IQRs and medians. Results are displayed for each of the original 5-point Likert scale options available to participants.

Figure 2. IQRs, medians, and distribution of responses for mobile apps and wearables' functions preferences.



A second method of analysis was applied to the last section of the questionnaire, particularly to the data regarding elements considered important. The analysis evaluated the ratio of positive to negative answers, both with and without including neutral responses, as reported in Table 5.

XSL•FO
RenderX

Table . Ratio values of responses for mobile apps and wearables' functions preferences.

Device's characteristics	Positives/negatives	positives+neutralsnegatives+neutrals
User-friendliness	7145	3120
Get all app/device content for free	5852	2800
Loading speed	6143	2795
Icons clarity	3267	1904
The fact that the app/device motivates me to change my lifestyle in a positive way	1940	1469
The possibility of manually uploading data into the app/device	2176	1429
The fact that the app or device/wearable helps me effectively manage my health and well-being	2025	1425
The presence of educational content within the app/device	1722	1350
The design/graphics	1830	1326
The ability to graphically customize the app/device	1676	1282
The ability to download data from the app/device (eg, PDF, Excel)	1508	1219
The possibility to have a playful approach with the app (eg, forest or plant nanny)	1070	1032
The fact that the app/device helps me improve my lifestyle without understanding the benefits of it in depth	0.640	0.806
Not having community-related functions	0.421	0.771
Have the option to buy premium content that other users do not have	0.235	0.546

From this analysis, it can be observed that user-friendliness, loading speed, free access to all content, and icon clarity were confirmed as the most important elements for our sample. Moreover, the possibility of manually uploading data and the efficiency of the app or device in helping individuals manage their health were also indicated as potentially important. When neutral responses are also considered, the elements with the highest ratio values remain user-friendliness, free access to all content, loading speed, and icon clarity. Other ratio values were lower and may not warrant further investigation.

Discussion

Principal Findings

This study comprehensively investigated the use of apps and wearables to track and improve health variables among young adults. The average BMI of the sample fell within the normal range, consistent with national trends. In 2023, the Italian National Institute of Statistics reported that the majority of young adults, approximately 2.8 million individuals, had a BMI within the normal range [15]. While nearly half of the participants reported using these tools to track PA, only a minority used them to monitor diet or mental health. Conversely, almost half of the sample reported not using any app or wearable for health purposes. This might be due to a low perceived relevance of health monitoring, negative aspects associated with using such tools (eg, distress or fixation), or a lack of understanding or trust in these devices and the results they provide [28,29]. Another main issue in using apps and wearables for health is the manual input required for certain data, which leads to lower adherence and reduced usage of these tools [30-32]. The findings regarding diet are consistent with those reported by Hahn et al [33] in a study conducted on a comparable population. Similarly, the reported percentage of use for mental health apps aligned with previous literature [34,35], except for Borghouts et al [35], who reported a higher usage rate. Viewed through the lens of the HBM, PA may be the most frequently monitored health domain due to the greater perceived benefits of tracking it and the immediate, tangible feedback provided by step count, distance covered, or calories burned. In line with this, the perceived severity and susceptibility of problems related to sedentary behavior (eg, weight gain) may also serve as motivating factors. By contrast, monitoring diet and mental health presents greater barriers, as it often requires manual data entry [30], thereby reducing engagement, self-efficacy, and both the frequency and quality of data collection [31,32]. In particular, diet variables cannot be automatically monitored, especially from wearables. For mental health, automated tracking of variables such as sleep or stress through heart rate variability [36] provides an exception, as it requires less user interaction, fostering higher self-efficacy by simplifying behavior monitoring. To increase diet and mental health monitoring, digital assistants (eg, vocal assistants), artificial intelligence, and machine learning could be

XSL•FO
RenderX

implemented into apps and wearables to automate tasks that are now manually required [37], lowering the need for users' interaction and broadening the adoption also to other populations with lower digital skills.

The higher prevalence of app and wearable use among females, particularly for PA and mental health, might also be viewed through the HBM lens. Women's greater focus on prevention could reflect a higher perceived susceptibility and severity of health risks, as well as stronger cues to action, such as societal expectations and health information targeting females [38]. In line with this, a previous study highlighted that females were more likely to use well-being apps to improve their health, whereas males were more interested in tracking their variables and tended to lose interest in these apps more quickly [39]. Moreover, factors influencing young adults' adoption and use of apps for health purposes are shaped by socioeconomic status and educational level, creating discrepancies and barriers for individuals with low socioeconomic status and lower educational levels, even when they are interested in using them [40].

In contrast with other studies [30,41,42], a very high frequency of use was found among the majority of our sample (148/232, 63.8%, users of wearables for PA), who reported using mobile apps and wearables to measure their health variables on an everyday basis. According to the HBM, this preference may be due to the immediate perceived benefits of short-term feedback, which reinforces behavior change. However, 38 out of 144 (26.4%) participants reported the need for multiple apps to track PA, highlighting a potential barrier: fragmented tools may lower perceived practicality and increase user burden. Integrating comprehensive solutions that address all necessary variables could enhance adoption and self-efficacy. Moreover, it can be argued that the wider adoption of apps for PA might be due to their greater market presence, supported by the widespread promotion of well-known brands such as Nike or Puma. The daily use of apps and wearables may indicate their potential in health interventions that require frequent monitoring over time, as well as the possibility of tracking small daily results to keep users motivated toward long-term goals.

Survey respondents demonstrated a clear preference for wearables over apps, especially for PA and mental health tracking. According to HBM's self-efficacy principle, the automatic data collection provided by wearables reduces the effort required for health monitoring. This ease of use increases confidence and adherence, as wearables simplify behavior tracking [43] while offering precise data [44]. Nonetheless, further analysis of the response ratios revealed a positive interest in the option of manually uploading data, which could represent an additional personalization feature of apps or wearables. Conversely, diet monitoring remains underutilized due to barriers such as manual input requirements, which reduce both perceived benefits and self-efficacy. Additionally, in our sample, wearables were also used to monitor sleep, mindfulness, and meditation practices. The latter is an interesting aspect, as another study [45] reported the use of a wearable to assess meditation practices, but only to identify abrupt movements during meditation rather than to monitor variables such as breathing or heartbeat. By contrast, participants in our sample reported using wearables to monitor actual meditation and mindfulness practices.

Comparison With Prior Studies

Regarding the motivations for use in our population, apps and wearables were mostly adopted for monitoring health variables and, additionally, for receiving personal advice based on individual data collected. These results align with previous literature [31] on young populations [43], which also reported a similar rationale behind the use of technological tools to support health. Considering these motivations and the HBM theory, the importance of cues to action and personal advice emerges. Hence, based on the results obtained, health interventions for young adults might benefit from the use of apps and wearables to monitor health variables while also providing personalized advice to support daily tasks aimed at improving health outcomes over time, particularly in the PA domain. A study conducted on university students showed that PA apps, in particular, are valued and considered useful; therefore, people are more likely not only to use them but also to promote their use [46], further supporting the hypothesis that apps may be beneficial in long-term health interventions beginning in youth.

Characteristics of mobile apps and wearables identified as important by the respondents were (1) user-friendliness, (2) free access to all content, and (3) fast loading speed of the app/wearable and its contents. These elements are key enablers of technology adoption and should therefore be carefully addressed by app and wearable developers [47]. User-friendliness, in particular, may increase self-efficacy, tool effectiveness [48], and young people's adoption of technology [49]. Previous evidence has already highlighted the importance of user-friendliness for young populations, especially in relation to fitness apps [50]. Moreover, free content may reduce financial barriers to access for people with low socioeconomic positions, thereby enabling the reach of broader populations [51,52], especially considering that app prices can vary depending on the specific app and market [53]. In line with this assumption, almost all participants were not interested in purchasing premium content. Interestingly, all the abovementioned aspects relate only to device characteristics and do not consider the specific types of content provided (ie, educational, motivational, informative) [38]. Regarding loading speed, this feature has been emphasized in previous literature [47,54,55] as a crucial element for technology acceptance, as users typically spend very little time deciding whether to use an app on their devices [56]. Hence, to ensure young adults' use of apps and wearables in health interventions, the choice should favor fast-loading tools that are inexpensive or free, to minimize reported barriers.

Interestingly, only half of the sample acknowledged the importance of including educational content within apps [57,58], as well as the importance of technology assisting users in understanding the benefits of improving health variables [47]. Similar results were observed for the motivational aspect: half of the users considered it important that apps or wearables support them in improving their health. This may be considered controversial, given the nature of the apps and wearables investigated and previous studies reporting a positive attitude

XSL•FO
RenderX

toward the technology's supportive role in improving health variables [40,59]. Nonetheless, this finding is consistent with responses on motivations for use, which highlighted that the primary function of apps and wearables was perceived as monitoring health variables. While the HBM emphasizes that understanding the benefits of healthy behaviors can motivate action, these results indicate a need to better communicate such benefits through engaging, user-centered app features.

Regarding the aesthetics of apps and wearable devices, the overall design and graphical aspects were not considered as significant as the clarity of the icons [52,53]. This particular element has been highlighted as a crucial factor in attracting users to the technology [60]. Additionally, the possibility of personalizing the app or wearable interface was deemed quite important, in line with existing evidence [61-63]. Such personalization may enhance users' perceived self-efficacy and reduce barriers to adoption, making the technology more adaptable to individual needs.

Despite evidence supporting gamification strategies for enhancing engagement and health outcomes [64], only one-third of our respondents expressed interest in playful features. This may suggest that young adults already feel sufficiently motivated to use these tools, reducing the perceived need for additional cues to action. A deeper analysis of the response ratios highlighted the significance of community functions. Previous literature has reported these features as a positive strategy for promoting challenging behaviors and motivation [65], serving as optional cues to action in health interventions. However, while participants in our study acknowledged community functions as important, they also reported not using them, raising questions about their actual role in motivating users to achieve their goals. In the literature, community features are often described as useful and appreciated in motivating individuals to improve their behaviors, particularly among young women [39], but mostly when they can view and compare results with people they personally know [50,66]. Nonetheless, a scoping review reported mixed results regarding their effectiveness in improving PA outcomes, largely due to the poor methodological quality of the studies evaluated [67]. Therefore, despite some positive results, further research is needed to strengthen the evidence on the use of community and social features in health programs targeting PA [68]. These features could be incorporated into health apps and wearables as optional tools, allowing users to decide whether to engage with them based on individual preferences. Thus, considering both the results obtained from our study and those of previous research, it would be useful for future studies to further investigate the effectiveness of community and social functions in young adult populations through both qualitative and quantitative approaches.

Given that mobile apps and wearables are frequently used by young adults for monitoring purposes, they could be integrated as valuable tools in strategies for early active aging [69,70]. These technologies may support the monitoring of health variables while also educating individuals about the importance of improving them. To further promote adoption and more efficient use of apps and wearables for health purposes, it is important to involve young adults in their design and development processes [63,71,72]. Finally, clear guidelines should be provided on how to develop effective health-related apps and wearables [73].

Strengths and Limitations

A strength of this study is the specific population investigated and the broad inclusion criteria adopted, which allowed for a more accurate description of the patterns of young adults who use apps and wearables for health purposes. Moreover, the broad inclusion criteria enabled heterogeneity in participants' characteristics, supporting the generalization of findings to the broader population. Another strength was the large number of questionnaires collected, which ensured sufficient statistical power to perform different types of analysis on the data and even examine subgroups (eg, gender). Despite this being the first study conducted in Italy on this population, some limitations should be acknowledged. First, the public representative who participated in the study was a student with prior professional experience in different roles (eg, social media manager, press officer). While her status as a student may have been an influencing factor, her communication skills and educational background provided a privileged perspective, making her contribution appropriate and constructive rather than biased. However, it cannot be excluded that her high level of education may limit the representativeness of individuals with lower educational backgrounds. Second, although this study investigated commercial mobile apps, the apps mentioned by participants were predominantly from renowned tech brands, with only a limited number of independent, nonprofit-oriented apps being reported. Future studies should aim to include a larger number of users of apps and wearables not linked to major brands to capture a broader range of opinions on different types of health tools. Third, data collection was conducted exclusively through digital procedures, which may have posed a barrier for individuals with poor digital literacy. However, given the young age of participants and the nature of the topic under investigation, such barriers were considered unlikely to represent a significant hindrance. Fourth, it cannot be excluded that, due to the digital nature of data collection, our sample may have been skewed toward individuals more comfortable with technology. Fifth, it is also possible that participants who completed the questionnaire were more interested in the topic investigated, which could have introduced distortion in the results. Sixth, the final sample included a high proportion of students, which may have influenced the findings; including more young adults who are not students might yield different results. Seventh, despite researchers' efforts to ensure the reliability of responses (as explained in the "Methods" section), the anonymous nature of the web survey could have introduced inaccuracies and self-reporting bias. In particular, participants may have overestimated their healthy behaviors (eg, social desirability bias). However, given the large number of responses and the nature of the topic, it is unlikely that responses were falsified. This likelihood is further supported by the study's aim, which focused on investigating the prevalence of app and wearable use rather than assessing their effectiveness. Eight, as it fell outside the study's aims, objective health outcomes were not directly assessed in relation to app or wearable usage; instead, only the outcomes reported and collected by apps and

XSL•FO
RenderX

wearables were considered, given that this was an exploratory study aimed at assessing the prevalence of use in young adults. Finally, socioeconomic status and geographic location data were not collected, making it impossible to examine their potential interactions with other variables. These aspects should be addressed in future research.

Future Directions

Further studies should assess the motivation, expectations, and user experience associated with apps and wearables for health variables, exploring other aspects that young adults might find helpful in improving their health. Barriers and facilitators to the use of these technologies in young adults should also be investigated to strengthen the evidence base and enhance their adoption. Moreover, future research should involve a larger proportion of individuals not engaged in educational programs and include those with weaker digital skills. Similar prevalence assessments should also be conducted in other age groups, such as adults and older adults. Finally, studies targeting health in young adults, particularly improvements in PA, should prioritize wearables over apps.

Conclusions

This study highlighted that young Italian adults track health-related variables more often via wearables than through apps, although both are consulted daily. Based on the results, young adults, particularly females, predominantly monitor PA-related variables, as these can be more easily tracked, while mental health and diet are monitored less. Moreover, apps and wearables are mainly adopted for monitoring rather than for improving health variables. Regarding the most important characteristics and features of these tools for young adults, user-friendliness, free access to all content, and the loading speed of the app/wearable software and content emerged as key. These characteristics should therefore be considered when developing or optimizing health interventions supported by apps and wearables in this population.

Acknowledgments

This work was carried out with the kind collaboration of the following institutions: Università degli Studi di Genova, Liceo Don Bosco Alassio, Liceo Statale Calasanzio, Istituto Ferraris Pancaldo Savona, Università degli Studi G. d'Annunzio Chieti, and Università di Parma, as well as Professors Fabrizio Sors and Cristiana Simonetti. We sincerely thank Stella Leuzzi for her support in pursuing this work. ChatGPT (OpenAI, Inc) was used solely for grammar and spelling checks in the manuscript. It was not involved in the study design, data collection, data analysis, generation of ideas for any section of the manuscript, or the creation of data visualizations. All participants provided consent for publication.

Data Availability

Data available on reasonable request from the authors.

Authors' Contributions

GL, MJ, and MT contributed to study conceptualization; GL and MJ to methodology. GL conducted the investigation and data extraction. GL, MJ, and AS performed the analysis. GL drafted the original manuscript, and GL, MJ, AS, and MT contributed to reviewing and editing. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Survey.
[DOCX File, 23 KB - humanfactors_v12i1e64629_app1.docx]

Multimedia Appendix 2
Study advertisement.
[PNG File, 680 KB - humanfactors_v12i1e64629_app2.png]

Multimedia Appendix 3
Number of parameters monitored for physical activity, diet, and mental health; motivations for use of apps and wearables for physical activity, diet, and mental health.
[JPG File, 77 KB - humanfactors_v12i1e64629_app3.jpg]

Multimedia Appendix 4
Use of premium contents and community functions.
[DOCX File, 15 KB - humanfactors_v12i1e64629_app4.docx]

XSL•FO
RenderX

Multimedia Appendix 5
Binomial logistic regression results.
[DOCX File, 41 KB - humanfactors_v12i1e64629_app5.docx]

References

1. Walker A. A strategy for active ageing. Int Social Security Review 2002 Jan;55(1):121-139. [doi: 10.1111/1468-246X.00118]
2. Cai T, Verze P, Bjerklund Johansen TE. The quality of life definition: where are we going? Uro 2021 Feb;1(1):14-22. [doi: 10.3390/uro1010003]
3. Active ageing: a policy framework. World Health Organization. 2022. URL: http://www.who.int/hpr/ [accessed 2023-01-31]
4. Walker A, Maltby T. Active ageing: a strategic policy solution to demographic ageing in the European Union. Int J Soc Welfare 2012 Oct;21(s1):S117-S130. [doi: 10.1111/j.1468-2397.2012.00871.x]
5. Park YT. Emerging new era of mobile health technologies. Healthc Inform Res 2016 Oct;22(4):253-254. [doi: 10.4258/hir.2016.22.4.253] [Medline: 27895955]
6. Robbins TD, Lim Choi Keung SN, Arvanitis TN. e-Health for active ageing: a systematic review. Maturitas 2018 Aug;114(34–40):34-40. [doi: 10.1016/j.maturitas.2018.05.008]
7. Klonoff DC, King F, Kerr D. New opportunities for digital health to thrive. J Diabetes Sci Technol 2019 Mar;13(2):159-163. [doi: 10.1177/1932296818822215]
8. Muro-Culebras A, Escriche-Escuder A, Martin-Martin J, et al. Tools for evaluating the content, efficacy, and usability of mobile health apps according to the consensus-based standards for the selection of health measurement instruments: systematic review. JMIR Mhealth Uhealth 2021 Dec 1;9(12):e15433. [doi: 10.2196/15433] [Medline: 34855618]
9. Llorens-Vernet P, Miró J. Standards for mobile health-related apps: systematic review and development of a guide. JMIR Mhealth Uhealth 2020 Mar 3;8(3):e13057. [doi: 10.2196/13057] [Medline: 32130169]
10. Bonnie RJ, Stroud C, Breiner H. Investing in the Health and Well-Being of Young Adults: National Academies Press; 2015. [Medline: 25855847]
11. Committee on Improving the Health, Safety, and Well-Being of Young Adults, Board on Children, Youth, and Families, Institute of Medicine, National Research Council. Introduction. In: Bonnie RJ, Stroud C, Breiner H, editors. Investing in the Health and Well-Being of Young Adults: National Academies Press; 2015. URL: https://www.ncbi.nlm.nih.gov/books/NBK284791 [accessed 2025-08-26]
12. Ge Y, Xin S, Luan D, et al. Association of physical activity, sedentary time, and sleep duration on the health-related quality of life of college students in Northeast China. Health Qual Life Outcomes 2019 Jul 16;17(1):124. [doi: 10.1186/s12955-019-1194-x] [Medline: 31311564]
13. Schramme T. Health as complete well-being: the WHO definition and beyond. Public Health Ethics 2023 Nov;16(3):210-218. [doi: 10.1093/phe/phad017] [Medline: 38333767]
14. Sheng J, Gong L, Zhou J. Exercise health belief model mediates the relationship between physical activity and peer support among Chinese college students: a cross-sectional survey. Front Psychol 2023;14:1103109. [doi: 10.3389/fpsyg.2023.1103109] [Medline: 36814667]
15. Istat. URL: https://www.istat.it/ [accessed 2024-12-16]
16. Gordon WJ, Landman A, Zhang H, Bates DW. Beyond validation: getting health apps into clinical practice. NPJ Digit Med 2020;3:14. [doi: 10.1038/s41746-019-0212-z] [Medline: 32047860]
17. Al-Nawaiseh HK, McIntosh WA, McKyer LJ. An m-health intervention using smartphone app to improve physical activity in college students: a randomized controlled trial. Int J Environ Res Public Health 2022 Jun 13;19(12):7228. [doi: 10.3390/ijerph19127228] [Medline: 35742477]
18. Holtz BE, McCarroll AM, Mitchell KM. Perceptions and attitudes toward a mobile phone app for mental health for college students: qualitative focus group study. JMIR Form Res 2020;4(8):e18347. [doi: 10.2196/18347]
19. Laranjo L, Ding D, Heleno B, et al. Do smartphone applications and activity trackers increase physical activity in adults? Systematic review, meta-analysis and metaregression. Br J Sports Med 2021 Apr;55(8):422-432. [doi: 10.1136/bjsports-2020-102892] [Medline: 33355160]
20. de Ieeuw D, Hox J, Dillman DA. International Handbook of Survey Methodology (European Association of Methodology Series) 1. URL: https://www.researchgate.net/publication/46706288_International_Handbook_Of_Survey_Methodology_2008 [accessed 2022-05-31]
21. WMA Declaration of Helsinki: ethical principles for medical research involving human subjects. WMA – The World Medical Association. URL: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/ [accessed 2023-09-01]
22. GDPR official legal text. General Data Protection Regulation (GDPR). URL: https://gdpr-info.eu/ [accessed 2023-09-01]
23. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE). URL: https://www.strobe-statement.org/ [accessed 2023-09-01]
24. Sadowski PK, Battista S, Leuzzi G, Sansone LG, Testa M. Low back pain in people with lower limb amputation: a cross-sectional study. Spine (Phila Pa 1976) 2022 Nov 15;47(22):1599-1606. [doi: 10.1097/BRS.0000000000004422] [Medline: 35920468]

XSL•FO
RenderX

25. Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Educ 2011 Jun 27;2:53-55. [doi: 10.5116/ijme.4dfb.8dfd] [Medline: 28029643]
26. Greenhalgh T, Hinton L, Finlay T, et al. Frameworks for supporting patient and public involvement in research: systematic review and co-design pilot. Health Expect 2019 Aug;22(4):785-801. [doi: 10.1111/hex.12888] [Medline: 31012259]
27. Giannarou L, Zervas E. Using Delphi technique to build consensus in practice. IJBSAM 2014;9(2):65-82 [FREE Full text] [doi: 10.69864/ijbsam.9-2.106]
28. Leuzzi G, Recenti F, Giardulli B, Scafoglieri A, Testa M. Exploring digital health: a qualitative study on adults' experiences with health apps and wearables. Int J Qual Stud Health Well-being 2025 Dec 31;20(1). [doi: 10.1080/17482631.2024.2447096]
29. Zhou L, Bao J, Watzlaf V, Parmanto B. Barriers to and facilitators of the use of mobile health apps from a security perspective: mixed-methods study. JMIR Mhealth Uhealth 2019 Apr 16;7(4):e11223 [FREE Full text] [doi: 10.2196/11223] [Medline: 30990458]
30. Rasche P, Wille M, Bröhl C, et al. Prevalence of health app use among older adults in Germany: national survey. JMIR Mhealth Uhealth 2018 Jan 23;6(1):e26. [doi: 10.2196/mhealth.8619] [Medline: 29362211]
31. Hilmarsdóttir E, Sigurðardóttir Á, Arnardóttir RH. A digital lifestyle program in outpatient treatment of type 2 diabetes: a randomized controlled study. J Diabetes Sci Technol 2021 Sep;15(5):1134-1141. [doi: 10.1177/1932296820942286] [Medline: 32680441]
32. Choe EK, Abdullah S, Rabbi M, et al. Semi-automated tracking: a balanced approach for self-monitoring applications. IEEE Pervasive Comput 2017;16(1):74-84. [doi: 10.1109/MPRV.2017.18]
33. Hahn SL, Hazzard VM, Larson N, Klein L, Loth KA, Neumark-Sztainer D. Correlates of weight-related self-monitoring application use during emerging adulthood in a population-based sample. Eat Weight Disord 2022 Aug;27(6):2107-2119. [doi: 10.1007/s40519-021-01349-4]
34. Melcher J, Camacho E, Lagan S, Torous J. College student engagement with mental health apps: analysis of barriers to sustained use. J Am Coll Health 2022;70(6):1819-1825. [doi: 10.1080/07448481.2020.1825225] [Medline: 33048626]
35. Borghouts J, Eikey E, Mark G, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. J Med Internet Res 2021 Mar 24;23(3):e24387. [doi: 10.2196/24387] [Medline: 33759801]
36. Benedetti D, Olcese U, Frumento P, et al. Heart rate detection by Fitbit ChargeHRTM: a validation study versus portable polysomnography. J Sleep Res 2021;30:e13346. [doi: 10.1111/jsr.13346] [Medline: 33837981]
37. Howell MD, Corrado GS, DeSalvo KB. Three epochs of artificial intelligence in health care. JAMA 2024 Jan 16;331(3):242-244. [doi: 10.1001/jama.2023.25057] [Medline: 38227029]
38. Hiller J, Schatz K, Drexler H. Gender influence on health and risk behavior in primary prevention: a systematic review. J Public Health 2017 Aug;25(4):339-349. [doi: 10.1007/s10389-017-0798-z]
39. Antezana G, Venning A, Smith D, Bidargaddi N. Do young men and women differ in well-being apps usage? Findings from a randomised trial. Health Informatics J 2022;28(1):14604582211064825. [doi: 10.1177/14604582211064825] [Medline: 35128952]
40. Simons D, De Bourdeaudhuij I, Clarys P, De Cocker K, Vandelanotte C, Deforche B. A smartphone app to promote an active lifestyle in lower-educated working young adults: development, usability, acceptability, and feasibility study. JMIR Mhealth Uhealth 2018;6(2):e44 [FREE Full text] [doi: 10.2196/mhealth.8287]
41. Jembai JVJ, Wong YLC, Bakhtiar N, et al. Mobile health applications: awareness, attitudes, and practices among medical students in Malaysia. BMC Med Educ 2022 Dec;22(1). [doi: 10.1186/s12909-022-03603-4]
42. Kc B, Alrasheedy AA, Hing Goh B, et al. The types and pattern of use of mobile health applications among the general population: a cross-sectional study from Selangor, Malaysia. Patient Prefer Adherence 2021;15:1755-1762. [doi: 10.2147/PPA.S325851] [Medline: 34408408]
43. Thornton L, Gardner LA, Osman B, et al. A multiple health behavior change, self-monitoring mobile app for adolescents: development and usability study of the Health4Life App. JMIR Form Res 2021 Jan;5(4):e25513. [doi: 10.2196/25513]
44. Vijayan V, Connolly JP, Condell J, McKelvey N, Gardiner P. Review of wearable devices and data collection considerations for connected health. Sensors (Basel) 2021 Aug 19;21(16):5589. [doi: 10.3390/s21165589] [Medline: 34451032]
45. Rodriguez VH, Medrano CT, Plaza I. Wearable sensors for measuring movement in short sessions of mindfulness sitting meditation: a pilot study. J Healthc Eng 2018;2018:7275049. [doi: 10.1155/2018/7275049] [Medline: 29854363]
46. Jabour AM, Rehman W, Idrees S, Thanganadar H, Hira K, Alarifi MA. The adoption of mobile health applications among university students in health colleges. J Multidiscip Healthc 2021;14:1267-1273. [doi: 10.2147/JMDH.S310539] [Medline: 34103927]
47. Willems SJ, Coppieters MW, Pronk Y, et al. A clinical journey mobile health app for perioperative patients: cross-sectional study. JMIR Hum Factors 2021 Feb 8;8(1):e20694 [FREE Full text] [doi: 10.2196/20694] [Medline: 33555262]
48. Schneider T, Baum L, Amy A, Marisa C. I have most of my asthma under control and I know how my asthma acts: users' perceptions of asthma self-management mobile app tailored for adolescents. Health Informatics J 2020 Mar;26(1):342-353. [doi: 10.1177/1460458218824734] [Medline: 30732520]
49. Ndayizigamiye P, Kante M, Shingwenyana S. An adoption model of mHealth applications that promote physical activity. Cogent Psychol 2020 Dec 31;7(1):1764703. [doi: 10.1080/23311908.2020.1764703]

XSL•FO
RenderX

50. Goodyear VA, Armour KM, Wood H. Young people learning about health: the role of apps and wearable devices. Learn Media Technol 2019 Apr 3;44(2):193-210. [doi: 10.1080/17439884.2019.1539011]
51. Arora S, ter Hofstede F, Mahajan V. The implications of offering free versions for the performance of paid mobile apps. J Mark 2017 Nov;81(6):62-78. [doi: 10.1509/jm.15.0205]
52. Liu H, Liu S. Research on advertising and quality of paid apps, considering the effects of reference price and goodwill. Mathematics 2020 Jun;8(5):733. [doi: 10.3390/math8050733]
53. García JS, Alonso SG, de la Torre Díez I, et al. Reviewing mobile apps to control heart rate in literature and virtual stores. J Med Syst 2019 Feb 19;43(4):80. [doi: 10.1007/s10916-019-1202-z] [Medline: 30783824]
54. Zhao W, Ge Y, Qu W, Zhang K, Sun X. The duration perception of loading applications in smartphone: effects of different loading types. Appl Ergon 2017 Nov;65:223-232. [doi: 10.1016/j.apergo.2017.06.015]
55. Newton A, Bagnell A, Rosychuk R, et al. A mobile phone-based app for use during cognitive behavioral therapy for adolescents with anxiety (MindClimb): user-centered design and usability study. JMIR Mhealth Uhealth 2020 Dec 8;8(12):e18439. [doi: 10.2196/18439] [Medline: 33289671]
56. Wang Y, Huang Y, Li J, Zhang J. The effect of mobile applications' initial loading pages on users' mental state and behavior. Displays 2021 Jul;68:102007. [doi: 10.1016/j.displa.2021.102007]
57. Timmers T, Janssen L, van der Weegen W, et al. The effect of an app for day-to-day postoperative care education on patients with total knee replacement: randomized controlled trial. JMIR Mhealth Uhealth 2019 Oct 21;7(10):e15323. [doi: 10.2196/15323] [Medline: 31638594]
58. Harmon DJ, Burgoon JM, Kalmar EL. Development and assessment of an integrated anatomy mobile app. Clin Anat 2022 Jul;35(5):686-696. [doi: 10.1002/ca.23895] [Medline: 35452135]
59. McCall T, Threats M, Pillai M, Lakdawala A, Bolton CS. Recommendations for design of a mobile application to support management of anxiety and depression among Black American women. Front Digit Health 2022;4:1028408. [doi: 10.3389/fdgth.2022.1028408] [Medline: 36620185]
60. Cao Y, Zhang Y, Ding Y, Duffy VG, Zhang X. Is an anthropomorphic app icon more attractive? Evidence from neuroergonomomics. Appl Ergon 2021 Nov;97:103545. [doi: 10.1016/j.apergo.2021.103545] [Medline: 34352470]
61. Bol N, Høie NM, Nguyen MH, Smit ES. Customization in mobile health apps: explaining effects on physical activity intentions by the need for autonomy. Digit Health 2019;5:2055207619888074. [doi: 10.1177/2055207619888074] [Medline: 31807312]
62. Liu N, Yin J, Tan SSL, Ngiam KY, Teo HH. Mobile health applications for older adults: a systematic review of interface and persuasive feature design. J Am Med Inform Assoc 2021 Oct 12;28(11):2483-2501. [doi: 10.1093/jamia/ocab151] [Medline: 34472601]
63. Kruzan KP, Reddy M, Washburn JJ, Mohr DC. Developing a mobile app for young adults with nonsuicidal self-injury: a prototype feedback study. Int J Environ Res Public Health 2022 Dec 2;19(23):16163. [doi: 10.3390/ijerph192316163] [Medline: 36498234]
64. Pérez-López IJ, Navarro-Mateos C, Mora-Gonzalez J. "STAR WARS: The first Jedi" gamification program: use of a mobile app to improve body composition in college students. Games Health J 2022 Oct;11(5):321-329. [doi: 10.1089/g4h.2022.0046] [Medline: 35944270]
65. Anderson K, Burford O, Emmerton L. Mobile health apps to facilitate self-care: a qualitative study of user experiences. PLoS ONE 2016;11(5):e0156164. [doi: 10.1371/journal.pone.0156164] [Medline: 27214203]
66. Tong HL, Coiera E, Laranjo L. Using a mobile social networking app to promote physical activity: a qualitative study of users' perspectives. J Med Internet Res 2018 Dec 21;20(12):e11439. [doi: 10.2196/11439] [Medline: 30578201]
67. Arigo D, Brown MM, Pasko K, Suls J. Social comparison features in physical activity promotion apps: scoping meta-review. J Med Internet Res 2020 Mar 27;22(3):e15642. [doi: 10.2196/15642] [Medline: 32217499]
68. Tong HL, Laranjo L. The use of social features in mobile health interventions to promote physical activity: a systematic review. NPJ Digit Med 2018;1(1):43. [doi: 10.1038/s41746-018-0051-3] [Medline: 31304323]
69. Helbostad JL, Vereijken B, Becker C, et al. Mobile health applications to promote active and healthy ageing. Sensors (Basel) 2017 Mar 18;17(3):622. [doi: 10.3390/s17030622] [Medline: 28335475]
70. Baer NR, Vietzke J, Schenk L. Middle-aged and older adults' acceptance of mobile nutrition and fitness apps: a systematic mixed studies review. PLoS ONE 2022;17(12):e0278879. [doi: 10.1371/journal.pone.0278879] [Medline: 36520839]
71. Chettri S, Wang V, Balkin EA, Rayo MF, Lee CN. Development of a mobile app for clinical research: challenges and implications for investigators. JMIR Mhealth Uhealth 2022 Apr 1;10(4):e32244. [doi: 10.2196/32244] [Medline: 35363154]
72. Bautista J, Schueller SM. Understanding the adoption and use of digital mental health apps among college students: secondary analysis of a national survey. JMIR Ment Health 2023 Mar 22;10:e43942. [doi: 10.2196/43942] [Medline: 36947115]
73. McCallum C, Rooksby J, Gray CM. Evaluating the impact of physical activity apps and wearables: interdisciplinary review. JMIR Mhealth Uhealth 2018 Mar 23;6(3):e58. [doi: 10.2196/mhealth.9054] [Medline: 29572200]

Abbreviations

HBM: health belief model

XSL•FO
RenderX

mHealth: mobile health
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Kushniruk; submitted 22.07.24; peer-reviewed by B Pouls, H Wieser, S Mitra; revised version received 09.05.25; accepted 27.05.25; published 03.09.25.

Please cite as:
Leuzzi G, Job M, Scafoglieri A, Testa M
Smartphone Apps and Wearables for Health Parameters in Young Adulthood: Cross-Sectional Study
JMIR Hum Factors 2025;12:e64629
URL: https://humanfactors.jmir.org/2025/1/e64629
doi:10.2196/64629

© Gaia Leuzzi, Mirko Job, Aldo Scafoglieri, Marco Testa. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 3.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Developing a Data Trust Model (Not Only) for Sleep Research: Conceptual Study and Quantitative Survey

Raphael Jan Dressle[1], PhD; Dieter Riemann[1], PhD; Nicole Thoma[1], MSc; Christina Erler[2], MSc, PhD; Rodger Burmeister[2], Dipl-Inf; Bianka Jogwitz[3], MA; Katharina Domschke[1,4], MD, PhD; Kai Spiegelhalder[1], MD, PhD; Joachim Boldt[5,6], PhD; Svenja Wiertz[5], DPhil; Bernd Feige[1], PhD

[1]Department of Psychiatry and Psychotherapy, Medical Center - University of Freiburg, Faculty of Medicine, University of Freiburg, Albert-Ludwigs-Universität Freiburg, c/o Medizinische Fakultät, Klinik für Psychiatrie und Psychotherapie, Hauptstraße 5, Freiburg, Germany

[2]FZI Research Center for Information Technology, Karlsruhe, Germany

[3]Medical Center - University of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany

[4]German Center for Mental Health (DZPG), Partner site Berlin/Potsdam, Berlin, Germany

[5]Department of Medical Ethics and the History of Medicine, University of Freiburg, Freiburg, Germany

[6]Center for Philosophy and Ethics of Health, University of Southern Denmark, Odense, Denmark

Corresponding Author:
Raphael Jan Dressle, PhD
Department of Psychiatry and Psychotherapy, Medical Center - University of Freiburg, Faculty of Medicine, University of Freiburg, Albert-Ludwigs-Universität Freiburg, c/o Medizinische Fakultät, Klinik für Psychiatrie und Psychotherapie, Hauptstraße 5, Freiburg, Germany

Abstract

Background: A large amount of data are generated in health care facilities, yet it is rarely made available for secondary research use. The reasons are manifold. Most importantly, different stakeholders' needs must be balanced. However, there are currently hardly any feasible solutions for this.

Objective: This study aimed to develop a data trust model with supporting user interface applications to provide a legally and ethically sound framework for secondary use of medical data. The development was based on extensive surveys of various stakeholders.

Methods: Semistructured interviews were conducted with researchers (data users) and institutional representatives of the Medical Center–University of Freiburg, and online questionnaires were administered to patients (data subjects), data users, and institutional representatives. The questionnaire for data subjects covered the dimensions of trust (measured with a 5-point Likert scale), quality of interaction and involvement (measured with a 4-point Likert scale), subjective and objective understanding, and usability (measured with the user version of the Mobile Application Rating Scale). For all other stakeholder groups, the questionnaire focused on usability measured using the user version of the Mobile Application Rating Scale. The surveys comprised a requirement elicitation followed by two rounds of evaluation. Independent-samples Welch *t* tests were used to compare group means between the first and second evaluations.

Results: We devised SouveMed, a framework for secondary use of medical data, applied to the use case of sleep research data. The model includes secure onboarding of data subjects and using digital consent and a digital interface for data users for onboarding, defining research aims, querying the amount of available data, and, finally, either downloading data or having algorithms run on it. At its core is a data trust entity that matches descriptions, consents, and constraints of all stakeholders using digital representations and constraint-solving techniques. Fourteen participants took part in the requirements elicitation, 22 in the first evaluation, and 16 in the second evaluation. In both the first and the second evaluations, data subjects showed a high level of trust in the concept, with mean ratings on the trust scale of 4.23 (SD 0.46) in the first and 4.23 (SD 0.68) in the second evaluation ($t_{15.78}$=0.03, P=.97). Regarding usability, the mean functionality score of the data user system increased from 3.56 (SD 0.77) in the first to 4.58 (SD 0.38) out of 5 points in the second evaluation ($t_{10.69}$=−3.28, P=.008). The mean functionality score of the data subject system increased from 4.30 (SD 0.41) in the first to 4.50 (SD 0.74) in the second evaluation ($t_{13.99}$=−0.75, P=.46).

Conclusions: The SouveMed concept provides a comprehensive framework for the secondary use of medical data. The developed processes can be adapted to other areas of medical research.

Trial Registration: German Clinical Trial Register DRKS00031093; https://drks.de/search/en/trial/DRKS00031093/details

(JMIR Hum Factors 2025;12:e66513) doi:10.2196/66513

XSL•FO
RenderX

KEYWORDS

health data; data trust; sleep research; informed consent; consent management; digital infrastructure; patient participation; sleep; survey; trust

Introduction

Digitalization in health care holds great potential for data-driven research and scientific progress for the benefit of patients and the general public. This includes, for example, automated systems for disease prediction [1-3], as well as innovative approaches that enhance data exchange between practitioners and enable better patient monitoring [4]. However, the use of data for purposes outside the specific treatment context (ie, secondary use) is hindered by several aspects. While an increasing amount of data is being digitally recorded during treatment, the common notion of *data silos*, where data is in principle readily accessible for research purposes, is too simplistic. This is because many processes in routine clinical care are not yet sufficiently digitized, and many analog, paper-based processes remain [5,6]. Furthermore, even digitized data is often not sufficiently preprocessed to be directly accessible for research purposes. They are documented using different software systems that lack interoperability and are stored in a variety of file formats [5-7]. In clinical practice, for example, questionnaires are still frequently handed out in paper form. Digitizing these questionnaires can simply mean that they are scanned and saved in PDF format. A major challenge in this context is the fact that electronic health records, while principally useful for research, are primarily intended to document treatment and reimbursement. Separate measures are necessary to render health record data *fit for purpose* regarding research and to check data quality [8,9].

Legal restrictions are another major obstacle. In the European Union, the introduction of the General Data Protection Regulation (GDPR) has been celebrated as a milestone for the protection of data subjects' rights but has also been criticized for creating obstacles to research [10-13]. For example, it has been criticized that the GDPR sets higher standards for anonymization than was previously the case in many European Union member states [12,14]. Recital 26 of the GDPR states that pseudonymized data "should be considered to be information on an identifiable natural person." This has been interpreted to mean that, in principle, all key-coded data must be considered and treated as identifiable personal data, even if the researcher holding the data is not in possession of the key to re-identify it [14]. Therefore, a legal basis is required for the processing of this data, which generally requires greater efforts from researchers and their infrastructure in terms of data protection.

Article 9 of the GDPR outright prohibits the processing of sensitive personal data, which includes health-related data. An exception is made in Article 9 (2) (a) if the patient (data subject) has given explicit consent to the processing of the personal data. Another exception refers directly to scientific research. Under Article 9 (2) (j), processing of sensitive data is allowed for research purposes if appropriate safeguards are in place and a lawful basis for the processing of data is given in the Union or the relevant member state in accordance with Article 6 [11,15].

With regard to consent, there are different consent models to choose from, ranging from broad consent (ie, consent for future studies whose exact research questions and specifications are unknown) to specific consent for a particular study. Hybrid models between broad and specific consent have been developed that allow participants to specify the terms of use without having to give consent for each individual study (tiered consent; for a general overview on different consent models, see [16]). However, in the area of secondary use of medical data, it may be impractical to obtain explicit consent (eg, due to the need to *reconsent* for data already generated), and it has been suggested that it would often be useful to consider other possible legal bases (eg, [11,14]).

The secondary use of health data must, however, also comply with other legal and ethical principles that are not fully taken into account in the GDPR [14]. While from the perspective of the GDPR, obtaining active consent might be replaced by referring to the special purpose of scientific research (if the required legal basis in the relevant member state exists, as mentioned above), consent is still a central ethical standard in medical research. Of particular interest in the field of medical research are the ethical principles formulated in the Declaration of Helsinki (DOH) [17] and Taipei (DOT) [18] of the World Medical Association. The DOH formulates general ethical principles for medical research involving human subjects, while the DOT focuses specifically on principles for databases containing health data. Both statements emphasize the importance of explicit consent for the secondary use of health data. The DOH states that "physicians or other qualified individuals must obtain free and informed consent from research participants for the collection, processing, storage, and foreseeable secondary use of biological material and identifiable or re-identifiable data" [17]. The DOT sets out further rules, such as the right to receive information about the use of the data and the right to change or withdraw consent at any time [18]. Indeed, these principles are in line with the spirit of the GDPR, which formulates and generally strengthens key data subject rights centered on purpose limitation, sovereignty, and transparency, including the right to be informed about processing and the rights to erasure and restriction of processing (see Articles 5 and 6) [10,15,19].

In addition to these basic requirements for the secondary use of research data, it should not be neglected that the secondary use of medical and research data affects many different interests, which must be carefully balanced in order to achieve the acceptance of all stakeholders. Besides the data subject and data users, these stakeholders crucially include the data-generating side, which is typically a clinic or practitioner who cannot waive their care obligations regarding the patient, including the use of sensitive data. With respect to the data subjects, it has been shown that patients are generally very willing to allow access to their health data for public research purposes [20-24]. Nevertheless, many patients are concerned about a lack of privacy, transparency, and data security [25,26]. To address

XSL•FO
RenderX

these concerns, a sufficient level of trust is required [27,28]. In this regard, appropriate patient involvement and communication appears to be crucial [25,26]. Adequate information about the secondary use of health data should be provided in an understandable and, at best, personalized form [26], including information about specific projects and findings related to data use [28]. Although there are large interindividual differences in the level of involvement favored by patients, many studies show that a large number of patients prefer a high level of control over their health data [28-32].

As noted above, data users are another important stakeholder group. There is a widespread view that considers health data to be the private property of patients. However, this approach falls short. Concepts of ownership cannot easily be applied to data, and health data can be considered as coconstructed, involving both the data subject and the clinicians and researchers with the generating institution, who contribute their efforts in collecting, classifying, interpreting, processing, and storing the data [33-35]. Because of this crucial contribution and the care obligations of the generating institutions noted above, the needs and constraints of institutions and researchers involved in the generation of health data must be equally considered. Correspondingly, the willingness of institutions and researchers to share health data without appropriate incentives, as well as trust and legal security, is often low [36,37]. At the same time, data users formulate their own requirements for the composition of data sets, utilization options, and acceptable expenditures.

The secondary use of health data, therefore, takes place in a complex environment: legal norms (relating to, but not limited to, data protection regulation), ethical standards, and the interests and needs of different stakeholder groups must be taken into account and balanced. At the same time, a progressive increase in the amount of data available can be expected, which necessitates time- and cost-efficient processes to make use of this data. While initiatives such as the envisaged European Health Data Space (EHDS), the centralized Finnish health data registries, and the German *Medizininformatikinitiative* (medical informatics initiative), among others, prove that the secondary use of health data is on the political and societal agenda, overall, the development is still in its infancy and the great potential that lies in the secondary use of data is currently not being realized. As a possible solution to the problems outlined above, the possibility of data trust models is being discussed (eg, [38-41]). In a report for the United Kingdom government by Hall and Pesenti [41] on the future use of artificial intelligence, data trusts are defined in an overarching manner as “a set of relationships underpinned by a repeatable framework, compliant with parties’ obligations, to share data in a fair, safe and equitable way” (p. 46). Nevertheless, as comprehensively summarized by Lauf et al [42], there is currently no universally accepted definition of a data trust. Instead, based on an analysis of applications across various domains, four *archetypes* of data trusts have been proposed, each with distinct characteristics: data brokerage trustees, data processing trustees, data aggregation trustees, and data custody trustees [42]. In view of the strict data protection requirements associated with the use of health data, existing studies in the medical domain have examined technical features designed to mitigate potential risks. For example, based on a security-by-design approach, the modular separation of demographic and medical data has been proposed to improve data protection within a data trust [43]. In the context of the German medical informatics initiative, the secondary use of medical data is being facilitated by the so-called data integration centers at several university hospitals. Access to the data is controlled by *use and access committees*, which manually review and evaluate data requests before granting access [44,45]. However, beyond these technical and organizational considerations, the successful implementation of a data trust model also depends on stakeholder acceptance and its effective integration into existing clinical workflows.

The aim of the present study was to design a comprehensive legally and ethically sound data trust model from the early conceptualization, through the technical implementation of a prototype system, to its evaluation from the perspectives of various stakeholders. In this study, a data trust model was conceptualized as a neutral entity that facilitates data use while accounting for the interests of all involved stakeholders. Importantly, the data trust—by our definition—does not make decisions on behalf of others. At its core, the system includes a matching service that automatically compares the rules and requirements specified by each stakeholder. When these align, access is granted, without the system itself exercising intentionality. Development prioritized preserving data subjects’ sovereignty while ensuring that system processes remain scalable.

The evaluation was preregistered in a clinical trials registry (German Clinical Trials Registry DRKS00031093). A particular focus was on the use case of sleep research. Sleep research seemed particularly suitable as it is a highly relevant field that warrants further intensified research efforts due to the high prevalence and disease burden of sleep disorders. For example, up to 10% of the general population experiences an insomnia disorder [46] and between 9% and 38% (dependent on age group and methodological approach used) from obstructive sleep apnea [47]. In addition, the field of sleep research is particularly suitable for data-based approaches, as investigating sleep generates a large amount of data, such as sensor data (from the conventionally used polysomnography to modern lifestyle devices such as smartwatches) and subjective data [48]. Besides that, this field is especially sensitive because of its links to psychiatry and is underrepresented in current data-sharing initiatives.

Methods

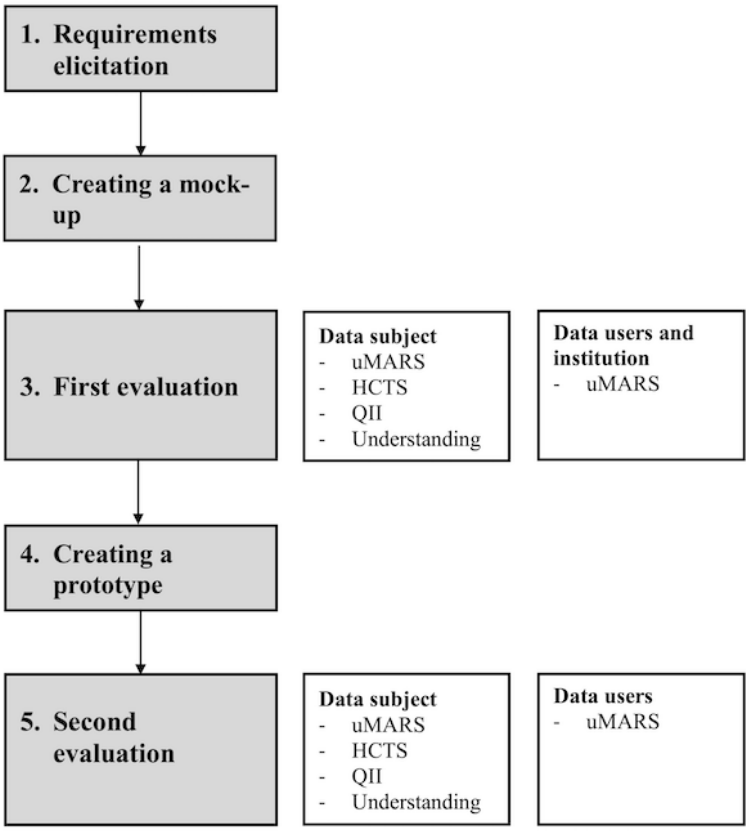
Overview of Study Design

The conceptualization of the data trust model was carried out in five steps (see Figure 1). First, a participatory approach was used to identify relevant needs regarding a data trust model from the perspective of different stakeholders and to define basic requirements for a data trust (requirement elicitation). This included a survey of data subjects, patient representatives, and sleep researchers (data users). An initial conceptual model was then developed, which included a refinement of the relevant procedures and first interactive mock-ups of user interfaces (UI). In a second step, these UIs and procedures were evaluated

XSL•FO
RenderX

from the perspective of data subjects, data users, and institutional representatives (including data protection and ethics experts; first evaluation). The feedback was used to further refine the system prototype that was developed in parallel. This prototype was then evaluated again in a second evaluation, which was similar to the process used in the first round of evaluation.

Figure 1. Overview of the study design. HCTS: Human-Computer Trust Scale; QII: Quality of Interaction and Involvement; uMARS: user version of the Mobile Application Rating Scale.



Requirements Elicitation

Separate questionnaires were developed for data subjects/patient representatives and data users. The questions were developed on the basis of a structured consensus process in the project group with the participation of IT experts and sleep researchers and included questions on general attitudes toward data sharing and preferred consent models as well as technical and organizational requirements, depending on the target group. The full questionnaire used for the requirements elicitation is available upon request.

Evaluation

Overview

The aim of the evaluation study was twofold. First, we wanted to gain some general insights into attitudes toward data sharing with a data trust model from the perspective of different stakeholder groups. These included data subjects from the sleep laboratory, Department of Psychiatry, Medical Center–University of Freiburg, as well as data users/representatives from data-generating institutions. We also aimed at evaluating our specific data sharing process and data trust model from both perspectives. To this end, we conducted a series of anonymous surveys using a semistructured interview and questionnaires (see below). The relevant constructs for the development of the questionnaires were identified in internal project workshops as well as on the basis of a systematic literature search and the needs analysis.

Semistructured Interviews

Semistructured interviews were conducted with data users and institutional representatives to gain a more comprehensive picture of the views on data sharing and the corresponding needs. This consisted of a total of 10 predefined questions (see Multimedia Appendix 1).

Questionnaire

In total, the questionnaire for data subjects comprised 51 items, which included 4 sociodemographic questions, 5 open questions, 33 Likert scale items, and 9 statement items. The questionnaire for data users/institutional representatives included a total of 30 items, which included 2 sociodemographic questions, 6 open questions, 16 Likert scale items, and 6 statement items.

Trust was assessed using the *human-computer trust scale* (HCTS) [49]. The HCTS is a 12-item questionnaire for assessing user trust based on the *human-computer trust model* [50]. The HCTS includes questions on three dimensions that have been shown to measure trust. These include benevolence (ie, the belief that a system acts in the interest of the user and provides sufficient help if necessary), competence (ie, whether the system

XSL•FO
RenderX

has all features and functions that are necessary to fulfill the intended tasks), and perceived risk (individual evaluation regarding the probability of adverse consequences when using the system). All questions were answered on a 5-point Likert scale from 1 ("strongly disagree") to 5 ("strongly agree").

Quality of interaction and involvement (QII) was assessed using 7 questions developed by Abdelhamid et al [51]. The questions were translated into German and adapted to relate specifically to the SouveMed process (eg, "Have your feelings and emotions been given the attention they deserve?"). Data subjects had to answer on a 4-point Likert scale ("yes, always" to "no, never").

The German version of the Mobile Application Rating Scale (MARS) [52] was used to assess usability. Three of four sections that are part of the user version of the MARS (uMARS) [53] were presented. These include the following sections: functionality (eg, "How easy is it to learn how to use the SouveMed app; how clear are the menu labels, icons and instructions?"), aesthetics (eg, "Visual appeal: How good does the SouveMed app look?"), and subjective quality (eg, "Would you recommend the SouveMed app to people who might benefit from it?"). We decided to use only a subset of the sections to avoid overlap with other elements of our comprehensive questionnaire and because some sections were not applicable to the objectives and content of our specific system. All questions were answered on a 5-point Likert scale, with responses reflecting judgments ranging from "inadequate" to "excellent." Higher scores, therefore, reflect higher quality ratings for the respective domain.

Sociodemographic information and understanding were assessed using 13 items of the questionnaire developed by Richter et al [21]. This included a self-assessment of understanding, subjective reasons for possible poor understanding, and questions to assess the content-related understanding objectively.

Participants

Current patients and participants of other studies in the sleep laboratory at the Department of Psychiatry, Medical Center–University of Freiburg, as well as data users, that is, sleep researchers and representatives of the Medical Center–University of Freiburg with special expertise in data protection or ethics, were approached and invited to participate voluntarily in the study. Participants who were 18 years or older were included and provided informed consent. We excluded pregnant participants and patients with severe mental or somatic disorders that would have made it unreasonable or impracticable to conduct the survey.

Ethical Considerations

Approval of the ethics committee of the University of Freiburg, Freiburg, Germany, was obtained (case number 23-1047-S2). The study was conducted in accordance with applicable German and EU data protection regulations as well as the DOH. Participation was voluntary and not compensated. All participants were informed about the aims, methods, risks, and potential benefits of the study and provided documented informed consent. No personally identifying information was collected, ensuring that all study data were fully anonymized.

Statistical Analysis

The results were averaged, and questionnaire scores are presented as mean (SD) values where possible. Analyses were performed using the statistics software R (version 4.5.0). In the case of missing data, subjects were excluded only from analyses involving the variables for which data were missing (pairwise deletion). Independent-samples Welch t tests were used to compare group means between the first and second evaluation, and Hedges g was calculated as a measure of effect size. P values ≤.05 were considered statistically significant.

Results

Requirements Elicitation and the SouveMed Data Trust

Requirements Elicitation

We received feedback from a total of 14 participants, including 10 data subjects/patient representatives and 4 data users. From the data subject's perspective, the survey revealed a clear preference for the possibility of personal contact as part of the informed consent process. There was also a preference for a consent model that allows for further refinement of the conditions under which data sharing is permitted. However, the majority of respondents were critical of giving separate consent for each individual data request. In addition, there was a high general willingness to provide data, although sufficient transparency and data security were often formulated as the most important conditions for this. Furthermore, 5 (50%) of the data subjects considered it "very important" or "important" to receive information about study results that were obtained using their own data. With 6 data subjects (60%), a majority considered it "very important" or "important" to receive a brief description of study projects, and 5 (50%) considered it "very important" or "important" to receive more detailed information about the status of these studies. A large proportion of the data subjects surveyed were therefore interested in comprehensive participation.

From the perspective of data users, there was a clear preference for the European Data Format for the raw polysomnography data. The availability of a range of relevant metadata (eg, number of people, age, diagnosis statistics, etc) was also seen as essential to the attractiveness of the system.

The SouveMed Data Trust

Based on the needs analysis and workshops by the project group, we developed the *SouveMed* concept, a comprehensive workflow for the secondary use of health data (see Figure 2) with the following key features from the perspective of different stakeholder groups. From the perspective of data subjects, the following functions exist:

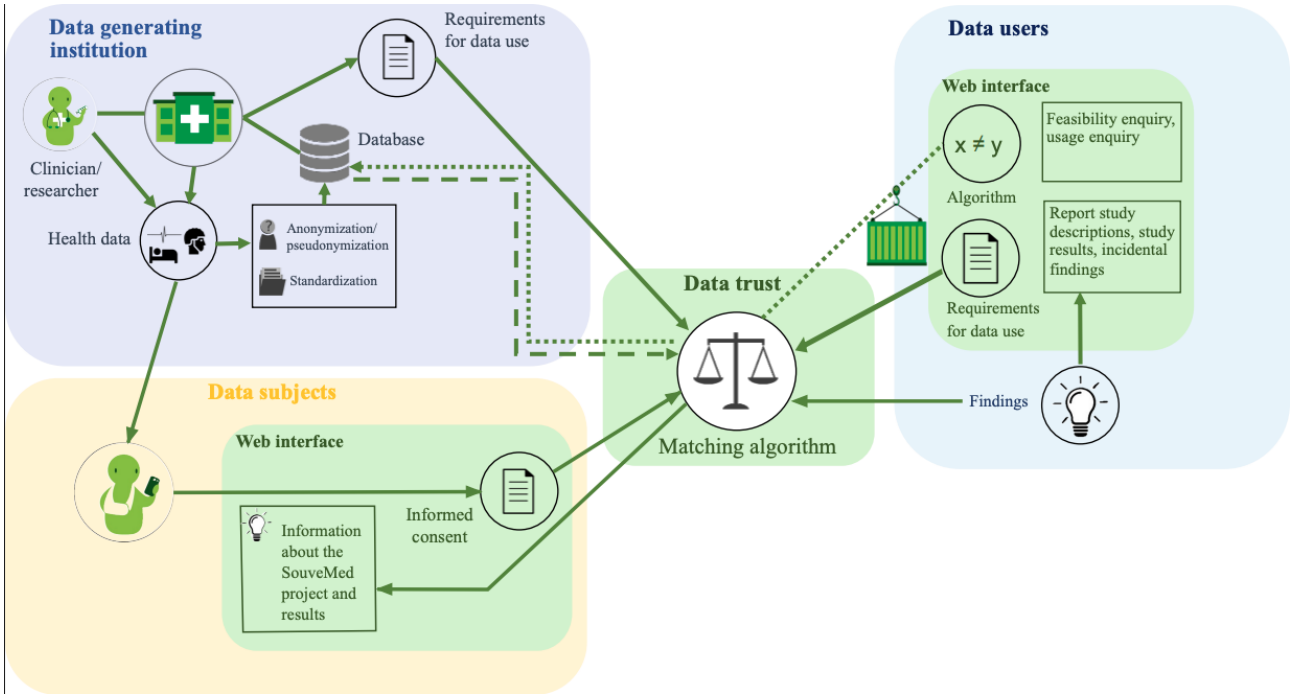
- *Onboarding:* A key element of the SouveMed concept is the separation of the onboarding process, that is, the willingness to participate in SouveMed and manage consents with the SouveMed system, from the act of consenting to or disallowing the use of data. It is therefore envisaged that data subjects will only be approached during their visit to the clinical facility to obtain information about

XSL•FO
RenderX

SouveMed and to register on the platform. This involves establishing the data subject's identity and generating a SouveMed ID, that is, a number that can be used together with a password to log in to the platform.

- *Providing consent:* In order to allow participants the necessary time for their decision as well as independence from the clinical setting, decisions are laid down digitally on the SouveMed platform. Consent is designed as a tiered consent (see Multimedia Appendix 2): It applies to a specific visit in the clinical facility and can be further specified, for example, regarding user groups (public or commercial research) and type of shared data (individual or aggregated). The currently valid consent can be viewed at any time on the digital platform, and consent can be withdrawn just as easily as given, amounting to a revocation of data use.
- *Participating in data usage:* The SouveMed platform enables participation in the use of data. Data subjects can obtain information about projects that have accessed their data (including a clear description of the main objectives and results of the studies) and have an overview of current and past consents via a consent archive. Should concerns about data protection or other questions arise, they can easily find the relevant contact information on the SouveMed platform.

Figure 2. The SouveMed concept.



From the perspective of data users, the following functions exist:

- *Feasibility inquiry:* Users can submit a feasibility inquiry for a defined research project (ie, specify the inclusion and exclusion criteria and receive the number of available matches with these criteria), as a low-threshold way to determine whether a research project with SouveMed data is realizable.
- *Usage inquiry:* If the result of the feasibility inquiry appears promising, data users can request data usage directly via the SouveMed platform. They can also obtain an overview of their former and current inquiries.
- *Data analysis on the platform.* Users can submit their own scripted algorithms and evaluate datasets in a controlled data trust environment using the curious container technique (for more details, see [54]). This allows for increased data protection, as only aggregated data are shared, including data from participants who only consented to the sharing of aggregated data.
- *Enabling participation of data subjects:* Users can use the SouveMed platform to report incidental findings, as well as provide study outlines and results to data subjects.

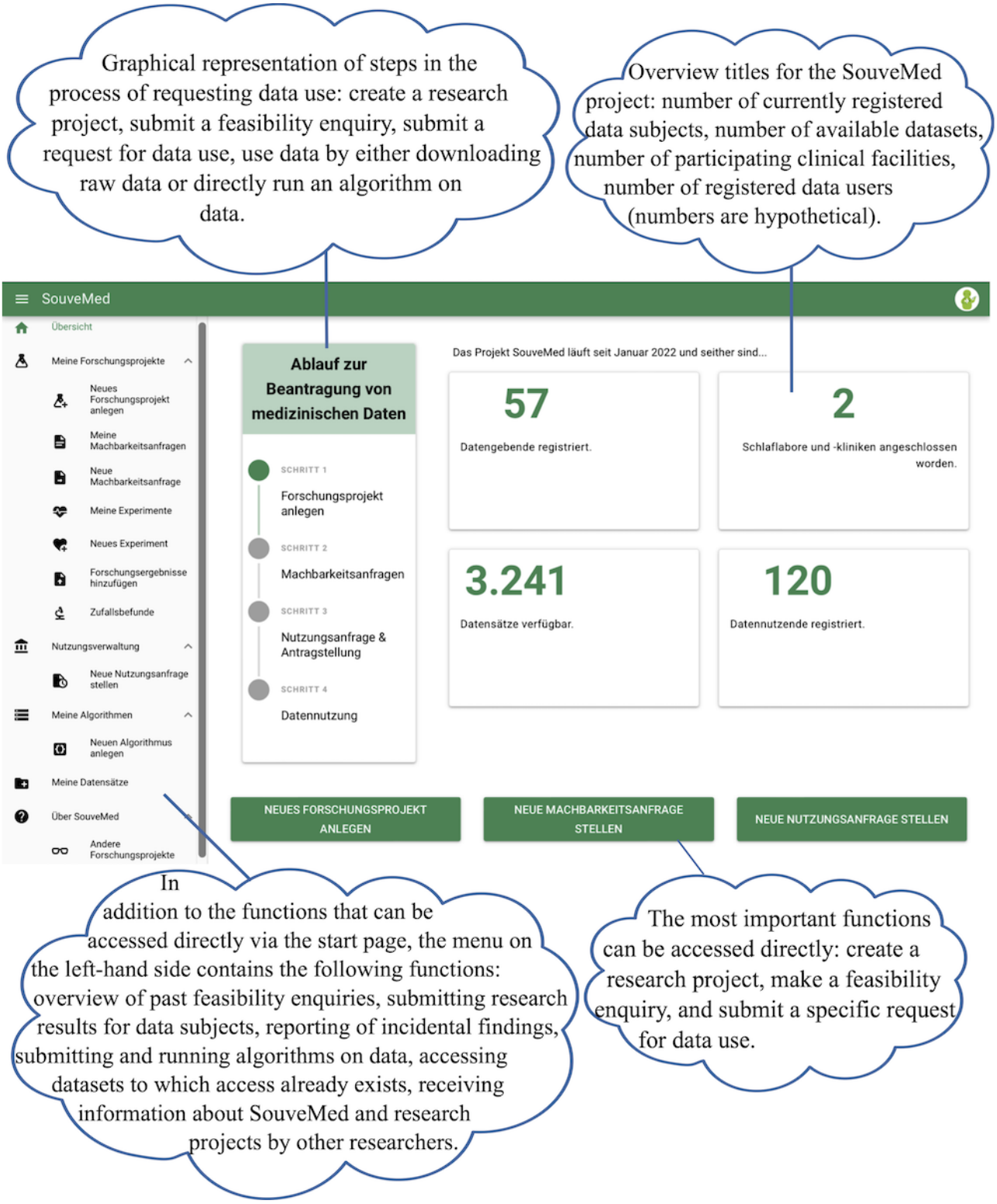
Finally, from the perspective of the data-generating institutions:

- Institutions can define requirements for data use, for example, embargos.
- Institutions can monitor the data flow decisions of the platform and ensure conformance with requirements.
- Technical prerequisites were established for this control functionality for data-generating institutions, but it was not implemented as a web interface; SouveMed is designed to be decentralized with a data access component that is under control of the generating institution.

In the SouveMed project, separate UIs were developed for data users (see Figure 3) and data subjects (see Figure 4). Conceptually, it is intended that data users must create their own user account, whereby a contract between the affiliated institution and SouveMed is a prerequisite for this. A key element of the SouveMed system is the *data trust component*, which automatically matches the descriptions, preferences, and requirements of all stakeholders (data subject, researcher, and institution) using digital contract representations. For details on the technical implementation of the described functions of the SouveMed concept, please refer to [54] and [55].

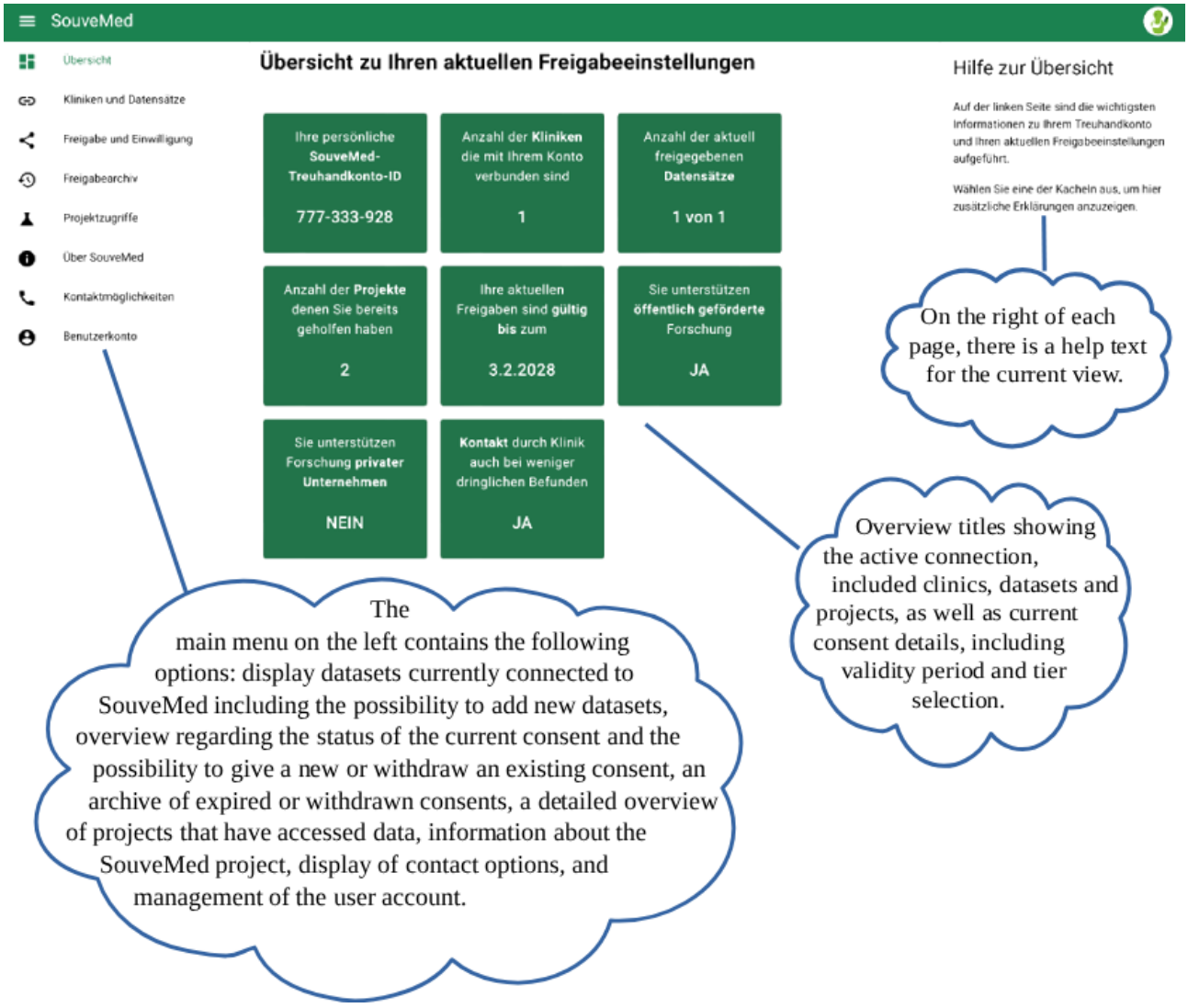
XSL•FO
RenderX

Figure 3. User interface for data users.



XSL•FO
RenderX

Figure 4. User interface for data subjects.



Evaluation

Sample Characteristics

A total of 22 participants (14 females and 8 males) were studied in the first evaluation. These included 11 patients with sleep disorders and 11 data users/institutional representatives. In the second evaluation, a total of 16 participants (9 females and 7 males) were examined. These included 10 patients with sleep disorders and 6 data users. Of the 22 participants in the first evaluation, 5 (23%) were between 18 and 30 years old, 10 (45%) between 31 and 45 years old, 4 (18%) between 46 and 60 years old, and 3 (14%) between 61 and 75 years old. Of the 16 participants in the second evaluation, 2 (13%) fell into the age range of 18 to 30 years, 7 (44%) into the age range of 31 to 45 years, 5 (31%) into the age range of 46 to 60 years, and 2 (13%) into the age range of 61 to 75 years.

The general willingness to make data available for secondary use was assessed in the data subject group with a mean value of 4.35 (SD 0.58) on a 5-point Likert scale (“very low”=1 point to “very high”=5 points; for the first and second evaluations combined). The need for a corresponding system was assessed in the group of data users/institutional representatives with a mean value of 4.29 (SD 0.77) on a 5-point Likert scale (“very low”=1 point to “very high”=5 points; for the first and second evaluations combined).

Semistructured Interview

Data users and institutional representatives described the system as clear and equipped with all the functions they required. It was pointed out that thorough explanations of the secondary use of research data are also important for this stakeholder group. Furthermore, it was uniformly emphasized that the need for a system that facilitates the secondary use of research data is very high.

Previous contact with the concept of data trust models was mentioned by only two participants. With regard to the required functionalities, it was emphasized that access to the system should be easy and that all legal requirements should be met. When the representatives of the institutions were asked whether they would agree to the use of SouveMed in their institution, all gave positive feedback. One person emphasized the importance of an integrated model for all areas of health care in order to avoid isolated solutions for different areas.

XSL•FO
RenderX

User Version of the MARS

In the first evaluation, 8 data users/institutional representatives and 10 data subjects provided complete data on the uMARS. In the second evaluation, 6 data users and 10 data subjects provided complete data on the uMARS.

Functionality

The mean functionality score of the data user system increased from 3.56 (SD 0.77) in the first evaluation to 4.58 (SD 0.38) in the second evaluation. This difference was statistically significant ($t_{10.69}$=−3.28, *P*=.008), with a large effect size (*g*=−1.51, 95% CI −2.76 to −0.26). The mean functionality score of the data subject system increased from 4.30 (SD 0.41) in the first evaluation to 4.50 (SD 0.74) in the second evaluation. This difference was not statistically significant ($t_{13.99}$=−0.75, *P*=.46; *g*=−0.32, 95% CI −1.23 to 0.58).

Aesthetics

The mean aesthetics score of the data user system increased from 4.08 (SD 0.68) in the first evaluation to 4.17 (SD 0.55) in the second evaluation. This difference was not statistically significant ($t_{11.91}$=−0.25, *P*=.80; *g*=−0.12, 95% CI −1.23 to 0.98). The mean aesthetics score of the data subject system increased from 3.93 (SD 0.44) in the first evaluation to 4.23 (SD 0.94) in the second evaluation. This difference was not statistically significant ($t_{12.72}$=−0.91, *P*=.38; *g*=−0.39, 95% CI −1.30 to 0.52).

Subjective Quality

In the first evaluation, the mean score for the recommendation item (“Would you recommend this app to people who might benefit from it?”) of the data user system was 4.63 (SD 0.52). In the second evaluation, the mean score increased to 4.83 (SD 0.41). This difference was not statistically significant ($t_{11.93}$=−0.84, *P*=.42; *g*=−0.42, 95% CI −1.52 to 0.70.

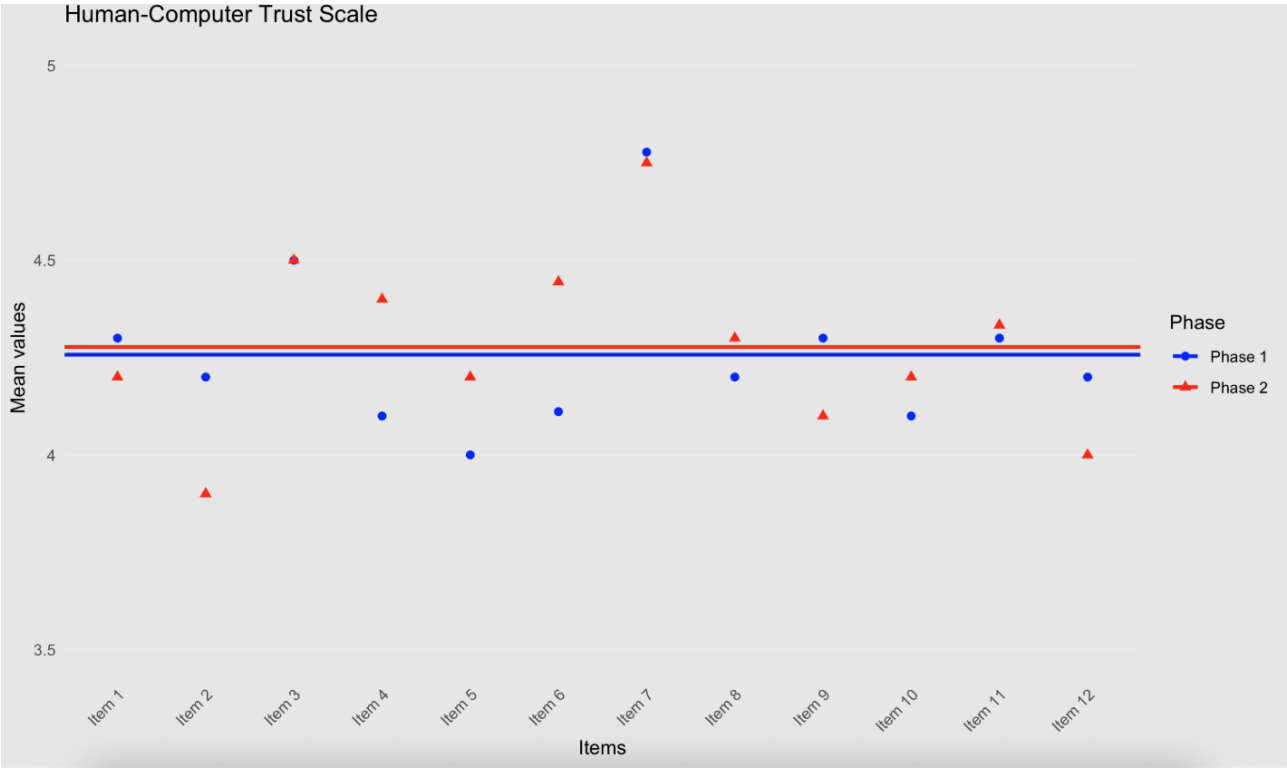
In the first evaluation, the mean score for the recommendation item of the data subject system was 3.90 (SD 0.74). In the second evaluation, the mean score increased to 4.20 (SD 0.79). This difference was not statistically significant ($t_{17.92}$=−0.88, *P*=.39; *g*=−0.38, 95% CI −1.28 to 0.53).

In the first evaluation, data users rated the system with an average of 3.75 out of 5 (SD 0.71) stars and in the second evaluation with 4.17 (SD 0.41) stars. This difference was not statistically significant ($t_{11.44}$=−1.39, *P*=.44; *g*=−0.65, 95% CI −1.78 to 0.48). Data subjects rated the system with an average of 3.90 (SD 0.57) stars in the first evaluation and 4.10 (SD 0.57) stars in the second evaluation. This difference was not statistically significant ($t_{18.00}$=−0.79, *P*=.44; *g*=−0.34, 95% CI −1.24 to 0.57).

Trust

In the first evaluation, the assessment of data subject trust resulted in a mean rating of 4.23 (SD 0.46) on the HCTS. One data subject did not respond to the trust scale and was excluded from the analysis. In the second evaluation, the mean rating was 4.23 (SD 0.68; see Figure 5 for item-level mean values). This difference was not statistically significant ($t_{15.78}$=0.03, *P*=.97; *g*=0.01, 95% CI −0.87 to 0.91).

Figure 5. Data subject trust.

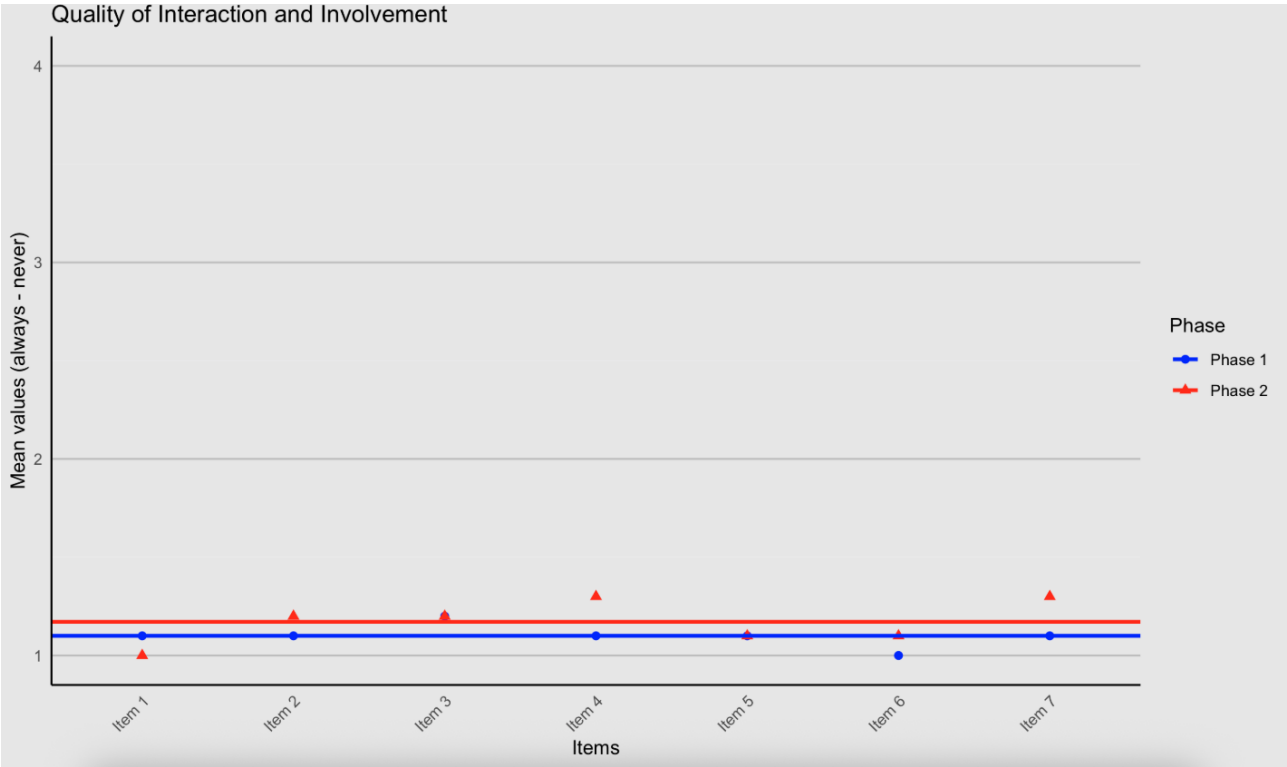


XSL•FO
RenderX

Quality of Interaction and Involvement

In the first evaluation, the assessment of QII resulted in a mean rating of 1.10 (SD 0.19), with lower scores indicating higher perceived quality. In the second evaluation, the mean rating was 1.17 (SD 0.34; see Figure 6 for item-level mean values). This difference was not statistically significant ($t_{14.12}$=−0.58, P=.57; g=−0.26, 95% CI −1.15 to 0.66). In the first evaluation, 1 data subject did not respond to the QII scale and was excluded from the analysis.

Figure 6. Quality of interaction and involvement.



Subjective Understanding

When participants were asked whether they thought the information about their data use was understandable, the mean rating was 4.40 (SD 0.97) in the first evaluation and 4.30 (SD 0.95) in the second evaluation. This difference was not statistically significant ($t_{17.99}$=0.23, P=.82; g=0.10, 95% CI −0.80 to 1.00). The scope/length of the information was rated as “just right” by 8 of 11 (73%) data subjects in the first evaluation, “too long” by 3 of 11 (27%) data subjects, and “too short” by none. In the second evaluation, it was rated as “just right” by 8 of 10 (80%), “too long” by 2 of 10 (20%) data subjects, and “too short” by none. The results on objective understanding can be found in the supplementary material (see Multimedia Appendix 3).

Perceived Risk

Data users and institutional representatives were asked to rate the perceived risk to their institution of using SouveMed. Sixteen (94%) out of 17 responded that they perceived no risk (for the first and second evaluation combined), and 1 of 17 (6%) perceived a risk (for the first and second evaluation combined). Data protection concerns were mentioned as a potential risk.

Discussion

Principal Findings

Making health data accessible for secondary research use is a major challenge in the upcoming decades for both health care institutions and research organizations. In addition to the increasing opportunities created by digitalization and the high willingness of data subjects to make health data available, as the present study and other works show [20-23], the potential is not being realized for a number of reasons. These include changing legal norms, inadequate preprocessing of data, and the existence of a large number of different, sometimes conflicting interests of the various stakeholders.

The SouveMed Concept

Data trust models are a promising solution to address this complexity. However, apart from a few small, heterogeneous initiatives, there are hardly any fully developed approaches in this regard in the field of medical research. We have therefore developed an ethically sound framework for the secondary use of medical data that is applied to the use case of sleep research. The model includes secure onboarding of data subjects and digital use of information, consent, and feedback channels; a digital interface for data users for onboarding, defining research aims and categorization, querying available data set numbers, and, finally, either downloading data or having algorithms run on it; and, crucially, a digital data trust that matches the

XSL•FO
RenderX

descriptions and preconditions of all stakeholders (data subject, data user, and institution) using digital contract representations. Data access is granted in our model only when all requirements and conditions set forth by the various stakeholders are satisfied. The explicit focus of our data trust model is on data subject sovereignty and participation, as data subject involvement and appropriate communication are critical to build trust [25,26] and were central to many data subjects in our needs analysis. Accordingly, a majority of data subjects emphasized the importance of a personal contact person in the consent process to be able to ask individual questions. In addition, transparency about data use was mentioned as an important prerequisite for the willingness to provide data, and there was a clear preference for a consent model that allows for further refinement of the conditions under which data sharing is permitted. Based on these results, the SouveMed model provides the opportunity for data subjects to give *real* informed consent by presenting them with comprehensible information material and encouraging them to ask individualized questions in the information process. By separating information and consent in terms of time and space, the ethically questionable linking of health care and the provision of informed consent is avoided. Participants also have the opportunity to receive information about study projects that use their data. The combination of information in face-to-face contact and via a digital platform enables both personalized information for data subjects and patient sovereignty over the amount of involvement, both key requirements demanded of a system for secondary use of health data (eg, [23,26,31]).

SouveMed uses an opt-in solution where data subjects explicitly decide whether or not they wish to share data. This is in line with generally recognized ethical standards for medical research [18,56]. In addition, the present finding regarding the usefulness of tiered consent provided in an electronic format is in line with previous studies investigating different consent models and modes of patient involvement (eg, [30,57-59]). In our approach, the use of a digital interface for consent management also allows data subjects to easily withdraw their consent, which fulfills the requirement in Article 7 of the GDPR that withdrawal of consent must be as easy as giving consent [15].

Evaluation of the SouveMed Concept

The present evaluation study showed a high degree of trust in our concept in both the first and second evaluations, as measured by the HCTS. It also showed a high level of satisfaction with our communication process, also in the first and second evaluations. Thus, the SouveMed system has proven to be effective in meeting data subjects' main concerns related to the secondary use of health data. The results of the usability assessment using the uMARS showed very good usability for both evaluations, with a slight increase in usability in the second evaluation. Good usability is key to increasing and maintaining data subjects' motivation to actively participate in SouveMed and is probably also related to trust [60].

From the perspective of data users and institutions, they may benefit from easy access to raw data and the ability to run algorithms on data in a secure environment. Both our requirements analysis and our evaluation study have shown that there is an urgent need for systems that enable secondary use of research data. The data on uMARS from our evaluation study showed excellent usability in all three subscales (functionality, aesthetics, and subjective quality) from the perspective of data users and institutions, which improved even further in the second evaluation. In this evaluation, all data users surveyed stated that the SouveMed system offers all the functionalities they desire from such a system. Except for a significant increase in the perceived functionality of the data user system, no statistically significant differences were observed between the first and second evaluations. This result is likely due to the limited sample size and the already high ratings in the first evaluation, which constrained the potential for further measurable improvements.

Overall, these results indicate that the SouveMed data trust effectively realizes the needs and requirements identified during the needs analysis and literature review.

Future Perspectives

Various legislative initiatives in the European Union (eg, EHDS [61]) and in Germany (eg, *Forschungsdatengesetz* [62] and *Gesundheitsdatennutzungsgesetz* [63]) are currently aimed at simplifying the secondary use of health data. This is to be welcomed in principle, as it contributes to the development of a viable research landscape and addresses a major challenge in medical research. Nevertheless, the solutions must be preceded by an analysis of the underlying problems and address them precisely. The results of this and numerous other studies indicate that the secondary use of health data is not hindered by a lack of willingness on the part of data subjects to provide data. It therefore does not seem necessary to pass over data subjects by allowing the use of health data without consent. There is also no lack of infrastructure for centralized data storage, which entails far more risks but only minimal benefits [64]. What is notably absent is a clear workflow for the secondary use of health data that can be easily integrated into routine clinical care, that takes data subjects' needs into account and thereby gains their long-term acceptance, and that relieves researchers of complex legal considerations and negotiations.

In the European Union, a standardized legal framework has been established through the EHDS, which permits an opt-out mechanism for the secondary use of health data [61]. According to Article 54, this should be implemented via "…an easily understandable and accessible user-friendly mechanism to exercise that right to opt out…" Article 54 further emphasizes that it is "…imperative to provide natural persons with sufficient and complete information regarding their right to opt out…" Nevertheless, how this will be implemented in practice in the individual member states is still largely unclear. Although an opt-in solution is preferred in our project, the SouveMed concept could serve as a blueprint for a system that enables data subjects to receive information about the secondary use of their health data and to express their preferences, including a possible opt-out.

Sleep research was selected as a use case because large volumes of data are typically generated in this field (eg, through polysomnography). In this context, it is important to consider that the likelihood of reidentifying individuals grows with the amount of data collected, making true anonymization increasingly difficult. Similar challenges are expected in other

XSL•FO
RenderX

areas of research due to expanding data volumes. Therefore, we argue that the active involvement of data subjects and other stakeholders is becoming ever more critical.

A key advantage of SouveMed is also the algorithmic trust component. Future data exchange will increasingly have to rely on algorithmic decisions to manage large amounts of requests in a time- and cost-efficient manner. Data use and access committees would oversee the rules of data use rather than each single application.

Limitations

There are some limitations to this study. First, the sample size was small and therefore probably not sufficiently representative. Therefore, surveys of larger patient and other stakeholder groups that include other diagnostic entities are needed. In addition, no *real* implementation of the SouveMed system was carried out as there were insufficient resources to achieve this in the current project. Although we tried to keep the framework of the evaluation as natural as possible, everyone involved was aware that it was not about real data sharing, which could have influenced the results.

Many aspects of our model were developed not only conceptually but also technically (see [54,55]). Nevertheless, further studies focusing on the technical implementation of digital contract representations are needed, as this is a key element of an effectively functioning matching algorithm that was not fully elaborated technically in the SouveMed project.

Funding

This study was funded by the German Ministry of Education and Research (grant 16DTM115A). We acknowledge support by the Open Access Publication Fund of the University of Freiburg.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Semistructured interview (interview guide, English version).
[DOCX File, 14 KB - humanfactors_v12i1e66513_app1.docx]

Multimedia Appendix 2
Providing consent via the SouveMed system.
[PNG File, 619 KB - humanfactors_v12i1e66513_app2.png]

Multimedia Appendix 3
Results on objective understanding.
[PNG File, 165 KB - humanfactors_v12i1e66513_app3.png]

References

1. Saberi ZA, Sadr H, Yamaghani MR. An intelligent diagnosis system for predicting coronary heart disease. Presented at: 2024 10th International Conference on Artificial Intelligence and Robotics (QICAR); Feb 29, 2024. [doi: 10.1109/QICAR61538.2024.10496601]
2. Nazari M, Emami H, Rabiei R, Hosseini A, Rahmatizadeh S. Detection of cardiovascular diseases using data mining approaches: application of an ensemble-based model. Cogn Comput 2024 Sep;16(5):2264-2278. [doi: 10.1007/s12559-024-10306-z]
3. Sadr H, Salari A, Ashoobi MT, Nazari M. Cardiovascular disease diagnosis: a holistic approach using the integration of machine learning and deep learning models. Eur J Med Res 2024 Sep 11;29(1):455. [doi: 10.1186/s40001-024-02044-7] [Medline: 39261891]
4. Nazari M, Moayed Rezaie S, Yaseri F, Sadr H, Nazari E. Design and analysis of a telemonitoring system for high-risk pregnant women in need of special care or attention. BMC Pregnancy Childbirth 2024 Dec 18;24(1):817. [doi: 10.1186/s12884-024-07019-4] [Medline: 39696006]
5. Benson T, Grieve G. Why interoperability is hard. In: Benson T, Grieve G, editors. Principles of Health Interoperability: FHIR, HL7 and SNOMED CT:21-40. [doi: 10.1007/978-3-030-56883-2_2]
6. Pohlmann S, Kunz A, Ose D, et al. Digitalizing health services by implementing a personal electronic health record in Germany: qualitative analysis of fundamental prerequisites from the perspective of selected experts. J Med Internet Res 2020 Jan 29;22(1):e15102. [doi: 10.2196/15102] [Medline: 32012060]
7. Edmondson ME, Reimer AP. Challenges frequently encountered in the secondary use of electronic medical record data for research. Comput Inform Nurs 2020 Jul;38(7):338-348. [doi: 10.1097/CIN.0000000000000609] [Medline: 32149742]

XSL•FO
RenderX

8. Declerck J, Kalra D, Vander Stichele R, Coorevits P. Frameworks, dimensions, definitions of aspects, and assessment methods for the appraisal of quality of health data for secondary use: comprehensive overview of reviews. JMIR Med Inform 2024 Mar 6;12(1):e51560. [doi: 10.2196/51560] [Medline: 38446534]
9. Robertson ARR, Fernando B, Morrison Z, Kalra D, Sheikh A. Structuring and coding in health care records: a qualitative analysis using diabetes as a case study. jhi 2014;22(2):275-283. [doi: 10.14236/jhi.v22i2.90]
10. de Hert P, Papakonstantinou V. The new General Data Protection Regulation: still a sound system for the protection of individuals? Comput Law Security Rev 2016 Apr;32(2):179-194. [doi: 10.1016/j.clsr.2016.02.006]
11. Dove ES. The EU General Data Protection Regulation: implications for international scientific research in the digital era. J Law Med Ethics 2018;46(4):1013-1030. [doi: 10.1177/1073110518822003]
12. Rumbold JMM, Pierscionek B. The effect of the General Data Protection Regulation on medical research. J Med Internet Res 2017 Feb 24;19(2):e47. [doi: 10.2196/jmir.7108] [Medline: 28235748]
13. Vukovic J, Ivankovic D, Habl C, Dimnjakovic J. Enablers and barriers to the secondary use of health data in Europe: General Data Protection Regulation perspective. Arch Public Health 2022 Apr 9;80(1):115. [doi: 10.1186/s13690-022-00866-7] [Medline: 35397557]
14. Peloquin D, DiMaio M, Bierer B, Barnes M. Disruptive and avoidable: GDPR challenges to secondary research uses of data. Eur J Hum Genet 2020 Jun;28(6):697-705. [doi: 10.1038/s41431-020-0596-x] [Medline: 32123329]
15. Regulation (EU) 2016/679 of the European Parliament and of the council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive 95/46/EC (General Data Protection Regulation). European Union. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679 [accessed 2024-02-27]
16. Wiertz S, Boldt J. Evaluating models of consent in changing health research environments. Med Health Care Philos 2022 Jun;25(2):269-280. [doi: 10.1007/s11019-022-10074-3] [Medline: 35286521]
17. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human participants. JAMA 2025 Jan 7;333(1):71-74. [doi: 10.1001/jama.2024.21972]
18. WMA Declaration of Taipei on ethical considerations regarding health databases and biobanks. The World Medical Association. URL: https://www.wma.net/policies-post/wma-declaration-of-taipei-on-ethical-considerations-regarding-health-databases-and-biobanks/ [accessed 2024-02-27]
19. Mondschein CF, Monda C. The EU's General Data Protection Regulation (GDPR) in a research context. In: Kubben P, Dumontier M, Dekker A, editors. Fundamentals of Clinical Data Science: Springer International Publishing:2019. [doi: 10.1007/978-3-319-99713-1_5]
20. Richter G, Borzikowsky C, Lesch W, et al. Secondary research use of personal medical data: attitudes from patient and population surveys in The Netherlands and Germany. Eur J Hum Genet 2021 Mar;29(3):495-502. [doi: 10.1038/s41431-020-00735-3] [Medline: 33005018]
21. Richter G, Krawczak M, Lieb W, Wolff L, Schreiber S, Buyx A. Broad consent for health care-embedded biobanking: understanding and reasons to donate in a large patient sample. Genet Med 2018 Jan;20(1):76-82. [doi: 10.1038/gim.2017.82] [Medline: 28640237]
22. Richter G, Borzikowsky C, Lieb W, Schreiber S, Krawczak M, Buyx A. Patient views on research use of clinical data without consent: legal, but also acceptable? Eur J Hum Genet 2019 Jun;27(6):841-847. [doi: 10.1038/s41431-019-0340-6] [Medline: 30683927]
23. Garrison NA, Sathe NA, Antommaria AHM, et al. A systematic literature review of individuals' perspectives on broad consent and data sharing in the United States. Genet Med 2016 Jul;18(7):663-671. [doi: 10.1038/gim.2015.138] [Medline: 26583683]
24. Hutchings E, Loomes M, Butow P, Boyle FM. A systematic literature review of attitudes towards secondary use and sharing of health administrative and clinical trial data: a focus on consent. Syst Rev 2021 May 4;10(1):132. [doi: 10.1186/s13643-021-01663-z] [Medline: 33941282]
25. Hutchings E, Loomes M, Butow P, Boyle FM. A systematic literature review of health consumer attitudes towards secondary use and sharing of health administrative and clinical trial data: a focus on privacy, trust, and transparency. Syst Rev 2020 Oct 9;9(1):235. [doi: 10.1186/s13643-020-01481-9] [Medline: 33036664]
26. Cumyn A, Ménard JF, Barton A, Dault R, Lévesque F, Ethier JF. Patients' and members of the public's wishes regarding transparency in the context of secondary use of health data: scoping review. J Med Internet Res 2023 Apr 13;25:e45002. [doi: 10.2196/45002] [Medline: 37052967]
27. Park YJ, Chung JE. Health privacy as sociotechnical capital. Comput Human Behav 2017 Nov;76:227-236. [doi: 10.1016/j.chb.2017.07.025]
28. Damschroder LJ, Pritts JL, Neblo MA, Kalarickal RJ, Creswell JW, Hayward RA. Patients, privacy and trust: patients' willingness to allow researchers to access their medical records. Soc Sci Med 2007 Jan;64(1):223-235. [doi: 10.1016/j.socscimed.2006.08.045] [Medline: 17045717]
29. Holm S, Kristiansen TB, Ploug T. Control, trust and the sharing of health information: the limits of trust. J Med Ethics 2021 Dec;47(12):e35-e35. [doi: 10.1136/medethics-2019-105887]

XSL•FO
RenderX

30. Broes S, Verbaanderd C, Casteels M, Lacombe D, Huys I. Sharing of clinical trial data and samples: the cancer patient perspective. Front Med 2020 Feb;7:33. [doi: 10.3389/fmed.2020.00033]
31. Cumyn A, Dault R, Barton A, Cloutier AM, Ethier JF. Citizens, research ethics committee members and researchers' attitude toward information and consent for the secondary use of health data: implications for research within learning health systems. J Empir Res Hum Res Ethics 2021 Jul;16(3):165-178. [doi: 10.1177/1556264621992214] [Medline: 33710932]
32. King T, Brankovic L, Gillard P. Perspectives of Australian adults about protecting the privacy of their health information in statistical databases. Int J Med Inform 2012 Apr;81(4):279-289. [doi: 10.1016/j.ijmedinf.2012.01.005] [Medline: 22306206]
33. Liddell K, Simon DA, Lucassen A. Patient data ownership: who owns your health? J Law Biosci 2021;8(2):lsab023. [doi: 10.1093/jlb/lsab023] [Medline: 34611493]
34. Ballantyne A. How should we think about clinical data ownership? J Med Ethics 2020 May;46(5):289-294. [doi: 10.1136/medethics-2018-105340] [Medline: 31911499]
35. Hummel P, Braun M, Dabrock P. Own data? Ethical reflections on data ownership. Philos Technol 2021 Sep;34(3):545-572. [doi: 10.1007/s13347-020-00404-9]
36. Perrier L, Blondal E, MacDonald H. The views, perspectives, and experiences of academic researchers with data sharing and reuse: a meta-synthesis. PLoS ONE 2020;15(2):e0229182. [doi: 10.1371/journal.pone.0229182] [Medline: 32106224]
37. Tenopir C, Allard S, Douglass K, et al. Data sharing by scientists: practices and perceptions. PLoS ONE 2011;6(6):e21101. [doi: 10.1371/journal.pone.0021101] [Medline: 21738610]
38. Paprica PA, Sutherland E, Smith A, et al. Essential requirements for establishing and operating data trusts: practical guidance co-developed by representatives from fifteen Canadian organizations and initiatives. Int J Popul Data Sci 2020 Aug 24;5(1):1353. [doi: 10.23889/ijpds.v5i1.1353] [Medline: 33644412]
39. Milne R, Sorbie A, Dixon-Woods M. What can data trusts for health research learn from participatory governance in biobanks? J Med Ethics 2022 May;48(5):323-328. [doi: 10.1136/medethics-2020-107020] [Medline: 33741681]
40. Hogan WR, Shenkman EA, Robinson T, et al. The OneFlorida Data Trust: a centralized, translational research data infrastructure of statewide scope. J Am Med Inform Assoc 2022 Mar 15;29(4):686-693. [doi: 10.1093/jamia/ocab221] [Medline: 34664656]
41. Hall W, Pesenti J. Growing the artificial intelligence industry in the UK. Australian Policy Online. 2017 Oct 15. URL: https://apo.org.au/node/114781 [accessed 2024-03-04]
42. Lauf F, Scheider S, Friese J, Kilz S, Radic M, Burmann A. Exploring design characteristics of data trustees in healthcare - taxonomy and archetypes. Presented at: Proceedings of the thirty-first European Conference on Information Systems; Jun 11-16, 2023 URL: https://aisel.aisnet.org/ecis2023_rp/323/ [accessed 2025-11-18]
43. Poschen C, Herres B, Knorr K. A threat-driven design of a data-trustee infrastructure for medical data. Presented at: 2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM); Dec 3-6, 2024. [doi: 10.1109/BIBM62325.2024.10822788]
44. Prokosch HU, Gebhardt M, Gruendner J, et al. Towards a national portal for medical research data (FDPG): vision, status, and lessons learned. Stud Health Technol Inform 2023 May 18;302:307-311. [doi: 10.3233/SHTI230124] [Medline: 37203668]
45. Semler S, Wissing F, Heyder R. German Medical Informatics Initiative: a national approach to integrating health data from patient care and medical research. Methods Inf Med 2018 Jul;57(S 01):e50-e56. [doi: 10.3414/ME18-03-0003]
46. Riemann D, Espie CA, Altena E, et al. The European Insomnia Guideline: an update on the diagnosis and treatment of insomnia 2023. J Sleep Res 2023 Dec;32(6):e14035. [doi: 10.1111/jsr.14035]
47. Senaratna CV, Perret JL, Lodge CJ, et al. Prevalence of obstructive sleep apnea in the general population: a systematic review. Sleep Med Rev 2017 Aug;34:70-81. [doi: 10.1016/j.smrv.2016.07.002] [Medline: 27568340]
48. Perez-Pozuelo I, Zhai B, Palotti J, et al. The future of sleep health: a data-driven revolution in sleep science and medicine. NPJ Digit Med 2020;3(1):42. [doi: 10.1038/s41746-020-0244-4] [Medline: 32219183]
49. Gulati S, Sousa S, Lamas D. Design, development and evaluation of a human-computer trust scale. Behav Inf Technol 2019 Oct 3;38(10):1004-1015. [doi: 10.1080/0144929X.2019.1656779]
50. Sousa S, Lamas D, Dias P. A model for human-computer trust: contributions towards leveraging user engagement. In: Zaphiris P, Ioannou A, editors. Learning and Collaboration Technologies Designing and Developing Novel Learning Experiences: Springer International Publishing; 2014:128-137. [doi: 10.1007/978-3-319-07482-5_13]
51. Abdelhamid M, Gaia J, Sanders GL. Putting the focus back on the patient: how privacy concerns affect personal health information sharing intentions. J Med Internet Res 2017 Sep 13;19(9):e169. [doi: 10.2196/jmir.6877] [Medline: 28903895]
52. Messner EM, Terhorst Y, Barke A, et al. The German version of the Mobile App Rating Scale (MARS-G): development and validation study. JMIR Mhealth Uhealth 2020 Mar 27;8(3):e14479. [doi: 10.2196/14479] [Medline: 32217504]
53. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the user version of the Mobile Application Rating Scale (uMARS). JMIR Mhealth Uhealth 2016 Jun 10;4(2):e72. [doi: 10.2196/mhealth.5849]

XSL•FO
RenderX

54. Burmeister R, Erler C, Gauger F, Dressle RJ, Feige B. Advancing sleep research through dynamic consent and trustee-based medical data processing. 2024 Presented at: Proceedings of the Eighteenth International Conference on Digital Society; May 26-30, 2024 URL: https://www.iaria.org/conferences2024/ICDS24.html [accessed 2025-11-06]
55. Erler C, Bauer AM, Gauger F, Stork W. Decision model to design trust-focused and blockchain-based health data management applications. Blockchains 2024 Jun;2(2):79-106. [doi: 10.3390/blockchains2020005]
56. WMA Declaration of Helsinki – ethical principles for medical research involving human subjects. The World Medical Association. 2013. URL: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/ [accessed 2024-02-27]
57. Spencer K, Sanders C, Whitley EA, Lund D, Kaye J, Dixon WG. Patient perspectives on sharing anonymized personal health data using a digital system for dynamic consent and research feedback: a qualitative study. J Med Internet Res 2016 Apr 15;18(4):e66. [doi: 10.2196/jmir.5011] [Medline: 27083521]
58. Wiertz S, Boldt J. Ethical, legal, and practical concerns surrounding the implemention of new forms of consent for health data research: qualitative interview study. J Med Internet Res 2024 Aug 7;26:e52180. [doi: 10.2196/52180] [Medline: 39110970]
59. Budin-Ljøsne I, Teare HJA, Kaye J, et al. Dynamic consent: a potential solution to some of the challenges of modern biomedical research. BMC Med Ethics 2017 Jan 25;18(1):4. [doi: 10.1186/s12910-016-0162-9] [Medline: 28122615]
60. Casaló LV, Flavián C, Guinalíu M. Generating trust and satisfaction in e-services: the impact of usability on consumer behavior. J Relationship Marketing 2010 Nov 30;9(4):247-263. [doi: 10.1080/15332667.2010.522487]
61. Regulation (EU) 2025/327 of the European Parliament and of the council of 11 February 2025 on the European Health Data Space and amending directive 2011/24/EU and regulation (EU) 2024/2847. European Union. 2025. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202500327 [accessed 2025-05-25]
62. Eckpunkte BMBF Forschungsdatengesetz. Bundesministerium für Bildung und Forschung. 2024. URL: https://www.bmbf.de/SharedDocs/Downloads/DE/gesetze/forschungsdatengesetz/sonstige/Eckpunktepapier.pdf?__blob=publicationFile&v=3 [accessed 2025-06-22]
63. Gesetz zur verbesserten Nutzung von Gesundheitsdaten. Bundesrepublik Deutschland. 2024. URL: https://www.recht.bund.de/bgbl/1/2024/102/VO.html [accessed 2025-06-22]
64. Hallock H, Marshall SE, 't Hoen PAC, et al. Federated networks for distributed analysis of health data. Front Public Health 2021;9:712569. [doi: 10.3389/fpubh.2021.712569] [Medline: 34660512]

Abbreviations

DOH: Declaration of Helsinki
DOT: Declaration of Taipei
EHDS: European Health Data Space
GDPR: General Data Protection Regulation
HCTS: Human-Computer Trust Scale
MARS: Mobile Application Rating Scale
QII: Quality of Interaction and Involvement
UI: user interface
uMARS: user version of the Mobile Application Rating Scale

Edited by A Kushniruk, Y Gong; submitted 15.09.24; peer-reviewed by B Aretz, H Sadr, M Nayak; revised version received 24.06.25; accepted 01.08.25; published 02.12.25.

Please cite as:
Dressle RJ, Riemann D, Thoma N, Erler C, Burmeister R, Jogwitz B, Domschke K, Spiegelhalder K, Boldt J, Wiertz S, Feige B
Developing a Data Trust Model (Not Only) for Sleep Research: Conceptual Study and Quantitative Survey
JMIR Hum Factors 2025;12:e66513
URL: https://humanfactors.jmir.org/2025/1/e66513
doi:10.2196/66513

© Raphael Jan Dressle, Dieter Riemann, Nicole Thoma, Christina Erler, Rodger Burmeister, Bianka Jogwitz, Katharina Domschke, Kai Spiegelhalder, Joachim Boldt, Svenja Wiertz, Bernd Feige. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 2.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic

XSL•FO
RenderX

information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Digital Health Literacy in Adults With Low Reading and Writing Skills Living in Germany: Mixed Methods Study

Saskia Muellmann[1,2], PhD; Rebekka Wiersing[1,3], MA; Hajo Zeeb[1,2,4], Prof Dr; Tilman Brand[1,2,3], PhD

[1]Leibniz Institute for Prevention Research and Epidemiology - BIPS, Achterstraße 30, Bremen, Germany

[2]Leibniz ScienceCampus Digital Public Health, Bremen, Germany

[3]Leibniz Living Lab, Bremen, Germany

[4]Faculty of Human and Health Sciences, University of Bremen, Bremen, Germany

Corresponding Author:
Saskia Muellmann, PhD
Leibniz Institute for Prevention Research and Epidemiology - BIPS, Achterstraße 30, Bremen, Germany

Abstract

Background: Digital health literacy is a key factor in enabling users to navigate in an increasingly digitalized health care system. Low levels of digital health literacy are associated with higher age, low education, and income, as well as low functional health literacy. Around 6.2 million adults living in Germany have low reading and writing skills. Due to their low literacy, this group is often underrepresented in research studies and therefore little is known about their digital health literacy and use of digital health tools.

Objective: The objectives of this study were to assess digital health literacy in adults with low reading and writing skills and to explore which digital health tools they use in daily life.

Methods: An interviewer-administered survey and focus groups were conducted with adult residents of Bremen, Germany, who were aged 18 - 64 years and had low reading and writing skills. In addition, a stakeholder workshop was held to derive recommendations on how digital health literacy could be improved. The survey questionnaire included 21 items addressing the use of digital health technologies and digital health literacy (eHealth Literacy Scale). Focus group participants completed several tasks on web-based health information and then discussed their experiences. Survey data were analyzed using descriptive statistics and linear regression. Qualitative content analysis was applied to analyze the focus group data and the written documentation of the stakeholder workshop.

Results: Survey participants (n=96) were on average 43 (SD 10.7) years old, 72% (69/96) were female, and 92% (88/96) were not born in Germany. Participants reported mainly using information-related digital health technologies such as health apps (40/96, 42%), health websites (30/96, 31%), or activity trackers (27/96, 28%). The mean digital health literacy score was 22 (SD 8) points, with 35% (34/96) of participants classified as having a low digital health literacy (score between 8-19/40 points). Digital health technology use was associated with higher digital health literacy. For participants in the 5 focus groups (total n=39; mean age 43, SD 12.6 years; n=34, 87% female), limited technical skills and language problems were the most important challenges. Furthermore, focus group participants reported that they favor videos when searching for web-based health information and prefer to seek support from family members or local organizations for health issues. Stakeholders (n=15) recommended that health websites should be available in multiple languages, contain simple and easy-to-read language, and use images, symbols, and videos.

Conclusions: While adults with low reading and writing skills use digital health technologies, many find it challenging to search for health information on the internet due to lacking technical skills and language problems. To ensure that adults with low reading and writing skills are not further left behind, future research should focus on developing tailored interventions to promote digital health literacy.

(JMIR Hum Factors 2025;12:e65345) doi:10.2196/65345

KEYWORDS

digital health literacy; digital health technologies; digital health; digital divide; mobile phone; adults; mixed methods study

Introduction

The digitalization of health care services, such as video consultations or prescribed digital health applications, is accompanied by a shift of prevention and health promotion activities and interventions to the digital environment (eg, wearables and online sport courses) [1,2]. Health information from the internet is easily and instantly available [3]. However,

XSL•FO
RenderX

the large amount of health information on the internet is often overwhelming for users, and it is particularly difficult to use the information found to make health decisions [4]. The level of digital health literacy is a key factor in enabling users to navigate in an increasingly digitalized health care system. According to one well-known definition, digital health literacy describes the ability to seek, find, understand, evaluate, and apply health information from digital sources to one's own health problem [5]. In addition, interaction with health apps (eg, for tracking physical activity) or interactions on social media platforms, including creating own content, are also included under current concepts of digital health literacy [6,7]. Hence, digital health literacy combines various subskills such as computer literacy (ie, use of computers and digital devices), information and media literacy (ie, critical use of information and media), science literacy (ie, knowledge and understanding of scientific principles, concepts, and processes), traditional literacy (ie, reading and writing skills), and health literacy [1]. Studies show that digital health literacy is unevenly distributed among population groups and influenced by sociodemographic factors such as age, education, income, and ethnicity. Older adults and persons with low levels of education and income have been found to have lower digital health literacy than their counterparts [6,8,9], whereas ethnicity was not linked with digital health literacy in the systematic review and meta-analysis by Estrela and colleagues [6]. One study found that being a Spanish-speaker in the United States was not associated with digital health literacy [10]. However, other studies found a lower level of digital health literacy among Arabic native speakers in Sweden [11] and non-English versus English-speaking parents of children with special health care needs in the United States [12]. The second Health Literacy Survey Germany did not find an association of digital health literacy with migration background per se, however, that survey showed an association of low functional health literacy, used as a proxy for general literacy (reading and writing skills and numeracy), with low digital health literacy [7,13]. This is important as an estimated 12% of the adult German population (ie, 6.2 million adults) have low reading and writing skills, meaning that they can only read and write simple sentences in German [14]. Around half of the participants with low reading and writing skills reported to have another native language than German. In addition, adults with low reading and writing skills are less likely to write emails or search for health-related information on the internet compared with adults with sufficient reading and writing skills [14]. Even though several studies have already investigated digital health literacy of the general population living in Germany [4,7,15] and various subgroups such as students [16], school children [17], teachers [18], physicians [19], or community inhabitants [20], very little is known about the digital health literacy of adults with low reading and writing skills and their use of digital tools for health purposes. Therefore, the main objectives of this study were (1) to assess the digital health literacy in adults with low reading and writing skills and (2) to explore which digital tools they use in daily life.

Methods

Study Design

This is an exploratory study using a mixed methods design consisting of a survey, focus group discussions, and a stakeholder workshop. The study report adheres to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) [21] (Checklist 1) and Consolidated Criteria for Reporting Qualitative Research (COREQ) [22] (Checklist 2) guidelines.

Participants

Inclusion criteria for the survey and focus group discussions were (1) age between 18 and 64 years; (2) living in Bremen, Germany; and (3) low reading and writing skills in German (ie, able to read and write simple sentences in German only; alpha level <4). Eligible participants could be native German speakers, adults who could read and write in their native language but not in German, or adults who could not read and write either in their native language or German.

Procedure

We aimed to conduct 70 survey interviews and 5 focus groups with 6-8 participants each, striving for a balanced ratio of participants with different literacy levels (ie, native German speakers, good literacy in native language but not in German, and low literacy in either native language and German). Survey and focus group participants were recruited via a snowball system. Health mediators employed at the research institution for the project duration recruited participants via their local community networks (eg, mothers' centers and integration courses). Flyers were displayed in various locations (eg, mothers' centers and meeting centers for citizens) and distributed during events organized by our community office (ie, Leibniz Living Lab) or cooperating city district initiatives. In addition, the project was presented in German language courses and integration courses, and eligible participants were approached personally via the health mediators and their local networks. The health mediators had a native language other than German (eg, Russian or Turkish) but were fluent German speakers and, due to their main occupation, were well-connected in the community and with the people who had immigrated to Germany living there. Participants were also recruited from German language courses provided by adult education centers. The language courses aimed to teach reading and writing skills to native German speakers and people who have immigrated to Germany and could understand and speak German.

All participants recruited by the health mediators were asked to complete a 9-item test to determine their literacy level [23]. Participants recruited from adult education centers did not have to complete this test as their literacy level had already been assessed for their language course attendance.

Survey Questionnaire

The structured interviewer-administered survey took approximately 30 minutes to complete and included 21 items covering three topic areas: (1) use and acceptance of digital health technologies (5 items), (2) digital health literacy (8 items),

XSL•FO
RenderX

and (3) sociodemographic characteristics (8 items). The items were self-developed or adapted from existing surveys or instruments. The German version of the questionnaire can be found in Multimedia Appendix 1.

Use and acceptance of digital health technologies were assessed based on purposes of digital devices, use of digital devices in health context, use of digital health technologies (ie, list with 12 digital health technologies), perceived advantages, and disadvantages [24,25].

Digital health literacy was measured using the German version of the eHealth Literacy Scale (eHEALS) [26]. Each item was rated on a 5-point Likert scale ranging from strongly disagree to strongly agree.

Sociodemographic characteristics included age, sex, country of birth, spoken language at home, self-assessed speaking, reading, and writing skills in native language, highest educational degree, and employment status. Items for age, sex, country of birth, spoken language at home, highest educational degree, and employment status were adapted from the "German Health Update" 2019/2020 study [27]. The item for self-assessed reading and writing skills in native language was self-developed.

The interviewer-administered questionnaire was filled out by a health mediator or a member of the research team during a personal interview conducted in our community office (ie, Leibniz Living Lab), mother centers, or adult education centers. Based on the participant's preference, the interview was conducted in German, English, Russian, Arabic, Turkish, Macedonian, Twi, or Tamil.

Focus Group Protocol

The approximately 90-minute long focus groups were structured into four parts: (1) welcome and introduction, (2) task course, (3) overarching discussion, and (4) summary and farewell.

During the welcome and introduction, participants received the study information in verbal and written form, gave informed consent to participate, and filled out a short questionnaire on sociodemographic information (ie, age, sex, and employment status). During the focus group, participants were presented with 4 tasks which were developed on the basis of the Digital Health Literacy Instrument (DHLI) by van der Vaart and Drossaert [28]. For each of the 7 DHLI dimensions (ie, operational skills, navigation skills, information searching, evaluating reliability, determining relevance, adding content, and protecting privacy), a performance-based item was developed that included the application of a particular skill in a fictional situation [28]. We adapted the proposed items for our context and target group, focusing on the dimensions of operational skills, navigation skills, information searching, evaluating reliability, and determining relevance (Textbox 1). The tasks were not intended to measure participants' digital health literacy. All tasks were formulated in simple language and were read out to the participants. In addition, pictures and symbols were used to describe the content of the tasks. Participants solved tasks individually on their smartphones or using a laptop provided by the study team. If necessary, participants were supported by the study team while completing the tasks. After each task, a group discussion on encountered difficulties and strategies to overcome difficulties was held (ie, what experiences did you have while working on the task? What was difficult for you? What did you do if you did not make any progress in the search?). In addition, the study team described a possible solution for each task. In the overarching discussion, participants discussed which digital devices they use, which formats (eg, videos) they prefer, and how they use web-based services for health-related purposes. At the end of the focus group, the study team provided a summary and gave out a handout to the participants with solutions to the tasks. Focus group discussions were conducted in German at our community office (ie, Leibniz Living Lab), mother centers, or adult education centers, and were moderated by at least 3 members of the study team. Focus group discussions were audio-recorded and transcribed verbatim. The focus group protocol is available from the authors upon request.

Textbox 1. Tasks for focus group discussion.

- Please open the internet on your laptop or smartphone. Please close the internet. Please open the website www.bremen.de (website with information and facts about the city of Bremen). Please open a search engine such as Google search (Digital Health Literacy Instrument [DHLI] dimensions: operational skills, navigation skills).
- Please search the internet for a general physician in your neighborhood. Note the telephone number to make an appointment (DHLI dimensions: information searching).
- Please open the website www.gesundheitsinformation.de (website for independent and evidence-based health information). When does a child with fever need to see a physician? (DHLI dimensions: information searching, determining relevance)
- Please search the internet. When do you need to see a physician if you have severe back pain? (DHLI dimensions: information searching, evaluating reliability, determining relevance)

Stakeholder Workshop

The main project results were presented and discussed with local stakeholders in a workshop conducted in our community office (ie, Leibniz Living Lab) using the world café method [29]. Stakeholders were recruited via a snowball system. Thus, stakeholders from civil and public services, local public authorities, and community social service providers or groups that work with adults with low reading and writing skills and are located in Bremen were asked to participate in the workshop. After presenting project results (ie, survey and focus group discussions), the stakeholders split into three groups (approximately 5 stakeholders per group) and discussed the following questions at group tables for 15 minutes each: (1) which competencies are needed by adults with low reading and writing skills to use digital health technologies, (2) how should

XSL•FO
RenderX

digital health technologies be designed for adults with low reading and writing skills, and (3) which and how stakeholders should be involved in making digital health technologies for adults with low reading and writing skills accessible. After 15 minutes, the next group went to the table until all 3 groups had discussed all the questions. Each group table was moderated by 1 researcher (SM, RW, or TB) and 1 - 2 health mediators. The results of the world café were documented on meta-plan paper. In addition, discussion notes for each group and photo documentation were completed. All participants received a copy of the discussion notes and world café results.

Data Analysis

Filled-out survey questionnaires were entered independently into the software LimeSurvey by 2 members of the study team (SM and Finola Goepfert; refer to Acknowledgments). The data were then imported into IBM-SPSS 24 (IBM Corp), cleaned, and prepared for statistical analysis. The responses to all survey items were analyzed descriptively using absolute and relative frequencies or means and SDs. Linear ordinary least square regression models were computed to investigate the association between digital health literacy and participant characteristics including age, sex, formal education, language skills, and use of digital health technologies. Based on the distribution of scores, formal education was categorized into 3 groups (low: less than 9 years of schooling, no formal degree; medium: 9 - 10 years of schooling, lower secondary education; and high: 12 - 13 years of schooling, university entrance qualification), employment status into 5 groups (full-time employed, part-time employed, care work, language course, and other), and use of digital health technologies into 2 groups (no use vs use of at least 1 digital health technology). Participants were grouped into three categories of language skills based on language spoken at home and self-assessed reading and writing skills in native language: (1) native German speakers, (2) good literacy in native language but not in German, and (3) low literacy in both native language and German. For digital health literacy, two variables were computed: (1) total eHEALS score, ranging from 8 (lowest digital health literacy) to 40 points (highest digital health literacy) and (2) a 3-category variable based on the study by De Santis and colleagues [24] (low: 8 - 19 points, medium: 20 - 29 points, and high: 30 - 40 points).

Qualitative content analysis [30] following a deductive-inductive approach was applied to analyze the focus group data. First, focus group transcripts were imported into MAXQDA 2020 (VERBI–Software Consult Sozialforschung GmbH) and relevant text units were coded into overarching categories and subcategories. Categories and subcategories were derived from the focus group protocol (deductive coding). New categories and subcategories were created if text units did not fit into any of the pre-existing categories or subcategories (inductive coding). Second, data were grouped and summarized into themes using Microsoft Excel. Coding, grouping, and summarizing of the data was performed by one researcher (SM) and checked by a second researcher (Meret Lakeberg; refer to Acknowledgments). Any discrepancies in the coding or reduction process were discussed until a consensus between researchers was reached. Illustrative quotes were selected from the existing material as anchor examples and translated into English. All translations were checked by a native English speaker.

As this is an exploratory study, no sample size calculation was conducted.

Ethical Considerations

This study was approved by the Ethics Committee of the University of Bremen, Germany, on December 6, 2022 (reference 2022 - 29), in accordance with the ethical guidelines for human participant research and the Declaration of Helsinki. All participants received verbal and written information about the study and gave informed consent (Multimedia Appendix 2) for their data to be used. All data were anonymized to ensure privacy and were securely stored on a password-protected server at the research institution accessible only to the research team. Participation in the survey and focus group discussions was voluntary, could be terminated at any time, and was reimbursed with €15 (US $17). No identifiable participant information has been included in the manuscript or supplementary materials.

Results

Survey

A total of 108 interviews were conducted between December 2022 and March 2023. Mean interview duration was 29 (range 6-75) minutes. Of the total, 12 interviews were excluded based on the following reasons: participants aged 65 years and above (n=5) or not meeting criteria for low literacy (n=7). The remaining 96 interviews were included in the analysis. Of these, around half were conducted in German (n=47, 49%), 13 in Russian (14%), 12 in Twi (13%), 8 in Arabic (8%), 6 in English (6%), and 4 in Turkish (4%). None of the interviews was conducted either in Macedonian or Tamil, respectively. For 6 interviews, no information on the conducted language is available. Interview participants were on average 43 (SD 10.7) years old, 72% (69/96) were female, 92% (88/96) were not born in Germany, 38% (36/96) reported a medium level of formal education, and 30% (29/96) attended a language course (Table 1). In total, 11 participants (12%) were German native speakers, 32 participants (33%) reported good literacy in their native language but not in German, and 53 participants (55%) had low literacy in both German and their native language. Mean digital health literacy score was 22.2 (SD 8) points ranging from 8 to 40 points. Digital health literacy was classified as low in 35% (n=34), medium in 44% (n=42), and high in 19% (n=18) of participants. In addition, 59 participants (62%) were digital health technology users.

Regression analysis found no associations between age, sex, and digital health literacy (Table 2). Digital health literacy was approximately 6 scale points lower among participants with low language skills in both their native language and German as compared with native speakers. Participants with a medium level of formal education reported a higher digital health literacy compared with those with no formal degree. Furthermore, the average digital health literacy score was approximately 10 points higher among digital health technology users compared with nonusers.

XSL•FO
RenderX

The 5 most commonly used health-related digital technologies were health apps (40/96, 42%), health-related websites or forums (30/96, 31%), fitness trackers (27/96, 28%), web-based appointment scheduling (24/96, 25%), and email contact with doctor's surgery or pharmacy (20/96, 21%) (Figure 1). Health care–related digital technologies, such as prescribed digital health applications (3/96, 3%), electronic sick leave notice (3/96, 3%), or electronic prescription (1/96, 1%), were used by very few participants. None of the participants reported accessing their electronic health record or using video consultations. Participants with low digital health literacy used fewer digital health technologies than those with medium or high digital health literacy, and specifically none of the health care–related options (Figure 1). Results stratified by language skills showed higher usage of digital health technologies by native German speakers and adults with good native language literacy compared with participants with both native language and German literacy (Multimedia Appendix 3).

Most participants (n=74, 77%) used a smartphone or tablet to access health-related information (Table 3). The most common reasons for using digital health technologies were time savings (n=44, 46%), access at anytime from anywhere (n=42, 44%), motivation to stay healthy (n=39, 41%), and cost-effectiveness (n=38, 40%). Reasons for not using digital health technologies included preference for personal advice (n=41, 43%) or paper-based information (n=33, 34%), concerns about data security (n=20, 21%), mistrust in health information (n=19, 20%), technical problems (n=9, 9%) and lack of a suitable device (n=6, 6%).

XSL•FO
RenderX

Table . Participant characteristics.

Characteristic		Survey (n=96)	Focus groups (n=39)[a]
Age (years), mean (SD)		43.1 (10.7)	43.1 (12.6)
Sex, n (%)			
	Female	69 (72)	34 (87)
	Male	27 (28)	5 (13)
Country of birth, n (%)			
	Germany	8 (8)	—[b]
	Other than Germany	88 (92)	—
Formal education[c], n (%)			
	Low	37 (39)	—
	Medium	36 (38)	—
	High	22 (23)	—
	Unknown	1 (1)	—
Employment status, n (%)			
	Full-time employed	15 (16)	3 (8)
	Part-time employed	10 (10)	6 (15)
	Care work	23 (24)	8 (21)
	Language course	29 (30)	18 (46)
	Other (eg, job seeking, unable to work)	17 (18)	4 (10)
	Unknown	2 (2)	0 (0)
Language skills, n (%)			
	Native German speakers	11 (12)	—
	Good literacy in native language but not in German	32 (33)	—
	Low literacy in both native language and German	53 (55)	—
Digital health literacy			
	Mean score (SD)	22.2 (8)	—
	Low (8 - 19 points), n (%)	34 (35)	—
	Medium (20 - 29 points), n (%)	42 (44)	—
	High (30 - 40 points), n (%)	18 (19)	—
	Unknown, n (%)	2 (2)	—
Use of digital health technologies, n (%)			
	User	59 (62)	—
	Nonuser	37 (38)	—

[a]Only age, sex, and employment status were assessed.

[b]Not assessed.

[c]Low: less than 9 years of schooling, no formal degree; medium: 9 - 10 years of schooling, lower secondary education; high: 12 - 13 years of schooling, university entrance qualification.

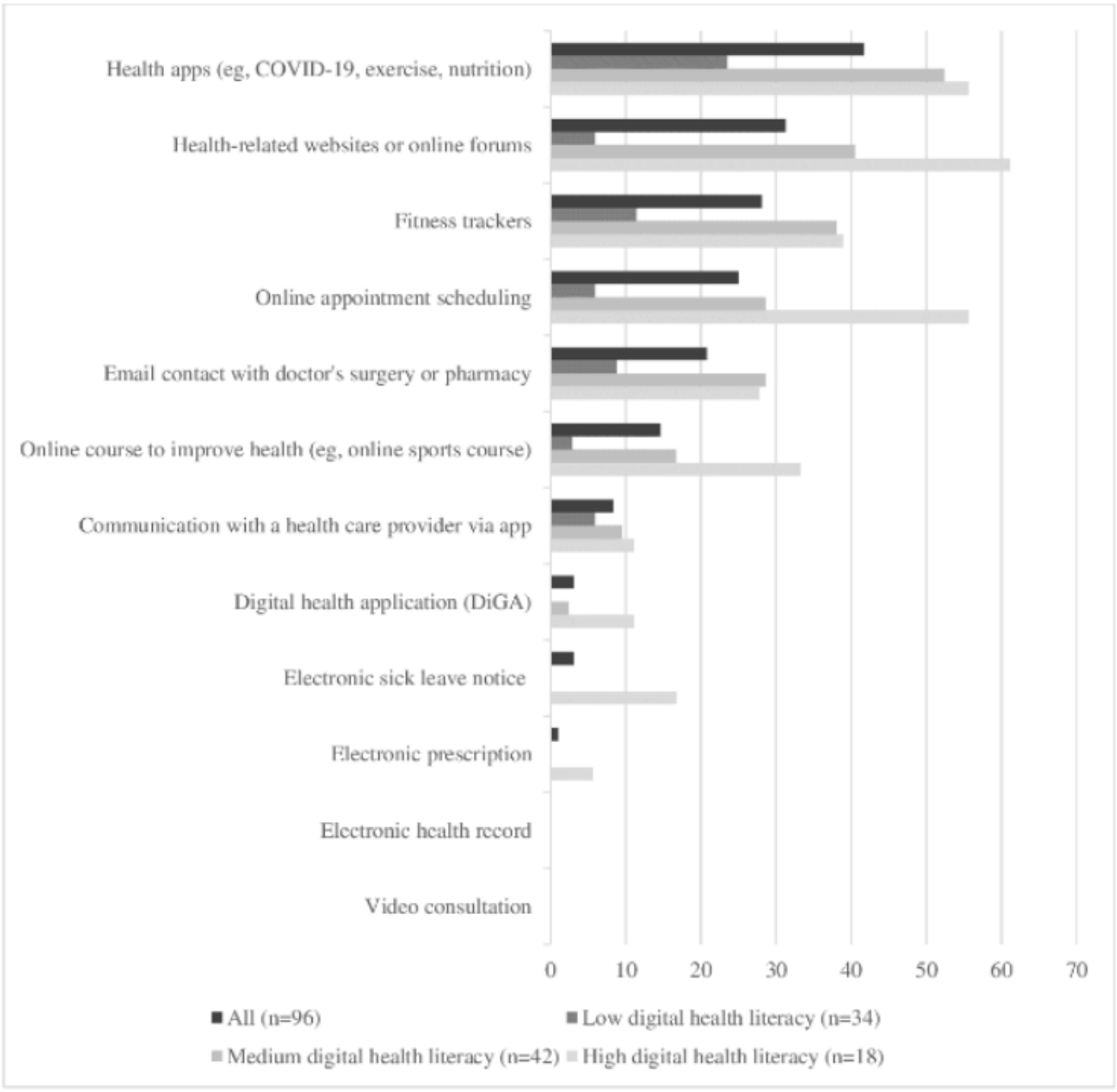
XSL•FO
RenderX

Table . Bivariate associations between digital health literacy and participant characteristics (linear regression models).

Participant characteristics		Unstandardized linear regression coefficient (95% CI)
Age		−0.05 (−0.21 to 0.10)
Sex		
	Male	Reference
	Female	−0.48 (−4.12 to 3.14)
Language skills		
	Native German speakers	Reference
	Good literacy in native language but not in German	1.59 (−6.69 to 3.51)
	Low literacy in both native language and German	−6.10 (−11.4 to −0.73)
Formal education		
	Low	Reference
	Medium	5.85 (2.30 to 9.40)
	High	1.77 (−2.28 to 5.81)
Use of digital health technologies		
	Nonuser	Reference
	User	9.94 (7.24 to 12.66)

XSL•FO
RenderX

Figure 1. Percentages of adults who used different digital health technologies for health-related purposes, by level of digital health literacy (multiple answers allowed).



XSL•FO
RenderX

Table . Use of digital devices in health context and reasons for use and non-use of digital health technologies.

Digital device use in health context and reasons		Value, n (%)
Use of digital devices in health context[a]		
	Computer or laptop	19 (20)
	Smartphone or tablet	74 (77)
	Activity tracker or smartwatch	19 (20)
	Gaming console	7 (7)
Reasons for use of digital health technologies[a]		
	It saves time	44 (46)
	I have access at any time and from anywhere	42 (44)
	It is cost-effective	38 (40)
	I want to stay healthy	39 (41)
	I am interested in digital health technologies	28 (29)
	I trust the information	25 (26)
	I otherwise have difficulties accessing healthcare services	14 (15)
	Other reasons	5 (5)
Reasons for nonuse of digital health technologies[a]		
	I prefer information on paper	33 (34)
	I prefer personal advice	41 (43)
	I have poor internet access	10 (10)
	I have technical problems	9 (9)
	I do not have a suitable device	6 (6)
	I do not know if my data are secure	20 (21)
	I do not trust the information	19 (20)
	I do not have interest in such offers	7 (7)
	Other reasons	20 (21)

[a]Multiple responses allowed.

Focus Groups

A total of 5 focus groups (total n=39 participants) were conducted between April and June 2023. Participants were on average 43 (SD 12.6) years old, 87% (34/39) participants were female, and 46% (18/39) attended a language course (Table 1).

Anchor examples for the reported difficulties and strategies to overcome difficulties can be found in Table 4. Due to technical difficulties in using digital devices (eg, computer) or the use of cookies on websites, searching for health information on the internet was difficult for participants. Most focus group participants reported language problems in comprehending technical and foreign words on health information websites and reading and writing German. The lack of translation options (eg, calling family members or friends with better German skills or unstable internet connection for using translation services) while filling out German-language forms in medical practices was also reported as a difficulty. In addition, participants reported difficulty in using search engines to find important and relevant information on the internet. Some participants mentioned that they were unsure on how to use or search on the internet because they were afraid of ordering something.

Most focus group participants stated that they seek advice from family members, friends, acquaintances, or local organizations when they encounter difficulties while searching for health information on the internet. Participants also tried to find a solution themselves by combining word and image searches or using translation tools. Some participants relied on their own experience (eg, if a child has fever) or on previous knowledge and thus did not consider searching for health-related information on the internet. In general, most focus group participants stated that they preferred personal or telephone contact with health care providers (eg, to arrange doctor's appointments or for acute medical concerns).

In the overarching discussion, focus group participants reported which digital devices, formats, languages, and digital health technologies they have used so far (Table 4). Participants used smartphones to search for information on the internet. Information was sought in German, in the native language, or in a combination of German and the native language.

XSL•FO
RenderX

Participants who could not or could barely read and write in their native language sought information in German, whereas participants who could read and write well in their native language sought information in their native language or a combination of German and their native language. Participants mainly used videos or a combination of different formats such as video and text. They also reported searching for images or using voice assistants.

XSL•FO
RenderX

Table . Examples of difficulties in using digital health technologies and strategies to overcome difficulties reported in focus group discussions.

Difficulties and strategies		Anchor examples
Difficulties		
	Use of digital device	• *If you can [work] with the computer, it's not difficult. […] I can with a smartphone. But with a laptop or computer I can't.* [ID02, task 1]
	Cookies	• *For me, cookie[s] were always [difficult] to accept. It was very, very difficult for me.* [ID03, task 3]
	Reading, writing and understanding German	• *For me, sometimes I have problem. Not all doctor or doctors can speak English, and my German [is] not so good. So sometimes when I look on the internet and sometimes my idea is: This doctor speaks English, or a little English? Or not? So sometimes is a bit difficult for me when I take a doctor from the internet.* [ID05, task 2] • *Internet not difficult, but sometimes we don't understand German. That's why [it's] a bit difficult for me. [...]* [ID04, task 2]
	Translation	• *And one problem is when first visiting, they give all the papers [to fill out]. And all is in German. And if someone can't understand everything, yes, once I [asked] the nurse, because I can't understand everything, my German is not good. But I can't help you, you have to do it on your own. And they don't have internet to translate. [...]* [ID05, task 2]
	Technical terms and foreign words	• *It wasn't difficult for me, but I don't understand the words, [for example] physiotherapy. The words. I don't know [if they] mean this or that, and I don't understand all of them. There's a lot of information, that's the problem.* [ID05, task 4]
	Searching and finding health information	• *It was a bit difficult this time. So you have to search and read at the same time. I also got a bit of help. [Interviewer: But what was more difficult?] So writing down this information. And to find it. [...]* [ID03, task 3]
	Uncertainty	• *So I can open it, I can see everything. But I'm afraid to order something. [...] That's why I don't dare to do that. [...]* [ID02, task 1]
Strategies used to overcome difficulties		

XSL•FO
RenderX

Difficulties and strategies		Anchor examples
	Seek support from family, friends, acquaintances, or local organizations	• *It's easy with me, but I'm a bit of a problem. I ask my colleague or my children, or. But now it's not difficult for me.* [ID02, task 2] • *[...] We can also enter, we have done a lot, because now Bremen is full of people from Ukraine and then you can enter directly into Google: Ukrainian-speaking doctors. Or we have another possibility and each of us, not just people from Ukraine, from Syria, from Persia, from Iran. We stick together, we share experiences. It's an organization called Bras. And when you go [there], you get in touch. Give your own details, then at a doctor's appointment, you get an interpreter.* [ID03, task 2]
	Try again	• *We'll try next day [laughs]. If I don't make it today, I'll try next day.* [ID02, task 2]
	Draw on own experiences or previous knowledge	• *In our home country our mother has a picture of fever, for fever goes down. That is a warm, a little warm, a little not warm water [Interviewer: Poultice?] Yes, every parent from Africa, for example Ghana, knows that if the child has a fever and you have no medication then you do that.* [ID05, task 3]
	Combine word and image search	• *Some words I don't understand well. I write on Google with German and then I see photos. Because I can't write in my own language. I write German and then photo comes up and I see photo and understand what it means.* [ID02, task 1]
	Use translation function	• *For me [it] is a bit difficult because I know many words in German - I don't understand it. So, I have to copy and then translate in Google Translator and sometimes the rest of the words from German it gives me wrong in English or different idea.* [ID05, task 4]
	Arrange a doctor's appointment by phone or in-person	• *When I'm looking for a doctor, I prefer to call, also because sometimes they want to have it that way. Then I make an appointment and at home I have a calendar of when I can go and then I do it.* [ID02, task 2]
	Personal contact to medical practice	• *You just go there and you have back pain or something. You don't have to call or google it first. It's much easier than going anywhere else.* [ID01, task 4]
Overarching discussion		

XSL•FO
RenderX

Difficulties and strategies		Anchor examples
	Use of digital device	• *It's faster and easier. Computers, some people don't know how to use computers, they can't cope, so it's easier with a smartphone.* [ID03, overarching discussion]
	Language	• *I do it in German. Because I speak Arabic and Kurdish, but I can't read and write because I didn't go to school in my country, only in Germany. I do everything in German.* [ID01, overarching discussion] • *And if there's something I don't understand, then I go to the dictionary or translation, so you can learn. If you always switch to your own language, you'll never learn German.* [ID03, overarching discussion] • *For health, I don't go to German. It's important and you have to understand everything. That's why I just go to English.* [ID05, overarching discussion]
	Format	• *Yes, video, because text is difficult for me a bit, because I don't read well and it's difficult to understand, but videos are easy and I always watch and have to understand (which teachings).* [ID04, overarching discussion] • *It depends on what you need […]. Sometimes text is enough, sometimes video, sometimes like this.* [ID05, overarching discussion] • *[...] If I don't feel like writing, I speak into the microphone. Then I do that, then it works well, yes.* [ID02, task 2]
	Use of digital health technologies	• *I usually go straight away. [Interviewer: And why directly to the doctor?] Yes, it's easier. I call for an appointment and then I go.* [ID03, overarching discussion] • *So I'll check the Internet first. If that gives me the right answer. If not, I sometimes call my mother and ask her. If she has any ideas. If that doesn't help, I go to the doctor.* [ID01, overarching discussion] • *So I just search the Internet when I need offers for health. It's much easier for me. [...] If I'm looking for a medicine, for example painkillers, ointment for back pain, they show it straight away and which pharmacies are close to me.* [ID03, overarching discussion]

Stakeholder Workshop

In total, 15 stakeholders (n=8 from civil and public services, n=1 from local public authorities, and n=6 from community social service providers or groups) and 8 members of the study team (n=3 researchers and n=5 health mediators) took part in the 150-minute stakeholder workshop at our community office (ie, Leibniz Living Lab) in October 2023. Discussions highlighted the following: First, to be able to use digital health technologies, adults with low reading and writing skills need to be curious and open to learning and trying out new things. They also need to be patient, have a high tolerance for frustration, and be willing to ask for support. Second, basic (German) language comprehension and interpretation of figures and symbols are as necessary as the financial resources to purchase digital devices. Third, health information websites should contain little, simple, and easy-to-read language, be available in multiple languages, barrier-free (eg, by integrating a read-aloud function), and compatible with smartphones. In addition, they should have easy navigation with few "clicks" to the searched content, no time limits, and use of images, symbols, and videos. Fourth, health programs should be named rather than using nondescriptive project acronyms. Fifth, web-based health services should refer to local contacts or organizations (eg, in the city district) and instead of a chat function, persons with low reading and writing skills may benefit

XSL•FO
RenderX

from a talking chatbot. Finally, stakeholders who should be involved are health professionals, health and educational institutions, physicians, health insurance companies, authorities, and stakeholders with ongoing health projects. To enable access to stakeholders not only digitally but also via analog channels, these should have fixed structures on site (eg, an office in the city district). Relevant stakeholders should be connected regularly, for instance, through the establishment of a working group. In addition, stakeholders should be educated and trained to pass on knowledge about web-based health services to adults with low reading and writing skills.

Discussion

Principal Results

This study assessed digital health literacy in adults with low reading and writing skills, explored which digital tools they use, and how digital health literacy can be improved from a stakeholder's perspective. Most participants had a low to medium digital health literacy. Digital health literacy was particularly low among those with low reading and writing skills both in their native language and in German and among those without a formal schooling degree. Participants reported using mainly information-related digital health technologies such as health apps, health websites, or activity trackers. Use of digital health technologies was associated with higher digital health literacy. Language problems and lack of technical skills made it difficult for participants to search for web-based health information, so support for health-related questions is mainly sought from family members, friends, or health care providers (eg, physicians).

Comparison With Previous Work

Adults with low reading and writing skills use digital health technologies, but less frequently compared with other population groups. For instance, compared with an online survey of adults insured at a statutory health insurance company in Germany [31], adults with low reading and writing skills reported lower usage of health websites (30/96, 31% vs 755/1678, 45%) and online appointment scheduling (24/96, 25% vs 1359/1678, 81%). In addition, similar to results reported by Schaeffer et al [7], use of digital health technologies in adults with low reading and writing skills is associated with higher levels of digital health literacy.

Compared with 2 cross-sectional telephone surveys of internet users living in Germany, the digital health literacy level was much lower in our study (mean eHEALS score of 22 points vs 31 points in survey 1 and 30 points in survey 2) [4,24]. This is also evident when comparing the eHEALS scores of the included studies in the systematic review by Estrela and colleagues [6] with our study. The included studies reported mean eHEALS scores ranging from 25 to 33 points. However, the authors stated that higher eHEALS scores were observed in studies targeting medical staff and that were conducted online or via phone. Furthermore, 2 more recently published systematic reviews and meta-analyses support these results [32,33]. In the reviews, mean eHEALS scores between 26 and 33 points [32] and between 20 and 29 points [33] were reported. Both reviews included studies with diverse population groups such as adolescents, university students, adults (ie, general adult population, army personnel, healthy adults, and adults with chronic diseases), or older adults. However, none of the included studies targeted adults with low reading and writing skills.

As several studies have shown, digital health literacy is unevenly distributed across populations with vulnerable groups such as older adults, persons with low levels of education and income reporting particularly low levels of digital health literacy [4,6-9,32]. More than one-third of our participants stated that they had no formal education. Low education and functional literacy appear to be cumulative resulting in a particularly low digital health literacy in adults with low reading and writing skills. Klinger and colleagues [13] come to similar conclusions; persons with low literacy skills and German proficiency have greater difficulties processing health information. This underlines the need for multilingual information in simple language and multimedia materials such as videos which was also supported by results from our focus group discussions and the stakeholder workshop. Best practices for researchers that should be considered when working with persons with limited literacy are, for instance, the use of graphics and visualizations, the consideration of cultural differences in the interpretation of graphics and phrases, the critical review of the wording of written materials, and the consideration of the digital (health) literacy of the target group [34]. However, even if reliable health information in various formats is available on the internet such as on the website for reliable and understandable information about health of the German Federal Ministry of Health [35], such information may not be accessible for persons with low and reading and writing skills due to lacking technical skills. Stakeholders responsible for community health projects and who are contacted for health purposes can improve (web-based) health information accessibility for adults with low reading and writing skills. However, to ensure that stakeholders can provide reliable information to adults with low reading and writing skills, they need sufficiently high levels of digital health literacy.

Our results are particularly important in the light of the ongoing shift of health and social services toward digital information exchange, contact provisions, and service inquiries. Given the lower digital competencies of persons with low reading and writing skills, more focus is needed on the risk of leaving less literate groups behind or reducing their options for interaction via their preferred, often nondigital pathways.

Strengths and Limitations

To our knowledge, this is the first study that assessed digital health literacy and digital technology use of adults with low reading and writing skills in Germany. The recruitment strategy via health mediators and adult education centers was successful and more participants than initially planned took part in the survey. However, this study has several limitations. First, selection bias occurred as we were not able to recruit a sufficient number of German native speakers. In addition, more females than males participated in the survey and focus group discussions. Second, some participants had difficulty answering the survey questions, in particular the eHEALS items. For the eHEALS items, differences between the individual items were hardly recognizable for participants, even if the interview was

XSL•FO
RenderX

conducted in their native language. Cutoff scores for the eHEALS sum score were built using the cutoffs proposed by De Santis and colleagues [24]. However, it should be noted that there are no uniformly used cutoff scores reported in the literature. Third, even though sufficient knowledge of German was required for participating in the focus groups, the discussions sometimes proceeded sluggishly due to language problems, and some participants may not have spoken as openly as they would have done if the focus group discussion had been done in their native language. Fourth, analysis of the qualitative data (ie, coding, grouping, and summarizing) was conducted by one researcher and checked by a second researcher. This ensures objectivity, but the gold standard would be to analyze the data or a part of the data by 2 researchers independently. However, this was not possible due to limited resources. Fifth, there are further methodological weaknesses (eg, no sample size calculation was conducted) inherent in the exploratory design of the study. This study aimed to gain first insights into the digital health technology use and digital health literacy of adults with low reading and writing skills in Germany. Based on the results, a methodologically more sophisticated study can be conducted in which, for instance, an intervention to promote digital health literacy in adults with low reading and writing skills is evaluated in a randomized controlled trial.

Future Directions

The results of this study indicate a need for studies that develop and evaluate interventions to promote digital health literacy for adults with low reading and writing skills. In total, 1 scoping review and 1 meta-analysis investigated interventions to enhance digital health literacy [8,33], but these reviews focused on study populations of children, adolescents, older adults, or adults with diabetes, cardiovascular diseases, or other health conditions, and not adults with low reading and writing skills. Most of the included interventions focused on educational training to improve health-related knowledge and skills and were delivered in-person, via the internet (eg, app, website, and online platform), or through a combination of in-person and web-based sessions [8,33]. Our findings also suggest that interventions to improve digital health literacy in adults with low reading and writing skills should focus on improving technical skills and how to search for web-based health information (eg, where and how to find reliable health information on the internet). In addition, practical sessions should be included similar to the tasks used in our focus group discussions. Based on learnings from our study, interventions to improve digital health literacy should be delivered using a combination of in-person and web-based modalities (eg, videos). Furthermore, our research underscores the importance of considering the low digital health literacy level of adults with low reading and writing skills. In addition, 2 previous literature reviews pointed out that digital technologies so far have not considered digital health literacy levels of the target group [36,37] and, in particular, persons with low reading skills who may have different reading and design requirements and needs [36].

An additional future research direction is to further develop the tasks used in the focus group discussions to create a more objective tool for assessing digital health literacy. There are several instruments for measuring digital health literacy available [38]. Existing instruments are based on self-reports, and some instruments, such as the widely used eHEALS scale, only relate to searching for health information on the internet (Web 1.0) and do not include interactions on social platforms (Web 2.0) [28,38]. For the focus group discussions, some of van der Vaart and Drossaert's performance-based items [28] were adapted to gain more in-depth knowledge on the challenges adults with low reading and writing skills face while searching for health-related information on the internet. The tasks gave valuable insights into the digital health literacy of our participants and can serve as a basis for further developing existing DHLI.

Conclusions

This mixed methods study showed that while adults with low reading and writing use digital health technologies, many have low to medium digital health literacy and find it challenging to search for health information on the internet due to lack of technical skills and language problems. Tailored interventions to promote digital health literacy are needed to ensure that adults with low reading and writing skills are not further left behind.

Acknowledgments

We gratefully acknowledge the support of Leibniz ScienceCampus Digital Public Health (LSC DiPH) Bremen [39], which is jointly funded by the Leibniz Association (W4/2018, W72/2022), the Federal State of Bremen, Germany, and the Leibniz Institute for Prevention Research and Epidemiology - BIPS, Bremen, Germany. LSC DiPH funded this study and did not influence its results. We would like to thank the health mediators (Patience Bonsu, Isil Dönmez, Elena Junck, Sabine Krliu, and Mariam Taraf) for recruiting participants, conducting the interviews, and supporting the focus group discussions and stakeholder workshop. We thank Finola Goepfert, BA, for entering the filled-out survey questionnaires into the software LimeSurvey. In addition, we would like to thank Meret Lakeberg, BA, for checking the coding, grouping, and summarizing of the focus group data. No generative AI was used for this manuscript.

Data Availability

All data generated and analyzed during the study are included in the manuscript and its supplementary materials. The data are not deposited in a public repository due to the sensitive nature of participant information.

XSL•FO
RenderX

Authors' Contributions

SM administered the project, conceptualized the study, developed the methodology, developed the questionnaire and focus group protocol, supervised and conducted the data collection, processed and analyzed the quantitative and qualitative data, visualized the results, wrote the first draft of the manuscript, and reviewed and edited the manuscript.
RW developed the methodology, developed the questionnaire and focus group protocol, conducted the data collection, and reviewed and edited the manuscript.
HZ cosupervised the project, acquired funding, and reviewed and edited the manuscript.
TB supervised and administered the project, acquired funding, conceptualized the study, developed the methodology, developed the questionnaire and focus group protocol, and reviewed and edited the manuscript.

No generative AI was used for this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
German version of survey questionnaire.
[PDF File, 384 KB - humanfactors_v12i1e65345_app1.pdf]

Multimedia Appendix 2
Informed consent form.
[PDF File, 382 KB - humanfactors_v12i1e65345_app2.pdf]

Multimedia Appendix 3
Percentages of adults who used different digital health technologies for health-related purposes, by language skills.
[PDF File, 58 KB - humanfactors_v12i1e65345_app3.pdf]

Checklist 1
STROBE (Strengthening the Reporting of Observational Studies in Epidemiology).
[DOCX File, 41 KB - humanfactors_v12i1e65345_app4.docx]

Checklist 2
COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist.
[DOCX File, 5653 KB - humanfactors_v12i1e65345_app5.docx]

References

1. Dratva J, Schaeffer D, Zeeb H. Digital health literacy in Germany: current status, concepts, and challenges. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2024 Mar;67(3):277-284. [doi: 10.1007/s00103-024-03841-5] [Medline: 38315221]
2. Stoumpos AI, Kitsios F, Talias MA. Digital transformation in healthcare: technology acceptance and its applications. Int J Environ Res Public Health 2023 Feb 15;20(4):3407. [doi: 10.3390/ijerph20043407] [Medline: 36834105]
3. Cline RJ, Haynes KM. Consumer health information seeking on the internet: the state of the art. Health Educ Res 2001 Dec;16(6):671-692. [doi: 10.1093/her/16.6.671] [Medline: 11780707]
4. De Santis KK, Jahnel T, Sina E, Wienert J, Zeeb H. Digitization and health in Germany: cross-sectional nationwide survey. JMIR Public Health Surveill 2021 Nov 22;7(11):e32951. [doi: 10.2196/32951] [Medline: 34813493]
5. Norman CD, Skinner HA. eHEALS: the eHealth literacy scale. J Med Internet Res 2006 Nov 14;8(4):e27. [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
6. Estrela M, Semedo G, Roque F, Ferreira PL, Herdeiro MT. Sociodemographic determinants of digital health literacy: a systematic review and meta-analysis. Int J Med Inform 2023 Sep;177:105124. [doi: 10.1016/j.ijmedinf.2023.105124] [Medline: 37329766]
7. Schaeffer D, Gille S, Berens EM, et al. Digital health literacy of the population in Germany: results of the HLS-GER 2. Gesundheitswesen 2023 Apr;85(4):323-331. [doi: 10.1055/a-1670-7636] [Medline: 34905785]
8. Arias López MDP, Ong BA, Borrat Frigola X, et al. Digital literacy as a new determinant of health: a scoping review. PLOS Digit Health 2023 Oct;2(10):e0000279. [doi: 10.1371/journal.pdig.0000279] [Medline: 37824584]
9. Milanti A, Chan DNS, Parut AA, So WKW. Determinants and outcomes of eHealth literacy in healthy adults: a systematic review. PLoS ONE 2023;18(10):e0291229. [doi: 10.1371/journal.pone.0291229] [Medline: 37792773]

XSL•FO
RenderX

10. Choi NG, Dinitto DM. The digital divide among low-income homebound older adults: internet use patterns, eHealth literacy, and attitudes toward computer/internet use. J Med Internet Res 2013 May 2;15(5):e93. [doi: 10.2196/jmir.2645] [Medline: 23639979]
11. Bergman L, Nilsson U, Dahlberg K, Jaensson M, Wångdahl J. Health literacy and e-health literacy among Arabic-speaking migrants in Sweden: a cross-sectional study. BMC Public Health 2021 Nov 25;21(1):2165. [doi: 10.1186/s12889-021-12187-5] [Medline: 34823499]
12. Knapp C, Madden V, Wang H, Sloyer P, Shenkman E. Internet use and eHealth literacy of low-income parents whose children have special health care needs. J Med Internet Res 2011 Sep 29;13(3):e75. [doi: 10.2196/jmir.1697] [Medline: 21960017]
13. Klinger J, Berens EM, Carol S, Schaeffer D. Health literacy of people with former Soviet Union and Turkish migration background in Germany. Gesundheitswesen 2023 Oct;85(10):887-894. [doi: 10.1055/a-2035-9107] [Medline: 37253365]
14. Grotlüschen A, Buddeberg K, Dutz G, Heilmann L, Stammer C. Hauptergebnisse Und Einordnung Zur LEO-Studie 2018 - Leben Mit Geringer Literalität: Springer; 2020:13-20.
15. König L, Kuhlmey A, Suhr R. Digital health literacy of the population in Germany and its association with physical health, mental health, life satisfaction, and health behaviors: nationally representative survey study. JMIR Public Health Surveill 2024 Feb 21;10:e48685. [doi: 10.2196/48685] [Medline: 38381497]
16. Dadaczynski K, Okan O, Messer M, et al. Digital health literacy and web-based information-seeking behaviors of university students in Germany during the COVID-19 pandemic: cross-sectional survey study. J Med Internet Res 2021 Jan 15;23(1):e24097. [doi: 10.2196/24097] [Medline: 33395396]
17. Dadaczynski K, Rathmann K, Schricker J, et al. Digital health literacy of pupils. Level and associations with physical activity and dietary behavior. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2022 Jul;65(7-8):784-794. [doi: 10.1007/s00103-022-03548-5] [Medline: 35652921]
18. Rangnow P, Fischer L, Hartmann A, et al. Digital health literacy among primary and secondary school teachers-a quantitative study. Front Public Health 2024;12:1334263. [doi: 10.3389/fpubh.2024.1334263] [Medline: 38912263]
19. Kirchberg J, Fritzmann J, Weitz J, Bork U. eHealth literacy of German physicians in the pre-COVID-19 era: questionnaire study. JMIR Mhealth Uhealth 2020 Oct 16;8(10):e20099. [doi: 10.2196/20099] [Medline: 33064102]
20. Cramer A, Keinki C, Saur F, Walter S, Hübner J. eHealth literacy, internet and eHealth service usage: a survey among a German municipality. J Public Health (Berl) 2025 Jan;33(1):237-248. [doi: 10.1007/s10389-023-01997-z]
21. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol 2008 Apr;61(4):344-349. [doi: 10.1016/j.jclinepi.2007.11.008] [Medline: 18313558]
22. Tong A, Sainsbury P, Craig J. Consolidated Criteria for Reporting Qualitative Research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
23. Ossner J, Rosebrock C, Scherf D. Alpha-Kurzdiagnostk. DVV-Rahmencurriculum Schreiben Und Lesen, 3rd edition: Deutscher Volkshochschul-Verband e.V; 2021.
24. De Santis KK, Muellmann S, Pan CC, et al. Digitisation and health: second nationwide survey of internet users in Germany. Digit Health 2024;10:20552076241301457. [doi: 10.1177/20552076241301457] [Medline: 39600383]
25. Muellmann S, De Santis KK, Pohlabeln H, Zeeb H. Use and acceptance of video consultation among adults insured by statutory health insurance providers in Germany: a nationwide online survey. Z Evid Fortbild Qual Gesundhwes 2025 Apr;193:36-44. [doi: 10.1016/j.zefq.2024.12.008] [Medline: 39848841]
26. Soellner R, Huber S, Reder M. The concept of eHealth literacy and its measurement. J Media Psychol 2014 Jan;26(1):29-38. [doi: 10.1027/1864-1105/a000104]
27. Robert Koch Institut. Fragebogen zur Studie Gesundheit in Deutschland aktuell: GEDA 2019/2020-EHIS [Article in German]. J Health Monit 2021;6(3):88-106. [doi: 10.25646/8560]
28. van der Vaart R, Drossaert C. Development of the Digital Health Literacy Instrument: measuring a broad spectrum of Health 1.0 and Health 2.0 skills. J Med Internet Res 2017 Jan 24;19(1):e27. [doi: 10.2196/jmir.6709] [Medline: 28119275]
29. Löhr K, Weinhardt M, Sieber S. The “World Café” as a participatory method for collecting qualitative data. Int J Qual Methods 2020 Jan 1;19:1-15. [doi: 10.1177/1609406920916976]
30. Mayring P. Qualitative inhaltsanalyse [Book in German]. In: Flick U, Kardoff E, Steinke I, editors. Qualitative Forschung: Ein Handbuch: Rowohlt; 2010:468-475. [doi: 10.1007/978-3-531-92052-8_42]
31. Muellmann S, Santis KK, Zeeb H. DigitHEALTHCARE – Eine Online-Befragung von Versicherten in Deutschland Zur Digitalisierung Des Gesundheitswesens Im Auftrag Der Hkk Krankenkasse Bremen [Book in German]: Leibniz-Institut für Präventionsforschung und Epidemiologie – BIPS; 2023.
32. Hua Z, Yuqing S, Qianwen L, Hong C. Factors influencing eHealth literacy worldwide: systematic review and meta-analysis. J Med Internet Res 2025 Mar 10;27:e50313. [doi: 10.2196/50313] [Medline: 40063939]
33. Barbati C, Maranesi E, Giammarchi C, et al. Effectiveness of eHealth literacy interventions: a systematic review and meta-analysis of experimental studies. BMC Public Health 2025 Jan 23;25(1):288. [doi: 10.1186/s12889-025-21354-x] [Medline: 39849354]

XSL•FO
RenderX

34. Reading Turchioe M, Mangal S. Health literacy, numeracy, graph literacy, and digital literacy: an overview of definitions, evaluation methods, and best practices. Eur J Cardiovasc Nurs 2024 May 28;23(4):423-428. [doi: 10.1093/eurjcn/zvad085] [Medline: 37590968]
35. Verlässliche informationen für ihre gesundheit [Web page in German]. German Federal Ministry of Health. URL: https://gesund.bund.de/ [accessed 2025-03-20]
36. Cheng C, Beauchamp A, Elsworth GR, Osborne RH. Applying the electronic health literacy lens: systematic review of electronic health interventions targeted at socially disadvantaged groups. J Med Internet Res 2020 Aug 13;22(8):e18476. [doi: 10.2196/18476] [Medline: 32788144]
37. Geukes C, Stark AL, Dockweiler C. eHealth literacy as a basis for developing digital technologies in health promotion and prevention? A systematic literature overview. Präv Gesundheitsf 2022;17:163-169. [doi: 10.1007/s11553-021-00858-5]
38. Lee J, Lee EH, Chae D. eHealth literacy instruments: systematic review of measurement properties. J Med Internet Res 2021 Nov 15;23(11):e30644. [doi: 10.2196/30644] [Medline: 34779781]
39. Leibniz ScienceCampus Digital Public Health. URL: https://www.digital-public-health.de/ [accessed 2025-05-08]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
DHLI: digital health literacy instrument
eHEALS: eHealth Literacy Scale
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Kushniruk; submitted 13.08.24; peer-reviewed by M Derksen, N Gordon; revised version received 20.03.25; accepted 27.03.25; published 22.05.25.

Please cite as:
Muellmann S, Wiersing R, Zeeb H, Brand T
Digital Health Literacy in Adults With Low Reading and Writing Skills Living in Germany: Mixed Methods Study
JMIR Hum Factors 2025;12:e65345
URL: https://humanfactors.jmir.org/2025/1/e65345
doi:10.2196/65345

© Saskia Muellmann, Rebekka Wiersing, Hajo Zeeb, Tilman Brand. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 22.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Chinese Version of the DigiHealthCom (Digital Health Competence) Instrument for Assessing Digital Health Competence of Health Care Professionals: Translation, Adaptation, and Validation Study

Lu Gao[1], BS; Meilian Chen[1], BS; Jingxin Wei[1], BS; Jinni Wang[1], MSN; Xiaoyan Liao[2,3], Prof Dr

[1]School of Nursing, Southern Medical University, Guangzhou, China

[2]Nanfang Hospital, Southern Medical University, Guangzhou, China

[3]School of Nursing, Guangzhou Medical University, Panyu Campus, No. 1 Xinzao Road, Panyu District, Guangzhou, China

Corresponding Author:
Xiaoyan Liao, Prof Dr
Nanfang Hospital, Southern Medical University, Guangzhou, China

Abstract

Background: Digital health competence is increasingly recognized as a core competence for health care professionals. A comprehensive evaluation of knowledge, skills, performance, values, and attitudes necessary to adapt to evolving digital health technologies is essential. DigiHealthCom (Digital Health Competence) is a well-established instrument designed to assess digital health competence across diverse health care professionals.

Objective: This study aimed to translate and culturally adapt DigiHealthCom into simplified Chinese (Mandarin) and verify its reliability and validity in assessing digital health competence of Chinese health care professionals.

Methods: DigiHealthCom was translated into Chinese following the guideline proposed by its original developers. The cultural adaptation involved expert review and cognitive interviewing. Internal consistency, test-retest reliability, content validity, convergent validity, discriminant validity, and factor structure were examined. Item analysis tested item discrimination, item correlation, and item homogeneity. Internal consistency was assessed using Cronbach α, and test-retest reliability was measured using the intraclass correlation coefficient. Content validity was assessed through both item and scale content validity indices. Convergent validity was measured by the Average Variance Extracted and Composite Reliability, while discriminant validity was measured by the heterotrait-monotrait ratio. A five-dimension model of DigiHealthCom was confirmed using confirmatory factor analysis.

Results: The finalized Chinese version of the DigiHealthCom was completed after addressing differences between the back-translations and the original version. No discrepancies affecting item clarity were reported during cognitive interviewing. The validation process involved 398 eligible health care professionals from 36 cities across 15 provinces in China, with 43 participants undergoing a retest after a 2-week interval. Critical ratio values (range 16.05 - 23.77, P<.001), item-total correlation coefficients (range 0.69 - 0.89), and Cronbach α if the item deleted (range 0.91 - 0.96) indicated satisfactory item discrimination, item correlation, and item homogeneity. Cronbach α for dimensions and the scale ranged from 0.94 to 0.98, indicating good internal consistency. The intraclass correlation coefficient was 0.90 (95% CI 0.81 - 0.95), indicating good test-retest reliability. Item content validity index ranged from 0.82 to 1.00, and the scale content validity index was 0.97, indicating satisfactory content validity. Convergent validity (average variance extracted: 0.60 - 0.79; composite reliability: 0.94 - 0.95) and divergent validity (heterotrait-monotrait ratio: 0.72 - 0.89) were satisfactory. Confirmatory factor analysis confirmed a well-fit five-dimension model (robust chi-square to df ratio=3.10, comparative fit index=0.91, Tucker-Lewis index=0.90, incremental fit index=0.91, root-mean-square error of approximation=0.07, standardized root-mean-square residual=0.05), with each item having a factor loading exceeding 0.40.

Conclusions: The Chinese version of DigiHealthCom has been proved to be reliable and valid. It is now available for assessing digital health competence among Chinese health care professionals. This assessment can be used to guide health care policy makers and educators in designing comprehensive and implementable educational programs and interventions.

(JMIR Hum Factors 2025;12:e65373) doi:10.2196/65373

KEYWORDS

competence; digital health; health care professionals; instrument; reliability; validity

XSL•FO
RenderX

Introduction

The World Health Organization (WHO) defines digital health as the field that involves the development and use of digital technologies to improve health outcomes [1]. In 2019, the WHO released the world's first guidelines for digital health interventions, outlining 10 ways to use digital health technologies to enhance health and primary services [2]. This concept extends beyond eHealth to include a wide array of smart devices and connected equipment, encompassing digital technologies such as the Internet of Things, big data, artificial intelligence, and robotics [2]. The release of these guidelines was followed by the Global Strategy for Digital Health (2020 - 2025), which emphasizes prioritized digital health strategies for global health care development [1]. China has incorporated measures for the development of a digitally healthy nation in the 14th Five-Year Plan [3]. With the global market projected to grow from US $211 billion in 2022 to US $809.2 billion by 2030 [4], digital health care is recognized as a rapidly expanding sector. Digital health has emerged as a significant trend in the evolution of global health care services. The digital transformation that the health care sector is currently undergoing is redefining the roles and responsibilities of health care professionals [5,6], creating an urgent need for digital health competence among them.

Given increasing prominence of digital health in global health care landscape, it is critical for health care professionals to possess sufficient digital health competence. Digital health competence is increasingly recognized as one of the core competencies for health care professionals, which would enable them to design and evaluate the impact of digital solutions on patient care and determine the best way to implement digital solutions in their work [7,8]. Although patients are becoming more accepting of and motivated to use digital health care services and tools, health care professionals face a digital skills shortage that impedes the adoption of digital solutions [9-11]. Inadequate digital health competence may lead to negative experiences and frustration with technology adoption among these professionals [12]. There is a strong association between health care professionals' digital health competence and their willingness to use such tools [13,14]. Studies have highlighted that the acceptance of digital health technologies by health care professionals significantly influences the adoption of digital solutions and emphasizes the critical role of digital health competence in ensuring patient safety [12,14]. Therefore, assessing digital health competence is essential to effectively providing digital health care solutions to the public.

Previous studies have attempted to explore digital health competence among health care professionals but have struggled to define it comprehensively due to the evolving nature of digital technologies [15]. Existing assessment tools primarily focus on informatics competence, digital health literacy, or skills related to the application of digital technologies. Examples include the Digital Health Literacy Instrument (DHLI, 2017) [16], the eHealth literacy questionnaire (eHLQ, 2018) [17], and the Nursing Digital Application Skill Scale (NDASS, 2024) [3]. The eHLQ is designed for eHealth user, especially individuals with low digital health literacy and those with chronic conditions [17]. The NDASS targets nurses' digital application skills in clinical settings [3]. Digital health literacy reflects users' knowledge and skills within their cultural, social, and institutional context [17]. Competence entails an integrative understanding of the knowledge, skills, performance, values, and attitudes essential for the effective execution of a given task [18]. Therefore, it is crucial to comprehensively evaluate the knowledge, skills, performance, and attitudes required for various health care professionals to adapt to the evolving digital health technologies.

Developed and validated among Finnish health care professionals in 2022, the DigiHealthCom (Digital Health Competence) instrument offers a more comprehensive scope than existing tools and is applicable to a wide range of health care professionals. In addition to assessing competence in using digital solutions and information and communication technology (ICT), it also explores previously unaddressed domains that reflect future requirements, such as acceptance of digital solutions, human-centered remote counseling, and ethical competence concerning digital solutions [19]. Furthermore, the instrument has been used to explore digital health competence profiles and associated factors in 817 health care professionals from 9 organizations in Finland [20]. It has been translated into 15 languages, and a large-scale international cross-sectional study on the digital health competence of health care professionals is currently in progress. Our team is part of this collaborative research effort. This study aimed to culturally adapt and validate the Chinese version of DigiHealthCom for Chinese health care professionals.

Methods

Study Design

A cross-sectional study was conducted.

Participants

Participants were recruited between May 2023 and April 2024 via convenience sampling. Recruitment posters with QR codes were disseminated on social networks specific to health care professionals. In addition, health care professionals attending local academic conferences were invited to participate. The inclusion criteria for participants were (1) employment within a health care organization and (2) consent to participate. The exclusion criteria were (1) individuals who were retired or had less than 1 year of work experience; (2) health care students; (3) individuals who completed the questionnaire in less than 150 seconds, as this indicated random clicking; or (4) individuals who displayed erratic response patterns.

According to the thumb rule, the sample size should be 5 to 10 times the number of items. With an anticipated 10% rate of invalid responses, a minimum of 231 samples was necessary. Finally, a total of 398 cases were included in the study. To assess the test-retest reliability of the instrument, 43 participants who provided contact information completed the questionnaire again after a 2-week interval.

XSL•FO
RenderX

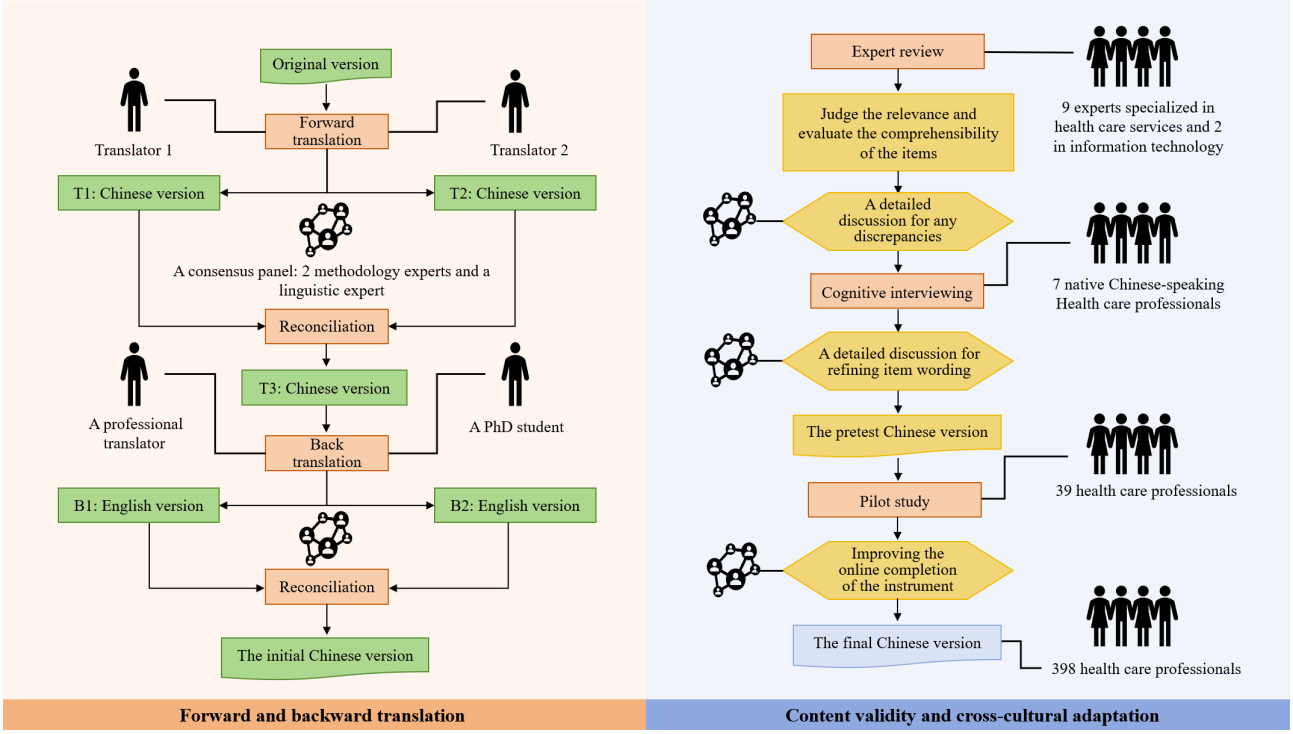
Instrument

The DigiHealthCom instrument comprises 5 domains, comprising a total of 42 items—competence in human-centered remote counseling (16 items), digital solutions as part of work (9 items), competence in ICT (5 items), competence in using and evaluating digital solutions (8 items), and ethical competence related to digital solutions (4 items). Each item was rated on a 4-point Likert scale (1=completely disagree, 2=partially disagree, 3=partially agree, and 4=completely agree). For each domain, a mean value of ≤2.49 indicated low competence, 2.50 - 3.49 indicated intermediate competence, and ≥3.50 indicated high competence [20]. DigiHealthCom has been validated among health care professionals across tertiary, primary, and private health care settings (n=817), demonstrating satisfactory internal consistency (Cronbach α=0.80 - 0.97) and content validity (item content validity index [I-CVI]: 0.77 - 1.00; S-CVI/Ave: 0.94) [19].

Translation and Cross-Cultural Adaptation

To ensure a high-quality Chinese translation of DigiHealthCom, a rigorous translation and cross-cultural adaptation process was followed [21]. Figure 1 illustrates the translation and cross-cultural adaptation process.

Figure 1. The translation and cross-cultural adaptation process of developing the Chinese version of the DigiHealthCom (Digital Health Competence) instrument.



Forward and Backward Translation

In total, 2 bilingual nursing professionals, who are native Chinese speakers (a nursing expert and a nursing graduate student), independently translated the English version of DigiHealthCom into Chinese, resulting in 2 forward translations (T1 and T2). A consensus panel reviewed these translations for conceptual equivalence, clarity, and comprehensibility, producing a reconciled version (T3). Subsequently, the T3 version was back-translated into English by a professional translator and a PhD student unfamiliar with the original version, both knowledgeable about Chinese and English-speaking cultures. The consensus panel reviewed and resolved discrepancies between the back-translations and the original version, finalizing the initial Chinese version. The consensus panel comprised 2 nursing researchers experienced in scale development and a linguistic expert.

Expert Review

To assess the content validity of the Chinese version of DigiHealthCom, 11 experts were invited to evaluate the relevance of its dimensions and items using a 4-point ordinal scale (1 [not relevant], 2 [weakly relevant], 3 [strongly relevant], and 4 [very relevant]). The expert panel consisted of 9 health care specialists and 2 IT experts. In addition, expert evaluated the comprehensibility of the items. The consensus panel, initially involved in the translation process, conducted a detailed discussion to resolve any potential discrepancies.

Cognitive Interviewing

Cognitive interviewing was used to evaluate the clarity and cultural suitability of the initial Chinese version. A total of 7 native Chinese speakers, including 2 doctors, 4 nurses, and 1 IT technician, were recruited. Participants were briefed on the study's objectives and methods before interviews, and their consent was obtained.

The first author, trained in qualitative research methods, conducted the interviews in a meeting room. Participants completed the Chinese version of DigiHealthCom independently and engaged in cognitive interviews. Interviewers evaluated whether participants found the items relevant to their condition and if they encountered any understanding difficulties. Field notes were taken. Modifications were deemed necessary if at

XSL•FO
RenderX

least 1 participant (1) found an item difficult to understand, (2) demonstrated mostly or completely inaccurate comprehension of an item, or (3) provided feedback indicating the need for improvements, especially regarding cultural relevance. The consensus panel determined whether to retain or alter item wording following expert review and cognitive interviewing. Discrepancies affecting item clarity were resolved to create the pretest version of the Chinese DigiHealthCom version. Specific item modifications were detailed in Table S1 in Multimedia Appendix 1.

Pilot Study

A pilot study was conducted with 39 health care professionals from a general hospital in Guangzhou, China. The pilot study confirmed that no further modifications were required. The final Chinese version of DigiHealthCom consists of 5 dimensions and 42 items.

Data Collection

Data were collected using the Wen Juan Xing (a Chinese web questionnaire platform) [22]. Respondents accessed the questionnaire by scanning a QR code or clicking a link, with 1 response allowed per IP address to prevent duplicates. The questionnaire platform performed completeness checks before submission. The Chinese version of DigiHealthCom was presented alongside a sociodemographic questionnaire. The questionnaire consisted of 2 pages and 73 items. Sex, age, education, service area, professional license, work experience, type of organization, and frequency of patient work were collected. Participants who voluntarily chose to provide contact information participated in a retest after a 2-week interval. An introduction outlined the study's purpose and provided questionnaire instructions.

Statistical Analysis

Item analysis was used to evaluate item discrimination, item correlation, and item homogeneity. For item discrimination, respondents were classified into high-score (top 27%) and low-score (bottom 27%) groups based on their total scores. An independent *t* test determined whether each item could significantly distinguish between these groups. Items with a critical ratio ($|t|$) less than 3.0 were considered for exclusion [23]. In addition, item homogeneity and item correlation were tested using Cronbach α if the item was deleted and corrected item-total correlation coefficient. Items with a corrected item-total correlation coefficient below 0.40 were considered for exclusion [23]. Cronbach α and Intraclass correlation coefficient (ICC) were used to assess the internal consistency and test-retest reliability of the instrument. A Cronbach α≥0.70 indicated good internal consistency, while ICC>0.70 indicated good time stability [23]. The I-CVI and the Scale CVI/Average (S-CVI/Ave) were used to evaluate content validity of the instrument. I-CVI≥0.78 and S-CVI/Ave≥0.90 indicate satisfactory content validity [24].

Construct validity was evaluated using confirmatory factor analysis (CFA) with a robust chi-square to df ratio (χ^2/df) less than 3, Tucker-Lewis Index (TLI), Incremental Fit Index (IFI), and Comparative Fit Index (CFI) greater than 0.90, and root-mean-square error of approximation (RMSEA) and standardized root-mean-square residual (SRMR) less than 0.08, indicating an acceptable data-model fit [25]. The average variance extracted (AVE) and composite reliability (CR) were used to assess convergent validity, with AVE greater than 0.50 and CR greater than 0.70 indicating good convergent validity [26]. Discriminant validity was performed using the heterotrait-monotrait ratio (HTMT), with a correlation matrix value <0.90 considered good [26].

Furthermore, participants were divided into 2 groups based on geographic location for the sensitivity analysis. The DigiHealthCom scores were compared to evaluate potential selection bias. The absence of a significant difference in digital health competence between the groups indicates that selection bias is unlikely in the study sample. Statistical analysis was conducted using IBM SPSS (version 27.0), IBM AMOS (version 29.0), and Smart PLS (Version 4.1.0.0; GmbH Corp). A *P* value of less than .05 was considered statistically significant.

Ethical Considerations

Ethical approval was obtained from the Ethical Committee at Nanfang Hospital, Southern Medical University, China (NFEC-2023 - 165). This research adhered to the principals of the Declaration of Helsinki. All participants provided informed consent and voluntarily completed the web questionnaire. Participants had the option to skip questions, review, and delete their responses. Participants had the right to withdraw from the survey at any time. Those who completed it received a random monetary reward ranging from CNY 2 to 5 (US $0.28 to 0.69). Participants' rights and researcher's contact information were provided on the first page of the web survey. Minimal sociodemographic was collected to maintain ethical standards. Participant information was kept confidential and anonymous. The CHERRIES (Checklist for Reporting Results of Internet E-Surveys; Checklist 1) was used to enhance the transparency of the study [27].

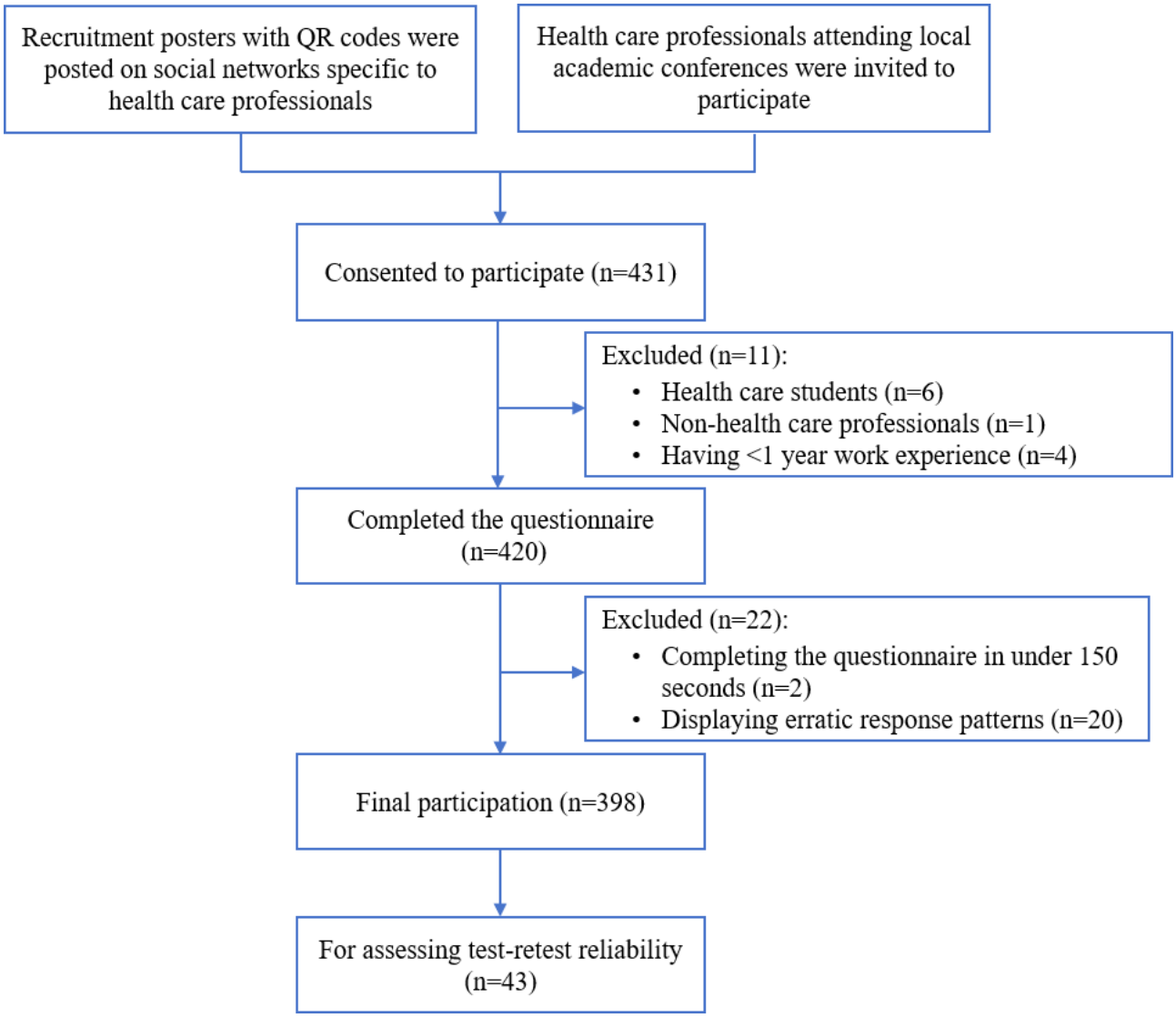
Results

Characteristics of the Participants

Initially, 431 participants were recruited for this study. Furthermore, 33 participants were subsequently excluded due to not meeting the inclusion criteria (n=7), having less than 1 year work experience (n=4), completing the questionnaire in under 150 seconds (n=2), and exhibiting erratic response patterns (n=20). Finally, 398 eligible health care professionals were included. Figure 2 illustrates the participant enrollment flowchart. The participants enrolled from 36 cities across 15 provinces, as identified by IP addresses. Among the participants, 249 (62.6%) were recruited from Guangzhou, a megacity in Guangdong province that contains 6125 registered medical facilities. As shown in Table 1, 364 (91.5%) participants were female, 386 (97%) worked in health care services, and 357 (89.7%) were nurses. The participants had an average of 13.7 (SD 9.4) years of work experience, with 281 (70.6%) working directly with patients for at least 5 days per week. Participant characteristics are summarized in Table 1.

XSL•FO
RenderX

Figure 2. The flowchart for the enrollment process of health care professionals.



XSL•FO
RenderX

Table . Demographic characteristics of the participants (N=398).

Variables		Value
Sex, n (%)		
	Female	364 (91.5)
	Male	34 (8.5)
Age (years), mean (SD)		36.1 (8.8)
Education, n (%)		
	Junior vocational qualification	58 (14.6)
	Bachelor's degree	273 (68.6)
	Master's degree	53 (13.3)
	Doctoral degree	14 (3.5)
Service area, n (%)		
	Health care service	386 (97)
	Social service	4 (1)
	Rehabilitation service	7 (1.8)
	Others	1 (0.3)
Location, n (%)		
	Southern China	351 (88.2)
	Northern and Western China	47 (11.8)
Type of organization, n (%)		
	Tertiary hospital	248 (62.3)
	Secondary hospital	36 (9)
	Community health care center	76 (19.1)
Professional license, n (%)		
	Nurse	357 (89.7)
	Doctor	24 (6)
	Midwife	10 (2.5)
	Others[a]	7 (1.8)
Working experience (years), mean (SD)		13.7 (9.4)
Patient work, n (%)		
	Daily (at least 5 days a week)	281 (70.6)
	Weekly (1-4 days per week)	62 (15.6)
	Monthly (a few times a month)	17 (4.3)
	Rarely (a few times in several months)	23 (5.8)
	I do not currently work with patient	15 (3.8)
Full-time, n (%)		398 (100)

[a]Including physiotherapist, paramedical technician, and pharmacist.

Results of Item Analysis

Item analysis showed a significant difference between high-score and low-score groups. The critical ratio values for all items were above 3.0 (range 16.05 - 23.77, P<.001; Table 2), indicating excellent item discrimination. All corrected item-total correlation coefficients exceeded 0.4 (Table 2). The Cronbach α if the item deleted (range 0.91 - 0.96) were acceptable (Table 2).

XSL•FO
RenderX

Table . Item analysis and content validity of the 42 items in the DigiHealthCom (Digital Health Competence; N=398).

Item	Critical ratio	CITC[a]	CID[b]	I-CVI[c]
RC1[d]	16.05	0.69	0.96	1
RC2	17.27	0.70	0.96	1
RC3	17.29	0.71	0.96	1
RC4	19.43	0.77	0.96	1
RC5	17.10	0.77	0.96	1
RC6	20.34	0.79	0.96	1
RC7	19.90	0.80	0.96	1
RC8	22.99	0.79	0.96	0.82
RC9	19.33	0.76	0.96	1
RC10	21.91	0.81	0.96	1
RC11	19.26	0.76	0.96	1
RC12	20.84	0.80	0.96	0.91
RC13	19.89	0.80	0.96	1
RC14	20.80	0.74	0.96	1
RC15	19.02	0.73	0.96	0.91
RC16	18.71	0.75	0.96	0.91
DS1[e]	21.53	0.77	0.96	1
DS2	20.63	0.81	0.95	1
DS3	19.82	0.84	0.95	1
DS4	21.53	0.86	0.95	1
DS5	22.19	0.89	0.95	0.91
DS6	22.00	0.86	0.95	1
DS7	22.57	0.87	0.95	0.91
DS8	20.62	0.71	0.96	0.91
DS9	23.77	0.81	0.95	0.91
ICT1[f]	20.62	0.84	0.92	1
ICT2	18.64	0.88	0.91	1
ICT3	22.03	0.88	0.91	1
ICT4	18.67	0.81	0.93	1
ICT5	18.88	0.75	0.94	0.82
UE1[g]	20.48	0.82	0.95	1
UE2	21.89	0.82	0.95	1
UE3	23.04	0.84	0.95	1
UE4	22.25	0.87	0.95	0.91
UE5	21.33	0.83	0.95	1
UE6	18.19	0.78	0.95	0.91
UE7	23.55	0.87	0.95	0.91
UE8	20.52	0.82	0.95	1
EC1[h]	21.40	0.84	0.92	1
EC2	20.79	0.81	0.93	0.91
EC3	20.48	0.89	0.91	1

XSL•FO
RenderX

Item	Critical ratio	CITC[a]	CID[b]	I-CVI[c]
EC4	21.44	0.86	0.91	1

[a]CITC: corrected item-total correlation.

[b]CID: Cronbach α if item deleted.

[c]I-CVI: item content validity index.

[d]RC: human-centered remote counseling competence.

[e]DS: digital solutions as part of work.

[f]ICT: information and communication technology competence.

[g]UE: competence in using and evaluating digital solutions.

[h]EC: ethical competence related to digital solution.

Internal Consistency and Test-Retest Reliability

As shown in Table 3, internal consistency (Cronbach α=0.94 - 0.98) and test-retest reliability were good (ICC=0.90; 95% CI 0.81 - 0.95).

Table . Internal consistency and test-retest reliability of the Chinese version of the DigiHealthCom (Digital Health Competence).

Dimension	Cronbach α (N=398)	ICC[a] (N=43)	95% CI of ICC
RC[b]	0.96	0.80	0.63-0.89
DS[c]	0.96	0.90	0.81-0.95
ICT[d]	0.94	0.79	0.62-0.89
UE[e]	0.96	0.80	0.63-0.89
EC[f]	0.94	0.81	0.65-0.90
Total	0.98	0.90	0.81-0.95

[a]ICC: intraclass correlation coefficient.

[b]RC: human-centered remote counseling competence.

[c]DS: digital solutions as part of work.

[d]ICT: information and communication technology competence.

[e]UE: competence in using and evaluating digital solutions.

[f]EC: ethical competence related to digital solutions.

Content Validity

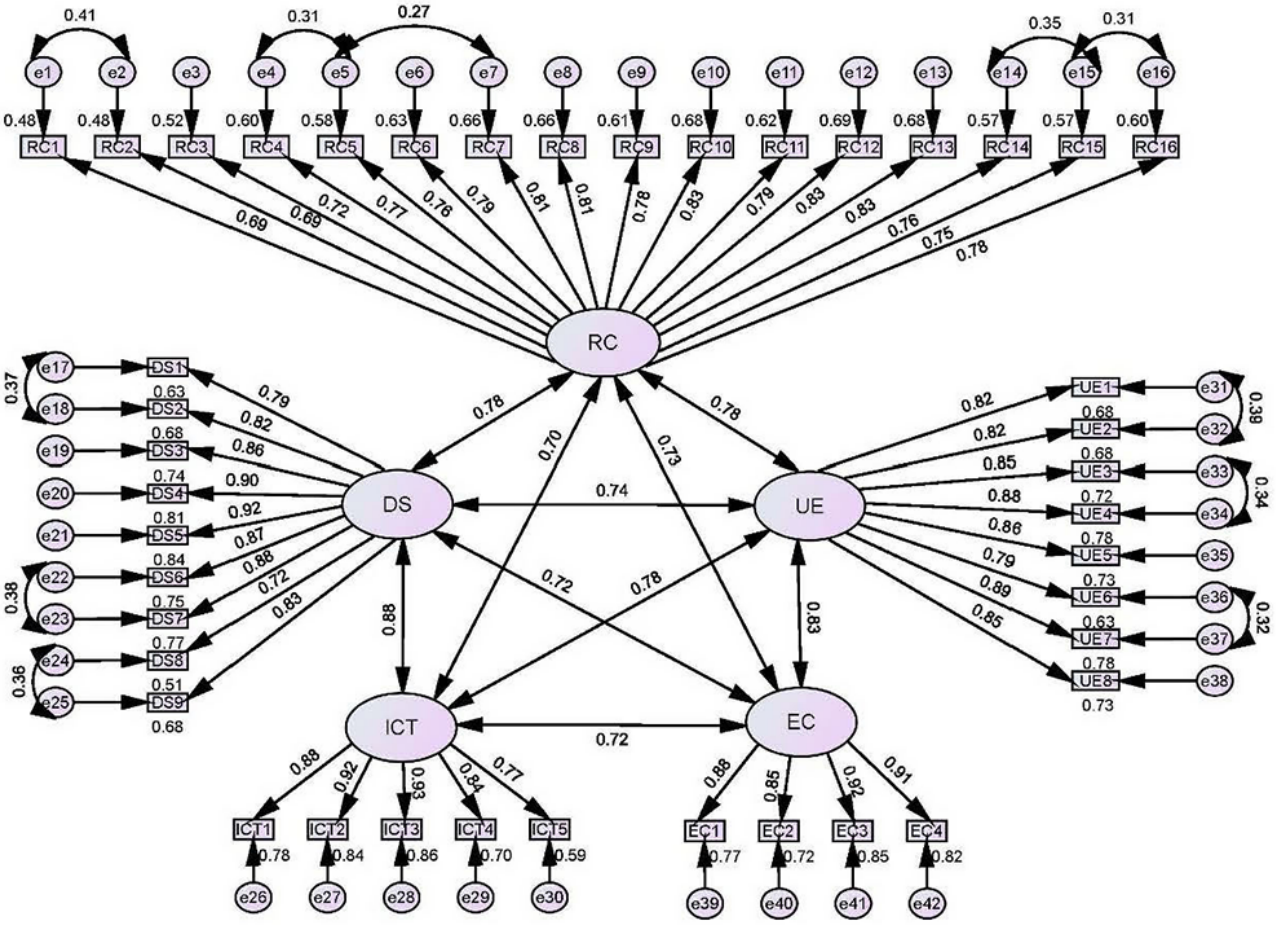
To test content validity, 11 experts evaluated the relevance of the items. The S-CVI/Ave for the instrument was 0.97, and the I-CVI ranged from 0.82 to 1.00 (Table 2), indicating satisfactory content validity.

Construct Validity

As illustrated in Figure 3, the results of CFA supported the 5-factors structure (digital solutions as part of work; ethical competence related to digital solutions; ICT competence; human-centered remote counselling competence; competence in using and evaluating digital solutions). The indices, including χ^2/df (3.10), CFI (0.91), TLI (0.90), IFI (0.91), RMSEA (0.07), and SRMR (0.05), indicated an acceptable model fit. The factor loadings of items ranged from 0.69 to 0.93 (Figure 3). The AVE values for each dimension exceeded 0.50 (range: 0.60 - 0.79), with the CR values >0.70 (range: 0.94 - 0.95), indicating excellent convergent validity. The HTMT values within the matrix were less than 0.90 (range: 0.72 - 0.89), indicating good divergent validity. These results collectively demonstrated satisfactory construct validity of the instrument.

XSL•FO
RenderX

Figure 3. The confirmatory factor model of the Chinese version of the DigiHealthCom (Digital Health Competence; N=398). DS: digital solutions as part of work; EC: ethical competence related to digital solutions; ICT: information and communication technology competence; RC: human-centered remote counselling competence; UE: competence in using and evaluating digital solutions.



Sensitivity Analysis

Participants were divided into 2 groups based on geographic location: Southern China and Northern and Western China. The DigiHealthCom scores were compared to evaluate potential selection bias. As illustrated in Table S3 in Multimedia Appendix 2, no significant difference was found in DigiHealthCom scores between the 2 groups (P>.05), indicating that selection bias is unlikely in the study sample.

Discussion

Principal Findings

In this study, we translated and culturally adapted the DigiHealthCom instrument into Chinese and assessed its reliability and validity among Chinese health care professionals. The Chinese version of DigiHealthCom demonstrated satisfactory internal consistency, test-retest reliability, content validity, and construct validity. To our knowledge, this study represents the first effort to validate a comprehensive Chinese tool for measuring digital health competence among health care professionals.

In this study, each item of the Chinese version of DigiHealthCom was translated and back-translated strictly following the dual direct-to-back translation model [21] to ensure alignment in semantic, conceptual, and content with the original English version. And then, we conducted expert review involving 11 experts in nursing and IT to examine content validity of the Chinese version. High content validity indices imply that the instrument provides a broad enough range of content to allow conclusions about the targeted construct. We also conducted cognitive interviewing and a pilot study to ensure its clarity and comprehensibility [28]. Issues related to semantic validation and understanding identified in the initial Chinese version were addressed after receiving feedback from the expert review and cognitive interviewing, leading to improvements such as clarifying complex terms and vague definitions. Item analysis revealed good differentiation and high correlations between the items in this study. Furthermore, this instrument underwent rigorous testing for internal consistency, test-retest reliability, and construct validity, exhibiting excellent internal consistency, time stability, and construct validity, resonating with the original study [19]. The types of reliability and validity assessed in this study represent essential psychometric properties for a measurement instrument.

Discrepancies between existing instruments largely stem from differing conceptual frameworks. The DHLI measures 7 individual skills: operation, navigation, information searching, evaluating reliability, determining relevance, adding self-generated content, and protecting privacy [16]. The eHLQ comprises 7 dimensions, addressing users' attributes, their interaction with technologies, and their experience with systems

XSL•FO
RenderX

[17]. Our CFA findings confirm a 5-factor structure, consistent with the exploratory factor analysis findings of the initial study [19]. Beyond competence in using and evaluating digital solutions and ICTs, the DigiHealthCom instrument addresses additional essential domains that were previously unaddressed, that is acceptance of digital solutions, human-centered remote counseling, digital interaction skills with patients and interprofessional teams, and ethical competence related to digital solutions [19]. Notably, acceptance of digital solutions significantly influences health care professionals' adoption of digital solutions [13,14]. Competence in person-centered remote consultations is crucial for fostering patient engagement and improving accessibility to equitable digital health services [8]. Furthermore, proficient digital interaction skills facilitate effective coordination in digital settings, thereby supporting decision-making and optimizing treatment plans [8]. As digital solutions become more prevalent, attention to data privacy and information security has intensified [29,30]. Ethical competence related to digital solutions ensures appropriate management of patient information, fostering trust in digital health technologies and enhancing the safety of digital health services [20,31]. These domains provide a unified theoretical framework for comprehensively measuring digital health competencies among various professionals [8,9,19], which are essential for delivering high-quality digital health services to meet future demands.

The Finnish health care system is renowned for its universality, robust public funding, and centralized digital infrastructure [32]. Regional variations exist in Finland; for example, northern districts have fewer professionals specializing in remote counseling and the integration of digital solutions compared with southern areas [20,33]. Strategies have been implemented to enhance health care professionals' digital health competence and improve education related to health care digitalization, including technology to support client engagement, digital services integrated into nursing work, and considerations of safety and ethics in the digital environment [20,34]. In contrast, China experiences regional disparities in digital health infrastructure and access to digital health services, especially in rural areas, despite the rapid growth of digital health services [35]. In addition, system compatibility across hospitals remains an unsolved issue [35]. These factors may influence the development of digital health competence among health care professionals. A cross-sectional study conducted in a central Chinese province reported that 49.9% (1690/3386) of clinical nurses demonstrate low telehealth readiness [36]. Consequently, evaluating, addressing, and reducing regional disparities in health care professionals' digital health competence is vital for promoting equity in health care service provision. Validating a comprehensive digital health competence measurement instrument for health care professionals is imperative to address challenges posed by diverse health care environments, such as China's. The Chinese version of DigiHealthCom is now ready for application in a wide range of settings and various health care professionals. Researchers, educators, health care providers, and policy makers can use it to evaluate digital health competence among diverse health care professionals, and develop digital health training curricula and policies based on the assessment.

Limitations

This study acknowledges several limitations. First, the primary participants were nurses. Future research should involve a broader spectrum of health care professionals. Second, although recruited from diverse health care institutions, the participants mainly came from southern China, which may introduce selection bias due to regional disparities in digital health infrastructure. To address this, participants were divided into 2 groups (Southern China versus Northern and Western China) for sensitivity analysis, and their DigiHealthCom scores were compared. We did not find significant differences in digital health competence between the 2 groups, indicating that selection bias is unlikely in the study sample. Future studies could benefit from employing a stratified sampling method. Third, the absence of a standardized method for assessing digital health competence among health care professionals hindered the evaluation of criterion validity. Assessments of digital health competence predominantly rely on subjective self-reports, which might be susceptible to response bias. However, relying exclusively on performance-based assessments or other objective measures to evaluate competence also presents challenges. Therefore, using a more holistic combination of methods could offer a viable solution. Future research should consider larger sample sizes and multicenter external evaluations to address these limitations.

Conclusions

In conclusion, we have successfully translated, culturally adapted, and validated DigiHealthCom for Chinese health care professionals. Our findings demonstrate that the Chinese version of DigiHealthCom is a reliable and valid instrument for assessing digital health competence among these health care professionals.

Acknowledgments

The authors would like to acknowledge all the participants who participated in this study; in particular, Dr Donglan Ling for help with the translation process. XL is supported by the National Natural Science Foundation of China (grant 82171202). The funding agents had no role in the study design; collection, analysis, and interpretation of the data; writing of the manuscript; or decision to submit the paper for publication.

Data Availability

The dataset collected in this study is available from the corresponding author on reasonable request.

XSL•FO
RenderX

Authors' Contributions

XL and LG had full access to all the data in this study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Concept and design were contributed by XL. Acquisition, analysis, or interpretation of data were contributed by LG, MC, and JW. Drafting of the manuscript was handled LG and XL. Critical revision of the manuscript for important intellectual content were contributed by XL, LG, JW, MC, and JW. Administrative, technical, or material support were contributed by JW. The authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Modified items of the Chinese version of the DigiHealthCom (Digital Health Competence) instrument after cultural adaption.
[PDF File, 132 KB - humanfactors_v12i1e65373_app1.pdf]

Multimedia Appendix 2
Comparison of DigiHealthCom (Digital Health Competence) scores between health care professionals in the Southern China group and in the Northern and Western China group.
[DOCX File, 16 KB - humanfactors_v12i1e65373_app2.docx]

Checklist 1
Checklist for Reporting Results of Internet E-Surveys (CHERRIES).
[XLSX File, 14 KB - humanfactors_v12i1e65373_app3.xlsx]

References

1. World Health Organization. Global strategy on digital health 2020-2025. 2021. URL: https://www.who.int/publications/i/item/9789240020924 [accessed 2024-08-04]
2. World Health Organization. WHO guideline recommendations on digital interventions for health system strengthening. 2019. URL: https://iris.who.int/handle/10665/311980 [accessed 2024-08-04]
3. Qin S, Zhang J, Sun X, et al. A scale for measuring nursing digital application skills: a development and psychometric testing study. BMC Nurs 2024 May 31;23(1):366. [doi: 10.1186/s12912-024-02030-8] [Medline: 38822276]
4. Grand View Research. Digital health market size, share & trends analysis report by technology (healthcare analytics, mhealth), by component (hardware, software, services), by application, by end-use, by region, and segment forecasts, 2024 - 2030. URL: https://www.grandviewresearch.com/industry-analysis/digital-health-market [accessed 2025-02-14]
5. Nazeha N, Pavagadhi D, Kyaw BM, Car J, Jimenez G, Tudor Car L. A digitally competent health workforce: scoping review of educational frameworks. J Med Internet Res 2020 Nov 5;22(11):e22706. [doi: 10.2196/22706] [Medline: 33151152]
6. Odendaal WA, Anstey Watkins J, Leon N, et al. Health workers' perceptions and experiences of using mHealth technologies to deliver primary healthcare services: a qualitative evidence synthesis. Cochrane Database Syst Rev 2020 Mar 26;3(3):CD011942. [doi: 10.1002/14651858.CD011942.pub2] [Medline: 32216074]
7. Bichel-Findlay J, Koch S, Mantas J, et al. Recommendations of the International Medical Informatics Association (IMIA) on Education in Biomedical and Health Informatics: Second Revision. Int J Med Inform 2023 Feb;170(104908):104908. [doi: 10.1016/j.ijmedinf.2022.104908] [Medline: 36502741]
8. Jarva E, Oikarinen A, Andersson J, et al. Healthcare professionals' perceptions of digital health competence: a qualitative descriptive study. Nurs Open 2022 Mar;9(2):1379-1393. [doi: 10.1002/nop2.1184] [Medline: 35094493]
9. Konttila J, Siira H, Kyngäs H, et al. Healthcare professionals' competence in digitalisation: a systematic review. J Clin Nurs 2019 Mar;28(5-6):745-761. [doi: 10.1111/jocn.14710] [Medline: 30376199]
10. Brown J, Pope N, Bosco AM, Mason J, Morgan A. Issues affecting nurses' capability to use digital technology at work: an integrative review. J Clin Nurs 2020 Aug;29(15-16):2801-2819. [doi: 10.1111/jocn.15321] [Medline: 32416029]
11. Kruk ME, Gage AD, Arsenault C, et al. High-quality health systems in the Sustainable Development Goals era: time for a revolution. Lancet Glob Health 2018 Nov;6(11):e1196-e1252. [doi: 10.1016/S2214-109X(18)30386-3] [Medline: 30196093]
12. Salahuddin L, Ismail Z. Classification of antecedents towards safety use of health information technology: a systematic review. Int J Med Inform 2015 Nov;84(11):877-891. [doi: 10.1016/j.ijmedinf.2015.07.004] [Medline: 26238706]
13. Wernhart A, Gahbauer S, Haluza D. eHealth and telemedicine: Practices and beliefs among healthcare professionals and medical students at a medical university. PLoS ONE 2019;14(2):e0213067. [doi: 10.1371/journal.pone.0213067] [Medline: 30818348]
14. Alam K, Mahumud RA, Alam F, Keramat SA, Erdiaw-Kwasie MO, Sarker AR. Determinants of access to eHealth services in regional Australia. Int J Med Inform 2019 Nov;131(103960):103960. [doi: 10.1016/j.ijmedinf.2019.103960] [Medline: 31518858]

XSL•FO
RenderX

15. Longhini J, Rossettini G, Palese A. Digital health competencies among health care professionals: systematic review. J Med Internet Res 2022 Aug 18;24(8):e36414. [doi: 10.2196/36414] [Medline: 35980735]
16. van der Vaart R, Drossaert C. Development of the Digital Health Literacy Instrument: Measuring a Broad Spectrum of Health 1.0 and Health 2.0 Skills. J Med Internet Res 2017 Jan 24;19(1):e27. [doi: 10.2196/jmir.6709] [Medline: 28119275]
17. Kayser L, Karnoe A, Furstrand D, et al. A multidimensional tool based on the eHealth Literacy Framework: development and initial validity testing of the eHealth Literacy Questionnaire (eHLQ). J Med Internet Res 2018 Feb 12;20(2):e36. [doi: 10.2196/jmir.8371] [Medline: 29434011]
18. Cowan DT, Norman I, Coopamah VP. Competence in nursing practice: a controversial concept--a focused review of literature. Nurse Educ Today 2005 Jul;25(5):355-362. [doi: 10.1016/j.nedt.2005.03.002] [Medline: 15904996]
19. Jarva E, Oikarinen A, Andersson J, Tomietto M, Kääriäinen M, Mikkonen K. Healthcare professionals' digital health competence and its core factors; development and psychometric testing of two instruments. Int J Med Inform 2023 Mar;171(104995):104995. [doi: 10.1016/j.ijmedinf.2023.104995] [Medline: 36689840]
20. Jarva E, Oikarinen A, Andersson J, Pramila-Savukoski S, Hammarén M, Mikkonen K. Healthcare professionals' digital health competence profiles and associated factors: a cross-sectional study. J Adv Nurs 2024 Aug;80(8):3236-3252. [doi: 10.1111/jan.16096] [Medline: 38323687]
21. Cruchinho P, López-Franco MD, Capelas ML, et al. Translation, cross-cultural adaptation, and validation of measurement instruments: a practical guideline for novice researchers. J Multidiscip Healthc 2024;17(2701-2728):2701-2728. [doi: 10.2147/JMDH.S419714] [Medline: 38840704]
22. Wenjuanxing. wjx. URL: https://www.wjx.cn [accessed 2024-11-07]
23. Zhang C, Yang Z, Zhang H. Psychometric evaluation of the Chinese version of occupational lowback pain prevention behaviors questionnaire among clinical nurses: a validation study. Front Public Health 2022;10(827604):827604. [doi: 10.3389/fpubh.2022.827604] [Medline: 35400039]
24. Polit DF, Beck CT. The content validity index: are you sure you know what's being reported? Critique and recommendations. Res Nurs Health 2006 Oct;29(5):489-497. [doi: 10.1002/nur.20147] [Medline: 16977646]
25. Jalali A, Naghibzadeh A, Mohammadi MM, et al. Translation and validation of the Persian version of the Nursing Practice Readiness Scale (NPRS) for new graduate nurses. BMC Nurs 2024 Oct 16;23(1):760. [doi: 10.1186/s12912-024-02434-6] [Medline: 39415162]
26. El-Ammari A, El Malki H, Moutawakkil SG, et al. Validation of the Center for Epidemiologic Studies Depression Scale (CES-D) in a Moroccan sample with substance use disorder. BMC Psychiatry 2023 Oct 6;23(1):723. [doi: 10.1186/s12888-023-05245-2] [Medline: 37803359]
27. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34. [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
28. Alelayan H, Liang L, Ye R, Aldosari N, Liao X. Translation and linguistic validation of the DISABKIDS chronic generic module into simplified Chinese (DCGM-37) for use among children with cancer. J Spec Pediatr Nurs 2022 Jul;27(3):e12374. [doi: 10.1111/jspn.12374] [Medline: 35415867]
29. Kaihlaniemi J, Liljamo P, Rajala M, Kaakinen P, Oikarinen A. Health care professionals' experiences of counselling competence in digital care pathways - a descriptive qualitative study. Nurs Open 2023 Jul;10(7):4773-4785. [doi: 10.1002/nop2.1729] [Medline: 36960773]
30. Nittari G, Khuman R, Baldoni S, et al. Telemedicine practice: review of the current ethical and legal challenges. Telemed J E Health 2020 Dec;26(12):1427-1437. [doi: 10.1089/tmj.2019.0158] [Medline: 32049608]
31. Kaihlanen AM, Elovainio M, Virtanen L, et al. Nursing informatics competence profiles and perceptions of health information system usefulness among registered nurses: A latent profile analysis. J Adv Nurs 2023 Oct;79(10):4022-4033. [doi: 10.1111/jan.15718] [Medline: 37243421]
32. Kaihlanen AM, Virtanen L, Buchert U, et al. Towards digital health equity - a qualitative study of the challenges experienced by vulnerable groups in using digital health services in the COVID-19 era. BMC Health Serv Res 2022 Feb 12;22(1):188. [doi: 10.1186/s12913-022-07584-4] [Medline: 35151302]
33. Ahonen O, Kouri P, Salanterä S, Liljamo P, Kinnunen UM, Saranto K, et al. Finnish nurses association's digital social and health services strategy. Finnish Nurses Association. URL: https://iris.who.int/handle/10665/366710 [accessed 2024-11-05]
34. Tynkkynen LK, Karanikolos M, Litvinova Y. Finland: Health system summary. 2023. URL: https://iris.who.int/handle/10665/366710
35. Wu G, Gong M, Wu Y, Liu L, Shi B, Zeng Z. Advancing digital health in China: aligning challenges, opportunities, and solutions with the Global Initiative on Digital Health (GIDH). Health Care Sci 2024 Oct;3(5):365-369. [doi: 10.1002/hcs2.118] [Medline: 39479272]
36. Yu-Tong T, Yan Z, Zhen L, Bing X, Qing-Yun C. Telehealth readiness and its influencing factors among Chinese clinical nurses: a cross-sectional study. Nurse Educ Pract 2022 Jan;58(103278):103278. [doi: 10.1016/j.nepr.2021.103278] [Medline: 34954659]

XSL•FO
RenderX

Abbreviations

AVE: average variance extracted
CFA: Confirmatory Factor Analysis
CFI: Comparative Fit Index
CHERRIES: Checklist for Reporting Results of Internet E-Surveys
CR: composite reliability
DHLI: Digital Health Literacy Instrument
DigiHealthCom: Digital Health Competence
eHLQ: eHealth literacy questionnaire
HTMT: heterotrait-monotrait ratio
I-CVI: item content validity index
ICC: Intraclass correlation coefficient
ICT: information and communication technology
IFI: Incremental Fit Index
NDASS: Nursing Digital Application Skill Scale
RSMEA: root-mean-square error of approximation
S-CVI/Ave: Scale CVI/Average
SRMR: standardized root-mean-square residual
TLI: Tucker-Lewis Index
WHO: World Health Organization

Edited by A Kushniruk; submitted 14.08.24; peer-reviewed by A Mainz, E Jarva, YC Uscanga; revised version received 03.02.25; accepted 10.02.25; published 21.03.25.

Please cite as:
Gao L, Chen M, Wei J, Wang J, Liao X
The Chinese Version of the DigiHealthCom (Digital Health Competence) Instrument for Assessing Digital Health Competence of Health Care Professionals: Translation, Adaptation, and Validation Study
JMIR Hum Factors 2025;12:e65373
URL: https://humanfactors.jmir.org/2025/1/e65373
doi:10.2196/65373

© Lu Gao, Meilian Chen, Jingxin Wei, Jinni Wang, Xiaoyan Liao. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 21.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Application of Clinical Department–Specific AI-Assisted Coding Using Taiwan Diagnosis-Related Groups: Retrospective Validation Study

An-Tai Lu[1,2], MSc; Chong-Sin Liou[3], MSc; Chia-Hsin Lai[3]; Bo-Tsz Shian[3]; Ming-Ta Li[2], BBA; Chih-Yen Sun[4], BSc; Hao-Yun Kao[1*], PhD; Hong-Jie Dai[5,6,7,8*], PhD; Ming-Ju Tsai[2,9,10*], PhD

[1]Department of Healthcare Administration and Medical Informatics, Kaohsiung Medical University, No.100, Shih-Chuan 1st Road, Sanmin Dist, Kaohsiung, Taiwan

[10]Department of Internal Medicine, School of Medicine, College of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan

[2]Department of Medical Records, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan

[3]Intelligent System Lab, College of Electrical Engineering and Computer Science, Department of Electrical Engineering, National Kaohsiung University of Science and Technology, Kaohsiung, Taiwan

[4]Department of Information Technology, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan

[5]Department of Electrical Engineering, National Kaohsiung University of Science and Technology, Kaohsiung, Taiwan

[6]School of Post-Baccalaureate Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan

[7]National Institute of Cancer Research, National Health Research Institutes, Tainan, Taiwan

[8]Center for Big Data Research, Kaohsiung Medical University, Kaohsiung, Taiwan

[9]Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan

[*]these authors contributed equally

Corresponding Author:
Hao-Yun Kao, PhD
Department of Healthcare Administration and Medical Informatics, Kaohsiung Medical University, No.100, Shih-Chuan 1st Road, Sanmin Dist, Kaohsiung, Taiwan

Abstract

Background: The accuracy of the *ICD-10-CM* (*International Classification of Diseases, Tenth Revision, Clinical Modification*) procedure coding system (PCS) is crucial for generating correct Taiwan diagnosis-related groups (DRGs), as coding errors can lead to financial losses for hospitals.

Objective: The study aimed to determine the consistency between an artificial intelligence (AI)-assisted coding module and manual coding, as well as to identify clinical specialties suitable for implementing the developed AI-assisted coding module.

Methods: This study examined the AI-assisted coding module from the perspective of health care professionals. The research period started in February 2023. The study excluded cases outside of Taiwan DRGs, those with incomplete medical records, and cases with Taiwan DRG disposals *ICD-10* (*International Statistical Classification of Diseases, Tenth Revision)* PCS. Data collection was conducted through retrospective medical record review. The AI-assisted module was constructed using a hierarchical attention network. The verification of the Taiwan DRGs results from the AI-assisted coding model focused on the major diagnostic categories (MDCs). Statistical computations were conducted using SPSS version 19. Research variables consisted of categorical variables represented by MDC, and continuous variables were represented by the relative weight of Taiwan DRGs.

Results: A total of 2632 discharge records meeting the research criteria were collected from February to April 2023. In terms of inferential statistics, κ statistics were used for MDC analysis. The infectious and parasitic diseases MDC, as well as the respiratory diseases MDC had κ values exceeding 0.8. Clinical inpatient specialties were statistically analyzed using the Wilcoxon signed rank test. There was not a difference in coding results between the 23 clinical departments, such as the Division of Cardiology, the Division of Nephrology, and the Department of Urology.

Conclusions: For human coders, with the assistance of the *ICD-10-CM* AI-assisted coding system, work time is reduced. Additionally, strengthening knowledge in clinical documentation enables human coders to maximize their role. This positions them to become clinical documentation experts, preparing them for further career development. Future research will apply the same method to validate the *ICD-10* AI-assisted coding module.

(JMIR Hum Factors 2025;12:e59961) doi:10.2196/59961

XSL•FO
RenderX

KEYWORDS

diagnosis-related group; artificial intelligence coding; International Classification of Diseases, Tenth Revision, Clinical Modification; ICD-10-CM; coding professionals

Introduction

The *International Statistical Classification of Diseases* (*ICD*) system was set up by the World Health Organization (WHO) for the purpose of tracking diseases globally. Over the past several decades, the WHO has made significant changes to both content and structure. It accompanies a new scientific understanding of diseases and new structures for organizing *ICD* codes to accommodate enhanced use and extensibility [1]. The WHO introduced the *ICD* in 1948, and it is a universal language used to categorize diseases or causes of death. The use of it is attributed to health care–related units in 194 countries and is generated by professional coders based on discharge records, with countries adjusting the *ICD* to their circumstances. In 2016, Taiwan adopted the international trend of switching from *ICD-9-CM* (*International Classification of Diseases, Ninth Revision, Clinical Modification*) to *ICD-10-CM* (*International Classification of Diseases, Tenth Revision, Clinical Modification*) procedure coding system (PCS) for coding hospital patient diagnoses, procedures, analysis and reimbursement. The National Health Insurance Administration (NHIA) under the Ministry of Health and Welfare has adopted the 2014 edition of *ICD-10-CM* PCS, with approximately 71,900 diagnosis codes and 78,500 procedure codes.

The use of the *ICD-10-CM* PCS involves coding and classifying morbidity data from health records, reimbursement claims, and administrative databases. Improving health care quality, monitoring public health, and conducting research are all benefits of the *ICD-10-CM* PCS in Taiwan and involves converting the physician's discharge diagnosis into *ICD-10-CM* codes by following the primary diagnosis selection principle announced by the NHIA. The diagnosis-related group (DRG) provides information such as health insurance reimbursement, relative weight, presence of comorbidities, and complications for the current hospitalization.

The accuracy of *ICD-10-CM* PCS coding is crucial for generating accurate Taiwan DRGs, as coding errors can lead to financial losses for hospitals [2,3]. According to the coding principles set forth by the NHIA and the Taiwan Society of Medical Records Management, coding is based on the inpatient and emergency room records of patients. In the past, this task was undertaken by trained and certified clinical coding professionals (referred to as coding professionals hereafter), but with the rapid advances of medical technology, the rules of disease classification have also evolved, and coding professionals must regularly accumulate relevant training hours to update their disease classification skills [4].

In recent years, artificial intelligence (AI) and natural language processing have shown exciting potential in the field of automatic clinical coding. In 2021, the disease coding scales in the United States were worth approximately 18 billion US dollars. Several technology companies in the United States have developed AI-assisted coding systems, and scholars believe that interdisciplinary collaboration and feedback from clinical coding professionals are essential to further refine the modules [5,6]. Research on AI-assisted coding consistently conclude that it improves quality and reduces error rates while saving costs [7,8]. AI-assisted *ICD-10-CM* PCS coding can be considered as a text classification task within the field of machine learning [9]. In recent years, studies in the machine learning text classification field have predominantly proposed using deep learning–based neural networks [10]. Many research papers have focused on AI assistance in *ICD-10* (*International Statistical Classification of Diseases, Tenth Revision)* coding [11-14], but few have examined the results of coding implementation from the perspective of disease classification personnel. The development and validation process of the AI-assisted coding model requires the involvement and feedback of clinical coders to enhance accuracy and correctness, aligning with user needs [15].

In Taiwan, several hospitals have also ventured into the development of AI-assisted coding for disease classification. However, due to variations in physicians' documentation of medical records across different hospitals, the AI-assisted coding systems developed are not universally applicable [11], necessitating the development and validation of customized AI-assisted coding systems. Medical coding personnel must review the discharge records meticulously and then translate the discharge diagnoses and procedures (interventions) recorded in the medical records into *ICD-10* codes. In the past, the most significant factor contributing to coding errors was handwritten medical records by physicians, which were difficult to decipher or included abbreviations, leading to mistakes [16]. In recent years, most medical centers in Taiwan have adopted electronic health records, resulting in a significant reduction in coding errors caused by handwritten records. Clinical coding personnel also encounter various pressures, including the need to accomplish all inpatient coding tasks within specified deadlines, optimize Taiwan DRGs assignment coding, enhance and maintain coding reliability and validity, and engage in discussions with clinical physicians regarding the content of medical record writing.

Recently, the global trend in AI coding has been on the rise [11,13,17-19]. In this study, we have developed an exclusive *ICD-10-CM* AI-assisted coding module. Coding professionals took part in the research and offered suggestions to improve the efficiency of coding operations. Consequently, this study focuses primarily on the following two research aims: (1) to verify the consistency between the AI-assistant coding module and a coding professional in encoding, based on the MDC results in Taiwan DRGs and (2) to find the clinical departments within the medical center that can benefit from using the developed AI-assisted coding module.

XSL•FO
RenderX

Methods

Data Description

This study used a total of 136,841 unstructured discharge summaries of patients who were hospitalized, recorded in Kaohsiung Medical University Chung-Ho Memorial Hospital from April 1, 2019, to December 31, 2020, as the primary data source. Textbox 1 displays an example of a discharge summary from the Kaohsiung Medical University Chung-Ho Memorial Hospital.

Textbox 1. Example discharge summary.

Chief complaint:

Abdominal pain for 1 day

Impression on admission:

Sepsis, focus on retroperitoneal abscess

Discharge diagnosis:

Sepsis, focus on retroperitoneal abscess due to surgical site infection

History on admission:

This time, according to the patient's statement, he suffered from recurrent abdominal dull pain after discharge. The pain was serious by jejunostomy feeding, and there was no relieving factor. The pain suddenly progressed...

This study verified the AI-assisted coding from the perspective of coding professionals. Since the AI-assisted coding system was introduced in the medical center in February 2023, the study period began in February 2023. The subjects of this study were selected based on the following exclusion criteria: non–Taiwan DRGs cases, cases with procedures (*ICD-10* PCS) in Taiwan DRGs, and cases with incomplete medical records. According to the study conditions, there were approximately 700 to 1000 cases per month. The coding by both the AI-assisted coding module and coding professionals were based on the electronic discharge summaries of a certain medical center each month.

Research Design

After each data entry was encoded by the AI-assisted coding module and verified by a coding professional, it was transmitted to a certain university's database. The results of both the AI-assisted coding and coding professional were compared using an Excel (Microsoft) file. Following the linkage to the NHIA's DRG calculation software, separate datasets for Taiwan DRGs were obtained for both the AI-assisted coding and coding personnel, with the consistency of the primary diagnosis coding between these two groups being examined. In cases of discrepancies, the medical records were scrutinized again by the coding professional to determine if the AI-assisted coding results met the criteria for primary coding as per the coding professional; the consistency results of the Taiwan DRGs data for both the AI-assisted coding and coding professional were adjusted accordingly.

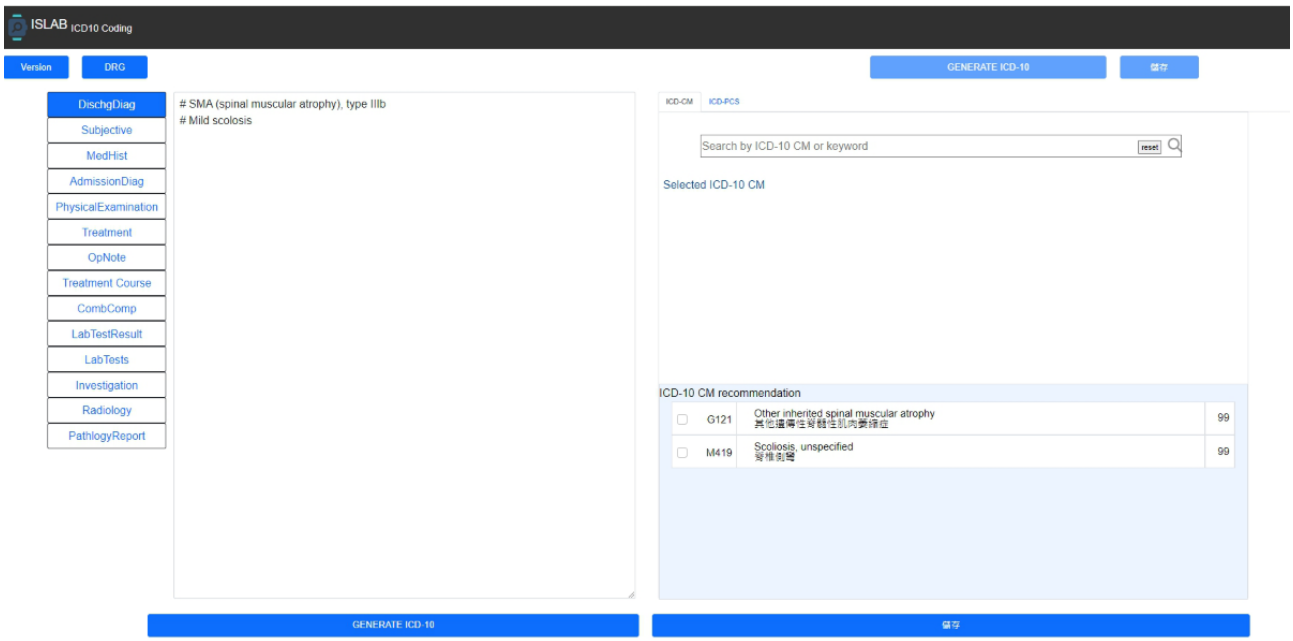
AI-Assisted Coding Construction Process

The AI-assisted *ICD-10-CM* coding system was developed by CSL, CHL, and BTS, based on approximately 110,000 discharge summaries collected from April 1, 2019 to December 31, 2020, in a medical center. The deidentified summary data were processed by segmenting sentences and filtering out meaningless delimiters and prefix symbols (eg, # or "") by using a clinical natural language processing tool [20]. The data were categorized into 21 groups based on the first 3 codes of the *ICD-10-CM*, and models were built using bidirectional encoder representations from transformers (BERTs) [21] and hierarchical attention networks (HANs) [22]. The results favored HANs, leading to the decision to adopt the HAN module. The precision, recall, and F1 scores of the developed HAN model were 0.55, 0.82, and 0.66, respectively. For the top 50 most frequent codes, the F1 score of the developed HAN model was 0.818.

Aside from module modeling, another time-consuming task was the design of the user interface for the coding professionals, as it needed to present discharge summaries, laboratory data, and imaging reports, as well as the *ICD-10-CM* codes predicted by the AI-assisted coding module. Coding professionals were actively involved in providing feedback during the interface design process. Figure 1 provides an illustration of the designed user interface, which provided suggestions automatic *ICD-10-CM* recommendation and fields for coding professionals to input the final codes. The developed AI-coding system was integrated into a medical center's hospital information systems in November 2022 and operated in February 2023.

XSL•FO
RenderX

Figure 1. User interface screenshot.



Ethical Considerations

This study was approved by Kaohsiung Medical University Ching-Ho Memorial Hospital (institutional review boards number: KMUHIRB-E(II)-20230214). The institutional review board approval covered secondary analysis without additional consent. Data was anonymized or deidentified. There was not any compensation provided to participants.

Statistical Analysis

The study involved an analysis incorporating descriptive statistics for exploration, as well as inferential statistics for investigating MDCs and relative weight. Statistical computations were conducted using SPSS version 19. Research variables consisted of categorical variables represented by MDC, and continuous variables were represented by the relative weight of Taiwan DRGs.

Results

Distribution of *ICD-10* Codes

The distribution of the *ICD-10* codes seen in the collected training dataset is shown in Multimedia Appendix 1. The first digit of the *ICD-10-CM* code consisted of English letters, so the alphabetical characters on the horizontal axis of the log data were the first digit of the *ICD-10-CM* code, showing diseases pertaining to different systems. According to Multimedia Appendix 1, data starting with codes C, E, and I in *ICD-10-CM* had the highest volume, with C representing neoplastic diseases; E for endocrinal, nutritional, and metabolic diseases; and I for diseases of the circulatory system. These were the body systems with the highest learning volumes for the AI-assisted coding module.

Descriptive Statistics

In the period from February to April 2023, a total of 15,756 discharges were recorded. Excluding cases with interventions, non–Taiwan DRG cases, and cases with incomplete medical records, there was a total of 2632 cases. The primary diagnosis was the key factor in deciding the main disease category, while secondary diagnoses only affected the distribution of Taiwan DRGs within the same primary disease category. According to disease classification rules, the primary diagnosis was based on the reason for the patient's admission, but only one disease could be selected as the primary diagnosis. If multiple diseases were treated during admission, selecting any one of them as the primary diagnosis was not considered an error. Therefore, the coding professional (author ATL) manually examined the discharged cases' notes to categorize the output of the AI-assisted system into one of the following categories. The results are shown in Table 1.

Table . Frequency distribution and percentage analysis of primary diagnoses.

Variable	Month of case, n (%)		
	February (n=748)	March (n=991)	April (n=893)
No primary diagnosis	181 (24.2)	277 (28)	285 (31.9)
Incorrect secondary diagnosis with a primary diagnosis	462 (61.8)	477 (48.1)	369 (41.3)
All correct	79 (10.6)	181 (18.3)	177 (19.8)
All incorrect	26 (3.5)	56 (5.7)	62 (7)

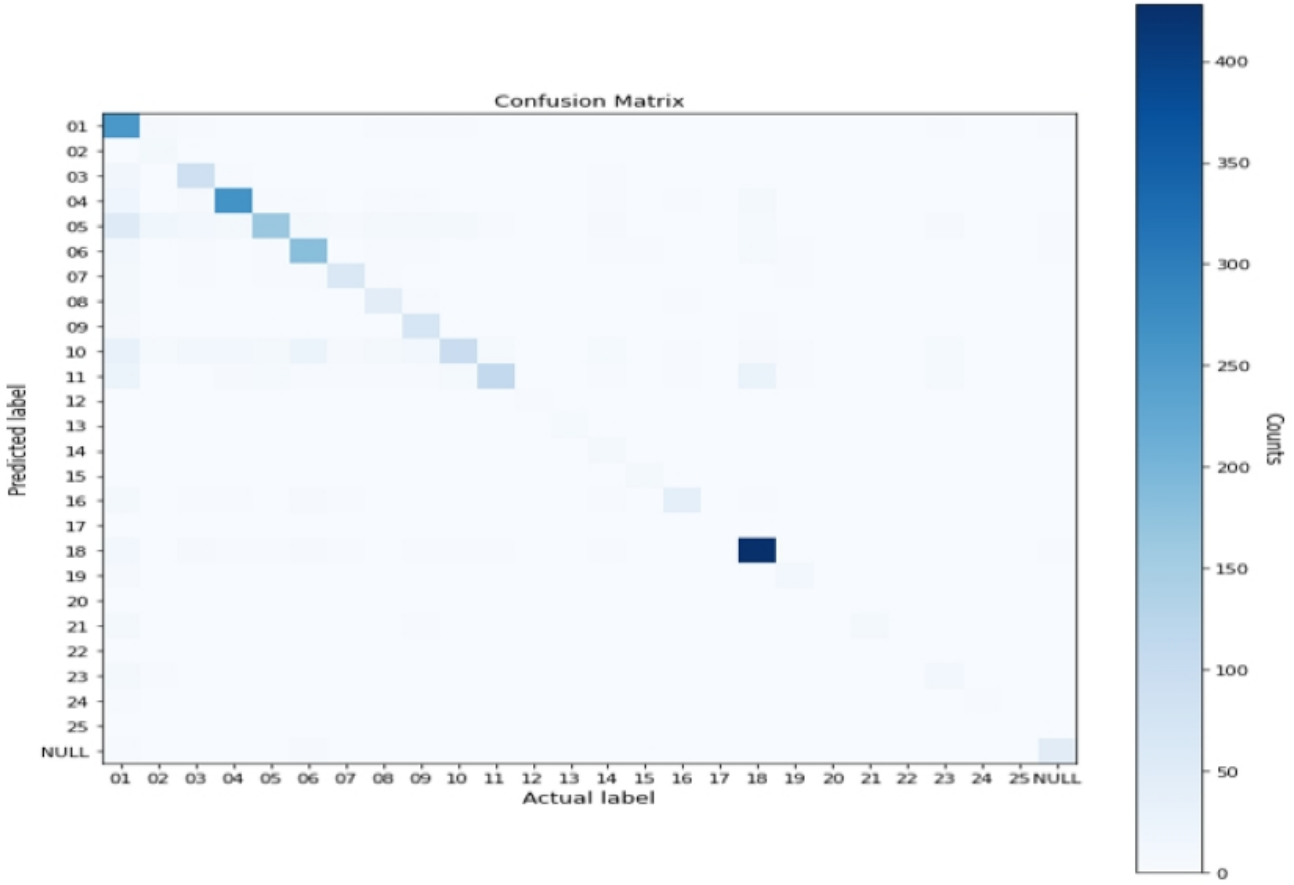
XSL•FO
RenderX

Operational definitions were as follows:

- No primary diagnosis: in comparison to the coding professional, a single hospitalization's predicted diagnosis codes did not include a primary diagnosis.
- Incorrect secondary diagnosis with a primary diagnosis: in comparison to the coding professional, a single hospitalization's predicted diagnosis codes included a primary diagnosis, but there was at least 1 error in the secondary diagnoses.
- All correct: all predicted diagnosis codes for a single hospitalization perfectly aligned with those given by the coding professional.
- All incorrect: in comparison to the disease classification personnel, none of the predicted diagnosis codes in a single hospitalization were the same.

In Figure 2, we analyzed the agreement of MDC classification between the AI-assisted coding module and the coding professional through a heat map analysis. The vertical and horizontal axes in Figure 2 represent MDCs coded by the AI-assisted coding module and MDCs coded by coding professionals, respectively. The intensity of color in the figure indicated a higher number of agreed MDCs between the AI-assisted coding module and professionals. As shown in Figure 2, MDC 1 (diseases and disorders of the nervous system), MDC 4 (diseases and disorders of the respiratory system), and MDC 18 (infectious and parasitic diseases and disorders) had the highest agreements.

Figure 2. MDC heat map analysis between AI-coding module and professionals.



The κ Coefficient Test

Furthermore, we assessed the MDC agreement between the AI coding module and coding professionals using the κ coefficient test. The κ values were broadly categorized into 5 groups based on various levels of agreement: extremely low agreement (0.00-0.20), fair agreement (0.21-0.40), moderate agreement (0.41-0.60), high agreement (0.61-0.80), and almost perfect agreement (0.81-1.0).

When analyzing the cumulative data for February to April 2023 (Table 2), the MDCs with the highest consistency were MDC 4 (diseases and disorders of the respiratory system) and MDC 18 (infectious and parasitic diseases and disorders), followed by MDC 1 (diseases and disorders of the nervous system), MDC 3 (diseases and disorders of the ear, nose, mouth and throat), MDC 6 (diseases and disorders of the digestive system), MDC 7 (diseases and disorders of the hepatobiliary system and pancreas), MDC 9 (diseases and disorders of the skin, subcutaneous tissue and breast), MDC 11 (diseases and disorders of the kidney and urinary tract), MDC 13 (diseases and disorders of the female reproductive system), MDC 15 (newborn and other neonates), and MDC 16 (diseases and disorders of the blood and blood forming organs and immunological disorders).

XSL•FO
RenderX

Table . Kappa tests for aggregation of major diagnostic category in the total counts for February to April 2023 (κ=0.592).

No.	Major diagnostic category	AI-assisted case coding (n=2362), n (%)	Cases coded by human coders (n=2362), n (%)	Kappa value
1	Diseases and disorders of the nervous system	280 (10.6)	509 (19.3)	0.670[a]
2	Diseases and disorders of the eye	9 (0.3)	38 (1.4)	0.300
3	Diseases and disorders of the ear, nose, mouth, and throat	113 (4.3)	132 (5)	0.689[a]
4	Diseases and disorders of the respiratory system	309 (11.7)	302 (11.5)	0.845[b]
5	Diseases and disorders of the circulatory system	310 (11.8)	184 (7)	0.607
6	Diseases and disorders of the digestive system	217 (8.2)	229 (8.7)	0.775[a]
7	Diseases and disorders of the hepatobiliary system and pancreas	90 (3.4)	84 (3.2)	0.710[a]
8	Diseases and disorders of the musculoskeletal system and connective tissue	66 (2.5)	87 (3.3)	0.576
9	Diseases and disorders of the skin, subcutaneous tissue, and breast	83 (3.2)	116 (4.4)	0.692[a]
10	Diseases and disorders of the endocrine, nutritional, and metabolic systems	237 (9)	132 (5)	0.505
11	Diseases and disorders of the kidney and urinary tract	205 (7.8)	120 (4.6)	0.648[a]
12	Diseases and disorders of the male reproductive system	7 (0.3)	3 (0.1)	0.362
13	Diseases and disorders of the female reproductive system	8 (0.3)	8 (0.3)	0.749[a]
14	Pregnancy, childbirth and puerperium	11 (0.4)	41 (1.6)	0.419
15	Newborn and other neonates (perinatal period)	9 (0.3)	15 (0.6)	0.635[a]
16	Diseases and disorders of the blood and blood forming organs and immunological disorders	64 (2.4)	57 (2.2)	0.624[a]
17	Myeloproliferative diseases and disorders (poorly differentiated neoplasms)	3 (0.1)	2 (0.1)	0.399
18	Infectious and parasitic diseases and disorders	465 (17.7)	505 (19.2)	0.870[b]
19	Mental diseases and disorders	23 (0.9)	0 (0)	—[c]
20	Alcohol or drug abuse or induced mental disorder	2 (0.1)	0 (0)	—
21	Injuries, poison, and toxic effects of drugs	24 (0.9)	21 (0.8)	0.404
22	Burns	0 (0)	1 (0)	—

XSL•FO
RenderX

No.	Major diagnostic category	AI-assisted case coding (n=2362), n (%)	Cases coded by human coders (n=2362), n (%)	Kappa value
23	Factors influencing health status and other contacts with health services	24 (0.9)	42 (1.6)	0.366
24	Multiple significant trauma	6 (0.2)	4 (0.2)	0.599
25	HIV infection	2 (0.1)	0 (0)	—
—	None	65 (2.5)	0 (0)	—

[a]High agreement (0.61-0.80).

[b]Almost perfect agreement (0.81-1.00).

[c]Kappa value was not calculated when there were 0 cases in a coding group.

Inferential Statistical Analysis: Wilcoxon Signed Rank Test

The κ coefficient test was used for a broad-scale analysis of MDCs. However, under the same MDC, it was possible to further classify the data into numerous Taiwan DRGs, with each having its own code and relative weight. Even within the same MDC, this might result in different Taiwan DRGs. Furthermore, some diseases could be treated across departments. Therefore, for the statistical analysis of relative weight, we first conducted a normality analysis of the relative weights obtained from both AI-assisted coding and coding professionals. The statistical results based on the Kolmogorov-Smirnov analysis yielded a significance level of less than .05, showing a nonnormal distribution. Given that the research sample consisted of paired data, the nonparametric Wilcoxon signed rank test was used to analyze whether there were differences in relative weight between AI-assisted coding and coder-assigned coding; the null hypothesis assumed that there was no difference in relative weight between AI-assisted coding and coder-assigned coding.

The Wilcoxon signed rank test, with clinical departments as the unit of analysis, identified differences in relative weight in the following 12 departments: Division of Endocrinology and Metabolism, Division of Hematology and Oncology, Division of General Internal Medicine, Division of Geriatrics and Gerontology, Division of Trauma, Division of Neurosurgery, Division of Cardiovascular Surgery, Division of General and Digestive Surgery, Division of Pediatric Neurology, Department of Otorhinolaryngology, Department of Neurology, and Department of Rehabilitation Medicine. As shown in Table 3, the overall statistical result with a *P* value of <.001 showed that there were still differences between AI-assisted coding and coder-assigned coding in this study.

XSL•FO
RenderX

Table . Wilcoxon signed rank test results across various clinical departments (P<.001).

Clinical department	Frequency of cases (n=2632), n (%)	Relative weight (95%CI)		*P* value
		AI[a] coding	Human coding	
Division of Gastroenterology	61 (2.3)	0.66 (0.58 - 0.74)	0.70 (0.63 - 0.77)	.12
Division of Hepatobiliary and Pancreatic Medicine	81 (3.1)	0.69 (0.63 - 0.75)	0.74 (0.69 - 0.79)	.07
Division of Cardiology	177 (6.7)	0.69 (0.65 - 0.73)	0.72 (0.68 - 0.75)	.39
Division of Chest Medicine	181 (6.9)	0.90 (0.85 - 0.93)	0.90 (0.87 - 0.94)	.50
Division of Nephrology	64 (2.4)	0.71 (0.65 - 0.78)	0.71 (0.64 - 0.78)	.91
Division of Endocrinology and Metabolism	36 (1.4)	0.62 (0.54 - 0.71)	0.68 (0.59 - 0.77)	.03
Division of Hematology and Oncology	42 (1.6)	0.79 (0.69 - 0.88)	0.87 (0.80 - 0.94)	.006
Division of Rheumatology, Immunology, and Allergology	36 (1.4)	0.68 (0.58 - 0.79)	0.71 (0.64 - 0.79)	.33
Division of Infectious Diseases	78 (3)	0.91 (0.85 - 0.98)	0.94 (0.90 - 0.99)	.39
Division of General Internal Medicine	205 (7.8)	0.69 (0.65 - 0.73)	0.73 (0.69 - 0.77)	.005
Division of Geriatrics and Gerontology	54 (2.1)	0.92 (0.85 - 0.99)	0.99 (0.94 - 1.04)	.003
Division of Trauma	16 (0.6)	0.43 (0.21 - 0.65)	0.64 (0.46 - 0.83)	.04
Division of Neurosurgery	151 (5.7)	0.53 (0.48 - 0.58)	0.68 (0.63 - 0.73)	<.001
Division of Cardiovascular Surgery	24 (0.9)	0.67 (0.54 - 0.80)	0.94 (0.82 - 1.06)	.002
Division of Chest Surgery	14 (0.5)	0.56 (0.43 - 0.69)	0.60 (0.45 - 0.75)	.40
Division of Pediatric Surgery	9 (0.3)	0.45 (0.31 - 0.59)	0.48 (0.33 - 0.63)	.28
Division of Plastic Surgery	9 (0.3)	0.64 (0.38 - 0.90)	0.68 (0.44 - 0.92)	.89
Division of Colorectal Surgery	46 (1.7)	0.54 (0.45 - 0.62)	0.59 (0.52 - 0.66)	.16
Division of Breast Oncology and Surgery	16 (0.6)	0.47 (0.32 - 0.62)	0.62 (0.48 - 0.75)	.05
Division of General and Digestive Surgery	56 (2.1)	0.56 (0.49 - 0.62)	0.63 (0.57 - 0.68)	.009
Department of Gynecology Obstetrics	60 (2.3)	0.47 (0.39 - 0.55)	0.46 (0.40 - 0.52)	.72
Division of Pediatric Hematology and Oncology	42 (1.6)	0.45 (0.38 - 0.52)	0.47 (0.40 - 0.55)	.44
Division of Pediatric Cardiology and Pulmonology	86 (3.3)	0.52 (0.37 - 0.66)	0.53 (0.43 - 0.63)	.24
Division of Pediatric Neurology	92 (3.5)	0.43 (0.34 - 0.52)	0.45 (0.40 - 0.50)	.008
Division of Neonatology	14 (0.5)	0.87 (0.28 - 1.47)	0.81 (0.43 - 1.18)	.69
Division of General Pediatrics	299 (11.4)	0.39 (0.36 - 0.41)	0.39 (0.37 - 0.41)	.56
Division of Pediatric Allergy Immunology	8 (0.3)	0.31 (0.17 - 0.46)	0.35 (0.25 - 0.46)	.32

XSL•FO
RenderX

Clinical department	Frequency of cases (n=2632), n (%)	Relative weight (95%CI)		*P* value
		AI[a] coding	Human coding	
Department of Otorhinolaryngology	53 (2)	0.57 (0.50 - 0.63)	0.51 (0.45 - 0.58)	.02
Ophthalmology Department	12 (0.5)	0.42 (0.34 - 0.50)	0.44 (0.36 - 0.52)	.44
Department of Orthopaedics	13 (0.5)	0.40 (0.25 - 0.55)	0.52 (0.39 - 0.65)	.16
Department of Urology	46 (1.7)	0.63 (0.56 - 0.70)	0.62 (0.55 - 0.69)	.95
Department of Dermatology	87 (3.3)	0.42 (0.37 - 0.48)	0.41 (0.35 - 0.46)	.18
Department of Neurology	366 (13.9)	0.66 (0.63 - 0.69)	0.72 (0.69 - 0.75)	<.001
Division of Family Medicine	49 (1.9)	0.96 (0.87 - 1.04)	1.00 (0.93 - 1.07)	.17
Department of Rehabilitation Medicine	47 (1.8)	0.85 (0.75 - 0.94)	1.21 (1.14 - 1.28)	<.001
Department of Psychiatry	1 (0)	—	—	—
Division of Oral Maxillofacial Surgery	1 (0)	—	—	—

[a]AI: artificial intelligence.

Discussion

Principal Results

For clinical coders, it is clear from the MDCs that AI-assisted coding can serve as a reference for disease systems. However, hospital administrators may require detailed statistical results from clinical departments to make judgments. In the individual clinical department analysis based on the Wilcoxon signed rank test, the Division of General Internal Medicine, the Department of Neurology, and the Division of Neurosurgery had the highest number of cases studied, but the statistical results were inconsistent with coder-assigned coding. However, in the κ coefficient test, the statistical results for the nervous system MDC were highly consistent. This is because patients admitted to the Department of Neurology and Neurosurgery do not exclusively have neurological conditions. According to further analysis shown in Table 4, the AI model's predictions for neurological system diseases were still highly consistent with those of the disease classification staff in the neurology department. However, respiratory and urinary system conditions have affected the AI model's coding performance for neurology and account for the discrepancies seen in both the Wilcoxon signed rank test and the κ coefficient test.

XSL•FO
RenderX

Table . Analysis of major diagnostic category for 366 neurology patients admitted from February to April 2023.

No.	MDC[a]	AI[b]-assisted case coding (n=366), n (%)	Cases coded by human coders (n=366), n (%)	Kappa value	*P* value
1	Diseases and disorders of the nervous system	281 (76.8)	289 (79)	0.62	<.001
2	Diseases and disorders of the eye	17 (4.6)	18 (4.9)	0.84	<.001
3	Diseases and disorders of the ear, nose, mouth, and throat	21 (5.7)	20 (5.5)	0.77	<.001
4	Diseases and disorders of the respiratory system	1 (0.3)	1 (0.3)	–0.03	.96
5	Diseases and disorders of the circulatory system	20 (5.5)	12 (3.3)	0.54	<.001
6	Diseases and disorders of the digestive system	1 (0.3)	1 (0.3)	1.00	<.001
8	Diseases and disorders of the musculoskeletal system, and connective tissue	6 (1.6)	8 (2.2)	0.56	<.001
9	Diseases and disorders of the skin, subcutaneous tissue, and breast	2 (0.6)	2 (0.6)	1.00	<.001
10	Diseases and disorders of the endocrine, nutritional, and metabolic systems	1 (0.3)	2 (0.6)	0.67	<.001
11	Diseases and disorders of the kidney and urinary tract	4 (1.1)	1 (0.3)	0.40	<.001
14	Pregnancy, childbirth, and puerperium	0 (0)	1 (0.3)	—[c]	—
16	Diseases and disorders of the blood and blood forming organs and immunological disorders	1 (0.3)	1 (0.3)	1.00	<.001
17	Myeloproliferative diseases and disorders (poorly differentiated neoplasms)	0 (0)	1 (0.3)	—	—
18	Infectious and parasitic diseases and disorders	4 (1.1)	0 (0)	—	—
19	Mental diseases and disorders	3 (0.8)	0 (0)	—	—
23	Factors influencing health status and other contacts with health services	1 (0.3)	9 (2.5)	–0.05	.87
—	None	3 (0.8)	0 (0)	—	—

[a]MDC: major diagnostic category.

[b]AI: artificial intelligence.

[c]Analysis was not performed when there were 0 cases in a coding group.

XSL•FO
RenderX

In the circulatory system, the statistical results for the Division of Cardiology and Division of Cardiovascular Surgery in the Wilcoxon signed rank test were also markedly different. Upon closer examination of the data from the exploratory study, it was discovered that in the Division of Cardiovascular Surgery, half of the cases helped by AI-coding modules did not have the main diagnosis coded, which could be attributed to differences in how physicians document medical records. For example, after carefully reviewing the 24 cases of data collected by the Cardiac Surgery Department, it was found that 12 Taiwan DRGs were inconsistent. All of these did not follow the disease classification coding rules and did not include the main diagnosis (Table 5).

Table . Discussion on writing medical records in the Division of Cardiovascular Surgery.

Number	Excerpt of discharge diagnosis	Cause analysis
Case 1	"Chest tightness for a week Acute heart failure with reduced ejection fraction * 2023/04/24 Thallium 201 (Stress SPECT (single-photon emission computed tomography) imaging): mild myocardial ischemia in inferolateral wall of LV (left ventricle)"	The examination results were attached to the discharge diagnosis, leading to coding confusion.
Case 2	"Type A aortic dissection - post TEVAR (Thoracic endovascular aneurysm repair)+ stenting grafts in the ascending to descending aorta, left common carotid and subcalvian arteries on 2023/02/07 # Suspect gastroparesis related to relative gastric malperfusion"	The discharge diagnosis showed coding confusion with previously treated conditions.
Case 3	"Type B dissection, intramural hemorrhage - 2023/03/10 Chest CT (computed tomography) angiography:1) Suspect intramural hematoma in the descending aorta.2) Suspect thrombus formation in the bilateral femoral arteries.3) Suspect a thrombosed aneurysm in the right internal iliac artery"	The anatomical location was not clearly documented, making correct coding impossible.

Furthermore, in MDC 14 (pregnancy, childbirth, and puerperium) and MDC 21 (injuries, poisonings, and toxic effects of drugs), there are specific coding rules. Clinical coders need to synthesize the entire medical record information and apply the coding rules, which could result in diagnoses different from those presented in the discharge summary.

Limitations

The AI-coding module was trained on inpatient data from April 2019 to December 2020. Advancements in medical care might lead to variations in the diseases of admitted patients. Taken together, these show situations where the AI-coding assistance module might not capture the main diagnosis, as observed in the Dermatology Department.

Conclusions

With the rapid advancements in global medical technology and the evolving challenges of diseases, the development of DRG-based hospital payment systems in various countries is also meeting significant challenges. Key areas for future research include determining the flexibility of DRG payments, balancing payment structures, and aligning with disease management goals [23]. The Taiwan DRGs system, like those in other countries, aims to prevent medical institutions from delivering excessive services and causing unnecessary waste, all while safeguarding patient rights. It looks to strengthen management mechanisms to improve the quality and efficiency of care and ensure fair payments among peers.

In this context, AI-assisted coding emerges as a powerful tool [24]. A recent study used cross-random control methods to prove that AI-assisted coding reduces the coding workload [25]. The focuses of our research were the practical applications of AI models, with two main goals. The first was to investigate the consistency between the AI-assisted coding module and coding professionals and the second was to find the departments suitable for using the AI-assisted coding module. The research results showed that the highest consistency in MDC classification was seen in diseases of the respiratory system, as well as infectious and parasitic diseases. In the analysis of various inpatient specialties, departments such as the Division of Cardiology, Division of Nephrology, and Department of Urology showed no significant difference from coder-assigned coding results; accordingly, consideration could be given to integrating the AI-assisted coding module into the hospital information system, allowing physicians to reference Taiwan DRGs assignments for hospitalized patients, thus effectively controlling medical expenses.

However, upon analyzing the entire hospital department, discrepancies were observed in alignment with disease categorizations and personnel coding, so the research team is actively working on continuous improvements. Nevertheless, AI-assisted coding indeed served as a valuable reference by reducing human errors, as during the research period, it was found that the error rate detected by human coders (number of coding errors by human coders/total cases) was 1.9% (50/2632). Given the regular updates to the tool book by the Department of Health and the revisions in coding rules, the coding assistance module undoubtedly proves to be a powerful tool.

XSL•FO
RenderX

The development of AI-assisted coding for the *ICD-10-CM* PCS is just the beginning for intelligent health care in disease classification. Many operational aspects of hospitals are closely related to the *ICD-10-CM* PCS, including inpatient coding monitoring, discharge preparation services, and infectious disease surveillance, among others. For hospital administrators, the goal of AI-assisted coding is to achieve best operational revenue. For human coders with the assistance of an *ICD-10-CM* AI coding system, work time is reduced. Additionally, strengthening knowledge in clinical documentation improvement enables human coders to maximize their role, positioning them to become documentation experts [15] and preparing them for further career development.

Acknowledgments

This study was supported by grants from Kaohsiung Medical University Hospital (SI11315).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Distribution of *ICD-10* (*International Statistical Classification of Diseases, Tenth Revision*) codes.
[JPEG File, 41 KB - humanfactors_v12i1e59961_app1.jpeg]

References

1. Hirsch JA, Nicola G, McGinty G, et al. ICD-10: history and context. AJNR Am J Neuroradiol 2016 Apr;37(4):596-599. [doi: 10.3174/ajnr.A4696] [Medline: 26822730]
2. Ayub S, Scali ST, Richter J, et al. Financial implications of coding inaccuracies in patients undergoing elective endovascular abdominal aortic aneurysm repair. J Vasc Surg 2019 Jan;69(1):210-218. [doi: 10.1016/j.jvs.2018.04.027] [Medline: 29937283]
3. Toner E, Khaled A, Ramesh A, Qureshi MK, Al Suyyagh K, Dunkow P. Financial impact of inaccurate coding plus cost-effectiveness analysis for surgically managed patients with periprosthetic fractures. Cureus 2021 Feb 1;13(2):e13060. [doi: 10.7759/cureus.13060] [Medline: 33680602]
4. Otero Varela L, Doktorchik C, Wiebe N, et al. International Classification of Diseases clinical coding training: an international survey. Hlth Inf Manag 2024 May;53(2):68-75. [doi: 10.1177/18333583221106509] [Medline: 35838185]
5. U.S. medical coding market size, share & trends analysis report by classification system (ICD, HCPCS, CPT), by component, by medical specialty, and segment forecasts, 2023 - 2030. Grand View Research. URL: https://www.grandviewresearch.com/industry-analysis/us-medical-coding-market [accessed 2025-01-06]
6. Venkatesh KP, Raza MM, Kvedar JC. Automating the overburdened clinical coding system: challenges and next steps. NPJ Digit Med 2023 Feb 3;6(1):16. [doi: 10.1038/s41746-023-00768-0] [Medline: 36737496]
7. Tong S, Jia N, Luo X, Fang Z. The Janus face of artificial intelligence feedback: deployment versus disclosure effects on employee performance. Strat Manag J 2021 Sep;42(9):1600-1631. [doi: 10.1002/smj.3322]
8. Vrontis D, Christofi M, Pereira V, Tarba S, Makrides A, Trichina E. Artificial intelligence, robotics, advanced technologies and human resource management: a systematic review. Int J Hum Resour Manag 2022 Mar 26;33(6):1237-1266. [doi: 10.1080/09585192.2020.1871398]
9. Kister TS, Remmler J, Schmidt M, et al. Acute kidney injury and its progression in hospitalized patients-results from a retrospective multicentre cohort study with a digital decision support system. PLoS ONE 2021;16(7):e0254608. [doi: 10.1371/journal.pone.0254608] [Medline: 34252151]
10. Wu S, Roberts K, Datta S, et al. Deep learning in clinical natural language processing: a methodical review. J Am Med Inform Assoc 2020 Mar 1;27(3):457-470. [doi: 10.1093/jamia/ocz200] [Medline: 31794016]
11. Chen PF, Wang SM, Liao WC, et al. Automatic ICD-10 coding and training system: deep neural network based on supervised learning. JMIR Med Inform 2021 Aug 31;9(8):e23230. [doi: 10.2196/23230] [Medline: 34463639]
12. Chraibi A, Delerue D, Taillard J, Chaib Draa I, Beuscart R, Hansske A. A deep learning framework for automated ICD-10 coding. Stud Health Technol Inform 2021 May 27;281:347-351. [doi: 10.3233/SHTI210178] [Medline: 34042763]
13. Diao X, Huo Y, Zhao S, et al. Automated ICD coding for primary diagnosis via clinically interpretable machine learning. Int J Med Inform 2021 Sep;153:104543. [doi: 10.1016/j.ijmedinf.2021.104543] [Medline: 34391016]
14. Chen Y, Chen H, Lu X, Duan H, He S, An J. Automatic ICD-10 coding: deep semantic matching based on analogical reasoning. Heliyon 2023 Apr;9(4):e15570. [doi: 10.1016/j.heliyon.2023.e15570] [Medline: 37151662]
15. Campbell S, Giadresco K. Computer-assisted clinical coding: a narrative review of the literature on its benefits, limitations, implementation and impact on clinical coding professionals. Hlth Inf Manag 2020 Jan;49(1):5-18. [doi: 10.1177/1833358319851305] [Medline: 31159578]

XSL•FO
RenderX

16. Asadi F, Ahmadi Hosseini M, Gomar T, Sabahi A. Factors affecting clinical coding errors. Shiraz E-Med J 2022;23(9). [doi: 10.5812/semj-122161]
17. Islam MM, Li GH, Poly TN, Li YCJ. DeepDRG: performance of artificial intelligence model for real-time prediction of diagnosis-related groups. Healthcare (Basel) -> Healthc (Basel) 2021 Nov 25;9(12):1632. [doi: 10.3390/healthcare9121632] [Medline: 34946357]
18. Shuai Z, Xiaolin D, Jing Y, et al. Comparison of different feature extraction methods for applicable automated ICD coding. BMC Med Inform Decis Mak 2022 Jan 12;22(1):11. [doi: 10.1186/s12911-022-01753-5] [Medline: 35022039]
19. Subotin M, Davis AR. A method for modeling co-occurrence propensity of clinical codes with application to ICD-10-PCS auto-coding. J Am Med Inform Assoc 2016 Sep;23(5):866-871. [doi: 10.1093/jamia/ocv201] [Medline: 26911826]
20. Chang NW, Dai HJ, Jonnagaddala J, Chen CW, Tsai RTH, Hsu WL. A context-aware approach for progression tracking of medical concepts in electronic medical records. J Biomed Inform 2015 Dec;58 Suppl(Suppl):S150-S157. [doi: 10.1016/j.jbi.2015.09.013] [Medline: 26432355]
21. Devlin J, Chang MW, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. Presented at: 2019 Conference of the North American Chapter of the Association for Computational Linguistics; Jun 2-7, 2019; Minneapolis, MN p. 4171-4186. [doi: 10.18653/v1/N19-1423]
22. Yang Z, Yang D, Dyer C, He X, Smola A, Hovy E. Hierarchical attention networks for document classification. Presented at: 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies; Jun 12-17, 2016; San Diego, CA p. 1480-1489. [doi: 10.18653/v1/N16-1174]
23. Liu Y, Wang G, Qin TG, Kobayashi S, Karako T, Song P. Comparison of diagnosis-related groups (DRG)-based hospital payment system design and implementation strategies in different countries: the case of ischemic stroke. Biosci Trends 2024 Mar 19;18(1):1-10. [doi: 10.5582/bst.2023.01027] [Medline: 38403739]
24. Wang H, Gao C, Dantona C, Hull B, Sun J. DRG-LLaMA: tuning LLaMA model to predict diagnosis-related group for hospitalized patients. NPJ Digit Med 2024 Jan 22;7(1):16. [doi: 10.1038/s41746-023-00989-3] [Medline: 38253711]
25. Chomutare T, Lamproudis A, Budrionis A, et al. Improving quality of ICD-10 (International Statistical Classification of Diseases, Tenth Revision) coding using AI: protocol for a crossover randomized controlled trial. JMIR Res Protoc 2024 Mar 12;13:e54593. [doi: 10.2196/54593] [Medline: 38470476]

Abbreviations

AI: artificial intelligence
BERT: bidirectional encoder representations from transformers
DRG: diagnosis-related group
HAN: hierarchical attention network
***ICD*:** *International Classification of Diseases*
***ICD-10*:** *International Statistical Classification of Diseases, Tenth Revision*
***ICD-10-CM*:** *International Classification of Diseases, Tenth Revision, Clinical Modification*
***ICD-9-CM*:** *International Classification of Diseases, Ninth Revision, Clinical Modification*
MDC: major diagnostic category
NHIA: National Health Insurance Administration
PCS: procedure coding system
WHO: World Health Organization

Edited by A Kushniruk; submitted 30.04.24; peer-reviewed by S Mao, YT Chang; revised version received 28.09.24; accepted 01.12.24; published 12.02.25.

Please cite as:
Lu AT, Liou CS, Lai CH, Shian BT, Li MT, Sun CY, Kao HY, Dai HJ, Tsai MJ
Application of Clinical Department–Specific AI-Assisted Coding Using Taiwan Diagnosis-Related Groups: Retrospective Validation Study
JMIR Hum Factors 2025;12:e59961
URL: https://humanfactors.jmir.org/2025/1/e59961
doi:10.2196/59961

© An-Tai Lu, Chong-Sin Liou, Chia-Hsin Lai, Bo-Tsz Shian, Ming-Ta Li, Chih-Yen Sun, Hao-Yun Kao, Hong-Jie Dai, Ming-Ju Tsai. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 12.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human

XSL•FO
RenderX

Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Enhancing Early Language Disorder Detection in Preschools: Evaluation and Future Directions for the Gades Platform

María Dolón-Poza[1*], MSc; Ana-Marta Gabaldón-Pérez[1*], PhD; Santiago Berrezueta-Guzman[2*], PhD; David López Gracia[1*], MSc; María-Luisa Martín-Ruiz[1*], PhD; Iván Pau De La Cruz[1*], PhD

[1]Grupo de Investigación Innovación Tecnológica para las Personas (InnoTep), Departamento de Ingeniería Telemática y Electrónica, ETSIS Telecomunicación, Campus Sur, Universidad Politécnica de Madrid, Madrid, Spain

[2]Technical University of Munich, Heilbronn, Germany

[*]all authors contributed equally

Corresponding Author:
Santiago Berrezueta-Guzman, PhD
Technical University of Munich
BildungsCampus 2
Heilbronn, 74076
Germany
Phone: 49 15159184710
Email: s.berrezueta@tum.de

Abstract

Background: Language acquisition is a critical developmental milestone, with notable variability during the first 4 years of life. Developmental language disorder (DLD) often overlaps with other neurodevelopmental disorders or simple language delay (SLD), making early detection challenging, especially for primary caregivers.

Objective: We aimed to evaluate the effectiveness of the Gades platform, an adaptive screening tool that enables preschool teachers to identify potential language disorders without direct support from nursery school language therapists (NSLTs).

Methods: The study took place in a nursery school and an early childhood educational and psychopedagogical center in Madrid, Spain, involving 218 children aged 6 to 36 months, 24 preschool teachers, and 2 NSLTs. Initially, NSLTs conducted informational sessions to familiarize teachers with DLDs and how to identify them. Following this, the teachers used the Gades platform to conduct language screenings independently, without ongoing support from NSLTs. The Gades platform was enhanced to collect detailed profiles of each child and implemented an adaptive screening model tailored to account for variability in language development. This setup allowed preschool teachers, who are not language experts, to observe and assess language development effectively in natural, unsupervised educational environments. The study assessed the platform's utility in guiding teachers through these observations and its effectiveness in such settings.

Results: Gades identified language difficulties in 19.7% (43/218) of the children, with a higher prevalence in boys (29/218, 13.3%) than in girls (14/218, 6.4%). These challenges were most frequently observed in children aged 15 to 27 months. The platform demonstrated a high accuracy rate of 97.41%, with evaluators largely agreeing with its recommendations. Teachers also found Gades to be user friendly and a valuable tool for supporting language development observations in everyday educational settings.

Conclusions: Gades demonstrates potential as a reliable and accessible tool for early detection of language disorders, empowering educators to identify DLD and SLD in the absence of NSLTs. However, further refinement of the platform is required to effectively differentiate between DLD and SLD. By integrating Gades into routine preschool assessments, educators can facilitate timely interventions, bridging gaps in early childhood education and therapy.

Trial Registration: Pan-African Clinical Trial Registry (PACTR) PACTR202210657553944; https://pactr.samrc.ac.za/TrialDisplay.aspx?TrialID=24051

(JMIR Hum Factors 2025;12:e60424) doi:10.2196/60424

KEYWORDS

developmental language disorder; simple language delay; adaptive screening system; early childhood education; pervasive therapy

XSL•FO
RenderX

Introduction

Screening and Prevalence of Developmental Language Disorder

Communication plays a fundamental role in children's cognitive development. Through communication, children acquire knowledge, express their thoughts and emotions, develop their cognitive ability, and establish relationships with others. Adequate communication development enables them to learn and participate in various social contexts. Therefore, language acquisition during childhood is one of the most critical developmental milestones but with the most substantial interindividual variability in the first 4 years of life. Children start consolidating their language learning process from the age of 4 to 6 years [1-7].

The *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (*DSM-5*), published by the American Psychiatric Association, introduced a new category known as neurodevelopmental disorders (NDDs) that groups different conditions that emerge early in child development before starting elementary school. These conditions eventually persist throughout adulthood [8-12].

The NDDs include communication disorders related to language, speech, and communication impairments. Developmental language disorder (DLD) is classified as a communication disorder. Children with DLD experience challenges in the acquisition and use of language. Language proficiency depends on receptive and expressive abilities. Therefore, children diagnosed with DLD encounter difficulties in both domains [12]. Thus, vocabulary is more limited than expected, and sentences are shorter and less complex with grammatical errors or with speech alterations in narration and comprehension or production of sentences. In addition, they may appear shy or prefer to communicate only with their family members. Children with DLD not only have symptomatology in communication skills but also have difficulties in cognitive and sociocultural processes [5,13-16].

Moreover, a family history of language disorders is often present, and DLD occurs with other disorders, such as autism spectrum disorder and attention-deficit/hyperactivity disorder (ADHD), in 30% of cases [5,17,18]. This makes early detection of DLD in these children more challenging for primary care pediatricians due to the significant comorbidity and similar symptoms with other NDDs [19].

On the other hand, early detection of DLD can be masked by misidentifying it as simple language delay (SLD). DLD is a persistent disorder with a slow rate of improvement and significant variability. This differentiates it from SLD [1,5,20]. However, this indicator does not allow for early detection of DLD, as we cannot identify whether the child has a one-off delay or is showing symptoms of DLD. To differentiate among them, difficulties must be identified in other areas, such as semantics, pragmatics, morphosyntax, and phonology [3]. Therefore, language difficulties manifest through the aforementioned skills that are measurably below what is expected for age, significantly impacting academic achievement, work performance, communication, and socialization [8-11].

DLD is diagnosed at the age of 4 years and is usually stable over time, as interindividual differences in language ability are reduced [1]. However, there may be warning signs that manifest themselves earlier. A study carried out by Vall d'Hebron Hospital Universitari in Barcelona, Spain, highlights a significant portion of children aged 5 to 17 years who exhibit clear symptoms of NDDs but have not been previously diagnosed with DLD [21]. Moreover, communication disorders are one of the significant NDDs with more prevalence in Spanish schools (1.05%-3.42%), along with ADHD and learning disorders [21,22]. Other studies estimated that the population prevalence of language disorders, without any relation to other intellectual disabilities, is 7.58% for children aged 4 to 5 years [17] and 6.4% at the age of 10 years [23]. To summarize, within a preschool setting of 30 children, it can be anticipated that about 2 children may exhibit DLD manifestations [17]. In total, 10% of preschool children present difficulties in language acquisition, of whom 5% to 7% end up being diagnosed with DLD, and the rest are diagnosed with SLD [5].

Before the detection and diagnosis of DLD, there is a prevention phase. Here, the nursery school plays a key role, as it is within this environment that the child spends most of their time and is involved in different interactions [1]. These interactions contribute to developing their communicative skills (communicative intention, nonverbal communication, imitation, waiting, etc) with peers of a similar age and teachers. Research has demonstrated a correlation between children with DLD and their academic performance. In fact, 88% of these children fail to meet the necessary curricular standards during their first year of school [14,17]. Hence, there is a pressing need to support education professionals with enough training and knowledge to identify early indicators of potential language disorders [1,24-27].

In the detection phase, the aim is to identify early whether a child is suspected of having some NDDs. To achieve adequate detection, tools such as screening forms and questionnaires allow information to be compared with previously defined risk indicators [20,28-32]. Some regions, such as Madrid (Spain), have early childhood educational and psychopedagogical guidance teams supporting nursery schools' detection phase. These teams engage in preventive actions and collaborate to detect developmental issues during the initial years of a child's life. If an educational case is identified, it is referred to the early intervention centers responsible for the assessment and intervention phase [33].

The assessment and intervention cover different aspects. During the assessment phase, it is necessary to analyze the family environment through interviews with relatives. In addition, hearing tests should be administered to eliminate potential associated psychological disorders. Evaluation of communication, playing with peers, comprehension, language production, and psychomotor skills are also essential. Finally, this information is correlated with medical history [6,14,15,20]. The intervention phase for DLD should be multidisciplinary, involving teachers, occupational therapists, speech therapists,

XSL•FO
RenderX

and psychologists skilled in children's language development techniques [7,24,27,34].

Gades: A Screening Tool for DLD

Gades functions as a clinical decision support system designed to grant expert knowledge regarding language disorders to preschool teachers. It offers a screening form to assess the language acquisition process during the early stages of a child's development. Gades allows the implementation of screening systems in natural environments for the child, thereby embracing the principles of pervasive therapy [35].

Pervasive or ubiquitous therapy is a therapeutic approach supervised by professionals that aims to promote engagement in activities focused on improving the individual's life within their natural environment through the support of information and communication technologies. It aims to be person centered instead of focusing on a specific clinical pathology. It especially seeks to identify how a problem impacts and manifests across various contexts [18,36]. This approach has the same phases of prevention, detection, assessment, and intervention as normative therapy. Moreover, using information and communication technologies ensures consistency and continuity of the therapeutic process within the person's natural setting [32,37-39].

For children's pervasive therapy, its application extends beyond the clinical center to environments like the child's school or home. Nursery schools are included in these settings. Being places where children spend much of their time, they are crucial for the early detection of DLD [24]. During this educational phase, teachers or educators and specialists in therapeutic pedagogy, such as nursery school language therapists (NSLTs), are typically involved. However, many nursery schools lack NSLTs, and even when they are available, fewer than half of the children suspected of having a language disorder are referred to or receive support from these specialists [1,7,17,23,27].

Martín Ruiz et al [35] developed and evaluated a previous version of Gades. The platform's cornerstone is a knowledge base (KB) that includes 106 milestones related to children's language acquisition based on age, defined and agreed upon by language experts. The previous version of this KB covered children aged 1 month to 72 months [35]. Building upon the findings of the previous experiment, increased instances of suspected DLD were identified in children aged 0 to 3 years. Consequently, in the current experiment, we focus on prioritizing the implementation of Gades during this developmental period. To extend this goal, Gades has undergone modifications and enhancements to its functionalities to gather additional pertinent information concerning language acquisition in children. These adaptations include integrating inquiries specific to the bilingual context of the children and their family backgrounds and evaluating results from the Gades across different language areas.

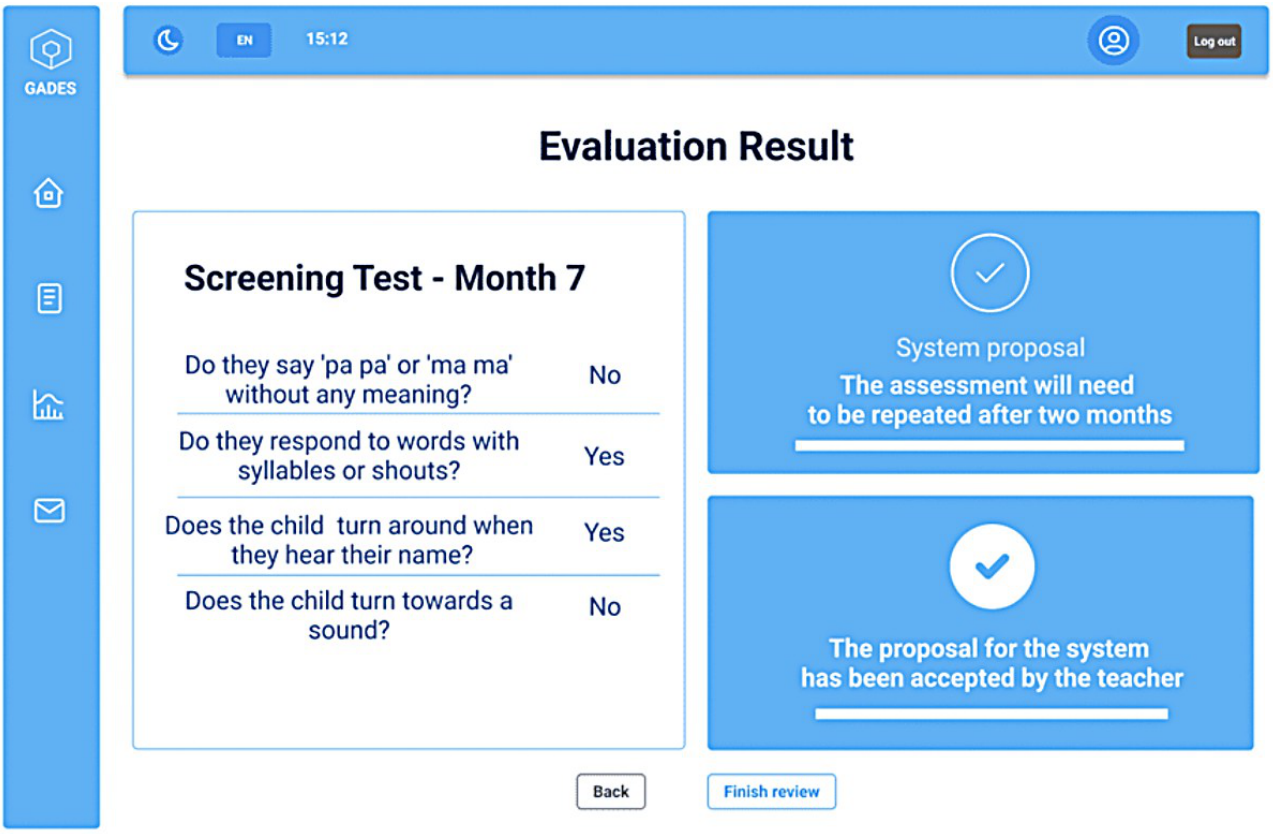
When using Gades, the corresponding milestones, tailored to the child's condition, are incorporated into a screening form through questions that education professionals use to assess a child's language development level. The milestones are categorized into 2 main types: "warning milestones," which prompt a reevaluation, and "alarm milestones," indicating the need for direct referral to health professionals. Gades' KB assesses 4 areas of speech and language development: sensory reception, speech perception, production, and pragmatic.

After the form adapted to the child is completed, Gades generates an evaluation result, including a suggestion for action if deviations from typical child development stages are detected. Therefore, it provides education professionals with expert knowledge on language acquisition, enhancing the early detection of DLD. Furthermore, this screening form acts as a psychopedagogical report for early intervention health professionals. Gades can aid in identifying the language development profile of a child suspected of having DLD through the assessment of processing-oriented and performance-based tasks.

Figure 1 depicts the graphical user interface of Gades, showcasing the evaluation results of consultation for a child aged 7 months. The platform displays the outcomes of the screening form, including the posed questions and the teacher's responses. Response options include "yes" if the child performs the specified action, "no" if not, and "do not know/no answer" for unknown answers. This form is completed based on the preschool teacher's observations of the child without direct interaction. After review, the system recommends that the assessment be repeated in 2 months, a suggestion that the overseeing teacher has validated. Gades also features additional tools for child registration, language evaluation, and user profile management.

XSL•FO
RenderX

Figure 1. Gades' evaluation result page for a child aged 7 months. In total, 4 questions were asked based on the child's age, and the system suggested repeating the evaluation in 2 months. The preschool teacher concurred with the system's recommendation.



Goal of This Study

For effective early detection and intervention of DLD, it is essential to offer sufficient services to these children to minimize the impact on their social, emotional, and educational development. Thus, language disorders often require a multidisciplinary approach among professionals from diverse fields such as education, speech-language therapy, and medicine to extend their reach into various settings [17,18,24]. Supporting training programs and systems for health and education professionals fosters collaboration. This collaboration facilitates the creation of adequate evaluation and intervention plans that meet the diverse needs of individuals with language difficulties, enhancing the quality of care and inclusivity [22,26,30,31,38,40].

Within this collaboration, the early detection of warning signs, which are potential indicators of language disorders in children, should be an integral part of the daily work of those interacting with children [1,7]. Supporting early childhood education professionals is crucial to ensure that they possess adequate knowledge and tools for preventing and detecting early language development issues in children aged 0 to 3 years because behaviors and warning signs related to language disorders can be identified from the age of 2 years onward [1,3,4,24]. It facilitates appropriate early diagnosis of DLD or SLD through subsequent assessment at early intervention centers. Moreover, given the absence of an NSLT in some schools, preschool teachers must have access to these resources.

Considering the higher incidence of suspected DLD observed in children aged 0 to 36 months in the previous experiment [17,23], this paper aims to emphasize the use of Gades during this period, targeting kindergarteners aged 6 to 36 months in a nursery school and an early childhood educational and psychopedagogical center. Gades has been adapted and expanded to gather more relevant information on language acquisition, including questions tailored to bilingual contexts and family backgrounds. Therefore, our primary goal is to evaluate Gades' ability to differentiate between typical and atypical language development for the target ages with these minor adjustments.

As the previous experiment consisted of a controlled environment, this paper also presents the relationship between preschool teachers and the Gades platform. Specifically, how Gades serves as a valuable tool for guiding nonexperts, such as preschool teachers, in conducting language observations. Therefore, it evaluates Gades' effectiveness in a real setting and its adherence to clinical criteria without direct supervision or intervention of language experts or technical staff. A potential challenge is also associated with integrating these technological tools into unsupervised educational settings to support therapeutic interventions [41].

Moreover, this screening process aims to identify the patterns indicative of a child with problems in language development, including SLD or DLD. Considering the updated *DSM-5* criteria, language development in children occurs within heterogeneous and dynamic contexts. Hence, it is imperative to use screening tools to identify significant patterns and milestones in the language development of a child with DLD, facilitating early detection and intervention. To accomplish this, the screening

XSL•FO
RenderX

tools must be intersectional, considering influential parameters, such as clinical conditions and the sociocultural and linguistic environment surrounding the child. Consequently, the final goal is to outline the framework for an adaptive screening model for language disorders that reflects the child's current developmental and social context. This approach ensures a comprehensive and nuanced understanding of each child's unique situation, enhancing the accuracy and effectiveness of early detection methods. Therefore, this study hypothesizes that the Gades platform can effectively support preschool teachers in the early detection of language development difficulties, including DLD and SLD, within natural educational settings, achieving high accuracy rates and providing actionable insights without requiring direct supervision from language disorder specialists.

Methods

Overview

This study used a prospective observational design to evaluate the updated Gades platform in real-world educational settings. The design involved monitoring children aged 6 to 36 months and their evaluators—preschool teachers and early childhood educators—over a defined period, collecting data as they naturally occurred during regular classroom activities. This approach was selected to ensure that the platform's performance could be assessed in authentic, uncontrolled conditions, reflecting its practical application. By focusing on how educators integrated Gades into their daily routines, the study aimed to capture unbiased insights into its usability and effectiveness in identifying early signs of language development difficulties.

Requirement Analysis

Individuals encounter varied communicative situations involving diverse entities and methods in everyday life. Current diagnoses and treatments for various disorders frequently overlook key characteristics, such as diversity, heterogeneity, and the role of socialization. The *DSM-5* and numerous studies highlight the necessity to consider interindividual differences in language acquisition early on. These sources suggest that assessments of language, speech, and communication should consider sociocultural aspects, linguistic context (including dialects), and socioeconomic status, as these aspects directly influence language development [1,12,16,17,19,21,31,32,42-44].

Bosch et al [21] demonstrated that 1.05% of the participants had communication disorders, with the most significant factors being foreign origin, genre, socioeconomic status, and age. Consequently, the diagnosis of language impairment should consider the individual's background, bilingual context, direct clinical observations in settings such as home or school, and results of standardized tests to assess the severity of the disorder.

The New Gades Platform

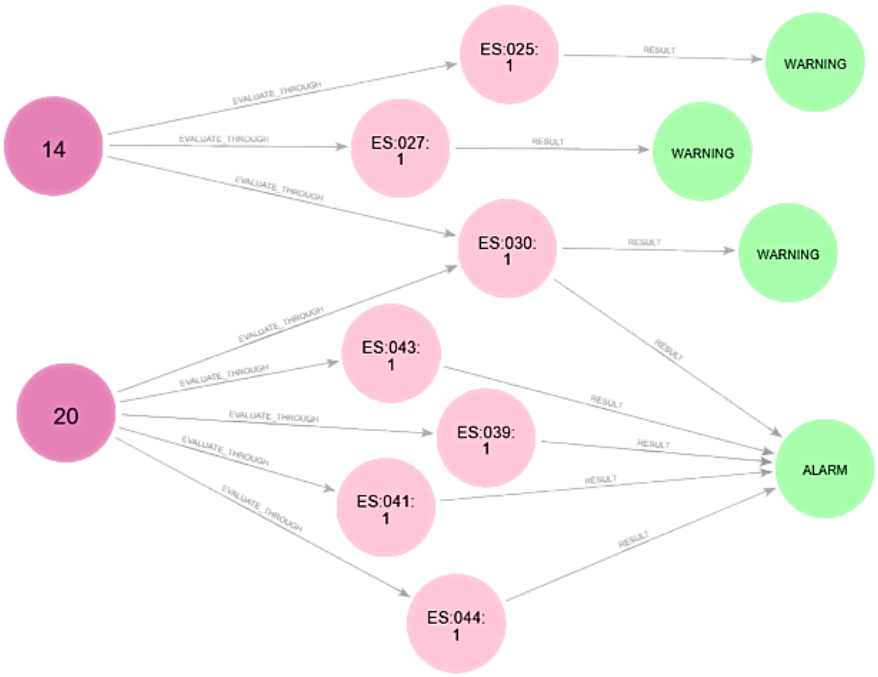
Building on the previous experiment with the Gades platform [35], we updated it with new technologies and features. These enhancements enable the collection of additional information from children to conduct more accurate and suitable assessments of a child's language development. This includes gathering details about the child's family background, linguistic environment, and medical history, all of which contribute to a more comprehensive understanding of each child's unique developmental context. Therefore, incorporating these minor adjustments delineates the structure of an adaptable screening model that considers the variability in language development. This methodology guarantees a thorough and nuanced comprehension of each child's individual circumstances, enhancing the precision and efficacy of techniques for identifying potential issues in language development.

Implementing Neo4j (Neo4j Inc), a graph database technology, is crucial in optimizing our Gades platform, specifically in enhancing the accuracy of language disorder detection. Neo4j allows for a more natural and flexible representation of complex data and interconnected relationships between them, which is essential for modeling the intricate interactions and dependencies observed in child language development [45]. In the context of Gades, Neo4j facilitates the dynamic integration and analysis of large volumes of data related to language development milestones, assessment responses, and individual characteristics of children. For example, the graph structure enables us to directly link children's responses to specific questions with language development patterns and automatically flag connections that indicate potential delays or deviations from typical development.

Therefore, the KB has also been migrated to Neo4j graph databases. As illustrated in Figure 2, each milestone is represented through a month (dark pink node) with several questions (pale pink nodes) and a final suggestion (green nodes). Thus, in the figure, a dark pink node is represented each month, corresponding to different questions designed to evaluate the child's language proficiency, depicted by pale pink nodes. Notably, some questions are linked to multiple months, reflecting the dynamic nature of language development assessment. Different action suggestions can be generated based on the answers to these questions, such as confirming typical development, recommending reassessment, or advising referral to early childhood care services (green nodes represent these). This graph-based approach facilitates a nuanced analysis of the child's language development, allowing for tailored interventions sensitive to each child's unique progression.

XSL•FO
RenderX

Figure 2. Gades' knowledge base in Neo4j: milestone months (dark pink node), questions identifiers (pale pink nodes), and suggestions (green nodes). Each month has various questions to evaluate language acquisition (for a child aged 14 months or 20 months) with suggestions such as typical development, repeat evaluation, or advising referral to early childhood care services.



With the updates to Gades, the system has addressed feedback from educators seeking more information on DLD. Previously, questions were categorized by language development areas and were only accessible to the system's technical staff. However, with the recent improvements, information about various skills is now accessible to the evaluation staff. This improvement means that by using Gades, preschool teachers and other professionals can gain deeper insights into the specific language development areas being assessed. Such transparency not only empowers educators with a better understanding of language development but also enables them to contribute more effectively to detecting warning signs.

Table 1 illustrates the classification and number of questions across different categories. The sensory reception category is administered in the earliest months of life, primarily to assess for hearing issues, and gradually transitions into language perception. This aspect entails the capacity to receive, process, and comprehend linguistic information through sensory channels, whether auditory or visual. Pragmatics includes conversational skills, coherence, cohesion within discourse, and the functional and social use of the language. Finally, language production encompasses phonology, morphosyntax, semantics, and nonverbal communication. This structured approach allows for a comprehensive assessment of a child's language development, targeting specific areas crucial for early detection and intervention of language disorders.

Table 1. Classification of questions according to language category [35].

Language development domains	Number of questions
Sensory reception	3
Language perception	32
Production	48
Pragmatics	25

Hypothesis and Experiment

We aimed to validate the reliability and effectiveness of the new Gades platform in a real setting without the direct supervision or intervention of language experts or technical staff. We explored the interaction between preschool teachers and the Gades platform, highlighting how Gades functions as a valuable tool for facilitating language observations to detect warning signs of language difficulties by nonexperts. The Gades platform is specifically designed to assist in detecting early language difficulties that may signal potential DLD. By focusing on these early warning signs, Gades supports the early detection and monitoring of language development challenges in real-world educational settings.

The Gades KB was established in a previous study [35] and is grounded in standardized tests commonly used to assess DLD. This ensures that Gades relies on validated methodologies to identify and interpret potential language difficulties.

The study's population sample consisted of 218 children aged 6 to 36 months from 2 educational centers in the city of Madrid: a preschool center and an early childhood educational and

XSL•FO
RenderX

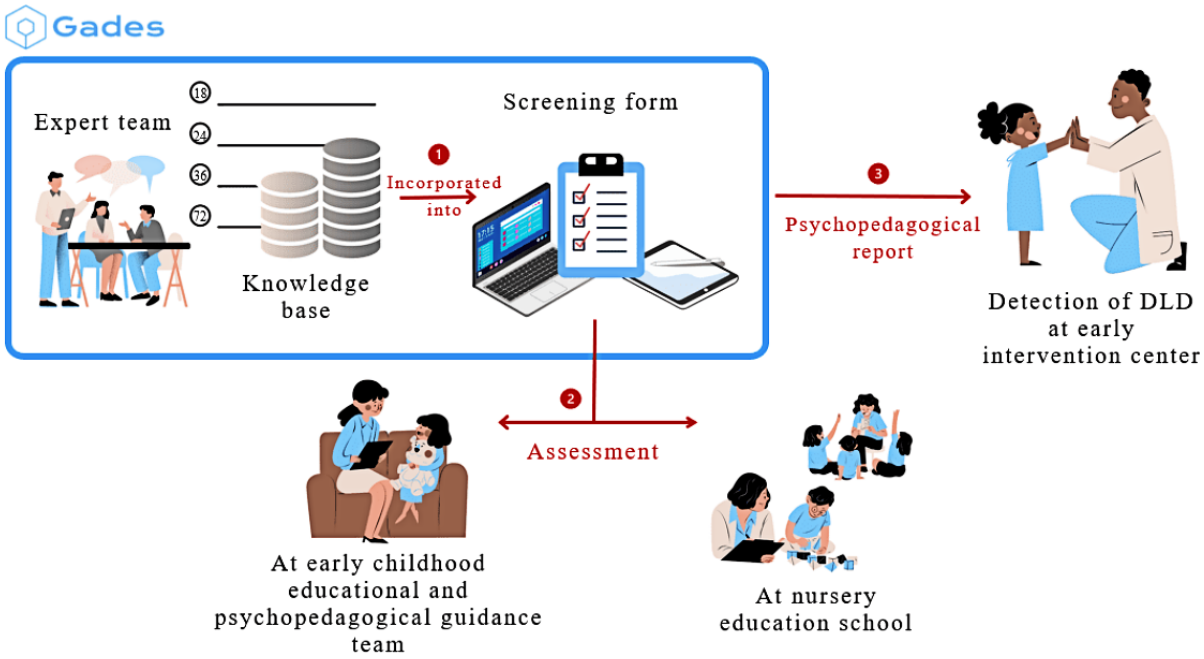
psychopedagogical center. All children enrolled in the first cycle of early childhood education, which aligns with this age range, participated in the study. This selection was based on recent evidence highlighting the critical importance of identifying language difficulties during this developmental period [1,3,4,24]. The screening data were collected from March to June 2023, providing an initial assessment of Gades' effectiveness in a real-world educational setting. Notably, access was limited to Spanish-speaking and some bilingual children.

The screening assessments for DLD involved the active participation of 24 nursery school educators and preschool teachers and a clinical professional from the early childhood center. Throughout this study, we will refer to them as evaluators. Hence, 2 categories of evaluators were involved in the experiment. The initial cohort consisted of professionals from early childhood centers, who, while not specialized in DLD, possessed adequate knowledge of NDDs. The second group comprised nursery school educators and preschool teachers, who may lack specialized language expertise but possess valuable insights into the children's competencies under their care. Before starting the experiment, the evaluators received training from qualified professionals affiliated with the Specific Language Disorder Association of Madrid. This training aimed to deepen their understanding and expertise regarding DLD and its identification with children aged 6 to 36 months. The training, delivered in a concise 2-hour format, was designed to ensure efficiency while providing comprehensive insights. By the session's conclusion, it was anticipated that the educators would have developed a more nuanced understanding of typical language milestones and a heightened capacity to discern early indications of language difficulties in children under their care.

During the evaluations, they operated without direct assistance from the training professionals or the Gades technical support team. Consequently, the experiment was conducted in an uncontrolled setting, relying exclusively on the evaluators' direct observation of the children's behavior for further evaluation through Gades. The evaluations were conducted following the guidelines recommended by Gades, aiming to ensure that the process was as effective and informative as possible within the constraints of the study's design.

Figure 3 outlines the Gades' framework, starting with an initial phase where a series of meetings were conducted to review and validate the KB established in the previous experiment. This phase involved experts in language disorders and the technical personnel from Gades to ensure the KB's relevance and accuracy. The panel of experts, who were also involved in constructing the KB as part of a previous study [35], recommended that including 3 to 6 questions per month would constitute an adequate measure for detecting language development issues. In the context of this study, the same experts conducted a review and confirmed the continued validity and applicability of the content of the KB. Once the KB was validated, the second phase, the screening process, began.

Figure 3. Gades' detailed framework. The process starts with expert validation of the knowledge base, integrated into a screening form with control and language questions. Evaluators assessed the children, and Gades provided recommendations, such as reassessment or referral, forming the basis of a psychopedagogical report. DLD: developmental language disorder.



The assessments were carried out by experienced female educators and professionals proficient in using digital platforms and information technologies. They used mobile devices and workplace computers to conduct evaluations, strategically scheduling them during the children's nap time around midday to minimize distractions. This approach involved each evaluator conducting assessments of the children under her care, which consisted of 2 parts. The first part included control questions related to the child's medical history and linguistic context, serving as a baseline for understanding each child's unique background. The second part consisted of the language difficulties screening questions. These questions were

XSL•FO
RenderX

meticulously developed in collaboration with experts to cover risk parameters for the early detection of DLD.

After the screening form was completed, the Gades system generated a recommendation to the evaluator, such as suggesting a repeat assessment in the upcoming months or indicating the need for a referral to an early intervention center. If a referral was recommended, nursery schools in Madrid forwarded the result to the expert through a psychopedagogical report. The designated expert, a language professional, was responsible for conducting a standardized evaluation tailored to the child's needs. Meanwhile, the educator played a supportive role by closely observing the child's progress in class and paying particular attention to areas of difficulty. Regular communication between schools and the expert was maintained with biweekly visits to ensure follow-up and support.

The application of the Gades platform was deliberately structured to be as nonintrusive as possible, allowing evaluators to focus on their primary duties while still contributing to the study. On average, each child underwent 1.24 (SD 0.49) assessments, indicating a consistent approach to the number of evaluations. The average duration of an assessment was 70 (SD 41.8) seconds per child, demonstrating the platform's efficiency in collecting data within a brief timeframe. To validate the Gades platform, both quantitative and qualitative analyses were conducted.

Quantitative validation focused on assessing the platform's accuracy and reliability. Descriptive statistics were used to summarize the duration and frequency of assessments. Agreement between the system's recommendations and evaluators' decisions was measured using concordance rates, calculated as the proportion of evaluations where the evaluators accepted Gades' recommendations. Statistical significance was set at $P<.05$. In addition, the platform's accuracy rate (97.41%) was determined by dividing the number of accepted suggestions (n=263) by the total number of evaluations (N=270). Instances of disagreement were analyzed qualitatively to identify patterns or recurring issues.

Qualitative validation focused on the integration of Gades into educational settings and its usability as perceived by preschool teachers. Evaluators were asked to provide feedback via a standardized questionnaire, following a validated system usability scale [41]. Responses were anonymized and collected outside the Gades platform to reduce potential bias. Evaluators also provided open-ended feedback to capture their experiences and insights, highlighting the platform's strengths and areas for improvement. The qualitative data were analyzed thematically to identify trends and recurring themes.

This combination of statistical and thematic analyses ensured a comprehensive validation of Gades. By comparing the platform's outcomes against educator feedback and established clinical criteria, the study aimed to validate Gades' ability to aid in the early detection of language development issues, such as SLD or DLD. Furthermore, comparisons between the initial version of Gades (2014) [35] and the current version highlighted improvements in adherence to clinical standards and system usability.

Ethical Considerations

This study, including the validation of the Gades platform and the experiment conducted through it, received ethics approval from the Ethics Committee of the Polytechnic University of Madrid in October 2023 GYGSDAALTD-MLMR-DATOS-20231010 and GYGSDAALTD-MLMR-HUMANOS-20231010). This approval ensures that both the use of the Gades platform and the specific experimental protocol comply with ethical standards, including data protection, participant rights, and research transparency.

For this study, an opt-out consent model was implemented. Participants were informed in advance about the study's objectives, the nature of the data collection, and the measures taken to ensure anonymity. They were given the opportunity to decline participation or withdraw their data at any point without providing a justification. Consent was implied by the participants' choice to continue with the evaluations, as no identifiable personal data were collected. Anonymized data were securely provided to the research team.

The professionals overseeing the evaluations maintained confidentiality in accordance with professional secrecy standards. Only the principal investigator had access to the evaluation results, which were securely stored in an encrypted database, accessible solely through authorized credentials.

Results

Sample Summary

The sample size for this study was calculated using a finite population correction formula to ensure adequate representation and statistical power. On the basis of demographic data from Madrid, the estimated population of children aged 6 to 36 months was approximately 150,000 [46]. With an assumed prevalence of DLD of 7%—a conservative estimate derived from epidemiological studies [17]—a 95% CI (Z=1.96), and a margin of error of 5%, the required sample size was calculated. The formula incorporates the finite population adjustment to account for the limited size of the target population, yielding a minimum required sample size of approximately 100 participants. This calculation also considered the variability in language development within the population.

The study's population sample comprised 218 Spanish-speaking children aged 6 to 36 months, more than double the calculated minimum sample size, ensuring robust statistical power and enhancing the reliability of the results. It featured a control group of 134 children from the nursery school (center 1) without any previous diagnosis or visible language difficulties and an experimental group of 84 children from the early childhood educational and psychopedagogical center (center 2), suspected of NDDs. Both groups were matched by gender to maintain homogeneity. Each child was assigned a random identifier for data anonymization and to facilitate subsequent analysis by the research team.

In Madrid's nursery schools, for each child, the evaluations were conducted by 2 assigned preschool teachers, also called an "educational pair." The concept of an "educational pair"

XSL•FO
RenderX

refers to 2 professionals sharing responsibilities and collaborating for the children's development and well-being, enabling the distribution of evaluation tasks between them. Therefore, in classrooms with children aged <12 months, each teacher evaluated an average of 4 (SD 0) children. For groups with children aged from 13 to 24 months, the responsibility increased to an average of 6.75 (SD 0.577) children per teacher. For those aged >24 months in the classroom, the average was 5.9 (SD 6.610) children per preschool teacher. Evaluations at the early childhood educational and psychopedagogical center were conducted by the teacher specializing in therapeutic pedagogy (NSLT).

Table 2 presents an overview of the study's sample, detailing the distribution of the children by age and gender: 45.4% (99/218) of the girls and 54.6% (119/218) of the boys participated. Within this group, 36% (36/99) of the girls and 40.3% (48/119) of the boys came from the early childhood center, initially identified with potential NDD concerns. Specifically, for participants aged >3 years, the study included 4 boys and 2 girls, all aged 37 months, except for 1 girl who was aged 38 months. The second year of the initial preschool cycle had the highest representation, with the largest number of children aged between 18 and 28 months, averaging 10 (SD 3.045) children per month. The average age in months per grade was distributed as follows: up to 12 months, the average age was 10.020 (SD 1.136) months; from 13 to 24 months, the average age was 20.331 (SD 1.108) months; and from 25 to 36 months, the average age was 31.578 (SD 1.094) months.

Table 2. Distribution of children by center, gender, and age range[a].

Grade	Up to 12 mo (n=26), n (%)	13-24 mo (n=102), n (%)	25-36 mo (n=84), n (%)
Children from center 1			
Female	10 (38.4)	22 (21.6)	31 (36.9)
Male	8 (30.7)	24 (23.5)	39 (46.4)
Children from center 2			
Female	5 (19.2)	22 (21.6)	9 (10.7)
Male	3 (11.5)	34 (33.3)	11 (13.1)

[a]Center 1 is the nursery school, and center 2 is a psychopedagogical center. The age ranges align with the preschool grades. The second year of the initial preschool cycle had the highest representation (102 children).

The screening form's initial section included questions regarding the child's medical history and linguistic context. Birth complication cases were more prevalent in center 1, particularly among male participants. Results indicated that 8.2% (11/134) of the children with no previous diagnosis of NDD experienced birth complications, including prenatal, perinatal, and preterm risks. 2.9% (4/134) were reported among females while among males, 5.2% (7/134) were observed. In center 2, only 2% (2/84) cases were reported in males and no cases in females. Notably, 3 children faced more than 2 of these risks, such as family history combined with perinatal risk. The mean gestational weeks remained relatively consistent across both centers and genders, with slightly more variation among girls. The average gestational age was 39.518 (SD 5.363) weeks, female children had an average of 38.984 (SD 2.36) weeks in center 1 and 40 (SD 0) weeks in center 2, while males had 39.563 (SD 1.29) weeks in center 1 and 39.791 weeks (SD 1.44) in center 2. A baby is considered premature if born before the 37th week of pregnancy. According to the study, only 3.7% (8/218) of the children were born preterm, highlighting specific early life factors that could influence developmental outcomes. Instances with a family history or bilingual cases were limited. Among the 218 cases studied, only 2 reported a family history of other NDDs such as autism spectrum disorder. Information regarding the bilingual context was exclusively available for children from center 1 (N=134). 7 (5.2%) cases were found among females, and 3 (2.2%) cases were recorded in males. It suggests that the language development impact of being raised in a bilingual environment could not be thoroughly assessed across the entire sample.

Overview of the Obtained Health Results

This section presents the results obtained from the evaluations that were conducted. A total of 270 assessments were completed, accounting for instances where children underwent multiple assessments. Retaining only the latest assessment conducted for each child (218 evaluations), there were 19.7% (43/218) children, which resulted in a suggestion to refer them to an early intervention center for a specific assessment by a professional. These types of suggestions were classified as alarms. That was the most severe system decision. Table 3 shows the results of the final assessments. The system has recommended repeating the remaining assessments in subsequent months to further evaluate and discard language issues. Suggestions that indicate repeat evaluation are classified as warning type.

XSL•FO
RenderX

Table 3. Distribution of Gades' final suggestionsa (N=218).

Description	Distribution, n (%)
Typical developmental	117 (53.7)
Refer to early childhood care	43 (19.7)
Repeat assessment (within 3 mo)	28 (12.8)
Repeat assessment (within 2 mo)	19 (8.7)
Repeat assessment (within 1 mo)	11 (5)

[a]More than half of the sample (n=117, 53.7%) were identified as typical language development. However, 19.7% (43/218) of assessments identified language difficulties as indicative of possible simple language delay or developmental language disorder.

Table 4 displays the distribution of assessment types by gender. It reveals that more alarm cases were identified in boys (29/218, 13.3%) compared to girls (14/218, 6.4%). In addition, both genders had a similar frequency of warning assessments and neurotypical results.

Table 4. Distribution of evaluations by suggestion type and gender. More alarm cases were identified in boys (29/218, 13.3%) compared with girls (14/218, 6.4%). Both genders had a similar rate of warning assessments and neurotypical results (N=218).

Gender and evaluation	Evaluations, n (%)
Female	
Alarm	14 (6.4)
Warning	29 (13.3)
No findings	56 (25.9)
Male	
Alarm	29 (13.3)
Warning	29 (13.3)
No findings	61 (27.9)

Gades identified more cases of alarm and warning in children already undergoing evaluation by specialists than from the center (alarm: 29.7%; warning: 32.1%), but we do not know whether these children had language disorders. Hence, among children already experiencing some difficulties, it identified that there may be some indication of language difficulties (ie, DLD or SLD).

Table 5 shows the distribution of evaluations that were classified as an alarm because it was detected that the assessed child did not meet the language acquisition milestones expected for their age. The incidence of language impairment from center 2 was comparable between boys and girls (15:10) and 30% (25/84) of the cases were detected as alarms. Meanwhile, for children without a suspicion of NDD coming from nursery school (center 1), difficulties in language and communication were detected in 13.4% (18/134) of the cases. However, a notable discrepancy was observed between boys and girls (14:4).

Table 5. Distributions of alarm evaluations by gender, month, and center[a].

Month	9	10	11	14	15	16	18	19	20	24	25	26	27	28	30	32	37	Total
Children center 1																		
Female	0	1	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	4
Male	0	1	0	1	0	0	4	0	0	1	1	2	1	0	1	1	1	14
Total	0	2	1	1	0	0	4	0	1	2	1	2	1	0	1	1	1	18
Children center 2																		
Female	1	0	0	0	2	0	1	3	0	0	0	1	1	1	0	0	0	10
Male	0	0	0	0	3	1	4	2	2	2	1	0	0	0	0	0	0	15
Total	1	0	0	0	5	1	5	5	2	2	1	1	1	1	0	0	0	25

[a]For children coming from center 1, language issues were detected in 13.4% (18/134) of the cases. The incidence of language issues from center 2 was comparable between male and female participants (15:10), with alarms detected in 30% (25/84) of the cases.

Notably, most of the suspected cases were detected in months 15, 18, 19 and 24. All these months correspond to the second year, encompassing children aged between 1 and 2 years. Upon thorough analysis of the evaluations during these months,

XSL•FO
RenderX

difficulties were discerned in the language production category for month 15. Half (23/40, 50%) of the children resulted in alarms. When compared to assessments from previous months, several observations emerged. First, most questions in this category were not been asked in previous months, as the child's age renders them too young to possess these competencies. Second, only 1 question, "first disyllables (one word in addition to dada and mama)?" was repeated in months 12 and 14, with half (8/14, 57%) of the assessed children answering negatively in both months. For month 15, all respondents answered negatively.

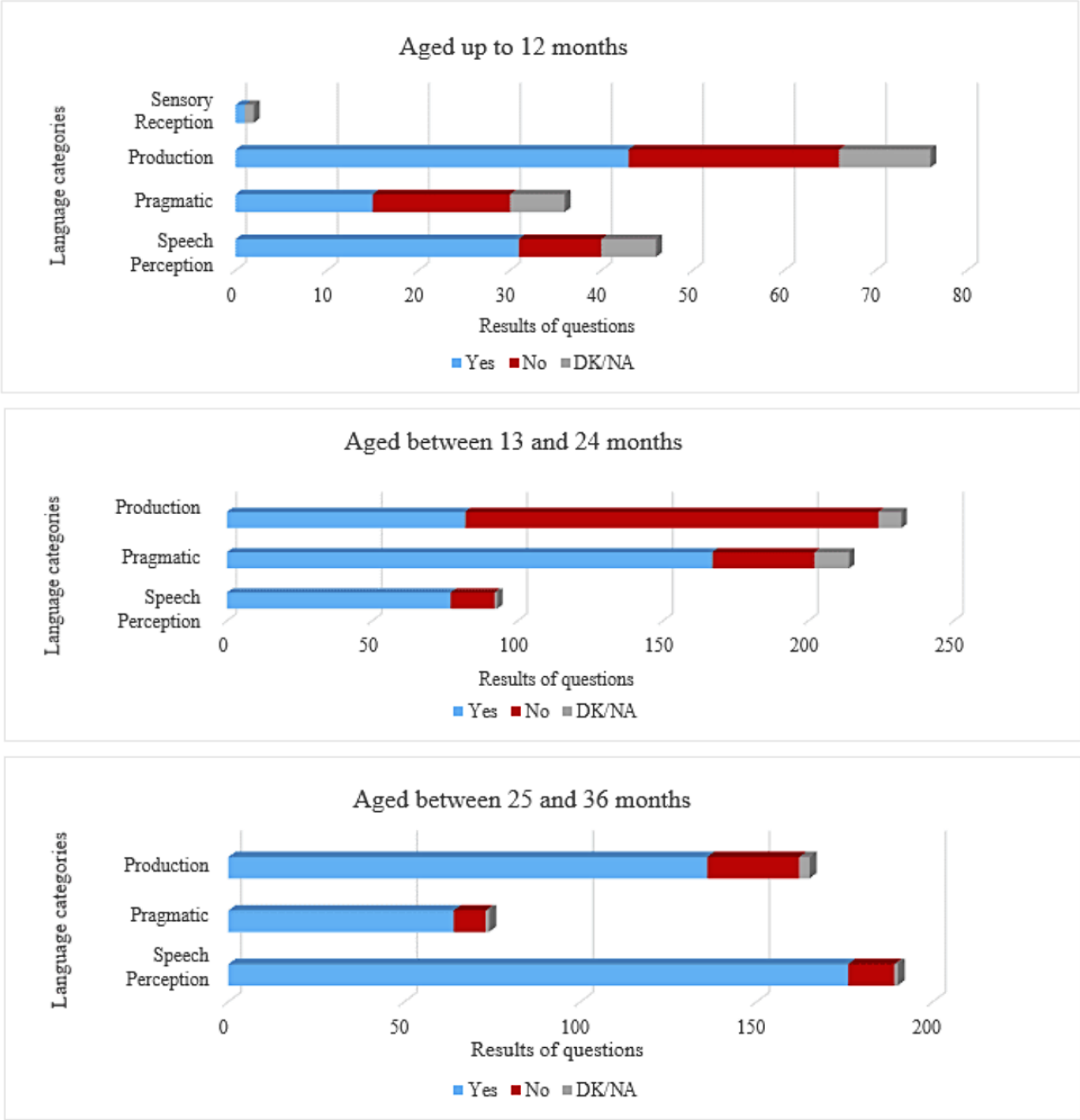
Month 18 had the highest number of detected suspected cases (9/11, 82%). During one of the meetings, educators expressed that they had observed more issues in language production, citing example, such as the child's inability to articulate certain words. In addition, similar questions asked in previous months also resulted in alarms. For instance, the one with the most error answers was "first bisyllables (two words in addition to dada and mama)?" In subsequent months, specifically month 19, identical pragmatics and language production questions were asked, and negative responses were received.

In month 24, 40% (4/10) of the evaluations raised alarms. Upon analyzing the questions, 3 were related to pragmatics and language comprehension, which were previously asked in earlier months. Most children satisfactorily answered these questions, indicating alignment with their current developmental stage. Therefore, the cases prompting alarms may signify potential instances of children with SLD or DLD.

Figure 4 depicts the results of questions categorized by language development areas. The sensory reception category was evaluated only for children aged up to 12 months to discard hearing issues. Children aged <12 months encountered more challenges in the production and pragmatics categories. Furthermore, there were no cases of hearing problems. The pragmatics category showed an equal balance between successes and failures. Conversely, children in the second year demonstrated notable difficulty in answering questions within the language production category, with a higher-than-average rate of incorrect responses. Other categories showed satisfactory performance.

XSL•FO
RenderX

Figure 4. Distribution of results based on language-category questions. Each question is categorized according to language category (refer to Table 1 for more information) and can be answered with a yes, no, or do not know/no answer (DK/NA). Children aged from 13 to 24 months demonstrated notable difficulty in production questions. Children aged >24 months exhibited positive responses across most categories.

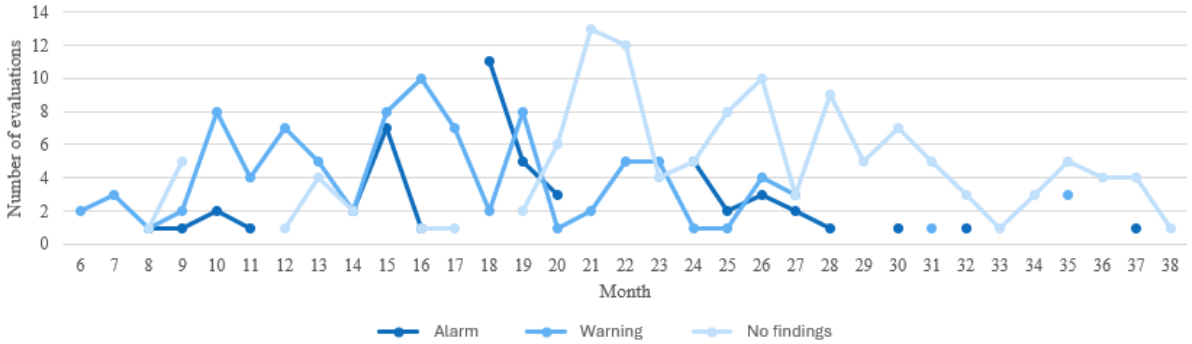


Finally, children aged >24 months exhibited predominantly positive responses across most categories. Hence, it becomes evident that children aged <24 months exhibit a higher frequency of errors, attributable to the inherent variability in language acquisition during this developmental stage. Consequently, there is a pressing need for greater flexibility in language milestones and extending their coverage across months.

As illustrated in Figure 5, this distribution reveals that instances of children with difficulty in language development are prevalent between months 15 and 27, gradually decreasing in the final months until 3 years of age.

XSL•FO
RenderX

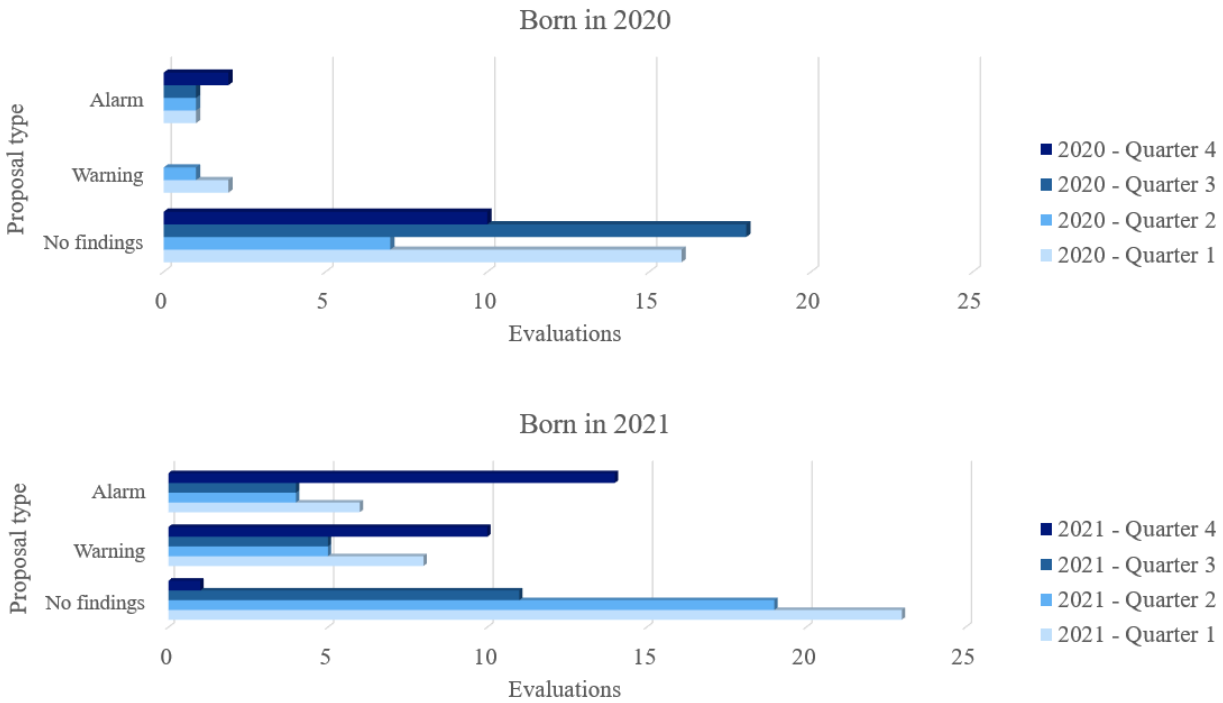
Figure 5. Distribution of suggestion type per month. Cases of children with language development issues were prevalent between months 15 and 27 (dark blue). The graphic reveals a gradual decline during the final months, with more typical development cases (pale blue).



In contrast, the evaluations were analyzed considering the control questions asked and the correlation between language development and the child's date of birth. In Madrid's nursery school classes, children were grouped based on their birth year. Consequently, at the beginning of the school year in September, a significant age gap existed between children born in the first quarter and those born in the last months. The latter group comprised the youngest in the class, resulting in a notable developmental difference between them.

The assessment outcomes for the youngest children in each class are presented in Figure 6. The youngest children were defined as those born from October to December 2020 for the age range of 25 to 36 months and from October to December 2021 for the age range of 13 to 24 months. This categorization was not applicable for children aged <12 months, as we did not have children from the last trimester. Hence, upon comparing the alarm for children born in 2020 and those born in 2021, it consistently appeared higher for those in the previous quarter. The youngest children tend to yield poorer results using Gades, as language development differs.

Figure 6. Assessment results by school quarter and suggestion type. For children born in 2020 (age between 25 and 36 months) alarm cases were more prevalent among children from quarter 4. However, these results are tempered by many more typical development cases at the same age. By contrast, for children born in 2021, the youngest children (quarter 4) tended to have alarm suggestions using the Gades platform.



The results of the assessment outcomes in relation to the control questions can be observed in Table 6. It is noteworthy that the 2 children who disclosed a family history of other NDD in the control questions were identified as cases with difficulties in language development (ie, SLD or DLD cases). Among cases with reported issues during pregnancy (ie, perinatal, prenatal, or premature), a notable percentage resulted in difficulties in language development (9/13, 69% including warnings). In the case of preterm children, only 8 (3.7%) of the 218 cases were identified as cases with difficulties in language development. Only one of these preterm cases, a girl, was assessed as an alarm. The 3 cases of preterm boys without previous NDD

XSL•FO
RenderX

demonstrated typical development. The remaining alarms in boys corresponded to cases of prenatal (3/218, 1.4%) and perinatal (4/218, 1.8%) risks. A similar scenario extended to bilingual cases, where only 1 child exhibited an alarm of language development difficulties.

Table 6. The child's evaluations were distributed by control information, gender, and suggestion typea (N=218).

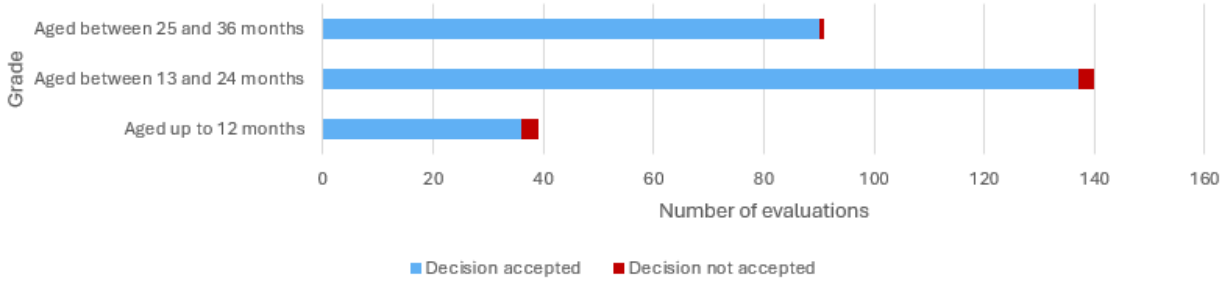
Gender and evaluation	Birth complications, n (%)		Family history, n (%)		Bilingual cases, n (%)	
	Center 1	Center 2	Center 1	Center 2	Center 1	Center 2
Female						
Alarm	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Warning	3 (1.4)	0 (0)	0 (0)	0 (0)	3 (1.4)	0 (0)
Typical development	1 (0.5)	0 (0)	0 (0)	0 (0)	4 (1.8)	0 (0)
Male						
Alarm	4 (1.8)	2 (0.9)	1 (0.5)	1 (0.5)	1 (0.5)	0 (0)
Warning	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Typical development	3 (1.4)	0 (0)	0 (0)	0 (0)	2 (0.9)	0 (0)

[a]Among children with birth complications, 69% (9/13) of them showed language development difficulties. Those with a family history of neurodevelopmental disorder also faced challenges, while only 1 bilingual child showed an alarm.

Gades' Performance

Gades' performance is assessed by the extent to which the system's decisions align with the perspectives of evaluators who support its recommendations. If there was disagreement with a recommendation, evaluators were prompted to justify dissent. Evaluators based their agreement or disagreement with the system's decision on their understanding of the child, considering factors such as age, family environment, and communicative intent. According to the results, Gades exhibited an accuracy rate of 97.4% (263/270). Out of 270 evaluations, only 7 (2.6%) suggestions were declined by evaluators. Figure 7 shows a higher disagreement rate with the system's decisions in children aged <24 months. Specifically, the accuracy rate was 92.3% up to 12 months, compared to 97.8% in the second grade (aged between 13 and 24 months) and 98.9% in the third grade (aged between 25 and 36 months).

Figure 7. Gades acceptance per grade. Out of 270 evaluations, only 7 proposals were declined (ie, 3 for evaluations of the first and second grades and 1 for evaluations in children aged >24 months). Even so, the acceptance of Gades remains above 90% in all courses.



Considering the total of 270 assessments conducted, including those where children were evaluated twice, the accuracy of Gades' decisions is depicted in Table 7. Among the suggestions provided by the Gades system, where evaluators disagreed (7/270, 2.6%), some suggested referring the child to early intervention for a specialized assessment (5/270, 1.9%).

Table 7. Gades' accuracy by the suggestion typea.

Suggestion description	Accepted suggestion, n (%)	
	No (n=7)	Yes (n=263)
Refer to early childhood care	5 (1.9)	45 (16.7)
Typical development	0 (0)	125 (46.3)
Repeat assessment (within 2 mo)	2 (0.74)	29 (10.7)
Repeat assessment (within 3 mo)	0 (0)	44 (16.3)
Repeat assessment (within 1 mo)	0 (0)	20 (7.4)

[a]Evaluators did not agree with Gades' suggestion in 2.6% (7/270) of evaluations. Most of these unaccepted suggestions were for the most severe cases identified by Gades ("refer to early childhood care"). Gades exhibit an accuracy rate of 97.41%.

XSL•FO
RenderX

When the evaluator disagreed with the Gades' decision, they were required to provide a comment stating the reason for nonacceptance. Thus, several reasons explain why the evaluators did not accept Gades' decision (Table 8). In 4 instances (ie, for months 10, 12, 15, and 18), they noted that the questions were not tailored to the child's developmental stage. In 2 cases from 18 months, the evaluator rejected the suggestion. Moreover, in case "The questions do not align with the child's developmental stage," the assessor disagreed with the suggestion and conducted an immediate reassessment. This subsequent assessment issued a warning (ie, to repeat the evaluation in the next months). In addition, in 2 cases involving 2 boys, the evaluators intended to continue observing the child's progress as they had identified other communicative skills. Finally, there was no disagreement for typical development cases.

Table 8. Evaluator's reasons for rejecting the final suggestion from Gadesa.

Reason	Children's gender	Evaluation month
Refer to early childhood care		
"We will monitor the situation closely. If necessary, we will consult the Early Assessment Team for guidance."	Male	26
"There are alternative forms of communication besides verbal language, and he demonstrates proficiency in them. We will observe his development and see how it progresses."	Male	18
"The questions do not align with the child's developmental stage."	Female	10
"The questions do not align with the child's developmental stage."	Male	10
"The questions do not align with the child's developmental stage."	Female	15
Repeat assessment (within 2 mo)		
"The questions do not align with the child's developmental stage."	Female	12
"We consider that it is not necessary at this stage."	Male	18

[a]The suggestions that were not accepted relate to "refer to early childhood care and repeat assessment (within 2 months)". Most of the reasons were that the assessment conducted by Gades was not aligned with the child's developmental stage.

Overview of the Obtained Functional Results

Numerous interviews were conducted with the evaluation staff, and they completed a web-based form assessing the usefulness and practicality of the Gades tool in uncontrolled educational settings, following the guidelines outlined in the referenced study [41]. None of the evaluators had previous experience using this specific platform. It was noted that preschool educators and professionals in early childhood educational and psychopedagogical centers found Gades highly acceptable. Most (18/24, 75%) of the evaluators found the system easy to use, although it required dedicated time. They experienced occasional technological issues such as student search processes or time sessions. A quarter felt previous technological knowledge was necessary.

The educators used an observational methodology during the school day to respond to the form's questions, considering it noninvasive and conducive to respectful support within an appropriate environment for the child. They suggested that implementing the Gades assessment during the first term of the school year, particularly during the adaptation period, could be challenging due to their limited familiarity with the child. In addition, they proposed conducting evaluations every 3 months for cases initially indicating "typical development" to ensure ongoing progress.

While control questions were valuable for accurate results from the Gades system, access to certain information, particularly related to birth difficulties, was limited for many children. Regarding the comprehension and appropriateness of the KB questions, educators encountered comprehension difficulties in only 2 out of the 108 questions, which they resolved by consulting the Gades' technical staff. They noted that some questions were unsuitable for the child's developmental stage. They highlighted the absence of questions about family context, communication dynamics, and disruptive behaviors in the second and third years.

Regarding recommendations, 75% (18/24) of the evaluators reported that they would suggest Gades to other nursery schools, emphasizing the need for previous DLD training to observe children effectively. They reported high knowledge acquisition and insight into their students' language development. For instance, they had observed that girls aged <12 months initiated communication earlier and more fluently.

Discussion

Principal Findings

This study's core findings underscore the Gades platform's effectiveness as a robust tool for screening language disorders in preschool environments, successfully identifying language difficulties in 19.7% (43/218) of the children assessed. This rate, with a notable prevalence in boys (29/218, 13.3%) compared to girls (14/218, 6.4%), is significantly higher than the traditional prevalence estimates of 2% to 7% for DLDs [8,17,21,23,43]. This discrepancy suggests that Gades may be particularly sensitive to early signs of language difficulties or may also capture cases of SLD. As a screening tool rather than a diagnostic system, Gades aims to identify children at higher risk for DLD or SLD, which might lead to elevated detection rates compared to studies focused strictly on diagnosis. Comprehensive assessments by specialists at early intervention

XSL•FO
RenderX

centers are required for a definitive diagnosis, highlighting the necessity of further evaluation following initial screening outcomes.

The study validated Gades' technical and functional capabilities in uncontrolled educational settings without direct intervention by an NSLT. The platform's utility in guiding nonexpert users, such as preschool teachers, through the screening process was particularly noteworthy. Teachers could use Gades effectively to observe and assess language development, reflecting its potential to empower educators with limited specialized training in language disorders.

Moreover, Gades demonstrated a high accuracy rate (97.41%) in aligning with the educators' assessments, underscoring its reliability and the robustness of its underlying algorithms and KB. The discrepancies between the system's recommendations and the educators' judgments were minimal, indicating a high level of agreement and trust in the platform's diagnostic suggestions.

The findings also illuminated specific age ranges (between 15 and 27 months) where language difficulties are most prevalent, reinforcing the importance of targeted early intervention during this critical developmental period. The ability of Gades to adapt its screening approach based on the child's age and specific educational context was a key factor in its effectiveness.

These promising results suggest that integrating Gades into regular preschool assessments could significantly enhance the early detection of language disorders, potentially leading to more timely and effective interventions.

Extending Detection to Natural Environments

This study highlights the importance of checking for developmental disorders similar to DLD in natural settings, such as preschools, where children naturally spend much time. DLD, similar to autism and ADHD, is among the most common developmental challenges but often goes unrecognized because its symptoms may resemble simple speech delays or overlap with other developmental disorders, leading to delayed diagnoses. Equipping educators, teachers, and school professionals with the right knowledge and support is crucial because they can spot early signs of these disorders. Early detection, significantly enhanced through collaboration with a multidisciplinary team, is key to improving a child's prospects for developing language skills effectively.

Applicability and Scalability

The Gades platform is designed to be versatile and scalable, making it a valuable tool for educators worldwide, especially in regions with limited access to specialized language disorder services. This flexibility is crucial for adapting to diverse linguistic and cultural environments across global educational systems. To effectively implement Gades in these diverse settings, developing training programs sensitive to local languages and educational standards is essential. These programs should equip educators with the skills to use the platform effectively and to recognize early signs of language disorders, even without specialized training.

Considering the technological limitations in different regions, Gades should be designed to operate in low-bandwidth environments and comply with local data protection regulations. This will make the platform accessible and trustworthy. Enhancing the platform's database to support multiple languages and reflect cultural specifics can make Gades a more inclusive tool that accurately reflects diverse child development patterns.

Practical implementation strategies include collaborating with local educational authorities to incorporate Gades into routine screening processes. This might involve tailoring the platform to meet local developmental benchmarks and language nuances. Partnering with local educational institutions can provide continuous support and feedback, helping to refine the platform.

Gades could offer versions that range from sophisticated, high-tech applications to more basic, web-accessible formats to accommodate varying technological capabilities. Establishing a community of practice among Gades users would encourage sharing best practices and provide essential peer support, enhancing the platform's overall effectiveness and user experience.

Limitations

One significant limitation of the Gades system is its reliance on the accuracy of the input provided by educators, who may not have specialized training in identifying language disorders. This dependency can introduce biases or inaccuracies in the screening process, as nonexpert interpretations of language development may vary widely, affecting the reliability of the outcomes. To mitigate these potential biases, it is essential to implement comprehensive training programs for educators using the platform. These programs should focus on familiarizing them with common language development milestones and the specific signs of language disorders. In addition, incorporating automated prompts and guidelines within the Gades system can help standardize recorded observations, reducing subjective interpretation errors.

Further refining the interface and feedback mechanisms of Gades can also enhance the accuracy of the data collected. For instance, the platform could include illustrative examples or short training videos on typical versus atypical language behaviors to guide educators' assessments before they input data. Regular updates and calibrations based on user feedback and new research in language development could further improve the system's precision and adaptability to real-world educational environments.

The study's limitations extend beyond the technical aspects of the Gades platform to include concerns about the sample size and the uncontrolled implementation settings. The small sample size may not adequately represent the broader population of preschool children, limiting the generalizability of the findings. In addition, while providing real-world insights into the use of Gades, the uncontrolled setting might introduce variability in how the platform is used across different environments, potentially influencing the consistency and applicability of the results. These factors could lead to higher variability in outcomes, which might not fully indicate the platform's efficacy under different or more controlled conditions.

XSL•FO
RenderX

Future Directions

Future research should include larger and more diverse populations to address the limitations mentioned to ensure that findings are robust and widely applicable. Controlled trials, where variables can be more tightly managed, would also help understand the platform's effectiveness across varied educational settings and populations. This would allow for more reliable validation of the platform's capabilities and identify specific conditions under which it performs best, guiding more targeted improvements.

To enhance Gades' effectiveness and reliability, it is imperative to refine the tool by expanding the KB to better differentiate between DLD and SLD at earlier stages. Adapting the system to align more closely with the actual ages of children rather than their academic year will provide more accurate assessments. Implementing educator training modules will equip them to effectively recognize early signs of language difficulties. Future versions of Gades should include a nuanced approach for assessing the youngest children in the class, considering their specific developmental stages and potential delays without prematurely recommending specialist evaluations. These improvements will ensure that Gades supports early detection and integrates smoothly into regular preschool assessments, potentially enhancing the effectiveness of interventions for language disorders.

Conclusions

This study confirms that DLD remains a commonly overlooked condition with significant variability and potential overlaps with other developmental disorders. Importantly, our findings reinforce the critical nature of age-specific screenings, with the most alarms for language difficulties arising in second-year grades, supporting the hypothesis that this period is crucial for detecting DLD risk factors. Despite the conservative and generic screening process, which identified language difficulties in 19% of cases—higher than the 7% prevalence typically noted in other studies—some of these instances may represent SLD. This suggests a need for refining language milestones within the Gades system to more effectively distinguish between DLD and SLD.

The qualitative outcomes from the Gades platform demonstrate its value in designing tailored educational programs and therapeutic interventions. By integrating such tools without the need for direct support from NSLT, Gades facilitates a deeper understanding of individual language development patterns. This insight enables educators, particularly those without specialized training in language disorders, to identify and address developmental challenges more effectively and earlier. The platform has been well-received in terms of usability and practicality, enhancing educators' ability to actively monitor and support language development.

Moreover, implementing Gades has provided substantial learning opportunities for preschool teachers. Educators with no previous expertise in DLD have gained significant knowledge and skills in informal screening processes, which has had a transformative impact on their ability to recognize and respond to language development issues. This empowerment of teachers underscores the potential of Gades to serve as a foundational tool in early education settings, enhancing early detection and intervention strategies.

To extend the utility of Gades beyond this study context, it is crucial to provide practical recommendations for its implementation in various educational settings. Tailoring the platform to accommodate different regional educational standards and linguistic backgrounds can maximize its effectiveness and reach. In addition, the successful deployment of Gades is contingent upon comprehensive training programs for educators. These programs should focus on enhancing understanding of language milestones, screening techniques, and the specific functionalities of the Gades platform. Effective training could be implemented through web-based modules, workshops, and ongoing support systems to ensure educators can proficiently use the tool.

In conclusion, the Gades platform represents a significant advancement in the field of educational technology for screening language disorders. By facilitating detailed observation and reporting of language development, Gades supports educators in their daily interactions with children and contributes to a broader strategy for addressing DLDs in early childhood education. The adoption of such tools, accompanied by adequate training and tailored implementation strategies, holds the promise of significantly improving outcomes for children at risk of DLD and SLD across diverse educational landscapes.

Acknowledgments

The authors would like to thank the team of educators and preschool teachers from Escuela Infantil Sabina and Patricia Calcerrada Muñoz from Equipo de Orientación Educativa y Psicopedagógica de Atención Temprana de Parla (Madrid, Spain) for their collaboration in performing the experiment presented here. In addition, the authors give special thanks to Esther González Dávila and Elena Jambrina Prieto from the Asociación del Trastorno Específico del Lenguaje de Madrid for their support in this study.

Data Availability

The datasets generated and analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

MD-P and MLM-R were responsible for conceptualization. MD-P and AMG-P were responsible for data curation, formal analysis, and methodology. SB-G was responsible for funding acquisition. MD-P, DLG, and MLM-R were responsible for investigation.

XSL•FO
RenderX

MLM-R was responsible for resources. MLM-R and IPDLC were responsible for project administration. MD-P and DLG were responsible for software. MLM-R and IPDLC were responsible for supervision. MD-P, MLM-R, and AMG-P were responsible for validation. MD-P and SB-G were responsible for visualization. MD-P, SB-G, IPDLC, and MLM-R were responsible for writing the original draft and review and editing.

Conflicts of Interest

None declared.

References

1. Kidd E, Donnelly S. Individual differences in first language acquisition. Annu Rev Linguist 2020 Jan 14;6(1):319-340. [doi: 10.1146/annurev-linguistics-011619-030326]
2. Kuhl PK. Early language acquisition: cracking the speech code. Nat Rev Neurosci 2004 Nov;5(11):831-843. [doi: 10.1038/nrn1533] [Medline: 15496861]
3. Nayeb L, Lagerberg D, Westerlund M, Sarkadi A, Lucas S, Eriksson M. Modifying a language screening tool for three-year-old children identified severe language disorders six months earlier. Acta Paediatr 2019 Sep 25;108(9):1642-1648. [doi: 10.1111/apa.14790] [Medline: 30896050]
4. Sansavini A, Favilla ME, Guasti MT, Marini A, Millepiedi S, Di Martino MV, et al. Developmental language disorder: early predictors, age for the diagnosis, and diagnostic tools. A scoping review. Brain Sci 2021 May 17;11(5):654 [FREE Full text] [doi: 10.3390/brainsci11050654] [Medline: 34067874]
5. Alás Rupérez A, Ramos Sánchez I, Machado Casas IS, Martín Fernández-Mayoralas D, Gortázar Díaz M, Aguilera Albesa S. Trastornos del lenguaje, del habla y de la comunicación. Conceptos, clasificación y clínica. Protoc Diagn Pediatr 2022;1:19-30 [FREE Full text]
6. Rowe ML. A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. Child Dev 2012 Jun 20;83(5):1762-1774 [FREE Full text] [doi: 10.1111/j.1467-8624.2012.01805.x] [Medline: 22716950]
7. Asorey MJ, Fernández MP. Dificultades de Aprendizaje y Trastornos del Desarrollo: Manual Didáctico. Madrid, Spain: Ediciones Pirámide; 2014.
8. Tomblin JB, Zhang X, Buckwalter P, O'Brien M. The stability of primary language disorder: four years after kindergarten diagnosis. J Speech Lang Hear Res 2003 Dec;46(6):1283-1296. [doi: 10.1044/1092-4388(2003/100)] [Medline: 14700355]
9. Tomblin JB, Records NL, Buckwalter P, Zhang X, Smith E, O'Brien M. Prevalence of specific language impairment in kindergarten children. J Speech Lang Hear Res 1997 Dec;40(6):1245-1260 [FREE Full text] [doi: 10.1044/jslhr.4006.1245] [Medline: 9430746]
10. Clegg J, Hollis C, Mawhood L, Rutter M. Developmental language disorders--a follow-up in later adult life. Cognitive, language and psychosocial outcomes. J Child Psychol Psychiatry 2005 Feb;46(2):128-149. [doi: 10.1111/j.1469-7610.2004.00342.x] [Medline: 15679523]
11. Fradley K, Oldfield J, Marshall J, Toseeb U. Early life risk factors for adolescent mental health difficulties for individuals at risk of developmental language disorder. JORSEN 2024 Mar 22;24(3):713-728. [doi: 10.1111/1471-3802.12654]
12. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. Washington, DC: American Psychiatric Publishing; 2013.
13. Lara EM. Trastorno Específico Del Lenguaje (TEL): Avances en El Estudio De Un Trastorno Invisible. Madrid, Spain: Ediciones Pirámide; 2016.
14. Bichay-Awadalla K, Qi CH, Bulotsky-Shearer RJ, Carta JJ. Bidirectional relationship between language skills and behavior problems in preschool children from low-income families. J Emot Behav Disord 2019 Jun 06;28(2):114-128. [doi: 10.1177/1063426619853535]
15. Romeo RR, Leonard JA, Robinson ST, West MR, Mackey AP, Rowe ML, et al. Beyond the 30-million-word gap: children's conversational exposure is associated with language-related brain function. Psychol Sci 2018 May 14;29(5):700-710 [FREE Full text] [doi: 10.1177/0956797617742725] [Medline: 29442613]
16. Kohnert K, Windsor J, Ebert KD. Primary or "specific" language impairment and children learning a second language. Brain Lang 2009 May;109(2-3):101-111 [FREE Full text] [doi: 10.1016/j.bandl.2008.01.009] [Medline: 18313136]
17. Norbury CF, Gooch D, Wray C, Baird G, Charman T, Simonoff E, et al. The impact of nonverbal ability on prevalence and clinical presentation of language disorder: evidence from a population study. J Child Psychol Psychiatry 2016 Nov;57(11):1247-1257 [FREE Full text] [doi: 10.1111/jcpp.12573] [Medline: 27184709]
18. López IJ, Förster J. Trastornos del neurodesarrollo: dónde estamos hoy y hacia dónde nos dirigimos. Revista Médica Clínica Las Condes 2022 Jul;33(4):367-378. [doi: 10.1016/j.rmclc.2022.06.004]
19. Francés L, Quintero J, Fernández A, Ruiz A, Caules J, Fillon G, et al. Current state of knowledge on the prevalence of neurodevelopmental disorders in childhood according to the DSM-5: a systematic review in accordance with the PRISMA criteria. Child Adolesc Psychiatry Ment Health 2022 Mar 31;16(1):27 [FREE Full text] [doi: 10.1186/s13034-022-00462-1] [Medline: 35361232]

XSL•FO
RenderX

20. Auza BA, Murata C, Peñaloza C. "Early detection of Spanish-speaking children with developmental language disorders: concurrent validity of a short questionnaire and a screening test". J Commun Disord 2023;104:106339. [doi: 10.1016/j.jcomdis.2023.106339] [Medline: 37247522]
21. Bosch R, Pagerols M, Rivas C, Sixto L, Bricollé L, Español-Martín G, et al. Neurodevelopmental disorders among Spanish school-age children: prevalence and sociodemographic correlates. Psychol Med 2021 Jan 13;52(14):3062-3072. [doi: 10.1017/s0033291720005115]
22. Carballal Mariño M, Gago Ageitos A, Ares Alvarez J, del Rio Garma M, García Cendón C, Goicoechea Castaño A, et al. Prevalence of neurodevelopmental, behavioural and learning disorders in paediatric primary care. Anales de Pediatría (English Edition) 2018 Sep;89(3):153-161. [doi: 10.1016/j.anpede.2017.10.005]
23. Calder SD, Brennan-Jones CG, Robinson M, Whitehouse A, Hill E. The prevalence of and potential risk factors for Developmental Language Disorder at 10 years in the Raine Study. J Paediatr Child Health 2022 Nov 03;58(11):2044-2050 [FREE Full text] [doi: 10.1111/jpc.16149] [Medline: 35922883]
24. Girolamo TM, Rice ML, Selin CM, Wang CJ. Teacher educational decision making for children with specific language impairment. Am J Speech Lang Pathol 2022 May 10;31(3):1221-1243. [doi: 10.1044/2021_ajslp-20-00366]
25. Bloder T, Eikerling M, Rinker T, Lorusso ML. Speech and language therapy service for multilingual children: attitudes and approaches across four European countries. Sustainability 2021 Nov 03;13(21):12143. [doi: 10.3390/su132112143]
26. Bishop DV, Snowling MJ, Thompson PA, Greenhalgh T, CATALISE Consortium. CATALISE: a multinational and multidisciplinary Delphi consensus study. Identifying language impairments in children. PLoS One 2016;11(7):e0158753 [FREE Full text] [doi: 10.1371/journal.pone.0158753] [Medline: 27392128]
27. Waine H, Bates S, Frizelle P, Oh TM. UK speech and language therapists' assessment of children's expressive language, and functional impairment and impact, following the CATALISE publications. Int J Lang Commun Disord 2023 Apr 27;58(5):1570-1587. [doi: 10.1111/1460-6984.12883] [Medline: 37102426]
28. Díaz MG, Sánchez IR, Casas IS, Rupérez AA, Fernández-Mayoralas DM, Aguilera Albesa S, et al. Trastornos del lenguaje, del habla y de la comunicación: desarrollo del lenguaje, signos de alerta, enfoque diagnóstico y tratamiento. Protoc Diagn ter Pediatr 2022;1:31-42 [FREE Full text] [doi: 10.14201/gredos.83174]
29. Gabaldón-Pérez AM, Martín-Ruiz ML, Díez-Muñoz F, Díez-Muñoz M, Máximo-Bocanegra N, Pau de la Cruz I. The potential of digital screening tools for childhood ADHD in school environments: a preliminary study. Healthcare 2023;11(20):2795. [doi: 10.3390/healthcare11202795]
30. Eikerling M, Lorusso M. A web-platform for DLD screening in Italian-German-speaking children: preliminary data on concurrent and predictive validity. Lingue e Linguaggi 2022;21(1):193-212 [FREE Full text] [doi: 10.1418/104455]
31. Eikerling MR, Bloder TS, Lorusso ML. A nonword repetition task discriminates typically developing Italian-German bilingual children from bilingual children with developmental language disorder: the role of language-specific and language-non-specific nonwords. Front Psychol 2022;13:826540 [FREE Full text] [doi: 10.3389/fpsyg.2022.826540] [Medline: 35719570]
32. Eikerling M, Andreoletti M, Secco M, Luculli B, Cha G, Castro S, et al. Remote screening for developmental language disorder in bilingual children: preliminary validation in Spanish–Italian speaking preschool children. Appl Sci 2023 Jan 21;13(3):1442. [doi: 10.3390/app13031442]
33. Equipos de orientación educativa y psicopedagógica. Comunidad de Madrid. URL: https://www.comunidad.madrid/servicios/educacion/equipos-orientacion-educativa-psicopedagogica#panel-135799 [accessed 2024-02-14]
34. Fiestas CE, Lugo-Neris MJ, Pratt AS, Peña ED, Bedore LM. Spanish language and literacy intervention for bilingual children at risk for developmental language disorder: a pilot study. Top Lang Disord 2021;41(4):309-321. [doi: 10.1097/TLD.0000000000000263]
35. Martín Ruiz ML, Valero Duboy M, Torcal Loriente C, Pau de la Cruz I. Evaluating a web-based clinical decision support system for language disorders screening in a nursery school. J Med Internet Res 2014 May 28;16(5):e139 [FREE Full text] [doi: 10.2196/jmir.3263] [Medline: 24870413]
36. Bowman R, Cooney O, Thieme A, Cowan BR, Doherty G. Pervasive therapy: designing conversation-based interfaces for ecological momentary intervention. IEEE Pervasive Comput 2022 Oct 1;21(4):55-63. [doi: 10.1109/mprv.2022.3209068]
37. Berrezueta-Guzman J, Martin-Ruiz ML, Pau I, Krusche S. A user-centered methodology approach for the development of robotic assistants for pervasive unsupervised occupational therapy. In: Proceedings of the 8th International Conference on Robotics and Artificial Intelligence. 2022 Presented at: ICRAI '22; November 18-20, 2022; Singapore, Singapore. [doi: 10.1145/3573910.3573925]
38. Robles-Bykbaev V, López-Nores M, García-Duque J, Pazos-Arias JJ, Arévalo-Lucero D. Evaluation of an expert system for the generation of speech and language therapy plans. JMIR Med Inform 2016 Jul 01;4(3):e23 [FREE Full text] [doi: 10.2196/medinform.5660] [Medline: 27370070]
39. Toki EI, Zakopoulou V, Tatsis G, Plachouras K, Siafaka V, Kosma EI, et al. A game-based smart system identifying developmental speech and language disorders in child communication: a protocol towards digital clinical diagnostic procedures. In: Proceedings of the Interactive Mobile Communication, Technologies and Learning. 2021 Presented at: IMCL 2021; November 4-5, 2021; Virtual Event. [doi: 10.1007/978-3-030-96296-8_50]

XSL•FO
RenderX

40. Bishop DV, Snowling MJ, Thompson PA, Greenhalgh T, the CATALISE-2 Consortium. Phase 2 of CATALISE: a multinational and multidisciplinary Delphi consensus study of problems with language development: terminology. J Child Psychol Psychiatry 2017 Oct;58(10):1068-1080 [FREE Full text] [doi: 10.1111/jcpp.12721] [Medline: 28369935]
41. Sevilla-Gonzalez MD, Moreno Loaeza L, Lazaro-Carrera LS, Bourguet Ramirez B, Vázquez Rodríguez A, Peralta-Pedrero ML, et al. Spanish version of the system usability scale for the assessment of electronic tools: development and validation. JMIR Hum Factors 2020 Dec 16;7(4):e21161 [FREE Full text] [doi: 10.2196/21161] [Medline: 33325828]
42. Fernald A, Marchman VA, Weisleder A. SES differences in language processing skill and vocabulary are evident at 18 months. Dev Sci 2013 Mar 08;16(2):234-248 [FREE Full text] [doi: 10.1111/desc.12019] [Medline: 23432833]
43. Nayeb L, Lagerberg D, Sarkadi A, Salameh EK, Eriksson M. Identifying language disorder in bilingual children aged 2.5 years requires screening in both languages. Acta Paediatr 2021 Jan;110(1):265-272. [doi: 10.1111/apa.15343] [Medline: 32869381]
44. O'Toole C, Gatt DR. Risk and Protective Environmental Factors for Early Bilingual Language Acquisition. Milton Park, UK: Taylor & Francis; 2019.
45. What is a graph database. Neo4j. URL: https://neo4j.com/docs/getting-started/graph-database/ [accessed 2024-02-14]
46. Población residente por fecha, sexo y generación (edad a 31 de diciembre). Instituto Nacional de Estadística. URL: https://ine.es/jaxiT3/Datos.htm?t=56941 [accessed 2024-12-14]

Abbreviations

ADHD: attention-deficit/hyperactivity disorder
DLD: developmental language disorder
DSM-5: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
KB: knowledge base
NDD: neurodevelopmental disorder
NSLT: nursery school language therapist
SLD: simple language delay

Edited by C Jacob; submitted 10.05.24; peer-reviewed by A Cruz-Santos, MÁV Duboy, D Casado-Mansilla, R Melguizo, R Soto-Varela; comments to author 18.11.24; revised version received 20.11.24; accepted 20.01.25; published 14.03.25.

Please cite as:
Dolón-Poza M, Gabaldón-Pérez AM, Berrezueta-Guzman S, López Gracia D, Martín-Ruiz ML, Pau De La Cruz I
Enhancing Early Language Disorder Detection in Preschools: Evaluation and Future Directions for the Gades Platform
JMIR Hum Factors 2025;12:e60424
URL: https://humanfactors.jmir.org/2025/1/e60424
doi:10.2196/60424
PMID:

©María Dolón-Poza, Ana-Marta Gabaldón-Pérez, Santiago Berrezueta-Guzman, David López Gracia, María-Luisa Martín-Ruiz, Iván Pau De La Cruz. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 14.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Developing a Novel Mobile App to Support HIV Testing and Pre-Exposure Prophylaxis Uptake Among Men Who Have Sex With Men: Formative and Technical Pilot Study

Albert Y Liu[1,2], MPH, MD; Thiago S Torres[3], PhD; Cat-Dancing Alleyne[1], BA; Janie Vinson[1], BA; Kelly Bojan[4,5], C-FNP, APRN, DNP; Pedro Alonso Serrano[4,6], MPH; Temitope Oyedele[4], MD; Amayvis Garcia[7], MA; Elizabeth Enriquez-Bruce[7], MD; Patricia Emmanuel[7], MD; Jeb Jones[8], MPH, MS, PhD; Kathryn E Muessig[9], PhD; Susan P Buchbinder[1,2], MD; Patrick Sullivan[8], DVM, PhD; Lisa B Hightow-Weidman[9], MPH, MD; Hyman Scott[1,2], MPH, MD

[1]Bridge HIV, San Francisco Department of Public Health, San Francisco, CA, United States

[2]Department of Medicine, University of California San Francisco, San Francisco, CA, United States

[3]Instituto Nacional de Infectologia Evandro Chagas, Fundação Oswaldo Cruz (INI-Fiocruz), Rio de Janeiro, Brazil

[4]The Ruth M. Rothstein CORE Center, John H. Stroger Jr. Hospital, Chicago, IL, United States

[5]Center for Dissemination and Implementation Science, University of Illinois, Chicago, IL, United States

[6]Department of Medical Social Sciences, Northwestern University, Chicago, IL, United States

[7]Department of Pediatrics, Morsani College of Medicine, University of South Florida, Tampa, FL, United States

[8]Department of Epidemiology, Emory University, Atlanta, GA, United States

[9]College of Nursing, Florida State University, Tallahassee, FL, United States

Corresponding Author:
Albert Y Liu, MPH, MD
Bridge HIV
San Francisco Department of Public Health
25 Van Ness Avenue, Suite 100
San Francisco, CA
United States
Phone: 1 6282177408
Fax: 1 415 431 7029
Email: albert.liu@sfdph.org

Abstract

Background: Young sexual minority men (YSMM) are disproportionately impacted by HIV in the United States. HIV or sexually transmitted infection (STI) testing rates and pre-exposure prophylaxis (PrEP) uptake are low in this priority population. Novel strategies are needed to increase access to HIV and STI prevention services among YSMM.

Objective: This study aims to describe the development and assess the feasibility and acceptability of LYNX, a mobile app to increase HIV testing and PrEP uptake among YSMM.

Methods: Informed by the Information-Motivation-Behavioral Skills model, the LYNX app was refined through 4 iterative focus groups in 2 US cities among YSMM aged 15 to 24 years. The LYNX app includes SexPro, an innovative tool that provides a personalized sexual health protection score, a sex diary to track sexual partners, HIV and STI testing information and reminders, access to home HIV and STI test kits, and geospatial-based testing and PrEP clinic site information. The refined app was then tested for feasibility and acceptability in a 2-month technical pilot. Baseline and 2-month follow-up assessments and exit interviews were completed. Self-reported app acceptability and use based on paradata were reported.

Results: In iterative focus groups among 30 participants (age: mean 20, SD 3 years; Black: 12/30, 40%; Hispanic or Latinx: 13/30, 43%), the app's design was well-received. Participants recommended providing information on how the SexPro score was calculated and how they could improve their score, changes to the language in the sex diary tailored for YSMM, providing a chat feature to facilitate communication between staff and app users, and gamification features to increase overall youth engagement with the app. These recommendations were incorporated into the app. In the technical pilot among 17 participants (age: mean 22.4, SD 1.6 years; Black: 4/17, 24%; Hispanic or Latinx: 8/17, 47%), the mean system usability score was 70 out of 100, falling

XSL•FO
RenderX

in the “good” range. Use of the app was high over the 2-month pilot (app opened an average of 8.5, SD 8.0 times with an average duration of 3.8, SD 3.2 min/session), indicating good feasibility. The most commonly used features included the testing feature (n=15, 100%), activity calendar (n=14, 93%), and diary (n=13, 86%). Overall, 11 (79%) participants were likely to continue using LYNX, and 10 (71%) participants were likely to recommend it to a friend. In exit interviews, there was a high level of acceptability of the content, interface, and features of the LYNX app.

Conclusions: Following a user-centered design approach, we tailored the LYNX app to increase HIV and STI testing and PrEP uptake among YSMM in the United States. Our positive findings support further testing of this mobile health tool in an upcoming effectiveness trial in broader youth populations.

Trial Registration: ClinicalTrials.gov NCT03177512; https://clinicaltrials.gov/study/NCT03177512

International Registered Report Identifier (IRRID): RR2-10.2196/10659

(JMIR Hum Factors 2025;12:e62830) doi:10.2196/62830

KEYWORDS

HIV testing; sexually transmitted infection testing; pre-exposure prophylaxis; youth; men who have sex with men; sexual minority men; mobile health; HIV prevention; STI

Introduction

Background

Young sexual minority men (YSMM) are disproportionately impacted by HIV in the United States. In 2021, 19% of the estimated 32,100 new HIV cases in the United States were among young people aged 13-24 years, with YSMM accounting for 82% of new HIV diagnoses in this age group [1]. YSMM of color are particularly impacted by HIV, with Black and Latino sexual minority men (SMM) accounting for 53% and 28% of infections among YSMM, respectively, in 2021 [1]. There is an urgent need to expand access to effective HIV prevention strategies in this priority population.

HIV Testing and Pre-Exposure Prophylaxis Use Among YSMM

HIV testing is the gateway to the HIV treatment and prevention continua, and it affords an important opportunity to offer timely treatment for YSMM living with HIV and for linkage to effective preventive strategies for those who test HIV negative. Although the Centers for Disease Control and Prevention recommends at least yearly HIV testing for SMM, in a recent national web-based survey, only 40% of YSMM reported testing in the past year, and 45% had never tested in their lifetime [2]. Young people are the least likely to be aware of their HIV status, with only 56% of US youth living with HIV aware of their diagnosis in 2022 [3]. Despite YSMM having among the highest annual sexually transmitted infection (STI) rate among any age group [4], STI testing rates are also low [5], with only a third of YSMM reporting STI testing in the last year [6]. Low perceived risk, lack of symptoms, limited knowledge of STIs, and lower access to health care providers have been identified as barriers to HIV and STI testing [7-9].

Numerous studies have demonstrated the effectiveness of pre-exposure prophylaxis (PrEP) for HIV prevention [10-12], now one of the pillars of the Ending the HIV Epidemic Initiative in the United States; however, uptake of PrEP has been low among YSMM. According to national prescription data, youth aged 15 to 24 years had the greatest unmet need for PrEP among all age groups, with a PrEP-to-need ratio of 9, indicating that for every person in this age group diagnosed with HIV, there were only 9 people using PrEP, compared to a PrEP-to-need ratio of 17 for people aged 35 to 44 years [13]. Demonstration projects also highlight challenges with PrEP uptake. In the Adolescent Medicine Trials Network for HIV Interventions (ATN), 110 study of YSMM aged 18 to 22 years, PrEP uptake was only 16%, and PrEP adherence was lower among Black YSMM and declined overall during follow-up, particularly with less frequent visits [14]. A number of factors contribute to low PrEP uptake, including low awareness and knowledge about PrEP, limited access to PrEP providers and clinics, deficits in self-perceived risk, HIV- and PrEP-related stigma, and medical mistrust, especially among YSMM of color, and highlight the importance of innovative strategies to increase testing and PrEP uptake tailored for YSMM [15-17].

Potential of Mobile Health Technologies

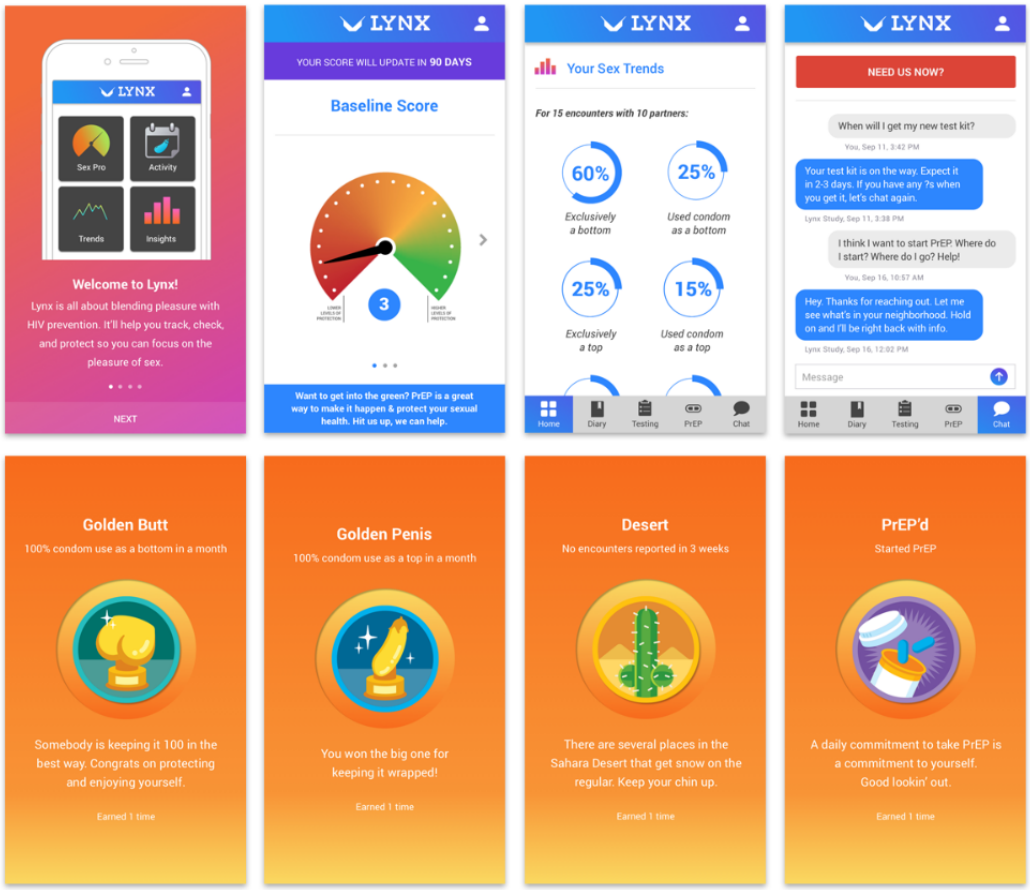
Mobile technologies have enormous potential to reach and engage YSMM in HIV prevention [18-22]. Mobile phone use is nearly ubiquitous among youth, and youth are more likely to use their mobile devices for more activities, such as downloading mobile apps, internet access, social media, and accessing health information, compared to people in older age groups [23,24]. The expansion of smartphones has increased the possibilities of dynamic, mobile phone–based HIV prevention interventions. Particularly, a mobile phone app designed to increase HIV testing and PrEP uptake could be highly scalable and has great potential to expand access to and uptake of prevention services among youth.

Using the Information-Motivation-Behavioral Skills (IMB) model [25], we previously developed the LYNX app, a highly interactive mobile app to promote accurate risk perception and increase HIV and STI testing and linkage to care, which has been previously described [26]. Components of the app are shown in Figure 1 and include (1) SexPro (Sexual Health Promotion), an innovative tool that provides a personalized HIV risk score displayed on a “speedometer” (1-20), with a higher score representing higher levels of protection [27]; (2) a sex diary to help users track their sexual partners and encounters, which are then displayed in a calendar; (3) HIV and STI testing information and reminders; and (4) ability to order home HIV

XSL•FO
RenderX

point-of-care self-test kits (Oraquick) and home STI testing collection kits (self-collected specimens mailed to a central laboratory for testing) through the app and access the geospatial-based testing site and linkage to HIV care information. To protect participant privacy, the app is accessed through a secure password-protected log-in, and information collected through the app is encrypted and stored in a secure environment that supports Health Insurance Portability and Accountability Act (HIPAA) compliance.

Figure 1. LYNX app screenshots. PrEP: pre-exposure prophylaxis.



In this study, we describe the results of a 2-phase study to develop and refine the LYNX app tailored for YSMM through iterative focus groups and an open technical pilot. This study was conducted within the National Institutes of Health's ATN as part of the University of North Carolina and Emory Center for innovative technology (iTech), which had the overall goal to develop technology-focused interventions addressing the HIV prevention and care continuum for youth [28].

Methods

Study Population

Eligible participants included cisgender men who were aged 15 to 24 years, self-reported being HIV negative or HIV status unknown at screening (technical pilot only), not currently taking PrEP (technical pilot, as participants are testing an app, which assists with PrEP uptake), fluent in English, owned an iOS or Android mobile phone, and had self-reported evidence of being at risk for HIV acquisition (detailed risk criteria are described in the study by Liu et al [26]). Focus group participants were able to participate in the technical pilot if they remained eligible.

Through the National Institutes of Health's ATN iTech U19, participants were recruited across 2 sites—Chicago, Illinois (study site: CORE Center), and Tampa, Florida (study site: University of South Florida). Participants for focus groups and the technical pilot were recruited through a variety of strategies, including web-based and social media postings (eg, Craigslist, Grindr, and Facebook advertisements); distributing posters, flyers, and palm cards about the study; direct outreach at local venues frequented by YSMM (eg, community-based organizations, schools, social clubs, health fairs, and balls); clinic-based recruitment; and participants in prior studies. To ensure inclusion of youth most impacted by HIV, we oversampled YSMM of color, with a goal of enrolling two-thirds of the cohort as YSMM of color and at least one-quarter as Black YSMM. Oversampling was accomplished by focusing recruitment efforts on venues more likely to be frequented by YSMM of color.

Iterative Focus Groups for App Refinement

After an initial prototype of the LYNX app was developed by the study team, we conducted 2 rounds of formative work through a series of 4 focus groups and individual interviews with 30 YSMM across our 2 iTech sites, which informed iterative development of the LYNX app. YSMM participated in 1 of the 4 groups, which each met 2 times. Participants unable to be scheduled for a focus group were offered an individual

XSL•FO
RenderX

interview, with the same content covered in both focus groups and interviews. Groups were divided by age strata (15-18 years, 19-24 years, etc), and focus groups were conducted in a private room at each site by a research staff member trained in qualitative methods and group facilitation. All participants completed a written consent or assent process before initiation of the focus group. During the focus group discussion, participants were shown how to download and install the app and were given a demonstration of key features of the LYNX app. The home HIV and STI testing process was also explained, and participants were shown a test kit. New app components were developed to increase PrEP uptake and included a PrEP Facts page with information about PrEP, a map showing local PrEP clinics in an area, PrEP video testimonials of youth sharing their experiences about taking PrEP, and a feature of SexPro showing how PrEP can improve one's score. Participants were asked to give feedback on each component of the app and suggestions for improvement. At the end of the focus group, participants completed a brief web-based survey to assess demographic characteristics and use of technology to help contextualize the focus group discussion.

Members of the iTech Analytic Core reviewed transcripts and identified emergent themes. We conducted thematic analysis using a deductive approach to code data for app acceptability overall and for each component, with a focus on functionality, appearance, usability, and privacy and confidentiality. Key findings are described in the text using representative quotes and were used to refine the LYNX app before initiation of the technical pilot.

Changes to the LYNX App

On the basis of feedback from the focus groups and interviews, we made a list of recommended changes to the app. The study team then prioritized changes based on the perceived impact of the changes on the usability and acceptability of LYNX, the likelihood of increasing HIV and STI testing and PrEP uptake, and the technical feasibility of the modifications. We then worked with our app developer to implement these revisions to the LYNX app.

Technical Pilot to Assess Feasibility and Acceptability of the LYNX App

After refining the LYNX app based on our formative work, we conducted a 2-month technical pilot among 17 participants to assess the feasibility and acceptability of the LYNX app and to elicit additional feedback for app improvement. Eligible participants who enrolled in the study completed a web-based survey on sociodemographics, risk behaviors, and use of technology. Study staff assisted participants in downloading the LYNX app and were shown how to use key features of the app using an onboarding guide. Participants were encouraged to use all components of the app over the next 2 months, including ordering and using the HIV and STI test kit.

Upon completion of the technical pilot at month 2, participants were asked to complete a brief web-based survey to assess preliminary usability and acceptability of the app using the System Usability Scale (SUS), a validated tool assessing various domains of the app with demonstrated high internal consistency across several studies [29]. The SUS is a 10-item scale that measures a system's usability on a scale of 0 to 100. Feasibility was assessed using app analytics to determine whether the app was used, how many times it was used, and which components were used most and least frequently. Feasibility was a priori defined as having at least 60% of individuals open the LYNX app at least once after introduction to the app by research staff.

Each participant also completed an exit interview to provide feedback on functionality, technical performance, errors and bugs encountered, overall experiences using the app, feasibility and acceptability of methods to confirm HIV testing results, and feedback for further refinement. Exit interviews were conducted remotely with qualitative-trained study staff from the iTech Analytic Core using a HIPAA-compliant videoconferencing technology. Debriefing reports capturing the quality and key findings were completed after each interview. Interviews were also transcribed and analyzed using the same approach outlined earlier.

Ethical Considerations

The study procedures were reviewed and approved by the University of North Carolina Institutional Review Board (IRB) as a single IRB-of-Record (approval number: 17-0170). IRB authorization agreements with all participating research entities were enacted. Written informed consent was obtained before conducting any study activities. All study data were identified by a coded number only to maintain participant confidentiality. The LYNX protocol is registered at ClinicalTrials.gov (NCT03177512). For iterative focus groups, participants received US $50-$60 for completing each focus group or interview, depending on local site standards. For the technical pilot, participants received US $50-$60 for completing the baseline visit, US $25-$30 for completing the 2-month survey, and US $50-$60 for completing the final exit interview. Additionally, at one site, a $30 gas card for round trip transportation was provided. Incentive amounts varied by site and were based on local site standards.

Results

Iterative Focus Groups for App Refinement

Characteristics of the 30 focus group participants, 18 (60%) in Chicago and 12 (40%) in Tampa, are shown in Table 1. Mean age was 20 (SD 3.0) years; 13 (43%) self-identified as Hispanic or Latinx, 18 (60%) completed high school or lower, and 13 (43%) were currently in school. A substantial proportion reported having health insurance (n=21, 70%), ever tested for HIV (n=26, 87%), tested for other STIs in the past year (n=22, 73%), and ever taken PrEP (n=16, 53%). Most participants reported using iOS (n=20, 67%), receiving health information through a website or an app on a mobile phone (n=23, 77%), and that they thought receiving health information through a mobile app would be of interest to YSMM (n=27, 90%).

XSL•FO
RenderX

Table 1. Characteristics of focus group participants for the development of an HIV prevention mobile app from Chicago, Illinois, and Tampa, Florida (N=30).

Characteristics	Values
City, n (%)	
Chicago	18 (60)
Tampa	12 (40)
Age (y), mean (SD)	20 (3.0)
Sexual orientation, n (%)	
Gay	21 (70)
Bisexual	9 (30)
Race , n (%)	
Asian	3 (10)
Black	12 (40)
Hispanic or Latinx	13 (43)
Mixed race	4 (13)
White	6 (20)
No answer	1 (3)
Schooling, n (%)	
Lower than high school	7 (23)
Completed high school	11 (37)
Some college	8 (27)
College completed	4 (13)
Currently in school	13 (43)
Health insurance, n (%)	21 (70)
Insurance coverage through parent or guardian, n (%)	10 (33)
Primary partner, n (%)	10 (33)
Number of male partners (past 6 months), mean (SD)	5 (6.0)
Number of male sexual partners met through internet (in the past 6 months), mean (SD)	5 (8.0)
HIV testing, n (%)	
Never	4 (13)
Past 3 months	18 (60)
>3 months	8 (27)
Tested for other STIs[a] (in the past year), n (%)	22 (73)
PrEP[b] awareness before the study, n (%)	29 (97)
Ever taken PrEP, n (%)	
Never	14 (47)
Yes, but not currently taking PrEP	5 (17)
Yes, I am currently taking PrEP	11 (37)
How likely to use PrEP in the future (extremely likely or likely)[c], n (%)	18 (60)
Previously been in a PrEP research study, n (%)	21 (70)
Operating system, n (%)	
iOS	20 (67)
Android	10 (33)
Share phone with someone, n (%)	2 (7)

XSL•FO
RenderX

Characteristics	Values
Phone disconnected (past 12 months), n (%)	14 (47)
Phone lost or stolen (past 12 months), n (%)	9 (30)
Ever used internet for, n (%)	
Tracking health behaviors	20 (67)
Acquiring information about HIV or other STIs	13 (43)
Acquiring other health or medical information	15 (50)
Sending reminders	17 (57)
Most common ways for getting health information , n (%)	
Website or app on mobile phone	23 (77)
Website or app on laptop or desktop computer	15 (50)
Health care providers	21 (70)
Think receiving health information through a mobile app would be of interest to young MSM[d], n (%)	27 (90)

[a]STI: sexually transmitted infection.

[b]PrEP: pre-exposure prophylaxis.

[c]Among all participants.

[d]MSM: men who have sex with men.

Overall, the app was well-received by participants, and the design was thought to be aesthetically pleasing and intuitive. Participants saw the app's utility and thought that it would be a helpful tool for people to engage in their sexual health. A participant in Tampa stated the following:

> *I see this as a fitness health app, but for sexual health. So that does seem useful.*

Participants suggested that the goal of the app be made clear as part of the onboarding process, which was added after the first round of focus groups.

Regarding SexPro, participants felt the numeric score and the color coding (red for low score indicating low protection from HIV, yellow or orange for intermediate score indicating moderate protection, and green for a high score indicating high protection) were important to provide a more robust understanding of what their score meant. One participant in Chicago described it as follows:

> *The higher the number the better, even 12 versus 9 just sounds and feels better. Like if I have a 12, I feel like I'm getting there, I can start to get better. If I see a 7, I just think "fuck!" I'm screwed. Especially versus if you just showed me that scale with no numbers, I would be worried, I would think I'm not protected. I would think crap, I'm screwed.*

Another participant in Chicago stated the following:

> *Yeah, if it's orange I wanna do something to get it to go up (to green). But if it's red you might think you are going to have to make drastic lifestyle changes to get out of the red zone.*

Participants thought that seeing the score could help motivate a change in behavior, as a participant from Tampa stated the following:

> *Seeing the score would be good to motivate you to maybe change your life, or your habits, so you can increase your score.*

Several participants suggested adding a summary to help users understand what went into their score, as well as suggestions for how to improve their score. A participant in Tampa said the following:

> *I think it would be useful to see, rather than a number, but like why I got a number. If I compared my score to a friend's, and they had a different number—we could compare, like, this is what I did, this is what you did, that's why a summary page would be more useful.*

Participants also liked the sexual diary component of the app. A participant in Chicago said the following:

> *I do think that's kind of useful, so you can keep track of what you're doing, or you know, if you need to "slow your ho down"—I feel like you can be able to hide it, so people don't see it, or call your 5-star person up.*

Another participant in Chicago commented the following:

> *The diary makes it fun, and not just a health app. I would look forward to looking back at it because of all the fun comments I made in it. It makes being safe fun to me.*

One participant liked having a separate place to track their sexual partners without having the information living in their phone's contact list, saying the following:

> *Let's just say I was raped, or I had lost my phone or whatever, I could download the app on my new phone where I had backed it up and I would still have that person's information in the app, because I don't*

XSL•FO
RenderX

> *usually add people in my phone contact list. It's good information to have and keeps it all private.*

Participants also liked the notes section of the diary, as it made the diary seem less like a tool of objectification, with a participant in Tampa saying the following:

> *People judge people anyway, but having like the notes makes it more personal... you can go back and look at the things you said to remember, or things about that person you want to recall later, even if it's like, you know, their apartment was dirty.*

Participants did recommend changing the terminology for describing partner types in the diary. For example, they found "no strings attached" to be confusing and suggested changing this to "one night stand."

Most participants felt comfortable using the home HIV and STI self-collection kits and saw value in this service, although YSMM aged 15 to 18 years expressed some concerns regarding potential errors with using the STI test kits and felt the shipping process was cumbersome and may be a deterrent for youth. Participants remarked that testing services were limited in their area (particularly Tampa). They reported that the few places that offered services had limited hours that conflicted with work and school schedules. A participant in Chicago said the following:

> *This option is a lot better especially if you are scared to go to a doctor. That's reassuring.*

They highlighted the importance of confidentiality. A participant in Tampa said the following:

> *As long as the packaging is confidential that came to my home. I'm totally fine to do a care kit like that. Especially since it's free to request and ship.*

However, some participants felt that testing at home would mean they would miss out on the opportunity to build rapport with their medical provider. In addition, some participants reported concerns with collecting the specimens and preparing the kits for shipment, as a participant in Chicago said the following:

> *For me it's easier to go to the doctor's office, then I don't have to worry about messing something up or anything like that.*

If someone were to receive a positive HIV test result, participants wanted immediate access to a provider who could offer supportive counseling. They also suggested adding a chat function to communicate with the clinic.

Participants appreciated the PrEP components of LYNX, including the ability to find local PrEP clinics in their area; however, they recommended that the goal of increasing PrEP uptake be made clearer in the onboarding section of the app. One focus group participant in Tampa said the following:

> *As far as like the HIV and like the prep, thing when I first downloaded it, I didn't really see it as that being like the focus. I thought like the focus was more of tracking your partners.*

While some participants liked showing how one's SexPro score increases on PrEP, they recommended including information about other non-PrEP–related factors that could influence their SexPro score, as described by another participant in Tampa:

> *Yeah like if swiping right included other things like let's see what your score would look like if you used like more condoms, or you know decreased your number of partners as well. Like different factors not always just, well if you were on PrEP this is what would like... obviously PrEP is not the only thing that's going to bring up your level and increase your protection.*

Participants were also interested in gamification and receiving feedback from the app based on the data they entered into the app. When shown examples of badges in the focus groups, one participant said the following:

> *It's a pat on the back. It's cute! You are being rewarded for being a ho! (followed by laughter)*

They liked the golden penis badge and suggested having a "gold butt" badge as a complement (Figure 1). They also recommended creating badges to keep track of PrEP pill taking and having a "desert badge" when they are not having sex. They also liked the idea of having a trends and insights page, which would provide additional feedback on their trends in sexual behaviors.

Changes to the LYNX App

On the basis of the feedback from the focus groups, we made several revisions and additions to the LYNX app, guided by the IMB model. Regarding "information," several welcome screens were added to the onboarding process to clarify the goals of the app, including the message "to marry pleasure with prevention," which was a theme that emerged from the focus groups, as well as clarifying the goal of increasing PrEP uptake. To provide more information about PrEP, a PrEP Facts page was added that included answers to frequently asked questions about PrEP. Regarding SexPro, we added information on behaviors contributing to the calculation of their SexPro score, with an option to click on "Learn why these matter," which provides a more detailed description of how different behaviors change one's risk for HIV acquisition. We also added information about how participants could improve their score and how their score would be impacted by taking PrEP. To provide additional feedback to participants on the data they entered into LYNX, we also created a trends page that summarized different sexual behaviors and encounters (eg, you were exclusively a bottom 50% of the time) and an insights page that summarized characteristics of their sexual partners (gender identity, partner type, partner HIV status, and number of repeat partners). Regarding "motivation," we added video testimonials of YSMM on their experience taking PrEP to help motivate YSMM to consider taking PrEP. We also added several gamification features to the app, including the ability to earn different types of badges (Figure 1), and a "top 5" feature for the 5 most highly rated partners and 5 most highly rated sexual encounters. In addition, the language in the sex diary was modified based on input from the focus groups (eg, changing "no strings attached" to "one night stand") and the notes section was included in the

XSL•FO
RenderX

final design to allow participants to enter free text to capture anything they would like to remember about the partner. Regarding Behavioral skills, we added the PrEP locator function to allow participants to find a local PrEP clinic that served youth. We also added a chat function to allow participants and study staff to send and receive in-app messages for additional guidance and support, as well as a "need us now" button in red at the top of the chat screen, which would allow participants to call an on-call provider available at all times. This chat function was real time and monitored by study staff (including study clinicians), who responded to participants within 24 business hours. Finally, we engaged a copywriter who simplified the language, using fewer words, and made the language in the app more appealing for youth. There were a few suggestions that were not incorporated into the app due to feasibility considerations but noted for future iterations of the app, including log-in via touch ID or social media log-in (instead of passcode), adding a group sex option in the sex diary, and adding dosing reminders for people who started taking PrEP.

Technical Pilot to Assess Feasibility and Acceptability of the LYNX App

From April to June 2018, a total of 17 participants were enrolled in the pilot study: 8 (47%) in Chicago and 9 (53%) in Tampa. Table 2 describes baseline characteristics of participants included in the technical pilot study. Overall, the mean age was 22.4 (SD 1.6) years; 8 (47%) self-identified as Hispanic or Latinx, and 5 (29%) completed high school or lower. The majority were currently employed (15/17, 88%), reported low income (12/17, 71%), lived with another person (14/17, 82%), currently had health insurance (10/17, 59%), and did not have a primary health care provider (12/17, 71%). The mean number of partners was 3.1 (SD 2.5) in the past 3 months, and 11 (65%) participants reported condomless sex. A total of 5 (29%) participants never tested for HIV, and only 1 (6%) tested in the past 3 months. For other STIs, 9 (53%) participants had never tested before and 8 (47%) tested >3 months ago. Regarding prevention technologies, 11 (65%) and 15 (88%) heard about HIV self-testing and PrEP before the study, respectively. Most participants (n=10, 59%) reported using the iOS mobile phone operating system, had a phone contract billed monthly (n=9, 53%), and used apps more than once per week (n=14, 82%).

A total of 15 (N=17, 88%) participants completed the follow-up visit. The mean SUS score was 70.4 (SD 17.2), corresponding to a rating in the good range (Table 3) [30]. The highest score was on item "I thought the system was easy to use" (mean 4.3, SD 0.9), and the lowest score was on item "I needed to learn a lot of things before I could get going with LYNX" (mean 2.2, SD 1.4). Most reported using the LYNX app at least once a week (8/14, 57%); that they would be somewhat likely, likely, or very likely to continue using the app (11/14, 79%); and would definitely recommend it to a friend in need of similar help with getting tested for HIV and STIs and accessing PrEP (10/14, 71%). A total of 43% (6/14) of participants ordered an HIV test kit, and 36% (5/14) ordered an STI test kit.

Most participants rated all components of LYNX as not at all difficult or a little difficult to use (Figure 2). Almost 86% (12/14) of participants reported the components "Taking SexPro Quiz" and "Entering information as part of the onboarding process" as not at all difficult. More than half of the participants considered all components at least somewhat helpful (Figure 3). "Ordering test kits and safer sex supplies" was considered extremely helpful for 64% (9/14) of respondents. More than half of the participants strongly agreed or agreed with statements related to IMB model impact (Figure 4). More than 90% (13/14) strongly agreed or agreed that "the LYNX app helped me keep track of my sexual partners and the types of sex I have" (13/14, 93%). More than 70% (10/14) strongly agreed or agreed that "The LYNX app helped me understand my risk for getting HIV" (10/14, 72%) and "The LYNX app helped me understand whether PrEP would be a good fit for me" (11/14, 79%) and half strongly agreed or agreed that the app helped them in starting PrEP (7/14, 50%).

Participants opened the app, on average, 8.5 (SD 8.0, range 1-31) times during the 2-month pilot intervention period, with an average duration of 3.8 (SD 3.2, range 0.2-11.8) minutes per session. Across all participants, the cumulative time spent in the app ranged from 0.6 to 183.7 minutes, with a median of 16.5 (IQR 5.45-40.8) minutes. The mean time between enrollment and first app log-in was 5.7 (SD 17.2, range 0-67) days. The mean number of features accessed was 9.5 (SD 4.8, range 0-14) out of 16 features. The "testing" feature was used by all participants (15/100, 15%), followed by "Activity Calendar" (14/15, 93%) and "Diary" (13/15, 87%; Table 4). The most visited features were "Home Screen" (240 visits), "Diary" (189 visits), and "Testing" (174 visits).

XSL•FO
RenderX

Table 2. Baseline characteristics of technical pilot participants for an HIV prevention mobile app from Chicago, Illinois, and Tampa, Florida (N=17).

Characteristics	Values
City, n (%)	
Chicago	8 (47)
Tampa	9 (53)
Age (y), mean (SD)	22.4 (2)
Sexual orientation, n (%)	
Gay	11 (65)
Bisexual	5 (29)
Queer	1 (6)
Race, n (%)	
Asian	1 (6)
Black	4 (24)
Hispanic or Latinx	8 (47)
White	10 (63)
Schooling, n (%)	
Completed high school or lower	5 (29)
Some college	7 (41)
Completed college	5 (29)
Currently in school	7 (41)
Current employed, n (%)	15 (88)
Income (self-defined), n (%)	
Low	12 (71)
Middle	4 (24)
Declined to answer	1 (6)
Relationship status, n (%)	
Single	4 (24)
Casually dating	5 (29)
Boyfriend, girlfriend, partner, or lover	8 (47)
People living with, n (%)	
Alone	3 (18)
1-2	5 (29)
3-5	7 (41)
>5	2 (12)
Health insurance, n (%)	
No	7 (41)
I have my own	6 (35)
I am covered by my parent or guardian	4 (24)
Primary health care provider, n (%)	5 (29)
Number of male partners (past 3 months), mean (SD)	3.1 (3)
Condomless sex (past 3 months), n (%)	11 (65)
HIV testing, n (%)	
Never	5 (29)
Past 3 months	1 (6)

XSL•FO
RenderX

Characteristics	Values
>3 months	11 (65)
Ever tested for other STIs[a], n (%)	
Never	9 (53)
Past 3 months	0 (0)
>3 months	8 (47)
HIV self-testing, n (%)	
Never heard about it	6 (35)
Heard before, but never used it	9 (53)
Used previously	2 (12)
PrEP[b] awareness before the study, n (%)	15 (88)
Interested in using PrEP, n (%)	
Not at all interested	2 (12)
A little interested	5 (29)
Somewhat interested	4 (24)
Very interested	4 (24)
Extremely interested	1 (6)
Operating system, n (%)	
iOS	10 (59)
Android	7 (41)
Share phone with someone, n (%)	3 (18)
Service, n (%)	
Monthly	9 (53)
Prepaid	2 (12)
Shared plan	6 (35)
Phone disconnected (past 12 months), n (%)	3 (18)
How often use apps, n (%)	
≤1 per week	3 (18)
>1 per week	14 (82)

[a]STI: sexually transmitted infection.

[b]PrEP: pre-exposure prophylaxis.

XSL•FO
RenderX

Table 3. System Usability Scale (SUS) and app use among technical pilot participants in Tampa and Chicago (N=14).

SUS[a]	Values
Items, mean (SD)	
1: I think that I would like to use LYNX on a regular basis.	3.5 (1.3)
2: I found LYNX unnecessarily complex.	2.4 (1.3)
3: I thought the system was easy to use.	4.3 (0.9)
4: I think that I would need the support of a technical person to be able to use LYNX.	2.3 (1.4)
5: I found the various functions in LYNX were well integrated.	4.0 (1.4)
6: I thought there was too much inconsistency between different parts of LYNX.	2.6 (1.2)
7: I would imagine that most people would learn to use LYNX very quickly.	4.2 (1.1)
8: I found LYNX very cumbersome to use.	2.5 (1.2)
9: I felt very confident using LYNX.	3.9 (1.1)
10: I needed to learn a lot of things before I could get going with LYNX.	2.2 (1.4)
SUS total score, standardized (0-100), mean (SD)	70.4 (17.2)
In the past 2 months, on average, how often did you use the LYNX app?, n (%)	
Once a month or less	1 (7)
A few times a month	5 (36)
Once a week	5 (36)
A few times a week	2 (14)
Once a day	1 (7)
How likely would you be to continue using the LYNX app if it were available?, n (%)	
Very unlikely	2 (14)
Unlikely	1 (7)
Somewhat unlikely	0 (0)
Neutral	0 (0)
Somewhat likely	4 (29)
Likely	4 (29)
Very likely	3 (21)
If a friend were in need of similar help with getting tested for HIV or STIs[b] and accessing PrEP[c], would you recommend the LYNX app to him?, n (%)	
No, definitely not	0 (0)
No, I do not think so	0 (0)
Yes, I think so	4 (29)
Yes, definitely	10 (71)

[a]Scores range from 1 (strongly disagree) to 5 (strongly agree).

[b]STI: sexually transmitted infection.

[c]PrEP: pre-exposure prophylaxis.

XSL•FO
RenderX

Figure 2. Ratings of ease of use of the components of the LYNX app among technical pilot participants in Tampa and Chicago. PrEP: pre-exposure prophylaxis; STI: sexually transmitted infection.

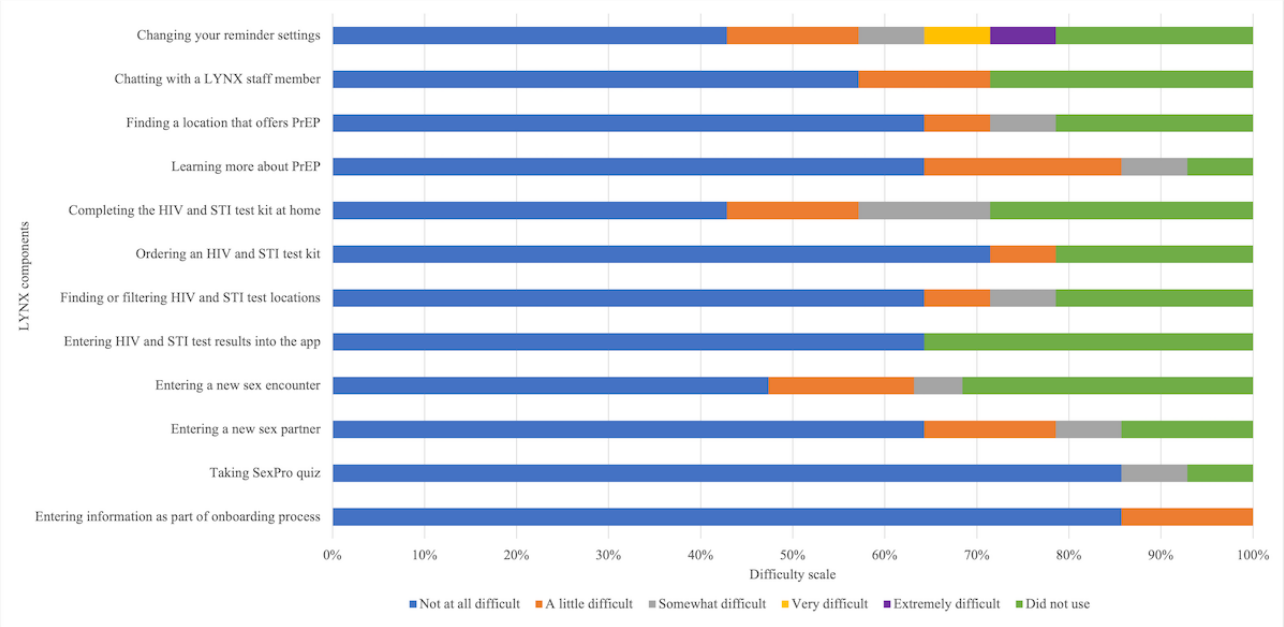
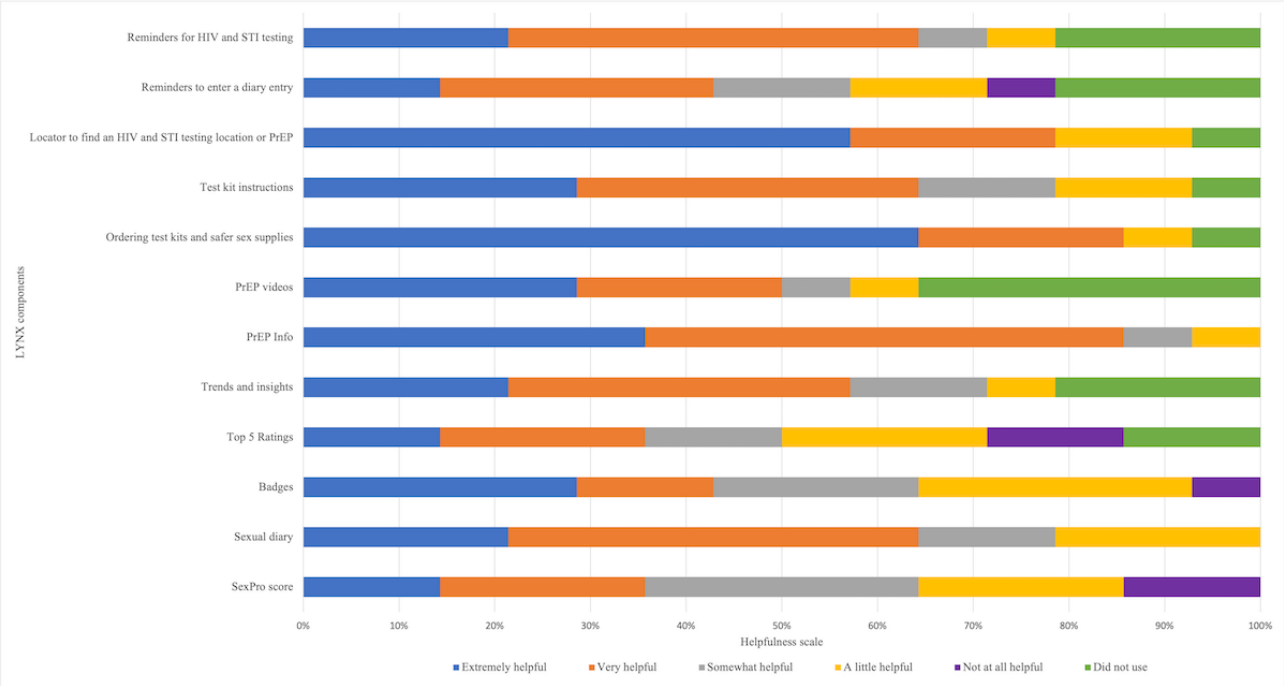


Figure 3. Ratings of helpfulness of the components of the LYNX app among technical pilot participants in Tampa and Chicago. PrEP: pre-exposure prophylaxis; STI: sexually transmitted infection.



XSL•FO
RenderX

Figure 4. Impact of the LYNX app on Information-Motivation-Behavioral Skills model constructs among technical pilot participants in Tampa and Chicago. Responses ranged from strongly agree to strongly disagree. No participants indicated "strongly disagree" to any of the response items. PrEP: pre-exposure prophylaxis.

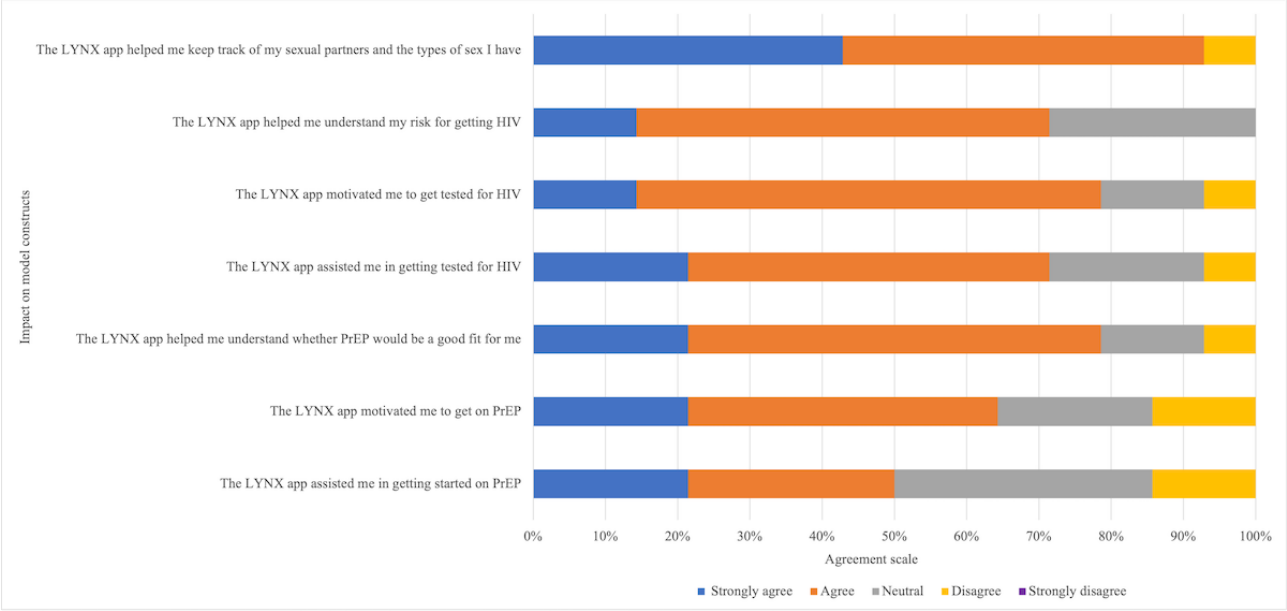


Table 4. Frequency of features being viewed or accessed by participants in a pilot evaluation of a mobile HIV prevention app from Tampa and Chicago (N=15).

Feature	Unique users, n (%)	Visits, n (%)
Testing page for logging HIV and STI[a] testing and ordering a test kit	15 (100)	175 (17)
Activity Calendar showing partner encounters	14 (93)	41 (4)
Diary to view and add new partners and encounters	13 (86)	189 (18)
Insights on partner types and HIV status	12 (80)	25 (2)
Navigation Chat: click to page to chat with LYNX team	12 (80)	72 (7)
Navigation PrEP[b]: click to PrEP information page	12 (80)	107 (10)
Sex trends, including sexual activities and condom use	11 (73)	31 (3)
Earned badges based on app use	10 (66)	23 (2)
Profile, including account details and setting reminders	10 (66)	19 (1)
SexPro page to view the current score and explanation	9 (60)	27 (2)
Top 5 partners and encounters	9 (60)	19 (1)
Encounter: recording a new entry in sex diary	6 (40)	24 (2)
PrEP Facts, including frequently asked questions about PrEP	4 (26)	15 (1)
PrEP Videos: testimonials from YSMM[c] on their PrEP experience	4 (26)	12 (1)
Testing and PrEP Map to search for testing and PrEP clinics	0 (0)	0 (0)

[a]STI: sexually transmitted infection.

[b]PrEP: pre-exposure prophylaxis.

[c]YSMM: young sexual minority men.

Among the 15 participants enrolled in the technical pilot, 9 (60%) participants completed an HIV test during the pilot, including 4 (27%) who tested at home. In addition, 4 (27%) participants started PrEP during the pilot.

Among 17 pilot study participants, 14 (82%) completed follow-up exit interviews. Overall, there was a high level of acceptability of content, interface, and features of the LYNX app, although some participants reported that they did not use all the app features (eg, sex diary) due to being in a monogamous relationship. One participant in Tampa noted the following:

> *I think that this app is so beneficial for like young gay men...they are scared and not confident in talking about this stuff and they don't have anyone to talk to...I think that's one thing that's really like important*

XSL•FO
RenderX

> *for this app because it has the answers, it has where to go, it has everything you need to know when you don't have the answers.*

Most participants liked the SexPro feature and thought that it accurately reflected their risk; however, most did not look back at the score during the pilot and did not believe that the SexPro score changed their level of interest in PrEP. Most participants liked the diary, trends, and insights features, as they helped them keep track of their partners and encounters and reflect on their sexual behaviors. One participant in Tampa said the following:

> *Yeah. I actually really thought the app is great. I lost tracking, my—like just my encounters and thought that was really beneficial especially like going back in time to kind of look at it because I don't know, it just keeps your eyes open about what you're doing.*

However, one participant felt that the diary was judgmental and objectifying of partners and was concerned that it would hurt a partner's feelings if they somehow saw it. Participants felt the diary was more helpful in tracking one-time encounters with casual partners, and this feature was less helpful for those in monogamous relationships. One participant suggested being able to look at data from a specific time frame (rather than cumulative partners and encounters), and several participants wanted the ability to edit previously entered entries. While most participants did not use the chat function, they liked that the feature was available, especially the "need us now" button. Many participants did not recall receiving notifications and reminders for HIV and STI testing, and would have preferred either weekly or monthly reminders. A few participants reported technical issues, such as receiving notifications at unscheduled times, and noted some bugs with the chat feature. Several participants who used the test kit reported being fearful of collecting the blood sample, and one participant had trouble with sending the kit back as they had lost the instruction sheet. Several participants suggested having video tutorials to help with the specimen collection and shipping process. One Chicago participant suggested adding resources for mental health, healthy relationships, and connections to the community. When asked whether they would use the app in the future, one participant in Chicago said the following:

> *Yes, 100%... It's really—it's good. I would–like yeah, I would advocate for it.*

Discussion

Principal Findings

Sexual minority youth face numerous barriers to accessing HIV and STI prevention services, including stigma and lack of access to prevention clinics, and have expressed high levels of interest in mobile health (mHealth) interventions to support HIV and STI prevention, which may help address some of these barriers [31,32]. While a growing number of mobile apps are now being designed and tested to support HIV and STI testing, linkage to care, and PrEP uptake [33], few have been tailored specifically for youth. This study describes findings from 2 phases of formative research to develop and tailor the LYNX app to increase HIV and STI testing and PrEP uptake among YSMM in the United States and demonstrated high initial acceptability and feasibility of the app.

To increase user engagement with mobile apps, it has been suggested that developers use an iterative data collection process to obtain input from the target audience about app preferences [34]. Similar to other studies that used focus groups or theater testing as part of formative development [35-38], we conducted iterative focus groups with a cohort of YSMM to efficiently elicit feedback and inform the development of updated prototypes of the LYNX app. Overall, the app's design was well-received by YSMM, and we received feedback to improve key components of the LYNX app, including the SexPro score and sexual diary, as well as recommendations for additional features to facilitate timely communication between users and staff and increase overall engagement and use of LYNX. In a 2-month technical pilot of the app, the revised LYNX app demonstrated feasibility based on use of the app as assessed via paradata and good acceptability as assessed via self-report.

Sanabria et al [39] used a multimethod, user-centered design approach, including input from a youth advisory board, a think-aloud protocol, and use case scenarios, to evaluate the usability of a mobile app to increase HIV testing among sexual and gender minority youth. This app, featuring HIV prevention information and an imaging algorithm for interpreting home-based HIV tests, demonstrated high user satisfaction and perceived usability that was independent of education and health literacy levels [40,41].

Several HIV risk scores have been developed and tested to assist in identifying individuals who may benefit from PrEP [42-45]; however, these have primarily been designed to be administered by providers and can be problematic, because clients may not be comfortable disclosing behaviors that put them at increased risk for HIV [46], and because identifying people at risk can be stigmatizing [47]. In this study, we tailored a patient-facing sexual health score (using a positive frame, with a higher score indicating higher sexual protection) specifically for use in youth. YSMM in our focus groups liked having both a color and numerical representation of protection from HIV; however, they wanted additional information on how the SexPro score was calculated and how they could improve their score. On the basis of these findings, key adaptations included adding information on the factors contributing to their score and how they could increase their score by changing behaviors or initiating PrEP.

Prior HIV research studies have also used sex diaries to capture detailed event-level sexual encounter data and to understand different profiles of HIV risk [48,49]; however, few have incorporated diaries into mobile apps as part of an mHealth intervention. A mobile app designed to collect sexual behavior in the Amsterdam PrEP Project was frequently used by SMM, with sexual behavior data collected by the app consistent with questionnaire data among those who used the app consistently [50]. In addition, in a study by Glick et al [51], completing sexual diaries was associated with reduced frequencies of reported anal sex, acquisition of new partners, and lower incident HIV and STI diagnoses. In our study, the sexual diary was included as part of the LYNX app as a strategy to help YSMM better understand and track their sexual behaviors, as well as

XSL•FO
RenderX

increase engagement with the use of the app. In focus groups and the technical pilot, participants liked the diary feature; however, they made recommendations to tailor the language used in the sexual diary for YSMM and recommended an open-text section so the youth could customize their diary entries. The diary also appeared to be more useful for YSMM with a greater number of casual sexual partners. It will be important in future studies to characterize the types of partners YSMM enter into the app, for whom the diary is most useful, and whether diary use is sustained over time.

A key finding in our study was the importance of incorporating gamification features to increase engagement with the LYNX app by the youth. In a recent systematic review on the impact of mHealth-based gamification on the HIV prevention and care continuum in SMM by Luo et al [52], gamification interventions were shown to reduce condomless anal sex acts and were associated with an increase in PrEP adherence. Across the 26 studies included in the review, the most frequently used game components were points, followed by challenge, discussion forums, and mentorship or character narrative. In the LYNX app, we incorporated several gamification features based on feedback from YSMM in focus groups, including the ability to earn different badges, a "top 5 feature" for the most highly rated partners and sexual encounters, and a trends and insights page to synthesize the partner data they entered into the app into useful summaries. In other studies, the youth have also suggested that tangible versus virtual rewards (eg, a gift card) may be more likely to increase engagement [53], which should be explored in future studies.

Similar to other studies, most LYNX participants expressed interest in home HIV and STI self-collection kits, which could help overcome barriers to HIV testing among YSMM [31]. Convenience, privacy, and rapid result delivery have been cited by YSMM as principal motivators for using HIV self-tests [54]. In our study, YSMM expressed some concerns with collecting and preparing the specimens for shipment. The youth may benefit from videos demonstrating the specimen collection process and check-ins from staff on whether they need assistance with completing the test kits. On the basis of feedback in focus groups, we added a chat feature to facilitate communication between YSMM and staff and a button to call a clinician (eg, in the event of a positive HIV test), app features which have also been recommended in other studies of youth [55,56].

Limitations and Strengths

This study had several limitations. First, the technical pilot was small and not designed to evaluate the efficacy of the intervention. In addition, due to the short duration of the pilot, we were unable to evaluate long-term feasibility and acceptability. Nonetheless, formative work through focus groups and optimizing usability in a technical pilot are critical steps in developing and tailoring interventions before testing efficacy in large randomized controlled trials. Social desirability may have impacted YSMM responses in the focus groups and technical pilot; however, we used a self-administered web-based survey to assess usability and acceptability and paradata metrics from the LYNX app to objectively assess feasibility. In addition, participants in focus groups and exit interviews were reminded that there were no right or wrong answers and were encouraged to provide honest feedback that would help us best improve the app. For both phases of this study, participants were enrolled in Chicago and Tampa, potentially limiting the generalizability of the findings, although both regions are in high-priority "ending the HIV epidemic" jurisdictions, and 1 site is located in the South, the region accounting for over half of new HIV diagnoses nationally. Despite these limitations, our study had several strengths, including enrolling a diverse cohort of YSMM across both sites; the user-centered approach, incorporating input from our participants into the final design of our intervention; and the mixed methods approach to evaluate feasibility and acceptability in our technical pilot.

Conclusions

Using a user-centered design approach via iterative focus groups, we tailored the LYNX app to support HIV and STI testing and PrEP uptake among YSMM in the United States. Preliminary testing in a technical pilot demonstrated high acceptability and feasibility of the refined version of the LYNX app and supports further evaluation in an upcoming effectiveness trial.

Acknowledgments

We would like to thank Amayvis Garcia for assistance with recruiting participants and collecting study data. This study was funded by the National Institutes of Health Adolescent Medicine Trials Network for HIV Interventions (grant U01HD068040; principal investigators: LHW and PS). The authors thank all the participants for their helpful insights and feedback in this study. TST was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil—Finance Code 001, Conselho Nacional de Desenvolvimento Científico e Tecnológico, and Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro.

Conflicts of Interest

AYL has received funding for investigator-sponsored research studies from Gilead Sciences and Viiv Healthcare and has led studies in which the study drug has been donated by Gilead Sciences. SPB has received funding for research studies from GSK, Gilead Sciences, and Merck. TST received consulting fees from Gilead. TO is a full time employee of AstraZeneca.

References

1. HIV surveillance report: diagnoses of HIV infection in the United States and dependent areas, 2021. Centers for Disease Control and Prevention. 2023 Jan 5. URL: https://stacks.cdc.gov/view/cdc/149071 [accessed 2023-06-26]

XSL•FO
RenderX

2. Wiatrek S, Zlotorzynska M, Rai R, Sullivan P, Sanchez T. The annual American men's internet survey of behaviors of men who have sex with men in the United States: key indicators report 2018. JMIR Public Health Surveill 2021 Mar 04;7(3):e21812 [FREE Full text] [doi: 10.2196/21812] [Medline: 33496669]
3. HV surveillance supplemental report: monitoring selected national HIV prevention and care objectives by using HIV surveillance data United States and 6 territories and freely associated states, 2022. Centers for Disease Control and Prevention. 2024. URL: https://stacks.cdc.gov/view/cdc/156511 [accessed 2024-05-25]
4. Sexually transmitted disease surveillance 2019. Centers for Disease Control and Prevention. 2021 Apr. URL: https://www.cdc.gov/sti-statistics/media/pdfs/2024/07/std-surveillance-2019.pdf [accessed 2023-10-09]
5. Shannon CL, Klausner JD. The growing epidemic of sexually transmitted infections in adolescents: a neglected population. Curr Opin Pediatr 2018 Feb;30(1):137-143 [FREE Full text] [doi: 10.1097/MOP.0000000000000578] [Medline: 29315111]
6. de Voux A, Bernstein KT, Kirkcaldy RD, Zlotorzynska M, Sanchez T. Self-reported extragenital chlamydia and gonorrhea testing in the past 12 months among men who have sex with men in the United States—American men's internet survey, 2017. Sexual Trans Dis 2019 Jun 11;46(9):563-570. [doi: 10.1097/olq.0000000000001032]
7. Balán IC, Lopez-Rios J, Dolezal C, Rael CT, Lentz C. Low sexually transmissible infection knowledge, risk perception and concern about infection among men who have sex with men and transgender women at high risk of infection. Sex Health 2019 Nov;16(6):580-586 [FREE Full text] [doi: 10.1071/SH18238] [Medline: 31699208]
8. MacKellar DA, Valleroy LA, Secura GM, Behel S, Bingham T, Celentano DD, et al. Perceptions of lifetime risk and actual risk for acquiring HIV among young men who have sex with men. AIDS Behav 2007 Mar;11(2):263-270. [doi: 10.1007/s10461-006-9136-0] [Medline: 16791527]
9. Dorell CG, Sutton MY, Oster AM, Hardnett F, Thomas PE, Gaul ZJ, et al. Missed opportunities for HIV testing in health care settings among young African American men who have sex with men: implications for the HIV epidemic. AIDS Patient Care STDS 2011 Nov;25(11):657-664. [doi: 10.1089/apc.2011.0203] [Medline: 21923415]
10. Grant RM, Lama JR, Anderson PL, McMahan V, Liu AY, Vargas L, et al. Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. N Engl J Med 2010 Dec 30;363(27):2587-2599 [FREE Full text] [doi: 10.1056/NEJMoa1011205] [Medline: 21091279]
11. McCormack S, Dunn DT, Desai M, Dolling DI, Gafos M, Gilson R, et al. Pre-exposure prophylaxis to prevent the acquisition of HIV-1 infection (PROUD): effectiveness results from the pilot phase of a pragmatic open-label randomised trial. Lancet 2016 Jan 02;387(10013):53-60 [FREE Full text] [doi: 10.1016/S0140-6736(15)00056-2] [Medline: 26364263]
12. Baeten JM, Donnell D, Ndase P, Mugo NR, Campbell JD, Wangisi J, et al. Antiretroviral prophylaxis for HIV prevention in heterosexual men and women. N Engl J Med 2012 Aug 02;367(5):399-410 [FREE Full text] [doi: 10.1056/NEJMoa1108524] [Medline: 22784037]
13. AIDSVu releases new PrEP data and launches PrEPVu.org, a new PrEP equity platform. AIDSVu. 2024 Jun 25. URL: https://tinyurl.com/44xku9yx [accessed 2025-04-16]
14. Hosek SG, Rudy B, Landovitz R, Kapogiannis B, Siberry G, Rutledge B, et al. An HIV preexposure prophylaxis demonstration project and safety study for young MSM. J Acquir Immune Defic Syndr 2017 Jan 01;74(1):21-29 [FREE Full text] [doi: 10.1097/QAI.0000000000001179] [Medline: 27632233]
15. Mayer KH, Agwu A, Malebranche D. Barriers to the wider use of pre-exposure prophylaxis in the United States: a narrative review. Adv Ther 2020 May;37(5):1778-1811 [FREE Full text] [doi: 10.1007/s12325-020-01295-0] [Medline: 32232664]
16. Bonacci RA, Smith DK, Ojikutu BO. Toward greater pre-exposure prophylaxis equity: increasing provision and uptake for Black and Hispanic/Latino individuals in the U.S. Am J Prev Med 2021 Nov;61(5 Suppl 1):S60-S72 [FREE Full text] [doi: 10.1016/j.amepre.2021.05.027] [Medline: 34686293]
17. Haggipavlou L, Hamshaw RJ. Barriers to PrEP uptake in young U.K. men who have sex with men. J Prev Health Promot 2023 Nov 27;4(3-4):404-433. [doi: 10.1177/26320770231199419]
18. Gold J, Lim MS, Hocking JS, Keogh LA, Spelman T, Hellard ME. Determining the impact of text messaging for sexual health promotion to young people. Sex Transm Dis 2011 Apr;38(4):247-252. [doi: 10.1097/OLQ.0b013e3181f68d7b] [Medline: 20966830]
19. Holloway IW, Cederbaum JA, Ajayi A, Shoptaw S. Where are the young men in HIV prevention efforts? Comments on HIV prevention programs and research from young men who sex with men in Los Angeles county. J Prim Prev 2012 Dec;33(5-6):271-278 [FREE Full text] [doi: 10.1007/s10935-012-0282-z] [Medline: 23132515]
20. Eyrich-Garg KM. Mobile phone technology: a new paradigm for the prevention, treatment, and research of the non-sheltered "street" homeless? J Urban Health 2010 May;87(3):365-380 [FREE Full text] [doi: 10.1007/s11524-010-9456-2] [Medline: 20397058]
21. Riley WT, Rivera DE, Atienza AA, Nilsen W, Allison SM, Mermelstein R. Health behavior models in the age of mobile interventions: are our theories up to the task? Transl Behav Med 2011 Mar;1(1):53-71 [FREE Full text] [doi: 10.1007/s13142-011-0021-7] [Medline: 21796270]
22. Muessig KE, Pike EC, Fowler B, LeGrand S, Parsons JT, Bull SS, et al. Putting prevention in their pockets: developing mobile phone-based HIV interventions for Black men who have sex with men. AIDS Patient Care STDS 2013 Apr;27(4):211-222 [FREE Full text] [doi: 10.1089/apc.2012.0404] [Medline: 23565925]

XSL•FO
RenderX

23. Anderson M. Technology device ownership: 2015. Pew Research Center. 2015 Oct 29. URL: http://www.pewinternet.org/2015/10/29/technology-device-ownership-2015 [accessed 2025-04-11]
24. Smith A. U.S. smartphone use in 2015. Pew Research Center. 2015 Apr 1. URL: https://www.pewresearch.org/internet/2015/04/01/us-smartphone-use-in-2015/2015/ [accessed 2025-04-25]
25. Fisher JD, Amico KR, Fisher WA, Harman JJ. The information-motivation-behavioral skills model of antiretroviral adherence and its applications. Curr HIV/AIDS Rep 2008 Nov;5(4):193-203. [doi: 10.1007/s11904-008-0028-y] [Medline: 18838059]
26. Liu A, Coleman K, Bojan K, Serrano PA, Oyedele T, Garcia A, et al. Developing a mobile app (LYNX) to support linkage to HIV/sexually transmitted infection testing and pre-exposure prophylaxis for young men who have sex with men: protocol for a randomized controlled trial. JMIR Res Protoc 2019 Jan 25;8(1):e10659 [FREE Full text] [doi: 10.2196/10659] [Medline: 30681964]
27. Scott H, Vittinghoff E, Irvin R, Liu A, Nelson L, Del Rio C, et al. Development and validation of the personalized sexual health promotion (SexPro) HIV risk prediction model for men who have sex with men in the United States. AIDS Behav 2020 Jan;24(1):274-283 [FREE Full text] [doi: 10.1007/s10461-019-02616-3] [Medline: 31352633]
28. Hightow-Weidman LB, Muessig K, Rosenberg E, Sanchez T, LeGrand S, Gravens L, et al. University of North Carolina/Emory Center for Innovative Technology (iTech) for addressing the HIV epidemic among adolescents and young adults in the United States: protocol and rationale for center development. JMIR Res Protoc 2018 Aug 03;7(8):e10365 [FREE Full text] [doi: 10.2196/10365] [Medline: 30076126]
29. Brooke J. SUS: a 'quick and dirty' usability scale. In: Usability Evaluation In Industry. Boca Raton, FL: CRC Press; 1996.
30. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009;4(3):1114-1123 [FREE Full text]
31. Marks SJ, Merchant RC, Clark MA, Liu T, Rosenberger JG, Bauermeister JA, et al. Barriers to HIV testing and opportunities for expansion using home-based HIV self-testing: results of a national study of higher HIV risk young men who have sex with men. Sage Open 2021 May 25;11(2). [doi: 10.1177/21582440211016835]
32. Jones J, Edwards OW, Merrill L, Sullivan PS, Stephenson R. Interest in HIV prevention mobile phone apps: focus group study with sexual and gender minority persons living in the rural southern United States. JMIR Form Res 2022 Jun 13;6(6):e38075 [FREE Full text] [doi: 10.2196/38075] [Medline: 35699980]
33. Mhando F, Nyankomo M, Hall C, Olughu K, Hlongwa M, Janson S, et al. Digital intervention services to promote HIV self-testing and linkage to care services: a bibliometric and content analysis-global trends and future directions. Public Health Rev 2024;45:1606354 [FREE Full text] [doi: 10.3389/phrs.2024.1606354] [Medline: 38434540]
34. Muessig KE, Pike EC, Legrand S, Hightow-Weidman LB. Mobile phone applications for the care and prevention of HIV and other sexually transmitted diseases: a review. J Med Internet Res 2013 Jan 04;15(1):e1 [FREE Full text] [doi: 10.2196/jmir.2301] [Medline: 23291245]
35. Biello KB, Hill-Rorie J, Valente PK, Futterman D, Sullivan PS, Hightow-Weidman L, et al. Development and evaluation of a mobile app designed to increase HIV testing and pre-exposure prophylaxis use among young men who have sex with men in the United States: open pilot trial. J Med Internet Res 2021 Mar 24;23(3):e25107 [FREE Full text] [doi: 10.2196/25107] [Medline: 33759792]
36. Zapata JP, Hirshfield S, Nelson K, Horvath K, John SA. Preferred characteristics for mHealth interventions among young sexual minoritized men to support HIV testing and PrEP decision-making: focus group study. JMIR Form Res 2023 Oct 16;7:e51103 [FREE Full text] [doi: 10.2196/51103] [Medline: 37713640]
37. Pingel ES, Thomas L, Harmell C, Bauermeister J. Creating comprehensive, youth centered, culturally appropriate sex education: what do young gay, bisexual and questioning men want? Sex Res Social Policy 2013 Dec 01;10(4):10.1007/s13178-013-0134-5 [FREE Full text] [doi: 10.1007/s13178-013-0134-5] [Medline: 24348222]
38. Goldenberg T, McDougal SJ, Sullivan PS, Stekler JD, Stephenson R. Building a mobile HIV prevention app for men who have sex with men: an iterative and community-driven process. JMIR Public Health Surveill 2015;1(2):e18 [FREE Full text] [doi: 10.2196/publichealth.4449] [Medline: 27227136]
39. Sanabria G, Scherr T, Garofalo R, Kuhns LM, Bushover B, Nash N, et al. Usability evaluation of the mLab app for improving home HIV testing behaviors in youth at risk of HIV infection. AIDS Educ Prev 2021 Aug;33(4):312-324 [FREE Full text] [doi: 10.1521/aeap.2021.33.4.312] [Medline: 34370566]
40. Dos Santos F, Garofalo R, Kuhns L, Wood O, Scherr T, Schnall R. Usability of a mobile health technology for HIV testing among sexual minorities. Stud Health Technol Inform 2024 Jul 24;315:744-745. [doi: 10.3233/SHTI240309] [Medline: 39049409]
41. Wood OR, Garofalo R, Kuhns LM, Scherr TF, Zetina APM, Rodriguez RG, et al. A randomized controlled trial of an mHealth intervention for increasing access to HIV testing and care among young cisgender men and transgender women: the mLab App study protocol. BMC Public Health 2021 Oct 29;21(1):1959 [FREE Full text] [doi: 10.1186/s12889-021-12015-w] [Medline: 34715833]
42. Smith DK, Pals SL, Herbst JH, Shinde S, Carey JW. Development of a clinical screening index predictive of incident HIV infection among men who have sex with men in the United States. J Acquir Immune Defic Syndr 2012 Aug 01;60(4):421-427. [doi: 10.1097/QAI.0b013e318256b2f6] [Medline: 22487585]

XSL•FO
RenderX

43. Wilton J, Kain T, Fowler S, Hart TA, Grennan T, Maxwell J, et al. Use of an HIV-risk screening tool to identify optimal candidates for PrEP scale-up among men who have sex with men in Toronto, Canada: disconnect between objective and subjective HIV risk. J Int AIDS Soc 2016;19(1):20777 [FREE Full text] [doi: 10.7448/IAS.19.1.20777] [Medline: 27265490]
44. Hoenigl M, Weibel N, Mehta SR, Anderson CM, Jenks J, Green N, et al. Development and validation of the San Diego Early Test Score to predict acute and early HIV infection risk in men who have sex with men. Clin Infect Dis 2015 Aug 01;61(3):468-475 [FREE Full text] [doi: 10.1093/cid/civ335] [Medline: 25904374]
45. Dijkstra M, Lin TC, de Bree GJ, Hoenigl M, Schim van der Loeff MF. Validation of the San Diego early test score for early human immunodeficiency virus infection among Amsterdam men who have sex with men. Clin Infect Dis 2020 May 06;70(10):2228-2230 [FREE Full text] [doi: 10.1093/cid/ciz895] [Medline: 31499523]
46. Oral PrEP risk assessment tools at a glance. PrEPWatch. 2018. URL: https://www.prepwatch.org/wp-content/uploads/2018/10/Oral_PrEP_Risk_Assessment-Tools_at_Glance.pdf [accessed 2024-05-11]
47. Golub SA. PrEP stigma: implicit and explicit drivers of disparity. Curr HIV/AIDS Rep 2018 Apr;15(2):190-197 [FREE Full text] [doi: 10.1007/s11904-018-0385-0] [Medline: 29460223]
48. Yu YF, Wu HJ, Ku SW, Huang PH, Li CW, Huang P, et al. Condomless anal sex associated with heterogeneous profiles of HIV pre-exposure prophylaxis use and sexual activities among men who have sex with men: a latent class analysis using sex diary data on a mobile app. J Med Internet Res 2021 Dec 23;23(12):e33877 [FREE Full text] [doi: 10.2196/33877] [Medline: 34941560]
49. Stalgaitis C, Glick SN. The use of web-based diaries in sexual risk behaviour research: a systematic review. Sex Transm Infect 2014 Aug 10;90(5):374-381 [FREE Full text] [doi: 10.1136/sextrans-2013-051472] [Medline: 24723619]
50. Finkenflügel RN, Hoornenborg E, Achterbergh RC, Marra E, Davidovich U, de Vries HJ, et al. A mobile application to collect daily data on preexposure prophylaxis adherence and sexual behavior among men who have sex with men: use over time and comparability with conventional data collection. Sex Transm Dis 2019 Jun;46(6):400-406 [FREE Full text] [doi: 10.1097/OLQ.0000000000000999] [Medline: 30882717]
51. Glick SN, Winer RL, Golden MR. Web-based sex diaries and young adult men who have sex with men: assessing feasibility, reactivity, and data agreement. Arch Sex Behav 2013 Oct 28;42(7):1327-1335 [FREE Full text] [doi: 10.1007/s10508-012-9984-9] [Medline: 22926939]
52. Luo Q, Zhang Y, Wang W, Cui T, Li T. mHealth-based gamification interventions among men who have sex with men in the HIV prevention and care continuum: systematic review and meta-analysis. JMIR Mhealth Uhealth 2024 Apr 15;12:e49509 [FREE Full text] [doi: 10.2196/49509] [Medline: 38623733]
53. Giovenco D, Muessig KE, Horvitz C, Biello KB, Liu AY, Horvath KJ, et al. Adapting technology-based HIV prevention and care interventions for youth: lessons learned across five U.S. Adolescent Trials Network studies. Mhealth 2021 Apr;7:21 [FREE Full text] [doi: 10.21037/mhealth-20-43] [Medline: 33898590]
54. Hurt CB, Soni K, Miller WC, Hightow-Weidman LB. Human immunodeficiency virus testing practices and interest in self-testing options among young, Black men who have sex with men in North Carolina. Sex Transm Dis 2016 Sep;43(9):587-593 [FREE Full text] [doi: 10.1097/OLQ.0000000000000484] [Medline: 27513387]
55. Holloway IW, Winder TJ, Lea CH, Tan D, Boyd D, Novak D. Technology use and preferences for mobile phone-based HIV prevention and treatment among Black young men who have sex with men: exploratory research. JMIR Mhealth Uhealth 2017 Apr 13;5(4):e46 [FREE Full text] [doi: 10.2196/mhealth.6436] [Medline: 28408360]
56. Biello KB, Horvitz C, Mullin S, Mayer KH, Scott H, Coleman K, et al. HIV self-testing and STI self-collection via mobile apps: experiences from two pilot randomized controlled trials of young men who have sex with men. Mhealth 2021;7:26 [FREE Full text] [doi: 10.21037/mhealth-20-70] [Medline: 33898595]

Abbreviations

ATN: Adolescent Medicine Trials Network for HIV Interventions
HIPAA: Health Insurance Portability and Accountability Act
IMB: Information-Motivation-Behavioral Skills
IRB: Institutional Review Board
iTech: innovative technology
mHealth: mobile health
PrEP: pre-exposure prophylaxis
SMM: sexual minority men
STI: sexually transmitted infection
SUS: System Usability Scale
YSMM: young sexual minority men

XSL•FO
RenderX

Edited by A Kushniruk; submitted 01.06.24; peer-reviewed by J Lee, T Scherr; comments to author 25.11.24; revised version received 22.02.25; accepted 27.03.25; published 29.05.25.

Please cite as:
Liu AY, Torres TS, Alleyne CD, Vinson J, Bojan K, Serrano PA, Oyedele T, Garcia A, Enriquez-Bruce E, Emmanuel P, Jones J, Muessig KE, Buchbinder SP, Sullivan P, Hightow-Weidman LB, Scott H
Developing a Novel Mobile App to Support HIV Testing and Pre-Exposure Prophylaxis Uptake Among Men Who Have Sex With Men: Formative and Technical Pilot Study
JMIR Hum Factors 2025;12:e62830
URL: https://humanfactors.jmir.org/2025/1/e62830
doi:10.2196/62830
PMID:

©Albert Y Liu, Thiago S Torres, Cat-Dancing Alleyne, Janie Vinson, Kelly Bojan, Pedro Alonso Serrano, Temitope Oyedele, Elizabeth Enriquez-Bruce, Patricia Emmanuel, Jeb Jones, Kathryn E Muessig, Susan P Buchbinder, Patrick Sullivan, Lisa B Hightow-Weidman, Hyman Scott. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 29.05.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Designing a Visual Analytics Tool to Support Data Analysis Tasks of Digital Mental Health Interventions: Case Study

Gyuwon Jung[1], PhD; Heejeong Lim[2], MS; Kyungsik Han[3,4], PhD; Hyungsook Kim[3,5], PhD; Uichin Lee[1], PhD

[1]School of Computing, KAIST, Daejeon, Republic of Korea

[2]Graduate School of Data Science, KAIST, Daejeon, Republic of Korea

[3]Department of Data Science, Hanyang University, Seoul, Republic of Korea

[4]Department of Artificial Intelligence, Hanyang University, Seoul, Republic of Korea

[5]Hanyang Digital Healthcare Center, Hanyang University, Seoul, Republic of Korea

Corresponding Author:
Uichin Lee, PhD
School of Computing
KAIST
291 Daehak-ro, Yuseong-gu
Daejeon, 34141
Republic of Korea
Phone: 82 42 350 3544
Email: uclee@kaist.edu

Abstract

Background: Digital health interventions (DHIs) are widely used to manage users' health in everyday life through digital devices. The use of DHIs generates various data, such as records of intervention use and the status of target symptoms, providing researchers with data-driven insights for improving these interventions even after deployment. Although DHI researchers have investigated these data, existing analysis practices have been fragmented, limiting a comprehensive understanding of the data.

Objective: We proposed an analysis task model to help DHI researchers analyze observational data from a holistic perspective. This model was then used to prototype an interactive visual analytics tool. We aimed to evaluate the suitability of the model for DHI data analysis and explore task support using a visual analytics tool.

Methods: We constructed a data analysis task model using 3 key components (ie, user grouping criteria) for DHI data analysis: user characteristics, user engagement with DHIs, and the effectiveness of DHIs on target symptoms based on comparisons before and after the intervention. On the basis of this model, we designed Maum Health Analytics, a medium-fidelity prototype of an interactive visual analytics tool. Each feature of the prototype was mapped one-to-one to the analysis task described in the model. To investigate whether the proposed model adequately reflects real-world DHI analysis needs, we conducted a preliminary user study with 5 groups of researchers (N=15). Participants explored the tool through scenario-based analysis tasks using in-the-wild data collected from a mobile DHI service targeting depressive symptoms. Following the session, we conducted interviews to assess the appropriateness of the defined tasks and the usability and practical utility of the visual analytics tool.

Results: DHI researchers responded positively to both the analysis task model and the visual analytics tool. In the interviews, participants noted that the tool supported the identification of users who needed additional care, informed content recommendations, and helped analyze intervention effectiveness in relation to user characteristics and engagement levels. They also appreciated the tool's role in simplifying analytic tasks and supporting communication across multidisciplinary teams. Additional suggestions included improvements for continuity across tasks and more detailed engagement metrics.

Conclusions: We proposed an analysis task model and designed an interactive visual analytics tool to support DHI researchers. Our user study showed that the model allows a holistic investigation of DHI data by integrating key analysis components and that the prototype tool simplifies analytic tasks and enhances communication among researchers. As DHIs grow, the proposed model and tool can effectively meet the data analysis requirements of researchers and improve efficiency.

(JMIR Hum Factors 2025;12:e64967) doi:10.2196/64967

XSL•FO
RenderX

KEYWORDS

digital health interventions; visual analytics; data analysis tasks; user characteristics; user engagement; effectiveness; mental health; observational data; user experience; human-data interaction

Introduction

Background

Digital devices, such as smartphones and wearables, are extensively used in health care to deliver interventions beyond traditional medical settings [1]. These interventions, known as digital health interventions (DHIs), are designed to address various health issues and promote healthy behaviors, such as physical activity and smoking cessation, or to manage chronic conditions, such as depression and diabetes, in daily life [2,3].

DHIs use various intervention strategies, including behavior change techniques, which have proven to be effective in altering health behaviors [4-6]. These techniques include monitoring user behavior, setting goals, providing social support, and incorporating gamification. Furthermore, DHIs can be classified into different categories, such as digital health, medicine, and therapeutics, based on clinical evidence and real-world outcomes [7]. Previous studies [8-11] have investigated the delivery procedures of DHIs, identified barriers in their delivery paths, and determined the opportune moments for providing DHIs. All these studies share the common objective of effectively improving user behavior or symptoms targeted by DHIs.

DHIs are commonly designed as mobile apps, such as those in mobile health, which enable individuals to install and use them on smart devices. When users engage with the features and content offered by DHIs, various data are collected. For instance, several types of log data are passively gathered during user interactions with DHIs. Most of these log data consist of detailed records of user activities, including the frequency of DHI app use; individual DHI content accessed; and user interactions, such as taps, swipes, and text entries, within particular intervention content [12]. In addition, some data are actively provided by users through manual inputs, which are essential for capturing information that may be challenging to track using passive log data alone. These data often include basic user information, such as sociodemographic details, and periodic self-reports on their physical or mental health status.

Previous studies have performed various analyses to better understand the data collected from DHIs and assess either user engagement with the intervention content or the effectiveness of the interventions in supporting target users. Researchers studying DHIs have explored user engagement with interventions, assessing either subjective experiences about how immersed users are or objective behaviors, such as use frequency [11]. When quantitatively assessing engagement, researchers can analyze macroindicators, such as the number of log-ins to the DHI, frequency of content access, time spent, sequence of content use, as well as microindicators, such as the number of clicks and swipes [13-16]. A recent study [17] suggested similar metrics for measuring engagement, categorizing them into individual-level metrics, such as launch counts, use durations, and long-term use patterns, and population-level metrics, including the ratio of users who open the app at least once and the number of completers.

In addition, researchers have evaluated the effectiveness of DHIs on target symptoms by comparing health status before and after the use of DHIs in natural settings. The evaluation of DHIs typically progresses from measuring efficacy with a small group of users in controlled environments (eg, randomized controlled trials) to assessing effectiveness with a large group of users in uncontrolled environments [1]. By evaluating effectiveness without predefined treatment and control groups, researchers can investigate whether DHIs can yield the intended health changes or real-world evidence by analyzing real-world data (RWD), although numerous confounding factors exist [18,19]. Several metrics, such as user engagement and DHI effectiveness, can indicate how actively users have used DHIs and whether they have experienced improved health conditions, thus providing reasonable criteria for evaluating DHIs. A relevant example is the SilverCloud platform proposed by Doherty et al [20], which explores ways to encourage users to engage more actively in online mental health interventions. Using this platform, researchers have proposed approaches for visualizing log data to understand the engagement levels of individuals or groups of users [21] and predict the clinical outcomes of interventions using machine learning techniques [22].

Nevertheless, existing analyses of DHI data have often been performed separately based on individual researchers' interests, which limits their ability to understand data from various perspectives. Consequently, researchers may miss meaningful relationships among various indicators that can be extracted from DHI data. Moreover, the lack of a comprehensive analysis of DHI use patterns and changes in health status makes it challenging to determine how to improve DHIs according to the needs and preferences of specific user groups. Given that data analyses are repeated throughout the development and evaluation life cycle of DHIs [23], it is necessary to integrate the diverse analytical approaches that researchers can use with DHI data.

Objectives

Therefore, we proposed an analysis task model to help DHI researchers analyze data from a holistic perspective, enabling them to uncover interactions and patterns that are not visible when these factors are analyzed independently. We constructed this model based on common analysis practices observed in previous DHI research, particularly those identified in a meta-analytic review by Moshe et al [24].

To achieve this goal, we developed an analysis task model to help researchers explore DHI data from multiple perspectives. This model guided us to implement a prototype for an interactive visual analytics tool named Maum Health Analytics, which was designed to support the practical application of the model. The tool was evaluated using RWD collected from Maum Health,

XSL•FO
RenderX

a mobile DHI service for individuals experiencing depressive symptoms.

This study aimed to evaluate whether the proposed analysis task model is suitable for DHI data analysis and explore how visual analytics can support such tasks. Our preliminary user study identified several benefits of the Maum Health Analytics tool from the human-computer interaction perspective. We found that the proposed tool helped DHI research teams investigate user engagement and the effectiveness of the Maum Health DHI service from various angles, gain insights for recommending intervention content, and better understand users who may need additional care. Furthermore, Maum Health Analytics facilitated communication among various stakeholders, streamlined repetitive analysis tasks, and exhibited potential for integration with existing analytic systems.

Overall, our contribution can be summarized as follows:

1. We proposed an analysis task model for DHIs, aimed at assisting researchers in understanding DHI data from diverse perspectives.
2. We designed an interactive visual analytics tool based on the proposed model and explored its feasibility and design implications to better support researchers in analyzing DHI data.

Methods

This section describes the development of the analysis task model, the visual analytics tool, as well as the procedure of the preliminary user study conducted with the prototype. The subsequent sections present the quantitative and qualitative study results and discuss their implications for future research and the design of DHI analytics.

Analysis Task Model

Overview

As part of our previous work on DHIs [25], we conducted an in-depth analysis of RWD collected from a mobile DHI service named Maum Health, which supports individuals with depression. Through this process, we identified recurring analytical needs among DHI researchers, motivating us to design a dedicated visual analytics tool for this study. Within this service, users could assess their depression symptoms using the 10-item Center of Epidemiologic Studies Depression Scale (CES-D-10) questionnaire [26] and access intervention content, such as art therapy (named Mandala), physical activity promotion (named Geunsimtapa), and cognitive-emotional games (named Finding Blue) to alleviate symptoms. Following a 3-month deployment of the DHI, we analyzed depression states and content use logs stored on the server to extract data-driven insights to enhance the intervention service.

As a multidisciplinary DHI research team comprising intervention content designers, clinicians, and system developers, our analysis addressed several research questions, including differences in depression changes based on user characteristics (eg, initial depression level) and the association between DHI use and changes in depression. During the DHI data analysis, we made 3 key observations regarding the analysis tasks.

First, DHI researchers primarily focused on exploring user engagement with DHIs and the effectiveness of these interventions when evaluating existing DHI services. They determined that the DHI service can be considered as well designed if it is used more actively and if it leads to improvements in target symptom levels compared with before its use. In addition, given that the DHI was deployed to assess whether it supports the achievement of the intended outcome (ie, improving the depressive state) in the real world, researchers usually measure the effectiveness of a DHI instead of its efficacy. Unlike efficacy, which is typically measured in clinical trials involving random user assignments to control and treatment groups (controlled research setting), effectiveness is evaluated using data from observational studies, allowing researchers to observe how interventions are performed in natural settings (uncontrolled nonresearch setting) [1]. This measurement is relevant to RWD in health care and medical domains [18], where data are collected without randomly assigning participants to specific treatment conditions, and researchers conduct observational studies [27,28].

Second, researchers have attempted to investigate whether certain groups of users would exhibit higher (or lower) levels of the aforementioned metrics, aiming to examine the relationships between user groups and the metrics. They analyzed the engagement and effectiveness of user groups specified based on certain criteria and compared them across different groups [8]. Although these analyses are valuable, they are often conducted separately, limiting a comprehensive understanding of how different factors, such as user characteristics, engagement levels, and effectiveness, interact. This fragmented approach can result in missing critical insights from a holistic perspective.

To address this gap, we constructed an analysis task model that aimed to integrate these typical analysis task components and facilitate a more thorough understanding of DHI data. Specifically, this model was constructed based on a meta-analytic review by Moshe et al [24], which assessed multiple studies on DHIs developed to address depressive symptoms. The review outlined four factors that previous studies analyzed regarding their association with the effectiveness of DHIs: (1) the characteristics of participants, (2) the presence of guidance in using DHIs, (3) engagement with DHIs, and (4) study design and quality. Because several existing studies have performed similar analyses [24], we determined that these analysis tasks were common and should be integrated into the proposed model. However, we excluded two factors from the proposed model: the presence of guidance in using DHIs and study design and quality. This decision was made because our DHI service did not involve any human expert support, and we did not aim to demonstrate the efficacy of DHIs through clinical trials.

Consequently, we constructed our analysis task model with three key components: (1) user characteristics, (2) user engagement with DHIs, and (3) the effectiveness of DHIs on target symptoms. On the basis of these key components, we

XSL•FO
RenderX

identified a set of exploratory analysis tasks by referring to the practices of existing studies, as explained in subsequent sections. These tasks are intended to examine each component and explore how they relate to segmented user groups. For instance, the model supports not only the analysis of overall effectiveness across all users but also comparisons across user subgroups segmented by user characteristics or engagement levels. In this model, effectiveness is treated as an outcome variable to be examined within specific segments instead of as a factor used to define those segments. Although the model enables the exploration of the relationships among components within specific user segments, it does not assume causal or recursive dependencies. Instead, it provides a structured framework for performing an in-depth analysis of specific user segments.

The overall structure of the proposed model is illustrated in Figure 1, and further details of each component and the associated analysis tasks are described in the subsequent subsections and Table 1.

Figure 1. Overview of the proposed analysis task model for digital health intervention (DHI) data.

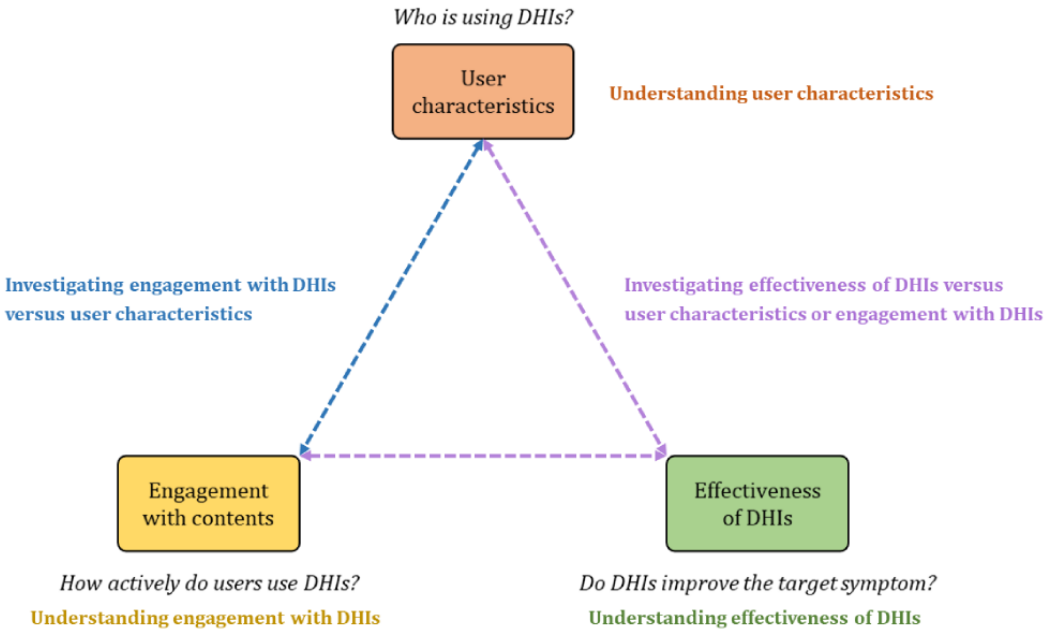


Table 1. Mapping of model-defined analysis tasks to interface, visualizations, and interactions in the Maum Health Analytics tool.

Model components, interface, and task IDs	Task description	Visualization and interactions
User characteristics (interface: User page)		
T1.1	Track the number of users over time	Line chart
T1.2	View user distribution by characteristics	Bar chart
T1.3	Explore individual user records	Table; filtering (by user characteristics)
User engagement with DHIs[a] (interface: Engagement page)		
T2.1	View engagement level across all users	Histogram
T2.2	View engagement level across user subgroups	Histogram; filtering (by user characteristics)
T2.3	Compare the engagement level among user subgroups	Histogram; group selection (by user characteristics)
T2.4	View user characteristics across engagement subgroups	Bar chart; filtering (by user engagement)
Effectiveness of DHIs (interface: Effectiveness page)		
T3.1	View effectiveness across all users	Histogram
T3.2	View effectiveness across user and engagement subgroups	Histogram; filtering (by user characteristics and engagement)
T3.3	Compare effectiveness across user and engagement subgroups	Histogram; group selection (by user characteristics and engagement)
T3.4	View user or engagement across effectiveness subgroups	Bar chart; filtering (by effectiveness)

[a]DHI: digital health intervention.

XSL•FO
RenderX

User Characteristics

Understanding the characteristics of DHI users is crucial because different user groups may lead to differences in DHI use behaviors and changes in target symptoms. As demonstrated in a large-scale cross-study evaluation by Pratap et al [29], user characteristics, such as gender, age, and geographic or racial or ethnic diversity, may influence the sustained use of DHIs. Similarly, other studies have shown that these characteristics could impact dropout rates from DHIs and changes in outcomes [30,31].

User characteristics for DHI data analysis can vary, including demographics (eg, age, gender, and ethnicity), physical states (eg, height and weight), and psychological states (eg, mood, depression, stress, and motivation) [11]. In addition, factors such as personality traits, digital health literacy, and the availability of time and space for using DHIs can be considered important constructs of user characteristics [9].

Regarding user characteristics, the proposed model comprises 3 analysis tasks, identified based on a review of existing DHI research and the authors' deployment experiences (as illustrated in the Methods section). First, researchers track the changes in the number of DHI users over time (task T1.1). This task is essential to understanding the overall adoption and sustained use of DHI services. Researchers can assess whether the DHI services are effectively used by monitoring the number of enrolled and active users. Next, researchers examine the distribution of DHI users based on their characteristics (eg, gender, age, and baseline symptom levels), as suggested in previous studies (task T1.2). This task involves sorting and filtering users based on various characteristics to explore the composition of the user population. It helps to identify trends and patterns in user engagement and the demographic reach of the DHI service. Finally, researchers review the detailed records of individual DHI users belonging to specific groups determined by user characteristics (task T1.3). This task allows for a more granular analysis, providing insights into the behavior and engagement levels of users within specific segments of the population. By following these tasks, researchers can gain comprehensive insights into who uses their DHI services through diverse visualizations, ranging from a broad overview to detailed individual records.

User Engagement With DHIs

User engagement is an essential metric in DHI research, as highlighted in previous studies. This component is crucial not only for understanding the current activeness of users interacting with DHIs but also for exploring strategies to maintain and enhance their sustained use [32,33]. Low engagement levels in DHIs are analogous to situations in which patients do not properly take medications. Therefore, this metric should be closely monitored to assist DHI users in maintaining their use until desired health outcomes are achieved.

As reviewed by Pham et al [34], previous studies have used various indicators to measure engagement levels, including the frequency of log-ins, accessed DHI features and modules, and duration of DHI use. Among these metrics, we included the frequency (ie, launch counts) and duration (ie, use time) of DHI content use in the proposed analysis task model, as they were commonly used in existing studies [11,35-37]. Furthermore, if a structured activity exists within each intervention, its completion level can be measured and evaluated as a detailed user engagement metric [38].

For user engagement with DHIs, we included 4 analysis tasks in the model. First, researchers assess the engagement level of each DHI content item across all users (task T2.1). This task involves evaluating various engagement metrics for each DHI content item, such as launch counts, use time, and completion levels, as highlighted in previous studies [11]. This provides an overview of how each content item is used.

However, considering that user characteristics may influence engagement with DHIs, researchers assess the engagement level of each DHI content item for user groups specified by user characteristics (task T2.2). This task allows for a segmented analysis based on attributes, such as age, gender, or baseline symptom levels, which helps understand how different user groups interact with DHI content. Furthermore, they compare the engagement level of each DHI content item across different user groups (task T2.3). For instance, to determine whether age affects engagement levels, researchers can compare the distribution of engagement metrics among DHI users of different age groups. This comparative analysis helps identify patterns and variations in engagement across user characteristics.

This group-level analysis can also be performed in reverse; researchers assess the user characteristics distribution for a user group where they attain a certain level of engagement (task T2.4). By performing these analytical tasks, DHI researchers can evaluate the extent to which DHI content is actively used, identify differences in engagement levels among various user groups, and understand the distribution of user characteristics at certain engagement levels.

Effectiveness of DHIs

As the primary objective of DHIs is to improve target symptoms, evaluating their effectiveness is necessary. To evaluate the effectiveness of DHIs in real-world settings, researchers analyze the changes in depression symptoms by comparing self-reported symptom levels before and after DHI use. This pre-post comparison allows us to determine whether the intervention has led to a statistically significant improvement in target symptoms, thus estimating the effectiveness of DHIs. As explained in the previous section, effectiveness is measured in an uncontrolled setting to observe how the DHI influences the target symptoms in natural environments. For depression assessment, the following instruments are widely used [39,40]: the 9-item Patient Health Questionnaire [41], the CES-D [42], the Beck Depression Inventory—II [43], and the Patient-Reported Outcomes Measurement Information System [44]. These assess depressive symptoms through items targeting self-reported negative emotions, self-perception, social interaction, and diminished positive affect.

Similar to the analysis of user engagement with DHIs, the analysis task model included 4 tasks for evaluating the effectiveness of the DHIs. The first task is to assess the changes in depression levels across all users based on their self-reported

XSL•FO
RenderX

depression states (task T3.1). This task provides an overall measure of the impact of the DHI on target symptoms.

The next task is to assess the changes in depression levels for the user groups (task T3.2) and compare these changes across different user groups (task T3.3). As illustrated previously, user characteristics and engagement levels can influence the effectiveness of DHIs. Hence, the user groups in tasks T3.2 and T3.3 can be specified based on user characteristics or the engagement level of each DHI content item. For instance, effectiveness can be evaluated for a user group comprising women who use a specific DHI content for more than an hour, or it can be compared among user groups divided by age. This segmented and comparative analysis helps to identify how the effectiveness of DHIs varies across user characteristics and engagement levels.

Group-level analysis can be performed in reverse, such as task T2.4; researchers assess user characteristics and engagement level distribution for a user group where they attain a certain change in depression levels (task T3.4). This task is crucial for understanding user characteristics and engagement behaviors associated with significant changes in depression levels. These tasks assist researchers in exploring the effectiveness of DHIs from multiple perspectives. They allow for the analysis of effectiveness while considering moderating factors, thereby enabling a nuanced understanding of how various user characteristics and engagement levels influence outcomes. Moreover, they provide methods to tailor DHIs to different users based on the observed effectiveness across diverse user groups.

Interactive Visual Analytics

We designed Maum Health Analytics, a prototype of an interactive visual analytics tool, to facilitate researchers in conducting DHI data analysis tasks from multiple perspectives. This prototype was developed based on the analysis task model proposed in the previous section, with each feature directly mapped to a specific analytical task defined in the model.

The primary goal of this prototype was to evaluate whether the proposed model adequately supported the typical analysis workflow used by DHI researchers. To achieve this, we developed a medium-fidelity prototype that implemented the tasks defined in the model. Rather than building a complete system, we focused on creating a functional prototype suitable for the initial evaluation.

The prototype was created using Figma and designed to enable users to explore DHI data through task-specific features embedded in predefined analysis scenarios. The development followed an iterative design process. We created an initial version that mapped model-defined tasks to concrete interface features and functionalities. This prototype was thereafter refined through multiple feedback sessions with 2 domain experts who were clinicians with experience in developing their own DHIs and analyzing real-world DHI data. They participated in expert heuristic evaluations and provided feedback on the clarity of the task flow, adequacy of the visual representations, and interpretability of the analytic results. On the basis of their input, we made several adjustments to the prototype to improve its usability and relevance. In the final version, we incorporated the actual analysis results derived from the RWD collected through our previously deployed DHI, Maum Health. This addition provided users with a realistic context during the preliminary user study and allowed us to better evaluate the usability of the tool.

Each analysis task defined in the model was mapped to a corresponding interface component and visualized in the prototype. Table 1 summarizes the mapping, including the key visual elements and interaction types used to support each task.

Consequently, Maum Health Analytics features 3 main pages. Each page corresponds to 1 of the 3 key analysis components: user characteristics, user engagement with DHIs, and the effectiveness of DHIs. On each page, researchers can select the specific user group conditions that they wish to investigate, and the tool displays the analyzed results interactively. These results are primarily presented through visual elements, such as bar charts, line charts, tables, and histograms, aiding researchers in quickly grasping the overall trends or differences. Moreover, detailed analysis results, such as statistical testing, are provided together to assist researchers in interpreting these findings.

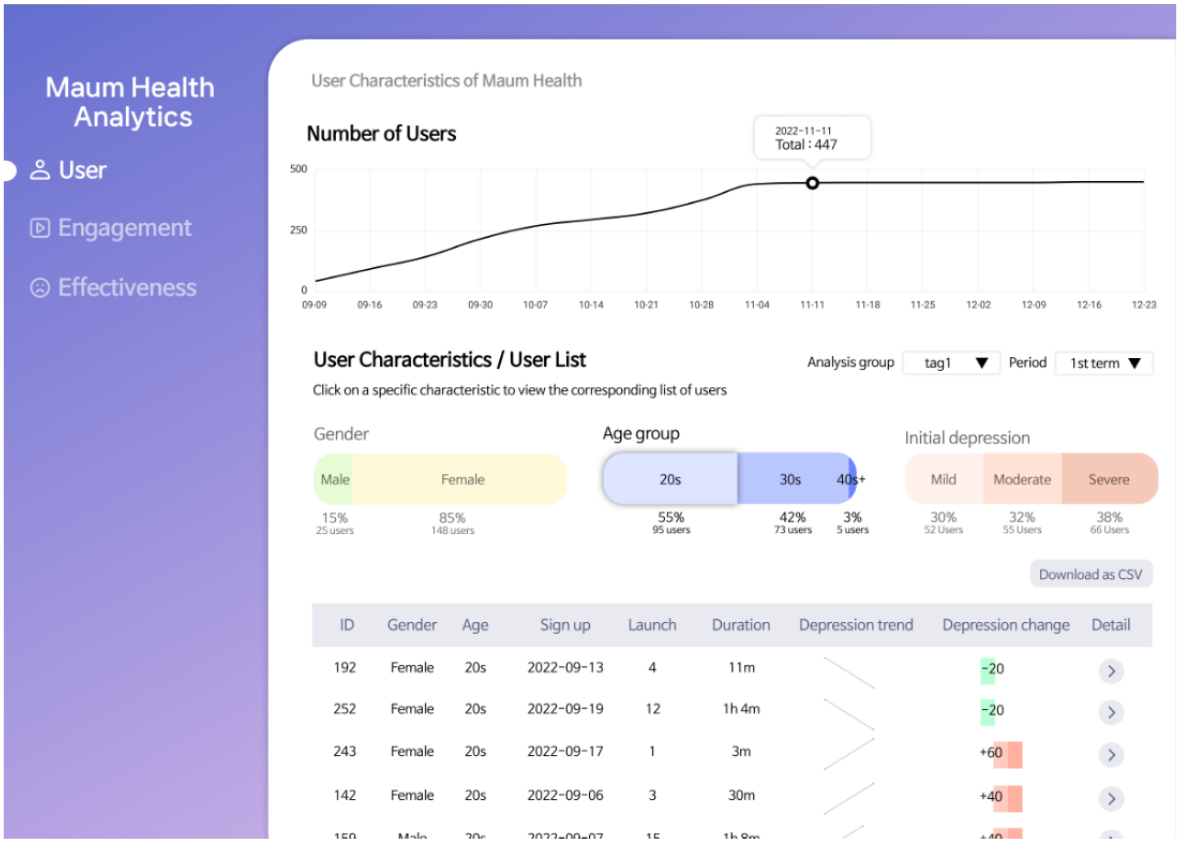
To enhance usability, we included a tag feature throughout the visual analytics tool. Tagging allows the development of researcher-defined groups. For instance, a researcher can create a tag for a user group characterized by high engagement levels and severe initial depression. Once saved, these tags enable quick and easy access to researcher-defined user groups and function as custom favorite lists. This allows researchers to efficiently revisit and analyze the same user groups without having to redefine the conditions, streamlining the data analysis process with Maum Health Analytics.

In the subsequent section, we provide brief explanations for each page of the Maum Health Analytics tool.

The User page (Figure 2) supports tasks associated with user characteristics, presenting visualizations of the number of DHI users over time (task T1.1) and the distribution of users by characteristics (task T1.2). Researchers can select specific user characteristics from the horizontal bar chart in the middle; users matching these criteria are displayed in the table that appears under User Characteristics and User List. Furthermore, they can select a specific user from the table to review detailed records (task T1.3), including the user's basic information, engagement levels with DHIs, changes in depression levels, and individual DHI content use records.

XSL•FO
RenderX

Figure 2. The User page of the Maum Health Analytics tool that supports tasks associated with user characteristics.

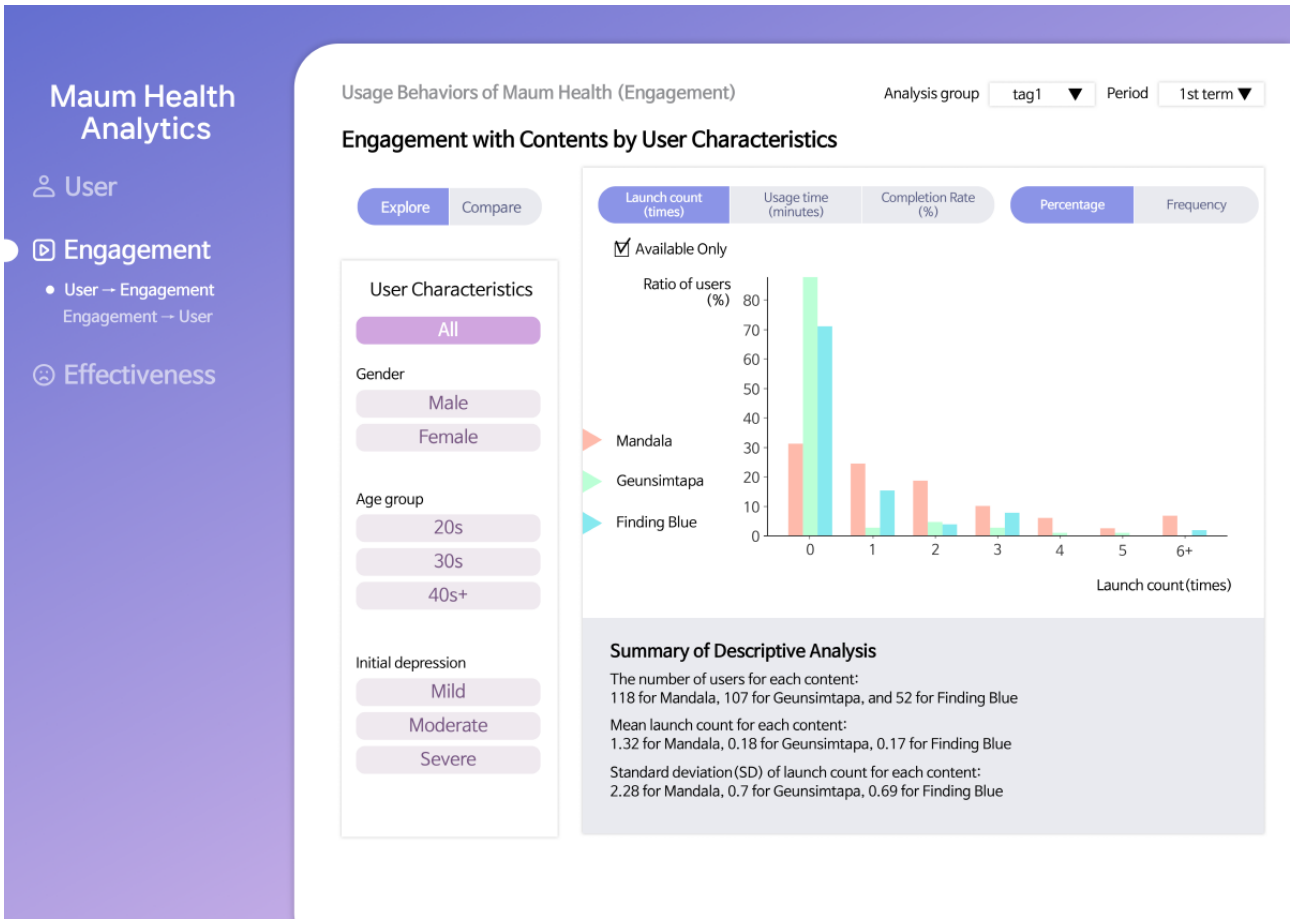


Line charts were selected to effectively show temporal trends in user participation (task T1.1), whereas bar charts allowed for a clear comparison of categorical variables, such as age or gender (task T1.2). The use of a table enables detailed record-level exploration for individual users (task T1.3), thereby supporting granular analysis.

On the Engagement page (Figure 3), researchers can explore the engagement level of each DHI content item across all users (task T2.1) or within selected user groups (task T2.2) specified by user characteristics. As shown in the figure, the histogram illustrates the distribution of engagement levels within the selected user group, providing insights into the overall use behavior for each DHI content item. In addition, researchers are allowed to compare the engagement level across different user groups (task T2.3) to identify the users who are more engaged with each DHI content item. Furthermore, user groups can be formed based on specific engagement levels to examine the user characteristic distribution within these groups (task T2.4).

XSL•FO
RenderX

Figure 3. The Engagement page of the Maum Health Analytics tool that supports tasks associated with user engagement with digital health interventions (DHIs).

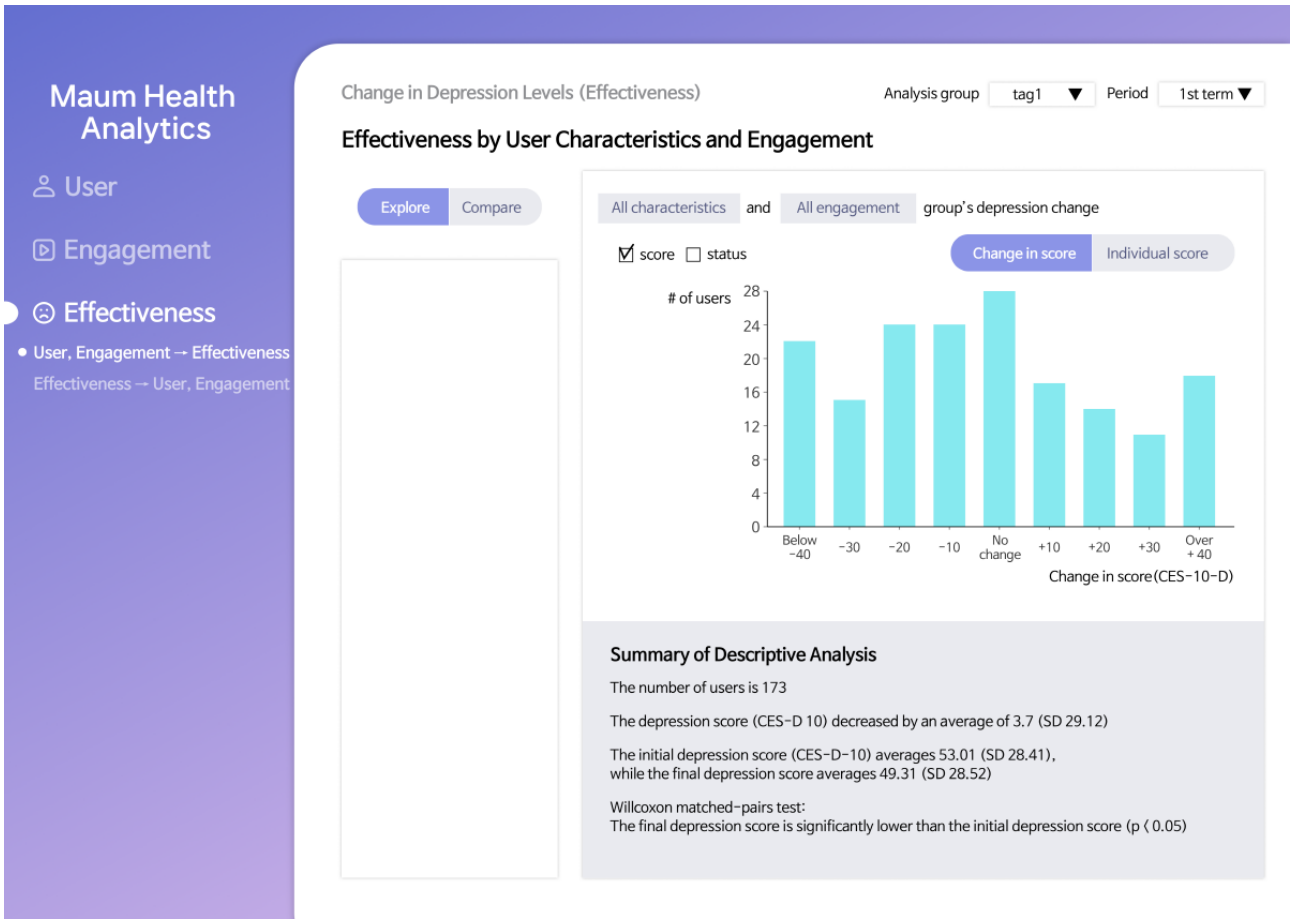


Histograms were used to visualize engagement distributions across user groups (tasks T2.1 to T2.3) as they effectively represented frequency patterns and enabled quick comparisons among subgroups. Bar charts were selected for task T2.4 to show how demographic characteristics vary by engagement level, thereby supporting the group-based analysis.

Finally, researchers can explore the changes in depression levels on the Effectiveness page (Figure 4) either across all users (task T3.1) or within selected user groups (task T3.2). When specifying user groups for investigation, researchers are allowed to select user characteristics or user engagement with the DHI content, and the results are shown in a histogram. Similar to the Engagement page, researchers can compare changes in depression levels across different user groups (task T3.3). In addition, they can create user groups based on depression changes to understand the distribution of user characteristics and engagement levels within these groups (task T3.4).

XSL•FO
RenderX

Figure 4. The Effectiveness page of the Maum Health Analytics tool that supports tasks associated with the effectiveness of digital health interventions (DHIs) on the target symptom.



Histograms were used to visualize changes in depressive symptoms before and after the intervention (task T3.1) as they are well suited for showing the distribution of continuous outcome variables and identifying general patterns of improvement or decline. For tasks T3.2 and T3.3, histograms were used to view or compare symptom change distributions across subgroups according to user characteristics or engagement, allowing for an intuitive visual comparison of group-level outcome variations. For task T3.4, bar charts were selected to summarize user characteristics and engagement within the effectiveness-based subgroups, thus supporting the identification of traits associated with greater or lesser effectiveness.

Further details about the Maum Health Analytics tool are provided in Multimedia Appendix 1.

Preliminary User Study

Maum Health Dataset

As previously mentioned, we decided to provide the analysis results derived from the data collected through our DHI service, Maum Health, when evaluating the interactive visual analytics. Here, we provide a brief description of Maum Health and the data it collected.

Maum Health

Maum Health is a DHI service developed to improve depressive symptoms and is offered as a mobile app. Similar to typical mental health mobile apps, Maum Health provides various intervention content beneficial for depressive symptoms, including art therapy (Mandala), physical activity (Geunsimtapa), and a cognitive-emotional screening game (Finding Blue). Each intervention content item consists of sessions comprising activities that users can perform on their own. For instance, users can make color drawings, perform walking and stretching exercises, and play interactive games when engaging with each content item. Moreover, it assesses users' depression levels every 2 weeks using the CES-D-10, a well-established and validated instrument for assessing depressive symptoms in mental health research. The results are converted to a 100-point scale, where a higher score indicates more severe depression.

Dataset

The Maum Health DHI service was distributed through a public counseling center located in Seoul, South Korea, which operates under the Seoul Metropolitan Government. The center is situated in an urban residential area with high accessibility via public transportation and provides mental health counseling services specifically for young adults, supported by licensed psychiatric professionals. As part of the counseling process, the center recommended Maum Health as a supplementary digital tool for managing mental health. Consequently, this DHI was used by 529 people over approximately 3 months, starting in September 2022. During this period, self-reported data entered by the users

XSL•FO
RenderX

and log data automatically recorded based on the use of the intervention content were collected.

The self-reported data in Maum Health consisted of one-time basic user information and periodically recorded levels of depressive symptoms. When users first registered for Maum Health, they provided demographic information, such as gender and age, along with information that might be related to depressive symptoms (eg, marital status and alcohol or smoking experience). Moreover, while using Maum Health, users reported their depressive symptom levels biweekly through the CES-D-10 survey.

As users engaged with the 3 different types of intervention content available on Maum Health, log data were recorded automatically. Whenever a user finished a session with specific content, the time stamps for the start and end points of the session were recorded. Furthermore, for each session, the completion rate was recorded to indicate how well the user performed the given activity. We used features from both self-reported and log data in the design of the Maum Health Analytics tool, as summarized in Table 2.

Table 2. Features of the Maum Health DHI service data used in the design and evaluation of the Maum Health Analytics tool.

Analysis components and categories	Features
User characteristics	
Basic information	Gender, age group, and initial depression level
Additional information	Marital status, cohabitant, occupation, education, economic status, drinking, smoking, army experiences, and handedness
Medical history	Depression history, medication, and physical illness
User engagement with DHIs[a]	
Mandala (art therapy)	Total launch counts, total use time, and average completion rate
Geunsimtapa (physical activity promotion)	Total launch counts, total use time, and average completion rate
Finding Blue (cognitive-emotional games)	Total launch counts, total use time, and average completion rate
Effectiveness of DHIs	
Depression level	CES-D-10[b] score

[a]DHI: digital health intervention.

[b]CES-D-10: 10-item Center of Epidemiologic Studies Depression Scale.

Considering the potential quality issues with RWD, we established inclusion and exclusion criteria for the data to be used when evaluating the Maum Health Analytics tool. We selected users who had at least 2 depression score records, which allowed us to observe changes in depression levels. Among them, we included users with a gap of 2 to 4 weeks between evaluations, as they were considered to maintain their DHI use sufficiently well. Consequently, we used data from 173 (32.7%) of the 529 users. Recognizing that the initial use time of the Maum Health DHI service could vary among users, we considered the relative use period starting with each user's first day of use. As a result, the final dataset used to evaluate the Maum Health Analytics tool included 3470 data points, comprising user-provided demographic information, preprocessed log records of DHI content use, and periodic self-reports of depressive symptoms.

Study Procedure

Participants

To evaluate the Maum Health Analytics tool, we recruited 5 groups of experts with experience in analyzing DHI data. Each group comprised 3 to 5 experts, including clinicians, intervention content designers, and system developers, who collaborated as a team. In total, 15 DHI researchers participated in the study (woman researchers: n=6, 40% and man researchers: n=9, 60%), who were aged between 26 and 44 years (mean 34.1, SD 6.4 years). All participants had previous experience in developing DHIs and analyzing real-world DHI data. Although they had not used the Maum Health DHI service as end users, they became familiar with its structure and features through their involvement in related research and evaluation activities. The details of these expert groups are listed in Table 3, and we will refer to these experts as "DHI researchers" hereinafter.

Table 3. Composition of participants by digital health intervention research teams.

Group IDs	Participant IDs	Expertise of the group
A	A1, A2, and A3	Clinicians and intervention content designers
B	B1 and B2	Clinicians and intervention content designers
C	C1, C2, C3, C4, and C5	Intervention content designers
D	D1, D2, and D3	Clinicians, intervention content designers, and system developers
E	E1 and E2	System developers

XSL•FO
RenderX

Evaluation

Textbox 1 provides an overview of the evaluation procedure, comprising the research introduction, scenario-based use session, qualitative interviews, and quantitative usability assessment. During the user study, we first briefly introduced our research and explained the Maum Health Analytics tool, along with the dataset used for the evaluation, to the participants. Then, we asked them to explore the various features of our tool. Given that our tool was a medium-fidelity prototype, the participants were only able to navigate predefined functions, interactions, and corresponding analysis results.

Textbox 1. Overview of the study procedure for evaluating the Maum Health Analytics tool.

Research introduction

- The authors introduced the study and provided an overview of the Maum Health Analytics tool.

Scenario-based use session

- Participants explored the Maum Health Analytics tool for 30 min.
- Realistic data analysis scenarios were provided to support system evaluation during this session.

Qualitative evaluation

- Semistructured interviews were conducted for 60 to 90 min.
- Participants evaluated the analysis tasks and discussed their applicability in real-world settings.

Quantitative evaluation

- The Post-Study System Usability Questionnaire was administered to assess the system's usability.

In the evaluation session, we presented participants with specific DHI data analysis scenarios, enabling them to test all the tasks supported by Maum Health Analytics. For instance, if the scenario was "The expert investigates the distribution of user characteristics among those who frequently used Mandala above average," the participants would follow several steps within the tool to analyze the data. Considering the limited functionality of the prototype, these scenarios were selected based on cases with sufficient Maum Health data to showcase the results completely. We visualized the analysis results from real-world DHI data to make the evaluation realistic, ensuring that the participants could interact with and assess the prototype in a meaningful context.

After the participants used the Maum Health Analytics tool, we conducted an interview session and asked key questions from the perspective of analyzing the Maum Health DHI service data: (1) Were the analysis tasks defined in our tool appropriate? (2) How could the information provided be used in practice? Each session lasted approximately 60 to 90 minutes, depending on the participant group's pace and discussion depth. All interview sessions were recorded with the participants' consent and transcribed to thoroughly examine their responses. Thereafter, we performed inductive analysis [45] while repeatedly reading the interview transcripts to identify key phrases, ideas, and themes. Next, we conducted affinity diagramming to group similar themes derived from the transcripts and reviewed the themes iteratively until all researchers agreed on the final themes.

We also used the Post-Study System Usability Questionnaire (PSSUQ) [46], which is based on a 7-point Likert scale, to evaluate the overall usability. The PSSUQ is specifically designed for scenario-based usability studies and includes a subscale for information quality, allowing us to quantitatively assess whether the defined analysis tasks are suitable. According to Lewis [47], the PSSUQ scores for real-world systems commonly ranged between 2.5 and 3.0 across subscales, with lower scores indicating better usability. Although no absolute benchmark has been proposed, scores within this range have been interpreted as indicating acceptable usability in practical settings. We used the Korean-translated version of the PSSUQ previously developed and validated by Jeon and Park [48]. The participants in our study did not report any difficulties in interpreting or responding to the translated version during the usability evaluation.

We analyzed the responses using descriptive statistics and computed the mean and SD for each of the 3 PSSUQ subscales: system usefulness, information quality, and interface quality. Two items related to error recovery (items 7 and 8) were excluded because the user scenarios used in this study did not involve any error-handling tasks.

Ethical Considerations

This study was approved by the Institutional Review Board of the KAIST university (KH2022-098), and written informed consent was obtained from all participants.

Results

Overview

The previous section introduced the analysis task model and visual analytics tool developed for DHI research. On the basis of them, we conducted a preliminary user study to evaluate the usability of the system. This section presents the results of this evaluation, including both quantitative and qualitative findings.

Quantitative Evaluation

The PSSUQ comprises 16 items across 3 subscales (ie, system usefulness, information quality, and interface quality) rated on

XSL•FO
RenderX

a 7-point Likert scale (1=strongly agree, 4=neutral, 7=strongly disagree), with lower scores indicating better usability.

For the Maum Health Analytics tool, the mean score of the PSSUQ was 2.89 (SD 1.02), and that of the overall usability was 2.80 (SD 1.57). Each subscale had similar mean scores: system usefulness 2.94 (SD 1.01), information quality 2.87 (SD 1.21), and interface quality 2.84 (SD 1.31). These scores fell within the range commonly observed in previous usability studies (2.5-3.0) and are generally interpreted to reflect acceptable usability levels [47].

Qualitative Evaluation

Analysis Task Model

Our interviews revealed participants' opinions on the analysis task model and its components proposed in this study.

User Characteristics

Participants found that tasks related to user characteristics helped enhance their understanding of user groups, particularly those sharing specific characteristics. They were interested in examining individual use records as concrete examples of the findings identified at the group level. By reviewing representative users within a group, they can understand broader trends and verify their interpretations:

> *When I discovered that females used Mandala more, I became curious about each female user's usage patterns. Reviewing their specific usage logs could help me identify common characteristics of this user group.* [A1]

The participants further noted that such an in-depth exploration was challenging in their previous practice, particularly when user interactions were recorded only on personal devices and were not easily accessible:

> *We could analyze DHI usage statistics by group, but it was challenging to check detailed usage logs since we had to take their phones for analysis.* [A3]

These comments highlighted that visualizing individual-level use data supported a deeper and more nuanced understanding of group-level patterns.

In addition, the participants reported that tasks in this category were useful for identifying DHI users (or user groups) who may require additional attention. They emphasized the importance of evaluating whether users met a minimum level of engagement (in launch counts or use time) within a given period and examining specific users who fell short of these thresholds:

> *For instance, if there is a weekly recommended use, I'd like to see how many users in the group meet that goal.* [B1]

The participants reviewed group-level engagement metrics and paid particular attention to individuals with low engagement within those groups. Among the participants, clinicians focused on improving users' target symptoms and discussed how such insights could inform intervention decisions:

> *I was wondering why this user didn't use Maum Health contents that much, and why her status got worse even while using intervention contents.* [A1]

They planned to intervene with users with low engagement to increase their engagement:

> *Perhaps I need to decide how to manage the low engagement users, for instance, send a text message, make a call, and so on.* [B1]

Some participants also mentioned that tagging could help monitor users who required closer attention:

> *In our study, there are individuals whom doctors have to care for more intensively. Previously, we had to find such users whenever we analyzed them. But here, I think it would be easier and more convenient because we can mark them with tags and monitor them accordingly.* [C2]

E1 viewed tags as customizable user characteristics and noted that using them as favorites made it easier to track certain users. Together, these findings suggested that being able to review individual user records within user characteristic tasks helped to identify users at both the group and individual levels who might require further intervention to improve their symptoms.

The participants also mentioned that tasks supported by Maum Health Analytics could help track new incoming users and monitor the retention of existing ones. They noted that understanding the characteristics of both groups was essential for delivering appropriate intervention content:

> *There is a difference between user groups by their retention rate. Based on that information, we could determine the frequency of exposing specific content for each user group, which might be available here, too.* [D1]

However, participants emphasized the need to distinguish "active users" from merely registered ones, identifying this as a potential area for improvement. For intervention content designers, knowing who actively engaged with the system was considered more important than simply tracking overall registrations:

> *I'd like to see how many users are actively using DHI, not just registered. For designers, it is important to see whether they are using DHIs well enough.* [C1]

To better determine active use, the participants suggested including session-level dropout information on individual user pages:

> *I think the quitting ratio while using content is also important. The intervention content may be interesting at the beginning, but it gets boring as time goes* [C3]

These findings implied that supporting both retention tracking and active user identification contributed to a comprehensive understanding of DHI users and facilitated effective user management planning.

User Engagement With DHIs

Participants reported that Maum Health Analytics could help them identify user groups that exhibited high engagement with

XSL•FO
RenderX

specific DHI content, serving as useful information for making more effective content recommendations. Tasks related to user engagement, such as task T2.4, provided a distribution of user characteristics among highly engaged users. For instance, when extracting users who used Mandala above the average level, participants observed an increased proportion of woman users, which they found helpful in decision-making:

> *After examining the characteristics of users who use a particular content often, we can recommend it to similar users, such as those in the same age group. Maybe we can promote the DHI service by targeting similar groups of people through social media.* [A1]

The participants assumed that users with similar characteristics would demonstrate similar engagement levels. They considered this assumption valuable for improving recommendation strategies for specific groups of users, particularly because DHI content is typically recommended based solely on target symptom severity (eg, depression scores) without considering user preferences or engagement patterns:

> *I think this tool can be used to make feedback more customized while monitoring the changes in DHI engagement.* [A2]

Conversely, some participants indicated that users with low engagement should be further examined to design new content tailored to them. For instance, a participant noted the following:

> *If the proportion of females in users with high engagement is significant, we should create another content targeting men. Otherwise, the male user group may not be supported well and might leave the Maum Health app.* [A3]

However, B1 noted that recommendations based solely on engagement should be used cautiously and that the actual improvement in depressive symptoms should also be considered. These findings demonstrated that examining engagement patterns in conjunction with user characteristics can support more personalized and effective content recommendations while also highlighting the importance of considering intervention effectiveness alongside DHI use behavior.

Some participants mentioned that presenting the overall distribution of engagement with the intervention content was useful, as it could help establish guidelines for DHI use. They noted that interventions such as traditional cognitive behavioral therapy can be easily evaluated in terms of engagement because the number of sessions is predefined. However, setting evaluation criteria is challenging when users use the content independently without specific guidelines.

After exploring the overall engagement levels and those of specific user groups through the Maum Health Analytics tool, participants expected to establish engagement guidelines using this tool as more DHI data accumulated over time:

> *If the data becomes sufficient, values such as average or top 25% could be more meaningful. Observing the overall engagement level with this tool, we can provide engagement guidelines for each step or set a recommended use.* [B1]

This indicated the potential for establishing optimal use guidelines by analyzing long-term DHI engagement patterns, which currently lack established benchmarks.

In addition, the participants mentioned the need to evaluate user engagement more rigorously. In particular, the system developer group emphasized the importance of selecting appropriate use metrics, such as launch counts and use time, when measuring engagement. Compared with other groups, they primarily focused on improving usability and maintaining overall engagement levels rather than evaluating the effectiveness of DHIs. Consequently, they prioritized analyzing detailed content use logs and identifying metrics that more accurately reflected the actual user behavior:

> *For example, there are cases when users stop using content when getting a call or eating. In those cases, they stayed in the content but there was no interaction recorded in the usage log.* [E1]

The participants also highlighted the importance of distinguishing what constitutes "effective" engagement, thereby emphasizing the need to focus on use behaviors that indicate meaningful interactions with content:

> *It is necessary to analyze how much users actively participated in the content before investigating its effectiveness. Then, we have to further examine what was behind if DHI turned out to be ineffective.* [E2]

This implied that measuring engagement in DHIs should go beyond basic metrics, such as launch counts or use time, and instead incorporate indicators that more accurately capture users' actual interaction behaviors.

Effectiveness of DHIs

Participants responded that the Maum Health Analytics tool was helpful in systematically evaluating the effectiveness of DHIs by examining changes in target symptoms. Because DHIs are intended to support users' mental health, the participants found it meaningful to analyze symptom improvement in relation to engagement with each type of intervention content:

> *Of course, showing individual indicators can be a meaningful analysis for almost all apps. However, for DHIs, we should investigate which intervention features interact with each other and how they improve symptoms, requiring a relational analysis of multiple factors.* [D1]

> *For example, for game apps, the longer the usage time, the better. But in the case of DHIs, we need to see not only how long it has been used, but also how effective it is. I think this tool supports this process well enough.* [E1]

The participants compared Google Analytics with our tool, highlighting a key difference in their purposes. They explained that the key distinction of the proposed analytics tool was the inclusion of health-related indicators, making it more suitable for analyzing DHI services:

> *Unlike Google Analytics, which mainly focuses on usage behavior and user retention, we can analyze*

XSL•FO
RenderX

> *DHI including the users' health indicators in this tool. This is critical for health-related services.* [D1]

This suggested that evaluating the effectiveness of DHIs required integrating health outcomes with use metrics and that Maum Health Analytics enabled such multidimensional analysis more effectively than existing analysis practices.

We observed that evaluating content effectiveness from multiple perspectives using Maum Health Analytics was considered beneficial by participants, particularly in the context of designing or studying the intervention content. They explained that the tool simplified comparisons that would otherwise require repeated statistical analyses and manual setup under each condition:

> *Previously, it was inconvenient to repeatedly perform statistical analysis for each condition, so we sometimes skipped examining some questions in detail. Perhaps, we lost opportunities to find something new.* [C1]

By supporting comparisons based on user characteristics and engagement levels, the tool enabled users to explore effectiveness across different subgroups more efficiently:

> *Like the analysis in research, I was able to create control groups and experimental groups based on several criteria and easily compare the difference in depression change between the groups. I could see the relationship between engagement and effectiveness through this process.* [C2]

The participants also highlighted the tag feature as a practical way to track and compare user-defined groups. This allowed them to monitor symptom changes in relation to factors such as the app version or intervention updates:

> *I like the tag because we can immediately see how much DHI has improved after updates if we label tags based on the version of DHI.* [C1]

C1 added that using tags to compare groups could help the research team formulate and evaluate new hypotheses, similar to A/B testing.

In addition, participants found it useful to evaluate other indicators based on the change in depression levels. As DHI services are still in their early stages, no clear evidence exists that increased use leads to better outcomes. Therefore, B1 suggested that it might be more insightful to investigate other indicators based on the effectiveness of the DHI. These findings suggested that enabling flexible group comparisons and supporting exploratory evaluations of DHI effectiveness can help researchers uncover new patterns, generate hypotheses, and refine intervention strategies based on subgroup insights.

Interactive Visual Analytics

We also explored participants' experiences with the interactive visual analytics tool during the interview sessions.

Simplifying the Analysis Tasks of DHI

Participants mentioned that using the Maum Health Analytics tool helped simplify several of the manual and repetitive tasks they typically faced in DHI data analysis. In particular, they appreciated not having to download the entire dataset each time, recognizing this as one of the most practical features:

> *Our lab is developing a similar DHI app, but we have to download and analyze the entire dataset since there's no analytics tool like this. But here, we can check what happened to the user right away without such tasks.* [A3]

Because most existing dashboards were built for experiment management instead of analysis, participants found it helpful that our tool efficiently supported basic analytical tasks.

Participants also appreciated the ability to view analysis results in real time:

> *I have experience in developing chatbot-based interventions, and I wanted to know in real-time whether users are using them well or how they respond. I was satisfied to see that information on this tool.* [A3]

> *Real-time information may be meaningful to those who provide counseling services based on analyzed results, those who manage certain groups of users, and those who monitor the entire service.* [B1]

The participants also appreciated that the tool allowed them to extract and export user groups for further analysis. In particular, A2 and A3 highlighted the usefulness of isolating specific subgroups, such as outliers with unusual engagement or symptom patterns, and downloading their data to examine them in more detail using external analytical tools. These findings highlighted how Maum Health Analytics can reduce the burden of manual data handling and facilitate a more efficient and targeted analysis workflow.

Supporting Communication Among Various Experts

The participants assessed that the Maum Health Analytics tool would facilitate communication among stakeholders during the DHI improvement process, particularly when experts from different domains collaborate. They emphasized that the tool could bridge the gaps between those with differing levels of expertise in data analysis and interpretation.

Previously, content designers had to create separate documents to explain analysis results to other domain experts, which they found to be inefficient:

> *As a content designer, I always had to prepare additional materials about the analysis results to explain it to mental health counselors. However, with these visual explanations, I think they can understand the results easily without any further documents.* [A3]

They added that even stakeholders unfamiliar with data analysis could interpret the visual results and answer relevant questions using Maum Health Analytics:

> *Since this tool showed the results visually, I thought it would be suitable for those who are not familiar with how to analyze data well. With this visual analytics tool, they would be able to answer various questions related to DHI usage.* [A2]

XSL•FO
RenderX

Participants also noted that in cases where data analysis and user recruitment were handled by separate individuals or teams, the tool could help track and understand the user base:

> *If someone else is recruiting DHI users, there will be difficulties in the analysis since we have to examine the data without understanding the participants. This tool displays that information right away, making it easy to track DHI users' information even if we don't recruit the participants by ourselves.* [C1]

These findings suggested that Maum Health Analytics could serve as a shared platform to enhance communication and understanding among DHI stakeholders with diverse roles and technical backgrounds.

Providing Statistical Analysis and Interpretation of Results

The participants reported that the Maum Health Analytics tool was useful in providing statistical analysis results when evaluating engagement and effectiveness. They emphasized that statistical testing added an objective basis for interpreting and presenting findings:

> *There must be people who think statistics are important. For example, in the public domain of managing these studies, it is important to leave a numerical rationale.* [B1]

> *If we say that this result is statistically significant, I think we can clearly communicate it to experts in other fields.* [C3]

The statistical analysis was meaningful, particularly because most of the results were presented visually in graphs. However, the participants suggested that these results should be accompanied by more insightful explanations to better support DHI improvement. Although descriptive statistics, such as means and SDs, can be roughly interpreted from visualizations, some participants noted that providing statistical testing results along with practical interpretations would be more beneficial:

> *When considering people who are not familiar with statistics may see statistical analysis, it would be nice to provide interpretations related to decision-making, such as what this result means and how to utilize it in updating DHIs.* [C1]

These findings indicated that providing statistical results together with clear, practical interpretations could improve the usability of analytical tools and help stakeholders better understand and apply insights.

Enhancing Continuity Across Analysis Tasks

Participants commented that the transition between different analysis tasks could be improved by allowing the currently selected user groups to be pinned. They expressed the need to fix a group of interests and view all subsequent analysis results for that group:

> *For example, if I went to another page after checking the engagement distribution of Mandala, I wanted to view the analysis results of the same user group first.* [B1]

One participant specifically emphasized the need for a feature that allowed saving filtered user groups:

> *We first narrowed the user group down by adding more filters, but we had to find and select the conditions again if we moved to other pages. As a result, the analysis tasks seemed somewhat disconnected.* [E1]

They suggested that the ability to save and pin a specified user group could serve as a bridge between analysis tasks, enabling a more seamless experience when working with DHI data. They expected this feature to help researchers maintain continuity across tasks and pages by preserving their focus on the same user group. This implies that ensuring continuity between the analysis steps is a critical design consideration for supporting exploratory workflows in DHI data analysis.

Discussion

Principal Findings

Overview

We proposed an analysis task model for practitioners to explore and analyze RWD generated by DHIs. Building on previous studies related to DHIs, we developed a task model comprising key components, including user characteristics, user engagement with DHIs, and the effectiveness of DHIs. The model comprised data analysis tasks that explored not only each component individually but also their interrelationships. In addition, we designed an interactive visual analytics tool based on the proposed model to investigate how the model and tool could support diverse DHI data analysis practices. To make the evaluation more realistic, we visualized the analysis results from RWD collected from Maum Health, a currently operational DHI service, via the prototype and examined the DHI researchers' user experience while using it.

Overall, the participants responded positively to the proposed model and the visual analytics tool. They noted that this supported analyses that were previously difficult or time consuming. Features such as reviewing individual user records, examining engagement and effectiveness from multiple perspectives, and comparing user-defined subgroups were helpful in improving the design and evaluation of DHIs. This evaluation also revealed several areas for improvement. The participants suggested including more detailed engagement metrics that reflected actual user interactions, adding functionality to retain selected user groups across views, and providing clearer interpretations of statistical results to support collaboration among researchers with different levels of data literacy.

In subsequent sections, we provide several insights into the analysis task model and the interactive visual analytics tool and suggest design implications for supporting researchers when analyzing DHI data.

Analysis Task Model

The analysis task model allowed the DHI research team to analyze the data holistically, addressing the limitations of existing studies that analyzed each key component separately.

XSL•FO
RenderX

The proposed model supported the analysis of the relationships between key components by specifying user groups based on certain conditions (eg, gender) and comparing target outcomes (eg, changes in depression). During this process, the research team identified meaningful insights for improving DHIs, such as suggesting content to potential target users, setting criteria for DHI use, and determining better interventions. This aligns with recent studies investigating which DHI use behaviors (ie, engagement with DHIs) affect target outcomes [8,49], highlighting the importance of further analyses of these relationships.

In real-world settings, DHIs can be distributed through app stores to individuals with varying characteristics, contexts of use, and engagement levels [50]. Therefore, as suggested by our model, evaluating the DHI use behaviors and changes in target symptoms for each group of similar users is critical. Group-level analysis has been widely used in previous studies on large-scale health data [13,24,51]. These studies explored the overall trends or patterns at the user group level and derived insights to address common issues within the group. Given that DHI data can be collected from several users, DHI researchers may follow similar analytical practices.

In addition, the records of individual users can be provided as sample cases to complement the understanding of a specific user group. Because it is challenging to review all records, the analysis task model can be improved to suggest the representative users of that group. For instance, they could be users showing mean or median engagement levels or those with the most frequently reported user characteristics. This approach helps researchers gain a more nuanced understanding of user behavior and outcomes, allowing for more tailored and effective interventions.

Considering the researchers' feedback regarding our prototype that they would set guidelines after exploring the engagement levels for each user group, group-level analysis could also be used to determine the optimal dosage. In practice, determining the dosage of typical DHIs is difficult because of various behavioral and contextual factors influencing dose-response relationships [14,16]. The proposed model could support researchers in investigating how much DHI use is required to achieve a certain level of improvement in the target symptoms. On the basis of this insight, they can further improve existing DHIs by adjusting the dosage or revising the difficulty of tasks to be performed in the intervention content while monitoring how these changes affect the effectiveness of DHIs.

Interactive Visual Analytics

The interactive visual analytics for DHIs proposed in our study demonstrated the potential of allowing researchers to perform analysis tasks efficiently. The findings of this study revealed that researchers often conducted similar DHI data analyses; however, there is a lack of research on visual analytics tools to streamline this process. The Maum Health Analytics tool has enabled researchers to analyze DHI data from diverse perspectives in real time through simple interactions.

The proposed analytics can be used at various stages of the development and evaluation life cycle of DHI services, which typically follow an agile software development process [23]. This process involves several stages of design and testing, with results from each testing fed back into the previous cycle to continuously improve DHIs. Therefore, researchers need to monitor and evaluate DHI data, and our tool can make this iterative process more efficient.

We also discovered that interactive visual analytics can be useful for DHI research teams comprising experts from diverse domains and backgrounds, such as clinicians, intervention designers, and system developers. As shown in this study, the purpose of data analysis may vary owing to background differences. For instance, clinicians value whether target symptoms have improved, whereas system developers focus more on the interaction log generated from DHI.

In such situations, the analyses provided by our analytics tool allow researchers to explore unfamiliar aspects of DHI data and extend their knowledge of the data. This facilitates communication among researchers from different fields, generating diverse approaches to better support users and improve existing intervention components. Furthermore, the analytical tool can be used to quickly share results with stakeholders outside the research team, such as funding bodies or external project managers, thus aiding the decision-making process.

Design Implications

Our findings offer further ways to quantify DHI use behaviors and diversify the metrics used to represent user engagement with interventions. As reported in this study, some DHI users may simply run the intervention content without actively engaging with it. This implies that relying solely on traditional high-level engagement metrics, such as launch counts or use time, may not distinguish this type of passive use, which could negatively affect the analysis quality.

Therefore, we propose incorporating fine-grained engagement metrics to better understand how active users interact with DHIs. For instance, rather than simply measuring how often a DHI is launched or how long it is used, we could track specific actions, such as touches, swipes, or clicks, within the app to obtain a clearer picture of user engagement [12]. Moreover, if the DHI includes activities that users are supposed to perform, the system could track their adherence to these activities and incorporate these data into the engagement metrics. By doing so, we can identify more meaningful (or effective) engagement and analyze its impact on improving target symptoms. This approach provides a more nuanced understanding of user behavior and the effectiveness of interventions.

In addition, we suggest design implications of interactive visual analytics tools for DHI researchers. Given that multiple user groups can be formed from data, determining the groups that should be investigated is challenging. Therefore, this tool can be designed to proactively identify the important user groups that should be examined by researchers. For instance, an analytics tool can recommend user groups with very high engagement or very low effectiveness, as these may provide critical insights. Leveraging the data analysis history stored in the analytics tool can provide analysis results that researchers

XSL•FO
RenderX

are interested in or guide them to discover other important user groups.

Furthermore, we can support user group selection using natural language, making it easier for researchers to locate the groups they want to explore. This could be a practical alternative, particularly if numerous metrics are available for analyzing user engagement with DHIs and their effectiveness. In addition, enabling researchers to explore various factors based on a specified user group (ie, the tag function in this study) would enhance the user experience of visual analytics by allowing for seamless DHI data analysis.

When designing a visual analytics tool, both the exploratory and explanatory aspects should be carefully considered. Members of a multidisciplinary research team may have varying levels of knowledge and experience in data analysis. Therefore, interpreting the results is crucial for applying data-driven insights in practice. For instance, our tool can be further improved by explaining the statistical results and offering practical guidelines based on them.

With recent advances in large language models, visual analytics can generate reports based on data analysis histories thus enhancing researchers' understanding of DHI data [52]. As suggested in a recent study [53], the data presented in visual elements can also be accompanied by a data-driven narrative generated by large language models, which effectively aids data communication. By integrating these features, the visual analytics tool can become a powerful and user-friendly resource for multidisciplinary research teams, facilitating deeper insights and more effective decision-making in the development and evaluation of DHIs.

Limitations and Future Works

We conducted a preliminary user study on the interactive visual analytics tool designed based on the findings from our research experience and previous studies. The analytics tool presented in this study was a preliminary medium-fidelity prototype, with limited interaction capabilities intended for exploratory evaluation. In future studies, it will be necessary to develop a fully functional system based on this prototype and compare it with existing analytical tools. During this process, it is important to include DHI data collected over extended periods to allow users to evaluate how engagement and effectiveness evolve over time, as frequently observed in real-world DHI implementation.

Investigating how this analytics tool and the underlying analysis task model are applied in real-world settings is another important direction for future research, particularly for supporting DHI research teams more effectively. By addressing these aspects, we aim to create comprehensive and robust analytics that contribute to the field of DHIs.

In addition, the analytic task model and visual analytics tool proposed in this study were developed based on data from a mobile DHI service focused on mental health, specifically targeting depressive symptoms. Thus, the scope and tasks of the model may be influenced by characteristics unique to this domain. When applying the model to DHIs in other health domains, it is necessary to adapt the analytical components to reflect domain-specific requirements. For instance, physical and metabolic health interventions may rely on objective indicators, such as step count, blood glucose levels, or sleep duration, which require different approaches for modeling and evaluation. Similarly, engagement in these domains may involve different types of intervention activities and their corresponding completion metrics. In such cases, an analytical framework is needed to accommodate more nuanced metrics that better reflect user engagement and health outcomes. Furthermore, interventions involving real-time tracking or context-aware features may produce logs that require different forms of temporal or event-based analyses. Expanding the model to support such flexible representations and metrics is critical to enhance its applicability across diverse types of DHIs.

Conclusions

We proposed an analysis task model to support DHI researchers by incorporating analysis practices from existing DHI research. On the basis of this model, we designed a prototype of an interactive visual analytics tool to explore how the proposed task model and tool can aid in DHI data analysis tasks. A preliminary user study demonstrated that the proposed model enabled a holistic investigation of DHI data by integrating the 3 key components and examining their relationships. Moreover, the analytics tool demonstrated the potential to simplify repetitive tasks and facilitate communication among researchers from diverse backgrounds and interests.

As the use of DHIs continues to grow, research on enhancing data analysis and supporting decision-making for improving existing DHIs has become increasingly important. In this context, we expect that the model and tool proposed in our study will effectively meet the data analysis needs of DHI researchers and make the process more efficient.

Acknowledgments

This research was supported by the Bio and Medical Technology Development Program of the National Research Foundation, funded by the Korean government (Ministry of Science and Information and Communication Technology [MSIT]; 2021M3A9E4080780); the National IT Industry Promotion Agency, funded by the Korean government (MSIT; H1001-24-1002); and the Institute of Information and communications Technology Planning and Evaluation grant funded by the Korean government (MSIT; 2022-0-00064).

XSL•FO
RenderX

Authors' Contributions

All authors participated in the conception of the study. GJ, HL, and UL developed the task analysis model and designed the interactive visual analytics, with contributions from KH and HK. KH and HK collected the Maum Health data, while GJ, HL, and UL analyzed and prepared it. The preliminary user study was conducted by GJ and HJ, with feedback from UL, KH, and HK. GJ prepared the manuscript, and all authors read, commented on, and contributed to the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Description of the analysis tasks supported by the Maum Health Analytics tool.
[DOCX File , 50478 KB - humanfactors_v12i1e64967_app1.docx]

References

1. Monitoring and evaluating digital health interventions: a practical guide to conducting research and assessment. World Health Organization. URL: https://www.who.int/publications/i/item/9789241511766 [accessed 2025-05-29]
2. Kowatsch T, Otto L, Harperink S, Cotti A, Schlieter H. A design and evaluation framework for digital health interventions. It Inf Technol 2019 Oct;61(5-6):253-263. [doi: 10.1515/itit-2019-0019]
3. Murray E, Hekler EB, Andersson G, Collins LM, Doherty A, Hollis C, et al. Evaluating digital health interventions: key questions and approaches. Am J Prev Med 2016 Nov;51(5):843-851 [FREE Full text] [doi: 10.1016/j.amepre.2016.06.008] [Medline: 27745684]
4. Dugas M, Gao GG, Agarwal R. Unpacking mHealth interventions: a systematic review of behavior change techniques used in randomized controlled trials assessing mHealth effectiveness. Digit Health 2020;6:2055207620905411 [FREE Full text] [doi: 10.1177/2055207620905411] [Medline: 32128233]
5. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95 [FREE Full text] [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
6. Michie S, Wood CE, Johnston M, Abraham C, Francis JJ, Hardeman W. Behaviour change techniques: the development and evaluation of a taxonomic method for reporting and describing behaviour change interventions (a suite of five studies involving consensus methods, randomised controlled trials and analysis of qualitative data). Health Technol Assess 2015 Nov;19(99):1-188 [FREE Full text] [doi: 10.3310/hta19990] [Medline: 26616119]
7. Lee U, Jung G, Ma EY, Kim JS, Kim H, Alikhanov J, et al. Toward data-driven digital therapeutics analytics: literature review and research directions. IEEE/CAA J Autom Sin 2023 Jan;10(1):42-66. [doi: 10.1109/jas.2023.123015]
8. Alshurafa N, Jain J, Alharbi R, Iakovlev G, Spring B, Pfammatter A. Is more always better?: discovering incentivized mHealth intervention engagement related to health behavior trends. Proc ACM Interact Mob Wearable Ubiquitous Technol 2018 Dec 27;2(4):1-26 [FREE Full text] [doi: 10.1145/3287031] [Medline: 32318650]
9. Borghouts J, Eikey E, Mark G, De Leon C, Schueller SM, Schneider M, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. J Med Internet Res 2021 Mar 24;23(3):e24387 [FREE Full text] [doi: 10.2196/24387] [Medline: 33759801]
10. Choi W, Park S, Kim D, Lim YK, Lee U. Multi-stage receptivity model for mobile just-in-time health intervention. Proc ACM Interact Mob Wearable Ubiquitous Technol 2019 Jun 21;3(2):1-26. [doi: 10.1145/3328910]
11. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
12. Cole-Lewis H, Ezeanochie N, Turgiss J. Understanding health behavior technology engagement: pathway to measuring digital behavior change interventions. JMIR Form Res 2019 Oct 10;3(4):e14052 [FREE Full text] [doi: 10.2196/14052] [Medline: 31603427]
13. Chen AT, Wu S, Tomasino KN, Lattie EG, Mohr DC. A multi-faceted approach to characterizing user behavior and experience in a digital mental health intervention. J Biomed Inform 2019 Jun;94:103187 [FREE Full text] [doi: 10.1016/j.jbi.2019.103187] [Medline: 31026595]
14. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011 Aug 05;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]
15. Fleming T, Bavin L, Lucassen M, Stasiak K, Hopkins S, Merry S. Beyond the trial: systematic review of real-world uptake and engagement with digital self-help interventions for depression, low mood, or anxiety. J Med Internet Res 2018 Jun 06;20(6):e199 [FREE Full text] [doi: 10.2196/jmir.9275] [Medline: 29875089]

XSL•FO
RenderX

16. Mclaughlin M, Delaney T, Hall A, Byaruhanga J, Mackie P, Grady A, et al. Associations between digital health intervention engagement, physical activity, and sedentary behavior: systematic review and meta-analysis. J Med Internet Res 2021 Feb 19;23(2):e23180 [FREE Full text] [doi: 10.2196/23180] [Medline: 33605897]
17. Lipschitz JM, Van Boxtel R, Torous J, Firth J, Lebovitz JG, Burdick KE, et al. Digital mental health interventions for depression: scoping review of user engagement. J Med Internet Res 2022 Oct 14;24(10):e39204 [FREE Full text] [doi: 10.2196/39204] [Medline: 36240001]
18. Kim M, Patrick K, Nebeker C, Godino J, Stein S, Klasnja P, et al. The digital therapeutics real-world evidence framework: an approach for guiding evidence-based digital therapeutics design, development, testing, and monitoring. J Med Internet Res 2024 Mar 05;26:e49208 [FREE Full text] [doi: 10.2196/49208] [Medline: 38441954]
19. Gilbert S, Pimenta A, Stratton-Powell A, Welzel C, Melvin T. Continuous improvement of digital health applications linked to real-world performance monitoring: safe moving targets? Mayo Clin Proc Digit Health 2023 Sep;1(3):276-287 [FREE Full text] [doi: 10.1016/j.mcpdig.2023.05.010] [Medline: 40206630]
20. Doherty G, Coyle D, Sharry J. Engagement with online mental health interventions: an exploratory clinical study of a treatment for depression. In: Proceedings of the 2012 SIGCHI Conference on Human Factors in Computing Systems. 2012 Presented at: CHI '12; May 5-10, 2012; Austin, TX p. 1421-1430 URL: https://dl.acm.org/doi/10.1145/2207676.2208602 [doi: 10.1145/2207676.2208602]
21. Morrison C, Doherty G. Analyzing engagement in a web-based intervention platform through visualizing log-data. J Med Internet Res 2014 Nov 13;16(11):e252 [FREE Full text] [doi: 10.2196/jmir.3575] [Medline: 25406097]
22. Thieme A, Hanratty M, Lyons M, Palacios J, Marques RF, Morrison C, et al. Designing human-centered AI for mental health: developing clinically relevant applications for online CBT treatment. ACM Trans Comput Hum Interact 2023 Mar 17;30(2):1-50 [FREE Full text] [doi: 10.1145/3564752]
23. Wilson K, Bell C, Wilson L, Witteman H. Agile research to complement agile development: a proposal for an mHealth research lifecycle. NPJ Digit Med 2018;1:46 [FREE Full text] [doi: 10.1038/s41746-018-0053-1] [Medline: 31304326]
24. Moshe I, Terhorst Y, Philippi P, Domhardt M, Cuijpers P, Cristea I, et al. Digital interventions for the treatment of depression: a meta-analytic review. Psychol Bull 2021 Aug;147(8):749-786. [doi: 10.1037/bul0000334] [Medline: 34898233]
25. Kim M, Han K. Relationship between depression level and mental health app activity log. HCI Soc Korea 2023:1003-1007 [FREE Full text]
26. Andresen EM, Malmgren JA, Carter WB, Patrick DL. Screening for depression in well older adults: evaluation of a short form of the CES-D. Am J Prev Med 1994 Mar;10(2):77-84. [doi: 10.1016/S0749-3797(18)30622-6]
27. Liu X. Real-world data for the drug development in the digital era. J Artif Intell Technol 2022 Mar 31;2(2):42-46. [doi: 10.37965/jait.2022.0091]
28. Taur SR. Observational designs for real-world evidence studies. Perspect Clin Res 2022;13(1):12-16 [FREE Full text] [doi: 10.4103/picr.picr_217_21] [Medline: 35198423]
29. Pratap A, Neto EC, Snyder P, Stepnowsky C, Elhadad N, Grant D, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. NPJ Digit Med 2020;3:21 [FREE Full text] [doi: 10.1038/s41746-020-0224-8] [Medline: 32128451]
30. Karyotaki E, Kleiboer A, Smit F, Turner DT, Pastor AM, Andersson G, et al. Predictors of treatment dropout in self-guided web-based interventions for depression: an 'individual patient data' meta-analysis. Psychol Med 2015 Oct;45(13):2717-2726 [FREE Full text] [doi: 10.1017/S0033291715000665] [Medline: 25881626]
31. Karyotaki E, Riper H, Twisk J, Hoogendoorn A, Kleiboer A, Mira A, et al. Efficacy of self-guided internet-based cognitive behavioral therapy in the treatment of depressive symptoms: a meta-analysis of individual participant data. JAMA Psychiatry 2017 Apr 01;74(4):351-359 [FREE Full text] [doi: 10.1001/jamapsychiatry.2017.0044] [Medline: 28241179]
32. Baumel A, Muench F, Edan S, Kane JM. Objective user engagement with mental health apps: systematic search and panel-based usage analysis. J Med Internet Res 2019 Sep 25;21(9):e14567 [FREE Full text] [doi: 10.2196/14567] [Medline: 31573916]
33. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
34. Pham Q, Graham G, Carrion C, Morita PP, Seto E, Stinson JN, et al. A library of analytic indicators to evaluate effective engagement with consumer mHealth apps for chronic conditions: scoping review. JMIR Mhealth Uhealth 2019 Jan 18;7(1):e11941 [FREE Full text] [doi: 10.2196/11941] [Medline: 30664463]
35. Short CE, DeSmet A, Woods C, Williams SL, Maher C, Middelweerd A, et al. Measuring engagement in eHealth and mHealth behavior change interventions: viewpoint of methodologies. J Med Internet Res 2018 Nov 16;20(11):e292 [FREE Full text] [doi: 10.2196/jmir.9397] [Medline: 30446482]
36. Sieverink F, Kelders SM, van Gemert-Pijnen JE. Clarifying the concept of adherence to eHealth technology: systematic review on when usage becomes adherence. J Med Internet Res 2017 Dec 06;19(12):e402 [FREE Full text] [doi: 10.2196/jmir.8578] [Medline: 29212630]
37. Taki S, Lymer S, Russell CG, Campbell K, Laws R, Ong KL, et al. Assessing user engagement of an mHealth intervention: development and implementation of the growing healthy app engagement index. JMIR Mhealth Uhealth 2017 Jun 29;5(6):e89 [FREE Full text] [doi: 10.2196/mhealth.7236] [Medline: 28663164]

XSL•FO
RenderX

38. Molloy A, Anderson PL. Engagement with mobile health interventions for depression: a systematic review. Internet Interv 2021 Dec;26:100454 [FREE Full text] [doi: 10.1016/j.invent.2021.100454] [Medline: 34621626]
39. Choi SW, Schalet B, Cook KF, Cella D. Establishing a common metric for depressive symptoms: linking the BDI-II, CES-D, and PHQ-9 to PROMIS depression. Psychol Assess 2014 Jun;26(2):513-527 [FREE Full text] [doi: 10.1037/a0035768] [Medline: 24548149]
40. Wahl I, Löwe B, Bjorner JB, Fischer F, Langs G, Voderholzer U, et al. Standardization of depression measurement: a common metric was developed for 11 self-report depression measures. J Clin Epidemiol 2014 Jan;67(1):73-86. [doi: 10.1016/j.jclinepi.2013.04.019] [Medline: 24262771]
41. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
42. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. Appl Psychol Meas 1977 Jun 01;1(3):385-401. [doi: 10.1177/014662167700100306]
43. Beck AT, Steer RA, Ball R, Ranieri WF. Comparison of beck depression inventories -IA and -II in psychiatric outpatients. J Pers Assess 1996 Dec;67(3):588-597. [doi: 10.1207/s15327752jpa6703_13] [Medline: 8991972]
44. Pilkonis PA, Choi SW, Reise SP, Stover AM, Riley WT, Cella D, PROMIS Cooperative Group. Item banks for measuring emotional distress from the Patient-Reported Outcomes Measurement Information System (PROMIS®): depression, anxiety, and anger. Assessment 2011 Sep;18(3):263-283 [FREE Full text] [doi: 10.1177/1073191111411667] [Medline: 21697139]
45. Patton MQ. Enhancing the quality and credibility of qualitative analysis. Health Serv Res 1999 Dec;34(5 Pt 2):1189-1208 [FREE Full text] [Medline: 10591279]
46. Sauro J, Lewis JR. Quantifying the User Experience: Practical Statistics for User Research. New York, NY: Morgan Kaufmann; 2012.
47. Lewis JR. Psychometric evaluation of the PSSUQ using data from five years of usability studies. Int J Hum Comput Interact 2002;14:463-488. [doi: 10.1207/s15327590ijhc143&4_11]
48. Jeon E, Park HA. Development of a smartphone application for clinical-guideline-based obesity management. Healthc Inform Res 2015 Jan;21(1):10-20 [FREE Full text] [doi: 10.4258/hir.2015.21.1.10] [Medline: 25705553]
49. Strauss G, Flannery JE, Vierra E, Koepsell X, Berglund E, Miller I, et al. Meaningful engagement: a crossfunctional framework for digital therapeutics. Front Digit Health 2022;4:890081 [FREE Full text] [doi: 10.3389/fdgth.2022.890081] [Medline: 36052316]
50. Grundy Q. A review of the quality and impact of mobile health apps. Annu Rev Public Health 2022 Apr 05;43(1):117-134 [FREE Full text] [doi: 10.1146/annurev-publhealth-052020-103738] [Medline: 34910582]
51. Kelders SM, Bohlmeijer ET, Van Gemert-Pijnen JE. Participants, usage, and use patterns of a web-based intervention for the prevention of depression within a randomized controlled trial. J Med Internet Res 2013 Aug 20;15(8):e172 [FREE Full text] [doi: 10.2196/jmir.2258] [Medline: 23963284]
52. Cao SX, Chen Q, Cao N. Visual narrative for data journalism based on user experience. J Visual 2024 Jun 27;27(6):1195-1211. [doi: 10.1007/s12650-024-01005-w]
53. Zhao Y, Zhang Y, Zhang Y, Zhao X, Wang J, Shao Z, et al. LEVA: using large language models to enhance visual analytics. IEEE Trans Vis Comput Graph 2025 Mar;31(3):1830-1847. [doi: 10.1109/TVCG.2024.3368060] [Medline: 38437130]

Abbreviations

CES-D-10: 10-item Center of Epidemiologic Studies Depression Scale
DHI: digital health intervention
PSSUQ: Post-Study System Usability Questionnaire
RWD: real-world data

Edited by A Kushniruk; submitted 31.07.24; peer-reviewed by JH Lee, M Lee, YA Alshehhi, JH Park; comments to author 07.11.24; revised version received 08.04.25; accepted 12.05.25; published 02.07.25.

Please cite as:
Jung G, Lim H, Han K, Kim H, Lee U
Designing a Visual Analytics Tool to Support Data Analysis Tasks of Digital Mental Health Interventions: Case Study
JMIR Hum Factors 2025;12:e64967
URL: https://humanfactors.jmir.org/2025/1/e64967
doi:10.2196/64967
PMID:

XSL•FO
RenderX

©Gyuwon Jung, Heejeong Lim, Kyungsik Han, Hyungsook Kim, Uichin Lee. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 02.07.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Acceptability and Usability of a Socially Assistive Robot Integrated With a Large Language Model for Enhanced Human-Robot Interaction in a Geriatric Care Institution: Mixed Methods Evaluation

Lauriane Blavette[1,2,3], MSc; Sébastien Dacunha[1,2,3], MSc; Xavier Alameda-Pineda[4], PhD; Daniel Hernández García[5], PhD; Sharon Gannot[6], Prof Dr; Florian Gras[7], MSc; Nancie Gunson[5], PhD; Séverin Lemaignan[8], PhD; Michal Polic[9], PhD; Pinchas Tandeitnik[6], MSc; Francesco Tonini[10], MSc; Anne-Sophie Rigaud[1,2,3], Prof Dr Med; Maribel Pino[1,2,3], PhD

[1]Institut national de la santé et de la recherche médicale - Optimisation thérapeutique en pharmacologie OTEN U1144, Université Paris Cité, Paris, France

[2]Assistance Publique - Hôpitaux de Paris, Hôpital Broca, Centre Mémoire de Ressources et Recherches Île-de-France-Broca, Service gériatrie 1&2, Paris, France

[3]Broca Living Lab, Hôpital Broca, Paris, France

[4]Institut national de recherche en sciences et technologies du numérique de l'Université Grenoble Alpes, Grenoble, France

[5]Interaction Lab, Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh, United Kingdom

[6]Faculty of Engineering, Bar-Ilan University, Bar Ilan, Israel

[7]ERM Automatismes, Carpentras, France

[8]PAL Robotics, Barcelona, Spain

[9]Czech Institute of Informatics, Robotics and Cybernetics, Czech Technical University in Prague, Prague, Czech Republic

[10]Department of Information Engineering and Computer Science, University of Trento, Trento, Italy

Corresponding Author:
Lauriane Blavette, MSc
Institut national de la santé et de la recherche médicale - Optimisation thérapeutique en pharmacologie OTEN U1144
Université Paris Cité
Faculté de Pharmacie de Paris
4, avenue de l'observatoire
Paris, 75006
France
Phone: 33 660125406
Email: Lauriane.blavette@aphp.fr

Abstract

Background: Socially assistive robots (SARs) hold promise for supporting older adults (OAs) in hospital settings by promoting social engagement, reducing loneliness, and enhancing emotional well-being. They may also assist health care professionals by delivering information, managing routines, and alleviating workload. However, their acceptability and usability remain major challenges, particularly in dynamic real-world care environments.

Objective: This study aimed to evaluate the acceptability and usability of a SAR in a geriatric day care hospital (DCH) and to identify key factors influencing its adoption by OAs and their informal caregivers.

Methods: Over the course of 1 year, 97 participants (n=65, 67%, OA patients and n=32, 33%, informal caregivers) took part in a mixed methods evaluation of ARI, a socially assistive humanoid robot developed by PAL Robotics. ARI was deployed in the waiting area of a geriatric day care robot in Paris (France), where it interacted with users through voice-based dialogue. After each session, participants completed 2 standardized assessments, the Acceptability E-scale (AES) and the System Usability Scale (SUS), administered orally to ensure accessibility. Open-ended qualitative feedback was also collected to capture subjective experiences and contextual perceptions.

XSL•FO
RenderX

Results: Acceptability scores significantly increased across waves (wave 1: mean 15.4/30, SD 5.81; wave 2: mean 20.9/30, SD 5.25; wave 3: mean 22.5/30, SD 4.23; *P*<.001). Usability scores also improved (wave 1: mean 47.9/100, SD 24.18; wave 2: mean 57.4/100, SD 22.46; wave 3: mean 69.3/100, SD 16.03; *P*<.001). A strong positive correlation was observed between acceptability and usability scores (r=0.664, *P*<.001). Qualitative findings indicated improved ease of use, clarity, and user satisfaction over time, particularly following the integration of a large language model (LLM) in wave 2, leading to more coherent, natural, and context-aware interactions.

Conclusions: Successive system enhancements, most notably the integration of an LLM, led to measurable gains in usability and acceptability among patients and informal caregivers. These findings underscore the importance of iterative, user-centered design in deploying SARs in geriatric care environments.

Trial Registration: Approved by the French national ethics committee (CPP Ouest II, IRB: 2021/20) as it did not involve randomization or clinical intervention

(JMIR Hum Factors 2025;12:e76496) doi:10.2196/76496

KEYWORDS

socially assistive robot; hospital environment; gerontology; older adults; informal caregivers; acceptability; usability; large language model; human-robot interaction

Introduction

The rapid aging of the population worldwide is one of the major demographic challenges of the 21st century. According to United Nations projections, the number of people aged 65 years and over worldwide is expected to increase from 727 million in 2020 to over 1.5 billion by 2050 [1]. This demographic evolution is accompanied by a growing need for long-term care and increased use of specialized services, particularly in geriatric institutions [2]. Consequently, geriatric institutions will face significant challenges in managing a rising number of residents, many of whom present multiple chronic conditions that require comprehensive and personalized care [3]. In response to this societal challenge, the adoption of innovative technologies and, in particular, socially assistive robots (SARs), is increasingly considered a complementary solution to improve the well-being and autonomy of older adults (OAs) [4,5]. SARs are robotic entities designed to interact with humans in socially and emotionally engaging ways. According to Dautenhahn [6], a SAR is "a robot that can interact with humans in a social context, while possessing communication and learning capabilities that mimic, to some extent, human behaviors." These robots integrate artificial intelligence (AI) and natural language processing to a certain extent, enabling them to interact more naturally with users, leading to enhanced engagement and adaptation to different social contexts [7]. Within geriatric institutions, SARs have the potential to alleviate caregiver workloads, provide cognitive and social stimulation for patients, and assist with daily tasks [8]. However, for these innovations to be successfully implemented and fully beneficial, 2 critical ergonomic dimensions must be addressed: their acceptability and their usability in human-robot interaction (HRI).

The effectiveness of SARs in geriatric institutions largely depends on the quality of HRI [5,9]. Acceptability refers to how willing users—including patients, their informal caregivers, and health care professionals—are to adopt these new tools. This willingness is shaped by their beliefs, needs, and trust in the system [10,11]. Usability, in this context, refers to the extent to which a system supports users in accomplishing their tasks effectively, efficiently, and with satisfaction, within a defined environment. Beyond the system's ergonomic characteristics, such as interface layout, feedback modalities, and physical interaction, it also encompasses aspects such as learnability, error tolerance, and cognitive demand, which are particularly critical when designing for OA populations. [12-14].

Despite their potential, SARs still face significant limitations regarding the quality of HRI and user engagement with the technology. Studies highlight that SARs often struggle to convey emotions effectively, understand the context, and predict users' behaviors accurately [15-17]. These shortcomings can lead to rigid and unnatural communication, hindering the robot's ability to interpret human intentions and emotions [18]. Additionally, SARs often lack adaptability to individual preferences, which can result in a gradual decline in the user's interest and engagement [19]. These challenges highlight the importance of designing HRIs that are intuitive, customizable and aligned with the expectations of OAs to enhance both acceptability and successful integration into geriatric institutions. For example, OAs expect robots to use clear, slow speech; maintain a polite, emotionally supportive tone; minimize technical vocabulary; and favor voice-based over touchscreen-based interaction, all of which have been repeatedly identified as essential for promoting acceptance and usability among OAs [5,19-23].

To overcome these limitations, recent research has explored the integration of large language models (LLMs) into SARs to enhance their communication and interaction capabilities [24-26]. LLMs allow for greater conversational flexibility, improved comprehension of complex requests, and enhanced contextual coherence in responses [27,28]. However, these advancements also introduce new challenges, including biases in generated responses, lack of transparency in robot decision-making, and difficulties in real-time interaction management [29,30].

Given the cognitive and physical limitations often present in OAs, SARs must feature intuitive and user-centered HRIs to maximize their effectiveness [20]. SAR acceptability encompasses not only perceived ease of use but also alignment with user needs, including preferences related to the communication style, autonomy, affective support, and

XSL•FO
RenderX

interaction dynamics [10,31-35]. Among OAs, factors such as low digital literacy, distrust of technology, and sensory or cognitive impairments can influence acceptance [21,36]. Barriers to adoption may include reluctance regarding the robot's appearance, the communication style, or ethical concerns about reduced human interaction [11].

Simultaneously, usability focuses on how efficiently, effectively, and satisfactorily a system can be used to perform expected tasks [13,14]. In geriatric institutions, usability considerations extend to physical ergonomics (eg, size, weight, tactile interfaces, voice interaction), user safety, and emotional comfort [37]. Evaluating SAR usability through standardized methods, such as the System Usability Scale (SUS), helps obtain insights into adoption barriers, assess the ease of operation, and guide iterative technological improvements [12,38]. These aspects emphasize the necessity of a user-centered approach tailored to the specific needs of elderly users [20].

Several studies have investigated the acceptability and usability of SARs by OAs in real-world contexts, including institutional settings, such as nursing homes and hospitals [19,37,39]. However, many of these evaluations rely on scripted interactions or exclude contextual constraints found in routine care environments. Although increasing attention has been paid to the perceptions of professional caregivers, relatively few studies have examined the opinions of informal caregivers (family or friends), despite their role in health care decision-making and daily care activities [21,36]. These gaps highlight the need for in situ evaluations that incorporate the perspectives of both patients and informal caregivers, while considering institutional realities.

To tackle these challenges, empirical research is essential to assess the acceptability and usability of SARs in real-world conditions. This study aimed to evaluate the performance, acceptability, and usability of a SAR among patients and their informal caregivers in a geriatric day care hospital (DCH) in Paris (France). It also aimed to identify the factors that facilitate or hinder SAR adoption in geriatric institutions. The findings will provide deeper insights into the expectations and concerns of OAs and their informal caregivers regarding the use of SAR in health care, while offering a set of recommendations to guide future technological and ergonomic advancements in this field.

Methods

Participants

This study involved 2 populations: OA patients attending consultations at a geriatric DCH and their informal caregivers.

Inclusion criteria for patients were (1) being ≥60 years old, (2) having a Mini-Mental State Examination (MMSE) score above 10 (indicating the absence of severe cognitive impairment; based on standard thresholds, scores of 25-30 are considered normal, 21-24 as mild, 10-20 as moderate, and below 10 as severe) [40], and (3) not exhibiting symptoms of altered reality and understanding and speaking French fluently. Informal caregivers were family members or friends (primarily spouses or children) aged ≥18 years, accompanying the patient and speaking French fluently.

No *exclusion criteria* were applied based on gender, socioprofessional backgrounds, or ethnicity.

Recruitment

Recruitment was carried out using the DCH database, and participants were prescreened prior to enrollment, contacted over the phone, and invited to participate in the study the day of their next consultation. An information letter was sent by post and informed consent was collected onsite.

Setting

The study was conducted between May 2023 and July 2024 in the geriatric DCH of the French Memory Clinic of a geriatric hospital in Paris (France). The DCH provides specialized outpatient care for OAs with physical or cognitive impairments, offering a wide range of consultations, including neurology, oncology, cardiology, psychiatry, and memory assessments. Three waves of data collection were carried out during this period, with a different sample of participants included in each iteration.

Study Design

A mixed methods design was used, combining qualitative and quantitative methods to provide an in-depth understanding of the acceptability and usability of a SAR in a geriatric institution. Mixed methods are particularly suited to exploring complex phenomena in the health and social sciences, as they enable the integration of complementary perspectives and strengthen the validity of results [41,42]. By combining quantitative rating scales with semistructured interviews, this approach identifies both general trends and nuances specific to users' perceptions [43,44]. Mixed methods also enable data triangulation, essential for understanding factors that influence SAR adoption in real-world contexts [45]. This approach, widely recognized in health service evaluation, guarantees a richer, contextualized analysis of the dynamics at play [46,47].

Materials

ARI Robot

The SAR used in this study was ARI, developed by PAL Robotics (Spain) [48]. The robot is 1.65 m (5 ft 5 in) tall and weighs 50 kg (Figure 1). The robot moves by rolling and is equipped with articulated arms that are not designed for load bearing. Its interface includes a touchscreen located on the torso for dialogue transcription, animated eyes with a gaze-tracking module, luminous ears, and an emergency stop button on the back of the robot. ARI supports autonomous operation for 8-12 hours prior to recharging. It is equipped with wired and wireless connectivity capabilities for flexible integration into a variety of environments. For visualization, ARI is equipped with 3 wide-angle cameras positioned on the head, chest, and back, giving it a wide and versatile view of its surroundings. In terms of audio, the robot features 4 microphones, facilitating voice recognition and the efficient capture of ambient sounds.

XSL•FO
RenderX

Figure 1. Photos of the ARI robot: front and side (photo credit: PAL Robotics).



For this study, the ARI robot was programmed with a set of modules explicitly developed as part of the European H2020 SPRING (Socially Pertinent Robots in Gerontological Healthcare) project described by Alameda-Pineda et al [49]. Among these modules, the conversational system was initially developed before recent LLM advances (eg, ChatGPT), relying on a "traditional" modular architecture that combines retrieval-based responses, rule-based intent handling, and open-domain generation (wave 1) [50]. To take advantage of LLMs' abilities for solving complex language-related tasks, the conversational system was redeveloped (wave 2) and refined (wave 3) with an LLM-based architecture, based on the Vicuna model with 13 billion parameters (Vicuna-13b-v1.5 [51]). For deployment in hospital settings, it was adapted to function offline and integrated with a custom prompt targeting health care–related scenarios in French. This design ensured safe use without requiring an internet connection, while enabling more coherent and context-sensitive interactions with patients and informal caregivers.

At the time of the study, ARI had not yet been commercially deployed in hospitals or long-term care institutions beyond the scope of research projects. Its use in this evaluation should therefore be considered as part of an exploratory, precommercial experimentation phase within controlled health care settings.

Introduction of ARI's Capabilities

The SAR was deployed in the waiting area of a geriatric DCH to welcome and assist patients and their informal caregivers. Although participants were free to interact spontaneously with the robot, the researcher introduced 5 illustrative use cases at the beginning of the session to showcase the robot's potential functionalities: (1) greeting and welcoming the user, (2) recalling hygiene and infection prevention procedures, (3) providing information about the consultation procedure, (4) offering orientation and guidance about hospital services, and (5) providing entertainment activities. These examples were intended to guide participants' understanding of the robot's capabilities without restricting the content of their interactions.

Ethical Considerations

The study was approved by the French National Ethics Committee Comité de Protection des Personnes, CPP Ouest II, Maison de la Recherche Clinique-CHU Angers (Institutional Review Board [IRB]: 2021/20) and complied with the General Data Protection Regulation (GDPR; DPO: 20210114153645, AP-HP register). It did not involve randomization or clinical intervention. Informed consent was obtained from all participants onsite. Participants were informed that they could stop their participation at any time. The original consent covered the secondary analysis without further consent. Participant data were anonymized. No compensation was provided to the participants.

XSL•FO
RenderX

Assessment Scales

To assess the acceptability and usability of the ARI robot in a DCH, we used 2 standardized scales:

- The Acceptability E-scale (AES): Acceptability was assessed using the AES, French version [52]. Acceptability is defined as the psychological determinants that shape an individual's intention to use a technology prior to any direct experience with the system. The AES scale comprises 6 items rated on a 5-point Likert scale, yielding a total score ranging from 6 to 30. For marketable products, the AES's acceptability threshold is set at 25.81/30.
- The System Usability Scale (SUS): Usability was assessed using the SUS [53,54]. Usability refers to the degree to which a system can be used by specified users to achieve specific goals with effectiveness, efficiency, and satisfaction in a specified context of use. Ease of use influences user performance and satisfaction, while acceptability determines actual usage [55]. The SUS consists of a 10-item scale designed to assess the overall usability of a system, generating an overall score out of 100, where a higher score reflects better usability. For this scale, experimenters are asked to respond to statements on a 5-point Likert scale, with ratings ranging from "strongly disagree" to "strongly agree." The usability threshold for marketable products is 72/100 for this scale.

The adapted versions of both scales used in this study are provided in Multimedia Appendix 1. For each scale, participants were invited to comment on their choices in a discussion with the researcher.

Assessment Procedure

Each session (~45 minutes) took place in a dedicated room. After free interaction with the robot, participants completed 2 validated instruments: the AES and the SUS. Both were administered orally and supported by open-ended discussions to elicit qualitative data. Finally, participants were accompanied back to the waiting area.

Data Analysis

To ensure a comprehensive understanding of the research topic, both qualitative and quantitative data were collected. Each session was audio-recorded and subsequently transcribed for analysis.

Descriptive statistics (means, SDs, and percentages) were used to describe the sample characteristics and the scores obtained on the AES and SUS.

To assess statistical differences between waves and between groups, several statistical tests were performed. First, Kruskal-Wallis tests were used to compare multiple groups, and Mann-Whitney tests were used to compare 2 scores or 2 groups. Shapiro-Wilk tests were used for normality assumption. When significant differences were found, Tukey post hoc tests were conducted to further explore the data. $P<.05$ was considered statistically significant in all analyses.

Qualitative data were analyzed using inductive thematic analysis [56], allowing themes to emerge from the transcripts of the interviews with the researcher.

System Evolution for Each Experimental Wave

To evaluate the evolution of the ARI robotic system with the help of feedback from participants, we carried out system updates. Over the evaluation period, the ARI robot received 2 updates, resulting in 3 waves of experimentation (Figure 2). Table 1 shows the overall differences in system performance over the 3 test waves, and the full list is described in SPRING Deliverable D1.6 [57].

XSL•FO
RenderX

Figure 2. Illustration of experiments.



Table 1. System evolution based on participants' feedback.

Aspect	Wave 1 (May-July 2023)	Wave 2 (September-December 2023)	Wave 3 (March-May 2024)
Robustness of the robot's responses	Basic diagnostic tools to prevent operational failures; first version of a modular dialogue system	Improved accuracy of diagnostic tools; reduced memory/computational load for vision modules; integration of an extended LLM[a]	Enhanced general system stability; improved speech recognition robustness
Feature display and interactive functionalities	Display of dialogues' transcription on the robot's screen	Optimized display output; improved display model for transcribed speech	Conversation start/end display; operator speech control; refined LLM prompt tailored to a hospital context
Perception and tracking	Depth estimation of the environment; person tracking; facial identification modules	Integration of a module for associating vocal input to individual users (Ecapa model); optimization of facial recognition processing time	Addition of head movements enabling the robot to follow participants' gaze; optimization of gaze estimation; improved audiovisual tracking; enhanced speaker identification (Ecapa2 model); improved voice activity detector

[a]LLM: large language model.

XSL•FO
RenderX

Results

Participants' Sociodemographic Data

Descriptive Statistics

Over the 3 experimental waves, 110 participants agreed to take part in the study. Of these, only ninety-seven (88.2%) completed the experiment. Incomplete robot assessment sessions occurred when participants were called away for their scheduled medical appointments. The results presented reflect the feedback from these 97 participants across the 3 experimental waves (wave 1: n=14, 14.4%; wave 2: n=43, 44.3%; wave 3: n=40, 41.3%). Table 2 shows the sociodemographic data for the participants in each wave. Participants were categorized as patients (P) or informal caregivers (IC).

Table 2. Sociodemographic data for each experimental wave.

Profile and characteristics	Wave 1 (n=14)	Wave 2 (n=43)	Wave 3 (n=40)	Total (N=97)
P[a]				
Participants, n/N (%)	11/14 (78.6)	30/43 (69.8)	24/40 (60.0)	65/97 (67.0)
Males, n/N (%)	2/11 (18.2)	9/30 (30.0)	6/40 (25.0)	17/65 (26.2)
Females, n/N (%)	9/11 (82%)	21/30 (70.0)	18/40 (75.0)	48/65 (73.8)
Age (years), mean (SD)	78.4 (7.1)	78.3 (6.6)	81.2 (6.4)	79.4 (6.7)
Education (years), mean (SD)	13.6 (2.1)	12.9 (3.5)	11.4 (3.6)	12.5 (3.4)
MMSE[b] score, mean (SD)	27.8 (1.9)	25.2 (4.5)	25.5 (3.9)	25.7 (4.1)
IC[c]				
Participants, n/N (%)	3/14 (21.4)	13/43 (30.2)	16/40 (40.0)	32/97 (33.0)
Males, n/N (%)	1/3 (33.3)	5/13 (38.5)	9/16 (56.3)	15/32 (46.9)
Females, n/N (%)	2/3 (66.7)	8/13 (61.5)	7/16 (43.7)	17/32 (53.1)
Age (years), mean (SD)	75.3 (11.6)	63.9 (19.4)	67.9 (12.9)	67 (15.7)
Education (years), mean (SD)	12.0 (3.0)	15.0 (0)	13.2 (4.1)	13.8 (3.2)
MMSE score, mean (SD)	N/A[d]	N/A	N/A	N/A

[a]P: patients.

[b]MMSE: Mini-Mental State Examination.

[c]IC: informal caregivers.

[d]N/A: not applicable.

Sociodemographic data analysis using nonparametric tests (justified by significant Shapiro-Wilk tests indicating nonnormal distributions: $P<.001$ for socioeducational level, $P=.008$ for MMSE score, and $P\leq.003$ for age across waves 2 and 3) revealed no significant difference in terms of age across the 3 waves ($\chi^2_2=0.388$, $P=.82$) and no significant difference in the socioeducational level ($\chi^2_2=3.330$, $P=.19$), suggesting a relatively homogeneous distribution of these characteristics among the groups. Furthermore, no significant difference was observed regarding the MMSE scores of patients across waves ($\chi^2_2=4.064$, $P=.13$).

A correlation analysis was carried out between AES and SUS scores in order to explore the consistency of participants' responses across the 2 scales. The results revealed a strong positive correlation between the 2 variables (r=0.664, $P<.001$), indicating that higher acceptability scores are associated with higher usability scores.

Acceptability E-Scale

The mean AES scores over the 3 waves of observation showed a positive progression between waves (Figure 3): 15.4/30 (SD 5.81) for wave 1, 20.9/30 (SD 5.25) for wave 2, and 22.5/30 (SD 4.23) for wave 3; this indicated a steady increase over time. Using the Kruskal-Wallis test, the mean AES scores showed a significant increase between waves ($\chi^2_2=13.4$, $P<.001$). The Mann-Whitney test revealed no significant differences in AES scores between male and female participants (U=960, $P=.71$) and between patients and informal caregivers (U=905, $P=.57$).

XSL•FO
RenderX

Figure 3. AES scores across the 3 experimental waves. AES: Acceptability E-scale.

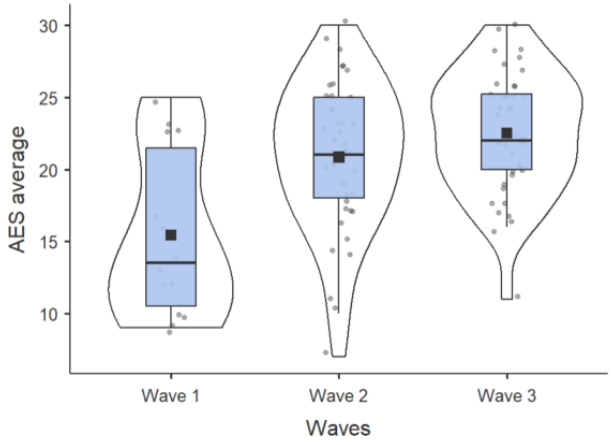
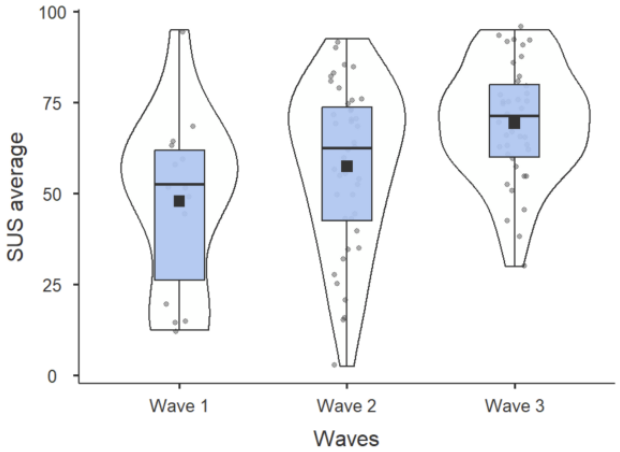


Figure 4. SUS scores across the 3 experimental waves. SUS: System Usability Scale.



Analysis of AES scores across the 3 experimental waves showed a significant difference between waves 1 and 2 (P<.01), as well as between waves 1 and 3 (P<.001). Additionally, no significant difference was observed between waves 2 and 3.

Based on these findings, an item-by-item analysis was conducted to identify specific differences across themes related to system acceptability. Tukey post hoc tests revealed significant differences in 4 of the 6 items: robot's usability, robot's perceived usefulness in the hospital setting, robot's response time, and overall satisfaction with the robot.

Regarding satisfaction with the robot, a significant mean difference of –0.956 was observed between waves 1 and 2 (P=.02) and a difference of –1.336 between waves 1 and 3 (P<.001). With respect to the robot's perceived usefulness in the hospital, a significant mean difference of –1.289 was found between waves 1 and 3 (P=.004). The robot's response time showed a significant difference of –1.86 between waves 1 and 2 and a significant difference of –2.079 between waves 1 and 3 (P<.001). Finally, on overall robot satisfaction, a mean difference of –0.815 was observed between waves 1 and 2 (P=.042), as well as a mean difference of –1.246 between waves 1 and 3 (P<.001).

System Usability Scale

The mean SUS scores over the 3 waves of observation showed a positive progression between waves: 47.9/100 (SD 24.18) for wave 1, 57.4/100 (SD 22.46) for wave 2, and 69.3/100 (SD 16.03) for wave 3; this indicated a steady increase over time. Using the Kruskal-Wallis test, the mean SUS scores showed a significant increase between waves (χ^2_2=11.4, P<.001). The Mann-Whitney test revealed no significant differences in SUS scores between men and women (U=848, P=.14) and between patients and informal caregivers (U=855, P=.19).

Analysis of SUS scores across the 3 experimental waves revealed significant differences between waves 1 and 3 (P<.01) and between waves 2 and 3 (P<.05). No significant difference was found between waves 1 and 2 (P>.05).

Based on these findings, an item-by-item analysis was conducted to examine specific differences in responses related to system usability. Tukey post hoc tests revealed significant differences in 5 of the 10 items: conversation complexity, assistance, inconsistency, ease of use, and user confidence.

Regarding the complexity of conversations, a significant difference of –1.089 was observed between waves 1 and 3 (P=.01). Regarding the assistance required to interact with the robot, a significant difference of –1.182 was observed between waves 1 and 3 (P=.02). Regarding the design of the robot's

XSL•FO
RenderX

functionality, a significant difference of –1.039 was observed between waves 1 and 3 (P=.005), as well as a mean difference of –0.569 between waves 2 and 3 (P=.04). Robot inconsistencies showed a significant difference of –1.28 between waves 1 and 2 (P=.003) and –1.632 between waves 1 and 3 (P<.001). Finally, regarding confidence using the robot, a significant difference –1.050 was observed between waves 1 and 3 (P=.01).

Qualitative Data

Thematic Qualitative Analysis

A thematic qualitative analysis of participants' feedback was conducted to explore perceptions of the ARI robot's acceptability and usability. Verbatim responses from patients and informal caregivers were reviewed and categorized into 5 themes representing key aspects of user experience: (1) ease of use, (2) comprehension of the robot's responses, (3) perceived usefulness of the robot in a hospital setting, (4) quality of interaction, and (5) overall system satisfaction. These themes correspond to established dimensions of acceptability (eg, perceived usefulness, satisfaction) and usability (eg, ease of use, comprehension, efficiency). Each theme is illustrated next with representative quotes translated from French to English.

Ease of Use

The robot's ease of use, an essential component of acceptability, was progressively enhanced throughout the study. Although its physical ergonomics remained constant, updates to the dialogue system facilitated more fluid and coherent interactions. These improvements contributed to participants' perceptions of the robot as simpler and more intuitive to operate. Several participants pointed out that the novelty of the robot could, at first, generate a certain amount of reluctance, suggesting that perceived ease of use also depended on experience acquired through interaction. For example, in wave 1, some found the voice interface practical:

> *The voice is helpful because I can't see well [referring to the transcript on the robot's screen]. You just have to ask it a question.* [Informal caregiver, wave 1]

Others found it unintuitive:

> *Not at all [easy to use]!* [Patient, wave 1]

However, some participants pointed out the limit of novelty when faced with a robot:

> *I mean, we're not used to it. When you're used to it, it's easy, but not the first time. It's a bit confusing. You have to get used to it.* [Patient, wave 1]

This need for an adaptation period was most prominent in wave 1, when the system was less refined, but it also emerged in a few cases in waves 2 and 3. This suggests that although improved usability facilitated smoother interactions, user familiarity remained a factor influencing initial comfort in engaging with the robot.

As waves progressed, more participants acknowledged the simplicity of interacting with the robot:

> *It's easy...you just need to ask questions, and if the question is relevant, it answers.* [Patient, wave 2]

However, some participants reported challenges in formulating their questions appropriately:

> *I had to force myself to phrase my questions differently from how I would naturally say them.* [Informal caregiver, wave 3]

Understanding the Robot's Responses

Although voice synthesis remained consistent, updates to the dialogue system, particularly the integration of an LLM facilitated clearer and more coherent exchanges. Participants generally found the language comprehensible; however, some reported mismatches between their questions and the robot's responses, suggesting that perceived comprehension remained sensitive to the contextual specificity of the interactions. In wave 1, participants' opinions ranged from finding the robot's speech fully understandable ("They were completely understandable and easy" [patient, wave 1]) to reporting inconsistencies in comprehension. In wave 2, many agreed that the robot spoke clearly:

> *It was perfectly understandable.* [Patient, wave 2]

However, some responses were found to be off-topic:

> *Its explanations are clear, but some responses were off the mark.* [Informal caregiver, wave 2]

In wave 3, participants generally found the language clear:

> *I found it very clear and precise.* [Patient, wave 3]

However, some noted occasional mismatches between their questions and the robot's responses:

> *Not entirely understandable, because it made me ask the same question multiple times.* [Patient, wave 3]

Perceived Usefulness of the Robot in a Hospital Setting

Perceptions of the robot's usefulness evolved over the 3 waves. Although some participants found it useful for providing basic information, others expected a more interactive experience. Over time, the focus shifted to the accuracy of the robot's responses, with some seeing it as a "glimpse into the future," while others emphasized its current limitations. In wave 1, some found it practical for basic information:

> *Just telling me [the location needed] it's next door was enough for me.* [Informal caregiver, wave 1]

However, others expected a more interactive experience:

> *If it [the robot] had come with me [to show me the path], it would have been better.* [Patient, wave 1]

In wave 2, usefulness was conditional on the accuracy of responses:

> *Plenty useful, but only if it understood the question and answered it correctly.* [Informal caregiver, wave 2]
>
> *It did the job, but a good signpost could do it too.* [Informal caregiver, wave 2]

By wave 3, opinions remained divided, with some participants perceiving the robot as indicative of future technological developments:

XSL•FO
RenderX

> *It's an opening into something futuristic. So, it's useful to me.* [Informal caregiver, wave 3]

However, for others, opinions remained mixed with respect to the robot:

> *Well, it annoyed me quite a bit [...] Every time I ask it a question, it tells me to go to the reception.* [Informal caregiver, wave 3]

Some participants highlighted its efficiency and consistency compared to human interactions:

> *It answered all those questions 100%. [Informal caregiver, wave 3)*
>
> *Even if some people might find a human more pleasant, humans can get tired or irritated—whereas ARI [the robot] never does.* [Informal caregiver, wave 3]

Quality of Interaction

Throughout the study, updates to the robot enhanced the fluidity of interactions and the quality of user feedback—2 factors critical to its acceptability. These improvements, particularly in the response speed, interruption management, and response richness, were reflected in participants' feedback. During wave 1, the robot demonstrated inconsistent response patterns, as reported by participants:

> *I asked a question, but it [the robot] responded with something completely unrelated—it was totally incoherent.* [Patient, wave 1]

Some participants also perceived that the robot had a limited range of programmed responses:

> *I have the impression that it is programmed only to answer certain questions.* [Patient, wave 1]

Additionally, participants noted that the robot often repeated suggestions unrelated to their queries:

> *But it does offer to go to lunch a lot.* [Patient, wave 1]

In wave 3, some participants noted that the robot responded promptly:

> *You ask the question, and it answers almost immediately.* [Informal caregiver, wave 3]

However, others reported challenges in interrupting the robot while it was delivering a response:

> *You can't interrupt it, and that bothered me a bit.* [Informal caregiver, wave 3]

Furthermore, some participants expressed a preference for the robot's responses over those provided by human receptionists:

> *I prefer your robot's answers to those of the ladies at the reception.* [Patient, wave 3]
>
> *There are humans who aren't as direct and precise [...] I'm surprised.* [Patient, wave 3]

General Satisfaction

Overall satisfaction with the experimentation of the robot reflected growing acceptance in terms of both functional effectiveness and ease of use, with clear progression observed from the first to the final wave. In wave 1, some participants appreciated the robot's novelty:

> *I think it is really nice. I think it's fun and funny.* [Informal caregiver, wave 1]
>
> *I like it—it's nice. Really nice. It's great to see a robot. But a human is still nicer.* [Informal caregiver, wave 1]

Conversely, some participants expressed dissatisfaction:

> *I'm sorry if I sound upset, but I didn't feel any satisfaction.* [Patient, wave 1]
>
> *I didn't get any satisfaction—the robot didn't help me at all. It's actually a bit scary.* [Patient, wave 1]

In wave 2, reactions remained mixed; some participants found the robot engaging and entertaining:

> *I find it amazing, it's super.* [Patient, wave 2]
>
> *It's impressive, it's the first time it's happened to me, I didn't know robots at all before today, and it's impressive to see a machine that you ask questions to, that answers you.* [Patient, wave 2]

However, others highlighted its limitations:

> *I'm sure the robot has potential, but at the moment, it's not performing well.* [Informal caregiver, wave 2]

In wave 3, participant feedback was generally more favorable. Participants described the robot as "very interesting, very instructive because it's very knowledgeable. It answered me perfectly well" (patient, wave 3) and declared that "it's funny because you don't expect to find something like this, and it does have quite a lot of answers" (informal caregiver, wave 3). Others expressed an overall appreciation:

> *Oh, I like it a lot. I think it's nice, it has beautiful eyes.* [Patient, wave 3]
>
> *I like this thing.* [Patient, wave 3]

Discussion

Principal Findings

This study provides one of the first in situ evaluations of a SAR equipped with an LLM in a geriatric hospital setting. Quantitative data showed a significant increase in both acceptability and usability of the ARI robot, corroborating the qualitative findings. Participants reported improved ease of use, comprehension, and perceived usefulness across experimental waves, as reflected in the significant increases in AES and SUS scores. Acceptability scores increased from 15.4/30 in wave 1 to 22.5/30 in wave 3 (P<.001), while usability scores rose from 47.9/100 to 69.3/100 (P=.003). Verbatim feedback reflected a growing appreciation for the robot's functionality and ease of interaction, aligning with observed improvements in satisfaction and perceived usefulness. These improvements are largely attributable to technical upgrades, most notably the integration of an LLM, which enhanced the naturalness, contextual relevance, and consistency of interactions. The experimental conversational system, developed collaboratively within the

XSL•FO
RenderX

European H2020 SPRING project [58], consists of an audiovisual perception pipeline that captures speech and behavioral cues, processed by a dialogue system powered by an LLM to synthesize responses into sound. Collectively, these developments contributed to a more fluid, engaging, and user-centered experience with the robot.

Although the ARI robot is commercially available, the conversational system developed in this study remains experimental. The final scores did not meet established thresholds for acceptability (AES score>25.81) and usability (SUS score>72), underscoring the need for further refinement prior to large-scale hospital implementation, such as transitioning to full-duplex communication, adding multilingual support, and incorporating nonverbal feedback. Nonetheless, the upward trajectory of these scores is encouraging, suggesting strong potential for future clinical deployment as performance approaches benchmark criteria.

Limitations

This study presents several limitations related to both the experimental design and the technical capabilities of the robotic system.

From a methodological perspective, the study was conducted at a single site within an innovation-friendly clinical environment. Although this facilitated implementation, it may limit the generalizability of the findings to institutions with different organizational structures or cultural contexts. Moreover, the recruitment strategy relied on voluntary participation, introducing a potential self-selection bias.

Another limitation concerns the ecological validity of the experimental conditions. Although the ARI robot was deployed in a hospital setting, the interactions took place in a quiet, controlled room rather than in a dynamic waiting area. This may have minimized external distractions and reduced communication challenges typically present in real-world hospital contexts. In addition, the interaction occurred only once per participant. As a result, the findings largely reflect first impressions and do not capture how user perceptions evolve with repeated exposures or long-term use. Longitudinal studies are needed to assess the durability of engagement and trust over time.

On a technical dimension, although the integration of an LLM improved conversational coherence, the system remained prone to occasional misinterpretations or generic responses, especially when confronted with vague or ambiguous input. This limitation is consistent with known issues in LLM-based dialogue systems, which often struggle with generating contextually appropriate and nongeneric responses when input is vague or underspecified [59]. Furthermore, the robot's speech recognition system showed sensitivity to user-specific features, such as accent, the speech rate, and background noise. Finally, the absence of multilingual support reduced inclusiveness for non–French-speaking users, an important consideration in multicultural health care environments.

Comparison With Prior Work

Our findings align with the existing literature showing that AI-powered SARs can significantly enhance user engagement when interactions are intuitive and socially appropriate [5,19,26]. The integration of an LLM between waves 1 and 2 notably improved the ARI robot's conversational fluency and contextual relevance, features widely recognized as essential for effective HRI [1,2]. Although natural language processing has long been identified as a key component of user engagement, few studies have directly captured the impact of integrating an LLM into SARs in a health care context. Most prior evaluations were either conducted before the emergence of transformer-based language models (pre-2019) or focused on single-instance deployments with fixed systems. Moreover, although most existing studies on SARs have been conducted in controlled environments, such as laboratories or nursing homes [3,4], few have examined how real-time system refinements affect user perceptions over successive implementation phases. This study addressed this gap by documenting how progressive improvements to the LLM-based dialogue system influence usability and acceptability in situ over 3 successive waves.

Future systems should integrate multimodal strategies (eg, gesture recognition, facial expression analysis, affective cues) to support more natural and empathetic communication in health care environments [16,49].

Our results also align with previous research advocating for voice-based interaction as particularly suitable for OAs, especially in contexts where digital literacy may vary [22,23]. Participants consistently favored spoken communication over touchscreen input. Furthermore, the inclusion of real-time transcription improved accessibility, confirming prior findings on the value of combining auditory and visual cues for inclusive design [60].

Confidence in using the system was closely linked to interaction quality as although wave 3 participants, who received consistent and context-aware responses, gained confidence, those in earlier waves, confronted with erratic outputs, voiced concerns about the robot's reliability. Interestingly, some participants also blamed themselves, reflecting a well-documented phenomenon in HRI literature, automation bias, where users tend to overestimate the capabilities of intelligent systems and assume personal responsibility for communication failures [39,61]. These findings underscore the critical role of transparency and predictability in AI-driven behaviors as key determinants of user confidence, reinforcing prior research in HRI [16,62,63].

Finally, our results reaffirm the need for user familiarization with voice-based robotic interaction in health care contexts. For the majority of participants, this study marked their first direct interaction with a SAR, and many indicated the need for a period of adjustment. Initial hesitation often gave way to positive engagement, echoing previous research showing that structured onboarding and repeated exposure are key to long-term acceptability [11].

XSL•FO
RenderX

Recommendations

Drawing on the empirical findings of this study, the following recommendations are intended to inform the design and implementation of SARs tailored to the specific interactional needs of OAs in geriatric care settings. These evidence-based insights aim to guide researchers and industry stakeholders in developing socially assistive technologies that are both user centered and contextually appropriate within health care environments.

Improve Natural Language Processing, Speech Recognition, and Multilingual Support

To enhance the quality of HRI in geriatric care, it is recommended that future SARs integrate more advanced natural language processing capabilities than those available in current open models to ensure greater coherence and contextual relevance in responses. Speech recognition systems should be optimized to account for prosodic variation, regional accents, and articulation differences (eg, slower speech rate, reduced vocal intensity, hesitations, or imprecise consonant production) commonly observed among OAs due to age-related changes in respiratory and orofacial motor control [64]. In addition, the incorporation of multilingual functionality is essential to accommodate the linguistic diversity of users in health care settings and to promote equitable access to robotic services.

Develop Clear Onboarding Protocols and Autonomous Robot Self-Presentation

To facilitate user engagement and support initial interaction, it is recommended that future deployments of SARs include a structured onboarding protocol. This may involve brief demonstrations or the distribution of printed materials (eg, leaflets with example questions) to familiarize users with the robot's capabilities. However, recognizing that real-world hospital settings may not consistently offer human assistance at the point of interaction, SARs should also be equipped with a robust self-introduction feature. This feature should autonomously guide users through the robot's purpose, functionalities, and appropriate use cases, clarifying when and how to engage and what types of questions it can address. Such improvements are essential for promoting user autonomy, reducing uncertainty, and improving the perceived usefulness of SARs in health care environments.

Such improvements are essential for OAs, who often face reduced digital literacy, cognitive decline or impairments, or heightened anxiety when interacting with unfamiliar technologies. Research has shown that guided introductions, whether verbal or visual, enhance initial confidence, reduce hesitation, and support sustained engagement [65,66].

Ensure Compliance With Privacy, Ethical, and Regulatory Standards

It is essential that the development and deployment of SARs in health care settings align with existing ethical guidelines and data protection regulations. Particular attention must be given to the collection, processing, and storage of sensitive health-related information. Developers and implementers should establish clear protocols for obtaining informed consent, ensuring transparency regarding data usage, and enabling users to understand the scope and limitations of data handling. Additionally, SARs should be designed to support confidential interactions, especially when used with vulnerable populations, such as OAs, by incorporating features that safeguard user privacy and uphold professional standards of care. Early integration of ethical and legal considerations will be critical to building trust and ensuring responsible implementation.

Enhance Interruption Management and Conversational Fluidity

Participants highlighted the difficulty of interrupting the ARI robot while it was speaking, a limitation that disrupted the natural flow of dialogue. To address this issue, future iterations of SARs should improve turn-taking mechanisms and enable responsive interruption handling, especially in multiparty or fast-paced care environments. These refinements are essential for supporting user-led interaction, minimizing frustration, and improving conversational naturalness.

Enable Autonomous Navigation and Spatial Guidance

Several participants expressed the desire for the ARI robot to not only provide verbal instructions but also physically accompany them to their destination. Incorporating autonomous navigation capabilities would expand the robot's role from an informational assistant to an active guide, offering greater support to users with mobility or orientation difficulties in complex care environments, such as hospitals. To ensure safe and effective guidance, particular attention should be paid to obstacle detection and navigation in dynamic environments, as well as to adapting the robot's speed to the walking pace of OAs.

Conclusion

This study showed that SARs, when iteratively enhanced and evaluated in real-world conditions, can reach growing levels of acceptability and usability among OAs and their informal caregivers. The integration of an LLM was a turning point in improving the ARI robot's ability to engage in meaningful, coherent interactions. These findings suggest that SARs, when designed with user needs in mind, hold strong potential for supporting care delivery in geriatric hospital settings.

Future work should focus on further leveraging the capabilities of LLMs to enable adaptive, personalized interactions that respond to individual users' communication styles, memory, and emotional cues. Additionally, the integration of multimodal communication features, such as gaze tracking, gesture and facial expression recognition, and targeted dialogue strategies, will be essential to enhancing the ARI robot's relevance, trustworthiness, and effectiveness in complex health care environments.

XSL•FO
RenderX

Acknowledgments

This research was funded by the European H2020 SPRING (Socially Pertinent Robots in Gerontological Healthcare) program (grant agreement 871245) [58]. The authors would like to sincerely thank all the participants who took part in this study, as well as all the health care professionals whose involvement contributed to the successful conduct of the SPRING project.

Data Availability

Requests to access the datasets should be sent to the corresponding author.

Authors' Contributions

Study conception and design were handled by LB, SD, XA-P, DHG, SG, FG, NG, SL, MP, PT, FT, A-SR, and MP; data collection by LB; analysis and interpretation of results by LB and SD; and draft manuscript preparation by LB, SD, MP, and A-SR. All authors have reviewed the results and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Evaluation scales: AES and SUS (translated from French to English). AES: Acceptability E-scale; SUS: System Usability Scale.
[DOCX File , 16 KB - humanfactors_v12i1e76496_app1.docx]

References

1. United Nations Department of Economic and Social Affairs, Population Division. World population ageing 2020 highlights. United Nations. 2020. URL: https://www.un.org/development/desa/pd/news/world-population-ageing-2020-highlights [accessed 2025-04-16]
2. Ageing and health. World Health Organization. 2023. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2025-04-16]
3. Health at a glance: Europe 2020. OECD. 2020. URL: https://www.oecd.org/en/publications/health-at-a-glance-europe-2020_82129230-en.html [accessed 2025-04-16]
4. Chu L, Chen H, Cheng P, Ho P, Weng I, Yang P, et al. Identifying features that enhance older adults' acceptance of robots: a mixed methods study. Gerontology 2019 Mar 7;65(4):441-450. [doi: 10.1159/000494881] [Medline: 30844813]
5. Naneva S, Sarda Gou M, Webb TL, Prescott TJ. A systematic review of attitudes, anxiety, acceptance, and trust towards social robots. Int J Soc Robot 2020 Jun 03;12(6):1179-1201. [doi: 10.1007/s12369-020-00659-4]
6. Dautenhahn K. Socially intelligent robots: dimensions of human-robot interaction. Philos Trans R Soc Lond B Biol Sci 2007 Apr 29;362(1480):679-704 [FREE Full text] [doi: 10.1098/rstb.2006.2004] [Medline: 17301026]
7. Mataric M. Socially assistive robotics: human-robot interaction methods for creating robots that care. 2014 Presented at: 9th ACM/IEEE International Conference on Human-Robot Interaction; March 3-6, 2014; Bielefeld, Germany p. 333. [doi: 10.1145/2559636.2560043]
8. Saadatzi MN, Logsdon MC, Abubakar S, Das S, Jankoski P, Mitchell H, et al. Acceptability of using a robotic nursing assistant in health care environments: experimental pilot study. J Med Internet Res 2020 Nov 12;22(11):e17509 [FREE Full text] [doi: 10.2196/17509] [Medline: 33180024]
9. David D, Thérouanne P, Milhabet I. The acceptability of social robots: a scoping review of the recent literature. Comput Hum Behav 2022 Dec;137:107419. [doi: 10.1016/j.chb.2022.107419]
10. de Jong C, Kühne R, Peter J, van Straten CL, Barco A. Intentional acceptance of social robots: development and validation of a self-report measure for children. Int J Hum-Comput Stud 2020 Jul;139:102426. [doi: 10.1016/j.ijhcs.2020.102426]
11. Mishra N. Exploring potential and acceptance of socially intelligent robot. In: Thalmann NM, Zhang JJ, Ramanathan M, Thalmann D, editors. Intelligent Scene Modeling and Human-Computer Interaction. Cham: Springer International Publishing; 2021:259-282.
12. Martín Rico F, Rodríguez-Lera FJ, Ginés Clavero J, Guerrero-Higueras Á, Matellán Olivera V. An acceptance test for assistive robots. Sensors (Basel) 2020 Jul 14;20(14):3912 [FREE Full text] [doi: 10.3390/s20143912] [Medline: 32674372]
13. Holden RJ. A simplified System Usability Scale (SUS) for cognitively impaired and older adult. Proc Int Symp Hum Factors Ergon Health Care 2020 Sep 16;9(1):180-182. [doi: 10.1177/2327857920091021]
14. Bevilacqua R. Robot-Era project: preliminary results on the system usability. In: Marcus A, editor. Design, User Experience, and Usability: Interactive Experience Design. Cham: Springer International Publishing; 2015:553-561.
15. Bishop L, van Maris A, Dogramadzi S, Zook N. Social robots: the influence of human and robot characteristics on acceptance. Paladyn J Behav Robot 2019;10(1):346-358. [doi: 10.1515/pjbr-2019-0028]
16. de Graaf MM, Ben Allouch S. Exploring influencing variables for the acceptance of social robots. Robot Auton Syst 2013 Dec;61(12):1476-1486. [doi: 10.1016/j.robot.2013.07.007]

XSL•FO
RenderX

17. Eyssel F, Kuchenbrandt D, Bobinger S. Effects of anticipated human-robot interaction and predictability of robot behavior on perceptions of anthropomorphism. 2011 Presented at: HRI '11: 6th International Conference on Human-Robot Interaction; March 6-9, 2011; Lausanne, Switzerland. [doi: 10.1145/1957656.1957673]
18. Heerink M. The influence of a robot's social abilities on acceptance by elderly users. 2006 Presented at: 15th IEEE International Symposium on Robot and Human Interactive Communication; September 6-8, 2006; Hatfield, UK p. 521. [doi: 10.1109/roman.2006.314442]
19. Whelan S, Murphy K, Barrett E, Krusche C, Santorelli A, Casey D. Factors affecting the acceptability of social robots by older adults including people with dementia or cognitive impairment: a literature review. Int J of Soc Robot 2018 Mar 1;10(5):643-668. [doi: 10.1007/s12369-018-0471-x]
20. Mahmoudi Asl A, Toribio-Guzmán JM, Castro-González Á, Malfaz M, Salichs MA, Franco Martín M. Evaluating the user experience and usability of the MINI robot for elderly adults with mild dementia and mild cognitive impairment: insights and recommendations. Sensors (Basel) 2024 Nov 08;24(22):7180 [FREE Full text] [doi: 10.3390/s24227180] [Medline: 39598957]
21. Beuscher LM, Fan J, Sarkar N, Dietrich MS, Newhouse PA, Miller KF, et al. Socially assistive robots: measuring older adults' perceptions. J Gerontol Nurs 2017 Dec 01;43(12):35-43 [FREE Full text] [doi: 10.3928/00989134-20170707-04] [Medline: 28700074]
22. Granata C, Pino M, Legouverneur G, Vidal J, Bidaud P, Rigaud A. Robot services for elderly with cognitive impairment: testing usability of graphical user interfaces. THC 2013 Jun 01;21(3):217-231. [doi: 10.3233/thc-130718]
23. Jakob D, Wilhelm S, Gerl A, Ahrens D, Wahl F. Adapting voice assistant technology for older adults: a comprehensive study on usability, learning patterns, and acceptance. Digital 2025 Jan 31;5(1):4. [doi: 10.3390/digital5010004]
24. Hanschmann L. A LLM-based social robot for human-like sales conversations. In: Følstad A, Araujo T, Papadopoulos S, Law ELC, Luger E, Goodwin M, et al, editors. Chatbot Research and Design. Cham: Springer Nature Switzerland; 2024:61-76.
25. Atuhurra J. Leveraging large language models in human-robot interaction: a critical analysis of potential and pitfalls. arXiv. Preprint posted online 2024. [doi: 10.48550/arXiv.2405.00693] . [doi: 10.48550/arXiv.2405.00693]
26. Kim C. Understanding large-language model (LLM)-powered human-robot interaction. 2024 Presented at: HRI '24: 2024 ACM/IEEE International Conference on Human-Robot Interaction; March 11-15, 2024; Boulder, CO. [doi: 10.1145/3610977.3634966]
27. Shibi S, Zaidi S. STEM approach to enhance robot-human interaction through AI large language modelsreinforcement learning. 2024 Presented at: 2024 IEEE Integrated STEM Education Conference (ISEC); March 9, 2024; Princeton, NJ. [doi: 10.1109/isec61299.2024.10665163]
28. Sobrín-Hidalgo D, González-Santamarta MA, Guerrero-Higueras AM, Rodríguez-Lera FJ, Matellán-Olivera V. Explaining autonomy: enhancing human-robot interaction through explanation generation with large language models. arXiv. Preprint posted online 2024. [doi: 10.48550/arXiv.2402.04206] . [doi: 10.48550/arXiv.2402.04206]
29. Zhang B, Soh H. Large language models as zero-shot human models for human-robot interaction. 2023 Presented at: 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); October 1-5, 2023; Detroit, MI p. 7961-7968. [doi: 10.1109/iros55552.2023.10341488]
30. Zhang C, Chen J, Li J, Peng Y, Mao Z. Large language models for human–robot interaction: a review. Biomim Intell Robot 2023 Dec;3(4):100131. [doi: 10.1016/j.birob.2023.100131]
31. Liberman-Pincu E, Oron-Gilad T. Exploring the effect of mass customization on user acceptance of socially assistive robots (SARs). 2022 Presented at: HRI '22: 2022 17th ACM/IEEE International Conference on Human-Robot Interaction; March 7-10, 2022; Sapporo, Hokkaido p. 880-884. [doi: 10.1109/hri53351.2022.9889466]
32. Irfan B, Kuoppamäki S, Hosseini A, Skantze G. Between reality and delusion: challenges of applying large language models to companion robots for open-domain dialogues with older adults. Auton Robot 2025 Mar 10;49(1):9. [doi: 10.1007/s10514-025-10190-y]
33. Lu H, Chan SS, Fung AW, Lam LC. Utility of Montreal Cognitive Assessment (Hong Kong Version) in the Diagnosis of Mild Neurocognitive Disorders (NCD): NCD due to Alzheimer Disease (NCD-AD) and NCD due to Vascular Disease (NCD-Vascular). J Am Med Dir Assoc 2016 Apr 01;17(4):366-367.
34. Fasola J, Mataric M. A socially assistive robot exercise coach for the elderly. J Hum Robot Interact 2013;2(2):3-32. [doi: 10.5898/jhri.2.2.fasola]
35. Heerink M, Kröse B, Evers V, Wielinga B. Assessing acceptance of assistive social agent technology by older adults: the Almere model. Int J of Soc Robot 2010 Sep 4;2(4):361-375. [doi: 10.1007/s12369-010-0068-5]
36. Becchimanzi C, Iacono E, Brischetto A. Acceptability of assistive robotics by older adults: results from a human-centred qualitative study. 2022 Presented at: AHFE 2022 International Conference on Human Factors in Accessibility and Assistive Technology; July 24-28, 2022; New York, NY p. 37. [doi: 10.54941/ahfe1001637]
37. Olde Keizer RACM, van Velsen L, Moncharmont M, Riche B, Ammour N, Del Signore S, et al. Using socially assistive robots for monitoring and preventing frailty among older adults: a study on usability and user experience challenges. Health Technol 2019 Apr 9;9(4):595-605. [doi: 10.1007/s12553-019-00320-9]

XSL•FO
RenderX

38. Olatunji S, Potenza A, Oron-Gilad T, Kiselev A, Loutfi A, Edan Y. Usability testing for the operation of a mobile robotic telepresence system by older adults. Proc Hum Factors Ergon Soc Annu Meet 2021 Feb 09;64(1):1191-1195. [doi: 10.1177/1071181320641284]
39. Jones-Jang S, Park Y. How do people react to AI failure? Automation bias, algorithmic aversion, and perceived controllability. J Comput Mediat Commun 2022;28(1):zmac029. [doi: 10.1093/jcmc/zmac029]
40. Wada K, Shibata T. Robot therapy in a care house - its socio-psychological and physiological effects on the residents. 2006 Presented at: 2006 IEEE International Conference on Robotics and Automation; May 15-19, 2006; Orlando, FL p. 3966-3971. [doi: 10.1109/robot.2006.1642310]
41. Almeida F. Strategies to perform a mixed methods study. Eur J Educ Stud 2018;5(1):137-151. [doi: 10.5281/zenodo.1406214]
42. Pluye P, Hong QN. Combining the power of stories and the power of numbers: mixed methods research and mixed studies reviews. Annu Rev Public Health 2014 Mar;35:29-45. [doi: 10.1146/annurev-publhealth-032013-182440] [Medline: 24188053]
43. Curry LA, Krumholz HM, O'Cathain A, Plano Clark VL, Cherlin E, Bradley EH. Mixed methods in biomedical and health services research. Circ Cardiovasc Qual Outcomes 2013 Jan 01;6(1):119-123 [FREE Full text] [doi: 10.1161/CIRCOUTCOMES.112.967885] [Medline: 23322807]
44. Scammon D, Tomoaia-Cotisel A, Day RL, Day J, Kim J, Waitzman NJ, et al. Connecting the dots and merging meaning: using mixed methods to study primary care delivery transformation. Health Serv Res 2013 Dec;48(6 Pt 2):2181-2207 [FREE Full text] [doi: 10.1111/1475-6773.12114] [Medline: 24279836]
45. Shorten A, Smith J. Mixed methods research: expanding the evidence base. Evid Based Nurs 2017 Jul 14;20(3):74-75 [FREE Full text] [doi: 10.1136/eb-2017-102699] [Medline: 28615184]
46. Terrell S. Mixed-methods research methodologies. TQR 2015 Jan 20;17:254-265. [doi: 10.46743/2160-3715/2012.1819]
47. McKim C. The value of mixed methods research: a mixed methods study. J Mixed Methods Res 2016 Jul 08;11(2):202-222. [doi: 10.1177/1558689815607096]
48. ARI. Pal Robotics. URL: https://pal-robotics.com/robot/ari/ [accessed 2025-07-22]
49. Alameda-Pineda X, Addlesee A, Hernández García D, Reinke C, Arias S, Arrigoni F, et al. Socially pertinent robots in gerontological healthcare. arXiv. Preprint posted online 2024. [doi: 10.48550/arXiv.2404.07560] . [doi: 10.48550/arXiv.2404.07560]
50. Gunson N, Hernández García D, Sieinska W, Dondrup C, Lemon O. Developing a social conversational robot for the hospital waiting room. 2022 Presented at: 2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN); August 29-September 2, 2022; Naples, Italy. [doi: 10.1109/ro-man53752.2022.9900827]
51. Wei-Lin C, Zhuohan L, Zi L, Ying S, Zhanghao W, Hao Z, et al. Vicuna: an open-source chatbot impressing GPT-4 with 90%* ChatGPT quality. LMSYS. 2023. URL: https://lmsys.org/blog/2023-03-30-vicuna/ [accessed 2025-07-22]
52. Micoulaud-Franchi J, Sauteraud A, Olive J, Sagaspe P, Bioulac S, Philip P. Validation of the French version of the Acceptability E-scale (AES) for mental E-health systems. Psychiatry Res 2016 Mar 30;237:196-200. [doi: 10.1016/j.psychres.2016.01.043] [Medline: 26809367]
53. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the System Usability Scale. Int J Hum-Comput Interact 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]
54. Brooke J. SUS: a quickdirty usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland IL, editors. Usability Evaluation in Industry. London, UK: Taylor & Francis; 1996:189-194.
55. Shackel B. What is usability? 1991 Presented at: 4th International Conference on Human-Computer Interaction; 1991; Stuttgart, Germany p. 3-15. [doi: 10.1016/B978-0-444-88464-4.50003-4]
56. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
57. SPRING D1.6 – user feedback from the final validation in relevant environments. SPRING Consortium. 2024. URL: https://spring-h2020.eu/wp-content/uploads/2024/07/SPRING_D1.6_-User-feedback-from-the-final-validation-relevant-environments_vFinal_31.05.2024.pdf [accessed 2025-07-22]
58. Socially pertinent robots in gerontological healthcare. SPRING. 2025. URL: https://spring-h2020.eu/ [accessed 2025-07-22]
59. Mielke C. Towards explainable spoken dialogue systems for robots. 2020 Presented at: 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations; July 5-10, 2020; Online p. 234-241. [doi: 10.18653/v1/2020.acl-demos]
60. Chen C, Lifset ET, Han Y, Roy A, Hogarth M, Moore AA, et al. Screen or no screen? Lessons learnt from a real-world deployment study of using voice assistants with and without touchscreen for older adults. 2023 Oct Presented at: ASSETS '23: 25th International ACM SIGACCESS Conference on Computers and Accessibility; October 22-25, 2023; New York, NY p. 1-21. [doi: 10.1145/3597638.3608378]
61. Petit E. The human/robot empathic relationship: normal or pathological? Implications Philosophiques. 2021 Jul 21. URL: https://tinyurl.com/3ptcasyk [accessed 2005-07-22]
62. Olatunji A. Investigating elderly perceptions of explainable robotic agents in elder care settings. 2023 Presented at: Workshop on Modeling Robots for Aging and Eldercare (MoRoBAE); 2023; Stockholm, Sweden.

XSL•FO
RenderX

63. Yamaji Y. How predictability of robot behaviors affects users' acceptance and trust in HRI: a study of timing control strategies in human-robot collaboration. IEEE Trans Haptics 2021;14(4):708-719. [doi: 10.1109/TOH.2021.3116294]
64. Klaassen W, van Dijk B, Spruit M. A review of challenges in speech-based conversational AI for elderly care. arXiv. Preprint posted online 2024. [doi: 10.48550/arXiv.2412.07388] . [doi: 10.48550/arXiv.2412.07388]
65. Flandorfer P. Population ageing and socially assistive robots for elderly persons: the importance of sociodemographic factors for user acceptance. Int J Populat Res 2012;2012:829835. [doi: 10.1155/2012/829835]
66. Borghouts J, Eikey EV, De Leon C, Schueller SM, Schneider M, Stadnick NA, et al. Understanding the Role of Support in Digital Mental Health Programs With Older Adults: Users' Perspective and Mixed Methods Study. JMIR Form Res 2022 Dec 13;6(12):e43192-e43174 [FREE Full text] [doi: 10.2196/43192] [Medline: 36512387]

Abbreviations

AES: Acceptability E-scale
AI: artificial intelligence
DCH: day care hospital
HRI: human-robot interaction
LLM: large language model
MMSE: Mini-Mental State Examination
OA: older adult
SAR: socially assistive robot
SUS: System Usability Scale

Edited by A Kushniruk; submitted 29.04.25; peer-reviewed by K Gopalan, C Getson, SB Guo; comments to author 24.05.25; revised version received 12.06.25; accepted 09.07.25; published 01.08.25.

Please cite as:
Blavette L, Dacunha S, Alameda-Pineda X, Hernández García D, Gannot S, Gras F, Gunson N, Lemaignan S, Polic M, Tandeitnik P, Tonini F, Rigaud AS, Pino M
Acceptability and Usability of a Socially Assistive Robot Integrated With a Large Language Model for Enhanced Human-Robot Interaction in a Geriatric Care Institution: Mixed Methods Evaluation
JMIR Hum Factors 2025;12:e76496
URL: https://humanfactors.jmir.org/2025/1/e76496
doi:10.2196/76496
PMID:

©Lauriane Blavette, Sébastien Dacunha, Xavier Alameda-Pineda, Daniel Hernández García, Sharon Gannot, Florian Gras, Nancie Gunson, Séverin Lemaignan, Michal Polic, Pinchas Tandeitnik, Francesco Tonini, Anne-Sophie Rigaud, Maribel Pino. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 01.08.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Providers' Perspective on the Feasibility of Digital Self-Management of Blood Pressure in Refugees: Mixed Methods Study

Carol Gonzalez[1], BS; Marcos Real[2], BS, MS; Nargis Ahmadi[3], BS; Raghad Aljenabi[3], BS; Lana Bridi[2], BS; Nour Makarem[4], PhD; Job Godino[1,5], PhD; Tala Al-Rousan[1], MD, MPH

[1]Herbert Wertheim School of Public Health, University of California San Diego, 9500 Gilman Drive, La Jolla, CA, United States

[2]School of Medicine, University of California San Diego, La Jolla, CA, United States

[3]University of California San Diego, La Jolla, CA, United States

[4]Mailman School of Public Health, Columbia University Irving Medical Center, New York, NY, United States

[5]Family Health Centers of San Diego, San Diego, CA, United States

Corresponding Author:
Tala Al-Rousan, MD, MPH
Herbert Wertheim School of Public Health, University of California San Diego, 9500 Gilman Drive, La Jolla, CA, United States

Abstract

Background: Mass displacement is a grand public health challenge. Refugees and immigrants experience a disparate hypertension burden and disparities in self-management. Successful hypertension self-management is key for improving outcomes, but research on its feasibility in refugee and immigrant health care settings is limited.

Objective: This study aimed to identify clinic staff–perceived barriers to and facilitators of implementing a digital intervention for hypertension self-management among refugee and immigrant patients and to identify its feasibility and usability.

Methods: Primary care physicians and medical assistants who care for refugees and immigrants in San Diego were interviewed using human-centered semistructured methods (n=18). Interviews were analyzed using an inductive approach. Usability testing for the software (Med Pro Care) was conducted with participants (n=15) to test the feasibility of real-time tracking of blood pressure (BP) home readings for hypertension self-management. Clinical staff rated their satisfaction on the System Usability Scale and the NASA Task Load Index, which measured mental workload.

Results: For refugee and immigrant patients self-managing hypertension, clinical staff identified barriers and facilitators in the following areas: (1) social determinants of health increase hypertension burden among refugee and immigrant patients, (2) clinical staff face challenges to effective hypertension care for refugee and immigrant patients, (3) perceived benefits of potential intervention for self-management, and (4) perceived barriers to potential intervention for self-management. Primary care physicians completed 90% of the tasks, and medical assistants completed 83% of the tasks successfully. Most clinic staff found the software system for monitoring BP to be easy to use with an average score for usability of 4.1 of 5.0 (SD 0.4).

Conclusions: Addressing identified barriers to and facilitators of self-management of hypertension is crucial to designing effective interventions in real-world refugee and immigrant health care settings. Telemonitoring interventions using software that transfers BP readings to clinical staff in real time may be feasible from the perspective of clinic staff and can address hypertension disparities in marginalized populations, such as immigrants and refugees. Addressing identified barriers to and facilitators of self-management is crucial to designing effective interventions in real-world refugee and immigrant health care settings. Our findings suggest that clinic staff view digital telemonitoring as both feasible and supportive of patient empowerment, health literacy, and improved communication—factors essential to addressing hypertension disparities in marginalized populations.

(JMIR Hum Factors 2025;12:e66176) doi:10.2196/66176

KEYWORDS

hypertension; refugee health; self-management; digital health; digital self-management; feasibility; blood pressure; refugee; mixed methods; public health; telemonitoring; telehealth; physician; medical assistants; real-time tracking; telemonitoring intervention; hypertension disparities

XSL•FO
RenderX

Introduction

Background

Hypertension holds global significance as one of the most important modifiable risk factors for cardiovascular disease. Almost one-third of the global population is diagnosed with hypertension [1], which causes 7.5 million deaths annually, and its increasing prevalence has led to greater morbidity for the global population from its sequelae, such as heart disease [2,3]. Notably, forcibly displaced populations, whose numbers are at all-time highs, including immigrants and refugees, experience a heightened burden of hypertension [4,5]. The challenges of fleeing conflict, interrupted health care access, adapting to new environments, and enduring persecution and trauma amplify the stress experienced by refugees and contribute to a higher hypertension risk [6-8]. In addition, in host societies, immigrants and refugees confront multifaceted barriers to engaging in health care, encompassing language and cultural barriers, economic barriers, and acculturation difficulties [9]. These factors contribute to this minoritized group's poor control of hypertension [10-14]. This is reflected in empirical evidence consistently pointing to elevated hypertension rates and almost double the cardiovascular disease risk in refugees compared to other immigrants or the hosting population, irrespective of their country of origin, compared to host country citizens [5,8].

Self-management of hypertension, including both self-monitoring of blood pressure (BP) and self-titration of medication, is an evidence-based, cost-effective method for improving BP control [15,16]. Patients who self-monitor are more likely to lower their BP [17,18], have fewer BP-related office visits [19,20], improve communication with their providers regarding lifestyle factors [21], and incur significant health care cost savings compared to usual care groups [22], regardless of the number of other comorbidities [23]. In addition to the reported accuracy of self-monitoring, many providers share positive perspectives on self-titration, emphasizing its potential benefits, including improved patient-provider relationships and enhanced self-efficacy [24-26]. Given the ongoing self-management demanded by chronic conditions such as hypertension, self-titration could provide crucial support for clinics in efficiently caring for their patients and decreasing clinic visits [7,21,27].

Objective

Providers are often responsible for health education, telemonitoring, and regular follow-up with patients for hypertension self-management, but acceptability and efficacy of digital health technologies, integration with existing electronic health records, usability, and technology support have emerged as key concerns [15,17,24,28]. Our study examined the feasibility of using technology for BP monitoring when caring for refugee and immigrant patients, and the barriers to and facilitators of this model of care. We also assessed the usability of a specific software (Med Pro Care; Withings) attached to connected BP monitors (Withings) that can help clinic staff implement telemonitoring in their clinics while receiving real-time BP readings from their patients at home [29].

Methods

Design

This study was part of a larger investigation into refugee and immigrant patients' barriers to hypertension care and opportunities to improve self-management. Patient participants in that study were provided a free and validated Withings BP monitor, a Food and Drug Administration–approved, cellular-connected home BP monitor [29], and were instructed to maintain a BP diary. The data from the BP monitor were transmitted to a web-based dashboard called Med Pro Care [29]. Our part of the study examined the clinic staff's perspective and experience with the Med Pro Care software using the data recorded from the patients in the larger study. Clinic staff participants in our study were previously trained by our staff on the use of the Med Pro Care dashboard, where they could view the charts for the patients in the larger study and where to find their BP readings in the software. The study was designed to determine the feasibility and usability of the Med Pro Care software for future research on longitudinal telemonitoring.

Our study used a human-centered design, mixed methods approach to gathering data concerning the perspectives of clinic staff on self-management of hypertension for refugee and immigrant patients and the usability of Med Pro Care for use in the clinic to help clinical staff telemonitor patients' home BP readings. We administered semistructured interviews, a demonstration of the Med Pro Care software for reviewing patient charts, and a usability test of Withing's Med Pro Care software, all to both medical assistants (MAs) and primary care providers (PCPs). Qualitative interviews aimed to identify clinic staff's experiences with treating refugee and immigrant patients with hypertension to better understand the social aspect of their health care (Multimedia Appendix 1), whereas the quantitative data collected from the usability test demonstrates the effectiveness of a possible solution addressing identified social factors (Multimedia Appendices 2-4). This study was conducted at federally qualified health centers, university clinics, and private clinics across southern San Diego County, California, where most refugees and immigrants in the region have resettled [30].

Participants

Participants included the following clinic staff: (1) MAs and (2) PCPs in San Diego County who regularly provided care for refugee and immigrant patients. We chose to ask both the same questions and combine their responses for data analysis because of growing evidence showing that multilevel interventions that involved integrated multidisciplinary care and promotion of self-management behaviors, which is done by diverse clinic staff, including providers, nurses, pharmacists, and MAs, were associated with the largest BP reductions [31,32]. Clinical staff were included if they self-reported regularly seeing refugee or immigrant patients in the San Diego region. Individuals were excluded if they did not care for patients with hypertension or self-reported not regularly caring for refugee or immigrant patients. While some participants self-identified as immigrants themselves, our study did not include a formal inquiry about previous or current citizenship or legal status.

XSL•FO
RenderX

Recruitment took place between August 2021 and November 2022 through 4 sites: (1) the Family Health Centers of San Diego (FHCSD), a federally qualified health center system that provides care for a large portion of refugees in San Diego County; (2) University of California San Diego Health System; and (3) 2 private practices in San Diego County. Study staff received contact information for potential participants through hospital directories and contacted clinic staff via email or phone. MAs were recruited via mass email to all MAs at the federally qualified health center, to which 13 MAs responded, and 2 were excluded for not working with patients with hypertension. Eleven MAs consented and participated, with 1 MA not completing the usability assessment. PCPs were referred by senior study staff who were aware of their practice. All 7 PCPs referred consented and participated, with 2 not completing the usability assessment.

Data Collection

Study staff created standardized operating procedures to ensure all interviews were conducted the same, regardless of the study staff who conducted the interview. In this process, the study team reviewed available literature about barriers to the implementation of self-management of hypertension for refugee and immigrant patients and created questions that could illuminate the clinic staff's perspective. The principal investigator then reviewed and approved these questions (Multimedia Appendix 1). The study coordinator (CG) then trained incoming study staff by reviewing the procedures and having new study staff shadow an interview.

Interviews were conducted in the participant's clinic (n=15), except for 1 PCP who participated in an office belonging to research staff and 2 PCPs who participated in the qualitative interview portion of the study over Zoom and did not complete the usability testing. Interviews typically lasted about 60 minutes in total and were audio recorded and later transcribed for qualitative analysis. Demographic information was collected from participants. Interview questions were the same for both MAs and PCPs and asked about their understanding of barriers to and facilitators of health care for refugee and immigrant patients, patient–clinic staff relationships, and the clinic staff's opinions of self-management of hypertension, including both self-monitoring and self-titration of antihypertensive medication. Although self-titration was not tested in the larger study, thus data were not available for our usability tests, participants were asked about their opinions on whether this was something that could help their refugee and immigrant patients because it has been previously shown to help self-management efficacy, such as in the TASMIN (Targets and Self-Management for the Control of Blood Pressure) trial where self-monitoring BP and an individualized self-titration algorithm improved systolic BP for patients in the United Kingdom [15,16]. Interview questions were the same for both MAs and PCPs because of their similar exposure to the social factors that were examined by the questions. While the focus was on refugee patients, question wording was modified to include immigrants for clinic staff who expressed more experience with immigrant patients rather than specifically refugees. Study staff then demonstrated the Med Pro Care software, which included administrative and monitoring tasks. Participants were then asked to complete a usability test of the software, where MAs were tested on more administrative tasks, such as registering patients and leaving notes on accounts, and PCPs were tested on more in-depth review of BP readings. Both MAs and PCPs were asked to complete 5 tasks. The usability test was screen and audio recorded using the TURF 4.0 software (McWilliams School of Biomedical Informatics) [33]. Participants then completed the System Usability Scale [34] and the NASA Task Load Index [35] for self-reported satisfaction and usability. Study staff then asked participants if they had any additional feedback on how the software could be used in their own clinics.

Data Analysis

Qualitative and quantitative data were analyzed separately. The semistructured interviews were analyzed using an inductive thematic approach to capture participants' experiences as comprehensively and accurately as possible [36]. To improve the credibility of the analysis, investigator triangulation was implemented as follows [37]: (1) 2 study staff (CG and MR) independently coded a set of transcripts (n=3), (2) then had multiple meetings to define a codebook and established intercoder agreements through the subjective assessment method [36]. The codebook was then reviewed by the principal investigator, finalized, and distributed to other researchers (NA and RA) who coded the remaining transcripts. Two study staff (CG and LB) analyzed the codes for recurrent themes following Crabtree and Miller's 5-step interpretive process [38]. Qualitative data were analyzed using Atlas. Using Microsoft Excel, 1 research team member (CG) performed basic descriptive statistical analysis on the demographic data, task completion, total clicks (collected from TURF software), and satisfaction questionnaires. For task completion, participants were given 1 point for successfully completing a task and half a point for partially completing a task. A final score was given as a percentage of tasks completed. Averages were then taken of all MAs scores and all PCPs scores separately. For the items that were negatively worded in the System Usability Scale, responses were reverse coded so that all items were aligned in the same direction, with higher scores being more favorable.

Ethical Considerations

The University of California San Diego Human Research Protection Program (#200063) granted institutional review board approval. All participants consented before participation and were compensated with a US $35 Visa gift card for their time. All participant data were fully deidentified prior to analysis to protect individual privacy and confidentiality. As part of our study, participants examined BP measurements from a larger investigation into refugee and immigrant patients' hypertension self-management (IRB#200063).

Results

Participant Characteristics

Data on participant demographics are listed in Table 1. MAs made up 61.1% (n= 11) of the 18 participants. Most participants were female (n=15, 83.3%). The 3 male participants were all PCPs. One PCP was a nurse practitioner, whereas all others were physicians. All MAs were associated with FHCSD,

XSL•FO
RenderX

whereas 42.9% (3/7) of PCPs were associated with FHCSD. The average age of participants was 38.9 (SD 11.9) years. Most participants self-identified as White (n=6, 33.3%), Hispanic (n=4, 22.2%), or Middle Eastern/North African (n=3, 16.7%). Besides English, the most commonly spoken languages were Spanish (n=9, 50%) and Arabic (n=6, 33.3%).

XSL•FO
RenderX

Table . Demographics of participants (N=18).

Demographics	Values, n (%)
Position	
PCP[a]	7 (38.9)
MA[b,c]	11 (61.1)
Clinic	
FHCSD[d]	14 (77.8)
Other clinics	4 (22.2)
Gender	
Male	3 (16.7)
Female	15 (83.3)
Race/ethnicity	
White (non-Hispanic)	6 (33.3)
Black (non-Hispanic)	2 (11.1)
Hispanic	4 (22.2)
Other	6 (33.3)
Middle Eastern/North African	3 (16.7)
Indian American	1 (5.6)
Multiracial	2 (11.1)
Languages spoken	
English only	4 (22.2)
Spanish	9 (50.0)
Arabic	6 (33.3)
French	2 (11.1)
Aramaic	1 (5.6)
Chaldean	1 (5.6)
German	1 (5.6)
Russian	1 (5.6)
Turkish	1 (5.6)
Level of education	
High school graduate or equivalent	1 (5.6)
Some college, no degree	6 (33.3)
Associate’s degree	2 (11.1)
Bachelor’s degree	2 (1.1)
Master’s degree[e]	1 (5.6)
Medical Doctor	5 (27.8)
Doctor of Osteopathic Medicine	1 (5.6)

[a]PCP: primary care physicians.

[b]All MAs were associated with FHCSD.

[c]MA: medical assistant.

[d]FHCSD: Family Health Centers of San Diego.

[e]Nurse practitioner.

XSL•FO
RenderX

Barriers and Facilitators Identified From Interviews

The interviews from this study, both MAs and PCPs, yielded 4 themes for barriers and facilitators: (1) social determinants of health increase hypertension burden among refugee and immigrant patients, (2) clinical staff face challenges to effective hypertension care for refugee and immigrant patients, (3) perceived benefits of potential intervention for self-management, and (4) perceived barriers of potential intervention for self-management.

Social Determinants of Health Influence Hypertension Burden Among Refugee and Immigrant Patients

All participants, including both MAs and PCPs, described hypertension as a burden among many of their refugee patients and described various factors unique to these patients influencing its management. Some described how patients' attitudes toward the disease limit care, including how patients may lack self-efficacy or motivation to make necessary changes to their care regimen. For example, when asked if refugee patients often followed the advice or directions given by the physician, an MA stated:

> *some [patients] are stuck in their ways. And they don't feel like changing is gonna help them either way, so they just continue to do whatever they're doing.* [MA]

Participants also acknowledged how the experience of forced displacement along with its associated stress and trauma can exacerbate hypertension:

> *Especially in a refugee population, we do see stress causing high blood pressure pretty directly. And so, to talk about any traumas they may have experienced, and have a way to treat those, [is important]... just really create an open conversation about that because I think it's not often talked about.* [PCP]

They described the health care barriers refugee and immigrant patients face such as language and cultural barriers:

> *...oftentimes they come untreated. I feel like it's similar to other patient populations, but the difference is that the language barrier and education has hindered treatment in the past, or they stop treatment for periods of time.* [PCP]

Participants also noted that poor health literacy and low socioeconomic status can contribute to difficulty with managing hypertension:

> *Patients don't have the supplies to take a proper blood pressure or they don't have the correct information [to self-manage hypertension].* [PCP]

Clinical Staff Face Challenges to Effective Hypertension Care for Refugee and Immigrant Patients

Common barriers faced by clinical staff when treating refugees and immigrants included effective communication, health literacy, adherence, and monitoring. Participants noted physician-patient communication was a critical variable in delivering effective care. For example, when a PCP was asked to mention a specific worry regarding their patients and the care they receive, they noted:

> *Yeah, the language barrier, them not understanding the education part of it, taking medication incorrectly...* [PCP]

Furthermore, both MAs and PCPs expressed concern that not all their refugee and immigrant patients completely understand the instructions that are given, or clinical staff do not have enough time to explain everything. This leads to miscommunication and confusion for patients, which can lead to uncontrolled and untreated hypertension, worsening patient health. Even with assistance from interpreters, participants highlighted that patients would often struggle to understand the concept and how to conduct BP checks. This miscommunication was emphasized when an MA stated that:

> *...even with the interpreters they do struggle to check the BP. I know some people receive the machine but they don't know how to use the machine and they don't let us know they don't know how to use it so there is always miscommunication between patient and provider or medical staff.* [MA]

Clinical staff were also concerned that their refugee and immigrant patients are nonadherent to treatment for a multitude of reasons, including access to monitoring technology, ability and willingness to take medication, and willingness to make lifestyle changes. Indeed, participants discussed how some patients would veer off the established plan and take medications in a manner that was not agreed on with the physician. When discussing the burden of hypertension within refugee patients, a PCP mentioned:

> *I actually have a lot of patients that they just go off the plan on their own because they think like they know it and they stop medications randomly. They take extra from others without like running it by me and they don't even admit it until I try to get it out of an individual.* [PCP]

Another significant variable that may exacerbate the prevalence of hypertension among refugee patients was concerns over patient adherence to treatment. Many refugee patients had fears regarding the consequences of undergoing medical intervention from clinic staff and thus were noted as being less compliant. This was highlighted when an MA was asked about concerns over medical treatments toward refugees and stated that:

> *...not wanting to take the medication because they are scared, or I know some patients get scared of taking medications and think it might affect the organs.* [MA]

Perceived Benefits of Potential Intervention for Self-Management

When asked about participation in a hypertension self-management intervention, participants largely endorsed potential benefits for their patients, with 15 participants expressing interest in this type of intervention. Interviewed participants often expressed the desire to not only educate their patients but also encourage active participation in treatment

XSL•FO
RenderX

planning and self-management of their condition. To facilitate patient adherence and hypertension self-management, participants would often try to educate refugee and immigrant patients on the significance of checking BP consistently at home and the concept of following trends. Physicians would also explain how hypertension does not always present with noticeable symptoms and can eventually lead to severe consequences if not managed on time. When discussing methods that help patients adhere to their medications, a physician stated the following:

> *Explaining to them the rationale behind checking the blood pressure at home and the trend and also educating the patient about hypertension being a silent killer. And that most people with high blood pressure have no symptoms initially.* [PCP]

With regard to the Med Pro Care device, most of the participants interviewed accepted self-management for hypertension to include both self-monitoring and self-titration. They believed that it more actively involves and empowers patients to be more in control of their health. In addition, self-management would reduce clinic burden, as patients would not have to come to the clinic as much because they would be monitored more consistently. For example, after trying out the Med Pro Care device, a physician was asked whether the device would be beneficial in the clinic and the physician responded with:

> *I think, overall, that is a really beneficial way to empower patients, and to not have to have them come back to the clinic as frequently when access is already pretty hard. And then, just to feel like they are in more control of their medications.* [PCP]

As both self-monitoring and self-titration have been successful for the management of other diseases, such as diabetes, most participants believed that with standardized protocols, it can also be successful for self-management of hypertension. Furthermore, self-management through Med Pro Care was overall noted as being a safe method to monitor and manage hypertension with the possibility of a reduction in adverse outcomes in the long term. This sentiment was explicitly noted by a physician when questioned about the potential benefits of the implementation of the Med Pro Care device:

> *I think that'll make for quicker, I guess, management of their high blood pressure. Minimize, you know, the adverse effects of like, you know, strokes, kidney damage, other things like that... I think it's very safe, this method of blood pressure monitoring by the patient themselves. It's safe and natural, and I think if we can implement more of that make that more available overall it's going to improve patient care and lessen adverse outcomes.* [PCP]

Despite acknowledging factors that lead to poor BP control, participants noted how family support and caregivers of refugee and immigrant patients play a significant role in their care. For example, families often attend appointments with patients and assist with BP monitoring, facilitating self-management and improved BP control:

> *I think immigrant patients tend to have a lot of family support and their kids often just help them and so they can pretty much take care of like downloading the app and checking their BP every day.* [MA]

Perceived Barriers to Potential Intervention for Self-Management

Two participants expressed concerns about implementing both self-monitoring and self-titration, mainly noting lack of success in patients self-titrating their hypertension medication. In addition, there was some concern over a possible increased workload because of monitoring an increased number of available readings. One participant was specifically concerned about legal liability if the clinic missed concerning BP readings and the patient did not reach out to the clinic or go to the hospital:

> *My concern is that... Let's say you did not see that alert. I mean, I have patients every day. I'm going to be honest. I don't know if I have time to check this every day... I don't want extra liability.* [PCP]

There was also a concern that patients would not be motivated enough to commit to their treatment plans without coming into the doctor's office regularly and would not remain adherent to medication and monitoring enough to maintain control over their BP. Although these concerns were raised, most participants remained optimistic that self-management of hypertension was feasible with protocols and training.

Results of Task Study

MAs were able to complete approximately 69% of the 5 tasks given to them, which included more administrative tasks such as how to register patients in the system and filtering through patients. Technical errors occurred with patient registration during many of the usability tests, resulting in all but 2 MAs being unable to complete the registration task, which could have influenced MA ratings of usability. Excluding this task, MAs were able to complete approximately 83% of tasks, with the next lowest score being in finding specific BP readings. PCPs completed approximately 90% of the 5 tasks, which were more focused on monitoring patients' readings over time. MAs completed all tasks with an average of 90.75 (SD 41.82) clicks, and PCPs completed all tasks with an average of 129.20 (SD 85.86) clicks. The average time to complete tasks for MAs was 6 minutes and 5 seconds (SD 2 min and 28 s), whereas the average time for PCPs was 4 minutes and 35 seconds (SD 48 s). With an average usability score of 4.1 of 5.0 (SD 0.4), participants believed the system was usable, as presented in Table 2, with the highest ratings in being able to quickly learn the software and it being easy to use. Results from the Task Load Index (Table 3) also indicate that participants believed that the tasks were not burdensome or stressful. MAs and PCPs had similar ratings of usability and satisfaction, apart from 1 MA who rated the physical demand and how hard they had to work as higher than others.

XSL•FO
RenderX

Table . System Usability Scale results.

System Usability Scale[a]	MA[b], mean (SD)	PCP[c], mean (SD)	All participants, mean (SD)
I think that I would like to use Med Pro Care frequently.	4.4 (1.2)	4.0 (1.0)	4.3 (1.0)
I found Med Pro Care unnecessarily complex.[d]	4.2 (1.4)	3.0 (0.3)	3.8 (0.9)
I thought Med Pro Care was easy to use.	4.4 (1.4)	4.4 (1.2)	4.4 (1.3)
I think that I would need the support of a technical person to be able to use Med Pro Care.[d]	4.3 (1.2)	3.6 (0.8)	4.1 (1.0)
I found the various functions in Med Pro Care were well integrated.	4.3 (1.1)	4.2 (1.4)	4.3 (1.1)
I thought there was too much inconsistency in Med Pro Care.[d]	4.2 (1.0)	4.2 (1.4)	4.2 (1.1)
I would imagine that most people would learn to use Med Pro Care very quickly.	4.6 (1.3)	4.4 (1.2)	4.5 (1.2)
I found Med Pro Care very cumbersome to use.[d]	3.1 (0.4)	4.0 (0.9)	3.4 (0.5)
I felt very confident using Med Pro Care.	4.4 (1.2)	3.4 (1.4)	4.1 (1.2)
I needed to learn a lot of things before I could get going with Med Pro Care.[d]	4.0 (1.0)	4.2 (1.3)	4.1 (1.0)

[a]All questions on a scale from 1 to 5, 1=strongly disagree, 2=disagree, 3=neither disagree nor agree, 4=agree, and 5=strongly agree.

[b]MA: medical assistant.

[c]PCP: primary care provider.

[d]Negatively worded items were reverse coded so that higher scores indicate more positive responses.

Table . NASA Task Load Index.

NASA Task Load Index[a]	MA[b], mean (SD)	PCP[c], mean (SD)	All participants, mean (SD)
How mentally demanding were the tasks?	0.0 (0.0)	6.8 (7.8)	5.4 (7.4)
How physically demanding were the tasks?	9.5 (13.4)	3.5 (4.7)	5.5 (7.7)
How hurried or rushed was the pace of the tasks?	2.7 (2.5)	5.3 (6.7)	4.1 (5.1)
How successful were you in accomplishing what you were	6.2 (7.4)	5.3 (8.5)	5.8 (7.4)
How hard did you have to work to accomplish your level of performance?	9.0 (12.7)	4.8 (6.1)	6.0 (7.5)
How insecure, discouraged, irritated, stressed, and annoyed were you?	0.7 (0.6)	1.3 (2.5)	1.0 (1.8)

[a]All items were measured on a scale from 0 to 20, with 0 being the least amount of demand or effort and 20 being the greatest amount of demand or effort, except the performance item where perfect is 0 and failure is 20.

[b]MA: medical assistant.

[c]PCP: primary care provider.

On the basis of the follow-up questions following the usability testing, participants were generally satisfied with the software, especially in areas such as ease of use and being able to learn the system easily. Participants expressed concern over logistical issues such as how well this would integrate with currently existing electronic health records, who was responsible for

XSL•FO
RenderX

keeping up with the increased number of measurements, and legal liability.

Discussion

Principal Findings

There was a general understanding by local clinic staff that refugees and immigrants experience pre- and postmigration stressors that negatively impact health, including increasing the risk of uncontrolled hypertension, and complicating treatment. Our study showed that addressing these unique barriers is essential for effective hypertension management. The Med Pro Care software was generally accepted by our sample and can be used in the clinic to facilitate self-monitoring of BP. Digital self-monitoring can enhance patient-clinic communication and support chronic disease self-management. Our study adds to existing literature by examining the clinic staff's perspective in treating refugee and immigrant patients and what can be done in the clinic to better address their specific needs.

Participants emphasized cultural and linguistic barriers, access, and patient empowerment as key priorities. Consistent with prior research, an approach that is sensitive to the cultural context and acknowledges the unique experiences of refugee and immigrant patients is pivotal in establishing effective communication and fostering patient engagement [39]. For refugees and immigrants, who must navigate unfamiliar health care systems in new countries, strategies that improve communication and enhance patient self-efficacy are crucial to long-term chronic disease self-management [40].

Clinic staff generally believed that self-management of hypertension can help empower patients by improving communication, self-efficacy, and health literacy, thereby potentially enhancing hypertension management. This aligns with other studies, which show self-management enhances health awareness and literacy, although few have focused on refugee and immigrant populations [41,42]. Our results, therefore, provide a unique perspective because the clinic staff interviewed have first-hand accounts of the barriers and facilitators toward health management in refugee and immigrant populations. Further research should directly examine the perspective of refugee and immigrant patients to inform clinical implementation.

Self-management has shown great potential in controlling hypertension; however, scarce resources and logistical concerns have slowed its adoption into many clinics [26]. Clinic staff noted that relying on self-management could reduce clinic burden by reducing the need for patients to come into the clinic as frequently. The usability of the Withing's Med Pro Care software was overall well received and considered easy to use and requiring minimal training, which could support efficient integration into clinical practice.

Perceived challenges of self-management included the clinic's ability to handle an increased load of information and safety concerns over self-titration of medications. Participants noted that Med Pro Care's ability to flag concerning readings and display trends could address some of these concerns. Liability concerns may be mitigated by patient education and standardized protocols. In addition, clinical trials have shown self-titration to be feasible and safe for patients with hypertension [28,40,43]. Med Pro Care can facilitate communication between patients and clinics through regular BP readings and can increase the safety of self-titration of antihypertensive medications.

Our study recruited participants from 3 different types of clinics: a federally qualified health center, a university clinic, and 2 private practices. Despite differences in electronic medical records and resources, clinic staff across settings shared similar perspectives, showing shared experience and suggesting broad applicability of Med Pro Care.

Recent studies from China have supported multimodel digital interventions for the management of hypertension [44,45]. A similar intervention in Pakistan based on the Health Belief Model included reminders, education, and 24/7 individual support, which improved medication adherence and systolic BP [46]. These studies share similarities in addressing social factors by addressing individual needs for care and support but lack a focus on the unique needs of refugee and immigrant patients. Our study highlights this gap and supports future testing of digital hypertension self-management platforms tailored to refugee and immigrant populations.

Limitations

Despite being among the few in the field who researched refugee and immigrant population providers, our study had notable limitations. First, our sample was limited to 1 county and may not reflect clinics elsewhere. Second, clinic staff with heavier patient loads may have been underrepresented due to the lengthy interview process. Nonetheless, San Diego's large refugee and immigrant population provides valuable insights, and future studies should evaluate digital self-management interventions in broader settings to avoid exacerbating disparities.

Conclusions

While self-management interventions have been shown to be effective in managing hypertension, it is essential to incorporate input from clinic staff that care for refugee and immigrant populations to understand the real-world implications of eHealth interventions. Clinic staff generally found the Med Pro Care software easy to use and feasible for integration into care, supporting the role of digital tools in addressing hypertension management among refugees and immigrants. Our study demonstrated that the refugee and immigrant communities have unique needs in hypertension management related to access, language, and clinic support. The implementation of self-management is supported by clinic staff because of its potential to empower patients by facilitating autonomy and a more active role in patient engagement in health care. Hypertension self-management interventions should focus on increasing health literacy, self-efficacy, improving access to health care, and leveraging digital health to streamline workflows for clinics and enhance communication with patients. Future studies should focus on the development and evaluation of contextually tailored hypertension self-management interventions that are responsive to the unique needs of the heterogeneous US refugee and immigrant population.

XSL•FO
RenderX

Acknowledgments

First and foremost, the authors would like to thank our participants for their time, trust, and perspectives. The authors are grateful to the various staff members and administrative leadership at Family Health Centers of San Diego and the Majdal Community Center in El Cajon, San Diego, for providing in-kind support for the development of study procedures, onboarding of study staff, participant recruitment, and participant interface. The authors are also grateful to Withing's RPM for Solutions and the San Diego division of the American Heart Association for their donation of blood pressure monitors used in our research. Finally, the authors would like to thank the Al-Rousan laboratory members for their assistance with data collection, interview transcription, and feedback throughout the research process and manuscript preparation.

This project was made possible by funding from the National Health Institute (National Heart, Lung, and Blood Institute [NHLBI] grant K23HL148530). NM is funded by NHLBI grant R00-HL148511, National Institute on Minority Health and Health Disparities grants P50MD017341 (subproject 8126) and 3P50MD017341-03S2.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Qualitative interview.
[PNG File, 94 KB - humanfactors_v12i1e66176_app1.png]

Multimedia Appendix 2
Tasks completed by participants.
[PNG File, 100 KB - humanfactors_v12i1e66176_app2.png]

Multimedia Appendix 3
System Usability Survey.
[PNG File, 45 KB - humanfactors_v12i1e66176_app3.png]

Multimedia Appendix 4
NASA Task Load Index.
[PNG File, 84 KB - humanfactors_v12i1e66176_app4.png]

References

1. Mills KT, Stefanescu A, He J. The global epidemiology of hypertension. Nat Rev Nephrol 2020 Apr;16(4):223-237. [doi: 10.1038/s41581-019-0244-2] [Medline: 32024986]
2. Indicator metadata registry list. World Health Organization. URL: https://tinyurl.com/wpdj654v [accessed 2024-03-19]
3. Lu Y, Lan T. Global, regional, and national burden of hypertensive heart disease during 1990-2019: an analysis of the global burden of disease study 2019. BMC Public Health 2022 Apr 27;22(1):841. [doi: 10.1186/s12889-022-13271-0] [Medline: 35473603]
4. Golub NI, Seplaki CL, Stockman DL, Thevenet-Morrison K, Fernandez D, Fisher SJ. Impact of length of residence in the United States on risk of diabetes and hypertension in resettled refugees. J Immigr Minor Health 2018 Apr;20(2):296-306. [doi: 10.1007/s10903-017-0636-y] [Medline: 28852919]
5. Zibara V, Costanian C, Al Haddad N, Kilani H, Tohme F, Aoun Bahous S. Epidemiology and management of hypertension among refugees in the Middle East: a review of the literature. J Hum Hypertens 2021 Aug;35(8):657-666. [doi: 10.1038/s41371-021-00508-5] [Medline: 33654233]
6. Fazel M, Wheeler JG, Danesh J. Prevalence of serious mental disorder in 7000 refugees resettled in western countries: a systematic review. Lancet 2005;365(9467):1309-1314. [doi: 10.1016/S0140-6736(05)61027-6] [Medline: 15823380]
7. Nicholson BL. The influence of pre-emigration and postemigration stressors on mental health: a study of Southeast Asian refugees. Soc Work Res 1997 Mar 1;21(1):19-31. [doi: 10.1093/swr/21.1.19]
8. Al-Rousan T, AlHeresh R, Saadi A, et al. Epidemiology of cardiovascular disease and its risk factors among refugees and asylum seekers: systematic review and meta-analysis. Int J Cardiol Cardiovasc Risk Prev 2022 Mar;12:200126. [doi: 10.1016/j.ijcrp.2022.200126] [Medline: 35199106]

XSL•FO
RenderX

9. Morris MD, Popper ST, Rodwell TC, Brodine SK, Brouwer KC. Healthcare barriers of refugees post-resettlement. J Community Health 2009 Dec;34(6):529-538. [doi: 10.1007/s10900-009-9175-3] [Medline: 19705264]
10. Alawa J, Zarei P, Khoshnood K. Evaluating the provision of health services and barriers to treatment for chronic diseases among Syrian refugees in Turkey: a review of literature and stakeholder interviews. Int J Environ Res Public Health 2019 Jul 25;16(15):2660. [doi: 10.3390/ijerph16152660] [Medline: 31349639]
11. Mulugeta W, Xue H, Glick M, Min J, Noe MF, Wang Y. Disease burdens and risk factors for diabetes, hypertension, and hyperlipidemia among refugees in Buffalo, New York, 2004-2014. J Health Care Poor Underserved 2019;30(3):1119-1131. [doi: 10.1353/hpu.2019.0077] [Medline: 31422992]
12. Marshall GN, Schell TL, Wong EC, et al. Diabetes and cardiovascular disease risk in Cambodian refugees. J Immigr Minor Health 2016 Feb;18(1):110-117. [doi: 10.1007/s10903-014-0142-4] [Medline: 25651882]
13. Albahsahli B, Bridi L, Aljenabi R, et al. Impact of United States refugee ban and discrimination on the mental health of hypertensive Arabic-speaking refugees. Front Psychiatry 2023;14:1083353. [doi: 10.3389/fpsyt.2023.1083353] [Medline: 37636820]
14. Asgary R, Segar N. Barriers to health care access among refugee asylum seekers. J Health Care Poor Underserved 2011 May;22(2):506-522. [doi: 10.1353/hpu.2011.0047] [Medline: 21551930]
15. Martínez-Ibáñez P, Marco-Moreno I, Peiró S, et al. Home blood pressure self-monitoring plus self-titration of antihypertensive medication for poorly controlled hypertension in primary care: the ADAMPA randomized clinical trial. J Gen Intern Med 2023 Jan;38(1):81-89. [doi: 10.1007/s11606-022-07791-z] [Medline: 36219303]
16. McManus RJ, Mant J, Haque MS, et al. Effect of self-monitoring and medication self-titration on systolic blood pressure in hypertensive patients at high risk of cardiovascular disease: the TASMIN-SR randomized clinical trial. JAMA 2014 Aug 27;312(8):799-808. [doi: 10.1001/jama.2014.10057] [Medline: 25157723]
17. Bray EP, Holder R, Mant J, McManus RJ. Does self-monitoring reduce blood pressure? Meta-analysis with meta-regression of randomized controlled trials. Ann Med 2010 Jul;42(5):371-386. [doi: 10.3109/07853890.2010.489567] [Medline: 20504241]
18. Madsen LB, Kirkegaard P, Pedersen EB. Blood pressure control during telemonitoring of home blood pressure. A randomized controlled trial during 6 months. Blood Press 2008;17(2):78-86. [doi: 10.1080/08037050801915468] [Medline: 18568696]
19. Bradbury K, Morton K, Band R, et al. Understanding how primary care practitioners perceive an online intervention for the management of hypertension. BMC Med Inform Decis Mak 2017 Jan 9;17(1):5. [doi: 10.1186/s12911-016-0397-x] [Medline: 28069041]
20. Soghikian K, Casper SM, Fireman BH, et al. Home blood pressure monitoring. Effect on use of medical services and medical care costs. Med Care 1992 Sep;30(9):855-865. [Medline: 1518317]
21. Fletcher BR, Hinton L, Hartmann-Boyce J, Roberts NW, Bobrovitz N, McManus RJ. Self-monitoring blood pressure in hypertension, patient and provider perspectives: a systematic review and thematic synthesis. Patient Educ Couns 2016 Feb;99(2):210-219. [doi: 10.1016/j.pec.2015.08.026] [Medline: 26341941]
22. Yeh PT, Rhee DK, Kennedy CE, et al. Self-monitoring of blood pressure among women with hypertensive disorders of pregnancy: a systematic review. BMC Pregnancy Childbirth 2022 May 31;22(1):454. [doi: 10.1186/s12884-022-04751-7] [Medline: 35641913]
23. Sheppard JP, Tucker KL, Davison WJ, et al. Self-monitoring of blood pressure in patients with hypertension-related multi-morbidity: systematic review and individual patient data meta-analysis. Am J Hypertens 2020 Mar 13;33(3):243-251. [doi: 10.1093/ajh/hpz182] [Medline: 31730171]
24. Bobrie G, Postel-Vinay N, Delonca J, Corvol P, SETHI Investigators. Self-measurement and self-titration in hypertension: a pilot telemedicine study. Am J Hypertens 2007 Dec;20(12):1314-1320. [doi: 10.1016/j.amjhyper.2007.08.011] [Medline: 18047923]
25. Cheng C, Studdiford JS, Diamond JJ, Chambers CV. Primary care physician beliefs regarding usefulness of self-monitoring of blood pressure. Blood Press Monit 2003 Dec;8(6):249-254. [doi: 10.1097/00126097-200312000-00005] [Medline: 14688555]
26. Gondi S, Ellis S, Gupta M, et al. Physician perceived barriers and facilitators for self-measured blood pressure monitoring-a qualitative study. PLoS One 2021;16(8):e0255578. [doi: 10.1371/journal.pone.0255578] [Medline: 34415946]
27. Shahin W, Kennedy GA, Stupans I. A qualitative exploration of the impact of knowledge and perceptions about hypertension in medication adherence in Middle Eastern refugees and migrants. Explor Res Clin Soc Pharm 2021 Sep;3:100038. [doi: 10.1016/j.rcsop.2021.100038] [Medline: 35480607]
28. Jones MI, Greenfield SM, Bray EP, et al. Patient self-monitoring of blood pressure and self-titration of medication in primary care: the TASMINH2 trial qualitative study of health professionals' experiences. Br J Gen Pract 2013 Jun;63(611):e378-e385. [doi: 10.3399/bjgp13X668168] [Medline: 23735408]
29. Lanssens D, Vandenberk T, Smeets CJ, et al. Remote monitoring of hypertension diseases in pregnancy: a pilot study. JMIR Mhealth Uhealth 2017 Mar 9;5(3):e25. [doi: 10.2196/mhealth.6552] [Medline: 28279948]
30. Demographics. SanDiegoCounty.gov. URL: https://www.sandiegocounty.gov/hhsa/statistics_demographics.html [accessed 2022-08-14]

XSL•FO
RenderX

31. Ogedegbe G, Teresi JA, Williams SK, et al. Home blood pressure telemonitoring and nurse case management in Black and Hispanic patients with stroke: a randomized clinical trial. JAMA 2024 Jul 2;332(1):41-50. [doi: 10.1001/jama.2024.6609] [Medline: 38842799]
32. Victor RG, Lynch K, Li N, et al. A cluster-randomized trial of blood-pressure reduction in Black barbershops. N Engl J Med 2018 Apr 5;378(14):1291-1301. [doi: 10.1056/NEJMoa1717250] [Medline: 29527973]
33. Khairat S, Coleman C, Newlin T, et al. A mixed-methods evaluation framework for electronic health records usability studies. J Biomed Inform 2019 Jun;94:103175. [doi: 10.1016/j.jbi.2019.103175] [Medline: 30981897]
34. Sauro J. Measuring usability with the System Usability Scale (SUS). MeasuringU. 2011 Feb 3. URL: https://measuringu.com/sus/ [accessed 2021-06-24]
35. NASA TLX: paper/pencil version. NASA Ames Human Systems Integration Division. URL: https://humansystems.arc.nasa.gov/groups/tlx/tlxpaperpencil.php [accessed 2021-06-24]
36. Guest G, MacQueen KM, Namey EE. Applied Thematic Analysis: SAGE Publications, Inc; 2012. [doi: 10.4135/9781483384436]
37. Korstjens I, Moser A. Series: practical guidance to qualitative research. Part 4: trustworthiness and publishing. Eur J Gen Pract 2018 Dec;24(1):120-124. [doi: 10.1080/13814788.2017.1375092] [Medline: 29202616]
38. Crabtree BF, Miller WL. A template approach to text analysis: developing and using codebooks. In: Doing Qualitative Research: SAGE Publications, Inc; 1992:93-109.
39. Selvan K, Leekha A, Abdelmeguid H, Malvankar-Mehta MS. Barriers adult refugees face to community health and patient engagement: a systematic review. Glob Public Health 2022 Dec;17(12):3412-3425. [doi: 10.1080/17441692.2022.2121846] [Medline: 36074889]
40. Jones MI, Greenfield SM, Bray EP, et al. Patients' experiences of self-monitoring blood pressure and self-titration of medication: the TASMINH2 trial qualitative study. Br J Gen Pract 2012 Feb;62(595):e135-e142. [doi: 10.3399/bjgp12X625201] [Medline: 22520791]
41. Munyungula J, Shakwane S. Self-monitoring of blood pressure for preeclampsia patients: knowledge and attitudes. Curationis 2021 Sep 21;44(1):e1-e8. [doi: 10.4102/curationis.v44i1.2195] [Medline: 34636622]
42. Shahin W, Kennedy GA, Cockshaw W, Stupans I. The role of refugee and migrant migration status on medication adherence: mediation through illness perceptions. PLoS One 2020;15(1):e0227326. [doi: 10.1371/journal.pone.0227326] [Medline: 31923264]
43. Monahan M, Jowett S, Nickless A, et al. Cost-effectiveness of telemonitoring and self-monitoring of blood pressure for antihypertensive titration in primary care (TASMINH4). Hypertension 2019 Jun;73(6):1231-1239. [doi: 10.1161/HYPERTENSIONAHA.118.12415] [Medline: 31067190]
44. Zhou H, Wang X, Yang Y, et al. Effect of a multicomponent intervention delivered on a web-based platform on hypertension control: a cluster randomized clinical trial. JAMA Netw Open 2022 Dec 1;5(12):e2245439. [doi: 10.1001/jamanetworkopen.2022.45439] [Medline: 36477479]
45. Wang Y, Guo F, Wang J, et al. Efficacy of a WeChat-based multimodal digital transformation management model in new-onset mild to moderate hypertension: randomized clinical trial. J Med Internet Res 2023 Dec 4;25:e52464. [doi: 10.2196/52464] [Medline: 38048156]
46. Arshed M, Mahmud A, Minhat HS, Lim PY, Zakar R. Effectiveness of a multifaceted mobile health intervention (multi-aid-package) in medication adherence and treatment outcomes among patients with hypertension in a low- to middle-income country: randomized controlled trial. JMIR Mhealth Uhealth 2024 Jun 19;12:e50248. [doi: 10.2196/50248] [Medline: 38896837]

Abbreviations

BP: blood pressure
FHCSD: Family Health Centers of San Diego
MA: medical assistant
PCP: primary care provider
TASMIN: Targets and Self-Management for the Control of Blood Pressure

Edited by A Kushniruk, M Waite; submitted 05.09.24; peer-reviewed by B Olujobi, K Culhane-Pera, Y Wang; revised version received 21.08.25; accepted 28.08.25; published 24.10.25.

Please cite as:
Gonzalez C, Real M, Ahmadi N, Aljenabi R, Bridi L, Makarem N, Godino J, Al-Rousan T
Providers' Perspective on the Feasibility of Digital Self-Management of Blood Pressure in Refugees: Mixed Methods Study
JMIR Hum Factors 2025;12:e66176
URL: https://humanfactors.jmir.org/2025/1/e66176
doi:10.2196/66176

XSL•FO
RenderX

© Carol Gonzalez, Marcos Real, Nargis Ahmadi, Raghad Aljenabi, Lana Bridi, Nour Makarem, Job Godino, Tala Al-Rousan. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 24.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Factors Determining Acceptance of Internet of Things in Medical Education: Mixed Methods Study

Khadija Alhumaid[1*], PhD; Kevin Ayoubi[2*], PhD; Maha Khalifa[1*], PhD; Said Salloum[3*], PhD

[1]Research & Innovation Division, Rabdan Academy, Abu Dhabi, United Arab Emirates

[2]English Faculty, Higher Colleges of Technology – CERT, Al Ain, United Arab Emirates

[3]School of Computing, Skyline University College, University City, Sharjah, United Arab Emirates

[*]all authors contributed equally

Corresponding Author:
Said Salloum, PhD
School of Computing, Skyline University College, University City, Sharjah, United Arab Emirates

Abstract

Background: The global increase in the Internet of Things (IoT) adoption has sparked interest in its application within the educational sector, particularly in colleges and universities. Previous studies have often focused on individual attitudes toward IoT without considering a multiperspective approach and have overlooked the impact of IoT on the technology acceptance model outside the educational domain.

Objective: This study aims to bridge the research gap by investigating the factors influencing IoT adoption in educational settings, thereby enhancing the understanding of collaborative learning through technology. It seeks to elucidate how IoT can facilitate learning processes and technology acceptance among college and university students in the United Arab Emirates.

Methods: A questionnaire was distributed to students across various colleges and universities in the United Arab Emirates, garnering 463 participants. The data collected were analyzed using a hybrid approach that integrates structural equation modeling (SEM) and artificial neural network (ANN), along with importance-performance map analysis to evaluate the significance and performance of each factor affecting IoT adoption.

Results: The study, involving 463 participants, identifies 2 primary levels at which factors influence the intention to adopt IoT technologies. Initial influences include technology optimism (TOP), innovation, and learning motivation, crucial for application engagement. Advanced influences stem from technology acceptance model constructs, particularly perceived ease of use (PE) and perceived usefulness (PU), which directly enhance adoption intentions. Detailed statistical analysis using partial least squares–SEM reveals significant relationships: TOP and innovativeness impact PE (β=.412, *P*=.04; β=.608, *P*=.002, respectively), and PU significantly influences TOP (β=.381, *P*=.04), innovativeness (β=.557, *P*=.003), and learning motivation (β=.752, *P*<.001). These results support our hypotheses (H1, H2, H3, H4, and H5). Further, the intention to use IoT is significantly affected by PE and usefulness (β=.619, *P*<.001; β=.598, *P*<.001, respectively). ANN modeling enhances these findings, showing superior predictive power (R^2=89.7%) compared to partial least squares–SEM (R^2=86.3%), indicating a more effective identification of nonlinear associations. Importance-performance map analysis corroborates these results, demonstrating the importance and performance of PU as most critical, followed by technology innovativeness and optimism, in shaping behavioral intentions to use IoT.

Conclusions: This research contributes methodologically by leveraging deep ANN architecture to explore nonlinear relationships among factors influencing IoT adoption in education. The study underscores the importance of both intrinsic motivational factors and perceived technological attributes in fostering IoT adoption, offering insights for educational institutions considering IoT integration into their learning environments.

(JMIR Hum Factors 2025;12:e58377) doi:10.2196/58377

KEYWORDS

collaborative learning; student; college; university; education; Internet of Things; IoT; technology acceptance model; technology optimism; TAM; experience; attitude; opinion; perception; perspective; acceptance; adoption; survey; questionnaire; ANN; deep learning; structural equation modeling; neural network; intent; use; medical education; artificial neural network; technology innovation

XSL•FO
RenderX

Introduction

Background

The use of Internet of Things (IoT) applications has been implemented in medical education projects [1]. IoT has the ability to completely transform the educational landscape by offering a more adaptable and quantifiable educational system that unites teachers and students under a single technological roof [2]. IoT innovation plays a significant role in transforming training at all levels, from school and college to university education [3]. Everyone, including students, instructors, and college campuses, can benefit from this innovation. Educators and administrators can leverage the power of IoT to connect people with devices and data, enabling them to gain valuable insights that have not been previously used in education. The traditional human-centered educational system has been transformed into an IoT-based one [4-6] using IoT. IoT has been leveraged to change the conventional personalized schooling system to an IoT-based system [4-6]. Figure 1 illustrates a cohort of university students actively engaging with IoT technology within an academic context. These individuals are adorned with sophisticated wearable devices, including smart glasses and smartwatches, which project information into their visual field, indicative of a digital and augmented educational milieu. The students are depicted with a palpable focus, using a portable electronic apparatus, likely for the manipulation of or interaction with the data overlay provided by their IoT devices.

The IoT has the potential to transform institutional practices and enhance learning capabilities across various levels and domains. University lecturers, students, and support staff can leverage large IoT platforms successfully. However, there is room for improvement in the utilization of IoT technology across different educational institutions. Researchers, scholars, and students can collaborate to develop IoT systems, devices, applications, and services, leading to the evolution of the educational environment as an increasingly dynamic and globally relevant subject. Better IoT deployments in colleges and universities have significantly improved positively to the creation of efficient and useful educational resources [7,8].

Figure 1. Students using augmented reality and Internet of Things technology (generated by Openart AI).



XSL•FO
RenderX

This research aims to explore how IoT technology is adopted and used by educators and students in universities across the Middle East. By integrating the technology acceptance model (TAM) and additional external factors, the study seeks to evaluate the effectiveness of IoT as an educational tool from the perspectives of both students and educators. Using the TAM as a metric, along with external variables, the research intends to identify and analyze the factors that influence the acceptance and use of IoT in medical education. The analysis uses a hybrid framework that combines structural equation modeling (SEM) and artificial neural networks (ANN) to examine how intrinsic motivational factors and perceived technological attributes affect IoT adoption. The SEM-ANN approach was specifically chosen to leverage the strengths of both methodologies [9,10]. SEM is highly effective in assessing the relationships between observed and latent variables, providing clarity on both the direct and indirect effects within the hypothesized model [11]. This method allows for robust statistical analysis capabilities, making it ideal for hypothesis testing and understanding the structural relationships among the theoretical constructs [12]. Conversely, ANN is used for its superior ability to model complex nonlinear relationships between variables, which are often not adequately captured by traditional linear models like SEM. ANN's data-driven nature allows it to directly learn and adapt to these relationships from the data, thereby enhancing the model's predictive accuracy and robustness. By integrating SEM with ANN, the study not only validates the theoretical framework through SEM's rigorous statistical analysis but also enhances the predictive power and generalization capabilities of the model with ANN's computational intelligence. This hybrid approach is particularly effective in exploring deeper, nonlinear interactions within the data, offering a more comprehensive understanding of the factors influencing IoT adoption. Given the complexities of modern datasets, which often exhibit nonlinear and digital behaviors among variables, this methodological integration is well-suited to achieving the research objectives and provides a justified, robust approach for the study.

The proposed hypotheses are: (1) intrinsic motivational factors like technology optimism (TOP), innovation, and learning motivation (LMT) significantly impact students' intention to adopt IoT; (2) perceived ease of use (PE) and perceived usefulness (PU), fundamental components of the TAM, strongly predict IoT adoption intentions among students in the United Arab Emirates.

This paper is organized as follows: "Literature Review" section delves into the existing research related to the adoption of IoT technologies, highlighting key theories and previous findings that set the groundwork for this study. Section "Methodology" details the research design, sampling methods, data collection procedures, and analytical techniques used to investigate the hypotheses. Section "Findings" presents the results of the data analysis, offering quantitative insights into the factors influencing IoT adoption. Section "Discussion" interprets the findings in the context of the existing literature, discussing the implications for theory and practice. Finally, the section "Conclusions" summarizes the study's main contributions, outlines its limitations, and suggests directions for future research. This structure is designed to provide a clear and logical progression through the topics covered, facilitating a comprehensive understanding of the study's scope and conclusions.

Related Work

The review of existing literature on the topic of IoT has explored both practical and theoretical aspects, suggesting a correlation between IoT and other factors such as self-efficacy, technology utilization, motivation, security, privacy, training, and more [13-19], indicating a relationship between the IoT and other elements including self-reliance, technological use, inspiration, safety, privacy, schooling, and beyond [13-19]. Likewise, research has examined the impact of IoT in conjunction with TAM and other external factors, and IoT with the help of TAM and external factors [20,21].

As IoT applications become more intricate, they can significantly impact learning. The difficulty is the IoT technology's quick development, which requires diverse skills ranging from developing IoT applications to incorporating devices into management systems that analyze device-generated data [14,22]. Previous research has highlighted the significance of IoT in addressing challenges students encounter when using modern IoT apps and gadgets. Potential solutions include focusing on computational thinking education, assisting students in solving challenges, and providing clear instructions and training to facilitate the integration of new students with IoT devices and encourage training in it.

A different approach was taken in another study, where a workshop was offered to address the importance of IoT. The results showed that students found the workshop highly satisfactory for learning about IoT, improving problem-solving skills, and enhancing problem-solving capabilities, while also finding it enjoyable [13-15].

Several outside variables, including drive, contentment, ease of use, effectiveness, involvement, and interest, have been investigated to explore the relationship between IoT and students' attitudes. The most significant factors are motivation and enjoyment regarding IoT technology, which are essential for its acceptance among students. While satisfaction and performance indirectly affect the application of IoT, greater contentment levels are not proportional to higher interactions [18,19]. Recent studies have used qualitative and quantitative methods to examine the uptake of IoT and have identified favorable attitudes, ease of use, contentment, affordability, basic knowledge, security, and privacy as crucial factors that influence its adoption. Apart from motivation and enjoyment, training and experience are also crucial in uptaking IoT devices, and educational training workshops can significantly improve the learning curve, digital learning, real-life applications, and problem-solving skills [16,17,23].

Previous research has explored the various benefits of implementing IoT technologies in educational settings, including the ability for teachers to gain insight into students' performance and knowledge levels, as well as the potential for improved teaching quality. As a result, the use of IoT has the potential to greatly impact learning environments, leading to a more

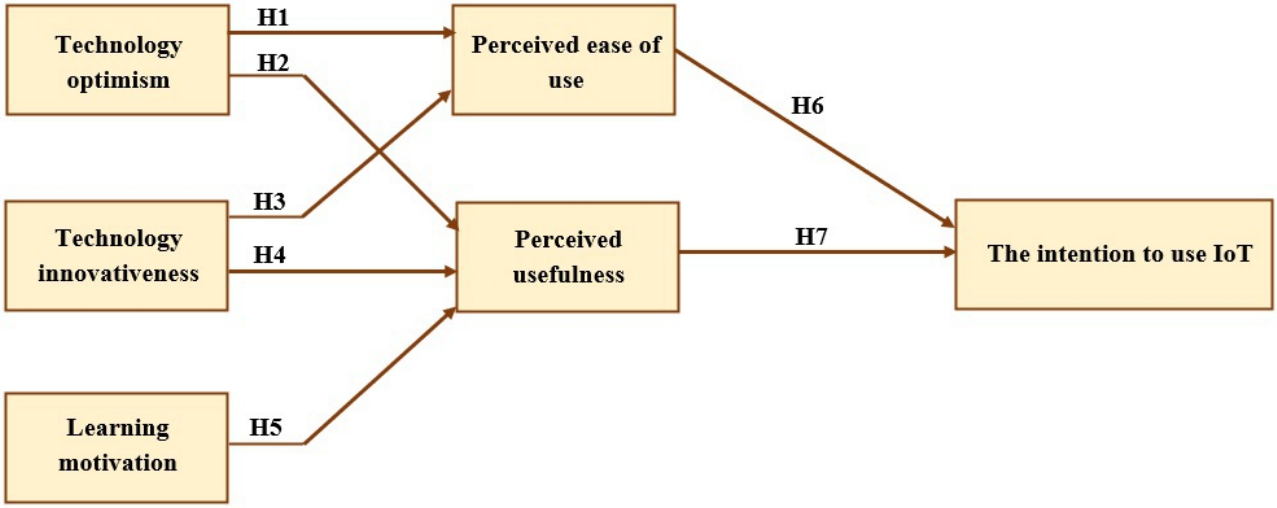
XSL•FO
RenderX

advanced educational landscape. This could also lead to changes in how we interact and collaborate as a society, as IoT technology continues to connect us in new ways [24,25]. While previous studies have investigated students' attitudes toward IoT from different angles [14,22,26], there has been limited research on the importance of IoT technology from the perspective of teachers' knowledge of the significance of IoT innovation from the educational viewpoint of instructors and how it impacts pedagogy in the classroom. Therefore, this study aims to address this gap by developing a model that combines both students' and teachers' attitudes to examine the effectiveness and efficiency of IoT in educational environments. By creating a framework that incorporates the viewpoints of learners and educators, this research intends to close this imbalance by analyzing the usefulness and efficacy of IoT in classrooms.

Developing Hypotheses and Theoretical Framework

The study paradigm shown in Figure 2 describes how the inspiration for learning, technological exuberance, and technological advancements influence learners' perceptions of the PU and ease of use (PE) of IoT devices. Such concepts have not yet been looked into in relation to IoT devices and applicability. Although earlier research has examined how these factors may affect IoT adoption plans [27], their individual effects on students' attitudes toward this technology have not been analyzed before.

Figure 2. Research model. IoT: Internet of Things.



LMT is a crucial factor impacting students' behavior in conventional and digital educational settings. It is an outside variable of the suggested approach and can be influenced by multiple factors, including the learning environment, expectations, and social values. Research has shown that LMT significantly impacts students' academic achievements [28]. Students must therefore be driven to acquire knowledge in order to achieve the suggested intended goals [29,30]. Highly motivated learners can achieve their learning goals spontaneously and willingly. Social cognitivism emphasizes the significance of requirements, conducts, and ideologies in education.

Researchers are investigating the influence of TOP and innovation on students' adoption of IoT in the education sector. The study aims to understand students' attitudes and behavior toward technology by using 3 theoretical concepts: inventiveness, technological exuberance, and learning incentive. The researchers have combined the (TAM) constructs with the aforementioned variables to accomplish their study's objectives. Several researchers have used the TAM model to evaluate the effectiveness of IoT technology and its applications [31,32]. The model explains why people accept technology based on their attitudes and beliefs and how these beliefs impact their behavioral intention toward technology use [33,34]. It is an established concept that supports embracing technology in various settings and strongly links it to technological progress [35-37]. The TAM model can forecast variables like technology elation and inventiveness that influence how well a certain technology is received [38,39]. TOP refers to users' favorable perception of technology, while technology innovativeness (TIN) refers to users' willingness to adopt technology early and lead to its use.

The PE and PU constructs of TAM are closely related to TOP and TIN. In the academic setting, students' perceptions of technology can be influenced by their peers and instructors. If the student's immediate academic circle has a positive view of a particular technology, the student will likely also develop a favorable opinion. Similarly, tech-savvy students tend to have a positive self-perception. Early on in the acceptance of technology, students are usually eager to pioneer using advanced technologies [40-44].

Furthermore, a student's optimism about technology is linked to their level of involvement in guiding its use. Similarly, technological optimism can significantly impact a student's attitude. IoT acceptance is greater for learners who are enthusiastic about experimenting with novel innovations. Technological innovators infrequently perceive new innovations as being challenging or outside their comprehension. Individuals are more inclined to regret not having the freedom to play around with novel technologies [45]. As a result, a number of hypotheses are put forth:

XSL•FO
RenderX

- H1: TOP affects positively the PE.
- H2: TOP affects positively the PU.
- H3: TIN affects positively the PE.
- H4: TIN affects positively the PU.
- H5: LMT affects positively the PU.
- H6: PE affects positively the intention to use Internet of Things (INT).
- H7: PU affects positively the INT.

By evaluating the association between TAM and its connected factors, this research seeks to add to a collection of current work. It has been said that it is important to measure these variables and analyze their relationship to TAM to assess the efficiency of IoT in classrooms. Past research and literature have mostly focused on preservice teachers, and there is a need to conduct more studies involving in-service teachers to enhance the practicality of the TAM model [46,47]. Thus, this study investigates the correlation between teachers' levels of TAM and their attitudes toward IoT acceptance when working with teachers from different fields.

Methods

Data Collection

The data collection for this study was carried out between January 20 and March 20, 2023, throughout the academic year 2023 - 2024's winter semester, at educational institutions in the United Arab Emirates. The research team used web-based surveys to collect data, with 500 questionnaires randomly distributed. Of these, 463 surveys were answered, resulting in a response rate of 93%. Some questionnaires were rejected due to missing values. Since a few surveys had no responses, they were discarded, leaving 769 usable questionnaires, which is considered an appropriate acceptable sample size according to Krejcie and Morgan [48]. Although the sample size exceeded the minimum requirements, the research team used SEM SmartPLS (version 3.2.7; SmartPLS GmbH) and SPSS Statistics (version 23; IBM Corp) to evaluate the hypotheses and confirm the relationship between variables. It is important to note that the hypotheses were based on previous theories related to IoT, which formed the foundation of this study.

Ethical Considerations

Ethical considerations were meticulously adhered to throughout the research process. All procedures involving human participants were approved by the institutional review board of the host universities in the United Arab Emirates (#RAREC00065), ensuring compliance with ethical standards. Informed consent was obtained from all participants before data collection, and they were informed of their rights to withdraw from the study at any time. Privacy and confidentiality were strictly maintained, with all data being anonymized and securely stored to prevent unauthorized access. Participants were not compensated for their participation, as the study involved minimal risk and was conducted as part of educational activities within the institutions involved.

Student's Personal Information

The demographic data of the respondents (N=463) are presented in Table 1. The data showed that 70% (n=325) of participants were female and 30% (n=138) were male. In terms of age, 42% (n=193) of participants were between 18 and 29 years, 35% (n=163) were between 30 and 39 years, 21% (n=99) were between 40 and 49 years, and 2% (n=8) were between 50 and 59 years. Regarding education, 80% (n=372) of participants held a bachelor's degree, 14% (n=62) held a master's degree, and 6% (n=29) held a doctoral degree. To obtain participants' willingness to participate, the research team used a purposive sampling approach. The participants came from various universities, academic levels, and programs relevant to this research. SPSS Statistics was used to analyze the demographic data.

Table . Demographic data of the respondents (N=463).

Category	Value, n (%)
Sex	
Female	325 (70)
Male	138 (30)
Age (years)	
Between 18 and 29	193 (42)
Between 30 and 39	163 (35)
Between 40 and 49	99 (21)
Between 50 and 59	8 (2)
Educational qualification	
Bachelor's degree	372 (80)
Master's degree	62 (14)
Doctorate	29 (6)

XSL•FO
RenderX

Study Instrument

This research proposed a survey instrument for validating the hypotheses. To assess the 6 constructs of the questionnaire, an additional 18 questions were included in the survey. The origins and histories of these constructs are presented in Table 2. To make the research more relevant, the researchers modified the questions from previous studies.

XSL•FO
RenderX

Table . Measurement items.

Constructs	Items	Definition	Instrument	Sources
TIN[a]	• TIN1 • TIN2 • TIN3	The term "technology innovativeness" describes a user's perception that they are at the forefront of technology use. Users who are pioneers in adopting new technologies typically do not view them as complicated or difficult to comprehend. Such users may feel a sense of regret if they miss the chance to experiment with new technologies.	• I accept IoT[b] technology to be used in my daily classes. • I am the only one. There is only me prepared to use IoT technology among my fellow students. • I am ready to use. I am prepared to use and experiment with the latest information technologies.	[45]
TOP[c]	• TOP1 • TOP2 • TOP3	A person's readiness to use technology is known as technological optimism.	• I am ready to test. Prepared to take the test IoT technology. • To complete my assignments to finish my homework, I will be using IoT. • I will learn more with the help of my preparedness to use IoT.	[2,49]
LMT[d]	• LMT1 • LMT2 • LMT3	The concept of learning motivation is used to measure the behavioral intention to use technology. Motivation learning is composed of 4 key components, which are attention, relevance, confidence, and satisfaction. These components have been identified in previous studies [50,51].	• I can improve my focus by using IoT for my daily classes. • Using IoT makes me feel more confident. • Using IoT for study purposes satisfies me.	[50,51]
PE[e]	• PE1 • PE2 • PE3	The TAM[f] was introduced by Davis [52] as a means of assessing the effectiveness and acceptance of technology. The model includes the concept of PE, which refers to the user's perception of how effortless it is to use the technology.	• IoT technology being simple will polish my skills. • I can improve my learning achievements by using IoT technology. • IoT is simple and easy to use.	[52]
PU[g]	• PU1 • PU2 • PU3	Usefulness refers to PU that the users of technology may see.	• IoT technology will hugely benefit me. • IoT will make my abilities and skills better. • Using IoT for my daily classes is beneficial.	[52]
INT[h]	• INT1 • INT2 • INT3	An individual's view of what others think about a certain behavior is known as a behavioral intention to use.	• IoT will be my go-to for daily tasks. • In the future, I will be using IoT. • IoT technology will be my recommendation to every student.	[53]

[a]TIN: technology innovativeness.

[b]IoT: Internet of Things.

XSL•FO
RenderX

[c]TOP: technology optimism.
[d]LMT: learning motivation.
[e]PE: perceived ease of use.
[f]TAM: technology acceptance model.
[g]PU: perceived usefulness.
[h]INT: intention to use Internet of Things.

Survey Structure

The questionnaire survey given to the students has 3 sections. Three components make up the survey that is provided to learners: (1) personal data is the focus of the first section, (2) the general question related to the "Intention to Use IoT" is the second section, and (3) 15 items that deal with TIN, TOP, LMT, PE, and PU is present in the third section.

A 5-point Likert scale was used to assess the 18 items, with response options ranging from 1=strongly disagree to 2=disagree, 3=neutral, 4=agree, and 5=strongly agree.

Results

Data Analysis

In this study, the gathered data were analyzed using SmartPLS (version 3.2.7) software through the partial least squares–structural equation modeling (PLS-SEM) technique [54-56]. The evaluation consisted of 2 stages: the measurement model and the structural model [57,58]. PLS-SEM was selected for this research after considering several factors.

The selection of PLS-SEM for this study was based on several reasons. First, PLS-SEM is preferred when the study aims to build on an existing theory [59]. Second, PLS-SEM is effective in handling complex models in exploratory research. Third, PLS-SEM analyzes the entire model as a single entity rather than dividing it into components [60]. Finally, PLS-SEM allows for the simultaneous analysis of structural and measurement models, leading to more accurate results [61].

Convergent Validity

Hair et al [57] suggested that to assess the measurement model, it is important to examine the construct reliability (including Cronbach ɑ and composite reliability) and validity (including convergent and discriminant validity). Table 3 shows that the Cronbach ɑ readings are greater than the suggested criterion of 0.7, and vary from 0.797 to 0.858 [62], indicating good construct reliability. Similarly, the proposed criterion is also greater than the composite reliability numbers, which vary from 0.735 to 0.858 [63]. Factor loading and average-variance extracted (AVE) analysis are required to evaluate convergent validity [57]. The factor loading values in Table 3 are higher than the recommended threshold of 0.7, and the AVE readings are greater than the suggested limit of 0.5, spanning from 0.556 to 0.712. These results suggest that there is convergent validity.

XSL•FO
RenderX

Table . Convergent validity results which assures acceptable values (factor loading, CA[a], CR[b], Dijkstra-Henseler's ρ ≥0.70, and AVE[c] >0.5).

Constructs and items	Factor loading	CA	CR	AVE
TIN[d]		0.856	0.824	0.608
TIN1	0.723			
TIN2	0.873			
TIN3	0.858			
TOP[e]		0.842	0.858	0.701
TOP1	0.804			
TOP2	0.816			
TOP3	0.801			
LMT[f]		0.815	0.851	0.712
LMT1	0.765			
LMT2	0.844			
LMT3	0.758			
PE[g]		0.797	0.735	0.662
PE1	0.858			
PE2	0.825			
PE3	0.758			
PU[h]		0.858	0.853	0.556
PU1	0.801			
PU2	0.829			
PU3	0.732			
INT[i]		0.825	0.843	0.612
INT1	0.812			
INT2	0.721			
INT3	0.749			

[a]CA: Cronbach ɑ.

[b]CR: composite reliability.

[c]AVE: average-variance extracted.

[d]TIN: technology innovativeness.

[e]TOP: technology optimism.

[f]LMT: learning motivation.

[g]PE: perceived ease of use.

[h]PU: perceived usefulness.

[i]INT: intention to use Internet of Things.

Discriminant Validity

To assess discriminant validity, 2 criteria were recommended: the Heterotrait-Monotrait ratio and the Fornell-Larcker criterion [57]. According to Table 4, the Fornell-Larcker criterion is met, as the AVE and its square root for each construct exceed its correlation with other constructs [64]. Table 5 displays the Heterotrait-Monotrait ratio results, which demonstrate that each construct's value is below the threshold value of 0.85 [65]. This suggests that discriminant validity exists, and the measurement model's reliability and validity were confirmed without any issues. As a result, the collected data can be used for analyzing the structural model.

XSL•FO
RenderX

Table . Fornell-Larcker scale.

	TIN[a]	TOP[b]	LMT[c]	PE[d]	PU[e]	INT[f]
TIN	0.864[g]	—[h]	—	—	—	—
TOP	0.675	0.883[g]	—	—	—	—
LMT	0.182	0.263	0.723[g]	—	—	—
PE	0.664	0.245	0.236	0.861[g]	—	—
PU	0.664	0.283	0.373	0.313	0.812[g]	—
INT	0.540	0.573	0.275	0.407	0.286	0.890[g]

[a]TIN: technology innovativeness.

[b]TOP: technology optimism.

[c]LMT: learning motivation.

[d]PE: perceived ease of use.

[e]PU: perceived usefulness.

[f]INT: intention to use Internet of Things.

[g]These values represent the square root of the average variance extracted for each construct, according to the Fornell-Larcker criterion. They are placed diagonally to demonstrate discriminant validity. A construct should share more variance with its indicators than with other constructs (off-diagonal correlations).

[h]Not applicable.

Table . Heterotrait-Monotrait ratio.

	TIN[a]	TOP[b]	LMT[c]	PE[d]	PU[e]	INT[f]
TIN	—[g]	0.355	0.473	0.113	0.741	0.336
TOP	0.355	—	0.406	0.512	0.579	0.512
LMT	0.473	0.406	—	0.702	0.559	0.021
PE	0.113	0.512	0.702	—	0.328	0.363
PU	0.741	0.579	0.559	0.328	—	0.486
INT	0.336	0.512	0.021	0.363	0.486	—

[a]TIN: technology innovativeness.

[b]TOP: technology optimism.

[c]LMT: learning motivation.

[d]PE: perceived ease of use.

[e]PU: perceived usefulness.

[f]INT: intention to use Internet of Things.

[g]Not applicable.

Hypotheses Testing Using PLS-SEM

To assess whether the theoretical constructs of the structural model are interconnected, the study used Smart PLS with maximum likelihood estimation to create a structural equation model [66]. The proposed hypotheses were then analyzed using this model. The results indicated a high level of predictive power for the model, with 86.3% of the variance in INT being accounted for, as shown in Figure 3 and Table 6.

Table 7 provides information on the β values, *t* values, and *P* values of all developed hypotheses based on the findings produced using the PLS-SEM technique. The researchers have confirmed each hypothesis. The empirical data supported H1, H2, H3, H4, H5, H6, and H7 following the data analysis hypotheses. This study demonstrates that TOP and TIN have a noteworthy influence on PE with respective regression coefficients (β=.412, *P*=.04), and (β=.608, *P*=.002), supporting H1 and H3. Moreover, the results indicate that PU is significantly impacted by TOP (β=.381, *P*=.04), TIN (β=.557, *P*=.003), and LMT (β=.752, *P*<.001), supporting hypotheses H2, H4, and H5, respectively. Finally, the study reveals that the relationship between PE and PU significantly affects INT with respective regression coefficients (β=.619, *P*<.001) and (β=.598, *P*<.001), supporting H6 and H7.

XSL•FO
RenderX

Figure 3. The path coefficients of the research model. IoT: Internet of Things. *P<.05; **P<.01; ***P<.001.

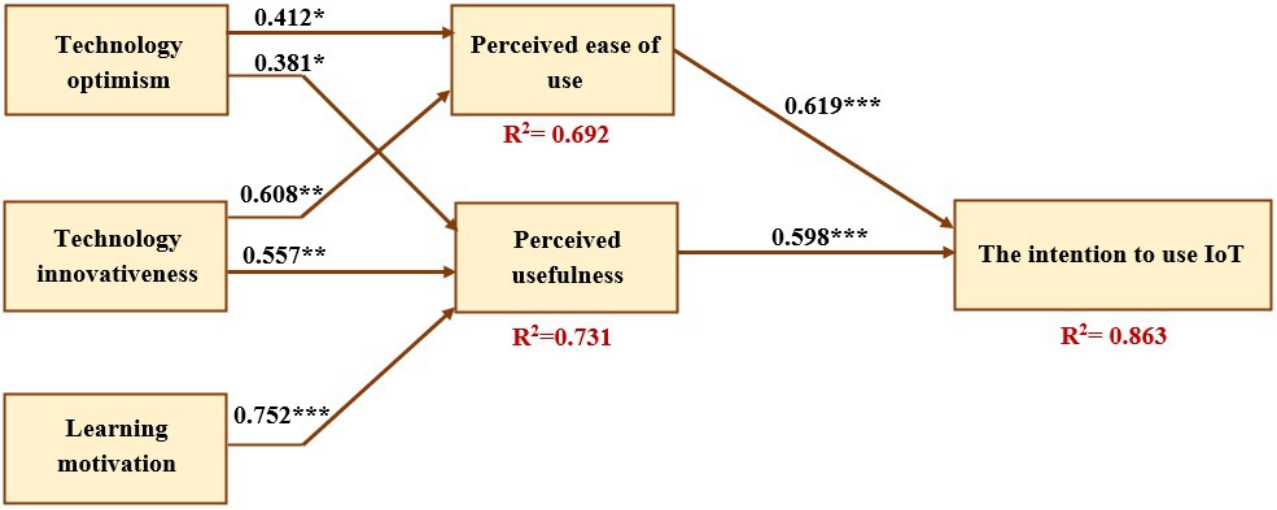


Table . The R^2 values for the endogenous latent variables.

Construct	R^2	Results
PE[a]	0.692	High
PU[b]	0.731	High
INT[c]	0.863	High

[a]PE: perceived ease of use.

[b]PU: perceived usefulness.

[c]INT: intention to use Internet of Things.

Table . Results of hypotheses testing for the research model at significance levels of .01 and .05.

H	Relationship	Path	*t* test (*df*)	*P* value	Direction	Decision
H1	TOP[a] -> PE[b]	0.412	5.552 (461)	.048	Positive	Supported
H2	TOP -> PU[c]	0.381	4.843 (461)	.04	Positive	Supported
H3	TIN[d] -> PE	0.608	10.247 (461)	.002	Positive	Supported
H4	TIN -> PU	0.557	9.358 (461)	.003	Positive	Supported
H5	LMT[e] -> PU	0.752	14.450 (461)	<.001	Positive	Supported
H6	PE -> INT[f]	0.619	16.753 (461)	<.001	Positive	Supported
H7	PU -> INT	0.598	14.195 (461)	<.001	Positive	Supported

[a]TOP: technology optimism.

[b]PE: perceived ease of use.

[c]PU: perceived usefulness.

[d]TIN: technology innovativeness.

[e]LMT: learning motivation.

[f]INT: intention to use Internet of Things.

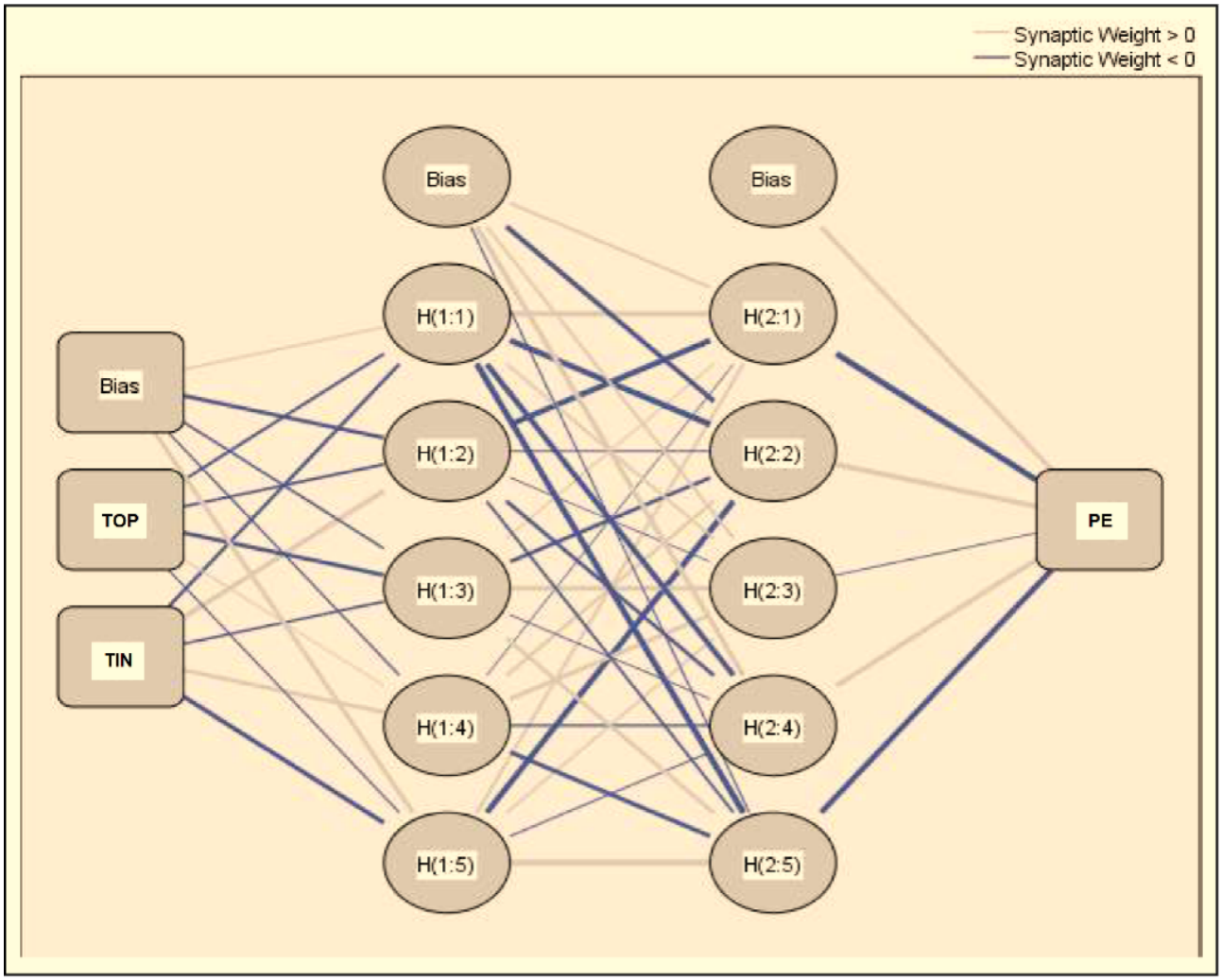
ANN Results

The predictors identified during the PLS-SEM analysis are further investigated through ANN analysis using SPSS software. The PLS-SEM analysis identified TIN, TOP, LMT, PE, and PU as critical factors; therefore, ANN analysis also considers these 3 factors only. The structure of the ANN model is based on behavioral intention as an output neuron and TIN, TOP, LMT, PE, and PU as input neurons (Figures 4-6). ANN model supported deep learning in all the output neuron modes through its 2-hidden layer deep structure [67,68]. The researcher applied the activation function of the sigmoid function to hidden neurons, as well as output neurons, keeping the values of input and output neurons between [0, 1]; this allowed the researcher to obtain better performance from the research model [69,70]. He also applied the 10-fold cross-validation method to training

XSL•FO
RenderX

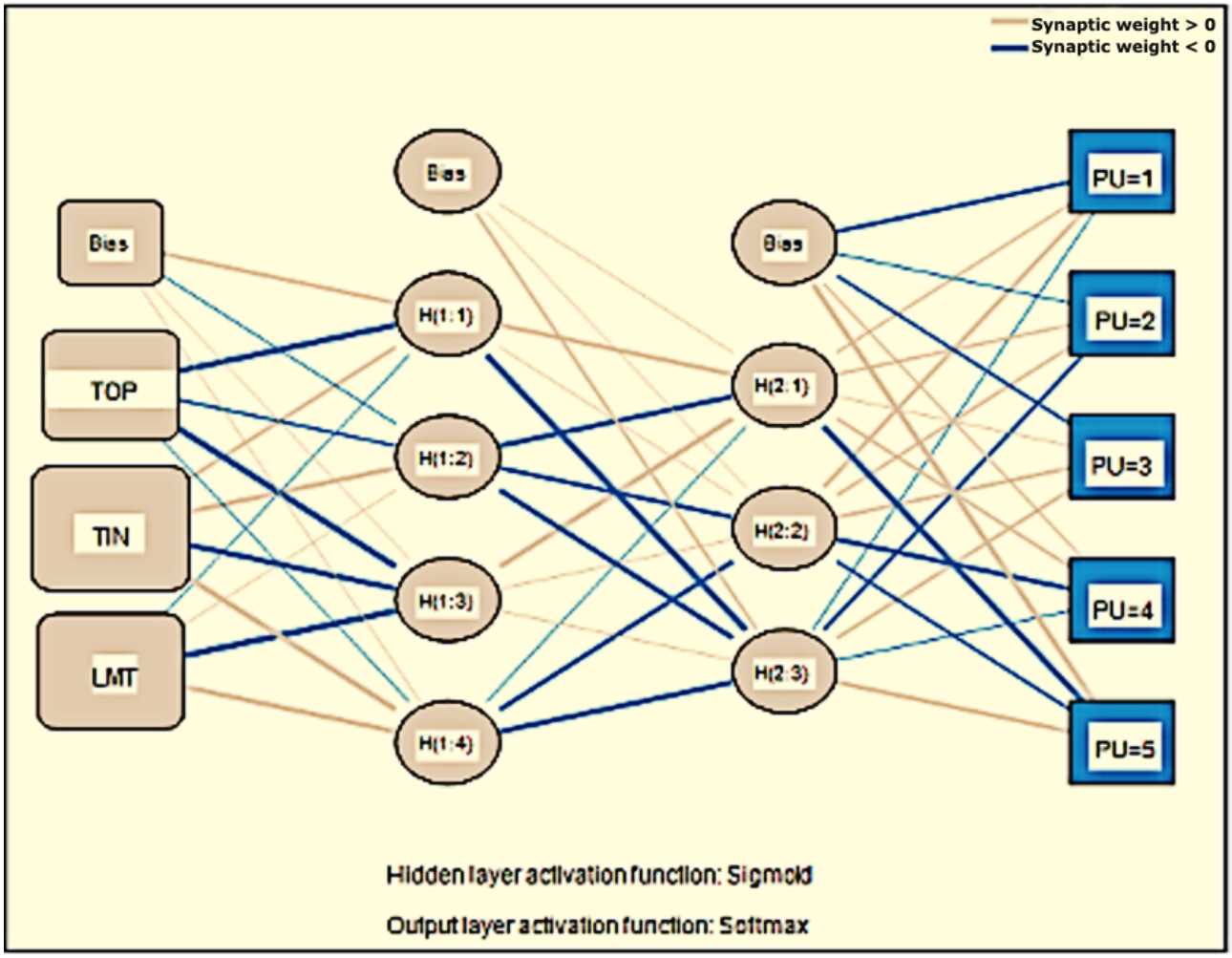
and testing data in the ratio of 80:20 to ensure that there is no overfitting in the ANN model [71]. The researcher evaluated the root mean square of error (RMSE) to test the neural network model for accuracy. The training data showed an RMSE value of 0.1388 for the ANN model while the testing data showed an RMSE value of 0.1439. The training and testing data showed only slight variance in the RMSE values and SD values (ie, 0.0043 and 0.0096). Hence, we can infer that using the ANN model enhances the accuracy of the research model.

Figure 4. ANN model for predicting PE. ANN: artificial neural network; PE: perceived ease of use; TIN: technology innovativeness; TOP: technology optimism.



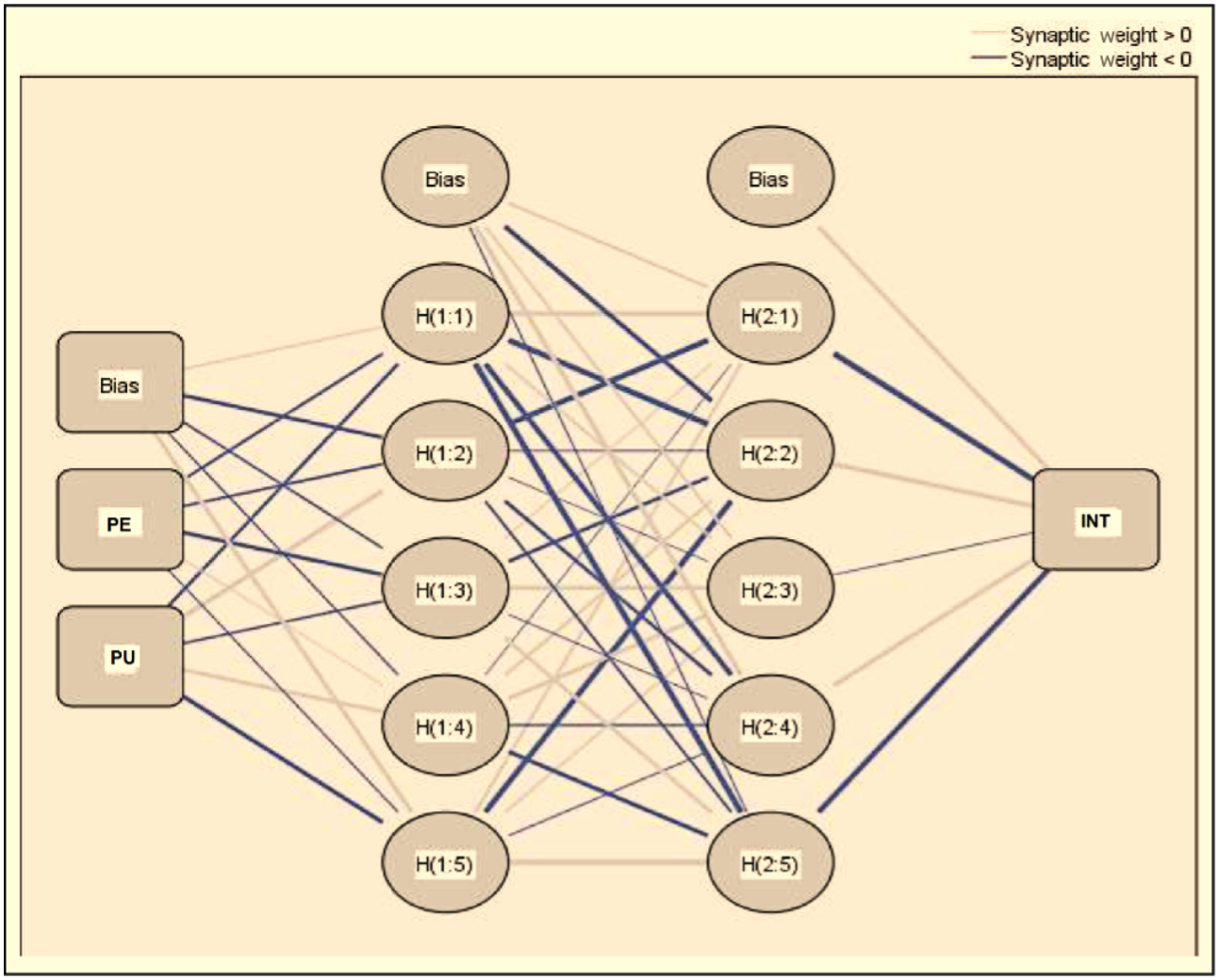
XSL•FO
RenderX

Figure 5. ANN model for predicting PU. ANN: artificial neural network; LMT: learning motivation; PU: perceived usefulness; TIN: technology innovativeness; TOP: technology optimism.



XSL•FO
RenderX

Figure 6. ANN model for predicting INT. ANN: artificial neural network; PE: perceived ease of use; PU: perceived usefulness; INT: intention to use Internet of Things.



Sensitivity Analysis

The researcher estimated the value of normalized importance by comparing each predictor's average value with the maximum mean value depicted among all predictors stated in percentage form. The values of normalized importance and mean importance computed for the predictors involved in ANN modeling are recorded in Table 8. This table also depicts the outcomes of the sensitivity analysis which identifies the predictor of PU to have the most significant impact on behavioral intention; the second most significant impact was imposed by TOP while the least impact was imposed by PE. ANN application was additionally assessed for its accuracy and performance by computing the goodness-of-fit for authentication and validation of the application. The ANN application uses goodness-of-fit just as PLS-SEM analysis uses R^2 [72,73]. The predictive powers of both the applications are compared where ANN analysis outperforms the other one with a predictive power of (R^2=89.7%) against the PLS-SEM predictive power of (R^2=86.3%). Hence, ANN analysis can explain endogenous constructs more effectively than PLS-SEM. Moreover, ANN analysis is based on deep learning and has the potential to better identify nonlinear associations among constructs which results in variances in the predictive powers of the ANN and PLS-SEM methods.

XSL•FO
RenderX

Table . Independent variable importance.

	Importance	Normalized importance (%)
TIN[a]	0.335	83.1
TOP[b]	0.463	94.8
LMT[c]	0.361	77.5
PE[d]	0.118	19.2
PU[e]	0.539	100

[a]TIN: technology innovativeness.

[b]TOP: technology optimism.

[c]LMT: learning motivation.

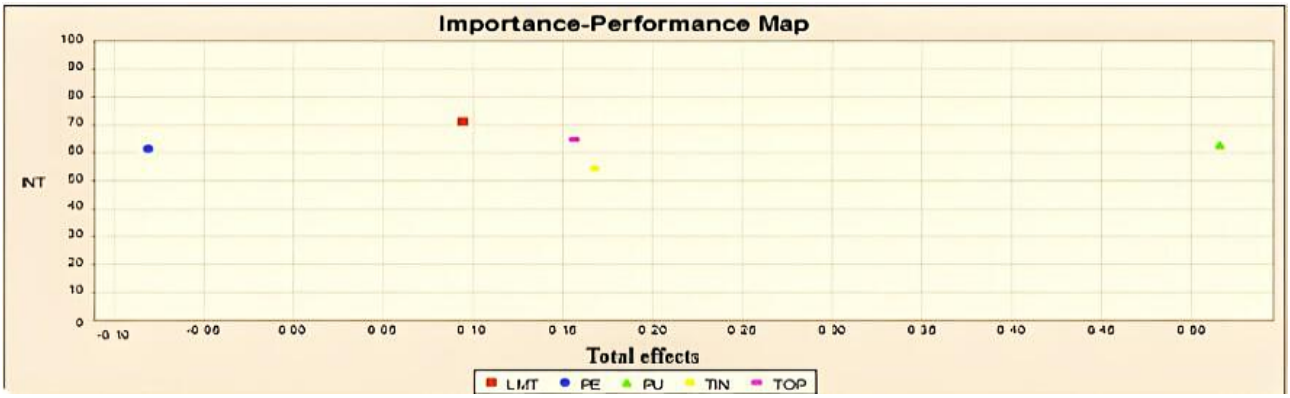
[d]PE: perceived ease of use.

[e]PU: perceived usefulness.

Importance-Performance Map Analysis

This study has used importance-performance map analysis (IPMA) as an advanced approach in PLS-SEM, which used behavioral intention as the main variable. As suggested by Ringle and Sarstedt [74], a better interpretation of the results of PLS-SEM is possible by using IPMA. A substitute way to only test the path coefficients (ie, importance measure), the average value of the latent constructs, and their indicators (ie, performance measure) are also included in the IPMA [74]. According to IPMA, the total effects reflect the predecessor factors' importance in developing the target factor (ie, behavioral intention), while the average of latent constructs' values is a reflection of their performance. The IPMA findings are reported in Figure 7. The estimation of the importance and performance of the 5 factors (ie, TIN, TOP, LMT, PE, and PU) has been shown in this table. According to the findings, the PU has been reported to have the largest values in terms of both importance and performance measures. Furthermore, it can be clearly seen that TIN has the second largest values in terms of both importance and performance measures. The third largest value was reported in the case of TOP in terms of the importance measure; however, it has the smallest value on the performance measure. Relatively, the opposite scenario was reported in the case of PE, as it had the lowest value on the importance measure.

Figure 7. Importance-performance map analysis results. INT: intention to use Internet of Things; LMT: learning motivation; PE: perceived ease of use; PU: perceived usefulness; TIN: technology innovativeness; TOP: technology optimism.



Discussion

Principal Findings

The main findings of this study indicate that the intention to adopt IoT technologies is positively influenced by both intrinsic motivational factors and TAM constructs. Specifically, TOP, innovation, and LMT play crucial roles at the initial level of influencing IoT adoption. At a secondary level, the perceived ease of use and PU, core elements of the TAM, directly enhance the INT technologies. Further analysis using ANNs and IPMA highlighted PU as a particularly significant predictor of IoT use intentions.

Based on these results, it is evident that technology features are critical in shaping users' PE and PU toward IoT applications and tools. The positive influence of TOP, innovation, and LMT is reflected across all model variables, suggesting that well-designed technology features can significantly boost IoT acceptance. The TAM framework effectively provides a user-friendly experience that meets users' value expectations and fosters positive emotional responses, which in turn, positively impacts their INT. Additionally, the significant positive impact of technology features on TOP boosts users' trust in the system's quality, further influencing their willingness to engage with IoT systems. The social aspect of technology use, where familiarity influences adoption, also plays a pivotal role in the INT features. Technology innovation contributes

XSL•FO
RenderX

significantly to perceived ease and usefulness, enhancing trust and satisfaction during the user experience, thus fostering a conducive environment for IoT adoption.

The confirmation of hypotheses H1-H5 supports the proposed conceptual model and the proposed hypotheses [75-77]. Previous research has also produced results that are consistent with these findings. However, other studies have shown that the lack of sufficient security and privacy are major challenges that may hinder the deployment of IoT in education [63,64], and to reduce these obstacles, future efforts to implement IoT in education must consider these factors. Although IoT has not been widely adopted in resource-limited countries, scholars need to examine the factors affecting its adoption to enable effective deployment. Therefore, scholars need to investigate the use and adoption of these technologies in other domains.

Managerial Implications

The adoption of IoT in educational settings demands that administrators and educational leaders ensure their faculty is proficient in using IoT technologies effectively in the classroom. This includes competencies in handling relevant technology tools, understanding pedagogical integration, and applying these technologies within various teaching scenarios. To support this, institutions should offer targeted professional development that focuses on both the technical and educational aspects of IoT. Additionally, management should consider the infrastructure upgrades necessary to support IoT technologies, such as improved wireless networks and enhanced security measures to protect student data [78,79].

Practical Implications

The practical applications of this study highlight the necessity for educational curricula to evolve alongside technological advancements. Institutions offering programs in computer science and engineering should integrate IoT courses to prepare students for the demands of the workforce, which increasingly relies on IoT technologies. Moreover, schools should align their IoT strategies with real-world applications, providing students with hands-on opportunities to work with IoT in context. This could include partnerships with IoT companies or practical projects that allow students to solve real problems using IoT solutions [80,81].

Theoretical Implications

This study contributes to the academic understanding of technology adoption by confirming the significant role of intrinsic motivational factors and TAM constructs in the adoption of IoT technologies. By highlighting the dual influence of personal motivation and perceived technological attributes, this research extends existing models of technology acceptance. Furthermore, the use of advanced analytical methods such as ANN and IPMA provides a deeper insight into the nonlinear relationships among the constructs, offering a nuanced perspective that can inform future research in technology adoption theories. This could encourage scholars to explore how different educational contexts or cultural backgrounds influence the adoption and effective use of emerging technologies like IoT [82,83].

Conclusions

IoT technology has fundamentally altered the tech and business industries, laying the groundwork for the creation of intelligent societies and advancing social and economic development. The IoT has advanced quickly and significantly. The research team developed the PE and PU frameworks for Arab customers using path estimation and modeling of structural equations depending on their responses to analyze individuals' acceptance of IoT. In addition to additional factors including technology exuberance, advancements in technology, and academic drive, the research additionally looked at the effects of TAM components on these factors. These factors and TAM components were found to be directly related to the study. According to the research, the incentive to learn, technological advancement, and technological positiveness all had a substantial influence on PE and PU. While earlier research has demonstrated that TAM components affect the motivation to use IoT, this research has investigated how other external factors, including social factors like LMT, can affect these constructs. The study used a conceptual model to examine users' attitudes toward IoT adoption, with 2 levels of analysis. The first level focused on social attitudes, particularly the impact of the incentive to learn as a standalone factor. The second stage looked at how personal traits influenced technological positivism and inventiveness, influencing customer needs and IoT interactions. Subsequent research could look at how individual characteristics affect incentives to learn and look at extra technological aspects that affect IoT uptake. Subsequent research could improve the assessment and give more insight into the value of IoT by including mediating factors between personal characteristics and technological attributes. The theoretical framework could also be used with cutting-edge technology like metaverse systems and artificial intelligence.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

K Alhumaid wrote the manuscript and conducted the formal analysis. SS validated the results and reviewed and edited the manuscript. MK developed the software and created the visualizations. K Ayoubi contributed to the investigation and validation. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

XSL•FO
RenderX

References

1. Ali M, Bilal HSM, Razzaq MA, et al. IoTFLiP: IoT-based flipped learning platform for medical education. Digital Commun Netw 2017 Aug;3(3):188-194. [doi: 10.1016/j.dcan.2017.03.002]
2. Almaiah M, Alfaisal R, Salloum S, et al. Integrating teachers' TPACK levels and students' learning motivation, technology innovativeness, and optimism in an IoT acceptance model. Electronics (Basel) 2022;11(19):3197. [doi: 10.3390/electronics11193197]
3. Al-Adwan AS, Al-Debei MM. The determinants of Gen Z's metaverse adoption decisions in higher education: integrating UTAUT2 with personal innovativeness in IT. Educ Inf Technol 2024 Apr;29(6):7413-7445. [doi: 10.1007/s10639-023-12080-1]
4. Banica L, Burtescu E, Enescu F. The impact of internet-of-things in higher education. Sci Bull Sci 2017;16(1):53-59 [FREE Full text]
5. Gul S, Asif MA, Ahmad S, et al. A survey on role of internet of things in education. Int J Comput Sci Netw Secur 2017;17(5):159-165 [FREE Full text]
6. Wakefield KJ. How the Internet of Things is transforming manufacturing. Forbes. 2014 Jul 1. URL: https://www.forbes.com/sites/ptc/2014/07/01/how-the-internet-of-things-is-transforming-manufacturing/ [accessed 2025-03-26]
7. Aldowah H, Ul Rehman S, Ghazal S, Naufal Umar I. Internet of Things in higher education: a study on future learning. J Phys: Conf Ser 2017 Sep 21;892(1):012017. [doi: 10.1088/1742-6596/892/1/012017]
8. Ning H, Hu S. Technology classification, industry, and education for future Internet of Things. Int J Commun 2012 Sep;25(9):1230-1241. [doi: 10.1002/dac.2373]
9. Kurdi BA, Alshurideh A, Nuseir M, Aburayya A, Salloum SA. The effects of subjective norm on the intention to use social media networks: an exploratory study using PLS-SEM and machine learning approach. In: Advanced Machine Learning Technologies and Applications (AMLTA 2021) 2021:581-592. [doi: 10.1007/978-3-030-69717-4_55]
10. Almarzouqi A, Aburayya A, Salloum SA. Determinants predicting the electronic medical record adoption in healthcare: a SEM-artificial neural network approach. PLoS One 2022;17(8):e0272735. [doi: 10.1371/journal.pone.0272735] [Medline: 35972979]
11. Almarzouqi A, Aburayya A, Salloum SA. Determinants of intention to use medical smartwatch-based dual-stage SEM-ANN analysis. Inform Med Unlocked 2022;28:100859. [doi: 10.1016/j.imu.2022.100859]
12. Alhumaid K, Alnazzawi N, Akour I, Khasoneh OA, Alfaisal R, Salloum S. An integrated model for the usage and acceptance of stickers in WhatsApp through SEM-ANN approach. Int J Data Netw Sci 2022;6(4):1261-1272. [doi: 10.5267/j.ijdns.2022.6.008]
13. Malhotra S, Kumar A, Dutta R. Effect of integrating IoT courses at the freshman level on learning attitude and behaviour in the classroom. Educ Inf Technol 2021 May;26(3):2607-2621. [doi: 10.1007/s10639-020-10376-0]
14. Lai YH, Chen SY, Lai CF, Chang YC, Su YS. Study on enhancing AIoT computational thinking skills by plot image-based VR. Interact Learn Environ 2021 Apr 3;29(3):482-495. [doi: 10.1080/10494820.2019.1580750]
15. Mavroudi A, Divitini M, Gianni F, Mora S, Kvittem DR. Designing IoT applications in lower secondary schools. Presented at: 2018 IEEE Global Engineering Education Conference (EDUCON); Apr 17-20, 2018; Santa Cruz de Tenerife, Spain.
16. Yamamoto T, Horita T. Analysis of teacher training students' understanding on learning objectives in elementary school programming classes using IoT materials. Presented at: Society for Information Technology & Teacher Education International Conference; Apr 6-10, 2020; New Orleans, LA, United States.
17. Akbar MA, Rashid MM, Embong AH. Technology based learning system in Internet of Things (IoT) education. Presented at: 2018 7th International Conference on Computer and Communication Engineering (ICCCE); Sep 19-20, 2018; Kuala Lumpur. [doi: 10.1109/ICCCE.2018.8539334]
18. Lee S, Cho K. A study on the structural relationship between IoT usage and life satisfaction among university students. J Internet Things Converg 2021;7(2):55-63 [FREE Full text]
19. Glaroudis D, Iossifides A, Spyropoulou N, Zaharakis ID. Investigating secondary students' stance on IoT driven educational activities. Presented at: European Conference on Ambient Intelligence; Nov 12-14, 2018; Larnaca, Cyprus.
20. Yildiz Durak H. Modeling of relations between K-12 teachers' TPACK levels and their technology integration self-efficacy, technology literacy levels, attitudes toward technology and usage objectives of social networks. Interact Learn Environ 2021 Oct 3;29(7):1136-1162. [doi: 10.1080/10494820.2019.1619591]
21. Yang J, Wang Q, Wang J, Huang M, Ma Y. A study of K-12 teachers' TPACK on the technology acceptance of E-schoolbag. Interact Learn Environ 2021 Oct 3;29(7):1062-1075. [doi: 10.1080/10494820.2019.1627560]
22. Pastor-Vargas R, Tobarra L, Robles-Gómez A, Martin S, Hernández R, Cano J. A WoT platform for supporting full-cycle IoT solutions from edge to cloud infrastructures: a practical case. Sensors (Basel) 2020 Jul 5;20(13):3770. [doi: 10.3390/s20133770] [Medline: 32635632]
23. Akbari M, Rezvani A, Shahriari E, Zúñiga M, Pouladian H. Acceptance of 5 G technology: mediation role of trust and concentration. J Eng Technol Manag 2020 Jul;57:101585. [doi: 10.1016/j.jengtecman.2020.101585]
24. Seiler L, Kuhnel M, Honal A, Ifenthaler D. Mobile learning analytics: potenziale für lernen und lehren am beispiel hochschule. In: Handbuch Mobile Learning: Springer; 2018:585-608.

XSL•FO
RenderX

25. Majeed A, Ali M. How internet-of-things (iot) making the university campuses smart? QA higher education (QAHE) perspective. Presented at: 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC); Jan 8-10, 2018; Las Vegas, NV, United States. [doi: 10.1109/CCWC.2018.8301774]
26. Singh UG, Msibi A. Students' perceptions towards IoTs adoption: a case study. J Phys: Conf Ser 2021 Jan 1;1714(1):012044. [doi: 10.1088/1742-6596/1714/1/012044]
27. Liu Y, Gan Y, Song Y, Liu J. What influences the perceived trust of a voice-enabled smart home system: an empirical study. Sensors (Basel) 2021 Mar 13;21(6):2037. [doi: 10.3390/s21062037] [Medline: 33805702]
28. Hampton NZ, Mason E. Learning disabilities, gender, sources of efficacy, self-efficacy beliefs, and academic achievement in high school students. J Sch Psychol 2003 Mar;41(2):101-112. [doi: 10.1016/S0022-4405(03)00028-1]
29. McCown RR, Driscoll M, Roop P. Facilitating student motivation. In: Educational Psychology: A Learning-Centered Approach to Classroom Practice: Allyn and Bacon; 1996:278-309.
30. Stipek D, Feiler R, Daniels D, Milburn S. Effects of different instructional approaches on young children's achievement and motivation. Child Dev 1995 Feb;66(1):209. [doi: 10.2307/1131201]
31. Sırakaya M, Sırakaya DA. Augmented reality in STEM education: a systematic review. Interact Learn Environ 2020:1-14 [FREE Full text]
32. Alsyouf A, Lutfi A, Alsubahi N, et al. The use of a technology acceptance model (TAM) to predict patients' usage of a personal health record system: the role of security, privacy, and usability. Int J Environ Res Public Health 2023 Jan 11;20(2):1347. [doi: 10.3390/ijerph20021347] [Medline: 36674105]
33. Esteban-Millat I, Martínez-López FJ, Pujol-Jover M, Gázquez-Abad JC, Alegret A. An extension of the technology acceptance model for online learning environments. Interact Learn Environ 2018 Oct 3;26(7):895-910. [doi: 10.1080/10494820.2017.1421560]
34. Huang HM, Liaw SS, Lai CM. Exploring learner acceptance of the use of virtual reality in medical education: a case study of desktop and projection-based display systems. Interact Learn Environ 2016 Jan 2;24(1):3-19. [doi: 10.1080/10494820.2013.817436]
35. Parasuraman A. Technology Readiness Index (TRI) a multiple-item scale to measure readiness to embrace new technologies. J Serv Res 2000;2(4):307-320. [doi: 10.1177/109467050024001]
36. Al-Adwan AS. Investigating the drivers and barriers to MOOCs adoption: the perspective of TAM. Educ Inf Technol 2020 Nov;25(6):5771-5795. [doi: 10.1007/s10639-020-10250-z]
37. Alsyouf A, Lutfi A, Al-Bsheish M, et al. Exposure detection applications acceptance: the case of COVID-19. Int J Environ Res Public Health 2022 Jun 14;19(12):7307. [doi: 10.3390/ijerph19127307] [Medline: 35742560]
38. Chin WW. The partial least squares approach to structural equation modeling. In: Modern Methods for Business Research: Psychology Press; 1998, Vol. 295:295-336.
39. Alhumaid K, Rizk RY, Snášel V, Abdel-Kader RF. Predicting the intention to use audi and video teaching styles: an empirical study with PLS-SEM and machine learning models. Presented at: International Conference on Advanced Machine Learning Technologies and Applications; May 5-7, 2022; Cairo, Egypt. [doi: 10.1007/978-3-031-03918-8_23]
40. Ngafeeson MN, Sun J. E-Book acceptance among undergraduate students: a look at the moderating role of technology innovativeness. Int J Web-Based Learn Teach Technol 2015;10(3):36-51. [doi: 10.4018/ijwltt.2015070103]
41. Li CH, Wu JJ. The structural relationships between optimism and innovative behavior: understanding potential antecedents and mediating effects. Creat Res J 2011 Apr;23(2):119-128. [doi: 10.1080/10400419.2011.571184]
42. Ziyae B, Mobaraki MH, Saeediyoun M. The effects of psychological capital on organizational performance with an emphasis on the mediatory role of innovation in information technology: a case study of the branches of agriculture bank in Tehran. Innov Manag J 2015;3(4):77-100 [FREE Full text]
43. Issham I, Siti NA, Nizuwan A. Accessing innovativeness of distance learners toward their readiness in embracing technology. African J Bus Manag 2011;5(33):12768-12776. [doi: 10.5897/AJBM11.824]
44. Al-Adwan AS, Li N, Al-Adwan A, Abbasi GA, Albelbisi NA, Habibi A. Extending the Technology Acceptance Model (TAM) to predict university students' intentions to use metaverse-based learning platforms. Educ Inf Technol (Dordr) 2023 Apr 28:1-33. [doi: 10.1007/s10639-023-11816-3] [Medline: 37361794]
45. Karahanna E, Straub DW, Chervany NL. Information technology adoption across time: a cross-sectional comparison of pre-adoption and post-adoption beliefs. MIS Q 1999 Jun;23(2):183. [doi: 10.2307/249751]
46. Kabakci Yurdakul I, Odabasi HF, Kilicer K, Coklar AN, Birinci G, Kurt AA. The development, validity and reliability of TPACK-deep: a technological pedagogical content knowledge scale. Comput Educ 2012 Apr;58(3):964-977. [doi: 10.1016/j.compedu.2011.10.012]
47. Schmidt DA, Baran E, Thompson AD, Mishra P, Koehler MJ, Shin TS. Technological pedagogical content knowledge (TPACK) the development and validation of an assessment instrument for preservice teachers. J Res Technol Educ 2009;42(2):123-149. [doi: 10.1080/15391523.2009.10782544]
48. Krejcie RV, Morgan DW. Determining sample size for research activities. Educ Psychol Meas 1970 Sep;30(3):607-610. [doi: 10.1177/001316447003000308]
49. Chung N, Han H, Joun Y. Tourists' intention to visit a destination: the role of augmented reality (AR) application for a heritage site. Comput Human Behav 2015 Sep;50:588-599. [doi: 10.1016/j.chb.2015.02.068]

XSL•FO
RenderX

50. Huang HMM, Rauch U, Liaw SSS. Investigating learners' attitudes toward virtual reality learning environments: based on a constructivist approach. Comput Educ 2010 Nov;55(3):1171-1182. [doi: 10.1016/j.compedu.2010.05.014]
51. Huang HM, Liaw SS. An analysis of learners' intentions toward virtual reality learning based on constructivist and technology acceptance approaches. Int Rev Res Open Distrib Learn 2018;19(1):1. [doi: 10.19173/irrodl.v19i1.2503]
52. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
53. Fishbein M, Ajzen I. Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research: Addison-Wesley; 1977.
54. Ringle CM, Wende S, Becker JM. SmartPLS 3. : SmartPLS; 2015.
55. Mouzaek E, Marzouqi A, Alaali N, Salloum S, Aburayya A, Suson O. An empirical investigation of the impact of service quality dimensions on guests satisfaction: a case study of Dubai hotels. J Contemp Issues Bus Gov 2021 Apr 2;27(3):1187. [doi: 10.47750/cibg.2021.27.03.160]
56. Salloum SA, Mhamdi C, Kurdi BA, Shaalan K. Factors affecting the adoption and meaningful use of social media: a structural equation modeling approach. Int J Inf Technol Lang Stud 2018;2(3):96-109.
57. Hair J, Hollingsworth CL, Randolph AB, Chong AYL. An updated and expanded assessment of PLS-SEM in information systems research. Ind Manag Data Syst 2017 Apr 10;117(3):442-458. [doi: 10.1108/IMDS-04-2016-0130]
58. Salloum SA, Al-Emran M, Habes M, Alghizzawi M, Ghani MA, Shaalan K. What impacts the acceptance of E-learning through social media? An empirical study. Recent Adv Technol Accept Model Theor 2021:419-431. [doi: 10.1007/978-3-030-64987-6_24]
59. MacKenzie SB, Podsakoff PM, Jarvis CB. The problem of measurement model misspecification in behavioral and organizational research and some recommended solutions. J Appl Psychol 2005 Jul;90(4):710-730. [doi: 10.1037/0021-9010.90.4.710] [Medline: 16060788]
60. Goodhue DL, Lewis W, Thompson R. Does PLS have advantages for small sample size or non-normal data? MIS Q 2012;36(3):981. [doi: 10.2307/41703490]
61. Barclay D, Higgins C, Thompson R. The partial least squares (PLS) approach to casual modeling: personal computer adoption and use as an illustration. Technol Stud 1995 [FREE Full text]
62. Nunnally JC, Bernstein IH. Psychometric Theory: McGraw Hill; 1994.
63. Kline RB. Principles and Practice of Structural Equation Modeling: Guilford Publications; 2015.
64. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. J Mark Res 1981 Feb;18(1):39-50. [doi: 10.1177/002224378101800104]
65. Henseler J, Ringle CM, Sarstedt M. A new criterion for assessing discriminant validity in variance-based structural equation modeling. J Acad Mark Sci 2015 Jan;43(1):115-135. [doi: 10.1007/s11747-014-0403-8]
66. Almaiah MA, Alfaisal R, Salloum SA, et al. Examining the impact of artificial intelligence and social and computer anxiety in E-learning settings: students' perceptions at the university level. Electronics (Basel) 2022;11(22):3662. [doi: 10.3390/electronics11223662]
67. Lee VH, Hew JJ, Leong LY, Tan GWH, Ooi KB. Wearable payment: a deep learning-based dual-stage SEM-ANN analysis. Expert Syst Appl 2020 Nov;157:113477. [doi: 10.1016/j.eswa.2020.113477]
68. Elareshi M, Habes M, Youssef E, Salloum SA, Alfaisal R, Ziani A. SEM-ANN-based approach to understanding students' academic-performance adoption of YouTube for learning during Covid. Heliyon 2022 Apr;8(4):e09236. [doi: 10.1016/j.heliyon.2022.e09236] [Medline: 35434400]
69. Liébana-Cabanillas F, Marinkovic V, de Luna IR, Kalinic Z. Predicting the determinants of mobile payment acceptance: a hybrid SEM-neural network approach. Technol Forecast Soc Change 2018 Apr;129:117-130. [doi: 10.1016/j.techfore.2017.12.015]
70. Aburayya A, Salloum SA, Alderbashi KY, et al. SEM-machine learning-based model for perusing the adoption of metaverse in higher education in UAE. Int J Data Netw Sci 2023;7(2):667-676. [doi: 10.5267/j.ijdns.2023.3.005]
71. Sharma SK, Sharma M. Examining the role of trust and quality dimensions in the actual usage of mobile banking services: an empirical investigation. Int J Inf Manage 2019 Feb;44:65-75. [doi: 10.1016/j.ijinfomgt.2018.09.013]
72. Leong LY, Hew TS, Ooi KB, Lee VH, Hew JJ. A hybrid SEM-neural network analysis of social media addiction. Expert Syst Appl 2019 Nov;133:296-316. [doi: 10.1016/j.eswa.2019.05.024]
73. Salloum SA, AlAhbabi NMN, Habes M, Aburayya A, Akour I. Predicting the intention to use social media sites: a Hybrid SEM - machine learning approach. Adv Intell Syst Comput 2021;1339:324-334. [doi: 10.1007/978-3-030-69717-4_32]
74. Ringle CM, Sarstedt M. Gain more insight from your PLS-SEM results: the importance-performance map analysis. Ind Manage Data Syst 2016. [doi: 10.1108/IMDS-10-2015-0449]
75. Zaidi MFA. The IoT readiness of SMEs in Malaysia: are they worthwhile for investigation? Presented at: International Conference on International Business, Marketing and Humanities (ICIBMAH 2017); Aug 26-27, 2017; Alor Setar, Malaysia.
76. Cadenas GA, Cantú EA, Lynn N, Spence T, Ruth A. A programmatic intervention to promote entrepreneurial self-efficacy, critical behavior, and technology readiness among underrepresented college students. J Vocat Behav 2020 Feb;116:103350. [doi: 10.1016/j.jvb.2019.103350]

XSL•FO
RenderX

77. Wang Y, Liu C, Tu YF. Factors affecting the adoption of AI-based applications in higher education. Educ Technol Soc 2021;24(3):116-129 [FREE Full text]
78. Chweya R, Ajibade SSM, Buba AK, et al. IoT and big data technologies: opportunities and challenges for higher learning. Int J Recent Technol Eng (IJRTE) 2020;9(2):909-913. [doi: 10.35940/ijrte.B3943.079220]
79. Rajamohan K, Rangasamy S, Manoj CN, Mary L, Sabu R, Jose SS. Internet of things (IoT) as a game changer to the education sector. In: Advanced IoT Technologies and Applications in the Industry 40 Digital Economy: CRC Press; 2024:251-274. [doi: 10.1201/9781003434269-14]
80. Abraham S. Using internet of things (IoT) as a platform to enhance interest in electrical and computer engineering. Presented at: 2016 ASEE Annual Conference & Exposition; Jun 26-29, 2016; New Orleans, Louisiana. [doi: 10.18260/p.27149]
81. Almodaires A. Technology-supported reflection: towards bridging the gap between theory and practice in teacher education. : University of Twente; 2009 URL: https://research.utwente.nl/en/publications/technology-supported-reflection-towards-bridging-the-gap-between--2 [accessed 2025-03-26]
82. Marikyan M, Papagiannidis P. Unified Theory of Acceptance and Use of Technology: a review. In: TheoryHub Book 2021.
83. Syed H, Malik AN. Comparative study of effect of culture on technology adoption in Pakistan and USA. Bus Manag Rev 2014;5(1):42 [FREE Full text]

Abbreviations

ANN: artificial neural network
AVE: average-variance extracted
INT: intention to use Internet of Things
IoT: Internet of Things
IPMA: importance-performance map analysis
LMT: learning motivation
PE: perceived ease of use
PLS-SEM: partial least squares–structural equation modeling
PU: perceived usefulness
RMSE: root mean square of error
SEM: structural equation modeling
TAM: technology acceptance model
TIN: technology innovativeness
TOP: technology optimism

Edited by A Kushniruk; submitted 14.03.24; peer-reviewed by A Lutfi, AS Al-Adwan, A Arbabisarjou, G Farid, R Alfaisal; revised version received 01.09.24; accepted 15.01.25; published 10.04.25.

Please cite as:
Alhumaid K, Ayoubi K, Khalifa M, Salloum S
Factors Determining Acceptance of Internet of Things in Medical Education: Mixed Methods Study
JMIR Hum Factors 2025;12:e58377
URL: https://humanfactors.jmir.org/2025/1/e58377
doi:10.2196/58377

© Khadija Alhumaid, Kevin Ayoubi, Maha Khalifa, Said Salloum. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 10.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Impact of Trust and the Role of the Opt-Out Mechanism in Willingness to Share Health Data via Electronic Health Records in Germany: Telephone Survey Study

Felix Wilke, Prof Dr

Department of Social Work, University of Applied Sciences Jena, Carl-Zeiss-Promenade 2, Jena, Germany

Corresponding Author:
Felix Wilke, Prof Dr
Department of Social Work, University of Applied Sciences Jena, Carl-Zeiss-Promenade 2, Jena, Germany

Abstract

Background: Electronic health records (EHRs) offer a valuable resource for research and health care improvement. However, public acceptance of sharing personal health data is critical to the success of such initiatives. In Germany, automatic data sharing via EHRs will be implemented unless people opt out.

Objective: This study aims to assess the willingness of the German population to share health data via EHRs and to explore the role of trust in influencing these attitudes.

Methods: A computer-assisted telephone interview study was conducted in December 2023, with 1004 respondents aged 18 years and older, representative of the German population. Descriptive statistics and multivariate linear regression models were used to analyze the data.

Results: The survey shows that 43.4% (n=432) of respondents would be willing to share their health data via EHR, and a significant 34% (n=338) remain undecided. While the population is open to adoption of the EHR for personal health issues (n=483, 53% show interest in using it), the opt-out model for data sharing is viewed critically, with 44.7% (n=438) of respondents rejecting it. Socioeconomic status significantly influences the willingness to share data, with higher income, education, and digital literacy being associated with greater openness to data sharing. However, trust emerged as the most significant factor. Additionally, experiences with digital technologies increase the willingness to share personal health data.

Conclusions: The German population shows general openness toward EHRs and data sharing. Trust plays a critical role in promoting willingness to share health data. The findings highlight challenges in Germany's transition to an opt-out system.

(JMIR Hum Factors 2025;12:e65718) doi:10.2196/65718

KEYWORDS

data sharing; health; citizens; electronic health record; trust; digitalization; opt-out

Introduction

Legislative Context and Research Questions

Health data stored in electronic health records (EHRs) are considered an extremely valuable source for advancing medical science and health care. EHRs have the potential to provide high-quality longitudinal data from a broad spectrum of the population. It is therefore not surprising that many countries are implementing data-sharing options to make use of these data. In Germany, the EHR ("elektronische Patientenakte") was introduced on January 1, 2021, and is available to all legally insured persons through their health insurance company. It can be used to store and manage medical diagnoses and treatment data. The use of the EHR also allows control over who can view and access data. However, by 2024, only 1% - 2% of the eligible population have an EHR. To increase the prevalence and benefits of the EHR, the German Bundestag passed the Digital Act ("Digital-Gesetz") and the Health Data Use Act ("Gesundheitsdatennutzungsgesetz") on December 14, 2023. This means that starting in 2025, for all people with statutory health insurance in Germany, an EHR will be automatically created unless they opt out. The Health Data Use Act further allows for the use of pseudonymized health data for research purposes by public and private institutions by amending §303e SGB V. That means data stored in the EHR will be available for research purposes including commercial purposes unless people opt out. The data will be made available through the German Health Research Data Center ("Forschungsdatenzentrum"). The legislative changes mark a paradigm shift in the way health data are handled: the creation of electronic patient records and the sharing of health data will be automatic unless people actively opt out. The German federal minister of health, Karl Lauterbach, stated that the paradigm shift will be a "gamechanger" for the development of health [1]. The new legal framework marks a step toward digitization

XSL•FO
RenderX

that many experts have been calling for for a long time [2]. At the same time, the extensive and passive sharing of health data has been criticized by various stakeholders.

This paradigm shift raises some serious questions about the social implications. There are a number of challenges to the "data is gold" perspective [3]. Some social scientists criticize that digitization in health bears potential for "panoptic surveillance" [4]. Other critics [5] have highlighted data security issues that can harm citizens in many ways. Furthermore, the risks, but also the opportunities to benefit from digital health, are unevenly distributed across society. There is ample empirical evidence of a so-called "digital divide," meaning that people with lower socioeconomic status have lower chances of using digital (health) services and face higher risks [4,6,7]. The digitization of health therefore needs to be critically monitored. Citizens should be aware of the potential benefits as well as the potential dangers of sharing personal health data. This becomes even more urgent with the implementation of an opt-out mechanism. Opt-out is an instrument of nudging [8]. It aims to increase the use of EHRs and the acceptance of data sharing by "choice architecture" [9]. Citizens retain their freedom of choice but are nudged toward using the EHR. However, an opt-out mechanism relies on informed citizens to take deliberative action, especially when data sharing can be harmful. The policy of choice architecture and the use of EHR data for health research make clear that health data sharing cannot be reduced to an individual action but must be analyzed in its social context. It raises questions of legitimacy, social inequality, and the diffusion of knowledge among citizens.

This paper focuses on public opinion on EHR data sharing in Germany. It answers three research questions: (1) Is there sufficient willingness in society to share health data that justifies the introduction of an opt-out scheme? (2) Does the willingness to share data differs with socioeconomic status? (3) How can individual willingness to share EHR data be explained?

State of the Art

This section summarizes the research on data sharing and evidence on EHR adoption. International research finds widespread support for the idea of EHRs among citizens and institutions in the health care sector [10-12]. However, "evidence indicates that consumers who have been offered access to PEHRs have not widely used them and the number of active users has remained low" [12]. For example, a recent study shows that in the United States, the number of active users of the EHR is about 26% (2019), despite its increasing integration in medical practice [13]. In Germany, until 2024, according to data from public health insurance, less than 2% of the insured actively requested an EHR [14]. At the same time, only about half of the population is even aware of the EHR [15]. Several studies have shown that interest and use of the EHR show a digital divide. Younger and more digital literate people show a higher probability of using EHRs, while older and low-income groups have less often access [12,16,17]. Regarding socioeconomic factors, lower educational and employment status, as well as belonging to an ethnic minority, have a negative impact on individuals' desire to access and use EHRs [12,18].

Several studies have examined citizens' motivations for using EHRs. The perceived advantages and disadvantages strongly influence the intention to use and the actual use. Key factors here are concerns about data protection and the fear of being discriminated by health insurers or employers on the one hand, and the benefits of better information and experience with the portals or providers on the other [12,19,20]. Data protection concerns are a key factor because health issues are relevant in many areas of daily life [10,12,19]. This can be seen directly when respondents report fears of stigmatization as a result of digitally disseminated health information from the EHR [21,22]. For people with chronic diseases, the balance of advantages and disadvantages seems to be clearly in favor of the advantages. Several studies have reported higher levels of interest in providing data due to the hope of treatment progress [23,24]. For people with chronic diseases, there are already care programs in place that people can benefit from, such as disease management programs [25].

While the research literature stresses that EHRs are potentially an important tool to transform "passive recipients of treatment to empowered consumers" [12], the empirical findings are rather disappointing. Several international studies have shown that the overall use of EHRs by patients remains low [12]. Furthermore, there is little evidence of empowerment effects [12]. Even people who have an EHR often do not know which data are stored in the health record and how to interpret the information [12,26]. When users experience that they do not understand the EHR or the information stored in it, their willingness to use it decreases significantly [23]. Thus, we can expect a significant gap between interest in EHRs and actual use. This gap is likely to be even more pronounced in Germany, with its high standards for data security and complicated routines for accessing the EHR.

Data sharing is an essential part of EHRs. Many EHR systems allow anonymized health data to be used for scientific research. While this can lead to progress in the health system, there is no clear-cut advantage for an individual sharing its data, since anonymization makes individual feedback regarding new health insights hard. That might be the reason, why in public discourse, data sharing is often labeled as "data donation" [27]. Research regarding data sharing has shown that a considerable part of the population (often exceeding 50%) is willing to share its health data [27-30]. There are considerable differences between different countries. For example, the Scandinavian countries show the highest willingness to share health data, while people in Southern Europe are rather skeptical about it [31]. For Germany, recent research suggests that 8 of 10 citizens are willing to share their health data for scientific purposes [15,32,33]. Notably, this high approval rate also seems to hold true for the planned opt-out regulation [33]. However, these high approval rates have to be interpreted with caution, since limited knowledge about health data sharing is widespread and some studies have methodological limitations. Like many other surveys in this area, the high approval rates rely on web-based surveys that have proven to be of less quality regarding the research results compared to face-to-face or telephone interviews [34]. One reason is that they show a higher tendency to produce biased results [35]. Since there is an empirical correlation

XSL•FO
RenderX

between digital literacy and the willingness to share health data, the results are likely to be too optimistic. Furthermore, some questionnaires tend to highlight the positive impacts of data sharing and thereby inflating the approval rate (The question of the survey conducted in Germany was [15,32]: “Assuming that in the future your personal health data, such as your medical history, examination results, X-ray images, etc. can be stored online in a digital health record. Would you agree that your personal health data is used anonymously and free of charge for the purpose of medical research so that in the future diseases can be better recognized and new treatments developed?” And “At present, the use of patient data for medical research requires the consent of patients. However, it is often difficult or even impossible to obtain such consent. Therefore, there are considerations in Germany to legally allow medical research without consent on encrypted patient data after an independent review of the research project. In return, it should be possible for every citizen to simply object to this so-called ‘data donation’ How would you find such a legal regulation ... for publicly funded research?”).

Empirical studies have revealed that the willingness to share data differs among different groups. The findings are comparable to the overall use of EHR: among those with higher age, educational level, and income, a higher willingness to share data can be observed [30,36-39]. Some studies investigate potential individual causes for the willingness to share health data for scientific purposes. These studies find that concerns about confidentiality breaches and data misuse are important factors that prevent data sharing [28,29]. Unsurprisingly, privacy protection, anonymization, and transparency have proven to be key factors for data sharing [29,40]. Furthermore, people are motivated by potential benefits, including more altruistic ones, like advancements in science and public health [29,40]. Another key factor is the use of health data. While international research has shown relatively high public support for the secondary use of health data, this support decreases significantly when shared data can be used by the private sector [41-43].

Another source of insight comes from studies that focus on the social contexts in which data are shared. They have identified 2 other relevant aspects: digital literacy and trust. Health data sharing relies on digital tools. Therefore, people who are more familiar with digital tools may be less reluctant to share data. In addition, technology companies such as Google and Apple have created many services that make use of data sharing in everyday life. Hence, digitally literate people are more involved in data sharing in general. In line with this argument, empirical studies show a higher propensity to share health data among the more digitally savvy people [36,37,40,44,45]. On a more abstract level, it can be argued that trust plays an essential role in health in general [46] and in data sharing in particular. Since the sharing of data is a socially highly indeterminate constellation for social action, trust is a crucial resource. The risks taken by data donors as well as the potential benefits depend strongly on other actors and can hardly be known in advance. For example, the question of anonymizing health data can only be answered in relation to existing knowledge and techniques. The benefits of data sharing are particularly uncertain because there is no guarantee of medical progress, and even if progress does occur, it is unclear whether data donors themselves benefit. Given these theoretical arguments, it is not surprising that empirical studies show an important influence of trust. They can show that trust in the institutions involved in data sharing is relevant, as is general trust in science [15,29,36,37,40,44,45].

Methods

Ethical Considerations

Ethics approval was obtained from the ethics department of the German Aerospace Center (21/23) before the interviews were conducted. Participants gave verbal informed consent to participate in the study. Anonymized interview data were provided to the author by the survey institute. No compensation was given to the interviewees.

Survey Design and Sample

The analysis is based on a dedicated telephone survey (computer-assisted telephone interview [CATI]) about health data conducted between December 7 and 21, 2023. The sampling was carried out by the survey institute drei.fakt (Erfurt) on behalf of the Ernst Abbe University of Applied Sciences Jena. For the study, a representative randomized stratified random sample was drawn based on randomized landline and cell phone numbers (30/70) [47]. Random landline numbers were called based on regional quotas. The mobile phone numbers were sampled and stratified by telephone provider. A total of 11,451 residential numbers were called. In total, 1004 respondents aged 18 years and older participated in an interview. In contrast to many web-based surveys in this area, CATI allows people without internet access to be able to take part in the survey. To ensure the quality of the sample, it was compared with the population for age, gender, and region during the field period. Statistics from the 2019 German microcensus served as a data basis for comparing the distribution of the sample with the distribution of the characteristics in the population. Compared to population statistics, the unweighted sample underrepresents immigrants and slightly overrepresents people with higher education. This selection bias of telephone interviews is well-known from survey research. Therefore, all analyses use a weighting variable in order to adequately represent the population in terms of demographic and socioeconomic characteristics.

Questionnaire

The questionnaire comprised 32 questions regarding various topics about health data sharing. The questionnaire included information on digital skills, the willingness to share personal health data in general, motives for and against data sharing, questions regarding trust and data protection, as well as attitudes toward the EHR and socioeconomic characteristics. The CATI lasted approximately 20 minutes. Most questions were based on a 5-point Likert scale or binary response options and included multiple response categories. The final questionnaire included feedback from 2 separate pretests. In this pretest, the survey instrument was tested for its practical suitability in the field and potential misunderstandings. The full questionnaire and survey data are available in German via the GESIS repository [47].

XSL•FO
RenderX

Data Analysis

Descriptive statistics and multivariate linear regression models are used for analysis. The dependent variable in the multivariate regression models is the attitude toward sharing health data via EHR. People were asked on a 5-point scale (1=yes, definitely and 5=no, definitely not): "Would you be willing to share your health data from your EHR anonymized for medical research purposes?" To minimize potential misunderstandings about the EHR, the question was preceded by a brief explanation of EHRs ("Since January 1, 2021, all people with statutory health insurance can use the EHR via an app of the insurance provider. The EHR contains medical findings and information from previous treatments."). It is important to note that this question is hypothetical, as only a small fraction of people actually had an electronic personal health record at the time of the fieldwork (see Introduction section). However, in order to measure the willingness to share personal health information and to take into account the new legal framework, the question was phrased in a hypothetical manner. This makes it possible to measure the willingness to share extensive and potentially sensitive personal health information rather than actual planned behavior. Therefore, the question is also analyzed for respondents without an EHR in order to reflect overall opinions about the new legislative approach.

Results

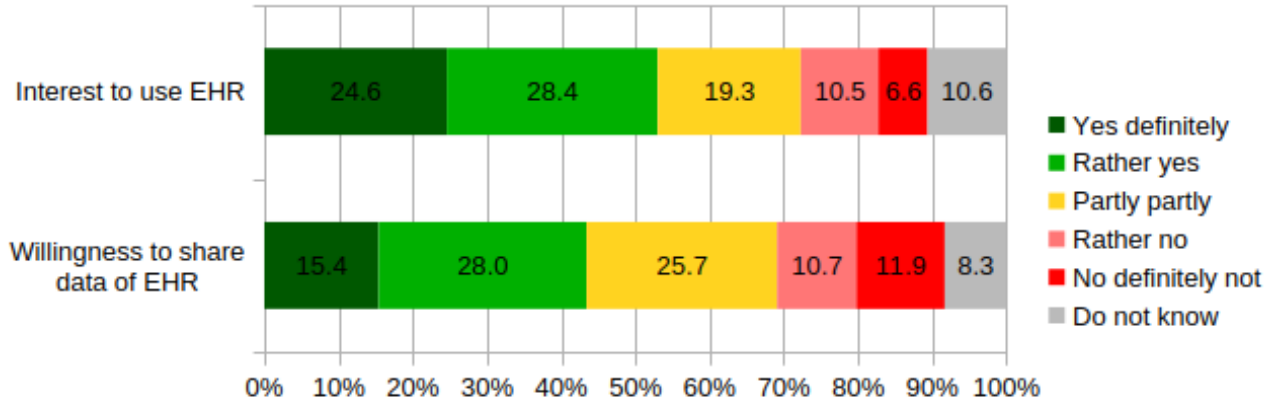
The first step is to analyze the approval rates for the use of the German EHR. In general, the population is very open to the EHR. In total, 53% (n=483) of the population who claim to have no EHR yet show interest in using it. At the same time, 8.9% (n=89) of the sample state that they already have an EHR, in comparison to the 1% reported in official statistics. Therefore, it has to be assumed that some people lack a proper understanding of the EHR (see Limitations and Further Research Questions section). Only a minority of 17.1% (n=156) refuse to use it. The proportion of people with no clear opinion (partly partly or do not know) is significantly higher (n=272, 29.9%).

The legislative proposals also introduce the sharing of health data from the EHR for research purposes, as long as individuals do not actively object. Figure 1 summarizes whether respondents would be willing to share their personal health data stored in an EHR. The figure summarizes the attitudes of respondents with and without EHR. Respondents are more cautious about data sharing compared to the interest in using the EHR. In total, 43.4% (n=432) of respondents are willing to share their health data (yes, definitely or rather yes). Nearly 1 in 4 (n=225, 22.6%) respondents oppose data sharing. In addition, a large proportion of respondents are undecided (partly or not at all or do not know). A high interest in using the EHR corresponds empirically with a high willingness to share health data via EHR (no figure). The bivariate correlation is considerably large with 0.56 (Spearman).

The attitude toward health data sharing varies considerably with the data user respondents have in mind (Figure 2). The survey included a question regarding health data sharing with different institutions. It has to be noted that this question did not mention the EHR directly but addressed the willingness to share health data more generally. While 28.7% (n=285) and 18.8% (n=186) do not want to share data with public research institutions or public health insurances, about 38% (n=377) object to share their personal health data with private research institutions.

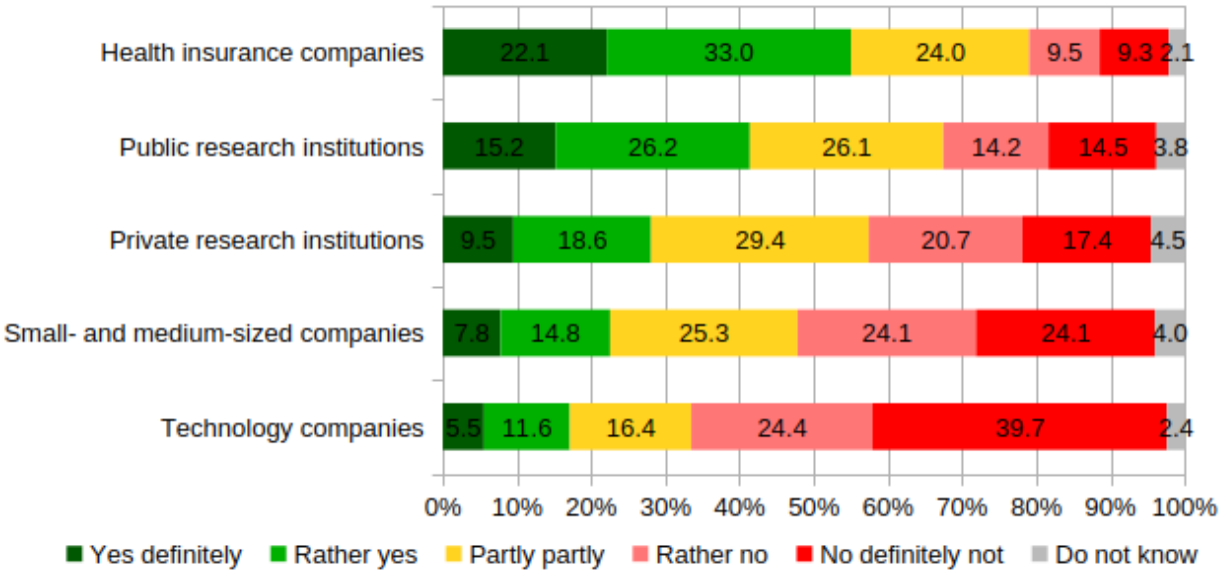
In the next step, the interest to use the EHR and the willingness to share health data for research purposes are measured for different socioeconomic groups. Table 1 shows the percentage of respondents in each group who intend to use the EHR and how many would agree with data sharing (yes and rather yes). In addition, respondents were asked for their opinions on different consent mechanisms. Figure 1 shows the proportion of people who explicitly reject the opt-out procedure as it will be introduced by law in 2025—compared to those who agree or are undecided (do not know).

Figure 1. Opinions about the use of electronic health record (EHR) and data sharing for research purposes (in percentages). Own calculation and weighted results for the following questions: "Would you like to use the EHR to view and manage your digital health data?" (n=911; nonresponse and people with EHR excluded) and "Would you be willing to share your health data from your EHR anonymized for medical research purposes?" (n=995; nonresponse excluded).



XSL•FO
RenderX

Figure 2. Willingness to share personal health data by data user (in percentages). Own calculation, weighted results for the question, “Would you be willing to share your health data anonymously for medical research purposes with the following institutions?” (nonresponse excluded; n=993).



XSL•FO
RenderX

Table . Socioeconomic differences in opinions about electronic health record (EHR)[a].

	Intention to use EHR, n (%)	Willingness to share data of EHR, n (%)	Disagree with opt-out, n (%)
Region			
West Germany	386 (52.1)	350 (43.2)	372 (46.8)
East Germany	97 (57)	82 (44.3)	67 (36)
Gender			
Woman	237 (50.2)	200 (39.1)	228 (45)
Man	245 (56)	232 (47.8)	210 (44.4)
Migrant (not born in Germany)			
No	484 (56.6)	430 (46)	409 (44.5)
Yes	18 (37.3)	18 (33.5)	25 (48.7)
Income quartile			
First (<1250)	67 (48.6)	50 (33.4)	57 (38.5)
Second (<1833)	120 (54.4)	104 (43.9)	90 (38.8)
Third (<2500)	103 (56.5)	97 (49.2)	91 (47.3)
Fourth	134 (61.4)	142 (57.3)	119 (48)
Age (years)			
18 - 39	120 (48.8)	116 (41.2)	121 (44)
40 - 59	164 (52.7)	140 (42.1)	163 (49.5)
Above 60	201 (56.6)	176 (46.5)	154 (41.2)
Educational level			
Low	99 (51.8)	79 (38.6)	70 (35.3)
Medium	182 (52.9)	172 (46.8)	148 (40.6)
High school degree	84 (49.4)	68 (35.9)	98 (51.9)
University degree	118 (57.7)	116 (50.8)	131 (58.1)
Total sample	483 (53)	432 (43.4)	438 (44.7)
Total respondents, n	911	995	980

[a]Own calculations, weighted percentages for columns 1 and 2 (share of a sample with statement yes definitely and rather yes) and column 3 (share of a sample who disagrees opt-out) in response to the question: “The German government is planning to use data from the electronic health records for medical research, eg, for research on new diseases and therapeutic approaches. Which of the following consents do you agree with? ... My data will be shared until I object.” Options to answer: agree or disagree or do not know or nonresponse; nonresponse excluded; do not know is interpreted as no interest to use EHR, not in favor of data sharing and agree with opt-out.

There are clear socioeconomic differences in interest in using the EHR and willingness to share health data, which are similar for both variables. The higher a person’s socioeconomic position (like income or education), the more open respondents are to using the EHR and sharing their personal health data for research purposes. Older people and men also tend to be more open. There are pronounced differences by migration status. The final column summarizes attitudes toward the opt-out process. First of all, it is striking that the procedure implemented by the legislator is explicitly rejected by a significant part of the population. In the entire sample, 44.7% (n=438) say that they are against an opt-out. Additional analyses not presented here show that the remaining respondents either agree (n=400, 40.8%) or have no opinion (n=142, 14.5%). Interestingly, the opt-out model is mainly rejected by people in a higher socioeconomic position. There are hardly any differences with regard to the other socioeconomic variables, except for a lower level of rejection in the eastern part of Germany. Overall, the descriptive statistics suggest that citizens are open to EHR and data sharing but want a high degree of autonomy in dealing with their health data. Additional analyses not shown here indicate that active consent is supported by over 80% (n=788) of respondents.

In the next step, the willingness to share data via the EHR is examined in a multivariate linear regression model (Table 2). The dependent variable is the willingness to share personal health data expressed on a 5-point scale (Figure 1). To prevent an excessive reduction in the sample, people who answered “do not know” were assigned to the “partly partly” category. The findings presented here remain stable even with a smaller sample. The determinants identified in the State of the Art section are added as independent variables to the model. Special

XSL•FO
RenderX

attention is given to the role of trust. The first model includes socioeconomic parameters. As shown in Table 1, the multivariate model confirms socioeconomic differences. However, some more differentiated statements can be made. Controlling for other parameters, there is a curvilinear relationship between age and willingness to share. The positive squared age term indicates that the willingness to share is highest among the very young and the very old. The willingness to share is lowest among people in their 50s. On a 5-point scale, their willingness is about 0.5 points lower than that of 20-year-old people. Income has a positive significant influence on willingness to share, which confirms the descriptive findings. The level of education also has an influence. However, when controlling for income, medium levels of education are associated with the highest willingness to share. Women and people born abroad show no different opinion regarding health data sharing after controlling for income and education. The model also includes a dummy variable indicating the presence of chronic diseases. On the one hand, people with chronic diseases may be at a higher risk of data misuse, but on the other hand, they may be more likely to benefit from digitally supported medical progress. In fact, interest in sharing data appears to be 0.2 points higher for people with chronic physical diseases.

The second model adds comfort with digital technologies. As in comparable studies from other countries, there are significant effects. People who find it easy to use digital technology are more open to sharing their health data. However, the effect disappears when the trust variables are included (model 3). Trust turns out to be a very strong influencing factor. Trust in science has the strongest influence. This is particularly noteworthy because the trust indicator does not specifically address the EHR or the institutions behind it but science as a whole. Controlling for socioeconomic parameters, a person who mistrusts the scientific system has a willingness to share that is more than 2 points lower on a scale of 5 than a person who fully trusts the scientific system. In addition, trust in the big technology firms also has a highly significant influence on the hypothesized direction. The strong influence of trust can be seen directly in the sharp increase in the fit of the model (R^2). The model also includes a proxy for altruistic tendencies in the form of a general willingness to donate (eg, money). The dummy for people who donate on a regular basis is significant. In the last model, motivations and concerns that may be associated with data sharing were included. However, the last model should be interpreted with caution. Because the wording of the questions in the independent variables refers to data sharing, the coefficients may be biased by self-reference (eg, if you care about doing good for society by sharing your data, you will more be willing to share). Nevertheless, the coefficients provide important clues about the motives that drive data sharing. The multivariate results show an interplay between altruistic motives (doing good for society) and cost-benefit considerations (potential benefits for personal and fear of data breaches that expose personal health data to the public).

XSL•FO
RenderX

Table . Willingness to share data of electronic health record for research purposes—linear regression model[a].

	Model 1 (n=828)		Model 2 (n=828)		Model 3 (n=828)		Model 4 (n=828)	
	Coefficient	*P* value	Coefficient	*P* value	Coefficient	*P* value	Coefficient	*P* value
Migrant (0/1[b])	−0.110	.57	−0.116	.54	−0.0190	.92	−0.0431	.83
Gender (reference=woman and 1 nonbinary person)								
Man (0/1)	0.153	.09	0.120	.19	0.0947	.24	0.140	.07
Age (years)	−0.0530[c]	.002	−0.0543[c]	.002	−0.0254	.11	−0.0133	.38
Age² (years)	0.000515[c]	.002	0.000542[c]	.001	0.000293	.06	0.000163	.27
Educational level (reference=medium)								
Low (0/1)	−0.233	.05	−0.217	.07	−0.184	.08	−0.151	.13
High school (0/1)	−0.321[d]	.02	−0.326[d]	.01	−0.319[c]	.01	−0.251[d]	.03
University (0/1)	0.00286	.98	−0.0100	.93	−0.0246	.83	−0.00558	.95
Income[e]	0.000130[c]	.002	0.000128[c]	.002	0.000136[f]	<.001	0.000116[f]	.001
Chronic disease (0/1)	0.201[d]	.03	0.206[d]	.03	0.118	.19	0.0804	.31
Mental illness (0/1)	0.0810	.56	0.0855	.54	0.0441	.73	−0.00517	.96
Rare internet use (0/1)[g]	__[h]	—	−0.229	.09	−0.235	.07	−0.0940	.42
Digital openness[i]	—	—	0.107[d]	.02	0.0330	.42	0.0379	.33
Trust in ...[j]								
Sciences	—	—	—	—	0.417[f]	<.001	0.191[f]	<.001
Big technology companies	—	—	—	—	0.181[f]	<.001	0.0948[d]	.02
Donate regularly (0/1)[k]	—	—	—	—	0.183[d]	.026	0.0474	.54
Motives for data sharing								
Doing good for society[l]	—	—	—	—	—	—	0.206[f]	<.001
Improve my health[m]	—	—	—	—	—	—	0.293[f]	<.001
Afraid of data breach (0/1)[n]	—	—	—	—	—	—	−0.206[c]	.007
Constant	4.307[f]	<.001	3.927[f]	<.001	1.448[c]	.002	0.571	.19

[a]Own calculations, weighted results. Model 1: R^2=0.059; model 2: R^2=0.072; model 3: R^2=0.270; and model 4: R^2=0.430.

[b]0/1 points to dummy variables.

[c]*P*<.01.

[d]*P*<.05.

[e]Net household equivalence income, imputed (5 times).

[f]*P*<.001.

[g]Once a week or less.

[h]Not applicable; regression model with fewer covariates.

[i]Using digital hardware and software feels easy for me (1.5=agree totally).

XSL•FO
RenderX

[j]Trust (1.5=trust totally) in technology companies like Google, Facebook, Apple or in sciences.

[k]Donate for charity on a regular basis (monthly or yearly).

[l]Subjective importance that with data sharing I am doing good for society (1.5=very important).

[m]Subjective importance that with data sharing I profit regarding my personal health.

[n]I am worried that my data will be made public (0=disagree or 1=agree 1).

Discussion

Principal Findings

The results of this study shed light on attitudes toward the use and sharing of EHRs for research purposes in Germany, particularly in the context of the new legislation based on the opt-out mechanism. Overall, the study reveals a nuanced picture of the German population's attitudes. The empirical results show an openness to EHRs (about half of the nonusers express a desire to use their EHR, signaling significant potential for increased adoption). At the same time, a significant share of the population remains skeptical or undecided. The overall openness is in line with international research regarding interest in EHR [12]. Given these high levels of interest, the introduction of an opt-out model does not seem unjustified. However, especially a large fraction of undecided people need to be properly informed about the consequences of having their health data stored electronically. While there is general support for sharing health data, with 43.4% (n=432) of respondents indicating a willingness to do so, a significant portion (n=225, 22.6%) remain opposed, and a large number are undecided. The results indicate a significantly lower willingness to share health data than has been measured in comparable studies in Germany [15,33]. Possible explanations for these differences are (1) different timing: other studies collected their data at a time when there was no legal framework for health data sharing and therefore risks, etc, had not yet been widely discussed in public. (2) Methodological: this study is based upon a telephone survey rather than web-based and explicitly allows people to have no specific attitude by including categories like refuse to answer, do not know, and neither nor. Last but not least, the comparatively low willingness to share health data in this study compared to international research [27,28] may be due to the emphasis on the personal character of the data potentially shared. The ambivalent results highlight the need for careful consideration of how the opt-out mechanism is implemented and communicated. Although there is openness in principle, "soft paternalism" must be viewed critically precisely because of the high proportion of undecided citizens. In contrast to the principle of informed consent, in which the processes and risks of data sharing have to be explained before citizens allow data sharing, an opt-out solution presupposes that the population is sufficiently informed in other ways. This becomes even more clear after considering attitudes toward the opt-out mechanism, which 44.7% (n=438) disapprove. Especially people who are positive about EHRs and the sharing of their health data reject the opt-out mechanism proposed by the legislator more often. Presumably, "choice architecture" is seen as too much interference in their own decisions.

The study also confirms a digital divide, as socioeconomic factors significantly influence both the willingness to use EHRs and the willingness to share personal health data. The results are in line with international research findings [16,30]. Higher income, education, and digital literacy are associated with greater openness, while those with lower socioeconomic status are more reluctant. However, socioeconomic factors are not the most important in explaining openness to data sharing. Trust plays a critical role in determining the willingness to share health information. While other studies have pointed to the role of trust as a significant factor [15,40,44,45], this study shows that institutional trust has a decisive influence compared to other factors. The influence of socioeconomic variables, as well as other factors such as digital literacy, diminishes when trust variables are taken into account, highlighting trust as a fundamental element in the acceptance of digital health initiatives. The observed socioeconomic disparities reinforce the need for targeted interventions and information to ensure equitable access and participation in digital health.

Policy Implications

The paradigm shift to an opt-out system is expected to increase participation rates because it relies on passive inclusion rather than active consent. However, this comes at the cost of potentially creating a population of uninformed and passive citizens who are enrolled in the system without fully understanding the implications. This raises critical concerns about consumer autonomy and challenges the notion of the autonomous consumer who actively engages with—and consents to—data-sharing practices. With a large portion of the population still undecided about EHR adoption and data sharing, the need for population-wide education becomes even more apparent. Ideally, all citizens, regardless of their socioeconomic background, should be able to make informed and autonomous decisions about their participation. The study also shows the complex mechanisms of the digital divide, with those in higher socioeconomic positions being more open to the adoption of EHRs but also more likely to oppose the opt-out system. This suggests that the paradigm shift is perceived as a restriction of autonomy. Those with greater resources and digital literacy want to make informed choices and therefore resist mechanisms that limit their freedom of choice. While those with greater resources are likely to be able to opt out if they perceive the risks of data sharing to be too high, they could trigger serious political resistance to the opt-out mechanism. In contrast, individuals with lower socioeconomic status may be less able to navigate these systems and therefore are more vulnerable to passive inclusion.

The transition to an opt-out system for EHRs in Germany marks a significant policy shift with the potential to transform the health sector. However, for this shift to be successful, it is essential to address the underlying concerns of the population, particularly around data security, autonomy, and the digital divide. Continuous monitoring of the implementation of the opt-out system will be crucial to keep and build trust within the population.

XSL•FO
RenderX

Limitations and Further Research Questions

In contrast to web-based surveys, which are often used in EHR research, the CATI method used here avoids many problems of self-reporting, sample bias, and inference [48,49]. However, the quality of data collected through telephone interviews also experiences low response rates. It seems highly plausible that nonresponse is not random. People with lower levels of trust and willingness are supposedly less likely to participate in the survey. Thus, despite the weighting procedure applied, the sample may be biased toward more trusting people. This problem is common to other data used in comparable research.

Another limitation concerns the correct understanding of the EHR in Germany. During the field phase, little more than 1% of the insured population opted for the EHR. It is possible that many respondents do not have a proper understanding of the EHR or have no knowledge at all. The survey provides some evidence in this direction. A question about knowledge of the EHR reveals that 13.9% (n=139) have never heard of it. For this group, the evaluation of data sharing is probably less substantiated. However, at the same time, a large fraction says that they have heard of the EHR (n=720, 71.7%) or even use it (n=89, 8.9%). On the one hand, these results support the assertion that a large proportion of the sample has a basic understanding of the EHR. During the field phase, there was also an extensive media discourse about the EHR, as the Gesundheitsdatennutzungsgesetz was being discussed in parliament in December 2023. On the other hand, the self-reported use of the EHR does not match the official statistics. Some respondents probably confused the EHR with other apps provided by their health insurance. This problem of overreporting has also been observed in other studies [15].

The study focuses on the interest in using the EHR and the willingness to share health data. At the time of the field phase, EHR adoption was very low, and a legal framework for health information exchange was in its formative stages. Research has already shown that there is a gap between attitudes and actual behavior [12,42]. So the results have to be interpreted with some caution. In addition, the high proportion of undecided respondents indicates that reported attitudes are not very salient and therefore susceptible to survey effects. However, other studies share this disadvantage. Moreover, the consequences of limited interest in EHR do not only affect surveys but also have far-reaching practical implications, especially when using opt-out regulations.

Future research should focus on the gap between interest in using the EHR and actual behavior. With a focus on Germany, the potential burdens of adoption should be analyzed in more detail. With regard to trust, the potential ways to build trust within the health care system without relying too much on general institutional trust also point to new research directions. Furthermore, there are interesting research questions regarding self-responsibility in a system that becomes largely based on passive consent (opt-out).

Acknowledgments

This work was supported by the German Federal Ministry of Education and Research and the European Union—Next Generation (grant 16KISA015). The funded project is called AVATAR ("Anonymisierung persönlicher Gesundheitsdaten durch Erzeugung digitaler Avatare") and consists of several research institutions that aim to develop new methods for anonymizing patient data for research. The funding agency had no role in the design, conduct, analysis, interpretation, or approval of the study. The content is the sole responsibility of the author. The author would like to thank Elias Kühnel, Julia Berghäuser, Sven Bock, and Anna Weber for their thorough work on the questionnaire.

Conflicts of Interest

None declared.

References

1. Elektronische Patientenakte bringt Entbürokratisierung und entlastet Ärzte und Pflegekräfte [Web page in German]. Bundesministerium für Gesundheit. URL: https://www.bundesgesundheitsministerium.de/presse/interviews/interview/spiegel-10-11-23-epa-und-ki [accessed 2025-03-05]
2. Digitalisierung für Gesundheit: Ziele und Rahmenbedingungen eines dynamisch lernenden Gesundheitssystems [Web page in German]. Gutachten. 2021. URL: https://www.svr-gesundheit.de/fileadmin/Gutachten/Gutachten_2021/SVR_Gutachten_2021.pdf [accessed 2025-04-08]
3. Strotbaum V, Pobiruchin M, Schreiweis B, Wiesner M, Strahwald B. Your data is gold—data donation for better healthcare? it - Information Technology 2019 Oct 25;61(5-6):219-229. [doi: 10.1515/itit-2019-0024]
4. Lupton D. Digital Health: Routledge; 2018.
5. Caspar J. Wir Datensklaven [Book in German]: Econ; 2023.
6. Matzat U, van Ingen E, Maasen S, Passoth JH. Social inequality and the digital transformation of western society: what can stratification research and digital divide studies learn from each other? In: Maasen S, Passoth JH, editors. Soziologie des Digitalen—Digitale Soziologie: Nomos; 2020:381-397.
7. Merkel S, Hess M. The use of internet-based health and care services by elderly people in Europe and the importance of the country context: multilevel study. JMIR Aging 2020 Jun 3;3(1):e15491. [doi: 10.2196/15491] [Medline: 32490837]
8. Thaler RH, Sunstein C. Nudge Improving Decisions about Health, Wealth and Happiness: Penguin; 2009.

XSL•FO
RenderX

9. Sunstein CR. Choosing Not to Choose: Understanding the Value of Choice: Oxford University Press; 2015.
10. Hoerbst A, Kohl CD, Knaup P, Ammenwerth E. Attitudes and behaviors related to the introduction of electronic health records among Austrian and German citizens. Int J Med Inform 2010 Feb;79(2):81-89. [doi: 10.1016/j.ijmedinf.2009.11.002] [Medline: 20031482]
11. Kruse CS, Stein A, Thomas H, Kaur H. The use of electronic health records to support population health: a systematic review of the literature. J Med Syst 2018 Sep 29;42(11):214. [doi: 10.1007/s10916-018-1075-6] [Medline: 30269237]
12. Crameri KA, Maher L, Van Dam P, Prior S. Personal electronic healthcare records: what influences consumers to engage with their clinical data online? A literature review. Health Inf Manag 2022 Jan;51(1):3-12. [doi: 10.1177/1833358319895369] [Medline: 31918578]
13. Zheng H, Jiang S. Frequent and diverse use of electronic health records in the United States: a trend analysis of national surveys. Digit Health 2022;8:20552076221112840. [doi: 10.1177/20552076221112840] [Medline: 35832476]
14. Seifert A. Nur wenige Deutsche nutzen bislang die Elektronische Patientenakte [Web page in German]. mdr. 2024. URL: https://www.mdr.de/nachrichten/deutschland/gesellschaft/wenige-elektronische-patientenakten-bisher-100.html [accessed 2025-03-05]
15. Haug S, Schnell R, Raptis G, Dotter C, Weber K. Wissen und Einstellung zur Speicherung und Nutzung von Gesundheitsdaten: Ergebnisse einer Bevölkerungsbefragung [Article in German]. Z Evid Fortbild Qual Gesundhwes 2024 Mar;184:50-58. [doi: 10.1016/j.zefq.2023.11.001]
16. Kim EH, Kim Y. Digital divide: use of electronic personal health record by different population groups. Presented at: 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology; Aug 31 to Sep 4, 2010; Buenos Aires, Argentina. [doi: 10.1109/iembs.2010.5626732]
17. Halmdienst N, Pruckner GJ, Winter-Ebmer R. Complexities of health and acceptance of electronic health records for the Austrian elderly population. Eur J Health Econ 2023 Feb;24(1):53-66. [doi: 10.1007/s10198-022-01451-z] [Medline: 35287235]
18. Paccoud I, Baumann M, Le Bihan E, et al. Socioeconomic and behavioural factors associated with access to and use of personal health records. BMC Med Inform Decis Mak 2021;21:18. [doi: 10.1186/s12911-020-01383-9]
19. Henkenjohann R. Role of individual motivations and privacy concerns in the adoption of German electronic patient record apps—a mixed-methods study. Int J Environ Res Public Health 2021 Sep 10;18(18):9553. [doi: 10.3390/ijerph18189553] [Medline: 34574475]
20. Griesser A, Bidmon S. A process related view on the usage of electronic health records from the patients' perspective: a systematic review. J Med Syst 2022 Dec 29;47(1):2. [doi: 10.1007/s10916-022-01886-0] [Medline: 36580132]
21. Stablein T, Hall JL, Pervis C, Anthony DL. Negotiating stigma in health care: disclosure and the role of electronic health records. Health Sociol Rev 2015 Sep 2;24(3):227-241. [doi: 10.1080/14461242.2015.1078218]
22. Kalckreuth N, Prümper AM, Feufel MA. The influence of health data on the use of the electronic health record (EHR): a mixed methods approach. 2023 Presented at: AMCIS 2023 Proceedings; Aug 10, 2023; Panama.
23. Eriksson-Backa K, Hirvonen N, Enwald H, Huvila I. Enablers for and barriers to using My Kanta— a focus group study of older adults' perceptions of the National Electronic Health Record in Finland. Inform Health Soc Care 2021 Dec 2;46(4):399-411. [doi: 10.1080/17538157.2021.1902331] [Medline: 33787438]
24. Greenberg AJ, Falisi AL, Finney Rutten LJ, et al. Access to electronic personal health records among patients with multiple chronic conditions: a secondary data analysis. J Med Internet Res 2017;19(6):e188. [doi: 10.2196/jmir.7417]
25. Nolte D. Vom Daten-Grab zum Daten-Schatz? Strukturierte Behandlungsprogramme (Disease Management Programme) für chronisch Kranke. Zugluft 2021(2):82-87 [FREE Full text]
26. Caine K, Kohn S, Lawrence C, Hanania R, Meslin EM, Tierney WM. Designing a patient-centered user interface for access decisions about EHR data: implications from patient interviews. J Gen Intern Med 2015 Jan;30(Suppl 1):S7-16. [doi: 10.1007/s11606-014-3049-9] [Medline: 25480719]
27. Skatova A, Goulding J. Psychology of personal data donation. PLOS ONE 2019;14(11):e0224240. [doi: 10.1371/journal.pone.0224240] [Medline: 31747408]
28. Howe N, Giles E, Newbury-Birch D, McColl E. Systematic review of participants' attitudes towards data sharing: a thematic synthesis. J Health Serv Res Policy 2018 Apr;23(2):123-133. [doi: 10.1177/1355819617751555]
29. Kalkman S, van Delden J, Banerjee A, Tyl B, Mostert M, van Thiel G. Patients' and public views and attitudes towards the sharing of health data for research: a narrative review of the empirical evidence. J Med Ethics 2022 Jan;48(1):3-13. [doi: 10.1136/medethics-2019-105651] [Medline: 31719155]
30. Hirst Y, Stoffel ST, Brewer HR, Timotijevic L, Raats MM, Flanagan JM. Understanding public attitudes and willingness to share commercial data for health research: survey study in the United Kingdom. JMIR Public Health Surveill 2023;9:e40814. [doi: 10.2196/40814]
31. Kühnel E, Wilke F. Potentiale der Digitalisierung heben: Die Bereitschaft zur Spende von Gesundheitsdaten in Deutschland und Finnland [Chapter in German]. In: Heckes KT, Lorke M, Siegler M, Wrona KJ, editors. Soziotechnische Transformationen im Sozial- und Gesundheitswesen: kollaborativ, divers, barrierefrei und sozialräumlich: Juventa; 2025.

XSL•FO
RenderX

32. Richter G, Borzikowsky C, Hoyer BF, Laudes M, Krawczak M. Secondary research use of personal medical data: patient attitudes towards data donation. BMC Med Ethics 2021 Dec 15;22(1):164. [doi: 10.1186/s12910-021-00728-x] [Medline: 34911502]
33. Richter G, Trigui N, Caliebe A, Krawczak M. Attitude towards consent-free research use of personal medical data in the general German population. Heliyon 2024 Mar 30;10(6):e27933. [doi: 10.1016/j.heliyon.2024.e27933] [Medline: 38509969]
34. Lee H, Kim S, Couper MP, Woo Y. Experimental comparison of PC web, smartphone web, and telephone surveys in the new technology era. Soc Sci Comput Rev 2019 Apr;37(2):234-247. [doi: 10.1177/0894439318756867]
35. Milton AC, Ellis LA, Davenport TA, Burns JM, Hickie IB. Comparison of self-reported telephone interviewing and web-based survey responses: findings from the Second Australian Young and Well National Survey. JMIR Ment Health 2017 Sep 26;4(3):e37. [doi: 10.2196/mental.8222] [Medline: 28951382]
36. Moon LA. Factors influencing health data sharing preferences of consumers: a critical review. Health Policy Technol 2017 Jun;6(2):169-187. [doi: 10.1016/j.hlpt.2017.01.001]
37. Karampela M, Ouhbi S, Isomursu M. Connected health user willingness to share personal health data: questionnaire study. J Med Internet Res 2019 Nov 27;21(11):e14537. [doi: 10.2196/14537] [Medline: 31774410]
38. Weng C, Friedman C, Rommel CA, Hurdle JF. A two-site survey of medical center personnel's willingness to share clinical data for research: implications for reproducible health NLP research. BMC Med Inform Decis Mak 2019 Apr 4;19(Suppl 3):70. [doi: 10.1186/s12911-019-0778-z] [Medline: 30943963]
39. Broekstra R, Aris-Meijer JL, Maeckelberghe ELM, Stolk RP, Otten S. Motives for withdrawal of participation in biobanking and participants' willingness to allow linkages of their data. Eur J Hum Genet 2022 Mar;30(3):367-377. [doi: 10.1038/s41431-021-00997-5] [Medline: 34803164]
40. Mählmann L, Schee gen. Halfmann S, von Wyl A, Brand A. Attitudes towards personal genomics and sharing of genetic data among older Swiss adults: a qualitative study. Public Health Genomics 2018 Apr 10;20(5):293-306. [doi: 10.1159/000486588]
41. Braunack-Mayer A, Fabrianesi B, Street J, et al. Sharing government health data with the private sector: community attitudes survey. J Med Internet Res 2021 Oct 1;23(10):e24200. [doi: 10.2196/24200] [Medline: 34596573]
42. Benevento M, Mandarelli G, Carravetta F, et al. Measuring the willingness to share personal health information: a systematic review. Front Public Health 2023;11:1213615. [doi: 10.3389/fpubh.2023.1213615] [Medline: 37546309]
43. Kühnel E, Wilke F. Health data sharing in Germany: individual preconditions, trust and motives. Front Public Health 2025;13:1538106. [doi: 10.3389/fpubh.2025.1538106] [Medline: 39975784]
44. Weitzman ER, Kaci L, Mandl KD. Sharing medical data for health research: the early personal health record experience. J Med Internet Res 2010 May 25;12(2):e14. [doi: 10.2196/jmir.1356] [Medline: 20501431]
45. Maiorana A, Steward WT, Koester KA, et al. Trust, confidentiality, and the acceptability of sharing HIV-related patient data: lessons learned from a mixed methods study about Health Information Exchanges. Implement Sci 2012 Apr 19;7:34. [doi: 10.1186/1748-5908-7-34] [Medline: 22515736]
46. Meyer S, Ward P, Coveney J, Rogers W. Trust in the health system: an analysis and extension of the social theories of Giddens and Luhmann. Health Sociol Rev 2008 Aug;17(2):177-186. [doi: 10.5172/hesr.451.17.2.177]
47. Wilke F, Kühnel E. Eine Befragung zur Bereitschaft zum Teilen von Gesundheitsdaten in Deutschland—Einflussfaktoren, Präferenzen und Motive. : GESIS, Köln; 2024.
48. McPhee C, Barlas F, Brigham N, Darling J, Dutwin D, et al. Data quality metrics for online samples: considerations for study design and analysis. : American Association for Public Opinion Research; 2022 URL: https://aapor.org/wp-content/uploads/2023/02/Task-Force-Report-FINAL.pdf [accessed 2025-04-01]
49. Pokorny S, Hirndorf D. Online, offline oder beides? Umfragemethoden im Praxistest [Article in German]. Analysen Argumente 2024(519):1-13 [FREE Full text]

Abbreviations

CATI: computer-assisted telephone interview
EHR: electronic health record

Edited by A Kushniruk; submitted 23.08.24; peer-reviewed by S Haug, SH Sung; revised version received 04.02.25; accepted 04.03.25; published 15.04.25.

Please cite as:
Wilke F
The Impact of Trust and the Role of the Opt-Out Mechanism in Willingness to Share Health Data via Electronic Health Records in Germany: Telephone Survey Study
JMIR Hum Factors 2025;12:e65718
URL: https://humanfactors.jmir.org/2025/1/e65718
doi:10.2196/65718

XSL•FO
RenderX

© Felix Wilke. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 15.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Collecting at-Home Biometric Measures for Longitudinal Research From the i3C: Feasibility and Acceptability Study

Marta Russell[1,2], BA; Erin Cain[1,2], BS; Lydia Bazzano[3], MD, PhD; Ileana De Anda[3], MD, MPH; Jessica G Woo[2,4], PhD; Elaine M Urbina[1,2], MD, MS

[1]The Heart Institute, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States

[2]Department of Pediatrics, University of Cincinnati College of Medicine, 3333 Burnet Ave, MLC 7002, Cincinnati, OH, United States

[3]Department of Epidemiology, School of Public Health, Tulane University, New Orleans, LA, United States

[4]Division of Biostatistics and Epidemiology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States

Corresponding Author:
Elaine M Urbina, MD, MS
The Heart Institute, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States

Abstract

Background: The use of individual wearable devices or internet-based applications to collect biometric data from research participants is popular, but several devices may be needed to replace a full set of research measurements.

Objective: In this study, we assessed the feasibility of a "Virtual Home Clinic" within the context of long-term epidemiologic studies.

Methods: Participants from 3 study cohorts were recruited. Devices were sent to the home to measure anthropometrics, resting metabolic rate, blood pressure (BP), heart rate (HR), heart rhythm, oxygen saturation, glucose, total cholesterol, physical activity, diet, sleep duration or quality, and arterial stiffness over the course of 1 week. Stool and saliva were also self-collected for microbiome, DNA, and cotinine. Feasibility and acceptability of collecting measurements using home devices were assessed.

Results: A total of 134 participants were enrolled (87% female, 31% Black; mean age 54.2, SD 8.4 years). Furthermore, 91% (N=122) performed at least one of the home tests. At least two-thirds of participants were able to complete all of the requested readings for glucose, electrocardiogram, BP, diet record, and resting metabolic rate. The scale that measured weight, body composition, and pulse wave velocity (PWV) was more difficult to use (113/134, 84% participants recorded at least one weight and 84/134, 63% recorded a PWV). The device to measure total cholesterol was least successful (32/134, 24% participants completed all readings, 72/134, 54% provided at least one result). Return of biospecimens was highly successful (115/134, 86% returned saliva and 113/134, 84% returned stool). Of 95 who responded to the user acceptability survey, 38 (40%) participants preferred home assessment, 36 (38%) preferred clinic, and 21 (22%) did not have a preference. The mean user acceptability score across devices for ease of use was 4.3 (SD 1.0), for instructions was 4.5 (SD 0.7), and for time to use was 3.9 (SD 1.1; scale of 1 - 5, with higher scores indicating greater acceptability). The study team documented several regulatory or IT, connectivity or account, data retrieval, and logistical issues encountered during the study.

Conclusions: Despite several complications involved with managing multiple devices and applications, most of the components of the virtual home clinic were reasonably feasible and acceptable to participants.

(JMIR Hum Factors 2025;12:e71103) doi:10.2196/71103

KEYWORDS

mHealth; cardiovascular risk factors; epidemiology; wearable devices; feasibility; acceptability; biometric

Introduction

Early life factors play an important role in successful aging [1,2]. However, prospectively collected data is needed to construct accurate trajectories of health status (eg, BMI, blood pressure, and cholesterol levels) and lifestyle factors (eg, diet, physical activity, alcohol, and tobacco use) across the lifespan to determine their association with cognitive and physical function later in life. Many barriers exist in collecting longitudinal data across the lifespan [3]. Importantly, as adolescents age into adulthood, many move away from their initial recruitment site which may reduce generalizability of findings of the remaining cohort from adult follow-up visits. Mobile health (mHealth) approaches have been proposed as a means to overcome some of these barriers [4], but the feasibility of collecting a broad range of measurements using at-home devices has not been sufficiently evaluated in this setting. Therefore, we conducted a study to evaluate the feasibility of a "Virtual Home Clinic" by sending biometric measurement devices to participants' homes. We also sought to determine

XSL•FO
RenderX

the acceptability of the virtual home clinic approach compared with a research clinic visit. Separately, we will seek to correlate measures obtained in the virtual home clinic to those in the research clinic. Our goal is to establish the feasibility and acceptability of using a suite of wearable technology and online applications to conduct "remote" study visits in epidemiologic studies to improve the sustainability of longitudinal studies whose study population has dispersed over time.

Methods

Population

Three cohorts that are members of the International Childhood Cardiovascular Cohort Consortium (i3C) [5] participated in this study: the Bogalusa Heart Study (BHS), the National Heart, Lung, and Blood Institute (NHLBI) Growth and Health Study (NGHS), and the Princeton Lipid Research Study (PLRS). The most recent phone follow-up of these cohorts this project was completed in 2019 [6].

The BHS which began in 1973 conducted 9 cross-sectional cardiovascular risk factor screenings of children aged 3 to 18 years between 1973 and 1992 in a community of Black and White men and women of Bogalusa, LA, with participation rates of 80% to 93%. In addition, 10 cross-sectional screenings of young adults aged 19 - 52 years who had previously been examined as children have been conducted to date, with participation rates of 72% to 93%. Cardiovascular risk factors (adiposity, BP, and cholesterol levels) and noninvasive measures of cardiac and vascular health (carotid intima media thickness and arterial stiffness) [7].

The NGHS was initiated in 1987 as a study of racial (Black-White) differences in the development of cardiovascular risk factors in girls in Cincinnati, Ohio [7]. All participants were between the ages of 9 and 10 years at baseline and followed annually or biannually to age 28 years. Cardiovascular risk factors, echocardiography and dual-energy x-ray absorptiometry (DXA) and abdominal magnetic resonance imaging (MRI) for adiposity were collected.

The PLRS was initiated in 1973 as part of the NHLBI Lipid Research Clinics Study [8]. The study enrolled children in first to 12th grades in the Princeton City School District of Cincinnati, Ohio, plus a 50% random sample of the students' parents (by household). Three study visits were conducted between 1973 and 1978 to collect a variety of cardiovascular risk factors. Between 1998 and 2003, 911 former school children participated in another in-person visit at mean age 36.

For the current feasibility study of the virtual home clinic, 50 participants from each cohort living near the study sites, with access to Wi-Fi at home, and who were willing to attend an in-person visit at the study sites were invited to participate. Procedures were explained, and informed consent was obtained electronically. These 50 per cohort were invited from a subset of 150 per cohort who had previously enrolled in a pilot study testing the feasibility of phone administration of neurocognitive evaluations. Enrollment into this virtual home clinic study thus represents a convenience sample of contactable and willing participants. The study protocol was approved by the institutional review boards of the participating institutions.

Data Collection

Participants were asked to complete the virtual home clinic and an in-person research clinic visit for the validation portion of the study. The virtual home clinic study kit was either mailed to a participant (if they conducted the virtual home clinic C first) or given to the participant at the end of their in-person visit. The virtual home clinic kit included detailed instructions, staff contact information, an iPad (Apple Inc) with device apps and instructional videos preloaded, a color-coded schedule of activities, and shipping materials for returning the kit. The virtual home clinic study protocol included 6 devices, 1 iPad app and 3 sample collections (Table 1).

XSL•FO
RenderX

Table . Home devices and online applications included in the virtual home clinic.

Mobile technology devices and measurement domains and variables		Data retrieval method
Withings Body Cardio Scale		Log into participant's account in a browser and request data download. Click link in email to download a zip folder containing data in three separate CSV[a] files.
	Weight	
	Body composition: %body fat, fat mass, lean mass	
	Arterial stiffness: PWV[b]	
Breezing Pro device		
	Resting metabolic rate: resting energy expenditure (measured and predicted)	Log into participant's account in app. Select option to export entire history. Follow link in email to download CSV file.
AliveCor KardiaMobile 6L		
	6-lead ECG[c]: heart rate and rhythm (normal, sinus, bradycardia, and atrial fibrillation)	Log into participant's account in app. Download PDF of each ECG completed, then export each PDF via email. Save PDFs and enter results in study database.
QardioArm Home BP[d] Monitor		
	Blood pressure; systolic and diastolic blood pressure and heart rate	Log into participant's account in app. Choose the "Send History" option and enter an email address. CSV file of results will be emailed to that address.
AimStrip Tandem		
	Laboratory assays: glucose and total cholesterol	Use buttons on meter to view each test result. Enter results in study database.
GB HealthWatch 360		
	Dietary intake: mean total calories per day, % calories from fat, total sodium per day, and total added sugars per day	Log into study's HealthWatch portal in browser. Generate and download report for relevant dates. Data exported in two CSV files in a zip folder.
Garmin Vivosmart 4		
	Physical activity: number of steps, minutes of intense movement, and heart rate	Log into participant's account in browser. Under "Reports," generate and download CSV files for each measure. Date formats may need to be adjusted.
	Sleep: time in REM[e] + deep sleep and oxygen saturation	In browser, view sleep and oxygen saturation data and save a screenshot for each night. Enter results in study database.
DNA Genotek OMNIgene GUT OMR-200 stool collection		
	Microbiome (16s): % collected	Sample collection status recorded in spreadsheet.
Abbott Quantisal saliva collection		
	Cotinine	Samples sent to LabCorp for processing. Retrieve results from study's LabCorp portal and enter results in study database.
DNA Genotek Oragene OG-500 saliva collection		
	DNA: % collected	Sample collection status recorded in spreadsheet.

[a]CSV: comma separated values.

[b]PWV: pulse wave velocity.

[c]ECG: electrocardiogram.

[d]BP: blood pressure.

[e]REM: rapid eye movement

XSL•FO
RenderX

Together, the devices were selected to measure key cardiovascular parameters that would be included in a standard in-person follow-up visit. Requested measurements included 3 glucose and total cholesterol readings with the AimStrip Tandem (AimStrip), 3 electrocardiogram tracings with the Alivecor KardiaMobile 6L (Kardia), 7 weight or body composition and 3 pulse wave velocity (PWV) readings with the Withings Body Cardio scale (Withings), and 3 resting metabolic rate measurements with the Breezing Pro device (Breezing; Cincinnati cohorts only). One week of daily diet tracking was requested using the GB HealthWatch 360 app (HealthWatch 360), and participants were asked to wear a Garmin Vivosmart 4 fitness tracker watch (Garmin) for the entire week to measure activity (steps, minutes with elevated heart rate), oxygen saturation during sleep, and sleep duration and quality. In addition, 3 sets of 3 BP and HR measurements (total of 9) were requested using the QardioArm Home BP monitor (Qardio). Three biologic samples were requested: Abbott Quantisal saliva collection (Quantisal), DNA Genotek Oragene OG-500 saliva collection (Oragene) and DNA Genotek OMNIgene GUT OMR-200 stool collection (OMNIgene). Subjects were asked to return the kit and samples after 1 week. Once device data was retrieved, a full reset was performed on all hardware to ensure privacy between participants. For additional security, new iPad app accounts were created for each participant using only their unique study identifier and the demographics required by the app (eg date of birth, sex, height, and weight). Reset procedures were rigorously tested to confirm effective deletion of all data, and were documented in detail in study manuals. All kit components were thoroughly cleaned using hospital protocols before and after each virtual home clinic.

A questionnaire to assess acceptability, ease of use, and functionality of each device was sent to all participants electronically after completion of the virtual home clinic. User acceptability was assessed with the questions “Was this device or app easy to use?” (1=Very difficult to use; 2=Somewhat difficult; 3=Neither difficult nor easy; 4=Somewhat easy; and 5=Very easy); “Were the instructions easy to follow?” (same scoring); “How do you feel about the time it took to use the device or app?” (1=Took much longer than I would have liked; 2=Took somewhat longer; 3=Took about as long as I expected; 4=Took somewhat less time than I expected; and 5=Took very little time).

Statistical Analyses

Feasibility was assessed as proportion of requested in-home tests completed, and percent of requested saliva and stool samples returned. Mean scores by device and overall were calculated. Due to small sample size, differences by cohort, sex and race were not assessed.

Ethical Considerations

Each center obtained approval from their local institutional review board (IRB) and all participants signed an informed consent. All data was identified with only a participant ID number (deidentified). Participants were compensated $200 for completion of all arms of the study.

Results

Population

A total of 134 participants consented to participate (Table 2).

XSL•FO
RenderX

Table . Description of the study population. Three participants did not have laboratory evaluation.

	BHS[a] (n=41)	NGHS[b] (n=52)	PLRS[c] (n=41)	All (n=134)
Age (years), mean (SD)	57.1 (6.7)	45.5 (0.6)	62.3 (3.4)	54.2 (8.4)
Race (White), n (%)	34 (82.9)	25 (48.1)	34 (82.9)	93 (69.4)
Sex (male), n (%)	7 (17.1)	0 (0)	11 (26.8)	18 (13.4)
In-clinic measurements, mean (SD)				
BMI (kg/m^2)	31.3 (6.8)	32.4 (10.6)	30.8 (6.4)	31.6 (8.3)
Total cholesterol (mg/dL)	181.1 (35.4)	182.2 (48.4)	191.6 (35.8)	184.6 (41.0)
LDL-C[d] (mg/dL)	105.6 (31.8)	103.4 (38.8)	107.0 (28.0)	105.2 (33.5)
HDL-C[e] (mg/dL)	54.6 (16.9)	53.4 (16.7)	54.8 (15.2)	54.2 (16.2)
Triglycerides (mg/dL)	116.4 (52.5)	107.5 (109.3)	99.9 (52.9)	108.1 (79.6)
Glucose (mg/dL)	109.9 (45.1)	104.5 (32.9)	105.6 (21.0)	106.5 (34.3)
SBP[f] (mm Hg)	122.7 (17.8)	115.0 (13.8)	126.2 (16.7)	120.8 (16.6)
DBP[g] (mm Hg)	73.8 (10.0)	81.2 (11.2)	79.5 (8.8)	78.4 (10.6)
HR[h] (beats/min)	73.0 (14.6)	75.3 (11.0)	67.7 (10.2)	72.3 (12.3)

[a]BHS: Bogalusa Heart Study.

[b]NGHS: National Heart, Lung, and Blood Institute Growth and Health Study.

[c]PLRS: Princeton Lipid Research Study.

[d]LDL-C:

[e]HDL-C:high-density lipoprotein cholesterol.

[f]SBP: systolic blood pressure.

[g]DBP: diastolic blood pressure.

[h]HR: heart rate.

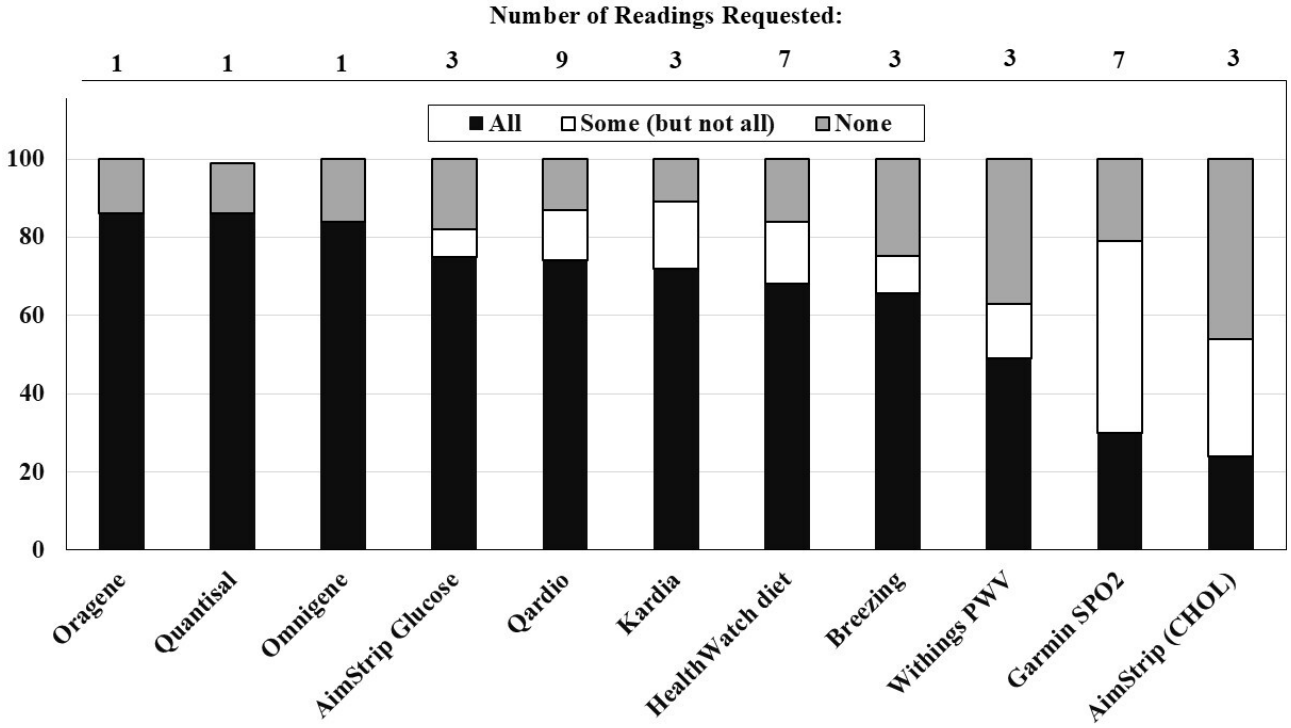
The mean age of the participants was 54.2 (SD 8.4) years, which differed somewhat by cohort. Males (13%) were less likely to participate than females (87%), partially due to the NGHS cohort consisting entirely of women. White individuals comprised 69% (93/134) of the participants and Black individuals were 31% (41/134). In-clinic measurements showed a mean BMI in the obese range (31.6, SD 8.3 kg/m^2), slightly elevated LDL (105.2, SD 33.5 mg/dL), glucose (106.5, SD 34.3 mg/dL), and systolic blood pressure (120.8, SD 16.6 mm Hg), and normal diastolic blood pressure (78.4, SD 10.6 mm Hg), total cholesterol (184.6, SD 41.0 mg/dL), and triglycerides (108.1, SD 79.6 mg/dL).

Data Collected

Table S1 in Multimedia Appendix 1 lists issues related to device or iPad application set up. General issues encountered included institutional firewalls requiring permission to load apps on an iPad; apps requiring Wi-Fi or a unique email address; complicated instructions for participants; different processes for account setup, data retrieval, cleaning or disinfection by device; need to delete old data or resetting for each device and participant; and differing exported data file formats by device (eg, CSV, PDF requiring data entry, or data entry directly from the device) requiring extensive programming to import into statistical software. Issues encountered during execution of the study (Table S2 in Multimedia Appendix 1) included IRB issues, changes in devices or apps during the study by the manufacturer, issues with batteries, and issues with passwords for apps. Success of completion of the virtual home clinic protocol by device or app type are displayed in Figure 1.

XSL•FO
RenderX

Figure 1. Completion success by device (N=134). PWV: pulse wave velocity, SpO_2: peripheral capillary oxygen saturation.



Overall, 122 out of 134 participants (91%) performed at least one of the home tests. Return of biospecimens was highly successful (115/134, 86% returned saliva and 113/134, 84% returned stool). At least 91 of the 134 (68%) participants were able to complete all of the requested readings for glucose, electrocardiogram, BP, and diet record. All requested resting metabolic rate measurements were obtained by 61 of the 93 (66%) NGHS and PLRS participants. The Withings scale that provided weight, body composition, and PWV was more difficult to use but 113 out of 134 (84%) participants provided at least one reading of weight, 109 (81%) provided at least one body composition reading, and 84 (63%) recorded a PWV. Physical activity (steps) and sleep quality and quantity were assessed in 113 out of 134 (84%) participants and oxygen saturation in 106 (79%) participants using the Garmin watch. The AimStrip device to measure total cholesterol resulted in the most calls to the research team and only 32 out of 134 (24%) participants completed all readings. At least one cholesterol result was provided by 72 out of 134 (54%)participants.

Participants were asked whether they preferred home or in-clinic assessment. Of the 95 who responded to this question, 38 (40%) preferred home assessment, 36 (38%) clinic and 21 (22%) did not have a preference. The mean user acceptability score (Table 3) across devices for ease of use was 4.3 (SD 1.0), for instructions was 4.5 (SD 0.7) and for time to use was 3.9 (SD 1.1). The Garmin, Qardio, and Kardia had the highest overall scores (4.7, SD 0.6; 4.6, SD 0.6; and 4.6, SD 0.7) followed by the Withings (4.2, SD 0.9), Breezing and HealthWatch 360 app (4.0, SD 0.7, and 3.9, SD 0.9), with the AimStrip achieving the lowest overall acceptability (3.6, SD 1.2). Greater success with a device was generally not correlated with higher acceptability ratings. This analysis was limited by the fact that participants completing few or no readings were less likely to complete the acceptability survey.

XSL•FO
RenderX

Table . Mean user acceptability score stratified by device.

Responses (n=100)	Withings	Breezing	Kardia	Qardio	Aim strip	HealthWatch 360	Garmin
Was this device or app easy to use?	4.1 (1.1)	4.1 (0.9)	4.6 (0.9)	4.7 (0.6)	3.6 (1.4)	4.0 (1.1)	4.7 (0.7)
Were the instructions for this device easy to follow?	4.4 (0.8)	4.5 (0.6)	4.8 (0.5)	4.8 (0.5)	4.1 (1.2)	4.5 (0.9)	4.8 (0.6)
How do you feel about the time it took to use the device/app?	3.9 (1.3)	3.4 (1.0)	4.4 (1.1)	4.3 (1.0)	3.2 (1.3)	3.4 (1.3)	4.6 (0.9)
Overall Acceptability Score (mean rating for the 3 factors assessed)	4.2 (0.9)	4.0 (0.7)	4.6 (0.7)	4.6 (0.6)	3.6 (1.2)	3.9 (0.9)	4.7 (0.6)

Discussion

Principal Findings

Our study demonstrated that, with careful implementation, collecting at-home biometric measures is feasible and generally acceptable to participants in a longitudinal study. Recently, attention has focused on the use of wearable technology and internet-based applications to capture data that otherwise could only be captured in the clinic. Although some technology, such as accelerometers, have been in use for years [9], there is increasing opportunity to use commercially-available wearable technology and other send-out or mail-back methods to collect epidemiologic data and biospecimens in lieu of clinic visits, even for measurements that have historically been exclusively conducted by medical personnel. However, there remains a considerable gap in knowledge about whether such devices are simple enough to facilitate the collection of valid research-quality data at home and whether these devices are acceptable in terms of ease of use and time commitment for participants participating in mHealth research studies. If these commercially available devices are feasible to use and acceptable to participants, the use of such technologies in extending the reach of epidemiologic studies to widely dispersed populations could revolutionize the field, reduce place-based follow-up bias, and improve the extent of measured data available in follow-up studies.

Our study found 122 out of our 134 (91%) participants performed at least one of the home tests thus demonstrating feasibility. Furthermore, return of biospecimens was highly successful. Home data collection was also highly acceptable with a mean user acceptability score across devices for ease of use of 4.3 out of 5. Interestingly, only 38 out of 95 (40%) participants assessed preferred a home assessment and 36 (38%) preferred in-clinic testing.

There were multiple general implementation issues identified, the majority of which arose because the devices or apps were developed for a single person's use, while this study required devices to be set up for multiple participants serially over the course of the study. The requirement for a unique email address for many devices, for example, quickly became an issue in this context, as individual participant data needed to be kept separate, but creation of an email address for each participant was not feasible. Device- or app-specific rules about creation or deletion of accounts required detailed protocols for study management. Similarly, data from many of the devices was primarily developed to be viewed in an app or portal by the individual using the device, rather than being exported and used in analyses. Thus, the processes for retrieving the data, and the formats available for the data exports, differed not only across but also within devices, creating complicated data management workflows. While work-arounds were developed to overcome nearly all of the issues, the burden on the study team was higher than for standard clinic measurements.

In terms of the specific devices we tested, our study is the first to date to demonstrate the feasibility and acceptability of the Withings scale to measure body composition in the home. This study also evaluated use of the Breezing in the home setting. All participants repeated the Breezing testing multiple times, and none reported complaints or issues; however, the sample size was small (n=20) and many did occasionally forget to perform the test in the fasting state. The present study did not ask whether participants were fasting for all tests, which would be a modification of study protocol in future studies. Another study evaluated the Breezing in the home in pregnant women [10]. They found 68% of participants performed all the requested readings, although 44% had technical issues [10]. User acceptability was modest with 56% reported they felt the device was enjoyable or had neutral feelings about it, while 44% reported it was not enjoyable [10]. We had higher completion rates and user acceptability but our population was not experiencing any pregnancy-related issues.

There are many studies validating apps to detect arrythmias such as the Kardia [11-14]; however, none of these studies provided user acceptability data. In a study using a different device to detect atrial fibrillation [15], adherence to the study

XSL•FO
RenderX

protocol (daily use for a year) was only 44%. Age, severity of symptoms, and BMI influenced adherence but no user acceptability data was provided [15].

There has been a rapid increase in commercially available direct-to-consumer home BP devices of which many, including the Qardio monitor used in this study, have undergone validation studies [16]. No data on acceptability has been published previously. The situation with home oxygen saturation measurements is similar. The Garmin Vivosmart had excellent agreement with a pulse oximeter at rest and 30 seconds after a sit-to-stand and 6-minute walk test [17], but no user acceptability data was provided and the device was not tested in the home. For more complicated hemodynamic measures such as PWV, there are no other commercially available devices for the home besides the Withings [18], and no user-acceptability data has been previously collected.

There are also a variety of commercially available products to measure lifestyle habits such as sleep, physical activity, and diet. The device we used (Garmin Vivosmart 4) has limited published validation data [19], but no studies on user-acceptability data. A study that used a Fitbit for activity monitoring found that 80% of participants provided at least some data during the 3-week study [20]. A few studies have evaluated smartphone apps to track dietary consumption [21], but none have published user acceptability data.

A few studies have examined general participant acceptability of mHealth apps in general. Harmon et al [22] summarized many of the barriers to implementation. For instance, older participants may not have a smartphone or experience in mobile apps. Waning participation over time is also noted when much of the data collection requires active participation from study subjects. Similar to our study, one group found that more than one-third of participants did not wear the device during the study period [23]. Push notifications can assist in reducing this issue but if the device app does not have this feature, sending reminders can be labor intensive. Also, reporting abnormal results requires specific infrastructure and supervision [22]. One study specifically measured acceptability based on a design where a single or multiple devices were used [24]. Although the use of multiple devices gave participants greater interest in the study and more perceived benefit, they had lower confidence in completing the measurements [24]. However, that study was a survey about attitudes toward mHealth studies and did not actually evaluate any specific devices.

Although our sample size is smaller than most epidemiologic studies, it is still larger than many of the mHealth papers that include user acceptability data published to date. We also had a predominance of females, few minorities and the requirement for Wi-Fi access for the participant to complete the at home visit. Although digital literacy was not directly assessed, our sample likely represents a reasonable cross-section of technological proficiency. The most willing and engaged participants likely self-selected themselves for participation as this was a subset of the full cohort. Participants were given some familiarity with the Breezing device during their in-clinic visit, so participants may have felt more comfortable with this device at home than would be the case in a truly remote visit where the participants would only have the study materials provided.

Conclusion

Our study demonstrated that a virtual home clinic is reasonably feasible and acceptable for remote assessments in longitudinal studies. Challenges for mHealth projects include manufacturer changes to the user interface or data extraction protocols over time, research-specific concerns such as maintaining unique profiles and emails for device apps without requiring the participant to use their personal information, and IRB and institutional concerns about device and participant safety. In addition, over the course of a longer study, the devices themselves might become obsolete and/or no longer supported by the manufacturer. Despite encountered challenges, successful data collection and positive user feedback suggest the potential for revolutionizing epidemiologic studies, especially for widely dispersed populations, and reducing geographic follow-up bias.

Acknowledgments

Funding was provided by the National Institutes of Health (National Institute of Aging; grant R21/ R33AG057983).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplemental tables 1 and 2.
[DOCX File, 22 KB - humanfactors_v12i1e71103_app1.docx]

References

1. Ryder JR, Northrop E, Rudser KD, et al. Accelerated early vascular aging among adolescents with obesity and/or type 2 diabetes mellitus. JAHA 2020 May 18;9(10):32370578. [doi: 10.1161/JAHA.119.014891]

XSL•FO
RenderX

2. Barker DJ. The fetal origins of coronary heart disease. Eur Heart J 1997 Jun;18(6):883-884. [doi: 10.1093/oxfordjournals.eurheartj.a015368] [Medline: 9183571]
3. Dwyer T, Magnus P, Olsen J. In the Aftermath of the National Children's Study: is large birth cohort data still a priority? JAMA Pediatr 2017 Mar 1;171(3):214-215. [doi: 10.1001/jamapediatrics.2016.3968] [Medline: 28068436]
4. Eapen ZJ, Turakhia MP, McConnell MV, et al. Defining a mobile health roadmap for cardiovascular health and disease. J Am Heart Assoc 2016 Jul 12;5(7):e003119. [doi: 10.1161/JAHA.115.003119] [Medline: 27405809]
5. Dwyer T, Sun C, Magnussen CG, et al. Cohort Profile: the international childhood cardiovascular cohort (i3C) consortium. Int J Epidemiol 2013 Feb;42(1):86-96. [doi: 10.1093/ije/dys004] [Medline: 22434861]
6. Jacobs DR, Sinaiko AR, Woo JG. Childhood risk factors and adult cardiovascular events. Reply. N Engl J Med 2022 Aug 4;387(5):473-474. [doi: 10.1056/NEJMc2208135] [Medline: 35921464]
7. Sinaiko AR, Jacobs DR Jr, Woo JG, et al. The International Childhood Cardiovascular Cohort (i3C) consortium outcomes study of childhood cardiovascular risk factors and adult cardiovascular morbidity and mortality: Design and recruitment. Contemp Clin Trials 2018 Jun;69:55-64. [doi: 10.1016/j.cct.2018.04.009] [Medline: 29684544]
8. LipidResearchClinics. Plasma lipid distributions in selected North American populations: the Lipid Research Clinics Program Prevalence Study. The lipid research clinics program epidemiology committee. Circulation 1979 Aug;60(2):427-439. [doi: 10.1161/01.CIR.60.2.427]
9. Edwards NM, Kalkwarf HJ, Woo JG, Khoury PR, Daniels SR, Urbina EM. Child physical activity associations with cardiovascular risk factors differ by race. Pediatr Exerc Sci 2016 Aug;28(3):397-406. [doi: 10.1123/pes.2015-0189]
10. Leonard KS, Pauley AM, Guo P. Feasibility and user acceptability of Breezing™, a mobile indirect calorimetry device, in pregnant women with overweight or obesity. Smart Health (Amst) 2023. [doi: 10.1016/j.smhl.2022.100372]
11. Tarakji KG, Wazni OM, Callahan T, et al. Using a novel wireless system for monitoring patients after the atrial fibrillation ablation procedure: the iTransmit study. Heart Rhythm 2015 Mar;12(3):554-559. [doi: 10.1016/j.hrthm.2014.11.015] [Medline: 25460854]
12. Lau JK, Lowres N, Neubeck L, et al. iPhone ECG application for community screening to detect silent atrial fibrillation: a novel technology to prevent stroke. Int J Cardiol 2013 Apr 30;165(1):193-194. [doi: 10.1016/j.ijcard.2013.01.220] [Medline: 23465249]
13. Pengel LKD, Robbers-Visser D, Groenink M, et al. A comparison of ECG-based home monitoring devices in adults with CHD. Cardiol Young 2023 Jul;33(7):1129-1135. [doi: 10.1017/S1047951122002244] [Medline: 35844104]
14. Wegner FK, Kochhäuser S, Ellermann C, et al. Prospective blinded evaluation of the smartphone-based AliveCor Kardia ECG monitor for atrial fibrillation detection: The PEAK-AF study. Eur J Intern Med 2020 Mar;73(72-5):72-75. [doi: 10.1016/j.ejim.2019.11.018] [Medline: 31806411]
15. Mekhael M, Ho C, Noujaim C, et al. Compliance challenges in a longitudinal COVID-19 cohort using wearables for continuous monitoring: observational study. J Med Internet Res 2023 Apr 5;25:e43134. [doi: 10.2196/43134] [Medline: 36763647]
16. Goldberg EM, Levy PD. New approaches to evaluating and monitoring blood pressure. Curr Hypertens Rep 2016 Jun;18(6):27137524. [doi: 10.1007/s11906-016-0650-9] [Medline: 27137524]
17. Støve MP, Graversen AH, Sørensen J. Assessment of noninvasive oxygen saturation in patients with COPD during pulmonary rehabilitation: smartwatch versus pulse oximeter. Respir Care 2023 Aug;68(8):1041-1048. [doi: 10.4187/respcare.10760] [Medline: 37193599]
18. Campo D, Khettab H, Yu R, et al. Measurement of aortic pulse wave velocity with a connected bathroom scale. Am J Hypertens 2017 Sep 1;30(9):876-883. [doi: 10.1093/ajh/hpx059] [Medline: 28520843]
19. Shcherbina A, Mattsson CM, Waggott D, et al. Accuracy in wrist-worn, sensor-based measurements of heart rate and energy expenditure in a diverse cohort. J Pers Med 2017 May 24;7(2):28538708. [doi: 10.3390/jpm7020003] [Medline: 28538708]
20. Wu C, Fritz H, Bastami S, et al. Multi-modal data collection for measuring health, behavior, and living environment of large-scale participant cohorts. Gigascience 2021 Jun 21;10(6):giab044. [doi: 10.1093/gigascience/giab044] [Medline: 34155505]
21. Siniarski A, Sobieraj P, Samel-Kowalik P, Sińska B, Milewska M, Bzikowska-Jura A. Nutrition-related mobile applications - Should they be used for dietary prevention and treatment of cardiovascular diseases? Nutr Metab Cardiovasc Dis 2022 Nov;32(11):2505-2514. [doi: 10.1016/j.numecd.2022.07.010] [Medline: 36064687]
22. Harmon DM, Noseworthy PA, Yao X. The digitization and decentralization of clinical trials. Mayo Clin Proc 2023 Oct;98(10):1568-1578. [doi: 10.1016/j.mayocp.2022.10.001] [Medline: 36669937]
23. Steinhubl SR, Waalen J, Edwards AM, et al. Effect of a home-based wearable continuous ECG monitoring patch on detection of undiagnosed atrial fibrillation: the mSToPS randomized clinical trial. JAMA 2018 Jul 10;320(2):146-155. [doi: 10.1001/jama.2018.8102] [Medline: 29998336]
24. Materia FT, Smyth JM. Acceptability of intervention design factors in mHealth intervention research: experimental factorial study. JMIR Mhealth Uhealth 2021 Jul 26;9(7):e23303. [doi: 10.2196/23303] [Medline: 34309563]

XSL•FO
RenderX

Abbreviations

BHS: Bogalusa Heart Study
DXA: dual-energy x-ray absorptiometry
I3C: International Childhood Cardiovascular Cohort Consortium
IRB: institutional review board
mHealth: mobile health
MRI: magnetic resonance imaging
NGHS: National Heart, Lung, and Blood Institute Growth and Health Study
NHLBI: National Heart, Lung, and Blood Institute
PLRS: Princeton Lipid Research Study
PWV: pulse wave velocity

Edited by K Adapa; submitted 09.01.25; peer-reviewed by B Senst, J Walsh; revised version received 16.04.25; accepted 17.04.25; published 18.06.25.

Please cite as:
Russell M, Cain E, Bazzano L, De Anda I, Woo JG, Urbina EM
Collecting at-Home Biometric Measures for Longitudinal Research From the i3C: Feasibility and Acceptability Study
JMIR Hum Factors 2025;12:e71103
URL: https://humanfactors.jmir.org/2025/1/e71103
doi:10.2196/71103

© Marta Russell, Erin Cain, Lydia Bazzano, Ileana De Anda, Jessica G Woo, Elaine M Urbina. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 18.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Assessing Postoperative Pain in Patients Who Underwent Total Knee Arthroplasty Using an Automated Self-Logging Patient-Reported Outcome Measure Collection Device: Retrospective Cohort Study

Prabjit Ajrawat[1], BKin, MSc; Blaine Price[2], BSc, MSc; Daniel Gooch[2], BSc, MSc, PhD; Rudolf Serban[3], MBBS; Ruqaiya Al-Habsi[4], MBBS, MSurg, MSc, MRCS; Oliver Pearce[3], MBBS

[1]University of Buckingham Medical School, Buckingham, United Kingdom

[2]School of Computing and Communications, The Open University, Jenny Lee Building, Walton Hall, Milton Keynes, United Kingdom

[3]Trauma and Orthopaedics Department, Milton Keynes University Hospital, Milton Keynes, United Kingdom

[4]Trauma and Orthopaedics Department, Barts Health NHS Trust, London, United Kingdom

Corresponding Author:
Daniel Gooch, BSc, MSc, PhD
School of Computing and Communications, The Open University, Jenny Lee Building, Walton Hall, Milton Keynes, United Kingdom

Abstract

Background: Patient-reported outcome measures (PROMs) are tools for assessing symptoms and the quality of care. Despite their growing use, conventional data collection methods limit widespread PROM implementation. In orthopedics, pain is a frequent patient complaint and a common PROM, especially following total knee arthroplasty (TKA). Although TKA is generally successful, some patients still report postoperative pain, potentially due to tourniquet use. Using an improved PROM data-gathering technique may help to address tourniquet use during a TKA procedure and its impact on postoperative pain. The PainPad, an automated self-logging device, was developed to capture patient pain levels accurately.

Objective: The aim of the study is to assess the feasibility and effectiveness of the PainPad device in quantifying in-hospital postoperative pain following TKA with or without tourniquet use.

Methods: A retrospective study with 234 patients who underwent TKA from 2018 to 2021 at Milton Keynes University Hospital was conducted. Patients were categorized as receiving TKA with an intraoperative tourniquet (tourniquet group) or TKA without a tourniquet (nontourniquet group). Postoperative pain during the first 24 hours was self-reported every 2 hours using the PainPad device. From both groups, data on hospital length of stay, total tourniquet time, and the presence of postoperative deep vein thrombosis were also collected.

Results: There were 115 TKAs with tourniquets (72/115, 62.6% female patients; mean age 69.26, SD 9.93 years) and 119 TKAs without tourniquets (91/119, 76.4% female patients; mean age 70.97, SD 9.01 years). When assessing 24-hour mean postoperative pain scores, the PainPad device data indicated no significant difference (P=.53; 95% CI −0.76 to 0.39) between the tourniquet (mean pain score 3.31, SD 2.34) and nontourniquet groups (mean pain score 3.12, SD 2.15). There was no correlation between tourniquet times and the pain scores retrieved from the PainPad device. A subgroup analysis comparing longer (>90 minutes) versus shorter (<90 minutes) tourniquet times showed no significant difference in terms of pain and length of stay.

Conclusions: The PainPad device is a feasible and effective method for collecting and evaluating in-hospital postoperative pain following TKA, allowing for the quantification of individual pain levels. This study aligns with the current health care trend toward leveraging innovative technologies and personalized data to enhance patient-centered care.

(JMIR Hum Factors 2025;12:e65271) doi:10.2196/65271

KEYWORDS

electronic patient-reported outcome measure; total knee arthroplasty; tourniquet; PainPad; knee arthroplasty; automated self-logging; collection device; retrospective cohort study; patient-reported outcome; quality of care; assessment; data collection; postoperative; feasibility; effectiveness; electronic patient report; patient-centered care; innovative technology

XSL•FO
RenderX

Introduction

Patient-reported outcome measures (PROMs) are tools that have become increasingly used to systematically assess symptom progression, satisfaction, and the overall quality of care provided from the patient's perspective [1,2]. Originally designed for research purposes, incorporating PROMs into clinical care represents a transition toward a more evidence-based and patient-centered health care system [3]. Implementing PROMs enables patients to quantify their subjective experience and allows physicians to monitor symptoms and make informed treatment decisions when combined with other objective measures.

Among surgical specialties, PROMs have been most extensively administered in orthopedics [4]. Since April 2009, the National Health Service (NHS) mandated all publicly funded health care providers of inpatient departments to collect PROMs for certain elective procedures, including hip and knee arthroplasties [5]. Despite the increasing use and evidence supporting PROMs, widespread implementation still faces several challenges among orthopedic practices. The additional time, cost, and administrative burden along with the risks of low patient participation, data errors, and incorrect interpretation have been frequently cited as limitations in integrating PROMs into surgical practices [3,5,6]. These limitations may be attributed to the current inefficiencies of conventional paper-based data collection methods. As initiatives strive to sustain PROM use, innovative data collection technologies tailored to the complexities of surgical workflows will be needed to alleviate these logistic and administrative constraints.

Within orthopedic practices, in-hospital postoperative pain is a regularly assessed PROM following surgical interventions, as it is a frequently cited patient complaint [7,8]. While some PROMs can be objectively measured, pain is inherently subjective and difficult to accurately capture. Current methods, such as patient pain diaries, are often prone to inaccuracies due to illegible handwriting, incomplete data, and recall bias [9-11]. Additionally, distributing, collecting, and processing paper-based PROM methods require significant administrative efforts and costs [12]. The manual data handling and entry further increase the risk of transcription errors, compromising the reliability of the collected data. The physical forms of paper-based methods also require substantial storage space, which can limit the timely retrieval of patient data [12]. The emergence of electronic devices and platforms presents an alternative approach to collecting pain-related data. Evidence suggests that electronic data collection methods may be more accurate, provide greater compliance, and yield fewer errors compared to conventional methods [13,14].

Even with total knee arthroplasty (TKA) being one of the most successful orthopedic procedures, 20% of patients were unsatisfied due to their postoperative pain levels [15]. A potential cause for this pain may be due to tourniquet use during a TKA procedure; however, this topic remains controversial among orthopedic surgeons. However, by using an improved PROM data-gathering technique, it may help to address the conflicting topic of tourniquet use during a TKA procedure and its impact on postoperative pain.

The PainPad is a handheld automated self-logging device that was developed to capture patient pain levels [11]. While other electronic pain collection methods have focused on remote monitoring in outpatient settings, the PainPad device was designed for inpatient hospital use to provide instantaneous postsurgical pain evaluation [11]. Thus, to showcase the clinical utility and effectiveness of assessing postoperative pain, the purpose of this study was to use the PainPad device to determine if tourniquet use during TKA has a detrimental effect on postoperative pain levels in the context of a modern multimodal pain pathway and early mobilization.

Methods

Ethical Considerations

This research was approved under Integrated Research Application System (229503) and the Open University Human Research Ethics Committee (17/LO/1404). Informed consent was provided by participants, and data were deidentified.

Study Design and Patients

This retrospective cohort study was conducted at Milton Keynes University Hospital (MKUH), United Kingdom, and reviewed 234 patients who underwent TKA. Patients undergoing a primary unilateral TKA between 2018 and 2021 at MKUH were included in the study. Any patients undergoing revision TKA were excluded.

Surgical Procedure and Postoperative Care

All cases were performed by 7 arthroplasty consultants at MKUH, and all patients underwent the multimodal analgesia program protocol for their inpatient rehabilitation. This modern multimodal analgesia protocol was specifically designed to enhance the recovery process, decrease the amount and frequency of opioid use, and reduce the length of stay (LOS) [16]. All patients were assessed preoperatively by the physiotherapy team to collect PROMs and range of movement.

Patients were categorized as receiving TKA with an intraoperative tourniquet (tourniquet group) or TKA without a tourniquet (nontourniquet group). For patients receiving the tourniquet, a conical tourniquet (85 cm long×8.5 cm wide) was inflated to 300 mm Hg and kept inflated from the beginning of the operation to the end of the procedure. This duration was considered the tourniquet time. Aside from tourniquet use, all TKAs were performed through the same standard surgical approach, and all patients received the same postoperative rehabilitation protocol.

The PainPad Device

Patients reported their postoperative pain levels using an automated self-logging device named "PainPad" (Figure 1). A previous study has outlined PainPad's design process and user testing with patients and hospital staff [11]. The handheld PainPad device (9.5×6.5×3 cm) consists of a large keypad allowing patients to rate their pain on a scale of 0 - 10 (0=no pain and 10=worst pain imaginable), mimicking the visual

XSL•FO
RenderX

analog scale (VAS). Patients were shown how to enter their pain level using the PainPad device. The physical box-like design of the PainPad allowed for easy and quick disinfectant and sanitary practices by hospital personnel. The PainPad device consists of 2 LEDs and a beeper to notify patients to enter their pain score every 2 hours to ensure compliance with pain logging [11]. These pain scores were then linked to the patient via their unique identifier. The combination of PainPad's pain database and the nurse's entered pain scores on the electronic patient record was combined to gain a maximal number of pain scores for each inpatient episode for analysis.

Figure 1. The PainPad device (9.5×6.5×3 cm) is a 3D-printed rectilinear box made of acrylonitrile butadiene styrene (ABS) plastic. It features a red and green surface-mounted LED light on the internal circuit board, which projects through light tubes on the surface, and a 3×4 membrane keypad recessed into the front face. Beneath the ABS shell is a 70-dB speaker that emits sounds prompting the user to self-report their pain level by pressing a button (0 - 10), which beeps when pressed. The device also plays distinct sounds to signal success or failure when clinicians assign it to a patient.



Outcome Measures

The primary outcome of this study was the self-reported 24-hour postoperative pain scores (mean, maximum, and minimum) on the PainPad device in both tourniquet and nontourniquet groups. Secondary outcomes included measuring hospital LOS, total tourniquet time, and the presence of postoperative deep vein

XSL•FO
RenderX

thrombosis (DVT) between both tourniquet and nontourniquet groups.

Statistical Analysis

Descriptive statistics were used to analyze patient demographics and general trends in outcome measures. Two-tailed *t* tests were used to determine the differences in patient demographics, 24-hour postoperative pain scores from the PainPad device, and hospital LOS between tourniquet and nontourniquet patients. Additionally, 2-tailed *t* tests were used to assess PainPad pain data and hospital LOS between a subgroup of patients with longer (>90 minutes) or shorter (<90 minutes) tourniquet times. The Spearman rank correlation was used to compare the tourniquet time with PainPad pain scores and hospital LOS. For our statistical analysis, the 2-tailed *t* test was suitable in comparing 2 groups of patients (ie, with and without tourniquet), with a Spearman rank determining a correlation between tourniquet use and pain scores. Our data met the criteria for both the 2-tailed *t* test (independence of observations, approaching normal distribution, and homogeneity of variances assessed through the Levene test) and Spearman rank (monotonic, paired, and interval data). All statistical testing was conducted using the SPSS software (version 21; IBM Corp).

Results

A total of 234 consecutive patients who underwent TKA (n=163, 70% female patients) were included (mean age 69.26, SD 9.93 years) and 119 underwent TKA without a tourniquet (91/119, 76.4% female patients; mean age 70.97, SD 9.01 years). Regarding patient demographics, there was a significant difference in female sex (*P*=.02) favoring the nontourniquet group (Table 1). The median hospital LOS was 2 (IQR 1 - 18) days, and there was no significant difference in hospital LOS between the 2 groups (*P*=.22; 95% CI −0.24 to 1.04). From the cohort, only 1 of 234 patients experienced a DVT event, which was from the nontourniquet group.

Table . Characteristics of included patients who underwent total knee arthroplasty using the PainPad device in the nontourniquet group and tourniquet groups.

Patient characteristics	Nontourniquet (n=119)	Tourniquet (n=115)	*P* value
Sex, n (%)			.02
Male	28 (23.5)	43 (37.3)	
Female	91 (76.4)	72 (62.6)	
Age (years), mean (SD)	70.97 (9.01)	69.26 (9.93)	.17

The PainPad device was used to assess patient-reported postoperative pain between the 2 groups. When assessing 24-hour mean postoperative pain scores, the PainPad device data indicated that there was no significant difference between the tourniquet (mean pain score 3.31, SD 2.34) and nontourniquet group (mean pain score 3.12, SD 2.15; *P*=.53; 95% CI −0.76 to 0.39). The PainPad device data further suggested no significant difference in the maximum pain score 24 hours postoperatively between the tourniquet (mean maximum pain score 6.34, SD 2.81) and nontourniquet group (mean maximum pain score 6.12, SD 2.87; *P*=.55; 95% CI −0.95 to 0.51). There was also no significant difference in the minimum pain score 24 hours postoperatively between the tourniquet (mean minimum pain score 0.66, SD 1.72) and nontourniquet (mean minimum pain score 1.03, SD 2.22; *P*=.15; 95% CI −0.89 to 0.13).

Tourniquet time was recorded from 110 of 115 patients with a median time of 73 (IQR 19 - 120) minutes. There was no correlation between tourniquet time and hospital LOS (Spearman ρ=−0.09; *P*=.38). When correlating tourniquet time with 24-hour postoperative pain scores retrieved from the PainPad device, there were no correlations for the mean (Spearman ρ=−0.01; *P*=.90), maximum (Spearman ρ=−0.18; *P*=.07), and minimum pain scores (Spearman ρ=0.12; *P*=.20). In correlating the VAS with 24-hour postoperative mean pain scores, there was a very low correlation (*r*=0.139) and an extremely low correlation with maximum pain score 24 hours postoperatively (*r*=0.078). This indicates that preoperative and postoperative pain were not correlated. A subgroup analysis of patients with longer tourniquet times (>90 minutes) and shorter tourniquet times (<90 minutes) was also evaluated. From this, there were no significant differences between longer and shorter tourniquet times in terms of hospital LOS and self-reported pain data as collected from the PainPad device (Table 2).

Table . Comparison of recorded PainPad postoperative pain levels and hospital length of stay (LOS) between long duration (>90 minutes) and short duration (<90 minutes) tourniquet groups.

Outcomes	Long duration tourniquet (n=91), mean (SD)	Short duration tourniquet (n=19), mean (SD)	*P* value	95% CI
Minimum pain	0.86 (2.01)	2.16 (3.04)	.09	−2.81 to 0.21
Maximum pain	6.29 (2.87)	6.37 (2.79)	.91	−1.54 to 1.37
Mean pain	3.14 (2.23)	4.00 (2.86)	.23	−2.30 to 0.58
Hospital LOS (days)	2.74 (2.49)	2.53 (1.81)	.67	−0.49 to 1.21

XSL•FO
RenderX

Discussion

Principal Findings

PROMs represent a standardized method for monitoring patient symptoms. The value of monitoring patients' self-reported pain levels during the first hours in the postoperative period is essential to ensure optimal analgesia and patient satisfaction [17,18]. This study provides evidence that the PainPad, an automated self-logging PROM device, is feasible in collecting pain levels after a major orthopedic surgical intervention. Since our institution uses a modern multimodal analgesia protocol to enhance the recovery process and limit the use of postoperative opioids, the PainPad device provided surgeons with real-time pain analytics within the 24-hour postoperative period to ensure that appropriate analgesia and prompt intervention were administered.

PROMs have become increasingly important in the context of orthopedics, as many surgical interventions aim to improve subjective patient outcomes such as pain. Incorporating PROMs, such as pain logging, into routine care can harness critical patient information to improve clinical decision-making and overall support patient-centered care. Generally, self-reported pain logging in an outpatient clinic setting is completed by paper-based methods in the form of questionnaires, forms, or patient diaries [10]. In contrast, during an inpatient hospital stay, nursing staff typically collect pain levels at regular intervals through monitoring or by having patients complete a numerical pain rating scale [19-21]. However, these methods are often time-consuming and impose additional burdens on clinical staff [11]. In some instances, staffing limitations can sometimes prevent sufficiently frequent pain log entries [11]. Additionally, studies have indicated that pain levels recorded by nursing staff are often incomplete and potentially inaccurate, as patients may feel reluctant to fully express their pain [20,22]. Orthopedic surgeons have also expressed various logistical and technical challenges in collecting pain levels and other PROMs. For instance, surgeons indicated that the additional administrative tasks involved in data collection and the prevalence of missing data along with poor visual display and the complexity of completing PROMs hindered their ability to administer PROMs to their patients [6]. Collectively, these limitations of conventional paper-based methods coupled with the earlier-mentioned barriers further emphasize the need for innovative technologies to seamlessly integrate self-logging of pain and PROMs.

Currently, the NHS PROM program administers paper-based collection methods, incurring an annual cost of US $1.1 million [23]. However, as the NHS transitions toward a paperless system and the PROM program expands, innovative technologies will be required to reduce costs and streamline data administration [24]. Studies have shown that electronic patient-reported outcome measure (ePROM) tools and devices are equivalent to their original paper-based counterpart [25,26]. With regard to paper-based methods, poor response rate is a limiting factor in assessing PROMs [27,28]. Yet, administrating ePROM tools has been shown to circumvent poor response rates, as automation permits outcomes to be collected at the scheduled intervals. Gurland et al [29] showed that administering ePROMs on a tablet to 103 patients who are undergoing surgery resulted in a 96% response rate in comparison to 25% with a paper-based collection method. Furthermore, ePROMs have the potential to enhance patient-physician communication by offering real-time pain tracking and avoiding possible recall bias [30]. The practicality of ePROM modalities, along with their advantages of automated data capture, lower long-term costs, faster completion times, and reduced administrative errors [12], demonstrates the potential role of ePROM tools within the NHS to ensure widespread data capturing.

Most self-logging pain and ePROM technologies are designed for remote or at-home monitoring. In contrast, the PainPad was developed as an automated device to enable patients to provide their pain ratings during their hospital stay. Previous research with the PainPad showed that it improved the frequency and compliance of self-reported pain logging among inpatients recovering from ambulatory orthopedic surgery compared to pain scores reported by nursing staff [11]. The PainPad offers a convenient user experience for both patients and staff. It was designed with patient feedback, ensuring an easy-to-use interface. In one study, authors noted that 29% of patients had difficulties completing the paper-based VAS form due to visual impairment and physical restriction [31].

The PainPad device has several advantages over other pain assessment methods in a hospital setting. Paper-based or verbal pain assessments can be limited by patient fatigue, cognitive impairment, and verbal communication difficulties and often require manual transcription that could lead to errors [32,33]. Mobile apps and tablets often require patients to navigate various user interfaces and provide typed or touched responses on a touchscreen that may be difficult for older patients, those with dexterity issues, or those experiencing severe postsurgical pain [34]. Other mobile apps such as the PainChek app use facial expressions for pain assessment; however, these tend to be used in long-term care settings for patients with moderate to severe dementia. Unlike other ePROM methods, the lightweight PainPad features larger buttons that make it accessible for those with mobility or vision issues, and the visual cues it provides minimize response errors and ensure accurate data capture. Additionally, the physical device allows patients who are unable to verbalize their pain to report it accurately. Incorporating tactile, auditory, and visual senses into the device is believed to minimize typical patient errors associated with paper-based or single-dimensional methods, thereby providing a more standardized assessment of individualized pain levels. This user-friendly interface minimizes the risk of noncompliance while ensuring immediate, real-time accurate pain logging, making it more suitable for hospital settings where rapid pain assessment is critical. Given the importance of hospital infection control, the device was designed for easy disinfection within the hospital environment and to seamlessly integrate into clinical workflows [11]. However, a drawback of the PainPad is that it has not been assessed in postsurgical pediatric patients, where adaptations may be required.

In our study, most were female participants, and research has shown that pain perception differs between sexes. Female participants often experience higher pain sensitivity and report

XSL•FO
RenderX

higher postoperative pain scores compared to male participants following TKA [35]. Additionally, patient age may also influence pain perception, as aging reduces pain sensitivity for lower pain intensities [36]. These sex and age-based differences in pain perception may influence how patients interact with and report pain using the PainPad device. Given these differences in pain perception, it is ideal for patients to self-report their pain at its most intense moments using accessible and patient-friendly tools. Prior research suggests that older patients might prefer alternative methods for self-logging their PROMs or struggle with handheld technologies [12,37,38]. Research has shown that younger patients have a greater preference for ePROM methods [12,38-40]. However, the mean age in our study was 69.2 years, indicating that older patients can effectively use this device. When comparing the PainPad interface to 2 tablet-based alternatives, older adults preferred the tangible PainPad interface for reporting their pain in the hospital setting [11].

In this study, we compared tourniquet use in TKA to showcase the clinical utility and feasibility of the PainPad device in collecting pain levels. Using tourniquets in TKA remains a topic of considerable debate. Some orthopedic surgeons advocate for their use, citing advantages such as improved surgical field visualization, enhanced cementation through increased interdigitation, reduced intraoperative blood loss, and shorter operative times [41]. On the other hand, studies have indicated a slightly elevated risk of DVT, quadriceps weakness, nerve injury, decreased range of movement, higher transfusion rate, longer LOS, and increased postoperative pain [42-49].

Our results suggest that the use of a tourniquet during TKA procedures has no impact on the postoperative VAS score in the context of a modern multimodal pain pathway with early mobilization. Furthermore, there was no significant difference between patients with a tourniquet for <90 minutes and those with a tourniquet for >90 minutes with regard to postoperative pain and hospital LOS. These findings contrast previous reports that demonstrated increased pain among patients who underwent TKA with a tourniquet [50,51]. However, this can be attributed to the predate use of modern multimodal analgesia protocols and early mobilization pathways, which aim to reduce postoperative pain levels and hospital LOS, increase patient mobility, and limit opioid analgesia use [52]. Moreover, the absence of precise PROM data collection methods may have impeded and further exacerbated the true efficacy of this technique. By providing patients with the self-logging PainPad device, we were able to gather sufficient data to demonstrate that tourniquet use is associated with positive patient outcomes, including effective pain management. By making a digital self-logging of PROMs more accessible, we suggest that better evidence will be available to improve other surgical and clinical procedures.

Limitations

There are limitations to this study that need to be considered. First, this was a retrospective study in a single academic center with 234 patients who underwent TKA, which limits the generalizability of our findings. The retrospective nature of our study may have also potentially introduced inherent selection, recall, and ascertainment biases, as the patient inclusion was based on available data from our institution's database. However, since the PainPad device collects pain scores in an instantaneous manner, this would ideally limit recall bias. Furthermore, our statistical analysis indicated no significant differences in major baseline confounding variables between groups.

Second, patient comorbidities and preoperative pain data were not collected, which may influence postoperative pain levels. Third, TKA procedures were completed by 7 orthopedic consultants, which may contribute to discrepancies in operative technique. However, all surgeons followed the same postoperative pain management protocol. Finally, this study only assessed 24-hour postoperative pain; however, within the 24 hours, there were multiple self-reported pain entries from patients (ie, completed bihourly) using the PainPad device. Further, large-scale studies that include various acute and chronic patient populations with diverse demographics (ie, pediatric patients) along with appropriate control groups will help ensure the generalizability of the results. Additionally, it will be important to directly evaluate the cost-effectiveness of using the PainPad device in a clinical setting.

Conclusions

The PainPad device is a feasible and effective method for collecting and evaluating in-hospital postoperative pain following TKA, allowing for precise quantification of individual pain levels. This study aligns with the current health care trend toward leveraging innovative technologies and personalized data to enhance patient-centered care.

Conflicts of Interest

OP and BP hold a patent on the PainPad device used for data collection. Neither OP nor BP have received compensation in any form for the use of the device in this study or any other setting. DG is an associate editor for *JMIR Human Factors* at the time of this publication.

References

1. Chow A, Mayer EK, Darzi AW, Athanasiou T. Patient-reported outcome measures: the importance of patient satisfaction in surgery. Surgery 2009 Sep;146(3):435-443. [doi: 10.1016/j.surg.2009.03.019] [Medline: 19715800]
2. Kotronoulas G, Kearney N, Maguire R, et al. What is the value of the routine use of patient-reported outcome measures toward improvement of patient outcomes, processes of care, and health service outcomes in cancer care? A systematic review of controlled trials. J Clin Oncol 2014 May 10;32(14):1480-1501. [doi: 10.1200/JCO.2013.53.5948] [Medline: 24711559]

XSL•FO
RenderX

3. Driscoll B, Leonard LD, Kovar A, et al. Surgeon perceptions of the integration of patient-reported outcome measures into clinical practice. J Surg Res 2022 Dec;280:486-494. [doi: 10.1016/j.jss.2022.07.038] [Medline: 36067535]
4. Briffa N. The employment of patient-reported outcome measures to communicate the likely benefits of surgery. Patient Relat Outcome Meas 2018;9(263-6):263-266. [doi: 10.2147/PROM.S132746] [Medline: 30147385]
5. Black N. Patient reported outcome measures could help transform healthcare. BMJ 2013 Jan 28;346:f167. [doi: 10.1136/bmj.f167] [Medline: 23358487]
6. Whitebird RR, Solberg LI, Ziegenfuss JY, et al. What do orthopaedists believe is needed for incorporating patient-reported outcome measures into clinical care? A qualitative study. Clin Orthop Relat Res 2022 Apr 1;480(4):680-687. [doi: 10.1097/CORR.0000000000002059] [Medline: 34846308]
7. Arefayne NR, Tegegne SS, Gebregzi AH, Mustofa SY. Incidence and associated factors of post-operative pain after emergency orthopedic surgery: a multi-centered prospective observational cohort study. International Journal of Surgery Open 2020;27:103-113. [doi: 10.1016/j.ijso.2020.10.003]
8. Edgley C, Hogg M, De Silva A, Braat S, Bucknill A, Leslie K. Severe acute pain and persistent post-surgical pain in orthopaedic trauma patients: a cohort study. Br J Anaesth 2019 Sep;123(3):350-359. [doi: 10.1016/j.bja.2019.05.030] [Medline: 31248645]
9. Schneider S, Stone AA, Schwartz JE, Broderick JE. Peak and end effects in patients' daily recall of pain and fatigue: a within-subjects analysis. J Pain 2011 Feb;12(2):228-235. [doi: 10.1016/j.jpain.2010.07.001] [Medline: 20817615]
10. Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient non-compliance with paper diaries. BMJ 2002 May 18;324(7347):1193-1194. [doi: 10.1136/bmj.324.7347.1193] [Medline: 12016186]
11. Price BA, Kelly R, Mehta V, McCormick C, Ahmed H, Pearce O. Feel my pain: design and evaluation of painpad, a tangible device for supporting inpatient self-logging of pain. 2018 Presented at: CHI '18: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems; Apr 21-26, 2018; Montreal, Quebec, Canada. [doi: 10.1145/3173574.3173743]
12. Meirte J, Hellemans N, Anthonissen M, et al. Benefits and disadvantages of electronic patient-reported outcome measures: systematic review. JMIR Perioper Med 2020 Apr 3;3(1):e15588. [doi: 10.2196/15588] [Medline: 33393920]
13. Jamison RN, Raymond SA, Levine JG, Slawsby EA, Nedeljkovic SS, Katz NP. Electronic diaries for monitoring chronic pain: 1-year validation study. Pain 2001 Apr;91(3):277-285. [doi: 10.1016/S0304-3959(00)00450-4] [Medline: 11275385]
14. Jibb LA, Khan JS, Seth P, et al. Electronic data capture versus conventional data collection methods in clinical pain studies: systematic review and meta-analysis. J Med Internet Res 2020 Jun 16;22(6):e16480. [doi: 10.2196/16480] [Medline: 32348259]
15. Scott CEH, Howie CR, MacDonald D, Biant LC. Predicting dissatisfaction following total knee replacement: a prospective study of 1217 patients. J Bone Joint Surg Br 2010 Sep;92(9):1253-1258. [doi: 10.1302/0301-620X.92B9.24394] [Medline: 20798443]
16. Parkes RJ, Ayeko O, Brunton L, Griffiths-Jones W, Ungvari Z, Goss H. Revolutionising rapid recovery: a quality improvement project in hip and knee replacement. BMJ Open Qual 2021 Apr;10(2):33879494. [doi: 10.1136/bmjoq-2020-001249] [Medline: 33879494]
17. Goel S, Deshpande SV, Jadawala VH, Suneja A, Singh R. A comprehensive review of postoperative analgesics used in orthopedic practice. Cureus 2023 Nov;15(11):e48750. [doi: 10.7759/cureus.48750] [Medline: 38094554]
18. Gordon DB, de Leon-Casasola OA, Wu CL, Sluka KA, Brennan TJ, Chou R. Research gaps in practice guidelines for acute postoperative pain management in adults: findings from a review of the evidence for an American Pain Society clinical practice guideline. J Pain 2016 Feb;17(2):158-166. [doi: 10.1016/j.jpain.2015.10.023] [Medline: 26719073]
19. Choinière M, Melzack R, Girard N, Rondeau J, Paquin MJ. Comparisons between patients' and nurses' assessment of pain and medication efficacy in severe burn injuries. Pain 1990 Feb;40(2):143-152. [doi: 10.1016/0304-3959(90)90065-L] [Medline: 2308761]
20. de Rond ME, de Wit R, van Dam FS, Muller MJ. A pain monitoring program for nurses: effects on communication, assessment and documentation of patients' pain. J Pain Symptom Manage 2000 Dec;20(6):424-439. [doi: 10.1016/s0885-3924(00)00209-8] [Medline: 11131261]
21. Puntillo K, Neighbor M, O'Neil N, Nixon R. Accuracy of emergency nurses in assessment of patients' pain. Pain Manag Nurs 2003 Dec;4(4):171-175. [doi: 10.1016/s1524-9042(03)00033-x] [Medline: 14663795]
22. Francke AL, Theeuwen I. Inhibition in expressing pain. A qualitative study among Dutch surgical breast cancer patients. Cancer Nurs 1994 Jun;17(3):193-199. [Medline: 8055489]
23. National patient reported outcome measures (PROMs) programme consultation. NHS England. 2016. URL: https://www.engage.england.nhs.uk/consultation/proms-programme [accessed 2025-05-08]
24. Cutting reliance on paper will make patients safer, says NHS england. NHS England. 2015. URL: https://www.england.nhs.uk/2015/09/cutting-reliance-on-paper [accessed 2025-05-08]
25. Gwaltney CJ, Shields AL, Shiffman S. Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: a meta-analytic review. Value Health 2008;11(2):322-333. [doi: 10.1111/j.1524-4733.2007.00231.x] [Medline: 18380645]

XSL•FO
RenderX

26. VanDenKerkhof EG, Goldstein DH, Blaine WC, Rimmer MJ. A comparison of paper with electronic patient-completed questionnaires in a preoperative clinic. Anesth Analg 2005 Oct;101(4):1075-1080. [doi: 10.1213/01.ane.0000168449.32159.7b] [Medline: 16192524]
27. Using electronic patient-reported outcome measures (PROMs) to improve patient outcomes. NHS England. 2017. URL: https://transform.england.nhs.uk/key-tools-and-info/digital-playbooks/rheumatology-digital-playbook/using-electronic-patient-reported-outcome-measures-proms-to-improve-patient-outcomes [accessed 2025-05-08]
28. Gibbs D, Toop N, Grossbach AJ, et al. Electronic versus paper patient-reported outcome measure compliance rates: a retrospective analysis. Clin Neurol Neurosurg 2023 Mar;226(107618):107618. [doi: 10.1016/j.clineuro.2023.107618] [Medline: 36773533]
29. Gurland B, Alves-Ferreira PC, Sobol T, Kiran RP. Using technology to improve data capture and integration of patient-reported outcomes into clinical care: pilot results in a busy colorectal unit. Dis Colon Rectum 2010 Aug;53(8):1168-1175. [doi: 10.1007/DCR.0b013e3181d87468] [Medline: 20628281]
30. Mowlem FD, Elash CA, Dumais KM, et al. Best practices for the electronic implementation and migration of patient-reported outcome measures. Value Health 2024 Jan;27(1):79-94. [doi: 10.1016/j.jval.2023.10.007] [Medline: 37879401]
31. Briggs M, Closs JS. A descriptive study of the use of visual analogue scales and verbal rating scales for the assessment of postoperative pain in orthopedic patients. J Pain Symptom Manage 1999 Dec;18(6):438-446. [doi: 10.1016/s0885-3924(99)00092-5] [Medline: 10641470]
32. Hjermstad MJ, Fayers PM, Haugen DF, et al. Studies comparing Numerical Rating Scales, Verbal Rating Scales, and Visual Analogue Scales for assessment of pain intensity in adults: a systematic literature review. J Pain Symptom Manage 2011 Jun;41(6):1073-1093. [doi: 10.1016/j.jpainsymman.2010.08.016] [Medline: 21621130]
33. Younger J, McCue R, Mackey S. Pain outcomes: a brief review of instruments and techniques. Curr Pain Headache Rep 2009 Feb;13(1):39-43. [doi: 10.1007/s11916-009-0009-x] [Medline: 19126370]
34. Stinson JN, Jibb LA, Nguyen C, et al. Development and testing of a multidimensional iPhone pain assessment application for adolescents with cancer. J Med Internet Res 2013 Mar 8;15(3):e51. [doi: 10.2196/jmir.2350] [Medline: 23475457]
35. Sodhi N, Qilleri A, Aprigliano C, Danoff JR. One size does not fit all: women experience more pain than men after total knee arthroplasty. J Arthroplasty 2025 Apr;40(4):880-886. [doi: 10.1016/j.arth.2024.09.028] [Medline: 39307204]
36. Lautenbacher S, Peters JH, Heesen M, Scheel J, Kunz M. Age changes in pain perception: a systematic-review and meta-analysis of age effects on pain and tolerance thresholds. Neurosci Biobehav Rev 2017 Apr;75:104-113. [doi: 10.1016/j.neubiorev.2017.01.039] [Medline: 28159611]
37. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021 Aug 17;21(1):1556. [doi: 10.1186/s12889-021-11623-w] [Medline: 34399716]
38. Keurentjes JC, Fiocco M, So-Osman C, et al. Hip and knee replacement patients prefer pen-and-paper questionnaires: Implications for future patient-reported outcome measure studies. Bone Joint Res 2013;2(11):238-244. [doi: 10.1302/2046-3758.211.2000219] [Medline: 24203164]
39. Engan HK, Hilmarsen C, Sittlinger S, Sandmæl JA, Skanke F, Oldervoll LM. Are web-based questionnaires accepted in patients attending rehabilitation? Disabil Rehabil 2016 Dec;38(24):2406-2412. [doi: 10.3109/09638288.2015.1129449] [Medline: 26800715]
40. Richter JG, Becker A, Koch T, et al. Self-assessments of patients via Tablet PC in routine patient care: comparison with standardised paper questionnaires. Ann Rheum Dis 2008 Dec;67(12):1739-1741. [doi: 10.1136/ard.2008.090209] [Medline: 18647853]
41. Tai TW, Lin CJ, Jou IM, Chang CW, Lai KA, Yang CY. Tourniquet use in total knee arthroplasty: a meta-analysis. Knee Surg Sports Traumatol Arthrosc 2011 Jul;19(7):1121-1130. [doi: 10.1007/s00167-010-1342-7] [Medline: 21161177]
42. Zak SG, Yeroushalmi D, Long WJ, Meftah M, Schnaser E, Schwarzkopf R. Does the use of a tourniquet influence outcomes in total knee arthroplasty: a randomized controlled trial. J Arthroplasty 2021 Jul;36(7):2492-2496. [doi: 10.1016/j.arth.2021.02.068] [Medline: 33795174]
43. Bhalchandra Londhe S, Vinod Shah R, Sanjay Londhe S, Agrawal PO, Antao NA, Churhe S. Comparison of local pain and tissue reaction between conventional pneumatic tourniquet and disposable silicone ring tourniquet during total knee arthroplasty. J Clin Orthop Trauma 2021 Apr;15(152-5):152-155. [doi: 10.1016/j.jcot.2020.09.005] [Medline: 33717930]
44. Alexandersson M, Wang EY, Eriksson S. A small difference in recovery between total knee arthroplasty with and without tourniquet use the first 3 months after surgery: a randomized controlled study. Knee Surg Sports Traumatol Arthrosc 2019 Apr;27(4):1035-1042. [doi: 10.1007/s00167-018-5196-8] [Medline: 30328495]
45. Dennis DA, Kittelson AJ, Yang CC, Miner TM, Kim RH, Stevens-Lapsley JE. Does tourniquet use in TKA affect recovery of lower extremity strength and function? A randomized trial. Clin Orthop Relat Res 2016 Jan;474(1):69-77. [doi: 10.1007/s11999-015-4393-8] [Medline: 26100254]
46. Kumar N, Yadav C, Singh S, Kumar A, Vaithlingam A, Yadav S. Evaluation of pain in bilateral total knee replacement with and without tourniquet; a prospective randomized control trial. J Clin Orthop Trauma 2015 Jun;6(2):85-88. [doi: 10.1016/j.jcot.2015.01.095] [Medline: 25983513]
47. Abdel-Salam A, Eyres KS. Effects of tourniquet during total knee arthroplasty. A prospective randomised study. J Bone Joint Surg Br 1995 Mar;77(2):250-253. [Medline: 7706340]

XSL•FO
RenderX

48. Dreyer HC. Tourniquet use during knee replacement surgery may contribute to muscle atrophy in older adults. Exerc Sport Sci Rev 2016 Apr;44(2):61-70. [doi: 10.1249/JES.0000000000000076] [Medline: 26829246]
49. Zhao HY, Yeersheng R, Kang XW, Xia YY, Kang PD, Wang WJ. The effect of tourniquet uses on total blood loss, early function, and pain after primary total knee arthroplasty: a prospective, randomized controlled trial. Bone Joint Res 2020 Jun;9(6):322-332. [doi: 10.1302/2046-3758.96.BJR-2019-0180.R3] [Medline: 32670565]
50. Xu H, Yang J, Xie J, et al. Tourniquet use in routine primary total knee arthroplasty is associated with a higher transfusion rate and longer postoperative length of stay: a real-world study. BMC Musculoskelet Disord 2020 Sep 18;21(1):620. [doi: 10.1186/s12891-020-03623-5] [Medline: 32948173]
51. Vandenbussche E, Duranthon LD, Couturier M, Pidhorz L, Augereau B. The effect of tourniquet use in total knee arthroplasty. Int Orthop 2002;26(5):306-309. [doi: 10.1007/s00264-002-0360-6] [Medline: 12378360]
52. Malviya A, Martin K, Harper I, et al. Enhanced recovery program for hip and knee replacement reduces death rate. Acta Orthop 2011 Oct;82(5):577-581. [doi: 10.3109/17453674.2011.618911] [Medline: 21895500]

Abbreviations

DVT: deep vein thrombosis
ePROM: electronic patient-reported outcome measure
LOS: length of stay
MKUH: Milton Keynes University Hospital
NHS: National Health Service
PROM: patient-reported outcome measure
TKA: total knee arthroplasty
VAS: visual analog scale

Edited by A Kushniruk; submitted 24.08.24; peer-reviewed by A Panahi, M O'Connor; revised version received 21.03.25; accepted 23.03.25; published 10.07.25.

Please cite as:
Ajrawat P, Price B, Gooch D, Serban R, Al-Habsi R, Pearce O
Assessing Postoperative Pain in Patients Who Underwent Total Knee Arthroplasty Using an Automated Self-Logging Patient-Reported Outcome Measure Collection Device: Retrospective Cohort Study
JMIR Hum Factors 2025;12:e65271
URL: https://humanfactors.jmir.org/2025/1/e65271
doi:10.2196/65271

© Prabjit Ajrawat, Blaine Price, Daniel Gooch, Rudolf Serban, Ruqaiya Al-Habsi, Oliver Pearce. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 10.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Factors Influencing Health Workers' Acceptance of Guideline-Based Clinical Decision Support Systems for Preventive Services in Thailand: Questionnaire-Based Study

Tullaya Sitasuwan[1], MD; Prapat Suriyaphol[2], PhD; Saranath Lawpoolsri[3], MD, PhD; Ngamphol Soonthornworasiri[3], PhD; Wirichada Pan-ngum[3,4], PhD

[1]Department of Medicine, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand

[2]Division of Medical Bioinformatics, Department of Health Research and Development, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand

[3]Department of Tropical Hygiene, Faculty of Tropical Medicine, Mahidol University, 8th floor and 9th floor Tranakchit Harinasuta Building, 420/6 Ratchawithi Road, Bangkok, Thailand

[4]Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand

Corresponding Author:
Wirichada Pan-ngum, PhD
Department of Tropical Hygiene, Faculty of Tropical Medicine, Mahidol University, 8th floor and 9th floor Tranakchit Harinasuta Building, 420/6 Ratchawithi Road, Bangkok, Thailand

Abstract

Background: A guideline-based clinical decision support system (CDSS) is a knowledge-based system designed to collect crucial data from electronic medical records to generate decision-making based on system data requirements and inputs from standard guidelines. Despite the potential to enhance health care delivery, the adoption rate of CDSSs in clinical practice remains suboptimal.

Objective: This study aimed to evaluate the determinants influencing the intention to use a new CDSS in preventive care within clinical practice.

Methods: A single-center, questionnaire-based, cross-sectional study was conducted among physicians and medical students responsible for providing comprehensive preventive services at the Continuity of Care Clinic, Siriraj Hospital, Thailand.

Results: In total, 89 participants were enrolled. Relationships between factors impacting the adoption of CDSSs were analyzed using correlation and regression analysis. We found that physicians' intentions to adopt the CDSS for preventive care were high, with 79% (70/89) of participants expressing their intention to use the system. According to the study's conceptual framework, modified from the original unified theory of acceptance and use of technology model, physicians' positive attitudes toward CDSS use in preventive services and a high level of effort expectancy emerged as crucial factors influencing the intention to use the new CDSS. The odds ratios for these factors were 5.44 (95% CI 1.62 - 18.34, *P*=.006) and 7.60 (95% CI 1.55 - 31.37, *P*=.01), respectively. Similar results were observed for medical students and for physicians who had graduated. The most prevalent barriers to CDSS implementation were related to physicians' attitudes, followed by issues such as the accuracy and burden of data input, time constraints for clinicians, and the risk of workflow disruption.

Conclusions: There was a high intention to adopt the CDSS in preventive care. Positive physician attitudes toward CDSS use in preventive services and effort expectancy were found to be critical factors influencing the intention to use the new CDSS.

(JMIR Hum Factors 2025;12:e57314) doi:10.2196/57314

KEYWORDS

clinical decision support systems; preventive health services; medical informatics; technology acceptance; continuity of patient care

Introduction

Background

In contemporary medical practice, evidence-based guidelines play a crucial role in integrating the best available evidence from various studies. These guidelines serve to support and guide physicians in making informed decisions regarding further investigation and treatment, ultimately improving the quality of care. Preventive care, among other areas, has numerous guidelines from various professional societies, recommending vaccinations and screening for diseases such as cardiometabolic diseases, cancer, and dementia. However, previous studies have revealed that clinical practice guidelines are not made full use

XSL•FO
RenderX

of in real-world clinical settings [1,2]. Barriers to implementing these guidelines include a lack of a comprehensible structure and local applicability, complexity in recommended clinical actions, a lack of awareness among physicians, insufficient time for recommended preventive services, a lack of awareness among patients, and financial constraints [2-5]. To improve the implementation of standard guidelines, clinicians value a user-friendly format, clear and valid evidence, details about competency and training requirements, and guidance on how to integrate experience with evidence when tailoring recommendations to individual patients.

The advent of information technology (IT) has had a major impact on the health care sector, with 1 notable advance in clinical practice being clinical decision support systems (CDSSs). CDSSs have been defined as systems "designed to aid directly in clinical decision making, in which characteristics of individual patients are used to generate patient-specific assessments or recommendations that are then presented to clinicians for consideration" [6]. CDSSs can range in complexity, from simple alerts and reminders to diagnostic assistance, therapy planning, and prescription decision support. A guideline-based CDSS is a knowledge-based system that collects crucial data from electronic medical records (EMRs), matches these data with the system's data requirements, and generates reasoning and decision-making based on input from established guidelines. This type of CDSS is considered a mid-level CDSS. Using a guideline-based CDSS can help to overcome obstacles associated with using traditional paper-based guidelines and enhance physicians' adherence to recommendations [6-8].

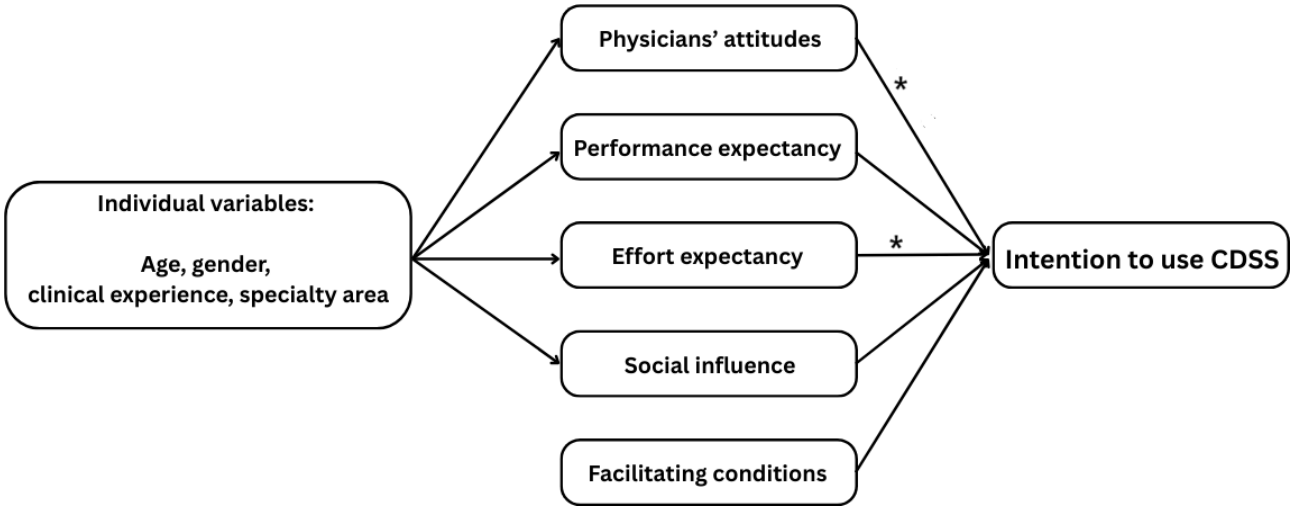
Literature Review

Despite increasing evidence for the benefits of guideline-based CDSSs, their adoption among outpatient physicians remains limited [9-11]. The challenges encountered during CDSS implementation stem from human factors related to system use, information quality, availability of EMR implementation, system quality, and system maintenance [11,12]. The main barriers to the implementation of mid-level CDSSs in low- and middle-income countries are inefficient CDSS-EMR system design, perception of use by physicians, and the cost of CDSS acquisition and maintenance. Additional barriers include negative beliefs about the impact of CDSSs on the clinician-patient relationship and physicians' reluctance to adopt these systems [13-15]. To investigate the successful adoption and acceptance of CDSSs, a comprehensive understanding of the interactions between human, technological, and organizational factors is crucial. Although numerous models exist for evaluating IT acceptance and use, in this study, we used the Unified Theory of Acceptance and Use of Technology (UTAUT). The UTAUT is formulated based on 4 core determinants of intentions and usage—performance expectancy (PE), effort expectancy (EE), social influence (SI), and facilitating conditions (FC). It also includes 4 moderators of key relationships with technology—gender, age, experience, and voluntariness of use. The UTAUT model distills the critical factors and contingencies related to predicting behavioral intention to use technology, particularly technology used primarily in organizational contexts. There are a lot of theories that underlie the UTAUT model, including the diffusion of innovation theory, the theory of reasoned action, the theory of planned behavior, the motivation theory, the technology acceptance model, the personal computer utilization model, and the social cognitive theory [16].

A meta-analysis found that the 4 primary constructs of the UTAUT model exhibit varying degrees of predictive power, with PE emerging as the strongest predictor of intention to use a new technology [17]. Previous studies have confirmed the validity of the UTAUT model in health care settings and for CDSSs, with 1 study conducted in Thailand [18-21]. Furthermore, weak positive associations have been observed between the intention to use IT and reported IT use in low- and middle-income countries, while a negative relationship has been noted between the perception of a threat to professional autonomy and the intention to adopt CDSSs [14,15]. In this study, we developed a conceptual framework based on the UTAUT model, aiming to provide a comprehensive understanding of the adoption of guideline-based CDSSs in preventive care within the health care system of an upper middle-income country. The conceptual framework encompassed two key components: (1) physicians' attitudes toward CDSS usage, which are positively influenced by the perceived importance of preventive services, the belief that CDSSs will enhance the quality of care by improving adherence to clinical practice guidelines, and physician involvement in the CDSS development process, but negatively affected by perceived threats to professional autonomy; and (2) factors associated with CDSS PE, EE, SI, and FC, that is, the 4 factors that arise from the UTAUT model. The conceptual framework is illustrated in Figure 1.

XSL•FO
RenderX

Figure 1. Conceptual framework used in this study * indicates a significant association at the 95% level of intention to use the CDSS. CDSS: clinical decision support system.



The objective of this study was to assess user acceptance of guideline-based CDSSs in preventive care within clinical practice using the expanded UTAUT model and evaluate the readiness of medical staff and students to use this system in the Continuity of Care Clinic at Siriraj Hospital, Thailand.

Methods

Study Design and Participants

This study had a single-center, questionnaire-based, cross-sectional design. The participants comprised physicians and medical students who provided comprehensive preventive services at the Continuity of Care Clinic, Department of Medicine, Siriraj Hospital, between 2018 and 2021. The clinic primarily focuses on delivering ongoing care to adult patients with chronic noncommunicable diseases, aligning with the most up-to-date guidelines while promoting preventive measures such as screening for newly developed cardiometabolic diseases, cancer screening, counseling for individuals who engage in risky health behaviors, and appropriate vaccinations.

The sample size was determined based on an appropriate percentage of the finite population. The target population comprised 16 experts, 200 in-training physicians, and 300 medical students. Our primary aim was to survey a minimum of 15% of the target population, which would require a total of 75 participants to provide a sufficiently representative sample. To ensure the inclusion of individuals with valuable expertise, we planned to recruit all ambulatory medicine experts who served as permanent physicians at the Continuity of Care Clinic. This approach was intended to maximize the breadth of knowledge represented in our study, recognizing the importance of these experts' contributions within the broader target population.

Data Collection

Data collection was conducted using a newly developed online questionnaire, created in Google Forms, to assess user acceptance and examine the factors influencing the adoption into clinical practice of guideline-based CDSSs in preventive care. The questionnaire was pretested among a small group of physicians to ensure clarity and technical functionality. Minor revisions were made based on their feedback.

The first page of the Google Form presented participants with detailed information about the study, including its purpose, modes of contact, and data confidentiality, as well as the fact that participation was voluntary. Participants were recruited via institutional email, targeted LINE (LY Corporation; messaging app) groups, and in-person invitations. The inclusion criteria were medical students and physicians who had experience working in the Continuity of Care Clinic between 2018 and 2021 and who were currently working at Siriraj Hospital. Each participant could only respond once, and submission was allowed only after all mandatory fields had been completed. No IP addresses, email addresses, or other identifying information were collected.

Ethical Considerations

The research protocol received approval from the Mahidol University Multi-faculty Cooperative Institutional Review Board (protocol TMEC21-052, MU-MOU CoA 608/2021). Informed consent to participate was obtained electronically by requiring participants to click "Next" before they proceeded to the questionnaire. All responses were stored securely in a password-protected institutional Google Drive account, accessible only by the research team.

Routine Workflow for Preventive Services in the Continuity of Care Clinic at Siriraj Hospital

Outpatient services at Siriraj Hospital operate using a hybrid EMR system, in which medical records and prescriptions are initially handwritten on paper and subsequently scanned into an EMR for archival and review purposes. Laboratory results, radiological study reports, and medication prescription data are managed through separate electronic systems within the laboratory, radiology, and pharmacy departments, respectively, and are made viewable through EMR-integrated plug-ins. This system includes a basic-level CDSS, which provides pop-up reminders for drug allergies, renal function, and high-alert medications such as warfarin. A CDSS that retrieves data directly from the EMR is not applicable in this system.

XSL•FO
RenderX

At the Continuity of Care Clinic, the primary focus is on maintaining continuity of care for patients with chronic noncommunicable diseases, while simultaneously promoting preventive services. Due to the limitations of the EMR system, paper-based guidance for preventive care is incorporated into clinical workflows. This approach aims to ensure the consistent delivery of preventive services to each patient, despite the rotation of different physicians.

Online Questionnaire

The questionnaire consisted of 22 items, including questions regarding demographic details and items related to physicians' attitudes, EE, PE, SI, and FC, formulated in accordance with the structured questions of the UTAUT model [16]. Questions pertaining to physicians' attitudes were derived from factors identified in previous studies as being influential on usage intentions [14,15,22]. Responses were recorded on a 5-point Likert scale ranging from "strongly agree" to "strongly disagree." Among these constructs, physicians' attitudes and EE had 1 item that was negatively worded. The questionnaire was initially developed in English and subsequently translated into Thai using the translation-back-translation method. Before completing the questionnaire, participants were asked to view a demonstration video outlining the functionality and use of the proposed CDSS.

Introductory Video for the Proposed CDSS

The study participants were introduced to the proposed CDSS through an informative online video. The video was expertly created by an ambulatory medicine specialist who possessed extensive expertise in integrating preventive care into clinical practice and was also highly familiar with the EMR system employed at the study site. This video provided participants with invaluable insights into the prevailing challenges within the workflow and offered a firsthand experience of the system's performance. The demonstration included a comprehensive overview of data entry methods, advanced data visualization capabilities, and personalized recommendations for preventive services. In addition, the video covered essential information concerning support services and regular maintenance of the clinical rules associated with these systems. A summary of the video narration and selected snapshots from the demonstration video are provided in Multimedia Appendix 1.

Statistical Analysis

Data were analyzed using the R programming language (R Core Team). The descriptive statistics are presented as mean, median, SD, minimum, and maximum. Logistic regression analysis was used to identify factors associated with the acceptance of guideline-based CDSSs in preventive care. The final model was selected through the application of forward and backward selection methods in multiple logistic regression analysis. The statistical significance of the final model was evaluated using the Wald test.

Results

Demographic Data

A total of 89 participants took part in this study. Their baseline characteristics are presented in Table 1. Among the participants, 45 (51%) were male. The median age of participants was 25 (range 22-64) years. The majority of participants (49/89, 55%) were medical students, with 28 participants (32%) in their fifth year and 21 participants (24%) in their sixth year. Furthermore, 40 participants (45%) were medical graduates with varying levels of experience; 4 (5%) were general practitioners, 11 (12%) had undergone training in an internal medicine residency program, and 25 (28%) were internal medicine specialists, including 16 individuals (18%) who specialized in ambulatory medicine. The medical graduate participants had clinical experience ranging from 2 to 40 (median 8) years.

Table . Demographic data of study participants.

Characteristic	Statistical value (N=89)
Age (years), median (range)	25 (22 - 64)
Sex (male), n (%)	45 (51)
Clinical experience, n (%)	
Medical student	49 (55)
Fifth year	28 (32)
Sixth year	21 (24)
Graduated physician	40 (45)
General practitioner	4 (5)
Internal medicine resident	11 (12)
Internal medicine specialist	9 (10)
Ambulatory medicine expert	16 (18)
Duration of medical doctors' professional practice (years), median (range)	8 (2-40)

XSL•FO
RenderX

User Acceptance Measures

After viewing the demonstration video of the proposed system, 70 participants (79%) expressed a definitive intention to use the system if it were available. However, 18 participants expressed reasonable concerns about its usage, while 1 participant refused outright to use the new system. Notably, there were no differences in the intention to use the proposed system based on the participants' baseline characteristics. Nevertheless, within the medical student group, it was observed that sixth-year medical students exhibited a higher comfort level with adopting the new system compared with their fifth-year counterparts, with acceptance rates of 95% and 68%, respectively.

The mean (SD) and median (range) scores for each question are presented in Table 2. The scoring system for each question had a maximum value of 5. For negative questions, appropriate score conversions were implemented before calculating the total score for each construct.

XSL•FO
RenderX

Table . Scoring of each question assessing user acceptance.

Determinants of intentions and usage	Full score[a]	Mean (SD)	Median (range)
PA[b]			
Perceived importance of disease prevention	5	4.89 (0.34)	5 (3-5)
Perceived improved quality of care by increased adherence to clinical practice guidelines	5	4.708 (0.53)	5 (3-5)
Perceived threat to professional autonomy[c]	5	4 (1)	4 (1-5)
Involvement in the CDSS[d] development process	5	4.17 (0.84)	4 (2-5)
Total PA score	20	17.78 (1.74)	18 (12 - 20)
PE[e]			
Perceived usefulness	5	4.64 (0.59)	5 (3-5)
Job-fit	5	4.67 (0.54)	5 (3-5)
Outcome expectations	5	4.55 (0.74)	5 (2-5)
Total PE score	15	13.87 (1.60)	15(9-15)
EE[f]			
Ease of use	5	4.29 (0.84)	4 (1-5)
Complexity[c]	5	3.75 (1.20)	4 (1-5)
Total EE score	10	8.05 (1.78)	8 (2-10)
SI[g]			
Subjective norm	5	4.19 (0.75)	4 (3-5)
Social factors	5	4.11 (0.92)	4 (1-5)
Image	5	3.76 (1.16)	4 (1-5)
Total SI score	15	12.07 (2.18)	12 (5 - 15)
FC[h]			
Perceived behavioral control	5	4.40 (0.75)	4 (2-5)
Facilitating conditions	5	4.26 (0.76)	4 (2-5)
Compatibility	5	4.25 (0.88)	4 (2-5)
Total FC score	15	12.91 (1.98)	13 (6 - 15)
Total score	75	64.67 (7.25)	65 (42 - 75)

[a]Measurement scale: 1=strongly disagree, 2=disagree, 3=neutral, 4=agree, 5=strongly agree.

[b]PA: physician's attitude.

[c]Indicates a negative question; a higher score means less concern about these items.

[d]CDSS: clinical decision support system.

[e]PE: performance expectancy.

[f]EE: effort expectancy.

[g]SI: social influence.

[h]FC: facilitating conditions.

The correlation between each construct and the intention to use the proposed system is presented in Table 3. A significant positive correlation was found between the physicians' attitudes, PE, EE, and their intention to use the system.

XSL•FO
RenderX

Table . The correlation matrix between user acceptance measurement and intention to use the new clinical decision support system.

Variable	PA[a]	PE[b]	EE[c]	SI[d]	FC[e]	IU[f]
PA						
Spearman rank	1	0.58	0.45	0.43	0.48	0.40
P value	—[g]	<.001	<.001	<.001	<.001	<.001
PE						
Spearman rank	0.58	1	0.62	0.51	0.67	0.29
P value	<.001	—	<.001	<.001	<.001	0.01
EE						
Spearman rank	0.45	0.62	1	0.25	0.40	0.44
P value	<.001	<.001	—	0.02	<.001	<.001
SI						
Spearman rank	0.43	0.51	0.25	1	0.68	–0.04
P value	<.001	<.001	0.02	—	<.001	0.69
FC						
Spearman rank	0.48	0.67	0.40	0.68	1	0.14
P value	<.001	<.001	<.001	<.001	—	0.20
IU						
Spearman rank	0.40	0.29	0.44	–0.04	0.14	1
P value	<.001	0.01	<.001	0.69	0.20	—

[a]PA: physician's attitude.

[b]PE: performance expectancy.

[c]EE: effort expectancy.

[d]SI: social influence.

[e]FC: facilitating conditions.

[f]IU: intention to use the clinical decision support system.

[g]Not applicable.

Factors Associated With Intention to Use the CDSS

The outcome from the logistic regression analysis is presented in Table 4. Physicians' positive attitudes toward preventive medicine and the integration of new technology into clinical practices were found to be significant factors associated with the adoption and use of the new CDSS, with adjusted odds ratios (OR) of 8.41 (95% CI 2.54 - 27.82). The EE was also associated with the intention to adopt the CDSS (adjusted OR 11.71, 95% CI 2.45 - 55.95). However, PE, SI, and FC did not reach statistical significance in their association with the adoption of the proposed system. In addition, individual factors, such as age, sex, clinical experience, and duration of clinical practice, did not appear to influence the likelihood of adopting the proposed CDSS.

XSL•FO
RenderX

Table . Logistic regression results.

Determinants of intentions and usage	Crude OR[a] (95% CI)	*P* value	Adjusted OR[b] (95% CI)	Final model OR (95% CI)	*P* value
Physician's attitude[c]	8.09 (2.55 - 25.63)[d]	<.001	8.41 (2.54 - 27.82)[d]	5.44 (1.62-18.34)[e]	.006
Performance expectancy[c]	2.78 (0.98 - 7.85)	.05	2.78 (0.94 - 8.26)	—[f]	—
Effort expectancy[c]	11.33 (2.43-52.85)[e]	.002	11.71 (2.45-55.95)[e]	7.60 (1.55-31.37)[e]	.01
Social influence[c]	0.97 (0.35 - 2.72)	.96	0.84 (0.29 - 2.43)	—	—
Facilitating conditions[c]	1.73 (0.62 - 4.82)	.30	1.63 (0.57 - 4.66)	—	—
Age	0.96 (0.91 - 1.01)	.12	—	—	—
Sex[g]	1.11 (0.40 - 3.07)	.84	—	—	—
Clinical experience					
Medical student	1	—	—	—	—
IM[h] resident and internist	0.93 (0.26-3.39)	.91	—	—	—
Ambulatory medicine expert	0.51 (0.15-1.81)	.30	—	—	—
Duration of clinical practice					
0 - 3 years	1	—	—	—	—
4 - 10 years	0.97 (0.27 - 3.50)	.96	—	—	—
>10 years	0.41 (0.11 - 1.49)	.18	—	—	—

[a]OR: odds ratio.

[b]Adjusted by age and clinical experience group because the result from the univariable analysis showed a *P* value <.20.

[c]Where values lower than the mean of the total score were treated as baseline, that is, odds ratio=1.

[d]*P*<.001.

[e]*P*<.05.

[f]Not applicable.

[g]Where being male was treated as the baseline, that is, odds ratio=1.

[h]IM: internal medicine.

The final model consensus was established through the implementation of the forward selection and backward selection methods, using the Wald test. The integrated final model encompassed all determinants associated with the intention to adopt the CDSS, along with the demographic variables indicated in Table 4. The factors influencing the adoption and use of the proposed CDSS were identified as the physicians' attitudes (odds ratio 5.44, 95% CI 1.62 - 18.34, *P*=.006) and EE (OR 7.60, 95% CI 1.55 - 31.37, *P*=.01), visualized in the study's conceptual framework in Figure 1.

Subgroup analysis demonstrated that there were no significant differences in the logistic regression results between medical students and graduated physicians, as indicated in Multimedia Appendix 2. Nevertheless, it is important to highlight the notable factors associated with the intention to adopt the CDSS, which were identified as being physicians' attitudes and EE.

Benefits of and Barriers to CDSS Adoption in Clinical Practice

All study participants agreed that the proposed CDSS would bring benefits to their clinical practice. Specifically, 84 participants believed that the proposed system encompassed multiple components that would prove advantageous in their clinical practices. A summary of the participants' preferences and barriers to their adopting the CDSS is provided in Table 5.

XSL•FO
RenderX

Table . The benefits and barriers of the proposed clinical decision support system.

Benefits and barriers	Frequency, n (%)
Benefits of CDSS[a]	
Data gathering and data visualization	79 (89)
Provides supplementary information for recommended disease screening	70 (79)
Provides additional information regarding recommended vaccines	68 (76)
Individualized suggestions for further preventive services	61 (69)
Integration with the doctor's order sheet	50 (56)
Information to improve clinical decision-making	49 (55)
Provides flexibility in use, not mandatory for every patient	37 (42)
Barriers to CDSS adoption	
Other physicians usually take care of prevention services	6 (7)
The use of the CDSS would require them to allocate more time to each patient	5 (6)
Potential disruption of the current workflow with the implementation of the CDSS	4 (5)
The burden of data input outweighs the benefit of the CDSS	3 (3)
Leakage of patients' personal data	3 (3)
The CDSS will not improve my performance	2 (2)
The accuracy of data input	2 (2)

[a]CDSS: clinical decision support system.

The majority of participants (70/89, 79%) reported having no concerns about implementing the new system in their clinical practice, while 19 participants raised concerns regarding the proposed CDSS. Among them, 4 participants had multiple concerns, while the rest expressed a single specific concern about the new system. The most prevalent barrier identified for the adoption of the new system was the participants' perception that they do not play a primary role in preventive services, which subsequently led to their decision not to use it.

Discussion

Principal Results and Comparison With Previous Work

This study suggests that a majority of medical professionals, including students and graduates, have a favorable attitude toward and intention to use a CDSS in their clinical practice. Factors, such as physicians' attitudes toward the use of CDSSs in preventive services and its perceived ease of use, play an important role in the adoption and use of the proposed system. The primary barrier to CDSS adoption arises from physician-related factors. If physicians are not actively involved in preventive services or do not prioritize them, they would be less likely to employ the proposed CDSS.

This study was undertaken to evaluate the acceptance of a new CDSS in preventive services. It was postulated that physicians differ from other users of IT due to the impact on their professional work of implementing an automated system that influences the comprehensiveness of their clinical management and adherence to clinical practice guidelines. However, physicians may perceive this disruption as a threat to their professionalism [23].

Our model indicates that a physician's attitude was one of the most influential factors determining their intention to use the new system. From the mean scores of each question included in this construct, we found that the highest score was for physicians recognizing the importance of disease prevention, which aligns with the primary function of the CDSS. The perception that the system can enhance the quality of health care delivery by promoting adherence to clinical practice guidelines ranked second. This finding is consistent with previous research emphasizing that physicians' decision to use IT relies on the usefulness and compatibility of any new technology with their clinical practice [24]. Physicians are more inclined to accept a system that they think will improve their performance. However, the perceived threat to professional autonomy, which has been identified as a significant deterrent to CDSS adoption among physicians [15,23], received the lowest mean score and exhibited the widest range of scores in this

XSL•FO
RenderX

construct. This suggests that participants had different concerns regarding the threat to their professional autonomy posed by the proposed CDSS. Nonetheless, most participants paid little attention to this factor, as evidenced by the median score of 4 out of 5. The level of clinical decision support provided by the CDSS can explain this discrepancy. In our study, the CDSS is intended to serve as a supportive tool for physicians' decision-making according to well-established clinical practice guidelines. The function of the proposed CDSS primarily involves clinical data gathering from EMR, data visualization, and recommendations for routine preventive services such as cancer screening and age-appropriate vaccination. These tasks do not rely heavily on individual clinical expertise, and the system allows physicians to make their own choices based on their experience and doctor-patient discussions. The low-to-medium level of clinical decision support provided by the CDSS may account for the limited concern expressed regarding the threat to professional autonomy in this study.

EE emerged as another influential factor determining the intention to use the new system. This finding is consistent with previous studies [15,20,21] and aligns with measures in the DeLone and McLean [25] information system success model that suggest ease of use is related to user satisfaction and adoption. The end user interface arrangement, which should be simple, user-friendly, and matched with a physician's thinking process, also constitutes a crucial factor contributing to system adoption [26].

The highest positive response on PE (mean score of 13.87 out of 15; 93%) reflected that participants expressed a strong belief in the system's potential usefulness. This is likely influenced by the demonstration video about the function and use of the proposed CDSS, which was made by an experienced clinician involved in preventive services at the study site to highlight how the system could address existing workflow gaps. Furthermore, the proposed CDSS has the potential to enhance physicians' performance, improve their effectiveness, and reduce the time spent on preventive services during patient encounters. The observed marginal positive relationship between PE and the intention to use the system (P=.05) was consistent with the expectation that perceived usefulness is a key determinant of adoption intention.

SI was not found to significantly affect the intention to use the CDSS in our study. This finding is consistent with previous literature suggesting that physicians differ from other groups of IT users, as they tend to prioritize a system's perceived usefulness over ease of use or peer influence within their organization [23]. However, this result may be attributable to the preimplementation context of our study, in which participants had no previous experience using the system and limited exposure to peer usage or established institutional norms surrounding its adoption.

Previous studies have indicated that SI may exert a greater effect, particularly among medical students, once a system has been actively implemented and its use becomes more visible within the learning or clinical environment [27]. Although our study was conducted during medical students' clinical clerkship phase, when decisions are likely shaped by clinical supervisors and perceived task relevance, peer and faculty endorsement may become increasingly influential as the CDSS becomes embedded in routine clinical workflows. Further research following system implementation is warranted to better understand the evolving role of social dynamics in CDSS adoption.

Subgroup analysis revealed that the factors influencing the intention to adopt the new CDSS did not differ significantly between medical students and physicians who had graduated. Correlation analysis indicated a significant positive correlation between physician's attitude, PE, EE, and intention to use the CDSS, consistent with several previous studies [15,20,21,28].

The rate of intention to adopt the proposed CDSS in this study was remarkably high (79%). No statistically significant differences were observed in the baseline characteristics of physicians related to the intention to adopt. However, there was a disparity in intention to adopt rates within the medical student group. Sixth-year medical students exhibited a higher comfort level with the new system compared with fifth-year medical students, with acceptance rates of 95% versus 68% (P=.04). These findings warrant further discussion. During medical education in Thailand, fifth-year medical students have had limited clinical exposure (less than 2 years) and are under the continuous supervision of medical staff. Consequently, they may not consider clinical judgment to be an integral part of their responsibilities. In contrast, sixth-year medical students engage in clinical clerkship, which distinguishes them from fifth-year medical students. Their mindset may differ, resulting in contrasting attitudes toward the benefits and usefulness of technology in their job, which could in turn influence their adoption rates. Furthermore, greater clinical exposure enables sixth-year medical students to recognize the substantial cognitive burden of being a general practitioner. They may perceive the system as being more valuable in alleviating their cognitive load than fifth-year medical students.

Qualitative analysis of free-text comments revealed that physicians regarded the system's ability to gather data from various sources in EMR and present these data in a recognizable and comprehensible format as the most beneficial aspect of the proposed CDSS. In addition, the system provides patient counseling information on preventive services, which was highly valued. These results align with those of previous studies, in that physicians perceive the benefit of CDSSs mainly in terms of accessibility to standardized information and the facilitation of patient discussion [19,21]. Automatic data gathering and visualization have also frequently been cited as favorable aspects of a CDSS [29]. These findings suggest that preventive services have relatively low priority in current clinical practice, primarily due to challenges in acquiring individual patient data and obtaining comprehensive information for patient counseling. A CDSS can effectively bridge this gap.

Regarding system performance, the results indicated that more than half of the study participants would prefer a single integrated system that incorporates a patient's history, laboratory results, vaccination records, and a linked appointment system. However, some physicians favored the use of the CDSS as an optional tool for each patient visit. This finding should be taken

XSL•FO
RenderX

into account during system design. An “add-on” system is perceived as being advantageous as it allows physicians to exercise their own choice. Nonetheless, this feature should be considered carefully, as it could pose a threat to system implementation due to low system usage [30].

Most study participants reported that they would use the proposed CDSS without any concerns. However, some participants expressed apprehension about adopting the CDSS. Regarding physician-related factors, physicians’ attitudes emerged as the primary barrier. As preventive services are not obligatory for all physicians, those who are not involved in such tasks or do not prioritize health promotion may be reluctant to use the proposed CDSS, as it could increase the time required for each patient visit and disrupt their routine workflow.

Regarding patient-related factors, the respondents expressed concerns about the accuracy of data input, as the system allows manual input of health maintenance data obtained elsewhere. Consequently, unverified data may lead to incorrect data being entered, which could subsequently impact the recommendations made during the use of the CDSS. This concern has been identified in previous studies with regard to information quality [31,32].

Finally, system-related factors were also identified as barriers to adoption. The burden of data input was the primary concern from this perspective. In addition, time constraints for clinicians and the potential disruption to workflow were considered important system design factors. These findings align with previous qualitative studies suggesting that a failure to balance the data input burden, the presentation of clinically relevant information, and integration into the workflow can negatively impact the adoption of CDSSs [29,30,33].

Study Implications

This study suggests that a carefully designed CDSS, featuring a simple, user-friendly interface, would be readily adopted by physicians. A potential application of this study is the development and implementation of the proposed CDSS described in the demonstration video. For successful implementation, several key factors should be considered, such as ensuring interoperability with existing EMR systems, developing a user-friendly interface, and integrating standardized security protocols. Streamlining data input and involving patients, while allowing physicians to verify and adjust data, may reduce the documentation burden.

Beyond system design, physician attitudes play a critical role in the adoption and use of a CDSS in the long term. Therefore, it is imperative to emphasize the incorporation of preventive services into the holistic care model during medical education, particularly in the context of chronic disease follow-up in ambulatory settings. In addition, long-term use of a CDSS requires institutional support, including ongoing training and effective change management strategies to facilitate integration into clinical practice.

Future Work

While this study demonstrates there would be a high rate of adoption of the proposed CDSS, some concerns remain among physicians, particularly regarding potential disruptions to clinical workflow, increased time spent per patient, and the burden of data input. Future qualitative studies will be important to guide the development of an improved system and to inform strategies for effective organizational implementation. In addition, addressing concerns regarding autonomy and emphasizing the role of CDSSs in supporting clinical decision-making may further encourage adoption. Long-term multicenter studies evaluating the system’s impact on clinical outcomes, as well as the perceptions of physicians and patients during clinical visits, would be valuable. In particular, these studies should explore how the use of a CDSS to guide clinical decision-making affects physician attitudes, patient trust, and the overall quality of the clinical encounter.

Strengths and Limitations

This study had several strengths. It represents the first evaluation of the intention of physicians to adopt a guideline-based CDSS in Thailand. A questionnaire was administered, incorporating determinants from the UTAUT model and factors identified to be salient in Southeast Asia, including physician attitudes toward new technology. Furthermore, a unique video was used to introduce the proposed system, featuring an experienced clinician demonstrating preventive services and addressing current workflow challenges. The video showcased the end user interface, data entry methods, CDSS performance, and integration with the existing EMR system, providing participants with an experience to aid their decision-making process.

However, the study also had some limitations. First, it was conducted at a single center, Thailand’s largest medical school hospital, potentially limiting the generalizability of our findings to other types of hospitals with differing EMR systems and limited IT support. Second, the inclusion of participants familiar with the continuity of care model incorporating preventive services resulted in a skewed-left distribution and wide CIs in the conceptual framework and logistic regression analysis.

Furthermore, this cross-sectional study examined preadopters of the CDSS, with self-reported constructs based on the demonstration video, impeding the determination of causal relationships. It is crucial to acknowledge potential differences in attitudes between preadopters and postadopters, limiting the generalizability of the study’s results to the early stages of CDSS implementation.

Challenges During the Research

This study was conducted during the COVID-19 pandemic, which posed several challenges. Hospital services for nonemergency cases were reduced by approximately 50%, and medical students transitioned to online learning. As a result, it was difficult to recruit study participants, particularly medical students, as they were not physically present at the clinic and were dispersed across various off-campus locations. Similarly, many graduated physicians were reassigned from the Continuity of Care Clinic to COVID-19 patient care duties. Furthermore, the closure of outpatient services greatly limited clinical exposure for medical students. Consequently, only approximately half of the fifth-year medical students had the opportunity to gain experience in providing comprehensive

XSL•FO
RenderX

preventive services at the Continuity of Care Clinic during the study period.

Conclusions

The use of CDSS in clinical management has the potential to enhance the quality of health care delivery. This study reveals a noteworthy inclination among physicians to embrace a guideline-based CDSS specifically tailored for preventive care. Furthermore, there were no significant differences in the intention to adopt the system based on diverse baseline characteristics of physicians.

Drawing upon the study's comprehensive conceptual framework, it was evident that positive physician attitudes toward incorporating CDSSs in preventive services, along with their perceived ease of use, emerged as the most pivotal determinants influencing the intention to use the novel CDSS. Conversely, factors such as PE, SI, and FC had minimal influence over the intention to adopt the proposed system. This study identified a range of common obstacles to the adoption of a CDSS, including physicians' attitudes, data input accuracy and burden, physicians' time constraints, and the potential disruption to existing workflow patterns.

Acknowledgments

The authors gratefully acknowledge all study participants who contributed their helpful and informative suggestions for the use of clinical decision support systems in preventive services. This research project was supported by the Master's Program in Biomedical and Health Informatics, Faculty of Tropical Medicine, Mahidol University, Thailand. WP is funded by Wellcome (grant 220211). For the purpose of open access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission.

Data Availability

Data are available upon reasonable request to the corresponding author.

Authors' Contributions

The authors' contributions are as follows. Conception or design, acquisition, analysis or interpretation of data were contributed by TS, PS, SL, NS, and WP. Drafting the work or revising and final approval of the manuscript was contributed by TS, PS, SL, NS, and WP. Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved was handled by TS, PS, SL, NS, and WP.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary material.
[DOCX File, 723 KB - humanfactors_v12i1e57314_app1.docx]

Multimedia Appendix 2
Logistic regression of the subgroup analysis.
[DOCX File, 16 KB - humanfactors_v12i1e57314_app2.docx]

References

1. Gagliardi AR, Brouwers MC, Palda VA, Lemieux-Charles L, Grimshaw JM. How can we improve guideline use? A conceptual framework of implementability. Implement Sci 2011 Mar 22;6:26. [doi: 10.1186/1748-5908-6-26] [Medline: 21426574]
2. Srivanichakorn W, Asavathitanonta K, Washirasaksiri C, et al. Prescribing rate of influenza vaccine among internal medicine residents for outpatient continuum care. J Med Assoc Thai 2014 Dec;97(12):1281-1289. [Medline: 25764635]
3. Thanapirom K, Treeprasertsuk S, Rerknimitr R. Awareness of colorectal cancer screening in primary care physicians. J Med Assoc Thai 2012 Jul;95(7):859-865. [Medline: 22919978]
4. Cochrane LJ, Olson CA, Murray S, Dupuis M, Tooman T, Hayes S. Gaps between knowing and doing: understanding and assessing the barriers to optimal health care. J Contin Educ Health Prof 2007;27(2):94-102. [doi: 10.1002/chp.106] [Medline: 17576625]
5. Foy R, MacLennan G, Grimshaw J, Penney G, Campbell M, Grol R. Attributes of clinical recommendations that influence change in practice following audit and feedback. J Clin Epidemiol 2002 Jul;55(7):717-722. [doi: 10.1016/s0895-4356(02)00403-1] [Medline: 12160920]

XSL•FO
RenderX

6. Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. BMJ 2005 Apr 2;330(7494):765. [doi: 10.1136/bmj.38398.500764.8F] [Medline: 15767266]
7. Shalom E, Shahar Y, Parmet Y, Lunenfeld E. A multiple-scenario assessment of the effect of a continuous-care, guideline-based decision support system on clinicians' compliance to clinical guidelines. Int J Med Inform 2015 Apr;84(4):248-262. [doi: 10.1016/j.ijmedinf.2015.01.004] [Medline: 25649843]
8. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. NPJ Digit Med 2020;3:17. [doi: 10.1038/s41746-020-0221-y] [Medline: 32047862]
9. Romano MJ, Stafford RS. Electronic health records and clinical decision support systems: impact on national ambulatory care quality. Arch Intern Med 2011 May 23;171(10):897-903. [doi: 10.1001/archinternmed.2010.527] [Medline: 21263077]
10. Camacho J, Zanoletti-Mannello M, Landis-Lewis Z, Kane-Gill SL, Boyce RD. A conceptual framework to study the implementation of clinical decision support systems (BEAR): literature review and concept mapping. J Med Internet Res 2020 Aug 6;22(8):e18388. [doi: 10.2196/18388] [Medline: 32759098]
11. Gold R, Larson AE, Sperl-Hillen JM, et al. Effect of clinical decision support at community health centers on the risk of cardiovascular disease: a cluster randomized clinical trial. JAMA Netw Open 2022 Feb 1;5(2):e2146519. [doi: 10.1001/jamanetworkopen.2021.46519] [Medline: 35119463]
12. Meunier PY, Raynaud C, Guimaraes E, Gueyffier F, Letrilliart L. Barriers and facilitators to the use of clinical decision support systems in primary care: a mixed-methods systematic review. Ann Fam Med 2023;21(1):57-69. [doi: 10.1370/afm.2908] [Medline: 36690490]
13. Anugoolpracha S, Booncherd C. The effect of computerized clinical decision support system on prescribing NSAIDs in patients with stage 3 to 5 chronic kidney disease at a community hospital. Isan Journal of Pharmaceutical Sciences 2023;19(4):1-14. [doi: 10.14456/ijps.2023.19]
14. Esmaeilzadeh P, Sambasivan M, Kumar N, Nezakati H. Adoption of clinical decision support systems in a developing country: antecedents and outcomes of physician's threat to perceived professional autonomy. Int J Med Inform 2015 Aug;84(8):548-560. [doi: 10.1016/j.ijmedinf.2015.03.007] [Medline: 25920928]
15. Sambasivan M, Esmaeilzadeh P, Kumar N, Nezakati H. Intention to adopt clinical decision support systems in a developing country: effect of physician's perceived professional autonomy, involvement and belief: a cross-sectional study. BMC Med Inform Decis Mak 2012 Dec 5;12(1):142. [doi: 10.1186/1472-6947-12-142] [Medline: 23216866]
16. Venkatesh V, Morris M, Davis G, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425. [doi: 10.2307/30036540]
17. Khechine H, Lakhal S, Ndjambou P. A meta‐analysis of the UTAUT model: eleven years later. Can J Adm Sci 2016 Jun;33(2):138-152 [FREE Full text] [doi: 10.1002/cjas.1381]
18. Kijsanayotin B, Pannarunothai S, Speedie SM. Factors influencing health information technology adoption in Thailand's community health centers: applying the UTAUT model. Int J Med Inform 2009 Jun;78(6):404-416. [doi: 10.1016/j.ijmedinf.2008.12.005] [Medline: 19196548]
19. Litvin CB, Ornstein SM, Wessell AM, Nemeth LS, Nietert PJ. Adoption of a clinical decision support system to promote judicious use of antibiotics for acute respiratory infections in primary care. Int J Med Inform 2012 Aug;81(8):521-526. [doi: 10.1016/j.ijmedinf.2012.03.002] [Medline: 22483528]
20. Chang IC, Hwang HG, Hung WF, Li YC. Physicians' acceptance of pharmacokinetics-based clinical decision support systems. Expert Syst Appl 2007 Aug;33(2):296-303. [doi: 10.1016/j.eswa.2006.05.001]
21. Laka M, Milazzo A, Merlin T. Factors that impact the adoption of clinical decision support systems (CDSS) for antibiotic management. Int J Environ Res Public Health 2021 Feb 16;18(4):1901. [doi: 10.3390/ijerph18041901] [Medline: 33669353]
22. Ahlan AR, Ahmad BI. User acceptance of health information technology (HIT) in developing countries: a conceptual model. Procedia Technology 2014;16:1287-1296. [doi: 10.1016/j.protcy.2014.10.145]
23. Walter Z, Lopez MS. Physician acceptance of information technologies: role of perceived threat to professional autonomy. Decis Support Syst 2008 Dec;46(1):206-215. [doi: 10.1016/j.dss.2008.06.004]
24. Chau PYK, Hu PJ. Examining a model of information technology acceptance by individual professionals: an exploratory study. Journal of Management Information Systems 2002 Mar;18(4):191-229. [doi: 10.1080/07421222.2002.11045699]
25. DeLone WH, McLean ER. Measuring e-commerce success: applying the DeLone & McLean information systems success model. International Journal of Electronic Commerce 2004 Oct;9(1):31-47. [doi: 10.1080/10864415.2004.11044317]
26. Laing S, Mercer J. Improved preventive care clinical decision-making efficiency: leveraging a point-of-care clinical decision support system. BMC Med Inform Decis Mak 2021 Nov 11;21(1):315. [doi: 10.1186/s12911-021-01675-8] [Medline: 34763691]
27. Alrashed FA, Ahmad T, Almurdi MM, et al. Incorporating technology adoption in medical education: a qualitative study of medical students' perspectives. Adv Med Educ Pract 2024;15:615-625. [doi: 10.2147/AMEP.S464555] [Medline: 38975614]

XSL•FO
RenderX

28. Jeng DJF, Tzeng GH. Social influence on the use of clinical decision support systems: revisiting the unified theory of acceptance and use of technology by the fuzzy DEMATEL technique. Computers & Industrial Engineering 2012 Apr;62(3):819-828. [doi: 10.1016/j.cie.2011.12.016]
29. Khairat S, Marc D, Crosby W, Al Sanousi A. Reasons for physicians not adopting clinical decision support systems: critical analysis. JMIR Med Inform 2018 Apr 18;6(2):e24. [doi: 10.2196/medinform.8912] [Medline: 29669706]
30. Zheng K, Padman R, Johnson MP, Diamond HS. Understanding technology adoption in clinical care: clinician adoption behavior of a point-of-care reminder system. Int J Med Inform 2005 Aug;74(7-8):535-543. [doi: 10.1016/j.ijmedinf.2005.03.007] [Medline: 16043083]
31. Hsiao JL, Wu WC, Chen RF. Factors of accepting pain management decision support systems by nurse anesthetists. BMC Med Inform Decis Mak 2013 Jan 29;13(1):16. [doi: 10.1186/1472-6947-13-16] [Medline: 23360305]
32. Phansalkar S, Weir CR, Morris AH, Warner HR. Clinicians' perceptions about use of computerized protocols: a multicenter study. Int J Med Inform 2008 Mar;77(3):184-193. [doi: 10.1016/j.ijmedinf.2007.02.002] [Medline: 17409016]
33. Moxey A, Robertson J, Newby D, Hains I, Williamson M, Pearson SA. Computerized clinical decision support for prescribing: provision does not guarantee uptake. J Am Med Inform Assoc 2010;17(1):25-33. [doi: 10.1197/jamia.M3170] [Medline: 20064798]

Abbreviations

CDSS: clinical decision support system
EE: effort expectancy
EMR: electronic medical record
FC: facilitating conditions
IT: information technology
OR: odds ratio
PE: performance expectancy
SI: social influence
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by A Kushniruk; submitted 25.02.24; peer-reviewed by M Skouroliakou, M Abouzahra, R Panicker, S Sarbadhikari; revised version received 24.05.25; accepted 28.05.25; published 16.07.25.

Please cite as:
Sitasuwan T, Suriyaphol P, Lawpoolsri S, Soonthornworasiri N, Pan-ngum W
Factors Influencing Health Workers' Acceptance of Guideline-Based Clinical Decision Support Systems for Preventive Services in Thailand: Questionnaire-Based Study
JMIR Hum Factors 2025;12:e57314
URL: https://humanfactors.jmir.org/2025/1/e57314
doi:10.2196/57314

© Tullaya Sitasuwan, Prapat Suriyaphol, Saranath Lawpoolsri, Ngamphol Soonthornworasiri, Wirichada Pan-ngum. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 16.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluating the Usability, Technical Performance, and Accuracy of Artificial Intelligence Scribes for Primary Care: Competitive Analysis

Emily Ha[1,2], MSc; Isabelle Choon-Kon-Yune[2,3], MPH; LaShawn Murray[4], MS; Siying Luan[2], MSc; Enid Montague[4], PhD; Onil Bhattacharyya[2,5,6], MD, PhD; Payal Agarwal[2,6], MSc, MD

[1]Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

[2]Women's College Hospital Institute for Health System Solutions and Virtual Care, Women's College Hospital, 76 Grenville Street, 6th Floor, Toronto, ON, Canada

[3]Institute for Better Health, Trillium Health Partners, Mississauga, ON, Canada

[4]Department of Mechanical & Industrial Engineering, University of Toronto, Toronto, ON, Canada

[5]Institute for Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada

[6]Department of Family and Community Medicine, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Emily Ha, MSc
Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

Abstract

Background: Primary care providers (PCPs) face significant burnout due to increasing administrative and documentation demands, contributing to job dissatisfaction and impacting care quality. Artificial intelligence (AI) scribes have emerged as potential solutions to reduce administrative burden by automating clinical documentation of patient encounters. Although AI scribes are gaining popularity in primary care, there is limited information on their usability, effectiveness, and accuracy.

Objective: This study aimed to develop and apply an evaluation framework to systematically assess the usability, technical performance, and accuracy of various AI scribes used in primary care settings across Canada and the United States.

Methods: We conducted a systematic comparison of a suite of AI scribes using competitive analysis methods. An evaluation framework was developed using expert usability approaches and human factors engineering principles and comprises 3 domains: usability, effectiveness and technical performance, and accuracy and quality. Audio files from 4 standardized patient encounters were used to generate transcripts and SOAP (Subjective, Objective, Assessment, and Plan)–format medical notes from each AI scribe. A verbatim transcript, detailed case notes, and physician-written medical notes for each audio file served as a benchmark for comparison against the AI-generated outputs. Applicable items were rated on a 3-point Likert scale (1=poor, 2=good, 3=excellent). Additional insights were gathered from clinical experts, vendor questionnaires, and public resources to support usability, effectiveness, and quality findings.

Results: In total, 6 AI scribes were evaluated, with notable performance differences. Most AI scribes could be accessed via various platforms (n=4) and launched within common electronic medical records, though data exchange capabilities were limited. Nearly all AI scribes generated SOAP-format notes in approximately 1 minute for a 15-minute standardized encounter (n=5), though documentation time increased with encounter length and topic complexity. While all AI scribes produced good to excellent quality medical notes, none were consistently error-free. Common errors included deletion, omission, and SOAP structure errors. Factors such as extraneous conversations and multiple speakers impacted the accuracy of both the transcript and medical note, with some AI scribes producing excellent notes despite minor transcript issues and vice versa. Limitations in usability, technical performance, and accuracy suggest areas for improvement to fully realize AI scribes' potential in reducing administrative burden for PCPs.

Conclusions: This study offers one of the first systematic evaluations of the usability, effectiveness, and accuracy of a suite of AI scribes currently used in primary care, providing benchmark data for further research, policy, and practice. While AI scribes show promise in reducing documentation burdens, improvements and ongoing evaluations are essential to ensure safe and effective use. Future studies should assess AI scribe performance in real-world settings across diverse populations to support equitable and reliable applications.

(JMIR Hum Factors 2025;12:e71434) doi:10.2196/71434

XSL•FO
RenderX

KEYWORDS

artificial intelligence; AI; AI scribes; clinical documentation; administrative burden; primary care; competitive analysis; evaluation

Introduction

Overview

The majority of primary care providers (PCPs) in Canada and the United States report feelings of burnout, largely attributable to an increase in clinical documentation and administrative tasks [1,2], changes in practice patterns, and workforce shortages [3-5]. Since the widespread implementation of electronic medical records (EMRs) and evolving regulations and policies on medical documentation, PCPs, including family physicians and nurse practitioners, spend nearly half of their workday and an additional 40 hours per month completing administrative tasks after hours, such as clinical documentation of patient encounters and inbox management [6-10]. This increase in administrative burden has decreased provider satisfaction [2,10,11], impacted access to care as some providers have reduced clinical hours to manage administrative workload, and affected the patient-provider relationship [11,12]. Patient experience and satisfaction have also been impacted as PCPs have increasingly spent more time working in their EMR than on direct patient-facing care [12].

Advancements in artificial intelligence (AI) and machine learning have sparked interest in how automated processes can streamline and standardize clinical documentation in primary care. Several perspectives [13,14] and reviews [15,16] have described the need for AI scribes (also known as ambient scribes or digital scribes) to reduce administrative burden in primary care as well as other health care sectors. AI scribes use automatic speech recognition (ASR) and natural language processing (NLP) to record, transcribe, and automate clinical documentation. These tools are generally accessible via web browsers, mobile apps, or desktop applications, with some being integrated into EMRs while others function independently. The workflow typically involves clinicians activating the AI scribe during a patient encounter, where it listens and transcribes conversations and then generates a medical note thereafter. Some AI scribes also have capabilities to fill in billing codes or extract information to support diagnostic decision-making [17].

Several companies have developed AI scribes for use in primary care, promising to streamline workflows and reduce documentation time. However, there remain several concerns, including (1) the impact that AI scribes may have on the workflow of PCPs; (2) the effectiveness of these tools; and (3) the accuracy and quality of their outputs. To date, there are no evaluations that systematically compare the effectiveness of multiple AI scribes currently being used in primary care while also considering factors related to usability, technical performance, and accuracy [18-20]. In addition, given the risks primary care professionals and organizations may assume when adopting new technologies that potentially access, store, and retain personal health information (PHI), a thorough investigation of AI scribes currently being used in primary care is essential to ensure their safety and effectiveness.

Objective

The study aimed to develop and apply an evaluation framework to systematically compare the usability, effectiveness, technical performance, and accuracy and quality of various AI scribes used in primary care settings in Canada and the United States.

Methods

Study Design

We conducted a systematic comparison of AI scribes using competitive analysis methods in the Virtual Care Lab at Women's College Hospital (WCH) in Toronto, Ontario, Canada [21]. Competitive analysis methods are used to evaluate and compare the strengths, weaknesses, opportunities, and threats of competing products within a specific market to provide insights into market dynamics, identify gaps, and inform strategic decisions [22]. The analysis also incorporated expert usability approaches (eg, heuristic evaluation) and principles of human factors engineering [23-26]. Expert usability approaches enable the identification and examination of the current capabilities across several digital health tools to identify areas for improvement or features that are beneficial but may be missing [26]. Human factors engineering examines the interaction between systems, products, and their environment to enhance human performance, safety, and well-being [27]. Combining these methods enabled a comprehensive assessment and comparison of a suite of AI scribes while ensuring that the unique workflows of PCPs, their scope of work, and how they may be interacting with AI scribes within their clinical environments were considered.

The AI scribes included in this study were part of the Clinical Evaluation of Artificial Intelligence and Automation Technology to Reduce Administrative Burden in Primary Care project funded by the Ontario Ministry of Health through a Transfer Payment Agreement between Ontario Health and OntarioMD [28]. Only vendors who were compliant with Ontario's Personal Health Information Protection Act (PHIPA) and adhered to best practices in data privacy and security were selected for participation in the clinical evaluation. These practices included supporting PCPs in ensuring that PHI was collected, used, and disclosed for authorized purposes only; supporting PCPs in securing valid consent for using PHI; implementing stringent safeguards against unauthorized access of PHI; and managing PHI retention and disposal appropriately. Vendors were also required to store and process data within Canada or, if stored outside, to provide adequate notice to health information custodians to support regulatory compliance and notification obligations.

Evaluation Framework

Overview

Although there are instruments to assess physician note quality [29,30], a valid and reliable instrument that evaluates the usability, effectiveness, technical performance, and quality of

XSL•FO
RenderX

the medical notes generated by AI does not exist. We developed an evaluation framework that adapted existing tools related to the usability (eg, user control, flexibility and efficiency of use, aesthetic and design, and help and documentation) [24] and provider documentation quality (eg, comprehensiveness, organization, conciseness, and usefulness) [29-31]. The framework was also informed by medical documentation policies and guidelines developed by governing and regulatory bodies in Canada, including the College of Physicians and Surgeons of Ontario (CPSO) [32] and the Canadian Medical Protective Association (CMPA) [33].

The evaluation framework has 12 items grouped into 3 domains: usability, effectiveness and technical performance, and accuracy and quality (Multimedia Appendix 1). The framework includes both quantitative and qualitative measures to capture various features and functions of AI scribes. For applicable items, a 3-point Likert scale (1=poor, 2=good, 3=excellent) was used to streamline the evaluation process. Definitions for each measure and level of the Likert scale were predefined to ensure consistent application of the evaluation framework.

Usability

Usability refers to the efficiency and satisfaction with which users can complete tasks on a given interface or platform [34]. Key usability factors for health care technologies in health care literature include user-software interaction, task alignment, and navigation [35]. Incorporating these factors, the usability domain in our evaluation framework assessed user interface, EMR compatibility and integration, and process flow of the AI scribes. Compatibility and integration measured the level of integration with common EMR systems in primary care in Ontario, focusing on efficient data exchange, interoperability, and standardized format adherence. User interface refers to the accessibility of each AI scribe across platforms and the graphical and interactive features users interact with. Process flow measured the steps (ie, mouse clicks and keystrokes) and time needed to sign in and launch the AI scribe. All measurements of time were conducted by team members in the same location with a consistent internet connection. Timing was averaged over 3 trials for each AI scribe to account for minor variability and excluded time spent entering credentials or textual information. Both user interface and EMR compatibility were evaluated using the 3-point Likert scale (with 1 representing poor usability or EMR compatibility and 3 representing excellent usability and EMR compatibility).

Effectiveness and Technical Performance

To assess effectiveness, average documentation time was measured as the time required to generate a complete medical note following a standardized 15-minute clinical encounter. This was calculated as the amount of time elapsed from the moment the recording stopped to the time in which a finalized medical note was generated by the AI scribe, excluding any additional edits or review by a PCP. For each AI scribe, documentation time was averaged across 12 appointments (4 standardized encounters, each repeated in 3 trials).

Technical performance refers to the ability of an AI scribe to generate reliable transcripts and medical notes when challenged with complicating factors, such as interruptions, loud background noise, or multiple speakers. Secondary audio files were used to introduce these elements during standardized patient-PCP interactions. Background noises included nonverbal sounds such as typing, paper shuffling, and ambient office sounds, as well as conversational noises from nearby voices. Interruptions and third-party speakers included brief interjections or conversations from other voices, mimicking real-world scenarios where individuals may briefly speak or disrupt the main patient-PCP interaction or if a caregiver is present. Both the transcripts and medical notes generated by each of the AI scribes were evaluated against the 3-point Likert scale (1=poor performance, 2=good performance, 3=excellent performance). “Excellent” performance was defined as accurately attributing statements to the correct speaker, maintaining the logical flow of conversation, and minimizing omissions or distortions. “Good” performance allowed for minor attribution errors or brief lapses that did not meaningfully impact the clinical content. “Poor” performance was characterized by major attribution errors, significant omissions, disorganized conversation flow, or outputs that impaired the clinical utility of the transcript or note.

Accuracy and Quality in Documentation

The accuracy and quality of the medical notes generated by the AI scribes were assessed across 7 items: accuracy, comprehensiveness, care plan, organization, comprehension, conciseness, and usefulness (Textbox 1). These items were informed by medical documentation policies and guidelines developed by the CPSO [32] and the CMPA [33], validated quality frameworks that assessed the accuracy and quality of medical notes written by physicians and medical scribes [30,36], and early evaluations on the accuracy of AI scribes in other health care settings [31]. Each item was measured against a 3-point Likert scale and scores were averaged to determine a single performance score (with 1 representing poor accuracy and quality and 3 representing excellent accuracy and quality). The transcript was also assessed for overall accuracy and quality.

XSL•FO
RenderX

Textbox 1. Description of items to evaluate the accuracy and quality of medical notes generated by an artificial intelligence (AI) scribe.

Accuracy

The medical note contains information that is true and free from errors or hallucinations.

Comprehensiveness

The medical note includes complete documentation of all relevant patient information, including medical history, examination findings, diagnostic results, and treatment plans without omissions.

Care Plan

The medical note provides a holistic understanding of the patient's health status. The medical note allows health care professionals to readily interpret the patient's health status and develop a plan of care.

Organization

The medical note is well-structured, adhering to the SOAP (Subjective, Objective, Assessment, and Plan) format.

Comprehension

The medical note is accessible and devoid of ambiguity or difficult-to-understand terms, phrases, or sections. The medical note allows health care professionals to readily interpret the patient's clinical status and make informed decisions.

Conciseness

The medical note succinctly and effectively conveys all essential information, avoiding unnecessary elaboration or redundancy.

Usefulness

The medical note presents pertinent clinical information in a clear, concise, and actionable manner, facilitating effective communication, decision-making, and continuity of care among health care professionals involved in the patient's care.

We also conducted an error analysis to identify the types of mistakes made by the AI scribes when generating transcripts and medical notes using error classification schemes developed for ASR systems [37-39]. Error types included deletion or omission errors and errors in names, proper nouns, numbers, punctuation, medication names, medical terminology, homonyms, and the Subjective, Objective, Assessment, and Plan (SOAP) format.

Applying the Evaluation Framework

To conduct the competitive analysis, 4 research team members (EH, IC-K-Y, LM, and SL) had an "unlimited" or "premium" license for each AI scribe. Each team member played the audio files from 4 different simulated clinical encounters between a PCP and a standardized patient to generate transcripts and medical notes in the SOAP format from each AI scribe. The audio files were obtained from the College of Family Physicians of Canada Certification Examination in Family Medicine website and were selected because they portrayed common clinical encounters and patient populations observed in primary care. Each audio file also had a verbatim transcript, detailed case notes, and SOAP-format notes written by PCPs, all of which served as a rubric for comparing to the AI scribe-generated outputs. The use of audio files helped minimize discrepancies and challenges related to variations in clinical presentations; thus, providing uniform encounters to each AI scribe and allowing researchers to verify results against the original source for more accurate assessments. The SOAP format was selected for documentation as it was a universally available template across all AI scribe products and is one of the most recommended and widely used methods for documenting a patient encounter in primary care [32].

Team members individually evaluated items in the framework, collaborating weekly to resolve any rating discrepancies through discussion and consensus. If an agreement could not be reached, a fifth team member adjudicated. Clinical experts and family physicians also reviewed each measure, adding insights on usability, effectiveness, technical performance, and output accuracy and quality for each AI scribe. Their feedback was incorporated through a second round of review and consensus meetings among the core evaluation team. In addition, data from a questionnaire completed by the AI scribe vendors about their products' capabilities was analyzed, and publicly available resources, including company websites, product documentation, user manuals, whitepapers, case studies, and user testimonials were reviewed.

Ethical Considerations

The study was formally reviewed by institutional authorities at WCH and received research ethics approval from the WCH Assessment Process for Quality Improvement Projects pathway (APQIP #2023 - 0059). The study was conducted within a controlled simulation environment using audio recordings of standardized clinical encounters between a PCP and a standardized patient. The audio recordings are publicly available to use for educational and research purposes. Results generated as part of this study were based solely on these standardized patient encounters and did not include human participants or the collection of any PHI. All research data were securely stored on password-protected internal servers with access restricted to authorized study personnel.

Results

Overview

A total of 6 AI scribes were selected for evaluation based on their market availability and compliance with data privacy and security regulations in Ontario (ie, PHIPA). To highlight the general capabilities of AI scribes, specific product names are

XSL•FO
RenderX

not used; however, when referring to a particular AI scribe product, they are labeled as AI scribes #1 through #6. Table 1 provides an overview of the health care sectors and users for each of the AI scribes. While 2 AI scribes were exclusively used in primary care (AI scribes #1 and #4), the remaining 4 were also deployed in other health care sectors, including acute care, home care, palliative care, and community services. Of these 4 AI scribes, 3 were used by a broader range of health care professionals, including nurses, pharmacists, dentists, veterinarians, and administrative staff, in addition to physicians. The number of active, paid subscribers for each AI scribe varied significantly, from 20 for AI scribe #4 to over 21,000 users for AI scribe #6. Half of the AI scribes supported languages other than English (AI scribes #2, #3, and #5), while the remaining 3 only supported English but reported that they were developing additional language options (eg, French and Spanish). The performance of these AI scribes in non-English languages was not evaluated in this analysis.

Table . Overview of supported health care settings and users across different artificial intelligence (AI) scribe products.

Context	AI scribes					
	AI scribe #1	AI scribe #2	AI scribe #3	AI scribe #4	AI scribe #5	AI scribe #6
Care settings[a]	Primary care	Primary care, acute and ambulatory care, social and community services, home care, palliative care, mental health and psychiatry, pediatrics, rehabilitation services, dental care, optometry, surgery, and veterinary medicine	Primary care, acute and ambulatory care, palliative care, and home care	Primary care	Primary care, acute and ambulatory care, community services, and mental health and psychiatry	Primary care and acute and ambulatory care
Main users[b]	Physicians	Physicians, nurses, other regulated health care professionals[c], and administrative staff	Physicians, nurses, other regulated health care professionals[c], and administrative staff	Physicians	Physicians, nurses, and other regulated health care professionals[c]	Physicians
Number of users[d]	100	13,000	360	20	4000	21,000
Supported languages	English	English, French, and 50 other languages	English, French, and 20 other languages	English	English and French	English

[a]All health care sectors that the AI scribe was used in.

[b]Types of clinicians and health care professionals that the AI scribe was marketed to, and who had an active, paid subscription to use the AI scribe (ie, users).

[c]Regulated health professionals include pharmacists, dentists, occupational therapists, physiotherapists, social workers, psychologists, and therapists.

[d]Approximate number of current users with an active, paid subscription for the AI scribe.

Usability

Table 2 presents the results of the usability assessment. Of the 6 AI scribes evaluated, 4 were accessible on multiple platforms (desktop or laptop and mobile or tablet devices); thus, receiving an excellent form factor rating. In contrast, AI scribes #4 and #6 could only be accessed on a single platform, potentially limiting accessibility in diverse clinical settings. Four of the AI scribes also demonstrated some level of EMR integration. AI scribes #1 and #3 exhibited the highest degree of integration as they were compatible with multiple EMRs used in primary care in Canada and could exchange metadata (eg, schedules, visit details, previous notes, and summaries) and patient information (eg, name, date of birth, and health card number) when launched from within the EMR. AI scribe #5 could be launched within a single EMR, but no data exchange was available. AI scribe #6 did not connect directly to common EMRs in primary care but allowed users to transfer notes through a separate application. To add the medical note to the EMR, only AI scribe #5 offered a seamless drag-and-drop note transfer into the EMR with a single mouse click. For AI scribes #1 - #4, users would have to toggle between the AI scribe and EMR interfaces to manually copy-paste notes, adding to the number of mouse clicks or keystrokes required to use the AI scribe. AI scribe #6 used a separate program to transfer the transcript and medical note from the mobile app to the EMR.

XSL•FO
RenderX

Table . Comparison of usability features across different artificial intelligence (AI) scribe products.

Usability measures	AI scribes					
	AI scribe #1	AI scribe #2	AI scribe #3	AI scribe #4	AI scribe #5	AI scribe #6
User interface						
Main platforms used to access the AI scribe[a]	Supported on Google Chrome, Mozilla Firefox, Apple Safari, and Microsoft Edge	Supported on all web browsers	Supported on Google Chrome, Mozilla Firefox, and Microsoft Edge	Desktop application	Google Chrome extension but also available on other Web browsers, although not formally supported	Mobile app
Supported mobile and tablet devices	Mobile-friendly via phone's web browser	Mobile-friendly via phone's web browser	Mobile-friendly via phone's web browser	N/A	Native application for iOS and Android devices	Native application available for iOS and Android devices
Form factor[b]	Excellent	Excellent	Excellent	Poor	Excellent	Poor
EMR integration						
EMR integration	Excellent	Poor	Excellent	Poor	Excellent	Good
Ability to transfer medical note	Good	Good	Good	Good	Excellent	Poor
Process flow						
Steps to sign-in and launch the AI scribe	5	7	6	6	4	4
Average time to sign-in and launch the AI scribe[c]	10.01 sec[d]	13.58 sec	26.36 sec	14.59 sec	11.63 sec	7.21 sec
Factor authentication (FA)	1FA[e]	1FA	2FA[f]	1FA	2FA	1FA
Ease of Restarting	Good	Good	Good	Good	Good	Good

[a]Medium or technology through which users can access and interact with the AI scribe, considering the various types of interfaces and devices that enable user access.

[b]Design and accessibility of platforms required to use the AI scribe, particularly in relation to its ease of access in primary care settings.

[c]Approximate amount of time needed to sign in and launch the AI scribe for use during a patient encounter. This measure was averaged across 3 runs and did not include time spent typing when entering credentials or textual information.

[d]sec: seconds.

[e]1FA: single-factor authentication.

[f]2FA: 2-factor authentication.

Other processes affecting usability, such as sign-in and launch, varied across products. All AI scribes required user authentication for sign-in, with 4 AI scribes using single-factor authentication and 2 implementing 2-factor authentication. There did not appear to be a link between the number of steps and the amount of time it took to sign in and launch the AI scribe. The encounter initiation processes also varied across AI scribes; some required preloading appointment data (eg, patient name and appointment time), while others could start recording immediately. Each AI scribe featured a visible "start encounter" button, and all except AI scribe #4 provided live transcription. Transitioning between patient encounters required users to save the medical note before moving on to the next patient.

Effectiveness and Technical Performance

On average, for a 15-minute clinical encounter between a patient and PCP, nearly all AI scribes generated a SOAP-format medical note within 1 minute of pressing the stop recording button, with AI scribe #4 taking almost twice as long (Table 3). Documentation time for all AI scribes was affected by the encounter length and the complexity of the topics discussed. Encounters involving multiple or complex topics, or extended, nonlinear conversations—including rapport-building exchanges not directly related to medical care—generally increased the documentation time of the AI scribes.

XSL•FO
RenderX

Table . Effectiveness and technical performance across different artificial intelligence (AI) scribe products.

Performance measures	AI scribes					
	AI scribe #1	AI scribe #2	AI scribe #3	AI scribe #4	AI scribe #5	AI scribe #6
Effectiveness						
Average documentation time	41.87 sec[a]	53.54 sec	42.8 sec	1:40 min[b]	32.4 sec	20.66 sec
Technical performance						
Background noise	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent
Interruptions	Excellent	Good	Excellent	Good	Good	Poor
Multiple speakers	Good	Poor	Good	Good	Poor	Good

[a]sec: seconds.

[b]min: minutes.

The performance of the AI scribes also varied when faced with complicating factors. All AI scribes successfully omitted nonconversational background noises. In contrast, only AI scribe #1 effectively filtered out interruptions and extraneous conversations, ensuring only the patient-PCP dialogue appeared in the transcript and medical note. AI scribes #2 - #5 included brief phrases from interruptions or extraneous conversations in the transcript but excluded them from the medical note. AI scribe #6, however, mistakenly incorporated information from the interruption into both the transcript and the medical note.

The ability of the AI scribes to manage conversations among 3 or more speakers also varied. AI scribes #1, #3, #4, and #6 received a “good” rating, effectively distinguishing between speakers and accurately assigning dialogue to speakers in the transcripts. In contrast, AI scribes #2 and #5 received a “poor” rating, often failing to differentiate between more than 2 speakers when generating the transcript. These errors led to inaccuracies in the medical note, where statements made by a third speaker were incorrectly attributed to the patient.

Accuracy and Quality in Documentation

Table 4 summarizes the evaluation findings on documentation accuracy and quality. None of the AI scribes consistently produced fully accurate and error-free transcripts or medical notes, although the types of errors varied. While most patient-PCP conversations were transcribed accurately, frequent grammatical and syntactical errors, such as repeated or omitted words, were observed, particularly among AI scribes #1 and #3 - #5. These AI scribes received an overall rating of “good” for transcript accuracy and quality, as these issues could affect the transcripts’ usefulness, making it harder for users to reference them quickly during later reviews, although it did not affect overall content.

XSL•FO
RenderX

Table . Evaluation of the accuracy and quality of the transcript and medical note generated by the artificial intelligence (AI) scribe.

Accuracy measures	AI scribe					
	AI scribe #1	AI scribe #2	AI scribe #3	AI scribe #4	AI scribe #5	AI scribe #6
Accuracy and quality						
Transcript	Good	Excellent	Good	Good	Good	Excellent
Medical note	Excellent	Excellent	Excellent	Good	Good	Good
Error analysis[a,b]						
Deleted words or omissions[c]	✓	✓	✓	✓	✓	✓
Generalizations[d]				✓	✓	✓
SOAP[e] structure[f]		✓	✓	✓		

[a]Checkmarks indicate that at least one occurrence of the error was observed.

[b]Error types not observed include errors in names, proper nouns, numbers, punctuation, medication names, or homonyms.

[c]Error where essential information is left out or not included, which can lead to incomplete understanding or documentation of the intended content or context.

[d]Error where specific information is simplified or replaced with a broader, less precise term, potentially leading to a loss of detail or nuance that could affect the accuracy or clarity of the information.

[e]SOAP: Subjective, Objective, Assessment, or Plan.

[f]Documentation error where information is incorrectly categorized within the SOAP sections of a medical note, potentially leading to confusion or misinterpretation of the clinical details.

For medical note accuracy and quality, 3 AI scribes (#4 - #6) produced notes rated as "good" quality, while the remaining 3 (AI scribes #1 - #3) produced "excellent" quality notes. An "excellent" rating indicated strong relative performance among the evaluated AI scribes rather than complete accuracy. The "excellent" notes were comprehensive, well-organized, and free from hallucinations or contradictory information, providing sufficient detail to accurately reflect the patient's clinical status and support informed decision-making. Interestingly, AI scribes that generated good-quality transcripts were still capable of producing excellent-quality medical notes (AI scribes #1 and #3), and the reverse was also true (AI scribe #6).

In the error analysis, errors involving names, proper nouns, numbers, punctuation, medication names, or homonyms were not observed. However, deletion or omission errors were observed across all AI scribes. The most frequent omissions involved details about the history of the presenting illness, social history, or lifestyle factors, with the extent of omissions varying across AI scribes. For example, in one clinical scenario, AI scribes #2, #3, and #4 omitted dietary information relevant to the encounter but still received a good rating, as this information could have been documented elsewhere in the patient's chart. Conversely, AI scribe #6 consistently omitted new social history details directly relevant to the presenting illness. AI scribes #4 - #6 also made generalization errors, simplifying medical terms in ways that lost specificity or nuance, such as reducing "Colles fracture" to "fracture" or even "dog fracture." In addition, AI scribes #2 - #4 displayed SOAP structure errors, such as misclassifying information in the Objective section despite the absence of a physical examination.

Discussion

Principal Findings

We conducted one of the first studies, to our knowledge, that systematically evaluated the usability, effectiveness, technical performance, and accuracy of 6 AI scribes developed for or used in primary care settings in Canada, including some tools that are also used in the United States. By leveraging competitive analysis methods, expert usability approaches, and human factors engineering, we developed a comprehensive evaluation framework that revealed a rapidly evolving AI scribe market and notable differences in performance across AI scribes.

Of the AI scribes included in this evaluation, most were accessible across multiple platforms, including desktops, laptops, mobile devices, and tablets, and could be launched directly within the EMR for easier access, although data exchange was minimal. In general, nearly all AI scribes generated a SOAP-format medical note in approximately 1 minute for a standardized 15-minute encounter, although documentation length increased based on the length of the encounter and complexity of topics discussed. While all AI scribes produced good to excellent quality medical notes, persistent deletion and omission errors highlighted the continued need for review and editing by PCPs. In addition, several factors were found to affect the quality and accuracy of both transcripts and medical notes, including extraneous conversations, encounters with multiple speakers, and the complexity of conditions discussed. Interestingly, AI scribes that produced good-quality transcripts could generate excellent-quality medical notes, and conversely, transcripts of excellent quality did not always result in equally high-quality medical notes. Despite these promising capabilities, limitations related to usability, technical performance, and

XSL•FO
RenderX

accuracy remain that must be addressed to ensure AI scribes continue to reduce administrative burdens effectively.

Comparison With Other Work

While studies on the usability of AI scribes in primary care remain limited, our findings are consistent with broader research on the usability of digital health technologies. For example, inadequate access, poor usability of platforms, and lack of workflow integration are known factors that contribute to PCP burnout, whereas enhanced accessibility and robust EMR integration are associated with improved PCP well-being and productivity [40,41]. As demonstrated in this study, flexible access to AI scribes across multiple platforms—such as mobile devices, tablets, and desktops—may enhance their usability by aligning better with PCPs' varied workflows and preferences. In addition, as highlighted in several commentaries, AI scribes remain an unregulated digital health tool [42,43]. While the current level of EMR integration may be sufficient in the absence of regulation, data privacy and security laws, like those in Ontario, Canada, place the responsibility on PCPs and institutions to understand privacy implications, assess risks, and ensure that AI scribe software functions as expected [42,43]. These considerations underscore the need for regulatory guidance and transparency in the development of AI scribe technology to support usability, enhance EMR integration, and maintain patient safety in primary care settings.

Studies on the accuracy and quality of transcripts and medical notes produced by AI scribes are emerging. For instance, Tierney et al [31] used a modified 9-item Physician Documentation Quality Instrument (PDQI-9) to evaluate the quality of notes produced by an AI scribe used in primary care, pediatric, hospital, mental health, surgical, and emergency care settings. Their regional pilot demonstrated that while the AI scribe produced high-quality notes, minor hallucinations and omission errors were common—similar to the error types observed in our study. In addition, a study on a commercially available Dutch AI scribe found that fully automated notes, without clinician editing, had poorer PDQI-9 scores, higher word counts, and reduced lexical diversity compared to those reviewed by clinicians [44]. Our evaluation of 6 AI scribes expanded upon this work and demonstrated that various AI scribes can produce good to excellent quality notes, though none were entirely error-free and likely require some editing by users. Compared with the PDQI-9 and other tools used to evaluate note quality, our framework consists of 7 items that focus on characteristics unique to AI-generated outputs, such as hallucinations, redundancy, and bias. Our framework also considers the ability of AI-generated notes to capture relevant patient information in a concise and actionable manner, with a particular emphasis on attributes like the Care Plan. It is designed to be suitable for SOAP-format medical notes, though it can be adapted to other formats, highlighting both similarities and differences in focus and scope.

Despite the promise of AI scribes, recent evaluations of large language models (LLMs), such as GPT-3.5 and Open AI's ChatGPT-4, urge caution before integrating AI-powered tools into clinical documentation workflows. In these studies, LLMs were fed medical evidence or transcripts of audio-recorded patient-provider conversations to generate summaries [45,46]. Findings revealed that LLMs frequently produce factually inconsistent summaries, with error types varying across replicates of the same case [45]. Performance also differed significantly across different LLMs [46]. In contrast to general purpose LLMs, such as ChatGPT, the AI scribes evaluated in our study primarily relied on conventional pipelines combining ASR systems and NLP models that were specifically fine-tuned for clinical documentation tasks. While some AI scribe vendors may be beginning to integrate LLM-based approaches, most tools included in this evaluation were optimized separately for ASR and NLP functions. Notably, the distinct functions of ASR systems and NLP within AI scribes likely explain why 'excellent' transcripts may sometimes accompany 'good' medical notes, or vice versa, as observed in our study. ASR systems are optimized to capture spoken words accurately, while NLP focuses on summarizing, organizing, and emphasizing clinically relevant information [14]. Consequently, some AI scribes may excel in extracting key details for notes despite minor transcript inaccuracies, while others may produce precise transcripts but lack effective summarization. A key takeaway is that the performance of AI scribes is expected to vary depending on the specific ASR and NLP systems used, and performance will likely change over time as these models are retrained on new data. The variability in performance within and across AI scribes underscores the need for ongoing evaluation and oversight to ensure that these tools continue to meet clinical documentation needs reliably [47]. Collectively, these findings highlight the importance of PCPs reviewing and editing AI-generated medical notes to ensure documentation accuracy and quality.

Overall, our study has important implications for research, policy, and practice. For research, the findings highlight the need for further large-scale studies to evaluate the usability, accuracy, and long-term impact of AI scribes, especially in real-world clinical environments. This research also raises important equity considerations, as the performance of various AI scribes may differ across PCP and patient populations or when used to support languages other than English [46,48]. For policy and practice, our findings provide benchmark data on the current performance of 6 AI scribes used in primary care settings, guiding users, institutions, and other stakeholders on both the strengths and areas needing improvement. This work can also set the foundation for broader implementation strategies that balance innovation with safety, equity, and accountability in digital health technology. Furthermore, there is a critical need for the ongoing evaluation of AI scribes and other AI tools in a manner that is both timely and rigorous. With the rapid evolution of AI technology, continuous assessment is necessary to ensure that these tools remain relevant, effective, and safe for diverse clinical contexts. Evaluations must be methodologically robust to address potential biases, assess long-term impacts, and ensure that the tools meet the needs of all patient and provider populations equitably to build trust, maintain accountability, and maximize the transformative potential of AI in health care.

XSL•FO
RenderX

Limitations

There are limitations to consider when interpreting findings. First, the evaluation was conducted in a controlled laboratory setting, which may not fully capture the nuances and complexities of real-world clinical environments. While we used standardized audio files representing typical primary care encounters, these simulations cannot entirely replicate live, unscripted interactions. It is also possible that some AI scribes were partially trained or fine-tuned using publicly available clinical audio datasets, which could introduce overlap with the types of audio files used in our evaluation. Although we cannot confirm the specific training sources used by each vendor (as this information was often proprietary), we acknowledge this as a limitation and recognize that model performance may differ when evaluated using de novo, real-world clinical encounter data. Second, we evaluated 6 AI scribes, although there are a countless number of AI scribes available globally. While the selected vendors reflect those used in primary care settings in Canada and the United States, the results may not capture the full range of features offered by all AI scribes. In addition, AI scribes and automation technologies are rapidly evolving. Although the evaluation framework may be applied to continuously evaluate AI scribes, as the underlying ASR and NLP models are retrained, findings represent the relative performance of each AI scribe at a single point in time and may not reflect any subsequent versions released since the time of the study. Finally, we assessed documentation quality using the SOAP format to ensure consistent evaluation of outputs across all AI scribes that were evaluated. While this may not represent the variety of templates or the specialized expertise required to select the most appropriate template for different clinical scenarios, the SOAP format remains the most widely recommended method for documenting clinical encounters in primary care, and its use in this study provided a consistent basis for comparison.

Conclusion

In conclusion, our findings demonstrate that developing a comprehensive framework to evaluate AI scribes is feasible, and findings from the competitive analysis underscore the value of AI scribes as a promising tool to alleviate administrative burden in primary care. However, findings also emphasize the need for ongoing enhancements to improve the usability, technical performance, and accuracy of AI scribes. Although AI scribes can generate good to excellent quality medical notes, user editing and review remain essential. Greater EMR integration and personalized customization are also needed to align with the varied documentation practices of PCPs. Overall, ongoing evaluation is crucial to ensure these technologies effectively support PCPs while enhancing safety, accuracy, and efficiency for both providers and patients. Future research should also explore AI scribe performance in dynamic, real-world settings.

Acknowledgments

The authors would like to thank the Ontario Ministry of Health for funding this study and OntarioMD and Amplify Care (formerly eHealth Centre of Excellence) for supporting this study and facilitating access to the artificial intelligence (AI) scribes. The funders had no role in study design, data collection, and analysis. PA is a recipient of the 2023 AMS Healthcare Fellowship in Compassion and AI.

Data Availability

The data generated and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

EH, EM, OB, and PA contributed to the conceptualization of the study. EH, IC-K-Y, LM, and SL were responsible for data curation and formal analysis. Funding acquisition was handled by OB and PA. EH, IC-K-Y, LM, and SL carried out the investigation. EH, LM, EM, OB, and PA contributed to the development of the methodology. EH managed the project administration. OB and PA provided the necessary resources. Software development was carried out by EH, IC-K-Y, LM, and SL. Supervision and validation were conducted by EM, OB, and PA. EH, IC-K-Y, LM, and SL contributed to data visualization. EH prepared the original draft of the manuscript. All authors contributed to the review and editing of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Overview of evaluation framework domains, items, and assessment method.
[DOCX File, 21 KB - humanfactors_v12i1e71434_app1.docx]

References

1. CMA 2021 national physician health survey. : Canadian Medical Protective Association; 2022 Aug URL: https://digitallibrary.cma.ca/link/digitallibrary17 [accessed 2025-07-16]

XSL•FO
RenderX

2. Shanafelt TD, West CP, Sinsky C, et al. Changes in burnout and satisfaction with work-life integration in physicians and the general US working population between 2011 and 2017. Mayo Clin Proc 2019 Sep;94(9):1681-1694. [doi: 10.1016/j.mayocp.2018.10.023] [Medline: 30803733]
3. Kralj B, Islam R, Sweetman A. Long-term trends in the work hours of physicians in Canada. CMAJ 2024 Mar 24;196(11):E369-E376. [doi: 10.1503/cmaj.231166] [Medline: 38527745]
4. Schultz SE, Glazier RH. Identification of physicians providing comprehensive primary care in Ontario: a retrospective analysis using linked administrative data. CMAJ Open 2017 Dec 19;5(4):E856-E863. [doi: 10.9778/cmajo.20170083] [Medline: 29259018]
5. Changes in practice patterns of family physicians in canada. CIHI. URL: https://www.cihi.ca/en/changes-in-practice-patterns-of-family-physicians-in-canada [accessed 2025-06-16]
6. New survey shows full-blown crisis in family medicine: urgent need to address red tape and administrative burden. Ontario College of Family Physicians. 2023. URL: https://ontariofamilyphysicians.ca/news/new-survey-shows-full-blown-crisis-in-family-medicine/ [accessed 2025-06-16]
7. Tai-Seale M, Olson CW, Li J, et al. Electronic health record logs indicate that physicians split time evenly between seeing patients and desktop medicine. Health Aff (Millwood) 2017 Apr 1;36(4):655-662. [doi: 10.1377/hlthaff.2016.0811] [Medline: 28373331]
8. Sinsky C, Colligan L, Li L, et al. Allocation of physician time in ambulatory practice: a time and motion study in 4 specialties. Ann Intern Med 2016 Dec 6;165(11):753-760. [doi: 10.7326/M16-0961] [Medline: 27595430]
9. Arndt BG, Beasley JW, Watkinson MD, et al. Tethered to the EHR: primary care physician workload assessment using EHR event log data and time-motion observations. Ann Fam Med 2017 Sep;15(5):419-426. [doi: 10.1370/afm.2121] [Medline: 28893811]
10. Woolhandler S, Himmelstein DU. Administrative work consumes one-sixth of U.S. physicians' working hours and lowers their career satisfaction. Int J Health Serv 2014;44(4):635-642. [doi: 10.2190/HS.44.4.a] [Medline: 25626223]
11. Friedberg MW, Chen PG, Van Busum KR, et al. Factors affecting physician professional satisfaction and their implications for patient care, health systems, and health policy. Rand Health Q 2014;3(4):1. [doi: 10.7249/RB9740] [Medline: 28083306]
12. Farber NJ, Liu L, Chen Y, et al. EHR use and patient satisfaction: what we learned. J Fam Pract 2015 Nov;64(11):687-696. [Medline: 26697540]
13. Quiroz JC, Laranjo L, Kocaballi AB, et al. Challenges of developing a digital scribe to reduce clinical documentation burden. NPJ Digit Med 2019;2(1):114. [doi: 10.1038/s41746-019-0190-1] [Medline: 31799422]
14. Coiera E, Kocaballi B, Halamka J, et al. The digital scribe. NPJ Digit Med 2018;1(1):58. [doi: 10.1038/s41746-018-0066-9] [Medline: 31304337]
15. Ghatnekar S, Faletsky A, Nambudiri VE. Digital scribe utility and barriers to implementation in clinical practice: a scoping review. Health Technol (Berl) 2021;11(4):803-809. [doi: 10.1007/s12553-021-00568-0] [Medline: 34094806]
16. Falcetta FS, de Almeida FK, Lemos JCS, et al. Automatic documentation of professional health interactions: a systematic review. Artif Intell Med 2023 Mar;137:102487. [doi: 10.1016/j.artmed.2023.102487] [Medline: 36868684]
17. van Buchem MM, Boosman H, Bauer MP, et al. The digital scribe in clinical practice: a scoping review and research agenda. NPJ Digit Med 2021 Mar 26;4(1):57. [doi: 10.1038/s41746-021-00432-5] [Medline: 33772070]
18. Haberle T, Cleveland C, Snow GL, et al. The impact of nuance DAX ambient listening AI documentation: a cohort study. J Am Med Inform Assoc 2024 Apr 3;31(4):975-979. [doi: 10.1093/jamia/ocae022] [Medline: 38345343]
19. Crampton NH. Ambient virtual scribes: Mutuo Health's AutoScribe as a case study of artificial intelligence-based technology. Healthc Manage Forum 2020 Jan;33(1):34-38. [doi: 10.1177/0840470419872775] [Medline: 31522566]
20. Cao DY, Silkey JR, Decker MC, et al. Artificial intelligence-driven digital scribes in clinical documentation: pilot study assessing the impact on dermatologist workflow and patient encounters. JAAD Int 2024 Jun;15:149-151. [doi: 10.1016/j.jdin.2024.02.009] [Medline: 38571698]
21. Women's College Hospital launches first of its kind virtual care laboratory. Hospital WC. URL: https://www.womenscollegehospital.ca/womens-college-hospital-launches-first-of-its-kind-virtual-care-laboratory/ [accessed 2025-06-16]
22. Fleisher CS, Bensoussan BE. Strategic and Competitive Analysis: Methods and Techniques for Analyzing Business Competition: Prentice Hall; 2003.
23. Rosenbaum S. Usability evaluations versus usability testing: when and why? IEEE Trans Profess Commun 1989;32(4):210-216. [doi: 10.1109/47.44533]
24. Nielsen J. Heuristic evaluation. In: Usability Inspection Methods: Wiley; 1994:25-62. [doi: 10.1201/9780849375477-633]
25. Gray WD, Salzman MC. Damaged merchandise? A review of experiments that compare usability evaluation methods. Hum-Comput Interact 1998 Sep;13(3):203-261. [doi: 10.1207/s15327051hci1303_2]
26. Jaspers MWM. A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence. Int J Med Inform 2009 May;78(5):340-353. [doi: 10.1016/j.ijmedinf.2008.10.002] [Medline: 19046928]
27. Salvendy G, Karwowski W. Handbook of Human Factors and Ergonomics: John Wiley & Sons; 2021:28-56.
28. Women's College Hospital Institute for Health System Solutions and Virtual Care. Clinical evaluation of artificial intelligence and automation technology to reduce administrative burden in primary care. : Women's College Hospital; 2024 Jul.

XSL•FO
RenderX

29. Stetson PD, Morrison FP, Bakken S, et al. Preliminary development of the physician documentation quality instrument. J Am Med Inform Assoc 2008;15(4):534-541. [doi: 10.1197/jamia.M2404] [Medline: 18436914]
30. Stetson PD, Bakken S, Wrenn JO, et al. Assessing electronic note quality using the Physician Documentation Quality Instrument (PDQI-9). Appl Clin Inform 2012;3(2):164-174. [doi: 10.4338/aci-2011-11-ra-0070] [Medline: 22577483]
31. Tierney AA, Gayre G, Hoberman B, et al. Ambient artificial intelligence scribes to alleviate the burden of clinical documentation. NEJM Catalyst 2024 Feb 21;5(3). [doi: 10.1056/CAT.23.0404]
32. Medical records documentation. College of Physicians and Surgeons of Ontario. 2020. URL: https://www.cpso.on.ca/Physicians/Policies-Guidance/Policies/Medical-Records-Documentation [accessed 2025-06-16]
33. Documentation and record keeping. Canadian Medical Protective Association. URL: https://www.cmpa-acpm.ca/en/education-events/good-practices/physician-patient/documentation-and-record-keeping [accessed 2025-06-16]
34. Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs) — Part 11: Guidance on Usability: International Organization for Standardization; 1998.
35. Hollin I, Griffin M, Kachnowski S. How will we know if it's working? A multi-faceted approach to measuring usability of a specialty-specific electronic medical record. Health Informatics J 2012 Sep;18(3):219-232. [doi: 10.1177/1460458212437008] [Medline: 23011817]
36. Podder V, Law V, Ghassemzadeh S. SOAP Notes: StatPearls Publishing; 2024. URL: https://pubmed.ncbi.nlm.nih.gov/29489268/ [Medline: 29489268]
37. Zafar A, Mamlin B, Perkins S, et al. A simple error classification system for understanding sources of error in automatic speech recognition and human transcription. Int J Med Inform 2004 Sep;73(9-10):719-730. [doi: 10.1016/j.ijmedinf.2004.05.008] [Medline: 15325329]
38. Goss FR, Zhou L, Weiner SG. Incidence of speech recognition errors in the emergency department. Int J Med Inform 2016 Sep;93:70-73. [doi: 10.1016/j.ijmedinf.2016.05.005] [Medline: 27435949]
39. Blackley SV, Schubert VD, Goss FR, et al. Physician use of speech recognition versus typing in clinical documentation: a controlled observational study. Int J Med Inform 2020 Sep;141:104178. [doi: 10.1016/j.ijmedinf.2020.104178] [Medline: 32521449]
40. Carayon P, Hoonakker P. Human factors and usability for health information technology: old and new challenges. Yearb Med Inform 2019 Aug;28(1):71-77. [doi: 10.1055/s-0039-1677907] [Medline: 31419818]
41. Borges do Nascimento IJ, Abdulazeem HM, Vasanthan LT, et al. The global effect of digital health technologies on health workers' competencies and health workplace: an umbrella review of systematic reviews and lexical-based and sentence-based meta-analysis. Lancet Digit Health 2023 Aug;5(8):e534-e544. [doi: 10.1016/S2589-7500(23)00092-4] [Medline: 37507197]
42. AI scribes: answers to frequency asked questions. Canadian Medical Protective Association. 2023. URL: https://www.cmpa-acpm.ca/en/advice-publications/browse-articles/2023/ai-scribes-answers-to-frequently-asked-questions [accessed 2025-06-16]
43. Agarwal P, Lall R, Girdhari R. Artificial intelligence scribes in primary care. CMAJ 2024 Sep 15;196(30):E1042. [doi: 10.1503/cmaj.240363] [Medline: 39284604]
44. van Buchem MM, Kant IMJ, King L, et al. Impact of a digital scribe system on clinical documentation time and quality: usability study. JMIR AI 2024 Sep 23;3:e60020. [doi: 10.2196/60020] [Medline: 39312397]
45. Tang L, Sun Z, Idnay B, et al. Evaluating large language models on medical evidence summarization. npj Digit Med ;6(1):158. [doi: 10.1038/s41746-023-00896-7]
46. Kernberg A, Gold JA, Mohan V. Using ChatGPT-4 to create structured medical notes from audio recordings of physician-patient encounters: comparative study. J Med Internet Res 2024 Apr 22;26:e54419. [doi: 10.2196/54419] [Medline: 38648636]
47. Meskó B, Topol EJ. The imperative for regulatory oversight of large language models (or generative AI) in healthcare. NPJ Digit Med 2023 Jul 6;6(1):120. [doi: 10.1038/s41746-023-00873-0] [Medline: 37414860]
48. Tran BD, Mangu R, Tai-Seale M, et al. Automatic speech recognition performance for digital scribes: a performance comparison between general-purpose and specialized models tuned for patient-clinician conversations. ANIA Annu Symp Proc 2023:1072-1080. [Medline: 37128439]

Abbreviations

AI: artificial intelligence
APQIP: Assessment Process for Quality Improvement Projects
ASR: automatic speech recognition
CMPA: Canadian Medical Protective Agency
CPSO: College of Physicians and Surgeons of Ontario
EMR: electronic medical record
LMM: large language model
NLP: natural language processing
PCP: primary care provider

XSL•FO
RenderX

PDQI-9: 9-item Physician Documentation Quality Instrument
PHI: personal health information
PHIPA: Personal Health Information Protection Act
SOAP: Subjective, Objective, Assessment, and Plan
WCH: Women's College Hospital

Edited by K Adapa; submitted 17.01.25; peer-reviewed by D Chrimes, J Zeng; revised version received 05.05.25; accepted 21.05.25; published 23.07.25.

Please cite as:
Ha E, Choon-Kon-Yune I, Murray L, Luan S, Montague E, Bhattacharyya O, Agarwal P
Evaluating the Usability, Technical Performance, and Accuracy of Artificial Intelligence Scribes for Primary Care: Competitive Analysis
JMIR Hum Factors 2025;12:e71434
URL: https://humanfactors.jmir.org/2025/1/e71434
doi:10.2196/71434

© Emily Ha, Isabelle Choon-Kon-Yune, LaShawn Murray, Siying Luan, Enid Montague, Onil Bhattacharyya, Payal Agarwal. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 23.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Clinician Perspectives of a Magnetic Resonance Imaging–Based 3D Volumetric Analysis Tool for Neurofibromatosis Type 2–Related Schwannomatosis: Qualitative Pilot Study

Shelby T Desroches[1], MS; Alice Huang[1], MS; Rithvik Ghankot[1], MS; Steven M Tommasini[1], PhD; Daniel H Wiznia[1,2], MD; Frank D Buono[3], PhD

[1]Department of Orthopaedics and Rehabilitation, Yale School of Medicine, New Haven, CT, United States

[2]Department of Mechanical Engineering, Yale University, New Haven, CT, United States

[3]Department of Psychiatry, Yale School of Medicine, 300 George Street, New Haven, CT, United States

Corresponding Author:
Frank D Buono, PhD
Department of Psychiatry, Yale School of Medicine, 300 George Street, New Haven, CT, United States

Abstract

Background: Accurate monitoring of tumor progression is crucial for optimizing outcomes in neurofibromatosis type 2–related schwannomatosis. Standard 2D linear analysis on magnetic resonance imaging is less accurate than 3D volumetric analysis, but since 3D volumetric analysis is time-consuming, it is not widely used. To shorten the time required for 3D volumetric analysis, our lab has been developing an automated artificial intelligence–driven 3D volumetric tool.

Objective: The objective of the study was to survey and interview clinicians treating neurofibromatosis type 2–related schwannomatosis to understand their views on current 2D analysis and to gather insights for the design of an artificial intelligence–driven 3D volumetric analysis tool.

Methods: Interviews examined for the following themes: (1) shortcomings of the currently used linear analysis, (2) utility of 3D visualizations, (3) features of an interactive 3D modeling software, and (4) lack of a gold standard to assess the accuracy of 3D volumetric analysis. A Likert scale questionnaire was used to survey clinicians' levels of agreement with 25 statements related to 2D and 3D tumor analyses.

Results: A total of 14 clinicians completed a survey, and 12 clinicians were interviewed. Specialties ranged across neurosurgery, neuroradiology, neurology, oncology, and pediatrics. Overall, clinicians expressed concerns with current linear techniques, with clinicians agreeing that linear measurements can be variable with the possibility of two different clinicians calculating 2 different tumor sizes (mean 4.64, SD 0.49) and that volumetric measurements would be more helpful for determining clearer thresholds of tumor growth (mean 4.50, SD 0.52). For statements discussing the capabilities of a 3D volumetric analysis and visualization software, clinicians expressed strong interest in being able to visualize tumors with respect to critical brain structures (mean 4.36, SD 0.74) and in forecasting tumor growth (mean 4.77, SD 0.44).

Conclusions: Clinicians were overall in favor of the adoption of 3D volumetric analysis techniques for measuring vestibular schwannoma tumors but expressed concerns regarding the novelty and inexperience surrounding these techniques. However, clinicians felt that the ability to visualize tumors with reference to critical structures, to overlay structures, to interact with 3D models, and to visualize areas of slow versus rapid growth in 3D would be valuable contributions to clinical practice. Overall, clinicians provided valuable insights for designing a 3D volumetric analysis tool for vestibular schwannoma tumor growth. These findings may also apply to other central nervous system tumors, offering broader utility in tumor growth assessments.

(JMIR Hum Factors 2025;12:e71728) doi:10.2196/71728

KEYWORDS

neurofibromatosis type 2; vestibular schwannoma; 3D volumetric analysis; magnetic resonance imaging

Introduction

Neurofibromatosis type 2 (NF2) is an autosomal dominant disorder caused by mutations in the NF2 gene on chromosome 22q12 [1,2]. This disorder is primarily characterized by the growth of vestibular schwannomas (VS), noncancerous brain tumors that arise from Schwann cells along the vestibulocochlear nerve [3]. Depending on the location of VS tumor growth, NF2 patients may develop a range of symptoms, including hearing loss, tinnitus [4], loss of balance, and dizziness [5]. As these tumors progress, they can become life-threatening if they

XSL•FO
RenderX

impinge on critical brain structures, such as cranial nerves and the brainstem [6].

Treatment for these tumors typically involves stereotactic radiosurgery, resection surgery [7], or, more recently, Brigatnib, an oral drug showing promising outcomes in those with neurofibromatosis type 2–related schwannomatosis (NF2-SWN) [8]. The decision to pursue these treatments depends on factors such as tumor size, growth, location, and mass, underscoring the importance of accurately monitoring and visualizing tumor progression over time [9]. T1-weighted cranial magnetic resonance imaging (MRI) with intravenous gadolinium has proven effective in visualizing VS tumors in cerebral space [10]. This imaging technique also helps distinguish VS tumors from other anomalies, including cysts, meningiomas, aneurysms, and facial nerve schwannomas, supporting more informed clinical decision-making [11].

Tumor progression can be monitored using various metrics, with the change in tumor size being the most common. In current standard clinical practice, linear analysis involves measuring the largest tumor diameter in two dimensions, usually in an axial or coronal view. Although linear measurements are derived from a patient's MRI planar slices, the complex geometry of VS tumors can introduce uncertainty in these measurements [7,9]. Larger VS tumors develop characteristic features that may not be fully captured in linear measurements, such as the round body and tail (the "ice cream cone" geometry) [12].

Approximate volumetric analysis addresses some of the limitations of linear analysis by incorporating an extra dimension from 2D to 3D. By approximating the tumor as an ellipsoid in each MR scan slice, the volume of the tumor can be approximated by summing the total areas derived from each slice. While this measure has been shown to be more sensitive to tumor growth than 2D measurements, it has also been found to overestimate tumor growth. As VS tumors continue to develop into the cerebellopontine angle, their shape in each MR slice begins to deviate from that of an ellipsoid [13].

The 3D volumetric analysis addresses the limitations of 2D linear and approximate volumetric methods by calculating tumor volume through the summation of tumor areas from each MR slice. Unlike approximating the tumor as an ellipsoid in each slice, this technique involves directly segmenting the tumor borders, which improves the specificity of the tumor's boundaries [14]. Although 3D volumetric analysis yields accurate tumor size measurements, it requires manual segmentation, a time-consuming process that can take hours per tumor, necessitating the use of automated volumetric analysis tools [15]. Furthermore, for 3D volumetric analysis to have utility in clinical practice, the 3D visualization resulting from this analysis, which describes the digital 3D model of the segmented tumor in 3D space, must be easily interpretable for clinicians to understand where the tumor exists with reference to the 3D mapping of the patient's brain.

The 3D modeling technologies such as OsiriX Lite (version 7.0.3; Pixmeo SARL), ITK-Snap (version 2.4.0; ITK-Snap), and BrainLab iPlan (version 4.5; BrainLab) have been described in the literature as performing semiautomated 3D volumetric analysis of VS tumors [16]. While these technologies show promising results with low coefficients of variability, they fail to provide adequate visualization of the segmented tumors in 3D space, limiting their clinical utility. Specifically, these technologies do not have the capabilities to overlay 2D planes on the 3D segmented tumor to demonstrate where different sections of the tumor are located within the patient's brain scan. Additionally, these technologies do not have features that allow for the overlaying of segmented tumors between chronological scans to investigate areas of tumor growth. We envision that the incorporation of visualization features like these has the potential to broaden the clinical utility of 3D volumetric analysis and visualization software and reduce the need for advanced 3D reconstruction when viewing 3D segmented tumors.

To make 3D volumetric analysis adopted more widely in clinical settings, a thorough understanding of clinical needs related to the analysis methodology (volumetric vs linear) and the 3D visualization must exist. Therefore, the aim of this study was to gain a deeper understanding of clinicians' current methods for measuring tumor growth, their recommendations for key features in a potential 3D volumetric analysis tool (to be developed), and their concerns regarding how the tool would be clinically integrated. To achieve this, we conducted surveys and interviews with clinicians who treat NF2. We hypothesized that insights from these experienced clinicians would guide the development of a US Food and Drug Administration–regulated auto-segmentation and visualization software to improve the diagnosis and measurement of VS tumors.

Methods

Study Sample

This study involved surveys and interviews with clinicians specializing in the treatment and diagnosis of NF2-related schwannomatosis. To be eligible for inclusion, participants had to be clinicians directly involved in the care of NF2-SWN patients and willing to complete a survey or participate in an interview. NF2-SWN patients were not included, as the software is intended primarily as a clinical tool rather than a visualization tool for patients. Clinicians with direct experience in diagnosing and treating NF2-SWN across various specialties, including neurology, neuro-oncology, neurosurgery, neuroradiology, and pediatrics, whose names were found on the Children's Tumor Foundation website were recruited for this study. Upon identifying clinicians' names and contact information, emails were subsequently sent to these clinicians, outlining the proposed 3D segmentation software and inviting them to participate in a survey and an interview.

Ethical Considerations

Ethics approval was obtained from the Yale University Institutional Review Board (approval number: 2000034426) through an exempt review. Verbal informed consent was obtained from all interviewed and surveyed participants before the collection of data, and the institutional review board waived the need for written consent. The clinicians providing the data were deidentified but were not anonymous. No vulnerable populations were included in the data collection of this study. Additional consent was not necessary for any secondary analyses

XSL•FO
RenderX

of the collected data. No compensation was provided for participation in this study.

Survey Data Collection and Analysis

Surveys were conducted via Qualtrics and included 25 statements organized into eight categories as follows: (1) linear/volumetric analysis, (2) markers of VS tumor progression, (3) surgical planning for 2D MRI scans of VS tumors, (4) 3D volumetric analysis, (5) 3D visualization of VS tumors, (6) surgical planning using 3D models of VS tumors, (7) postintervention monitoring, and (8) general questions (Textbox 1). The survey questions were structured as a Likert scale questionnaire, where clinicians indicated their level of agreement with each statement on a scale of 1 to 5, ranging from "completely disagree" to "completely agree." Additionally, clinicians were asked demographic questions about their clinical specialties, years of experience, and the number of VS cases they had treated.

XSL•FO
RenderX

Textbox 1. A total of 25 total statements were addressed in both the interviews and in the questionnaire. In the surveys, all clinicians selected 1 out of 5 choices based on how much they agreed with each statement. In the interviews, clinicians had the opportunity to share open-ended responses to the statements.

Category 1: 2D linear/volumetric analysis

1. Linear analysis of vestibular schwannoma (VS) tumors reflects tumor size

2. Volumetric analysis of tumors reflects tumor size

3. Linear measurements are sensitive to small changes in tumor diameter

4. Two different clinicians can review the same magnetic resonance imaging (MRI) but calculate two different tumor sizes.

5. The orientations of bi- and uni-dimensional linear measurements can change from one MRI to the next for a single patient.

6. Patient position in MRI machines influences the orientations of MRI scans, affecting linear measurements.

7. It can be difficult to make a decision about tumor growth from a series of linear measurements due to the lack of clear progression thresholds.

Category 2: markers of VS tumor progression

8. Change in tumor size is the most important marker of VS tumor progression

9. Edema and mass effects on surrounding critical structures are good markers of VS tumor progression.

Category 3: surgical planning with 2D MRI scans of VS tumors

10. It is simple to determine the best approach to tumor resection before operation (location or angle of entry, how to best maximize extent of resection while avoiding functional areas) from 2D MRI scans.

11. The cranial nerves and blood vessels that need to be avoided during intervention can easily be determined through analysis of MRI scans.

Category 4: 3D volumetric analysis

12. 3D volumetric measurements would be more helpful in determining clearer thresholds of tumor growth than linear measurements.

13. 3D volumetric analysis would standardize measurements by reducing variation in measurements of tumor size taken by different clinicians.

14. 3D volumetric analysis would lead to more informed clinical decision-making.

Category 5: 3D visualization of VS tumors

15. 3D volumetric measurements must be combined with 3D visualization of the VS tumor to be clinically meaningful.

16. The ability to interact with 3D models of VS tumors in the cerebral space (rotate, zoom, adding or removing structures from view) would provide information helpful for surgical planning.

17. Growth in tumor boundaries observed by overlaying 3D models of VS tumors from MRI scans collected over time is easier to interpret than comparing the same MRI scans in 2D.

18. 3D visualization of VS tumors and their location relative to critical brain structures (cranial nerves, brainstem, etc.) would be helpful in patient care.

19. Knowing regions of rapid versus slow tumor growth on VS tumors would be helpful in treatment planning.

Category 6: surgical planning with 3D models of VS tumors

20. 3D visualization of VS tumors in the surrounding cerebral space (relative to critical brain structures) would aid in surgical planning.

21. 3D models of VS tumors would lead to more informed decisions of the type of intervention to perform (eg, different approaches of tumor resection and gamma knife)

Category 7: postintervention monitoring

22. Being able to distinguish between post-op changes and residual tumor progression would significantly improve clinical decision-making (ex. After gamma knife).

23. Automatic diagnosis of tumor progress, regression/response, or stability according to changes in size over time would aid treatment planning.

Category 8: general

24. Having the ability to forecast tumor growth will significantly improve clinical decision-making.

25. 3D visualizations of VS tumors would significantly improve patient education by allowing patients to see their tumors in 3D space.

In the end, individual survey statement responses were statistically analyzed and means and SD were reported. Higher means indicated that clinicians agreed more with the statement. Following statistical analysis, survey statements relating to tangible features or capabilities of the software were grouped together, and the top four features that clinicians favored based on mean agreement levels were determined.

Interview Data Collection and Analysis

Two members of the research team (AH and DW) with prior training in conducting semistructured interviews and focus

XSL•FO
RenderX

groups jointly facilitated each focus group discussion (FGD) session; the senior author (FB) conducted all the individual interviews. The topics discussed within these discussions mirrored the 25 survey statements reframed as open-ended questions (Textbox 1). The FGD sessions lasted approximately one hour, and the individual interviews lasted approximately 30 minutes. Audio recordings from the FGDs and individual interviews were then transcribed verbatim by Zoom's transcription feature and returned to the research team for proofreading and subsequent analysis.

The data analysis team (AH, DW, and FB) met regularly to develop the codebook and note potential emergent themes. Interview sessions and data analysis continued simultaneously and iteratively until the team determined that data saturation had been reached [17]. Beginning with an initial set of codes derived from the domains included in the Interview Guide, the coding team members independently coded each transcript to identify potential additional codes. The team then discussed the merits of each existing or new code and created definitions for each for inclusion in the final version of the codebook. The codes were inductively developed, grounded in the data, iteratively reviewed, and revised based on the team's discussion of additional independently coded transcripts. The total interobserver agreement of the reviewers was assessed over the course of the reviews. The overall total interobserver agreement was 92% across the study.

The codes assigned to each quote were then reviewed for a final time by the first author as part of entering the coded transcripts using Dedoose. Using thematic analysis, the team identified patterns across the entire data set to arrive at a final, parsimonious set of themes. The analysis also examined the data for any "negative" instances where the data did not fit the domains or themes or where no information on specific topics queried was obtained. In addition to the qualitative analysis, descriptive statistics were generated to describe the study sample.

Results

Study Sample

A total of 14 clinicians were surveyed from the neurosurgery (n=5), neurology (n=7), oncology (n=1), and general pediatrics (n=1) departments. Clinicians came from 4 different medical institutions and had a range of general clinical experience ranging from 3 to 41 years (mean 21.9, SD 12.1). The average number of VS cases that clinicians treated each year ranged from 0 to 100 (mean 24.6, SD 28.9). The clinicians who did not directly treat VS cases still possessed considerable knowledge about VS treatment through their experience working with NF2-SWN patients.

12 additional clinicians from the neurosurgery (4) neuroradiology (4), neurology (2), and neuro-oncology (2) departments were interviewed to provide open-ended responses to the survey statements. This number was determined based on the amount of interviews needed to reach saturation [17].

Survey Data Collection and Analysis

The results for all survey statements are reported in Table 1.

XSL•FO
RenderX

Table . From the Likert questionnaire, clinicians voted with how much they agreed with each statement by choosing (1) “Completely disagree”, (2) “Mostly disagree”, (3) “Neither agree nor disagree”, (4) “Mostly agree”, or (5) “Completely agree”. For the statistical analysis, these metrics of agreement were given numerical values 1 through 5, with 1 aligning with “Completely disagree” and 5 aligning with “Completely agree.”

Statements	Values, mean (SD)
Category 1: 2D linear/volumetric analysis	
Linear analysis of vestibular schwannoma (VS) tumors reflects tumor size.	3.50 (0.94)
Volumetric analysis of tumors reflects tumor size.	4.43 (0.51)
Linear measurements are sensitive to small changes in tumor diameter.	2.43 (1.22)
Two different clinicians can review the same MRI[a] but calculate two different tumor sizes.	4.64 (0.49)
The orientations of bi- and uni-dimensional linear measurements can change from one MRI to the next for a single patient.	4.21 (0.97)
Patient position in MRI machines influences the orientations of MRI scans, affecting linear measurements.	4.71 (0.47)
It can be difficult to make a decision about tumor growth from a series of linear measurements due to the lack of clear progression thresholds.	3.79 (1.25)
Category 2: markers of VS tumor progression	
Change in tumor size is the most important marker of VS tumor progression	3.50 (0.94)
Edema and mass effects on surrounding critical structures are good markers of VS tumor progression.	4.00 (0.78)
Category 3: surgical planning with 2D MRI scans of VS tumors	
It is simple to determine the best approach of tumor resection before operation (location or angle of entry, how to best maximize extent of resection while avoiding functional areas) from 2D MRI scans.	3.25 (1.16)
The cranial nerves and blood vessels that need to be avoided during intervention can easily be determined through analysis of MRI scans.	3.13 (0.99)
Category 4: 3D volumetric analysis	
3D volumetric measurements would be more helpful in determining clearer thresholds of tumor growth than linear measurements	4.50 (0.52)
3D volumetric analysis would standardize measurements by reducing variation in measurements of tumor size taken by different clinicians.	4.21 (0.58)
3D volumetric analysis would lead to more informed clinical decision making	4.14 (0.66)
Category 5: 3D visualization of VS tumors	
3D volumetric measurements must be combined with 3D visualization of the VS tumor to be clinically meaningful.	3.07 (0.92)
The ability to interact with 3D models of VS tumors in cerebral space (rotate, zoom, adding or removing structures from view) would provide information helpful for surgical planning.	3.86 (1.17)
Growth in tumor boundaries observed by overlaying 3D models of VS tumors from MRI scans collected over time is easier to interpret than comparing the same MRI scans in 2D	4.00 (0.68)
3D visualization of VS tumors and their location relative to critical brain structures (cranial nerves, brainstem, etc.) would be helpful in patient care.	4.36 (0.74)
Knowing regions of rapid versus slow tumor growth on VS tumors would be helpful in treatment planning.	3.71 (0.91)
Category 6: surgical planning with 3D models of VS tumors	
3D visualization of VS tumors in the surrounding cerebral space (relative to critical brain structures) would aid in surgical planning.	3.20 (1.30)

XSL•FO
RenderX

Statements	Values, mean (SD)
3D models of VS tumors would lead to more informed decisions of the type of intervention to perform (ex. Different approaches to tumor resection, gamma knife)	3.00 (1.41)
Category 7: postintervention monitoring	
Being able to distinguish between post-op changes and residual tumor progression would significantly improve clinical decision-making (eg, after gamma knife).	4.29 (0.73)
Automatic diagnosis of tumor progress, regression/response, or stability according to changes in size over time would aid treatment planning	3.86 (0.66)
Category 8: general	
Having the ability to forecast tumor growth will significantly improve clinical decision-making.	4.77 (0.44)
3D visualizations of VS tumors would significantly improve patient education by allowing patients to see their tumors in 3D space.	3.92 (0.76)

[a]MRI: magnetic resonance imaging.

Category 1: 2D Linear/Volumetric Analysis

Statistical analyses of statement responses revealed that clinicians felt that 3D volumetric analyses (mean 4.43, SD 0.51) offered a slightly greater reflection of tumor size than 2D linear analyses (mean 3.50, SD 0.94). Clinicians' responses additionally revealed that 2D analyses were generally not sensitive to small changes in tumor growth (mean 2.43, SD 1.22). Statements addressing variability in linear measurements highlighted a consensus around the potential for inconsistency due to factors like clinician interpretation (mean 4.64, SD 0.49), MRI orientation (mean 4.21, SD 0.97), and patient orientation (mean 4.71, SD 0.47), with high agreement that these factors impact linear measurements. Additionally, there was moderate agreement (mean 3.79, SD 1.25) surrounding the statement that linear measurements were difficult to interpret.

Category 2: Markers of VS Tumor Progression

Clinicians generally agreed that secondary effects, such as edema and impact on surrounding structures, serve as valuable markers for tracking VS tumor progression (mean 4.0, SD 0.78) and demonstrated a moderate agreement that tumor size is the most important marker above these secondary effects (mean 3.5, SD 0.94). This demonstrated the opinion that all metrics, including tumor size and secondary effects, may hold relatively equal importance to clinicians.

Category 3: Surgical Planning With 2D MRI Scans of VS Tumors

Responses indicated skepticism about the utility of 2D MRI scans in preoperative planning, with clinicians leaning towards neither disagreeing nor agreeing that these scans alone allow clear determination of surgical approach (mean 3.25, SD 1.16) and that critical structures can easily be determined through 2D MRI analyses (mean 3.13, SD 0.99).

Category 4: 3D Volumetric Analysis

Clinicians supported the use of 3D volumetric analysis over linear measurements, suggesting it would enhance consistency across clinicians (mean 4.21, SD 0.58), help establish clearer growth thresholds (mean 4.50, SD 0.52), and improve clinical decision-making (mean 4.14, SD 0.66).

Category 5: 3D Visualization of VS Tumors

In terms of features important for a 3D visualization software, most clinicians somewhat agreed that interactions with 3D models (rotating, zooming, or hiding structures) (mean 3.86, SD 1.17) and knowing areas of rapid versus slow growth (mean 3.71, SD 0.91) would be important for treatment planning. Clinicians were overall in agreement that overlaying chronological tumors to visualize growth (mean 4.00, SD 0.68) and viewing tumors relative to critical brain structures (mean 4.36, SD 0.74) were important for understanding tumor growth and improving patient care. However, clinicians generally neither agreed nor disagreed that visualizations were necessary for 3D analyses to be clinically useful (as opposed to volumes alone) (mean 3.07, SD 0.92).

Category 6: 3D Visualization of VS Tumors

In general, clinicians did not strongly agree or disagree regarding statements that claimed that 3D visualization would be a helpful aid for surgical planning (mean 3.20, SD 1.30) and would help make more informed clinical decisions (mean 3.00, SD 1.41).

Category 7: Postintervention Monitoring

Clinicians somewhat agreed that automatic diagnosis of tumor changes (regression/response or stability) according to changes in size would aid in treatment planning (mean 3.86. SD 0.66).

There was a greater agreement that the ability to distinguish between postoperative changes and residual tumor progression through chronological tumor overlaying would improve clinical decision-making (mean 4.29, SD 0.73)

Category 8: General

Clinicians strongly agreed that the ability to forecast tumor growth would be a significant improvement in clinical decision-making (mean 4.77, SD 0.44). In addition, clinicians somewhat agreed that 3D visualizations would be helpful for patient education (mean 3.92, SD 0.76).

XSL•FO
RenderX

The top four features that clinicians felt would be most beneficial in a 3D analysis and visualization software are listed in Textbox 2. These four features are listed in the order of priority based on the mean agreement levels in Table 1.

Textbox 2. List of 4 main features, ranked in order of clinician's priority, for 3D tumor visualization interface.

1. Forecast tumor growth based on previous 3D growth trends
2. Visualization of tumors with reference to other critical brain structures
3. Overlaying chronological 3D tumor models from the same patient to visualize growth
4. Interacting with 3D models, including rotating, zooming, and adding or removing tumors from view.

Interview Data Collection and Analysis

Interview findings were analyzed and sorted into four corresponding themes as follows: (1) shortcomings of the currently used linear analysis, (2) utility of 3D visualizations, (3) features of an interactive 3D modeling software, and (4) lack of a gold standard to assess the accuracy of 3D volumetric analysis.

Theme 1: Shortcomings of the Currently Used Linear Analysis

Clinicians observed that most medical professionals tracking the size or growth of VS tumors rely on linear analysis, primarily using uni- or bi-dimensional measurements. As such, linear analysis is considered the current "gold standard."

In general, discussions on the limitations of current linear analysis methods generally focused on the variability in measurements when conducted at different times or by different individuals. They highlighted several factors contributing to variability in these measurements. First, there is a lack of standardization in the linear analysis procedure, leading to variability in how clinicians perform measurements. Additionally, many clinicians will perform linear analysis by first determining the longest tumor diameter slice, a process that is subjective. Clinicians also noted that variations in patient orientation during MRI scans can result in inconsistent tumor presentations in axial and coronal views across different scans. Consequently, comparing linear measurements from one scan to the corresponding slice in another may be inaccurate. Finally, they emphasized that due to the subjectivity and potential inaccuracies in these measurements, tumor growth is more easily detected in scans taken over longer intervals, while frequent scans make it harder to discern changes.

Theme 2: Utility of 3D Visualizations

The clinicians interviewed agreed that visualizing VS tumors in relation to critical brain structures would significantly improve diagnosis and treatment planning. Currently, 2D MRI images must be mentally reconstructed by clinicians to understand how tumors interact with these critical structures in a 3D space. This mental reconstruction requires an abstract and advanced understanding of 3D cerebral anatomy, which typically comes with experience.

Clinicians further suggested that creating 3D models of tumors, along with color gradients to represent tumor growth over time, could enhance the visualization of tumor progression and help bridge the knowledge gap between experienced and less experienced clinicians. They also noted that additional markers—such as edema, necrotic core growth, and mass effects like compression on surrounding brain structures—are important in assessing tumor progression. Several clinicians emphasized that segmenting and visualizing these markers in 3D, alongside the tumor itself, could provide critical information for optimizing tumor management.

Theme 3: Features of an Interactive 3D Modeling Software

The clinicians agreed that adding interactive features to 3D modeling software would enable a more comprehensive understanding of each patient's case compared with simply viewing static 3D models on a screen. They suggested several interactive capabilities, including (1) adjusting the transparency of tumor models, (2) zooming in and out, and (3) rotating the models. These interactive functions were seen as providing significantly more information than traditional static 2D imaging. Clinicians also proposed incorporating animations of tumor growth over time, with newer tumors overlaid on older ones, to improve clinical understanding of disease progression. Additionally, they recommended incorporating a coordinate grid within the 3D tumor model to aid in tumor localization within the cerebral space, which could assist in guiding stereotactic interventions.

Theme 4: Lack of a Gold Standard to Assess the Accuracy of 3D Volumetric Analysis

In addition to providing suggestions regarding features that should be included in 3D visualization software, clinicians shared concerns regarding the incorporation of this type of software into clinical practice. These concerns primarily surrounded the novelty of this approach and the difficulty of validating measurements.

The clinicians interviewed noted that because linear measurements are currently the most common metric of tumor size, most experienced clinicians are adept at interpreting linear measurements to determine tumor management and treatment. By combining these linear measurements with the Response Evaluation Criteria in Solid Tumors (RECIST), clinicians can evaluate tumor growth. Thus, in the interviews, the clinicians expressed the concern that due to the novelty of using volume as a metric of tumor size, they may be unfamiliar with the clinical implication of these measurements and thus may be unable to take advantage of the additional information this 3D tool provides. Furthermore, clinicians expressed concerns regarding the lack of a gold standard by which to assess the accuracy of volumetric measurements. While no gold standard exists to determine the accuracy of linear measurements, the

XSL•FO
RenderX

novelty of volumetric analysis necessitates an additional metric of accuracy.

Discussion

Principal Findings

In this study, surveys and clinical interviews were conducted to evaluate the efficacy of current 2D tumor analysis protocols and to assess the clinical interest, recommendations, and concerns regarding the incorporation of 3D volumetric analysis for the treatment of NF2-related schwannomatosis. In summary, the consensus among clinicians was that a 3D volumetric analysis and visualization tool could greatly benefit the treatment of NF2-related schwannomatosis, but it must be implemented with caution. We believe that these clinician perspectives are extremely important as 3D volumetric analyses are increasingly being discussed within the literature, and for these technologies to have clinical relevance, the perspectives of clinicians must be at the forefront.

Clinicians were enthusiastic about the prospect of using 3D volumetric analysis in lieu of traditionally used linear analysis. This was largely due to several shortcomings of linear analysis that were highlighted through both survey results and interviews, with the two main concerns being varying patient orientations during MR imaging procedures causing variability in 2D image presentation and the subjectivity of defining the maximum tumor diameter in linear assessments. Despite efforts to collect linear measurements across scans as uniformly as possible by locating anatomical landmarks for individual patients, it can be challenging to ensure the measurement is collected in the same location. These limitations described by the interviewed clinicians align with those described by the Response Evaluation in Neurofibromatosis and Schwannomatosis consortium, something that led to their recommendation of using 3D volumetric analysis to assess tumor size initially [18].

From the interview and survey findings, there were several features of 3D analysis and visualization software that were believed to have the potential to benefit clinical practice. From interviews, the main consensus was that the ability to visualize the relationship between critical brain structures and tumors in 3D space would be the largest benefit of a 3D analysis and visualization tool as it would reduce the need for 3D mental reconstruction of where the tumor exists within 3D space. This feature was largely supported by the survey results as well, with "visualizing tumors with respect to critical brain structures" being the second highest feature that clinicians believed would be beneficial (Textbox 2).

Other features highlighted in both surveys and interviews demonstrated that the software should be interactable by incorporating zoom and rotation, allow for the adding/removing of the tumor from view, and have transparency adjustment features. These features are even more important when paired with the capability to view multiple chronological tumors at once, one of the other main features identified by both interviewed and surveyed clinicians as being important. Specifically, clinicians expressed enthusiasm for the prospect that these chronological tumor overlays could eventually lead to the capability of forecasting tumor growth. In Textbox 2 this "forecasting feature" ranked as the highest priority feature for clinicians demonstrating a belief that this could have a major positive impact on clinical treatment. Overall, the clinical opinions regarding important features demonstrate that current 3D modeling technologies such as OsiriX Lite, ITK-Snap, and BrainLab iPlan may not be heavily integrated into clinical practice due to shortcomings in their visualization capabilities, specifically their inability to overlay chronological tumors. Jester et al [19] provide a visualization of this chronological tumor overlay, demonstrating how it can give clinicians an interpretable vision of tumor growth over time.

Some statements regarding the benefits or potential use of 3D volumetric analysis or 3D visualizations were met with hesitancy, with average agreement levels nearing 3.0 (neither disagree nor agree). These statements claimed that 3D volumetric measurements must be combined with visualization to be clinically meaningful (mean 3.07, SD 0.92), that 3D models of VS tumors would lead to more informed decisions about the type of intervention needed (mean 3.00, SD 1.41) and that 3D visualization of VS tumors in the surrounding cerebral space would aid in surgical planning (mean 3.20, SD 1.30). This hesitancy demonstrates that while visualizations are important to clinicians, more precise volumes to compare between chronological scans may still be more important. Additionally, the hesitancy surrounding the use of 3D visualization to better define treatment approaches could also be attributed to a lack of understanding of how these visualizations may be used in this setting, something that could be better demonstrated to clinicians with a more detailed overview of the software's design.

Despite recommendations to move towards volumetric analyses, the interviewed clinicians also expressed concerns about how they would determine the clinical implications of volume measurements. These clinicians felt that they could comfortably determine the clinical implications of linear measurements by assessing the trends of uni- and bi-dimensional linear measurements of tumor size, but they lacked a similar knowledge of trends for volumetric measurements. As a future step to increase clinical confidence in the interpretation of tumor volumes, correlation studies can be conducted to determine the approximate relationship between linear and volumetric measurements. Additionally, future development of this diagnostic software can integrate linear analysis measurements as well, automating assessments of uni- and bi-dimensional diameters. Through such studies, clinicians can gain exposure to using segmentation software to perform volumetric analysis and interpret volume measurements while collecting familiar linear measurements. Pairing these correlation studies with currently used RECIST criteria can encourage the adoption of 3D volumetric analysis in the clinical space.

For full adoption and validation of 3D volumetric analysis in the clinical space, multi-center validation studies must take place with clinicians from multiple institutions performing 3D volumetric analysis on a diverse array of MRIs. DICE scores can be used as a metric to evaluate the differences in voxel selection within each scan between clinicians. To generate RECIST-compatible standards, current RECIST protocols

XSL•FO
RenderX

should be used as a basis to form corresponding 3D-specific protocols. This will require collaboration from multi-institutional research and clinical communities interested in the full adoption of 3D volumetric analysis.

Limitations

Clinical interviews incorporated in this study were conducted with a limited group of clinicians from just four institutions. Incorporating clinical feedback from a larger and broader group of clinicians experienced in the diagnosis and treatment of NF may more clearly elucidate the strengths and limitations of a novel segmentation and visualization software. Specifically, including perspectives from clinicians from institutions with varying levels of expertise, resources, and knowledge in treating NF and understanding how these varied departmental structures impact perspectives was not done in this study, but could aid in generating a tool more applicable to all clinicians, not just those from high-resource medical institutions. Additionally, though the sole user of this software is intended to be clinicians, other users, such as researchers, may use this software. However, within this study, interviews were limited to just clinicians, potentially limiting the information gathered.

Conclusions

Clinicians were overall in favor of the adoption of 3D volumetric analysis techniques for measuring VS tumors but expressed concerns regarding the novelty and inexperience surrounding these techniques. However, clinicians felt that the ability to visualize tumors with reference to critical structures, to overlay structures, to interact with 3D models, and to visualize areas of slow versus rapid growth in 3D would be valuable contributions to clinical practice. Overall, clinicians provided valuable insights for designing a 3D volumetric analysis tool for VS tumor growth. These findings may also apply to other central nervous system tumors, offering broader utility in tumor growth assessments.

Conflicts of Interest

None declared.

References

1. Asthagiri AR, Parry DM, Butman JA, et al. Neurofibromatosis type 2. Lancet 2009 Jun 6;373(9679):1974-1986. [doi: 10.1016/S0140-6736(09)60259-2] [Medline: 19476995]
2. Starnoni D, Giammattei L, Cossu G, et al. Surgical management for large vestibular schwannomas: a systematic review, meta-analysis, and consensus statement on behalf of the EANS skull base section. Acta Neurochir (Wien) 2020 Nov;162(11):2595-2617. [doi: 10.1007/s00701-020-04491-7] [Medline: 32728903]
3. Slattery WH. Neurofibromatosis type 2. Otolaryngol Clin North Am 2015 Jun;48(3):443-460. [doi: 10.1016/j.otc.2015.02.005] [Medline: 26043141]
4. Halliday D, Emmanouil B, Pretorius P, et al. Genetic severity score predicts clinical phenotype in NF2. J Med Genet 2017 Oct;54(10):657-664. [doi: 10.1136/jmedgenet-2017-104519] [Medline: 28848060]
5. Carlson ML, Link MJ. Vestibular schwannomas. N Engl J Med 2021 Apr 8;384(14):1335-1348. [doi: 10.1056/NEJMra2020394] [Medline: 33826821]
6. Anderson JL, Gutmann DH. Neurofibromatosis type 1. Handb Clin Neurol 2015;132(75-86):75-86. [doi: 10.1016/B978-0-444-62702-5.00004-4] [Medline: 26564071]
7. Nguyen D, de Kanztow L. Vestibular schwannomas: a review. Appl Radiol 2019;48(3):22-27. [doi: 10.37549/AR2579]
8. Plotkin SR, Yohay KH, Nghiemphu PL, et al. Brigatinib in NF2-related schwannomatosis with progressive tumors. N Engl J Med 2024 Jun 27;390(24):2284-2294. [doi: 10.1056/NEJMoa2400985] [Medline: 38904277]
9. Selleck AM, Rodriguez JD, Brown KD. Vestibular schwannoma measurements-is volumetric analysis clinically necessary? Otol Neurotol 2021 Jul 1;42(6):906-911. [doi: 10.1097/MAO.0000000000003055] [Medline: 33577243]
10. Jackler RK, Shapiro MS, Dillon WP, Pitts L, Lanser MJ. Gadolinium-DTPA enhanced magnetic resonance imaging in acoustic neuroma diagnosis and management. Otolaryngol Head Neck Surg 1990 Jun;102(6):670-677. [doi: 10.1177/019459989010200608] [Medline: 2115653]
11. Lin EP, Crane BT. The management and imaging of vestibular schwannomas. AJNR Am J Neuroradiol 2017 Nov;38(11):2034-2043. [doi: 10.3174/ajnr.A5213] [Medline: 28546250]
12. Lees KA, Tombers NM, Link MJ, et al. Natural history of sporadic vestibular schwannoma: a volumetric study of tumor growth. Otolaryngol Head Neck Surg 2018 Sep;159(3):535-542. [doi: 10.1177/0194599818770413] [Medline: 29685084]
13. Ho HH, Li YH, Lee JC, et al. Vestibular schwannomas: accuracy of tumor volume estimated by ice cream cone formula using thin-sliced MR images. PLoS One 2018;13(2):e0192411. [doi: 10.1371/journal.pone.0192411] [Medline: 29438424]
14. Morris KA, Parry A, Pretorius PM. Comparing the sensitivity of linear and volumetric MRI measurements to detect changes in the size of vestibular schwannomas in patients with neurofibromatosis type 2 on bevacizumab treatment. Br J Radiol 2016 Sep;89(1065):20160110. [doi: 10.1259/bjr.20160110] [Medline: 27452262]
15. MacKeith S, Das T, Graves M, et al. A comparison of semi-automated volumetric vs linear measurement of small vestibular schwannomas. Eur Arch Otorhinolaryngol 2018 Apr;275(4):867-874. [doi: 10.1007/s00405-018-4865-z] [Medline: 29335780]

XSL•FO
RenderX

16. Kollmann P, Mautner VF, Koeppen J, et al. MRI based volumetric measurements of vestibular schwannomas in patients with neurofibromatosis type 2: comparison of three different software tools. Sci Rep 2020 Jul 14;10(1):11541. [doi: 10.1038/s41598-020-68489-y] [Medline: 32665659]
17. Guest G, Bunce A, Johnson L. How many interviews are enough?: An experiment with data saturation and variability. Field Methods 2006;18(1):59-82. [doi: 10.1177/1525822X05279903]
18. Dombi E, Ardern-Holmes SL, Babovic-Vuksanovic D, et al. Recommendations for imaging tumor response in neurofibromatosis clinical trials. Neurology (ECronicon) 2013 Nov 19;81(21 Suppl 1):S33-S40. [doi: 10.1212/01.wnl.0000435744.57038.af] [Medline: 24249804]
19. Jester N, Singh M, Lorr S, Tommasini SM, Wiznia DH, Buono FD. The development of an artificial intelligence auto-segmentation tool for 3D volumetric analysis of vestibular schwannomas. Sci Rep 2025 Feb 18;15(1):5918. [doi: 10.1038/s41598-025-88589-x] [Medline: 39966622]

Abbreviations

MRI: magnetic resonance imaging
NF2: neurofibromatosis type 2
NF2-SWN: neurofibromatosis type 2–related schwannomatosis
RECIST: Response Evaluation Criteria in Solid Tumors
VS: vestibular schwannoma

Edited by A Kushniruk; submitted 27.01.25; peer-reviewed by JM Immanuvel Arockiasamy, RT Potla, Y Murali Mohan Babu; revised version received 04.06.25; accepted 09.06.25; published 30.07.25.

Please cite as:
Desroches ST, Huang A, Ghankot R, Tommasini SM, Wiznia DH, Buono FD
Clinician Perspectives of a Magnetic Resonance Imaging–Based 3D Volumetric Analysis Tool for Neurofibromatosis Type 2–Related Schwannomatosis: Qualitative Pilot Study
JMIR Hum Factors 2025;12:e71728
URL: https://humanfactors.jmir.org/2025/1/e71728
doi:10.2196/71728

© Shelby T Desroches, Alice Huang, Rithvik Ghankot, Steven M Tommasini, Daniel H Wiznia, Frank D Buono. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 30.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Leveraging Dual Usability Methods to Evaluate Clinical Decision Support Among Patients With Traumatic Brain Injury: Mixed Methods Study

Rubina F Rizvi[1,2,3], MD, PhD; Sameen Faisal[4], BS; Mark Sussman[5], MPAS; Patricia Mendlick[6], CMA; Sam Brown[6], BA; Elizabeth Lindemann[2,6], MHA; Sean Switzer[6], DO; Genevieve B Melton[1,2,3,6], MD, PhD; Christopher J Tignanelli[1,2], MD, MS

[1]Department of Surgery, University of Minnesota, 420 Delaware Street South East, Mayo Mail Code 195, Minneapolis, MN, United States

[2]Center For Learning Health System Sciences, University of Minnesota, Minneapolis, MN, United States

[3]Institute for Health Informatics, University of Minnesota, Minneapolis, MN, United States

[4]College of Biological Sciences, University of Minnesota, Minneapolis, MN, United States

[5]Stanson Premier Inc, Charlotte, NC, United States

[6]Fairview Health Services, Minneapolis, MN, United States

Corresponding Author:
Rubina F Rizvi, MD, PhD
Department of Surgery, University of Minnesota, 420 Delaware Street South East, Mayo Mail Code 195, Minneapolis, MN, United States

Abstract

Background: Patients with traumatic brain injury are at an increased risk of developing venous thromboembolism. Clinical decision support systems (CDSSs) may improve the use of venous thromboembolism prophylaxis protocols, yet suffer from poor compliance among end users due to a lack of user-centered design.

Objective: The objective of this work was to improve the content, design, and workflow integration of a traumatic brain injury–CDSS based on feedback from experts and end users.

Methods: The CDSS was evaluated leveraging a dual usability approach. A set of usability experts (n=3) and trauma providers (n=5) performed heuristic evaluations and usability testing by end users. Data was collected through a triangulation of methods and analyzed using qualitative (thematic) and quantitative (descriptive) analyses.

Results: Among the 145 total issues identified across both methods, 66 issues were found to be unique. Of the 66, a total of 17 issues were found by heuristic evaluations, 43 by usability testing by end users, and 6 were found across both methods. Thematic analysis was conducted on the 66 unique issues, which were further assigned to themes and subsequent subthemes. We identified 13 unique themes. The 3 most prevalent themes of 66 issues were lack of supporting evidence (n=17, 26%), operational barriers arising from the test environment (n=11, 17%), formatting inconsistencies, and lack of following standards (n=8, 12%). The system's usability scale survey score was 77.5 (SD 16, 95% CI 57.6-97.4), interpreted as an acceptable or good usability range. The mean response score for Single Ease Questions for all tasks was 5.9 (SD 0.53).

Conclusions: Combining expert and end user–driven usability evaluation methods identified a more comprehensive list of issues. This can facilitate the optimization of the traumatic brain injury–CDSS, resulting in improved usability and care management.

(JMIR Hum Factors 2025;12:e60268) doi:10.2196/60268

KEYWORDS

traumatic brain injury; venous thromboembolism; clinical decision support system (CDSS); heuristic evaluations; end user–based usability testing; decision support system

Introduction

Traumatic brain injury is one of the leading causes of mortality and disability worldwide. In the United States alone, over 69,000 traumatic brain injury–related deaths occurred in 2021 [1]. Patients with traumatic brain injury are at an increased risk of developing complications secondary to venous thromboembolism [2]. Fifty-four percent of patients with a traumatic brain injury will develop a venous thromboembolism in the absence of appropriate anticoagulation intervention [3]. Despite the significance of timely venous thromboembolism prophylaxis administration, there has been a reluctance to

XSL•FO
RenderX

implement anticoagulation in such patients, thus resulting in delays in timely intervention [4-7].

Optimal use of venous thromboembolism prophylaxis protocols manages the risk of thromboembolism with the competing risk of progression of intracranial bleeding in patients following trauma [7]. However, there is variability in venous thromboembolism prophylaxis protocols across trauma centers [2,7,8]. In an attempt to optimize thromboembolism prophylaxis for the patient, institutional implementation of anticoagulation protocols has been shown to improve clinical outcomes [2,7,8]. To improve prophylactic anticoagulation of patients with traumatic brain injury without adversely affecting outcomes, integrating a clinical decision support system (CDSS) has emerged as an indispensable solution to supplement providers with evidence-based venous thromboembolism prevention guidelines [9]. Existing literature cites improved patient outcomes when evidence-based practices are followed, along with improved compliance with these practices through the use of CDSSs [10-12].

CDSSs, whether stand-alone or integrated with an electronic health records system, are considered a monumental step forward in health care delivery as they serve as reliable and readily available aids to decision makers. CDSSs serve as a platform for clinicians by comparing individual patient data and integrating evidence-based recommendations into their daily workflow [13,14]. When adhered to, these tools serve as an effective means of changing clinician behavior [15-17]. The use of CDSSs varies by purpose and context. However, the resultant benefits of CDSSs are found to consistently positively impact various elements of health care delivery for both patients and clinicians. For providers, it facilitates overall disease management by aiding complex decision-making that is mandated for accurate diagnosis, effective treatment, and efficient monitoring [18]. For patients, the use of CDSSs leads to fewer errors, improved patient safety, improved care quality, and enhanced health outcomes [19].

Despite these benefits, there are some risks and resistance to using CDSS [19]. This includes but is not limited to cost-related challenges, alert fatigue, physicians' autonomy and trust, liability issues, workflow disruptions, and usability challenges. The negative consequences of a tool to be developed and implemented could be mitigated if iterative evaluations are carried out earlier rather than in the later phase of the system development life cycle [20]. Often, due to the additional resources required (both time and money), iterative evaluations are not prioritized before deployment in clinical practice. However, predeployment usability evaluation could potentially save time and money by preventing the need for critical changes after system implementation [21]. This study highlights the criticality of early testing and engaging end users during the system development life cycle. Identifying and addressing users' concerns around tool optimization and workflow integration should be considered an integral component when building, implementing, and maintaining a CDSS.

Lack of user-centered design, interoperability, and implementation barriers hinder the widespread adoption of a CDSS [22-24]. There is a need for robust usability studies to address existing barriers and optimize the functionality of the CDSS according to users' needs. The objective of this study was to evaluate and optimize the usability of a traumatic brain injury CDSS integrated into a US Midwest American College of Surgeons–verified trauma center. Usability issues in a traumatic brain injury-venous thromboembolism-CDSS were identified by leveraging a dual usability evaluation approach. We hypothesized that combining expert and end user–driven methods of usability evaluation would help identify a wide range of unique issues specific to each method, along with common findings shared between the 2 methods.

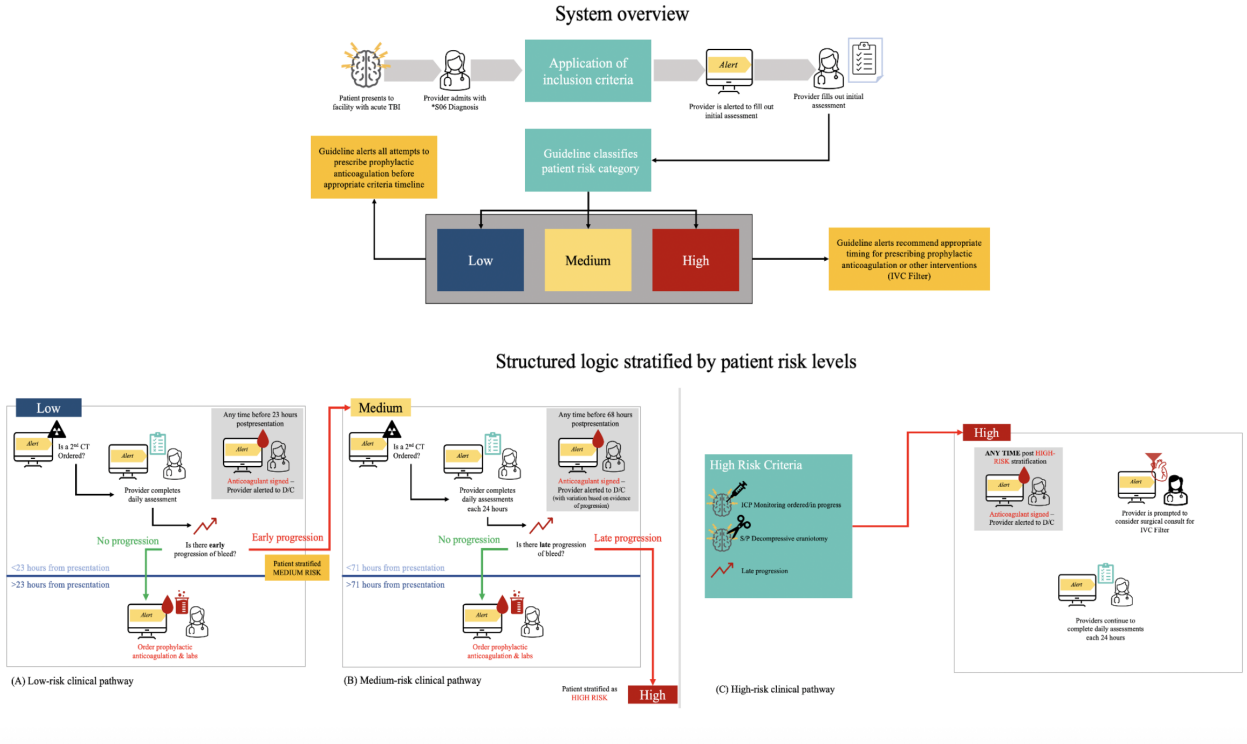
Methods

Overview of Traumatic Brain Injury–Venous Thromboembolism Prevention Guidelines

The traumatic brain injury–CDSS delivers a patient-centered outcomes research guideline for venous thromboembolism prevention in patients with traumatic brain injury [2,25], which was centered around the modified Berne-Norwood criteria [8]. Patients are categorized as being at low, moderate, or high risk for spontaneous progression of hemorrhage, with prophylactic treatment regimens tailored for each group, respectively. The traumatic brain injury–CDSS is a modification of an original, natively developed COVID-19 CDSS logic model, which is currently translated to interoperable standards [16] (Figure 1). We chose this model for two major reasons: (1) we have previous experience implementing the traumatic brain injury venous thromboembolism prevention guideline with success [2], and (2) many of the data elements and artifacts (~80%) overlap with minimal additional mappings and technology build. The final CPG-FHIR (clinical practice guidelines on Fast Healthcare Interoperability Resources) COVID-19 CDSS included 108 data elements. The traumatic brain injury clinical practice guideline added 21 data elements specifically related to traumatic brain injury (eg, intracranial pressure monitoring, neurosurgical consultation, craniotomy, head computed tomography imaging, subdural hematoma, contusion, and Glasgow Coma Scale).

XSL•FO
RenderX

Figure 1. Clinical decision support system logic model for venous thromboembolism prevention among patients with traumatic brain injury. *S06: 2024 *ICD-10-CM* Diagnosis Code S06. CT: computed tomography; D/C: discontinue anticoagulation order; *ICD-10-CM*: *International Statistical Classification of Diseases, Tenth Revision, Clinical Modification*; ICP: intracranial pressure; IVC: inferior vena cava; S/P: status post; TBI: traumatic brain injury.



Given the logic model, patients will enter the system by admission under predetermined diagnoses, generally falling under the traumatic brain injury category. Our target population is adult patients admitted with an acute traumatic brain injury, as defined by *ICD-10-CM* (*International Statistical Classification of Diseases, Tenth Revision, Clinical Modification*): S06.1-S06.9 or S06.A. Patients who died within 24 hours of hospital admission and patients documented as "comfort cares" during the first 72 hours of hospitalization will be excluded. Additionally, patients with a pre-existing venous thromboembolism or inferior vena cava filter at the time of admission and patients with a mechanical heart valve or ventricular assist device will be excluded from the final analysis.

The CDSS is provided in the subsequent 72 hours (dependent on risk stratification) and administered in the form of seven alerts (Textbox 1). Textbox 1 provides descriptions of the seven alerts deployed at various points in time, stratified by level of risk.

Textbox 1. Description of seven alerts in traumatic brain injury–clinical decision support system (CDSS).

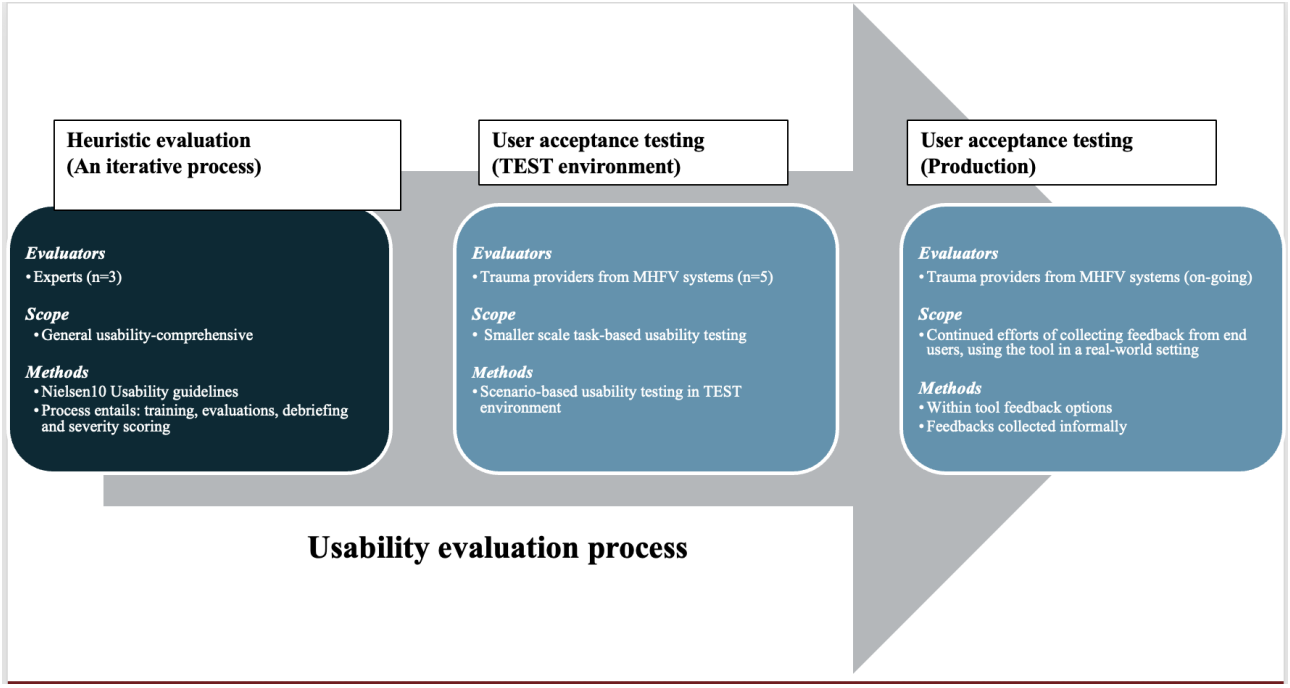
- **Initial assessment:** provides initial risk stratification.
- **Follow-up computed tomography (CT) scan alert:** Guide proper timing of a follow-up CT head scan based on stratification
- **Daily assessments:** Track patient course and modify stratification in the event of critical conditions; that is, progression of bleed, placement of intracranial pressure (ICP) monitor, decompressive craniotomy, etc.
- **Anticoagulation alert:** Guides the health care provider to prescribe the correct dose of anticoagulation medication at the appropriate time, and not too soon.
- **Discontinue anticoagulation order alert:** Any attempt to prescribe anticoagulation before the appropriate clinical decision support system calculated time triggers an alert to discontinue the medication.
- **Laboratory order alert:** Prompts order of appropriate laboratories to monitor side effects of the anticoagulation medications, that is, Anti-Xa, platelets, etc.
- **Laboratory level alerts:** Additional prompts to monitor laboratory levels for appropriateness.

Study Settings

Our foundational usability evaluation-related research work was conducted by a team of researchers from the University of Minnesota (UMN) and associated health systems, such as the M (Minnesota) Health Fairview and affiliated hospitals, which consisted of 9 state-verified trauma centers, including the flagship UMN as an American College of Surgeons–verified level 2 trauma center. The usability evaluation methodology is reported in Figure 2.

XSL•FO
RenderX

Figure 2. A process diagram of the methodology. MHFV: M (Minnesota) Health Fairview; TEST: test environment.



Data Collection

Expert-Based Usability Evaluations

Heuristic evaluations leveraging Nielsen 10 Usability Heuristics for User Interface Design were used iteratively during the whole development cycle, where issues identified were addressed sequentially [26]. The initial round of heuristic evaluations was conducted on screenshots of the user interface that researchers independently evaluated.

Three experts with varied backgrounds, that is, a physician informatics or usability expert (RFR); a hospitalist with content and CDSS-related expertise (SS); and an undergraduate in neuroscience (SF) conducted the evaluation independently. We then had a debriefing session to walk through findings from individual users, removing any duplicates. Once we had a final, consolidated, clean list of unique issues, each evaluator did the severity ranking of each issue. A 0 to 4 rating scale was used to rate the severity of usability problems (0=I do not agree that this is a usability problem at all; 1=cosmetic problem only: need not be fixed unless extra time is available on project; 2=minor usability problem: fixing this should be given low priority; 3=major usability problem: important to fix, so should be given high priority; 4=usability catastrophe: imperative to fix this before product can be released) [27]. While assigning the severity score to usability problems, 3 factors were considered: the frequency with which the problem occurs, the impact of the problem, and the persistence of the problem. The team reconvened in a severity scoring session where a scoring consensus was reached for each problem after mutual deliberations.

End Users' Based Usability Testing

End users' based usability testing was conducted with actual end users, that is, trauma providers from M Health Fairview and affiliated hospital systems. Purposive sampling was used for recruitment purposes by sharing information about the traumatic brain injury–CDSS usability study through emails, discussions during meetings, word of mouth, etc. Scenario–based usability testing was done with individual end users along with our usability team comprised of up to 5 team members, including a moderator (n=1), facilitators–information technology (IT) experts both from the MHVF and the vendors who helped with building (n=2 - 3); and a note taker (n=1). Each session was conducted on Zoom (Zoom Inc.) and lasted up to 90 minutes.

In the scenario-based guided usability testing, we asked users to perform certain tasks on 3 unique test patients corresponding to 3 risk groups: low, medium, and high. Two to four scenarios were associated with each test patient for a total of 9 scenarios, which required each user to perform a few predefined, sequential tasks. Participants were asked to think out loud as they interacted with the traumatic brain injury–CDSS in the Epic testing environment (Epic Inc). Once each scenario and relevant task were completed, participants were asked a few questions to elaborate on their experience and complete a single ease survey questionnaire (REDCap [Research Electronic Data Capture] Consortium; Vanderbilt University). At the end of the session, participants completed a System Usability Scale survey (SUS) [28] and Single Ease Question (SEQ) [29] and were asked to share their overall experience with the tool.

Building Test Patients and Testing in Epic Test Environment

Our team of experts was comprised of builders or IT specialists (SB, PM, and MS) and clinicians (SS, RFR, and CJT), who helped with building the 3 master test patients stratified by risk level (Textbox 2). Scenarios were generated by a team of clinicians having diverse sets of experiences (ie, CJT, a trauma surgeon, and project principal investigator, SS, a hospitalist, and RFR, a physician informaticist with user acceptance testing experiences) and builders or IT specialists (SB, PM, and MS)

XSL•FO
RenderX

after in-depth discussions. The goal was to have scenarios that are commonly encountered and require trauma providers to perform certain high-priority tasks. Copies of the master patients were created to be used by each study participant..

Textbox 2. Patient risk stratification modified the Berne-Norwood criteria.

Test patient # 1: low risk

- Subdural hematoma 2 mm.
- No moderate or high-risk criteria.

Recommendation: initiate pharmacologic prophylaxis if repeat head computed tomography (CT) was stable at 24 hours.

Test patient # 2: moderate risk

- Subdural or epidural hematoma >8 mm.
- Contusion or intraventricular hemorrhage >2 cm.
- Multiple contusions per lobe.
- Subarachnoid hemorrhage with abnormal CT.
- Evidence of progression at 24 hours.

Recommendation: initiate pharmacologic prophylaxis if the head CT was stable at 72 hours.

Test patient # 3: high risk

- Placement of an intracranial pressure monitor, craniotomy, or both.
- Evidence of progression at 72 hours.

Recommendation: consider placement of an inferior vena cava filter.

The purpose of the testing was to receive user feedback effectively and efficiently, ensuring that the testing context matches as much as possible with the user interactions in the real world. Leveraging the Epic test environment to create a simulated environment came with a few constraints described under limitations.

Data Analysis

Overview

The respective data was analyzed both qualitatively and quantitatively as described.

Qualitative-Thematic Analysis

We compiled all the issues identified by each method into 1 consolidated list. This included data collected through heuristic evaluations and usability testing by end users. First, a subset of the issue was coded independently by 2 individuals (RFR and SF) at a more granular level, and a codebook was generated after mutual agreement between both coders and applied to the full list of issues. Each analyst coded the rest of the transcripts for various usability problems, using the codebook generated earlier. We cataloged the lowest-level codes into higher-level subthemes and themes to gain a more holistic understanding of traumatic brain injury–CDSS.

Quantitative Analysis

The primary focus was on data collected from the SEQ survey and SUS. SEQ is an effective and efficient tool comprised of a single question asked immediately after task completion. SUS is the most widely used and accepted standardized tool to measure subjective usability. We performed descriptive analysis, reporting means and SDs.

Ethical Considerations

This component of the project is not a clinical trial. This study protocol (STUDY00017107) was submitted to the UMN Institutional Review Board and given the determination of “exempt” as secondary research for which consent is not required. The mixed methods investigation was determined as a “not human research” as a quality improvement activity. Signed and written informed consent was obtained to screen and audio record the session, with the speech-to-text transcription feature turned on. Each recording was saved on a secured and encrypted UMN device. Each participant later received a US $50 gift card for their participation.

Results

End Users’ Characteristics

Five trauma providers participated in the individual usability testing session. The detailed characteristics are described in Table 1.

XSL•FO
RenderX

Table . Demographics of participants.

Characteristics	Trauma providers (n=5), n (%)
Sex	
Male	2 (40)
Female	3 (60)
Role	
MD[a]	4 (80)
APP[b]	1 (20)
Specialty	
Trauma surgery	3 (60)
Neurosurgery	1 (20)
Neurocritical care	1 (20)
Comfort with technology	
Expert	4 (80)
Intermediate	1 (20)

[a]MD: medical doctor.

[b]APP: advanced practice provider.

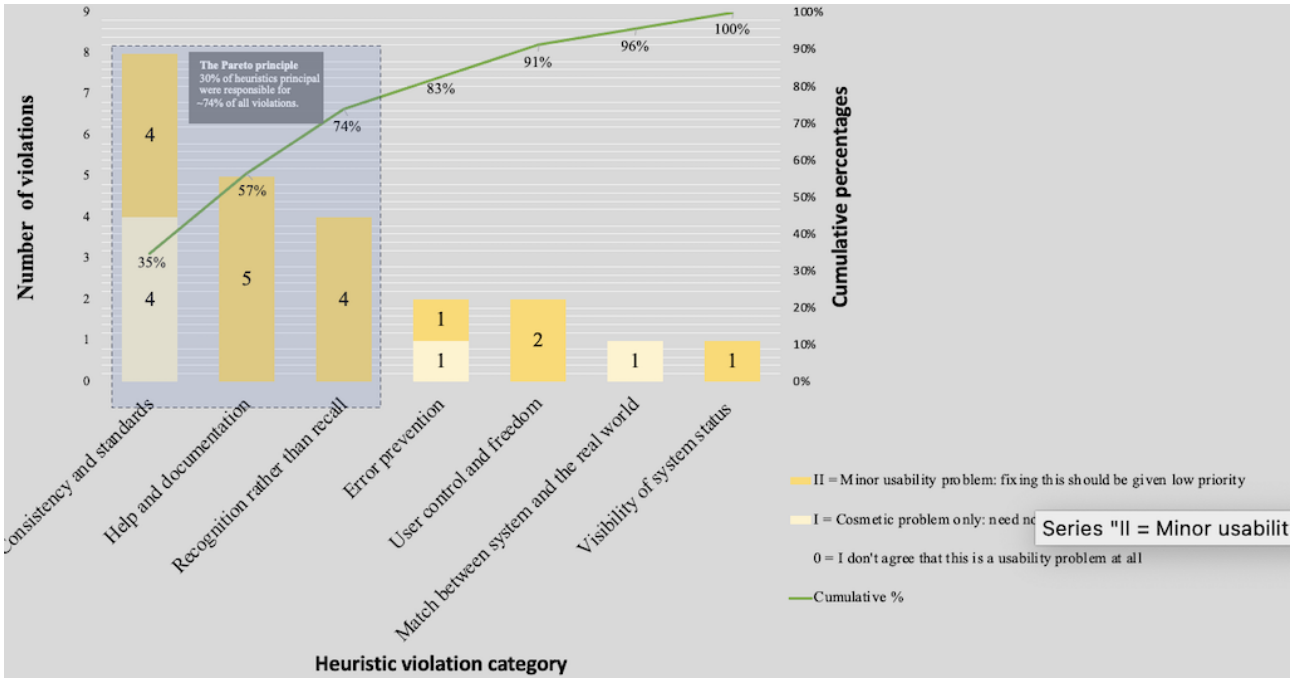
Usability-Related Findings From the 2 Methods

We identified a set of both common and unique issues using 2 methods: heuristic evaluation with experts (domain experts, ie, physicians, usability experts, or both) and usability testing with end users (the actual users).

Findings From Heuristics Evaluation

Thirty-seven heuristic violations (23 unique and 14 duplicates) were identified by 3 experts. The results showed that 30% of possible heuristics were responsible for ~74% of all violations, the Pareto principle (Figure 3). Three heuristics that represented the most violations were (1) consistency and standards (n=8, 35%), (2) help and documentation (n=5, 22%), and (3) recognition rather than recall (n=4, 17%). Of 23 unique, we found cosmetic (n=6) and minor (n=17) issues only.

Figure 3. The Pareto chart of heuristics principles violation and the severity ranking.

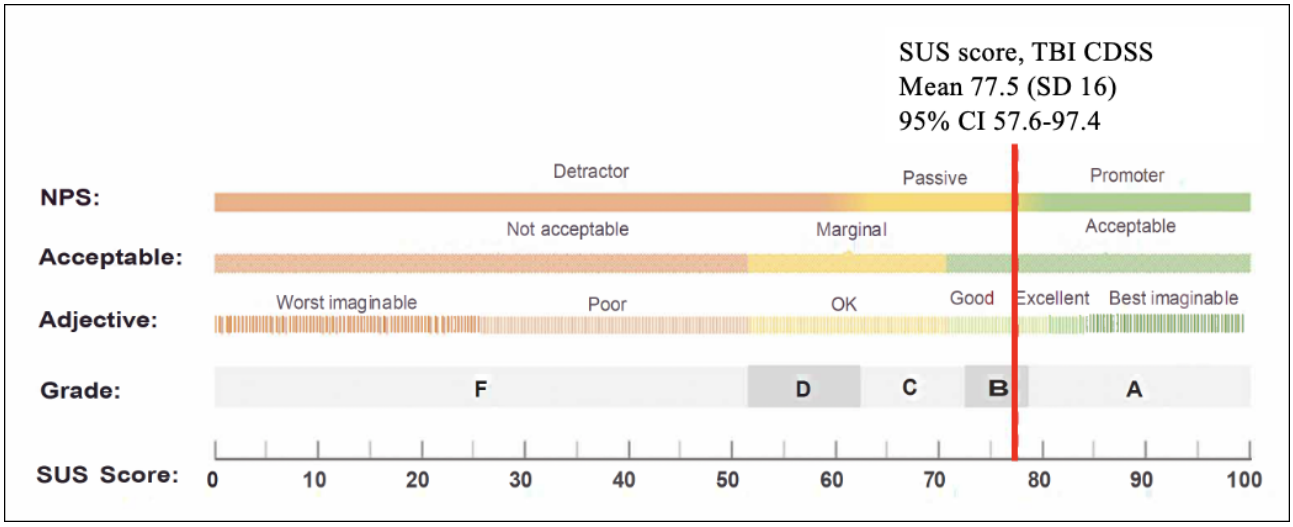


XSL•FO
RenderX

Findings From Usability Testing

We identified 108 issues (49 unique and 59 duplicates). Of 49 issues, we found cosmetic (n=6), minor (n=20), major (n=10), and catastrophic (n=2) issues. The remaining 11 were considered operational and not ranked. The mean responses to SEQ for all tasks were 5.9 (SD 0.53), ranked from 1 (very difficult) to 7 (very easy). The mean SUS score was 77.5 (SD 16) interpreted as an acceptable or good usability range (Figure 4). The 95% CI for the mean SUS score is 77.5 (95% CI 57.6 - 97.4).

Figure 4. The mean Systems Usability Scale score is interpreted as being within acceptable or good usability range (adapted from Lewis and Sauro [30], with permission from Jim Lewis). CDSS: clinical decision support system; NPS: net promoter score; SUS: Systems Usability Scale survey



Between the 2 methods, we identified a total of 145 issues, 66 of which were unique: 17 from heuristic evaluations, 43 from end users, and 6 common issues shared between both methods. Thematic analysis was performed on all unique issues, eliminating those arising from the constraints of thfsuye Epic test environment (66 - 11=55).

Ranking of Usability Issues

We ranked 55 unique usability issues on a severity scale from 0 - 4, where 20% (n=11) were marked as cosmetic; 58% (n=32) were identified as minor; 18% (n=10) as major, and 4% (n=2) as catastrophic issues. We identified both cosmetic and minor issues through heuristic evaluations and usability testing with end users. However, severe and catastrophic issues were solely identified during the usability sessions with end users.

Thematic Analysis

Thematic analysis was conducted on 66 unique issues identified through heuristic evaluations and usability testing. Eleven issues arising from Epic test environment constraints were eliminated. We identified 13 unique themes, including lack of supporting evidence, operational barriers arising from the test environment, formatting inconsistencies and lack of following standards, suboptimal data retrieval and display, lack of user control, lacking language clarity, order set-related suboptimal usability, language inconsistencies, best practice alert lacking patient-specific details, clinicians missing critical information or action, unfamiliarity with the full clinical context, infinite ordering loop leading to error, and having redundant options. The 3 most identified themes were lack of supporting evidence (n=17, 26%), operational barriers arising from the test environment (n=11, 17%), formatting inconsistencies, and lack of following standards (n=8, 12 %; Table 2).

XSL•FO
RenderX

Table . Thematic analysis distribution.

Theme	Unique issues (n=66)
Lack of supporting evidence	17
Operational barriers arising from the test environment	11
Formatting inconsistencies and lack of following standards	8
Suboptimal data retrieval and display	6
Lack of user control	5
Lacking language clarity	4
Order-set related suboptimal usability	4
Language inconsistencies	3
Best Practice Advisory lacking patient-specific details	2
Clinicians missing critical information, action, or both	2
Unfamiliarity with the full clinical context	2
Infinite ordering loop leading to error	1
Redundant options	1

Lack of supporting evidence refers to the end users' request to integrate and access evidence-based citations throughout the CDSS. Of the 17 usability issues about evidence-based medicine, 14 were identified by end users and 3 were identified by experts. Examples include users disagreeing with threshold values and requesting evidence-based medicine citations, and users requesting the ability of the CDSS tool to be explained.

Formatting consistency and a lack of following standards refer to the system's ability to present information clearly and concisely, with a consistent interface to ensure easy readability for the user. Of the 8 usability issues about formatting consistency and standards, 2 were identified by end users, and 6 were identified by experts. Examples include inconsistent and difficult-to-read text usage, style, color, and sizes.

Suboptimal data retrieval and display refer to the system's ability to present information in a manner that reduces user cognitive burden by increasing smooth and efficient navigation by auto-populating data (such as laboratory results). Of the 6 usability issues about easy retrieval and optimal display, all were identified by end users. Examples include users requesting that laboratory or imaging values and administration time be incorporated into the CDSS.

Lack of user control refers to the flexible design of the system to ensure smooth interaction with the user. Of the 5 usability issues about user control, 2 were identified by end users and 3 were identified by experts. Examples include users requesting the ability to navigate away from the tool and forgiveness when using the CDSS.

Lacking language clarity refers to specifying verbiage to omit any confusion the user may have. Of the 4 usability issues about language clarity, 3 were identified by end users and 1 was identified by experts. Examples include specifying verbiage used throughout the tool and using specific phrases (ie, the use of "active bleed" vs "new bleed" may imply 2 different ideas).

Order set-related suboptimal usability refers to any suboptimal interactions the user may encounter when interacting with the order set. Of the 4 usability issues related to order set-related suboptimal usability, all were identified by end users. Examples include users requesting the tool display prompts regarding its functionality on a case-by-case basis and easy access to the admission order set.

Language inconsistencies refer to the consistent use of verbiage throughout the system. Of the 3 usability issues about language consistency, they were all identified by experts. Examples include discrepancies between verbiage used in the storyboard or title and the consistent use of terms implying the same definition (ie, either "snooze" or "lockout time" should be used, not both).

Best Practice Advisory (BPA) lacking patient-specific details refers to information that the system does not include, which users would prefer to be included. Of the 2 usability issues about BPA lacking patient-specific details, both were identified by end users. Examples include adding patient risk stratification to the BPA (ie, changing the tool to be more diagnostic) and including crucial patient-specific laboratory values.

Clinicians missing critical information refers to ensuring a system design that ensures easy readability for the clinician. Of the 2 usability issues about eliminating the chance of clinicians missing critical information, 1 was identified by end users and 1 was identified by experts. Examples include bolding information and making information more accessible to improve readability.

Unfamiliarity with the context refers to additional guidance presented to the clinician on the system. Of the 2 usability issues about unfamiliarity with context, both were identified by end users. Examples include ensuring proper education and introduction of the tool to clinicians to ensure it does not interrupt their workflow.

A redundant option refers to options already included as part of the system. Of the 1 usability issue about redundant options, it was identified by end users. Examples include omitting the "already ordered complete" button.

XSL•FO
RenderX

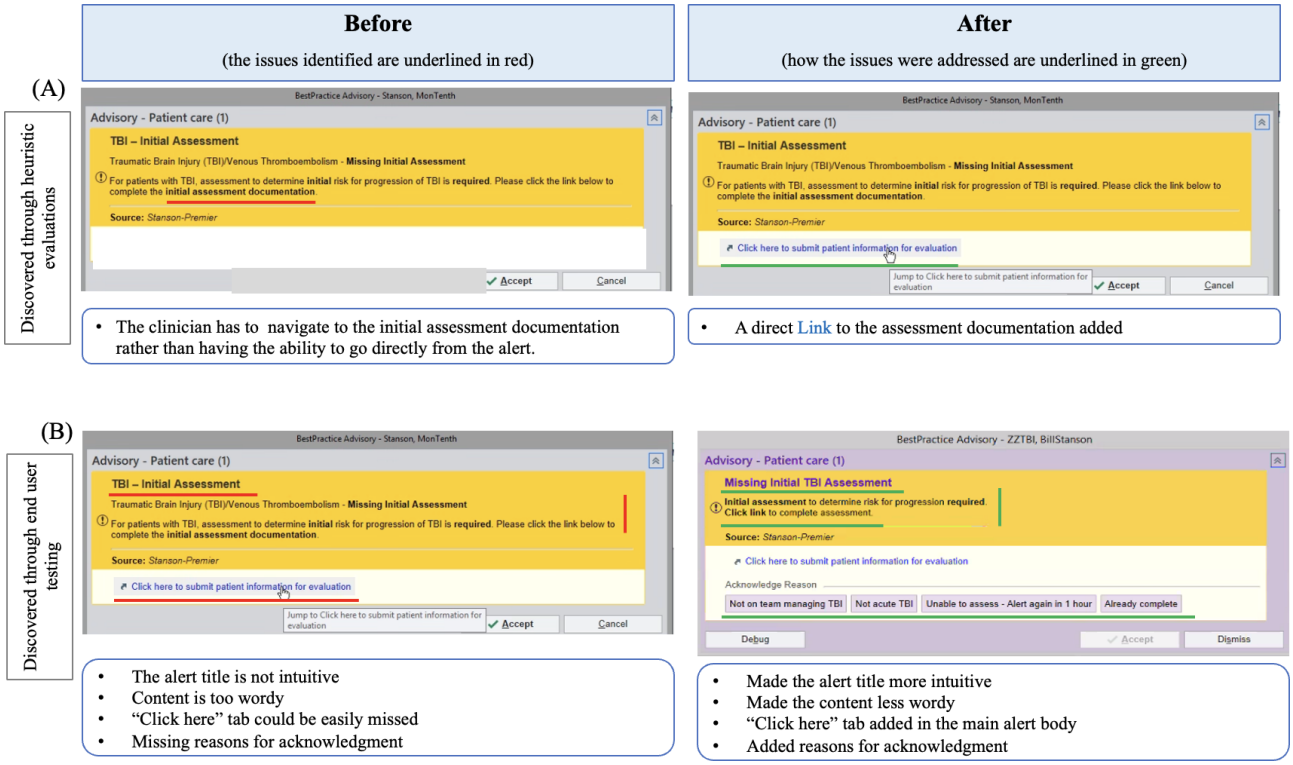
An infinite ordering loop leading to error refers to an infinite loop that occurs when using the system, during the beginning of its launch. It has since been resolved as of May 16, 2023. The 1 usability issue about the infinite ordering loop was identified by end users. Examples include users noting their warranted frustration with the ordering process.

Clinicians most often identified issues relating to the readily apparent integration of evidence-based medicine into the CDSS, recognizing a need for the use of citations and principles. Clinicians frequently noted how the CDSS might disrupt their workflow. For practical feasibility, request the integration of critical laboratory or imaging values and consistency in formatting throughout the tool to reduce user cognitive burden and ensure an efficient workflow. A compilation of unique usability issues categorized by theme is included in Multimedia Appendix 1.

Examples of Usability Modifications

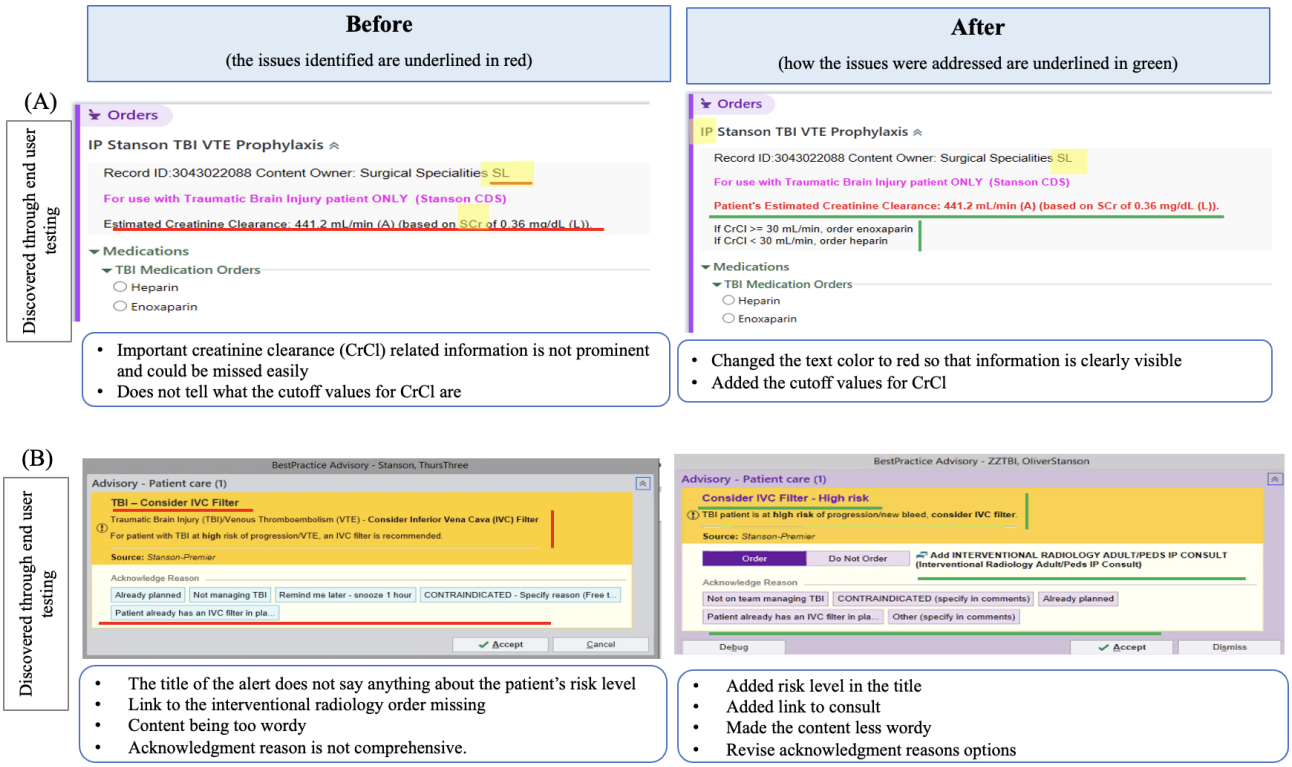
Figures 5 and 6 present a few examples of modifications made based on insights from heuristic evaluations and usability testing by end users. The full list of identified usability issues is provided as a supplementary table.

Figure 5. Examples of modifications. TBI: traumatic brain injury. Copyright 2024 Epic Systems Corporation. Used with permission.



XSL•FO
RenderX

Figure 6. Examples of modifications. Images included with approval from Epic. CrCl: creatinine clearance; IP: inpatient; IVC: inferior vena cava; SL: service line; SCr: serum creatinine; T/PEDS: ADULT/PEDS; TBI: traumatic brain injury; VTE: venous thromboembolism. Copyright 2024 Epic Systems Corporation. Copyright 2024 Epic Systems Corporation. Used with permission.



Discussion

Summary and Interpretation of Findings

Using a dual usability evaluation approach, namely expert-driven heuristic evaluation and end user–driven usability testing, we identified a variety of both similar and unique issues across the traumatic brain injury venous thromboembolism CDSS. Of the 66 unique issues identified, the percentage attributed to each method was as follows: usability testing (43 issues, 65%), heuristic evaluations (17 issues, 26%), and both methods together (6 issues, 9%). Although the mean severity scores of the identified problems did not differ significantly between the 2 methods, issues that were ranked as severe and catastrophic were exclusively found through usability testing by the end users. The most common issues identified across both methods were related to a lack of supporting evidence, operational barriers, and formatting consistency or lack of following standards. Heuristic evaluations primarily identified issues related to formatting consistency or lack of following standards, user control, suboptimal language clarity, and consistency. In contrast, usability testing revealed a broader range of issues, where lack of supporting evidence, operational barriers, and suboptimal data retrieval and display were found to be the most common.

Using a combination of these 2 methods can identify a wider range of usability issues. This aligns with prior evidence suggesting that combining 2 methods of usability evaluation can facilitate the development of an optimal CDSS [24]. We found that the majority of issues identified by usability testing were only discovered by end users, as compared to the experts. Although the experts identified fewer issues, they pinpointed more granular issues related to the layout, formatting, and consistency. In contrast, end users were more concerned with larger, more consequential issues, such as the lack of supporting evidence to increase their trust and confidence and fragmented workflow (primarily sending form test environment-related constraints). The findings from this study validate many of the claims made in previous comparisons of heuristic evaluations and usability testing [31,32] such as heuristic evaluations being more cost-effective in terms of time and resources, and the ability of each method to identify distinct sets of usability issues with very small convergence rates (9% in our study). Heuristic evaluation findings were limited to identifying issues, whereas end users engaged in usability testing provided both strengths and weaknesses. For example, end users appreciated the clarity of the premier questionnaire (which covered the critical issue for risk profiling), as well as the straightforward alert verbiage. They also appreciated that the creatinine clearance value was apparent in the laboratory order results, eliminating the need to navigate to find it.

Limitations

There are some limitations to this study. First, the usability evaluation was conducted at a single institution (M Health Fairview) using a single electronic health system (Epic). We had 5 clinicians participate in the end user testing. Several usability evaluation resources suggest that testing with 5 individuals is often considered appropriate for identifying usability issues, especially in qualitative analysis [26]. However, a larger sample size would have strengthened the generalizability of the findings for quantitative analysis. Additionally, we did not examine whether participants' amount of experience in their

XSL•FO
RenderX

respective fields or their overall familiarity with digital technology influenced their insights.

Usability testing being conducted in the Epic test environment introduced some constraints that users had to adjust to, such as manually shortened refresh time frames to decrease testing time, which resulted in some unusual system behaviors. To simulate the passage of time, users had to right-click anywhere and hit "refresh" after responding to alerts. For the testing session, test results and imaging reports were simplified and appeared shorter in length. Consequently, the insights gathered during the simulation may not completely represent the realities and challenges faced by clinicians working in high-pressure environments, caring for their patients who are critically ill. Finally, the IT team faced challenges due to system constraints, as the systems were hardwired with limited ability for upgrades. We were, however, able to incorporate several build-related upgrades through close collaborations. We are still in the process of upgrading the tool based on the observations made during our evaluations and the feedback collected from end users since the tool went live.

Implications for Future Directions

Concerning further CDSS implementation in different hospital settings, we plan to scale and assess the effectiveness of the CDSS for venous thromboembolism prophylaxis guidelines in patients with traumatic brain injury. This will consist of a stepped wedge hybrid effectiveness implementation trial to scale the CDSS across 9 sites within 4 trauma systems.

Conclusions

By using dual usability evaluation methods, leveraging both expert and end user input during the early phases of system development in an iterative manner, we were able to identify a diverse range of usability issues. These issues varied in severity and were relevant to each evaluator based on their respective experience and role. Through conducting this study and our continued efforts to gather feedback from end users, we have been able to optimize the design and functionality of the traumatic brain injury–CDSS to align more effectively with clinicians' needs and workflow. This comprehensive approach has provided us with valuable insights in refining CDSS, ultimately improving its usability and effectiveness in clinical practice.

Acknowledgments

We acknowledge M (Minnesota) Health Fairview information technology and data sciences, Vincent Peta, PhD, and the staff at Stanson (Stanson Premier Inc) for their support in these efforts. We are grateful to clinicians at the Minneapolis location who facilitated our running the pilot and enabling dissemination of the decision support solution. This research was supported by the Agency for Healthcare Research and Quality (AHRQ; R18HS028583, CJT and GBM-M [principal investigators], and the AHRQ and Patient-Centered Outcomes Research Institute [PCORI; grant P30HS029744], Minnesota-Learning Health System scholar program [MN-LHS], and the National Institutes of Health's [NIH] National Center for Advancing Translational Sciences [UM1TR004405, K12TR004373, T32HL007741]). Additional support for MN-LHS is provided by the University of Minnesota Office of Academic Clinical Affairs, Clinical Translational Science Institute, and Center for Learning Health System Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of AHRQ, PCORI, NIH, or MN-LHS.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Usability issues by theme.
[DOCX File, 26 KB - humanfactors_v12i1e60268_app1.docx]

References

1. Get the facts about TBI | concussion | traumatic brain injury. CDC Injury Center. 2023. URL: https://www.cdc.gov/traumatic-brain-injury/data-research/facts-stats/?CDC_AAref_Val=https://www.cdc.gov/traumaticbraininjury/get_the_facts.html [accessed 2025-06-21]
2. Tignanelli CJ, Gipson J, Nguyen A, et al. Implementation of a prophylactic anticoagulation guideline for patients with traumatic brain injury. Jt Comm J Qual Patient Saf 2020 Apr;46(4):185-191. [doi: 10.1016/j.jcjq.2019.11.007] [Medline: 31899154]
3. Geerts WH, Code KI, Jay RM, Chen E, Szalai JP. A prospective study of venous thromboembolism after major trauma. N Engl J Med 1994 Dec 15;331(24):1601-1606. [doi: 10.1056/NEJM199412153312401] [Medline: 7969340]
4. Norwood SH, McAuley CE, Berne JD, et al. Prospective evaluation of the safety of enoxaparin prophylaxis for venous thromboembolism in patients with intracranial hemorrhagic injuries. Arch Surg 2002 Jun;137(6):696-701. [doi: 10.1001/archsurg.137.6.696] [Medline: 12049541]
5. Wong H, Lovett N, Curry N, Shah K, Stanworth SJ. Antithrombotics in trauma: management strategies in the older patients. J Blood Med 2017;8:165-174. [doi: 10.2147/JBM.S125209] [Medline: 29042825]

XSL•FO
RenderX

6. Phelan HA. Pharmacologic venous thromboembolism prophylaxis after traumatic brain injury: a critical literature review. J Neurotrauma 2012 Jul 1;29(10):1821-1828. [doi: 10.1089/neu.2012.2459] [Medline: 22651698]
7. Yorkgitis BK, Berndtson AE, Cross A, et al. American Association for the Surgery of Trauma/American College of Surgeons-Committee on Trauma Clinical Protocol for inpatient venous thromboembolism prophylaxis after trauma. J Trauma Acute Care Surg 2022 Mar 1;92(3):597-604. [doi: 10.1097/TA.0000000000003475] [Medline: 34797813]
8. Pastorek RA, Cripps MW, Bernstein IH, et al. The Parkland Protocol's modified Berne-Norwood criteria predict two tiers of risk for traumatic brain injury progression. J Neurotrauma 2014 Oct 15;31(20):1737-1743. [doi: 10.1089/neu.2014.3366] [Medline: 24945196]
9. Haut ER, Lau BD, Kraenzlin FS, et al. Improved prophylaxis and decreased rates of preventable harm with the use of a mandatory computerized clinical decision support tool for prophylaxis for venous thromboembolism in trauma. Arch Surg 2012 Oct;147(10):901-907. [doi: 10.1001/archsurg.2012.2024] [Medline: 23070407]
10. Macheel C, Reicks P, Sybrant C, et al. Clinical decision support intervention for rib fracture treatment. J Am Coll Surg 2020 Aug;231(2):249-256. [doi: 10.1016/j.jamcollsurg.2020.04.023] [Medline: 32360959]
11. Moja L, Friz HP, Capobussi M, et al. Effectiveness of a hospital-based computerized decision support system on clinician recommendations and patient outcomes: a randomized clinical trial. JAMA Netw Open 2019 Dec 2;2(12):e1917094. [doi: 10.1001/jamanetworkopen.2019.17094] [Medline: 31825499]
12. Clinical decision support. Agency for Healthcare Research and Quality. 2022. URL: https://www.ahrq.gov/cpi/about/otherwebsites/clinical-decision-support/index.html [accessed 2025-06-21]
13. Loftus TJ, Altieri MS, Balch JA, et al. Artificial intelligence-enabled decision support in surgery: state-of-the-art and future directions. Ann Surg 2023 Jul 1;278(1):51-58. [doi: 10.1097/SLA.0000000000005853] [Medline: 36942574]
14. Nguyen AS, Yang S, Thielen BV, et al. Clinical decision support intervention and time to imaging in older patients with traumatic brain injury. J Am Coll Surg 2020 Sep;231(3):361-367. [doi: 10.1016/j.jamcollsurg.2020.05.023] [Medline: 32561447]
15. Durieux P, Nizard R, Ravaud P, Mounier N, Lepage E. A clinical decision support system for prevention of venous thromboembolism: effect on physician behavior. JAMA 2000 Jun 7;283(21):2816-2821. [doi: 10.1001/jama.283.21.2816] [Medline: 10838650]
16. Shah S, Switzer S, Shippee ND, et al. Implementation of an anticoagulation practice guideline for COVID-19 via a clinical decision support system in a large academic health system and its evaluation: observational study. JMIR Med Inform 2021 Nov 18;9(11):e30743. [doi: 10.2196/30743] [Medline: 34550900]
17. Kawamoto K, Lobach DF. Clinical decision support provided within physician order entry systems: a systematic review of features effective for changing clinician behavior. AMIA Annu Symp Proc 2003;2003:361-365. [Medline: 14728195]
18. Grechuta K, Shokouh P, Alhussein A, et al. Benefits of clinical decision support systems for the management of noncommunicable chronic diseases: targeted literature review. Interact J Med Res 2024 Nov 27;13(1):e58036. [doi: 10.2196/58036] [Medline: 39602213]
19. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. NPJ Digit Med 2020;3(1):17. [doi: 10.1038/s41746-020-0221-y] [Medline: 32047862]
20. Cho H, Keenan G, Madandola OO, et al. Assessing the usability of a clinical decision support system: heuristic evaluation. JMIR Hum Factors 2022 May 10;9(2):e31758. [doi: 10.2196/31758] [Medline: 35536613]
21. Corrao NJ, Robinson AG, Swiernik MA, Naeim A. Importance of testing for usability when selecting and implementing an electronic health or medical record system. J Oncol Pract 2010 May;6(3):120-124. [doi: 10.1200/JOP.200017] [Medline: 20808553]
22. Liberati EG, Ruggiero F, Galuppo L, et al. What hinders the uptake of computerized decision support systems in hospitals? A qualitative study and framework for implementation. Implement Sci 2017 Sep 15;12(1):113. [doi: 10.1186/s13012-017-0644-2] [Medline: 28915822]
23. Khairat S, Marc D, Crosby W, Al Sanousi A. Reasons for physicians not adopting clinical decision support systems: critical analysis. JMIR Med Inform 2018 Apr 18;6(2):e24. [doi: 10.2196/medinform.8912] [Medline: 29669706]
24. Jones EK, Banks A, Melton GB, Porta CM, Tignanelli CJ. Barriers to and facilitators for acceptance of comprehensive clinical decision support system-driven care maps for patients with thoracic trauma: interview study among health care providers and nurses. JMIR Hum Factors 2022 Mar 16;9(1):e29019. [doi: 10.2196/29019] [Medline: 35293873]
25. Rappold JF, Sheppard FR, Carmichael Ii SP, et al. Venous thromboembolism prophylaxis in the trauma intensive care unit: an American Association for the Surgery of Trauma Critical Care Committee Clinical Consensus Document. Trauma Surg Acute Care Open 2021;6(1):e000643. [doi: 10.1136/tsaco-2020-000643] [Medline: 33718615]
26. 10 usability heuristics for user interface design. Nielsen Norman Group. 2023. URL: https://www.nngroup.com/articles/ten-usability-heuristics [accessed 2025-06-21]
27. Severity ratings for usability problems: article by Jakob Nielsen. Nielsen Norman Group. 2023. URL: https://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems [accessed 2025-06-21]
28. Lewis JR. The System Usability Scale: past, present, and future. Int J Hum-Comput Interact 2018 Jul 3;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]

XSL•FO
RenderX

29. Sauro J. 10 things to know about the Single Ease Question (SEQ). Measuring U. 2012. URL: https://measuringu.com/seq10/ [accessed 2025-06-21]
30. Lewis J, Sauro J. How to estimate SUS using the UX-lite. Measuring U. 2021. URL: https://measuringu.com/how-to-estimate-sus-with-ux-lite/ [accessed 2025-07-30]
31. Bailey RW, Allan RW, Raiello P. Usability testing vs. heuristic evaluation: a head-to-head comparison. Proc Hum Factors Soc Ann Meet 1992 Oct;36(4):409-413. [doi: 10.1177/154193129203600431]
32. Desurvire HW. Faster, cheaper!! Are usability inspection methods as effective as empirical testing? In: Usability Inspection Methods: John Wiley & Sons, Inc; 1994:173-202.

Abbreviations

BPA: Best Practice Advisory
CDSS: clinical decision support system
CPG-FHIR: clinical practice guidelines on Fast Healthcare Interoperability Resources
***ICD-10-CM*:** *International Statistical Classification of Diseases, Tenth Revision, Clinical Modification*
IT: information technology
SEQ: Single Ease Question
SUS: Systems Usability Scale
UMN: University of Minnesota

Edited by A Kushniruk; submitted 06.05.24; peer-reviewed by H Witteman, Y Quintana; revised version received 23.04.25; accepted 12.05.25; published 30.07.25.

Please cite as:
Rizvi RF, Faisal S, Sussman M, Mendlick P, Brown S, Lindemann E, Switzer S, Melton GB, Tignanelli CJ
Leveraging Dual Usability Methods to Evaluate Clinical Decision Support Among Patients With Traumatic Brain Injury: Mixed Methods Study
JMIR Hum Factors 2025;12:e60268
URL: https://humanfactors.jmir.org/2025/1/e60268
doi:10.2196/60268

© Rubina F Rizvi, Sameen Faisal, Mark Sussman, Patricia Mendlick, Sam Brown, Elizabeth Lindemann, Sean Switzer, Genevieve B Melton, Christopher J Tignanelli. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 30.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Digital Capability, Open-Source Use, and Interoperability Standards Within the National Health Service in England: Survey of Health Care Trusts

Matthew Russell Bennion[1,2], BEng, MBA, MSc, PhD; Ross Spencer[3], BSc, MA; Roger K Moore[1], BA, MSc, PhD; Richard Kenyon[4], BSc

[1]School of Computer Science, University of Sheffield, Regent Court, 211 Portobello, Sheffield, United Kingdom

[2]Digital Innovation Unit, NHS Midlands and Lancashire Commissioning Support Unit, Stoke-on-Trent, United Kingdom

[3]Unaffiliated, Leipzig, Germany

[4]Aston Business School, Aston University, Birmingham, United Kingdom

Corresponding Author:
Matthew Russell Bennion, BEng, MBA, MSc, PhD
School of Computer Science, University of Sheffield, Regent Court, 211 Portobello, Sheffield, United Kingdom

Abstract

Background: In 2016, the National Health Service (NHS) England sought to drive digital transformation within select NHS trusts through the Global Digital Exemplar (GDE) program. While the program did advance the NHS's integration with digital technologies, disparities in digital maturity persisted between GDE-funded and nonfunded NHS trusts. The Department of Health and Social Care (DHSC) launched a data strategy in 2022 that aimed to develop the appropriate technical infrastructure and data architecture to enable more effective and efficient use of its data. Given the diversity in digital capabilities, open-source adoption, and interoperability standards within NHS services, official guidance has continued to struggle to provide effective unification. Data about capabilities and technologies from application development teams in the NHS trusts, crucial for advancing these areas, remains insufficient.

Objective: This study aimed to further document the capabilities and technologies used in the NHS to develop digital capacity, comparing those with standard funding against those with additional GDE funding. This comparative analysis provides a foundational understanding for evaluating current practices and identifying potential areas for improvement in the NHS digital transformation efforts.

Methods: This study was conducted using Freedom of Information (FOI) requests and systematic website searches. The Freedom of Information Act (FOIA) allows individuals to request information held by public authorities. This process supports transparency and accountability by ensuring public access to government data. Data were compiled from two sources: (1) FOI requests submitted to NHS trusts between July 2020 and December 2020, and (2) systematic website searches for technology conducted between August 2020 and July 2021. A series of chi-square tests was conducted to validate and strengthen the robustness of the FOI questions.

Results: A total of 191 (84.5%) of the then 226 NHS trusts completed the FOI request, and 161 of the 191 (84%) had software and app development, website, or innovation teams. A total of 112 (69.6%) teams developed front-facing service user websites and apps. Out of 191, 150 (93.2%) worked with clinical staff to formulate innovative ideas, 55 (34.2%) carried out developments for other trusts and external entities, 35 (21.7%) had attempted to secure an innovation grant, and 138 (86%) disclosed the technologies they use. A total of 25 (15.5%) said they always used open-source technology, and 24 (17%) disclosed technologies associated with interoperability standards in their responses.

Conclusions: The NHS must adopt a cohesive strategy and refine policies to ensure the success of its digital, open-source technology and interoperability standards initiatives. Five recommendations toward greater organizational interoperability are made by the authors. Future research should examine digital innovation across NHS trusts, focusing on barriers such as limited resources, organizational culture, and technical expertise. Identifying these challenges is essential for developing strategies to reduce disparities and promote equal progress.

(JMIR Hum Factors 2025;12:e66398) doi:10.2196/66398

XSL•FO
RenderX

KEYWORDS

NHS; healthcare; open source; technologies; digital capability; capabilities; GDE funding; interoperability; surveys; England; Global Digital Exemplar; comparative analysis; document; initiative; Department of Health and Social Care; National Health Service

Introduction

Background

In 2022, the Department of Health and Social Care (DHSC) released its data strategy, "Data Saves Lives: Reshaping Health and Social Care with Data," to improve how the National Health Service (NHS) uses data [1]. A key goal of the strategy was to develop the appropriate technical infrastructure to ensure that the data architecture underpinning the NHS works in unison, enabling more effective and efficient use of data. The data strategy would be carried out via a commitment to develop and publish a draft of the standards alongside an interoperability strategy to ensure that fit-for-purpose standards are widely adopted across health and adult social care. The anticipated standards would reuse and build upon international standards where applicable. However, the NHS's current ability to fulfill this strategy across hospitals and other facilities is not well understood due to a lack of visibility into its digital development capabilities.

The Global Digital Exemplar (GDE) program was conceived in 2016 by NHS England to achieve digital transformation in select NHS trusts. It was hoped that the program would create a knowledge-sharing ecosystem to spread learning to other NHS trusts [2]. The program resulted in the creation of 16 GDE acute trusts [3,4], 3 GDE ambulance trusts [5,6], and 7 GDE mental health trusts [7].

While the GDE program improved the use of digital within the NHS, studies carried out since have highlighted that NHS trusts and social care providers still lack digital maturity [8], and that a digital divide has emerged between NHS trusts with and without GDE funding [9,10].

The "Data Saves Lives" strategy uses user stories to describe various systems that can work together in a secure, safe, accessible, and interoperable manner. Within this strategy, the UK Health Security Agency (UKHSA) aims to prioritize FAIR (findability, accessibility, interoperability, and reusability) principles [11]. In the context of FAIR, interoperability is defined as "the data need to interoperate with applications or workflows for analysis, storage, and processing." However, interoperability should not be viewed solely from a data perspective; it must also encompass system interfaces, the methods by which data are produced and consumed across the NHS ecosystem. While a universal approach to the architecture of NHS systems is impractical, achieving interoperability requires a commitment to key design considerations across data and application interfaces. This ensures the secure, safe, and effective use and reuse of data and information across system and organizational boundaries.

To complement interoperability, NHS England's Transformation Directorate is also in pursuit of making source code within the NHS freely available and downloadable through the use of open-source. Both open-source and interoperability can save NHS services from duplicating effort and help them build better services faster through shared resources [12]. The NHS Transformation Directorate presently promotes the open-source languages Python (Python Software Foundation) and R (R Foundation for Statistical Computing) through their network and engagement communities as common standards for programming [13].

Current Study

The digital capabilities of NHS services in England are diverse, with ever-evolving services providing differing levels of digital capability, open-source use, and interoperability standards. Official recommendations and advice struggle to unify the NHS effectively. NHS trust application development teams play a vital role in uptake in these areas, but in the present landscape, the NHS has limited visibility of these teams and their capabilities. NHS trust application development teams are focused on the creation and deployment of digital technologies designed to improve clinical care, patient outcomes, health system efficiency, and user engagement. This includes the development of software applications for clinical workflows or patient self-management, websites that provide access to health information or services, and innovative technological solutions such as artificial intelligence tools, remote monitoring systems, or digital diagnostics that address existing challenges or create new capabilities within health care delivery. The key objective of this paper is to illuminate the current state of NHS trust application development teams in England by documenting their capabilities and technologies, while comparing standard trusts with those that have received extra funding via the GDE program, thus providing a starting point for the evaluation of current practice.

Methods

Design

We documented the capabilities and technologies used in the NHS to develop digital products and services. Our data source was Freedom of Information (FOI) requests sent to NHS trusts in England. In our FOIs to NHS trusts, we asked for the disclosure of all departments and teams handling application and software development, websites, and technical innovation, along with the number of team members and roles. Further information regarding the teams was requested, covering technologies and software design methodologies used, use of open source, development of service user websites and apps, innovative collaboration with clinical staff, tenders for system development, developments for others, selling systems commercially, and use of innovation grants. Eleven of the 14 FOI questions were designed to elicit clear binary responses (1=yes, 0=no), which made them straightforward to answer and reduced the potential for variability or interpretation bias across NHS trusts. Respondents could still provide additional detail if

XSL•FO
RenderX

available, but the core data remained consistent (Table 1). To ensure the information received was relevant, focused, and reflective of genuine NHS digital capabilities, all trusts were appraised against the inclusion criterion of having teams and departments handling application and software development, websites, and technical innovation. To meet this criterion, the NHS trust had to respond "yes" to question 1 (Table 1). The explicit exclusion criteria were trusts without dedicated teams handling application and software development, websites, or technical innovation; trusts that did not respond or declined to provide information; and those that submitted their FOI response after December 31, 2020.

Table . Questions asked of National Health Service trusts.

Question	Wording
1	Does your trust have teams/departments that handle any of the following: application/software development, websites, or technical innovation? (Yes/No). If yes, please name these teams/departments.
2	How many members are in these teams/departments? What roles are the teams/departments made up of?
3	What technologies do the teams/departments use? For example: HTML5, C#, SQL, and .NET Core 2.0.
4	Are the developments of the teams/departments open source? (Yes/No). If yes, please provide any details you may be able to disclose regarding this.
5	What software methodology do the teams/departments use? (For example, one or more of the following: PRINCE2, AGILE, or Waterfall).
6	Do the teams/departments develop front-facing service user websites/apps? (Yes/No). If yes, please give any details you may be able to disclose regarding this.
7	Do the teams/departments work with clinical staff to formulate any innovative ideas they may have? (Yes/No). If yes, please give any details you may be able to disclose regarding this (digital innovation refers to the application of digital technology to existing business problems).
8	Have the teams/departments ever gone for external tenders for health care system developments? (Yes/No). If yes, please give any details you may be able to disclose regarding this.
9	Have the teams/departments ever done developments for other trusts/external entities? (Yes/No). If yes, please provide any details you may be able to disclose regarding this.
10	Have the teams/departments ever sold a development they produced commercially? (Yes/No). If yes, please provide any details you may be able to disclose regarding this.
11	Have the teams/departments ever attempted to secure an innovation grant? (Yes/No). If yes, please provide any details you may be able to disclose regarding this.
12	Were the teams/departments used during the COVID-19 crisis? (Yes/No). If yes, please provide any details you may be able to disclose regarding this. If no, please explain why they were not used.
13	Does your service provide mental health services? (Yes/No). If yes, have the teams/departments been involved in developing these digitally/online?
14	Do the teams/departments feel they have good visibility within the trust regarding the services they can offer? (Yes/No). Please provide any details you may be able to disclose regarding this.

All technologies were appraised against the inclusion criterion of being a programming language, framework, library, database, query language, interoperability standard, application programming interface, web service, or version control and collaboration platform. To meet this criterion, it had to be possible to find the technology via a Google search when the technology's name was entered as the search term, and the returned results identified the technology as matching one of the inclusion criteria. Examples of matches that were deemed acceptable included legitimate sources such as developer websites, corporations, and media outlets.

Procedure

On June 8, 2020, a list of NHS trusts in the United Kingdom was requested through an FOI email to NHS England, asking for the contact details of all NHS trusts within the United Kingdom. This yielded a list of 226 NHS trusts. During July 2020, the FOI email address and FOI contact form for each trust

XSL•FO
RenderX

were identified, and an FOI email request was sent to the Freedom of Information Act (FOIA) officer or department of each of the 226 trusts. The questions asked are reported in Table 1. Each initial question had a further open-ended subquestion asking for more details when a trust confirmed the existence of any associated information. Since FOI responses vary in format and structure across NHS trusts, the data had to be interpreted on a case-by-case basis rather than through a uniform extraction template. According to the FOIA, requests must be answered within 20 working days of receipt [14]. Due to the COVID-19 pandemic, the author allowed until December 31, 2020, for a response to keep all data within the same year. Retrospective ethical approval was requested from the Aston University Department of Business and Social Sciences Ethics Committee.

Content and Data Analysis

A conceptual content analysis methodology was used to conduct the data analysis. This approach facilitated the extraction of both quantitative and qualitative insights from the open-ended responses obtained via FOI requests. Specifically, it enabled the generation of a high-level statistical overview of digital capabilities across NHS Trusts, as well as a more granular examination of the technologies and methodologies reported by these institutions in relation to the study objectives.

To support systematic analysis, a codebook was developed to categorize the FOI responses (Textbox 1). Based on this codebook, the lead author designed and implemented a bespoke data management system through which the FOI response data were processed. Following data entry, the coded responses were exported as a complete dataset for statistical analysis using SPSS (IBM Corp).

XSL•FO
RenderX

Textbox 1. Codebook.

Question 1: Does your trust have teams/departments that handle any of the following: application/software development, websites, technical innovation?

Response options:

1. No
2. Yes
3. N/A

Number of teams: if option 1 was coded, calculate the total number of teams based on the response.

Question 2: How many members are in these teams/departments?

Response options:

1. Undisclosed
2. Disclosed
3. Ambiguous
4. Not applicable
5. Unanswered
6. Unheld information
7. Declined under a section of the Freedom of Information Act.

Number of team members: If 1 response 1 was coded, calculate the total number of team members based on the response.

Question 3: What technologies do the teams/departments use? For example: HTML5, C#, SQL, and .NET Core 2.0.

Response options:

1. Undisclosed
2. Disclosed
3. Ambiguous
4. Not applicable
5. Unanswered
6. Unheld information
7. Declined under a section of the Freedom of Information Act.

Technologies and technology categories: if response 1 was coded, list all technologies in the response and code each technology against one of the following categories:

- Programming language: a language used by a programmer to communicate with computers. They are mainly used to develop desktop, website, and mobile apps.
- Framework: a structure upon which software can be built. They serve as a foundation, so a programmer does not have to start entirely from scratch every time they write a new program.
- Library: a collection of prewritten code that programmers can use to optimize tasks.
- Database: an organized collection of structured information, typically stored electronically in a computer system.
- Query language: a language used to make queries in databases and information systems.
- Interoperability standard: a standard that enables the operational processes, underlying exchange, and sharing of information between different systems.
- API or web service: a technology that enables the transfer of data between separate software applications.
- Version control and collaboration platforms: facilitate coordination, sharing, and collaboration across the entire software development team. Version control software enables teams to work in distributed and asynchronous environments, manage changes and versions of code and artifacts, and resolve merge conflicts and related anomalies.
- Other

Question 4: Are the developments of the teams/departments open source? (Yes/No). If yes, please give any details you may be able to disclose regarding this.

Response options:

XSL•FO
RenderX

1. No
2. Yes
3. Sometimes
4. Ambiguous
5. Not applicable
6. Unanswered
7. Unheld information
8. Declined under a section of the Freedom of Information Act.

Question 5: What software methodology (eg, PRINCE2, AGILE, or Waterfall) do the teams/departments use?

Response options:

1. Undisclosed
2. Disclosed
3. Ambiguous
4. Not applicable
5. Unanswered
6. Unheld information
7. Declined under a section of the Freedom of Information Act.

Software design methodologies: if response 1 was coded, all methodologies should be appraised and coded against the inclusion criterion of being a software design methodology.

Question 6: Do the teams/departments develop front-facing service user websites/apps? (Yes/No). If yes, please give any details you may be able to disclose regarding this.

Response options:

1. No
2. Yes
3. Ambiguous
4. Not applicable
5. Unanswered
6. Unheld information
7. Declined under a section of the Freedom of Information Act.

Question 7: Do the teams/departments work with clinical staff to formulate any innovative ideas they may have? (Yes/No). If yes, please give any details you may be able to disclose regarding this.

Response options:

1. No
2. Yes
3. Ambiguous
4. Not applicable
5. Unanswered
6. Unheld information
7. Declined under a section of the Freedom of Information Act.

Question 8: Have teams and departments ever done developments for other trusts or external entities?

Response options:

1. No
2. Yes

XSL•FO
RenderX

3. Ambiguous
4. Not applicable
5. Unanswered
6. Unheld information
7. Declined under a section of the Freedom of Information Act.

Questions 9: Have teams and departments ever attempted to secure an innovation grant?

Response options:

1. No
2. Yes
3. Ambiguous
4. Not applicable
5. Unanswered
6. Unheld information
7. Declined under a section of the Freedom of Information Act.

A detailed classification of the types of data extracted from the FOI responses is provided in Table 2, as categorized using the codebook. The dataset predominantly consisted of dichotomous variables.

Table . Collected data types from content analysis.

Question	Collected data types
1	• Yes or no (dichotomous) • Number of teams (numerical) • Details (qualitative)
2	• Number of members (numerical) • Details (qualitative)
3	• Technology names (nominal) • Technology categories (nominal) • Technology companies (nominal) • Uses legacy tech (dichotomous) • Details (qualitative)
4	• Yes or no (dichotomous)
5	• Methodology names (nominal) • Details (qualitative)
6	• Yes or no (dichotomous) • Details (qualitative)
7	• Yes or no (dichotomous) • Details (qualitative)
8	• Yes or no (dichotomous) • Details (qualitative)
9	• Yes or no (dichotomous) • Details (qualitative)

To establish whether the answers to these questions differed as a function of trust funding (standard or GDE), tests of association were used. A series of chi-square tests was conducted as an additional layer of validation to enhance the robustness of the findings for questions 1, 3, 4, 6, 7, 9, and 11. For analysis, “No” and “N/A” responses were collapsed into a single category, and only “Yes” and “No” responses were included in the final tests. Assumptions were checked, and when violated, a Fisher extract was used instead. All analyses were conducted using SPSS for Windows (version 25; IBM Corp).

XSL•FO
RenderX

Ethical Considerations

The Aston University Department of Business and Social Sciences Ethics Committee has reviewed and granted approval for the study. The study used UK FOIA data from NHS trusts, a process that supports transparency and public access to government-held information. The ethical considerations primarily focused on ensuring adherence to FOIA guidelines and that the requests were legitimate, appropriate, and respectful of the privacy rights of individuals and NHS trusts involved.

As the study did not involve human participants directly, and the data requested were publicly available, there were no concerns regarding private or sensitive data. Ethical issues related to informed consent and participant confidentiality were not applicable. However, all data requests were made in alignment with the principles of transparency, fairness, and public accountability set out by the FOIA.

In addition, the study adhered to best practices for data handling and reporting, ensuring that the requested data were used responsibly and in compliance with data protection laws where applicable. Any personal or sensitive information was either excluded or anonymized to protect confidentiality. To ensure fairness and safeguard the identity of NHS trusts, all reporting was done in a manner that made it impossible to trace data back to individual organizations, unless the FOI responses were requested directly from the NHS Trusts.

Results

Overview

We present the data pertaining to (1) response rates, (2) capabilities, (3) technologies, (4) open source, and (5) interoperability standards.

Response Rates

Out of 226 trusts, 191 (84.5%) completed the FOI request, 23 (10.2%) acknowledged receiving the FOI request but never provided followed up with answers, 7 (3.1%) had merged with another NHS trust, 2 (0.9%) sent partial information consisting of an organizational plan, 2 (0.9%) refused the FOI because it would take too long to complete, and 1 (0.4%) FOI email address no longer accepted emails. In total, 161 (84.3%) met the inclusion criteria by having teams or departments handling application and software development, websites, and technical innovation. The final sample comprised 83% (134/161) standard trusts and 17% (27/161) GDE trusts.

Capabilities

Have Software and App Development, Website, and Innovation Teams

Overall, 161 out of 191 (84%) trusts said they had software and app development, website, and innovation teams. The chi-square assumptions were marginally violated, with 25% of the cells (1 out of 5) having an expected count of less than 5; the minimum expected count was 4.87, so a Fisher exact test was run. According to the Fisher exact test, there was a nonsignificant association between trust type and having teams (P=.79). This suggests no difference in team presence between standard and GDE trusts.

Developing Front-Facing Service User Websites and Apps

Overall, 112 out of 161 (69.6%) trust teams developed front-facing service user websites and apps. A total of 43 (26.7%) trusts said no, 3 (1.9%) said the question was not applicable, 1 (0.6%) declined to answer the question on grounds of security, 1 (0.6%) stated the information was uncaptured, and 1 (0.6%) did not answer the question. There was a nonsignificant association between trust type and developing front-facing service user websites and apps (χ^2_1=1.47; P=.23). This suggests no difference between standard and GDE trusts.

Work With Clinical Staff to Formulate Innovative Ideas

Overall, 150 out of 161 (93.2%) of trust teams work with clinical staff to formulate innovative ideas. A total of 8 (5%) trusts said no, 2 (1.2%) said the question was not applicable, and 1 (0.6%) stated the information was uncaptured. The chi-square assumptions were marginally violated, with 25% of the cells having expected counts less than 5, and the minimum expected count was 1.69, so a Fisher exact test was run. According to the test, there was a nonsignificant association between trust type and the work with clinical staff to formulate innovative ideas (P=.21). This suggests no difference between standard and GDE trusts.

Carried Out Developments for Other Trusts and External Entities

Overall, 55 out of 161 (34.2%) trust teams carry out developments for other trusts or external entities. One hundred (62.1%) of trusts said no, 3 (1.9%) said the question was not applicable, 1 (0.6%) declined to answer the question on grounds of security, 1 (0.6%) stated the information was uncaptured, and 1 (0.6%) did not answer the question. There was a nonsignificant association between trust type and carrying out developments for other trusts and external entities (χ^2_1=0.031; P=.86). This suggests no difference between standard and GDE trusts.

Attempted to Secure an Innovation Grant

Overall, 35 out of 161 (21.7%) trust teams have attempted to secure an innovation grant. A total of 118 (73.3%) trusts said no, 3 (1.9%) said the question was not applicable, 4 (2.5%) stated the information was uncaptured, and 1 (0.6) declined to answer on the grounds of security. There was a nonsignificant association between trust type and attempting to secure an innovation grant (χ^2_1=1.57; P=.21). This suggests no difference in attempts to secure innovation grants between standard and GDE trusts.

Technologies

Overall, 138 out of 161 (86%) trusts with teams disclosed the technologies they used. A total of 84 technologies were identified: 27 programming languages, 19 frameworks, 10 libraries, 11 databases, 1 query language, 7 interoperability standards, 5 application programming interfaces or web services,

XSL•FO
RenderX

and 4 version control and collaboration platforms. All technologies are summarized in Table 3.

XSL•FO
RenderX

Table . All technologies reported to be used by the National Health Service for digital development.

Category and technology	Developer or organization	NHS[a] trusts using (N=138), n (%)
Query language		
SQL[b]	ISO[c]	88 (63.8%)
Programming languages		
C#	Microsoft	81 (58.7%)
JavaScript	ECMA[d] International	54 (39.1%)
PHP[e]	PHP Group	20 (14.5%)
VB.NET[f]	Microsoft	15 (10.9%)
VB[g]	Microsoft	14 (10.1%)
Java	Oracle Corporation	10 (7.2%)
Python	Python Software Foundation	7 (5.1%)
VBA[h]	Microsoft	4 (2.9%)
Delphi	Embarcadero Technologies, Incorporated	3 (2.2%)
C++	ISO and IEC[i]	3 (2.2%)
Ruby	Open Standard	3 (2.2%)
CFML[j]	Adobe	2 (1.4%)
XSLT[l]	W3C[k]	2 (1.4%)
R	R Core Team	2 (1.4%)
ObjectScript	InterSystems	2 (1.4%)
PowerShell	Microsoft	2 (1.4%)
Oracle PL/SQL[m]	Oracle Corporation	2 (1.4%)
TypeScript	Microsoft	1 (0.7%)
Meditech Magic	Medical Info Tech, Incorporated	1 (0.7%)
ASP.NET[n] Razor	Microsoft	1 (0.7%)
C	ANSI[o]	1 (0.7%)
ECMAScript[p]	ECMA International	1 (0.7%)
XAML[q]	Microsoft	1 (0.7%)
Elixir	Plataformatec	1 (0.7%)
VB 6	Microsoft	1 (0.7%)
Lua	Tecgraf[r]	1 (0.7%)
Perl	Larry Wall	1 (0.7%)
Frameworks		
.NET	Microsoft	42 (30.4%)
.NET Core	Microsoft	40 (29.0%)
ASP.NET MVC[s]	Microsoft	22 (15.9%)
ASP.NET	Microsoft	21 (15.2%)
Classic ASP[t]	Microsoft	8 (5.8%)
Angular	Google	8 (5.8%)
Entity Framework	Microsoft	5 (3.6%)
Xamarin	Microsoft	5 (3.6%)

XSL•FO
RenderX

Category and technology	Developer or organization	NHS[a] trusts using (N=138), n (%)
ASP.NET WebForms	Microsoft	4 (2.9%)
AngularJS	Google	3 (2.2%)
Vue.js	Evan You	2 (1.4%)
Ionic	Drifty	2 (1.4%)
WPF[u]	Microsoft	2 (1.4%)
Blazor	Microsoft	1 (0.7%)
Symfony	Symfony community	1 (0.7%)
Hibernate	Red Hat Software	1 (0.7%)
LINQ[v]	Microsoft	1 (0.7%)
Spring	VMware	1 (0.7%)
Oracle Forms	Oracle Corporation	1 (0.7%)
Libraries		
jQuery	The jQuery Team	21 (15.2%)
React	Meta and community	4 (2.9%)
AJAX[w]	Adaptive Path	3 (2.2%)
Knockout.js	Knockout	1 (0.7%)
SignalR	Microsoft	1 (0.7%)
SweetAlert	Tristan Edwards	1 (0.7%)
TensaFlow	Google Brain Team	1 (0.7%)
chart.js	Chart.js Team	1 (0.7%)
jQuery UI[x]	The jQuery Team	1 (0.7%)
WinForms	Microsoft	1 (0.7%)
Databases		
Microsoft SQLServerr	Microsoft	29 (21.0%)
Caché and Ensemble	InterSystems	12 (8.7%)
MySQL[y]	Oracle Corporation	4 (2.9%)
PostgreSQL	PostgreSQL Global Group	4 (2.9%)
Oracle	Oracle Corporation	3 (2.2%)
PouchDB	Apache Software Foundation	1 (0.7%)
MariaDB	MariaDB Foundation	1 (0.7%)
MongoDB	MongoDB, Incorporated	1 (0.7%)
CouchDB	Apache Software Foundation	1 (0.7%)
SQlite	Richard Hipp	1 (0.7%)
NoSQL	Unidentifiable	1 (0.7%)
Interoperability		
HL7[z]	HL7 International	11 (8.0%)
JSON[aa]	ECMA International	10 (7.2%)
XML[ab]	W3C	7 (5.1%)
Arden Syntax	HL7 International	3 (2.2%)
Mirth Connect	NextGen Healthcare	3 (2.2%)
FHIR[ac]	HL7 International	3 (2.2%)

XSL•FO
RenderX

Category and technology	Developer or organization	NHS[a] trusts using (N=138), n (%)
NextGen Connect	NextGen Healthcare	1 (0.7%)
APIs[ad] and web services		
Web Services	Unidentifiable	3 (2.2%)
RESTful APIs	Unidentifiable	3 (2.2%)
WebAPI[ae]	Microsoft	3 (2.2%)
SOAP[af] XML Web services	Unidentifiable	2 (1.4%)
APIs	Unidentifiable	1 (0.7%)
Version control and collaboration platform		
Git	Junio Haman	5 (3.6%)
Azure DevOps[ag]	Microsoft	4 (2.9%)
TFS[ah]	Microsoft	2 (1.4%)
GitLab	GitLab Incorporated	1 (0.7%)

[a]NHS: National Health Service.
[b]SQL: Structured Query Language.
[c]ISO: International Organization for Standardization.
[d]ECMA: European Computer Manufacturers Association.
[e]PHP: Hypertext Preprocessor.
[f]VB.NET: Visual Basic .NET.
[g]VB: Visual Basic.
[h]VBA: Visual Basic for Applications.
[i]IEC: International Electrotechnical Commission.
[j]CFML: ColdFusion Markup Language.
[k]W3C: World Wide Web Consortium.
[l]XSLT: Extensible Stylesheet Language Transformations.
[m]PL/SQL: Procedural Language extensions to the Structured Query Language.
[n]ASP.NET: Active Server Pages .NET.
[o]ANSI: American National Standards Institute.
[p]ECMAScript: European Computer Manufacturers Association Script.
[q]XAML: Extensible Application Markup Language.
[r]Tecgraf: Computer Graphics Technology Group.
[s]MVC: model–view–controller.
[t]ASP: Active Server Pages.
[u]WPF: Windows Presentation Foundation.
[v]LINQ: Language-Integrated Query.
[w]AJAX: Asynchronous JavaScript and XML.
[x]UI: user interface.
[y]MySQL: My Structured Query Language.
[z]HL7: Health Level Seven.
[aa]JSON: JavaScript Object Notation.
[ab]XML: Extensible Markup Language.
[ac]FHIR: Fast Healthcare Interoperability Resources.
[ad]API: Application Programming Interface.
[ae]WebAPI: Web Application Programming Interface.
[af]SOAP: Simple Object Access Protocol.
[ag]DevOps: Development and Operations.
[ah]TFS: Team Foundation Server.

Python, the programming language used and promoted by NHS England and the NHS Transformation Directorate, was disclosed by 7 out of 138 (5.1%) trusts with teams that disclosed technologies. R, the programming language used for statistics

XSL•FO
RenderX

and data science, was disclosed by 2 out of 138 (1.4%) trusts with teams that disclosed technologies. The most disclosed query language was Structured Query Language with 88 out of 138 (63.8%) trusts, and the most disclosed programming language was C# with 81 out of 138 (58.7%) trusts.

Open Source

Overall, 25 out of 138 (15.5%) trust teams said they always used open-source technology, and 9 out of 138 (5.6%) said they used open-source sometimes. In total, 119 (73.9%) trusts said no, 6 (3.7%) said the question was not applicable, and 2 (1.2%) declined to answer on the grounds of security. There was a nonsignificant association between team type and open-source use (χ^2_1=0.014; *P*=.91). This suggests no difference in the use of open source between standard and GDE teams. Eight out of 138 (5.8%) trusts with teams that disclosed technologies indicated they used version control and collaboration platforms. A total of 4 platforms were identified in this review. The most used was Git, reported by 5 out of 138 (3.6%) trusts with teams that disclosed technologies.

Interoperability Standards

Overall, 24 out of 138 (17%) trusts with teams that disclosed technologies referred to interoperability standards in their responses. A total of 7 interoperability standards were identified. The most disclosed interoperability technology was HL7, mentioned by 11 out of 138 (8%) trusts.

Discussion

Principal Findings

This study is the first attempt to document the capabilities and technologies of NHS trust application development teams at a particular point in time. While these capabilities and technologies are changeable, the paper provides future researchers, commissioners, and policy makers with a baseline from which to build. The data presented raise several interesting questions relating to capabilities, open-source use, interoperability use, and programming languages.

Data Accessibility, Quality, and Limitations

The study was reliant on the provisions of the UK FOIA (2000) for data collection. Approximately 85% of NHS trusts responded to the request. Although this represents a large sample of trusts, it is difficult to ascertain the extent to which they are representative of all trusts. FOI requests rely on the expertise of those responsible for handling them in any given NHS trust, and the author was unable to determine how accurate and thorough individual responses were. During the data collection period, trusts were under immense pressure due to the COVID-19 pandemic, meaning that many may not have had the capacity to respond accurately. While most trusts addressed the questions posed, some declined to respond on security grounds. These data gaps, however, were minimal.

Capabilities

Digital maturity across NHS trusts is presently considered to be underdeveloped, with previous reports highlighting the need for the NHS to bridge the digital divide to avoid the creation of a 2-tier secondary care system [10]. While this may still be true in some cases, the data presented here highlights that in-house processes to develop digital capability are well established in 84% of responding trusts. Overall, no statistical differences were found between trust types and these results. This may indicate that increasing in-house capability to carry out digital projects was not a consideration when trusts received GDE funding. These results also highlight an issue with digital maturity measures such as HIMSS EMRAM (Healthcare Information and Management Systems Society: Electronic Medical Record Adoption Model). The measure exclusively focuses on technological functionality, such as electronic health records, and does not consider human and organizational capabilities [2]. The apps lead to an incomplete picture of a trust's digital maturity. While no significant differences were found between trust types, the overall results for each question highlight some interesting findings.

A review by Sood and McNeil [15] highlighted that it would be impossible for the NHS to become a modern, effective, and efficient health care system without fully embracing the digital agenda. The results of this survey reveal that trusts are working toward overcoming Sood and McNeil's [15] concern, with 69.6% of respondents reporting the creation of websites and apps to help service users.

A study by Asthana et al [16] investigated how, despite having a tax-funded NHS and a strong policy toward eHealth innovations, the English NHS had been limited in the extent to which it exploited digital potential. The survey results indicate that 93% (150/161) of respondents help clinical staff formulate innovative ideas. So, despite the suggestion of limited exploitation, NHS trusts are actively working to enable their clinical staff to innovate. However, the data do not reveal why, if this practice is commonplace, it has not been exploited to its full potential. Possible reasons may include limited resources regarding time, staffing, and funding.

The NHS Global Exemplar program was implemented to establish a knowledge-sharing ecosystem to spread learning among NHS trusts [2]. The survey results indicate 34% (55/161) carry out this practice. NHS England needs to do more to promote digital collaboration between trusts. Encouraging more digital development between organizations would help innovation and the leveling up of digital NHS services. Making trust teams aware of collaboration platforms such as the FutureNHS, that is used to support the right technology solutions to health and social care [17], could enable better sharing of technical innovations between trusts.

Innovation is critical in NHS England achieving the ambitions set out in the mandate, to increase the pace and scale of change. The NHS's Accelerated Access Collaborative (AAC) reported 2708 innovations being developed in 2019/2020 [18]. Survey data indicate that only 21.7% of respondents had applied for innovation grants. NHS England needs to do more to promote the innovation funding available to trusts. Academic Health Science Networks (AHSNs), now known as the Health Innovation Network, were established by NHS England in 2013 to spread innovation, improve health outcomes, and generate economic growth [19]. If the Health Innovation Network were

XSL•FO
RenderX

made more accessible to trust development teams, the internal potential of trusts to unlock innovation could help to maximize their benefit and better fund digital maturity. In addition, trusts could use the funding streams advised by the Health Innovation Network to independently create new digital innovations. Unfortunately, there is a limited amount of investment available, and not all trusts can benefit from this type of funding, but more should be done to encourage trusts to apply for these grants.

Open-Source Technology

NHS England has tried to instill open-source technology within the NHS for nearly a decade through multiple strategic programs such as the Open-Source Program and Code4Health. The latter aimed to serve as a vessel for promoting open-source technology through clinicians [20]. The survey results revealed that 73.9% of trust development teams do not use open-source approaches, suggesting that previous initiatives have had limited cultural influence, with developers still opting for closed-source methods. Recent publications, such as the report by Goldacre et al [21], have helped to put forward proposals for using open-source implementation and draw focus to NHS data analysts.

GitHub is the NHS's preferred platform for publishing open-source code [13]. However, the results of the survey show that only 5.8% of respondents listed using version control and collaboration platforms, and none of these gave reference to GitHub. NHS England needs to better communicate its desire for the NHS to use GitHub and open source and to work to promote this across all services. Without a unified approach to implementation, GitHub and open-source practices will continue to be used inconsistently across the NHS. The NHS data strategy has committed to making all new source code that is produced or commissioned open and reusable by default [1]. For this commitment to succeed in the NHS, the DHSC needs to go beyond publication of policy and digital guidance. Training and tools will be required to enable this at a regional level, meaning significant investment would be required, as services currently do not have the funding capacity.

There are presently 3 network and engagement communities run by NHS England's Transformation Directorate: Analyst X, NHS-R, and NHS Python [22]. The NHS Python community is focused on the promotion of open-source practices. However, this is done purely in conjunction with the Python programming language. Survey results showed that only 5.1% of respondents disclosed using Python. For open-source to be more widely adopted, a new open-source program of activity is required for the NHS. This should be aimed at the full spectrum of roles within the data process cycle and focus on dominant technologies used by NHS services. This, in turn, will make the community more inclusive to all NHS services and help to increase the adoption of open-source technology. The program should be coordinated at a regional level, with open-source champions that have both technical knowledge and understanding of data processing at the helm. The actions required to carry out this recommendation would be a large undertaking, and whether NHS England has the time or capacity to do this in the current climate should be considered; it may not be feasible. It will also be reliant on trusts being willing to involve themselves in the program.

Interoperability Standards

The objective to improve interoperability has been included in several NHS plans, and the end benefits of advanced interoperability are well recognized [23]. The primary benefit of interoperability is the ability to offer safe and reliable information transfer across the care pathway, reducing the possibility of adverse clinical outcomes [24]. The survey showed 17% listed interoperability standards in the technologies they disclosed. The DHSC has emphasized that it wants the NHS to use Fast Healthcare Interoperability Resources (FHIR) to achieve interoperability (DHSC, 2020). Unfortunately, the most disclosed standard was HL7, the predecessor to FHIR, disclosed by 8% of trust teams. Better promotion of the NHS Digital Interoperability Toolkit would benefit NHS trusts by helping them to standardize their technology and interoperability specifications [25]. Uptake, however, would be reliant on trust and willingness to use the toolkit.

Technologies

Interoperability is reliant on the NHS using technologies that are capable of communicating with one another. The most used programming language named in the survey was C# with 58.7%, a language that differs from NHS England's preferred programming language of Python, which was disclosed by 5.1% of those in the study. The difference in programming language standards puts the Transformation Directorate's push for open source and code sharing at odds with the wider NHS. There is still a benefit to code sharing, but some of this is lost because of the overhead involved in migrating between programming languages. A focus on the most consistently used languages across the NHS services would help to unify the NHS. There will be a strong reliance on the languages they have become accustomed to, and moving away from these is difficult to implement. Any coding examples NHS England develops should be written in the programming languages most appropriate for NHS services. Ensuring the use of modern languages will enable the NHS to keep a competitive advantage and help to retain staff looking to work on more relevant and modern languages over time.

Recommendations

Some of the insights derived from the study and related discussion suggest that the NHS has a solid foundation in certain areas. While trusts may not yet be fully aligned in their use of shared code, source control systems, or in leveraging all available opportunities, such as relevant grants, there is clear evidence of technology utilization across the organization. To build upon these foundations, this paper offers the following recommendations as steps toward achieving a more comprehensive digital maturity.

Promoting Digital Collaboration

Enhancing digital development between organizations can foster innovation and elevate digital NHS services. By making trust development teams aware of collaborations, such as the FutureNHS platform, we can enhance the sharing of technical innovations across trusts.

XSL•FO
RenderX

Leveraging the Health Innovation Network for Innovation

Increasing the accessibility of the Health Innovation Network to trusted development teams could unlock significant internal innovation potential, aiding in the funding and advancement of digital maturity. Trusts could independently develop new digital innovations by using funding streams advised by the Health Innovation Network.

Promoting Open-Source Usage

NHS England should better communicate its advocacy for using GitHub and open-source technologies and actively promote this initiative across all services. The NHS Data Strategy aims to make all new source code open and reusable by default. However, to achieve this, in addition to publishing policy and digital guidance, substantial investment in training and tools is required at a regional level, as current services lack sufficient funding capacity.

Increasing Open-Source Adoption

The survey revealed that only 5.1% of respondents use Python, indicating a need for a comprehensive open-source program across the NHS. This program should target all roles within the data process cycle and focus on the prevalent technologies used in NHS services. By making the community more inclusive, the adoption of open-source technology can be increased. The program should be regionally coordinated, with open-source champions possessing both technical expertise and an understanding of data processing leading the initiative.

Aligning Programming Language Standards

The variation in programming language standards presents a challenge to the Transformation Directorate's push for open source and code sharing across the NHS. Although there are benefits to code sharing, the overhead involved in migrating between different programming languages diminishes some of these benefits. Focusing on the most widely used languages across NHS services could unify efforts. It is essential to ensure that coding examples provided by NHS England are developed in languages most appropriate for NHS services, thereby maintaining a competitive advantage and retaining staff interested in working with modern technologies.

Future Research Needs

Future research should analyze digital innovation across NHS trusts, focusing on barriers such as limited resources, organizational culture, and technical expertise. Understanding these challenges is crucial for creating strategies to reduce disparities and promote equitable progress. Research should also explore ways to increase the adoption of open-source technologies, evaluating programs such as NHS Python and Analyst X, and expanding their reach. Investigating the programming languages and digital tools used across NHS trusts, aligned with NHS England's standards, can enhance interoperability and improve code sharing. In addition, assessing the feasibility and impact of unified programming standards within the NHS's digital infrastructure is essential. Research in these areas could significantly accelerate the NHS's digital transformation, ensuring services remain efficient, innovative, and inclusive, thereby improving health care delivery and stakeholder satisfaction.

Conclusion

Interoperability principles enable the use of data and information across system and organizational boundaries. Embracing interoperability addresses the unrealistic expectation for a one-size-fits-all system for the NHS, or that heterogeneous systems will communicate with one another out of the box. Interoperability reduces the need for greater oversight across the spectrum and allows the NHS to demonstrate maturity at crucial sinews and joints between systems and services.

Without interoperability, much-needed digital maturity will not be gained, and the divide between NHS trusts with and without GDE funding will continue to grow. As the NHS continues to advance its plans for the use of digital, open-source technology and interoperability standards, it is essential that a unified approach is followed. This study has provided a snapshot of the digital development technology landscape and the areas of digital implementation policy that need the most refinement to improve their success.

Acknowledgments

This work was supported by the NHS through a Senior Leader Master's Degree Apprenticeship (SLMDA) awarded to the first author.

Disclaimer

The content is solely the responsibility of the authors and does not represent the views of NHS England.

Authors' Contributions

MRB was responsible for conceptualization, data curation, formal analysis, investigation, methodology, project administration, resource management, software development, validation, visualization, and writing of both the original draft and the review and editing. RS contributed to writing the original draft and participated in the review and editing process. RM was involved in reviewing and editing the manuscript. RK provided supervision for the project.

Conflicts of Interest

None declared.

XSL•FO
RenderX

References

1. Data saves lives: reshaping health and social care with data. : Department of Health and Social Care; 2022 URL: https://www.gov.uk/government/publications/data-saves-lives-reshaping-health-and-social-care-with-data/data-saves-lives-reshaping-health-and-social-care-with-data [accessed 2025-07-21]
2. Cresswell K, Sheikh A, Krasuska M, et al. Reconceptualising the digital maturity of health systems. Lancet Digit Health 2019 Sep;1(5):e200-e201 [FREE Full text] [doi: 10.1016/S2589-7500(19)30083-4] [Medline: 33323267]
3. New plans to expand the use of digital technology across the NHS. Department of Health and Social Care. 2016. URL: https://www.gov.uk/government/news/new-plans-to-expand-the-use-of-digital-technology-across-the-nhs [accessed 2025-07-21]
4. Stevens L. Four more acute global digital exemplars confirmed by NHS england. Digital health. 2017. URL: https://www.digitalhealth.net/2017/03/four-more-acute-global-digital-exemplars-confirmed-by-nhs-england/ [accessed 2025-07-21]
5. Crouch H. Ambulance trust gdes confirmed by NHS england. Digital health. 2018. URL: https://www.digitalhealth.net/2018/01/ambulance-trust-gde-nhs-england/ [accessed 2025-07-21]
6. Ambulance global digital exemplars. NHS England. 2022. URL: https://www.england.nhs.uk/digitaltechnology/connecteddigitalsystems/exemplars/ambulance-global-digital-exemplars/ [accessed 2025-07-21]
7. Mental health global digital exemplars. NHS England. 2022. URL: https://www.england.nhs.uk/digitaltechnology/connecteddigitalsystems/exemplars/mental-health-global-digital-exemplars/ [accessed 2025-07-21]
8. Transformation directorate. NHS England. 2022. URL: https://transform.england.nhs.uk/digitise-connect-transform/ [accessed 2025-07-21]
9. Crouch H. GDE evaluation calls for action to “address digital divide”. Digital health. 2021. URL: https://www.digitalhealth.net/2021/10/gde-evaluation-digital-divide/ [accessed 2025-07-21]
10. Cresswell K, Williams R, Sheikh A. Bridging the growing digital divide between NHS England’s hospitals. J R Soc Med 2021 Mar;114(3):111-112 [FREE Full text] [doi: 10.1177/0141076820974998] [Medline: 33259251]
11. FAIR principles. GO FAIR. 2025. URL: https://www.go-fair.org/fair-principles/ [accessed 2025-07-21]
12. 12 make new source code open. NHS Digital Service Manual. 2019. URL: https://service-manual.nhs.uk/service-standard/12-make-new-source-code-open [accessed 2025-07-21]
13. Executive summary. Github. 2022. URL: https://github.com/nhsx/open-source-policy/blob/main/open-source-policy.md [accessed 2025-07-21]
14. Introduction. ICO Information Commissioner’s Office. 2025. URL: https://ico.org.uk/for-organisations/foi/guide-to-managing-an-foi-request/introduction/ [accessed 2025-07-21]
15. Sood HS, McNeil K. How is health information technology changing the way we deliver NHS hospital care? Future Healthc J 2017 Jun;4(2):117-120. [doi: 10.7861/futurehosp.4-2-117] [Medline: 31098447]
16. Asthana S, Jones R, Sheaff R. Why does the NHS struggle to adopt eHealth innovations? A review of macro, meso and micro factors. BMC Health Serv Res 2019 Dec 21;19(1):984. [doi: 10.1186/s12913-019-4790-x] [Medline: 31864370]
17. Futures for you. Futures. 2020. URL: https://future.nhs.uk/about#about_futurenhs [accessed 2025-07-21]
18. 2019/20 AAC impact figures. NHS England. 2020. URL: https://www.england.nhs.uk/aac/how-are-we-doing/2019-20-aac-impact-figures/ [accessed 2025-07-21]
19. The health innovation network. NHS England. 2025. URL: https://www.england.nhs.uk/aac/what-we-do/innovation-for-healthcare-inequalities-programme/health-innovation-network/ [accessed 2025-07-21]
20. NHS code4health programme will improve digital literacy for clinicians. NHS England. 2015. URL: https://www.england.nhs.uk/2015/03/code4health/ [accessed 2025-07-21]
21. Goldacre B, Bardsley M, Benson T, et al. Bringing NHS data analysis into the 21st century. J R Soc Med 2020 Oct;113(10):383-388. [doi: 10.1177/0141076820930666] [Medline: 32672131]
22. A plan for digital health and social care. Department of Health and Social Care. 2022. URL: https://www.gov.uk/government/publications/a-plan-for-digital-health-and-social-care/a-plan-for-digital-health-and-social-care [accessed 2025-07-21]
23. Zhang J, Sood H, Harrison OT, Horner B, Sharma N, Budhdeo S. Interoperability in NHS hospitals must be improved: the Care Quality Commission should be a key actor in this process. J R Soc Med 2020 Mar;113(3):101-104. [doi: 10.1177/0141076819894664] [Medline: 31904306]
24. Oyeyemi A, Scott P. Interoperability in health and social care: organisational issues are the biggest challenge. BMJ Health Care Inform 2018 Jul;25(3):196-197 [FREE Full text] [doi: 10.14236/jhi.v25i3.1024]
25. Interoperability toolkit. NHS England. 2020. URL: https://digital.nhs.uk/services/interoperability-toolkit [accessed 2025-07-21]

Abbreviations

AAC: Accelerated Access Collaborative
AHSN: Academic Health Science Network
DHSC: Department of Health and Social Care
FAIR: findability, accessibility, interoperability, and reusability

XSL•FO
RenderX

FHIR: Fast Healthcare Interoperability Resources
FOI: Freedom of Information
FOIA: Freedom of Information Act
GDE: Global Digital Exemplar
HIMSS EMRAM: Healthcare Information and Management Systems Society: Electronic Medical Record Adoption Model
NHS: National Health Service
SQL: structured query language
UKHSA: UK Health Security Agency

Edited by E Borycki; submitted 12.09.24; peer-reviewed by S Steele, W Tam; revised version received 21.05.25; accepted 23.06.25; published 12.08.25.

Please cite as:
Bennion MR, Spencer R, Moore RK, Kenyon R
Digital Capability, Open-Source Use, and Interoperability Standards Within the National Health Service in England: Survey of Health Care Trusts
JMIR Hum Factors 2025;12:e66398
URL: https://humanfactors.jmir.org/2025/1/e66398
doi:10.2196/66398

© Matthew Russell Bennion, Ross Spencer, Roger K Moore, Richard Kenyon. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 12.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Health-Related Internet Use Among Outpatients Undergoing Cancer Treatment During the COVID-19 Pandemic: Cross-Sectional Survey Study

Nicole Erickson[1], PhD; Daniel Nasseh[1], PhD; Sarah Seynstahl[1], MD; Nicole Jost[1], MSc; Rachel Wuerstlein[1,2], MD; Jozefina Casuscelli[3], MD; Sebastian Theurich[4], MD; Volker Heinemann[1,5], MD; Theres Fey[1,5], PhD

[1]Comprehensive Cancer Center, Ludwig Maximilian University Hospital Munich, Pettenkoferstrße 8a, Munich, Germany

[2]Department of Obstetrics and Gynecology, Ludwig Maximilian University Hospital Munich, Munich, Germany

[3]Department of Urology, Ludwig Maximilian University Hospital Munich, Munich, Germany

[4]Department of Medicine III, Ludwig Maximilian University Hospital Munich, Munich, Germany

[5]Bavarian Cancer Research Center, Munich, Germany

Corresponding Author:
Theres Fey, PhD
Comprehensive Cancer Center, Ludwig Maximilian University Hospital Munich, Pettenkoferstrße 8a, Munich, Germany

Abstract

Background: Health care professionals and patients frequently use the internet to access medical information. However, health-related mobile apps have yet to become fully integrated into routine clinical practice. Although the demand for eHealth apps has grown only modestly, studies suggest that such tools hold significant potential to enhance medical care and improve the quality of life, particularly for patients with cancer. To successfully implement these technologies in everyday practice, it is essential to understand the specific needs and preferences of the target population.

Objective: The aim of this study was to assess internet and eHealth use among patients with cancer receiving chemotherapy at an outpatient ward and to evaluate how the COVID-19 pandemic influenced internet and eHealth app use.

Methods: Between May 2021 and September 2021, a total of 303 patients receiving outpatient care at the hemato-oncology and gynecology departments of a Comprehensive Cancer Center were surveyed using a 21-item paper-based questionnaire, adapted from a validated information and communication technology–use survey. The questionnaire provided patient-reported information related to general internet abilities and use rates including health-related apps and changes during the COVID-19 pandemic. Data analysis was conducted using descriptive statistics, chi-square tests, Mann-Whitney *U* tests, and Spearman correlations.

Results: In total, 98.7% (299/303) of participants reported regular internet use, 72.6% (217/299) reported using the internet to search for health-related information, and 79.1% (235/297) expressed readiness to communicate digitally with health care providers. Decreasing age and higher internet literacy correlated with a more frequent use of eHealth apps (P<.001). A total of 24.7% (68/275) reported increased internet use during the pandemic.

Conclusions: The majority of patients were regular internet users and expressed an openness to eHealth apps. Factors such as internet literacy and average age are important to consider when implementing new eHealth apps in a clinical setting. Despite the positive influence of the pandemic on internet use, there remains a gap between self-reported readiness and real use of eHealth apps.

(JMIR Hum Factors 2025;12:e72614) doi:10.2196/72614

KEYWORDS

eHealth; mobile health; health-related internet use; health literacy; patients with cancer; COVID-19; oncology

Introduction

The internet and its wide scope of application possibilities is indispensable in our everyday life. In Germany, approximately 92% of the population uses the internet daily, and the number of nonusers decreases steadily every year [1,2]. While internet use is more common among the younger age groups, data show that already over 90% of people younger than 65 years of age use the internet daily [2]. Internet use among the older generation is concurrently steadily increasing. According to recent data, 27.7% of the internet users are ≥60 years of age [3,4].

At the same time, the importance of digital technologies in health care is rising. This trend is captured under the term

XSL•FO
RenderX

"eHealth," defined as the cost-effective and secure use of information and communication technologies in support of health and health-related fields [5,6]. In the medical context, the internet is used particularly frequently by medical professionals and patients to search for information [7-9]. However, the adoption of more advanced digital health tools—particularly mobile health apps—remains limited. These apps extend beyond mere information retrieval and offer functionalities such as patient-provider communication, symptom monitoring, therapy management, and the collection of patient-reported outcomes [10]. Practical examples of these apps include the Corona-Tracing-App [11], the digital COVID-19 European Union–vaccination certificate [12], or cancer-specific solutions such as CANKADO [13-15]. Despite their potential, such tools have not yet been widely integrated into routine clinical care [16]. Nevertheless, van Eenbergen et al [17] documented that the demand for eHealth apps among patients has increased to a very small extent in the years between 2005 and 2017 [9,17].

Among patients with cancer, internet use is often driven by the desire to understand their disease and treatment options and to connect with other patients or peer support groups [8,18]. While approximately three-quarters of patients with cancer turn to the internet for health-related information, only about half of these users are considered autonomous—able to navigate the internet independently without external assistance [18]. Despite this widespread use for informational purposes, only a small fraction of patients actively engage with more advanced eHealth apps beyond basic web-based searches [7,18-21].

Nevertheless, emerging evidence indicates that the use of eHealth apps has the potential to improve medical care and increase quality of life among patients with cancer. Some studies even demonstrate a link between eHealth app use and an improvement in overall and progression-free survival [22,23]. However, before introducing and using such tools in clinical practice, it is important to understand the target group. Therefore, the aim of this cross-sectional single-point cohort study was to determine if patients with cancer receiving outpatient treatment within our Comprehensive Care Center have access to, and are open to, the use of digital health apps.

Methods

Recruitment

This single-point cohort study was conducted at the outpatient hemato-oncology and gynecology departments of the Comprehensive Cancer Center at LMU Klinikum (Ludwig-Maximilians-University Hospital) in Munich, Germany. LMU Klinikum is one of the largest university hospitals in Europe, serving a diverse patient population from both urban and rural regions across Bavaria.

Between May 1, 2021, and September 30, 2021, patients receiving outpatient care in the aforementioned departments were invited to participate. Study personnel screened patients daily to assess eligibility. All patients aged 18 years or older were eligible for inclusion. Individuals with significant language barriers or functional impairments, as assessed using the Eastern Cooperative Oncology Group performance status, were excluded.

Eligible patients were approached in person during their scheduled visits and received verbal information about the purpose of the study. The survey was paper-based, completed on-site, and required approximately 5 to 10 minutes to complete. Assistance was offered if needed.

Participation was entirely voluntary. Due to the anonymous nature of the data collection, no personal or demographic information was recorded for nonparticipants. As a result, no formal comparison between participants and nonparticipants could be conducted. However, we acknowledge the potential for selection bias—particularly a tendency toward participation among individuals with higher digital literacy or a greater interest in digital health. This may limit the generalizability of the findings to populations with lower levels of internet access or digital affinity.

Questionnaires

The 21-item questionnaire was a modified version of the extended "use of information and communication technologies questionnaire (ICT) developed and validated by Seifert et al [24] for people aged 65 years and older in Switzerland in 2015" (Multimedia Appendix 1).

The first 3 of the 21 items collected basic demographic data including gender, year of birth, and diagnosis. The second part, the core of the questionnaire, consisted of 9 items relating to aspects of the patient's internet use. Analogous to Seifert et al [24], we collected data regarding which internet-enabled devices they own in the household (ie, computer and mobile), which one of these devices they most frequently use, and how often they use the internet. Patients who did not use the internet were asked for reasons. Furthermore, we gathered information regarding the age, the operating system, and the software of the mobile device. The remaining items referred to the patient's ability to use the internet and install apps.

The third part of the questionnaire consisted of 6 items related specifically to eHealth apps. We asked the patients to specify which eHealth apps they already use and which type of eHealth apps they would be willing to try to use.

The final 3 items focused on internet use during the COVID-19 pandemic. The first 2 items referred to the use of the "Corona-Warn-App," which was developed and released by the German government in order to simplify and expand contact tracing [25]. The last item used self-reported data to elicit information regarding how the pandemic has changed the frequency of internet use and if they have begun using eHealth apps during the pandemic. Patients were defined as "completers" when 90% of the items on the questionnaire were filled out (ie,≤2 items of 21 items were missing). All data are therefore presented based on completers only.

In future studies, the inclusion of open-ended questions—such as patient suggestions for desired features or concerns about digital health tools—may provide valuable qualitative insights and complement the quantitative findings.

XSL•FO
RenderX

Statistical Analysis

All data analysis was conducted with SPSS (version 26; IBM Corp) for Windows. Demographics and recruitment data were analyzed using descriptive statistics. Descriptive statistics were also used for analyzing internet user patterns of patients. The 2-tailed *t* test and Mann-Whitney *U* test were used to analyze group differences for numeric variables (ie, number of used apps). The chi-square test of independence and contingency analyses were used to analyze categorical variables. The Spearman correlation analysis was chosen to explore associations between ordinal variables (eg, perceived digital confidence and willingness to use eHealth) and continuous variables (eg, age), as many of the relevant variables were nonnormally distributed or measured on an ordinal scale. After analyzing the complete dataset, the data were then reanalyzed for differences between the 2 cohorts of patient groups from the 2 different departments (hemato-oncology and gynecology) using Mann-Whitney *U* tests and chi-square tests of independence. All tests were 2-sided, and the significance level was set to .05.

Ethical Considerations

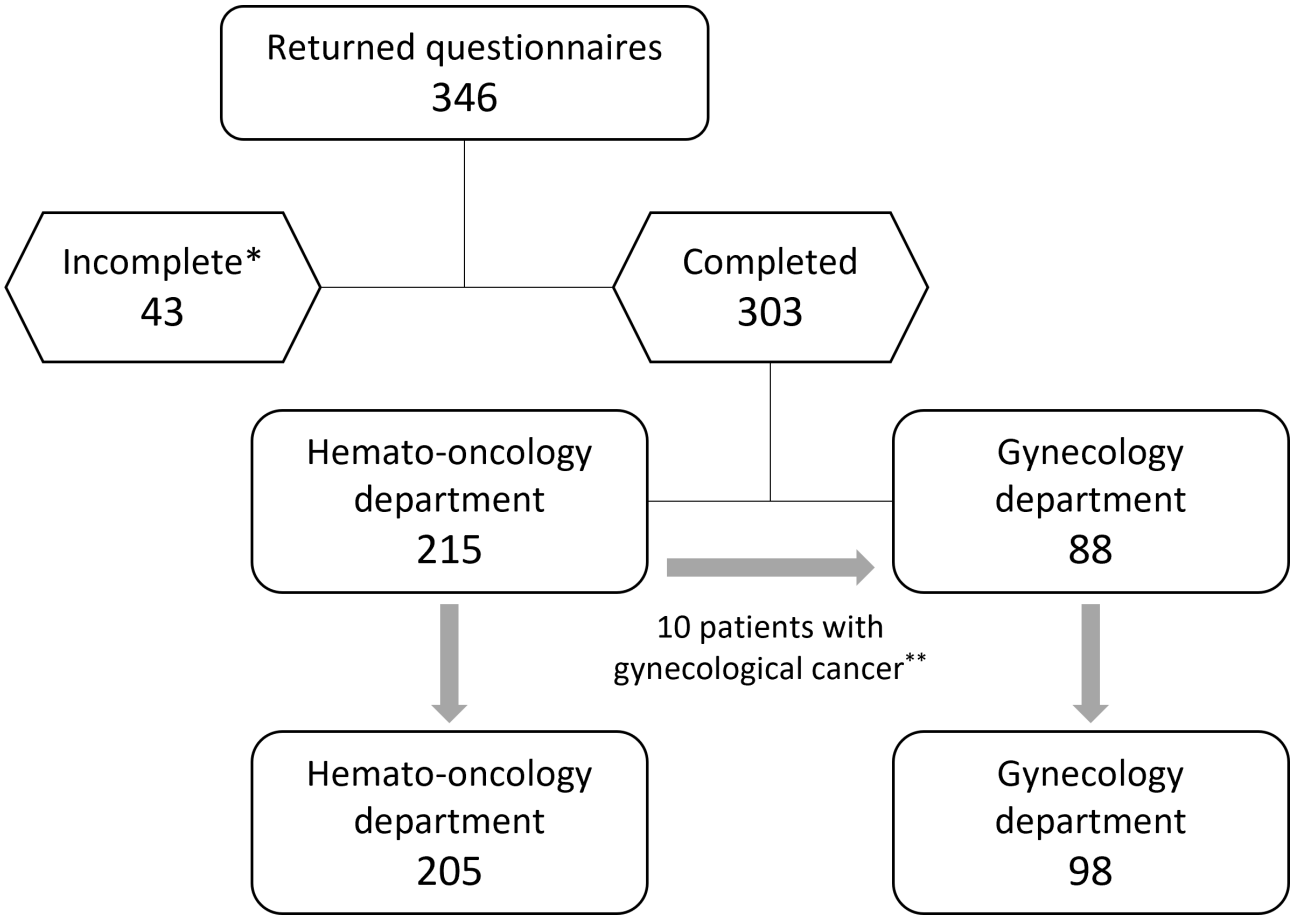
The Medical Ethics Committee of the University of Munich determined that no formal ethics approval was required for this study, as the data were collected entirely anonymously and no personal identifiers were retained (project: KB 20/015). This decision is in accordance with institutional policies for research involving fully anonymized and nonreversibly deidentified data. Participation in the study was entirely voluntary. Patients were informed verbally about the purpose and content of the questionnaire during their outpatient visit, and the completion of the anonymous questionnaire was considered as implied informed consent. All data were collected and processed in a way that ensured full anonymity; no personally identifying information was recorded, and responses were nonreversibly anonymized prior to analysis. No compensation was provided to participants for their participation in the study.

Results

Overview

Of 403 eligible patients approached, 346 returned the questionnaire, yielding a response rate of approximately 85.9%. In total, 303 (87.6%) of them were categorized as complete. A total of 97.3% (285/293) were oncology patients. As a few patients with multiple sclerosis were receiving chemotherapy in the hemato-oncology department, our collective also included 2.7% (8/293) of patients without cancer. In total, 10 patients did not specify their diagnosis, and due to the anonymous nature of the questionnaire, it was not possible to obtain this information. Figure 1 provides an overview of the recruitment and participation process.

Figure 1. Study implementation. *Questionnaires with 2 or fewer missing variables were defined as “complete.” **For the purpose of this analysis, 10 patients with gynecological cancer who were receiving care at the hemato-oncology were moved into the group derived from the gynecology department.



XSL•FO
RenderX

Demographics

The mean age of the respondents was 59 (SD 14.064) years. In total, 55.8% (169/303) were women, and 44.2% (134/303) were men. Table 1 shows the distribution of the participants' primary diagnoses. The most common primary diagnosis was a gynecological tumor, followed by gastrointestinal tumors (defined as a tumor on an organ located in abdominal cavity, with the exception of kidney).

Table . Participants' gender and diagnoses.

	Values, n (%)
Gender (n=303)	
Women	169 (55.8)
Men	134 (44.2)
Diagnosis (n=294)	
Gastrointestinal tumor[a]	88 (29.9)
Gynecological tumor[b]	96 (32.7)
Lymphoma	25 (8.5)
Neuroendocrine tumor[c] (undefined)	13 (4.4)
Sarcoma	9 (3.1)
Urological tumor[d]	17 (5.8)
Bronchial carcinoma	6 (2.0)
Leukemia	5 (1.7)
Oropharyngeal tumor	4 (1.4)
Others	31 (10.5)

[a]All tumors of the following organs were designated as gastrointestinal tumors: liver, gall bladder, bile ducts, stomach, intestine (small and large), pancreas, esophagus, and any neuroendocrine tumors located in these mentioned organs.

[b]A total of 60.4% (58/96) of these were mamma carcinoma; all participants with other gynecology-related cancers were counted to "others."

[c]Other neuroendocrine tumor not located in abdominal cavity.

[d]Urological tumors we defined kidney, ureter, bladder, urethra, and prostate. No participants had penile and testicle cancer.

Access to and Use of the Internet

Only 3 patients did not have an internet-enabled device in their household, while 88.4% (268/303) of the patients reported having 2 or more internet-enabled devices. The most commonly used device was the mobile phone (199/288, 69.1%). In contrast, just over a third of the households had and used a stationary computer (Table 2).

Table . Ownership of web-enabled devices (n=303).

Web-enabled device available in the household	Values, n (%)
Smartphone	279 (92.1)
Notebook	209 (69)
Tablet	169 (55.8)
PC	110 (36.3)

Internet Use Patterns and Digital Confidence

Most of the patients reported using the internet several times a day (244/299, 81.6%), and 91.6% (274/299) reported using the internet at least once a day (Table 3). In total, 17 (5.6%) patients (average age 73, SD 9.5 years) reported technical difficulties and no perceived benefit as their reasons for hardly using the internet. Additionally, 4 patients declared themselves as noninternet users and were therefore excluded from the complete final dataset. In total, 299 of the 303 questionnaires were included in the final results for this part of the evaluation.

The frequency of internet use was age dependent and increased with decreasing age (ρ=0.211; P<.001). In total, 40% (119/298) of the internet users always felt confident navigating the internet. A total of 18.5% (55/298) reported feeling confident in their ability to use the internet most of the time, and 33.9% (101/298) reported feeling confident using simple apps. Only 7.7% (23/298) reported that they did not feel at all confident using the internet. In total, 70% (208/297) reported that they have installed an app, and 9.4% (28/297) estimated their ability for this purpose as sufficient. A total of 17.5% (52/297) required help from others when installing apps; yet, only 3% (9/297)

XSL•FO
RenderX

were of the opinion that they do not have the skills required to install an app.

Among the internet users, the majority (254/293, 86.7%) could receive updates and install them and have the current software version installed (shown in Table 4). These data are also reflected by the age of the devices owned by most patients. More than half of the devices (201/299, 67.2%) were at most 3 years old.

Table . Internet use.

Frequency of internet use	Values (n=299), n (%)	Age (years), mean (SD)
Minimum 1 time per day	274 (91.6)	57.77 (14.082)
Minimum 1 time per week	18 (6)	64.39 (11.387)
Less often	7 (2.3)	65.29 (10.242)

Table . Characteristics of most used devices.

	Values, n (%)
Age of device (years) (n=299)	
<1	49 (16.4)
1 - 3	152 (50.8)
3 - 5	71 (23.7)
>5	27 (9)
Operating system[a] (n=292)	
Android	110 (37.7)
iOS	98 (33.6)
Windows	89 (30.5)
Other	19 (6.5)
Current software version (n=293)	
Updates available	254 (86.7)
Updates available but not installed	24 (8.2)
No updates available	15 (5.1)

[a]The questionnaire items targeted the primary internet-enabled device used by each participant. Nonetheless, some participants gave multiple answers.

Use of the Internet and eHealth Apps

In total, 72.6% (217/299) of the patients reported using the internet for medical-related queries. These patients most frequently searched for information regarding their medical condition (170/299, 56.9%), followed by communication with the health insurance (112/299, 37.5%). Women tended to use the internet more frequently compared to men to communicate with other survivors of cancer (ϕ=0.184; *P*=.001). More than a third of the internet users already used 2 or more of the eHealth apps (Table 5). Younger participants tended to use eHealth apps more often than older ones (ρ=0.295; *P*<.001). A correlation was also found between the number of used apps and the increasing perceived competence in navigating the internet (ρ=0.320; *P*<.001) and navigating apps (ρ=0.330; *P*<.001). Perceived confidence navigating the internet and apps was significantly associated with a principal readiness to use more eHealth offers in the future (ρ=0.373 and 0.490, respectively; *P*<.001). Table 5 illustrates that patients most frequently reported readiness to use web-based platforms for communication with health care professionals and their health insurance. The least popular apps among this patient group were eHealth apps for the purpose of prevention. In total, 87.3% (261/299) of the internet users did at least specify that they are prepared to use 1 or more of the listed eHealth apps. However, increasing age correlated negatively with the number of eHealth apps that the participants were willing to use in the future (ρ=0.390; *P*<.001). In fact, significantly younger patients reported readiness to digitally manage health data for the following purposes: communicating with health care professionals, health insurance companies, or platforms certified by the government (*P*≤.001). Similar gender-related patterns could be detected. Between both collectives, our data revealed that men rather than women prefer to manage their health data digitally (ϕ=0.141; *P*=.02).

XSL•FO
RenderX

Table . Use of eHealth.

	Values, n (%)
Patients reported using the following eHealth apps (n=299)	
Searching for information	170 (56.9)
Communication with health care providers	40 (13.4)
Communication with health insurance	112 (37.5)
Communication with other survivors[a]	23 (7.7)
Keeping a symptom diary	22 (7.4)
Tracking vital parameters	27 (9)
Therapy apps[b]	7 (2.3)
Readiness to use the following eHealth apps (n=297)	
Organization of health data	200 (67.3)
Communication with health care providers	235 (79.1)
Communication with health insurance[c]	222 (74.5)
Use of a medical app approved by the German government[d]	190 (67.4)
eHealth apps for the purpose of (n=299)	
Health care[e]	179 (59.9)
Administrative processes	193 (64.5)
Prevention	110 (36.8)
Research	161 (53.8)
Therapy programs	137 (45.8)

[a]Definition of cancer survivor: “One who remains alive and continues to function during and after overcoming a serious hardship or life-threatening disease. In cancer, a person is considered to be a survivor from the time of diagnosis until the end of life” [26].

[b]For example, Diabetes- or Tinnitus-Trainer.

[c]n=298.

[d]n=282.

[e]For example, a videoconference with a clinician.

The type of operating system also seemed to affect patients’ affinity toward eHealth apps. For the following analysis, the 3 most common operating systems were considered: Android, iOS, and Windows. Patients with iOS on their device more often reported the use of therapy apps (ϕ=0.125; *P*=.03)—including examples such as Diabetes-Trainer and Tinnitus-Trainer [27,28]—and, in comparison with Android users, were also tended to use the internet more often to search for medical information (ϕ=0.148; *P*=.04). Here, no association could be shown with age or gender.

Digital Medicine and COVID-19

In total, 36% (107/298) of the patients had installed the German government–issued “Corona-Warn-App” (Corona-Tracing-App) on their mobile phones at the point of participation in this investigation. Patients who had installed the app were on average 55 (SD 14.327) years of age and were thereby significantly younger than those who had not yet installed the app (mean 60, SD 13.388 years; 2-tailed *t* test; *P*=.001). Among the 170 patients who did not yet have the app installed, 51.1% (87/170) reported that they did not use the “Corona-Warn-App” because they did not see a benefit. However, 27.6% (47/170) of patients reported that they did not use the “Corona-Warn-App” due to technical problems (incompatible operating system and problems with the installation), and 15.3% (26/170) of patients had not heard about the app. In total, 10% (17/170) of patients did not use the “Corona-Warn-App” because they did not want to share their infection status through a digital means. An additional 2.9% (5/170) of the patients did not want to know about their own infection status. A total of 24.7% (68/275) of participants reported that their internet use changed due to the pandemic. In fact, 5.8% (16/275) began to use eHealth apps during the pandemic. In total, 62.5% (10/16) of these patients had installed the government-issued “Corona-Warn-App” as the first eHealth app on their mobile phone. While decreasing age correlated with increased eHealth app use during the pandemic (ρ=0.367; *P*<.001), women tended to be more likely to increase their use during the pandemic, while men did not report any pandemic-related changes (Cramer ϕ=0.241; *P*=.001).

Comparison Between the Collective of Gynecology and Hemato-Oncology

Patients from the gynecology and hemato-oncology departments were also analyzed separately and compared. Patients from the gynecology department were significantly younger than those of the hemato-oncology (56 years compared to 60 years; *P*=.01).

XSL•FO
RenderX

When comparing the 2 collectives, patients from the gynecology department reported feeling more competent using the internet (P=.02), used the internet more frequently (P=.04), and also tended to use more eHealth apps (P=.004). Gynecology patients also reported using web-based forums more often compared to other patients (ϕ=0.183; P=.002). Furthermore, they tended to measure and report vital parameters more frequently within the apps (ϕ=0.139; P=.02). Additionally, patients from the gynecology department communicated with their health insurance through digital apps more frequently (ϕ=0.136; P=.02) and kept digital symptom diaries more often when compared to the patients from the hemato-oncology department (ϕ =0.140; P=.01).

Discussion

Principal Findings

The fact that the overwhelming majority of patients own and use an internet-enabled device provides the foundation for widespread use of eHealth apps. In fact, the majority of the patients reported feeling confident navigating the internet and apps most of the time. Among those who reported lower levels of internet literacy, almost all could receive help when required. It can thus be ascertained that the fundamental conditions, the technical resources, and the know-how for the implementation of digital health apps are all given. In fact, in comparison with other studies, access to and use of the internet is rather high for our collective [20,21,29]. This allows the assumption that implementing eHealth may be even simpler among our patients when compared to other patient settings.

Our patients reported using the internet most frequently to search for and obtain medical information. While these results are consistent with existing literature [18,19], our results do not show quite a high tendency to use the internet to seek information about their disease as previous literature reports. This could be due to the fact that patients from both the gynecology and hemato-oncology departments are routinely made aware of the problems of misinformation through the internet and explicitly instructed not to seek information about their disease through a Google search. Instead, they are provided with evidence-based websites they can refer to when necessary. The next most common reasons cited for internet use among our collective were as follows: communication with health insurance companies (112/299), communication with physicians (40/299), and tracking of vital parameters (27/299). The number of patients who communicate digitally with physicians among our collective is 6.2% higher than that which Makowsky et al [30] report among their collective. However, Makowsky et al [30] investigated internet use among healthy patients, and this may explain the discrepancy. Furthermore, 1 in 13 patients reported using web-based forums to interact with other survivors. These findings align with prior studies, suggesting that patients with cancer seek digital-based peer communication and emotional support, which may serve as a coping mechanism and source of empowerment [8,31,32]. This postulation may also be supported by the fact that women tended to be more likely to participate in such forums than men.

In congruence with previous results [9,21], age influences use as well as openness to eHealth. As expected, the younger the patient, the more frequent the internet use, and the more confident the patients are when navigating the internet. Generally, younger patients are more open to digital health apps. Nevertheless, studies show that the older generation is also becoming increasingly familiar with the internet [33], indicating that the target group for the use of eHealth will continue to grow. Our data reflect this trend by demonstrating that almost all of our patients are regular internet users even though the average age was 59 years.

Whereas a younger age is consistently associated with increased internet use and competency, data regarding gender differences are not as consistent. Our results show that women are more likely than men to use internet forums to communicate with other survivors, but men report a higher readiness for managing health data digitally. These results are consistent with Harrison et al [34], who also demonstrated that, on average, women seek social support more than men. Harrison et al [34] explained this difference in reference to the classical gender stereotype that women are more communicative than men. This may also explain why women are more likely to search the web for support or more ready to participate in support groups [31,32]. However, our data did not confirm gender as an overall predictor of eHealth app use. To explore this further and rule out gender bias, future studies should apply multivariate analyses that control for age and internet literacy.

Among the 3 most frequently used operating systems (Android, iOS, and Windows), our results showed that iOS users tend to be more open-minded toward eHealth apps in comparison to patients who use other operating systems. In fact, iOS users were significantly younger than patients who used other operating systems. However, we found no relationship between age or the current operating system with regard to the use of therapy apps and using the internet to seek information. This may suggest an existing relationship between the operating system and the readiness for using digital medicine. For example, Ubhi et al [35] demonstrated that iOS users more often installed an app to help them stop smoking.

In line with existing literature, our data show that internet literacy is a strong predictor of eHealth readiness, often moderated by sociodemographic factors such as age, gender, and education [26-28]. Therefore, we recommend that before the implementation of eHealth apps, it is important to be familiar with the potential patient's internet literacy. However, if time is lacking, our results, as well as results from previous studies, indicate that studies with patients younger than the age of 65 years may be more successful [2,9,21].

Both internet use in general and the use of eHealth apps increased during the pandemic. This could be due to the fact that when faced with social distancing restrictions and fear of infection, the domain of eHealth and digital health experienced an upswing [36]. Algumzi [37] confirmed this postulation, reporting that during the pandemic, the use of eHealth apps such as telemedicine and digital health increased in acceptance and use. These authors suggest that the reason could be due to fear of infection. This is also reflected in the fact that during the

XSL•FO
RenderX

pandemic, many health care facilities, including our clinical facilities, reduced in-person consultations and switched to telehealth or digital options [38,39]

Although studies with similar timelines report that between 46% and 66% of the population use the government-issued "Corona-Warn-App," only 36% (107/298) of our collective installed the Corona-Warn-App. The main reason cited for not using the "Corona-Warn-App" among all the studies was that the patients did not perceive any benefit from the app. Whereas previous studies cite privacy concerns and the perception of being monitored by the government, our patients named these reasons only sporadically. Nevertheless, our patients were more likely to cite technical problems (47/170, 27.6%) as an obstacle compared to 3.5% reported by Horstmann et al [40]. Our results did reflect a previously reported relationship between decreasing age and use of the government-issued Corona-Warn-App [40,41].

Our data revealed an existing gap between reported readiness to use eHealth apps and the reality of current use. While 79.1% (235/297) of our population reported readiness to use digital apps to communicate with health care professionals, only 13.4% (40/299) reported that they currently use a digital app to communicate with health care professionals. Similarly, while 74.5% (222/298) of the patients reported readiness to use eHealth apps to communicate with health insurance companies, only half (112/222) were currently using such methods, although they are widely available.

Comparable studies reflect these results [19]. In fact, consistent evidence shows that although health care professionals see great potential for eHealth apps to augment care, the step to real use is still hindered by various practical factors such as data protection, time restraints, and resistance to change among health care professionals [42]. Beyond these structural and technical barriers, additional challenges must be considered—particularly concerns about data privacy, trust in digital tools, and the influence of physician encouragement, all of which play a critical role in patients' willingness to engage with eHealth solutions [43].

To bridge this gap, several practical strategies could be implemented: physician-mediated introductions to digital tools during consultations may foster trust and promote engagement; intuitive, low-threshold app designs can reduce usability barriers [44,45]; and tailored support—such as one-on-one training or digital health navigators as introduced in a national study in Germany [46]—could help patients build both competence and confidence. System-level integration of eHealth into routine care processes may further normalize use and promote sustainable adoption.

As our participant pool was limited to patients receiving outpatient therapy in hemato-oncology and gynecology, further studies are needed to explore whether these observations hold true across other clinical settings.

Practice Implications

The gap between readiness and actual use of eHealth apps among oncology patients is decreasing. A slow and careful introduction of eHealth apps into the clinical routine and understanding of the population increases the likelihood of success. In general, this population seems open to eHealth tools; however, dedicated resources for support and assistance may promote growth in real-time use.

Limitations

This study has several limitations. First, convenience sampling was used to recruit patients, which may have led to participation only by those patients who were interested and open to the topic of digital health. To rule out skewed results, patients who did not participate were screened to rule out a lack of knowledge or use of the internet. Furthermore, patients who were not interested or did not use the internet were not excluded from completing the questionnaire. However, a small number of patients (n=4) reported that they did not use the internet. The significantly younger age of the gynecology patients [47,48] must be considered as a possible confounding factor when analyzing our results. However, although there was a mean age difference between the 2 groups, we did not find a relationship with age or gender difference in the willingness to use digital apps to keep symptom diaries in the entire collective. Therefore, it can be postulated that we accounted for this confounding factor adequately. Second, due to the cross-sectional design of this study, no causal inferences can be made regarding the relationship between age, digital literacy, and actual or intended eHealth use. Future longitudinal research is needed to better understand the temporal dynamics and potential causality of these associations. Finally, more demographic data—such as education level, socioeconomic status, or digital access barriers—could have allowed us to account for the influence of education level on patients' responses, and this factor could be considered a limitation when interpreting our results.

Conclusions

Our study demonstrates the existence of the necessary basic conditions for implementing the use of eHealth apps among our collective. In fact, the majority of our patients already use the internet regularly and reported a high readiness to use eHealth. Nevertheless, as only a small percentage of the population currently uses a digital app, although they have access to and are aware of their existence, this study identified a gap between reported readiness to use eHealth apps and the reality of current use. More comprehensive research is necessary in order to identify ways to close this gap. Nevertheless, reported acceptance of different apps is high in almost all domains of eHealth. Positive predictors for use of the internet and eHealth include declining age, higher internet literacy, current use of the iOS operating system, and the number of currently used eHealth apps. While men are more likely to use eHealth for organizational functions, women tended to search for emotional support. As the pandemic caused a surge in eHealth use and acceptance, this may be an impulse for further use of eHealth apps and create an ideal environment for conducting further studies on this topic.

XSL•FO
RenderX

Acknowledgments

The authors received no financial support for this study's research, authorship, and publication.

Data Availability

The datasets generated and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

NE was primarily responsible for the study design (conceptualization and methodology and manuscript writing and revision). SS and NJ led the data collection (investigation). SS contributed to writing the original draft and project administration. SS conducted formal data analysis along with TF, NE, and DN. TF, NE, and DN provided substantial revisions to the manuscript (reviewing and editing), with DN also contributing expertise in digital health (supervision). RW, ST, JC, and VH supervised the clinical components and contributed to the overall study concept (supervision, conceptualization, and resources). All authors approved the final version of the manuscript.

Conflicts of Interest

NE has received speaker honoraria from Audi BKK, Baxter, Beiersdorf, CSL Behring, Fresenius, Freiraum, Fraunhofer Institut, Janssen-Cilag GmbH, Cognitando GmbH, and GHD. She served on the expert advisory board at Baxter and received consulting fees from Fresenius. None of these activities were related to the content of this paper. RW received honoraria for consultancy, advisory boards, and travel support from Agendia, Amgen, APOGHEVA, Aristo, Astra Zeneca, Bayer, Celgene, Clinsol, Daiichi-Sankyo, Eisai, Esteve, Exact Sciences, Gilead, Hexal/Sandoz, Lilly, MSD, Mundipharma, Mylan, Nanostring, Novartis, Onkowissen, Paxman, Palleos, Pfizer, Pierre Fabre, PINK, Riemser, Roche, Seagen, Sidekick, Stemline, Tesaro Bio, Teva, Veracyte, Viatris, Wiley, Thieme, Digimed, Eickeler, FOMF, Art Tempi, Aurikamed, Clinsol, Pomme Med, iMEDInstitut, medconcept, MCI, MediSeminar, Medicultus, H+O Communications, New Oncology Concepts, SPCC, streamed up, medupdate GmbH, and institutions: SPCC, WSG, AGO, and DKG.

Multimedia Appendix 1
Questionnaire.
[PDF File, 196 KB - humanfactors_v12i1e72614_app1.pdf]

References

1. Use of information technologies; computer and internet use by persons in the time comparison; information and communication technology in households. Destatis-Statistisches Bundesamt. 2020. URL: https://tinyurl.com/mtpaku8k [accessed 2025-07-09]
2. Use of information technologies; average internet use by individuals by age; information and communication technology in households 2020. Destatis-Statistisches Bundesamt. 2020. URL: https://www.destatis.de/EN/Themes/Society-Environment/Income-Consumption-Living-Conditions/Use-Information-Technologies/Tables/use-internet-age-ikt.html [accessed 2025-07-09]
3. Anzahl der internetnutzer nach altersgruppen in Deutschland in den jahren 2003 bis 2020. Statista. 2021. URL: https://de.statista.com/statistik/daten/studie/36151/umfrage/anzahl-der-internetnutzer-in-deutschland-nach-altersgruppen-seit-1997 [accessed 2025-07-09]
4. Huxhold O, Otte K. Internetzugang und internetnutzung in der zweiten lebenshälfte [Report in German]. : Deutsches Zentrum für Altersfragen; 2019 URL: https://www.dza.de/fileadmin/dza/Dokumente/DZA_Aktuell/DZAAktuell_Internetnutzung_fin.pdf [accessed 2025-07-24]
5. Alt R, Zimmermann HD. The digital transformation of healthcare—an interview with Werner Dorfmeister. Electron Mark 2021;31(4):895-899. [doi: 10.1007/s12525-021-00476-1] [Medline: 35599690]
6. The World Health Assembly held its fifty-eighth session from 16 to 25 May this year. East Mediterr Health J 2005 May;11(3):256-257. [Medline: 16602444]
7. van Uden-Kraan CF, Jansen F, Lissenberg-Witte BI, Eerenstein SEJ, Leemans CR, Verdonck-de Leeuw IM. Health-related and cancer-related Internet use by patients treated with total laryngectomy. Support Care Cancer 2020 Jan;28(1):131-140. [doi: 10.1007/s00520-019-04757-6] [Medline: 30993449]
8. Holmes MM. Why people living with and beyond cancer use the internet. Integr Cancer Ther 2019;18:1534735419829830. [doi: 10.1177/1534735419829830] [Medline: 30741026]
9. Tarricone R, Cucciniello M, Armeni P, et al. Mobile health divide between clinicians and patients in cancer care: results from a cross-sectional international survey. JMIR Mhealth Uhealth 2019 Sep 6;7(9):e13584. [doi: 10.2196/13584] [Medline: 31493318]

XSL•FO
RenderX

10. Weldring T, Smith SMS. Patient-reported outcomes (PROs) and patient-reported outcome measures (PROMs). Health Serv Insights 2013;6:61-68. [doi: 10.4137/HSI.S11093] [Medline: 25114561]
11. Ellmann S, Maryschok M, Schöffski O, Emmert M. The German COVID-19 digital contact tracing app: a socioeconomic evaluation. Int J Environ Res Public Health 2022 Nov 2;19(21):14318. [doi: 10.3390/ijerph192114318] [Medline: 36361198]
12. EU Digital COVID Certificate. European Comission. URL: https://commission.europa.eu/strategy-and-policy/coronavirus-response/safe-covid-19-vaccines-europeans/eu-digital-covid-certificate_en [accessed 2025-07-09]
13. Gieseler F, Schinköthe TN, Overkamp F, Weislau, M, Würstlein, R. CANKADO: first open eHealth platform for oncology. Presented at: German Cancer Conference; Feb 24-27, 2016; Berlin.
14. Fey T, Erickson N, Stahler A, et al. Should routine risk reduction procedures for the prevention and control of pandemics become a standard in all oncological outpatient clinics? The prospective COVID-19 cohort study: protect-CoV. Med Oncol 2022 Apr 10;39(6):104. [doi: 10.1007/s12032-022-01700-4] [Medline: 35397689]
15. Erickson N, Schinkoethe T, Eckhardt C, et al. Patient-reported outcome measures obtained via E-Health tools ease the assessment burden and encourage patient participation in cancer care (PaCC Study). Support Care Cancer 2021 Dec;29(12):7715-7724. [doi: 10.1007/s00520-021-06351-1] [Medline: 34159428]
16. Marwaha JS, Landman AB, Brat GA, Dunn T, Gordon WJ. Deploying digital health tools within large, complex health systems: key considerations for adoption and implementation. NPJ Digit Med 2022 Jan 27;5(1):13. [doi: 10.1038/s41746-022-00557-1] [Medline: 35087160]
17. van Eenbergen M, Vromans RD, Boll D, et al. Changes in internet use and wishes of cancer survivors: a comparison between 2005 and 2017. Cancer 2020 Jan 15;126(2):408-415. [doi: 10.1002/cncr.32524] [Medline: 31580497]
18. Borges U, Riese C, Baumann W. Internet use by oncology outpatients: results of a survey in Germany. Gesundheitswesen 2018 Dec;80(12):1088-1094. [doi: 10.1055/s-0043-104213] [Medline: 28505698]
19. Drewes C, Kirkovits T, Schiltz D, et al. EHealth acceptance and new media preferences for therapy assistance among breast cancer patients. JMIR Cancer 2016 Sep 14;2(2):e13. [doi: 10.2196/cancer.5711] [Medline: 28410189]
20. Dee EC, Muralidhar V, Butler SS, et al. General and health-related internet use among cancer survivors in the United States: a 2013-2018 cross-sectional analysis. J Natl Compr Canc Netw 2020 Nov;18(11):1468-1475. [doi: 10.6004/jnccn.2020.7591] [Medline: 33152707]
21. Potdar R, Thomas A, DiMeglio M, et al. Access to internet, smartphone usage, and acceptability of mobile health technology among cancer patients. Support Care Cancer 2020 Nov;28(11):5455-5461. [doi: 10.1007/s00520-020-05393-1] [Medline: 32166381]
22. Penedo FJ, Oswald LB, Kronenfeld JP, Garcia SF, Cella D, Yanez B. The increasing value of eHealth in the delivery of patient-centred cancer care. Lancet Oncol 2020 May;21(5):e240-e251. [doi: 10.1016/S1470-2045(20)30021-8] [Medline: 32359500]
23. Basch E, Deal AM, Dueck AC, et al. Overall survival results of a trial assessing patient-reported outcomes for symptom monitoring during routine cancer treatment. JAMA 2017 Jul 11;318(2):197-198. [doi: 10.1001/jama.2017.7156] [Medline: 28586821]
24. Seifert AS, Rudolf H. Digitale Senioren Nutzung von Informations- und Kommunikationstechnologien (IKT) Durch Menschen Ab 65 Jahren in der Schweiz im Jahr 2015 [Book in German]: Pro Senectute Verlag; 2015. URL: https://www.dza.de/fileadmin/dza/Dokumente/DZA_Aktuell/DZAAktuell_Internetnutzung_fin.pdf [accessed 2025-07-24]
25. Corona-Warn-App: frequently asked questions. Die Bundesregierung. URL: https://www.bundesregierung.de/breg-de/themen/corona-warn-app/corona-warn-app-englisch/corona-warn-app-faq-1758636 [accessed 2025-07-24]
26. NCI Dictionary of Cancer Terms. National Cancer Institute. URL: https://www.cancer.gov/publications/dictionaries/cancer-terms/def/survivor [accessed 2025-07-09]
27. MySugr-Diabetes Management App. URL: https://www.mysugr.com/en [accessed 2025-08-05]
28. The ReSound Relief app. ReSound. URL: https://www.resound.com/en-us/hearing-aids/apps/relief [accessed 2025-07-09]
29. Leader AE, Capparella LM, Waldman LB, et al. Digital literacy at an urban cancer center: implications for technology use and vulnerable patients. JCO Clin Cancer Inform 2021 Aug;5(872-80):872-880. [doi: 10.1200/CCI.21.00039] [Medline: 34428075]
30. Makowsky MJ, Jones CA, Davachi S. Prevalence and predictors of health-related internet and digital device use in a sample of South Asian adults in Edmonton, Alberta, Canada: results from a 2014 community-based survey. JMIR Public Health Surveill 2021 Jan 8;7(1):e20671. [doi: 10.2196/20671] [Medline: 33416506]
31. Goodwin BC, Chambers S, Aitken J, et al. Cancer-related help-seeking in cancer survivors living in regional and remote Australia. Psychooncology 2021 Jul;30(7):1068-1076. [doi: 10.1002/pon.5643] [Medline: 33534193]
32. Krizek C, Roberts C, Ragan R, Ferrara JJ, Lord B. Gender and cancer support group participation. Cancer Pract 1999;7(2):86-92. [doi: 10.1046/j.1523-5394.1999.07206.x] [Medline: 10352066]
33. Adults' media use and attitudes. : Ofcom; 2017 URL: https://www.ofcom.org.uk/siteassets/resources/documents/research-and-data/media-literacy-research/adults/adults-literacy-2017/adults-media-use-attitudes-2017.pdf?v=322525 [accessed 2025-07-24]
34. Harrison J, Maguire P, Pitceathly C. Confiding in crisis: gender differences in pattern of confiding among cancer patients. Soc Sci Med 1995 Nov;41(9):1255-1260. [doi: 10.1016/0277-9536(94)00411-l] [Medline: 8545678]

XSL•FO
RenderX

35. Ubhi HK, Kotz D, Michie S, van Schayck OCP, West R. A comparison of the characteristics of iOS and Android users of a smoking cessation app. Transl Behav Med 2017 Jun;7(2):166-171. [doi: 10.1007/s13142-016-0455-z] [Medline: 28168609]
36. Brody JE. A pandemic benefit: the expansion of telemedicine. The New York Times. 2020. URL: https://www.nytimes.com/2020/05/11/well/live/coronavirus-telemedicine-telehealth.html [accessed 2025-07-09]
37. Algumzi A. Evolving factors influencing consumers' attitudes towards the use of eHealth applications: implications on the future of Neom. Int Health 2022 Mar 2;14(2):152-160. [doi: 10.1093/inthealth/ihab020] [Medline: 34038551]
38. Mehrotra A, Chernew M, David, HH, Cutler DA, Schneider EC. The impact of COVID-19 on outpatient visits in 2020: visits remained stable, despite a late surge in cases 2021. The Commonwealth Fund. URL: https://www.commonwealthfund.org/publications/2021/feb/impact-covid-19-outpatient-visits-2020-visits-stable-despite-late-surge [accessed 2025-07-24]
39. Richards M, Anderson M, Carter P, Ebert BL, Mossialos E. The impact of the COVID-19 pandemic on cancer care. Nat Cancer 2020 Jun;1(6):565-567. [doi: 10.1038/s43018-020-0074-y] [Medline: 35121972]
40. Horstmann KT, Buecker S, Krasko J, Kritzler S, Terwiel S. Who does or does not use the "Corona-Warn-App" and why? Eur J Public Health 2021 Feb 1;31(1):49-51. [doi: 10.1093/eurpub/ckaa239] [Medline: 33340328]
41. Munzert S, Papoutsi M, Nowak H. Ein jahr digitale kontaktpersonennachverfolgung mit der corona-warn-app [Report in German]. : Opus; 2021 URL: https://opus4.kobv.de/opus4-hsog/frontdoor/index/index/docId/3999 [accessed 2025-07-24]
42. Wangler J, Jansky M. The use of health apps in primary care-results from a survey amongst general practitioners in Germany. Wien Med Wochenschr 2021 May;171(7-8):148-156. [doi: 10.1007/s10354-021-00814-0] [Medline: 33570692]
43. Schreiweis B, Pobiruchin M, Strotbaum V, Suleder J, Wiesner M, Bergh B. Barriers and facilitators to the implementation of eHealth services: systematic literature analysis. J Med Internet Res 2019 Nov 22;21(11):e14197. [doi: 10.2196/14197] [Medline: 31755869]
44. Zhou L, Bao J, Setiawan IMA, Saptono A, Parmanto B. The mHealth App Usability Questionnaire (MAUQ): development and validation study. JMIR Mhealth Uhealth 2019 Apr 11;7(4):e11500. [doi: 10.2196/11500] [Medline: 30973342]
45. Van Baelen F, De Regge M, Larivière B, Verleye K, Schelfout S, Eeckloo K. Role of social and app-related factors in behavioral engagement with mHealth for improved well-being among chronically ill patients: scenario-based survey study. JMIR Mhealth Uhealth 2022 Aug 26;10(8):e33772. [doi: 10.2196/33772] [Medline: 36018618]
46. Hinneburg J, Zacher S, Berger-Höger B, et al. Enhancing transsectoral interdisciplinary patient-centered care for patients with rare cancers: protocol for a mixed methods process evaluation. JMIR Res Protoc 2023 Oct 12;12:e49731. [doi: 10.2196/49731] [Medline: 37824180]
47. Cancer causes and prevention; age and cancer risk. National Cancer Institute. 2021. URL: https://www.cancer.gov/about-cancer/causes-prevention/risk/age [accessed 2025-07-24]
48. Gynecologic cancer incidence, United States—2012-2016. Centers for Disease Control and Prevention. 2019. URL: https://stacks.cdc.gov/view/cdc/81916 [accessed 2025-07-24]

Abbreviations

AGO: Arbeitsgemeinschaft Gynäkologische Onkologie
DKG: German Cancer Society
SPCC: Sharing Progress in Cancer Care
WSG: Westdeutsche Studiengruppe

Edited by A Kushniruk; submitted 13.02.25; peer-reviewed by H Maheshwari, L Zhang, MA Zoubi, VS Kishore Kancharla; revised version received 28.05.25; accepted 03.06.25; published 14.08.25.

Please cite as:
Erickson N, Nasseh D, Seynstahl S, Jost N, Wuerstlein R, Casuscelli J, Theurich S, Heinemann V, Fey T
Health-Related Internet Use Among Outpatients Undergoing Cancer Treatment During the COVID-19 Pandemic: Cross-Sectional Survey Study
JMIR Hum Factors 2025;12:e72614
URL: https://humanfactors.jmir.org/2025/1/e72614
doi:10.2196/72614

© Nicole Erickson, Daniel Nasseh, Sarah Seynstahl, Nicole Jost, Rachel Wuerstlein, Jozefina Casuscelli, Sebastian Theurich, Volker Heinemann, Theres Fey. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 14.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information,

XSL•FO
RenderX

a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

User-Centered Delivery of AI-Powered Health Care Technologies in Clinical Settings: Mixed Methods Case Study

Meredith Schreier[1], MA; Randall Brandt[2], PA-C; Hien Brown[1], MHI, RN; T Saensuksopa[1], MEng, MHCI; Christine Silva[3], BS; Laura M Vardoulakis[1], PhD

[1]Google (United States), 1600 Amphitheater Parkway, Mountain View, CA, United States

[2]Mile Bluff Medical Center, Mauston, United States

[3]MEDITECH, Canton, MA, United States

Corresponding Author:
Laura M Vardoulakis, PhD
Google (United States), 1600 Amphitheater Parkway, Mountain View, CA, United States

Abstract

Background: Providers spend a large percentage of their day using electronic health record (EHR) technology and frequently report frustration when EHR tasks are time-consuming and effortful. To solve these challenges, artificial intelligence (AI)–based enhancements to EHR technology are increasingly being deployed. However, AI-based implementations for EHR features often lack user-centered evaluation.

Objective: This study evaluates, using a user-centered approach, the implementation of an AI-powered search and clinical discovery tool within an EHR system.

Methods: We conducted a mixed methods study consisting of interviews, observations, and surveys for 5 months.

Results: High adoption rates for the AI-based features (163/176, 93% users after 3 months) and significant increases across key metrics, including user satisfaction (U=49; P<.001) and perception of time saved (U=49; P<.001), demonstrated that the AI-based features were not only successfully integrated into various clinical workflows but also improved the user experience for clinicians.

Conclusions: Our results underscore the feasibility and effectiveness of using a user-centered approach for the deployment of clinical AI tools. High adoption rates and positive user experiences were driven by our user-centered research program, which emphasized close collaboration with users, rapid incorporation of feedback, and tailored user training. This study program can be used as a starting framework for the design and integration of human-centered research methods for AI tool deployment in clinical settings.

(JMIR Hum Factors 2025;12:e76241) doi:10.2196/76241

KEYWORDS

user-centered research; software design; delivery of healthcare; implementation science; electronic health record

Introduction

Artificial intelligence (AI)–based tools have permeated culture and industry, generating excitement, hesitation, and visions of opportunity [1-3]. This is especially true in health care: AI-based tools have transformative potential, despite known challenges spanning implementation, security, infrastructure, user experience, and product performance [4-7]. While electronic health records (EHRs) have been transformative in moving health care professionals to a more digital and connected way of working, challenges of use can lead to increased time spent using the system, increased feelings of frustration, and general dissatisfaction among users [8-14]. As EHR systems incorporate AI-based features into their products, questions remain about whether these solutions will improve the clinical user experience.

The incorporation of human-centered research methods, including the engagement of end users throughout the development process, is critical to the development of usable products that meet clinical users' needs [15,16] and has a long history within the field of human-computer interaction [17,18]. Previous research has demonstrated the value and feasibility of user-centered research methods during the development of AI-based clinical tools [19-21]. In this paper, we build upon previous research by demonstrating how user-centered evaluations can extend to the evaluation of clinical AI tools during the final stages of implementation.

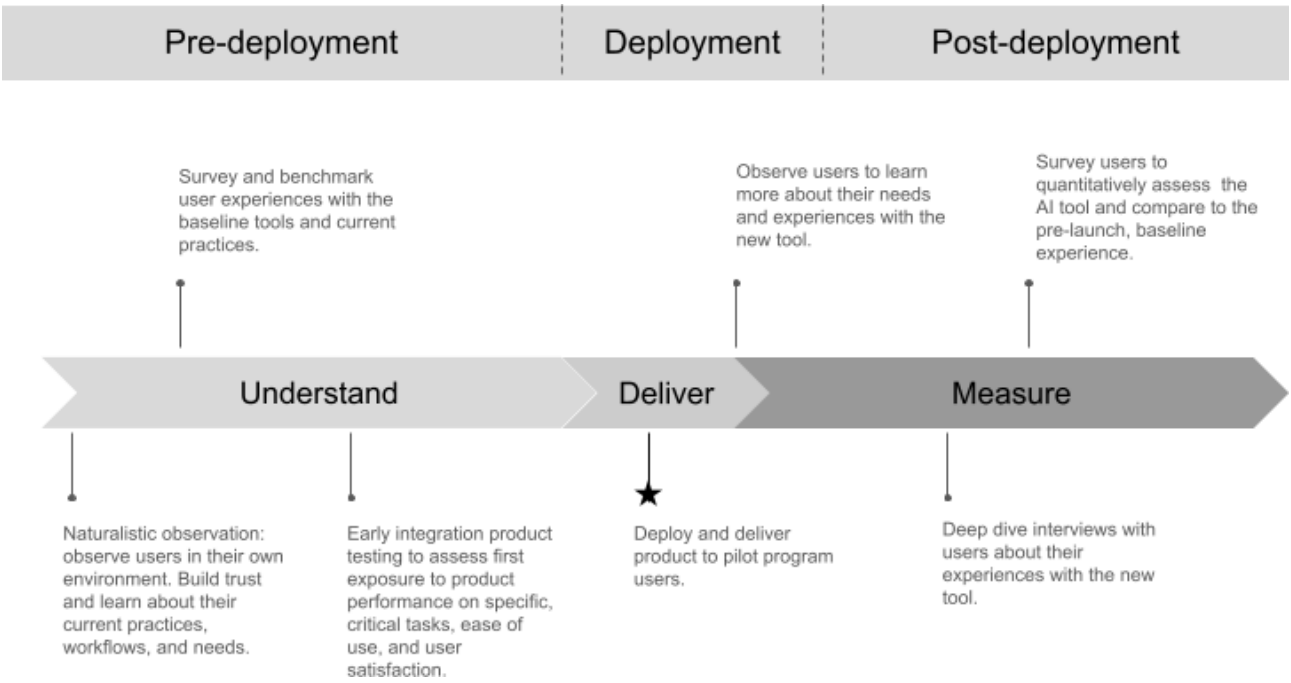
Through this research, we sought to understand if an AI-based search and summarization tool [22], previously used side by side with an EHR, could be successfully integrated directly within the EHR. To test the implementation approach and measure the experiential impact on users at a health care system, we developed a multitouchpoint user-centered study program

XSL•FO
RenderX

that emphasized strong partnerships, close collaboration with users, and robust evaluations focused on newly unlocked product capabilities. The research methods draw from long-established practices in user experience design and research [23-25], as well as newer guidelines for AI-based research [26], which were adapted for use within the clinical environment.

This paper makes two primary contributions. First, we introduce a collection of research activities that form a novel, user-centered framework for implementing and evaluating AI-based clinical tools (Figure 1). This framework emphasizes the importance of conducting both qualitative and quantitative methods to evaluate the user experience and collecting signals from users throughout multiple stages of implementation and deployment. Further, there is flexibility in the specific methods that can be used in a collection such as this, making it an adaptable means of conducting human-centered research.

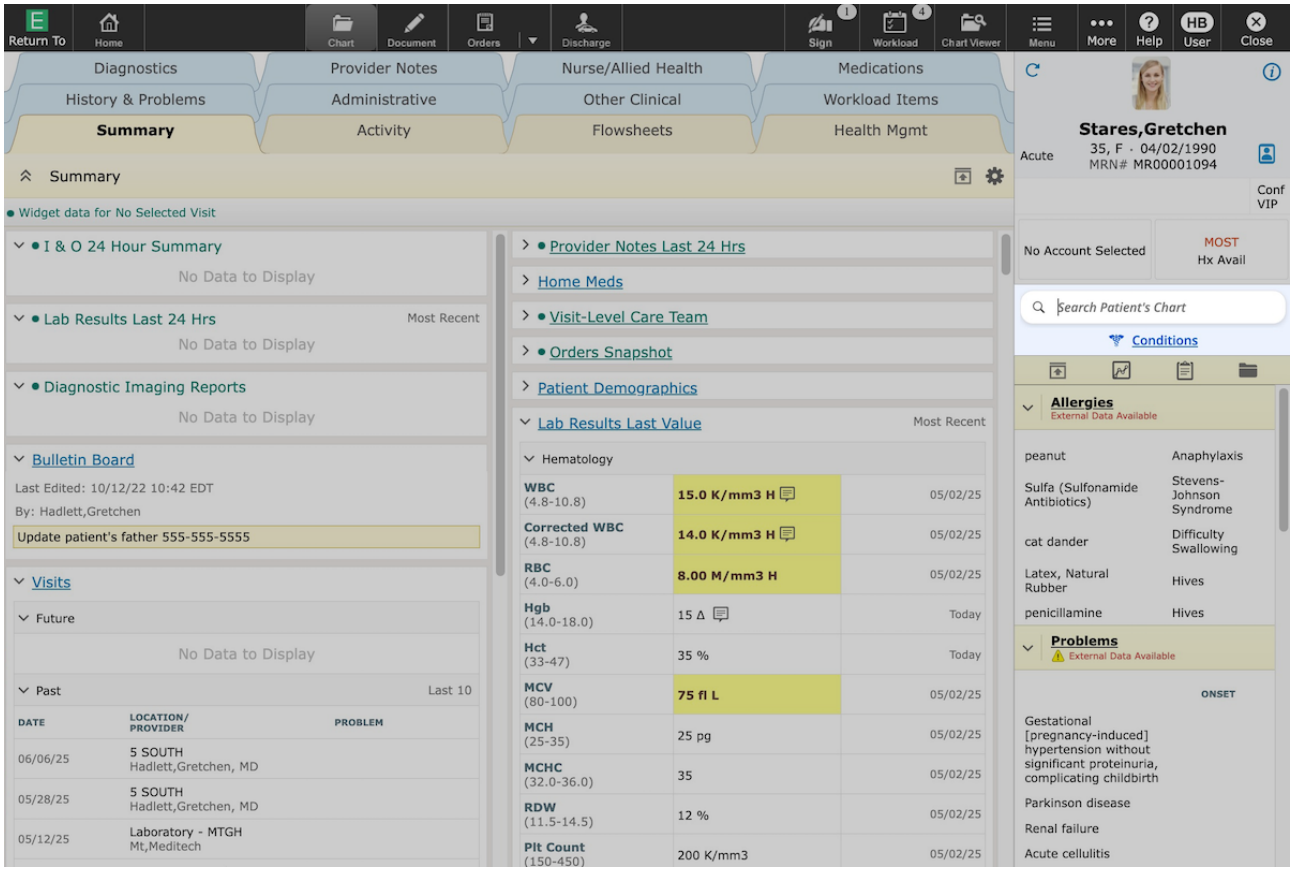
Figure 1. User-centered study program timeline. AI: artificial intelligence.



Second, we apply this framework through a case study of the delivery of Expanse Navigator (Google and MEDITECH; Figures 2 and 3; Multimedia Appendix 1), an AI-powered search and summarization tool accessible within MEDITECH's Expanse EHR. Through this case study, we show both the feasibility of the approach and how it enabled rapid incorporation of insights into the delivery strategy, including factors influencing use, potential risks, and user satisfaction.

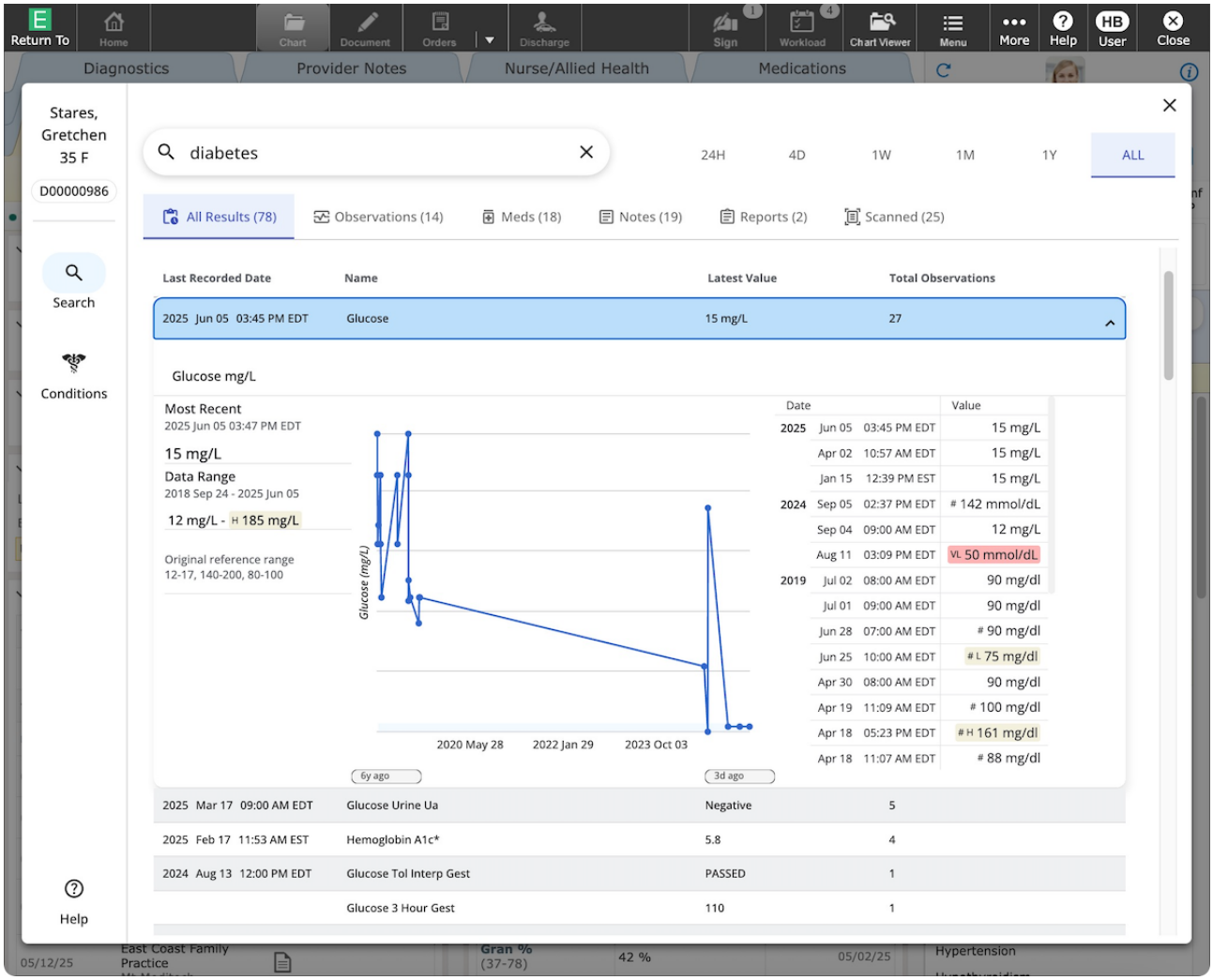
XSL•FO
RenderX

Figure 2. Artificial intelligence (AI)–powered search integration within MEDITECH's Expanse, highlighted. All data shown in this figure are synthetic (ie, realistic but not real) patient data. Participants in the study interacted with real patient data.



XSL•FO
RenderX

Figure 3. Search results experience. All data shown in this figure are synthetic (ie, realistic but not real) patient data. Participants in the study interacted with real patient data.



Methods

Overview

A multistep, mixed methods, user-centered study was developed and run for 5 months. The goal of this research program was to measure the qualitative and quantitative experiential effects of an AI-based search and summarization tool (referred to as Expanse Navigator) within MEDITECH's Expanse and to understand the feasibility of incorporating user-centered methods during the implementation phases of AI-based clinical tools.

Program Site and Participants

Mile Bluff Medical Center is a 40-bed acute-care hospital and outpatient care center in Mauston, Wisconsin, United States. A total of 176 health care professionals were provided access to a clinical pilot of Expanse Navigator in January 2024; this includes 53 medical doctors or advanced practice providers, 61 registered nurses, 48 allied health professionals, and 14 health information management specialists across 24 distinct clinical and operational departments.

All pilot participants were invited to participate in the user-centered research program. A sample of 128 out of the 176 (72%) pilot program members opted into the program.

Materials

Expanse Navigator offers advanced search capabilities and an AI-generated, comprehensive view of a patient's health history (Figures 2 and 3; Multimedia Appendix 1). It works together with and is accessible directly within MEDITECH's Expanse. The tool offers new capabilities, unlocks valuable data (eg, PDFs and faxes), and makes them quickly and easily accessible to users via a modern search experience. The search capabilities handle misspellings and abbreviations, provide suggested search terms, and give results ranked by relevance. The tool is powered by AI rather than keyword matching.

Naturalistic Observation

The first component of the user-centered program was clinical observation. Before the deployment of Expanse Navigator, 21 representative users from the hospital, ambulatory care, outpatient clinics, and the emergency department were observed in their normal clinical environments (outside of patient care). Observations provided an understanding of users' existing tools, workflows, environments, and experiences before the deployment of Expanse Navigator. Further, this effort helped to establish collaboration and build rapport with clinical users.

XSL•FO
RenderX

Survey Program

The second component of the user-centered program focused on experiential measurement through the use of pre- and postdeployment surveys. To develop the survey measures, MEDITECH and Google jointly developed a set of primary metrics: user satisfaction, product helpfulness across specified tasks (Multimedia Appendix 2), and perceived time savings. Survey questions used fully labeled 5-point scales, following industry best practices [27], and are available in the Multimedia Appendix 3. Survey questions were optional; respondents were not required to answer all questions.

Before deploying Expanse Navigator, we measured prelaunch metrics of MEDITECH's Expanse via a baseline survey. Four weeks following deployment, we distributed a postlaunch survey to collect identical metrics on users' experiences with Expanse Navigator.

Surveys were administered to all pilot program participants. There were 94 out of 176 (73%) responses and 73 out of 176 (57%) responses to the baseline survey and postlaunch survey, respectively.

Early Integration Product Testing

The third component of the user-centered program involved early integration of product testing. The purpose of this assessment was to test the integrated product's ability to support key clinical tasks, to collect feedback based on users' first exposure to the product, and to ensure that no issues that could prevent easy and efficient use were present ahead of widespread launch.

During a single 60-minute session, participants viewed 3 minutes of prerecorded onboarding content and had 5 minutes of unstructured exploration with the product. Immediately thereafter, participants were prompted with 12 structured tasks to complete using Expanse Navigator features (Multimedia Appendix 4). Tasks were scored as "completed," "completed with difficulty," or "not completed." The tasks were informed by findings from naturalistic observation sessions and designed to reflect responsibilities fulfilled in the course of daily work. Sessions were conducted until data confidently showed that users could complete key tasks. In our case, this research phase ended after sessions with 3 users showed high levels of success. This methodology facilitated unique feedback related to intuitiveness and ease of use and also highlighted areas of focus for user onboarding and education.

User Interviews

The final component of the user-centered program was qualitative interviews. Approximately 6 weeks after launching Expanse Navigator, we conducted 60-minute interviews with 4 users. We sought to capture qualitative data around the impact of the product, users' experiences, high-priority use cases, and self-reported engagement. Participants spanned distinct roles and specialties and had not previously participated in one-on-one user-centered research sessions.

Data Analysis

Completed survey responses were analyzed using descriptive statistics (eg, percentage of users reporting satisfaction, perceived helpfulness, and time savings). To measure the experiential impact of the AI tool, a within-subjects Mann-Whitney *U* test (1-tailed) was run on each metric, using data from respondents who completed both baseline and postlaunch surveys. Qualitative interviews were analyzed by conducting a thematic analysis of detailed interview notes [28].

Ethical Considerations

This research was reviewed by the Advarra institutional review board and determined to be exempt from IRB oversight. All participants provided informed consent before participating.

Results

Early integration testing demonstrated that users completed tasks with relative ease, having received minimal instructions, and no major obstacles to efficient and easy use were observed, clearing the way for further deployment. Results from this research stage highlighted the need to explicitly onboard users to the new structure and categories by which information was organized.

Postdeployment, our 3 primary metrics for Expanse Navigator, described above in the "Introduction" section and Figures 2 and 3, saw strong ratings from users (Table 1): 62 out of 72 (86%) respondents reported high satisfaction, 45 out of 60 (75%) respondents described high helpfulness, and 63 out of 69 (91%) respondents reported that Expanse Navigator felt faster than the baseline system.

Table . Experiential ratings of the artificial intelligence (AI)–based experience from participants who completed the postlaunch survey: primary metric scores as percentages.

Metric	Postlaunch survey, n (%)
Satisfaction (N=72)	
Users who reported being "somewhat" or "very" satisfied	62 (86)
Helpfulness (N=60)	
Average % of users who rated the product as "very" or "extremely" helpful in accomplishing the 7 specific tasks	45 (75)
Perceived time savings (N=69)	
Users who rated the system as "somewhat" or "much" faster than the previous system	63 (91)

XSL•FO
RenderX

Notably, when compared to baseline results, users rated the AI-based experience more favorably across all metrics (Tables 2-4). Qualitative analyses showed that clinicians had use for the product at all points of care, identified the key value derived, and provided data to inform potential future features.

Table . Within-subjects comparison of ratings among participants who completed both baseline and postlaunch surveys: primary metric scores as percentages.

Metric	Baseline survey, n (%)	Postlaunch survey, n (%)
Satisfaction, N=19		
Users who reported being “somewhat” or “very” satisfied	2 (11)	15 (79)
Helpfulness, N=15		
Users who rated the product as “very” or “extremely” helpful in accomplishing the 7 specific tasks averaged across tasks on average	2 (13)	8 (53)
Perception of time savings, N=18		
Users who rated the system as “somewhat” or “much” faster than the previous system	9 (50)	17 (94)

Table . Within-subjects comparison of user satisfaction and perceived time savings: 1-tailed Mann-Whitney *U* test results. Effect size is given by the rank biserial correlation.

Mann-Whitney *U* test	U	*P* value	Rank-biserial correlation (*r*)	SE of rank-biserial correlation
Satisfaction score	49.00	<.001	−0.729	0.188
Time savings score	74.50	.002	−0.540	0.193

Table . Within-subjects comparison of helpfulness: 1-tailed Mann-Whitney *U* test results. Effect size is given by the rank biserial correlation.

Helpfulness scores per task	U	*P* value	Rank-biserial correlation (*r*)	SE of rank-biserial correlation
Task 1	35.00	<.001	−0.689	0.211
Task 2	68.00	.03	−0.396	0.211
Task 3	46.50	.003	−0.587	0.211
Task 4	32.00	<.001	−0.716	0.211
Task 5	88.00	.15	−0.218	0.211
Task 6	73.00	.08	−0.305	0.215
Helpfulness task #7	51.50	.009	−0.510	0.215

Results from the program at large demonstrated that the AI-based experience was highly adaptable across various clinical and operational settings. In addition, the user-centered approach provided the clinical and AI teams with rapid data on the tool’s effectiveness and identified valuable focus areas for user training. This data-driven approach, paired with positive word-of-mouth and user experiences, fueled rapid delivery and adoption. When measured 3 months post rollout, 163 out of 176 (93%) participants had accessed Expanse Navigator, and over 4000 searches had been performed.

Discussion

Principal Findings

The results of this study not only demonstrate that the AI-powered search and clinical discovery tool led to significant improvements in user satisfaction, helpfulness, and perception of time saved, but also demonstrate the feasibility of evaluating the clinical experience through a mixed method, user-centered approach.

AI-based tools are becoming increasingly common in clinical settings. As these technologies continue to advance, so will user expectations for high-quality product experiences. It will become increasingly important to measure the user experience of these tools with robust approaches that can be flexibly applied within a clinical environment. This research program was intentionally designed to evaluate product experience, refine understanding of user mental models, and optimize our onboarding program for the greatest impact when deploying an AI tool.

Incorporating user-centered research during predeployment stages allowed for the collection of benchmark measures of user satisfaction, helpfulness, and speed. It also informed the design of onboarding experiences and custom training for cohorts with

XSL•FO
RenderX

specialized workflows. Furthermore, the activities during this stage led to a strong partnership with hospital leadership, which, in turn, increased awareness about the upcoming product launch.

The deployment and postdeployment research allowed the team to measure the impact of the AI tool from a clinical user's point of view. The results of our case study demonstrate that the AI-based search and summarization experience, accessible directly within an EHR, provides users with new capabilities and increased levels of satisfaction, helpfulness, and speed.

Both the lift and feasibility of a user-centered approach were demonstrated through strong user participation, survey completion rates, and the overall duration of data collection (60 minutes per one-on-one study session).

The design of this study program follows previous work showing the importance of assessing the user experience during all phases of the product development life cycle: design, development, and deployment [29,30], and addresses calls for an increase in human-centered evaluative research alongside AI clinical deployments [31]. It extends past research that outlines approaches to user-centered research during design and development stages of AI-based clinical tools [19-21], and provides a framework for assessments that can occur during the later stages of clinical implementation.

Limitations

There are several limitations to this research. First, while the research involved a large percentage of users for the AI-based tool within MEDITECH's Expanse, across a wide range of clinical roles and specialities, it did so with users from a single medical center. Though we demonstrate feasibility for the user-centered study program presented in this paper, future research is needed to demonstrate generalizability. Second, while the findings from the case study show positive improvements in the user experience for this AI-based EHR integration, further research would be needed to demonstrate the impact across additional clinical measures.

Conclusions

By developing and executing a mixed methods, user-centered program for AI tool deployment, we enabled trust, collaboration, and frequent communication between clinical end users and AI practitioners, and a robust collection of key product metrics. This data-driven and user-centered approach championed end users' experiences and surfaced those findings pre- and postproduct launch, highlighting specific areas for curated user training and optimizing the opportunity for rapid deployment. Integrating user-centered research as part of the deployment process ensures that AI tools are not only technologically advanced but are also relevant, usable, useful, and safe for use in any setting; notably, all of these benefits are of critical importance for tools used in health care and clinical settings.

Prioritizing a human-centered research program for AI-based tool deployment supports effective product integration and user satisfaction; further, it supports the thoughtful delivery of products and services in high-stakes health care environments where patient safety is the utmost priority.

Acknowledgments

The authors thank Lynsi Garvin and Pete Clardy for their assistance with this study, and Rory Sayres for his helpful comments on the manuscript. This study was funded by Google LLC.

Authors' Contributions

MS, LMV, and HB contributed to the conceptualization and methodology of the study. MS and HB were responsible for the data curation and investigation. LMV was responsible for funding acquisition and study supervision. TS was responsible for the design of the software and visualization. HB and CS led project administration, and RB provided resources. MS prepared the original draft, and CS, HB, LMV, RB, and TS contributed to the review and editing of the manuscript.

Conflicts of Interest

MS, HB, TS, and LMV are employees of Google LLC, and CS is an employee of MEDITECH, representing a potential conflicts of interest. Throughout the study design, data collection, and analysis, efforts were made to reduce potential bias and to report findings based solely on objective study data.

Multimedia Appendix 1
Summarization experience. All data shown in this figure are synthetic (ie, realistic but not real) patient data. Participants in the study interacted with real patient data.
[PNG File, 346 KB - humanfactors_v12i1e76241_app1.png]

Multimedia Appendix 2
Survey program: helpfulness tasks.
[PDF File, 61 KB - humanfactors_v12i1e76241_app2.pdf]

Multimedia Appendix 3

XSL•FO
RenderX

Survey measures. Prelaunch surveys assessed baseline experiences with “Search Patient’s Chart,” and postlaunch surveys assessed experiences with “Search and Summarization.”
[PDF File, 71 KB - humanfactors_v12i1e76241_app3.pdf]

Multimedia Appendix 4
Early integration testing task set.
[PDF File, 62 KB - humanfactors_v12i1e76241_app4.pdf]

References

1. Goldberg CB, Adams L, Blumenthal D, et al. To do no harm - and the most good - with AI in health care. Nat Med 2024 Mar;30(3):623-627. [doi: 10.1038/s41591-024-02853-7] [Medline: 38388841]
2. Rajpurkar P, Chen E, Banerjee O, Topol EJ. AI in health and medicine. Nat Med 2022 Jan;28(1):31-38. [doi: 10.1038/s41591-021-01614-0] [Medline: 35058619]
3. Loh E. ChatGPT and generative AI chatbots: challenges and opportunities for science, medicine and medical leaders. BMJ Lead 2023 May 2;8(1):e000797. [doi: 10.1136/leader-2023-000797] [Medline: 37192124]
4. Seneviratne MG, Shah NH, Chu L. Bridging the implementation gap of machine learning in healthcare. BMJ Innov 2020 Apr;6(2):45-47. [doi: 10.1136/bmjinnov-2019-000359]
5. Khan B, Fatima H, Qureshi A, et al. Drawbacks of artificial intelligence and their potential solutions in the healthcare sector. Biomed Mater Devices 2023 Feb 8;1(2):1-8. [doi: 10.1007/s44174-023-00063-2] [Medline: 36785697]
6. Bizzo BC, Dasegowda G, Bridge C, et al. Addressing the challenges of implementing artificial intelligence tools in clinical practice: principles from experience. J Am Coll Radiol 2023 Mar;20(3):352-360. [doi: 10.1016/j.jacr.2023.01.002] [Medline: 36922109]
7. Mullankandy S, Mukherjee S, Ingole BS. Applications of AI in electronic health records, challenges, and mitigation strategies. Presented at: 2024 International Conference on Computer and Applications (ICCA); Dec 17-19, 2024; Cairo, Egypt p. 1-7. [doi: 10.1109/ICCA62237.2024.10927863]
8. Meigs SL, Solomon M. Electronic health record use a bitter pill for many physicians. Perspect Health Inf Manag 2016;13(Winter):1d. [Medline: 26903782]
9. Wawrzyniak C, Marcilly R, Baclet N, Hansske A, Pelayo S. EHR usage problems: a preliminary study. Stud Health Technol Inform 2019;257:484-488. [Medline: 30741244]
10. Kellogg KM, Fairbanks RJ, Ratwani RM. EHR usability: get it right from the start. Biomed Instrum Technol 2017;51(3):197-199. [doi: 10.2345/0899-8205-51.3.197] [Medline: 28530885]
11. How doctors feel about electronic health records. The Harris Poll. 2018. URL: https://med.stanford.edu/content/dam/sm/ehr/documents/EHR-Poll-Presentation.pdf [accessed 2025-06-10]
12. Walsh SH. The clinician’s perspective on electronic health records and how they can affect patient care. BMJ 2004 May 15;328(7449):1184-1187. [doi: 10.1136/bmj.328.7449.1184] [Medline: 15142929]
13. Kroth PJ, Morioka-Douglas N, Veres S, et al. Association of electronic health record design and use factors with clinician stress and burnout. JAMA Netw Open 2019 Aug 2;2(8):e199609. [doi: 10.1001/jamanetworkopen.2019.9609] [Medline: 31418810]
14. Mishra V, Liebovitz D, Quinn M, Kang L, Yackel T, Hoyt R. Factors that influence clinician experience with electronic health records. Perspect Health Inf Manag 2022;19(1):1f. [Medline: 35440924]
15. Harte R, Glynn L, Rodríguez-Molinero A, et al. A human-centered design methodology to enhance the usability, human factors, and user experience of connected health systems: a three-phase methodology. JMIR Hum Factors 2017 Mar 16;4(1):e8. [doi: 10.2196/humanfactors.5443] [Medline: 28302594]
16. Melles M, Albayrak A, Goossens R. Innovating health care: key characteristics of human-centered design. Int J Qual Health Care 2021 Jan 12;33(Supplement_1):37-44. [doi: 10.1093/intqhc/mzaa127] [Medline: 33068104]
17. Allen CD, Ballman D, Begg V, et al. User involvement in the design process: why, when & how? Presented at: CHI ’93: Proceedings of the INTERACT ’93 and CHI ’93 Conference on Human Factors in Computing System; Apr 24-29, 1993; Amsterdam, NL p. 251-254 URL: https://doi.org/10.1145/169059.169203 [accessed 2025-08-16] [doi: 10.1145/169059.169203]
18. Delgado F, Yang S, Madaio M, Yang Q. The participatory turn in ai design: theoretical foundations and the current state of practice. 2023 Oct 30 Presented at: EAAMO ’23: Proceedings of the 3rd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization; Oct 30 to Nov 1, 2023; Boston, MA p. 1-23 URL: https://dl.acm.org/doi/proceedings/10.1145/3617694 [accessed 2025-08-18] [doi: 10.1145/3617694.3623261]
19. Seneviratne MG, Li RC, Schreier M, et al. User-centred design for machine learning in health care: a case study from care management. BMJ Health Care Inform 2022 Oct;29(1):e100656. [doi: 10.1136/bmjhci-2022-100656] [Medline: 36220304]
20. Auernhammer J. Human-centered AI: the role of human-centered design research in the development of AI. Presented at: DRS2020; Aug 11-14, 2020; Online URL: https://doi.org/10.21606/drs.2020.282 [accessed 2025-08-16] [doi: 10.21606/drs.2020.282]

XSL•FO
RenderX

21. Wang L, Zhang Z, Wang D, et al. Human-centered design and evaluation of AI-empowered clinical decision support systems: a systematic review. Front Comput Sci 2023;5. [doi: 10.3389/fcomp.2023.1187299]
22. MEDITECH and Google Health collaborate for AI-powered Expanse search and summarisation. Medical Information Technology. 2024. URL: https://ehr.meditech.com/news/meditech-and-google-health-collaborate-for-ai-powered-expanse-search-and-summarisation [accessed 2025-06-10]
23. Law ELC, Roto V, Hassenzahl M, Vermeeren A, Kort J. Understanding, scoping and defining user experience. 2009 Apr 4 Presented at: CHI '09: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; Apr 4-9, 2009; Boston, MA p. 719-728 URL: https://dl.acm.org/doi/proceedings/10.1145/1518701 [accessed 2025-07-29] [doi: 10.1145/1518701.1518813]
24. Nielsen J. Usability Engineering: Academic Press; 1993.
25. Portigal S. Interviewing Users: How to Uncover Compelling Insights, 2nd edition: Rosenfeld Media, LLC; 2023.
26. What's the PAIR guidebook? Google PAIR. URL: https://pair.withgoogle.com/guidebook [accessed 2025-06-10]
27. Thomander SD, Krosnick JA. Question and questionnaire design. In: Edlund JE, Nichols AL, editors. The Cambridge Handbook of Research Methods and Statistics for the Social and Behavioral Sciences: Volume 2: Performing Research: Cambridge University Press; 2024:352-370.
28. Braun V, Clarke V. Thematic Analysis: A Practical Guide: SAGE Publications; 2021.
29. Baxter K, Courage C, Caine K. Understanding Your Users: A Practical Guide to User Research Methods: Morgan Kaufmann; 2015.
30. Rosenzweig E. Successful User Experience: Strategies and Roadmaps: Morgan Kaufmann; 2015.
31. Beede E, Baylor E, Hersch F, et al. A human-centered evaluation of a deep learning system deployed in clinics for the detection of diabetic retinopathy. 2020 Apr 21 Presented at: CHI '20: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; Apr 25-30, 2020; Honolulu, HI p. 1-12 URL: https://dl.acm.org/doi/proceedings/10.1145/3313831 [accessed 2025-08-18] [doi: 10.1145/3313831.3376718]

Abbreviations

AI: artificial intelligence
EHR: electronic health record

Edited by A Kushniruk; submitted 18.04.25; peer-reviewed by S Banerjee, V Jawa, V Alemede; revised version received 10.06.25; accepted 26.06.25; published 26.08.25.

Please cite as:
Schreier M, Brandt R, Brown H, Saensuksopa T, Silva C, Vardoulakis LM
User-Centered Delivery of AI-Powered Health Care Technologies in Clinical Settings: Mixed Methods Case Study
JMIR Hum Factors 2025;12:e76241
URL: https://humanfactors.jmir.org/2025/1/e76241
doi:10.2196/76241

© Meredith Schreier, Randall Brandt, Hien Brown, T Saensuksopa, Christine Silva, Laura M Vardoulakis. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 26.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Screening Workers for Occupational Exposure to Respirable Crystalline Silica: Development and Usability of an Electronic Data Capture Tool

Fiona Hore-Lacy[1,2]; Christina Dimitriadis[1]; Ryan F Hoy[1,2]; Javier Jimenez-Martin[1]; Malcolm R Sim[1]; Jane Fisher[1]; Deborah C Glass[1]; Karen Walker-Bone[1]

[1]School of Public Health and Preventive Medicine, Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, Australia
[2]Department of Respiratory Medicine, Alfred Health, Melbourne, Australia

Corresponding Author:
Fiona Hore-Lacy
School of Public Health and Preventive Medicine, Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, Australia

Abstract

Background: Cases of the occupational lung disease silicosis have been identified in workers processing artificial stone in the stone benchtop industry (SBI). In the Australian state of Victoria, the Regulator commissioned a screening program for all workers in this industry.

Objective: To facilitate systematic data collection, including high-quality exposure assessment, an electronic data capture tool (EDCT) was developed.

Methods: A multidisciplinary team developed an EDCT using Research Electronic Data Capture (REDCap; Vanderbilt University). The needs of the EDCT were (1) data entry by multiple clinicians and the workers attending for screening and (2) systematic collection of data for clinical and research purposes. The comprehensibility and utility of the tool were investigated with a sample of workers, and the EDCT was subsequently refined.

Results: The EDCT was used in clinical practice, with capacity for data extraction for research. Testing of comprehension and utility was undertaken with 15 workers, and the refined version of the Occupational Silica Exposure Assessment Tool (OSEAT) was subsequently developed.

Conclusions: The refined OSEAT has been determined to be comprehensible to workers and capable of collecting exposure data suitable for assessment of risk of silicosis. It was developed for workers in the SBI in Australia and is adaptable, including translation into other languages. It can also be modified for SBI workers in other countries and for use by workers from other industries (mining, construction) at risk of silica exposure, including in lower-income settings.

(JMIR Hum Factors 2025;12:e64111) doi:10.2196/64111

KEYWORDS

silicosis; occupational history; electronic data capture tool (EDCT); REDCap; occupational respiratory screening; occupational hazard; exposure; silica; fibrotic lung disease; lung disease; respirable crystalline silica; mining; construction; workers; occupational lung disease; occupational; Australia; screening

Introduction

Silicosis is an incurable, potentially life-threatening, form of fibrotic lung disease caused by inhalation of respirable crystalline silica (RCS) [1]. The disease has been recognized globally for over 100 years, and lung disease screening is recommended for high-risk industries, including mining and construction [2]. Cases of silicosis were identified in 2010 among workers in the stone benchtop (countertop) industry (SBI) working with artificial stone (AS; Textbox 1) [3-12]. Subsequently, a number of cases of artificial stone-associated silicosis were diagnosed in Australia [3,13]. AS has a very high crystalline silica content, often over 90% [14]. Processing AS by drilling, polishing, cutting, or grinding generates fine particles of dust containing RCS, which can cause silicosis when inhaled [15].

XSL•FO
RenderX

Textbox 1. Timeline of artificial stone (AS) and silicosis in Australia

Early 2000s: AS introduced to Australia
2010: First case of silicosis associated with AS reported in Italy
2015: First case of silicosis associated with AS reported in Australia (conference abstract)
2017: First case of silicosis associated with AS reported in Australia
2019: Screening program of stone benchtop industry workers begins in Victoria, Australia (paper-based data collection)
2021: Screening program first incorporates electronic data capture tool (EDCT)
2021-2023: Refinement of EDCT informed by data cleaning and evaluation study with workers

An investigation of the effects of RCS exposure in the SBI was commissioned by the Victorian regulator, WorkSafe Victoria [16], and developed into a screening program by Monash University. It included (1) exposure assessment from a detailed occupational history; (2) collection of respiratory symptoms; (3) recording of investigations including spirometry, chest x-ray, and high-resolution CT of the chest; (4) screening for comorbidities associated with silica exposure, including autoimmune diseases and tuberculosis [4,17]; and (5) a mental health instrument.

The initial paper-based questionnaire was developed by a multidisciplinary team, including respiratory physicians, an occupational hygienist, and occupational physicians [16]. Simplicity was prioritized, as many industry workers were born outside Australia and spoke English as an additional language [18].

Up to 6 jobs in the SBI could be recorded in the occupational history. The proportion of time spent on specified *tasks* in each job and the proportion of time spent on dry cutting of stone and working near someone doing dry cutting were recorded. *Exposure control measures* (ventilation and respirator use) were identified for each job. Other information collected included the country, start and (if relevant) finish date; days per week worked; number of people in the organization; and type of stone predominantly worked with (AS or natural stone). Other silica-associated occupations (eg, mining, quarrying) and any non-occupational activities that involved dust exposure (including hobbies and home repairs, eg, tiling, plastering) were also recorded.

Data were collected from multiple users, including respiratory physicians, multidisciplinary team, workers, and administrative staff, and capture all the elements listed earlier. The data were cleaned and entered into an electronic data capture tool (EDCT) held on the secure REDCap platform [19,20]. The data were used for both clinical and research purposes. In 2021, screening was centralized and carried out at a single site, which led to a need for direct data entry to the EDCT by the worker.

Exposure calculations from the occupational data have been used to identify roles within the SBI with greater RCS exposure, such as factory machinists and installers, and that exposure intensity and cumulative exposure were associated with dyspnea and radiological abnormalities consistent with silicosis [21]. The screening program data has also been used to describe the numbers of cases of silicosis diagnosed to date [18], the rates and determinants of psychological distress, and the psychometric properties of the mental health instrument [22,23].

The aim of this study was to describe the development of the EDCT, present its refined content, and describe the results of an audit of its clinical utility undertaken with a sample of workers.

Methods

Overview

The original team reviewed the exposure questionnaire, which overall had been well understood, and identified items that required substantial data cleaning. The team redeveloped the questions, which included, for example, adding illustrations of dry and wet cutting examples and the types of respirator and ventilation options that were sourced from workplace health and safety organizations [24-26]. Further, additional optional responses were added from free text replies, for example, water jet cutting.

In the first draft questionnaire, participants were asked to apportion their tasks (Figure 1).

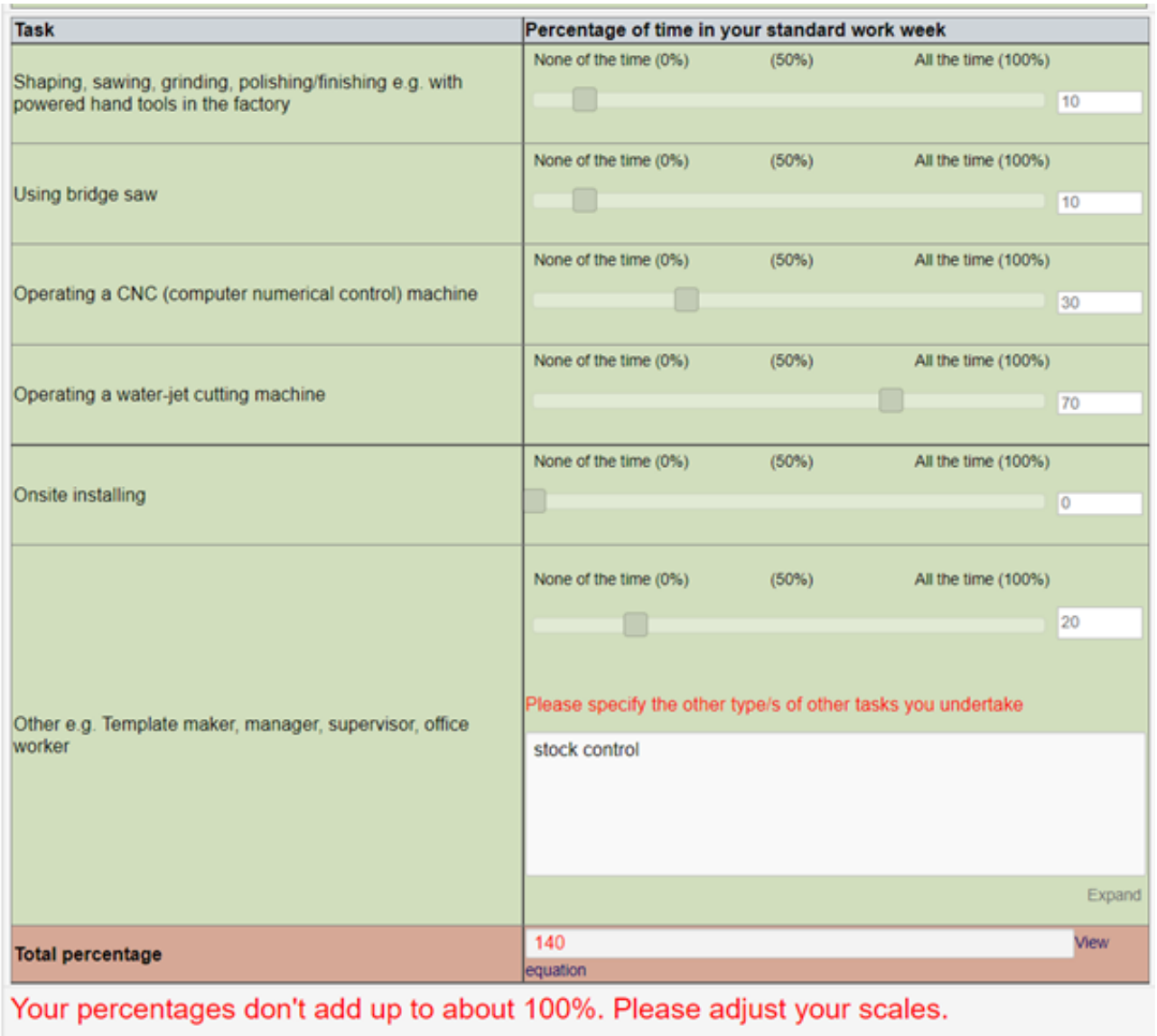
The percentages seldom added to 100%, as shown in the example in Figure 1, so in the revised questionnaire, a sliding bar was provided that provided visual input of the proportions (Figure 2). A pop-up trigger was included if the task proportions were out of range, as shown in Figure 2.

XSL•FO
RenderX

Figure 1. Original paper questionnaire asking workers to estimate proportions of work time spent doing specific tasks in their workplace. CNC: Computer Numerical Control____.

15.	What do you do in this job? (Tick all that apply)	
TASK		Estimate percentage of time in a typical week undertaking this task (e.g. 10%)
Shaping e.g. with powered hand tools	☑	20
Sawing e.g. with bridge saw	☑	50
Using CNC machine	☐	
Polishing/Finishing	☑	50
General labouring	☑	50
Maintenance	☑	40
Cleaning the tools, surfaces and/or work space	☑	50
Onsite Installing	☑	100
Other eg Template maker, manager, supervisor, office worker	☐	
If other, specify:_____		

Figure 2. EDCT (electronic data capture tool) version of the task estimation section of the occupational history with a total that adds to a proportion of >100% of the time and the warning message provided to the worker.



Prior to attending for screening, workers were emailed a link to the EDCT containing the revised questions in order to

XSL•FO
RenderX

complete their occupational history. The final page completed by workers included optional electronic consent for sharing data with Monash University.

Other data collected during screening included medical history, smoking status, respiratory symptoms, physical examination findings, diagnosis, return-to-work assessment, and results of all relevant investigations, including chest x-ray, high-resolution CT, pathology, and spirometry.

In 2023, an investigation of the comprehensibility, feasibility, and face validity of the occupational history section of the EDCT was completed with workers.

A total of 15 workers participated in the investigation, which was conducted between February and May 2023. Workers were interviewed about their responses on the EDCT, using a pre-developed proforma to prompt feedback. They were asked: "Were there any words you did not know?" "Did you understand what was meant by 'dry work'" "Are there any other tasks in your workplace that expose you to dust?" and "Did the list of ventilation options include what you use in your workplace?" For all questions, respondents could reply yes or no and provide additional comments. If a worker reported no dry cutting in their current or most recent job but had exposure to dry cutting from previous jobs, they were asked to respond to questions referring to their former job.

Ethical Considerations

Workers were eligible to participate if they had completed their occupational history using the EDCT, did not require an interpreter for their appointment, and had provided electronic consent to share data with Monash University. All participants were provided with a written information sheet and asked to provide verbal consent for participation. Responses were deidentified, and workers did not receive compensation for their participation. Approval was granted by the Alfred Hospital Research Ethics Committee as a substudy of project 292/21.

Results

The demographics of the workers are presented in Table 1. Most participants were male, and they covered a range of ages and years of employment in the SBI. They included machinists, installers, and office workers.

Table . Demographics of participants.

Characteristics		Values
Males, n (%)		14 (93)
Age (years), mean (SD)		38.1 (10.9)
Years in stone benchtop industry, mean (SD)		9.6 (7.85)
Born in Australia, n (%)		5 (33)
Language other than English spoken at home, n (%)		6 (40)
Most recent SBI job held, n		
	Director	3
	Installer	4
	Stonemason	5
	Foreman	2
	Other	1

All participants were asked to report about their *comprehension* of an introductory statement, and one worker commented that it took a while to read and understand, and did not feel like he understood it fully. All workers reported that they comprehended what was meant by "dry work." When asked about *tasks* that exposed them to dust that were not already listed on the EDCT, 2 workers identified new relevant tasks: emptying bins containing benchtop fragments and cleaning of the final benchtop product, onto which the dust-containing water used in wet cutting had dried. In total, 4 workers reported that their *ventilation* option was not included in the list on the EDCT, nor was it pictured. However, after discussion, the alternative options they were describing were "air conditioning," "garage door," "no ventilation," and "ventilation in the wall," all of which were listed.

For *respirators*, 14 workers identified the type they used from the descriptions and pictures in the EDCT, and the only worker that did not see their device described use of one similar to that depicted. Of 8 workers who were asked if they could easily remember and estimate the percentage of time they spent wearing a respirator, 7 responded in the affirmative. One worker had worn his respirator for 6 hours out of an 8-hour shift (approximately 75%), but had estimated that he wore it for 35% of the day. One worker commented that he wore his respirator all day, regardless of the task, whilst another pointed out that each job was different, with some jobs being "perfect" (ie, not requiring any adjustments), whereas others required adjustments onsite, adding to the difficulty of responding to this question accurately.

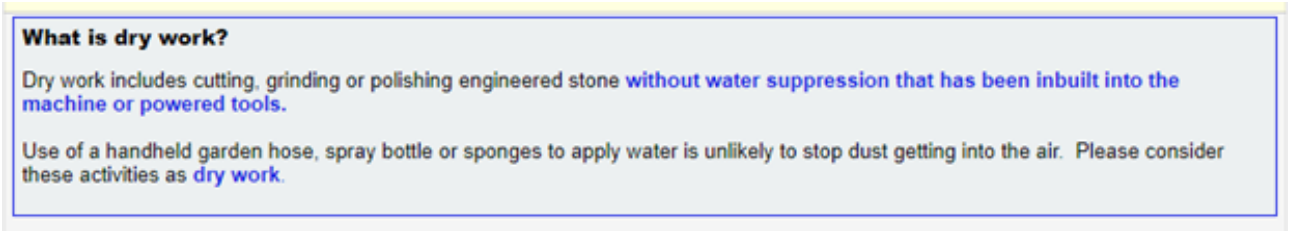
If a worker reported dry cutting in their current job, they were asked whether or not practices around dry cutting had changed (this was because dry cutting without suitable protection had been officially banned recently in Australia). Of the 11 workers who completed this question, 10 understood the question and were able to complete it accurately, but one worker expressed some confusion around the wording of the question. Workers

XSL•FO
RenderX

also described measures other than ventilation and respirators that their employer had introduced including changing clothes at work and using water systems for dust suppression when loading stone.

Subsequently, modifications were made to improve accessibility for the workers: simplification of the language used in descriptive statements and instructions, modification of the color scheme to improve readability, and addition of commonly reported responses (eg, "home maintenance," "tiling," and "plastering") as a prompt on the non-occupational (eg, hobby) dust exposure history section. One of the simplified statements was an introductory statement that defined "dry work" (Figure 3). An image accompanied the text with examples of dry and wet cutting [26]:

Figure 3. Refined EDCT for clarification of the definition of dry work.



The current version of the Occupational Silica Exposure Assessment Tool from the EDCT is presented in Multimedia Appendix 1.

Discussion

Principal Findings

Since 2019, over 1000 SBI workers have undergone screening for silicosis in Victoria, Australia, through a protocolized screening program. In this study, we described the development of the screening questionnaire into an EDCT, the Occupational Silica Exposure Assessment Tool, and have described how it was refined as a result of our experiences and after assessment of its acceptability and comprehension among this worker population. The results of the investigation suggested that the usability and comprehension of the refined EDCT are acceptable among English-speaking workers.

The benefits of EDCTs for improving patient care [27]; improving accuracy of data collection compared to paper methods [28]; and facilitating data collection from multiple users, including patients and health care providers [29], have been established in many settings. Moreover, EDCTs are able to capture and retain large volumes of data, maximizing cost- and time-efficiency in clinical and research settings [29,30]. As demonstrated, the development of this EDCT has already provided all of these benefits and enabled us to create a streamlined and efficient screening program for workers in the artificial stone benchtop industry.

Benchtops made from AS are a popular kitchen product globally, and there are concerns that cases of silicosis among workers who produce them are underreported in the literature [31,32]. Globally, silica deaths were estimated to be more than 12.9 thousand in 2019 [33], and the highest rates were recorded in low- and middle-income countries [33,34]. Silicosis has been seen in a range of industries, including construction, jewelry production, quarrying, tunneling, dental material manufacturing, denim jean production, and ceramic and pottery manufacturing [35]. There is therefore an urgent need for occupational screening for large numbers of workers exposed to RCS, for which reliable instruments are needed. The Occupational Silica Exposure Assessment Tool can be deployed in settings in which workers are exposed to RCS, whether for workers in the SBI or modified for other occupational settings.

In addition to assessing SBI workers at risk of silicosis, the data collected from the Occupational Silica Exposure Assessment Tool (OSEAT) can be used to estimate an individual's level of RCS exposure. In previous work, occupational history data collected using the OSEAT were used to group SBI workers by extent of silica exposure, using a combination of the proportion of time working with AS and the proportion of time spent dry cutting [21]. Both cumulative exposure and exposure intensity were found to be associated with symptoms of dyspnea and chest x-ray abnormalities [21]. This illustrates the ready utility of having added the e-consent function to the EDCT within REDCap, facilitating extracting data from the OSEAT to use for research purposes.

One consideration when introducing a REDCap-based EDCT in a clinical setting is its reliance on workers having adequate internet connection. This has been a limiting factor for the utilization of similar instruments, in which open-source software that was not reliant on internet connection was preferred [28,36]. In some resource-poor settings, REDCap was the preferred platform [37,38], and REDCap has now developed a mobile app that can be used offline that may overcome internet connection limitations [20].

A limitation of this study was the small number of workers included, none of whom required interpreters. Its comprehensibility is therefore unknown among those who require interpreters. Furthermore, these workers were recruited consecutively and recently from the clinic and consequently are not necessarily representative of the wider workforce. Another limitation was that the original instrument was not co-designed with consumers (workers in the SBI), something that we addressed in this study. We were unable to investigate floor or ceiling effects within the scope of this work, which would be a necessary step if the tool is to be used in translated versions.

A strength of the OSEAT is that it has been used and improved for close to 3 years with demonstrably good comprehension by SBI workers.

Future development of the OSEAT will include its translation into other languages using the REDCap multilanguage module.

XSL•FO
RenderX

The commonest languages other than English among the workers attending screening at our center have been Vietnamese, Persian, Chinese, and Arabic, and are therefore a priority for translation and inclusion into the module.

Conclusions

This study has presented the development, comprehension, utility, and refinement of the OSEAT, a purpose-built EDCT for use among SBI workers undergoing assessment for silicosis that included input from workers and has the capacity for modification and use within other silica-exposed occupational settings.

Acknowledgments

This project was funded by WorkSafe Victoria, and FHL was supported by a PhD project stipend from WorkSafe Victoria. The manuscript was prepared by the Monash Centre for Occupational and Environmental Health, Monash University. The opinions, findings, and conclusions expressed in this publication are those of the authors and not necessarily those of WorkSafe Victoria.

Data Availability

The datasets generated and analyzed during the current study are not publicly available for ethical and privacy reasons. Patient consent did not include sharing of data external to Monash University. If data need to be shared, additional approval from Monash University HREC and Alfred Hospital HREC would be required.

Authors' Contributions

FHL provided the concept and study design. Input into study design was provided by KWB and DCG. The EDCT was built by CD with input from RH, DG, FHL, MRS, JJM, and KWB. Data acquisition was performed by FHL. Data analysis was completed by FHL, with interpretation by FHL, KWB, and DCG. The manuscript was drafted by FHL, with significant contributions by KWB, RFH, and DCG, and additional editing provided by MRS, JJM, CD, and JF.

Conflicts of Interest

FHL was supported by a PhD project stipend from WorkSafe Victoria. The authors have no further interests to declare.

Multimedia Appendix 1
Occupational history data collection template: stone benchtop industry, and other silica and nonsilica exposed jobs.
[PDF File, 538 KB - humanfactors_v12i1e64111_app1.pdf]

References

1. Barnes H, Glaspole I. Occupational interstitial lung diseases. Immunol Allergy Clin North Am 2023 May;43(2):323-339. [doi: 10.1016/j.iac.2023.01.006] [Medline: 37055091]
2. Austin EK, James C, Tessier J. Early detection methods for silicosis in Australia and internationally: a review of the literature. Int J Environ Res Public Health 2021 Jul 31;18(15):8123. [doi: 10.3390/ijerph18158123] [Medline: 34360414]
3. Hoy RF, Baird T, Hammerschlag G, et al. Artificial stone-associated silicosis: a rapidly emerging occupational lung disease. Occup Environ Med 2018 Jan;75(1):3-5. [doi: 10.1136/oemed-2017-104428] [Medline: 28882991]
4. Shtraichman O, Blanc PD, Ollech JE, et al. Outbreak of autoimmune disease in silicosis linked to artificial stone. Occup Med 2015 Aug;65(6):444-450. [doi: 10.1093/occmed/kqv073] [Medline: 26070814]
5. Leso V, Fontana L, Romano R, Gervetti P, Iavicoli I. Artificial stone associated silicosis: a systematic review. Int J Environ Res Public Health 2019 Feb 16;16(4):568. [doi: 10.3390/ijerph16040568] [Medline: 30781462]
6. Kramer MR, Blanc PD, Fireman E, et al. Artificial stone silicosis [corrected]: disease resurgence among artificial stone workers. Chest 2012 Aug;142(2):419-424. [doi: 10.1378/chest.11-1321] [Medline: 22383661]
7. Grubstein A, Shtraichman O, Fireman E, Bachar GN, Noach-Ophir N, Kramer MR. Radiological evaluation of artificial stone silicosis outbreak: emphasizing findings in lung transplant recipients. J Comput Assist Tomogr 2016;40(6):923-927. [doi: 10.1097/RCT.0000000000000454] [Medline: 27680410]
8. Rosengarten D, Fox BD, Fireman E, et al. Survival following lung transplantation for artificial stone silicosis relative to idiopathic pulmonary fibrosis. Am J Ind Med 2017 Mar;60(3):248-254. [doi: 10.1002/ajim.22687] [Medline: 28145560]
9. Pérez-Alonso A, Córdoba-Doña JA, Millares-Lorenzo JL, Figueroa-Murillo E, García-Vadillo C, Romero-Morillos J. Outbreak of silicosis in Spanish quartz conglomerate workers. Int J Occup Environ Health 2014;20(1):26-32. [doi: 10.1179/2049396713Y.0000000049] [Medline: 24804337]
10. Pascual S, Urrutia I, Ballaz A, Arrizubieta I, Altube L, Salinas C. Prevalence of silicosis in a marble factory after exposure to quartz conglomerates. Arch Bronconeumol 2011 Jan;47(1):50-51. [doi: 10.1016/j.arbres.2010.09.004] [Medline: 21190766]

XSL•FO
RenderX

11. Pascual Del Pobil Y Ferré MA, García Sevila R, García Rodenas MDM, Barroso Medel E, Flores Reos E, Gil Carbonell J. Silicosis: a former occupational disease with new occupational exposure scenarios. Rev Clin Esp (Barc) 2019;219(1):26-29. [doi: 10.1016/j.rce.2018.06.006] [Medline: 30293675]
12. Pérez-Alonso A, Córdoba-Doña JA, León-Jiménez A. Silicosis caused by artificial quartz conglomerates: keys to controlling an emerging disease. Arch Bronconeumol (Eng Ed) 2019 Jul;55(7):394-395. [doi: 10.1016/j.arbr.2019.01.002]
13. Newbigin K, Parsons R, Deller D, Edwards R, McBean R. Stonemasons with silicosis: Preliminary findings and a warning message from Australia. Respirology 2019 Dec;24(12):1220-1221. [doi: 10.1111/resp.13672] [Medline: 31407419]
14. Kumarasamy C, Pisaniello D, Gaskin S, Hall T. What do safety data sheets for artificial stone products tell us about composition? A comparative analysis with physicochemical data. Ann Work Expo Health 2022 Aug 7;66(7):937-945. [doi: 10.1093/annweh/wxac020] [Medline: 35411922]
15. Barnes H, Goh NSL, Leong TL, Hoy R. Silica-associated lung disease: an old-world exposure in modern industries. Respirology 2019 Dec;24(12):1165-1175. [doi: 10.1111/resp.13695] [Medline: 31517432]
16. Hoy RF, Glass DC, Dimitriadis C, Hansen J, Hore-Lacy F, Sim MR. Identification of early-stage silicosis through health screening of stone benchtop industry workers in Victoria, Australia. Occup Environ Med 2021 Apr;78(4):296-302. [doi: 10.1136/oemed-2020-106897] [Medline: 33115923]
17. Ehrlich R, Akugizibwe P, Siegfried N, Rees D. The association between silica exposure, silicosis and tuberculosis: a systematic review and meta-analysis. BMC Public Health 2021 May 20;21(1):953. [doi: 10.1186/s12889-021-10711-1] [Medline: 34016067]
18. Hoy RF, Dimitriadis C, Abramson M, et al. Prevalence and risk factors for silicosis among a large cohort of stone benchtop industry workers. Occup Environ Med 2023 Aug;80(8):439-446. [doi: 10.1136/oemed-2023-108892] [Medline: 37328266]
19. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381. [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
20. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208. [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
21. Glass DC, Dimitriadis C, Hansen J, Hoy RF, Hore-Lacy F, Sim MR. Silica exposure estimates in artificial stone benchtop fabrication and adverse respiratory outcomes. Ann Work Expo Health 2022 Jan 7;66(1):5-13. [doi: 10.1093/annweh/wxab044] [Medline: 35015818]
22. Hore-Lacy F, Hansen J, Dimitriadis C, et al. Predictors of psychological stress in silica-exposed workers in the artificial stone benchtop industry. Respirology 2022 Jun;27(6):455-461. [doi: 10.1111/resp.14257] [Medline: 35421270]
23. Hore-Lacy F, Gwini S, Glass DC, et al. Psychometric properties of the Perceived Stress Scale (PSS-10) in silica-exposed workers from diverse cultural and linguistic backgrounds. BMC Psychiatry 2024 Mar 5;24(1):181. [doi: 10.1186/s12888-024-05613-6] [Medline: 38439053]
24. Health and Safety Executive. Controlling Airborne Contaminants at Work: A Guide to Local Exhaust Ventilation (LEV) - HSG258: Health and Safety Executive; 2017:111 URL: https://www.hse.gov.uk/pubns/books/hsg258.htm [accessed 2025-01-30]
25. Health and Safety Executive. Respiratory Protective Equipment at Work: A Practical Guide - HSG53: Health and Safety Executive; 2013:59 URL: https://www.hse.gov.uk/pubns/books/hsg53.htm [accessed 2025-01-30]
26. WorkSafe Victoria. Managing Exposure to Crystalline Silica: Engineered Stone Compliance Code - WSV1999/02/0822: WorkSafe Victoria; 2022:68 URL: https://www.worksafe.vic.gov.au/resources/archived-compliance-code-managing-exposure-crystalline-silica-engineered-stone [accessed 2025-01-30]
27. Denton E, Hore-Lacy F, Radhakrishna N, et al. Severe Asthma Global Evaluation (SAGE): an electronic platform for severe asthma. J Allergy Clin Immunol Pract 2019;7(5):1440-1449. [doi: 10.1016/j.jaip.2019.02.042] [Medline: 30954467]
28. Dillon DG, Pirie F, Rice S, et al. Open-source electronic data capture system offered increased accuracy and cost-effectiveness compared with paper methods in Africa. J Clin Epidemiol 2014 Dec;67(12):1358-1363. [doi: 10.1016/j.jclinepi.2014.06.012] [Medline: 25135245]
29. OME Cleveland Clinic Orthopaedics. Implementing a scientifically valid, cost-effective, and scalable data collection system at point of care: The Cleveland Clinic OME cohort. J Bone Joint Surg Am 2019 Mar 6;101(5):458-464. [doi: 10.2106/JBJS.18.00767] [Medline: 30845040]
30. Brown J, Bhatnagar M, Gordon H, Goodner J, Cobb JP, Lutrick K. An electronic data capture tool for data collection during public health emergencies: development and usability study. JMIR Hum Factors 2022 Jun 9;9(2):e35032. [doi: 10.2196/35032] [Medline: 35679114]
31. Fazio JC, Gandhi SA, Harrison RJ. Workers at risk of silicosis-ongoing overexposure and lack of medical surveillance-reply. JAMA Intern Med 2024 Feb 1;184(2):225-226. [doi: 10.1001/jamainternmed.2023.6626] [Medline: 38109093]
32. Wagner GR, Michaels D. Preventing the continuing tragedy of silicosis. JAMA Intern Med 2024 Feb 1;184(2):223-224. [doi: 10.1001/jamainternmed.2023.6629] [Medline: 38109105]
33. Chen S, Liu M, Xie F. Global and national burden and trends of mortality and disability-adjusted life years for silicosis, from 1990 to 2019: results from the Global Burden of Disease study 2019. BMC Pulm Med 2022 Jun 21;22(1):240. [doi: 10.1186/s12890-022-02040-9] [Medline: 35729551]

XSL•FO
RenderX

34. Wang H, Abbas KM, Abbasifard M, et al. Global age-sex-specific fertility, mortality, healthy life expectancy (HALE), and population estimates in 204 countries and territories, 1950–2019: a comprehensive demographic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct;396(10258):1160-1203. [doi: 10.1016/S0140-6736(20)30977-6] [Medline: 33069325]
35. Hoy RF, Chambers DC. Silica-related diseases in the modern world. Allergy 2020 Nov;75(11):2805-2817. [doi: 10.1111/all.14202] [Medline: 31989662]
36. Syzdykova A, Malta A, Zolfo M, Diro E, Oliveira JL. Open-source electronic health record systems for low-resource settings: systematic review. JMIR Med Inform 2017 Nov 13;5(4):e44. [doi: 10.2196/medinform.8131] [Medline: 29133283]
37. Maré IA, Kramer B, Hazelhurst S, et al. Electronic Data Capture System (REDCap) for health care research and training in a resource-constrained environment: technology adoption case study. JMIR Med Inform 2022 Aug 30;10(8):e33402. [doi: 10.2196/33402] [Medline: 36040763]
38. Marroquin Rivera A, Rosas-Romero JC, Castro SM, et al. Implementing a Redcap-based research data collection system for mental health. Rev Colomb Psiquiatr (Engl Ed) 2021 Jul;50 Suppl 1(Suppl 1):110-115. [doi: 10.1016/j.rcpeng.2021.06.004] [Medline: 34257053]

Abbreviations

AS: artificial stone
CNC: Computer Numerical Control
CT: computed tomography
EDCT: electronic data capture tool
OSEAT: Occupational Silica Exposure Assessment Tool
RCS: respirable crystalline silica
REDCap: Research Electronic Data Capture
SBI: stone benchtop industry

Edited by A Kushniruk; submitted 09.07.24; peer-reviewed by C Berriault, RN Carey, SJL Linde; revised version received 14.11.24; accepted 24.11.24; published 24.02.25.

Please cite as:
Hore-Lacy F, Dimitriadis C, Hoy RF, Jimenez-Martin J, Sim MR, Fisher J, Glass DC, Walker-Bone K
Screening Workers for Occupational Exposure to Respirable Crystalline Silica: Development and Usability of an Electronic Data Capture Tool
JMIR Hum Factors 2025;12:e64111
URL: https://humanfactors.jmir.org/2025/1/e64111
doi:10.2196/64111

© Fiona Hore-Lacy, Christina Dimitriadis, Ryan F Hoy, Javier Jimenez-Martin, Malcolm R Sim, Jane Fisher, Deborah C Glass, Karen Walker-Bone. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 24.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Collecting Real-World Data via an In-Home Smart Medication Dispenser: Longitudinal Observational Study of Survey Panel Persistency, Response Rates, and Psychometric Properties

Benjamin Ogorek[1], PhD; Thomas Rhoads[1], MBA; Erica Smith[2], PhD

[1]Spencer Health Solutions Inc, 2501 Aerial Center Pkwy, Suite 100, Morrisville, NC, United States

[2]WCG, Princeton, NJ, United States

Corresponding Author:
Benjamin Ogorek, PhD
Spencer Health Solutions Inc, 2501 Aerial Center Pkwy, Suite 100, Morrisville, NC, United States

Abstract

Background: A smart medication dispenser called "spencer" is a novel generator of longitudinal survey data. The patients dispensing medication act as a survey panel and respond to questions about quality of life and patient-reported outcomes.

Objectives: Our goal was to evaluate panel persistency, survey response rates, reliability, and validity of surveys administered via spencer to 4138 polychronic patients residing in the United States and Canada.

Methods: Patients in a Canadian health care provider's program were included if they were dispensing via spencer in the June 2021 to February 2024 time frame and consented to have their data used for research. Panel persistency was estimated via discrete survival methods for 2 years and survey response rates were computed for 1 year. Patients were grouped by mean response rates in the 12th month (<90% vs ≥90%) to observe differential response rate trends. For reliability and validity, we used a spencer question about recent falls with ternary responses value-coded −1, 0, and 1. For reliability, we computed Pearson correlation between mean scores over 2 years of survey responses, and transitions between mean score intervals of [0, 0.5), [−0.5, 0.5), and [0.5, 1]. For validity, we measured the association between the falls question and known factors influencing fall risk: age, biological sex, quality of life, physical and emotional health, and use of selective serotonin reuptake inhibitors or serotonin-norepinephrine reuptake inhibitors, using repeated-measures regression for covariates and Kendall τ for concomitant spencer questions.

Results: From 4138 patients, dispenser persistency was 68.3% (95% CI 66.8% - 69.8%) at 1 year and 51% (95% CI 49% - 53%) at 2 years. Within the cohort observed beyond 1 year, 82.3% (1508/1832) kept surveys enabled through the 12th month with a mean response rate of 84.1% (SD 26.4%). The large SD was apparent in the subgroup analysis, where a responder versus nonresponder dichotomy was observed. For 234 patients with ≥5 fall risk responses in each of the first 2 years, the Pearson correlation estimate between yearly mean scores was 0.723 (95% CI 0.630 - 0.798). For mean score intervals [0, 0.5), [−0.5, 0.5), and [0.5, 1], self-transitions were the most common, with 59.8% (140/234) of patients starting and staying in [0.5, 1]. Fall risk responses were not significantly associated with sex (P=.66) or age (P=.76) but significantly related to selective serotonin reuptake inhibitor or serotonin-norepinephrine reuptake inhibitor usage, quality of life, depressive symptoms, physical health, disability, and trips to the emergency room (P<.001).

Conclusions: A smart medication dispenser, spencer, generated years of longitudinal survey data from patients in their homes. Panel attrition was low, and patients continued to respond at high rates. A fall risk measure derived from the survey data showed evidence of reliability and validity. An alternative to web-based panels, spencer is a promising tool for generating patient real-world data.

(JMIR Hum Factors 2025;12:e60438) doi:10.2196/60438

KEYWORDS

real-world data; real-world evidence; patient-reported outcomes; longitudinal studies; survey methods

Introduction

Background

The use of patient data collected in real-world settings has never been more impactful. The US Food and Drug Administration's Real-World Evidence (RWE) Program has elevated real-world data (RWD) as a tool to support new indications for already approved drugs [1-3], the European Medicines Agency has published their RWE framework [4], and Canada's Drug and Health Technology Agency has published their guidance document [5].

XSL•FO
RenderX

RWD may take the form of claims records, electronic health records, registries, or patient-generated data, with patient-reported outcomes (PROs) as an important subset. Longitudinal surveys, where patients are surveyed at 2 or more points in time, generate data that allow for the analysis of within-unit change as well as aggregations over time [6]. This results in greater "causal leverage" than cross-sectional surveys [7] and is ideal for submissions to regulatory bodies.

Web-based panels, or "registered persons who have agreed to take part in online studies on a regular basis," rose in attractiveness with the proliferation of the web [8]. Recently, however, shortcomings of longitudinal studies based on web-based panels have undermined their reputation as a high-quality data source. Panel attrition, where subjects in earlier waves cease to respond in later waves, has become worse since the 1990s [7,9-11]. While web-based panel data are also prone to quality problems (eg, false answers, careless responses, and multiple panel memberships [12]), these problems have been exacerbated by innovations in automation and improvements in large language models, where human reviewers are unable to consistently detect automated responses [13]. This has become a corrupting force in web-based survey sampling [14].

Amazon Mechanical Turk (MTurk) was considered a representative and convenient source of web-based longitudinal survey data [15] but has seen its reputation deteriorate within the last decade. For example, a study that used MTurk to build a diabetes panel failed after only 5.8% (13/224) were deemed eligible for future survey research [16]. Researchers noted declines in MTurk data quality starting around summer 2018, as evidenced by degraded psychometric properties of well-understood metrics [17]. A warning was issued in the journal *Perspectives on Psychological Science* after an exercise revealed that only 2.6% (14/529) of MTurk samples were valid [18].

Alternatives to web-based panels exist in populations of patients using web-connected hardware, also known as "smart" products. One interesting subset is the population of patients using smart medication dispensers, as these products sit in the home amidst a public health need for digital adherence solutions [19]. A 2023 review of smart medication adherence products reviewed the features of 51 products without mention of survey administration capabilities [20]. One of these products, a dispenser named "spencer" [21,22] (manufactured by Spencer Health Solutions, Inc), has a touch screen display that allows it to administer survey questions following on-time medication dispenses (Figure 1).

At the time of writing, in-home spencer devices have generated more than 3 million longitudinal responses from more than 4000 unique patients to quality of life and PRO measures from a polychronic population residing in the United States and Canada. These are patients of Canadian health care provider Custom Health, Inc, a company offering "a personalized, connected service that goes beyond medication management and ensures medications are working as they should" [23]. Patients or caregivers can express interest directly via a collection of sign-up forms [24,25] or they may be directed to spencer via their health plans that have partnered with Custom Health [26]. When a health plan partners with Custom Health, their services are provided to members "who require a high degree of clinical oversight, those managing multiple medications or those experiencing medication adherence challenges" [27].

XSL•FO
RenderX

Figure 1. Key components of the spencer smart medication dispenser.



XSL•FO
RenderX

Objectives

The study's aim was to evaluate the spencer smart medication dispenser as a longitudinal survey platform for a polychronic patient population. Panel persistency, survey response rates, and measurement reliability and validity were assessed.

Methods

Recruitment

Patients of Custom Health were included if they met the criteria enumerated in Textbox 1.

Textbox 1. Inclusion criteria for patients in the study.

1. The patient entered Custom Health's intake process either by self-selection or based on the recommendation of a health care provider.
2. Custom Health professionals decided to pair the patient with a spencer smart medication dispenser.
3. The patient agreed to the Custom Health consent form.
4. The patient agreed to the Spencer Health Solutions End User License Agreement, permitting his or her deidentified data to be used for research purposes. This occurred on the spencer unit's touch screen.
5. The patient's first scheduled medication dose was between June 3, 2021, and February 14, 2024.
6. The patient dispensed a medication dose by March 14, 2024. In this paper, dispensing medication refers to dispensing multidose packs containing oral solids.

After completing Custom Health's intake process, spencer devices were shipped to patients' homes. Once set up, the devices displayed both current local time and the scheduled time of the next medication dispense via a touch screen display. Refills containing medication strips (multidose adherence packaging) prepared by a pharmacy were shipped to the patients' homes and were inserted by the patient or care nurse into the top of the unit via an electronically controlled door. At scheduled dosing times, the unit alerted the patient through sound, light, and a message on the touch screen display. After the patient pressed the dispense button on the touch screen, the unit dispensed 1 or more medication pouches. After an on-time dispense, a question was presented to those patients who had not explicitly opted out of surveys.

Data Generation and Processing

The question and response mechanism is further elaborated here. If a dose was dispensed on time and the patient had not opted out of surveys, 1 question was displayed on-screen. To answer, a single button press was needed to select from a multiple-choice answer set. This was followed by a review step (also serving as the completeness check) where the patient could confirm the selection or go back and change the answer. If a patient did not make the confirmation in the review step, the questionnaire would not be submitted to the database and later analyzed as an instance of nonresponse. If left attended to, a question would remain on the screen until the next scheduled dose.

In collaboration with health care professionals at Custom Health, 35 survey items were designed to measure the spencer experience, quality of life, and PROs. To avoid copyright infringement, these questions were not taken from any existing validated scale. Questions were scheduled one-to-one with doses in a predefined sequence that repeated indefinitely. Response options were consistently listed from most positive sentiment (eg, "Excellent") toward the top of the screen to least positive sentiment (eg, "Poor") toward the bottom of the screen. Questions were manually answered on test devices in a quality assurance laboratory before being released to patients, and patients could call into a support line to provide feedback regarding the questions or to request that they be turned off.

As is typical of web-based panels, the panel formed by the selection criteria in Textbox 1 constitutes a convenience sample. The target population best described by the sample is polychronic patients taking multiple medications daily. Since the surveys were administered as part of routine patient monitoring, no institutional review board (IRB) approval was needed.

Survey responses were sent to the application database through either cellular connection (the default) or Wi-Fi. In cases where the spencer lost connectivity, a store and forward mechanism sent data to the cloud database once connectivity was reestablished. The database is managed by Spencer Health Solutions that has received both ISO27001 [28] and Data Privacy Framework [29] certifications.

Date of birth, biological sex, and residential postal code were entered into a web-based portal by health care providers when patients were recruited. These fields were retrieved from the application database March 4, 2024. Dates of birth that were within 2 years of the database entry date were replaced with missing values, and age was computed as the difference between the first dispense date and date of birth. Within the United States, 5-digit postal codes were converted into US states via the *zipcodeR* R package. For Canadian postal codes, a function was written that maps the first letter of the postal code to the associated province. Prescription information was created by the pharmacies at the time of refill creation and sent to the database.

Statistical Analysis

Panel Attrition

Patients may leave the spencer dispensing platform for multiple reasons, including life transitions to higher care services or natural death. Leaving the dispensing platform is the primary mechanism of spencer panel attrition. To estimate dispenser persistency, we used the discrete survival analysis framework described by Allison [30], where the periods start on the first

XSL•FO
RenderX

day a patient is scheduled, are 30 days in length, and an attrition event occurs when a patient is not scheduled during an entire 30-day period. There is a *resurrection* mechanism: when a patient is scheduled in a later period after previously meeting the definition for an attrition event, the attrition flag is reset for all previous periods.

For readability, we will refer to a 30-day period as a *month*, 12 thirty-day periods as a *year*, and so on. Furthermore, we will refer to the time in years between the first scheduled dose via spencer and the analysis date as *tenure*. A patient's tenure represents the amount of experience a patient has had with the spencer platform as of an analysis date.

Beyond dispenser attrition, the second source of panel attrition is when patients request that their questions no longer appear on-screen following a dispensing event. To study this phenomenon, we computed rates of requested question discontinuation for the first 12 months of tenure for patients who remained dispenser persistent for more than a year.

When pursuing the subset of patients who were dispenser persistent for more than a year, the subset taken was patients who remained on the spencer platform through the 14th month. In addition to our operational month being shorter on average than a calendar month (by a fraction of a day), the 2 additional months of persistency provided a buffer against the decreased interaction with the device that often precedes full platform discontinuation.

Survey Response Rates

On the survey platform, nonresponse occurs when patients do not enter a response after a question is displayed and the question is cleared. We computed rates of nonresponse by month and plotted the resulting series. We knew from prior analyses that some patients consistently respond to the questions, and we wanted to observe this phenomenon. For patients who were still receiving surveys in the 12th month, we created 2 groups: those with 12th-month response rates of <90% and those with 12th-month response rates of ≥90%. For both groups, we computed the frequency of patients, plotted response rates by month, and provided a qualitative description of the patterns observed.

Psychometric Analysis

Reliability

A reliability analysis in the context of a platform requires a narrowing of focus to a specific measure, as both reliable and unreliable measures may be generated from any platform. Inspired by the Falls Efficacy Scale-International, a reliable measure of fear of falling known to be related to both past and future falls [31,32], we chose an existing question from our rotation that asks the patient about recent falls. Hereafter referred to as Q_FALL, the question text read "Have you experienced a fall in the past month?" The response options were "No," "Not Sure," and "Yes" (a 1-letter variation in capitalization occurring after September 2022, where "Not sure" was replaced with "Not Sure"), which were value-coded as 1, 0, and −1, respectively.

One conceptualization of reliability is test-retest reliability and can be quantified using Pearson correlation between a measure's values at 2 time points [33]. For a comparative baseline in the literature, Falls Efficacy Scale-International measurements taken by the same patients at different time intervals had Pearson correlations ranging from 0.66 to 0.83 for measurements taken up to a year apart [34].

The Pearson correlation coefficient is known to suffer bias when distributional assumptions are violated, a concern because Q_FALL has only 3 response levels and there were different response counts between patients and years. The use of averages and bootstrap resampling were thus employed to address these factors. First, we limited attention to a subset of 234 patients from the persistency analysis who answered Q_FALL at least 5 times in both a full first year and a second year of tenure, hereafter referred to as year 1 and year 2. Second, we used the bootstrap to obtain a bias-corrected estimate of the Pearson correlation along with a nonparametric 95% CI [35]. This allowed us to perform inference on the coefficient of determination (R^2) for the equivalent regression of the year 2 means regressed on the year 1 means.

Averaging the ternary scores allowed us to work on a continuum where rare fallers and never fallers appear close together on the resulting scale, a notion supported by similarities between these groups in a 1-year cohort study [36]. To circumvent the limitations of a linear correlation analysis, we performed an additional discrete state transition analysis. We examined the frequency of transitions to and from mean Q_FALL scores of [−1, −0.5), [−0.5, 0.5), and [0.5, 1.0] in year 1 and year 2, expecting self-transitions to be the most frequent.

Validity

To assess the convergent validity of the recent falls question administered via spencer, the mean scores for year 1 and year 2 were compared with the following established risk factors of fall risk: increased age, biological sex, previous fall frequency, low quality of life, depressive symptoms, physical impairment, and medication use [31,34,37]. Many patients in this population were prescribed selective serotonin reuptake inhibitors (SSRIs) or serotonin-norepinephrine reuptake inhibitors (SNRIs), and these are associated with falls in the older adults [38,39]. A meta-analysis found that 95% (70/74) of studies reported gender or sex differences in fall-related outcomes with females at a higher risk than males [40]. Canonically, increased age is a risk factor for falls [41]. The validity analysis was split into 2 parts, each based on the 234-patient subset from the reliability analysis.

We first conducted an analysis of the relationship between raw Q_FALL values and covariates age, sex, and SSRI or SNRI usage, as these were known before any responses were received (medication can be discontinued but medication classes tend to be stable within patient). To accommodate the repeated measures received from each patient, we used a generalized estimating equation approach to model the relationship between the coded value of Q_FALL and an exchangeable working correlation structure. This was accomplished with the *geepack* package in R, which reports SEs that are robust to both the choice of working correlation structure and nonnormality of the response. For age and SSRI or SNRI usage, we expected to see negative

XSL•FO
RenderX

relationships. For biological sex, we expected that female patients would be associated with lower mean Q_FALL than male patients.

For evidence of association between Q_FALL and other relevant variables, including quality of life, depressive symptoms, and hospital visits, we selected the questions listed in Table 1 as contemporaneous survey-based measures that had face validity for concepts of interest. Their responses are integer-coded and arranged by sentiment, and thus we expected positive correlations with Q_FALL.

The robust Kendall τ measure was used to test for associations, as the sample sizes of the questions from Table 1 may be arbitrarily small within patient. Kendall τ is more appropriate for ties and has an accompanying 2-sided nonparametric test for testing the null hypothesis of zero association [35]. For a nonparametric 95% CI on τ, we used the *kendall.ci* function from the R package *NSM3* [42], which provides a bootstrap CI.

Table . Standard spencer questions relating to known risk factors of falling.

Question text	Possible responses	Values coded	Construct[a]
Rate your recent quality of life.	Excellent | Very good | Good | Fair | Poor	5 | 4 | 3 | 2 | 1	Quality of life
How is your emotional health today?[b]	Excellent | Very good | Good | Fair | Poor	5 | 4 | 3 | 2 | 1	Depression
How would you rate your physical health today?	Excellent | Very good | Good | Fair | Poor	5 | 4 | 3 | 2 | 1	Physical health
Rate your ability to perform activities today.	Excellent | Very good | Good | Fair | Poor	5 | 4 | 3 | 2 | 1	Ability or disability
Are you able to accomplish what you have planned today?	Completely | Mostly | Moderately | A little | Not at all	5 | 4 | 3 | 2 | 1	Ability or disability
Hospital, ER[c], or urgent care in the past month?	No | Not sure | Yes	1 | 0 | −1	ER visits from falls

[a]Construct is based on face validity of the spencer standard questions.

[b]This question has been in rotation for multiple years, but in September 2022, the number of responses changed from 3 (“Poor,” “Good,” and “Excellent”) to 5 (“Poor,” “Fair,” “Good,” “Very good,” and “Excellent”). We coded the 3-response set as 1, 3, and 5, and the 5-response set as 1, 2, 3, 4, and 5, respectively.

[c]ER: emergency room.

Ethical Considerations

This study used operational data collected from a commercial medication dispensing system used in routine patient care and was not subject to IRB review requirements, so IRB approval was not pursued. Users of the spencer device provided consent for data collection through the End User License Agreement, which covers the collection of medication adherence data and responses to quality of life and PRO surveys as part of the system’s standard operation. No additional compensation was provided to users beyond the normal terms of their device usage agreement. All data analyzed in this study were deidentified prior to analysis. Spencer Health Solutions has achieved both ISO27001 [28] and Data Privacy Framework [29] certifications, and the system uses industry-standard encryption and security measures.

This research analyzed data collected during standard clinical care and device usage. All results are presented as anonymous aggregate statistics. The original data collection occurred as part of routine clinical practice, with patients providing consent for research use through the device terms of service and care management agreement. Under Canadian TCPS 2 Article 2.4, research ethics board review is not required for research that relies exclusively on secondary use of anonymous information where the process does not generate identifiable information. Under US regulation 45 CFR 46.104(d)(4)(ii), IRB review is not required when information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects, the investigator does not contact the subjects, and the investigator will not reidentify subjects.

Results

Patient Population

The patient population was majority female (2552/4133, 61.7%), with 0.1% (5/4138) of the biological sex fields missing. The mean age was 54.4 years (SD 19.9, range 5-104 years). Most patients (3736/4138, 90.3%) resided in Canada, with 1 address unmapped at the country level. Patients were scheduled to take multiple drugs per day (mean 9.6, SD 5.1). Of the 2805 unique compounds scheduled during the observation window, 70.2% (1970/2805) were mapped to an Anatomical Therapeutic Chemical classification system second-level code, with a modified “Vitamins & Supplements” that included dietary supplements. Table 2 contains the 20 most frequently observed subgroups observed during the observation window.

XSL•FO
RenderX

Table . Patient demographics, geographic distribution, and medication usage among the 4138 patients studied.

Section and variable	Patients (N=4138)
Patient demographics, n (%)	
Sex	
Female	2552 (61.7)
Male	1581 (38.2)
Missing	5 (0.1)
Age (years)	
Valid records, n (%)	4123 (99.6)
Invalid records, n (%)	15 (0.4)
Mean (SD)	54.4 (19.9)
IQR	39 - 70
Age range	5 - 104
Geographic distribution, n (%)	
Country	
Canada	3736 (90.3)
United States	401 (9.7)
Missing or other	1 (0)
Canadian provinces	
Ontario	2083 (50.3)
British Columbia	1149 (27.8)
Saskatchewan	485 (11.7)
Other	19 (0.5)
US states	
Tennessee	221 (5.3)
Missouri	117 (2.8)
California	32 (0.8)
Ohio	24 (0.6)
Other	7 (0.2)
Medication usage (Anatomic Therapeutic Chemical codes, second level), n (%)	
Psychoanaleptics	3040 (73.5)
Vitamins & supplements	1917 (46.3)
Lipid-modifying agents	1848 (44.7)
Drugs for acid-related disorders	1788 (43.2)
Antiepileptics	1717 (41.5)
Agents acting on the renin-angiotensin system	1664 (40.2)
Psycholeptics	1503 (36.3)
Drugs used in diabetes	1317 (31.8)
Beta-blocking agents	1107 (26.8)
Antithrombotic agents	1050 (25.4)
Calcium channel blockers	893 (21.6)
Diuretics	822 (19.9)
Thyroid therapy	741 (17.9)

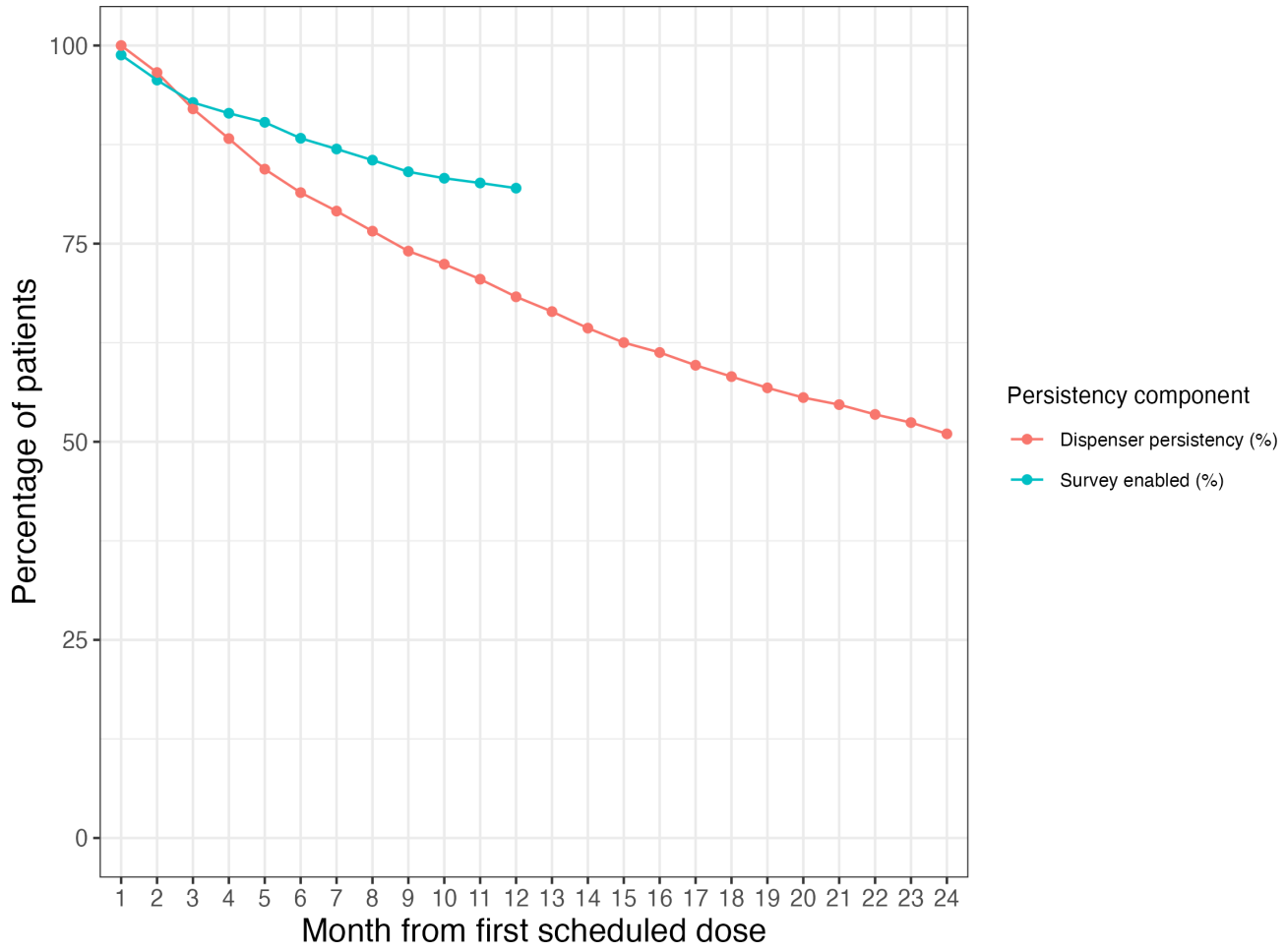
XSL•FO
RenderX

Section and variable	Patients (N=4138)
Urologicals	592 (14.3)
Analgesics	543 (13.1)
Anti-inflammatory and antirheumatic products	378 (9.1)
Antihypertensives	345 (8.3)
Anti-Parkinson drugs	309 (7.5)
Drugs for constipation	295 (7.1)
Antihistamines for systemic use	289 (7)

Panel Attrition

From 4138 patients, dispensing persistency was estimated to be 68.3% (95% CI 66.8% - 69.8%) at year 1 and 51% (95% CI 49% - 53%) at year 2. Among the patients who stayed on the dispensing platform past year 1, 82.3% (1508/1832) kept surveys enabled through the 12th month. The rates of question opt-out slowed during the year, as can be seen in Figure 2.

Figure 2. Dispenser persistency and percentage of patients with surveys enabled by month on platform.

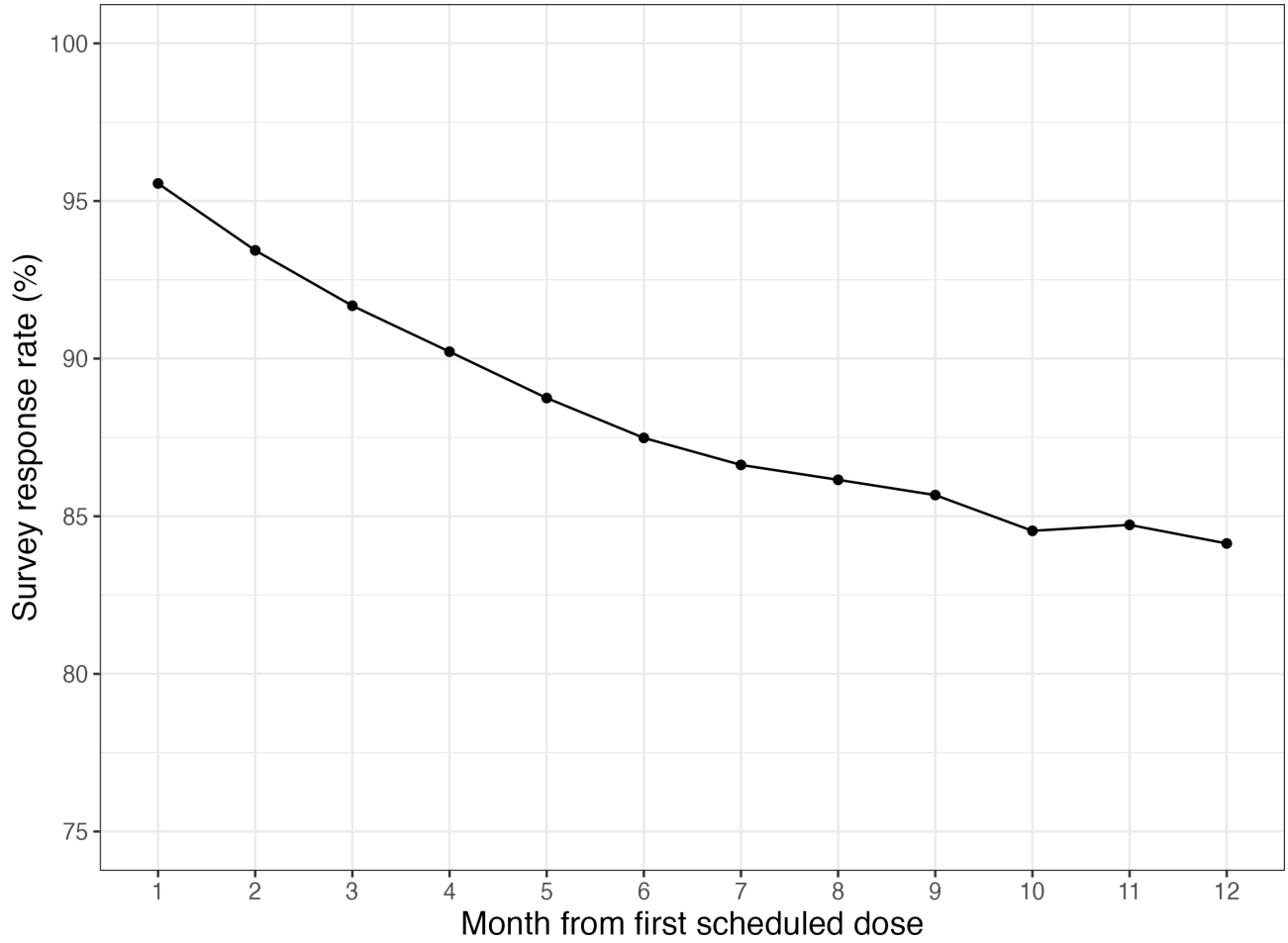


Survey Response Rates

Among the 1508 patients who kept their surveys enabled through year 1, the mean response rate was 95.6% (SD 11.9%) in the first month and 84.1% (SD 26.4%) in the 12th month, with the rate of decline slowing in the second half of the year (Figure 3). For patients with surveys enabled in the 12th month, 67.9% (1024/1508) had response rates at or above 90% and 32.1% (484/1508) had response rates below 90%. Figure 4 shows the trajectories of both groups, where the high-response group maintained near perfect response rates while the low-response group experienced a substantial decline by the 12th month.

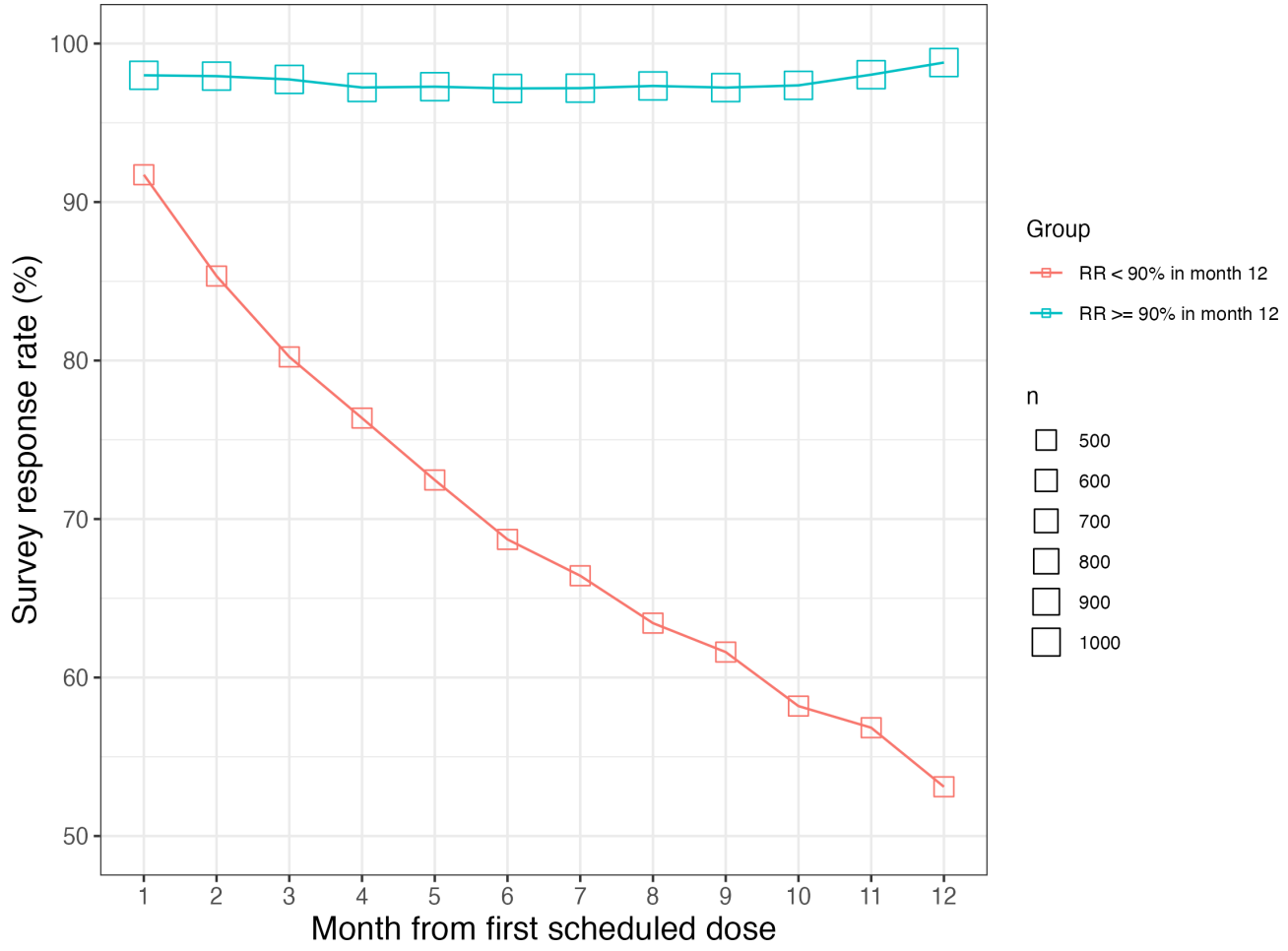
XSL•FO
RenderX

Figure 3. Mean survey response rate (%) for patients with enabled surveys by month on platform.



XSL•FO
RenderX

Figure 4. Mean survey RR by month on platform, grouped by RR in the 12th month. RR: response rate.



Reliability

Among 234 patients, the bootstrap estimate of the Pearson correlation between the year 1 and year 2 mean Q_FALL was 0.723 (95% CI 0.630 - 0.798), and the estimate of R^2 for the equivalent regression on year 2 versus year 1 means was 0.523 (95% CI 0.397 - 0.637).

Mean Q_FALL interval transitions from year 1 to year 2 are shown in Table 3. As hypothesized, self-transitions were the most common, with 59.8% (140/234) of patients starting and staying in [0.5, 1]. These rare fallers in the [0.5, 1.0] interval in year 1 remained in this interval during year 2 in 83.8% (140/167) of cases. For the frequent fallers in the [−1.0, −0.5) interval during year 1, 66.7% (12/18) remained during year 2.

XSL•FO
RenderX

Table . Mean score range transitions from year 1 to year 2[a].

Year 1 and year 2 (n=234)		Frequency, n (%)
Frequent fallers [−1.0, −0.5) (n=18)		
	[−1.0, −0.5)	12 (66.7)
	[−0.5, 0.5)	5 (27.8)
	[0.5 to 1.0]	1 (5.6)
Occasional fallers [−0.5, 0.5) (n=49)		
	[−1.0, −0.5)	5 (10.2)
	[−0.5, 0.5)	24 (49)
	[0.5 to 1.0]	20 (48.8)
Rare fallers [0.5, 1.0] (n=167)		
	[−1.0, −0.5)	1 (0.6)
	[−0.5, 0.5)	26 (15.6)
	[0.5 to 1.0]	140 (83.8)

[a]A mean score of +1.0 would be a "perfect score" of no reported falls, a mean score of −1.0 indicates that all responses indicated that a recent fall had occurred, and scores in between span the interval (−1.0, 1.0). One patient who moved from [−1.0, −0.5) to [0.5, 1.0] had a score of exactly 0.5 (the boundary), with only 6 measurements in year 2.

Validity

Covariates

For 232 patients, the generalized estimating equation model results of Q_FALL on these covariates are shown in Table 4. The coefficients associating biological sex and age to Q_FALL were not significantly different from zero; this was unexpected. However, the coefficient indicating the presence of an SSRI or SNRI medication was negative and highly significant, indicating more falls in the SSRI or SNRI group adjusted for sex and age.

Table . Generalized estimating equation linear model summary for sex, age, and whether the patient was prescribed a selective serotonin reuptake inhibitor or serotonin-norepinephrine reuptake inhibitor during the observation window.

Coefficient	Estimate	SE	Wald	*P* value
Intercept	0.684	0.127	28.857	<.001[a]
Sex (male)	−0.030	0.069	0.195	.66[b]
Patient age	0.001	0.002	0.092	.76[b]
SSRI[c] or SNRI[d]	−0.232	0.062	14.092	<.001[b]

[a]The *P* value corresponding to the hypothesis of the intercept being zero is included by convention but is not a meaningful statistic.

[b]These *P* values correspond to the 2-sided test of the hypothesis of a zero regression coefficient.

[c]SSRI: selective serotonin reuptake inhibitor.

[d]SNRI: serotonin-norepinephrine reuptake inhibitor.

Contemporaneous Outcomes From the Spencer

The analysis of correlations between Q_FALL and other contemporaneous spencer questions, based on Kendall τ, is shown in Table 5. Interpretation of correlation coefficients varies, for example, 0.2 may be characterized as either "weak" or "poor," and 0.3 as "weak," "moderate," or "fair" [43], and in the bivariate normal case, a τ value of 0.200 corresponds to a Pearson correlation of 0.309 [44]. While the strength of association between mean Q_FALL and the contemporaneous response outcomes was consistently weak to moderate, *P* values were uniformly small, indicating positive relationships of these questions with the measure of recent falls.

XSL•FO
RenderX

Table . Contemporaneous association between survey questions administered via spencer.

Question	Patients, n[a]	Average responses per patient, n	Kendall τ	95% bootstrap CI for τ	*P* value[b]
Rate your recent quality of life.	233	14	0.15	0.058-0.244	<.001
How is your emotional health today?	233	54	0.21	0.120-0.293	<.001
How would you rate your physical health today?	164	8	0.18	0.071-0.296	<.001
Rate your ability to perform activities today.	232	35	0.23	0.140-0.316	<.001
Are you able to accomplish what you have planned today?	197	9	0.18	0.083-0.276	<.001
Hospital, ER[c], or urgent care in the past month?	192	9	0.20	0.096-0.299	<.001

[a]Number of unique patients who responded to each question at least once and also responded at least 5 times to Q_FALL in year 1 and year 2.

[b]Derived from Kendall τ test, a nonparametric hypothesis test used to measure the ordinal association between 2 variables.

[c]ER: emergency room.

Discussion

Principal Findings

Although its primary function is dispensing medication, the spencer platform doubled as a web-based longitudinal panel where polychronic patients answered survey questions at high rates and exhibited low panel attrition over years of platform tenure. Measures generated from the responses where stable through time (ie, evidence of reliability) and were associated with other theoretically related variables (ie, evidence of validity). For polychronic patients residing in the US and Canada, the home medication dispenser is a promising source of reliable and valid measures of important health constructs.

As with all survey panels, there was attrition and nonresponse. Panel attrition could be decomposed into attrition from the dispensing platform and survey opt-outs for patients remaining on the dispensing platform. These losses were cumulative. Based on the estimates presented, starting with 100 patients, 68 would still be dispensing via spencer by the end of the first year, with 56 still receiving questions following their dispenses.

In our literature review, persistency was often a serious issue in the context of longitudinal patient studies. For 8 remote digital studies conducted between 2014 and 2019, researchers found that more than half of all participants discontinued their participation within the first week of the study [9]. In a web-based study during the COVID-19 pandemic, of 2734 participants who completed wave 1, only 964 participated in wave 3 [10]. In a study of smartphone app usage to improve oral anticoagulation adherence, a retention rate of 27% at 6 months was reported [11]. Considering these results, keeping more than half of the initial patients actively participating in surveys at the end of the first year represents favorable retention.

The rate of new survey opt-outs also decreased substantially through the year, setting up milder losses in year 2.

By the end of first year, the average survey response rate for patients taking surveys on spencer was 84%. While 80% has been considered excellent in the context of primary care research studies [45], multi-item surveys administered at a single point in time are an imperfect benchmark. Ecological momentary assessment, a survey methodology that addresses phenomena as they occur, typically sees compliance rates from 50% to 90% [46]. By either standard, the response rates observed in spencer surveys were good.

We can speculate on why some patients responded to fewer spencer questions over time than others. Survey fatigue is a well-known phenomenon that occurs when respondents become weary of repeated survey tasks [47], and although surveys administered via spencer are brief, they are frequent. In addition to fatigue, some patients may not have been aware of how the questions were being used to monitor their well-being. Developing interventions to improve survey response rates is a topic for future research.

Noncoverage in web-based surveys, often defined as lack of access to the web, is thought to be a more serious problem than nonresponse, which is an unwillingness to participate [48]. Since every participant has a connection to the web through the device itself (the spencer units have both cell and Wi-Fi connections), there is no noncoverage in the sense of lack of web access, although machines do go offline for varying durations.

With sufficiently high panel persistency and response rates, attention focuses on the quality of the data that are generated. We showed that a measure about recent falls generated from a spencer question exhibited temporal stability, a form of reliability. The measure showed expected associations with most theoretically related variables. One exception was the

XSL•FO
RenderX

demographic factors of age and biological sex, which failed to achieve significance in a regression where medication use was significant ($P<.001$). While additional data may reveal the expected relationships, we surmise that in a polychronic population taking many medications, demographic factors may be weaker predictors of falling than in the general population.

Limitations

First, since the patients studied in this paper were polychronic patients enrolled in a care management program and residing in the United States and Canada, inferences to other populations may not be warranted.

Second, our validity analyses were limited to data collected entirely within the spencer ecosystem. While correlations between spencer survey responses suggest meaningful patterns, these questions, although having face validity, lack validation against established instruments. The observed correlations might partly reflect the consistent presentation format on the device, where responses are always ordered from most to least positive sentiment. Our validity arguments would be substantially stronger with independent measurements, particularly comparisons between spencer responses and validated traditional instruments measuring the same constructs.

Third, this study did not consider sensitivity to change, which is important in the context of RWE because it allows researchers and clinicians to detect change resulting from a minimal intervention [49]. While we focused on reliability by treating fall risk as a stable construct, and although fall risk is sufficiently stable to support our reliability analysis, treating it as static was a limitation of this research. Future research could explore methods for estimating changing states from longitudinal survey data, building on established approaches in the literature [50,51].

Conclusions

Administering longitudinal surveys via spencer, a smart medication dispenser, effectively generated high-quality RWD from patients in their homes. Patients persisted on the platform for years and maintained high response rates. A measure derived from longitudinal surveys assessing fall risk demonstrated both reliability and validity. The performance of spencer as a longitudinal survey platform offers a promising alternative at a time when web-based survey data quality is deteriorating.

Because medication dispensing is a fundamental component of the survey-generating mechanism, RWD from the spencer platform offers an ideal opportunity to study medication effectiveness and health outcomes, providing evidence to support new drug indications and demonstrate relationships between health outcomes and economic factors.

Acknowledgments

This work was funded by Spencer Health Solutions, Inc, a private corporation headquartered in Morrisville, North Carolina. The authors wish to thank University of Maryland professor Greg Hancock for his perspective on presenting this new source of patient-reported outcomes. The authors thank Devi Kodavanti, a member of the data team at Spencer Health Solutions, for computing many reliability coefficients in the past that gave them confidence to do this research. Their resident designer Lew Wilson created the table of contents image. Finally, the authors thank Alan Menius, the scientist behind the evolution of the spencer standard questions.

Data Availability

The dataset analyzed during this study is not publicly available to protect patient privacy and comply with health care privacy regulations, since it contains patient-reported outcomes collected through medication-dispensing devices. While the data are deidentified, they represent longitudinal health-related information of single patients alongside demographic information; it is difficult to guarantee that reidentification, however unlikely, is impossible. However, controlled access to specific portions of the data are available from the corresponding author on reasonable request, which should include (1) a brief research proposal outlining their intended use of the data, and (2) documentation of their institution's data security protocols.

Authors' Contributions

All authors participated in the conceptualization, review, and revision of the final manuscript draft for submission. BO proposed the methodology, performed the formal analysis, and wrote the original draft. TR supervised the research through all stages and performed validation activities.

Conflicts of Interest

Authors BO and TR are current employees of Spencer Health Solutions, Inc, the developer of the spencer platform. Author ES was employed at Spencer Health Solutions during the initial research and manuscript development phase.

References

1. Framework for FDA's real‐world evidence program. U.S. Food and Drug Administration. 2018. URL: https://www.fda.gov/media/120060/download [accessed 2025-01-23]

XSL•FO
RenderX

2. Klonoff DC. The New FDA Real-World Evidence Program to support development of drugs and biologics. J Diabetes Sci Technol 2020 Mar;14(2):345-349. [doi: 10.1177/1932296819832661] [Medline: 30862182]
3. Considerations for the use of real-world data and real-world evidence to support regulatory decision-making for drug and biological products. U.S. Food and Drug Administration. 2023. URL: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/considerations-use-real-world-data-and-real-world-evidence-support-regulatory-decision-making-drug [accessed 2025-01-23]
4. Real-world evidence framework to support EU regulatory decision-making: report on the experience gained with regulator-led studies from september 2021 to february 2023. European Medicines Agency. 2023. URL: https://www.ema.europa.eu/system/files/documents/report/real-world-evidence-framework-support-eu-regulatory-decision-making-report-experience-gained_en.pdf [accessed 2025-01-23]
5. Guidance for reporting real-world evidence. Canada's Drug Agency. 2023. URL: https://www.cadth.ca/sites/default/files/RWE/MG0020/MG0020-RWE-Guidance-Report-Secured.pdf [accessed 2025-01-23]
6. Lynn P. Methodology of Longitudinal Surveys: John Wiley & Sons, Ltd; 2009. [doi: 10.1002/9780470743874]
7. Hillygus DS, Snell S. In: Atkeson LR, Alvarez RM, editors. Longitudinal Surveys: Issues and Opportunities: Oxford University Press; 2015. [doi: 10.1093/oxfordhb/9780190213299.013.7]
8. Göritz AS, Reinhold N, Batinic B. Online panels. In: Online Social Sciences: Hogrefe & Huber Publishers; 2002:27-47.
9. Pratap A, Neto EC, Snyder P, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. NPJ Digit Med 2020;3:21. [doi: 10.1038/s41746-020-0224-8] [Medline: 32128451]
10. Yu T, Chen J, Gu NY, Hay JW, Gong CL. Predicting panel attrition in longitudinal HRQoL surveys during the COVID-19 pandemic in the US. Health Qual Life Outcomes 2022 Jul 6;20(1):104. [doi: 10.1186/s12955-022-02015-8] [Medline: 35794553]
11. Senoo K, Miki T, Ohkura T, et al. A smartphone app to improve oral anticoagulation adherence in patients with atrial fibrillation: prospective observational study. JMIR Mhealth Uhealth 2022 Jan 7;10(1):e30807. [doi: 10.2196/30807] [Medline: 34894626]
12. Hays RD, Liu H, Kapteyn A. Use of Internet panels to conduct surveys. Behav Res Methods 2015 Sep;47(3):685-690. [doi: 10.3758/s13428-015-0617-9] [Medline: 26170052]
13. Lebrun B, Temtsin S, Vonasch A, Bartneck C. Detecting the corruption of online questionnaires by artificial intelligence. Front Robot AI 2023;10:1277635. [doi: 10.3389/frobt.2023.1277635] [Medline: 38371744]
14. Research in the crowdsourcing age, a case study. Pew Research Center. 2016. URL: https://www.pewresearch.org/internet/2016/07/11/research-in-the-crowdsourcing-age-a-case-study/ [accessed 2025-01-23]
15. Buhrmester M, Kwang T, Gosling SD. Amazon's Mechanical Turk: a new source of inexpensive, yet high-quality, data? Perspect Psychol Sci 2011 Jan;6(1):3-5. [doi: 10.1177/1745691610393980] [Medline: 26162106]
16. Lin YK, Newman S, Piette J. Response consistency of crowdsourced web-based surveys on type 1 diabetes. J Med Internet Res 2023 Aug 18;25:e43593. [doi: 10.2196/43593] [Medline: 37594797]
17. Chmielewski M, Kucker SC. An MTurk crisis? Shifts in data quality and the impact on study results. Soc Psychol Personal Sci 2020 May;11(4):464-473. [doi: 10.1177/1948550619875149]
18. Webb MA, Tangney JP. Too good to be true: bots and bad data from Mechanical Turk. Perspect Psychol Sci 2022 Nov:174569162211200. [doi: 10.1177/17456916221120002] [Medline: 36343213]
19. Hafez G, Aarnio E, Mucherino S, et al. Barriers and unmet educational needs regarding implementation of medication adherence management across Europe: insights from COST Action ENABLE. J Gen Intern Med 2024 Jun;39(15):2917-2926. [doi: 10.1007/s11606-024-08851-2] [Medline: 38941058]
20. Faisal S, Samoth D, Aslam Y, et al. Key features of smart medication adherence products: updated scoping review. JMIR Aging 2023 Dec 19;6:e50990. [doi: 10.2196/50990] [Medline: 38113067]
21. Patel T, Ivo J, Pitre T, Faisal S, Antunes K, Oda K. An in-home medication dispensing system to support medication adherence for patients with chronic conditions in the community setting: prospective observational pilot study. JMIR Form Res 2022 May 19;6(5):e34906. [doi: 10.2196/34906] [Medline: 35587371]
22. Baby B. User Experience with Medication Adherence Technology: Determining Usability by Capabilities [Graduate thesis]. : University of Waterloo; 2024 URL: http://hdl.handle.net/10012/20613 [accessed 2025-01-23]
23. Personal solutions. Custom Health. URL: https://customhealth.com/solutions/personal.html [accessed 2025-01-23]
24. Sign up. Custom Health. URL: https://customhealth.com/signup/ [accessed 2025-01-23]
25. i'm spencer: the future of connected care!. Custom Health. URL: https://customhealth.com/spencer.html [accessed 2025-01-23]
26. Professional solutions. Custom Health. URL: https://customhealth.com/solutions/professional-solutions.html [accessed 2025-01-23]
27. Custom health to provide services to health plan of San Mateo members. PR Newswire. 2022 Nov 29. URL: https://www.prnewswire.com/news-releases/custom-health-to-provide-services-to-health-plan-of-san-mateo-members-301689232.html [accessed 2025-01-23]

XSL•FO
RenderX

28. ISO/IEC 27001:2022: information security, cybersecurity and privacy protection — information security management systems — requirements. International Organization for Standardization. 2022. URL: https://www.iso.org/standard/27001 [accessed 2025-01-23]
29. Data Privacy Framework Program. URL: https://www.dataprivacyframework.gov/ [accessed 2025-01-23]
30. Allison PD. Discrete-time methods for the analysis of event histories. Sociol Methodol 1982;13:61. [doi: 10.2307/270718]
31. Yardley L, Beyer N, Hauer K, Kempen G, Piot-Ziegler C, Todd C. Development and initial validation of the Falls Efficacy Scale-International (FES-I). Age Ageing 2005 Nov;34(6):614-619. [doi: 10.1093/ageing/afi196] [Medline: 16267188]
32. Chen Y, Du H, Song M, et al. Relationship between fear of falling and fall risk among older patients with stroke: a structural equation modeling. BMC Geriatr 2023 Oct 11;23(1):647. [doi: 10.1186/s12877-023-04298-y] [Medline: 37821821]
33. Jhangiani RS, Chiang IC, Cuttler C, Leighton DC. Research Methods in Psychology, 4th edition: Kwantlen Polytechnic University; 2019.
34. Delbaere K, Close JCT, Mikolaizak AS, Sachdev PS, Brodaty H, Lord SR. The Falls Efficacy Scale International (FES-I). A comprehensive longitudinal validation study. Age Ageing 2010 Mar;39(2):210-216. [doi: 10.1093/ageing/afp225] [Medline: 20061508]
35. Bishara AJ, Hittner JB. Reducing bias and error in the correlation coefficient due to nonnormality. Educ Psychol Meas 2015 Oct;75(5):785-804. [doi: 10.1177/0013164414557639] [Medline: 29795841]
36. Asai T, Oshima K, Fukumoto Y, Yonezawa Y, Matsuo A, Misu S. The association between fear of falling and occurrence of falls: a one-year cohort study. BMC Geriatr 2022 May 5;22(1):393. [doi: 10.1186/s12877-022-03018-2] [Medline: 35509040]
37. Dewan N, MacDermid JC. Fall Efficacy Scale-International (FES-I). J Physiother 2014 Mar;60(1):60. [doi: 10.1016/j.jphys.2013.12.014] [Medline: 24856947]
38. Darowski A, Chambers SA, Chambers DJ. Antidepressants and falls in the elderly. Drugs Aging 2009;26(5):381-394. [doi: 10.2165/00002512-200926050-00002] [Medline: 19552490]
39. Park H, Satoh H, Miki A, Urushihara H, Sawada Y. Medications associated with falls in older people: systematic review of publications from a recent 5-year period. Eur J Clin Pharmacol 2015 Dec;71(12):1429-1440. [doi: 10.1007/s00228-015-1955-3] [Medline: 26407688]
40. Sebastiani C, Wong JYX, Litt A, et al. Mapping sex and gender differences in falls among older adults: a scoping review. J Am Geriatr Soc 2024 Mar;72(3):903-915. [doi: 10.1111/jgs.18730] [Medline: 38147460]
41. Ambrose AF, Paul G, Hausdorff JM. Risk factors for falls among older adults: a review of the literature. Maturitas 2013 May;75(1):51-61. [doi: 10.1016/j.maturitas.2013.02.009] [Medline: 23523272]
42. Schneider G, Chicken E, Becvarik R. NSM3: functions and datasets to accompany Hollander, Wolfe, and Chicken—Nonparametric Statistical Methods, Third Edition. The Comprehensive R Archive Network. 2023. URL: https://CRAN.R-project.org/package=NSM3 [accessed 2025-01-23]
43. Akoglu H. User's guide to correlation coefficients. Turk J Emerg Med 2018 Sep;18(3):91-93. [doi: 10.1016/j.tjem.2018.08.001] [Medline: 30191186]
44. Gilpin AR. Table for conversion of Kendall's tau to Spearman's rho within the context of measures of magnitude of effect for meta-analysis. Educ Psychol Meas 1993 Mar;53(1):87-92. [doi: 10.1177/0013164493053001007]
45. Booker QS, Austin JD, Balasubramanian BA. Survey strategies to increase participant response rates in primary care research studies. Fam Pract 2021 Sep 25;38(5):699-702. [doi: 10.1093/fampra/cmab070] [Medline: 34213547]
46. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
47. de Koning R, Egiz A, Kotecha J, et al. Survey fatigue during the COVID-19 pandemic: an analysis of neurosurgery survey response rates. Front Surg 2021 Aug 12;8:690680. [doi: 10.3389/fsurg.2021.690680] [Medline: 34458314]
48. Couper MP, Kapteyn A, Schonlau M, Winter J. Noncoverage and nonresponse in an internet survey. Soc Sci Res 2007 Mar;36(1):131-148. [doi: 10.1016/j.ssresearch.2005.10.002]
49. Hays RD, Hadorn D. Responsiveness to change: an aspect of validity, not a separate dimension. Qual Life Res 1992 Feb;1(1):73-75. [doi: 10.1007/BF00435438] [Medline: 1301117]
50. Le H, Schmidt FL, Putka DJ. The multifaceted nature of measurement artifacts and its implications for estimating construct-level relationships. Organ Res Methods 2009 Jan;12(1):165-200. [doi: 10.1177/1094428107302900]
51. Ozkok O, Zyphur MJ, Barsky AP, Theilacker M, Donnellan MB, Oswald FL. Modeling measurement as a sequential process: Autoregressive Confirmatory Factor Analysis (AR-CFA). Front Psychol 2019;10:2108. [doi: 10.3389/fpsyg.2019.02108] [Medline: 31616338]

Abbreviations

IRB: institutional review board
MTurk: Amazon Mechanical Turk
PRO: patient-reported outcome
Q_FALL: question about recent falls asked on spencer

XSL•FO
RenderX

RWD: real-world data
RWE: real-world evidence
SNRI: serotonin-norepinephrine reuptake inhibitor
SSRI: selective serotonin reuptake inhibitor

Edited by K Adapa; submitted 10.05.24; peer-reviewed by J Misko, LB Iantovics; revised version received 06.11.24; accepted 03.12.24; published 03.02.25.

Please cite as:
Ogorek B, Rhoads T, Smith E
Collecting Real-World Data via an In-Home Smart Medication Dispenser: Longitudinal Observational Study of Survey Panel Persistency, Response Rates, and Psychometric Properties
JMIR Hum Factors 2025;12:e60438
URL: https://humanfactors.jmir.org/2025/1/e60438
doi:10.2196/60438

© Benjamin Ogorek, Thomas Rhoads, Erica Smith. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 3.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Comparisons of Physicians', Nurses', and Social Welfare Professionals' Experiences With Participation in Information System Development: Cross-Sectional Survey Study

Susanna Martikainen[1], PhD; Johanna Viitanen[2], DSc; Samuel Salovaara[3], PhD; Ulla-Mari Kinnunen[1], PhD; Tinja Lääveri[2,4], MD, PhD

[1]Department of Health and Social Management, Faculty of Social Sciences and Business Studies, University of Eastern Finland, Kuopio, Finland

[2]Department of Computer Science, School of Science, Aalto University, Espoo, Finland

[3]Department of Social Work, Faculty of Social Sciences, University of Lapland, Rovaniemi, Finland

[4]Infectious Diseases, Inflammation Center, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

Corresponding Author:
Susanna Martikainen, PhD
Department of Health and Social Management, Faculty of Social Sciences and Business Studies, University of Eastern Finland, Kuopio, Finland

Abstract

Background: The integration of health care and social welfare services together with the consolidation of health care information systems (HISs) and client information systems (CISs) has become a timely topic. Despite this development, there is a scarcity of systematic research on physicians', registered nurses' (RNs) and social welfare professionals' (SWPs) experiences of participating in the development of HISs and CISs.

Objective: This study aimed to examine how physicians, RNs and SWPs experience collaboration with HIS or CIS vendors, and what kinds of end users have participated in HIS or CIS development.

Methods: National cross-sectional usability surveys were conducted in Finland among RNs and SWPs in 2020 and physicians in 2021. Questions concerning participation experiences were analyzed by professional group, working sector, managerial position, and age.

Results: In total, 4683 physicians, 3610 RNs, and 990 SWPs responded to the surveys. In all 3 professional groups, those working in nonmanagerial positions and the youngest respondents participated least in HIS or CIS development, and 76% (n=3528) of physicians, 78% (n=2814) of RNs and 67% (n=664) of SWPs had not participated at all. When comparing the groups, physicians were least aware of feedback processes and least satisfied with vendors' interest in end-user feedback and the manner and speed of HIS development. Those who had dedicated working time for HIS or CIS development were less critical of vendors' interest and responsiveness to development ideas than those who had not participated at all. In all 3 professional groups, the youngest were most dissatisfied with HIS and CIS vendor collaboration.

Conclusions: Experiences of participation in HIS and CIS development were relatively negative across all 3 professional groups, with physicians being the most critical. Dialogue and collaboration between developers and end users—also the youngest ones and frontline workers—need improvement; simply increasing allotted working time is unlikely to produce more positive participation experiences.

(JMIR Hum Factors 2025;12:e51495) doi:10.2196/51495

KEYWORDS

participation; development; usability; user experience; physician; nurse; social worker; information system; national survey; system development; users; user feedback; cross-sectional survey; Finland; Finnish

Introduction

Background

The increasing collaboration between health care and social care and the consequent need for integrated information systems (ISs) warrants studies on the end-user participation viewpoints of all major user groups. Physicians [1-5], nurses [1,2], and social welfare professionals (SWPs) [6] are not satisfied with the usability or daily work support of health care information systems (HISs) and client information systems (CISs) [7]. Furthermore, these professionals experience collaboration with HIS and CIS vendors and developers to be unsatisfactory. However, the majority would be willing to participate in

XSL•FO
RenderX

information system development, but suitable means are lacking [1,2,6,8-11].

End users' participation in IS development is considered to lead to a better user experience and increased user acceptance [12,13]. However, the characteristic inherent complexity of HISs and CISs complicates the implementation of many user participation methods [14,15]. Although end-user participation is regarded as essential in HIS and CIS development [16], without careful management throughout the software development process, participation alone does not guarantee system success [17-19]. End-user participation may even cause more problems than benefits, particularly if the ISs are expected to solve organizational problems [20]. Users may experience that their participation does not affect IS functionalities in a desired manner [1,4,6,16]. In addition, emphasizing administrative information needs and wishes instead of the needs of frontline professionals can complicate work processes by increasing the requirements for data entry and thus impair the workflows [21]. IS users in leadership positions have different needs for ISs than frontline professionals [22].

Comprehensive HISs and CISs are complex ISs used by dozens of user roles in a wide variety of use contexts [23,24]. Many countries, including Finland, are integrating health and social welfare services and consequently ISs [25-27]. One of the rationales behind this development is that those with high numbers of visits to health care often need social welfare services and vice versa [28]. From the point of view of patients and clients, treatment and service packages often include both health care and social services [29-33], which emphasizes the need for fluent information exchange between professional groups to guarantee high-quality and safe care. Consequently, end users from the major professional groups are needed in the IS development processes.

Context of the Study: Health Care, Social Welfare and HISs and CISs in Finland

Until 2023, municipalities (n=309 in 2022) were responsible for organizing social welfare services and primary health care (health centers) in Finland. A total of 20 hospital districts, jointly owned by the municipalities of the region, organized specialized medical care; 5 university hospitals provided tertiary care. Although one-third of outpatient visits to physicians are to private providers (eg occupational health care), the variety of services provided by the private sector is narrow; for example, there are no private intensive care units (ICUs) or labor and delivery units [34]. In social welfare, municipalities or federations of municipalities often purchase some services from private service providers and non-governmental organizations (n=3971 in 2017) [27]. In 2018, there were 19,627 working-age physicians, 70,198 RNs, and 34,523 SWPs in Finland [35-37].

In Finland, the first HISs and CISs were implemented in the 1970s [27]. While only every tenth US hospital used electronic health records (EHRs) as late as in 2010 [38], in Finnish public health care, HIS coverage had already reached 100% by 2007, and by 2014 CISs covered almost all public social services [39,40]. By contrast, in 2020, a quarter of nonpublic social welfare organizations still operated on paper [27,41].

All public hospitals and health centers had joined the Kanta services (national patient data repository and electronic prescription system) by 2015 [42,43], but implementation of the national data repository for social welfare services only began in 2020 [44]. This has required considerable resources from both health care and social welfare organizations and IS vendors over the years [45].

During 2020‐21, in public health care, 2 leading specialized care and 2 leading primary care EHR brands were in wide use. In addition, 4 EHR brands covered both primary and specialized care, of which 1 also covered tertiary care (including functionalities for eg, operating rooms, ICUs, radiology, and emergency departments) and 6 out of the 7 nationally defined social welfare service lines. In addition, 1 EHR brand covered most private sector health care. In public social welfare, 2 CIS brands were in use in most municipalities [46]. In addition, EHR brands were also used in social welfare [41]. In 2018‐21, a new IS was deployed in Southern Finland with 47,000 end users.

Research Questions

In this study, we examined the experiences of physicians, registered nurses (RNs) and SWPs of participating in HIS or CIS development. The data were gathered in 3 large Finnish national surveys in 2020 and 2021. The research questions were as follows:

1. What experiences do physicians, RNs and SWPs have of collaboration with HIS and CIS vendors?
2. Do participation experiences vary by managerial position, employment sector, or age?
3. What types of physicians, RNs and SWPs have participated in HIS and CIS development?

Methods

Survey

This study was part of large national cross-sectional HIS and CIS usability surveys conducted among SWPs and RNs in 2020, and physicians in 2021 [47]. The survey questionnaires and data are available online [47]. The surveys were based on the validated National Usability-Focused HIS Scale [48] and included a section on end users' experiences of participation in HIS or CIS development (Table 1). The statements were originally created and piloted for the national physician survey in 2010 [4,5], and the same statements have been used in later surveys for physicians [1], for RNs [8,49], and SWPs [6]. The survey method and the questionnaire have been described in detail previously [4,5,48].

XSL•FO
RenderX

Table . Questionnaire statements.

Questions and statement/option designations		Statements and options
Question (1) What has been your experience of providing feedback on the information systems you use and their development? please assess the following statements based on your experience. Response options: Fully agree / Somewhat agree / Neither agree nor disagree / Somewhat disagree / Fully disagree		
	Statement A	I know how and to whom I can send feedback about the system if I wish to do so.
	Statement B	The system vendor is interested in feedback about the system provided by the end users.
	Statement C	The system vendor implements corrections and change requests according to the suggestions of the end users.
	Statement D	Corrections and change requests are implemented within a reasonable time frame.
Question (2) Have you participated in information systems development work?		
	Option A	Yes, some of my working time has been allocated for such development work
	Option B	Yes, in addition to my work
	Option C	No

The link to the survey questionnaire was sent to the members of the Finnish Medical Association (FMA) (>90%) (email address available, n=19,142), RN members of the Finnish Nursing Association, the National Association of Health and Welfare Professionals, and the National Professional Association (n=58,276), and SWPs with at least a Bachelor's degree who were members of the following trade unions: the Union of Professional Social Workers, the Trade Union for the Public and Welfare Sectors, or the Social Science Professionals union (n=12,471) [9,50,51] .

The section addressing end-user participation experiences in HIS or CIS development (Table 1) was identical for the physicians and the SWPs with 5-point Likert scale response statements (fully agree; somewhat agree; neither agree nor disagree; somewhat disagree; or fully disagree). However, for the 2020 RNs' survey, a sixth option "Prefer not to respond / Don't know" was added. Furthermore, unlike in the physicians' and SWPs' surveys, it was not possible for RNs to not respond at all. Up to 25% of RNs chose this sixth option. To make the surveys more comparable, we formed a sixth category also for physicians and SWPs of those who had not responded to the statements. In addition, for the descriptive statistics in Figures 1-4 and Table S1 in Multimedia Appendix 1, we combined the responses "Fully agree" and "Somewhat agree" to form a new category "Agree" and included those from the sixth category in the denominator.

The statement on having allocated time to participate in IS development was identical in all 3 surveys, and the background questions included in this study were all optional in all 3 surveys, so nonrespondents were not included in these numbers.

XSL•FO
RenderX

Figure 1. Participation experiences of Finnish physicians, RNs and SWPs by leadership position (leaders versus others). RN: registered nurse; SWP: social welfare professional.

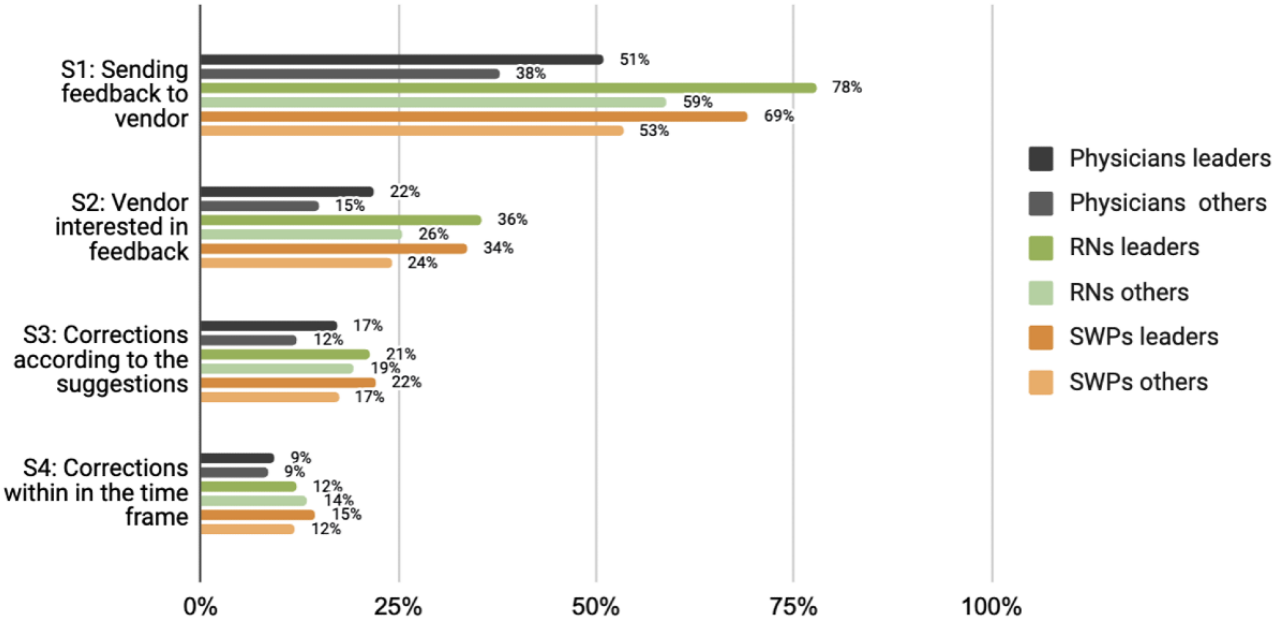
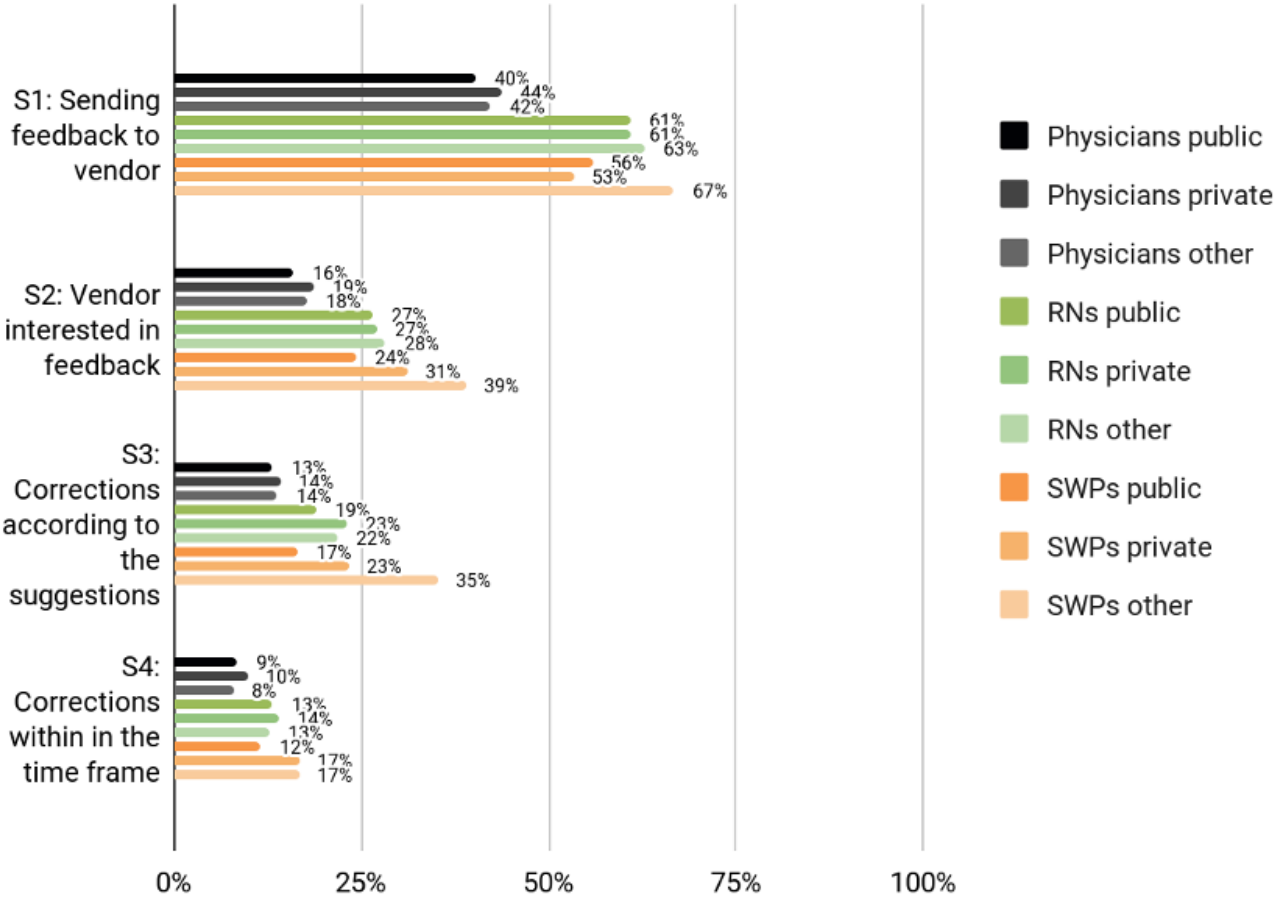
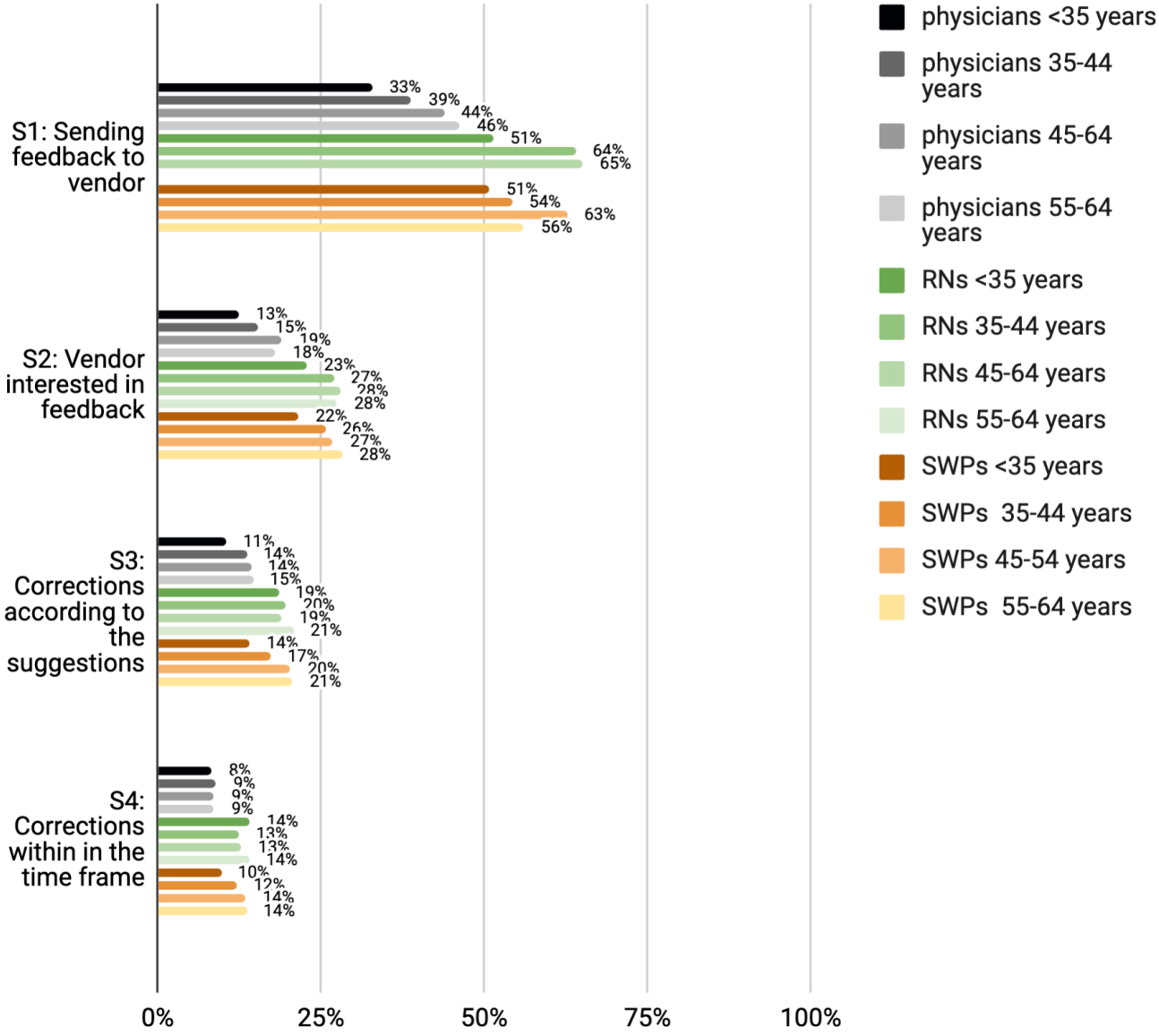


Figure 2. Participation experiences of Finnish physicians, RNs and SWPs by working sector. RN: registered nurse; SWP: social welfare professional.



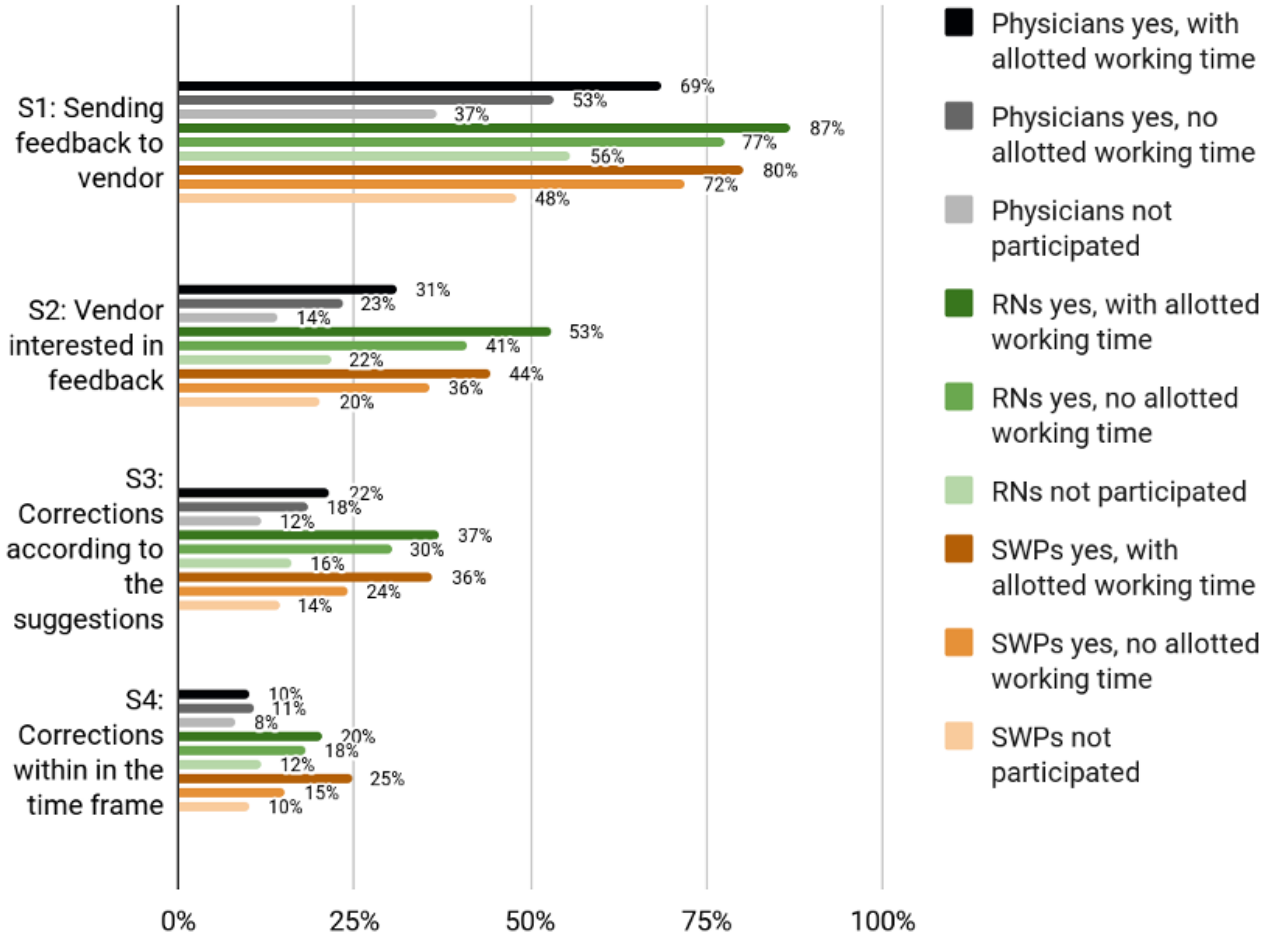
XSL•FO
RenderX

Figure 3. Participation experiences of Finnish physicians, RNs and SWPs by age group. RN: registered nurse; SWP: social welfare professional.



XSL•FO
RenderX

Figure 4. Participation experiences of Finnish physicians, RNs and SWPs by working time allotted for participation. RN: registered nurse; SWP: social welfare professional.



Statistical Analyses

Statistical analyses were carried out with SPSS 28 (IBM Corp). The χ^2 test or Fisher exact test was used as appropriate. Statistical significance was determined as $P<.05$.

Ethical Considerations

According to the national ethical instructions for research, the studies did not require ethical approval (Finnish Advisory Board on Research Integrity 2023) [52].

The autonomy of research subjects was respected, there was informed consent, no harm was possible to the participants and confidentiality of the subjects, and research data were protected. The researchers were not able to identify individual respondents. However, as the data for the RN and SWP studies were collected by a national authority (Finnish Insititute for Health and Welfare), the ethical approval (THL482/6.02.01/2020) was provided by its institutional review board.

Results

Respondent Characteristics

The demographics of the respondents to all 3 surveys are provided in Table 2. In 2021, 4683/19,142 physicians (24.5% of email invitation recipients) participated in the survey, and in 2020, 3610/58,276 RNs (6.2% of email invitation recipients) and 990/12,471 SWPs (7.9% of the theoretical target group) participated in the survey [35-37].

XSL•FO
RenderX

Table . Respondent characteristics.[a]

	Physicians (n=4683), n (%)	Registered nurses (n=3610), n (%)	Social welfare professionals (n=990), n (%)
Working sector			
Public sector	3654 (78)	3076 (85.2)	846 (85.5)
Private	775 (16.5)	456 (12.6)	90 (9.1)
Other	253 (5.4)	78 (2.2)	54 (5.5)
Age group (years)			
<35	949 (20.3)	739 (20.5)	185 (18.7)
35 - 44	1215 (25.9)	833 (23.1)	346 (34.9)
45 - 54	1161 (24.8)	1108 (30.7)	260 (26.3)
55 - 64	1315 (28.1)	921 (25.5)	198 (20)
Leadership position			
Works in a leading or managerial position	1139 (24.3)	406 (11.2)	172 (17.4)
Works in other positions	3543 (75.7)	3204 (88.8)	818 (82.6)

[a]The email invitation was received by 19,142 physicians, 58,276 registered nurses, and 12,471 social welfare professionals. Of these, 4683 (24.5%) physicians, 3,610 (6.2%) registered nurses, and 990 (7.9%) social welfare professionals responded to the survey.

Participation Experiences of the 3 Professional Groups

Physicians appeared to be least knowledgeable of how and where to send feedback, with only 41% (1920/4683) agreeing with statement 1, as compared to 61% (2204/3610) of RNs and 56% (556/990) of SWPs (Figure 5). In terms of vendors' interest in end-user feedback and the manner and speed of IS development, physicians were also least satisfied, with 16% (776/4683), 14% (672/4683), and 9% (407/4683) of them, respectively, agreeing with statements 2, 3, and 4, as compared to 26% (960/3610), 19% (707/3610), and 13% (481/3610), respectively, of RNs and 25% (225/990), 19% (180/990), and 13% (122/990), respectively, of SWPs (Figure 5). In total, 76% (3528/4683) of physicians, 78% (2814/3610) of RNs and 67% (664/990) of SWPs had not participated at all in HIS or CIS development (Figure 6).

XSL•FO
RenderX

Figure 5. Experiences of physicians, RNs and SWPs of collaboration with health care information system and client information system developers. RN: registered nurse; SWP: social welfare professional.

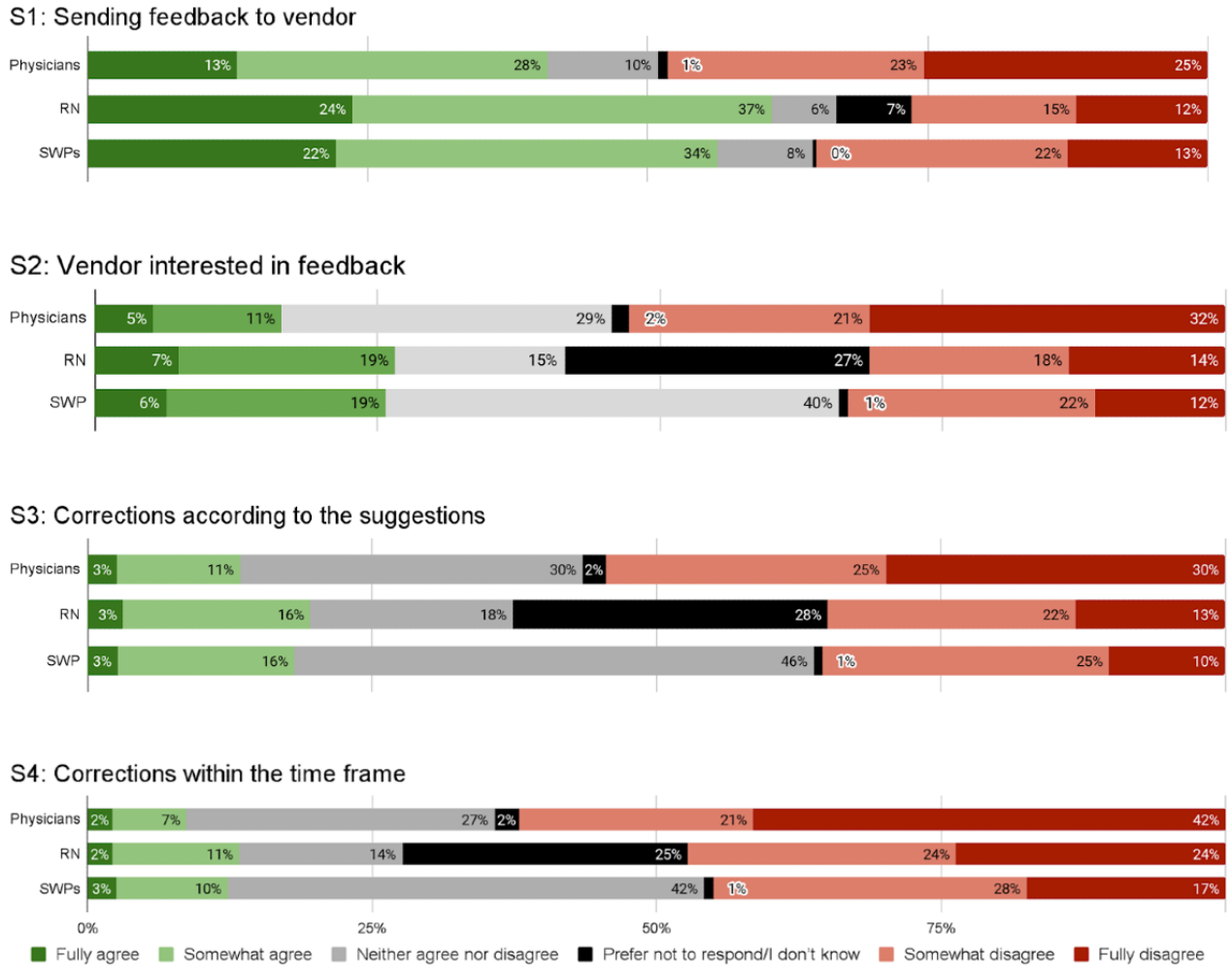
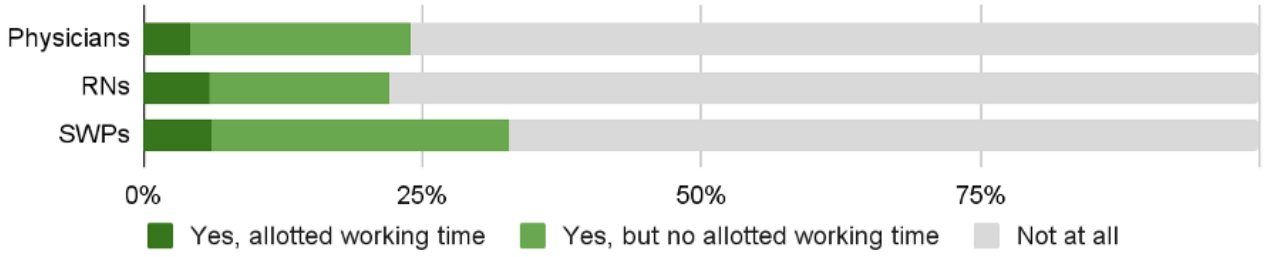


Figure 6. Participation in information system development work by professional group. RN: registered nurse; SWP: social welfare professional.



Factors Associated With Participation Experiences

Leaders, particularly among RNs, were more aware of how and where to send system feedback (statement 1) than those working in nonmanagerial positions (Figures 1-4). Leaders in all 3 professional groups were more satisfied with the system vendor collaboration (statements 2 - 4) than the others. The working sector did not impact end users' experiences. In all professional groups, the youngest appeared least aware of how and where to send feedback and least satisfied with the collaboration. Those who had participated in HIS or CIS development considered IS vendors more interested in feedback and were more satisfied with the manner and speed of system improvements and corrections.

Factors Associated With Having Participated in HIS or CIS Development

In all 3 professional groups, leaders had participated more in IS development than their colleagues in nonleadership positions (Figure 6).

Among physicians and RNs, but not SWPs, those working in the private sector had participated less than their public sector colleagues (Table 3). In all 3 professional groups, the youngest had participated the least (Table 3).

XSL•FO
RenderX

Table . Participation in development by allocated working time.

Participated in development	Yes, allotted working time, n (%)[a]	Yes, but no allotted working time, n (%)[a]	Not at all, n (%)[a]
Leadership position			
Physicians			
Leaders	85 (7.5)	403 (35.7)	640 (56.7)
Others	115 (3.3)	514 (14.6)	2887 (82.1)
RNs[b]			
Leaders	54 (13.3)	131 (32.3)	221 (54.4)
Others	162 (5.1)	449 (14)	2593 (80.9)
SWPs[c]			
Leaders	11 (6.4)	83 (48.3)	78 (45.3)
Others	50 (6.1)	180 (22.1)	586 (71.8)
Working sector			
Physicians			
Public	161 (4.4)	764 (21)	2705 (74.5)
Private	26 (3.4)	104 (13.5)	639 (83.1)
Other	13 (5.3)	49 (20)	183 (74.7)
RNs			
Public	188 (6.1)	500 (16.3)	2388 (77.6)
Private	23 (5)	65 (14.3)	368 (80.7)
Other	5 (6.4)	15 (19.2)	58 (74.4)
SWPs			
Public	53 (6.3)	218 (25.8)	573 (67.9)
Private	6 (6.7)	24 (26.7)	60 (66.7)
Other	2 (3.7)	21 (38.9)	31 (57.4)
Age group (years)			
Physicians			
Age group<35	20 (2.1)	93 (9.9)	827 (88)
Age group 35 - 44	59 (4.9)	227 (18.8)	920 (76.3)
Age group 45 - 54	68 (5.9)	299 (25.9)	787 (68.2)
Age group 55 - 64	52 (4)	289 (22.2)	962 (73.8)
RNs			
Age group<35	34 (4.6)	89 (12)	616 (83.4)
Age group 35 - 44	53 (6.4)	134 (16.1)	646 (77.6)
Age group 45 - 54	83 (7.5)	186 (16.8)	839 (75.7)
Age group 55 - 64	45 (4.9)	166 (18)	710 (77.1)
SWPs			
Age group<35	13 (7.1)	32 (17.5)	138 (75.4)
Age group 35 - 44	23 (6.6)	97 (28)	226 (65.3)
Age group 45 - 54	13 (5)	84 (32.3)	163 (62.7)
Age group 55 - 64	12 (6.1)	50 (25.3)	136 (68.7)

[a]Denominators for calculating percentages are the sum of n values for each row

[b]RN: registered nurse

XSL•FO
RenderX

[c]SWP: social welfare professional

Discussion

Overview

To our knowledge, this is the first study to assess how the major professional groups in the health care and social welfare sector, that is, physicians, RNs, and SWPs, view HIS and CIS development participation. The responses were analyzed by managerial position, employment sector, and age group. Furthermore, we examined which types of professionals have participated in HIS and CIS development.

RNs and SWPs Highly Aware of to Whom and How to Send Development-Related Feedback

The majority of RN (2204/3610, 61%) and SWP (556/990, 56%) respondents knew how and to whom to send development-related feedback; the respective proportion for physicians was 41% (1920/4683). The difference may be explained by mentoring or superuser and training programs during and after HIS implementations among RNs [9,53]. On the other hand, since the response rate among RNs and SWPs was relatively low, those who responded were probably more interested in ISs and thus more aware of HIS and CIS development than those who did not. The findings concur with our earlier studies, which have shown that physicians tend to be more critical towards their HIS compared with RNs and SWPs [2,5].

Leadership and its competence play an important role in the implementation of ISs and investment of resources in digitalization [22,54].

In this study, leaders in all professions and those with allotted working time for HIS or CIS development were more aware of feedback processes than others. These are usually responsible for the orientation of personnel, furthermore, they are often the ones to whom other personnel report development ideas or problems with HIS or CIS use. In all 3 professional groups the youngest were least aware of the feedback processes; whereas the youngest often have mentors who also help with HIS or CIS related problems, it is also probable that the currently available means are not suitable for the younger generations.

Leaders Critical About Cooperation

Although leaders had more often dedicated working time and participated more often in the development of HISs or CISs than the others, they were critical about cooperation with vendors. Their information needs, informatics competencies, and partly also the ISs differ from those working in nonmanagerial positions [55]. The particularly negative viewpoints of physician leaders may be impacted by most of them regularly using HISs for direct patient care, unlike RN or SWP managers who mainly use HISs or CISs for managerial purposes [56]. The most recent study shows that RN managers are able to use HISs for their managerial duties. Due to poor system integration, they need to gather data from different systems for management, which wastes resources inefficiently [22].

Those With Dedicated Working Time Less Dissatisfied With Vendor Cooperation

Those who had dedicated working time for HIS or CIS development were less dissatisfied with vendors' interest and responsiveness to development ideas than those who had not participated at all. They are likely to be more aware of the development processes and timelines of their respective HISs or CISs. Furthermore, since they have been chosen by their respective organizations as participants in HIS or CIS development, their ideas are more likely to become realized. Earlier studies have also found that user participation increases acceptance and active use of HISs contributes to the acceptance and increased active use of HISs [57]. Although not all development ideas are suitable for execution and not all end users can be expected to spend considerable working time on HIS or CIS development, to achieve better engagement in the use of ISs, users need to experience that they are heard and understood [57].

The Youngest Least Satisfied With Vendor Cooperation

Similar to the findings of our previous studies [1,4,6], the youngest were the most dissatisfied with vendors' interest in feedback. This is a particularly important finding as it suggests that the current ways of engaging professionals will not become more suitable or even acceptable to future generations. Barchielli et al [58] also found that younger nurses rely on their colleagues' opinions of health technology use, while older nurses rely on their own experiences.

What Kinds of Users Have Participated in HIS or CIS Development?

Previous studies have shown that impactful participation in IS development requires dedicated working time [58]. Of those working in nonmanagerial positions, 72% - 82% responded that they have not participated at all in HIS or CIS development, whereas the respective proportion for leaders was 45% - 57%. Physician and RN leaders were most likely to have allotted working time for HIS development. Our findings agree with several studies suggesting that managerial viewpoints are likely to become overrepresented in IS development [59,60]. Although the data produced by the ISs is essential for leadership and management purposes, if the participating leaders are not engaged in clinical work or practice the solutions may end up not supporting the needs of frontline workers [61,62].

Health care professionals working in the private sector participated less than their public sector colleagues, among SWPs the differences were minimal. As the majority of Finnish private sector physicians work as private practitioners, their participation would usually result in decreased earnings. It is also likely that the lack of most complex patients in private healthcare reduces the need for HIS development.

In all 3 professional groups, the youngest participated the least. This may be because they are at the stage of learning the clinical content of their work and their employers may not want to invest their time in HIS or CIS development. Khairat et al [62] also

XSL•FO
RenderX

report underrepresentation of physicians in specialization training in HIS development groups. Although not possessing advanced professional skills, the youngest are not burdened with old, often paper-based workflows, which could assist in redesigning processes and enhance the use of newer technologies [62].

How to Improve Satisfaction in HIS and CIS Development

The development processes of large-scale complex systems such as HIS and CIS are typically dominated by cooperative activities involving multiple stakeholders [63-65]. Different information needs must be identified, prioritized, and communicated clearly enough to the IS designers and developers [65].

It remains challenging to increase user input and select appropriate participants and human-centered design methods through the different phases of the participatory development cycle [15,20]. Identifying other factors that influence user experiences, such as decisions made by regulators, policymakers, and administrators, may assist in developing better HISs and CISs [66]. Previous studies have recognized the importance of clinical informaticists who also use HISs or CISs in clinical work in communicating end users' needs and feedback to designers and developers [6,67-75]. From the organizational perspective, the benefits can be seen beyond the IS implementation phase [76]: informatics competent social and health care professionals have been found to be able to improve patient safety and patient care outcomes [67]. In social welfare, however, this role is still being developed [74,75].

Limitations

Our study has some limitations. First, compared with physicians, the lower response rates among RNs and SWPs are likely to have resulted in the selection of more involved and interested participants among these professional groups.

Second, the response rates were highest for physicians. The FMA, which was responsible for collecting the physician data, has a long history of conducting surveys among physicians, the results of which are used in FMA policies. However, the number of participating professionals from all 3 professional groups in the national level studies can be considered high compared to other similar studies [50].

The RNs questionnaire respondents, comprising nurses from various sectors including hospitals, health centers, private practice, and social care, were found to align with the target population according to Statistics Finland's employment data for nurses, midwives, and community nurses [9].

During data collection with SWPs questionnaire, incomplete contact information in membership registries limited survey outreach. To compensate, survey invitations were distributed also through sector networks and social media. The study's final sample comprised 990 SWP respondents, with an estimated response rate of 8%. It's important to note this limitation when interpreting findings. Nonetheless, the sample size was considerable and exhibited diversity in age, service backgrounds, and geographical representation across Finland [51].

Third, our questionnaire did not cover how end users were involved in HIS or CIS development and whether the means of participation would impact satisfaction with the process and the end results. Further research is needed on best practices of user participation in the development of complex HIS and CIS systems.

Conclusion

The fluent use of HISs and CISs is a prerequisite for efficient and safe health care and social welfare, as currently professionals spend a considerable part of their working time with ISs and rely on them as their primary source of information. User participation of all major professional groups—physicians, RNs and SWPs—and their involvement in development are essential for the success of complex HIS or CIS. Compared with RNs and SWPs, physicians appeared to be more critical towards IS vendors and the success of participatory HIS development. As even those with allotted working time were mostly dissatisfied with vendor cooperation, it is evident that simply allocating more end users' working time for HIS and CIS development will not guarantee satisfaction; rather, dialogue between end users and developers needs improvement. New means are needed to better engage all end-user groups, particularly the youngest ones and those working in nonmanagerial positions.

Acknowledgments

This study was part of larger projects supported by the Strategic Research Council of the Academy of Finland (projects 303607 and 327145) and the Ministry of Social Affairs and Health (project 112241). We thank the Finnish Medical Association, the University of Oulu, University of Eastern Finland, the Union of Health and Social Care Professionals in Finland, the Finnish Nurses Association, the National Association of Health and Welfare Professionals, the National Professional Association, the Union of Professional Social Workers, the Trade Union for the Public and Welfare Sectors, and the Social Science Professionals Union for their collaboration. We would also like to thank the respondents to the questionnaires. According to the local and national ethical instructions for research (Finnish Advisory Board on Research Integrity) instructions, this study did not require ethical approval.

Authors' Contributions

SM was responsible for the main supervision, writing and visualization of the research output. TL was mainly responsibile for the statistical analysis. All authors contributed to the analysis and writing of the research output. In addition, SM, JV, and TL

XSL•FO
RenderX

contributed to the design of methodology of all 3 studies, SS to the design, data curation and investigation of the SWP study, and UMK to the design of the registered nurse study.

Conflicts of Interest

SM has been previously employed by 2 of the health care information system (HIS) and client information system (CIS) software vendors included in the study. SS has been previously employed by a publicly owned in-house HIS and CIS vendor included in the study. TL has been previously employed by a publicly owned in-house HIS and CIS software vendor included in the study. The employers did not provide any support, financial or otherwise, to the study. Furthermore, these vendors were not involved in the design of the study or in the collection, analysis or interpretation of the data.

Multimedia Appendix 1
Experiences of participation in system development.
[DOCX File, 20 KB - humanfactors_v12i1e51495_app1.docx]

References

1. Martikainen S, Kaipio J, Lääveri T. End-user participation in health information systems (HIS) development: physicians' and nurses' experiences. Int J Med Inform 2020 May;137:104117. [doi: 10.1016/j.ijmedinf.2020.104117] [Medline: 32179254]
2. Kaipio J, Kuusisto A, Hyppönen H, Heponiemi T, Lääveri T. Physicians' and nurses' experiences on EHR usability: comparison between the professional groups by employment sector and system brand. Int J Med Inform 2020 Feb;134:104018. [doi: 10.1016/j.ijmedinf.2019.104018] [Medline: 31835158]
3. Kaipio J, Lääveri T, Hyppönen H, et al. Usability problems do not heal by themselves: national survey on physicians' experiences with EHRs in Finland. Int J Med Inform 2017 Jan;97:266-281. [doi: 10.1016/j.ijmedinf.2016.10.010] [Medline: 27919385]
4. Martikainen S, Viitanen J, Korpela M, Lääveri T. Physicians' experiences of participation in healthcare IT development in Finland: willing but not able. Int J Med Inform 2012 Feb;81(2):98-113. [doi: 10.1016/j.ijmedinf.2011.08.014] [Medline: 21956004]
5. Viitanen J, Hyppönen H, Lääveri T, Vänskä J, Reponen J, Winblad I. National questionnaire study on clinical ICT systems proofs: physicians suffer from poor usability. Int J Med Inform 2011 Oct;80(10):708-725. [doi: 10.1016/j.ijmedinf.2011.06.010] [Medline: 21784701]
6. Martikainen S, Salovaara S, Ylönen K, et al. Social welfare professionals willing to participate in client information system development - results from a large cross-sectional survey. Inform Health Soc Care 2022 Oct 2;47(4):389-402. [doi: 10.1080/17538157.2021.2010736] [Medline: 34877899]
7. Khairat S, Coleman C, Ottmar P, Jayachander DI, Bice T, Carson SS. Association of electronic health record use with physician fatigue and efficiency. JAMA Netw Open 2020 Jun 1;3(6):e207385. [doi: 10.1001/jamanetworkopen.2020.7385] [Medline: 32515799]
8. Kinnunen UM, Hyppönen H, Liljamo P, Saranto K. Nurses' experiences of health and social care information systems. In: Vehko T, Ruotsalainen S, Hyppönen H, editors. E-Health and E-Welfare of Finland: National Institute for Health and Welfare (THL); 2019:130-147.
9. Kinnunen UM, Heponiemi T, Rajalahti E, Ahonen O, Korhonen T, Hyppönen H. Factors related to health informatics competencies for nurses-results of a national electronic health record survey. Comput Inform Nurs 2019 Aug;37(8):420-429. [doi: 10.1097/CIN.0000000000000511] [Medline: 30741730]
10. Topaz M, Ronquillo C, Peltonen LM, et al. Nurse informaticians report low satisfaction and multi-level concerns with electronic health records: results from an international survey. AMIA Annu Symp Proc 2016;2016:2016-2025. [Medline: 28269961]
11. Melnick ER, West CP, Nath B, et al. The association between perceived electronic health record usability and professional burnout among US nurses. J Am Med Inform Assoc 2021 Jul 30;28(8):1632-1641. [doi: 10.1093/jamia/ocab059] [Medline: 33871018]
12. Barki H, Hartwick J. Measuring user participation, user involvement, and user attitude. MIS Q 1994 Mar;18(1):59. [doi: 10.2307/249610]
13. Kujala S. User involvement: a review of the benefits and challenges. Behav Inf Technol 2003 Jan;22(1):1-16. [doi: 10.1080/01449290301782]
14. Lin WT, Shao BBM. The relationship between user participation and system success: a simultaneous contingency approach. Inf & Manag 2000 Sep;37(6):283-295. [doi: 10.1016/S0378-7206(99)00055-5]
15. Kushniruk A, Nøhr C. Participatory design, user involvement and health IT evaluation. Stud Health Technol Inform 2016;222:139-151. [Medline: 27198099]
16. Bano M, Zowghi D. A systematic review on the relationship between user involvement and system success. Inf Softw Technol 2015 Feb;58:148-169. [doi: 10.1016/j.infsof.2014.06.011]

XSL•FO
RenderX

17. Bano M, Zowghi D, da Rimini F. User satisfaction and system success: an empirical exploration of user involvement in software development. Empir Softw Eng 2017 Oct;22(5):2339-2372. [doi: 10.1007/s10664-016-9465-1]
18. Johnson L. Effective organizational communication for large-scale healthcare information technology initiatives. In: McCormick KA, Gugerty B, Mattison JE, editors. Healthcare Information Technology Exam Guide for CHTS and CAHIMS Certification: McGraw Hill Education; 2018.
19. Rahimi B, Karimian S, Ghaznavi A, Heydarlou MJ. Development and evaluation of dental image exchange and management system: a user-centered perspective. arXiv. Preprint posted online on Jun 4, 2022. [doi: arXiv:2206.01966]
20. Gillingham P. Electronic information systems in human service organisations: the what, who, why and how of information. Br J Soc Work 2015 Jul;45(5):1598-1613. [doi: 10.1093/bjsw/bcu030] [Medline: 27559213]
21. Gillingham P. Electronic information systems and human services organisations: avoiding the pitfalls of participatory design. Br J Soc Work 2015 Mar 1;45(2):651-666. [doi: 10.1093/bjsw/bct126] [Medline: 27559213]
22. Saranto K, Koponen S, Vehko T, Kivekäs E. Nurse managers' opinions of information system support for performance management: a correlational study. Methods Inf Med 2023 Jun;62(S 01):e63-e72. [doi: 10.1055/a-1978-9727] [Medline: 36379471]
23. Sittig DF, Wright A, Coiera E, et al. Current challenges in health information technology-related patient safety. Health Informatics J 2020 Mar;26(1):181-189. [doi: 10.1177/1460458218814893] [Medline: 30537881]
24. Gillingham P. Developments in electronic information systems in social welfare agencies: from simple to complex. Br J Soc Work 2019 Jan 1;49(1):135-146. [doi: 10.1093/bjsw/bcy014]
25. Vilpponen H, Grundström M, Abrahamsson P, University of Jyväskylä, Finland and City of Kerava, Finland. Combining social service and healthcare as the first country in the world: exploring the impacts on information systems. J Adv Infor Tech 2018;9(4):84-88. [doi: 10.12720/jait.9.4.84-88]
26. Atherton IM, Lynch E, Williams AJ, Witham MD. Barriers and solutions to linking and using health and social care data in Scotland. Br J Soc Work 2015 Jul;45(5):1614-1622. [doi: 10.1093/bjsw/bcv047]
27. Vehko T, Ruotsalainen S, Hyppönen H, editors. E-health and e-welfare of Finland. Check point 2018. Julkari.: National Institute for Health and Welfare (THL); 2019. URL: https://www.julkari.fi/bitstream/handle/10024/138244/RAP2019_7_e-health_and_e-welfare_web_4.pdf?sequence=4&isAllowed=y [accessed 2025-01-06]
28. Hunt KA, Weber EJ, Showstack JA, Colby DC, Callaham ML. Characteristics of frequent users of emergency departments. Ann Emerg Med 2006 Jul;48(1):1-8. [doi: 10.1016/j.annemergmed.2005.12.030] [Medline: 16781914]
29. Smits FTM, Mohrs JJ, Beem EE, Bindels PJE, van Weert HCPM. Defining frequent attendance in general practice. BMC Fam Pract 2008 Apr 15;9(21):21. [doi: 10.1186/1471-2296-9-21] [Medline: 18412954]
30. Vedsted P, Christensen MB. Frequent attenders in general practice care: a literature review with special reference to methodological considerations. Public Health 2005 Feb;119(2):118-137. [doi: 10.1016/j.puhe.2004.03.007] [Medline: 15694959]
31. Social determinants of health. World Health Organization. 2022. URL: https://www.who.int/health-topics/social-determinants-of-health#tab=tab_1 [accessed 2025-01-06]
32. Jørgensen JT, Andersen JS, Tjønneland A, Andersen ZJ. Determinants of frequent attendance in Danish general practice: a cohort-based cross-sectional study. BMC Fam Pract 2016 Jan 28;17(9):9. [doi: 10.1186/s12875-016-0412-4] [Medline: 26821807]
33. Report to Congress; social risk factors and performance under medicare's value-based purchasing programs. Office of the Assistant Secretary for Planning and Evaluation. 2016. URL: https://aspe.hhs.gov/reports/report-congress-social-risk-factors-performance-under-medicares-value-based-purchasing-programs [accessed 2025-01-06]
34. Social welfare and health care system in Finland, responsibilities. Ministry of Social Affairs and Health. 2019. URL: https://stm.fi/en/social-and-health-services/responsible-agencies [accessed 2025-01-06]
35. Lääkärityövoima [Web page in Finnish]. Finnish Medical Association. 2022. URL: https://www.laakariliitto.fi/laakariliitto/tutkimus/laakarityovoima/ [accessed 2025-01-06]
36. Terveys- ja sosiaalipalvelujen henkilöstö 2021 [Web page in Finnish]. Finnish Institute for Health and Welfare. URL: https://thl.fi/fi/tilastot-ja-data/tilastot-aiheittain/sosiaali-ja-terveydenhuollon-resurssit/terveys-ja-sosiaalipalvelujen-henkilosto [accessed 2025-01-06]
37. Terveys- ja sosiaalipalvelujen henkilöstö 2018: tuleva sote-uudistus koskee 7,3 prosenttia työssäkäyvistä [Web page in Finnish]. Julkari. 2021. URL: https://urn.fi/URN:NBN:fi-fe2021121460373 [accessed 2025-01-06]
38. Non-federal acute care hospital electronic health record adoption. Office of the National Coordinator for Health Information Technology. 2017. URL: https://www.healthit.gov/data/quickstats/non-federal-acute-care-hospital-electronic-health-record-adoption [accessed 2025-01-06]
39. Kuusisto-Niemi S, Hyppönen H, Ruotsalainen S. Availability and use of e-welfare in finland. In: Vehko T, Ruotsalainen S, Hyppönen H, editors. E-Health and E-Welfare of Finland Check Point 2018: National Institute for Health and Welfare (THL); 2019. URL: https://www.julkari.fi/bitstream/handle/10024/138244/RAP2019_7_e-health_and_e-welfare_web_4.pdf?sequence=4&isAllowed=y [accessed 2025-01-06]
40. Kärki J, Ryhänen M. Use of Information and Communications Technology in Social Services in 2014: National Institute for Health and Welfare (THL); 2015.

XSL•FO
RenderX

41. Salovaara S, Silén M, Vehko T, Kyytsönen M, Hautala S. Use of Information and Communication Technology in Social Welfare Services in 2020: National Institute for Health and Welfare (THL); 2021.
42. Citizens. Kanta. 2022. URL: https://www.kanta.fi/en/citizens [accessed 2025-01-06]
43. Jormanainen V. Large-scale implementation and adoption of the Finnish national Kanta services in 2010–2017: a prospective, longitudinal, indicator-based study. FinJeHeW 2018;10(4):381-395. [doi: 10.23996/fjhw.74511]
44. Client data repository for social welfare services. Kanta. 2022. URL: https://www.kanta.fi/en/professionals/client-data-archive-for-social-welfare-services [accessed 2025-01-06]
45. Jormanainen V, Reponen J. CAF and CAMM analyses on the first 10 years of national Kanta services in Finland. FinJeHeW 2020;12(4):302-315. [doi: 10.23996/fjhw.98548]
46. Jormanainen V, Rötsä M, Parhiala K. Social care client information systems in Finnish municipality social services in 2017. Finn J EHealth E Welfare 2019;11(1-2):125-138. [doi: 10.23996/fjhw.76835]
47. Implementation of the monitoring of information system services in social welfare and healthcare. Finnish Institute for Health and Welfare. 2023. URL: https://thl.fi/en/topics/information-management-in-social-welfare-and-health-care/information-management-in-finland/follow-up-of-the-information-system-services-in-social-welfare-and-health-care/implementation [accessed 2025-01-06]
48. Hyppönen H, Kaipio J, Heponiemi T, et al. Developing the National Usability-Focused Health Information System Scale for physicians: validation study. J Med Internet Res 2019 May 16;21(5):e12875. [doi: 10.2196/12875] [Medline: 31099336]
49. Kaihlanen AM, Gluschkoff K, Saranto K, Kinnunen UM. The associations of information system's support and nurses' documentation competence with the detection of documentation-related errors: Results from a nationwide survey. Health Informatics J 2021;27(4):14604582211054026. [doi: 10.1177/14604582211054026] [Medline: 34814758]
50. Viitanen J, Reponen J, Lääveri T, et al. Potilastietojärjestelmissä on yhä paljon parannettavaa: lääkärien kokemukset käytettävyydestä vuosilta 2010-2021. Usability of electronic health record systems needs improvements – physicians' experiences from 2010-21. Finn Med J 2024(41):79 [FREE Full text]
51. Salovaara S, Hautala S, Silén M. Sosiaalialan ammattilaisten kyvykkyys tietojärjestelmien käyttäjinä ja kokemukset tietojärjestelmien tuesta työlle [Article in Finnish]. Fin Je He W 2022;14(2):208-225. [doi: 10.23996/fjhw.110119]
52. Finnish Advisory Board on Research Integrity (TENK). URL: https://tenk.fi/en/tenk [accessed 2025-01-13]
53. Rhodes B, Short A, Shaben T. Effectiveness of training strategies that support informatics competency development in healthcare professionals. In: Shachak A, Borycki EM, Reis SP, editors. Health Professionals' Education in the Age of Clinical Information Systems, Mobile Computing and Social Networks: Academic Press; 2017:299-322. [doi: 10.1016/B978-0-12-805362-1.00015-2]
54. Booth RG, Strudwick G, McBride S, O'Connor S, Solano López AL. How the nursing profession should adapt for a digital future. BMJ 2021;373:1190. [doi: 10.1136/bmj.n1190]
55. Hübner U, Shaw T, Thye J, et al. Technology Informatics Guiding Education Reform – TIGER: an international recommendation framework of core competencies in health informatics for nurses. Methods Inf Med 2018;57(Suppl 1):e30-e42. [doi: 10.3414/ME17-01-0155] [Medline: 29956297]
56. Backonja U, Langford LH, Mook PJ. How to support the nursing informatics leadership pipeline: recommendations for nurse leaders and professional organizations. Comput Inform Nurs 2022 Jan 1;40(1):8-20. [doi: 10.1097/CIN.0000000000000827] [Medline: 34996883]
57. Vahteristo A, Jylhä V. Effects of user participation in the development of health information systems on their evaluation within occupational health services. In: Integrated Citizen Centered Digital Health and Social Care: IOS Press; 2020:207-211. [doi: 10.3233/SHTI200724]
58. Barchielli C, Marullo C, Bonciani M, Vainieri M. Nurses and the acceptance of innovations in technology-intensive contexts: the need for tailored management strategies. BMC Health Serv Res 2021 Jul 3;21(1):639. [doi: 10.1186/s12913-021-06628-5] [Medline: 34215228]
59. Gillingham P. From bureaucracy to technocracy in a social welfare agency: a cautionary tale. Asia Pac J Soc Work Dev 2019 Apr 3;29(2):108-119. [doi: 10.1080/02185385.2018.1523023]
60. Devlieghere J, Roose R. Documenting practices in human service organisations through information systems: when the quest for visibility ends in darkness. Soc Inclu 2019;7(1):207-217. [doi: 10.17645/si.v7i1.1833]
61. Høstgaard AM, Bertelsen P, Nøhr C. Methods to identify, study and understand end-user participation in HIT development. BMC Med Inform Decis Mak 2011 Sep 28;11(1):1-11. [doi: 10.1186/1472-6947-11-57] [Medline: 21955493]
62. Khairat S, Burke G, Archambault H, Schwartz T, Larson J, Ratwani RM. Focus section on health IT usability: perceived burden of EHRs on physicians at different stages of their career. Appl Clin Inform 2018 Apr;09(02):336-347. [doi: 10.1055/s-0038-1648222]
63. Saleem N, Steel D, Gercek G, Chandra A. Significance of user participation in a hospital information system success. Health Care Manag 2017;36(2):199-205. [doi: 10.1097/HCM.0000000000000156]
64. Sommerville I, Cliff D, Calinescu R, et al. Large-scale complex IT systems. Commun ACM 2012 Jul;55(7):71-77. [doi: 10.1145/2209249.2209268]
65. Dalsgaard P, Frich J, Riisgaard T. Four challenges of designing digital services to support participatory design. Presented at: Nordic Human-Computer Interaction Conference; Oct 8, 2022; Aarhus Denmark. [doi: 10.1145/3546155.3547273]

XSL•FO
RenderX

66. Tutty MA, Carlasare LE, Lloyd S, Sinsky CA. The complex case of EHRs: examining the factors impacting the EHR user experience. J Am Med Inform Assoc 2019 Jul 1;26(7):673-677. [doi: 10.1093/jamia/ocz021] [Medline: 30938754]
67. Lagsten J, Andersson A. Use of information systems in social work – challenges and an agenda for future research. Eur J Soc Work 2018 Nov 2;21(6):850-862. [doi: 10.1080/13691457.2018.1423554]
68. AMIA's role in CIS and history. American Medical Informatics Association. URL: https://amia.org/careers-certifications/clinical-informatics-subspecialty/amias-role-cis-and-history [accessed 2025-01-13]
69. Singer JS, Cheng EM, Baldwin K, Pfeffer MA, UCLA Health Physician Informaticist Committee. The UCLA Health Resident Informaticist Program - a novel clinical informatics training program. J Am Med Inform Assoc 2017 Jul 1;24(4):832-840. [doi: 10.1093/jamia/ocw174] [Medline: 28115427]
70. Sieja A, Markley K, Pell J, et al. Optimization sprints: improving clinician satisfaction and teamwork by rapidly reducing electronic health record burden. Mayo Clin Proc 2019 May;94(5):793-802. [doi: 10.1016/j.mayocp.2018.08.036] [Medline: 30824281]
71. Lehmann CU, Gundlapalli AV, Williamson JJ, et al. Five years of clinical informatics board certification for physicians in the United States of America. Yearb Med Inform 2018 Aug;27(1):237-242. [doi: 10.1055/s-0038-1641198] [Medline: 29681038]
72. Cummins MR, Gundlapalli AV, Murray P, Park HA, Lehmann CU. Nursing informatics certification worldwide: history, pathway, roles, and motivation. Yearb Med Inform 2016 Nov 10;25(1):264-271. [doi: 10.15265/IY-2016-039] [Medline: 27830261]
73. Conrad JB, Magsamen‐Conrad K. Understanding the impact of the coronavirus pandemic on families involved in the child welfare system: technological capital and pandemic practice. Child Fam Soc Work 2022 Feb;27(1):11-21. [doi: 10.1111/cfs.12876]
74. Parker-Oliver D, Demiris G. Social work informatics: a new specialty. Soc Work (Stell) 2006 Apr 1;51(2):127-134. [doi: 10.1093/sw/51.2.127]
75. Kuusisto-Niemi S, Saranto K, Rissanen S. Informatics in social services: research, developments, and outcomes from the finnish perspective. In: Tavana M, Ghapanchi AH, Talaei-Khoei A, editors. Healthcare Informatics and Analytics Emerging Issues and Trends: IGI Global; 2015:1-26. [doi: 10.4018/978-1-4666-6316-9.ch002]
76. HIMSS Nursing Informatics Workforce Survey. HIMSS. 2020. URL: https://www.himss.org/resources/himss-nursing-informatics-workforce-survey [accessed 2025-01-06]

Abbreviations

CIS: client information system
FMA: Finnish Medical Association
HIS: health care information system
IS: information system
RN: registered nurse
SWP: social welfare professional

Edited by A Kushniruk; submitted 02.08.23; peer-reviewed by A Sieja, W Tesfaye; revised version received 02.05.24; accepted 08.11.24; published 22.01.25.

Please cite as:
Martikainen S, Viitanen J, Salovaara S, Kinnunen UM, Lääveri T
Comparisons of Physicians', Nurses', and Social Welfare Professionals' Experiences With Participation in Information System Development: Cross-Sectional Survey Study
JMIR Hum Factors 2025;12:e51495
URL: https://humanfactors.jmir.org/2025/1/e51495
doi:10.2196/51495

© Susanna Martikainen, Johanna Viitanen, Samuel Salovaara, Ulla-Mari Kinnunen, Tinja Lääveri. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 22.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Patient-Generated Collections for Organizing Electronic Health Record Data to Elevate Personal Meaning, Improve Actionability, and Support Patient–Health Care Provider Communication: Think-Aloud Evaluation Study

Drashko Nakikj[1], PhD; David Kreda[1], BA; Karan Luthria[1], BCS; Nils Gehlenborg[1], PhD

Department of Biomedical Informatics, Harvard Medical School, Boston, MA, United States

Corresponding Author:
Nils Gehlenborg, PhD
Department of Biomedical Informatics
Harvard Medical School
10 Shattuck Street
Suite 514
Boston, MA, 02115
United States
Phone: 1 8572725075
Email: nils@hms.harvard.edu

Abstract

Background: Through third party applications, patients in the United States have access to their electronic health record (EHR) data from multiple health care providers. However, these applications offer only a predefined organization of these records by type, time stamp, or provider, leaving out meaningful connections between them. This prevents patients from efficiently reviewing, exploring, and making sense of their EHR data based on current or ongoing health issues. The lack of personalized organization and important connections can limit patients' ability to use their data and make informed health decisions.

Objective: To address these challenges, we created Discovery, an experimental app that enables patients to organize their medical records into collections, analogous to placing pictures in photo albums. These collections are based on the evolving understanding of the patients' past and ongoing health issues. The app also allows patients to add text notes to collections and their constituent records. By observing how patients used features to select records and assemble them into collections, our goal was to learn about their preferred mechanisms to complete these tasks and the challenges they would face in the wild. We also intended to become more informed about the various ways in which patients could and would like to use collections.

Methods: We conducted a think-aloud evaluation study with 14 participants on synthetic data. In session 1, we obtained feedback on the mechanics for creating and assembling collections and adding notes. In session 2, we focused on reviewing collections, finding data patterns within them, and retaining insights, as well as exploring use cases. We conducted reflexive thematic analysis on the transcribed feedback.

Results: Collections were useful for personal use (quick access to information, reflection on medical history, tracking health, journaling, and learning from past experiences) and clinical visits (preparation and raising physicians' awareness). Assembling EHR data into reliable collections could be difficult for typical patients due to considerable manual work and lack of medical knowledge. However, automated collection building could alleviate this issue. Furthermore, having EHR data organized in collections may have limited use. However, augmenting them with patient-generated data, which are entered with flexible richness and structure, could add context, elevate meaning, and improve actionability. Finally, collections might produce a misconstrued health picture, but bringing the physician in the loop for verification could increase their clinical validity.

Conclusions: Collections can be a powerful tool for advancing patients' proactivity, awareness, and self-advocacy, potentially facilitating patient-centered care. However, patients need better support for incorporating their own everyday data and adding meaningful annotations for future reference. Improvements in the comprehensiveness, efficiency, and reliability of the collection assembly process through automation are also necessary.

(JMIR Hum Factors 2025;12:e50331) doi:10.2196/50331

XSL•FO
RenderX

KEYWORDS

mobile health; patients; electronic health records; sensemaking; data organization; collections; awareness; proactivity; self-advocacy; patient–health care provider communication

Introduction

Background

Digital technologies play a pivotal role in facilitating patient-centered care that focuses on understanding patients' needs and fostering shared decision-making with their health care providers [1]. However, this approach relies heavily on well-informed patients and on effective patient-provider communication [2]. In the United States, a major step toward meeting these requirements was enabling patients to access their medical records from multiple providers through third-party applications [3,4]. While this is a significant achievement, enabling patients to engage in making sense of these data and turn the insights they gain from this process into actionable steps remains a challenge. Consequently, many patients lack a satisfactory understanding of their health, feel discouraged to self-advocate, and have mediocre communication with their providers, which is at odds with the core values of patient-centered care. Therefore, addressing the challenges around sensemaking and the usability of health data will be important to advancing patient-centered care and empowering patients to take an active role in their health journey.

Making sense of data, or *sensemaking*, is a cyclic process that involves cognitive activities for answering complex questions [5]. These activities involve repeated access to artifacts, identifying relevant information, finding information relationships, and presenting the answers in an understandable format [6]. Patients face a plethora of sensemaking challenges to manage their health. They need to assemble health information from different providers and identify outliers, correlations, and trends to become educated on health topics, drive decision-making, and formulate discussion points with health care providers [7]. Unfortunately, robust platforms to support patients in making sense of their clinical data are lacking.

Several commercial mobile apps such as Apple Health Records [8], iBlueButton [9], OneRecord [10], and 1upHealth [11], along with the academic web application Discovery [12], advance patient sensemaking by offering data visualizations and specialized views for data exploration. These views help patients uncover interesting patterns related to prevalence, periodicity, co-occurrence, and pre-post analysis of medical events. Typically, these solutions organize electronic health record (EHR) data by type, time stamp, or provider. However, such approaches provide very little support for patients in finding deeper connections between their medical records that are required for understanding health issues, reflecting on medical history, or preparing for clinical encounters.

Moreover, these apps do not allow patients to annotate their medical records or save their sensemaking progress, forcing them to remember findings or record them elsewhere. This necessitates patients to revisit the same data repeatedly to refresh inferences and recreate mental notes. Such work is typically tedious and frustrating, leading to anxiety and missing crucial information. Consequently, patients can form skewed health impressions, resulting in poor decisions or risky actions. In clinical visits, the inability to communicate the sensemaking insights to physicians may hinder optimal treatment, leading to repeated tests or medical errors.

To address these limitations, we explored an alternative solution that organizes EHR data into collections based on health issues and ongoing problems [13]. Inspired by findings that a problem-based view of EHR data improves clinician awareness, prioritization, and decision-making in the intensive care unit [14], we adapted a similar approach for patients anchored in the *data-frame* sensemaking theory [15]. Data frames (ie, structured mental models) or *collections* of health data, as we refer to them in our previous work [13], systematically break down problems and help in answering complex questions. These capabilities of the collections hold significant potential for managing health data effectively for all sorts of patients, particularly those with complex medical histories or multiple comorbidities [13]. By organizing abundant data around health issues, collections help patients avoid fragmented health impressions, a common challenge for those with multiple comorbidities. Patients who see multiple specialists can use collections to track the development of specific issues and share insights across providers, raising awareness and improving care coordination.

More precisely, our proposed concept of *collections* allows patients to dynamically organize, adapt, and explore their health information based on evolving needs and available data. For example, a patient managing cardiovascular issues might create a Blood Pressure collection to consolidate related records, which could later branch into more specific collections such as Extreme Blood Pressures or Blood Pressure Lab Work. These refined groupings help uncover patterns and dependencies among factors such as BMI, diet, cholesterol levels, and blood pressure, enhancing understanding and facilitating proactive management of health conditions. Gathering insights from the collections may also help patients have more productive discussions with their providers.

In our study, we extended the capacity to transform EHR data into collections and facilitated reasoning regarding them. As patients' sensemaking of health data is driven by finding outliers, correlations, and trends [7,16], we enabled capabilities to identify data patterns within the collections. In addition, we supported the assembly of relevant medical records for the collections by helping patients visually explore, find temporal patterns [17-20], and make sense of their EHR data within a single, context-preserving view [12]. Acknowledging patient requests for automation [13], we offered manually assembled and system-assembled collections. We also allowed for personal data input through free-text notes, fulfilling previously identified patient needs [13].

A notable advancement in our work is moving from mock-ups [13] to a fully functional mobile app, Discovery. The app

XSL•FO
RenderX

provides patients with a realistic platform to interact with collections, uncovering deeper insights into preferred mechanisms for creating, refining, and using these collections. Moreover, we intended to motivate patients to see collections as tools for self-advocacy during clinical visits, which was identified as a key use case in our earlier research [13].

Objectives

In this study, we used the Discovery app to explore patients' needs, preferences, and desired interactions for organizing EHR data and delve into potential use cases. More concretely, we asked the following research questions (RQs):

- What are the needs and feature preferences for organizing EHR data from multiple providers? (RQ 1)
  - What are the patients' experiences with creating and building collections?
  - How effectively does our app allow patients to organize their EHR data using the concept of collections?
  - How can we better meet patients' needs for meaningful EHR data organization?
- What purpose would organizing the EHR data in collections have? (RQ 2)

To answer these RQs, we conducted a qualitative evaluation study with 14 participants.

Methods

Description of Discovery

General Overview

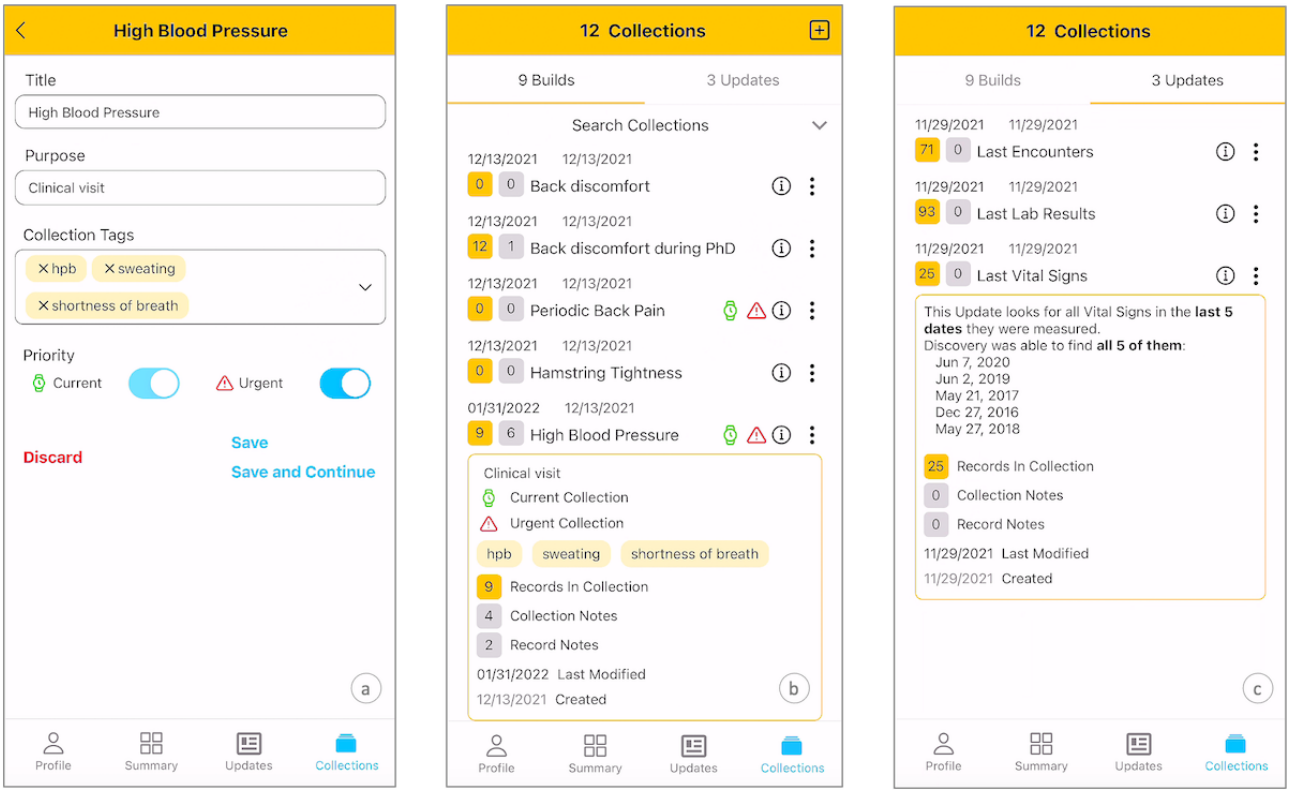
Discovery is a noncommercial iPhone app designed to help patients make sense of their EHR data. It introduces a novel concept that organizes EHR data into personalized, problem-based collections. In addition, it allows patients to add their own observations and insights, providing context and complementing their medical records with patient-generated data (PGD).

For this study, the app was restricted to accessing synthetic EHR data through a Fast Healthcare Interoperability Resources (FHIR) format [21] from the SMART Health IT repository [22]. Discovery accesses only structured EHR data and relies on a 2-level hierarchy. At the highest level, there are the record categories ("Conditions," "Immunizations," and "Vital Signs"). Each record category has multiple record types ("Vital Signs: Body Height," "Body Weight," and "Blood Pressure"), and each record type can have multiple instances ("Blood Pressure: systolic: 125, diastolic: 90," and "date: 02/01/2022"). Each instance of data in our app is called a record and corresponds to an FHIR resource with standardized attributes.

Organizing Records in Collections

The manually created collections are called Builds (creation shown in Figure 1A and list shown in Figure 1B), whereas the app-assembled collections are called Updates (Figure 1C).

Figure 1. (A) Creation of a new collection; (B) list of patient-built collections—Builds; (C) list of system-generated collections—Updates.



Patients can manually create a new collection (Figure 1A) by entering a name and additional metadata. To prioritize and distinguish the collections, we introduced descriptors: purpose, tags, and priority. For example, in Figure 1A, the patient created a High Blood Pressure collection with the purpose Clinical Visit for an upcoming appointment. They also added tags such as

XSL•FO
RenderX

hpb, sweating, and shortness of breath for reference. The priority descriptor indicates that the collection addresses a current and urgent issue. Patients can modify any of this information as the collection develops and changes.

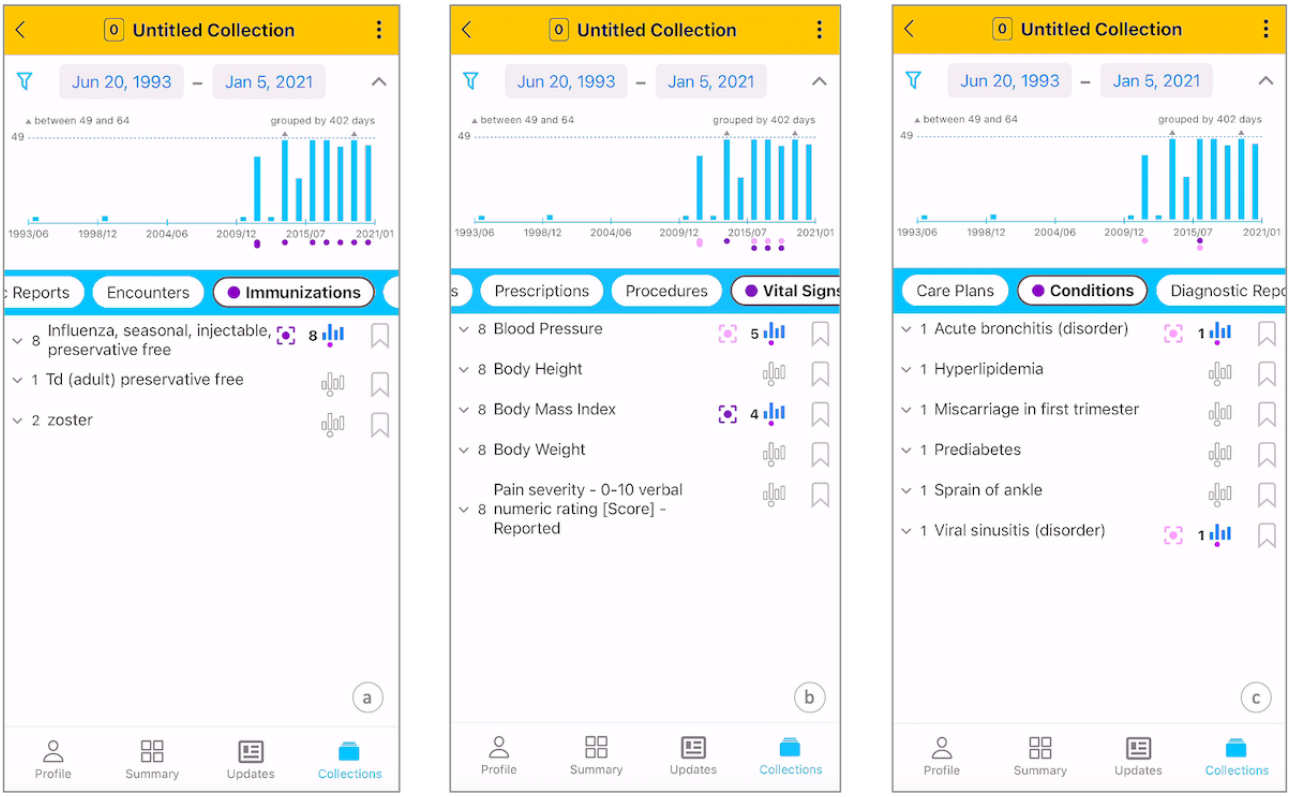
As patients repeatedly create collections, their EHR data transform into a problem-based list, as shown in Figure 1B, which includes health issues such as back discomfort, hamstring tightness, and high blood pressure. Collections are displayed with their name, creation and last modification dates, record count, number of patient-added notes, and labels for current (green clock) and urgent (danger sign) issues. An information button provides a summary with explanations. For example, the High Blood Pressure collection was created on December 31, 2021, and last modified on January 31, 2022. An information panel summarizes its contents: 9 records, 6 notes (4 for the collection and 2 for individual records), and tags for detailed search (hpb, sweating, and shortness of breath).

Updates automatically familiarize patients with the latest relevant events without requiring any action on their part (Figure 1C). The app scans all records and matches predefined templates, such as recent encounters, laboratory test results, and vital signs. In Figure 1C, 3 Updates are listed—Last Encounters, Last Lab Results, and Last Vital Signs—based on the last 5 dates when corresponding records were logged by the provider. The list entry follows the same structure as that of the Builds without the labels for currency and urgency. Although records in an Update cannot be changed, patients can add and remove notes. Patients can also clone an Update into a new Build and rename it if needed, allowing them to reuse and customize the assembled records for specific purposes. The reason for having this distinction between Updates and Builds is to delineate what the system and the user have produced and which entity is responsible for the collection that might have led to certain actions.

Identifying Relevant Records and Patterns for Collections

Discovery offers an interactive visualization to find patterns within records. The Timeline depicts record counts in equal time intervals, showing the prevalence of medical events. A dotted horizontal line marks the threshold above which the volume of records is considered abnormal, corresponding to the mean record count per interval. Gray triangle glyphs indicate values between the mean and 2 SDs, whereas red triangles highlight values of >2 SDs. By highlighting individual records or entire record types, patients can explore patterns that may be saved in existing collections or trigger the creation of new collections. For example, Figure 2A shows periodicity of influenza shots (when and how frequently influenza shots were administered), Figure 2B shows the co-occurrence of high blood pressure and high BMI, and Figure 2C shows the absence of respiratory conditions after an amoxicillin prescription.

Figure 2. Finding medical event patterns: (A) periodicity of influenza shots (showing when and how frequently influenza shots were administered), (B) co-occurrence of high blood pressure and high BMI, and (C) absence of respiratory conditions after an amoxicillin prescription.



The FHIR resources (ie, medical records) are represented with Record Cards, which display the clinical information in human-readable format. Patients can use a Filter and Date Picker to narrow down record categories and time frames for displayed records. Selected record categories appear in a Sliding Tabs control, allowing patients to swipe left or right for immediate access. Clicking on a record category in the Sliding Tabs organizes it by record type, represented with Accordions (eg,

XSL•FO
RenderX

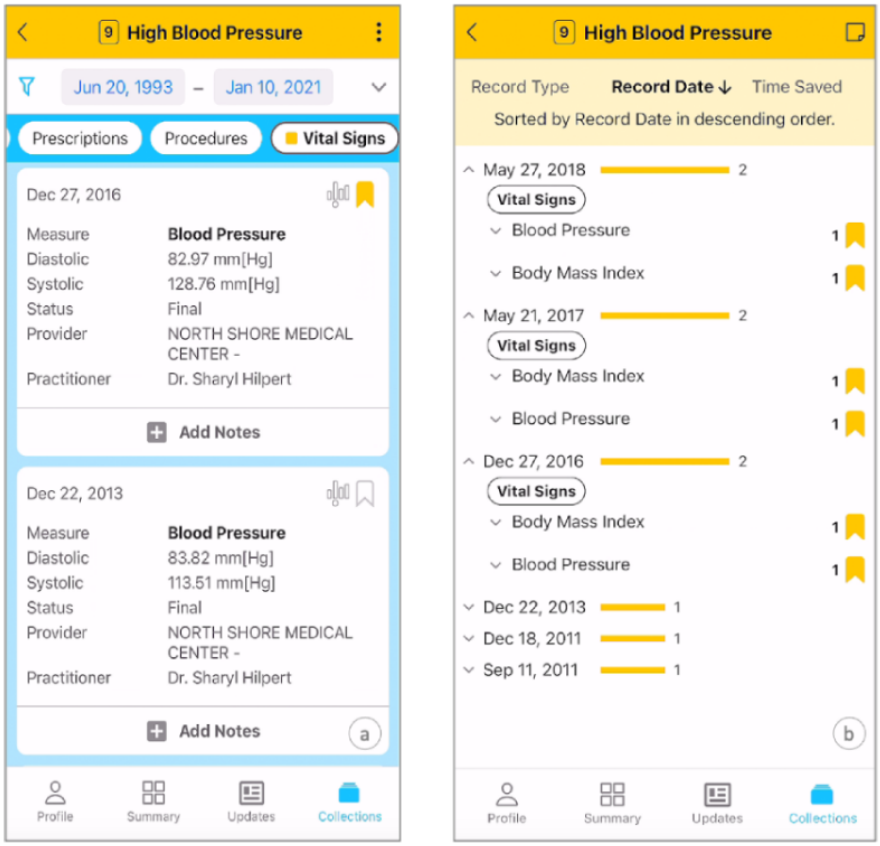
the Accordion for Immunizations category will have 3 sections for the Flu shot, Tdap, and Zooster record types). Patients can expand the Accordion sections to scroll through individual Record Cards. For example, Figure 2A shows 8 Record Cards for the influenza shot under the Immunization Accordion section. Accessing and revisiting records involves swiping the Sliding Tabs and selecting record categories, with Accordions retaining their expanded or collapsed state and scrolling position. This method is more context preserving compared to existing solutions, which require repeated back-and-forth navigation through different views for each record category and record type.

Producing Insights for the Collection

Patients can save individual records or entire record types by tapping the bookmark icon in the corresponding Record Card or Accordion section. Figure 3A demonstrates adding individual Blood Pressure records to a collection using a Record Card. Records can be removed from collections by tapping the selected bookmark again.

To identify data relationships and produce insights, patients can inspect a collection in the Collection Review (Figure 3B). Here, records can be viewed by type, date recorded in the provider's EHR system, or time added to the collection. When sequence matters, records can be ordered chronologically. Patients can also remove records in the Collection Review by tapping the selected bookmark. For example, in Figure 3B, the patient views records grouped by recording dates in descending order, with yellow bars indicating record counts by date. They observe high blood pressure and high BMI co-occurring 3 times, suggesting a pattern that may need further investigation or discussion with a physician.

Figure 3. (A) Saving a record to a collection and (B) reviewing the collection and inspecting data patterns.

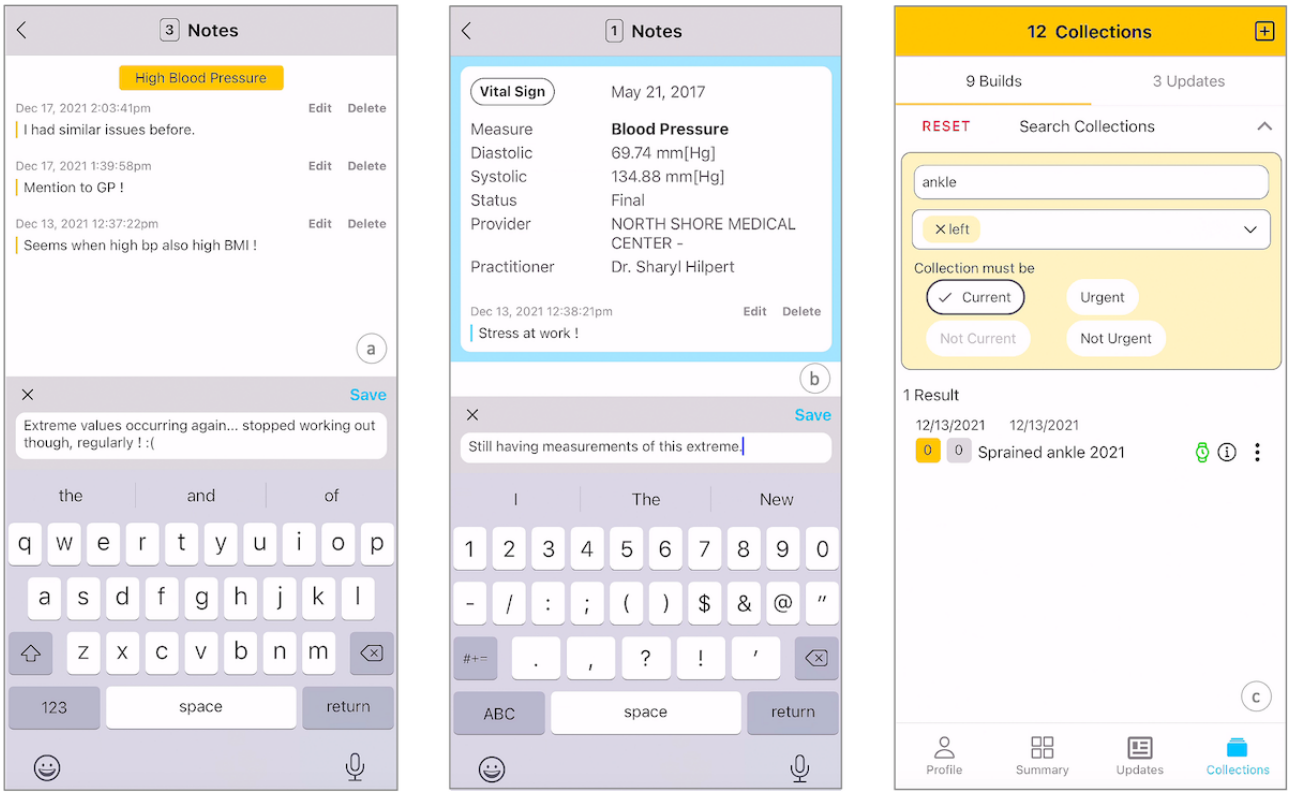


Supporting PGD

Patients can enter notes for a collection (Figure 4A) or individual records within it (Figure 4B) to add personal insights, progress, observations, details, or disease journal entries. Notes can be modified or removed at any time. In Figure 4A, the first note provides context on past high blood pressure experiences. The second note serves as a reminder to mention occasional high blood pressure to the general practitioner. The currently created note adds context about noticing changes in blood pressure after stopping regular workouts. In Figure 4B, the patient contextualizes a high blood pressure measurement taken during a stressful period at work and notes the recurrence of high values.

XSL•FO
RenderX

Figure 4. (A) Adding a note to a collection, (B) adding a note to a record, and (C) searching the collections.



Searching the Collections

Free-text search targets the collection name, purpose, and notes (Figure 4C), as well as tags and priority. Results dynamically update as the search query is constructed.

Study Design

Participants

We recruited 14 participants from our email list compiled from previous recruiting efforts and Craigslist. This number is sufficient to uncover usability issues and provide rich findings as per current design research practices [23] and literature on user feedback quality [24]. We balanced the sample by age, gender, and medical history (including healthy individuals, those with acute episodes, and those with chronic illnesses). Eligibility criteria included adults fluent in English; possessing an iPhone (iPhone 6 or above) and a laptop or desktop computer (screen size of 13“ or more) with a stable, fast internet connection for both devices; and with normal or corrected vision, no color blindness, and medical records from one or more providers. Medical history was self-reported.

Table 1 illustrates the detailed participant demographics collected using the questionnaire from Table S1 in Multimedia Appendix 1. Our 14 adult study participants included 10 (71%) women and 4 (29%) men aged 24 to 61 years (mean age 35.6, SD 12.6; median 30.5 years). All had some college experience: half (7/14, 50%) held bachelor’s degrees, 14% (2/14) had some graduate experience, and 14% (2/14) had completed master’s degrees. Participants had between 2 and 15 health care providers, with half (7/14, 50%) having ≥6. The 29% (4/14) of the participants who were healthy saw physicians a few times a year. Those with chronic illnesses had been managing their diseases for 1 to 20 years.

All participants were comfortable with daily technology use, and 21% (3/14) had work experience in data analytics. Most (11/14, 79%) used third-party apps to track mental health, weight loss, sleep, and exercise. All but 1 (13/14, 93%) used provider-patient portals to review test and laboratory results, refill prescriptions, and schedule appointments. However, participants found it cumbersome to remember multiple passwords and difficult to find specific information due to interface issues. Data sharing among providers was often slow or impossible, forcing participants to print and assemble records for clinical visits.

XSL•FO
RenderX

Table 1. Participant demographics.

Participant ID	Age (y)	Sex	Educational level	Medical issues	Health care providers, N
P1	61	Male	2-year college	High blood pressure, sleep apnea, fatty liver (stage 3), and liver transplant	7-8
P2	45	Female	Some college	High blood pressure, prediabetes, obesity, and hypothyroidism	6-7
P3	51	Male	Some graduate school (incomplete degree)	Hypertension and gout	2
P4	24	Male	Bachelor's degree	Childhood asthma, recent hernia surgery, and mental health therapy	5
P5	24	Female	Bachelor's degree	No chronic issues	2
P6	27	Female	Bachelor's degree	No chronic issues	3-4
P7	28	Male	Graduate degree	Physical and mental (20 years)	Approximately 13
P8	37	Female	Master's degree	Thyroid condition (many years)	3-5
P9	29	Female	3 years (incomplete degree)	Chronic migraines (10 years) and hemiplegic migraines (1 year)	≥12
P10	59	Female	Bachelor's degree	Yes—5 years	10-15
P11	25	Female	Bachelor's degree	No chronic issues	3-5
P12	33	Female	Master's degree	Anemia (2 years) and IBS[a] (10 years)	3
P13	19	Female	Some college (incomplete degree)	Asthma (since she was a baby) and EDD[b] (10 years)	>5
P14	32	Female	Master's degree	No chronic issues	3

[a]IBS: irritable bowel syndrome.

[b]EDD: eosinophilic digestive disease.

Procedures

Inspired by existing patient portal usability studies, the study procedures were tailored to our RQs [23,25,26] and are presented in Figure 5. The study was conducted remotely via Zoom (Zoom Video Communications) meetings. Participants downloaded our app on their iPhone using Apple TestFlight for beta testing [27] and mirrored their iPhone screen on their laptop using AirServer (App Dynamic ehf.) [28]. The laptop screen was then shared with the researcher for observation.

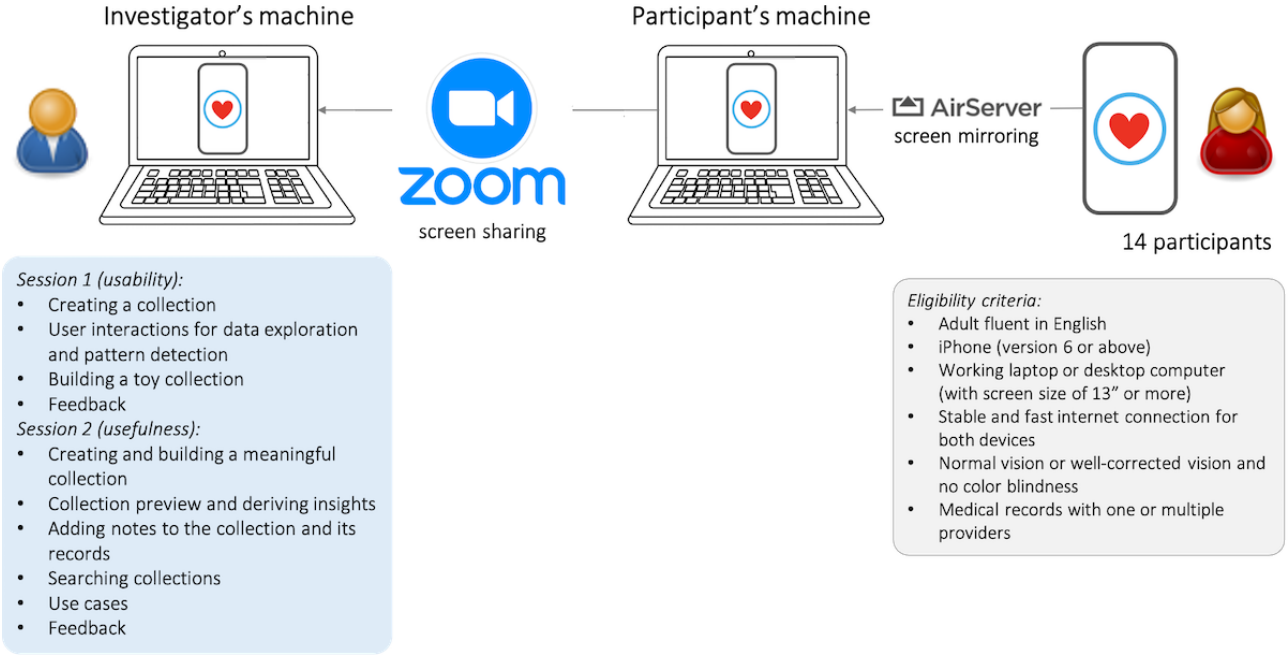
Participants attended an initial 60-minute session (session 1) and a follow-up 45-minute session (session 2). Both sessions used the concurrent think-aloud protocol to gather rich qualitative data on user needs, perceptions, and preferences efficiently [29]. Textbox 1 illustrates 2 example topics and detailed tasks for creating a toy collection and realistic collection from sessions 1 and 2, respectively. The full set of 13 topics for both sessions and their detailed tasks that follow a similar structure are shown in Tables S2 and S3 in Multimedia Appendix 1.

In session 1, the researcher first collected demographics and digital health consumer information from participants (Table S1 in Multimedia Appendix 1). The session then evaluated the usability of the app (RQ 1). Participants were introduced to the app's key concepts and features and were then given tasks to learn the interactions for creating and assembling relevant records in collections. After each task block, participants provided feedback. Tasks included creating a toy collection, adding descriptors, exploring scoping mechanisms for record categories and time ranges, navigating records using the Sliding Tabs, and inspecting the interactive Timeline visualization. Participants also completed tasks related to finding periodicity, co-occurrence, and pretest-posttest analysis of medical events using the Timeline; saving records to a collection; and reviewing it. The session concluded with feedback on the intuitiveness, usability, and usefulness of creating, building, and reviewing collections, as well as data exploration and pattern detection features (Textbox 2).

Session 2 focused on the usefulness of collections in real-life settings (RQ 2). Participants performed tasks involving creating a collection tied to a specific issue (eg, high blood pressure) and purpose (eg, clinical visit). After each task block, participants provided feedback. Tasks included creating a High Blood Pressure collection, adding descriptors for a clinical visit, selecting records related to high blood pressure, identifying patterns in selected records, adding notes to the collection and specific records for clinical context, and using the search feature to find other collections. The session concluded with feedback on the usefulness of collections, brainstorming real-life use case scenarios, and suggesting potential improvements (Textbox 2).

XSL•FO
RenderX

Figure 5. Study design—14 participants for 2 one-on-one sessions with a researcher. Session 1 focused on usability, and session 2 examined usefulness. Feedback was collected on how participants would use collections for their own data and needs.



Textbox 1. Example tasks from sessions 1 and 2.

Session 1: creating a collection (5 min)

- For the purposes of learning the basic mechanics around building a Collection from scratch, first Create a new Collection and name it "Toy Collection."
- Now, let's add some descriptors about the Collection.
  - Please add a purpose to the Collection, something related to learning about this app.
  - Now, add a couple of tags to further describe, summarize or annotate the Collection for future quick access.
  - Finally, specify the priority of this Collection by marking it as urgent.
- What was your experience with creating the new Collection?
  - How intuitive was it? Have you seen similar interactions elsewhere?
  - How useful do you find the descriptors for the Collection?
- Now, let's go back to the list of Collections. Let me know how can you see the details about the Collection you just created?
  - How intuitive was it?
- What are possible improvements?

Session 2: building a collection (15 min)

- You will now create a Collection that is more realistic and meaningful for use in a real-life scenario. We will assume that you are preparing for an upcoming visit to your physician's office related to potential issues with high blood pressure. Create a collection called "High Blood Pressure."
  - Add the purpose for the Collection
  - Add a few tags
  - Mark its priority
- Add the Records with blood pressure with systolic value over 120 and BMI over 30.
- What was your experience with assembling the Records for the Collection?
  - How laborious was it?
  - What are some ways in which we can make this assembling process more efficient?
  - How do you feel about having the system prepopulate the Collection for you and let you modify it afterwards?

XSL•FO
RenderX

Textbox 2. Feedback collected at the end of the study sessions.

Session 1 example tasks

- What was your impression of this app?
  - What did you like?
  - What did you dislike?
- How intuitive was the app?
- How easy or hard was it to explore the data?
- How useful were the features in the app to identify patterns in the data?
- How did you like the mechanism for saving Records in the Collection?
- What are some improvements you would like to see?

Session 2 example tasks

- How useful do you think the Collections can be for you?
- What are some use cases for the Collections that you can think of?
- What are some improvements you would like to see for the Collections?
  - Automatic support for building Collections?
  - Automatically finding data patterns in the Collections?
  - Patient-generated data?

Data Collection

We recorded the Zoom meetings for audio and video capture of the entire interaction. The first author also took notes during the meetings.

Data Analysis

Audio recordings were transcribed using Rev (Rev.com, Inc). We analyzed video recordings for additional context and a deeper understanding of participant comments during the think-aloud protocol. Video annotations were added to the transcripts and session notes [23] for reflexive thematic analysis [30]. The first author began by open coding the textual data. Emerging categories were reconciled in meetings with the second and last authors to identify use cases and detailed approaches to organizing and annotating EHR data. These themes were validated in a group meeting with researchers unfamiliar with the collection concept and with our app.

Ethical Considerations

The Harvard Faculty of Medicine Institutional Review Board approved this study (protocol IRB20-1757). Participants signed a consent form, which also allowed them to opt out of the study at any point. Each participant received a US $40 Amazon gift card as compensation. The data obtained from the study sessions did not include any identifiable information about the participants and were stored on a password-protected computer with encryption. Only the research team had access to these data.

Results

Overview

The results from our study are organized and analyzed around five qualitative themes, two applicable to RQ 2 and three applicable to RQ 1: (1) using collections for personal benefit (RQ 2), (2) using collections in a clinical setting (RQ 2), (3) creating and building collections (RQ 1), (4) enhancing collections (RQ 1), and (5) accessing collections (RQ 1). In the remainder of the Results section, we characterize the participants and report on these themes using 16 quotes from 11 different participants. The participants are labeled as P1 to P14.

Purposes for Organizing the EHR Data in Collections

Using Collections for Personal Benefit

Quick Access to Information

Participants perceived the Collections feature as a way to index information at a suitable level of granularity for health issues, topics of interest, or conditions to monitor. They expected that this would give them quick access to current or urgent problems that were being managed:

> *Personally I would like to monitor my asthma because I am using medication for that, the typical inhaler, but I would like to monitor, these days I had certain attacks or shortness of breath, so collections, having it back for that specific condition is very useful to me, because I wouldn't have to speculate about when my last attack was or when my last appointment date was, it's right here for me to access.* [P13]

XSL•FO
RenderX

Reflection on Medical History

Participants viewed collections as a tool to construct and understand their medical history. They expected collections to serve as a repository for the health issues they faced, aiding in reflecting on the existence, prevalence, and development of their health conditions:

> *This will be able to find what happens when I have, in my case, headaches related to my high blood pressure and nothing else, or if you're sick for the flu, influenza and you have fever, it's just related to the influenza, not to the COVID.* [P1]

Keeping Track of Health Status

Most participants envisioned using collections to track their health status proactively. This included monitoring urgent issues needing immediate attention, unstable conditions requiring frequent observation, and treatments needing careful monitoring. Participants also wanted collections to track abnormal laboratory test results and vital sign values across various health issues:

> *Well, for me, it's kind of good [to have medical records organized in collections], especially, for example, blood samples, especially those with high triglycerides or something. Maybe I can collect them and see whether there's a trend for this month, or for January, I'm high in this one. Then second month, I'm also high, so maybe I can lower it down...For collections, just categorize those. Which are high, which are low.* [P5]

Journaling Daily Events

Most of the participants envisioned using collections as a personal diary for coping with diseases and logging measurements and their effects on lifestyle. They wanted to track challenges, successes, and progress toward finding solutions and monitor disease developments:

> *What I've done is I've taken all of my videos and stuff since February. Like I said, I've been to eight different doctors and I've shown them, this is the progress of what's happened from...I've had two surgeries before this surgery where they lanced it, cut it, drained it. Nothing happened. The cyst came back and then it went into my bone. So I'm able to bring these photos, I'm able to bring this timeline, I'm able to bring my frustrations and show this doctor within 30 seconds [using a collection], look, this is what it looked like. And this is my own diary, my own history. It's very important because they're a doctor. They don't know me, they don't know what it looked like on day one.* [P10]

Learning From Past Experiences

Nearly all participants wanted to use collections to identify trends and patterns in their ongoing health experiences, including co-occurring symptoms and treatment effects. They also intended to log food intake, sleep, activity, or stress factors to find triggers for symptoms:

> *I think I could definitely use them there. It's a lot easier now because I could highlight the certain event, and put like my triggers down with it, like migraine on the third was this, you could even put what you took with it. So for me, I would look back and be like, "Oh, I can tell my doctor that, I've had 10 migraines. I took a medication with these three. What's my options." So I think that would be great, it's a great tool that I can actually do that with this app.* [P9]

Using Collections in a Clinical Setting

Preparation for the Clinical Visit

Most participants would use collections to prepare talking points for clinical visits combining personal measurements and notes with medical records from other providers (eg, laboratory and test results). This was crucial because their physicians often did not have access to external data:

> *For collections, I would say that if I'm meeting with multiple providers about one health issue, I could see myself combining all my records there so that multiple providers can see each other's records...I would probably [use the notes], if I needed to jot down a certain time that I took a measurement, or if my doctor told me keep track of what you ate that day, or kind of anything that I would want to have the details for the next time that I go and see the provider.* [P6]

Relying on Collections During Clinical Visits

All participants wanted to share collections with their physicians during clinical visits to establish ground truths, raise awareness of other providers' information, and provide transparent talking points. They believed that adding PGD to collections could describe what happened between visits and raise their physicians' awareness:

> *If you have a condition, you need to check on your blood pressure. You need to communicate that with your doctor, so you can add a note [in the collection] saying like, "Latest, highest blood pressure from this week," from a date.* [P14]

In addition, most participants felt that taking in-visit notes and saving them in a collection could help them understand care plans and take appropriate actions afterward:

> *Maybe [taking notes in the collection for the clinical visit] just for your own personal reference or if you wanted to bring it up later on in another appointment or something, or just maybe, I guess, just general recording of something that happened during that visit.* [P11]

All participants saw physicians as essential partners in reviewing collections and deciding on actions based on their contents. This was primarily because most participants doubted their expertise in determining what should go into collections or which collections to create. While they were very open to include their physicians in the collection curation, some feared that they might overburden physicians with verification inquiries:

> *I guess I would definitely do that [look for patterns in the data and store them in a collection] just*

XSL•FO
RenderX

> *because I can actually consult the doctor. Is this actually correlated or will I have to change my diet because it affects this? At least you can ask the doctor, or confirm whether that is true or not.* [P5]

Needs and Feature Preferences for Organizing EHR Data in Collections

Creating and Building Collections

Manual Workflow for Creating and Building Collections

Overall, all participants expressed satisfaction with the clarity and simplicity of the mechanics to create collections and save records in them. However, those with less medical knowledge and disease experience (relatively healthy and recently diagnosed individuals) thought that initiating and building collections was challenging. For them, it was not always clear what issues deserved separate collections, what records to include in a collection due to delicate dependencies, and why they should invest substantial time and effort in assembling collections. In contrast, the more experienced (chronic) patients were relatively confident in their ability to carve a personalized view of their EHR data. However, some did acknowledge that they might not be as exhaustive and reliable in their data organization.

Participants quickly mastered using the Sliding Tabs with Accordions and appreciated the context-preserving record exploration. However, most of the participants found it challenging to assess the relevance of individual records due to unexplained clinical language and attribute values. Participants requested explanations, prominent visual cues for abnormal values to attract their attention, and time-series visualizations for additional context and noticing trends. They also wanted to be able to sort by date or attribute value, with filtering capabilities that also included adding the filtered records in bulk to a collection.

Some participants desired collections with richer internal structures, recognizing the need to identify subsets of records and their importance within a collection. They suggested linking groups or individual records using specific annotations for easy identification during visual inspection or search:

> *...maybe if certain records are related to each other. So I would want to mark that. And then maybe just have a way of sorting down based on certain labels.* [P8]

Automatic Support for Building Collections

Participants were highly receptive to discussing ideas that could automate building collections. They were interested in seeing their records automatically put into collections based on provenance (provider, physician, hospital, location, and date) and clinical meaning (condition, disease, organ, organ system, and abnormality of the values across records). Some also suggested grouping records into collections based on personal annotations. For all these groupings, they wished to be able to edit the collection manually:

> *If I had neurological problems, neurology collection. If I had urological problems, urology collection. I think that for me at least would seem a more straightforward way to categorize them. But from the categories I've already seen, I find those useful.* [P4]

Several participants suggested using a "seed" to automate collection building, such as naming a collection, adding keywords, or including a few records:

> *And I think a lot of patients don't know where to start, what data to begin with. So if it's something that's already preset, they say, "Okay, I'm suffering from depression or I have diabetes." And the system pulls the different data points that they would need to look at for someone who's diabetic or someone who's dealing with depression, I think that's helpful. Because sometimes, the problem is you don't know where to start and you don't know what to look for.* [P11]

Finally, most of the participants wanted to receive automated help to add or remove records for a collection that had already been created. Few described wanting to choose from a list of suggested records based on the existing content of a collection, revealing records and record patterns otherwise invisible to them. Others saw this automated record offering more as an idea generation approach—needing some follow-up validation, including taking it up with their physicians:

> *When I work, I want to listen to some music and then I'm like, okay. I just don't know what's next. I need something similar to this, the same vibe, but I just can't think about that. And then there is suggestions. And yes, some of it's weird. But maybe like the doctor can also have some help here, and when you review the collections together, they might say, "Hey, listen, this is what the system gave you and that's great. Let's remove a few things. I would suggest you add a couple others. And whatever you put there, it's also fine. And let's keep it that way."* [P6]

Enhancing Collections Through Personally Provided Data

Making Collections Complete

Participants strongly expressed the need to complement their EHR data with daily entries from sensors; self-monitoring devices; and manual measurements of symptoms, treatments, and outcomes in various formats (text, photos, videos, and scanned documents):

> *Actually, I think you can sort of restructure the whole core of the collections on top of two main pillars. The first one would be all of the doctor's data, which is basically hard data, which allows you to diagnose, allows you to run statistical analysis...That could be part of the core data, but all of the context, maybe I'm getting this shortness of breath in my home, watching my TV, might be added by the notes. You have these two types of data. By adding the user data, would allow me to get context, give context, which is important and will allow me to, on a daily basis, keep a record, which in case of data like shortness of breath, I'm having, I'm not having. Would allow the*

XSL•FO
RenderX

> *doctor to have a really unbiased input on symptoms I'm having.* [P7]

Participants wanted to log detailed observations and measurements, pairing treatments with outcomes and symptoms with triggers. They suggested dedicating a special PGD record category for these data, with some preferring complex structures and others favoring simple data entry options:

> *I would probably use the notes quite often just to maybe outline the symptoms I was experiencing and the steps I took to alleviate those symptoms or which doctors I contacted.* [P13]

Making Collections Distinguishable

Participants liked the existing collection descriptors and suggested additional ones. They wanted labels for clarity (clear, unclear, or potential issue), stability (stable or unstable), progress toward resolution, development stages (nonthreatening or threatening), and a list of involved providers and physicians. When collections were related to clinical visits, participants wanted to specify the targeted physicians.

Making Collections Actionable

Participants believed that organizing medical records by health issues in collections was a good start but thought that actionability could be improved with specific insight notes and annotations applied to entire collections:

> *And then as far as the purpose of adding a note to the whole category, I would say that, like you said, if you happen to notice any patterns when you're looking at the data, or basically I would use it for any general or bigger-picture takeaway that I wanted to tell my doctor, "Hey, I noticed this" or something and I wanted to bring it to their attention.* [P6]

Participants envisioned using collection-wide notes to summarize contents or purpose, track progress, describe issue development, and highlight special events. They also wanted notes representing care plans and actions prioritized in a to-do list. Participants intended to use collections to prepare for clinical visits with questions, reminders, and critical measurements. They also saw value in adding collection notes about visit outcomes, key takeaways, and next steps.

Some participants wanted to annotate and highlight keywords or add tags to free-text notes for organized review and pattern identification.

Accessing the Collections

All participants emphasized the importance of fast, reliable access to collections and their contents. They primarily relied on collection descriptors but also desired a deep search feature that would scan through individual records, notes, and annotations within collections.

Discussion

Principal Findings

We identified 3 principal findings of our study. First, participants embraced the collection concept. Unrestrictedly organizing EHR data into collections that map medical records to health issues and track ongoing concerns gave participants a sense of ownership. They felt empowered by developing personalized health narratives that could aid in self-management and communication with their physicians, enhancing their self-advocacy. Second, while participants easily mastered the interface for initiating and adding records to collections, they found the process laborious. They lacked confidence in selecting appropriate records due to limited medical knowledge and requested additional visual cues, explanations, and automatic collection features. There was concern about potential self-misguidance without physician verification. Third, collections would need richer PGD capabilities for adding contextual information not found in participants' EHR data, logging observations, and labeling data. This would enhance the comprehensiveness and accuracy of their health narratives and support foraging, sensemaking, and action taking.

Interpretation of the Findings and Contributions

Overview

On a broader scale, this work contributes to patient-centered care. This is achieved by demonstrating potential to enhance patients' grasp of their health, encourage self-advocacy, and improve patient-provider communication. More accurately, there are several concrete contributions of our work that can be considered as proxies toward achieving the aforementioned objectives: (1) encouraging patient ownership of their EHR data by organizing them into personalized, health issue–based collections; (2) understanding patients' perceptions and preferences for creating, building, and using these collections; and (3) offering design insights for automating collections, integrating rich PGD, enhancing access to collection contents, and using collections to facilitate patient-provider communication.

Going forward, we will situate our findings within a sensemaking framework and discuss contributions related to 3 key patient needs: increasing awareness through independent health sensemaking, proactivity through efficient action taking, and self-advocacy through incorporating evidence-supported patient perspectives into patient-provider communication. We will elaborate on how collections can meet these needs and offer design implications to enhance their capabilities.

The Role of the Collections in Supporting Sensemaking

To explain the collections' role in sensemaking, we used the model by Pirolli and Card [6], which divides the sensemaking process into 2 subcycles: the foraging loop and the sensemaking loop. On the basis of this model, collections can be described as a space for assembling relevant data about a topic, finding relationships between them, and storing outcomes from the sensemaking. In the foraging loop, patients gather relevant records to answer questions such as the following—"Is there a relationship between my weight and blood pressure?"—and save them in a collection, such as Weight vs. Blood Pressure. In the sensemaking loop, patients identify information relationships within the collection that they capture in notes, such as instances where there was co-occurrence of high blood pressure and high body mass. These notes help argue hypotheses such as the following: "My blood pressure is high when I'm

XSL•FO
RenderX

overweight." The outcome of this sensemaking process could be a comprehensive note for a clinical visit.

While, in their current form, collections respond to the needs of the sensemaking model by Pirolli and Card [6], improvements can be made to make this more efficient. This study revealed that medical records alone are not enough for reliable sensemaking. Adding PGD such as symptoms, measurements, outcomes, and everyday events is essential for creating comprehensive collections. The foraging loop can be made less laborious and time-consuming if there are additional visual cues, medical explanations, filtering capabilities, and automatic support to improve the relevance and reduce the effort of assembling collections. The sensemaking loop could be improved by adding more schematization capabilities such as annotating medical records and PGD to identify patterns later (eg, symptom triggers, medication effects, and correlations) and grouping records within collections, labeling those groups, and establishing group relationships with explanations (eg, linking "cholesterol lab results" with "food intake" as "food effects on cholesterol" in a High Blood Pressure collection).

Reliability of Collections

According to our study, there are 4 main factors that can influence the reliability of the collections: robust coverage of health issues, provision of PGD, grouping and linking of medical records and PGD within a collection, and verifying the contents of the collections. This reliability is related to collections' capability to aid in creating personalized but realistic health issue narratives, support self-management, and stimulate awareness and proactivity.

Robust Coverage of Health Issues: Relevance Assessment

Collections should ensure that patients can create collections for their most important health issues to support awareness and proactivity. Participants desired visual cues, explanations, and automatic support to determine which collections to create and what records to include.

While participants found the context-capturing data exploration using Sliding Tabs and Timeline convenient, they needed more to identify relevant records quickly. Future tools could incorporate Accordions that summarize record types, graph values over time, and highlight abnormal or extreme values. In addition, patients should be able to expand individual records to see explanations of clinical terms and clinical meaning interpretations. While there is existing work related to visualizing time series of EHR data [17-20] and automatic provision of short explanations [31], this study shows the need for combining them in a new way to support relevance assessment for a novel purpose—constructing collections. Finally, patients should be able to order and filter records by attribute for quicker browsing and bulk addition to a collection.

Automatic support should also be provided for creating and building collections to save time and ensure robust, reliable coverage of health issues. Our previous work highlighted the need for automating collections [13], and this study highlights a clear preference for automatically grouping medical records by clinical meaning—whether *thematically* or *guided by patient input*. Thematic collections would be those that tie records together based on conditions (bronchitis or diabetes) and procedures (stent placement or appendix removal) or with respect to organs (heart or kidneys) or organ systems (cardiovascular or renal). Alternatively, patients could specify a seed by setting parameters such as the collection's name, tags, or initial records. The system can offer candidate records to include or delete with confidence scores and explanations. Patients could then refine these system-generated collections by adding or removing records, PGD, and tags.

Addressing automatic support for collections may be challenging due to subtle relationships between medical records [32-34]. However, starting with easier constructs such as *time stamps* (eg, medical records from the same day, week, or month), *FHIR links* (eg, medical records from the same encounter or physician), *abnormal values* based on well-established clinical guidelines (eg, high or low blood pressure or cholesterol), *test findings* (eg, positive and negative), and *patient tags* (eg, triggers or pivotal events) is a feasible approach.

Provision of PGD: Sensemaking Data Completeness

Collections should include PGD to improve data completeness for sensemaking. Previous research has suggested that maintaining consistent PGD logging over time is difficult [35]. However, this should not be considered a barrier or a prerequisite for the collections' success. Patients' motivation and preferences for PGD logging intensity vary based on their disease self-management state [36]. When setting goals and learning strategies, patients prefer meticulous data collection. Once goals and strategies are in place, logging intensity decreases. In addition, if physicians require PGD logging for treatment planning, patients are motivated to engage in it [36,37].

Thus, collections should enable flexible and efficient PGD logging. Disease-specific contexts such as irritable bowel syndrome [38], diabetes [7], and migraine [39] have explored health sensemaking without focusing on diverse data types. This contrasts with patients' desire to log PGD for various medical issues within a single application [40,41] using a universal logging model for different observations [42,43]. To address this issue, we propose a straightforward workflow where patients initiate free-text entries and use tags to specify the type, quality, or other details. This mechanism allows for quick data capture and embellishment at convenient times.

Tags can classify PGD as *clinical observations*, *everyday life events*, or *notes*. Further specification can be added using tags such as *symptom*, *measurement*, *treatment*, and *outcome* for observations; *meal*, *exercise*, *meeting*, and *deadline* for life events; and *context*, *personal note*, and *visit note* for notes. Additional tags such as *absent*, *normal*, *high*, *low*, *extreme*, *improvement*, *deterioration*, *pivotal event*, *trigger*, or *relaxer* can be used for further detail that captures the quality and importance of the logged data. In addition to these system-offered tags, patients can also create their own custom tags for better personalization.

XSL•FO
RenderX

Grouping and Linking Within Collections: Schematization Capabilities

Patients need to connect medical records and PGD within collections for easier sensemaking. Future tools should add structure by enabling record grouping and linking of groups or individual records. This helps highlight important subsets of records and trace major conclusions as collections grow.

We recommend using a simple yet powerful tagging concept. Records sharing the same tag can form a group, whereas links between groups can be specified using related tags. The same mechanism can link individual records with other records or groups, providing nuanced sensemaking. This approach aligns well with the proposed PGD tagging model that can be applied to medical records as well.

Collection Verification

Collections represent the patient's personalized perception of their health and issues. As such, they should undergo occasional verification by the patient's physician for safe decision-making and action taking. While collection verification may add to the physician's workload, it can inspire and enable patients to manage health issues more independently. That said, patients should consider the physician's workload before requesting verification [44].

Future tools could allow for the conversion of a collection into a well–laid-out PDF document capturing all its contents. This document can be printed for review during a clinical visit or shared as a PDF attachment in the patient portal for verification at the physician's convenience.

Taking Actions Based on Collections

While annotating PGD is known to aid learning and disease self-management [7,45], our findings reveal that annotations can also enhance EHR data, creating synergy with PGD. Participants expressed a desire to annotate their data for various purposes: identifying triggers to avoid or encourage certain behaviors, marking pivotal events to remind them of shifts in health attitudes and management, and labeling outcomes as desired or undesired to evaluate treatments and strategies. These capabilities can be easily implemented using the previously elaborated tag-based design for linking records.

To increase the awareness and prioritization of collections, we previously proposed *collection descriptors* such as topic, urgency, currency, and sentiment [13]. Participants found value in these descriptors but expressed a need for additional ones that can be classified as *patient specified* and *data driven*. Future tools may include patient-specified descriptors for *clarity* (eg, is the diagnosis clear?), *stability* (eg, is the treatment working consistently?), *severity* (eg, is there a significant medical risk?), and *progress markers* (eg, is the issue substantially resolved?). Data-driven descriptors could be derived from the collection data, indicating the *time span* (from the oldest to the latest record) or listing the *physicians involved* (the providers and physicians the records came from). Both types of descriptors should be optional for patients to use as needed.

Providing an inner structure, enabling annotations, and describing collections can improve information access and expedite decision-making. Powerful search engines can use these metadata to allow patients direct and easy access not only to individual collections but also to their specific contents.

While these features can enhance collections' actionability, it is important to note that collections are not meant for making independent clinical decisions by patients. Collections should be verified by a physician to serve as reliable tools for sensemaking and health self-management. However, collections can always be invaluable tools for patients to understand their health; organize thoughts, hypotheses, and insights; and communicate effectively with their physicians.

Collection Use for Patient-Provider Communication

As observed in our findings, patients can use collections to prepare for a clinical visit by devising checklists and organizing thoughts supported by evidence. During the visit, collections can be used for *note taking* and, afterward, for *recording reminders* and *follow-up actions*. These uses indicate how collections can start addressing known challenges during clinical visits, such as problem presentation [46], information retention [47], setting common ground, aligning goals, and understanding instructions [48]. To effectively tackle such challenges, the Collections feature should support richer note capturing and collaborative data analysis in a collocated setting [49,50].

Future improvements in capturing PGD could make collections more appealing to physicians. For physicians, PGD play a crucial role in understanding the boundaries and context for accurate diagnosis and optimal treatment [51]. However, physicians often face problems with PGD, such as incomplete data, inconsistent data structures, and insufficient time for reviewing due to poor organization [52]. These issues arise because patients use disjointed platforms to log their data, lacking consistent models for logging different types of data [52,53], and face challenges in efficiently using these platforms [42,52]. Collections can help overcome these issues by providing a single platform for logging PGD for various health issues in a simple, universal way that allows for flexibility, organization, and standardization.

An alternative approach to enhancing collections as a communication tool and fostering physician collaboration in their creation and verification is to introduce them as a *shared resource* similar to Google Docs [54,55]. While this may seem unconventional, it builds on the principles of OpenNotes [56]. OpenNotes provides access to and transparency regarding clinical notes, enabling patients to improve their treatment and EHR data quality by taking an active role in detecting errors, raising concerns, asking questions, or seeking clarifications [56,57].

Similar to OpenNotes, *shared collections* would follow the principle of asynchronous communication and transparency. However, shared collections could eliminate the expressiveness constraints and lack of efficient ways to provide granular and tailored context observed in existing messaging systems [58,59]. In addition, shared collections would enable direct editing of underlying data in collaborative ways, minimizing communication overhead.

XSL•FO
RenderX

Moreover, shared collections would introduce a new communication channel between patients and providers outside the traditional patient portal. Synchronizing the digital traces of care in collections with the provider's EHR system to avoid discrepancies and legal issues should be a top consideration in future design iterations of the shared collection concept.

A Glimpse Into the Future: Collections and Generative Artificial Intelligence

In the future, we should explore the potential of generative artificial intelligence (GenAI) models to support patient sensemaking through collections. Tools such as ChatGPT [60] and Med-PaLM [61], which have demonstrated substantial medical knowledge [62-64], can replace the need for custom-made machine learning algorithms for knowledge-intensive tasks.

In particular, GenAI tools can aid in automatic and iterative collection construction with explanations and guidance. They can analyze the data within collections for insights, including medical records, PGD, notes, and tags. In addition, GenAI can assist in composing case narratives and talking points for clinical encounters. By offering this level of automation, GenAI can help tackle the significant knowledge challenges while lowering the labor barrier for patients' sensemaking activities.

Using GenAI models, collection construction could rely on natural language instructions such as the following: "Group my EHR data by condition," "Find all records related to my bronchitis," or "Identify records that don't belong to this collection and those that are missing." GenAI models could also deliver context, explaining why certain records are included or excluded and providing educational material such as term definitions and clinical implications.

In addition, patients can issue commands for identifying relationships within their annotated data, such as "List all triggers for my headaches over the last year." Finally, they can ask for help in constructing case presentations for clinical visits (eg, "Based on my 'High Blood Pressure' collection notes, write a 100-word summary").

To be useful for sensemaking, GenAI tools do not need to achieve complete accuracy. While still striving for maximum reliability, their main value should come from providing an environment that enables and encourages patients to refine artificial intelligence–generated outputs. As such, the contribution of GenAI toward sensemaking would be evaluated on its ability to help the patient efficiently produce a satisfactory solution with minimal physician input.

Finally, existing approaches for supporting sensemaking through search and interactive visualizations should not be disregarded. Exploring the integration of GenAI, search, and visualization is a prudent strategy as different sensemaking tasks related to collection assembly, editing, and analysis may require diverse approaches based on complexity, patient skills, and artificial intelligence reliability.

Limitations

This study has several limitations. First, the cohort skewed younger, likely due to recruitment via Craigslist (less popular with older adults) and the complexity of the remote setup. Second, participants used data from a fictitious patient, which may have reduced their motivation to learn the app and their ability to suggest real-life use cases. Third, participants had limited time to learn how to interact with collections, possibly affecting their perceptions of usability and utility. Future studies should have participants use their own data with automatic interaction logging. Despite these limitations, this study provided valuable insights into designing patient-facing sensemaking tools for organizing and augmenting EHR data.

Conclusions

Collections can potentially improve patient-centered care by involving patients more in decision-making and encouraging self-advocacy. Current assumptions often expect patients to have the necessary skills, tools, and motivation. We believe that collections can lower these barriers, encouraging patients to *increase engagement* with their health data, better *educate themselves*, and *communicate more effectively* with their care providers.

Our study suggests that EHR data can be better used and more useful for patients through *improved organization* and *annotation*. This approach can incentivize patients to engage more deeply with their EHR data, develop insights, and reflect on their experiences. Patients felt that this empowered their awareness, resourcefulness, and proactivity regarding health issues, making them more prepared and better informed for clinician interactions.

These findings support our premise that collections are a crucial step toward *patient empowerment* and *self-advocacy*. With appropriate improvements, collections can enhance patients' expertise by facilitating sensemaking activities and enabling insightful discussions with their physicians. First, collections motivate patients to construct health models based on their issues and ongoing problems. Second, patients gain medical education by actively participating in the evolution of collections through independent or system-assisted assembly and editing. Finally, patients acquire additional medical knowledge by engaging in meaningful discussions with their physicians and considering their feedback on collection verification.

Our study highlighted the importance of integrating PGD with EHR data. We envision a synergy in which patients use clinical data as a foundation, augmenting them with their observations, notes, and annotations to create personalized health narratives that support better health management and provider communication.

In the future, we should explore GenAI models to support patient sensemaking through collections. These models could help patients build collections, analyze the data within them, and produce health narratives more efficiently. Such enhancements may also reduce physicians' workload for verifying collection contents, leading to more focused, evidence-driven discussions during visits.

Promising ideas from this work should be advanced carefully, with gradual design improvements tested in real-life settings. Respecting existing clinical practices and workflows can facilitate quicker adoption and more significant changes in the

XSL•FO
RenderX

future. We believe that collections can revolutionize how patients interact with their medical records and communicate with their providers.

Acknowledgments

This work was funded by Scripps Research through the Technology to Empower Changes in Health Network Participant Technologies Center award (grant 5U24OD023176).

Conflicts of Interest

NG is a cofounder and equity owner of datavisyn.

Multimedia Appendix 1
Demographic and participant characterization questions and the evaluation tasks for the 2 study sessions—session 1 and session 2.
[DOCX File , 25 KB - humanfactors_v12i1e50331_app1.docx]

References

1. Penedo FJ, Oswald LB, Kronenfeld JP, Garcia SF, Cella D, Yanez B. The increasing value of eHealth in the delivery of patient-centred cancer care. Lancet Oncol 2020 May;21(5):e240-e251. [doi: 10.1016/s1470-2045(20)30021-8]
2. Kwame A, Petrucka PM. A literature-based study of patient-centered care and communication in nurse-patient interactions: barriers, facilitators, and the way forward. BMC Nurs 2021 Sep 03;20(1):158 [FREE Full text] [doi: 10.1186/s12912-021-00684-2] [Medline: 34479560]
3. New applicability dates included in ONC interim final rule. The Office of the National Coordinator for Health Information Technology. URL: https://tinyurl.com/bdhxedu2 [accessed 2025-01-14]
4. ONC's cures act final rule. National Coordinator for Health Information Technology. URL: https://tinyurl.com/3upcepm4 [accessed 2025-01-14]
5. Naumer C, Fisher KE, Dervin B. SenseMaking: a methodological perspective. In: Proceedings of the CHI Conference on Human Factors in Computing Systems. 2008 Presented at: CHI '08; April 5-10, 2008; Florence, Italy.
6. Pirolli P, Card SK. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In: Proceedings of International Conference on Intelligence Analysis. 2005 Presented at: ICIA 2005; May 2-6, 2005; McLean, VA. [doi: 10.1145/1358628.1358972]
7. Raj S, Lee JM, Garrity A, Newman MW. Clinical data in context: towards sensemaking tools for interpreting personal health data. Proc ACM Interact Mob Wearable Ubiquitous Technol 2019 Mar 29;3(1):1-20. [doi: 10.1145/3314409]
8. Healthcare. Apple Store. URL: https://www.apple.com/healthcare/health-records/ [accessed 2024-09-08]
9. iBlueButton homepage. iBlueButton. URL: https://www.ibluebutton.com/ [accessed 2024-09-08]
10. OneRecord homepage. OneRecord. URL: https://onerecord.com [accessed 2024-09-08]
11. 1upHealth homepage. 1upHealth. URL: https://1up.health/product/patient [accessed 2024-09-08]
12. Nakikj D, Kreda D, Gehlenborg N. New ways for patients to make sense of their electronic health record data using the discovery web application: think-aloud evaluation study. JMIR Form Res 2023 Apr 03;7:e41346 [FREE Full text] [doi: 10.2196/41346] [Medline: 37010887]
13. Nakikj D, Kreda D, Gehlenborg N. Alerts and collections for automating patients' sensemaking and organizing of their electronic health record data for reflection, planning, and clinical visits: qualitative research-through-design study. JMIR Hum Factors 2023 Aug 21;10:e41552 [FREE Full text] [doi: 10.2196/41552] [Medline: 37603400]
14. Zhang M, Ehrmann D, Mazwi M, Eytan D, Ghassemi M, Chevalier F. Get to the point! Problem-based curated data views to augment care for critically ill patients. In: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. 2022 Presented at: CHI '22; April 29-May 5,2022; New Orleans, LA. [doi: 10.1145/3491102.3501887]
15. Klein G, Phillips JK, Rall EL, Peluso DA. A data-frame theory of sensemaking. In: Hoffman RR, editor. Expertise Out of Context: Proceedings of the Sixth International Conference on Naturalistic Decision Making. Mahwah, NJ: Lawrence Erlbaum Associates Publishers; 2007.
16. Choe EK, Lee B, Zhu H, Riche NH, Baur D. Understanding self-reflection: how people reflect on personal data through visual data exploration. In: Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare. 2017 Presented at: PervasiveHealth '17; May 23-26, 2017; Barcelona, Spain. [doi: 10.1145/3154862.3154881]
17. Plaisant C, Mushlin R, Snyder A, Li J, Heller D, Shneiderman B. LifeLines: using visualization to enhance navigation and analysis of patient records. Proc AMIA Symp 1998:76-80 [FREE Full text] [Medline: 9929185]
18. Wang TD, Plaisant C, Quinn AJ, Stanchak R, Murphy S, Shneiderman B. Aligning temporal data by sentinel events: discovering patterns in electronic health records. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2008 Presented at: CHI '08; April 5-10, 2008; Florence, Italy. [doi: 10.1145/1357054.1357129]

XSL•FO
RenderX

19. Rind A, Wang TD, Aigner W, Miksch S, Wongsuphasawat K, Plaisant C. Interactive Information Visualization to Explore and Query Electronic Health Records. Norwell, MA: Now Publishers Inc; 2013.
20. Gotz D, Borland D. Data-driven healthcare: challenges and opportunities for interactive visualization. IEEE Comput Graph Appl 2016;36(3):90-96. [doi: 10.1109/MCG.2016.59] [Medline: 28113160]
21. Welcome to FHIR®. Health Level 7 Fast Healthcare Interoperability Resources. URL: https://www.hl7.org/fhir/ [accessed 2024-09-07]
22. SMART health IT homepage. SMART Health IT. URL: https://smarthealthit.org/ [accessed 2024-09-01]
23. Martinez W, Threatt AL, Rosenbloom ST, Wallston KA, Hickson GB, Elasy TA. A patient-facing diabetes dashboard embedded in a patient web portal: design sprint and usability testing. JMIR Hum Factors 2018 Sep 24;5(3):e26 [FREE Full text] [doi: 10.2196/humanfactors.9569] [Medline: 30249579]
24. Faulkner L. Beyond the five-user assumption: benefits of increased sample sizes in usability testing. Behav Res Methods Instrum Comput 2003 Aug;35(3):379-383. [doi: 10.3758/bf03195514] [Medline: 14587545]
25. Haggstrom DA, Saleem JJ, Russ AL, Jones J, Russell SA, Chumbler NR. Lessons learned from usability testing of the VA's personal health record. J Am Med Inform Assoc 2011 Dec;18 Suppl 1(Suppl 1):i13-i17 [FREE Full text] [doi: 10.1136/amiajnl-2010-000082] [Medline: 21984604]
26. Yen PY, Walker DM, Smith JM, Zhou MP, Menser TL, McAlearney AS. Usability evaluation of a commercial inpatient portal. Int J Med Inform 2018 Feb;110:10-18. [doi: 10.1016/j.ijmedinf.2017.11.007] [Medline: 29331248]
27. Developer. Apple Store. URL: https://developer.apple.com/testflight/ [accessed 2024-09-16]
28. AirServer connect homepage. AirServer Connect. URL: https://www.airserver.com/ [accessed 2024-09-16]
29. Alhadreti O, Mayhew P. Rethinking thinking aloud: a comparison of three think-aloud protocols. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI '18; April 21-26, 2018; Montreal, QC. [doi: 10.1145/3173574.3173618]
30. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport Exerc Health 2019 Jun 13;11(4):589-597. [doi: 10.1080/2159676x.2019.1628806]
31. Wilcox L, Morris D, Tan D, Gatewood J, Horvitz E. Characterizing patient-friendly "micro-explanations"of medical events. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2011 Presented at: CHI '11; May 7-12, 2011; Vancouver, BC. [doi: 10.1145/1978942.1978948]
32. Hirsch JS, Tanenbaum JS, Lipsky Gorman S, Liu C, Schmitz E, Hashorva D, et al. HARVEST, a longitudinal patient record summarizer. J Am Med Inform Assoc 2015 Mar;22(2):263-274 [FREE Full text] [doi: 10.1136/amiajnl-2014-002945] [Medline: 25352564]
33. Khedr AM, Aghbari ZA, Ali AA, Eljamil M. An efficient association rule mining from distributed medical databases for predicting heart diseases. IEEE Access 2021 Jan 19;9:15320-15333. [doi: 10.1109/ACCESS.2021.3052799]
34. Sánchez-de-Madariaga R, Martinez-Romo J, Escribano JM, Araujo L. Semi-supervised incremental learning with few examples for discovering medical association rules. BMC Med Inform Decis Mak 2022 Jan 24;22(1):20 [FREE Full text] [doi: 10.1186/s12911-022-01755-3] [Medline: 35073885]
35. Clawson J, Pater JA, Miller AD, Mynatt ED, Mamykina L. No longer wearing: investigating the abandonment of personal health-tracking technologies on craigslist. In: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2015 Presented at: UbiComp '15; September 7-11, 2015; Osaka, Japan. [doi: 10.1145/2750858.2807554]
36. Li I, Dey AK, Forlizzi J. Understanding my data, myself:supporting self-reflection with ubicomp technologies. In: Proceedings of the 13th international conference on Ubiquitous computing. 2011 Presented at: UbiComp '11; September 17-21, 2011; Beijing, China. [doi: 10.1145/2030112.2030166]
37. Andersen T, Bjørn P, Kensing F, Moll J. Designing for collaborative interpretation in telemonitoring: re-introducing patients as diagnostic agents. Int J Med Inform 2011 Aug;80(8):e112-e126. [doi: 10.1016/j.ijmedinf.2010.09.010] [Medline: 21067968]
38. Karkar R, Schroeder J, Epstein DA, Pina LR, Scofield J, Fogarty J, et al. TummyTrials: a feasibility study of using self-experimentation to detect individualized food triggers. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017 Presented at: CHI '17; May 6-11, 2017; Denver, CO. [doi: 10.1145/3025453.3025480]
39. Schroeder J, Chung CF, Epstein D, Karkar R, Parsons A, Murinova N, et al. Examining self-tracking by people with migraine: goals, needs, and opportunities in a chronic health condition. In: Proceedings of the 2018 Designing Interactive Systems Conference. 2018 Presented at: DIS '18; June 9-13, 2018; Hong Kong, China. [doi: 10.1145/3196709.3196738]
40. Tollmar K, Bentley F, Viedma C. Mobile health mashups: making sense of multiple streams of wellbeing and contextual data for presentation on a mobile device. In: Proceedings of the 6th International Conference on Pervasive Computing Technologies for Healthcare. 2012 Presented at: PervasiveHealth 2012; May 21-24, 2012; San Diego, CA. [doi: 10.4108/icst.pervasivehealth.2012.248698]
41. Hsieh CK, Tangmunarunkit H, Alquaddoomi F, Jenkins J, Kang J, Ketcham C, et al. Lifestreams: a modular sense-making toolset for identifying important patterns from everyday life. In: Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems. 2013 Presented at: SenSys '13; November 11-15, 2013; Rome, Italy. [doi: 10.1145/2517351.2517368]

XSL•FO
RenderX

42. Ancker JS, Witteman HO, Hafeez B, Provencher T, Van de Graaf M, Wei E. "You get reminded you're a sick person": personal data tracking and patients with multiple chronic conditions. J Med Internet Res 2015 Aug 19;17(8):e202 [FREE Full text] [doi: 10.2196/jmir.4209] [Medline: 26290186]
43. Jones SL, Hue W, Kelly RM, Barnett R, Henderson V, Sengupta R. Determinants of longitudinal adherence in smartphone-based self-tracking for chronic health conditions: evidence from axial spondyloarthritis. Proc ACM Interact Mob Wearable Ubiquitous Technol 2021 Mar 30;5(1):1-24. [doi: 10.1145/3448093]
44. North F, Luhman KE, Mallmann EA, Mallmann TJ, Tulledge-Scheitel SM, North EJ, et al. A retrospective analysis of provider-to-patient secure messages: how much are they increasing, who is doing the work, and is the work happening after hours? JMIR Med Inform 2020 Jul 08;8(7):e16521 [FREE Full text] [doi: 10.2196/16521] [Medline: 32673238]
45. Mamykina L, Mynatt E, Davidson P, Greenblatt D. MAHI: investigation of social scaffolding for reflective thinking in diabetes management. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2008 Presented at: CHI '08; April 5-10, 2008; Florence, Italy. [doi: 10.1145/1357054.1357131]
46. Robinson JD, Heritage J. The structure of patients' presenting concerns: the completion relevance of current symptoms. Soc Sci Med 2005 Jul;61(2):481-493. [doi: 10.1016/j.socscimed.2004.12.004] [Medline: 15893061]
47. Unruh KT, Skeels M, Civan-Hartzler A, Pratt W. Transforming clinic environments into information workspaces for patients. Proc SIGCHI Conf Hum Factor Comput Syst 2010 Apr;2010:183-192 [FREE Full text] [doi: 10.1145/1753326.1753354] [Medline: 21654895]
48. Raj S, Newman MW, Lee JM, Ackerman MS. Understanding individual and collaborative problem-solving with patient-generated data: challenges and opportunities. Proc ACM Hum Comput Interact 2017 Dec 06;1(CSCW):1-18. [doi: 10.1145/3134723]
49. Chung CF, Wang Q, Schroeder J, Cole A, Zia J, Fogarty J, et al. Identifying and planning for individualized change: patient-provider collaboration using lightweight food diaries in healthy eating and irritable bowel syndrome. Proc ACM Interact Mob Wearable Ubiquitous Technol 2019 Mar 29;3(1):1-27 [FREE Full text] [doi: 10.1145/3314394] [Medline: 31080941]
50. Schroeder J, Hoffswell J, Chung CF, Fogarty J, Munson S, Zia J. Supporting patient-provider collaboration to identify individual triggers using food and symptom journals. In: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing. 2017 Presented at: CSCW '17; February 25-March 1, 2017; Portland, OR URL: https://doi.org/10.1145/2998181.2998276 [doi: 10.1145/2998181.2998276]
51. Chung CF, Dew K, Cole A, Zia J, Fogarty J, Kientz JA, et al. Boundary negotiating artifacts in personal informatics: patient-provider collaboration with patient-generated data. In: Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing. 2016 Presented at: CSCW '16; February 27-March 2, 2016; San Francisco, CA URL: https://doi.org/10.1145/2818048.2819926 [doi: 10.1145/2818048.2819926]
52. West P, Van Kleek M, Giordano R, Weal MJ, Shadbolt N. Common barriers to the use of patient-generated data across clinical settings. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI '18; April 21-26, 2018; Montreal, QC URL: https://doi.org/10.1145/3173574.3174058 [doi: 10.1145/3173574.3174058]
53. Choe EK, Lee NB, Lee B, Pratt W, Kientz JA. Understanding quantified-selfers' practices in collecting and exploring personal data. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2014 Presented at: CHI '14; April 26-May 1, 2014; Toronto, ON URL: https://doi.org/10.1145/2556288.2557372 [doi: 10.1145/2556288.2557372]
54. Google Docs. Facebook. URL: https://www.facebook.com/GoogleDocs/ [accessed 2024-08-01]
55. Jung YW, Lim YK, Kim MS. Possibilities and limitations of online document tools for design collaboration: the case of Google Docs. In: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing. 2017 Presented at: CSCW '17; February 25-March 1, 2017; Portland, OR URL: https://doi.org/10.1145/2998181.2998297 [doi: 10.1145/2998181.2998297]
56. Walker J, Leveille S, Bell S, Chimowitz H, Dong Z, Elmore JG, et al. OpenNotes after 7 years: patient experiences with ongoing access to their clinicians' outpatient visit notes. J Med Internet Res 2019 May 06;21(5):e13876 [FREE Full text] [doi: 10.2196/13876] [Medline: 31066717]
57. Fossa AJ, Bell SK, DesRoches C. OpenNotes and shared decision making: a growing practice in clinical transparency and how it can support patient-centered care. J Am Med Inform Assoc 2018 Sep 01;25(9):1153-1159 [FREE Full text] [doi: 10.1093/jamia/ocy083] [Medline: 29982659]
58. Huang M, Fan J, Prigge J, Shah ND, Costello BA, Yao L. Characterizing patient-clinician communication in secure medical messages: retrospective study. J Med Internet Res 2022 Jan 11;24(1):e17273 [FREE Full text] [doi: 10.2196/17273] [Medline: 35014964]
59. Eschler J, Liu LS, Vizer LM, McClure JB, Lozano P, Pratt W, et al. Designing asynchronous communication tools for optimization of patient-clinician coordination. AMIA Annu Symp Proc 2015;2015:543-552 [FREE Full text] [Medline: 26958188]
60. Introducing ChatGPT. OpenAI. 2022 Nov 30. URL: https://openai.com/blog/chatgpt [accessed 2024-06-21]
61. Med-PaLM. Google. URL: https://sites.research.google/med-palm/ [accessed 2024-06-21]

XSL•FO
RenderX

62. Gilson A, Safranek CW, Huang T, Socrates V, Chi L, Taylor RA, et al. How does ChatGPT perform on the United States Medical Licensing Examination (USMLE)? The implications of large language models for medical education and knowledge assessment. JMIR Med Educ 2023 Feb 08;9:e45312 [FREE Full text] [doi: 10.2196/45312] [Medline: 36753318]
63. Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepaño C, et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. PLOS Digit Health 2023 Feb 9;2(2):e0000198 [FREE Full text] [doi: 10.1371/journal.pdig.0000198] [Medline: 36812645]
64. Singhal K, Tu T, Gottweis J, Sayres R, Wulczyn E, Hou L, et al. Towards expert-level medical question answering with large language models. arXiv Preprint posted online on May 16, 2023 [FREE Full text] [doi: 10.48550/arXiv.2305.09617]

Abbreviations

EHR: electronic health record
FHIR: Fast Healthcare Interoperability Resources
GenAI: generative artificial intelligence
PGD: patient-generated data
RQ: research question

Edited by A Kushniruk; submitted 01.09.23; peer-reviewed by S Young, N Finn; comments to author 15.11.23; revised version received 05.02.24; accepted 25.11.24; published 03.02.25.

Please cite as:
Nakikj D, Kreda D, Luthria K, Gehlenborg N
Patient-Generated Collections for Organizing Electronic Health Record Data to Elevate Personal Meaning, Improve Actionability, and Support Patient–Health Care Provider Communication: Think-Aloud Evaluation Study
JMIR Hum Factors 2025;12:e50331
URL: https://humanfactors.jmir.org/2025/1/e50331
doi:10.2196/50331
PMID:

©Drashko Nakikj, David Kreda, Karan Luthria, Nils Gehlenborg. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 03.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Mental Health Providers' Challenges and Solutions in Prescribing Over Telemedicine: Content Analysis of Semistructured Interviews

Julia Ivanova[1], PhD; Mollie R Cummins[1,2], RN, PhD; Hiral Soni[1], MS, MBA, PhD; Triton Ong[1], PhD; Brian E Bunnell[1,3], PhD; Esteban López[1], MD, MBA; Brandon M Welch[1,4], PhD

[1]Doxy.me Research, Doxy.me Inc, Charleston, SC, United States

[2]Department of Biomedical Informatics, College of Nursing and Spencer Fox Eccles School of Medicine, University of Utah, Salt Lake City, UT, United States

[3]Department of Psychiatry and Behavioral Neurosciences, Morsani College of Medicine, University of South Florida, Tampa, FL, United States

[4]Biomedical Informatics Center, Medical University of South Carolina, Charleston, SC, United States

Corresponding Author:
Julia Ivanova, PhD
Doxy.me Research
Doxy.me Inc
3rd Floor Suite 6 & 7
18 Broad Street
Charleston, SC, 29401
United States
Phone: 1 8444369963
Email: julia.ivanova@doxy.me

Abstract

Background: In response to the COVID-19 pandemic, the United States extended regulatory flexibilities to make telemedicine more accessible to providers and patients. Some of these flexibilities allowed providers to intake patients over telemedicine and prescribe certain scheduled medications without an in-person visit.

Objective: We aim to understand providers' parameters for their comfort in prescribing over telemedicine and report on solutions providers have adopted in response to potential barriers and challenges in prescribing via telemedicine.

Methods: As part of a larger mixed methods study between February and April 2024, we conducted 16 semistructured interviews with mental health providers who prescribe via telemedicine within the United States. We used the results of a web-based, cross-sectional survey to develop a codebook and support recruitment. We analyzed a subsection of the 16 interviews using content analysis to capture comfort, barriers, and workarounds in telemedicine prescribing. We reported codes by frequency and by provider.

Results: Participants were typically male (11/16, 69%), provided care mostly or completely over telemedicine (11/16, 69%), and were psychiatrists (8/16, 50%) or other physician (3/16, 19%). Providers' primary states (10/16, 62%) of practice included Oregon, Texas, New York, and California. The content analysis yielded a total of 234 codes, with three main codes—comfort (98/234, 41.9%), barriers or challenges (85/234, 36.3%), and workarounds or solutions (27/234, 11.5%)—and two subcodes—uncomfortable prescribing (30/98, 31%) and comfortable prescribing (68/98, 69%) over telemedicine. Participants reported being comfortable prescribing over telemedicine as long as they could meet their main parameters of working within their expertise, having access to needed patient health information, and being compliant with rules and regulations. Participants reported frustrations with e-prescription workflows and miscommunications with pharmacies. Solutions to ease frustrations and alleviate discomforts in prescribing over telemedicine included developing workflows to help patients complete laboratory tests and physical examinations and directly communicating with pharmacies.

Conclusions: By applying content analysis to the semistructured provider interviews, we found that physicians are comfortable prescribing via telemedicine when they feel they are practicing within their personal parameters for safety. While many providers experience frustrations such as miscommunication with pharmacies, these barriers appear to not prevent them from telemedicine prescribing. With expected changes in 2024 and 2025 to the US laws and regulations for telemedicine prescribing, we may see changes in provider comfort in prescribing.

XSL•FO
RenderX

(JMIR Hum Factors 2025;12:e65419) doi:10.2196/65419

KEYWORDS

telemedicine; telehealth; prescribe; prescription; drug; pharmacology; pharmacotherapy; pharmaceutical; medication; barrier; buprenorphine; mental health; digital health; informatics; qualitative analysis; content analysis; provider perspective; provider; experience; attitude; opinion; perception; perspective

Introduction

During the height of the COVID-19 pandemic, health care visits moved to telemedicine when possible to help stem the spread of infection [1]. In the United States, the public health emergency (PHE) and regulatory flexibilities were instrumental in streamlining this effort by waiving specific requirements, such as the providers' ability to see patients outside the states where they are licensed and to prescribe over telemedicine with fewer restrictions [2,3]. Due to these temporary changes, providers could prescribe via telemedicine less restrictively and determine what level of care they felt comfortable providing over telemedicine.

The US Drug Enforcement Administration extended prescribing flexibilities past the May 11, 2023, PHE deadline and is in the process of finalizing a rule regarding telemedicine prescribing [4,5]. There has been concern that waiving elements of the Ryan Haight Act, which previously required providers to meet with patients in person before prescribing certain controlled substances, would adversely impact patient outcomes [6,7]. The concurrent opioid epidemic has also provided a new level of scrutiny in the space of telemedicine and prescribing scheduled medication for the treatment of substance use disorder (SUD) [8-10]. As US providers await the promulgation of this final rule, evidence continues to grow regarding positive patient outcomes via mental health prescribing over telemedicine [11]. One Medicaid data study following over 90,000 patients showed remarkable positive outcomes from telemedicine initiation of buprenorphine treatment, including better odds of 90-day treatment retention than if initiation occurred in person [12]. In a scoping review from 2008 to March 18, 2021, reviewers found that telehealth technology in SUD treatment increased access and adherence to buprenorphine and generally showed higher patient satisfaction and comparability to in-person retention rates [11]. Especially within the space of buprenorphine prescription, evidence supports that telemedicine offers comparable treatment to in-person care and can be an effective method for increasing patient access to mental health treatments [8,11].

With the flexibilities, providers can now choose the extent to which they use telemedicine. One study reported psychiatrists felt telemedicine use in a hybrid scenario allowed for treatment adaptation based on client needs and access [13]. The study's participants emphasized their need to consider each patient's unique case in determining whether they could successfully leverage telemedicine to initiate care or ensure continued visits over time [13]. For example, one psychiatrist in the study noted the following:

> *When you have these patients who are at risk for falling out of care, if you don't offer them telemedicine, but [they] are also at risk for getting slightly suboptimal care when you do offer them telemedicine, it is a very case-by-case judgment call in terms of the risks and benefits of enabling the telemedicine.*

Such sentiments support the call for greater provider autonomy and options, as well as the focus on including telemedicine as part of the holistic health care model [13-15].

Though there is evidence supporting the safety of prescribing certain scheduled substances over telemedicine (even without an initial in-person visit) [8,11,12], there is a need for greater insight into mental health providers' experiences in prescribing scheduled medications during and after the PHE. A better understanding of providers' perspectives, experiences, and comfort in prescribing for their patients over telemedicine would build a more accurate understanding of what providers view as actual challenges and limitations. This knowledge would add to the discourse regarding updates in policy and guidelines, especially necessary for policy makers as they deliberate over proposed rules and regulations [16]. To the best of our knowledge, no previous studies have focused on mental health providers' experiences regarding general prescribing—including scheduled medications—over telemedicine. In early 2024, we conducted a mixed methods study that included a web-based, cross-sectional survey and semistructured interviews of mental health providers who prescribe via telemedicine. The survey aimed to assess provider perceptions, experiences, comfort, and perceived patient safety in prescribing medications over telemedicine [17]. The interviews aimed to get a better understanding of providers' general telemedicine use, challenges and solutions specific to prescribing over telemedicine, and providers' knowledge and experiences regarding compliance with prescribing over telemedicine. In this study, our objectives were (1) to understand better providers' parameters for safe prescribing over telemedicine and (2) to report practices or workflows providers have adopted to accomplish safe prescribing via telemedicine. The results of the study will help stakeholders reassess long-standing policies—made at the beginning of these technological advances—that may now be causing unintended harm as the technology has improved and become more accessible to laypeople.

Methods

Study Design

We performed exploratory content analysis of semistructured interviews with tele–mental health care providers as part of a larger, mixed methods study. We used an explanatory sequential design, building upon a previously reported, web-based, cross-sectional survey [17].

XSL•FO
RenderX

Interview Development

After an initial literature review, we developed the semistructured interview guide with consultation and pretesting from 5 providers. We estimated that interviews would take 60 minutes. The interview attempts to capture context regarding what factors affect providers' views and comfort in prescribing medications via different scenarios and combinations of in-person and telemedicine visits. The interview guide covers the topics of (1) the provider's practice, such as specialty, telemedicine use, patients, and general difficulties with telemedicine; (2) challenges and solutions concerning prescribing over telemedicine; and (3) knowledge pertaining to laws and regulations of prescribing to telemedicine (see Multimedia Appendix 1 for the interview guide). This study focuses specifically on answers to the challenges and solutions concerning prescribing over telemedicine (Multimedia Appendix 1).

Participants and Recruitment

We recruited participants for the semistructured interviews from among participants in a research survey. The survey sampling, recruitment, and results are reported elsewhere [17]. In brief, we recruited a sample of US tele–mental health care providers from a telemedicine research panel, TelehealthEngage [18]. A total of 115 participants fully completed the survey. Participants were predominantly White (83/115, 72.2%), non-Hispanic or Latino (97/115, 84.3%), female (66/115, 57.4%) providers who were physicians (69/115, 60%) seeing their clients most to all of the time (50%-100% of visits) over telemedicine (82/115, 71.3%) [17].

Participants who agreed to be contacted at the end of the survey were invited to participate in follow-up interviews. Of the 115 completed surveys, 59 (51.3%) participants showed interest in being interviewed. Research team member JI (a White, immigrant female individual with a background in medical anthropology, psychology, and biomedical informatics) sent out interview invitations and completed interviews between February and April 2024. We aimed to complete 20 interviews or until we reached content saturation [19].

Data Analysis

One author (JI) performed an exploratory content analysis of 6 interviews. The codebook for content analysis, consisting of three main codes and two subcodes, was specified on the basis of prior survey findings (Table 1) [17]. The codebook was developed to test our expectations that providers will feel comfortable when their personal parameters for patient safety have been met and when the telemedicine visit is within the scope of provider practice and expertise. We used content analysis of the interviews to identify codes related to personal parameters for comfort in prescribing over telemedicine, named barriers and challenges of telemedicine prescribing, and the types of workarounds providers have found useful within those situations [19].

Table 1. Demographics as reported by providers during the interview[a].

Demographics for content analysis	Value, n (%)
Sex (n=16)	
Male	11 (69)
Female	5 (31)
How often telemedicine is used (n=16)	
None to seldom	1 (6)
Fairly often (about half the time)	4 (25)
Mostly or completely	11 (69)
Provider type (n=16)	
Psychiatry	8 (50)
Family medicine or general practice physician	3 (19)
Nurse practitioner	3 (19)
Physician assistant	2 (13)
Type of practice (n=21)	
Private practice	12 (57)
Statewide health system	4 (19)
Community health center	2 (10)
Solo practice	1 (5)
City-run program	1 (5)
Outpatient rehabilitation center	1 (5)

[a]Some demographic values sum to more than 16 as some providers worked in multiple practices.

XSL•FO
RenderX

We transcribed and deidentified completed interviews using Dovetail software (Dovetail). We confirmed transcripts for accuracy prior to deleting video data of interviews and uploaded transcripts to MAXQDA 2024 software (MAXQDA) for analysis. A single author (JI) coded the interviews over three iterations using the previously developed codebook, and coauthor HS (an Asian, immigrant female individual with a background in biomedical informatics and human factors research in behavioral health) reviewed the codes and codebook for validity and accuracy. We resolved discrepancies through consensus to ensure consistency and accuracy in the qualitative method [20,21]. The unit of analysis was meaningful phrases. Although we calculated the total frequency of code instances, we primarily focused on the frequency of codes by individual participants [19]. Additionally, contextual subtopics within the main reported codes were often double coded—as participants provided multiple examples, explanations, and reasoning—resulting in percentages over 100. We reported participant characteristics in the aggregate to ensure participants' confidentiality. For context, we included participants' numbers (1 through 16) after each quote.

Ethical Considerations

This study was reviewed and approved by the BRANY Institutional Review Board (IRB00010793). We attained verbal informed consent for the interviews in addition to the written informed consent from the survey protocol. Participants had the opportunity to opt out of our research. Participants were compensated with a US $75 e-gift card for their time. While there is always a risk of selection and other compensation-related biases, the amount was deemed appropriate for US mental health professionals who were the focus of recruitment. We use deidentified data for the purposes of this research.

Results

Demographics

We completed 16 semistructured interviews between February 22 and April 26, 2024. There were 11 (69%) male and 5 (31%) female participants (see Table 1 for demographics). Participants primarily practiced in the following states: Oregon (n=3, 19%), Texas (n=3, 19%), New York (n=2, 12%), Illinois (n=2, 12%), California (n=1, 6%), Florida (n=1, 6%), Michigan (n=1, 6%), Washington (n=1, 6%), Indiana (n=1, 6%), and Tennessee (n=1, 6%). The majority (n=11, 69%) provided care mostly or completely over telemedicine (75%-100% of visits), 4 (25%) providers reported an even mix (40%-60% of telemedicine and in-person visits) in their hybrid practice often depending on the weather, and 1 (6%) provider reported almost complete in-person care except for emergencies. Providers included psychiatrists (n=8, 50%), family medicine or general practice physicians (n=3, 19%), nurse practitioners (n=3, 19%), or physician assistants (n=2, 12%). Providers reported working in private practice (n=12, 75%), statewide health systems or clinics (n=4, 25%), community health centers (n=2, 12%), a solo practice (n=1, 6%), a city-run program (n=1, 6%), and an outpatient rehabilitation center (n=1, 6%). Some providers worked in multiple positions or practice types.

Content Analysis

A total of 234 code instances were identified, corresponding to three main codes and two subcodes. Table 2 shows the frequency of code instances corresponding to each main code and subcode. Frequencies of instances of *generally uncomfortable* and *generally comfortable* subcodes add up to the total frequencies of the *comfort* main code instances.

Of the 98 (41.9%) out of 234 total code instances under *comfort* in prescribing medications over telemedicine, 69% (68/98 code instances) were related to describing positive comfort in prescribing from all 16 providers. However, the majority of providers (10/16, 62%) also reported varying discomfort (30/98, 31% code instances) in prescribing over telemedicine even though it was less frequently mentioned.

Providers reported (85/234, 36.3%) varied barriers and challenges (85 code instances, 14 providers), providing information regarding problems that both providers and their patients face when prescribing medications over telemedicine. Providers reported fewer (27/234, 11.5%) workarounds and solutions (27 code instances, 11 providers).

XSL•FO
RenderX

Table 2. Content analysis results with the frequency of instances by main codes and subcodes[a].

Name	Frequency of code instances, n	Definition	Example
Comfort (main code)	98	Any direct discussion regarding feelings of comfort or ease in prescribing over telemedicine	"I just don't wanna like, yeah, miss something. And like, actually probably more so with the Benzos than with, than with stimulants because I feel like the stimulants you can: see if people are kind of like … when people are on Benzos they might be more drowsy, that you're just not picking up with, on the, the, the, the virtual…" [Participant 1]
Generally uncomfortable (subcode)	30	Any direct discussion regarding feelings of discomfort or unease in prescribing over telemedicine	"I feel less comfortable with over telemedicine because sometimes just like the cadence of a conversation is hard. Sometimes if there's a lag in telehealth or like whatever it may be. As for prescriptions, I would say, like your level twos and level threes are ones that I would not consider prescribing unless I'd at least seen them once in person…" [Participant 3]
Generally comfortable (subcode)	68	Any direct discussion regarding positive feelings of comfort or ease in prescribing over telemedicine	"And I would say that it's really about the same, the actual prescribing is pretty darn smooth. So, I feel like the mechanics of the prescribing are the same. I do not prescribe differently because I, I've got a sense, I don't know this person as well because I'm not me. No, I feel like I know my patients really well. Certainly, before I'm gonna be prescribing, I know what the story is in my mind. I may be wrong, but I've got that level of certainty. So, I don't think the telehealth platform impacts my decision-making or the mechanics around prescribing significantly." [Participant 6]
Barriers and challenges (main code)	85	Any direct discussion where a provider mentions a problematic situation that affects their decision-making to prescribe over telemedicine or affects a patient's ability to receive treatment	"But I would say, yes, it's difficult to feel comfortable, directing someone in another state to something that they may need without really knowing where that resource, who that resource is." [Participant 7]
Workarounds and solutions (main code)	27	Any direct discussion where the provider describes a way they adapt their process or decision-making due to a barrier or challenge	"So, a lot of times I give them information on how to find a particular how to vet another practitioner, but that's about the best [I can do]." [Participant 7]

[a]Main code and subcode definitions were developed based on survey analysis.

Parameters for Comfort in Prescribing Over Telemedicine

Overview

Providers brought up three main parameters as they contextualized their comfort or discomfort with prescribing medications over telemedicine: (1) limits of telemedicine and personal expertise, (2) knowledge of patients and access to their relevant health information, and (3) liability concerns.

Limits of Telemedicine and Personal Expertise

Five (31%) of the 16 providers established that they felt comfortable when the interaction stayed within their perceived limits of telemedicine and occurred within their realm of expertise. One provider pointed out that these parameters essentially establish criteria for a successful telemedicine visit:

> *So the direct answer to the question is I feel very comfortable [prescribing over telemedicine]. But I think the reason that I feel very comfortable has to do with the patients: who I see and who I won't see on telemedicine. So again, doing this is like a philosophical thing. I'm a big believer in getting people to the right place. And a lot of times that's me and a lot of times that's not me when it comes to mental health and telehealth, I try to, there are things that I, I don't personally feel comfortable or think would be appropriate to, to treat. So in general, if someone is like in a manic state, if they're psychotic, if they're actively suicidal, like suicidal with a plan intending to act on it, that sort of stuff, I won't see them through telehealth or if I have a patient who starts to experience those things, I will refer them for a higher level of care and we'll be pretty insistent on it. Not that I'm not gonna see them anymore, but it's just I want to get you to the best place.* [Participant 13]

Regarding the parameter of experience and expertise, 3 (19%) providers reported feeling uncomfortable prescribing over telemedicine when they thought they did not have the expertise to treat a particular diagnosis. One provider described feeling nervous about such a scenario:

> *I don't have anybody that I'm treating for narcotic abuse or opiate stuff. So I, not that I wouldn't, but I don't have anybody and I don't have a lot of experience with that. So that would make me nervous, especially over telemed. I just because I don't do it much in person either.* [Participant 9]

XSL•FO
RenderX

In this case, the provider notes they would feel uncomfortable treating such a patient in person and that seeing such a case over telemedicine would increase that feeling.

Knowledge of Patients and Access to Their Relevant Health Information

This parameter encompassed several topics including working with established patients and accessing needed patient health information such as labs or physical exams. A total of 4 (25%) providers identified knowledge of a patient's history—especially in regard to them being an established patient—as a crucial element of feeling comfortable prescribing for them.

> *So, certain medication that requires good monitoring for that level that may make me not want to prescribe medications on this platform. But I don't think it has anything to do with virtual platforms, any providers or psychiatrists would do so in, in person also. So I'm not, I'm trying to think if, unless the only time I don't prescribe is if someone makes the first appointment and say I'm on this medicine and I need refill, I don't do that because that's not a good practice. It's like, OK, unless I have some data or previous records or not a very strong medication, then you can continue. But if someone comes with a specific request for medicines that are not appropriate, then I wouldn't consider that as an OK to do it as a virtual practice.* [Participant 10]

This provider noted that in the scenario provided, they would need more information about the patient before prescribing, regardless of whether the visit is in person or via telemedicine. Meanwhile, one of these providers mentioned the importance of generally knowing their patients and their health status as they determine whether telemedicine visits are sufficient for care:

> *...If you have somebody who's a fragile diabetic, you need to have a different level of observation for that person. But if you have a stable diabetic, then you can just continue to prescribe that whatever it might be…* [Participant 15]

Inherent in this explanation is that the provider felt they have sufficient knowledge regarding a patient's current and past health status and feeling personally able to diagnose and treat in the presenting situation.

A total of 4 (25%) providers commented on the importance of having established patients in a scenario to make providers feel more comfortable prescribing over telemedicine. One of these providers underlined the importance of having such a relationship, especially when seeing patients for SUD treatment, as they may be able to pick up on potential concerns requiring more oversight via telemedicine or in person:

> *I mean, people with substance use problems. Now that brings a whole host of challenges itself because of confidence in, in history, you know, and confidence of just general reporting ... So I have a special, I have a different kind of relationship with [specific patient with SUD]. I can tell when somebody's bullshitting. And I can call them pretty well on it. I rarely get blindsided because I've been doing this for a long time. So I come with enough experience, I think also that allows that to happen. But suffice to say that if I have an issue with someone there, I'll put more controls over it, and we'll go to task if needed. If they need to come in, I need to see them in person...* [Participant 15]

In this case, the provider leveraged not only their years of experience in the specific field (expertise parameter) and having known the patient for a long time but also their hybrid practice, as they can have the patient come for an in-person visit.

A total of 3 (19%) providers mentioned they feel comfortable prescribing over telemedicine because they have such an option. In addition to this, 4 (25%) providers also pointed out that having a way to check on their patient's safety, whether it be having local resources for the patient or being able to view a Prescription Drug Monitoring Program, influences their comfort positively. One of these 4 providers described that knowing they can get patient health information quickly and seamlessly made them comfortable prescribing for an individual:

> *If they have established care, either someone who can update me on their vitals or regular monthly visits and have some kind of formal evaluation for me to backtrack the diagnosis. And then there is a way to order urine drug screens to make sure they're not abusing any other drugs on the top of. So, if that makes it an easy flow for us, then, yeah. Yeah. I think that's not a big problem. It's just how to coordinate that care sometimes makes it harder...* [Participant 10]

Additionally, 4 (25%) providers stated the importance of developing and adhering to a protocol with their patients. Providers discussed this concept within the patient information parameter in everything from the formal quality of care processes to simply expecting to see a patient in person for all intakes. Adhering to protocols appears to be driven by the need for patient health status and history, as providers often discussed continuity of care and procurement of labs and exams.

> *So, I think patients who are, really complex, I feel less comfortable with prescribing over telemedicine because sometimes just, like, the cadence of a conversation is hard. Sometimes if there's a lag in telehealth or whatever it may be. As for prescriptions, I would say, your level two's and level three's are ones that I would not consider prescribing unless I'd at least seen them once in person. I know personally, our clinic does random drug tests on our patients...* [Participant 3]

Liability Concerns

A total of 5 (31%) providers noted that they maintain their comfort in prescribing over telemedicine by simply refusing to prescribe certain controlled substances or not accepting specific diagnoses in their practice, especially due to their wariness of laws and regulations. One of these providers explained that while they treat SUDs, they do not prescribe certain medications and are aware of prescribing barriers:

XSL•FO
RenderX

> *...I do see a lot of patients with substance use disorder...just the fact that it's [buprenorphine] more accessible now is amazing for these patients. I've heard there are still some barriers with methadone in particular...it hasn't been an issue for me because I don't prescribe that medication in particular.* [Participant 16]

Indeed, 2 (12%) providers indicated they avoid prescribing over telemedicine unless they feel they have to for the patient.

> *I generally refrain from prescribing anything scheduled over telemedicine ... just because of the, the insane liability with it ... the unbalanced liability, I would say that physicians shoulder in this industry.* [Participant 7]

Here, the provider also touched on the issue of compliance and liability inherent in this situation.

Barriers and Workarounds in Prescribing Over Telemedicine

Of the 85 (36.3%) out of the 234 total code instances that describe barriers (14 providers), 37 (44%) of the 85 instances specifically mentioned that the e-prescription platform and the pharmacies caused challenges in prescribing medicines or picking up medicines (13 providers). There were 17 (20%) instances of providers describing difficulties with the actual e-prescription platform they use.

> *But Epic, they started putting in all these like hard blocks like you need to click this before you can do what you're trying to...I think the way that they went about it sometimes, it was really like a hassle for the workflow when you add it all together.* [Participant 2]

Providers also often noted their patients' prescription order was not received or not fulfilled by the pharmacy due to shortages in medicines or misunderstandings in prescribing over telemedicine (20/85, 24%). One provider noted how such a situation acted as a barrier for patients receiving their necessary prescriptions:

> *But I had a pharmacy once, call me that said that...they weren't going to allow my patient to fill a pack of the prescription because I hadn't provided an estimated glomerular filtration rate or creatinine clearance with my prescription. But like that's because they now have the ability to prescribe that on their own. But I sent the prescription. I'm assuming the responsibility. I already checked that all you're doing is throwing up another barrier to this patient, getting their medication when you do that, which I get. You don't want to be audited or whatever, but you're not the one prescribing it. I am, you know what I mean? I just like the pendulum has swung so far the other way since the opioid epidemic that they're really trying to tighten the noose and on all this stuff, and I just think it's going to end up harming patients who are calling for their beta blockers or, you know, whatever.* [Participant 2]

Providers reported shortages of certain medicines, usually stimulants, in pharmacies, requiring patients to find pharmacies with stock and providers to resend prescriptions. Providers' complaints regarding the e-prescription platforms and pharmacies' inability to provide the medications ranged from workflow nuisances to concerns that their patients could not receive necessary medications promptly.

Other mentioned barriers (49/85, 58%) were related to general telemedicine concerns and issues such as patient health concerns (17/85, 20% codes; 11/16, 69% providers), rules and regulations (16/85, 19% codes; 7/16, 44% providers), and reimbursement concerns (9/85, 11% codes; 2/16, 13% providers). Patient safety concerns mainly touched on worries that patients are not having their follow-up appointments, exams, and labs in a timely manner, which can affect patient health outcomes, as well as the ability of providers to prescribe medications for them.

> *I will put a child on a medication and I'll say I need to have you call me in two or three days because this medicine doesn't take a long time to start working. And I need to know so we can make adjustments. And invariably 99% of the time they wait till the next month, scheduled appointment and we could have made some, you know, needed changes. And that was difficult even when we were in the office...* [Participant 11]

Concerns regarding rules and regulations more closely align with providers seeing patients across state borders. Some providers felt such regulations hindered their ability to treat patients but also created unnecessary frustrations.

> *So, yeah, it's interesting navigating all of that because, and honestly, I think it's just so stupid having state-dependent licensing system when they all use the same criteria.* [Participant 7]

While providers recognized the need for laws and regulations, their current struggles often placed them in positions having to choose between seeing their (often established) patients or complying with regulations.

> *It reminds me of another patient, he's in college in North Carolina and, he had really good, he had a psychiatrist through his school but then the school had canceled the contract with that [psychiatrist] and he's like, what now am I supposed to do? And like, he wasn't exactly, he wasn't unstable but he wasn't stable. Like, you know, in that gray zone and he's like, can't you see me?...And I would love to see you. I feel my heart goes out to you, but like, you're in North Carolina, like, I...I, I will see you all through Christmas break, summer break. But like, I don't know what to do. Like, if you're having a crisis, definitely call me like, and I will, you know, I will just override it. But like, if you're going there and, you know, like you need the help, like you gotta, I don't know, I don't know what to tell you. Like, I feel terrible if you.* [Participant 1]

XSL•FO
RenderX

Providers often reported that mobile patients such as college students are the ones most strongly affected by these types of laws and regulations.

Providers reported far fewer workarounds or solutions (27/234, 11.5% codes; 11/16, 69% providers) to mentioned barriers relating to prescribing over telemedicine. Direct responses to barriers and challenges included—as subtopics—streamlining how they receive patient health data (8/27, 30%), finding new pharmacies for patients to use (5/27, 19%), using reminders and new ways to connect with patients quickly (5/27, 19%), responding to regulatory pressures (3/27, 11%), and making individual adaptations in their practice (6/27, 22%; due to double coding of subtopics, percentages add up over 100) The most common solution providers reported was finding an easier way to receive laboratory tests, examinations, and vitals from their patients (8/27, 30% codes; 5/16, 31% providers). While providers were able to direct their patients to local laboratories or request that their patients invest in a blood pressure cuff, one provider noted their practice found certain tests to be beneficial for telemedicine visits:

> *Hm, we do drug screens online. So, we do saliva testing online and drug screens. And so that's if we don't always have to have people go pee somewhere or go to the lab, they can do those at home. So, I guess that's something that we found just as a, not really a workaround, well, kind of a workaround, and I guess to allow them to stay at home, but just to incorporate it as something that's a little bit more convenient.* [Participant 5]

Providers further noted the emphasis on streamlining patient processes by using novel ways to communicate with them regarding simple questions or requests. For example, one provider noted the following:

> *[Patients] can just text me a question...A few patients of mine have learned that they come and sit in the waiting room if they have a question. And so, I see them sitting in my waiting room, and sometimes I'm done. I'm like, "What's up?" And they say, "Yes, you know, hey, I need a refill.*" [Participant 12]

In this case, the provider sought to have secure, simple messaging with a patient. Other solutions related to streamlining processes with patients included actively talking with pharmacies in the patients' area (5/27, 19%), moving to a new e-prescription platform (1/27, 4%), and creating a protocol for helping their out-of-state patients find health resources (1/27, 4%).

Providers mentioned two differing approaches to handling the barriers created by laws and regulations in prescribing over telemedicine. Some (2/16, 13%) providers chose to unequivocally adopt the strictest state and federal regulations within their practice, and another explained that they would continue with a process that works for their patients but was admittedly in the gray zone legally.

Content analysis of barriers affecting provider comfort in prescribing over telemedicine identified 85 code instances over 16 interviews. With 27 code instances of workarounds and solutions directly relating to these noted barriers, the qualitative analysis identifies areas in telemedicine where progress still needs to be made.

Discussion

Principal Findings

By leveraging the results of our more extensive mixed methods study [17], we framed our content analysis to provide needed context for understanding mental health providers' comfort in prescribing over telemedicine. This qualitative study identified vital parameters that constitute comfort for mental health providers who prescribe via telemedicine. Our results show that providers feel most comfortable prescribing over telemedicine when they are practicing within its inherent limitations and their own professional expertise, with access to patients' relevant health information, and within clearly defined legal and regulatory confines. In addition, we identified related barriers that affect these parameters and the potential solutions that providers have used to alleviate these challenges. Content analysis revealed providers found the e-prescription platforms and the miscommunications with pharmacies receiving prescriptions as substantial sources of frustration and concern in the prescribing process. Ultimately, providers prescribe medications over telemedicine within the bounds of their perceived comfort parameters but would like to see changes in streamlining the prescribing workflow, assurance of timely patient examinations and laboratory tests, and clarity in the laws and regulations surrounding prescribing over telemedicine.

Overall, even while providers reported feeling comfortable prescribing over telemedicine, they were aware their comfort was due to practicing within their established personal parameters. Providers noted their comfort in prescribing over telemedicine was attributable to their specialty and providing care to patients who require specific types of medications that providers may be generally hesitant to prescribe. Research using claims data between 2020 and 2023 shows that certain specialties (ie, psychiatry) were seen to initiate treatment for alcohol use disorder over telemedicine more than others (ie, primary care physicians) [22]. Within the same study, researchers observed a different rate of telemedicine initiations dependent on the types of medications being prescribed (eg, naltrexone at 14.6% vs topiramate at 11.8%). Our results support these findings as psychiatrists are better versed in treating mental health disorders and have tacit knowledge regarding certain medications and patient symptoms. Unsurprisingly, in this study, some providers emphasized that their comfort in prescribing decreased when faced with diagnoses they have little experience treating. From a 2023 review of pandemic-era research on telemedicine use in SUD treatment in the United States, authors noted providers would vary the number of days (supply) of prescribed medication based on their judgment of each patient interaction and context [23-26]. These findings are supported by our results as providers work within their personal parameters and determine how to proceed with prescribing based on their interaction with each patient.

Additionally, providers' personal parameters likely echo current guidelines regarding telemedicine use and limitations [27,28].

XSL•FO
RenderX

Such guidelines proffered protocols and practices based on earlier telemedicine research and stricter rules and regulations [29]. Considering that guidelines emphasize compliance with state and federal laws and regulations, we may be seeing providers' personal parameters reflecting these policies even when they feel comfortable and capable of prescribing over telemedicine [30]. Updates of such guidelines would help reframe providers' understandings of telemedicine prescribing limitations and liabilities; however, lagging legal progress limits the ability for such revisions [31,32]. With the loosening of laws and regulations regarding telemedicine prescribing, new evidence has shown that providers can care for their patients effectively while maintaining positive patient outcomes [29].

Providers also emphasized the importance of access to patient's health information, especially timely updates of exams and labs needed to prescribe certain medications such as stimulants. Certainly, the content analysis showed multiple workarounds focused on streamlining the process of accessing patient health information. Understandably, providers also noted that they feel more comfortable working with established patients, where patient health history and a personal repertoire with the patient can provide context to a telemedicine visit. Interview results suggest that an established patient—via telemedicine or in-person visit—is a central factor in how providers view their needs for timely patient health updates. This parameter is relevant for in-person and telemedicine visits. Nevertheless, with the inflated fear of liability when using telemedicine, providers show increased concern when operating outside their parameters for comfort in telemedicine prescribing [33]. Similar findings show providers may feel that additional assessment, such as physical examination, reinforces their decision-making and allays liability fears [34].

Our content analysis shows that the ultimate tipping point in whether providers prescribe via telemedicine is whether the clinical situation falls within the provider's personal parameters for comfort in prescribing. The obstacles we identified in the study, such as the e-prescription platform errors, were predominantly situations that caused frustrations in the prescribing process but were insufficient reasons for providers to discontinue prescribing medications. Of course, particular barriers, such as miscommunications with pharmacies and not receiving patient examinations or laboratory tests promptly, directly impact patient safety and outcomes. Our results support prior research showing that providers and patients faced challenges with e-prescriptions where pharmacies faced a dearth of certain medications such as buprenorphine [23,35,36]. Providers emphasized the importance of patient safety with their noted solutions: the most frequently mentioned solutions revolved around streamlining the receipt of patient health information and ensuring patient access to needed medications. Given the current, typical support infrastructure in the United States, telemedicine is not as effective for treating disorders that require laboratory examinations [37]. However, telemedicine is becoming a permanent mode of health care delivery, and successful solutions for assessment, such as diagnostic testing and information access, will be a priority for providers and other stakeholders. As a result, there is a growing support infrastructure including at-home laboratory options ranging from technicians coming to your home to home tests that can be done during a telemedicine visit [38,39]. While large health institutions such as Yale School of Medicine may have such resources available for their patients [38], small clinics or solo providers face difficulties in connecting their patients to these resources, especially in relation to an individual patient's insurance coverage [39].

Our results show that providers prescribe within the scope of their personal parameters determined by perceived limitations of telemedicine and professional expertise, provider access to relevant patient health information, and liability concerns. Once providers met their individual parameters, they expressed higher comfort in prescribing over telemedicine. With an increase in legal clarity and simplification of prescribing over telemedicine, providers may feel more capable and comfortable with the process, as guidelines would reflect the loosening of restrictions [30]. These personal parameters exist in the context of current legal and regulatory considerations; therefore, making PHE telemedicine-related policy flexibilities permanent may help expand the situations in which providers feel comfortable prescribing. Such progress may also improve patient access to mental health care services—providers may feel more comfortable expanding their patient base for telemedicine, thus alleviating access and inequity issues for mental health care treatment [40].

Limitations

This study used interviews from providers recruited through nonprobability convenience sampling via the TelehealthEngage research panel. Our findings are not generalizable to all US mental health providers. It reflects the view of interview participants, including providers from solo, small private practices; statewide health systems and clinics; and other types of mental health care settings. Additionally, as we included all prescribing mental health providers, the prescribing challenges, workarounds, and level of comfort we captured in the study may differ for other providers based on their practice focus. Future research should include a larger sample of participants including those practicing in larger health systems and academic health sciences centers, who were underrepresented in our sample.

Conclusions

Through content analysis of semistructured provider interviews, we determined that the participating providers feel comfortable prescribing over telemedicine when they practice within perceived limits of telemedicine and their own expertise, with sufficient access to patient health information, and when they do not have liability concerns. Providers mentioned multiple hindrances to prescribing over telemedicine. However, the ultimate reason they refer patients to other providers or convert encounters from telemedicine to in-person visits is that their comfort parameters are not met, and thus, patient safety is at risk. Future research, including inductive-deductive qualitative analysis of the interviews, will help create a more robust understanding of provider perspectives on telemedicine prescribing and inform future implementation and policy of prescribing over telemedicine.

XSL•FO
RenderX

Acknowledgments

We gratefully acknowledge Beju Shah for his assistance in developing the interview guide. Funding for this research was provided through Doxy.me, Inc.

Conflicts of Interest

BMW is a shareholder, and JI, HS, TO, BEB, EL, and MRC are employees of Doxy.me Inc, a commercial telemedicine company.

Multimedia Appendix 1
Semistructured interview guide.
[DOCX File , 15 KB - humanfactors_v12i1e65419_app1.docx]

References

1. Fiscal year 2024: budget in brief. U.S. Department of Health and Human Services. URL: https://www.hhs.gov/sites/default/files/fy-2024-budget-in-brief.pdf [accessed 2024-02-01]
2. CARES Act: AMA COVID-19 pandemic telehealth fact sheet. American Medical Association. URL: https://www.ama-assn.org/delivering-care/public-health/cares-act-ama-covid-19-pandemic-telehealth-fact-sheet [accessed 2024-04-05]
3. VanderWerf M, Bernard J, Barta DT, Berg J, Collins T, Dowdy M, et al. Pandemic action plan policy and regulatory summary telehealth policy and regulatory considerations during a pandemic. Telemed e-Health 2022;28(4):457-466 [FREE Full text] [doi: 10.1089/tmj.2021.0216] [Medline: 34265216]
4. DEA hosts public listening sessions on telemedicine regulations. United States Drug Enforcement Administration. URL: https://www.dea.gov/press-releases/2023/08/07/dea-hosts-public-listening-sessions-telemedicine-regulations [accessed 2023-09-21]
5. DEA extends telemedicine flexibilities for prescribing of controlled medications: second time is the charm. Foley. URL: https://www.foley.com/insights/publications/2023/10/dea-extends-telemedicine-prescribing-controlled/ [accessed 2024-04-10]
6. Huskamp HA, Riedel L, Uscher-Pines L, Busch AB, Barnett ML, Raja P, et al. Initiating opioid use disorder medication via telemedicine during COVID-19: implications for proposed reforms to the ryan haight act. J Gen Intern Med 2022;37(1):162-167 [FREE Full text] [doi: 10.1007/s11606-021-07174-w] [Medline: 34713386]
7. Niyibizi A, Haveric A, Irio G. Telehealth in opioid use disorder treatment: policy considerations for expanding access to care. J Osteopath Med 2024;124(4):179-181 [FREE Full text] [doi: 10.1515/jom-2023-0215] [Medline: 37976470]
8. Hailu R, Mehrotra A, Huskamp HA, Busch AB, Barnett ML. Telemedicine use and quality of opioid use disorder treatment in the US during the COVID-19 pandemic. JAMA Netw Open 2023;6(1):e2252381 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.52381] [Medline: 36692880]
9. Mattocks KM, Moore DT, Wischik DL, Lazar CM, Rosen MI. Understanding opportunities and challenges with telemedicine-delivered buprenorphine during the COVID-19 pandemic. J Subst Abuse Treat 2022;139:108777 [FREE Full text] [doi: 10.1016/j.jsat.2022.108777] [Medline: 35346533]
10. Lin LA, Fortney JC, Bohnert AS, Coughlin LN, Zhang L, Piette JD. Comparing telemedicine to in-person buprenorphine treatment in U.S. veterans with opioid use disorder. J Subst Abuse Treat 2022;133:108492 [FREE Full text] [doi: 10.1016/j.jsat.2021.108492] [Medline: 34175175]
11. Guillen AG, Reddy M, Saadat S, Chakravarthy B. Utilization of telehealth solutions for patients with opioid use disorder using buprenorphine: a scoping review. Telemed e-Health 2022;28(6):761-767. [doi: 10.1089/tmj.2021.0308] [Medline: 34714172]
12. Hammerslag LR, Mack A, Chandler RK, Fanucchi LC, Feaster DJ, LaRochelle MR, et al. Telemedicine buprenorphine initiation and retention in opioid use disorder treatment for medicaid enrollees. JAMA Netw Open 2023;6(10):e2336914 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.36914] [Medline: 37851446]
13. Uscher-Pines L, Parks AM, Sousa J, Raja P, Mehrotra A, Huskamp HA, et al. Appropriateness of telemedicine versus in-person care: a qualitative exploration of psychiatrists' decision making. Psychiatr Serv 2022;73(8):849-855 [FREE Full text] [doi: 10.1176/appi.ps.202100519] [Medline: 35080419]
14. Cummins MR, Ivanova J, Ong T, Soni H, Barrera JF, Wilczewski H, et al. Will the United States pass on telemedicine progress? JAMIA Open 2024;7(1):ooae016 [FREE Full text] [doi: 10.1093/jamiaopen/ooae016] [Medline: 38410742]
15. Meese KA, Hall AG, Feldman SS, Colón-López A, Rogers DA, Singh JA. Physician, nurse, and advanced practice provider perspectives on the rapid transition to inpatient and outpatient telemedicine. Telemed Rep 2022;3(1):7-14 [FREE Full text] [doi: 10.1089/tmr.2021.0034] [Medline: 35720449]
16. Telemedicine can change care for the better—with the right rules. Harvard Medical School. URL: https://hms.harvard.edu/news/telemedicine-can-change-care-better-right-rules [accessed 2024-04-25]

XSL•FO
RenderX

17. Cummins M, Ivanova J, Soni H, Robbins Z, Bunnell BE, López E, et al. Telemedicine prescribing: a national, cross-sectional survey of U.S. mental health care providers. JMIR Preprints Preprint posted online on December 21, 2024 [FREE Full text] [doi: 10.2196/preprints.63251]
18. TelehealthEngage. Doxy.me. URL: https://doxy.me/en/ [accessed 2025-01-11]
19. Bernard HR, Wutich A, Ryan GW. Analyzing Qualitative Data: Systematic Approaches. Thousand Oaks, CA: SAGE Publications; 2016.
20. Campbell JL, Quincy C, Osserman J, Pedersen OK. Coding in-depth semistructured interviews. Sociol Methods Res 2013;42(3):294-320. [doi: 10.1177/0049124113500475]
21. Raskind IG, Shelton RC, Comeau DL, Cooper HLF, Griffith DM, Kegler MC. A review of qualitative data analysis practices in health education and health behavior research. Health Educ Behav 2019;46(1):32-39 [FREE Full text] [doi: 10.1177/1090198118795019] [Medline: 30227078]
22. Huskamp HA, Uscher-Pines L, Raja P, Normand ST, Mehrotra A, Busch AB. Telemedicine for initiation of alcohol use disorder medications. JAMA Netw Open 2024;7(9):e2431594 [FREE Full text] [doi: 10.1001/jamanetworkopen.2024.31594] [Medline: 39230907]
23. Lin C, Pham H, Zhu Y, Clingan SE, Lin L, Murphy SM, et al. Telemedicine along the cascade of care for substance use disorders during the COVID-19 pandemic in the United States. Drug Alcohol Depend 2023;242:109711 [FREE Full text] [doi: 10.1016/j.drugalcdep.2022.109711] [Medline: 36462230]
24. Uscher-Pines L, Sousa J, Raja P, Mehrotra A, Barnett M, Huskamp HA. Treatment of opioid use disorder during COVID-19: experiences of clinicians transitioning to telemedicine. J Subst Abuse Treat 2020;118:108124 [FREE Full text] [doi: 10.1016/j.jsat.2020.108124] [Medline: 32893047]
25. Nordeck CD, Buresh M, Krawczyk N, Fingerhood M, Agus D. Adapting a low-threshold buprenorphine program for vulnerable populations during the COVID-19 pandemic. J Addict Med 2021;15(5):364-369 [FREE Full text] [doi: 10.1097/ADM.0000000000000774] [Medline: 33177436]
26. Duncan A, Sanders N, Schiff M, Winkelman TN. Adaptations to jail-based buprenorphine treatment during the COVID-19 pandemic. J Subst Abuse Treat 2021;121:108161 [FREE Full text] [doi: 10.1016/j.jsat.2020.108161] [Medline: 33371945]
27. Smith K, Ostinelli E, Macdonald O, Cipriani A. COVID-19 and telepsychiatry: development of evidence-based guidance for clinicians. JMIR Ment Health 2020;7(8):e21108 [FREE Full text] [doi: 10.2196/21108] [Medline: 32658857]
28. Gough F, Budhrani S, Cohn E, Dappen A, Leenknecht C, Lewis B, et al. ATA practice guidelines for live, on-demand primary and urgent care. Telemed e-Health 2015;21(3):233-241 [FREE Full text] [doi: 10.1089/tmj.2015.0008] [Medline: 25658882]
29. Ivanova J, Ong T, Wilczewski H, Cummins M, Soni H, Barrera J, et al. Telemental health care guidelines from the first 2 years of the COVID-19 pandemic: a scoping review using content and thematic analyses. JMIR Preprints Preprint posted online on January 18, 2024 [FREE Full text] [doi: 10.2196/preprints.56534]
30. Krawczyk N, Rivera BD, King C, Dooling BCE. Pandemic telehealth flexibilities for buprenorphine treatment: a synthesis of evidence and policy implications for expanding opioid use disorder care in the United States. Health Aff Sch 2023;1(1):qxad013 [FREE Full text] [doi: 10.1093/haschl/qxad013] [Medline: 38145115]
31. Double secret protection: bridging federal and state law to protect privacy rights for telemental and mobile health users. Duke Law Journal. URL: https://dlj.law.duke.edu/article/double-secret-protection-sherman-vol67-iss5/ [accessed 2023-05-22]
32. Tong KW. Telehealth as a double-edged sword: lessons from court cases to gain understanding of medico-legal risks. Med Law 2019;38(1):85-106 [FREE Full text]
33. Nguyen M, Waller M, Pandya A, Portnoy J. A review of patient and provider satisfaction with telemedicine. Curr Allergy Asthma Rep 2020;20(11):72. [doi: 10.1007/s11882-020-00969-7] [Medline: 32959158]
34. Riley PE, Fischer JL, Nagy RE, Watson NL, McCoul ED, Tolisano AM, et al. Patient and provider satisfaction with telemedicine in otolaryngology. OTO Open 2021;5(1):2473974X20981838. [doi: 10.1177/2473974x20981838]
35. Castillo M, Conte B, Hinkes S, Mathew M, Na CJ, Norindr A, et al. Implementation of a medical student-run telemedicine program for medications for opioid use disorder during the COVID-19 pandemic. Harm Reduct J 2020;17(1):88 [FREE Full text] [doi: 10.1186/s12954-020-00438-4] [Medline: 33203460]
36. Tofighi B, McNeely J, Walzer D, Fansiwala K, Demner A, Chaudhury CS, et al. A telemedicine buprenorphine clinic to serve New York City: initial evaluation of the NYC public hospital system's initiative to expand treatment access during the COVID-19 pandemic. J Addict Med 2022;16(1):e40-e43 [FREE Full text] [doi: 10.1097/ADM.0000000000000809] [Medline: 33560696]
37. Haleem A, Javaid M, Singh RP, Suman R. Telemedicine for healthcare: capabilities, features, barriers, and applications. Sens Int 2021;2:100117 [FREE Full text] [doi: 10.1016/j.sintl.2021.100117] [Medline: 34806053]
38. Corcoran TP. Mobile lab helps Yale pathology establish relationships, build trust with community. Yale School of Medicine. URL: https://medicine.yale.edu/pathology/news-article/mobile-lab-helps-yale-pathology-establish-relationships-build-trust-with-community/ [accessed 2024-07-10]
39. Drake K. At-home blood tests are becoming a trend. What it means for healthcare. Healthline. URL: https://www.healthline.com/health-news/at-home-blood-tests-are-becoming-a-trend-what-it-means-for-healthcare [accessed 2024-07-10]

XSL•FO
RenderX

40. Garett R, Young SD. Potential effects of digital inequality on treatment seeking for opioid use disorder. Int J Mental Health Addict 2023;21(2):890-895 [FREE Full text] [doi: 10.1007/s11469-021-00629-5] [Medline: 34466133]

Abbreviations

PHE: public health emergency
SUD: substance use disorder

Edited by A Kushniruk; submitted 14.08.24; peer-reviewed by J Ong, T Khodaveisi; comments to author 14.10.24; revised version received 10.12.24; accepted 20.12.24; published 20.03.25.

Please cite as:
Ivanova J, Cummins MR, Soni H, Ong T, Bunnell BE, López E, Welch BM
Mental Health Providers' Challenges and Solutions in Prescribing Over Telemedicine: Content Analysis of Semistructured Interviews
JMIR Hum Factors 2025;12:e65419
URL: https://humanfactors.jmir.org/2025/1/e65419
doi:10.2196/65419
PMID:

©Julia Ivanova, Mollie R Cummins, Hiral Soni, Triton Ong, Brian E Bunnell, Esteban López, Brandon M Welch. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 20.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Health Care Professionals' Perspectives on Using eHealth Tools in Advanced Home Care: Qualitative Interview Study

Eric Vincent Rivas[1], MMSc, MD; Ulf Lesley[1], MSc, MMSc; Nadia Davoody[1], MSc, PhD

Health Informatics Centre, Department of Learning, Informatics, Management and Ethics, Karolinska Institutet, Stockholm, Sweden

Corresponding Author:
Nadia Davoody, MSc, PhD
Health Informatics Centre
Department of Learning, Informatics, Management and Ethics
Karolinska Institutet
Tomtebodavägen 18a
Stockholm, SE-171 77
Sweden
Phone: 46 852486486
Email: nadia.davoody@ki.se

Abstract

Background: The rising demand for advanced home care services, driven by an aging population and the preference for aging in place, presents both challenges and opportunities. While advanced home care can improve cost-effectiveness and patient outcomes, gaps remain in understanding how eHealth technologies can optimize these services. eHealth tools have the potential to offer personalized, coordinated care that increases patient engagement. However, research exploring health care professionals' (HCPs) perspectives on the use of eHealth tools in advanced home care and their impact on the HCP-patient relationship is limited.

Objective: This study aims to explore HCPs' perspectives on using eHealth tools in advanced home care and these tools' impact on HCP-patient relationships.

Methods: In total, 20 HCPs from 9 clinics specializing in advanced home care were interviewed using semistructured interviews. The discussions focused on their experiences with 2 eHealth tools: a mobile documentation tool and a mobile preconsultation form. The data were analyzed using content analysis to identify recurring themes.

Results: The data analysis identified one main theme: optimizing health care with eHealth; that is, enhancing care delivery and overcoming challenges for future health care. Two subthemes emerged: (1) enhancing care delivery, collaboration, and overcoming adoption barriers and (2) streamlining implementation and advancing eHealth tools for future health care delivery. Five categories were also identified: (1) positive experiences and benefits, (2) interactions between HCPs and patients, (3) challenges and difficulties with eHealth tools, (4) integration into the daily workflow, and (5) future directions. Most HCPs expressed positive experiences with the mobile documentation tool, highlighting improved efficiency, documentation quality, and patient safety. While all found the mobile preconsultation form beneficial, patient-related factors limited its utility. Regarding HCP-patient relationships, interactions with patients remained unchanged with the implementation of both tools. HCPs successfully maintained their interpersonal skills and patient-centered approach while integrating eHealth tools into their practice. The tools allowed more focused, in-depth discussions, enhancing patient engagement without affecting relationships. Difficulties with the tools originated from tool-related issues, organizational challenges, or patient-related complexities, occasionally affecting the time available for direct patient interaction.

Conclusions: The study underscores the importance of eHealth tools in enhancing advanced home care while maintaining the HCP-patient relationship. While eHealth tools modify care delivery techniques, they do not impact the core dynamics of the relationships between HCPs and patients. While most of the HCPs in the study had a positive attitude toward using the eHealth tools, understanding the challenges they encounter is crucial for improving user acceptance and success in implementation. Future development should focus on features that not only improve efficiency but also actively enhance HCP-patient relationships, such as facilitating more meaningful interactions and supporting personalized care in the advanced home care setting.

(JMIR Hum Factors 2025;12:e60582) doi:10.2196/60582

XSL•FO
RenderX

KEYWORDS

eHealth; mobile health; mHealth; advanced home care; content analysis; nurse; staff-patient relationship; aging population; patient engagement; personalized care; patient experience

Introduction

Advanced Home Care

Advanced home care, also known as advanced health care in the home, involves providing medical care or treatment directly at a patient's home, serving as an alternative care to inpatient care. The World Health Organization defines home care as "an array of health and social support services provided to clients in their residence. Such coordinated services may prevent, delay, or be a substitute for temporary or long-term institutional care" [1]. According to a study by Barakat et al [2], the components of home care include preventive actions and assessments, postdischarge actions, and evaluations, with objectives focused on enhancing or maintaining the quality of life, optimizing functional health status, and promoting independence. The composition of health care professionals (HCPs) providing home care varies depending on the patient's needs and the extent of services they require. This team may involve various HCPs, such as physicians, nurses, assistant nurses, physical therapists, occupational therapists, speech-language pathologists, home health aides, home infusion nurses (specialized nurses who administer intravenous medications in patients' homes), hospice caregivers, and medical social workers [3]. Patients receiving home care may have a wide range of health conditions and diseases, including chronic conditions, such as diabetes, hypertension, and heart disease; neurological conditions, such as Parkinson disease, dementia, and stroke; respiratory conditions, such as chronic obstructive pulmonary disease, asthma, and sleep apnea; cancer; individuals requiring wound care; and those in need of palliative care [3].

The increasing demand for home care services is driven not only by an aging population but also by the growing preference among many individuals to age in place, remaining in their own homes. The primary goal of advanced home care is to enhance patients' quality of life and clinical outcomes while simultaneously reducing health care costs and hospital readmissions [4].

Use of Electronic Health in Advanced Home Care

Organizational suppliers of health care are increasingly delivering medical care directly to patients' homes, which has created a growing reliance on medical technologies and eHealth tools [1] to support health care staff in managing not only older people care but also palliative and end-of-life care and patients with multiple comorbidities or complex diseases. The increased presence of eHealth tools also highlights the importance of considering the competencies and requirements of HCPs for using eHealth tools in older people care [2].

eHealth's vital role in advanced home care includes using telehealth, remote patient monitoring, and mobile health apps to deliver health care services to patients with complex medical needs in their homes [5]. eHealth tools facilitate communication between patients and HCPs, support medication management, monitor vital signs and symptoms, and offer educational resources to patients and caregivers. Furthermore, integrating eHealth into advanced home care can also enable organizational suppliers of health to deliver personalized and coordinated care, increase patient engagement and self-management, and improve overall patient satisfaction [2].

In Sweden, the national policy emphasizes home assistance and home care over institutionalized care, with various programs supporting long-term care for older people, assistance for people with disabilities, end-of-life care, palliative care in hospitals or hospices, and advanced palliative home care administered by municipalities. In 2017, home care met 72% of long-term care needs, while institutions served the remaining 28% [6]. Given the substantial reliance on home care for health care delivery, there is a significant need for resources, making eHealth services, such as those mentioned earlier, highly beneficial [7].

A study by Rydenfält et al [8] outlined the implementation of eHealth services in Swedish home care nursing, listing national patient summaries, mobile documentation, digital locks, digital medical lists, digital security alarms, and camera supervision as the most commonly used services. Organizational suppliers of health care had implemented mobile documentation most widely among these services at the time of the study, with plans for its continued use in the future for home care nursing.

Traditional Documentation

Health care staff have long used traditional methods, such as pen-and-paper notes, to document patient encounters, including observations, assessments, signs, symptoms, and communications during clinic consultations and home care visits. While some remnants of this traditional documentation method may persist in current practices, health care organizations have adopted information and communication technology solutions [9] and notable mobile communications [10] to address the inherent challenges of traditional documentation. Traditional documentation methods occasionally lead HCPs to omit crucial information inadvertently, record data inaccurately, produce illegible handwriting, and misunderstand patients' accounts, while requiring significant time for documentation [11].

Patient Experiences With eHealth

Researchers have extensively explored patient experiences with eHealth, particularly in home and palliative care settings, and have revealed a range of benefits and challenges. Steindal et al [3] found that telehealth applications in palliative home care enhance access to HCPs, helping patients feel more secure and safe. Similarly, Widberg et al [4] highlighted that eHealth applications facilitate better communication between patients and HCPs in palliative care. Karlsen et al [12] focused on older adults' use of telecare in home care services and emphasized their desire to age in place. These studies show that patients view technology as a valuable tool to achieve this goal, although

XSL•FO
RenderX

some resist adopting new technologies, struggle with digital proficiency, and worry about privacy and security [5].

Older patients face notable challenges when using eHealth tools. These include their preference for direct contact with HCPs, age-related cognitive decline, the stigmatization of telemonitoring devices, and the potential loss of social interactions, which are essential for well-being [5,13]. Experts emphasize the importance of balancing in-person care with eHealth use for certain populations, such as older patients with chronic conditions. Despite these obstacles, studies highlight that eHealth improves care accessibility and patient autonomy [3,5,12].

Specific eHealth applications directly benefit patients. For instance, an eHealth app tested in home care settings allowed patients to report health concerns and increased their sense of security [10] by reducing the frequency of phone calls to nurses, enabling them to prioritize care based on the app's alerts. However, frequent reporting through the app led HCPs to perceive minor health issues as significant, which increased their workload [10]. These findings illustrate the nuanced impact of eHealth on patient experiences and underline the need to carefully address patients' needs and contexts.

The Acceptance and Perspectives of HCPs on eHealth

HCPs play a pivotal role in adopting and successfully implementing eHealth solutions, as their acceptance and perspectives directly influence outcomes. Li et al [14] identified key factors that affect HCPs' acceptance of eHealth. These factors include HCP characteristics, voluntariness of use, performance and effort expectancy, and how organizational conditions either support or hinder eHealth adoption. HCPs emphasize patient autonomy, personalization, and continuity of professional support as facilitators of eHealth adoption, while they note advanced patient age as a primary barrier [15].

Qualitative studies offer additional insights into HCPs' attitudes toward eHealth. In an interview study, general practitioners expressed positive attitudes toward eHealth as a means to promote healthy lifestyles for patients and themselves [16]. They showed confidence in transitioning from traditional paper-based approaches to digital solutions, particularly those incorporating patient-reported outcome measures. Similarly, Das et al [17] examined the impact of an eHealth portal on HCP-patient interactions and reported that HCPs value the portal for its ability to provide comprehensive patient information, foster accountability, and serve as a clinical tool. However, they also identified organizational challenges, such as a lack of incentives and time constraints, which hinder eHealth integration.

Understanding HCPs' experiences and perceptions is essential for optimizing eHealth tools. While eHealth enhances health care quality, efficiency, and accessibility, its success depends on addressing the barriers that HCPs face. Targeted training and strategies to overcome organizational constraints can improve HCP satisfaction and ensure seamless integration of eHealth into clinical practice [18].

Aim of the Study

Despite extensive research in eHealth, there remains a limited understanding of how eHealth tools affect the relationship between HCPs and patients in advanced home care settings. While some studies have explored patient perspectives, fewer researchers have focused on HCPs' experiences and perceptions.

The aim of this study is to explore HCPs' perspectives on using eHealth tools in advanced home care and their impact on HCP-patient relationships. This research is particularly relevant in the context of Sweden's robust home care system.

Methods

Methodology Overview

This study used qualitative research methods, collecting data through semistructured interviews with HCPs working in advanced home care at Aleris in Sweden. Aleris is one of the major private organizational suppliers of health care in Sweden, offering a range of medical services, including advanced home care and other specialist medical treatments [19]. This approach suited the study's objective, which aimed to gain a deep understanding of HCPs' perspectives on using eHealth tools in advanced home care and their impact on relationships with patients. The flexibility of qualitative research allowed nuanced meanings to emerge from participants, captured through direct quotations from interview transcripts. The study adopted an exploratory design to further explain the role of eHealth tools in advanced home care as perceived by HCPs and examined how these tools influence their interactions with patients. This design proved beneficial in uncovering connections between ideas that were either underrepresented or not previously demonstrated in the literature [20]. An inductive approach [21] analyzed the data, aligning with the study's objectives.

The eHealth Tools

The eHealth tools used in the study are part of the SwipeCare health care process management system, which partially integrates into TakeCare, the electronic health record (EHR) system used by HCPs in their daily practice. TakeCare serves as the main repository for patient information and is used for routine clinical documentation. SwipeCare functions as an overlay on TakeCare, enhancing the existing EHR system by adding functionalities, such as iPad (Apple Inc)-based questionnaires, calculations, creating and sending questionnaires to patients, and patient reminders [22]. SwipeCare controls and coordinates the patient care process from diagnosis and treatment to aftercare and follow-up. It allows the organizational suppliers of health to define and manage the entire health care process, initiating actions within that process, some performed by SwipeCare itself (like sending out questionnaires) and others by HCPs or other units. SwipeCare interacts with TakeCare only when it needs to read from or write to the patient's record, thus complementing rather than replacing the EHR.

The 2 tools chosen for the evaluation were a mobile documentation tool and a mobile preconsultation form. The tools were chosen for their complementary nature and potential to capture a wide range of HCP-patient interactions. While the mobile documentation tool was used by the HCP, or

XSL•FO
RenderX

collaboratively by HCP and patients, during in-person visits, the mobile preconsultation form was completed independently by patients at home. This diversity in tool use allowed a comprehensive exploration of eHealth's impact on HCP-patient relationships across different contexts.

The iPad-based mobile documentation tool consists of different predefined standardized forms, which HCPs, either themselves or together with the patient, fill in. Such a form could, for example, cover the required information related to the admission of the patient to a ward or a checklist for evaluating the patient's physical status after an adjustment in the treatment. The questions in the form consist of predefined categories designed to guide HCPs in documenting patient information. Caregivers have the discretion to determine which answers are available for HCPs to fill in. In some cases, there is a designated "comments" section where HCPs can provide additional explanations or insights. However, in other instances, caregivers may choose to limit responses to predefined alternatives. This limitation mirrored traditional documentation methods, such as pen-and-paper systems, where options were similarly restricted. In addition, certain calculations, such as the Mini Nutritional Assessment (MNA) [23] or the National Early Warning Score 2 (NEWS2) [24], necessitate numerical input and cannot accommodate text responses. This further underscored the importance of predefined answer alternatives within the tool. In addition, the forms ensure no relevant aspects are missed, as questions can be made mandatory to answer.

The second tool, the mobile preconsultation form, is part of the patient portal in the system, which enables the patient to access their own portal, or web page, on their mobile or equivalent digital device. The patient first receives an SMS text message with a link to their patient portal. The patient logs in using a secure method and can then access the preconsultation form. The patient answers the questions at their own pace and time and finally sends the answers back to the caregiver. This enables HCPs to review the status of the patient before traveling to the patient's home, avoiding the time spent on questioning and filling in answers. Instead, it allows that time to be spent caring for the patient where the patient has its needs, for example, pain in the foot or problems with digestion. The intention is to free HCPs from administration to devote time to the care of the patient and encourage patients to take a more active role in their own care.

Data Collection and Participant Selection

The recruitment process involved contacting selected participants via email, facilitated by managers from each of the 9 participating clinics in Stockholm, Sweden. Clinics were selected to ensure geographic diversity across the Stockholm region, including clinics from the north, south, and other areas. Clinic heads collaborated in identifying suitable participants with experience using the eHealth tools. This selection strategy aimed to capture a comprehensive range of perspectives, strengthening the validity of the findings. Inclusion criteria for participants comprised individuals aged ≥18 years, officially employed at these clinics, and experienced in using the eHealth tools in their professional roles. The diversity in HCPs' roles and experience levels allowed for capturing a broad view of eHealth tool use in advanced home care.

Interviews took place between March and May 2023. Participants received the informed consent form via email before the interviews, which they reviewed and signed in person at the beginning of the interview session. Audio recordings and notes documented each 30- to 45-minute interview. A structured interview guide provided consistency while allowing the interviewer flexibility to explore topics relevant to HCPs' perspectives further. For example, one question from the interview guide asked participants to describe how they adapted their use of eHealth tools based on individual patient needs, providing deeper insights into their experiences and perceptions regarding the patient-HCP relationship.

The characteristics of the 20 HCPs who participated in the interviews varied across the 9 clinics. Each clinic contributed between 1 and 4 HCPs, with a median of 2, ensuring a good spread of perspectives across the clinics involved. Participants were aged between 24 to 59 years, with a mean of 45 (SD 9.3; IQR 41-52) years.

Table 1 presents participants' characteristics, categorized by profession, years of practice, duration of using mobile documentation tools, and duration of using mobile preconsultation forms. All (20/20, 100%) participants used the mobile documentation tool, while only 60% (12/20) of them used the mobile preconsultation form. This difference in tool use reflects the phased implementation approach common in health care settings, allowing for the gradual adoption and refinement of new technologies. The relatively short duration of eHealth tool use for many participants (0-2 years) illustrates the dynamic nature of health care staffing, characterized by a high turnover rate [25]. HCPs joining from other organizational suppliers of health care or units typically lacked previous experience with these specific eHealth tools. This limited experience duration accurately represents the reality found in many health care units and provides valuable insights into the initial adoption phase of eHealth tools in advanced home care.

These factors—the selection process, varied tool familiarity, and relatively short use duration—shaped the findings by offering a realistic snapshot of eHealth tool adoption in health care units where staff turnover is common. While this scenario limits observations on long-term impacts, the study offers a pragmatic view of eHealth tool implementation, helping stakeholders set realistic expectations and develop strategies that account for the realities of staff turnover and varying levels of tool familiarity in health care settings.

XSL•FO
RenderX

Table 1. Participant characteristics.

Participant	Duration of clinical practice[a] (y)	Duration of the use of mobile documentation tools (y)	Duration of the use of mobile preconsultation form (y)
Nurse 1	6-10	0-2	—[b]
Nurse 2	16-20	0-2	—
Nurse 3	6-10	0-2	0-2
Nurse 4	6-10	0-2	0-2
Nurse 5	6-10	0-2	0-2
Nurse 6	0-5	0-2	0-2
Nurse 7	6-10	3-5	3-5
Nurse 8	0-5	0-2	0-2
Nurse 9	0-5	3-5	3-5
Nurse 10	0-5	0-2	0-2
Nurse 11	6-10	>5	0-2
Nurse 12	11-15	0-2	—
Nurse 13	0-5	0-2	0-2
Nurse 14	6-10	0-2	0-2
Nurse 15	0-5	0-2	0-2
Assistant nurse 1	6-10	0-2	—
Assistant nurse 2	6-10	0-2	—
Assistant nurse 3	0-5	0-2	—
Assistant nurse 4	0-5	3-5	—
Operational therapist 1	>20	0-2	—

[a]Clinical practice: total time spent as health care professionals across all settings, including advanced home care clinics.

[b]Not applicable.

Data Analysis

The interview, recorded in audio format, underwent verbatim transcription and subsequent review for accuracy and completeness. Following the transcription of the data, participants' responses were systematically organized and categorized. The analysis was conducted using a content analysis approach grounded in the methodology described by Graneheim and Lundman [26]. The process began with a thorough data familiarization. To gain a deep understanding of its content, we immersed ourselves in the material, reading it carefully multiple times. Once a comprehensive understanding of the data was achieved, we identified *meaning units*—specific segments of text that conveyed key information relevant to the study's objectives. These meaning units were then condensed to retain their core meaning while eliminating unnecessary details. Each condensed meaning unit was assigned a code, which served as a concise label capturing its essential message. The codes were reviewed and grouped into *subcategories* based on shared patterns and similarities, providing a more detailed and refined understanding of the data. These subcategories were subsequently combined into broader *categories* that represented the distinct dimensions of our study. We then merged related categories and identified key subthemes. Ultimately, we consolidated these findings into an overarching theme. This theme captured the core insights from the data, highlighting both the benefits and challenges of eHealth tool use in advanced home care from HCPs' perspectives.

Ethical Considerations

This research was conducted in Sweden. According to the Swedish Ethical Review Act (SFS 2003:460) [27,28], this study does not require ethics approval as it does not involve the handling of sensitive personal information, as defined by the European General Data Protection Regulation (GDPR), Regulation (EU) 2016/679 [29]. However, we acknowledge that ethical principles still apply and have been followed in accordance with relevant regulations and research guidelines.

Prospective study participants received an invitation email outlining the study's purpose. Those who agreed to participate were contacted via email, SMS text messages, or phone calls to schedule interviews according to their preferences. Before beginning the interviews, participants were provided with a comprehensive explanation of the study objectives, interview procedures, and their rights. Both verbal and written consent were obtained from each participant, with informed consent signed by the participant and the researcher. The consent form included information regarding the aim of the study, potential risks and discomforts, potential benefits, how confidentiality would be handled, the right to withdraw, and consent

XSL•FO
RenderX

procedures. It also assured participants of their confidentiality. All participants were informed of their right to withdraw from the study at any stage. Furthermore, no monetary or nonmonetary compensation was provided to any participants. Finally, they were assured that all raw data, comprising audio recordings and interview notes, would be stored on a secured laptop with biometrics log-in (fingerprint) for enhanced data security.

Results

Overview of Data Analysis

The data analysis identified one main theme—*optimizing health care with eHealth: enhancing care delivery and overcoming challenges for future health care*. Two subthemes emerged: (1) enhancing care delivery, collaboration, and overcoming adoption barriers and (2) streamlining implementation and advancing eHealth tools for future health care delivery. Five categories were also identified: (1) positive experiences and benefits, (2) interactions between HCPs and patients, (3) challenges and difficulties with eHealth tools, (4) integration into the daily workflow, and (5) future directions. Table 2 shows the subcategories, categories, subthemes, and themes that emerged after coding.

The overarching theme highlighted the study's central focus on the HCPs' perspectives on the use and integration of eHealth tools in their daily practice. It emphasized how these tools impact the dynamics of health care delivery, including their influence on the relationships between HCPs and patients, and how they shape the overall patient care experience.

Table 2. Overview of the *optimizing health care with electronic health: enhancing care delivery and overcoming challenges for future health care* theme, subthemes, categories, and subcategories.

Subtheme and categories	Subcategories
Enhancing care delivery, collaboration, and overcoming adoption barriers	
Positive experiences and benefits	• Efficiency • Better documentation and enhanced patient safety
Interactions between HCPs[a] and patients	• Unaffected HCP-patient relationships • Enhanced patient engagement • Communication barrier
Challenges and difficulties with eHealth tools	• Technical and resource challenges • Patient-related factors
Streamlining implementation and advancing eHealth tools for future health care delivery	
Integration into the daily workflow	• Ease of use • System integration and application
Future directions	• Staff recommendations • Continuity of use

[a]HCP: health care professional.

Enhancing Care Delivery, Collaboration, and Overcoming Adoption Barriers

Positive Experiences and Benefits

The participants highlighted several benefits of using the mobile documentation tool. One significant advantage was the decreased documentation time during patient consultations. The mobile tool allows HCPs to document directly during patient encounters, which increase *efficiency* by eliminating the need for double documentation [30] and reducing the time spent on paperwork. Approximately two-thirds (13/20, 65%) of the HCPs indicated that the mobile documentation tool helped them save time on documentation:

> *I think it [the mobile documentation tool] decreases my work time [taking clinical notes during consultation].* [Nurse 6]

> *It saves me time because I document while I am at the patient's house.* [Nurse 9]

Another benefit mentioned was the standardization of content for clinical note-taking. The use of standardized templates within the mobile tool ensured that all relevant information was captured consistently, which is crucial for effective patient care. Participants also expressed an enhanced sense of safety and security for patients when using the mobile documentation tool. This improvement was attributed to more accurate and timely documentation of health status, which contributed to better patient outcomes. The time saved through efficient documentation potentially translated to improved patient care quality. With reduced administrative burden, HCPs could devote more time to meaningful patient interactions, allowing more in-depth discussions about health concerns and treatment plans.

While most (13/20, 65%) participants perceived the mobile documentation tool as beneficial, some noted challenges. A few (7/20, 35%) participants reported that it disrupted their routine workflow by prolonging consultation times. They found that carrying the iPad during home care visits added to their equipment load, making it inconvenient:

XSL•FO
RenderX

> *I think it [the mobile documentation tool] increases my work time. My usual consultation is about 20-30 minutes, with it [the mobile documentation tool], it can take 45-50 minutes. We carry a lot of equipment during home care visits, including a laptop, and it can be heavy. The iPad is an additional weight.* [Nurse 1]

> *Because we cannot do everything we need in the [mobile documentation] tool, it kind of doubles the work for us.* [Nurse 4]

Among users of the mobile preconsultation form, participants highlighted several advantages, including knowing the patient's medical concerns in advance and saving time on preparations. This previsit work done by patients reduced the administrative workload on HCPs, allowing for a more efficient use of consultation time.

Standardizing documentation through the mobile documentation tool and preconsultation form may lead to more consistent and thorough patient assessments across different HCPs. This standardization ensures that all relevant information is captured systematically, potentially improving the continuity of care and reducing the risk of overlooking critical patient data.

The use of eHealth tools in advanced home care led to *better documentation practices while enhancing patient safety*. The majority (13/20, 65%) of the participants found the standardized content of the mobile documentation tool beneficial, noting that it ensured they would not overlook critical information:

> *It is kind of nice that the [mobile documentation] tool has a lot of information that can help with taking notes. Also, the setup is the same for everyone, so I think the note-taking is sort of standardized.* [Nurse 1]

This standardized approach to documentation not only improved consistency across different HCPs but also potentially enhanced the quality of patient care. By ensuring comprehensive data collection, the tool might lead to more informed clinical decision-making and improved continuity of care. The structured format guided HCPs through specific assessment points, helping to ensure that all relevant aspects of patient care were addressed during consultations. Improved documentation through eHealth tools contributed to patient safety and security. Participants noted that accurate information recorded in EHRs reassures patients, helping them feel more secure in their care:

> *When we use [the mobile documentation tool], we can check the necessary things we have to ask the patient. We will not forget anything because we have a guide, I think that is good for the patient...it can increase patient security.* [Nurse 13]

> *I think they [patients] feel safer when I can write everything in the [mobile documentation] tool while I am with them, and they can check what I write.* [Nurse 8]

The ability for patients to review clinical notes during consultations further enhanced their confidence, fostering a sense of transparency in the health care process and further strengthening the relationship. This aligned with the principle of "equal care" in Sweden, which emphasizes the importance of providing consistent quality care to all patients, regardless of their location or the HCPs they encounter [31].

By standardizing documentation practices, eHealth tools contributed to this principle by ensuring that all relevant patient information is consistently captured and accessible to different HCPs. This standardization not only improved continuity of care but also enhanced trust between patients and HCPs. The structured format of the mobile documentation tool facilitated comprehensive data collection, allowing HCPs to focus on critical issues during consultations. This approach optimized patient interactions and strengthened the overall quality of care provided, aligning with Sweden's commitment to equitable health care access.

Interactions Between HCPs and Patients

All (20/20, 100%) participants indicated that their *relationship with patients remained unaffected* by the implementation of both the mobile documentation tool and the mobile preconsultation form. This finding suggested that HCPs were able to maintain their interpersonal skills and patient-centered approach while integrating eHealth tools into their practice:

> *I think nothing has changed in how I interact with patients when using it [the mobile documentation tool].* [Nurse 5]

The consistency in relationships despite introducing new technology highlighted the adaptability of HCPs in maintaining effective communication with patients. While the eHealth tools introduced new elements to patient interactions, such as the physical presence of devices, HCPs appeared to successfully navigate these changes without compromising the quality of their patient relationships. Furthermore, the mobile preconsultation form may have enhanced patient engagement by allowing patients to provide more comprehensive information before consultations. This previsit work could potentially lead to more focused and in-depth discussions during face-to-face interactions, without negatively impacting the HCP-patient relationship.

The mobile preconsultation form emerged as a tool that potentially *enhances patient engagement* in their care. Some (3/12, 25%) participants noted that certain patients provided more comprehensive information through written responses on the mobile preconsultation form than during face-to-face interactions with the HCPs. This suggested that the tool may lead to more informed and collaborative interactions during consultations:

> *If you send out the form, I think you can get more information from the patients because they [patients] have more time to think about it rather than when we ask them in person.* [Nurse 7]

The previsit work done by patients allowed them to elaborate on both current problems and positive aspects of their health status. This more comprehensive written information might shift the dynamics of face-to-face interactions, potentially allowing more focused and in-depth discussions during consultations. Furthermore, the mobile preconsultation form's ability to highlight patients' medical concerns in advance

XSL•FO
RenderX

enabled HCPs to prepare more effectively for consultations. This preparation may contribute to more efficient and patient-centered interactions, as evidenced by one participant's comment:

> *We send out the form once every four weeks for when we do the team rounds, it can help us in doing the team rounds, making it faster and more efficient, in my experience.* [Nurse 5]

By facilitating more comprehensive information gathering and allowing HCPs to focus on critical issues, the mobile preconsultation form may enhance patient engagement and potentially improve the quality of HCP-patient interactions in advanced home care settings.

Participants mentioned *communication barrier* as a challenge. While the eHealth tools offered numerous benefits, 20% (4/20) of the participants observed noted challenges in their use during patient interactions. These HCPs noted that the iPad (mobile documentation tool) could sometimes be a barrier to effective communication:

> *When I am holding the iPad in front of the patient, it kind of becomes a barrier between us.* [Nurse 6]

> *When using the iPad, you tend to look down instead of looking at the patient directly.* [Nurse 9]

> *Maybe some patients find it rude [using an iPad during consultation], but most of them understand that we do it to get everything right. Also, I think sometimes I have a less open conversation with patients.* [Nurse 10]

Most (16/20, 80%) participants did not report communication issues, suggesting that the majority of HCPs were able to integrate the eHealth tools without significant impact on patient interactions. Compared to traditional pen-and-paper methods, the eHealth tool presented both advantages and challenges for patient-HCP communication. While pen-and-paper notes may feel less intrusive during patient interactions, the eHealth tool offers improved legibility, standardization, and ease of data retrieval, potentially enhancing the quality of information shared with patients. However, the physical presence of the device could create a perceived barrier for some HCPs and patients. The eHealth tool's structured format ensures comprehensive documentation, potentially reducing the risk of omitting crucial information compared to freeform paper notes. This standardization may lead to more consistent and thorough patient assessments across different HCPs, indirectly benefiting patient care and communication. However, this structure might sometimes limit the flexibility in capturing nuanced patient narratives.

Regarding focus on issues and medical decision-making, the eHealth tool's ability to provide instant access to patient history and standardized assessments can potentially enhance HCP's ability to focus on critical issues during patient interactions. However, as noted by some participants, the tool might occasionally divert attention from direct patient interaction. These observations highlighted the need for balanced use of eHealth tools, combining their benefits with maintaining effective patient communication. HCPs may need to develop strategies to integrate these tools seamlessly into their patient interactions, ensuring that technology enhances rather than hinders the HCP-patient relationship.

Challenges and Difficulties With the eHealth Tools

While most (14/20, 70%) participants reported no technical issues, some HCPs encountered *technical and resource challenges* when using eHealth tools in advanced home care. Specifically, 30% (6/20) of the participants faced technical problems with the mobile documentation tool, including log-in difficulties, frequent software updates, and issues importing clinical notes to the patient's EHR:

> *First, when you come in to work, you have to log in to the [mobile documentation tool]. Sometimes, it's a problem. Now, with the frequent updates—they do updates like every week—there is often something wrong after each update...* [Nurse 3]

Notably, 100% (12/12) of the participants who used the mobile preconsultation form reported no technical issues at all. Resource limitations were a more widespread concern. All (20/20, 100%) participants noted the need to share iPads among staff members, potentially limiting access and flexibility in documentation:

> *The staff has to share iPads [with the mobile documentation tool] within the team.* [Nurse 9]

These challenges highlighted the importance of robust IT infrastructure and adequate resource allocation in eHealth implementation. While technical issues affected only a minority (6/20, 30%) of users, device sharing impacted all (20/20, 100%) participants, potentially influencing HCP efficiency and, consequently, patient care quality.

Patient-related factors, such as patient characteristics, influenced the adoption and use of eHealth tools in advanced home care. Approximately 55% (11/20) of the participants noted that the patient's age and health conditions affected receptiveness to eHealth tools, particularly the mobile preconsultation form:

> *Not all patients are open to using the [mobile preconsultation] form. Usually, older patients or those with difficult diseases are not open to it.* [Nurse 12]

HCPs demonstrated adaptability by adjusting their approach based on patient characteristics. This flexibility was crucial for maintaining quality care across diverse patient populations. However, it also raised concerns about potential disparities in care delivery between technologically proficient patients and those less comfortable with eHealth tools. Some patients expressed skepticism about using the mobile preconsultation form, particularly due to privacy concerns:

> *There are patients who are skeptical about using it [the mobile preconsultation form], as it involves using their [personal information] BankID [a Swedish electronic identification system].* [Nurse 11]

On the other hand, a subset of patients, predominantly younger individuals, demonstrated an openness to using eHealth tools, according to 55% (11/20) of the participants. This varying receptiveness highlighted the need for a personalized approach in implementing eHealth tools in advanced home care settings,

XSL•FO
RenderX

ensuring that all patients receive appropriate care regardless of their technological proficiency.

Streamlining Implementation and Advancing eHealth Tools for Future Health Care Delivery

Integration Into the Daily Workflow

Regarding *ease of use*, 75% (15/20) of the participants found the mobile documentation tool easy to use. The initial introduction and demonstration of these eHealth tools to HCPs lasted an hour and was perceived as sufficient by all the 12 (12/12, 100%) participants who received it. Some (4/20, 20%) participants had the opportunity to familiarize themselves with the tool beforehand as part of the implementation team, while 12 (12/20, 60%) participants received information during introductory seminars at their clinics. Some (4/20, 20%) participants only underwent introductory training after using the mobile documentation tool:

> *I have tried it [the mobile documentation tool], honestly, without any training... I tried to use it... I think it is easy.* [Assistant Nurse 2]

> *My colleague taught me how to use it [the mobile documentation tool].* [Nurse 8]

In addition, 25% (5/20) of the participants reported that the log-in process and navigating through all the components of the mobile documentation tool required some time to become familiar with, but they gradually improved with more use:

> *As I continue to use it [the mobile documentation tool], I learn more about it, which makes me better at using it.* [Nurse 2]

Regarding *system integration and adaption* all (20/20, 100%) participants highlighted challenges in integrating eHealth tools into existing systems and their practical *application* in advanced home care. A key limitation was the incomplete integration of the mobile documentation tool with the existing EHR system, TakeCare. While clinical notes could be uploaded to the patient's journal, other crucial data remained inaccessible through the mobile tool, requiring staff to rely on computers for a complete view of patient information:

> *It [the mobile documentation tool] doesn't have all the functions that you have in the EHR.* [Nurse 7]

> *[B]ecause with the computer, I can do and see everything on the patient's EHR, but not with the iPad.* [Nurse 4]

The lack of full system integration disrupted seamless workflows and raised concerns about potential inconsistencies in documentation and delays in accessing critical information during consultations. Participants expressed a strong desire for improved interoperability to ensure that all necessary data are accessible on the mobile device.

Regarding practical application, HCPs generally found the mobile preconsultation form useful for integrating into their existing workflows. However, its applicability was perceived as limited for certain patient groups and care scenarios. For example, HCPs noted that sending the form before every consultation was unnecessary for patients with chronic conditions requiring long-term care as their primary clinical concerns often remained consistent:

> *For example, I have had this patient for a very long time, and the patient has a chronic condition. I don't need to send the form [before] every visit because the status is still the same.* [Nurse 11]

Despite these limitations, participants acknowledged the value of the preconsultation form for new patients or those with changing health conditions. It allowed HCPs to gather detailed information in advance, facilitating more focused consultations. These findings underscored the importance of improving system integration to enhance workflow efficiency while adapting eHealth tools to meet diverse patient needs in advanced home care settings.

Future Directions

HCPs shared *recommendations* on enhancing the effectiveness of the eHealth tools in their daily workflow in advanced home care. Most (16/20, 80%) participants expressed that completely integrating the mobile documentation tool with the patient's EHR system would improve efficiency. They suggested that accessing the same information as in the patient's EHR directly on the mobile device would eliminate the need to switch between the iPad and the computer for clinical documentation, improving their workflow:

> *Lab results and checking off the to-do list, that is what is still missing, I think...and an easier login.* [Nurse 9]

All (12/12, 100%) HCPs who used the mobile preconsultation form found it adequate. However, 17% (2/12) of the participants suggested further enhancements to the interface, including languages other than Swedish:

> *The [mobile preconsultation] form does not look very good, in my opinion. It can be better.* [Nurse 11]

Regarding equipment use, half (10/20, 50%) of the participants expressed that having sufficient iPads for all staff members would be advantageous, reducing the need for frequent logging in and out.

Approximately two-thirds (14/20, 70%) of participants anticipated *continuity of use* of the 2 eHealth tools in advanced home care, acknowledging the emergence of the new technologies. In addition, 45% (9/20) of the HCPs expressed optimism regarding the increased patient acceptance of technology in health care:

> *I think they [patients] will use them [eHealth tools] more and more in the future.* [Nurse 7]

Discussion

Principal Findings

The aim of this study was to explore HCPs' perspectives on using eHealth tools in advanced home care and their impact on HCP-patient relationships. In total, 20 HCPs from 9 different advanced home care clinics participated in the interviews, leading to the identification of five categories: (1) positive experiences and benefits, (2) interactions between HCPs and

XSL•FO
RenderX

patients, (3) challenges and difficulties with the eHealth tool, (4) integration into the daily workflow, and (5) future directions.

Overall, most (13/20, 65%) HCPs expressed a positive perspective regarding the 2 eHealth tools used in the study: the mobile documentation tool and the mobile preconsultation form. The results showed that HCPs identified several benefits for themselves and their patients. Notably, 65% (13/20) of the participants reported that the mobile documentation tool allowed them to document directly during patient encounters, eliminating double documentation and saving time. Standardized templates within the tool ensured consistency in clinical note-taking, which participants believed enhanced patient safety by reducing errors and omissions.

While most (13/20, 65%) HCPs reported increased efficiency, 35% (7/20) experienced the opposite, suggesting that user acceptance of eHealth is multifaceted. These participants reported that the mobile documentation tool disrupted their routine workflow by prolonging consultation times and adding to their equipment load during home care visits. This highlights the importance of considering individual user experiences in eHealth implementation. The ease of use of the eHealth tools was a significant factor in their integration into daily workflows. Most (15/20, 75%) of the participants found the mobile documentation tool easy to use, indicating successful integration. This high level of usability likely contributed to the overall positive reception of the tools and suggested that well-designed eHealth solutions can be effectively adopted in advanced home care settings.

However, HCPs noted challenges, both tool-related and patient-related. Some (6/20, 30%) of the participants faced technical problems with the mobile documentation tool, including log-in difficulties and issues importing clinical notes to the EHR. These technical challenges highlight the need for robust IT infrastructure and ongoing technical support in eHealth implementation. Resource limitations were a widespread concern, with all participants noting the need to share iPads among staff members. This sharing potentially limits access and flexibility in documentation, impacting workflow efficiency. Future eHealth implementations should consider resource allocation to ensure adequate device availability for all HCPs.

Patient characteristics influenced the adoption of eHealth tools, particularly the mobile preconsultation form. Approximately 55% (11/20) of the participants noted that patient age and health conditions affected receptiveness to these tools. This finding underscores the importance of tailoring eHealth implementations to diverse patient populations and considering alternative approaches for patients less comfortable with technology. The mobile preconsultation form emerged as a tool that potentially enhances patient engagement. Some (3/12, 25%) of the participants using this tool noted that certain patients provide more comprehensive information through written responses than face-to-face interactions with HCPs. This previsit work allows for more focused and efficient consultations, potentially improving the quality of HCP-patient interactions.

Regarding the HCP-patient relationship, all (20/20, 100%) participants reported that the use of eHealth tools had no significant impact on the relationship with patients. This suggests that HCPs successfully maintained their interpersonal skills and patient-centered approach despite introducing new technology. However, 20% (4/20) of the participants noted that using devices such as iPads could occasionally act as a communication barrier during consultations, highlighting the need for strategies to seamlessly integrate technology into patient interactions.

Looking to the future, 70% (14/20) of the participants anticipated continued use of the eHealth tools in advanced home care, reflecting their perceived long-term value despite existing challenges. In addition, 45% (9/20) of the HCPs expressed optimism about increased patient acceptance of technology in health care. While participants did not report changes in HCP-patient interactions, they emphasized that current eHealth tools successfully maintain existing relationships without negatively impacting communication or rapport.

Comparison With Prior Work

Mathijssen et al [32] highlighted a similar finding of positive perception of eHealth by HCPs in home care. Despite recognized challenges such as inadequate training and support, data privacy concerns, and the necessity for more user-friendly technology, eHealth in home care is seen as a promise to enhance care quality. A study focusing on a specific eHealth tool, electronic medication administration records, emphasized the importance of cautious implementation due to potential unintended consequences [33]. Participants noted that the electronic medication administration records introduced new tasks alongside existing nursing tasks. In contrast, our study found no perceived changes in the workflow or work situation due to eHealth tools, potentially contributing to their overall positive reception.

Carlqvist et al [34] conducted a qualitative interview study examining how an eHealth application can serve as a value-creating resource from the perspective of HCPs. Their findings indicated that while such applications can enhance proactive communication and support patient engagement in self-care, challenges remain. Notably, patients' difficulties in using the application or performing measurements sometimes led to value destruction, requiring time-consuming recovery efforts by professionals. This contrasts with our study, where HCPs reported no significant workflow disruptions due to eHealth tools, potentially contributing to their positive acceptance.

Nakrem et al [35] explored the introduction of digital medicine dispensers in home health care services and their influence on patient-caregiver relationships. They found that while such technologies can improve efficiency and enhance patient independence, they may also strain patient-caregiver relationships if not aligned with patients' needs and safety concerns. However, in our study, HCPs did not perceive eHealth tools as affecting their relationships with patients, suggesting that the specific type of technology and its implementation context play crucial roles in acceptance and impact.

Furthermore, Ramachandran et al [36] conducted a review examining the impact of eHealth on patient-HCP and HCP-HCP relationships in primary care. They found that eHealth can have

XSL•FO
RenderX

both positive and negative effects on relationships and trust, influenced by factors, such as technology design, patient demographics, and organizational implementation strategies. This aligns with our findings, where the successful integration of eHealth tools without perceived workflow disruptions may be attributed to effective implementation strategies and user-centered design.

Future Research

Future studies could explore strategies for seamlessly integrating eHealth tools into patient interactions without compromising communication quality. In addition, research into personalized eHealth solutions that accommodate diverse patient needs and preferences could address challenges related to patient characteristics. Investigating the long-term impacts of eHealth tools on care quality, patient outcomes, HCP-patient relationships, and HCP job satisfaction would provide valuable insights for future implementations.

Another important area for investigation is the development and implementation of features that actively enhance HCP-patient relationships in advanced home care environments. For instance, improving patient education modules or creating shared decision-making tools tailored to home care settings could address some of the gaps identified in this study, such as maintaining meaningful interactions while using eHealth tools.

Furthermore, exploring the role of relatives in supporting patients with eHealth tools could offer new avenues for enhancing patient engagement and care delivery [37]. Relatives could assist with completing preconsultation forms or managing digital tools for patients who face challenges due to age or health conditions. Recognizing relatives as key stakeholders alongside HCPs may improve overall patient support in advanced home care settings [38].

Limitations

This study focused specifically on advanced home care settings in Sweden, which may limit the applicability of findings to other health care contexts. The unique characteristics of advanced home care, such as the need for mobile documentation and remote patient monitoring, shaped the experiences reported by HCPs. In addition, the exclusive focus on HCPs' interviews represents another limitation. Incorporating patients' perspectives could offer valuable insights and contribute to a more comprehensive understanding of the implementation of eHealth tools in advanced home care.

Conclusions

HCPs in advanced home care recognize the value of eHealth tools in their daily work while maintaining the quality of their relationships with patients. The study provides insights into both the experiences of HCPs using eHealth tools and how these tools influence their interactions with patients in advanced home care settings. Positive experiences related to eHealth use and the level of integration of eHealth tools into daily work motivate staff to use these tools, potentially allowing for more time and focus on patient interactions. By contrast, negative experiences with the eHealth tools limit their use and acceptance.

Notably, HCPs reported no perceived changes in their relationships with patients when using the mobile documentation tool and the preconsultation form tool. This suggests that eHealth tools can be integrated into care delivery without compromising the personal aspect of HCP-patient interactions. Although most of the staff in the study had a positive view of their use of the eHealth tools, understanding the challenges they encounter is essential to increase acceptance and success in implementation further. Future development should focus on features that not only improve efficiency but also actively enhance HCP-patient relationships in the advanced home care setting.

Acknowledgments

The authors would like to thank all the health care staff members who participated in this study for their time and valuable insights. Artificial intelligence was used for writing assistance. ND was funded by the Health Informatics Center at Karolinska Institutet. The funding did not influence the study's design, data collection, analysis, or manuscript writing.

Authors' Contributions

EVR, ND, and UL participated in the study design. Data collection was conducted by EVR. EVR performed the initial analysis, which was then reviewed and refined through discussions with ND. The first draft of the manuscript was prepared by EVR, with all authors contributing to subsequent writing and review processes.

Conflicts of Interest

During this study, UL was employed by the medical technology company providing the eHealth tool. Therefore, UL did not participate in the data collection, analysis, and manuscript writing of the study and was limited to the initial part of the design phase and preparing and reviewing the manuscript.

References

1. Ten Haken I, Ben Allouch S, van Harten WH. The use of advanced medical technologies at home: a systematic review of the literature. BMC Public Health 2018 Feb 26;18(1):284 [FREE Full text] [doi: 10.1186/s12889-018-5123-4] [Medline: 29482550]

XSL•FO
RenderX

2. Barakat A, Woolrych RD, Sixsmith A, Kearns WD, Kort HS. eHealth technology competencies for health professionals working in home care to support older adults to age in place: outcomes of a two-day collaborative workshop. Med 2 0 2013;2(2):e10 [FREE Full text] [doi: 10.2196/med20.2711] [Medline: 25075233]
3. Steindal SA, Nes AA, Godskesen TE, Dihle A, Lind S, Winger A, et al. Patients' experiences of telehealth in palliative home care: scoping review. J Med Internet Res 2020 May 05;22(5):e16218 [FREE Full text] [doi: 10.2196/16218] [Medline: 32369037]
4. Widberg C, Wiklund B, Klarare A. Patients' experiences of eHealth in palliative care: an integrative review. BMC Palliat Care 2020 Oct 14;19(1):158 [FREE Full text] [doi: 10.1186/s12904-020-00667-1] [Medline: 33054746]
5. Magdalena M, Bujnowska-Fedak M, Grata-Borkowska U. Use of telemedicine-based care for the aging and elderly: promises and pitfalls. Smart Homecare Technol Telehealth 2015 May;3:91-105 [FREE Full text] [doi: 10.2147/shtt.s59498]
6. Glenngård AH. The Swedish health care system. In: Tikkanen R, Osborn R, Mossialos E, Djordjevic A, Wharton G, editors. International Profiles of Health Care Systems. Adelaide, Australia: The Commonwealth Fund; 2020:181-189 URL: https://www.commonwealthfund.org/sites/default/files/2020-12/International_Profiles_of_Health_Care_Systems_Dec2020.pdf
7. Lindberg B, Nilsson C, Zotterman D, Söderberg S, Skär L. Using information and communication technology in home care for communication between patients, family members, and healthcare professionals: a systematic review. Int J Telemed Appl 2013;2013:461829 [FREE Full text] [doi: 10.1155/2013/461829] [Medline: 23690763]
8. Rydenfält C, Persson J, Erlingsdottir G, Johansson G. eHealth services in the near and distant future in Swedish home care nursing. Comput Inform Nurs 2019 Jul;37(7):366-372. [doi: 10.1097/CIN.0000000000000536] [Medline: 31135467]
9. Lucas B, Schladitz P, Schirrmeister W, Pliske G, Walcher F, Kulla M, et al. The way from pen and paper to electronic documentation in a German emergency department. BMC Health Serv Res 2019 Aug 09;19(1):558 [FREE Full text] [doi: 10.1186/s12913-019-4400-y] [Medline: 31399096]
10. Göransson C, Eriksson I, Ziegert K, Wengström Y, Langius-Eklöf A, Brovall M, et al. Testing an app for reporting health concerns- experiences from older people and home care nurses. Int J Older People Nurs 2018 Jun 05;13(2):e12181. [doi: 10.1111/opn.12181] [Medline: 29210218]
11. Akhu-Zaheya L, Al-Maaitah R, Bany Hani S. Quality of nursing documentation: paper-based health records versus electronic-based health records. J Clin Nurs 2018 Feb 20;27(3-4):e578-e589. [doi: 10.1111/jocn.14097] [Medline: 28981172]
12. Karlsen C, Ludvigsen MS, Moe CE, Haraldstad K, Thygesen E. Experiences of community-dwelling older adults with the use of telecare in home care services: a qualitative systematic review. JBI Database System Rev Implement Rep 2017 Dec;15(12):2913-2980. [doi: 10.11124/JBISRIR-2017-003345] [Medline: 29219874]
13. Roberts A, Philip L, Currie M, Mort A. Striking a balance between in-person care and the use of eHealth to support the older rural population with chronic pain. Int J Qual Stud Health Well-being 2015 Sep 02;10(1):27536 [FREE Full text] [doi: 10.3402/qhw.v10.27536] [Medline: 26336996]
14. Li J, Talaei-Khoei A, Seale H, Ray P, Macintyre CR. Health care provider adoption of eHealth: systematic literature review. Interact J Med Res 2013 Apr 16;2(1):e7 [FREE Full text] [doi: 10.2196/ijmr.2468] [Medline: 23608679]
15. Cohen Rodrigues TR, de Buisonjé DR, Keesman M, Reijnders T, van der Geer JE, Janssen V, et al. Facilitators of and barriers to lifestyle support and eHealth solutions: interview study among health care professionals working in cardiac care. J Med Internet Res 2021 Oct 15;23(10):e25646 [FREE Full text] [doi: 10.2196/25646] [Medline: 34652280]
16. Brandt CJ, Søgaard GI, Clemensen J, Sndergaard J, Nielsen JB. General practitioners' perspective on eHealth and lifestyle change: qualitative interview study. JMIR Mhealth Uhealth 2018 Apr 17;6(4):e88 [FREE Full text] [doi: 10.2196/mhealth.8988] [Medline: 29666045]
17. Das A, Faxvaag A, Svanæs D. The impact of an eHealth portal on health care professionals' interaction with patients: qualitative study. J Med Internet Res 2015 Nov 24;17(11):e267 [FREE Full text] [doi: 10.2196/jmir.4950] [Medline: 26601678]
18. Enam A, Torres-Bonilla J, Eriksson H. Evidence-based evaluation of eHealth interventions: systematic literature review. J Med Internet Res 2018 Nov 23;20(11):e10971 [FREE Full text] [doi: 10.2196/10971] [Medline: 30470678]
19. Om Aleris. Aleris. URL: https://www.aleris.se/om-aleris/ [accessed 2024-12-16]
20. Green J, Thorogood N. Qualitative Methods for Health Research. 4th edition. Thousand Oaks, CA: Sage Publications; 2018.
21. Thomas DR. A general inductive approach for analyzing qualitative evaluation data. Am J Eval 2006 Jun 01;27(2):237-246. [doi: 10.1177/1098214005283748]
22. SwipeCare - an award-winning system for digital efficiency and automation of care. Strikersoft. URL: https://strikersoft.com/en/focusareas/ehealth/ [accessed 2025-03-10]
23. Vellas B, Guigoz Y, Garry PJ, Nourhashemi F, Bennahum D, Lauque S, et al. The Mini Nutritional Assessment (MNA) and its use in grading the nutritional state of elderly patients. Nutrition 1999 Feb;15(2):116-122. [doi: 10.1016/s0899-9007(98)00171-3] [Medline: 9990575]
24. Smith GB, Redfern OC, Pimentel MA, Gerry S, Collins GS, Malycha J, et al. The National Early Warning Score 2 (NEWS2). Clin Med (Lond) 2019 May;19(3):260 [FREE Full text] [doi: 10.7861/clinmedicine.19-3-260] [Medline: 31092526]
25. Eriksson A, Jutengren G, Dellve L. Job demands and functional resources moderating assistant and registered nurses' intention to leave. Nurs Open 2021 Mar 20;8(2):870-881 [FREE Full text] [doi: 10.1002/nop2.694] [Medline: 33570298]

XSL•FO
RenderX

26. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today 2004 Feb;24(2):105-112. [doi: 10.1016/j.nedt.2003.10.001] [Medline: 14769454]
27. Swedish Ethical Review Act (SFS 2003:460). Swedish Code of Statutes. 2004. URL: https://www.riksdagen.se/sv/dokument-och-lagar/dokument/svensk-forfattningssamling/lag-2003460-om-etikprovning-av-forskning-som_sfs-2003-460/ [accessed 2025-03-06]
28. What the Act says. Swedish Ethical Review Authority. 2025. URL: https://etikprovningsmyndigheten.se/en/what-the-act-says/#fl-main-content [accessed 2025-03-06]
29. EUR-Lex. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). 2016. URL: https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng [accessed 2025-03-06]
30. Laitinen H, Kaunonen M, Åstedt-Kurki P. The impact of using electronic patient records on practices of reading and writing. Health Informatics J 2014 Dec;20(4):235-249 [FREE Full text] [doi: 10.1177/1460458213492445] [Medline: 25411220]
31. Towards a good and equitable health. Public Health Agency of Sweden. 2021. URL: https://www.folkhalsomyndigheten.se/contentassets/bb50d995b033431f9574d61992280e61/towards-good-equitable-health.pdf [accessed 2024-04-29]
32. Mathijssen E, de Lange W, Bleijenberg N, van Houwelingen T, Jaarsma T, Trappenburg J, et al. Factors that influence the use of eHealth in home care: scoping review and cross-sectional survey. J Med Internet Res 2023 Mar 09;25:e41768 [FREE Full text] [doi: 10.2196/41768] [Medline: 36892935]
33. Karnehed S, Erlandsson LK, Norell Pejner M. Nurses' perspectives on an electronic medication administration record in home health care: qualitative interview study. JMIR Nurs 2022 Apr 22;5(1):e35363 [FREE Full text] [doi: 10.2196/35363] [Medline: 35452400]
34. Carlqvist C, Hagerman H, Fellesson M, Ekstedt M, Hellström A. Health care professionals' experiences of how an eHealth application can function as a value-creating resource - a qualitative interview study. BMC Health Serv Res 2021 Nov 05;21(1):1203 [FREE Full text] [doi: 10.1186/s12913-021-07232-3] [Medline: 34740340]
35. Nakrem S, Solbjør M, Pettersen IN, Kleiven HH. Care relationships at stake? Home healthcare professionals' experiences with digital medicine dispensers - a qualitative study. BMC Health Serv Res 2018 Jan 15;18(1):26 [FREE Full text] [doi: 10.1186/s12913-018-2835-1] [Medline: 29334953]
36. Ramachandran M, Brinton C, Wiljer D, Upshur R, Gray CS. The impact of eHealth on relationships and trust in primary care: a review of reviews. BMC Prim Care 2023 Nov 03;24(1):228 [FREE Full text] [doi: 10.1186/s12875-023-02176-5] [Medline: 37919688]
37. Nilsson Å, Skär L, Söderberg S. Nurses' views of shortcomings in patent care encounters in one hospital in Sweden. J Clin Nurs 2015 Oct 16;24(19-20):2807-2814. [doi: 10.1111/jocn.12886] [Medline: 26177676]
38. Jazieh AR, Volker S, Taher S. Involving the family in patient care: a culturally tailored communication model. Glob J Qual Saf Healthc 2020;1(2):33-37. [doi: 10.4103/JQSH.JQSH_3_18]

Abbreviations

EHR: electronic health record
HCP: health care professional
MNA: Mini Nutritional Assessment
NEWS2: National Early Warning Score 2

Edited by C Jacob; submitted 17.05.24; peer-reviewed by R Marshall, C Göransson; comments to author 06.12.24; revised version received 11.01.25; accepted 07.02.25; published 24.03.25.

Please cite as:
Rivas EV, Lesley U, Davoody N
Health Care Professionals' Perspectives on Using eHealth Tools in Advanced Home Care: Qualitative Interview Study
JMIR Hum Factors 2025;12:e60582
URL: https://humanfactors.jmir.org/2025/1/e60582
doi:10.2196/60582
PMID:

©Eric Vincent Rivas, Ulf Lesley, Nadia Davoody. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 24.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information,

XSL•FO
RenderX

a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Need Analysis of Clinician-Oriented Integrated Precision Oncology Decision Support Tools: Qualitative Descriptive Study

Sheng Zheng[1*], BA; Yi Feng[2*], PhD; Jieran Long[3], MD; Xiang Li[4], MD; Meng Zhang[2], PhD; Hui Chen[1], BA; Xue Du[2], PhD; Huilong Duan[1], PhD; Shuqin Jia[5], PhD; Nan Wu[6], MD; Xudong Lu[1], PhD

[1]College of Biomedical Engineering & Instrument Science, Zhejiang University, Hangzhou, China

[2]Key laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Center for Molecular Diagnosis, Peking University Cancer Hospital & Institute, Beijing, China

[3]Department of Thoracic Oncology, Peking University Cancer Hospital & Institute, Beijing, China

[4]Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Department of Thoracic Surgery II, Peking University Cancer Hospital & Institute, Beijing, China

[5]State Key Laboratory of Holistic Integrative Management of Gastrointestinal Cancers, Beijing Key Laboratory of Carcinogenesis and Translational Research Center for Molecular Diagnosis, Peking University Cancer Hospital & Institute Beijing, Beijing, China

[6]State Key Laboratory of Molecular Oncology, Beijing Key Laboratory of Carcinogenesis and Translational Research, Department of Thoracic Surgery II, Peking University Cancer Hospital & Institute, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Xudong Lu, PhD
College of Biomedical Engineering & Instrument Science
Zhejiang University
Xihu District, Zheda Road 38
Hangzhou, 310058
China
Phone: 86 13957118891
Email: lvxd@zju.edu.cn

Abstract

Background: The rapid advancement of next-generation sequencing has significantly expanded the landscape of precision medicine. However, health care professionals face increasing challenges in keeping pace with the growing body of oncological knowledge and integrating it effectively into clinical workflows. Precision oncology decision support (PODS) tools aim to assist clinicians in navigating this complexity, yet their current functionalities only partially address clinical needs. A lack of comprehensive needs assessment may result in unaddressed requirements, limiting the effectiveness of these tools in real-world practice.

Objective: This study aimed to explore clinicians' needs and expectations regarding the functionalities of integrated PODS tools, providing insights into essential features that could enhance their usability and impact.

Methods: We conducted a qualitative investigation at Peking University Cancer Hospital to explore clinicians' needs and expectations for the functions of integrated PODS tools. Data were collected through 143 structured participant observations during multidisciplinary team meetings and 17 in-depth semistructured interviews with a diverse group of oncology specialists, including physicians, surgeons, molecular biologists, radiotherapists, radiologists, and pathologists. Thematic analysis was applied to identify key functional requirements, and a requirements framework was formed.

Results: Three overarching functional needs emerged: (1) better access to oncological knowledge, including support for therapy selection (guidelines, conferences, and consensuses), clinical trials, drug and treatment information, and complex case knowledge, as well as improved diagnostic and prognostic insights; (2) clinical contextualization and resource navigation, referring to the process of contextualizing scientific knowledge within real-world clinical settings, including access to clinical trials and drugs, along with predictive models for treatment response; and (3) support abilities in the decision-making process, highlighting the need for integration of flexible biological knowledge and phenotypic data; automated patient information synthesis; improved data visualization; and optimized retrieval, recommendation, and question-answering functionalities. A functional framework for integrated PODS tools was proposed based on these findings.

XSL•FO
RenderX

Conclusions: The study conducted a qualitative descriptive observation and interview in the use, needs of integrated PODS tools. PODS tools serve as complex, multilevel decision support systems. A clear understanding of clinicians' actual needs is crucial for their refinement and practical adoption. By capturing perspectives directly from oncology professionals, this study provides actionable insights into the functional enhancements required for PODS tools, ultimately aiming to bridge the gap between genomic advancements and clinical decision-making in precision oncology.

(JMIR Hum Factors 2025;12:e67476) doi:10.2196/67476

KEYWORDS

precision oncology decision support; needs analysis; structured observation; semistructured in-depth interview; functional framework; artificial intelligence

Introduction

Background

Advancements in next-generation sequencing have paralleled the rise of precision oncology, leading to the discovery of numerous diagnostic biomarkers for cancer and the approval of many targeted drugs by regulatory authorities, such as the US Food and Drug Administration, European Medicines Agency, and National Medical Products Administration, as well as recommendations by prestigious organizations such as the National Comprehensive Cancer Network), European Society for Medical Oncology, and Chinese Society of Clinical Oncology (eg, tepotinib [1] and sunvozertinib [2]). Despite this progress, health care professionals face challenges in keeping pace with these advancements and determining the most effective treatment strategies, as well as handling complex genomic information, such as regularly updating information and understanding the clinical significance of genomic alterations [3].

Currently, clinicians mainly rely on genetic testing reports to acquire the genomic profile of patients with cancer and to understand the clinical significance of certain genetic variants [4]. With rapidly growing knowledge of precision oncology, the content of genetic testing reports has greatly expanded, which also brings the problem of inconvenience in reading and searching [5]. Moreover, the heavy clinical workload significantly limits the time available to review the full contents of genetic testing reports [6]. A fast and effective way to display comprehensive genomic information may help clinicians locate the needed information easily [7].

Under these circumstances, computer-based interactive precision oncology decision support (PODS) tools have emerged. In oncology, tools with high interactivity and interoperability, such as knowledge and clinical trial (CT) databases [8-10], therapy recommendation engines [11], literature search tools [12], and predictive algorithm–based models [13], have been developed and are anticipated as valuable resources for facilitating the timely update and informed interpretation of genomic data from bioinformatics research to clinical statistical analysis [14-17]. The literature has described the properties, opportunities, and future effects of these tools [18,19].

However, it should be noted that although these tools may bring broad prospects in oncology research, they may not fully meet the needs of practical clinical application [20]. There are already systems integrating some of these functions of PODS tools for clinicians [21,22], but 2 key issues of current clinician-oriented PODS tools remain unclear.

The first issue is that clinician-oriented PODS functions are scattered across different tools. When somatic mutations and germline variants were detected in patients with cancer, to make a more appropriate and individualized treatment plan based on the patients' genomic profile, some functions, such as (1) whether these variants are pathogenic or oncogenic [23], (2) how they are involved in tumorigenesis and cancer progression [24,25], (3) what their biological functions and clinical significance are [26,27], (4) to what extent they potentially affect the patients' diagnosis and treatment [28], and (5) whether they will influence the choice of certain therapeutic strategy or clinical management, may be clinicians' interests [29-32]. However, currently, a single available knowledge base could only meet 1 or 2 aspects of the abovementioned clinical requirements. To answer all these questions, clinicians must search in multiple knowledge bases using various search engines to achieve 1 goal, as different evidence may be scattered across diverse databases [33]. Such a process of switching tools normally costs a large amount of time with limited achievement.

The second issue is that some clinician needs regarding PODS tools remain unclear. Clinicians from different specialties may require different PODS functions, and while some commonalities could be found through needs analysis, the current lack of need analysis results in current PODS tools not yet meeting all practical requirements from clinicians. This fact restrains the clinical implementation of PODS tools and underestimates the clinical value of genomic data from patients with cancer [7]. Moreover, the lack of need analysis exposes the problem that there is currently no appropriate evaluation metric for the measurement of the quality of developing such metrics for PODS tools [34].

Objectives

This study aimed to explore clinicians' perceptions of functional needs in PODS tools. It also seeks to identify the existing urgent needs not met by current tools and outline a functional need framework for PODS tools.

Methods

Study Design and Theoretical Framework

This qualitative descriptive study, informed by a naturalistic perspective and grounded theory [35,36], investigates how clinicians use PODS tools in multidisciplinary teams (MDTs)

XSL•FO
RenderX

and routine clinical practice. Comprising experts such as molecular biologists, physicians, surgeons, radiotherapists, radiologists, pathologists, and specialists from other departments (Table 1), MDT plays a crucial role in integrating clinical data and molecular profiling to tailor targeted therapies [37-39]. In routine clinical practice, such as in outpatient services, clinicians see many patients with simple to complex medical conditions. When the course of disease and treatments has either no clinical guideline to be followed or lacks evidence to support standard treatments, the case is raised for discussion in an MDT. Therefore, structured observation of MDT discussion was selected to investigate the need for PODS tools for clinicians and to propose elements for a framework of PODS tools. These elements were then validated and further verified by in-depth semistructured interviews, which might not only reflect the perspective of clinicians and quickly capture their views as a complement to objective observation but also constantly explore the deep reasons behind the ideas for unclear clinical requirements.

Table 1. Characteristics of MDT[a] records.

Site	Thoracic Oncology Center	Gastrointestinal Oncology Center
Observations, n	26	20
Records, n	86	57
Participants, n	15-40	10-20
Patients with genetic testing, n (%)	39 (45)	12 (21)
Frequency	Once a week	Once a week
Duration (min)	10-120	20-40
Discussion duration per patient (min)	10-20	10-15
Speakers per discussion, n	4-10	3-6
Participating specialties		
Surgeon	✓	✓
Physician	✓	✓
Radiotherapist	✓	✓
Radiologist	✓	✓
Pathologist	✓	✓
MDC[b] staff[c]	✓	✓[d]
Nursing	✓	—[e]
TCM[f]	✓	—

[a]MDT: multidisciplinary team.

[b]MDC: Molecular Diagnostic Center.

[c]Molecular biologists and geneticists from the Molecular Diagnostic Center.

[d]The molecular biologists and geneticists are present, but because of the genetic test information from the patients, they seldom speak in multidisciplinary teams and the Traditional Chinese Medicine Department.

[e]Not applicable.

[f]TCM: traditional Chinese medicine.

The study adhered to the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist [40]. Conducted from June 2023 to June 2024, the study used a multiple qualitative approach, combining structured participant observation and in-depth semistructured interviews to comprehensively assess the use and integration requirements of PODS tools [41].

Setting and Data Collection

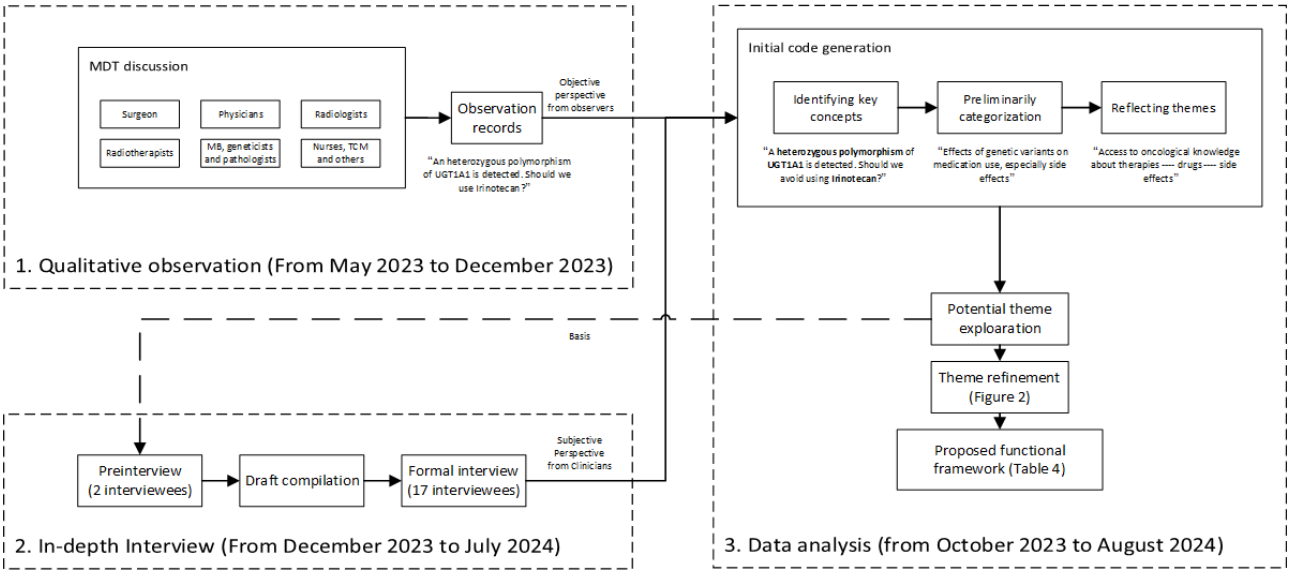
Structured participant observations and in-depth semistructured interviews were conducted at Peking University Cancer Hospital (PUCH). Participant observation data were collected during MDT meetings held by the Thoracic Oncology Center and Gastrointestinal Oncology Center; these meetings occurred weekly and involved large discussion groups, with an average of 19.92 (SD 5.51) participants per session.

The process of the research could be divided into 3 parts: qualitative observation, in-depth semistructured interview, and data analysis (Figure 1). MDT principals and interviewees all signed informed consent. Interviewees were selected using sampling theory and snowball sampling [42], based on the following criteria: (1) holding a medical certificate, (2) working in tumor-related clinical practice, (3) participating in MDTs, and (4) having experience using PODS tools or having heard of or observed the operation of PODS tools. Data saturation was defined as the point at which <5% of the total codes were new across 3 consecutive MDT observations. Snowball sampling

XSL•FO
RenderX

was terminated when no new codes or ideas emerged from the 2 most recent interviews. Each interview took place at the clinicians' lounge.

Figure 1. The process of the research. MB: molecular biologist; MDT: multidisciplinary team; TCM: traditional Chinese medicine.



Qualitative Structured Participant Observation

From May to December 2023, 2 team members, SZ and YF, observed MDTs at the Thoracic Oncology Center and Gastrointestinal Oncology Center. YF participated fully in MDT discussions as a member of the Molecular Diagnostic Center (MDC) and a participant-as-observer, while SZ, a doctoral candidate in medical informatics trained in qualitative research, took part as an observer-as-participant. The combination of the 2 observers balanced the perspectives of both clinical practice and informatics. SZ made preliminary records of the observations (Multimedia Appendix 1 [43-45]), and these records were subsequently refined in collaboration with YF. The MDT at PUCH can be divided into three parts: (1) patient information preparation before the MDT (*preparation*), (2) patient information briefing during the MDT (*briefing*), and (3) multidisciplinary discussion (*discussion*). The observation of this research mainly focused on the discussion part, which was recorded in detail to identify PODS needs related to genomic interpretation, diagnosis, therapeutic recommendation, and possible functional requirements of PODS tools that emerged during the discussion (a description of the phenomena, the process of MDTs, and detailed observation of the other 2 parts are provided in Multimedia Appendix 1). To ensure methodological rigor, we implemented a data saturation verification protocol. For proactive monitoring, we tracked emerging codes through real-time coding logs during observations and conducted weekly codebook audits using NVivo's (Lumivero) code frequency matrices to ensure that it achieved semantic saturation [46]. The slides presented during MDTs were obtained after the meeting to refine the records. The team observed 26 Thoracic Oncology Center MDTs (86 records) and 20 Gastrointestinal Oncology Center MDTs (57 records), with each session lasting approximately 35 minutes on average.

In-Depth Semistructured Interview

Semistructured, in-depth qualitative interviews were conducted from December 2023 to July 2024 by SZ and YF, with SZ serving as the main interviewer and having no prior interaction with the participants. YF, a colleague of the interviewees, assisted SZ in refining and complementing questions. The semistructured interview questions were developed based on the potential themes that emerged from the initially coded observation record data. Before the formal interview, 2 interviewees were invited to the preinterview session, during which the draft interview guide was compiled based on the initial coding of the qualitative observation records. Both interviewers maintained reflexive journals to document emergent assumptions and field observations. The formal interview guide was later standardized to focus on clinicians' use of genetic test reports and their needs and expectations for PODS tools (Multimedia Appendix 1). Clinicians were encouraged to provide additional perspectives on PODS and to share their experiences using PODS tools through open-ended questions during the interview. After each interview, potential themes were extracted. If data saturation had not been reached, additional interviewees were recruited. The demographic details are summarized subsequently. The interviews were conducted in Chinese, audiotaped with a digital recorder, transcribed verbatim, and edited by SZ. A total of 12 hours of individual interviews were transcribed, resulting in approximately 48,000 words. The interviews lasted approximately 40 (range 26-70) minutes.

Data Analysis

The observation and interview data were thematically analyzed following the reflexive thematic analysis by Braun and Clark [47], which includes familiarization with the data, generating initial codes, exploring potential themes, reviewing themes, defining and refining themes, and finalizing findings. The data analysis was divided into 2 steps. Between the observation and interview phases, the observation record data were initially

XSL•FO
RenderX

coded, and the potential themes that emerged formed the basis of the interview. After each interview, the themes explored in the interview were compared with the initial potential themes as a form of constant comparison and validation [38]. This was considered part of the review process. Subsequently, the 2 classes of themes were combined and refined. The functional framework of PODS tools was extracted through the themes and concluded through a team discussion. The interviewees were invited to confirm whether their perspectives were fully represented during the construction of the framework. NVivo 14 software was used for the analysis, with coding performed by SZ and reviewed by YF. Cohen κ was calculated to confirm strong coding consistency. Discrepancies were resolved through team discussions.

Ethical Considerations

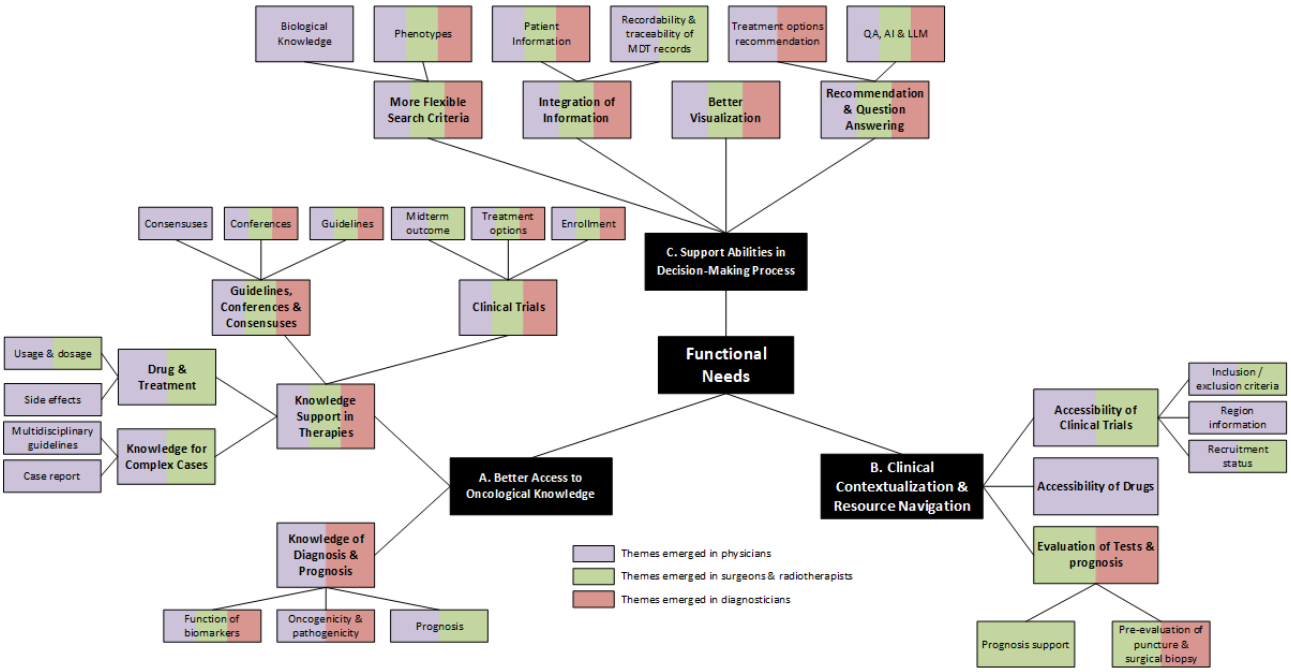
The study received ethics approval from the ethics committee of PUCH (2023KT35). Informed consent was obtained from all interview participants for both the interviews and the secondary analysis of the interview data. All observation and interview data were recorded anonymously. The interviewees received a small thank-you gift of approximately CNY ¥50 (US $7.00) after the interview.

Results

Qualitative Observation and Interviewee Characteristics

A total of 46 MDT discussions were qualitatively observed, involving treatment decisions for 143 patients (Table 1). For the 46 MDT observations, saturation was reached by the 41st MDT, and no new codes emerged in the 45th and 46th MDT (Figure S1 in Multimedia Appendix 1). After 6 months of observation, the research team began the first round of exploring potential themes. Potential needs for PODS tools emerged from our observations (Tables S2-S6 in Multimedia Appendix 1), which were refined and validated by the in-depth semistructured interviews and further constructed into the themes of functional needs in PODS (Figure 2). The MDT in the Thoracic Oncology Center had significantly more detailed information and longer discussion times for genetic testing and personalized targeted therapy compared to the MDT in the Gastrointestinal Oncology Center. In addition, the use of PODS tools was more prevalent in the Thoracic Oncology Center.

Figure 2. The 3 key themes emerging from the qualitative observation and interviews: (1) better access to oncological knowledge: clinicians emphasized the importance of precision oncology decision support (PODS) tools in facilitating access to high-level precision medicine evidence, including guidelines, treatment information (eg, use, dosage, and side effects), clinical trials, multidisciplinary guidelines, and case reports; (2) clinical contextualization and resource navigation: PODS tools aim to integrate fragmented clinical information, such as trial eligibility, drug accessibility, and evaluation models, enabling clinical contextualization and resource navigation for practical application; (3) support abilities in the decision-making process: clinicians anticipated that the PODS design would enhance decision-making efficiency. They also highlighted advancements in information systems, including user-friendly interfaces, enhanced search functions, chatbots, and large language model applications, as key features to improve workflow and clinical decision-making. AI: artificial intelligence; LLM: large language model; QA: question answering.



On the basis of these findings, we focused our interviews with clinicians in the Thoracic Oncology Center on their use of PODS tools. We also conducted interviews with clinicians from the Radiotherapy Department and the MDC.

A total of 17 clinicians participated in in-depth semistructured interviews (Table 2). For the 17 interviews, saturation was achieved by interview 15, with no new codes appearing in interviews 16 and 17 (Figure S2 in Multimedia Appendix 1). All 17 clinicians had experience using genetic testing reports, while 59% (n=10) had personal experience using at least 1 interactive PODS tool. However, all 17 clinicians had heard of or observed the use of PODS tools. The potential needs were screened and confirmed during the interview. After data

XSL•FO
RenderX

analysis, 3 themes emerged from our observations and interviews (Figure 2; Figures S3 and S4 in Multimedia Appendix 1; and Tables S2-S6 in Multimedia Appendix 1). Themes that emerged from the interview data and were organized by knowledge, feasibility, and support ability levels are described subsequently and presented in Table 3. From these 3 themes, the functional framework of integrated PODS tools is described in detail in the subsequent sections.

Table 2. Characteristics of the interviewees (N=17)[a].

Characteristics	Interviewees, n (%)
Sex	
Male	10 (59)
Female	7 (41)
Years of working	
<10	6 (35)
10-20	8 (47)
>20	3 (18)
Clinical duty	
Physician	9 (53)
Surgeon	5 (29)
Molecular biologist	2 (12)
Radiotherapist	1 (6)
Professional title	
Chief	2 (12)
Associate chief (deputy chief)	7 (41)
Attending chief (supervising)	5 (29)
Attending	3 (18)
Preparer	6 (41)
Had used at least 1 PODS[b] tool	10 (59)

[a]The characteristic “preparer” refers to a role in the multidisciplinary team, responsible for collecting patient information, preparing multidisciplinary team briefing slides, and presenting the information during the meeting. Further description of the preparer is provided in Multimedia Appendix 1.

[b]PODS: precision oncology decision support.

XSL•FO
RenderX

Table 3. Sources of themes from key informants, translated into English.

Theme and subtheme	Illustrative quote and quote number
Better access to oncological knowledge	
Knowledge support in therapies	
Guidelines, conferences, and consensuses	• "Before the MDT, Dr. [name] would like to share recent progress of [progress in treatment of specific disease]..." [Observation 008, 028, and 079] (Quote 1) • "The way I get access to the oncological knowledge are usually from WeChat official accounts, going to conferences, and newly updated information shared by peer expertise. Sometimes I also search for literatures or reviews in my spare time...What I'm most interested in is the change of the latest version update guideline, from which I could learn a lot...would be great if there is a notification of the changes..." [005, associate chief physician] (Quote 2) • "We usually discuss the latest progress of conferences before the MDT starts. If PODS tools could notify us this information in our spare time, it would help a lot in summarizing this information for discussion..." [006, associate chief physician] (Quote 3)
Clinical trials	• "About 25-50% of patients in our hospital would require enrollment in CTs...develop a real-time updating CT recommendation tool is the best..." [004, attending physician] (Quote 4) • "...The information in the website of CTs updates frequently...requires us to retrieve and summarize the latest eligibility criteria and enrollment every half month to one month...so time-consuming..." [009, associate chief surgeon] (Quote 5) • "...Latest progresses are often discussed in MDT...not so convenient to check...we want future PODS tools could capture these in real time..." [005, associate chief physician] (Quote 6) • "Before the MDT, Dr. [name] would like to share the midterm result of [progress of specific CT] for us..." [Observation 033 and 082] (Quote 7)
Drugs and treatment options	• "Genetic test reports that vormetinib is effective in LUAD patients with EGFR exon 20 insertion mutation...the descriptions are not particularly clear...we sometimes want to understand the mechanism of drugs..." [003, chief surgeon] (Quote 8) • "Osimertinib, vormetinib and almonertinib are commonly used among the third generation of EGFR-TKIs...there are differences in their use, which requires years of clinical experience...We hope PODS tool will support us in this regard..." [010, attending physician] (Quote 9) • "...Some drugs are needed for four cycles, but some patients' health condition is poor...We want to know whether can I reduce the dosage...sometimes drug suppliers would tell us some information, but that's not enough..." [005, associate chief physician] (Quote 10) • "...The two most common and risky side effects are vomiting and leukopenia...need prompt intervention...if PODS tools could have an alert or reminder function of these side effects, it would help us reduce unnecessary prescriptions to alleviate these side effects..." [001, associate chief physician] (Quote 11)
Knowledge for complex cases	• "...when the patient has other cancers...we would invite colleagues from the corresponding departments for MDT discussion, but it is brief...If we can learn easily from other disciplines, we can make treatment decisions more quickly..." [006, associate chief physician] (Quote 12) • "Today's discussion is about the patient with malignant pleural mesothelioma for whom conventional chemotherapy drugs did not respond...if any doctors have seen relevant case reports?" [Observation 046] (Quote 13)
Knowledge support in diagnosis and prognosis	• "...some colleagues asked me the differences between germline variants and somatic mutations...PODS tools could be a useful way to help my colleagues understand the biological knowledge like this..." [014, supervising molecular biologist] (Quote 14) • "...we sometimes want to know the biological function of molecular alterations, understanding why the alteration happens could better guide us to understand the mechanism of the disease..." [013, associate chief surgeon] (Quote 15) • "...previously encountered a LUAD patient with co-mutations in TP53, RB1, and EGFR...no appropriate standard treatment or diagnosis options but continued observation at the time...an article said that the concurrence of three variants may be induced to small cell cancer transformation, which is an important reference basis to do puncture for diagnosis...hope to have a better mechanism to search evidences..." [011, attending physician] (Quote 16[a]) • "...we often evaluate the prognosis between different therapies, like radiotherapies and targeted therapies...the knowledge of which could help us make decision better..." [017, attending chief radiotherapist] (Quote 17)
Clinical contextualization and resource navigation	

XSL•FO
RenderX

Theme and subtheme	Illustrative quote and quote number
Accessibility of CTs[b]	• "When there is a patient needs to be enrolled, we will look at the inclusion and exclusion criteria in the CTs we have printed (take out some hard copy of CTs) and compare the patient's medical status with the criteria item by item...it is very time consuming...Our CT screening usually follow this sequence: 'First, CTs held in our ward; Second, CTs held in our hospital; Third, CTs held in Beijing's Hospital; Fourth, different places other than Beijing in China; Lastly, we consider overseas'...CT enrollment is usually based on the screening order of inside the department, hospital, Beijing, China and foreign countries...We need to manually screen it ourselves...want an automatically sort..." [002, attending surgeon] (Quote 18) • "Sometimes we know there is an enrollment of CT, but we don't know the up-to-date recruitment status. Future PODS tools or CT website should update the status more frequently." [006, associate chief physician] (Quote 19)
Accessibility of drugs	• "...Osimertinib is covered in Drug Catalog, but vormetinib and almonertinib need to be paid at patients' own expense..." [001, associate chief physician] (Quote 20) • "Clear mark of drug catalog and price information of drugs can make us discuss the most suitable treatment option for with patient him/herself." [012, associate chief physician] (Quote 21)
Prediction models: the necessity of ordering diagnostic tests and prediction of prognosis	• "...some genetic tests are expensive...I have known that there are some radiology technologies to auxiliary identify whether there is a certain risk of mutations..." [012, associate chief physician] (Quote 22) • "...examinations like puncture examinations are invasive...we want to use PODS tools to help estimate the necessity of such invasive examinations..." [013, associate chief surgeon] (Quote 23) • "...common mutations like TP53 have a poor prognosis...We want to combine the molecular alteration with prognostic prediction..." [002, attending surgeon] (Quote 24)
Support abilities in the decision-making process	
Using more flexible biological knowledge and phenotypes as search criteria	• "...we once had a patient with MEK mutation and ERK wild type...there were not many drugs targeting MEK...wanted to find some CTs targeting ERK...had to reset the search criteria which was inconvenient...would like to have a flexible retrieval mechanism..." [008, associate chief physician] (Quote 25) • "...would be great if PODS tools could automatically match the inclusion criteria with patient's information and tell us directly whether this patient is a qualified candidate...helps us reduce a lot of comparison work..." [007, chief physician] (Quote 26)
Automatic integration of patient information	• "...would be great if we just input the ID of patient and click a button...all the information scattered in HIS, PACS...is organized..." [011, attending physician] (Quote 27) • "Acquiring the radiology information is inconvenient currently. We have patients all over the country, and they have a variety of formats and types of image information. We hope that the future of PODS tools could help us to reduce the burden of identification and input of these data." [010, attending physician] (Quote 28) • "Suppose there is a timeline of patient in MDT, and all the events of this patient are on the line. Click one button and the details of the event will be shown. That's amazing..." [002, attending physician] (Quote 29) • "Some patients will be discussed for several times. We want the PODS tools to record our discussion every time." [004, attending physician] (Quote 30) • "This patient was discussed [time] ago. Last time we decided to use [treatment option], but the disease progressed now..." [Observation 033] (Quote 31)
Better visualization of information	• "...Switching slides back and forth to watch the texts is inconvenient...hard to sort out the therapeutic relationships..." [010, attending physician] (Quote 32) • "...automatically generated timeline of history is more logical and visualizable..." [002, attending surgeon] (Quote 33) • "...when there is a patient with rare mutation...may want to find previous patients who have the same or similar mutation...may be helpful for diagnosis and treatment...some simple automated generated statistical distribution plots are intuitive to see and discuss..." [014, supervising molecular biologist] (Quote 34)
Optimization in retrieval, recommendation, and question answering	• "...would be great if PODS tools could automatically match the inclusion criteria with the patient information..." [005, associate chief physician] (Quote 35) • "...don't have enough time to read and search the evidences...we may just input some key search information to AI, and ask what latest information it finds..." [015, associate chief physician] (Quote 36) • "I wonder if there's a day if AI could be a real assistant for me. Nowadays there is chatGPT, I believe in future there will be a chatbot in oncology area, knows everything, tells us what is the best treatment option...future PODS tools may be like this..." [002, attending surgeon] (Quote 37)

[a]The interviewee referenced the study by Lai et al [43].

[b]CT: clinical trial.

XSL•FO
RenderX

The percentages in brackets in the sections below represent the percentage of interviewees who reported the corresponding functional requirements. The percentages reflect only the proportion of participants who mentioned a specific need during the interviews and do not indicate the relative importance of the needs.

Theme 1: Better Access to Oncological Knowledge

Oncological knowledge plays an important role in PODS, and the most frequently mentioned need for PODS was faster and better access to the latest oncological information. Precisely, comprehensively, and timely integrating the knowledge of diagnosis, treatment, and prognosis from various sources, such as knowledge bases, academic conferences, and guidelines, is the most valuable function of PODS tools.

Knowledge Support in Therapies (17/17, 100%)

All interviewees emphasized the importance of the function of knowledge acquisition for therapies in support of clinical practice in oncology.

Guidelines, Conferences, and Consensuses (12/17, 71%)

Clinicians paid much attention to the latest progress in oncology and updates released in guidelines, conferences, and consensuses. However, the current access to this information was mainly through manual search and information shared by colleagues (quote 1). To meet these requirements, a news update session is held before the MDT discussions in the Thoracic Oncology Center to review and interpret the latest updates released at academic conferences or by professional associations. However, once-a-week knowledge updates provided by the attending MDT may not meet the urgent needs during daily clinical practice. Clinicians hoped that this latest information could be updated in PODS tools in a timely manner (quotes 2 and 3).

CT Need (14/17, 82%)

All physicians referred to this need. Physicians had a greater need in this aspect, as most patients in the Thoracic Oncology Department present with advanced-stage conditions and require systemic treatment. After the failure of guideline-recommended therapies, appropriate CTs had to be considered. Approximately 25% to 50% of patients would require enrollment in CTs (quote 4), necessitating that physicians have a comprehensive understanding of the available options. The 2 major channels to access up-to-date CT information were manual searches of CT registration websites and information shared by physicians responsible for CT recruitment and data collection. However, frequent retrieval of this information imposed a burden on their workload. Timely and automatic updates of CT information could help alleviate this burden (quotes 4 and 5). Midterm outcomes of CTs were also given more attention (quotes 6 and 7).

Drugs and Treatments (8/17, 47%)

Surgeons and radiotherapists hoped to access the latest knowledge on targeted therapy, immunotherapy, and chemotherapy drugs through PODS tools (quote 8). Physicians had greater demands for drug information, including their use and dosage (quotes 9 and 10). Moreover, clinicians paid close attention to the side effects of the treatment regimens. It was important to be aware of and informed about the potential side effects of certain treatment options for every patient. At present, the prediction of side effects for certain therapeutic strategies still relies on clinicians' experience, while current PODS tools provide limited support (quote 11).

Knowledge for Complex Cases (4/17, 24%)

In some complex cases, physicians expressed a need to acquire more evidence, such as multidisciplinary guidelines and case reports. When encountering patients with a history of multiple primary cancers, convenient access to multidisciplinary guidelines was identified as an important need (quote 12). For example, a patient undergoing targeted therapy for lung cancer was found to have a mass during colonoscopy, but it remained unclear whether the lesion was a primary colorectal tumor or a metastasis from the lung cancer. Furthermore, clinicians hoped to find similar case reports when managing patients with rare tumor types (quote 13).

Knowledge Support in Diagnosis and Prognosis (6/17, 35%)

In addition, interviewees sometimes mentioned diagnostic and prognostic evidence as important for the selection of treatment options. Especially in the area of diagnosis, experts from MDC and the Pathology Department were often asked to explain the functions, oncogenicity, and pathogenicity of specific biomarkers and molecular alterations to interpret their potential clinical significance (quote 14). PODS tools could help collect and filter related evidence from various knowledge bases and assist in a more comprehensive explanation and interpretation (quote 15). Moreover, current PODS tools lack reliable search functions for these purposes (quotes 16 and 17).

Theme 2: Clinical Contextualization and Resource Navigation

When a treatment option was acknowledged or proposed, evaluating its real-world applicability, including insurance coverage criteria, availability of CTs and drugs, and relevant prognostic models, was necessary. Much of this work was repetitive and could be automated by the PODS tools, thereby accelerating decision-making and improving care coordination.

The Accessibility of CTs (9/17, 53%)

Whether patients had the channel to access appropriate CTs was as important as meeting the inclusion and exclusion criteria (quote 18). Geographic barrier seemed to be an important influencing factor that might directly affect CTs' accessibility (quote 18). Furthermore, the up-to-date recruitment status was another essential factor that could help clinicians filter the available CTs and improve the effectiveness of patients' enrollment (quote 19).

The Accessibility of Drugs (5/17, 29%)

The patient's affordability was 1 of the biggest factors, after medical condition, when considering treatment options (mentioned by all interview participants). The drugs mentioned in knowledge bases or clinical guidelines might not be practically accessible because of cost. In such cases, clinicians wanted to know whether the cost of certain therapeutic drugs

XSL•FO
RenderX

could be covered by the Chinese medical insurance policy or by commercial insurance, and they considered patient affordability when making decisions about a reasonable therapeutic regime (quotes 20 and 21).

Prediction Models: Necessity of Ordering Diagnostic Tests and Prediction of Prognosis (4/17, 24%)

Currently, genomic information can be acquired through genetic testing of a peripheral blood sample. However, tissue samples are sometimes needed to obtain more precise test results. The methods for obtaining tissue are usually invasive, such as puncture, bronchoscopy, or surgery, which often cause pain to patients. Clinicians were cautious about ordering invasive and unaffordable tests and preferred a less expensive way to assess the need for these tests (quotes 22 and 23). In addition, some surgeons reported using prognostic prediction models to help estimate the risk of surgery (quote 24).

Theme 3: Support Abilities in the Decision-Making Process

During MDTs, there was also a clear need for support in the decision-making process.

Using More Flexible Biological Knowledge and Phenotypes as Search Criteria (10/17, 59%)

Clinicians expressed a demand to incorporate biological knowledge, such as the function of variants or signaling pathways, into their searches for information (quote 25). They also wanted to combine phenotypes in the process of identifying better treatment options. The most frequently mentioned phenotypes were stage, number of treatment lines (ie, the treatment course), location of lesion, presence of multiple primary cancers, and physical conditions such as Eastern Cooperative Oncology Group score and complications. Automatic matching of patients based on phenotype and inclusion and exclusion criteria was also a frequently mentioned need (quote 26).

Automatic Integration of Information (8/17, 47%)

Summarizing information during the preparation phase was described as repetitive and time-consuming. All preparers mentioned that it would be beneficial to have an automatic patient information integration function (quotes 27-29). The ability to record the discussions in MDT was also considered a valuable function, as some patients were discussed several times. Tracking not only the patient's diagnostic and treatment history but also the previous MDT discussion conclusion was seen as extremely helpful (quotes 30 and 31).

Better Visualization of Information (3/17, 18%)

Some clinicians noted the inconvenience of slide-based presentations, especially when a patient's medical history was long and had to be spread across several pages. In such cases, it was not uncommon for gaps in the patient's diagnostic and treatment information to occur. They preferred a better visualization of patient information (quotes 32 and 33). Some clinicians were willing to see an automatic chart display function added to PODS tools to show statistical results from previous real-world genomic data more clearly. This improved data display function could potentially assist clinicians in viewing cancer mutation profile and understanding their relationship with clinical phenotype, therapeutic effectiveness, and prognosis more directly. Furthermore, such a function could not only help clinicians optimize clinical decisions based on previous in-house clinical evidence but also support scientists in designing clinical studies based on objective analysis (quote 34).

Optimization in Retrieval, Recommendation, and Question Answering (5/17, 29%)

In addition to search optimization (mentioned in Better Access to Oncological Knowledge section), clinicians wanted PODS tools to rank treatment options and CTs intelligently and to adjust recommendations as search conditions change (quote 35).

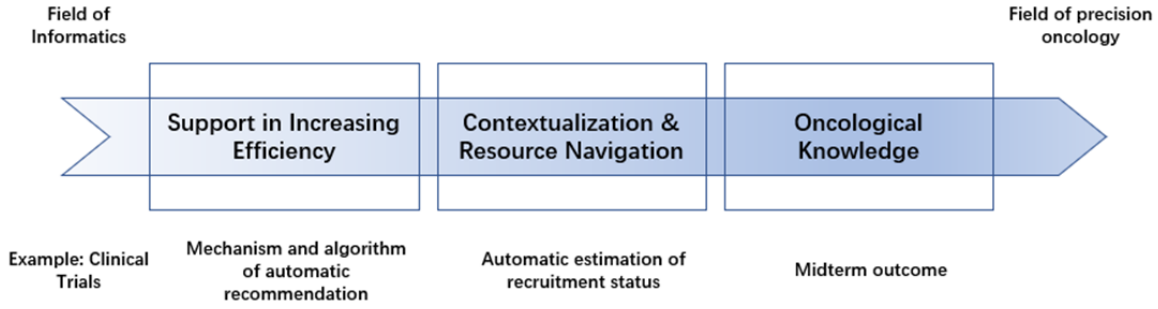
The popularity of language models has led clinicians to hope that artificial intelligence (AI) could help interpret clinical evidence in the literature through question answering (quotes 36 and 37).

The Categorization of the 3 Main Thematic Needs

During the thematic analysis, we maintained a central objective: to leverage rapidly advancing information systems technology to support decision-making in precision oncology. This objective naturally extends to 2 fundamental scenarios: 1 focused on assisting clinicians in understanding precision oncology knowledge, which is more aligned with the medical field, and the other focused on directly using advancements in IT to help clinicians improve diagnostic and treatment efficiency, which leans more toward informatics. During the data analysis process, we found that clinicians may seek data unrelated to clinical knowledge but still critically important or infer implicit clinical decisions from the data. These 2 scenarios may lie at the intersection of medicine and informatics. On the basis of this, we categorized the 3 main thematic needs according to the "direction of flow" from the field of informatics to the field of precision oncology (Figure 3). These themes are hierarchically layered and functionally independent. The theme, *better access to oncology knowledge*, forms the foundational theme, supplying the raw, up-to-date evidence base. Another theme, *clinical contextualization and resource navigation*, builds on that knowledge by translating it into locally relevant, feasible options—evaluating insurance coverage, resource availability, and the patient's socioeconomic context. The third theme, *decision support abilities in increasing efficiency*, leverages both themes to deliver actionable recommendations, visualizations, and workflow-integrated tools that streamline the clinician decision-making process.

XSL•FO
RenderX

Figure 3. Precision oncology decision support–related needs following the "direction of flow" from the informatics domain to the precision oncology domain, highlighting the interdependence among the themes.



Variety of Needs Among Specialties, Professional Titles, and Roles in MDT

We found that clinicians from different specialties exhibit varying preferences in their needs (Table S3 in Multimedia Appendix 1), which are likely influenced by the nature of their departments and the types of patients they encounter. Physicians are more likely to treat patients with advanced or progressive-stage conditions, and therefore, their needs for oncological knowledge are most pronounced, particularly in areas such as CTs and the management of complex cases. In contrast, surgeons tend to prioritize PODS-related knowledge related to the perioperative period, such as surgical risks and postoperative complication assessments.

There are also varying needs based on different professional titles. For senior physicians, such as department heads, their extensive clinical experience allows them to make rapid diagnostic and treatment decisions. Thus, they expect PODS to provide more up-to-date knowledge, particularly regarding cross-disciplinary insights and the latest drug research. In contrast, younger physicians tend to seek more support in oncology pharmacology and case reports.

As for the roles in MDT, preparers (Multimedia Appendix 1) seek PODS to support the data preparation and interface presentation phases of MDT. They have a more extensive need for visualized data and convenient retrieval functionalities.

The Functional Framework of Needs for Integrated PODS Tools

After the themes were extracted and confirmed, a functional framework of needs for integrated PODS tools was proposed. In the framework, the needs for PODS tools were basically divided into three dimensions: (1) to provide clinicians with up-to-date and straightforward oncology knowledge; (2) to enable clinical contextualization and resource navigation, offering suggestions for applicable and feasible clinical management; and (3) to offer comprehensive and efficient support for the decision-making process (Table 4).

XSL•FO
RenderX

Table 4. The functional framework of needs for integrated PODS[a] tools.

Dimension and subcomponents	Functions of needs	Recommendation
Up-to-date and straightforward oncology knowledge		
High-level evidence	• Access to high-level clinical evidence (eg, guidelines, conferences, and consensuses) • Concise overview of evidence, containing a general idea of medical information (eg, drug, therapy, and test results) • Labeling of credibility or describing the clinical significance or referability • Source of literature	High-level evidence (eg, guidelines, conferences, and consensuses) is always a priority for clinicians when making decisions about any patient. Clear labeling of the evidence level is also helpful for clinicians in determining the credibility of the evidence [44].
CTs[b]	• Information on inclusion and exclusion criteria (related to the accessibility of the CT function) • Provide midterm results, if available, for CTs that are ongoing and not recruiting • Provide enrollment information for recruiting CTs • Comprehensive description of medication information as outlined in the CT protocol	After the failure of guideline-recommended therapies, appropriate CTs should be considered.
Biological knowledge in diagnosis, treatment, and prognosis	• Description of basic biological information of biomarkers, such as nomenclature (eg, alias and transcript numbering), type (eg, point mutation, fusion, and copy number variation), structural information (eg, mutation site and location in chromosome), and basic biological function • Description of the medical function of biomarkers related to oncology, such as the role in tumor development (eg, oncogenic or tumor suppressor and driver or passenger), mechanism of involvement in tumor development (eg, kinase activation and DNA damage repair), and frequency of expression and mutation in different cancers • Description of the medical function of biomarkers directly related to patient information, such as clinical significance (eg, diagnosis, treatment, and prognosis) and pathogenicity • Description of inclusion in public evidence repositories of biomarkers • Source of evidence	A better explanation of biological knowledge will help clinicians understand the pathological mechanism. The integration will also improve the interpretation of the clinical significance of genomic testing results [45].
Drug and therapy information	• Information on the use, dosage, and side effects of drugs and therapy options from prior medication and literature	The use, dosage, and side effects information of drugs and therapy options help clinicians implement treatment accurately.
Case reports and preclinical evidence	• Source of the case reports related to the patient • Phase 1 to 2 CTs and animal experiments of drugs	Case reports and preclinical evidence are important references for complex or rare cases.
Multidisciplinary guidelines	• Clinical guidelines for different cancer areas	Although the proportion of patients with multiple primary cancers is small, the number of discussions about these patients is gradually increasing. These guidelines can be used to keep physicians updated on other cancer types outside of MDTs.
Clinical contextualization and resource feasibility support		
Accessibility of CTs	• Automatically match patient information with inclusion and exclusion criteria (refer to the information on the inclusion and exclusion criteria function) • Automatically estimate recruitment status • Provide an automatic geographic accessibility for eligible CTs	For patients, the enrollment criteria of CT are as important as the therapy itself. Letting the clinician know whether a certain CT could be accessed also matters.

XSL•FO
RenderX

Dimension and subcomponents	Functions of needs	Recommendation
Accessibility of drugs and treatment options	• Information on the insurance coverage status of drugs • Cost of drugs and treatment options • Approval information of drugs and treatment options	This function helps clinicians know about drug accessibility and could save time in screening available CTs and communicating with patients.
Application of prediction models	• Convenient interface to access prediction models	There are many prediction models in the oncology domain, such as operative risk assessment and risk of recurrence. However, the current PODS tools have not integrated these models yet. Integrating and applying these models could be helpful in decision-making.
Comprehensive and efficient support abilities in the decision-making process		
Individualized search criteria	• Incorporate biological knowledge (eg, the function of variants or signaling pathways) • Incorporate phenotypic features (eg, stage, number of treatment lines [ie, the treatment course], and the location of the lesion) into search criteria	Optimization in search criteria, such as biological knowledge and phenotypes, could help clinicians search for treatment options more precisely.
Automatic integration of patient information	• Automatically summarize the patient's diagnosis and treatment history • Automatically integrate the patient's health information from different systems in the hospital, such as the HIS[c], PACS[d], and LIS[e].	The automatic summarization of patient information could help reduce the *preparation* time.
Recordability and traceability	• An interface for recording MDTf discussion, preferably automatically	A timely recording of the MDT discussion function could help clinicians trace back, summarize, and analyze previous data easily.
Better visualization	• Chronological patient history • Logical display • Organization of information by priority • Clear hierarchy of information	Automatically generated timelines help clinicians read patients with a long history and complex treatment course straightforwardly, whereas graphs and tables can help them understand data better and interpret clinical significance more easily.
Treatment option recommendation	• Personalized ranking of drugs and treatments based on their advantages and disadvantages • Provide automated treatment recommendations	A treatment recommendation function with evidence classification, data matching, and a sorting algorithm could provide extra comprehensive support to clinicians during decision-making.
Question answering and dialogue	• Automatic dialogue and question answering • Large language model agents that automatically retrieve information or tools to support decision-making	The development of artificial intelligence and large language models may implement the function of multiturn question answering and dialogue, as well as provide precise information.

[a]PODS: precision oncology decision support.

[b]CT: clinical trial.

[c]HIS: hospital information system.

[d]PACS: picture archiving and communication system.

[e]LIS: laboratory information system.

[f]MDT: multidisciplinary team.

Discussion

This qualitative descriptive study aimed to gain a more in-depth understanding of clinician-oriented integrated PODS tools. The data were collected with structured observations and semistructured in-depth interviews.

Principal Findings

Needs in PODS

Three main themes of needs in PODS were extracted: (1) the *oncology knowledge* theme underscores the necessity of a continuously updated, structured knowledge resource that synthesizes genomic, therapeutic, and guideline data into clinically digestible formats; (2) the *clinical contextualization and resource navigation* theme highlights the importance of situating this knowledge within real-world practice to ensure that treatment recommendations are feasible and equitable; and (3) the *support abilities in the decision-making process* theme streamlines clinician decision-making and improve care efficiency.

XSL•FO
RenderX

Prospects for PODS Tools Based on the Proposed Framework

A functional framework for the needs of integrated PODS tools was proposed. Providing timely and updated oncology information was the most frequently mentioned need by clinicians. Currently, most clinicians still obtain the latest cancer knowledge through literature reviews; data released and presented at academic activities and conferences; or information shared by, for example, fellow experts and drug suppliers. An integrative PODS tool that properly meets the clinical requirements for up-to-date knowledge summarization and presentation is not yet available. However, clinicians are not satisfied with the tools that are currently available. The main issues include delayed updates, poor retrieval conditions, and insufficient integration with molecular biological evidence. PODS tools are supposed to incorporate the interpretation of available molecular testing results to meet the requirements of the precision medicine concept and to aid the molecular tumor board [21,22].

The framework suggests that integrated PODS tools related to oncology knowledge should include a timely updated and comprehensive knowledge base [48], incorporating drug approval information from official administrative authorities; clinical guidelines; and consensus from well-recognized organizations, released CTs' results from academic conferences, preclinical evidences from literature reviews, and open CTs' updates through official channels. This knowledge base should be highly structured and user-friendly to meet the advanced retrieval needs of clinicians, focusing on information regarding genetic variant data, molecular functional significance, evidence on therapies and prognosis, multidisciplinary guidelines, and case reports. The genomic profiles of patients with cancer are typically presented in genetic testing reports, which contain information on potential actionable targets with variant classification based on well-recognized guidelines and recommendations. However, clinicians were generally unfamiliar with the variant classification criteria, such as the 4-class systems for somatic mutations [49] and the five-class system for germline variants [44], and their clinical significance. Therefore, using a straightforward and clinically oriented way to display genetic results may help improve the effectiveness of the decision-making process. Furthermore, each clinical evidence for a given genomic variant is often presented separately in genetic testing reports, whereas clinicians often prefer a more integrated and comprehensive summary that synthesizes multiple related pieces of evidence, even when discrepancies exist.

In addition, most clinicians express a need for support with feasibility considerations in the diagnosis and treatment process. Although current drug insurance and CT information query systems are well developed, their integration with patient information is insufficient. Future feasibility-support PODS tools should integrate patient information from local or in-house hospital information system data with medication use, making it easier for clinicians to access relevant data. In addition, diagnostic and prognostic algorithms or models have broad potential applications in clinical practice, supporting clinicians in optimizing the decision-making process in various situations [45,50].

Support in the decision-making process is a less developed yet widely demanded aspect of existing PODS tools. Clinicians are interested in improvements in integrating patient information, optimizing information visualization, and enhancing search and recommendation methods [51] for a more efficient decision-making process. In particular, oncology knowledge–related PODS tools should feature an optimized recommendation function [52] and a matching algorithm based on the level of evidence and CTs to enhance clinicians' confidence in using these tools [53]. With the development of language models, knowledge acquisition through question answering or dialogue is also expected to become a future direction for knowledge-related PODS tools [54]. Currently, the integration of patient information in PODS tools is limited and often requires manual input [55]. Automatic integration of patient information and improved visualization would offer significant assistance in MDT meetings and routine clinical practice [10,56].

When adapting the framework to hospitals of different sizes, resource levels, and IT infrastructures, some previous experiences and frameworks may bring value to the development and deployment of a system in a real environment. The Nonadoption, Abandonment, Scale-Up, Spread, and Sustainability (NASSS) framework for complex health technologies will assist in predicting and evaluating the success of a PODS deployment [57]. The Exploration, Preparation, Implementation, and Sustainment framework will be a guideline in PODS implementation [58].

Comparison With Prior Work

PODS can be understood as a specialized form of a clinical decision support system (CDSS). Our work directly builds upon a substantial body of former research. The model by Hendriks et al [59] identified critical barriers, including guideline maintenance and electronic health record interoperability, aligning with our findings on the necessity of dynamic knowledge updating. However, our work extends this paradigm by systematically addressing 3 dimensions in PODS.

Walsh et al [17] examined existing oncology-focused CDSS platforms, identifying critical dimensions, such as data integration, AI capabilities, and stakeholder concerns, and highlighting how technical sophistication alone is insufficient without addressing clinicians' real-world concerns about data quality, user burden, and trust in AI-driven recommendations. Liao et al [60] documented the rapidly growing intersection of AI with clinical oncology workflows, suggesting that advanced natural language interfaces could improve information retrieval and decision support in tumor boards.

Since Kurnit et al [7] proposed the vision of PODS for harnessing genomic and molecular testing, alongside targeted therapy and immunotherapy insights to enable precision oncology at the point of care in 2018, a variety of PODS tools have emerged, driven by advances in next-generation sequencing, biomarker-driven therapeutics, and data-sharing initiatives [23,24,28,61,62]. However, despite these

XSL•FO
RenderX

technological strides, there is no comprehensive synthesis of what characteristics and functions PODS should offer or a clear mapping of clinicians' practical needs for such systems.

In contrast to prior studies that have predominantly focused on theoretical frameworks or technical prototypes, our study adopts a descriptive and user-centered approach. Through intensive qualitative observations of MDT discussions and in-depth interviews with clinicians of varied specialties and experience levels, we systematically identified and categorized the functional requirements of PODS.

Strengths and Limitations

Our contribution complements existing CDSS and PODS literature by moving beyond prototype descriptions and generic recommendations, instead offering a pragmatic, empirically derived framework for future PODS development that aligns tightly with the lived experiences of oncology practitioners.

However, the needs analysis should be considered in light of certain study limitations. First, the rapid development of precision oncology means that additional needs may arise in the future. Convening a multidisciplinary advisory group from different specialties from time to time to evaluate emerging technologies, therapies, and policy changes as stakeholder review cycles occur and using structured Delphi rounds to reach consensus on necessary revisions may help maintain a living functional needs framework for PODS, which could be applied to integrate perspectives from different departments to validate the prioritization outcomes. Moreover, the MoSCoW (must have, should have, could have, won't have) method is effective in distinguishing critical needs from optional ones during the PODS implementation phase. Second, this study was conducted in only 1 hospital in China, highlighting the need for broader investigations across more regions and clinical units. In addition, a follow-up survey across a larger sample and additional quantitative validations may help validate the identified needs and their relative importance. In practice, institutions can conduct a rapid workshop with local stakeholders to map existing workflows, data systems, and patient populations against the 3 framework dimensions proposed before PODS implementation. Prioritizing the needs when implementing PODS tools is also needed for further study. Third, the observation and interview processes were based on the subjective perspectives of researchers and interviewees, which may introduce some bias. While the study focuses on clinicians, the needs framework may change when considering other stakeholders (eg, hospital administrators and IT developers). A bigger study integrating their perspectives will be needed in the future. Future quantitative studies—such as developing questionnaires focused on areas of likely interest to nonclinical stakeholders, including cost management, data interoperability, IT infrastructure readiness, and organizational feasibility—can help identify key domains and contribute to broader integration, potentially leading to new thematic categories within the framework. Fourth, clinicians may express various needs for specific content, such as CTs, which might be dispersed across these 3 themes. We acknowledge that this is unavoidable. Nonetheless, we anticipate that this classification will provide a new perspective for PODS system developers, offering a way to consolidate the diverse needs of clinicians.

Future Directions

Although the development of PODS tools is rapid, the traditional genetic testing report will likely remain an irreplaceable support for genomic interpretation well into the future. The combination of the genetic testing reports and PODS tools may offer complementary advantages, providing new paradigms in oncology decision-making, which could be explored in further research. Advancements in AI, such as generalist models or large language model agents, may provide improved optimization and interaction capabilities for clinicians and PODS tool developers.

Conclusions

The rapid advancement of precision oncology has created an increasing demand for decision support tools to assist clinicians in navigating complex treatment choices. This study provides valuable insights into the current landscape of PODS tools, identifying and categorizing the needs and challenges that clinicians face in using these tools effectively. Our findings highlight key areas for improvement, including better access to oncological knowledge, clinical contextualization and resource navigation, as well as support abilities in the decision-making process. In addition, we propose a functional framework that outlines the essential features of an integrated PODS tool, which could enhance clinical decision-making in precision oncology. The results offer a foundation for future research and development, emphasizing the need for tools that are not only technologically advanced but also aligned with the practical needs of clinicians in diverse health care settings.

Acknowledgments

This study was conducted with support from the National Key Research and Development Program of China (code: 2022YFC2406804). The authors gratefully acknowledge the participants of this study for their time and insights.

Data Availability

The datasets generated or analyzed during this study are not publicly available due the privacy of patients and physicians in the structural observation and in-depth interviews, but are available from the corresponding author on reasonable request.

XSL•FO
RenderX

Authors' Contributions

SJ and NW are the co-corresponding authors (SJ: email: shuqin_jia@hsc.pku.edu.cn, phone: +8610-88196755; NW: email: nanwu@bjmu.edu.cn, phone: +8610-88196755).
SZ contributed to the conceptualization of the study and the writing of the original draft. SZ and YF contributed to data curation, formal analysis, and investigation of the study. SZ and X Lu contributed to the study methodology. SZ, YF, X Lu, JL, and X Li contributed to the review and editing of the draft. SZ, YF, X Lu, NW, and SJ contributed to project administration. YF, X Lu, DH, NW, and SJ were involved in funding acquisition. YF, X Lu, JL, X Li, MZ, XD, NW, and SJ contributed to the study resources. X Lu, HD, NW, and SJ contributed to the supervision of the study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary materials.
[DOCX File , 154 KB - humanfactors_v12i1e67476_app1.docx]

References

1. Mathieu LN, Larkins E, Akinboro O, Roy P, Amatya AK, Fiero M, et al. FDA approval summary: Capmatinib and Tepotinib for the treatment of metastatic NSCLC harboring MET exon 14 skipping mutations or alterations. Clin Cancer Res 2022 Jan 15;28(2):249-254. [doi: 10.1158/1078-0432.CCR-21-1566] [Medline: 34344795]
2. Wang M, Yang JC, Mitchell PL, Fang J, Camidge DR, Nian W, et al. Sunvozertinib, a selective EGFR inhibitor for previously treated non-small cell lung cancer with EGFR exon 20 insertion mutations. Cancer Discov 2022 Jul 06;12(7):1676-1689 [FREE Full text] [doi: 10.1158/2159-8290.CD-21-1615] [Medline: 35404393]
3. Scott EC, Baines AC, Gong Y, Moore Jr R, Pamuk GE, Saber H, et al. Trends in the approval of cancer therapies by the FDA in the twenty-first century. Nat Rev Drug Discov 2023 Aug 21;22(8):625-640. [doi: 10.1038/s41573-023-00723-4] [Medline: 37344568]
4. Pritchard D, Goodman C, Nadauld LD. Clinical utility of genomic testing in cancer care. JCO Precis Oncol 2022 Jan;6:e2100349 [FREE Full text] [doi: 10.1200/PO.21.00349] [Medline: 35085005]
5. Tamborero D, Dienstmann R, Rachid MH, Boekel J, Baird R, Braña I, Cancer Core Europe consortium, et al. Support systems to guide clinical decision-making in precision oncology: the Cancer Core Europe Molecular Tumor Board Portal. Nat Med 2020 Jul 06;26(7):992-994 [FREE Full text] [doi: 10.1038/s41591-020-0969-2] [Medline: 32632195]
6. Sackett DL, Straus SE. Finding and applying evidence during clinical rounds: the "evidence cart". JAMA 1998 Oct 21;280(15):1336-1338. [doi: 10.1001/jama.280.15.1336] [Medline: 9794314]
7. Kurnit KC, Dumbrava EE, Litzenburger B, Khotskaya YB, Johnson AM, Yap TA, et al. Precision oncology decision support: current approaches and strategies for the future. Clin Cancer Res 2018 Jun 15;24(12):2719-2731 [FREE Full text] [doi: 10.1158/1078-0432.CCR-17-2494] [Medline: 29420224]
8. Kurnit KC, Bailey AM, Zeng J, Johnson AM, Shufean MA, Brusco L, et al. "Personalized cancer therapy": a publicly available precision oncology resource. Cancer Res 2017 Nov 01;77(21):e123-e126 [FREE Full text] [doi: 10.1158/0008-5472.CAN-17-0341] [Medline: 29092956]
9. Patterson SE, Liu R, Statz CM, Durkin D, Lakshminarayana A, Mockus SM. The clinical trial landscape in oncology and connectivity of somatic mutational profiles to targeted therapies. Hum Genomics 2016 Jan 16;10(1):4 [FREE Full text] [doi: 10.1186/s40246-016-0061-7] [Medline: 26772741]
10. Krysiak K, Danos AM, Saliba J, McMichael JF, Coffman A, Kiwala S, et al. CIViCdb 2022: evolution of an open-access cancer variant interpretation knowledgebase. Nucleic Acids Res 2023 Jan 06;51(D1):D1230-D1241 [FREE Full text] [doi: 10.1093/nar/gkac979] [Medline: 36373660]
11. Zeng J, Shufean MA, Khotskaya Y, Yang D, Kahle M, Johnson A, et al. OCTANE: oncology clinical trial annotation engine. JCO Clin Cancer Inform 2019 Jul;3:1-11 [FREE Full text] [doi: 10.1200/CCI.18.00145] [Medline: 31265323]
12. Xu Q, Liu Y, Hu J, Duan X, Song N, Zhou J, et al. OncoPubMiner: a platform for mining oncology publications. Brief Bioinform 2022 Sep 20;23(5):383. [doi: 10.1093/bib/bbac383] [Medline: 36058206]
13. Uppal S, Shrivastava PK, Khan A, Sharma A, Shrivastav AK. Machine learning methods in predicting the risk of malignant transformation of oral potentially malignant disorders: a systematic review. Int J Med Inform 2024 Jun;186:105421. [doi: 10.1016/j.ijmedinf.2024.105421] [Medline: 38552265]
14. Tsimberidou AM, Kahle M, Vo HH, Baysal MA, Johnson A, Meric-Bernstam F. Molecular tumour boards - current and future considerations for precision oncology. Nat Rev Clin Oncol 2023 Dec 16;20(12):843-863. [doi: 10.1038/s41571-023-00824-4] [Medline: 37845306]
15. Pawloski PA, Brooks GA, Nielsen ME, Olson-Bullis BA. A systematic review of clinical decision support systems for clinical oncology practice. J Natl Compr Canc Netw 2019 Apr 01;17(4):331-338 [FREE Full text] [doi: 10.6004/jnccn.2018.7104] [Medline: 30959468]

XSL•FO
RenderX

16. Yung A, Kay J, Beale P, Gibson KA, Shaw T. Computer-based decision tools for shared therapeutic decision-making in oncology: systematic review. JMIR Cancer 2021 Oct 26;7(4):e31616 [FREE Full text] [doi: 10.2196/31616] [Medline: 34544680]
17. Walsh S, de Jong EE, van Timmeren JE, Ibrahim A, Compter I, Peerlings J, et al. Decision support systems in oncology. JCO Clin Cancer Inform 2019 Feb;3:1-9 [FREE Full text] [doi: 10.1200/CCI.18.00001] [Medline: 30730766]
18. Bosserman LD, Mambetsariev I, Ladbury C, Barzi A, Johnson D, Morse D, et al. Pyramidal decision support framework leverages subspecialty expertise across enterprise to achieve superior cancer outcomes and personalized, precision care plans. J Clin Med 2022 Nov 14;11(22):6738 [FREE Full text] [doi: 10.3390/jcm11226738] [Medline: 36431215]
19. Meric-Bernstam F, Johnson A, Holla V, Bailey A, Brusco L, Chen K, et al. A decision support framework for genomically informed investigational cancer therapy. J Natl Cancer Inst 2015 Jul;107(7):e98 [FREE Full text] [doi: 10.1093/jnci/djv098] [Medline: 25863335]
20. Hong N, Sun G, Zuo X, Chen M, Liu L, Wang J, et al. Application of informatics in cancer research and clinical practice: opportunities and challenges. Cancer Innov 2022 Jun 15;1(1):80-91 [FREE Full text] [doi: 10.1002/cai2.9] [Medline: 38089452]
21. Dias-Santagata D, Heist RS, Bard AZ, da Silva AF, Dagogo-Jack I, Nardi V, et al. Implementation and clinical adoption of precision oncology workflows across a healthcare network. Oncologist 2022 Nov 03;27(11):930-939 [FREE Full text] [doi: 10.1093/oncolo/oyac134] [Medline: 35852437]
22. Tamborero D, Dienstmann R, Rachid MH, Boekel J, Lopez-Fernandez A, Jonsson M, Cancer Core Europe consortium, et al. Author correction: the molecular tumor board portal supports clinical decisions and automated reporting for precision oncology. Nat Cancer 2022 May 21;3(5):649 [FREE Full text] [doi: 10.1038/s43018-022-00378-x] [Medline: 35449310]
23. Chakravarty D, Gao J, Phillips S, Kundra R, Zhang H, Wang J, et al. OncoKB: a precision oncology knowledge base. JCO Precis Oncol 2017 Jul;2017:11 [FREE Full text] [doi: 10.1200/PO.17.00011] [Medline: 28890946]
24. Sondka Z, Dhir NB, Carvalho-Silva D, Jupe S, Madhumita, McLaren K, et al. COSMIC: a curated database of somatic variants and clinical data for cancer. Nucleic Acids Res 2024 Jan 05;52(D1):D1210-D1217 [FREE Full text] [doi: 10.1093/nar/gkad986] [Medline: 38183204]
25. Landrum MJ, Chitipiralla S, Brown GR, Chen C, Gu B, Hart J, et al. ClinVar: improvements to accessing data. Nucleic Acids Res 2020 Jan 08;48(D1):D835-D844 [FREE Full text] [doi: 10.1093/nar/gkz972] [Medline: 31777943]
26. de Andrade KC, Lee EE, Tookmanian EM, Kesserwan CA, Manfredi JJ, Hatton JN, et al. The TP53 Database: transition from the International Agency for Research on Cancer to the US National Cancer Institute. Cell Death Differ 2022 May 29;29(5):1071-1073 [FREE Full text] [doi: 10.1038/s41418-022-00976-3] [Medline: 35352025]
27. Cline MS, Liao RG, Parsons MT, Paten B, Alquaddoomi F, Antoniou A, BRCA Challenge Authors, et al. BRCA challenge: BRCA exchange as a global resource for variants in BRCA1 and BRCA2. PLoS Genet 2018 Dec 26;14(12):e1007752 [FREE Full text] [doi: 10.1371/journal.pgen.1007752] [Medline: 30586411]
28. Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, et al. The GeneCards suite: from gene data mining to disease genome sequence analyses. Curr Protoc Bioinformatics 2016 Jun 20;54:1.30.1-1.3033. [doi: 10.1002/cpbi.5] [Medline: 27322403]
29. O'Regan P, Hoskins R, Grave C, Stevenson J, Frost H, Graham DM, et al. Digital ECMT cancer trial matching tool: an open source research application to support oncologists in the identification of precision medicine clinical trials. JCO Clin Cancer Inform 2023 Jan;7:e2200137. [doi: 10.1200/CCI.22.00137] [Medline: 36657092]
30. Xu Q, Liu Y, Sun D, Huang X, Li F, Zhai J, et al. OncoCTMiner: streamlining precision oncology trial matching via molecular profile analysis. Database (Oxford) 2023 Nov 04;2023:baad077 [FREE Full text] [doi: 10.1093/database/baad077] [Medline: 37935585]
31. Li P, Chen T, Yu J, Shih S, Su C, Lin Y, et al. pubmedKB: an interactive web server for exploring biomedical entity relations in the biomedical literature. Nucleic Acids Res 2022 Jul 05;50(W1):W616-W622 [FREE Full text] [doi: 10.1093/nar/gkac310] [Medline: 35536289]
32. Wei C, Allot A, Lai P, Leaman R, Tian S, Luo L, et al. PubTator 3.0: an AI-powered literature resource for unlocking biomedical knowledge. Nucleic Acids Res 2024 Jul 05;52(W1):W540-W546 [FREE Full text] [doi: 10.1093/nar/gkae235] [Medline: 38572754]
33. Khairat S, Marc D, Crosby W, Al Sanousi A. Reasons for physicians not adopting clinical decision support systems: critical analysis. JMIR Med Inform 2018 Apr 18;6(2):e24 [FREE Full text] [doi: 10.2196/medinform.8912] [Medline: 29669706]
34. Laka M, Carter D, Milazzo A, Merlin T. Challenges and opportunities in implementing clinical decision support systems (CDSS) at scale: interviews with Australian policymakers. Health Policy Technol 2022 Sep;11(3):100652. [doi: 10.1016/j.hlpt.2022.100652]
35. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000 Aug;23(4):334-340. [doi: 10.1002/1098-240x(200008)23:4<334::aid-nur9>3.0.co;2-g] [Medline: 10940958]
36. Glaser BG, Strauss AL. Discovery of Grounded Theory: Strategies for Qualitative Research. New York, NY: Routledge; 1999.

XSL•FO
RenderX

37. Saini KS, Taylor C, Ramirez AJ, Palmieri C, Gunnarsson U, Schmoll HJ, et al. Role of the multidisciplinary team in breast cancer management: results from a large international survey involving 39 countries. Ann Oncol 2012 Apr;23(4):853-859 [FREE Full text] [doi: 10.1093/annonc/mdr352] [Medline: 21821551]
38. Prades J, Remue E, van Hoof E, Borras JM. Is it worth reorganising cancer services on the basis of multidisciplinary teams (MDTs)? A systematic review of the objectives and organisation of MDTs and their impact on patient outcomes. Health Policy 2015 Apr;119(4):464-474 [FREE Full text] [doi: 10.1016/j.healthpol.2014.09.006] [Medline: 25271171]
39. Zeng Y, Zhu S, Wang Z, Chen J, Dai J, Liu Z, et al. Multidisciplinary Team (MDT) discussion improves overall survival outcomes for metastatic renal cell carcinoma patients. J Multidiscip Healthc 2023;16:503-513 [FREE Full text] [doi: 10.2147/JMDH.S393457] [Medline: 36865786]
40. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
41. Burchett N. Qualitative Inquiry and Research Design: Choosing Among Five Approaches. 3rd edition. Thousand Oaks, CA: Sage Publication; 2014.
42. Thompson SK. Adaptive cluster sampling. J Am Stat Assoc 1990 Dec 21;85(412):1050-1057. [doi: 10.2307/2289601]
43. Lai L, Meng W, Wei J, Zhang X, Tan Z, Lu Y, et al. Transformation of NSCLC to SCLC after 1st- and 3rd-generation EGFR-TKI resistance and response to EP regimen and erlotinib: 2 CARE-compliant case reports. Medicine (Baltimore) 2021 Mar 12;100(10):e25046 [FREE Full text] [doi: 10.1097/MD.0000000000025046] [Medline: 33725888]
44. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015 May 5;17(5):405-424 [FREE Full text] [doi: 10.1038/gim.2015.30] [Medline: 25741868]
45. Yackel HD, Halpenny B, Abrahm JL, Ligibel J, Enzinger A, Lobach DF, et al. A qualitative analysis of algorithm-based decision support usability testing for symptom management across the trajectory of cancer care: one size does not fit all. BMC Med Inform Decis Mak 2024 Mar 05;24(1):63 [FREE Full text] [doi: 10.1186/s12911-024-02466-7] [Medline: 38443870]
46. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018 Mar;52(4):1893-1907 [FREE Full text] [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
47. Braun V, Clarke V. Is thematic analysis used well in health psychology? A critical review of published research, with recommendations for quality practice and reporting. Health Psychol Rev 2023 Dec 19;17(4):695-718. [doi: 10.1080/17437199.2022.2161594] [Medline: 36656762]
48. Wagner AH, Walsh B, Mayfield G, Tamborero D, Sonkin D, Krysiak K, Variant Interpretation for Cancer Consortium. A harmonized meta-knowledgebase of clinical interpretations of somatic genomic variants in cancer. Nat Genet 2020 Apr 03;52(4):448-457 [FREE Full text] [doi: 10.1038/s41588-020-0603-8] [Medline: 32246132]
49. Li MM, Datto M, Duncavage EJ, Kulkarni S, Lindeman NI, Roy S, et al. Standards and guidelines for the interpretation and reporting of sequence variants in cancer: a joint consensus recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists. J Mol Diagn 2017 Jan;19(1):4-23 [FREE Full text] [doi: 10.1016/j.jmoldx.2016.10.002] [Medline: 27993330]
50. Gong Y, Pang H, Yu Z, Wang X, Li P, Zhang Q. Construction of inflammatory associated risk gene prognostic model of NSCLC and its correlation with chemotherapy sensitivity. Ann Med 2023 Dec 21;55(1):2200034 [FREE Full text] [doi: 10.1080/07853890.2023.2200034] [Medline: 37083272]
51. Keller RB, Mazor T, Sholl L, Aguirre AJ, Singh H, Sethi N, et al. Programmatic precision oncology decision support for patients with gastrointestinal cancer. JCO Precis Oncol 2023 Jan;7:e2200342 [FREE Full text] [doi: 10.1200/PO.22.00342] [Medline: 36634297]
52. Andre F, Mardis E, Salm M, Soria JC, Siu LL, Swanton C. Prioritizing targets for precision cancer medicine. Ann Oncol 2014 Dec;25(12):2295-2303 [FREE Full text] [doi: 10.1093/annonc/mdu478] [Medline: 25344359]
53. Li MJ, Yao H, Huang D, Liu H, Liu Z, Xu H, et al. mTCTScan: a comprehensive platform for annotation and prioritization of mutations affecting drug sensitivity in cancers. Nucleic Acids Res 2017 Jul 03;45(W1):W215-W221 [FREE Full text] [doi: 10.1093/nar/gkx400] [Medline: 28482068]
54. Kresevic S, Giuffrè M, Ajcevic M, Accardo A, Crocè LS, Shung DL. Optimization of hepatological clinical guidelines interpretation by large language models: a retrieval augmented generation-based framework. NPJ Digit Med 2024 Apr 23;7(1):102 [FREE Full text] [doi: 10.1038/s41746-024-01091-y] [Medline: 38654102]
55. Khumalo AC, Kane BT. Perspectives on record-keeping practices in MDT meetings and meeting record utility. Int J Med Inform 2022 May;161:104711 [FREE Full text] [doi: 10.1016/j.ijmedinf.2022.104711] [Medline: 35189482]
56. Ansoborlo M, Gaborit C, Grammatico-Guillon L, Cuggia M, Bouzille G. Prescreening in oncology trials using medical records. Natural language processing applied on lung cancer multidisciplinary team meeting reports. Health Informatics J 2023 Mar 24;29(1):14604582221146709 [FREE Full text] [doi: 10.1177/14604582221146709] [Medline: 36964666]

XSL•FO
RenderX

57. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
58. Aarons GA, Hurlburt M, Horwitz SM. Advancing a conceptual model of evidence-based practice implementation in public service sectors. Adm Policy Ment Health 2011 Jan;38(1):4-23 [FREE Full text] [doi: 10.1007/s10488-010-0327-7] [Medline: 21197565]
59. Hendriks MP, Jager A, Ebben KC, van Til JA, Siesling S. Clinical decision support systems for multidisciplinary team decision-making in patients with solid cancer: composition of an implementation model based on a scoping review. Crit Rev Oncol Hematol 2024 Mar;195:104267 [FREE Full text] [doi: 10.1016/j.critrevonc.2024.104267] [Medline: 38311011]
60. Liao J, Li X, Gan Y, Han S, Rong P, Wang W, et al. Artificial intelligence assists precision medicine in cancer treatment. Front Oncol 2022 Jan 4;12:998222 [FREE Full text] [doi: 10.3389/fonc.2022.998222] [Medline: 36686757]
61. Hammer RD, Fowler D, Sheets LR, Siadimas A, Guo C, Prime MS. A digital tumor board solution impacts case discussion time and postponement of cases in tumor boards. Health Technol 2021 Feb 27;11(3):525-533. [doi: 10.1007/s12553-021-00533-x]
62. Krupinski EA, Comas M, Gallego LG, GISMAR Group. A new software platform to improve multidisciplinary tumor board workflows and user satisfaction: a pilot study. J Pathol Inform 2018;9:26 [FREE Full text] [doi: 10.4103/jpi.jpi_16_18] [Medline: 30167341]

Abbreviations

AI: artificial intelligence
CDSS: clinical decision support system
COREQ: Consolidated Criteria for Reporting Qualitative Research
CT: clinical trial
MDC: Molecular Diagnostic Center
MDT: multidisciplinary team
MoSCoW: must have, should have, could have, won’t have
NASSS: Nonadoption, Abandonment, Scale-Up, Spread, and Sustainability
PODS: precision oncology decision support
PUCH: Peking University Cancer Hospital

Edited by A Kushniruk; submitted 14.10.24; peer-reviewed by A Huertas-Zurriaga, Y Chu; comments to author 13.12.24; revised version received 03.01.25; accepted 26.06.25; published 08.09.25.

Please cite as:
Zheng S, Feng Y, Long J, Li X, Zhang M, Chen H, Du X, Duan H, Jia S, Wu N, Lu X
Need Analysis of Clinician-Oriented Integrated Precision Oncology Decision Support Tools: Qualitative Descriptive Study
JMIR Hum Factors 2025;12:e67476
URL: https://humanfactors.jmir.org/2025/1/e67476
doi:10.2196/67476
PMID:

©Sheng Zheng, Yi Feng, Jieran Long, Xiang Li, Meng Zhang, Hui Chen, Xue Du, Huilong Duan, Shuqin Jia, Nan Wu, Xudong Lu. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 08.09.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Personal Health Record Software for Neuroendocrine Tumors: Patient-Centered Design Approach

Juan Pablo Hourcade[1*], PhD; Michael O'Rorke[2,3*], PhD; Elizabeth Chrischilles[2,3], PhD; Nicholas J Rudzianski[2], BS; Brian Gryzlak[2], MSW, MA; Kerry Peterman[1], BA; Chris Ortman[4], BS; Sam Wolstencroft[4], BA; Margaret Bean[5], MEd; Elyse Gellerman[6], MHS; Josh A Mailman[7], MBA; Maryann Wahmann[8]

[1]Department of Computer Science, University of Iowa, 14 MacLean Hall, Iowa City, IA, United States

[2]College of Public Health, University of Iowa, Iowa City, IA, United States

[3]Holden Comprehensive Cancer Center, University of Iowa, Iowa City, IA, United States

[4]Institute for Clinical and Translational Science, University of Iowa, Iowa City, IA, United States

[5]Healing NET, Nashville, TN, United States

[6]Neuroendocrine Tumor Research Foundation, Boston, MA, United States

[7]NorCal CarciNET Community, San Jose, CA, United States

[8]Neuroendocrine Cancer Awareness Network, Fort Mill, SC, United States

[*]these authors contributed equally

Corresponding Author:
Juan Pablo Hourcade, PhD
Department of Computer Science, University of Iowa, 14 MacLean Hall, Iowa City, IA, United States

Abstract

Background: Personal health record (PHR) software has the potential of aiding with patient engagement and data collection in longitudinal research to better understand the long-term impact of treatments on patients with rare medical conditions. Neuroendocrine tumors (NETs) represent a rare condition with unique challenges related to symptom management, treatment tracking, and patient-provider communication.

Objective: This study aimed to design, develop, and evaluate PHR software tailored for patients with NETs as part of a longitudinal research study. Our goal was to create a patient-centered PHR that supports both self-management and research data collection.

Methods: This included activities spanning the entire development lifecycle from identifying user requirements through focus groups and surveys, iterative prototype refinement via cognitive walkthroughs, and usability testing of the functional PHR system. Feedback from patient advocacy organizations and clinical experts further informed PHR development.

Results: The resulting PHR allows patients with NETs to access condition-specific information, track symptoms, monitor treatment regimens, and share data with health care providers. Patients valued the ability to visualize personal health trends and patterns over time, enhancing both self-management and communication with medical teams. Usability testing indicated high levels of patient satisfaction with the system's functionality and design.

Conclusions: The development of this PHR demonstrates the value of engaging patients in the design process to ensure that health technologies address real-world needs. Our approach provides a model for designing PHR systems for other rare conditions, highlighting the importance of patient-centered design in supporting both clinical care and longitudinal research.

(JMIR Hum Factors 2025;12:e68788) doi:10.2196/68788

KEYWORDS

personal health records; user-centered design; neuroendocrine tumors; rare disease; longitudinal research

Introduction

The past 20 years have seen a revolution in the digitization of health records [1-3]. There are 2 aspects to this digitization that have been occurring in parallel. One is the advent of electronic health records in hospitals and doctor's offices [1,3], which are often accompanied by health portals providing patients with access to their health records [4,5]. The data in these systems almost exclusively originate from hospitals, laboratories, and doctor's offices. The other is patients using technology to capture or record personal health data. This latter aspect has involved automated capture of information (eg, steps taken from smartwatches or smartphones) [6], the use of smart health devices (eg, blood pressure monitors that connect with

XSL•FO
RenderX

smartphones or computers) [7], and the use of personal health record (PHR) software that enables patients to record information that may otherwise not be tracked (eg, symptoms) [8,9].

PHR software has also been proposed as a tool to facilitate medical research [10]. Many medical studies require long-term follow-up of patients with self-reported measures. The data collection necessary for such studies has traditionally involved mailing paper questionnaires to patients using the postal system or calling them on the phone. Using PHR software for the same purpose has the potential of lowering costs by avoiding mailings, transcriptions, or time spent on phone calls, while increasing data quality by avoiding errors in transcription and enabling patients to contribute data at a time and location and on a device that is most convenient for them [11,12]. In order for PHR software to successfully play a data gathering role for research studies it needs to provide a valuable service to patients, otherwise patients are less likely to use it [9,13]. Ideally, PHR software used for research studies would follow a model similar to that followed by free online services: users get a valuable service in exchange for their personal data. The difference in this case is that patients would explicitly consent to provide their data for a specific purpose ie, medical research.

One specific area of medical research that we believe could benefit from the use of PHR software for gathering data from patients is research on rare medical conditions. Knowledge gaps remain prevalent for most rare medical conditions due to systemic and structural challenges in health care and research. These include small, geographically dispersed patient populations, limited funding for research, and frequent diagnostic delays. The latter arise from the complexity of these diseases, limited practitioner training when compared to more common conditions [14], and the evolving understanding of disease characteristics [15]. There is also often a need for longitudinal research to better understand the long-term impact of treatments on patients with rare medical conditions [16] and PHR software could play a role in repeated data collection for such research. The challenge is that because these conditions are rare, it is unlikely that there are commercial PHR systems that are a good match for patient needs, unlike the case for more common conditions, such as asthma, diabetes, glaucoma, HIV, hyperlipidemia, and hypertension [14,17]. The sole example we found of a PHR designed for patients with a rare condition (idiopathic thrombocytopenic purpura) resulted from an adaptation of a generic PHR system based on advice from expert physicians [18]. We recommend, in addition to consulting with experts, engaging with patients in user-centered design activities, positioning patients as research partners.

In this article, we present our experience engaging with patients with neuroendocrine tumors (NETs) [19,20], a rare medical condition, throughout the design process of PHR software currently being used as part of a prospective cohort study of people with NETs [19-21]. NETs include a broad variety of tumors that can appear throughout the body, but are most common in the small intestine, pancreas, and lung [19,20].

Methods

Participant Recruitment

For all phases of the project, we recruited patients with NETs through 4 partnering NET advocacy groups in the United States (The Neuroendocrine Tumor Research Foundation, The Neuroendocrine Cancer Awareness Network, The Healing NET Foundation and Northern California CarciNET Community). Inclusion and exclusion criteria for volunteers in this study included adults (≥18 years old, any gender) with a confirmed diagnosis of any NET, able to read, speak, and understand English to participate in focus groups, surveys, and usability testing, a willingness to share personal experiences regarding symptom tracking, treatment management, and interactions with health care providers related to their NET, and access to a computer, tablet, or smartphone with internet connectivity to interact with the PHR system and participate in online activities. There were no other demographic or other requirements for volunteer participants including previous PHR experience. Overall, 89 patients with NETs participated in at least 1 activity (60 female, 18 male, 11 unknown gender; 75 White or Caucasian, 1 Black or African American, 1 Asian or Asian American, 12 unknown race; 2 Hispanic or Latino, 80 not Hispanic or Latino, 7 unknown; 32 with postgraduate degrees, 27 with college degrees, 16 with some college, 2 with a high school degree or GED, 1 without a high school degree or GED, 11 unknown educational level; 7 aged 18 - 40, 30 aged 41 - 60, 49 aged 61 - 80, and 3 aged >80 years). Participants who provided their home address (n=75) joined activities from 34 different states and one location outside of the United States.

Outline of Research Activities

Research activities involved a comprehensive, multiphase process designed to engage patients with NETs at every stage of the PHR development including: (1) 9 online design sessions with patients with NETs to develop requirements and a survey to validate these requirements with a broader set of patients, (2) another 9 online sessions to iteratively improve a low-fidelity interactive prototype, and (3) three sets of online usability testing sessions to refine the implemented PHR software (refer to Table 1 for details). The patient activities described in this report were reviewed and approved by the University of Iowa Institutional Review Board.

XSL•FO
RenderX

Table . Session activities organized by phase, strategy, format, and number of participants.

Strategy	Format	Participants, n	Activities
Phase 1: identifying requirements			
Brainstorming	Online focus groups (3 groups of people, first of 3 sessions)	16	• Imagine a superhero that can help with their condition – what would the superhero do? (Superhero construct frees patients from perceived constraints) • Research team compiles a list of potential features from participant discussions
Organizing	Online focus groups (same 3 groups of people, second session)	16	• Participants rank desirable or undesirable features • Prioritize features and provide more details on what they want • Research team reviews results and identifies what categories to implement in the NET[a] PHR[b] in two areas: (1) providing information on NETs; (2) tracking condition-related data
Organizing	Survey of focus group attendees	16	• Further prioritization of features and details on requirements
Confirmation	Online focus groups (same 3 groups of people, third session)	16	• Discuss survey results with each group to better understand the reasons behind survey answers • Probe what patients would expect from a PHR that implemented these requirements
Ranking and attitudes	Validation survey	52	• Prioritize categories of information or advice (high, intermediate, and low) • Preference of categories to track
Phase 2: prototype refinement			
Rapid prototyping	Cognitive walkthrough with users	27 (3 groups of 9)	• Rapid prototyping with wireframe software (Balsamiq) by research team • Patients provide feedback on prototype based on specific tasks • Iteratively improve between group activities until optimal design reached (3 rounds with each group of 9 patients)

XSL•FO
RenderX

Strategy	Format	Participants, n	Activities
Phase 3: usability testing			
Functional prototype	Usability testing (tasks, survey)	10	• Use functional prototype on Zoom with team • Patients complete prepared tasks using a test version of the tool, thinking aloud • Provide feedback via survey—NASA task load index administered after each task • Discussion at end on what worked well, what did not, and what they would change • Iteratively improve based on feedback (3 rounds of usability testing)

[a]NET: neuroendocrine tumors.

[b]PHR: personal health record.

Ethical Considerations

We obtained ethics approval from the University of Iowa's Institutional Review Board (202110283). All participants received an information and consent letter before enrollment and participation. Participants were compensated US $50 for Phase 1 and Phase 2 Zoom sessions, US $10 for the Phase 1 survey, and US $20 for Phase 3 usability Zoom sessions. To preserve participants' privacy and confidentiality, we took notes rather than record sessions and these notes are not associated with specific participants. In addition, the Phase 1 surveys were designed to ensure anonymity of participants by not collecting personally identifiable information.

Phase 1: Identifying Requirements (Online Focus Groups, 9 Sessions)

To identify requirements, we recruited 16 patients with NETs, divided them into 3 groups, and conducted 3 1-hour online sessions with each group using Zoom, for a total of 9 sessions.

For the first session, our goal was to obtain ideas for PHR software for patients with NETs without any constraints. After introducing the project and our overall goal, our instructions to patients during the first session were to imagine that there was a superhero that could help them with their medical condition, who could do anything except give them a cure. The idea behind the superhero discussion was to remove any constraints patients could think of with respect to what would be possible with technologies. When we met with subsequent groups, we started by getting ideas from them and then introducing ideas by previous groups if they had not mentioned them to obtain further feedback. We had previously used the superhero activity, albeit face-to-face, while developing a PHR focused on elder medication safety [13] and others later used it face-to-face with older adults as well [22].

During each superhero session, at least 3 researchers took notes on participants' ideas. To analyze these notes, we used an inductive coding process, similar to the one proposed by Gioia et al [23]. First, 2 of the researchers (JPH and KP) went through the notes and created 1 physical sticky note per unique concept, for a total of 51. Using these sticky notes, they then organized patient ideas, developing an affinity diagram, from which they produced a hierarchically organized list of requirements. The high-level themes in this hierarchy of requirements were information on NETs (eg, support groups, advocacy, education, and ideas for relief), tracking condition-related personal information, and managing logistics. These requirements were the basis for the second session with each group. We used this second session to validate the set of requirements, begin to prioritize them, and obtain more details on the requirements. We had pursued a similar approach in previous research, in particular with respect to information tracking, to understand the difference between information patients would like to track and information they would be willing to enter [13].

The research team met after the second set of sessions to consider the priorities identified by patients and the practical realities of what we could accomplish within the constraints of the project. One important area in which patients were interested that we were unable to pursue was managing logistics, from better integration of electronic health records across providers to managing prescriptions and appointments. Based on these considerations, we identified a set of requirements on which to focus. We then developed a web-based Qualtrics survey sent to the 16 participating patients with NETs to obtain more details on requirements and further prioritize them. We used the third set of sessions to discuss the results of the survey with each of the groups to better understand the reasons behind survey answers and get a sense for what patients would expect from PHR software that implemented these requirements.

XSL•FO
RenderX

After the third set of sessions, we created an electronic survey to be distributed to a broader set of patients with NETs to validate the requirements we developed with 16 patients. We believe it is important to validate requirements from participatory sessions with a larger group of patients due to potential biases related to who can be available and feel comfortable participating in such activities. Others have followed similar approaches to balance the depth of discussion in participatory sessions with the greater representativeness afforded by a survey [24]. Questions focused on preferred data tracking features, information needs, and feedback or output options (eg, personal trends and data visualizations). The 52 patients with NETs who answered the survey helped us validate and further prioritize the set of requirements we had developed. The result of Phase 1 was a list of specifications that would guide prototype development.

Phase 2: Prototype Refinement (Cognitive Walkthroughs, 3 Rounds, and 27 Participants)

Our next step was to design an interactive wireframe prototype using Balsamiq [25]. Balsamiq enables screen designs that look as if they were sketched with pen and paper and includes commonly used user interface elements (eg, buttons and images). The sketches are created so feedback on them can focus on those elements and navigation through the system, as opposed to aesthetic aspects. Screen sketches can be linked to each other, so the prototype is navigated by clicking on user interface elements (eg, menus and links) like a normal website.

We then sought to iteratively improve the prototype through 3 rounds of feedback from patients with NETs over Zoom using cognitive walkthrough with users techniques [26], which have previously been applied to wireframe prototypes [27]. We obtained feedback from 3 different groups of 9 patients with NETs, for a total of 27 subjects. Patient participation varied per round. For the first round, we had 24 patients participate, 20 for the second, and 17 for the third. During each session, a research team member would work with 2 - 3 patients at a time (ie, small group sessions), asking them how they would navigate the system for a given set of tasks. Participants verbalized their thought process (“think-aloud” protocol), highlighting usability issues and areas of confusion. Feedback from each round informed changes to the prototype (ie, iterative refinement), which were then reevaluated in subsequent rounds.

Phase 3: Usability Testing (3 Rounds, 10 Participants)

The following step was to develop a working web application based on the design of the Balsamiq prototype. This design incorporated more detailed functionality that could not be tested in Balsamiq and aesthetic aspects, such as fonts and colors. To implement the design, we used an existing system we previously developed, the Iowa PHR, and customized it based on the outcomes of Phases 1 and 2. The Iowa PHR is a Java 8 [28] Spring application [29] with a Microsoft SQL Server database [30]. It is hosted with Jetty [31] and fronted with Apache [32].

To iteratively improve this working system, we conducted 3 rounds of usability testing [33] with a total of 10 patients with NETs, 4 participated in the first session, 3 in the second, and 3 in the third. There were about 3 weeks between the first and second round of sessions, and about 6 months between the second and third rounds (in part due to competing commitments for the software developers), enabling our developers to make changes to the system.

At each session we asked participants to complete tasks involving the system’s most important features, including finding specific types of information about NETs and tracking a variety of data. After a group of tasks pertaining to similar features, we asked patients to fill out a NASA Task Load Index survey [34] to assess the ease or difficulty of those tasks. At the end of each session, patients provided qualitative feedback on what worked well, what did not work well, and what they would like to change about the system. From this feedback, after each round of usability sessions, we developed a list of changes for the PHR, which the development team completed before the following round of usability sessions.

Our patient advocacy organizations provided “test-drive” feedback through credentialed access to the PHR, followed by structured surveys (via Research Electronic Data Capture [REDCap]) [35] to capture their insights. In a similar way, useful feedback was also obtained from our NET specialists (ie, working oncologists and endocrinologists) who reviewed the PHR to suggest improvements related to medical terminology, treatments, information accuracy, and ordering or presentation of data.

After we incorporated all feedback into the PHR, we made it available to patients with NETs enrolled in an ongoing research study in March 2024. Subjects enrolled at the time were sent an email notification introducing the availability of those features. Further outreach is planned to promote awareness and gather more comprehensive feedback on long-term utility.

Results

Phase 1: Identifying Requirements

Identified Requirements

The requirements we identified fell into 2 categories: information on NETs and tracking condition-related personal information. The types of information in which patients were interested were related to the fact that NETs are rare, there are different subtypes of NETs, and research is still ongoing regarding optimum therapeutic management and treatment sequencing [19,20].

In terms of tracking personal information, most patients had a strong interest in tracking symptoms because some (eg, diarrhea) have a significant impact on quality of life [20]. Together with symptoms, they expressed interest in tracking factors possibly triggering symptoms including consumption of certain foods, medications, episodes of stress, and quality of sleep. They sought to share this information with their medical team but also visualize patterns and potential interactions over time.

Validated Requirements

In total, 52 patients with NETs completed the requirements validation survey. The survey focused on the 2 key aspects that emerged from our participatory sessions: providing information

XSL•FO
RenderX

on NETs and tracking condition-related data. Through the survey, we were able to both validate and prioritize the requirements obtained during Zoom sessions. Multimedia Appendix 1 shows preferred types of data to track and Table 2 lists preferred features related to tracking data. Multimedia Appendix 2 shows preferred types of information, education, or advice about NETs.

Table . How useful are certain feedback and output options for patients with neuroendocrine tumors.

Feedback or output option	Most useful, n	Least useful, n
Personal trends over time	39	1
Patterns	29	3
Comparisons of data from other users	13	19
Ability for other users to enter data	14	11
Simple graphs or charts	26	5
Two or multi axis graphs or charts between items	11	14
Printable data, graphs, or charts	25	6
Download raw data	11	21
Detailed reports and summaries	25	6
Allow users to select type of output	26	3
Snapshots or focused summaries	17	8
Option to save editable preferences	15	12
Flexibility in timeframe of outputs	12	13
Generic notifications or nudges	13	18

Phase 2: Prototype Refinement

The most difficult challenges we addressed through these sessions involved information architecture for the broad range of information patients thought should be provided about NETs. Due to the wealth of information available, challenges included decisions on the most reliable or updated data sources, and how to organize and categorize the information. In addition, we also obtained feedback on basic tracking capabilities during the first prototyping round, and on custom tracking features (enabling users to add their own custom tracking categories) during the second and third rounds.

Iteratively improving these interactive wireframe prototypes in Balsamiq enabled us to improve the design of the system at a low cost, since it is quick and easy to modify Balsamiq prototypes when compared to modifying a web application.

Phase 3: Usability Testing

Patients were provided with a set of 10 predetermined tasks, presented in an order of increasing complexity, concerning information seeking, data entry, and editing and deleting entries. At the end of each set of related tasks, patients were instructed to complete the NASA Task Load Index survey [34] which evaluated patients' responses across 6 dimensions including mental, physical and temporal demand as well as ease of performance, effort involved and frustration with the set tasks. Overall, the findings indicated a generally manageable task load, with the majority of patients scoring mental, physical and temporal demands across tasks negatively (ie, low demand). As expected, higher order tasks were associated with an increase in demands of effort, but most patients reported minimal frustration with the assigned tasks.

Together with the feedback obtained from our patient advocacy organization partners and NET clinician groups, usability testing feedback helped us identify minor bugs, make some information easier to find, make it easier for users to start tracking data, simplify some tracking options, and otherwise confirm that the system was ready for use.

The final version of the PHR app includes options for tracking data related to the condition. We identified 18 categories of data from our interactions with stakeholders (eg, skin flushing, food intake, pain, and medication intake), which are ready to be used with a design that aims to minimize the steps required to enter data. In addition, users can create their own custom category to track, which addresses the broad range of needs within NETs. An instructional video is available to patients explaining the PHR app's tracking capabilities and how to use them. The system also enables users to track the medications they are currently taking or have previously taken, with a user interface developed during previous research [10,13]. In addition, the PHR app also provides access to various types of information relevant to patients with NETs organized into tabs: about NETs, treatments for NETs, tips for managing care (many of these came from stakeholders and patients), and NET resources (eg, patient support groups and clinical trials).

A video walkthrough of the final PHR is available in the Multimedia Appendix 3.

Utility Assessment

Our PHR can now be accessed by patients with NETs enrolled in a research study where the emphasis is on completing study-related surveys. Patients enrolled in the study received one email announcing the newly designed features of the PHR.

XSL•FO
RenderX

These 1070 patients had access to the PHR, mostly as they waited completion of a subsequent study survey, with 110 logging in just to use the PHR. These 110 patients were the most active in using it. We are planning to further advertise the PHR to promote greater awareness for eligible patients.

Discussion

Principal Findings

As a result of extensive focus group-based participatory design and usability testing, we learned that patients with NETs were interested in keeping track of their symptoms and accessing many types of information about their disease, as well as the means to identify personal patterns, interactions and trends in these factors over time. Having the capacity to share these observations more efficiently with their health care providers was of high importance to them. They frequently expressed that they had to become advocates for their own disease management and that these features would help them communicate better with their providers. These findings were different from our previous study designing a PHR with older adults [10,13]. Participants often voiced their appreciation for being treated as partners in the design process, their feedback helped identify navigation issues, data visualization preferences, and feature requests. Their input and improvements based on real-time feedback were crucial to optimizing the PHR.

Rare diseases present a unique set of challenges for patients, health care providers, researchers, and society [36,37]. Rare diseases are often difficult to diagnose because many medical professionals are unfamiliar with them. Patients frequently go through a “diagnostic odyssey” that can take years, seeing multiple specialists before receiving an accurate diagnosis [38]. Moreover, symptoms of rare diseases often overlap with more common conditions, leading to misdiagnosis and inappropriate treatment. Since rare diseases affect a small percentage of the population, there is often less financial incentive for pharmaceutical companies to develop treatments, a phenomenon known as the “orphan drug” problem [39]. Many rare diseases have no approved treatments or cures, leaving patients with limited options for management and quality of life improvement. Specialists for rare diseases are often concentrated in a few medical centers worldwide (ie, centers of excellence), meaning patients may need to travel extensively to receive appropriate care or consultation. In particular, rural areas may lack access to these experts, further isolating patients and delaying necessary care. Furthermore, the rarity of these diseases can mean less public and governmental awareness, leading to fewer resources and funding opportunities for research, support services, and patient care. This has promoted the evolution and development of many patient advocacy organizations, often championed and spearheaded by patients with rare disease and their families in response to their own patient experiences. These groups are essential for raising awareness, driving research and elevating health-care standards, but often have limited resources themselves [40]. Unfortunately, this forces many patients with rare disease to advocate for their own disease, and their particular needs, whether they be informational, expertise or treatment related, physical or emotional. Due to the small patient populations, and often limited data available on them, protecting patient privacy while sharing valuable information for research also presents unique challenges in rare disease studies, and highlights the importance and potential of a coordinated effort between health care organizations, researchers, patient advocacy groups and industry to address these challenges.

NETs in particular are an excellent case-study among rare conditions for the development of a PHR. They are typically slow growing in nature with an unusually prolonged survival and significant symptom burden [41]. Their vague antecedent signs and symptoms cause significant delay and difficulty in their diagnosis and detection [19]. Metastatic disease is observed in up to 40% of NETs at initial diagnosis [42], and many patients with liver or other distant metastases experience carcinoid syndrome, a direct result of serotonin hypersecretion [43] manifesting clinical symptoms ranging from watery diarrhea and flushing to bronchospasm, hypotension and right-sided cardiac deficits (carcinoid heart disease). Thus, patients with NETs typically experience prolonged survival with active disease, and many have significant symptom burdens that may vary daily [44]. Treatment strategies for NETs depend on tumor grade, stage, location, and functional status. Usual treatment for low to intermediate grade gastroenteropancreatic (GEP-NETs, the most common NET subtype) remains surgery, but the requirement for adjuvant therapy is questionable [45-47]. Therapy with somatostatin analogues are established first-line agents for low-grade GEP-NETs, with demonstrated improvements in overall survival [48,49]. However, there are no clear consensus guidelines as to the optimum sequencing of other therapeutic options [50-52], leading many patients to read widely, and become advocates for their own disease and its appropriate management.

To our knowledge, only one previous study has been published evaluating the expansion of a web-based PHR in another rare disease setting. This was a 6-month prospective pilot study involving 43 patients diagnosed with chronic idiopathic thrombocytopenic purpura (ITP) in Marseille, France [18]. The primary objective of the study was to evaluate the usability of the Sanoia tool (a freely accessible web-based PHR) that was expanded with ITP-specific tools including links to validated ITP-related web resources, an emergency protocol based on updated national recommendations, and a section for personal notes, allowing patients to record symptoms, medications, and other relevant events. Health evaluations undertaken and assessed with the Sanoia Interface included quality of life (via an ITP Patient Assessment Questionnaire) and usability of the tool (measured by the frequency of tool usage and patient engagement with the personal notes section). The pilot study demonstrated good usability of the customized Sanoia interface for patients with ITP. While no significant impact on quality of life was observed, the tool facilitated patient engagement and information sharing with health care providers [18].

In comparison, our adaptation of the Iowa PHR was more extensive – whilst we similarly linked out to resources for patients with NETs, we spent more effort in development of new user interfaces to physically capture, visualize and document disease symptoms, track and compare these features over time, and record past and present medication use. An

XSL•FO
RenderX

advantage of the Sanoia interface is that patients could create their PHR, receive a unique personal identifier, and share this identifier with health care providers to facilitate access to their medical information in multiple languages.

When compared to generic PHRs, ours shares functionalities including medication management, education, and self-health monitoring [53]. Unlike some generic PHRs, it is not integrated with a hospital's or health system's records and therefore does not enable patients to access their health records, communicate directly with health care providers, or make appointments [53]. The patients with whom we interacted, all based in the United States, typically had to interact with multiple providers who often used different electronic health records systems. Hence, a PHR that could integrate with all these systems would require significant resources to implement and maintain, making these features beyond the scope of our project. Our PHR, however, addresses an aspect that is missing from most PHRs: personalization [53]. Our PHR is personalized not only to the specific needs of patients with NETs but also gives them the opportunity to further customize what information to track. Addressing patients with NETs needs not only makes the PHR more relevant to patients with NETs, but also more usable by reducing the number of steps necessary to access relevant information and functionality, when compared to a generic PHR.

Anticipated long-term impacts of the PHR could include future updates to extend the current tracking capabilities including refining symptom tracking, medication logs, and data visualization tools, making them more customizable and user-friendly. Adaptive user interfaces may also be possible, offering personalized dashboards and automated data entry suggestions. Similar to the Sanoia interface, a major future development would be integrating with hospital and clinic-based EHRs, allowing seamless data sharing between patients and providers. The PHR's ability to collect patient-reported data over time could contribute significantly to other research studies on rare diseases (ie, future cross-disease adaptations). Establishing ongoing engagement through user feedback surveys, focus groups, and iterative testing to refine features in response to user experiences would close the feedback loop and augment on-going user support and engagement. It may also be beneficial to consider further data privacy and security enhancements (ie, strengthening encryption, access control, and data-sharing policies to ensure continued regulatory compliance).

Overall, developing PHR software tailored to rare disease patients can offer numerous benefits, improving health care experiences, treatment outcomes, and providing research opportunities. Patients with rare disease and their caregivers often play an active role in managing their care. A PHR could provide them direct access to their medical data making it easier to monitor changes, track symptoms, and communicate with health care providers. Importantly, a PHR has the potential to consolidate all relevant health data—such as diagnoses, test results, medications, and treatment plans—into a single, easily accessible platform (ie, centralized medical information). This empowers patients with access to their own health information, promotes self-advocacy and informed decision-making.

Strengths

Through patient-centered design approaches, this project prioritized patient involvement throughout, ensuring that the final PHR addressed real patient needs effectively. Multiple phases of testing and refinement (focus groups, surveys, and usability testing) lead to a well-validated and user-friendly PHR system adapted for a NET population. The final PHR includes symptom tracking, custom tracking options, and access to relevant information and resources for patients with NETs, addressing both immediate and long-term patient information needs. Use of Balsamiq for wireframe development allowed for efficient and cost-effective testing of ideas before full implementation. Our emphasis on NETs addresses a critical gap in health care for rare medical conditions, which are often underserved. As a result, patients in the study have gained a tool that improves their understanding and management of their condition, empowering self-advocacy. The final adapted PHR facilitates longitudinal data collection, benefiting both medical research and patient care, a platform that could easily be tailored to other rare disease applications in similar research settings.

Limitations

A limitation is that the study PHR does not have interoperability with patient medical records or smart health devices. Features that integrate patient reported information with systems-derived clinical records should be developed. Incidentally, patients who participated in Phase 1 often had multiple providers and medical record systems. They reported significant challenges with lack of interoperability across their medical records from different providers.

NETs are not necessarily representative of other rare medical conditions. Simpler or more complex patient needs could result in a need to adjust the methods we presented. Despite the features, only 110 out of 1070 patients logged in specifically to use the PHR, suggesting a need for greater awareness and engagement efforts which we plan to promote. Although valuable, the insights may be skewed by the demographics and availability of the participants involved, potentially overlooking the perspectives of less active or digitally less literate patients.

Conclusions

PHR software can be a useful tool to collect data from patients with their permission as part of research studies while providing them with a valuable service. This strategy could be particularly useful for studies of rare medical conditions, where commercial PHR software may not be available and there is often a need to identify the most effective treatments and optimal therapy sequencing. We presented our experience designing and developing PHR software designed for patients with NETs, a rare medical condition. Our online, patient-centered approach altogether involved 89 patients with NETs in identifying requirements through participatory activities, validating them through a survey, providing feedback on an interactive wireframe prototype, and usability testing a fully implemented system. Not only was patient input valuable in helping us design a useful, usable system, but many of the participating patients enjoyed the activities, having their opinions valued, and in many cases meeting other patients with NETs for the first time. We

XSL•FO
RenderX

expect that the methods we used would work well for similar projects.

Acknowledgments

Research reported in this paper is funded through a Patient-Centered Outcomes Research Institute Award (RD-2020C2-20329), by the National Center for Advancing Translational Sciences of the National Institutes of Health (award UM1TR004403), and supported by Cancer Center Support Grant (P30 CA086862). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Ranking of interest in tracking data (first, second, and third preferences) by patients with neuroendocrine tumor during validation of requirements.
[PNG File, 58 KB - humanfactors_v12i1e68788_app1.png]

Multimedia Appendix 2
Ranking of the types of information and education and advice that could be provided to newly diagnosed patients.
[JPEG File, 71 KB - humanfactors_v12i1e68788_app2.jpeg]

Multimedia Appendix 3
NETPRO guide.
[MP4 File, 203705 KB - humanfactors_v12i1e68788_app3.mp4]

References

1. Apathy NC, Holmgren AJ, Adler-Milstein J. A decade post-HITECH: critical access hospitals have electronic health records but struggle to keep up with other advanced functions. J Am Med Inform Assoc 2021 Aug 13;28(9):1947-1954. [doi: 10.1093/jamia/ocab102] [Medline: 34198342]
2. Jha AK, DesRoches CM, Campbell EG, et al. Use of electronic health records in U.S. hospitals. N Engl J Med 2009 Apr 16;360(16):1628-1638. [doi: 10.1056/NEJMsa0900592]
3. Ribelles N, Alvarez-Lopez I, Arcusa A, et al. Electronic health records and patient registries in medical oncology departments in Spain. Clin Transl Oncol 2021 Oct;23(10):2099-2108. [doi: 10.1007/s12094-021-02614-9] [Medline: 33864619]
4. Irizarry T, DeVito Dabbs A, Curran CR. Patient portals and patient engagement: a state of the science review. J Med Internet Res 2015 Jun 23;17(6):e148. [doi: 10.2196/jmir.4255] [Medline: 26104044]
5. Martínez Nicolás I, Lê Cook B, Flores M, et al. The impact of a comprehensive electronic patient portal on the health service use: an interrupted time-series analysis. Eur J Public Health 2019 Jun 1;29(3):413-418. [doi: 10.1093/eurpub/cky257] [Medline: 30544169]
6. King CE, Sarrafzadeh M. A survey of smartwatches in remote health monitoring. J Healthc Inform Res 2018 Jun;2(1-2):1-24. [doi: 10.1007/s41666-017-0012-7] [Medline: 30035250]
7. Kang M, Park E, Cho BH, Lee KS. Recent patient health monitoring platforms incorporating internet of things-enabled smart devices. Int Neurourol J 2018 Jul;22(Suppl 2):S76-S82. [doi: 10.5213/inj.1836144.072] [Medline: 30068069]
8. Ozok AA, Wu H, Gurses AP. Exploring patients' use intention of personal health record systems: implications for design. International Journal of Human–Computer Interaction 2017 Apr 3;33(4):265-279. [doi: 10.1080/10447318.2016.1277637]
9. Taha J, Czaja SJ, Sharit J, Morrow DG. Factors affecting usage of a personal health record (PHR) to manage health. Psychol Aging 2013;28(4):1124-1139. [doi: 10.1037/a0033911]
10. Chrischilles EA, Hourcade JP, Doucette W, et al. Personal health records: a randomized trial of effects on elder medication safety. J Am Med Inform Assoc 2014;21(4):679-686. [doi: 10.1136/amiajnl-2013-002284] [Medline: 24326536]
11. Andrikopoulou E. A mixed methods study of electronic personal health record design and use for people with long-term conditions [Dissertation]. : University of Portsmouth; 2021.
12. Weitzman ER, Kelemen S, Kaci L, Mandl KD. Willingness to share personal health record data for care improvement and public health: a survey of experienced personal health record users. BMC Med Inform Decis Mak 2012 May 22;12(1):39. [doi: 10.1186/1472-6947-12-39] [Medline: 22616619]
13. Hourcade JP, Chrischilles EA, Gryzlak BM, et al. Design lessons for older adult personal health records software from older adults. In: International Conference on Universal Access in Human-Computer Interaction: Springer; 2011:176-185. [doi: 10.1007/978-3-642-21663-3_19]

XSL•FO
RenderX

14. Nestler-Parr S, Korchagina D, Toumi M, et al. Challenges in research and health technology assessment of rare disease technologies: report of the ISPOR rare disease special interest group. Value Health 2018 May;21(5):493-500. [doi: 10.1016/j.jval.2018.03.004] [Medline: 29753344]
15. Boulanger V, Schlemmer M, Rossov S, Seebald A, Gavin P. Establishing patient registries for rare diseases: rationale and challenges. Pharmaceut Med 2020 Jun;34(3):185-190. [doi: 10.1007/s40290-020-00332-1] [Medline: 32215853]
16. Griggs RC, Batshaw M, Dunkle M, et al. Clinical research for rare disease: opportunities, challenges, and solutions. Mol Genet Metab 2009 Jan;96(1):20-26. [doi: 10.1016/j.ymgme.2008.10.003] [Medline: 19013090]
17. Price M, Bellwood P, Kitson N, Davies I, Weber J, Lau F. Conditions potentially sensitive to a personal health record (PHR) intervention, a systematic review. BMC Med Inform Decis Mak 2015 Apr 18;15(1):32. [doi: 10.1186/s12911-015-0159-1] [Medline: 25927384]
18. Chiche L, Brescianini A, Mancini J, Servy H, Durand JM. Evaluation of a prototype electronic personal health record for patients with idiopathic thrombocytopenic purpura. Patient Prefer Adherence 2012;6:725-734. [doi: 10.2147/PPA.S36320] [Medline: 23077409]
19. Darbà J, Marsà A. Exploring the current status of neuroendocrine tumours: a population-based analysis of epidemiology, management and use of resources. BMC Cancer 2019 Dec;19(1):1-7. [doi: 10.1186/s12885-019-6412-8]
20. Kulke MH, Benson AB 3rd, Bergsland E, et al. Neuroendocrine tumors. J Natl Compr Canc Netw 2012 Jun 1;10(6):724-764. [doi: 10.6004/jnccn.2012.0075] [Medline: 22679117]
21. O'Rorke M, Chrischilles E, NET‐PRO Study Investigators. Making progress against rare cancers: a case study on neuroendocrine tumors. Cancer 2024 May 1;130(9):1568-1574. [doi: 10.1002/cncr.35184] [Medline: 38244195]
22. Hendriks N, Truyen F, Duval E. Designing with dementia: guidelines for participatory design together with persons with dementia. In: Kotzé P, Marsden G, Lindgaard G, Wesson J, Winckler M, editors. Human-Computer Interaction – INTERACT 2013: Springer; 2013:649-666. [doi: 10.1007/978-3-642-40483-2_46]
23. Gioia DA, Corley KG, Hamilton AL. Seeking qualitative rigor in inductive research. Organ Res Methods 2013 Jan;16(1):15-31. [doi: 10.1177/1094428112452151]
24. An evaluation of three democratic, community-based approaches to citizen participation: surveys, conversations with community groups, and community dinners. Soc Nat Resour 2001 Feb;14(2):107-126. [doi: 10.1080/089419201300000526]
25. Balsamiq. URL: https://balsamiq.com/ [accessed 2022-04-21]
26. Mahatody T, Sagar M, Kolski C. State of the art on the cognitive walkthrough method, its variants and evolutions. Int J Hum Comput Interact 2010 Jul 30;26(8):741-785. [doi: 10.1080/10447311003781409]
27. Roth RE, Hart D, Mead R, Quinn C. Wireframing for interactive & web-based geographic visualization: designing the NOAA Lake Level Viewer. Cartogr Geogr Inf Sci 2017 Jul 4;44(4):338-357. [doi: 10.1080/15230406.2016.1171166]
28. Information about Java 8. Java. URL: https://www.java.com/en/download/help/java8.html [accessed 2025-02-11]
29. Spring. URL: https://spring.io/ [accessed 2025-02-11]
30. Microsoft SQL Server. Microsoft. URL: https://www.microsoft.com/en-us/sql-server [accessed 2025-02-11]
31. The Eclipse Jetty Project. URL: https://jetty.org/index.html [accessed 2025-02-11]
32. The Apache HTTP Server Project. URL: https://httpd.apache.org/ [accessed 2025-02-11]
33. Dumas JS, Redish J. A practical guide to usability testing. . 1999.
34. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research advances in psychology. In: Advances in Psychology: Elsevier; 1988, Vol. 52:139-183. [doi: 10.1016/S0166-4115(08)62386-9]
35. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208. [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
36. Stoller JK. The challenge of rare diseases. Chest 2018 Jun;153(6):1309-1314. [doi: 10.1016/j.chest.2017.12.018] [Medline: 29325986]
37. Tidman J. Why rare diseases are posing a massive societal challenge. Elsevier. 2024 Feb 26. URL: https://www.elsevier.com/connect/why-rare-diseases-are-posing-a-massive-societal-challenge [accessed 2024-11-12]
38. Wu AC, McMahon P, Lu C. Ending the diagnostic odyssey-is whole-genome sequencing the answer? JAMA Pediatr 2020 Sep 1;174(9):821-822. [doi: 10.1001/jamapediatrics.2020.1522] [Medline: 32597967]
39. Aronson JK. Rare diseases and orphan drugs. Br J Clin Pharmacol 2006 Mar;61(3):243-245. [doi: 10.1111/j.1365-2125.2006.02617.x] [Medline: 16487216]
40. Taruscio D, Gahl WA. Rare diseases: challenges and opportunities for research and public health. Nat Rev Dis Primers 2024 Feb 29;10(1):13. [doi: 10.1038/s41572-024-00505-1] [Medline: 38424095]
41. Singh S, Granberg D, Wolin E, et al. Patient-reported burden of a neuroendocrine tumor (NET) diagnosis: results from the first global survey of patients with NETs. J Glob Oncol 2017 Feb;3(1):43-53. [doi: 10.1200/JGO.2015.002980] [Medline: 28717741]
42. Hermans BCM, de Vos-Geelen J, Derks JL, et al. Unique metastatic patterns in neuroendocrine neoplasms of different primary origin. Neuroendocrinology 2021;111(11):1111-1120. [doi: 10.1159/000513249] [Medline: 33227805]
43. Oronsky B, Ma PC, Morgensztern D, Carter CA. Nothing but NET: a review of neuroendocrine tumors and carcinomas. Neoplasia 2017 Dec;19(12):991-1002. [doi: 10.1016/j.neo.2017.09.002] [Medline: 29091800]

XSL•FO
RenderX

44. Adams JR, Ray D, Willmon R, Pulgar S, Dasari A. Living with neuroendocrine tumors: assessment of quality of life through a mobile application. JCO Clin Cancer Inform 2019 Jul;3:1-10. [doi: 10.1200/CCI.19.00025] [Medline: 31283354]
45. Caplin ME, Baudin E, Ferolla P, et al. Pulmonary neuroendocrine (carcinoid) tumors: European Neuroendocrine Tumor Society expert consensus and recommendations for best practice for typical and atypical pulmonary carcinoids. Ann Oncol 2015 Aug;26(8):1604-1620. [doi: 10.1093/annonc/mdv041] [Medline: 25646366]
46. Garcia-Carbonero R, Sorbye H, Baudin E, et al. ENETS consensus guidelines for high-grade gastroenteropancreatic neuroendocrine tumors and neuroendocrine carcinomas. Neuroendocrinology 2016;103(2):186-194. [doi: 10.1159/000443172] [Medline: 26731334]
47. Lamarca A, Clouston H, Barriuso J, et al. Follow-up recommendations after curative resection of well-differentiated neuroendocrine tumours: review of current evidence and clinical practice. J Clin Med 2019 Oct 5;8(10):1630. [doi: 10.3390/jcm8101630] [Medline: 31590343]
48. Tsoli M, Chatzellis E, Koumarianou A, Kolomodi D, Kaltsas G. Current best practice in the management of neuroendocrine tumors. Ther Adv Endocrinol Metab 2019;10:2042018818804698. [doi: 10.1177/2042018818804698] [Medline: 30800264]
49. Stueven AK, Kayser A, Wetz C, et al. Somatostatin analogues in the treatment of neuroendocrine tumors: past, present and future. Int J Mol Sci 2019 Jun 22;20(12):3049. [doi: 10.3390/ijms20123049] [Medline: 31234481]
50. Hicks RJ, Kwekkeboom DJ, Krenning E, Antibes Consensus Conference participants, et al. ENETS consensus guidelines for the standards of care in neuroendocrine neoplasia: peptide receptor radionuclide therapy with radiolabeled somatostatin analogues. Neuroendocrinology 2017;105(3):295-309. [doi: 10.1159/000475526] [Medline: 28402980]
51. Zaknun JJ, Bodei L, Mueller-Brand J, et al. The joint IAEA, EANM, and SNMMI practical guidance on peptide receptor radionuclide therapy (PRRNT) in neuroendocrine tumours. Eur J Nucl Med Mol Imaging 2013 May;40(5):800-816. [doi: 10.1007/s00259-012-2330-6]
52. Passhak M, McNamara MG, Hubner RA, Ben-Aharon I, Valle JW. Choosing the best systemic treatment sequence for control of tumour growth in gastro-enteropancreatic neuroendocrine tumours (GEP-NETs): What is the recent evidence? Best Pract Res Clin Endocrinol Metab 2023 Sep;37(5):101836. [doi: 10.1016/j.beem.2023.101836]
53. Harahap NC, Handayani PW, Hidayanto AN. Functionalities and issues in the implementation of personal health records: systematic review. J Med Internet Res 2021 Jul 21;23(7):e26236. [doi: 10.2196/26236] [Medline: 34287210]

Abbreviations

GEP-NET: gastroenteropancreatic neuroendocrine tumor
ITP: idiopathic thrombocytopenic purpura
NET: neuroendocrine tumor
PHR: personal health record
REDCap: Research Electronic Data Capture

Edited by A Kushniruk; submitted 27.11.24; peer-reviewed by E Okoyeocha, T Awofala; revised version received 03.04.25; accepted 04.04.25; published 03.06.25.

Please cite as:
Hourcade JP, O'Rorke M, Chrischilles E, Rudzianski NJ, Gryzlak B, Peterman K, Ortman C, Wolstencroft S, Bean M, Gellerman E, Mailman JA, Wahmann M
Personal Health Record Software for Neuroendocrine Tumors: Patient-Centered Design Approach
JMIR Hum Factors 2025;12:e68788
URL: https://humanfactors.jmir.org/2025/1/e68788
doi:10.2196/68788

© Juan Pablo Hourcade, Michael O'Rorke, Elizabeth Chrischilles, Nicholas J Rudzianski, Brian Gryzlak, Kerry Peterman, Chris Ortman, Sam Wolstencroft, Margaret Bean, Elyse Gellerman, Josh A Mailman, Maryann Wahmann. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 3.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Using Social Media Platforms to Raise Health Awareness and Increase Health Education in Pakistan: Structural Equation Modeling Analysis and Questionnaire Study

Malik Mamoon Munir, PhD; Nabil Ahmed, MPH

Integral Global Consulting LLC, 3541 Habersham at Northlake Parkway Road, Building E, Tucker, GA, United States

Corresponding Author:
Malik Mamoon Munir, PhD
Integral Global Consulting LLC, 3541 Habersham at Northlake Parkway Road, Building E, Tucker, GA, United States

Abstract

Background: Current health care education methods in Pakistan use traditional media (eg, television and radio), community health workers, and printed materials, which often fall short of reach and engagement among most of the population. The health care sector in Pakistan has not yet used social media effectively to raise awareness and provide education about diseases. Research on the impact social media can have on health care education in Pakistan may expand current efforts, engage a wider audience, and reduce the disease burden on health care facilities.

Objective: This study aims to evaluate the perceptions of health care professionals and paramedic staff regarding social media use to raise awareness and educate people about diseases as a potential means of reducing the disease burden in Pakistan.

Methods: The study used two-stage structural equation modeling (SEM). Data analysis used AMOS 26.0 software, adopting scales from previous literature. Four-item scales for each social media usefulness and health awareness construct and 8-item scales for health care education constructs were adopted on the basis of their higher loading in alignment with psychometric literature. A 7-point Likert scale was used to measure each item. Data collection used convenience sampling, with questionnaires distributed to more than 450 health care professionals and paramedic staff from 2 private hospitals in Lahore, Pakistan. There were 389 useful responses received. However, 340 completed questionnaires were included in the data analysis.

Results: The study found that all the squared multiple correlation (SMC) values were greater than 0.30. Furthermore, convergent validity was measured using (1) standardized factor loading (found greater than 0.5), (2) average variance explained (found greater than 0.5), and (3) composite reliability (found greater than 0.7). The confirmatory factor analysis (CFA) of the measurement model indicated the fitness of the constructs (Chi-square minimum [CMIN]=357.62; CMIN/degrees of freedom [DF]=1.80; Goodness of Fit [GFI]=0.90; Adjusted Goodness of Fit Index [AGFI]=0.89; Buntler-Bonett Normed Fit Index [NFI]=0:915; Comparative Fit Index [CFI]=0:93; Root Mean Square Residual [RMR]=0:075; Root Mean Square Error of Approximation [RMSEA]=0:055). Moreover, the structural model fitness was also confirmed (CMIN=488.6; CMIN/DF=1.85; GFI=0.861; AGFI=0.893; NFI=0.987; CFI=0.945; RMR=0:079; RMSEA=0.053). Hence, the results indicated that social media usefulness has a positive and significant effect on health awareness (hypothesis 1: β=.669, *P*<.001), and health awareness has a positive and significant effect on health care education in Pakistan (hypothesis 2: β=.557, *P*<.001).

Conclusions: This study concludes that health care professionals and paramedic staff in private hospitals support the use of social media to raise awareness and provide health care education. It is considered an effective tool for reducing the disease burden in Pakistan. The study results also revealed that young health care professionals are more inclined toward social media usage and express the need for legislation to support it and establish a monitoring process to avoid misinformation.

(JMIR Hum Factors 2025;12:e65745) doi:10.2196/65745

KEYWORDS

social media; health awareness; health education; innovation diffusion theory; structural equation modeling; disease burden; healthcare facilities; health professionals; misinformation; cost effective

Introduction

Background

In Pakistan, the health care system is resource-strapped, leading to challenges in preventing infectious diseases. The country has a high maternal mortality rate (186 deaths per 100,000 live births) and a high infant mortality rate (56 deaths per 1000 live births) [1]. Many complications arise from a lack of awareness and education about prenatal care, safe delivery practices, and postnatal care [2]. Moreover, poor understanding of nutrition has led to widespread malnutrition among children and pregnant

XSL•FO
RenderX

women in Pakistan [1,2]. Malnutrition in children leads to stunted growth, weakened immunity, and impaired cognitive development [2]. The prevalence of stunting in Pakistan is among the highest in the world, affecting approximately 38% of children under 5 years of age [3].

In this context, health awareness and education campaigns potentially lead to behavioral changes and alleviate the burden on health facilities [4]. Public health campaigns, particularly those focused on preventive measures like vaccination, hygiene, and lifestyle changes, have substantially reduced the incidence of various diseases [3,5]. For example, handwashing campaigns alone have reduced respiratory infections by up to 23% [6]. Awareness campaigns aimed at managing chronic conditions, such as diabetes and hypertension, through education on medication adherence and lifestyle modifications have reduced complications and hospital visits [7]. These interventions have been shown to improve health outcomes and reduce health care costs by up to 25% [8].

Among various tools and health interventions, social media has attained the most global attention to enhance health awareness and education to a wider population and discovered its significant impact in addressing health care issues [9,10]. According to social media statistics, social media users worldwide reached 5.04 billion in January 2024, representing 63% of the world's population, with users growing 6% annually [11]. On average, social media users spend 2 hours and 23 minutes daily on social media platforms and see various content [11,12]. Social media tools, such as Facebook (Meta), TikTok (ByteDance), Twitter (subsequently rebranded as X), Instagram (Meta), and YouTube (Google), facilitate health care professionals in educating and influencing perceptions, including networking, sharing ideas, disseminating information, demonstrating, coaching, consulting, and advertising [12,13]. Existing literature in the health care sector has exhibited the transformative potential of social media and its usefulness in spreading health awareness and education [9,10,12,13].

However, the existing literature lacks comprehensive insights into the role of social media in fostering health awareness and education within developing countries, particularly Pakistan. This study addresses this gap by applying Innovation Diffusion Theory to examine the mediating effect of health awareness and empirically testing the impact of social media on health care education among health care professionals.

Therefore, the primary objective of this study is to investigate the role of social media in enhancing health awareness and its subsequent impact on health care education in Pakistan. By addressing the gap in the literature concerning the adoption of social media for health care communication in developing countries, this research aims to provide empirical evidence and practical insights for leveraging digital platforms to improve health outcomes.

This section comprehensively reviews existing literature on social media usefulness, health awareness, health care education, and health care sector challenges in Pakistan.

Social Media Usefulness

The global adoption of social media platforms has led to scalable public health interventions, increased precision and effectiveness, and improved resource mobilization [14]. Social media has encouraged healthy behaviors and enabled informed choices through sharing tips, challenges, and success stories [9]. Health care professionals can provide accurate, evidence-based information as trusted sources [11]. Their involvement is vital due to their expertise, credibility, and ability to reach a wide audience. Campaigns through social media on vaccination, hygiene practices, mental health awareness, and other preventive measures have been helpful in countries like Germany and France [9,10].

Moreover, social media has been effectively used to combat the pandemic and other diseases during the COVID-19 pandemic [15]. The social media data were instrumental in tracking the spread of the COVID-19 virus and disseminating accurate information [16]. Campaigns on platforms such as Facebook and Twitter have been used to promote vaccination, leading to increased uptake in various regions [12]. Tailored messaging on HIV prevention for at-risk populations has shown to be more effective than generalized campaigns [13]. Moreover, existing literature has shown examples of successful public health campaigns, such as:

1. #SmearForSmear Campaign: this social media campaign aimed to raise awareness about cervical cancer screening and successfully increase screening rates by leveraging the reach of platforms such as Twitter and Instagram [17].
2. Truth Initiative: an antismoking campaign that effectively used social media to reduce teen smoking rates through engaging content and interactive platforms [18].

Health Awareness

Health awareness is termed as "Enhancing the ability of individuals to understand and use health information to make informed decisions about their health" [13]. Efforts for health awareness are varied and multifaceted, encompassing a wide range of activities aimed at educating the public about health issues, promoting healthy behaviors, and preventing diseases [16]. Health awareness campaigns lead to early detection and timely intervention, which benefits individuals by lowering their health care expenses and reducing the financial burden on the health care system [18]. Health awareness is generally performed using mass media, social media, posters, billboards, workshops, seminars, local outreach programs, influencers, and public figures [12]. However, social media is most useful in the current era due to its cost-effectiveness and wider audience engagement [17]. Social media allows targeted messaging based on demographics, interests, and online behavior [19]. This ensures that health messages reach the people most likely to benefit from them. Moreover, the effectiveness of health awareness through social media can be measured in a few easy ways, such as the number of impressions, likes, shares, comments, hashtag usage, and so on [18,19]. Hence, medical professionals can analyze the efficacy of their awareness messages. There have been several successful health awareness campaigns conducted through social media that have achieved significant outcomes (Textbox 1).

XSL•FO
RenderX

Textbox 1. Health care awareness campaigns.

- **ThisGirlCan (2015)**
  - Objective: Encourage women of all ages and backgrounds to be more physically active, regardless of their shape, size, or fitness level [20].
  - Execution: The campaign featured real women participating in various physical activities, promoting the message that women should feel confident about being active regardless of societal pressures or body image concerns.
  - Platforms used: Instagram, Twitter, Facebook, and YouTube.
  - Outcomes:
    - Widespread engagement: The hashtag #ThisGirlCan was used millions of times, with significant engagement from women sharing their own stories and photos.
    - Behavioral change: The campaign was credited with encouraging over 2.8 million women in the United Kingdom to become more active.
    - Long-term impact: The campaign continued beyond its initial phase, evolving into a broader movement that still influences public attitudes towards women and fitness.
- **#BellLetsTalk (Ongoing)**
  - Objective: Raise awareness about mental health issues and reduce stigma around mental illness in Canada [21].
  - Execution: For every tweet using the hashtag #BellLetsTalk, every text message sent by Bell customers, and every Facebook video view, Bell Canada donates 5 cents to mental health initiatives.
  - Platforms used: Twitter, Facebook, Instagram, and Snapchat.
  - Outcomes:
    - Record participation: The 2021 campaign saw 159 million interactions, raising nearly CAD $8 million (equivalent US $5.56 million) in a single day.
    - Sustained impact: Since its inception, the campaign has raised more than CAD $121 million (equivalent US $84.5 million) for mental health initiatives.
    - Increased conversations: The campaign has significantly increased public dialogue about mental health in Canada, contributing to reducing stigma and promoting mental wellness.

Health Care Education

Health care education refers to a deeper understanding of specific health topics, including the causes, symptoms, prevention, and treatment of diseases and how to maintain or improve health [20]. It involves having factual information and comprehension of health-related subjects [17]. Health care education encourages regular check-ups, screenings, and vaccinations, leading to early detection of diseases and more effective prevention strategies [19]. For those with chronic conditions such as diabetes, hypertension, or asthma, health care education is essential for managing symptoms, adhering to treatment plans, and avoiding complications [13].

In resource-limited economies, insufficient health care education has been reported among the population, which impedes individuals' ability to understand, access, and apply health information effectively [21]. Moreover, low-income individuals may have less access to health resources, education, and information, which hinders their ability to acquire and apply health education [22]. Addressing these barriers in resource-limited countries requires targeted strategies such as improving health literacy, offering culturally sensitive health care education, increasing access to technology, and combating misinformation [21].

Among these targeted strategies, social media has gained the most attention in the high-income world [20]. Health care education through social media has become an increasingly effective method for reaching diverse audiences with health information. Studies suggest that health care education campaigns on social media can achieve behavioral change success rates ranging from 20% to 40% [16,17]. For instance, social media campaigns promoting COVID-19 vaccination have seen varying success, with some countries reporting a 15% to 25% increase in vaccine uptake attributable to social media efforts [15]. Moreover, some campaigns that promote health services (eg, vaccination drives or mental health counseling) report conversion rates (actual service uptake) of 5% to 20%, depending on the call to action and the ease of access to the service [18]. The existing literature has discussed a few notable health care education campaigns executed through social media and validated their impact. One such campaign for spreading health care education is The Amyotropic Lateral sclerosis (ALS) Ice Bucket Challenge (Textbox 2).

XSL•FO
RenderX

Textbox 2. Health care education campaigns through social media.

1. The ALS Ice Bucket Challenge (2014)

- Overview: Participants dumped buckets of ice water over their heads, shared videos on social media, and nominated others to do the same, all to raise awareness and funds for Amyotrophic Lateral Sclerosis (ALS) research [23].
- Impact:
  - Financial: Raised over US $115 million for the ALS Association in just a few months.
  - Awareness: Dramatically increased global awareness of ALS, with millions participating worldwide.
  - Research Advancement: Funds contributed to significant research advancements, including the discovery of new ALS genes.

Pakistan Context

In Pakistan, the Federal and provincial governments jointly administer 60% of the health care system, with the private sector contributing 30%. Autonomous bodies support the remaining 10 percent [24]. The diverse structure and the country's economic challenges pose unique obstacles to effective health care delivery. The country's health care spending is at 2.95% of its Gross Domestic Product, and the Pakistani government always disregarded the economic survey data and faced criticism from the apex medical organization as it necessitated health allocations in line with global guidelines [25]. To provide quality health care, the Pakistan Medical Association (PMA) expressed the need for 6% of the country's Gross Domestic Product allocation, as recommended by the World Health Organization [17], which the country is unable to meet due to economic challenges [26].

Due to these financial limitations, there are reported disparities in health awareness, with urban areas generally having better access than rural regions [2]. Thus, communicable diseases, such as waterborne diseases and vector-borne diseases, remain a more significant concern in rural regions [1,3]. Moreover, with limited access to maternal health care and family planning services, Pakistan witnesses high maternal and child mortality rates [2]. Diseases such as malaria, tuberculosis, and hepatitis are prevalent, which exacerbates the situation [24]. Moreover, stigma about sexually transmitted infections results in inadequate prevention and treatment efforts that result in a psychological attack on patients [25]. Poor sanitation and hygiene contribute to outbreaks of diseases like cholera and dysentery in Pakistan and many preventable diseases, such as polio and measles, persist due to low vaccination rates [23].

Amid financial crises, health care awareness and education are known to be effective in reducing the disease burden rate [8]. There are several traditional efforts made by the Government of Pakistan Ministry of Health in the past to improve disease awareness and reduce the patient burden [24]. These include the following:

1. Engaging community health workers in disseminating health information, especially in rural areas. However, the number of community health workers to cover the target population is very limited and requires substantial spending [24].
2. Use of radio and television for health awareness programs, public service announcements, and talk shows but requires financial resources [27].
3. Newspapers, magazines, and pamphlets are also used to spread health information, but considering the low literacy rate and cost of print media, these are not found to be much effective for the target population [27].

However, among the modern techniques, social media has gained much attention from health care communities globally [7,10]. Pakistan has also witnessed the widespread of social media but in academics, e-commerce, entertainment, media, politics, sports, and religious sectors. In Pakistan, 29.5% of the population is on social media, whereas 77.8% has active mobile connections, meaning these users also have access to social media platforms [28].

The health care sector is structured to depend on government approval for any initiative [29]. The government has not devised any legislation to promote the use of social media for awareness [24,27]. This mainly restricts the government's health departments and health care professionals from using this tool to publicize health messages. Some efforts have been made to digitalize the health care sector using the eHealth concept and implement a few apps to initiate telehealth and telemedicine [30]. Also, mobile messaging and caller voice tunes are found to be significant in health care services awareness [30]. However, the interpretation and usage of social media tools are still scarce [31]. For this reason, it is imperative to investigate the social media usefulness among health care professionals in Pakistan and to identify whether these health care professionals support the use of social media for health awareness and education.

Theoretical Foundation

A few theories explain the behaviors toward acceptance of any new digitalized tool and address their attitude [32], including the innovation diffusion theory proposed by Rogers in 1962, the Theory of Reasoned Action proposed by Fishbein and Ajzen in 1967, the Self-Efficacy Theory proposed by Bandura in 1977, the Theory of Planned Behavior proposed by Ajzen in 1985, the Social Cognitive Theory proposed by Bandura in 1986, and the Technology Acceptance Model proposed by Davis in 1986 [32]. Among these various behavioral theories, the Innovation Diffusion Theory (IDT) offers a framework for understanding how new ideas, behaviors, or innovations spread within a population [33].

In public health, innovation diffusion theory helps explain how new health interventions, practices, or policies are adopted and disseminated within communities [33]. The theory identifies 5

XSL•FO
RenderX

stages individuals and communities pass when adopting an innovation: knowledge, persuasion, decision, implementation, and confirmation [34]. Figure 1 explains these five stages of the innovation-decision process.

To understand social media acceptance for health awareness and education, the theory has categorized individuals within a population into subgroups based on their readiness to adopt innovations. Innovators and early adopters are likely to adopt innovations early, while most of the population follows suit over time. Laggards are the last to adopt [33]. Figure 2 explains these subgroups of individuals.

Figure 1. Innovation diffusion process.

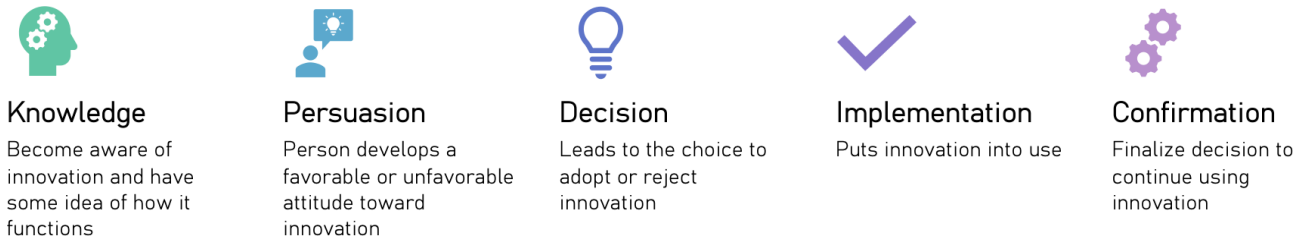
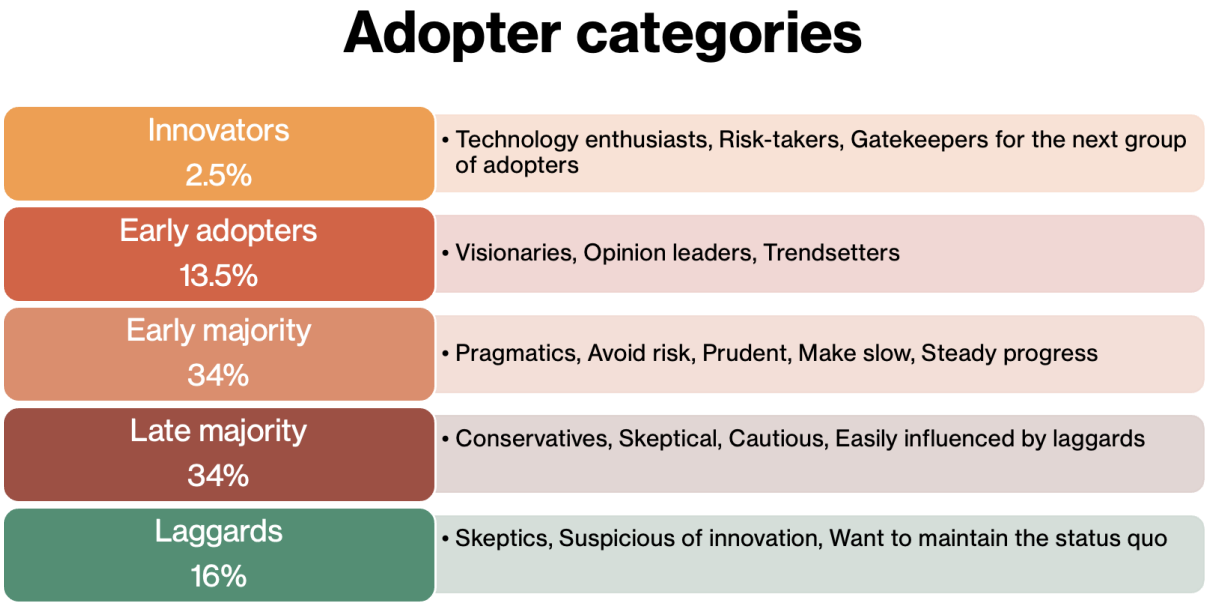


Figure 2. Adopter categories.



In health care education and awareness, the diffusion of innovation is used to fast-track the acceptance of crucial public health digitalized interventions that ordinarily intend to influence the behavior of a social system [34]. Thus, this theory provides the foundation for understanding social media usage and its adoption attitude in a particular community.

In Pakistan, the ineffective innovation management approaches and deprived diffusion of innovation strategies are hurdles to realizing the importance of social media and its usage for health care education [24,25]. The deprived diffusion of innovation is a significant factor in the inability to attain the desired output. Even for digitalized hospital tools other than social media, the country needs more innovative and user-friendly equipment in the hospital sector [25]. The notion of consistent health care challenges and inadequate health resources in Pakistan is also evident from the recent Joint External Evaluation report [24].

For this reason, it is imperative to investigate the impact of social media's effective use on health awareness and health care education in Pakistan using IDT. This may result in highlighting and understanding the effectiveness of social media tools in reducing the disease burden in the country.

Figure 3 explains the theoretical framework of the current study.

XSL•FO
RenderX

Figure 3. Conceptual framework of the study.



Hence, the current study hypotheses are:

1. Hypothesis 1 (H1): social media usefulness positively and significantly impacts health awareness in Pakistan.
2. Hypothesis 2 (H2): health awareness has a significant positive impact on health care education in Pakistan.

Methods

Instrument Development

A quantitative technique was applied to test the study hypothesis and validate the proposed model using the structural equation modeling (SEM) technique [35]. SEM is deemed the most suitable due to its ability to simultaneously test complex relationships among multiple variables and validate the proposed conceptual model. Data analysis was carried out using AMOS 26.0 software (IBM Corp) [36]. Since the study aims to identify social media's impact on health awareness leading to health care education and uses the IDT as a theoretical foundation, scales from previous literature have been adopted. The main advantage of adopting these scales is that they have already been validated in the health industry's acceptance behavior of new technological tools. For instance, the measurement scale for Innovative Technology (social media) usefulness was adopted from a previous study [37], and measurement scales for health awareness have been adopted from another study [38]. Aligning with the psychometric literature, 4-item scales for each construct based on their higher loadings have been adopted [37]. The modified 8-item scale for measuring health care education was adopted based on the questionnaire of Ho et al [39]. For the measurement of each item, a 7-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree) was used [38].

Data Collection

A questionnaire was designed based on the variables of the present study, utilizing the aforementioned scales, which are provided in Multimedia Appendix 1. A personally administered questionnaire was used for data collection [32]. For data collection, the current research to eliminate targeted health care professionals and paramedic staff of 2 private hospitals in Lahore, Pakistan. The questionnaire was distributed to more than 450 health care professionals and paramedic staff using convenience sampling and requested their valuable feedback. However, only 389 useful responses were received.

Ethical Considerations

As a matter of ethical concerns, prior permission from the hospital management was obtained, and the purpose of the study was explained to all respondents through "informed consent." It was assured that respondents' privacy would be ensured through confidentiality. Ethical approval was not required for this study as it did not involve identifiable personal data, or interventions falling under the scope of institutional or national ethics review requirements. In accordance with the Declaration of Helsinki (World Medical Association, 2013) research ethics policy, studies that do not involve sensitive data are exempt from formal ethics board approval.

Results

Primary Data Screening

Initial screening was carried out among 389 useful responses to eliminate unusual responses, mainly incomplete questionnaires and respondents who marked similar scores on all items. Furthermore, questionnaires with more than 4 missing values were also excluded. The remaining missing values were replaced with the series mean/average for the rest of the responses. Therefore, the final useful sample was reduced to 340.

Table 1 provides the demographics of the study respondents. Among 340 respondents, 58% (197/340) were male, and the rest were female (143/340, 42%). A total of 91% (309/340) of the respondents were between 25 and 35 years of age, which may lead to the fact that most health care professionals and paramedic staff in private hospitals in Lahore are young. However, we may be unable to neglect the 6% of the population under 25 years old and more eager to start their career in the health care sector. The remaining 3% were more than 35 years of age and had extensive health industry work experience.

XSL•FO
RenderX

Table . Sociodemographic characteristics of study respondents (N=340).

Characteristics	Values, n (%)
Sex	
Male	197 (58)
Female	143 (42)
Age (years)	
<25	21 (6)
25-35	309 (91)
>35	10 (3)
Social media experience (years)	
<1	34 (10)
2-5	207 (61)
>5	99 (29)
Monthly income[a]	
Less than PKR[b] 40,000 (US $142.98)	101 (30)
PKR 40,000-80,000 (US $142.98-US $285.97)	197 (58)
More than 80,000 (US $285.97)	42 (12)
Education	
Less than secondary school	63 (18)
Secondary school to graduation	259 (76)
Medical (MBBS[c]/BDS)[d] or Masters degree	18 (6)
Health services experience (years)	
<3	21 (6)
3-6	225 (66)
>6	94 (28)

[a]A currency exchange rate of US $1=280 PKR is applicable.

[b]PKR: Pakistani rupees.

[c]MBBS: Bachelor of Medicine, Bachelor of Surgery.

[d]BDS: Bachelor of Dental Surgery.

Data Analysis

An exploratory factor analysis was initially performed on the 16 measurement items to confirm the underlying relationships [36]. For sample adequacy, the KMO value (0.850) was greater than the recommended value (0.60), and the significance value of *P*0.001 confirmed the Bartlett test of sphericity, as shown in Table S1 in Multimedia Appendix 2. All those factors were retained for data analysis, with factor loading greater than 0.50 and Eigenvalue greater than 1.0. Hence, this technique provided 3 factors that explained 72.2% of the variance after varimax rotation. Furthermore, scale reliability coefficients were greater than the acceptable value (0.70) [40], and Shapiro-Wilk tests were conducted to check the data normality [40], as shown in Table S2 in Multimedia Appendix 2 . The results of both tests were significant, meaning there was no normality issue in the data.

For further data analysis, a two-stage structural equation modeling (SEM) technique was used [36-38]. This approach allows researchers to evaluate the measurement and structural models separately using two dissimilar subsamples. For this reason, a sample of 340 participants was divided into two parts: a sample of 170 individuals was used for the measurement model, and a similar sample size was used for the structural model assessment to attain impartial results.

Reliability and Validity Measures

According to Hair et al [35], finding the reliability of each item and construct in the research study is essential. Therefore, squared multiple correlation (SMC) was used to find the reliability of each reliability item of each measurement item. SMC represents “the amount of variance explained by an individual indicator/construct of its respective factor and measured by the square of its (indicator’s) standardized factor loading” [40].

Hair [35] indicated that the cutoff value of SMC is 0.30. Hence, it is evident from Table 2 that all the SMC values are greater than 0.30. For measuring the reliability of each variable, the

XSL•FO
RenderX

Cronbach alpha value was used. As suggested in the literature, the cutoff value of Cronbach alpha is 0.70 [36,37]; it is evident from Table 2 that all variables' Cronbach alpha values are more than 0.70. Furthermore, Hair et al [35] also highlighted that for measuring the convergent validity, there are three common approaches: (1) standardized factor loading (0.5 or greater), (2) average variance explained (0.5 or higher), and (3) composite reliability (0.7 or above).

Table . Measurement of reliability and validity.

Constructs and item	Values, mean (SD)	Standardized factor loading	Squared multiple correlation	Composite reliability	Cronbach alpha	Average variance explained
Social media usefulness (SMU)						
SMU1	5.19 (1.65)	0.781	0.607	0.788	0.774	0.589
SMU2	5.33 (1.71)	0.780	0.687			
SMU3	5.01 (1.12)	0.777	0.671			
SMU4	5.05 (1.32)	0.787	0.646			
Health awareness (HA)						
HA1	5.45 (1.56)	0.859	0.677	0.769	0.773	0.592
HA2	5.98 (1.45)	0.874	0.681			
HA3	5.39 (1.43)	0.881	0.675			
HA4	5.51 (1.39)	0.867	0.654			
Health care education (HED)						
HED1	5.23 (1.77)	0.719	0.651	0.783	0.789	0.678
HED2	5.22 (1.74)	0.775	0.629			
HED3	5.60 (1.64)	0.721	0.622			
HED4	5.41 (1.49)	0.739	0.637			
HED5	5.22 (1.15)	0.787	0.663			
HED6	5.11 (1.23)	0.772	0.654			
HED7	5.09 (1.43)	0.759	0.691			
HED8	5.32 (1.59)	0.768	0.609			

Hence, each standardized factor loading was statistically significant (P<.001), and values ranged from 0.719 to 0.881, thus validating adequate convergent validity as shown in Table 2. In addition, the average value explained (AVE) values and construct reliability were also more than their cut-off level of 0.5 and 0.7, respectively, and were statistically significant. Therefore, these measures confirm sufficient convergent validity.

A comparison of shared variance between factors with the average variance explained by individual factors was ensured for measuring the discriminant validity. The diagonal value should be greater than the non-diagonal value to confirm adequate discriminant validity [40]. Hence, results given in Table S3 in Multimedia Appendix 2indicates the correlation matrix of constructs, where non-diagonal elements are correlated among constructs and diagonal elements are the square root of average variance explained (AVE) by that construct and clearly explain that all three constructs differ.

Measurement Model

Using statistical software AMOS 26.0, the measurement model's confirmatory factor analysis (CFA) was performed [40,41]. This provided a passable model fit for the primary measurement model (CMIN=1220.23; CMIN/DF=1.98; GFI=0.72; AGFI=0.80; NFI=0.81; CFI=0.826; RMR=0.089; RMSEA=0.082); however, the number of indicators per item was large; for instance, the number of indicators for health care education were 8. Consequently, 2 items were deleted for further refinement to obtain an adequate model fit through the CFA of the measurement model. This refinement was conducted by deleting items one by one, based on their standardized residual; that is, that item was first deleted, which had a larger error variance than their measurement items. Each item was carefully reviewed before deleting it to ensure that, from a theoretical viewpoint, its error variance also seemed rational. The refinement and assessment process for every construct was first evident by Churchill [41]. Churchill defined this process as; "Though this application may be satisfactory during the early stages of research on a construct, the use of factor analysis in a confirmatory fashion would seem better at later stages." Furthermore, Gerbing and Anderson's [42] study also provided support to Churchill's argument and stated that; "to demonstrate that an explicit evaluation of Unidimensionality is accomplished with a confirmatory factor analysis of the individual measures as specified by a multiple-indicator measurement model. Coefficient alpha is important in assessing reliability but does not assess dimensionality. Although item-total correlations and exploratory factor analysis can provide useful preliminary

XSL•FO
RenderX

analyses, particularly in the absence of a sufficiently detailed theory, they do not directly assess unidimensionality. The reason is that a confirmatory factor analysis assesses the internal consistency and external consistency criteria of unidimensionality implied by the multiple-indicator measurement model.

Hence, this refinement process provided an adequate model fit (CMIN=357.62; the ratio of χ^2 to degree of freedom value (1.80) is remarkably less than its recommended value (5.0). CMIN/DF=1.80; GFI=0.90; AGFI=0.89; NFI=0:915; CFI=0:93; RMR=0:075; RMSEA=0:055) as shown in Table 3.

Table . Structural equation modeling fit indices for the confirmatory factor analysis model.

Fit indices	Cut-off criteria	Results obtained
Absolute fit indices		
Chi-square (*df*)	N/A[a]	357.62 (170)
χ^2/df (CMIN/DF)	<5.00	1.800
Root Mean Square Error of Approximation	<0.06	0.055
Goodness of Fit Index	>0.85	0.900
Adjusted Goodness of Fit Index	>0.85	0.890
Incremental fit indices		
Buntler-Bonett Normed Fit Index	>0.90	0.915
Comparative Fit Index	>0.93	0.930
Tucker Lewis Index	>0.90	0.941
Incremental Fit Index	>0.90	0.932
Parsimonious fit indices		
Parsimony Goodness-Fit Index	>0.50	0.798
Parsimony Normed Fit Index	>0.50	0.848

[a]N/A: not applicable.

Structural Model

The research hypotheses were tested through structural model estimation [40]. Therefore, a second subsample (n=170) was used and provided the adequate structural model fit (CMIN=488.6; CMIN/DF=1.85; GFI=0.861; AGFI=0.893; NFI=0.987; CFI=0.945; RMR=0:079; RMSEA=0.053) as shown in Table 4.

Table . Hypothesized structural model fit indices.

Fit index	Cut-off criteria	Results obtained
Absolute fit indices		
Chi-square (*df*)	N/A[a]	488.6 (170)
χ^2/df (CMIN/DF)	<5.00	1.850
Root Mean Square Error of Approximation	<0.06	0.053
Goodness of Fit Index	>0.85	0.861
Adjusted Goodness of Fit Index	>0.85	0.893
Incremental fit indices		
Buntler-Bonett Normed Fit Index	>0.90	0.987
Comparative Fit Index	>0.93	0.945
Tucker Lewis Index	>0.90	0.976
Incremental Fit Index	>0.90	0.943
Parsimonious fit indices		
Parsimony Goodness-Fit Index	>0.50	0.680
Parsimony Normed Fit Index	>0.50	0.609

[a]N/A: not applicable.

XSL•FO
RenderX

Moreover, the significance of the hypothesis is shown in Table 5. Testing of H1 revealed that social media usefulness has a significant effect on health awareness (H1: β=.669, *P*<.001), supporting the IDT proposed by Rogers. Testing of H2 also supplied similar significant outcomes to the current research model (H2: β=.557, *P*<.001), confirming the findins of prior work.

Table . Hypothesis testing results.

S.no.	Impact of	Impact on	Hypothesis	Path coefficient	*P* value	Outcome
1	SMU[a]	HA[b]	H1[c]	0.669[d]	.001	Significant impact
2	HA	HED[e]	H2[f]	0.557[d]	.001	Significant impact

[a]SMU: social media usefulness.

[b]HA: health awareness.

[c]H1: hypothesis 1.

[d]*P*<.001.

[e]HED: health care education.

[f]H2: hypothesis 2.

Discussion

Principal Findings

This study examined the impact of social media on health awareness and its role in health care education, using IDT as the theoretical framework in Pakistan. A conceptual model was developed to assess the mediating effect of health awareness on the relationship between social media usage and health care education. The hypothesis testing results demonstrated significant relationships, with social media positively influencing health awareness (H_1: path coefficient=0.669, *P*<.001) and health awareness positively impacting health care education (H_2: path coefficient=0.557, *P*<.001). These findings highlight the potential of social media as an effective tool for promoting health care education when strategically leveraged by health care stakeholders.

The study also revealed that younger and more educated health staff are more inclined to use social media, while older groups, though recognizing its usefulness, engage less actively. This demographic insight underlines the need for tailored engagement strategies to maximize participation. Limited research in developing countries, including Pakistan, has tested the application of IDT to health care professionals' adoption of social media for health care education. The standardized direct effects from hypothesis testing affirm the significant relationship between social media usefulness, health awareness, and health care education in the Pakistani context.

Given that developing countries are often classified as late adopters of technology, according to IDT, the findings emphasize the importance of region-specific social media content on platforms, such as Facebook, WhatsApp, or locally developed alternatives, for health care interventions. IDT further suggests that simplifying innovations can facilitate broader adoption, making it easier for health care professionals and the general population to engage with digital health initiatives. These insights highlight the need for cost-effective, user-centric solutions to enhance healthcare outcomes in Pakistan.

The government of Pakistan can also play a significant role in spreading health awareness. They can use social media to run campaigns on maternal health, child nutrition, and disease prevention. The health care professionals in private hospitals also indicated that live sessions and webinars on platforms like Facebook Live and Instagram Live to educate the patients and hospital staff on various health issues and answer their real-time questions have been very useful.

Using social media for health care education in Pakistan could be a game-changer, surpassing the expectations of spreading awareness regarding health needs in a particular community. This potential is not just theoretical; it has been proven in developed countries, such as the US, Jordan, the United Kingdom, and Europe, where social media was used during the COVID-19 pandemic to combat the pandemic and other notable diseases [15]. Moreover, social media platforms have been used to promote telemedicine services, enabling health care professionals to connect with patients remotely for consultations, follow-ups, and monitoring of chronic conditions [16,30].

Overall, this study validates the use of social media to improve public health care education through public awareness. Therefore, the study concludes that health care professionals should use social media tools to inform the wider public and address healthcare issues in Pakistan.

Conclusion

The findings of this study underscore the significant role that social media plays in enhancing health care awareness and education among various populations in developing countries like Pakistan. With their wide reach and interactive features, social media platforms have proven to be effective tools for disseminating health-related information and engaging individuals in meaningful health conversations. Social media breaks down geographical barriers, allowing health information to reach a global audience instantly. This broadens the scope of health care awareness campaigns and makes information accessible to a diverse demographic, including those in remote or underserved areas. The interactive nature of social media facilitates active engagement between health care providers and the public. This interaction fosters a sense of community, encourages the sharing of personal health experiences, and allows for immediate feedback and clarification of health information. Also, social media platforms enable the

XSL•FO
RenderX

dissemination of personalized health information tailored to specific audiences' needs. Moreover, the real-time nature of these platforms ensures that information is up to date, which is crucial during health emergencies or outbreaks.

Health care education through social media can help understand how to prevent common diseases such as malaria, tuberculosis, hepatitis, and polio through vaccination, sanitation, and hygiene in Pakistan. Awareness campaigns can encourage healthier lifestyles, reducing the incidence of chronic diseases such as diabetes, hypertension, and heart disease. Educating women about prenatal and postnatal care, safe delivery practices, and child nutrition can significantly lower maternal and infant mortality rates. Awareness through social media about proper nutrition, immunization, and hygiene practices helps improve children's health and development in the rural outskirts of Pakistan. Also, awareness about communicable diseases and their transmission can help in preventing outbreaks of diseases like HIV/AIDS, tuberculosis, and waterborne illnesses.

Governmental policies provide a framework for regulating the content shared on social media, ensuring that the information disseminated is accurate, reliable, and in line with public health guidelines. Information shared by government-endorsed social media accounts is more likely to be trusted by the public. Official policies lend credibility and legitimacy to health messages, increasing the likelihood of public acceptance and compliance. All active organizations (govt and private) working in the health industry of Pakistan seek prior permission or ethical clearance to initiate any new work. Without government support, the organizations lack the confidence to spread and advertise any health-related content due to fear that they may be subject to local resistance.

Despite its benefits, the use of social media in health care awareness also presents challenges, such as the spread of misinformation and the need for privacy and data protection. Health care organizations must implement strategies to verify information and ensure the credibility of the content shared. Hence, the study suggests that Pakistan's health care organizations should continue leveraging social media to enhance health communication strategies and education. Ongoing research is also needed to explore new ways to maximize social media's benefits while mitigating its risks.

In conclusion, social media is a powerful tool for health care awareness, offering unprecedented opportunities to reach and engage with a wide audience. By harnessing its potential in developing economies, health care organizations can improve public health literacy, promote healthy behaviors, and ultimately contribute to better health outcomes. Collaborative efforts between health care professionals, social media platforms, and policymakers will be crucial in leveraging social media's full potential for public health advancement.

Future Research

Future research should move beyond compliance with existing social media policies to explore strategies and variables that demonstrate tangible improvements in the health care sector. This could include investigating how social media can influence enhanced patient outcomes through better communication and engagement, broaden health awareness campaigns by targeting diverse and underserved populations, and improve access to health care services through innovative digital solutions. Additionally, longitudinal studies examining the long-term impact of social media interventions on public health indicators would provide valuable insights. Identifying and analyzing these variables will contribute to a more comprehensive understanding of social media's transformative potential in health care beyond mere policy adherence.

Acknowledgments

The authors would like to thank Mr Nabil Ahmed, Director at Integral Global Consulting LLC, Tucker, GA, US for providing the necessary resources and facilities that enabled the successful execution of this research. The authors also want to extend their special thanks to all the respondents and hospital management in Lahore, Pakistan, for their feedback and the collaborative environment that fostered productive discussions and the exchange of ideas.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Questionnaire.
[DOCX File, 19 KB - humanfactors_v12i1e65745_app1.docx]

Multimedia Appendix 2
Additional tables.
[DOCX File, 17 KB - humanfactors_v12i1e65745_app2.docx]

References

1. Shaeen SK, Tharwani ZH, Bilal W, Islam Z, Essar MY. Maternal mortality in Pakistan: challenges, efforts, and recommendations. Ann Med Surg (Lond) 2022 Sep;81:104380. [doi: 10.1016/j.amsu.2022.104380] [Medline: 36042926]

XSL•FO
RenderX

2. Hanif M, Khalid S, Rasul A, Mahmood K. Maternal mortality in rural areas of Pakistan: challenges and prospects. Rural Health 2021;7. [doi: 10.5772/intechopen.96934]
3. Soofi SB, Khan A, Kureishy S, et al. Determinants of stunting among children under five in Pakistan. Nutrients 2023 Aug 7;15(15):3480. [doi: 10.3390/nu15153480] [Medline: 37571417]
4. Bul K, Holliday N, Bhuiyan MRA, Clark CCT, Allen J, Wark PA. Usability and preliminary efficacy of an artificial intelligence-driven platform supporting dietary management in diabetes: mixed methods study. JMIR Hum Factors 2023 Aug 9;10:e43959. [doi: 10.2196/43959] [Medline: 37556198]
5. Siddiqi DA, Ali RF, Shah MT, et al. Evaluation of a mobile-based immunization decision support system for scheduling age-appropriate vaccine schedules for children younger than 2 years in Pakistan and Bangladesh: lessons from a multisite, mixed methods study. JMIR Pediatr Parent 2023 Feb 17;6:e40269. [doi: 10.2196/40269] [Medline: 36800221]
6. Mo Y, Pham TM, Lim C, et al. The effect of hand hygiene frequency on reducing acute respiratory infections in the community: a meta-analysis. Epidemiol Infect 2022;150:e79. [doi: 10.1017/S0950268822000516]
7. Eaton MC, Probst YC, Smith MA. Characterizing the discourse of popular diets to describe information dispersal and identify leading voices, interaction, and themes of mental health: social network analysis. JMIR Infodemiology 2023 May 5;3:e38245. [doi: 10.2196/38245] [Medline: 37159259]
8. Allegranzi B, Tartari E, Pittet D. 'Seconds SAVE LIVES – clean your hands': the 5 May 2021 World Health Organization's SAVE LIVES: Clean Your Hands campaign. Int J Infect Control 2021;17(1):21418. [doi: 10.3396/ijic.v17.21418]
9. Subramanyam C, Becker A, Rizzo J, Afzal N, Nong Y, Sivamani R. Visibility of board-certified dermatologists on TikTok. JMIR Dermatol 2024 Jan 5;7:e46085. [doi: 10.2196/46085] [Medline: 38180786]
10. Griffis HM, Kilaru AS, Werner RM, et al. Use of social media across US hospitals: descriptive analysis of adoption and utilization. J Med Internet Res 2014 Nov 27;16(11):e264. [doi: 10.2196/jmir.3758] [Medline: 25431831]
11. Global social media statistics. DATAREPORTAL. URL: https://datareportal.com/social-media-users [accessed 2025-03-25]
12. Terrasse M, Gorin M, Sisti D. Social media, e-health, and medical ethics. Hastings Cent Rep 2019 Jan;49(1):24-33. [doi: 10.1002/hast.975] [Medline: 30790306]
13. Smith MA. NodeXL: simple network analysis for social media. Presented at: Proceedings of the 2013 International Conference on Collaboration Technologies and Systems (CTS); May 20-24, 2013; San Diego, CA, USA URL: https://ieeexplore.ieee.org/document/6567211
14. Bazzano AN, Patel T, Nauman E, Cernigliaro D, Shi L. Optimizing telehealth for diabetes management in the deep south of the United States: qualitative study of barriers and facilitators on the patient and clinician journey. J Med Internet Res 2024 Jan 1;26:e43583. [doi: 10.2196/43583] [Medline: 37976468]
15. Balki E, Holland C, Hayes N. Use and acceptance of digital communication technology by older adults for social connectedness during the COVID-19 pandemic: mixed methods study. J Med Internet Res 2023 Aug 2;25:e41535. [doi: 10.2196/41535] [Medline: 37531187]
16. Santos DS, Batistelli CRS, Lara MMDS, Ferreira EDS, Moreira TR, Cotta RMM. The effectiveness of the use of telehealth programs in the care of individuals with hypertension and, or diabetes mellitus: systematic review and meta-analysis. Diabetol Metab Syndr 2022 May 28;14(1):76. [doi: 10.1186/s13098-022-00846-5] [Medline: 35643495]
17. Lenoir P, Moulahi B, Azé J, Bringay S, Mercier G, Carbonnel F. Raising awareness about cervical cancer using Twitter: content analysis of the 2015 #SmearForSmear campaign. J Med Internet Res 2017 Oct 16;19(10):e344. [doi: 10.2196/jmir.8421] [Medline: 29038096]
18. Niederdeppe J, Farrelly MC, Haviland ML. Confirming "truth": more evidence of a successful tobacco countermarketing campaign in Florida. Am J Public Health 2004 Feb;94(2):255-257. [doi: 10.2105/ajph.94.2.255] [Medline: 14759936]
19. Maher CA, Lewis LK, Ferrar K, Marshall S, De Bourdeaudhuij I, Vandelanotte C. Are health behavior change interventions that use online social networks effective? A systematic review. J Med Internet Res 2014 Feb 14;16(2):e40. [doi: 10.2196/jmir.2952] [Medline: 24550083]
20. Bauman A, McNeil N, Nicholson M, et al. Impact of the first year of the "This girl can" physical activity and sport mass media campaign in Australia. BMC Public Health 2023 Feb 15;23(1):333. [doi: 10.1186/s12889-023-15091-2] [Medline: 36793043]
21. Booth RG, Allen BN, Bray Jenkyn KM, Li L, Shariff SZ. Youth mental health services utilization rates after a large-scale social media campaign: population-based interrupted time-series analysis. JMIR Ment Health 2018 Apr 6;5(2):e27. [doi: 10.2196/mental.8808] [Medline: 29625954]
22. Stronks K, van de Mheen HD, Mackenbach JP. A higher prevalence of health problems in low income groups: does it reflect relative deprivation? J Epidemiol Community Health 1998 Sep;52(9):548-557. [doi: 10.1136/jech.52.9.548] [Medline: 10320855]
23. Song P. The Ice Bucket Challenge: the public sector should get ready to promptly promote the sustained development of a system of medical care for and research into rare diseases. Intractable Rare Dis Res 2014 Aug;3(3):94-96. [doi: 10.5582/irdr.2014.01015] [Medline: 25364651]
24. Khan SJ, Asif M, Aslam S, Khan WJ, Hamza SA. Pakistan's healthcare system:a review of major challenges and the first comprehensive universal health coverage initiative. Cureus 2023 Sep;15(9):e44641. [doi: 10.7759/cureus.44641] [Medline: 37799252]

XSL•FO
RenderX

25. Naz F, Asghar K. Patients satisfaction and healthcare delivery system in public sector hospitals. Pak Soc Sci Rev 2023 Jul;7(3):352-362. [doi: 10.35484/pssr.2023(7-III)28]
26. PMA expresses grave concern over deteriorating health indicators. BOL News. URL: https://www.pmacentre.org.pk/new/Newscuttings%2010-06-2023%20-%20Health%20Indicators.pdf [accessed 2025-03-25]
27. Saleem A, Rizvi WR, Farrukh M, Saleem M. Influence of radio programs on development of small & medium enterprises in central Punjab Pakistan. PalArch's J Archaeol Egypt/ Egyptol 2021;18(10):116-140.
28. Global Social Media Statistics. Pakistan. URL: https://datareportal.com/reports/digital-2024pakistan#:~:text=Pakistan%20was%20home%20to%2071.70,percent%20of%20the%20total%20population [accessed 2025-03-25]
29. Shaikh BT, Ali N. Universal health coverage in Pakistan: is the health system geared up to take on the challenge? Global Health 2023 Jan 12;19(1):4. [doi: 10.1186/s12992-023-00904-1] [Medline: 36635698]
30. Bilal W, Qamar K, Siddiqui A, Kumar P, Essar MY. Digital health and telemedicine in Pakistan: Improving maternal healthcare. Ann Med Surg (Lond) 2022 Sep;81:104425. [doi: 10.1016/j.amsu.2022.104425] [Medline: 36046714]
31. Ittefaq M, Seo H, Abwao M, Baines A. Social media use for health, cultural characteristics, and demographics: a survey of Pakistani millennials. Digit Health 2022;8:20552076221089454. [doi: 10.1177/20552076221089454] [Medline: 35401998]
32. Ho KF, Chang PC, Kurniasari MD, Susanty S, Chung MH. Determining factors affecting nurses' acceptance of a care plan system using a modified technology acceptance model 3: structural equation model with cross-sectional data. JMIR Med Inform 2020 May 5;8(5):e15686. [doi: 10.2196/15686] [Medline: 32369033]
33. Dearing JW. Applying diffusion of innovation theory to intervention development. Res Soc Work Pract 2009 Sep 1;19(5):503-518. [doi: 10.1177/1049731509335569] [Medline: 20976022]
34. Zhang X, Yu P, Yan J, Ton A M Spil I. Using diffusion of innovation theory to understand the factors impacting patient acceptance and use of consumer e-health innovations: a case study in a primary care clinic. BMC Health Serv Res 2015 Feb 21;15:71. [doi: 10.1186/s12913-015-0726-2] [Medline: 25885110]
35. Hair J, Hult G, Ringle C, Sarstedt M. A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM): Sage Publications; 2013.
36. Verdam MGE, Oort FJ, Sprangers MAG. Using structural equation modeling to investigate change and response shift in patient-reported outcomes: practical considerations and recommendations. Qual Life Res 2021 May;30(5):1293-1304. [doi: 10.1007/s11136-020-02742-9] [Medline: 33550541]
37. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manage Sci 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
38. Venkatesh V. Determinants of perceived ease of use: integrating control, intrinsic motivation, and emotion into the technology acceptance model. Information Systems Research 2000 Dec;11(4):342-365. [doi: 10.1287/isre.11.4.342.11872]
39. Ho KF, Ho CH, Chung MH. Theoretical integration of user satisfaction and technology acceptance of the nursing process information system. PLoS One 2019;14(6):e0217622. [doi: 10.1371/journal.pone.0217622] [Medline: 31163076]
40. Hair J. Multivariate Data Analysis: Pearson Prentice Hall; 2006. URL: https://drnishikantjha.com/papersCollection/Multivariate%20Data%20Analysis.pdf [accessed 2024-08-21]
41. Churchill GA. A paradigm for developing better measures of marketing constructs. Journal of Marketing Research 1979 Feb;16(1):64-73. [doi: 10.1177/002224377901600110]
42. Anderson JC, Gerbing DW. The meffect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factor analysis. Psychometrika 1984 Jun;49(2):155-173. [doi: 10.1007/BF02294170]

Abbreviations

AGFI: Adjusted Goodness of Fit Index
AVE: average value explained
CFA: confirmatory factor analysis
CFI: Comparative Fit Index
CMIN: chi-square minimum
DF: degrees of freedom
GFI: Goodness of Fit
H1: hypothesis 1
H2: hypothesis 2
IDT: Innovation Diffusion Theory
NFI: Buntler-Bonett Normed Fit Index
PMA: Pakistan Medical Association
RMR: Root Mean Square Residual
RMSEA: Root Mean Square Error of Approximation
SEM: structural equation modeling
SMC: squared multiple correlation

XSL•FO
RenderX

Edited by A Kushniruk; submitted 24.08.24; peer-reviewed by A Bilal, M Khairulbahri, M Javed; revised version received 09.02.25; accepted 10.02.25; published 07.04.25.

Please cite as:
Munir MM, Ahmed N
Using Social Media Platforms to Raise Health Awareness and Increase Health Education in Pakistan: Structural Equation Modeling Analysis and Questionnaire Study
JMIR Hum Factors 2025;12:e65745
URL: https://humanfactors.jmir.org/2025/1/e65745
doi:10.2196/65745

© Malik Mamoon Munir, Nabil Ahmed. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 7.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Patients' and Physicians' Experience With and Acceptability of a Telemedicine Cabin: Mixed Methods Study

Caroline Villela Galvão de França[*], MD; Paola Boaro Segalla[*], MD; Felipe Sebastião de Assis Reis[*], MD, MSc; José Ricardo Silveira Pereira[*], MBA; Alexandre Oliveira de Mattos[*], BSc; Roberta de Moura Ferron[*], BSc; Cleyton Zanardo de Oliveira[*], MSc; Jéssica Bassani Borges[*], MBA, PhD; Lilian Quintal Hoffmann[*], MBA; Edmundo Di Giaimo Caboclo[*], MD

Department Occupational Medicine, Hospital BP (A Beneficência Portuguesa de São Paulo), Rua Maestro Cardin, 769, São Paulo, Brazil

[*]all authors contributed equally

Corresponding Author:
Edmundo Di Giaimo Caboclo, MD
Department Occupational Medicine, Hospital BP (A Beneficência Portuguesa de São Paulo), Rua Maestro Cardin, 769, São Paulo, Brazil

Abstract

Background: Telemedicine represents an essential tool with the potential to reduce health costs, thus avoiding patient displacement and improving patient care outcomes, positioning it as a significant social technology.

Objective: This study aims to analyze the implementation of a telehealth cabin at BP Hospital (A Beneficência Portuguesa de São Paulo), focusing on the evaluation of the experiences of both patients and health care professionals, as well as the acceptability of this tool.

Methods: A mixed methods study was conducted with 229 participants, divided into 2 phases. The first phase involved 40 apparently healthy individuals to assess the usability, experience, and satisfaction of this group for the later safe application in the group with clinical complaints. The second phase included 189 participants, with complaints to assess the usability, experience, and satisfaction of patients and doctors. In both phases, participants completed screening questionnaires (to assess the eligibility criteria), a socioeconomic demographic questionnaire before using the cabin, and a questionnaire including the System Usability Scale and the Net Promoter Score (NPS) after using the cabin.

Results: The data analysis of the first phase showed high acceptance of the telehealth cabin, which supported the progression to the second phase. In the second phase, a high usability score was observed among participants with clinical complaints (mean System Usability Scale score of 85.97, SD 15.50) and a high favorability rating (NPS score of 9.4). Health care professionals also reported favorable results, with a usability score of 67.8 and an NPS of 8.0.

Conclusions: The results of this study reinforce the potential for scaling up this practice based on usability outcomes, and highlight its relevance for the development of public policies aimed at expanding access to quality health care in Brazil. This approach improves the interaction of patients with the health care system, while providing professionals with an extended view of clinical conditions through integrated devices, particularly in areas with limited access to medical care.

(JMIR Hum Factors 2025;12:e55430) doi:10.2196/55430

KEYWORDS

telemedicine cabin; telehealth; teleservice; e-health; connected offices

Introduction

Health is a fundamental right of the population and must be guaranteed to everyone. It is even included in the Universal Declaration of Human Rights of 1948, in Article 25, which defines that everyone has the right to a standard of living that ensures health, well-being, and health care [1]. Although access to health is often presented as a goal in the health policy, there are several challenges to ensure that right into the Brazilian reality, such as inequalities related to the access and use of health services, discontinuities in the geographical distribution of services, especially of medium and high complexity, in addition to the articulation between health systems of several care levels [2,3].

Among the various proposals developed to improve health care at the national level, the use of technological and telecommunication resources for the exchange of information across different levels of health care, between health professionals, and between doctors and patients, has gained prominence.

XSL•FO
RenderX

According to the Ministry of Health, telehealth uses information and communication technologies to promote the expansion and improvement of medical services. Telemedicine is included in telehealth, related only to remote care through technology. According to the World Health Organization, telemedicine is defined as the "delivery of health services where distance to health professionals is a critical factor, through the use of information and communication technologies to the exchange of valid information for the diagnosis, treatment, and prevention of disease and injuries, research and assessment, and to the continuing education of the providers of healthcare and patients."

Telehealth has been used in Brazil since the 1990s, emerging in a decentralized and fragmented way in the health sites [4]. Telemedicine has the potential to reduce health costs, avoid patient displacement, and even improve patient care outcomes [5-7]. Although many studies are being published, especially as a consequence of the COVID-19 pandemic, the implementation of telemedicine can still be challenging, mainly due to technological barriers and poor computer knowledge, and even due to resisting change and the patient's education level [8]. Despite the increasing interest in telehealth, a significant part of the research focuses on the technological aspects rather than on the tool's adoption and acceptance.

One of the limitations of telemedicine is the lack of possibility to monitor the patient throughout medical care, and some telemedicine platform initiatives have been developed for this purpose. However, today, they provide more complete solutions through a digital health ecosystem, which offers medical instruments to be used by the patient, such as a temperature sensor, stethoscope, dermatoscope, oximeter, otoscope, sphygmomanometer, scale, and height sensor. These new functionalities aim to create resources that can improve even more access and affect the quality of life by bringing doctors and patients closer together.

This research aimed to assess the human aspects of the experience and usability when using telehealth cabins with built-in medical devices, thus expanding the diagnostic and interventional capacity of health professionals in clinical, diagnostic, and interventional capacity practice in its remote format.

To this end, the socioeconomic and demographic profile of the population, the experience and usability of the population in the cabin, the doctor's experience in providing care through telehealth cabins, and compliance in handling the equipment available in the self-examination cabin were evaluated.

Methods

Recruitment

This is a mixed methods study using (1) the System Usability Scale (SUS), (2) the Net Promoter Score (NPS), and (3) a thematic analysis based on participant' perceptions. The aim was to assess the usability and favorability of health care provided in a connected cabin for symptomatic and asymptomatic employees of the BP Hospital (Beneficência Portuguesa de São Paulo).

Ethical Considerations

Ethics approval (CAAE: 58070622.9.0000.5483) for this study was provided by the institutional review board of the BP Hospital (Beneficência Portuguesa de São Paulo) on January 31, 2023. All eligible patients signed an informed consent form for this research protocol.

The data was collected on the RedCap platform, which has access control to ensure the security of the data collected in the research. In addition, all data was anonymized to ensure the privacy of the participants.

Sample Size

The Binomial test was used to calculate the sample size for both parts of the study. For the sampling of asymptomatic individuals (phase 1), an 80% test power was stipulated, a 70% acceptance rate of the cabin under the null hypothesis, and a 90% expected in the sampling rate were specified, resulting in 41 asymptomatic patients.

For the sampling of symptomatic individuals (phase 2), we opted for a conservative scenario. Thus, we set a power of 90%, an acceptance rate of 80% under a null hypothesis, and a 90% expected proportion in the sampling, resulting in 169 patients. Anticipating the possibility of a 12% sample loss, the researchers decided to increase the sampling to 190 symptomatic patients.

Inclusion and Exclusion Criteria

Phase 1: Asymptomatic Patients

The study's first phase involved 40 apparently healthy individuals with no clinical complaints or decompensated chronic diseases (exclusion criteria for the first phase) to assess the safety and usability of the cabin.

During this phase, all hospital staff received an email with information about the study, inviting asymptomatic individuals to participate. Interested individuals could schedule an appointment by phone to use the cabin.

Phase 2: Symptomatic Patients

The second phase included 189 individuals who could be with clinical complaints but did not meet the exclusion criteria, which were: inability to understand and answer the screening questionnaire; patients with signs and clinical symptoms indicating the need for urgent medical care, such as cardiac arrest, shock of any origin, reduced level of consciousness, focal neurological signs, epileptic seizure, chest pain, deep wound, and heavy bleeding.

During this phase, individuals who sought medical care at the institution's employee health center and who met the inclusion criteria were invited to participate. Those who agreed to participate were enrolled in the study.

Telemedicine Cabin

The telemedicine cabin Diagnostica, manufactured in Argentina, was used. It is made of plastic and fiberglass and measures 250×150×230cm dimensions. It has a forced air ventilation and filtering system, environmental lighting with intensity and color control, and an audio and video system for videoconferences (Figure 1).

XSL•FO
RenderX

Figure 1. Telemedicine cabin at the location of the research project.



The participant used his cell phone or the device available in the cabin to receive care and to interact with the person through the app developed for Android and IOS. The patient was accommodated in the cabin and selected the self-examination devices according to their needs.

The following devices were available in the cabin: an otoscope, a digital dermatoscope, and a high-definition camera, all manufactured by Firefly Global in 2021; a stethoscope Riester manufactured in 2021; and an oxyhemoglobin saturation monitor (oximeter), an electrocardiography, a heart rate and temperature monitor, and a sphygmomanometer, models PM6100, all manufactured by Berry, 2021. The patient was instructed to use the app on his cell phone, and the light signals in the cabin were added to a video display to operate the devices for the self-examination.

The cabin recorded the measurements taken by the available devices. The patient had the option, at their discretion, to join the medical consultation at any time during the use of the cabin, even if he did not use the devices. During the synchronous service, the doctor could request that the patient use any of the medical devices to obtain additional clinical information.

Doctors were recruited through direct invitations sent to professionals associated with the hospital. Selection criteria included previous experience with similar technologies and availability to participate in all stages of the study.

Data between the patient and the doctor was transmitted via the Doc24 telehealth platform, implanted in the data processing centers, and connected to the internal service network.

XSL•FO
RenderX

Study Procedure

Before starting the service, a nursing technician applied the informed consent form. After the participant's agreement, a screening questionnaire focusing on the inclusion and exclusion criteria and a questionnaire to collect demographic and socioeconomic data before using the cabin were administered. Individuals who met the exclusion criteria or had acute alterations requiring face-to-face assessment were sent to the Employee Health Center for immediate medical care.

The cabin was equipped with a Wi-Fi network and the patient should download an app to access the cabin and the interface (application) which was customized for the project.

During the appointment, the patient, seated in the cabin, can choose between selecting devices for self-examination or opting for a consultation with a doctor. At the patient's discretion, they have the option to join the medical consultation at any time during the use of the cabin, even without having used the devices.

If the participant chooses to attend a medical consultation during the service, the doctor may ask the patient to use some of the equipment to collect clinical information, or to reuse some of the equipment due to observing poor quality results initially measured by the patient.

Equipment-assisted measurements were available for medical assessment, if required, or later, in the outpatient treatment.

At the end of the service, the research participant was invited to answer the usability questionnaires translated and validated including the SUS, which is a method to measure the usability of several products and services through 10 questions answered on a scale of 1 to 5, where 1 is total disagreement and 5 is total agreement [9], and the NPS, which aims to analyze the patient's experience and satisfaction [10].

After assigning the NPS score, participants were asked to explain their ratings. Inductive thematic analysis was used to identify a rating [11]. Two authors carefully read each response individually and grouped the responses. Any disagreement was discussed with the first author, and then the response was labeled into 6 categories: "service," "usability," "difficulty of use," "innovation," "technology resistance," and "unspecific." The final step was to group the responses into potential themes.

In order to check the usability of health professionals, the participating doctors also completed the same questionnaires at the end of the daily sessions.

Statistical Analysis

Study data were collected and managed by an electronic data capture tool, REDCap (Research Electronic Data Capture), and hosted on the BP server [12,13]. REDCap is a secure software web-based platform designed to support data collection for research studies, providing (1) an intuitive interface for the validated data collection, (2) audit trails to track data manipulation and export procedures, (3) automated export procedures for further data downloads to standard statistical packages, and (4) procedures for data integration and interoperability with external sources. Once collected, data were described considering the mean and SD for the numerical variables and the absolute and relative frequencies to the categorical variables. SUS scores were interpreted using the Sauro and Lewis Curved Grading Scale [14].

Patients were classified into 2 groups according to the SUS scale score "Acceptable usability" (SUS ≥68) and "Usability issues" (SUS <68), and the percentage of patients was compared with 70% (for the Symptomatic sample), and 80% (for the Symptomatic sample) by the Binomial test. The cut-off point of 68 was selected because, despite variability in acceptability, this is the threshold at which the technology is considered to have acceptable usability.

The chi-square test or Fisher exact test was used to compare the groups' qualitative characteristics, and the Mann-Whitney Test was used for the numeric characteristics. The significance level of 0.05 was used throughout the study, and the analyses were performed using SPSS software (version 25; IBM corporation). Graphs were generated using R (version 4.3.2; R studio Team) software and the *ggplot2* package.

Results

Overview

The study was divided into 2 phases: the first occurred from March 6, 2023, to March 16, 2023, and 40 asymptomatic patients were included. The second occurred from March 17, 2023, to June 1, 2023, and 189 symptomatic patients were included.

Assessment of Asymptomatic Patients

The demographic distribution of the asymptomatic patients is shown in Table 1. Most were female, White, married, university-educated, and of higher socioeconomic status.

Concerning the use of the cabin, most of the patients requested support to use the cabin at some point during use. Only 3 patients described the reasons for requesting assistance as difficulties in using the equipment. The appointment had a duration between 10 and 52 minutes.

Upon assessing the satisfaction (Figure 2), we observed that the SUS had a mean value corresponding to Grading A+ (Table 2). When defining the cut-off point for SUS, we identified that only 5 (12.5%) patients had a score of <68 points, whereas 35 (87.5%) had a score of ≥68 points. Thus, the acceptability of 87.5% was statistically different from 70% (P=.009) and had a 95% CI 75.5%-100%, justifying the acceptability of the asymptomatic sample and allowing the study to proceed with the sampling of symptomatic patients.

XSL•FO
RenderX

Table . Distribution of demographic variables for the sampling of asymptomatic patients.

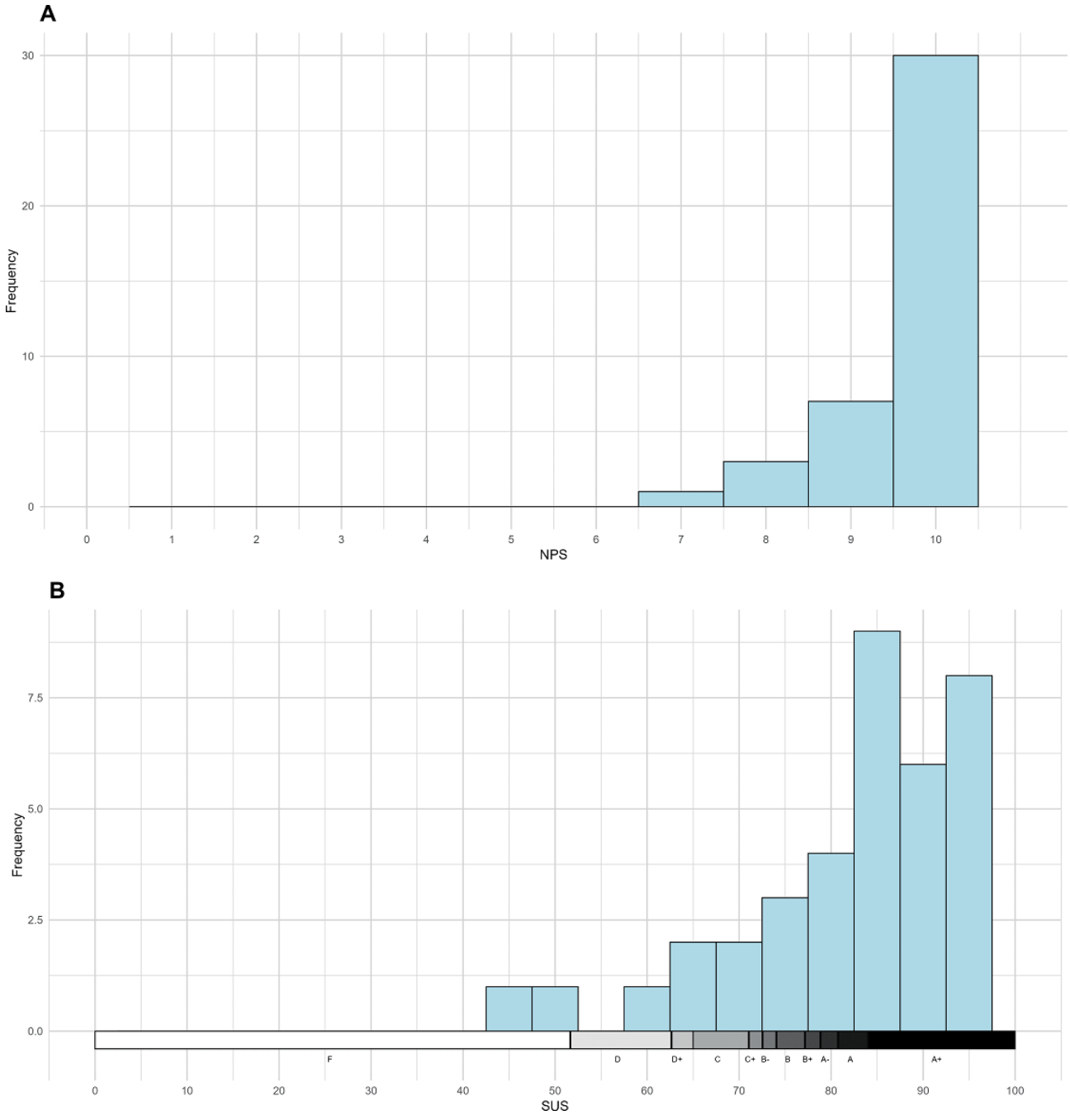
	Values (N=40)
Sex, n (%)	
Male	15 (38)
Female	25 (63)
Ethnicity and race, n (%)	
Asian	2 (5)
White	24 (60)
Mixed	8 (20)
Black	6 (15)
Education, n (%)	
High school	6 (15)
College	34 (85)
Marital status, n (%)	
Single	15 (38)
Married, common-law marriage, or partnered	24 (60)
Judicially separated or divorced	1 (3)
Monthly household income, n (%)	
Up to 3 minimum wages	8 (20)
4 - 6 minimum wages	6 (15)
7 - 11 minimum wages	13 (33)
Above 11 minimum wages	13 (33)
Care function, n (%)	
No	36 (90)
Yes	4 (10)
Has the patient requested support to use the cabin outside the previously set times?, n (%)	
No	31 (78)
Yes	9 (23)
At what point did the patient request medical advice?, n (%)	
Medical advice requested after the tests	24 (60)
Medical advice requested at the beginning; tests performed upon medical orientation	7 (18)
Did not request medical orientation	9 (23)
Age (years), mean (SD)	37.9 (10.9)
Appointment duration (minutes), mean (SD)	27.0 (10.5)
NPS[a], mean (SD)	9.6 (0.7)
SUS[b], mean (SD)	84.9 (13.0)

[a]NPS: Net Promoter Score.

[b]SUS: System Usability Scale.

XSL•FO
RenderX

Figure 2. Distribution for asymptomatic patients for the variables: (A) “How likely is it that you would recommend it to a friend or colleague.” (B) Final SUS Score upon graduation using the Curved Grading Scale. NPS: Net Promoter Score; SUS: System Usability Scale.



XSL•FO
RenderX

Table . Distribution of patients by Curved Grading Scale grades of asymptomatic and symptomatic groups.

SUS score range	Grading	Asymptomatic (n=40), n (%)	Symptomatic (n=189), n (%)
84.1 - 100	A+	26 (65)	128 (68.1)
80.8 - 84.0	A	3 (7.5)	8 (4.3)
78.9 - 80.7	A−	1 (2.5)	3 (1.6)
77.2 - 78.8	B+	1 (2.5)	11 (5.9)
74.1 - 77.1	B	2 (5)	8 (4.3)
72.6 - 74.0	B−	0 (0)	0 (0)
71.1 - 72.5	C+	1 (2.5)	2 (1.1)
65.0 - 71.0	C	3 (7.5)	11 (5.9)
62.7 - 64.9	C−	0 (0)	0 (0)
51.7 - 62.6	D	2 (5)	6 (3.2)
0.0 - 51.6	F	1 (2.5)	11 (5.9)

Assessment of Symptomatic Patients

The demographic distribution of the symptomatic patients is shown in Table 2. Most were female, white, single, had a college degree, had a household income up to 3 minimum wages, did not perform care functions, and had a mean age was 35.1 years.

Only 5 (2.7%) patients requested support to use the cabin outside the previously established times. The 3 reasons described for requesting support were inconsistencies in performing the cabin and difficulties in using the instruments (otoscope and oximeter) available in the cabin.

The most of participants requested medical orientation and performed the measurement under medical orientation. The duration of the appointment ranged from 2 to 58 minutes. Regarding satisfaction (Figure 3), the variable SUS mean corresponds to Grading A+ (Table 2). When assessing the participants' score in relation to the cut-off point of the SUS scale, we found that only 21 (11.1%) patients had scored <68 points, while 167 (88.8%) had scored ≥68 points. Thus, 88.8% acceptability was statistically different from 80% (P=.001) and had a 95% CI 84.3%-100%.

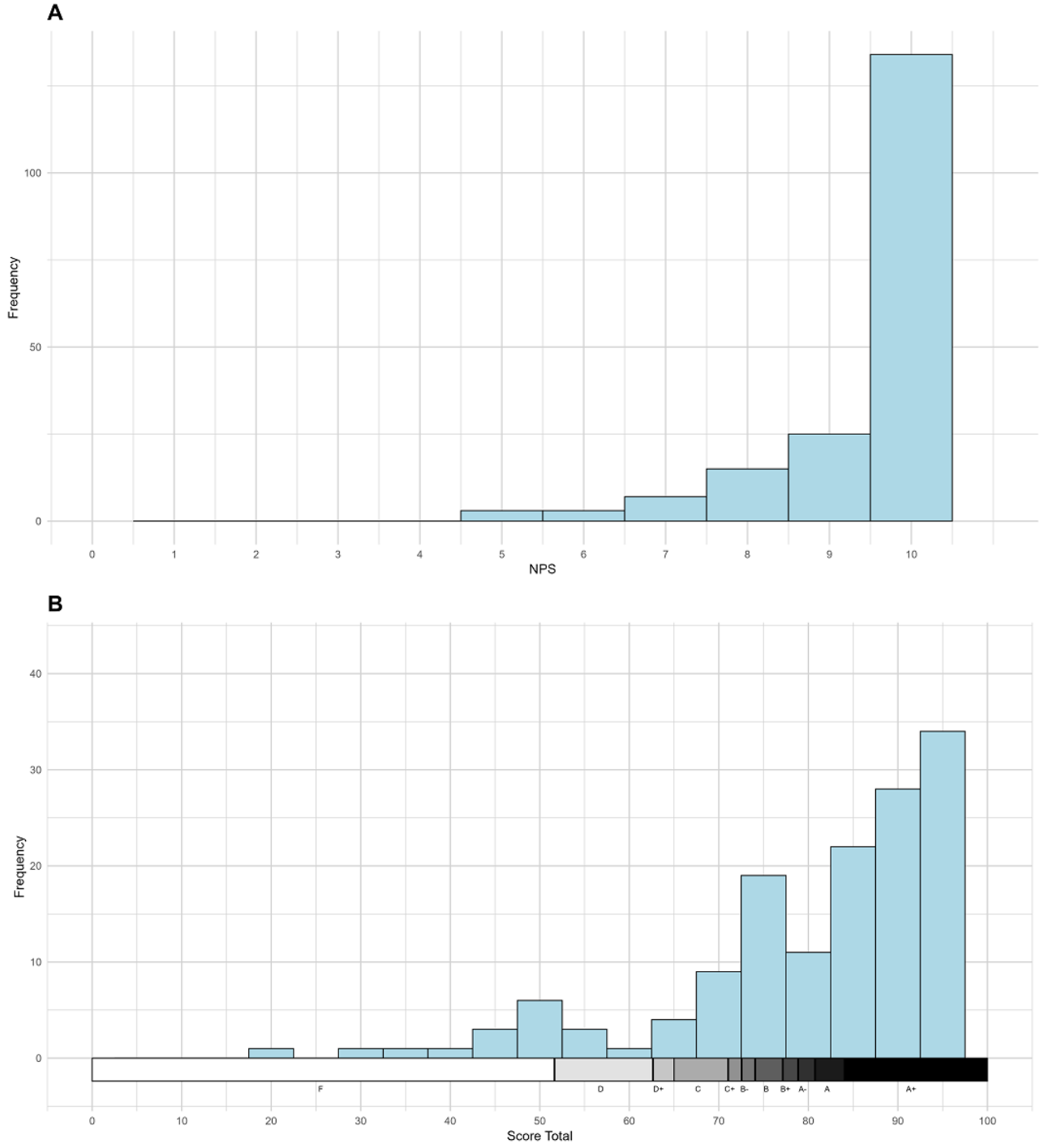
The 21 patients with a score of 68 or less on the SUS were considered in the group "Usability issues" at the cabin. Eleven (52.4%) were female, 13 (61.9%) self-declared White, 11 (52.4%) had a college degree, 13 (61.9%) were single, 5 (45.5%) were married, 7 (63.6%) earned up to 3 minimum wages, 8 (72.7%) did not carry out welfare activities, and had a mean age of 32.64 years (SD 8.43).

Regarding the use of the cabin by these participants, none of them required assistance from the nursing technician, 8 (38.1%) participants requested medical instruction to perform the tests, 7 (33.3%) participants requested medical orientation and made the tests under medical orientation, while 6 (28.6%) participants did not request medical orientation.

The mean appointment duration was 20.38 (SD 10.29) minutes, ranging from 6 to 40 minutes. When comparing the characteristics of patients with SUS ≥68 related to SUS <68, all of them were not statistically significant (attaining a value equal to or higher than P=.05) (Table 3), meaning that we did not identify characteristics that could have influenced the difference in the SUS score.

XSL•FO
RenderX

Figure 3. Distribution for asymptomatic patients for the variables: (A) “How likely is it that you would recommend it to a friend or colleague.” (B) Final SUS Score upon graduation using the Curved Grading Scale. NPS: Net Promoter Score; SUS: System Usability Scale.



XSL•FO
RenderX

Table . Distribution of the variables for sampling symptomatic patients and comparison System Usability Scale (SUS) <68 versus ≥68.

		Patient's number	SUS[a]		*P* value
			<68	≥68	
Sex, n (%)					.05[b]
	Male	55 (29.1)	10 (47.6)	45 (26.9)	
	Female	134 (70.9)	11 (52.4)	122 (73.1)	
Ethnicity and race, n (%)					.99[c]
	Asian	3 (1.6)	0 (0)	3 (1.8)	
	White	113 (59.8)	13 (61.9)	100 (59.9)	
	Mixed	44 (23.3)	5 (23.8)	38 (22.8)	
	Black	29 (15.3)	3 (14.3)	26 (15.6)	
Education, n (%)					.62[c]
	Elementary school	6 (3.2)	1 (4.8)	5 (3)	
	High school	91 (48.1)	9 (42.9)	81 (48.5)	
	College	92 (48.7)	11 (52.4)	81 (48.5)	
Marital status, n (%)					.32[c]
	Single	93 (49.2)	13 (61.9)	80 (47.9)	
	Married, common-law marriage, or partnered	84 (44.4)	6 (28.6)	77 (46.1)	
	Widowed	2 (1.1)	0 (0)	2 (1.2)	
	Judicially separated or divorced	10 (5.3)	2 (9.5)	8 (4.8)	
Monthly household income, n (%)					.88[c]
	Up to 3 minimum wages	95 (50.3)	12 (57.1)	82 (49.1)	
	4 - 6 minimum wages	59 (31.2)	5 (23.8)	54 (32.3)	
	7 - 11 minimum wages	23 (12.2)	3 (14.3)	20 (12)	
	Above 11 minimum wages	12 (6.3)	1 (4.8)	11 (6.6)	
Care function, n (%)					.59[c]
	No	145 (76.7)	15 (71.4)	129 (77.2)	
	Yes	44 (23.3)	6 (28.6)	38 (22.8)	
Has the patient requested support to use the cabin outside the previously set times[a]?, n (%)					.99[c]
	No	184 (97.3)	21 (100)	161 (97)	
	Yes	5 (2.7)	0 (0)	5 (3)	
At what point did the patient request medical advice?, n (%)					.39[b]
	Medical advice requested after the tests	68 (36)	8 (38.1)	59 (35.3)	
	Medical advice requested at the beginning; tests performed upon medical orientation	85 (45)	7 (33.3)	78 (46.7)	
	Did not request medical orientation	36 (19)	6 (28.6)	30 (18)	
Age (years), mean (SD)		35.1 (10.9)	35.0 (10.2)	35.2 (11.0)	.93
Appointment duration (minutes), mean (SD)		21.8 (11.8)	20.4 (10.3)	22.0 (12.0)	.58

XSL•FO
RenderX

	Patient's number	SUS[a]		*P* value
		<68	≥68	
NPS[d], mean (SD)	9.4 (1.3)	7.7 (2.5)	9.6 (0.8)	<.01
SUS, mean (SD)	86.0 (15.5)	51.2 (11.5)	90.3 (9.1)	<.01

[a]There was 1 missing value for the variable "SUS" and 1 missing value for the variable "Has the patient requested support to use the cabin outside the previously set times?".

[b]Chi-square test.

[c]Fisher exact test.

[d]NPS: Net Promoter Score.

Regarding the reasons for their ratings, only 3 participants chose not to answer this question. After careful reading of each response by 2 researchers, the answers were grouped according to the following categories:

- Service: assessment of the interaction between the cabin's professionals and the patient, considering the attention to the needs, clarification of doubts, and resolution of problems. Examples: (1) "Doctor's attention and understanding of my problem"; (2) "The speed and attention of the professionals who cared for me"; and (3) "Spectacular service. All my questions were answered. I thought it was excellent."
- Usability: an assessment of the ease of use of both the interface and the equipment. Examples: (1) "The experience was excellent; I enjoyed doing the tests myself and talking to the doctor about them"; (2) "Super practical, efficient, easy to use the equipment to diagnose, and talk to the doctor afterwards. Loved it!"; and (3) "I found it practical, accessible, and easy to use."
- Difficulty of use: difficulty in using equipment and tools to locate and understand the functions. Examples: (1) "Difficulty in accurately showing images and pain sites to the professional and the client's lack of affinity with the equipment can make it difficult to reach a diagnosis"; (2) "Some people may have difficulty with technology"; and (3) "Conflicting instructions between the cell phone and the screen at the self-diagnosis."
- Innovation: considerations about new products and technologies that can be aggregated to a business or project to bring about improvements. Examples: (1) "Great technology"; (2) "Very advanced technology"; and (3) "For being an innovative, versatile, and practical project."
- Technology Resistance: resistance to technological change. Example: "In no way does it replace the personal contact with the doctor, as there is no way of knowing for sure what you are feeling."
- Nonspecific: answers that were not intended or that did not belong to a predetermined group or situation, in this case, to classes of sentences mentioned above. Examples: (1) "Ok"; (2) "I loved it"; and (3) "Top."

The distributions of the classifications are shown in Table 4.

Most opinions are classified as related to "Usability." When comparing the classes of sentences of those who scored SUS equal or less than 68 (usability issues) with the remaining (acceptable usability), we found statistical differences in the class "Difficult Usability." In the group "Usability issues," 5 (27.8%) scored in this class, while 8 (4.8%) scored in the group "Acceptable usability."

XSL•FO
RenderX

Table . Distribution of the classes of sentences.

	Patient's number, n (%)	System Usability Scale (SUS), n (%)		*P* value
		<68	≥68	
Service				.06[a]
No	106 (57)	14 (77.8)	92 (55.1)	
Yes	80 (43)	4 (22.2)	75 (44.9)	
Usability				.23[a]
No	89 (47.8)	11 (61.1)	77 (46.1)	
Yes	97 (52.2)	7 (38.9)	90 (53.9)	
Difficulty of use				<.01[b]
No	173 (93)	13 (72.2)	159 (95.2)	
Yes	13 (7)	5 (27.8)	8 (4.8)	
Innovation				.99[b]
No	175 (94.1)	17 (94.4)	157 (94)	
Yes	11 (5.9)	1 (5.6)	10 (6)	
Technological resistance				.10[b]
No	185 (99.5)	17 (94.4)	167 (100)	
Yes	1 (0.5)	1 (5.6)	0 (0)	
Nonspecific				.17[b]
No	157 (84.4)	13 (72.2)	143 (85.6)	
Yes	29 (15.6)	5 (27.8)	24 (14.4)	

[a]Chi-square test.

[b]Fisher exact test.

Medical Assessment

At the end of the day's work, the doctor performing the appointments was oriented to answer the NPS and SUS, evaluating his or her day's work. The assessment was performed by 2 doctors, reducing the variability of the influence of the medical service on usability; 89% of the services were performed by one of the doctors, and 11% were performed by the other. The distribution of the NPS and SUS is shown in Table 5.

The main difficulties were related to technical difficulties, such as software updates and network connection problems, besides difficulties in adjusting the cabin's technology (problems with the sensors for measuring the weight and height; problems with the image or screen in the cabin).

Figures 4 and 5 show the evolution of the allocation of medical grades over time.

Table . Distribution of the Net Promoter Score (NPS) and System Usability Scale (SUS) values answered by the doctor at the end of each day of service.

	Patient's number (N)	Mean (SD)	95% CI	Min[a]	1oQ[b]	Medium	3oQ[c]	Max[d]
NPS	41	8.0 (1.4)	7.6-8.4	5.0	7.0	8.0	9.0	10.0
SUS	41	67.8 (13.6)	63.5-72.1	45.0	57.5	65.0	72.5	95.0

[a]Min: minimum.

[b]1oQ: first quartile.

[c]3oQ: third quartile.

[d]Max: maximum.

XSL•FO
RenderX

Figure 4. Evolution of the NPS answered by the doctor at the end of each day of service. NPS: Net Promoter Score.

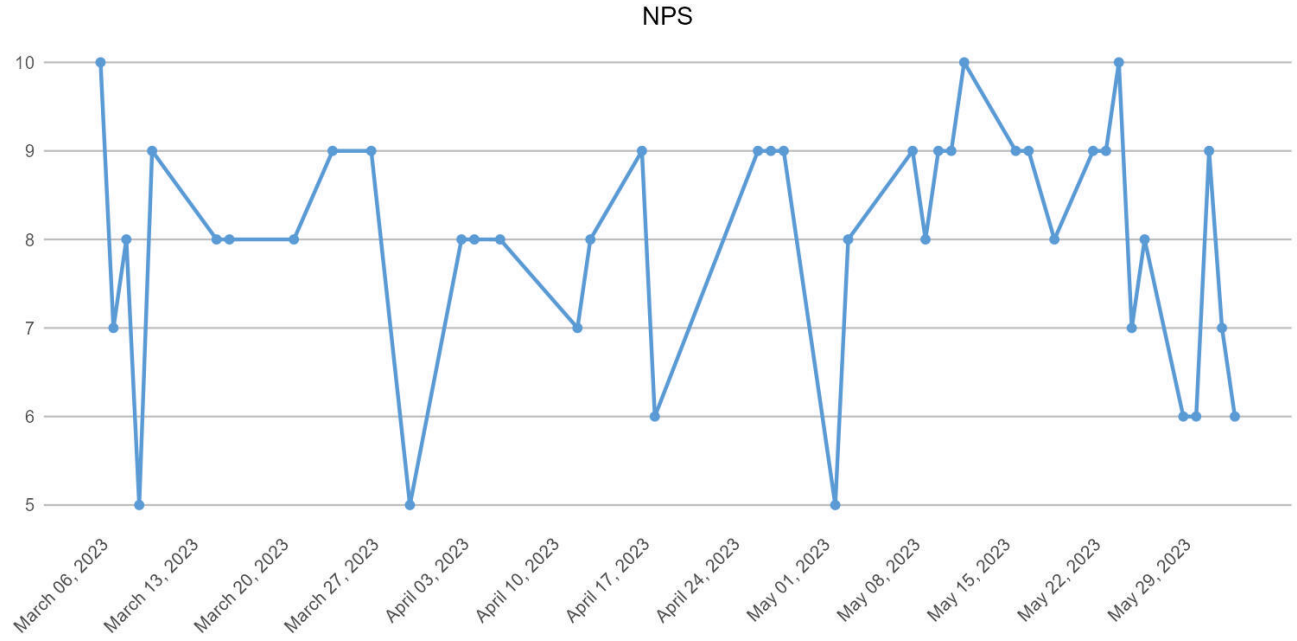
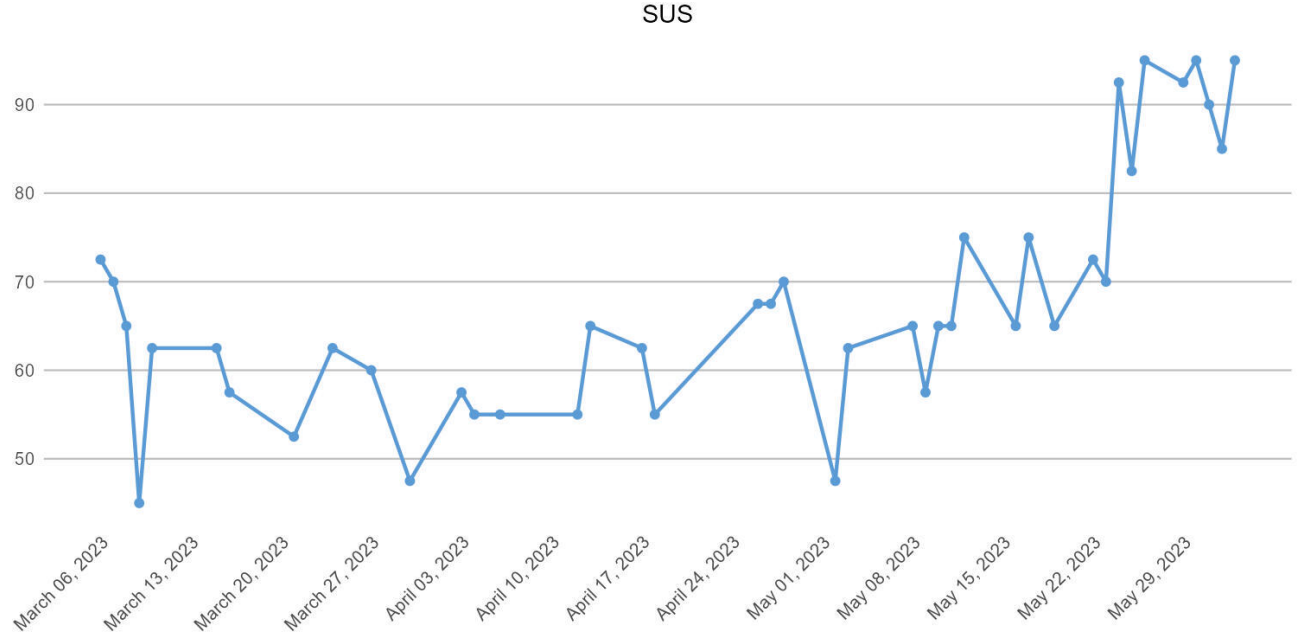


Figure 5. Evolution of the SUS score answered by the doctor at the end of each day of service. SUS: System Usability Scale.



Discussion

Principal Findings

The assessment of the implementation of the telehealth cabin in this study aimed to observe the level of favorability and usability of users (doctors and patients) to the cabin and the self-examination devices available, added to their experience of remote service through the technologies available.

The analysis of the results showed that the participants of the first phase of the research (asymptomatic patients) had a different socioeconomic profile than those of the second phase (symptomatic patients), with higher education and monthly income, probably due to the new technology, which attracted leaders and directors of the institution to the research. Although this group was only used to assess the acceptability of both the cabin and the self-examination devices in order to continue the study with symptomatic patients, the outcomes of the favorability and usability were similar between the groups.

Therefore, despite the disparity in access to technology among populations of different economic levels, such a result suggests that the telehealth cabin may have broad applicability for several publics due to the increasing availability and use of mobile devices among the whole population, thus promoting health the equity and reducing social inequalities.

A study with similar equipment performed with French students also concluded that the personal relationship with the technology did not influence the intended use of the telemedicine cabin [15].

XSL•FO
RenderX

A significant portion of the participants did not work in the hospital care sector and therefore did not routinely use the self-examination devices, suggesting that lack of familiarity with the use of the self-examination devices may not impact the use of the cabin. However, this study was performed with professionals working in a hospital, which may bring a certain level of familiarity with those devices, even for those not directly working in the care sector.

Still, we observed that more than half of the patients requested an interview with the doctor, either before or after self-examination, indicating the relevant role of the doctor in such a kind of service.

High favorability and usability were observed for both patient groups (symptomatic and asymptomatic). No statistically significant differences were found when comparing the socioeconomic and cultural profile of those patients with the group of symptomatic who considered the cabin to have acceptable usability, which shows that the usability and favorability of the cabin are unrelated to socioeconomic and cultural factors. However, concerning comments from this group show that some aspects were considered positive even in this group. Furthermore, only one patient described a comment that could be classified in the category "Resistance to Technology," suggesting that most of the population tested is interested in learning and adopting technologies that improve their routines and experiences.

The findings of this study support the work by Scott Kruse et al [8], asserting that the principal barriers to telemedicine are technology specific. They can be overcome by continuous improvement of these technologies, adequate training of technology users, and personal interaction between the patient and the care provider [8].

Concerning the favorability and usability of doctors who performed the cabin services cabin, we found that the answer was also positive, but with a slightly lower level of usability than that of the patients. This fact seems to agree with other studies that assert that some health professionals may be reluctant. However, this resistance may be related to the difficulty of the technology. Therefore, both the favorability and usability can improve with the advancement of technology and the due training of doctors in web-based physical tests [16,17]. This phenomenon highlights the importance of training programs as a strategy to improve user experience and expedite the adoption of this technology.

A limitation of this study is the participation of only individuals without respiratory symptoms (a consequence of the COVID pandemic) or without symptoms that could indicate a medical emergency. Another limiting factor is the poor accessibility of the cabin for people with disabilities. Furthermore, because the total sampling included people employed at the hospital, and so, in the productive age group and predominantly female [18], along with workers in the day shift, this study may have introduced a bias in the perception of favorability and usability of the participants (for instance, older men are less prone to participate in several telehealth activities [19]).

Finally, we emphasize that the results obtained from health professionals should also be interpreted with caution due to the limited number of professionals who provided care for the study.

It should also be noted that no studies using this type of technology were identified in South America to compare the results obtained. In addition, the fact that this study was conducted after the pandemic may have influenced the results given the impact of the COVID-19 pandemic on people's lives, which may have affected perceptions.

Based on the favorable assessment of the technology by the participants in this study (patients and doctors) new studies that extend its use to different populations, without any link to the hospital and from different sectors of activity, different work shifts, sex, and different age groups, may bring new inputs to the application of such technology, expanding the service types and favoring the inclusion of the service in remote locations and for diverse populations.

Conclusions

The authors conclude that the telehealth cabin had good usability and favorability by patients, regardless of the socioeconomic and cultural profile of the population. Doctors who performed the cabin services attested to the cabin's good usability cabin but with a usability index slightly lower than patients. Most patients used the devices for self-examination, with little need for technical assistance. However, a SUS score greater than 68 does not necessarily mean that there are no usability issues to be addressed.

The outcomes of this research allow us to assess as feasible the expansion of this kind of practice from the point of view of usability and that this study can contribute as a subsidy to the construction of public policies to expand the access of the Brazilian population to qualified medical care, promoting the engagement of patients in their health, providing health professionals with an extended view of the clinical conditions through embedded devices even in areas with a shortage of this type of assistance.

Acknowledgments

All authors would like to thank the Votorantim Institute for subsidizing the possibility of this project and for partnering with DOC24 and Diagnóstica, which brought a new way of experiencing the health of the future through innovative equipment.

Conflicts of Interest

None declared.

XSL•FO
RenderX

References

1. Universal Declaration of Human Rights [Web page in Portugese]. UNICEF. 1948. URL: https://www.unicef.org/brazil/declaracao-universal-dos-direitos-humanos [accessed 2022-02-20]
2. Jacobs B, Ir P, Bigdeli M, Annear PL, Van Damme W. Addressing access barriers to health services: an analytical framework for selecting appropriate interventions in low-income Asian countries. Health Policy Plan 2012 Jul;27(4):288-300. [doi: 10.1093/heapol/czr038] [Medline: 21565939]
3. Viacava F, Laguardia J, Ugá MAD, Porto S. PROADESS - performance assessment of the Brazilian health system: indicators for monitoring (research report). PROADESS. URL: http://www.proadess.icict.fiocruz.br/Relatorio_Proadess_08-10-2012.pdf [accessed 2023-07-18]
4. Silva AB, da Silva RM, Ribeiro GDR, et al. Three decades of telemedicine in Brazil: mapping the regulatory framework from 1990 to 2018. PLoS ONE 2020;15(11):e0242869. [doi: 10.1371/journal.pone.0242869] [Medline: 33237947]
5. Paré G, Moqadem K, Pineau G, St-Hilaire C. Clinical effects of home telemonitoring in the context of diabetes, asthma, heart failure and hypertension: a systematic review. J Med Internet Res 2010 Jun 16;12(2):e21. [doi: 10.2196/jmir.1357] [Medline: 20554500]
6. Polisena J, Coyle D, Coyle K, McGill S. Home telehealth for chronic disease management: a systematic review and an analysis of economic evaluations. Int J Technol Assess Health Care 2009 Jul;25(3):339-349. [doi: 10.1017/S0266462309990201] [Medline: 19619353]
7. Russo JE, McCool RR, Davies L. VA telemedicine: an analysis of cost and time savings. Telemed J E Health 2016 Mar;22(3):209-215. [doi: 10.1089/tmj.2015.0055] [Medline: 26305666]
8. Scott Kruse C, Karem P, Shifflett K, Vegi L, Ravi K, Brooks M. Evaluating barriers to adopting telemedicine worldwide: a systematic review. J Telemed Telecare 2018 Jan;24(1):4-12. [doi: 10.1177/1357633X16674087] [Medline: 29320966]
9. Brooke J. SUS: a quick and dirty usability scale. In: Usability Evaluation In Industry: CRC Press; 1995:189-194 URL: http://hell.meiert.org/core/pdf/sus.pdf [accessed 2025-04-03]
10. Reichheld FF. The one number you need to grow. Harv Bus Rev 2003 Dec;81(12):46-54. [Medline: 14712543]
11. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
12. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381. [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
13. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95(103208):103208. [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
14. Sauro J, Lewis JR. Quantifying the User Experience: Practical Statistics for User Research: Morgan Kaufmann; 2016.
15. Baudier P, Kondrateva G, Ammi C. The future of telemedicine cabin? The case of the French students' acceptability. Futures 2020 Sep;122(102595):102595. [doi: 10.1016/j.futures.2020.102595] [Medline: 32834075]
16. Lisboa KO, Hajjar AC, Sarmento IP, Sarmento RP, Gonçalves SHR. The history of telemedicine in Brazil: challenges and advantages. Saude Soc 2023;32(1):e210170pt. [doi: 10.1590/s0104-12902022210170pt]
17. Ladin K, Porteny T, Perugini JM, et al. Perceptions of telehealth vs in-person visits among older adults with advanced kidney disease, care partners, and clinicians. JAMA Netw Open 2021 Dec 1;4(12):e2137193. [doi: 10.1001/jamanetworkopen.2021.37193] [Medline: 34870680]
18. Machado MH, Wermelinger M, Machado AV, et al. Profile and working conditions of health professionals in times of Covid-19: the Brazilian reality. In: Portela MC, Reis LGC, Lima SML, editors. Covid-19: Challenges for the Organization and Repercussions on Health Systems and Services: Fiocruz Publishing House; 2022:283-295.
19. Kontos E, Blake KD, Chou WYS, Prestin A. Predictors of eHealth usage: insights on the digital divide from the health information national trends survey 2012. J Med Internet Res 2014 Jul 16;16(7):e172. [doi: 10.2196/jmir.3117] [Medline: 25048379]

Abbreviations

BP Hospital: Beneficência Portuguesa de São Paulo
NPS: Net Promoter Score
REDCap: Research Electronic Data Capture
SUS: System Usability Scale

XSL•FO
RenderX

Edited by A Kushniruk; submitted 21.12.23; peer-reviewed by C Udeh, J Brooke, NS Sood; revised version received 02.01.25; accepted 05.01.25; published 16.04.25.

Please cite as:
de França CVG, Segalla PB, Reis FSDA, Pereira JRS, de Mattos AO, Ferron RDM, de Oliveira CZ, Borges JB, Hoffmann LQ, Caboclo EDG
Patients' and Physicians' Experience With and Acceptability of a Telemedicine Cabin: Mixed Methods Study
JMIR Hum Factors 2025;12:e55430
URL: https://humanfactors.jmir.org/2025/1/e55430
doi:10.2196/55430

© Caroline Villela Galvão de França, Paola Boaro Segalla, Felipe Sebastião de Assis Reis, José Ricardo Silveira Pereira, Alexandre Oliveira de Mattos, Roberta de Moura Ferron, Cleyton Zanardo de Oliveira, Jéssica Bassani Borges, Lilian Quintal Hoffmann, Edmundo Di Giaimo Caboclo. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 16.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Gender-Inclusive Language in Public-Facing Labor and Delivery Web Pages in the New York Tristate Area: Cross-Sectional Study

Sarah Mohsen Isaac[1], BA; Mark Dawes[2], BSc; Emily Ruth Howell[1,3], DO; Antonia Francis Oladipo[1,4], MSCI, MD

[1]Hackensack Meridian School of Medicine, 123 Metro Blvd, Nutley, NJ, United States

[2]Cornell University, Ithaca, NY, United States

[3]Department of Obstetrics and Gynecology, Hackensack University Medical Center, Hackensack, NJ, United States

[4]Division on Maternal-Fetal Medicine, Department of Obstetrics and Gynecology, Hackensack University Medical Center, Hackensack, NJ, United States

Corresponding Author:
Sarah Mohsen Isaac, BA
Hackensack Meridian School of Medicine, 123 Metro Blvd, Nutley, NJ, United States

Abstract

Background: Transgender and nonbinary (TGNB) individuals are increasingly intentionally becoming pregnant to raise children, and hospital websites should reflect these trends. For prospective TGNB parents, a hospital website is the only way they can assess their safety from discrimination while receiving perinatal care. Cisnormativity enforced by communication gaps between medical institutions and TGNB patients can and has caused delays in receiving urgent care during their pregnancy.

Objective: The aim of this study was to evaluate the current prevalence of gender-inclusive terminology among labor and delivery services in the New York tristate area.

Methods: The labor and delivery web pages of 189 hospitals from New York, New Jersey, and Connecticut were examined for gender-inclusive language. “Fully inclusive” websites explicitly acknowledged lesbian, gay, bisexual, transgender, queer, intersex, and asexual plus other gender- and sexual-oriented (LGBTQIA+) parents, “inclusive” websites did not use gendered terminology for parents, and “noninclusive” websites used gendered terms at least once in the text reviewed. The hospitals’ web pages were further stratified by Healthcare Equality Index scores and population classifications defined by the 2013 National Center for Health Statistics Urban-Rural classification given to the county that each hospital was located in.

Results: Of the 300 hospital websites reviewed, only 189 websites met the criteria for inclusion. Overall, only 6.3% (n=12) of labor and delivery web pages were “inclusive” or “fully inclusive.” No geographic areas (*P*=.61) or Healthcare Equality Index scores (*P*=.81) were associated with inclusive or fully inclusive language.

Conclusions: Hospitals need to use inclusive language to help TGNB people identify hospitals where their existence and needs are acknowledged and thus feel more comfortable in their transition to parenthood.

(JMIR Hum Factors 2025;12:e53057) doi:10.2196/53057

KEYWORDS

OBGYN; transgender; nonbinary; pregnancy; maternity; transmasculine; observational study; gestational; perinatal care; communication; labor; USA; United States; New York City; sexual orientation; inclusion; parents; obstetrician gynecologist; delivery

Introduction

Within the last 2 decades, as transgender and nonbinary (TGNB) people have gained greater legal and social recognition, those in the TGNB community who want to become pregnant have become increasingly more common, as have the number of families with same-sex, transgender, or nonbinary parents [1]. Despite this trend, the field of medicine has mostly maintained the heteronormative model of a “mother” and “father” as opposed to a more fluid, freeform reality that accompanies the rise of lesbian, gay, bisexual, transgender, queer, intersex, and asexual plus other gender- and sexual-oriented (LGBTQIA+) couples who raise children.

As pregnant TGNB people attend numerous prenatal visits, they make repeated contact with heterosexist health care systems without the ability to hide their transgender status. Pharr [2] explains that heterosexism is not an active form of discrimination but rather “a belief that the world is and must be heterosexual.” According to the heterosexist worldview, every couple contains—or should contain—only 2 gender-conforming people of the opposite sex [3].

Heterosexism works with homophobia to make health care inaccessible for LGBTQIA+ populations. Current literature has

XSL•FO
RenderX

found that these parents and couples are often invalidated and marginalized throughout the health care process through obstacles like registration forms, comments from ancillary staff, and physicians who are unprepared to deal with LGBTQIA+ couples [4-6]. These experiences can take a deep mental and emotional toll—doubly so during the sensitive transition to parenthood [7].

Beyond these mental health impacts, the invisibility of TGNB parents can negatively affect pregnancy outcomes. LGBTQIA+ patients were less likely to trust providers and divulge important medical information when they received heteronormative medical treatment [8,9]. Patients may also experience delays receiving urgent pregnancy care due to the systematic heterosexism built into the health care system.

For example, Parker et al [6] discuss the experiences of a transmasculine patient who was delayed in seeing the doctor because the receptionist argued that he was not the intended person. These patterns are not limited to individual heterosexist health care workers but also integrated into hospital software. Berger et al [10] describe another scenario where a transmasculine patient's care was delayed, this time because the hospital's electronic medical record required that he reregister as female to document the pregnancy, regardless of his actual gender identity. A delay in care to circumvent cisnormative systems can be dangerous for all pregnant TGNB people. Ultimately, discrimination has often necessitated that prospective TGNB parents discern a "safe" hospital before seeking care.

Historically, LGBTQIA+ people have relied on word of mouth from their personal social circles to find safer health care [11]. However, younger LGBTQIA+ people, especially ones without LGBTQIA+ support networks, also rely on the internet to search for health information and providers that are inclusive of LGBTQIA+ people [12,13]. This vetting of hospitals, along with the increase of patient choice and consumerism for perinatal care in general, has prompted hospitals to advertise their unique benefits, such as low cesarean rates, "baby-friendly" designations, and private rooms [14-19]. Hospitals have tailored their advertisements for other demographics around them, but there is a dearth of literature showing how hospitals advertise their services for LGBTQIA+ populations, who rely on publicly available information to find inclusive care and preserve their health and safety [20].

Purdie-Vaughns et al [21] point to purposeful word choice as one safety cue that, when recognized, signals protection from identity-based discrimination. Hospitals might therefore attract pregnant LGBTQIA+ parents by crafting more inclusive obstetrical web pages. These pages could signal inclusivity by explicitly referencing LGBTQIA+ care or by avoiding gender-exclusive language like "mother and baby," "mom," or presumptive she/her pronouns for parents. Through the words chosen on these public-facing web pages, hospitals thus enable parents to choose to give birth in places where their existence is actively supported during the physically dangerous and psychologically difficult transition to parenthood.

The states surrounding New York City—New York, New Jersey, and Connecticut—boast an exceptionally high population density of LGBTQIA+ individuals, who make up between 3% to 5% of the total adult population [22]. This geographical region is viewed as more inclusive towards LGBTQIA+ people than average, so hospitals may have more incentive to provide inclusive care [23,24]. This study aims to evaluate the current prevalence of LGBTQIA+ inclusion and gender-inclusive terminology among labor and delivery (L&D) service web pages in the New York tristate area.

Methods

Study Design

The targeted words used to assess gender-inclusiveness for this study were largely adapted from the Jennings et al [25] study of gender-inclusive language on National Health Service websites.

The official public-facing obstetric web pages of nonfederal, short-term, acute-care hospitals from Connecticut, New Jersey, and New York were analyzed (n=300). Hospitals without L&D services or web content describing these services were excluded (n=189). Websites were reviewed from late November 2022 to January 2023.

Hospitals were categorized by state, 2013 National Center for Health Statistics Rural-Urban classification, and Healthcare Equality Index (HEI) score. The National Center for Health Statistics Rural-Urban classification is a tool used to identify urban and rural areas of the United States. It was used to analyze any association between urbanization and hospital-based inclusiveness of LGBTQIA+ people. The HEI score is the national LGBTQIA+ benchmarking criterion that assesses health care facilities' policies and practices regarding equity and inclusion of LGBTQIA+ patients, visitors, and employees. It was also used to identify if there was any association between a hospital's publicly perceived LGBTQIA+ inclusivity and word choice on the web pages.

For each hospital, at least 1 web page was examined alongside up to 2 additional pages as supplementation for language analysis. The gendered language used was recorded and analyzed by a single reviewer. The complete L&D-related text was analyzed and the types of gendered language used were recorded. (Explicit discussion of related services, such as chestfeeding, within the same site was excluded.) Any non–gender-inclusive descriptors for the name of the building or third-party services were also excluded from analysis, as these are often not controlled by hospital administration.

Language Analysis

Each web page was reviewed independently by the chief reviewer to minimize any discrepancies. Each hospital's L&D web page was rated as "fully inclusive," "inclusive," or "noninclusive." "Fully inclusive" websites explicitly acknowledged LGBTQIA+ or TGNB parents. "Inclusive" websites did not use gendered terminology or pronouns for prospective parents. "Noninclusive" websites used the terms "woman" or "women"; "mom" or "mother"; other terms for women; "father" or "dad"; or she/her pronouns at least once in the text reviewed.

XSL•FO
RenderX

Statistical Analysis

Categories were analyzed using χ^2 tests presented as frequencies with percentages. *P* values <.05 were considered statistically significant, and all tests were 2-sided.

Ethical Considerations

Ethics and insitutional review board approval were not required since the study did not include human or animal subjects and all data were collected from publicly available websites.

Results

Of the 300 hospital websites reviewed, 111 hospital websites did not have a L&D web page or did not have content describing their L&D services (Multimedia Appendix 1). Of the remaining 189 websites analyzed, 12 (6.3%) of them used fully inclusive or inclusive language (Table 1). Only 1 hospital (0.5%) was considered fully inclusive because it acknowledged "same-sex" couples in its L&D content. The most common noninclusive terms used were "mom" or "mother" (n=166, 87.8%) and "woman" or "women" (n=94, 49.7%). No geographic areas (*P*=.61) or HEI scores (*P*=.81) were associated with inclusive or fully inclusive language (Tables 2-4).

Table . Labor and delivery web pages that used each type of language (N=189).

	Fully inclusive, n (%)	Inclusive, n (%)	Noninclusive, n (%)	Total, n (%)
Websites	1 (0.5)	11 (5.8)	177 (93.7)	189 (100)
Language used				
woman or women	1 (1.1)	0 (0)	93 (49.2)	94 (49.7)
mother or mom	0 (0)	0 (0)	166 (87.8)	166 (87.8)
she/her (parent)	0 (0)	0 (0)	42 (22.2)	42 (22.2)
she/her (staff)	0 (0)	0 (0)	9 (4.8)	9 (4.8)
synonyms for women (ladies, etc)	0 (0)	0 (0)	1 (0.5)	1 (0.5)
father or dad	0 (0)	0 (0)	33 (17.5)	33 (17.5)

Table . Summary of HEI[a] scores and gender-inclusive language used on L&D[b] web pages .

	Fully inclusive (n=1), n (%)	Inclusive (n=11), n (%)	Noninclusive (n=177), n (%)	Total (N=189), n (%)
HEI score=100%	0 (0)	5 (45.5)	60 (33.9)	65 (34.8)
HEI score <100%	0 (0)	1 (9.1)	25 (14.1)	26 (13.9)
HEI score not applicable	1 (100)	5 (45.5)	92 (52)	98 (52.4)

[a]HEI: Healthcare Equality Index.

[b]L&D: labor and delivery.

Table . Summary of gender-inclusive language used on labor and delivery web pages and population data.

	Fully inclusive (n=1), n (%)	Inclusive (n=11), n (%)	Noninclusive (n=177), n (%)	Total (N=189), n (%)
Large central metro	0 (0)	6 (54.5)	53 (30.6)	59 (31.2)
Large fringe metro	0 (0)	3 (27.3)	65 (37.6)	68 (36)
Medium metro	1 (100)	1 (9.1)	31 (17.9)	33 (17.5)
Small metro	0 (0)	0 (0)	7 (4)	7 (3.7)
Micropolitan	0 (0)	1 (9.1)	17 (9.8)	18 (9.5)
Noncore	0 (0)	0 (0)	4 (3.9)	4 (2.1)

Table . Examples of suggested gender-inclusive language [19,25].

Non–gender-inclusive language	Gender-inclusive language
"Mothers"	"Birthing parents" OR "women and birthing parents"
"Pregnant woman"	"Pregnant patient" OR "pregnant person"
"Mother and baby unit"	"Maternity unit" OR "birthing unit"

XSL•FO
RenderX

Discussion

Principal Findings

These results demonstrate that there is a large barrier for TGNB parents to search for and identify potentially inclusive pregnancy care. Out of the 12 inclusive and fully inclusive L&D web pages, 92% were inclusive not because they included gender-additive language or LGBTQIA+ topics but rather because they omitted the pregnant person's gender altogether by addressing the reader in the second person. The websites may have been inclusive not by intention but by coincidence. In stark contrast, there are multiple private reproductive endocrinology and infertility clinics that specifically target LGBTQIA+ couples using specific gender-inclusive language on their websites [26]. TGNB parents who are accustomed to a purposefully inclusive experience during their fertility journey and early pregnancy may be caught off guard by the sudden invisibility of their identities as they progress further through their pregnancy.

Interestingly, none of the hospitals that are acknowledged for their excellence in LGBTQIA+ care in other specialties discussed serving prospective TGNB parents for L&D care on their websites. This likely reflects a wider societal trend of "repronormativity," by which society at large does not recognize reproductive sex between TGNB parents as possible or legitimate [27].

The accessibility of websites and web-based platforms is important for TGNB people to find services and connect to similar parents. Our above findings suggest less than 10% of hospitals use gender-inclusive language when representing their services. Thus, TGNB parents who are not connected to a wider LGBTQIA+ community may struggle to find inclusive prenatal care and delivery services due to the lack of representation.

Limitations

The use of multiple surrogate end points may limit this study. The analyzed web pages, while used as a proxy for the culture in L&D departments, may not fully represent institutional attitudes and practices once parents start using their providers. This is exacerbated by delays between institutional attitude changes and hospital website updates. HEI scores were ineffective in predicting the LGBTQIA+ inclusivity of hospitals' web pages because HEI scores are determined purely through institutional measures like nondiscrimination policies; they do not directly address subtler, underlying heterosexism that hopeful TGNB parents try to avoid in their health care. Finally, the methodology used in this study makes it impossible to establish a causal link between gendered terminology and the quality of LGBTQIA+ inclusive care.

Additionally, it is important to note that individual TGNB people may feel varying levels of dysphoria around maternal terms; some TGNB parents may not consider the words "mom" or "mother" to be exclusively for women. However, using gender-inclusive language and terminology is an important step towards providing a more welcoming and inclusive environment for TGNB parents, regardless of those individuals' personal dysphoria triggers.

Conclusions

Using gender-inclusive language and terminology is the first step towards providing a more welcoming and inclusive environment for pregnant TGNB parents. Hospitals that want to be recognized as more inclusive towards LGBTQIA+ people can integrate gender-additive language into their L&D web pages (eg, "mothers and birthing parents") rather than omit mentions of gender entirely [28]. US hospitals should consider expanding this language to meet the needs of a growing group of people who are having children. Future research should be done including LGBTQIA+ patient advocate groups on the use of inclusive language within health care providers' obstetrical and gynecology departments, specifically in L&D units and on how this language impacts TGNB parents' health outcomes and rapport with physicians.

Conflicts of Interest

None declared.

Multimedia Appendix 1
List of hospitals reviewed for analysis, along with a list of hospitals where labor and delivery pages were missing for various reasons.
[XLSX File, 77 KB - humanfactors_v12i1e53057_app1.xlsx]

References

1. Tornello SL, Bos H. Parenting intentions among transgender individuals. LGBT Health 2017 Apr;4(2):115-120. [doi: 10.1089/lgbt.2016.0153] [Medline: 28212056]
2. Pharr S. Homophobia: A Weapon of Sexism: Chardon Press; 1988.
3. Hudak N. "Who's the Mom?": heterosexism in patient-provider interactions of queer pregnant couples. Health Commun 2023 Jan;38(1):114-123. [doi: 10.1080/10410236.2021.1936752] [Medline: 34096428]
4. Wojnar DM, Katzenmeyer A. Experiences of preconception, pregnancy, and new motherhood for lesbian nonbiological mothers. J Obstet Gynecol Neonatal Nurs 2014;43(1):50-60. [doi: 10.1111/1552-6909.12270] [Medline: 24354595]

XSL•FO
RenderX

5. Light AD, Obedin-Maliver J, Sevelius JM, Kerns JL. Transgender men who experienced pregnancy after female-to-male gender transitioning. Obstet Gynecol 2014 Dec;124(6):1120-1127. [doi: 10.1097/AOG.0000000000000540] [Medline: 25415163]
6. Parker G, Ker A, Baddock S, Kerekere E, Veale J, Miller S. "It's Total Erasure": trans and nonbinary peoples' experiences of cisnormativity within perinatal care services in aotearoa New Zealand. Womens Reprod Health (Phila) 2022 Dec 26;10(4):591-607. [doi: 10.1080/23293691.2022.2155496]
7. Greenfield M, Darwin Z. Trans and non-binary pregnancy, traumatic birth, and perinatal mental health: a scoping review. Int J Transgend Health 2021;22:203-216. [doi: 10.1080/26895269.2020.1841057] [Medline: 34806082]
8. Utamsingh PD, Richman LS, Martin JL, Lattanner MR, Chaikind JR. Heteronormativity and practitioner-patient interaction. Health Commun 2016;31(5):566-574. [doi: 10.1080/10410236.2014.979975] [Medline: 26421354]
9. Seelman KL, Colón-Diaz MJP, LeCroix RH, Xavier-Brier M, Kattari L. Transgender noninclusive healthcare and delaying care because of fear: connections to general health and mental health among transgender adults. Transgend Health 2017;2(1):17-28. [doi: 10.1089/trgh.2016.0024] [Medline: 28861545]
10. Berger AP, Potter EM, Shutters CM, Imborek KL. Pregnant transmen and barriers to high quality healthcare. Proc Obstet Gynecol 2015;5(2):1-12. [doi: 10.17077/2154-4751.1285]
11. Hudak N, Bates BR. In pursuit of "queer-friendly" healthcare: an interview study of how queer individuals select care providers. Health Commun 2019 Jul;34(8):818-824. [doi: 10.1080/10410236.2018.1437525] [Medline: 29482359]
12. Charest M, Kleinplatz PJ, Lund JI. Sexual health information disparities between heterosexual and LGBTQ+ young adults: implications for sexual health. Can J Hum Sex 2016;25(2):74-85. [doi: 10.3138/cjhs.252-A9]
13. Mitchell KJ, Ybarra ML, Korchmaros JD, Kosciw JG. Accessing sexual health information online: use, motivations and consequences for youth with different sexual orientations. Health Educ Res 2014 Feb;29(1):147-157. [doi: 10.1093/her/cyt071] [Medline: 23861481]
14. Fischer S, Pelka S, Riedl R. Understanding patients' decision-making strategies in hospital choice: literature review and a call for experimental research. Cog Psych 2015 Dec 15;2(1). [doi: 10.1080/23311908.2015.1116758]
15. Mehrotra A, Wolfberg A, Shah NT, et al. Impact of an educational program and decision tool on choice of maternity hospital: the delivery decisions randomized clinical trial. BMC Pregnancy Childbirth 2022 Oct 10;22(1):759. [doi: 10.1186/s12884-022-05087-y] [Medline: 36217115]
16. Witteveen AB, De Cock P, Huizink AC, et al. Pregnancy related anxiety and general anxious or depressed mood and the choice for birth setting: a secondary data-analysis of the DELIVER study. BMC Pregnancy Childbirth 2016 Nov 22;16(1):363. [doi: 10.1186/s12884-016-1158-7] [Medline: 27871257]
17. Lally S, Lewis V. Maternity care patient engagement strategies. California Health Care Foundation. 2014. URL: https://www.chcf.org/wp-content/uploads/2017/12/PDF-MaternityCarePtEngagementStrategiesIHA.pdf [accessed 2024-12-13]
18. Tayyari Dehbarez N, Lou S, Uldbjerg N, Møller A, Gyrd-Hansen D, Søgaard R. Pregnant women's choice of birthing hospital: a qualitative study on individuals' preferences. Women Birth 2018 Dec;31(6):e389-e394. [doi: 10.1016/j.wombi.2017.11.006] [Medline: 29198502]
19. Maurer M, Firminger K, Dardess P, Ikeler K, Sofaer S, Carman KL. Understanding consumer perceptions and awareness of hospital-based maternity care quality measures. Health Serv Res 2016 Jun;51 Suppl 2(Suppl 2):1188-1211. [doi: 10.1111/1475-6773.12472] [Medline: 26927831]
20. Purcarea VL. The impact of marketing strategies in healthcare systems. J Med Life 2019;12(2):93-96. [doi: 10.25122/jml-2019-1003] [Medline: 31406509]
21. Purdie-Vaughns V, Steele CM, Davies PG, Ditlmann R, Crosby JR. Social identity contingencies: how diversity cues signal threat or safety for African Americans in mainstream institutions. J Pers Soc Psychol 2008 Apr;94(4):615-630. [doi: 10.1037/0022-3514.94.4.615] [Medline: 18361675]
22. Herman JL, Flores AR, O'Neill KK. How many adults and youth identify as transgender in the United States? UCLA Williams Institute. 2022. URL: https://williamsinstitute.law.ucla.edu/publications/trans-adults-united-states [accessed 2024-12-13]
23. Hasenbush A, Flores AR, Kastanis A, Sears B, Gates GJ. The LGBT divide: a data portrait of LGBT people in the midwestern, mountain & southern states. UCLA Williams Institute. 2014. URL: https://williamsinstitute.law.ucla.edu/lgbtdivide/ [accessed 2024-12-13]
24. 2022 State Equality Index. Human Rights Campaign. 2023. URL: https://reports.hrc.org/2022-state-equality-index [accessed 2024-12-13]
25. Jennings L, Goût B, Whittaker PJ. Gender inclusive language on public-facing maternity services websites in England. Br J Midwifery 2022 Apr 2;30(4):208-214. [doi: 10.12968/bjom.2022.30.4.208]
26. Wu HY, Yin O, Monseur B, et al. Lesbian, gay, bisexual, transgender content on reproductive endocrinology and infertility clinic websites. Fertil Steril 2017 Jul;108(1):183-191. [doi: 10.1016/j.fertnstert.2017.05.011] [Medline: 28579417]
27. Weissman AL. Repronormativity and the reproduction of the nation-state: the state and sexuality collide. J GLBT Fam Stud 2016 Sep 9;13(3):277-305. [doi: 10.1080/1550428X.2016.1210065]

XSL•FO
RenderX

28. Green H, Riddington A. Gender inclusive language in perinatal services: mission statement and rationale. Brighton and Sussex University Hospitals. 2020. URL: https://www.bsuh.nhs.uk/maternity/wp-content/uploads/sites/7/2021/01/Gender-inclusive-language-in-perinatal-services.pdf [accessed 2023-07-16]

Abbreviations

HEI: Healthcare Equality Index
L&D: labor and delivery
LGBTQIA+: lesbian, gay, bisexual, transgender, queer, intersex, and asexual plus other gender- and sexual-oriented identities
TGNB: transgender and nonbinary

Edited by K Adapa; submitted 24.09.23; peer-reviewed by F Storey, GP Quinn; revised version received 13.07.24; accepted 30.07.24; published 06.01.25.

Please cite as:
Isaac SM, Dawes M, Howell ER, Oladipo AF
Gender-Inclusive Language in Public-Facing Labor and Delivery Web Pages in the New York Tristate Area: Cross-Sectional Study
JMIR Hum Factors 2025;12:e53057
URL: https://humanfactors.jmir.org/2025/1/e53057
doi:10.2196/53057

© Sarah Mohsen Isaac, Mark Dawes, Emily Ruth Howell, Antonia Francis Oladipo. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 6.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Leveraging Smart Telemedicine Technology to Enhance Nursing Care Satisfaction and Revolutionize COVID-19 Care: Prospective Cohort Study

You-Lung Chang[1*], MD; Chi-Ying Lin[1*], MD; Jiun Hsu[2], MD, PhD; Sui-Ling Liao[3], RN; Chun-Ti Yu[3], RN, HN; Hung-Chueh Peng[1], MBA; Chung-Yu Chen[1,4], MD, PhD; Matthew Huei-Ming Ma[1,4], MD, PhD; Juey-Jen Hwang[5], MD, PhD

[1]Department of Internal Medicine, National Taiwan University Hospital Yunlin Branch, Douliou, Yunlin County, Taiwan

[2]Medical Information Technology Office, National Taiwan University Hospital Yunlin Branch, Yunlin County, Taiwan

[3]Department of Nursing, National Taiwan University Hospital Yunlin Branch, Yunlin County, Taiwan

[4]College of Medicine, National Taiwan University, Taipei, Taiwan

[5]Department of Cardiovascular Medicine, Fu Jen Catholic University Hospital, New Taipei, Taiwan

[*]these authors contributed equally

Corresponding Author:
Chung-Yu Chen, MD, PhD
Department of Internal Medicine, National Taiwan University Hospital Yunlin Branch, Douliou, Yunlin County, Taiwan

Abstract

Background: Telemedicine has been utilized in the care of patients with COVID-19, allowing real-time remote monitoring of vital signs. This technology reduces the risk of transmission while providing high-quality care to both self-quarantined patients with mild symptoms and critically ill patients in hospitals.

Objective: This study aims to investigate the application of telemedicine technology in the care of patients with COVID-19, specifically focusing on usability, effectiveness, and patient outcomes in both home isolation and hospital ward settings.

Methods: The study was conducted between January 2022 and December 2022. More than 800 cases were monitored using the QOCA remote home care system, a telemedicine platform that enables remote monitoring of physiological data—including heart rate, blood pressure, temperature, and oxygen levels—through Internet of Things devices and a 4G-connected tablet. Of these, 27 patients participated in thie study: the QOCA remote home care system was deployed 36 times in the isolation ward and 21 times to those in home isolation. The QOCA remote care system monitored isolated cases through remote care packages and a 4G tablet. Case managers and physicians provided telemedicine appointments and medications. Innovative methods were developed to enhance usage, including online health education, remote care equipment instructions via QR core links, and video consultations for patients without smartphones.

Results: A clinical nurse satisfaction survey revealed that most respondents found the content of the remote care package comprehensive and the interface easy to learn. They expressed a desire to continue using the system. The majority also agreed that using the remote care system and package would reduce their workload and that patients and caregivers could easily learn to use the package. While some respondents expressed concerns about network and Bluetooth connectivity, the majority (24/27, 89%) agreed to include the remote device as part of their routine equipment, with an average score of 84.8 points.

Conclusions: The integration of telemedicine technology improves the quality of care while reducing the workload and exposure of health care workers to viruses.

(JMIR Hum Factors 2025;12:e53456) doi:10.2196/53456

KEYWORDS

COVID-19; telemedicine; smart home care; intelligent medical care; nursing care; medical care; remote monitoring; vital signs; quarantine; home-care; hospital-care; screening; treatment; mHealth; mobile health; digital health; health education; remote equipment; smartphone; video consultation; remote care; caregiver; patient; quality of care; medical staff

Introduction

COVID-19, a global pandemic, has posed significant challenges to health care systems worldwide [1,2]. While Taiwan has achieved success in epidemic prevention, the reopening of borders raises concerns about potential gaps in prevention measures and strains on health care resources. Addressing these

XSL•FO
RenderX

challenges, this study identifies critical needs in the context of COVID-19 care.

The need for home quarantine and home medical care is vital. National Taiwan University Hospital Yunlin Branch has pioneered a telecommunication-based remote consultation model [3], providing video consultations for individuals under home quarantine. This initiative has not only met their basic health care needs but also minimized the risk of infection by reducing hospital visits. However, with the majority of cases being mild or asymptomatic, and a higher proportion of symptomatic cases among middle-aged and older individuals [4], it is crucial to develop a comprehensive telemedicine model for continuous medical care during home quarantine or isolation and for patients with mild COVID-19 undergoing home treatment [5-7].

During the COVID-19 pandemic, home monitoring of patients has gained popularity. There are two main research designs: (1) assessing patients before hospital admission to reduce virus exposure and identify deterioration and (2) postdischarge monitoring to provide continued care [8]. Monitoring methods include online platforms, paper-based operations with telephone interviews, or wearable devices [9-11]. Paper-based operations with telephone interviews are more inclusive due to limited digital literacy among the target population [12,13]. Most studies focus on vital signs rather than psychological issues. Wearing a pulse oximeter is important for monitoring, but it may cause anxiety and has limitations for certain patients [14-16]. Home monitoring during COVID-19 has benefits and can reduce health care costs [17]. Improving accessibility and convenience is key for future adoption.

Due to the pandemic, health care providers are unable to provide patients with the usual standard of care during hospitalization, leading to various challenges and impacts. Health care workers face psychological stress and staffing shortages, while patients experience isolation and mental health issues. Family members are unable to visit their loved ones in critical condition [18,19]. Technology can offer some assistance in addressing these issues, such as self-monitoring of blood oxygen levels and the integration of vital sign monitoring [20]. Collaboration between technology companies and health care providers is crucial, especially as the pandemic becomes normalized in the community. The development of information-assisted solutions is urgent in order to improve health care delivery [21].

There is a need for low-contact health care for hospitalized, isolated patients. To mitigate the risk of infection among patients with confirmed COVID-19 infection or those at high risk of being infected in dedicated isolation wards, the integration of smart ward solutions and telemedicine is proposed. Real-time remote physiological monitoring can minimize exposure risks and reduce the workload of health care staff, optimize management processes, and preserve valuable isolation rooms and personal protective equipment. Furthermore, obtaining real-time patient information enables early intervention and improves work efficiency, thereby enhancing the safety of patients with confirmed COVID-19 infection.

The study was conducted as a collaboration between National Taiwan University Hospital Yunlin Branch and Quanta Computer, utilizing their expertise in big data analysis and telecommunication telemedicine technology. The project focuses on the COVID-19 response hospitals, with the National Taiwan University Hospital Yunlin Branch serving as the testing ground. Specifically, it aims to provide assistance to patients with confirmed COVID-19 infection or those at high risk of being infected in the dedicated isolation ward, using advanced telemedicine equipment for clinical care. By implementing real-time remote physiological monitoring, the project aims to reduce the exposure risk and workload of health care personnel, simplify management processes, and optimize the utilization of valuable isolation rooms and personal protective equipment. Furthermore, by promptly assessing patient conditions and intervening as needed, the project seeks to enhance work efficiency and ensure the safety of patients with confirmed COVID-19 infection.

Methods

Ethical Considerations

The Institutional Review Board (IRB) of National Taiwan University Hospital approved this study (202009106RIPA). Questionnaires were collected after obtaining informed consent.

All participant data were anonymized or deidentified to ensure privacy. Participants voluntarily provided informed consent, and no personal identifiers were included in the analysis or dissemination. Additionally, anonymous surveys were used to ensure there were no risks of information leakage.

Participants in the study were not compensated monetarily or otherwise. Their involvement was voluntary, and all necessary support for participation, such as the provision of remote care packages and telemedicine tools, was offered free of charge. This ensured equitable participation without financial coercion.

Study Design, Data Sources, and Population

The study was conducted at the National Taiwan University Hospital Yunlin Branch. The inclusion criteria comprised individuals diagnosed with COVID-19, while exclusion criteria included moderate to severe COVID-19 cases, patients unwilling to participate in telemedicine monitoring, and individuals deemed unsuitable for telemedicine evaluation by physicians. Participants were required to be 20 years of age or older and provide informed consent approved by the IRB. The study enrollment period spanned from January to December 2022.

Home Quarantine and Home Medical Care

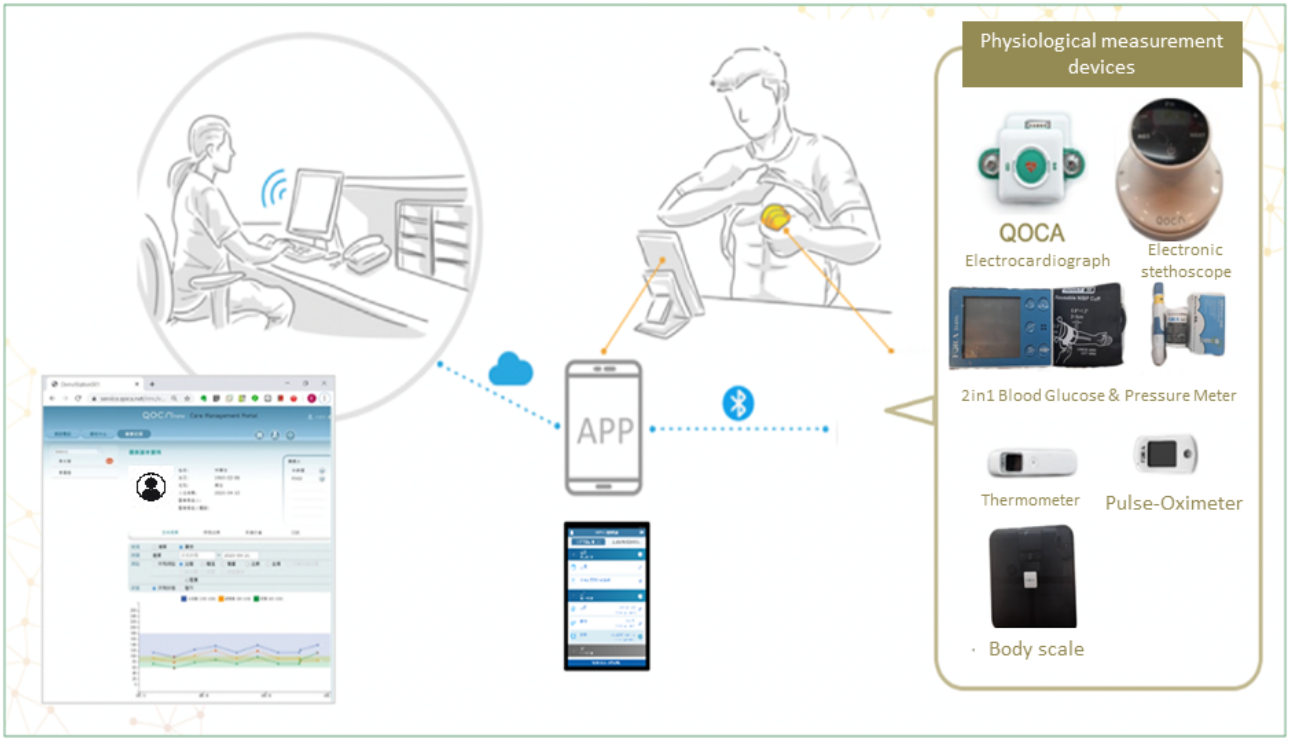
This study included individuals who have tested positive for COVID-19 and were classified as mild cases or have been determined by the Center for Disease Control, Taiwan, to not require immediate hospitalization due to severe conditions. These individuals can undergo self-isolation monitoring and were either required to quarantine or be monitored at home or quarantine hotels.

The QOCA home system, developed by Quanta Computer, is the core of the remote health care platform. It utilizes a 4G-connected tablet computer for patients, offering features like remote video communication, physiological signal measurement transmission, and remote physical examination (Figure 1). Home monitoring occurs at the patient's residence

XSL•FO
RenderX

or quarantine hotel, providing audiovisual communication and simple consciousness assessment. Vital signs, such as heart rate, blood pressure, temperature, and oxygen level, are measured and transmitted using Internet of Things (IoT) devices. A remote physical examination system, including electronic stethoscope devices, captures and uploads important findings like lung and heart sounds. An interactive care app records visual aspects and clinical conditions, which can be uploaded for health care providers. An online questionnaire system tracks COVID-19 symptoms, comorbidity control, and psychological well-being. Medication assessment and personalized care planning are provided. The hospital-side system manages home isolation quarantine and home-based COVID-19 care. Case managers track patient conditions through a dashboard, maintain daily health logs, and determine if hospital treatment is needed. Information and communication technology–based remote health care ensures continuous medical care and psychological support for individuals in home quarantine or isolation and home-based patients with COVID-19. Remote care packages enable self-isolated individuals to monitor and report symptoms and vital signs, with data transmitted back to the hospital. Dedicated care managers communicate with patients, report to physicians, and provide necessary support in case of discomfort or abnormalities.

Figure 1. The QOCA home system by Quanta Computer is central to the remote health care platform. Utilizing a 4G tablet, it enables remote communication, physiological data transmission, and virtual exams. Monitoring occurs at homes or quarantine locations, allowing audiovisual contact and basic assessment. IoT devices measure and send vital signs, while an electronic stethoscope captures and uploads lung and heart sounds. An interactive app records visual and clinical data for health care providers. An online questionnaire tracks COVID-19 symptoms, comorbidity control, and psychological well-being and includes medication assessment and personalized care planning. IoT: Internet of Things.



Smart Hospital Ward

Patients with COVID-19 admitted to negative pressure isolation wards or intensive care units at National Taiwan University Hospital Yunlin Branch were monitored using the QOCA home system developed by Quanta Computer. Tablet computers, IoT physiological monitoring devices, and video cameras were installed in patient rooms for remote monitoring. Patient physiological data, including heart rate, respiration, body temperature, blood pressure, and blood oxygen levels, were automatically recorded and uploaded to the cloud. The care team accessed these data without entering the rooms. The smart ward included an electronic stethoscope device for recording heart and lung sounds. A central monitoring module and a patient dynamic dashboard were established at the nursing station, providing alerts for abnormal conditions. Patient laboratory reports and images were integrated, enabling health care providers to interact with patients through the video system. The study aimed to utilize information and communication technology, including 4G-connected tablets and IoT monitoring platforms, to facilitate remote health care for patients with COVID-19 in isolation units. This system enabled continuous medical care, psychological support, and remote consultations with physicians; nursing care; psychological counseling; and rehabilitation therapy. It allowed for monitoring of vital signs and COVID-19 symptoms, promoting patient well-being without direct contact.

Satisfaction Survey for the QOCA Smart Ward Remote Care System: Clinical Nursing Staff and General Population

This questionnaire aimed to understand the relevant settings of the remote care system and its clinical application in dedicated wards for COVID-19. The information gathered will be used for future improvements and optimization. This survey was anonymous, and there were no concerns about personal

XSL•FO
RenderX

information leakage. The questionnaire included aspects of communication quality, user-friendliness, overall satisfaction, and suggestions. Each option consisted of strongly agree, agree, neutral, disagree, and strongly disagree (Multimedia Appendix 1).

Results

Overview

Before officially enrolling patients in the monitoring care package, thorough explanations and assistance were provided to each individual in the isolation ward to ensure their understanding and obtain their consent to participate in the research study. Following this, every isolated patient was equipped with a set of remote care packages. This became standard practice in the ward, significantly reducing the frequency of health care personnel entering patient rooms to don and remove isolation garments. Patients were empowered to self-measure their vital signs whenever they felt unwell, providing immediate access to physiological data that could alert health care professionals to any discomfort or changes in their condition. This approach not only alleviated patients' psychological anxiety about their disease progression but also allowed health care staff to spend less time in isolation garments while maintaining a real-time understanding of the patients' current status, ultimately improving the quality of clinical care.

Enhancing Hardware Infrastructure for Remote Care System in Isolation Ward

To enhance the hardware aspect of the remote care system implemented in the isolation ward, the original 7-bed negative pressure ward was expanded to accommodate the installation of 32 sets of care packages, with one set allocated to each bed. Additionally, the nursing treatment cart and the nursing station computer were equipped with the QOCA remote care system platform and individual nursing staff accounts. This setup empowered the nursing personnel to monitor the patients' vital sign data in real time, which were automatically uploaded by the system after patient measurements. Continuous monitoring of vital signs was made possible throughout the patients' entire hospitalization and isolation period. Moreover, the QOCA system facilitated real-time video communication when necessary, enabling consultations with psychiatric physicians to address any illness-related anxieties that the patients may have. By extending the focus of care beyond the patients' physiological needs, the hospitalization and treatment process also addressed their psychological well-being.

Remote Home Care Package for Individual Who Are COVID-19 Positive Under Self-Isolation

For individuals who have tested positive for COVID-19 and can self-isolate at home without requiring hospitalization, a comprehensive remote home care package was provided. This package enabled individuals to independently monitor their health, even without monitoring devices at home. It included daily self-observation of symptoms and measurement of vital signs, with the numerical data automatically transmitted to the hospital through a tablet computer connected to 4G.

To ensure prompt attention to any discomfort or adverse conditions, a dedicated care manager was assigned to each patient. The care manager promptly reported any concerns to the responsible physician and maintained regular contact to inquire about the patient's condition throughout the home isolation period. If individuals experience symptoms, they could also schedule telemedicine appointments to closely monitor changes in their condition.

In cases where symptom management was necessary, the responsible physician could prescribe appropriate medications to help alleviate symptoms. This comprehensive remote home care package aimed to provide effective care and support to individuals with COVID-19, ensuring their well-being while minimizing the risk of further transmission.

Advancing Care Through QOCA Telecare System for COVID-19 Monitoring and Isolation

The number of cases monitored using the QOCA remote home care system exceeded 800. Among them, 29 individuals voluntarily participated in this research study and provided their consent by signing the IRB consent form. The QOCA remote home care system was deployed 36 times in the isolation ward and 21 times to those in home isolation, as indicated in Table 1. In the ward, patients primarily monitor themselves or receive assistance from family members for monitoring, while a small number of individuals without family support or hired caregivers receive help from nursing staff for measurements.

The participants in this study represented a diverse range of age groups, with the majority being aged 41-60 years (13/29, 45%), followed by the 61-80 years age group (9/29, 31%), and the fewest participants in the 21-40 years (3/29, 10%) and 80+ years age groups (4/29, 14%). The sample was predominantly male, with 27 (93%) male participants and only 2 (7%) female participants. Regarding the need for assistance, 11 (38%) participants required help during the study, while 18 (62%) managed independently. The participants' education levels varied, with most having elementary or junior high education (20/29, 69%), while a smaller number had completed high school or higher education (9/29, 31%). In terms of technology usage, 21 (72%) participants used digital devices daily, while 8 (28%) reported no prior experience. Health engagement levels were mixed, with the majority (20/29, 69%) not actively seeking information about their conditions, while some participants regularly consumed health-related content. Additionally, 17 (59%) participants had prior experience with health monitoring devices, further illustrating the range of familiarity with telecare technology among the group.

In contrast to traditional inpatient monitoring methods, the ward implemented a specialized remote home care system designed specifically for the isolation care required during the COVID-19 pandemic. This system enables patients to independently measure their vital signs and automatically transmit the data to the nursing staff. To ensure effective utilization of the remote home care system by patients and their accompanying family members, innovative care methods have been developed, including the following:

XSL•FO
RenderX

1. Online health education materials: With the widespread use of smartphones due to technological advancements, disease-specific health education materials are provided through QR codes during hospitalization. This allows patients to access relevant information whenever needed.
2. Guidelines for operating remote care devices: Some individuals may be unfamiliar with monitoring devices. Even after receiving instructions from health care professionals, patients may require repeated practice. To facilitate practice sessions without overwhelming the nursing staff, instructional videos demonstrating device operation can be accessed through QR codes. Patients can practice as many times as necessary to ensure accurate measurements.
3. Tablet-based video consultations for remote care: Patients without smartphones can still access important information through tablet-based video consultations offered by the remote care system.

The ward continues to fully utilize the remote home care system for clinical monitoring, and the number of users continues to grow. Health care professionals not only provide personalized care to patients but also make adjustments to the care process based on the implementation of the remote home care system. This approach reduces the risk of viral exposure and eases the workload for health care staff, streamlines management procedures, and ultimately enhances the quality of care provided.

Table . Results of case enrollment.

	Value
QOCA remote home care system, n	
Isolation ward	36
Home isolation	21
Total	57
Goal	50
Total utilization (cases), n	>800
Execution rate	100%

Clinical Nursing Staff Satisfaction Survey

This survey aimed to assess the satisfaction and experiences of the clinical nursing staff in using the telecare system in the dedicated COVID-19 isolation wards. The questionnaire was divided into 3 main categories: *communication quality*, *user-friendliness*, and *overall experience*. The objective was to identify and analyze the challenges and difficulties faced by users in relation to equipment, signals, and interface operations.

A total of 27 questionnaires were collected exclusively from nursing staff working in the dedicated COVID-19 wards, specifically Ward 7A and Ward 7B. Regarding the transmission signals of the equipment, 44% (n=12) of the respondents believed that the internet connection of the telecare system was stable, while 37% (n=10) found the Bluetooth connection between the tablet and the devices in the package to be stable. When using the telecare system for video calls, 56% (n=15) stated that the image quality was clear, and 56% (n=15) stated that the sound quality was clear (Table 2).

In terms of usability, following training on the telecare system's interface, 89% (n=24) of the respondents found it easy to learn, and an equal percentage (n=24, 89%) expressed their willingness to continue using the telecare system to assist in clinical care. Furthermore, 67% (n=18) believed that patients or caregivers could quickly grasp the usage of the telecare package. Moreover, 81% (n=22) of the respondents indicated that the telecare package contained all the necessary components and fulfilled the requirements for clinical monitoring. Additionally, 89% (n=24) agreed that the use of the telecare system and package could help alleviate the burden of clinical care (Table 3).

Overall, 89% (n=24) of the respondents agreed to incorporate the telecare device as a regular piece of equipment in isolation wards, with an average satisfaction score of 84.8. The majority (24/27, 89%) of the responses highlighted the short monitoring time of the device and the instability of signal transmission. They expressed a desire for improved stability in wireless internet and Bluetooth signals, as it would lead to a smoother user experience (Table 4).

XSL•FO
RenderX

Table . Communication quality survey responses regarding the telemedicine system.

Question	Score[a] (n=27), n (%)				
	1	2	3	4	5
I think the network connection of this telecare system is stable	1 (4)	11 (41)	12 (44)	3 (11)	0 (0)
The Bluetooth connection between the tablet in the care bag and the instrument is stable	2 (7)	8 (30)	12 (44)	5 (19)	0 (0)
When using the nursing system, the image is very clear.	3 (11)	12 (44)	6 (22)	5 (19)	1 (4)
When using the nursing system, the sound is very clear.	2 (7)	13 (48)	8 (30)	3 (11)	1 (4)

[a]Score: 1=strongly agree, 2=agree, 3=neutral, 4=disagree, and 5=strongly disagree.

Table . User-friendliness survey responses regarding the telemedicine system.

Question	Score[a] (n=27), n (%)				
	1	2	3	4	5
The interface of this telecare system is easy for me to use after being taught.	15 (56)	9 (33)	1 (4)	1 (4)	1 (4)
I find the equipment in this telecare package difficult to use.	1 (4)	3 (11)	1 (4)	6 (22)	16 (59)
I think most patients or caregivers can quickly learn to use the telecare package.	7 (26)	11 (41)	4 (15)	3 (11)	2 (7)
I think the contents of this telecare package are complete and meet the needs of clinical monitoring	12 (44)	10 (37)	3 (11)	1 (4)	1 (4)

[a]Score: 1=strongly agree, 2=agree, 3=neutral, 4=disagree, and 5=strongly disagree.

XSL•FO
RenderX

Table . Overall satisfaction with the QOCA remote care system in smart wards.

Question and score	Value (n=27), n (%)
I think the use of telecare systems and care packages can help reduce the load of clinical care.[a]	
1	15 (56)
2	9 (33)
3	2 (7)
4	0 (0)
5	1 (4)
I would like to continue to use this telecare system to assist clinical care.[a]	
1	15 (56)
2	9 (33)
3	2 (7)
4	0 (0)
5	1 (4)
I think the remote care system and care package can become one of the routine equipment used by the unit.[a]	
1	15 (56)
2	6 (22)
3	4 (14)
4	1 (4)
5	1 (4)
If a full score is 100, how much would you like to rate the telecare system and care package?	
60	1 (4)
70	2 (7)
75	1 (4)
80	6 (22)
85	3 (11)
90	11 (41)
95	2 (7)
99	1 (4)

[a]Score: 1=strongly agree, 2=agree, 3=neutral, 4=disagree, and 5=strongly disagree.

Discussion

Principal Findings

The enhancement of the hardware infrastructure for the remote care system in the isolation ward has significantly improved the quality of care and patient outcomes [22,23]. By expanding to accommodate more care packages, each patient receives continuous monitoring and support. The integration of the QOCA remote care system into nursing carts and computers allows real-time access to vital sign data, enabling prompt interventions and ensuring patient well-being. The system also facilitates video consultations with psychiatric physicians, addressing patients' anxieties and holistic needs. This approach benefits both patients and nursing staff, who can efficiently monitor multiple patients, respond promptly to changes, and access patient data remotely, thus saving time and reducing errors.

The implementation of a remote home care package for individuals who are COVID-19 positive in self-isolation offers effective care and support while reducing transmission risk [24]. Patients can independently monitor their health by observing symptoms and measuring vital signs, actively participating in their care. A tablet connected to 4G enables real-time transmission of data to health care providers, facilitating informed decisions. Dedicated care managers provide personalized support, regularly checking in on patients and promptly reporting concerns. Telemedicine appointments allow close monitoring, adjustment of treatment plans, and direct access to medical professionals, thus reducing anxiety. Remote prescription of medications ensures symptom management without visiting a health care facility. However, challenges include subjective self-reporting and disparities in internet access and technological skills [23,25]. Education and efforts to address disparities are essential for equitable care.

XSL•FO
RenderX

The implementation of the QOCA remote home care system for COVID-19 monitoring and isolation has been successful in advancing care and improving patient outcomes. With over 800 cases monitored, the system allows patients to independently measure vital signs and transmit data to nursing staff, reducing the need for frequent in-person monitoring. Innovative methods, such as online health education materials and instructional videos, ensure optimal utilization of the system. Tablet-based video consultations cater to patients without smartphones. By fully utilizing the remote home care system, the ward provides personalized care, streamlines management procedures, reduces viral exposure, and eases the workload for staff, ultimately enhancing care quality [26]. Ongoing monitoring and evaluation are crucial for system refinement, and its success in COVID-19 care opens doors for its application in other health care settings, showcasing the transformative potential of technological advancements in patient care.

During the COVID-19 pandemic, nursing professionals have made substantial achievements in delivering adequate patient care by utilizing telehealth [8,12,24]. The results of the clinical nursing staff satisfaction survey regarding the telecare system in COVID-19 isolation wards provide valuable insights into their experiences and challenges [27]. The majority of respondents found the telecare package to be usable and to have met the requirements for clinical monitoring. After training, the respondents stated that the interface was easy to learn, and they expressed their willingness to continue using the system. A substantial percentage believed that patients and caregivers could quickly grasp its usage, and that the use of the telecare system helped alleviate the burden of clinical care [12]. Regarding transmission signals, opinions were mixed. While a portion of respondents found the internet connection stable, others highlighted issues with Bluetooth connectivity. In terms of video calls, a majority found the image and sound quality to be clear. Overall, the survey indicates high satisfaction with the telecare system, with a majority of respondents expressing their support for incorporating it as a regular equipment in isolation wards. The average satisfaction score was positive [13]. However, feedback emphasized the need for improvements in signal stability, particularly in wireless internet and Bluetooth connections [24]. The short monitoring time of the device was also a concern.

The difficulties and obstacles in telemedicine for COVID-19 care include insufficient hardware environment and the short monitoring time of equipment [28]. The negative pressure isolation environment in the ward, with multiple walls and controlled doors, leads to weak signals. Installing signal amplifiers, as suggested by the hospital's IT department, can improve signal stability. The monitoring time of the equipment is too short, but through discussions with engineers and the manufacturer, it is believed that adjusting the monitoring time according to individual patient needs can align it with clinical requirements, enabling more accurate monitoring of values. Addressing these challenges will enhance the effectiveness and reliability of telemedicine in COVID-19 care.

One potential limitation of this study is the relatively small sample size, which may affect the generalizability of the findings. With only 36 deployments in the isolation ward and 21 deployments in home isolation, the results may not fully represent the broader population of patients with COVID-19 or other health care contexts. Additionally, the voluntary nature of participation could introduce selection bias, as those who agreed to participate may differ from those who declined in terms of health status or access to technology. Another limitation is that the study was conducted within a single hospital, potentially limiting the applicability of the results to other settings with different resources or patient demographics. Lastly, the short-term focus of the study may not capture the long-term effectiveness or sustainability of telemedicine interventions, warranting further research in this area.

Conclusion

The implementation of the QOCA telemedicine system demonstrated significant benefits in managing patients with COVID-19 who have mild symptoms, both in isolation wards and home quarantine settings. By enabling continuous remote monitoring of vital signs and offering real-time communication between patients and health care providers, the system not only improved the quality of care but also reduced the workload and exposure risk for medical staff.

However, the study had some limitations, including a relatively small sample size and the potential for selection bias, which may affect the generalizability of the findings. Additionally, the short-term nature of the study may not fully capture the long-term impacts of telemedicine in patient care.

Despite these limitations, our findings suggest that telemedicine systems like QOCA can play a crucial role in pandemic response and future health care delivery. Further research with larger, more diverse populations and extended follow-up periods is recommended to fully explore the potential of telemedicine in improving patient outcomes and health care efficiency.

Acknowledgments

We sincerely thank Quanta Computer for their invaluable support in establishing the following health care systems: the Smart Ward Cloud Care System, QOCA Home, Intelligent Remote Health Care System, and Remote Home Care Package. This study was supported by the National Science and Technology Council, Taiwan, for the research (MOST 109－2823－8－002－013-CV).

Data Availability

All data generated or analyzed during this study are included in this published paper and its supplementary materials.

XSL•FO
RenderX

Authors' Contributions

Conceptualization: MHMM, JJH
Data curation: YLC, CYL
Investigation: JH, SLL, CTY, HCP
Methodology: JH, SLL, CTY, HCP, MHMM, JJH
Software: JH
Supervision: CYC, MHMM, JJH
Validation: CYC
Visualization: JH, SLL, CTY
Writing – original draft: YLC, CYL
Writing – review & editing: CYC

Conflicts of Interest

None declared.

References

1. Mahase E. COVID-19: WHO declares pandemic because of “alarming levels” of spread, severity, and inaction. BMJ 2020 Mar 12;368:m1036. [doi: 10.1136/bmj.m1036] [Medline: 32165426]
2. Zhu N, Zhang D, Wang W, et al. A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med 2020 Feb 20;382(8):727-733. [doi: 10.1056/NEJMoa2001017] [Medline: 31978945]
3. Chang KC, Hung SH, Hsueh SJ, et al. Development of the Houston-Apollo model for older people living in remote areas in Taiwan. Geriatr Gerontol Int 2021 Mar;21(3):299-305. [doi: 10.1111/ggi.14130] [Medline: 33527540]
4. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet 2020 Mar 28;395(10229):1054-1062. [doi: 10.1016/S0140-6736(20)30566-3] [Medline: 32171076]
5. Hollander JE, Carr BG. Virtually Perfect? telemedicine for COVID-19. N Engl J Med 2020 Apr 30;382(18):1679-1681. [doi: 10.1056/NEJMp2003539] [Medline: 32160451]
6. COVID-19 telehealth program (invoices & reimbursements). Federal Communications Commission. URL: https://www.fcc.gov/covid-19-telehealth-program [accessed 2021-12-12]
7. Telehealth.HHS.gov. US Department of Health and Human Services. URL: https://www.hhs.gov/coronavirus/telehealth/index.html [accessed 2021-12-12]
8. Vindrola-Padros C, Singh KE, Sidhu MS, et al. Remote home monitoring (virtual wards) for confirmed or suspected COVID-19 patients: a rapid systematic review. E Clin Med 2021 Jul;37:100965. [doi: 10.1016/j.eclinm.2021.100965] [Medline: 34179736]
9. O’Keefe JB, Tong EJ, Taylor TH Jr, O’Keefe GAD, Tong DC. Use of a telemedicine risk assessment tool to predict the risk of hospitalization of 496 outpatients with COVID-19: retrospective analysis. JMIR Public Health Surveill 2021 Apr 30;7(4):e25075. [doi: 10.2196/25075] [Medline: 33667174]
10. Margolius D, Hennekes M, Yao J, et al. On the front (phone) lines: results of a COVID-19 hotline. J Am Board Fam Med 2021 Feb;34(Suppl):S95-S102. [doi: 10.3122/jabfm.2021.S1.200237] [Medline: 33622824]
11. Hutchings OR, Dearing C, Jagers D, et al. Virtual health care for community management of patients with COVID-19 in Australia: observational cohort study. J Med Internet Res 2021 Mar 9;23(3):e21064. [doi: 10.2196/21064] [Medline: 33687341]
12. Greenhalgh T, Knight M, Inda-Kim M, Fulop NJ, Leach J, Vindrola-Padros C. Remote management of COVID-19 using home pulse oximetry and virtual ward support. BMJ 2021 Mar 25;372:677. [doi: 10.1136/bmj.n677] [Medline: 33766809]
13. Agarwal P, Mukerji G, Laur C, et al. Adoption, feasibility and safety of a family medicine-led remote monitoring program for patients with COVID-19: a descriptive study. CMAJ Open 2021 Apr 1;9(2):E324-E330. [doi: 10.9778/cmajo.20200174] [Medline: 33795222]
14. COVID-19 clinical management: living guideline. World Health Organization. 2021 Jan 25. URL: https://apps.who.int/iris/bitstream/handle/10665/338882/WHO-2019-nCoV-clinical-2021.1-eng.pdf [accessed 2025-01-06]
15. Nunan J, Clarke D, Malakouti A, et al. Triage Into the Community for COVID-19 (TICC-19) patients pathway - service evaluation of the virtual monitoring of patients with COVID pneumonia. Acute Med 2020;19(4):183-191. [Medline: 33215171]
16. Feiner JR, Severinghaus JW, Bickler PE. Dark skin decreases the accuracy of pulse oximeters at low oxygen saturation: the effects of oximeter probe type and gender. Anesth Analg 2007 Dec;105(6 Suppl):S18-S23. [doi: 10.1213/01.ane.0000285988.35174.d9] [Medline: 18048893]
17. Poncia HD, Ryan J, Carver M. Next day telephone follow up of the elderly: a needs assessment and critical incident monitoring tool for the accident and emergency department. J Accid Emerg Med 2000 Sep;17(5):337-340. [doi: 10.1136/emj.17.5.337] [Medline: 11005403]

XSL•FO
RenderX

18. Deressa W, Worku A, Abebe W, Gizaw M, Amogne W. Risk perceptions and preventive practices of COVID-19 among healthcare professionals in public hospitals in Addis Ababa, Ethiopia. PLoS ONE 2021 Jun 25;16(6):e0242471. [doi: 10.1371/journal.pone.0242471] [Medline: 34170910]
19. Wakam GK, Montgomery JR, Biesterveld BE, Brown CS. Not dying alone - modern compassionate care in the COVID-19 pandemic. N Engl J Med 2020 Jun 11;382(24):e88. [doi: 10.1056/NEJMp2007781] [Medline: 32289215]
20. Sutherland AE, Stickland J, Wee B. Can video consultations replace face-to-face interviews? palliative medicine and the COVID-19 pandemic: rapid review. BMJ Support Palliat Care 2020 Sep;10(3):271-275. [doi: 10.1136/bmjspcare-2020-002326] [Medline: 32457086]
21. Ohannessian R, Duong TA, Odone A. Global telemedicine implementation and integration within health systems to fight the COVID-19 pandemic: a call to action. JMIR Public Health Surveill 2020 Apr 2;6(2):e18810. [doi: 10.2196/18810] [Medline: 32238336]
22. Shen YT, Chen L, Yue WW, Xu HX. Digital technology-based telemedicine for the COVID-19 pandemic. Front Med (Lausanne) 2021 Jul 6;8:646506. [doi: 10.3389/fmed.2021.646506] [Medline: 34295908]
23. Haleem A, Javaid M, Singh RP, Suman R. Telemedicine for healthcare: capabilities, features, barriers, and applications. Sens Int 2021;2:100117. [doi: 10.1016/j.sintl.2021.100117] [Medline: 34806053]
24. Monaghesh E, Hajizadeh A. The role of telehealth during COVID-19 outbreak: a systematic review based on current evidence. BMC Public Health 2020 Aug 1;20(1):1193. [doi: 10.1186/s12889-020-09301-4] [Medline: 32738884]
25. Coffey JD, Christopherson LA, Williams RD, et al. Development and implementation of a nurse-based remote patient monitoring program for ambulatory disease management. Front Digit Health 2022 Dec 14;4:1052408. [doi: 10.3389/fdgth.2022.1052408] [Medline: 36588748]
26. Bardhan R, Byrd T, Boyd J. Workforce management during the time of COVID-19—lessons learned and future measures. COVID 2023;3(1):1-27. [doi: 10.3390/covid3010001]
27. Wu Y. Utilization of telehealth and the advancement of nursing informatics during COVID-19 pandemic. Int J Nurs Sci 2021 Oct 10;8(4):367-369. [doi: 10.1016/j.ijnss.2021.09.004] [Medline: 34552807]
28. Zhang T, Mosier J, Subbian V. Identifying barriers to and opportunities for telehealth implementation amidst the COVID-19 pandemic by using a human factors approach: a leap into the future of health care delivery? JMIR Hum Factors 2021 Apr 9;8(2):e24860. [doi: 10.2196/24860] [Medline: 33779566]

Abbreviations

IoT: Internet of Things
IRB: institutional review board

Edited by E Borycki; submitted 07.10.23; peer-reviewed by CY Chang, S Setia; revised version received 02.10.24; accepted 09.12.24; published 21.01.25.

Please cite as:
Chang YL, Lin CY, Hsu J, Liao SL, Yu CT, Peng HC, Chen CY, Ma MHM, Hwang JJ
Leveraging Smart Telemedicine Technology to Enhance Nursing Care Satisfaction and Revolutionize COVID-19 Care: Prospective Cohort Study
JMIR Hum Factors 2025;12:e53456
URL: https://humanfactors.jmir.org/2025/1/e53456
doi:10.2196/53456

© You-Lung Chang, Chi-Ying Lin, Jiun Hsu, Sui-Ling Liao, Chun-Ti Yu, Hung-Chueh Peng, Chung-Yu Chen, Matthew Huei-Ming Ma, Juey-Jen Hwang. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 21.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Examining Individuals' Use of the Internet for Health Care Activities Over Time: Results from the US National Health Interview Survey

Zachary Junkins[1], MS; Nusrath Zahan[1], BS; David Neyens[1], MPH, PhD

Department of Industrial Engineering, Clemson University, Clemson, SC, United States

Corresponding Author:
David Neyens, MPH, PhD
Department of Industrial Engineering
Clemson University
100 Freeman Hall
Clemson, SC, 29634
United States
Phone: 1 8646564719
Email: dneyens@clemson.edu

Abstract

Background: Telehealth is an increasingly important component of health care services. Telehealth services may present an opportunity to increase the equity, accessibility, and effectiveness of health care. As such, it is critical that telehealth design focuses on reducing the barriers to access and usability that may impair some telehealth users.

Objective: Our goal was to identify different demographic characteristics, behaviors, or opinions that may predict groups who are likely to face a barrier to using telehealth services.

Methods: We used data from the National Health Interview Survey and multiple logit regression models focused on different aspects of telehealth to examine three different avenues of telehealth service: looking up health information using the internet, scheduling an appointment using the internet, and communicating with a care provider through email using the internet in order to consider the ways in which different telehealth services may face different barriers.

Results: Our results suggest that middle-aged (36-55 years old) and older adult (56-85 years old) respondents were significantly less likely to look up health information using the internet or schedule an appointment using the internet versus younger individuals (18-35 years old). Specifically, our analysis found that middle-aged adults were found to have a higher odds ratio than older adults (0.83 vs 0.65) for looking up health information using the internet. We also found that there were differences in age groups for using technology to perform health care–related tasks. In terms of searching for health information using the internet and scheduling appointments using the internet, we found differences between men and women, with women being significantly more likely than men to look up health information using the internet, schedule an appointment using the internet, and communicate with a care provider through email using the internet. Across all the investigated variables, we found that the rates of using the internet for looking up health information, scheduling an appointment, and communicating with a care provider over email increased substantially across the study period. The impact of costs was inconsistent across the different models in our analysis. We also found that there is a strong correlation between respondents' collaboration in their personal health and the likelihood that they would use telehealth services to meet these needs.

Conclusions: This analysis provides an exploratory look at the data to highlight barriers that may impact a user's ability to access telehealth services in the context of other potential predictor variables to account for the real-world variability that these may present. Future work should examine the complex relationships of those variables and understand how these interactions are correlated with the respondents' use of telehealth.

(JMIR Hum Factors 2025;12:e58362) doi:10.2196/58362

KEYWORDS

internet; web search; internet search; internet use; searching behavior; access to health information; telemedicine; telehealth; virtual care; virtual health; virtual medicine; logistic regression model; regression model; National Health Interview Survey; NHIS

XSL•FO
RenderX

Introduction

Background

With the advancement of digital technologies, the use of technology in health care has grown [1,2]. While there have been several attempts to better understand the use of digital technologies in health care, much of the work has been done in silos and is sometimes limited in scope [2,3]. Generally, telehealth refers to health care communication through technology, often used in conjunction with telemedicine and eHealth [3-5]. Mobile health is an additional term that describes telehealth services in the context of mobile devices [6]. Telehealth is increasingly being used in growing populations, offering distinct advantages and potential barriers for different patient populations and their care.

Benefits Associated With Telehealth

The increased adoption of telehealth services has allowed for its impacts to be examined on actual patients receiving telehealth care [7]. It has been shown that when telehealth is implemented with current best practices, it may help to balance the health care supply and demand disparity [8], improve patient access to care [8,9], and reduce the cost of care [8,9]. Due to increasing demand with changes in population sizes and demographics, as well as decreasing supply as care providers retire or change careers [8], telehealth services and telehealth programs have helped bridge this gap [10,11]. Telehealth also has been shown to improve patients' ability to access care, thus enabling patients to receive efficient and cost-effective care [8,10] by reducing the impacts of geographical barriers [12] and cost barriers to care [13]. Telehealth may also improve the cost-effectiveness of care, with [14] finding that the "all-cause" cost of providing care to older patients (older than 65 years) decreased from US $937.25 to US $491.52 with the inclusion of telehealth services. Telehealth can also offer support for individuals seeking mental health care where mental health specialties are distant or when there may be patient privacy issues in obtaining mental health care [15].

Barriers to the Use of Telehealth Services

There are several potential barriers that have been identified related to the use and effective implementation of telehealth services [7]. A notable barrier to telehealth services is patients' health literacy [16-18]. Individuals with lower levels of health literacy have shown lower comfort levels with using telehealth technologies [17,19]. Older people are using as well as providing telehealth services [20]. Yet, it has been suggested that older patients (older than 60 years) might have greater difficulty using telehealth services [21]. Additionally, technological infrastructure issues (low-quality, limited internet access) may reduce the ability of individuals to access telehealth services [22]. Individuals' access to electronic devices (ie, smartphones, computers, and tablets) may also be a barrier to engaging in telehealth services [22]. In addition to the access issues, the time and costs required to engage in the services (both for the patients as well as clinicians) may be a barrier to rolling out large programs or integrating them into the workflow processes for clinical staff [23].

Changes in Health Care Populations Over Time

There are several factors that have been shown to impact the use of web-based resources for health care and other telehealth services [24]. A 2018 study suggests that demographic factors, such as living in rural areas, age, and insurance types, impact telehealth implementation and use in health care facilities [24]. There have also been racial and ethnic discrepancies reported in the literature, with Black and Hispanic patients preferring emergency departments to telemedicine compared to White patients [25]. Age is also an important factor as the use of telehealth differs among different age groups [14,25]. In addition, health insurance access, and having internet access also influence the use of telehealth [14].

It is important to consider how user needs with telehealth change over time [26]. Changing population demographics, such as age, have major impacts on the use and function of health care [27]. Other additional factors, such as socioeconomic factors and changes in the prevalence of different health conditions, also impact the function of the health care system [27]. Over time, population demographics, access, and technology can change significantly, leading to incorrect conclusions or missed critical trends due to observing only one moment. It is important to consider all these elements when examining which predictors may indicate barriers to telehealth access. The objective of this research is to examine how individuals used the internet to support their health care over 7 years (2012-2018) using the National Health Interview Survey (NHIS) data. Specifically, we examined how different demographic variables and respondents' perspectives influenced how individuals looked up health information using the internet, scheduled appointments using the internet, and emailed to communicate with a care provider using the internet.

Methods

Ethical Considerations

An ethics board review was not required for this analysis as this study used publicly available data.

Overview

Each year the Centers for Disease Control and Prevention (CDC), a government agency assigned to monitor the general public's behavior and health status conducts the NHIS, which is an annual cross-sectional survey designed to gather data on a variety of health-related topics throughout the United States with oversampling of certain demographic groups in a way that is nationally representative [28]. The survey weighting variable is poststratified based on the US Census data to represent national population characteristics [28]. These survey weights are identified for all combinations of persons within the variable WTFA_SA in the NHIS data. Due to the complex survey sampling strategy, the survey person weights must be used in analyzing the data to avoid substantial bias in the results [28].

For this analysis, the NHIS sample adult files for each year from 2012 to 2018 were used. For the files from 2012 to 2014, the American Standard Code for Information Interchange data file was combined using the stringr, stringi, foreign, and RCurl packages in R (R Foundation for Statistical Computing) to run

XSL•FO
RenderX

the associated Statistical Analysis System statements to produce a .csv version of the sample adult file. For the years 2015-2018, a csv file was provided for the sample adult survey data. We created a single file by combining responses for each year using the R functions rbindlist from the data.table package to match each survey year on the corresponding variable columns. Only identical questions related to our research objectives were included in the combined data set.

All responses recorded as the option "not ascertained" were recoded as NA for all variables. All the variables were recoded by assigning binary values or group responses to facilitate the data analysis. The variables used as outcome variables were recoded to become binary variables: health information seeking using the internet (HIT1A), scheduling health care appointments using the internet (HIT3A), and communication with health care providers through email using the internet (HIT4A). Demographic variables were also recoded, including the respondent's race (RACERPI2), Hispanic ethnicity (HISPANI_I), sex (SEX), region (REGION), marital status (R_MARITL), internet use (AWEBUSE) and the frequency of using the internet (AWEBOFNO and AWEBOFTP), email use (AWEBEML) and the frequency of using email (AWEBMTP), whether a respondent would go to a clinic or doctor's office (AUSUALPL and APLKIND), and where a respondent goes to seek preventative care (AHCPLKND). Additionally, whether a respondent skipped medication doses to save money (ARX12_1), whether a respondent took less medication to save money (ARX12_2), if a respondent had delayed filling a prescription (ARX12_3), preferred low-cost medication (ARX12_4), reported buying prescription drugs from another country to save money (ARX12_5), used alternative therapies to save money (ARX12_6), affordability of prescribed medication (AHCAFYR1), affordability of mental health care or counseling (AHCAFYR2), affordability of dental care (AHCAFYR3), affordability of eyeglasses (AHCAFYR4), and worries about paying medical bills (AWORPAY).

All responses where the data were coded as "not ascertained" or "missing" were recoded as "NAs" for all of the variables and were dropped from the analysis. Whether or not a survey respondent had looked up health information using the internet in the last 12 months was recoded to a binary response of 1 to indicate if the respondent had looked up health information on the internet or 0 if they did not indicate looking up health information on the internet (all other responses). Similarly, all binary variables were recoded in the same manner. A survey respondent's race was identified within the data by the RACERPI2 variable from the NHIS data. The variable was recoded to indicate if the survey respondent reported their race as being White, Black or African American, American Indian or Alaska Native, Asian, or multiple races. A survey respondent's Hispanic ethnicity was recoded to indicate if the survey respondent reported being Hispanic, multiple Hispanic, Puerto Rican, Mexican, Mexican-American, Cuban or Cuban American, Dominican (Republic), Central or South American, other Latin American (type not specified), other Spanish, Hispanic or Latino or Spanish (nonspecific type), Hispanic or Latino or Spanish (type refused), or were not Hispanic (a response of 12, not Hispanic or Spanish origin). A survey respondent's region was identified within the data using the variable REGION from the NHIS data. The variable was recoded to indicate if the survey respondent reported residing in the Northeastern, Midwestern, Southern, or Western United States. A survey respondent's marital status was recoded to indicate if the survey respondent reported living with a spouse or partner or not living with a spouse or partner (all other responses). A survey respondent's internet use was identified in the data using the AWEBUSE variable from the NHIS data. The variable was recoded to indicate if the survey respondent reported using the internet, or not. Additionally, a survey respondent's internet use frequency was identified in the data using the AWEBOFNO and AWEBOFTP variables. The variables were recoded to indicate frequently using the internet (responses of once per day or more frequently) or not using the internet frequently (all other response combinations). The variable APLKIND was recoded to indicate if a respondent goes to a clinic or doctor's office (a response of clinic or health center, or doctor's office or health maintenance organization) or somewhere other than a clinic or doctor's office (all other responses). These variables were combined to indicate if a respondent goes to a clinic or doctor's office when sick, goes somewhere other than a clinic or a doctor's office when sick, or does not indicate going anywhere when sick. Where a respondent goes to seek preventative care was identified using the variable AHCPLKND in the NHIS data. The variable was recoded to indicate if a respondent goes to a clinic or doctor's office for preventative care (a response of clinic or health center, or doctor's office or health maintenance organization) or does not go to a clinic or doctor's office for preventative care (all other responses).

We used logit regression models to examine the relationship between our predictor variables and each of our 3 separate questions about the survey participants' use of telehealth: looking up health information using the internet, scheduling an appointment with a health care provider using the internet, and communicating with a provider over email using the internet. In the field of machine learning and statistics, a wide range of computer models can be used for predicting clinical outcomes such as logit regression, decision trees, artificial neural networks, and Bayesian networks. We have selected Logit regression because it is a well-known statistical fitting model that is frequently used for modeling medical problems where it is needed to identify the relation between a binary response variable and a set of independent predictor variables [29]. The models were built using the svyglm function within the survey package in R version 4.2.1 (R Foundation for Statistical Computing) in R Studio 2023.03.0+386 (Posit PBC). The svyglm function includes the ability to account for a weighting variable in the data to facilitate population estimates. The stepAIC function within the MASS package was used to identify the best-fit model for this data.

Results

Overview

Young adults (18-35 years old) made up between 32.1% (n=10,140, unweighted) and 31.26% (n=6902, unweighted) of

XSL•FO
RenderX

weighted survey respondents in 2012 and 2017, respectively, as shown in Table 1, while older adults (56-85 years old) accounted for 32.37% (n=12,662, unweighted) in 2012 and 35.63% (n=11,483, unweighted) in 2018. Women made up between 51.72% (n=13,867, unweighted) and 51.87% (n=19,252, unweighted) of the weighted survey respondents over the years.

Table 1. Demographic information broken down by year; given as count of response and percentage of weighted totals[a].

Parameter	2012	2013	2014	2015	2016	2017	2018
Age, n (%)							
Younger adult	10,140 (32.10)	10,183 (31.92)	10,431 (31.85)	9230 (31.73)	8662 (31.68)	6902 (31.26)	6177 (31.34)
Middle-aged	11,723 (35.53)	11,521 (35.31)	12,065 (34.64)	10,894 (34.23)	10,094 (33.60)	8176 (33.27)	7757 (33.02)
Older adult	12,662 (32.37)	12,853 (32.77)	14,201 (33.51)	13,548 (34.05)	14,272 (34.72)	11,664 (35.47)	11,483 (35.63)
Sex, n (%)							
Male	15,273 (48.13)	15,440 (48.16)	16,398 (48.20)	15,071 (48.20)	14,991 (48.23)	12,096 (48.24)	11,550 (48.28)
Female	19,252 (51.87)	19,117 (51.84)	20,299 (51.80)	18,601 (51.80)	18,037 (51.77)	14,646 (51.76)	13,867 (51.72)
Race, n (%)							
AIAN[b]	349 (0.82)	360 (0.83)	377 (0.81)	392 (0.95)	357 (1.01)	307 (1.20)	295 (1.13)
Asian	2183 (5.35)	2153 (5.59)	2129 (5.74)	1983 (5.94)	1670 (6.07)	1402 (6.37)	1350 (6.41)
Black or African American	5319 (11.91)	5361 (12.03)	5173 (12.26)	4673 (12.32)	3685 (12.31)	2980 (12.42)	2974 (12.38)
Multiple races	659 (1.67)	662 (1.57)	734 (1.61)	699 (1.74)	687 (1.93)	529 (1.99)	563 (2.32)
White	25,939 (80.25)	25,935 (79.98)	28,209 (79.57)	25,831 (79.05)	26,524 (78.68)	21,472 (78.93)	20,173 (77.75)
Region, n (%)							
Northeast	5774 (18.20)	5645 (17.52)	5919 (17.31)	5580 (17.45)	5590 (18.30)	4348 (18.31)	4143 (17.34)
Midwest	7193 (22.72)	7070 (22.68)	7809 (22.99)	7102 (22.42)	7345 (22.17)	6350 (21.81)	5949 (21.98)
South	12,536 (36.43)	12,813 (36.93)	12,896 (37.24)	11,646 (37.12)	11,487 (35.65)	9860 (36.21)	9312 (36.90)
West	9022 (22.65)	9029 (22.88)	10073 (22.46)	9344 (23.01)	8606 (23.88)	6184 (23.67)	6013 (23.78)
Living with spouse or partner, n (%)	14,371 (48.15)	14,199 (48.27)	15,424 (48.08)	14,213 (48.11)	14,160 (48.07)	11,400 (48.66)	11,031 (48.89)
Unweighted sample size	34,525	34,557	36,697	33,672	33,028	26,742	25,417
Weighted sample size	234,920,670	237,394,354	239,688,457	242,500,657	245,142,225	246,657,271	249,455,533

[a]The table shows the unweighted sample size for each variable and the related percentage once the weights are applied to the data.

[b]AIAN: American Indian or Alaska Native.

The percentage of frequent internet users increased from 58.77% (n=18,016, unweighted) in 2012 to 73.15% (n=17,153, unweighted) in 2018. The percentage of email users increased from 64.46% (n=20,038, unweighted) in 2012 to 74.37% (n=17,635, unweighted) in 2018. The number of individuals who reported living with their spouse or partner ranged from 48.07% (n=14,160, unweighted) of weighted survey respondents in 2016 to 48.89% (n=11,031, unweighted) in 2018.

As shown in Figure 1A, older adult respondents tend to look up health information less than the middle-aged or younger adult groups over the years, while the percentage of the respondents looking up health information using the internet grew at similar rates for all groups. As shown in Figure 1B, men tend to look up health information using the internet less than women across all years of the study. Older adult respondents tended to report scheduling an appointment using less than younger adults or middle-aged respondents (Figure 2A). Across the years included in this study, the proportion of younger and middle-aged respondents tended to grow at similar rates. As shown in Figure 2B, men tended to report scheduling an appointment using the internet less often than women. As shown in Figure 3A, a higher percentage of middle-aged adults tended to report communicating with their care provider through email than older adults; a higher percentage of older adults tended to report it than younger adults. As shown in Figure 3B, men tended to report communicating with their care provider through email less often than women.

XSL•FO
RenderX

Figure 1. Percentage of respondents that looked up health information using the internet (A) for each age group and (B) by sex.

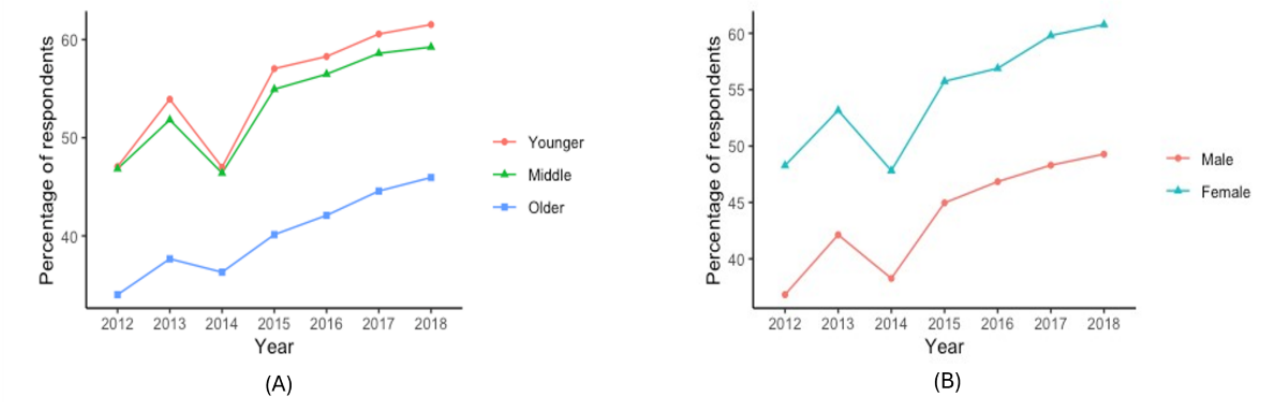


Figure 2. Percentage of respondents that scheduled an appointment using the internet (A) for each age group and (B) by sex.

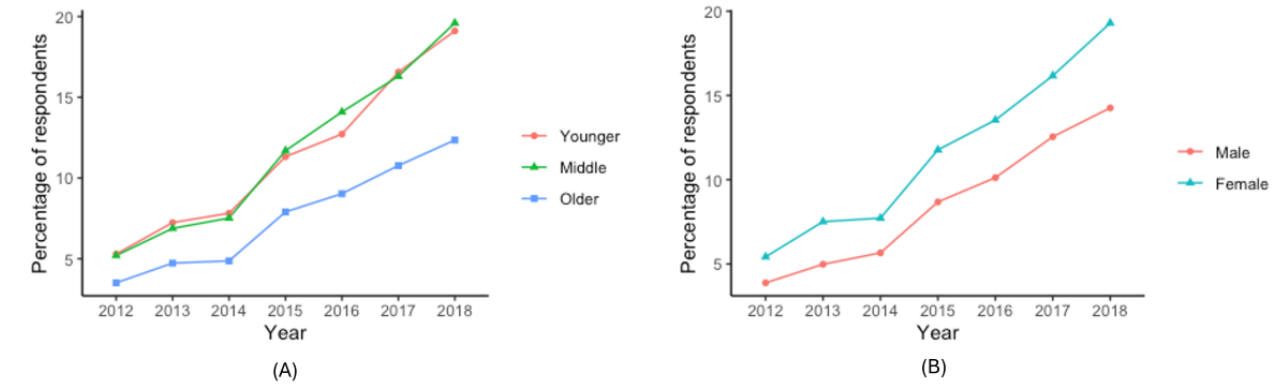
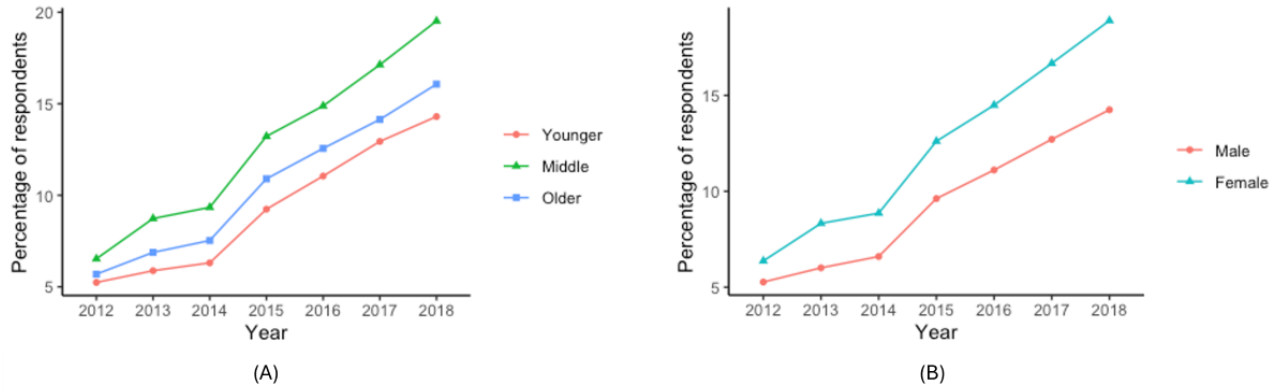


Figure 3. Percentage of respondents that emailed their care provider using the internet (A) for each age group and (B) by sex.



The number of respondents who accessed health information by using the internet increased from 42.76% (n=13,621, unweighted) in 2012 to 55.22% (n=13,677, unweighted) in 2018 (Table 2). In 2018, the percentage of respondents who scheduled an appointment using the internet rose from 4.68% (n=1463, unweighted) in 2012 to 16.86% (n=3962, unweighted). The number of respondents emailing their provider using the internet increased from 5.84% (n=1800, unweighted) in 2012 to 16.65% (n=4176, unweighted) in 2018. The number of respondents seeking medical care who visited a clinic or doctor's office increased from 79.95% (n=27,085, unweighted) in 2012 to 82.03% (n=21,092, unweighted) in 2018. The number of respondents seeking preventive care who visited a clinic or doctor's office increased from 35.96% (n=2706, unweighted) in 2012 to 45.16% (n=2133, unweighted) in 2018. The number of respondents concerned about paying medical bills decreased from 50.24% (n=17,267, unweighted) in 2012 to 43.54% (n=10,637, unweighted) in 2018. The use of alternative therapies to save money increased from 4.07% (n=1454, unweighted) in 2012 to 5.14% (n=1294, unweighted) in 2018.

XSL•FO
RenderX

Table 2. Response variables broken down by year; given as the count of responses and percentage of weighted totals[a].

Parameter	2012	2013	2014	2015	2016	201	2018
Uses internet frequently (>daily), n (%)	18,016 (58.77)	19,699 (63.28)	21,463 (65.04)	19,380 (67.12)	21,013 (69.45)	17,596 (71.05)	17,153 (73.15)
Uses email, n (%)	20,038 (64.46)	21,255 (67.51)	22,697 (68.24)	20,008 (68.84)	21,873 (71.44)	18,198 (72.51)	17,635 (74.37)
Search health information using the internet, n (%)	13,621 (42.76)	15,241 (47.86)	14,783 (43.21)	15,917 (50.55)	16,543 (52.04)	14,097 (54.25)	13,677 (55.22)
Schedule appointment using the internet, n (%)	1463 (4.68)	1973 (6.29)	2099 (6.73)	2974 (10.28)	3542 (11.89)	3543 (14.42)	3962 (16.86)
Email care providers, n (%)	1800 (5.84)	2252 (7.21)	2550 (7.77)	3295 (11.17)	3994 (12.86)	3822 (14.76)	4176 (16.65)
Skip medication to save money, n (%)	2293 (6.31)	1745 (7.72)	1654 (6.93)	1400 (6.07)	1340 (5.78)	1071 (5.96)	1003 (5.71)
Take less medication to save money, n (%)	2418 (6.67)	1875 (8.21)	1733 (7.21)	1519 (6.46)	1390 (6.01)	1118 (6.08)	1074 (5.95)
Request lower-cost medication, n (%)	6320 (18.72)	5186 (24.39)	4918 (21.35)	4259 (19.45)	4460 (19.65)	3456 (18.74)	3436 (19.20)
Uses international medications to save money, n (%)	663 (1.90)	588 (1.57)	554 (1.45)	483 (1.31)	477 (1.62)	397 (1.51)	388 (1.58)
Alternate therapies to save money, n (%)	1454 (4.07)	1559 (4.20)	1587 (4.04)	1358 (3.78)	1411 (4.26)	1161 (4.20)	1294 (5.14)
Cannot afford medications, n (%)	3040 (8.29)	2838 (7.81)	2623 (6.89)	2243 (6.35)	2151 (6.19)	1661 (6.07)	1689 (6.30)
Cannot afford mental care, n (%)	944 (2.51)	798 (2.10)	766 (1.92)	682 (1.87)	686 (1.95)	582 (2.11)	665 (2.65)
Cannot afford dental care, n (%)	4776 (13.23)	4748 (12.85)	4490 (11.60)	3913 (10.90)	3498 (10.24)	2891 (10.77)	2922 (11.26)
Cannot afford vision care, n (%)	2860 (7.80)	2734 (7.33)	2574 (6.47)	2398 (6.40)	1971 (5.64)	1607 (5.81)	1676 (6.33)
Worried about paying medical bills, n (%)	17,237 (50.24)	17,106 (49.76)	16,912 (46.87)	15,136 (45.55)	13,816 (43.72)	11,296 (44.44)	10,637 (43.54)
Goes to a clinic when sick, n (%)							
No where	5660 (16.17)	5325 (15.27)	5033 (13.69)	4518 (13.80)	3941 (13.07)	3167 (12.98)	3137 (13.69)
Somewhere else	1442 (3.88)	1551 (3.90)	1543 (3.97)	1304 (3.59)	1108 (3.27)	947 (3.48)	1047 (4.29)
Goes to clinic	27,085 (79.95)	27,486 (80.83)	29,841 (82.34)	27,587 (82.60)	27,757 (83.66)	22,452 (83.54)	21,092 (82.03)
Visits clinic for preventive care	2706 (35.96)	2478 (33.78)	2607 (35.74)	2495 (39.04)	2640 (44.76)	2191 (46.10)	2133 (45.16)
Unweighted sample size	34,525	34,557	36,697	33,672	33,028	26,742	25,417
Weighted sample size	234,920,670	237,394,354	239,688,457	242,500,657	245,142,225	246,657,271	249,455,533

[a]The table shows the unweighted sample size for each variable and the related percentage once the weights are applied to the data.

Looking Up Health Information Using the Internet

We constructed a model to examine the relationship between the predictor variables and whether or not the survey respondent accessed health information using the internet (Table 3). With each successive year, respondents were more likely to look up health information using the internet (odds ratio [OR] 1.12, 95% CI 1.10-1.15). Middle-aged adults (OR 0.83, 95% CI 0.75-0.93) and older adults (OR 0.65, 95% CI 0.57-0.73) were less likely in comparison to younger adults to look up health information using the internet. Women were more likely to look up health information using the internet in comparison to men (OR 1.69, 95% CI 1.54-1.85). Black or African American respondents were less likely to look up health information in comparison to White respondents (OR 0.77, 95% CI 0.67-0.89). The survey respondents in the South region were less likely to look up health information using the internet compared to those in the

XSL•FO
RenderX

North region (OR 0.73, 95% CI 0.63-0.85). A frequent internet user was more likely to look up health information (OR 3.40, 95% CI 3.00-3.85). Similarly, respondents who used email regularly were more likely to look up health information using the internet than those who reported not using email frequently (OR 3.88, 95% CI 3.41-4.42). Respondents who reported asking their clinicians for lower-cost medications in order to save money were more likely across all years to look up health information (OR 1.55, 95% CI 1.38-1.75). Respondents using alternative therapies in order to save money were almost twice as likely to look up health information using the internet (OR 1.92, 95% CI 1.57-2.36). Respondents who reported going to a clinic or doctor's office when they were sick were more likely to look up health information using the internet than those who did not report going there (OR 1.21, 95% CI 1.08-1.36). Respondents who reported going to a clinic or doctor's office as part of preventative care were also more likely to look up health information using the internet than those who did not seek out preventative care at a clinic or doctor's office (OR 1.18, 95% CI 1.07-1.31).

Table 3. The best-fit model predicting the likelihood that an individual look up health information using the internet.

Parameter	Parameter estimate (SE)	*t* test	*P* value	OR[a] (95% CI)
Intercept	–238.10 (21.92)	–10.86	<.001	N/A[b]
Year	0.12 (0.01)	10.77	<.001	1.12 (1.10-1.15)
Middle-aged	–0.18 (0.05)	–3.35	<.001	0.83 (0.75-0.93)
Older adult	–0.44 (0.06)	–6.81	<.001	0.65 (0.57-0.73)
Female	0.52 (0.05)	11.25	<.001	1.69 (1.54-1.85)
Black or African American	–0.26 (0.07)	–3.62	<.001	0.77 (0.67-0.89)
AIAN[c]	–0.32 (0.24)	–1.33	.18	NS[d]
Asian	–0.28 (0.11)	–2.64	.008	NS
Multiple race	0.00 (0.14)	0.03	.98	NS
Midwest	–0.14 (0.08)	–1.65	.10	NS
South	–0.31 (0.076)	–4.16	<.001	0.73 (0.63-0.85)
West	–0.11 (0.08)	–1.31	.19	NS
Not living with spouse	–0.15 (0.049)	–3.10	.002	NS
Uses internet frequently	1.23 (0.06)	19.29	<.001	3.40 (3.00-3.85)
Uses email	1.36 (0.07)	20.60	<.001	3.88 (3.41-4.42)
Used lower-cost medication to save money	0.44 (0.06)	7.19	<.001	1.55 (1.38-1.75)
Used drugs from other countries to save money	–0.30 (0.15)	–1.95	.05	NS
Used alternate therapies to save money	0.65 (0.10)	6.28	<.001	1.92 (1.57-2.36)
Cannot afford mental care	0.26 (0.11)	2.46	.01	NS
Cannot afford dental care	0.20 (0.07)	2.88	.004	NS
Cannot afford eye care	0.12 (0.08)	1.44	.15	NS
Worried about paying for medical bills	0.08 (0.05)	1.70	.09	NS
Goes somewhere other than a clinic or doctor's office when sick	0.13 (0.07)	1.82	.07	NS
Goes to a clinic or doctor's office when sick	0.19 (0.05)	3.36	<.001	1.21 (1.08-1.36)
Goes to a clinic or doctor's office for preventative care	0.17 (0.05)	3.39	<.001	1.18 (1.07-1.31)

[a]OR: odds ratio.

[b]N/A: not applicable.

[c]AIAN: American Indian or Alaska Native.

[d]NS: parameter was not significant at α=.001.

XSL•FO
RenderX

Scheduling an Appointment Using the Internet

We constructed a model to examine the relationship between the predictor variables and whether or not the survey respondent scheduled an appointment using the internet (Table 4). With each successive year, respondents were more likely to schedule an appointment using the internet (OR 1.30, 95% CI 1.26-1.34). Older adults were less likely in comparison to younger adults to schedule an appointment using the internet (OR 0.63, 95% CI 0.51-0.78). Women were more likely than men to schedule an appointment using the internet (OR 1.61, 95% CI 1.40-1.85). Respondents who use the internet frequently were more likely to schedule an appointment than those not using the internet frequently (OR 2.75, 95% CI 2.02-3.74). Survey respondents using email were more likely than respondents not using email to schedule an appointment using the internet (OR 3.92, 95% CI 2.78-5.53). Respondents who used alternative therapies in order to save money were more likely to schedule an appointment using the internet than those who did not report using alternative therapies to save money (OR 1.70, 95% CI 1.34-2.16). Respondents who were worried about paying medical bills in the last 12 months were less likely than those who did not report being worried about paying medical bills to schedule an appointment using the internet (OR 0.77, 95% CI 0.66-0.89). Respondents going to a clinic or doctor's office for preventative care were more likely to schedule an appointment using the internet than those who did not report going to a clinic or doctor's office for preventative care (OR 1.55, 95% CI 1.33-1.82).

Table 4. The best-fit model predicting the likelihood that an individual schedules a health care appointment using the internet.

Parameter	Parameter estimate (SE)	*t* test	*P* value	OR[a] (95% CI)
Intercept	–537.38 (32.31)	–16.63	<.001	N/A[b]
Year	0.26 (0.02)	16.49	<.001	1.30 (1.26-1.34)
Middle-aged	–0.19 (0.08)	–2.37	.02	NS[c]
Older adult	–0.46 (0.11)	–4.22	<.001	0.63 (0.51-0.78)
Female	0.48 (0.07)	6.63	<.001	1.61 (1.40-1.85)
Black or African American	–0.08 (0.11)	–0.66	.51	NS
AIAN[d]	–0.71 (0.53)	–1.34	.18	NS
Asian	0.40 (0.14)	2.88	.004	NS
Multiple race	0.20 (0.23)	–0.89	.37	NS
Midwest	–0.37(0.12)	–3.01	.003	NS
South	–0.10 (0.11)	–0.92	.36	NS
West	0.03 (0.11)	0.26	.79	NS
Not living with spouse or partner	–0.08 (0.07)	–1.10	.28	NS
Uses internet frequently (>1 per day)	1.01 (0.16)	6.46	<.001	2.75 (2.02-3.74)
Uses email	1.37 (0.18)	7.80	<.001	3.92 (2.78-5.53)
Used lower-cost medication to save money	0.26 (0.09)	2.96	.003	NS
Used alternate therapies to save money	0.53 (0.12)	4.39	<.001	1.70 (1.34-2.16)
Cannot afford prescription medicine	–0.23 (0.14)	–1.67	.09	NS
Cannot afford mental care	0.37 (0.17)	2.21	.03	NS
Cannot afford dental care	–0.28 (0.11)	–2.40	.02	NS
Worried about paying for medical bills	–0.26 (0.08)	–3.46	<.001	0.77 (0.66-0.89)
Goes somewhere other than a clinic or doctor's office when sick	–0.07 (0.12)	–0.55	.58	NS
Goes to a clinic or doctor's office when sick	0.18 (0.08)	2.16	.03	NS
Goes to a clinic or doctor's office for preventative care	0.44 (0.08)	5.50	<.001	1.55 (1.33-1.82)

[a]OR: odds ratio.

[b]N/A: not applicable.

[c]NS: parameter was not significant at α=.001.

[d]AIAN: American Indian or Alaska Native.

XSL•FO
RenderX

Communicating With a Care Provider Through Email Using the Internet

We constructed a model to examine the relationship between the predictor variables and whether or not the survey respondent communicated with a care provider through email using the internet (Table 5). With each successive year, respondents were more likely to communicate with a care provider through email using the internet (OR 1.24, 95% CI 1.20-1.28). Women were more likely than men to communicate with a care provider through email using the internet (OR 1.48, 95% CI 1.28-1.71). Respondents who use the internet frequently were more likely to communicate with a care provider through email than those not using the internet frequently (OR 2.08, 95% CI 1.56-2.76). Survey respondents using email were more likely than respondents not using email to communicate with a care provider through email (OR 5.56, 95% CI 3.84-8.05). Respondents who could not afford mental care in the last 12 months were more likely to communicate with a care provider through email using the internet than those who could afford mental care in the last 12 months (OR 1.92, 95% CI 1.39-2.66). Respondents who reported going to a clinic or doctor's office when they were sick were more likely to communicate with a care provider through email using the internet than those who did not report going there when sick (OR 1.32, 95% CI 1.12-1.55). Respondents who reported going to a clinic or doctor's office as part of preventative care were also more likely to communicate with a care provider through email using the internet than those who did not seek out preventative care at a clinic or doctor's office (OR 1.57, 95% CI 1.34-1.85).

XSL•FO
RenderX

Table 5. The best-fit model predicting the likelihood that an individual communicated with a care provider through email using the internet.

Parameter	Parameter estimate (SE)	*t* test	*P* value	OR[a] (95% CI)
Intercept	–442.69 (32.92)	–13.45	<.001	N/A[b]
Year	0.22 (0.02)	13.29	<.001	1.24 (1.20-1.28)
Middle-aged	0.16 (0.08)	1.88	.06	NS[c]
Older adult	0.18 (0.11)	1.71	.09	NS
Female	0.39 (0.07)	5.34	<.001	1.48 (1.28-1.71)
Black or African American	0.02 (0.13)	0.14	.89	NS
AIAN[d]	–0.37 (0.42)	–0.90	.37	NS
Asian	0.37 (0.15)	2.50	.01	NS
Multiple race	–0.44 (0.24)	–1.80	.07	NS
Midwest	–0.09 (0.13)	–0.72	.47	NS
South	–0.09 (0.12)	–0.79	.43	NS
West	0.32 (0.12)	2.71	.007	NS
Not living with spouse or partner	–0.19 (0.07)	–2.55	.01	NS
Uses internet frequently (>1 per day)	0.73 (0.14)	5.04	<.001	2.08 (1.56-2.76)
Uses email	1.72 (0.19)	9.10	<.001	5.56 (3.84-8.05)
Skipped medication to save money	0.31 (0.20)	1.55	.12	NS
Took less medication to save money	–0.30 (0.21)	–1.43	.15	NS
Used lower-cost medication	0.30 (0.09)	3.18	.002	NS
Used alternate therapies	0.41 (0.13)	3.11	.002	NS
Cannot afford medication	–0.24 (0.16)	–1.50	.13	NS
Cannot afford mental care	0.65 (0.16)	3.98	<.001	1.92 (1.39-2.66)
Cannot afford dental care	–0.30 (0.13)	–2.26	.02	0.74 (0.57-0.96)
Cannot afford eye care	–0.19 (0.17)	–1.14	.26	NS
Worried about a medical bill	–0.18 (0.08)	–2.35	.02	NS
Goes somewhere other than a clinic or doctor's office when sick	0.10 (0.12)	0.85	.39	NS
Goes to a clinic or doctor's office when sick	0.28 (0.08)	3.32	<.001	1.32 (1.12-1.55)
Goes to a clinic or doctor's office for preventative care	0.45 (0.08)	5.50	<.001	1.57 (1.34-1.85)

[a]OR: odds ratio.

[b]N/A: not applicable.

[c]NS: parameter was not significant at α=.001.

[d]AIAN: American Indian or Alaska Native.

Discussion

Principal Findings

This study used NHIS data, which is weighted to represent national characteristics, to examine the use of technology for performing some health care–related tasks over time: looking up health information using the internet, scheduling an appointment using the internet, and communicating with a care provider over email using the internet. Across all our models, we found that middle-aged and older adult respondents were significantly less likely to use technology to look up health information using the internet or schedule an appointment using the internet versus younger individuals. It has been shown that some older adults may need additional special usability requirements due to their inexperience in the use of technology [21], yet this population is a growing user of technology [30]. This also reaffirmed the results found in the existing literature [26]. Across all of the variables we investigated, we found that the rates of looking up health information using the internet, scheduling an appointment using the internet, and communicating with a care provider over email using the internet increased substantially across the study period. This demonstrates that there is an increasing use and need for these

XSL•FO
RenderX

services to support larger populations of users including devices and abilities to engage with technology.

Specifically, our analysis found that middle-aged adults were found to have a higher OR than older adults (0.83 vs 0.65) for looking up health information using the internet. We also found that there were differences in age groups for using technology to perform health care–related tasks. In terms of searching health information and scheduling appointments using the internet, we found differences between men and women, with women being significantly more likely than men to look up health information using the internet, schedule an appointment using the internet, and communicate with a care provider through email using the internet. There are conflicting results in the literature related to sex differences in the use of the internet and technology associated with health-related tasks. Newhouse et al [31] found that men were more likely to use email communication for health care than women; however, Baumann et al [32] found that women were more likely to use the internet and technology for health-related tasks such as scientific literature review, communicating with their physician or medical team, and interpreting the diagnostic test results and medications used in treatment. These sex differences may be related to the social construction and perceptions of technology [33]. Future research needs to explore why these trends are occurring and what factors are associated with the differences between men and women especially since there are conflicting results in the literature about use but also about the factors associated with the differences in use.

The impact of costs was inconsistent across the different models in our analysis. With respect to looking up health information using the internet, there was a significant association with using lower-cost medications and alternative therapies to save money. For individuals scheduling a health care appointment using the internet, respondents who indicated using alternative therapies to save money were more likely to schedule appointments using the internet, and surprisingly those who indicated being worried about paying for medical bills were less likely to schedule appointments using the internet. Individuals using email to communicate with a care provider had two different cost-related variables that were significant. Respondents who reported not being able to afford mental care were more likely to email, while those who indicated not being able to afford dental care were less likely to email. In some ways, the included cost-related variables did not seem to indicate a consistent response around cost. This reinforces the notion that determining the extent to which the cost of care is a barrier is challenging and unclear [34]. By examining the results across several cost-of-care–related variables and different elements of health care, it appears that there may be other barriers whose interaction with cost impacts the effect of cost as a barrier. As suggested by Clarke et al [34], and supported by our results, telehealth research should continue to investigate the ways in which different types of costs are intertwined with the use of telehealth; especially as it relates to different avenues of care.

Going to a clinic or doctor's office for preventative care was associated with a greater likelihood of looking up health information using the internet, schedule a health care appointment using the internet, and communicate with a care provider through email using the internet. Going to a clinic or doctor's office when sick was associated with using technology to look up health information and communicating with a care provider through email. This suggests that there is a strong correlation between respondents' collaboration in their personal health and the likelihood that they would use telehealth services to meet these needs. In fact, Sawesi et al [35] found that information technology platforms can enhance patient engagement and improve health outcomes. It is crucial that telehealth research investigate ways in which telehealth can be used to either support individuals in developing an engagement in their own health or identify ways to encourage users to develop an engagement in their own health.

Limitations and Future Work

There are several factors that limit the generalizability of our analyses. One notable limitation of this research is that it relies on self-reported survey data and may not fully capture specific perspectives and opinions about the "why" for some of the activities reported in the NHIS. Yet, the NHIS data collected and maintained by the US CDC is widely used [36-38] and the data, the data design, and the imputation for national representation are widely documented [39,40]. Future research should expand on these results to try to identify if these trends continue or what factors may be driving the differences identified over time.

This research uses responses from specific questions from 2012 to 2018, as those were the only consecutive years where these questions were consistent. As a result, being able to project current and future use of technology for telehealth is limited. Additionally, the COVID-19 pandemic had a major impact on health care as it presented a unique situation that resulted in overcoming many of the traditional barriers to telehealth adoption [41]. COVID-19 also emphasized the use cases of a variety of different technological tools that can be used in telehealth such as chatbots [42]. As our data was limited to data collected before the COVID-19 pandemic, it is important that future work examine the impact that the pandemic has had on the use of telehealth services and if these effects have been sustained in the following years. Future research should also investigate the ways in which the pandemic might have impacted or altered the trends and barriers examined in this research. For example, it may be interesting to evaluate if the exposure and use of these tools during the pandemic have had lasting impacts on the use of the internet tools. To do this, we need similar data from 2020-2023 or later from the NHIS on the use of telehealth services to be able to apply time series modeling techniques to evaluate specific trends in how using telehealth services have changed over time. Interestingly, some initial work has suggested that there is a decline in the use of telemedicine between 2021 and 2022 in the United States, [43] and future work should evaluate if that trend continues or if the trend prior to the COVID-19 pandemic continues.

In order to develop a fundamental understanding of the important barriers, this research did not examine the complex relationships between demographic and response variables. Future research should focus on understanding how these interactions (eg, older women vs middle-aged women vs

XSL•FO
RenderX

younger women) might be correlated with the respondents' use of telehealth. Our models also did not investigate telehealth use for specific socioeconomic levels. Instead, it includes questions regarding a respondent's ability to afford different health care services, which is different from financial security or socioeconomic conditions [44].

Conclusions

As telehealth is increasingly becoming an important component of health care services, it is important to focus on different aspects of telehealth to determine the demographic characteristics, behaviors, or opinions that may predict or influence groups that are likely to face a barrier to using telehealth services. This study used NHIS data to examine the use of technology for performing some health care–related tasks over time: looking up health information using the internet, scheduling an appointment using the internet, and communicating with a care provider over email using the internet. From this analysis, we have found some potential barriers that may impact a user's ability to access telehealth services, as well as differences in the use of these tools for different groups. Understanding those who are using the internet for health care–related activities and the barriers that they may face is important for the design and implementation of these systems to be as effective as possible. Systems and technology designers, as well as health care providers, should be aware of these differences and the impacts they may have on engaging individuals in their health care.

Conflicts of Interest

None declared.

References

1. van Dyk L. A review of telehealth service implementation frameworks. Int J Environ Res Public Health 2014;11(2):1279-1298 [FREE Full text] [doi: 10.3390/ijerph110201279] [Medline: 24464237]
2. Hu Y. Health communication research in the digital age: a systematic review. J Commun Healthcare 2015;8(4):260-288. [doi: 10.1080/17538068.2015.1107308]
3. Doarn CR, Pruitt S, Jacobs J, Harris Y, Bott DM, Riley W, et al. Federal efforts to define and advance telehealth--a work in progress. Telemed J E Health 2014 May;20(5):409-418. [doi: 10.1089/tmj.2013.0336] [Medline: 24502793]
4. Maheu M, Whitten P, Allen A. E-Health, telehealth, and telemedicine. J Healthcare Qual 2002;24(4):52. [doi: 10.1111/j.1945-1474.2002.tb00449.x]
5. Tuckson RV, Edmunds M, Hodgkins ML. Telehealth. N Engl J Med 2017;377(16):1585-1592. [doi: 10.1056/NEJMsr1503323] [Medline: 29045204]
6. Gogia S. Rationale, history, and basics of telehealth. Fundam Telemed Telehealth 2020:11-34. [doi: 10.1016/b978-0-12-814309-4.00002-1]
7. Hoffman DA. Increasing access to care: telehealth during COVID-19. J Law Biosci 2020;7(1):lsaa043 [FREE Full text] [doi: 10.1093/jlb/lsaa043] [Medline: 32843985]
8. Mahtta D, Daher M, Lee MT, Sayani S, Shishehbor M, Virani SS. Promise and perils of telehealth in the current era. Curr Cardiol Rep 2021;23(9):115 [FREE Full text] [doi: 10.1007/s11886-021-01544-w] [Medline: 34269884]
9. Curfman AL, Hackell JM, Herendeen NE, Alexander JJ, Marcin JP, Moskowitz WB, et al. Telehealth: improving access to and quality of pediatric health care. Pediatrics 2021;148(3):e2021053129 [FREE Full text] [doi: 10.1542/peds.2021-053129] [Medline: 34462339]
10. Finley BA, Shea KD. Telehealth: disrupting time for health care quantity and quality. Nurs Adm Q 2019;43(3):256-262. [doi: 10.1097/NAQ.0000000000000357] [Medline: 31162344]
11. Uscher-Pines L, Rudin RS. Lessons learned from the MAVEN project pilot: using physician volunteers to increase access to care via telehealth. Rand Health Q 2020;9(1):9 [FREE Full text] [Medline: 32742751]
12. Alford-Teaster J, Wang F, Tosteson ANA, Onega T. Incorporating broadband durability in measuring geographic access to health care in the era of telehealth: a case example of the 2-step virtual catchment area (2SVCA) method. J Am Med Inform Assoc 2021;28(11):2526-2530 [FREE Full text] [doi: 10.1093/jamia/ocab149] [Medline: 34414437]
13. Boothe E, Kaplan J. Using telemedicine in Mississippi to improve patient access to genetic services. J Genet Couns 2018;27(2):320-322. [doi: 10.1007/s10897-017-0192-6] [Medline: 29264692]
14. Chen YH, Lin YH, Hung CS, Huang CC, Yeih DF, Chuang PY, et al. Clinical outcome and cost-effectiveness of a synchronous telehealth service for seniors and nonseniors with cardiovascular diseases: quasi-experimental study. J Med Internet Res 2013;15(4):e87 [FREE Full text] [doi: 10.2196/jmir.2091] [Medline: 23615318]
15. Langarizadeh M, Tabatabaei M, Tavakol K, Naghipour M, Rostami A, Moghbeli F. Telemental health care, an effective alternative to conventional mental care: a systematic review. Acta Inform Med 2017;25(4):240-246 [FREE Full text] [doi: 10.5455/aim.2017.25.240-246] [Medline: 29284913]
16. Brown SH, Griffith ML, Kripalani S, Horst SN. Association of health literacy and area deprivation with initiation and completion of telehealth visits in adult medicine clinics across a large health care system. JAMA Netw Open 2022;5(7):e2223571 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.23571] [Medline: 35877125]

XSL•FO
RenderX

17. Irfan A, Lever JM, Fouad MN, Sleckman BP, Smith H, Chu DI, et al. Does health literacy impact technological comfort in cancer patients? Am J Surg 2022;223(4):722-728 [FREE Full text] [doi: 10.1016/j.amjsurg.2021.08.006] [Medline: 34384588]
18. Smith B, Magnani JW. New technologies, new disparities: the intersection of electronic health and digital health literacy. Int J Cardiol 2019;292:280-282 [FREE Full text] [doi: 10.1016/j.ijcard.2019.05.066] [Medline: 31171391]
19. Sullivan-Baca E, Babicz MA, Choudhury TK, Miller BI. The relationship between health literacy and comfort with teleneuropsychology in a veteran sample. Arch Clin Neuropsychol 2022;37(2):292-301 [FREE Full text] [doi: 10.1093/arclin/acab079] [Medline: 34599332]
20. Demiris G, Charness N, Krupinski E, Ben-Arieh D, Washington K, Wu J, et al. The role of human factors in telehealth. Telemed e-Health 2010;16(4):446-453. [doi: 10.1089/tmj.2009.0114] [Medline: 20420540]
21. Dhillon JS, Ramos C, Wunsche BC, Lutteroth C. Designing a web-based telehealth system for elderly people: an interview study in New Zealand. : IEEE; 2011 Presented at: 24th International Symposium on Computer-Based Medical Systems (CBMS); June 27, 2011; Bristol, UK. [doi: 10.1109/cbms.2011.5999157]
22. Zhang T, Mosier J, Subbian V. Identifying barriers to and opportunities for telehealth implementation amidst the COVID-19 pandemic by using a human factors approach: a leap into the future of health care delivery? JMIR Hum Factors 2021;8(2):e24860 [FREE Full text] [doi: 10.2196/24860] [Medline: 33779566]
23. Snoswell CL, Taylor ML, Comans TA, Smith AC, Gray LC, Caffery LJ. Determining if telehealth can reduce health system costs: scoping review. J Med Internet Res 2020;22(10):e17298 [FREE Full text] [doi: 10.2196/17298] [Medline: 33074157]
24. Lin CCC, Dievler A, Robbins C, Sripipatana A, Quinn M, Nair S. Telehealth in health centers: key adoption factors, barriers, and opportunities. Health Aff (Millwood) 2018;37(12):1967-1974. [doi: 10.1377/hlthaff.2018.05125] [Medline: 30633683]
25. Weber E, Miller SJ, Astha V, Janevic T, Benn E. Characteristics of telehealth users in NYC for COVID-related care during the coronavirus pandemic. J Am Med Inform Assoc 2020;27(12):1949-1954 [FREE Full text] [doi: 10.1093/jamia/ocaa216] [Medline: 32866249]
26. Hung L, Lyons JG, Wu C. Health information technology use among older adults in the United States, 2009-2018. Curr Med Res Opin 2020;36(5):789-797. [doi: 10.1080/03007995.2020.1734782] [Medline: 32096650]
27. Haughton B, Stang J. Population risk factors and trends in health care and public policy. J Acad Nutr Diet 2012;112(3 Suppl):S35-S46. [doi: 10.1016/j.jand.2011.12.011] [Medline: 22709860]
28. About the National Health Interview Survey. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/nhis/about_nhis.htm [accessed 2024-10-18]
29. Ayer T, Chhatwal J, Alagoz O, Kahn CE, Woods RW, Burnside ES. Informatics in radiology: comparison of logistic regression and artificial neural network models in breast cancer risk estimation. Radiographics 2010;30(1):13-22 [FREE Full text] [doi: 10.1148/rg.301095057] [Medline: 19901087]
30. Vaportzis E, Clausen MG, Gow AJ. Older adults perceptions of technology and barriers to interacting with tablet computers: a focus group study. Front Psychol 2017;8:1687 [FREE Full text] [doi: 10.3389/fpsyg.2017.01687] [Medline: 29071004]
31. Newhouse N, Lupiáñez-Villanueva F, Codagnone C, Atherton H. Patient use of email for health care communication purposes across 14 European countries: an analysis of users according to demographic and health-related factors. J Med Internet Res 2015;17(3):e58 [FREE Full text] [doi: 10.2196/jmir.3700] [Medline: 25798912]
32. Baumann E, Czerwinski F, Reifegerste D. Gender-specific determinants and patterns of online health information seeking: results from a representative German health survey. J Med Internet Res 2017;19(4):e92 [FREE Full text] [doi: 10.2196/jmir.6668] [Medline: 28377367]
33. Noone JH, Stephens C. Men, masculine identities, and health care utilisation. Sociol Health Illn 2008;30(5):711-725 [FREE Full text] [doi: 10.1111/j.1467-9566.2008.01095.x] [Medline: 18564976]
34. Clarke M, Fursse J, Brown-Connolly NE, Sharma U, Jones R. Evaluation of the National Health Service (NHS) direct pilot telehealth program: cost-effectiveness analysis. Telemed e-Health 2018;24(1):67-76. [doi: 10.1089/tmj.2016.0280] [Medline: 28723244]
35. Sawesi S, Rashrash M, Phalakornkule K, Carpenter JS, Jones JF. The impact of information technology on patient engagement and health behavior change: a systematic review of the literature. JMIR Med Inform 2016;4(1):e1 [FREE Full text] [doi: 10.2196/medinform.4514] [Medline: 26795082]
36. Peprah E, Gyamfi J, Lee JT, Islam F, Opeyemi J, Tampubolon S, et al. Analysis of the 2007-2018 National Health Interview Survey (NHIS): examining neurological complications among children with sickle cell disease in the United States. Int J Environ Res Public Health 2023;20(12):6137 [FREE Full text] [doi: 10.3390/ijerph20126137] [Medline: 37372724]
37. Blumberg SJ, Parker JD, Moyer BC. National Health Interview Survey, COVID-19, and online data collection platforms: adaptations, tradeoffs, and new directions. Am J Public Health 2021;111(12):2167-2175. [doi: 10.2105/AJPH.2021.306516] [Medline: 34878857]
38. Gyamfi J, Tampubolon S, Lee JT, Islam F, Ojo T, Opeyemi J, et al. Characterisation of medical conditions of children with sickle cell disease in the USA: findings from the 2007-2018 National Health Interview Survey (NHIS). BMJ Open 2023;13(2):e069075 [FREE Full text] [doi: 10.1136/bmjopen-2022-069075] [Medline: 36854589]

XSL•FO
RenderX

39. Blewett LA, Dahlen HM, Spencer D, Rivera Drew JA, Lukanen E. Changes to the design of the National Health Interview Survey to support enhanced monitoring of health reform impacts at the state level. Am J Public Health 2016;106(11):1961-1966 [FREE Full text] [doi: 10.2105/AJPH.2016.303430] [Medline: 27631739]
40. Kindratt TB. National Health Interview Survey. Mavs Open Press. 2022. URL: https://uta.pressbooks.pub/bigdataforepidemiology/chapter/chapter6-nhis/ [accessed 2024-08-09]
41. Hoffman DA. Increasing access to care: telehealth during COVID-19. J Law Biosci 2020;7(1):lsaa043 [FREE Full text] [doi: 10.1093/jlb/lsaa043] [Medline: 32843985]
42. Espinoza J, Crown K, Kulkarni O. A guide to chatbots for COVID-19 screening at pediatric health care facilities. JMIR Public Health Surveill 2020;6(2):e18808 [FREE Full text] [doi: 10.2196/18808] [Medline: 32325425]
43. Lucas JW, Wang X. Declines in telemedicine use among adults: United States, 2021 and 2022. Natl Health Stat Report 2024(205):1-11 [FREE Full text] [Medline: 38912919]
44. Williams C, Shang D. Telehealth usage among low-income racial and ethnic minority populations during the COVID-19 pandemic: retrospective observational study. J Med Internet Res 2023;25:e43604 [FREE Full text] [doi: 10.2196/43604] [Medline: 37171848]

Abbreviations

CDC: Centers for Disease Control and Prevention
NHIS: National Health Interview Survey
OR: odds ratio

Edited by K Adapa; submitted 13.03.24; peer-reviewed by S Sawesi, EP Greenup; comments to author 11.07.24; revised version received 09.08.24; accepted 23.08.24; published 26.02.25.

Please cite as:
Junkins Z, Zahan N, Neyens D
Examining Individuals' Use of the Internet for Health Care Activities Over Time: Results from the US National Health Interview Survey
JMIR Hum Factors 2025;12:e58362
URL: https://humanfactors.jmir.org/2025/1/e58362
doi:10.2196/58362
PMID:

©Zachary Junkins, Nusrath Zahan, David Neyens. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 26.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Development and Systematic Evaluation of a Progressive Web Application for Women With Cardiac Pain: Usability Study

Monica Parry[1], RN-EC, PhD; Tony Huang[1], BScH; Hance Clarke[2,3], MD, PhD, FRCPC; Ann Kristin Bjørnnes[4], RN, PhD; Paula Harvey[3,5], BMBS, PhD, FRACP; Laura Parente[6], BDES; Colleen Norris[7], RN, PhD; Louise Pilote[8], MD, PhD, FRCPC; Jennifer Price[5], RN, PhD; Jennifer N Stinson[1,9], RN-EC, PhD; Arland O'Hara[1], BSc; Madusha Fernando[1], BSc; Judy Watt-Watson[1], RN, PhD; Nicole Nickerson[10]; Vincenza Spiteri DeBonis[11], MES; Donna Hart[11], BA; Christine Faubert[11]

[1]Lawrence Bloomberg Faculty of Nursing, University of Toronto, Toronto, ON, Canada
[2]Pain Research Unit, University Health Network, Toronto, ON, Canada
[3]University of Toronto, Toronto, ON, Canada
[4]Department of Nursing and Health Promotion, Oslo Metropolitan University, Oslo, Norway
[5]Women's College Hospital, Toronto, ON, Canada
[6]Healthcare Human Factors, University Health Network, Toronto, ON, Canada
[7]Faculty of Nursing, University of Alberta, Edmonton, AB, Canada
[8]Department of Medicine, McGill University, Research Institute of the McGill University Health Centre, Montreal, QC, Canada
[9]Research Institute, The Hospital for Sick Children, Toronto, ON, Canada
[10]Patient Advisor, LaHave, NS, Canada
[11]Patient Advisor, Toronto, ON, Canada

Corresponding Author:
Monica Parry, RN-EC, PhD
Lawrence Bloomberg Faculty of Nursing
University of Toronto
Office 276
155 College Street
Toronto, ON, M5T 1P8
Canada
Phone: 1 416 946 3561
Fax: 1 416 978 8222
Email: monica.parry@utoronto.ca

Abstract

Background: Cardiac pain has been widely considered to be the primary indicator of coronary artery disease. The presentation of cardiac pain and associated symptoms vary in women, making it challenging to interpret as cardiac, possibly cardiac, or noncardiac. Women prefer to consult with family and friends instead of seeking immediate medical care.

Objective: This study aimed to assess the user performance (ie, ease of use, efficiency, and errors) and user satisfaction (System Usability Scale; SUS) of a progressive web application for women with cardiac pain.

Methods: Following ethics approval, a purposive sample of women aged >18 years with cardiac pain or associated symptoms lasting >3 months and able to speak and read English was recruited to participate in 2 iterative usability testing cycles. The first cycle assessed the performance of and satisfaction with *at heart* using a web application, and the second cycle assessed the performance of and satisfaction with *at heart* across various Android and iOS devices. In total, 2 investigators recorded user comments and documented problems. At the end of the testing session, the participants completed the SUS and 4 semistructured interview questions.

Results: In total, 10 eligible women participated in usability testing from March 31, 2020, to April 17, 2020 (cycle 1), and from November 17, 2020, to November 30, 2020 (cycle 2). Women across usability testing cycles had a mean age of 55.6 (SD 7.3) years, and most (9/10, 90%) were well educated. In total, 50% (5/10) were employed full or part time, and 60% (6/10) earned >CAD $70,000 (US $48,881.80) annually. Participants across 2 testing cycles reported the overall usability of the *at heart*

XSL•FO
RenderX

progressive web application as highly acceptable (mean SUS score 81.75, SD 10.41). In total, 90% (9/10) of participants rated the user-friendliness of *at heart* as good or excellent. All participants (10/10, 100%) thought *at heart* was easy to use and efficient. Only 2 testing errors were noted as high priority; these were low contrast or small font and clarification that the chatbot was not a real person. User satisfaction was assessed using themes that emerged from the debrief and 4 semistructured interview questions; *at heart* was engaging, comprehensive, understandable, credible, relevant, affirming, personalized, and innovative.

Conclusions: This study provides initial support for the *at heart* progressive web application for women living with cardiac pain and symptoms. Ongoing evaluations in phases 3 and 4 should aim to examine the feasibility and acceptability of and the extent of engagement with the *at heart* core feature set: Heart Check, Wellness Check, and the library. In addition to assessing effectiveness in the phase-4 effectiveness-implementation hybrid trial (type I), describing and better understanding the context for implementation (eg, race and ethnicity and geography) will be necessary.

International Registered Report Identifier (IRRID): RR2-10.1136/bmjopen-2019-033092

(JMIR Hum Factors 2025;12:e57583) doi:10.2196/57583

KEYWORDS

digital health; chatbot; women; cardiac pain; usability testing; self-management; artificial intelligence; AI

Introduction

Background

Cardiovascular diseases constitute the leading cause of mortality worldwide, exerting a substantial economic burden on the health care system [1-4]. As the most prevalent form of cardiovascular disease, coronary artery disease (CAD) is estimated to be associated with 8.93 million deaths annually worldwide across all ages [1]. CAD is a complex condition that varies in clinical presentation across sexes, with both obstructive (macrovascular) and nonobstructive (microvascular) CAD being associated with cardiac pain and other symptoms [5]. Although cardiac pain has been widely considered to be the primary indicator of CAD [6], cardiac pain and associated symptoms reported by women with CAD differ markedly from those reported by men [7,8]. The presentation of cardiac pain and associated symptoms may vary in frequency, pattern, and distribution in women [8], making it challenging to interpret as cardiac specific [9]. Similarly, women that undergo a percutaneous coronary intervention or cardiac surgery report higher prevalence of persistent pain of moderate to severe intensity after treatment as compared to men [10-12]. This cardiac pain is often described by women as sharp and burning and may present with varying extent of dyspnea, fatigue, anxiety, and discomfort that may radiate to the jaws, shoulders, back, and arms [6,8,13]. The manifestations of cardiac pain and associated symptoms contribute to substantial morbidity and impairments in health-related quality of life (HRQoL) in women [14,15].

The complex presentation of CAD in women poses a challenge for timely recognition and management of symptoms. Recent data have shown that women often delay seeking medical care when experiencing acute cardiac pain or associated symptoms, with time between symptom onset and emergency department arrival being 85 to 320 minutes [16]. Women report experiencing difficulties in interpreting, understanding, and attributing their cardiac pain or symptoms to CAD, preferring to consult with family and friends instead of seeking immediate medical care [17]. Moreover, women often hesitate to seek medical care for cardiac pain or associated symptoms as a result of their gendered roles as caregivers [17,18]. Women describe having gendered roles that they are unable to delegate, such as providing care for dependent family members [17,18]. The cumulative effect of symptom underrecognition and hesitancy in women leads to delayed care-seeking behaviors and an increased risk of major adverse cardiac events or mortality as compared to men [6]. As such, it is imperative to promote proper recognition, assessment, and management of symptoms in women to improve health outcomes and HRQoL.

Self-management programs are designed to engage users as active participants in the management of their conditions; they are key predictors of successful behavior change [19,20]. These interventions generally use educational strategies designed to assist users in achieving optimal knowledge, understanding beliefs, and skills, as well as providing meaningful social supports [21]. Digital health–based self-management programs have been developed and effectively used to help women manage weight [22-25], increase physical activity [26], monitor for perinatal depression, and assist with postpartum smoking cessation [27]. Many women describe digital health interventions as being novel and supportive [22] and effective in motivating healthy behaviors, reducing symptoms [28], and improving HRQoL [23]. However, there is a lack of evidence-informed digital health self-management programs specifically for women with CAD living with cardiac pain and associated symptoms [7], demonstrating a clear need for a digital self-management program.

At heart (formerly HEARTPA♀N) [29], a self-management progressive web application, was developed for women with CAD using a sequential phased approach recommended by the Medical Research Council (MRC) [30-32]. In phase 1, an integrated mixed methods systematic review was conducted to evaluate the current evidence related to the self-management of cardiac pain and associated symptoms (eg, dyspnea and fatigue) in women [7,29]. The results of the review suggested that self-management interventions could reduce cardiac pain and associated symptoms if they targeted a greater proportion of women (standardized mean difference [SMD]=−0.01; SE 0.003; *P*=.02), goal setting (SMD=−0.26, 95% CI −0.49 to −0.03), and collaboration or support from a health care provider (SMD=−0.57, 95% CI −1.00 to −0.14) [33]. The review also identified a lack of self-management interventions targeted specifically for cardiac pain and associated symptoms in women

XSL•FO
RenderX

[33]. In phase 2A, the content and core feature set, chatbot, and symptom triage algorithms were co-designed with health care professionals and women with CAD [29]. In phase 2B, the usability of *at heart* was evaluated to ensure that the platform was intuitive and acceptable for women with cardiac pain, and it is the focus of this paper. In phase 3, a process and preliminary efficacy evaluation using a 2-group parallel pilot randomized controlled trial (RCT) will be undertaken, and then a phase-4 effectiveness-implementation hybrid trial is planned (Figure 1).

Figure 1. Phases of the *at heart* progressive web application development.



Objectives

The objectives of this study were to assess the user performance (ie, ease of use, efficiency, and errors) and user satisfaction (interviews and System Usability Scale [SUS]) of a progressive web application for women with cardiac pain.

Methods

Participant Selection

Following ethics approval, a purposive sample of women was recruited from (1) an ambulatory care hospital focused on women's health, (2) an adult tertiary care transitional pain clinic; (3) the Alberta Provincial Project for Outcome Assessment in Coronary Heart Disease registry; (4) the CorHealth Ontario Cardiac Registry; (5) established strategies through the Canadian Pain Coalition; (6) the Ontario Women's Health Network's listserve, which reaches >1900 women and community organizations in Ontario, Canada; and (7) social media. The Ontario Women's Health Network regularly brought women who lived in rural and remote areas together, including women who were disabled, women of color, women with a low income, and Indigenous and older women. Women were eligible to participate in the usability testing if they (1) were aged >18 years, (2) had been diagnosed with obstructive or nonobstructive CAD pain or pain after a percutaneous coronary intervention or cardiac surgery lasting >3 months, (3) were able to speak and read English, and (4) had not previously used or tested the *at heart* progressive web application. Women were excluded from the study if they had (1) severe cognitive impairment, assessed using the Six-Item Screener administered via telephone; or (2) a major comorbid medical or psychiatric condition that would preclude their ability to participate in the usability testing. The Six-Item Screener is a brief instrument for identifying individuals with cognitive impairment, and its diagnostic properties are similar to those of the Mini-Mental State Examination [38,39].

Usability testing is an important phase in the development of digital health interventions [34,35]. End users test a prototype in iterative cycles; they provide feedback about what works, what does not work, and where gaps might exist in the information and functionality [34,35]. These factors contribute to the frequency of use, understanding, and acceptability and enhance the likelihood that users will use the end product [35,36]. Testing the usability of the intervention also serves to assess the suitability of the platform interface and content [37].

Ethical Considerations

The phase-2B usability testing was approved by the Health Sciences Research Ethics Board (REB; 36415) at the University of Toronto on November 26, 2018, and the Health Sciences and Affiliated Teaching Hospitals REB (6026830) at Queen's University on June 20, 2019. An amendment was granted by the Health Sciences REB (36415) at the University of Toronto on March 26, 2020, to use secure video- or web conferencing (Zoom; Zoom Video Communications) for the usability testing sessions due to COVID-19 pandemic restrictions, which prohibited in-person participant contact. Informed consent and a demographic and clinical information form (Multimedia Appendix 1) were obtained from participants before each iterative cycle. Participants were identified using a study ID, and data were stored on a secure OneDrive folder (Microsoft Corp) accessible only to the principal investigator and research officer. All participants were compensated with a user-identified CAD $25 (US $17.52) gift card.

Study Design and Procedures

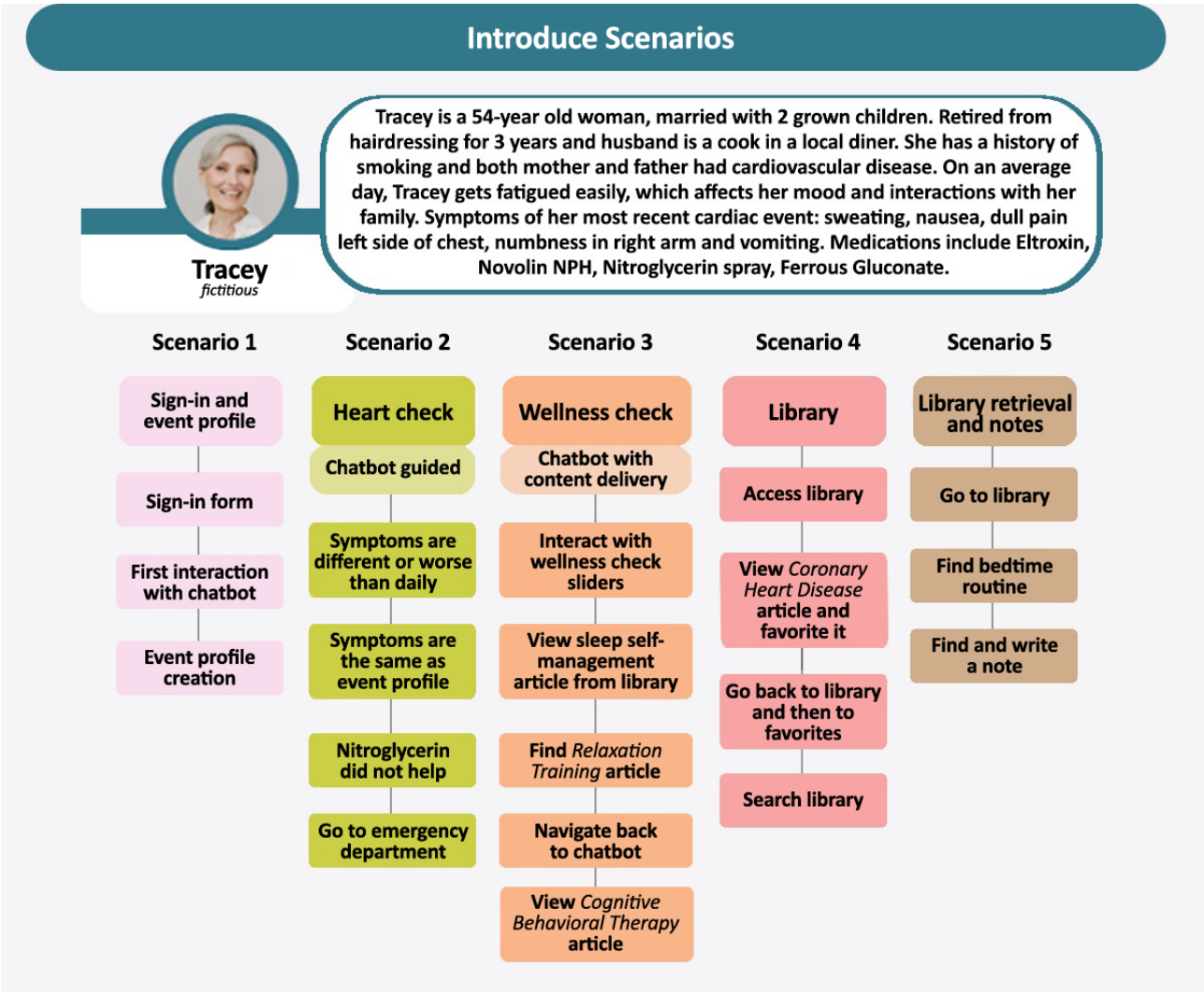
The usability of the *at heart* progressive web application focused on a think-aloud scenario-based approach to assess user performance (ie, ease of use, efficiency, and errors) and user satisfaction with *at heart's* content and functionality assessed qualitatively through the short scenario debrief questions and the interviews and quantitatively using the SUS [40]. This multi-method approach to usability testing is supported by a recent scoping review [41]. A sample of 10 eligible women was recruited to participate in the iterative usability testing cycles (n=5, 50% of the women participated in each of 2 testing cycles). This sample size was based on the experience of others [42-44], as well as recommendations that usability testing by 3 to 5 users can find approximately 85% of interface usability problems [45,46]. The first cycle tested user performance and satisfaction on a desktop computer, and the second cycle assessed user performance and satisfaction across various Android and iOS devices (ie, smartphones and tablets). Participants were provided with a brief explanation of the study and the *at heart* progressive web application before undergoing a 60- to 90-minute one-on-one observation period conducted and recorded through

XSL•FO
RenderX

the Zoom videoconferencing platform. Participants were introduced to a case of a woman aged 54 years with symptoms that were "possibly cardiac" [47] and asked to sign in, complete an event profile (scenario 1), and then progress through a set of standardized scenarios that incorporated each core feature of *at heart*: Heart Check, Wellness Check, and the library (scenarios 2-5; Figure 2).

The "think-aloud" approach was used to capture users' thought process and problem-solving as they progressed through *at heart's* core features in a systematic manner. After each scenario, the participants answered two short debrief questions: (1) "What did you like about the content and functionality of this specific scenario?" (2) "What did you dislike about the content and functionality of this specific scenario?"

Figure 2. Usability testing workflow.



In total, 2 investigators also recorded user comments and documented any problems encountered during each scenario using the usability testing error and efficiency documentation form (Multimedia Appendix 2). Three error types were captured on this form: (1) navigation errors (ie, failure to locate a function or follow the recommended screen flow), (2) presentation errors (ie, selection errors due to labeling ambiguities), and (3) control use errors (ie, improper entry field errors). Four semistructured interview questions were posed following the think-aloud scenarios to assess the overall satisfaction with *at heart's* content and functionality: (1) "What was your overall impression of the *at heart* progressive web application?" (2) "What did you like or dislike about the *at heart* progressive web application?" 3) "Is there anything that could be improved or changed?" (4) "Is there anything missing from the *at heart* progressive web application?" Participants then completed the SUS on the web using REDCap (Research Electronic Data Capture; Vanderbilt University), a secure web application. The SUS is a widely used scale that quantifies the usability of digital health applications. It consists of 10 questions with Likert scales that range from *strongly agree* to *strongly disagree*, and it has been validated across a range of interfaces, including web pages and web applications [40,48,49]. Scores range from 0 to 100, with an accepted benchmark mean SUS score of 68 (SD 12.5) [40]. One 7-point adjective-anchored Likert scale was added to the bottom of the SUS so that participants could rate the overall user-friendliness of *at heart* as *worst imaginable*, *awful*, *poor*, *ok*, *good*, *excellent*, or *best imaginable*. This adjective rating scale helps inform the absolute usability of a product [50-52], such as the *at heart* progressive web application. Numerical equivalents of 1 (*worst imaginable*) to 7 (*best imaginable*) were assigned to the adjectives for scoring. After the first cycle,

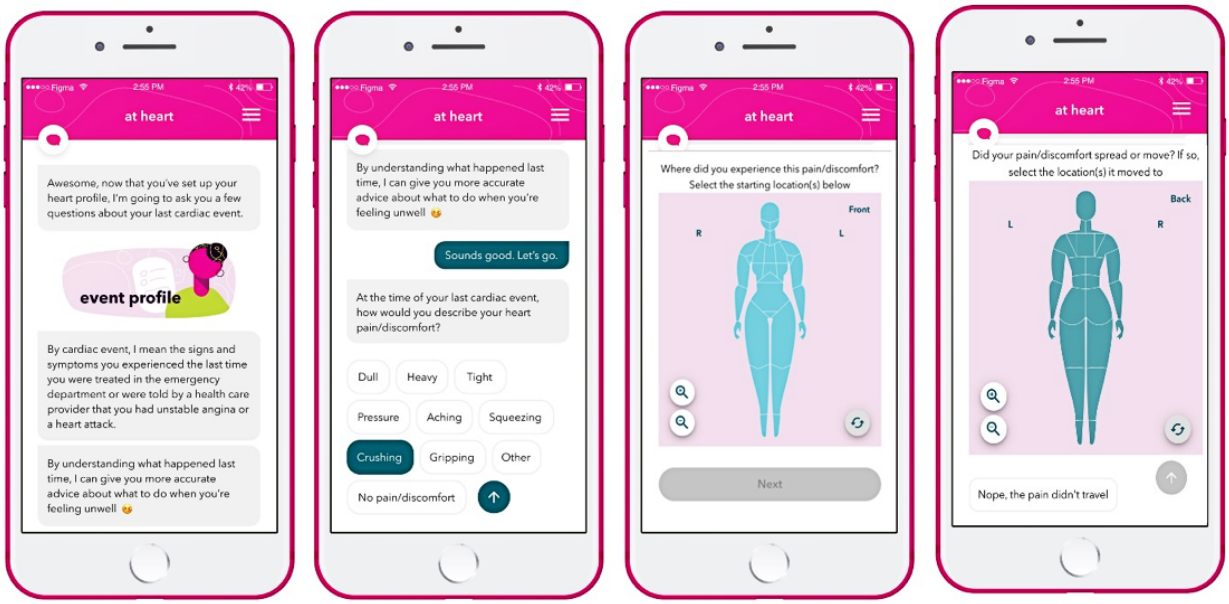
XSL•FO
RenderX

revisions were made to the prototype, and the revised prototype was tested in a second usability testing cycle.

At Heart Progressive Web Application

At heart is a novel progressive web application that consists of 3 core features: Heart Check, Wellness Check, and a library. Users are guided through the core feature set by a rule-based chatbot that manages content and conversations. An Event Profile is created on the initial log-in to the progressive web application (Figure 3). The Event Profile contains individualized or personalized data that include the quality and location of cardiac pain or associated symptoms experienced at the time of the participant's last heart event (ie, treated in the emergency department). Female front and back full-body maps were specifically developed for *at heart* using the chest pain or associated symptom locations most commonly described in the literature [6,8,13,47,53].

Figure 3. Screenshots of the *at heart* feature Event Profile prototype.



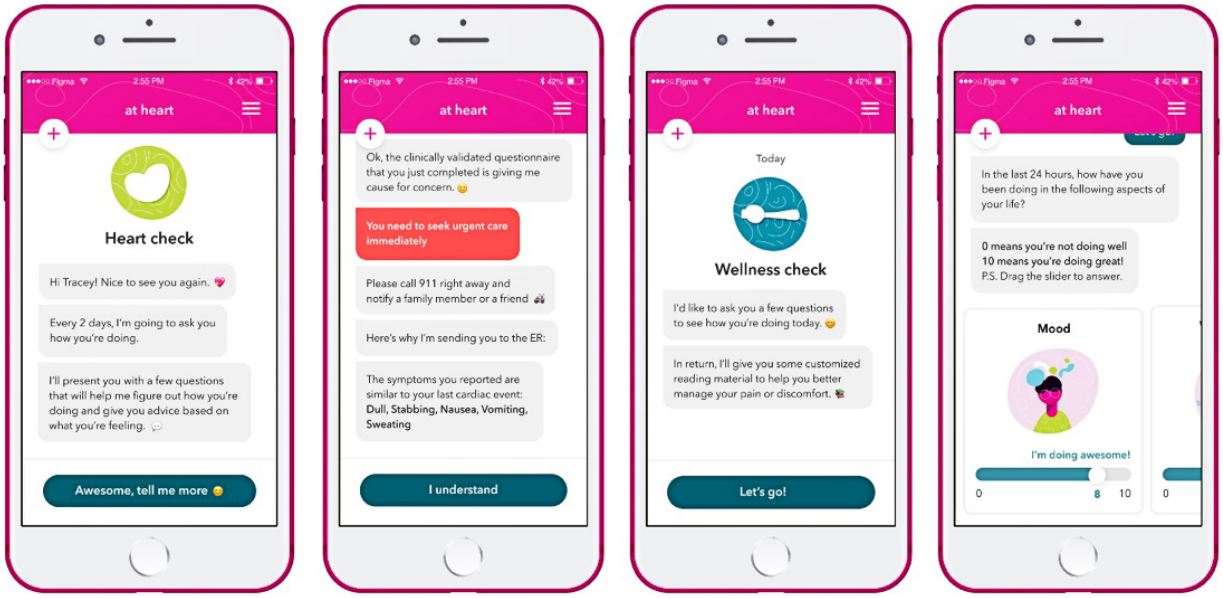
Chatbot Holly asks women the same series of cardiac pain and symptom assessment questions in the Heart Check, which is programmed into the *at heart* progressive web application to occur every 3 days (Figure 2). The Heart Check can occur more frequently simply by asking chatbot Holly for a Heart Check. Rule-based symptom triage algorithms compare each Heart Check to the individually stored Event Profile to make recommendations on an appropriate level of care. Level of care recommendations are based on the similarity and timing of current cardiac pain or discomfort to that stored Event Profile, noting similarities between the current and previous descriptions of the quality (ie, heavy, tight, or pressure), location, radiation, and associated symptoms. Recommendations are assigned to one of three categories: (1) *red* or high risk (ie, similar or high-risk symptoms that have occurred within 24 hours), (2) *yellow* or moderate risk (ie, similar or high-risk symptoms that have occurred beyond 24 hours but within 7 days), and (3) *green* or low risk (ie, no similar or high-risk symptoms). Women who are categorized as *red* are encouraged to notify a family member of their current symptoms and call 911 (ie, to seek appropriate and urgent assessment of their symptoms). Women who are categorized as *yellow* are encouraged to see their primary care provider within 48 hours, and women who are categorized as *green* are permitted access to the *at heart* progressive web application. High-risk symptoms are defined as ≥3 new typical symptom features (ie, dull, heavy, or tight chest pain) [53] or any new associated symptoms, including shortness of breath; palpitations; a racing heart rate; or feeling lightheaded, faint, or dizzy [47].

Women complete the Wellness Check, informed using features of the Brief Pain Inventory–Interference Subscale [54], to indicate the degree to which cardiac pain or associated symptoms interfere with 7 domains of life, which include general activities, paid and unpaid work, walking, mood, relations with others, sleep, and overall enjoyment of life (Figure 4). Wellness Checks are programmed to occur every 7 days at a minimum, and if the score in any domain is below a threshold, the chatbot uses rule-based algorithms to deliver educational content from the library.

The library contains scientific papers, each with a lay summary, covering educational content about CAD in women, with self-management advice to promote good health and well-being [29]. The library also contains videos and podcasts from women with lived experiences of CAD. *At heart* was evaluated by Health Canada and deemed not to be a medical device as defined in the Guidance Document: Software as a Medical Device (SaMD): Classification Examples [55]. *At heart* provides self-management support to women who have CAD and, through a chatbot, guides women to the most appropriate form of assessment (ie, primary care and emergent care) based on their medical symptoms.

XSL•FO
RenderX

Figure 4. Screenshots of the Heart Check and Wellness Check features of the *at heart* prototype.



Data Analysis

Quantitative descriptive data from the demographic and clinical information form and the SUS were analyzed using SPSS Statistics (version 29; IBM Corp). Debrief feedback from each of the 5 scenarios with data from the usability testing error and efficiency documentation form was reviewed after each of the 2 iterative usability testing cycles. Conventional content analysis as outlined by Sandelowski [56] was used to analyze qualitative data obtained from the interview questions. In total, 2 researchers independently assigned codes to label key actions, thoughts, and ideas. Related codes were grouped into categories, and any disagreements were discussed between the 2 researchers. A list of prototype issues was generated and prioritized based on risk to patients and ease of improvement [57]. Modifications were communicated to the website design and development team. Prototype changes made after the first cycle were evaluated in the second cycle, and prototype changes following the second cycle were completed before the phase-3 RCT.

Results

Participant Characteristics

A total of 10 women meeting the eligibility criteria were invited to participate in the usability testing from March 31, 2020, to April 17, 2020 (cycle 1), and from November 17, 2020, to November 30, 2020 (cycle 2). Women across the usability testing cycles had a mean age of 55.6 (SD 7.3) years, and most had a diploma (5/10, 50%), an undergraduate degree (2/10, 20%), or a graduate degree (2/10, 20%). Most of the women were employed full (3/10, 30%) or part time (2/10, 20%) and earned >CAD $70,000 (US $48,881.80) annually (6/10, 60%); 50% (5/10) lived with a disability, 50% (5/10) were the primary household earners, and 80% (8/10) were the primary persons responsible for housework in the home. The women reported cardiac pain or associated symptoms for 1 to 2 years (3/10, 30%), 2 to 5 years (3/10, 30%), or >5 years (3/10, 30%). The most common comorbid conditions included depression (4/10, 40%) and anxiety (4/10, 40%), and 40% (4/10) of the women had had a previous myocardial infarction. Participant demographics by usability testing cycle are reported in Table 1.

All women (10/10, 100%) had a computer at home and were comfortable (3/10, 30%) or very comfortable (7/10, 70%) using a computer, comfortable (1/10, 10%) or very comfortable (9/10, 90%) using the internet, and comfortable (3/10, 30%) or very comfortable (7/10, 70%) using smartphone or tablet apps. Participants' comfort levels and use of the computer, the internet, and smartphone or tablet apps are reported in Table 2. All women selected their own devices for usability testing. Women in the first cycle used only the Google Chrome browser (5/5, 100%), and women in the second cycle used an iPad mini version 9.3.5 (1/5, 20%), Samsung Galaxy S10e (1/5, 20%), iPhone 8 (2/5, 40%), and iPhone 11 (1/5, 20%).

XSL•FO
RenderX

Table 1. Participant characteristics.

Characteristic	Cycle 1 (n=5)	Cycle 2 (n=5)
Age (y), mean (SD)	58.4 (4.827)	52.8 (8.872)
Gender identity (woman), n (%)	5 (100)	5 (100)
Indigenous, n (%)	0 (0)	0 (0)
Identifying as a visible minority, n (%)	1 (20)	1 (20)
Living with a disability, n (%)	2 (40)	3 (60)
Gendered roles (housework hours), mean (SD)	6.6 (5.1)	10.2 (8.0)
Primary person responsible for housework, n (%)	4 (80)	4 (80)
Educational level, n (%)		
High school	0 (0)	1 (20)
Diploma	2 (40)	3 (60)
Bachelor's degree	2 (40)	0 (0)
Master's degree	1 (20)	1 (20)
Employment status, n (%)		
Full time	1 (20)	2 (40)
Part time	1 (20)	1 (20)
Unemployed	3 (60)	2 (40)
Income, n (%)		
<CAD $15,000 (US $10,474.70)	0 (0)	1 (20)
CAD $15,000-$29,900 (US $10,474.70-$20,879.50)	0 (0)	0 (0)
CAD $30,000-$49,900 (US $20,949.40-$34,845.80)	0 (0)	1 (20)
CAD $50,000-$69,900 (US $34,915.60-$48,812)	1 (20)	0 (0)
CAD $70,000-$99,900 (US $48,881.80-$69,761.40)	3 (60)	1 (20)
>CAD $100,000 (US $69,831.20)	1 (20)	1 (20)
Not reported	0 (0)	1 (20)
Primary earner, n (%)	1 (20)	4 (80)

XSL•FO
RenderX

Table 2. Use of the computer, the internet, and smartphone or tablet apps.

Characteristic	Cycle 1 (n=5)	Cycle 2 (n=5)
Computer use—home, n (%)	5 (100)	5 (100)
Computer use—work, n (%)	2 (40)	3 (60)
Computer hours per week, mean (SD)	6.4 (2.302)	8 (0.000)
Comfortable using computers, n (%)		
Comfortable	1 (20)	2 (40)
Very comfortable	4 (80)	3 (60)
Internet use, n (%)	5 (100)	5 (100)
Internet hours per week, mean (SD)	6.8 (2.168)	8 (0.000)
Comfortable using the internet, n (%)		
Comfortable	1 (20)	0 (0)
Very comfortable	4 (80)	5 (100)
Smartphone or tablet app use, n (%)	5 (100)	5 (100)
Smartphone or tablet app hours per week, mean (SD)	6.4 (2.302)	7 (1.414)
Comfortable using smartphone or tablet apps, n (%)		
Comfortable	1 (20)	2 (40)
Very comfortable	4 (80)	3 (60)

User Performance

User performance was assessed through ease of use, efficiency, and observation of testing errors (ie, navigation, presentation, and control use errors) through each scenario and cycle of usability testing. Testing errors included those related to navigation (difficulties moving through or locating content), presentation (selection errors due to labeling), or control use (improper entry field errors) across the 5 scenarios in 2 usability testing cycles (Figure 2). In scenario 1 of cycle 1 (sign-in, chatbot, and Event Profile), participants reported a low contrast between the text and the background and small font at sign-in. Participants also wanted clarification that the chatbot was not a real person. These were flagged as high priority, and subsequent revisions included larger font and chatbot and user text boxes identified using different colors, and a page was added to the *at heart* progressive web application to describe the chatbot, including its development and function. Specific text included the following: “let me tell you a little more about myself. I’m an educated chatbot, thoughtfully designed by a group of women’s heart health experts. I’ll be here to help you understand more about your heart pain/discomfort through regular check-ins, and I’ll send you tips and tricks for managing that pain. If you want to learn more about my makers, the same ones who named me Holly, go to the menu above...” A summary of suggestions and subsequent changes made to the progressive web application during usability testing cycle 1 is shown in Table 3.

The scenarios were repeated in the second cycle to assess user performance and user satisfaction across various Android and iOS devices (ie, smartphones and tablets). In scenario 1 of cycle 2 (sign-in, chatbot, and Event Profile), participants requested clarification on the terminology used in some of the questions at sign-in, so descriptors and more answers were added to these questions. Participants also had difficulty choosing specific areas on the body map, so the body maps were subsequently amended. In scenario 2 of cycle 2 (chatbot-guided workflow for the Heart Check), participants appeared confused with the order and speed of the questions delivered by the chatbot. The chat time stamp was subsequently increased for desktop, Android, and iOS devices. This allowed women ample time to read and respond appropriately to the Heart Check questions, especially those with longer messages and prompts. Finally, the Heart Check button was reported to be too small and gray, so the size and contrast were adjusted. The chatbot-guided workflow for the Wellness Check (scenario 3, cycle 2) required minor adjustments to ensure that the chatbot selected appropriate resources from the Wellness Library based on wellness scores of ≤4. One participant suggested that “anaphylaxis” be added to the Share My Data section and related articles be added to the library. As a result, lay summaries were generated, and scientific papers were added to the library for Kounis syndrome (n=3) and mast cell activation syndrome (n=2). Kounis syndrome is a complex multisystem arterial condition caused by mast cell activation and T-lymphocyte and macrophage interactions [58]. The resulting allergic, hypersensitivity, and anaphylactic reactions cause coronary symptoms (ie, cardiac pain), and this is referred to as Kounis syndrome [59]. A summary of suggestions and subsequent changes made to the progressive web application during usability testing cycle 2 is also shown in Table 3.

XSL•FO
RenderX

Table 3. Summary of suggestions and subsequent changes during usability testing.

Suggestions	Changes
Cycle 1 (n=5)	
Scenario 1—sign-in, chatbot, and Event Profile	
There was a low contrast between the text and the background and small font at sign-in.	Larger font and chatbot and user text boxes identified using different colors were implemented.
Clarification that the chatbot is not a real person was requested.	A page was added to the *at heart* progressive web application to describe the chatbot, including its development and function.
Participants selected only 1 descriptor on the body map and reported lack of clarity that more than one descriptor could be selected.	Directions for completing the body map were enhanced.
Scenario 2—Heart Check	
None	None
Scenario 3—Wellness Check	
None	None
Scenario 4—library	
Participants had difficulty locating the Favorites section of the library.	The Favorites section was moved to the top of the screen and identified with a larger icon that represented a heart.
Scenario 5—library retrieval and notes	
None	None
Other	
None	None
Cycle 2 (n=5)	
Scenario 1—sign-in, chatbot, and Event Profile	
Participants requested clarification on the terminology used in some of the questions at sign-in.	Descriptors and more answers were added to the questions for clarity.
Participants had difficulty choosing specific areas on the body map to describe their symptoms.	The body maps were subsequently amended to include axillae and body map descriptors.
Scenario 2—Heart Check	
Participants appeared confused with the order and speed of the questions delivered by the chatbot.	The chat speed was reduced for desktop, Android, and iOS devices.
The Heart Check button was reported to be too small and gray.	The size and contrast of the “Heart Check” button were adjusted.
Scenario 3—Wellness Check	
The chatbot says the following: “based on what you’ve told me, here are some topics...” However, the chatbot presented all topics even for wellness scores that were excellent (ie, the readings from the library did not appear tailored to the needs of each participant).	Minor adjustments were made to the rules to ensure that the chatbot selected appropriate resources from the wellness library based on wellness scores of ≤4.
Scenario 4—library	
None	None
Scenario 5—library retrieval and notes	
None	None
Other	
Add anaphylaxis to the Share My Data section.	This was added.
Add anaphylaxis articles to the library as they related to women and heart disease.	Lay summaries were generated and scientific papers were added to the library for Kounis syndrome (n=3) and mast cell activation syndrome (n=2).
Participants requested to have a more robust Share My Data section, with opportunity to record medications and other comorbid conditions.	The introduction and the content of the Share My Data section were enhanced.

XSL•FO
RenderX

User Satisfaction

Overview

User satisfaction was assessed using the SUS and the 4 semistructured interview questions. Overall, 10 participants across 2 cycles of testing reported the overall usability of the *at heart* progressive web application as highly acceptable (mean SUS score 81.75, SD 10.41). There was no difference in SUS scores between cycle 1 (mean SUS score 81.5, SD 8.59) and cycle 2 (mean SUS score 82.0, SD 13.04; t_8=−0.72, 95% CI −16.6 to 15.6). A total of 90% (9/10) of the participants rated the user-friendliness of *at heart* as good or excellent. All participants (10/10, 100%) thought that *at heart* was easy to use (SUS question 3; ie, most women could learn to use *at heart* very quickly):

> *I think everything was pretty straightforward. It was overall really, really good.* [Cycle 2; participant 3]

None of the participants indicated that they would need the support of a technical person to be able to use *at heart* (SUS question 4), and 80% (8/10) of the participants found the various functions in *at heart* well integrated (SUS question 5) with minimal inconsistencies (SUS question 6):

> *Well...I was a bit worried it might be confusing, but it's not confusing at all. It's very intuitive, it, you've built it so it's similar to a lot of other apps out there, right? So, clicking here, you know, submitting the button. So, it seems to be pretty normal as far, it looks, the same look and feel as other apps that are out there which are good. It seems to be easy. I'm not a tech person, so I wanted to make sure that—I was nervous that I wasn't going to be able to manage it because I, I'm not great with technology, but it was easy. So, yeah. I mean, I think it's a great tool.* [Cycle 2; participant 5]

A total of 8 themes emerged from the interview data (ie, engaging, comprehensive, understandable, credible, relevant, affirming, personalized, and innovative) and can be found in Multimedia Appendix 3.

Engaging

Participants commented on the layout, visual appeal, language, and name and logo of the progressive web application. They found the content easy to read and maneuver. Important content was highlighted, and participants reported this to be helpful. Article summaries were presented in larger font and used lay language. At the outset of this study, the progressive web application was named HEARTPA♀N [29]. Participants recommended a name change to one that was less focused on pain as many women describe their cardiac pain as discomfort and have other associated symptoms (eg, dyspnea and fatigue). They felt that the name *at heart* was more personally welcoming and less focused on pain. The heart logo was simple yet distinctive and recognizable.

Comprehensive

Participants indicated that the progressive web application provided necessary information for helping them make decisions about their cardiac symptoms. They liked that it was specific to women and that they could search through the library to improve their knowledge about wellness and women's heart health. Participants also commented that the chatbot assisted them in making decisions:

> *...the bot kind of identifies for me what the big blind spots are that I am perhaps not noticing.*

Understandable

Participants indicated that *at heart* was intuitive, the content made sense, and it flowed well. The tone and the language were suitable and clear. Participants commented on the appropriateness of the associated symptom descriptors included in the Event Profile and Heart Check.

Credible

Participants liked that *at heart* was developed as a living progressive web application. They particularly liked the breadth and depth of scientific articles and lay summaries included in the *at heart* library. Participants valued having an evidence-informed digital health self-management program specifically designed for women with CAD living with cardiac pain and associated symptoms.

Relevant

Participants commented on the relevance and accessibility of content in the *at heart* library. The chatbot asked relevant questions and was able to use rule-based algorithms to provide advice to women about timely assessment in the emergency department. They particularly noted the relevance of having their Event Profile and Heart Check information stored so that they could refer to it when visiting their health care providers. The content was applicable to their everyday lives.

Affirming

Participants looked for affirmation from *at heart's* chatbot when they were having cardiac symptoms. They indicated that they would have otherwise consulted with family members or Facebook group members for guidance. Affirmation of symptoms was important to women, and they appreciated the high-alert messages to seek urgent care immediately.

Personalized

Participants liked the chatbot. They found it to be personal, immediately interactive, accessible, calming, and friendly. They commented that the chatbot may help make them feel less alone if they went to the emergency room because it checked in with them to see how they were doing. The chatbot provided comfort at times when participants felt scared. Participants also valued the videos and podcasts as they also helped them feel that they were not alone in their experience with heart disease.

Innovative

Participants felt that *at heart* provided an opportunity for further research focused on women's heart health. They were willing to share anonymized data on socioeconomics, risk factors, gynecological or obstetrics history, lifestyle, medications, and other conditions (eg, arthritis or depression) to improve outcomes for women at risk of or with heart disease. Participants also valued the ability to journal and take notes in the

XSL•FO
RenderX

progressive web application. They valued the opportunities that *at heart* could provide to women who live in more isolated areas or those who live alone. The *at heart* progressive web application was described as all-encompassing, covering necessary content to improve knowledge and decision-making, with a chatbot to ask questions and deliver information in a comforting manner.

Discussion

Principal Findings

There is growing support for the importance of usability testing for mobile health and eHealth innovations [60-62]. Usability testing incorporates an iterative process of testing and refining to meet end-user needs [63]. The objectives of this study were to assess the user performance (ie, ease of use, efficiency, and errors) and user satisfaction (SUS) of a progressive web application for women with cardiac pain. User performance was assessed through ease of use, efficiency, and observation of testing errors (ie, navigation, presentation, and control use errors) through 5 brief scenarios and 2 cycles of usability testing.

The overall usability of the *at heart* progressive web application was rated highly by 10 participants who completed the usability testing cycles; the progressive web application was easy to use and efficient. However, 2 high-priority testing errors were identified during the usability testing cycles. The first involved small font at sign-in and low contrast between the text and background, and these were corrected by enlarging the font and making alterations to the text and background colors. The second was to clarify that the chatbot was not a real person. A significant amount of time was committed to addressing potential language challenges during the development of the chatbot, and there were no technical, design, or language challenges identified during the usability testing cycles. In fact, the design and language or conversation of the chatbot appeared to mimic human-to-human interaction so closely that participants requested confirmation that the chatbot was actually nonhuman.

The *at heart* chatbot is a simple rule-based conversational agent designed to mimic human-to-human interaction at sign-in and creation of an Event Profile and during completion of the Heart Check and Wellness Check. Our previous systematic review and meta-analysis indicated that self-management programs were more effective in reducing cardiac pain and associated symptoms (eg, dyspnea and fatigue) in women if they included collaboration and support from health care providers [33]. Chatbots are conversational agents able to promote health and provide education and support [64]. Their use in health care is still in a developmental stage; they could improve access to care and health care provider–patient communication, but more evidence is needed. Technical, design (ie, lack of empathy), and language challenges [65] can impede the integration of chatbot technologies into health care [64]. Adopting user-centered and theory-based designs, optimizing user experiences, and addressing patient concerns can improve chatbot uptake and use [64]. Our chatbot relies on scripted computational algorithms, with specific rules for its text-based conversations [66]. More advanced human-machine conversation is now based on natural language processing and large language models that use artificial intelligence methods to learn, understand, and produce structured language content [66,67]. Health care delivery is rapidly changing and is driven by social, scientific, and technological change; our future chatbot may need to emulate person-to-person conversation through dialogue and body movements, with appropriate expressions of empathy and compassion [68]. We are really at a cusp in health care; growth in chatbot use will be driven by a desire for health and wellness and 24-hour access to care, with a growing number of platforms from which to build an intelligent and emotive chatbot in the future [69].

At heart is the first of its kind; no smartphone- or web-based self-management program has been co-designed and systematically developed with women who have lived experience (phase 2A) and then tested with women who have cardiac pain (phase 2B). We used the individual and family self-management theory [70,71], mobile device functionality, and the pervasive information architecture of mobile health interventions [72] and followed the MRC's guidance for developing complex interventions [30-32]. *At heart* was designed as a progressive web application aimed at delivering nativelike user experiences regardless of the browser or the mobile device [73]. This study identified that the comprehensiveness and credibility of the information was important to women. The information helped them understand more about CAD and wellness and provided them with guidance in decision-making. At inception, women clearly articulated the need for a web application that was accessible across Android and iOS operating systems, including computers, smartphones, and tablets. In phase 1, women identified the need to access the library anytime and anywhere (eg, while waiting for dental appointments). As women were involved at the outset in co-design, no other significant user performance or satisfaction–related issues were identified. Female front and back full-body maps were specifically developed for *at heart* using the chest pain or associated symptom locations most commonly described in the literature [6,8,13,47,53]. These required only minor refinements during usability testing (ie, more precise identification of the axillae). Importantly, participants also viewed *at heart* as a way to contribute to the future of women's heart health. Women feel "stopped at the gate," and they want to take charge and advocate for better awareness, education, diagnosis, and management [74].

Limitations and Strengths

First, usability testing was conducted on a predominantly affluent sample of participants, which limits the generalizability of our results. A total of 60% (6/10) of the participants who took part in usability testing earned >CAD $70,000 (US $48,881.80) annually. They were also well educated, and most (8/10, 80%) were White. Canada is known for its diversity in race and ethnicities— over 450 ethnic or cultural origins were reported in the 2021 census [75]. A total of 1 out of 5 people in Canada is born elsewhere; the 3 largest visible ethnic groups are South Asian, Chinese, and Black individuals, representing 60% of the Canadian racial and ethnic populations [76]. Although this was a limitation in the sample for the usability testing, this is not a limitation specific to the *at heart* progressive web application. In fact, the *at heart* library contains heart

XSL•FO
RenderX

disease and heart wellness scientific papers and lay summaries related to South Asian and Indigenous people. These include risk and traditional practices such as drumming and healing and talking circles. However, it will be necessary to include a more diverse sample of women that encompasses those in less populous areas where age-standardized cardiovascular-related deaths in women are highest [77]. It will also be important to target recruitment of women from South Asian and Afro-Caribbean races and ethnicities, those with disabilities, those of a lower socioeconomic status, and those across age and gender identities who also experience greater risk of cardiovascular disease [77-79].

In addition, heterogeneous use of platforms and devices may also be a limitation to the usability testing. Progressive web applications are designed to work across platforms, eliminating the cost, fragmentation, and need to develop the same application several times for multiple platforms. They are designed to behave like mobile apps, working seamlessly across Android and iOS devices. However, past research indicates that progressive web applications may not be as smooth as mobile apps on Android devices [80]. Our sample size was limited, and only 10% (1/10) of the participants used an Android device for testing. Further research is needed to evaluate the usability of the *at heart* progressive web application across various platforms and devices. The phase-3 pilot RCT will determine the feasibility of implementing the *at heart* intervention, including the *feasibility* of randomization, recruitment, and retention; *acceptability* and barriers to implementing the intervention; and the extent of engagement with the intervention. The phase-4 effectiveness-implementation hybrid trial (type I) will investigate the effectiveness and implementation of *at heart* among women living in Canada using the practical, robust implementation and sustainability model [81]. The MRC [30-32] discusses 4 phases of complex intervention research (eg, development, feasibility, evaluation, and implementation) and the importance of refinements with each phase transition depending on the context [70,71] in which the intervention is evaluated and implemented. The practical, robust implementation and sustainability model [81] aligns contextual needs with intervention design using an inequity, systems thinking, and cocreation engagement approach. Context will be addressed by considering the individual perspectives on the intervention, external environment, characteristics of those receiving the intervention, and supports and resources needed to deliver the intervention [81]. Recruiting representative populations can be improved through strengthened community partnerships in governance and decision-making [82-84], and this will be a focus in subsequent phases of this research.

Second, remote usability testing methods were used during the COVID-19 pandemic. Testing a technology's usability remotely has inherent advantages as well as limitations. While participants can prefer remote testing to in-person testing, it can take longer for participants to complete tasks during remote testing, more errors can be made compared to in-person testing sessions, and context and indirect cues can be missed during remote testing sessions [85]. However, the approach taken to usability testing in this study was iterative, comprehensive, and aligned with methods used by others recently reported in a scoping review [41]. Of 133 articles included in this scoping review, 105 used questionnaires, 45 used a think-aloud approach, and 37 used interviews in their usability testing methods [41]. Moreover, many studies used a combination of 2 (n=46) or 3 (n=30) methods of testing, and the SUS was the most reported and validated questionnaire used to assess satisfaction and usability [41]. Finally, the *at heart* chatbot is rule based and designed for specific functions within the progressive web application. It is not led by artificial intelligence technologies (ie, natural language processing or large language models) and, therefore, may have inherent limitations that were not identified during the 2 cycles of usability testing (ie, personality, flexibility, dialogue structure, and conversation complexity or flow) [65].

Conclusions

This study does provide initial support for the *at heart* progressive web application for women with cardiac pain. Participants rated the performance of and satisfaction with this progressive web application as high. Ongoing evaluations in phases 3 and 4 should aim to examine the feasibility and acceptability and the extent of engagement with the *at heart* core feature set: Heart Check, Wellness Check, and the library. In addition to assessing effectiveness in an effectiveness-implementation hybrid trial (type I), describing and better understanding the context for implementation (eg, race and ethnicity and geography) will be necessary.

Acknowledgments

The authors thank the women who were participants in this study. They also thank the women with lived experience who collaborated in the co-design of *at heart* and the Heart and Stroke Foundation of Canada for their partnership. Phases 2 and 3 of the *at heart* program of research are supported by the Canadian Institutes of Health Research (389044) Project Grant: Fall 2017 competition.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Participant demographic and clinical information form.
[PDF File (Adobe PDF File), 94 KB - humanfactors_v12i1e57583_app1.pdf]

XSL•FO
RenderX

Multimedia Appendix 2
Usability testing error and efficiency documentation form.
[PDF File (Adobe PDF File), 72 KB - humanfactors_v12i1e57583_app2.pdf]

Multimedia Appendix 3
Summary of participant comments by theme.
[PDF File (Adobe PDF File), 103 KB - humanfactors_v12i1e57583_app3.pdf]

References

1. GBD 2017 Causes of Death Collaborators. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Nov 10;392(10159):1736-1788 [FREE Full text] [doi: 10.1016/S0140-6736(18)32203-7] [Medline: 30496103]
2. Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, et al. Global burden of cardiovascular diseases and risk factors, 1990-2019: update from the GBD 2019 study. J Am Coll Cardiol 2020 Dec 22;76(25):2982-3021 [FREE Full text] [doi: 10.1016/j.jacc.2020.11.010] [Medline: 33309175]
3. Vaduganathan M, Mensah GA, Turco JV, Fuster V, Roth GA. The global burden of cardiovascular diseases and risk: a compass for future health. J Am Coll Cardiol 2022 Dec 20;80(25):2361-2371 [FREE Full text] [doi: 10.1016/j.jacc.2022.11.005] [Medline: 36368511]
4. Mensah GA, Roth GA, Fuster V. The global burden of cardiovascular diseases and risk factors: 2020 and beyond. J Am Coll Cardiol 2019 Nov 19;74(20):2529-2532 [FREE Full text] [doi: 10.1016/j.jacc.2019.10.009] [Medline: 31727292]
5. Pepine CJ, Ferdinand KC, Shaw LJ, Light-McGroary KA, Shah RU, Gulati M, et al. Emergence of nonobstructive coronary artery disease: a woman's problem and need for change in definition on angiography. J Am Coll Cardiol 2015 Oct 27;66(17):1918-1933 [FREE Full text] [doi: 10.1016/j.jacc.2015.08.876] [Medline: 26493665]
6. McSweeney J, Cleves MA, Fischer EP, Moser DK, Wei J, Pettey C, et al. Predicting coronary heart disease events in women: a longitudinal cohort study. J Cardiovasc Nurs 2014;29(6):482-492 [FREE Full text] [doi: 10.1097/JCN.0b013e3182a409cc] [Medline: 24231895]
7. Parry M, Bjørnnes AK, Clarke H, Cooper L, Gordon A, Harvey P, et al. Self-management of cardiac pain in women: an evidence map. BMJ Open 2017 Nov 25;7(11):e018549 [FREE Full text] [doi: 10.1136/bmjopen-2017-018549] [Medline: 29175891]
8. Canto JG, Canto EA, Goldberg RJ. Time to standardize and broaden the criteria of acute coronary syndrome symptom presentations in women. Can J Cardiol 2014 Jul;30(7):721-728. [doi: 10.1016/j.cjca.2013.10.015] [Medline: 24530216]
9. Kirchberger I, Heier M, Wende R, von Scheidt W, Meisinger C. The patient's interpretation of myocardial infarction symptoms and its role in the decision process to seek treatment: the MONICA/KORA Myocardial Infarction Registry. Clin Res Cardiol 2012 Nov;101(11):909-916. [doi: 10.1007/s00392-012-0475-8] [Medline: 22618810]
10. Kok MM, van der Heijden LC, Sen H, Danse PW, Löwik MM, Anthonio RL, et al. Sex difference in chest pain after implantation of newer generation coronary drug-eluting stents: a patient-level pooled analysis from the TWENTE and DUTCH PEERS trials. JACC Cardiovasc Interv 2016 Mar 28;9(6):553-561 [FREE Full text] [doi: 10.1016/j.jcin.2015.10.043] [Medline: 26947384]
11. Choinière M, Watt-Watson J, Victor JC, Baskett RJ, Bussières JS, Carrier M, et al. Prevalence of and risk factors for persistent postoperative nonanginal pain after cardiac surgery: a 2-year prospective multicentre study. CMAJ 2014 Apr 15;186(7):E213-E223 [FREE Full text] [doi: 10.1503/cmaj.131012] [Medline: 24566643]
12. Katz J, Weinrib A, Fashler SR, Katznelzon R, Shah BR, Ladak SS, et al. The Toronto General Hospital Transitional Pain Service: development and implementation of a multidisciplinary program to prevent chronic postsurgical pain. J Pain Res 2015;8:695-702 [FREE Full text] [doi: 10.2147/JPR.S91924] [Medline: 26508886]
13. Sella YO, Manistamara H, Apriliawan S, Lukitasari M, Rohman MS. Characteristic differences of chest pain in male and female patients with acute coronary syndrome: a pilot study. J Public Health Res 2021 Apr 14;10(2):2242 [FREE Full text] [doi: 10.4081/jphr.2021.2242] [Medline: 33855424]
14. Altintas E, Yigit F, Taskintuna N. The impact of psychiatric disorders with cardiac syndrome X on quality of life: 3 months prospective study. Int J Clin Exp Med 2014;7(10):3520-3527 [FREE Full text] [Medline: 25419392]
15. Conradie A, Atherton J, Chowdhury E, Duong M, Schwarz N, Worthley S, et al. Health-Related Quality of Life (HRQoL) and the effect on outcome in patients presenting with coronary artery disease and treated with Percutaneous Coronary Intervention (PCI): differences noted by sex and age. J Clin Med 2022 Sep 05;11(17):5231 [FREE Full text] [doi: 10.3390/jcm11175231] [Medline: 36079161]
16. Sabbag A, Matetzky S, Gottlieb S, Fefer P, Kohanov O, Atar S, et al. Recent temporal trends in the presentation, management, and outcome of women hospitalized with acute coronary syndromes. Am J Med 2015 Apr;128(4):380-388. [doi: 10.1016/j.amjmed.2014.10.041] [Medline: 25460528]

XSL•FO
RenderX

17. Sjöström-Strand A, Fridlund B. Women's descriptions of symptoms and delay reasons in seeking medical care at the time of a first myocardial infarction: a qualitative study. Int J Nurs Stud 2008 Jul;45(7):1003-1010. [doi: 10.1016/j.ijnurstu.2007.07.004] [Medline: 17803997]
18. Hadid LA, Al Barmawi M, Al Hmaimat NA, Shoqirat N. Factors associated with prehospital delay among men and women newly experiencing acute coronary syndrome: a qualitative inquiry. Cardiol Res Pract 2020;2020:3916361 [FREE Full text] [doi: 10.1155/2020/3916361] [Medline: 32454998]
19. Lorig KR, Holman H. Self-management education: history, definition, outcomes, and mechanisms. Ann Behav Med 2003 Aug;26(1):1-7. [doi: 10.1207/S15324796ABM2601_01] [Medline: 12867348]
20. Osborne RH, Batterham R, Livingston J. The evaluation of chronic disease self-management support across settings: the international experience of the health education impact questionnaire quality monitoring system. Nurs Clin North Am 2011 Sep;46(3):255-70, v. [doi: 10.1016/j.cnur.2011.05.010] [Medline: 21791261]
21. Barlow J, Wright C, Sheasby J, Turner A, Hainsworth J. Self-management approaches for people with chronic conditions: a review. Patient Educ Couns 2002;48(2):177-187. [doi: 10.1016/s0738-3991(02)00032-0] [Medline: 12401421]
22. Knight-Agarwal C, Davis DL, Williams L, Davey R, Cox R, Clarke A. Development and pilot testing of the Eating4two mobile phone app to monitor gestational weight gain. JMIR Mhealth Uhealth 2015 Jun 05;3(2):e44 [FREE Full text] [doi: 10.2196/mhealth.4071] [Medline: 26048313]
23. Willcox JC, Campbell KJ, McCarthy EA, Wilkinson SA, Lappas M, Ball K, et al. Testing the feasibility of a mobile technology intervention promoting healthy gestational weight gain in pregnant women (txt4two) - study protocol for a randomised controlled trial. Trials 2015 May 07;16:209 [FREE Full text] [doi: 10.1186/s13063-015-0730-1] [Medline: 25947578]
24. Waring ME, Moore Simas TA, Xiao RS, Lombardini LM, Allison JJ, Rosal MC, et al. Pregnant women's interest in a website or mobile application for healthy gestational weight gain. Sex Reprod Healthc 2014 Dec;5(4):182-184 [FREE Full text] [doi: 10.1016/j.srhc.2014.05.002] [Medline: 25433828]
25. Fukuoka Y, Komatsu J, Suarez L, Vittinghoff E, Haskell W, Noorishad T, et al. The mPED randomized controlled clinical trial: applying mobile persuasive technologies to increase physical activity in sedentary women protocol. BMC Public Health 2011 Dec 14;11:933. [doi: 10.1186/1471-2458-11-933] [Medline: 22168267]
26. Llanos AA, Krok JL, Peng J, Pennell ML, Vitolins MZ, Degraffinreid CR, et al. Effects of a walking intervention using mobile technology and interactive voice response on serum adipokines among postmenopausal women at increased breast cancer risk. Horm Cancer 2014 Apr;5(2):98-103 [FREE Full text] [doi: 10.1007/s12672-013-0168-4] [Medline: 24435584]
27. Wen KY, Miller SM, Kilby L, Fleisher L, Belton TD, Roy G, et al. Preventing postpartum smoking relapse among inner city women: development of a theory-based and evidence-guided text messaging intervention. JMIR Res Protoc 2014 Apr 03;3(2):e20 [FREE Full text] [doi: 10.2196/resprot.3059] [Medline: 24698804]
28. Licskai C, Sands TW, Ferrone M. Development and pilot testing of a mobile health solution for asthma self-management: asthma action plan smartphone application pilot study. Can Respir J 2013;20(4):301-306 [FREE Full text] [doi: 10.1155/2013/906710] [Medline: 23936890]
29. Parry M, Dhukai A, Clarke H, Bjørnnes AK, Cafazzo JA, Cooper L, et al. Development and usability testing of HEARTPA ♀ N: protocol for a mixed methods strategy to develop an integrated smartphone and web-based intervention for women with cardiac pain. BMJ Open 2020 Mar 09;10(3):e033092 [FREE Full text] [doi: 10.1136/bmjopen-2019-033092] [Medline: 32156763]
30. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008 Sep 29;337:a1655 [FREE Full text] [doi: 10.1136/bmj.a1655] [Medline: 18824488]
31. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. Framework for the development and evaluation of complex interventions: gap analysis, workshop and consultation-informed update. Health Technol Assess 2021 Sep;25(57):1-132 [FREE Full text] [doi: 10.3310/hta25570] [Medline: 34590577]
32. Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, et al. Framework for design and evaluation of complex interventions to improve health. BMJ 2000 Sep 16;321(7262):694-696 [FREE Full text] [doi: 10.1136/bmj.321.7262.694] [Medline: 10987780]
33. Parry M, Bjørnnes AK, Victor JC, Ayala AP, Lenton E, Clarke H, et al. Self-management interventions for women with cardiac pain: a systematic review and meta-analysis. Can J Cardiol 2018 Apr;34(4):458-467. [doi: 10.1016/j.cjca.2017.12.011] [Medline: 29477931]
34. Wichansky AM. Usability testing in 2000 and beyond. Ergonomics 2000 Jul;43(7):998-1006. [doi: 10.1080/001401300409170] [Medline: 10929833]
35. Currie LM. Evaluation frameworks for nursing informatics. Int J Med Inform 2005 Dec;74(11-12):908-916. [doi: 10.1016/j.ijmedinf.2005.07.007] [Medline: 16099711]
36. Ritterband LM, Gonder-Frederick LA, Cox DJ, Clifton AD, West RW, Borowitz SM. Internet interventions: in review, in use, and into the future. Prof Psychol Res Pract 2003 Oct;34(5):527-534. [doi: 10.1037/0735-7028.34.5.527]
37. Gustafson DH, Wyatt JC. Evaluation of ehealth systems and services. BMJ 2004 May 15;328(7449):1150. [doi: 10.1136/bmj.328.7449.1150] [Medline: 15142895]

XSL•FO
RenderX

38. Callahan CM, Unverzagt FW, Hui SL, Perkins AJ, Hendrie HC. Six-item screener to identify cognitive impairment among potential subjects for clinical research. Med Care 2002 Sep;40(9):771-781. [doi: 10.1097/00005650-200209000-00007] [Medline: 12218768]
39. Carpenter CR, DesPain B, Keeling TN, Shah M, Rothenberger M. The Six-Item Screener and AD8 for the detection of cognitive impairment in geriatric emergency department patients. Ann Emerg Med 2011 Jun;57(6):653-661 [FREE Full text] [doi: 10.1016/j.annemergmed.2010.06.560] [Medline: 20855129]
40. Hyzy M, Bond R, Mulvenna M, Bai L, Dix A, Leigh S, et al. System usability scale benchmarking for digital health apps: meta-analysis. JMIR Mhealth Uhealth 2022 Aug 18;10(8):e37290 [FREE Full text] [doi: 10.2196/37290] [Medline: 35980732]
41. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: a scoping review. Int J Med Inform 2019 Jun;126:95-104. [doi: 10.1016/j.ijmedinf.2019.03.018] [Medline: 31029270]
42. Fleming JB, Hill YN, Burns MN. Usability of a culturally informed mHealth intervention for symptoms of anxiety and depression: feedback from young sexual minority men. JMIR Hum Factors 2017 Aug 25;4(3):e22. [doi: 10.2196/humanfactors.7392] [Medline: 28842389]
43. Srikesavan C, Williamson E, Cranston T, Hunter J, Adams J, Lamb SE. An online hand exercise intervention for adults with rheumatoid arthritis (mySARAH): design, development, and usability testing. J Med Internet Res 2018 Jun 27;20(6):e10457 [FREE Full text] [doi: 10.2196/10457] [Medline: 29950288]
44. Voncken-Brewster V, Moser A, van der Weijden T, Nagykaldi Z, de Vries H, Tange H. Usability evaluation of an online, tailored self-management intervention for chronic obstructive pulmonary disease patients incorporating behavior change techniques. JMIR Res Protoc 2013 Jan 16;2(1):e3 [FREE Full text] [doi: 10.2196/resprot.2246] [Medline: 23612363]
45. Macefield R. How to specify the participant group size for usability studies: a practitioner's guide. J Usab Stud 2009:34-45 [FREE Full text]
46. Nielsen J, Landauer TK. A mathematical model of the finding of usability problems. In: Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems. 1993 May 01 Presented at: CHI '93; April 24-29, 1993; Amsterdam, The Netherlands URL: https://dl.acm.org/doi/abs/10.1145/169059.169166 [doi: 10.1145/169059.169166]
47. Gulati M, Levy PD, Mukherjee D, Amsterdam E, Bhatt DL, Birtcher KK, et al. 2021 AHA/ACC/ASE/CHEST/SAEM/SCCT/SCMR guideline for the evaluation and diagnosis of chest pain: a report of the American College of Cardiology/American Heart Association Joint Committee on clinical practice guidelines. Circulation 2021 Nov 30;144(22):e368-e454 [FREE Full text] [doi: 10.1161/CIR.0000000000001029] [Medline: 34709879]
48. Brooke J. SUS: a 'quick and dirty' usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation in Industry. Boca Raton, FL: CRC Press; 1996.
49. Peres SC, Pham T, Phillips R. Validation of the system usability scale (SUS): SUS in the wild. Proc Hum Factors Ergon Soc Annu Meet 2013 Sep 30;57(1):192-196. [doi: 10.1177/1541931213571043]
50. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009;4(3):114-123.
51. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Hum Comput Interact 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]
52. Gräve E, Buchner A. Is less sometimes more? An experimental comparison of four measures of perceived usability. Hum Factors 2025 Jan;67(1):32-48 [FREE Full text] [doi: 10.1177/00187208241237862] [Medline: 38482806]
53. Ferry AV, Anand A, Strachan FE, Mooney L, Stewart SD, Marshall L, et al. Presenting symptoms in men and women diagnosed with myocardial infarction using sex‐specific criteria. JAHA 2019 Sep 03;8(17). [doi: 10.1161/jaha.119.012307]
54. Gjeilo KH, Stenseth R, Wahba A, Lydersen S, Klepstad P. Validation of the brief pain inventory in patients six months after cardiac surgery. J Pain Symptom Manage 2007 Dec;34(6):648-656 [FREE Full text] [doi: 10.1016/j.jpainsymman.2007.01.010] [Medline: 17629665]
55. Health Canada. Guidance document: Software as a Medical Device (SaMD): classification examples. Government of Canada. Ottawa, ON: Health Canada; 2019 Dec 18. URL: https://www.canada.ca/content/dam/hc-sc/documents/services/drugs-health-products/medical-devices/application-information/guidance-documents/software-medical-device-guidance/examples/examples.pdf [accessed 2025-03-26]
56. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000 Aug;23(4):334-340. [doi: 10.1002/1098-240x(200008)23:4<334::aid-nur9>3.0.co;2-g] [Medline: 10940958]
57. Rogers ML, Patterson E, Chapman R, Render M. Usability testing and the relation of clinical information systems to patient safety. In: Henriksen K, Battles JB, Marks ES, Lewin DI, editors. Advances in Patient Safety: From Research to Implementation (Volume 2: Concepts and Methodology). Rockville, MD: Agency for Healthcare Research and Quality (US); 2005.
58. Kounis NG. Coronary hypersensitivity disorder: the Kounis syndrome. Clin Ther 2013 May;35(5):563-571. [doi: 10.1016/j.clinthera.2013.02.022] [Medline: 23490289]
59. Kounis NG, Koniari I, Velissaris D, Tzanis G, Hahalis G. Kounis syndrome—not a single-organ arterial disorder but a multisystem and multidisciplinary disease. Balkan Med J 2019 Jul 11;36(4):212-221 [FREE Full text] [doi: 10.4274/balkanmedj.galenos.2019.2019.5.62] [Medline: 31198019]

XSL•FO
RenderX

60. Hales S, Turner-McGrievy G, Fahim A, Freix A, Wilcox S, Davis RE, et al. A mixed-methods approach to the development, refinement, and pilot testing of social networks for improving healthy behaviors. JMIR Hum Factors 2016 Feb 12;3(1):e8 [FREE Full text] [doi: 10.2196/humanfactors.4512] [Medline: 27026535]
61. Kaufman DR, Patel VL, Hilliman C, Morin PC, Pevzner J, Weinstock RS, et al. Usability in the real world: assessing medical information technologies in patients' homes. J Biomed Inform 2003 Feb;36(1-2):45-60 [FREE Full text] [doi: 10.1016/s1532-0464(03)00056-x] [Medline: 14552846]
62. Kushniruk A. Evaluation in the design of health information systems: application of approaches emerging from usability engineering. Comput Biol Med 2002 May;32(3):141-149. [doi: 10.1016/s0010-4825(02)00011-2] [Medline: 11922931]
63. Snodgrass A, Coyne R. Models, metaphors and the hermeneutics of designing. Design Issues 1992;9(1):56-74. [doi: 10.2307/1511599]
64. Nadarzynski T, Miles O, Cowie A, Ridge D. Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: a mixed-methods study. Digit Health 2019 Aug 21;5:2055207619871808 [FREE Full text] [doi: 10.1177/2055207619871808] [Medline: 31467682]
65. Fadhil A, Schiavo G. Designing for health chatbots. arXiv Preprint posted online on February 24, 2019 [FREE Full text]
66. Hirschberg J, Manning CD. Advances in natural language processing. Science 2015 Jul 17;349(6245):261-266. [doi: 10.1126/science.aaa8685] [Medline: 26185244]
67. Lopez-Jimenez F, Attia Z, Arruda-Olson AM, Carter R, Chareonthaitawee P, Jouni H, et al. Artificial intelligence in cardiology: present and future. Mayo Clin Proc 2020 May;95(5):1015-1039. [doi: 10.1016/j.mayocp.2020.01.038] [Medline: 32370835]
68. Martínez-Miranda J, Martínez A, Ramos R, Aguilar H, Jiménez L, Arias H, et al. Assessment of users' acceptability of a mobile-based embodied conversational agent for the prevention and detection of suicidal behaviour. J Med Syst 2019 Jun 25;43(8):246. [doi: 10.1007/s10916-019-1387-1] [Medline: 31240494]
69. Chakraborty C, Pal S, Bhattacharya M, Dash S, Lee SS. Overview of Chatbots with special emphasis on artificial intelligence-enabled ChatGPT in medical science. Front Artif Intell 2023;6:1237704 [FREE Full text] [doi: 10.3389/frai.2023.1237704] [Medline: 38028668]
70. Fawcett J, Watson J, Neuman B, Walker PH, Fitzpatrick JJ. On nursing theories and evidence. J Nurs Scholarsh 2001;33(2):115-119. [doi: 10.1111/j.1547-5069.2001.00115.x] [Medline: 11419305]
71. Ryan P, Sawin KJ. The individual and family self-management theory: background and perspectives on context, process, and outcomes. Nurs Outlook 2009 Jul;57(4):217-25.e6 [FREE Full text] [doi: 10.1016/j.outlook.2008.10.004] [Medline: 19631064]
72. Danaher BG, Brendryen H, Seeley JR, Tyler MS, Woolley T. From black box to toolbox: outlining device functionality, engagement activities, and the pervasive information architecture of mHealth interventions. Internet Interv 2015 Mar 01;2(1):91-101 [FREE Full text] [doi: 10.1016/j.invent.2015.01.002] [Medline: 25750862]
73. Tandel SS, Jamadar A. Impact of progressive web apps on web app development. Int J Innov Res Sci Eng Technol 2018;7(9):9439-9444. [doi: 10.15680/IJIRSET.2018.0709021]
74. Colella TJ, Hardy M, Hart D, Price JA, Sarfi H, Mullen KA, et al. The Canadian Women's Heart Health Alliance Atlas on the epidemiology, diagnosis, and management of cardiovascular disease in women-chapter 3: patient perspectives. CJC Open 2021 Mar;3(3):229-235 [FREE Full text] [doi: 10.1016/j.cjco.2020.11.020] [Medline: 33778439]
75. The Canadian census: a rich portrait of the country's religious and ethnocultural diversity. Statistics Canada. Ottawa, ON: Statistics Canada; 2022 Oct 26. URL: https://www150.statcan.gc.ca/n1/daily-quotidien/221026/dq221026b-eng.htm [accessed 2024-01-23]
76. Immigration and ethnocultural diversity in Canada. Statistics Canada. Ottawa, ON: Statistics Canada; 2011. URL: https://www12.statcan.gc.ca/nhs-enm/2011/as-sa/99-010-x/99-010-x2011001-eng.cfm [accessed 2025-03-26]
77. Jaffer S, Foulds HJ, Parry M, Gonsalves CA, Pacheco C, Clavel MA, et al. The Canadian Women's Heart Health Alliance Atlas on the epidemiology, diagnosis, and management of cardiovascular disease in women-chapter 2: scope of the problem. CJC Open 2021 Jan;3(1):1-11 [FREE Full text] [doi: 10.1016/j.cjco.2020.10.009] [Medline: 33458627]
78. Norris CM, Yip CY, Nerenberg KA, Clavel M, Pacheco C, Foulds HJ, et al. State of the science in women's cardiovascular disease: a Canadian perspective on the influence of sex and gender. J Am Heart Assoc 2020 Feb 18;9(4):e015634. [doi: 10.1161/JAHA.119.015634] [Medline: 32063119]
79. Parry M, Van Spall HG, Mullen KA, Mulvagh SL, Pacheco C, Colella TJ, et al. The Canadian Women's Heart Health Alliance Atlas on the epidemiology, diagnosis, and management of cardiovascular disease in women - chapter 6: sex- and gender-specific diagnosis and treatment. CJC Open 2022 Jul;4(7):589-608 [FREE Full text] [doi: 10.1016/j.cjco.2022.04.002] [Medline: 35865023]
80. Fournier C. Comparison of smoothness in progressive web apps and mobile applications on android. KTH Royal Institute of Technology. 2020. URL: https://www.diva-portal.org/smash/get/diva2:1474729/FULLTEXT01.pdf [accessed 2025-03-25]
81. Jolles MP, Fort MP, Glasgow RE. Aligning the planning, development, and implementation of complex interventions to local contexts with an equity focus: application of the PRISM/RE-AIM Framework. Int J Equity Health 2024 Feb 27;23(1):41 [FREE Full text] [doi: 10.1186/s12939-024-02130-6] [Medline: 38408990]

XSL•FO
RenderX

82. Herington J, Connelly K, Illes J. Ethical imperatives for working with diverse populations in digital research. J Med Internet Res 2023 Sep 18;25:e47884 [FREE Full text] [doi: 10.2196/47884] [Medline: 37721792]
83. Parry M, Owadally T, O'Hara A, Nickerson N, Hart D. Community- and patient-partner engagement in women's cardiovascular disease research: a rapid review of the evidence. CJC Open 2024 Feb;6(2Part B):485-502 [FREE Full text] [doi: 10.1016/j.cjco.2023.12.016] [Medline: 38487065]
84. Alkureishi MA, Choo ZY, Rahman A, Ho K, Benning-Shorb J, Lenti G, et al. Digitally disconnected: qualitative study of patient perspectives on the digital divide and potential solutions. JMIR Hum Factors 2021 Dec 15;8(4):e33364 [FREE Full text] [doi: 10.2196/33364] [Medline: 34705664]
85. Hill JR, Brown JC, Campbell NL, Holden RJ. Usability-in-place-remote usability testing methods for homebound older adults: rapid literature review. JMIR Form Res 2021 Nov 02;5(11):e26181 [FREE Full text] [doi: 10.2196/26181] [Medline: 34726604]

Abbreviations

CAD: coronary artery disease
HRQoL: health-related quality of life
MRC: Medical Research Council
RCT: randomized controlled trial
REB: Research Ethics Board
REDCap: Research Electronic Data Capture
SMD: standardized mean difference
SUS: System Usability Scale

Edited by K Adapa; submitted 20.02.24; peer-reviewed by EM Noack, SY Fang, S Ania, V Quinten; comments to author 23.05.24; revised version received 10.12.24; accepted 31.03.25; published 17.04.25.

Please cite as:
Parry M, Huang T, Clarke H, Bjørnnes AK, Harvey P, Parente L, Norris C, Pilote L, Price J, Stinson JN, O'Hara A, Fernando M, Watt-Watson J, Nickerson N, Spiteri DeBonis V, Hart D, Faubert C
Development and Systematic Evaluation of a Progressive Web Application for Women With Cardiac Pain: Usability Study
JMIR Hum Factors 2025;12:e57583
URL: https://humanfactors.jmir.org/2025/1/e57583
doi:10.2196/57583
PMID:

©Monica Parry, Tony Huang, Hance Clarke, Ann Kristin Bjørnnes, Paula Harvey, Laura Parente, Colleen Norris, Louise Pilote, Jennifer Price, Jennifer N Stinson, Arland O'Hara, Madusha Fernando, Judy Watt-Watson, Nicole Nickerson, Vincenza Spiteri DeBonis, Donna Hart, Christine Faubert. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 17.04.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Co-Designing, Developing, and Testing a Mental Health Platform for Young People Using a Participatory Design Methodology in Colombia: Mixed Methods Study

Laura Ospina-Pinillos[1], MD, PhD; Débora L Shambo-Rodríguez[1], BSc; Mónica Natalí Sánchez-Nítola[1], PhD; Alexandra Morales[1], BSc; Laura C Gallego-Sanchez[1], MD; María Isabel Riaño-Fonseca[1], BSc; Andrea Carolina Bello-Tocancipá[1], BSc; Alvaro Navarro-Mancilla[2,3], eMBA, MPhil, MD; Jaime A Pavlich-Mariscal[4], PhD; Alexandra Pomares-Quimbaya[4], PhD; Carlos Gómez-Restrepo[1,5], MD, PhD; Ian B Hickie[6,7], MD; Jo-An Occhipinti[6,7,8], PhD

[1]Department of Psychiatry and Mental Health, School of Medicine, Pontificia Universidad Javeriana, Bogota, Colombia

[2]Keralty Medical Centers, Bogota, Colombia

[3]Medicalnet SAS, Bogota, Colombia

[4]Department of Systems Engineering, School of Engineering, Pontificia Universidad Javeriana, Bogota, Colombia

[5]Department of Clinical Epidemiology and Biostatistics, School of Medicine, Pontificia Universidad Javeriana, Bogota, Colombia

[6]Faculty of Medicine and Health, The University of Sydney, Sydney, Australia

[7]Mental Wealth Initiative, Brain and Mind Centre, Faculty of Medicine and Health, The University of Sydney, Sydney, Australia

[8]Computer Simulation and Advanced Research Technologies (CSART), Sydney, Australia

Corresponding Author:
Laura Ospina-Pinillos, MD, PhD
Department of Psychiatry and Mental Health
School of Medicine
Pontificia Universidad Javeriana
Carrera 7 N° 40-62
Bogota, 110311
Colombia
Phone: 57 3208320 ext 2705
Email: laura.ospina@javeriana.edu.co

Abstract

Background: Globally, mental health (MH) problems increasingly affect young people, contributing significantly to disability and disease. In low- and middle-income countries, such as Colombia, barriers to accessing care exacerbate the treatment gap. In addition, the lack of widespread digital interventions further deepens the digital health divide between the Global North and Global South, limiting equitable access to innovative MH solutions.

Objective: This study aims to co-design and develop an MH platform using participatory design methodologies and conduct a 15-month naturalistic observational trial to assess its feasibility among Colombian youth.

Methods: This study used a mixed methods approach within a structured research and development cycle. To ensure a user-centered design, we began with a series of co-design workshops, where stakeholders collaboratively identified key user needs. Following this, usability testing was conducted in 2 stages, alpha and beta, using the System Usability Scale (SUS) to assess functionality and user experience. To capture real-world interactions, a naturalistic observational trial ran from July 2022 to October 2023, collecting data on user engagement and system performance. This study integrated quantitative and qualitative analyses.

Results: A total of 146 individuals participated in the co-design process, with 110 (75.3%) contributing to the development of platform components and 36 (24.7%) participating in usability testing. The co-designed platform integrated several key features, including social media and advertising, an MH screening tool, registration, targeted psychoeducational resources, automated tailored recommendations, and a “track-as-you-go” feature for continuous MH monitoring. Additional elements included user-friendly follow-up graphs, telecounseling integration, customizable well-being nudges, an emergency button, and gamification components to enhance engagement. During usability testing, the beta prototype received a median SUS score of 85.0 (IQR 80-92.5), indicating high usability. In the subsequent observational trial, which ran from July 2022 to October 2023, a total of

XSL•FO
RenderX

435 users interacted with the platform—314 (72.2%) as registered users and 121 (27.8%) anonymously. Emotional distress was prevalent, with 63.7% (200/314) of the registered users and 61.2% (74/121) of the anonymous users reporting distress, as measured by the 6-item Kessler Psychological Distress Scale. Despite 102 users requesting telecounseling, only 26.5% (27/102) completed a session. While usability scores remained high, engagement challenges emerged, with only 18.8% (59/314) of the users continuing platform use beyond the first day.

Conclusions: This study explored the development and user experience of a youth MH platform in Colombia, demonstrating that a cocreation approach is both feasible and effective. By actively involving users throughout the design process, the platform achieved high usability and incorporated features that resonated with its target audience. However, sustaining long-term engagement remains a challenge, as does addressing privacy concerns, particularly for younger users. These findings highlight the importance of continuous user-centered refinement to enhance both accessibility and retention in digital MH interventions.

(JMIR Hum Factors 2025;12:e66558) doi:10.2196/66558

KEYWORDS

low- and middle-income countries; Latin America; telemedicine; medical informatics; eHealth; mental health; help seeking; community-based participatory research; patient participation; patient satisfaction

Introduction

Background

Globally, 1 in 7 young people aged 10 to 19 years are estimated to experience a mental illness [1]. Depression, anxiety, and behavioral disorders are considered some of the main contributors to the burden of disability and disease in young populations [2,3]. Likewise, suicide has been found to be one of the main causes of death in this population [1,4]. Unsurprisingly, the economic and human costs of mental ill-health represent a public health concern where urgent actions are required to improve health outcomes, reduce years of life lost due to preventable mortality and disability, and save millions in health care spending [4]. The impacts of mental ill-health are particularly challenging for low- and middle-income countries, which are home to 90% of the world's youth [5].

Evidence suggests that there can be up to a decade-long delay between the first symptoms of a mental health (MH) disorder and initial contact with services, especially among young people [6]; consequently, the use of digital health tools (DHTs) has gained attention as a potential solution, as young people often seek information online about their MH before discussing it with family members or engaging with MH services [7]. Young people have a preference to find a middle ground between formal and informal sources of information, which opens the opportunity to engage them with MH services or with sources of care [8]. Notwithstanding the value of DHTs, challenges to their adoption exist. First, the internet and digital technologies might not be equally accessible across and within countries [9,10]. Second, even in countries where DHTs are readily available, retention of users and usability are still their main challenges, as it is reported that their use decreases throughout the trials, and most evaluations are limited to short periods [11-13]. To promote the adoption of DHTs [14], participatory design approaches [15,16] offer a promising solution by centering users' needs. This user-driven methodology not only enhances usability but also fosters long-term engagement, ensuring that the technology remains relevant and effectively integrated into everyday practice [17].

Colombia is an upper–middle-income country [18] that faces many societal and economic challenges, further exacerbated by the COVID-19 pandemic [18]. The latest MH survey indicates that 26.5% of adolescents and 26.2% of adults aged 18 to 44 years are likely to experience >2 MH problems [19]. Despite this high prevalence, service use remains low, with only 12.4% of the individuals in urban areas with MH disorders having accessed any health services, according to the World Health Organization [20]. The substantial treatment gap in MH care underscores the critical need for enhanced access to services, particularly given the high prevalence of multiple MH issues among adolescents and young adults. This gap is intrinsically linked to the help-seeking process, as barriers, such as stigma, lack of awareness, and limited availability of resources, prevent individuals from seeking and receiving necessary support [21].

This Study

To the best of our knowledge, there are no comprehensive DHTs available in Bogotá, the capital city, or elsewhere in Colombia that assist young people with their MH. Using participatory design methodologies with end users, this research study explored the ideal design of an MH platform for young people in Colombia, with a focus on essential functionalities and visual appearance to effectively address their MH needs and preferences. We aimed to describe the co-design process and outcomes, outline the development process, and conduct preliminary testing of the minimum viable product (MVP) of this youth MH platform.

Methods

Overview

In line with the World Health Organization framework for developing digital interventions [22], through a well-tested research and development (R&D) cycle [15,23,24], a youth MH platform was co-designed, developed, and tested.

Recruitment involved convenience and snowball sampling, using peer groups and advisory representatives. Participants represented end users of the platform (young people, people with lived experience, their supportive others [eg, carers, family,

XSL•FO
RenderX

friends, and educators], and health professionals [eg, MH professionals and allied health professionals]).

Ethical Considerations

The study was approved by the human research ethics committee of Pontificia Universidad Javeriana and Hospital Universitario San Ignacio (protocol number FM-CIE-0103-21). Using REDCap (Research Electronic Data Capture; Vanderbilt University), informed consent was obtained from adult participants and legal representatives of minors, who also provided their informed assent, as per national regulations. Data were deidentified and safely stored in institutional cloud servers (OneDrive), only accessible to the principal investigator and research officers. Participants in the study received a grocery store voucher in recognition for their time and expertise. For codesign workshops, nonmedical staff and community participants (including young people, people with lived experience, and their supportive others) received a COP $50,000 voucher, while health professionals who participated outside of regular working hours received a COP $100,000 voucher. For user testing sessions, the same groups received COP $40,000 and COP $60,000 Colombian, respectively. These differences in voucher amounts reflected the varying levels of time and effort required for each activity (COP $1000=US $0.26, on average, during project activities).

Co-Design

Co-design workshops (CDWs) aimed to determine the needs and wants of end users as well as understand the functionalities and interface ideas they had for an MH platform. The information collected in these workshops was used to shape the actual development of the platform. What was developed and what it could look like was first determined in a knowledge translation session held after the workshops. The meeting included a multidisciplinary team composed of researchers, psychologists, psychiatrists, legal team members, engineers, and young people and representatives of people with lived experience. Mock-ups were then iteratively refined across each workshop. Furthermore, in between these meetings, several sessions with the reference group were held to always have the end users engaged in the decision-making process.

The CDWs were run using the Google Meet platform [25] in a 2-hour period with around 6 to 8 participants. Information collection was undertaken via surveys, word clouds, Cartesian planes, sketching pads, and scribble sheets [25].

Qualitative data were analyzed using NVivo (Lumivero) and a thematic analysis framework with deductive and inductive coding. An iterative approach guided each phase, and consensus methods were applied at the end of each phase (eg, counts and frequencies). Coding frequencies (CFs) were calculated for all subcategories within the main themes. Interrater reliability was measured using NVivo's coding comparison feature, with 2 researchers independently coding CDW transcriptions until achieving at least a fair agreement based on Cohen κ statistic [26]. Once reliability was reached, the remaining material was individually coded.

Development

Guided by young people's wants and needs and considerations from the multidisciplinary team, several software prototypes were developed as responsive web applications for mobile and desktop devices. Considerations included the assessment of technical feasibility, resource availability, complexity, prioritization, and compliance requirements, among others.

Usability testing was divided into 2 stages. The alpha prototype testing involved 90-minute, one-on-one sessions using a think-aloud protocol, with task difficulty measured by the Single Ease Question scale [27]. Following this, feedback on corrections and reasonable adjustments was provided to the engineering team for the beta prototype. As this version offered a more refined MVP, beta testing followed the same procedures as the alpha sessions but included a more detailed usability evaluation using the System Usability Scale [28].

For data analysis of user testing, sets of notes were generated for each session according to the interviewer's guide. Each section of all sets of notes was read for the first time to identify common themes or recommendations made by participants. Once themes were identified, sections were examined a second time to calculate frequencies (ie, to identify how many participants gave the same recommendation or expressed a similar opinion). Careful attention was given to negative or divergent opinions, for which the number of occurrences was also calculated. This process was carried out for each section of the notes until all sections were completed. Notes from alpha tests were examined first, followed by notes from beta tests. Microsoft Excel was used for frequency calculation.

Testing

To test the MVP, a naturalistic trial was run from July 2022 to October 2023, collecting data on use in real-world settings, including feature use, survey responses, and sociodemographic information.

The quantitative analysis used descriptive statistics to summarize the data. For skewed distributions, the median was used as the preferred measure of central tendency. Variability was assessed through SDs and IQRs, while frequencies and percentages were used to summarize categorical variables. The relationships between categorical variables were analyzed using the chi-square test with a 95% CI. The analysis was conducted using SPSS (version 23; IBM Corp) and Microsoft Excel.

Results

Overview

A total of 146 participants were part of the co-design (n=110, 75.3%) and development (n=36, 24.7%) of the platform's MVP. Most participants were female (n=103, 70.5%), with a median age of 18 (IQR 16-29) years. Middle (tiers 3 and 4) and low (tiers 1 and 2) socioeconomic tiers were the most represented in the sample, with 130 (89%) individuals. Vulnerable populations, such as survivors of armed conflict (n=6, 4.1%), ethnic minority individuals (n=5, 3.4%), and people with disabilities (n=1, 0.7%), were also part of the study.

XSL•FO
RenderX

Co-Design

Overview

A total of 19 CDWs were run, 125 source documents were collected, and 1309 annotations were coded. The deductive coding included 3 themes based on the framework by Ospina-Pinillos et al [15,23]: functionality (actions that can be performed within the platform; annotations: n=504, 38.5%, including all subcategories), user interface (annotations: n=420, 32.09%), and privacy and data management (annotations: n=86, 6.57%). In total, 2 themes emerged, namely technology use for health (annotations: n=111, 8.48%) and youth MH problems (annotations: n=188, 14.36%). Interrater reliability demonstrated a moderate agreement (Cohen κ=0.490).

Results on the use of technology for health and youth MH problems, presented subsequently, highlighted the potential of a technological platform for young people and provided key insights for the development stage. Specific wants and needs regarding functionality, user interface, privacy, and data management are also described subsequently.

Technology Use for Health

Most participants said they used the internet for information seeking (eg, symptoms, tips for managing health concerns, information about healthy lifestyles, and reviews about health professionals), teleconsultation (both from the health care provider and patient perspectives), education, and low intensity treatments (yoga, meditation, and breathing exercises). Information was obtained both from formal sources (eg, e-books) and informal sources, including peer-generated information, influencers, and social media. Acknowledging the amount of misinformation on the internet, participants mentioned developing diverse techniques to triangulate the veracity of the information.

Youth MH Problems

During workshops, each participant contributed to word clouds of youth MH problems, which were used to identify 123 concerns and guide platform development. The most frequently mentioned difficulties, shown in Table 1, involved social relations (n=39, 31.7% mentions), followed by MH disorders (n=28, 22.8% mentions) and problems with self-esteem and self-concept (n=23, 18.7% mentions).

XSL•FO
RenderX

Table 1. Most frequently mentioned youth mental health problems in Bogotá, as reported by co-design workshop participants (N=123).

Category	Mentions, n (%)
Problems with social relationships	39 (31.7)
Family problems	10 (8.1)
Peer problems	8 (6.5)
Bullying (including cyberbullying)	3 (2.4)
Couple problems	2 (1.6)
Others	16 (13)
MH[a] disorders	28 (22.8)
Substance use	9 (7.3)
Depression	9 (7.3)
Anxiety	6 (4.9)
Eating disorders	2 (1.6)
Attention deficit	2 (1.6)
Problems with self-esteem and self-concept	23 (18.7)
Self-esteem	10 (8.1)
Self-concept	7 (5.7)
Life project	6 (4.9)
Emotional difficulties	16 (13)
Emotion regulation	4 (3.3)
Loneliness	4 (3.3)
Stress	3 (2.4)
Resilience	1 (0.8)
Recklessness	1 (0.8)
Hope	1 (0.8)
Boredom	1 (0.8)
Sadness	1 (0.8)
Sexual health	5 (4.1)
Economic difficulties	5 (4.1)
Self-harm and suicide	3 (2.4)
Academic difficulties	3 (2.4)
Pandemic	1 (0.8)

[a]MH: mental health.

Functionality

This theme included 12 subcategories, for which we calculated the CF.

Reliable Information

Participants wanted to receive reliable, targeted, customizable, and periodically updated information (annotations: 52/504, CF=10.3%) delivered in multimedia format. They suggested adapting the platform to users' needs by including an initial test to personalize content and featuring an animated character to guide users and deliver information.

In terms of topics, participants believed the platform should provide general information about MH (eg, warning signs); specific recommendations for healthy lifestyles (eg, nutrition); instructions or tutorials for mental hygiene activities (eg, breathing exercises); and infographics, worksheets, and information on leisure activities (eg, hiking groups). Where relevant, the platform should also redirect users to other, more specialized resources (eg, fitness apps). Information about helplines, spirituality, alcohol and drugs, as well as sexual health was also mentioned. For informative pieces on high-risk topics, such as sexual abuse or suicidal ideation, allied health and MH professionals stated that clear information should be provided about the formal service routes available in Colombia.

XSL•FO
RenderX

Registration

Participants highlighted the value of choosing to explore the platform anonymously or creating an account (annotations: 36/504, CF=7.1%). They proposed to have different user profiles (eg, young people, health professionals, and caretakers), which would receive targeted content.

They proposed the inclusion of different user profiles (eg, young people, health professionals, and caretakers), acknowledging that a young person's MH care benefits from a collaborative support system involving both school and family. By incorporating multiple profiles, the platform would facilitate better interaction and coordination among these key stakeholders, ensuring more comprehensive support for young people:

> *I would create the account so that, let's say, you would have more options, right? Because if you say, "Well, I'm a woman, a teenager, a girl," it would give you more tailored options, right?* [Adolescent from a CDW]

Virtual Care

Participants emphasized the importance of being heard, suggesting the platform should offer options for discussing MH with professionals via videoconference, call, or chat while also highlighting the need for both crisis support and follow-up care (annotations: 33/504, CF=6.5%).

Participants expressed a desire for counseling, guidance, and practical advice tailored to their concerns, with some suggesting virtual care should include professional-led activities such as meditation or life project discussions:

> *I liked the idea of the calls because something similar was done in another project. Two young psychology students would call us monthly or weekly, depending on our schedule, to check in on how we were doing. They even had a project to help us create a life plan, guiding us on what we wanted and providing orientation.* [Adolescent from a CDW]

Others preferred simpler interventions, such as having professionals available to answer specific questions. A blended approach, combining virtual care with face-to-face interactions for initial meetings or home visits, was recommended by some to build rapport and foster empathy.

Screening Tests and Continuous Monitoring of MH Areas

Participants prioritized the ability to evaluate their well-being through tests (annotations: 153/504, CF=30.4%) with immediate results to track progress and express their feelings:

> *It works well because it might help the person express how they're feeling at that moment. Maybe they don't have someone to talk to and share those feelings, like saying, "I haven't been sleeping well" or "I'm feeling sad" or "This happened to me yesterday." It's a way to express yourself while also getting help to feel better and reflect, asking yourself, "What can I do to sleep better?" I'd say it's a good thing.* [Adolescent from CDW]

They wanted diverse response formats (eg, text, video, and audio), customizable survey reminders, and coverage of various topics, such as loneliness, physical health, and emotions. MH professionals suggested tracking coping strategies and integrating the platform into therapy. They emphasized that test results should include actionable feedback or resource links to support improvement or maintain progress.

Gamification

Participants emphasized the value of gamification (annotation: 42/504, CF=8.3%) to engage and motivate young people, suggesting elements, such as video games, avatars, digital medals, and incentives, such as vouchers, or proposing combinations, such as avatars, integrated into games or tied to rewards.

Social Media and Advertising

Participants suggested integrating the platform with social media (annotations: 68/504, CF=13.5%) through visually appealing and relatable advertisements on platforms, such as Facebook, Instagram, TikTok, and Twitter (subsequently rebranded X), focusing on common concerns, such as sleep and mood. They proposed using hashtags for awareness campaigns, adding a share button for easy promotion, and collaborating with influencers or YouTubers to donate content to engage young people and normalize MH concerns.

Well-Being Nudges

Participants acknowledged the diversity of MH concerns and coping strategies among young people. Suggestions included a calendar to track results, a virtual trainer to encourage engagement in activities, and notifications with motivational messages and activity reminders (annotations: 36/504, CF=7.1%).

Emergencies

Participants expressed a desire to either talk to someone immediately or receive guidance from the platform, but they also wanted the option to act independently or support a friend in need. In emergency situations, they emphasized the importance of quickly accessing help (annotations: 3/504, CF=0.6%). In addition, they suggested the ability to notify the platform if another user appeared to require urgent care.

Community Participation

Participants suggested adding blogs and forums (annotations: 60/504, CF=11.9%) moderated by health professionals for safety. They wanted to discuss MH and shared interests, such as soccer or movies, emphasizing that interaction with others could foster support, bonding, and mood improvement.

Chatbot

Participants suggested including artificial intelligence (annotations: 6/504, CF=1.2%) in the platform to provide care during after-hours or weekends, offering general advice, or linking users to specialized resources. They noted that chatbots could also benefit introverted users who may find it difficult to share their concerns.

XSL•FO
RenderX

Interoperability With Local Services

Participants recommended integrating the platform with local services (annotations: 12/504, CF=2.4%) either by referring users to specialized city services or incorporating it into existing programs offered by private and public organizations.

Research and Public Health Surveillance

Participants, particularly MH and allied health professionals, highlighted the research and public health value of an MH platform (annotations: 3/504, CF=0.6%). They suggested using data from continuous monitoring for epidemiological purposes to improve care for young people and integrating it into health surveillance systems.

User Interface

User interface refers to the visual presentation of content and functionalities. Participants stated that the platform should be visually appealing, engaging, and responsive, using inclusive language. Lay terms were preferred over clinical jargon. Customization and interactivity were essential characteristics to be considered across all platform functionalities.

The most important aspect for end users was an eye-catching, aesthetically pleasing appearance. Participants consistently mentioned using images, cartoons, and icons but emphasized that the platform should not appear childish.

Privacy and Data Management

Regarding data management, participants preferred clear, concise information about how their data would be used, favoring simple language over lengthy, jargon-filled documents while still adhering to privacy standards. Concerns were raised about obtaining consent for children aged <14 years, who could not legally provide consent in Colombia. While some younger participants were open to involving their legal guardians, most preferred to engage independently without parental authorization. In addition, most parents expressed a desire to access their children's data:

> *If she were to use the platform, I would want to know absolutely everything—the good, the risky, the threatening, the happy—everything.* [Parent from a CDW]

They acknowledged that children may be hesitant to share their concerns if parents had access to the data, and although most would allow them to use the platform, for some, it could be a sign of poor or declining MH. Some parents indicated they would like to use the platform too, which showcased the multiple ways in which they wished to be involved:

> *Each child is a different world, everyone has their own way of thinking, and not everything in life is complete. I believe that both children and us as parents have our flaws or gaps. When looking at the platform, I think it would be great to use it to explore those gaps or the things we sometimes don't understand about them or wish we could know.* [Parent from a CDW]

Development

Functionality

Reliable Information

To expand access to Spanish-language psychoeducation resources, we integrated our platform with Mental Punto de Apoyo [29], the largest repository of MH materials in the region. This integration allows users to explore a wide range of resources, including videos, pamphlets, and podcasts on MH topics, along with practical techniques to enhance well-being (eg, meditation).

On our platform, these resources are presented with their title and images. When users click on an image, they are redirected to the Mental Punto de Apoyo website for full access. The repository is regularly updated, ensuring users have access to the latest information.

In addition, registered users benefit from tailored recommendations based on periodic survey results and can track a chronological record of all suggested resources, strengthening the connection between both platforms and enhancing the user experience.

Registration

We developed a registration feature that allows users to either create an account or continue anonymously. Anonymous users can access general MH resources but must provide basic demographic details and accept the terms of use. However, registered users gain access to additional features on their home page, including an SOS button, a customizable well-being plan, and avatar selection.

During development, parents and caregivers expressed interest in having their own profiles to better support their children. However, this feature was not included in the MVP.

Virtual Care

To provide access to virtual care, we integrated the platform with Mentes Colectivas, a free web-based program providing counseling and support for individuals aged ≥14 years. The program, accessible on a website [30], offers a safe space where users can seek guidance on emotional well-being, sexuality, and family planning. Users can connect with a team of students, professionals, and volunteers through their preferred digital contact method—telephone call, web-based chat, or web-based call. Appointments can be scheduled within 72 hours, with up to 5 follow-up sessions available as needed. For emotional well-being counseling, support is provided directly through the platform, including follow-up sessions. Users requiring urgent assistance outside of service hours are redirected to emergency services. For sexuality and family planning counseling, users are seamlessly transferred to a designated website using their platform credentials, with unlimited follow-up sessions available if needed [31].

Screening Tests and Continuous Monitoring of MH Areas

An MH screening and continuous monitoring feature ("track-as-you-go" functionality) was developed based on key concerns identified by CDW participants (Table 1). First-time users, whether anonymous or registered, complete an initial

XSL•FO
RenderX

screener using the survey by Kessler et al [32] to assess concerns and distress levels, receiving tailored feedback and psychoeducation through automated algorithms. Registered users can access the continuous monitoring function, selecting from 13 areas (stress, loneliness, well-being, emotions, physical activity, sleep, subjective health status, self-esteem, relationships, eating, alcohol and substance use, bullying, and sexual and reproductive health) to track their progress. Surveys can be scheduled daily, weekly, or biweekly, with a maximum of 2 available at a time to encourage mindful participation. Registered users also gain access to visual progress charts, including line graphs, gauges, and calendars, with options to filter results by date, view them alongside resource histories, and export them as PDFs.

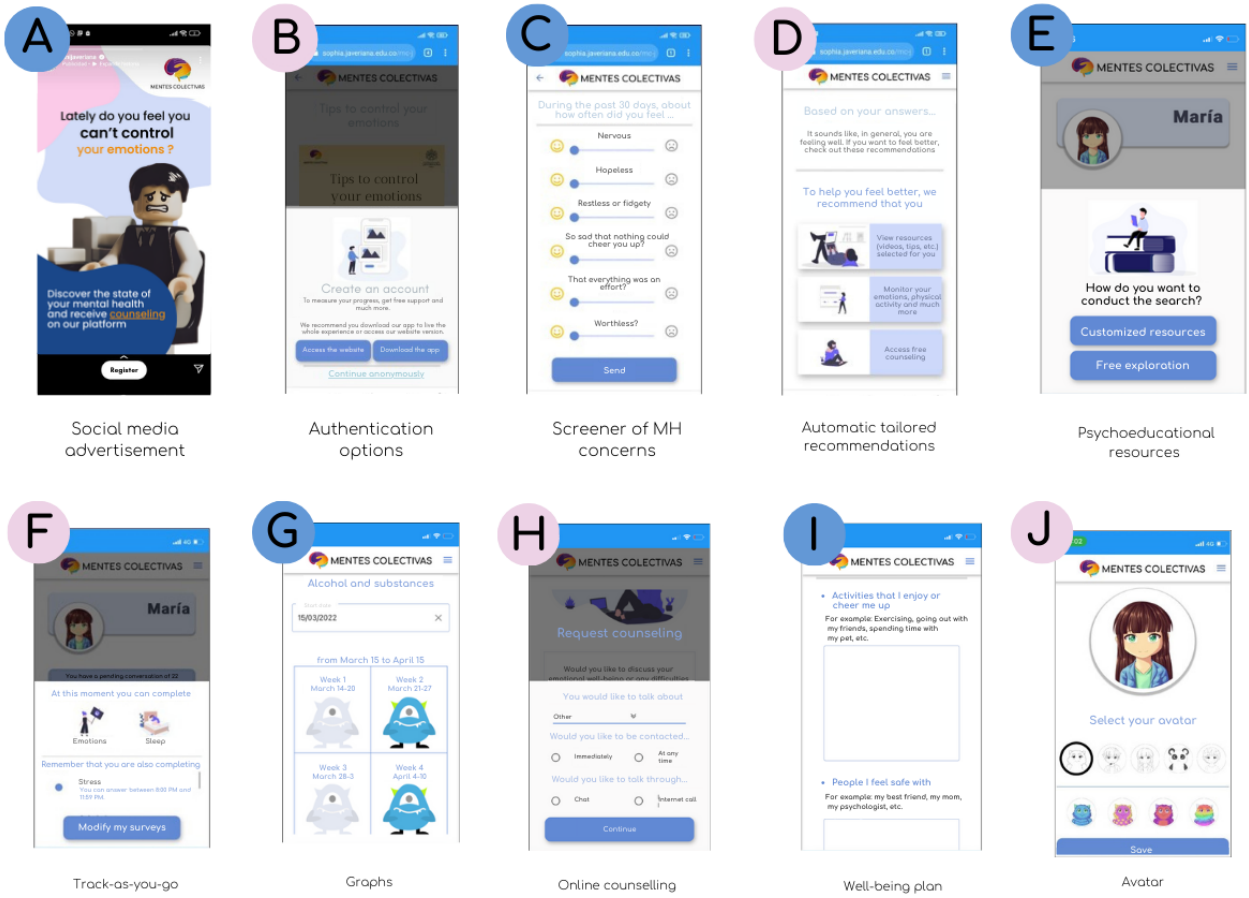
Gamification

During knowledge translation sessions, we recognized the importance of gamification in enhancing user engagement. However, due to the complexity of developing video games and incentive systems, we focused on implementing avatars as the first step. Users can choose from a set of predefined avatars and update them at any time, allowing a degree of personalization. While this feature laid the foundation for a more interactive experience, future iterations may explore additional gamification elements to further enhance user motivation and engagement.

Social Media and Advertising

Recognizing the importance of outreach, we developed a social media campaign featuring advertisements with cartoons to promote the platform and attract users (Figure 1). While awareness campaigns and influencer collaborations could further enhance visibility, they were not included at this stage. This initial approach allowed us to gauge user engagement and refine strategies for future promotional efforts.

Figure 1. User journey of the platform. Platform screenshots were translated from Spanish to English for clarity. The app version of the platform is currently under development. MH: mental health.



From Well-Being Nudges to Well-Being Plan

To encourage young people to engage in MH-promoting activities, we aimed to incorporate well-being nudges throughout the platform. Some of these nudges were successfully integrated into the monitoring graphs, providing gentle prompts to encourage reflection and engagement. However, expanding nudges across the entire platform proved challenging.

As an alternative, we developed the well-being plan, a structured tool inspired by safety plans. It helps users identify uplifting activities, supportive people, and emotional warning signs. Users can personalize their plan with open-text responses and are encouraged to practice beneficial activities when facing challenges. The platform also provides examples of warning signs, such as weight loss or insomnia, ensuring accessibility at all times.

Given development constraints, we prioritized the Wellbeing Plan as the core feature for the MVP, with the potential for further integration of nudges and additional tools in future iterations.

XSL•FO
RenderX

Emergencies

To respond to this need, we designed the telecounseling feature to allow participants to speak with someone immediately from Monday to Friday, 7 AM to 7 PM. During off-hours, the button provided continuous access to emergency local services information (eg, gender-based violence and suicide-related Colombian helplines), ensuring support was always available. Given that the platform is web based, automatic dialing could not be integrated.

Deferred Functionalities

The remaining functionalities were not developed for this MVP due to resource constraints and implementation complexities. Community participation was deprioritized to ensure platform sustainability and focus on essential features such as virtual care, as moderation and content creation required dedicated budgets and personnel. Artificial intelligence functionalities were also not included due to technical challenges and safety concerns, as current large language models in Spanish required significant refinements to ensure reliability and user safety. Interoperability with local services was set aside due to the complexity of institutional agreements, data management, and privacy concerns. Finally, while research and public health surveillance were recognized as valuable, they fell outside the MVP's scope, as they required advanced data management, administrative agreements, and rigorous validation for public health use. Future developments may reassess the feasibility of these features to enhance user experience and support services.

User Journey

The user journey (Figure 1) begins by attracting potential users through targeted advertisements on social media. When users click on an advertisement, they are redirected to the platform's landing page, where they find general information about the platform, MH, and well-being. From there, they can choose to log in and create an account for full access or continue anonymously with limited functionality. If they opt to register, they are guided through a straightforward sign-up process where they provide informed consent, accept the platform's terms and conditions, and acknowledge the habeas data regulations for data protection.

After completing registration, users take a screening test designed to assess their MH concerns. Once they finish the test, they immediately receive feedback on their results and are invited to explore targeted and specialized resources based on their responses. They can also start monitoring their MH through the track-as-you-go feature, which allows them to select areas of interest and continuously track them over time. On the basis of their tracking results, further specialized algorithms provide tailored recommendations and psychoeducational resources. Users can also request a web-based counseling session if needed.

Registered users can access additional features on their home page, including the SOS button, a customizable well-being plan, and avatar selection. As users continue engaging with the platform, they receive increasingly personalized information, whether by tracking additional areas or adjusting their preferences over time. This ensures that the platform evolves with their needs, offering continuous and adaptive support for their mental well-being.

Usability Results

A total of 36 individuals participated in usability sessions for the alpha prototype (n=16, 44%) and the beta prototype (n=20, 56%).

Task Completion Results

In alpha tests, the median task completion time was 43 (IQR 23.5-66.5) seconds, with a median score of 1.00 (IQR 1-1) on the Single Ease Question scale, showing minimal difficulty in performing tasks. The beta tests showed similar results, with a median completion time of 28 (IQR 16-48) seconds and a score of 1.00 (IQR 1-2; Table 2). The track-as-you-go surveys generally showed favorable results with short completion times, except for the physical activity test, which had the longest time and highest difficulty. Accessing resources, returning to profiles, and requesting telecounseling increased completion time without impacting the experience (Table 3).

The beta prototype had a System Usability Scale median score of 85.0 (IQR 80-92.5), indicating acceptable usability. Nearly all participants (18/20, 90%) found the monitoring surveys easy to use. Most (14/20, 70%) found the avatar functionality visually appealing, though some (7/20, 35%) suggested more customization options. The SOS feature was considered useful by all (14/20, 70%), with participants suggesting it should be made more visible.

Improvements between the alpha and beta builds mostly included corrections to the back end of the platform to solve bugs and reduce malfunctioning. Front-end improvements were limited to changes in the wording of instructions, feedback from test results, headings, or descriptions.

XSL•FO
RenderX

Table 2. Completion times and Single Ease Question (SEQ) scores for tasks conducted in alpha and beta user testing sessions. The number of participants who performed the task or responded to the SEQ is presented for each task, as time restrictions or platform malfunctioning impeded all participants from completing all activities[a].

Task	Alpha test		Beta test	
	Completion time (s), median (IQR)	SEQ score, median (IQR)	Completion time (s), median (IQR)	SEQ score, median (IQR)
Creating an account	67.50 (47-89.75)[b]	1.00 (1-1)[b]	61.50 (53.25-91.75)[c]	1.00 (1-1)[c]
Filling the screener of concerns	43.00 (34.25-50.75)[b]	1.00 (2-1)[d]	34.00 (30.5-44.5)[c]	2.00 (3-1)[c]
Accessing MH[e] resources	21.00 (15.25-70)[b]	1.00 (1-1.25)[f]	59.50 (35-83)[c]	1.00 (1-1)[c]
Going to my profile	5.00 (2.75-13)[f]	1.00 (1-1)[g]	6.00 (5-13)[h]	1.50 (1-3)[i]
Requesting a web-based counseling session	44.00 (39-72.5)[j]	1.00 (1-1.5)[j]	76.00 (64-101)[h]	1.00 (1-1)[i]

[a]Median values were used, considering the small samples, to provide a more accurate representation of central tendency.
[b]n=16.
[c]n=20.
[d]n=15.
[e]MH: mental health.
[f]n=14.
[g]n=11.
[h]n=19.
[i]n=18.
[j]n=13.

XSL•FO
RenderX

Table 3. Completion times and Single Ease Question (SEQ) scores for tasks conducted only in beta user testing sessions[a].

Task	Completion time (s), median (IQR)	SEQ score, median (IQR)
Filling out the sociodemographic survey	11.00 (10-14.75)[b]	1.00 (1-1)[c]
MH[d] monitoring[e]		
Stress	23.00 (17–29)[c]	1.00 (1–1)[c]
Loneliness	16.00 (13.75–24)[c]	1.00 (1.75–1)[c]
Well-being	37.50 (29–49.75)[c]	1.00 (2–1)[f]
Emotions	47.00 (32–63)[c]	1.00 (2–1)[c]
Physical activity	51.50 (37.25–68)[c]	3.00 (1.25–5)[c]
Sleep	18.50 (15–24.25)[c]	1.00 (2–1)[c]
Subjective health status	16.00 (11.5–17.25)[g]	1.00 (1–1)[g]
Self-esteem	26.50 (16.25–33.5)[c]	1.00 (3–1)[c]
Relationships	30.00 (26.25–44.75)[c]	1.00 (2–1)[c]
Eating	27.00 (20.25–34.75)[c]	1.00 (1–1)[c]
Alcohol and substance use	35.00 (29.25–41.5)[c]	1.00 (1–1)[c]
Bullying	10.50 (9–15)[c]	1.00 (1–1)[c]
Filling the well-being plan	102.00 (149-86.5)[h]	2.00 (1-2)[h]
Using the SOS button	15.50 (26.5-10)[i]	1.00 (1-1)[i]

[a]Median values were used, considering the small samples, to provide a more accurate representation of central tendency.

[b]n=20.

[c]n=20.

[d]MH: mental health.

[e]During beta sessions, all MH monitoring surveys that the platform offered were tested. An additional test (sexual and reproductive health) was developed after beta sessions had concluded and thus is not included in this table.

[f]n=19.

[g]n=18.

[h]n=13.

[i]n=14.

Reliable Information

Participants (13/16, 81%) from alpha user testing sessions found the topics of the tips useful and relevant. Furthermore, they (12/16, 75%) liked the images and tips design and described them as striking and pleasant. In total, 12% (2/16) of the participants believed the tips were generic or random, and 6% (1/16) of them found the design boring. In the beta sessions, participants (16/20, 80%) expressed their satisfaction with the tips and their content. However, most of them (16/20, 80%) suggested design changes, including color and image updates, and adding more information to the tips by providing links to additional resources. Tips were not accepted by 15% (3/20) of the participants who did not find them useful and suggested changes to the text and more detailed explanations.

Registration

Most (14/16, 88%) of the participants from the alpha testing sessions found the account creation process easy. Only 1 participant was confused by the registration process. In terms of registration options, 31% (5/16) of the participants favored the web version, 31% (5/16) preferred the app version, 12% (2/16) opted for the anonymous route, while the rest (4/16, 25%) did not express any preferences. Most participants (14/16, 88%) understood the sociodemographic questions, and more than half (11/16, 69%) found the images and questions design aesthetically pleasing. The others suggested making changes, especially regarding the gender question and the text description of the platform. Moreover, 19% (3/16) of the participants thought that the survey should collect information about their motive for consultation. The platform's description was changed based on feedback; however, the motive for consultation was not included due to the platform's focus on promotion and prevention, and no changes were made to the gender questions due to limitations in illustration availability.

In the beta sessions, 9 (45%) of the 20 participants preferred logging in to the platform through the web, and 6 (30%) of the 20 participants chose to access the platform anonymously. All participants (20/20, 100%) found the survey easy and clear and

XSL•FO
RenderX

considered the gender question to be acceptable, inclusive, and noninvasive. They made some recommendations related to including more questions or additional answer options.

Virtual Care

In alpha user testing sessions, all participants (13/13, 100%) were able to easily access the feature. They valued its shortness and simplicity but suggested adding more MH topics in the request form. A participant stated that requesting a session involved too many steps. Due to parameterization issues, no topics were added to the request form. In beta user testing sessions, almost all interviewees (18/19, 95%) liked the web-based counseling functionality for its usefulness and flexibility.

Screening Tests and Continuous Monitoring of MH Areas

In alpha user testing, 62% (10/16) of the participants found the screener test easy but suggested adding instructions and more intermediate options. Most (11/16, 69%) participants understood the recommended functionalities, though 25% (4/16) of the participants had trouble with the monitoring feature. Most (12/16, 75%) participants had no issues completing surveys or sharing information but suggested improving the design, adding options to the emotions survey, and making the survey list more prominent. For the beta build, instructions were revised, but intermediate options were excluded to preserve consistency with the original scales. The appearance of the feature on the main profile was also revised, changing the wording of the title to increase clarity.

In beta sessions, nearly all participants (18/20, 90%) liked the screener of concerns, finding it easy, useful, and interactive. They suggested changes to the visibility of the scale, such as making the slider larger, making the answer options visible, or adding instructions (13/20, 65%). More than half of the interviewees (12/20, 60%) understood the 3 recommendations given after the test, with the monitoring feature recommendation being the least understood (7/20, 35%). Nearly all of the participants (18/20, 90%) found most surveys easy to use, and the feedback was positive overall. The most difficult survey was physical activity (6/20, 30%), where participants had trouble making estimates in minutes. They suggested changing the survey to ask about hours; however, due to parameterization restrictions, it was not possible to accommodate this request. Participants also recommended including more targeted resources and additional answer options for the emotions, bullying, and self-esteem surveys, among others. To comply with these recommendations, changes would have to be made to algorithms, which require investments in time and resources. Thus, this recommendation will be considered for further developments. Graphs were also tested in beta sessions, where the calendar type with icons was the most accepted. Participants suggested changes to the icons for most of the graphs, especially the well-being graph (12/20, 60%), followed by the loneliness graph (10/20, 50%).

Gamification

To prioritize other functionalities, the avatar selection feature was only tested in beta sessions. Most of the participants (14/18, 78%) found it visually appealing. Some of them (7/18, 39%) suggested more options, such as color customization and adding more avatar types. In total, 1 participant did not like the anime style.

Integration With Social Media and Advertising

Alpha user testing participants (9/16, 56%) suggested changes in the image and colors to make the advertising more striking and more related to the MH topic. Among the options, the most liked one used cartoons (12/20, 60%) instead of pictures. This advertisement was used in beta sessions where slightly more than half of the participants (13/20, 65%) found it appealing.

Well-Being Nudges

The well-being plan was tested in beta sessions. Participants liked the tool and found it useful and easy to manage. They suggested adding examples of signs of distress or default options to choose from.

Emergencies

The SOS feature was only tested in beta sessions, where all participants (14/14, 100%) found it useful. They valued the variety of helplines and contact options and suggested making it more visible in the main profile.

Testing

Although the platform was designed for young people, feedback from CDWs indicated interest from support individuals, such as parents, caregivers, or health professionals, who expressed a desire to explore the platform before their children or patients used it. In response, we decided to collect demographic information from all users, specifying their role or profile. During the naturalistic trial, no age restrictions were imposed, allowing us to analyze data from all users, including adults, as part of the platform evaluation.

A total of 452 individuals entered the platform. Due to technical difficulties (and hence incomplete information), 17 (3.8%) records were excluded from the analysis. There were 3 registration peaks: the first in November 2022, the second in March 2023, and the third in June 2023. The platform reached 26 of the country's 32 departments; most of the users were from Bogotá (139/314, 44.3%), Cundinamarca (25/314, 8%), Boyacá (23/314, 7.3%), and Antioquia (15/314, 4.8%). During the observed period, 435 individuals used the platform, with 314 (72.2%) users registering and 121 (27.8%) using it anonymously. Most anonymous users were female (89/121, 73.6%) as well as identified to be a young person (31/121, 25.6%). Importantly, more than half (74/121, 61.2%) of the users presented some degree of emotional distress according to the 6-item Kessler Psychological Distress scale (mild distress: 29/121, 24%; moderate distress: 33/121, 27.3%; and severe distress: 12/121, 9.9%) [32].

In relation to those who registered, the median age was 34 (IQR 23-46) years, and most of them (232/314, 73.9%) were female. About the role in the platform, 21% (66/314) of those who registered were young individuals, 26.8% (84/314) were carers, 22% (69/314) were educators, and 20.7% (65/314) were health professionals. Most users (200/314, 63.7%) presented some degree of emotional distress (mild distress: 68/314, 21.7%;

XSL•FO
RenderX

moderate distress: 70/314, 22.3%; and severe distress: 62/314, 19.7%).

To evaluate the relationship between the level of distress and the selected role, a chi-square test of independence was executed. In the registered group, the relationship was statistically significant (χ^2_9=37.9; *P*<.001), and observed frequencies revealed that young individuals had more severe, moderate, and mild levels of distress, while health professionals appeared more frequently with low levels of distress. In addition, an effect size was calculated using Cramér V, which was found to be 0.20. While this effect size is statistically significant, it is small in magnitude. In the group of anonymous users, the relationship was statistically significant (χ^2_9=17.0; *P*=.048), and the Cramér V test showed a weak effect (0.22). The observed frequencies showed that young people had more severe levels of distress, carers and young people had more moderate levels of distress, carers and educators had more mild levels of distress, and health professionals had lower levels of distress more frequently.

Regarding user engagement, 81.2% (255/314) of the individuals used the platform for just 1 day, 8.9% (28/314) used it for 15 days, 2.5% (8/314) used it for 1 month, and 7.3% (23/314) used it for over a month up to 1 year. Among the functionalities, the screening test (314/314, 100%) was the most used as it was mandatory, followed by track-as-you-go (118/314, 37.6%), telecounseling request (102/314, 32.5%), well-being plan (32/314, 10.2%), SOS button (15/314, 4.8%), and finally avatar (14/314, 4.5%). A chi-square test of independence was executed between the level of distress and the platform's use. The relationship was statistically significant (χ^2_{12}=24.5; *P*=.02), and the effect size was calculated using Cramér V, which was found to be 0.16, which meant a weak effect size. The observed frequencies indicated that people who only used the platform on the day of registration were more likely to report no discomfort in the emotional distress screening. In addition, severe discomfort was the most common outcome for people who used the platform during their first month of registration. People who used the platform for 2 months or between 3 and 6 months most often experienced moderate discomfort. By contrast, those who used the platform for >6 months had higher scores in severe and mild discomfort.

Among the individuals who used the tracking component, the most frequently used survey was the emotion tracker (32/118, 27.1%), followed by sleep (25/118, 21.2%), stress (21/118, 17.8%), and physical activity (10/118, 8.5%). The rest of the surveys (loneliness feeling: 6/118, 5.1%; well-being levels: 6/118, 5.1%; self-esteem evaluation: 5/118, 4.2%; risk of eating disorders: 4/118, 3.4%; health perception: 3/118, 2.5%; sexual health: 3/118, 2.5%; relationships: 2/118, 1.7%; and bullying: 1/118, 0.8%) represented less than 5% use. In addition, only a small number of participants opted to continuously monitor any of these areas, with 47% (15/32) tracking their emotions, 44% (11/25) tracking sleep, 24% (5/21) tracking stress, and 20% (2/10) tracking physical activity beyond the initial use.

Concerning the baseline (first time) responses, neutral emotion was the most reported with 41% (13/32), followed by sadness 22% (7/32), fear 19% (6/32), anger 12% (4/32), and joy 6% (2/32). In terms of sleep quality, 96% (24/25) of the users who answered the sleep survey reported sleep problems. With respect to the perceived stress, 71% (15/21) of the participants experienced high levels of stress, 24% (5/21) had some degree of stress, and 5% (1/21) had low stress. In relation to physical activity, participants reported spending a median of 10 (IQR 0-37.5) minutes doing moderate activity per day, with just 10% (2/10) of the users doing it for >60 minutes daily. Median sitting time was 210 (IQR 47.5-345) minutes.

Of the 102 participants who requested telecounseling, only 27 (26.5%) had a complete counseling session. Regarding the reasons for consulting, the most frequent problem was anxiety, reported by 67% (18/27) of the participants, followed by sadness at 11% (3/27), and everyday problems at 11% (3/27). In terms of the form of communication, the platform offered 5 types: immediate chat, chat later, immediate internet call, internet call later, and video call and traditional phone call. Chat later and immediate call options were the main forms chosen by users, with 48% (13/27) and 33% (9/27), respectively. Moreover, the immediate chat option and internet call me later option were less popular among participants, with only 19% (5/27) and 15% (4/27). Finally, video call (1/27, 4%) was the least used option. During the counseling session, only 4% (1/27) of the individuals reported self-harm and suicide plans, which required referral to emergency services.

Discussion

Principal Findings

This study aimed to co-design, develop, and test an MH platform tailored for youth in Bogotá, Colombia. Using a rigorous R&D cycle, we actively engaged a diverse sample, ensuring representation across different ages, socioeconomic backgrounds, and marginalized or hard-to-reach populations. Findings revealed a strong interest in a youth-focused MH platform that prioritized a visually appealing and user-friendly interface, clear privacy protections, and a range of essential features. These included registration, access to reliable information, virtual care options, MH screening and continuous monitoring, gamification, social media integration, personalized well-being plans, emergency support, connections to local services, chatbot interaction, and support for research and public health surveillance. Through knowledge translation sessions, we refined these insights into a platform with 11 key features, including social media integration, registration, MH screening tools, psychoeducational resources, automated recommendations, real-time MH tracking across 13 areas, follow-up graphs, telecounseling integration, customizable well-being nudges, an emergency button, and gamification components. Usability testing confirmed high functionality and acceptance, and a naturalistic trial further evaluated real-world engagement and use patterns. This extensive participatory approach not only ensured that the platform was contextually relevant and user driven but also demonstrated its acceptability, usability, and potential impact in supporting youth MH.

A notable aspect of our research is the extensive engagement of end users throughout the co-design process, with a sample size of 146 participants. Existing literature indicates that most

XSL•FO
RenderX

studies using participatory methodologies in the development of MH DHTs are conducted in high-income countries and often involve relatively small sample sizes (<50 participants). In Latin America, a unique instance of co-design was conducted in Colombia, featuring culturally adapted approaches [24]. Despite the inclusive methodology, men and other genders represented <30% of the sample. This level of participation aligns with research showing that men tend to seek less MH support [33] and engage less in research [34]. This disparity highlights a research gap in understanding the engagement of men and gender-diverse individuals in MH initiatives, which may also impact their adoption and use of DHTs [16]. Future work is needed to address this gap and ensure these technologies are truly inclusive and accessible to all [16,35].

Moreover, the ample sample size of this study allowed for the inclusion of individuals who may act as gatekeepers in young people's MH help-seeking journeys, offering valuable insights into how their involvement can support youth. In addition, the study captured a broad spectrum of perspectives, including diverse and sometimes conflicting opinions, not only regarding design preferences but also in the roles and expectations of caregivers versus young people. These conflicting perspectives highlight a fundamental tension in designing DHTs, that is, balancing the autonomy and preferences of young users with the concerns and involvement of caregivers. However, this diversity of opinions enriches the research process by prompting critical discussions and leading to more adaptable, inclusive, and responsive solutions. Future work should further explore these tensions and develop strategies to ensure that DHTs effectively support both young users and their caregivers while respecting youth autonomy and privacy [36].

This paper provides a comprehensive description of the participatory processes, thereby enriching the literature, as many studies do not provide comprehensive details, thus limiting the reader's understanding of the methodologies used and the extent of user involvement [37]. Our study sought to provide detailed information on the participation of end users at every stage of the development process, including design, feedback provision, knowledge translation, decision-making, and testing. These results highlight the need for end users to be at the center of the design and implementation process rather than just being mere consultants [22,23]. In the same line, the process involved young people and people with lived experience in a meaningful way, avoiding tokenism and promoting genuine participation [38].

In terms of use, engagement is a significant challenge in the eHealth field, with research consistently showing high attrition and low engagement rates [11]. For example, studies indicate that retention rates for mental well-being apps are as low as 3.9% after 15 days and 3.3% after 30 days [39]. In our study, we observed a particularly striking trend on the first day of use, with only 18.8% (59/314) of the users continuing to engage with the platform after the first day. Understanding and improving the first-day experience is crucial for enhancing overall retention and ensuring users continue to benefit from the platform. In our study, participants with higher levels of distress demonstrated a tendency to use the platform for extended periods, aligning with existing evidence [40]. This suggests that the platform may be particularly well-suited for individuals already engaged in the MH care process. Engagement remains a key area for future R&D to address the substantial drop-off after initial use. While our user engagement rates were better (longer periods of engagement) than those reported in previous research, possibly due to our thorough co-design process tailored to users' needs and preferences, achieving long-term engagement requires sustained efforts and innovative strategies to keep users motivated and invested over time. Some alternatives were highlighted by our participants, which referred to integration with local services (in person or virtual), which are already available in Colombia and are known to communities. This synergistic approach could provide more customized follow-ups for service users while expanding the platform's reach to communities outside of social media, creating a blended approach to engagement. However, integration represents a challenge as it requires coordinated responses as well as administrative and data protection agreements.

Our study aimed to contribute to reducing the significant delay in seeking help for MH problems, which averages >10 years [6] from the first symptom to professional contact. By implementing our platform, we provide a potential tool to facilitate earlier access to support, addressing the urgent need for more accessible and timely MH interventions. Internet-based interventions have shown promise in this regard, offering a cost-effective alternative to traditional care [16]. A meta-analysis indicated that these interventions are marginally more effective than usual care in terms of quality-adjusted life years gained, have similar prices, and have a net benefit of US $255 (95% CI US $91-US $419; P=.002) [41]. Furthermore, research has shown that targeted prevention efforts are generally more cost-effective than universal approaches in the youth population [42]. Specifically, interventions that are deployed in community settings (eg, schools or workplaces) and coupled with specific interventions (eg, psychotherapy) have proven highly cost-effective for preventing MH disorders in children and adolescents [42].

Our findings support the need for rethinking services and care pathways, as two-thirds of the users (both anonymous and registered) showed some level of emotional distress. Young users exhibited higher distress levels, with a notable relationship between distress levels and user roles (χ^2_9=37.9; P<.001) as demonstrated in the Results section. These insights emphasize the potential of our platform to identify individuals in need of immediate support. By integrating such technologies with traditional MH services, we can enhance early detection and intervention, ultimately improving MH outcomes. Our findings advocate for the continued innovation, development, and implementation of DHT to detect early signs of distress and integrate these tools with conventional care systems to better support young people and those at risk early in the course of illness.

Limitations

In academia, it is difficult to compete with commercial DHTs that offer highly intuitive and engaging user experiences, even though their content often lacks solid evidence-based support [43,44]. A significant limitation of our study is the absence of

XSL•FO
RenderX

collaboration with industry partners, which could provide valuable resources and expertise in user-centered design and technology development at commercial standards. Partnerships that leverage industry expertise could result in more usable and high-fidelity technologies [44] while also facilitating improved data collection methods, such as the use of phone sensors or wearable devices, to enhance data accuracy and enable more timely notifications, advancing to precision medicine [45]. Furthermore, partnerships with industry could help address the critical challenges of innovation, sustainability, and scalability. Moving forward, it is essential to combine rigorous evidence-based methods with the engaging design features typical of commercial products, along with the advantages provided by industry collaboration, to develop effective, sustainable, and scalable digital health interventions.

Implementation in the Colombian Context

Although participants reported good acceptability of the platform, there are several barriers to its implementation in the Colombian context. A key issue is the discrepancy between young people's needs and legal requirements. Colombian regulations mandate parental consent for access to most health services [46]. While some young people were willing to involve their guardians, most preferred using the platform independently. Parents, while desiring access to their children's information, recognized this might deter openness. To address this, an anonymous browsing option was created, requiring registration only for specialized functions, such as receiving targeted information or telecounseling.

The co-design process was conducted in Bogotá, which, though diverse, may not represent all populations. Further adaptations are needed to tailor the platform to the MH needs of different Colombian communities, including minority groups.

Conclusions

In conclusion, this study successfully co-designed and tested an MH platform tailored for youth in Bogotá, Colombia, to promote help-seeking behaviors for MH issues. Engaging a diverse sample of 146 participants from various backgrounds, the platform addressed the preferences of users across different ages and socioeconomic statuses, demonstrating high acceptability and usability. Unlike many studies with smaller sample sizes, our inclusive methodology underscored the importance of involving end users throughout the development process. However, engagement remains a challenge, with only 18.8% (59/314) of the users continuing to engage with the platform after the first day. Future work must enhance the initial user experience and consider differences in age, sex, region, and socioeconomic background to improve long-term engagement.

Despite the promising results, a significant limitation was the lack of collaboration with industry partners, which could provide valuable resources and expertise in user-centered design and technology development. Partnerships could improve data collection methods, address challenges of sustainability and scalability, and integrate rigorous evidence-based methods with engaging design features. While the platform was well received, implementing it in the Colombian context presents challenges, such as legal requirements for parental consent, which can deter young people from seeking help. Further research and adaptations are needed to tailor the platform to the unique characteristics and MH needs of diverse Colombian communities, ensuring it effectively supports early detection and intervention.

Acknowledgments

This research was funded by Fondation Botnar.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

LO-P, JAO, and CG-R contributed to conceptualization. LOP contributed to supervision. MNS-N, DLS-R, MIR-F, and AM contributed to data collection. MNS-N, DLS-R, AM, MIR-F, and ACB-T contributed to data curation. LO-P, MNS-N, DLS-R, MIR-F, AM, and ACB-T contributed to formal analysis. LO-P contributed to the methodology. DLS-R and MIR-F contributed to validation. LO-P, DLS-R, AM, MNS-N, ACB-T, and LCG-S contributed to writing the original draft. JAO, IBH, CG-R, MIR-F, MNS-N, AP-Q, JAP-M, AN-M, and LCG-S contributed to reviewing and editing the manuscript. AP-Q and JAP-M led the platform development. JAO and LOP contributed to funding acquisition, and CG-R to resources. All authors read and agreed to the published version of the manuscript.

Conflicts of Interest

JAO is the head of systems modeling and simulation and codirector of the Mental Wealth Initiative at the University of Sydney's Brain and Mind Centre (BMC). She is also the managing director of Computer Simulation and Advanced Research Technologies (CSART) and acts as an advisor to the Brain Capital Alliance. IBH is the codirector, health and policy at the BMC, University of Sydney. The BMC operates early-intervention youth services at Camperdown under contract to Headspace. He is the chief scientific advisor to, and a 3.2% equity shareholder in, Innowell Pty Ltd, which aims to transform mental health services through the use of innovative technologies. All other authors declare no conflicts of interest.

XSL•FO
RenderX

References

1. Mental health of adolescents 2021. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health [accessed 2021-11-17]
2. Otto C, Reiss F, Voss C, Wüstner A, Meyrose AK, Hölling H, et al. Mental health and well-being from childhood to adulthood: design, methods and results of the 11-year follow-up of the BELLA study. Eur Child Adolesc Psychiatry 2021 Oct 12;30(10):1559-1577 [FREE Full text] [doi: 10.1007/s00787-020-01630-4] [Medline: 32918625]
3. Patel V, Saxena S, Lund C, Thornicroft G, Baingana F, Bolton P, et al. The Lancet Commission on global mental health and sustainable development. Lancet 2018 Oct 27;392(10157):1553-1598. [doi: 10.1016/S0140-6736(18)31612-X] [Medline: 30314863]
4. GBD 2019 Adolescent Mortality Collaborators. Global, regional, and national mortality among young people aged 10-24 years, 1950-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2021 Oct 30;398(10311):1593-1618 [FREE Full text] [doi: 10.1016/S0140-6736(21)01546-4] [Medline: 34755628]
5. Kieling C, Baker-Henningham H, Belfer M, Conti G, Ertem I, Omigbodun O, et al. Child and adolescent mental health worldwide: evidence for action. Lancet 2011 Oct 22;378(9801):1515-1525. [doi: 10.1016/S0140-6736(11)60827-1] [Medline: 22008427]
6. Wang PS, Berglund PA, Olfson M, Kessler RC. Delays in initial treatment contact after first onset of a mental disorder. Health Serv Res 2004 Apr;39(2):393-415 [FREE Full text] [doi: 10.1111/j.1475-6773.2004.00234.x] [Medline: 15032961]
7. Zaremohzzabieh Z, Ismail IA, Ismail N, Omar SZ, Bolong J. Information and communication technologies and sustainable livelihood opportunities: a study of information behavior among Malaysian youth. Asian Soc Sci 2021 Oct 14;17(11):57. [doi: 10.5539/ass.v17n11p57]
8. Scott J, Hockey S, Ospina-Pinillos L, Doraiswamy P, Alvarez-Jimenez M, Hickie I. Research to clinical practice-youth seeking mental health information online and its impact on the first steps in the patient journey. Acta Psychiatr Scand 2022 Mar;145(3):301-314 [FREE Full text] [doi: 10.1111/acps.13390] [Medline: 34923619]
9. Duggal M, El Ayadi A, Duggal B, Reynolds N, Bascaran C. Editorial: challenges in implementing digital health in public health settings in low and middle income countries. Front Public Health 2022 Jan 10;10:1090303 [FREE Full text] [doi: 10.3389/fpubh.2022.1090303] [Medline: 36703825]
10. Statistics. International Telecommunication Union. URL: https://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx [accessed 2025-04-15]
11. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
12. Vanegas-Osorio JH, Krause-Jacob M, Bauer S, Moessner M, Espinosa Duque D, Carrasco Guzmán Á, et al. Uso de e-mental health para el seguimiento posterapia de pacientes con depresión en Colombia. Revista Virtual Universidad Católica del Norte 2017(52):41-56 [FREE Full text]
13. Baumel A, Muench F, Edan S, Kane JM. Objective user engagement with mental health apps: systematic search and panel-based usage analysis. J Med Internet Res 2019 Sep 25;21(9):e14567 [FREE Full text] [doi: 10.2196/14567] [Medline: 31573916]
14. Vial S, Boudhraâ S, Dumont M. Human-centered design approaches in digital mental health interventions: exploratory mapping review. JMIR Ment Health 2022 Jun 07;9(6):e35591 [FREE Full text] [doi: 10.2196/35591] [Medline: 35671081]
15. Ospina-Pinillos L, Davenport TA, Navarro-Mancilla AA, Cheng VW, Cardozo Alarcón AC, Rangel AM, et al. Involving end users in adapting a Spanish version of a web-based mental health clinic for young people in Colombia: exploratory study using participatory design methodologies. JMIR Ment Health 2020 Feb 08;7(2):e15914 [FREE Full text] [doi: 10.2196/15914] [Medline: 32027313]
16. The Lancet Digital Health. Children must co-design digital health research. Lancet Digit Health 2023 May;5(5):e248 [FREE Full text] [doi: 10.1016/S2589-7500(23)00071-7] [Medline: 37032201]
17. Bucher A, Chaudhry BM, Davis JW, Lawrence K, Panza E, Baqer M, et al. How to design equitable digital health tools: a narrative review of design tactics, case studies, and opportunities. PLOS Digit Health 2024 Aug 22;3(8):e0000591. [doi: 10.1371/journal.pdig.0000591] [Medline: 39172776]
18. Poverty and Shared Prosperity 2022: Correcting Course. Washington, DC: The World Bank; 2022.
19. Encuesta nacional de salud mental. Ministerio de Salud Colombia. 2015. URL: https://www.minjusticia.gov.co/programas-co/ODC/Publicaciones/Publicaciones/CO031102015-salud_mental_tomoI.pdf [accessed 2025-04-15]
20. Borges G, Aguilar-Gaxiola S, Andrade L, Benjet C, Cia A, Kessler RC, et al. Twelve-month mental health service use in six countries of the Americas: a regional report from the World Mental Health Surveys. Epidemiol Psychiatr Sci 2019 Aug 27;29:e53. [doi: 10.1017/s2045796019000477]
21. Cruz A. Poor mental health, an obstacle to development in Latin America. World Bank Group. 2015 Jul 13. URL: https://tinyurl.com/3p4z44es [accessed 2025-04-15]
22. Youth-centred digital health interventions: a framework for planning, developing and implementing solutions with and for young people. World Health Organization. 2021 Jul 19. URL: https://www.who.int/publications/i/item/9789240011717 [accessed 2025-04-15]

XSL•FO
RenderX

23. Ospina-Pinillos L, Krausz RM, Hickie IB. The potential of health information technologies to transform mental health care for young people in low- and middle-income countries. In: Okpaku SO, editor. Innovations in Global Mental Health. Cham, Switzerland: Springer; 2021.
24. Ospina-Pinillos L, Davenport T, Mendoza Diaz A, Navarro-Mancilla A, Scott EM, Hickie IB. Using participatory design methodologies to co-design and culturally adapt the Spanish version of the mental health eClinic: qualitative study. J Med Internet Res 2019 Aug 02;21(8):e14127 [FREE Full text] [doi: 10.2196/14127] [Medline: 31376271]
25. Calvo-Valderrama MG, Marroquín-Rivera A, Burn E, Ospina-Pinillos L, Bird V, Gómez-Restrepo C. Adapting a mental health intervention for adolescents during the COVID-19 pandemic: web-based synchronous focus group study. JMIR Form Res 2021 Nov 03;5(11):e30293 [FREE Full text] [doi: 10.2196/30293] [Medline: 34637395]
26. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. Fam Med 2005 May;37(5):360-363. [Medline: 15883903]
27. Sauro J, Dumas JS. Comparison of three one-question, post-task usability questionnaires. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2009 Presented at: CHI '09; April 4-9, 2009; Boston, MA. [doi: 10.1145/1518701.1518946]
28. Brooke J. SUS: a 'quick and dirty' usability scale. In: Usability Evaluation In Industry. Boca Raton, FL: CRC Press; 1996.
29. Principal - APS - Punto de Apoyo. Pontificia Universidad Javeriana. URL: https://www.javeriana.edu.co/mentalpuntodeapoyo/ [accessed 2025-04-15]
30. Mentes colectivas. Pontificia Universidad Javeriana. URL: https://www.javeriana.edu.co/mentescolectivas/ [accessed 2025-04-15]
31. LaMonica H, Bettancourt Niño PN, Gómez-Restrepo C, Uribe-Restrepo JM, Colón-Llamas T, Escobar Altare A, et al. Providing digital mental health support and guidance across Colombia: an observational study. Digit Health 2025 Apr 13;11. [doi: 10.1177/20552076251330766]
32. Kessler RC, Green JG, Gruber MJ, Sampson NA, Bromet E, Cuitan M, et al. Screening for serious mental illness in the general population with the K6 screening scale: results from the WHO World Mental Health (WMH) survey initiative. Int J Methods Psychiatr Res 2010 Jun;19 Suppl 1(Suppl 1):4-22 [FREE Full text] [doi: 10.1002/mpr.310] [Medline: 20527002]
33. Doherty DT, Kartalova‐O'Doherty Y. Gender and self‐reported mental health problems: predictors of help seeking from a general practitioner. British J Health Psychol 2011 Jan 07;15(1):213-228. [doi: 10.1348/135910709x457423]
34. Otufowora A, Liu Y, Young H2, Egan KL, Varma DS, Striley CW, et al. Sex differences in willingness to participate in research based on study risk level among a community sample of African Americans in North Central Florida. J Immigr Minor Health 2021 Feb 23;23(1):19-25 [FREE Full text] [doi: 10.1007/s10903-020-01015-4] [Medline: 32328873]
35. Naderbagi A, Loblay V, Zahed IU, Ekambareshwar M, Poulsen A, Song YJ, et al. Cultural and contextual adaptation of digital health interventions: narrative review. J Med Internet Res 2024 Jul 09;26:e55130 [FREE Full text] [doi: 10.2196/55130] [Medline: 38980719]
36. Wies B, Landers C, Ienca M. Digital mental health for young people: a scoping review of ethical promises and challenges. Front Digit Health 2021 Sep 6;3:697072 [FREE Full text] [doi: 10.3389/fdgth.2021.697072] [Medline: 34713173]
37. Bevan Jones R, Stallard P, Agha SS, Rice S, Werner-Seidler A, Stasiak K, et al. Practitioner review: co-design of digital mental health technologies with children and young people. J Child Psychol Psychiatry 2020 Aug;61(8):928-940 [FREE Full text] [doi: 10.1111/jcpp.13258] [Medline: 32572961]
38. Moitra M, Owens S, Hailemariam M, Wilson KS, Mensa-Kwao A, Gonese G, et al. Global mental health: where we are and where we are going. Curr Psychiatry Rep 2023 Jul;25(7):301-311 [FREE Full text] [doi: 10.1007/s11920-023-01426-8] [Medline: 37256471]
39. Nwosu A, Boardman S, Husain MM, Doraiswamy PM. Digital therapeutics for mental health: is attrition the Achilles heel? Front Psychiatry 2022 Aug 5;13:900615 [FREE Full text] [doi: 10.3389/fpsyt.2022.900615] [Medline: 35982936]
40. Christensen H, Mackinnon A. The law of attrition revisited. J Med Internet Res 2006 Sep 29;8(3):e20; author reply e21 [FREE Full text] [doi: 10.2196/jmir.8.3.e20] [Medline: 17032636]
41. Rohrbach PJ, Dingemans AE, Evers C, Van Furth EF, Spinhoven P, Aardoom JJ, et al. Cost-effectiveness of internet interventions compared with treatment as usual for people with mental disorders: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2023 Jan 05;25:e38204 [FREE Full text] [doi: 10.2196/38204] [Medline: 36602854]
42. Le LK, Esturas AC, Mihalopoulos C, Chiotelis O, Bucholc J, Chatterton ML, et al. Cost-effectiveness evidence of mental health prevention and promotion interventions: a systematic review of economic evaluations. PLoS Med 2021 May 11;18(5):e1003606 [FREE Full text] [doi: 10.1371/journal.pmed.1003606] [Medline: 33974641]
43. Lyon AR, Dopp AR, Brewer SK, Kientz JA, Munson SA. Designing the future of children's mental health services. Adm Policy Ment Health 2020 Sep 06;47(5):735-751 [FREE Full text] [doi: 10.1007/s10488-020-01038-x] [Medline: 32253634]
44. Bakker D, Kazantzis N, Rickwood D, Rickard N. Mental health smartphone apps: review and evidence-based recommendations for future developments. JMIR Ment Health 2016 Mar 01;3(1):e7 [FREE Full text] [doi: 10.2196/mental.4984] [Medline: 26932350]

XSL•FO
RenderX

45. Babu M, Lautman Z, Lin X, Sobota MH, Snyder MP. Wearable devices: implications for precision medicine and the future of health care. Annu Rev Med 2024 Jan 29;75:401-415 [FREE Full text] [doi: 10.1146/annurev-med-052422-020437] [Medline: 37983384]
46. Sentencia C-246/17. República de Colombia. 2017. URL: https://www.corteconstitucional.gov.co/relatoria/2017/c-246-17.htm [accessed 2025-04-14]

Abbreviations

CDW: co-design workshop
CF: coding frequency
DHT: digital health tool
MH: mental health
MVP: minimum viable product
R&D: research and development
REDCap: Research Electronic Data Capture

Edited by A Parush; submitted 17.09.24; peer-reviewed by P Jackson, D Rickwood; comments to author 20.12.24; revised version received 10.03.25; accepted 07.04.25; published 06.05.25.

Please cite as:
Ospina-Pinillos L, Shambo-Rodríguez DL, Sánchez-Nítola MN, Morales A, Gallego-Sanchez LC, Riaño-Fonseca MI, Bello-Tocancipá AC, Navarro-Mancilla A, Pavlich-Mariscal JA, Pomares-Quimbaya A, Gómez-Restrepo C, Hickie IB, Occhipinti JA
Co-Designing, Developing, and Testing a Mental Health Platform for Young People Using a Participatory Design Methodology in Colombia: Mixed Methods Study
JMIR Hum Factors 2025;12:e66558
URL: https://humanfactors.jmir.org/2025/1/e66558
doi:10.2196/66558
PMID:

©Laura Ospina-Pinillos, Débora L Shambo-Rodríguez, Mónica Natalí Sánchez-Nítola, Alexandra Morales, Laura C Gallego-Sanchez, María Isabel Riaño-Fonseca, Andrea Carolina Bello-Tocancipá, Alvaro Navarro-Mancilla, Jaime A Pavlich-Mariscal, Alexandra Pomares-Quimbaya, Carlos Gómez-Restrepo, Ian B Hickie, Jo-An Occhipinti. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 06.05.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Use of Online Tools for Mental Health Among Racially and Ethnically Diverse College Students: Mixed Methods Study

Sarah Z Hamza[1], BS; Yesenia Aguilar Silvan[1], MA; Lauren C Ng[1], PhD

University of California, Los Angeles, Los Angeles, CA, United States

Corresponding Author:
Lauren C Ng, PhD
University of California, Los Angeles
1285 Franz Hall
Box 951563
Los Angeles, CA, 90095
United States
Phone: 1 310 794 9137
Email: laurenng@ucla.edu

Abstract

Background: Anxiety and depression symptoms have been rising among college students, with many increasingly meeting the criteria for 1 or more mental health problems. Due to a rise in internet access and lockdown restrictions associated with the COVID-19 pandemic, online mediums, such as teletherapy, repositories for mental health information, discussion forums, self-help programs, and online screening tools, have become more popular and used by college students to support their mental health. However, there is limited information about individual-level factors that lead college students to use these online tools to support their mental health.

Objective: This mixed methods study aimed to examine the associations between demographics, symptom severity, mental health literacy, stigma, attitudes, and self-efficacy and the use of online tools to seek psychological information and services among racially and ethnically diverse college students. This study also aimed to qualitatively characterize types of online tools used, reasons for using tools or lack thereof, and perceived helpfulness of tools.

Methods: Undergraduate students (N=123) completed validated measures and provided open-ended descriptions of the types of online tools they used to seek psychological information and services and their reasons for using those tools. Logistic regression analyses were used to test associations of online tool use to seek mental health information and hypothesized predictors. Descriptive statistics were conducted to examine online tool types, reasons for using online tools, and helpfulness explanations.

Results: In total, 49.6% (61/123) of the participants used online tools (eg, search engines) to seek mental health information, while 30.1% (37/123) used online tools (eg, medical websites) to seek mental health services. Mental health literacy (P=.002; odds ratio 1.14, 95% CI 1.05-1.24) was associated with greater use of online tools to seek mental health information. None of the hypothesized variables predicted online tool use to seek mental health services. In total, 82% (50/61) of participants who sought information found online tools somewhat helpful, while 49% (18/37) of participants who sought services found online tools very helpful. Of the students who did not use online tools to seek information, 19% (12/62) reported it was because they did not know which online tools to use and 31% (19/62) stated they would be encouraged to use online tools if it was recommended by professionals, therapists, family, or friends. Of the students who did not use online tools to seek services, 33% (28/86) reported it was because they did not think mental health help was necessary.

Conclusions: These findings highlight the use of online tools to provide mental health information and connect to professional services, suggesting that online tools are widely used to access mental health support.

(JMIR Hum Factors 2025;12:e60628) doi:10.2196/60628

KEYWORDS

mental health; online information; online tool; college students; help-seeking behavior; survey; qualitative; literacy; experience; attitude; perspective; self-efficacy; diverse

XSL•FO
RenderX

Introduction

Mental Health Prevalence and Service Use

Due to an increase in the prevalence of mental disorders among college students over the past decade, exploring and addressing mental health concerns has become highly necessary. In 2024, approximately 70% of college students reported mild to severe depression or anxiety, which is a substantial increase in depression and anxiety symptoms since 2013 [1,2]. College students have reported that mental health problems affected their academic success, with depression symptoms predicting their grade point average and likelihood of dropping out, further highlighting the need to increase access to and use of effective mental health treatments among this population [3,4]. While past-treatment use for mental health problems has increased among college students, young people (aged 18-25 years) have the lowest rates of mental care use among the adult population [1,2,5]. Furthermore, racial or ethnic minority (eg, Black, multiracial, and Arab American) college students have lower rates of mental health care use, including psychological and pharmacological treatment, compared to non-Hispanic White college students, resulting in a need to examine novel tools that can increase mental health service use rates and potentially offset barriers to care for racial or ethnic minority college students [2].

Digital Help-Seeking Behaviors

One way to increase mental health service use rates is to improve help-seeking behaviors, which refers to the act of seeking out and using formal sources (eg, clinical psychologist), semiformal sources (eg, teachers and clergy), informal sources (eg, friends and family), and self-help (eg, unguided website use) sources to address mental health problems [6]. Due to a rise in internet access and lockdown restrictions associated with the COVID-19 pandemic worldwide, online mediums such as virtual counseling and therapy services, repositories for mental health information and resources, discussion forums, self-help programs, and online screening tools have become more popular among college students to support their mental health [7,8]. Global initiatives (eg, World Mental Health International College Student initiative) have prioritized focusing on research that leverages digital mental health interventions to support young peoples' mental health, further demonstrating the potential of digital technologies and online resources to increase service use among college students and potentially address inequities in care access [9,10]. Despite the increased use and research prioritization of online resources to support college students' mental health, there is limited information about individual-level factors that lead college students to use online tools for mental health support. In addition, there is little research examining why college students might opt out of using online tools for mental health support.

Predictors of Digital Tool Use

Past Therapy and Symptom Severity

Research shows that young adults who had previously used formal counseling services started to use those services digitally during the COVID-19 pandemic to support their mental health as they transitioned to online spaces [8]. Positive past experiences with mental health services were associated with greater engagement in digital mental health interventions [11]. Furthermore, a systematic review found that globally, students who had received a mental health diagnosis in the past and wanted to monitor their mental status were more likely to use online mediums compared to those without a diagnosis [7]. Even among students without a mental health diagnosis, being curious about their symptoms or having severe symptoms increased students' likelihood of using online tools, such as educational materials, self-help interventions, screening tools, and online therapies delivered by professionals, for their mental health [7,12,13].

Stigma and Attitudes

Worldwide, stigma is a well-documented barrier to traditional mental health help seeking, often deterring individuals from pursuing care due to concerns about disclosure and social judgment [14]. However, the anonymity, immediacy, and confidentiality of online platforms may reduce concerns about disclosure and social stigma, possibly making them a more accessible option for help seeking [14]. Unlike traditional help seeking, research suggests that stigma does not significantly impact the use of online self-help resources [15]. Similarly, negative attitudes toward mental health or external influences from others do not appear to deter online help-seeking intentions among young people [16,17]. Nevertheless, more research is needed to understand why these factors may not have a strong influence on online help-seeking behaviors.

Mental Health Literacy and Self-Efficacy

Despite the accessibility and anonymity of online tools, barriers such as low mental health literacy continue to limit their use among college students [7,14]. Students with limited mental health literacy may lack the knowledge needed to recognize their symptoms and identify available services. Without this essential knowledge, students may struggle to initiate the online search process, including identifying relevant keywords or finding evidence-based mental health resources and tools online. Another barrier to using online tools may be due to students' self-efficacy [18]. While studies worldwide have explored self-efficacy as a predictor of traditional help-seeking behaviors within this population, research examining the role of self-efficacy in online tool use remains limited [19]. Further investigation is needed to better understand this relationship.

Race or Ethnicity, Gender, Age, and Non–US Born Status

It is possible that there are disparities in the use of online tools based on college students' demographics. One study found that racial and ethnic minority students expressed more interest in teletherapy and online self-help guided mental health support when they were free compared to non-Hispanic White students, but their interest declined when a cost was involved [16]. However, another study found that racial and ethnic minority students were just as likely as their non-Hispanic White counterparts to enroll in and initiate teletherapy [20]. There are also mixed results on the association between gender and use of online tools for mental health, with some studies reporting

XSL•FO
RenderX

no gender differences and others reporting that women are more likely to use online tools than men [7,20,21]. In terms of age, younger adults are more likely to use smartphone apps for their mental health or well-being compared to older adults [22]. Finally, while non–US born immigrants are more reluctant to partake in traditional mental health help-seeking behaviors due to stigma and language barriers, research exploring the relationship between non–US born status and digital help-seeking behavior is limited [23]. For immigrants and refugees, digital tools may still impose barriers to help seeking due to limited language services and the cultural appropriateness of content [24]. These mixed and limited findings highlight the need to explore disparities in the use of digital mental health tools among students from diverse racial and ethnic, gender, and non–US backgrounds.

Objectives of the Study

The relationship between individual-level factors such as symptom severity, stigma, mental health literacy, and online tool use among university students still needs to be further explored to better understand how digital tools can be leveraged to promote mental health equity among college students [21,24]. Consequently, we first examined the association between race and ethnicity, non–US born status, gender, symptom severity, mental health literacy, stigma, self-efficacy, and attitudes toward mental health treatment and the use of online tools for mental health. Then, we identified themes related to factors influencing online tool use and perceptions of the helpfulness of online tools by exploring the types of online tools used, reasons for using or not using online tools, helpfulness of online tools, and motivators for using online tools in the future.

Hypotheses of the Study

This study had four primary hypotheses: (1) we hypothesized that racial and ethnic minority and non–US born students would be more likely to use online tools to seek help due to digital tools addressing stigma-related barriers hindering traditional help seeking among these students [14,25,26]; (2) we also hypothesized that younger students and women would be more likely to use online tools due to existing research showcasing that they are more likely to use online mental health tools, such as smartphone apps, discussion forums, and internet searches [21,22]; (3) we hypothesized that more severe symptoms of depression, anxiety, and trauma would each be independently associated with a greater likelihood of using online tools to seek help as well because symptom severity predicts students' use of different online mental health resources [7,12]; (4) finally, we hypothesized that lower mental health stigma, higher mental health literacy, greater self-efficacy, and more positive attitudes toward help seeking would each be independently associated with a greater likelihood of using online tools for mental health support, as these factors predict traditional help-seeking behaviors [14,18,19,25].

Methods

Study Design and Sample

This study is a cross-sectional mixed methods analysis derived from a larger study examining the use of direct-to-consumer strategies to increase initial engagement in mental health services. Participants (N=123) were recruited through the University of California, Los Angeles (UCLA) Psychology Subject Pool, a university system that allows researchers to recruit participants for research studies, from May 2021 to December 2021. Data collection occurred during the COVID-19 pandemic, during which there were lockdown restrictions in the United States, such as school and nonessential business closures, large gathering bans, and stay-at-home orders [27]. The study was designed to be completed entirely online. The UCLA Psychology Subject Pool system was used to post study advertisements with the inclusion criteria and allow participants to sign up for the study on a secure online platform. The inclusion criteria included undergraduate students at UCLA who were interested in seeking mental health services, were aged at least 18 years, and had corrected to normal vision. The participants represent UCLA's student body, which is racially and ethnically diverse, with approximately 75% of the undergraduate population reporting a racial or ethnic minority identity [28]. In addition, around 15% of the student population at UCLA has indicated the need to use mental health services in the past [29].

Ethics Approval

All study procedures were approved by the UCLA Institutional Review Board (IRB#20-002082). After signing up for the study on the UCLA Psychology Subject Pool system, participants met with an undergraduate-level research assistant through Health Insurance Portability and Accountability Act–compliant Zoom (Zoom Communications, Inc) video call to discuss consent procedures. Only after consenting to participate in the study were participants allowed to complete a series of surveys. Students received 1 course credit for their participation in completing the surveys. Collected data were deidentified and securely stored. Only authorized personnel had access to the data.

Measures

Demographic Form

Participants completed a 14-item self-reported demographic form. Demographic variables included age, gender, assigned sex at birth, race, ethnicity, place of birth, total household income during childhood, and past experiences with medication and the use of therapy or counseling for mental health treatment. Race or ethnicity were categorized as racial and ethnic minorities and White for our sample. Middle Eastern was categorized as a racial or ethnic minority, alongside Asian, Latinx, Black or African American, and multiracial. This categorization aligns with research findings indicating that the designation of "White" does not correspond to Middle Easterners' lived experiences or others' perceptions of them [30]. One participant did not report their race or ethnicity, resulting in 1 missing value for the race or ethnicity demographic variable.

Patient Health Questionnaire-9

The Patient Health Questionnaire-9 comprises 9 questions designed to detect the 9 diagnostic criteria for major depressive disorder according to the *Diagnostic and Statistical Manual of Mental Disorders, 4th Edition* [31]. Responses to these questions

XSL•FO
RenderX

were measured on a 4-point scale from 0 to 3, where 0 indicated "never" and 3 represented "nearly every day." To comply with the institutional review board requirements, item 9, which assesses suicidality, was removed. The total score was calculated by combining reported values from all 8 questions. In this sample, reliability was fairly high (α=.88) [32]. There were no missing data for this variable.

General Anxiety Disorder-7

The General Anxiety Disorder-7 questionnaire consists of 7 questions used to measure each 1 of the 7 core symptoms of generalized anxiety disorder [33]. Responses to these questions were measured using a 4-point scale from 0 to 3, where 0 indicated "never" and 3 represented "nearly every day." The total score was calculated by combining reported values from all 7 questions. In this sample, reliability was fairly high (α=.90) [32]. There were no missing data for this variable.

PTSD Checklist for DSM-5

The PTSD Checklist for the DSM-5 comprises 20 questions, each designed to measure 1 of the 20 corresponding posttraumatic stress disorder (PTSD) symptom criteria outlined in the DSM-5 [34]. Responses for these questions were measured using a 4-point scale from 0 to 3, with 0 indicating "not at all" and 4 referring to "extremely" in terms of how bothered the individual feels by the given PTSD symptom over the past month. The total score was calculated, combining reported values from all 20 questions. In this sample, reliability was deemed fairly high (α=.95) [32]. There were no missing data for this variable.

Mental Health Literacy Scale

The mental health literacy scale consists of 31 questions used to measure the mental health literacy of individuals [35]. The first 15 items were measured using a 4-point scale from 0 to 3, with 0 indicating "very unlikely" or "very unhelpful" and 3 referring to "very likely" or "very helpful." The rest of the questions were scored on a 5-point scale from 0 to 4, with 0 indicating "strongly disagree" and 4 indicating "strongly agree." The total score was calculated by combining reported values from all 31 questions. Our sample indicated fairly high reliability (α=.82) [32]. This scale was incorporated into the original study at a later stage, resulting in missing data for 29% (36/123) of the participants who did not receive this measure.

Internalized Stigma of Mental Illness Scale-10

The Internalized Stigma of Mental Illness-10 scale comprises 10 questions designed to measure the extent of internalized stigma among individuals with psychiatric disorders; responses are rated on a 4-point Likert scale, ranging from 1 (strongly disagree) to 4 (strongly agree) [36]. The total score was calculated by combining reported values for the 10 questions. Our sample had high reliability (α=.74) [32]. There were no missing data for this variable.

Attitudes Toward Mental Health Treatment Scale

The Attitudes Toward Mental Health Treatment Scale consists of 20 questions and is used to assess internalized stigma and attitudes and behaviors related to seeking mental health treatment [37]. Responses to these questions were measured using a 4-point Likert scale, where participants ranked their level of agreement with statements about mental health treatment, with 1 indicating "strongly disagree" and 4 representing "strongly agree". The total score was calculated by combining reported values from all 20 questions. Our sample also indicated fairly high reliability (α=.78) [32]. There were no missing data for this variable.

Self-Efficacy in Seeking Mental Health Care Scale

The Self-Efficacy in Seeking Mental Health Care Scale consists of 9 questions used to measure self-efficacy for mental health care by assessing 2 subscales: confidence in knowledge and confidence in coping [38]. Responses to these questions were measured using a Likert response scale ranging from 1 to 10, where 1 represents no confidence and 10 indicates complete confidence in the individual's ability to conduct a given behavior. For this study, only the total score was calculated by combining reported values from all 9 questions. Finally, our sample had fairly high reliability (α=.89) [32]. There were no missing data for this variable.

Online Tool Use

Questions from the California Health Interview Survey were used to assess online tool use [39]. Participants were asked 2 dichotomous (ie, yes or no) questions: "Have you used an online tool (eg, mobile apps, texting services, webpages, or forums) to learn about problems with your mental health, emotions, nerves, or your use of alcohol/drugs?" and "Have you used online tools (eg, mobile apps, texting services, or webpages) to find, be referred to, contact, or connect with a mental health professional/service?" In addition, participants were asked to rate the helpfulness of the online tools, with response options being very helpful, somewhat helpful, or not at all helpful. There were no missing data for these variables.

To obtain more detailed information about online tool use, participants who responded "yes" were asked to provide open-ended descriptions of the type of online tools used, reasons for their use, and why they were helpful. Those who responded "no" were asked to share open-ended descriptions of reasons for not using online tools and the motivators that would encourage them to use such tools for learning about mental health problems or connecting with mental health services. Open-ended responses varied in length, with some participants providing brief 3- to 4-word answers, while others offered more detailed responses consisting of a few sentences.

Statistical Analyses

Mixed Methods Design

A Qual+QUAN mixed methods design was used to analyze the data [40]. This design incorporated simultaneous collection and analysis of both quantitative and qualitative data for confirmation and hypothesis testing. Quantitative data analysis was the primary method, examining a priori constructs that could impact online tool use. The qualitative data expanded upon the quantitative findings, such that it provided a deeper understanding of the relationship between a priori constructs and online tool use. Specifically, the qualitative analysis aimed to identify themes related to factors influencing online tool use

XSL•FO
RenderX

and perceptions of their helpfulness by exploring the types of online tools participants used and the reasons behind their use.

Quantitative Analysis

To analyze sample characteristics, descriptive analysis was conducted using jamovi (version 2.3; the jamovi project [41]). In addition, associations between student-level factors and the use of online tools for mental health were examined using multiple logistic regressions. These regressions assessed the relationship between demographics, symptom severity, mental health literacy, stigma, attitudes toward mental health treatment, self-efficacy, and engagement with online tools for seeking mental health information and professional services. Finally, due to previous research showcasing past therapy use as one of the few consistent predictors of online tool use for mental health, we controlled for past therapy use for the regression analyses [7,8,11].

Qualitative Analysis

Qualitative data underwent an inductive ("bottom-up") approach, creating codes based directly on the participants' open-ended text responses without applying a preexisting theoretical framework or a priori codes [42]. The first author familiarized themselves with the data, taking notes and highlighting important points. Using 20% of the data, initial codes were generated based on emerging concepts, such as reasons for using or not using online tools. These codes were categorized based on how responses were related (eg, "gave me helpful techniques to calm myself down" and "it taught me how to breathe in stressful situations" were combined into the code "coping mechanisms") and refined iteratively by the first and second authors in weekly coding meetings throughout March 2023. Definitions were developed for each code and category, which were then grouped into themes to create a final codebook. The final codebook consisted of a label, definitions, and exemplar statements derived from the data. The first author then coded the open-ended responses based on the final codebook. Coding decisions were reviewed during consensus meetings held with the second author, which is an approach commonly used to address coding biases and ensure coding reliability [43]. Open-ended responses that were not consistent with the codebook definitions were also discussed during consensus meetings. When responses were too vague (eg, "online resource" and "I don't know"), it was decided to categorize those statements as "vague" rather than forcing them into previously established categories in the final codebook. After coding all the responses, the codes and categories were interpreted at the theme level. Thematic analysis allowed us to identify overarching patterns within the data to better understand factors influencing online tool use and perceptions of the helpfulness of online tools.

Results

Sample Characteristics

A total of 123 participants completed the survey and were aged on average 21 (SD 3.39) years. The majority self-identified as women (104/123, 84.6%) and belonged to racial or ethnic minority groups (98/123, 79.7%). Racial or ethnic minority participants were primarily Asian (47/123, 38.2%) followed by Middle Eastern (19/123, 15.4%), Latinx (16/123, 13%), multiracial (11/123, 9%), and Black or African American (5/123, 4.1%). Furthermore, a substantial proportion of participants were born outside the United States (32/123, 26%) and spoke non-English languages at home (55/123, 44.7%). Finally, more than half (65/123, 52.8%) of the participants had received therapy or counseling for a mental health condition in the past, but a majority (90/123, 73.2%) had not taken medication for a mental health condition in the past. Refer to Table 1 for more demographic information.

Using the symptom checklist cutoffs for the Patient Health Questionnaire-9, General Anxiety Disorder-7, and PTSD Checklist for the DSM-5, participants on average reported mild depression (mean 8.16, SD 5.74) and mild anxiety (mean 7.06, SD 5.60) and did not meet criteria for a probable PTSD diagnosis (mean 20.6, SD 18.0). The mean mental health literacy score was 117 (SD 9.14), indicating high mental health literacy. The mean score for internalized mental illness stigma was 13.2 (SD 7.33), suggesting low stigma. The mean score for attitudes toward mental health treatment was 57.4 (SD 6.48), indicating more positive attitudes toward mental health. The mean score for self-efficacy in seeking mental health care was 64.2 (SD 15.1), suggesting high self-efficacy.

XSL•FO
RenderX

Table 1. Sociodemographic characteristics of participants grouped by online tool use.

	Full sample (N=123), n (%)	Online tool use to seek mental health information		Online tool use to connect to professional service	
		Yes (n=61), n (%)	No (n=62), n (%)	Yes (n=37), n (%)	No (n=86), n (%)
Gender					
Man	18 (15)	8 (44)	10 (56)	6 (33)	12 (67)
Woman	104 (85)	52 (50)	52 (50)	30 (29)	74 (71)
Agender	1 (0.8)	1 (100)	0 (0)	1 (100)	0 (0)
Race[a]					
Racial/ethnic minority	98 (80)	45 (46)	53 (54)	28 (29)	70 (71)
White	24 (20)	15 (63)	9 (37)	9 (37)	15 (63)
Ethnicity[b]					
Asian	47 (39)	20 (43)	27 (57)	12 (26)	35 (74)
Black or African American	5 (4)	1 (20)	4 (80)	1 (20)	4 (80)
Latinx	16 (13)	9 (56)	7 (44)	4 (25)	12 (75)
Middle Eastern	19 (16)	9 (47)	10 (53)	5 (26)	14 (74)
Multiracial	11 (9)	6 (55)	5 (45)	6 (55)	5 (45)
Non–US born status					
No	91 (74)	47 (52)	44 (48)	31 (34)	60 (66)
Yes	32 (26)	14 (44)	18 (56)	6 (19)	26 (81)
Language spoken at home					
English	68 (55)	35 (51)	33 (49)	23 (34)	45 (66)
Non-English	55 (45)	26 (47)	29 (53)	14 (25)	41 (75)
Previous therapy use					
Yes	65 (53)	38 (58)	27 (42)	27 (42)	38 (58)
No	58 (47)	23 (40)	10 (17)	10 (17)	48 (83)
Previous medication use					
Yes	33 (27)	24 (73)	16 (48)	16 (48)	17 (52)
No	90 (73)	37 (41)	21 (23)	21 (23)	69 (77)

[a]Missing value for 1 participant (n=122).

[b]Breakdown of ethnicities for the category racial and ethnic minority (n=98).

Quantitative Results

In total, 49.6% (61/123) of the participants used online tools to seek mental health information, while 30.1% (37/123) used online tools to find, be referred to, contact, or connect with a mental health professional or service.

Predictors of Online Tool Use to Seek Mental Health Information

The multiple logistic regression model was not statistically significant for online information seeking (N=86, χ^2_{12}=19.2; *P*=.08). Age, gender, race and ethnicity, non–US born status, past therapy use, symptom severity, stigma, self-efficacy, and attitudes toward mental health treatment did not predict online information seeking. However, mental health literacy was associated with an increase in the likelihood of online information seeking, such that odds of online tool use to seek mental health information increased by 14% for every additional increase in mental health literacy (odds ratio 1.14, 95% CI 1.05-1.24).

Predictors of Online Tool Use to Seek Mental Health Services

The multiple logistic regression model was not statistically significant for seeking mental health services online (N=86, χ^2_{12}=20.4; *P*=.06). None of the hypothesized predictors were associated with online help-seeking behaviors (Table 2).

XSL•FO
RenderX

Table 2. Regression analyses of predictors of online tool use.

	β (SE)	Wald χ^2 (*df*)	*P* value	Odds ratio (e^{β}; 95% CI)
Predictors of online tool use to seek mental health information				
Age	−.07 (.07)	1.1 (12)	.30	0.93 (0.81-1.1)
Gender	.68 (.77)	0.8 (12)	.37	1.98 (0.44-8.9)
Race or ethnicity	−.22 (.68)	0.1 (12)	.74	0.80 (0.21-3.0)
Non–US born	−.10 (.62)	0.03 (12)	.87	0.90 (0.27-3.1)
Previous therapy	.59 (.64)	0.9 (12)	.35	1.8 (0.52-6.3)
Depression	.02 (.08)	0.04 (12)	.85	1.0 (0.87-1.2)
Anxiety	−.12 (.10)	1.6 (12)	.22	0.89 (0.73-1.1)
PTSD[a]	.03 (.03)	1.0 (12)	.34	1.0 (0.97-1.1)
Mental health literacy	.13 (.04)	12.4 (12)	.002	1.14 (1.1-1.2)
Internalized stigma	.01 (.05)	0.1 (12)	.74	1.02 (0.92-1.1)
Attitudes toward mental health treatment	−.03 (.06)	0.3 (12)	.59	0.97 (0.87-1.1)
Self-efficacy	−.01 (.02)	0.2 (12)	.69	0.99 (0.94-1.0)
Predictors of online tool use to connect with professional mental health services				
Age	.13 (.09)	2.7 (12)	.15	1.14 (0.95-1.4)
Gender	.43 (.84)	0.3 (12)	.61	1.54 (0.30-8.0)
Race or ethnicity	.78 (.79)	1.0 (12)	.33	2.18 (0.46-10.3)
Non–US born	.82 (.76)	1.2 (12)	.28	2.26 (0.51-10.1)
Previous therapy	−.31 (.74)	0.2 (12)	.67	0.73 (0.17-3.1)
Depression	−.19 (.10)	4.5 (12)	.05	0.83 (0.68-1.0)
Anxiety	.04 (.11)	0.1 (12)	.74	1.04 (0.84-1.3)
PTSD	.06 (.03)	3.9 (12)	.07	1.07 (0.99-1.2)
Mental health literacy	−.02 (.04)	0.3 (12)	.61	0.98 (0.91-1.1)
Internalized stigma	.04 (.06)	0.5 (12)	.47	1.04 (0.93-1.2)
Attitudes toward mental health treatment	.08 (.07)	1.6 (12)	.22	1.09 (0.95-1.2)
Self-efficacy	.01 (.03)	0.2 (12)	.63	1.01 (0.96-1.1)

[a]PTSD: posttraumatic stress disorder.

Qualitative Results

Types of Online Tools Used

Qualitative results indicated that participants used search engines, medical websites (eg, BetterHelp and Psychology Today), informational websites, apps (eg, MindDoc, LiveHealth, and Calm), and university-affiliated tools to seek mental health information and connect to services. Refer to Table 3 for a summary of qualitative responses.

XSL•FO
RenderX

Table 3. Summary of qualitative responses on types of online tools used for seeking mental health information and connecting with professional services (N=123).

Tool type codes	Definition	Seek mental health information[a] (n =61), n (%)	Connect with professional services[b] (n =37), n (%)
Search engines	Tools that carry out web searches to identify websites that correspond to given keywords	20 (48)	3 (8)
Medical websites	Websites with information related to mental health diagnosis, symptoms, or professional services	12 (20)	11 (30)
Informational websites	Websites that provide general background information and research on mental health disorders and coping mechanisms or exercises	6 (10)	2 (5)
Apps	Downloadable apps that can be used for mental health information, services, or resources	5 (8)	4 (11)
Discussion forums	Internet forums or message boards where an individual can post about issues or reply to other individual's posts that they may find relatable	4 (7)	—[c]
Social media	Apps that can be used to share content	1 (2)	—
University-affiliated tools	Tools created by universities to provide mental health services or information to students	1 (2)	9 (24)
Call, video, and SMS text messaging	Platforms that can be used to connect with family, friends, or professionals through voice, video, or messages	—	4 (11)
Insurance	Tools created by insurance companies to be used for insurance-covered services and information	—	3 (8)
Vague statements	Statements that are too vague to code under a given category	3 (5)	1 (3)

[a]Codes endorsed for seeking mental health information online.

[b]Codes endorsed for connecting with professional mental health services online.

[c]Not applicable.

Reasons for Using Online Tools

Qualitative findings yielded 3 themes related to the reasons why participants used online tools. Participants used online tools because they (1) sought psychoeducation; (2) recognized a mental health need; and (3) perceived online tools as beneficial due to increased accessibility, usability, and availability of support.

A primary reason participants used online tools was to increase their knowledge and to access psychoeducation about mental health. Online tools, such as discussion forums and social media, were frequently used for seeking mental health information, including gaining a better understanding of their own mental illness (eg, understanding PTSD and anxiety symptoms) and due to having a general interest in mental health topics. For example, participants stated that they wanted to "better understand PTSD and ways to cope with anxiety" and to learn more about their "symptoms at the time and to see if they can explain any other bodily functions (ie, pain or aches, inconsistent menstruation cycles, and headaches)." Participants also used online tools to learn about coping skills that they could use to manage their mental health concerns. Participants reported wanting to learn about a variety of coping skills, such as exercise, sleep, nutrition (eg, "vitamin intake"), and "emotion regulation and meditation skills."

Another reason participants used online tools to seek mental health information and connect with a mental health service was due to perceived need concerns, such as recognizing a mental health need and noticing that their symptoms were becoming more severe. For example, a participant reported, "I had been having anxiety that led to a panic attack and decided I needed to seek help in managing my anxiety," while another mentioned that "I was having depersonalization panic attacks and did not understand what was happening to me." Participants also used online tools, such as voice, video, and messaging tools, along with tools created by insurance companies, to connect with mental health services. Participants reported actively seeking help, expressing their intent to "find a therapist" and choosing online platforms because they "knew this is where" they "could find a mental health professional." Participants stated that these online tools were often "recommended by someone trusted," such as friends or family, who were aware of the participants' mental health needs.

Participants also used online tools to find mental health information and connect with services because they found them beneficial. Many participants cited accessibility as a key factor, noting that they often could not find other traditional support and resources in a timely manner and instead turned to online tools because "it was fast, easy, and free." Other participants highlighted usability as a key factor, describing online tools as convenient and intuitive, with a participant stating, "It's convenient and I don't need to ask someone else" and another expressing, "It's quick and easy, and I can look for any information and get different input." In addition, some participants used online tools to connect with others who shared similar mental health experiences, seeking a sense of community and emotional support "to feel less alone." Refer to Table 4 for summary of themes and codes related to reasons for using online tools.

XSL•FO
RenderX

Table 4. Summary of qualitative themes and codes related to reasons for using online tools and perceived helpfulness (N=123).

Theme and code	Definition	Seek mental health information[a] (n =61), n (%)	Connect with professional services[b] (n =37), n (%)
Reasons for using online tools			
Sought psychoeducation			
Knowledge about personal symptoms	Wanting to learn more about mental health disorders and their symptoms	19 (31)	—[c]
Generalized interest	Had a generalized interest in learning more about mental health	10 (16)	—
Coping skills	Searching for techniques, strategies, or exercises to help cope with mental health issues or boost self-efficacy	9 (15)	—
Perceived benefits: accessibility, usability, and availability of support			
Accessibility	Increased access to mental health services, information, or resources and usability	8 (13)	9 (24)
Lack of availability	Other resources being unavailable	1 (2)	1 (3)
Connecting with community	Wanting to connect with others with similar mental health issues, wanting to not feel alone, or searching for a sense of community	4 (7)	—
Perceived need			
Symptom severity	Feeling a heightened personal struggle with mental health, elevated symptoms, or using tools as a last resort	6 (10)	6 (16)
Testimonial	Receiving a recommendation to use an online tool from another individual	1 (2)	2 (5)
Help-seeking behaviors	Vague statements wanting to generally seek help from professionals or services	1 (2)	17 (46)
Vague statements	Statements were too vague to code under a given category	2 (3)	2 (5)
Helpfulness of online tools			
Provided valuable psychoeducation			
General information	Provided broad information about mental health	18 (31)	3 (9)
Knowledge from personal experience	Provided information about mental health disorders and their symptoms, so the individual was able to better understand	16 (27)	—
Coping mechanisms	Provided techniques, strategies, or exercises to help cope with mental health issues	11 (19)	—
Normalizing mental health	Normalized mental health disorders, symptoms, and treatments and boosted the individual's confidence in their ability to improve mental health	3 (5)	—
Confirmed previous knowledge	Provided information that confirmed what they already knew about mental health issues	1 (2)	—
Help-seeking behaviors	Provided resources to seek help through a professional or service	1 (2)	15 (44)
Attainable and convenient support			
Accessibility	Provided resources or services in an easily accessible manner	1 (2)	8 (24)
Specialty filter	Provided a tool to filter through website content to cater it to individuals' needs	—	2 (6)
Delayed care	Eventually provided connection to care but had some challenges and barriers	—	3 (9)
Connections with community	Provided a connection to others with similar mental health issues or made the individual feel less alone	5 (9)	—
Vague statements	Statements that are too vague to code under a given category	3 (5)	3 (9)

[a]Codes endorsed for seeking mental health information online.

[b]Codes endorsed for connecting with professional mental health services online.

[c]Not applicable.

XSL•FO
RenderX

Helpfulness of Online Tools

Most participants reported that the online tools they used to seek mental health information were somewhat (50/61, 82%) and very (9/61, 15%) helpful. Online tools used to connect to services were also reported to be somewhat (16/37, 43%) and very (18/37, 49%) helpful.

Qualitative findings yielded 2 themes related to the reasons why participants found online tools helpful. Online tools were perceived as helpful because they (1) provided valuable psychoeducation and (2) made obtaining support more attainable and convenient.

Students found online tools to be helpful because they provided information about symptoms of mental health disorders (eg, anxiety, depression, and panic attacks), coping mechanisms (eg, techniques to calm down), general mental health facts (eg, "Provided me with more information than I originally had"), and ways to seek help (eg, resources on using university counseling and psychological services). For example, a participant reported that online tools "allowed me to have more understanding of myself before seeking treatment" and "they helped me realize that these feelings were symptoms of anxiety and depression." Given that psychoeducation was the primary reason for both using online tools and finding them helpful, this indicates that most participants who used online tools were able to find the information they were initially searching for. Some participants expressed that online tools were also helpful because the psychoeducation normalized mental health (eg, "They helped normalize my behavior and gave me the confidence to make progress with my CBT") and confirmed previous knowledge they had about mental health (eg, "Didn't tell me new information, just confirmed what I already had expected").

Participants expressed that online tools were helpful because they made obtaining support more attainable and convenient, with a participant mentioning, "It was a lot easier to schedule a time," and other participants stating, "it gave me support when I needed it." Participants also expressed that online tools were easy to use, with a participant explaining, "The site was extraordinarily simple to use." Certain features, such as specialty filters, made support more attainable, with a participant stating, "I was able to specify my search via provider specialty and modality." However, a couple of participants experienced delays in accessing care, as a participant mentioned, "It took so long to go through the list, but I eventually found someone." In addition, online tools made it easier to connect to others who were dealing with similar feelings or symptoms, resulting in participants feeling less alone. For example, a participant shared, "It is nice to know that others feel the same way that you do and that you can feel more normal because of it." On the basis of these responses, participants who had aimed to connect to community and mental health services using the online tools found that they were able to effectively do so. Refer Table 4 for themes and codes related to online tool helpfulness.

Reasons for Not Using Online Tools

In total, 50.4% (62/123) of the participants reported that they did not use online tools to seek mental health information, and 69.9% (86/123) of the participants reported that they did not use online tools to connect with professional mental health services.

Qualitative findings yielded 3 themes related to why students did not use online tools. Participants did not use online tools due to (1) a lack of awareness about existing online tools, (2) negative attitudes toward online tools, and (3) a lack of perceived need.

Many participants refrained from using online tools to seek mental health information or connect with services because they did not know they existed. Participants reported a lack of awareness about suitable platforms, with some explaining they "have not seen them" or "didn't know of any" online tools for mental health information or connecting to services.

For participants who were aware of online tools, they expressed not using online tools because they did not "trust them" and because they believed that online tools were "not personalized or accurate." Along with finding online tools unreliable or unhelpful, some participants disclosed a preference for traditional help, such as already having a "personal therapist" or receiving in-person referrals from trusted individuals (eg, primary care providers, family, and friends).

Another commonly reported reason for not using online tools to find mental health information or connect to services was due to a lack of perceived need, as participants reported that mental health help was unnecessary or that they were not dealing with severe symptoms. For example, participants stated, "I don't believe that I have needed to connect with a professional/service" and "I don't think it has reached a point for me where I need to seek professional help." Others expressed avoidance and self-reliance tendencies, explaining that they did not "put much thought into learning about mental health" or "how to feel better," and that "I tend to deal with any problems I have by myself." Refer to Table 5 for a summary of themes and codes explaining why online tools were not used.

XSL•FO
RenderX

Table 5. Summary of themes and codes explaining why online tools were not used to seek mental health information and connect with professional services, as well as factors motivating increased use (N=123).

Theme and code	Definition	Seek mental health information (n =62), n (%)	Connect with professional services (n = 86), n (%)
Reasons for the lack of use			
Lack of awareness			
Lack of awareness	Reasoning related to not knowing tools existed or which tools to use	12 (19)	10 (12)
Attitudes toward online tools			
Unreliable	Reasoning related to not finding tools to be reliable or trustworthy	8 (13)	2 (2)
Inaccessible and unhelpful	Reasoning related to ease of tool use or believing the tools would not be helpful	5 (8)	4 (5)
Traditional help preference	Reasoning related to preferring to talk to a professional or trusting a professional more	10 (16)	19 (23)
Lack of perceived need			
Mental health help is unnecessary	Reasoning related to broadly believing mental health help or services to be unneeded, so no online tool was used	11 (18)	28 (34)
Low symptom severity	Reasoning related to not experiencing mental illness or any problems with mental health	10 (16)	10 (12)
Avoidance	Reasoning related to not wanting to resurface negative feelings or deal with problems	2 (3)	—[a]
Self-reliance	Reasoning related to dealing with mental issues by themselves or on their own	1 (2)	1 (1)
Vague statements	Statements that are too vague to code under a given category	3 (5)	9 (11)
Motivating factors			
Reliability			
Recommendation	Suggestions related to referral from friends, family, or professionals; support and encouragement; and positive reviews and testimonials	19 (31)	19 (22)
Trustworthiness	Suggestions related to reliability and confidentiality	4 (7)	4 (5)
Accessibility, availability, and usability of support			
Specific features	Suggestions related to specific website and app features, technological services, personalization, and usability	10 (16)	7 (8)
Accessibility	Suggestions related to finances, time, and other aspects of increased accessibility	6 (10)	10 (12)
Awareness	Suggestions related to advertising or promoting awareness	4 (7)	8 (9)
Perceived need			
Symptom severity	Suggestions related to an increase in negative symptoms or behaviors or an increase in negative impact on everyday life	6 (10)	13 (15)
Help seeking	Suggestions related to broadly wanting to seek help	2 (3)	3 (4)
Personal preferences			
No motivators	Includes reasoning related to individuals currently in therapy or services	4 (7)	8 (9)
Vague statements	Statements that are too vague to be coded under a given category	6 (10)	14 (16)

[a]Not applicable.

XSL•FO
RenderX

Motivators for Using Online Tools in the Future

Of participants who reported not using online tools, a majority expressed motivation to use online tools to seek mental health information (58/62, 94%) or connect with services (78/86, 91%).

Among those who did not use online tools, 3 key factors emerged as potential motivators for future use. Participants expressed they would be motivated to use online tools for their mental health in the future if (1) they perceived a mental health need, (2) they trusted that the tools were reliable, and (3) the tools were accessible and easy to use.

Participants expressed that they would be motivated to use online tools if they found their symptoms to be severe or if their symptoms started to impact their functioning. For example, students stated that they would use online tools "If stress/depressed state impacted my other responsibilities and friends/family" and "if my anxiety keeps getting worse after the school year ends."

Participants expressed that recommendations from trusted individuals, such as "a professional," "family or friends," or a "therapist," would motivate them to use online tools, as recommendations would indicate that the online tool was reliable and trustworthy. One participant stated they would use an online tool "if I heard a good review from people I know," while another added, "if a professional gave me some to use." Other participants noted encouragement from others would also increase the likelihood of using online tools to seek help, such as "knowing that it's encouraged to reach out for help may make me feel more inclined to reach out if I found myself in the position."

Regarding accessibility, participants reported that they would be open to using online tools if they were "free" or "not expensive" and if the online tools used "easy to understand language" and provided "easily accessible information." Many participants also emphasized that raising awareness of the online tools would improve their accessibility, stating "awareness of their existence and that they have proven to be effective" and "more advertisement" as key motivators. Finally, students suggested that specific features improving the usability of the online tool, such as "video chats, scheduled appointments, surveys," or "a comprehensive search filter on the app so you can find the best therapist for your needs" would make them more willing to use the online tool in the future. They also emphasized that the online tool should have "no association to emails, names, or locations" to improve usability. Refer to Table 5 for additional motivating factors that would encourage the use of online tools for mental health.

Discussion

Principal Findings and Comparison to Previous Work

Due to the high prevalence of mental health problems, low use of mental health services among university students, and the rise of the internet, there has been an increase in online tools available for mental health, including online tools to reach and connect youth to mental health services [2,5]. However, the availability of digital tools for mental health has not translated to widespread use [44]. Our study found that about half of the participants (61/123, 49.6%) used online tools to seek mental health information, and only a third (37/123, 30.1%) used online tools to connect with mental health professionals or services. Regarding the types of online tools used, our findings are consistent with previous reports stating that a high proportion of young people use search engines to locate mental health resources, with smaller percentages using platforms such as discussion forums and mental health apps [8]. Our study also found much lower rates of using these online tools to actively seek professional mental health help, in comparison to mental health information [8].

We found that students who had higher mental health literacy were more likely to use online tools to seek information but not professional help. This might suggest that students with more knowledge on mental illness, such as the ability to recognize and label their symptoms or attribute them to mental health rather than other causes, may be better equipped to seek help online. For example, students with higher mental health literacy may know what specific key terms to use to search for relevant information. In addition, those knowledgeable about coping skills (eg, relaxation skills for anxiety) may be more likely to use online tools to find more information on how to practice these skills. Although higher mental health literacy is associated with increased intentions to seek help, it might not translate to actual service use [45]. Considering the cross-sectional nature of our study, it is also possible that greater use of online tools may improve the mental health literacy of students by increasing exposure to online information. The lack of association between mental health literacy and the use of online tools to connect with professional services may be due to individuals interpreting online psychoeducational resources and coping strategies as sufficient, leading them to conclude that their symptoms did not require professional care [44,46]. This may also reflect students' perception of digital mental health tools as either a more accessible alternative to traditional services or as a complement to conventional care [7].

Our study found that mental health symptoms, including depression, anxiety, and trauma symptoms, did not predict online tool use. Past research has found mixed findings on this association, with 1 study finding anxiety symptoms to predict app use, while depression did not [12]. It is possible that we did not find this association because participants were asked about overall online tool use and not specific online tools, such as apps. Moreover, students may be using different types of online tools available (eg, websites, apps, and smart devices) depending on different symptom-related concerns. For example, some individuals may prefer using websites rather than an app for information seeking and psychoeducation on symptoms, while others may prefer self-help apps that provide tips to regulate symptoms and promote relaxation [12]. Although mental health symptoms were not statistically significant predictors of online tool use, given that some associations trended toward significance (depression: P=.05, PTSD: P=.07), it may be possible that the study was too underpowered to detect a statistically significant relationship between mental health symptoms and online tool use.

Furthermore, our sample reported mild depression and anxiety on average, with very few PTSD symptoms. On the basis of the

XSL•FO
RenderX

qualitative findings, a common reason for not using online tools was a lack of perceived need. Therefore, it is possible that the absence of an association between mental health symptoms and online tool reflects students' belief that their symptoms did not warrant significant attention. Other evidence indicates that young people with more severe symptoms, especially those with higher depression severity, may avoid traditional care, preferring to use anonymous electronic sources (ie, internet forums, support groups, and chat rooms) and digital tools to better understand and cope with their symptoms on their own [47]. Considering students who have more severe symptoms of depression may not seek care compared to those with more moderate symptoms due to coping with specific depression symptomology (ie, lack of motivation, low energy, increased fatigue, and reduced problem-solving abilities), such self-reliance may further reduce their perceived need for professional mental health services [48,49]. Future research should examine what factors may play a role in students' perceived need of mental health care.

Participants in our study reported that online tools to seek information were accessible, easily providing fast information and answers to questions quickly, which is consistent with past research [21,50,51]. While previous research had mixed results about existing disparities in online tool use based on race or ethnicity, non–US born status, and gender, our study found no differences in online tool use based on these factors [7,16,23]. However, it is important to note that our sample was highly educated, had lower stigma, and had more positive views and knowledge about mental health than the general population. These factors likely facilitated the use of online tools for mental health among our participants. Although we did not find associations between stigma, attitudes toward mental health treatment, or self-efficacy and online tool use among our sample, this may be a result of the privacy provided by the digital nature of these tools. While support from family and friends may be incorporated in digital tools to promote the management and treatment of mental health disorders, for those dealing with more social or family-based stigma, the anonymity and privacy of online tools may provide more autonomy to decide who they would like to involve in this process, if any [7,52,53]. For those who do not feel comfortable using online tools, it may be due to harboring certain attitudes and beliefs about digital tools, such as finding them untrustworthy or unreliable, that resulted in their unwillingness to use them. As for self-efficacy, digital self-efficacy, such as students' confidence in navigating digital spaces, may have a greater impact on their willingness to use online tools rather than their confidence in seeking mental health support itself [18]. Therefore, stigma that exists around the confidentiality of online data, negative attitudes toward the use of digital platforms, and students' confidence in their ability to navigate them effectively may play a role in online help seeking. This highlights the importance of research examining the different types of attitudes and beliefs that exist about digital mental health tools, specifically.

Recommendations to Improve the Implementation of Online Tools

Due to possible differences in barriers faced by students across a variety of school settings, we recommend that researchers consider the needs and priorities of their target population to increase the relevance and fit of the digital tools. It is important to recognize that students from other educational settings, aside from 4-year institutions, may experience barriers to using online tools for mental health. For example, research has found that community college students face many barriers to using digital tools (eg, mental health apps), including financial challenges, privacy concerns, self-reliance, skepticism about the seriousness of symptoms, and concerns about others' perceptions of them [51]. This suggests that accessibility and stigma barriers may exert a more pronounced impact on community college students than on their university counterparts [51]. Thus, it is important to consider the target population and setting when implementing digital tools. To develop digital tools that are acceptable to college students from a variety of educational backgrounds, we recommend that formative research be conducted with the population of interest, such as identifying primary mental health concerns and including user content and interface suggestions (ie, crisis text line, telehealth, websites for connecting to services, SMS text messaging with counselors, personalization, anonymity, privacy, security, and peer engagement) that may promote service use before the implementation and dissemination of a digital tool [8,54,55].

Our qualitative results demonstrated that many students are unaware of how digital tools could be used to connect with care. There is a need to improve the digital tool's user interface because many online tools do not have the ability to connect users directly to a coach or therapist [44,56]. Given that students view the ability to connect to care as important, we recommend that online tools help users assess symptoms and include a direct connection to professional services, as it may increase use of mental health services among college students with unmet mental health needs [12]. For example, rather than having students navigate help seeking on large search engines, which may be a dissatisfying process, digital tools that could triage help should be promoted and researched, as highlighted by our participants' requests for features that personalize the help-seeking process. Young people have reported higher satisfaction when online tools such as Link, a navigation website that connects youth to treatment based on reported mental health symptom type and severity, were tailored specifically to their needs (ie, immediate return of personalized options for help seeking, stories from individuals dealing with similar symptoms, and self-care tips) [57]. Digital interventions, such as the Screening and Treatment for Anxiety and Depression program that screens college students for mental health symptoms and reliably connects students to evidence-based care (eg, self-guided wellness program, sessions with peer coaches, or individual therapy) based on the severity of depressive symptoms, might act as a gateway to connect students to care [49]. This type of online tool could circumvent barriers such as misjudging perceived need and self-reliance, facilitating a more efficient pathway for students to access the necessary mental health support [49].

Qualitative findings also indicated that students used online tools for psychoeducation and general information about mental health and found them helpful. Although there are websites that provide general psychoeducation, such online tools may lack

XSL•FO
RenderX

detailed information about evidence-based treatments and guidance on connecting with professionals [8,12,58]. Without adequate psychoeducation on available treatments and how to access them, students may struggle to take the next step in pursuing professional services, ultimately leaving their mental health needs unmet. College students have reported that online platforms, such as clinic websites, should provide more comprehensive information, including clear explanations on how to get started with services, detailed descriptions of services available for specific mental health disorders, and mental health professionals' background information, which in turn can increase a students' willingness to seek mental health care [59]. Thus, we recommend that digital tools provide explicit psychoeducation on how to receive evidence-based treatments, as that may increase students' likelihood of using online tools to connect to mental health services. Another recommendation for content creators who design mental health websites is to include resources for crisis lines, which was suggested by participants who did not use online tools to connect to care as a possible motivating factor for future use. Crisis lines have been shown to connect people with immediate help and promote engagement with long-term mental health services (ie, referrals to mental health professionals) [60,61].

Recommendations to Increase the Dissemination of Online Tools

Our qualitative results demonstrated that a major reason why college students did not use online tools to support their mental health was because they were unaware of what digital tools were available. Indeed, previous studies have found that many participants are unaware of formal online resources (ie, charities and health services) available to support them [8]. Considering we found mental health literacy to predict online tool use, video interventions and digital media campaigns targeting young people with low mental health literacy should also include information on available online tools for mental health support. These interventions can enhance mental health literacy, reduce stigma, and promote help seeking by increasing awareness of various online tools for accessing care [62].

Referrals from friends, family, or health care professionals could help raise awareness about available online tools. As support from loved ones often encourages help seeking, recommendations from these sources signal to participants that the tools are more reliable and trustworthy [7,44,55]. Although numerous digital tools have become more readily available, proper evaluation and validation of these tools and their implementation is limited, resulting in students needing to deduce whether a tool is trustworthy or not [63]. While young people may be motivated to search the internet for mental health help, their ability to deduce the quality and reliability of the mental health information may prove to be more challenging with limited mental health literacy, further hindering their willingness to use online tools. It is recommended that mental health professionals receive training on available digital tools so they can confidently recommend them to patients, who may then share these tools with friends and family [43,63,64]. Therefore, referrals and recommendations by trusted individuals can spread awareness of credible digital tools, and in turn, reach underserved populations who may have negative perceptions of the internet and digital mediums or find the digital tools unreliable due to privacy concerns [13,14,45,55,65].

Limitations and Future Directions

As data for this study were collected in 2021, findings should be interpreted within the unique context of the COVID-19 pandemic. Although our study found that a third of participants (37/123, 30.1%) used online tools to connect with mental health professionals or services, which was consistent with online help-seeking rates before the COVID-19 pandemic [21], general use of online tools to seek mental health information likely increased during the pandemic [8,25,27]. At that time, lockdowns and social distancing measures significantly limited in-person mental health services, which may have influenced students' decisions to seek information and support through online tools. These conditions may have contributed to both increased exposure to and reliance on online mental health resources. Consequently, patterns of online help seeking captured in this study may reflect a period of heightened necessity rather than sustained preference. Future research is needed to examine how students' attitudes and behaviors toward online mental health tools have evolved after the pandemic, when the use of digital resources is increasingly optional rather than mandated.

We used standardized procedures to code the qualitative responses; however, some open-ended responses were too vague to be coded (eg, "it was something" was provided as a response to "Why was this online tool helpful?"). In addition, our small sample size limited our ability to quantitatively examine patterns in types of online tools used, perceived helpfulness, and motivating factors across different demographic groups. To better tailor these tools to populations considered marginalized, we believe that future research should aim to identify if there are differences in these outcomes based on demographic characteristics. Related to our sample size, post hoc power analyses suggest the study may have been underpowered to detect small to moderate associations between various key predictors and outcomes. For example, analyses examining the relationship between mental health symptoms (depression, anxiety, and trauma symptoms) and online tool use were underpowered (eg, depression: power=0.69, anxiety: power=0.10, and trauma: power=0.63), potentially limiting our ability to detect significant effects. Our study sample also had a positively skewed distribution of mental health symptoms, with most participants reporting mild symptoms, as well as a negatively skewed distribution for self-efficacy and attitudes, where most participants expressed high self-efficacy and more positive attitudes toward mental health. Future research with larger, more diverse samples and greater variability in responses is needed to clarify whether these null findings reflect a true absence of association or limitations in statistical power. Due to the cross-sectional design of our study, the directionality of associations between our key predictor variables and online tool use also could not be established. Future longitudinal or experimental research is needed to clarify causal pathways and better understand the temporal sequencing of these relationships.

Furthermore, this study did not measure the following factors that are associated with the use of online tools for mental health:

XSL•FO
RenderX

access to reliable technology, digital literacy (ie, ability to locate, evaluate, and communicate information using digital mediums), internet self-efficacy (ie, confidence in one's ability to master new technology), attitudes toward digital mental health (ie, positive and negative views relating to trustworthiness, credibility, and efficacy), and preferences for in-person help seeking [11,12,15,65,66]. Future research should examine how these factors play a role in the use of online tools for mental health purposes, as highlighted by our qualitative responses. In addition, participants in our sample were recruited from students enrolled in psychology courses and may be more knowledgeable about mental health and have less stigma than students with other majors [67,68]. These findings are limited to the study sample, which included university students attending a 4-year institution, and may not be applicable to students from other educational backgrounds.

Conclusions

There is a need to bridge the gap between effective online tools for mental health and users. Research exploring digital mental health tool design, acceptability, and feasibility can promote access to psychological information and professional services among college students [10]. Our study shows that university students find online tools to be an acceptable and feasible method to seek mental health information; however, more research is needed on how to optimize online tools to help students connect to mental health care. To optimize the utility of online tools, it is crucial to conduct research that will yield rich user insight to identify and implement user-suggested components, such as psychoeducation about disorders, treatments, and services; personalized content; interactive features; and connections to professional care [9,69-71].

Conflicts of Interest

None declared.

References

1. Healthy minds study data report, (2023-2024). Healthy Minds Network. URL: https://healthymindsnetwork.org/wp-content/uploads/2024/09/HMS_national_report_090924.pdf [accessed 2025-05-29]
2. Lipson SK, Zhou S, Abelson S, Heinze J, Jirsa M, Morigney J, et al. Trends in college student mental health and help-seeking by race/ethnicity: findings from the national healthy minds study, 2013-2021. J Affect Disord 2022 Jul 01;306:138-147 [FREE Full text] [doi: 10.1016/j.jad.2022.03.038] [Medline: 35307411]
3. Eisenberg D, Golberstein E, Hunt JB. Mental health and academic success in college. B E J Econom Anal Policy 2009;9(1):1935-1682. [doi: 10.2202/1935-1682.2191]
4. Sontag-Padilla L, Woodbridge MW, Mendelsohn J, D'Amico EJ, Osilla KC, Jaycox LH, et al. Factors affecting mental health service utilization among California public college and university students. Psychiatr Serv 2016 Aug 01;67(8):890-897. [doi: 10.1176/appi.ps.201500307] [Medline: 27032662]
5. Key substance use and mental health indicators in the United States: results from the 2021 national survey on drug use and health. Substance Abuse and Mental Health Services Administration. URL: https://www.samhsa.gov/data/report/2021-nsduh-annual-national-report [accessed 2025-05-29]
6. Rickwood D, Thomas K. Conceptual measurement framework for help-seeking for mental health problems. Psychol Res Behav Manag 2012;5:173-183 [FREE Full text] [doi: 10.2147/PRBM.S38707] [Medline: 23248576]
7. Montagni I, Tzourio C, Cousin T, Sagara JA, Bada-Alonzi J, Horgan A. Mental health-related digital use by university students: a systematic review. Telemed J E Health 2020 Mar 01;26(2):131-146. [doi: 10.1089/tmj.2018.0316] [Medline: 30888256]
8. Pretorius C, Coyle D. Young people's use of digital tools to support their mental health during COVID-19 restrictions. Front Digit Health 2021;3:763876 [FREE Full text] [doi: 10.3389/fdgth.2021.763876] [Medline: 34927133]
9. White house report on mental health research priorities. The White House. URL: https://www.whitehouse.gov/ostp/news-updates/2023/02/07/white-house-report-on-mental-health-research-priorities/ [accessed 2025-05-29]
10. Ebert DD, Mortier P, Kaehlke F, Bruffaerts R, Baumeister H, Auerbach RP, WHO World Mental Health-International College Student Initiative collaborators. Barriers of mental health treatment utilization among first-year college students: first cross-national results from the WHO World Mental Health International College Student Initiative. Int J Methods Psychiatr Res 2019 Jul 09;28(2):e1782 [FREE Full text] [doi: 10.1002/mpr.1782] [Medline: 31069905]
11. Borghouts J, Eikey E, Mark G, De Leon C, Schueller SM, Schneider M, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. J Med Internet Res 2021 Mar 24;23(3):e24387 [FREE Full text] [doi: 10.2196/24387] [Medline: 33759801]
12. Bautista J, Schueller SM. Understanding the adoption and use of digital mental health apps among college students: secondary analysis of a national survey. JMIR Ment Health 2023 Mar 22;10:e43942 [FREE Full text] [doi: 10.2196/43942] [Medline: 36947115]
13. Salamanca-Sanabria A, Jabir AI, Lin X, Alattas A, Kocaballi AB, Lee J, et al. Exploring the perceptions of mHealth interventions for the prevention of common mental disorders in university students in Singapore: qualitative study. J Med Internet Res 2023 Mar 20;25:e44542 [FREE Full text] [doi: 10.2196/44542] [Medline: 36939808]

XSL•FO
RenderX

14. Pretorius C, Chambers D, Coyle D. Young people's online help-seeking and mental health difficulties: systematic narrative review. J Med Internet Res 2019 Dec 19;21(11):e13873 [FREE Full text] [doi: 10.2196/13873] [Medline: 31742562]
15. Levin ME, Krafft J, Levin C. Does self-help increase rates of help seeking for student mental health problems by minimizing stigma as a barrier? J Am Coll Health 2018 Apr 19;66(4):302-309. [doi: 10.1080/07448481.2018.1440580] [Medline: 29447600]
16. Ahuvia IL, Sung JY, Dobias ML, Nelson BD, Richmond LL, London B, et al. College student interest in teletherapy and self-guided mental health supports during the COVID-19 pandemic. J Am Coll Health 2024 May;72(3):940-946. [doi: 10.1080/07448481.2022.2062245] [Medline: 35427460]
17. ten Have M, de Graaf R, Ormel J, Vilagut G, Kovess V, Alonso J, ESEMeD/MHEDEA 2000 Investigators. Are attitudes towards mental health help-seeking associated with service use? Results from the European Study of Epidemiology of Mental Disorders. Soc Psychiatry Psychiatr Epidemiol 2010 Mar;45(2):153-163 [FREE Full text] [doi: 10.1007/s00127-009-0050-4] [Medline: 19381427]
18. Tetri B, Juujärvi S. Self-Efficacy, internet self-efficacy, and proxy efficacy as predictors of the use of digital social and health care services among mental health service users in Finland: a cross-sectional study. Psychol Res Behav Manag 2022;15:291-303 [FREE Full text] [doi: 10.2147/PRBM.S340867] [Medline: 35210878]
19. Yamawaki N, Green J. Stigma and self-efficacy as predictors of intention to seek help among American and Japanese college students. Int J Psychol Couns 2020;12(3):63-72. [doi: 10.5897/IJPC2020.0614]
20. Kodish T, Lau AS, Gong-Guy E, Congdon E, Arnaudova I, Schmidt M, et al. Enhancing racial/ethnic equity in college student mental health through innovative screening and treatment. Adm Policy Ment Health 2022 Mar 09;49(2):267-282 [FREE Full text] [doi: 10.1007/s10488-021-01163-1] [Medline: 34505211]
21. Kauer SD, Mangan C, Sanci L. Do online mental health services improve help-seeking for young people? A systematic review. J Med Internet Res 2014 Mar 04;16(3):e66. [doi: 10.2196/JMIR.3103]
22. Vera Cruz G, Aboujaoude E, Khan R, Rochat L, Ben Brahim F, Courtois R, et al. Smartphone apps for mental health and wellbeing: a usage survey and machine learning analysis of psychological and behavioral predictors. Digit Health 2023 Jan 22;9:20552076231152164 [FREE Full text] [doi: 10.1177/20552076231152164] [Medline: 36714544]
23. Mohammadifirouzeh M, Oh KM, Basnyat I, Gimm G. Factors associated with professional mental help-seeking among U.S. immigrants: a systematic review. J Immigr Minor Health 2023 Oct 31;25(5):1118-1136 [FREE Full text] [doi: 10.1007/s10903-023-01475-4] [Medline: 37000385]
24. Johnson JA, Sanghvi P, Mehrotra S. Technology-based interventions to improve help-seeking for mental health concerns: a systematic review. Indian J Psychol Med 2022 Jul 21;44(4):332-340 [FREE Full text] [doi: 10.1177/02537176211034578] [Medline: 35949632]
25. Wies B, Landers C, Ienca M. Digital mental health for young people: a scoping review of ethical promises and challenges. Front Digit Health 2021 Sep 6;3:697072 [FREE Full text] [doi: 10.3389/fdgth.2021.697072] [Medline: 34713173]
26. Clement S, Schauman O, Graham T, Maggioni F, Evans-Lacko S, Bezborodovs N, et al. What is the impact of mental health-related stigma on help-seeking? A systematic review of quantitative and qualitative studies. Psychol Med 2015 Jan;45(1):11-27. [doi: 10.1017/S0033291714000129] [Medline: 24569086]
27. Bergquist S, Otten T, Sarich N. COVID-19 pandemic in the United States. Health Policy Technol 2020 Dec;9(4):623-638 [FREE Full text] [doi: 10.1016/j.hlpt.2020.08.007] [Medline: 32874854]
28. Facts and figures: an overview of the data that makes UCLA unique. University of California. URL: https://www.ucla.edu/about/facts-and-figures [accessed 2025-05-29]
29. CAPS data overview UCLA counseling and psychological services. University of California. URL: https://counseling.ucla.edu/about-us/caps-data-overview [accessed 2025-05-29]
30. Maghbouleh N, Schachter A, Flores RD. Middle Eastern and North African Americans may not be perceived, nor perceive themselves, to be White. Proc Natl Acad Sci U S A 2022 Mar 15;119(7):e2117940119 [FREE Full text] [doi: 10.1073/pnas.2117940119] [Medline: 35131945]
31. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Oct;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
32. Taber KS. The use of Cronbach's alpha when developing and reporting research instruments in science education. Res Sci Educ 2017 Jun 7;48(6):1273-1296. [doi: 10.1007/S11165-016-9602-2]
33. Johnson SU, Ulvenes PG, Øktedalen T, Hoffart A. Psychometric properties of the general anxiety disorder 7-item (GAD-7) scale in a heterogeneous psychiatric sample. Front Psychol 2019 Aug 6;10:1713 [FREE Full text] [doi: 10.3389/fpsyg.2019.01713] [Medline: 31447721]
34. Blevins CA, Weathers FW, Davis MT, Witte TK, Domino JL. The posttraumatic stress disorder checklist for DSM-5 (PCL-5): development and initial psychometric evaluation. J Trauma Stress 2015 Dec;28(6):489-498. [doi: 10.1002/jts.22059] [Medline: 26606250]
35. O'Connor M, Casey L. The Mental Health Literacy Scale (MHLS): a new scale-based measure of mental health literacy. Psychiatry Res 2015 Oct 30;229(1-2):511-516. [doi: 10.1016/j.psychres.2015.05.064] [Medline: 26228163]

XSL•FO
RenderX

36. Boyd JE, Otilingam PG, Deforge BR. Brief version of the Internalized Stigma of Mental Illness (ISMI) scale: psychometric properties and relationship to depression, self esteem, recovery orientation, empowerment, and perceived devaluation and discrimination. Psychiatr Rehabil J 2014 Mar;37(1):17-23. [doi: 10.1037/prj0000035] [Medline: 24660946]
37. Conner KO, Koeske G, Brown C. Racial differences in attitudes toward professional mental health treatment: the mediating effect of stigma. J Gerontol Soc Work 2009 Oct 24;52(7):695-712. [doi: 10.1080/01634370902914372] [Medline: 19787528]
38. Moore CD, Schofield C, van Rooyen DR, Andersson LM. Development and preliminary validation of a scale to measure self-efficacy in seeking mental health care (SE-SMHC). Springerplus 2015 Jul 11;4(1):339 [FREE Full text] [doi: 10.1186/s40064-015-1109-1] [Medline: 26185741]
39. California health interview survey. UCLA Center for Health Policy Research. URL: https://healthpolicy.ucla.edu/our-work/california-health-interview-survey-chis [accessed 2025-05-29]
40. Palinkas LA, Aarons GA, Horwitz S, Chamberlain P, Hurlburt M, Landsverk J. Mixed method designs in implementation research. Adm Policy Ment Health 2011 Jan 22;38(1):44-53 [FREE Full text] [doi: 10.1007/s10488-010-0314-z] [Medline: 20967495]
41. The jamovi project. Version 2.3. jamovi. URL: https://www.jamovi.org/ [accessed 2025-05-29]
42. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Dec;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
43. Cascio MA, Lee E, Vaudrin N, Freedman DA. A team-based approach to open coding: considerations for creating intercoder consensus. Field Methods 2019 Mar 27;31(2):116-130. [doi: 10.1177/1525822X19838237]
44. Liu M, Schueller SM. Moving evidence-based mental health interventions into practice: implementation of digital mental health interventions. Curr Treat Options Psych 2023 Oct 03;10(4):333-345. [doi: 10.1007/S40501-023-00298-2]
45. Webb M, Burns J, Collin P. Providing online support for young people with mental health difficulties: challenges and opportunities explored. Early Interv Psychiatry 2008 May;2(2):108-113. [doi: 10.1111/j.1751-7893.2008.00066.x] [Medline: 21352141]
46. Oexle N, Ajdacic-Gross V, Müller M, Rodgers S, Rössler W, Rüsch N. Predicting perceived need for mental health care in a community sample: an application of the self-regulatory model. Soc Psychiatry Psychiatr Epidemiol 2015 Oct;50(10):1593-1600. [doi: 10.1007/s00127-015-1085-3] [Medline: 26084865]
47. Czyz EK, Horwitz AG, Eisenberg D, Kramer A, King CA. Self-reported barriers to professional help seeking among college students at elevated risk for suicide. J Am Coll Health 2013 Oct;61(7):398-406 [FREE Full text] [doi: 10.1080/07448481.2013.820731] [Medline: 24010494]
48. Schraufnagel TJ, Wagner AW, Miranda J, Roy-Byrne PP. Treating minority patients with depression and anxiety: what does the evidence tell us? Gen Hosp Psychiatry 2006 Jan;28(1):27-36. [doi: 10.1016/j.genhosppsych.2005.07.002] [Medline: 16377362]
49. Wolitzky-Taylor K, LeBeau R, Arnaudova I, Barnes-Horowitz N, Gong-Guy E, Fears S, et al. A novel and integrated digitally supported system of care for depression and anxiety: findings from an open trial. JMIR Ment Health 2023 Jul 24;10:e46200 [FREE Full text] [doi: 10.2196/46200] [Medline: 37486735]
50. Li J, Theng Y, Foo S. Predictors of online health information seeking behavior: changes between 2002 and 2012. Health Informatics J 2016 Dec;22(4):804-814 [FREE Full text] [doi: 10.1177/1460458215595851] [Medline: 26261218]
51. Borghouts J, Eikey EV, Mark G, De Leon C, Schueller SM, Schneider M, et al. Understanding mental health app use among community college students: web-based survey study. J Med Internet Res 2021 Oct 14;23(9):e27745 [FREE Full text] [doi: 10.2196/27745] [Medline: 34519668]
52. Steinkamp JM, Goldblatt N, Borodovsky JT, LaVertu A, Kronish IM, Marsch LA, et al. Technological interventions for medication adherence in adult mental health and substance use disorders: a systematic review. JMIR Ment Health 2019 Mar 12;6(3):e12493 [FREE Full text] [doi: 10.2196/12493] [Medline: 30860493]
53. Cho S, Bastien L, Petrovic J, Böke BN, Heath NL. The role of mental health stigma in university students' satisfaction with web-based stress management resources: intervention study. JMIR Form Res 2024 May 04;8:e50018 [FREE Full text] [doi: 10.2196/50018] [Medline: 38573758]
54. Oti O, Pitt I. Online mental health interventions designed for students in higher education: a user-centered perspective. Internet Interv 2021 Dec;26:100468 [FREE Full text] [doi: 10.1016/j.invent.2021.100468] [Medline: 34703772]
55. Chan JK, Farrer LM, Gulliver A, Bennett K, Griffiths KM. University students' views on the perceived benefits and drawbacks of seeking help for mental health problems on the internet: a qualitative study. JMIR Hum Factors 2016 Jan 19;3(1):e3 [FREE Full text] [doi: 10.2196/humanfactors.4765] [Medline: 27026140]
56. Lagan S, D'Mello R, Vaidyam A, Bilden R, Torous J. Assessing mental health apps marketplaces with objective metrics from 29,190 data points from 278 apps. Acta Psychiatr Scand 2021 Aug 29;144(2):201-210. [doi: 10.1111/acps.13306] [Medline: 33835483]
57. Sanci L, Kauer S, Thuraisingam S, Davidson S, Duncan A, Chondros P, et al. Effectiveness of a mental health service navigation website (link) for young adults: randomized controlled trial. JMIR Ment Health 2019 Oct 17;6(10):e13189 [FREE Full text] [doi: 10.2196/13189] [Medline: 31625945]

XSL•FO
RenderX

58. King S, Ritchie KC, McGonnell M, Doe C, Corkum PV, Côté E, et al. Paging Dr. Google: availability and reliability of online evidence-based treatment information about ADHD. Evid Based Pract Child Adolesc Ment Health 2021 Mar 24;6(2):277-289. [doi: 10.1080/23794925.2021.1901632]
59. Aguilar Silvan Y, Hamza S, Fardeheb S, Bird C, Ng LC. Marketing mental health services: a mixed-methods analysis of racially and ethnically diverse college students' engagement with and perspectives on U.S. university mental health clinics' websites. BMC Health Serv Res 2024 Oct 02;24(1):1163 [FREE Full text] [doi: 10.1186/s12913-024-11652-2] [Medline: 39354461]
60. Hoffberg AS, Stearns-Yoder KA, Brenner LA. The effectiveness of crisis line services: a systematic review. Front Public Health 2019 Jan 17;7:399 [FREE Full text] [doi: 10.3389/fpubh.2019.00399] [Medline: 32010655]
61. Knapp AA, Cohen K, Nicholas J, Mohr DC, Carlo AD, Skerl JJ, et al. Integration of digital tools into community mental health care settings that serve young people: focus group study. JMIR Ment Health 2021 Aug 19;8(8):e27379 [FREE Full text] [doi: 10.2196/27379] [Medline: 34420928]
62. Janoušková M, Tušková E, Weissová A, Trančík P, Pasz J, Evans-Lacko S, et al. Can video interventions be used to effectively destigmatize mental illness among young people? A systematic review. Eur Psychiatry 2017 Mar;41:1-9 [FREE Full text] [doi: 10.1016/j.eurpsy.2016.09.008] [Medline: 28049074]
63. Smith KA, Blease C, Faurholt-Jepsen M, Firth J, Van Daele T, Moreno C, et al. Digital mental health: challenges and next steps. BMJ Ment Health 2023 Mar 17;26(1):e300670 [FREE Full text] [doi: 10.1136/bmjment-2023-300670] [Medline: 37197797]
64. Nogueira-Leite D, Diniz JM, Cruz-Correia R. Mental health professionals' attitudes toward digital mental health apps and implications for adoption in Portugal: mixed methods study. JMIR Hum Factors 2023 Jul 02;10:e45949 [FREE Full text] [doi: 10.2196/45949] [Medline: 37266977]
65. Bird MD, Chow GM, Yang Y. College students' attitudes, stigma, and intentions toward seeking online and face-to-face counseling. J Clin Psychol 2020 Oct 09;76(9):1775-1790. [doi: 10.1002/jclp.22956] [Medline: 32271953]
66. Tartaglia J, Jaghab B, Ismail M, Hänsel K, Meter AV, Kirschenbaum M, et al. Assessing health technology literacy and attitudes of patients in an urban outpatient psychiatry clinic: cross-sectional survey study. JMIR Ment Health 2024 Dec 30;11:e63034 [FREE Full text] [doi: 10.2196/63034] [Medline: 39753220]
67. Ruiz JC, Fuentes-Durá I, López-Gilberte M, Dasí C, Pardo-García C, Fuentes-Durán MC, et al. Public stigma profile toward mental disorders across different university degrees in the University of Valencia (Spain). Front Psychiatry 2022 Aug 12;13:951894 [FREE Full text] [doi: 10.3389/fpsyt.2022.951894] [Medline: 36032229]
68. Miles R, Rabin L, Krishnan A, Grandoit E, Kloskowski K. Mental health literacy in a diverse sample of undergraduate students: demographic, psychological, and academic correlates. BMC Public Health 2020 Nov 13;20(1):1699-1613 [FREE Full text] [doi: 10.1186/s12889-020-09696-0] [Medline: 33187487]
69. Regnault A, Willgoss T, Barbic S, International Society for Quality of Life Research (ISOQOL) Mixed Methods Special Interest Group (SIG). Towards the use of mixed methods inquiry as best practice in health outcomes research. J Patient Rep Outcomes 2017 Apr 11;2(1):19 [FREE Full text] [doi: 10.1186/s41687-018-0043-8] [Medline: 29757311]
70. Williams L, Pinto M, Abrahim H, Bhattacharya A, Harrison T, Burton C, et al. Designing an evidence-based mental health intervention alongside university students. In: Proceedings of the HCI 2023, Held as Part of the 25th HCI International Conference on Human-Computer Interaction: Thematic Area. 2023 Presented at: HCI/HCII '23; July 23-28, 2023; Copenhagen, Denmark p. 180-202 URL: https://doi.org/10.1007/978-3-031-35572-1_13 [doi: 10.1007/978-3-031-35572-1_13]
71. Kornfield R, Meyerhoff J, Studd H, Bhattacharjee A, Williams JJ, Reddy M, et al. Meeting users where they are: user-centered design of an automated text messaging tool to support the mental health of young adults. Proc SIGCHI Conf Hum Factor Comput Syst 2022 May;2022:329 [FREE Full text] [doi: 10.1145/3491102.3502046] [Medline: 35574512]

Abbreviations

PTSD: posttraumatic stress disorder
UCLA: University of California, Los Angeles

Edited by A Kushniruk; submitted 16.05.24; peer-reviewed by K Schroeder, V Harrison, N Wong; comments to author 22.12.24; revised version received 11.04.25; accepted 13.05.25; published 13.06.25.

Please cite as:
Hamza SZ, Aguilar Silvan Y, Ng LC
Use of Online Tools for Mental Health Among Racially and Ethnically Diverse College Students: Mixed Methods Study
JMIR Hum Factors 2025;12:e60628
URL: https://humanfactors.jmir.org/2025/1/e60628
doi:10.2196/60628
PMID:

XSL•FO
RenderX

©Sarah Z Hamza, Yesenia Aguilar Silvan, Lauren C Ng. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 13.06.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Investigating How Clinicians Form Trust in an AI-Based Mental Health Model: Qualitative Case Study

Anthony Kelly[1,2], PhD; Niharika Bhardwaj[1], BE; Trine Theresa Holmberg Sainte-Marie[3], MSc; Pepijn Van de Ven[1,2], PhD; Ruth Melia[2,4], DClinPsych; John Eustis Williams[5], PhD; Kim Mathiasen[3,6], PhD; Amalie Søgaard Nielsen[6], PhD

[1]Department of Electronic and Computer Engineering, University of Limerick, Casteltroy, Limerick, Ireland

[2]Health Research Institute, University of Limerick, Castletroy, Limerick, Ireland

[3]Department of Psychology and Behavioural Sciences, Aarhus University, Aarhus, Denmark

[4]Department of Psychology, University of Limerick, Limerick, Ireland

[5]Center for Digital Psykiatri, Forskningsenhed, Denmark

[6]Research Unit of Digital Psychiatry, Department of Clinical Research, University of Southern Denmark, Odense, Denmark

Corresponding Author:
Anthony Kelly, PhD
Department of Electronic and Computer Engineering, University of Limerick, Casteltroy, Limerick, Ireland

Abstract

Background: Trust in artificial intelligence (AI) remains a critical barrier to the adoption of AI in mental health care. This study explores the formation of trust in an AI mental health model and its human-computer interface among clinicians at a web-based mental health clinic in the Region of Southern Denmark with national coverage.

Objective: This study aims to explore clinicians' perspectives on how trust is built in the context of an AI-supported mental health screening model and to identify the factors that influence this process.

Methods: This was a qualitative case study using semistructured interviews with clinicians involved in the pilot of a mental health AI model. Thematic analysis was used to identify key factors contributing to trust formation.

Results: Clinicians' initial attitudes toward AI were shaped by prior positive experiences with AI and their perception of AI's potential to reduce cognitive load in routine screening. Trust development followed a sequential pattern resembling a "trust journey": (1) sense-making, (2) risk appraisal, and (3) conditional decision to rely. Trust formation was supported by the explainability of the model, particularly through (1) visualization of confidence and uncertainty through violin plots, aligning with the clinicians' expectations of decision ambiguity; (2) feature attribution for and against predictions, which mirrored clinical reasoning; and (3) use of pseudo-sumscores in the AI model, which increased interpretability by presenting explanations in familiar clinical formats. Trust was contextually bounded to low-risk clinical scenarios, such as preinterview patient screening, and contingent on safety protocols (eg, suicide risk flagging). The use of both structured and unstructured patient data was seen as key to expanding trust into more complex clinical contexts. Participants also expressed a need for ongoing evaluation data to reinforce and maintain trust.

Conclusions: Clinicians' trust in AI tools is contextually and sequentially constructed, influenced by both model performance and alignment with clinical reasoning. Interpretability features were essential in establishing intrinsic trust, particularly when presented in ways that resonate with clinical norms. For broader acceptance and responsible deployment, trust must be supported by rigorous evaluation data and the inclusion of clinically relevant data types in model design.

(JMIR Hum Factors 2025;12:e79658) doi:10.2196/79658

KEYWORDS

artificial intelligence; AI; explainable AI; XAI; trust; machine learning; ML; mental health; decision support system; intrinsic trust

Introduction

Background

Effective and accessible mental health treatment is crucial in addressing the growing global burden of mental health disorders. Untreated mental health disorders not only impact individual well-being but also result in substantial economic and societal strain through lost productivity, increased health care costs, and reduced quality of life. With the economic burden estimated to be US $5 trillion in 2019 [1], the need for timely and appropriate intervention has never been more urgent.

XSL•FO
RenderX

Like many aspects of society, the treatment of mental health disorders stands to be transformed by artificial intelligence (AI). By recognizing complex patterns and analyzing vast amounts of data, AI has the potential to enhance early detection, personalize treatment, and deliver digital therapy at scale [2,3].

AI adoption depends on factors related to trust, such as system reliability, robust performance, and perceived competence, which is especially important in medical [4] and psychological decision making. In mental health, lack of trust is seen as a key barrier to adoption of AI-based decision support systems [5,6]. Explainability is an important factor in fostering trust in AI systems. When users understand why an AI makes certain decisions (explainability) and can clearly comprehend the reasoning behind those explanations (interpretability), they are more likely to adopt AI.

Trust has also been identified as one of the most critical elements of AI adoption in clinical settings. In a scoping review, Hassan, Kushniruk, and Boryck [7] highlight that a lack of trust remains a major barrier to implementation in health care. Transparency, system reliability, and alignment with clinical reasoning are identified as key facilitators. This reinforces the idea that explainability needs to be considered from the perspective of those expected to use the model in practice. Other stakeholders, such as regulators, also play a role in ensuring that digital mental health tools are safe, effective, and trusted through robust regulation and evaluation, which builds public trust [8].

When an AI model predicts a classification with near 100% probability, this implies that the model is near 100% confident that its prediction is correct. As such, the prediction probability is known as the confidence of the prediction [9]. Uncertainty occurs when an AI model is unsure about its confidence prediction, ie, predictive uncertainty [10]. This may occur due to aleatoric uncertainty (uncertainty in the data), or epistemic uncertainty (uncertainty in the model or a shift in the statistical distribution of the data) [11]. The quantification of predictive uncertainty depends on the type of model being evaluated. In neural network models, a single prediction vector is typically output for each data point. The uncertainty is quantified as the deviation of the true-class probability from 100% confidence, measured as the negative log likelihood, or the deviation of all probabilities in the prediction vector, from the true-class probability, measured as the static calibration error or the Brier score. When predictive uncertainty distributions are available, as in Bayesian, Ensemble, or Monte Carlo dropout models, uncertainty manifests itself as the shape of the probability distribution associated with the prediction, with a narrow, peaked distribution indicating certainty, and a wide, flat distribution indicating uncertainty.

Although confidence and uncertainty estimation are considered essential for safe clinical deployment of AI [12], most machine learning methods in the recent medical literature are said to neglect the important issue of model uncertainty [10]. This is an issue, as it is well known that machine learning models can be confidently wrong [13]. In the absence of an uncertainty estimate, such confidently wrong model predictions may undermine end user trust in the model.

Accordingly, there is a need to understand how trust in AI manifests in the interaction between clinical users and an AI system that predicts both treatment probability and uncertainty, in the context of mental health decision support. This includes understanding how such trust affects decisions to use the system, and in which contexts. This qualitative study explores aspects of trust in clinicians' use of an AI model as a part of a decision support system for treatment prediction.

Clinician Trust in AI and the Role of Explainability

Although trust is a prerequisite for adoption, it must be appropriate. The review by Asan et al [14] emphasizes reliability, transparency, and alignment with clinical reasoning as central determinants of clinicians' trust in AI and argues that trust should be appropriate rather than maximized, in order to avoid unwarranted trust. In mental health specifically, the sensitivity of decisions and the potential for harm heighten the need for trustworthy behavior under distribution shift and within workflow constraints.

Explainability is often proposed as the main lever for calibrating trust; however, its effects are context-dependent. The systematic review by Rosenbacke et al [15] reports mixed effects of explainable artificial intelligence (XAI) on clinicians' trust. Explanations can increase trust when they are clear, concise, and clinically relevant, but they can also have no effect or reduce trust when they are confusing, cognitively heavy, or poorly matched to the task.

Evidence from applied settings shows similar patterns. The empirical study by Wysocki et al [16] examines explainability, utility, and trust in clinical decision making and describes a mixed role for explanations. Benefits include help with sense-making in ambiguous cases and support for less experienced staff, while drawbacks include added interaction effort and the risk of confirmation bias or over-reliance if limitations are not clear. These findings underline the relevance of the design and presentation of explanations in context.

Accordingly, explanation content should be clinically aligned and should be paired with uncertainty communication, such as calibrated confidence or abstention, in order to discourage unwarranted trust [14,15]. Altogether, trust improves when explanations are concise, fit the clinical task, acknowledge limitations, and respect clinicians' expertise; conversely, trust erodes when explanations are opaque, verbose, or misaligned with workflow [14-16].

Explainable AI and Trust

XAI refers to the ability of AI systems to provide explanations of their predictions in a way that is understandable and clear [17]. XAI techniques help bridge the gap between AI algorithms and user comprehension. In their transparency and interpretability for understandability framework, Joyce et al [18] describe how XAI relates to the application of AI in mental health. They argue that trust is built when clinicians and patients feel confident that an AI system's reasoning is transparent and is seen to align with clinical understanding—a concept similar to face validity. Jacovi et al [19] suggest that users tend to trust an AI if its decision-making process aligns with their

XSL•FO
RenderX

expectations of rational behavior. Transparency and interpretability are central to the evaluation of such a rationale.

However, the terms surrounding the explainability, interpretability, and transparency of AI models and their operation lack consensus definitions [20,21]. Explainability refers to the ability to understand why a model has made a specific prediction. Interpretability may be thought of as empowering explainability: providing the ability to understand how a prediction was made. Being able to convey the meaning of an AI model's operation to a human in understandable terms operationalizes the "how" and is therefore an important aspect of interpretability [21]. Interactivity with the AI model and visualization of its operations also play a role in interpretability [22]. Li et al [23] point out that the user interface is a key factor influencing user experience, acting as the main medium through which an AI system communicates its internal reasoning and decision-making process. The way explanations are presented, therefore, plays an important role in how users form trust in the AI.

Transparency, which is related to interpretability, emphasizes a shared, user-centered understanding of the meaning of model features and their operation [19,21]. Transparency implies that the whole model can be understood [20], or in a less strict sense, that the model features and parameters are understandable [17]. In clinical models, transparency implies that the meaning of features has a direct clinical interpretation [18].

Lacking a consensus and in consideration of the various aspects of explainability, involving trust, understanding, confidence, etc, Barredo Arrieta et al [21] suggest that explainability depends on who is required to understand the model, suggesting the following definition: "Given an audience, an explainable Artificial Intelligence is one that produces details or reasons to make its functioning clear or easy to understand." In a clinical sense, this may be thought of as AI coformulation.

Understandability is used as a proxy for explainability in the transparency and interpretability for understandability framework of XAI [18]. In this framework, transparency and interpretability are the constituent components of understandability (and therefore, explainability). Consistent with the description of those terms above, interpretability is seen as being determined by the model architecture, structure, and how the prediction is presented (similar to the concept of algorithmic transparency); transparency is determined by the features and data used by the model. Taking logistic regression as an example, interpretability comes from the linear model structure and probabilistic presentation of model predictions inherent in logistic regression. The data and features of such a model can be designed so that the meaning of features has a direct clinical interpretation, fulfilling the requirement for transparency.

Jacovi et al [19] frame trust as an explicit contract between the AI and users in a context in which there is a risk to the user if the AI breaks the contract. The contract lays out what the model is being trusted to provide: transparency, predictability, and fairness (for example). Predictability in this context means that model performance is known to the user and accepted. This may involve an overall performance metric, but it may also involve contextual conditionality, implying that the user knows in which context to trust the model performance and in which contexts the model should not be trusted (eg, on certain data subgroups or out-of-domain data). The mechanisms by which an AI model gains a user's trust can be understood through a differentiation into intrinsic and extrinsic factors. Intrinsic trust is based on the AI's reasoning process aligning with human expectations (eg, interpretable models, logical decision-making). Therefore, intrinsic trust depends on the user, for example, a clinician, being capable of interpreting the meaning of the features of a transparent clinical model. Extrinsic trust is based on external behavior or evaluations of the model, such as observing consistent past performance or validation of performance reliability. Extrinsic trust is therefore a trust in the model's evaluation quality. In this way, trust is seen as having both human and quantitative factors, upon which the user judges whether to trust the model or not.

Trust in AI systems often involves a degree of control relinquishment, yet this trust is highly contextual. As Lipton [20] points out, interpretability is not a one-size-fits-all goal, and trust arises not just from transparency but from the alignment of model behavior with user expectations in specific settings. In many cases, users are willing to trust AI when its judgments match their own, whether or not those judgments are objectively correct. That is, both humans and AI may classify a case as a "mistake" or "non-mistake" depending on the category and context, reinforcing or undermining trust.

This context-sensitive nature of trust highlights the importance of aligning AI outputs with human reasoning, not just through outcomes but through explanations. AI models typically associate data features with predictions based on correlations, whereas humans tend to understand scientific models through cause-and-effect relationships. To bridge this conceptual gap, Holzinger et al [24] introduced the concept of causability, defined as the extent to which AI explanations support users in forming causal understanding from inherently correlational outputs.

In the context of AI recommender systems, Shin [25] found that users' trust increased when they were able to comprehend explanations in causal terms. Rather than offering purely statistical associations ("feature X correlates with outcome Y"), causability emphasizes understanding why such correlations arise. As such, it is considered a prerequisite for effective explainability. A comparable shift can be observed in high-stakes fields such as suicide prevention, where decision-making must be both evidence-informed and explainable to clinicians. In suicide research, for example, there has been a call to move from risk factors to predictive algorithms [26], and clinical guidelines now advocate a shift away from categorical risk classification toward individualized, formulation-based approaches [27,28].

According to Shin [25], trust in AI systems develops through a dual process: a heuristic route, where users form quick judgments based on prior knowledge or biases, and a systematic route, where trust emerges from deeper evaluation, particularly when explanations are clear, meaningful, and causally informative.

XSL•FO
RenderX

Since explainability depends on humans seeking the explanation, the views of users in the clinical context are central. The purpose of this qualitative study is to explore what affects the trust as perceived by clinical users of an AI model aimed at a mental health decision support system. Using thematic analysis, this study explores how users understand the operation of the model, how they interpret its outputs, and what influences their trust. By focusing on user perspectives, it aims to develop a context-sensitive understanding of explainability that is grounded in clinical reasoning and practice.

Methods

Study Design and Procedure

The clinical setting for this study was the Internet Psychiatry Clinic, an internet-based cognitive behavioral therapy (iCBT) service operated by the Centre for Digital Psychiatry within the Mental Health Services of the Region of Southern Denmark [29]. This publicly funded, free-of-charge service is integrated into routine mental health care in Denmark. Individuals self-refer and are screened in 2 stages: an initial request form is followed by a clinical assessment conducted by trained psychologists. The screening referred to in this study is the first part of the clinical assessment, with the second part being a clinical interview. Eligibility criteria include being aged 18 years or older and meeting diagnostic criteria for one of the following conditions: mild to moderate depressive disorder, single phobia, social phobia, or panic disorder with or without agoraphobia [30]. The center delivers its iCBT treatment with a nationwide coverage of approximately 4.8 million adults. The AI model and human-computer interface (HCI) used in the moderated cognitive walkthrough is the Probabilistic Integrated Scoring Model (PrISM) originally reported by Kelly et al [31] (further described below), that predicts the probabilities of the 4 available treatment options—depression, panic disorder, social phobia, and specific phobia; based on patient questionnaires in step 1 of the clinical assessment.

In this study, following informed consent, semistructured interviews were conducted with the clinical psychologist participants, guided by the frameworks of perceived trust in AI discussed earlier. The interviews took place in the clinic in Odense, Denmark, over a period of 2 days. Interviews were conducted one-on-one with one participant and one researcher present in the room.

A trust framework-driven interview guide (Multimedia Appendix 1) was used to ensure that key dimensions of trust were systematically explored across all participants, while also allowing space for participants to raise personal or contextual factors that may have influenced their perception of AI [32]. Questions focused on concepts such as explainability, interpretability, confidence, uncertainty, and perceived trust. The flexible nature of the semistructured format made it possible to follow up on unexpected or participant-driven topics that emerged during the discussion.

To support the interviews, a moderated cognitive walkthrough was carried out. During this process, participants interacted with the AI system while responding to questions. The interaction involved navigating synthetic patient records, using the AI to generate treatment predictions, and reviewing both the prediction and its explanation through the user interface. This method allowed for the observation of trust-related responses during actual system use, aiming to minimize the influence of the interview structure on participant behavior, and helped to identify usability issues that may influence trust [33]. The cognitive walkthrough is particularly useful in assessing the user experience and understanding how people cognitively process the interface [34]. In this study, it also served to situate participants' reflections in a realistic decision-support context, linking system outputs to users' thought processes in real time.

Individual interviews provided insight into the cognitive, emotional, and experiential factors shaping trust in AI [32], while the cognitive walkthrough allowed for an examination of how perceptions evolved during live interaction with the system so that the interview structure affected behavior as little as possible [34]. Together, these methods enable a deeper exploration of the factors influencing trust, providing both subjective self-reports and objective behavioral data.

This qualitative study used reflexive thematic analysis (RTA) [35,36] as an individual case study [37] to explore participants' experiences of using an AI-enabled decision support tool in a Danish web-based mental health service. The aim was to investigate how clinicians perceived trust and explainability in the AI model during real-time interaction. Data were generated using think-aloud protocols embedded within the semistructured interviews, allowing participants to verbalize their thoughts and responses as they navigated the AI interface. This approach made it possible to capture both spontaneous reactions and more reflective evaluations of the model's performance and trustworthiness.

Individual interviews were selected to provide participants with a private and uninterrupted setting to share their views. This was considered particularly important given the sensitivity of topics such as trust and clinical judgment, where participants might feel more comfortable expressing concerns or critiques in a one-on-one context [32]. The interviews were audio recorded consensually, for transcription purposes. The semistructured interview format allowed for flexibility in exploring participants' responses while ensuring a consistent set of topics related to perceived trust and explainability in AI [38].

The RTA was conducted by 2 researchers independently, with complementary lenses. Reflexively, one of these (AK) brought an AI trust perspective to the analysis. The other (TTHSM) brought a clinical perspective. Our aim was interpretive, focusing on meaning-making rather than reliability metrics. Consistent with RTA, we did not seek intercoder agreement, did not hold intercoder dialogue aimed at consensus, and did not conduct member-checking. The analysis followed Braun and Clarke's [35] six-phase process: (1) familiarization with the data, (2) initial coding, (3) generating initial themes, (4) reviewing themes, (5) defining and naming themes, and (6) producing the report. The analysis was conducted as an iterative, reflexive, and nonlinear process, in line with Braun and Clarke's [36] guidance for RTA [39]. This involved the formulation of

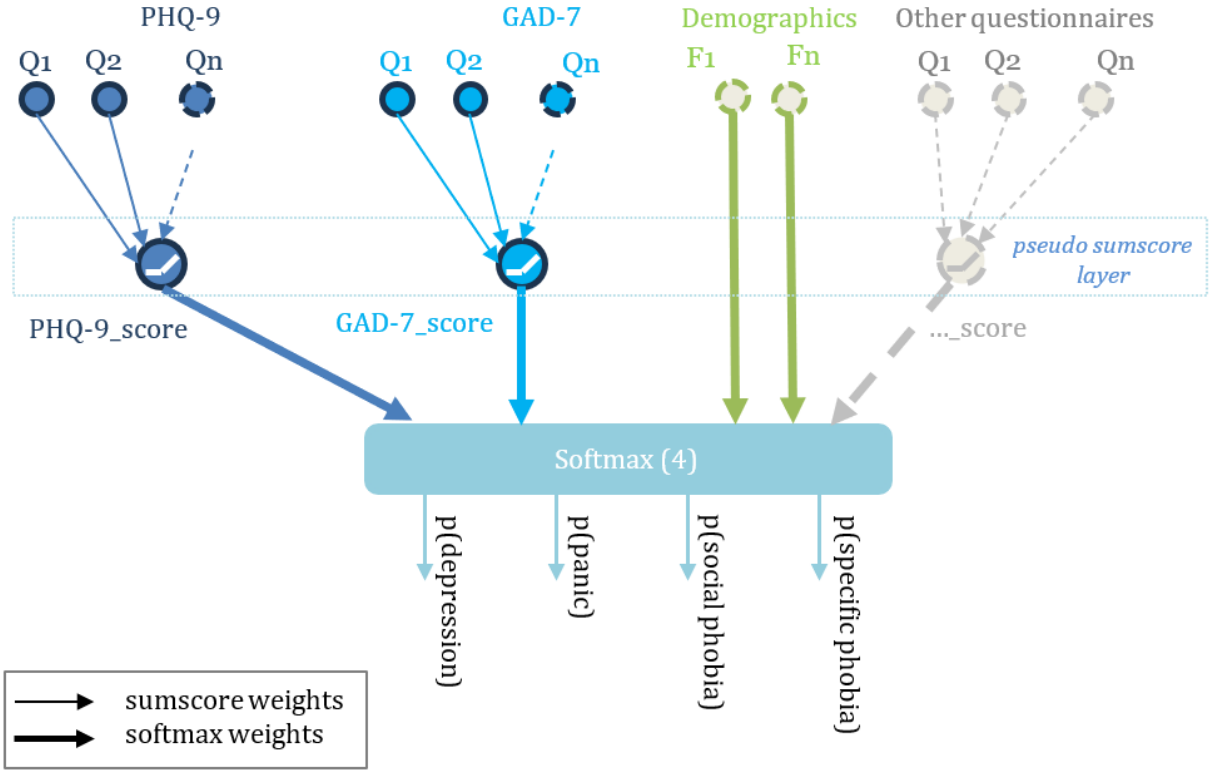
XSL•FO
RenderX

semantic codes of surface meaning, but also where an underlying meaning emerged, latent codes were formulated, involving a more subjective interpretation.

AI Model

The PrISM AI model [31] (Figure 1), used in the moderated cognitive walkthrough, is a 2-level feed-forward neural network that emulates the traditional "sum-score" logic of mental-health screening instruments while retaining full differentiability and probabilistic predictions. Item-level responses from 6 validated questionnaires—Patient Health Questionnaire-9 (PHQ-9) [40], the Generalized Anxiety Disorder Questionnaire-7 [41], the Social Interaction Anxiety Scale (SIAS) [42], the Panic Disorder Severity Scale [43], and the Fear Questionnaire [44], and selected MANSA items—together with 8 sociodemographic variables constitute the input vector. In the pseudosum layer, each questionnaire's items are combined by nonnegative, L1-regularized weights and passed through a rectified-linear activation, yielding learned questionnaire scores that remain directly interpretable to clinicians. These scores, augmented by the raw demographic features via a skip connection, are then supplied to a softmax output layer that performs multiclass logistic regression across 4 treatment categories (depression, panic disorder, social phobia, and specific phobia).

Figure 1. Structure of the Probabilistic Integrated Scoring Model (PrISM). GAD-7: Generalized Anxiety Disorder Questionnaire-7; PHQ-9: Patient Health Questionnaire-9.



Uncertainty was quantified at inference time through Monte Carlo dropout: a 5% dropout mask is retained during every forward pass, and the network is sampled 100 times, producing an empirical posterior distribution over class probabilities. This procedure approximates a fully Bayesian logistic model while incurring negligible computational overhead and enables case-level reporting of both confidence and epistemic uncertainty.

The user interface to the PrISM model, presented to the participants, is shown in Figure 2. The top panel shows violin plots of the posterior distribution over class probabilities for the treatment prediction for a specific patient example, indicating a prediction of depression. The explanation of the prediction is shown in the middle panel. This shows graphs of ranked feature importance; the features that explain the class prediction selected (left) and the features that counter that prediction (right). The "explain prediction" dropdown above the panel allows the user to select which of the 4 class predictions (depression, panic disorder, social phobia, and specific phobia) to explain. The bottom plot shows individual questionnaire answers that relate to a selected questionnaire.

XSL•FO
RenderX

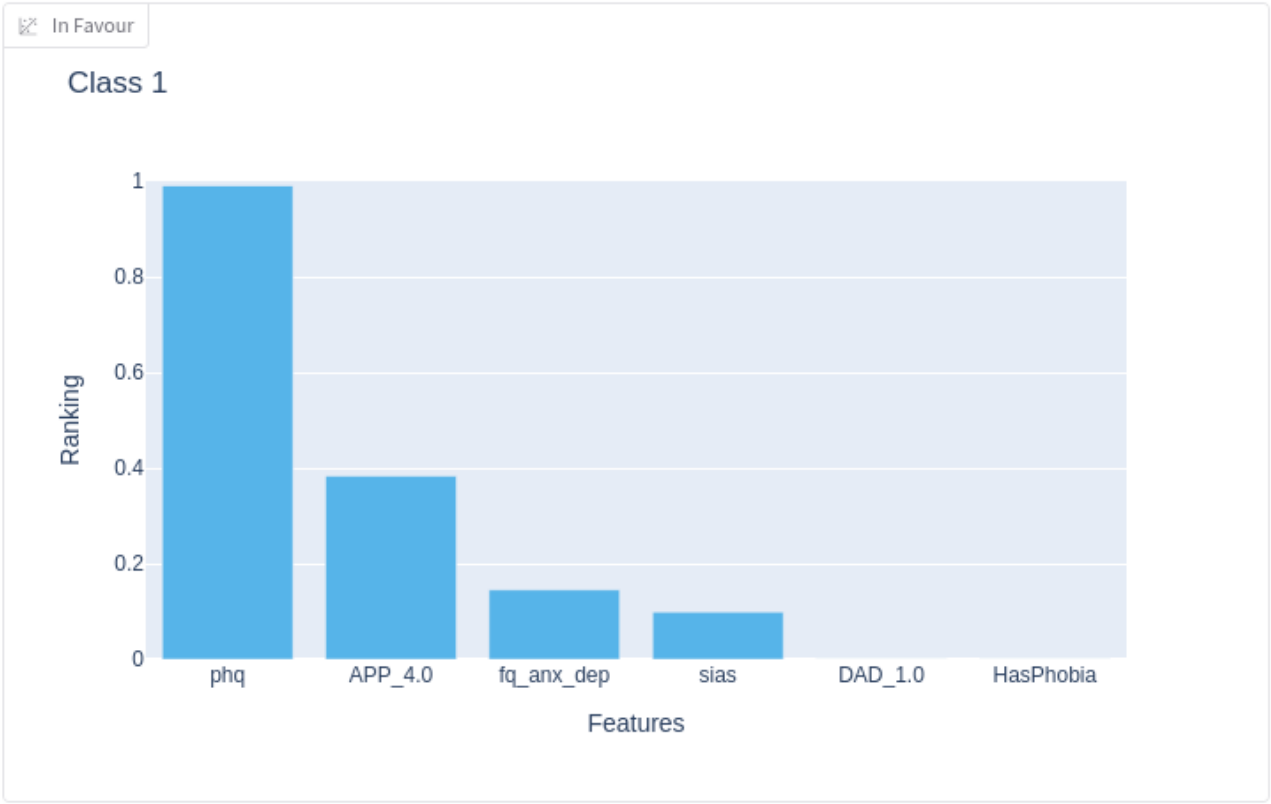
Figure 2. Illustration of the artificial intelligence human-computer interface. The top plot shows the prediction of the 4 classes (y-axis), with the probability distribution of the predictions on the x-axis, illustrated as violin plots. The middle plots show the features that explain the class prediction selected (left) and the features that counter that prediction (right). The bottom plot shows individual questionnaire answers that relate to a selected questionnaire.



The detail of the prediction explanation for a specific patient is shown in further detail in Figure 3. The chart indicates that the PHQ-9 score (phq), the demographic feature “doesn’t know what treatment to apply for” (APP_4.0), and the fear questionnaire anxiety-depression subscale (fq_anx_dep) have the most influence on the prediction of the depression class (class 1) in this specific patient example. When interacting with the AI HCI, hover-overs provided text explaining the meaning of the features and score totals where relevant, for improved usability.

XSL•FO
RenderX

Figure 3. Details of the features that explain the class prediction selected (left middle panel of the artificial intelligence human-computer interface).



The reference implementation (TensorFlow 2.19, Python 3.11) was trained on retrospectively collected data from 1068 adults referred to the Danish national internet-psychiatry service. Records were stratified by diagnosis and partitioned 75%/25% into development and held-out test sets. Optimization used the Adam algorithm (batch=32, learning rate=0.001) for 2000 epochs, with L1-regularization (λ=0.001) promoting sparsity and retaining only clinically salient inputs.

On the unseen test cohort, the model achieved a balanced accuracy of 0.79 and per-class AUC ranging from 0.91 to 0.98. This architecture, therefore, provides two properties considered essential for the safe deployment of machine-learning systems in frontline mental health care: (1) transparent, score-level explanations that align with routine clinical reasoning and (2) probability distributions that facilitate the identification of low-confidence or uncertain cases.

Data

A dataset comprising patient questionnaire answers and demographics was used to train the PrISM AI model used in the cognitive walkthrough semistructured interviews. The data were gathered between November 14, 2019, and December 31, 2022, from patients of the “Internetpsykiatrien” web-based treatment services delivered by the Centre for Digital Psychiatry, Denmark. The center delivers routine care and internet-delivered cognitive behavioral therapy with nationwide coverage [29]. The data consisted of answers to the PHQ-9, the Generalized Anxiety Disorder Questionnaire-7, the SIAS, the Panic Disorder Severity Scale, and the Fear Questionnaire. Additional features included demographic information and a brief medical history, such as previously diagnosed conditions. The ground truth of the dataset was the choice of treatment selected by the psychologists in the clinic for each patient. Further details can be found in [31].

Ethical Considerations

Data for the AI model were extracted from Internetpsykiatrien after approval from the Regional Council in Southern Denmark. Since this was a secondary data analysis, separate informed consent was not required. The Regional Committees on Health Research Ethics for Southern Denmark were informed about the study and were provided with the case number S-20232000 - 65. In accordance with Danish national ethical guidelines, no additional ethics approval was needed.

Ethical review and approval were not required for the study on human participants in accordance with the local legislation. The patients/participants provided their written informed consent to participate in this study. The study was reported to the Danish Data Protection Agency. Data were anonymized before being provided to the researchers. No compensation was provided.

Participants

The participants were clinical psychologists working in the Region of Southern Denmark. The group comprised 5 participants (N=5; 2 female and 3 male), with a mean age of 35.8 years. All participants were members of the screening team at an iCBT clinic within the service. Their role involves screening patients for interview prior to possible treatment, as well as conducting clinical interviews as part of the treatment assessment process.

Participants were recruited based on their clinical role in the service, with a range of experience levels represented, from junior to senior staff. The study included 55% (5/9) of the entire

XSL•FO
RenderX

screening team, meaning that the majority of clinicians responsible for initial patient triage were included. This provides a strong representation of the perspectives within the clinical setting and enhances the validity of the findings. To preserve anonymity within this small team, specific age and experience ranges are not disclosed.

In a review of sample size requirements for usability studies, Lewis [45] concluded that 5 participants are typically sufficient to reveal around 80% of usability issues. This implies a sample size of 5 is considered appropriate for a qualitative study focused on human-AI interaction. Thematic saturation appeared to be reached, and the inclusion of most of the staff involved in the relevant clinical task provides additional justification for the adequacy of the sample.

Structured Interview Guide

A structured interview guide was developed to explore participants' experiences with the AI model and interface. The interview's semistructured format guided the participants through a process of (1) enquiring about attitudes to AI in mental health care and in their role; (2) a cognitive walkthrough of using the AI system with various patient scenarios; and (3) an in-depth discussion on trust and explainability (Table 1). The guide consisted of open-ended questions designed to elicit detailed responses about usability and trust. Once introduced to the system (Table 1), the Cognitive Walkthrough and Interface Interaction section allowed for an open use of the AI system with the synthetic patient profiles, allowing an observation of natural behavior with the system, in order to minimize the influence of the interview structure on participant behavior. Since the interviews were limited to 1 hour, further follow-up questions were answered in written form. These included the topics: (1) usability and workflow integration, and (2) feedback and improvement.

XSL•FO
RenderX

Table . Description of the structured interview.

Topic and subtopic	Explanation/example question
Opening and background	
Reconfirm the informed consent	—[a]
Participant demographics	If not collected ahead of time
Initial attitudes toward AI[b] in health care	When you hear "AI" in the context of patient care, what comes to mind?
Perceived usefulness of AI in health care	Based on your experience, do you think AI could be useful in patient care or treatment decision-making?
Introduction to the AI system	A brief explanation of the system's function as a treatment recommendation system for mental health and set expectations for the cognitive walk-through.
Introduction to AI definitions	
Overview of the AI model	
Instructions for cognitive walkthrough	
Cognitive walkthrough and interface interaction	
First impressions	Take a look at the interface. What are your initial impressions?
Clinical assessment	Looking at the patient record, what would your assessment of the treatment recommendation be for this patient?
Locating and understanding the recommendation	Please use the AI tool to make a prediction for the patient.
Interpreting confidence and uncertainty	Do you feel you understand whether the model is "very sure" or "not very sure" about its recommendation?
Accessing and reviewing the explanation	Can you find or access the explanation for why the AI made this recommendation?
Explanation depth and satisfaction	Does the system provide enough detail about how it arrived at its decision?
Trust and ease of use	So far, do you feel comfortable trusting the system's recommendations? Why or why not?
In-depth discussion on trust and explainability	
Reliability and past evidence	Would it help your trust if you saw data on how often the AI's predictions were correct in previous cases of this type?
Transparency	How important is it for you to see exactly how the AI weighs different questionnaire scores?
Interpretability	Would it help to understand how this model generally works, beyond just one case?
Interactivity	Do you feel the interface is interactive enough for you to explore the patient record and model explanations?
Fairness	Do you have concerns about the AI being biased or missing certain patient subtleties?
Closing	
Summarize final thoughts and close	What would be your key takeaway from this experience with the AI system?

[a]Not applicable.

[b]AI: artificial intelligence.

To clarify the definitions of the terms used in the interviews, the definitions in Table 2 were provided to participants. Although confidence and uncertainty are somewhat ambiguous, we took confidence to be aligned with aleatoric uncertainty and uncertainty to be aligned with epistemic uncertainty.

XSL•FO
RenderX

Table . Definitions of terms provided to participants.

Term	Description
Explainable (AI[a])	Relates to understanding "why" AI models make their decisions.
Interpretable (AI)	Understanding "how" an AI model makes its prediction.
Transparent (AI)	How the prediction is made is clear from the model and features.
Confidence	The "probability" of the prediction.
Uncertainty	How "wide" the probability distribution is.

[a]AI: artificial intelligence.

Patient Profiles

Four patient profiles were developed to support the moderated cognitive walkthrough and provide a basis for exploration during the interviews. These profiles were designed to represent a diversity of treatment predictions and comorbidities that are typical of treatment decisions routinely seen in the clinic (Table 3). The patient profiles were accompanied by corresponding data suitable for presentation to the AI model. The data involves the answers to patient questionnaires and demographic data for the various patient scenarios [31]. The data was presented to the participants via a patient record view in the AI interface. The written description, provided in Table 3 for context, was not presented until the postinterview debrief.

XSL•FO
RenderX

Table . Patient profile scenarios.

Scenarios	Patient profile	Clinical rationale
Scenario 1	• Jane presents with persistent low mood, a significant decrease in energy levels, and an overwhelming sense of hopelessness. She reports difficulty concentrating at work and has recently been reprimanded for missing deadlines, which has exacerbated feelings of inadequacy. She also experiences intense anxiety in social and professional settings, including meetings and casual social gatherings, which has led her to withdraw from colleagues and friends. Jane's symptoms began approximately 8 months ago, following a critical performance review at work. Her lack of energy and motivation caused her low work performance, she believes. Initially, she noticed an increased reluctance to speak up in meetings due to fear of judgment. Over time, this fear extended to informal social settings, making her avoid after-work social events and even casual conversations. Concurrently, she began experiencing a pervasive sense of sadness, feeling "empty" most days, with episodes of crying without a clear trigger. • She reports difficulty sleeping, often lying awake ruminating over interactions where she feels she may have been perceived as inadequate. She has lost interest in hobbies she once enjoyed, such as painting and attending book club meetings, citing both a lack of motivation and a fear of being judged by others. Previous episodes: Jane recalls experiencing a similar period of low mood and social avoidance during her late teens but did not seek treatment. Family history: A history of depression in her father. No significant medical conditions or history of substance abuse.	• Ambivalent case where either depression or social anxiety could be primary. • Would tend to treat depression as a priority.

XSL•FO
RenderX

Scenarios	Patient profile	Clinical rationale
Scenario 2	• Susan M, a 34-year-old administrative assistant, presents with a 6-month history of persistent low mood, chronic worry, and reduced functioning. She describes feeling "stuck" and unable to engage fully in her personal and professional life. Her symptoms began following a period of prolonged work-related stress, which has since resolved, but her distress has persisted and gradually intensified. • Susan reports a constant sense of unease that permeates most of her day, particularly in the mornings. She wakes up feeling restless and on edge, with racing thoughts about her responsibilities and potential shortcomings. These worries often spiral into feelings of inadequacy, leaving her preoccupied with the fear of disappointing others. Despite recognizing that many of her concerns are unlikely to materialize, she struggles to control or dismiss them. She also reports difficulty concentrating, which she attributes to her mental preoccupation and lack of energy. There is some anxiety and depression history in her family, with her brother having had severe depression at some point and her mother being an anxious and overbearing person.	• Primary depression with features of generalized anxiety.
Scenario 3		• Panic disorder.

XSL•FO
RenderX

Scenarios	Patient profile	Clinical rationale
	• John D is a 42-year-old accountant who is currently unemployed. He describes experiencing sudden episodes of intense fear, accompanied by physical symptoms such as chest tightness, shortness of breath, sweating, and dizziness. He reports catastrophic thoughts about having a heart attack. These episodes, which John recognizes as panic attacks, occur unpredictably, often when he is driving or in crowded spaces like grocery stores. The fear of these episodes has led him to limit his outings and rely increasingly on his wife to run errands, though he continues to apply for jobs where he can work from home. John admits that he has started avoiding situations where he feels escape might be difficult, such as long car rides, crowded events, or unfamiliar environments. While he can still attend essential outings, such as dropping his kids off at school, he often feels on edge and hypervigilant for signs of an impending attack. These restrictions have caused strain in his family life, as he worries about burdening his wife and missing family outings. To cope with his anxiety, John has started drinking small amounts of alcohol in the evenings, which he finds calms his nerves and helps him unwind. He is careful not to exceed the recommended daily intake, consuming 1 to 2 drinks a night, and denies any history of problematic drinking. However, he acknowledges that he relies on this habit to manage his discomfort and worries about its long-term implications. John has experienced a general sense of fatigue and low mood. While he denies feeling hopeless or experiencing anhedonia, he describes difficulty finding joy in his usual hobbies, such as woodworking and jogging, because he feels “too keyed up” or “too drained” to engage. His sleep is often restless, disrupted by worry about his health and family, though he sleeps sufficiently most nights. He also reports a pervasive sense of guilt, believing that his anxiety is a weakness and that he is failing to be a strong role model for his children. When John went to the gymnasium, he received a diagnosis of anxiety, but he doesn’t remember much about that, and he did not receive treatment other than antianxiety medication.	
Scenario 4		• Single phobia

XSL•FO
RenderX

Scenarios	Patient profile	Clinical rationale
	• Mark, a 37-year-old software developer, presents with a persistent fear of bridges and heights that has mildly impacted his daily life for the past several years. His fear began in his late twenties after a stressful experience driving across a high suspension bridge during a storm, which left him feeling uneasy and lightheaded. While he did not experience a full-blown panic attack, the incident caused him to develop a wariness of high bridges. Over time, this fear generalized to other situations involving heights, such as tall balconies or observation decks. Mark avoids bridges when convenient but does not let his fear entirely dictate his life. He reports experiencing mild physical symptoms, such as increased heart rate and sweating, but these subside once he is off the bridge. Mark also feels discomfort in high places where there is an open or exposed view, such as glass elevators or balconies with low railings. He avoids these situations when possible but can manage if necessary, often relying on grounding techniques like focusing on a fixed point or staying close to a wall. Mark is seeking therapy now because he wants to feel more confident in situations involving heights, particularly as his daughter has expressed interest in visiting landmarks like bridges and observation towers. He is motivated to address his fear so that he can fully participate in family activities without hesitation or discomfort.	

Results

Thematic Analysis

Consistent with Braun and Clarke's [36] reflexive approach, we aimed for reflexive interpretive richness, rather than intercoder reliability, with each coder's distinct point of view contributing to a wider perspective interpreting the data. Therefore, 2 analysts with complementary skills (AK and TTHSM) undertook the RTA, each bringing a distinct epistemic lens.

Coder A, with a nonclinical background in AI trust and HCI, designs and evaluates decision-support tools, holding a pragmatic stance that emphasizes usability, explainability, and uncertainty communication. They adopted a predominantly inductive, bottom-up stance. First, descriptive codes were generated and clustered into provisional subthemes. Since a "trust journey" emerged in the subthemes, a top-down, deductive pass interrogated the data through trust frameworks to articulate the journey in terms of established models. Any latent emerging themes were also preserved.

Coder B, who practices in psychology, prioritizes patient safety, workflow fit, and risk management. They worked with a context-led, primarily semantic focus, ensuring that emergent themes remained grounded in front-line priorities such as areas of clinical hesitancy and reassurance, and how AI intersects with their duty of care.

Braun and Clarke [36] identify this analytic-storytelling mode within reflexive TA when the aim is to surface latent, practice-based logics that a purely code-centric read might miss.

The RTA generated 4 sequential themes. Together they outline a notional "trust journey" that begins with a first encounter and culminates in an informed decision to adopt the tool in real-world screening.

1. "Why engage with AI?": participants articulate the anticipated utility that motivates engagement.
2. "Making sense of the model": they attempt to follow and investigate the system's reasoning, building intrinsic trust and causability.
3. "Should I rely on it?": they appraised the AI by interpreting probabilistic outputs and explanations in the light of their clinical judgment of the patient record and current work practices.
4. "Deciding to use the model in practice": they elaborated on when, where, and under what safeguards they would rely on the AI in clinical screening.

In Figure 4, the themes are illustrated as a process of cognitive steps in the process of establishing trust:

1. Anticipated utility: "Why engage with AI?"

XSL•FO
RenderX

2. Sense making: “Making sense of the model.”
3. Appraisal: “Should I rely on it?”
4. Deciding to rely: “Deciding to use the model in practice”

Figure 4. Stages of artificial intelligence–mediated trust.



In the themes that follow, quotations are referenced by participant ID (P1-P5); square brackets denote clarifications.

Theme 1: “Why Engage With AI?”—Anticipated Utility

Although none of the clinicians had deployed AI tools in routine practice, four of the five were already experimenting with generative AI in their personal lives. This prior exposure produced a sense of guarded optimism about AI: as one participant put it, in relation to generative AI:

> *I don’t necessarily trust the idea it provides; I trust my own interpretation of the idea.* [P5]

The perceived value of an AI-enabled screening tool is centered around three expectations:

1. Improved efficiency and reduced bureaucracy: Highlighting that AI could examine “Multiple variables at a quick pace—perhaps faster than an experienced clinician could do” (P1); and that screening “Needs to go fast” (P4).
2. Greater consistency and mitigation of human bias: “I think a model like this could maybe also fight...prejudice or bias to some degree” (P2); “I would hope it would be more sure than a person” (P3).
3. Reallocation of clinician time to higher-value tasks: “If it frees me to do more clinical work, that’s very nice” (P2); “Less time looking at numbers” (P5).

Overall, participants framed AI as a complementary aid rather than a replacement for clinical judgment. They were willing to engage with the technology because they believed it could accelerate mundane evaluation tasks, apply rules impartially, and create time for deeper patient engagement. This motivational backdrop set the tone for the subsequent trust-building process described in the following themes.

Theme 2: “Making Sense of the Model”—Sense-Making Through Intrinsic Trust and Causability

Theme 2 reflects the initial stage of trust-formation where participants inspected violin plots, feature-level importance, and question-level explanations to decide whether the AI’s logic aligned with their mental model of clinical reasoning. Since explanations are only effective if they are understandable to the user [25], this theme captures how participants build intrinsic trust in the AI by attempting to follow and investigate the system’s reasoning.

This theme comprises four subthemes that relate to sense-making and trust-building: (1) orientation and navigation, capturing how participants initially find their way around the interface and data; (2) explainability-by-design, capturing aspects of the model or UI design that improve understanding, currently or in the future; (3) AI-driven insight, capturing clinical or data insights revealed by the model that the user had not initially noticed; (4) UI improvement/usability, capturing improvements that may make the interface easier to use.

During “Orientation and Navigation,” participants navigated the interface easily, spending time on the patient record tab to make their own assessment of the treatment required (as directed in the structured interview). The questionnaire total scores on this tab (eg, PHQ-9, SIAS) were familiar and therefore aided sense-making, often in the form of a patient story:

> *And then there’s a graph...And it tells me that there’s...both social phobia and Depression score. There’s symptoms of generalised anxiety and...could be...multiple diagnosis, but it could also be...an image of someone who's depressed and...feels very self-conscious about...dealing with depression. They might be very worried about the future...* [P2]

However, in the “Explainability-by-Design” subtheme, the y-axis scaling of the bar chart of the questionnaire total scores caused some confusion:

> *The bar graph on the scores page may mislead people because that scale has a larger range, it looks like, statistically, it may be much higher, but the range for the depression scale is much smaller.* [P1]

Other common understandability improvements involved the provision of hover-overs giving more information on the violin plots*:*

> *I wish there was some hover over’s here [points at the violin plots].* [P1]

The graphical explainability, which was enabled by the model transparency, was seen as a major bonus of the AI:

> *the graphics impressed me the most, but I think it’s very nice that you can go down and see how the model works* [P2]

While using the AI model for prediction and subsequently exploring the explanations of the predictions in the interface, some novel insights were revealed, as captured by the “AI-Driven Insight” subtheme. For example, participants found scenario 4 (Table 3) challenging to assess because the questionnaire total scores, which they typically relied on, were uniformly low in this case. The recommended scenario treatment was Specific Phobia, which was explained mainly by the

XSL•FO
RenderX

demographic question “Has Phobia.” When the AI prediction explanation revealed this, many participants expressed surprise:

> *So, they answered yes. Oh, didn’t see that* [P3]

> *So looking at this...What I'm thinking is I should probably have checked the fear questionnaire for specific phobias, OK, because that's nice.* [P2]

> *So, I will try to explain...to see what it says.[gets explanation from UI] They have actually said that they have a phobia...I think again, it makes sense, if it is...some kind of phobia.* [P4]

In one case, the prediction in the form of the violin plots, rather than the model explanation, revealed the insight:

> *So this one, this definitely helped me this graph [violin plot]...More so than the other ones, because well, because I didn't see the fear...the specific phobia.* [P1]

As well as highlighting unseen aspects of the patient record, the violin plots also revealed insights about uncertainty, which is typified by this comment from participant 1:

> *The probability level wouldn't really matter. It’s just the uncertainty. If they are overlapping too much...definitely it would need to be reviewed by the team.* [P1]

In the subtheme of “UI Improvement/Usability”, participant 2 saw the Violin plots as the main aspect of sense-making:

> *As a Screener I would probably...just look at the photo [violin plots] and then I would...If it wasn’t as clear cut as this. I would click around. But you know it needs to just needs to be easy to work with.* [P2]

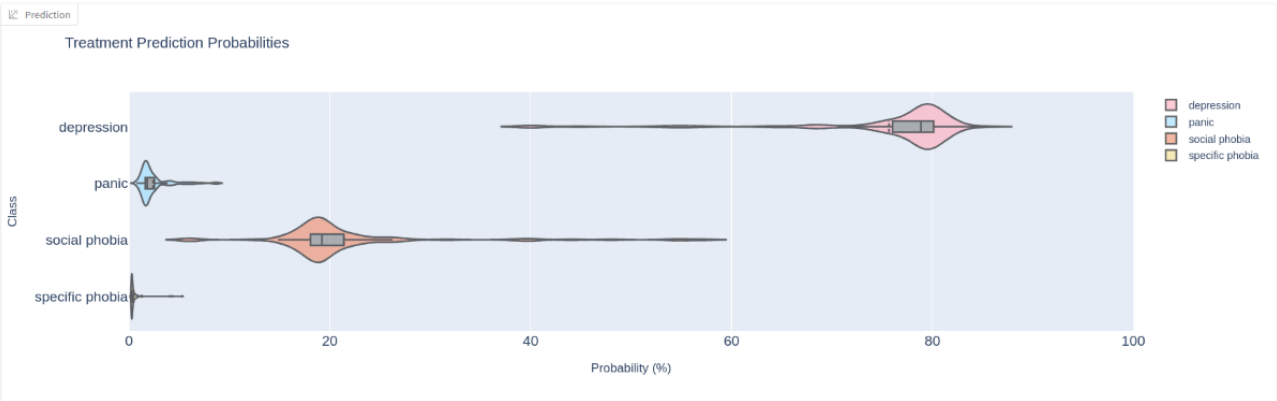
Theme 3: “Should I Rely on It”—Risk and Uncertainty Appraisal

Since AI adoption depends on user trust [25], theme 3 captured the bridge between making sense of the AI model and deciding to rely on it.

The guided interviews started by defining confidence and uncertainty to set the foundation for the concepts to be discussed. However, the terminology was hard to remember, and the terms were often confused with each other. Nevertheless, participants demonstrated understanding of the concepts from the violin plots, often articulated using their own terms. For example, when referring to confidence, participant 1 says, “Maybe even 70, I think would be sufficient for probability if we don’t have a lot of uncertainty either.” (P1); when referring to uncertainty, participant 3 correctly identifies it as relating to the dispersion of the violin plot, articulating it as being “unsure” “And then social phobia is somewhere between 20% and 40%, but it seems a lot more unsure about what this is actually about” (P3). This shows us that the terminology may be confusing, but the concepts were clear. For example, participants intuitively interpret the violin plots (Figure 5) to reveal the confidence (probability) and uncertainty (dispersion). Participant 4 articulates this well:

> *I think I would do it a little bit primitively, just look at where is this big round shape and where is it on the probability percent and yeah, that’s probably how I would do it...And then...about how certain it is? Yeah. OK, so I don’t know if this is very wide or not because I think that this fat bit is quite concentrated...But I think that to me, intuitively I would look at the round bit.* [P4]

Figure 5. Treatment prediction violin plot for hypothetical patient 1: treatment-depression.



All participants saw the absence of free-text input as a significant limitation of the AI model. In their normal workflow, free text provides significant information:

> *looking at what did the person actually write? I think that that’s actually where we get the most information.* [P2]

Participant 1 stressed that the free text provides important clarifying information over and above the questionnaire scores:

> *This...number three here [patient example 3] I think would need to be reviewed because there could be information in the free text that would either clarify this or make us think there’s...it’s too complicated or something else is going on that’s not picked up by our screening measures or the AI prediction.* [P1]

Although free text was seen as complementary to certain questionnaire items—for example, the Fear questionnaire specific phobia question where the free text describes the specific phobia (P1)—the lack of free-text input was also seen as a limitation to adoption:

> *I wouldn’t be comfortable not having the patient’s own words...because it can mean something else.* [P4]

XSL•FO
RenderX

> *Sometimes it can be a very clear-cut depression on the scores alone, but then we read the text, and this is a clear rejection because what they're applying for is not correct. So, I think if the AI sort of interpret the qualitative text as well and that be even better...I'd be even more in favour of it* [P5]

Theme 4: "Deciding to Use the Model"—Extrinsic Trust/Reliance

Theme 4 captures how participants moved from appraising risk to deciding to use the model in practice. In this decision phase, participants looked for external confirmation in the form of agreement with their own judgment, evidence of validation, and consistent performance, before deciding if the AI model could be relied upon, and in what circumstances. Therefore, this theme captured how the AI model fits in the process of patient screening, and the limits of perceived competence of the AI and the data it is trained on.

As patient scenarios were examined, participants demonstrated skill in the use of the AI system: making predictions, drilling down into explanations, then jumping back to the record for demographics. Participant 4 illustrates this fluidity as they gain practice assessing patient scenarios, as though they are conversing with the AI as they think-aloud:

> *We start with just panic and say 'why did you choose this' and they say that the PDSS score is a significant, which makes sense I think, and also generalised anxiety.* [P4]

The AI aligning with clinical judgment was an important factor in deciding to trust the AI. Participant 1 provided commentary commensurate with having adopted a trusting attitude whilst staying alert to clinical incongruence:

> *I haven't seen anything to suggest that I would not trust it. Everything fits with my own clinical judgment from what I've seen.* [P1]

Participant 2 highlighted the role of data in trust, implying that an AI system should come to the same conclusions as a person, given the same data:

> *[I would trust the AI] To the same degree that I would trust the system we have now because it's, you know, it's the same data.* [P2]

Participant 2 decided to adopt the AI prediction as their own fairly early in the process, when they were somewhat undecided (scenario 1), showing early willingness to put faith in the system. However, in general, trust was seen to be conditional on confidence, certainty, and context. The model could be trusted in screening clear cases to go to an interview, especially for certain cases, but not for making a diagnosis. Participants expressed this as follows:

> *If it only comes to screening, then I would probably trust it, especially if it's certain. And then if it's not certain, I would maybe spend a little longer to look at it* [P4]

> *For the clear yeses that we don't have to use our time for that, but they can just go straight to our screening interviews.* [P3]

> *I think the model thinks the same way as a Screener does. But it doesn't think like a clinician does. So as a screening tool, something that might work automatically, I can see how something like this could probably be implemented like soon. As clinical tool, I wouldn't...with what I see now, trust it to the degree that it could make a diagnosis.* [P2]

Contextually, participants wanted reassurance that the AI would take into account patient risk factors before they would recommend the AI as a screening tool.

> *We cannot have someone with a high suicidal risk. So, but if I knew this prediction took that into account, then that would be OK.* [P3]

And then only to allow a patient to proceed to the interview, never to exclude a patient:

> *I think that's long as it confirms and does not reject, I think ethically I'm less hesitant...I think as soon as we start rejecting people or limiting people's access to treatment based on automatic AI responses. That would be problematic.* [P5]

Screeners routinely assess the severity of presentation (eg, severe depression), and so any decision to put patients through to interview should take account of severity, as well as prediction confidence and uncertainty: "It doesn't say the degree of the illness or treatment or disorder." (P3). In addition, some patients (eg, suicidal or severe cases) should be rejected at the interview stage, and an AI putting patients through to an unnecessary interview could be unethical. This re-emphasizes the need to include free text in the prediction, so any prediction is as complete as possible:

> *It's a waste of patient time if they specifically write about something that in their application that I, after an hour's interview with them, says we can't help you because of this. But they already informed me about it, so they wouldn't have to sit there and open up, and get a rejection if I had just...read that before.* [P4]

The issue of causability was evident when participants questioned the reasons why certain demographics were associated with prediction explanations. Even though the participants all expressed trust in the AI generally, there was a sense that trust was suspended or withheld when explanations were associative but did not lend themselves easily to a causal explanation:

> *Yeah, and I understand that it can't necessarily explain why the data says it, but the more it can explain why, at least guess why the data says, the better. And that would definitely increase my confidence as well...So, it's not random, but it doesn't tell the story why the item is important.* [P5]

Increased use of the AI tended to build trust further:

> *Yeah, but I think I trust it more because I agree with it, or it agrees with me. Yeah, I think because I first looked at it and made my mind up from very limited data, actually...But who knows if I'm wrong and this is wrong as well. But I think I definitely trust it more*

XSL•FO
RenderX

> *than now that I've interacted a bit more with it, yeah.* [P4]

Once participants had decided to trust the AI, not agreeing with the AI's prediction, after some experience with it, evoked a suspension but not a revocation of trust:

> *With the specific phobias, there was the factors against, so I might not agree, but that doesn't make me trust the model less.* [P2]

The violin plots were seen as effective in communicating predictions, but the AI outputting notifications of when to trust/when not to trust, based on the model's own uncertainty, was thought to be required for deployment.

There was a view that the AI could be more trustworthy than humans because of its systematic reasoning:

> *It concerns me when individual clinicians look at an elevation on that score and interpret it, giving it too much weight, but I don't have a fear that the algorithm is doing that.* [P1]

However, since the participants were not given data on the verifiable validity of the model as part of the guided interview process, and although trust was evident intrinsically, future use was conditional on the validity being evidenced extrinsically:

> *I don't know how reliable or valid the model is. I think it would be an improvement of the current system.* [P2]
>
> *Perhaps we will have the data to support this in the future when the appropriate studies have been completed.* [P1]

Discussion

Thematic Analysis

The cautiously positive initial attitudes toward AI set a positive tone and likely eased movement toward reliance. If the baseline attitude had been skeptical, it is possible that the same interface features might have been interpreted more critically. Concerns about algorithmic bias were muted because the model was trained on the same data the screeners themselves used, and it was viewed as representative of the patient population and therefore acceptable.

As participants worked through the interface, a 4-stage "trust journey" (Figure 4) became apparent, providing a useful support to the literature on AI trust that conceptualizes trust as a staged, context-sensitive process [19,25]. First, participants oriented themselves to the user interface layout and information; next, they made sense of the model's operation; then they appraised risk and uncertainty. Finally, they decided whether to rely on the output, under what safeguards, and in which contexts. Progress along this trust journey was shaped by 2 elements: how specific visualizations supported sense-making and risk appraisal, and how duty-of-care requirements set boundaries for reliance. It is notable that this journey appeared to emerge in the context of using the AI in a free manner to explore various patient scenarios, rather than being enforced by the interview structure.

Understanding confidence and uncertainty in the predictions was central to trust building, with the violin plots communicating both concepts intuitively. Even though the terminology surrounding these concepts was sometimes confused, the understanding was clearly demonstrated and articulated. In use, violin plots were the main aspect that participants referred to and said they would refer to in the future, given the limited time given to screening. In moving back and forth between record, prediction, and explanation, clinicians repeatedly tested the AI's reasoning against their own judgment and, in doing so, built intrinsic trust.

The dataset available from the clinic did not include the patients' free-text inputs. From the clinical safety perspective emphasized by Coder B, trust depended on narrative completeness and safeguarding. The absence of free text limited the assessment of suicidality, comorbidity, and context, which in turn constrained willingness to rely on the tool. Integrating patient free text and providing risk-related cues in summary form would extend the contexts in which reliance is acceptable. Relatedly, questions of causability surfaced when demographic variables appeared in explanation panels: correlations were accepted as "in the data," but clinicians still sought a story that linked those variables to clinical reasoning. Adopting a pedagogical/causability approach, adding ceteris paribus ("what-if") sensitivity plots to the interface (eg, partial dependence and individual conditional expectation), could be beneficial. These displays let clinicians vary a single input while holding others constant and observe the corresponding change in the model score, supporting understanding.

Once clinicians felt that the system "thought like a screener," occasional disagreement with a prediction did not revoke trust outright; rather, it triggered a momentary suspension while they compared the model's output and explanations with their own assessment. An implication is that early alignment between user judgment and model output may accelerate reliance, but it may also seed confirmation bias. A more skeptical starting stance may have led to alignment being scrutinized more stringently.

Willingness to rely on the AI remained conditional. It was strong when the model expressed high confidence with straightforward cases but tempered in the presence of uncertainty or missing contextual information. Participants indicated that the AI may expedite low-risk referrals, but they resisted allowing it to exclude patients or establish diagnoses without human oversight.

Intrinsic trust accrued through interaction with visualizations and explanations; extrinsic trust depended on evidence of validity and calibration. The reliability analysis presented earlier addresses this requirement and should sit alongside uncertainty displays and clear guardrails as deployment criteria.

The clinical perspective (coder B) clarifies 3 conditions for acceptable use: practice-legible explanations, inclusion of narrative data for risk detection, and strict boundaries where stakes or uncertainty are high.

First, explainability must map to everyday reasoning. Violin plots conveyed the prediction confidence effectively, but transparency had to translate into explanations that clinicians

XSL•FO
RenderX

could understand in relation to their practice. The graphical explanations (Figure 3) were also effective in this regard.

Second, trust was built on the belief that the AI model and the clinician relied on the same data and had the same "understanding" of it. Repeatedly, participants returned to the absence of free-text in the patient data, believing that short-form questionnaires alone were insufficient for capturing detail, suicidality, or the subtleties of comorbidity. This concern underlined a priority for clinical safeguarding so that the AI builds trust, rather than eroding it by ignoring critical information.

Third, automation boundaries should reflect accountability. The model should be used to confirm low-risk referrals and escalate ambiguity, not to exclude patients or assign diagnoses without human oversight.

Taken together, the clinical perspective aligns with the main themes that depict a pragmatic balancing of promised efficiency against the need for interpretability, contextual trust, and ethically defensible deployment. Only when these conditions are met can the AI become dependable in the screening process. In settings where clinicians are skeptical about AI, explanations may need to be deeper and more informative, uncertainty cues may need to be more explicit, and evidence of external validation may need to be available up front, to avoid premature dismissal or resistance.

Overall, the findings portray trust as contextual and conditional. Clinicians recognize the efficiency and consistency that AI can bring to mental health screening, but require interpretability, narrative completeness, and contextual safeguards before embracing the tool fully. Trust, in this setting, is a negotiated, context-sensitive stance that evolves as users see the model's reasoning align, or fail to align, with their own.

Principal Findings

This qualitative study offers several insights into how clinicians in a web-based mental health clinic in Southern Denmark, with national coverage, formed trust in an AI-mental health model and HCI.

The participants' initial attitude to AI in mental health was framed by their positive interactions in the past with AI (eg, generative AI), and their view of the possibilities for AI to improve productivity or reduce the burden of low-expertise tasks.

Framed by the anticipated utility of AI, trust developed across a further 3 sequential themes that outlined a notional "trust journey": sense-making, risk appraisal, and a conditional decision to rely. This reflects the literature on AI trust: moving between (1) building intrinsic trust by understanding the model contextually in relation to potential use cases and user experience; (2) deciding the model can fulfill the trust contract; (3) extrinsic trust based on the functional performance. This "trust journey" provides a useful support to the literature on AI trust that conceptualizes trust as a staged, context-sensitive process.

The explainability of the AI was of particular importance in trust building. This involved several elements. First, displaying prediction confidence and uncertainty in violin plots effectively conveyed the ambiguity inherent in differential clinical decision-making. This was found to be of critical importance in building intrinsic trust, underscoring the role of interpretability in building trust in relation to the expert users' expectations. Second, the presentation of key features influencing the predicted class—including those supporting and opposing the outcome—allowed users to see that the model reasoned in a manner similar to a clinical screener. This underscores the role of interpretability in allowing an understanding to form, that the model "thought" like a clinical screener. Third, the PriSM model's feature structure, incorporating pseudo-sumscores, improved interpretability by presenting explanations in familiar terms.

Trust was contextually confined to low-risk areas of clinical practice, such as screening patients prior to interview, and with a guarantee that safety protocols would be incorporated (eg, suicidality indicator checks). Incorporating free-text patient data, which clinicians typically rely on in screening, in addition to tabular data, would increase the contextual area of trust because the model would then have exactly the same data available as the screener.

Limitations

This study offers a focused and detailed, but necessarily bounded view of clinician trust in an AI model and HCI for mental health treatment prediction. Several constraints should be acknowledged.

The participant pool was deliberately small and clinic-specific: 5 screeners from a single web-based mental health clinic in Southern Denmark, providing national coverage. While this concentration allowed for deep, reflexive engagement with the clinic context—and included the majority of its screeners—it also limits transferability, as attitudes, workflows, and organizational norms may differ in other clinical settings or professional cultures.

The encounters took place with a high-fidelity prototype, not a fully deployed system. Although the interface replicated key functions, it operated on prepared case scenarios under interview conditions. Real-world contingencies—time scarcity, partial records, competing clinical duties—were only indirectly represented. The trust trajectory observed here may evolve when the tool is embedded in day-to-day use.

The think-aloud, semistructured protocol channeled attention toward predefined topics. While this ensured coverage, it may have shaped the sequence of concerns, potentially different from those that would surface in naturalistic use. In particular, this may have influenced the notion of a trust journey that surfaced in the analysis. However, it is believed that the protocol reflected a natural workflow wherein clinicians examine the patient record and assess the treatment required. The inclusion of the AI in this workflow seemed to flow naturally during the interviews, and this gives confidence that the semistructured protocol did not overly influence the results.

The available dataset and prototype excluded free-text input, a feature that clinicians regarded as important for detecting suicidality, subtle comorbidity, and patient voice. As a result,

XSL•FO
RenderX

participants evaluated a system that they acknowledged as limited; their conditional trust therefore reflects both the strengths of transparent visualizations and explainability and the recognized absence of narrative data.

No quantitative performance metrics, such as accuracy, calibration, or validation results, were provided during the sessions. Participants formed judgments on the basis of interactive experience alone. The trust documented here thus stands without the extrinsic evidence being available to the participants.

Finally, the thematic account is explicitly reflexive. This approach, building on the expertise of 2 complementary perspectives, prioritizes interpretive richness but also means that alternative analytic lenses might organize the material differently. Because the 2 analysts conducted separate reflexive analyses, some interpretive tensions may remain; we view these as theoretically informative.

Future Work

Future work should address the concerns of participants in evaluating the AI model's conditional domain of trust. In the contractual view of trust [19], an AI model should "know what it doesn't know" and communicate this distinction to the user. The importance of known model uncertainty can be appreciated when the data distribution shifts from the statistical distributions of the training data. During these data distribution shifts, the model's predictions should be more uncertain. Thus, evaluating confidence-based metrics (eg, negative log likelihood, Brier score, static calibration error) in this scenario is an important next step in answering whether the model's uncertainty estimates are responsive to a shift even when its class predictions are unaffected, which is a desirable property in high-stakes or safety-critical applications where knowing when not to trust a prediction is as important as the prediction itself.

Once trust has been built, participants expressed a desire that less interaction with the AI would be required, so it could fulfill the participants' hopes for AI to save time. To fulfill this, we require a method of expressing confidence and uncertainty that does not rely upon inspecting violin plots. To this end, approaches such as conformal prediction should be explored as a principled method to express these factors, which are important for the context of trust. In conformal prediction, conformal sets contain the predicted class labels with a mathematical guarantee such that sets that contain more than one label, for example,{depression, panic}, can be interpreted as being uncertain [46]. Accordingly, the size of the prediction set can drive a traffic-light trust indicator: singleton (green), multilabel (amber), and abstention/empty (red).

The integration of free text in the dataset and as a model feature will be another important next step. Architecting a hybrid model that associates the text with the existing model's tabular and demographic data would bring closure to the issue of detecting suicidality, subtle comorbidity, and patient voice in the patient record. In addition, the severity of conditions expressed in text may be taken into account. We plan to address this limitation by seeking to add the free-text data to the dataset and building a hybrid model, bridging the text and tabular data into a common feature space for model processing.

Conclusions

This study supports the literature on AI trust by providing insights into how clinicians form trust in AI tools within the context of web-based mental health care. Trust was not given unconditionally but developed over time through a process of understanding, risk evaluation, and conditional reliance. This journey mirrors the AI trust literature in which AI users approach AI through trust heuristics [25] based on the alignment between the perceived explanations and their mental models [24]; ie, causability. Graphical explanations, especially violin-plot probability distributions, were readily interpreted and became the primary vehicle for intrinsic trust.

The explainability of the AI system played a critical role in this trust journey, particularly when interpretability features matched clinicians' cognitive models and expectations. Trust was seen as contextual [20]; strongest in low-risk, well-defined contexts, and was contingent on safeguards and transparency.

As such, trust was constrained by three factors with varying emphasis: (1) narrative data completeness (the absence of free-text limited perceived safety), (2) causability (clinicians would like an explanatory story, not correlations alone, but were willing to trust the data); (3) external evidence (formal validation metrics were still expected before full explicit trust would be realized). However, once the contextual risk was appraised, using the AI in the areas of trust was evident, leading to explicit trust building in the screening domain.

For AI systems to gain broader acceptance in mental health practice, they must not only be accurate but also interpretable, context-aware, and subject to ongoing evaluation. Expanding the types of patient data the AI can process—especially free-text notes—could extend the boundaries of trust and enhance clinical adoption.

Acknowledgments

The authors wish to thank the clinical staff, administrative staff, and patients at the Center for Digital Psychiatry, Denmark.

Funding

This work has been supported by the Innovation Fund, Denmark.

XSL•FO
RenderX

Authors' Contributions

AK designed the artificial intelligence model, developed the human-computer interface, contributed to the methodology and interview guide, conducted the interviews and thematic analysis, and drafted the manuscript. ASN developed the methodology, interview guide, and reviewed the manuscript. TTHSM conducted the thematic analysis and developed the patient profiles. NB contributed to the literature review and interview transcription. NB, PVDV, RM, JEW, KM, and ASN reviewed the manuscript and provided valuable input on various sections. PVDV contributed to the artificial intelligence model development. KM provided support for funding and data collection.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study interview guide.
[DOCX File, 404 KB - humanfactors_v12i1e79658_app1.docx]

References

1. Arias D, Saxena S, Verguet S. Quantifying the global burden of mental disorders and their economic value. EClinicalMedicine 2022 Dec;54:101675. [doi: 10.1016/j.eclinm.2022.101675] [Medline: 36193171]
2. Olawade DB, Wada OZ, Odetayo A, David-Olawade AC, Asaolu F, Eberhardt J. Enhancing mental health with artificial intelligence: current trends and future prospects. Journal of Medicine, Surgery, and Public Health 2024 Aug;3:100099. [doi: 10.1016/j.glmedi.2024.100099]
3. D'Alfonso S. AI in mental health. Curr Opin Psychol 2020 Dec;36:112-117. [doi: 10.1016/j.copsyc.2020.04.005] [Medline: 32604065]
4. Choung H, David P, Ross A. Trust in AI and its role in the acceptance of AI technologies. International Journal of Human–Computer Interaction 2023 May 28;39(9):1727-1739. [doi: 10.1080/10447318.2022.2050543]
5. Higgins O, Short BL, Chalup SK, Wilson RL. Artificial intelligence (AI) and machine learning (ML) based decision support systems in mental health: an integrative review. Int J Ment Health Nurs 2023 Aug;32(4):966-978. [doi: 10.1111/inm.13114] [Medline: 36744684]
6. Antoniou G, Papadakis E, Baryannis G. Mental health diagnosis: a case for explainable artificial intelligence. Int J Artif Intell Tools 2022 May;31(03). [doi: 10.1142/S0218213022410032]
7. Hassan M, Kushniruk A, Borycki E. Barriers to and facilitators of artificial intelligence adoption in health care: scoping review. JMIR Hum Factors 2024 Aug 29;11:e48633. [doi: 10.2196/48633] [Medline: 39207831]
8. Hopkin G, Branson R, Campbell P, et al. Building robust, proportionate, and timely approaches to regulation and evaluation of digital mental health technologies. Lancet Digit Health 2025 Jan;7(1):e89-e93. [doi: 10.1016/S2589-7500(24)00215-2] [Medline: 39550311]
9. Baan J, Fernández R, Plank B, Aziz W. Interpreting predictive probabilities: model confidence or human label variation? Presented at: Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Volume 2: Short Papers); Mar 2024; St Julian's, Malta p. 268-277. [doi: 10.18653/v1/2024.eacl-short.24]
10. Kompa B, Snoek J, Beam AL. Second opinion needed: communicating uncertainty in medical machine learning. NPJ Digit Med 2021 Jan 5;4(1):4. [doi: 10.1038/s41746-020-00367-3] [Medline: 33402680]
11. Malinin A, Gales M. Predictive uncertainty estimation via prior networks. Presented at: NIPS'18: Proceedings of the 32nd International Conference on Neural Information Processing Systems; Dec 3-8, 2018; Montréal, QC p. 7047-7058. [doi: 10.5555/3327757.3327808]
12. Begoli E, Bhattacharya T, Kusnezov D. The need for uncertainty quantification in machine-assisted medical decision making. Nat Mach Intell 2019;1(1):20-23. [doi: 10.1038/s42256-018-0004-1]
13. Zhu F, Zhang XY, Cheng Z, Liu CL. Revisiting confidence estimation: towards reliable failure prediction. IEEE Trans Pattern Anal Mach Intell 2024 May;46(5):3370-3387. [doi: 10.1109/TPAMI.2023.3342285] [Medline: 38090830]
14. Asan O, Bayrak AE, Choudhury A. Artificial intelligence and human trust in healthcare: focus on clinicians. J Med Internet Res 2020;22(6):e15154. [doi: 10.2196/15154]
15. Rosenbacke R, Melhus Å, McKee M, Stuckler D. How explainable artificial intelligence can increase or decrease clinicians' trust in AI applications in health care: systematic review. JMIR AI 2024 Oct 30;3:e53207. [doi: 10.2196/53207] [Medline: 39476365]
16. Wysocki O, Davies JK, Vigo M, et al. Assessing the communication gap between AI models and healthcare professionals: explainability, utility and trust in AI-driven clinical decision-making. Artif Intell 2023 Mar;316:103839. [doi: 10.1016/j.artint.2022.103839]
17. Bernardo V. EDPS TechDispatch on explainable artificial intelligence. Eur Data Prot Superv 2023. [doi: 10.2804/132319]

XSL•FO
RenderX

18. Joyce DW, Kormilitzin A, Smith KA, Cipriani A. Explainable artificial intelligence for mental health through transparency and interpretability for understandability. NPJ Digit Med 2023 Jan 18;6(1):6. [doi: 10.1038/s41746-023-00751-9] [Medline: 36653524]
19. Jacovi A, Marasović A, Miller T, Goldberg Y. Formalizing trust in artificial intelligence: prerequisites, causes and goals of human trust in AI. Presented at: FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency; Mar 3-10, 2021; Virtual Event Canada p. 624-635. [doi: 10.1145/3442188.3445923]
20. Lipton ZC. The mythos of model interpretability. Queue 2018 Jun;16(3):31-57. [doi: 10.1145/3236386.3241340]
21. Barredo Arrieta A, Díaz-Rodríguez N, Del Ser J, et al. Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion 2020 Jun;58:82-115. [doi: 10.1016/j.inffus.2019.12.012]
22. Ali S, Abuhmed T, El-Sappagh S, et al. Explainable artificial intelligence (XAI): what we know and what is left to attain trustworthy artificial intelligence. Information Fusion 2023 Nov;99:101805. [doi: 10.1016/j.inffus.2023.101805]
23. Li H, Zhang R, Lee YC, Kraut RE, Mohr DC. Systematic review and meta-analysis of AI-based conversational agents for promoting mental health and well-being. NPJ Digit Med 2023 Dec 19;6(1):236. [doi: 10.1038/s41746-023-00979-5] [Medline: 38114588]
24. Holzinger A, Carrington A, Müller H. Measuring the quality of explanations: the system causability scale (SCS). Künstl Intell 2020 Jun;34(2):193-198. [doi: 10.1007/s13218-020-00636-z]
25. Shin D. The effects of explainability and causability on perception, trust, and acceptance: implications for explainable AI. Int J Hum Comput Stud 2021 Feb;146:102551. [doi: 10.1016/j.ijhcs.2020.102551]
26. Belsher BE, Smolenski DJ, Pruitt LD, et al. Prediction models for suicide attempts and deaths: a systematic review and simulation. JAMA Psychiatry 2019 Jun 1;76(6):642-651. [doi: 10.1001/jamapsychiatry.2019.0174] [Medline: 30865249]
27. Pisani AR, Murrie DC, Silverman MM. Reformulating suicide risk formulation: from prediction to prevention. Acad Psychiatry 2016 Aug;40(4):623-629. [doi: 10.1007/s40596-015-0434-6] [Medline: 26667005]
28. Staying safe from suicide. NHS England. URL: https://www.england.nhs.uk/long-read/staying-safe-from-suicide/ [accessed 2025-06-25]
29. Mathiasen K, Riper H, Andersen TE, Roessler KK. Guided internet-based cognitive behavioral therapy for adult depression and anxiety in routine secondary care: observational study. J Med Internet Res 2018 Nov 28;20(11):e10927. [doi: 10.2196/10927] [Medline: 30487118]
30. Tarp K, Nielsen SL, Holmberg TT, et al. Therapist perceptions of the implementation of a new screening procedure using the ItFits-toolkit in an iCBT routine care clinic: a mixed-methods study using the consolidated framework for implementation research. Front Psychiatry 2023;14:1104301. [doi: 10.3389/fpsyt.2023.1104301] [Medline: 37091699]
31. Kelly A, Jensen EK, Grua EM, Mathiasen K, Van de Ven P. An interpretable model with probabilistic integrated scoring for mental health treatment prediction: design study. JMIR Med Inform 2025 Mar 26;13(1):e64617. [doi: 10.2196/64617] [Medline: 40138679]
32. Bryman A. Social Research Methods, 5th edition: Oxford University Press; 2016. [doi: 10.4135/9781849209939]
33. Hine C. Ethnography for the Internet: Routledge; 2020. [doi: 10.4324/9781003085348-6]
34. Mahatody T, Sagar M, Kolski C. State of the art on the cognitive walkthrough method, its variants and evolutions. Int J Hum Comput Interact 2010 Jul 30;26(8):741-785. [doi: 10.1080/10447311003781409]
35. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
36. Braun V, Clarke V. One size fits all? what counts as quality practice in (reflexive) thematic analysis? Qual Res Psychol 2021 Jul 3;18(3):328-352. [doi: 10.1080/14780887.2020.1769238]
37. Yin RK. Introduction: The Case Study as a Research Strategy Case Study Res Des Methods: SAGE Publications; 1994.
38. Hick A, Ziefle M. A qualitative approach to the public perception of AI. Int J Cybern Inform 2022 Aug 27;11(4):1-17. [doi: 10.5121/ijci.2022.110401]
39. Byrne D. A worked example of Braun and Clarke's approach to reflexive thematic analysis. Qual Quant 2022 Jun;56(3):1391-1412. [doi: 10.1007/s11135-021-01182-y]
40. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9. J Gen Intern Med 2001 Sep;16(9):606-613. [doi: 10.1046/j.1525-1497.2001.016009606.x]
41. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
42. Mattick RP, Clarke JC. Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. Behav Res Ther 1998 Apr;36(4):455-470. [doi: 10.1016/s0005-7967(97)10031-6] [Medline: 9670605]
43. Houck PR, Spiegel DA, Shear MK, Rucci P. Reliability of the self-report version of the panic disorder severity scale. Depress Anxiety 2002;15(4):183-185. [doi: 10.1002/da.10049] [Medline: 12112724]
44. Marks IM, Mathews AM. Brief standard self-rating for phobic patients. Behav Res Ther 1979;17(3):263-267. [doi: 10.1016/0005-7967(79)90041-x] [Medline: 526242]
45. Lewis JR. Sample sizes for usability studies: additional considerations. Hum Factors 1994 Jun;36(2):368-378. [doi: 10.1177/001872089403600215] [Medline: 8070799]

XSL•FO
RenderX

46. Angelopoulos AN, Bates S. A gentle introduction to conformal prediction and distribution-free uncertainty quantification. : arXiv Preprint posted online on Jul 15, 2021. [doi: 10.48550/arXiv.2107.07511]

Abbreviations

AI: artificial intelligence
HCI: human-computer interface
iCBT: internet-based cognitive behavioral therapy
PHQ-9: Patient Health Questionnaire-9
PrISM: Probabilistic Integrated Scoring Model
RTA: reflexive thematic analysis
SIAS: Social Interaction Anxiety Scale
XAI: explainable artificial intelligence

Edited by A Kushniruk; submitted 25.06.25; peer-reviewed by I Adefolaju, SK Nandipati, Z Lin; revised version received 10.09.25; accepted 20.11.25; published 19.12.25.

Please cite as:
Kelly A, Bhardwaj N, Holmberg Sainte-Marie TT, Van de Ven P, Melia R, Williams JE, Mathiasen K, Nielsen AS
Investigating How Clinicians Form Trust in an AI-Based Mental Health Model: Qualitative Case Study
JMIR Hum Factors 2025;12:e79658
URL: https://humanfactors.jmir.org/2025/1/e79658
doi:10.2196/79658

© Anthony Kelly, Niharika Bhardwaj, Trine Theresa Holmberg Sainte-Marie, Pepijn Van de Ven, Ruth Melia, John Eustis Williams, Kim Mathiasen, Amalie Søgaard Nielsen. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 19.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

TeleAllergy: Potential of Telemedicine in Management of Patients With Allergies

Hanna Lindemann[1,2], MD; Emil Hammer[1,2,3]; Luca Bonifacio[1,2,3]; Christian Greis[4*], MD; Karin Hartmann[1,2,3*], MD

[1]Division of Allergy, Department of Dermatology, University Hospital Basel and University of Basel, Burgfelderstrasse 101, Basel, Switzerland

[2]Department of Clinical Research, University Hospital Basel and University of Basel, Basel, Switzerland

[3]Department of Biomedicine, University Hospital Basel and University of Basel, Basel, Switzerland

[4]Department of Dermatology, University Hospital Zurich and University of Zurich, Zurich, Switzerland

[*]these authors contributed equally

Corresponding Author:
Hanna Lindemann, MD
Division of Allergy, Department of Dermatology, University Hospital Basel and University of Basel, Burgfelderstrasse 101, Basel, Switzerland

Abstract

Background: The growing prevalence of allergic diseases alongside a shortage of trained allergists creates significant challenges in delivering timely care, especially for underserved populations. Telemedicine presents a promising solution, offering remote care through digital tools. While telemedicine has been widely adopted in other fields, its use in allergy care remains underexplored.

Objective: This study aimed to assess the potential of telemedicine in managing allergic diseases by examining patient preferences and experiences.

Methods: A survey of 27 questions was distributed to adult patients (>18 y) with allergic diseases attending the outpatient allergy clinic at the Division of Allergy, University Hospital Basel, Basel, Switzerland, between May and August 2024. The survey covered demographic information, prior use of telemedicine, and preferences for teleconsultation modalities. It also assessed patients' willingness to share various types of clinical data, including images and written reports, and explored which allergic diseases were considered appropriate for telemedicine.

Results: A total of 102 patients participated in the survey, with a mean age of 44.4 years (SD 16.7 y). For further analysis, the patients were stratified into four age groups: 18 - 34 years (36/102), 35 - 49 years (26/102), 50 - 64 years (31/102), and ≥65 years (9/102). Among them, 44% (41/94; *P*=.22) had previously used telemedicine services, with 34% (32/94; *P*=.04) specifically using it for allergic diseases. When asked about consultation formats, 49% (49/100) of patients preferred in-person visits, while 41% (41/100) favored a hybrid model combining telemedicine and in-person care. Regarding telemedicine tools, 57% (51/89) preferred telephone consultations with a doctor. Patients would use telemedicine preferentially for mild compared to severe allergic diseases as well as for chronic compared to acute conditions. The spectrum of diseases for which patients would use telemedicine comprised a wide range of allergic conditions, with allergic rhinoconjunctivitis (16%; 14/85), Hymenoptera venom allergy (13%; 11/85), and food allergy/intolerance (13%; 11/85) cited most frequently. Only 7% (6/85) of patients indicated they would not use telemedicine for any allergic disease.

Conclusions: This study emphasizes the growing adoption and importance of telemedicine in allergy care, with a significant proportion of patients already having experience using it for managing allergic diseases. Patients' inclination toward multiple communication formats underscores the growing need for individualized management of allergic diseases.

(JMIR Hum Factors 2025;12:e75483) doi:10.2196/75483

KEYWORDS

teleallergy; telemedicine; healthcare access; allergy management; teledermatology; synchronous communication; asynchronous communication; digital health; allergy diagnosis; teleconsultation; healthcare technology; patient-centered care

Introduction

Challenges in Allergy Care

The increasing prevalence of allergic diseases, combined with a persistent shortage of trained allergists, presents significant challenges in delivering timely and accessible allergy care. This issue is especially pronounced in underserved rural areas and among vulnerable populations, such as older adults and those with mobility impairments [1,2].

Allergic diseases encompass a wide spectrum of conditions, from common diseases such as allergic rhinitis, eczema, and food allergy, to rare and frequently severe disorders such as

XSL•FO
RenderX

mastocytosis and hereditary angioedema. Despite their varying severity, these diseases share the need for specialized expertise and ongoing, long-term management [1,2].

Potential of Telemedicine

Telemedicine, particularly during the COVID-19 pandemic, has emerged as a promising and cost-effective solution to address these challenges, utilizing digital technologies to connect patients with health care providers. It offers two primary modes of communication: synchronous, where patients and providers interact in real-time (such as video consultations) and asynchronous, which allows communication at different times. A widely used asynchronous method is store-and-forward technology, where clinical data, such as images or medical records, are uploaded in a digital patient folder for later review. This approach has been particularly effective in teledermatology, where visual assessments are crucial [3-5].

In addition to communication tools, various digital applications and technologies are available to support health care management, such as e-diaries ("digital companion") that track symptoms and therapy adherence, as well as electronic prescriptions and appointment scheduling systems that streamline health care processes [6].

While telemedicine is well-established in many medical specialties, there is limited data on its application for patients with allergies. In this study, we employed a patient-centered questionnaire to evaluate the current use of telemedicine in allergy care, to identify patients' preferred communication tools and the specific allergic conditions for which they would be open to using telemedicine.

Objective

This study aimed to assess telemedicine's role in allergy care by evaluating patient preferences, usage patterns, and conditions considered suitable for remote management.

Methods

Study Design and Participants

To evaluate patient experiences and preferences with telemedicine, we conducted a survey of 27 questions targeting adult patients (over 18 y) with various allergic diseases prior to their consultations at our outpatient allergy clinic. Exclusion criteria included language barriers, cognitive impairment, and lack of informed consent.

Inclusion and Exclusion Criteria

The questionnaire, developed in German, French, Italian, and English, collected demographic data, educational background, travel time to the nearest allergist, prior use of telemedicine services, the extent of health insurance coverage for telemedicine, and patients' preferences regarding teleconsultation modalities. It also explored which digital communication tools and clinical data types (eg, images or medical reports) patients were open to sharing. Furthermore, the survey assessed which allergic diseases (categorized as mild or severe and acute or chronic based on patients' subjective evaluation of their symptoms) patients considered appropriate for telemedicine. Questions were formatted as binary responses, multiple-choice items, visual analog scales (0 - 10), and open-ended text fields to capture comprehensive insights.

Statistical Analysis

To analyze the data collected through the questionnaire across different demographic categories, the 102 patients were categorized into four age groups: 18 - 34 years, 35 - 49 years, 50 - 64 years, and over 65 years. For continuous variables such as age, the mean and standard deviation were calculated, while categorical variables are presented as absolute frequencies (n) and percentages (%). Group differences in telemedicine usage and preferences were evaluated using χ^2 tests, with Bonferroni correction applied to adjust the family-wise error rate in multiple pairwise comparisons. Statistical significance was defined as a two-sided P value <.05. All analyses were conducted using R software (version 4.2.0; R Foundation for Statistical Computing).

Ethical Considerations

The study was reviewed by the Ethics Committee of Northwest and Central Switzerland (EKNZ). According to their assessment (Ref: Req-2024 - 00650), the project does not fall under the Swiss Human Research Act, as only anonymized data were processed, and therefore no formal approval was required.

Results

Demographics

Between May 29, 2024, and August 8, 2024, a total of 102 patients completed the questionnaire (mean age 44.4, SD 16.7 y). Age distribution was as follows: 35% participants (36/102) were 18 - 34 years old, 25% (26/102) were 35 - 49 years, 30% (31/102) were 50 - 64 years, and 9% (9/102) were ≥65 years. 61% were women (62/102). Most of the respondents lived in suburban areas with populations between 10,000 and 1,00,000 (37/101; 36%) or in rural areas with fewer than 10,000 inhabitants (36/101; 35%). Approximately, 57% participants (58/101) reported a travel time of 10 to 29 minutes to reach an allergist, and 49% (50/102) held a university or college degree (Table S1 in Multimedia Appendix 1).

Experience With Telemedicine

44% of the respondents (41/94; P=.22) reported previous use of telemedicine services. Prior telemedicine use was significantly associated with age group (P<.001), as was prior use of tele-allergy services (P<.001). The highest rate of telemedicine use was observed in the 35 - 49 age group (16/26; 62%), significantly higher than in the 50 - 64 group (P<.001) and the ≥65 group (P=.04) (Figure 1A). 39% patients (37/94; P=.04) had used telemedicine specifically for allergy-related care (Figure 1B). Prior tele-allergy use was significantly associated with age group (P<.001), and was most frequent in the 50 - 64 age group (14/30; 47%), significantly more than in the 18 - 34 group and the ≥65 group (both P<.001).

XSL•FO
RenderX

Figure 1. Patients' experience and preferences in telemedicine. (A) Patients (%) with previous experience in telemedicine. (B) Patients (%) with previous experience in telemedicine consultations for allergic conditions. (C) Patients (%) with insurance coverage for telemedicine. (D) Patients (%) by preferred consultation method. Levels of statistical significance are indicated as follows: * P< .05; ** P< .01; *** P< .001.

XSL•FO
RenderX

Awareness of Insurance Coverage

Uncertainty regarding insurance coverage for telemedicine was high, with 71% (73/102) unsure about reimbursement. This proportion was significantly greater than among those who responded with "yes" or "no" (P<.001). Awareness differed by age group (P<.001), with only 12% (3/26) of the 35 - 49 age group reporting uncertainty compared to all other age groups (P<.001) (Figure 1C).

Preferences for Consultation Formats

49% respondents (49/100) preferred exclusively in-person consultations (P=.04), whereas 41% (41/100) favored a hybrid model (P=.07) compared to telemedicine. Age-related differences in preference were significant (P<.001), with the ≥65 group being significantly less inclined to use telemedicine than all younger groups (P<.001) (Figure 1D).

Preferred Telemedicine Communication Tools and Consultation Formats

A significantly greater number of patients (70/92; 76%; P<.001) expressed interest in sharing photos and images with health care providers, whereas 61% patients (56/92; P=.01) preferred sharing brief written reports (<5 pages) (Table S2 in Multimedia Appendix 1).

A statistically significant association was observed between age and the preference for sharing photos or brief written reports (P<.001). All younger age groups showed a significantly higher preference for these communication methods (P<.001) compared to participants aged ≥ 65 years (data not shown).

The most favored telemedicine communication formats included phone calls with doctors (51/89; 57%; P=.02). Asynchronous modes such as email/chat with trained staff (27/89; 30%; P=.002) or physicians (24/89; 27%; P=.0014), video consultations with trained staff (33/89; 37%, P=.03), and digital patient information like instruction videos (27/89; 30%; P=.002) and symptom diaries (24/89; 27%; P=.0014) were significantly less preferred. Of these, preferences for phone calls with doctors differed significantly by age (P<.001). Only 4% (2/89) of participants aged ≥65 endorsed phone calls this consultation format, compared to those aged 18 - 34 (P<.001) (data not shown).

Allergic Conditions in Telemedicine

Telemedicine was considered especially helpful for managing mild allergic conditions (70/100; 70%; P=.007) and chronic conditions (48/100; 48%; P=.84), while considerably fewer participants supported its use for severe (11/100; 11%; P<.001) or acute conditions (27/100; 27%; P<.001).

Participants aged ≥65 were significantly less likely to consider telemedicine appropriate for mild allergic conditions compared to all younger age groups (Figure 2A).

For severe allergies, individuals aged 35 - 49 reported a significantly higher willingness to use telemedicine compared to all other age groups (all P<.001) (Figure 2B).

For acute conditions, adults aged ≥65 again reported a significantly lower willingness to use telemedicine compared to all other age groups (all P<.001) (Figure 2C).

Regarding chronic diseases, patients aged between 18 - 34 years, reported a significantly higher willingness to use telemedicine compared to individuals aged 50 - 64 and ≥65 (both P<.001) (Figure 2D).

XSL•FO
RenderX

Figure 2. Patients' willingness to use telemedicine based on the type of allergic disease. (A) Patients (%) willing to use telemedicine for mild allergic diseases. (B) Patients (%) willing to use telemedicine for severe allergic diseases. (C) Patients (%) willing to use telemedicine for acute allergic diseases. (D) Patients (%) willing to use telemedicine for chronic allergic diseases. Levels of statistical significance are indicated as follows: *P< .05; **P< .01; ***P< .001.

XSL•FO
RenderX

The spectrum of allergic diseases for which patients would use telemedicine encompassed allergic rhinoconjunctivitis (16.47%; 14/85), Hymenoptera venom allergy (12.94%; 11/85), and food allergy/intolerances (12.94%; 11/85) as most frequent diseases, followed by various less frequent allergic conditions (Multimedia Appendix 2).

Only a minority of patients (7.06%; 6/85) indicated they would not use telemedicine for any allergic disease (data not shown).

Discussion

Principal Findings

Our study demonstrates the increasing role of telemedicine in allergy care. We found high adoption rates of telemedicine, especially among younger patients, and a preference for hybrid models that combine telemedicine with in-person consultations. Telemedicine in allergy care is seen by patients as especially useful for managing mild conditions.

Demographics

Demographic data on age and gender in allergy patient surveys are limited, but our findings align with previous telemedicine studies, showing a mean age of 44.4 years and a higher participation rate among women [3]. The proportion of urban residents in our cohort was slightly lower compared to data from the Swiss Federal Statistical Office [7], likely due to the lower density of allergists in rural areas, requiring travel to urban centers for consultations [8].

Experience With Telemedicine

A high percentage of respondents reported using telemedicine services (Figure 1), consistent with findings from other medical fields [3,9]. This increased adoption of telemedicine in our cohort may partly be explained by the higher educational attainment often associated with higher usage rates [9]. Patients aged 35 - 49 years were the most frequent users of telemedicine (Figure 1), aligning with previous research. Interestingly, those aged 65 and older used telemedicine services more than those aged 50 - 65 years, suggesting an unmet need for digital health services among older, potentially less mobile patients—a trend also observed in other studies [10]. However, given the small sample of patients aged 65 and older (9 patients), this finding should be interpreted cautiously.

Although we could not determine the specific share of allergy services within telemedicine offerings, allergy services were not typically emphasized in studies assessing usage rates across specialties [11]. The high utilization rates in our study may also be influenced by selection bias, as participants were surveyed in the waiting area of the outpatient allergy clinic.

Awareness of Insurance Coverage

Despite 29 of the 37 Swiss health insurance providers offering telemedicine policies and basic insurance (OKP) covering telemedicine services when proven effective, 71% (73/102) of patients in our cohort were unaware of their coverage [12]. This suggests an unmet need for allergy services, as many patients may be paying out-of-pocket for telemedicine. On the other hand, younger patients, who often opt for insurance plans with higher deductibles, may find telemedicine more cost-effective than in-person visits. This trend is likely influenced by the Swiss health care system, which allows for flexible deductible options.

Older patients tend to prefer in-person consultations, likely due to concerns about remote care quality, limited familiarity with technology, and reliance on face-to-face interactions for building trust [13,14]. In contrast, 41% (41/100) of all patients preferred a hybrid model that combines telemedicine with in-person visits, while only 8% (8/100) favored telemedicine exclusively. This suggests that digital services are seen as complementary to, rather than a replacement for, traditional in-person care, enhancing access and alleviating pressure on health care systems [4,9]. Younger patients (18 - 34 y) preferred telephone consultations, particularly when integrated with video or in-person visits for comprehensive assessment. In contrast, patients over 50 years favored video consultations by a doctor (44%), likely due to their preference for face-to-face interactions [5,15].

Digital health tools, including online appointment scheduling and e-prescriptions, were widely embraced by patients aged 18 - 49, indicating higher comfort with technology in these age groups. However, adoption was lower among patients over 65, likely due to barriers such as unfamiliarity with digital platforms and concerns about data security [15-17].

The field of allergy care lacks substantial evidence on the most useful communication types for telemedicine consultations [5]. In our study, most patients expressed interest in uploading photos, a preference similar to findings from dermatology, where visual information is crucial for remote consultations [4]. Over half of the patients in our study preferred sharing detailed written reports, reflecting the complex and heterogeneous nature of allergic diseases, which may require more individualized data exchanges.

Allergic Conditions in Telemedicine

Patients in our survey expressed a preference for telemedicine in managing mild and chronic allergic diseases (Figure 2), aligning with broader trends in telemedicine across specialties [18]. While a small proportion of patients expressed concerns about using telemedicine for severe allergic diseases, some studies suggest that telemedicine can positively impact outcomes for diseases like anaphylaxis by providing specific management instructions [19]. However, telemedicine should not be considered a substitute for emergency care. It can play an important role in managing allergies, but cannot replace emergency responses or in-person care. Given that many allergic diseases are chronic and benefit from ongoing monitoring, telemedicine represents an effective approach for long-term management, a view that was supported by most participants in our study.

Limitations

While our study offers valuable preliminary insights into telemedicine use in allergy care, several limitations should be considered when interpreting the findings.

First, the sample size of 102 participants, though sufficient for initial exploration, limits the statistical power, particularly for

XSL•FO
RenderX

subgroup analyses by diagnosis and age groups. The small number of respondents aged 65 and older further restricts conclusions for this demographic, which is particularly relevant given their lower engagement with telemedicine.

Second, recruitment was conducted at a single center and participation was voluntary, which may have introduced self-selection bias. Individuals with a greater interest or familiarity with digital health tools may have been more likely to participate, potentially skewing the findings and limiting generalizability to broader patient populations.

Third, the study population included both new and follow-up patients. While this reflects real-world clinical diversity, differences in experience and expectations between these groups may have influenced their responses, particularly regarding telemedicine preferences and awareness of insurance coverage. To address these limitations, future studies should include larger and more diverse samples across multiple sites.

Conclusion

In conclusion, our study highlights the growing role of telemedicine in allergy care. High adoption rates, especially among younger patients, suggest that telemedicine can complement in-person care. Future retrospective research is needed to refine teleconsultation strategies, enhancing diagnostic accuracy, therapeutic effectiveness, and patient satisfaction, thereby maximizing patients' benefits from telemedicine in allergy care.

Acknowledgments

We sincerely thank all patients in this study for their time and effort in completing the questionnaire, which provided invaluable insights.

Conflicts of Interest

All authors declare that economic interests had no influence on the preparation of this manuscript. H.L. has provided a lecture and received honoraria sponsored by Allergopharma, Almirall, BioCryst, Blueprint, and Cogent, participated in a congress in collaboration with Almirall and serves as a member of the medical team at derma2go. E.H. and L.B. declare no conflict of interest. C.G. has consulted for and received honoraria from Almirall AG, Eli Lilly (Suisse) SA, Janssen-Cilag AG, L'Oréal S.A, L'Oréal Deutschland GmbH, LEO Pharma GmbH, Leo Pharmaceutical Products Sarath Ltd, MSD Merck Sharp & Dohme AG, Merz Pharma Schweiz, Novartis Pharma Schweiz AG, Viatris Pharma GmbH, Dr. August Wolff GmbH & Co. KG Arzneimittel and is the Medical Director at/ owns shares in derma2go. K.H. has consulted for and received honoraria from ALK-Abello, Allergopharma, Almirall, BioCryst, Blueprint, Cogent, Galderma, KalVista, Leo, Menarini, Novartis, Pfizer, Sanofi, Takeda, and Thermo Fisher, and owns shares in derma2go.

Multimedia Appendix 1
Distribution of patients by age group, gender, residence, distance to allergist, and education level and patients' preferences for digital communication and telemedicine consultations for allergic diseases.
[DOCX File, 23 KB - humanfactors_v12i1e75483_app1.docx]

Multimedia Appendix 2
Allergic diseases for which patients would use telemedicine.
[PNG File, 260 KB - humanfactors_v12i1e75483_app2.png]

References

1. Gutowska-Ślesik J, Samoliński B, Krzych-Fałta E. The increase in allergic conditions based on a review of literature. Postepy Dermatol Alergol 2023 Feb;40(1):1-7. [doi: 10.5114/ada.2022.119009] [Medline: 36909919]
2. Diwakar L, Cummins C, Lilford R, Roberts T. Systematic review of pathways for the delivery of allergy services. BMJ Open 2017 Feb 7;7(2):e012647. [doi: 10.1136/bmjopen-2016-012647] [Medline: 28174222]
3. Lindemann H, Bonetzki T, Frank J. Diagnostic spectrum and therapeutic efficiency in teledermatology-Results of the largest cohort study to date. J Dermatol 2023 Jun;50(6):800-803. [doi: 10.1111/1346-8138.16769] [Medline: 36890665]
4. Seuren LM, Ilomäki S, Dalmaijer E, Shaw SE, Stommel WJP. Communication in telehealth: a state-of-the-art literature review of conversation-analytic research. Res Lang Soc Interact 2024;57(1):73-90. [doi: 10.1080/08351813.2024.2305045] [Medline: 38741749]
5. Brinker TJ, Hekler A, von Kalle C, et al. Teledermatology: comparison of store-and-forward versus live interactive video conferencing. J Med Internet Res 2018 Oct 24;20(10):e11871. [doi: 10.2196/11871] [Medline: 30355564]
6. Konopik J, Blunck D. Development of an evidence-based conceptual model of the health care sector under digital transformation: integrative review. J Med Internet Res 2023 Jun 8;25:e41512. [doi: 10.2196/41512] [Medline: 37289482]
7. Die Bevölkerung der Schweiz im Jahr 2022. Bundesamt für statistik. URL: https://www.bfs.admin.ch/asset/de/28425361 [accessed 2025-03-25]

XSL•FO
RenderX

8. Arztpraxen und Grundversorgung in der Schweiz 2018–2021. Bundesamt für Statistik. 2023 Nov. URL: https://www.bfs.admin.ch/news/de/2023-0177 [accessed 2025-03-25]
9. Maul LV, Jahn AS, Pamplona GSP, et al. Acceptance of telemedicine compared to in-person consultation from the providers' and users' perspectives: multicenter, cross-sectional study in dermatology. JMIR Dermatol 2023 Aug 11;6:e45384. [doi: 10.2196/45384] [Medline: 37582265]
10. Miyawaki A, Tabuchi T, Ong MK, Tsugawa Y. Age and social disparities in the use of telemedicine during the COVID-19 pandemic in Japan: cross-sectional study. J Med Internet Res 2021 Jul 23;23(7):e27982. [doi: 10.2196/27982] [Medline: 34259641]
11. Patel SY, Mehrotra A, Huskamp HA, Uscher-Pines L, Ganguli I, Barnett ML. Variation In telemedicine use and outpatient care during the COVID-19 pandemic in the United States. Health Aff (Millwood) 2021 Feb;40(2):349-358. [doi: 10.1377/hlthaff.2020.01786] [Medline: 33523745]
12. Vergütung von digitalen Gesundheitsanwendungen im Rahmen der OKP. Bundesamt für gesundheit. 2024. URL: https://www.bag.admin.ch/dam/de/sd-web/m9zwn3pfzBfc/faktenblatt-digtale-anwendungen-vers.11.2024.pdf [accessed 2025-03-25]
13. Bhatia R, Gilliam E, Aliberti G, et al. Older adults' perspectives on primary care telemedicine during the COVID-19 pandemic. J American Geriatrics Society 2022 Dec;70(12):3480-3492 [FREE Full text] [doi: 10.1111/jgs.18035] [Medline: 36350837]
14. Savira F, Gupta A, Gilbert C, et al. Virtual care initiatives for older adults in Australia: scoping review. J Med Internet Res 2023 Jan 18;25:e38081. [doi: 10.2196/38081] [Medline: 36700113]
15. Smolinska S, Popescu FD, Izquierdo E, et al. Telemedicine with special focus on allergic diseases and asthma-Status 2022: An EAACI position paper. Allergy 2024 Apr;79(4):777-792. [doi: 10.1111/all.15964] [Medline: 38041429]
16. Odebunmi OO, Hughes TD, Waters AR, et al. Findings from a national survey of older US adults on patient willingness to use telehealth services: cross-sectional survey. J Med Internet Res 2024 May 23;26:e50205. [doi: 10.2196/50205] [Medline: 38780994]
17. Sten-Gahmberg S, Pedersen K, Harsheim IG, Løyland HI, Abelsen B. Experiences with telemedicine-based follow-up of chronic conditions: the views of patients and health personnel enrolled in a pragmatic randomized controlled trial. BMC Health Serv Res 2024 Mar;24(1). [doi: 10.1186/s12913-024-10732-7] [Medline: 38511245]
18. Greis C, Meier Zürcher C, Djamei V, Moser A, Lautenschlager S, Navarini AA. Unmet digital health service needs in dermatology patients. J Dermatolog Treat 2018 Nov;29(7):643-647. [doi: 10.1080/09546634.2018.1441488] [Medline: 29455570]
19. Anto A, Sousa-Pinto B, Bousquet J. Anaphylaxis and digital medicine. Curr Opin Allergy Clin Immunol 2021 Oct 1;21(5):448-454. [doi: 10.1097/ACI.0000000000000764] [Medline: 34292178]

Edited by J Waterson; submitted 04.04.25; peer-reviewed by MJ Goikoetxea Lapresa, V Kvedariene; revised version received 30.07.25; accepted 25.09.25; published 06.11.25.

Please cite as:
Lindemann H, Hammer E, Bonifacio L, Greis C, Hartmann K
TeleAllergy: Potential of Telemedicine in Management of Patients With Allergies
JMIR Hum Factors 2025;12:e75483
URL: https://humanfactors.jmir.org/2025/1/e75483
doi:10.2196/75483

© Hanna Lindemann, Emil Hammer, Luca Bonifacio, Christian Greis, Karin Hartmann. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 6.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring Young Adults' Attitudes Toward AI-Driven mHealth Apps: Qualitative Study

Ali Aboueldahab, MSc; Gabriele Damaschi, MSc; Marco D'Addario, PhD; Patrizia Steca, PhD

Department of Psychology, University of Milano-Bicocca, Piazza dell'Ateneo Nuovo, 1, Milano, Italy

Corresponding Author:
Ali Aboueldahab, MSc
Department of Psychology, University of Milano-Bicocca, Piazza dell'Ateneo Nuovo, 1, Milano, Italy

Abstract

Background: Artificial intelligence (AI)–driven mobile health (mHealth) apps are emerging as a promising tool for health management, yet little is known about users' psychological perceptions and attitudes toward these technologies. Understanding these aspects is crucial for both the appropriate design and the effective use of these technologies, ensuring the psychological and physical well-being of potential end users.

Objective: This study aimed to investigate the attitudes and perceptions of young adults toward a possible use of AI-driven mHealth apps, focusing on the perceived benefits and potential concerns related to their future adoption.

Methods: A qualitative focus group methodology was used. Fifteen participants (12 men, 3 women; mean age 27 years, range: 25 - 34 years) were recruited. Data were analyzed using thematic analysis to identify key themes influencing engagement with these technologies.

Results: Four main themes emerged: "Usability," which emphasized the importance of user-friendly, personalized experiences; "Innovation and Reliability," where participants expressed both enthusiasm and skepticism towards AI's potential; "Affectivity and Interaction with AI," highlighting mixed opinions on the emotional impact of AI interactions; and "Perceived Risks," which focused on concerns regarding data privacy and the need for human supervision. These factors contributed to ambivalent attitudes toward AI-driven mHealth apps, with some participants being open to adoption, while others remained cautious.

Conclusions: To foster greater engagement with AI-driven mHealth apps, developers should prioritize usability, trust, emotional support, and privacy issues, considering users' psychological needs and expectations. The findings offer valuable insights for designing more user-oriented mHealth solutions. Further research should explore how perceptions evolve with direct experience and long-term use.

(JMIR Hum Factors 2025;12:e76075) doi:10.2196/76075

KEYWORDS

AI-driven mHealth apps; user attitudes; thematic analysis; perceived risks; usability; innovation and reliability; emotional interaction; privacy concerns; mobile health; artificial intelligence

Introduction

In recent years, the rapid advancement of artificial intelligence (AI) has had a profound impact on various sectors of society, shaping industries, decision-making processes, and everyday experiences [1-3]. One of the most significant developments has been the rise of generative AI models, such as ChatGPT, Midjourney, and Gemini, which have gained widespread attention for their ability to generate human-like text, images, and other creative outputs. This rapid proliferation of AI-driven tools has sparked discussions about their integration into multiple aspects of daily life, from education and entertainment to professional services and health care [4].

Among these domains, health care has emerged as a particularly critical area for AI innovation, with growing interest in leveraging AI technologies to improve medical diagnostics, treatment recommendations, and patient support systems [5-7]. A specific and increasingly relevant subset of AI applications within health care is mobile health (mHealth) [8]—digital platforms designed to support health-related services and patient engagement via smartphones and other mobile devices. mHealth apps have gained popularity due to their accessibility and convenience, enabling users to track their health metrics, receive medical advice, and engage with health care providers remotely [9].

The integration of AI into mHealth apps is gaining momentum as developers seek to enhance functionalities through AI-powered tools, such as chatbots for patient interaction, symptom checkers, and personalized health recommendations [10,11]. In this context, AI-driven mHealth apps—defined as mHealth platforms that integrate AI-based tools to process health data, provide interactive support, and deliver tailored

XSL•FO
RenderX

recommendations [8]— are emerging as a new frontier in digital health care. However, given the current state of development, only a small number of users have had a direct experience with these beta versions that incorporate AI-driven features. This limited exposure reflects the early stages of these apps and the challenges in widespread implementation. Despite this momentum, the introduction of AI in such a sensitive domain as personal health management raises several concerns and uncertainties [12,13]. Unlike traditional health care solutions, AI-driven mHealth apps rely on machine learning algorithms, raising concerns about privacy, security, transparency, and potential biases. While these technologies offer promising advancements, current discussions in the literature and public discourse reveal a landscape marked by both enthusiasm for innovation and apprehension regarding ethical, technical, and social implications [14,15].

One of the key gaps in the literature concerns the public perception toward the integration of AI within mHealth apps. While previous studies have explored the technical feasibility and the general acceptance of AI in health care settings, including hospitals and clinical diagnostics, there is limited research focusing specifically on mHealth apps incorporating AI-powered tools—such as interactive chatbots handling personal data and health-related information—and the way everyday users perceive their benefits and risks [16-18]. Given that these apps are designed to operate in highly personal and private contexts—potentially serving as a first point of contact for health-related inquiries—it is crucial to explore what expectations, concerns, and attitudes individuals hold toward them, even in the absence of direct experience or exposure to these tools, and what factors influence their acceptance or skepticism [19]. Unlike AI adoption in traditional medical environments, where health care professionals act as intermediaries, mHealth AI tools are envisioned to function autonomously, making user trust and acceptance even more critical key factors to investigate [20].

Existing research on AI acceptance has examined factors like trust, perceived usefulness, and data privacy concerns [12,14,21], but studies specifically addressing these aspects within mHealth apps remain limited. Given the relatively early stage of AI implementation in mHealth, understanding public attitudes is crucial for shaping ethical, effective, and user-centered AI applications in health care. Issues related to data privacy, trust in AI-generated recommendations [12], potential biases in AI algorithms [19], and the perceived reliability of AI-driven health tools remain underexplored in relation to users' willingness to adopt such innovations [6].

Given the novelty of AI integration in mHealth, this study specifically examines young adults (25‐35 years old) as a strategically relevant demographic. This cohort represents both the primary users of digital health technologies [22-24] and digital natives who are more likely to engage with emerging AI interfaces compared with older aged groups [25,26]. Focusing on this population helps minimize age-related technology adoption biases that might confound perceptions in older groups, while providing insights from individuals most likely to shape early adoption patterns.

Against this backdrop, this study aims to explore the key factors shaping individuals' attitudes toward the integration of AI tools within mHealth. Specifically, it investigates perceived risks, doubts, and concerns, as well as potential benefits and innovations that users associate with AI-powered health technologies.

The remainder of this article is structured as follows: the next section outlines the materials and methods used in this study, detailing the research design, data collection process, and analytical framework. This is followed by a comprehensive presentation of the results, highlighting key findings related to public perception, concerns, and expectations. The discussion section then interprets these findings in the context of existing literature, drawing comparisons with previous studies and identifying potential implications for developers, health care professionals, and policymakers. Finally, the conclusion summarizes the primary insights derived from this research and outlines possible directions for future studies in this rapidly evolving field.

Methods

Study Design

The study used a qualitative research approach based on focus group discussions. Two focus groups were conducted in January 2025 to explore participants' perspectives, experiences, and opinions on the topic. Focus groups were chosen as they foster discussion, allowing for the emergence of insights that might not surface in individual interviews [27]. This format enables participants to engage with one another, facilitating a more comprehensive understanding of shared and divergent viewpoints [28,29].

Ethical Considerations

The study design, procedures, and informed consent form were evaluated and approved by the local commission for minimal-risk studies of the Psychology Department of the University of Milan-Bicocca (protocol number RM-2024‐854). Participants did not receive any compensation for their involvement.

Study Participants

The study involved 15 participants (12 males and 3 females) aged 25 to 34 years old. The first focus group included 7 participants (6 males and 1 female) aged 25 to 28 years old, while the second group consisted of 8 participants (6 males and 2 females) aged 25 to 34 years old. Participants were selected through a snowball sampling method via messaging apps (eg, WhatsApp [Meta]) to ensure diversity while maintaining relevance to the research topic. Each session lasted approximately 120 minutes.

Eligibility criteria required participants to: (1) be young adults or digital natives (18-35 years old), (2) have varying levels of familiarity with mobile health (mHealth) apps and AI, assessed through a preliminary question during recruitment, and (3) be over the legal age. The participants were evenly distributed between the 2 groups based on their levels of familiarity with and usage of mHealth apps and AI, in order to create as

XSL•FO
RenderX

homogeneous groups as possible and encourage a more productive discussion.

While recruitment efforts sought a balanced sample, the final composition resulted in a gender imbalance (12 men vs 3 women), primarily due to participant availability and recruitment challenges. This limitation is further addressed in the discussion section, where its potential impact on the findings is considered, along with other constraints of this study.

Procedure

Each focus group followed a structured discussion guide to ensure consistency while allowing for open-ended exploration of key themes. The sessions were moderated by a researcher and lasted approximately 120 minutes. The focus groups were conducted remotely via WebEx, ensuring accessibility for all participants.

At the beginning of each session, participants were welcomed and thanked for their participation. The moderator briefly introduced himself and provided an overview of the study's objectives. A general definition of mHealth apps was presented, outlining their main types and uses. The concept of a focus group was explained, including its purpose and ground rules, such as mutual respect, open expression of opinions, and the absence of right or wrong answers. Participants were reassured that their insights were valuable and encouraged to share their perspectives freely.

To create a comfortable and engaging atmosphere, the discussion began with an icebreaker activity, where participants introduced themselves by sharing their names, age, and professional background. This step aimed to foster interaction and establish a collaborative dynamic. The discussion proceeded with a series of structured interactive exercises. First, a free association exercise (mHealth apps), in which participants were asked to list 3 words that came to mind when thinking about mHealth apps. Their responses were visualized in a word cloud using Mentimeter. Each participant then explained their choices, and an open discussion followed to explore shared and contrasting views. Second, anchored association exercise (mHealth apps and AI); the previous exercise was repeated, this time focusing on AI-driven mHealth apps. Participants provided 3 words, which were compiled into another word cloud. The discussion centered on emerging themes and differences in perception between traditional and AI-enhanced mHealth apps. Third, a role-playing scenario (using an AI-integrated mHealth app), where participants were asked to imagine using an AI-integrated mHealth app and to identify one advantage and one disadvantage of such technology. Their responses were collected, and a group discussion was conducted to explore the potential benefits and challenges of AI in health care apps. Fourth, preferred communication style for AI chatbots; participants described their ideal AI chatbot communication style by selecting three descriptive words, which were visualized in a word cloud. They then explained their choices, leading to a broader discussion on expectations and preferences regarding AI-driven health communication. Fifth, ranking exercise (AI chatbot communication features), in which participants were asked to rank the words from the previous exercise in order of importance, reflecting their priorities for chatbot interaction. The final group ranking was analyzed collectively, and participants shared their reflections on the results. The focus groups concluded with a summary of key discussion points, followed by a debriefing on the study's objectives and the intended use of the collected data. Participants were invited to ask any final questions or provide additional comments before the session ended.

Data Analysis

The 2 focus group sessions were audio-recorded and subsequently transcribed. The transcripts were analyzed qualitatively using inductive thematic analysis, following the 6-phase approach outlined by Braun and Clarke [29]: (1) familiarizing with the data, (2) generating of initial codes, (3) searching for themes, (4) reviewing themes, (5) defining and naming themes and, (6) producing the report. This analytical approach is inherently recursive rather than strictly linear, allowing for flexibility in revisiting and refining earlier phases as new insights emerge. Thematic analysis was chosen as it enables an in-depth exploration of participants' perspectives, capturing both shared viewpoints and individual divergences, thereby facilitating a nuanced understanding of the topic under investigation.

To ensure a rigorous and unbiased analytical process, 2 researchers independently reviewed the transcripts multiple times, noting initial impressions and generating preliminary codes. This parallel approach was intended to minimize researcher bias and provide a broader interpretative perspective. The degree of agreement between the 2 researchers was calculated directly, resulting in an agreement rate of approximately 82.76%, indicating a very high level of consistency. After calculating the agreement, the researchers presented their identified codes, explaining the rationale behind their choices and highlighting areas of consensus and divergence. Through discussion, they worked toward a shared interpretation, reconciling differences where necessary. This collaborative process enabled the researchers to refine the thematic structure, identifying the main themes for the thematic analysis.

Once the themes were clearly defined and named, representative verbatim excerpts were selected and organized in a table, alongside their corresponding codes and main themes. Data were analyzed in their original language to preserve participants' intended meanings, while coding and theme development were conducted in English. The selected transcript examples were translated into English for reporting purposes.

Results

Overview of the Findings

The participants were divided into the 2 groups considering their levels of familiarity with and usage of mHealth apps and AI. Specifically, 8 out of 15 (mean 53.3%, SD 12.9%) reported frequent use or a positive attitude toward mHealth apps, while 7 out of 15 (mean 46.7%, SD 12.9%) used them less often or not at all. Similarly, all participants had used AI at least once, with 9 out of 15 (mean 60.0%, SD 12.6%) being frequent users,

XSL•FO
RenderX

while the remaining 6 (mean 40.0%, SD 12.6%) used AI sporadically or were opposed to it.

Four themes were identified in relation to the research question: "Usability," "Innovation and Reliability," "Affectivity and Interaction with AI," and "Perceived risks." The "Usability" theme included 3 subthemes: personalization, accessibility and cost, and AI communication style. The "Innovation and Reliability" theme encompassed subthemes related to reliability and understanding, perceived innovation and trust in technology, and use as a monitoring tool. The "Affectivity and Interaction with AI" theme explored the emotional and interactive dynamics with AI, with subthemes of psychological impact and human-AI interaction. Finally, the "Perceived Risks" theme addressed concerns about data privacy and management, the need for human supervision, and uncertainties surrounding the use of AI in mHealth apps, with subthemes including privacy and data security, human oversight and control, and uncertainty and knowledge gaps (refer to Table 1 for a brief description of the themes and subthemes).

Table . Summary of the themes and subthemes that emerged from the thematic analysis.

Themes and subthemes	Description
Usability	This theme addresses how participants perceive the features of mHealth[a] apps with AI[b], focusing on aspects like personalization, accessibility, and AI communication style.
Personalization	The extent to which AI adapts to individual needs, progress, and personal data, offering tailored support and interactions.
Accessibility and cost	Reflections on how easy it is to access and use these apps, along with considerations of their affordability.
AI communication style	Preferences regarding the tone, manner, and ideal interaction style of the AI when addressing users.
Innovation and Reliability	This theme explores the level of trust participants have in mHealth apps with AI, their perception of innovation, and how these technologies are used.
Reliability and understanding	How reliable participants perceive AI to be in providing accurate and useful responses.
Perceived innovation and trust in technology	How participants perceive AI as either a necessary advancement or a technology to be approached with skepticism.
Use as monitoring tool	Reflections on the role of mHealth apps with AI primarily as tools for tracking personal health data.
Affectivity and Interaction with AI	This theme examines the emotional dynamics, uncertainties, and interactions between users and AI in mHealth apps.
Psychological impact	Emotions evoked by interacting with an AI chatbot in a health context, including motivation, anxiety, trust, and skepticism.
Human-AI interaction	Reflections on the comparison between AI chatbots and human operators, highlighting the advantages of personalization and constant availability, but also the need for human interaction for trust and understanding.
Perceived Risks	This theme addresses concerns related to data privacy, the need for human oversight, and uncertainties regarding the use of AI in mHealth apps.
Privacy and data security	Concerns about who manages the collected data, the level of access AI has to personal information, and the differences between traditional health apps and those with AI.
Human supervision and control	The need for human oversight to ensure safety, particularly in managing sensitive health data and situations.
Uncertainty and knowledge gaps	Difficulties in understanding how AI-powered apps work, concerns about the technology, and broader issues such as environmental impact.

[a]mHealth: mobile health.

[b]AI: artificial intelligence

Usability

Overview

Participants emphasized the practical advantages of AI-driven mHealth apps, particularly their precision, ease of use, and ability to deliver personalized, real-time support that enhances user experience. In addition, affordability was identified as an important factor that might influence the adoption of such technologies. However, opinions diverged regarding the preferred communication style, with some imagining a preference for a formal and direct approach, while others anticipated a more friendly and engaging interaction.

XSL•FO
RenderX

> *Like, just thinking about AI doing an initial screening for acceptance—it could speed things up. Or you ask it something, and it gives you an answer in 30 seconds instead of having to book an appointment six months later. But that's only if it's really accurate. I don't know, maybe for very basic things. I'm not sure where the line is between what you can actually ask a chatbot in an app and what really needs a doctor in a proper visit.* [Female #1]

> *So, it turns out to be more precise and easier to use compared to a regular app—more targeted and more user-friendly. I think this would make people more likely to use it.* [Male #8]

> *Artificial intelligence offers more engaging suggestions compared to a regular application that lacks this feature.* [Male #10]

Subtheme 1: Personalization

A major perceived advantage of AI-driven mHealth apps was thought to be their ability to provide a highly personalized experience. Participants highlighted the value of dynamic personalization, imagining apps that refine recommendations based on progress and align with individual needs and goals to enhance engagement. This dynamic personalization was considered a key differentiator from traditional mHealth apps.

> *Personalization is the key because, imagining this app, I see a section where I can specify my interests, what I enjoy doing, and my goals. […] In my view, the main advantage of artificial intelligence in this context is a better match between user interests and objectives […] For example, I imagine an app that suggests hiking trails—places I wasn't aware of but that, thanks to artificial intelligence, it can recommend based on my activity level and the preferences I've set.* [Male #7]

> *The ability to customize a tool definitely adds value—it can do the work for us, access more data, process it, and provide what I'd consider a better answer.* [Female #2]

Subtheme 2: Accessibility and Cost

Participants valued the immediacy of AI-powered responses, considering them a key advantage for health and fitness management. However, some raised concerns that the speed of response might compromise the depth and accuracy of the provided insights. Given these potential limitations, participants agreed that these services should ideally be financially accessible, ensuring that they could remain a cost-effective alternative to professional consultations.

> *The cost is a concern—it shouldn't be more expensive than consulting a nutritionist, that's for sure.* [Male #9]

> *The ability to receive immediate information is always a strong point.* [Male #2]

> *Speed is crucial—the fact that I don't have to wait for a human response would be a fundamental advantage.* [Male #7]

Subtheme 3: AI Communication Style

There was no consensus on the preferred communication style for AI-powered chatbots, which are considered a key tool for the envisioned AI-driven mHealth apps in this study. Some participants favored a pragmatic, direct approach—emphasizing clarity and efficiency in responses—could be more effective. Others preferred a friendlier, more conversational tone, which might be preferable, especially when discussing sensitive health-related topics, where the ideal tone might depend on context. While a professional, detached tone could enhance credibility in medical discussions, a warmer and more engaging approach could be better suited for fitness or lifestyle guidance.

> *It should be pragmatic—it needs to provide concise and direct answers to the issues I bring up.* [Male #10]

> *(it should be) Formal and straightforward, because I'm not looking for casual chat or small talk. If I use an app or chatbot, I don't need friendliness—I just want something fast, efficient, and simple, without being overly technical.* [Female #1]

> *I don't think I'd want a friendly interaction. In fact, that kind of interaction bothers me.* [Male #12]

> *I'd like the chatbot to address me in a friendly manner, but not as if it were a doctor or an ultimate expert trying to teach me how to live.* [Male #4]

> *Whatever role this bot is designed to take on, I believe each role requires a different communication style. If it is meant to provide medical advice, it should, like human doctors, maintain a professional and detached approach to preserve its credibility.* [Male #1]

Innovation and Reliability

Overview

Participants expressed mixed opinions regarding the reliability of AI-driven mHealth apps, particularly in terms of the accuracy of the information these apps might provide. While some recognized their potential, others emphasized the need to verify AI-generated insights before relying on them. Trust in AI technologies varied among participants, with some seeing it as an intriguing and potentially valuable tool for health management, while others remained skeptical about their consistency and applicability across different health-related contexts.

> *I don't mind if the app has a deeper understanding of me—this doesn't scare me; on the contrary, it intrigues me. In a way, it even adds a bit of creativity to using the app.* [Male #11]

> *How reliable is the information provided by artificial intelligence? It might be useful in certain areas, but I'm not sure it's equally reliable across all applications.* [Male #3]

Subtheme 1: Reliability and Understanding

A recurring concern among participants was the degree of trust they might place in AI-generated responses. While some recognized advancements in AI technology, they remained

XSL•FO
RenderX

cautious about the possibility of errors or misleading recommendations. Several participants highlighted that the reliability of these apps could depend on the sophistication of the AI model used, acknowledging the significant progress made in 2025. Despite these improvements, many participants indicated that they would still prefer consulting a human expert, particularly when dealing with complex or highly personalized health concerns—a topic that will be further explored in relation to theme 4: perceived risks, subtheme 2: human supervision and control.

> *A chatbot can make serious mistakes. Trusting a bot instead of a human is risky—of course, humans can also make errors, but a person can look you in the eye and read between the lines, whereas a bot operates in a more uncontrolled manner.* [Female #2]
>
> *The risk is that it might promote certain behaviors that are not suitable for everyone.* [Male #12]
>
> *The chatbot's level is important. If we're talking about something as advanced as ChatGPT or Gemini, I think we've reached a very high standard by 2025.* [Male #9]

Subtheme 2: Perceived Innovation and Trust in Technology

Participants expressed divergent views on the role of AI in health care, oscillating between enthusiasm for its potential and skepticism about its long-term impact. Some perceived AI as a temporary trend, doubting that its applications would revolutionize health management in a meaningful way. Others, however, regarded AI as an evolving tool with growing significance, predicting that it would become increasingly integrated into daily life and professional settings.

> *I wrote down "innovation" and "trend" … But to be honest, I have some doubts about calling it "innovation"—I see it as more of a trend, in my opinion.* [Male #5]
>
> *I don't agree with calling it a trend. Even though AI is experiencing a huge boom right now, it's being used everywhere, and I don't think it's just a passing fad. I believe AI is here to stay, to improve further, and, in a way, to increasingly permeate our lives.* [Male #4]

Subtheme 3: Use as Monitoring Tool

One of the most frequently mentioned applications of AI-driven mHealth apps was their role in health monitoring, with participants seeing AI tools as a way to enhance traditional mHealth apps they were already familiar with. These apps were envisioned as valuable for tracking health parameters like medical health, sleep patterns, nutrition, and physical activity, offering the ability to record, manage, analyze, and visualize personal data, especially when combined with AI-driven personalization and real-time feedback.

> *What I like most about these apps [traditional mHeatlh apps] is their ability to track progress. It's not so much about receiving notifications to remind me to move […] I would appreciate If they provided me an overview of how things have been going over time—whether there have been improvements or declines.* [Male #11]
>
> *I used an mHealth app for running, mainly for tracking my progress. […] If the app included also an AI chatbot that guides my training and suggests specific workouts it would be even easier.* [Male #5]
>
> *The fact that an app can monitor my physical condition, diet, and habits is really convenient. I can see changes over time and adjust my habits immediately.* [Male #4]

Affectivity and Interaction With AI

Overview

The focus groups highlighted various concerns and reflections regarding the psychological implications of interacting with AI-driven mHealth apps, especially if chatbot-based consultations were to replace human specialists. Participants expressed mixed feelings about this potential shift, recognizing both potential benefits and drawbacks.

> *Anxiety and motivation go hand in hand. I haven't used these apps [mHealth traditional apps] much, but when I did, they always started with motivation but ended up making me anxious.* [Male #2]
>
> *If it's about something sport-related—like asking the chatbot how many kilometers I should run to train for something—that's fine. But if it's a deeper health-related conversation about symptoms or medical issues, it would feel strange to discuss it with a chatbot. I would find it a bit distant.* [Female #1]

Subtheme 1: Psychological Impact

From an emotional perspective, participants described a range of potential psychological effects associated with these apps. Some raised concerns about increased stress levels, particularly due to frequent health-related notifications, which could inadvertently heighten anxiety rather than alleviate it. Participants also raised concerns about how AI-generated recommendations might affect motivation and emotional well-being, suggesting that the absence of human interaction could reduce accountability or increase stress in some users. Conversely, several participants viewed AI-driven support as an opportunity for greater personal empowerment. In this regard, they suggested that receiving recommendations from an app—without the perceived pressure of a direct human interaction—might foster a sense of autonomy and reduce stress.

> *I already don't sleep much. If I had an app telling me "It's 8 PM, time to go to bed," I'd get even more anxious, and, in the end, I'd probably stay up until 3 AM.* [Female #3]
>
> *My fear is that if this AI provides recommendations, support, and guidance, I might lose autonomy—especially regarding diet and exercise. Since the chatbot is always available, it's different from dealing with a professional.* [Male #11]

XSL•FO
RenderX

> *Thinking of something that is a pro and con at the same time... For example, besides sports, another thing mentioned earlier is the area of diets, meal plans, etc. Maybe you don't feel the anxiety of saying, "Oh no, I didn't stick to this," if you just report it to an app. Indicating that you didn't meet a certain goal doesn't come with the fear of judgment or reactions. [...] But the downside is that you might lack some motivation, because you think, "Who cares if I write on my phone that I binged on sweets?"—compared to having to admit it to a real person, like a professional you're seeing.* [Female #1]

Subtheme 2: Human AI Interaction

The nature of the interaction between humans and AI was another key theme that emerged. Some participants reported that chatbot-based consultations could potentially lack warmth, emotional support, and the relational aspects that characterize human interactions, particularly in health care settings. This perceived absence of a human element was viewed as a limitation, potentially reducing the effectiveness of AI-driven consultations. However, others saw the use of AI as an advantage, as it might provide a neutral and judgment-free environment for discussing personal or sensitive health concerns. In this sense, rather than creating a sense of detachment, AI-driven apps could offer a new and alternative way of engaging with health-related issues.

> *The real loss would be the human element—the fact that we would be handing over this work to something that isn't human.* [Male #5]

> *If I imagine consulting a human—a doctor or a nutritionist—they can provide reassurance and emotional support. I'm not sure a chatbot could offer the same kind of relational interaction.* [Male #11]

> *There are now online platforms offering certified psychotherapy sessions. On one hand, these services involve professionals, but on the other, the screen can be both a barrier and a tool that helps people open up.* [Female #2]

Perceived Risks

Overview

The adoption of AI-driven mHealth apps could bring various uncertainties that would need to be addressed before these technologies can become widely implemented with minimal risks to users. Among the most prominent concerns raised during the focus group discussions were issues related to data privacy, the potential need for human oversight, and broader knowledge gaps regarding AI's role in health care.

> *The risk is that the app might promote behaviors that are not suitable for all individuals or body types and fail to acknowledge differences.* [Male #12]

> *The issue of privacy is significant. How much data do I have to provide to get a personalized AI-powered service? That's something I worry about.* [Female #2]

Subtheme 1: Privacy and Data Security

A significant portion of participants voiced concerns regarding the collection and management of sensitive personal data, particularly health-related information. They questioned the extent of the data being requested, the entities responsible for processing it, and the specific purposes for which it might be used. Some participants advocated for state or governmental oversight of data management, believing that public regulation would provide greater protection than private organizations. However, a shared expectation across the discussions was the need for transparency in data handling. Participants emphasized that sensitive health data should not be exploited for commercial purposes, such as market profiling.

> *Speaking of privacy, I remember an incident—last year, I think—in the United States. In some states where abortion had been made illegal, women using period-tracking apps found out that their sensitive data had been sold. That's definitely worrying.* [Female #3]

> *In my opinion, the data shouldn't be, let's say, managed by private companies, but by the national healthcare system. So, basically, the management should be as state-run as possible.* [Female #2]

> *Privacy is a big concern. You hear people talking about it—what happens to our data when we use these apps?.* [Female #1]

Subtheme 2: Human Supervision and Control

Many participants agreed that AI-driven mHealth apps should be envisioned not as standalone tools but rather as supplementary aids under the guidance of health care professionals. These apps were viewed as potentially useful in assisting specialists but not as substitutes for expert judgment. The presence of a medical professional was regarded as crucial to ensure the accuracy of recommendations, assist users in interpreting AI-generated advice, and safeguard sensitive data.

> *Losing the human element is dangerous, especially in healthcare. I would feel more comfortable talking to a doctor who can guide me through my health journey.* [Male #5]

> *As long as there is human oversight, AI in these apps is fine.* [Female #3]

Subtheme 3: Uncertainty and Knowledge Gaps

The discussions also highlighted various uncertainties and knowledge gaps surrounding AI technology and its integration into health care applications. Participants expressed uncertainty about how well AI could handle the complexity of health-related issues, particularly when distinguishing between everyday concerns and more serious conditions. They also raised unexpected but resonant concerns about the long-term implications of widespread AI adoption, such as its environmental footprint.

> *For very basic issues, AI might work. But where is the boundary between what a chatbot can handle and what requires a real doctor's visit?* [Female #1]

XSL•FO
RenderX

> *I think the risk of errors decreases over time, but another concern that hasn't been mentioned is the environmental impact. AI consumes a huge amount of energy, and if we start using it for everything, including these apps, we might face sustainability issues.* [Male #7]

Discussion

This section presents the principal findings regarding individuals' attitudes toward AI-driven mHealth apps, followed by a discussion contextualizing these results within existing literature. Both consistencies with previous work and novel insights from the study are highlighted, followed by an examination of the limitations and suggestions for future research directions.

Key Findings

This study explored how potential users perceive and anticipate AI-powered mHealth applications, identifying key factors that could influence usability, reliability, emotional interaction, and perceived risks. The findings suggest that participants associated personalization, accessibility, and data monitoring with significant advantages, whereas primary concerns centered around privacy, data security, and the reliability of AI-generated information. In addition, aspects such as communication style and the role of human interaction elicited mixed opinions, highlighting an ambivalent attitude toward adopting these technologies. These attitudes seemed closely linked to past experiences, with factors like monitoring and cost being evaluated through familiarity with traditional mHealth apps, while aspects like speed and automation were framed within broader AI-related expectations.

Participants appreciated the potential accuracy and adaptability of AI-driven mHealth apps, recognizing their ability to enhance user experience through tailored recommendations and real-time updates. However, preferences regarding AI's communication style varied: while some users favored a clear and professional tone, others found a more friendly and engaging approach to be more effective. To accommodate this diversity, app developers could consider integrating customizable tone settings or adaptive algorithms that adjust communication style based on user preferences or behavior [30,31]. Furthermore, despite acknowledging the accessibility and cost-effectiveness of these technologies, some participants expressed concerns about the quality and depth of AI-provided responses compared to professional consultations. This highlights communication style as a key design consideration and suggests that tone flexibility may improve user comfort and engagement.

Reliability emerged as another critical issue. While some participants viewed technological advancements as a positive factor, many emphasized the need to verify AI-generated information before fully relying on it. This concern reflects broader uncertainties about AI's role in health management, balancing its potential as a supportive tool with skepticism about its ability to replace human judgment, particularly in complex or sensitive situations.

Given the novelty of AI-driven mHealth apps and the limited direct experience with these tools, participants' attitudes were largely shaped by past interactions with traditional mHealth apps for certain factors (eg, monitoring and cost) and by their broader knowledge or experiences with AI for others (eg, speed and automation). This suggests that their perceptions stem from an attempt to integrate these 2 reference points. Future research could further investigate how these general attitudes develop, whether certain factors weigh more heavily in shaping acceptance or rejection, and what individual differences contribute to the ambivalence observed in responses to these emerging technologies.

Relation to Existing Literature

While very few studies have directly examined users' perceptions of the risks and benefits of AI-driven mHealth apps [8,32], existing research has explored related aspects by focusing either on AI in health care more broadly or on the characteristics and user expectations of traditional mHealth apps. Some findings from these studies are consistent with those of this study, whereas others reveal key differences that emerge when AI is integrated into mHealth solutions.

Usability emerges as a key factor influencing attitudes toward AI-driven mHealth apps, with personalization playing a particularly strong role in fostering positive perceptions. Participants anticipated AI's ability to offer tailored recommendations, real-time feedback, and enhanced accessibility—features that, when combined with the affordability of mHealth solutions (eg, diet planning or personalized fitness programs), were imagined to make AI integration particularly appealing. These findings align with existing literature on user engagement in mHealth, which emphasizes customization as a critical driver of adoption [9,27,28]. However, this study also reveals a degree of ambivalence: while many users envisioned AI as efficient and convenient, others expressed uncertainty about its communication style and interaction patterns. This variability in perception suggests that attitudes are shaped not by direct experience with these technologies but by an attempt to extrapolate from existing knowledge of AI and mHealth.

The theme of innovation and reliability elicited mixed reactions, with attitudes varying significantly based on users' previous knowledge of AI and mHealth apps rather than the actual use. Some participants imagined AI as a natural progression in digital health care, viewing it as a promising tool for health management—an attitude that may reflect a form of pro-innovation bias [33], where technological advancements are perceived as inherently beneficial. Conversely, others remain skeptical, particularly regarding the accuracy and reliability of AI-generated recommendations. This skepticism appeared more connected to general concerns about AI technology rather than direct experience with AI-driven mHealth apps [12,21], suggesting that participants' attitudes reflected pre-existing beliefs about AI's trustworthiness rather than firsthand interactions [10,34,35].

Reliability concerns also intersected with how users perceive mHealth apps as a monitoring tool. While some individuals expected AI-driven mHealth apps to enhance traditional tracking

XSL•FO
RenderX

features through advanced data analysis and personalization, others remained cautious about their role in delivering medical advice. The qualitative data suggests that trust in AI-driven recommendations depended more on participants' general exposure to AI-based tools rather than direct usage of AI-driven mHealth apps. Participants familiar with AI-powered technologies appeared more open to their guidance, whereas those with limited knowledge expressed greater hesitation—a dynamic consistent with patterns observed in AI adoption literature [19,21,36]. These findings reinforce the notion that trust in technology is not only shaped by technical performance but also by users' broader conceptualizations, expectations, and personal experiences of AI in different contexts.

The psychological and emotional impact of AI-driven mHealth apps presents a particularly nuanced dynamic, as the same features were imagined to elicit contrasting reactions among different users. One central factor in this divergence was human-AI interaction, particularly in hypothetical scenarios where AI replaced direct human engagement. Some users expected AI-based coaching to reduce social pressure, finding it less intimidating than interactions with human professionals. For them, AI was envisioned as a judgment-free space for self-improvement, enhancing feelings of control and autonomy—a benefit highlighted in studies examining patient-clinician dynamics [37]. However, others express discomfort with the absence of human empathy, speculating that AI may fail to provide the emotional support and nuanced understanding required in sensitive health-related contexts, as noted in research on AI's limitations in therapeutic settings [38].

This tension is closely connected to the broader discussion on human supervision and control. Participants who are skeptical of AI's role in health care emphasize the need for human oversight, particularly in situations where AI-generated advice could have serious implications. This perspective aligns with frameworks proposing hybrid human-AI systems [39], and is supported by findings showing concerns about fully automated health solutions in the absence of clear human intervention mechanisms [40]. These observations collectively reinforce the perspective that AI-driven medical solutions were imagined as more readily accepted as assistive tools rather than autonomous decision-makers [6,12].

Privacy concerns and data governance emerged as dominant themes, revealing participants' surface-level awareness of legal risks surrounding sensitive health data management. While users expressed concrete anxieties about data access controls [18], storage security protocols, and risk-benefit tradeoffs of AI convenience versus potential breaches, their understanding of specific regulatory frameworks remained limited. This highlights the need for clearer communication around data governance. Developers could help bridge these knowledge gaps through evidence-based user education strategies, such as (1) interactive onboarding tutorials with just-in-time explanations, (2) "privacy nutrition labels" for transparent policy summaries, and (3) participatory consent processes that visualize data flows—all proven to enhance comprehension and trust in health AI contexts [41,42]. Critically, these approaches move beyond static disclosures to dynamic, user-centered designs that align with real-world decision-making needs [42]. These findings mirror broader literature identifying privacy paradoxes in digital health adoption [14], where expressed concerns often outweigh practical protective behaviors. Notably, these apprehensions coexisted with broader skepticism about AI's reliability in health care contexts, leading some users to stress the perceived necessity of human supervision, framing it as both a quality control mechanism and an ethical safeguard. This concern permeated multiple discussion topics in the findings, suggesting it represents a core barrier to trust in AI-driven health care tools.

This landscape of uncertainty spanned multiple dimensions—both technical and conceptual. Participants questioned not only how AI-driven systems function algorithmically but also how they comply with existing regulations, what monitoring systems are in place, and who holds accountability for decision-making. Notably, these knowledge gaps extended beyond core functionality to include unexpected concerns, such as the potential environmental impact of AI infrastructure. Although mentioned only once in focus groups, this issue quickly resonated with other participants, suggesting that ecological considerations (eg, energy consumption of AI models [43]) may subconsciously influence perceptions of mHealth solutions, even among users without direct technical knowledge. Such broad uncertainty, encompassing both technical and nontechnical factors, frequently characterizes public reception of emerging health technologies [18], particularly when transparency about decision-making processes appears insufficient. These concerns frequently translated into demands for clearer accountability frameworks [44] and verifiable regulatory compliance [45], as fundamental requirements for trust-building.

The analysis reveals that attitudes toward AI-driven mHealth apps stem from multiple interconnected factors. These include: (1) theoretical expectations about technical app characteristics (eg, usability and personalization features), (2) individual tendencies to trust or distrust technology, and (3) broader ethical-legal concerns. Crucially, these factors do not operate in isolation but interact dynamically, creating unique perceptual configurations for different users. This complex interplay ultimately determined the spectrum of observed responses—from enthusiastic anticipation to cautious skepticism or outright rejection.

Limitations

While this study provides valuable exploratory insights, several limitations should be acknowledged. First, the gender imbalance within the sample (12 men vs 3 women) may have influenced the findings. Although existing evidence refers separately to mHealth apps and to AI apps in health care, studies suggest that gender can affect attitudes toward these technologies, and similar dynamics may also apply to AI-driven mHealth apps. For example, women tend to report higher health literacy and greater concerns about privacy, while men may place more emphasis on response efficacy and perceive these technologies as easier to use [46,47]. This imbalance may have shaped the salience of certain themes—such as emotional interaction, communication style, and privacy concerns—highlighting the importance of considering gendered perspectives when interpreting qualitative findings [48]. Future research should therefore aim to achieve

XSL•FO
RenderX

more balanced gender representation and could more rigorously explore gendered dimensions by incorporating stratified sampling strategies, gender-sensitive interview protocols, or comparative analysis frameworks designed to detect gender-specific themes [49].

Regarding the age distribution, the study focused on young adults (mean age: 27 years; range: 25 - 34 years), a demographic that represents a particularly relevant target group for AI-mHealth given their high engagement with digital health tools and greater trust in AI technologies [22-24]. However, this choice necessarily limits the generalizability of findings to other age groups, whose needs, digital literacy, and expectations may differ [50,51]. Future research could explore these age-related differences through targeted sampling and quantitative approaches. Finally, as an exploratory qualitative study, the methodological approach is inherently limited by potential group dynamics, such as dominant voices or social desirability [52]. Future research could benefit from robust mixed methods designs to address some of the limitations mentioned earlier.

Conclusions and Future Directions

This study offers valuable insights into users' attitudes toward the potential integration of AI in mHealth apps, focusing on key advantages such as personalization, accessibility, cost-effectiveness, and their potential as monitoring tools that leverage AI for tailored recommendations. The immediacy and affordability of such apps were seen as notable benefits, suggesting that they could make health management more accessible compared with traditional professional consultations. However, concerns about data management, transparency, and users' understanding of these technologies emerged as significant barriers. Given the novelty of AI-driven mHealth apps, many participants formed opinions based on general perceptions of AI or traditional mHealth apps rather than direct experience, highlighting a broader uncertainty about their functionality and reliability.

Ambivalent attitudes were particularly evident regarding the perceived reliability of AI-generated recommendations, the need for human supervision, and the psychological impact of AI interactions. These factors elicited mixed reactions, with responses largely influenced by previous knowledge, pre-existing opinions about AI, and attitudes toward digital health tools. In addition, certain themes, although not central to the main discussion, warrant further exploration. For instance, a pro-innovation bias was observed in some users, indicating a tendency to view technological advancements as inherently beneficial. Furthermore, concerns about the environmental impact of AI systems—though mentioned only briefly—resonated with a broader segment of participants, highlighting a potential influence on consumer attitudes.

Developers could focus on highlighting the positive aspects that emerged in this study, such as the increased personalization options offered by AI-driven mHealth apps, their cost-effectiveness, and their strong monitoring capabilities for health data. At the same time, they should address potential risks or concerns, such as enhancing the clarity and transparency of data management practices, particularly around privacy, and ensuring systems of periodic human supervision. Furthermore, developers could provide greater customization for those ambivalent factors by allowing users to select the type of interaction with the AI, adapting the intensity of interactions to suit personal preferences—for example, adjusting the tone to be more formal or informal, scientific or motivational, and so on.

Future research should address the limitations of this study and explore these underlying factors in greater depth. Using quantitative or mixed-methods approaches could provide deeper insight into the psychological and behavioral drivers that influence attitudes toward the integration of AI in mHealth apps. This would ultimately guide the development of more user-centered and ethically responsible strategies for these technologies, ensuring a vision that not only protects and supports end users in the most comprehensive way possible, but also addresses the delicate nature of health-related matters. Furthermore, as AI-driven mHealth apps become more widely available, longitudinal studies tracking user perceptions before and after direct experience with such technologies could offer valuable insights into the evolution of attitudes toward these technologies over time [27,28]. Given the novelty of these technologies, such contributions can make a real difference, both in the academic landscape and in practical terms, by helping design safer, more effective, and widely accepted AI-driven mHealth apps that truly meet the needs of end users.

Acknowledgments

This research was funded by the Italian Ministry of University and Research under the scheme Research Projects of National Interest (PRIN 20227BY88A).

Authors' Contributions

AA and MD contributed to conceptualization. AA and GD managed methodology. AA and GD conducted formal analysis. AA handled the investigation. MD, AA, and PS managed resources. GD and AA handled data curation. AA and GD contributed to writing—original draft preparation. AA, MD, and PS handled writing—review and editing. GD and AA managed visualization. MD and PS conducted supervision. MD and PS managed project administration. All authors have read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

XSL•FO
RenderX

References

1. Mishra D, Mishra R, Agarwal R. Impacts of artificial intelligence on society. Journal of Science Research International (JSRI) 2025 Mar 18;11:29-43. [doi: 10.5281/zenodo.15044020]
2. Nadikattu RR. The emerging role of artificial intelligence in modern society. International Journal of Creative Research Thoughts 2016;4(4) [FREE Full text]
3. Howard J. Artificial intelligence: Implications for the future of work. American J Industrial Med 2019 Nov;62(11):917-926. [doi: 10.1002/ajim.23037]
4. Sengar SS, Hasan AB, Kumar S, Carroll F. Generative artificial intelligence: a systematic review and applications. Multimed Tools Appl 2024 Aug 14;84(21):23661-23700. [doi: 10.1007/s11042-024-20016-1]
5. Jiang F, Jiang Y, Zhi H, et al. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol 2017 Dec;2(4):230-243. [doi: 10.1136/svn-2017-000101]
6. Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. Nat Biomed Eng 2018;2(10):719-731. [doi: 10.1038/s41551-018-0305-z]
7. Kelly S, Kaye SA, Oviedo-Trespalacios O. What factors contribute to the acceptance of artificial intelligence? A systematic review. Telematics and Informatics 2023 Feb;77:101925. [doi: 10.1016/j.tele.2022.101925]
8. Deniz-Garcia A, Fabelo H, Rodriguez-Almeida AJ, et al. Quality, usability, and effectiveness of mHealth apps and the role of artificial intelligence: current scenario and challenges. J Med Internet Res 2023 May 4;25:e44030. [doi: 10.2196/44030] [Medline: 37140973]
9. Birkmeyer S, Wirtz BW, Langer PF. Determinants of mHealth success: an empirical investigation of the user perspective. Int J Inf Manage 2021 Aug;59:102351. [doi: 10.1016/j.ijinfomgt.2021.102351] [Medline: 36540093]
10. Sharma V, Gupta M, Jangir K, Chopra P, Pathak N. The impact of post-use consumer satisfaction on smart wearables repurchase intention in the context of AI-based healthcare information. In: Enhancing Customer Engagement Through Location-Based Marketing: IGI Global; 2023:77-101. [doi: 10.4018/978-1-6684-8177-6.ch007]
11. Ara A, Mifa AF. Integrating artificial intelligence and big data in mobile health: a systematic review of innovations and challenges in healthcare systems. JALHE 2024;3(1). [doi: 10.62304/jbedpm.v3i01.70]
12. Richardson JP, Smith C, Curtis S, et al. Patient apprehensions about the use of artificial intelligence in healthcare. NPJ Digit Med 2021 Sep 21;4(1):140. [doi: 10.1038/s41746-021-00509-1] [Medline: 34548621]
13. Perc M, Ozer M, Hojnik J. Social and juristic challenges of artificial intelligence. Palgrave Commun 2019;5(1) [FREE Full text] [doi: 10.1057/s41599-019-0278-x]
14. Esmaeilzadeh P. Use of AI-based tools for healthcare purposes: a survey study from consumers' perspectives. BMC Med Inform Decis Mak 2020 Jul 22;20(1):170. [doi: 10.1186/s12911-020-01191-1] [Medline: 32698869]
15. Muthineni SR. AI in mobile health apps: transforming chronic disease management. IJSRCSEIT 2025 [FREE Full text]
16. Khalifa M, Albadawy M. Artificial intelligence for clinical prediction: exploring key domains and essential functions. Computer Methods and Programs in Biomedicine Update 2024;5:100148. [doi: 10.1016/j.cmpbup.2024.100148]
17. Yildirim N, Zimmerman J, Preum S. Technical feasibility, financial viability, and clinician acceptance: on the many challenges to AI in clinical practice. Presented at: AAAI Fall 2021 Symposium in Human Partnership with Medical Artificial Intelligence; Nov 4-6, 2021 URL: http://star.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-3068/short5.pdf [accessed 2025-04-01]
18. Čartolovni A, Tomičić A, Lazić Mosler E. Ethical, legal, and social considerations of AI-based medical decision-support tools: a scoping review. Int J Med Inform 2022 May;161:104738. [doi: 10.1016/j.ijmedinf.2022.104738] [Medline: 35299098]
19. Albahri AS, Duhaim AM, Fadhel MA, et al. A systematic review of trustworthy and explainable artificial intelligence in healthcare: Assessment of quality, bias risk, and data fusion. Information Fusion 2023 Aug;96:156-191. [doi: 10.1016/j.inffus.2023.03.008]
20. Rowland SP, Fitzgerald JE, Holme T, Powell J, McGregor A. What is the clinical value of mHealth for patients? NPJ Digit Med 2020;3(1):4. [doi: 10.1038/s41746-019-0206-x]
21. Gillath O, Ai T, Branicky MS, Keshmiri S, Davison RB, Spaulding R. Attachment and trust in artificial intelligence. Comput Human Behav 2021 Feb;115:106607. [doi: 10.1016/j.chb.2020.106607]
22. Silva BMC, Rodrigues J, de la Torre Díez I, López-Coronado M, Saleem K. Mobile-health: a review of current state in 2015. J Biomed Inform 2015 Aug;56:265-272. [doi: 10.1016/j.jbi.2015.06.003] [Medline: 26071682]
23. Nadarzynski T, Miles O, Cowie A, Ridge D. Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: a mixed-methods study. Digit HEALTH 2019;5:2055207619871808. [doi: 10.1177/2055207619871808] [Medline: 31467682]
24. Lupton D. Young people's use of digital health technologies in the global north: narrative review. J Med Internet Res 2021 Jan 11;23(1):e18286. [doi: 10.2196/18286] [Medline: 33427684]
25. Anderson M, Perrin A. Technology use among seniors. : Pew Research Center for Internet & Technology; 2017 May.
26. Helsper EJ, Eynon R. Digital natives: where is the evidence? British Educational Res J 2010 Jun;36(3):503-520. [doi: 10.1080/01411920902989227]
27. Rivera-Romero O, Gabarron E, Ropero J, Denecke K. Designing personalised mHealth solutions: an overview. J Biomed Inform 2023 Oct;146:104500. [doi: 10.1016/j.jbi.2023.104500]

XSL•FO
RenderX

28. Vo V, Auroy L, Sarradon-Eck A. Patients' perceptions of mHealth apps: meta-ethnographic review of qualitative studies. JMIR Mhealth Uhealth 2019 Jul 10;7(7):e13817. [doi: 10.2196/13817] [Medline: 31293246]
29. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
30. Germanakos P, Tsianos N, Lekkas Z, Mourlas C, Samaras G. Realizing comprehensive user profile as the core element of adaptive and personalized communication environments and systems. Comput J 2009 Oct 1;52(7):749-770. [doi: 10.1093/comjnl/bxn014]
31. Troussas C, Krouska A, Sgouropoulou C. The Role of Communication Styles in the Learning Process Human-Computer Interaction and Augmented Intelligence: Springer Nature Switzerland; 2025:235-303. [doi: 10.1007/978-3-031-84453-9_6]
32. Khan F, Tulon TA, Masrur N, Hasan M, Badrul T, Islam A. Development and evaluation of ShasthoBondhu: mHealth app for guiding emergency remote healthcare with wearable IoT and AI fusion. Presented at: 2024 Advances in Science and Engineering Technology International Conferences (ASET); Jun 3-5, 2024; Abu Dhabi, United Arab Emirates. [doi: 10.1109/ASET60340.2024.10708738]
33. Rogers EM. New product adoption and diffusion. J Consum Res 1976 Mar;2(4):290. [doi: 10.1086/208642]
34. Sheppard MK. mHealth apps: disruptive innovation, regulation, and trust—a need for balance. Med Law Rev 2020 Aug 1;28(3):549-572. [doi: 10.1093/medlaw/fwaa019] [Medline: 32638001]
35. Horgan D, Romao M, Morré SA, Kalra D. Artificial intelligence: power for civilisation – and for better healthcare. Public Health Genomics 2020 Feb 11;22(5-6):145-161. [doi: 10.1159/000504785]
36. Oksanen A, Savela N, Latikka R, Koivula A. Trust toward robots and artificial intelligence: an experimental approach to human–technology interactions online. Front Psychol 2020;11:568256. [doi: 10.3389/fpsyg.2020.568256]
37. Lupton D. Data selves: more-than-human perspectives. : Polity Press Cambridge; 2020 URL: https://quod.lib.umich.edu/j/jcms/images/18_61.3ram-lupton.pdf [accessed 2025-04-01]
38. Topol E. Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again: Hachette UK; 2019. URL: https://tinyurl.com/4nsmk6we [accessed 2025-04-01]
39. Holzinger A, Langs G, Denk H, Zatloukal K, Müller H. Causability and explainability of artificial intelligence in medicine. WIREs Data Min & Knowl 2019 Jul;9(4):e1312. [doi: 10.1002/widm.1312]
40. Laranjo L, Dunn AG, Tong HL, et al. Conversational agents in healthcare: a systematic review. J Am Med Inform Assoc 2018 Sep 1;25(9):1248-1258. [doi: 10.1093/jamia/ocy072]
41. Kalkman S, van Delden J, Banerjee A, Tyl B, Mostert M, van Thiel G. Patients' and public views and attitudes towards the sharing of health data for research: a narrative review of the empirical evidence. J Med Ethics 2022 Jan;48(1):3-13. [doi: 10.1136/medethics-2019-105651]
42. Bietz MJ, Bloss CS, Calvert S, et al. Opportunities and challenges in the use of personal health data for health research. J Am Med Inform Assoc 2016 Apr 1;23(e1):e42-e48. [doi: 10.1093/jamia/ocv118]
43. Strubell E, Ganesh A, McCallum A. Energy and policy considerations for modern deep learning research. Presented at: AAAI Conference on Artificial Intelligence 2020; Feb 7-12, 2020; New York Hilton Midtown, New York, USA URL: https://ojs.aaai.org/index.php/AAAI/article/view/7123 [accessed 2025-04-01] [doi: 10.1609/aaai.v34i09.7123]
44. Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. Nat Mach Intell 2019;1(9):389-399. [doi: 10.1038/s42256-019-0088-2]
45. Gerke S, Minssen T, Cohen G. Ethical and legal challenges of artificial intelligence-driven healthcare. In: Artificial Intelligence in Healthcare: Elsevier; 2020:295-336 URL: https://www.sciencedirect.com/science/article/pii/B9780128184387000125 [accessed 2025-04-01] [doi: 10.1016/B978-0-12-818438-7.00012-5]
46. Li AKC, Rauf IA, Keshavjee K. Knowledge is not all you need to generate trust in AI use in healthcare. medRxiv. Preprint posted online on Jan 11, 2024. [doi: 10.1101/2024.01.10.24301097]
47. Luo Y, Mou J. Understanding gender differences in mHealth apps continuance: a modified protection motivation theory. JEBDE 2022 Dec 16;1(1/2):225-247. [doi: 10.1108/JEBDE-09-2022-0032]
48. Cirillo D, Solarz SC, Guney E. Sex and Gender Bias in Technology and Artificial Intelligence: Biomedicine and Healthcare Applications: Academic Press; 2022. URL: https://tinyurl.com/2ntd5dth [accessed 2025-09-04]
49. Lindsay S, Kolne K. Understanding clinicians' strategies for providing gender-sensitive care: an exploration among pediatric rehabilitation health care providers. Disabil Rehabil 2022 Jun;44(11):2437-2447. [doi: 10.1080/09638288.2020.1836270] [Medline: 33096004]
50. Schreurs K, Quan-Haase A, Martin K. Problematizing the digital literacy paradox in the context of older adults' ICT use: aging, media discourse, and self-determination. Canadian Journal of Communication 2017 May 23;42(2):359-377. [doi: 10.22230/cjc.2017v42n2a3130]
51. Doyle S, Pavlos R, Carlson SJ, et al. Efficacy of digital health tools for a pediatric patient registry: semistructured interviews and interface usability testing with parents and clinicians. JMIR Form Res 2022 Jan 17;6(1):e29889. [doi: 10.2196/29889] [Medline: 35037889]
52. O.Nyumba T, Wilson K, Derrick CJ, Mukherjee N. The use of focus group discussion methodology: insights from two decades of application in conservation. Methods Ecol Evol 2018 Jan;9(1):20-32. [doi: 10.1111/2041-210X.12860]

XSL•FO
RenderX

Abbreviations

AI: artificial intelligence
mHealth: mobile health

Edited by A Kushniruk; submitted 16.04.25; peer-reviewed by C Ochoa-Zezzatti, C Obianyo; revised version received 04.06.25; accepted 09.06.25; published 26.09.25.

Please cite as:
Aboueldahab A, Damaschi G, D'Addario M, Steca P
Exploring Young Adults' Attitudes Toward AI-Driven mHealth Apps: Qualitative Study
JMIR Hum Factors 2025;12:e76075
URL: https://humanfactors.jmir.org/2025/1/e76075
doi:10.2196/76075

© Ali Aboueldahab, Gabriele Damaschi, Marco D'Addario, Patrizia Steca. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 26.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Digital Health Intervention for Patient Monitoring in Immune-Mediated Inflammatory Diseases: Cocreation and Feasibility Study of the IMIDoc Platform

Diego Benavent[1*]; Jose M Iniesta-Chamorro[2,3*]; Marta Novella-Navarro[4]; Miguel Pérez-Martínez[5]; Nuria Martínez-Sánchez[6]; Mónica Kaffati[4]; Manuel Juárez-García[4]; Marina Molinari-Pérez[4]; Andrea González-Torbay[7]; Mariana Gutiérrez[4]; Natalia López-Juanes[4]; Victoria Navarro-Compán[4]; Irene Monjo-Henry[4]; Germán Rodríguez-Rosales[8]; Javier Bachiller[9]; Enrique Calvo-Aranda[4,10]; Xabier Michelena[11]; Laura Berbel-Arcobé[1]; Alejandro Balsa[4]; IMIDOC Research Team[13]; Enrique J Gómez[2,3,12]; Chamaida Plasencia-Rodríguez[4]

[1]Hospital Universitari de Bellvitge, L'Hospitalet de Llobregat, Carrer de la Feixa Llarga, Barcelona, Spain

[10]Department of Medicine, Universidad Complutense de Madrid, Madrid, Spain

[11]Hospital Universitari Vall d'Hebron, Barcelona, Spain

[12]Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina, CIBER, Madrid, Spain

[13]

[2]Biomedical Engineering and Telemedicine Centre, ETSI, Center for Biomedical Technology, Universidad Politécnica de Madrid, Madrid, Spain

[3]Instituto de Investigación Hospital 12 de Octubre (imas12), Hospital Universitario 12 de Octubre, Madrid, Spain

[4]Rheumatology Department, Hospital Universitario La Paz, Madrid, Spain

[5]Kairós Digital Solutions, Madrid, Spain

[6]Obstetrics and Gynecology Department, Hospital Universitario La Paz, Madrid, Spain

[7]Clinical Immunology Department, Hospital Universitario La Paz, Madrid, Spain

[8]Orthopedics & Traumatology Department, Hospital Universitario del Henares, Madrid, Spain

[9]Rheumatology Department, Hospital Universitario Ramón y Cajal, Madrid, Spain

[*]these authors contributed equally

Corresponding Author:
Diego Benavent
Hospital Universitari de Bellvitge, L'Hospitalet de Llobregat, Carrer de la Feixa Llarga, Barcelona, Spain

Abstract

Background: Immune-mediated inflammatory diseases, such as rheumatoid arthritis and spondyloarthritis, pose challenges due to recurrent flares and gaps in patient monitoring. Traditional health care models often fail to capture disease progression effectively.

Objective: This study aimed to describes the structured cocreation of the IMIDoc platform, an interdisciplinary initiative aimed at improving patient monitoring, education, and health care provider decision-making.

Methods: IMIDoc was cocreated through an interdisciplinary team involving clinical experts, biomedical engineers, and technical developers, using user-centered design principles. The development process included the identification of unmet clinical needs, user-centered app design, implementation of medication management features, patient data recording capabilities, and educational content. A 3-month feasibility and functionality testing was performed to evaluate the usability and technical performance of the apption.

Results: During the feasibility testing, 111 entries were logged for the patient mobile app, comprising 76 errors identified and corrected, 16 improvements addressing functionality, usability, and performance, and 10 evolutionary suggestions. The professional interface received 45 entries, identifying 40 errors and 5 evolutionary suggestions. Ten iterative updates significantly enhanced the user interface intuitiveness and medication reminder functionality, aligning the solution closely with clinical workflows and user needs.

Conclusions: The IMIDoc platform, developed by a multidisciplinary cocreation methodology, shows potential to improve the management of immune-mediated inflammatory diseases ithrough enhanced communication and monitoring. A multicenter clinical study with 360 patients across 5 Spanish hospitals will further evaluate its impact.

Trial Registration: ClinicalTrials.gov NCT06273306; https://tinyurl.com/4t6ubcsf

XSL•FO
RenderX

(JMIR Hum Factors 2025;12:e58095) doi:10.2196/58095

KEYWORDS

immune-mediated inflammatory diseases; cocreation; health care provider; digital health solution; inflammatory; inflammatory disease; monitoring; care model; patient care; development; app; user centered; patient monitoring; decision-making; user needs

Introduction

Immune - mediated inflammatory diseases such as rheumatoid arthritis (RA) or spondyloarthritis(SpA) are chronic diseases characterized by joint or spinal inflammation with sporadic flares, leading to significant functional impairment in affected individuals [1,2]. Managing these diseases often involves complex therapeutic regimens, necessitating close monitoring and effective communication between patients and health care providers to promptly identify changes in disease progression or treatment-related concerns. Traditional care models, which rely heavily on scheduled clinical consultations, may fail to provide a comprehensive and continuous view of disease activity, resulting in gaps in patient care [3].

Advancements in digital health technologies present promising opportunities to enhance the management of immune - mediated inflammatory diseases. Telemonitoring, in particular, offers the potential to bridge the gaps between face-to-face consultations by enabling real-time disease activity tracking and facilitating timely interventions during periods of disease instability [4]. Smartphone data collection offers enhanced research and clinical care capabilities, notably with electronic patient-reported Outcomes (ePROs) [5]. Such integration aims to optimize the quality of patient care and facilitate effective communication between health care professionals and their patients.

In this context, our research group designed the mixed care model (MAM, from the Spanish Modelo Atencion Mixta), which incorporates the use of information and communication technologies, and also designed a previous pilot study called Digireuma, with the goal of developing a digital health tool for monitoring ePROs in patients with RA and SpA[6]. The Digireuma study demonstrated that while remote monitoring via digital tools could reduce the need for physical consultations, challenges such as declining patient adherence, with around half of the patients dropping out during the follow-up of 6 months, and usability issues (primarily related to the login process), remained persistent. A targeted survey administered at the study's conclusion, gathering insights from both patients and professionals, highlighted these challenges and helped identify critical areas for improvement.

Based on previous results, the medical team decided to develop a digital health solution specifically tailored to the needs of both health care professionals and patients. Feedback indicated a demand for reliable educational resources, leading to the development of tailored modules incorporating interactive texts, infographics, and videos created in collaboration with health care professionals. To enhance adherence, the app integrates personalized medication management tools with automated reminders and scheduled messages for patients demonstrating positive clinical evolution, thereby reinforcing a sense of continuous clinical follow - up. In addition, bidirectional communication between patients and providers facilitates real - time feedback, ensuring that the solution effectively bridges the gaps observed in previous digital health interventions for RA and SpA.

In the current manuscript, we delineate the collaborative effort among a clinical team, a group of biomedical engineering experts, and a dedicated team of developers, in cocreating an application tailored to the unique needs of patients with immune - mediated inflammatory diseases. Drawing upon interdisciplinary expertise, their efforts culminated in the IMIDoc app, an innovative digital health solution, which has also led to invaluable insights and lessons for future collaborative health-tech initiatives.

Methods

Development of IMIDoc

In the development of our digital health app, we adopted a structured cocreation process encompassing 2 distinct phases. The initial "planning phase" involved the clinical team meticulously reviewing their previous experiences and identifying prevalent unmet needs in disease management. Concurrently, the clinical team engaged in a comprehensive assessment of potential funding strategies. Then, preliminary discussions with the engineers to outline the envisioned features and functionalities were initiated. This process culminated in the articulation of a clear statement of purpose. The subsequent "development phase" involved the technical and development teams, who adhered to Agile methodologies, fostering a dynamic collaboration between clinicians and engineers [7]. Central to this phase was a rigorous evaluation of pivotal factors influencing technology selection, the formulation of user profiles, as well as the definition and development of the functionalities to be incorporated into the digital solution. In addition, paramount attention was devoted to assessing the correct functioning of the digital solution by end users through pilot testing.

Planning Phase

Overview With the Developers and Statement of Purpose

Upon successfully securing a sustainable funding source, we initiated consultations with several development teams, leveraging our collaboration with the biomedical engineering team from the Universidad Politécnica de Madrid. Following a public tender process, the project was awarded to Kairos DS [8]. During several collaborative sessions with Kairos DS, we meticulously crafted a statement of purpose for our digital health solution, including a description of its main functionalities and features.

XSL•FO
RenderX

Funding

To secure funding for a new medical app, efforts were diversified across private funding contests, collaboration with pharmaceutical companies [9,10], and exploration of public grants focused on health care innovation [11], tapping into multiple sources to maximize potential financial support.

Analysis of Previous Experience

An analysis of the previous remote monitoring experience, the Digireuma study [6], was conducted. For this, semistructured telephone interviews were conducted by an external expert with a randomly selected subset of patients, in order to gain a comprehensive understanding of user experiences, capturing both the strengths and potential areas for improvement of the application and implementation methodology. Clinicians who participated in the previous project completed a questionnaire to evaluate the usability of the medical interface, the usefulness of the solution and its utility in medical practice. Further information on these questionnaires can be found in Text S1 inMultimedia Appendix 1.

Development Phase

Throughout the development phase, interdisciplinary work was shared among the clinicians, the technical software and biomedical engineering team. This also involved intensive dialogue among engineers and end-users, including both patients and clinicians.

Due to the nature of the project, Scrum was adopted by the development team and the product development stakeholders [12]. This framework is based on small, iterative cycles or sprints with each iteration adding value to the product, so that corrective measures can be taken without compromising the entire project [13].

An initial kick-off meeting between clinicians and the technical team aimed at establishing shared objectives and product expectations was held. The development schedule included 15 sprints, with each sprint ending with a review.

Definition of User Profiles and Functionalities

We used a user-centered design approach [14], for which by definition must consider the needs and experiences of end users via a continuous series of evaluations [15,16]. Initial user profiles were established based on interviews with clinicians and patients [6], categorizing users into health care providers and patients. For health care providers, functionalities such as patient data monitoring, alert resolution, and telematic contact were developed. For patients, features included ePROs registration, medication management, reminders, and access to educational content covering disease information, management strategies, and relevant hospital information.

Technical Requirements

As a starting point for designing the architecture of the system, the MAM protocol and the agents that comprise the proposed clinical intervention were considered, combining face-to-face and telematic follow-up protocols [12]. The profiles, needs and functionalities previously defined were then considered. With all of this in mind, the system and components to be developed were then defined: a mobile app for patients, an application for health professionals, and services for communication, data processing and storage.

Throughout the design and development process, it was necessary to evaluate possible technologies and their deployment (Text S2 in Multimedia Appendix 1). The system architecture was designed using cloud-based technologies, specifically Amazon Web Services (AWS), to optimize resource usage and ensure data security [17-19]. To ensure cost-effectiveness, scalability, and ease of maintenance, we selected hybrid and cross-platform technologies [20], allowing IMIDoc to be deployed on both Android and iOS platforms efficiently.

In addition, to carry out the clinical evaluation study, a series of requirements had to be met and specific documentation had to be submitted to the Information Systems Security Office of the Madrid Health Service.

Preclinical Testing Phase of IMIDoc

During the first sprints of the project execution and once the main components and specifications of the user applications were identified, those activities associated with the interface design were carried out—that is, generating wireframes, sketches, and mock-ups. This process is key in user centered-design (UCD) and allowed us to preliminarily evaluate the structure of the applications, the composition of the scenarios and the interactions before the development phase [17].

An initial phase of testing was conducted by a multidisciplinary team of clinicians, engineers and patients, generating discussions in 3 online meetings. A subsequent digital solution testing phase with active users was planned for 4 weeks. Both patients and health care professionals were actively involved. A comprehensive guide detailing functionality, data input methods, and troubleshooting steps was provided to all participants before the testing period. The guide aimed to standardize the usage of the application for all users, thereby minimizing variation and enhancing the validity of the pilot results. During the 4-week period, feedback was collected through forms and informal interviews, focusing on usability, feature relevance, and overall experience.

Ethical Considerations

The protocol of the study was approved by the local ethics committee, CEIm Hospital La Paz (PI-4519) and has been registered as a clinical trial (ClinicalTrials.gov Identifier: NCT06273306).

Results

Planning Phase

Scope and Main Objective of IMIDoc

In a dedicated meeting with the multidisciplinary development team, the scope of the digital solution was outlined by conducting a scoping analysis to delineate its potential capabilities and limitations. The name IMIDoc was selected for the digital health solution proposed in this project. The primary purpose of IMIDoc was defined as addressing an unmet need in the care of patients with RA and SpA, which is the providing

XSL•FO
RenderX

of continuous monitoring. Conventional clinical practices often allow for discrete, scheduled medical consultations, leaving gaps in patient care when immediate attention is most needed. The IMIDoc app was designed to bridge this gap, ensuring timely medical oversight of patients.

Primary Source of Financing

This project was financially supported by a grant from the Fondo de Investigaciones Sanitarias (FIS) of the Instituto de Salud Carlos III (PI22/00777). The grant specifically facilitated the technical development of our digital solution and will allow us to carry out the clinical-technical evaluation in a clinical study.

Analysis of Previous Experience and Unmet Needs

In the platform ideation phase, relevant aspects and features that the digital solution should incorporate were identified. During this phase we considered the opinion of 41 patients and 4 clinicians about their experience with the previous digital solution (Table S1 in Multimedia Appendix 1). Regarding the patients, usability ranged from 85.3% - 97.5%; most found the digital solution easy to use, and well adapted. Over 90% (38/41) of participants found the application was well adapted to its objective, with an adequate level of detail. Most patients reported considered that it allowed them to record their current health situation and would recommend it. The areas patients most frequently cited as needing improvement were the application's design and the amount and detail of informational content related to the disease. On the other hand, clinicians rated the overall usability and design (0 - 10, from worse to best) of the application with a mean (SD) of 8.50 (1.22), the app utility with 8.50 (1.16), and the feasibility with 7.56 (2.41). They primarily expressed concerns regarding the feasibility of the implementation in clinical practice, particularly in terms of time and logistics, followed by the performance of the solution.

Development Phase

The iterative development process was carried out over 15 sprints from January to August 2023, each punctuated by a bi-weekly update meeting and culminating in a final review meeting (Supplementary Figure 1 in Multimedia Appendix 1). The initial meetings laid the groundwork by focusing on service architecture and basic functionalities like login and user registration. Subsequent meetings were geared toward enhancing the user interface, developing workflows in the digital solution, and integrating questionnaire designs. In these bi-weekly review meetings, the development team assessed progress and recalibrated, when necessary, thereby allowing them to respond to changing requirements and deliver a high-quality product efficiently. During each sprint, different tasks were assumed by the development team with the aim of achieving the sprint goal, which was previously set in an online meeting by the clinical team. The middle phase emphasized refining communication features, resource and medication development, and incident management. Data visualization tools like historical graphs and evolution screens were also integrated. The concluding meetings targeted system debugging, specifically in messaging and error management, before a comprehensive final review was conducted.

The collaborative and interdisciplinary process made it possible not only to define the instruments to be implemented for the collection of data of interest, but also to design and develop the applications for the users; that was , mobile apps for patients with RA or SpA, and a Web application for clinicians.

Set of Variables and Instruments

In accordance with the objectives of the clinical study, in which the implementation of the MAM model in clinical practice through IMIDoc will be evaluated, Textbox 1 specifies the clinical and ePROS variables that the solution was designed to capture and that comprise the monitoring and follow-up of patients with RA and SpA.

In addition to the capture of clinical variables, instruments have been implemented to evaluate the usability of the applications, satisfaction, and user experience: System Usability Scale (SUS) and the Net-Promoter Score (NPS) indicator [18,19]. These were evaluated at 10 weeks and 34 weeks, respectively.

XSL•FO
RenderX

Textbox 1. Clinical variables and electronic patient-reported outcomes collected by the IMIDoc app.

1) General characteristics

- Sociodemographics and treatment information including age, sex, toxic habits, date of diagnosis, date of initiation of treatment, medication, etc.
- Electronic patient-reported outcomes (ePROs)
  - a) Patient Global Health (PGH)
  - b) Visual Analogue Scale (VAS)
  - c) Tender Joint Count (TJC) and Swollen Joint Count (SJC)
- Events or incidents
  - a) Flare
  - b) Infection
  - c) Medication problem
  - d) Other reasons
- Adherence

a) Compliance Questionnaire on Rheumatology, CQR

2) Rheumatoid arthritis (ePROs)

- Disease Activity Score 28 (DAS28)
- Health Assessment Questionnaire (HAQ)

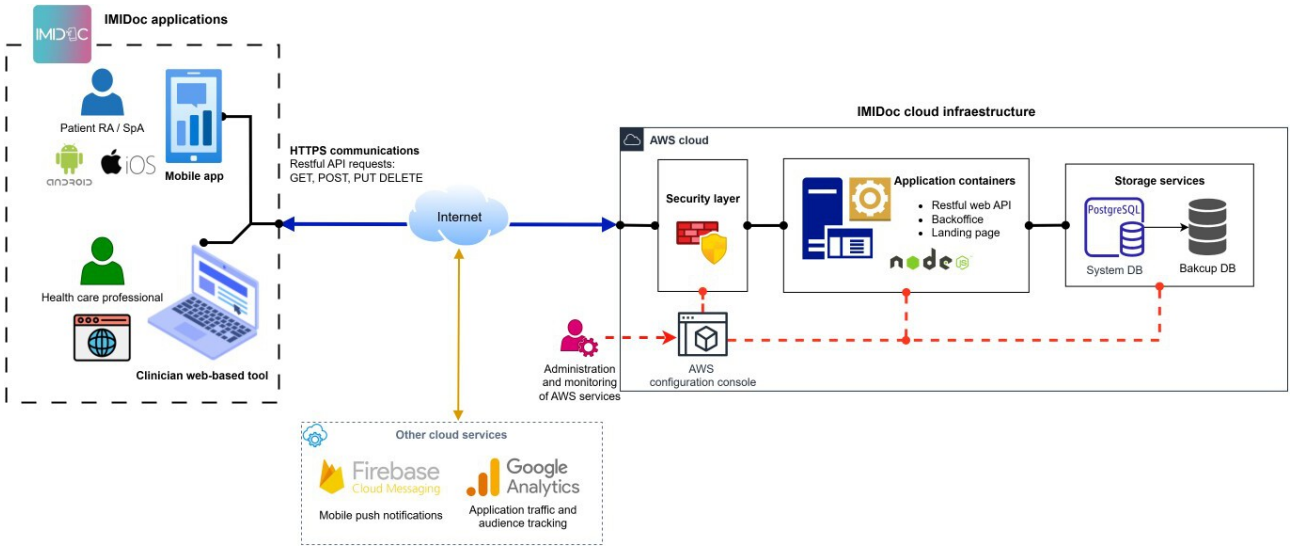
3) Spondyloarthritis (ePROs)

- Bath Ankylosing Spondylitis Disease Activity Index (BASDAI)
- Axial Spondyloarthritis Disease Activity Score (ASDAS)
- Assessment of SpondyloArthritis International Society Health Index
- Painful enthesis count

System Architecture

The analysis and design phase of the project made it possible to identify the user profiles and the functionalities and characteristics of the applications included in the IMIDoc digital health solution. Figure 1 shows a schematic of the main agents, components, and services of the system architecture using the MAM applied to RA and SpA. The figure shows the users (patients and health care professionals) accessing their respective applications and communicating with the service infrastructure deployed in the cloud.

Figure 1. Diagram of the architecture of the IMIDoc solution. RA: rheumatoid arthritis; SpA: Spondyloarthritis; AWS: amazon web services; DB: database; API: Application Programming Interface.



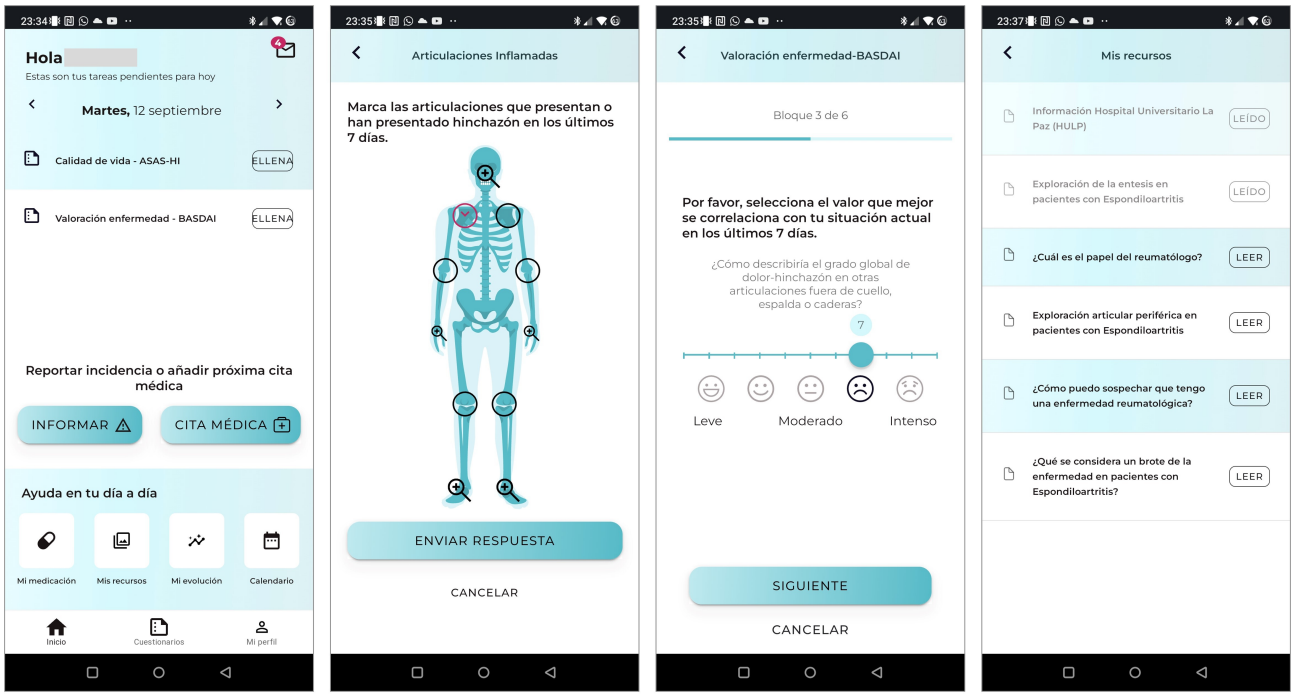
XSL•FO
RenderX

IMIDoc Applications

The first steps in the development of the IMIDoc applications (the mobile app for patients and the web interface for professionals) were focused on addressing the interface design. For this purpose, we used the Figma platform, which is a cloud prototyping platform that facilitates collaborative work.

The IMIDoc mobile app (Figure 2) was designed to incorporate functionalities for recording disease activity, as well as some other domains potentially affected by RA or SpA.

Figure 2. Screenshots of the Spanish version of IMIDoc app designed for patients with rheumatoid arthritis (RA) and spondyloarthritis. (a) Home screen displaying pending tasks, including daily questionnaires, and options to report incidents, access medication details, educational resources, disease progression, and a calendar. (b) Interface for marking swollen joints by patient's perception. (c) Example of a Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) score questionnaire item for disease assessment. (d) List of informational and educational resources available within the app.



(a) (b) (c) (d)

In addition, the IMIDoc app included scenarios for medication management. To achieve this, the professional team carefully consulted the latest clinical practice guidelines for RA and SpA. This comprehensive review of current guidelines informed the creation of tools for monitoring disease activity and managing medication within the app.

A section with educational and informative content to help patients understand their disease and treatment options was also included; specifically, 26 educational articles were generated by clinicians with multimedia components, images and 12 educational videos. Specialists in various fields, including immunology and obstetrics, were engaged to contribute to the development of educational resources. The material produced included illustrated resources designed to improve patient understanding of their condition and the treatment options available. In addition, a diverse range of formats was used to adapt to different preferences. This included the creation of infographics, interactive texts, and videos. Each format was carefully selected to best suit the content being presented, with the aim of enhancing user engagement and understanding.

Another feature of the app involves task and medical appointment reminders in a personalized questionnaire that is complemented by push notifications.

The IMIDoc clinician web platform is a web-based application for health care professionals, primarily clinical rheumatologists [21]. The IMIDoc clinician web platform includes scenarios for monitoring ePROs in a prespecified time frame and generates events or alerts. These features will enable clinicians to respond in near real-time to significant changes in disease activity. Figure 3 shows an image of the screen displaying the data recorded by a patient.

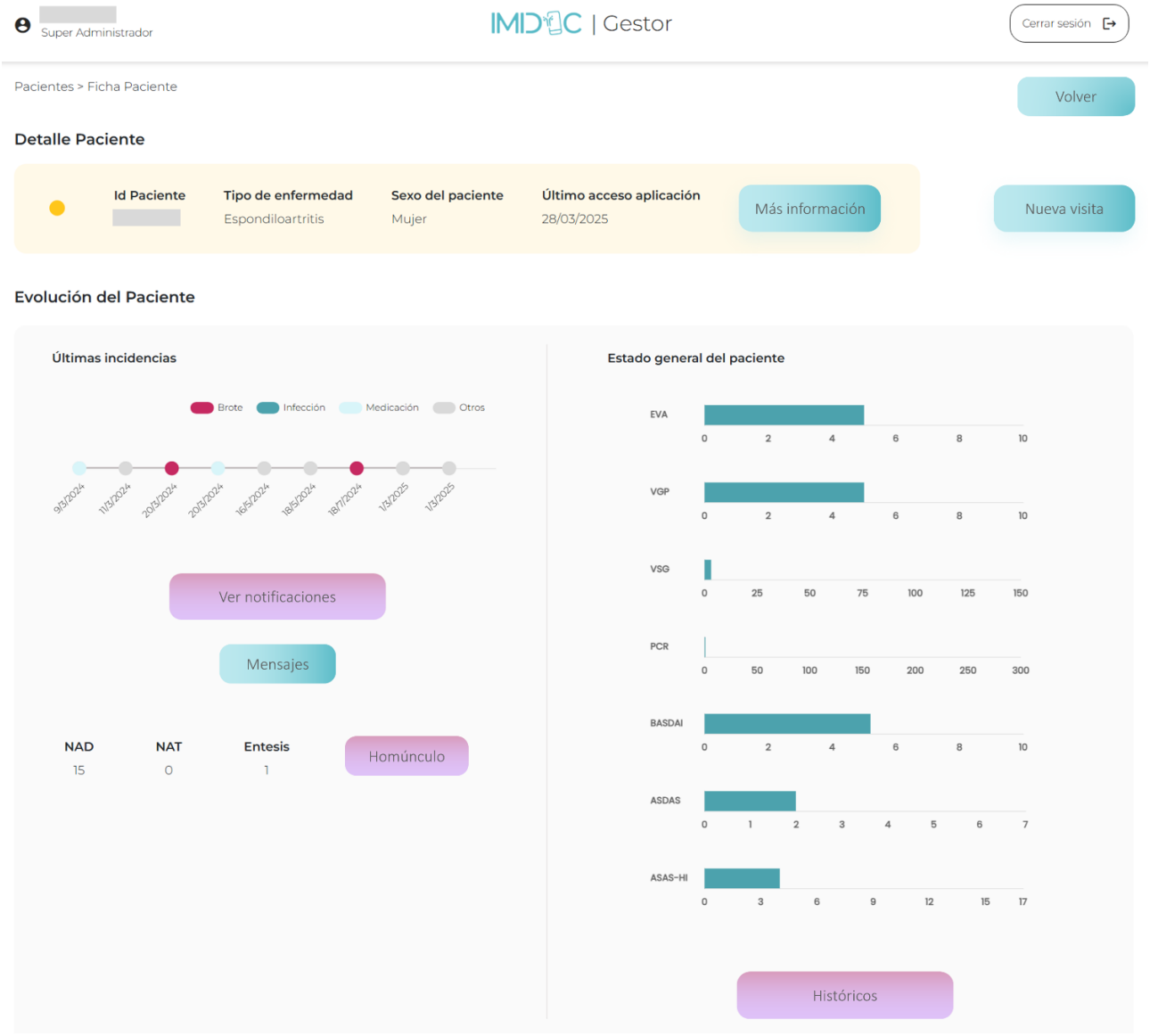
The top section displays basic patient information, including patient code, sex, and last application access. Below, the left panel shows recent incidents (flares, infections, medication issues, or others), the number of swollen joints (NAT, Spanish acronym for “Número de Articulaciones Tumefactas”), the number of tender joints (NAD, Spanish acronym for “Número de Articulaciones Dolorosas”), enthesitis, and access to a homunculus visualization, along with options to view patient messages. The right panel below features bar graphs tracking patient progress, including pain scale (EVA, Spanish acronym for “Escala Visual Analógica”), patient global assessment (VGP, Spanish acronym for “Valoración Global del Paciente”), erythrocyte sedimentation rate (VSG, Spanish acronym for “Velocidad de Sedimentación Globular”), C-reactive protein level (PCR, Spanish acronym for “Proteína C Reactiva”), Bath Ankylosing Spondylitis Disease Activity Index score (BASDAI), and Axial Spondyloarthritis Disease Activity Score (ASDAS).

Through IMIDoc, clinicians could quickly identify the health status of patients and their level of adherence based on the

XSL•FO
RenderX

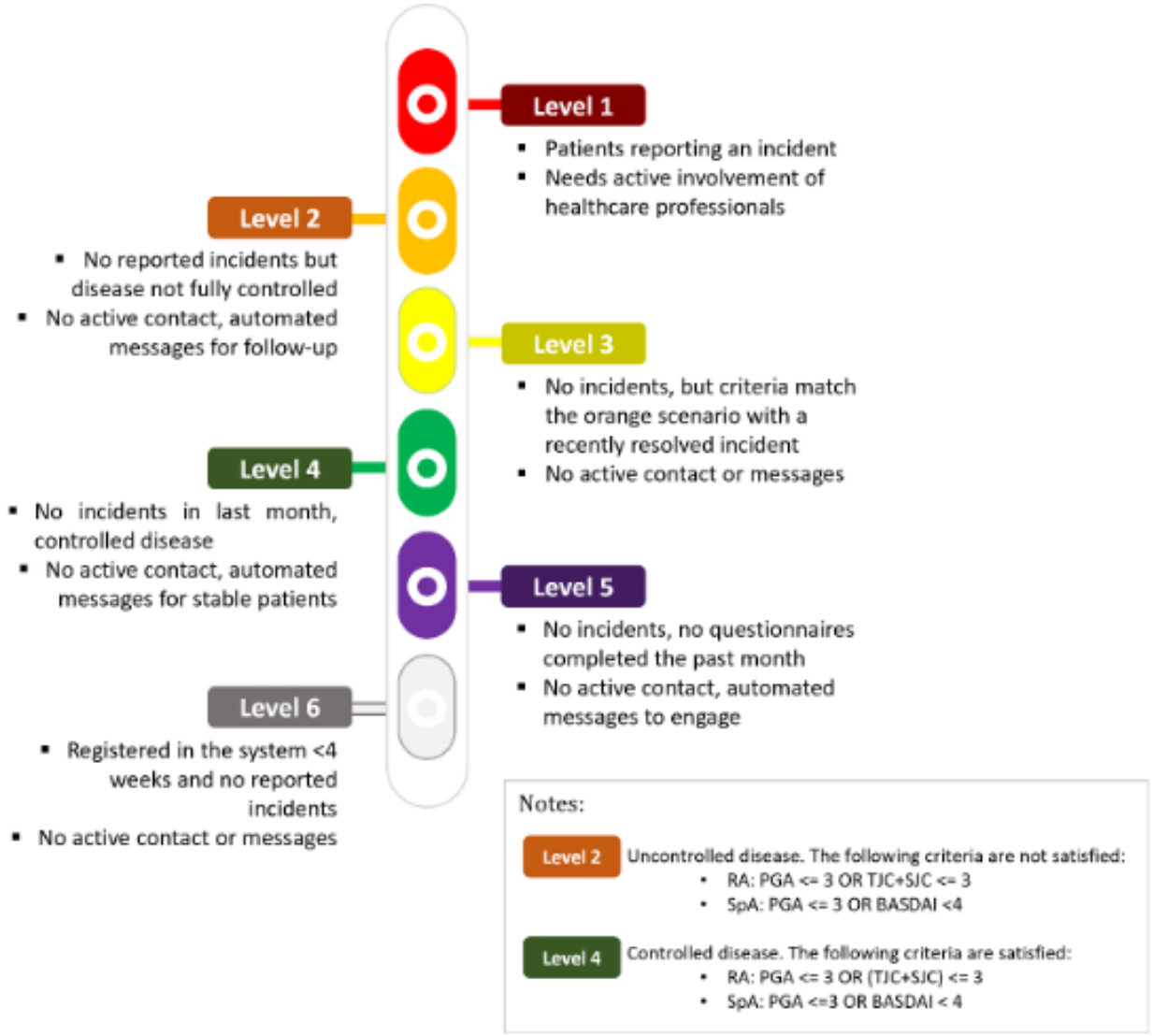
automatic processing of data. The IMIDoc classification algorithm assigns 6 different levels to patients as shown in Figure 4.

Figure 3. Screenshot of the Spanish version of IMIDoc web-based tool for monitoring patients with rheumatoid arthritis and spondyloarthritis.



XSL•FO
RenderX

Figure 4. Classification of the health status and clinical adherence of patients with rheumatoid arthritis and spondyloarthritis. PGA: patient global assessment, TJC: tender joint count, SJC: swollen joint count, BASDAI: bath ankylosing spondylitis disease activity index.



IMIDoc Backend

As mentioned above, there are several platforms available from technology providers such as Amazon Web Services (Amazon), Azure (Microsoft), Google Cloud (Google) and IBM Cloud Computing (IBM), among others. For the implementation of IMIDoc, we selected Amazon Web Services (AWS) to deploy the backend components. In addition, AWS offered a number of features of interest for our implementation [15,16]. We used AWS Fargate which is a serverless computing engine for containers, in which a Node.js-based application has been deployed that features a RESTful API to respond to requests using GET, POST, PUT, DELETE methods via HTTPS. For the storage of ePROs data and clinical actions, a relational database PostgreSQL with Amazon RDS was implemented and deployed in the European region (Spain). An important aspect to consider was security, for which the Security Hub service was enabled. This integrates functionalities to detect and manage security events, in addition to other components such as a Firewall and distributed denial of serviceprotection. Moreover, AWS complies with various certifications, accreditations, European laws and regulations such as ENS (high category), DESC CSP, CISPE and the General Data Protection Regulation (GDPR). It should be noted that these characteristics, in terms of security and data processing, are requirements mandated by the Information Systems Security Office for the approved use of the application in health care. Finally, AWS allows us to use a continuous integration and delivery service and, in this way, allows us to maintain and prepare system updates for subsequent release in production.

The system has modules that run automatically, monitoring the data, establishing the patient's health status, and generating notifications that are sent to the patients' app through the Google Firebase service. Table 1 specifies the types of notifications included in IMIDoc.

Google Analytics was used to track and process dimensions and metrics related to the use of both applications: country, device, operating system, app version and user interactions with the app (number of sessions, duration of sessions, screen views, and interaction times). Google Analytics service and mechanisms were integrated to record the usage metrics of both applications.

XSL•FO
RenderX

Table . IMIDoc mobile app notifications.

Dimension	Trigger	Title	Body
Messages	Response from a health care professional to an incident.	IMIDoc[a] notification	"You have a new message".
Medication	Time of medication intake	IMIDoc medication	"You have medication pending confirmation".
Resources	New reading resource available	IMIDoc multimedia content	"You have new educational resources available".
Questionnaire(s)	New questionnaire(s) to fill out	IMIDoc questionnaires	"You have new questionnaire(s) to fill in".
Medical appointment	One day before the medical appointment	IMIDoc medical appointment	"Remember that you have a doctor's appointment tomorrow, please check that you have provided all the necessary information".
Adherence support	One week without logging into the app	IMIDoc misses you	"We have noticed that you have not been using the application for a while, we encourage you to give us another chance".

[a]IMIDoc: Immune-mediated inflammatory diseases

Preclinical Testing Phase of IMIDoc

The involvement of the multidisciplinary team in the initial testing phase led to several modifications to the digital solution, primarily to the mobile solution. These mainly focused on the user interface, the wording used within the application, and its distribution mechanisms, aiming to ensure that the application was intuitive for both patients and health care professionals.

During the subsequent 12-week digital solution testing phase with 10 active users (7 professionals and 3 patients), several functional issues and areas for potential improvement were identified. Detailed functionality reports for the mobile app recorded 111 entries—76 errors, 16 improvements, and 10 evolutionary suggestions—while the professional tool generated 45 entries, with 40 errors and 5 evolutionary suggestions. Based on this feedback, users' reported difficulties with navigation and data entry prompted a redesign of the interface to reduce logging steps and incorporate clearer prompts. In addition, the need for customizable alerts led to refinements in medication and appointment reminders and improvements in the messaging system to enable prompt clinician responses. These iterative enhancements, delivered over 10 update cycles, set the stage for a multicenter clinical study involving 360 patients from 5 Spanish hospitals.

Discussion

This study underscores the efficacy of using a structured cocreative approach to crafting the IMIDoc digital health application tailored for rheumatology practice. Over the years, the rheumatology community has voiced the demand for refined apps and tools to yield meaningful clinical outcomes [20,22]. Despite advancements in mobile health technology (mHealth) and treatment in rheumatology, there persists a notable gap in continuous patient monitoring [23,24]. Recognizing this unmet need in rheumatology care, particularly during intervals between scheduled consultations, informed the conceptualization of the digital health solution presented here.

Our initial needs assessment drew from past experiences, incorporating feedback from clinicians and patients to identify key features and functionalities of IMIDoc [6]. This information informed requirements and specifications, supporting a participatory design approach complemented by UCD principles to enhance usability, acceptability, and user experience [17]. Embracing such techniques is crucial for crafting more effective health interventions as it incorporates diverse perspectives [25]. A recent systematic review underscores the increasing adoption of cocreated technology solutions in health care [26]. This aligned with our own approach in developing IMIDoc, in which the representation of both patients and clinicians ensured that real needs were met. The involvement of all relevant stakeholders, including biomedical engineers and technicians, throughout design and development fostered communication, transparency, and a design that aligned with users' priorities.

Furthermore, the application of agile methodologies during the development phase, with iterative cycles, proved invaluable. This approach facilitated continuous adjustments and input, minimizing risks and adapting to emerging changes, and proved particularly valuable in a rheumatology clinical setting [27]. While previous literature has acknowledged the benefits of Agile methodologies [28], our study empirically substantiates its efficacy in this specific context. The pilot testing phase provided valuable insights and led to necessary adjustments, highlighting the dynamic and responsive nature of a development cycle based on agile methodologies. As an example, participants frequently reported difficulties in navigating the app and entering data. In response, the interface was redesigned to minimize the number of steps needed to log information and incorporate clearer prompts to guide users. Additional feedback highlighted the need for customizable alerts, leading us to refine medication and appointment reminders, and to improve the messaging system so clinicians could rapidly respond to patient messages.

Challenges in managing the project scope and addressing various complexities led to a shift toward a more strategic and modular

XSL•FO
RenderX

approach, prioritizing essential features and refining functionalities based on end-user needs. To ensure a rigorous assessment of usability and patient satisfaction, instruments such as the SUS and NPS have been implemented and will be applied in the upcoming multicenter clinical study involving 360 patients from 5 Spanish hospitals.

To ensure the long-term sustainability and continuous improvement of IMIDoc, the project team is actively pursuing additional funding avenues beyond the initial grant. This includes applying for public research grants, seeking private partnerships with health care technology companies, and exploring collaborations with pharmaceutical firms interested in supporting digital health initiatives. In addition, we are developing a strategic plan that outlines regular updates and maintenance schedules for the IMIDoc platform to keep pace with technological advancements and evolving clinical needs. By diversifying our funding sources and establishing a robust maintenance framework, we aim to maintain the app's functionality, incorporate user feedback, and expand its features to better serve patients and health care providers in the future.

Previous literature has provided insights into patient perspectives on digital health interventions and has revealed diverse opinions among patients. Moreover, it has highlighted both advantages and concerns regarding telemedicine in rheumatology. A mixed-methods approach study that explored the acceptability and safety of telemedicine in rheumatology during the COVID-19 pandemic highlighted the need for careful patient selection and further consultation with clinicians to improve telemedicine's acceptability and safety [29]. A prospective study including patients referred to a Veterans Affairs rheumatology clinic found that telerheumatology was well-received by patients, and accurately identified all inflammatory rheumatic conditions at the telerheumatology visits. This suggests that remote monitoring may be suitable for screening and prioritizing patients for in-person rheumatology clinics [30]. Recent research has demonstrated the potential for digital tools to yield positive outcomes in health care. Two randomized trials have been conducted in RA. Pers et al [31] compared a smartphone application connected to a physician's interface for monitoring RA with conventional monitoring. The connected monitoring group had significantly fewer physical visits while maintaining tight control of disease activity and improving quality of life [31]. Salaffi et al [32] investigated the effectiveness of a telemonitoring strategy in early RA, showing higher rates of remission by CDAI within a 1 year compared with the conventional group (38.1% vs 25%) [32].

Our approach has certain limitations. At present, the focus is restricted to RA and SpA, which may limit the generalizability of IMIDoc to other rheumatologic disorders. The small number of participants in the pilot, without extensive demographic characterization, constrains broad interpretations of feasibility and usability. Continued funding is crucial for further feature enhancements and to maintain robust data protection measures, ensuring compliance with evolving regulations. Future large-scale evaluations are planned to validate the efficacy of IMIDoc, broaden its applicability to wider patient populations, and strengthen its long-term sustainability.

While patient opinions on telemedicine may vary, these studies have shown the potential for digital tools to improve clinical outcomes, such as achieving disease remission and reducing clinical visits. Therefore, it is essential to embrace these technological advancements by implementing digital health solutions in real-world rheumatology practice. To this end, it is essential to carefully select patients, and work closely with clinicians to ensure that digital health interventions are effectively integrated into clinical practice, ultimately enhancing the care and well-being of rheumatology patients.

Conclusions

The development of IMIDoc underscores the significance of a cocreative design approach in health care, on that can lead to a product that not only excels in functionally, but that is also contextually relevant to rheumatology practice. Addressing unmet needs such as the integration of ePRO collection, tailored alerts, and a robust messaging system into a single platform, IMIDoc provides a digital health platform for comprehensive disease activity tracking, medication management, patient education, and empowerment in self-health management. This innovation is pivotal to navigating the complexities of managing patients with immune‐mediated inflammatory diseases, offering personalized and effective management strategies while reducing the gap between clinic visits. This work strives to be a breakthrough in rheumatology care, and the proposed multicenter study aims to further validate and support these findings.

Acknowledgments

The authors would like to thank all of the health professionals, researchers, and technicians who collaborated in the development of IMIDoc and the patients who participated in the technical validation tests; specifically, We acknowledge IMIDOC Research group and the members: Leticia Lojo-Oliveira, Verónica García, Claudia María Gómez, Carlos Cañado, Jaime Arroyo, Helena Borrell, Mariana Díaz for their dedicated involvement and support throughout the research process and thanks to Laura Miguel Berenguel, Ricardo Cuesta Martín de la Cámara, Ana Martínez Feito, Gabriela Torres, and Mariela Uyaguari for their invaluable contribution to the development of educational materials. The authors would like to acknowledge the PI22/00777 "Hacia la Telemonitorización en Enfermedades Inflamatorias Inmunomediadas: Implementación de un Modelo Asistencial Mixto (MAM) en una Unidad de Terapias Complejas" Project, financed by the Fondo de Investigaciones Sanitarias (FIS) of the Instituto de Salud Carlos III (ISCIII) and cofinanced with EU Structural Funds. We extend our gratitude to the Rheumatology Society of the COmmunity of Madrid (SORCOM), the Catalan Society of Rheumatology (SCR), and the patient associations that contributed to our survey, including CEADE and ConArtritis. Furthermore, we wish to express our appreciation to the dedicated individuals

XSL•FO
RenderX

who participated in the pilot phase of this study, sharing invaluable insights into their experiences with RA and spondyloarthritis. The authors thank the Spanish Foundation of Rheumatology for providing editorial assistance during the preparation of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Patient questionnaire.
[DOCX File, 282 KB - humanfactors_v12i1e58095_app1.docx]

References

1. Cross M, Smith E, Hoy D, et al. The global burden of rheumatoid arthritis: estimates from the global burden of disease 2010 study. Ann Rheum Dis 2014 Jul;73(7):1316-1322. [doi: 10.1136/annrheumdis-2013-204627] [Medline: 24550173]
2. Taurog JD, Chhabra A, Colbert RA. Ankylosing spondylitis and axial spondyloarthritis. N Engl J Med 2016 Jun 30;374(26):2563-2574. [doi: 10.1056/NEJMra1406182] [Medline: 27355535]
3. Stenzel R, Hadaschik K, May S, et al. Digitally-supported patient-centered asynchronous outpatient follow-up in rheumatoid arthritis - an explorative qualitative study. BMC Health Serv Res 2022 Oct 28;22(1):1297. [doi: 10.1186/s12913-022-08619-6] [Medline: 36307779]
4. Seppen BF, Wiegel J, Nurmohamed MT, Bos WH, Ter Wee MM. Facilitators and barriers to adhere to monitoring disease activity with ePROs: a focus group study in patients with inflammatory arthritis. Rheumatol Int 2023 Apr;43(4):677-685. [doi: 10.1007/s00296-022-05263-5] [Medline: 36627396]
5. Austin L, Sharp CA, van der Veer SN, et al. Providing “the bigger picture”: benefits and feasibility of integrating remote monitoring from smartphones into the electronic health record. Rheumatology (Oxford) 2020 Feb 1;59(2):367-378. [doi: 10.1093/rheumatology/kez207] [Medline: 31335942]
6. Benavent D, Fernández-Luque L, Núñez-Benjumea FJ, et al. Monitoring chronic inflammatory musculoskeletal diseases mixing virtual and face-to-face assessments-Results of the digireuma study. PLOS Digit Health 2022 Dec;1(12):e0000157. [doi: 10.1371/journal.pdig.0000157] [Medline: 36812651]
7. Manifesto for agile software development. URL: https://agilemanifesto.org/ [accessed 2023-09-11]
8. Kairós DS. Business agility | new technologies. URL: https://www.kairosds.com/ [accessed 2023-01-03]
9. Becas leonardo a investigadoresy creadores culturales 2022- red leonardo. URL: https://www.redleonardo.es/becas/becas-leonardo-investigadores-creadores-culturales-2022/ [accessed 2023-09-11]
10. Advancing science through pfizer investigator research exchange research awards. URL: https://www.aspireresearch.org/ [accessed 2023-09-11]
11. Páginas - solicitudes. URL: https://www.isciii.es/QueHacemos/Financiacion/solicitudes/Paginas/default.aspx [accessed 2023-09-11]
12. Home | scrum.org. URL: https://www.scrum.org/ [accessed 2023-01-03]
13. Kokol P. Agile software development in healthcare: a synthetic scoping review. Appl Sci (Basel) ;12(19):9462. [doi: 10.3390/app12199462]
14. ISO 9241-210:2019 - Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems. URL: https://www.iso.org/standard/77520.html [accessed 2024-03-30]
15. Koumpouros Y. User-centric design methodology for mHealth apps: the painApp paradigm for chronic pain. Technologies (Basel) 2022;10(1):25. [doi: 10.3390/technologies10010025]
16. Farao J, Malila B, Conrad N, Mutsvangwa T, Rangaka MX, Douglas TS. A user-centred design framework for mHealth. PLoS ONE 2020;15(8):e0237910. [doi: 10.1371/journal.pone.0237910] [Medline: 32813711]
17. Saparamadu A, Fernando P, Zeng P, et al. User-centered design process of an mHealth app for health professionals: case study. JMIR Mhealth Uhealth 2021 Mar 26;9(3):e18079. [doi: 10.2196/18079] [Medline: 33769297]
18. Adams C, Walpola R, Schembri AM, Harrison R. The ultimate question? Evaluating the use of net promoter score in healthcare: a systematic review. Health Expect 2022 Oct;25(5):2328-2339. [doi: 10.1111/hex.13577] [Medline: 35985676]
19. System usability scale (SUS). Agency for Healthcare Research and Quality. 2013 Sep 6. URL: https://digital.ahrq.gov/sites/default/files/docs/survey/systemusabilityscale%2528sus%2529_comp%255B1%255D.pdf [accessed 2025-04-17]
20. Lalloo C, Jibb LA, Rivera J, Agarwal A, Stinson JN. “There’s a Pain App for That”: review of patient-targeted smartphone applications for pain management. Clin J Pain 2015 Jun;31(6):557-563. [doi: 10.1097/AJP.0000000000000171] [Medline: 25370138]
21. IMIDoc. IMIDoc: Home. URL: https://imidoc.com/landing [accessed 2025-04-01]
22. Gossec L, Guyard F, Leroy D, et al. Detection of flares by decrease in physical activity, collected using wearable activity trackers in rheumatoid arthritis or axial spondyloarthritis: an application of machine learning analyses in rheumatology. Arthritis Care Res (Hoboken) 2019 Oct;71(10):1336-1343. [doi: 10.1002/acr.23768] [Medline: 30242992]

XSL•FO
RenderX

23. Seppen BF, den Boer P, Wiegel J, et al. Asynchronous mHealth interventions in rheumatoid arthritis: systematic scoping review. JMIR Mhealth Uhealth 2020 Nov 5;8(11):e19260. [doi: 10.2196/19260] [Medline: 33151161]
24. Mollard E, Michaud K. A mobile app with optical imaging for the self-management of hand rheumatoid arthritis: pilot study. JMIR Mhealth Uhealth 2018 Oct 29;6(10):e12221. [doi: 10.2196/12221] [Medline: 30373732]
25. Stock C, Dias S, Dietrich T, Frahsa A, Keygnaert I. Editorial: how can we co-create solutions in health promotion with users and stakeholders? Front Public Health 2021;9:773907. [doi: 10.3389/fpubh.2021.773907] [Medline: 34957026]
26. Merchán-Baeza JA, Borralleras Andreu C, Minobes-Molina E, Grau Carrión S, Romero-Mas M, Ramon-Aribau A. Co-created technological solutions for caregivers in health care: systematic review. J Med Internet Res 2023 May 1;25:e41260. [doi: 10.2196/41260] [Medline: 37126384]
27. Khambari M, Rofie FSM. Augmented reality mobile application co-design experience: delineating the nuances from the lens of a system developer. ASS 2021;17(11):198. [doi: 10.5539/ass.v17n11p198]
28. Tsangaris E, Edelen M, Means J, et al. User-centered design and agile development of a novel mobile health application and clinician dashboard to support the collection and reporting of patient-reported outcomes for breast cancer care. BMJ Surg Interv Health Technol 2022;4(1):e000119. [doi: 10.1136/bmjsit-2021-000119] [Medline: 35464815]
29. Sloan M, Lever E, Harwood R, et al. Telemedicine in rheumatology: a mixed methods study exploring acceptability, preferences and experiences among patients and clinicians. Rheumatology (Oxford) 2022 May 30;61(6):2262-2274. [doi: 10.1093/rheumatology/keab796] [Medline: 34698822]
30. Nguyen-Oghalai TU, Hunter K, Lyon M. Telerheumatology: the VA experience. South Med J 2018 Jun;111(6):359-362. [doi: 10.14423/SMJ.0000000000000811] [Medline: 29863228]
31. Pers YM, Valsecchi V, Mura T, et al. A randomized prospective open-label controlled trial comparing the performance of A connected monitoring interface versus physical routine monitoring in patients with rheumatoid arthritis. Rheumatology (Oxford) 2021 Apr 6;60(4):1659-1668. [doi: 10.1093/rheumatology/keaa462] [Medline: 33020846]
32. Salaffi F, Carotti M, Ciapetti A, et al. Effectiveness of a telemonitoring intensive strategy in early rheumatoid arthritis: comparison with the conventional management approach. BMC Musculoskelet Disord 2016 Apr 2;17(1):146. [doi: 10.1186/s12891-016-1002-2] [Medline: 27038788]

Abbreviations

AWS: Amazon Web Services
ePROs: electronic Patient-Reported Outcomes
GDPR: General Data Protection Regulation
MAM: Mixed Care Model
mHealth: mobile health
NPS: Net-Promoter Score
RA: rheumatoid arthritis
SUS: System Usability Scale

Edited by A Kushniruk; submitted 08.03.24; peer-reviewed by A Saraux, J Zhao, M Gasmi, T Orlemann; revised version received 17.02.25; accepted 04.03.25; published 21.04.25.

Please cite as:
Benavent D, Iniesta-Chamorro JM, Novella-Navarro M, Pérez-Martínez M, Martínez-Sánchez N, Kaffati M, Juárez-García M, Molinari-Pérez M, González-Torbay A, Gutiérrez M, López-Juanes N, Navarro-Compán V, Monjo-Henry I, Rodríguez-Rosales G, Bachiller J, Calvo-Aranda E, Michelena X, Berbel-Arcobé L, Balsa A, IMIDOC Research Team, Gómez EJ, Plasencia-Rodríguez C Digital Health Intervention for Patient Monitoring in Immune-Mediated Inflammatory Diseases: Cocreation and Feasibility Study of the IMIDoc Platform
JMIR Hum Factors 2025;12:e58095
URL: https://humanfactors.jmir.org/2025/1/e58095
doi:10.2196/58095

© Diego Benavent, Jose M Iniesta-Chamorro, Marta Novella-Navarro, Miguel Pérez-Martínez, Nuria Martínez-Sánchez, Mónica Kaffati, Manuel Juárez-García, Marina Molinari-Pérez, Andrea González-Torbay, Mariana Gutiérrez, Natalia López-Juanes, Victoria Navarro-Compán, Irene Monjo-Henry, Germán Rodríguez-Rosales, Javier Bachiller, Enrique Calvo-Aranda, Xabier Michelena, Laura Berbel-Arcobé, Alejandro Balsa, IMIDOC Research Team, Enrique J Gómez, Chamaida Plasencia-Rodríguez. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 21.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors,

XSL•FO
RenderX

is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Understanding Dermatologists' Acceptance of Digital Health Interventions: Cross-Sectional Survey and Cluster Analysis

Patrick Reinders, MSc; Matthias Augustin, Prof Dr Med; Marina Otten, PhD

Institute for Health Services Research in Dermatology and Nursing (IVDP), University Medical Center Hamburg-Eppendorf (UKE), Martinistraße 52, Hamburg, Germany

Corresponding Author:
Marina Otten, PhD
Institute for Health Services Research in Dermatology and Nursing (IVDP), University Medical Center Hamburg-Eppendorf (UKE), Martinistraße 52, Hamburg, Germany

Abstract

Background: Digital health interventions (DHIs) have the potential to enhance dermatological care by improving quality, patient empowerment, and efficiency. However, adoption remains limited, particularly in Germany.

Objective: This study explores German dermatologists' attitudes toward DHIs, clustering them by acceptance levels and analyzing differences in sociodemographics and current and future DHI use.

Methods: We conducted a cross-sectional survey, randomly inviting 1000 dermatologists in Germany to participate. The questionnaire consisted of Likert scale items rating the acceptability of DHIs from 1 to 5. Items on the current and future use of DHIs were also included. Exploratory factor analysis was used to identify factors and reduce data as input for a 2-step clustering algorithm.

Results: The survey with 170 dermatologists (mean age 50.8, SD 10.3 y; 74/167, 55.7% female) identified four factors through the exploratory factor analysis: (1) "Positive Expectancies and Acceptability of DHIs," (2) "Dermatologists' Digital Competencies," (3) "Negative Expectancies and Barriers," and (4) "Dermatologists' Perspectives on Patients' Acceptability and Competencies." The analysis identified three distinct clusters: (1) Indecisives (n=69)—moderate intentions to use DHIs and moderate negative expectations toward them; (2) Adopters (n=60)—high intentions to use DHIs and high digital competencies; and (3) Rejectors (n=26)—low intentions to use DHIs and low digital competencies. Adopters were significantly younger, more often based in urban centers, and exhibited the highest adoption rates of DHIs compared to the other clusters. Across all clusters, inadequate reimbursement and perceived structural barriers were cited as significant challenges to DHI adoption. Still, only one-third of the Adopters used DHIs including teledermatology or artificial intelligence.

Conclusions: Dermatologists in Germany exhibited varied levels of acceptance and readiness for DHIs, with demographic and structural factors influencing adoption. Addressing barriers such as reimbursement and investing in digital literacy could promote wider use, potentially reducing health inequalities by improving access to digital health care.

(JMIR Hum Factors 2025;12:e59757) doi:10.2196/59757

KEYWORDS

dermatology; cross-sectional survey; cluster analysis; acceptability; eHealth; Germany; attitude; dermatologist; digital health intervention; teledermatology

Introduction

Dermatology is marked by the presence of prevalent chronic conditions, including acne, psoriasis, atopic dermatitis, and skin cancer [1,2]. In Germany, approximately 27% of adults have at least one dermatological condition requiring specialized care each year, resulting in long waiting times for patients [3,4]. Additionally, limited capacity also impacts comprehensive patient consultations, and the demographic shift could exacerbate these challenges [5-7].

Digital health interventions (DHIs) are proposed as one mitigating solution to sustain health care by enhancing the quality of care, efficiency, and empowering patients [8,9]. In this study, DHIs are defined as health services that use information and communication technologies for patients, consumers, or health care professionals [9,10]. In dermatology, numerous promising DHIs have already been developed [11], including artificial intelligence (AI) [12], teledermatology, diagnoses or monitoring tools [13-16], and patient-targeted apps for medication adherence [17]. The 2020 consensus guideline for teledermatology recommends its use for certain indications but also stresses the necessity for stronger evidence to evaluate benefits and risks [18]. Currently, no guidance exists for other apps.

XSL•FO
RenderX

Digitalization of health care in Germany lags behind that of other European nations [19]. Teledermatology, for example, is less frequently adopted in Germany [20]. The complexity of the situation is grounded in the intricacies of the German health care system, encompassing aspects such as missing interoperability, missing incentives for physicians, and high standards of data protection [21,22].

Apart from the system perspective, the reluctance of medical professionals to embrace digital health may be another potential barrier for the implementation [21,22], but can also function as a facilitator. Their acceptability can compensate for low patient demand or financial resource limitations and can additionally improve patient adherence to these interventions [23-26]. Dermatologists' overall acceptability of DHIs appears to be lower than that of patients with a dermatological condition [27]. Nevertheless, it is noteworthy that dermatologists differ in their perspectives and adoption of DHIs, shaped by various factors including younger age, practicing in urban settings, positive prior experiences with DHIs, or the social influence of colleagues [28-30].

The objective of this study was to explore the acceptability and current and future use of DHIs among dermatologists. Additional objectives included identifying and describing clusters of dermatologists based on their acceptability and examining differences between these clusters and sociodemographic variables, as well as the current and future use of DHIs.

Decision makers or developers of DHIs can use these results to tailor strategies for a successful implementation.

Methods

The consensus-based Checklist for Reporting of Survey Studies and the Checklist for Reporting Results of Internet E-Surveys were applied [31,32]. The full checklist is displayed in Checklist 1.

Study Design and Questionnaire

An anonymous cross-sectional survey among German dermatologists was conducted in June 2022. The questionnaire was based on existing questionnaires from the literature, and in particular, on preconducted focus groups among dermatologists, patients, and nurses (not yet published). The focus groups dealt with the acceptability, barriers, and facilitators of DHIs in dermatology using the Unified Theory of Acceptance and Use of Technology [33,34]. The questionnaire development resulted in 27 items related to the acceptability of DHIs in general (eg, "I could imagine using more digital applications for the care of my patients"). Five items on the potential future use of specific DHIs (eg, "medication reminder application") and four items on the benefits and costs of the nationwide eHealth Strategy (eg, "The eHealth Strategy has strengthened patient care") were added to the survey. All items were answered on a 5-point Likert scale (5=strongly agree; 1=strongly disagree). The survey also included 10 items about dermatologists' current use (>1 per wk) of available DHIs (eg, live-interactive teledermatology or electronic doctor letters) with a simple "check all that apply" option.

In addition, sociodemographic information covering age, sex, type of employment (eg, working in an outpatient clinic or medical practice), and postcode of the clinic or practice were collected. The postcode was used to assign additional regional data to participants (eg, urban or rural areas). The questionnaire was tested with 5 dermatologists of the Institute for Health Services Research in Dermatology and Nursing (German: Institut für Versorgungsforschung in der Dermatologie und bei Pflegeberufen). The final questionnaire is provided in Multimedia Appendix 1.

The offline version of this survey was designed to fit two DIN-A4 pages, while the web-based version was adapted to six screens or pages.

Study Population, Recruitment, and Data Entry

In order to conduct an exploratory factor analysis (EFA; statistical analysis), we aimed for a minimum sample size of 150 dermatologists [35]. To reach the required sample size, the questionnaire was randomly distributed to 1000 members of the German Dermatological Academy (German: Deutsche Dermatologische Akademie [DDA] with ~3500 active dermatologists) in June 2022. The DDA comprises members of the Federal Association of German Dermatologists (German: Berufsverband der Deutschen Dermatologen), representing dermatologists from medical practices, and of the German Dermatological Society (German: Deutsche Dermatologische Gesellschaft), representing dermatologists from both outpatient clinics and medical practices. Dermatologists were asked to participate via email (Unipark), receiving one reminder or offline via mail on paper. The web-based survey was an open survey (no password protection). Additionally, participants could review their responses before submission. The survey was voluntary, and no incentives were offered. The survey was terminated after 8 weeks. The paper-based data were digitized, double-checked, corrected, and combined with the data of the web-based version. To prevent multiple entries by a single dermatologist in the web-based survey, participation was tracked using Unipark's cookie system. To identify and eliminate further duplicate entries from the web-based and offline survey, demographic information and postcodes were cross-referenced and thoroughly checked. No duplicate entries were identified.

Ethical Considerations

The study was submitted to the local ethics committee and was waived due to its noninvasive and anonymous nature. The study was reviewed and approved by the Lokale Psychologische Ethikkommission am Zentrum für Psychosoziale Medizin (LPEK), under the reference LPEK-0407. The study was conducted in accordance with Good Scientific Practice and the Declaration of Helsinki [36,37]. Due to the anonymous nature of the study, informed consent was not collected. Participants received no compensation for their participation and had the option to opt out of the survey at any time.

Statistical Analysis

After conducting a descriptive analysis of all sociodemographic and geographic parameters, an EFA including all 27 items related to acceptability was performed for two reasons: (1) for data reduction by identifying underlying factors, also commonly

XSL•FO
RenderX

used and recommended prior to running a cluster analysis; and (2) to use the underlying factors to closer describe and explore dermatologists' acceptability toward DHIs [38,39].

Adequacy of the EFA for factor analysis was assumed with a Kaiser-Meyer-Olkin criterion (KMO) ≥0.8 and Bartlett's test of sphericity <0.05. All factors with an eigenvalue ≥1.0 were considered. Criteria for retaining an item were a substantial factor loading (>|0.4|) on one factor and no strong cross-loading onto other factors. Cross-loading was defined as loadings (>|0.3|) on another factor and a difference of at least 0.2 between the loading on the main factor and the strongest loading on any other factor [40]. Internal consistencies of the factors were measured by Cronbach α, where values >0.7 are acceptable, >0.8 are good, and >0.9 are considered excellent.

To identify distinct clusters of dermatologists that differ in their acceptability, a combination of hierarchical and K-means cluster analysis was applied. The mean scores (using Likert scale ratings: 1–5) of the factors were used as input variables for the 2 approaches. The hierarchical cluster approach was used to identify the optimal number of clusters. Subsequently, a K-means cluster analysis was conducted with this optimal number. To test on differences between the identified clusters based on mean scores, a one-way ANOVA was performed. In order to create a profile of the identified clusters, first, each item of the EFA (acceptability items) was binary coded (1=strongly agree or agree; 0=neither, disagree, strongly disagree). Then, a descriptive analysis was conducted, analyzing each item for each cluster. Next, chi-square tests were applied for categorical variables to identify differences in sociodemographic parameters, as well as current and potential future use of DHIs between clusters. In case more than 20% of cells had a count below five, Fisher's exact tests were used. For continuous variables, one-way ANOVAs were used. A significance level of .05 was applied. Missing rates were reported for all variables. No imputation methods or weighting methods were applied.

Results

Participant Characteristics

In the survey, 170 dermatologists participated, resulting in a response rate of 17% (170/1000 invited dermatologists). In 131 (77.1%) cases, participation took place online. The online participation rate was 75.7%, calculated as the ratio of unique visitors who agreed to participate (131) to the total number of unique visitors who viewed the first survey page (173). Participants had a mean age of 50.8 (SD 10.3) years, 55.7% (74/167) were female, and for 84.2% (117/139) of participants, the practice or clinic was located in an urban county or city (Table 1). Missing data on all items were below 5% (9/170), except for regional variation (31/170, 18.2%) and outpatient clinic or medical practice (17/170, 10%).

XSL•FO
RenderX

Table . Sociodemographic characteristics of the total population and identified clusters.

	Total population[a] (n=170)	Rejectors (Cluster 1; n=26)	Indecisives (Cluster 2; n=69)	Adopters (Cluster 3; n=60)	*P* value
Age group (years), n (%)					<.001
25 - 39	24 (14.9)	0 (0)	5 (7.8)	17 (29.3)	
40 - 49	33 (20.5)	2 (8.3)	14 (21.9)	13 (22.4)	
50 - 59	71 (44.1)	9 (37.5)	39 (60.9)	18 (31)	
60 and older	33 (20.5)	13 (54.2)	6 (9.4)	10 (17.2)	
Missing	9 (5.3)	2 (7.7)	5 (7.2)	2 (3.3)	
Age, mean (SD)	50.8 (10.3)	58.8 (6.7)[b,c]	51.4 (7.9)[b,d]	46.9 (11.7)[b,d]	<.001
Sex, n (%)					.33
Female	74 (44.3)	12 (46.2)	25 (37.9)	32 (53.3)	
Male	93 (55.7)	14 (53.8)	41 (62.1)	28 (46.7)	
Missing	3 (1.8)	0 (0)	3 (4.3)	0 (0)	
Regional variation, n (%)					.01
Urban county or city	117 (84.2)	11 (61.1)	46 (82.1)	51 (91.1)	
Rural county	22 (15.8)	7 (38.9)	10 (17.9)	5 (8.9)	
Missing	31 (18.2)	8 (30.8)	13 (18.8)	4 (6.7)	
Outpatient clinic or medical practice, n (%)					<.001
Medical practice	108 (70.6)	26 (100)	54 (80.6)	28 (46.7)	
Outpatient clinic	45 (29.4)	0 (0)	13 (19.4)	32 (53.3)	
Missing	17 (10.0)	0 (0)	2 (2.9)	0 (0)	

[a]15 participants could not be used in the cluster analysis, due to missing values of the used subscores.

[b]Different from Cluster 2 (Adopters); $P<.05$.

[c]Different from Cluster 1 (Indecisives); $P<.05$.

[d]Different from Cluster 3 (Rejectors); $P<.05$.

Factors of Acceptability

The KMO criterion (0.92) and Bartlett's test of sphericity ($\chi^2_{210}=2264$; $P<.001$) revealed that the data were appropriate to conduct an EFA. Four factors with an eigenvalue >1 were identified (Table S1 in Multimedia Appendix 2). Cronbach ɑ values (0.93, 0.90, 0.73, and 0.80) indicated at least an acceptable internal consistency for all factors. The explained total variance of the EFA was 67.6%. A total of 6 items were excluded with a factor loading below 0.4. We named the factors as follows:

- Factor 1: "Positive Expectancies and Acceptability of DHIs" comprised of 10 items.
- Factor 2: "Dermatologists' Digital Competencies" comprised of 5 items.
- Factor 3: "Negative Expectancies and Barriers of DHIs" comprised of 4 items.
- Factor 4: "Dermatologists' Perspectives on Patients' Acceptability and Competencies" comprised of 2 items.

The four mean scores of the factors ranged between 2.9 (SD 0.8) for Factor 4 and 3.5 (SD 0.8) for Factor 2 (Table 2). Ranging from 1 to 5, higher scores indicate a higher agreement with the items of a factor, except for Factor 3, where it is the other way around.

XSL•FO
RenderX

Table . Differences in identified mean scores on acceptability between the clusters.

	Total population (n=170), mean (SD)[a]	Rejectors (Cluster 1; n=26), mean (SD)	Indecisives (Cluster 2; n=69), mean (SD)	Adopters (Cluster3; n=60), mean (SD)	*P* value (ANOVA)
Factor 1: Positive expectancies and acceptability of DHIs[b]	3.4 (0.9)	2.0 (0.6)[c,d]	3.2 (0.6)[c,e]	4.1 (0.3)[c,e]	<.001
Factor 2: Dermatologists' digital competencies	3.5 (0.8)	1.9 (0.7)[c,d]	3.4 (0.6)[c,e]	4.3 (0.4)[c,e]	<.001
Factor 3: Negative expectancies and barriers	3.4 (0.5)	3.8 (0.3)[c,d]	3.5 (0.4)[c,e]	3.0 (0.5)[c,e]	<.001
Factor 4: Dermatologists' perspectives on patients' acceptability and competencies	2.9 (0.8)	2.2 (0.7)[c,d]	2.7 (0.6)[c,e]	3.5 (0.6)[c,e]	<.001

[a]15 participants could not be used in the cluster analysis, due to missing values of the used subscores.

[b]DHI: digital health intervention.

[c]Different from Cluster 2 (Adopters); *P*<.05.

[d]Different from Cluster 1 (Indecisives); *P*<.05.

[e]Different from Cluster 3 (Rejectors); *P*<.05.

Clusters of Dermatologists According to Acceptability Levels

After exploring 2-, 3-, and 4-cluster solutions, the hierarchical cluster analysis recommended the 3-cluster solution as it provides the best fit based on the K-means evaluation. The performed ANOVA and subsequent post hoc tests revealed that all differences between the four mean scores of the three clusters were significant.

Cluster 1 (n=26, 16.7%) exhibited the lowest willingness (Factor 2: mean 1.9, SD 0.7) and digital competencies (Factor 2: mean 1.9, SD 0.7). Correspondingly, this cluster demonstrated the highest negative expectancies of DHIs (Factor 3: mean 3.8, SD 0.3) and low expectations regarding patients' adequate competencies and willingness to use DHIs (Factor 4: mean 2.2, SD 0.7).

Cluster 2 (n=69, 44.5%) showed moderate positive expectancies toward DHIs (Factor 1: mean 3.2, SD 0.6), along with moderate digital competencies (Factor 2: mean 3.4, SD 0.6) and also moderate negative expectancies (Factor 3: mean 3.5, SD 0.4). Cluster 2's expectations of patients' intention to use DHIs and of patients' digital competencies were low (Factor 4: mean 2.7, SD 0.6).

Cluster 3 (n=60, 38.7%) had higher intentions (Factor 1: mean 4.1, SD 0.3), greater competencies (Factor 2: mean 4.3, SD 0.4), and lower, though still moderate, negative expectancies (Factor 3: mean 3.0, SD 0.5) compared to Clusters 1 and 2. The cluster's opinion about patients' intention and competencies was higher, though still moderate (Factor 4: mean 3.5, SD 0.6).

Based on these differences, we will henceforth refer to Cluster 1 as Rejectors, Cluster 2 as Indecisives, and Cluster 3 as Adopters in the subsequent sections.

Differences in Sociodemographic Information Between Clusters

The Adopters (mean age 46.9, SD 11.7 y) were significantly younger compared to both the Indecisives (mean age 51.4, SD 7.9 y; *P*=.03) and Rejectors (mean age 58.8, SD 6.7 y; *P*<.001). No significant differences in sex were observed among the clusters (*P*=.33). Notably, an association between regional variation (urban or rural) and the three clusters was identified (*P*=.01): the Rejectors had the highest proportion of participants from rural counties (7/18, 38.9%), while the Adopters had the highest percentage of participants from urban counties or cities (51/56, 91.1%). In addition, Indecisives (54/67, 80.6%) and Rejectors (26/26, 100%) performed their services primarily in medical practices, whereas Adopters (32/60, 53.3%) performed theirs in outpatient clinics (*P*<.001).

Acceptability Profiles of the Clusters

The analysis of individual items revealed distinct profiles of the three clusters on their attitude toward DHIs. The following paragraphs describe selected items for each factor and cluster. A comprehensive overview of all item ratings is presented in Figure 1.

Nearly all Adopters expressed openness to using (Item 2: 59/60, 98.3%) and recommending DHIs (Item 3: 58/60, 96.7%) in the future (Figure 1; Factor 1). However, a lower proportion was willing to pay for DHIs (Item 10: 41/60, 68.3%). They rated themselves highly in digital media skills (Item 12: 55/60, 91.7%) and expressed confidence in the easy integration of DHIs into daily routines (Item 14: 53/60, 88.3%; Factor 2). About 45% (27/60; Item 17) feared an increase in the time required for patient care due to DHIs, but only a minority saw the effort for IT as a burden (Item 19: 10/60, 16.7%; Factor 3). A majority agreed that patients would accept DHIs (Item 20: 36/60, 60.1%), while a minority agreed that patients would find using DHIs easy (Item 21: 19/60, 31.7%; Factor 4).

XSL•FO
RenderX

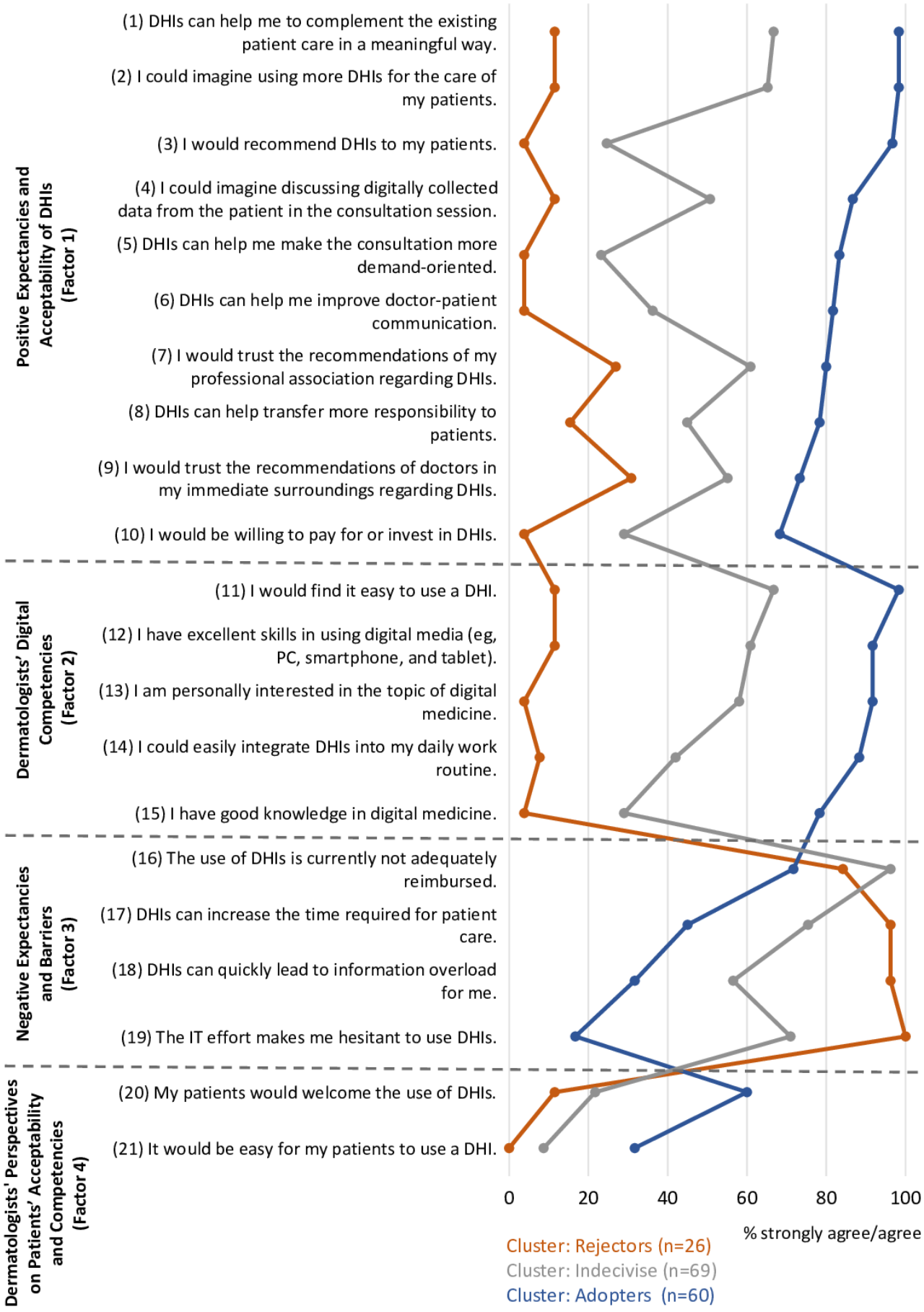
About two-thirds of Indecisives were willing to use DHIs in the future (Item 2: 45/69, 65.2%), but a minority would recommend (Item 3: 17/69, 24.6%) or invest in DHIs (Item 10: 20/69, 29%; Factor 1). Although the majority would find it easy to use a DHI (Item 11: 46/69, 66.7%), few considered themselves knowledgeable in digital medicine (Item 15: 20/69, 29%; Factor 2). The majority perceived IT effort as a burden (Item 19: 49/69, 71%), expressed concerns about DHIs increasing consultation time (Item 17: 52/69, 75.4%), and causing information overload (Item 18: 39/69, 56.5%; Factor 3). Only 21.7% (15/69) believed patients would welcome DHIs (Item 20), and 8.7% (6/69; Item 21) thought patients would find them easy to use (Factor 4).

Rejectors scored low in Factor 1, with the highest agreement rates in trust toward recommendations of other physicians (Item 9: 8/26, 30.8%) and professional associations (Item 7: 7/26, 26.9%; Figure 1; Factor 1). Only 11.5% (3/26; Item 12) agreed to have high competencies in digital medicine (Factor 2). Almost all Rejectors worried about increased time requirements for patients (Item 17: 25/26, 96.2%) and information overload (Item 18: 25/26, 96.2%). They perceived high IT effort as a barrier to DHIs (Item 19: 26/26, 100%; Factor 3). Only 11.5% (3/26; Item 20) believed patients would welcome DHIs, and none (Item 21: 0%) thought patients would find them easy to use (Factor 4).

Overall, ratings on items in Factors 1, 2, and 3 were distinctly different among the three clusters, with the exception that the majority within all clusters agreed that DHIs are inadequately reimbursed (Figure 1; Factor 3). For Factor 4, Indecisives and Rejectors showed similar low agreement rates on a per-item level (Figure 1).

XSL•FO
RenderX

Figure 1. Cluster profile analyses on an item-by-item level. DHI: digital health intervention.



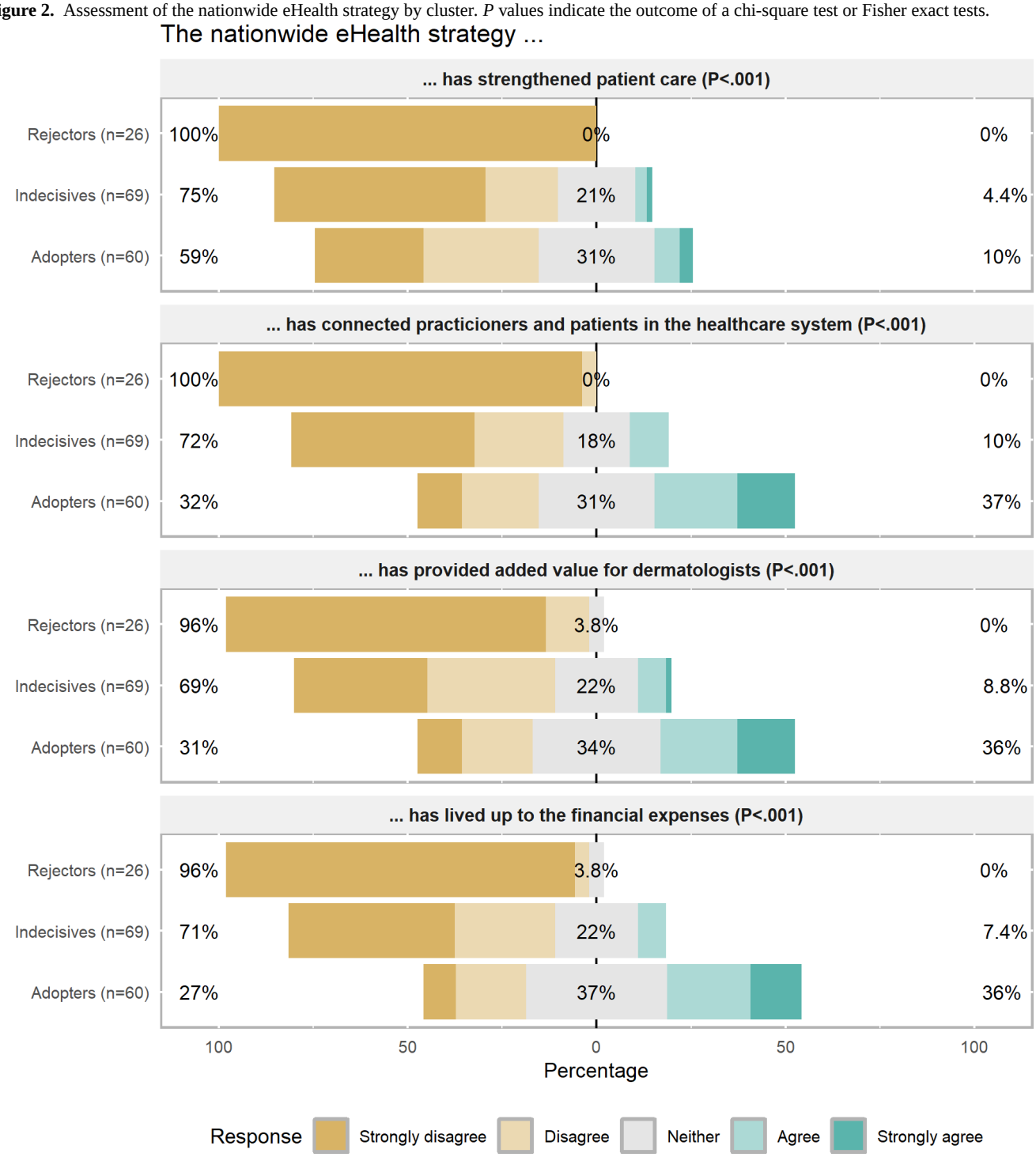
XSL•FO
RenderX

Assessment of the Nationwide eHealth Strategy

In the assessment of the nationwide eHealth Strategy, most dermatologists across all clusters did not perceive the strategy as successful (Figure 2). A majority within each cluster concluded that the eHealth Strategy did not enhance patient care (Indecisives: 52/69, 75%; Adopters: 35/60, 58.3%; and Rejectors: 26/26, 100%).

However, significant variations could still be observed among the three clusters. All Rejectors disagreed on the strategy being effective in fostering connections between patients and providers (26/26, 100%). The Indecisives shared similar doubts (50/69, 72.4%). This behaved similarly with regard to the strategy being cost-effective (49/69, 71% disagreement for Indecisives; 26/26, 100% for Rejectors) and providing added value to dermatologists (48/69, 69.1% disagreement for Indecisives; 25/26, 96.2% for Rejectors). In contrast, the “Adopters” were more divided in their assessment: one-third of the cluster agreed with the statements, another third disagreed, and the remaining third was inconclusive. For example, 37.1% (22/59) agreed, 32.3% (19/59) disagreed, and 30.6% (18/59) were inconclusive on the statement regarding whether the strategy has connected patients and physicians.

Figure 2. Assessment of the nationwide eHealth strategy by cluster. *P* values indicate the outcome of a chi-square test or Fisher exact tests.

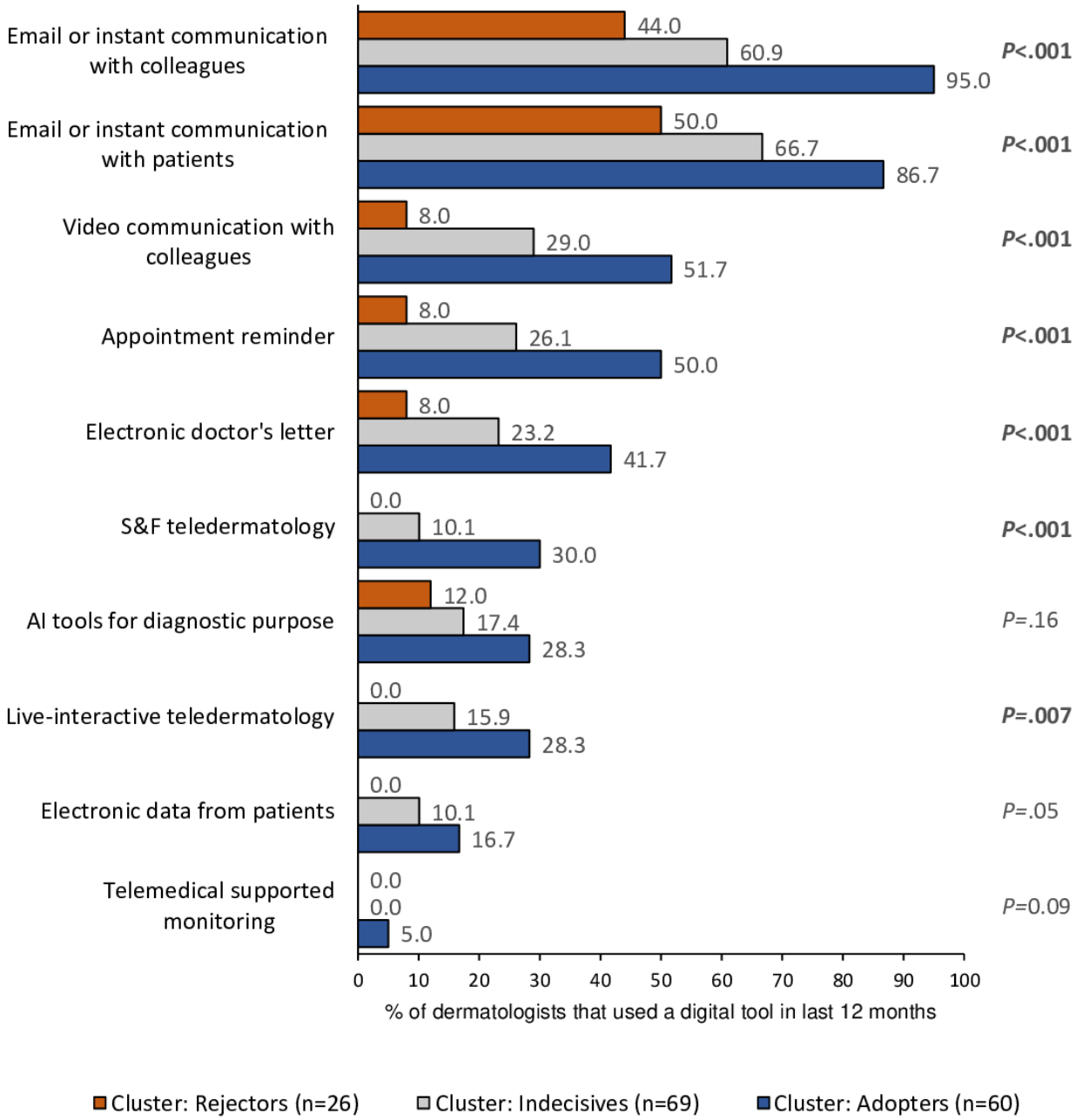


XSL•FO
RenderX

Current Use of DHIs Within Identified Clusters

The use of common DHIs varied significantly between the three clusters, except for AI tools for diagnostic purposes (range: Rejectors 12%-Adopters 28.3%; *P*=.15), electronic patient data (range: Rejectors 0%-Adopters 16.7%; *P*=.05), and telemedical supported monitoring (range: Rejectors 0%-Adopters 5%; *P*=.09; Figure 3). The Adopters demonstrated the highest and the Rejectors the lowest utilization rates across all presented DHIs. A total of 95% (57/60) of the Adopters reported to communicate via email or instant messaging with colleagues, and 86.7% (52/60) used email to communicate with patients. In contrast, only 44% (11/25) and 50% (13/26) of the Rejectors engaged in these communication methods. Apart from email communication, less than half of the Adopters and one-third of the Indecisives used other DHIs. For example, "Appointment Reminders" were used by 50% (30/60) of the Adopters and 26.1% (18/69) of the Indecisives. The Rejectors did not use many of the DHIs at all, such as store-and-forward (0%) and live-interactive teledermatology (0%). Additionally, their utilization of AI for diagnostic purposes was very low (3/25, 12%).

Figure 3. Use of digital technology in the last 12 months by clusters. *P* values indicate the outcome of a chi-square test or Fischer exact tests. AI: artificial intelligence; S&F: store-and-forward.

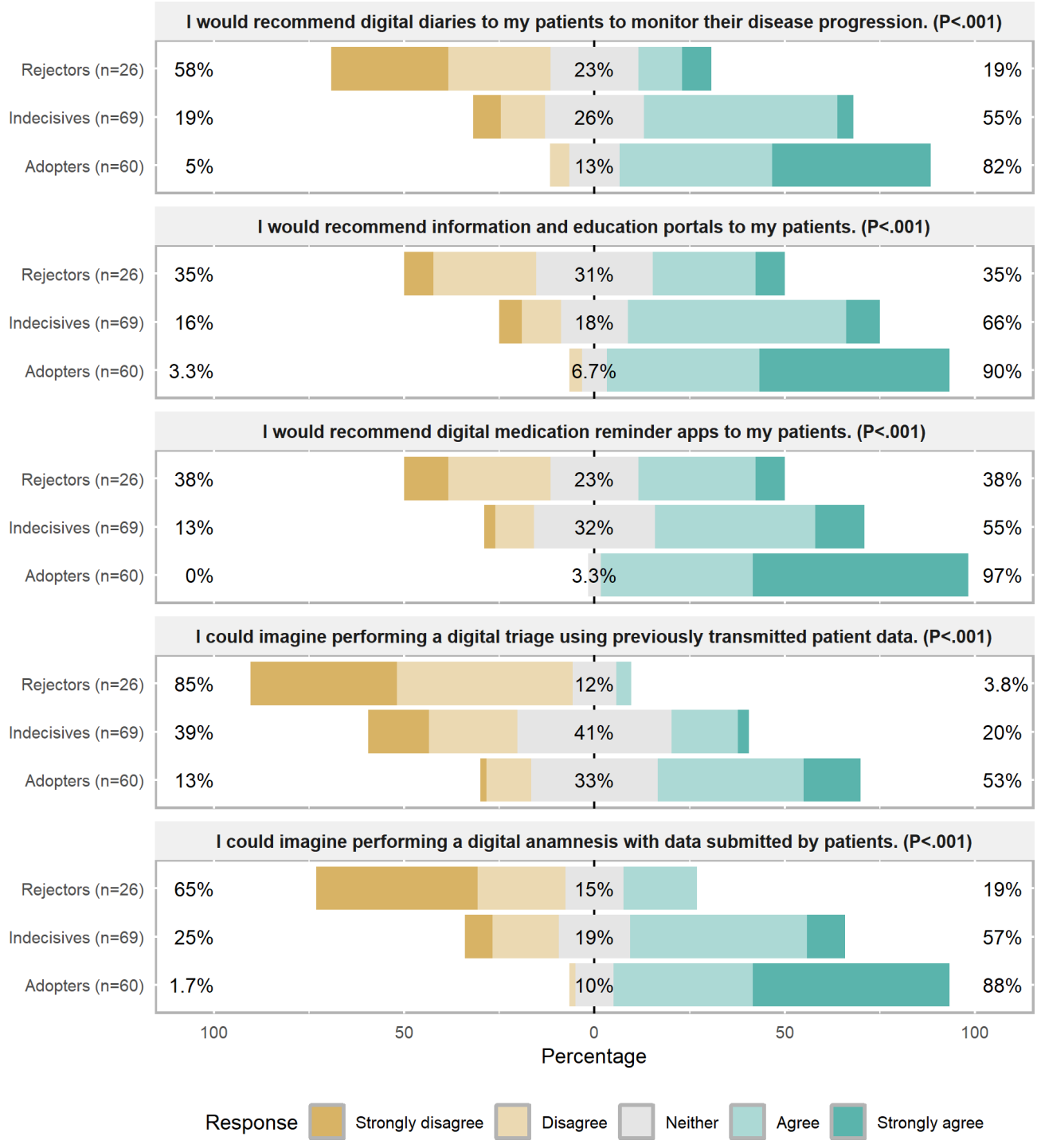


XSL•FO
RenderX

Future Intention to Use DHIs Within Identified Clusters

When considering the potential use of DHIs for their patients in the future, significant differences were identified among the three clusters (Figure 4). Between 80% (48/60; patient diaries) and 97% (58/60; medication reminder) of Adopters were willing to use 4 out of 5 described interventions. In contrast, only 55% (38/69: medication reminders) to 66% (45/68; education portals) of Indecisives can imagine using or recommend specific DHIs more often than currently. Rejectors had low potential use rates ranging from 19% (5/26; for digital anamnesis) to 35% (9/26; for education portals). While a slight majority of Adopters (32/60, 53%) could envision using digital triage, 33% (20/60) were undecided. However, 20% (14/69) of the Indecisives could imagine using digital triage, while 41% (28/69) were undecided. In contrast, 85% (22/26) of the Rejectors disagreed on using digital triage in the future.

Figure 4. Potential future use of digital health interventions by clusters. *P* values indicate the outcome of a chi-square test.



XSL•FO
RenderX

Discussion

Principal Findings

This study aimed to explore dermatologists' acceptability of DHIs, as well as their current and potential future use. Further, clusters were identified according to their level of acceptability, and differences were examined between the clusters and sociodemographics, as well as DHI current and future use. The current use barely exceeded 30% for interventions like teledermatology, AI for diagnostic purposes, and electronic patient data, even among willing dermatologists.

Overall, acceptability was moderate but varied significantly among the identified clusters. We recognized a cluster of dermatologists with high acceptability (Adopters), another similar-sized cluster with moderate acceptability (Indecisives), and a third minority with low acceptability (Rejectors). Indecisives and Rejectors perceived negative expectancies following the introduction of DHIs (eg, workload increase) and mostly questioned patients' willingness and competence to use them. Adopters were younger, and more commonly practicing in urban settings and outpatient clinics compared to Rejectors. Satisfaction with the nationwide eHealth strategy was generally low, although perspectives among Adopters were more divided. Current and future use of DHIs depended on the type of DHI and cluster, with Adopters showing higher rates.

Significant inequalities in DHI adoption between clusters indicate the presence of a second-order digital divide (disparities in digital competencies and acceptability toward DHIs) [41,42]. The more common adoption of DHIs among younger and urban dermatologists has already been identified [30]. This difference can contribute to existing health inequities by impacting access to health care and potentially even health outcomes [43]. It is crucial to closely monitor the impact of further DHI adoption in dermatology and address any potential health disparities.

In general, inadequate knowledge, as perceived in the survey, is a recognized barrier to the adoption of DHIs [22,44]. Integrating education on digital medicine into medical curricula and continuing medical education programs is, therefore, a potential approach to increase use and mitigate the digital divide [45]. As the health care system undergoes digital transformation, investing in digital health education is crucial [46].

The survey identified structural barriers to the widespread adoption of DHIs, including inadequate reimbursement and incentives, workload concerns, and IT infrastructure challenges. In addition, the inadequate nationwide eHealth infrastructure [47], as also noted by dermatologists in our survey, hinders the seamless integration of DHIs, for example, in teledermatology due to problems with electronic prescriptions and lack of access to patient records [21,48]. These factors may help explain why the adoption rate of innovative technologies, including teledermatology, remained below 30% even among willing adopters. To address these challenges, developers and decision makers must prioritize aligning the eHealth infrastructure with practitioners' needs, as seen in other European countries [19]. Incentivizing physicians, alongside involving them in the development process more frequently to reduce concerns, can be a focus to increase DHI uptake [21].

The limited availability of DHIs may also contribute to the relatively low adoption rates, even among those categorized as Adopters. In Germany, for instance, store-and-forward teledermatology platforms primarily focus on diagnosis [15,16,49], with limited emphasis on triage or referrals [14] and follow-up of patients with chronic conditions [13]. This limitation restricts the broader use of the technology for different patient groups. Additionally, patient electronic health records remain incompletely implemented [19], and there is a shortage of patient apps capable of either generating (eg, heart rate) or supporting patients in generating (eg, patient-reported outcomes) health data. Conversely, the swift uptake of AI apps, even among those categorized as Rejectors, underscores the benefits of this technology for dermatologists [50]. Our research suggests that most Adopters and the majority of Indecisives express a desire to use and recommend specific DHIs more frequently, highlighting the need for increased availability of evidence-based DHIs.

The perspective that patients would not welcome or find DHIs easy to use aligns with evidence that over 70% of the German population lacks adequate digital health competencies [51]. Patients also demonstrate low to moderate acceptability toward DHIs [34,52-54]. Other research contradicts those findings, suggesting that patients with dermatological conditions even have a higher acceptability toward DHIs in comparison to dermatologists [27]. Nevertheless, physicians and nurses can provide vital guidance, given their significant influence on patients [23,25,26]. Notably, Adopters recognized this role, expressing a willingness to recommend DHIs despite acknowledging patients' overall lack of digital competencies. This proactive stance should be complemented by health insurers, who also have a responsibility to promote DHIs and enhance digital literacy to ensure their appropriate utilization [55]. Such efforts are aligned with insurers' interests in supporting cost-effective health care services [55, 56].

One key distinction we found is that Adopters are more commonly affiliated with a hospital-connected outpatient clinic, while the majority of the other clusters involve self-employed dermatologists in medical practices. This may explain the differences in opinions on costs, reimbursements, or investments. Physicians in outpatient clinics may be less concerned with IT costs and IT investments than physicians in private practices, as they do not have to organize or pay themselves [57].

Similar to other studies on acceptability toward digitalization, it is important to consider the Intention-Behavior gap when interpreting the survey results. Here, intending to adopt DHIs in the future may not guarantee increased actual adoption in the future [58]. Additionally, while an association has been observed between a group exhibiting high acceptability (Adopters) and their current adoption of DHIs, the cross-sectional nature of the study means that the direction of this association remains ambiguous and cannot be extrapolated into the future.

XSL•FO
RenderX

Limitations

The survey has further limitations. Only 170 dermatologists participated out of over 6000 practicing dermatologists, of which 3500 were organized in the DDA and 1000 were invited [59]. Additionally, while dermatologists were randomly invited, our survey exhibited an overrepresentation of dermatologists in the medium age group (50–59 years), with younger dermatologists (40–49 years) and older dermatologists (60+) being underrepresented [59]. This observation is particularly significant considering the age differences identified between the clusters, where younger participants are more commonly affiliated with the Adopters cluster and older participants with the Rejectors cluster. The survey is also prone to voluntary participation, which might have attracted dermatologists with more extreme views on health care digitalization. Hence, the sizes of the clusters and other findings may not be fully generalizable to the real world. We also did not account for the potential integration of certain DHIs, such as the use of AI to preanalyze images submitted through store-and-forward teledermatology. Future research should therefore explore the impact of AI on perceptions and attitudes toward DHIs [54].

A factor analysis was conducted, as recommended, to reduce data dimensionality before performing the cluster analysis [39]. EFA was chosen instead of confirmatory factor analysis, as no theoretical assumptions about item interactions were made. The approach limits the comparability with other research on acceptability [33,34]. In addition, it is often recommended to conduct a confirmatory factor analysis following an EFA [60]. However, our sample size was not large enough to split the dataset to allow for a subsequent analysis. Despite this limitation, EFA calculations were deemed appropriate based on the KMO criterion and Bartlett's test of sphericity, resulting in factors with acceptable internal consistency [40]. To validate the findings of our research, the results could be further explored by conducting qualitative interviews with dermatologists and using a mixed methods approach [61].

Conclusions

The survey results indicate that dermatologists' overall acceptability toward DHIs is moderate. However, a sizable cluster expresses a high willingness to incorporate these interventions and recommend them to patients. Two other distinct clusters exist: one indecisive about health care digitalization and another rejecting DHIs.

The observed variation in actual adoption rates among clusters suggests a nuanced relationship between willingness and the implementation of DHIs. Dermatologists who express willingness are more inclined to adopt DHIs more frequently. However, despite this willingness, overall use rates remain moderate. This may be attributed, in part, to several structural barriers, including inadequate reimbursement, insufficient IT infrastructure, limitations of the nationwide eHealth infrastructure, and the limited availability of DHIs.

A notable perception among dermatologists, especially those with low acceptability, is the perceived low patient willingness and competency regarding digital health. This points to efforts needed to enhance digital literacy among patients. It is imperative for all stakeholders in the health care system, including health insurers, to invest in initiatives to improve patient education and empowerment in digital health.

Additionally, prioritizing education for health care providers on digital health and involving them early in the development of DHIs could mitigate potential disparities in digital health.

Addressing all the mentioned barriers could improve the acceptability and use of DHIs among dermatologists, particularly among those who are indecisive, thereby enhancing the integration of DHIs to support patient care.

Acknowledgments

The authors thank the Scientific Communication Team of the IVDP, in particular Merle Twesten, for copyediting. We acknowledge financial support from the Open Access Publication Fund of UKE—Universitätsklinikum Hamburg-Eppendorf and DFG—German Research Foundation. This text was revised using ChatGPT-3.5/4.0 to enhance readability and ensure grammatical accuracy. The research project was financially supported via a research grant by Novartis Pharma GmbH. Novartis did not influence the design, data collection, analysis, publication decision, or development of this study.

Conflicts of Interest

MO is a coauthor of the German Association of Scientific Medical Societies in Germany (AWMF) guideline on teledermatology. MA is a coauthor of the German AWMF guideline on teledermatology and a scientific advisor for the teledermatology platform derma2go AG, A+ Videoclinic GmbH, and Novartis Pharma GmbH.

Multimedia Appendix 1
Overview of the items of the survey.
[DOCX File, 33 KB - humanfactors_v12i1e59757_app1.docx]

Multimedia Appendix 2
Factor loadings of items on the acceptability: results from the exploratory factor analysis.
[DOCX File, 21 KB - humanfactors_v12i1e59757_app2.docx]

XSL•FO
RenderX

Checklist 1
Checklist for Reporting of Survey Studies (CROSS)
[DOCX File, 21 KB - humanfactors_v12i1e59757_app3.docx]

References

1. Svensson A, Ofenloch RF, Bruze M, et al. Prevalence of skin disease in a population-based sample of adults from five European countries. Br J Dermatol 2018 May;178(5):1111-1118. [doi: 10.1111/bjd.16248] [Medline: 29247509]
2. Cribier B, Aroman MS, Merhand S, et al. Prevalence of visible skin diseases: an international study of 13,138 people. J Eur Acad Dermatol Venereol 2023 Feb;37(2):e180-e182. [doi: 10.1111/jdv.18666] [Medline: 36251446]
3. Augustin M, Herberger K, Hintzen S, Heigel H, Franzke N, Schäfer I. Prevalence of skin lesions and need for treatment in a cohort of 90 880 workers. Br J Dermatol 2011 Oct;165(4):865-873. [doi: 10.1111/j.1365-2133.2011.10436.x] [Medline: 21623753]
4. Krensel M, Augustin M, Rosenbach T, Reusch M. Waiting time and practice organization in dermatology. J Dtsch Dermatol Ges 2015 Aug;13(8):812-814. [doi: 10.1111/ddg.12625] [Medline: 26213818]
5. Kis A, Augustin M, Augustin J. Regional healthcare delivery and demographic change in Germany—scenarios for dermatological care in 2035. J Dtsch Dermatol Ges 2017 Dec;15(12):1199-1209. [doi: 10.1111/ddg.13379] [Medline: 29228491]
6. Pilz AC, Zink A, Schielein MC, et al. Despite large choice of effective therapies: individuals with psoriasis still seem undertreated. J Dtsch Dermatol Ges 2021 Jul;19(7):1003-1011. [doi: 10.1111/ddg.14387] [Medline: 33955676]
7. van der Kraaij GE, Vermeulen FM, Smeets PMG, Smets EMA, Spuls PI. The current extent of and need for shared decision making in atopic dermatitis and psoriasis in the Netherlands: an online survey study amongst patients and physicians. J Eur Acad Dermatol Venereol 2020 Nov;34(11):2574-2583. [doi: 10.1111/jdv.16340] [Medline: 32163645]
8. Wildenbos GA, Peute LW, Jaspers MWM. Impact of patient-centered eHealth applications on patient outcomes: a review on the mediating influence of human factor issues. Yearb Med Inform 2016 Nov 10;2016(1):113-119. [doi: 10.15265/IY-2016-031] [Medline: 27830238]
9. Eysenbach G. What is e-health? J Med Internet Res 2001;3(2):E20. [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
10. Boogerd EA, Arts T, Engelen LJ, van de Belt TH. "What Is eHealth": time for an update? JMIR Res Protoc 2015 Mar 12;4(1):e29. [doi: 10.2196/resprot.4065] [Medline: 25768939]
11. Reinders P, Augustin M, Kirsten N, Fleyder A, Otten M. Digital health interventions in dermatology-mapping technology and study parameters of systematically identified publications. J Eur Acad Dermatol Venereol 2023 Dec;37(12):2440-2449. [doi: 10.1111/jdv.19392] [Medline: 37528462]
12. Liu Y, Jain A, Eng C, et al. A deep learning system for differential diagnosis of skin diseases. Nat Med 2020 Jun;26(6):900-908. [doi: 10.1038/s41591-020-0842-3] [Medline: 32424212]
13. Armstrong AW, Ford AR, Chambers CJ, et al. Online care versus in-person care for improving quality of life in psoriasis: a randomized controlled equivalency trial. J Invest Dermatol 2019 May;139(5):1037-1044. [doi: 10.1016/j.jid.2018.09.039] [Medline: 30481495]
14. Koch R, Rösel I, Polanc A, et al. TELEDerm: implementing store-and-forward teledermatology consultations in general practice: results of a cluster randomized trial. J Telemed Telecare 2024 May;30(4):647-660. [doi: 10.1177/1357633X221089133] [Medline: 35578544]
15. Abeck F, Kött J, Bertlich M, et al. Direct-to-consumer teledermatology in Germany: a retrospective analysis of 1,999 teleconsultations suggests positive impact on patient care. Telemed J E Health 2023 Oct;29(10):1484-1491. [doi: 10.1089/tmj.2022.0472] [Medline: 36862525]
16. Otten M, Greis C, Reinders P, Fleyder A, Reich K, Augustin M. Patient and physician outcomes of a store-and-forward teledermatology application in Germany. J Eur Acad Dermatol Venereol 2024 Feb;38(2):e148-e151. [doi: 10.1111/jdv.19501] [Medline: 37669866]
17. Svendsen MT, Andersen F, Andersen KH, et al. A smartphone application supporting patients with psoriasis improves adherence to topical treatment: a randomized controlled trial. Br J Dermatol 2018 Nov;179(5):1062-1071. [doi: 10.1111/bjd.16667]
18. Augustin M, Strömer K, et al. S2k-leitlinie teledermatologie. AWMF online. 2020. URL: https://www.awmf.org/leitlinien/detail/ll/013-097.html [accessed 2022-05-25]
19. Transfer of digital health applications into everyday healthcare (part 6) [Report in German]. : Bertelsmann Stiftung; 2019 URL: https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/transfer-von-digital-health-anwendungen-in-den-versorgungsalltag-teil-6
20. Elsner P. Teledermatology in the times of COVID-19—a systematic review. J Dtsch Dermatol Ges 2020 Aug;18(8):841-845. [doi: 10.1111/ddg.14180] [Medline: 33448667]
21. Nohl-Deryk P, Brinkmann JK, Gerlach FM, Schreyögg J, Achelrod D. Barriers to digitalisation of healthcare in Germany: a survey of experts [Article in German]. Gesundheitswesen 2018 Nov;80(11):939-945. [doi: 10.1055/s-0043-121010]

XSL•FO
RenderX

22. Schreiweis B, Pobiruchin M, Strotbaum V, Suleder J, Wiesner M, Bergh B. Barriers and facilitators to the implementation of eHealth services: systematic literature analysis. J Med Internet Res 2019 Nov 22;21(11):e14197. [doi: 10.2196/14197] [Medline: 31755869]
23. Wade VA, Eliott JA, Hiller JE. Clinician acceptance is the key factor for sustainable telehealth services. Qual Health Res 2014 May;24(5):682-694. [doi: 10.1177/1049732314528809] [Medline: 24685708]
24. Arsenijevic J, Tummers L, Bosma N. Adherence to electronic health tools among vulnerable groups: systematic literature review and meta-analysis. J Med Internet Res 2020 Feb 6;22(2):e11613. [doi: 10.2196/11613] [Medline: 32027311]
25. Rentrop V, Damerau M, Schweda A, et al. Predicting acceptance of e-Mental health interventions in patients with obesity by using an extended unified theory of acceptance model: cross-sectional study. JMIR Form Res 2022 Mar 17;6(3):e31229. [doi: 10.2196/31229] [Medline: 35297769]
26. Safi S, Danzer G, Schmailzl KJ. Empirical research on acceptance of digital technologies in medicine among patients and healthy users: questionnaire study. JMIR Hum Factors 2019 Nov 29;6(4):e13472. [doi: 10.2196/13472] [Medline: 31782741]
27. Klein TM, Augustin M, Kirsten N, Otten M. Attitudes towards using electronic health records of patients with psoriasis and dermatologists: a cross-sectional study. BMC Med Inform Decis Mak 2020 Dec 30;20(1):344. [doi: 10.1186/s12911-020-01302-y] [Medline: 33380329]
28. Hennemann S, Beutel ME, Zwerenz R. Ready for eHealth? Health professionals' acceptance and adoption of ehealth interventions in inpatient routine care. J Health Commun 2017 Mar;22(3):274-284. [doi: 10.1080/10810730.2017.1284286] [Medline: 28248626]
29. Sezgin E, Özkan-Yildirim S, Yildirim S. Understanding the perception towards using mHealth applications in practice. Inf Dev 2018 Mar;34(2):182-200. [doi: 10.1177/0266666916684180]
30. Augustin M, Reinders P, Janke TM, et al. Attitudes toward and use of eHealth technologies among German dermatologists: repeated cross-sectional survey in 2019 and 2021. J Med Internet Res 2024 Feb 12;26:e45817. [doi: 10.2196/45817] [Medline: 38345855]
31. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34. [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
32. Sharma A, Duc NTM, Thang TLL, et al. A consensus-based checklist for reporting of survey studies (CROSS). J Gen Intern Med 2021 Oct;36(10):3179-3187. [doi: 10.1007/s11606-021-06737-1] [Medline: 33886027]
33. Viswanath V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425. [doi: 10.2307/30036540]
34. Philippi P, Baumeister H, Apolinário-Hagen J, et al. Acceptance towards digital health interventions—model validation and further development of the unified theory of acceptance and use of technology. Internet Interventions 2021 Dec;26:100459. [doi: 10.1016/j.invent.2021.100459] [Medline: 34603973]
35. Anderson JC, Gerbing DW. Structural equation modeling in practice: a review and recommended two-step approach. Psychol Bull 1988;103(3):411-423. [doi: 10.1037/0033-2909.103.3.411]
36. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
37. Guidelines for safeguarding good research practice: code of conduct. : Deutsche Forschungsgemeinschaft eV; 2019 URL: https://zenodo.org/records/14281892 [accessed 2024-12-04]
38. Floyd FJ, Widaman KF. Factor analysis in the development and refinement of clinical assessment instruments. Psychol Assess 1995;7(3):286-299. [doi: 10.1037/1040-3590.7.3.286]
39. Gao CX, Dwyer D, Zhu Y, et al. An overview of clustering methods with guidelines for application in mental health research. Psychiatry Res 2023 Sep;327:115265. [doi: 10.1016/j.psychres.2023.115265] [Medline: 37348404]
40. Howard MC. A review of exploratory factor analysis decisions and overview of current practices: what we are doing and how can we improve? Int J Hum Comput Interact 2016 Jan 2;32(1):51-62. [doi: 10.1080/10447318.2015.1087664]
41. Elena-Bucea A, Cruz-Jesus F, Oliveira T, Coelho PS. Assessing the role of age, education, gender and income on the digital divide: evidence for the European Union. Inf Syst Front 2021 Aug;23(4):1007-1021. [doi: 10.1007/s10796-020-10012-9]
42. Households—level of internet access: online data code: isoc_ci_in_h. Eurostat. 2022. URL: https://ec.europa.eu/eurostat/databrowser/view/ISOC_CBS/default/table?lang=en&category=isoc.isoc_c.isoc_cb [accessed 2023-01-02]
43. Yao R, Zhang W, Evans R, Cao G, Rui T, Shen L. Inequities in health care services caused by the adoption of digital health technologies: scoping review. J Med Internet Res 2022 Mar 21;24(3):e34144. [doi: 10.2196/34144] [Medline: 35311682]
44. Rheingans A, Soulos A, Mohr S, Meyer J, Guse AH. The Hamburg integrated medical degree program iMED. GMS J Med Educ 2019;36(5):Doc52. [doi: 10.3205/zma001260] [Medline: 31815162]
45. Werner R, Henningsen M, Schmitz R, Guse AH, Augustin M, Gauer T. Digital health meets Hamburg integrated medical degree program iMED: concept and introduction of the new interdisciplinary 2nd track digital health. GMS J Med Educ 2020;37(6):Doc61. [doi: 10.3205/zma001354] [Medline: 33225053]
46. The Topol Review: preparing the healthcare workforce to deliver the digital future: an independent report on behalf of the secretary of state for health and social care February 2019. : NHS Health Education England; 2019 URL: https://topol.hee.nhs.uk/wp-content/uploads/HEE-Topol-Review-2019.pdf [accessed 2025-05-10]

XSL•FO
RenderX

47. Report from IGES on the practice barometer 2020 [Report in German]. : IGES URL: https://www.iges.com/sites/igesgroup/iges.de/myzms/content/e6/e1621/e10211/e24893/e26423/e26424/e26426/attr_objs26428/IGES_KBV_Digitalbaromenter_2020_ger.pdf [accessed 2024-12-04]
48. Keil A, Brück R, Hahn K. Transmission of vital data into the German electronic health record. Curr Direct Biomed Eng 2023 Sep 1;9(1):583-586. [doi: 10.1515/cdbme-2023-1146]
49. Sondermann W, von Kalle C, Utikal JS, et al. External scientific evaluation of the first teledermatology app without direct patient contact in Germany ("Online Dermatologist—AppDoc") [Article in German]. Hautarzt 2020 Nov;71(11):887-897. [doi: 10.1007/s00105-020-04660-w]
50. Nelson CA, Pachauri S, Balk R, et al. Dermatologists' perspectives on artificial intelligence and augmented intelligence—a cross-sectional survey. JAMA Dermatol 2021 Jul 1;157(7):871-874. [doi: 10.1001/jamadermatol.2021.1685] [Medline: 34037674]
51. Schaeffer D, Gille S, Berens EM, et al. Digital health literacy of the population in Germany: results of the HLS-GER 2. Gesundheitswesen 2023 Apr;85(4):323-331. [doi: 10.1055/a-1670-7636] [Medline: 34905785]
52. von Weinrich P, Kong Q, Liu Y. Would you zoom with your doctor? A discrete choice experiment to identify patient preferences for video and in-clinic consultations in German primary care. J Telemed Telecare 2024 Jul;30(6):969-992. [doi: 10.1177/1357633X221111975] [Medline: 35915997]
53. Reinders P, Augustin M, Otten M. General and dermatological population's use and acceptance of digital health in Germany—a representative survey. J Dtsch Dermatol Ges 2024 Sep;22(9):1221-1231. [doi: 10.1111/ddg.15454] [Medline: 39555644]
54. Busch F, Hoffmann L, Xu L, et al. Multinational attitudes towards AI in healthcare and diagnostics among hospital patients. Health Informatics [FREE Full text] [doi: 10.1101/2024.09.01.24312016]
55. Schaeffer D, Gille S. Health literacy in the age of digitalization [Article in German]. Präv Gesundheitsf 2022 May;17(2):147-155. [doi: 10.1007/s11553-021-00872-7]
56. Gerke S, Stern AD, Minssen T. Germany's digital health reforms in the COVID-19 era: lessons and opportunities for other countries. NPJ Digital Med 2020;3(94):94. [doi: 10.1038/s41746-020-0306-7] [Medline: 32685700]
57. Müller-Walden T, Leibner M, Kohler M. Special costs in practice management [Report in German]. : Zi-Paper; 2020 URL: https://www.zi.de/fileadmin/Downloads/Service/Publikationen/Zi-Paper_15-2020_Erhebung_Praxismanagement.pdf [accessed 2024-12-10]
58. Sheeran P, Webb TL. The intention–behavior gap. Soc Personality Psych 2016 Sep;10(9):503-518. [doi: 10.1111/spc3.12265]
59. Doctors statistics as of 31 [Report in German]. : Bundesaerztekammer; 2022 Dec URL: https://www.bundesaerztekammer.de/fileadmin/user_upload/BAEK/Ueber_uns/Statistik/AErztestatistik_2022_09062023.pdf [accessed 2024-12-04]
60. Flora DB, Flake JK. The purpose and practice of exploratory and confirmatory factor analysis in psychological research: decisions for scale development and validation. Can J Behav Sci 2017;49(2):78-88. [doi: 10.1037/cbs0000069]
61. Zhang W, Creswell J. The use of "mixing" procedure of mixed methods in health services research. Med Care 2013 Aug;51(8):e51-e57. [doi: 10.1097/MLR.0b013e31824642fd] [Medline: 23860333]

Abbreviations

AI: artificial intelligence
DHI: digital health intervention
EFA: exploratory factor analysis
KMO: Kaiser-Meyer-Olkin

Edited by A Kushniruk; submitted 22.04.24; peer-reviewed by B Eapen, H Thulesius; revised version received 17.12.24; accepted 31.12.24; published 21.05.25.

Please cite as:
Reinders P, Augustin M, Otten M
Understanding Dermatologists' Acceptance of Digital Health Interventions: Cross-Sectional Survey and Cluster Analysis
JMIR Hum Factors 2025;12:e59757
URL: https://humanfactors.jmir.org/2025/1/e59757
doi:10.2196/59757

© Patrick Reinders, Matthias Augustin, Marina Otten. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 21.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information,

XSL•FO
RenderX

a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Device Functionalities and Technology Acceptance for Innovations in Neonatal Ventilation and Enhanced, Immediate Newborn Care: International, Multicenter, Web-Based Survey Study

Anna-Sophie Käferböck[1], MSc; Meggy Hayotte[2], PhD; Daniel Sieber[3], Prof Dr; Martin Pillei[3,4], Prof Dr-Ing; Martin Wald[5], Apl Prof Dr med univ

[1]Medical Sciences, Paracelsus Medical University, Strubergasse 21, Salzburg, Austria

[2]LAMHESS, Université Côte d'Azur, Nice, France

[3]Department Medical and Health Technologies, MCI | The Entrepreneurial University, Innsbruck, Austria

[4]Industrial Engineering and Management, MCI | The Entrepreneurial University, Innsbruck, Austria

[5]Division of Neonatology, Department of Pediatrics and Adolescent Medicine, University Hospital Salzburg, Paracelsus Medical University, Salzburg, Austria

Corresponding Author:
Anna-Sophie Käferböck, MSc
Medical Sciences, Paracelsus Medical University, Strubergasse 21, Salzburg, Austria

Abstract

Background: A substantial number of newborns face postdelivery respiratory issues annually. Current ventilation devices in immediate newborn care lack integrated sensors and supporting mechanisms for medical professionals. This is a potential field of improvement, as safe ventilation relies on accurate pressure administration in current t-piece resuscitators. As the needed support during the process is currently limited, it highlights the demand for innovations in neonatal ventilation technology to improve efficacy and reduce potential errors.

Objective: The objective of the study was to facilitate collaboration between medical and engineering experts to evaluate the critical factors for the successful implementation of an innovative ventilation technology in clinical immediate newborn care. Incorporating the views of medical professionals into the survey is expected to offer valuable insights to engineers for subsequent technological refinement.

Methods: An international multicenter online survey was conducted among 51 neonatal health care professionals in the DACH region (Germany, Austria, and Switzerland) in order to (1) assess the specific functionalities required in a neonatal ventilation assistant in immediate newborn care from a medical technology viewpoint, (2) characterize the acceptance of such a device as support tool using the extended technology acceptance model, and (3) identify further steps toward integration of such technologies.

Results: According to the results, a visual representation of the current mask leakage and tidal volume is an essential feature. Integrating alarms in visual rather than audible form when limit values are exceeded is preferable. In contrast, medical professionals ranked an external control using a foot pedal as the least necessary feature. Based on the findings, acceptance constructs of the neonatal ventilation technology were moderately scored. Perceived usefulness (β=.76, P<.001) was the main predictor of the behavioral intention to use such a supportive instrument.

Conclusions: There is an evident willingness to integrate sophisticated support techniques into a neonatal ventilation device for immediate newborn care.

(JMIR Hum Factors 2025;12:e64701) doi:10.2196/64701

KEYWORDS

infant; neonatology; medical device; respiration disorders; pulmonary ventilation; medical engineering; technology acceptance model; technology acceptability; resuscitation; critical care

Introduction

The probability of survival of newborns depends on the extent to which they can adapt to the new environment and living conditions after delivery. Nevertheless, 3% of premature or newborn infants required respiratory support in the form of mask ventilation in the first few minutes of life [1], with approximately 1% of infants requiring postnatal resuscitation [2]. The lungs and associated tissue need to develop in the first few minutes of life before they are fully functional [3]. This period is highly critical, as excessive positive pressure ventilation can lead to damage to the lung tissue [3].

XSL•FO
RenderX

This issue is further compounded by the stress [4] experienced by medical personnel during a resuscitation scenario or ventilation procedure, which disrupts their typical level of physical performance and further exacerbates the situation. This can lead to inaccurate assessment of color [5] and heart rate [6]. One study [7] showed that despite a constant level of peak inspiratory pressure (PIP), maximum pressure on the lungs during inhalation, there was a range in the tidal volume administered, with the rationale that chest movement may not be accurately detected [8]. In such instances, the utilization of medical devices becomes paramount, as they have the potential to alleviate the psychological burden on health care providers.

The choice of ventilator used by medical personnel makes a difference in the actual quality of neonatal ventilation [9,10]. One study [11] has determined that the currently used technology (t-piece) may result in excessive ventilation pressures for newborns. Encounter emerging criticism regarding the inaccurate administration of ventilation pressures through t-piece devices [12], it is fundamental to design upgraded products in medical care to enhance safety in newborn life support.

This study aimed to assess essential device features for neonatal ventilation during immediate newborn care, as existing research addresses this aspect only in a limited way. The goal is not to replace the clinician but to enhance support through predefined functions that minimize potential errors and improve patient outcomes. The primary objective was to identify neonatologists' acceptance and need for technology that aids neonatal ventilation in the delivery room. The study sought to establish a collaborative interface between engineers and neonatologists, thereby aligning their perspectives and guiding the development of technology in this domain. A central objective was to address potential misunderstandings between medicine and technology proactively.

Methods

Recruitment

This survey was conducted online using the platform LimeSurvey (LimeSurvey Cloud Version 6.3.0, LimeSurvey GmbH). The first page of the survey stated the objective; the inclusion criteria, which were limited to individuals with (1) completed medical degrees and (2) a valid license to practice in neonatology; and the imperative need for their participation. The second page was dedicated to the information leaflet and the electronic consent form. Sociodemographic data and experiences with neonatal ventilation (ie, inquired about how participants would assess their personal experience in neonatal ventilation, the frequency of their exposure to this practice, and their familiarity with specific ventilation equipment used in the delivery room) were then asked, followed by the evaluation of particular requirements and assessment of the acceptance of a possible prototype of a medical device in the area of newborn ventilation in immediate newborn care.

Individual responder contacts were not available, therefore, cluster sampling was used to target the responders by their neonatology departments. The survey link was sent to a total of 351 neonatology departments in the DACH region (Germany, Austria, Switzerland), which provided contact information to the Department of Neonatology of the University Hospital Salzburg. Neonatology departments were asked to distribute the link to the survey to their medical experts using a prepared email. Departments were contacted 3 times to improve the response rate. Respondents were given the option to answer the questionnaire in either German or English. Data collection took place in September and October 2023.

Measurement

The survey consisted of two parts: (1) evaluation of specific requirements, suggested by a consensus of experts in neonatology of the University Hospital Salzburg, for a medical device in the area of newborn ventilation in immediate newborn care; and (2) evaluation of the acceptance of a possible prototype of this technology based on the extended technology acceptance model (TAM) [13].

To assess the acceptance of the ventilation device for premature and newborn patients, we used the TAM constructs [13], extended with the constructs of trust and risk [14], and the construct of social influence [15], which were adapted for this specific technology. In this adaptation, we substituted "information technology" [16] with "Neonatal Ventilation Device in Immediate Newborn Care." Data were gathered through an adapted TAM questionnaire containing 29 items distributed across 6 dimensions—perceived usefulness (6 items) [17], perceived ease of use (6 items) [18], behavioral intention to use (4 items) [16], social influence (5 items) [17], trust (3 items) [14], and perceived risk (5 items) [16]. The German version was translated based on existing versions [19] and adapted for the technology studies.

A pretesting phase preceding the study was conducted to ensure the comprehensibility of the included questionnaire items. Participants rated questions on a scale of 1 to 7, where "1" represented "not understandable" and "7" denoted "completely understandable." All items were above a cutoff of 4 at the end of the iterative phase, with an average rating of 5.4/7 (SD 1.3), indicating a good understanding of the individual items when applied to the response scale. Respondents (n=19), both academic and medical professionals, carried out this assessment.

The survey used a 7-point Likert scale for evaluation, where "1" indicated "strongly disagree" and "7" represented "strongly agree." The original survey version is attached in the Multimedia Appendix 1. In our sample, TAM dimensions showed high internal consistency (Cronbach α ranged from 0.73 to 0.94) [20]. Overall, the survey was structured according to that suggested by Burns et al [21] in their work with clinicians, with a cluster sampling design, a mixture of free-text and structured questions (mainly based on validated questionnaires), as well as a pretesting phase to ensure comprehensibility of the survey.

Statistical Analysis

The statistical analysis was undertaken using Excel (Microsoft Excel 2016, Microsoft Corp) and SPSS (IBM SPSS Statistics 29.0.0, IBM Corp). Means, medians, and SDs were used to present continuous measurements, and results from categorical measurements are shown as numbers and proportions. Multivariate ANOVAs were performed to examine differences

XSL•FO
RenderX

in the perceptions of the technology requirements as a function of the participants' characteristics (gender, age, and experience). Furthermore, 2 authors (ASK and DS) were involved in the analysis of the qualitative data by deductive analysis using the software MAXGDA (MAXGDA 24, VERBI – Software. Consult. Sozialforschung. GmbH).

Ethical Considerations

The study was conducted in accordance with local ethical guidelines as stipulated by the seventh revision of the Declaration of Helsinki. Ethical clearance was obtained from the Ethics Committee Salzburg (415-EALL/4/207/6 - 2024). Informed consent was obtained from all participants at the start of the study, and their data were analyzed anonymously. No forms of financial compensation were provided to the participants in this study.

Results

Sample Characteristics

A total count of 131 clicks was captured, 82 were identified as eligible attempts to respond by providing consent to participate, resulting in 51 valid responses. The survey was undertaken in German only. Physicians of pediatrics and adolescent medicine specializing in neonatology contributed to all responses. No anesthesiologists engaged in the field of neonatology participated, despite meeting the inclusion criteria. Participants reported their personal experience with neonatal ventilation with a mean of 6.0/7 (SD 1.2, median 6.0, IQR 5.00-7.00) and indicated the frequency of exposure with a mean of 5.8/7 (SD 1.4, median 6.0, IQR 5.00-7.00). Refer to Table 1 for detailed participant characteristics.

XSL•FO
RenderX

Table . Sample characteristics (n=51).

Variable and level		Count and proportion, n (%)
Sex		
	Male	27 (52)
	Female	24 (47)
Age (years)		
	18 - 29	1 (1)
	30-39	9 (17)
	40-49	19 (37)
	50-59	16 (31)
	60-69	6 (11)
	70-79	0 (0)
Educational level		
	Ongoing residency	8 (15)
	Completed residency	3 (5)
	CR[a] with ongoing specialization	4 (7)
	CR with completed specialization	35 (68)
Employment status		
	Full-time employment	38 (74)
	Part-time employment	12 (23)
	Retired	1 (1)
Ventilation experience (1=basic knowledge, 7=high expertise)		
	1	0 (0)
	2	0 (0)
	3	2 (3)
	4	5 (9)
	5	8 (15)
	6	13 (25)
	7	23 (45)
Ventilation frequency (1=never, 7=very often)		
	1	0 (0)
	2	1 (1)
	3	3 (5)
	4	6 (11)
	5	9 (17)
	6	8 (15)
	7	24 (47)

[a]CR: completed residency in pediatrics and adolescent medicine.

The ventilator from Paykel Healthcare, which operates under the name Perivent in German-speaking countries, but is also called Neopuff, was named the most frequently used device (n=44). The participants had the opportunity to select multiple devices simultaneously, depending on which ones they were currently using and with which they had previous experience. Furthermore, 2 of the devices mentioned were originally not designed by the manufacturers for ventilation of newborns in the delivery room, but for continuous ventilation in an intensive care setting. One of the devices cited is not intended for use in the immediate newborn care setting by the manufacturer, as it is specifically designed as a monitoring system for the neonatal ventilation process. Further details on the devices used for

XSL•FO
RenderX

ventilation in the delivery room setting are illustrated in Table 2.

Table . Currently used ventilation devices in immediate newborn care.

Device designation	Manufacturer	Country	Count, n
Perivent or Neopuff	Fisher & Paykel Healthcare	New Zealand	44
Giraffe Stand-alone Infant Resuscitation System	GE HealthCare	Canada	15
Giraffe Warmer	GE HealthCare	Canada	14
Neo-Tee	MedCare Visions GmbH	Germany	14
Resuscitaire	Drägerwerk AG & Co. KGaA	Germany	13
rPAP	Inspiration Healthcare	England	8
Stephan F120	Fritz Stephan GmbH	Germany	4
Panda Warmer ResusView	GE HealthCare	Canada	4
Babyroo TN300	Drägerwerk AG & Co. KGaA	Germany	3
Neo100[a]	Monivent AB	Sweden	3
Leonie Plus[b]	Löwenstein Medical SE & Co. KG	Switzerland	3
Hamilton-C1[b]	Hamilton Medical AG	Switzerland	2
Resusci Flow with Blender Unit 104 E Type	Atom Medical Corp	Japan	1
PNEUPAC BabyPAC100	Smiths Medical, Inc	United States	1

[a]Monitoring device for ventilation in immediate newborn care.

[b]Not designed for neonatal ventilation in immediate newborn care.

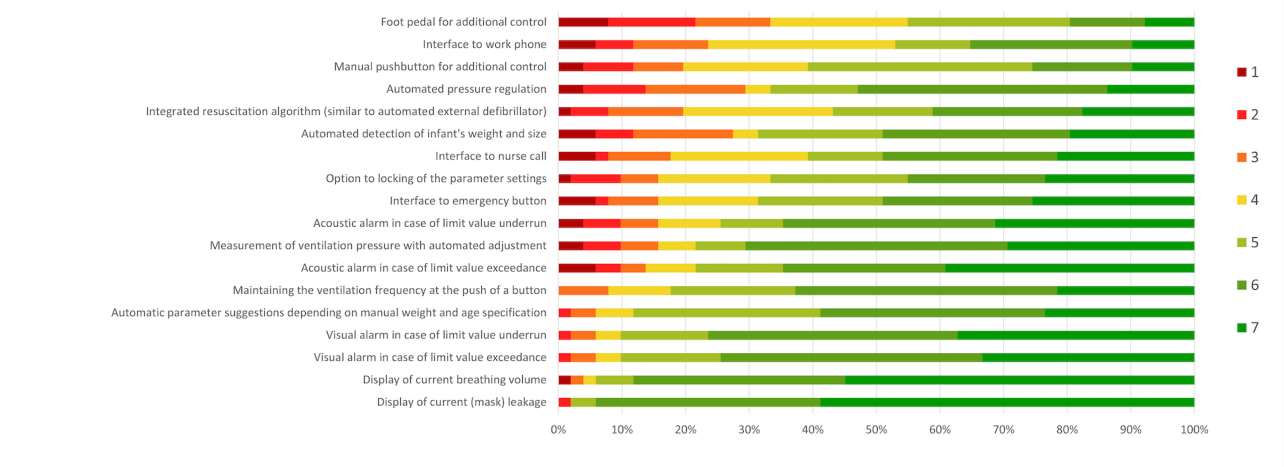
Device Functionalities

The participants rated the listed device features in terms of their future needs for implementation. Figure 1 illustrates the percentages of each rating score in a bar chart for each requirement. In addition to these predefined ratings, the participants were asked to what extent automation of workflows of neonatal ventilation at the push of a button might make sense: “1” represented a manual system, where the whole responsibility lies with the individual, whereas “7” indicated a fully automated system where the person is solely responsible for monitoring the device. The results show an average score of 4.2/7 (SD 1.7, median 4.0, IQR 3.00-5.00).

In addition to the predefined device requirements that participants were asked to rate, respondents were also provided with the opportunity to express their opinion on various requirements through a free-text field (n=18). Accordingly, a neonatal ventilation device used in the delivery room should include a “display of fraction of inspired oxygen” (n=3), along with the output of the “measurement of End-tidal carbon dioxide” (n=2) and the display of a “pressure-volume curve” (n=1). Another potential integration into such technology would be the capability for “humidification” (n=2) and “warming” (n=1) of breathing gases. Furthermore, features such as “time measurement” (n=1), “tidal volume per body weight” (n=1), and “alarms in the event of technical defects” (n=1) should be incorporated into this device. Another suggested feature could be an “automated start procedure with first ventilation” (n=1). A “memory function” (n=1) of the entire process would be beneficial for postintervention analyses. Finally, some participants expressed the preference “not to rely exclusively on the device with integrated technology” (n=2), emphasized the importance of considering “users’ experience” (n=1), and suggested that, in case of uncertainty, preset parameters can be “manually overwritten” (n=4).

XSL•FO
RenderX

Figure 1. Assessment of requirements for a ventilation device in immediate newborn care (n=51; no significant gender [$F_{18,32}$=1.40; P=.20], age [$F_{72,128}$=0.95; P=.60], or experience [$F_{72,128}$=1.01; P=.47] differences were shown based on multivariate ANOVAs).



Technology Acceptability

The participants expressed moderate levels of acceptance for the ventilation device in immediate newborn care based on the extended TAM (mean perceived usefulness 5.2, SD 1.1; mean perceived ease of use 4.8, SD 0.7; mean social influence 4.6, SD 1.1; mean trust 4.3, SD 1.2; mean risk 3.0, SD 0.9; mean behavioral intention 5.1, SD 1.2). Perceived usefulness, perceived ease of use, social influence, trust, and behavioral intention were all rated positively, while risk was rated as negligible.

Difficulties in Usage

Overview

In addition to the possible functions of such a device in the delivery room, participants were free to comment on any difficulties that might arise when using it. The ratings were divided into difficulties with first use (n=47), repeated use (n=25), and routine use (n=31).

First Use

Possible difficulties when using this type of technology for the first time can be divided into 6 categories based on the responses. Potential problems could lead to “incorrect operation” (n=10) during the process due to “personal uncertainties” (n=13) with the new technology. In addition, the “complexity of the system would require training” (n=9). There is also a risk of “relying on device support” (n=3) without additional follow-up checks of essential parameters. In addition, there is still the possibility of “technical malfunctions” (n=1) when using the system for the first time. In the “other” category (n=6), the following comments were made: on the one hand, that “curiosity to try something new” (n=1) arises personally when using the device for the first time, which allows “sensible setting of parameters” (n=1) and is “easy to use” (n=2). On the other hand, participants could imagine that the “individual assessment of the newborn” (n=1) plays a role in first-time use and that “focusing on the technology could lead to a reduction in observation of the child” (n=1).

Repeated Use

With repeated use of a new type of neonatal ventilation technology in the delivery room setting, the participants mentioned the difficulty that “clinical knowledge could be lost” (n=7) if a particular “reliance on the device” (n=3) was established. What would exist would be an “incorrect use” (n=2) with an “increased training effort” (n=4) in order to achieve a specific “safety in use” (n=4). However, after repeated use, the participants could already imagine having “fewer problems with increasing routine” (n=3). The “robustness of the device” (n=1) should, therefore, not be underestimated, as should “overly liberal use” (n=1) during this process.

Routine Use

In the course of routine use of such an assistance tool, the participants stated, among other things, the difficulty that users could “lose their clinical judgment” (n=11) due to increasing “familiarization with the technology” (n=5). However, “easy handling” (n=3) with little to “no” (n=5) issues in handling should arise in the course of routine. This technology could provide more “time for decisions relevant to care” (n=1), but there would still be the option of “no significant improvements” (n=1) in the overall process, for example, by “ignoring alarms” (n=2) because personal “expectations” (n=1) are different. In addition, “additional costs” (n=2) could be incurred for materials or upgrading an existing device.

Life-Threatening Malfunctions

In addition to the information on possible issues and difficulties in using such a respiratory assistant, the participants could also give their thoughts on life-threatening malfunctions (n=39). Comments were primarily related to “malfunctions in the area of ventilation pressure” (n=15), with incorrect outputs, both too high and too low, in the area of PIP and positive end-expiratory pressure (PEEP), as well as possible difficulties adjusting to the patient’s conditions. A “total failure” (n=9) of the entire system for no foreseeable reason and without a workaround (eg, switching to manual mode) was the second most common unacceptable malfunction. “Personal confusion” (n=8) when using such a tool can also lead to life-threatening situations, which must be considered during development. Incorrect

XSL•FO
RenderX

"measurements of the system" (n=5), starting with incorrect detection of the child's weight and the associated miscalculation of ventilation pressures, must be prevented at all costs. In addition, a "too long activation time" (n=1) before ventilation can take place, and "complicated handling" (n=1) can pose a life-threatening risk.

Additional Comments

In the final part of the questionnaire, there was also the option to provide general comments on medical technology in neonatal ventilation that the participants wished to contribute in the course of the survey. A total of 29 comments were received, with one response stating that "no" further issues were brought forward. In this section, the participants were concerned that "manual mode" (n=8) could be used to override any automated default settings or, if in doubt, switch them off. Such a system should be "easy to use" (n=6) in order to support both newcomers to the profession and neonatologists with many years of experience in their work. In addition, users should receive good and regular "training" (n=4) on the assistance tool in order to ensure that the familiarization phase and the assessment of therapy needs are of high quality. In the area of technical functionalities, functions such as "warming and humidification" (n=3) of the respiratory gas, "coupling with vital parameters" (n=1) and a "display of the ventilation strokes applied" (n=1) as well as a "sufficient battery" (n=1) and "alarm management" (n=1) should be considered. Depending on the different hospitals and clinics, there is also a difference as to which "group of people would use such technology" (n=1). The option to switch to a "mobile and portable version" (n=1) would avoid possible structural difficulties.

Discussion

Principal Results

This paper describes the results of a questionnaire survey in the field of ventilation technology for neonates in immediate newborn care. The responses highlighted the need for support mechanisms during the neonatal ventilation process, as participants rated the desired level of automation of such a technology as 4.1 out of 7, representing a medium to high degree of automation. The guidelines for neonatal care [22] further indicate that while situations requiring neonatal resuscitation are rare, the routine expertise to provide optimal quality is often lacking. Our findings should therefore serve as a starting point to support neonatologists in their practice in the best possible way, in terms of making high efforts in development in pressure administration, intuitive device design for easy handling, and display of critical vital parameters.

The current state-of-the-art neonatal ventilation in the delivery room uses a tube system containing a t-piece with a mask or tube plugged in. This system connects to an air-powered ventilator, where a specific oxygen saturation is set. A valve on the t-piece allows control of PEEP to manage lung pressure at the end of expiration. However, it lacks integrated monitoring of parameters such as tidal volume, ventilation pressure, in particular the PIP, and the determination of pressure loss due to mask leakage.

Our results are consistent and confirm a previous study [23] examining the influence of human factors during this process and improving the situation with assistive tools, such as integrating a display of current (mask) leakage, rated as the leading requirement. Free-text comments from individuals in our research support this trend. On the other hand, it is essential to highlight that one particular study [24] demonstrated that individuals with greater expertise in this domain exhibit an elevated mental workload and encounter higher cognitive demands throughout the process compared with inexperienced individuals using an assistive tool. In order to enhance personal performance during neonatal ventilation or resuscitation, the equipment to be used should be kept to a minimum in order to avoid fixation errors [25].

A topic addressed in the free-text comments pertained to the potential for "manual overwriting" of the device and the "incorrect usage during initial operation." According to the established standards, specifically the norm titled "Medical Devices -- Part 1: Application of Usability Engineering to Medical Devices" (IEC 62366-1:2015), such concerns should be anticipated during the manufacturing process and are fundamental prerequisites for any medical device. The operation of a medical device is intended to be seamless for the user and should minimize the awareness of the underlying technical interactions. It is recommended to carry out so-called usability tests to obtain such device properties. These tests are based on a specific scheme which, in addition to other considerations, should bring to light the problems mentioned earlier and provide opportunities to improve them [26].

It is important to note that this study focused on neonatologists, as they are the specialists in the ventilation of newborns. However, it is also worth considering that other medical staff, such as midwives, nurses, and anesthesiologists, could use the designed device. It is equally essential to acknowledge that some of the aforementioned devices for neonatal ventilation are not intended for immediate newborn care but rather for continuous ventilation in an intensive care setting or monitoring purposes only. This reinforces the importance of this work in providing an in-depth understanding of the needs of physicians for the development of assistive technologies specifically dedicated to neonatal ventilation in immediate newborn care.

Limitations

This work has some limitations, one of which is that the link to the survey was solely shared with neonatology facilities in German-speaking countries (DACH region). Therefore, only German-speaking neonatologists participated in the survey. No physicians with a background in anesthesia specialized in neonatology took part in this study. It could be speculated that they do not feel enough concerned by the survey. Second, to ensure a clear understanding of the questionnaire study, a pretesting phase was carried out on a small scale to assess the comprehensibility of the questions within the target group. Although considerable efforts were made to obtain a sufficient number of replies by broadly diffusing the survey, it ultimately resulted in a limited number of valid responses. It may be that not every center contacted and disseminated the survey due to

XSL•FO
RenderX

the high volume of inquiries that doctors receive. It is advisable to consider this potential limitation when interpreting the results.

Third, there was a huge difference in group sizes of the sample in relation to the educational level, where the group of fully qualified physicians of pediatrics and adolescent medicine with completed specialization in neonatology were overrepresented. Fourth, the method of free text response used to examine complementary user requirements could have been better examined through qualitative interviews. Based on these findings, future research could include qualitative interviews for a first prototype.

Comparison With Previous Work

Simulation training using a respiratory function monitor is already being conducted to further assist health care professionals by providing real-time information on the ventilation process [27]. However, certain studies [28-30] indicate a tendency for individuals to shift their focus from the infant to the monitor. This underlines the importance of a broad analysis of technologies in the delivery room. Our findings are in line with the idea that additional information, such as the display of current mask leakage and tidal volume, during the neonatal ventilation process may be beneficial, as well as with the experience of resuscitators using a respiratory function monitor [23]. However, additional information should not distract from the child nor overwhelm the caregiver. Considerations for individual functions should prioritize visual alarm functions over auditory ones to minimize potential distractions as well as to regain focus on the child.

Conclusions

The integration of supportive mechanisms into neonatal ventilation within the context of immediate newborn care is essential for the assistance of medical personnel. The acceptance of neonatologists of a device with additional functionalities and integrated sensors indicates a promising outlook. The objectives are clearly defined—to enhance the safety, efficacy, and quality of neonatal ventilation by minimizing human error and optimizing the overall care provided. The priorities encompass the improvement of device functionality, the assurance of user-friendliness, and the reduction of ventilatory complications. The analysis of technology acceptance reveals significant potential for future incorporation into existing systems.

The participation of physicians in the development process is of utmost importance. Their clinical expertise guarantees that the technology is aligned with real-world needs, remains intuitive, and integrates seamlessly into established protocols. Engaging with clinicians from the initial stages establishes a connection between medical and technological necessities, thereby ensuring that the technology improves patient care and aids clinicians in the delivery room.

Acknowledgments

ASK conceptualized the study, coordinated the data collection, conducted the statistical analysis, and prepared the first draft. MH supported within the methodology for the study design as well as the statistics. MW distributed the link to the survey within all the neonatology centers and supervised at all stages. DS checked the coding of the qualitative analysis. All authors critically reviewed the manuscript for improvement.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Detailed questionnaire for medical personnel in neonatology.
[PDF File, 47493 KB - humanfactors_v12i1e64701_app1.pdf]

References

1. Ersdal HL, Mduma E, Svensen E, Perlman JM. Early initiation of basic resuscitation interventions including face mask ventilation may reduce birth asphyxia related mortality in low-income countries: a prospective descriptive observational study. Resuscitation 2012 Jul;83(7):869-873. [doi: 10.1016/j.resuscitation.2011.12.011] [Medline: 22198423]
2. Antonucci R, Antonucci L, Locci C, Porcella A, Cuzzolin L. Current challenges in neonatal resuscitation: what is the role of adrenaline? Pediatr Drugs 2018 Oct;20(5):417-428. [doi: 10.1007/s40272-018-0300-6]
3. Wyllie J, Bruinenberg J, Roehr CC, Rüdiger M, Trevisanuto D, Urlesberger B. Die Versorgung und Reanimation des Neugeborenen [Article in German]. Notfall Rettungsmed 2015 Dec;18(8):964-983. [doi: 10.1007/s10049-015-0090-0]
4. Paliatsiou S, Boutsikou T, Xanthos T, et al. Stress in healthcare personnel involved in neonatal resuscitation – a systematic review. J Transl Sci 2021;7(5). [doi: 10.15761/JTS.1000447]
5. O'Donnell CPF, Kamlin COF, Davis PG, Carlin JB, Morley CJ. Clinical assessment of infant colour at delivery. Arch Dis Child Fetal Neonatal Ed 2007 Nov;92(6):F465-F467. [doi: 10.1136/adc.2007.120634] [Medline: 17613535]
6. Kamlin COF, O'Donnell CPF, Everest NJ, Davis PG, Morley CJ. Accuracy of clinical assessment of infant heart rate in the delivery room. Resuscitation 2006 Dec;71(3):319-321. [doi: 10.1016/j.resuscitation.2006.04.015] [Medline: 16989935]
7. Poulton DA, Schmölzer GM, Morley CJ, Davis PG. Assessment of chest rise during mask ventilation of preterm infants in the delivery room. Resuscitation 2011 Feb;82(2):175-179. [doi: 10.1016/j.resuscitation.2010.10.012] [Medline: 21074926]

XSL•FO
RenderX

8. Johnston ED, Stenson BJ. Am I getting chest wall movement? Arch Dis Child Fetal Neonatal Ed 2010 Nov;95(6):F391-F392. [doi: 10.1136/adc.2010.191254] [Medline: 20688864]
9. Restin T, Hönes M, Hummler HD, Bryant MB. Effective ventilation and chest compressions during neonatal resuscitation – the role of the respiratory device. J Matern Fetal Neonatal Med 2023 Dec 15;36(2):2276042. [doi: 10.1080/14767058.2023.2276042]
10. Hartung JC, Dold SK, Thio M, tePas A, Schmalisch G, Roehr CC. Time to adjust to changes in ventilation settings varies significantly between different T-piece resuscitators, self-inflating bags, and manometer equipped self-inflating bags. Am J Perinatol 2014 Jun;31(6):505-512. [doi: 10.1055/s-0033-1354562] [Medline: 24000108]
11. Hinder M, McEwan A, Drevhammer T, Donaldson S, Tracy MB. T-piece resuscitators: how do they compare? Arch Dis Child Fetal Neonatal Ed 2019 Mar;104(2):F122-F127. [doi: 10.1136/archdischild-2018-314860] [Medline: 29728414]
12. Hinder MK, Drevhammar T, Donaldsson S, Boustred M, Crott M, Tracy MB. T-piece resuscitators: can they provide safe ventilation in a low compliant newborn lung? Arch Dis Child Fetal Neonatal Ed 2021 Jan;106(1):25-30. [doi: 10.1136/archdischild-2019-318673] [Medline: 32546543]
13. Davis FD. A technology acceptance model for empirically testing new end-user information systems: theory and results. MIT Libraries. 1985. URL: https://dspace.mit.edu/handle/1721.1/15192 [accessed 2023-11-23]
14. Ortega Egea JM, Román González MV. Explaining physicians' acceptance of EHCR systems: an extension of TAM with trust and risk factors. Comput Human Behav 2011 Jan;27(1):319-332. [doi: 10.1016/j.chb.2010.08.010]
15. Barchielli C, Marullo C, Bonciani M, Vainieri M. Nurses and the acceptance of innovations in technology-intensive contexts: the need for tailored management strategies. BMC Health Serv Res 2021 Jul 3;21(1):639. [doi: 10.1186/s12913-021-06628-5] [Medline: 34215228]
16. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 1989 Sep;13(3):319-340 [FREE Full text] [doi: 10.2307/249008]
17. Holden M, Buck E, Clark M, Szauter K, Trumble J. Professional identity formation in medical education: the convergence of multiple domains. HEC Forum 2012 Dec;24(4):245-255. [doi: 10.1007/s10730-012-9197-6] [Medline: 23104548]
18. Zhang X, Han X, Dang Y, Meng F, Guo X, Lin J. User acceptance of mobile health services from users' perspectives: the role of self-efficacy and response-efficacy in technology acceptance. Inform Health Soc Care 2017 Mar;42(2):194-206. [doi: 10.1080/17538157.2016.1200053] [Medline: 27564428]
19. Harborth D, Pape S. German translation of the Unified Theory of Acceptance and Use of Technology 2 (UTAUT2) Questionnaire. SSRN Journal 2018. [doi: 10.2139/ssrn.3147708]
20. Taber KS. The use of Cronbach's alpha when developing and reporting research instruments in science education. Res Sci Educ 2018 Dec;48(6):1273-1296. [doi: 10.1007/s11165-016-9602-2]
21. Burns KEA, Duffett M, Kho ME, et al. A guide for the design and conduct of self-administered surveys of clinicians. Can Med Assoc J 2008 Jul 29;179(3):245-252. [doi: 10.1503/cmaj.080372]
22. Madar J, Roehr CC, Ainsworth S, et al. European Resuscitation Council Guidelines 2021: newborn resuscitation and support of transition of infants at birth. Resuscitation 2021 Apr;161:291-326. [doi: 10.1016/j.resuscitation.2021.02.014] [Medline: 33773829]
23. Kuypers K, van Zanten HA, Heesters V, et al. Resuscitators' opinions on using a respiratory function monitor during neonatal resuscitation. Acta Paediatr 2023 Jan;112(1):63-68. [doi: 10.1111/apa.16559] [Medline: 36177808]
24. Wagner M, Bibl K, Hrdliczka E, et al. Effects of feedback on chest compression quality: a randomized simulation study. Pediatrics 2019 Feb;143(2):e20182441. [doi: 10.1542/peds.2018-2441] [Medline: 30700565]
25. Yamada NK, Kamlin COF, Halamek LP. Optimal human and system performance during neonatal resuscitation. Semin Fetal Neonatal Med 2018 Oct;23(5):306-311. [doi: 10.1016/j.siny.2018.03.006] [Medline: 29571705]
26. Matern U, Büchel D. Usability of medical devices. In: Springer Handbook of Medical Technology: Springer; 2011:59-71. [doi: 10.1007/978-3-540-74658-4_6]
27. Dvorsky R, Rings F, Bibl K, et al. Real-time intubation and ventilation feedback: a randomized controlled simulation study. Pediatrics 2023 May 1;151(5):e2022059839. [doi: 10.1542/peds.2022-059839] [Medline: 37038898]
28. Wagner M, Gröpel P, Eibensteiner F, et al. Visual attention during pediatric resuscitation with feedback devices: a randomized simulation study. Pediatr Res 2022 Jun;91(7):1762-1768. [doi: 10.1038/s41390-021-01653-w] [Medline: 34290385]
29. Gröpel P, Wagner M, Bibl K, et al. Provider visual attention correlates with the quality of pediatric resuscitation: an observational eye-tracking study. Front Pediatr 2022;10:867304. [doi: 10.3389/fped.2022.867304] [Medline: 35685920]
30. Herrick H, Weinberg D, Cecarelli C, et al. Provider visual attention on a respiratory function monitor during neonatal resuscitation. Arch Dis Child Fetal Neonatal Ed 2020 Nov;105(6):666-668. [doi: 10.1136/archdischild-2020-319291] [Medline: 32616559]

Abbreviations

DACH region: Germany, Austria, and Switzerland
PEEP: positive end expiratory pressure
PIP: peak inspiratory pressure

XSL•FO
RenderX

TAM: technology acceptance model

Edited by C Jacob; submitted 30.07.24; peer-reviewed by EM Schwindt, F Rubulotta; revised version received 14.03.25; accepted 23.03.25; published 28.05.25.

Please cite as:
Käferböck AS, Hayotte M, Sieber D, Pillei M, Wald M
Device Functionalities and Technology Acceptance for Innovations in Neonatal Ventilation and Enhanced, Immediate Newborn Care: International, Multicenter, Web-Based Survey Study
JMIR Hum Factors 2025;12:e64701
URL: https://humanfactors.jmir.org/2025/1/e64701
doi:10.2196/64701

© Anna-Sophie Käferböck, Meggy Hayotte, Daniel Sieber, Martin Pillei, Martin Wald. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 28.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Pharmaceutical Analysis of Inpatient Prescriptions: Systematic Observation of Hospital Pharmacists' Practices in the Early User-Centered Design Phase

Jesse Butruille[1,2], PharmD, MEng; Natalina Cirnat[2], PharmD, MEng; Mariem Alaoui[2], MEng; Jérôme Saracco[1,3], PhD; Etienne Cousein[2,4], PharmD, PhD; Noémie Chaniaud[1,3], PhD

[1]ENSC – Ecole Nationale Supérieure de Cognitique, Bordeaux INP, 109 Av. Roul, Talence, France

[2]PharmIA SAS, Paris, France

[3]IMS – Laboratoire de l'Intégration du Matériau au Système, CNRS UMR5218, Talence, France

[4]ULR 7365 – GRITA, Université de Lille, Lille, France

Corresponding Author:
Jesse Butruille, PharmD, MEng
ENSC – Ecole Nationale Supérieure de Cognitique, Bordeaux INP, 109 Av. Roul, Talence, France

Abstract

Background: The health care sector's digital transformation has accelerated, yet adverse drug events continue to rise, posing significant clinical and economic challenges. Clinical decision support systems (CDSSs), particularly those related to medication, are crucial for improving patient care, identifying drug-related problems, and reducing adverse drug events. Hospital pharmacists play a key role in using CDSSs for patient management and safety. Human factors and ergonomics (HFE) methods are essential for designing effective, human-centered CDSSs. HFE involves 3 phases—exploration, design, and evaluation—with exploration being critical yet often overlooked in the literature. For medication-related CDSSs, understanding hospital pharmacists' tasks and challenges is vital for creating user-centered solutions.

Objective: This study aimed to explore the actual practices and identify the needs of hospital pharmacists analyzing electronic prescriptions. This study focused on the preliminary stage of the user-centered design of a pharmacist-centered CDSS.

Methods: The study involved observing 16 pharmacists across 5 hospitals in mainland France (a university hospital, 2 large general hospitals, a smaller general hospital, and a specialized clinic). Pharmacists were selected regardless of expertise. The observation method—systematic in situ observation with shadowing posture—involved following pharmacists as they analyzed prescriptions. Researchers recorded activities, tools used, verbalizations, behaviors, and interruptions, using an observation grid. Data analysis focused on modeling pharmacists' cognitive work, categorizing activities by action type, specificity, and information source. Sequential time data analysis and distance matrices were used to generate hierarchical clustering and identify similarity groups among the pharmacists' analyses. Each group was described using its typical sequences of analysis and related covariates.

Results: In total, 16 pharmacists analyzed and validated electronic prescriptions for 140 patients, averaging 5.48 minutes per patient. They spend 91% of their time searching for information rather than transmitting it. Most information comes from the list of prescriptions, but it is the time spent in electronic medical records (EMRs) that dominates at the heart of the analysis. Pharmaceutical interventions are most frequently transmitted in the last third of the sequence. The pharmaceutical analyses were grouped into 4 clusters: (cluster A, 22%) interventionist clinical analysis with extensive crossing of various sources of information and almost systematic pharmaceutical interventions; (cluster B, 52%) most common clinical analysis focusing on EMRs and biology results; (cluster C, 13%) logistical analysis, focusing on the pharmacy workflow and the medication circuit; and (cluster D, 13%) quick, trivial analyses based exclusively on the list of prescriptions.

Conclusions: The pharmaceutical analysis process is complex and multifaceted. Pharmacists are detectives, accessing a wealth of information to discriminate drug-related problems and respond accordingly. They also carry out different types of analysis, which lead to different needs and require different solutions from CDSSs. This exploratory study is an essential prerequisite for meeting the challenge of designing tools to support pharmaceutical analysis and pharmacists.

(JMIR Hum Factors 2025;12:e65959) doi:10.2196/65959

KEYWORDS

human factors; user experience research; user-centered design; professional practice; clinical pharmacy information systems; pharmacy service; systematic observation; pharmaceutical; pharmacy; pharmacist; France; prescription; user-centered; adverse

XSL•FO
RenderX

drug events; ADEs; clinical decision support system; CDSS; drug-related problems; patient management; safety; electronic medical record; EMR; pharmaceutical analysis; electronic health; eHealth; hospital

Introduction

Background

The digital transformation of the health care sector has accelerated over the past decade [1]. This digitization has reduced some risks associated with the use of health care products, such as handwriting-induced confusion [2]. However, adverse drug events continue to represent a significant clinical and economic burden for populations [3]. This paradox is due to the very diverse origins of adverse drug events and to the fact that the computerization of health care organizations corrects some drug-related problems but introduces many others, such as lack of ergonomics and alert fatigue [4,5]. Nowadays, health care providers must make more complex decisions [6], with an exponential increase in clinical and pharmaceutical information. Knowing that certain patient data are available, knowing how to access the data, having the time to search for information, and keeping up to date with the latest treatment recommendations are all challenges faced by professionals in their daily practices [7,8]. Out of this chaos came the interest in clinical decision support systems (CDSSs) to help clinicians make informed decisions and to derive full benefit from hospital information systems [9]. In particular, medication-related CDSSs have the greatest potential to benefit patient care and reduce drug-induced iatrogeny [10-12].

Hospital pharmacists are key users of CDSSs. The pharmacy department is responsible for the entire hospital's health care product circuit, and the pharmacist's activity is essential to patient safety and the reduction of health care expenditure [13]. Several studies have demonstrated the value of pharmacist-oriented CDSSs for patient management [10-12]. Of course, in health care, the patient's well-being is the main concern. Nevertheless, it is also necessary to consider the well-being of health care providers because they are responsible for the majority of medication-related errors [14]. This is why human factors and ergonomics (HFE) methods are fundamental in promoting the human-centered design of systems that support individuals and teams. As defined by Holden et al [15], "HFE is a scientific and practical human-centered discipline that studies and improves human performance in sociotechnical systems." It is also a standard of practice [16], with applications in many fields: aviation, surface transportation, military, and energy. In this respect, HFE methods are equally applicable to pharmacy challenges and medication-related CDSSs [15,17,18].

Using HFE to design digital products involves three phases [19]. (1) *Exploration* (or *analysis*) aimed at examining users in their natural environment, determining their tasks, activities, goals, constraints, and needs. Here, the real user problems are identified. (2) *Design* takes advantage of exploration data to provide solutions to identified problems. This ranges from simple wireframes to more complete prototypes. (3) *Evaluation* helps to assess ideas and the proposed design in order to optimize and correct it. It involves testing and scoring the user experience, ergonomics, and messages in both form and content. The first stage is fundamental to the success of the other two. However, HFE studies in the pharmacy sector tend to focus on the design or evaluation phase [20,21]. The same is true of studies on CDSSs [22-24]. Exploring the practices of hospital pharmacists would appear to be a crucial step, particularly for the development of medication-related CDSSs. Software editors need to study the complexity of the task of analyzing and validating electronic prescriptions in order to design products in accordance with the principles of human-centered design [16].

Prior Work

HFE methodologies are now well established in the health care sector, with a growing number of publications [25]. They all deal with exploration, design, and evaluation, but some phases of human-centered design are better represented than others. Evaluations are clearly the most represented [20,21,23], often involving products already designed or at the end of development, on which usability studies are carried out [26,27]. Unfortunately, these studies give little indication of the earlier stages, unless they also deal with design or co-design [19,22].

In fact, it is rare to find publications that focus solely on the analysis phase, especially in relation to health care providers and pharmacists. Studies focus on the patient [19,28,29] or on the patient's involvement in the health care system [30]. There are also studies on physicians [31-33], but these tend to focus on the usability of already implemented digital systems [34,35]. Finally, a few studies focus on the importance of the physician-pharmacist relationship for patient safety [12,36]. These studies use questionnaires, interviews, and focus groups, which are HFE methods that have the disadvantage of transcribing user subjectivity rather than actual behavior [37]. More objective methods, such as systematic observations, are better suited to identifying problems and needs for pharmacy research and clinical practice [15].

Several studies use observations as part of their methodology [38-46]. However, they tend to complement other methods in order to diversify data sources and reduce the risk of systematic bias based exclusively on reported behaviors and experiences [38]. Observations are treated here from a qualitative point of view, rather than as quantifying the activity of professionals. Moreover, as with all exploratory publications, the study participants are patients [38,39], physicians [40,41], nurses [42], and health care professionals as a whole [43-45], not just hospital pharmacists. Finally, they are not always user-centered [42], but rather focused on products [45,46] and their usability [39].

Ultimately, some studies have examined pharmacists' practices using quantitative field observations [47-51]. They do not necessarily focus on HFE methods but rather on the study of professional practice to improve the quality and safety of care [52]. One focuses on community pharmacists and the other 3 on hospital pharmacists. However, they do not deal specifically with the pharmaceutical analysis of prescriptions, but with the work pattern as a whole, and the impact of the medication

XSL•FO
RenderX

management system on this work pattern. In the context of developing a CDSS, the information provided by these studies is interesting but does not provide the exploratory data needed to design a prescription analysis support module. Moreover, it is difficult to retrieve exploratory data from the design and evaluation phases, as they only give access to the tip of the iceberg of product design: that is, "What was done?" and not "Why was it done?" To the best of our knowledge, there are no HFE or professional practice publications on prescription analysis by hospital pharmacists. There are only recommendations from professional associations, with useful elements for analyzing and improving practice [53,54].

Goal of This Study

This exploratory study focuses on the first stage of user-centered design: the phase of in-depth understanding of the user, his or her activity, and the context in which this activity takes place [15]. The aim is to explore hospital pharmacists' practices when analyzing and validating electronic prescriptions. It also aims to find their specific needs for the design of user-centered CDSSs.

We make several exploratory hypotheses: (1) there are different types of pharmaceutical analysis, (2) pharmacy logistics play a significant role in analysis, and (3) some information is difficult to access for pharmaceutical analysis.

Methods

Ethical Considerations

Ethics approval was not sought, as no identifiable health data were collected, and the pharmacists who were observed signed a free and informed consent form after they were informed of their personal data rights, in line with the General Data Protection Regulation (EU 2016/679). The observed pharmacists were not compensated for their participation in this study.

Recruitment

We visited 5 hospitals in mainland France for a total of 16 pharmacists observed: a university hospital with more than 2000 beds (3 pharmacists), 2 general hospitals with between 1000 and 2000 beds (4 pharmacists), 1 hospital with fewer than 1000 beds (5 pharmacists), and a specialized clinic with fewer than 200 beds (4 pharmacists).

Pharmacists are selected regardless of their expertise or status. This includes pharmacy residents, as they perform much of the validation of hospitalized patients. The only exclusion point is that the pharmacist's usual task is not to analyze and validate prescriptions (the activity must be performed at least once a week).

The number of participants observed depends on the discrepancy between pharmacists' practices: the greater the discrepancy, the more observations are needed to describe the different pharmaceutical analysis groups.

Observations

Protocol

Our exploration approach is based on systematic observations in an in situ context [37]. This method enables us to characterize the current state, analyze the environment, and identify the needs of hospital pharmacists. It will also describe how hospital pharmacists access, process, and return information over time in order to analyze and validate patients' electronic prescriptions.

Each observation consists of following a pharmacist analyzing and validating electronic prescriptions for about an hour. Pharmacists are informed of the objectives of the study before participating. They also receive some instructions, but the full methodological details are not given until the end of the observation to avoid influencing their behavior. Participants are instructed to act as they normally would. They are allowed to verbalize as much as they wish. They can move from one patient to another and can return to a patient who has already been validated. Task interruptions are natural in the professional practice of pharmacists [55], so the participant can interrupt his or her activity at any time.

At the start of each observation, the researcher notes generic information about the environment (date, time, site, and location) and the pharmacist being observed (area of expertise and status). At the end, he or she enters the time and comments, debriefing the participant. In 1 hour of observation, the pharmacist will analyze and validate several electronic prescriptions corresponding to different patients. For each patient, the researcher uses an observation grid to structure data collection.

Grid and Posture

An observation grid is completed for each patient (Multimedia Appendix 1). It is organized in the form of a successive list of activities, with each activity necessarily associated with the time required to perform it. Depending on the situation, other information may also be entered: (1) tools used, (2) verbalizations relevant to understanding the activity, (3) behaviors and reactions, and (4) the occurrence of a task interruption and its origin. The researcher records all the pharmacist's activities, taking care to provide sufficient detail for understanding, but without going too far, at the risk of losing the thread of the process.

The observation method requires the researcher to adopt a posture clearly defined by the study protocol. This is the shadowing technique [56,57]. It consists of following the user's activities as closely as possible, without intervening directly in the observed situation. However, the researcher may ask the participant questions in order to better understand the reasons for certain behaviors. Here, the researchers are hospital pharmacists familiar with user experience methods. This choice enhances the discernment of very specific professional activities and limits the questions asked of the participant.

Data Analysis

Observation Processing

In order to model pharmacists' cognitive work from raw observation data, each activity is assigned (1) a type of *action*

XSL•FO
RenderX

(search or transmit), (2) a level of *specificity* (generic or specific), and (3) a *source* of information (list of prescriptions, electronic medical records [EMRs], biology results, pharmaceutical interventions, pharmacy logistics, drug information, other information, or alert from prescriptions). Analysis times are represented in relative terms (percentage of total time) rather than absolute values to facilitate the comparison of activity sequences with wide time disparities.

The processing of temporal successions of activities is inspired by the processing of sociological data where populations change state as a function of time (eg, professional status: student → employee → unemployed → employee → retired). In particular, we drew on the work on exploratory analysis in a social context by Mariette et al [58]. To this end, we use the R package *TraMine*R [59] and multidomain case studies (in our case, activities are combinations of 3 domains) [60]. These methods are used to split up activities and generate comprehensible representations.

Clustering

Another advantage of these sequential data analysis methods is their ability to generate distance matrices between pharmaceutical analysis sequences [59,60]. This is followed by hierarchical clustering using the R package *cluster* [61] to distribute observations into different similarity groups. Finally, we describe the clusters by matching the graphs of typical sequences for each group (*TraMineR* package), with the description of covariates provided by the R package *FactoMineR* [62] and its class-based modality description function (contingency table, *V* test, and chi-square test).

Results

Observations and Sequences of Activities

Overview of Observations

The 16 pharmacists analyzed and validated the electronic prescriptions of 140 patients. The distribution of the time data for the 140 observations (mean 5.5; range 0.5-15; median 5, IQR 3.5-7.5 min) was nonnormal (Multimedia Appendix 2).

Pharmacists take an average of 5 minutes and 29 seconds to analyze one patient. The fastest analyses take 30 seconds, and the longest 15 minutes. The time spent on task interruptions is deducted from the total analysis time.

The observations show 17 different types of activity, formed by combining the 3 domains of *action*, *specificity*, and *source* (combinations are available in Multimedia Appendix 3). The representations of the activities (Table 1) separate the action, finesse, and source domains to better compare proportions.

The *action* domain is mostly represented by the search category, accounting for 90.7% of the total analysis time (Table 1). Pharmacists thus spend almost 91% of their time looking for information, and 9% transmitting it. In terms of *specificity*, information is more often generic (66%) than specific (34%). As for the *source* of information, the list of prescriptions is the most common (37.5%), along with the EMRs (28.2%); next in order of frequency are biology results (13.1%), pharmaceutical interventions (7.7%), pharmacy logistics (4.7%), drug information (4%), other information (3%), and alert from prescriptions (1.8%).

Table . Time distribution of categories in each area (action, specificity, and source). Results are expressed as a percentage of total analysis time.

Area and categories	Time distribution, %
Action	
Search	90.7
Transmit	9.3
Specificity	
Generic	66.0
Specific	34.0
Source	
List of prescriptions	37.5
Electronic medical records	28.2
Biology results	13.1
Pharmaceutical interventions	7.7
Pharmacy logistics	4.7
Drug information	4.0
Other information	3.0
Alert from prescriptions	1.8

Time Sequences

Time trends in activity categories are shown in the graphs of relative activity frequencies (Figure 1). They are based on graphs of the individual activity sequences of the 140 patients (Multimedia Appendix 4).

XSL•FO
RenderX

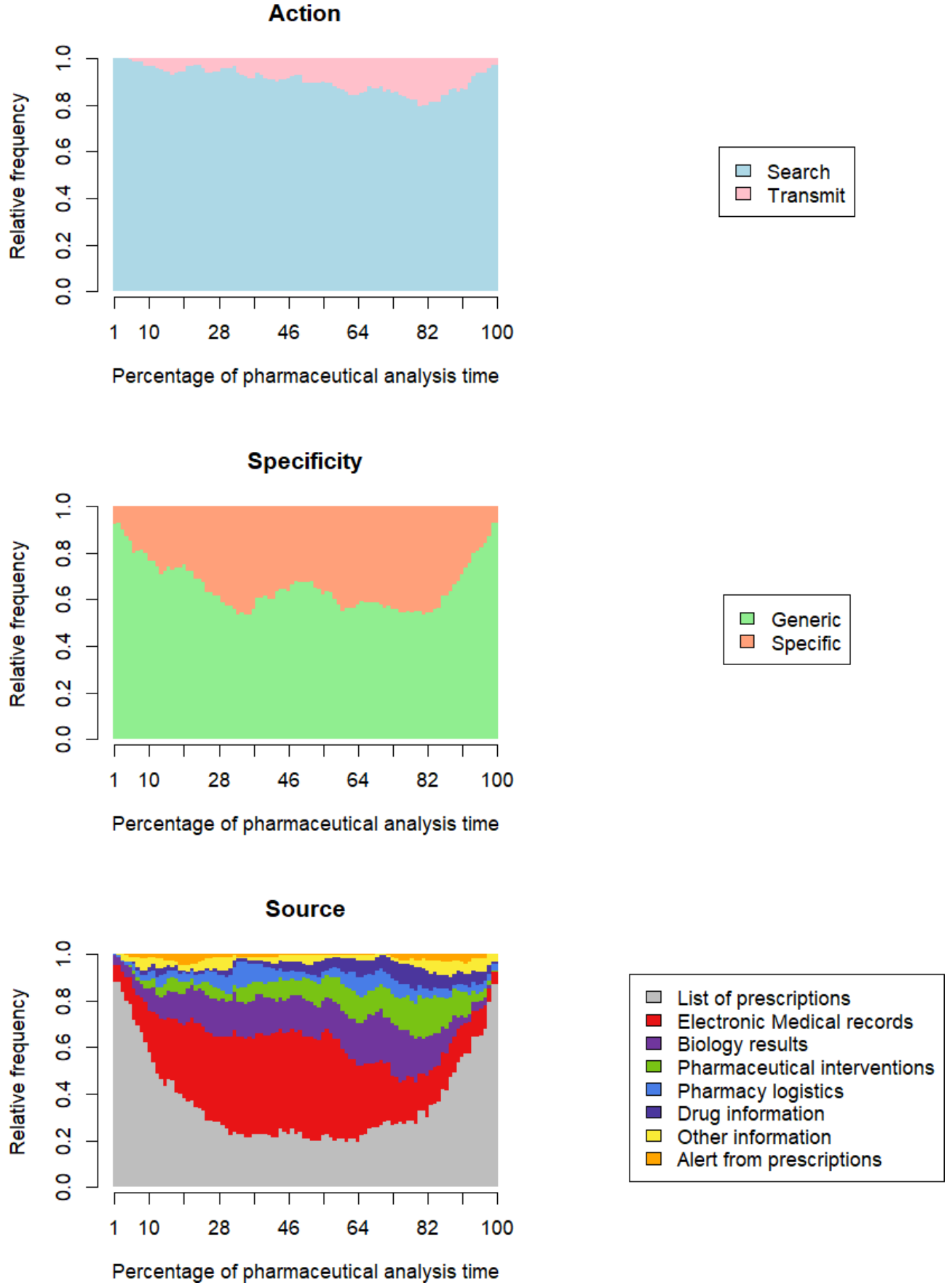
For the *action* domain, information search predominates over information transmission throughout the analysis. The transmit category rises slowly from the start of the analysis, peaking at the end, with a frequency of 20% (82% analysis time), before falling back to zero.

For the *specificity* domain, the beginning of the analysis mainly concerns generic information, with a gradual increase in the proportion of specific information as the analysis progresses. After 30% of the analysis time, the share of specific information stabilizes at around 40% of the total information consulted or transmitted. At the end of the analysis, pharmacists return to a large majority of generic information.

For the *source* domain, the list of prescriptions is consulted frequently throughout the analysis (basal frequency 20%), with 2 peaks at the beginning and end. Between 28% and 64% of the analysis time, the EMR is in the majority, representing around 40% of all activities performed during this period. The proportion of "biology results" remains fairly constant, as do "pharmacy logistics," "other information," and "alerts from prescriptions." Pharmaceutical interventions and drug information are poorly represented at the outset but increase to reach a peak of around 82% of analysis time. Details of the relative frequencies of the *source* domain for each observation are available in Multimedia Appendix 5.

XSL•FO
RenderX

Figure 1. Graphs showing the relative frequency of each type of activity in relation to the percentage of time spent analyzing the prescription.



XSL•FO
RenderX

Types of Pharmaceutical Analysis

The distance matrix between analysis sequences is generated by assigning greater weight to the source domain than to the action and specificity domains (ratio of 2:1). The reason behind this ratio is that the source domain is more objective than the specificity domain (less bias) and more discriminating than the action domain (improving the strength of the classification structure). The agglomeration coefficient of the hierarchical clustering is 0.88 (see dendrogram in Multimedia Appendix 6).

In order to facilitate interpretation and maximize class coherence, the clustering is divided into 4 groups (A, B, C, and D).

The characterization of pharmaceutical analysis typologies is done using graphs of the reference sequences for each sequence cluster (Figure 2). The results in this figure can be read using the tables describing the clusters according to covariates (Tables 2 and 3).

Figure 2. Reference activity sequences for each cluster and domain. Dist.: statistical distance (used to assess the dissimilarity or difference between two statistical objects).



XSL•FO
RenderX

Table . Description of each cluster by qualitative variables. Chi-square test *P* value: ***P*<.001; **P*<.01; *P*<.05.

Qualitative variables=Modality	Cluster/modality (%)	Modality/cluster (%)	Global (%)	*V* test
Cluster A (31 patients)				
Validation=With reservations	41.8	74.2	39.3	+4.4***
Status=Associate practitioner	37.5	67.7	40.0	+3.5***
Pharmacist=OBS08	55.6	16.1	6.4	+2.2*
Environment=Shared office	26.7	87.1	72.1	+2.1*
Patient=PAT002	43.8	22.6	11.4	+2.0*
Observation=OBS04	0.0	0.0	8.6	−2.0*
Status=Senior	12.9	29.0	50.0	−2.6**
Department=Medical	11.5	22.6	43.6	−2.7**
Center=CEN04	5.6	6.5	25.7	−2.9**
Validation=Yes	9.1	22.6	55.0	−4.1***
Cluster B (73 patients)				
Center=CEN04	86.1	42.5	25.7	+4.8***
Validation=Yes	70.1	74.0	55.0	+4.7***
Environment=Individual office	82.8	32.9	20.7	+3.7***
Status=Junior	92.9	17.8	10.0	+3.3**
Pharmacist=OBS09	92.9	17.8	10.0	+3.3**
Department=Medical	67.2	56.2	43.6	+3.1**
Pharmacist=OBS11	81.8	12.3	7.9	+2.0*
Patient=PAT002	25.0	5.5	11.4	−2.3*
Center=CEN02	16.7	2.7	8.6	−2.5*
Observation=OBS05	16.7	2.7	8.6	−2.5*
Department=Surgery	28.9	15.1	27.1	−3.3***
Environment=Shared office	42.6	58.9	72.1	−3.7***
Pharmacist=OBS04	0.0	0.0	8.6	−3.9***
Validation=With reservations	23.6	17.8	39.3	−5.5***
Cluster C (18 patients)				
Pharmacist=OBS04	83.3	55.6	8.6	+5.9***
Status=Senior	25.7	100	50.0	+4.9***
Validation=With reservations	29.1	88.9	39.3	+4.5***
Center=CEN01	29.7	61.1	26.4	+3.3**
Environment=Shared office	17.8	100	72.1	+3.1**
Department=Surgery	26.3	55.6	27.1	+2.7**
Center=CEN04	2.8	5.6	25.7	−2.2*
Environment=Individual office	0.0	0.0	20.7	−2.5*
Validation=Yes	2.6	11.1	55.0	−4.0***
Status=Associate practitioner	0.0	0.0	40.0	−4.1***

XSL•FO
RenderX

Qualitative variables=Modality	Cluster/modality (%)	Modality/cluster (%)	Global (%)	*V* test
Cluster D (18 patients)				
Environment=Open space	40.0	22.2	7.1	+2.2*
Pharmacist=OBS06	40.0	22.2	7.1	+2.2*
Validation=Yes	18.2	77.8	55.0	+2.1*
Validation=With reservations	5.5	16.7	39.3	−2.1*

Table . Description of each cluster by quantitative variables. Chi-square test *P* value: **P<.001; *P<.01.

Quantitative variables	Mean in cluster	Overall mean	SD in cluster	Overall SD	*V* test
Cluster A (31 patients)					
Analysis time (minutes)	7	5.5	2.7	3	+3.2*
Cluster B (73 patients)					
Null	N/A[a]	N/A	N/A	N/A	N/A
Cluster C (18 patients)					
Null	N/A	N/A	N/A	N/A	N/A
Cluster D (18 patients)					
Analysis time (minutes)	2.8	5.5	2	3	−4.1**

[a]N/A: not applicable.

Cluster A comprises 31 (22.1%) sequences of the 140 patients analyzed. The typical group A sequence includes a wide variety of information *sources*, with many alternating between them. The level of *specificity* is balanced between generic and specific information. The *action* of transmitting occurs mainly at the end of the analysis, in correlation with the "pharmaceutical interventions" *source*. Cluster A analyses took significantly longer for pharmacists. They are mostly carried out by assistant pharmacists, and their validations are mostly made with reservations (pharmaceutical interventions).

Cluster B involves the most patients, with 73 (52.1%) sequences of the 140 patients analyzed. The typical cluster B sequence is characterized by the *action* of searching for information (no transmission), and the *specificity* of the information is more generic than specific. The diversity of information *sources* is lower than in cluster A: pharmacists alternate between the list of prescriptions and the EMRs, with a few passages in biology results and drug information. Cluster B sequences contain almost all the sequences produced by junior pharmacists and give rise to unreserved validations (no pharmaceutical interventions). Cluster B patients are more likely to be hospitalized in medical departments than in surgical ones.

Cluster C contains 18 (12.9%) sequences of the 140 patients analyzed. The typical cluster C sequence devotes most of its analysis time to the *action* of transmitting. Most of the information is specific, with generic information used mainly at the beginning and end of the analysis. Activity *sources* are few when compared to cluster A, and alternate between list of prescriptions, pharmacy logistics, and pharmaceutical interventions. The sequences in cluster C relate more to the analysis practices of senior hospital practitioners who validate surgical services. Validation is generally carried out with reservations (pharmaceutical interventions).

Cluster D also comprises 18 (12.9%) sequences of the 140 patients analyzed. The typical cluster D sequence is the simplest of the 4 groups. It consists solely of search *actions*, with generic *specificity* and the list of prescriptions as its *source*. Cluster D analyses take significantly less time for pharmacists and result in unreserved validation (no pharmaceutical interventions). This type of analysis is more common among pharmacists working in open spaces.

Discussion

An Investigative Posture

The goal of this study is to understand how pharmacists analyze prescriptions, and in particular, whether analysis profiles exist. The 16 pharmacists who participated in the observations took an average of 5 minutes and 29 seconds to analyze and validate a patient. Although this value is close to the median (5 minutes), it should be noted that the total range of validation times is rather wide (between 30 seconds and 15 minutes) and that the distribution of values does not follow a normal distribution (Multimedia Appendix 2). We can therefore expect major disparities in analysis typologies, but we will come back to this point later.

From an overall point of view, the pharmacists' posture is overwhelmingly focused on seeking information from

XSL•FO
RenderX

prescriptions (Table 1). Whether they are identifying a drug-related problem, or justifying the absence of a problem, the professionals are constantly questioning the list of prescriptions (Figure 1). A new question leads to a new search, which in turn may lead to further questions. The pharmacists keep in mind (or physically note) the questions that come to mind when reading the list of prescriptions and then access the appropriate source of information. At the end of the analysis, as they return to the prescriptions for final validation, they make a final review in order to "not miss anything important." Overall, pharmacists are naturally at ease when interacting with the drug list, and it is during these periods that participants exchange the most with the researchers.

In terms of specificity, the results are less clear-cut than for the "action" domain. Specificity depends more on the nature of the information researched or transmitted by pharmacists. Information is more often generic in the list of prescriptions and EMR sources, and predominantly specific in the biology results, pharmaceutical interventions, and pharmacy logistics (Figure 2). The need for specific data tends to come at a more advanced stage in the pharmacist's cognitive process. In general, a specific activity follows generic research of information and is more likely to take place in a structured data source such as biology results. In the case of generic activities, we find structured information sources (prescription lists) and unstructured sources (EMRs). So, it is not the structure of the information that counts here but rather the fact of having a global view of a situation. Generic searches are therefore a way for pharmacists to retrieve as much information as possible about the drug and clinical context, and then focus on more specific activities.

Importance of Clinical Contextualization

Searching for information in a patient's medical record is by far the most cognitively demanding activity for professionals. As soon as participants immerse themselves in the clinical data, their verbalizations diminish, and their posture shifts toward reading and concentration. While the total time allocated to clinical contextualization is the second highest of all analyses (Table 1), it dominates between 20% and 70% of analysis time (Figure 1). Moreover, this activity is expressed as the most time-consuming by the participants and is probably the most critical point of the analysis. Pharmacists find themselves highly dependent on information provided by other medical professionals. The search for clinical data is made all the more difficult by the wide disparities in information architecture from one department to another and from one facility to another. Typically, pharmacists find less relevant information on treatments in surgical departments than in medical specialty departments. As for establishments, the way in which they search for clinical data depends on the software that is used for "electronic medical records" and its affinity with pharmacy practice. Finally, if we compare clinical record activities with biology activities, it is clear that the biology laboratory provides much more structured information, enabling pharmacists to target specific data and waste less time searching for it (Table 1). In short, the lack of structure in clinical data requires hospital pharmacists to make a more intense cognitive effort at the core of their analysis. This validates the hypothesis that access to relevant information is often difficult.

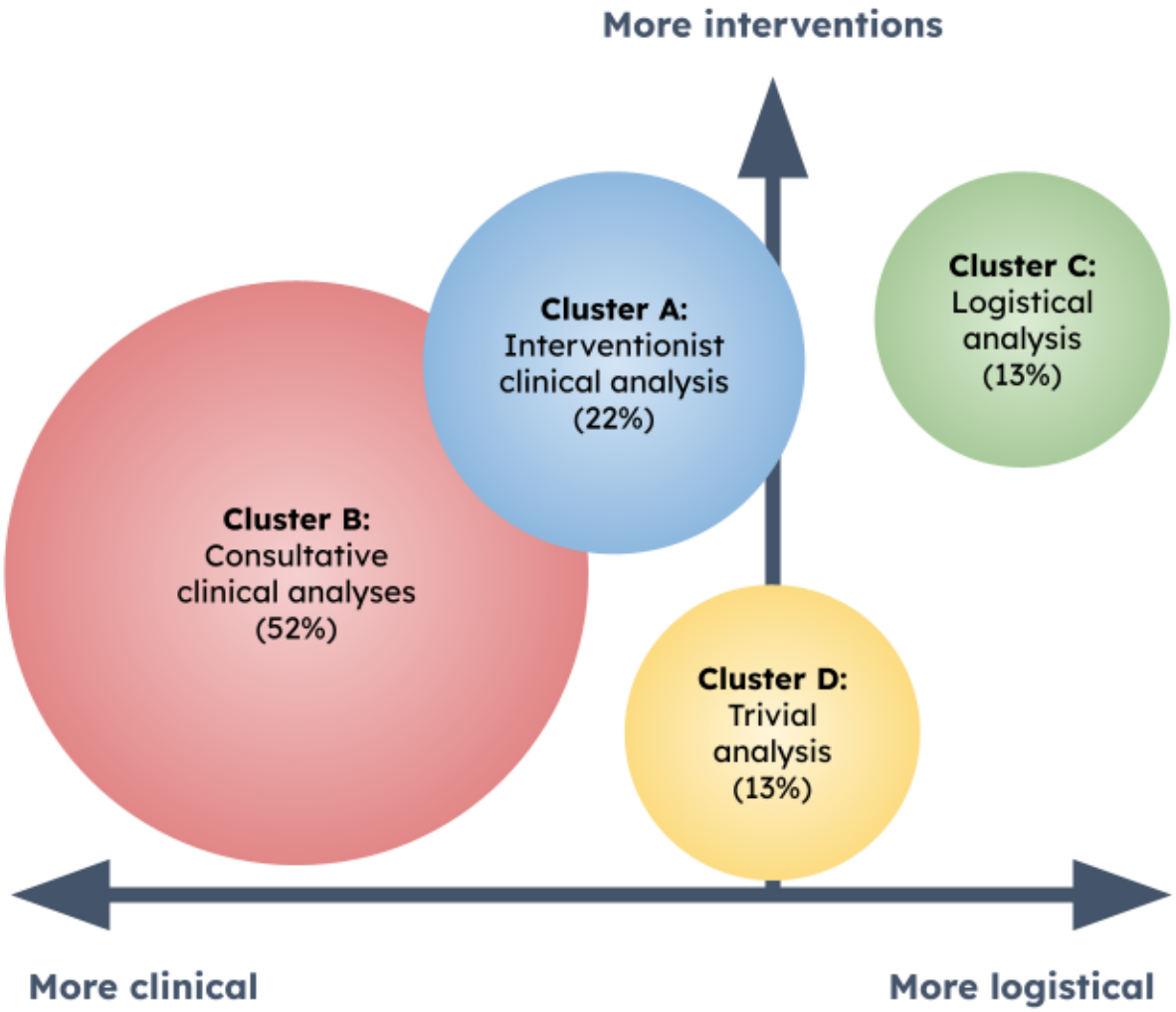
The use of other information sources is more sporadic but no less interesting. The search for drug information is correlated with the transmission of pharmaceutical interventions. Pharmacists have a strong need to verify and justify their transmission actions: Acting as drug experts who inspect prescribers' actions, they do not want to leave any room for uncertainty. In addition, they have a preventive posture toward their medical colleagues, and many participants double their written intervention with a call to ensure that it is properly transmitted. Curiously, iatrogenic alerts prompted by EMRs are rarely consulted, even though they are presented by vendors as aids to analysis and decision-making. When questioned on this subject, all the pharmacists answered that "these alerts are not appropriate," as they are "too general" and "irrelevant to the context." Finally, it is the activities related to pharmacy logistics that are the most difficult to describe from raw observations. A review by typology would be necessary to better understand the place of pharmacy logistics, as well as the disparities between the other analyses.

Several Types of Analysis

The classification of analysis sequences into 4 clusters highlights very different analysis typologies (Figure 2). This confirms the first hypothesis of the study. Groups A and B are referred to as "clinical," as they comprise sequences containing more information from the clinical and biological file than the other clusters. In total, these "clinical" sequences account for almost 75% of the analyses carried out. The remaining 25% is divided equally between clusters C and D. Although cluster D provides fewer elements for understanding professional processes (very rapid analysis, based only on prescriptions), it nonetheless illustrates situations that do not require in-depth expertise and the intervention of a pharmacist. Cluster D is ultimately a group of "trivial" analyses for drug experts. On the other hand, cluster C involves sequences that call on the logistical skills of hospital pharmacists, including their ability to intervene finely in the medication circuit. In the following, cluster C is referred to as the "logistical" group. The terms "clinical," "logistical," and "trivial" are introduced to qualify clusters more effectively and to give them a distinctive color. Analysis groups cannot be reduced to their names, as a clinical typology may include a logistical component, and vice versa. The following diagram provides a visual representation of the 4 analysis clusters (Figure 3).

XSL•FO
RenderX

Figure 3. Representation of the 4 pharmaceutical analysis clusters, according to their clinical or logistical trend, and their composition in pharmaceutical interventions. The size of the circles is proportional to the number of analyses concerned.



The 2 "clinical" clusters differ mainly in the overall dynamics of their analysis sequences. Cluster A is richer in sources of information, the data sought are more specific, and the sequence of activities is long and complex. Furthermore, these sequences regularly identify drug-related problems and generally lead to validations with reservations (pharmaceutical interventions at the end of the analysis). The value of pharmaceutical analysis of cluster A prescriptions is therefore indisputable, particularly in the fight against iatrogenesis.

Cluster A can be described as "clinical and interventionist," as opposed to cluster B, which is more "clinical and consultative." Cluster B's analysis sequences focus primarily on consulting the patient's medical file, with the addition of a few biological elements or information on treatments. It is hard to say whether consultative sequences require more cognitive effort from pharmacists than interventionist sequences. In any case, cluster B analyses concern more than 7 out of 10 "clinical" sequences and generally result in full validation (74% of cases). The impact of the clinical advisory analysis is consequently more limited. Finally, it should be noted that cluster B is devoid of logistical aspects, unlike cluster A. This is certainly because cluster A comprises a majority of associate practitioners, who are often more involved in logistics than juniors. In addition, type A analyses are generally more complex, both clinically and logistically, with an important alternation in the specificity and source of information.

The "logistics" cluster includes sequences produced exclusively by hospital practitioners (senior pharmacists). It reflects the increasing specialization of these professionals in the analysis and validation of prescriptions. In fact, their intention is to focus on the medication circuit rather than on iatrogenicity. Their interactions are therefore less with prescribers than with nurses, pharmacy assistants, and automated dispensing systems (dispensing robots and secure cabinets). The clinical aspect, while not completely omitted, tends to be secondary in this type of sequence. As expected, pharmacy logistics therefore play an integral role in prescription analysis.

Finally, "trivial" analyses are not linked to a particular category of pharmacist or hospital. Rather, they appear in relation to clinical situations that do not require the intervention of a pharmacist. This might be a straightforward surgical protocol, a prescription with a limited number of drugs, or standardized palliative care management. In all these situations, pharmacists

XSL•FO
RenderX

quickly realize that there are no drug-related problems. However, before opening the prescription list, pharmacists have no way of knowing whether these are "trivial" prescriptions. Given that it represents 13% of analyzed prescriptions, this raises the question of the place of this kind of analysis in the pharmacists' workflow.

Identifying Pharmacists' Needs

The diversity of pharmaceutical analysis typologies adds a further challenge to the design of pharmacist-centered decision support systems. Between "clinical," "logistical," and "trivial" validation intentions, the needs seem very different, and perhaps these intentions should be treated separately.

For "clinical" analysis, contextualization is a key concern. Pharmaceutical professionals need rapid access to relevant information, that is, without having to search through all the documents and unstructured data. Further, in the case of interventionist analyses, additional elements need to be included to facilitate the drafting and justification of pharmaceutical interventions.

For "logistics" validations, the need lies in the integration of stock management software and the import of highly specialized information (stock-outs, equivalences, and traceability). Since the act of validating a prescription means that it has been approved by the pharmacist, whatever the type of analysis performed, it is important that "logistical" analyses include some support for "clinical" analysis. This is to ensure that iatrogenic risk is minimized.

For "trivial" validations, the need for a decision-support system is more for prioritization. In fact, the time pharmacists devote to analysis is precious and should be used in priority for other types of validation.

Limitations

One of the first limitations of observation methods is their representativity [52]. In fact, this method requires a great deal of resources to collect field data and then to process and analyze the activities. This limits the number of observations made and calls into question the generalizability of the results. This is why the study of pharmaceutical analysis typologies is based on the analysis sequences of each patient (140 in all) rather than on each pharmacist (16 in all), and why the link between the analysis clusters and the pharmacists observed is not discussed. This also has to do with intraindividual variability: in the study, there was no indication of whether pharmacists changed their testing intention according to the time of day, the day of the week, or the month of the year. It is simply a 1-hour observation of their pharmacy practice at a given point in time.

Although our study includes 5 different hospitals, it would be necessary to generalize the results to other French hospitals and regions. Based on our data, it is not possible to assess the extent of variability attributable to different hospital workflows or pharmacist profiles. Indeed, it would be valuable to analyze the impact of hospital regions, facilities with varying workflows, or pharmacists with differing levels of experience on hospital pharmacy practices.

The aim of the study was to create a general model of pharmaceutical analysis and to form analysis groups. Unfortunately, by generalizing, it is not possible to detail every action performed by professionals. Moreover, in the health care environment, actions are rarely complete and linear [52], making it impractical to group very precise tasks into 140 different sequences. It is also for this reason that task interruption data were not exploited, as there were too many hazards and their use was not relevant to the study's objective.

Comparison With Prior Work

Few publications deal with the analysis of hospital pharmacists' practices and needs with regard to the pharmaceutical analysis of prescriptions. And most of these are usability tests rather than exploratory studies. Nevertheless, 4 studies have similar methodologies and subjects of observation [47-50]. They all use a pay-for computerized data collection tool, the validated Work Observation Method By Activity Timing (WOMBAT, Macquarie University) software [52]. This tool is very useful for observing complex situations, enabling the collection of multiple dimensions of work (who, what, when, why, and how). However, our situation is specific to a given practice, that is, analyzing prescriptions. In the publications cited above, the perception of tasks is more global, more general, and more exhaustive of the daily work of pharmacists. By combining the *action*, *specificity*, and *source* domains, we are able to better characterize tasks relating specifically to pharmaceutical analysis. On the other hand, we lose information on social interactions and other tasks such as dispensing and medication reconciliation. This difference in objectives explains why it is difficult to compare the times obtained in our study with those of other publications: the tasks studied are not homogeneous.

Three of the studies use observations to make comparisons. These are either before-and-after comparisons of the computerization of the prescription management system [50], sometimes adding a comparison between different countries [48], or comparing the activities of pharmacists with those of pharmacy technicians [47]. None of these studies attempt to group together or identify specific typologies of professional practice; the methodologies are based solely on descriptive statistics. Further, the observations are carried out in just 1 or 2 hospitals, with observed pharmacists working in the same ecosystem. This makes it more difficult to eliminate the bias associated with environmental conditions and to distinguish the needs associated with each type of pharmaceutical practice [49]. These studies are therefore not really to be considered as first steps in user-centered design.

Finally, there are some very interesting points of agreement between prior research and our results. Hospital pharmacists spend the majority of their time reviewing prescriptions, affirming the importance of the analysis activity in professional practice. We can also see from the publication by Lo et al [50] that the task of reading the patient's medical record ranks second in frequency, just after reviewing the prescription list, regardless of the presence or absence of an electronic medication management system. Prior works thus confirm the importance of clinical contextualization in making medicines safer. Even if it is difficult to compare our study to prior work, the problems

XSL•FO
RenderX

observed seem to be common. We have every reason to believe that pharmaceutical analysis is an activity whose workings do not differ fundamentally from one organization to another, and whose same needs persist over time. It should also be noted that we do not know if these needs are well addressed by CDSSs. It would be interesting to make this assessment in a future study, comparing it with the needs identified in our exploratory study.

Conclusions

This exploratory study focuses on pharmaceutical analysis and validation practices for intrahospital prescriptions. We can confirm that the process of analyzing is multifaceted and complex. It depends on the clinical or biological situation, the treatments involved, and the profile of the validating pharmacist. Given that all patients must benefit from the same quality of care, software product teams need to ensure that user needs coincide with the regulatory and ethical constraints of the health care sector. They must also take into account the diversity of pharmaceutical analyses and their specific needs. User-centered design solutions can emerge from this type of exploratory study, leaving designers free to define the right challenges to address and draw innovative ideas from them.

Beyond the interest of this study for the development of useful, usable, and acceptable digital solutions, the work presented here proposes a singular methodology for evaluating the activity of hospital pharmacists. We hope that this methodological contribution will provide a better understanding of the complexity of this medicotechnical profession and make it easier to measure professional practice evolutions, especially after the introduction of a pharmacist-oriented CDSS.

Acknowledgments

PharmIA SAS financed the study and paid the publication fee.

Authors' Contributions

JB wrote the article. N Chaniaud, EC, and N Cirnat helped with proofreading and corrections. JB and N Cirnat conducted the observations. JB carried out the data processing and analysis. JB and MA established the interest of the research topic. JB and N Chaniaud designed the study, and JS participated in the choice of statistical treatments.

Conflicts of Interest

JB, N Cirnat, MA, and EC work for PharmIA SAS, which funded the study. JS and N Chaniaud have no conflicts of interest to declare.

Multimedia Appendix 1
Observation grid for a prescription or patient.
[PNG File, 40 KB - humanfactors_v12i1e65959_app1.png]

Multimedia Appendix 2
Histogram of total analysis time versus number of patients. The horizontal whisker box represents quartiles, median, and limit values. The mean time is represented by the vertical line.
[PNG File, 76 KB - humanfactors_v12i1e65959_app2.png]

Multimedia Appendix 3
Different types of activity observed. They are formed by combining the categories of action, specificity, and source. The description column provides a better understanding of the activities resulting from these combinations.
[DOCX File, 22 KB - humanfactors_v12i1e65959_app3.docx]

Multimedia Appendix 4
Sequences of activities in each domain for the 140 patients analyzed. The abscissa represents the percentage of prescription analysis time, and the ordinate corresponds to the patient index.
[PNG File, 48 KB - humanfactors_v12i1e65959_app4.png]

Multimedia Appendix 5
Relative frequencies of source domain activities for each of the 16 observations (pharmacists).
[PNG File, 21 KB - humanfactors_v12i1e65959_app5.png]

Multimedia Appendix 6
Reference activity sequences for each cluster and domain.
[PNG File, 25 KB - humanfactors_v12i1e65959_app6.png]

XSL•FO
RenderX

References

1. Dal Mas F, Massaro M, Rippa P, Secundo G. The challenges of digital transformation in healthcare: an interdisciplinary literature review, framework, and future research agenda. Technovation 2023 May;123:102716. [doi: 10.1016/j.technovation.2023.102716]
2. Engstrom T, McCourt E, Canning M, et al. The impact of transition to a digital hospital on medication errors (TIME study). NPJ Digit Med 2023 Jul 25;6(1):133. [doi: 10.1038/s41746-023-00877-w] [Medline: 37491469]
3. Elliott RA, Camacho E, Jankovic D, Sculpher MJ, Faria R. Economic analysis of the prevalence and clinical and economic burden of medication error in England. BMJ Qual Saf 2021 Feb;30(2):96-105. [doi: 10.1136/bmjqs-2019-010206] [Medline: 32527980]
4. Magrabi F, Ammenwerth E, Hyppönen H, et al. Improving evaluation to address the unintended consequences of health information technology. Yearb Med Inform 2016 Aug;25(1):61-69. [doi: 10.15265/IY-2016-013]
5. Lehmann CU, Séroussi B, Jaulent MC. Troubled waters: navigating unintended consequences of health information technology. Yearb Med Inform 2016 Nov 10;25(1):5-6. [doi: 10.15265/IY-2016-056] [Medline: 27830225]
6. Carayon P, Wetterneck TB, Cartmill R, et al. Characterising the complexity of medication safety using a human factors approach: an observational study in two intensive care units. BMJ Qual Saf 2014 Jan;23(1):56-65. [doi: 10.1136/bmjqs-2013-001828]
7. Mamlin BW, Tierney WM. The promise of information and communication technology in healthcare: extracting value from the chaos. Am J Med Sci 2016 Jan;351(1):59-68. [doi: 10.1016/j.amjms.2015.10.015] [Medline: 26802759]
8. Barbieri C, Neri L, Stuard S, Mari F, Martín-Guerrero JD. From electronic health records to clinical management systems: how the digital transformation can support healthcare services. Clin Kidney J 2023 Nov;16(11):1878-1884. [doi: 10.1093/ckj/sfad168] [Medline: 37915897]
9. Wasylewicz ATM, Scheepers-Hoeks AMJW. Clinical decision support systems. In: Kubben P, Dumontier M, Dekker A, editors. Fundamentals of Clinical Data Science: Springer; 2019:153-169. [doi: 10.1007/978-3-319-99713-1_11]
10. Skalafouris C, Blanc AL, Grosgurin O, et al. Development and retrospective evaluation of a clinical decision support system for the efficient detection of drug-related problems by clinical pharmacists. Int J Clin Pharm 2023 Apr;45(2):406-413. [doi: 10.1007/s11096-022-01505-5] [Medline: 36515779]
11. Robert L, Quindroit P, Henry H, et al. Implementation of a clinical decision support system for the optimization of antidiabetic drug orders by pharmacists. Br J Clin Pharmacol 2024 Jan;90(1):239-246. [doi: 10.1111/bcp.15898] [Medline: 37657079]
12. Pontefract SK, Coleman JJ, Vallance HK, et al. The impact of computerised physician order entry and clinical decision support on pharmacist-physician communication in the hospital setting: a qualitative study. PLoS One 2018;13(11):e0207450. [doi: 10.1371/journal.pone.0207450] [Medline: 30444894]
13. Jermini M, Fonzo-Christe C, Blondon K, et al. Financial impact of medication reviews by clinical pharmacists to reduce in-hospital adverse drug events: a return-on-investment analysis. Int J Clin Pharm 2024 Apr;46(2):496-505. [doi: 10.1007/s11096-023-01683-w] [Medline: 38315303]
14. Karsh BT, Holden RJ, Alper SJ, Or CKL. A human factors engineering paradigm for patient safety: designing to support the performance of the healthcare professional. Qual Saf Health Care 2006 Dec;15(Suppl 1):i59-i65. [doi: 10.1136/qshc.2005.015974] [Medline: 17142611]
15. Holden RJ, Abebe E, Russ-Jara AL, Chui MA. Human factors and ergonomics methods for pharmacy research and clinical practice. Res Social Adm Pharm 2021 Dec;17(12):2019-2027. [doi: 10.1016/j.sapharm.2021.04.024] [Medline: 33985892]
16. ISO 9241-210:2019(en), Ergonomics of human-system interaction—Part 210: Human-centred design for interactive systems. International Organization for Standardization. 2019. URL: https://www.iso.org/obp/ui/fr/#iso:std:iso:9241:-210:ed-2:v1:en [accessed 2025-03-31]
17. Weir NM, Newham R, Bennie M. A literature review of human factors and ergonomics within the pharmacy dispensing process. Res Social Adm Pharm 2020 May;16(5):637-645. [doi: 10.1016/j.sapharm.2019.08.029] [Medline: 31444123]
18. Flood M, Ennis M, Ludlow A, et al. Research methods from human-centered design: potential applications in pharmacy and health services research. Res Social Adm Pharm 2021 Dec;17(12):2036-2043. [doi: 10.1016/j.sapharm.2021.06.015] [Medline: 34229952]
19. Srinivas P, Cornet V, Holden R. Human factors analysis, design, and evaluation of Engage, a consumer health IT application for geriatric heart failure self-care. Int J Hum Comput Interact 2017;33(4):298-312. [doi: 10.1080/10447318.2016.1265784] [Medline: 30429638]
20. Al Nemari M, Waterson J. The introduction of robotics to an outpatient dispensing and medication management process in Saudi Arabia: retrospective review of a pharmacy-led multidisciplinary Six Sigma performance improvement project. JMIR Hum Factors 2022 Oct 11;9(4):e37905. [doi: 10.2196/37905] [Medline: 36222805]
21. Chen ZJ, Liang WT, Liu Q, et al. Use of a remote oncology pharmacy service platform for patients with cancer during the COVID-19 pandemic: implementation and user acceptance evaluation. J Med Internet Res 2021 Jan 21;23(1):e24619. [doi: 10.2196/24619] [Medline: 33395398]
22. Weile KS, Mathiasen R, Winther JF, Hasle H, Henriksen LT. Hjernetegn.dk—the Danish central nervous system tumor awareness initiative digital decision support tool: design and implementation report. JMIR Med Inform 2024 Jul 25;12(1):e58886. [doi: 10.2196/58886] [Medline: 39052326]

XSL•FO
RenderX

23. Calvo-Cidoncha E, Verdinelli J, González-Bueno J, et al. An ontology-based approach to improving medication appropriateness in older patients: algorithm development and validation study. JMIR Med Inform 2023 Jul 13;11(1):e45850. [doi: 10.2196/45850] [Medline: 37477131]
24. Sharma V, McDermott J, Keen J, Foster S, Whelan P, Newman W. Pharmacogenetics clinical decision support systems for primary care in England: co-design study. J Med Internet Res 2024 Jul 23;26(1):e49230. [doi: 10.2196/49230] [Medline: 39042886]
25. Carayon P, Wooldridge A, Hose BZ, Salwei M, Benneyan J. Challenges and opportunities for improving patient safety through human factors and systems engineering. Health Aff (Millwood) 2018 Nov;37(11):1862-1869. [doi: 10.1377/hlthaff.2018.0723] [Medline: 30395503]
26. Martinez W, Threatt AL, Rosenbloom ST, Wallston KA, Hickson GB, Elasy TA. A patient-facing diabetes dashboard embedded in a patient web portal: design sprint and usability testing. JMIR Hum Factors 2018 Sep 24;5(3):e26. [doi: 10.2196/humanfactors.9569] [Medline: 30249579]
27. Broekhuis M, van Velsen L, Peute L, Halim M, Hermens H. Conceptualizing usability for the eHealth context: content analysis of usability problems of eHealth applications. JMIR Form Res 2021 Jul 27;5(7):e18198. [doi: 10.2196/18198] [Medline: 34313594]
28. van der Vaart R, van Driel D, Pronk K, et al. The role of age, education, and digital health literacy in the usability of internet-based cognitive behavioral therapy for chronic pain: mixed methods study. JMIR Form Res 2019 Nov 21;3(4):e12883. [doi: 10.2196/12883] [Medline: 31750839]
29. Leung CL, Alacapa J, Tasca BG, et al. Digital adherence technologies and differentiated care for tuberculosis treatment and their acceptability among persons with tuberculosis, health care workers, and key informants in the Philippines: qualitative interview study. JMIR Hum Factors 2024 Jul 23;11(1):e54117. [doi: 10.2196/54117] [Medline: 39042889]
30. St-Maurice JD, Burns CM. Modeling patient treatment with medical records: an abstraction hierarchy to understand user competencies and needs. JMIR Hum Factors 2017 Jul 28;4(3):e16. [doi: 10.2196/humanfactors.6857] [Medline: 28754650]
31. Blijleven V, Koelemeijer K, Wetzels M, Jaspers M. Workarounds emerging from electronic health record system usage: consequences for patient safety, effectiveness of care, and efficiency of care. JMIR Hum Factors 2017 Oct 5;4(4):e27. [doi: 10.2196/humanfactors.7978] [Medline: 28982645]
32. Ratanawongsa N, Matta GY, Bohsali FB, Chisolm MS. Reducing misses and near misses related to multitasking on the electronic health record: observational study and qualitative analysis. JMIR Hum Factors 2018 Feb 6;5(1):e4. [doi: 10.2196/humanfactors.9371] [Medline: 29410388]
33. Golay D, Sving C, Cajander Å. An emotion-driven approach to hospital physicians' work-related user experience. Presented at: NordiCHI '22: Nordic Human-Computer Interaction Conference; Oct 8-12, 2022; Aarhus, Denmark p. 1-12. [doi: 10.1145/3546155.3546684]
34. Vijayakumar S, Lee VV, Leong QY, Hong SJ, Blasiak A, Ho D. Physicians' perspectives on AI in clinical decision support systems: interview study of the CURATE.AI personalized dose optimization platform. JMIR Hum Factors 2023 Oct 30;10(1):e48476. [doi: 10.2196/48476] [Medline: 37902825]
35. Sääskilahti M, Ahonen R, Timonen J. Pharmacy customers' experiences of use, usability, and satisfaction of a nationwide patient portal: survey study. J Med Internet Res 2021 Jul 16;23(7):e25368. [doi: 10.2196/25368] [Medline: 34269687]
36. Mercer K, Burns C, Guirguis L, et al. Physician and pharmacist medication decision-making in the time of electronic health records: mixed-methods study. JMIR Hum Factors 2018 Sep 25;5(3):e24. [doi: 10.2196/humanfactors.9891] [Medline: 30274959]
37. Lallemand C, Gronier G. Méthodes de Design UX: 30 Méthodes Fondamentales Pour Concevoir Des Expériences Optimales, 2nd edition: Eyrolles; 2018.
38. Wu JP, Damschroder LJ, Fetters MD, et al. A web-based decision tool to improve contraceptive counseling for women with chronic medical conditions: protocol for a mixed methods implementation study. JMIR Res Protoc 2018 Apr 18;7(4):e107. [doi: 10.2196/resprot.9249] [Medline: 29669707]
39. Valaitis R, Cleghorn L, Vassilev I, et al. A web-based social network tool (GENIE) for supporting self-management among high users of the health care system: feasibility and usability study. JMIR Form Res 2021 Jul 12;5(7):e25285. [doi: 10.2196/25285] [Medline: 34255654]
40. Sturesson L, Groth K. Effects of the digital transformation: qualitative study on the disturbances and limitations of using video visits in outpatient care. J Med Internet Res 2018 Jun 27;20(6):e221. [doi: 10.2196/jmir.9866]
41. Østervang C, Vestergaard LV, Dieperink KB, Danbjørg DB. Patient rounds with video-consulted relatives: qualitative study on possibilities and barriers from the perspective of healthcare providers. J Med Internet Res 2019 Mar 25;21(3):e12584. [doi: 10.2196/12584] [Medline: 30907746]
42. Drews FA, Markewitz BA, Stoddard GJ, Samore MH. Interruptions and delivery of care in the intensive care unit. Hum Factors 2019 Jun;61(4):564-576. [doi: 10.1177/0018720819838090]
43. Hohl CM, Small SS, Peddie D, Badke K, Bailey C, Balka E. Why clinicians don't report adverse drug events: qualitative study. JMIR Public Health Surveill 2018 Feb 27;4(1):e21. [doi: 10.2196/publichealth.9282] [Medline: 29487041]

XSL•FO
RenderX

44. Cresswell K, Hinder S, Sheikh A, et al. Complex hospital-based electronic prescribing-based intervention to support antimicrobial stewardship: qualitative study. JMIR Form Res 2024 Jul 26;8(1):e54458. [doi: 10.2196/54458] [Medline: 39059001]
45. Mörike F, Spiehl HL, Feufel MA. Workarounds in the shadow system: an ethnographic study of requirements for documentation and cooperation in a clinical advisory center. Hum Factors 2024 Mar;66(3):636-646. [doi: 10.1177/00187208221087013] [Medline: 35471848]
46. MacKenna B, Bacon S, Walker AJ, Curtis HJ, Croker R, Goldacre B. Impact of electronic health record interface design on unsafe prescribing of ciclosporin, tacrolimus, and diltiazem: cohort study in English national health service primary care. J Med Internet Res 2020 Oct 16;22(10):e17003. [doi: 10.2196/17003] [Medline: 33064085]
47. Magee K, Fromont M, Ihle E, et al. Direct observational time and motion study of the daily activities of hospital dispensary pharmacists and technicians. J Pharm Pract Res 2023 Apr;53(2):64-72. [doi: 10.1002/jppr.1852]
48. Westbrook JI, Li L, Shah S, et al. A cross-country time and motion study to measure the impact of electronic medication management systems on the work of hospital pharmacists in Australia and England. Int J Med Inform 2019 Sep;129:253-259. [doi: 10.1016/j.ijmedinf.2019.06.011]
49. Lehnbom EC, Li L, Prgomet M, Lam WY, Westbrook JI. Little things matter: a time and motion study of pharmacists' activities in a paediatric hospital. In: Digital Health Innovation for Consumers, Clinicians, Connectivity and Community: IOS Press; 2016, Vol. 227:80-86. [doi: 10.3233/978-1-61499-666-8-80]
50. Lo C, Burke R, Westbrook JI. Electronic medication management systems' influence on hospital pharmacists' work patterns. J Pharm Pract Res 2010 Jun;40(2):106-110. [doi: 10.1002/j.2055-2335.2010.tb00516.x]
51. Cavaye D, Lehnbom EC, Laba TL, El-Boustani E, Joshi R, Webster R. Considering pharmacy workflow in the context of Australian community pharmacy: a pilot time and motion study. Res Social Adm Pharm 2018 Dec;14(12):1157-1162. [doi: 10.1016/j.sapharm.2018.01.003] [Medline: 29317188]
52. Tanzini M, Westbrook JI, Guidi S, Sunderland N, Prgomet M. Measuring clinical workflow to improve quality and safety. In: Donaldson L, Ricciardi W, Sheridan S, Tartaglia R, editors. Textbook of Patient Safety and Clinical Risk Management: Springer; 2021:393-402. [doi: 10.1007/978-3-030-59403-9_28]
53. Recommandations de bonnes pratiques: bonnes pratiques de pharmacie clinique. : Société Française de Pharmacie Clinique; 2022 Mar URL: https://sfpc.eu/wp-content/uploads/2022/03/Bonnes-Pratiques-de-pharmacie-Clinique-2022.pdf [accessed 2025-03-31]
54. American College of Clinical Pharmacy. Standards of practice for clinical pharmacists. J Am Coll Clin Pharm 2023 Oct;6(10):1156-1159. [doi: 10.1002/jac5.1873]
55. Pasupathy KS, Clinic M, Steege LM, Cho CC. Technology implementation and associated pharmacy interruptions. AMIA Annu Symp Proc 2020 Mar;2019:707-716. [Medline: 32308866]
56. Czarniawska B. Shadowing: And Other Techniques for Doing Fieldwork in Modern Societies: Copenhagen Business School Press; 2007.
57. Isaacs E, Szymanski MH. Capturing mobile telepresence through logging and video shadowing: a two-phase study design. Field Methods 2013 Aug;25(3):299-315. [doi: 10.1177/1525822X12466802]
58. Mariette J, Olteanu M, Vialaneix N. Kernel and dissimilarity methods for exploratory analysis in a social context. In: Daouia A, Ruiz-Gazen A, editors. Advances in Contemporary Statistics and Econometrics: Festschrift in Honor of Christine Thomas-Agnan: Springer; 2021:669-690. [doi: 10.1007/978-3-030-73249-3_34]
59. Gabadinho A, Ritschard G, Müller NS, Studer M. Analyzing and visualizing state sequences in R with TraMineR. J Stat Softw 2011 Apr;40(4):1-37. [doi: 10.18637/jss.v040.i04]
60. Ritschard G, Liao TF, Struffolino E. Strategies for multidomain sequence analysis in social research. Sociol Methodol 2023 Aug;53(2):288-322. [doi: 10.1177/00811750231163833]
61. cluster: "Finding Groups in Data": Cluster Analysis Extended Rousseeuw et al. R Project. 2023. URL: https://cran.r-project.org/web/packages/cluster/index.html [accessed 2025-03-31]
62. Lê S, Josse J, Husson F. FactoMineR: an R package for multivariate analysis. J Stat Softw 2008;25(1):1-18. [doi: 10.18637/jss.v025.i01]

Abbreviations

CDSS: clinical decision support system
EMR: electronic medical record
HFE: human factors and ergonomics
WOMBAT: Work Observation Method by Activity Timing

XSL•FO
RenderX

Edited by A Kushniruk, J Waterson; submitted 04.09.24; peer-reviewed by B Fadojutimi, D Duggal, O Egunlae; revised version received 23.11.24; accepted 25.11.24; published 25.04.25.

Please cite as:
Butruille J, Cirnat N, Alaoui M, Saracco J, Cousein E, Chaniaud N
Pharmaceutical Analysis of Inpatient Prescriptions: Systematic Observation of Hospital Pharmacists' Practices in the Early User-Centered Design Phase
JMIR Hum Factors 2025;12:e65959
URL: https://humanfactors.jmir.org/2025/1/e65959
doi:10.2196/65959

© Jesse Butruille, Natalina Cirnat, Mariem Alaoui, Jérôme Saracco, Etienne Cousein, Noémie Chaniaud. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 25.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Web-Based Delivery of an Effective Church-Based Intervention Program to Promote Cancer Screening (Community-based Health litEracy-focused intervention for breast and cervical Cancer Control) Among Korean Immigrant Women in the United States: Randomized Controlled Trial

Hae-Ra Han[1,2], PhD, MSN; Yoon-Jae Lee[3], DNP; Deborah Min[2,4], MPH; Joyline Chepkorir[1], PhD; DaSol Amy Hwang[1], FNP, MSN, MAT; Steve Chae[1], MSN, RN

[1]Johns Hopkins School of Nursing, 525 N. Wolfe St, Baltimore, MD, United States

[2]Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

[3]University of Pennsylvania School of Nursing, Philadelphia, PA, United States

[4]New York University Grossman School of Medicine, New York, NY, United States

Corresponding Author:
Hae-Ra Han, PhD, MSN
Johns Hopkins School of Nursing, 525 N. Wolfe St, Baltimore, MD, United States

Abstract

Background: Women with limited English proficiency continue experiencing an unequal cancer burden. Non-White immigrant women present with more advanced breast and cervical cancer than non-Hispanic Whites, attributed to significant cultural barriers as well as low health literacy in attempting to navigate the United States health care system for cancer screening. Community-based Health litEracy-focused intervention for breast and cervical Cancer Control (CHECC-uP) was an in-person, community health worker-led intervention, addressing both cultural and health literacy barriers through health literacy education and follow-up counseling with navigation assistance. The in-person program was tested in a large cluster-randomized trial and yielded high efficacy in promoting mammogram and Papanicolaou test screening in Korean-speaking women. With over 90% of Americans now having online access, the in-person program was adapted to web-based delivery.

Objective: This study aimed to evaluate the feasibility, acceptability, and preliminary efficacy of the web version of the intervention—e-CHECC-uP.

Methods: A randomized pilot trial was conducted. A total of 40 women were enrolled and randomized (20 per arm). The study intervention consisted of web-based health literacy education followed by phone counseling with navigation assistance. Study assessments were done at baseline, 3 months, and 6 months. The study's primary outcomes were cancer screening behaviors verified by medical record review. Upon completion of final study assessments, intervention participants were invited to join postintervention interviews.

Results: In total, 34 women (intervention: n=15; control: n=19) completed the assessment at 6 months, yielding a retention rate of 85%. The intervention participants were highly satisfied with e-CHECC-uP with a median rating of 8 on a 10-point scale. Between-group differences in screening rates were 34.6%, 47.9%, and 37.5%, respectively, for completion of the mammogram, Papanicolaou test, and both at 6 months.

Conclusions: We achieved a high retention rate and high participant satisfaction. Although the study was not powered for significance testing, the e-CHECC-uP intervention resulted in large group differences across all cancer screening outcomes in the pilot sample. Online technology can help address multiple logistical barriers associated with in-person intervention delivery. Our findings suggest that web-based delivery of CHECC-uP may be used to promote cancer screening among immigrant women with limited English proficiency, as a promising avenue to ultimately reduce health disparities in underserved communities.

Trial Registration: ClinicalTrials.gov NCT03726619; https://clinicaltrials.gov/study/NCT03726619

(JMIR Hum Factors 2025;12:e66092) doi:10.2196/66092

KEYWORDS

health literacy; digital health; cancer screening; community health workers; Korean Americans

XSL•FO
RenderX

Introduction

Despite significant progress in the early detection of breast and cervical cancer in the United States, certain racial and ethnic minoritized groups—particularly those with limited English proficiency (LEP)—continue to experience an unequal cancer burden, which is strongly associated with underusage of cancer screening. Non-White immigrant women present with more advanced disease than do non-Hispanic Whites [1,2], which is attributed to significant cultural barriers [1-3] as well as low health literacy [2-6] in attempting to navigate the health care system in the United States. Korean Americans are the fifth largest Asian population in the United States [7] and represent one of the most vulnerable groups for late cancer diagnosis due to significant language barriers and limited health literacy [6,8].

Community health worker (CHW)–led interventions have grown rapidly in health care as possible solutions to alleviating health inequalities. CHW-led interventions have demonstrated to be efficacious and are a commonly used approach in promoting health outcomes among vulnerable populations at risk for health disparities [9-11]. A systematic review revealed that CHW-led interventions promoting cancer screening resulted in 6% - 33% between-group differences for mammography and 7% - 29% for Papanicolaou screening [11]. Likewise, a randomized clinical trial involving Asian immigrant women showed that a CHW-led health literacy-focused intervention was highly efficacious in promoting breast and cervical cancer screening with a group difference of 46% for mammography and 45% for Papanicolaou screening [12].

eHealth—"the use of information and communication technologies for health" [13]—is an effective health communication and educational approach for preventive and self-care practices [14,15]. Today, it is estimated that over 90% of Americans have online access [16]. According to a recent community survey [17], 66% of Korean immigrant women have access to the Internet and use the internet as the main method to obtain health information; 98% of them have access to a mobile phone. In light of high digital accessibility, the potential of eHealth as an effective health communication and education tool has been demonstrated [15]. Although CHW-led interventions have shown efficacy in promoting cancer screening among underserved populations, there remains a critical gap in scalable, accessible delivery models—particularly digital formats—that can extend the reach and impact of these interventions for immigrant women with LEP. This study directly addresses this gap by adapting a previously successful CHW-led, health literacy-focused cancer screening intervention (CHECC-uP: Community-based Health litEracy-focused intervention for breast and cervical Cancer Control [12]) into a web-based format (e-CHECC-uP), leveraging the high rate of digital access among Korean American immigrant women. The purpose of this study was to evaluate the acceptability and feasibility of e-CHECC-uP. The study also aimed to address preliminary efficacy of the e-CHECC-uP intervention in addressing cancer screening and psychosocial outcomes.

Methods

Design and Sample

A pilot randomized trial was conducted to test the feasibility, acceptability, and preliminary efficacy of e-CHECC-uP in 40 Korean American women. Eligibility criteria included: (1) women self-identifying as Korean living in the Baltimore-Washington Metropolitan area, (2) aged 21 to 65 years, and (3) overdue for mammogram (>2 y since last mammogram for women 40 y and older) [18] or Papanicolaou smear screening (>3 y since last Papanicolaou test for women 21 y and older) [19]. Study participants were recruited from Korean churches in the Baltimore-Washington Metropolitan area. Using computer-generated random numbers, enrolled study participants who completed the baseline study survey were randomized into either the intervention or the control arms (n=20 each). This study was conducted in line with CONSORT (Consolidated Standards of Reporting Trials) guidelines (Checklist 1).

E-CHECC-uP Intervention

The study intervention consisted of web-based health literacy education and monthly phone counseling with navigation assistance over 6 months. The prototype intervention CHECC-uP consisted of health literacy education and follow-up phone counseling, both delivered by trained CHWs, and was successfully tested in Korean immigrant women from ethnic churches [12]. The health literacy education, which was delivered in-person, was designed to help Korean American women to better understand essential medical terminology used in mammogram and Papanicolaou test screening with relevant medical instructions, while increasing familiarity with the navigational processes to get these screenings done. In order to scale up, we transformed the print materials used for in-person health literacy education into a web-based platform. Details of the online module development will be described elsewhere. Briefly, we worked with a community advisory board to develop the web-based platform of the health literacy education component of CHECC-uP. The community advisory board consisted of 7 Korean American women aged 40 - 65 years residing in California or Maryland. Over 2 months, the advisory board members assessed module contents for its organization, relevance, and comprehension. Also, they reviewed both audio and visual quality of the modules and consistency and overall design of them. The final version consisted of 9 educational modules with an estimated completion time of approximately 30 to 60 minutes.

Procedures

After the online modules were finalized, a trained bilingual research team worked with 5 ethnic churches to identify potential study participants. Upon verbal verification of study eligibility, potential participants were given detailed information about the study and an informed consent form. As the participants signed the form and completed the baseline questionnaire, each participant was assigned to one of 2 groups. For the intervention group, trained study staff handed the participants a tablet along with a log-in instruction sheet with individualized ID and password. Once successfully logged in,

XSL•FO
RenderX

the participants were guided to touch the "Start" button on the screen when they were ready to begin and stated that the narrator would walk them through the modules. The staff stood by to answer any questions while the intervention participants completed the modules. Upon module completion, intervention participants received a hard copy version of the educational contents covered in the modules, along with detailed instructions to enable them to log in and view the modules at their convenience. Within 2 - 3 weeks after completing the modules, intervention women received monthly phone counseling for about 6 months. The purpose of phone counseling was for CHWs to address participant questions or concerns about mammogram and Papanicolaou test screening as well as monitor participant progress toward screening completion, while providing necessary navigation assistance. For both the intervention and control groups, a bilingual research assistant provided an informational brochure in Korean about breast and cervical cancer and mammogram and Papanicolaou smear screening.

Trained bilingual staff who were masked to the group assignment collected data at baseline, 3 months, and 6 months. During the follow-up phase of the study, our team developed Qualtrics surveys for participants who were not available to come to our scheduled in-person data collection visit at each community site. In addition, at 6 months, intervention participants were asked to join postintervention in-person or phone interviews to share their experiences with e-CHECC-uP. The main goal of these interviews was to better understand intervention participants' experiences so that the study team could identify areas for improvement and enhance future implementations. Example questions included: "Please share what you gained from participating in the e-CHECC-uP program. What was the most valuable part or lesson of the program? What made it more valuable than others?" or "What did you like least about the program? What changes would you suggest to help improve the e-CHECC-uP program?" All interviews were recorded and transcribed verbatim.

Measurement

A study questionnaire was developed to collect sociodemographic and medical characteristics of participants. Primary outcomes were age-appropriate mammogram and Papanicolaou test screening uptake [18,19], which were verified by medical record review. Secondary outcomes included health literacy, social support, cancer screening-related self-efficacy, and quality of life. The questionnaire was translated into Korean and available to participants. All study instruments were used at baseline, 3 months, and 6 months except for social support and quality of life, which were administered at baseline and 6 months. In addition, the 6-month survey included items addressing satisfaction with the e-CHECC-uP intervention (intervention group only).

Health literacy was measured with the Assessment of Health Literacy in Cancer Screening (AHL-C). The AHL-C includes a total of 52 items addressing five dimensions of health literacy in the context of breast and cervical cancer screening: familiarity, comprehension, reading ability, navigation, and numeracy [20]. Based on previous research [21], we chose 3 subscales addressing familiarity, comprehension, and reading ability. Example items are "Please read out loud the listed words below" and "Please find the correct meaning of each word listed below" [20]. The AHL-C has been validated among Korean women in the United States [20] and in Korea [22], and Black women living with HIV in the United States [23] with α coefficients ranging from 0.70 to 0.96.

We used the 8-item Medical Outcomes Study Social Support Survey (MOS-SS) to measure social support. This shorter version was modified from the original 19-item MOS-SS and includes the emotional and instrumental domains of social support on a 5-point Likert-type scale (1=none of the time to 5=all of the time), with higher scores indicating greater perceived social support [24]. An example item is "How often is someone available to take you to the doctor if you need it?" The shorter version had excellent reliability coefficients ranging from 0.88 to 0.92 [24].

Cancer screening-related self-efficacy was measured by the Korean translated Breast and Cervical Cancer Self-Efficacy scale which measures how confident a woman is in carrying out each of the 4 tasks in relation to mammogram and Papanicolaou tests, respectively. Each scale uses a 4-point Likert type item (1=not at all confident to 4=very confident) with higher scores indicating higher confidence in getting cancer screening [25]. The reliability of the Korean version had an α coefficient of 0.92.

We assessed quality of life using the Korean version of the EQ-5D [26]. The EQ-5D is a widely used validated instrument measuring quality of life [27]. It defines health in 5D: mobility, self-care, usual activities, pain or discomfort, and anxiety or depression. Each dimension has 3 response levels (no problems, some problems, and extreme problems) [27]. In the validation study, the Korean version of the EQ-5D had good reliability and validity through various methods including test-retest reliability using intra-class correlations and Kappa statistics along with a priori hypothesis testing [26].

Analysis

Final analysis included data from 34 women who completed 6 months assessments and whose cancer screening status was confirmed objectively by medical records. We performed descriptive statistics using means, SD, and frequencies to summarize sample characteristics and study variables and to compare the primary outcomes of the completion of a mammogram, a Papanicolaou test, or both at 6 months. We also used chi-square tests and independent samples *t* tests to compare intervention and control groups at baseline. Statistical significance was determined at P=.05. For secondary outcomes, effect sizes were calculated by dividing the difference in mean changes (from baseline to 3 mo and from baseline to 6 mo) by the pooled SD at baseline [28]. Finally, for postintervention interview data, we used a thematic analysis [29] to identify common themes related to participants' experiences with the e-CHECC-uP program. In total, 2 trained research assistants independently coded the interview data to identify themes, which were then reviewed to determine those that were most common and relevant. Any discrepancies in coding were resolved through team consensus.

XSL•FO
RenderX

Ethical Considerations

This study was approved by the Johns Hopkins Medicine institutional review board (IRB00180383). Every participant provided written informed consent before their participation, at which time they were informed of the study's aims, procedures, and the benefits and potential risks. The voluntary nature of their participation was underscored, and participants could withdraw from the study at any given time without any negative consequences. Participants were compensated US $20 in cash, each at baseline survey and 3-month follow-up, and US $40 in cash at 6-month follow-up. Those who participated in the postintervention interview were compensated an additional US $30 in cash at the end of the interview.

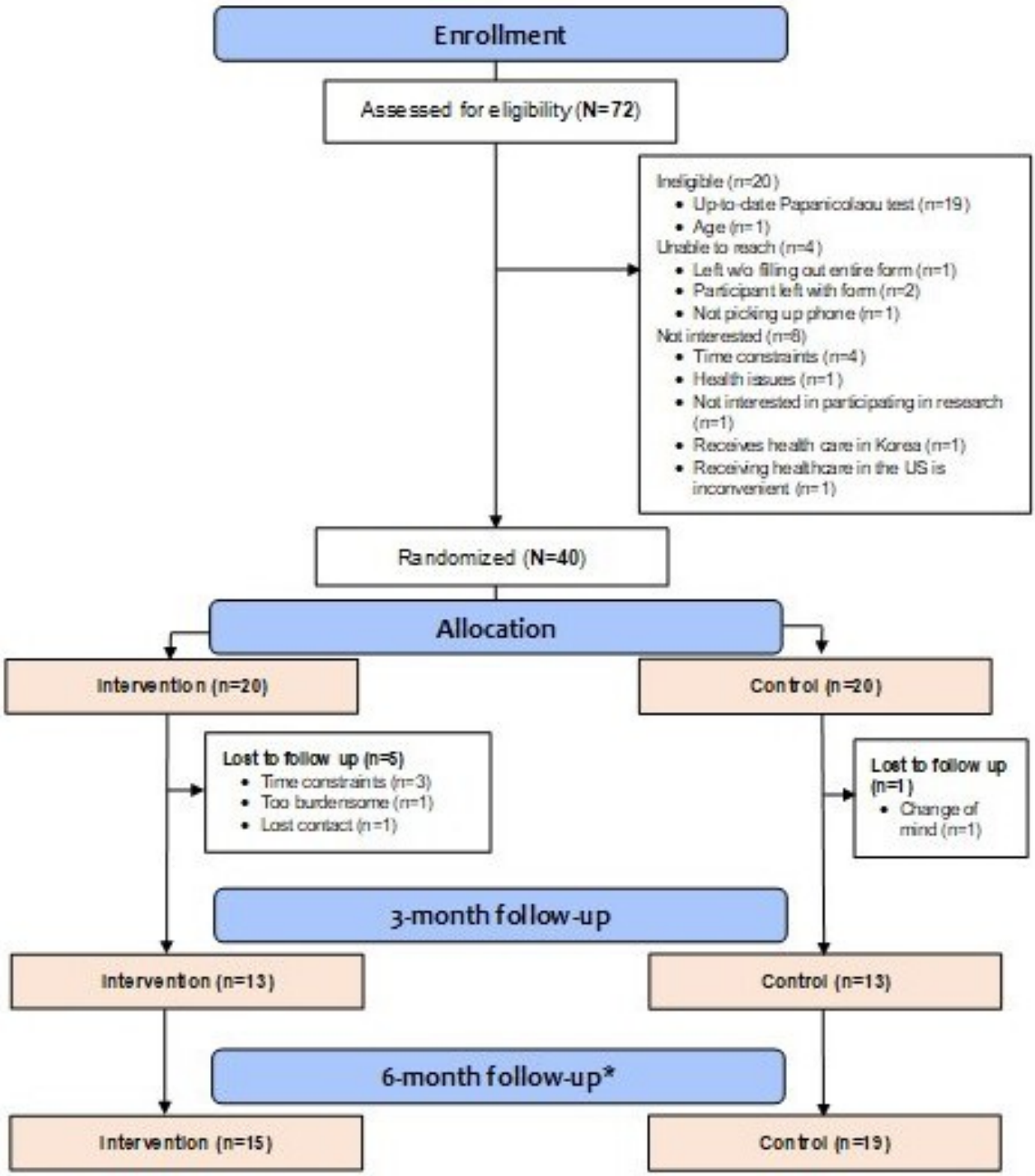
Results

Sample Characteristics

A total of 52 of the 72 women who were screened for this study were identified as eligible for enrollment. In total, 8 of those eligible declined to participate in our study. In addition, we were unable to reach 4 eligible women after initial contact, yielding a total of 40 women enrolled in the study (77% response rate). A total of 34 participants (intervention: n=15; control: n=19) completed the final study assessment at 6 months (retention rate=85%; refer to Figure 1 for details on recruitment and attrition). Table 1 summarizes participant characteristics at baseline. The final sample was mostly middle-aged (mean 46.3, SD 11.4 years), married (27/34, 79.4%), employed (29/34, 85.3%), and highly educated (years of education: mean 15.6, SD 3.0 years). While more than two-thirds (25/34, 73.5%) reported just okay, difficult, or very difficult to manage with current income, 76.5% (26/34) had limited English proficiency despite living in the United States for an average of 17.3 (SD 10.5) years. The majority of the sample had health insurance (27/34, 79.4%); nevertheless, less than one-third (11/34, 32.4%) indicated that they had a primary care provider at the time of the study. Nearly all (32/34, 94.1%) had access to the internet with 100% (34/34) of the sample reporting mobile phone access. Table 1 also compares the sample characteristics by group. The intervention and control groups were generally compatible at baseline except English proficiency: the intervention group had a significantly higher proportion of women with limited English proficiency (ie, speaking English less than proficient) than the control group (93.3%, 14/15 vs 63.2%, 12/19).

XSL•FO
RenderX

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flow diagram. The asterisk (*) here indicates that the participants missed the 3-month follow-up but completed the 6-month final survey (n=8).



XSL•FO
RenderX

Table . Sample characteristics at baseline.

Variable	Total (N=34)	Intervention (n=15)	Control (n=19)	*P* value (chi-square or *t* test)
Age (years; range 22 - 62 years), mean (SD)	46.3 (11.4)	48.9 (10.1)	44.3 (12.2)	.24
Married, %	79.4	86.7	73.7	.35
Employed full- or part-time, %	85.3	80.0	89.5	.44
Education (years; range 12 - 20 years), mean (SD)	15.6 (3.0)	16.0 (0.0)	15.3 (4.2)	.84
Just okay, difficult, or very difficult to manage with current income, %	73.5	80.0	68.4	.45
Speaking English: none, little, or good, %	76.5	93.3	63.2	.04
Years in the United States (years; range 0 - 38 years), mean (SD)	17.3 (10.5)	15.7 (11.9)	18.6 (9.2)	.43
Have health insurance, %	79.4	86.7	73.7	.35
Have a primary care provider, %	32.4	40.0	26.3	.40
Have internet access, %	94.1	93.3	94.7	.86
Have mobile phone access, %	100	100	100	—

Outcome Changes

Table 2 compares the study's primary outcomes of cancer screening behaviors between intervention and control groups at 6 months. Of 34 women included in the analysis sample, 26, 22, and 14 were overdue for a mammogram, Papanicolaou test, or both tests, respectively, at the time of study enrollment. There were noticeable differences in screening rates between intervention and control groups for completion of mammogram, Papanicolaou test, and both tests at 6 months (34.6%, 47.9%, and 37.5%, respectively).

Table . Comparison of behavioral outcomes between groups at 6 months (N=34).

Screening test	Intervention group, n/n (%)	Control group, n/n (%)	Difference between groups, %
Mammogram (n=26)	5/12 (41.7)	1/14 (7.1)	34.6
Papanicolaou test (n=22)	5/9 (55.6)	1/13 (7.7)	47.9
Both tests (n=14)	3/6 (50.0)	1/8 (12.5)	37.5

Table 3 presents mean scores of psychosocial outcomes at baseline, 3 months, and 6 months for the intervention and control groups, along with mean change scores and effect sizes. e-CHECC-uP resulted in small to large effect sizes across various psychosocial outcomes measured. Specifically, e-CHECC-uP yielded small effect sizes for reading ability and for both self-efficacy measures related to breast and cervical cancer screening (effect sizes=0.10, 0.06, and 0.06, respectively). One of the dimensions of health literacy—comprehension—resulted in an effect size of 0.27 at 3 months, but the effect was attenuated and in a negative direction at 6 months. Social support was another variable with a medium-size effect of 0.37. Finally, e-CHECC-uP was associated with medium to large effect sizes for a health literacy domain—familiarity—and quality of life with values for each being 0.58 and 0.65 at 6 months.

XSL•FO
RenderX

Table . Changes in psychosocial outcomes over 6 months (N=34).

Outcome	Values, mean (SD)						Mean change (SD) at 3 months[a]		Mean change (SD) at 6 months[b]		Effect size at 3 months[c]	Effect size at 6 months[d]
	Baseline		3 months		6 months							
	Intervention	Control	Intervention	Control	Intervention	Control	Intervention	Control	Intervention	Control		
Familiarity	22.0 (12.3)	28.9 (13.1)	26.4 (13.5)	33.2 (16.5)	31.1 (11.7)	31.1 (14.6)	3.1 (7.1)	1.7 (8.2)	9.1 (9.6)	1.6 (7.1)	0.11	0.58
Comprehension	5.7 (3.6)	4.4 (3.7)	7.6 (3.6)	5.6 (4.4)	7.7 (4.0)	7.3 (3.8)	2.2 (3.5)	1.2 (3.7)	2.1 (4.5)	2.9 (4.0)	0.27	−0.22
Reading ability	6.6 (3.2)	6.1 (4.2)	7.7 (2.4)	9.2 (3.1)	9.4 (2.1)	9.0 (3.1)	0.7 (2.9)	0.7 (4.6)	2.8 (2.4)	2.4 (3.8)	0	0.1
MOS-SS[e,f]	27.5 (6.0)	30.4 (7.1)	—[g]	—	29.8 (7.3)	30.2 (6.5)	—	—	2.3 (8.0)	−0.2 (3.7)	—	0.37
Self-Efficacy-B[h]	10.7 (1.4)	11.1 (1.9)	10.5 (1.7)	11.1 (2.1)	10.7 (1.7)	11.0 (1.8)	−0.1 (1.8)	0.0 (1.3)	0.0 (1.8)	−0.1 (1.6)	−0.06	0.06
Self-Efficacy-C[i]	10.5 (1.5)	11.2 (1.6)	10.4 (1.9)	11.3 (2.1)	10.7 (1.4)	11.3 (1.8)	0.1 (2.2)	0.1 (1.8)	0.3 (1.5)	0.2 (2.1)	0	0.06
EQ-5D[e]	8.7 (1.8)	9.7 (0.7)	—	—	7.9 (2.0)	8.1 (2.2)	—	—	−0.8 (2.0)	−1.7 (2.3)	—	0.65

[a]Mean change from baseline to 3 months.

[b]Mean change from baseline to 6 months.

[c]Difference in mean change from baseline to 3 months divided by the pooled SD at baseline.

[d]Difference in mean change from baseline to 6 months divided by the pooled SD at baseline.

[e]Assessment performed at baseline and 6 months only.

[f]MOS-SS: Medical Outcomes Survey-Social Support.

[g]Not applicable.

[h]Self-Efficacy-B: Self-Efficacy-Breast Cancer.

[i]Self-Efficacy-C: Self-Efficacy-Cervical Cancer.

Acceptability and Feasibility

e-CHECC-uP was well received by the intervention participants. When asked about how satisfied they were with the e-CHECC-uP overall, 11 of 15 (73.3%) rated 7 or higher on a 10-point scale (range 5 - 10, median 8, mean 7.6, SD 1.8). In total, 13 (86.7%) women were somewhat satisfied or satisfied with the information they learned about mammogram and Papanicolaou test screening, as well as with the way they learned about the screening; 2 women were not satisfied with either. Similarly, 13 (86.7%) women were somewhat satisfied or satisfied with phone counseling, while 2 women were not satisfied. When given an opportunity to comment further, the 2 participants who were not satisfied said the program was either "too tiresome" or "too difficult for a person who is not familiar or comfortable with English."

Using postintervention interview data, we identified a few themes addressing the experiences of intervention participants: Recognizing the importance of cancer screening, improved health literacy and self-confidence, and improved knowledge about screening procedures. In terms of recognizing the importance of cancer screening, one participant noted:

> *Personally, I knew that screening for breast and cervical cancer is important, but I would always skip it because I was too busy. Through this program, I was convinced and learned that I should actively get the screenings.*

Similarly, another intervention participant shared how e-CHECC-uP helped improve health literacy which led to boosting her self-confidence:

> *I am always hesitant to go to the hospital because of the language barrier like the words I am not familiar with, but I gained some confidence through this program.*

Finally, intervention participants talked about how their knowledge about cancer screening procedures improved as a

XSL•FO
RenderX

result of their participation in the intervention program. One woman said:

> *It was great that I learned about the [cancer] screening procedures and processes through the listening components [of the web modules].*

The trial was feasible in terms of recruitment, retention, and intervention delivery. Specifically, we were able to meet the recruitment goal of 40 women with a response rate of 77%, after accounting for 12 women out of 52 eligible women who either declined (n=8) or were unreachable (n=4). We achieved a retention rate of 85% at 6 months, after accounting for 6 women who were lost to follow-up due to time restraints, participation burden, or loss of contact (Figure 1). Finally, all of our intervention components, including online modules and phone follow-up, were delivered as planned.

Discussion

Principal Findings

e-CHECC-uP was highly acceptable and feasible with promising outcome changes. We were able to achieve 85% retention at 6 months with overall high satisfaction ratings in this pilot sample. Given the target population being underserved with limited English proficiency, low health literacy, and limited access to care [6,8], high feasibility and acceptability observed in the pilot sample is encouraging and suggests possible avenues for promising intervention programs to promote breast and cervical cancer screening in Korean immigrant women who often suffer from late diagnoses [1].

Our successful recruitment, retention, and intervention delivery are a result of multiple factors such as authentic care, community partnership, self-directed web-based intervention with human support, and diverse yet standardized modes of communication with study participants [30]. We recruited participants and collected baseline data mainly through face-to-face interactions. As noted by Felsen et al [31], adopting in-person methods may likely have led to successful recruitment as it helps recruiters build rapport with potential study participants. Throughout the study period, we provided various but streamlined approaches for communication. At the earlier stage of the study, phone calls were used as the main method of contact and communication with participants. However, we quickly learned that adding text messaging and email yielded higher rates of response from study participants. Indeed, nearly every woman in the study sample had internet and mobile phone access (Table 1). Additionally, in this sample, 85% participants used texting daily [30], which made it easier to adopt text messaging as one of the main methods of communication. The high rate of internet and mobile phone use among Korean immigrant women with LEP demonstrates strong digital readiness and supports the feasibility of scaling e-CHECC-uP to broader immigrant populations with similar technology access profiles. Taking into account the patterns of technology access and use among the target sample would also facilitate the process of recruitment, retention, and data collection.

e-CHECC-uP pilot resulted in promising group differences across all cancer screening outcomes (range 34.6%-47.9%). While traditional face-to-face CHW interventions designed to promote same cancer screening outcomes have resulted in about 6% - 33% differences [11], the in-person version of this health literacy-focused CHW-led intervention achieved a larger group difference of 46% for mammography and 45% for Papanicolaou test screening [12]. The results indicate that the health literacy-focused approach is effective and perhaps better than the traditional cancer knowledge education. e-CHECC-uP produced comparable preliminary cancer screening outcome changes relative to the original in-person version [12]. This suggests digital CHW-led interventions may serve as scalable alternatives to in-person programs. Additionally, our finding highlights the importance of multilingual digital platforms for delivering health education. Health agencies and payers should invest in scalable, culturally competent eHealth infrastructure to reduce screening disparities. A sufficiently powered, full-scale clinical trial is warranted to rigorously investigate the efficacy of e-CHECC-uP in addressing breast and cancer screening outcomes.

We observed medium to large effect sizes for some of the key psychosocial outcomes including familiarity, social support, and quality of life, whereas self-efficacy showed smaller effects. Additionally, the effect size for comprehension attenuated between the 3- and 6-month follow-up assessments. The reasons for the smaller effect size in self-efficacy and the attenuation in comprehension over time are not entirely clear. One possible explanation may be an artifact of the pilot sample, which included a higher proportion of women with LEP in the intervention arm compared to the control arm (Table 1). The comprehension domain of health literacy required participants to associate each term on the list with a corresponding word or picture, assessing their understanding of the meaning [20]. The reduction in effect size over time may also reflect memory decline related to the learned terms, suggesting a need for reinforcement of health literacy education and an opportunity for scalability through various strategies such as digital refreshers or SMS nudges to sustain health literacy gains during the follow-up period. Existing research documents that people with low health literacy have poorer health outcomes due to its impact on self-efficacy, disease knowledge, or usage of health services [32,33]. For example, diabetic patients with higher health literacy had better understanding of their conditions and self-care behaviors [34,35]. Similarly, other studies support the positive relationship between health literacy and cancer screening, in that people with better health literacy are likely to have a better understanding of cancer screening and improve care-seeking behavior, thereby increasing the use of cancer screening services [36-38]. Postintervention qualitative interviews revealed that e-CHECC-uP helped intervention participants recognize the importance of cancer screening, improved health literacy and self-confidence, and enhanced their knowledge about screening procedures. Future research should investigate the mechanism of e-CHECC-uP for its impact on behavioral outcomes through mediation analyses.

Limitations

Our study has a few limitations to note. First, due to the nature of a pilot study, there is a small sample size and insufficient power. Additionally, the limited number of intervention

XSL•FO
RenderX

participants constrained our ability to systematically integrate qualitative themes with corresponding quantitative findings. A fully powered trial is warranted to test the efficacy of the e-CHECC-uP intervention program and to enable a more robust integration of quantitative outcomes with participants' qualitative feedback. In addition, the study sample was recruited from ethnic churches in the east coast region of the United States; hence, the study findings may not be generalizable to women who live elsewhere or who do not attend church. We also note that intervention participants were provided tablet devices on-site with technological assistance from the study team as a part of the web-based health literacy education. Such aspects of the study may also prevent findings from being generalizable to those who may face barriers related to use of and access to technology and the internet [17].

Conclusions

Online technology allows and engages experiences for learners at relatively low cost, given that facilities and instructors are not required. While prior studies have consistently documented inadequate cancer screening in women with LEP [1,2], few have focused on methods to actually address far-reaching, modifiable barriers such as low health literacy and to improve screening participation. Our use of eHealth to deliver health literacy education may represent a promising strategy for promoting engagement and integrating health literacy skills training aimed at educating and empowering women with LEP to improve their mammogram and Papanicolaou test screening.

Acknowledgments

This study was supported by a Johns Hopkins Provost Discovery award.

Data Availability

All data that support this study's findings will be made available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Checklist 1
CONSORT-EHEALTH checklist (V 1.6.1).
[PDF File, 7799 KB - humanfactors_v12i1e66092_app1.pdf]

References

1. Cancer Facts & Figures 2024. American Cancer Society.: American Cancer Society; 2024. URL: https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2024/2024-cancer-facts-and-figures-acs.pdf [accessed 2024-08-13]
2. Cancer Facts & Figures for Hispanic/Latino People 2021-2023. American Cancer Society.: American Cancer Society; 2021. URL: https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/cancer-facts-and-figures-for-hispanics-and-latinos/hispanic-latino-2021-2023-cancer-facts-and-figures.pdf [accessed 2024-08-13]
3. Heintzman J, Hatch B, Coronado G, et al. Role of race/ethnicity, language, and insurance in use of cervical cancer prevention services among low-income Hispanic women, 2009-2013. Prev Chronic Dis 2018 Feb 22;15:E25. [doi: 10.5888/pcd15.170267] [Medline: 29470167]
4. Lee RJ, Madan RA, Kim J, Posadas EM, Yu EY. Disparities in cancer care and the Asian American population. Oncologist 2021 Jun;26(6):453-460. [doi: 10.1002/onco.13748] [Medline: 33683795]
5. Heberer MA, Komenaka IK, Nodora JN, et al. Factors associated with cervical cancer screening in a safety net population. World J Clin Oncol 2016 Oct 10;7(5):406-413. [doi: 10.5306/wjco.v7.i5.406] [Medline: 27777883]
6. Kim K, Xue QL, Walton-Moss B, Nolan MT, Han HR. Decisional balance and self-efficacy mediate the association among provider advice, health literacy and cervical cancer screening. Eur J Oncol Nurs 2018 Feb;32:55-62. [doi: 10.1016/j.ejon.2017.12.001] [Medline: 29353633]
7. American Community Survey 5-Year Data. US Census Bureau. 2023 Dec 7. URL: https://www.census.gov/data/developers/data-sets/acs-5year.html [accessed 2024-08-13]
8. Esterline C, Batalova J. Korean immigrants in the United States. Migration Policy Institute. 2022. URL: https://www.migrationpolicy.org/article/korean-immigrants-united-states#english [accessed 2024-08-13]
9. Blanchard AK, Prost A, Houweling TAJ. Effects of community health worker interventions on socioeconomic inequities in maternal and newborn health in low-income and middle-income countries: a mixed-methods systematic review. BMJ Glob Health 2019;4(3):e001308. [doi: 10.1136/bmjgh-2018-001308] [Medline: 31275619]

XSL•FO
RenderX

10. Vasan A, Morgan JW, Mitra N, et al. Effects of a standardized community health worker intervention on hospitalization among disadvantaged patients with multiple chronic conditions: A pooled analysis of three clinical trials. Health Serv Res 2020 Oct;55 Suppl 2(Suppl 2):894-901. [doi: 10.1111/1475-6773.13321] [Medline: 32643163]
11. Kim K, Choi JS, Choi E, et al. Effects of community-based health worker interventions to improve chronic disease management and care among vulnerable populations: A systematic review. Am J Public Health 2016 Apr;106(4):e3-e28. [doi: 10.2105/AJPH.2015.302987] [Medline: 26890177]
12. Han HR, Song Y, Kim M, et al. Breast and cervical cancer screening literacy among Korean American women: A community health worker-led intervention. Am J Public Health 2017 Jan;107(1):159-165. [doi: 10.2105/AJPH.2016.303522] [Medline: 27854539]
13. Global observatory for eHealth. World Health Organization. 2024. URL: https://www.who.int/observatories/global-observatory-for-ehealth [accessed 2024-08-13]
14. Miller DP Jr, Weaver KE, Case LD, et al. Usability of a novel mobile health iPad App by vulnerable populations. JMIR Mhealth Uhealth 2017 Apr 11;5(4):e43. [doi: 10.2196/mhealth.7268] [Medline: 28400354]
15. Hamine S, Gerth-Guyette E, Faulx D, Green BB, Ginsburg AS. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. J Med Internet Res 2015 Feb 24;17(2):e52. [doi: 10.2196/jmir.3951] [Medline: 25803266]
16. Internet, broadband fact sheet. Pew Research Center. 2024 Jan 31. URL: https://www.pewresearch.org/internet/fact-sheet/internet-broadband [accessed 2025-08-20]
17. Chae S, Lee YJ, Han HR. Sources of health information, technology access, and use among non-English-speaking immigrant women: Descriptive correlational study. J Med Internet Res 2021 Oct 29;23(10):e29155. [doi: 10.2196/29155] [Medline: 34714249]
18. Breast cancer: Screening. US Preventive Services Task Force. 2024 Apr 30. URL: https://www.uspreventiveservicestaskforce.org/uspstf/recommendation/breast-cancer-screening#bcei-recommendation-title-area [accessed 2024-08-13]
19. Cervical cancer: Screening. US Preventive Services Task Force. 2018 Aug 21. URL: https://www.uspreventiveservicestaskforce.org/uspstf/recommendation/cervical-cancer-screening [accessed 2024-08-13]
20. Han HR, Huh B, Kim MT, Kim J, Nguyen T. Development and validation of the assessment of health literacy in breast and cervical cancer screening. J Health Commun 2014;19 Suppl 2(0 2):267-284. [doi: 10.1080/10810730.2014.936569] [Medline: 25315598]
21. Han HR, Kim K, Cudjoe J, Kim MT. familiarity, navigation, and comprehension: key dimensions of health literacy in Pap test use among Korean American women. J Health Commun 2019;24(6):585-591. [doi: 10.1080/10810730.2019.1607955] [Medline: 31046641]
22. Shin HS, Chi E, Han HR. Validity and Reliability of the Korean Version of Assessment of Health Literacy in Breast and Cervical Cancer Screening. J Korean Acad Nurs 2021 Dec;51(6):769-781. [doi: 10.4040/jkan.21108] [Medline: 35023864]
23. Sun CA, Chepkorir J, Jennifer Waligora Mendez K, Cudjoe J, Han HR. A descriptive analysis of cancer screening health literacy among Black women living with HIV in Baltimore, Maryland. Health Lit Res Pract 2022 Jul;6(3):e175-e181. [doi: 10.3928/24748307-20220616-01] [Medline: 35858186]
24. Moser A, Stuck AE, Silliman RA, Ganz PA, Clough-Gorr KM. The eight-item modified Medical Outcomes Study Social Support Survey: psychometric evaluation showed excellent performance. J Clin Epidemiol 2012 Oct;65(10):1107-1116. [doi: 10.1016/j.jclinepi.2012.04.007] [Medline: 22818947]
25. Hogenmiller JR, Atwood JR, Lindsey AM, Johnson DR, Hertzog M, Scott JC. Self-Efficacy Scale for Pap smear screening participation in sheltered women. Nurs Res 2007;56(6):369-377. [doi: 10.1097/01.NNR.0000299848.21935.8d]
26. Kim MH, Cho YS, Uhm WS, Kim S, Bae SC. Cross-cultural adaptation and validation of the Korean version of the EQ-5D in patients with rheumatic diseases. Qual Life Res 2005 Jun;14(5):1401-1406. [doi: 10.1007/s11136-004-5681-z] [Medline: 16047514]
27. Kai JY, Xu Y, Li DL, Zhou M, Wang P, Pan CW. Impact of major age-related eye disorders on health-related quality of life assessed by EQ-5D: a systematic review and meta-analysis. Graefes Arch Clin Exp Ophthalmol 2023 Sep;261(9):2455-2463. [doi: 10.1007/s00417-023-06034-z] [Medline: 36917317]
28. Lichstein KL, Riedel BW, Grieve R. Fair tests of clinical trials: A treatment implementation model. Adv Behav Res Therapy 1994 Jan;16(1):1-29. [doi: 10.1016/0146-6402(94)90001-9]
29. Braun V, Clarke V. Is thematic analysis used well in health psychology? A critical review of published research, with recommendations for quality practice and reporting. Health Psychol Rev 2023 Dec;17(4):695-718. [doi: 10.1080/17437199.2022.2161594] [Medline: 36656762]
30. Hwang DA, Lee A, Song JM, Han HR. Recruitment and retention strategies among racial and ethnic minorities in web-based intervention trials: Retrospective qualitative analysis. J Med Internet Res 2021 Jul 12;23(7):e23959. [doi: 10.2196/23959] [Medline: 34255658]
31. Felsen CB, Shaw EK, Ferrante JM, Lacroix LJ, Crabtree BF. Strategies for in-person recruitment: lessons learned from a New Jersey primary care research network (NJPCRN) study. J Am Board Fam Med 2010;23(4):523-533. [doi: 10.3122/jabfm.2010.04.090096] [Medline: 20616295]

XSL•FO
RenderX

32. Dewalt DA, Berkman ND, Sheridan S, Lohr KN, Pignone MP. Literacy and health outcomes: a systematic review of the literature. J Gen Intern Med 2004 Dec;19(12):1228-1239. [doi: 10.1111/j.1525-1497.2004.40153.x] [Medline: 15610334]
33. Wayment A, Wong C, Byers S, Eley R, Boyde M, Ostini R. Beyond access block: Understanding the role of health literacy and self-efficacy in low-acuity emergency department patients. Ochsner J 2020;20(2):161-169. [doi: 10.31486/toj.19.0047] [Medline: 32612470]
34. Lee YJ, Shin SJ, Wang RH, Lin KD, Lee YL, Wang YH. Pathways of empowerment perceptions, health literacy, self-efficacy, and self-care behaviors to glycemic control in patients with type 2 diabetes mellitus. Patient Educ Couns 2016 Feb;99(2):287-294. [doi: 10.1016/j.pec.2015.08.021] [Medline: 26341940]
35. Bohanny W, Wu SFV, Liu CY, Yeh SH, Tsay SL, Wang TJ. Health literacy, self-efficacy, and self-care behaviors in patients with type 2 diabetes mellitus. J Am Assoc Nurse Pract 2013 Sep;25(9):495-502. [doi: 10.1111/1745-7599.12017] [Medline: 24170654]
36. Lindau ST, Tomori C, Lyons T, Langseth L, Bennett CL, Garcia P. The association of health literacy with cervical cancer prevention knowledge and health behaviors in a multiethnic cohort of women. Am J Obstet Gynecol 2002 May;186(5):938-943. [doi: 10.1067/mob.2002.122091] [Medline: 12015518]
37. Duong TV, Aringazina A, Baisunova G, et al. Measuring health literacy in Asia: Validation of the HLS-EU-Q47 survey tool in six Asian countries. J Epidemiol 2017 Feb;27(2):80-86. [doi: 10.1016/j.je.2016.09.005] [Medline: 28142016]
38. Mazor KM, Williams AE, Roblin DW, et al. Health literacy and pap testing in insured women. J Cancer Educ 2014 Dec;29(4):698-701. [doi: 10.1007/s13187-014-0629-7] [Medline: 24633725]

Abbreviations

AHL-C: Assessment of Health Literacy in Cancer Screening
CHECC-uP: Community-based Health litEracy-focused intervention for breast and cervical Cancer Control
CHW: community health worker
CONSORT: Consolidated Standards of Reporting Trials
e-CHECC-uP: web-version of Community-based Health litEracy-focused intervention for breast and cervical Cancer Control
LEP: limited English proficiency
MOS-SS: Medical Outcomes Study Social Support Survey

Edited by P Santana-Mancilla; submitted 03.09.24; peer-reviewed by L Raymond Guo, V Soto; revised version received 26.06.25; accepted 27.06.25; published 25.08.25.

Please cite as:
Han HR, Lee YJ, Min D, Chepkorir J, Hwang DA, Chae S
Web-Based Delivery of an Effective Church-Based Intervention Program to Promote Cancer Screening (Community-based Health litEracy-focused intervention for breast and cervical Cancer Control) Among Korean Immigrant Women in the United States: Randomized Controlled Trial
JMIR Hum Factors 2025;12:e66092
URL: https://humanfactors.jmir.org/2025/1/e66092
doi:10.2196/66092

© Hae-Ra Han, Yoon-Jae Lee, Deborah Min, Joyline Chepkorir, DaSol Amy Hwang, Steve Chae. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 25.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Following Up Patients With Chronic Pain Using a Mobile App With a Support Center: Unicenter Prospective Study

Marta Antonia Gómez-González[1*], MD; Nicolas Cordero Tous[1*], MD, PhD; Javier De la Cruz Sabido[1*], MD; Carlos Sánchez Corral[1*], MD; Beatriz Lechuga Carrasco[1*], RNC; Marta López-Vicente[1*], MD; Gonzalo Olivares Granados[1,2*], MD, PhD

[1]Department of Neurosurgery, Hospital Universitario Virgen de las Nieves, Av. Juan Pablo II s/n, Granada, Spain

[2]Department of Human Anatomy, University of Granada, Granada, Spain

[*]all authors contributed equally

Corresponding Author:
Nicolas Cordero Tous, MD, PhD
Department of Neurosurgery, Hospital Universitario Virgen de las Nieves, Av. Juan Pablo II s/n, Granada, Spain

Abstract

Background: Chronic pain is among the most common conditions worldwide and requires a multidisciplinary treatment approach. Spinal cord stimulation is a possible treatment option for pain management; however, patients undergoing this intervention require close follow-up, which is not always feasible. eHealth apps offer opportunities for improved patient follow-up, although adherence to these apps tends to decrease over time, with rates dropping to approximately 60%. To improve adherence to remote follow-up, we developed a remote follow-up system consisting of a mobile app for patients, a website for health care professionals, and a remote support center.

Objective: Our objective was to evaluate patient adherence to remote follow-up using a system that includes a mobile app and a remote support center.

Methods: After review of the literature and approval of the design of the follow-up system by a multidisciplinary committee, a team of experts developed a system based on a mobile app, a website for health care professionals, and a remote support center. The system was developed in collaboration with health care professionals and uses validated scales to capture patients' clinical data at each stage of treatment (ie, pretreatment phase, trial phase, and implantation phase). Data were collected prospectively between January 2020 to August 2023, including the number of total surveys sent, surveys completed, SMS text message reminders sent, and reminder calls made.

Results: A total of 64 patients were included (n=40 women, 62.5%) in the study. By the end of the study, 19 (29.7%) patients remained in the pretreatment phase, 8 (12.5%) patients had completed the trial phase, and 37 (57.8%) reached the implantation phase. The mean follow-up period was 15.30 (SD 9.43) months. A total of 1574 surveys were sent, along with 488 SMS text message reminders and 53 reminder calls. The mean adherence rate decreased from 94.53% (SD 20.63%) during the pretreatment phase to 65.68% (SD 23.49%) in the implantation phase, with an overall mean adherence rate of 87.37% (SD 15.37%) for the app. ANOVA showed that adherence was significantly higher in the earlier phases of treatment (P<.001).

Conclusions: Our remote follow-up system, supported by a remote support center improves adherence to follow-up in later phases of treatment, although adherence decreased over time. Further studies are needed to investigate the relationship between adherence to the app and pain management.

(JMIR Hum Factors 2025;12:e60160) doi:10.2196/60160

KEYWORDS

pain management; mobile health; mHealth; eHealth; chronic pain; support center; mobile phone app; survey; follow-up; pain control; prospective study

Introduction

Chronic pain is one of the most common conditions globally and is associated with reduced quality of life, increased medical expenses, and significant economic costs [1]. Its prevalence ranges between 2% - 40% [1-3], and annual health care costs due to pain amount to US $300 billion in the United States [1] and €12 billion (US $12.3 billion) in Norway [2].

The treatment of chronic pain requires a multidisciplinary approach, with spinal cord stimulation being a potential treatment option. This technique is safe and effective [4] and has also shown benefits in management of other chronic pain conditions, such as complex regional pain syndromes, low back

XSL•FO
RenderX

pain, and other forms of pain. However, ensuring consistent follow-up for these patients is challenging, as they require close monitoring, which is not always feasible.

New technologies have opened up a lot of possibilities to address this challenge. The World Health Organization defines eHealth as the cost-effective and secure use of information and communication technologies to support health and health-related areas, including health services, health monitoring, health literature, health education, and knowledge and research [5].

There are numerous tools designed to facilitate medical care, including mobile apps, websites, and other platforms. While there is limited evidence to support their use, preliminary results are promising. For example, mobile apps used alongside conventional treatments have shown better results in the management of chronic pain than conventional interventions alone [6,7].

eHealth apps also offer an opportunity for closer patient follow-up, although adherence to treatment tends to decline over time, dropping to approximately 60%, as previous studies have shown [8,9]. To address this challenge, we have developed a system that includes a mobile app for patients, a website for health care professionals, and a remote support center.

The aim of this study is to assess patient adherence to follow-up care using a mobile app supported by a remote support center.

Methods

Overview

In 2019, researchers at the Department of Pain Surgery at Hospital Virgen de las Nieves in Granada, Spain, initiated approaches to improve the remote monitoring of patients implanted with epidural spinal cord stimulators for chronic pain management. The aim was to reduce the number of face-to-face consultations or hospital visits while providing additional support to these patients. Therefore, we decided to develop an app complemented by a support center to improve patient monitoring as much as possible. After reviewing the relevant literature, we adapted the tool kit published by Marvel et al [10] to suit our needs, following the four key steps outlined in Table 1. As we already had a multidisciplinary group, we sought the agreement of existing team members and assembled a group of experts. Instead of accelerating the project, we adapted our current workflows and protocols to the new working framework. We also developed the follow-up system after adapting the protocols already in place and enrolled patients only after the system was finalized.

Table . Project phases.

Johns Hopkins tool steps[a]	Project workflow (current study)
Define the problem and the digital tool	Define the problem and the digital tool
Creation of a multidisciplinary group	Creation of an expert team, approval by a multidisciplinary group
Seeking opportunities to accelerate the project	Adaptation of the existing clinical and educational protocol into the framework of the remote follow-up system
Involving professionals	Designing the follow-up system, considering the needs of the involved professionals, and subsequent enrollment of patients
Consulting different partners	Performing quality assessment
Conducting a clinical validation	Conducting a clinical validation

[a]Marvel et al [10].

The concept of a remote follow-up system was presented to a multidisciplinary committee composed of anesthesiologists, neurologists, neurosurgeons, neurophysiologists, and rehabilitation specialists. This committee meets monthly to assess patients with chronic pain who may benefit from interventional therapies such as spinal cord stimulation when conventional treatments have failed. The target population was identified as patients who could benefit from spinal cord stimulation and had not yet started the trial phase, to enable data collection before and after implantation.

The inclusion criteria for patients eligible for spinal cord stimulation system implantation at our center are shown in Textbox 1.

Textbox 1. Inclusion criteria.

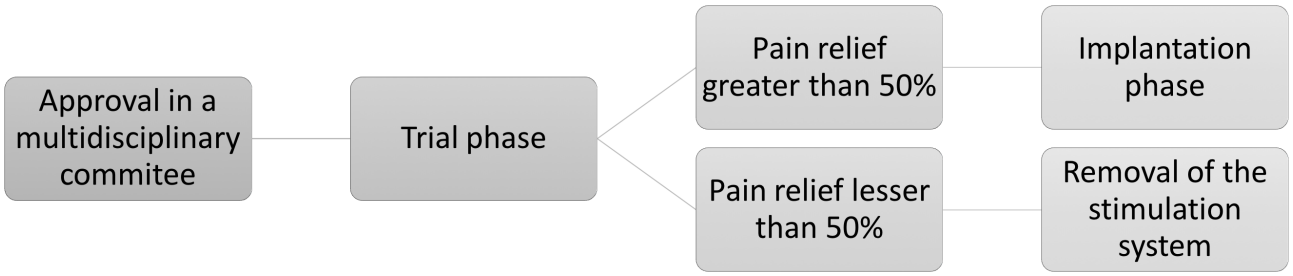
- Chronic pain in a specific area of the body (patients experiencing diffuse pain were not accepted)
- Failure of conservative treatments, including rehabilitation and local infiltration
- A favorable psychological evaluation confirming that the patient can wear a spinal cord stimulation system
- No contraindications to surgery or allergies to system components
- No history or current substance use, especially opioids

Once the committee had approved the implantation of a spinal cord stimulator, the procedure was performed in 2 phases, as shown in Figure 1. In the trial phase, one or more electrodes were implanted in the epidural space and connected to an external stimulator. The efficacy of the system was then assessed over a 4-week period using the verbal numerical rating score

XSL•FO
RenderX

(VNRS). The intervention was considered effective if the patient's reported subjective perception of pain relief was at least 50%. If the trial phase was successful, it was followed by the implantation phase, in which a permanent stimulator was implanted in the adipose tissue of the abdomen or lower back. If the 50% threshold was not reached, the system was removed completely.

Figure 1. Algorithm used to evaluate the implantation of the spinal cord stimulation system.



Ethical Considerations

Ethics approval was obtained from the Comité de Investigación Provincial de Granada (study: NC-D-01; ethics committee reference: SICEIA-2020 - 000438), and informed consent for the procedure and follow-up system was obtained from all included patients. To maintain privacy, each patient was assigned a code with a number so that only health care professionals could access the personal data. The platform meets all the standards of Regulation (EU) 2016/679 of the European Parliament and of the Council of April 27, 2016, and has been approved by local and state authorities. This ensures that only health care professionals can access all the data collected and that data security is contractually guaranteed. The study participants have not been compensated in any way.

Design of the Follow-Up System

The remote follow-up system consists of 3 main components: a mobile app for patients, a website for health care professionals, and a remote follow-up center. The mobile app was designed to ensure close monitoring during all phases of treatment, using feedback from all health care professionals involved. The app was developed in collaboration with Clinical Care Connect (Persei) and customized by our team to meet our specific needs. It offers various features including educational content such as brochures and videos, questionnaires with validated scales, pop-up notifications, and a form to contact the remote support center. The validated scales used include Doleur Neuropathique 4 (DN4) to assess neuropathic pain, Oswestry Disability Index/Neck Disability Index (ODI/NDI) to assess back pain, VNRS to assess pain, and 36-item Short Form Survey (SF-36) to assess quality of life. In addition, patient-reported experience measures were used to assess experience and satisfaction with the app. Since pain is a subjective perception, the use of validated scales can help to measure it and compare it between patients.

The website designed by Persei (Vivarium S.L., Spain) can be accessed from anywhere with the necessary authorizations. It also provides alerts using different colors depending on the urgency of the detected problem: red for infection, yellow for insufficient pain relief, and blue for low battery. These alerts are monitored by the remote support center, which contacts the specialist by making a telephone call if necessary, such as, in the event of an infection. For nonemergency issues, such as insufficient pain relief, the support center contacts the patient to analyze the cause, and a priority appointment for treatment is scheduled. This remote support center was set up to increase the efficiency of treatment and improve pain management. The aim was to maximize treatment adherence and ensure the sustainability of the remote follow-up system. Previously, follow-up was carried out 1 month after implantation, 1 year after implantation, and subsequently, as required.

Data Collection and Analysis

We collected data from all patients who were approved for spinal cord stimulation between January 2020 and August 2023 and enrolled in the remote monitoring program. If the committee approved the patient for spinal cord stimulation, informed consent for the remote follow-up system was obtained in addition to consent for implantation. Data were collected automatically and prospectively through surveys completed via the app. The questionnaires were distributed according to the protocol shown in Textbox 2.

XSL•FO
RenderX

Textbox 2. Protocol used to send validated scales.

Validated scales used in the protocol used by the Virgen de las Nieves Hospital

- Pretreatment phase: Doleur Neuropathique 4 (DN4), verbal numerical rating score (VNRS), Oswestry Disability Index/Neck Disability Index (ODI/NDI), 36-item Short Form Survey (SF-36)
- Trial phase: Daily VNRS and patient-reported experience measures (PREMs) at the end of the trial phase
- Implantation phase: DN4, VNRS, ODI, and SF-36 after 1 and every 3 months; the PREMs survey was also distributed after 12 months

The support center sends a notification each time a survey needs to be completed, for a maximum of 3 times. If the questionnaires were not completed, a notification was sent to the hospital so that a specialized caregiver can contact the patient. One of the biggest challenges was ensuring privacy. To achieve this, each patient was coded with a number so that only health care professionals had access to the personal data and that data security was contractually guaranteed.

The data were collected automatically when a survey was answered via the app and stored in a private cloud. This allowed us to minimize potential errors. In addition, the support center guaranteed that the data collected via the app belonged to our hospital's patients and was only accessed and analyzed when a specialist at our center requested it. Demographic information was collected for data analysis. Adherence was measured by calculating the percentage of completed surveys relative to the total number of surveys sent. We analyzed adherence rate during each treatment phase and for the overall study, and compared adherence by demographic variables using ANOVA, adjusting for gender, age, diagnosis, and phase. Age was categorized into 2 groups, with 40 years as the cutoff point, since previous studies have found significant differences in prevalence rates for chronic pain and quality of life [11]. Regarding completeness, we only considered a survey as answered if it was completed. If data were missing, a notification was sent to the patient to complete the survey.

Results

Between January 2020 and August 2023, a total of 64 patients were enrolled in the study, comprising 24 (37.5%) men and 40 (62.5%) women. While identifying the cause of pain, 32 (50%) patients had complex regional pain syndromes, 18 (28%) had failed back syndrome, and 14 (22%) had other diagnoses such as phantom limb syndrome, adhesive arachnoiditis, coccydynia, and axonotmesis. Before the onset of pain, 19 (28%) patients had no fixed occupation, while 10 (18%) were involved in nonphysical and 35 (55%) in physical occupations such as farming or construction (Table 2).

Table . Summary of epidemiological data.

Variables	Patients (N=64), n (%)
Gender	
Men	24 (38)
Women	40 (63)
Age interval	
<40 years	9 (14)
≥40 years	55 (86)
Diagnoses	
CRPS[a]	32 (50)
FBS[b]	18 (22)
Others	14 (18)
Occupation	
Physical occupation	35 (55)
Nonphysical occupation	19 (18)
No fixed occupation	10 (28)
Phase	
Pretreatment	19 (30)
Trial	8 (13)
Implantation	37 (58)

[a]CRPS: chronic regional pain syndrome.

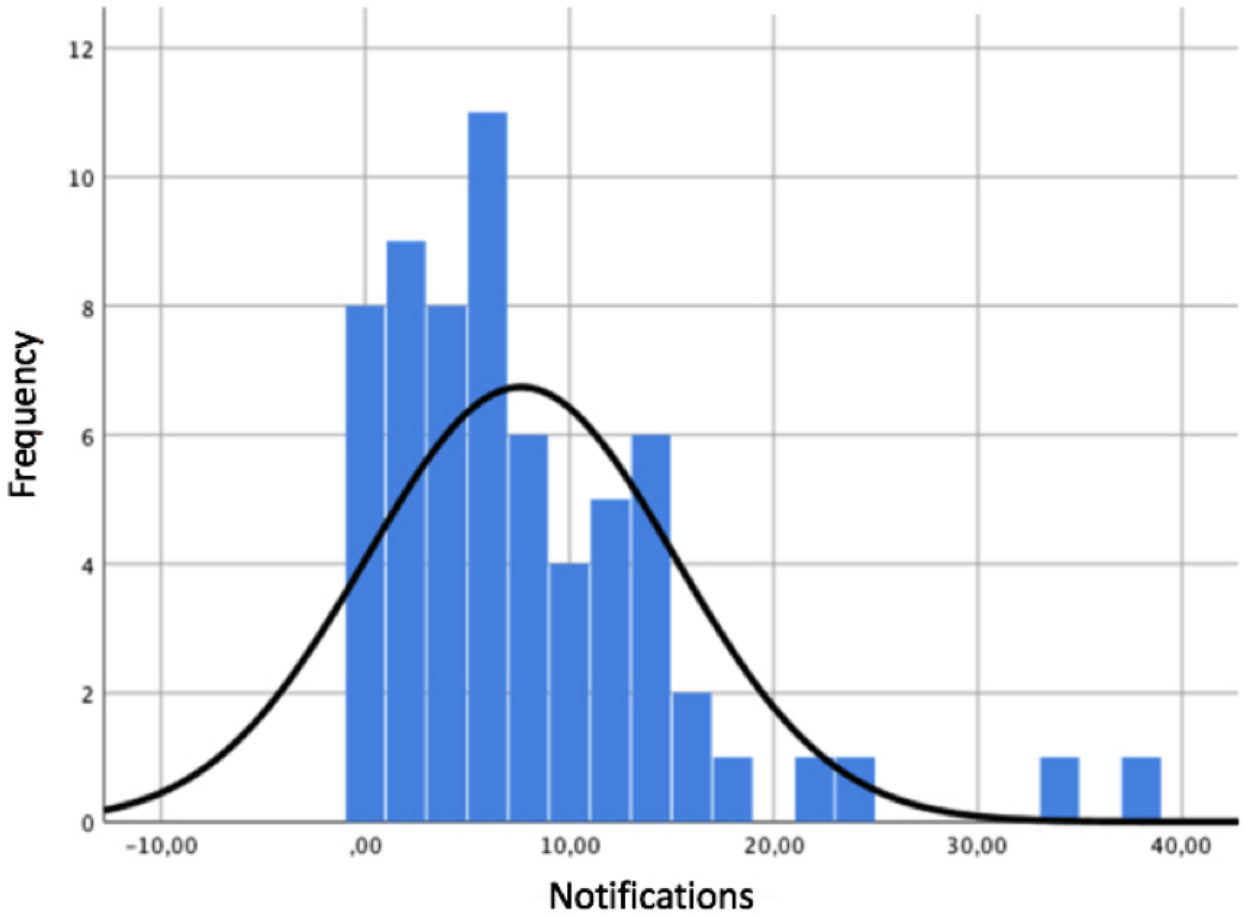
[b]FBS: failed back syndrome.

XSL•FO
RenderX

At the end of the study, 19 (30%) patients were in the pretreatment group, awaiting the start of the trial phase. A total of 8 (13%) patients reached the trial phase but were not transferred to the implantation phase due to insufficient improvement in pain management. Further, a total of 37 (58%) patients advanced to the implantation phase (Table 2).

A total of 1586 surveys were distributed during the study. SMS text message notifications had to be sent to the patients 488 times, which corresponded to a mean of 7.62 (SD 7.57) SMS text messages per patient. However, 8 (13%) patients did not require SMS text messages notifications, while the remaining (n=56, 88%) patients required at least one notification (Figure 2). It was observed that several patients required more than 30 notifications. In addition, a total of 53 phone calls were made to remind patients to complete the surveys; however, 43 (67%) patients did not require any reminder calls.

Figure 2. Histogram with a normality curve showing the number of patients in groups (frequency) based on the number of required notifications (notifications).



The mean follow-up time was 15.38 (SD 9.43) months. A total of 12 (19%) patients were lost to follow-up, and 3 (4%) patients discontinued using the app (ie, 1 during the pretreatment phase and 2 patients during the implantation phase). For patients who discontinued using the app, any surveys not answered by them were considered failures.

Adherence was analyzed for each phase and overall, as shown in Figure 3, which represents the total number of surveys sent and answered in each phase, as well as the adherence rate in each phase. During the pretreatment phase, 256 surveys were sent, of which 242 were completed, corresponding to a mean adherence rate of 94.53% (SD 20.63%). During the trial phase, 645 surveys were sent, of which 499 were answered, corresponding to a mean adherence rate of 77.36% (SD 37.01%). During the implantation phase, 624 surveys were sent, of which 422 were answered, corresponding to a mean adherence of 65.68% (SD 23.49%). A total of 61 satisfaction surveys were sent (excluding 3 patients who discontinued using the app), of which 55 were answered, corresponding to a mean adherence rate of 92.11% (SD 18.47%). The overall mean adherence rate for using the app was 87.37% (SD 15.37%) across all phases (Table 3).

XSL•FO
RenderX

Figure 3. Surveys and adherence rate.

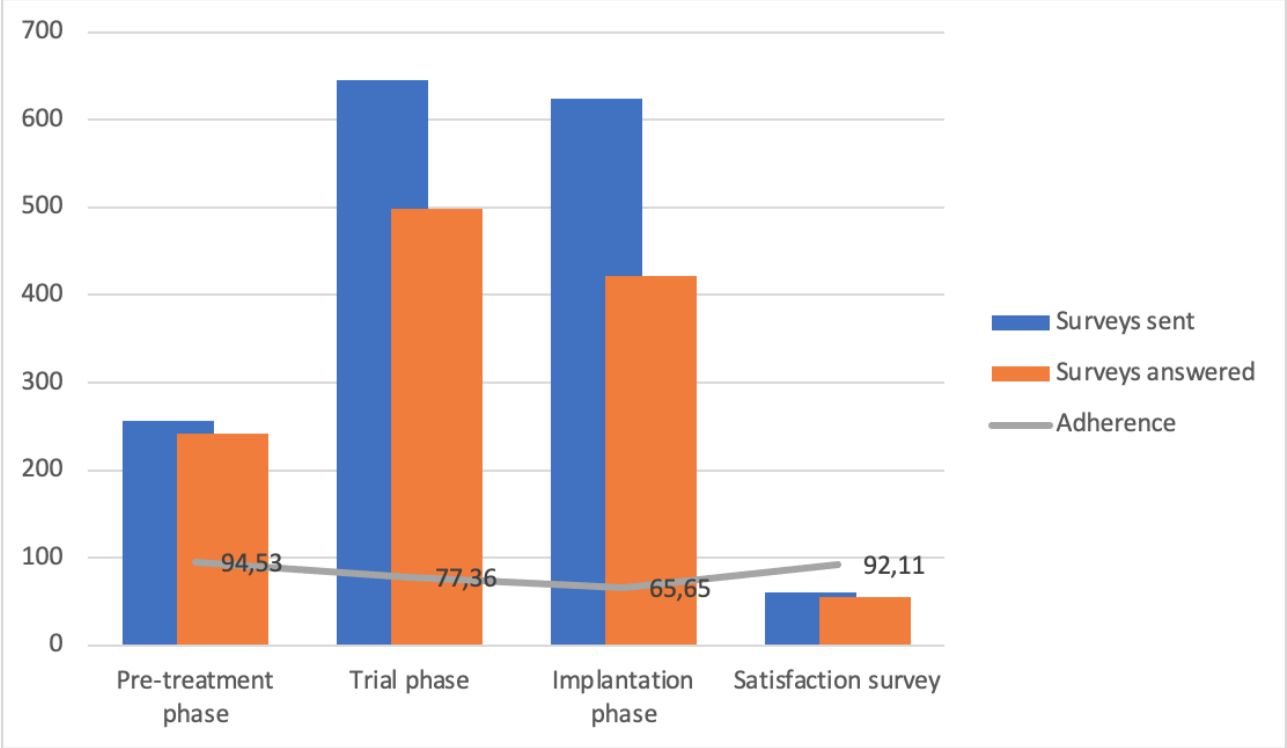


Table . Summary of surveys sent and answered and adherence in each phase.

Phase	Surveys sent, n[a]	Surveys answered, n[a]	Adherence rate[b] (%), mean (SD[c])
Pretreatment phase	256	242	94.53 (20.63)
Trial phase	645	499	77.36 (37.01)
Implantation phase	624	422	65.68 (23.49)
Satisfaction survey	61	55	92.11 (18.47)
Total adherence	1586	1218	87.37 (15.37)

[a]Number of surveys during each phase.

[b]Adherence rate was defined as the total surveys answered x 100/total surveys sent.

[c]The SD represents the variability of adherence rates among individual patients.

ANOVA revealed that adherence rates were significantly higher in the earlier phases of treatment (P<.001). There was also a trend indicating that women were more likely to complete satisfaction surveys (P=.17) and surveys in general (P=.64). However, women also required more SMS text message reminders (P=.70) and phone calls (P=.71). No significant differences in treatment adherence were found when comparing younger patients versus patients older than 40 years or when comparing adherence to treatment based on the cause of pain or type of occupation (Table 4).

XSL•FO
RenderX

Table . Summary of the ANOVA results.

Variables	Pretreatment phase adherence (%; n=64), mean (SD)	*P* value	Trial phase adherence (%; n=43), mean (SD)	*P* value	Implantation phase adherence (%; n=38), mean (SD)	*P* value	Satisfaction survey adherence (%; n=38), mean (SD)	*P* value	Global adherence (%), mean (SD)	*P* value
Gender		.81		.49		.18		.40		.80
Total	94.5 (20.6)		66.7 (23.5)		65.7 (23.5)		92.1 (18.5)		87.4 (15.4)	
Men	93.7 (22.4)		58.2 (31.3)		58.2 (31.3)		95.8 (14.4)		86.8 (19.1)	
Women	95.0 (19.8)		69.1 (18.6)		69.1 (18.6)		90.4 (20.1)		87.7 (12.9)	
Age (years)		.38		.29		.74		.90		.30
Total	94.5 (20.6)		77.36 (37.0)		65.7 (23.5)		92.1 (18.5)		83.4 (15.4)	
≤40	88.9 (33.3)		62.9 (42.2)		68.4 (19.9)		92.9 (18.9)		82.5 (14.9)	
>40	95.5 (18.1)		80.2 (34.6)		65.1 (25.6)		91.3 (18.7)		88.2 (15.4)	
Diagnoses		.42		.63		.57		.83		.35
Total	94.5 (10.6)		77.4 (37.0)		65.7 (23.5)		92.1 (18.5)		87.4 (15.6)	
CRPS[a]	100 (0)		72.8 (42.2)		63.1 (29.9)		92.8 (17.9)		85.7 (16.9)	
FBS[b]	94.5 (21.8)		81.7 (30.2)		73.5 (12.8)		93.8 (17.7)		92.7 (9.7)	
Others	90.3 (25.9)		85.3 (28.3)		64.8 (9.7)		88.9 (22.0)		86.2 (15.9)	
Phase		.61		.51		.60		.67		.001
Total	94.5 (20.6)		—[c]		—		—		87.4 (15.4)	
Pretreatment	96.1 (12.5)		77.8 (21.1)		77.8 (0)		100 (0)		96.1 (12.5)	
Trial	100 (0)		65.4 (4)		65.4 (23.8)		91.9 (18.7)		94.3 (9.2)	
Implantation	92.6 (25.6)		77.4 (37.0)		65.7 (23.5)		92.1 (18.5)		81.4 (15.2)	

[a]CRPS: complex regional pain syndromes.

[b]FBS: fail back syndrome.

[c]Not applicable.

Discussion

Comparison With Previous Research

In recent years, there has been increasing interest in the use of pain control apps for various purposes, such as postsurgical follow-up, pain self-management, and chronic pain management [6-10,12-14]. However, many of these tools have not been as effective as expected, often due to inadequate planning and a lack of preliminary studies to anticipate necessary roles [15]. However, some studies have demonstrated that these apps can outperform traditional follow-up methods [6,14]. Our study introduces a remote support center, a feature that we believe is not yet widely adopted internationally. The aim of this study is to determine whether the inclusion of a support center improves adherence compared to other systems.

An analysis of the medical apps across various app stores shows that approximately 86% of these apps were developed without the involvement of medical professionals [16]. Lalloo et al [17] examined 1019 apps developed for postoperative pain management or education and found that only 10 apps met the established criteria. When evaluating apps developed with the involvement of health care professionals, only 5 (0.49%) of them were deemed suitable. Similarly, Bhattarai et al [16] examined 373 apps focused on arthritis pain self-management with only 4 apps meeting the Stanford criteria for pain self-management. Portelli and Eldred [18] evaluated 195 apps for pain management against the guidelines for cognitive behavioral therapy. Of these, only 6 (3%) of these apps met the standards, leading to the conclusion that neither health care professionals nor patients were involved in their development. Despite these limitations, a recent meta-analysis of 4767 patients in 22 randomized trials found that these apps offer a small but significant improvement in long-term pain management [19].

For patients undergoing spinal cord stimulation, close follow-up is required to adjust the type and parameters of stimulation, which usually necessitates face-to-face visits. This increases the time and cost burdens for patients and caregivers, especially for those who live far from specialized centers [20]. Our remote follow-up system was developed to overcome these challenges and ensure optimal patient monitoring. The mobile app in combination with the remote support center reduces the need

XSL•FO
RenderX

for in-person examinations while ensuring a high standard of care. The app's alert system helps to detect problems at an early stage, such as infections or low battery levels so that immediate action can be taken.

In addition, the use of validated scales during an in-person consultation is challenging due to the limited time available. Consequently, this system has allowed a better understanding of the patient's condition. Furthermore, some patients have found that expressing themselves via the app helps them [20] in experiencing privacy and without any time pressure.

Adherence to medical app use is crucial as studies show that 75% of users stop using an app within 48 hours of downloading, and 15% delete it after initial use [21]. A review by Wikström et al [9] found that none of the studies focused on improving adherence or motivating users to continue using medical apps. The easiest way to increase motivation is through notifications, which is why our study focused on the role of the remote support center. We consider this as a key factor in the success of our study. However, numerous motivational elements in medical apps have been analyzed in previous studies, as summarized in Table 5.

Table . Motivational elements in eHealth apps.

Studies	Motivational elements								
	Tablet lent to the patient	Chat	Involvement of relatives	Educational texts, pictures, and/or videos	Follow-up of auto evaluation and graphic results	Personalized follow-up	Sharing media elements	Notifications and reminders	Alerts
Alam et al [20]		✓[a]		✓				✓	
Pecorelli et al [22]	✓			✓	✓			✓	
Perdoncini et al [23]		✓		✓					
Shah et al [24]		✓		✓					
Davidovitch et al [25]		✓		✓			✓		
Felbaum et al [26]		✓		✓			✓	✓	
Glauser et al [27]		✓		✓	✓		✓	✓	
Gustavell et al [28]		✓		✓	✓			✓	✓
Hou et al [29]		✓		✓		✓		✓	✓
Pickens et al [30]		✓	✓	✓				✓	
Timmers et al [31]				✓	✓		✓	✓	
van der Meij et al [32]				✓		✓			
Mundi et al [33]				✓				✓	
This study		✓	✓	✓	✓	✓	✓	✓	✓

[a]Element analyzed in the corresponding study.

Our remote follow-up system was developed based on relevant literature to maximize adherence and achieved an overall mean adherence rate of 87.37% (SD 15.37%). Comparison of this rate with previous studies is difficult, as most of these focused on adherence to a treatment or intervention rather than the use of an app. For example, Gomis-Pastor et al [34] showed that the use of an app significantly improved adherence to treatment and medication regimen, which led to significantly better symptom control, as found in a 2022 review [35]. However, Wikström et al [9] have shown that adherence to app use decreases over time, potentially due to improvement in health status or loss of interest [29], among other factors [26,29,32]. Our study aligns with these findings and shows a decrease in adherence from 94.53% in the pretreatment phase to 77.36% in the testing phase and 65.58% in the implantation phase [36-38].

XSL•FO
RenderX

Limitations

The main limitation of our study is the specificity of the target population, as spinal cord stimulation is a specialized treatment for patients experiencing chronic pain. These patients often display a significant emotional component that can influence the outcomes of the interventions, posing challenges during designing this study and its remote follow-up system. In addition, these systems rely on new technologies that can be challenging for older patients, potentially limiting their ability to fully benefit from them. By using self-reported data, we also acknowledged that omissions or misunderstandings in the surveys may be introduced, which we attempted to minimize using validated scales. Finally, setting up a system similar to this study requires a significant financial investment.

Conclusions

Digital systems are a part of medical care, and it is important that health care professionals are involved in the development of tools to ensure the achievement of desired standards. We have developed a remote monitoring system for patients undergoing spinal cord stimulation, based on scientific evidence and supported by a remote support center that improves treatment adherence. This system also provides us with detailed information about each patient. However, there is a tendency for patients to abandon the app usage over time, which could be related to long-term pain control. Further studies are needed to investigate the relationship between adherence and pain control.

Conflicts of Interest

We have received financial support from Medtronic IHS CareConnect Pain to cover publication costs for this article. However, we declare no conflicts of interest as we implant spinal cord stimulation systems from different brands and treat all equally. In addition, none of the authors work for Medtronic IHS CareConnect Pain or receive economic compensation from Medtronic IHS CareConnect Pain.

References

1. Gaskin DJ, Richard P. The economic costs of pain in the United States. J Pain 2012 Aug;13(8):715-724. [doi: 10.1016/j.jpain.2012.03.009] [Medline: 22607834]
2. Stubhaug A, Hansen JL, Hallberg S, Gustavsson A, Eggen AE, Nielsen CS. The costs of chronic pain–Long-term estimates. Eur J Pain 2024 Jul;28(6):960-977. [doi: 10.1002/ejp.2234] [Medline: 38214661]
3. Yong RJ, Mullins PM, Bhattacharyya N. Prevalence of chronic pain among adults in the United States. Pain 2022 Feb 1;163(2):e328-e332. [doi: 10.1097/j.pain.0000000000002291] [Medline: 33990113]
4. Kapural L, Peterson E, Provenzano DA, Staats P. Clinical evidence for spinal cord stimulation for failed back surgery syndrome (FBSS): systematic review. Spine (Phila Pa 1976) 2017 Jul 15;42 Suppl 14:S61-S66. [doi: 10.1097/BRS.0000000000002213] [Medline: 28441313]
5. eHealth. World Health Organization - Regional Office for the Eastern Mediterranean. URL: https://www.emro.who.int/health-topics/ehealth/ [accessed 2024-04-03]
6. Toelle TR, Utpadel-Fischler DA, Haas KK, Priebe JA. App-based multidisciplinary back pain treatment versus combined physiotherapy plus online education: a randomized controlled trial. NPJ Digit Med 2019;2:34. [doi: 10.1038/s41746-019-0109-x] [Medline: 31304380]
7. Rughani G, Nilsen TIL, Wood K, et al. The selfBACK artificial intelligence-based smartphone app can improve low back pain outcome even in patients with high levels of depression or stress. Eur J Pain 2023 May;27(5):568-579. [doi: 10.1002/ejp.2080] [Medline: 36680381]
8. Kerckhove N, Delage N, Cambier S, et al. eDOL mHealth app and web platform for self-monitoring and medical follow-up of patients with chronic pain: observational feasibility study. JMIR Form Res 2022 Mar 2;6(3):e30052. [doi: 10.2196/30052] [Medline: 35234654]
9. Wikström L, Schildmeijer K, Nylander EM, Eriksson K. Patients' and providers' perspectives on e-health applications designed for self-care in association with surgery - a scoping review. BMC Health Serv Res 2022 Mar 23;22(1):386. [doi: 10.1186/s12913-022-07718-8] [Medline: 35321707]
10. Marvel FA, Wang J, Martin SS. Digital health innovation: a toolkit to navigate from concept to clinical testing. JMIR Cardio 2018 Jan 18;2(1):e2. [doi: 10.2196/cardio.7586] [Medline: 31758761]
11. Rustøen T, Wahl AK, Hanestad BR, Lerdal A, Paul S, Miaskowski C. Age and the experience of chronic pain: differences in health and quality of life among younger, middle-aged, and older adults. Clin J Pain 2005;21(6):513-523. [doi: 10.1097/01.ajp.0000146217.31780.ef] [Medline: 16215337]
12. Jaensson M, Dahlberg K, Eriksson M, Grönlund Å, Nilsson U. The development of the recovery assessments by phone points (RAPP): a mobile phone app for postoperative recovery monitoring and assessment. JMIR Mhealth Uhealth 2015 Sep 11;3(3):e86. [doi: 10.2196/mhealth.4649] [Medline: 26362403]
13. Bhatia A, Kara J, Janmohamed T, et al. User engagement and clinical impact of the manage my pain app in patients with chronic pain: a real-world, multi-site trial. JMIR Mhealth Uhealth 2021 Mar 4;9(3):e26528. [doi: 10.2196/26528] [Medline: 33661130]

XSL•FO
RenderX

14. Priebe JA, Haas KK, Moreno Sanchez LF, et al. Digital treatment of back pain versus standard of care: the cluster-randomized controlled trial, Rise-uP. J Pain Res 2020;13:1823-1838. [doi: 10.2147/JPR.S260761] [Medline: 32765057]
15. Salazar A, de Sola H, Failde I, Moral-Munoz JA. Measuring the quality of mobile apps for the management of pain: systematic search and evaluation using the mobile app rating scale. JMIR Mhealth Uhealth 2018 Oct 25;6(10):e10718. [doi: 10.2196/10718] [Medline: 30361196]
16. Bhattarai P, Newton-John TRO, Phillips JL. Quality and usability of arthritic pain self-management apps for older adults: a systematic review. Pain Med 2018 Mar 1;19(3):471-484. [doi: 10.1093/pm/pnx090] [Medline: 28541464]
17. Lalloo C, Shah U, Birnie KA, et al. Commercially available smartphone apps to support postoperative pain self-management: scoping review. JMIR Mhealth Uhealth 2017 Oct 23;5(10):e162. [doi: 10.2196/mhealth.8230] [Medline: 29061558]
18. Portelli P, Eldred C. A quality review of smartphone applications for the management of pain. Br J Pain 2016 Aug;10(3):135-140. [doi: 10.1177/2049463716638700] [Medline: 27583140]
19. Rammo R, Gostkowski M, Rasmussen PA, Nagel S, Machado A. The need for digital health solutions in deep brain stimulation for Parkinson's disease in the time of COVID-19 and beyond. Neuromodulation 2021 Feb;24(2):331-336. [doi: 10.1111/ner.13307] [Medline: 33174292]
20. Alam M, Banwell C, Olsen A, Lokuge K. Patients' and doctors' perceptions of a mobile phone-based consultation service for maternal, neonatal, and infant health care in Bangladesh: a mixed-methods study. JMIR Mhealth Uhealth 2019 Apr 22;7(4):e11842. [doi: 10.2196/11842] [Medline: 31008716]
21. Whitehead L, Seaton P. The effectiveness of self-management mobile phone and tablet apps in long-term condition management: a systematic review. J Med Internet Res 2016 May 16;18(5):e97. [doi: 10.2196/jmir.4883] [Medline: 27185295]
22. Pecorelli N, Fiore JF Jr, Kaneva P, et al. An app for patient education and self-audit within an enhanced recovery program for bowel surgery: a pilot study assessing validity and usability. Surg Endosc 2018 May;32(5):2263-2273. [doi: 10.1007/s00464-017-5920-3] [Medline: 29098431]
23. Perdoncini NN, Schussel JL, Amenábar JM, Torres-Pereira CC. Use of smartphone video calls in the diagnosis of oral lesions: teleconsultations between a specialist and patients assisted by a general dentist. J Am Dent Assoc 2021 Feb;152(2):127-135. [doi: 10.1016/j.adaj.2020.10.013] [Medline: 33494867]
24. Shah SS, Gvozdanovic A, Knight M, Gagnon J. Mobile app-based remote patient monitoring in acute medical conditions: prospective feasibility study exploring digital health solutions on clinical workload during the COVID crisis. JMIR Form Res 2021 Jan 15;5(1):e23190. [doi: 10.2196/23190] [Medline: 33400675]
25. Davidovitch RI, Anoushiravani AA, Feng JE, et al. Home health services are not required for select total hip arthroplasty candidates: assessment and supplementation with an electronic recovery application. J Arthroplasty 2018 Jul;33(7S):S49-S55. [doi: 10.1016/j.arth.2018.02.048] [Medline: 29588123]
26. Felbaum DR, Stewart JJ, Anaizi AN, Sandhu FA, Nair MN, Voyadzis JM. Implementation and evaluation of a smartphone application for the perioperative care of neurosurgery patients at an academic medical center: implications for patient satisfaction, surgery cancelations, and readmissions. Oper Neurosurg (Hagerstown) 2018 Mar 1;14(3):303-311. [doi: 10.1093/ons/opx112] [Medline: 28541569]
27. Glauser G, Ali ZS, Gardiner D, et al. Assessing the utility of an IoS application in the perioperative care of spine surgery patients: the NeuroPath pilot study. Mhealth 2019;5(40):40. [doi: 10.21037/mhealth.2019.09.01] [Medline: 31620467]
28. Gustavell T, Sundberg K, Segersvärd R, Wengström Y, Langius-Eklöf A. Decreased symptom burden following surgery due to support from an interactive app for symptom management for patients with pancreatic and periampullary cancer. Acta Oncol 2019 Sep;58(9):1307-1314. [doi: 10.1080/0284186X.2019.1633473] [Medline: 31284797]
29. Hou J, Yang R, Yang Y, et al. The effectiveness and safety of utilizing mobile phone-based programs for rehabilitation after lumbar spinal surgery: multicenter, prospective randomized controlled trial. JMIR Mhealth Uhealth 2019 Feb 20;7(2):e10201. [doi: 10.2196/10201] [Medline: 30785406]
30. Pickens R, Cochran A, Tezber K, et al. Using a mobile application for real-time collection of patient-reported outcomes in hepatopancreatobiliary surgery within an ERAS® pathway. Am Surg 2019 Aug 1;85(8):909-917. [doi: 10.1177/000313481908500847] [Medline: 32051068]
31. Timmers T, Janssen L, van der Weegen W, et al. The effect of an app for day-to-day postoperative care education on patients with total knee replacement: randomized controlled trial. JMIR Mhealth Uhealth 2019 Oct 21;7(10):e15323. [doi: 10.2196/15323] [Medline: 31638594]
32. van der Meij E, Anema JR, Leclercq WKG, et al. Personalised perioperative care by e-health after intermediate-grade abdominal surgery: a multicentre, single-blind, randomised, placebo-controlled trial. Lancet 2018 Jul;392(10141):51-59. [doi: 10.1016/S0140-6736(18)31113-9]
33. Mundi MS, Lorentz PA, Grothe K, Kellogg TA, Collazo-Clavell ML. Feasibility of smartphone-based education modules and ecological momentary assessment/intervention in pre-bariatric surgery patients. Obes Surg 2015 Oct;25(10):1875-1881. [doi: 10.1007/s11695-015-1617-7] [Medline: 25702141]
34. Gomis-Pastor M, Mirabet Perez S, Roig Minguell E, et al. Mobile health to improve adherence and patient experience in heart transplantation recipients: the mHeart trial. Healthcare (Basel) 2021 Apr 14;9(4):463. [doi: 10.3390/healthcare9040463] [Medline: 33919899]

XSL•FO
RenderX

35. Collado-Borrell R, Escudero-Vilaplana V, Narrillos-Moraza Á, Villanueva-Bueno C, Herranz-Alonso A, Sanjurjo-Sáez M. Patient-reported outcomes and mobile applications. a review of their impact on patients' health outcomes. Farm Hosp 2022 Apr 27;46(3):173-181. [Medline: 36183211]
36. Wang TF, Huang RC, Yang SC, Chou C, Chen LC. Evaluating the effects of a mobile health app on reducing patient care needs and improving quality of life after oral cancer surgery: quasiexperimental study. JMIR Mhealth Uhealth 2020 Jul 27;8(7):e18132. [doi: 10.2196/18132] [Medline: 32716303]
37. Higgins J, Chang J, Hoit G, Chahal J, Dwyer T, Theodoropoulos J. Conventional follow-up versus mobile application home monitoring for postoperative anterior cruciate ligament reconstruction patients: a randomized controlled trial. Arthroscopy 2020 Jul;36(7):1906-1916. [doi: 10.1016/j.arthro.2020.02.045] [Medline: 32268161]
38. Armstrong KA, Coyte PC, Brown M, Beber B, Semple JL. Effect of home monitoring via mobile app on the number of in-person visits following ambulatory surgery: a randomized clinical trial. JAMA Surg 2017 Jul 1;152(7):622-627. [doi: 10.1001/jamasurg.2017.0111] [Medline: 28329223]

Abbreviations

DN4: Doleur Neuropathique 4
ODI/NDI: Oswestry Disability Index/Neck Disability Index
SF-36: 36-item Short Form Survey
VNRS: verbal numeral rating score

Edited by C Jacob; submitted 05.05.24; peer-reviewed by AM Pinard, Y Chu; revised version received 26.11.24; accepted 01.12.24; published 22.01.25.

Please cite as:
Gómez-González MA, Cordero Tous N, De la Cruz Sabido J, Sánchez Corral C, Lechuga Carrasco B, López-Vicente M, Olivares Granados G
Following Up Patients With Chronic Pain Using a Mobile App With a Support Center: Unicenter Prospective Study
JMIR Hum Factors 2025;12:e60160
URL: https://humanfactors.jmir.org/2025/1/e60160
doi:10.2196/60160

© Marta Antonia Gómez-González, Nicolas Cordero Tous, Javier De la Cruz Sabido, Carlos Sánchez Corral, Beatriz Lechuga Carrasco, Marta López-Vicente, Gonzalo Olivares Granados. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 22.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Exploring Older Adults' Needs for a Healthy Life and eHealth: Qualitative Interview Study

Paula Valkonen[1], MA (Art and Des); Sari Kujala[1*], PhD; Kaisa Savolainen[1*], MSc (Tech), DA (Art and Des); Riina-Riitta Helminen[2,3], MD, LLLic, DSc (Tech)

[1]Department of Computer Science, Aalto University, Espoo, Finland

[2]Suomen Terveystalo Oy, Suomen Terveystalo Oy, Helsinki, Finland

[3]Department of Industrial Engineering and Management, Aalto University, Espoo, Finland

[*]these authors contributed equally

Corresponding Author:
Paula Valkonen, MA (Art and Des)
Department of Computer Science
Aalto University
P.O.Box 15400
Espoo, FI-00076
Finland
Phone: 358 503862768
Email: paula.valkonen@aalto.fi

Abstract

Background: Aging brings physical and life changes that could benefit from eHealth services. eHealth holistically combines technology, tasks, individuals, and contexts, and all these intertwined elements should be considered in eHealth development. As users' needs change with life situations, including aging and retirement, it is important to identify these needs at different life stages to develop eHealth services for well-being and active, healthy lives.

Objective: This study aimed to (1) understand older adults' everyday lives in terms of well-being and health, (2) investigate older adults' needs for eHealth services, and (3) create design recommendations based on the findings.

Methods: A total of 20 older adults from 2 age groups (55 to 74 years: n=12, 60%; >75 years: n=8, 40%) participated in this qualitative interview study. The data were collected remotely using a cultural probes package that included diary-based tasks, sentence completion tasks, and 4 background questionnaires; we also performed remote, semistructured interviews. The data were gathered between the fall of 2020 and the spring of 2021 in Finland as a part of the Toward a Socially Inclusive Digital Society: Transforming Service Culture (DigiIN) project (2019 to 2025).

Results: In the daily lives of older adults, home-based activities, such as exercising (72/622, 11.6% of mentions), sleeping (51/622, 8.2% of mentions), and dining and cooking (96/622, 15.4% of mentions), promoted well-being and health. When discussing their needs for eHealth services, participants highlighted a preference for a chat function. However, they frequently mentioned barriers and concerns such as the lack of human contact, inefficiency, and difficulties using eHealth systems. Older adults value flexibility; testing possibilities (eg, trial versions); support for digital services; and relevant, empathetically offered content with eHealth services on short-term and long-term bases in their changing life situations.

Conclusions: Many older adults value healthy routines and time spent at home. The diversity of older adults' needs should be considered by making it possible for them to manage their health safely and flexibly on different devices and channels. eHealth services should adapt to older adults' life changes through motivation, personalized content, and appropriate functions. Importantly, older adults should still have the option to not use eHealth services.

(JMIR Hum Factors 2025;12:e50329) doi:10.2196/50329

KEYWORDS

older adults; eHealth; needs; retirement; well-being; cultural probes; sentence completion; human-computer interaction

XSL•FO
RenderX

Introduction

Background

Aging involves physical and life changes [1-4], where eHealth services have the potential to provide support and benefits [4-7]. The prevalence of many diseases increases with age [8,9], and multimorbidity is common among people aged ≥65 years [10,11]. Conversely, healthy behaviors can extend people's lives [12-14].

eHealth can help people nurture their health and well-being, potentially extending their lives. eHealth can also make treatment more accessible, promote treatment continuity, enhance communication, help shared decision-making, and enable patient self-management [10]. Older adults seem to be generally positive toward eHealth in a health management context [15]. They can benefit from eHealth services, as they often have complex health issues and engage in many self-management tasks supporting their health and well-being [1,5,16].

However, older adults sometimes face challenges with eHealth services [5,16,17], including a lack of skills and interest [3,16,18]. Hirvonen et al [19] reported that some older adults find the content or functionalities of eHealth services irrelevant (eg, [20,21]). Sometimes, older adults might even feel that digital services are not "meant for them" (eg, [20,22]). Thus, older adults might have a risk of exclusion from digital services [16,18,23,24] and society [16,23].

Older adults have not always been involved in technological development, especially in mainstream technology [25-27]. At the same time, the older population is growing [28], and health care costs are rising [29]. Stephanidis et al [30] encourage understanding emerging new technology possibilities and context-based user needs when designing digital solutions for the health and well-being (hereafter, *eHealth services* or *eHealth*) of older adults and a better understanding of the role of technology in older adults' lives [30]. Ideally, the technology should be harmonized with these adults' lives [15].

European older adults regularly engage in work, household chores, exercise, cultural activities, and tourism [31]. On the basis of the literature, many older adults in Europe live in rural areas, where they have easy access to nature and can engage in refreshing activities [31]. On the other hand, older adults with urban lifestyles access a variety of services [31]. For example, in Finland and Sweden, rural areas are in the north and distant from health care services, and eHealth services offer a promising option for taking care of one's health [32]. Physical activities might support older adults' health situation during aging. In 2017, in total 43.2% of Europeans aged 50 to 64 years and 44.5% of those aged 65 to 74 years old spent >3 hours per week engaging in physical activity [31]. On the other hand, older adults' daily routines differ. Older adults are a heterogeneous group of people, and aging is a process, not a static state [33]. Therefore, it is important to obtain qualitative insights into older adults' activities. These could help inform the idea and role of eHealth services and harmonize it with older adults' routines [15,30]. We know that older adults use eHealth services, for example, depending on their digital devices, capabilities, interests, or education [31,34]. Statistically, 100% of Europeans have, at least in theory, access to eHealth services or health information [35].

The best approach to eHealth service development means selecting the optimal approach case by case [36]. In eHealth, technology, tasks, individuals, and contexts are holistically combined, which should all be considered in developments [36-39]. eHealth services should genuinely support people pursuing healthier lifestyles and be accessible to all [10], although older adults may perceive the technology differently from other generations [1,40]. Humanity is needed alongside IT. Therefore, design decisions should also help ensure that as a part of eHealth services development, the human touch of personal care will not be removed [30,41].

Mapping user needs is important because without understanding needs, designing high-quality solutions is challenging [42]. User needs can be defined as the problems that prevent users from reaching their goals or possibilities to support them in reaching these goals [43]. User needs can be personal or more social in nature, and they can be related, for example, to available information or technology functions [44]. Poor understanding of users, user needs, and use contexts increases the risk of failure in eHealth service development [39]. On the other hand, user needs are not a stable phenomenon; they might change with age or the use of technology [45]. However, not much is known about what kind of user needs for eHealth services arise in older adults' lives, which is unfortunate because understanding user needs is a great innovation source [46]. In addition, Hirvonen et al [19] recommend further investigation into older adults' eHealth service use as a part of their daily lives.

Goal of This Study

We aimed to understand older adults' everyday lives in terms of well-being and health and their needs for eHealth services. We investigated 2 life stages: working at an older age and life after retirement, possibly with a chronic disease. We were interested in the health and well-being practices of 2 age groups to understand older adults' needs for eHealth services. Investigating these aspects allows developers and designers to better understand how eHealth services can meet older adults' needs and support their healthy living.

Users' needs change with life situations, including aging [47]. Therefore, if we want to harmonize eHealth services with older adults' everyday use, as is recommended by Cabrita et al [15], and to support technology development for well-being and an active, healthy lifestyle, it is important to identify user needs at different stages of life and daily routines in different life situations. The research questions were as follows:

1. What is everyday life like for older adults?
2. What kind of needs do older adults have for eHealth services?

Prior Work

Qualitative Methods for Investigation of Needs

User needs can be explored in many ways. In the human-computer interaction field [48,49], designers and

XSL•FO
RenderX

researchers have several methods for investigating user needs. Soraghan et al [50] and Dickinson et al [51] encourage researchers and designers to step into older adults' homes to assess technology use and empathize with respondents' lives, for example, through in-home interviews and observations. Understanding of older adults can be collected through remote, semistructured interviews [52]. Projective techniques, such as cultural probes [53-57] or sentence completion tasks [58], in which the participants provide information about themselves (instead of the researcher asking them direct questions), can be fruitful methods for gaining a deeper understanding of end users when the topic is challenging to verbalize or includes sensitive aspects [42]. For example, the sentence completion technique can be used to gather both an understanding of end users' values and needs as well as inspirational data for design and discussion activation [42,59,60].

Cultural probes are designed to trigger older adults to provide inspirational information about their lives to designers [53]. Following development, they have been widely used for various purposes, including collecting information about caregivers [56], investigating patient experiences with an eHealth service [54], and in other health care contexts [55,57]. Cultural probes are packages or toolkits of various tasks that end users complete and document [53]. While cultural probes show potential as a projective technique for gathering user needs [42], they appear to be underused in identifying older adults' needs for eHealth services.

Older Adults' Needs and Attitudes Toward Technology

Older adults' needs and attitudes toward technology vary. For example, some older adults do not trust computers and find the terminology and content of computers and digital services confusing [17,18,61]. They might even fear computers [61]. Their self-confidence with computers is sometimes fragile, and usability issues can be disastrous for their self-confidence [61]. In addition, older adults as a user group are much more diverse than many traditional user groups; they may have problems such as sensory loss, challenges with language, and attitudes to technology [61].

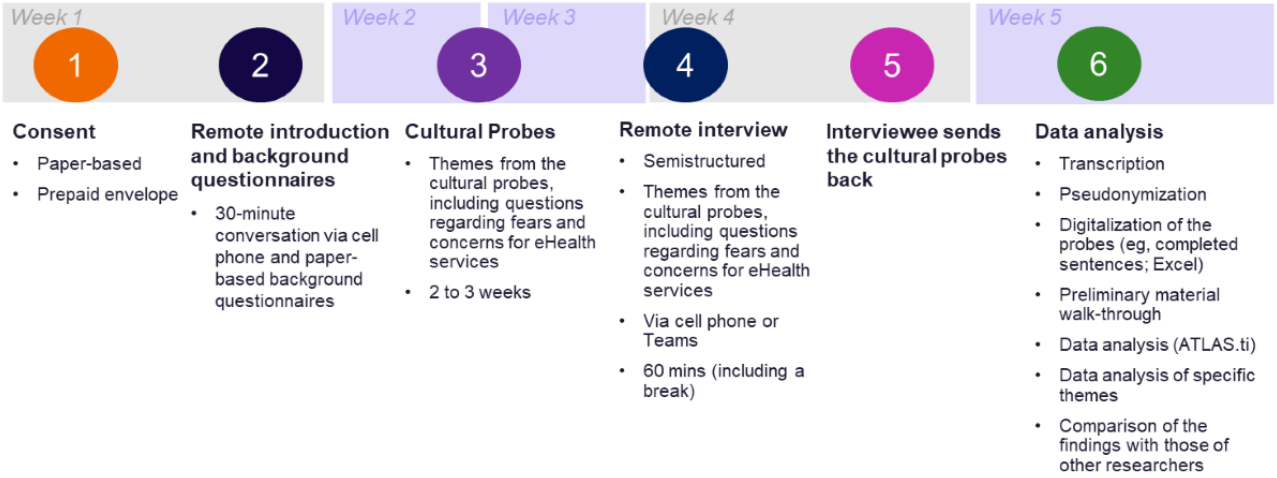
Many older adults appreciate easy access to eHealth services and their medical records, which should be offered in an easy, understandable, and secure way [5,52]. When medical information has compact and clearly labeled contents in well-organized menus [62], clear interface language, and linear navigation [20], this can assist older adults. Backonja et al [47] recommend eHealth services that adjust functionality and content based on the user's needs. Similarly, Cabrita et al [15] recommend personalization of design and functionality, thus lowering older end users' fears toward technology by allowing it to support its users and even offer empathy and sympathy to them.

Methods

Overview

This empirical qualitative study consisted of cultural probes [53,63], which included 4 background information questionnaires (*background questionnaires*), sentence completion tasks [42,58,59], and a diary-based exercise [64-66], and remote interviews [52]. Data collection included 6 phases, which all played an important role (Figure 1). Data collection started with filling out a consent form. After that, background information was collected, and a short remote interview was conducted. Then, the participant had time to respond to the cultural probe package (ie, sentence completion tasks and diary-based exercise). Participation ended with a long remote interview and the return of the probes package to the researcher. Finally, the data were analyzed.

Figure 1. Research procedure.



Sampling and Recruitment

The research was conducted in Finland between the fall of 2020 and the spring of 2021 (ie, during the COVID-19 pandemic) as part of the DigiIN project (2019 to 2025).

Defining an older adult purely by age is somewhat complex [67]. Generally, they are people in late adulthood. In line with the study by Ware et al [5], this study focused on older age groups associated with a higher risk of chronic disease, with participants from 2 age groups (those aged between 55 and 74

XSL•FO
RenderX

years and employed and those aged >75 years and retired). Group 1 participants (ie, those who were still working) were relevant because aging includes preparing for retirement [68].

The participants were selected through purposive sampling (digital survey) and snowball sampling [69]. Group 1 participants (12/20, 60%) were recruited via a digital survey [70]. Group 2 participants (8/20, 40%) were recruited in 2 ways: via the digital survey and snowball sampling. To ensure a diverse and engaged sample for this long-term study, participants' interest in digital services was assessed in advance. The groups were formed according to the life situation: group 1 participants were still working, and group 2 participants were retired. The age category of group 1 was between 55 and 74 years (later, 55-74 years). The age category of group 2 was between 75 and >90 years (later, >75 years). Participants in group 1 had different professions, such as yoga teacher, director, project manager, or entrepreneur. The participants were from different parts of Finland but did not fully cover the whole country.

After participants expressed potential interest in participating in the study, the researcher explained the study in a telephone call. This call lasted for 15 to 30 minutes, depending on the participant's questions. If the participant was still interested, the researcher mailed or emailed the information letter that described this study in detail and a consent form to the participants for their signature.

Ethical Considerations

The empirical study was conducted as part of the DigiIN project (2019 to 2025). The study protocol was reviewed and approved by the Ethical Review Board of Aalto University (95_03.04_2019_DigiIN). All participants provided their voluntary, informed, and written consent. Patients' ability and willingness to participate in the study were confirmed through a phone conversation before they signed the consent form. During the call, the study process was explained, and it was emphasized that participation was voluntary, with the option to withdraw at any time. This was also mentioned in the written consent form.

Research Procedure

The study comprised six phases (Figure 1):

1. First, the study purpose and research procedure were introduced.
2. Second, the cultural probe packages, including 4 background questionnaires, were delivered. The study procedure and contents of the package were explained again in detail by phone or Teams (version 1.6.00.11166; Microsoft Corporation).
3. Third, the participants filled out the cultural probes for 2 weeks, during which the researcher made 2 short phone calls to the participants, being present for possible questions and collecting specific day-based information about the participants' routines and activities related to health and well-being.
4. At the end of the study, the participants were interviewed remotely.
5. Then, they returned the probes package to the researcher in a prepaid envelope. The data collection took a maximum of 5 weeks. When the researcher received the probes package, the participants were sent a small thank-you gift.
6. Data were analyzed.

Cultural Probes

Overview

In this study, cultural probes made data collection of older adults' everyday lives possible in research environments, where remote data collection was important for safety reasons [52]. It helped in collecting data indirectly, as the participants themselves completed the cultural probes at their own pace as part of their everyday lives [42,52,53]. It also helped the participants prepare for the interview [52].

The cultural probe package, recruitment questionnaire, and recruitment letter were tested with 2 nonparticipating members of the target group. The probes, questionnaires, and letter were revised based on the feedback. The research materials were then tested again with another target group member. Finally, the entire study setup with semistructured interviews was tested with 2 target group members.

Background Questionnaires

Among other tasks, the cultural probe package included 4 questionnaires: a background information questionnaire, the Health Confidence Score (HCS) [71], the eHealth Literacy Scale (eHEALS) [72], and the European Health Literacy Survey Questionnaire (HLS-EU-Q16) translated into Finnish by Eronen et al [73]. The other questionnaires were translated into Finnish by the authors. Questions 3 to 10 in the eHEALS questionnaire were used in this study with the added option of "I don't know" [72].

In addition to demographic questions, the background information questionnaire asked about the participant's internet and device use, the eHealth services they had used, and how useful they found them. The purpose of the HCS, eHEALS, and HLS-EU-Q16 questionnaires was to obtain an overview of the participants' capabilities (or perceived skills) to understand health information and use eHealth services.

Diary-Based and Sentence Completion Tasks

The cultural probe package (1) motivated participants to notice and document their actions related to health and well-being as a part of their everyday lives and (2) worked as a triggering element for remote, semistructured interviews. In this remotely conducted study, cultural probes and their exercises helped make the researcher appear more relatable to the participant. This was especially important given the sensitive and personal topics, such as everyday life around health and well-being, discussed in the study, both from the researcher's perspective, and more importantly, from the participant's perspective [74]. The cultural probe tasks included sentence completion tasks [59] and a diary-based task [64-66]. In the diary-based task, participants were asked to make notes regarding their health routines, decisions, and actions for 5 days. In addition, 13 incomplete sentences were formulated following the best practices explained by Nurkka et al [42] and included in the

XSL•FO
RenderX

pretesting phase. In the sentence completion task, participants were asked to complete sentences (Multimedia Appendices 1 and 2) with 4 themes (Table 1).

Table 1. Themes of the sentence completion tasks.

Theme	Included in group 1	Included in group 2
eHealth services now	Yes	Yes
eHealth services in the future	Yes	Yes
Concerns and fears for eHealth services	No	Yes
Informal caregiver's eHealth services (excluded from the data)	Yes	Yes

The cultural probe packages had differences regarding incomplete sentence themes between the group 1 and group 2 participants. The study started with the younger participant group. As data began to accumulate, additional questions addressing concerns and fears were incorporated into the sentence completion tasks, as shown in Table 1. This could support the participants in preparing for the interview as well [52]. However, the interview covered the theme with both participant groups.

Remote, Semistructured Interviews

The themes of the cultural probes were investigated in more detail in remote, semistructured interviews [52], which took 1.5 hours maximum to complete with group 1 participants and 1 hour maximum with group 2 participants. If an interview took longer, the participants were offered the opportunity to continue the interview the next day. The researcher made phone calls using either the Teams call function or a mobile phone, making it easier for participants to take part. Two researchers conducted the interviews, but only 1 conducted each interview for trust-building purposes. The interviews were audio recorded and documented with field notes.

The interviews comprised different perspectives regarding the participants' everyday lives and activities concerning health and well-being. The interview structure (Multimedia Appendix 3) was created based on the themes of the cultural probes. The participants were asked to describe the contents of their cultural probes. The idea was to empower the participants and let them decide how much or in how much detail they were willing to express their thoughts on their well-being and everyday lives. In the interview design, the interview checklist by Tong et al [75] was followed.

Analysis

Overview

After the data were transcribed and pseudonymized, the background questionnaire was analyzed using descriptive statistics, and the qualitative data were analyzed following the inductive content analysis process [69,76,77]. User needs were formulated based on the content analysis of the interviews and the analysis of the cultural probe data. The cultural probe data included the participants' diary notes of their daily well-being activities and completed sentences. The data were analyzed according to data type. The different analysis phases are explained in the subsequent sections.

Preliminary Material Walk-Through

During the interviews, a researcher (PV) identified the preliminary content-based categories and, after each interview, categorized each participant's direct quotes in Excel (version 2208 Build 16.0.15601.20644; Microsoft Corporation). Due to the richness of the data, an additional preliminary review was necessary. The researcher (PV) manually reviewed the entire dataset through after receiving the cultural probes package, using traditional pen-and-paper data analysis methods [78].

After this preliminary data analysis phase, the analysis needs for the next phases were recognized. The analysis needs were identified, keeping in mind that the purpose was to understand user needs. Similar to the data analysis process proposed by Nielsen et al [79], the preliminary material walk-through process needed both systematic and circular practices.

Content Analysis for Interview Data

Afterward, 2 researchers (PV and KS) read the pseudonymized interview transcripts and coded the data in ATLAS.ti 9 (version 9.1.5.0; Scientific Software Development GmbH; Multimedia Appendix 4) to understand the data as a whole and to support the analysis of the cultural probes data later. The process proposed by Mayring [76] was followed in the analysis. When approximately 30% (11/40) of the transcripts were coded, the researchers (PV and KS) revised and compared their categories and agreed on common ones. Finally, the results were discussed again by 2 researchers (PV and KS), and the findings were reported.

Analysis for Background Information Questionnaire Data

The background information questionnaires were analyzed to understand the participants' characteristics and their internet and IT use habits. These questionnaires provided preliminary information about the sample. The questionnaire analysis was done in collaboration between 2 researchers (PV and SK). The means were calculated for the data from each of the 3 questionnaires (PV), and the results were discussed (PV and SK). Owing to the qualitative nature of the study and the small sample size, detailed quantitative analysis was not performed.

Content Analysis for Cultural Probe Data

The data from the cultural probe package were analyzed in collaboration between 3 researchers. Each sentence's main content was analyzed by a researcher (PV), and the results were discussed with 2 other researchers (KS and SK) to ensure the quality of the analysis. The diary analysis focused on the

XSL•FO
RenderX

participants' descriptions of their well-being activities. Activities were written on sticky notes, 1 per note, with the participant code and time of day. After 2 researchers (PV and KS) created the notes for 1 participant, their consistency was checked: a researcher (KS or PV) read the findings and another one (KS or PV) wrote the findings down. The analysis included 644 grouped sticky notes. The sticky notes were grouped by 2 researchers. The notes were taken during the analysis process.

Results

Overview

The results are discussed in three parts: (1) participant information; (2) description of older adults' everyday lives; and (3) user needs, which were expressed based on the findings from the sentence completion and remote, semistructured interview data.

The data comprised 19 cultural probe packages and 329 pages of transcripts of 20 interviews (in Verdana font, 11-point font size, single spaced). One participant did not return the cultural probe package.

Participants

The study included 20 participants (Table 2). The participants were divided into two groups: group 1, those who were still employed, and group 2, those who were retired. Both groups included active computer users with many different devices, but compared to group 1, group 2 included more participants who did not routinely use the internet and IT. However, the variation in device types was greater in group 2 compared to group 1. The results suggest that perhaps retired people use health and well-being services less frequently in general than those who are working.

Table 2. Participant background information (N=20).

	Group 1 (n=12)	Group 2 (n=8)
Age (y)	55 to 74	>75
Gender, n (%)		
Women	9 (75)	6 (75)
Men	3 (25)	2 (25)
Internet activity, n (%)		
Internet use daily or many times in a day	11 (92)	6 (75)
Internet use on a weekly basis	1 (8)	1 (12)
Do not know to use or do not use the internet	0 (0)	1 (12)
Internet habits, n (%)		
Most popular devices for internet use		
Computer	12 (100)	5 (63)
Cell phone	12 (100)	4 (50)
Tablet	9 (75)	4 (50)
Place of internet use		
At home	12 (100)	7 (88)
Outside home	10 (83)	3 (38)
The most popular reasons to use the internet		
Searching information	12 (100)	7 (88)
Banking	12 (100)	6 (75)
Health and well-being services	12 (100)	4 (50)
eHealth use, n (%)		
Most popular digital public health care services used		
Patient portal: checking health information and receipt renewal	12 (100)	4 (50)
Appointment booking	11 (92)	4 (50)
Chat, email, or SMS text message to health care professionals	8 (67)	—[a]
Visiting a health center's web pages	—	2 (25)

[a]Data not available.

XSL•FO
RenderX

All group 1 participants (12/12, 100%) and half (4/8, 50%) of the group 2 participants checked their health information and renewed their prescriptions via eHealth services. Internet users used it for banking and information searches, and 80% (16/20) of them also used it for health services.

Multimedia Appendix 5 presents the background questionnaire results. The eHEALS results (20/20, 100%) show the participants' familiarity with using digitally offered health information, with differences between the participant groups. The results of the HLS-EU-Q16 (19/20, 95%) on health literacy indicate that the participants seemed to have at least basic health literacy skills. One participant did not send the questionnaire back to the researchers. The HCS results (20/20, 100%) on patients' confidence in taking care of their own health show that the participants had at least a basic understanding of their health situation and how to operate in case of health challenges [63].

Everyday Life of Older Adults

The analysis of the data from the diary-based task and the interviews resulted in 15 themes in the participants' everyday activities. The themes were spending time at home, dining and cooking, routines, hobbies or exercising, sleep, mundane activities, work-related activities, medical treatment, cottage or nature, friends, pets, pampering and rest, communication with a relative, planning and controlling everyday life, and weather.

The most common activity was spending time at home, often watching television (102/622, 16.4%); spending time at home and dining and cooking (96/622, 15.4%) were the most popular regular activities of the participants. The importance of food for well-being was often mentioned in the data. Exercising was also popular (72/622, 11.6%), with walking and jogging being commonly practiced, including Nordic walking and climbing stairs. Walking could also be a social activity with a partner or a dedicated group. Sleep was mentioned in the data 51 out of 622 times (8.2%), as were mundane activities, such as going to the post office or having lunch or a coffee (51/622, 8.2%). The repetitive nature of routines, including spending time at home, the importance of food, and the variety of exercises, is reflected in the following quotes (the quotes are translated from Finnish):

> *From Monday to Friday, I wake up 5.50 am on two mornings, and on other mornings 6.50 am. And that 5.50 am means, that I go swimming. And on other mornings I eat my breakfast normally at home and go to work. I drink my morning coffee at the office.* [Group 1, participant U2]

> *There have been such beautiful days lately. Today it seems it is not so beautiful, or at least the sun is not shining yet. I feel I'm privileged that I'm able to live in my hometown, because this is such a beautiful city. I started with 30 min walking exercises, and now I do my walking exercise two times in week. Each walking exercise takes one and half hours nowadays. I explore those different terrains, and we have here a lot of a lot of stairs and a lot of small hills, which I use to improve my fitness.* [Group 1, participant U12]

> *I start my work at 9 am, then I eat breakfast, and walk with my dogs. I normally work at my computer until 6–8 pm. I eat, and drink coffee during short work breaks. At 4 pm I give food to my dogs. In the evening, I don't meet anybody else than other dog owners. I also meet new employees during my workday.* [Group 1, participant U6]

> *I eat breakfast in the morning, and at the same time, I read news from my iPad. My husband and I have not ordered any paper newspapers.* [Group 2, participant HX4]

> *I might visit quite often to meet my neighbor, but not on an everyday basis. But it is a part of my everyday life. ...And then, I go quite often after 6 pm a short walk, and at the same time, visit at my friend, who has two cats. ...Then I make some dinner and watch TV News. I do not usually follow any TV series, but I might sometimes watch something from TV. I go to sleep at around 12 am.* [Group 2, participant HX5]

> *I don't have so many meetings anymore, but my week starts with English lessons every Monday at 10 am. At 4 pm I go to an hour's outdoor exercise. In the evening, I often go to German language speaking exercise group. Next day, I go to aqua jogging, and after that we take a cup of coffee with the members of that group. Next day I participate in outdoor walking exercise group, with which we normally walk an hour together. And at the end of the week, I often do some housework, a little bit cleaning home or something.* [Group 2, participant U7]

Watching television was emphasized in both groups, and activities related to cleaning and housekeeping were popular. Overall, participants in both groups seemed to like routines. The everyday life activity themes are detailed in Multimedia Appendix 6.

The Needs of Older Adults for eHealth Services

Overview

The older adults' needs for eHealth services were collected based on the main findings from remote, semistructured interviews and sentence completion tasks. The themes discovered from the interviews, the main findings under each theme, and the participants' needs based on those findings are provided in Multimedia Appendix 7. In addition, sentence completion revealed the participants' wishes and needs for eHealth services in the future (Multimedia Appendices 1 and 2). The needs described in the following sections were expressed based on the data.

The Need for a Carefree Mind

Older adults need a carefree mind when taking care of their health and regarding their current health and well-being situation, which can be supported with eHealth services. They need to understand the safety of eHealth service use and be able to manage health-related matters well regardless of the channel (eg, mobile phone, chat, or remote appointment). Users' technical skills vary, and some use technical devices versatilely. Support and training possibilities help with this need. Feedback

XSL•FO
RenderX

on one's current health and well-being situation via eHealth services is appreciated. The safety needs are evident in the following quotes:

> *I avoid using all parts [of the user interface] that have even a little bit of a foreign language to me. Because it might happen there that I don't understand something and make a wrong choice somehow [in the user interface].* [Group 1, participant U2]

> *I think managing my health-related affairs should be safe.* [Group 1, participant U6]

The need to obtain health care service regardless of the service channel is reflected, for example, in the following comments:

> *I would like to be able to use eHealth services in parallel with other service channels in the future.* [Group 1, participant U1]

> *I look for information on the internet every day. I use all possible medical and health services that are available digitally. And the same thing with public services: everything that is possible, I do them in digital channels.* [Group 2, participant H3]

On the basis of the data, people's diversity is emphasized: being able to use the eHealth service can be a big challenge for one person, but for another, it brings benefits and ease. Owing to older adults' varying technology skills and device availability, health self-management was desired in varying ways and different channels, either face-to-face or digitally. In addition, group 1 participants seemed to be experienced or very experienced with technology and felt very comfortable using eHealth services. The clear variations between the participant groups should be considered in all design decisions regarding eHealth service development.

Varying skills and comfort levels with using eHealth services are demonstrated, for example, in the following quotes:

> *For me, using eHealth services is natural and easy.* [Group 1, participant U6]

> *For me, using eHealth services is easy and self-evident.* [Group 2, participant HX4]

> *eHealth services do not help me because I have neither the equipment nor the skills.* [Group 2, participant S8]

> *Learning something new [about digital services] is difficult because of the tricky terminology. I've never studied English at school, and systems use a special system language. ...Then these jungles of safety encryptions and passwords: remembering them is almost the most difficult part of it.* [Group 2, participant S9]

The Need for eHealth Service Adaptation Based on Current Life Situation, Use Contexts, and Everyday Life

Regarding different life situations, the participants needed support via eHealth with their health and well-being in changing situations, such as when retiring, regardless of their current technology skills or devices. eHealth could help them, for example, with mental well-being or keeping physical activity levels high enough. The changing situations are evident in the following quotes:

> *My working years are in the final phase. It's sure that I have a couple of years left until I retire officially. After that, my children are already adults as well. But currently, my life is quite work-oriented, long workdays, and I'm all the time busy.* [Group 1, participant U1]

> *I have done more or less work [in my life]. And now I'm retiring. That's a pretty big change.* [Group 1, participant U4]

> *I'd like to use eHealth services in the future even after I retire.* [Group 1, participant U12]

> *I have a little bit tricky stage of life. My son went to heaven two weeks ago, and I am currently going through grief. Five years ago, my husband died. So, that kind of stage of life I have now.* [Group 2, participant H2]

> *Yes, it was easier when I was healthy. After all, then I could do what I wanted and run if I liked and, in every way. ...That life was completely different then. I can't say, however, that none of these moments in this life have been completely unpleasant. ...When I think about this end of life, yes, I have had a good life the whole time, no matter what era it was. They have all been part of life that was then, and it was good at that moment.* [Group 2, participant S7]

Older adults desired eHealth services that could adapt to their current and evolving life situations, such as retirement or changes in health. These services should be tailored to their lives in terms of both content and functionality, including the user interface, and should be flexible enough to adjust according to the season, regardless of their location. Older adults want to take care of their health and well-being regardless of location (from places other than home or a clinic, such as a summer cottage). Service adaptation to everyday life makes eHealth services relevant.

However, the ease of use of an eHealth service becomes relevant only after it has been ensured that it can be accessed at all. One's health state affects everyday life and the ability and willingness to use eHealth services. For example, when a health situation improves due to following healthy daily routines, the content of the eHealth service should follow this new health situation and its user's new needs for the eHealth service. eHealth services should be adaptable to changes in life and health situations as well as fluctuations in motivation to use eHealth, both in the short and long term. eHealth service's role and value as a part of changing life situations and user needs regarding eHealth service's adaptability are reflected in the following quotes:

> *eHealth services help me very well in my health situation.* [Group 1, participant U12]

> *eHealth services help me plan schedules flexibly.* [Group 1, participant U7]

> *For me, the most important thing with eHealth services is that I don't have to commit to a specific time.* [Group 1, participant U10]

XSL•FO
RenderX

For me, the most important thing in eHealth services is the right information at the right time. [Group 2, participant S9]

For me, the most important thing with eHealth services is that I get information immediately without waiting. [Group 2, participant HX4]

Using them [eHealth services] saves time. [Group 1, participant U8]

The Need for a Holistic Perspective

Older adults need content that supports their health and well-being holistically. eHealth services should, for example, include perspectives on their current everyday lives and health situations. Group 1 members often wished that eHealth services would be one service among others, smoothly integrated into everyday life. For group 2 members, more research is needed to investigate their perspectives on a suitable package of service channels and content so that they receive the same holistic support as group 1 members. The example quotes behind the need for holistically offered health and well-being support as a part of older adults' everyday lives are as follows:

I fill in crosswords, and I read when I like to. I cook, but not every day now. Things like that I get pleasure from. After all, all in all, the fact that there are not so many financial worries is a part of that balance. Everything you need is here and now. Everyone can ask themselves the question whether they are happy. [Group 1, participant U11]

For me, the most important thing in eHealth services are the right health goals considering the right information; based on researched and measured information. [Group 1, participant U7]

eHealth services help me to plan flexible schedules. [Group 1, participant U7]

eHealth services help me take care of things related to my health at once. They save time. I get e.g., lab results through a digital service. [Group 1, participant U13]

eHealth services help me get an overall picture of my health. [Group 1, participant U1]

I read a lot. Now that I've been here sick and the moments, I've been awake, I'm reading something all the time. [Group 2, participant HX6]

And then jogging with the dog. It keeps me in balance. All my social interaction now takes place there, walking the dog, because there are many dogs here, and we have lived here 45 years, so there are so many people who know dogs. [Group 2, participant S8]

The Need to be Able to Avoid Using eHealth Services But to Use Specific eHealth Functions

Older adults, especially group 2 members, need to be able to select the service channel: eHealth services or face-to-face. On the other hand, there was a need for eHealth functions. For example, group 2 members often appreciated the opportunity to look at laboratory results without having to go anywhere. The COVID-19 pandemic era might have affected this result. The following example quotes bring insights into the older adults' needs for specific functions and the possibility of obtaining health care services via service channels other than eHealth services:

In principle, they [service providers] accept to do business [face-to-face], but if you ask for some advice there, they recommend that you "go to do it online or look it up online." They don't understand that not everything works perfectly on the internet for them either. I think it's pointless for them to guide you to do business online when it can't be done there online. [Group 1, participant U8]

Or you can use a video connection to contact the nurse! They are certainly good additions to the healthcare service. [Group 1, participant U1]

It's embarrassing that I don't know how to book appointments [digitally]. I don't know how to book a time for our laboratory or X-ray. I don't know how to make an appointment. And you should know that, but when there was no need, you didn't learn. [Group 2, participant HX6]

When I imagine myself retired, and when I no longer have a work computer, and if it is difficult to get from one place to another, then of course, they [eHealth services] will be useful. And then the chat service, where you can easily ask for help or advice!

I think managing my health-related affairs should be possible in person. [Group 2, participant S8]

This means that eHealth services should be accessible: (1) devices should be available and (2) the network connection should be stable enough, and the service must be available regardless of time. The service should be offered in a multichannel manner, including personal contact, phone service, or in writing (eg, chat).

Operations related to health and well-being, such as appointment booking through eHealth services, must be clear and easy to use for all user groups. The operations and the eHealth service should support long-term and short-term health and well-being planning in terms of the content and functions available to both end users and health care professionals. It should work reliably in all situations and support older adults' empowerment in managing and controlling their current and future health situations. This is evident, for example, in the following quotes:

I think operations related to my health and well-being should be easy. [Group 2, participant HX5]

eHealth services do not help me at all. [Group 2, participant HX6]

User Needs Regarding User Experiences

The completed sentences revealed that the older adults' experiences with eHealth services varied, with some finding them easy to use and others finding them challenging. The use of eHealth services was felt to be part of everyday life and routines, and it did not produce many "peak experiences," which are the most memorable experiences producing positive

XSL•FO
RenderX

emotions [70,80]. The variation in user experiences from eHealth use is apparent in the following example quotes:

> *For me, using eHealth services is easy, and a routine.* [Group 1, participant U13]
>
> *Using them [eHealth services] makes me a little irritated.* [Group 1, participant U1]
>
> *I think managing my health-related affairs should be confidential, and the response quick and empathetic.* [Group 1, participant U11]
>
> *Using them [eHealth services] puts me in a bad mood.* [Group 2, participant HX6]

Two (25%) out of 8 participants in group 2 had worrying experiences, such as a loss of self-confidence in the use of eHealth services and an increase in concern about the use of eHealth services. These worries are reflected in the following example quotes:

> *Using them [eHealth services] makes me feel insecure.* [Group 2, participant S7]
>
> *The thing that worries or scares me the most about eHealth services is that soon you won't be able to see a doctor anymore when everything will be done digitally.* [Group 2, participant S8]

However, eHealth services especially brought peace of mind to group 1 participants, seeming to improve their well-being. For both groups, eHealth services produced important information about their current health situation. eHealth services seemed to be used in both age groups, not only out of necessity but also to obtain information and improve the flexibility of taking care of their health. eHealth services provided longer-term insights into their own health status, and information obtained through eHealth services was highlighted at different points. Finding out about their own health status seemed important. With the help of eHealth services, the level of concern about their own health condition improved. Group 2 members had more negative views of eHealth services, and they also mentioned a lack of equipment as a barrier to eHealth service use. eHealth service's role in offering information, flexibility in taking care of one's health, or role as a calming factor is reflected in the following example quotes:

> *I use eHealth services if I book an appointment with my doctor, or check my health information* [Group 1, participant U12]
>
> *I use eHealth services whenever possible.* [Group 1, participant U6]
>
> *Using them gives me peace of mind when I can book appointments and find out about health-related issues at a time that suits me.* [Group 1, participant U13]

Other barriers to eHealth services were lack of human contact, inefficiency, and challenges with its use. eHealth services divided participant opinions regarding its benefits: group 1 participants did not have as many helpful experiences with eHealth services as group 2 participants. Both groups included individuals who felt that eHealth services did not provide holistic support during life changes. At the same time, group 2 participants felt that eHealth helped them acquire information about health-related issues, supported them in planning health-related issues, and helped them commit to lifestyle changes. On the other hand, they lacked competence and felt that eHealth services (use processes and user interfaces) was too complicated. The negative experiences and use barriers are evident in the following quotes:

> *With eHealth services, I am saddened by the lack of consistency and aimlessness, as well as the lack of genuine and cheerful support.* [Group 1, participant U7]
>
> *With eHealth services, I am frustrated by delays and waiting.* [Group 1, participant U6]
>
> *With eHealth services, I am saddened by my laziness and lack of time available to study the use of eHealth services.* [Group 2, participant H3]
>
> *With eHealth services, it saddens me that I don't know how to use them.* [Group 2, participant HX6]

The most important things in eHealth services use were efficiency, ease and smoothness of transactions, access to information, planning (where group 1 members had diverse answers), and support for well-being. Participants in group 2 had the same eHealth needs as those in group 1 but they were more skeptical (Multimedia Appendix 8).

The efficiency, ease, and support for well-being via eHealth services are apparent in the following quotes:

> *Using them [eHealth services] saves my time.* [Group 1, participant U6]
>
> *Using them [eHealth services] makes me think about my health based on researched information.* [Group 1, participant U10]
>
> *For me, the most important thing about eHealth services is that I don't have to commit to a specific time or date.* [Group 1, participant U10]
>
> *eHealth services help me renewing my medicine prescriptions.* [Group 2, participant S7]

Discussion

This qualitative exploratory study aimed to gain a more in-depth understanding of older adults' needs for eHealth services as part of their everyday lives in different life stages. The data were collected with interviews and a cultural probe package, including 4 background information questionnaires, sentence completion tasks, and a diary-based exercise.

Principal Findings

Everyday Life of Older Adults

Both participant groups valued time spent at home and home-based activities as part of their everyday lives. Healthy eating habits were an especially important aspect of the participants' well-being from their own perspective. These older adults enjoyed different activities, such as jogging and Nordic walking, which is in line with other literature [31].

Older adults generally appreciated bringing health and well-being into everyday life. They had weekly or daily

XSL•FO
RenderX

activities or hobbies that supported them to reach their long-term health targets. These routines varied greatly, including exercise, good night sleep, personal hygiene, repetitive food, pet care, and housework-related morning routines. Group 1 participants reported routines and activities that partly differed from those of group 2 participants. For example, all the routines related to the gym or exercising while commuting were reported by group 1 participants. Interestingly, the only mention of reading newspapers with a tablet was reported by a participant from group 2. Furthermore, a mention of doing morning exercise with television was reported by a group 2 participant. Cleaning routines were mainly mentioned by participants from group 2.

The eHealth service can both support health self-management on a long-term basis and activate and support older adults in ad hoc type of activities and operations related to health and well-being, for example, by offering information based on the current health situation or reminding them of a (daily or weekly) hobby.

The Needs of Older Adults for eHealth Services

On the basis of the interviews and sentence completion tasks, four needs for eHealth services were identified: (1) the need for a carefree mind; (2) the need for eHealth service adaptation based on the current life situation, use contexts, and everyday life; (3) the need for a holistic perspective; and (4) the need to avoid using eHealth services but to use specific eHealth service functions. These needs should be considered when designing eHealth services that support older adults' healthy living. In addition, they generally apply to both user groups but may be emphasized differently depending on the group.

On the basis of our findings, the older adults seemed to have generally positive attitudes toward eHealth, aligning with those identified by Cabrita et al [15] and Mielonen et al [28]. The infrastructure behind eHealth services must be stable; then, eHealth services can be offered in an accessible way and without limitations for everyone, assuming that end devices are available. Owing to varying end devices among older adults, eHealth services should be scalable regardless of the service channel or devices. Therefore, the design should support the reliability of eHealth services so that they work in all situations, supporting older adults' health self-management [15,70]. Our study adds the perspective of changing time spans: the service must function smoothly for both short-term and long-term use, accommodating variations in use duration.

Well-being, sickness, health, and illness seem to be subjective personal experiences. The need for eHealth services changes due to the user's current health and well-being situation: eHealth services become unnecessary when health and well-being increase. Therefore, it must be possible for older adults to easily try out whether eHealth services will benefit them. An eHealth service may also be used less frequently at times.

eHealth services can support older adults' health and well-being, particularly when the content of the service aligns with their everyday lives and interests. Therefore, we recommend that the eHealth service work as part of an older adult's healthy routines, which might include, for example, encouraging attitudes and functions for housework and other activities.

Comparison With Prior Work

Our results show that understanding the user's life stage and the key interests at that stage is important, as found in other literature [47]. Owing to older adults' changing health and life situations [1-3,10,11,47], an individual eHealth service should function as part of a wider service offering so that the service supports and adapts to changing life situations and health statuses. For example, in this study, it was recognized that both those who are still working and those who are retired use IT and eHealth services. IT use at work might support learning new skills and sometimes even updating them [70]. In addition, eHealth service flexibility is valued [4]. The need for eHealth services to adapt to the older adult's current life and health situation is in line with the study by Reiners et al [81].

Earlier research recognized a large variation in activities that support older adults' well-being, in line with our findings: exercising; cooking; and hobbies, such as language courses [82]. Even with poor health, older adults achieve well-being and positive effects from daily activities [82,83]. In addition to physical activities, daily activities, such as reading books or watching television, can also offer relaxation [81]. Owing to the differing routines and daily activities, the need for eHealth adaptation based on the current life situation, use contexts, and everyday life was recognized in this study, and we recommend that eHealth services work as a part of older adults' daily living smoothly. Previous research supports our results behind this user need: older adults are a heterogeneous group whose activity possibilities vary regarding their living conditions or current health situation [31]. However, the importance of studying the use of eHealth service as part of everyday life has been recognized [19].

The devices with which the eHealth services are used should be considered when designing these services. Given the possible effect of the small sample size, it can be seen from the data that group 2 participants seemed to use eHealth services somewhat less than group 1 participants. One reason for that could be that older adults who are still employed might use eHealth services for work reasons, as well [70]. On the other hand, it can probably be expected that, nowadays, older adults who are retired have used at least some IT at work.

Understanding the safety of eHealth services use was important to older adults, as in earlier research [4,5,10,18]. In line with previous literature [18,70], they needed available support and training for eHealth service use. In particular, the older participants in this study also needed not to be forced to use the eHealth service. However, this varied considerably between participants, which supports earlier research results [18].

In line with an earlier study [15], one of the most important design functions is to ensure that the eHealth service conveys empathy to older adults. This can be offered via an empathic tone of voice in the services and by offering them through service channels that are especially suitable to each older adult's current situation. The whole life span, including transitions to different life phases, such as retirement, should be considered in design decisions [2,47,70]. This can be done, for example, by offering relevant content related to health and well-being to

XSL•FO
RenderX

different older adult groups and by making sure that the content is relevant to each end user's unique life and health situation.

On the basis of our findings, both those still working and those who have already retired have suitable devices and sufficient technical skills, as also reported by Mielonen et al [28]. However, eHealth services were used only by half (4/8, 50%) of the older participant group. Our findings support the observations of Kruse et al [3] regarding common barriers to eHealth use among older adults: a lack of self-confidence, a fear of making irreversible mistakes, and content that does not meet their needs. Fear can be dispelled with a clearer service concept, implementation support, user support, and instructions, as well as reassurance, such as by highlighting safety both in the service content and in the presentation of the service. In addition to minimizing fear, strengthening the self-confidence of the end user of the service plays an important role. Another phenomenon that emerged from the data was that well-being, health, and illness seem to be subjective, personal experiences.

The need for eHealth services changes due to the user's current health and well-being situation; eHealth services become unnecessary when health and well-being improve. Therefore, it should be possible for older adults to easily try out whether the eHealth service will benefit them. This can be done, for example, by offering trial versions of services, which other research [15,18,84] recommends. Being able to safely try out an eHealth service could help reduce fears, build self-confidence, and better identify the current need for the service. Another helpful option would be to group service functions into easier basic functions and possibly other more complex functions that are recommended to be familiarized with time, as is recognized in the literature [15,47]. However, based on the study findings, the end user's age or life situation does not appear to be a decisive factor in determining the need for user interface content. That said, older adults who are retired may face slightly more challenges in using eHealth services compared to those who are still working.

As also found in other literature [18,70], individual eHealth service should function as part of a wider, multichannel service package so that it adapts to changing life situations and changes in health status. It should work as a channel for end users to receive empathy in challenging health situations. This could be done with effective but sensitive communication possibilities and with empathic content. Because eHealth services should support older adults' health and well-being, it is important to look at and connect their well-being in the IT context [84]. This can be done according to content and with easy and logically functioning user interfaces. In addition, this study's participants wished for an option to not use eHealth services or specific eHealth functions. This need is in line with other research findings [85].

Limitations

This qualitative exploratory study aimed to gain a more in-depth understanding of older adults' needs. The sample was small; the participants were from 1 country, and they were likely to be more interested than average in eHealth services. Older adults are a diverse group from the perspective of, for example, cognitive, motor, or technology use abilities [61]. Therefore, completely covering the population with a representative sample was not possible. This has potentially biased the results. In addition, the COVID-19 pandemic era could have influenced the results, but this was not especially followed as a part of this study.

Regarding the adaptation of eHealth services to different life and health situations, we do not know whether the service should always be the same or differ depending, for example, on whether the health situation has changed, in which case the role and meaning of the service changes from supporting well-being to treating illness. We recommend this as a future research area, especially in long-term studies. More research is needed on the best service channel combinations for older adults to obtain sufficient holistic health and well-being support. In the future, it would be worthwhile to test the design recommendations with representatives of the designer community and older adults, for example, in co-design workshops. In this study, the older adults' needs in different life situations were not compared to those of other age groups. However, this would be valuable to study in the future.

From a methodological perspective, the cultural probe package, including diary-based and sentence completion tasks, combined with remotely conducted semistructured interviews seemed to be effective in collecting information about participants' everyday lives and in mapping user needs for eHealth services. This is interesting because, as far as we know, not many cultural probes studies on this topic are available. The working participants completed more handwriting-based sentences. Therefore, cultural probes including many handwriting exercises might fit better for them. On the other hand, the data were collected using several methods, which confirmed the findings of the study and helped identify user needs. The methods seemed to work well in remote conditions during the COVID-19 pandemic era.

Conclusions

In this qualitative study, we collected information about the everyday lives of older adults in different life situations and identified their related needs for eHealth services. The results helped us understand how older adults see and experience health care with the help of eHealth services as part of everyday life. The older adults in this study often preferred a consistent routine in their daily practices including healthy eating and exercise. eHealth services could offer time saving and flexibility within these routines. A lack of devices or skills was often mentioned as a use barrier. On the basis of the results, four main needs of the older adults for eHealth services were identified: (1) the need for a carefree mind, in which older adults could receive status information on their current health and well-being situation via eHealth services; (2) the need for eHealth service adaptation based on the current life situation, use contexts, and everyday life; (3) the need for a holistic perspective, in which older adults wished to receive support for health issues from many perspectives and through several different service channels; and (4) the need to avoid using eHealth services but to use specific eHealth service functions, in which face-to-face support for their health was especially appreciated. On the basis of the results, eHealth services should be designed such that

XSL•FO
RenderX

they fit well in older adults' everyday lives and adapt to users' everyday practices and health statuses according to content. eHealth services should be accessible and reliable.

Acknowledgments

The authors would especially like to thank the study participants. The authors would also like to thank Marina Broman for completing 3 of the interviews and Marjaana Kanervo, Annetta Kurvinen, Tytti Natunen, Kirsi Pitkälä, Sanna Pussinen, Semra Räikkönen, Reija Saarinen, Sanna Solala, Arja Sunila, and Iina Vartia for the help with the research arrangements and collaboration. The work was supported by the Strategic Research Council at the Academy of Finland (grants 352501 and 352503) and NordForsk (project 100477).

Data Availability

The datasets generated and analyzed during this study are not publicly available due to the sensitive nature of the data, but the numeric data are available from the corresponding author on reasonable request.

Authors' Contributions

The study was conceptualized by PV, SK, and RRH. PV was involved in the methodology, investigation, project administration, data curation, visualization, and formal analysis of the study. KS contributed to the formal analysis and validation of the study. PV, KS, and SK were involved in preparing the original draft of the study. SK contributed to the methodology, resources, data analysis, validation, supervision, and funding acquisition of the study. RRH also contributed to the resource collection of the study. All authors were involved in writing, reviewing, and editing the manuscript.

The visualizations used in the cultural probes were under creative commons licenses. The creators' nick names were OpenClipart-Vectors, Emmie_Norfolk, Buntysmum, Clker-Free-Vector-Images, bibin9363pbr, 1,820,796, ArtsyBee, and b0red.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sentence completion examples.
[PDF File (Adobe PDF File), 191 KB - humanfactors_v12i1e50329_app1.pdf]

Multimedia Appendix 2
Sentence completion task by age group (current situation).
[PDF File (Adobe PDF File), 138 KB - humanfactors_v12i1e50329_app2.pdf]

Multimedia Appendix 3
Remote, semistructured interview (translated from Finnish).
[PDF File (Adobe PDF File), 121 KB - humanfactors_v12i1e50329_app3.pdf]

Multimedia Appendix 4
Interview data analysis code comparison and final codes after comparison.
[PDF File (Adobe PDF File), 166 KB - humanfactors_v12i1e50329_app4.pdf]

Multimedia Appendix 5
Health Confidence Score, eHealth Literacy Scale, and European Health Literacy Survey Questionnaire results by age group.
[PDF File (Adobe PDF File), 128 KB - humanfactors_v12i1e50329_app5.pdf]

Multimedia Appendix 6
Identified themes in life activities.
[PDF File (Adobe PDF File), 80 KB - humanfactors_v12i1e50329_app6.pdf]

Multimedia Appendix 7
Interview themes, example quotes, and user needs.
[PDF File (Adobe PDF File), 208 KB - humanfactors_v12i1e50329_app7.pdf]

XSL•FO
RenderX

Multimedia Appendix 8
Differences in the completed sentences between the 2 groups, focusing on the situation in the future.
[PDF File (Adobe PDF File), 131 KB - humanfactors_v12i1e50329_app8.pdf]

References

1. Czaja SJ, Lee CC. The impact of aging on access to technology. Univ Access Inf Soc 2006 Dec 8;5(4):341-349. [doi: 10.1007/s10209-006-0060-x]
2. Kornadt AE, Kessler EM, Wurm S, Bowen CE, Gabrian M, Klusmann V. Views on ageing: a lifespan perspective. Eur J Ageing 2020 Dec 11;17(4):387-401 [FREE Full text] [doi: 10.1007/s10433-019-00535-9] [Medline: 33380996]
3. Kruse C, Fohn J, Wilson N, Nunez Patlan E, Zipp S, Mileski M. Utilization barriers and medical outcomes commensurate with the use of telehealth among older adults: systematic review. JMIR Med Inform 2020 Aug 12;8(8):e20359 [FREE Full text] [doi: 10.2196/20359] [Medline: 32784177]
4. Hechinger M, Hentschel D, Aumer C, Rester C. A conceptual model of experiences with digital technologies in aging in place: qualitative systematic review and meta-synthesis. JMIR Aging 2022 Sep 09;5(3):e34872 [FREE Full text] [doi: 10.2196/34872] [Medline: 36083625]
5. Ware P, Bartlett SJ, Paré G, Symeonidis I, Tannenbaum C, Bartlett G, et al. Using eHealth technologies: interests, preferences, and concerns of older adults. Interact J Med Res 2017 Mar 23;6(1):e3 [FREE Full text] [doi: 10.2196/ijmr.4447] [Medline: 28336506]
6. Allmér HR. Digital wellness services' servicescape for young elderly. In: Proceedings of the 31th Bled eConference Digital Transformation - Meeting the Challanges. 2018 Presented at: BLED 2018; June 17-23, 2018; Bled, Slovenia URL: https://aisel.aisnet.org/bled2018/32/ [doi: 10.18690/978-961-286-170-4.10]
7. Carlsson C, Walden P. Digital wellness services: key to better quality of life for young elderly authors. In: Proceedings of the 31th Bled eConference Digital Transformation - Meeting the Challanges. 2018 Presented at: BLED 2018; June 17-23, 2018; Bled, Slovenia URL: https://aisel.aisnet.org/bled2018/26/ [doi: 10.18690/978-961-286-170-4.16]
8. Ageing and health. World Health Organization. 2024 Oct 1. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2024-12-01]
9. MacNee W, Rabinovich RA, Choudhury G. Ageing and the border between health and disease. Eur Respir J 2014 Nov;44(5):1332-1352 [FREE Full text] [doi: 10.1183/09031936.00134014] [Medline: 25323246]
10. Steele Gray C, Miller D, Kuluski K, Cott C. Tying eHealth tools to patient needs: exploring the use of eHealth for community-dwelling patients with complex chronic disease and disability. JMIR Res Protoc 2014 Nov 26;3(4):e67 [FREE Full text] [doi: 10.2196/resprot.3500] [Medline: 25428028]
11. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. Lancet 2012 Jul 07;380(9836):37-43 [FREE Full text] [doi: 10.1016/S0140-6736(12)60240-2] [Medline: 22579043]
12. Li Y, Schoufour J, Wang DD, Dhana K, Pan A, Liu X, et al. Healthy lifestyle and life expectancy free of cancer, cardiovascular disease, and type 2 diabetes: prospective cohort study. BMJ 2020 Jan 08;368:l6669 [FREE Full text] [doi: 10.1136/bmj.l6669] [Medline: 31915124]
13. Mehta N, Myrskylä M. The population health benefits of a healthy lifestyle: life expectancy increased and onset of disability delayed. Health Aff (Millwood) 2017 Jul 19:1495-1502 [FREE Full text] [doi: 10.1377/hlthaff.2016.1569] [Medline: 28724530]
14. Chudasama YV, Khunti K, Gillies CL, Dhalwani NN, Davies MJ, Yates T, et al. Healthy lifestyle and life expectancy in people with multimorbidity in the UK Biobank: a longitudinal cohort study. PLoS Med 2020 Sep;17(9):e1003332 [FREE Full text] [doi: 10.1371/journal.pmed.1003332] [Medline: 32960883]
15. Cabrita M, Tabak M, Vollenbroek-Hutten MM. Older adults' attitudes toward ambulatory technology to support monitoring and coaching of healthy behaviors: qualitative study. JMIR Aging 2019 Mar 12;2(1):e10476 [FREE Full text] [doi: 10.2196/10476] [Medline: 31518252]
16. Heponiemi T, Jormanainen V, Leemann L, Manderbacka K, Aalto AM, Hyppönen H. Digital divide in perceived benefits of online health care and social welfare services: national cross-sectional survey study. J Med Internet Res 2020 Jul 07;22(7):e17616 [FREE Full text] [doi: 10.2196/17616] [Medline: 32673218]
17. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021 Aug 17;21(1):1556 [FREE Full text] [doi: 10.1186/s12889-021-11623-w] [Medline: 34399716]
18. Kaihlanen AM, Virtanen L, Buchert U, Safarov N, Valkonen P, Hietapakka L, et al. Towards digital health equity - a qualitative study of the challenges experienced by vulnerable groups in using digital health services in the COVID-19 era. BMC Health Serv Res 2022 Feb 12;22(1):188 [FREE Full text] [doi: 10.1186/s12913-022-07584-4] [Medline: 35151302]
19. Hirvonen N, Enwald H, Känsäkoski H, Eriksson-Backa K, Nguyen H, Huhta AM, et al. Older adults' views on eHealth services: a systematic review of scientific journal articles. Int J Med Inform 2020 Mar;135:104031. [doi: 10.1016/j.ijmedinf.2019.104031] [Medline: 31918340]

XSL•FO
RenderX

20. Grindrod KA, Li M, Gates A. Evaluating user perceptions of mobile medication management applications with older adults: a usability study. JMIR Mhealth Uhealth 2014 Mar 14;2(1):e11 [FREE Full text] [doi: 10.2196/mhealth.3048] [Medline: 25099993]
21. Scheibe M, Reichelt J, Bellmann M, Kirch W. Acceptance factors of mobile apps for diabetes by patients aged 50 or older: a qualitative study. Med 2 0 2015 Mar 02;4(1):e1 [FREE Full text] [doi: 10.2196/med20.3912] [Medline: 25733033]
22. Williams V, Price J, Hardinge M, Tarassenko L, Farmer A. Using a mobile health application to support self-management in COPD: a qualitative study. Br J Gen Pract 2014 Jul;64(624):e392-e400 [FREE Full text] [doi: 10.3399/bjgp14X680473] [Medline: 24982491]
23. Seifert A. The digital exclusion of older adults during the COVID-19 pandemic. J Gerontol Soc Work 2020;63(6-7):674-676. [doi: 10.1080/01634372.2020.1764687] [Medline: 32401181]
24. Seifert A, Hofer M, Rössel J. Older adults' perceived sense of social exclusion from the digital world. Educ Gerontol 2019 Feb 08;44(12):775-785. [doi: 10.1080/03601277.2019.1574415]
25. Rice M, Newell A, Morgan M. Forum Theatre as a requirements gathering methodology in the design of a home telecommunication system for older adults. Behav Inf Technol 2007 Jul;26(4):323-331. [doi: 10.1080/01449290601177045]
26. Mannheim I, Schwartz E, Xi W, Buttigieg SC, McDonnell-Naughton M, Wouters EJ, et al. Inclusion of older adults in the research and design of digital technology. Int J Environ Res Public Health 2019 Oct 02;16(19):3718 [FREE Full text] [doi: 10.3390/ijerph16193718] [Medline: 31581632]
27. Mannheim I, Wouters EJ, Köttl H, van Boekel LC, Brankaert R, van Zaalen Y. Ageism in the discourse and practice of designing digital technology for older persons: a scoping review. Gerontologist 2023 Aug 24;63(7):1188-1200 [FREE Full text] [doi: 10.1093/geront/gnac144] [Medline: 36130318]
28. Mielonen J, Kinnunen UM, Saranto K, Kemppi A, Kuusisto H. The effect of chronic diseases on the use of health technology and digital services in the elderly population in Finland. Stud Health Technol Inform 2020 Nov 23;275:147-151. [doi: 10.3233/SHTI200712] [Medline: 33227758]
29. Schweitzer J, Synowiec C. The economics of eHealth. Results for Development Institute. 2010. URL: https://r4d.org/wp-content/uploads/Economics-of-eHealth.final_.3Nov1010_0.pdf [accessed 2024-12-01]
30. Stephanidis CC, Salvendy G, Antona M, Chen JY, Dong J, Duffy VG, et al. Seven HCI grand challenges. Int J Hum Comput Interact 2019 Jul 01;35(14):1229-1269. [doi: 10.1080/10447318.2019.1619259]
31. Eurostat. Ageing Europe: looking at the lives of older people in the EU: 2020 edition. Publications Office of the European Union. 2020. URL: https://ec.europa.eu/eurostat/web/products-statistical-books/-/ks-02-20-655 [accessed 2023-01-13]
32. Valokivi H, Carlo S, Kvist E, Outila M. Digital ageing in Europe: a comparative analysis of Italian, Finnish and Swedish national policies on eHealth. Ageing Soc 2021 Jul 08;43(4):835-856. [doi: 10.1017/s0144686x21000945]
33. Vines J, Pritchard G, Wright P, Olivier P, Brittain K. An age-old problem: examining the discourses of ageing in HCI and strategies for future research. ACM Trans Comput Hum Interact 2015 Mar 04;22(1):1-27. [doi: 10.1145/2696867]
34. Heponiemi T, Virtanen L, Kaihlanen AM, Kainiemi E, Koponen P, Koskinen S. Use and changes in the use of the internet for obtaining services among older adults during the COVID-19 pandemic: a longitudinal population-based survey study. New Media Soc 2022 Jun 08;26(6):3146-3167. [doi: 10.1177/14614448221097000]
35. European Commission: Directorate-General for Communications Networks, Content and Technology. 2024 digital decade eHealth indicator study: final report. Publications Office of the European Union. 2024. URL: https://op.europa.eu/en/publication-detail/-/publication/c04f6162-3833-11ef-b441-01aa75ed71a1/language-en [accessed 2024-08-13]
36. Kip H, Keizer J, da Silva MC, Beerlage-de Jong N, Köhle N, Kelders SM. Methods for human-centered eHealth development: narrative scoping review. J Med Internet Res 2022 Jan 27;24(1):e31858 [FREE Full text] [doi: 10.2196/31858] [Medline: 35084359]
37. Matthew-Maich N, Harris L, Ploeg J, Markle-Reid M, Valaitis R, Ibrahim S, et al. Designing, implementing, and evaluating mobile health technologies for managing chronic conditions in older adults: a scoping review. JMIR Mhealth Uhealth 2016 Jun 09;4(2):e29 [FREE Full text] [doi: 10.2196/mhealth.5127] [Medline: 27282195]
38. Kujala S, Ammenwerth E, Kolanen H, Ervast M. Applying and extending the FITT framework to identify the challenges and opportunities of successful eHealth services for patient self-management: qualitative interview study. J Med Internet Res 2020 Aug 12;22(8):e17696 [FREE Full text] [doi: 10.2196/17696] [Medline: 32784175]
39. Kayser L, Kushniruk A, Osborne RH, Norgaard O, Turner P. Enhancing the effectiveness of consumer-focused health information technology systems through eHealth literacy: a framework for understanding users' needs. JMIR Hum Factors 2015 May 20;2(1):e9 [FREE Full text] [doi: 10.2196/humanfactors.3696] [Medline: 27025228]
40. Charness N, Boot WR. Aging and information technology use: potential and barriers. Curr Dir Psychol Sci 2009 Oct 01;18(5):253-258. [doi: 10.1111/j.1467-8721.2009.01647.x]
41. Currie M, Philip LJ, Roberts A. Attitudes towards the use and acceptance of eHealth technologies: a case study of older adults living with chronic pain and implications for rural healthcare. BMC Health Serv Res 2015 Apr 16;15:162 [FREE Full text] [doi: 10.1186/s12913-015-0825-0] [Medline: 25888988]
42. Nurkka P, Kujala S, Kemppainen K. Capturing users' perceptions of valuable experience and meaning. J Eng Des 2009 Oct;20(5):449-465. [doi: 10.1080/09544820903158835]

XSL•FO
RenderX

43. Kujala S, Kauppinen M, Rekola S. Bridging the gap between user needs and user requirements. In: Advances in Human-Computer Interaction I. Patras, Greece: Typorama Publications; 2001:45-50.
44. Langefors B. Analysis of user needs. In: Proceedings of the 2nd Conference of the European Cooperation in Informatics. 1978 Presented at: ECI 1978; October 10-12, 1978; Venice, Italy. [doi: 10.1007/3-540-08934-9_71]
45. Peine A, Neven L. The co-constitution of ageing and technology – a model and agenda. Ageing Soc 2020 Jun 03;41(12):2845-2866. [doi: 10.1017/s0144686x20000641]
46. Fronemann N, Peissner M. User experience concept exploration: user needs as a source for innovation. In: Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational. 2014 Presented at: NordiCHI '14; October 26-30, 2014; Helsinki, Finland URL: https://dl.acm.org/doi/10.1145/2639189.2641203 [doi: 10.1145/2639189.2641203]
47. Backonja U, Hall AK, Thielke S. Older adults' current and potential uses of information technologies in a changing world. Int J Aging Hum Dev 2014 Dec 01;80(1):41-63 [FREE Full text] [doi: 10.1177/0091415015591109] [Medline: 26215298]
48. Fallman D. Design-oriented human-computer interaction. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2003 Presented at: CHI '03; April 5-10, 2003; Ft. Lauderdale, FL URL: https://doi.org/10.1145/642611.642652 [doi: 10.1145/642611.642652]
49. Sinha G, Shahi R, Shankar M. Human computer interaction. In: Proceedings of the 3rd International Conference on Emerging Trends in Engineering and Technology. 2010 Presented at: ICETET 2010; November 19-21, 2010; Goa, India. [doi: 10.1109/icetet.2010.85]
50. Soraghan CJ, Hermann S, Boyle G. User-centered design and older people: introductory considerations for engaging users. In: Hippe Z, Kulikowski J, Mroczek T, Wtorek J, editors. Human-Computer Systems Interaction: Backgrounds and Applications 3. Cham, Switzerland: Springer; 2014:221-231.
51. Dickinson A, Goodman J, Syme A, Eisma R, Tiwari L, Mival O, et al. Domesticating technology: in-home requirements gathering with frail older people? In: Proceedings 10th International Conference on Human - Computer Interaction HCI. 2003 Presented at: HCI 2003; June 22-27, 2003; Crete, Greece p. 827-831 URL: https://napier-repository.worktribe.com/preview/267243/mival2.pdf
52. Valkonen P, Karisalmi N, Kaipio J, Kujala S. Remote interviews and visual timelines with patients: lessons learned. Stud Health Technol Inform 2021 May 27;281:845-849. [doi: 10.3233/SHTI210298] [Medline: 34042793]
53. Gaver B, Dunne T, Pacenti E. Design: cultural probes. interactions 1999 Jan;6(1):21-29. [doi: 10.1145/291224.291235]
54. Lee A, Sandvei M, Asmussen HC, Skougaard M, Macdonald J, Zavada J, et al. The development of complex digital health solutions: formative evaluation combining different methodologies. JMIR Res Protoc 2018 Jul 16;7(7):e165 [FREE Full text] [doi: 10.2196/resprot.9521] [Medline: 30012548]
55. Van den Bulck SA, Hermens R, Slegers K, Vandenberghe B, Goderis G, Vankrunkelsven P. Designing a patient portal for patient-centered care: cross-sectional survey. J Med Internet Res 2018 Oct 01;20(10):e269 [FREE Full text] [doi: 10.2196/jmir.9497] [Medline: 30287416]
56. Bosch L, Kanis M, Dunn J, Stewart KA, Kröse B. How is the caregiver doing? Capturing caregivers' experiences with a reflective toolkit. JMIR Ment Health 2019 May 28;6(5):e13688 [FREE Full text] [doi: 10.2196/13688] [Medline: 31140435]
57. Danbjørg DB, Villadsen A, Gill E, Rothmann MJ, Clemensen J. Usage of an exercise app in the care for people with osteoarthritis: user-driven exploratory study. JMIR Mhealth Uhealth 2018 Jan 11;6(1):e11 [FREE Full text] [doi: 10.2196/mhealth.7734] [Medline: 29326092]
58. Kujala S, Nurkka P. Sentence completion for evaluating symbolic meaning. Int J Des 2012;6(3) [FREE Full text]
59. Kujala S, Walsh T, Nurkka P, Crisan M. Sentence completion for understanding users and evaluating user experience. Interact Comput 2013 Jul 12;26(3):238-255. [doi: 10.1093/iwc/iwt036]
60. Lallemand C, Mercier E. Optimizing the use of the sentence completion survey technique in user research: a case study on the experience of e-reading. In: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. 2022 Presented at: CHI '22; April 29-May 5, 2022; New Orleans, LA. [doi: 10.1145/3491102.3517718]
61. Newell A, Arnott J, Carmichael A, Morgan M. Methodologies for involving older adults in the design process. In: Proceedings of the 4th International Conference on Universal Access in Human-Computer Interaction. 2007 Presented at: UAHCI 2007; July 22-27, 2007; Beijing, China. [doi: /doi.org/10.1007/978-3-540-73279-2_110]
62. Chun YJ, Patterson PE. A usability gap between older adults and younger adults on interface design of an internet-based telemedicine system. Work 2012;41 Suppl 1:349-352. [doi: 10.3233/WOR-2012-0180-349] [Medline: 22316747]
63. Boehner K, Vertesi J, Sengers P, Dourish P. How HCI interprets the probes. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2007 Presented at: CHI '07; April 28-May 3, 2007; San Jose, CA p. 1077-1086. [doi: https://doi.org/10.1145/1240624.1240789]
64. Isomursu M, Tähti M, Väinämö S, Kuutti K. Experimental evaluation of five methods for collecting emotions in field settings with mobile applications. Int J Hum Comput Stud 2007 Apr;65(4):404-418. [doi: 10.1016/j.ijhcs.2006.11.007]
65. Ehn M, Eriksson LC, Åkerberg N, Johansson AC. Activity monitors as support for older persons' physical activity in daily life: qualitative study of the users' experiences. JMIR Mhealth Uhealth 2018 Feb 01;6(2):e34 [FREE Full text] [doi: 10.2196/mhealth.8345] [Medline: 29391342]

XSL•FO
RenderX

66. Hong MK, Lakshmi U, Do K, Prahalad S, Olson T, Arriaga RI, et al. Using diaries to probe the illness experiences of adolescent patients and parental caregivers. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. 2020 Presented at: CHI '20; April 25-30, 2020; Honolulu, HI p. 229-299. [doi: https://doi.org/10.1145/3313831.3376426]
67. Klausen SH. Understanding older adults' wellbeing from a philosophical perspective. J Happiness Stud 2019 Oct 31;21(7):2629-2648. [doi: 10.1007/s10902-019-00197-5]
68. Munteanu C, Axtell B, Rafih H, Liaqat A, Aly Y. Designing for older adults: overcoming barriers to a supportive, safe, and healthy retirement. In: Agnew J, Mitchell OS, editors. The Disruptive Impact of FinTech on Retirement Systems. Oxford, UK: Oxford University Press; 2019:104-126.
69. Moser A, Korstjens I. Series: practical guidance to qualitative research. Part 3: sampling, data collection and analysis. Eur J Gen Pract 2018 Dec;24(1):9-18 [FREE Full text] [doi: 10.1080/13814788.2017.1375091] [Medline: 29199486]
70. Valkonen P, Kujala S, Hörhammer I, Savolainen K, Helminen RR, Vartia I. Health self-management of older employees: identifying critical peak experiences of a patient portal. FinJeHeW 2023 Apr 20;15(2):114-129. [doi: 10.23996/fjhw.126837]
71. Benson T, Potts HW, Bark P, Bowman C. Development and initial testing of a Health Confidence Score (HCS). BMJ Open Qual 2019;8(2):e000411 [FREE Full text] [doi: 10.1136/bmjoq-2018-000411] [Medline: 31259277]
72. Norman CD, Skinner HA. eHEALS: The eHealth Literacy Scale. J Med Internet Res 2006 Nov 14;8(4):e27 [FREE Full text] [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
73. Eronen J, Paakkari L, Portegijs E, Saajanaho M, Rantanen T. Assessment of health literacy among older Finns. Aging Clin Exp Res 2019 Apr;31(4):549-556 [FREE Full text] [doi: 10.1007/s40520-018-1104-9] [Medline: 30578457]
74. Çerçi S, Cecchinato ME, Vines J. How design researchers interpret probes: understanding the critical intentions of a designerly approach to research. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. 2021 Presented at: CHI '21; May 8-13, 2021; Yokohama, Japan. [doi: 10.1145/3411764.3445328]
75. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
76. Mayring P. Qualitative content analysis: theoretical foundation, basic procedures and software solution. Social Science Open Access Repository. 2014. URL: https://www.ssoar.info/ssoar/handle/document/39517 [accessed 2024-12-15]
77. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
78. Maher C, Hadfield M, Hutchings M, de Eyto A. Ensuring rigor in qualitative data analysis: a design research approach to coding combining NVivo with traditional material methods. Int J Qual Methods 2018 Jul 10;17(1):1-13. [doi: 10.1177/1609406918786362]
79. Nielsen BK, Nielsen LA, Lasgaard M. Loneliness in emerging adulthood: a qualitative study using cultural probes and in-depth interviews. Curr Res Behav Sci 2024;7:100154. [doi: 10.1016/j.crbeha.2024.100154]
80. Kujala S, Miron-Shatz T. Emotions, experiences and usability in real-life mobile phone use. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2013 Presented at: CHI '13; April 27-May 2, 2013; Paris, France p. 1061-1070 URL: https://doi.org/10.1145/2470654.2466135 [doi: 10.1145/2470654.2466135]
81. Reiners F, Sturm J, Bouw LJ, Wouters EJ. Sociodemographic factors influencing the use of eHealth in people with chronic diseases. Int J Environ Res Public Health 2019 Feb 21;16(4):645 [FREE Full text] [doi: 10.3390/ijerph16040645] [Medline: 30795623]
82. Jarosz E. What makes life enjoyable at an older age? Experiential wellbeing, daily activities, and satisfaction with life in general. Aging Ment Health 2022 Jun;26(6):1242-1252. [doi: 10.1080/13607863.2021.1916879] [Medline: 33908290]
83. Ryu J, Heo J, Ellis G, Widmer MA. Leisure, eudaimonic behavior, physical activity and well-being of older adults. J Leis Res 2022 Jun 08;53(4):595-614. [doi: 10.1080/00222216.2022.2051115]
84. Zhang J, Umemuro H. Older adults' hedonic and eudaimonic well-being in information and communication technology activities. Gerontechnology 2012;10(4):231-243. [doi: 10.4017/gt.2012.10.4.004.00]
85. Heponiemi T, Kainiemi E, Virtanen L, Saukkonen P, Sainio P, Koponen P, et al. Predicting internet use and digital competence among older adults using performance tests of visual, physical, and cognitive functioning: longitudinal population-based study. J Med Internet Res 2023 May 05;25:e42287 [FREE Full text] [doi: 10.2196/42287] [Medline: 37145836]

Abbreviations

DigiIN: Toward a Socially Inclusive Digital Society: Transforming Service Culture Project
eHEALS: eHealth Literacy Scale
HCS: Health Confidence Score
HLS-EU-Q16: European Health Literacy Survey Questionnaire

XSL•FO
RenderX

Edited by C Jacob; submitted 28.06.23; peer-reviewed by L Heryawan, I Mannheim, S Jalil, L Simonse; comments to author 07.12.23; revised version received 01.02.24; accepted 11.11.24; published 08.01.25.

Please cite as:
Valkonen P, Kujala S, Savolainen K, Helminen RR
Exploring Older Adults' Needs for a Healthy Life and eHealth: Qualitative Interview Study
JMIR Hum Factors 2025;12:e50329
URL: https://humanfactors.jmir.org/2025/1/e50329
doi:10.2196/50329
PMID:

©Paula Valkonen, Sari Kujala, Kaisa Savolainen, Riina-Riitta Helminen. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 08.01.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Exploring the Impact of Digital Peer Support Services on Meeting Unmet Needs Within an Employee Assistance Program: Retrospective Cohort Study

Harpreet Nagra[1], PhD; Robert A Mines[2], PhD; Zara Dana[1], PhD

[1]Supportiv, Berkeley, CA, United States

[2]MINES & Associates, Littleton, CO, United States

Corresponding Author:
Harpreet Nagra, PhD
Supportiv
2222 Harold Way
Berkeley, CA, 94704
United States
Phone: 1 800 845 0015
Email: harpreet@supportiv.com

Abstract

Background: The World Health Organization estimates that 1 in 4 people worldwide will experience a mental disorder in their lifetime, highlighting the need for accessible support.

Objective: This study evaluates the integration of digital peer support (DPS) into an employee assistance program (EAP), testing 3 hypotheses: (1) DPS may be associated with changes in EAP counseling utilization within a 5-session model; (2) DPS users experience reduced sadness, loneliness, and stress; and (3) DPS integration generates a positive social return on investment (SROI).

Methods: The study analyzed EAP utilization within a 5-session model using pre-post analysis, sentiment changes during DPS chats via natural language processing models, and SROI outcomes.

Results: Among 587 DPS chats, 432 (73.6%) occurred after business hours, emphasizing the importance of 24/7 availability. A matched cohort analysis (n=72) showed that DPS reduced therapy sessions by 2.07 per participant (P<.001; Cohen d=1.77). Users' messages were evaluated for sentiments of sadness, loneliness, and stress on a 1-10 scale. Significant reductions were observed: loneliness decreased by 55.04% (6.91 to 3.11), sadness by 57.5% (6.84 to 2.91), and stress by 56.57% (6.78 to 2.95). SROI analysis demonstrated value-to-investment ratios of US $1.66 (loneliness), US $2.50 (stress), and US $2.58 (sadness) per dollar invested.

Conclusions: Integrating DPS into EAPs provides significant benefits, including increased access, improved emotional outcomes, and a high SROI, reinforcing its value within emotional health support ecosystems.

(JMIR Hum Factors 2025;12:e68221) doi:10.2196/68221

KEYWORDS

digital peer support; peer support; EAPs; cost-effectiveness; SROI

Introduction

Background

The World Health Organization estimates that 1 in 4 people globally will experience mental disorders in their lifetime [1], stressing the need for a readily available treatment ecosystem. Relatedly, the increasing popularity of remote work post–COVID-19, along with work- and workplace-related stress—stemming from long hours, high workloads, an inability to disconnect, and, in particular, isolation—has contributed to a rise in mental health problems such as anxiety and depression, musculoskeletal pain, sleep deprivation, and stress [2]. These stressors impact not only employees' well-being but also presenteeism, absenteeism, and work-life balance [2-4]. Burnout, in particular, arises from chronic stress characterized by emotional, physical, and mental exhaustion, leading to reduced productivity, increased absenteeism, and higher turnover rates [3].

XSL•FO
RenderX

While the Centers for Disease Control and Prevention's National Center for Health Statistics reports a notable increase in treatment-seeking behaviors between 2019 (19.2%) and 2021 (21.6%), a significant majority of working adults may still suffer alone [5]. The proportion of working adults aged 18-44 years receiving mental health treatment grew significantly from 18.5% in 2019 to 23.2% in 2021 [5], with women (23.8% to 28.6%, respectively) continuing to outpace men (13.1% to 17.8%, respectively) [5]. By 2021, adults aged 45-64 years (21.2%) and those aged 65 years and over (18.9%) ranked second and third, respectively, in seeking treatment [5].

Employee Assistance Programs

Mental and emotional health–focused employee assistance programs (EAPs), which are confidential, voluntary, organization-sponsored, and low-cost services for employees, offer personal or work-related counseling and support. The purpose of EAPs is to help employees manage stress, enhance problem-solving abilities, and, in turn, improve work productivity, reduce absenteeism, and promote overall well-being [6]. Counseling-based EAPs provide psychoeducation, teach coping strategies, and assist users in setting achievable personal goals that build confidence and self-efficacy [7-12]. Additionally, EAPs offer training in communication skills, such as assertiveness, to improve relationships and help users navigate challenges effectively. Patient activation, which refers to an individual's knowledge, skills, confidence, and willingness to manage their health and well-being [7,8], is further supported by connecting users to resources and support networks, including health care providers and peer support groups. Finally, EAPs empower users with self-management tools, such as digital apps and resource libraries, that allow them to track their progress and independently manage their well-being [12]. Together, these strategies foster self-reliance, enhancing users' ability to manage their mental health and overall well-being [7]. Depending on the EAP, a significant return on investment (ROI) may be observed, with some organizations seeing US $3-US $10 for every dollar spent [13], largely due to positive impacts on absenteeism, productivity, and employee turnover and retention [6,13]. EAPs are well-established assets to organizations, delivering emotional benefits to employees and productivity and financial benefits to employers [6,13].

Digital Peer Support

While counseling-based EAPs are known for their advantageous returns, not every emotional concern requires counseling support. A tiered support system for individuals who are mentally well but experiencing lower-acuity emotional distress may be necessary. Simultaneously, the US health care system currently faces a shortage of mental health professionals, limiting access to adequate support [14,15]. In 2021, 5930 areas were designated as health professional shortage areas, leaving 129.6 million people without access to affordable and accessible care [15]. President Biden's 2022 White House brief highlighted this issue and called for efforts to "strengthen system capacity," including the increased use of peer support specialists and paraprofessionals [16].

Peer support interventions involve individuals with similar lived experiences offering mutual support [17]. While expert-based professional support may evoke feelings of shame, failure, and mental health stigma, peer support can normalize emotional struggles and foster connection, validation, and empathy through shared experiences [18]. This "mutual empowerment" from peer-driven interventions grants individuals greater autonomy and competency [18]. However, some Americans may not have access to digital peer services due to a lack of universal internet access and low health literacy. Approximately 24 million people in the United States lack reliable broadband service or high-speed internet [19]. Historically, lower rates of high-speed internet use have been observed in households where the owner is 65 years or older, Hispanic, African American, American Indian, or Alaska Native [19]. Furthermore, in 2021, only about 5 in 10 households with incomes below US $25K used high-speed internet [19]. Additionally, approximately 90 million American adults struggle to understand and navigate complex health- and text-based tasks accurately and consistently, exhibiting low health literacy [20]. By addressing these disparities, digital peer support (DPS) services could become even more inclusive by expanding support to those currently excluded from traditional services.

Self-determination theory, as articulated by Ryan and Deci [21], offers insight into key features that make peer support impactful. The theory states that self-determination and emotional growth are driven by 3 core elements—autonomy, competence, and relatedness [21,22]. Autonomy refers to the ability to self-direct and control one's actions, competence involves mastering new skills, and relatedness is about forming meaningful connections with others [21,22]. Peer support provides a nonjudgmental, nonstigmatizing space that fosters autonomy, as there is no professional expert directing treatment. Competence is strengthened through the normalization of mistakes shared via lived experiences, as well as the application and practice of coping strategies. Relatedness is reinforced through the verbal and emotional connections formed with similar peers [18]. Research has consistently shown that relatedness in supportive relationships contributes significantly to emotional regulation, motivation, and the overall success of interventions. By reinforcing relatedness through peer-to-peer interactions, peer support programs address a fundamental psychological need that complements autonomy and competence, further validating their essential role in the model [17,18]. The interrelatedness of these dynamic factors is believed to enhance mental health and well-being [17-22]. The introduction of DPS platforms extends these benefits by offering accessible, mutually supportive communities to help alleviate feelings of loneliness and isolation [23]. Numerous studies have demonstrated the effectiveness of these digital interventions in addressing mental health needs [23-30].

Confirmed benefits of peer support interventions are evident through self-determination theory and digital platforms [23-30]; however, a significant gap remains in understanding the ROI, particularly when integrated into counseling-based EAPs. Studies suggest that peer support may lead to direct health care savings by equipping individuals with coping mechanisms and providing emotional support, which, in turn, reduces the risk

XSL•FO
RenderX

of crises and costly subsequent interventions (eg, inpatient hospitalization, emergency room use) [31-33]. Specifically, current research does not adequately address how incorporating 24/7 DPS into EAPs translates into measurable cost-effectiveness for organizations. Traditionally, cost-effectiveness and cost-benefit analyses are used to assess the value for money of health and social interventions, but the value derived from participating in 24/7 DPS can be subtle and difficult to measure [34]. As a result, there is a paucity of research on the broader social, economic, and environmental value. This gap underscores the need for a focused study to evaluate the social return on investment (SROI) of integrating 24/7 DPS into a counseling-based EAP, assessing both the economic impact and the efficiency of these interventions.

Social Return On Investment

SROI has become a recognized method for measuring the impact, outcomes, and value created by social-emotional–focused organizations [35,36]. Broadly, SROI uses a mixed-method design to assess the value of an intervention relative to the cost of enabling it. Beyond financial metrics, which may assign limited value to a program, SROI captures social, environmental, and economic elements to generate a comprehensive "social value."

Study Aim and Hypotheses

For this study, a 24/7 US-based DPS service was integrated with a US-based counseling-based EAP firm to explore potential changes in utilization, sentiment, and the SROI of the 24/7 DPS for the EAP's clientele. The following study hypotheses were examined: (1) The introduction of the 24/7 DPS service may lead to changes in the utilization of EAP counseling services within a 5-session model; (2) participants who utilize the DPS service will experience significant changes in sentiment (eg, reduced sadness, loneliness, and stress) over the course of their engagement with the service; and (3) the integration of DPS into the EAP will produce a positive SROI, reflecting the added value of peer support services for EAP clientele. Through these key objectives, the study aims to assess the effects of DPS on EAP utilization, emotional well-being, and the social and economic value added to the EAP.

Methods

Pre- and Post-DPS Integration: EAP Services and Implementation

Using aggregated participant data, pre- and postanalyses of adding DPS to EAP services were conducted. Pre-DPS was defined as the routine care provided within the EAP, including an initial intake, screening for clinical concerns, and offering service options such as counseling, legal resources, psychoeducation, and coaching. Before the introduction of DPS, participants either accessed the EAP provider's website using a company code to explore available services or completed an initial online or phone intake. They were screened for clinical concerns, including substance use and suicidal or homicidal ideation, by licensed clinicians or clinically supervised paraprofessionals. The EAP offers a comprehensive suite of behavioral health and wellness services designed to support mental health, work-life balance, and overall well-being. These offerings include free and confidential counseling, legal and financial benefits (including Medicare and Social Security consultation), wellness coaching, work-life referral services, and an online resource library. Additional services include virtual reality programs, as well as parenting, life, career, and work performance coaching. The EAP provides managed care 24/7, with crisis support delivered by licensed mental health professionals.

An anonymous US-based DPS service that provided 24/7/365 moderated, synchronous peer group chats was integrated into the US-based EAP's client support ecosystem. The DPS service utilized artificial intelligence (AI)–driven natural language processing to match users with peers facing similar issues in small groups, each facilitated by a trained human moderator who ensured a safe environment and directed users to professional services in crises. The DPS model allows for unlimited sessions, in both duration and frequency, while the EAP's 5-session model was selected for further analysis due to the prevalence of 5-session models over other session formats.

Participants

DPS: Peer Support Users

Nonidentifying participant data from the DPS provider, collected between June 2023 and May 2024, were used to evaluate the research objectives. DPS service participants had access to unlimited peer support chats and were included in the data analysis if they utilized the chat service at least once.

EAP: Counseling Service Participants

Inclusion Criteria

Pre-DPS (June 2022 to May 2023) and post-DPS (June 2023 to May 2024) participants were selected based on having active access to the 5-session EAP model. For the analysis of total utilization, no demographic factors were matched between the pre- and post-DPS groups. However, for the utilization analysis of the 5-session EAP model and to ensure comparable groups, participants in both the pre-DPS and post-DPS cohorts were matched based on demographic and emotional concern variables. Matching was conducted using propensity score matching, adjusting for potential confounders such as age, gender, and emotional concerns (eg, anxiety or depression). Participants who had utilized the full 5-session model in both groups were selected, with efforts made to capture diverse client trajectories by including multiple participants matched on these same variables. EAP participants were employed across a wide range of sectors, including health care, education, professional services, manufacturing, construction, building trades, nonprofits, technology, service industries (eg, restaurants), unions, mining, government, and others.

Cohort Exclusion Criteria

Participants were excluded from the analysis if they had missing demographic data (age or gender) or if they exceeded the authorized number of counseling sessions (eg, >5 sessions in the pre-DPS group). These exclusions were necessary to control for confounding variables, as participants requiring more intensive mental health support (beyond 5 sessions) may have

XSL•FO
RenderX

had significantly different needs compared with those who adhered to the 5-session model. This approach ensured that the study focused on participants whose engagement with EAP services reflected typical utilization patterns.

Peer Support Moderators

Peer support moderators (PSMs) are human moderators trained in digital safety monitoring who provide real-time, text-based support. PSMs receive up to 164 hours of training within the first 90 days of hire, focusing on engaging in digital conversations, facilitating synchronous group chats, and managing the psychological safety of chat users by monitoring for and removing trolls, as well as enacting safety protocols, such as referring users who may be in active crisis to appropriate services. Additionally, PSMs receive 36 hours of ongoing asynchronous and synchronous clinical supervision and consultation per quarter.

Procedures

Overview

This retrospective study evaluated the differences in the EAP's services before and after DPS to assess the programmatic impact and added social value.

Data Variables

The primary data sources for this study were aggregated engagement variables, such as chat time, number of chats, age, gender, user struggles or presenting concerns, and national-level cost-of-service value data.

Data Extraction and Management

The data extraction process involved retrieving DPS chat data, which included digital chat records, user interactions with the platform, and usage patterns. All collected data were deidentified and anonymized to ensure participant privacy and confidentiality. Data management procedures adhered to institutional and regulatory guidelines, ensuring data security and integrity, with access restricted to authorized personnel. Measures were in place to maintain data security, integrity, and confidentiality throughout the analysis process.

Statistical Analysis

Analysis and Group Comparability

Descriptive statistics were used to summarize participant characteristics and their engagement with DPS and EAP services. To assess differences in service utilization and sentiment scores between pre-DPS and post-DPS groups, paired *t* tests (2-tailed) were conducted for continuous data (eg, session frequency), and chi-square tests were used for categorical data (eg, gender, emotional concern). As DPS is anonymous, it was not possible to directly track which EAP participants used the service. Using volunteered age and gender data from DPS participants as a starting point, propensity score matching was applied to control for potential confounders such as age, gender, and emotional concern, ensuring group comparability before DPS integration.

Mood Categorization

Clinical counseling experts reviewed the emotional concerns reported by DPS participants, categorizing them into mood categories such as depression, anxiety, adjustment, and relational concerns to align DPS participants with the presenting emotional concerns of EAP participants. The counseling experts were 2 licensed psychologists with doctoral degrees in counseling psychology and a combined 50 years of experience in the field. The clinical interpretation process, guided by the Biopsychosocial Model [37], involved categorizing participants' emotional struggles into mood categories. This model integrates biological, psychological, and social dimensions to provide a comprehensive understanding of mental health. For example, participants' reports of feeling "trapped" or "isolated" were examined through a psychological lens to identify depressive thought patterns, while work-related stress (eg, "I hate my boss") was contextualized within the social environment to classify relational concerns [38]. By incorporating these factors, the model enabled clinical experts to holistically interpret participants' struggles, ensuring that the categorizations captured the complexity of their emotional experiences. The categorization process involved initially reviewing the raw data from participants' reported "struggles" and presenting concerns in therapy individually, followed by cross-validation between the 2 experts to ensure consistency. This collaborative step helped reconcile differences in categorization and address overlapping cases. It is important to note that neither the text-based struggle data from DPS nor the presenting concerns from EAP represented formal diagnoses; rather, they reflected participants' personal interpretations of their symptoms. Keywords such as *helpless* or *sadness* were flagged for depression-related presenting concerns, while terms related to *fear, excessive worry,* or *panic* were matched with anxiety-focused concerns. In cases where symptoms overlapped (eg, depression and anxiety), the experts aligned their categorizations based on the primary presenting concern reflected in the participant's statement. For example, if a participant's statement indicated feelings of hopelessness and despair alongside anxious thoughts, the struggle was categorized as depression. Alternatively, if worry or nervousness predominated, it was categorized as anxiety. In scenarios where a specific relationship was mentioned alongside anxious or depressive feelings, the struggle was categorized primarily as a relational concern. Symptom overlap was resolved by categorizing the statement according to the dominant theme present in the participant's experience. Table 1 provides examples of these interpretations.

XSL•FO
RenderX

Table 1. Clinical interpretation of comparable emotional concerns shared on digital peer support service with employee assistance program service.

Participant's struggle	Presenting concern[a]
"I am feeling pretty suicidal and trapped. I am new to my work from home job and I need help, resources, a plan something or this will only get worse. I am trapped with no transportation, no friends or any meaningful relationships in life. Just a work schedule and I'm isolated."	Depression
"I hate my job and my boss"	Relational concern
"I've been feeling a lot of anxiety around work and my personal life"	Anxiety/stress
"Alcohol addiction"	Substance use

[a]These are not diagnoses, they are layperson understandings.

Sentiment Analysis

Peer-to-peer conversations were analyzed to produce a quantitative measure of emotional change. Each participant's chat session, referred to as a "user-chat-session" or simply "user-chat," was treated as an independent unit of analysis, regardless of whether the same participant took part in multiple sessions. As the DPS service is anonymous, each user and their experience in a given chat session were considered separately, without combining data from different sessions for the same user. This approach ensures that each user's experience is analyzed in isolation, providing a more accurate, user-centric view of emotional changes.

Each user's message was analyzed in the context of preceding messages from other users to capture their emotional journey throughout the chat session. Context was defined as the concatenation of all preceding messages from other users in the same chat, starting from the latest message of the user of interest. A user's initial struggle and first message had no prior context and were evaluated with an empty context. For example, in the following conversation:

userA: message1

userB: message2

moderator: message3

userA: message4

userB: message5

This example assumes that message1 and message2 are the first messages of user A and user B, respectively, in the chat session. User A's emotional journey was quantified by evaluating message1 with an empty context and message4 in the context of message2 + message3. Similarly, for user B, message2 was evaluated without context, and message5 was evaluated in the context of message3 + message4.

Each user's message, along with its corresponding context, was evaluated for sentiments of sadness, loneliness, and stress using a scale of 1-10, where 1 indicates low intensity, 5 indicates moderate intensity, and 10 indicates high intensity of the sentiment. The analysis focused on chat sessions where participants exhibited at least moderate sentiment intensity (score≥5) from the beginning.

A third-party, public natural language processing model from OpenAI, called GPT-4, was used to assign sentiment scores to each user message and its preceding context [39]. A few-shot learning approach was applied, where the model was provided with 10 examples representing different sentiment intensities to calibrate its interpretations [40]. GPT-4 demonstrates an advanced capacity to interpret emotion, surpassing benchmarks from the general population and offering highly congruent performance in emotion interpretation [41]. See Table 2 for examples provided to GPT-4 to predict loneliness levels.

Sentiment analysis evaluated each user's chat messages in temporal order. The quantified emotions were then interpolated and extrapolated to generate a collective sentiment trend across peer-to-peer chats of varying lengths. A one-sided Mann-Whitney *U* test was used to compare emotion scores at the start and end of the conversations.

Table 2. Examples provided to GPT-4 to interpret loneliness intensities in a few-shot manner.

Intensities	Loneliness
1=low	"Take care everyone. Thanks for the great talk"
5=moderate	"I went on Monday and today. After today's class is when I felt worse as I chatted with a lady in class who told me that she got really depressed when she retired."
10=high	"I just always feel totally alone. Like no one understands what is going on in my head"

SROI Methodology

Historically, the SROI methodology has relied on qualitative methods to identify outcome themes most relevant to participants, followed by quantitative approaches to determine monetary proxies for the identified outcomes and related values [41].

With the advent of natural language processing models, significantly larger volumes of text-based data can now be analyzed, allowing qualitative data to be graphically represented. This study utilized text-based data from peer group participants—including open-ended questions, affirmations, reflective questions, and summaries—to examine changes in sentiments such as sadness, loneliness, and stress within

XSL•FO
RenderX

peer-support chats [42]. The SROI methodology was applied to assess the social value created by integrating DPS into the EAP. Qualitative data from peer-to-peer conversations, including open-ended questions, affirmations, and summaries, were analyzed to understand sentiment changes and emotional shifts within the peer group chats. Quantitative methods were used to calculate a monetary proxy for the emotional outcomes. The Sopact SROI calculator, available on the Sopact website [41], was utilized to determine the social value generated by the DPS service, incorporating factors such as enhanced emotional well-being and cost savings from reduced therapy usage. Given the novelty of DPS and its integration into the EAP, a small sample size was anticipated. This limitation was addressed through rigorous matching techniques and statistical analyses to ensure the findings remained valid and robust despite the smaller sample size. Moreover, it is acknowledged that some participants may have preferred individual counseling services over DPS, potentially introducing selection bias among those who opted for the DPS model.

Ethical Considerations

Informed consent was obtained from all participants before their involvement in the study. The consent process included clear explanations of the research's purpose, data collection methods, and the measures implemented to ensure participant privacy and confidentiality. All data collected were deidentified to safeguard participants' anonymity. The study complied with all institutional and regulatory guidelines related to data protection and participant privacy. To ensure participant safety within the DPS, real-time moderation was provided by trained human moderators specializing in digital safety monitoring. Furthermore, the Pearl Institutional Review Board reviewed and approved the study (approval ID 2024-0442).

Results

Total DPS Utilization

- Study hypothesis 1: The introduction of the 24/7 DPS service may be associated with utilization changes of EAP counseling services within a 5-session model, specifically among participants who engage with DPS.

Out of 587 total chats, 432 took place between 5 PM and 8 AM, hours when most therapy offices are typically closed. Chats lasting less than 5 minutes were excluded from the analysis, as the initial 5 minutes are generally used for introductions and setting expectations [30]. Consequently, 107 chats were omitted. A summary of the data is provided in Table 3.

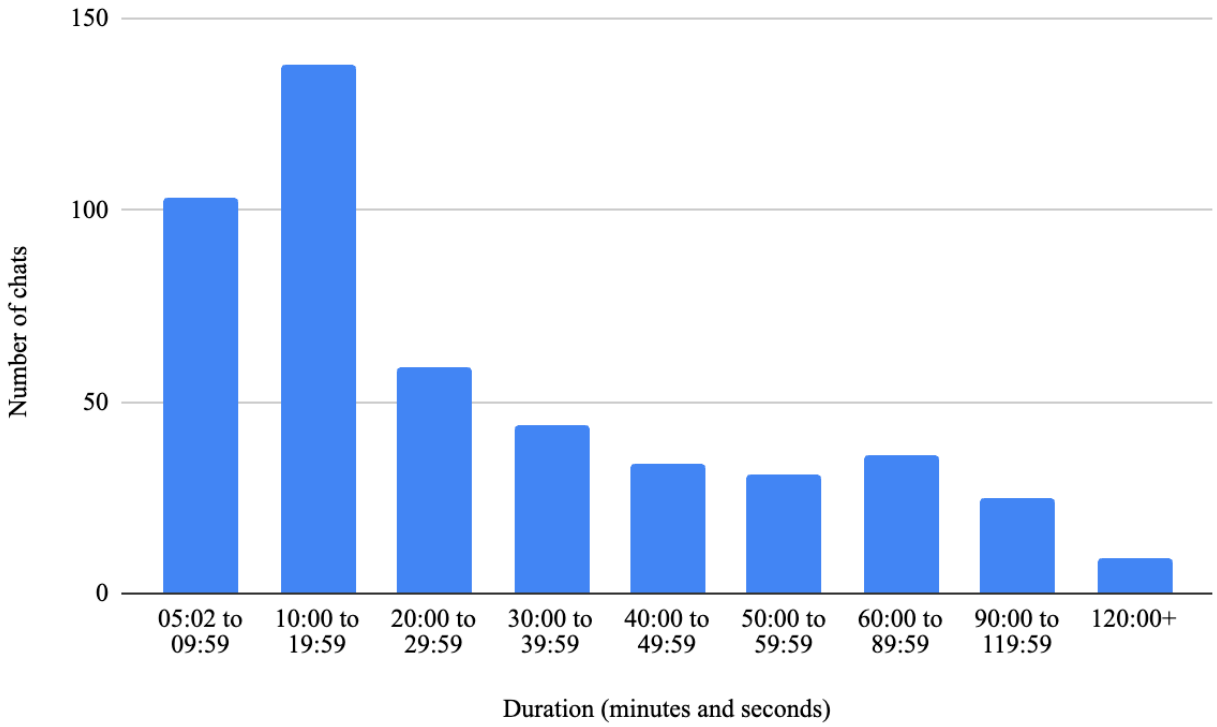
When categorized by duration, the data revealed that 103 chats lasted between 0 and 9 minutes 59 seconds, 138 chats between 10 and 19 minutes and 59 seconds, and 59 chats between 20 and 29 minutes and 59 seconds. Additionally, 44 chats lasted between 30 and 39 minutes and 59 seconds, 34 chats between 40 and 49 minutes and 59 seconds, and 31 chats between 50 and 59 minutes and 59 seconds. While longer chats were less common, they occurred consistently: 36 chats lasted between 60 and 89 minutes and 59 seconds, 25 chats between 90 and 119 minutes and 59 seconds, and 9 chats exceeded 120 minutes. Refer to Figure 1 for further details.

Table 3. Peer support program utilization data.

Metric	Value
Total chats, n	480
Total messages, n	7219
Total duration, hours:minutes:seconds	254:17:57
Average duration, hours:minutes:seconds	0:31:51
Shortest duration, hours:minutes:seconds	0:05:02
Longest duration, hours:minutes:seconds	3:19:11

XSL•FO
RenderX

Figure 1. Chat duration distribution.



Pre-Post Analysis

Matched Cohort Analysis

To evaluate the utilization impact of DPS on the EAP's 5-session model, semimatched cohorts were created for the periods before (June 2022 to May 2023) and after DPS implementation (June 2023 to May 2024). Because of the anonymous nature of the DPS service, only participants who voluntarily provided their age and gender were included in the matching process for analysis in the EAP's pre- and post-DPS cohorts. The matching variables comprised age, gender, and presenting emotional concerns.

Sample

Data on presenting emotional concerns were readily available based on participants' responses to the question, "What's your struggle?"—the sole query asked before matching them with the DPS service. A total of 45 unique users provided both their age and gender information voluntarily. These users ranged in age from 18 to 51 years, with a mean age of 34.97 (SD 9.71) years. Gender distribution was as follows: 18 (40%) males, 26 (58%) females, and 1 (2%) nonbinary individual.

The availability of participant age, gender, and emotional concern data from the DPS service enabled the creation of a matched pre-DPS EAP cohort. The frequency of emotional concerns in the peer support program sample is presented in Table 4.

Table 4. Emotional concern frequency in the digital peer support sample (n=45).

Emotional concern type	Value, *n* (%)
Adjustment	1 (2)
Depression	18 (40)
Anxiety	8 (18)
Posttraumatic stress disorder	2 (4)
Relational concern	14 (31)
Substance use	2 (4)

Statistical Analysis

Using age and gender data from the DPS service participant pool, a total sample of 72 EAP participants—36 pre-DPS and 36 post-DPS—was analyzed for differences. The ages of EAP participants ranged from 18 to 51 years, with a mean age of 33.86 (SD 8.18) years. Females comprised the majority of participants (n=68, 94%), while males accounted for a smaller proportion (n=4, 6%) in both the pre- and post-DPS cohorts.

To assess the impact of the DPS addition, a paired samples *t* test (2-tailed) was performed, comparing the number of therapy

XSL•FO
RenderX

sessions utilized by participants pre- and post-DPS within the matched cohort. The analysis revealed a mean difference of –2.07 (SD 1.77) sessions (t_{71}=9.92; P<.001), indicating a significant reduction in the number of therapy sessions used after the intervention. These findings suggest that the intervention may have been associated with a decrease in therapy session utilization among the sampled participants.

Effect Size

In addition to the *t* test, Cohen *d* was calculated to determine the effect size of the intervention. Based on Cohen guidelines, effect sizes are classified as small (0.2), medium (0.5), and large (0.8). The effect size in this study was 1.77, which is considered large. This suggests that the DPS service had a substantial impact on reducing the number of therapy sessions utilized by participants.

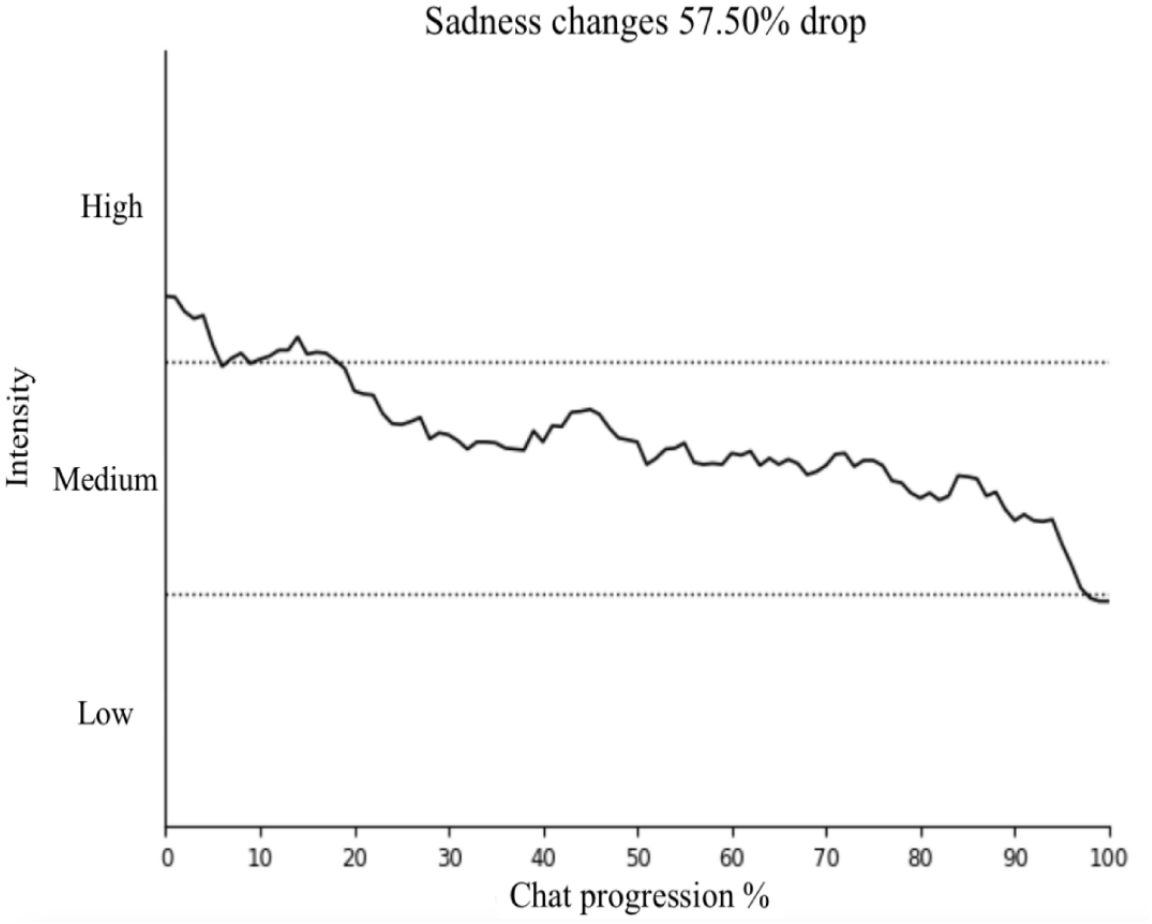
The CI for the mean difference in therapy session utilization was calculated to be between –1.47 and –0.87 at the 95% CI level. This range further supports the reliability of the observed reduction in session utilization, providing evidence that the DPS service may have effectively decreased the number of therapy sessions used by participants in the small sample group.

- Study hypothesis 2: Participants who utilize the DPS service will experience significant changes in sentiment (eg, reduced sadness, loneliness, and stress) throughout their engagement with the service.

Sentiment Change

From the DPS service's sampled cohort (n=45), sentiment changes for sadness, loneliness, and stress were investigated for each user chat. The results are presented in Figures 2-4, showing the changes in sentiment for these 3 emotions. All 3 emotions demonstrated statistically significant reductions (sadness, loneliness, and stress: P<.001) in sentiment intensity during the chat sessions. The sentiment intensity levels are categorized as follows: 1-4 as "low," 4-7 as "moderate," and 7-10 as "high."

Figure 2. Significant (P<.001) changes in sadness sentiment during chats.



XSL•FO
RenderX

Figure 3. Significant (*P*<.001) changes in loneliness sentiment during chats.

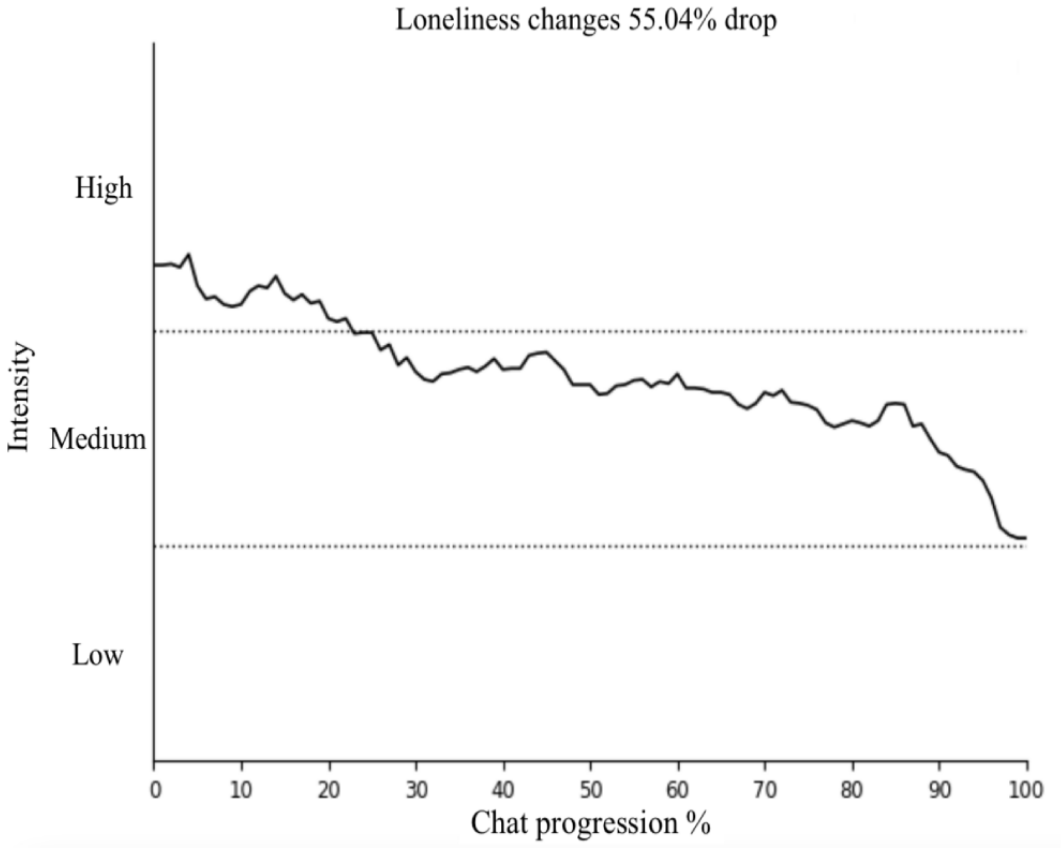
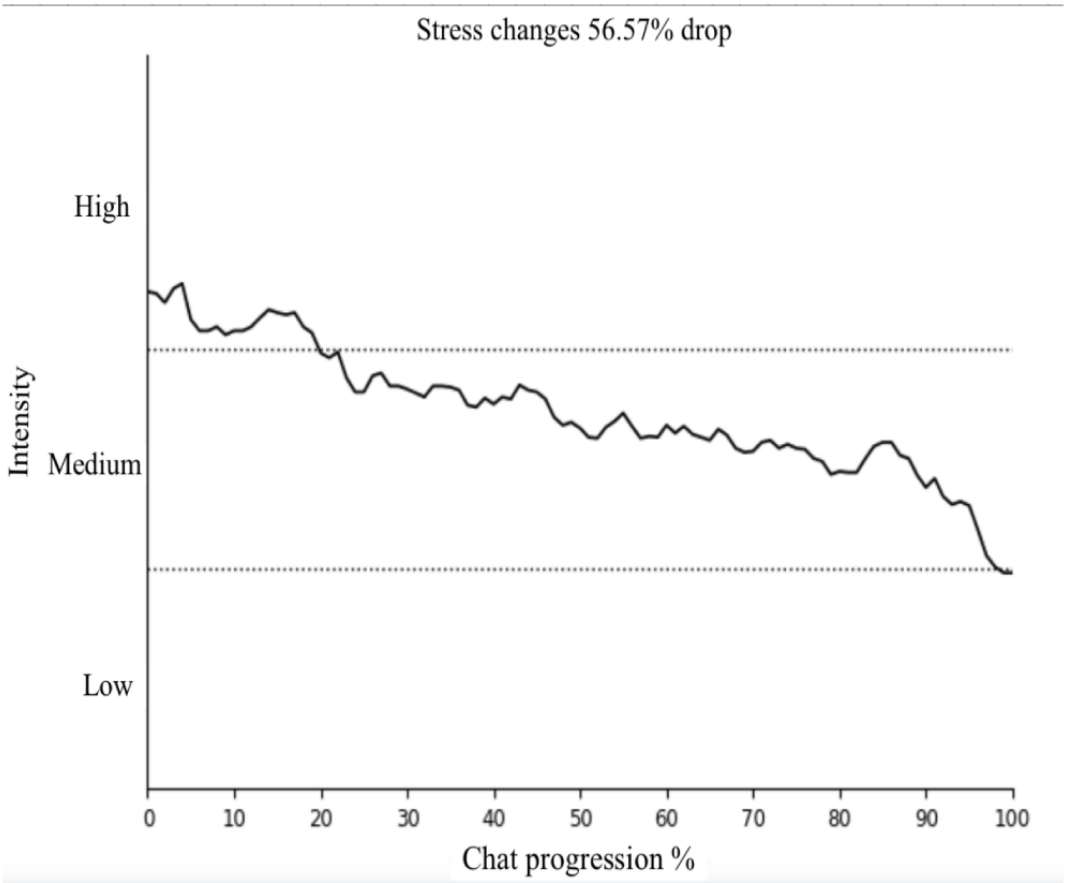


Figure 4. Significant (*P*<.001) changes in stress sentiment during chats.



XSL•FO
RenderX

Social Return on Investment Analysis

SROI assigns financial proxies to quantify the value of each identified sentiment change, enabling the attribution of a market price when no direct market value exists [41]. The proxies were carefully selected based on their relevance and supported by cited research studies to ensure the credibility of the assigned market prices. Table 5 presents the financial proxies used for each sentiment.

- Study hypothesis 3: The integration of DPS into the EAP will generate a positive SROI, reflecting the added value of peer support services for EAP clientele.

Table 5. Financial proxies for each sentiment.

Factors in which digital peer support users experience a reduction	Financial proxy	Cost
Sadness	Average unit cost of treating someone with depression (National Institutes of Health)	• PPPY[a]: US $10,074 (SD $25,694) • PPPM[b,c]: US $839.50
Loneliness and isolation	Average unit cost of treating someone with objective isolation and loneliness (National Institutes of Health)	• PPPY: US $1643 in Medicare spending (objective isolation) • PPPM: US $136.92
Stress (eg, generalized anxiety)	Average unit cost of treating someone with stress	• PPPY: US $6475 • PPPM: US $539.58

[a]PPPY: per person per year.

[b]PPPM: per person per month.

[c]PPPM calculations were made by dividing the PPPY value by 12 months.

SROI Analysis

The primary objective of this SROI analysis is to calculate and present a ratio that compares the cost of investment (US $) in an intervention with the total social, environmental, and economic value (US $) it generates. To prevent overestimation, the SROI methodology recommends certain considerations (see Table 6) [41].

The SROI analysis demonstrated that the DPS program generates positive social value, complementing overall treatment. The SROI values ranged from US $1.66 (loneliness) to US $2.50 (stress) to US $2.58 (sadness) for every dollar invested. Detailed calculations for the impact ratio are available in Multimedia Appendix 1.

Table 6. Factors considered for calculating the social return on investment impact ratio.

Factor	Description
Deadweight	A measure to describe the amount of an outcome that would have happened anyway, even if the DPS[a] service had not been offered. For example, the participants experiencing reductions in momentary loneliness may have a deadweight value of 10% from simply joining the DPS chat, but without further intervention from the peer support moderator or other peers.
Displacement	An assessment of what activities or services are displaced by the presence of the digital peer group. There was no evidence of displacement as, before the digital support option was added, there were no after-hours DPS chat services for employee assistance program members to access.
Attribution	A measure to consider how much of an identified theme is a result of the group studied or is influenced by external factors, for example, if participants also attend another service, such as counseling, which reduces their sense of chronic sadness, a conservative estimate of 20% of the identified momentary sentiment change may be attributed to the group studied.
Dropoff	The percentage of the outcome that decreases after the intervention is complete. For example, a conservative percentage of 10% re-emergence of stress may be expected after the user ends their DPS chat due to a return to their stressful environment.

[a]DPS: digital peer support.

Discussion

Principal Findings

The study examined the impact of integrating a DPS service within an EAP on utilization rates, sentiment change, and therapy session use over 1 year. Additionally, it assessed the social, environmental, and economic impact of the DPS service by estimating an SROI value. Results indicated that DPS service utilization primarily occurred after hours. In the matched sample, participants used fewer therapy sessions post-DPS integration. The SROI ranged from US $1.66 to US $2.50 to US $2.58 for every dollar invested for loneliness, stress, and sadness sentiments, respectively. These findings suggest that combining DPS with traditional EAP services may provide noncrisis

XSL•FO
RenderX

emotional support after hours, potentially reducing EAP therapy session use for lower-acuity emotional concerns. This, in turn, could allow EAPs to allocate funds toward additional ancillary support services and enable employers to benefit from a positive SROI.

The study's findings—that participants experience significant improvements in sentiment, such as reduced sadness, loneliness, and stress—align with existing literature on patient activation [7,8]. Research indicates that when individuals are activated—meaning they gain confidence and skills in managing their health—they tend to experience improvements in emotional well-being [8]. By providing peer support that fosters autonomy and self-efficacy, participants can develop coping strategies that reduce distress and promote mental health [43]. This aligns with findings that patient activation leads to better emotional outcomes through the development of personal coping mechanisms and self-management [8]. The positive SROI from integrating DPS into EAP services is also consistent with the literature on patient activation. Activated patients, who feel more in control of their health, contribute to more efficient resource utilization and improved outcomes [44]. The integration of peer support services such as DPS has been shown to enhance engagement with formal mental health services, creating value not only for individuals but also for the broader service system [44]. This supports the idea that empowering individuals through peer support can yield both personal and systemic benefits, ultimately leading to a positive SROI.

The study may also contribute to the quality of life literature, which is commonly used in cost-effectiveness studies. Health-related quality of life (HRQoL) refers to a person's physical, emotional, and social well-being as influenced by a health intervention [45]. While this study did not include standardized HRQoL assessments (eg, 36-item Short Form [SF-36] or EQ-5D) due to real-world data collection constraints and the anonymous nature of the DPS service, sentiment changes and therapy utilization could serve as HRQoL proxies [46]. Significant reductions in sadness, loneliness, and stress during DPS chats suggest improvements in emotional well-being, a core component of HRQoL. The reduction in momentary negative sentiments may indicate that participants used DPS as a coping mechanism for emotional stability. Additionally, the suggested decrease in therapy sessions implies that participants may have achieved better emotional regulation or relief through DPS, ultimately impacting their overall quality of life. Relatedly, the high utilization of DPS after hours addresses a gap in traditional mental health services and may contribute to participants' perceived control over their emotional health. Such improvements in HRQoL proxies could translate into better job performance [47] and employee satisfaction [48], while also serving as a valuable complement to therapy. This approach allows mental health providers to focus on higher-acuity cases while DPS services address lower-acuity concerns—key objectives for most EAP organizations [48]. Finally, the calculated SROI highlights tangible benefits for emotional health (eg, sadness, loneliness, stress), further demonstrating the intervention's effectiveness in enhancing HRQoL's socioemotional components.

The study's findings further align with existing literature in several key areas. The use of AI for matching participants with peers who have shared lived experiences, as well as for sentiment analysis, supports growing evidence that AI can enhance the personalization and efficiency of life struggle and nonmental health services [49-52] This study offers several novel contributions to the literature on mental health interventions and EAPs. It is among the first to demonstrate that integrating a DPS service with EAPs may enhance after-hours support, reduce the utilization of EAP therapy sessions—allowing EAPs to allocate resources more effectively for higher-acuity emotional concerns—and provide supplementary SROI. DPS services are unique within the EAP environment and may be particularly valuable for EAP participants who are not yet ready to confront the stigma or commitment often associated with therapy [53] but still benefit from peer support. By leveraging AI-driven natural language processing to match users based on their struggles and conducting sentiment analysis on chat narratives, this study highlights an innovative use of technology to enhance support for stress and life challenges. Conducted in a naturalistic setting with a diverse sample, the study's design strengthens ecological validity, demonstrating the practical feasibility and effectiveness of DPS in real-world organizational environments.

Study Limitations

This study faced several potential threats to both internal and external validity and implemented strategies to mitigate these issues. One threat to internal validity was the influence of external events (history) during the study period (June 2023 to May 2024) [54,55]. To address this, a comparison group from the previous year (June 2022 to May 2023) was used, and quarterly data monitoring helped identify and adjust for significant external changes [55]. Matched samples based on additional demographics were not possible due to the anonymity of DPS users. Future research using alternative study designs with nonanonymized data may be necessary.

Maturation effects, referring to the natural changes in participants' mental health over time, were addressed using a cohort design with propensity score matching. This approach ensured that participants in both the intervention and control groups were similar in age, gender, and presenting concerns [56]. By balancing the groups, it minimized the impact of maturation on the results [56]. Additionally, selection bias was mitigated through propensity score matching, which created comparable groups by matching participants in the pre-DPS and post-DPS periods based on key demographics and presenting concerns [57].

Experimental mortality, or dropout, was another concern. To address this, the study included multiple participants matched on the same variables, ensuring that the analysis could proceed even if some participants dropped out [57]. A diverse sample across age and gender was included to reduce the impact of selection biases on generalizability. Additionally, conducting the study in a naturalistic setting, where participants used the DPS and EAP services as they would in real life, helped minimize the reactive effects of experimental arrangements, enhancing ecological validity [55]. Furthermore, the availability

XSL•FO
RenderX

and normalization of the EAP program may have influenced participant engagement. EAP programs are voluntary for companies in the United States, Australia, Canada, and England, whereas Nordic countries have legal mandates requiring mandatory EAPs [58]. This distinction is relevant from the perspective of patient activation. Voluntary EAPs may require high awareness and active promotion to engage participants and may primarily attract individuals with higher activation or those who are more self-directed. As a result, individuals with lower activation or those in the early stages of their mental health concerns may remain underserved [58]. Additionally, factors such as program promotion and mental health stigma could contribute to uneven access across the workforce. By contrast, mandatory EAPs may help normalize access to mental health support, encourage early and preventive engagement, and promote continuity of care by integrating with public health systems, unlike the segmentation often seen with voluntary EAPs [59]. Limitations of the SROI process include the complexity of assigning financial proxies to sentiment change and the availability of data required for robust calculations, such as displacement and attribution values. A key risk in SROI analysis is an overemphasis on the ratio itself, without considering the underlying content, which provides deeper insight into the value created by different groups [58]. By acknowledging and addressing these challenges through design and analysis strategies, the study generated meaningful insights into the impact of integrating DPS services within an EAP.

Future Research

The study opens several avenues for future research. Long-term impact assessments are needed to evaluate how integrating DPS with EAP services influences therapy utilization and outcomes over extended periods. A broader demographic analysis could provide insights into how factors such as age, socioeconomic status, and cultural background affect engagement with these services. Additionally, incorporating qualitative insights may help illuminate user experiences and satisfaction, offering a deeper understanding of which aspects of the service are most effective or subjectively valued by clients.

Future research could also compare DPS with other digital mental health interventions to identify which models offer the greatest benefits in different contexts. Investigating how DPS integrates with other mental health services, such as primary care or community programs, could provide a more comprehensive understanding of its role within the broader health care system. While comparisons of single- versus multilayered support ecosystems fall beyond the scope of this study, they represent a valuable avenue for future exploration (eg, digital mental health vs in-person services or EAP vs DPS). Additionally, examining DPS's impact on different mental health conditions and conducting a detailed economic analysis of cost savings could further strengthen the case for its adoption. Moreover, this study did not calculate quality-adjusted life years; however, the observed improvements and associated cost reductions provide strong support for the intervention's economic value. Future research could explore additional methods to further quantify these benefits. These efforts would deepen our understanding of DPS' effectiveness and its potential applications across diverse settings.

Conclusions

This study demonstrates that integrating DPS services within an EAP may influence therapy utilization, allowing EAPs to serve higher-acuity clients while lower-acuity clients access supplementary support, such as DPS, for everyday emotional needs. Additionally, the research highlights the potential social, environmental, and broader economic benefits of incorporating DPS services into the emotional support ecosystem, with positive estimated SROI outcomes. By leveraging AI-driven natural language processing for user matching and sentiment analysis, the study underscores the potential of technology to enhance stress management, address life challenges, support mental health, and optimize resource allocation. The findings emphasize the need for further research into DPS effectiveness, particularly regarding long-term impacts, demographic variations, and comparative effectiveness with other interventions. Overall, this study provides valuable insights into the evolving landscape of digital mental health services, reinforcing their role in enhancing organizational wellness programs and identifying key areas for future exploration to maximize their potential.

Conflicts of Interest

Authors HN and ZD disclose an employment relationship at Supportiv. RAM discloses a founder, financial, and an employment relationship at MINES and Associates. Available precautions were taken by all authors to minimize the impact of these affiliations on the study design, data collection, analysis, or interpretation of the findings. All other authors declare no further conflicts of interest.

Multimedia Appendix 1
Breakdown of SROI calculations as calculated using the SROI calculator available on the Sopact website. SROI: social return on investment.
[DOCX File , 8 KB - humanfactors_v12i1e68221_app1.docx]

References

1. World Health Organization (WHO). The world health report. WHO. Geneva, Switzerland: WHO; 2001. URL: https://www.who.int/news/item/28-09-2001-the-world-health-report-2001-mental-disorders-affect-one-in-four-people [accessed 2025-02-14]

XSL•FO
RenderX

2. Nowrouzi-Kia B, Haritos AM, Long BS, Atikian C, Fiorini LA, Gohar B, et al. Remote work transition amidst COVID-19: impacts on presenteeism, absenteeism, and worker well-being-a scoping review. PLoS One 2024;19(7):e0307087 [FREE Full text] [doi: 10.1371/journal.pone.0307087] [Medline: 39024322]
3. Abramson A. Burnout and stress are everywhere. American Psychological Association. 2022 Jan 1. URL: https://www.apa.org/monitor/2022/01/special-burnout-stress [accessed 2024-10-23]
4. Ling LM. Prevalence of mental health and its impact on employee productivity. Reviews of Contemporary Business Analytics 2023;6(1):1-13 [FREE Full text]
5. Terlizzi E, Schiller J. Mental health treatment among adults aged 18-44: United States, 2019-2021. NCHS Data Brief 2022 Sep(444):1-8 [FREE Full text] [Medline: 36135999]
6. Wu A, Roemer E, Kent K, Ballard D, Goetzel R. Organizational best practices supporting mental health in the workplace. Journal of Occupational & Environmental Medicine 2021;63(12):e925-e931. [doi: 10.1097/jom.0000000000002407]
7. Hibbard JH, Stockard J, Mahoney ER, Tusler M. Development of the Patient Activation Measure (PAM): conceptualizing and measuring activation in patients and consumers. Health Serv Res 2004 Aug;39(4 Pt 1):1005-1026 [FREE Full text] [doi: 10.1111/j.1475-6773.2004.00269.x] [Medline: 15230939]
8. Hibbard JH, Greene J. What the evidence shows about patient activation: better health outcomes and care experiences; fewer data on costs. Health Aff (Millwood) 2013 Feb;32(2):207-214. [doi: 10.1377/hlthaff.2012.1061] [Medline: 23381511]
9. Miele DB, Scholer AA. Self-regulation of motivation. In: Handbook of Motivation at School. New York, NY: Routledge; 2016:375-396.
10. Burns D. The Feeling Good Handbook: The Groundbreaking Program With Powerful New Techniques and Step-by-Step Exercises to Overcome Depression, Conquer Anxiety, and Enjoy Greater Intimacy. Washington, D.C: Plume; 1999.
11. Koechlin H, Coakley R, Schechter N, Werner C, Kossowsky J. The role of emotion regulation in chronic pain: a systematic literature review. J Psychosom Res 2018 Apr;107:38-45. [doi: 10.1016/j.jpsychores.2018.02.002] [Medline: 29502762]
12. Chinman M, George P, Dougherty RH, Daniels AS, Ghose SS, Swift A, et al. Peer support services for individuals with serious mental illnesses: assessing the evidence. Psychiatr Serv 2014 Apr 01;65(4):429-441. [doi: 10.1176/appi.ps.201300244] [Medline: 24549400]
13. Berry LL, Mirabito AM, Baun WB. What's the hard return on employee wellness programs? Harvard Business Review. 2010 Dec 1. URL: https://hbr.org/2010/12/whats-the-hard-return-on-employee-wellness-programs [accessed 2024-10-23]
14. Senn S, Orgera K, Ramirez L. Who's your therapist? Examining the behavioral health workforce. Association of American Medical Colleges. 2023 Oct 26. URL: https://www.aamcresearchinstitute.org/our-work/data-snapshot/whos-your-therapist#:~:text=The%20behavioral%20health%20workforce%20typically,health%20care%20model%20and%20organization [accessed 2025-02-14]
15. Modi H, Orgera K, Grover A. Exploring barriers to mental health care in the U.S. Research and Action Institute. 2022 Oct 10. URL: https://www.aamcresearchinstitute.org/our-work/issue-brief/exploring-barriers-mental-health-care-us [accessed 2024-10-23]
16. Fact sheet: President Biden to announce strategy to address our national mental health crisis, as part of unity agenda in his First State of the Union. The White House. 2022 Mar 1. URL: https://www.whitehouse.gov/briefing-room/statements-releases/2022/03/01/fact-sheet-president-biden-to-announce-strategy-to-address-our-national-mental-health-crisis-as-part-of-unity-agenda-in-his-first-state-of-the-union [accessed 2024-10-23]
17. Tracy K, Wallace S. Benefits of peer support groups in the treatment of addiction. SAR 2016 Sep;Volume 7:143-154. [doi: 10.2147/sar.s81535]
18. Fortuna KL, Venegas M, Umucu E, Mois G, Walker R, Brooks JM. The Future of Peer Support in Digital Psychiatry: Promise, Progress, and Opportunities. Curr Treat Options Psych 2019 Jun 20;6(3):221-231. [doi: 10.1007/s40501-019-00179-7]
19. 24 Million Americans Still Lack Broadband Connectivity. The Pew Charitable Trusts. 2018. URL: https://www.pewtrusts.org/en/research-and-analysis/fact-sheets/2018/07/24-million-americans-still-lack-broadband-connectivity [accessed 2024-10-23]
20. Institute of Medicine. Health literacy: a prescription to end confusion. The National Academies Press. URL: https://nap.nationalacademies.org/catalog/10883/health-literacy-a-prescription-to-end-confusion?onpi_newsdoc04082004= [accessed 2024-10-23]
21. Ryan R, Deci EL. Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness. New York, NY: Guilford Publications; 2018.
22. Deci E, Ryan R. Handbook of Self-Determination Research. Rochester, NY: University Rochester Press; 2004.
23. Schwei RJ, Hetzel S, Kim K, Mahoney J, DeYoung K, Frumer J, et al. Peer-to-peer support and changes in health and well-being in older adults over time. JAMA Netw Open 2021 Jun 01;4(6):e2112441. [doi: 10.1001/jamanetworkopen.2021.12441] [Medline: 34129024]
24. Kaplan K, Salzer MS, Solomon P, Brusilovskiy E, Cousounis P. Internet peer support for individuals with psychiatric disabilities: A randomized controlled trial. Social Science & Medicine 2011 Jan;72(1):54-62. [doi: 10.1016/j.socscimed.2010.09.037]

XSL•FO
RenderX

25. Fortuna KL, Naslund JA, LaCroix JM, Bianco CL, Brooks JM, Zisman-Ilani Y, et al. Digital peer support mental health interventions for people with a lived experience of a serious mental illness: systematic review. JMIR Ment Health 2020 Apr 03;7(4):e16460. [doi: 10.2196/16460] [Medline: 32243256]
26. Rotondi AJ, Haas GL, Anderson CM, Newhill CE, Spring MB, Ganguli R, et al. A clinical trial to test the feasibility of a telehealth psychoeducational intervention for persons with schizophrenia and their families: intervention and 3-month findings. Rehabil Psychol 2005 Nov;50(4):325-336 [FREE Full text] [doi: 10.1037/0090-5550.50.4.325] [Medline: 26321774]
27. Rotondi AJ, Anderson CM, Haas GL, Eack SM, Spring MB, Ganguli R, et al. Web-based psychoeducational intervention for persons with schizophrenia and their supporters: one-year outcomes. Psychiatric Services 2010 Nov 01;61(11):1099-1105. [doi: 10.1176/appi.ps.61.11.1099]
28. Biagianti B, Quraishi SH, Schlosser DA. Potential benefits of incorporating peer-to-peer interactions into digital interventions for psychotic disorders: a systematic review. Psychiatr Serv 2018 Apr 01;69(4):377-388. [doi: 10.1176/appi.ps.201700283] [Medline: 29241435]
29. Naslund JA, Aschbrenner KA, Marsch LA, Bartels SJ. The future of mental health care: peer-to-peer support and social media. Epidemiol Psychiatr Sci 2016 Apr;25(2):113-122 [FREE Full text] [doi: 10.1017/S2045796015001067] [Medline: 26744309]
30. Simms L, Ottman KE, Griffith JL, Knight MG, Norris L, Karakcheyeva V, et al. Psychosocial peer support to address mental health and burnout of health care workers affected by COVID-19: a qualitative evaluation. Int J Environ Res Public Health 2023 Mar 03;20(5):4536 [FREE Full text] [doi: 10.3390/ijerph20054536] [Medline: 36901545]
31. Long KH, Moriarty JP, Mittelman MS, Foldes SS. Estimating the potential cost savings from the New York University Caregiver Intervention in Minnesota. Health Aff (Millwood) 2014 Apr;33(4):596-604. [doi: 10.1377/hlthaff.2013.1257] [Medline: 24711320]
32. Bassuk EL, Hanson J, Greene RN, Richard M, Laudet A. Peer-delivered recovery support services for addictions in the United States: a systematic review. J Subst Abuse Treat 2016 Apr;63:1-9. [doi: 10.1016/j.jsat.2016.01.003] [Medline: 26882891]
33. Rayland A, Andrews J. From Social Network to Peer Support Network: Opportunities to Explore Mechanisms of Online Peer Support for Mental Health. JMIR Ment Health 2023 Feb 28;10:e41855. [doi: 10.2196/41855]
34. Madaan A, Zhou S, Alon U, Yang Y, Neubig G. Language models of code are few-shot commonsense learners. In: Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing. New York, NY: Association for Computing Machinery; 2022 Presented at: The 2022 Conference on Empirical Methods in Natural Language Processing; December 7-11, 2022; Abu Dhabi, United Arab Emirates. [doi: 10.18653/v1/2022.emnlp-main.90]
35. Engel G. The need for a new medical model: a challenge for biomedicine. Science 1977 Apr 08;196(4286):129-136. [doi: 10.1126/science.847460] [Medline: 847460]
36. Nicholls J, Lawlor E, Neitzert E, Goodspeed T. A guide to social return on investment (SROI) (Revised). Socioeco.org. 2012. URL: https://www.socioeco.org/bdf_fiche-publication-929_en.html [accessed 2024-10-23]
37. Millar R, Hall K. Social Return on Investment (SROI) and Performance Measurement. Public Management Review 2013 Sep;15(6):923-941. [doi: 10.1080/14719037.2012.698857]
38. Engel GL. The clinical application of the biopsychosocial model. Am J Psychiatry 1980 May;137(5):535-544. [doi: 10.1176/ajp.137.5.535] [Medline: 7369396]
39. Elyoseph Z, Refoua E, Asraf K, Lvovsky M, Shimoni Y, Hadar-Shoval D. Capacity of Generative AI to Interpret Human Emotions From Visual and Textual Data: Pilot Evaluation Study. JMIR Ment Health 2024 Feb 06;11:e54369 [FREE Full text] [doi: 10.2196/54369] [Medline: 38319707]
40. Dana Z, Nagra H, Kilby K. Role of synchronous, moderated, and anonymous peer support chats on reducing momentary loneliness in older adults: retrospective observational study. JMIR Form Res 2024 Oct 25;8:e59501 [FREE Full text] [doi: 10.2196/59501] [Medline: 39453688]
41. Social return on investment (SROI) - new approach. Sopact. URL: https://www.sopact.com/guides/social-return-on-invesment-sroi#social-return-on-investment-sroi [accessed 2024-10-23]
42. Greene J, Hibbard JH. Why does patient activation matter? An examination of the relationships between patient activation and health-related outcomes. J Gen Intern Med 2012 May;27(5):520-526 [FREE Full text] [doi: 10.1007/s11606-011-1931-2] [Medline: 22127797]
43. Schunk DH, DiBenedetto MK. Motivation and social cognitive theory. Contemporary Educational Psychology 2020 Jan;60:101832. [doi: 10.1016/j.cedpsych.2019.101832]
44. Lin M, Weng W, Apriliyasari R, VAN Truong P, Tsai P. Effects of patient activation intervention on chronic diseases: a meta-analysis. J Nurs Res 2020 Oct;28(5):e116. [doi: 10.1097/jnr.0000000000000387] [Medline: 32649394]
45. Taylor V. Measuring Healthy Days: Population Assessment of Health-Related Quality of Life. PsycEXTRA Dataset 2000:40 [FREE Full text]
46. Revicki DA, Gnanasakthy A, Weinfurt K. Documenting the rationale and psychometric characteristics of patient reported outcomes for labeling and promotional claims: the PRO Evidence Dossier. Qual Life Res 2007 May 1;16(4):717-723. [doi: 10.1007/s11136-006-9153-5] [Medline: 17268927]

XSL•FO
RenderX

47. Golden-Biddle K. A Review of: "Positive Organizational Scholarship: Foundations of a New Discipline, edited by Kim S. Cameron, Jane E. Dutton, and Robert E. Quinn". International Public Management Journal 2007 Feb 24;9(2):199-202. [doi: 10.1080/10967490600767035]
48. Ilies R, Dimotakis N, De Pater I. Psychological and physiological reactions to high workloads: implications for well - being. Personnel Psychology 2010 May 12;63(2):407-436. [doi: 10.1111/j.1744-6570.2010.01175.x]
49. Cross S, Bell I, Nicholas J, Valentine L, Mangelsdorf S, Baker S, et al. Use of AI in Mental Health Care: Community and Mental Health Professionals Survey. JMIR Ment Health 2024 Oct 11;11:e60589-e60589. [doi: 10.2196/60589]
50. Bossewitch J, Brown LX, Gooding PM, Harris L, Horton J, Katterl S, et al. SSRN Journal 2022:4215994 [FREE Full text] [doi: 10.2139/ssrn.4215994]
51. Stigma, prejudice and discrimination against people with mental illness. Psychiatry. URL: https://www.psychiatry.org/patients-families/stigma-and-discrimination [accessed 2024-10-23]
52. Duckett LJ. Quantitative Research Excellence: Study Design and Reliable and Valid Measurement of Variables. J Hum Lact 2021 Aug 09;37(3):456-463. [doi: 10.1177/08903344211019285] [Medline: 34107788]
53. Mara C, Peugh J. Validity of data collected from randomized behavioral clinical trials during the COVID-19 pandemic. J Pediatr Psychol 2020 Oct 01;45(9):971-976. [doi: 10.1093/jpepsy/jsaa078] [Medline: 32968774]
54. Harder VS, Stuart EA, Anthony JC. Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. Psychol Methods 2010 Sep;15(3):234-249 [FREE Full text] [doi: 10.1037/a0019623] [Medline: 20822250]
55. ijones. Measurement error and bias. In: Epidemiology for the Uninitiated. London, UK: BMJ; 2020.
56. Marr B. Artificial Intelligence in Practice: How 50 Successful Companies Used AI and Machine Learning to Solve Problems. Hoboken, NJ: John Wiley & Sons; 2019.
57. Thieme A, Hanratty M, Lyons M, Palacios J, Marques RF, Morrison C, et al. Designing human-centered AI for mental health: developing clinically relevant applications for online CBT treatment. ACM Trans Comput-Hum Interact 2023 Mar 17;30(2):1-50. [doi: 10.1145/3564752]
58. Attridge M. A global perspective on promoting workplace mental health and the role of employee assistance programs. Am J Health Promot 2019 May 21;33(4):622-629. [doi: 10.1177/0890117119838101c] [Medline: 31006254]
59. Nordic cooperation. URL: https://www.norden.org/en/nordic-council-ministers [accessed 2025-01-05]

Abbreviations

AI: artificial intelligence
DPS: digital peer support
EAP: employee assistance program
HRQoL: health-related quality of life
PSM: peer support moderator
ROI: return on investment
SF-36: 36-item Short Form
SROI: social return on investment

Edited by A Kushniruk; submitted 31.10.24; peer-reviewed by I Pyykko, CI Sartorão Filho, KW Tay; comments to author 22.12.24; revised version received 09.01.25; accepted 28.01.25; published 25.02.25.

Please cite as:
Nagra H, Mines RA, Dana Z
Exploring the Impact of Digital Peer Support Services on Meeting Unmet Needs Within an Employee Assistance Program: Retrospective Cohort Study
JMIR Hum Factors 2025;12:e68221
URL: https://humanfactors.jmir.org/2025/1/e68221
doi:10.2196/68221
PMID:39998863

©Harpreet Nagra, Robert A Mines, Zara Dana. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 25.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Identifying Strategies for Home Management of Ostomy Care: Content Analysis of YouTube

Marketa Haughey[1*], BSc, PhD; David M Neyens[2*], PhD; Casey S Hopkins[3*], PhD; Christofer Gonzaga[2*], BS; Melinda Harman[1*], PhD

[1]Department of Bioengineering, Clemson University, Clemson, SC, United States

[2]Department of Industrial Engineering, Clemson University, Clemson, SC, United States

[3]School of Nursing, Clemson University, Clemson, SC, United States

[*]all authors contributed equally

Corresponding Author:
Melinda Harman, PhD
Department of Bioengineering
Clemson University
301 Rhodes Engineering Research Center
Clemson, SC, 29634
United States
Phone: 1 8646564140
Email: harman2@clemson.edu

Abstract

Background: The social media platform YouTube is a recognized educational resource for health information, but few studies have explored its value for conveying the lived experience of individuals managing chronic health conditions and end users' interactions with medical device technology. Our study explores self-care strategies and end user needs of people living with a stoma because patient education and engagement in ostomy self-care are essential for avoiding ostomy-related complications. Ostomy surgery creates a stoma (an opening) in the abdomen to alter the route of excreta from digestive and urinary organs into a detachable external pouching system. After hospital discharge, people who have undergone ostomies perform critical self-care tasks including frequent ostomy appliance changes and stomal and peristomal skin maintenance.

Objective: The purpose of this study was to systematically assess YouTube videos narrated by people who have undergone ostomies about their ostomy self-care in home (nonhospital) settings with a focus on identifying end user needs and different strategies used by people who have undergone ostomies during critical self-care tasks.

Methods: Using predefined search terms and clear inclusion and exclusion criteria, we identified YouTube videos depicting narrators who have undergone ostomies and their ostomy self-care in home settings. Using a consensus coding approach among 3 independent reviewers, all videos were analyzed to collect metadata, data of narrators who have undergone ostomies, and specific content data.

Results: There were 65 user-generated YouTube videos that met the inclusion and exclusion criteria. These videos were posted by 28 unique content creators representing a broad range of ages who used a variety of supplies. The common challenges discussed were peristomal skin complications, inadequate appliance adhesion and subsequent leakage, and supplies-related challenges. Narrators who have undergone ostomies discussed various expert tricks and tips to successfully combat these challenges.

Conclusions: This study used a novel approach to gain insights about end user interactions with medical devices while performing ostomy self-care, which are difficult to gain using traditional behavioral techniques. The analysis revealed that people who have undergone ostomies are willing to share their personal experience with ostomy self-care on the web and that these videos are viewed by the public. User-generated videos demonstrated a variety of supplies used, end user needs, and different strategies for performing ostomy self-care. Future research should examine how these findings connect to YouTube ostomy self-care content generated by health care professionals and organizations and to guidelines for ostomy self-care.

(JMIR Hum Factors 2025;12:e66634) doi:10.2196/66634

XSL•FO
RenderX

KEYWORDS

medical device usability; digital health; online support groups; living with chronic medical conditions; ostomy self-care; YouTube; patient education; user needs assessment; users experience; social media; ostomates; colostomy; ileostomy; usability; usefulness; utility; wearable device; medical device; support group; socials; social network; ostomy; digital; digital technology; digital intervention

Introduction

YouTube as Health Information Resource

Individuals are increasingly seeking health-related information on the web, and in some cases, patients rely on the internet as much or more than their physicians [1,2]. Patients have also reported using web-based resources to make health-related decisions and to manage chronic conditions [2,3].

YouTube, the largest on the web video-sharing platform worldwide for streaming a variety of user-generated content with 2.5 billion users worldwide [4] is a recognized educational resource for sharing and disseminating health-related information and influencing individuals' viewpoints related to health care topics such as disease prevention, treatment therapy, or immunizations [2,3]. Several studies highlight the informational flaws and biases in a wide variety of YouTube health information videos related to clinical decision-making, diagnosing, treatment recommendation, and health promotion. The analysis in these studies generally focuses on the health information provided rather than the individual user experience with the health issue [2,5-9].

Studies involving so-called "e-patients" who are motivated to share health-related personal experiences on the web reveal that sharing their lived experience and knowledge of their disease can empower and engage others while also providing (and gaining) social support [10,11]. While many health-related government organizations such as the World Health Organization, the National Institute of Health, the Center for Disease Control and Prevention, and the American Red Cross use YouTube to disseminate health information [1], user-generated videos are also used as an educational resource. This suggests that user-generated content decreases the gap between health information delivered by health care professionals and the lived experience of dealing with a disease.

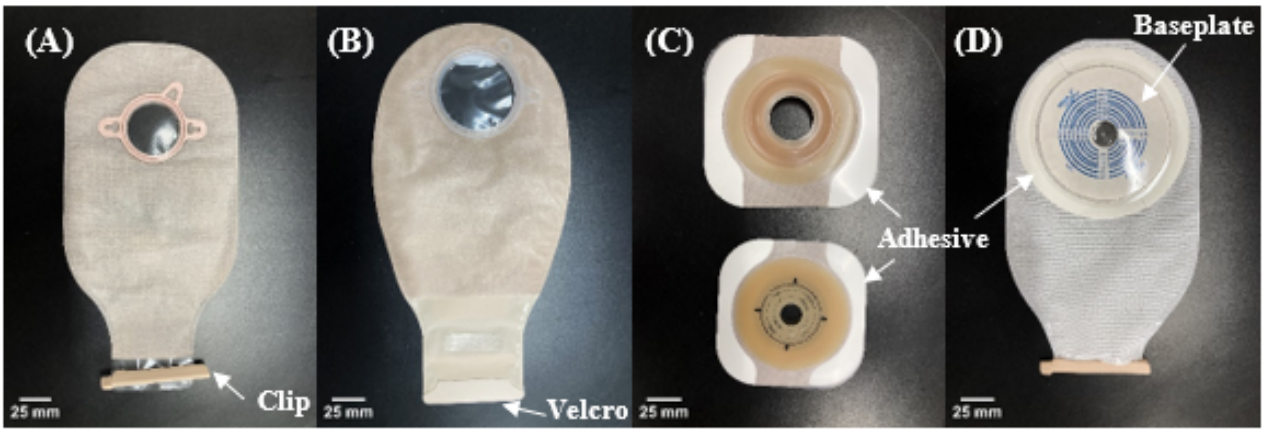
Medical Device Usability in the Home Setting

Usability studies examining the interaction between medical device technology and end users are essential for developing medical devices that are safe and effective [12]. These studies allow for actively including end users in the design process while considering end user competencies and environment, thus, continuously improving medical device design to fit user needs [13]. Capturing the real world, end user experience of people with chronic conditions and their interaction with complex medical device technology in home (nonhospital) settings is challenging [12], such as having a stoma and routinely using an ostomy appliance, and to our knowledge, these studies are limited [14].

Ostomy Surgery

There are about one million people living with ostomy in the United States and more than 150,000 Americans undergo ostomy surgery each year [15,16]. Ostomy surgery alters the route of excreta from internal digestive and urinary organs through a stoma in the abdomen and into a detachable external pouching system (Figure 1) that fits snugly around the stoma [16]. People who have undergone ostomies perform critical self-care tasks that include frequent changing of the pouching system (3-10 changes per week) and maintaining healthy stoma and peristomal skin [17]. Therefore, there is a need to understand the usability of medical device technology used by people who have undergone ostomies in ostomy appliance change procedures and self-care.

Figure 1. An ostomy appliance consists of an external pouching system that is attached to the abdomen of people who have undergone an ostomy using adhesive. There are various design features, including a (A) 2-piece pouch with clip closure, (B) 2-piece pouch with Velcro closure, (C) precut (top) and moldable (bottom) baseplate that is subsequently attached to the 2-piece pouch [in (A) or (B)] at the ring, and (D) 1-piece appliance with clip closure (pouch and baseplate being 1 part).



XSL•FO
RenderX

Ostomy Care Patient Education and Self-Management

It is widely recognized that patient education and engagement in ostomy self-care are essential for avoiding ostomy-related complications [18-20]. However, the education of new people who have undergone ostomies to gain the knowledge and skills needed for managing chronic conditions and performing self-care tasks is not well defined [18]. Ideally, according to the Wound, Ostomy, and Continence Nurses Society Clinical Guidelines and recommendations [21], new people who have undergone ostomies should obtain education from ostomy nurse specialists. Nevertheless, there is evidence of insufficient pre- and postoperative education provided to new people who have undergone ostomies related to stoma and peristomal skin care and complications [22-25]. Due to the short hospitalization trends and the limited availability of ostomy nurse specialists, there is limited teaching time and formal follow-up with the specialists [22-25]. Further, new people who have undergone ostomies may not be mentally and physically prepared for self-management of their ostomy, which requires gaining new knowledge, skills, and attitudes [26,27]. Most people who have undergone ostomies experience some long-term challenges concerning the management of daily self-care of their stoma that broadly impact their health-related quality of life and outcomes [28,29]. Peristomal skin complications are a common problem following ostomy surgery affecting over one-third of patients with ostomies within 90 days post surgery and up to 80% within 2 years [30-33]. While it is known that skin problems interfere with pouch adhesion, causing challenges with pouch leakage, and thus odor, and the necessity for frequent and unscheduled appliance changes [29], it is very difficult to study real-world responses to such challenges.

Study Objective

Given the complexity of ostomy appliance change procedure medical technology and self-care, there is a need to assess the usability of this technology and how self-management skills are developed by people who have undergone ostomies over time. We completed a systematic analysis of user-generated videos related to ostomy self-care published on YouTube and created by individuals with expertise in managing chronic conditions about ostomy self-care in home (eg, nonhospital) settings. Our objective was to identify end user experiences with ostomy medical device technology and the different strategies and procedures used by these experts to change their ostomy appliance and perform stomal and peristomal skin care.

Methods

Ethical Considerations

This study was excluded from an institutional ethics board review because it assessed only copyrighted, public source information and no individual user data was included in the manuscript.

Search Terms and Search Strategy

Search terms related to ostomy self-care were defined by the research team (Table 1) for identifying videos using the search tools within YouTube. During December 2023, YouTube was searched for relevant videos using search terms and Boolean operators under the search category “videos only” tab or “video” filter. All searches were conducted using an incognito or private browser that did not retain cookies to prevent suggested related videos from appearing on search results and to ensure that the search for each term could be replicated. The first 20 videos for each search term were reviewed and documented in a spreadsheet. No videos were included in the analysis if they appeared in the YouTube search result under the “people also watched,” “shorts,” or “sponsored websites” section as these videos may or may not have met the search term criteria. The initial search resulted in 960 videos.

Table 1. Search terms used to search for relevant user-generated videos published on YouTube.

Search term category	Search terms
Ostomy pouch change procedure	[a] + (pouch OR bag) + (care OR change OR procedure)
Ostomy tips and tricks	[a] + (pouch OR bag) + (pros and cons OR tips OR tips and tricks OR empty OR cleaning OR care OR leakage OR leakage prevention)
Ostomy self-care and home management	[a] + (pouch change OR bag change) + (home OR self-care OR care)
Ostomy skincare	[a] + (skin OR peristomal skin) + (care OR treatment OR irritation OR dermatitis)

[a]Is ostomy OR urostomy OR colostomy OR ileostomy OR stoma OR colorectal cancer OR Crohn disease OR inflammatory bowel disease OR IBD OR diverticulitis OR bladder cancer OR cystectomy.

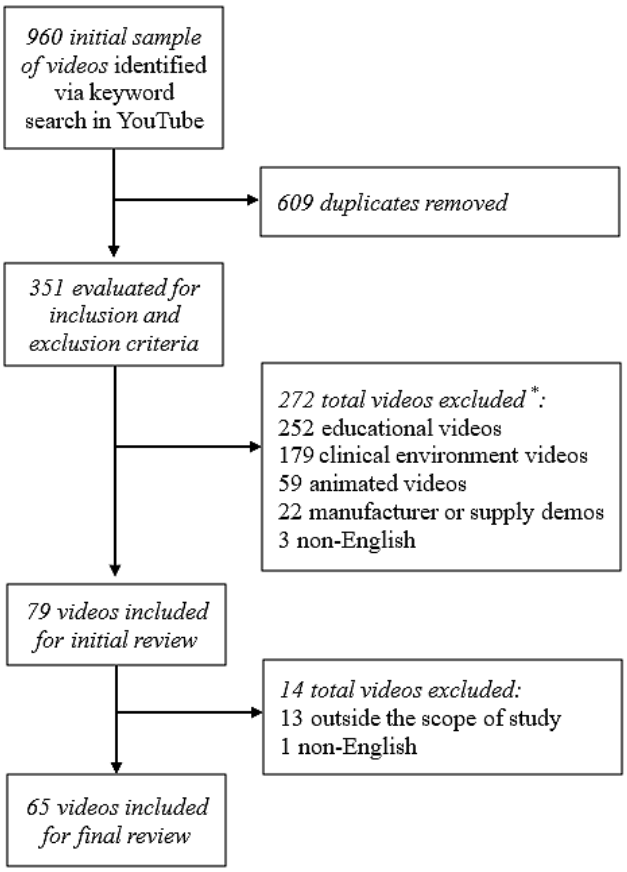
Inclusion and Exclusion Criteria

Overview

From the initial set of 960 videos, a final set of 65 videos (Figure 2) was identified for review after the removal of duplicates and the application of inclusion and exclusion criteria (Textbox 1). Duplicate videos (n=609) were identified and removed by channel grouping and comparing URLs and video titles. The inclusion and exclusion criteria were applied to eliminate educational videos (n=252), videos recorded in a clinical environment (n=179), animated videos (n=59), supply or manufacturer demonstration videos (n=22), videos in languages other than English (n=4) and those videos that were outside the scope of this study (n=13) such as broad discussions of clinical issues or ostomy support. Educational videos were defined as videos not discussing personal experiences related to ostomy self-care, for example, videos providing information related to ostomy belts.

XSL•FO
RenderX

Figure 2. Video search inclusion and exclusion process. *Some videos were excluded for multiple criteria.



Textbox 1. Specific inclusion and exclusion criteria.

Included videos

- Were found using search terms
- Included a user narrative
- Involved any channel, age, gender, race, ethnicity, year, or length
- Depicted ostomy appliance change procedures in a nonhospital setting
- Discussed ostomy self-care, tips, tricks, or challenges
- Involved a person living with a stoma as a narrator
- Were published in the English language

Excluded videos

- Were recorded in a clinical environment or were narrated by a medical professional (not having a stoma)
- Included experiences of users who have not undergone ostomies
- Were educational videos
- Were supply or manufacturer demonstration videos
- Were animated
- Duplicated videos from other search terms
- Were published in the non-English language
- Categorized as video “shorts”
- Were outside the scope of this study

XSL•FO
RenderX

Data Coding

In total, 3 categories of video-specific variables were recorded, including video metadata, narrator data, and specific content data (Table 2). The researchers used consensus coding of the included videos, which is a well-accepted method to enhance credibility and reliability in qualitative research [34]. Before video reviews, 3 general categories and 15 variables were named and defined (Table 2). Each video was then coded individually by 2 reviewers using the defined categories and any discrepancies between reviewers were discussed. If consensus was not reached, a third reviewer reviewed the video and data coding to address bias and reach consensus and to ensure accuracy in data coding. The category definitions were then revised as needed before a final consensus coding of all the included videos by 2 reviewers.

Table 2. Data coding for 3 categories of videos including specific variables, and their definition and type.

Category and variable	Definition
Video metadata	
Title	The title of the video by the content creator
URL	The web address of the video
Channel	The name of the page where videos are posted for viewing
Posted year	The year that video was publicly posted to the YouTube platform
Video length	The time duration of the video
View counts	The number of times the video was viewed as of January 4, 2024
Public ratings	The number of likes or dislikes as of January 4, 2024
Data of narrators who have undergone ostomies	
Presenting assumed gender	The assumed gender of the content creator based on gender markers such as physical build, voice, clothes, and hair
Presenting assumed age group	The assumed age group of the content creator mainly based on physical appearance and hair coloration
Specific video content data	
Supplies used	Any item, product, equipment, or accessory used or discussed as used by a narrator who has undergone an ostomy during the ostomy appliance change procedure
Appliance change frequency[a]	How frequently the narrator who has undergone an ostomy completed an appliance change process
Type of ostomy[a,b]	The type of ostomy disclosed by the narrator who has undergone an ostomy
Underlying condition for ostomy[a,b]	The medical condition experienced by the narrator who has undergone an ostomy and contributing to the ostomy surgery
Common challenges[a]	Description of the problems related to ostomy, ostomy appliance change procedures, and ostomy self-care
Expert tips and tricks	Advice or recommendation given by a narrator who has undergone an ostomy to the audience to address common challenges, aid in expediting or easing the appliance change process, or provide additional information to improve the overall user experience

[a]Variable was not communicated in every video.

[b]If not communicated in the video, the intro page and the about me page for the creator were reviewed.

Results

Video Metadata

After removing duplicate videos and applying the inclusion and exclusion criteria, there were 65 videos that explored different topics related to ostomy, ostomy appliance change procedures, and the self-care of people who have undergone ostomies. Included videos were posted between 2011 and 2023 and averaged about 9.5 minutes in length, over a quarter-million views, and more than 3000 likes (Table 3).

Table 3. The video metadata descriptive statistics for the videos included in the analysis.

Video metadata	Data
Posted year, n/N (%)	Between 2011 and 2023 (11/65, 17% and 12/65, 18% videos in 2020 and 2023, respectively)
Video length, mean (SD; range)	9:36 min (SD 6:30 s; 1:00-31:40 min)
View counts, mean (SD; range)	269,076 (SD 922,352; 1900-5,493,702) views
Public ratings, mean (SD; range)	3015 (SD 8761; 0-58,000) likes

XSL•FO
RenderX

Data of Narrators Who Have Undergone Ostomies

There were 28 unique content creators with several individual creators posting multiple videos. Further, 2 high volume narrators who have undergone ostomies uploaded almost half (31/65) of all the videos (24 and 7 videos from the 2 most frequent content creators). Most videos included a female narrator and a narrator in the 20-30 years age range followed by a middle-aged narrator, an older narrator, and a teenage narrator (Table 4). None of the narrators who have undergone ostomies communicated their gender or age in the videos. Therefore, the presenting assumed gender was determined by the reviewers based on the gender markers such as physical build, voice, clothes, and hair and the presenting assumed age based on the physical appearance and hair coloration.

Table 4. Data of narrators who have undergone ostomies for videos included in the final review.

Unique narrators who have undergone ostomies (n=28)	Data, n/N (%)
Presenting assumed gender	
Female	20/28 (71% narrators who have undergone ostomies)
Male	8/28 (29% narrators who have undergone ostomies)
Presenting assumed age group	
Infant[a]	1/65 (2% videos)
Teenager	4/65 (6% videos)
Aged 20-30 years	42/65 (65% videos)
Middle age	13/65 (20% videos)
Older	5/65 (8% videos)

[a]One video was presented about an infant with an ostomy that was narrated by the parent caregiver.

Specific Video Content Data

Ostomy Video Content

From the final set of 65 videos, 26/65 (40%) videos discussed ostomy self-care tips and tricks that focused most frequently on maintaining peristomal skin health, preventing leaks, pancaking, supply use, and general tips; 21/65 (32%) videos captured a full ostomy appliance change procedure and 14/65 (22%) videos included part of an ostomy appliance change procedure such as emptying, specific supply use, and skin preparation and appliance application, among others; and 4/65 (6%) videos were about the personal experience of a person living with a stoma (Table 5).

XSL•FO
RenderX

Table 5. Types of ostomy-related video content, its frequency, and topics discussed.

Video content and topic discussed	Video count, n	Video frequency, n/N (%)
Tips and tricks		26/65 (40%)
Peristomal skin health	6	
Leaks	5	
Pancaking	3	
Emptying or not emptying	3	
Specific supply use	2	
General	2	
Odor	1	
Intimacy	1	
Adhesion	1	
Pouch decoration	1	
Full ostomy appliance change		
Ostomy appliance change procedure	21	21/65 (32%)
Part of ostomy appliance change		14/65 (22%)
Empty an appliance	5	
Specific supply use	3	
Skin preparation and appliance application	2	
Change appliance after shower	1	
Remove appliance and clean peristomal skin	1	
Cut or fit ostomy baseplate	1	
Appliance change omitting details	1	
Experience		4/65 (6%)
Living with stoma	2	
Living in London	1	
Stories of people who have undergone ostomies	1	

Supplies Used

The narrators who have undergone ostomies did not all use identical supplies and there was a variety of types, numbers, and frequencies of supplies used during the ostomy appliance change procedures (Table 6). Among the 35 videos that captured a full or partial ostomy appliance change procedure, 25/35 (71%) included a 2-piece pouching system, and 9/35 (26%) included a 1-piece pouching system (Table 6). In 1/35 (3%) videos, the type of ostomy appliance could not be determined. Furthermore, the baseplate was cut to fit or precut by the supplier in 25/35 (71%) videos and molded by hand in 3/35 (9%) videos. In 7/35 (20%) videos, the method for sizing the baseplate by cutting or molding could not be determined.

Further, supplies used to clean the peristomal skin and remove any remaining adhesive or output included paper towel or toilet paper in 19/35 (54%) videos, adhesive remover wipes in 17/35 (49%) videos, adhesive remover spray in 14/35 (40%) videos, dry wipes in 11/35 (31%) videos, towel or washcloth in 9/35 (26%) videos, water in 7/35 (20%) videos, and wet wipes in 4/35 (11%) videos. Supplies used to protect the skin included a barrier ring in 17/35 (49%) videos, skin barrier in the form of wipe, spray, or cream in 15/35 (43%) videos, and ostomy paste in 6/35 (17%) videos. Supplies used to treat peristomal skin disorders included stoma powder in 12/35 (34%) videos and skin treatments (eg, corticosteroids, and antifungal or other anti-inflammatory medications) in 8/35 (23%) videos. Supplies used to aid adhesion between the pouch and peristomal skin included extenders in 5/35 (14%) videos and belts in 4/35 (11%) videos. Supplies used to prevent or minimize odor issues were odor eliminators in 9/35 (26%) videos and lubricating deodorant in 5/35 (14%) videos. Other general ostomy supplies that were used included plastic bags in 19/35 (54%) to discard the ostomy appliance before disposal, scissors in 16/35 (46%) videos to cut a baseplate to match the size of the stoma, some kind of bag in 4/35 (11%) videos to carry ostomy supplies, and some type of container in 4/35 (11%) videos to rinse the pouch with water.

XSL•FO
RenderX

Table 6. Supplies used or discussed in the 35 videos depicting full or partial ostomy appliance change procedures organized by supply category.

Supply category and supply	Videos[a], n (%)
Pouching system	
2-piece appliance	25 (71)
1-piece appliance	9 (26)
Not reported	1 (3)
Stoma or peristomal skin cleaning	
Paper towel or toilet paper	19 (54)
Adhesive remover wipes	17 (49)
Adhesive remover spray	14 (40)
Dry wipes	11 (31)
Towel or washcloth	9 (26)
Water	7 (20)
Wet wipes	4 (11)
Gauze	2 (6)
Skin cleanser	2 (6)
Soap	2 (6)
Saline	1 (3)
Peristomal skin protection	
Barrier ring	17 (49)
Skin barrier (wipe, spray, or cream)	15 (43)
Ostomy paste	6 (17)
Flow assist device	2 (6)
Peristomal skin treatment	
Stoma powder	12 (34)
Skin treatments	8 (23)
Pouch adhesion support	
Extenders	5 (14)
Belt	4 (11)
Skin adhesive	3 (9)
Equalizer	2 (6)
Heating pad	2 (6)
Odor elimination	
Odor eliminator (deodorant or air freshener)	9 (26)
Lubricating deodorant	5 (14)
Lubricant	2 (6)
General	
Plastic bags	19 (54)
Scissors	16 (46)
Supply bag or bag or ziplock	4 (11)
Plastic container or water bottle or cup	4 (11)
Marker	3 (9)
Size template	3 (9)
Disposable pad	1 (3)

XSL•FO
RenderX

Supply category and supply	Videos[a], n (%)
Q-tip	1 (3)
Shower stopper	1 (3)
Syringe	1 (3)

[a]Calculated as x/35×100%.

Underlying Condition and Type of Ostomy

One-third of the narrators who have undergone ostomies (9/28, 32%) did not disclose their underlying condition. Crohn disease was the most frequently reported underlying condition for 10/28 (36%) narrators who have undergone ostomies followed by ulcerative colitis (3/28, 11%), cancer (3/28, 11%), and unspecified inflammatory bowel disease (1/28). Further, 2 (7%) narrators who have undergone ostomies reported other underlying conditions for their ostomy surgery. Per the type of ostomy, ileostomies accounted for 79% (22/28) and colostomies accounted for 14% (4/28). Furthermore, 2 narrators who have undergone ostomies did not report their type of ostomy and no urostomy videos met inclusion and exclusion criteria for final review.

Appliance Change Frequency

More than half of narrators who have undergone ostomies (16/28, 57%) did not discuss appliance change frequency. For the remaining narrators who have undergone ostomies (12/28, 43%), there was considerable variability in ostomy appliance change frequency. In general, a majority of narrators who have undergone ostomies (n=10) reported changing their appliance within a 2-5-day interval, 2 every 7 and 15 days, and 1 when the pouch is full. Further, 1 narrator reported various changing frequencies across 4 videos (2-3 days, 3-4 days, every 4 days, and 4-5 days). Additionally, 2 narrators who have undergone ostomies commented on changing the appliance more frequently when having peristomal skin issues (every day or every other day) and emptying the ostomy appliance 5 to 8 times a day.

Common Challenges

About one-third of the content creators noted specific issues with supplies or processes that required workarounds or techniques to be implemented. Major challenges included peristomal skin complications (21/65 videos, 32%), leaks around the adhesive area (20/65 videos, 31%), inadequate appliance adhesion (13/65 videos, 20%), and no control over stoma effluent discharge making it difficult to keep the stoma and peristomal skin clean while changing the ostomy appliance (13/65 videos, 20%). Within the videos, the narrators who have undergone ostomies reported peristomal skin complications related to skin irritation and redness due to the output leakage, peristomal skin disorders such as granulomas, or allergic reactions toward specific ostomy supplies and these complications subsequently compromised appliance adhesion, caused pain, and led to further leaks.

Narrators who have undergone ostomies also expressed difficulties related to supplies in 17/65 (26%) videos. For example, supplies leaving residues on the skin compromising appliance adhesion, supplies containing alcohol causing a burning sensation on the wounded skin, variability in supplies' names causing confusion, and supplies not delivered on time or not working properly for a given individual. Challenges with the design of the ostomy appliance were noted, including the clip of the pouch being uncomfortable, difficulty with cleaning the end of the pouch after emptying, moisture absorption by the barrier ring leading to inadequate adhesion, or a 2-piece appliance being more difficult to apply. Finally, pouch pancaking with the output being trapped at the top of the pouch, pouch ballooning due to the gas accumulating in the ostomy appliance, cutting the baseplate to fit the stoma size, emptying an ostomy appliance and its cleaning, lifestyle adjustments (clothing, diet, relationships, intimacy, travelling, etc), and odor were additional challenges discussed in the videos.

Expert Tips and Tricks

The narrators who have undergone ostomies identified various strategies to address the challenges encountered. To overcome peristomal skin complications, narrators who have undergone ostomies (1) emphasized the importance of cleaning peristomal skin (eg, shower without an appliance and using adhesive remover to remove any output residue which could further irritate the skin); (2) avoided supplies containing alcohol to minimize burning sensations on the skin; (3) used a crusting technique which involves the application of stoma powder over sore skin followed by moistening the powder with skin barrier (wipe or spray), letting the powder dry, and repeating the process multiple times [16]; (4) used ostomy paste, or barrier ring to protect and allow healing of the wounded skin by creating a protective barrier between the wounded skin and appliance, or implemented more aggressive skin treatment products such as steroids or silver nitrate based on the consultation with ostomy care nurse specialists; and (5) recommended balance between changing an appliance too frequently or too rarely.

The tips of narrators who have undergone ostomies to avoid leakage included (1) properly fitting the appliance baseplate with the sizing of the opening for the stoma not to be too tight or too big; (2) using a barrier ring or ostomy paste to fill any abdominal gaps and divots; (3) using an ostomy belt, extenders, or medical tape to support the weight of the pouch; (4) trying different supplies (eg, convex pouches for retracted stomas) and different suppliers to identify supplies working properly for a given individual, while also emphasizing not using too many supplies; and (5) emptying an appliance regularly.

Narrators who have undergone ostomies believed that appliance adhesion was improved by (1) cleaning the skin properly using an adhesive remover followed by cleaning with water or some type of wipe to avoid leaving any residues that could compromise adhesion, (2) ensuring the peristomal skin is dry before applying an appliance, (3) wearing an ostomy belt, (4)

XSL•FO
RenderX

using an equalizer ring that applies an equal pressure around the stoma or ostomy paste that fills any divots or gaps between appliance and abdomen, (5) applying an appliance on a flat abdominal surface (eg, leaning backward or standing), (6) massaging the baseplate and running a finger around the stoma, and (7) using heat to stimulate baseplate adhesion to the abdomen (eg, hairdryer on low setting, heating pad, or holding hand over pouch). Narrators who have undergone ostomies also acknowledged there is not only one proper way of performing ostomy self-care and that what works for them may not work for someone else. Thus, they recommended trying different supplies and self-care strategies and consulting with ostomy nursing care specialists about supplies selection and their use for skin care.

Concerning managing a frequent and uncontrollable stoma output, some narrators who have undergone ostomies found it effective to eat marshmallows or to fast before changing an appliance to slow down the output, covering the stoma with some type of wipe, paper towel, or gauze while changing the appliance to protect the skin and make it easier to clean, and organizing or preparing all supplies and emptying the pouch before starting to change the appliance.

Narrators who have undergone ostomies conveyed strategies to prevent "pancaking" such as using a lubricant or some type of oil inside the pouch, emptying and rinsing the pouch with water, disabling an appliance filter, and blowing air into the pouch. Further, they used deodorants and "burped" (let the air out of the pouch) their appliance to prevent ballooning. They used an ostomy cover and deodorizing drops along with eating naturally deodorizing food such as kefir, yoghurt, or parsley to deal with the odor. To overcome the challenge of cutting a baseplate to the proper size, narrators who have undergone ostomies traced the stoma size with a marker on the template and saved the template with a marking for repeated use, or ordered the appliance precut from suppliers once their stoma size remained constant. Strategies used to adjust to the postostomy lifestyle included (1) finding and joining a support group, (2) talking to people who "get it," (3) learning about their stomas (eg, diversity of supplies, proper diet and hydration, or clothing tips), (4) exercising to reduce anxiety, and (5) developing an ability to perform ostomy self-care while also having someone else who can change their appliance.

Discussion

Principal Findings

This study uses a novel approach for gaining insight into ostomy care end user (person who has undergone an ostomy) experience and usability with medical device (ostomy appliance change) technology through the identification of strategies and supplies used, and common challenges and needs experienced by people who have undergone ostomies while changing the ostomy appliance and performing self-care. The needs, supply choices, and self-care strategies that people who have undergone ostomies used and discussed reflected the need to address a common challenge of peristomal skin complications following an ostomy surgery [21-23] and demonstrated characteristics of an expert understanding [24-27].

YouTube as Ostomy Self-Care Resource

This study examined YouTube videos related to the interaction of people who have undergone ostomies with medical device technology and self-care that were provided by and for the public, and the need for additional information related to long-term ostomy self-care in home (nonhospital) settings. In our study, we initially identified 960 ostomy-related videos and analyzed 65 videos with a narrator who has undergone an ostomy that met our inclusion criteria. Each of these videos was viewed 269,000 times on average, liked 3000 times on average, and posted by 28 unique content creators representing a broad range of ages and both, female and male genders. Our study supports current research suggesting that individuals are increasingly seeking health-related information on the web and patients use web-based resources to make health-related decisions and to manage chronic conditions [1-3,35,36]. Patient education and engagement in ostomy self-care are essential for performing self-care tasks and successful management of their chronic condition [18-20,35]. Nevertheless, research evidence shows the pre- and postoperative education provided to people who have undergone ostomies is insufficient [22-25] and new people who have undergone ostomies may not be prepared to gain new knowledge, skills, and attitudes necessary for self-care management of their ostomy immediately post surgery [26,27]. Therefore, our study suggests patients are looking for additional support and resources for ostomy self-care.

The video metadata and data of narrators who have undergone ostomies show that people who have undergone ostomies of various ages and genders are willing to share their personal experiences with ostomy self-care and appliance change procedures on the web and that these videos are viewed by others. Through YouTube, people who have undergone ostomies can share their stories, empower and help other people who have undergone ostomies, and develop a sense of community with other people who have undergone ostomies [10,11], which goes beyond the ability of health care providers not living with a stoma. In our study, narrators who have undergone ostomies created videos depicting ostomy self-care not to substitute health care professional guidelines and recommended practices, but rather to share a lived experience with a chronic illness, and their challenges and wins. They also commented on being able to develop a sense of community, learn from the lived experiences of other people who have undergone ostomies, and adjust to postostomy surgery lifestyle changes such as by finding and joining a support group or talking to people who "get it." Many of the narrators who have undergone ostomies made disclosures that they were not giving medical advice and commented on positive experiences with and encouraged interaction with ostomy care nurse specialists. The ostomy self-care-related videos depicted what it is like living life with a stoma, and what self-care strategies were found to be effective and what were not. This type of data would be difficult to gain using traditional observational techniques and surveys.

End User Experience With Ostomy Medical Device Technology

The narrators who have undergone ostomies used a variety of supplies at different frequencies to change their ostomy

XSL•FO
RenderX

appliance and discussed challenges related to supplies, challenges with peristomal skin health, appliance adhesion, leakage, and not having control over the stoma output which reflects their user needs. For example, peristomal skin complications were identified by narrators who have undergone ostomies in approximately one-third of analyzed ostomy self-care YouTube videos, and stoma powder (used to treat skin complications) and skin treatments were included and discussed in 12 (34%) and 8 (23%) of the videos depicting an ostomy appliance change procedures, respectively. This finding aligns with extensive research on peristomal skin complications following an ostomy surgery representing a common challenge [30,37,38] and negatively impacting the quality of life of people who have undergone ostomies [39]. The use of peristomal skin protection supplies such as barrier rings, skin barriers, and ostomy paste in 17 (49%), 15 (43%), and 6 (17%) of videos, respectively, and adhesion supplies such as extenders, ostomy belts, and skin adhesives in 5 (14%), 4 (11%), and 3 (9%) of the videos, respectively, manifests the need for properly fitting and adhering external pouching system [17,40]. Thus, the selection of supplies may have been associated with attempts to address appliance adhesion and leakage issues and thus peristomal skin complications.

Narrators who have undergone ostomies reported usability issues with supplies, such as the difficulty with out-of-package baseplate not adhering to "warm" body temperature skin, supplies that are effective in cleaning but compromise peristomal skin health and ostomy appliance adhesion and cause pain (eg, alcohol-based ostomy paste), which call for design improvements of ostomy supplies or redesign related to their properties (eg, baseplate adhesion to peristomal skin, or alcohol-free supplies).

Moldable technology is believed to improve peristomal skin health, reduce the incidence of irritant dermatitis, and be well-perceived by people who have undergone ostomies for its ease of use, ease of learning, and comfort [41,42]. However, in our study, only 3 (9%) of videos depicting the full or part of the ostomy appliance change procedure included a moldable baseplate compared to 25 (71%) of videos where the baseplate was cut to fit or precut and people who have undergone ostomies reported challenges with cutting the baseplate to size. If the stoma hole is too small, the stoma can be damaged, if the stoma hole is too big, the output is more likely to leak on the peristomal skin which in turn will lead to peristomal skin complications. Generally, there are 3 designs of baseplate, precut (cut by supplier), cut-to-size (cut by a person who has undergone an ostomy), and moldable. Moldable baseplate technology does not require cutting the baseplate with scissors but rather using fingers to roll flexible material to size and around ("turtle necking") the stoma and allows for adjusting the stoma hole to fit snuggly around the stoma while also creating a protective barrier and preventing output from reaching peristomal skin [41]. The narrators who have undergone ostomies did not elaborate on the reason for cutting versus molding baseplate to size, except for 1 video where the narrator who has undergone an ostomy had her husband change the appliance and commented on the molding being "easier to use at first." Therefore, further studies should be conducted to explain this trend and to more specifically identify user needs. Additionally, future work should determine if people who have undergone ostomies are aware of and educated on the benefits of using moldable baseplate technology versus using cut-to-fit or precut appliances and if moldable appliances are widely available to people who have undergone ostomies.

Experts' User-Generated Ostomy Self-Care Strategies

The tips and tricks discussed by the narrators who have undergone ostomies in the videos we examined provide evidence of problem-solving strategies useful for decision-making and ostomy self-management. Ostomy self-care requires complex daily decision-making to solve ostomy-related challenges such as compromised appliance adhesion, leaks, or peristomal skin complications discussed above. People who have undergone ostomies have to monitor their daily output, diet, and stomal or peristomal skin health; perform ostomy appliance changes; and plan future self-management. In previous studies of professional expertise, experts stood out in their ability to identify problems [43], understand functional relationships [44], and use problem-solving strategies [45].

Lippa et al [46] suggest that in the self-management of chronic conditions, such as diabetes, professional expertise may be useful for aiding decision-making and supporting patient's strategies. Lippa et al [46] also found that diabetic patients who used functional knowledge and problem-solving strategies reported higher levels of adherence to treatment and glycemic control and that descriptive knowledge about diabetes may not necessarily correlate with effective self-management [46]. In our study, narrators who have undergone ostomies demonstrated their expert understanding while elaborating and showing their self-care strategies and adaptation to ostomy-related challenges. These strategies relate to supplies use and organization, ostomy appliance adhesion, baseplate sizing, maintenance of healthy peristomal skin, pouch pancaking and ballooning, and adjustments to lifestyle changes.

Finally, our study of ostomy self-care and end user experiences with ostomy appliance change technology complements the "Ostomy Life Study," [47] a global effort to gain knowledge about what is it like to live with a stoma and the user experience with ostomy-related supplies. Our study identifies an efficient alternative to traditional observational and survey methods for obtaining end user data that have implications for addressing patient safety and quality of care issues.

Limitations

Several limitations should be considered when interpreting the results of this study and guiding the direction of future research. In our sample of videos, there were more ileostomies than colostomies, and there were no urostomies, therefore, our findings did not compare different types of ostomies. Further, the narrators who have undergone ostomies were essentially "performing" and demonstrating for the video, therefore, these videos may not capture the routine procedures people who have undergone ostomies use for self-care. In our data analysis, half of the narrators who have undergone ostomies were within an assumed age of 20-30 years and approximately one-third of narrators who have undergone ostomies were in their middle

XSL•FO
RenderX

age. This is likely due to most YouTube users falling between the ages of 18 and 44 years [48].

There were also limitations related to demographics in this study. All videos analyzed were narrated in English, but the global representation is unknown. This may limit the global generalizability of this study. Additionally, we did not examine any race, gender, or ethnicity differences in the experiences of people who have undergone ostomies.

Future Work

Further, it should examine differences in ostomy appliance change procedure and self-care strategies across all the types of ostomies, more specifically, whether ileostomies are more prone to peristomal skin complications considering ileostomies' output is frequent, more liquid, and high in digestive enzymes [32]; and the prevalence and decision-making of using a moldable versus cut-to-fit appliance technology among people who have undergone ostomies. Future research should also specifically study younger and older people who have undergone ostomies and their experiences, complement this work in other languages and for individuals with experiences in other health care systems, and examine if there are differences among various populations of people who have undergone ostomies.

We identified 28 individual narrators who have undergone ostomies, and their videos were widely viewed. In addition to counting likes, future research should examine who is viewing these videos, what they are learning from them, and the content of viewer comments and discussions related to the video content. The comments analysis could provide a rich discussion around the strategies and ostomy self-care provided in the ostomy self-care videos.

In our study, we purposely did not include any video content from medical professionals or health care professional organizations because it can contain different information about ostomy self-care strategies than what is practiced by people who have undergone ostomies in home settings [35]. We intended to evaluate end user experiences with ostomy self-care rather than how people engage with YouTube videos created by health care professionals. However, we recognize there is value in comparing user-generated information to what clinicians and health-related government organizations may communicate about ostomy self-care as they use YouTube to disseminate health information [1] that is considered to be the most reliable and accurate information [6]. Future research should examine the content in the ostomy self-care videos created by health care professionals and professional organizations and evaluate how they compare to the user-generated ostomy self-care YouTube videos examined in this study.

Conclusions

This study used a novel approach for ostomy care (YouTube videos) to gain insights into ostomy self-care in home settings and end user interactions with medical devices while performing self-care, which are difficult to gather using traditional behavioral techniques. Using YouTube or other publicly available content created by individuals dealing with a chronic condition is a low-cost means to identify successful strategies, potential design problems, and potential workaround for issues that would be difficult to obtain in other study designs and approaches. The analysis of 65 videos showed that people who have undergone ostomies are willing to share their personal experience with ostomy self-care on the web and that these videos are viewed by the public. The findings also identify user needs related to medical device technology and indicate that narrators who have undergone ostomies were able to use problem-solving strategies for decision-making and successful ostomy self-management. This study demonstrates the potential value of these videos for filling the gap between health information delivered by health care professionals and the lived experience of individuals dealing with a chronic condition.

Additional information related to long-term ostomy self-care in a home (nonhospital) setting and the differences between what is clinically recommended and how the procedures and supplies are actually used by individuals living with ostomies needs to be more clearly understood for the design and testing of future ostomy supplies as well as for patient education from clinicians. Future research could use these approaches to address other types of clinical information that other clinical populations need to learn or be able to access when at home as well as strategies to design better medical devices for home health care in the future.

Acknowledgments

This work was supported by a Transformative Seed Grant from the Prisma Health System Research Division, a Center of Biomedical Research Excellence (SC TRIMH) funded by NIH/NIGMS (P20GM121342), and institutional funding from the Creative Inquiry Program at Clemson University.

Conflicts of Interest

None declared.

References

1. Gooding LF, Gregory D. Descriptive analysis of YouTube music therapy videos. J Music Ther 2011;48(3):357-369. [doi: 10.1093/jmt/48.3.357] [Medline: 22097103]
2. Steinberg PL, Wason S, Stern JM, Deters L, Kowal B, Seigne J. YouTube as source of prostate cancer information. Urology 2010 Mar;75(3):619-622. [doi: 10.1016/j.urology.2008.07.059] [Medline: 19815255]

XSL•FO
RenderX

3. Madathil KC, Rivera-Rodriguez AJ, Greenstein JS, Gramopadhye AK. Healthcare information on YouTube: a systematic review. Health Informatics J 2015 Sep;21(3):173-194 [FREE Full text] [doi: 10.1177/1460458213512220] [Medline: 24670899]
4. Ceci L. Statista. 2024 Apr. URL: https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/ [accessed 2025-02-19]
5. Altunisik E, Firat YE, Kiyak Keceli Y. Content and quality analysis of videos about multiple sclerosis on social media: the case of YouTube. Mult Scler Relat Disord 2022 Sep;65:104024. [doi: 10.1016/j.msard.2022.104024] [Medline: 35841755]
6. Basim P, Argun D. A Qualitative Analysis of Ostomy-Related Patient Education Videos on YouTube. Adv Skin Wound Care 2021 Jun 01;34(6):314-320. [doi: 10.1097/01.ASW.0000744340.11915.93] [Medline: 33979820]
7. Cetin A. Evaluation of YouTube video content related to the management of hypoglycemia. Cureus 2021 Jan 06;13(1):e12525 [FREE Full text] [doi: 10.7759/cureus.12525] [Medline: 33564529]
8. Ruppert L, Køster B, Siegert AM, Cop C, Boyers L, Karimkhani C, et al. YouTube as a source of health information: analysis of sun protection and skin cancer prevention related issues. Dermatol Online J 2017 Jan 15;23(1):1-10 [FREE Full text] [doi: 10.5070/D3231033669] [Medline: 28329465]
9. Keelan J, Pavri-Garcia V, Tomlinson G, Wilson K. YouTube as a source of information on immunization: a content analysis. JAMA 2007 Dec 05;298(21):2482-2484. [doi: 10.1001/jama.298.21.2482] [Medline: 18056901]
10. Frohlich DO, Zmyslinski-Seelig A. The presence of social support messages on YouTube videos about inflammatory bowel disease and ostomies. J Health Commun 2012;27(5):421-428. [doi: 10.1080/10410236.2011.606524] [Medline: 21962112]
11. Gómez-Zúñiga B, Fernandez-Luque L, Pousada M, Hernández-Encuentra E, Armayones M. ePatients on YouTube: analysis of four experiences from the patients' perspective. Med 2 0 2012;1(1):e1 [FREE Full text] [doi: 10.2196/med2.2039] [Medline: 25075229]
12. Tase A, Vadhwana B, Buckle P, Hanna GB. Usability challenges in the use of medical devices in the home environment: A systematic review of literature. Appl Ergon 2022 Sep;103:103769 [FREE Full text] [doi: 10.1016/j.apergo.2022.103769] [Medline: 35430424]
13. Dul J, Bruder R, Buckle P, Carayon P, Falzon P, Marras WS, et al. A strategy for human factors/ergonomics: developing the discipline and profession. Ergonomics 2012;55(4):377-395 [FREE Full text] [doi: 10.1080/00140139.2012.661087] [Medline: 22332611]
14. Bitkina OV, Kim HK, Park J. Usability and user experience of medical devices: An overview of the current state, analysis methodologies, and future challenges. Int J Ind Ergon 2020 Mar;76:102932 [FREE Full text] [doi: 10.1016/j.ergon.2020.102932]
15. Davis BR, Valente MA, Goldberg JE, Lightner AL, Feingold DL, Paquette IM, Prepared on behalf of the Clinical Practice Guidelines Committee of the American Society of ColonRectal Surgeons. The American Society of Colon and Rectal Surgeons clinical practice guidelines for ostomy surgery. Dis Colon Rectum 2022 Oct 01;65(10):1173-1190. [doi: 10.1097/DCR.0000000000002498] [Medline: 35616386]
16. United Ostomy Associations of America. United Ostomy Associations of America. New ostomy patient guide. 2024 Apr. URL: https://www.ostomy.org/wp-content/uploads/2024/04/UOAA-New-Ostomy-Patient-Guide-2024-04.pdf [accessed 2025-02-19]
17. Herlufsen P, Olsen AG, Carlsen B, Nybaek H, Karlsmark T, Laursen TN, et al. Study of peristomal skin disorders in patients with permanent stomas. Br J Nurs 2006;15(16):854-862. [doi: 10.12968/bjon.2006.15.16.21848] [Medline: 17108855]
18. Krouse RS, Grant M, McCorkle R, Wendel CS, Cobb MD, Tallman NJ, et al. A chronic care ostomy self-management program for cancer survivors. Psychooncology 2016 May;25(5):574-581 [FREE Full text] [doi: 10.1002/pon.4078] [Medline: 26804708]
19. Millard R, Cooper D, Boyle MJ. Improving self-care outcomes in ostomy patients via education and standardized discharge criteria. Home Healthc Now 2020;38(1):16-23. [doi: 10.1097/NHH.0000000000000816] [Medline: 31895893]
20. Prinz A, Colwell JC, Cross HH, Mantel J, Perkins J, Walker CA. Discharge planning for a patient with a new ostomy: best practice for clinicians. J Wound Ostomy Continence Nurs 2015;42(1):79-82. [doi: 10.1097/WON.0000000000000094] [Medline: 25333690]
21. Wound, Ostomy Continence Nurses Society Task Force. Wound, ostomy, and continence nursing: scope and standards of WOC practice, 2nd edition: an executive summary. J Wound Ostomy Continence Nurs 2018;45(4):369-387. [doi: 10.1097/WON.0000000000000438] [Medline: 29698290]
22. Gemmill R, Kravits K, Ortiz M, Anderson C, Lai L, Grant M. What do surgical oncology staff nurses know about colorectal cancer ostomy care? J Contin Educ Nurs 2011 Feb;42(2):81-88. [doi: 10.3928/00220124-20101101-04] [Medline: 21053790]
23. Tallman NJ, Cobb MD, Grant M, Wendel CS, Colwell J, Ercolano E, et al. Colostomy irrigation: issues most important to wound, ostomy and continence nurses. J Wound Ostomy Continence Nurs 2015;42(5):487-493. [doi: 10.1097/WON.0000000000000170] [Medline: 26336046]
24. Celik SS, Tuna Z, Yildirim M. The experience of urostomists who do not have access to pre‐operative and post‐operative stoma care nursing intervention. Int J of Uro Nursing 2014 Dec 29;9(2):101-107 [FREE Full text] [doi: 10.1111/ijun.12068]

XSL•FO
RenderX

25. Braumann C, Müller V, Knies M, Aufmesser B, Schwenk W, Koplin G. Quality of life and need for care in patients with an ostomy: a survey of 2647 patients of the Berlin Ostomy-Study (BOSS). Langenbecks Arch Surg 2016 Dec;401(8):1191-1201. [doi: 10.1007/s00423-016-1507-z] [Medline: 27659022]
26. Metcalf C. Stoma care: empowering patients through teaching practical skills. Br J Nurs 1999;8(9):593-600. [doi: 10.12968/bjon.1999.8.9.6621] [Medline: 10711005]
27. White T, Watts P, Morris M, Moss J. Virtual postoperative visits for new ostomates. Comput Inform Nurs 2019 Feb;37(2):73-79. [doi: 10.1097/CIN.0000000000000498] [Medline: 30562169]
28. Nichols TR, Inglese GW. The burden of peristomal skin complications on an ostomy population as assessed by health utility and the physical component summary of the SF-36v2. Value Health 2018 Jan;21(1):89-94 [FREE Full text] [doi: 10.1016/j.jval.2017.07.004] [Medline: 29304945]
29. Nybaek H, Knudsen DB, Laursen TN, Karlsmark T, Jemec GBE. Quality of life assessment among patients with peristomal skin disease. Eur J Gastroenterol Hepatol 2010 Feb;22(2):139-143. [doi: 10.1097/MEG.0b013e32832ca054] [Medline: 19918185]
30. Hendren S, Vu J, Suwanabol P, Kamdar N, Hardiman K. Hospital variation in readmissions and visits to the emergency department following ileostomy surgery. J Gastrointest Surg 2020 Nov;24(11):2602-2612 [FREE Full text] [doi: 10.1007/s11605-019-04407-6] [Medline: 31754986]
31. LeBlanc K, Whiteley I, McNichol L, Salvadalena G, Gray M. Peristomal medical adhesive-related skin injury: results of an international consensus meeting. J Wound Ostomy Continence Nurs 2019;46(2):125-136 [FREE Full text] [doi: 10.1097/WON.0000000000000513] [Medline: 30844869]
32. Lyon CC, Smith AJ, Griffiths CEM, Beck MH. The spectrum of skin disorders in abdominal stoma patients. Br J Dermatol 2000 Dec;143(6):1248-1260. [doi: 10.1046/j.1365-2133.2000.03896.x] [Medline: 11122029]
33. Salvadalena G. Incidence of complications of the stoma and peristomal skin among individuals with colostomy, ileostomy, and urostomy: a systematic review. J Wound Ostomy Continence Nurs 2008;35(6):596-607; quiz 608. [doi: 10.1097/01.WON.0000341473.86932.89] [Medline: 19018200]
34. Hill CE, Knox S. Essentials of consensual research. In: American Psychological Association. Washington, DC: American Psychological Association; 2021.
35. Silva BWACD, O LBD, Araújo AKD, Medeiros MBC, Melo VLD, Sena JFD, et al. Analysis of self-care videos on YouTube about exchange of intestinal ostomy bags. Rev Rene 2020;21:e44275. [doi: 10.15253/2175-6783.20202144275]
36. Yin R, Neyens DM. Online health resource use by individuals with inflammatory bowel disease: analysis using the national health interview survey. J Med Internet Res 2020 Sep 24;22(9):e15352 [FREE Full text] [doi: 10.2196/15352] [Medline: 32969831]
37. Sheetz KH, Waits SA, Krell RW, Morris AM, Englesbe MJ, Mullard A, et al. Complication rates of ostomy surgery are high and vary significantly between hospitals. Dis Colon Rectum 2014 May;57(5):632-637 [FREE Full text] [doi: 10.1097/DCR.0000000000000038] [Medline: 24819104]
38. Taneja C, Netsch D, Rolstad BS, Inglese G, Eaves D, Oster G. Risk and economic burden of peristomal skin complications following ostomy surgery. J Wound Ostomy Continence Nurs 2019;46(2):143-149 [FREE Full text] [doi: 10.1097/WON.0000000000000509] [Medline: 30844870]
39. Baxter NN, Novotny PJ, Jacobson T, Maidl LJ, Sloan J, Young-Fadok TM. A stoma quality of life scale. Dis Colon Rectum 2006 Feb;49(2):205-212. [doi: 10.1007/s10350-005-0275-6] [Medline: 16392025]
40. Arumugam PJ, Bevan L, Macdonald L, Watkins AJ, Morgan AR, Beynon J, et al. A prospective audit of stomas-analysis of risk factors and complications and their management. Colorectal Dis 2003 Jan;5(1):49-52. [doi: 10.1046/j.1463-1318.2003.00403.x] [Medline: 12780927]
41. Szewczyk MT, Majewska G, Cabral MV, Hölzel-Piontek K. The effects of using a moldable skin barrier on peristomal skin condition in persons with an ostomy: results of a prospective, observational, multinational study. Ostomy Wound Manage 2014 Dec;60(12):16-26 [FREE Full text] [Medline: 25485549]
42. Liu G, Chen Y, Luo J, Liu A, Tang X. The application of a moldable skin barrier in the self-care of elderly ostomy patients. Gastroenterol Nurs 2017;40(2):117-120. [doi: 10.1097/SGA.0000000000000143] [Medline: 26987102]
43. Dreyfus H, Drey-fus SE. Mind over Machine: The Power of Human Intuition and Expertise in the Era of the Computer. New York: The Free Press; Jun 1987:110-111.
44. Chi MTH, Feltovich PJ, Glaser R. Categorization and representation of physics problems by experts and novices. Cognitive Science 1981;5(2):121-152. [doi: 10.1207/s15516709cog0502_2]
45. Klein G. Sources of Power: How people make decisions. United States: The MIT Press; 1999.
46. Lippa KD, Klein HA, Shalin VL. Everyday expertise: cognitive demands in diabetes self-management. Hum Factors 2008 Feb;50(1):112-120. [doi: 10.1518/001872008X250601] [Medline: 18354975]
47. Coloplast. Ostomy Life Study. Coloplast Corp. URL: https://www.coloplast.us/landing-pages/ols [accessed 2025-02-19]
48. GMI Research Team. GLOBAL MEDIA INSIGHT - Dubai Digital Interactive Agency. URL: https://www.globalmediainsight.com/blog/youtubeusersstatistics/#YouTube_Users_by_Year [accessed 2025-02-19]

XSL•FO
RenderX

Edited by A Kushniruk; submitted 18.09.24; peer-reviewed by C Milne, C Ruiz-Nunez; comments to author 07.11.24; revised version received 06.12.24; accepted 20.12.24; published 06.03.25.

Please cite as:
Haughey M, Neyens DM, Hopkins CS, Gonzaga C, Harman M
Identifying Strategies for Home Management of Ostomy Care: Content Analysis of YouTube
JMIR Hum Factors 2025;12:e66634
URL: https://humanfactors.jmir.org/2025/1/e66634
doi:10.2196/66634
PMID:40053741

©Marketa Haughey, David M Neyens, Casey S Hopkins, Christofer Gonzaga, Melinda Harman. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 06.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Development of a Patient-Centered Symptom-Reporting Application in Pharmacy Settings Using a Hierarchical Patient-Friendly Symptom List: Developmental and Usability Study

Seiya Watanabe[1], BSc; Hayato Kizaki[1], PhD; Satoko Hori[1], PhD

Division of Drug Informatics, Keio University Faculty of Pharmacy, Tokyo, Japan

Corresponding Author:
Satoko Hori, PhD
Division of Drug Informatics
Keio University Faculty of Pharmacy
1-5-30 Shibakoen, Minato-ku
Tokyo, 105-8512
Japan
Phone: 81 354002650
Email: satokoh@keio.jp

Abstract

Background: Effective symptom identification, a key responsibility for community pharmacists, requires patients to describe their symptoms accurately and comprehensively. However, current practices in pharmacies may be insufficient in capturing patient-reported symptoms comprehensively, potentially affecting the quality of pharmaceutical care and patient safety.

Objective: This study aimed to construct a new, hierarchical symptom list derived from the Patient-Friendly Term List of the Medical Dictionary for Regulatory Activities (MedDRA) and to develop and evaluate a mobile app incorporating this list for facilitating symptom reporting by patients in pharmacy settings. The study also aimed to assess the usability and acceptance of this app among potential users.

Methods: Subjective symptom-related terms were extracted from the Patient-Friendly Term List version 23.0 of the MedDRA. These terms were systematically consolidated and organized into a hierarchical, user-friendly symptom list. A mobile app incorporating this list was developed for pharmacy settings, featuring a symptom selection interface and a free-text input field for additional symptoms. The app included an instructional video explaining the importance of symptom reporting and guidance on navigation. Usability tests and semistructured interviews were conducted with participants aged >20 years. Interview transcripts were analyzed using the Unified Theory of Acceptance and Use of Technology (UTAUT) model to evaluate factors influencing the acceptance of technology.

Results: From the initial 1440 terms in the Patient-Friendly Term List, 795 relevant terms were selected and organized into 40 site-specific subcategories, which were then grouped into broader site categories (mental, head, trunk, upper limb, lower limb, physical condition, and others). These terms were further consolidated into 211 patient-friendly symptom terms, forming a hierarchical symptom list. The app's interface design limited options to 10 items per screen to assist with decision-making. A total of 5 adults participated in the usability test. Participants found the interface intuitive and easy to use, requiring minimal effort, and provided positive feedback regarding the potential utility of the app in pharmacy settings. The UTAUT analysis identified several facilitating factors, including ease of use and the potential for enhanced pharmacist-patient communication. However, concerns were raised about usability for older adults and the need for simplified technical terminology.

Conclusions: The user-friendly app with a hierarchically structured symptom list and complementary free-text entry has potential benefits for improving the accuracy and efficiency of symptom reporting in pharmacy settings. The positive user acceptance and identified areas for improvement provide a foundation for further development and implementation of this technology to enhance communication between patients and pharmacists. Future improvements should focus on addressing usability for older adults and simplifying technical terminology.

(JMIR Hum Factors 2025;12:e71439) doi:10.2196/71439

KEYWORDS

patient symptom monitoring; hierarchical symptom list; community pharmacy; interview survey; mobile application

XSL•FO
RenderX

Introduction

Patient-reported symptoms are critical indicators in health care, essential for maintaining patient safety and improving therapeutic outcomes in clinical settings [1,2]. In Japan, the Ministry of Health, Labour and Welfare launched "Pharmacy Visions for Patients" in 2015, clearly outlining the role of community pharmacies, including comprehensive monitoring of medications, as well as providing tailored drug management and guidance [3]. A key responsibility for community pharmacies in this context is to gather and monitor patient-reported symptoms, essential for identifying potential health issues and ensuring timely intervention.

For pharmacists, collecting patient-reported symptoms plays multiple critical roles in medication therapy management. First, it facilitates the early detection and monitoring of potential adverse drug reactions, enabling timely interventions to prevent serious complications. Second, comprehensive symptom information helps pharmacists assess therapeutic effectiveness and adjust medication regimens accordingly. Third, systematic symptom monitoring supports pharmacists in providing targeted patient education and improving medication adherence through a better understanding of patient experiences. These activities are fundamental to the pharmacist's role in ensuring medication safety and optimizing therapeutic outcomes.

While health care providers traditionally collect symptom information through direct questioning, a patient-centered approach that enables patients to report symptoms at their own pace using familiar terminology may improve the accuracy and comprehensiveness of symptom reporting. Such an approach acknowledges patients as active participants in their health care, potentially leading to better identification of health-related issues and more effective interventions. A cross-sectional study demonstrated that structured symptom reporting tools can help identify numerous patient-reported symptoms and their potential associations with medications, providing valuable information for medication reviews [4]. However, another previous study showed that patients often underreported symptoms to health care providers, either because they do not attribute symptoms to the medication or do not recognize the significance of the symptoms [5].

To facilitate patient reporting, various symptom questionnaires have been developed, enabling patients to describe their experiences [6]. However, many of these tools lack thorough validation. In Japan, Nojo et al [7] introduced the "Adverse Drug Reaction Signal Check Sheet," which lists subjective symptoms of side effects for patients taking high-risk drugs. Although limited to specific drugs, the sheet proved effective in prompting patients to communicate their symptoms during consultations with pharmacists. This approach suggests that selecting symptoms from a structured list can support patient reporting.

The Medical Dictionary for Regulatory Activities (MedDRA), a globally recognized dictionary of medical terminology, includes a Patient-Friendly Term List that reflects a wide range of symptoms reported by patients in safety databases [8]. MedDRA has a hierarchical structure with organ categories at the top, and its Patient-Friendly Term List is a supplementary list made up of the lowest-level terms of the MedDRA. This structured approach allows patients to progressively identify and report their symptoms. However, international research has indicated that patients do not fully use this list when reporting adverse drug reactions [9], suggesting challenges in the direct adoption of a symptom-reporting system.

This study aimed to construct a new hierarchical list of symptoms derived from the Patient-Friendly Term List of the MedDRA to simplify and facilitate incremental symptom selection by patients. Moreover, we aimed to develop an app incorporating this list to enhance communications between patients and pharmacists by facilitating more accurate symptom reporting.

Methods

Creation of the New Hierarchical, Patient-Friendly List of Symptoms

A hierarchical list of subjective symptoms was created using the Patient-Friendly Term List version 23.0 from the MedDRA. MedDRA, developed by the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use, is a standardized medical terminology used in pharmaceutical regulation. The MedDRA Patient-Friendly Term List was selected for this study due to its standardized structure and global adoption in pharmacovigilance. This list consists of patient-friendly terms derived from MedDRA, ensuring that the terminology is both medically accurate and comprehensible to nonexpert users. The structured format of this list was expected to facilitate consistent symptom categorization while remaining accessible to patients, supporting pharmacists in identifying potential adverse drug reactions (ADRs) and optimizing medication therapy management.

Terms associated with subjective symptoms were extracted from the Patient-Friendly Term List and categorized by body part. Similar terms were consolidated into "symptoms," forming the lowest level of a 3-tier hierarchical structure, with "broader site categories" and "site-specific subcategories" as the upper levels. This new hierarchical list was reviewed and refined through multiple rounds of consensus by three researchers (WS, KH, and HS). WS was an undergraduate student in the Faculty of Pharmaceutical Sciences, and KH and HS were researchers in the Faculty of Pharmaceutical Sciences.

Development of a Patient Symptom-Reporting Application

An app was developed to incorporate the newly created hierarchical list of symptoms. The app featured a symptom selection interface, an instructional video explaining the importance of symptom reporting, and guidance on how to navigate the app. To assist with decision-making, no more than 10 options were displayed on the screen at a time. In addition to the structured symptom list, the app included a complementary free-text entry, allowing users to input symptoms that may not be represented on the list or when the listed symptoms did not fit their condition. The app is intended for use in a pharmacy setting and was designed to capture a

XSL•FO
RenderX

wide range of symptoms, beyond those related to side effects of specific drugs. Apache Cordova, an open-source mobile app development framework, was used for the development. The operating system used was Android, and the app was run on a HUAWEI MediaPad T3. The content of the app was written in Japanese.

Usability Study and Interviews

A usability study was conducted from September to October 2020 with adults aged >20 years, recruited through research flyers. Participants first completed a questionnaire on their demographics, medication use, history of ADRs, and familiarity with digital devices. They were then briefed on the purpose of the app in a pharmacy setting and the interview format. Participants tested the app on a HUAWEI MediaPad T3, entering personal and symptom data on 5 hypothetical patient scenarios (Table S1 in Multimedia Appendix 1). Following the app trial, a semistructured interview was conducted, based on the interview guide (Table S2 in Multimedia Appendix 1), to gather feedback on the app. Interviews were conducted in Japanese, focusing on impressions of the app from participants and their awareness of reporting side effects. All interviews were recorded with the consent of participants and conducted by a single researcher (WS).

Analysis of the Interview

Thematic analysis was applied to the recorded interview transcripts using a deductive and theoretical approach. The Unified Theory of Acceptance and Use of Technology (UTAUT) model was used for this analysis. Initially proposed by Venkatesh et al [10] in 2003, the UTAUT model has been widely adopted in the medical field in recent years [11]. This model is particularly useful for understanding factors influencing the adoption of new technologies. It posits that performance expectancy (expected benefits of using new technology), effort expectancy (expectations of the ease of use and understanding of new technology), and social influence (how much the user's decisions about technology are influenced by others) determine behavioral intention to adopt the technology. Behavioral intention, along with facilitating conditions, influence actual technology use. In addition, gender, age, experience with similar technologies, and spontaneity of use serve as general adjustment variables (moderators). The translation of these constructs was based on the work of Ono [12]. In total, 3 researchers conducted analyses to ensure objectivity.

Ethical Considerations

On the day of the usability study and interview, the survey content, privacy protection, and plans for publishing the research results were comprehensively explained to the participants, and written informed consent to participate in the survey was obtained. The preinterview questionnaire was self-administered anonymously. Participants received compensation in the form of a gift card (Quo Card) valued at 1000 JPY for their participation in the study. This survey was conducted after being approved by the Ethics Committee for Research Involving Human Subjects, Keio University Faculty of pharmacy, following the Ethical Guidelines for Medical Research Involving Human Subjects (200213-3).

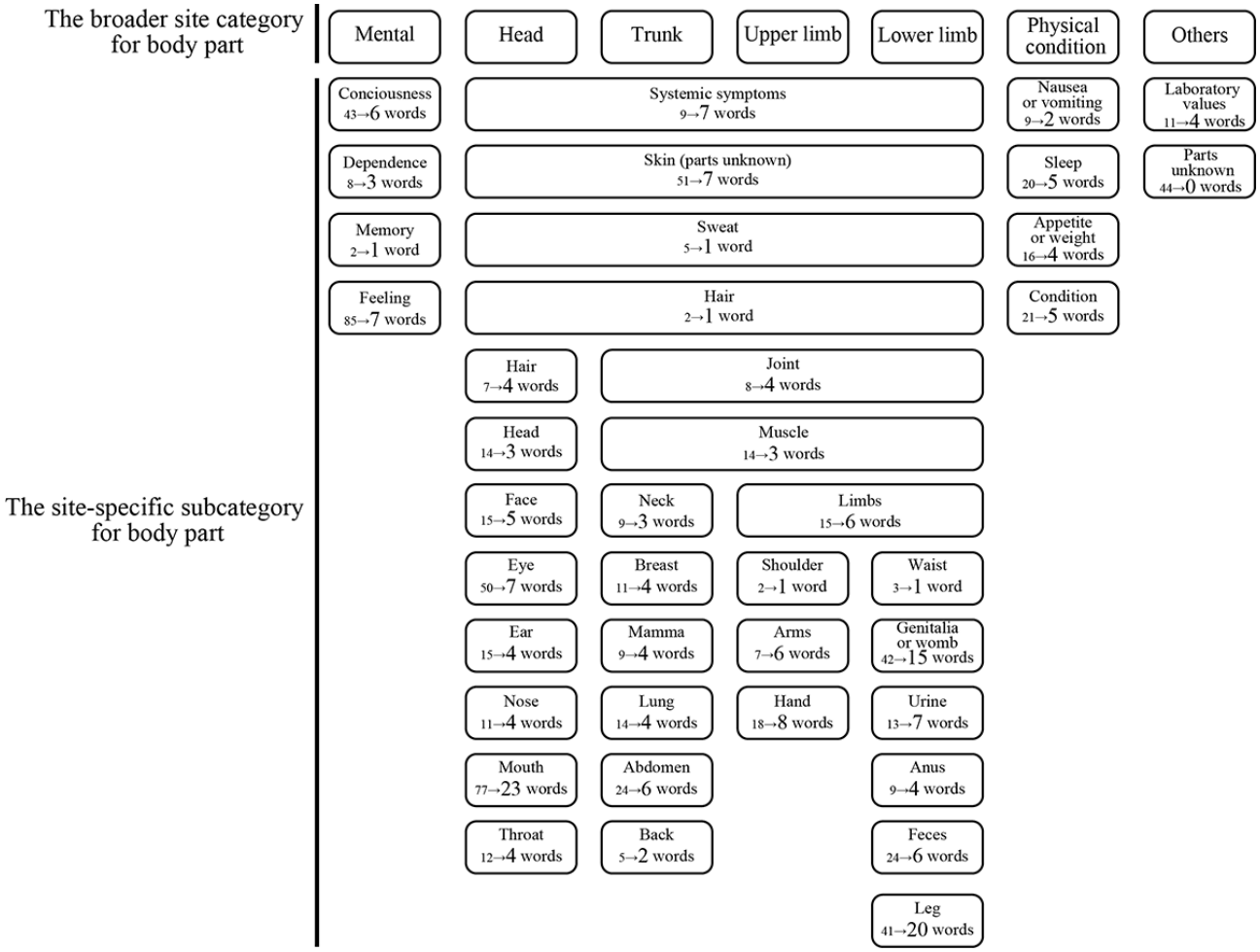
Results

Creation of the New Patient-Friendly Symptom List

From the 1440 terms in the Patient-Friendly Term List version 23.0, duplicates with the same Japanese translations were removed, resulting in 1288 terms. Subsequently, 12 product-related terms such as "suspected counterfeit products" were excluded. In addition, 450 terms that were difficult for patients to self-identify, including those requiring a medical diagnosis or those specific to certain conditions, were removed. Another 31 terms unrelated to medications were deleted, leaving a total of 795 terms. The 795 terms were systematically sorted into 40 site-specific subcategories based on information related to the locus of impact of each word. These site-specific subcategories were grouped into 7 broad site categories: mental, head, trunk, upper limb, lower limb, physical condition, and others (Figure 1). Within each site-specific subcategory, similar terms were consolidated by grouping related symptoms (eg, various types of arm discomfort such as "arm paralysis" and "arm discomfort") and creating standardized descriptive names (eg, "tingling and discomfort in the arm") that accurately represent the consolidated symptoms. This consolidation process aimed to simplify symptom selection while maintaining clinical relevance, yielding 211 subjective symptoms (Figure 1). Finally, a hierarchical list of patient-friendly symptoms was created using the hierarchical structure of site categories, subcategories, and symptoms.

XSL•FO
RenderX

Figure 1. The broader site categories and site-specific categories for body parts, as well as the word counts before and after consolidation. The 795 terms were systematically sorted into 40 site-specific subcategories based on information related to the locus of impact of each word. These site-specific subcategories were grouped into 7 broad site categories: mental, head, trunk, upper limb, lower limb, physical condition, and others. Within each site-specific subcategory, similar terms were consolidated by grouping related symptoms (eg, various types of arm discomfort such as "arm paralysis" and "arm discomfort") and creating standardized descriptive names (eg, "tingling and discomfort in the arm") that accurately represent the consolidated symptoms. This consolidation process aimed to simplify symptom selection while maintaining clinical relevance. In this figure, the number on the left side of each arrow represents the count of terms before consolidation and the number on the right indicates the count following this unification process.



To integrate this list into the app, we designed a question flow that allowed for easy symptom selection. Broader and site-specific subcategories were modified and recombined into major site categories and site subcategories for an intuitive question flow (Figure S1 in Multimedia Appendix 1). As an example, the hierarchy under the major site category "upper limbs" and its subcategories is detailed in Table 1. The app included a free-text entry field labeled as "other" at each categorical level to capture symptoms that users could not find in the structured list. This feature was implemented to ensure comprehensive symptom reporting and to collect data for future improvements of the symptom list.

XSL•FO
RenderX

Table 1. Details of selection items (excerpt for upper limbs only). The major site categories, site subcategories, lower-level terms, and symptoms after consolidation were described in Japanese. The table presented here has been translated into English for this publication.

Major site categories and subcategories of body parts (LLT[a])	LLT code	Symptoms after consolidation
Upper limb		
Shoulder		
Scapular pain	10040610	Shoulder pain
Shoulder pain	10040617	Shoulder pain
Arm		
Upper limb pain	10033421	Arm pain
Elbow pain	10033424	Elbow pain
Arm rash	10037875	Arm rash
Arm paralysis	10003098	Tingling and discomfort in the arm
Arm discomfort	10049877	Tingling and discomfort in the arm
Arm swelling	10042680	Arm swelling and edema
Arm weakness	10050379	Weakness in the arm
Hand		
Brittle nails	10006373	Abnormal nails
Nail discoloration	10028691	Abnormal nails
Finger swelling	10042694	Swelling and edema in fingers
Hand swelling	10042695	Swelling and edema in fingers
Hand rash	10019117	Hand rash
Cold hands	10009860	Cold sensation in hands
Finger pain	10033428	Hand pain
Hand pain	10033430	Hand pain
Itching of both hands	10023087	Itchy hands
Wrist pain	10048692	Wrist pain
Clumsiness	10009696	Tingling, trembling, and discomfort in fingers
Hand cramp	10011287	Tingling, trembling, and discomfort in fingers
Finger deformity	10061156	Tingling, trembling, and discomfort in fingers
Stiffness of fingers	10016695	Tingling, trembling, and discomfort in fingers
Tingling in fingers	10029837	Tingling, trembling, and discomfort in fingers
Tingling in hands	10049681	Tingling, trembling, and discomfort in fingers
Reduced dexterity	10067727	Tingling, trembling, and discomfort in fingers
Hand tremor	10040530	Tingling, trembling, and discomfort in fingers

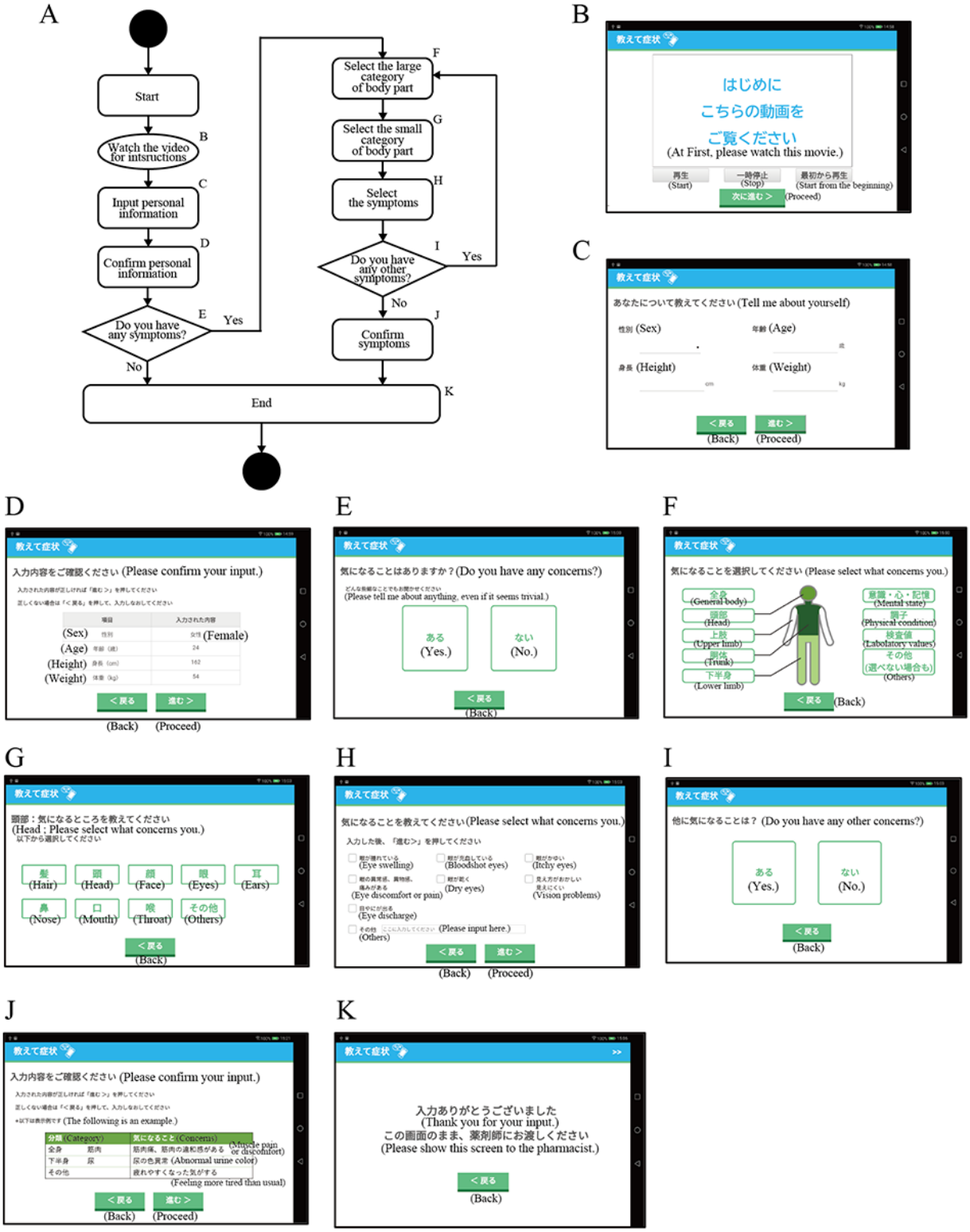
[a]LLT: lowest-level term.

Development of a Symptom-Reporting App

The screen transition diagram and display screens for the app are shown in Figure 2. Figure 2A shows the screen transition diagram of the app. The starting screen featured an instructional video explaining the purpose of the app and the importance of comprehensive symptom reporting (Figure 2B), followed by a series of intuitive selections for symptom identification (Figure 2E-J). The closed screen signaled the end of the session, prompting the user to pass the device on to the pharmacist (Figure 2K). To minimize cognitive overload, no more than 10 options were displayed per screen, making symptom selection easier and more manageable. An instructional video was included to explain the importance of symptom reporting, providing users with essential background knowledge before using the app. In addition, a free-text entry field was implemented to allow patients to describe symptoms not covered by the structured list, ensuring flexibility while maintaining structured reporting as the primary method. These features were designed to enhance the accuracy and completeness of symptom reporting, thereby improving communication between patients and pharmacists.

XSL•FO
RenderX

Figure 2. Screen transition diagram and display screen of the app. (A) Screen transition. (B) Start and video viewing section. (C) Input of personal information, (D) Confirmation of personal information. (E) Selection of the presence or absence of symptoms. (F) The large category of body parts. (G) The small category of body parts. (H) The selection of symptoms. (I) Selection of presence of other symptoms. (J) Confirmation of the selected symptoms. (K) End. The app screen in Japanese is shown since the actual app is in Japanese. English translations of the screen elements are provided in parentheses within the figure.



Usability of the Symptom-Reporting App

In total, 5 adults participated in the usability study and interviews. Table 2 presents the demographic characteristics of the interviewees. Overall, 4 of the 5 interviewees regularly used smartphones, and 4 of the 5 interviewees had previously experienced side effects. The interviews elicited valuable opinions on the usability of the app, as summarized in Table 3, with an analysis aligned with the UTAUT model presented in Figure S2 in Multimedia Appendix 1.

XSL•FO
RenderX

Table 2. Basic information of interviewees.

No.	Age (years)	Sex	Occupation	Types of medicines taken, n	Experience of side effects	Smartphone or tablet	
						Possession	Frequency
1	79	Female	Unemployed	9	Yes	No	Do not use
2	24	Male	Student	3	Yes	Yes	Frequently
3	23	Male	Office worker	None	No	Yes	Frequently
4	50	Male	Sole proprietorship	3	Yes	Yes	Frequently
5	54	Female	Medical office worker	1	Yes	Yes	Frequently

XSL•FO
RenderX

Table 3. Opinions on app usability based on the Unified Theory of Acceptance and Use of Technology (UTAUT) model.

UTAUT category	Details
Performance expectancy	
Ease of consultation	• Easier to answer than to be asked • Easier to tell something than in person • Gives a chance to talk • Unclear what the app is used for
Contribution for medication	• May contribute to the treatment of others
Feedback from health care professionals	• Good to have an opinion from a pharmacist • Good to have continuous monitoring • Good to have topics related to daily life • Good to have a comfortable relationship with the pharmacist
Other functions	• Good to provide information about medication • Good to enter one's medication status • Good to be able to discuss other issues besides symptoms • Good to be linked to other systems • Good to be used outside the pharmacy • Concerns about adding too many functions
Effort expectancy	
Ease of use	• Usable without problems • Easy to use selective forms • Easy to be guided by selective forms • Negative feelings though none of the options apply to me • Difficult to enter free-text fields • Difficult to check, change, or add information • Difficult to use for older people who are not used to using smartphones or tablets
Expression of question items	• Usable without problems • Very detailed items • Many specialized words and phrases • Difficult to understand the words of the site classification • Difficult to understand the way the questions are asked
Social influence	
Influence from health care professionals	• Easy to answer if a health care provider asks about symptoms • App that can be used by pharmacists in medication instruction
Information collection and management	• Resistance to data being stored
Facilitating conditions	
Appearance	• Easy-to-read layout • Easy-to-read colors • Easy-to-read font size • Difficult reading text for older people
Device	• Difficult to use keyboard • No problems with the size of device • Easy to use the device with horizontal screen display • Probably being too small for older people
Behavioral intention	
Use of app	• Want to use the app • Should be used also outside of pharmacies • Probably not being able to be used for older people alone • Sometimes difficult to tell something even with app

In terms of "performance expectancy," participants expressed a preference for using the app to consult with pharmacists because "it is easier to answer than when being asked by a person," "it is easier to communicate than doing it face-to-face,"

XSL•FO
RenderX

and “it gives me a chance to talk to someone.” Some participants suggested that “it would be beneficial to receive expert opinions as feedback from a pharmacist.” In terms of “effort expectancy” (expectations about the ease of use and understanding of new technology), usability and expressions were both deemed “usable without problems,” while issues such as “too many technical terms” and “I feel bad when there are no applicable choices” were identified. Regarding “social influence” (the extent to which the user’s technology decisions are influenced by others), opinions included “it is easy to answer when a medical professional asks about symptoms” and “pharmacists can use it in medication guidance,” indicating that the influence of medical professionals contributes to the behavioral intention to use the technology. Regarding the “facilitating conditions,” it was suggested that there should be no problems with the design and terminal. Based on the above, all participants were positive about their behavioral intention to use the app, and some suggested that the app should be used outside pharmacies. Concerns were raised regarding the use of the app by older adults.

Several factors that did not align with the UTAUT model but significantly affected how ADRs reported to health care providers were identified. These factors included “I recognize pharmacist as a person to consult, but do not want to consult,” “I will not consult if not sure of side effect,” “I will not consult if the side effect is minor or not necessary to discuss by myself,” and “I will not consult if I think it is a side effect.” In addition, factors specifically affecting the reporting of ADRs to health care providers comprised “I recognize that it is an ADR but do not want to discuss it with a pharmacist,” “I do not discuss it unless I am sure it is an adverse drug reaction,” “I do not discuss it unless I am sure it is an adverse drug reaction,” and “I feel uncomfortable reporting an adverse drug reaction.”

Discussion

Principal Findings

In this study, we developed a hierarchical list of symptoms based on the MedDRA Patient-Friendly Term List and created an app that enables patients to easily select symptoms. The potential use of the app in pharmacies was well received, indicating that it could significantly aid patients in communicating their symptoms more effectively.

Our newly created list comprised 211 symptoms, and to improve user-friendliness, we implemented a “questionnaire flow” within the app, based on the hierarchical structure. This approach allows users to systematically select symptoms by navigating through major site categories, site subcategories, and detailed symptoms. Participants using the app in 5 hypothetical patient scenarios praised its ease of use, confirming the feasibility of using the hierarchical list of symptoms and the app for effective symptom selection. Previous research revealed that approximately 20% of reports on side effects used patient-friendly terms, with the majority opting for free-text entries [5]. By limiting displayed options to 10 or fewer, while providing a free-text input field for unlisted symptoms, our app may improve the ease of symptom identification. This design feature was further supported by the overall positive feedback from the participants. Future updates to the symptom list, informed by the analysis of real-world, free-text entries, could further streamline symptom reporting.

Participants in our study expressed their opinions about reporting usual self-aware symptoms such as “I do not discuss subjective symptoms unless I am sure they are side effects” and “I do not mention side effects if they seem minor or unnecessary to report,” which aligns with reasons reported in previous research for underreporting symptoms to health care providers [5]. These responses highlight common barriers to symptom reporting in health care settings, including patients’ uncertainty about whether symptoms are drug-related and their hesitation to mention symptoms they perceive as minor. To address these universal challenges, our app was designed with a structured interface and a hierarchical organization of symptoms, guiding patients through a systematic symptom-reporting process. The hierarchical structure of the app enables patients to progressively identify and select symptoms, potentially increasing the comprehensiveness of reported symptoms while reducing psychological barriers to reporting even minor or uncertain symptoms. Positive feedback from participants, such as “easier to answer than being asked by someone,” “easier to communicate than doing it face-to-face,” and “becomes a trigger to talk,” suggests that this structured approach could effectively address these reporting barriers. In addition, participants’ feedback indicating the app’s effortlessness (“can use without a problem”) and willingness to use (“I want to use it”) suggests its potential for successful implementation in pharmacy practice.

From the perspective of health care providers, pharmacists can review these structured patient reports before medicine consultation, enabling more focused and efficient discussions. These features, combined with the ability of the app to document patient-reported outcomes systematically, provide a foundation for continuous monitoring of therapy effectiveness and early detection of drug-related problems.

While our usability study received positive feedback regarding the potential of the app, we also identified several usability issues, particularly among older users. A participant aged 79 years expressed concerns, stating “I might not be able to use it on my own,” and other feedback pointed to specific issues such as “The size may be too small for elderly users” and “The text might be difficult to read.” Given that many customers of pharmacies are older adults, future usability assessments should focus on this population.

Implementing our app in pharmacy settings strategically aligns with the evolving health care landscape in Japan, particularly in aging societies where pharmacies are expanding beyond traditional medication dispensing [3,13]. Pharmacies offer several unique advantages as ideal implementation sites, including their high accessibility, regular patient visits, and prescription preparation waiting time, which offers a natural opportunity for symptom reporting without disrupting workflow. While the app could be valuable in various health care settings, including clinics and hospitals, pharmacies offer unique advantages. Pharmacists’ expertise in medication, combined with their frequent patient interactions, makes them ideal providers for continuous symptom monitoring and self-care

XSL•FO
RenderX

support. Their role has evolved from traditional medication dispensing and adverse effect monitoring to more comprehensive patient care support. Integrating the app into pharmacy practice could particularly support patient engagement in their own health care by enabling systematic symptom reporting in a familiar, low-pressure environment. This approach not only supports pharmacists' evolving role in comprehensive patient care but also promotes active patient participation in medication management.

Future research should systematically evaluate the effectiveness of the app in achieving its intended outcomes. Key areas of investigation include the impact of the app on the number and types of symptoms reported. Quantitative measurements will be valuable in assessing the impact on pharmaceutical interventions, including the identification of drug-related problems and subsequent care recommendations. Further research opportunities lie in evaluating improvements in the quality of patient-pharmacist communication and the effect of the app on the efficiency of pharmacy workflow. These investigations would benefit from including a larger, more diverse patient population in real-world pharmacy settings, with particular attention to older adults who represent a significant portion of pharmacy patients.

Although our sample size of 5 participants is consistent with recommendations for initial usability testing suggesting that 5 users can identify approximately 80% of major usability issues [14], this small sample size limits the generalizability of our findings. This limitation highlights the need for additional testing with a larger and more diverse user group to ensure the app meets the needs of all potential users in pharmacy settings.

Conclusions

We successfully developed a new hierarchical list of symptoms and an accompanying app. The app, designed to enable users to select symptoms from a structured list, was positively accepted and showed strong potential for improving patient behavior in symptom reporting. Future improvements to the design of the app design, particularly for older users, will further enhance its utility in both pharmacy settings and broader health care environments.

Acknowledgments

We would like to thank all the participants in this study.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to privacy or ethical restrictions but are available from the corresponding author on reasonable request.

Authors' Contributions

WS contributed to conceptualization, methodology, software, investigation, formal analysis, validation, and writing—review and editing. KH handled conceptualization, methodology, project administration, validation, visualization, writing—original draft, and writing—review and editing. HS performed conceptualization, methodology, project administration, supervision, validation, writing—original draft, and writing—review and editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary materials including hierarchical structure of symptom selection categories, UTAUT model application results, model patient profiles, and interview guide. UTAUT: Unified Theory of Acceptance and Use of Technology.
[DOCX File , 162 KB - humanfactors_v12i1e71439_app1.docx]

References

1. Basch E, Deal AM, Kris MG, Scher HI, Hudis CA, Sabbatini P, et al. Symptom monitoring with patient-reported outcomes during routine cancer treatment: a randomized controlled trial. J Clin Oncol 2016;34(6):557-565. [doi: 10.1200/jco.2015.63.0830]
2. Lizán L, Pérez-Carbonell L, Comellas M. Additional value of patient-reported symptom monitoring in cancer care: a systematic review of the literature. Cancers (Basel) 2021;13(18):4615 [FREE Full text] [doi: 10.3390/cancers13184615] [Medline: 34572842]
3. "Pharmacy Vision for Patients" - From "Front of the Door" to "Family Care" and then to "Community" - has been formulated. Ministry of Health, Labour and Welfare. URL: https://www.mhlw.go.jp/stf/houdou/0000102179.html [accessed 2024-08-28]
4. Schoenmakers TWA, Teichert M, Wensing M, de Smet PAGM. Evaluation of potentially drug-related patient-reported common symptoms assessed during clinical medication reviews: a cross-sectional observational study. Drug Saf 2017;40(5):419-430 [FREE Full text] [doi: 10.1007/s40264-017-0504-7] [Medline: 28205099]

XSL•FO
RenderX

5. Martin ML, Chung H, Rydén A. Willingness to report treatment-related symptoms of immunotherapy among patients with non-small cell lung cancer. Qual Life Res 2022;31(4):1147-1155 [FREE Full text] [doi: 10.1007/s11136-021-02966-3] [Medline: 34383226]
6. Lim R, Ellett LK, Roughead EE, Cheah PY, Masnoon N. Patient-reported questionnaires to identify adverse drug reactions: a systematic review. Int J Environ Res Public Health 2021;18(22):11877 [FREE Full text] [doi: 10.3390/ijerph182211877] [Medline: 34831635]
7. Nojo Y, Yamanaka S, Son N, Koda K, Kitagaki K, Narukawa M, et al. Improving data collection in pharmacies on subjective symptoms experienced by patients on high-risk medication using "Adverse Drug Reaction Signal Check Sheet" [Article in Japanese]. Journal of Regulatory Science (RSMP) 2022 Sep;12(3):259-270. [doi: 10.14982/rsmp.12.259]
8. Patient-friendly term list. MedDRA. URL: https://www.meddra.org/patient-friendly-term-list [accessed 2025-01-09]
9. de Vries ST, Harrison J, Revelle P, Ptaszynska-Neophytou A, Radecka A, Ragunathan G, et al. Use of a patient-friendly terms list in the adverse drug reaction report form: a database study. Drug Saf 2019;42(7):881-886 [FREE Full text] [doi: 10.1007/s40264-019-00800-x] [Medline: 30725337]
10. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Quarterly 2003;27(3):425-478. [doi: 10.2307/30036540]
11. Kim S, Lee KH, Hwang H, Yoo S. Analysis of the factors influencing healthcare professionals' adoption of mobile electronic medical record (EMR) using the Unified Theory of Acceptance and Use of Technology (UTAUT) in a tertiary hospital. BMC Med Inform Decis Mak 2016;16:12 [FREE Full text] [doi: 10.1186/s12911-016-0249-8] [Medline: 26831123]
12. Ono A. A consumer theory of new technology acceptance [Article in Japanese]. Mita Business Review 2008 Apr;51(1):1-18 [FREE Full text]
13. Kawahara A, Nakai K, Kurata N, Kamei M. New role of Japanese pharmacists focusing on medication therapy management in a team care approach. J Pharm Pract Res 2020;50(1):112-113. [doi: 10.1002/jppr.1616]
14. Nielsen J, Landauer TK. A mathematical model of the finding of usability problems. 1993 Presented at: CHI '93: Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems; April 24-29, 1993; Amsterdam, the Netherlands. [doi: 10.1145/169059.169166]

Abbreviations

ADR: adverse drug reaction
MedDRA: Medical Dictionary for Regulatory Activities
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by A Kushniruk, G Eysenbach; submitted 17.01.25; peer-reviewed by H Hosono, H Okada; comments to author 30.01.25; revised version received 07.02.25; accepted 10.02.25; published 06.03.25.

Please cite as:
Watanabe S, Kizaki H, Hori S
Development of a Patient-Centered Symptom-Reporting Application in Pharmacy Settings Using a Hierarchical Patient-Friendly Symptom List: Developmental and Usability Study
JMIR Hum Factors 2025;12:e71439
URL: https://humanfactors.jmir.org/2025/1/e71439
doi:10.2196/71439
PMID:40053749

©Seiya Watanabe, Hayato Kizaki, Satoko Hori. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 06.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Illustrating User Needs for eHealth With Experience Map: Interview Study With Chronic Kidney Disease Patients

Paula Valkonen[1], MA (Art and Des); Sini Hölsä[1], MSc (Tech); Johanna Viitanen[1], MSc (Tech), LicSc (Tech), DSc (Tech); Sini Leinonen[1], MSc (Tech); Nina Karisalmi[1], MSc (Tech); Virpi Rauta[2,3], MD, PhD

[1]Department of Computer Science, Aalto University, Espoo, Finland

[2]Department of Nephrology, University of Helsinki, Helsinki University Central Hospital, Helsinki, Finland

[3]Helsinki University Hospital, Abdominal Center, Nephrology, Helsinki, Finland

Corresponding Author:
Paula Valkonen, MA (Art and Des)
Department of Computer Science
Aalto University
P.O. Box 15400
Espoo, 00076
Finland
Phone: 358 947001
Email: paula.valkonen@aalto.fi

Abstract

Background: Chronic kidney disease (CKD) is a common condition worldwide and home dialysis (HD) provides economic, quality of life, and clinical advantages compared to other dialysis modalities. Human-centered design aims to support the development of eHealth solutions with high usability and user experience. However, research on the eHealth needs of patients using HD is scarce.

Objective: This study aimed to support the design of eHealth for patients with CKD, particularly for patients using HD, by developing a kidney disease experience map that illustrates user needs, concerns, and barriers. The research questions were (1) what experiences do patients, particularly older adults, have in their everyday lives with CKD? (2) what user needs do patients with CKD have for HD eHealth? (3) how can these needs be illustrated using the experience map technique? The study focused on patients aged >60 years, as they are at a higher risk of chronic conditions. The study was conducted as part of the eHealth in HD project, coordinated by Hospital District of Helsinki and Uusimaa, Finland.

Methods: In total, 18 patients in different care modalities participated in retrospective interviews conducted between October 2020 and April 2021. The interviews included a preliminary task with patient journey illustrations and questions about their experiences and everyday lives with CKD. The data analysis was conducted using a thematic analysis approach and the process included several phases.

Results: On the basis of the thematic analysis, 5 categories were identified: healthy habits, concerns about and barriers to eHealth use, digital communication, patients' emotions, and everyday life with CKD. These were illustrated in the first version of the kidney disease experience map. The patients had different healthy habits regarding social life, sports, and other activities. They had challenges with poorly functioning eHealth software and experienced other factors, such as a lack of interest and lack of skills for eHealth use. Technical devices do not always meet the emotional or physical needs of their users. This caused feelings of frustration, worry, and fear in patients, yet also fostered situational awareness and hope.

Conclusions: The experience map is a promising method for illustrating user needs and communicating the patient's voice for eHealth development. eHealth offers possibilities to support patient's everyday life with chronic disease. The patient's situation and capacity to use eHealth solutions vary with their everyday challenges, opportunities, and their current stage of treatment. The kidney disease experience map will be used and further developed in the ongoing research project "Better Health at Home—Optimized Human-Centered Care of Predialysis and Home Dialysis Patients" (2022 to 2026).

(JMIR Hum Factors 2025;12:e48221) doi:10.2196/48221

KEYWORDS

user need; chronic illness; kidney disease; older adult; eHealth; experience map; human-centered design; home dialysis

XSL•FO
RenderX

Introduction

Background

Chronic kidney disease (CKD) is a common condition worldwide. It has been estimated that 11% to 13% of the population in high-income countries have CKD. The number of patients with advanced kidney disease is growing 5% to 7% per year [1,2]. The prevalence of CKD is highest among older adults, ranging from 38% to 44% in patients aged >65 years [3,4]. In total 2.05 million people were treated with dialysis worldwide in 2010 [2]. In 2017, 1.2 million people died globally because of CKD [5].

Patients with end-stage CKD need kidney transplants or dialysis to survive. Dialysis can be performed in a dialysis unit in a hospital (ie, in-center dialysis), in a satellite dialysis unit, or at home using peritoneal dialysis or hemodialysis. The dialysis modality may vary depending on the patient's current health and life situation. Both in-center and satellite dialysis can be laborious for patients [6] and impose a heavy financial burden on medical care [7]. Therefore, home dialysis (HD) provides better quality of life and clinical advantages and empowers patients by providing them with more flexibility in their everyday lives [6,8].

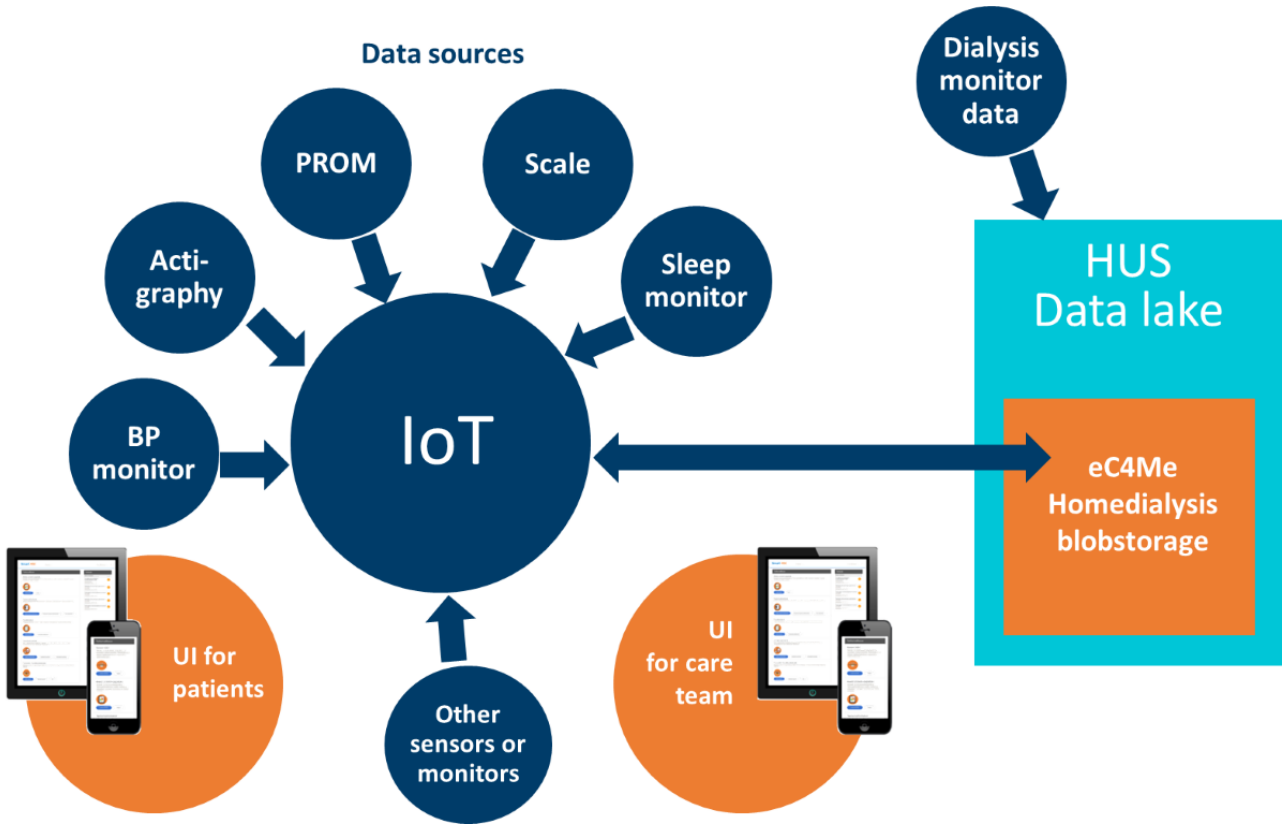
Even though bringing dialysis treatment to a patient's home might be burdensome and complicated for both the patient and the health care unit [8,9], there is a common understanding that HD prevalence needs to increase [1]. HD creates opportunities to improve the patient's safety and quality of life, as well as to support self-management of health [6,10]. HD has also proven to be feasible for older adults, even though challenges such as fear of needles or doubts about handling them, or other physical limitations have been recognized [8,11,12].

Research advises a holistic eHealth design that integrates technologies, end users, and use contexts [13-15]. Using human-centered design (HCD) approach [16], our study aimed to support the design of eHealth for patients with CKD, especially those on HD care, by investigating user needs, concerns, and digital service barriers, and illustrating those in the format of a kidney disease experience map. Due to the high prevalence of CKD among older adults [8], both in Finland and worldwide [17], the focus of the study was set on the perspective of older adults. The research questions were as follows:

1. What experiences do patients, particularly older adults, have in their everyday lives with CKD?
2. What user needs do patients with CKD have for HD eHealth?
3. How can these needs be illustrated using the experience map technique?

The study was part of the eHealth in HD project ("device research") [18] coordinated by the Hospital District of Helsinki and Uusimaa in Finland from 2020 to 2022. The project aimed to create a novel eHealth solution for patients with CKD undergoing HD (named "Home dialysis eHealth solution") and research devices that support patients both before and throughout dialysis therapy by gathering monitoring data, managing HD supply orders, and facilitating communication between the patient and the health care team (Figure 1).

Figure 1. eHealth solution and research devices for home dialysis. BP: blood pressure; HUS: Helsinki University Hospital; IoT: Internet of Things; PROM: patient-reported outcome measure; UI: user interface.



XSL•FO
RenderX

HCD of eHealth Solutions

HCD is an approach to interactive system development that aims to develop solutions with high usability and user experience [16]. The principles of HCD highlight the importance of involving end users in an iterative design process [16,19] and understanding their needs at the beginning of the development process [16,19].

When exploring user needs, researchers collect information, for example, about what features digital services should contain and how they should function from the perspectives of end users. The user needs exploration is important as it helps service designers to understand the role and meaning of the solutions for end users, in our case, patients using HD. eHealth is used to refer to digital health services, and health information offered through the internet or other technology solutions improving health care [20,21]. By designing in a human-centric way and involving end users in the development, eHealth can offer more value and new ways to provide health care and well-being services for patients and all citizens [13,21]. For patients with chronic illnesses, eHealth solutions can offer new communication possibilities between the patient and the health care team, as well as support empowerment and improve quality of life [21,22].

Patient Journey Maps and Experience Maps as Illustrations of User Needs

A patient journey map and an experience map are tools for 2 different purposes. In this study, patient journey maps were used as a part of patient interviews [23], and an experience map was used to communicate the results of the study holistically.

Patient journey maps as tools for visualizing patient journeys can be used to support the design of eHealth in several ways. The maps can help researchers and designers to identify how health care processes can be improved, for example, identify gaps and improve the processes by integrating eHealth solutions as part of those processes [24]. The patient journey maps sum up patients' experiences and care activities in a single, chronological, timeline-type visualization [24]. As the maps can help to explain what patients go through with their disease [23,24], they can help designers promote empathy, which is important in HCD [25-27]. Furthermore, they can help the sharing of knowledge between stakeholders as well as activate creative thinking [28].

Patient journey maps can be applied in various specialties regardless of the medical condition. For example, patient journey maps have been used to illustrate home isolation experiences of people with mild COVID-19 [29], hypertensive disorders in pregnancy [30], cancer [24], cervical dystonia [31], and kidney diseases [32]. In literature, the focus of patient journey maps varies from communicating patients' emotions and identifying process gaps to the importance of shared decision-making [33].

Experience maps are tools to combine situations, functions, emotions, and contacts in the same visualization [34]. Experience maps are valuable in the development of various eHealth, mobile health, and other apps, helping to communicate, for example, how cancer affects a patient's life and capturing patients' voices [35]. However, similar to patient journey maps, there is no standardized way to visualize and use them in an eHealth development context [33]. To our knowledge, no experience maps have been produced to represent the everyday life of patients with CKD using HD. Even though working with experience maps helps to identify users' needs [34], based on our literature review, few patient journey maps or experience maps focus on illustrating and communicating the needs as a basis for eHealth design and requirements specification.

Older Adults as eHealth Users

The European population is aging [36]. In Finland, >2 million citizens (ie, 36% of the population) are ≥55 years old [37]. Finnish older adults have good self-confidence in using digital services. In the 55 to 64 age group, >80% of Finns felt that their digital skills were at least on par with those of other Finns, and in the 65 to 74 age group, >60% of Finns felt the same [38]. At the same time, the number of patients in Finland who undergo dialysis, is expected to increase by >36% before 2040 [39]. Therefore, Finnish older adults seem to have both a need for and self-confidence in using CKD eHealth.

Older adults, due to an increase in diseases and complex health issues, could benefit substantially from eHealth [40]. Still, they experience several barriers to using eHealth, such as privacy concerns, lack of motivation to use digital solutions [41,42], and challenges to finding, accessing and understanding health-related information [40,43]. In addition, they have concerns about eHealth reducing the time with the physicians during appointments [42]. Although they can benefit from eHealth, they still face several challenges that could be solved with careful design.

The health situation also affects the use of the eHealth and user needs. Many eHealth tools are intended for the treatment of a specific disease. One or more chronic health conditions can lead to several regular self-management tasks [40], which can potentially affect a patient's everyday life and the use context of eHealth. Therefore, eHealth should be adapted to each user's special situation [44].

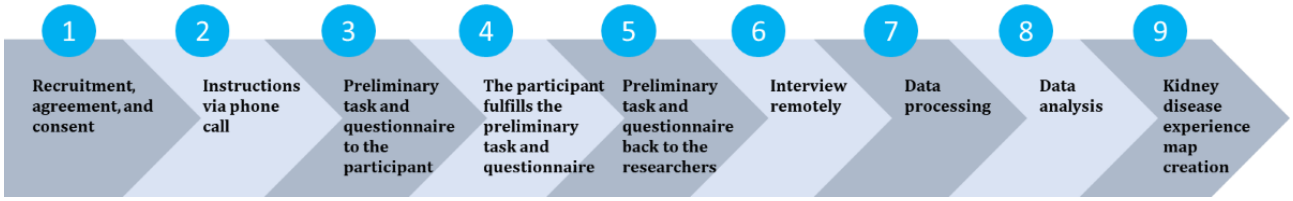
Methods

Overview

Combining different context-specific strategies is important when developing services in a complex environment, such as health care [15]. Our study on user needs of patients with CKD for eHealth used a qualitative research approach and semistructured interviews [23,45] as the primary method for data gathering. The interviews were supported with preliminary tasks, which included visualizing patient's journey in the form of a timeline drawing [23]. The procedure of the study included several phases starting from the recruitment and ending with the creation of the kidney disease experience map illustration (Figure 2).

XSL•FO
RenderX

Figure 2. Procedure of the study.



Ethical Considerations

The empirical study was conducted as part of the eHealth in HD project, which received permission from the ethical committee of the Hospital District of Helsinki and Uusimaa (HUS/1649/2020, Jarkko Ihalainen). All participants gave their voluntary, informed, and written consent. The patients' capabilities to participate in the study were ensured. While analyzing the data, the information of the patients participating in the study was pseudonymized and coded with the identification numbers. Only the researchers assigned to the study had access to the data. No compensation was paid to the participants for their participation.

Participants

In total, 18 patients aged ≥60 years participated in the study: 5 patients in predialysis phase, 4 patients in satellite dialysis, 5 patients in home peritoneal dialysis care, and 4 patients in home hemodialysis care (Table 1). Most of the participants were retired. The sample excluded in-center patients with dialysis, but included patients with dialysis in satellite units. The participants were recruited from the group of patients who were participating in the larger research project of HD eHealth solution development.

Table 1. Demographics of study participants (N=18).

Demographics	Frequency, n (%)
Dialysis	
Predialysis	5 (28)
Satellite dialysis	4 (22)
Home peritoneal dialysis	5 (28)
Home hemodialysis	4 (22)
Gender	
Woman	4 (22)
Man	14 (78)
Education	
Basic education	4 (22)
Upper secondary level	4 (22)
Bachelor's degree	5 (28)
Master's degree	3 (17)
Other	2 (11)
Technology skills (self-assessed)	
Good	6 (33)
Basic	11 (61)
Weak	1 (6)

The health care team recruited the patients by distributing materials about the study and asking about their interest in participation. Participation was voluntary and they did not get any compensation. The study participants were already familiar with the novel HD eHealth solution under development and had tested the first version of the solution. The research nurse contacted the potential participants first and informed them about the study. If the participants wanted to participate in the study, they signed the consent form, and the research nurse provided their contact information to the researchers. Then, the research material package, prepared by the researchers, consisting of a cover letter, a preliminary interview task, background information forms, and the responsible researcher's contact details were sent to the participants. After sending the package to the participants, the researcher called them to provide more detailed instructions, answer any questions, and schedule a time for the remote interview.

XSL•FO
RenderX

Preliminary Tasks and Retrospective Interviews

Interviews are useful in the phases of any development process related to eHealth [15]. In our study, retrospective interviews included questions about the patient's journey, treatments experiences, cooperation with the health care team, technology experiences, and visions for the future. The themes broadly covered patients' everyday experiences with the illness, practical questions, and the comprehensive timeline with CKD.

The idea of the visual timeline drawing as the preliminary task for the interview was to (1) help the interviewee to process and structure their multistage patient path even before the interview, and (2) help with communication between the interviewee and researcher during the remote interview [23]. The preliminary task was sent to the interviewees 2 weeks before each interview. In this task, patients were asked to identify and illustrate significant milestones, events, and experiences with their illness to a timeline. The patients returned the task to the researchers before the remote interview.

The interviews were conducted between October 2020 and April 2021 when the second wave of COVID-19 was underway in Finland. For safety reasons, the interviews were arranged remotely via Microsoft Teams. Two researchers were present in the interview session with the participating patient: one researcher being the interviewer and the other note-taker. Interviews were audio-recorded using recording functionality of Microsoft Teams. Data gathering was conducted in collaboration with 5 researchers (SL, NK, JV, SH, and PV): 2 doctoral researchers, 2 students, and a professor from the human-computer interaction field.

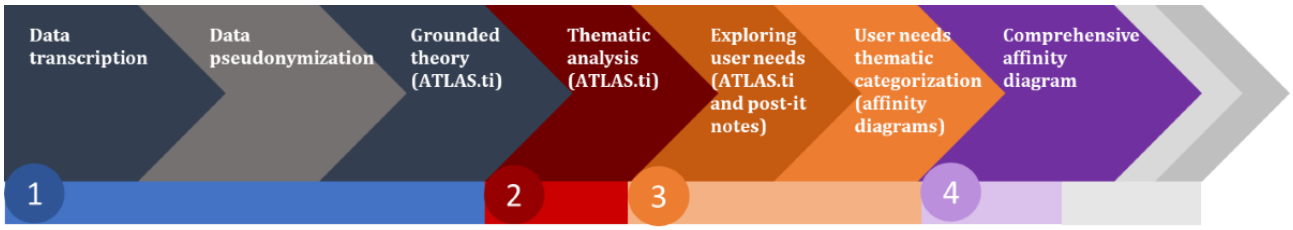
Data and Analysis

Overview

The data included recordings and notes from the interviews, as well as illustrations of patient journeys. The interview notes were finalized according to audio recordings and the transcriptions were pseudonymized. In total, the data consisted of 155 pages (ie, 58,429 words) of written notes and 18 paper-based patient journey visualizations (ie, visual timeline drawings). A total of 5 researchers (SL, NK, JV, SH, and PV) participated in the data analysis.

The data were analyzed in 4 phases (Figure 3): (1) grounded theory analysis, (2) thematic analysis, (3) user needs exploration and thematic categorization, and (4) comprehensive affinity diagram creation. The data included many different perspectives to understand the different nuances and ensure reliability, and the analysis was done in collaboration with several researchers in many phases.

Figure 3. Four phases of data analysis.



Phase 1: Grounded Theory Analysis

In the first phase [46], the pseudonymized interview transcripts were analyzed following the main structure of grounded theory [47,48] using ATLAS.ti software version 9.1.5.0 (ATLAS.ti Scientific Software Development GmbH, "ATLAS.ti") [46]. During the analysis, the interview data was coded using researcher-denoted concepts and open coding influenced by the grounded theory [46,47]. The first round of interview data analysis following grounded theory was conducted by SL.

Phase 2: Thematic Analysis

In the second phase, the principles of the thematic analysis method [48,49] were followed to analyze the interview data from 18 participants from an existentialist perspective, including references to death, well-being, and atmosphere in life (eg, emotions) [50-52]. In addition, the analysis used the holistic framework [13-15] of exploring the context of use, technology, and people.

On the basis of the analysis, 3 thematic categories were identified: healthy habits, concerns and barriers for eHealth use, and everyday life with CKD. Observations based on interview transcripts were written on post-it notes and thematically categorized following the phases of the affinity diagram method [48,49]. The data gathered using preliminary task of patient journeys were mapped in line with the interview transcripts. In addition to the initial 3 themes, 2 categories were formed based on thematic analysis: digital communication and patients' emotions. Concerns and use barriers for eHealth were investigated from the data by two researchers (PV and SH). The digital communication affinity diagram was done by three researchers (PV, SH, and JV).

The thematic analysis categories include several aspects as presented in Textbox 1.

XSL•FO
RenderX

Textbox 1. Categories of thematic analysis.

- Healthy habits: health-promoting activities and hobbies that were mentioned by interviewees, including the whole scale of activities from sports to eating habits.
- Concerns and use barriers for eHealth: observations related to everyday life challenges with the disease, use challenges and barriers of digital services, negative emotions, such as worries or dissatisfactions, and patient path challenges.
- Digital communication: observations concerning health care-related digital communication habits, channels, and experiences.
- Patients' emotions: observations of emotions and experiences interviewees mentioned in the interviews and patient journey visualizations. The emotions included the whole scale of emotions, from fear and confusion to happiness.
- Everyday life with chronic kidney disease: observations related to interviewees' thoughts, stories, and experiences of their everyday lives, including references to death. Death made the participants reflect from different perspectives: their own (forthcoming) death was compared to other deaths, the actions to avoid death were listed, and the inevitable nature of death was considered.

Finally, the frequency of observations per theme was calculated (Multimedia Appendix 1). Most of the observations were related to concerns and barriers to technology use. Emotions and healthy habits were also common topics in all the interviews.

Phase 3: The User Needs Exploration and Thematic Categorization

The third phase focused on analyzing user needs. All the gathered data were explored from the viewpoint of user needs from 2 perspectives: exploring digital communication needs and exploring overall user needs for interactive solution development. In the analysis, 2 researchers (PV and SH) explored the data labeled "digital communication," "user needs," and "to eHealth solution concept" in ATLAS.ti and formulated the observations as user needs. The second round resulted in 165 needs.

After this, the digital communication affinity diagram was created in collaboration with 3 researchers (SH, JV, and PV). The user needs were written on post-it-notes by PV, and analyzed, and regrouped to the affinity diagram in collaboration with SH, JV, and PV. In the affinity diagram, the identified needs (n=165, 100%) were thematically grouped under 5 categories: overall interaction and communication (n=77, 47% of all identified needs), patients' digital activities (n=28, 17%), instruments (n=27, 16%), inventory and ordering dialysis supplies (n=19, 12%), and digital communication with the health care team (n=14, 8%). While creating the affinity diagram of digital communication needs, doubles were removed.

When exploring overall user needs, the data transcriptions of 18 interviews were analyzed using ATLAS.ti. During this analysis, the data labeled with the following codes—"user needs," "healthy habits and hobbies," "suggestion," "communication," "digital services," "tasks," "care team," "communication," "social relationships," "equipment," "challenges," "positive experiences," and "device research"—were explored and formulated as user needs. In total, 287 user needs were identified and written on post-it notes as preparation for the comprehensive affinity diagram work.

Phase 4: Comprehensive Affinity Diagram

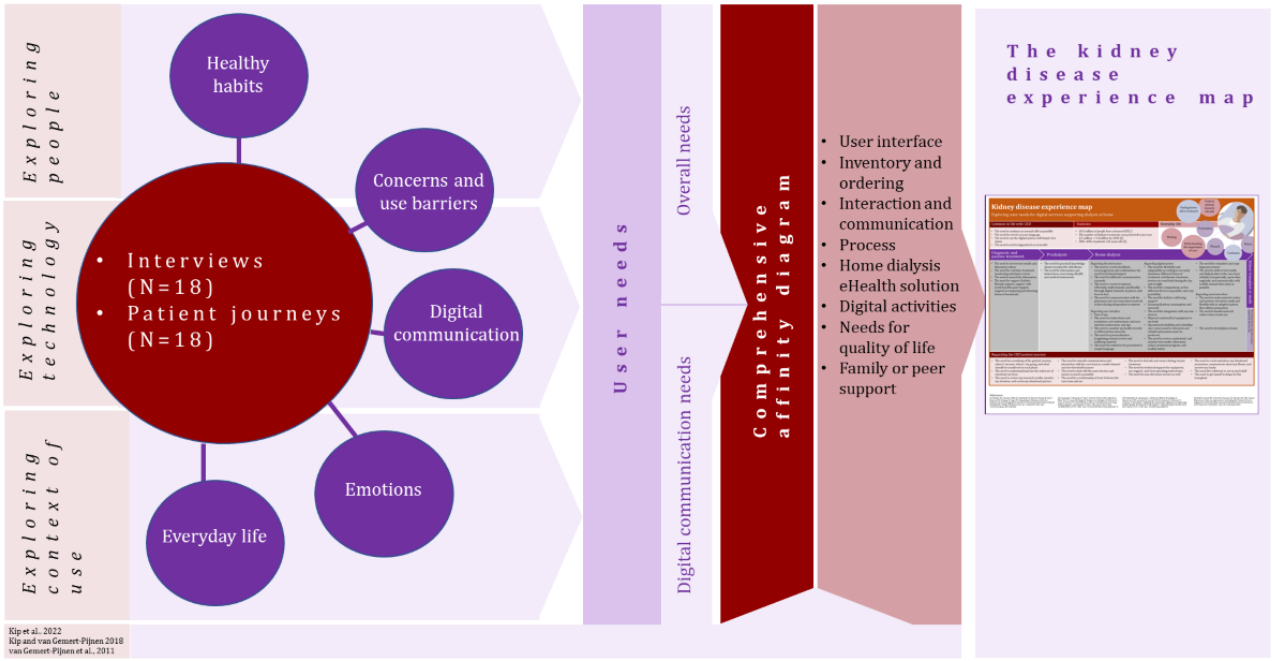
In the fourth phase, the affinity diagrams created in the previous analysis phase were merged to form a comprehensive affinity diagram. Three researchers recategorized the post-it notes to create a comprehensive affinity diagram of user needs. The main themes of the diagram were "eHealth user interfaces," "Inventory and ordering dialysis supplies," "Overall interaction and communication," "Process," "HD eHealth solution," "Family/peer-support," "Patients' digital activities," and "Needs for improving the quality of life" (Multimedia Appendix 2). These 5 were used as the leading themes to guide the design of the first version of the kidney disease experience map. Data analysis results were discussed between PV, SH, VR, and JV.

Creation of the First Version of the Kidney Disease Experience Map

An experience map is a tool that strings the perspectives of the context of use, people, and technology, together in a holistic manner [13-15]. The maps can be used to support the communication between stakeholders and to capture a patient's voice in eHealth design [35]. The first version of the kidney disease patient experience map (Figure 4) was made by combining remarks from existing example models particularly the cancer experience map [35]. In the first paper-based template of our map, the preliminary phases from the preliminary tasks (ie, patient journey maps) were illustrated: "before diagnosis," "diagnosis and passive treatment," "treatment method selection," "beginning of treatment/training," and "home dialysis." Post-it notes from the comprehensive affinity diagram were added under those phases. The first draft of the patient experience map was created in collaboration with 3 researchers.

XSL•FO
RenderX

Figure 4. Overview of the data analysis process leading to the creation of the first version of the kidney disease experience map.



Next, the structure of the experience map was refined based on the number of post-it notes in each phase. The duplicate post-it notes were removed and the content of the map was reorganized (PV, SH, and JV). To validate the contents of the draft map further, 2 researchers (PV and SH) reexamined the post-it notes after which 4 other researchers (JV, NK, SL, and VR) gave feedback on the next version of the map. After all, in total of 6 researchers (PV, SH, JV, NK, SL, and VR) participated at least somehow in the experience map visualization process. The published version of the first version of the kidney disease experience map was finalized by PV. An overview of the multiphase analysis process which led to the creation of this first version of the kidney disease experience map is illustrated in Figure 4.

Results

Overview

As the outcome of the study, the first version of the kidney disease experience map, which illustrates the user needs, concerns, and barriers of patients with CKD for eHealth, was created. This section explores the patients (ie, people), technology, and context of use perspectives [13-15]. First, patients' healthy habits are investigated and their concerns and barriers to technology are addressed. Then, the technology is examined, concentrating on digital communication. Finally, the context of use from the perspective of patients' everyday lives is explored. Due to sensitive personal issues, research results and citations are presented without identification information to ensure complete anonymity of the participants.

Exploring the People: Healthy Habits

The patients had different needs for healthy habits regarding social life, sports, and other activities (Multimedia Appendix 3). Patients mentioned that social life and taking care of close relatives, such as grandchildren, were important to them. Incidental exercises, including gardening or shoveling snow, were popular. Being in nature, picking mushrooms or berries, fishing, and hiking, along with other everyday life activities, such as club and society activities or studying, brought well-being to the participants. Animals, for example, dogs, cats, and summer chickens, also supported the participants' well-being. Handcrafting, art, and culture were also seen as important, but COVID-19 restricted those activities. The importance of many remote social contacts was mentioned in the data, and patients had to learn new eating habits, such as losing weight or cutting down on alcohol consumption.

Exploring the People: Concerns and Barriers for eHealth Use

Patients were challenged with poorly functioning available eHealth solutions and also experienced other negative factors, such as lack of interest or lack of skills (Multimedia Appendix 4). Technical devices did not always meet the emotional or physical needs of their users. This caused emotions like frustration, worry, and fear from patients. The patients' responses covered their available eHealth solutions, not focusing especially on the HD eHealth solution.

The solution or software used by patients for their CKD care was found to function poorly in many ways. For example, the devices as a part of the HD eHealth solution did not work as expected, or they worked differently than they expected:

> *I have no control over the information. If I measure my blood pressure, I press the BT button and then it should be delivered, but I don't know if it did. The same goes for the scale, *beep*, and maybe it synchronized.* [P18]

Alongside poorly functioning devices, in some cases, the HD eHealth app also did not work well. The software content (eg, one's results or visualizations of one's health situations) was challenging to understand or find. The HD eHealth app was challenging to use with small mobile phone screens and its user interface terminology caused problems for patients:

XSL•FO
RenderX

> *There have been a few problems with the phone. There will be updates and then it will freeze somehow. I don't understand much; the basic things are in Finnish, but then they are English words so... This was at least the third time that the nurse had visited us.* [P17]

> *Others have also worked well. Even the mobile phone. You can monitor your health from it, even though it has those [health status] curves that you don't understand much about.* [P13]

The participants doubted the functionality of the HD eHealth solution, both devices and the app. With the dialysis monitor, they were worried that something would go wrong:

> *I haven't used (ie, the devices and instruments) much, I've just tried. I'm a bit lazy. I doubt whether the instrument is still in good condition. There was an error message about a month ago. There was a notification that there was still something wrong. I thought I'd let it be and fix the instrument before I started.* [P24]

> *When the first treatments started, it was so exciting. But after a month had passed, it had become routine and was no longer exciting. Then, when I moved home from hospital, I started to get excited again.* [P22]

Exploring the Technology: Digital Communication

Overview

Patients commented on digital devices and the app from many different aspects in terms of the technical functionality and the emotion they cause. The contents of the solution's user interface (Figure 1), as well as other user interfaces, were also commented on. Patients linked and communicated with the health care team through digital channels and used technology independently as a part of their HD treatment during the research.

Many Groups Communicate With the Patient

Communication within the health care team is important, but also nurses, physicians, and patients share information (Multimedia Appendix 2). To get the supplies and instruments needed for HD care, the patient also needs to be in regular contact with the pharmacy or the health care team. Relatives must not be forgotten either, as patients may want to share information with them:

> *To be able to get in touch with the on-call nephrologist or the other way around, perhaps you could organize a little more frequent meeting. I think that many patients could have questions for the doctor that the nurses cannot or are not allowed to answer. Nephrologist on-call from the computer.* [P24]

Examples of Current Communication Channels

Depending on the patients' devices and patients' capabilities to use them, HD patients use different communication channels for following their health records and CKD-specific operations. They order HD supplies once per month. Some patients have created their digital tools (eg, Excel sheets), which calculate the number of needed supplies automatically. After that, the order will be sent via email to the nurse. Still, Excel sheets or other tools for ordering dialysis supplies are not an option for everybody. Therefore, an easy-to-use solution for ordering supplies and managing inventory is needed for HD patients:

> *Orders must be sent once a month. I have an Excel sheet that automatically keeps the balances. Then, there is a page where you can automatically see what needs to be ordered. I have made Excel sheets myself. It's quite a job when there are more than 30 different supplies. Syringes... you must know their numbers. Sometimes, I also check some information. If one item is true, then so are all the others. It's automatic. I then make a .pdf list from the Excel sheet, one A4, and then it goes [by e-mail to the nurse].* [P30]

In Finland, patients follow their health records via national or local patient portals (eg, "My Kanta Pages" [53] or "MyChart" [54]). Patients with chronic illness have many laboratory tests taken. They appreciate that all their health records are found in one place, but each record is shown in detail separately, too. Therefore, easy-to-read visualizations of test results and current health situations are needed:

> *MyChart is as hopeless as My Kanta, both of which break down the lab results separately, even if they want to see them as a whole. Applying for your results is difficult.* [P37]

> *I look at the results on 'My Kanta,' all diagnoses come there, as well as laboratory test results. From there, I can see them all. The information is readable and remains stored. I think 'My Kanta' is really good. There you can also see laboratory visits from a certain period if you want to follow a limited period.* [P21]

Positive Attitude Toward Digital Communication

All patients participating in the study seemed to have a positive attitude toward digital communication tools. For example, paper-based dialysis supply orderings may be forgotten and instead, digital tools can step in and help patients with their everyday lives. Patients felt that digital solutions enabled information-seeking and offered more flexibility in meetings with the health care team and other contacts:

> *Their version is that they hand out an A4 paper. You should fill it in; it's difficult manually. The risk is that you forget.* [P30]

> *We have internet and e-mail. My wife uses them more; I just read. The wife handles banking and other things remotely. I kind of stopped using the computer when I don't need to. You could probably handle things digitally. I haven't come across anything like that. I have had to deal with a lot of paper. I have not been approached digitally. Yes, it would be even easier. You could reach the staff at any time.* [P24]

XSL•FO
RenderX

Exploring the Context of Use: Everyday Life and Emotions

Overview

The main themes related to everyday life were thinking of death, the importance of care, changes in well-being after treatment started, and emotions (Multimedia Appendix 5).

The Presence of Death

With kidney disease, death is present in life. It is known that without dialysis or receiving a transplant, one dies. The logic of the disease is unequivocal: when dialysis no longer helps, the person knows they are going to die:

> *I think that the disease is deadly and must be treated.* [P21]
>
> *Everyone knows that life ends eventually, but I'm not afraid of it in any way.* [P1]

The General Well-Being of the Participants Varied

The patients understood the essentiality of treatment and often felt better after the treatment started. They wished to be able to continue life as normally as possible. One patient (P9) described experiencing leg cramps at night, which had caused them to sleep poorly. After the dialysis treatment started, the cramps stopped. They managed better than before and could even go for walks:

> *I hope I don't end up on dialysis. Of course, it is ahead, but hopefully as far as possible. It is a wish.* [P20]
>
> *It's just that now I feel like I'm getting older. I recently turned 70, so all kinds of unnecessary aches and pains will increase.* [P25]

Everyday Life Is Characterized by Waiting

The health situation of a patient can be improved either with dialysis or a transplant. Having to wait for transplant can have major setbacks; one patient (P24) had to be removed from the transplantation waiting list due to amputation and sepsis, which is why they now focus on keeping themselves in the best possible physical condition through physical activities. Patients were waiting to travel.

A Wide Range of Emotions in Everyday Life

A wide range of emotions appeared in the interview data about the patient's everyday life. Despite the challenges of the disease, most of the patients described their lives with the disease in somewhat positive terms. Everyday life with the disease seemed to go somewhat smoothly.

In addition to positive feelings, patients also expressed negative emotions, such as confusion, fear, irritation, anger, and frustration. Feelings of confusion were caused by changes in health status and events in the patient's journey, uncertainty of the future steps in the journey, or technology that worked unexpectedly. After receiving the diagnosis, one patient was confused because they did not fully understand what kind of disease the diagnosis was about. Several patients brought up fear in the interviews. Things related to the treatments were especially scary. Most of the patients also described irritation or anger at least sometimes during the patient journey. In addition, half of the patients reported frustration. Health challenges, the difficulty or even failure of a treatment, and the commitment to dialysis treatment were irritating. The most frustrating thing was when HD treatment did not go as well as planned. Another source of frustration was digital systems, equipment, and devices:

> *And a little fear. I am quite hopeful that I will get there on the transplantation waiting list and that it will be successful at some point. I hope that I can make it there (ie, from a health perspective), and if this continues, then there is nothing to worry about.* [P11]
>
> *Actigraphy, a part of HD eHealth solution, is inconvenient, but there is nothing wrong with the other devices, everything is fine. I have a strong belief that information will go forward. It's enough for me to see on the scale and blood pressure monitor that my weight and blood pressure are okay.* [P18]
>
> *This home dialysis takes up a lot of space. Do you see these boxes? A huge package comes once a month. The dialysis monitor also takes up quite a lot of space. And then the cabinets, where you put all the treatment consumables. And there is a bottle like this: hand sanitizer comes once a month. Quite a lot of supplies come from the hospital once a month. This room is completely reserved for dialysis.* [P30]

HD taking up too much space or other resources was both irritating and frustrating. HD requires much space at home, so some patients had to renovate their homes. However, almost half of the patients felt that their living conditions did not require changes due to HD. The patients felt it was important that the dialysis treatment could also be done at the summer cottage.

Positive emotions included joy, optimism, hope, and satisfaction. A third of the patients mentioned joy in the interviews and most patients expressed hope and optimism. They were satisfied with the professional staff and the stability of the current situation. The supportive and successful dialysis made them happy. Keeping their health situation stable was their biggest wish. Patients also mentioned their willingness to continue their hobbies.

The patients were satisfied with both their health condition and the information related to it. Half of them were satisfied with the treatment. The patients were pleased with the eHealth solution in use during the study and the HD treatment was experienced as proceeding smoothly. The nursing staff was considered friendly and knowledgeable:

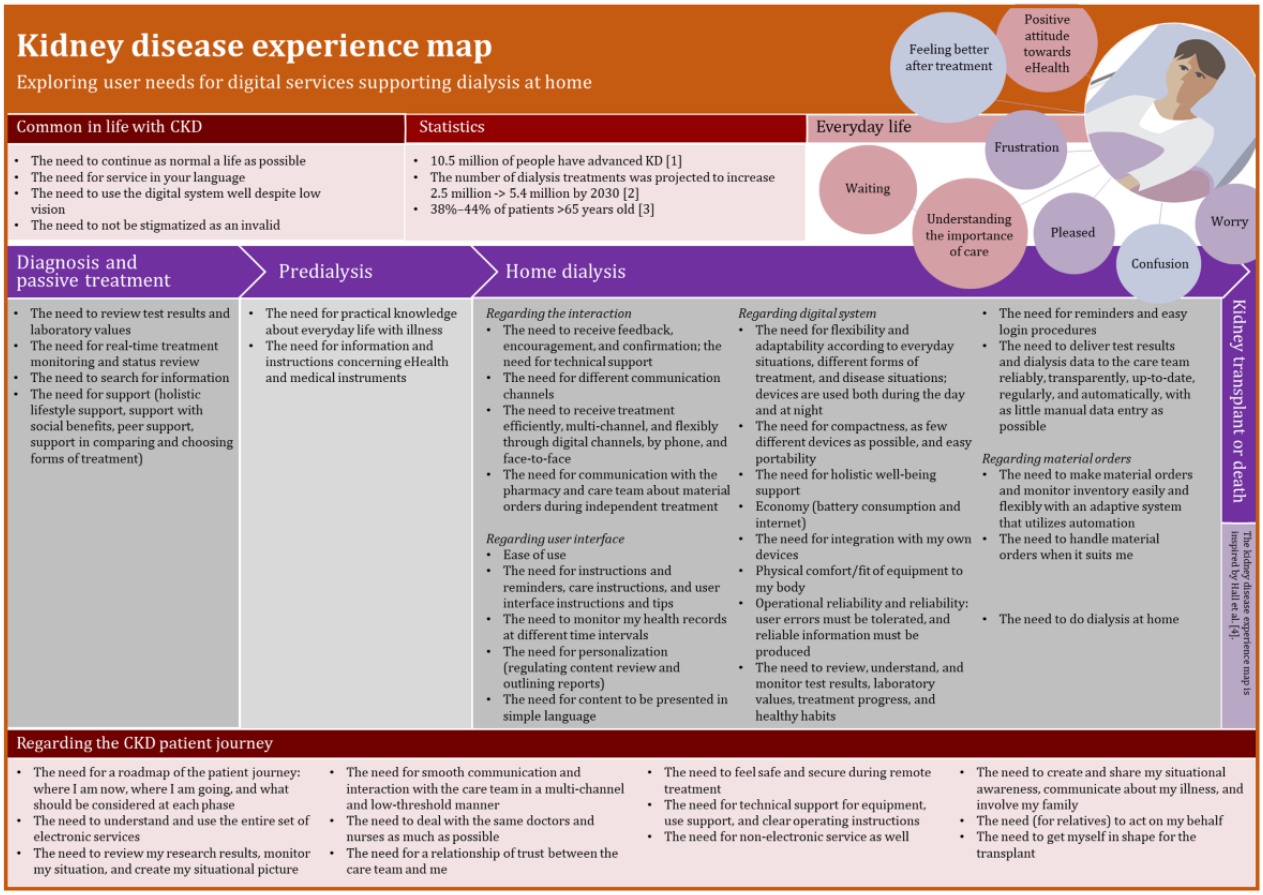
> *I have been very satisfied. They, the care team, have guided me, and when I have challenged them, they have accepted challenges and promised to sort things out.* [P37]
>
> *Home dialysis has gone surprisingly well—no major problems.* [P34]
>
> *At least now, this present moment looks bright.* [P13]

XSL•FO
RenderX

The First Version of the Kidney Disease Experience Map

The first version of the kidney disease experience map (Figure 5 [1-3,35]) was developed from the perspective of user needs of eHealth. It includes their needs for everyday life with CKD (ie, "Common needs"), needs relating to 4 phases of the journey of patients with CKD, and needs in all journey phases of patient with CKD from diagnosis to kidney transplant or death. The patient journey phases, which are inspired by the experience timeline of cancer experience map [35], are "Diagnosis and passive treatment," "Predialysis," "HD," and "Preparing for kidney transplant or death." In the top corner of our map kidney disease statistics and patients' everyday life experiences with CKD are illustrated.

Figure 5. The first version of the kidney disease experience map from a user needs perspective. A higher resolution image is available in Multimedia Appendix 6. CKD: chronic kidney disease. KD: kidney disease.



References:
[1] Mendu ML, Divino-Filho JC, Vanholder R, Mitra S, Davies SJ, Jha V, Damster KC, Gallego D, Seger M. Expanding Utilization of Home Dialysis: An Action Agenda From the First International Home Dialysis Roundtable. Kidney Medicine 2021 Jul 1;3(4):635–643. doi: 10.1016/j.xkme.2021.04.004
[2] Liyanage T, Ninomiya T, Jha V, Neal B, Patrice HM, Okpechi I, Zhao M, Lv J, Garg AX, Knight J, Rodgers A, Gallagher M, Kotwal S, Cass A, Perkovic V. Worldwide access to treatment for end-stage kidney disease: a systematic review. The Lancet 2015 May 16;385(9981):1975–1982. doi: 10.1016/S0140-6736(14)61601-9
[3] Vanholder R, Annemans L, Bello AK, Bikbov B, Gallego D, Gansevoort RT, Lameire N, Luyckx VA, Noruisiene E, Oostrom T, Wanner C, Wieringa F. Fighting the unbearable lightness of neglecting kidney health: the decade of the kidney*. Clinical Kidney Journal 2021 Jul 1;14(7):1719–1730. doi: 10.1093/ckj/sfab070
[4] Hall LK, Kunz BF, Davis EV, Dawson RJ, Powers RS. The Cancer Experience Map: An Approach to Including the Patient Voice in Supportive Care Solutions. Journal of Medical Internet Research 2015 May 28;17(5):e3652. doi: 10.2196/jmir.3652

The first version of the kidney disease experience map (Figure 5) includes aspects from all three perspectives [13-15]:

1. Context of use: the context of use perspective can be found especially in the patient journey phases and the "everyday life with CKD" part of the map.
2. People: information about people is available in the "statistics and everyday life with CKD" parts of the map.
3. Technology: technology aspects are mainly covered in the form of user needs.

Discussion

Principal Findings

On the basis of our study, eHealth as a part of the daily lives of the patients with CKD was perceived as being useful, but the patients also had challenges with devices and apps. Acknowledging that CKD is a long-lasting journey, it is essential to consider long-term user experiences. Finally, several user needs in different phases of the journey of patients with CKD were identified and illustrated with the first version of the kidney disease experience map.

Answers to Research Questions

Overview

This study had three research questions: (1) what experiences do patients, particularly older adults, have in their everyday lives with CKD; (2) what user needs do patients with CKD have for HD eHealth; and (3) how can these needs be illustrated using the experience map technique?

XSL•FO
RenderX

What Experiences Do Patients, Particularly Older Adults, Have in Their Lives With CKD?

Three key experiences related to daily life of patients with CKD were identified as depicted in Textbox 2

Textbox 2. Key experiences related to daily life of patients with chronic kidney disease (CKD).

- Patients' healthy habits: patients with CKD must adapt to the changes in their everyday lives resulting from dialysis care. For example, patients must learn new routines for their treatment at home. Patients felt that CKD limited their lives, but they dealt with it in different ways: some were frustrated and others continued to live as normally as possible. However, they understood the necessity of the treatment and continued the treatment as usual. Similar findings were reported from a study involving cancer patients [35], who expressed concerns with unknown, life-threatening, and life-changing experiences.
- Patients' concerns and barriers to use for eHealth: patients had challenges with user interface and technical devices, including dialysis monitors and research devices of HD eHealth solution (Figure 1), that did not always meet the users' needs.
- Digital communication: digital communication opportunities via the eHealth solution play an important role for patients. Digital solutions made information-seeking possible and offered more flexibility when meeting the health care team and other contacts. Information gaps did not accumulate as easily as for cancer [24].

What User Needs Do Patients With CKD Have for HD eHealth?

In our study, both general needs and needs related to different patient journey phases were identified and illustrated in the first version of the kidney disease experience map (Figure 5). We found that CKD was different from day to day, and thus the use context of eHealth varied. Poor usability burdens the everyday life of a person with disease even more. Similar observations were reported in the case of digital independent living supporting systems of older adults [34]: designers were important to understand the treatment stage, the transition from health to illness, and statuses between them. Furthermore, the needs and experiences of patients with cancer were found to vary during the care process [35].

The HD eHealth solution is targeted to all patients with CKD, as well as individuals with disabilities, such as low vision or even amputated limbs. The solution is to be used during the entire patient journey, which may last for years. Considering this, the HD eHealth solution is observed to play a remarkable role in the patient's life and highlights the need for research on long-term user experience.

How Can the User Needs Be Illustrated Using the Experience Map Technique?

Typically, user needs are illustrated with use cases or user needs tables, and they can be processed further to requirements format [55]. Findings from user research can also be presented as user profile illustrations, such as personas [15,34], stakeholder maps [15], experience maps [34], and visualizations of stakeholders [34].

In our study, we used the experience map method to illustrate user needs. As with the cancer experience map [35], we identified the phases of care and relevant themes for grouping the findings. We considered the experience map technique applicable for illustrating not only the phases but also the eHealth related needs of the patients. We found a group of general needs, which are common for all phases of care, but also needs that arise from specific phases of care. Understanding the care process and how user needs evolve is important for the designers and also for the health care team, who are responsible for introducing the services to patients and training them. Experience maps help to ensure the care continuum in service development [34].

In this study, the created experience map represents a common higher-level experience map, which can illustrate a specific persona, or the timeline could be narrower (ie, 1 month or 1 day). In the first version of the kidney disease experience map, the HD eHealth solution covers different wearable solutions and the software (ie, including the wearable solutions and software connected to the dialysis monitors); however, different physical dimensions (ie, dialysis monitors, packing materials of supplies ordered) are not covered.

Experience mapping can be seen as a promising but laborious method for illustrating user needs and communicating a patient's voice holistically to eHealth development, especially in a case where the experience map is based on extensive qualitative data. Following the example of cancer experience map [35], we captured patient's voices with their quotes. Similarly, we also focused on reporting patient quotes in the paper instead of the experience map. Therefore, the first version of the kidney disease experience map can be used both separately and with the paper. Our interpretation is that similar to the integrated patient journey map for hypertensive disorders in pregnancy [30] and the cancer experience map [35], the kidney disease experience map can help developers build effective eHealth solutions and promote empathy and shared understanding between the developing team members. It can also help empower patients and help with shared decision-making [33]. Like the patient journey map [24], the first version of the kidney disease experience map has the potential to help simplify and optimize processes for better health outcomes. In addition to that, similar to the cervical dystonia patient journey map [31], it can help to communicate the different phases of the care to the patients and help them to understand their needs at each phase.

Relevance of the Research

In the study, the experience map describing the patient's experience was formed based on the qualitative data from retrospective interviews and preliminary tasks. This first version of the kidney disease experience map can be used as a basis for other experience maps related to kidney diseases in the future. In addition, it can also bring perspective to the experience maps of other chronic diseases.

XSL•FO
RenderX

The presented results can be used to support the further development of HD eHealth. The results have already been used in HD eHealth solution development project in Finland (eg, to inform the redesign of the solution user interfaces).

Experience maps, such as the map created in the study, can help researchers, designers, and developers to empathize with the patients with chronic illnesses, and to understand their needs and use contexts for future eHealth innovations. We see our kidney disease experience map and other similar maps having the potential to help design the solutions from a more holistic perspective compared to currently prevailing practices [13,21], as well as to create a base for the new ecosystems and new ways to offer health care in the future [21]. We continue the work in the ongoing research project "Better Health at Home—Optimized Human-Centered Care of Predialysis and Home Dialysis Patients" (2022 to 2026).

Evaluation of the Study

Our study had some limitations. First, exploring the context of use was not allowed at patients' homes due to COVID-19 restrictions. Therefore, we needed to collect data remotely. This was partly problematic because the patient's home is the main context of using HD eHealth. In contrast, due to the patient journey preliminary task and participants' open-minded attitude during the interviews, we were able to gather rich data about user needs and the context of eHealth use. In our study, the possible impact of the COVID-19 period was considered.

Our sample size is quite small, however, typical for a qualitative interview study. The representativeness of our sample can be questioned as it is divided from a larger patient group based on age and treatment methods. To find as comprehensive a sample as possible, while ensuring the safety of the participants with their ongoing treatment, the health care team helped researchers in recruiting participants for the research study.

Due to the personal nature of health information and privacy, we could not release the participants' background information, such as specific ages or treatments in this paper. In addition, the indirect identifiers have been removed from the paper. The relative satisfaction of the patients in their lives despite CKD could have been influenced by various factors, besides the HD eHealth solution. First, the target group in this study included only active patients with relatively good physical conditions for research ethics reasons. Second, many of them were waiting for a kidney transplant, which brought them hope with the disease. Overall, there was a lot of variation in emotions, but deep gloom appeared in the data only rarely.

The first version of the kidney disease experience map was created based on a reasonable amount of qualitative data of the everyday lives on patients aged ≥60 years in Finland. Before participating in this study, the patients already knew that the study covered new eHealth solution testing. Therefore, perspectives from patients whose condition is very weak and those whose attitude is not open to new technology are missing in this study. In contrast, research shows that HD is optimal for those who live in areas with long distances to travel to an in-center dialysis unit [6,7]. In part, this can also be seen in the first version of the kidney disease experience map: the study did not include patient experiences during the kidney transplant phase—meaning that the user needs related to this essential phase of the patient journey are missing from our map. In future, the introduced version of the kidney disease experience map can be updated by including motivational issues for HD.

The study was conducted as an integrated part of the patient's normal CKD treatment, so the research setup was organized in line with the patients' typical daily conditions and treatments. The study included participants who started using the HD eHealth solution shortly before the interviews but also those who had already been using the solution for several months. Thus, the study covered both novice users and more experienced users. Patients' reported technology skills varied from weak to good.

Further Research

This study provided initial experiences of creating an experience map to illustrate patients' needs for eHealth. In addition, the patients' experiences from the HD eHealth solution development can offer important perspectives to other development projects regarding patients with CKD. Further research is needed to gather more experience about the applicability of the method. We aim to continue the research with patients with CKD and HD eHealth designs. We will use the illustrated first version of the kidney disease experience map and the listed user needs in further analysis, which will aim to describe design principles and guidelines for HCD of eHealth for patients using HD and other patients with chronic illnesses.

The outcome of the first version of the kidney disease experience map offers insights to the different phases of the patient journey for patients with kidney disease at the HD care. Due to the diverse experiences of patients in different phases of the care process, different types of kidney diseases could also be studied and visualized with experience maps. In addition, the next versions of the kidney disease experience maps could focus on specific days and months. The next versions of the kidney disease experience maps could also be created from the health care team's perspective.

The content of the first version of the kidney disease experience map was collected with qualitative research methods, which made the consideration of the many different nuances possible. Due to the qualitative nature of the data collection for the first version of the kidney disease experience map, we did not prioritize the available content, for example, based on the amount of the post-it notes. However, without careful prioritization, there is always a risk that the map is too challenging to use as a part of the practical software development work. Therefore, we will involve the patients, as well as the health care team, designers, and software developers, in prioritizing the content of the map in the future.

The perspectives of patients using in-center dialysis are missing from that data, which was the base of the first version of the kidney disease experience map. However, we believe the creation of experience maps for the eHealth needs of patients using in-center dialysis is important for the future as the dialysis modalities might change depending on the patient's health and life situation.

XSL•FO
RenderX

Our study pointed out the need for further research on the existentialism perspectives of the patients with chronic illnesses using eHealth. Deeper understanding is required on user needs and the role of eHealth as a part of patients' everyday lives, even when life is nearing its end [50-52].

Conclusions

eHealth offers possibilities to support the key needs of patients with chronic diseases, such as CKD; however, digital services need to adapt to a wide variety of situations and user needs because kidney diseases are unique to each patient. Patient experiences vary both between and within the stages of treatment. In addition, the patient's situation, needs, and capacity to use eHealth vary with everyday challenges and opportunities. Furthermore, with older adult patients, the problems that come with age increase the spectrum of barriers to use eHealth and associated challenges.

eHealth should enable proactive communication between the patient and the health care team. Furthermore, awareness of the status should be formed for both parties in real time, as much as possible. In this way, the patients feel safe, are more motivated in their treatment, and take responsibility for themselves.

The experience map is a promising method for illustrating user needs for eHealth, communicating a patient's voice and describing the longitudinal patient journey to support eHealth development. Our first version of the kidney disease experience map combines people, context of use, and technology perspectives holistically and illustrates the eHealth needs of patients with CKD for the future.

Acknowledgments

The project eHealth in Home Dialysis was part of the CleverHealth Network ecosystem and the project (eCare for Me Homedialysis) was funded by participating companies, the Hospital District of Helsinki and Uusimaa, and Business Finland. We thank our project partners Tiina Puurtinen and Pirgit Silvast-Äikäs from Helsinki University Hospital and Malin Ekholm from Helsinki University, and the patients who participated in this study.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to the ethical reasons as stated in the research permit application.

Authors' Contributions

The study was conceptualized by PV, SL, SH, NK, VR, and JV. Data were collected by five researchers (PV, SH, SL, NK, and JV). PV, SL, and SH were involved in the methodology, investigation, project administration, data curation, visualizations, and formal analysis of the study. PV finalized the experience map. JV contributed to the resources, validation, supervision, project administration, and funding acquisition of the study. VR contributed to funding acquisition, project administration, resource collection, supervision, and results' validation. All authors were involved in writing, reviewing, and editing the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Theme analysis: themes from the second phase of the analysis. CKD: chronic kidney disease.
[PDF File (Adobe PDF File), 82 KB - humanfactors_v12i1e48221_app1.pdf]

Multimedia Appendix 2
Main themes of the comprehensive affinity diagram (n=287). Some post-it notes included content for many themes.
[PDF File (Adobe PDF File), 106 KB - humanfactors_v12i1e48221_app2.pdf]

Multimedia Appendix 3
Healthy habits and hobbies, and their frequency (n=105, 18 participants).
[PDF File (Adobe PDF File), 83 KB - humanfactors_v12i1e48221_app3.pdf]

Multimedia Appendix 4
Barriers to eHealth use (n=288, 18 participants).
[PDF File (Adobe PDF File), 104 KB - humanfactors_v12i1e48221_app4.pdf]

Multimedia Appendix 5
Aspects of everyday life with chronic kidney disease, N=18 (aspects N>3 are excluded from the table)

XSL•FO
RenderX

[PDF File (Adobe PDF File), 134 KB - humanfactors_v12i1e48221_app5.pdf]

Multimedia Appendix 6
The first version of the kidney disease experience map from a user needs perspective. KD: kidney disease. CKD: chronic kidney disease.
[PNG File , 479 KB - humanfactors_v12i1e48221_app6.png]

References

1. Mendu ML, Divino-Filho JC, Vanholder R, Mitra S, Davies SJ, Jha V, International Home Dialysis Roundtable Steering Committee. Expanding utilization of home dialysis: an action agenda from the first international home dialysis roundtable. Kidney Med 2021;3(4):635-643 [FREE Full text] [doi: 10.1016/j.xkme.2021.04.004] [Medline: 34401729]
2. Liyanage T, Ninomiya T, Jha V, Neal B, Patrice HM, Okpechi I, et al. Worldwide access to treatment for end-stage kidney disease: a systematic review. Lancet 2015 May 16;385(9981):1975-1982. [doi: 10.1016/S0140-6736(14)61601-9] [Medline: 25777665]
3. Vanholder R, Annemans L, Bello AK, Bikbov B, Gallego D, Gansevoort RT, et al. Fighting the unbearable lightness of neglecting kidney health: the decade of the kidney. Clin Kidney J 2021 Jul;14(7):1719-1730 [FREE Full text] [doi: 10.1093/ckj/sfab070] [Medline: 34221379]
4. Brück K, Stel VS, Gambaro G, Hallan S, Völzke H, Ärnlöv J, European CKD Burden Consortium. CKD prevalence varies across the European general population. J Am Soc Nephrol 2016 Jul;27(7):2135-2147 [FREE Full text] [doi: 10.1681/ASN.2015050542] [Medline: 26701975]
5. GBD Chronic Kidney Disease Collaboration. Global, regional, and national burden of chronic kidney disease, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2020 Feb 29;395(10225):709-733 [FREE Full text] [doi: 10.1016/S0140-6736(20)30045-3] [Medline: 32061315]
6. Rygh E, Arild E, Johnsen E, Rumpsfeld M. Choosing to live with home dialysis-patients' experiences and potential for telemedicine support: a qualitative study. BMC Nephrol 2012 Mar 19;13:13 [FREE Full text] [doi: 10.1186/1471-2369-13-13] [Medline: 22429705]
7. Malmström RK, Roine RP, Heikkilä A, Räsänen P, Sintonen H, Muroma-Karttunen R, et al. Cost analysis and health-related quality of life of home and self-care satellite haemodialysis. Nephrol Dial Transplant 2008 Jun;23(6):1990-1996. [doi: 10.1093/ndt/gfm910] [Medline: 18223263]
8. Wu HH, Dhaygude AP, Mitra S, Tennankore KK. Home dialysis in older adults: challenges and solutions. Clin Kidney J 2023 Mar;16(3):422-431 [FREE Full text] [doi: 10.1093/ckj/sfac220] [Medline: 36865019]
9. Wang C, Ku E. eHealth in kidney care. Nat Rev Nephrol 2020 Jul;16(7):368-370 [FREE Full text] [doi: 10.1038/s41581-020-0271-z] [Medline: 32238923]
10. Hölsä S, Valkonen P, Viitanen J, Rauta V. Kohti käyttäjälähtöisiä kotidialyysin sähköisiä terveyspalveluita: Potilaskokemus ja ammattilaisten tarpeet suunnittelun lähtökohtana. FinJeHeW 2022 May 09;14(2):55. [doi: 10.23996/fjhw.114929]
11. Cornelis T, Tennankore KK, Goffin E, Rauta V, Honkanen E, Őzyilmaz A, et al. An international feasibility study of home haemodialysis in older patients. Nephrol Dial Transplant 2014 Dec;29(12):2327-2333. [doi: 10.1093/ndt/gfu260] [Medline: 25085237]
12. Muroma-Karttunen R. Almost anyone can have home therapy but can we overcome the barriers? J Ren Care 2013 Jan;39 Suppl 1:22-27. [doi: 10.1111/j.1755-6686.2013.00331.x] [Medline: 23464910]
13. Kip H, van Gemert-Pijnen L. Holistic development of eHealth technology. In: Kip H, Beerlage-de Jong N, van Gemert-Pijnen L, Sanderman R, Kelders SM, editors. eHealth Research, Theory and Development: A Multi-Disciplinary Approach. London, UK: Routledge; 2018:131-166.
14. van Gemert-Pijnen JE, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011 Dec 05;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
15. Kip H, Keizer J, da Silva MC, Beerlage-de Jong N, Köhle N, Kelders SM. Methods for human-centered eHealth development: narrative scoping review. J Med Internet Res 2022 Jan 27;24(1):e31858 [FREE Full text] [doi: 10.2196/31858] [Medline: 35084359]
16. ISO 9241-210:2019 Ergonomics of human-system interaction - part 210: Human-centred design for interactive systems. International Organization for Standardization (ISO). URL: https://www.iso.org/standard/77520.html [accessed 2024-04-29]
17. Webster AC, Nagler EV, Morton RL, Masson P. Chronic kidney disease. Lancet 2017 Mar 25;389(10075):1238-1252. [doi: 10.1016/S0140-6736(16)32064-5] [Medline: 27887750]
18. Development of home dialysis. CleverHealth Network. URL: https://www.cleverhealth.fi/en/development-of-home-dialysis/ [accessed 2023-01-13]
19. Gould JD, Lewis C. Designing for usability: key principles and what designers think. Commun ACM 1985 Mar;28(3):300-311 [FREE Full text] [doi: 10.1145/3166.3170]

XSL•FO
RenderX

20. Eysenbach G. What is e-health? J Med Internet Res 2001;3(2):E20 [FREE Full text] [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
21. van Gemert-Pijnen LJ, Kip H, Kelders SM, Sanderman R. Introducing eHealth. In: Kip H, Jong NB, van Gemert-Pijnen L, Sanderman R, Kelders SM, editors. eHealth Research, Theory and Development: A Multi-Disciplinary Approach. New York, NY: Routledge; 2018:3-26.
22. Brands MR, Gouw SC, Beestrum M, Cronin RM, Fijnvandraat K, Badawy SM. Patient-centered digital health records and their effects on health outcomes: systematic review. J Med Internet Res 2022 Dec 22;24(12):e43086 [FREE Full text] [doi: 10.2196/43086] [Medline: 36548034]
23. Valkonen P, Karisalmi N, Kaipio J, Kujala S. Remote interviews and visual timelines with patients: lessons learned. Stud Health Technol Inform 2021 May 27;281:845-849. [doi: 10.3233/SHTI210298] [Medline: 34042793]
24. Kushniruk AW, Borycki EM, Parush A. A case study of patient journey mapping to identify gaps in healthcare: learning from experience with cancer diagnosis and treatment. Knowl Manag E-Learn 2020;12(4):405-418 [FREE Full text] [doi: 10.34105/j.kmel.2020.12.022]
25. Ilen E, Groth C, Ahola M, Niinimäki K. Empathy in a technology-driven design process: designing for users without a voice of their own. Nordes 2019 May 10;8:1-11 [FREE Full text] [doi: 10.21606/nordes.2019.021]
26. Ector GI, Westerweel PE, Hermens RP, Braspenning KA, Heeren BC, Vinck OM, et al. The development of a web-based, patient-centered intervention for patients with chronic myeloid leukemia (CMyLife): design thinking development approach. J Med Internet Res 2020 May 15;22(5):e15895 [FREE Full text] [doi: 10.2196/15895] [Medline: 32412424]
27. Leary M, Cacchione PZ, Demiris G, Carthon JM, Bauermeister JA. An integrative review of human-centered design and design thinking for the creation of health interventions. Nurs Forum 2022 Nov;57(6):1137-1152. [doi: 10.1111/nuf.12805] [Medline: 36255150]
28. McCarthy S, O'Raghallaigh P, Woodworth S, Lim YL, Kenny LC, Adam F. An integrated patient journey mapping tool for embedding quality in healthcare service reform. J Decis Syst 2016 Jun 16;25(sup1):354-368. [doi: 10.1080/12460125.2016.1187394]
29. He Q, Du F, Simonse LW. A Patient journey map to improve the home isolation experience of persons with mild COVID-19: design research for service touchpoints of artificial intelligence in eHealth. JMIR Med Inform 2021 Apr 12;9(4):e23238 [FREE Full text] [doi: 10.2196/23238] [Medline: 33444156]
30. McCarthy S, O'Raghallaigh P, Woodworth S, Lim YY, Kenny LC, Adam F. Embedding the pillars of quality in health information technology solutions using "integrated patient journey mapping" (IPJM): case study. JMIR Hum Factors 2020 Sep 17;7(3):e17416 [FREE Full text] [doi: 10.2196/17416] [Medline: 32940610]
31. Benson M, Albanese A, Bhatia KP, Cavillon P, Cuffe L, König K, et al. Development of a patient journey map for people living with cervical dystonia. Orphanet J Rare Dis 2022 Mar 21;17(1):130 [FREE Full text] [doi: 10.1186/s13023-022-02270-4] [Medline: 35313909]
32. Kelly J, Wilden C, Herman K, Martin G, Russell C, Brown S. Botling knowledge and sharing it - using patient journey mapping to build evidence and improve aboriginal renal patent care. Ren Soc Australas J 2016;12(2):48-55 [FREE Full text] [doi: 10.3316/informit.182564296613440]
33. Joseph AL, Kushniruk AW, Borycki EM. Patient journey mapping: current practices, challenges and future opportunities in healthcare. Knowl Manag E-Learn 2020;12(4):387-404 [FREE Full text] [doi: 10.34105/j.kmel.2020.12.021]
34. Schäpers B, Krewer C, Yasuoka M, Becker D, Heller S, Kaiser M, et al. Determining design requirements for active aging: personas, experience maps, and stakeholders. Gerontechnology 2017 Oct 31;16(3):139-150 [FREE Full text] [doi: 10.4017/gt.2017.16.3.003.00]
35. Hall LK, Kunz BF, Davis EV, Dawson RI, Powers RS. The cancer experience map: an approach to including the patient voice in supportive care solutions. J Med Internet Res 2015 May 28;17(5):e132 [FREE Full text] [doi: 10.2196/jmir.3652] [Medline: 26022846]
36. Ageing Europe — looking at the lives of older people in the EU — 2020 edition. Eurostat. 2020. URL: https://ec.europa.eu/eurostat/documents/3217494/11478057/KS-02-20-655-EN-N.pdf/9b09606c-d4e8-4c33-63d2-3b20d5c19c91?t=1604055531000 [accessed 2023-01-13]
37. 11rc population according to age (5-year) and sex, 1865-2022. Statistics Finland. URL: https://pxdata.stat.fi/PxWeb/pxweb/en/StatFin/StatFin__vaerak/statfin_vaerak_pxt_11rc.px/ [accessed 2023-12-20]
38. Digital Skills Report 2023. Digital and Population Data Services Agency. URL: https://dvv.fi/documents/16079645/0/Digitaitoraportti_2023_saavutettava.pdf/4d2f74b1-0baf-a06a-f28e-954bb8df491d/Digitaitoraportti_2023_saavutettava.pdf?t=1702556379736 [accessed 2023-12-20]
39. Finnish registry for kidney diseases 2017. Munuais- ja maksaliitto. URL: https://muma.fi/wp-content/uploads/2024/02/Suomen-munuaistautirekisteri-vuosiraportti-2017.pdf [accessed 2023-12-20]
40. Ware P, Bartlett SJ, Paré G, Symeonidis I, Tannenbaum C, Bartlett G, et al. Using eHealth technologies: interests, preferences, and concerns of older adults. Interact J Med Res 2017 Mar 23;6(1):e3 [FREE Full text] [doi: 10.2196/ijmr.4447] [Medline: 28336506]

XSL•FO
RenderX

41. Backonja U, Hall AK, Thielke S. Older adults' current and potential uses of information technologies in a changing world: a theoretical perspective. Int J Aging Hum Dev 2014 Dec;80(1):41-63 [FREE Full text] [doi: 10.1177/0091415015591109] [Medline: 26215298]
42. Latulipe C, Gatto A, Nguyen HT, Miller DP, Quandt SA, Bertoni AG, et al. Design considerations for patient portal adoption by low-income, older adults. Proc SIGCHI Conf Hum Factor Comput Syst 2015 Apr;2015:3859-3868 [FREE Full text] [doi: 10.1145/2702123.2702392] [Medline: 27077140]
43. Reiners F, Sturm J, Bouw LJ, Wouters EJ. Sociodemographic factors influencing the use of eHealth in people with chronic diseases. Int J Environ Res Public Health 2019 Feb 21;16(4):645 [FREE Full text] [doi: 10.3390/ijerph16040645] [Medline: 30795623]
44. Scheibe M, Reichelt J, Bellmann M, Kirch W. Acceptance factors of mobile apps for diabetes by patients aged 50 or older: a qualitative study. Med 2 0 2015 Mar 02;4(1):e1 [FREE Full text] [doi: 10.2196/med20.3912] [Medline: 25733033]
45. Blandford A, Furniss D, Makri S. Qualitative HCI Research: Going Behind the Scenes. Cham, Switzerland: Springer; 2016.
46. Leinonen S. Design principles of eHealth services for chronically ill patients. Master of Science Thesis, Tampere University. 2021. URL: https://trepo.tuni.fi/handle/10024/130761 [accessed 2023-01-16]
47. Corbin J, Strauss A. Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. Thousand Oaks, CA: Sage Publications; 2014.
48. Boyatzis RE. Transforming Qualitative Information: Thematic Analysis and Code Development. Thousand Oaks, CA: Sage Publications; 1998.
49. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
50. Kaptelinin V. Making the case for an existential perspective in HCI research on mortality and death. In: Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems. 2016 Presented at: CHI EA '16; May 7-12, 2016; San Jose, CA p. 352-364 URL: https://dl.acm.org/doi/10.1145/2851581.2892585#:~:text=The%20paper%20discusses%20some%20of,and%20evaluation%20%2D%2D%20is%20not [doi: 10.1145/2851581.2892585]
51. Kaptelinin V. Technology and the givens of existence: toward an existential inquiry framework in HCI research. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI '18; April 21-26, 2018; Montreal, QC p. 1-14 URL: https://dl.acm.org/doi/10.1145/3173574.3173844 [doi: 10.1145/3173574.3173844]
52. Albers R, Sadeghian S, Laschke M, Hassenzahl M. Dying, death, and the afterlife in human-computer interaction. a scoping review. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. 2023 Presented at: CHI '23; April 23-28, 2023; Hamburg, Germany p. 1-16 URL: https://dl.acm.org/doi/10.1145/3544548.3581199 [doi: 10.1145/3544548.3581199]
53. My Kanta pages - citizens. Kanta Services. URL: https://www.kanta.fi/en/my-kanta-pages [accessed 2023-01-16]
54. Maisa makes life easier. Client Portal Maisa.: Apotti URL: https://www.apotti.fi/en/client-portal-maisa/ [accessed 2023-01-13]
55. Kujala S, Kauppinen M, Rekola S. Bridging the gap between user needs and user requirements. In: Proceedings of the 2021 Panhellenic Conference with International Participation in Human-Computer Interaction. 2021 Presented at: PC-HCI '21; July 24-29, 2021; Virtual Event p. 45-50 URL: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=62823e8ccba3c33fa89ddb74b45c1d4c49ac878d

Abbreviations

CKD: chronic kidney disease
HCD: human-centered design
HD: home dialysis

Edited by A Kushniruk; submitted 16.04.23; peer-reviewed by N Koizumi, J Eriksen; comments to author 18.09.23; revised version received 21.12.23; accepted 15.01.25; published 18.03.25.

Please cite as:
Valkonen P, Hölsä S, Viitanen J, Leinonen S, Karisalmi N, Rauta V
Illustrating User Needs for eHealth With Experience Map: Interview Study With Chronic Kidney Disease Patients
JMIR Hum Factors 2025;12:e48221
URL: https://humanfactors.jmir.org/2025/1/e48221
doi:10.2196/48221
PMID:

XSL•FO
RenderX

©Paula Valkonen, Sini Hölsä, Johanna Viitanen, Sini Leinonen, Nina Karisalmi, Virpi Rauta. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 18.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Best Practice Guide for Reducing Barriers to Video Call–Based Telehealth: Modified Delphi Study Among Health Care Professionals

Lena Rettinger[1,2], BSc, MSc; Lea Aichinger[1], BSc, MSc; Veronika Ertelt-Bach[3], MAS; Andreas Huber[4], MSc; Susanne Maria Javorszky[5], BSc, MSc; Lukas Maul[1], BSc, MSc; Peter Putz[6], Dr rer nat; Sevan Sargis[7], BSc, MSc; Franz Werner[1], Dr. techn; Klaus Widhalm[8], MSc; Sebastian Kuhn[2], Prof Dr, MME

[1]Research Center Digital Health and Care, FH Campus Wien, University of Applied Sciences Vienna, Vienna, Austria
[2]Institute of Digital Medicine, Philipps-University & University Hospital of Giessen and Marburg, Marburg, Germany
[3]Occupational Therapy, FH Campus Wien, University of Applied Sciences Vienna, Vienna, Austria
[4]Orthoptics, FH Campus Wien, University of Applied Sciences Vienna, Vienna, Austria
[5]Logopedics – Phoniatrics – Audiology, FH Campus Wien, University of Applied Sciences Vienna, Vienna, Austria
[6]Research Center Health Sciences, FH Campus Wien, University of Applied Sciences Vienna, Vienna, Austria
[7]Midwifery, FH Campus Wien, University of Applied Sciences Vienna, Vienna, Austria
[8]Physiotherapy, FH Campus Wien, University of Applied Sciences Vienna, Vienna, Austria

Corresponding Author:
Lena Rettinger, BSc, MSc
Research Center Digital Health and Care
FH Campus Wien, University of Applied Sciences Vienna
Favoritenstrasse 226
Vienna, 1100
Austria
Phone: 43 6066977 ext 4382
Email: lena.rettinger@fh-campuswien.ac.at

Abstract

Background: Telehealth has grown, especially during the COVID-19 pandemic, improving access for those in remote or underserved areas. However, its implementation faces technological, practical, and interpersonal barriers.

Objective: The aim of this study was to identify and consolidate best practices for telehealth delivery, specifically for video call sessions, by synthesizing the insights of health care professionals across various disciplines.

Methods: We first identified 15 common telehealth barriers from a preceding scoping review. Subsequently, a modified Delphi method was used, involving 9 health care professionals (physiotherapists, speech and language therapists, dietitians, and midwife) with telehealth experience in qualitative interviews and 2 iterative rounds of web-based surveys to form consensus.

Results: This study addressed 15 telehealth barriers and identified 105 best practices. Among these, 20 are technology-related and 85 concern health care practices. Emphasis was placed on setting up telehealth environments, ensuring safety, building relationships and trust, using nonmanual methods, and enhancing observation and assessment skills. Best practice recommendations for dealing with patients or caregiver skepticism or lack of telehealth-specific knowledge were developed. Further, approaches for unstable networks and privacy and IT security issues were identified. Areas with fewer best practices were the lack of technology skills or technology access, unreliability of hardware and software, increased workload, and a lack of caregiver support.

Conclusions: This guide of best practices serves as an actionable resource for health care providers to navigate the complexities of telehealth. Despite a small participant sample and the potential for profession-specific biases, the findings provide a foundation for improving telehealth services and inform future research for its application and education.

(JMIR Hum Factors 2025;12:e64079) doi:10.2196/64079

XSL•FO
RenderX

KEYWORDS

telehealth; best practices; video call; Delphi study; health communication; barriers; health care professionals; qualitative interviews; web-based survey; physiotherapists; speech therapists; language therapists; dietitians; midwife

Introduction

Video call–based telehealth refers to the use of information and communication technologies to deliver health care services remotely and synchronously, bridging the gap between patients and health care providers [1]. Over the past few years, telehealth has seen an increase in adoption due to its potential to improve access to health care services, particularly for individuals living in rural or remote areas [2]. The COVID-19 pandemic has further accelerated this trend, as the need for remote health care solutions to reduce the risk of infection transmission has become more urgent [3]. As a result, telehealth has emerged as an essential part of modern health care delivery, offering new opportunities and challenges for both patients and health care providers.

Telehealth has the potential to offer numerous advantages and benefits for patients, health care providers, and the overall health care system. Benefits summarized in previous reviews include, for example, increased access and convenience, cost savings, improved family involvement, and better management [4]. Increased access refers particularly to individuals living in remote or underserved areas [2], as well as to those with mobility limitations or chronic conditions that require frequent monitoring [5]. Furthermore, telehealth enables patients to receive care from the comfort of their own homes, reducing the need for travel and waiting times associated with in-person appointments and providing a more convenient way of service provision [6]. By reducing travel costs, decreasing hospital readmissions, and minimizing the need for in-person consultations, telehealth might also have the potential to reduce health care costs for both patients and providers [7,8].

Despite its numerous advantages, telehealth also faces several challenges and barriers that can hinder its successful implementation. Recent studies and reviews have highlighted various obstacles to delivering care via real-time video consultations. These include technical issues such as unstable internet connectivity [4,9,10], lack of technology skills among both patients and providers [4,9,11], and concerns over privacy and data security [9,12]. Moreover, health care professionals have reported challenges in establishing rapport and trust with patients via video call [9,11], and there exists skepticism about the effectiveness of remote care compared to in-person sessions [9]. Addressing these barriers is especially critical when considering the rapid shift to telehealth prompted by the COVID-19 pandemic, which has underscored the potential benefits of video consultations but also exposed the need for clearer guidelines to ensure quality of care [9,13]. Despite growing evidence on these challenges, the literature lacks a comprehensive set of actionable best practices to mitigate barriers associated with video consultations. This study therefore aims to fill that gap by drawing on multiple health care disciplines and using a modified Delphi process to achieve expert consensus on best practices to maximize the potential benefits of video call–based telehealth for patients, health care providers, and the overall health care system [14].

Methods

Overview

In this study, we used a modified Delphi method [15] to identify best practices to overcome common barriers in video call–based telehealth. The reporting follows the proposed Delphi reporting guideline by Spranger et al [16].

Research Paradigm

The epistemological positioning of this study is based on a pragmatic paradigm that recognizes the value of diverse viewpoints and the importance of practical solutions to real-world problems. Pragmatism has been identified as relevant and useful for qualitative research on organizational processes [17], as well as in patient-oriented research [18]. This approach also aligns with the goals of the modified Delphi method, which aims to synthesize expert knowledge into actionable best practices. It acknowledges the dynamic and iterative nature of knowledge creation and embraces subjective experiences to address the complex challenges of telehealth implementation.

Study Design

Our modified Delphi method was characterized by 2 main adaptations to the traditional Delphi process. First, we used preliminary expert interviews to gather an initial pool of best practices. These interviews guided the development of a list of 131 practice statements and 15 barriers, which were then presented to the expert panel for rating. Second, our survey structure diverged from a classic Delphi by using (1) a 4-point Likert scale and (2) an additional response option “This practice is not relevant for my profession.” This allowed professionals from various health care disciplines to opt out of rating statements they found inapplicable. We also set specific thresholds for acceptance (≥80% agreement after round 1; ≥75% agreement after round 2) to ensure each item reached a robust consensus without necessitating a third round. These modifications provided a more streamlined approach and helped account for the diverse professional backgrounds of the participants and time and resource limitations.

Recruitment

We recruited participants via a combination of email invitations, professional social media groups, and a snowball sampling strategy. Known telehealth practitioners were first contacted directly by the research team and then invited to share the invitation within their professional networks. To ensure broad representation across allied health professions, we distributed the study invitation to professional associations and specialized digital communities. To be eligible, participants had to (1) have at least 2 years of professional experience in a non-physician health care field (eg, nursing, dietetics, occupational therapy, speech-language therapy, physiotherapy, orthoptics, or

XSL•FO
RenderX

midwifery), (2) have delivered a minimum of 10 video call–based telehealth sessions in the preceding quarter, and (3) have excellent German language proficiency. All participants were given detailed information about the study objectives, procedures, and time commitments.

Participants

A total of 9 health care professionals (4 physiotherapists (PTs), 3 speech and language therapists (SLTs), 1 dietitian, and 1 midwife) were purposively recruited based on their experience with telehealth. All participants were licensed and currently practicing health care professionals with at least 2 years of experience in their field. They each had conducted a minimum of 10 video-based telehealth sessions in the preceding quarter and practiced in an area where telehealth is applicable. Their professional experience ranged from 2 to 32 years, and encompassed diverse health care settings such as neurorehabilitation, musculoskeletal care, adults, and pediatric care, bariatric medicine, mixed practice settings, respiratory therapy, independent midwifery, and pelvic health. In total, 4 panelists were female and 5 were male.

Interviews

The expert interviews were conducted using Zoom videoconferencing software (Zoom Video Communications) in September 2022. Specifically, 4 interview sessions took place: 3 small-group focus groups and 1 individual interview. Focus group 1 included 2 PTs and lasted 1 hour 34 minutes; focus group 2 included 2 SLTs and lasted 1 hour 3 minutes; and focus group 3 included 2 PTs, 1 SLT, and 1 dietitian, lasting 1 hour and 32 minutes. The one-on-one interview was conducted with a midwife and lasted 29 minutes. We used a semistructured interview guide to help focus the discussions, which was based on commonly reported barriers with video call–based telehealth, found in the existing research for these health care fields (Multimedia Appendix 1). All interviews were conducted by the lead author (LR), who is experienced in qualitative research methods and trained in facilitating interviews and focus groups.

The barriers discussed during the expert interviews were identified through a preceding scoping review [9]: Lack of technology access, unstable internet connections, lack of privacy & IT security, lack of technology skills, unreliability of hardware and software, reluctance of patients or caregivers to use telehealth, decreased establishment of relationships and trust, decreased validity of observations and assessments, lack of hands-on methods, decreased quality and effectiveness, lack of telehealth training, increased workload, lack of support for patients from caregivers, increased safety risks for patients, and unsuitability of setting for telehealth. Experts were asked to provide insights and best practices for addressing these barriers, based on their experience.

Each interview was transcribed verbatim by the lead author (LR) with support from an automated transcription software. After automated transcription, LR reviewed the transcripts to ensure accuracy and fidelity to the original recordings. The final transcripts were analyzed using qualitative content analysis, following the framework proposed by Kuckartz and Rädiker [19]. The first author (LR) conducted the initial deductive coding by extracting passages that contained strategies for overcoming telehealth barriers. These excerpts were then assigned to the corresponding barriers. For each identified strategy, the main characteristics and essential elements were extracted to form a clear and actionable statement. To enhance the reliability of the coding, the research team engaged in multiple team rounds to discuss the initial findings. During these sessions, the team identified similarities across different strategies, combined overlapping or related strategies to reduce redundancy, and formulated precise statements that accurately captured the essence of each strategy. Through iterative discussions, codes were refined and adjusted to better represent the data, thereby enhancing the credibility of the findings.

Ethical Considerations

Prior to participation, all individuals provided informed consent, thereby ensuring ethical compliance and voluntary engagement in the study. The ethics committee confirmed the adherence to ethical principles, and an ethical approval was not required as the study involved only expert consultations without collecting or processing sensitive patient data. All participants gave informed consent, and data collection did not involve patient data. The Ethics Committee for Research Activities of FH Campus Wien, University of Applied Sciences, confirmed that no formal approval was needed for this study (EK 243/2025). The participants did not receive any financial compensation.

Delphi Rounds

Following the qualitative interviews, we conducted Delphi rounds to achieve expert consensus on best practices for video call–based telehealth. The Delphi process was carried out in 2 rounds. Round 1 was initiated at the beginning of November 2022 and remained open for 2 weeks. During this round, panelists received a web-based survey containing a demographic data section, including name and profession, and 15 sections on barriers and corresponding strategies. Round 2 was conducted at the end of November and beginning of December 2022, also open for 2 weeks. In this round, panelists were presented with a survey of 11 sections with the remaining best practices that were lacking consensus.

Barriers

Barriers were rated for their perceived magnitude (1=very small barrier, 2=rather small barrier, 3=rather big barrier, 4=very big barrier) in round 1. Frequency distributions were calculated to determine the perceived magnitude of each barrier among panelists. These results serve to contextualize the forthcoming best practice recommendations.

Analysis of Best Practices

Best practices, on the other hand, were rated for their perceived helpfulness (1=not helpful, 2=rather not helpful, 3=rather helpful, 4=very helpful). In recognition of the diverse professional backgrounds of the panelists, the survey items included an additional response option: “This practice is not relevant for my profession.” Additionally, the survey invited the panelists to contribute further comments and propose new best practices not previously covered, thereby contributing to a more comprehensive dataset.

XSL•FO
RenderX

Response percentages were calculated, and the perceived helpfulness of each practice was quantified using frequency distributions for each response category. Following the first survey round, any practice that received a "helpful" or "very helpful" rating from ≥80% of participants and elicited no comments was automatically included in the final guide. Practices with ratings between 50% and 79% or those that received substantive comments moved on to the second round for reevaluation. Regarding practices rated "helpful" or "very helpful" by <50% of participants, we weighed the risk of prematurely discarding potentially relevant items against the need to maintain focus on practices demonstrating at least a minimal level of acceptance by the panel. After careful consideration, the study team decided to exclude these practices immediately following round 1, rather than carrying them forward. However, participants were given opportunities to propose modifications or new practices at the end of each round, ensuring that any critical practice could be reintroduced or reframed if panelists believed it had merit. The research team discussed all received comments and considered if any rephrasing based on the comments should be made. In case the comments led to a rephrasing of a best practice, both the original and revised versions were presented in the second round for further evaluation. During this round, panelists were provided with the remaining best practices along with the percentage of previous agreement and anonymized comments for each best practice in December 2022. Participants were offered 2 response options: "This practice should be included in the best practice guide" or "This practice should not be included in the best practice guide." When 2 variations of a practice were presented, panelists could select their preferred version or "none of both." Practices that carried over from round 1 required at least 75% agreement among participants to be included in the final best practice guide [20]. This slightly lower threshold accounted for the fact that these practices already showed moderate support or warranted further discussion due to comments. In case 2 versions of a best practice existed, the version receiving the higher preference was selected.

Results

Barriers and Best Practices

The qualitative content analysis on the expert interviews yielded a total of 422 coded text segments. From this process, an initial set of 131 practices was developed. The association of the practices with the discussed barriers can be found in Table 1, and the perceived importance of the 15 presented barriers is shown in Figure 1.

In the first Delphi round, 108 practices exceeded the 80% approval threshold for perceived helpfulness; 83 of these practices had no comments and were therefore included directly. In total, 48 practices met the criteria to proceed to the second round: 23 practices rated as "very helpful" or "helpful" by 50%-79% of the experts; 18 without changes but with comments from the first round; and 7 practices with alternative revision based on comments. One entirely new strategy was introduced (Figure 2). Upon completion of the second round, the final best practice guide emerged with 105 practices, covering practices addressing 5 technology barriers and 10 practical barriers (Table 2 and Textbox 1).

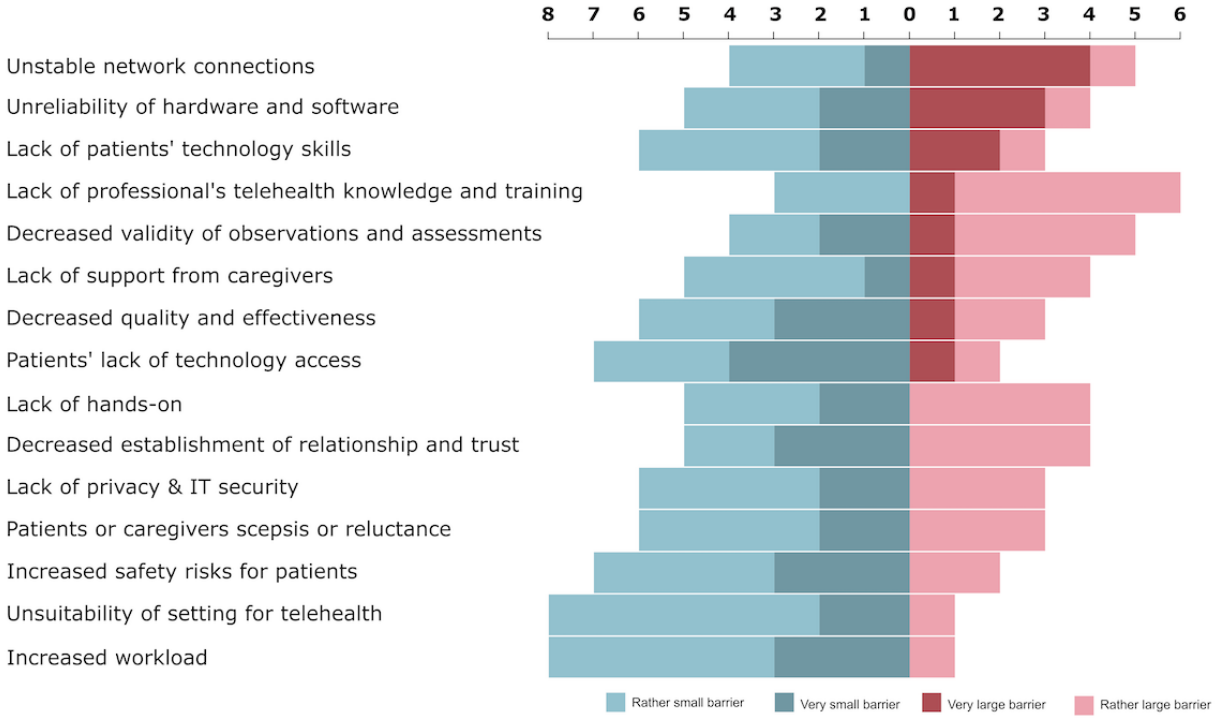
In total, we identified 105 best practices addressing 15 telehealth barriers, with 20 aimed at technology challenges and 85 focused on practice issues.

Table 1. Demographic characteristics of the panelists.

Panelist number	Gender	Profession	Years of professional experience	Professional experience
1	Male	Physiotherapist	15	Orthopedics, traumatology, neurolrehabilitation
2	Male	Physiotherapist	3	Orthopedics, musculoskeletal, training
3	Female	Speech and language therapist	20	Swallowing disorders, functional voice disorders
4	Female	Speech and language therapist	2	Childhood speech development disorders
5	Male	Physiotherapist	32	Orthopedics, respiratory therapy
6	Female	Dietitian	4	Bariatrics
7	Male	Speech therapist	2.5	General speech therapy, respiratory therapy
8	Female	Physiotherapist	20	Pelvic floor therapy
9	Female	Midwife	28	Birth preparation, support, and aftercare

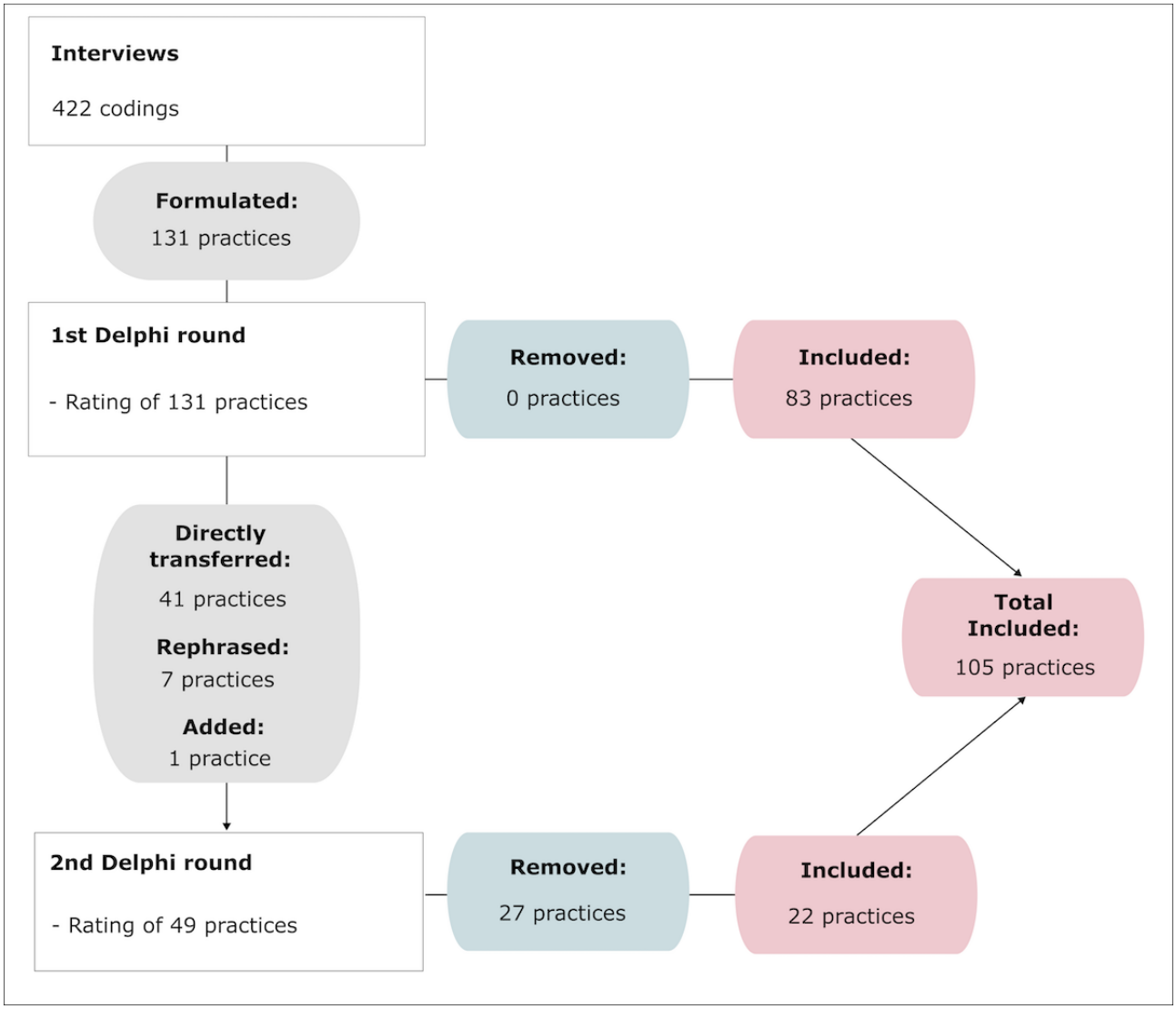
XSL•FO
RenderX

Figure 1. Panelists' (n=9) perceived magnitude of the 15 barriers presented (bars represent absolute frequencies).



XSL•FO
RenderX

Figure 2. Flowchart of the Delphi process.



XSL•FO
RenderX

Table 2. Number of best practices for each presented barrier.

Barrier	Number of practices after expert interviews	Number of practices, included after round 1	Number of practices included after round 2	Final number of practices
Technology barriers				
Unstable network connections	9	3	3	6
Unreliability of hardware and software	4	1	1	2
Lack of technology skills	6	2	2	4
Lack of privacy and IT security	7	2	3	5
Lack of technology access	3	0	3	3
Health care practice barriers				
Lack of professional's telehealth knowledge and training	8	6	0	6
Decreased validity of observations and assessments	13	8	2	10
Lack of support from caregivers	3	1	0	1
Lack of hands-on	11	7	3	10
Reluctance of patients or caregivers to use telehealth	8	5	2	7
Decreased establishment of relationship and trust	11	9	1	10
Decreased quality and effectiveness	10	9	1	10
Increased safety risks for patients	13	10	1	11
Unsuitability of setting for telehealth	21	18	0	18
Increased workload	4	2	0	2
Sum total	131	83	22	105

XSL•FO
RenderX

Textbox 1. Best practice guide for barriers to telehealth via video call.

Best practices to address technology barriers

A. Unstable network connections

1. Every participant should choose a location with a stable network connection for video call sessions. This should be communicated to the clients in advance.
2. Other devices, persons, or activities should not use the same network connection during the video call session.
3. In case of an unstable network connection, the device may be exchanged.
4. In case of an unstable network connection, switching off the camera for a short time could help.
5. In case of network disconnection, the client should be informed about further steps by telephone (full agreement from all participants).
6. In case of audio transmission problems, an additional phone connection can be used.

B. Unreliability of hardware and software

1. The video call software used should be installed and tested in a timely manner by all participants before the start of the session.
2. If clients experience technical issues, possible solutions can be discussed and solved directly during the video call session or via telephone (full agreement from all participants).

C. Lack of technology skills

1. Health care professionals should gain digital literacy skills in using different devices and video call software, as this can help in solving common technical problems encountered by patients (full agreement from all participants).
2. For technically non-proficient individuals, user-friendly video call software should be selected. Getting started via a direct link without prior installation should be favored (full agreement from all participants).
3. Telephone support can help inexperienced individuals get started with video call software (full agreement from all participants).
4. As required, the operation of the video call software can be supported at the beginning or ongoing by relatives or other persons on site.

D. Lack of privacy & IT security

1. Software used should ensure encrypted data transmission (full agreement from all participants).
2. One's own decisions regarding the handling of client data should be made and communicated in accordance with current data protection laws.
3. Written consent from clients should be obtained in advance.
4. When using video call software, access should only be possible for invited individuals (full agreement from all participants).
5. If software has not been recommended by a certification body for use in health care, it should be reviewed concerning its data processing.

E. Lack of technology access

1. Providing a checklist of minimum technical requirements for video call sessions can optimize conditions in advance.
2. Availability of equipment should be confirmed prior to the beginning of the session.
3. For sessions that focus on visual demonstration or perception, it is recommended to use a device with a large, high image quality screen.

F. Lack of telehealth knowledge and training

1. Watching video call software tutorials can provide information on available functions and give confidence in handling.
2. Learning by doing can help to gain confidence in conducting video call sessions.
3. A test run with individuals from the private environment can provide confidence in using the technology (full agreement from all participants).
4. Materials for use in video call sessions can be researched on the internet (full agreement from all participants).
5. Reading current articles on the topic of telehealth can increase knowledge (full agreement from all participants).
6. Exchanging experiences with colleagues about the use of video units can expand one's own repertoire (full agreement from all participants).

G. Decreased validity of observations and assessments

1. Important assessments should be conducted in a supplementary face-to-face unit if specific information cannot be gathered via a video call session.
2. Clients should be informed that their clothing of choice should allow detailed observations of the body (full agreement from all participants).
3. Cameras should be positioned in a way that enables a stable image (full agreement from all participants).

XSL•FO
RenderX

4. Detailed and guided instructions on how to position the body or involved body parts in front of the camera should be given to ensure high-quality observations (full agreement from all participants).
5. The background should provide visual reference points for guiding and observing movements.
6. The observation quality of details can be improved by moving the corresponding body section closer to the camera (full agreement from all participants).
7. The observation quality of details can be improved with targeted lighting (full agreement from all participants).
8. Assessments can be adapted with the help of using objects in the client´s room during video call sessions (full agreement from all participants).
9. Increased attention should be given to small nonverbal signals of the clients.
10. Involving self-observation and self-perception of clients can supplement limited on-screen observation possibilities (full agreement from all participants).

H. Lack of support from caregivers

1. Ensuring that a support person is within reach can help to run a video call session smoothly (full agreement from all participants).

I. Lack of hands-on

1. Education about functional deficits and how to deal with them in everyday life should be emphasized in video call sessions.
2. Training self-awareness and empowering clients in their self-efficacy can be part of video call sessions to achieve lasting effects (full agreement from all participants).
3. Proven effectiveness of therapy interventions aiming at lifestyle modifications should be explained to clients (full agreement from all participants).
4. Exercises that can be performed by the clients independently should be selected (full agreement from all participants).
5. Exercises should be guided verbally in such a manner that they can be understood even if visual guidance is limited (full agreement from all participants).
6. In advance, material needed for the video call session should be communicated to the clients or relatives, and it should be prepared in a timely manner (full agreement from all participants).
7. Selected self-treatments can be taught to clients during the video call session (full agreement from all participants).
8. Clients should be given specific advice for performing self-control of guided exercises (full agreement from all participants).
9. Clients' self-awareness can be promoted by feeling movements or structures with their own hands, wherein possible misinterpretations should be considered.
10. Material required for a video call session should be communicated in advance with clear information, for example, on size, quantity, and weight.

J. Patients or caregivers scepsis or reluctance

1. Clients can be offered a combination of presence and video call sessions (full agreement from all participants).
2. It can be explained to skeptical clients that some therapy content is suitable for video call sessions, others not (full agreement from all participants).
3. Demonstrating video call software and discussing a possible video call session can reduce insecurities and help with decision-making.
4. A single video call session for familiarization can be offered to skeptical clients, followed by a discussion about the experience and further procedure.
5. After a few face-to-face sessions, video call sessions can be suggested again.
6. Involving relatives or the social environment can be facilitated with video call sessions.
7. The reduction of travel and increase in time flexibility can be used as an argument for using video call sessions.

K. Decreased establishment of relationship and trust

1. Authenticity and professionalism can be conveyed through conscious design choices and insight into the premises in the background (full agreement from all participants).
2. Optimally adjusted lighting of one's face can create a friendly atmosphere and convey mimic information well (full agreement from all participants).
3. A well-adjusted camera angle and optimal positioning in front of the screen support communication on an equal footing (full agreement from all participants).
4. Facial expressions and gestures during a video call session should be used consciously (full agreement from all participants).
5. Since real eye contact is not possible via the screen, this can be imitated by consciously looking into the camera.
6. Maintaining one's own appearance can promote authenticity in front of the camera (full agreement from all participants).
7. Conscious choice of clothing supports the professional appearance.

XSL•FO
RenderX

8. Seeing one's own camera image can be used as feedback about one’s own appearance (full agreement from all participants).
9. Framing a video call session with introductory words and a deliberate end can provide orientation (full agreement from all participants).
10. Spontaneous playful elements and humor can loosen up video call sessions and promote the relationship (full agreement from all participants).

L. Decreased quality and effectiveness

1. Clients who can benefit from video call sessions should be selected purposefully (full agreement from all participants).
2. The limitations of video call sessions for certain medical conditions should be clearly communicated to clients (full agreement from all participants).
3. The use of video call sessions should be evaluated regularly and adjusted if necessary.
4. Strategies to ensure whether elements learned in the session can be implemented independently should be used.
5. Video call sessions can be used to maintain a regular appointment frequency (full agreement from all participants).
6. By collecting relevant health parameters at the beginning and the end of the session, the effectiveness of treatment interventions can be evaluated.
7. At the beginning of a video call session, expectations should be gathered. At the end they should be compared with the session's content and success.
8. The role of relatives in the treatment process can be strengthened through targeted counseling during video call sessions.
9. By setting a clear focus and sequence for the video call session and individualizing the duration, one’s ability to focus can be considered (full agreement from all participants).
10. The video call session should take place during a time of the day when clients have a good ability to focus (full agreement from all participants).

M. Increased safety risks for patients

1. Clear exclusion and termination criteria for video call sessions for clients with safety-critical characteristics should be established and adhered to (full agreement from all participants).
2. Competent relatives can increase clients’ safety.
3. For video call sessions with children, a legal guardian should be present at the child's side (full agreement from all participants).
4. Clients or family members should be educated about making the environment as safe as possible to reduce potential hazards such as tripping, falling, or slipping (full agreement from all participants).
5. Client’s pets that may interfere with free movement in the room should be kept out of the client's room during video call sessions.
6. Exercises instructed via video call sessions should be chosen according to clients' abilities (full agreement from all participants).
7. Video call sessions should not include unfamiliar exercises that may jeopardize client safety.
8. Exercise difficulty should be increased only step by step, adapted to the client (full agreement from all participants).
9. In case of uncertainties regarding perceived health parameters, referral to specialists should be made (full agreement from all participants).
10. In psychologically tense situations, one’s own voice can serve as an instrument for calming down (full agreement from all participants).
11. The daytime of the video call session, and therefore a potentially limited suitability of certain measures, should be taken into consideration.

N. Unsuitability of setting for telehealth

1. A quiet room without distractions should be chosen for video call sessions (full agreement from all participants).
2. Background noise from people, pets, or household equipment should be reduced (full agreement from all participants).
3. For video call sessions, a room with good sound quality and reduced hall should be chosen (full agreement from all participants).
4. The speaker used should be tuned sufficiently loud (full agreement from all participants).
5. Balanced lighting conditions allowing good contrasts should be selected (full agreement from all participants).
6. Devices should be positioned stable to not produce blurry images (full agreement from all participants).
7. If a tablet or smartphone is used, it should be positioned at a 90-degree angle to the surface with a suitable mount (full agreement from all participants).
8. Clients should be guided in the optimal camera setup and positioning in front of it (full agreement from all participants).
9. The camera section should be chosen so that only things that one is willing to share are visible in the background (full agreement from all participants).
10. Other people in the household or work premises should be informed that a video call session is taking place to ensure privacy (full agreement from all participants).
11.

XSL•FO
RenderX

It should be discussed with the clients in which area of their premises the video session will take place in order to avoid the concern of giving insight into things they want to keep private (full agreement from all participants).

12. Windows and doors should be closed.
13. High-quality microphones and headphones can improve acoustic transmission.
14. Advice should be given on how to use built-in or external microphones to ensure good transmission and avoid noise interference.
15. Doors in the camera's field of view should be avoided to prevent unwanted privacy intrusion.
16. A dedicated therapy location—for example, at a table—can promote attention and concentration.
17. Marking one or more places where the camera is well positioned facilitates a sustainably working set-up.
18. If the video call session involves movement across the room, a flexible camera setup for rapid changes should be facilitated.

O. Increased workload

1. Using quick-to-use treatment or exercise methods that require little pre-planning and explanation can reduce preparation time.
2. By combining preparation time for several video units, one's own working time can be used efficiently.

Best Practices to Address Technology Barriers

In total, 20 best practices were selected by the participants to address the barriers of unstable network connections, unreliability of hardware and software, lack of technology skills, lack of privacy and IT security, and lack of technology access (A-E in Textbox 1). For optimal video call sessions, each participant should ensure a stable network connection, free from other devices and activities that might compete for bandwidth. When issues such as an unstable connection arise, switching devices, temporarily turning off the camera, or making a phone call can provide continuity and effective communication. Participants should install and test the video call software well before the session, with technical issues being addressed directly during the call or via telephone if necessary. Health care professionals should develop digital literacy for various devices and software to assist patients with technical issues. For less tech-savvy individuals, user-friendly software with direct links, telephone support, as well as on-site help from relatives, can facilitate the use of video call applications. To ensure data privacy and security, video call software must encrypt data transmission and comply with data protection laws. Informed consent should be obtained from users, and access should be restricted to invited participants. If software has not been recommended by a certification body for use in health care, it should be reviewed concerning its data processing. To optimize video call sessions, a checklist of technical requirements should be provided, and equipment availability should be confirmed beforehand. For visually intensive sessions, a high-quality large screen is recommended.

Best Practices to Address Health Care Practice Barriers

Overall, 85 best practices addressed practice issues such as lack of telehealth knowledge and training, decreased validity of observations and assessments, lack of support from caregivers, lack of hands-on, patients or caregivers scepsis or reluctance, decreased establishment of relationship and trust, decreased quality and effectiveness, increased safety risks for patients, unsuitability of setting for telehealth, and increased workload (F-O in Textbox 1).

Reviewing current telehealth publications and internet resources to search for material to use in telehealth sessions, as well as watching software tutorials, can enhance proficiency and confidence in conducting video call sessions, complemented by practical experience and sharing knowledge with colleagues.

For comprehensive assessments in telehealth, certain procedures may need to be supplemented with in-person visits. Clients should be guided on how to prepare for video sessions with appropriate clothing, camera positioning, lighting, and background setup to enhance the quality of observations. Additionally, using objects for adapted assessments, paying attention to subtle nonverbal cues, and involving clients in self-observation can maximize the efficacy of remote evaluations. Having a support person readily available can facilitate a smooth video call session.

In video call sessions, educating clients on functional deficits and self-management in daily life is crucial, with a focus on training self-awareness and self-efficacy for sustainable outcomes. For effective video call sessions, exercises that clients can perform independently should be chosen, and clear verbal guidance for exercises with limited visual cues should be given. This can be added with the training of self-treatments, advice for self-control during exercises, and the promotion of clients' self-awareness through tactile experiences while addressing potential misinterpretations. Clear communication and preparation of materials is necessary.

To address skepticism, clients can be offered a blend of in-person and video call sessions, emphasizing the suitability of certain therapy contents for remote delivery, with demonstrations and trial sessions helping to alleviate skepticism and showcasing the convenience and flexibility that video calls offer. To convey professionalism and authenticity in video calls, one should make deliberate design and setting choices, ensure optimal lighting for a friendly atmosphere, and use a camera angle that supports equitable communication. Conscious use of facial expressions, gestures, and clothing, along with looking into the camera to simulate eye contact, enhances authenticity. Additionally, framing the session with a clear beginning and end, along with the use of humor, can improve the interaction.

XSL•FO
RenderX

Purposeful selection of clients for video call sessions is essential, with transparent communication regarding their limitations for certain conditions. Regular evaluations should be conducted to assess effectiveness and determine if expectations are being met. Strategies to ensure whether elements learned in the session can be implemented independently should be used. Sessions should be structured in a way that they have a clear focus and sequence, an appropriate duration, and are scheduled for when the client's focus can optimize the session. Incorporating relatives into the treatment process via video call sessions can provide targeted counseling. Establishing clear criteria for excluding or terminating video call sessions is vital for clients with safety-critical concerns. Competent relatives can enhance safety. Guardian presence for children, safety education, and gradual exercise progression can ensure the clients' well-being during remote sessions.

For effective video call sessions, it is essential to choose a quiet room with as little background noise as possible, ensuring good sound quality and balanced lighting. Proper camera setup, privacy considerations, and guidance on microphone use enhance the session quality. Additionally, creating a dedicated therapy location and facilitating flexible camera setups can improve the overall experience and engagement during sessions. Efficient use of time in video sessions can be achieved by opting for quick and easy treatment methods and combining preparations for multiple video units.

Discussion

Principal Findings

The findings from this study, which encompassed expert interviews and a modified Delphi process, provide a best practice guideline for telehealth in health care. The key results underscore the importance of addressing the challenges of technology accessibility, enhancing observations and assessments, creating a professional setup, building relationships and trust, and ensuring privacy and IT security. Although the majority of the panelists were PTs and SLTs, their combined professional experiences—spanning areas such as neurorehabilitation and pediatric care—have facilitated a comprehensive exploration of these practices across various health care settings, thereby complementing the aims and findings of previous studies and guidelines [21-25].

The best practice guideline offers a starting point towards a unified guide to improve telehealth across various health care settings. Prior publications have established telehealth frameworks in disciplines such as occupational therapy, physiotherapy, and speech and language therapy [21-24], which partially align with our findings. A high level of agreement is seen in these existing frameworks, particularly regarding fundamental principles of telehealth delivery. However, there remain clear gaps in the current literature on how to manage these barriers through concise best practice statements [24]. Grogan-Johnson [25] discusses the challenges of maintaining engagement in virtual settings but provides limited guidance on how to address this. Our best practice guide suggests trust-building strategies, such as non-verbal communication enhancements, environmental adjustments, and interactive session structuring. Davies et al [13] outline the importance of technological proficiency for PTs but do not provide detailed practices for managing unstable networks or unreliable hardware, where our study adds explicit recommendations for overcoming technological challenges, including strategies for low-bandwidth settings, selecting robust hardware, and providing patient education on basic troubleshooting. Other studies limit their focus to specific populations or pathologies [22,25]. Cottrell and Russell [22], for instance, provide an in-depth analysis of pre-implementation considerations and practical steps specifically within the domain of musculoskeletal physiotherapy. Their work examines technological and operational challenges, offering targeted strategies to enhance the successful deployment of telehealth services. Other guidelines have delineated key technical principles health care professionals are expected to follow to ensure a smooth and secure telehealth experience [26,27], focusing on adherence to laws and regulations and the implementation of privacy and security measures. Our study extends these principles by proposing practical, applicable strategies and emphasizes dual-focused education, targeting both health care providers and patients. It includes specific methods to address skepticism and enhance digital literacy, particularly for populations unfamiliar with telehealth.

Notably, our research revealed a strong consensus on the effectiveness of specific strategies across health care settings, such as providing detailed technical guidelines, enhancing visual and audio components of telehealth sessions, and incorporating elements that foster a sense of connection and trust between health care providers and clients. These findings offer new insights into the practical implementation and adaptation of telehealth services in diverse health care contexts.

In examining the distribution of best practices identified to address various barriers, we found notable differences. "Unsuitability of setting for telehealth" yielded the largest set of best practices despite being rated as a relatively small barrier. This contrast likely arose because many aspects of the setting (eg, lighting, device positioning, noise reduction) are easily modified by health care professionals. These direct, practitioner-driven solutions help frame the barrier as "small," in that it can be quickly addressed through practical steps. In other words, the perception of a barrier's magnitude may reflect how empowered providers feel to remedy it: settings are simpler to adjust, whereas more systemic issues (eg, workload, limited technology access) require external interventions and may thus seem more significant.

Additionally, a significant collection of best practices has been identified to address practical implementation barriers such as "increased safety risks for patients," "decreased validity of observations and assessments," "lack of hands-on," "decreased establishment of relationship and trust," and "decreased quality and effectiveness." These best practices provide pragmatic strategies to navigate the range of challenges presented by telehealth, ensuring a higher standard of patient care and treatment integrity. On the other side, there were barriers that received fewer best practices to address them, such as "lack of support from caregivers" and "increased workload," both at the lower end, with 1 and 2 identified best practices. These issues

XSL•FO
RenderX

likely require organizational or systemic changes that are beyond the immediate control of the individual practitioner. Similarly, "lack of technology access" and "unreliability of hardware and software" are countered by only 3 best practices each, possibly indicating a gap between the technological requirements of telehealth and the ability of health professionals to influence these parameters.

Reflections on the Delphi Method

Our modified Delphi approach relied on 2 rounds, underpinned by preliminary interviews that generated a well-defined set of barriers and best practices. This groundwork contributed to a high level of initial consensus—83 of 108 items achieved ≥80% agreement in round 1. While we acknowledge that time constraints and difficulty finding sufficient participants limited the scope of our panel (n=9), each participant brought substantial experience: all met strict inclusion criteria (eg, minimum 2 years' professional practice, at least 10 video-based telehealth sessions in the previous quarter), ensuring specialized and recent expertise.

We employed a 2-threshold system: items achieving ≥80% agreement in round 1 were immediately included, while items carried forward required ≥75% agreement in round 2. This tiered approach balanced efficiency and adaptability by refining practices with moderate support without burdening participants with repeated rounds. Items garnering <50% agreement were excluded after round 1, given their limited traction; however, participants could introduce revisions or propose new items if they felt critical information was overlooked. Although this may appear stringent, it enabled us to produce a focused, consensus-driven guide for reducing barriers to video call–based telehealth within the time and resource constraints of our project.

While our study does not claim to capture every possible facet of telehealth across all health care contexts, the methodology employed aligned with recommended standards.

Limitations

This study still has several limitations that need to be acknowledged. First, the sample size was small, consisting of only 9 eligible experts in specific health professions. Our participant sample was relatively small, yet it was purposefully recruited to encompass diverse professional backgrounds and extensive telehealth experience. This approach generated a robust consensus on best practices for video call–based telehealth. Future studies with larger or more varied panels could further enhance and validate the findings, but our current methods provide a strong foundation for ongoing research and practical implementation. Second, the heterogeneity of the participants' professions may be both an advantage and a disadvantage. On the one hand, it allowed for a broader range of perspectives and strategies to be considered than if we would have focused on only one profession. On the other hand, it could also have resulted in the exclusion of profession-specific strategies that may be essential for a successful telehealth implementation, which could be studied in the future. Each profession uses telehealth differently based on their specific clinical needs and methods. For instance, PTs may focus more on movement assessments and exercise guidance, whereas SLTs might emphasize communication techniques and auditory assessments. This variation means that certain best practices identified may be more applicable to some professions than others, potentially limiting the universal applicability of the recommendations. The small number of distinct professions included in the study may have resulted in the exclusion of specialized best practices pertinent to each field. Essential strategies unique to specific clinical scenarios or patient interactions within each profession might not have been fully captured or emphasized, thereby affecting the comprehensiveness of the best practice guide for individual disciplines.

Moreover, the findings reflect the current state and must be interpreted as such. They likely are subject to change due to the continuous evolution of professional expertise in health care and ongoing technological advancements that are likely to shape the future landscape of telehealth. Furthermore, this study relied on self-reported data from the experts. While this method is commonly used in qualitative research, it is important to acknowledge that the data collected may be subject to bias or limitations in recall. Another important consideration is the length of the final best practice list. Although it represents a comprehensive set of potential solutions, its sheer number may pose challenges for day-to-day clinical application.

Implication of Findings

The study's findings offer valuable insights for various stakeholders within the telehealth ecosystem, promoting enhanced effectiveness, accessibility, and sustainability of video call–based telehealth services. The best practices serve as a practical guide for clinicians to optimize telehealth delivery. Implementing these strategies can improve care quality, strengthen patient-provider relationships, and increase patient satisfaction while reducing operational burdens through streamlined session management and technical troubleshooting. Using enhanced privacy measures and user-friendly technologies ensures a secure and confident telehealth experience for patients, particularly those in remote or underserved areas. Insights from this study can also inform the development of regulations and standards that promote equitable telehealth access. By addressing practical barriers and supporting measures such as technology funding and training programs, policy makers can facilitate the broader adoption and integration of telehealth into standard health care practices. Finally, incorporating these best practices into health care curricula ensures that future practitioners are equipped with the necessary digital literacy and remote communication skills. Training programs can emphasize technical competencies and patient engagement strategies specific to telehealth, preparing health care workers for modern care delivery.

Further Research

While this Delphi study provided insights into potential solutions, it would be beneficial to evaluate the impact of these strategies on the successful implementation of telehealth in different real-world settings by exploratory and confirmatory studies. For example, while this study focused on barriers and solutions in the selected health care professions, other health care specialties and settings may face different challenges and

XSL•FO
RenderX

would come up with different best practices when implementing telehealth. Evaluating these barriers and solutions in different contexts would provide a more comprehensive understanding of telehealth implementation. Moving forward, further research could investigate strategies to streamline or prioritize these recommendations, identifying which are most beneficial for specific professional groups or patient populations. Further, it would be interesting to create educational programs to see how teaching these strategies to students would increase their telehealth proficiency.

Conclusions

This publication presents a modified Delphi study that compiles best practices for overcoming barriers in video call–based telehealth. It includes insights from various health care professionals and focuses on enhancing telehealth's quality and effectiveness. The study emphasizes technology accessibility, professional setups, and patient-centered care. Limitations include a small sample size and self-reported data, suggesting future research should assess these strategies' real-world effectiveness, explore barriers in different health care specialties, and develop educational programs to increase telehealth proficiency.

Acknowledgments

The authors wish to thank the participants of the Delphi study. This work was supported by the City of Vienna, Magistratsabteilung 23, Austria (grant MA23-338474-2021-2).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

LR contributed to project administration; LR and LA were involved in conceptualization, supported by all authors; LR and PP contributed to methodology; data curation was done by LR; formal analysis and investigation were done by LR and LM; LR involved in writing—original draft preparation; all authors contributed to writing—review and editing; funding acquisition was done by LR and FW; supervision was done by SK.

Conflicts of Interest

SK is the founder and shareholder of MED.digtal. The other authors declare no conflicts of interest.

Multimedia Appendix 1
Interview Guideline_Translated version_English.
[PDF File (Adobe PDF File), 63 KB - humanfactors_v12i1e64079_app1.pdf]

References

1. Mechanic OJ, Persaud Y, Kimball AB. Telehealth Systems. StatPearls Treasure Island, FL: StatPearls Publishing; 2022.
2. Orlando JF, Beard M, Kumar S. Systematic review of patient and caregivers' satisfaction with telehealth videoconferencing as a mode of service delivery in managing patients' health. PLoS One 2019;14(8):e0221848 [FREE Full text] [doi: 10.1371/journal.pone.0221848] [Medline: 31469865]
3. Monaghesh E, Hajizadeh A. The role of telehealth during COVID-19 outbreak: a systematic review based on current evidence. BMC Public Health 2020;20(1):1193 [FREE Full text] [doi: 10.1186/s12889-020-09301-4] [Medline: 32738884]
4. Almathami HKY, Win KT, Vlahu-Gjorgievska E. Barriers and facilitators that influence telemedicine-based, real-time, online consultation at patients' homes: systematic literature review. J Med Internet Res 2020;22(2):e16407 [FREE Full text] [doi: 10.2196/16407] [Medline: 32130131]
5. Gellis ZD, Kenaley B, McGinty J, Bardelli E, Davitt J, Ten Have T. Outcomes of a telehealth intervention for homebound older adults with heart or chronic respiratory failure: a randomized controlled trial. Gerontologist 2012;52(4):541-552. [doi: 10.1093/geront/gnr134] [Medline: 22241810]
6. Kruse CS, Krowski N, Rodriguez B, Tran L, Vela J, Brooks M. Telehealth and patient satisfaction: a systematic review and narrative analysis. BMJ Open 2017;7(8):e016242 [FREE Full text] [doi: 10.1136/bmjopen-2017-016242] [Medline: 28775188]
7. Snoswell CL, Taylor ML, Comans TA, Smith AC, Gray LC, Caffery LJ. Determining if telehealth can reduce health system costs: scoping review. J Med Internet Res 2020;22(10):e17298 [FREE Full text] [doi: 10.2196/17298] [Medline: 33074157]
8. Snoswell CL, Smith AC, Page M, Scuffham P, Caffery LJ. Quantifying the societal benefits from telehealth: productivity and reduced travel. Value Health Reg Issues 2022;28:61-66 [FREE Full text] [doi: 10.1016/j.vhri.2021.07.007] [Medline: 34800833]

XSL•FO
RenderX

9. Rettinger L, Kuhn S. Barriers to video call-based telehealth in allied health professions and nursing: scoping review and mapping process. J Med Internet Res 2023;25:e46715 [FREE Full text] [doi: 10.2196/46715] [Medline: 37526957]
10. Van Eerdenbrugh S, Schraeyen K, Leysen H, Mostaert C, D'haenens W, Vandenborre D. Delivery of speech-language therapy and audiology services across the world at the start of the COVID-19 pandemic: a survey. Perspect ASHA Spec Interest Groups 2022;7(2):635-646. [doi: 10.1044/2021_persp-21-00134]
11. Rosenstrøm S, Groth S, Risom S, Hove J, Brødsgaard A. Nurses' experiences with virtual consultations and home-monitoring in patients with cardiac disease: a systematic review and qualitative meta-synthesis of results. Nurs Health Sci 2024;26(4):e13180. [doi: 10.1111/nhs.13180] [Medline: 39437992]
12. Greenhalgh T, Shaw S, Wherton J, Vijayaraghavan S, Morris J, Bhattacharya S, et al. Real-world implementation of video outpatient consultations at macro, meso, and micro levels: mixed-method study. J Med Internet Res 2018;20(4):e150 [FREE Full text] [doi: 10.2196/jmir.9897] [Medline: 29625956]
13. Davies L, Hinman RS, Russell T, Lawford B, Bennell K, International Videoconferencing Steering Group. An international core capability framework for physiotherapists to deliver quality care via videoconferencing: a Delphi study. J Physiother 2021;67(4):291-297 [FREE Full text] [doi: 10.1016/j.jphys.2021.09.001] [Medline: 34521586]
14. Jennett PA, Andruchuk K. Telehealth: 'real life' implementation issues. Comput Methods Programs Biomed 2001;64(3):169-174. [doi: 10.1016/s0169-2607(00)00136-x] [Medline: 11226614]
15. Nasa P, Jain R, Juneja D. Delphi methodology in healthcare research: how to decide its appropriateness. World J Methodol 2021;11(4):116-129 [FREE Full text] [doi: 10.5662/wjm.v11.i4.116] [Medline: 34322364]
16. Spranger J, Homberg A, Sonnberger M, Niederberger M. Reporting guidelines for Delphi techniques in health sciences: a methodological review. Z Evid Fortbild Qual Gesundhwes 2022;172:1-11 [FREE Full text] [doi: 10.1016/j.zefq.2022.04.025] [Medline: 35718726]
17. Kelly LM, Cordeiro M. Three principles of pragmatism for research on organizational processes. Methodol Innov 2020;13(2):205979912093724. [doi: 10.1177/2059799120937242]
18. Allemang B, Sitter K, Dimitropoulos G. Pragmatism as a paradigm for patient-oriented research. Health Expect 2022;25(1):38-47 [FREE Full text] [doi: 10.1111/hex.13384] [Medline: 34748689]
19. Kuckartz U, Rädiker S. Qualitative Content Analysis: Methods, Practice and Software. Thousand Oaks: SAGE Publications; 2023.
20. Diamond IR, Grant RC, Feldman BM, Pencharz PB, Ling SC, Moore AM, et al. Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. J Clin Epidemiol 2014;67(4):401-409. [doi: 10.1016/j.jclinepi.2013.12.002] [Medline: 24581294]
21. Almubark BM, Majrashi N, Alghusun N, Alhammad M, Alhthifi F, Alyahya RS. Telehealth clinical practice guide for occupational therapy, physical therapy, and speech and language pathology: a Saudi and middle eastern guide. Telemed J E Health 2022;28(5):636-642. [doi: 10.1089/tmj.2021.0021] [Medline: 34529497]
22. Cottrell MA, Russell TG. Telehealth for musculoskeletal physiotherapy. Musculoskelet Sci Pract 2020;48:102193 [FREE Full text] [doi: 10.1016/j.msksp.2020.102193] [Medline: 32560876]
23. World Federation Of Occupational Therapists. World federation of occupational therapists' position statement on telehealth. Int J Telerehabil 2014;6(1):37-39 [FREE Full text] [doi: 10.5195/IJT.2014.6153] [Medline: 25945221]
24. Cason J, Hartmann K, Jacobs K, Richmond T. Telehealth in occupational therapy. Am J Occup Ther 2018;72(Supplement_2):7212410059p1-7212410059p18. [doi: 10.5014/ajot.2018.72S219] [Medline: 30674404]
25. Grogan-Johnson S. The five W's meet the three R's: the who, what, when, where, and why of telepractice service delivery for school-based speech-language therapy services. Semin Speech Lang 2021;42(2):162-176 [FREE Full text] [doi: 10.1055/s-0041-1723842] [Medline: 33725733]
26. Anil K, Freeman JA, Buckingham S, Demain S, Gunn H, Jones RB, et al. Scope, context and quality of telerehabilitation guidelines for physical disabilities: a scoping review. BMJ Open 2021;11(8):e049603 [FREE Full text] [doi: 10.1136/bmjopen-2021-049603] [Medline: 34385253]
27. Leone E, Eddison N, Healy A, Royse C, Chockalingam N. Exploration of implementation, financial and technical considerations within allied health professional (AHP) telehealth consultation guidance: a scoping review including UK AHP professional bodies' guidance. BMJ Open 2021;11(12):e055823 [FREE Full text] [doi: 10.1136/bmjopen-2021-055823] [Medline: 34969656]

Abbreviations

PT: physiotherapist
SLT: speech and language therapist

XSL•FO
RenderX

Edited by C Jacob; submitted 08.07.24; peer-reviewed by A Krithika, T van Houwelingen, L Davies; comments to author 02.12.24; revised version received 22.01.25; accepted 03.03.25; published 26.03.25.

Please cite as:
Rettinger L, Aichinger L, Ertelt-Bach V, Huber A, Javorszky SM, Maul L, Putz P, Sargis S, Werner F, Widhalm K, Kuhn S
Best Practice Guide for Reducing Barriers to Video Call–Based Telehealth: Modified Delphi Study Among Health Care Professionals
JMIR Hum Factors 2025;12:e64079
URL: https://humanfactors.jmir.org/2025/1/e64079
doi:10.2196/64079
PMID:

©Lena Rettinger, Lea Aichinger, Veronika Ertelt-Bach, Andreas Huber, Susanne Maria Javorszky, Lukas Maul, Peter Putz, Sevan Sargis, Franz Werner, Klaus Widhalm, Sebastian Kuhn. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 26.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluating the Needs and Characteristics of Individuals of Low Socioeconomic Status Using Digital Health Technology to Address Health-Related Social Needs: Mixed Methods Study With Patients and Care Providers

Kayla Li Haydon[1], MSc; Cynthia LeRouge[2], PhD, MS; Polina Durneva[3], PhD; Maria Paula Diaz Campo[4], BS; David Richard Brown[1], MD

[1]Herbert Wertheim College of Medicine, Florida International University, Miami, FL, United States

[2]School of Public Health and School of Medicine, University of Washington, Seattle, WA, United States

[3]Department of Information Systems and Business Analytics, Loyola Marymount University, Los Angeles, CA, United States

[4]Jarvis College of Computing and Digital Media, DePaul University, Chicago, IL, United States

Corresponding Author:
Kayla Li Haydon, MSc
Herbert Wertheim College of Medicine
Florida International University
11200 SW 8th St
Miami, FL, 33199
United States
Phone: 1 8134037898
Email: kli025@fiu.edu

Abstract

Background: Social determinants of health (SDOH) are the conditions in which people are born, grow, live, work, and age, encompassing social and economic factors that shape health outcomes. There is an increasing call to leverage digital health technology (DHT) to address SDOH and health-related social needs and establish connections to resources and services.

Objective: This study aimed to (1) identify the DHT-related characteristics of DHT users with low socioeconomic status (SES), (2) determine the needs and preferences of DHT users with low SES, and (3) explore how current SDOH-DHT addresses these needs and preferences in addressing their health-related social needs.

Methods: We used a multiphase, mixed method, user-centered design approach. In phase 1, we developed a user profile based on a literature review, aggregate data, interviews with 26 low-SES individuals, and focus groups with 28 professionals. In phase 2, we conducted a landscape analysis of 17 existing SDOH-DHTs.

Results: DHT users of low SES had diverse social and technology characteristics. Five key themes emerged regarding user needs and preferences: (1) user-centered design, including multilingual support, visual guidance, and customization; (2) efficient, solution-based assessment of social risks, assets, and needs; (3) e-caring support features; (4) user education and feedback mechanisms; and (5) trust, privacy, and security. The landscape analysis revealed that current SDOH-DHT features do not adequately meet these needs.

Conclusions: Discrepancies between target user needs and current DHT features represent missed opportunities in developing user-centered tools for individuals of low SES. Findings underscore the importance of inclusive, empowering, and responsive design in SDOH-DHT to bridge health disparities and advance public health.

(JMIR Hum Factors 2025;12:e69545) doi:10.2196/69545

KEYWORDS

social determinants of health; health-related social needs; consumer health informatics; mobile health; digital inclusion

XSL•FO
RenderX

Introduction

Background

Social determinants of health (SDOH) are "the conditions in which people are born, grow, live, work, and age," encompassing social and economic factors that shape health outcomes [1]. These determinants can function as assets or risk factors, with individual-level adverse social conditions being defined as social risk factors [2]. "Health-related social needs" (HRSN) differ from social risk factors by reflecting both the presence of risks and individuals' preferences for addressing them [2]. These immediate needs, such as food insecurity and social isolation, stem from upstream SDOH, including economic stability, education, health care access, and the physical environment [3]. The resulting social gradient creates a corresponding health gradient with particular impact on lower socioeconomic status (SES) populations [4-6].

Studies have highlighted the impact of SES on both health outcomes and technology use, underscoring the importance of considering SES in health technology research [4,7,8]. A broad definition of "low SES" includes a lack of resources (eg, education and income) or measures of poverty (eg, food insecurity and income-to-needs ratio) [9]. Individuals of low SES often lack the knowledge or resources to respond to social needs and maintain health [6]. Bourdieu's theory of capital—economic, social, and cultural—provides a useful framework for understanding how these determinants operate [10]. Economic capital refers to financial resources, social capital to networks and relationships, and cultural capital to knowledge and dispositions that confer social value. Each form of capital can impact health outcomes and access to health care [11]. Unmet social needs are closely associated with disparities in health outcomes, quality of life, and life expectancy, disproportionately affecting marginalized communities [12]. While SES, typically measured by income, education, and occupation, is a key social determinant of health, SDOH encompasses a broader range of social and environmental conditions beyond individual-level SES [9]. SES and related SDOH significantly impact health outcomes, with studies estimating that they contribute to around half of health outcome variation, while health behaviors account for about a third, and clinical care accounts for about a fifth of variation in health outcomes [13-15].

There is an increasing call to leverage technology, specifically digital health technology (DHT), to support social needs and establish connections to resources and services. DHTs such as mobile health (mHealth) apps, telemedicine and telehealth services, health information technology (HIT) systems, and virtual reality (VR) and augmented reality (AR) technologies play increasing roles in the modern health care system. mHealth apps provide accessible platforms for health monitoring and management [16], telemedicine and telehealth offer remote health care access [17], HIT systems facilitate the integration of social needs data with medical care [18], and VR/AR technologies enhance training and patient engagement [19]. These technologies can systematically identify and tailor resources to individual needs, promoting personal health care management and reducing health disparities among low-SES populations [20]. Ideally, DHT should focus on increasing factors that promote health-related behaviors (eg, education) and decreasing factors discouraging health-related behaviors (eg, resource access barriers).

A compelling case exists for targeting individuals of lower SES as SDOH-DHT primary users. This study focuses on SDOH-DHT among low-SES populations to capture the complex interplay of social, economic, and environmental factors influencing their health and health-related technology use. Although individuals of low SES have traditionally had limited access to digital technology [21], recent studies show a growing use of technology, particularly mobile phones, and mHealth apps among this population [8,22]. While the development of DHT apps has substantially increased, many do not meet the needs of people of low SES [8]. For example, engagement with digital health systems generally requires a basic level of digital health literacy, which is still in its early stages for many individuals of low SES [23]. Digital health literacy is "the ability to seek, find, understand, and appraise health information from electronic sources and apply the knowledge gained to addressing or solving a health problem [24]." Moreover, there is not currently a universal framework for assessing the effectiveness and usability of DHT in this population [8]. Furthermore, many SDOH-DHT apps are designed to address a single domain on SDOH, mostly commonly food and housing, but individuals of low SES must often address multiple SDOH domains that fluctuate throughout their lives [25] (Figure 1).

XSL•FO
RenderX

Figure 1. The 12 domains of social determinants of health (SDOH) used in this study.



Objectives

The manuscript aims to identify the characteristics, needs, and preferences of individuals with low SES using DHT to address HRSN. For digital SDOH management tools to best serve individuals of lower SES, we must use a user-centered design (UCD) approach that incorporates characteristics specific to the target user population, including contextual situations, preferences, and capabilities, to develop system requirements, goals, and design features [26,27]. In technology development, developers may be unfamiliar with the target user group, especially when specific user groups have historically been excluded from the design process [27,28]. UCD aims to shift developers' focus from their own mental models to those of the targeted users [28]. This allows for a more user-empathetic, inclusive, and flexible approach to DHT development so that the functions, features, and aesthetics of SDOH-DHT meet the target user group's needs [28].

Few studies have leveraged UCD to inform DHT addressing social needs among low-SES populations. In response to this gap, the first phase of this study uses UCD methods (eg, interview and focus group) and tools (eg, user profile) to address the following questions:

1. What are the DHT characteristics of SDOH-DHT users of low SES?

2. What are the needs and preferences of SDOH-DHT users of low SES?

The application of insights gained from exploring the characteristics, needs, and preferences of SDOH-DHT-targeted users raises the question of whether the current SDOH-DHT tools meet users "where they are." Thus, we explore the following research question in the second phase of the study:

3. What identified needs and preferences do current SDOH-DHT address?

XSL•FO
RenderX

We then discuss how our findings can be applied to developing SDOH apps, particularly referencing the Florida International University (FIU) Thrive app currently being developed.

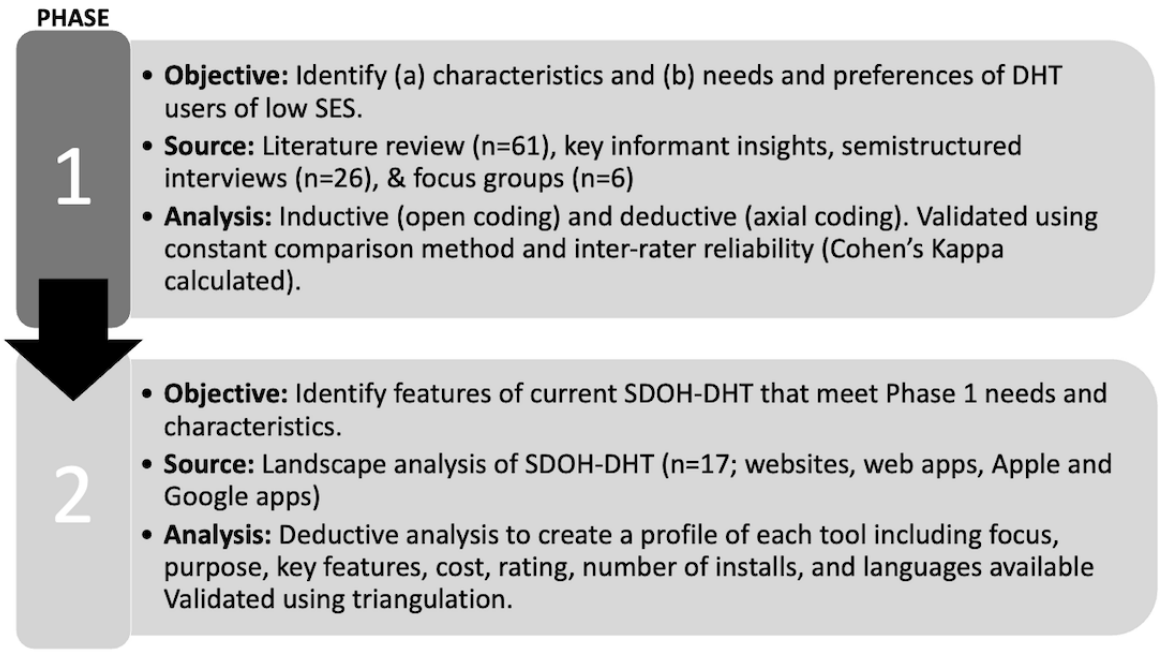
Overall, this study promotes an enhanced understanding of the target user population, desirable SDOH-DHT features and functions, and the status of existing SDOH-DHT in addressing those features and functions, and a discussion, with illustration, of how to collectively address desirable features with some focus on addressing differences between identified and current features and functions. As a result, findings and insights pave a pathway to advance SDOH-DHT design science and DHT product design.

Methods

Study Design

We used a multiphase, mixed method UCD approach following the consolidated criteria for reporting qualitative research (COREQ) [29] to address our research questions (Figure 2 depicts an illustration of the methods). UCD seeks to incorporate all relevant stakeholders in developing system requirements, goals, and design features [26].

Figure 2. Mixed method approach including objective, data source, and analysis for each phase. DHT: digital health technology; SDOH: social determinants of health; SES: socioeconomic status.



Ethical Considerations

As this research involved human participants, ethics approval was obtained from the Institutional Review Board at FIU (IRB-22-0206). After the screening process for the interviews and focus groups, participants received an information sheet detailing the study procedures, risks, and benefits. At the beginning of each interview, the interviewer explained the study procedures, reaffirmed verbal consent for participation, and invited questions from participants. Participants were informed that they were free to withdraw from the study at any time. The data obtained from participants have been anonymized and are not identifiable. Following their participation, interview participants received compensation of US $25 in the form of a gift card.

Phase 1: Understanding Users and Their Needs

Phase 1 occurred in the context of FIU Thrive, a trans-disciplinary design science research (DSR) study to use mHealth to address social risks and resulting disparities in vulnerable populations. DSR is a problem-solving approach that creates novel technologies and associated knowledge [30-32]. In this study, we report on our UCD efforts, which are integral to the more extensive FIU Thrive DSR study.

Phase 1 addressed our first 2 research questions on the characteristics, preferences, and needs among individuals of low SES for SDOH-DHT. Leveraging the expertise of UCD experts on our team and digital health literacy models (eg, Digital Literacies Domains and Relationships [33], ICT Digital Framework [34], and A Review of Digital Literacy Assessment Instruments [35]), we developed user profiles describing our target users based on their characteristics, prior experiences, and anticipated behavior to guide design choices [36]. The user profile was organized based on demographics, health/SDOH-related characteristics, and DHT-related characteristics of interest for this study and the overall FIU Thrive project (Multimedia Appendix 1 [8,35,37-77] shows categorized characteristics).

As a first step to completing the user profile framework, we conducted a focused narrative literature review to obtain generalized insight into the characteristics of interest for our target user population. Using multiple interdisciplinary databases

XSL•FO
RenderX

(eg, Web of Science and PubMed), our search parameters combined low-SES terms with targeted characteristics or characteristic types aligning with our SDOH domains (eg, technology, wellness, health care, and health education; Figure 1) or SDOH terms (eg, social risk and social need). We performed a full-text review of each study to ensure relevance to our profile characteristics of interest.

To add detail to our user profile, we used aggregated data (eg, demographics and some social and health characteristics) from the records of The Green Family Foundation Neighborhood Health Education Learning Program (NeighborhoodHELP), a university-led community program providing interprofessional household-centered care for individuals who face challenges addressing social needs and are open to assistance [78,79]. These program members were targeted as the initial FIU Thrive pilot users. Our study population closely aligned with the literature's definition of low SES: individuals or groups with limited income, education, and occupational prestige, often resulting in reduced access to resources and opportunities. NeighborhoodHELP aims to address social needs and maintain community connections, allowing us to address current and past priority needs with our participants. We aimed to leverage the firsthand experiences, needs, and preferences of individuals enrolled in the program to develop SDOH-DHT that effectively addresses their needs.

Guided by findings from our literature review and aggregate data, we tailored the development of target user interview and outreach worker focus group protocols to areas needing deeper exploration and gaps to be explored.

We constructed a semistructured interview protocol to explore SDOH-DHT user needs and preferences, adding depth and breadth to the user profile. We recognized three key aspects in developing a phenomenological interview guide: (1) contextualization, asking general questions related to participants' general SDOH/social needs experiences and their management of these concerns; (2) apprehending the phenomenon, asking questions focused on participants' experiences with using DHT to manage health and wellness concerns; and (3) clarifying the phenomenon, exploring areas for improvement in DHT for social concern management.

We developed a parallel structure for focus groups with targeted co-users, primarily NeighborhoodHELP outreach staff. The focus group protocols emphasized the attributes and deficits of currently used tools, the information needed to perform their jobs effectively, and their perception of the digital health capabilities of those they served.

The resulting interview and focus group protocols were reviewed and refined by a multidisciplined team involved with the Thrive project. This team included a researcher with substantial qualitative methods expertise (focused on information systems and health informatics) and familiarity with the study constructs, a primary care medical provider engaged in community outreach, a sociotechnical community health researcher experienced in developing technological platforms, and several team members with experience focused on UCD. A consensus was reached on the final protocols.

Participants (n=26, ages 20-71 years, mean age 50 years) were selected through purposive sampling to ensure representation across age decades within the lower SES demographic. The decision to include a diverse age range was guided by the recognition that age can significantly impact one's perspective on socioeconomic challenges and resilience strategies [80,81]. By incorporating voices from younger and older individuals, we aimed to capture a broader spectrum of experiences and insights related to lower SES. NeighborhoodHELP outreach staff were trained in the purposes of the study and recruited program members to participate in the interview process, verifying that age (between 18 and 75 years) and capability (basic technical skills and English comprehension) inclusion criteria were met. For those who had difficulty with technology, the outreach specialist provided support to access Zoom. Participants whose primary language was Spanish or Haitian Creole were included if they could speak enough English to complete the pre-enrollment questions in English. Before each interview, recruiters obtained information from qualified and consented participants via telephone to characterize the nature of participant information technology (IT) use questions. This preliminary information provided insight and helped set the appropriate starting point for interview inquiries regarding technology.

We conducted 26 one-hour semistructured interviews with individuals of low SES via Zoom to obtain the end user perspective. Sample size guidelines for qualitative research suggest that having at least 20 participants is sufficient for data saturation, where incremental learning is minimal [82,83]. Interviews were conducted by 2 trained team members, one serving as lead and the other as scribe. The scribe could interject clarifying questions periodically during the interview to facilitate a comprehensive understanding. The scribe also provided a summary of interview highlights at the end of each interview, inviting participant comments regarding accuracy, omissions, and further explanation (ie, member check) [84].

For the focus groups, we recruited NeighborhoodHELP care providers aged 18 to 75, including clinical, outreach, administrative, faculty, and student divisions. These individuals have experience using a variety of electronic medical records (EMRs) and IT systems, including the university's homegrown outreach health risk profile and outreach portal. Their familiarity with the challenges of collecting, maintaining, and reporting social risks, needs, goals, and patient-reported outcomes provided valuable insight into developing technological solutions. As with the focus groups, 2 research team members conducted the interviews as lead and scribe, performing a collective member check.

We conducted 6 confidential focus groups with a total of 28 participants. Typically, 3 to 5 focus groups per project are sufficient for achieving saturation [85]. We reached saturation by the fifth focus group, but conducted a sixth to ensure data saturation. Participants included physicians, outreach specialists, program coordinators, and social workers with an average of 3.8 years of experience working in the program. Most participants had prior experience working with underserved populations through other programs.

XSL•FO
RenderX

Data Analysis

To enhance the reliability of the results, we performed member checks at the end of each interview and focus group by sharing interview summaries and highlights [84]. During these checks, we presented participants with highlights of our initial data analysis for feedback on accuracy and resonance with their experiences [86]. Deidentified interview and focus group transcripts were uploaded to Dedoose (Sociocultural Research Consultants) qualitative analysis software.

We used a hybrid thematic analysis approach using inductive and deductive coding [87]. Following guidance from King [88], we developed an a priori coding schema consisting of high-level concepts (ie, categories of the user profile and functional categories of the planned FIU Thrive app) for deductive coding (eg, member characteristics, areas of need -risk factors; SDOH data intake survey; app mock-up reactions; promoting app awareness, and implementation of app). These high-level concepts were applied to conceptually meaningful sections of the transcripts (aka slips) [89]. For instance, if participants discussed their perception of the Thrive DHT mock-up screen, we used app mock-up reactions.

Next, we further analyzed the selected slips of transcript texts coded at the deductive level using an inductive, open-coding approach to provide a more detailed and specific understanding of each high-level concept [90] (eg, details of technology use and reactions to screen mock-up). Specifically, we re-examined relevant portions of the text previously coded deductively under each high-level concept to create more discreet, detailed subcodes needed to provide a depth of insight into our first 2 research questions through an open coding approach.

Finally, we used axial coding to organize the discrete codes related to our first 2 research questions (user characteristics and SDOH-DHT needs and preferences) into broader, generalizable categories of the core themes presented in our results section [91].

We aimed to capture the rich, diverse narratives of participants and ideas as a collective set of data without subordinating or elevating the importance of any particular narrative or group of narratives (narratives belonging to a particular demographic group), providing valuable insights that can inform policy and practice aimed at addressing the challenges faced by this subpopulation. Therefore, while we may have had more individuals in a certain age group than others (or other demographic bin), data from each individual transcript was considered towards building an aggregated data set.

To ensure the trustworthiness of thematic analysis (ie, internal validity and reliability), we used Guba's [84] strategies and Weber's [92] recommendations for conducting a rigorous qualitative research study. These include peer debriefing (ie, code review), the code-recode procedure, and researcher triangulation. Table 1 summarizes these procedures.

Table 1. Summary of validity and reliability procedures.

Trustworthiness criteria and procedures	Description
Internal validity	
Peer debriefing (code review)	A health communication expert reviewed high-level codes, their contextualized definitions, alignment between high-level codes and subcodes, the conceptual integrity of the coding schema, and 20% of data.
Reliability	
Reproducibility	A graduate assistant (not involved in data collection/analysis) coded 15% of the excerpts using high-level coding schema. Cohen κ value was 0.99, indicating a high level of agreement.
Stability	The lead researcher recoded 15% of data from the first 3 interviews approximately a month after coding. Cohen κ was 0.99, indicating a high level of coding stability.

Phase 2: Landscape Analysis

Phase 2 involved a landscape analysis that included a structured review of available SDOH management technologies to identify whether their features aligned with the needs and preferences of health consumers, as identified in phase 1. This analysis did not involve direct usability testing, but rather examined publicly available information (eg, product descriptions and user documentation) to infer alignment with the needs and preferences identified by our target population. This approach is consistent with UCD methodologies that emphasize understanding the current design landscape from the mental model of users to inform new solution development [93]. For example, a recent study on diabetes self-management tools used a similar strategy to map available features against user needs as part of a UCD process [94]. Notably, as part of UCD, landscape analyses are critical to understanding the existing market of DHTs and identifying missing features and design elements that should be incorporated into new DHT designs [93-95].

To understand the current environment for SDOH apps and tools, we searched for website tools and web apps, as well as the Google Play and Apple App Stores using the search criteria presented in Table 2.

XSL•FO
RenderX

Table 2. Social determinants of health digital health technology (SDOH-DHT) search strategy.

	Web-based tools	App
Search Terms	One or more of these SDOH term permutations: • Social determinants of health • Social services • Social services referral system • SDOH • Social Risks • Social factors • Social risk factors • Social determinants • Health - related social needs • Health - related social problems • Social needs • Social needs-informed care • Social needs-targeted care **AND** One or more of these digital term permutations: • Digital health intervention • Digital health technology • Client communication system • Community-based information system • Digital health tools • Interactive tools • Tool • Technology interventions • E-health • Website • App • Mobile health/mhealth • Platform • Connect	One or more of these SDOH term permutations: • Social determinants of health • Social services • Social services referral system • SDOH • Social Risks • Social factors • Social risk factors • Social determinants • Health - related social needs • Health - related social problems • Social needs • Social needs-informed care • Social needs-targeted care
Inclusion criteria	• Tool addresses more than one SDOH domain. • Primary target user includes those experiencing a social needs consumer directed • Interactive human-computer interaction (HCI) features included to provide SDOH education and awareness, assess social needs, obtain services, or manage social needs.	• Tool addresses more than one SDOH domain. • Primary target user includes those experiencing a social needs consumer directed • Interactive-(HCI) features included to assess needs, obtain services, or manage needs.
Exclusion criteria	• Tools that limit scope to one SDOH domain (eg, just health care needs). • Tools that target primarily policy makers, community leaders, and health care providers or administrators (eg, tools that provide data sets and strategies for community health improvement)	• Tools that limit scope to one SDOH domain (eg, just health care needs). • Tools that target primarily policy makers, community leaders, and health care providers or administrators (eg, tools that provide data sets and strategies for community health improvement)

We followed several steps to identify the available SDOH technology tools targeting people with social needs. For website tools and resources, we performed an internet search using permutations of "SDOH technology tools" as a search term (Textbox 1 contains a complete listing). Search terms were identified through team discussion and literature search of relevant terms (eg, SDOH lexicon work by Alderwick and Gottlieb [2]). We included websites that included an interactive SDOH feature or specific training or guidance on implementing or evaluating SDOH-based programs. Websites that only provided general education about SDOH without targeted tools were excluded.

Next, we searched for smartphone tools and resources on the Google Play Store and the App Store using permutations of SDOH search terms (Table 2 contains a complete listing). We also reviewed apps suggested on the pages of other apps. Additionally, if we found an app on one platform (eg, Google Play Store), we searched for the same app on the alternative store (eg, App Store). We included any app that assisted professionals and community organizations in providing education, raising awareness, or alleviating more than one domain for people facing social challenges.

XSL•FO
RenderX

Textbox 1. Summary of social determinants of health digital health technology (SDOH-DHT) included in landscape analysis.

Mobile app (Apple):

- Providers: EBT, debit, & more (Propel, Inc)
- What I Need (WIN) (OurCommunityLA)

Mobile app (Apple and Android):

- myCOMPASS PA (Pennsylvania State Government)

Website and mobile app (Apple and Android):

- FindHelp.org (FindHelo)
- HelpSteps (Boston Children's Hospital)
- Your Texas Benefits (TX Health and Human Services Commission)
- ebtEDGE (Fidelity National Information Services, Inc)

Website:

- ACCESS Florida (Florida Department of Children and Families)
- Washington Connection - Your link to services (Washington Connection)
- Washington Healthplanfinder (Washington Healthplanfinder)
- 211.org (United Way Worldwide)
- Benefit Finder (US Government)
- HHS.gov Social Services (US Department of Health and Human Services)

Web app:

- ACT.md (Activate Care)
- NowPow (As of July 2024, the NowPow website states that the platform is no longer supported and available.) (Unite Us)
- Pieces (Pieces Technologies)
- WellSky Social Care Coordination

Three rounds of the search process were performed during the study to account for the inclusion of recently released tools in the final results. Two authors (CL and PD) conducted the final searches and data extraction during October 2024. It is of note that many tools were excluded as they only covered one domain. Indeed, we discovered that most existing apps and websites designed for at-risk individuals connect users to social services that address only a narrow range of social issues, leading those with more than one social need to choose to use technology tools to rely on multiple DHTs to meet their needs. The search process resulted in analyzing 17 DHTs (Textbox 1). Further details of these tools can be found in Multimedia Appendix 2.

We assessed each app based on the inclusion of clearly identifiable features from phase 1. In some cases, we were not able to gain full accessibility to some tools reviewed for inclusion due to privacy features (required a health provider registration and identification). In these instances, we consulted reference material on the tool features to help assure that we were not missing key features in our analysis.

Results

Overview

Our results provide insights into the characteristics of SDOH-DHT users and how currently available SDOH-DHT addresses their needs and preferences. First, we highlight insights regarding target user SDOH-DHT characteristics in detail in our user profile (Multimedia Appendix 1 [8,35,37-77]). Next, we review our findings regarding the needs and preferences of SDOH-DHT users and the associated landscape analysis results. This analysis is organized according to 5 themes: UCD; efficient solution-based assessment of SDOH risk/need; trust, privacy, and security; e-caring and support; and user feedback and education. To support the general themes and findings from the interviews and focus groups, we provide a detailed evidence trace table in Multimedia Appendix 3, moving from theme to associated open codes to representative quotes.

Target User Characteristics

We defined our targeted profile based on the literature, focusing on individuals who experience a lack of resources such as food, education, income, and transportation, resulting in social risks and challenges. Phase 1 results revealed the DHT characteristics

XSL•FO
RenderX

and needs of DHT users with low SES, particularly early adopters, allowing us to complete the comprehensive user profile in Multimedia Appendix 1 [8,35,37-77]. We augment this artifact with the highlights and extended descriptions provided below.

Demographics

Our interview participants include individuals enrolled in NeighborhoodHELP to help address unmet social needs and risks. Table 3 contains the demographic information of our participants collected during screening.

Table 3. Interview participant demographic information.

Demographics	Values
Age (years), mean (range)	50 (20-71)
Household income, mean	US $18,519
Number of household members, mean	4
Sex, n (%)	
Male	8 (31)
Female	18 (69)
Ethnicity, n (%)	
Hispanic or Latino	5 (19)
Non-Hispanic or Latino	21 (81)
Race, n (%)	
African American	18 (69)
White	5 (19)
Mixed race	2 (8)
Asian	1 (4)

The age range of participants for our study was between 20 and 71 years, including representatives from each age decade to ensure a comprehensive understanding of how age influences the experiences of individuals within the low-SES population. The mean household income was US $18,519, and the average household size was 4 members. While English-speaking was required for interview participation, linguistic diversity is present within the overall NeighborhoodHELP population. Among the NeighborhoodHELP population, 40% of members primarily speak Spanish and 14% speak Haitian Creole. Educational attainment varied, with 77% having completed at least high school.

The Social Needs Landscape: Insights From Participants and Professionals

Participants reported a wide array of social needs, reflecting the complex challenges low-SES populations face. These needs spanned financial instability, food insecurity, housing concerns, transportation difficulties, unemployment, educational barriers, limited health care access, and inadequate social support. Interviewees provided examples of how issues in the social domains, such as education, finances, and food insecurity, negatively impacted their physical and emotional health.

Participants indicated that they attempted to address their social needs through various resources, such as governmental programs, churches, and local nonprofits. When trying to address these needs, participants encountered several obstacles, including fragmented services, burdensome application processes, financial constraints, and competing life demands that made it challenging to navigate support systems effectively. Participating professionals indicated that successful care plans addressing social needs must include a triad of individual patient priorities, the feasibility of meeting specific needs, and evidence-based guidelines. They also noted that identifying the target user's strengths alongside risk areas is beneficial as these strengths can serve as leverage points and motivators.

DHT Use

Studies show that individuals with low SES often struggle with IT literacy, making it difficult to adapt to new tools. However, our research reveals that during the COVID-19 pandemic, many reported improved digital health use. Focus group participants highlighted various programs that promoted the use of technology, particularly for completing applications and checking statuses. Despite low SES, participants demonstrated significant engagement with digital technologies. Most (81%, n=21) reported daily internet use, and 81% (n=21) regularly used smartphone apps. However, self-assessed digital skills varied widely among participants, indicating a range of comfort levels with technology. Still, about half of the study participants used DHT for various health-related social activities, including scheduling health care appointments, completing social need applications, reading health care reviews, monitoring symptoms, accessing EMRs, retrieving medical records, completing health surveys, and using health-related apps (Multimedia Appendix 4).

User Needs, Preferences, and Current Technology Solutions

Our user profile work and detailed qualitative analysis resulted in 5 key themes related to user needs and preferences: UCD,

XSL•FO
RenderX

solution-based assessment of social needs, trust, privacy and security, e-caring and support, and education and user feedback. The following sections explore these themes identified through qualitative analysis of target user needs and preference data and share the presence of features associated with these themes in current DHT offerings. Table 4 highlights key comparisons from our landscape analysis, showing how current DHT aligns with or diverges from the identified needs and preferences. This comparative narrative addresses our research question on how the current landscape meets—or fails to meet—these needs. Multimedia Appendix 2 includes detailed descriptions of each SDOH-DHT we explored for a cohesive depiction of the landscape.

Table 4. Overview of themes from literature review, interviews, focus groups, and landscape analysis.

User needs and preferences	Current DHT[a] identified
User-centered design	
Multilingual	Few offer language translation beyond Spanish (eg, Creole)
Use of images, icons, and videos over text to help address literacy and add to aesthetics	Limited use of images and videos; most DHTs are text-based with few illustrations
Customized needs prioritization, goal setting, messaging (eg, reminders and progress tracking with incentivization/encouraging messages)	No DHTs found that allow easy assessment and prioritization of needs and goals; users often cannot choose notification timing
Self-paced, streamlined survey completion process that meets users where they are and their schedules	DHTs break questions into sections but do not allow users to save progress; few offer prefilled responses to streamline processes
Efficient, solution-based assessment of social risk/need	
One-stop access to resources addressing multiple social domains; cross-domain referencing	No DHTs surveyed include guidance to assess and navigate all health-related social domains with direct access by users; users need multiple DHTs, specific ideas about what services they need, or a case manager to assess needs and address barriers
Multiple referrals filtered by goals, geography, open hours, eligibility, etc.	Resources are searchable by keyword or domain area; some offer zip code filtering
Track progress over time with outreach workers/administrators	Program-focused DHTs track eligibility, application progress, and benefits status but are limited to those receiving formal assistance
Easy sharing of data with health care providers/social service professionals	Patient SDOH status integration into other systems (eg, EMRs) is generally unavailable for patient use
E-caring and support	
Use of empathetic and cooperative tone and language; motivational messaging	Few DHTs show evidence of an empathetic tone or motivational messaging
Option for continued contact with outreach workers; surrogate users (household, friends, outreach workers)	Several DHTs are designed for health care providers, not the general public; some offer training/resources for addressing social needs, while others only provide search functionality for services
User education and feedback	
DHT skill training (eg, swipe function) and health information searching; education on taking action on social needs	Few DHTs offer education on taking action to address social needs or risks; existing training is often limited to basic use of DHT or provider training
Review existing feedback (reviews and ratings) and add personal feedback	Most SDOH-DHT apps have user reviews in app stores; few include feedback from users and outreach workers on past service experiences
Trust, privacy, and security	
Trustworthy DHT platform and sponsor; evidence of safeguards against hackers, scams, and spam	Few current SDOH-DHTs prominently display privacy and data security policies or trust language
Information is optional, particularly for sensitive topics, unless required for app processes	Many DHTs do not allow users to skip sensitive questions

[a]DHT: digital health technology.

User-Centered Design

The UCD process tailors app design and development to the needs and capabilities of the target user population. Our data highlighted 3 key areas of UCD needs and preferences: inclusiveness, customization, and human-computer interaction efficiency. Although some existing apps addressed 1 or 2 of these areas, most apps in our landscape analysis lacked comprehensive attention to all aspects of UCD.

XSL•FO
RenderX

Inclusiveness

Participants strongly emphasized the need for inclusive design in SDOH-DHT. They expressed a desire for multilingual options, particularly highlighting the importance of Spanish and Haitian Creole translations alongside English. This linguistic inclusivity was seen as crucial for ensuring that diverse user groups could effectively engage with the technology. Our participants preferred apps in English, Spanish, and Creole. While several DHTs offer Spanish translations (eg, Your Texas Benefit, ebtEDGE, myCOMPASS PA), very few include Creole (eg, ACCESS Florida).

Additionally, participants stressed the importance of using simple, accessible language to accommodate varying literacy levels within the target population. Participants also noted that simpler terminology could reduce on-screen text and suggested that icons, graphics, and videos could further enhance usability. Some participants used sites like WebMD for their illustrative content, but few SDOH-DHTs include such features. For instance, ACCESS Florida (Florida Department of Children and Families) offers limited visuals, and while ACT.md (Activate Care) and WellSky Social Care Coordination (Wellsky) provide videos, these are general overviews rather than guides for accessing specific services. Most DHTs in our analysis were text-heavy, offering questionnaires or service lists with minimal visual aids (eg, Benefit Finder).

Customization

Customization emerged as another critical aspect of UCD. Participants wanted the ability to prioritize their needs within the app, set personalized goals, and receive customized messaging. Current SDOH-DHTs for individual use often fail to account for personal factors and prioritized needs.

Participants envisioned features such as reminders and progress tracking accompanied by encouraging messages, which they felt would enhance their engagement and motivation to address their social needs. While many existing SDOH-DHTs, especially mobile apps, offer notifications, they often lack options for users to choose when notifications should appear. For example, the "Providers: EBT, debit, & more" app (Propel, Inc) allows users to receive notifications for benefits, offers, tax filing, and promotions, but does not let them customize the frequency of these alerts.

Efficiency

Efficiency in interactions was also a key concern. Participants expressed a preference for a self-paced, streamlined survey completion process. They indicated that ideally, survey sessions should be limited to 10-30 questions and last no more than 20 minutes. This preference reflects the time constraints and competing priorities many individuals of low SES face. However, some SDOH-DHTs, like the Benefit Finder (US government), require up to 30 minutes to complete a questionnaire, which does not align with these preferences. While about one-third of current SDOH-DHTs avoid repetitive and lengthy data collection, some still take up to 30 minutes to complete.

Furthermore, participants stressed the importance of completing surveys at their own pace, starting with the most critical sections and saving their progress to continue later. The ability to save progress helps avoid repeating processes, which participants found essential. However, most web-based DHTs do not allow users to save their progress. For example, ACCESS Florida (Florida Department of Children and Families), Benefit Finder (US Government), and Washington Connection require users to fill out entire questionnaires without the option to save and return later. While some DHTs, like ACCESS Florida (Florida Department of Children and Families), break down questions by sections (eg, household, income, expenses), they still do not allow users to save partially completed questionnaires.

Efficient Solution-Based Assessment of Health-Related Social Risks, Needs, and Goals

Users articulated a strong desire for comprehensive tools integrated with practical solutions. They envisioned a "one-stop" platform that could address multiple social domains simultaneously, recognizing the interconnected nature of many social needs. For example, one participant faced food insecurity but could not access a food pantry due to a lack of transportation. Others were ineligible for services because of their immigration status. Participants expressed frustration with fragmented services and saw great value in cross-domain referencing within a single platform. Tools with a broader scope of social needs typically cover housing, food, employment, health care, transportation, and childcare services. For example, FindHelp (FindHelp) spans domains including food, housing, transportation, health, money, education, work, and legal services.

Social needs navigation professionals also emphasized the importance of tracking progress in addressing specific social needs. However, current DHTs typically do not integrate social needs assessments with service referrals or track progress over time. Some platforms, like Pieces (Pieces Technologies), use artificial intelligence and advanced analytics to automate screening and address SDOH issues through predictive modeling of at-risk patients in EMR systems. Unfortunately, these advanced algorithms and processes are inaccessible to the general low-SES population.

Another key preference was the ability to receive filtered referrals based on specific criteria. Users wanted to be able to narrow down service options based on their personal goals, geographic location, service hours, and eligibility criteria. This level of customization was seen as essential for making the platform truly useful in addressing individual needs. Most SDOH-DHTs provide lists of service referrals for specific needs and guide users to access a single service, but do not automatically filter resources based on user eligibility and practical considerations like access. More comprehensive tools, such as HelpSteps (Boston Children's Hospital) in Massachusetts and FindHelp (FindHelp) nationally, serve as central resources for individuals but are primarily designed for resource searching rather than automatically filtering resources based on eligibility. These tools allow users to search for help by directly finding and connecting with services like food,

XSL•FO
RenderX

housing, health care, and other types of support available nearby, usually by entering a zip code.

Participants also emphasized the importance of easy data sharing with health care providers and social service professionals. They recognized the potential for improved care coordination if their social needs data could be seamlessly integrated with their medical care. Administrators and outreach workers suggested that integrating this functionality with EMR systems could be an effective solution. While some DHT platforms are already integrated with EMRs, these are typically designed to assist providers rather than directly support low-SES patients. Certain platforms, such as NowPow (Unite Us) and WellSky Social Care Coordination (WellSky), focus on building networks among community care providers.

E-Caring and Support

The concept of "e-caring" also emerged as a significant theme in our discussions. Participants valued an empathetic and cooperative tone in all interactions within the app. They expressed a desire for motivational messaging, particularly when completing tasks or reaching milestones, as a way to maintain engagement and boost morale. However, our landscape analysis revealed that few existing DHTs offer motivational messaging when users complete questionnaires or select different services.

Many participants also highlighted the importance of having options for continued contact with outreach workers or support from surrogate users, such as household members or friends. This human element was seen as a crucial complement to the digital platform, providing personalized support and guidance when needed. During our landscape analysis, we identified ACT.md (Activate Care), an app that integrates with EMRs. The ACT.md platform enables health care providers to manage patient needs and coordinate care with community partners by sharing tasks, messages, and data across organizations. Care teams can view patients' unmet social needs and collaborate with community organizations to address those needs. However, these platforms are only accessible to providers and patients within health care systems that have purchased them. Additionally, they are primarily designed for health care providers rather than individual users.

User Education and Feedback

Participants showed interest in educational content within the SDOH-DHT, but with a specific focus on actionable information. Rather than basic facts about SDOH, they wanted practical guidance on how to address their specific social needs. Practical guidance references included a desire for DHT skill training (such as instruction on navigation functions) and advice on how to search for health information effectively. Existing digital health technologies tend to focus on offering general information across various social domains (eg, Your Texas Benefits, HHS.gov Social Services) or screening users for eligibility for certain benefits (eg, Benefit Finder, Washington Connection - Your Link to Services) without providing step-by-step guidance on how to address the identified needs and risks effectively.

The ability to engage with feedback was another important aspect for users. Participants wanted to be able to review existing feedback, including ratings and reviews of services, as well as contribute their feedback based on their experiences. This 2-way flow of information was seen as valuable for making informed decisions about services and contributing to community knowledge. While most SDOH-DHT apps feature user reviews on the technology itself in the app store, only 2 apps included feedback from both users and outreach workers on past service experiences. For instance, ACT.md, a platform that enables health care providers to manage patient needs and coordinate care with community partners, includes a page dedicated to client experiences. This page details how the platform has assisted its clients and features quotes from clients about their experiences with the programs.

Trust, Privacy, and Security

Trust, privacy, and security emerged as foundational concerns for potential SDOH-DHT users. Participants emphasized the need for clear evidence of safeguards against possible threats such as hackers, scams, and spam. They were particularly sensitive to the possibility of personal information misuse leading to unwanted solicitations. Our landscape analysis showed that only a few current SDOH-DHTs prominently feature clear privacy and data security policies. One exception is the Pieces DHT, which offers users a dedicated page outlining its data privacy policy and contact information for any questions or concerns.

Regarding data sharing, participants strongly preferred granular control over their information. They wanted the option to selectively disclose sensitive information, with mandatory fields limited to only what was absolutely necessary for specific processes. Existing DHTs, especially those using questionnaires, often do not allow users to skip these sensitive questions. For example, the Benefit Finder (US government) requires users to disclose citizenship status, annual income, and employment status before allowing them to proceed with selecting services. This mandatory disclosure can deter some users from using the platform. Transparent privacy policies were seen as essential for building trust in the platform.

Discussion

Principal Findings

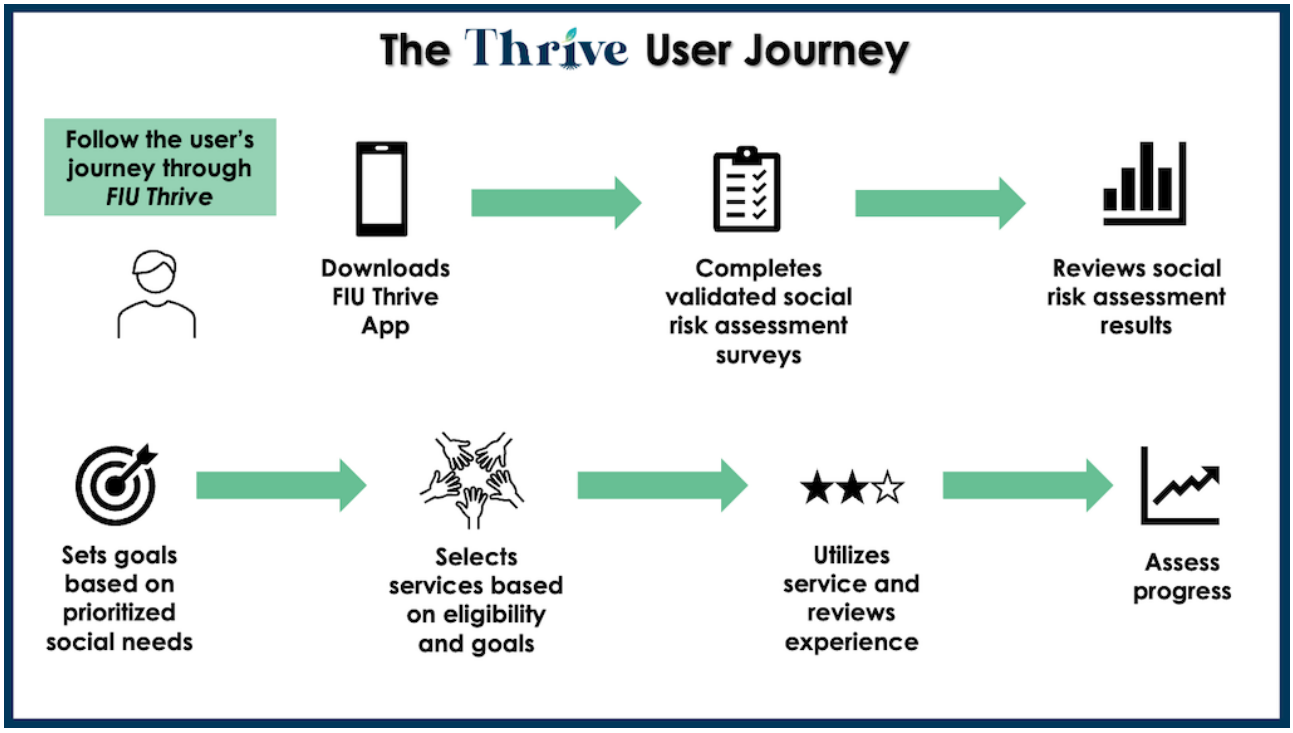
Our study reveals the characteristics, needs, and preferences of individuals of low SES using SDOH-DHT and how current DHTs align with or diverge from these needs and preferences. Our results, identified across the key themes such as UCD, solution-based features, e-caring support, education and feedback, and trust and security, have important implications for the effectiveness, adoption, and equity of digital health interventions targeting SDOH. The pervasive message users have communicated is the need for continuity across all themes.

The insights gained from this study are crucial for informing the iterative design process of the FIU Thrive mHealth tool. FIU Thrive aims to maintain continuity in addressing social needs by creating the user journey. This mHealth tool uniquely emphasizes identifying and addressing users' health risks, needs,

XSL•FO
RenderX

priorities, and goals. It integrates these elements into actionable insights, creating individualized care plans that enhance self-knowledge and empower individuals. Users can connect with recommended services based on their specified social risks, goals, and geographic area. FIU Thrive aims to address risks and needs across multiple SDOH domains. Overall, the FIU Thrive user will embark on a journey to understand their social needs and risks better and find evidence-based and personalized solutions to address them. The FIU Thrive User Journey is presented in Figure 3.

Figure 3. Thrive user journey. FIU: Florida International University.



FIU Thrive also seeks to address the needs and preferences of users of low SES within the different themes. Based on this UCD study, features designed into FIU Thrive to address the needs and preferences of our target users are presented in Textbox 2 and further described below.

The gaps in current SDOH-DHT may perpetuate barriers to access and engagement for individuals of low SES, who often face language, literacy, and cultural challenges [8,23,24,26-28]. Features for improving aspects of UCD to tailor app design and development to the needs and capabilities of the intended user population include multilingual support, use of graphics for visual guidance, customization, and self-paced and streamlined data collection. By interviewing participants aged 20-71 years, we captured a broad spectrum of experiences and insights related to lower SES. Decreased DHT adoption can be attributed to biopsychosocial factors that can be generalized to many age groups, including cognitive and physical impairments, lack of familiarity with technology, or lack of device access [96,97]. Based on current literature and our results, strategies to accommodate these needs include technology features such as larger font size and instructional videos, voice-to-text functions, and increased support from family, caregivers, or health care providers [96,97].

The fragmented nature of current SDOH-DHT resources, coupled with limited integration of referral and tracking features, may impose additional burdens on individuals of low SES who often navigate complex and disconnected social service systems [5,98]. The absence of comprehensive, coordinated support can lead to unmet social needs, delays in accessing care, and poorer health outcomes [3,27,99]. Improving solution-based assessment of HRSN by offering integrated, one-stop access to a wide range of social services, personalized referrals, eligibility screening for social services, and progress monitoring (for managing social risks) could help streamline connecting users to appropriate resources and support [6,20,100].

The lack of empathetic support and action-oriented education in existing tools may also undermine user engagement and motivation, particularly among individuals facing multiple stressors, competing priorities, and limited social support [27-29,99]. While using DHT platforms, our users emphasized the importance of conveying sensitivity, empathy, and motivation, as positive in-person interactions with health care providers do. Participants reported that features conveying e-caring included using empathetic and cooperative language, motivational messaging, and web-based chat or support group features.

XSL•FO
RenderX

Textbox 2. Florida International University (FIU) Thrive app features.

The FIU Thrive mHealth app is designed to address the specific needs and preferences of low socioeconomic status users identified in this study. The app is designed to incorporate the following features:

1. Multilingual support (user-centered design [UCD])

- In-app language selection option (eg, English, Spanish, and Haitian Creole)
- Automatic language detection based on the user's device settings

2. Simple and intuitive user interface (UCD)

- Clean, uncluttered design with minimal text
- Use of icons and visual cues to guide navigation

3. Customized user experience (UCD)

- Customizable user profile
- Ability to set personal goals and priorities
- Tailored content and recommendations based on user's social risks/needs assessment results

4. Solution-based assessment of social risk/need

- Brief, user-friendly questionnaires for each social domain
- Progress tracking and visualization using color-coded SDOH wheel
- Option to complete assessments in multiple sessions

5. Integrated referral system

- Personalized referrals to local resources and services based on social needs assessment results
- Detailed information about each referral, including eligibility criteria, contact information, and user ratings
- Ability to save and track referrals within the app

6. E-caring support features

- In-app messaging with outreach workers and peer support groups
- Encouraging notifications and reminders to support goal progress
- Web-based coaching and educational content tailored to user's needs

7. User education and feedback mechanisms

- User feedback on referral services
- In-app feedback form to report issues or suggest improvements
- Regular prompts to share progress and successes with outreach workers

8. Robust privacy and security measures

- Secure user authentication and data encryption
- Granular privacy settings to control data sharing
- Clear, concise privacy policy and terms of service

Regarding references of e-caring to their current technology, commonly used apps manifesting some aspect of e-caring that impacted our participants' health and wellness were spiritual and fitness apps. Participants reported continual use of spiritual and fitness apps because of the encouragement and motivation they received through notifications and messages, ensuring the user that they are moving in the right direction and imparting a sense of accomplishment. Participants reported that these messages helped them feel cared for and as if they received individualized attention to their actions, reducing the monotony of the DHT interaction. Ultimately, our participants appreciated that these messages allowed them to obtain reassurance and affirmation without physical interactions from a provider. Some participants commented that these elements of e-caring were an essential component of the long-term use of DHT. Our study supports that designing for e-caring and support (eg, "individualized attention"), which combines technology's efficiency with human interaction's empathy, may be particularly valuable in building trusting relationships and promoting sustained engagement [101].

XSL•FO
RenderX

Incorporating increased user education and feedback enhances engagement and impact of SDOH-DHT for this population [8,102,103]. For example, participants indicated that they desired the apps to enhance their knowledge and skills and provide useful feedback in the following areas: IT skill training, education on taking action to address HRSN, and a feedback/rating mechanism for social services consulted.

Privacy and security concerns emerge as a critical barrier to trust and adoption of SDOH-DHT among users of low SES who may have heightened sensitivities around the protection of their personal information, given experiences of discrimination, stigma, or exploitation [104,105]. The absence of clear privacy policies, granular data-sharing controls, and secure communication channels in many current tools may exacerbate these concerns and deter engagement [106]. Developing SDOH-DHT with robust privacy safeguards, transparent data practices, and user control over information sharing is essential to fostering trust, buy-in, and long-term use among low-SES populations [107,108].

Our findings highlight the importance of addressing HRSN among low-SES populations through various DHTs. mHealth apps provide accessible health monitoring and management platforms, crucial for low-SES populations facing barriers to traditional health care [16]. Telemedicine and telehealth services offer remote health care access, addressing transportation barriers and improving care for underserved communities [17]. HIT systems facilitate the integration of social needs data with medical care, enabling providers to better address the social determinants affecting patients' health [18]. EHR interoperability is crucial in achieving seamless communication and coordination of care. While VR and AR technologies are increasingly used in surgery [109] and occupational health to enhance training and patient engagement [19], their direct application to addressing HRSN among low-SES populations may be less clear. VR has been used in education related to SDOH to help providers recognize situations and increase empathy in management [110]. However, incorporating e-caring support features and personalized feedback mechanisms can improve the user experience and effectiveness of VR and AR applications for low SES populations [111].

In low and middle-income countries, barriers to DHT implementation include infrastructure limitations (unreliable electricity, limited internet), insufficient resources, and poor digital literacy among patients and providers [112,113]. Key facilitators include continuous on-the-job training, strong stakeholder commitment, and technologies designed with user-friendly interfaces [113]. Simpler DHTs such as telemedicine, mHealth, and EMR (including patient portal access or printouts) may be best initial DHT options in low and middle-income countries as they are easiest to implement due to lower infrastructure requirements, and easier integration into workflows [112]. Current literature has shown that efforts to bring DHT to such countries have been successful in improving the standard of health care service delivery, data management for decision-making, and boosting attendance at health facilities and use of services [112,113]. Therefore, we see promise in adapting our results in many settings, including low and middle-income countries, with adaptation as needed.

Overall, future work should focus on developing and evaluating SDOH-DHT that incorporates the user needs and preferences identified in this study. This could involve the creation of new tools (eg, FIU Thrive) or adapting existing platforms to better align with the requirements of those with low SES. Importantly, these development efforts require a solid commitment to UCD principles and the meaningful engagement of diverse individuals as co-designers and cocreators throughout the process to facilitate usability and acceptance. By involving target users as active partners in the conception, design, and iteration of SDOH-DHT, we can ensure that these tools are aligned with their unique needs, values, and lived experiences [37]. This participatory approach can help to identify and address potential barriers to adoption and use, enhance the cultural responsiveness and contextual relevance of the tools, and ultimately, improve their acceptance, usability, and impact among populations with low SES [114-116].

Furthermore, the development and implementation of SDOH-DHT must be situated within a broader ecosystem of supportive policies, programs, and partnerships that address the root causes of health disparities [3]. Digital tools alone cannot solve the complex, systemic issues that contribute to health disparities. Still, they can play a valuable role in connecting individuals to resources and facilitating communication and coordination among service providers [117].

Limitations

The purposive sampling method, while effective for ensuring representation, may limit the generalizability of our findings to the broader population of lower SES individuals. Additionally, the qualitative nature of the study means that the findings are inherently subjective and may not reflect the experiences of all individuals within this demographic. Limitations of our study also include the focus on participants speaking at a threshold level of English (speaking enough English as a primary or secondary language to complete the pre-enrollment questions in English) and the potential for incomplete identification of relevant SDOH-DHT given sample limitations. While we aimed to include a diverse sample of target users, the lack of non-English speaking participants may limit the generalizability of our findings to populations with limited English proficiency. Future research should use a cross-sectional design study with a larger and more representative sample of North American populations, including First Nations individuals and African immigrants, to capture the unique needs and preferences of a diverse patient population.

Additionally, despite our systematic search strategy, it is possible that some relevant SDOH-DHT were not identified or included in our landscape analysis. Moreover, due to the proprietary nature of some of the systems, we were unable to access some parts of some systems and include functionality of these parts (beyond what was described and referenced in information materials) in our landscape analysis. As the field of digital health is rapidly evolving, new tools and platforms may have emerged since our data collection in May 2024. Ongoing research should continue to monitor and evaluate the SDOH-DHT landscape to ensure that the most current and relevant tools are considered.

XSL•FO
RenderX

Conclusions

This study identifies characteristics, needs, and preferences related to SDOH-DHT among individuals of low SES compared with the current DHT landscape. Five themes emerged from our analysis: UCD, solution-based assessment of social risks and needs, e-caring and support, education and feedback, and trust, privacy, and security. This study also identified gaps between the needs and preferences of individuals of low SES for comprehensive social needs support and the features of current SDOH-DHT. Discrepancies between target user needs and current DHT features represent missed opportunities in developing user-centered tools to assist individuals with self-management of social needs. These findings underscore the need to design SDOH digital health interventions that are inclusive, empowering, and responsive to the unique challenges faced by these populations.

By creating digital health interventions that strive to truly meet the needs and preferences of users with low SES, we can harness the power of technology to bridge health disparities and promote public health. However, the development of user-centered SDOH-DHT is just one piece of the puzzle. We can only create the enabling environment needed for SDOH-DHT to achieve its full potential in advancing health and reducing health disparities by working together across disciplines, sectors, and communities.

Acknowledgments

The authors acknowledge Nancy Napolitano, Jessica Vallejo, and Carolina Cardenas for coordinating recruitment for this study. They also acknowledge Troy Stefano, Andres Rodriguez, Staci Morris, Rodolfo Padron, Nicholas Chung, and the rest of the College of Medicine IT team for their work in developing the FIU Thrive app platform. They also acknowledge Troy Stefano's leadership in conceiving and funding FIU Thrive. They also acknowledge Gabriela Torres' contribution to data analysis for the UCD project. Last, the authors recognize the Green Family Foundation Neighborhood Health Education Learning Program (NeighborhoodHELP) outreach team and household members without whom this program and study would not have been possible. Funded in part by CDC-RFA-OT21-2103, the National Initiative to Address COVID-19 Health Disparities Among Populations at High-Risk and Underserved, Including Racial and Ethnic Minority Populations and Rural Communities, through Florida Department of Health in Miami-Dade contracts DEX87 and DEY50.

Authors' Contributions

KLH contributed to the writing of the original draft and its review and editing, as well as to investigation, data curation, and analysis. CL contributed to conceptualization, administration, and supervision, and participated in the writing of the original draft and its review and editing, investigation, data curation, and analysis. PD contributed to the writing of the original draft and its review and editing, investigation, data curation, and analysis. MPDC contributed to the review and editing of the manuscript, investigation, data curation, and analysis. DRB contributed to conceptualization, administration, and supervision, and participated in the writing of the original draft and its review and editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1
User profile low-socioeconomic status individual as potential social determinants of health mobile health user.
[DOCX File , 67 KB - humanfactors_v12i1e69545_app1.docx]

Multimedia Appendix 2
Social determinants of health digital health technology landscape analysis.
[DOCX File , 21 KB - humanfactors_v12i1e69545_app2.docx]

Multimedia Appendix 3
Evidence trace table.
[DOCX File , 29 KB - humanfactors_v12i1e69545_app3.docx]

Multimedia Appendix 4
Digital health technology user characteristics from screening survey.
[DOCX File , 16 KB - humanfactors_v12i1e69545_app4.docx]

References

XSL•FO
RenderX

1. World Health Organization. Closing the gap in a generation: health equity through action on the social determinants of health. In: Commission on Social Determinants of Health Final Report. Geneva, Switzerland: World Health Organization; 2008.
2. Alderwick H, Gottlieb LM. Meanings and misunderstandings: a social determinants of health lexicon for health care systems. Milbank Q 2019;97(2):407-419 [FREE Full text] [doi: 10.1111/1468-0009.12390] [Medline: 31069864]
3. Braveman P, Gottlieb L. The social determinants of health: it's time to consider the causes of the causes. Public Health Rep 2014;129 Suppl 2(Suppl 2):19-31 [FREE Full text] [doi: 10.1177/00333549141291S206] [Medline: 24385661]
4. Marmot MG, Smith GD, Stansfeld S, Patel C, North F, Head J, et al. Health inequalities among British civil servants: the Whitehall II study. Lancet 1991;337(8754):1387-1393. [doi: 10.1016/0140-6736(91)93068-k] [Medline: 1674771]
5. Fiscella K, Williams DR. Health disparities based on socioeconomic inequities: implications for urban health care. Acad Med 2004;79(12):1139-1147. [doi: 10.1097/00001888-200412000-00004] [Medline: 15563647]
6. Pampel FC, Krueger PM, Denney JT. Socioeconomic disparities in health behaviors. Annu Rev Sociol 2010;36:349-370 [FREE Full text] [doi: 10.1146/annurev.soc.012809.102529] [Medline: 21909182]
7. Etindele Sosso FA, Holmes SD, Weinstein AA. Influence of socioeconomic status on objective sleep measurement: a systematic review and meta-analysis of actigraphy studies. Sleep Health 2021;7(4):417-428. [doi: 10.1016/j.sleh.2021.05.005] [Medline: 34266774]
8. Sharma S, Gergen Barnett K, Maypole J, Grochow Mishuris R. Evaluation of mHealth apps for diverse, low-income patient populations: framework development and application study. JMIR Form Res 2022;6(2):e29922 [FREE Full text] [doi: 10.2196/29922] [Medline: 35147502]
9. Western MJ, Armstrong MEG, Islam I, Morgan K, Jones UF, Kelson MJ. The effectiveness of digital interventions for increasing physical activity in individuals of low socioeconomic status: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2021;18(1):148 [FREE Full text] [doi: 10.1186/s12966-021-01218-4] [Medline: 34753490]
10. Pinxten W, Lievens J. The importance of economic, social and cultural capital in understanding health inequalities: using a Bourdieu-based approach in research on physical and mental health perceptions. Sociol Health Illn 2014;36(7):1095-1110. [doi: 10.1111/1467-9566.12154] [Medline: 25040507]
11. Schneider-Kamp A. Health capital: toward a conceptual framework for understanding the construction of individual health. Soc Theory Health 2021;19(3):205-219 [FREE Full text] [doi: 10.1057/s41285-020-00145-x] [Medline: 32837331]
12. White-Williams C, Rossi LP, Bittner VA, Driscoll A, Durant RW, Granger BB, American Heart Association Council on Cardiovascular Stroke Nursing, Council on Clinical Cardiology, Council on Epidemiology Prevention. Addressing social determinants of health in the care of patients with heart failure: a scientific statement from the American Heart Association. Circulation 2020;141(22):e841-e863 [FREE Full text] [doi: 10.1161/CIR.0000000000000767] [Medline: 32349541]
13. Son H, Zhang D, Shen Y, Jaysing A, Zhang J, Chen Z, et al. Social determinants of cardiovascular health: a longitudinal analysis of cardiovascular disease mortality in US counties from 2009 to 2018. J Am Heart Assoc 2023;12(2):e026940 [FREE Full text] [doi: 10.1161/JAHA.122.026940] [Medline: 36625296]
14. Hood CM, Gennuso KP, Swain GR, Catlin BB. County health rankings: relationships between determinant factors and health outcomes. Am J Prev Med 2016;50(2):129-135. [doi: 10.1016/j.amepre.2015.08.024] [Medline: 26526164]
15. Park H, Roubal AM, Jovaag A, Gennuso KP, Catlin BB. Relative contributions of a set of health factors to selected health outcomes. Am J Prev Med 2015;49(6):961-969. [doi: 10.1016/j.amepre.2015.07.016] [Medline: 26590942]
16. Maita KC, Maniaci MJ, Haider CR, Avila FR, Torres-Guzman RA, Borna S, et al. The impact of digital health solutions on bridging the health care gap in rural areas: a scoping review. Perm J 2024;28(3):130-143 [FREE Full text] [doi: 10.7812/TPP/23.134] [Medline: 39135461]
17. Dasari S, Mehreen R, Baker Spohn K, Ostrovsky A. Opportunities for CMS to improve healthcare access and equity through advancing technology-enabled startups and digital health innovations. NPJ Digit Med 2024;7(1):23 [FREE Full text] [doi: 10.1038/s41746-024-00997-x] [Medline: 38291101]
18. Zhang X, Hailu B, Tabor DC, Gold R, Sayre MH, Sim I, et al. Role of health information technology in addressing health disparities: patient, clinician, and system perspectives. Med Care 2019;57 Suppl 6 Suppl 2(Suppl 6 2):S115-S120 [FREE Full text] [doi: 10.1097/MLR.0000000000001092] [Medline: 31095049]
19. Urlings J, Sezer S, Ter Laan M, Bartels R, Maal T, Boogaarts J, et al. The role and effectiveness of augmented reality in patient education: a systematic review of the literature. Patient Educ Couns 2022;105(7):1917-1927 [FREE Full text] [doi: 10.1016/j.pec.2022.03.005] [Medline: 35341611]
20. Hassan A, Scherer EA, Pikcilingis A, Krull E, McNickles L, Marmon G, et al. Improving social determinants of health: effectiveness of a web-based intervention. Am J Prev Med 2015;49(6):822-831. [doi: 10.1016/j.amepre.2015.04.023] [Medline: 26215831]
21. Paccoud I, Baumann M, Le Bihan E, Pétré B, Breinbauer M, Böhme P, et al. Socioeconomic and behavioural factors associated with access to and use of personal health records. BMC Med Inform Decis Mak 2021;21(1):18 [FREE Full text] [doi: 10.1186/s12911-020-01383-9] [Medline: 33435970]
22. Kavanaugh AL, Puckett A, Tatar DG. Scaffolding technology for low literacy groups: from mobile phone to desktop PC? Int J Hum Comput Int 2013;29(4):274-288. [doi: 10.1080/10447318.2013.765766]

XSL•FO
RenderX

23. Livingood WC, Bautista MAB, Smotherman C, Azueta D, Coleman J, Grewal R, et al. Comparative study of different SES neighborhood clinics for health literacy and internet access. Digit Health 2022;8:20552076221123715 [FREE Full text] [doi: 10.1177/20552076221123715] [Medline: 36081750]
24. Arias López MDP, Ong BA, Borrat Frigola X, Fernández AL, Hicklent RS, Obeles AJT, et al. Digital literacy as a new determinant of health: a scoping review. PLOS Digit Health 2023;2(10):e0000279 [FREE Full text] [doi: 10.1371/journal.pdig.0000279] [Medline: 37824584]
25. Andermann A, CLEAR Collaboration. Taking action on the social determinants of health in clinical practice: a framework for health professionals. CMAJ 2016;188(17-18):E474-E483 [FREE Full text] [doi: 10.1503/cmaj.160177] [Medline: 27503870]
26. Norman DA, Draper SW. User Centered System Design: New Perspectives on Human-computer Interaction. Boca Raton: CRC Press; 1986.
27. Pichon A, Jackman K, Winkler I, Bobel C, Elhadad N. The messiness of the menstruator: assessing personas and functionalities of menstrual tracking apps. J Am Med Inform Assoc 2022;29(2):385-399 [FREE Full text] [doi: 10.1093/jamia/ocab212] [Medline: 34613388]
28. Faber JS, Al-Dhahir I, Kraal JJ, Breeman LD, van den Berg-Emons RJG, Reijnders T, et al. Guide development for eHealth interventions targeting people with a low socioeconomic position: participatory design approach. J Med Internet Res 2023;25:e48461. [doi: 10.2196/48461] [Medline: 38048148]
29. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
30. Hevner A. A three cycle view of design science research. Scand J Inf Syst 2007;19(2) [FREE Full text]
31. Hevner A, March ST, Park J, Ram S. Design science in information systems research. MIS Q 2004;28(1):75-105. [doi: 10.2307/25148625]
32. Hevner A, March ST. IT systems perspectives - the information systems research cycle. Computer 2003;36(11):111-113. [doi: 10.1109/MC.2003.1244541]
33. Peled Y, Kurtz G, Avidov-Ungar O. Pathways to a knowledge society: a proposal for a hierarchical model for measuring digital literacy among israeli pre-service teachers. Electron J e Learn 2021;19(3):118-132. [doi: 10.34190/ejel.19.3.2217]
34. Panel I. Digital Transformation: A Framework for ICT Literacy, a Report of the International ICT Literacy Panel. Princeton, New Jersey, U.S: Educational Testing Service; 2001.
35. Covello S. A review of digital literacy assessment instruments. IDE-712: analysis for human performance technology decisions. Academia. 2010. URL: https://www.academia.edu/7935447/A_Review_of_Digital_Literacy_Assessment_Instruments [accessed 2025-08-08]
36. LeRouge C, Ma J, Sneha S, Tolle K. User profiles and personas in the design and development of consumer health technologies. Int J Med Inform 2013;82(11):e251-e268. [doi: 10.1016/j.ijmedinf.2011.03.006] [Medline: 21481635]
37. Fitzpatrick L. The Time Is Now: The Case for Digital Health Innovation for the Poor and Underserved. New York City, U.S: The Commonwealth Fund; 2018.
38. Smith CA, Keselman A. Consumer health informatics : enabling digital health for everyone. First edition ed. Boca Raton, Florida: CRC Press; 2021.
39. Agness C, Murrell E, Nkansah N, Martin C. Poor health literacy as a barrier to patient care. Consult Pharm 2008;23(5):378-82, 385. [doi: 10.4140/tcp.n.2008.378] [Medline: 18540791]
40. Schillinger D. The intersections between social determinants of health, health literacy, and health disparities. Stud Health Technol Inform 2020;269:22-41 [FREE Full text] [doi: 10.3233/SHTI200020] [Medline: 32593981]
41. Towe VL, May LW, Huang W, Martin LT, Carman K, Miller CE, et al. Drivers of differential views of health equity in the U.S.: is the U.S. ready to make progress? results from the 2018 National Survey of Health Attitudes. BMC Public Health 2021;21(1):175 [FREE Full text] [doi: 10.1186/s12889-021-10179-z] [Medline: 33478438]
42. Elgaddal N, Kramarow EA, Reuben C. Physical activity among adults aged 18 and over: United States, 2020. NCHS Data Brief 2022(443):1-8. [Medline: 36043905]
43. High school graduation. Office of Disease Prevention and Health Promotion. URL: https://odphp.health.gov/healthypeople/priority-areas/social-determinants-health/literature-summaries/high-school-graduation [accessed 2024-11-22]
44. Trends in high school dropout and completion rates in the United States. National Center for Education Statistics. 2018. URL: https://nces.ed.gov/pubs2018/2018117.pdf [accessed 2024-11-22]
45. Food insecurity. Office of Disease Prevention and Health Promotion. URL: https://odphp.health.gov/healthypeople/priority-areas/social-determinants-health/literature-summaries/food-insecurity [accessed 2024-11-22]
46. Food insecurity in the U.S. - key statistics and graphics. Economic Research Service. 2024. URL: https://www.ers.usda.gov/topics/food-nutrition-assistance/food-security-in-the-u-s/key-statistics-graphics/ [accessed 2024-11-22]
47. 2021 national healthcare quality and disparities report. National Center for Biotechnology Information. 2021. URL: https://www.ncbi.nlm.nih.gov/books/NBK578529/ [accessed 2025-08-08]
48. Housing instability. Office of Disease Prevention and Health Promotion. URL: https://odphp.health.gov/healthypeople/priority-areas/social-determinants-health/literature-summaries/housing-instability [accessed 2024-11-22]

XSL•FO
RenderX

49. The state of the nation's housing 2020. Joint Center for Housing Studies of Harvard University. 2020. URL: https://www.jchs.harvard.edu/state-nations-housing-2020 [accessed 2024-11-22]
50. Republican gains in 2022 midterms driven mostly by turnout advantage. Pew Research Center. 2023. URL: https://www.pewresearch.org/politics/2023/07/12/republican-gains-in-2022-midterms-driven-mostly-by-turnout-advantage/ [accessed 2024-11-22]
51. Poverty. Office of Disease Prevention and Health Promotion. URL: https://odphp.health.gov/healthypeople/priority-areas/social-determinants-health/literature-summaries/poverty#cit3 [accessed 2024-11-22]
52. Shrider EA, Kollar M, Chen F, Semega J. Income and poverty in the United States: 2020. United States Census Bureau. 2020. URL: https://www.census.gov/library/publications/2021/demo/p60-273.html [accessed 2024-11-22]
53. Chetty R, Jackson M, Kuchler T, Stroebel J, Hendren N, Fluegge R, et al. Social capital I: measurement and associations with economic mobility. Nature 2022;608(7921):108-121 [FREE Full text] [doi: 10.1038/s41586-022-04996-4] [Medline: 35915342]
54. Vogels EA. Digital divide persists even as Americans with lower incomes make gains in tech adoption. Pew Research Center. 2021. URL: https://www.pewresearch.org/short-reads/2021/06/22/digital-divide-persists-even-as-americans-with-lower-incomes-make-gains-in-tech-adoption/ [accessed 2025-08-08]
55. Ng AE, Adjaye-Gbewonyo D, Dahlhamer JJ. Lack of reliable transportation for daily living among adults: United States, 2022. NCHS Data Brief 2024(490):1-8 [FREE Full text] [Medline: 38252448]
56. Weissman J, Pratt L, Miller E, Parker J. Serious psychological distress among adults: United States, 2009-2013. National Center for Health Statistics, Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/products/databriefs/db203.htm [accessed 2024-11-22]
57. A profile of the working poor, 2021. U.S. Bureau of Labor Statistics. 2023. URL: https://www.bls.gov/opub/reports/working-poor/2021/ [accessed 2024-11-22]
58. Wardle J, Steptoe A. Socioeconomic differences in attitudes and beliefs about healthy lifestyles. J Epidemiol Community Health 2003;57(6):440-443 [FREE Full text] [doi: 10.1136/jech.57.6.440] [Medline: 12775791]
59. Reshma P, Rajkumar E, John R, George AJ. Factors influencing self-care behavior of socio-economically disadvantaged diabetic patients: a systematic review. Health Psychol Open 2021;8(2):20551029211041427 [FREE Full text] [doi: 10.1177/20551029211041427] [Medline: 34552758]
60. LeRouge C, Austin E, Lee J, Segal C, Sangameswaran S, Hartzler A. ePROs in Clinical Care: Guidelines and Tools for Health Systems. Seattle, WA: CERTAIN, University of Washington; 2020.
61. Craig S, Bejan R, Muskat B. Making the invisible visible: are health social workers addressing the social determinants of health? Soc Work Health Care 2013;52(4):311-331. [doi: 10.1080/00981389.2013.764379] [Medline: 23581836]
62. Kangovi S, Mitra N, Norton L, Harte R, Zhao X, Carter T, et al. Effect of community health worker support on clinical outcomes of low-income patients across primary care facilities: a randomized clinical trial. JAMA Intern Med 2018;178(12):1635-1643 [FREE Full text] [doi: 10.1001/jamainternmed.2018.4630] [Medline: 30422224]
63. Collins ME, Mowbray CT, Bybee D. Establishing individualized goals in a supported education intervention: program influences on goal-setting and attainment. Res Soc Work Pract 1999;9(4):483-507. [doi: 10.1177/104973159900900405]
64. Partiwi N. Investigate the effectiveness, mechanisms, and scalability of community-based interventions designed to reduce health disparities. J Adv Med Med Res 2024;2(1):49-65 [FREE Full text] [doi: 10.60079/ahr.v2i1.371]
65. Brakefield WS, Olusanya OA, White B, Shaban-Nejad A. Social determinants and indicators of COVID-19 among marginalized communities: a scientific review and call to action for pandemic response and recovery. Disaster Med Public Health Prep 2022;17:e193 [FREE Full text] [doi: 10.1017/dmp.2022.104] [Medline: 35492024]
66. Masterson JM, Luu M, Dallas KB, Daskivich LP, Spiegel B, Daskivich TJ. Disparities in COVID-19 disease incidence by income and vaccination coverage - 81 communities, Los Angeles, California, July 2020-September 2021. MMWR Morb Mortal Wkly Rep 2023;72(26):728-731 [FREE Full text] [doi: 10.15585/mmwr.mm7226a5] [Medline: 37384567]
67. Ringlein GV, Ettman CK, Stuart EA. Income or job loss and psychological distress during the COVID-19 pandemic. JAMA Netw Open 2024;7(7):e2424601 [FREE Full text] [doi: 10.1001/jamanetworkopen.2024.24601] [Medline: 39078628]
68. Singu S, Acharya A, Challagundla K, Byrareddy SN. Impact of social determinants of health on the emerging COVID-19 pandemic in the United States. Front Public Health 2020;8:406 [FREE Full text] [doi: 10.3389/fpubh.2020.00406] [Medline: 32793544]
69. Liu P, Astudillo K, Velez D, Kelley L, Cobbs-Lomax D, Spatz ES. Use of mobile health applications in low-income populations: a prospective study of facilitators and barriers. Circ Cardiovasc Qual Outcomes 2020;13(9):e007031 [FREE Full text] [doi: 10.1161/CIRCOUTCOMES.120.007031] [Medline: 32885681]
70. Arpey NC, Gaglioti AH, Rosenbaum ME. How socioeconomic status affects patient perceptions of health care: a qualitative study. J Prim Care Community Health 2017;8(3):169-175 [FREE Full text] [doi: 10.1177/2150131917697439] [Medline: 28606031]
71. Thomas-Henkel C, Schulman M. Screening for social determinants of health in populations with complex needs: implementation considerations. Center for Health Care Strategies. 2017. URL: https://www.chcs.org/media/SDOH-Complex-Care-Screening-Brief-102617.pdf [accessed 2024-11-22]

XSL•FO
RenderX

72. Liu P, Astudillo K, Velez D, Kelley L, Cobbs-Lomax D, Spatz ES. Use of mobile health applications in low income populations: a prospective study of facilitators and barriers. Circ Cardiovasc Qual Outcomes 2020:e007031 [FREE Full text] [doi: 10.1161/circoutcomes.120.007031]
73. The smartphone difference. Pew Research Center. 2015. URL: https://www.pewresearch.org/internet/2015/04/01/us-smartphone-use-in-2015/ [accessed 2024-11-22]
74. Madden M. Privacy, security, and digital inequality. Data & Society. 2017. URL: https://datasociety.net/library/privacy-security-and-digital-inequality/ [accessed 2024-11-22]
75. Williams C, Shang D. Telehealth usage among low-income racial and ethnic minority populations during the COVID-19 pandemic: retrospective observational study. J Med Internet Res 2023;25:e43604 [FREE Full text] [doi: 10.2196/43604] [Medline: 37171848]
76. Gordon A, Liu Y, Tavitian K, York B, Finnell SM, Agiro A. Bridging health and temporary housing services for medicaid members experiencing homelessness: program impact on health care utilization, costs, and well-being. J Health Care Poor Underserved 2021;32(4):1949-1964. [doi: 10.1353/hpu.2021.0175] [Medline: 34803052]
77. Nielsen J. Enhancing the Explanatory Power of Usability Heuristics. 1994 Presented at: CHI '94: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; 1994 April 24; Massachusetts, Boston, USA p. 152-158 URL: https://doi.org/10.1145/191666.191729
78. Greer PJ, Brown DR, Brewster LG, Lage OG, Esposito KF, Whisenant EB, et al. Socially accountable medical education: an innovative approach at Florida International University Herbert Wertheim College of Medicine. Acad Med 2018;93(1):60-65 [FREE Full text] [doi: 10.1097/ACM.0000000000001811] [Medline: 28658020]
79. Rock JA, Acuña JM, Lozano JM, Martinez IL, Greer PJ, Brown DR, et al. Impact of an academic-community partnership in medical education on community health: evaluation of a novel student-based home visitation program. South Med J 2014;107(4):203-211. [doi: 10.1097/SMJ.0000000000000080] [Medline: 24937510]
80. Elder GH. Time, human agency, and social change: perspectives on the life course. Soc Psychol Q 1994;57(1):4-15. [doi: 10.2307/2786971]
81. Settersten RA, Ray B. What's going on with young people today? the long and twisting path to adulthood. Future Child 2010;20(1):19-41. [doi: 10.1353/foc.0.0044] [Medline: 20364620]
82. Creswell JW, Poth CN. Qualitative Inquiry and Research Design: Choosing Among Five Approaches. Los Angeles, London: SAGE Publications; 2017.
83. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907 [FREE Full text] [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
84. Guba EG. Criteria for assessing the trustworthiness of naturalistic inquiries. ECTJ 1981;29(2):75-91. [doi: 10.1007/bf02766777]
85. Morgan DL. Focus groups as qualitative research, 2nd ed. Thousand Oaks, CA, US: Sage Publications, Inc; 1997.
86. Birt L, Scott S, Cavers D, Campbell C, Walter F. Member checking: a tool to enhance trustworthiness or merely a nod to validation? Qual Health Res 2016;26(13):1802-1811. [doi: 10.1177/1049732316654870] [Medline: 27340178]
87. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis. Int J Qual Methods 2017;16(1):160940691773384. [doi: 10.1177/1609406917733847]
88. King N. Using templates in the thematic analysis of text. In: Essential Guide to Qualitative Methods in Organizational Research. London: Sage Publications; 2004:257-270.
89. Trauth EM, Jessup LM. Understanding computer-mediated discussions: positivist and interpretive analyses of group support system use. MIS Q 2000;24(1):43-79. [doi: 10.2307/3250979]
90. Boyatzis RE. Transforming Qualitative Information: Thematic Analysis and Code Development. Thousand Oaks, CA: Sage Publications, Inc; 1998.
91. Scott C, Medaugh M. Axial Coding. The International Encyclopedia of Communication Research Methods. Los Angeles, Washington DC: Sage; 2024:1-2.
92. Weber R. Basic content analysis. In: Quantitative Applications in the Social Sciences. Thousand Oaks, California: SAGE Publications, Inc; 1990.
93. Kongjit C, Nimmolrat A, Khamaksorn A. Mobile health application for Thai women: investigation and model. BMC Med Inform Decis Mak 2022;22(1):202 [FREE Full text] [doi: 10.1186/s12911-022-01944-0] [Medline: 35907950]
94. Bonet-Olivencia S, Carrillo-Leal J, Rao A, Sasangohar F. User-centered design of a diabetes self-management tool for underserved populations. J Diabetes Sci Technol 2024;18(1):22-29. [doi: 10.1177/19322968231212220] [Medline: 37978811]
95. Gabrielli S, Dianti M, Maimone R, Betta M, Filippi L, Ghezzi M, et al. Design of a mobile app for nutrition education (TreC-LifeStyle) and formative evaluation with families of overweight children. JMIR Mhealth Uhealth 2017;5(4):e48 [FREE Full text] [doi: 10.2196/mhealth.7080] [Medline: 28408361]
96. Greysen SR, Chin Garcia C, Sudore RL, Cenzer IS, Covinsky KE. Functional impairment and Internet use among older adults: implications for meaningful use of patient portals. JAMA Intern Med 2014;174(7):1188-1190 [FREE Full text] [doi: 10.1001/jamainternmed.2014.1864] [Medline: 24839165]

XSL•FO
RenderX

97. Mace R, Mattos M, Vranceanu A. Older adults can use technology: why healthcare professionals must overcome ageism in digital health. Transl Behav Med 2022;12(12):1102-1105 [FREE Full text] [doi: 10.1093/tbm/ibac070] [Medline: 36073770]
98. Brown DR, Hernández A, Saint-Jean G, Evans S, Tafari I, Brewster LG, et al. A participatory action research pilot study of urban health disparities using rapid assessment response and evaluation. Am J Public Health 2008;98(1):28-38. [doi: 10.2105/AJPH.2006.091363] [Medline: 18048802]
99. Solar O, Irwin A. A conceptual framework for action on the social determinants of health. World Health Organization. 2010. URL: https://www.who.int/publications/i/item/9789241500852 [accessed 2024-11-22]
100. McMaughan DJ, Oloruntoba O, Smith ML. Socioeconomic status and access to healthcare: interrelated drivers for healthy aging. Front Public Health 2020;8:231 [FREE Full text] [doi: 10.3389/fpubh.2020.00231] [Medline: 32626678]
101. Chi NTK, Hoang Vu N. Investigating the customer trust in artificial intelligence: the role of anthropomorphism, empathy response, and interaction. CAAI Trans Intel Tech 2022;8(1):260-273 [FREE Full text] [doi: 10.1049/cit2.12133]
102. Pedamallu H, Ehrhardt MJ, Maki J, Carcone AI, Hudson MM, Waters EA. Technology-delivered adaptations of motivational interviewing for the prevention and management of chronic diseases: scoping review. J Med Internet Res 2022;24(8):e35283 [FREE Full text] [doi: 10.2196/35283] [Medline: 35943775]
103. Smith CA, Keselman A. Consumer Health Informatics: Enabling Digital Health for Everyone. FL, USA: CRC Press; 2020.
104. Patel V, Dhopeshwarkar R, Edwards A, Barron Y, Likourezos A, Burd L, et al. Low-income, ethnically diverse consumers' perspective on health information exchange and personal health records. Inform Health Soc Care 2011;36(4):233-252. [doi: 10.3109/17538157.2011.554930] [Medline: 21851182]
105. Marwick A, Fontaine C, boyd D. “Nobody Sees It, Nobody Gets Mad”: social media, privacy, and personal responsibility among low-SES youth. Soc Media Soc 2017;3(2):205630511771045 [FREE Full text] [doi: 10.1177/2056305117710455]
106. Pew Research Center. Americans and privacy: concerned, confused and feeling lack of control over their personal information. Pew Research Center. 2019. URL: https://www.pewresearch.org/internet/2019/11/15/americans-and-privacy-concerned-confused-and-feeling-lack-of-control-over-their-personal-information/ [accessed 2024-11-22]
107. Rathbone A, Stumpf S, Claisse C, Sillence E, Coventry L, Brown R, et al. People with long-term conditions sharing personal health data via digital health technologies: a scoping review to inform design. PLOS Digit Health 2023;2(5):e0000264 [FREE Full text] [doi: 10.1371/journal.pdig.0000264] [Medline: 37224154]
108. Bernaerdt J, Moerenhout T, Devisch I. Vulnerable patients' attitudes towards sharing medical data and granular control in patient portal systems: an interview study. J Eval Clin Pract 2021;27(2):429-437. [doi: 10.1111/jep.13465] [Medline: 32886953]
109. Khor WS, Baker B, Amin K, Chan A, Patel K, Wong J. Augmented and virtual reality in surgery-the digital surgical environment: applications, limitations and legal pitfalls. Ann Transl Med 2016;4(23):454 [FREE Full text] [doi: 10.21037/atm.2016.12.23] [Medline: 28090510]
110. Brammer SV, Regan SL, Collins CM, Gillespie GL. Developing innovative virtual reality simulations to increase health care providers' understanding of social determinants of health. J Contin Educ Health Prof 2022;42(1):60-65. [doi: 10.1097/CEH.0000000000000400] [Medline: 34799522]
111. Mehrabi S, Muñoz JE, Basharat A, Boger J, Cao S, Barnett-Cowan M, et al. Immersive virtual reality exergames to promote the well-being of community-dwelling older adults: protocol for a mixed methods pilot study. JMIR Res Protoc 2022;11(6):e32955 [FREE Full text] [doi: 10.2196/32955] [Medline: 35700014]
112. Yew SQ, Trivedi D, Adanan NIH, Chew BH. Facilitators and barriers to the implementation of digital health technologies in hospital settings in lower- and middle-income countries since the onset of the COVID-19 pandemic: scoping review. J Med Internet Res 2025;27:e63482 [FREE Full text] [doi: 10.2196/63482] [Medline: 40053793]
113. Kaboré SS, Ngangue P, Soubeiga D, Barro A, Pilabré AH, Bationo N, et al. Barriers and facilitators for the sustainability of digital health interventions in low and middle-income countries: a systematic review. Front Digit Health 2022;4:1014375 [FREE Full text] [doi: 10.3389/fdgth.2022.1014375] [Medline: 36518563]
114. Hakobyan L, Lumsden J, O'Sullivan D. Participatory design: how to engage older adults in participatory design activities. Int J Mob Hum Comput Interact 2015;7(3):78-92 [FREE Full text]
115. Wilkinson C, De Angeli A. Applying user centred and participatory design approaches to commercial product development. Des Stud 2014;35(6):614-631 [FREE Full text] [doi: 10.1016/j.destud.2014.06.001]
116. Ellis RD, Kurniawan SH. Increasing the usability of online information for older users: a case study in participatory design. Int J Mob Hum Comput Interact 2000;12(2):263-276. [doi: 10.1207/s15327590ijhc1202_6]
117. Dixon BE, Embi PJ, Haggstrom DA. Information technologies that facilitate care coordination: provider and patient perspectives. Transl Behav Med 2018;8(3):522-525 [FREE Full text] [doi: 10.1093/tbm/ibx086] [Medline: 29800411]

Abbreviations

AR: augmented reality
COREQ: consolidated criteria for reporting qualitative research

XSL•FO
RenderX

DHT: digital health technology
DSR: design science research
EMR: electronic medical record
FIU: Florida International University
HIT: health information technology
HRSN: health-related social needs
IT: information technology
mHealth: mobile health
NeighborhoodHELP: Neighborhood Health Education Learning Program
SDOH: social determinants of health
SES: socioeconomic status
UCD: user-centered design
VR: virtual reality

Edited by A Kushniruk; submitted 02.12.24; peer-reviewed by FAE Sosso, B Chakalov, C Udensi; comments to author 27.03.25; revised version received 22.05.25; accepted 23.06.25; published 12.09.25.

Please cite as:
Haydon KL, LeRouge C, Durneva P, Diaz Campo MP, Brown DR
Evaluating the Needs and Characteristics of Individuals of Low Socioeconomic Status Using Digital Health Technology to Address Health-Related Social Needs: Mixed Methods Study With Patients and Care Providers
JMIR Hum Factors 2025;12:e69545
URL: https://humanfactors.jmir.org/2025/1/e69545
doi:10.2196/69545
PMID:

©Kayla Li Haydon, Cynthia LeRouge, Polina Durneva, Maria Paula Diaz Campo, David Richard Brown. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 12.09.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Collaborative Design and Development of a Patient-Centered Digital Health App for Supportive Cancer Care: Participatory Study

Sonia Difrancesco[1], PhD; Matthia Martina Bauert[2], MSc; Claude Lehmann[3], MSc; Steven Häsler[3], MSc; Yi Zhang[3], MSc; Sven Hirsch[4], Prof Dr; Philipp Ackermann[3], PhD; Kurt Stockinger[3], Prof Dr; Monika Reif[5], Prof Dr; Sunjoy Mathieu[4], MA; Anna Götz[6], PhD; Andreas Wicki[6], Prof Dr Med; Michael Krauthammer[1], Prof Dr; Claudia M Witt[7], Prof Dr Med

[1]Department of Quantitative Biomedicine, University of Zurich, Zurich, Switzerland

[2]Department of Health Science and Technology, ETH Zurich, Zurich, Switzerland

[3]Institute of Computer Science, ZHAW Zurich University of Applied Sciences, Winterthur, Switzerland

[4]Institute of Computational Life Sciences, ZHAW Zurich University of Applied Sciences, Wädenswil, Switzerland

[5]Institute of Applied Mathematics and Physics, ZHAW Zurich University of Applied Sciences, Winterthur, Switzerland

[6]Department of Medical Oncology and Hematology, University Hospital Zurich and University of Zurich, Zurich, Switzerland

[7]Institute for Complementary and Integrative Medicine, University Hospital Zurich and University of Zurich, Zurich, Switzerland

Corresponding Author:
Sonia Difrancesco, PhD
Department of Quantitative Biomedicine
University of Zurich
Schmelzbergstrasse 26
Zurich, 8006
Switzerland
Phone: 41 44 635 66 31
Email: sonia.difrancesco@uzh.ch

Abstract

Background: Digital health tools such as smartphone apps have the potential to improve supportive cancer care. Although numerous smartphone apps for supportive care are available, few are designed using a user-centered approach. Such an approach is crucial for successful implementation, as it may improve user engagement, usability, and adoption in clinical settings.

Objective: This study aimed to co-design and develop a digital health app for supportive cancer care in collaboration with patients with cancer and health care professionals and to explore factors influencing its future acceptance.

Methods: We conducted a participatory study with the major stakeholders at the University Hospital Zurich. Workshops, individual qualitative interviews, and focus groups were held with health care professionals, survivors of cancer, and patients with cancer. The co-design process was divided into 3 phases: predesign, generative phase, and prototyping. User-centered design methods included scoring cards and think-aloud protocols to co-create design ideas, identify important functionalities, and test usability. Qualitative data were analyzed using thematic analysis.

Results: Patients and health care professionals emphasized the need for a digital health app to improve patient–healthcare professional communication, digitalize supportive care screening and processes, and enhance self-efficacy. The resulting app, OncoSupport+, was co-designed and integrated into the clinical workflow for supportive cancer care. It consists of (1) a patient dashboard to record patient-reported outcome measures and to provide access to personalized supportive care information and contact details, and (2) a nurse dashboard to visualize patient data, which can be used during nursing consultations. Potential facilitators for adoption included ease of use, workflow integration, introduction by health care professionals, and technical support, whereas internet anxiety may be a potential barrier.

Conclusions: Collaborative development with patients and health care professionals is crucial for creating digital health tools that can be implemented successfully. Future research should evaluate the feasibility of long-term implementation and the real-world usability and effectiveness of OncoSupport+ for improving communication, self-efficacy, and quality of life.

(JMIR Hum Factors 2025;12:e73829) doi:10.2196/73829

XSL•FO
RenderX

KEYWORDS

cancer; digital health; mHealth; patient-reported outcome measures; supportive care

Introduction

Cancer is a highly prevalent condition with increasing incidence rates and is the leading cause of death worldwide [1]. As improvements in diagnosis, treatment, and surveillance have led to better survival rates, patients with cancer are living longer [2], though a considerable portion of patients experience significant morbidity and symptoms related to both the cancer and its treatment. Patients with cancer often exhibit physical, emotional, social, spiritual, and informational needs, also known as supportive care needs, and the aggregated impact may ultimately contribute to exacerbated health outcomes and diminished quality of life [3]. Supportive care, which focuses on the prevention and management of cancer symptoms and cancer treatment–related symptoms, has been widely acknowledged as an essential component of comprehensive cancer care [4]. It includes the management of physical and psychological symptoms and side effects across the cancer care trajectory from diagnosis through treatment to posttreatment care. Examples of supportive care include management of symptoms such as nausea, nutritional and exercise support, psychological support, and practical assistance for concerns such as transportation [4].

Several cancer centers have established specialized supportive care services, employing the expertise of a multidisciplinary team with diverse medical specializations such as nutritionists, physiotherapists, and psychologists. Supportive care has been shown to improve quality of life [5], tolerability of cancer treatments, survival, and systemic health economic benefits [6]. However, studies continue to report unmet supportive care needs, such as physical and informational needs, indicating that significant efforts are still required to ensure the effective implementation of evidence-based supportive care in clinical practice [7]. Specifically, the Multinational Association of Supportive Care in Cancer has emphasized the routine collection of patient-reported outcome measures (PROMs) to facilitate timely, individualized supportive care and empower patients through evidence-based education [7]. Commonly used PROMs include the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 (EORTC QLQ-C30) [8] and the National Comprehensive Cancer Network (NCCN) Distress Thermometer [9]. In this context, digital health tools, particularly mobile health (mHealth) apps, have emerged as promising solutions for delivering patient-centered supportive care interventions [10].

A growing number of digital health applications offer mental health support, symptom self-management, and educational resources. Among these, digital symptom monitoring and electronic patient-reported outcome measures (ePROMs) are commonly integrated interventions for supportive cancer care [11,12] and provide advantages over traditional paper-based PROMs or retrospective questionnaires, which are often limited by recall bias and low completion rates [13,14] . As digital health apps enable the remote collection of symptoms and supportive care needs in daily living environments, these tools allow for a more timely and accurate assessment of patient well-being. Research has shown that such applications can improve communication with health care professionals, facilitate connections with other patients, and enhance self-efficacy, thereby complementing clinical care [15]. Systematic reviews and meta-analyses have further suggested that digital health interventions may improve symptoms and quality of life [11,15]. Despite these findings, their implementation and adoption remain limited, particularly in cancer care [16]. A key reason is the limited involvement of key stakeholders, including patients and health care professionals, during development, leading to suboptimal adoption and effectiveness [10,11,17].

Co-design has been identified as a promising approach to ensure that digital health interventions are patient-centered, clinically relevant, and technically feasible [18]. By incorporating participatory and iterative stakeholder engagement, co-design can help tailor digital health interventions to meet the needs and preferences of both patients and health care providers while facilitating the early identification of implementation barriers and factors influencing adoption and long-term engagement. However, despite its potential, co-design is not always rigorously applied in digital health research, as studies often lack structured methodologies, fail to systematically integrate diverse stakeholder input, and do not adequately assess long-term feasibility [19]. Addressing these gaps is crucial for developing user-centered and sustainable digital health interventions in supportive cancer care.

The overall aim of this study was to collaboratively design and develop a digital health app for supportive cancer care together with the relevant stakeholders (patients with cancer, survivors of cancer, and health care professionals). Specifically, the first objective was to map the context of supportive cancer care at the University Hospital Zurich, including challenges and needs. The second objective was to identify specific functionalities and features of the app and gather the patients' and nurses' perspectives; as part of this objective, we also explored factors that can potentially influence the uptake and implementation of the digital health app for supportive cancer care. Finally, the third objective was to design and develop the prototype and user interface of the digital health app to be technically implemented.

Methods

Ethical Considerations

This study was reviewed by the Ethics Committee of the Canton of Zurich, which determined that it did not fall within the scope of the Swiss Human Research Act and therefore did not require formal approval (Req-2023-01095). According to the Swiss Human Research Act (SR 810.30), ethical approval is required for studies involving health-related personal data, biological material, or clinical interventions. This study involved collecting opinions and experiences from patients, advocates, and nurses to inform app development; no health-related personal data were analyzed. Patients, patient advocates, and nurses received written

XSL•FO
RenderX

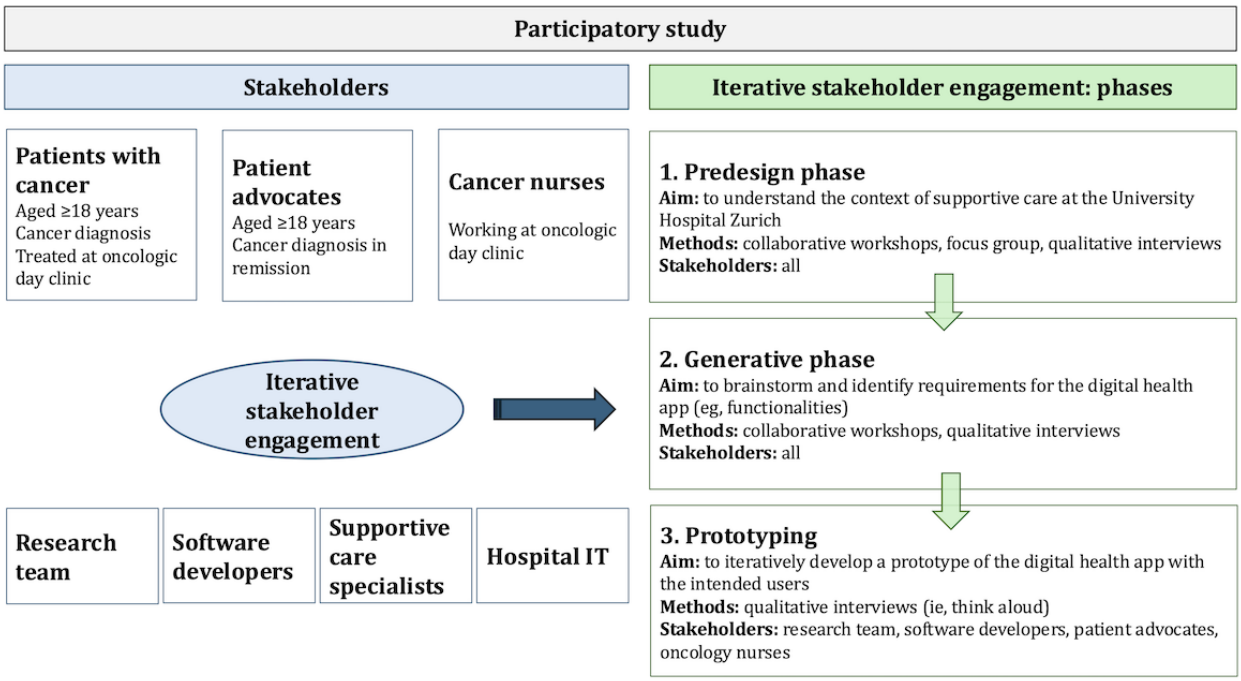
and oral information about the study objectives and procedures, and written informed consent was obtained before participation. Data were collected in a pseudonymized manner. Transcripts and digital records were deidentified before analysis, and all personal identifiers were removed. Access to study data was restricted to the research team. As patient advocates were part of Swiss Group for Clinical Cancer Research (SAKK), they received financial compensation for their contribution. No quotes or materials that could identify individual participants are included in the manuscript or supplementary materials.

Study Design

This participatory study was conducted between May 2023 and October 2024 as part of the Digital Health Zurich initiative [20]—a collaboration between the University of Zurich, the University Hospital of Zurich, and the Zurich University of Applied Sciences. The overarching goal of Digital Health Zurich is to create a sustainable, data-driven, and cross-institutional digital ecosystem by fostering collaboration between patients, health care professionals, technology developers, and researchers in Zurich and beyond. The initiative aims to translate research into real-world digital health solutions, with a strong focus on patient and provider needs, participatory development, cross-institutional cooperation, and innovation within the health care system. As a use case, this study was conducted at the University Hospital Zurich in collaboration with the Department of Oncology and Hematology and the Comprehensive Cancer Center Zurich.

The stakeholders included in the study were patients with cancer and patient advocates, cancer nurses, supportive care specialists, oncologists, and the research team; more details are listed in the Study Participants and Stakeholders section. Figure 1 shows an overview of the study design, using iterative stakeholder engagement, user-centered approaches, and design thinking techniques. The study can be divided into different phases, similar to the framework suggested by Noorbergen et al [18] for co-designing mHealth systems, which is a novel extension of the work by Sanders and Stappers [21]. The predesign phase aimed to understand the context of supportive cancer care at the University Hospital of Zurich, including the challenges faced by cancer nurses and patients in accessing supportive cancer care. The generative phase focused on brainstorming digital health app functionalities, as well as identifying potential factors that may impact future uptake of the technology. The prototyping phase, taking into account the findings from the previous phases, aimed to iteratively develop a prototype of the digital health app by gathering feedback from the intended users. Between June 2023 and October 2024, the research team held collaborative workshops, focus groups with cancer nurses, and qualitative interviews with patients and patient advocates.

Figure 1. Participatory study design using iterative stakeholder engagement, divided into 3 phases: predesign, generative, and prototyping.



Study Participants and Stakeholders

Patients With Cancer and Patient Advocates

Patients with cancer and patient advocates were recruited for the study. Inclusion criteria for patients with cancer were: (1) current treatment for cancer at the oncologic day clinic of the Department of Oncology and Hematology, (2) age ≥18 years, and (3) ability to speak German, Italian, or English. As patients with cancer often experience fatigue and may face challenges in participating in research activities, patient advocates from the SAKK (today Swiss Cancer Institute) were also included. These advocates were survivors of cancer in remission, enabling them to contribute insights from both patient and survivor perspectives. Their ability to engage more actively made them valuable contributors to the research. Inclusion criteria for patient advocates were (1) history of cancer, (2) age ≥18 years, and (3) ability to speak German, Italian, or English.

XSL•FO
RenderX

Cancer Nurses

Cancer nurses were included in the study. Inclusion criteria were employment at the Department of Oncology and Hematology of the University Hospital Zurich.

Supportive Care Specialists

Supportive care professionals, including a supportive care nurse, a nutritionist, 2 physiotherapists, and medical doctors with specializations in oncology, palliative care, and complementary medicine, were engaged in the participatory study to share their expertise and ensure that the app aligned with clinical practices in supportive cancer care.

Research Team

The research team consisted of one cancer nurse and one oncology doctor, researchers in digital health, computer science, computational health, and medical informatics, as well as a user experience designer. Two of the computer science researchers also contributed as software developers for the digital health app. Digital health researchers were responsible for organizing, moderating, and conducting the research activities. One of the computer science researchers supported data collection.

Recruitment of Patients With Cancer, Patient Advocates, and Cancer Nurses

Patients with cancer were recruited by cancer nurses from the Department of Oncology and Hematology at the University Hospital Zurich. Potential participants were introduced to the study through an informational flyer describing the study procedures and offering the option to participate in interviews while receiving systemic antineoplastic therapy at the oncologic day clinic. Interested patients could provide their contact details and suggest a convenient date and time for an interview with a researcher. Of the 8 patients who provided their contact details, a total of 6 agreed to participate in the study. Patient advocates were identified by the coordinator of patient advocates at SAKK, and 4 agreed to participate. A total of 5 nurses were recruited by a project team member who also works as an oncology nurse at the day clinic.

Study Procedure

Between June 2023 and October 2024, data were collected through collaborative workshops, focus groups with oncology nurses, and qualitative interviews with patients with cancer and patient advocates. Interviews and focus groups were conducted in German, English, or Italian, while workshops were held in English, the working language of the project team. Focus groups and interviews were audio-recorded, whereas workshops were documented by 2 researchers who took notes, cross-checked them, and shared the notes with the wider team.

Phase 1: Predesign

The predesign phase consisted of a collaborative workshop with the research team, one focus group with cancer nurses, and individual qualitative interviews with patients and patient advocates. The initial workshop aimed to map the context of supportive cancer care at the hospital and define the study problem statement. The subsequent 1.5-hour focus group with cancer nurses was guided by questions on their experiences with supportive care, challenges in daily practice, and opportunities for digital health to enhance supportive care. In parallel, interviews with patients with cancer and patient advocates explored their experiences with supportive care services and their views on how digital health tools could support their needs.

Phase 2: Generative

The generative phase included a second collaborative workshop with the research team, qualitative interviews with patients with cancer and patient advocates, and a workshop with supportive care specialists. Informed by the predesign phase, the research team specified potential functionalities for integration into the app. During the interviews, patients and patient advocates were asked to reflect on potential app functionalities, suggest additional features, and discuss factors that might influence the uptake of digital health interventions. To prioritize functionalities, a scoring exercise was included in which participants rated each feature on a scale from 1 to 10, with 10 indicating a “must-have” functionality. The collaborative workshop with the research team focused on brainstorming digital health solutions and proposing an initial structure for the app. As the preceding step highlighted the need to personalize supportive care for patients with cancer, a 2-hour workshop was conducted with supportive care specialists and digital health researchers, during which participants co-designed a rule-based algorithm to tailor supportive care.

Phase 3: Prototyping

From July to October 2024, the prototyping phase was conducted. Based on insights from the previous phases, a digital health researcher iteratively developed user requirements and specifications documented in Confluence [22], which were then translated into user interface designs with Figma [23] by the user experience designer. A user flow was created to outline app navigation, and this flow was tested by 2 cancer nurses and 4 patient advocates. The prototype of the user interface was subsequently developed and iteratively reviewed by the same group. Participants were encouraged to verbalize their impressions, identify challenges in navigation, and comment on the clarity and usefulness of the proposed functionalities. Finally, a workshop involving health care professionals and the research team was held to evaluate and further refine the prototype.

A meeting between the software developers and the IT department of the hospital was also held to assess different digital solutions and explore how these could potentially be integrated into the IT system of the hospital.

Data Analysis

The qualitative data from interviews and focus groups were transcribed, translated, and analyzed using thematic analysis, following Braun and Clarke [24] methodology.

All audio recordings were transcribed orthographically using OpenAI Whisper [25], performed locally to ensure data security. Transcripts were translated to English, initially using Whisper and then manually checked and corrected by bilingual researchers who conducted the interviews, ensuring high-quality transcription and translation. The corrected transcripts were

XSL•FO
RenderX

used for analysis. Thematic analysis was conducted following the six-step process [24]: (1) familiarization with the data, (2) generating codes, (3) searching for themes, (4) reviewing themes, (5) defining and naming themes, and (6) producing a report. To reduce bias, analyses were conducted by the first 2 coauthors of this study. During familiarization, the first 2 coauthors read transcripts multiple times and highlighted text segments relevant to 5 aspects: app and dashboard functionalities, barriers and facilitators to use, challenges with the health care system and supportive care, and experiences with digital devices. Initial codes were generated using an open coding approach, capturing meaningful segments without predefined categories. Codes were then grouped based on similarity to identify patterns and inform preliminary theme development.

Themes were iteratively reviewed, refined, and organized into coherent subthemes, with disagreements resolved through discussion between the first 2 coauthors. Theme names were defined to accurately reflect core concepts, and data extracts were used to support each theme. Results were summarized with mind maps, journey maps, and frequencies of reported themes. Median and ranges were reported to summarize scorecards on functionalities.

Unified Theory of Acceptance and Uptake of Technologies

As it is important to identify factors influencing acceptance early in the design phase to guide design choices and improve engagement, the Unified Theory of Acceptance and Use of Technology (UTAUT) framework was used to guide reporting of the results [26]. Originally developed to predict technology acceptance in organizational settings, the framework has been validated in health care contexts [27] and adapted for mHealth adoption in cancer care [28]. The validated framework by Philippi et al [28] identifies 3 key predictors of acceptance: performance expectancy, effort expectancy, and social influence. Internet anxiety moderates their effects.

Performance expectancy refers to the extent to which an individual believes the technology will be beneficial and is considered the strongest predictor of acceptance. Effort expectancy describes the perceived ease of use of the system, with higher effort expectancy (ie, lower ease of use) negatively impacting acceptance. Social influence refers to the degree to which users perceive that significant others encourage system usage, positively influencing adoption. Internet anxiety moderates the relationship between effort expectancy and social influence, potentially hindering adoption. Specifically, the UTAUT was used to report the identified themes related to barriers and facilitators to uptake.

Sample Size and the Principle of Saturation

Although the study initially aimed to enroll approximately 30 participants for qualitative interviews and focus groups, the principle of thematic and design saturation was applied [29]. Recruitment was stopped once no new themes emerged, ensuring that data collection was both rigorous and meaningful without overburdening participants. For instance, when 2 or more consecutive interviews or focus groups did not reveal new themes or suggest new design features, the recruitment was stopped. This approach is consistent with co-design studies in digital health, where sample sizes are often small and vary widely; a recent umbrella review reported that as few as 2 participants have been included in some co-design projects [30]. In addition, as mentioned by Arcia at al [31] what matters most is not the absolute number of participants, but the iterative design process and achievement of thematic and design saturation, both of which were ensured in this study.

Results

Demographics, Experience With Digital Tools, and Other Characteristics of Patients With Cancer, Patient Advocates, and Cancer Nurses

A total of 6 patients with cancer and 4 patient advocates participated in the study (Table 1). Their ages ranged from 29 to 75 years, with a mean of 50.3 years. Four participants were female. Most held advanced university degrees, with 6 out of 10 having completed a bachelor's degree or higher (ie, master's or PhD). None were employed full-time at the time of the study; 4 worked part-time, and 2 were retired. All participants owned both a smartphone and a computer. While all reported daily smartphone use, 7 also used a computer daily, and 6 owned a tablet. Half of the participants rarely needed assistance when using digital devices, 4 required help occasionally, and 1 needed assistance daily. All nurses who participated in the study were female, aged between 37 and 53 years, with at least 10 years of experience in oncology nursing.

XSL•FO
RenderX

Table 1. Sociodemographic characteristics, device ownership, daily use, and need for assistance with digital devices among patients with cancer and patient advocates (N=10).

Characteristic	Total sample (N=10)	Patient advocates (n=4)	Patients with cancer (n=6)
Age (years), mean (range)	50.3 (29-75)	51 (47-75)	47.5 (29-67)
Sex, n (%)			
Female	4 (40)	1 (25)	3 (50)
Male	6 (60)	3 (75)	3 (50)
Education level, n (%)			
Primary education	0 (0)	0 (0)	0 (0)
Secondary education	4 (40)	1 (25)	3 (50)
Tertiary education	6 (60)	3 (75)	3 (50)
Working status, n (%)			
Full time	0 (0)	0 (0)	0 (0)
Part time	4 (40)	2 (50)	2 (33.3)
Unable to work for health reasons	4 (40)	1 (25)	3 (50)
Retired	2 (20)	1 (25)	1 (16.7)
Unemployed	0 (0)	0 (0)	0 (0)
Ownership of digital devices, n (%)			
Computer or laptop	10 (100)	4 (100)	6 (100)
Smartphone	10 (100)	4 (100)	6 (100)
Tablet	5 (50)	3 (75)	2 (33.3)
Daily use of digital devices, n (%)			
Computer or laptop	7 (70)	4 (100)	3 (50)
Smartphone	10 (100)	4 (100)	6 (100)
Tablet	3 (30)	2 (50)	1 (16.7)
Need for assistance when using digital devices, n (%)			
Rarely	5 (50)	2 (50)	3 (50)
Sometimes	4 (40)	2 (50)	2 (33.3)
Daily	1 (10)	0 (0)	1 (16.7)

Phase 1 (Predesign): Supportive Care Pathway, Challenges, and Opportunities for Digital Health

Figure 2 shows the supportive care pathways at the University Hospital Zurich. Patients may access supportive care services through oncology nursing consultations, physician consultations, or direct contact. At the initial nursing consultation, supportive care needs are assessed using standardized paper-based PROMs, specifically the EORTC QLQ-C30 [8] and the NCCN Distress Thermometer [9]. Based on these assessments, nurses provide referrals to one or more of the 12 supportive care services offered, and patients additionally receive a supportive care brochure.

The predesign phase revealed 4 major themes regarding challenges and opportunities for digitalization in supportive cancer care at the University Hospital Zurich (Figure 3). Nurses, patients, and patient advocates highlighted the need for digitalization and standardization of supportive care, suggesting that digital platforms could replace paper-based PROMs, improve screening rates, and allow consultations to focus on discussing patient responses. As one cancer nurse explained:

> *At the moment, we give patients a paper-and-pencil version of the screening when they come to the hospital. […] So, the digital screening should actually improve this by allowing patients to answer in advance […]. This way, patients can complete it at home, and we can spend more time discussing their responses during the consultation.* [Cancer Nurse 2, focus group]

Communication challenges were also emphasized, including difficulties in reaching doctors, treatment-related fatigue, and language barriers. One patient advocate noted:

> *It's difficult [the communication with healthcare professionals] ... most of the time, my oncologist answers quite quickly. But sometimes he has too many things to do.* [Patient Advocate 3]

XSL•FO
RenderX

Patients reported information overload, describing the volume of paper-based brochures and consultations as overwhelming, and expressed a preference for simplified, trustworthy, and personalized digital resources.

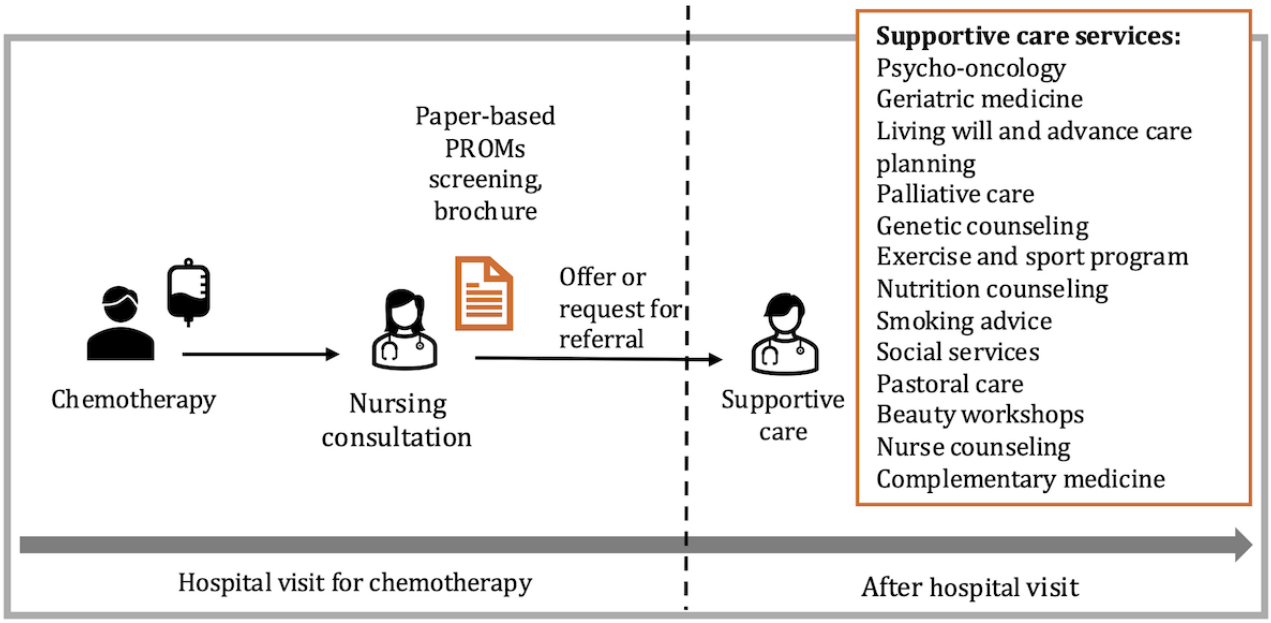
> *I have read a few things but, in the end, I was so overwhelmed with the conversations. Before each surgery, you have a conversation with the surgeon or the radiologist and then with the anesthetist and then again, a conversation and then you get a few more documents […]* [Cancer Patient 6]

Finally, the importance of connection and social support was highlighted. Patients valued peer support and digital opportunities to connect, but also emphasized that face-to-face consultations remain indispensable.

> *If I had an app for my illness, it would be good, but it would make everything so virtual, and you still need to connect with people in real life.* [Cancer Patient 4]

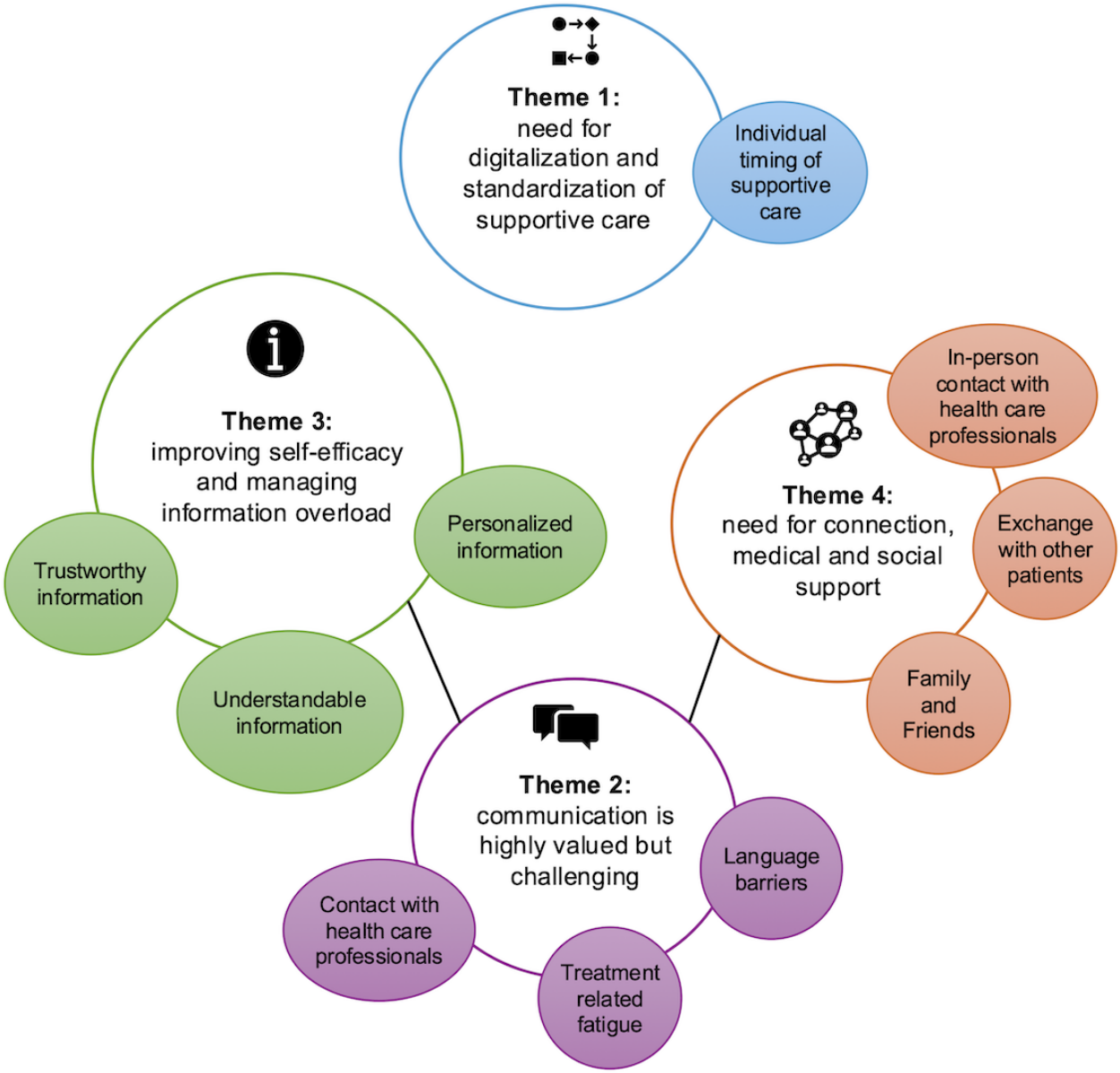
These themes and their relationships are summarized visually in Figure 3. More illustrative quotes can be found in Multimedia Appendix 1.

Figure 2. Mapping of supportive care pathways at the University Hospital Zurich.



XSL•FO
RenderX

Figure 3. Challenges and opportunities for digitalization in supportive cancer care at the University Hospital Zurich. Themes were derived from interviews and focus groups with patients with cancer, patient advocates, and oncology nurses, and informed subsequent app functionalities.



Phase 2 (Generative): Functionalities and Integration in the Supportive Care Pathway

Table 2 summarizes the functionalities of the digital health app, named OncoSupport+, that were either proposed by the research team or suggested by patients, mapped to the related themes identified in phase 1 (Figure 3). Functionalities developed by the research team were rated by patients on a scale from 1 to 10, where higher scores indicated greater perceived importance. Personalized information on supportive care (median 9.5) and digital PROMs screening (median 9.0) were rated as the most relevant features. Other functionalities, including self-help group information, a diary for consultations, and PROMs visualization, also received favorable ratings (all medians ≥7). Patients suggested several additional functionalities, including patient–health care professional chat (proposed by 6 participants), patient–patient chat (5 participants), self-management advice and symptom checking (5 participants), proactive contact from health care professionals (4 participants), reminders (4 participants), sharing data with family and friends (3 participants), information about cancer types (3 participants), patient stories (3 participants), and supportive care information for relatives (2 participants). Several of these patient-suggested features were integrated into the app (eg, reminders, patient stories, cancer type information, and peer support links), while others were not implemented due to scope limitations or regulatory considerations.

XSL•FO
RenderX

Table 2. Proposed functionalities for OncoSupport+, by source (research team vs patients), with brief descriptions, related themes identified in the predesign phase (Figure 3), ratings by patients and patient advocates.

Functionality	Description	Related theme (phase 1)	Median (range)	Frequency of suggestions,n	Included (yes/no)
Digital PROMs[a] screening[b]	Scheduled digital assessment of supportive care needs using standardized instruments: EORTC QLQ-C30[c] and NCCN[d] Distress thermometer (ePROMs[e]).	Digitalization and standardization; communication	9.0 (6-10)	—[f]	Yes
Personalized information on supportive care[b]	Personalized list of relevant supportive-care services and contact details generated from the patient profile and ePROMs results.	Information overload and self-efficacy	9.5 (7-10)	—	Yes
Self-help group information[b]	Link to verified cancer self-help and peer-support groups.	Information; connection, and peer support	7.5 (2-10)	—	Yes
Diary for consultations[b]	Free-text notes to record symptoms not captured by ePROMs and questions for nursing consultations.	Communication	7.5 (3-10)	—	Yes
PROMs visualization[b]	Dashboards for patients and cancer nurses displaying ePROMs and trends over time, with integration of diary notes.	Communication; self-efficacy	7.0 (2-10)	—	Yes
Patient–Health Care Provider chat[g]	Messaging between patients and cancer nurses.	Communication	—	6	No
Patient–patient chat[g]	Link to a peer-support platform provided by the cancer association.	Connection and peer support	—	5	Yes
Symptom checker and self-management advice[g]	Personalized symptom management advice based on reported symptoms and ePROMs results.	Self-efficacy	—	5	No
Proactive contact from health care provided (phone or email)[g]	Nurse-initiated outreach (phone or email) triggered by digital PROMs screening.	Communication	—	4	No
Appointment and PROMs reminders[g]	Automated reminders for clinic appointments and ePROMs completion by email, SMS, or mailed letter.	Self-efficacy	—	4	Yes
Sharing data with family and friends[g]	Patient-controlled sharing of selected PROMs summaries with family members and friends.	Social support	—	3	No
Information about cancer types[g]	Evidence-based educational content on cancer diagnoses, treatments, and common side effects.	Information overload; self-efficacy	—	3	Yes
Patient stories and experiences[g]	Positive stories and experiences from patients and survivors to support coping with cancer.	Emotional support and connection	—	3	Yes
Supportive care information for relatives[g]	Resources for relatives on available services and practical guidance.	Social support	—	2	No

[a]PROMs: patient-reported outcome measures.

[b]Research team suggestions.

[c]EORTC QLQ-C30: European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core 30.

[d]NCCN: National Comprehensive Cancer Network.

[e]ePROMs: electronic patient-reported outcome measures.

[f]Not available.

[g]Patient suggestions.

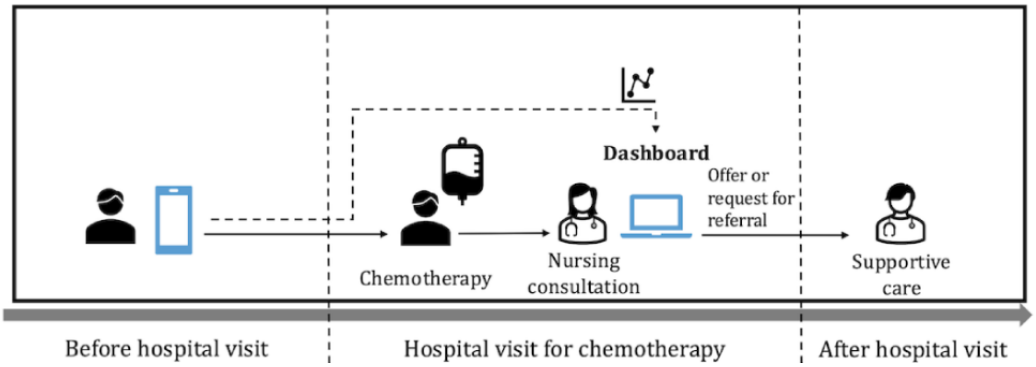
Figure 4 summarizes how OncoSupport+ and the included functionalities are integrated into the supportive care pathway. Patients complete digital PROMs screenings (Multimedia Appendix 2 lists the entire digital screening questionnaires, including the EORTC QLQ-C30 and the NCCN Distress Thermometer tool) in the week before their consultation, after which the app provides personalized supportive care information and other more general information, such as self-help groups and the peer-support platform. During the consultation, nurses access a dashboard displaying PROMs, diary notes, and

XSL•FO
RenderX

personalized information for supportive care and can recommend supportive care services. During the cancer treatment journey, patients use the app at least 3 times: at the beginning, halfway through treatment, and at the end.

To provide personalized supportive care information based on digital PROMs screening and the patient profile, a rule-based algorithm was developed in collaboration with supportive care specialists at the hospital. This algorithm prioritizes supportive care services and provides personalized information based on the symptoms and needs reported by patients and is based on medical knowledge and published scientific literature. For example, psycho-oncology is prioritized if a patient reports a distress score ≥5 on the NCCN Distress Thermometer and rates at least one emotional or social functioning item as "very much." A detailed description of the algorithm and its rules is provided in Multimedia Appendix 3.

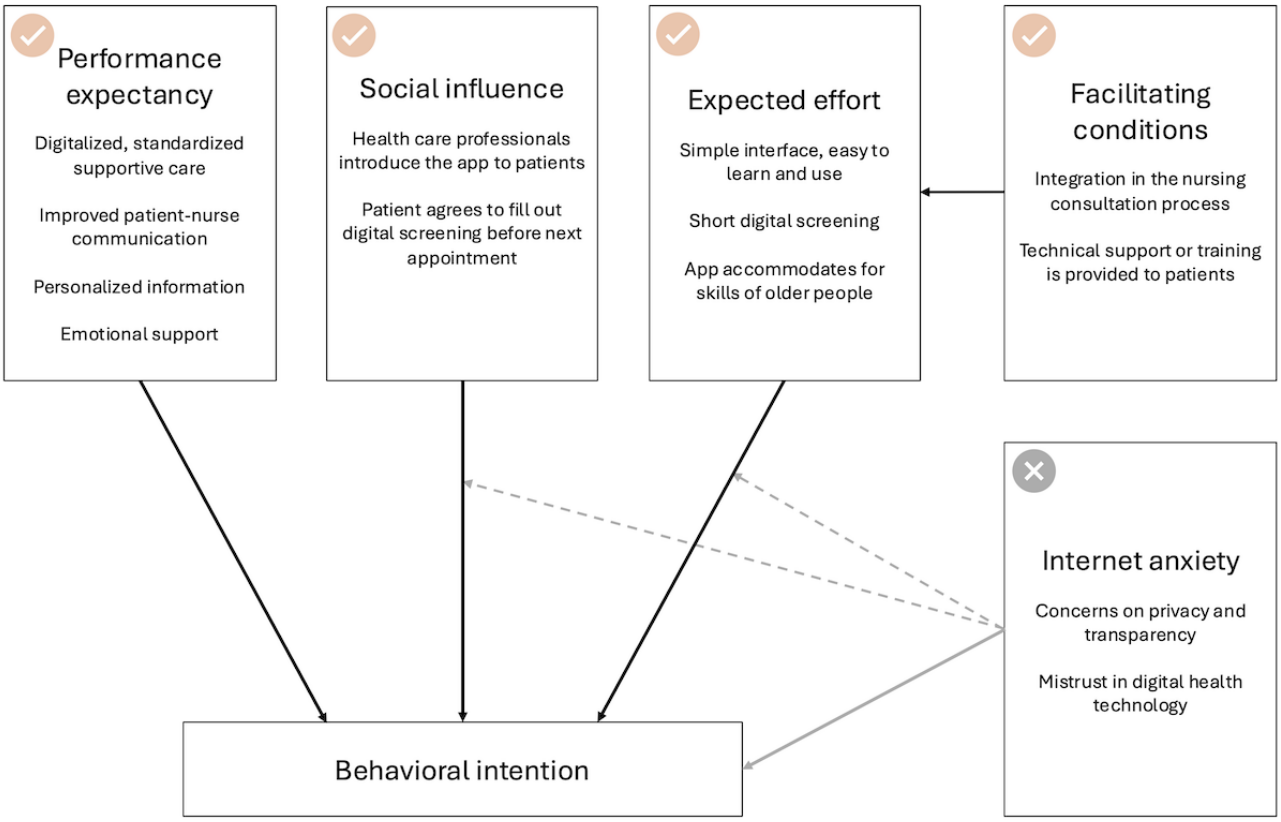
Figure 4. Integration of OncoSupport+ into the supportive care pathway.



Factors Potentially Impacting the Uptake of the Digital Health App OncoSupport+

Factors influencing the potential uptake of OncoSupport+ were mapped to the UTAUT framework (Figure 5; Multimedia Appendix 1 shows supporting quotes). Improved communication with health care professionals was the strongest driver of performance expectancy. Social influence was also important, as patients indicated they would be more likely to use the app if recommended during consultations. Effort expectancy was reflected in the preference for a simple, intuitive interface, particularly among older users or those with lower digital literacy. Facilitating conditions included integration into routine nursing consultations and the availability of technical support. Conversely, privacy concerns and internet anxiety emerged as barriers that may limit adoption.

Figure 5. Unified Theory of Acceptance and Use of Technology model illustrating factors affecting acceptance of OncoSupport+. Performance expectancy, social influence, and expected efforts positively correlate behavioral intention (black lines), whereas internet anxiety negatively correlates with behavioral intention (gray line) and moderates the relationships between social influence and expected effort (gray dotted line).



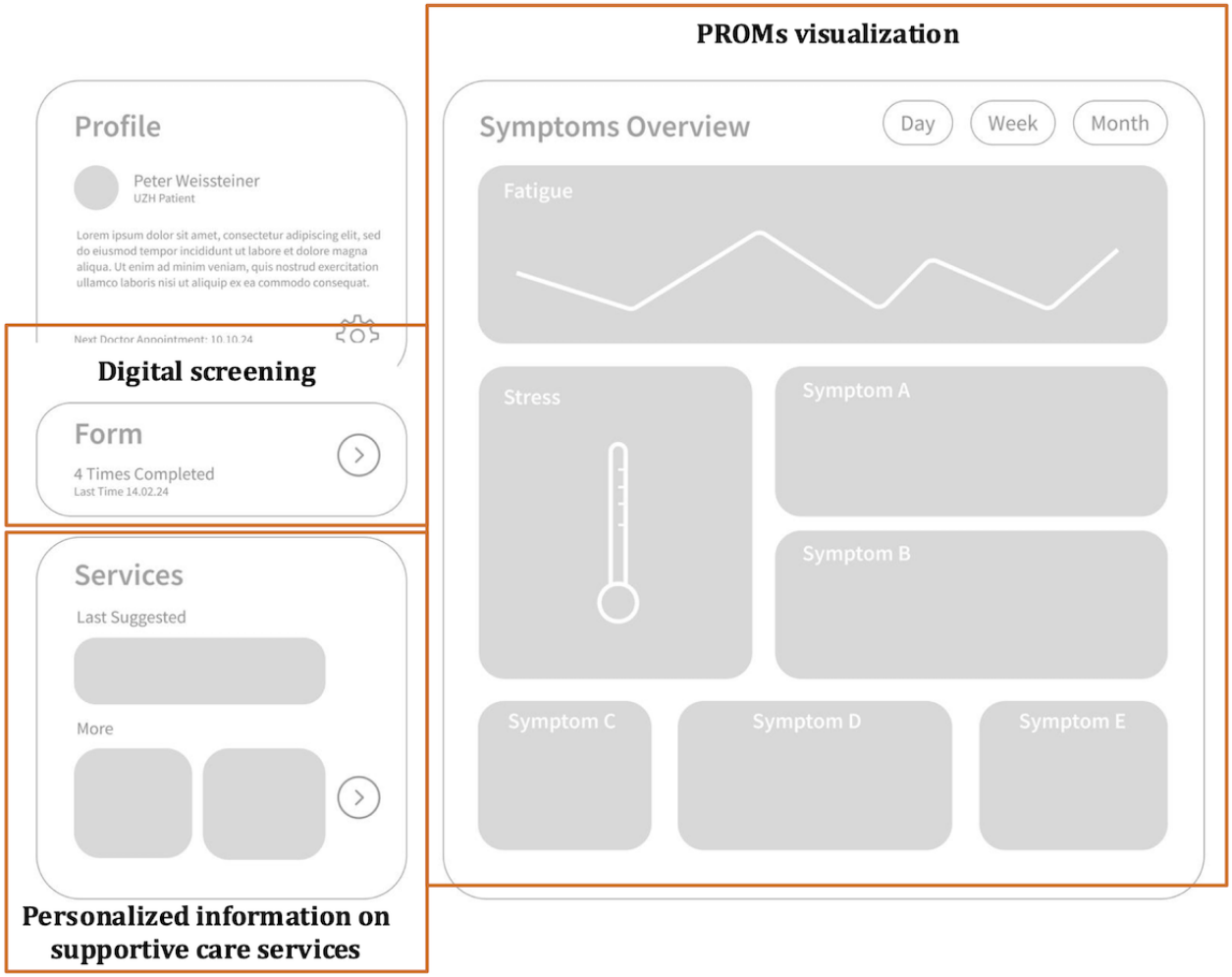
XSL•FO
RenderX

Prototyping

In the prototyping phase, the requirements identified in earlier stages were translated into user interface designs and iteratively refined with patient advocates and cancer nurses. A user flow was first created to map the patient and health care professional (ie, cancer nurses) journey through the app. Patients can self-register, whereas cancer nurse accounts, which allow access to multiple patient records, can only be created by the OncoSupport+ development team to ensure proper verification and compliance with data security protocols.

The first low-fidelity prototype (Figure 6) underwent usability testing, where some patient advocates expressed discomfort with frequently viewing symptom trajectories. This led to relocating the symptom overview to a dedicated section, while the digital screening and personalized information sections were emphasized.

Figure 6. First version of the low-fidelity prototype displaying the main functionalities: digital screening, personalized information on supportive care services, and patient-reported outcome measures (PROMs) visualization.

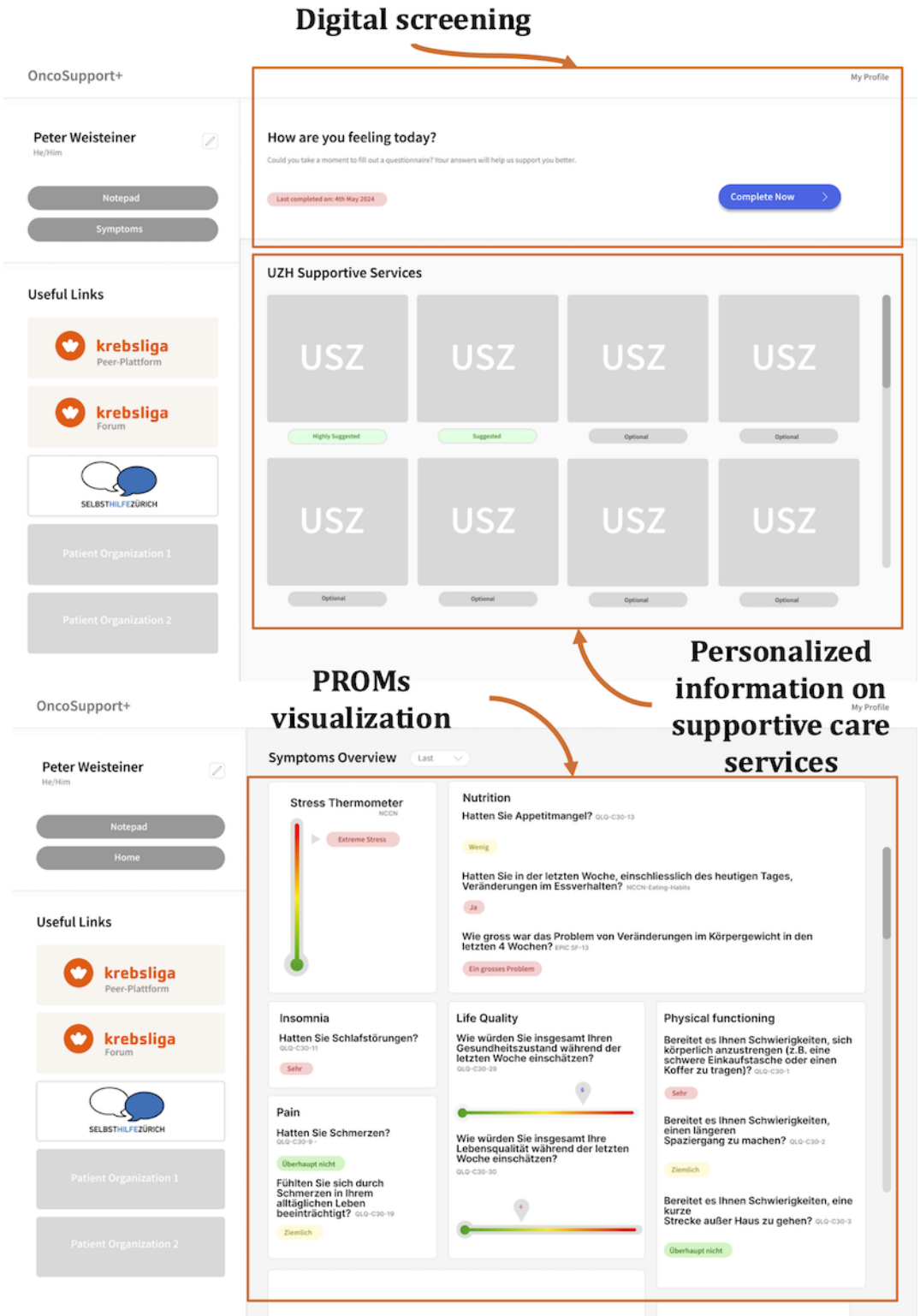


A second low-fidelity prototype (Figure 7) underwent further feedback from patients and one cancer nurse. The cancer nurse recommended replacing red and green color schemes with neutral tones and adding bar charts to facilitate intuitive interpretation of symptom trends. Patient advocates suggested replacing clinical images with neutral, calming visuals and simplifying terminology.

Based on this feedback, a high-fidelity prototype (Figure 8) with responsive design was developed, compatible with both smartphones and laptops. A landing page was introduced to increase transparency about data storage and use and to highlight the key benefits of the app, particularly its potential to enhance communication between patients and health care professionals. Calming images were also integrated to improve the user experience. More details about the IT infrastructure can be found in Multimedia Appendix 4.

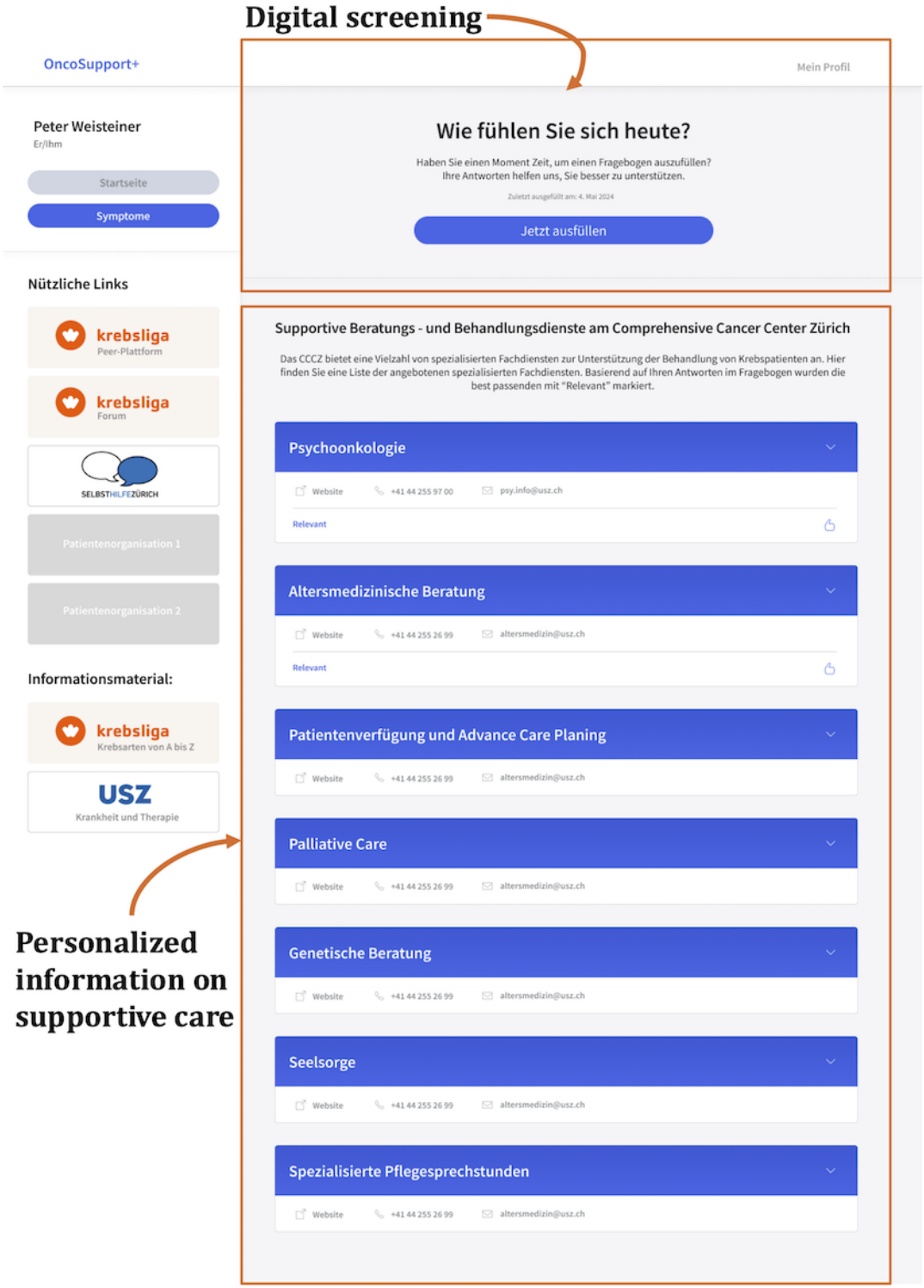
XSL•FO
RenderX

Figure 7. Second version of the low fidelity prototype displaying the main functionalities: digital screening, personalized information on supportive care services, and patient-reported outcome measures (PROMs) visualization.



XSL•FO
RenderX

Figure 8. High-fidelity prototype showing the digital screening functionality and the personalized information on supportive care functionality.



Discussion

Principal Results

This study reports the end-to-end co-design of OncoSupport+, a supportive cancer care app developed with patients with cancer and cancer nurses. The co-design approach identified challenges in supportive care and opportunities for digital health technology, generated a set of functionalities to address these challenges, and iteratively produced a prototype. As improving patient–health care professional communication and standardization and digitalization of supportive care were primary challenges, the OncoSupport+ app was designed to integrate a digital PROMs screening for supportive care using standardized oncology questionnaires (ie, EORTC QLQ-C30 and the NCCN Distress Thermometer) into the supportive care pathway. In addition, as enhancing self-efficacy, managing information overload, and receiving support from peers and

XSL•FO
RenderX

health care professionals appeared to be very important, OncoSupport+ prepares patients for in-person nursing consultations by providing personalized information, access to supportive care services, and links to self-help groups, patient-to-patient chats, and stories from other patients with cancer. These features may prepare patients effectively for consultations, thereby improving the overall quality of the interaction between patients and health care professionals and their sense of empowerment. While meeting the performance expectations of patients and health care professionals is a primary driver for the uptake of the technology [28], the development of an easy-to-use app, its introduction by health care professionals, and the provision of technical support may be critical for improving overall adoption. However, concerns related to interacting with a digital environment (eg, internet anxiety) may hinder this adoption. Future usability and feasibility studies will further evaluate the implementation feasibility, real-world usability, and acceptance of OncoSupport+ in clinical practice.

Comparison With Previous Work

The challenges identified in this study highlight the critical role of communication between patients and health care professionals, a domain where digital health technologies may play an important role. These technologies are not intended to replace health care professionals but rather to enhance and facilitate interactions with patients [32]. Consistent with previous research and systematic reviews [3], effective communication has been recognized as a key supportive care need. Not surprisingly, effective communication in cancer care has long been associated with improved medical and psychosocial outcomes, including reduced distress and increased satisfaction for both patients and clinicians [33]. Moreover, the European Society for Medical Oncology supports the integration of digital symptom monitoring, including the assessment of PROMs, as a vital element of cancer supportive care [34]. Evidence suggests that the use and digitalization of PROMs significantly enhance communication between patients and health care professionals [35]. For instance, a systematic review by Yang et al [35] suggests several mechanisms through which PROMs influence communication. These include increasing symptom awareness for both patients and health care professionals, prompting discussions (eg, addressing patient fears or uncertainty about disclosing symptoms), streamlining consultations (eg, using PROMs as a guide to direct health care professionals toward key patient-reported symptoms), and facilitating interprofessional communication between health care professionals (eg, enabling standardized patient tracking and optimizing the transfer of information). These mechanisms illustrate how PROMs may be a valuable tool for improving the quality of communication and, thereby, potentially enhancing clinical outcomes.

The findings of this study suggest that improving empowerment and self-efficacy in patients with cancer may be an important aspect of digital interventions in supportive care, consistent with previous research in this field. Recent systematic reviews [36,37] have indicated that digital health interventions, such as mobile apps and web-based platforms incorporating educational resources, interactive tools, medication reminders, and self-management features, may encourage patients to actively engage in their health care by enhancing empowerment and increasing self-efficacy. This enhanced sense of empowerment has been associated with better adherence to treatment regimens and more proactive health behaviors [37]. However, while digital health interventions that incorporate educational resources and self-management tools have been more consistently associated with increased self-efficacy, their direct impact on clinical outcomes, such as symptom reduction and improvements in quality of life, remains inconsistent across studies [38]. Although OncoSupport+ includes features related to informational and educational resources, appointment reminders, and tools to facilitate patient interaction, future research may explore the integration of self-management advice for symptom management to further enhance patient empowerment [39]. Additionally, while chat functions and telemedicine tools could further improve communication and self-efficacy, these features were not feasible for implementation in OncoSupport+. As patients value such tools, future research may focus on exploring the role of telemedicine solutions or artificial-intelligence–based chatbots in enhancing patient outcomes.

The digitalization of supportive care screenings and processes presents both significant opportunities and considerable challenges for hospitals. Digital transformation is reshaping health care by enhancing clinical workflows and promoting patient-centered care [40]. Implementing digital tools can improve efficiency and the quality of care, thereby reducing waste in health care costs. For instance, as reported by health care professionals interviewed in our participatory study, a separate study has shown that digital questionnaires significantly reduce the administrative burden on nurses compared to traditional paper-based questionnaires, allowing them to dedicate more time to patient care and resulting in substantial reductions in paper consumption [41]. Similarly, the adoption of electronic health record systems has been shown to facilitate better patient data management and streamline workflows. However, the transition to digital systems is not without challenges. Hospitals often face issues such as interoperability, strict data security regulations, the need for comprehensive staff training, and potential resistance to change among both health care professionals and patients [40]. This has been shown for cancer care as well [42]. Additionally, the lack of standardization across different IT systems at the hospital can hinder the seamless integration of digital tools. For example, the unsuccessful implementation of a nationwide digital patient dossier in Switzerland highlights the difficulties associated with a fragmented health care infrastructure and the absence of standardized frameworks [43]. Therefore, while digitalization offers promising benefits, careful consideration and strategic planning are essential to effectively navigate the associated challenges.

This study explored factors influencing the future acceptance of digital health apps like OncoSupport+, aligning with the UTAUT framework, which has been validated using quantitative data from digital health interventions [28]. An interesting finding concerns internet anxiety and the observation that age does not appear to be a barrier to technology acceptance. Internet anxiety,

XSL•FO
RenderX

defined as the fear or discomfort individuals experience when using online technologies, can hinder the adoption of digital health solutions [28]. Previous research indicates that while age is not a significant predictor of health technology acceptance [44], it correlates with internet anxiety—an important factor in adoption. Moreover, individuals with lower digital literacy [45] tend to exhibit higher levels of internet anxiety. Although these results appear to be in line with previous literature, a larger and more representative sample may confirm the role of internet anxiety in relation to digital health technology acceptance in patients with cancer. However, providing adequate training and support can help alleviate these concerns, thereby enhancing the adoption of digital health technologies.

Strengths and Limitations

This study has both strengths and limitations that should be considered when interpreting the results. The relatively small sample size (N=10 patients with cancer and patient advocates) may limit generalizability. However, the principle of thematic saturation was applied to ensure that no new themes emerged, supporting the adequacy of the sample. In addition, the inclusion of a significant number of female participants may account for gender-related biases in digital health acceptance. The patient group may not fully represent the broader population of patients with cancer, as most participants were aged below the average age of people with cancer and had a tertiary education level, higher than the Swiss national average. As education level correlates with digital literacy and some participants were highly experienced with digital health tools, acceptance of OncoSupport+ may differ in populations with lower digital experience. While individual interviews provided valuable insights, future research may benefit from workshops involving different user groups (eg, patients, nurses, family members, and caregivers) to foster mutual understanding and refine the co-design process. Although this study primarily examined user acceptance using the UTAUT framework, future implementation studies may apply frameworks such as Nonadoption, Abandonment, Scale-up, Spread, and Sustainability framework [46] or the Consolidated Framework for Implementation Research [47] to explore technology, organizational, and system-level factors affecting implementation and adoption. A key strength of this study is the direct integration of perspectives of patients, ensuring that OncoSupport+ is aligned with user needs and expectations, and the use of a structured co-design framework to guide development.

Conclusion

This study highlights the feasibility and importance of collaborative design in developing digital health applications and interventions, emphasizing the need for direct involvement of patients and health care professionals in the development process. The findings suggest that digital interventions such as OncoSupport+ may include features and functionalities to enhance patient–health care professional communication, facilitate the digitalization of supportive care screenings, and support patient empowerment and self-efficacy by providing personalized information on supportive care services, promoting connection with other patients, and offering self-management advice. However, while OncoSupport+ appears to meet the performance expectations of both patients and health care professionals, future research is needed to evaluate its implementation feasibility and determine whether it can effectively improve communication, enhance self-efficacy, and ultimately lead to better health outcomes, such as improved quality of life.

Acknowledgments

The authors thank Helén Johanna Luedtke for her assistance with the recruitment phase at the Department of Oncology and Hematology of the Zurich University Hospital and Bruno Ceolin Argenta for his help in creating Figma prototypes in the prototyping phase. We also thank the supportive care specialists working at the Zurich University Hospital who participated in our workshop and contributed to creating the knowledge-based algorithm: Maja Zenz, PhD; Prof Dr med David Blum; Christina Gassmann; Nicola Greco; and Anastasios Manettas. This study was supported by the Digitalization Initiative of the Zurich Higher Education Institutions (DIZH).

This study used OpenAI Whisper AI to transcribe qualitative data. To ensure accuracy, researchers manually reviewed and verified transcripts by cross-checking them with the original audio recordings. Additionally, generative AI–assisted tools were used solely to support grammar correction and language refinement, with the goal of improving readability. No content, data interpretation, or results were generated by AI. All final decisions regarding the wording, content, interpretation of results, and conclusions were made by the research team. The authors take full responsibility for the integrity and accuracy of the work presented.

The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

The dataset generated and analyzed during this study is not publicly available because it is being used for the ongoing development of the digital health app OncoSupport+ at the University Hospital Zurich, but it is available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: CMW, MK, AW, AG, SD

XSL•FO
RenderX

Data curation: SD, MMB, CL
Formal analysis: SD, MMB
Funding acquisition: KM, CMW, KS, SH, PA, SM, AW, AG
Investigation: SD, MMB
Methodology: CMW, SD
Project administration: SD
Software: KS, CL, SH, YZ
Visualization: SD, CL
Writing – original draft: SD
Writing – review & editing: MMB, CL, KS, CMW, MK, AG, AW, SH, PA, SM, MR

Conflicts of Interest

CMW has active research grants to the University for Digital Health Projects from the DIZH, the Swiss Cancer Research Foundation, the German Health Care Innovation Fund, and Newsense Lab GmbH and has received honoraria from Swiss hospitals for scientific presentations on the digitalization and AI in medicine. MK is cofounder and AI advisor at Oncobit but has no conflicts of interest related to this study. The remaining coauthors declare no conflicts of interest.

Multimedia Appendix 1
Thematic analysis and results.
[DOCX File , 4547 KB - humanfactors_v12i1e73829_app1.docx]

Multimedia Appendix 2
Digital screening questionnaires.
[DOCX File , 23 KB - humanfactors_v12i1e73829_app2.docx]

Multimedia Appendix 3
Knowledge-based algorithm to connect patient-reported outcome measures to supportive care services.
[DOCX File , 79 KB - humanfactors_v12i1e73829_app3.docx]

Multimedia Appendix 4
IT Infrastructure.
[DOCX File , 3734 KB - humanfactors_v12i1e73829_app4.docx]

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74(3):229-263 [FREE Full text] [doi: 10.3322/caac.21834] [Medline: 38572751]
2. De Angelis R, Demuru E, Baili P, Troussard X, Katalinic A, Chirlaque Lopez MD, EUROCARE-6 Working Group. Complete cancer prevalence in Europe in 2020 by disease duration and country (EUROCARE-6): a population-based study. Lancet Oncol 2024;25(3):293-307. [doi: 10.1016/S1470-2045(23)00646-0] [Medline: 38307102]
3. Evans Webb M, Murray E, Younger ZW, Goodfellow H, Ross J. The supportive care needs of cancer patients: a systematic review. J Cancer Educ 2021;36(5):899-908 [FREE Full text] [doi: 10.1007/s13187-020-01941-9] [Medline: 33492650]
4. Scotté F, Taylor A, Davies A. Supportive care: The "Keystone" of modern oncology practice. Cancers (Basel) 2023;15(15):3860 [FREE Full text] [doi: 10.3390/cancers15153860] [Medline: 37568675]
5. Keane D, Phillips G, Mitchell N, Connolly RM, Hegarty J. Improving quality of life and symptom experience in patients with metastatic breast cancer: a systematic review of supportive care interventions. Psychooncology 2023;32(8):1192-1207. [doi: 10.1002/pon.6183] [Medline: 37434307]
6. Basch E, Deal AM, Dueck AC, Scher HI, Kris MG, Hudis C, et al. Overall survival results of a trial assessing patient-reported outcomes for symptom monitoring during routine cancer treatment. JAMA 2017;318(2):197-198 [FREE Full text] [doi: 10.1001/jama.2017.7156] [Medline: 28586821]
7. Chan RJ, Knowles R, Ashbury FD, Bowen J, Chan A, Chin M, et al. Supportive care 2030 movement: towards unifying ambitions for global excellence in supportive cancer care – an international Delphi study. EClinicalMedicine 2024;76:102825 [FREE Full text] [doi: 10.1016/j.eclinm.2024.102825] [Medline: 39309723]
8. Fayers P, Bottomley A, EORTC Quality of Life Group, Quality of Life Unit. Quality of life research within the EORTC—the EORTC QLQ-C30. European Organisation for Research and Treatment of Cancer. Eur J Cancer 2002;38 Suppl 4:S125-S133. [doi: 10.1016/s0959-8049(01)00448-8] [Medline: 11858978]
9. Donovan KA, Grassi L, McGinty HL, Jacobsen PB. Validation of the distress thermometer worldwide: state of the science. Psychooncology 2014;23(3):241-250. [doi: 10.1002/pon.3430] [Medline: 25160838]

XSL•FO
RenderX

10. Aapro M, Bossi P, Dasari A, Fallowfield L, Gascón P, Geller M, et al. Digital health for optimal supportive care in oncology: benefits, limits, and future perspectives. Support Care Cancer 2020;28(10):4589-4612 [FREE Full text] [doi: 10.1007/s00520-020-05539-1] [Medline: 32533435]
11. Shaffer KM, Turner KL, Siwik C, Gonzalez BD, Upasani R, Glazer JV, et al. Digital health and telehealth in cancer care: a scoping review of reviews. Lancet Digit Health 2023;5(5):e316-e327 [FREE Full text] [doi: 10.1016/S2589-7500(23)00049-3] [Medline: 37100545]
12. Vaffis S, Whaley S, Axon DR, Hall-Lipsy E, Hincapie A, Slack M, et al. Features of cancer mHealth apps and evidence for patient preferences: scoping literature review. JMIR Cancer 2023;9:e37330 [FREE Full text] [doi: 10.2196/37330] [Medline: 37115587]
13. Meirte J, Hellemans N, Anthonissen M, Denteneer L, Maertens K, Moortgat P, et al. Benefits and disadvantages of electronic patient-reported outcome measures: systematic review. JMIR Perioper Med 2020;3(1):e15588 [FREE Full text] [doi: 10.2196/15588] [Medline: 33393920]
14. Mowlem FD, Elash CA, Dumais KM, Haenel E, O'Donohoe P, Olt J, Electronic Clinical Outcome Assessment Consortium. Best practices for the electronic implementation and migration of patient-reported outcome measures. Value Health 2024;27(1):79-94 [FREE Full text] [doi: 10.1016/j.jval.2023.10.007] [Medline: 37879401]
15. Fournier V, Duprez C, Grynberg D, Antoine P, Lamore K. Are digital health interventions valuable to support patients with cancer and caregivers? An umbrella review of web-based and app-based supportive care interventions. Cancer Med 2023;12(23):21436-21451 [FREE Full text] [doi: 10.1002/cam4.6695] [Medline: 37937812]
16. Parikh RB, Basen-Enquist KM, Bradley C, Estrin D, Levy M, Lichtenfeld JL, et al. Digital health applications in oncology: an opportunity to seize. J Natl Cancer Inst 2022;114(10):1338-1339 [FREE Full text] [doi: 10.1093/jnci/djac108] [Medline: 35640986]
17. Lazarou I, Krooupa A, Nikolopoulos S, Apostolidis L, Sarris N, Papadopoulos S, et al. Cancer patients' perspectives and requirements of digital health technologies: a scoping literature review. Cancers (Basel) 2024;16(13):2293 [FREE Full text] [doi: 10.3390/cancers16132293] [Medline: 39001356]
18. Noorbergen TJ, Adam MTP, Teubner T, Collins CE. Using co-design in mobile health system development: a qualitative study with experts in co-design and mobile health system development. JMIR Mhealth Uhealth 2021;9(11):e27896 [FREE Full text] [doi: 10.2196/27896] [Medline: 34757323]
19. Peters S, Guccione L, Francis J, Best S, Tavender E, Curran J, et al. Evaluation of research co-design in health: a systematic overview of reviews and development of a framework. Implementation Sci 2024;19(1):63 [FREE Full text] [doi: 10.1186/s13012-024-01394-4] [Medline: 39261956]
20. Digital Health Zurich. URL: https://www.digitalhealthzurich.com/en [accessed 2025-02-22]
21. Sanders EBN, Stappers PJ. Co-creation and the new landscapes of design. CoDesign 2008;4(1):5-18. [doi: 10.1080/15710880701875068]
22. Confluence. URL: https://www.atlassian.com/software/confluence [accessed 2025-02-22]
23. Figma Design. URL: https://www.figma.com [accessed 2025-02-04]
24. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2008;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
25. OpenAI Whisper. URL: https://openai.com/index/whisper/ [accessed 2025-02-27]
26. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425-478. [doi: 10.2307/30036540]
27. AlQudah AA, Al-Emran M, Shaalan K. Technology acceptance in healthcare: a systematic review. Appl Sci 2021;11(22):10537. [doi: 10.3390/app112210537]
28. Philippi P, Baumeister H, Apolinário-Hagen J, Ebert DD, Hennemann S, Kott L, et al. Acceptance towards digital health interventions - model validation and further development of the unified theory of acceptance and use of technology. Internet Interv 2021;26:100459 [FREE Full text] [doi: 10.1016/j.invent.2021.100459] [Medline: 34603973]
29. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907 [FREE Full text] [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
30. Kilfoy A, Hsu TC, Stockton-Powdrell C, Whelan P, Chu CH, Jibb L. An umbrella review on how digital health intervention co-design is conducted and described. NPJ Digit Med 2024;7(1):374 [FREE Full text] [doi: 10.1038/s41746-024-01385-1] [Medline: 39715947]
31. Arcia A, Stonbraker S, Mangal S, Lor M. A practical guide to participatory design sessions for the development of information visualizations: tutorial. J Particip Med 2024;16:e64508 [FREE Full text] [doi: 10.2196/64508] [Medline: 39671555]
32. Hong YA, Hossain MM, Chou WS. Digital interventions to facilitate patient-provider communication in cancer care: a systematic review. Psychooncology 2020;29(4):591-603. [doi: 10.1002/pon.5310] [Medline: 31834650]
33. Fallowfield L, Jenkins V. Effective communication skills are the key to good cancer care. Eur J Cancer 1999;35(11):1592-1597. [doi: 10.1016/s0959-8049(99)00212-9] [Medline: 10673967]
34. Di Maio M, Basch E, Denis F, Fallowfield L, Ganz P, Howell D, ESMO Guidelines Committee. Electronic address: clinicalguidelines@esmo.org. The role of patient-reported outcome measures in the continuum of cancer clinical care:

XSL•FO
RenderX

ESMO Clinical Practice Guideline. Ann Oncol 2022;33(9):878-892 [FREE Full text] [doi: 10.1016/j.annonc.2022.04.007] [Medline: 35462007]
35. Yang LY, Manhas DS, Howard AF, Olson RA. Patient-reported outcome use in oncology: a systematic review of the impact on patient-clinician communication. Support Care Cancer 2018;26(1):41-60. [doi: 10.1007/s00520-017-3865-7] [Medline: 28849277]
36. Tuominen L, Leino-Kilpi H, Poraharju J, Cabutto D, Carrion C, Lehtiö L, et al. Interactive digital tools to support empowerment of people with cancer: a systematic literature review. Support Care Cancer 2024;32(6):396 [FREE Full text] [doi: 10.1007/s00520-024-08545-9] [Medline: 38816629]
37. Ancona C, Caroppo E, De Lellis P. Digital solutions supporting the quality-of-life of European cancer patients and their caregivers: a systematic literature review. Health Technol 2025;15:243-272. [doi: 10.1007/s12553-025-00942-2]
38. Fournier V, Duprez C, Grynberg D, Antoine P, Lamore K. Are digital health interventions valuable to support patients with cancer and caregivers? An umbrella review of web-based and app-based supportive care interventions. Cancer Med 2023;12(23):21436-21451 [FREE Full text] [doi: 10.1002/cam4.6695] [Medline: 37937812]
39. Lim DSC, Kwok B, Williams P, Koczwara B. The impact of digital technology on self-management in cancer: systematic review. JMIR Cancer 2023;9:e45145 [FREE Full text] [doi: 10.2196/45145] [Medline: 37991831]
40. Stoumpos AI, Kitsios F, Talias MA. Digital transformation in healthcare: technology acceptance and its applications. Int J Environ Res Public Health 2023;20(4):3407 [FREE Full text] [doi: 10.3390/ijerph20043407] [Medline: 36834105]
41. Kodama K, Konishi S, Manabe S, Okada K, Yamaguchi J, Wada S, et al. Impact of an electronic medical record-connected questionnaire on efficient nursing documentation: usability and efficacy study. JMIR Nurs 2023;6:e51303 [FREE Full text] [doi: 10.2196/51303] [Medline: 37634203]
42. Ardito V, Golubev G, Ciani O, Tarricone R. Evaluating barriers and facilitators to the uptake of mHealth apps in cancer care using the consolidated framework for implementation research: scoping literature review. JMIR Cancer 2023;9:e42092 [FREE Full text] [doi: 10.2196/42092] [Medline: 36995750]
43. Elektronisches Patientendossier. Was gibt es noch zu tun bis zu seiner flächendeckenden Verwendung?. URL: https://www.parlament.ch/centers/eparl/curia/2018/20184328/Bericht%20BR%20D.pdf [accessed 2025-02-04]
44. Jokisch MR, Schmidt LI, Doh M. Acceptance of digital health services among older adults: findings on perceived usefulness, self-efficacy, privacy concerns, ICT knowledge, and support seeking. Front Public Health 2022;10:1073756 [FREE Full text] [doi: 10.3389/fpubh.2022.1073756] [Medline: 36582385]
45. Melhem SJ, Nabhani-Gebara S, Kayyali R. Digital trends, digital literacy, and E-Health engagement predictors of breast and colorectal cancer survivors: a population-based cross-sectional survey. Int J Environ Res Public Health 2023;20(2):1472 [FREE Full text] [doi: 10.3390/ijerph20021472] [Medline: 36674237]
46. Greenhalgh T, Abimbola S. The NASSS framework - A synthesis of multiple theories of technology implementation. Stud Health Technol Inform 2019;263:193-204. [doi: 10.3233/SHTI190123] [Medline: 31411163]
47. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated consolidated framework for implementation research based on user feedback. Implement Sci 2022;17(1):75 [FREE Full text] [doi: 10.1186/s13012-022-01245-0] [Medline: 36309746]

Abbreviations

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core 30
ePROM: electronic patient-reported outcome measure
mHealth: mobile health
NCCN: National Comprehensive Cancer Network
PROM: patient-reported outcome measure
SAKK: Swiss Group for Clinical Cancer Research
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by A Kushniruk; submitted 14.03.25; peer-reviewed by M Imperio, A Castro, V Schadewaldt; comments to author 27.08.25; revised version received 26.09.25; accepted 16.10.25; published 11.11.25.

Please cite as:
Difrancesco S, Bauert MM, Lehmann C, Häsler S, Zhang Y, Hirsch S, Ackermann P, Stockinger K, Reif M, Mathieu S, Götz A, Wicki A, Krauthammer M, Witt CM
Collaborative Design and Development of a Patient-Centered Digital Health App for Supportive Cancer Care: Participatory Study
JMIR Hum Factors 2025;12:e73829
URL: https://humanfactors.jmir.org/2025/1/e73829
doi:10.2196/73829
PMID:41217811

XSL•FO
RenderX

©Sonia Difrancesco, Matthia Martina Bauert, Claude Lehmann, Steven Häsler, Yi Zhang, Sven Hirsch, Philipp Ackermann, Kurt Stockinger, Monika Reif, Sunjoy Mathieu, Anna Götz, Andreas Wicki, Michael Krauthammer, Claudia M Witt. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 11.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Key Factors Shaping Successful Implementation of the Internet of Things (IoT) in Health Care: Qualitative Study

Klas Palm[1*], PhD; Carl Kronlid[1*], PhD; Marie Elf[2*], Prof Dr; Anders Brantnell[1*], PhD

[1]Department for Industrial Engineering and Management, Uppsala University, Lägerhyddsvägen 1, Uppsala, Sweden

[2]School of Health and Welfare, Dalarna University, Falun, Sweden

[*]all authors contributed equally

Corresponding Author:
Klas Palm, PhD
Department for Industrial Engineering and Management, Uppsala University, Lägerhyddsvägen 1, Uppsala, Sweden

Abstract

Background: The utilization of the Internet of Things (IoT) can significantly enhance health care. However, successful implementation of IoT requires a holistic approach including factors beyond technology alone.

Objective: This paper seeks to advance understanding of the factors influencing the successful implementation of IoT solutions in the health care sector, expanding beyond a purely technological focus.

Methods: Using data from 22 semistructured interviews with a diverse group of stakeholders—including health care professionals, researchers, municipal and regional officials, and private companies—this study examines 5 leading IoT projects in Sweden.

Results: Grounded in sociotechnical systems theory, the research identifies five critical subsystems impacting IoT implementation: (1) laws and regulations, which present challenges due to their complexity and misalignment with rapid technological advances; (2) organizational support, highlighting the essential commitment and resources from management to drive innovation; (3) user focus, emphasizing the importance of engaging end-users—such as patients and health care providers—in the design and implementation of IoT solutions; (4) resources, encompassing both financial investments and human capital needed for effective deployment; and (5) infrastructure, which addresses the technological foundations required to support IoT systems reliably.

Conclusions: By shifting attention from adoption to the complexities of implementation, this study fills a critical gap in the literature, which has largely emphasized adoption and technical aspects over practical implementation challenges. The findings provide a nuanced understanding of the primary factors influencing IoT implementation in health care, illuminating both the challenges and potential avenues for successful integration. Ultimately, this research advances the sociotechnical systems theory and also offers valuable insights for managers and policymakers tasked with driving digital transformation in health care systems.

International Registered Report Identifier (IRRID): RR2-10.2196/44562

(JMIR Hum Factors 2025;12:e71546) doi:10.2196/71546

KEYWORDS

IoT; Internet of Things; implementation; health care; system; sociotechnical systems; case study

Introduction

Background

The Internet of Things (IoT) and technologies, such as artificial intelligence, big data, and cloud computing [1], can significantly enhance the quality of health care [2]. IoT is a network of intelligent sensing devices and physical objects that are digitally connected for the collection, monitoring, and control of health care data [3]. It enables tracking, identification, authentication, data collection, and sensing [4]. These features can lead to improved patient outcomes, increased efficiency, remote monitoring, and enhanced collaboration between health care providers [5]. They are essential to addressing rising demands, such as an aging population, an increase in long-term conditions, limited resources, and growing expectations for personalized health care [6].

However, in exploring IoT's potential in health care, it is essential to distinguish between adoption and implementation. Adoption refers to the decision to use an innovation, while implementation involves the actual use of that innovation [7]. This distinction is often overlooked, as the 2 terms frequently used interchangeably or without clear differentiation [2,8-11]. IoT adoption and its influencing factors have been widely researched, typically focusing on technology [12-14], but also covering aspects related to individuals and the surrounding environment [15]. Adoption studies typically draw on frameworks such as the technology acceptance model (TAM) [16,17] and the Unified Theory of Acceptance and Use of Technology (UTAUT) [18,19].

XSL•FO
RenderX

A recent sociotechnical analysis of factors influencing IoT adoption in health care, based on a literature review and carried out by the Swedish researchers; Kronlid et al [20] identified a broad array of factors, 94 in total, that impact IoT adoption in health care. The comprehensive literature review by Kronlid et al serves as a starting point for this article, as it responds to previous calls for a deeper sociotechnical understanding of IoT adoption within health care and proposes a specialized sociotechnical systems framework tailored specifically to IoT adoption in health care settings. The review found no studies on the actual implementation of IoT in health care and only a few focusing on IoT adoption. Kronlid et al call for more rigorous empirical research on IoT adoption and implementation, grounded in frameworks that reflect sociotechnical realities.

Other previous research has shown that successful implementation requires consideration of factors beyond technology [21]. As early as 1951, Trist and Bamforth [22] emphasized the importance of social factors in realizing the full potential of technological innovations. Earlier research on implementing medical innovations shows that influencing factors extend beyond technical considerations to include managerial, clinical, financial, legal, and individual elements [23,24]. However, research exploring IoT implementation in health care through a sociotechnology lens remains scarce [20].

Rogers [7] defines adoption as "a decision to make full use of an innovation as the best course of action available," emphasizing the importance of the decision-making process. Research on IoT adoption in health care began to gain traction around 2013, with a predominant focus on technological aspects [25]. Dantu et al [25] identified 5 main topic clusters for adoption: security and privacy, cloud and smart health, wireless network technologies, data, and applications. IoT solutions in health care rely heavily on the transfer of large volumes of data (ie, big data) and device-to-device communication, making high-speed internet a crucial requirement [26]. A review of studies on IoT in elderly care also highlighted bandwidth and network connectivity (eg, WiFi) as essential preconditions for IoT adoption [27]. In addition, cloud services play a significant role in IoT adoption, providing big data storage and access, often described as the "backbone" of IoT solutions [12]. Several studies have explored the benefits and risks associated with cloud services [13,14,28-31].

The literature addressing nontechnical aspects of IoT adoption in health care is relatively limited, and the lack of focus on the social dimensions of IoT adoption has been noted by several researchers (eg, [8,25]. In a review of IoT adoption in health care [15], categorizing the factors influencing IoT adoption into 5 groups: individual factors (eg, social influence), technological factors (eg, perceived ease of use), security factors (eg, trust), health factors (eg, perceived health risk), and environmental factors (eg, financial cost). Many of the included studies applied the TAM or the UTAUT1 and UTAUT2 [15].

An interesting finding is that Alraja [32] found that gender influences the intention to use IoT technology among Generation Y, but not among Generation X. Furthermore, concerns about privacy risks did not impact the decision to use IoT technology. Additionally, Huarng et al [33] developed a model for the adoption of health care wearable devices, identifying data privacy, perceived ease of use, and perceived usefulness as key influencing factors. The review by Kronlid et al [20], which found 94 factors influencing IoT adoption, was grouped into eight subthemes: (1) financial circumstances (eg, reimbursement), (2) infrastructure (eg, data management), (3) people (eg, attitudes and feelings), (4) procedures (eg, capability development), (5) regulatory frameworks (eg, security and privacy), (6) stakeholders (eg, commercialization challenges), (7) technology (eg, complexity), and (8) other factors (eg, conflicts of interest). Of these, 34 factors were related to technology and infrastructure, accounting for approximately 42% of the total. The categories people and regulatory frameworks each counted for around 20%. Thus, this highlights technology, infrastructure, people, and regulatory frameworks as the most important areas. Notably, many of these factors were identified from research in low- and middle-income countries [20], raising questions about their relevance in high-income settings.

IoT Implementation in Health Care

Implementation occurs when an individual or organization actively begins using an innovation [7] and can be understood as "a deliberately initiated process, in which agents aim to bring into operation new or modified practices that are institutionally sanctioned and performed by themselves and other agents" [34]. A recent review did not find any studies specifically addressing IoT device implementation [20]. Moreover, distinctions between adoption and implementation are often blurred or unclear in the existing literature [20]. To gain insights into the range of nontechnical factors influencing IoT implementation, it is useful to consider findings from related fields, such as telemedicine, where technology integration in health care shares similarities with IoT [3]. In telemedicine research, the difference between adoption and implementation is more distinctly delineated; see, for example, Broens et al [35].

Telemedicine implementation is a well-established research area, with several reviews summarizing key findings. Broens et al [35] identified 5 primary factors influencing telemedicine implementation: technology (eg, quality), user acceptance (eg, attitudes), financial considerations (eg, reimbursement structure), organizational aspects (eg, work practices), and policy and legislation (eg, standardization). Scott Kruse et al [23] identified 5 common barriers to telemedicine implementation globally: technically challenged staff, resistance to change, cost, reimbursement issues, and demographic factors such as patient age and education. Recent studies validate these findings; for example, Khodadad-Saryazdi [24] emphasized that successful implementation requires adaptability in technology, strategy, and culture. Similarly, Cannavacciuolo et al [36] affirmed the 5 factors initially outlined by Broens et al [35]. Overall, successful implementation is a mix of technical and nontechnical factors.

Most previous empirical research on IoT implementation in health care is quantitative, using surveys to collect data [2,8,9,15] combined with adoption theories such as TAM or UTAUT. While useful, there are drawbacks that we have identified in previous research. The combination of quantitative

XSL•FO
RenderX

methodologies and adoption theories limits the research, as it lacks the capacity to include unexpected aspects of IoT implementation. We thus see the need to broaden the literature both in terms of methodology and theory: methodologically, with more qualitative studies that have more flexibility to let the data guide the research, and, theoretically, to identify and make sense of, for example, organizational and contextual factors that can play a major role in the implementation of IoT.

In summary, research demonstrates the importance of considering many factors when implementing new health care technologies. This highlights the need for more empirical studies on IoT implementation that draw on a sociotechnical systems framework to account for the complex interplay of these factors.

The Sociotechnical Systems Model

This study builds upon the sociotechnical systems framework proposed by Kronlid et al [20], which categorized key factors influencing IoT adoption in health care settings (see Figure 1). The framework is informed by and aligns with implementation research in related fields, such as telemedicine [23,35], thus providing a solid foundation for analyzing the multifaceted challenges and enablers associated with IoT implementation in health care.

The sociotechnical systems model, grounded in the work of Leavitt [37], highlights that any system is composed of multiple, interdependent subsystems. Changes in one subsystem, such as technology, will invariably impact others, including organizational culture, infrastructure, or processes. A more modern version of this approach, developed by Davis et al [38], outlines 6 key subsystems within an organizational system: goals, technology, culture, people, infrastructure, and processes. These subsystems interact with an external environment that includes regulatory contexts, financial factors, and other stakeholders. While this model has proven useful in many areas, Kronlid et al [20] propose that specific modifications might be needed in health care. Their review showed that the "culture" and "goals" subsystems of the model have not been well studied in health care research.

This study investigates the key factors influencing IoT implementation in health care, based on empirical evidence from 5 state-of-the-art IoT projects in Sweden. These projects were pilot projects; the focus of the study was on the early stages of the implementation process.

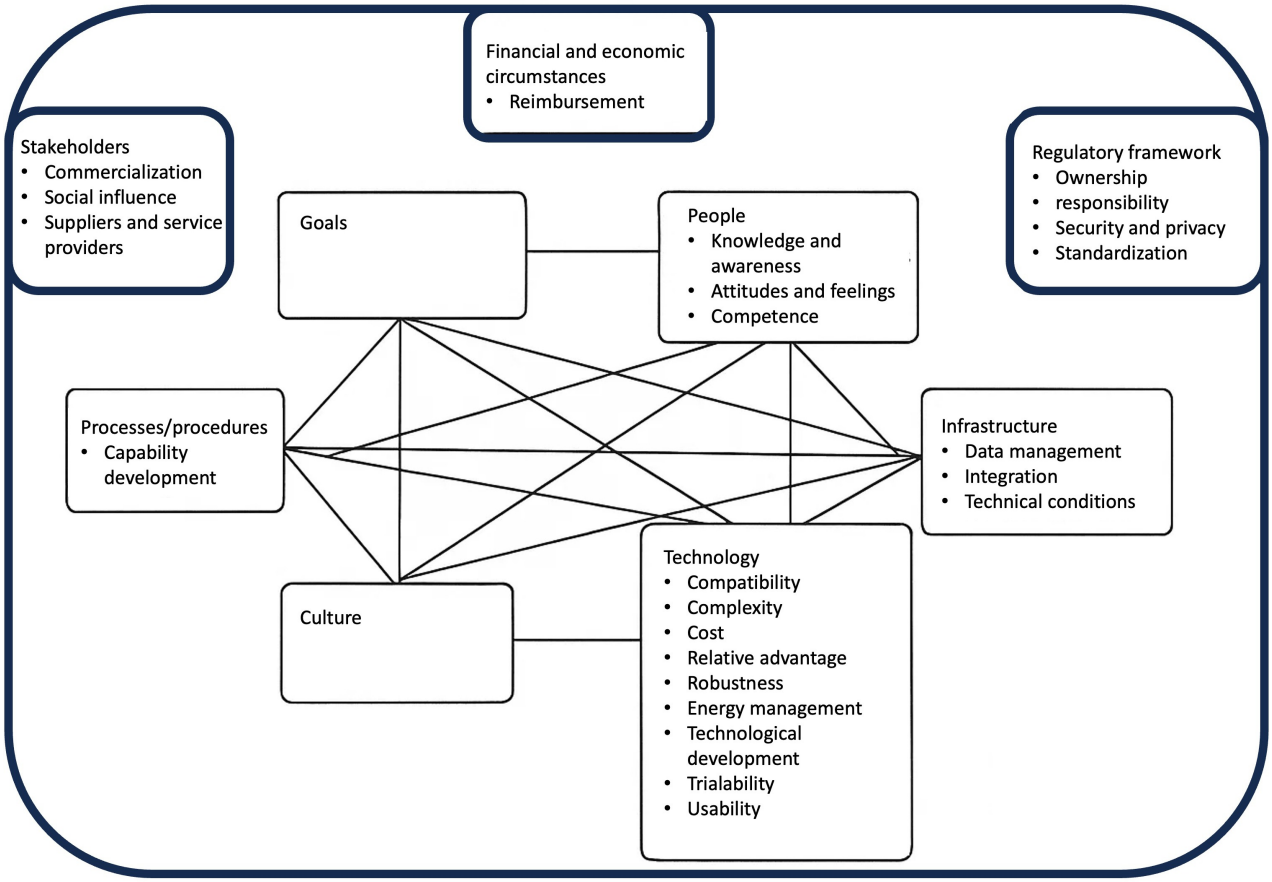
We explored how various factors, particularly organizational and societal, contributed to successful implementation of IoT. Using a comparative case study approach, we analyzed the key mechanisms driving the integration of IoT in health care. Through this analysis, we also aimed to refine sociotechnical systems theory as it pertains to IoT implementation in healthcare.

We moved beyond traditional adoption studies that often rely on frameworks like UTAUT, which, while useful, do not capture the complexity of implementing new technology in practice. Instead, we applied the 7 subsystems proposed by Kronlid et al [20], tailored for health care but flexible in terms of specific influencing factors within each subsystem. Since their framework was developed from a limited number of adoption-focused studies, we did not predefine factors within each subsystem. Instead, we remained open to identifying new factors, allowing a more comprehensive understanding of IoT implementation in health care.

In this article, we aimed to deepen theoretical insights into the critical factors affecting the early phase of implementation. We identified factors influencing the first phases of IoT implementation in health care through the lens of sociotechnical systems theory. Accordingly, our research questions were: (1) What are the most important factors that influence the early phase of implementation of IoT in health care? (2) How can sociotechnical systems theory contribute to understanding IoT implementation in health care?

XSL•FO
RenderX

Figure 1. Sociotechnical systems model (adapted from Kronlid et al [20], which is published under Creative Commons Attribution 4.0 International License [39]).



Methods

Study Design

The study uses a qualitative multiple-case approach [40], suitable for generating new insights and contributing to theory development. This approach is particularly useful for studying emerging areas such as IoT implementation in health care when an in-depth understanding of the phenomenon is needed [40,41]. Using multiple cases allows for replication logic where cases are compared with each other to create robust and reliable insights [41,42].

Case Selection

To identify appropriate sources of information, we reviewed 72 state-of-the-art IoT development initiatives funded by a Swedish agency specializing in IoT solutions. By analyzing the applications and mid-term reports, we identified 20 initiatives directly related to IoT applications within health care. Among these, 5 initiatives were selected as relevant for studying implementation.

These 5 cases represent small-scale IoT pilots, each testing practical applicability in health care. The pilot nature of these projects, with varying numbers of end users (patients), provided insights into implementation factors under controlled conditions before any larger-scale deployment. We initially aimed to include full implementation cases; our review showed that none had reached that stage. Instead, these pilots served as the closest point to practical implementation available for study.

Notably, none of the used IoT technologies had undergone the public procurement process required for broader implementation within Sweden's health and social care systems. Multimedia Appendix 1 provides an overview of these 5 cases.

Data Collection

We started the data collection by first reviewing the existing documentation (eg, grant applications and follow-up reports) concerning each of the projects to map all partner organizations in each of the projects and identify the key players and the project managers leading the initiative. Then, we conducted semistructured interviews with the project managers in each project and regarded them as key informants [43]. We also asked the project managers to identify other relevant partner organizations that we should interview to obtain a broad picture [41]. We conducted 22 interviews altogether, which lasted between 22 and 57 minutes. We followed an interview guide that covered background questions, questions concerning the initiative, and possible factors influencing implementation. The interviews were recorded and transcribed. Table S2 in Multimedia Appendix 2 provides an overview of interviews.

Data Analysis

The analysis of the transcribed interviews was conducted using an abductive method [44]. We started with an inductive systematic thematic analysis [45] to identify patterns grounded in the participants' perspectives. These identified themes were then organized according to the 7 subsystems proposed by

XSL•FO
RenderX

Kronlid et al [20], allowing us to relate the empirical findings to the theoretical framework.

The process involved the following 7 stages. First, following Braun and Clarke [45], we familiarized ourselves with the data by thoroughly reading and rereading the transcripts. This stage aimed to gain a comprehensive understanding of the content and context of the participants' responses, noting initial ideas and recurring topics that emerged throughout the interviews. Second, we identified and generated initial codes from the transcribed interviews. This involved systematically coding segments of text that reflected factors influencing the implementation of IoT solutions in health care. Each segment was labelled with a code that encapsulated its essence. The coding process was flexible and data-driven, ensuring that the codes emerged directly from the participants' responses rather than from preconceived notions. Two researchers coded the transcripts independently to capture diverse interpretations and perspectives, which were then discussed to reach consensus.

Third, we grouped the initial codes into broader themes. This involved examining the codes for patterns and relationships and identifying clusters that represented similar ideas or concepts. The entire research team collaborated to discuss and refine these potential themes, ensuring they accurately reflected the data and contributed to addressing the research questions. Fourth, once the initial themes were established, we engaged in a thorough review process to refine and validate them. This involved checking whether the themes accurately represented the coded data and aligned with the overall dataset. The researchers revisited the interview transcripts to ensure the themes were robust and reflective of the diversity of participant experiences. Any inconsistencies or gaps were addressed through discussions and iterative adjustments. Fifth, after finalizing the themes, we aligned them with the sociotechnical subsystems by Kronlid et al [20] to the extent possible. However, we used inductive labeling, which provided a more accurate reflection of the data. As a result, the findings are structured around five adapted subsystems drawn from Kronlid et al [20] framework: (1) regulatory framework, (2) organizational support, (3) user focus, (4) financial and economic circumstances, and (5) infrastructure. The remaining 2 subsystems in the original framework, "technology" and "stakeholders" [20], did not emerge from the data and are therefore not included in the findings.

Sixth, to identify the specific factors influencing implementation of IoT in health care, we revisited the data within each of the 5 subsystems. This allowed us to break down the data into distinct factors linked to each subsystem. For details, see Table S3 in Multimedia Appendix 2.

Seventh, we compared the subsystems and factors across cases, identifying similarities and differences [41]. Finally, we compiled the data into a coherent findings section, aiming to convey the rich insights derived from the analysis and highlighting how various factors influence the implementation of IoT solutions in health care.

Ethical Considerations

The Swedish Ethics Review Authority made a rejection of the request for ethical review as they considered that ethical review was not necessary, which means that a favorable ethical decision was obtained (exemption number: Dnr 2022-01259-01). The data were collected through interviews, and the participants chose to agree to the interviews. The original consent covered secondary analysis without additional consent. Data were anonymized. No compensation was provided to participants.

Results

Our findings show that factors within 5 subsystems influence implementation of IoT in health care.

Regulatory Framework

Respondents from all but one of the cases we studied identified laws and regulations as major factors influencing the implementation process. In the IoT Older Adults Home Care project, challenges related to laws and regulations were identified early on, and the project was designed for straightforward implementation. The primary issue concerned informed consent from the older adults. The project targeted older adults, including individuals with neurocognitive disorders such as dementia; the implementation faced difficulties due to the inability of older adults to provide informed consent. In one participating municipality (Municipality A), this was taken into account when selecting participants, and no major issues arose. However, in another municipality (Municipality B), the decision was made to include older adults who could not provide informed consent, which led to a failed implementation in that municipality.

> *But you could say that [Municipality B] actually failed due to issues with consent and so on. So, we made installations in [Municipality B], but it didn't work out. Because they wanted to include a user group that couldn't give their consent. Legally, it fell through because they couldn't study the individuals they wanted to include in the project.* [Respondent 11, IoT Older Adults Home Care & IoT Older Adults Care Home projects]

This issue, where those who might benefit the most from IoT technology are unable to provide informed consent, may stem from laws and regulations that have not kept pace with recent technological advancements. One respondent specifically mentioned that the laws and regulations are "outdated," emphasizing how this discrepancy creates challenges from a user perspective.

> *If you look at how laws are written and changed, the time we've had digital tools is very short compared to how long we've had the Patient Data Act. We have several... quite... or very strict laws that we work under when working in healthcare, and they are outdated when it comes to... […] Patient safety and user-friendliness don't really go hand in hand. So that's one of the major obstacles when talking about [IoT] implementation.* [Respondent 16, IoT CF project]

XSL•FO
RenderX

Some strategies were used to navigate existing laws and regulations. This was particularly evident in the IoT CF (cystic fibrosis) project, where creative approaches were used to deliver IoT services to patients and collaborate with companies. One approach involved "hiding behind studies," using ongoing research studies as a way to test and implement IoT technology. Research studies have less demanding regulatory requirements, so hospitals could not object to the activities conducted within them. In addition, the involved companies were not permitted to integrate with specific health care systems. As a workaround, they used the diabetes registry, which operates under different rules.

> *And we hid behind studies. Studies have much lighter requirements, and you can test all the technology. So, it turned out that we started new studies, and we have several ongoing studies because the hospital couldn't say anything. That way, we could hide behind them with the technology. […] My actual research is in something else. This is not my research area, even though I conduct studies, but it's just to hide behind them. As we said, it makes it easier to get around these technical...* [Respondent 15, IoT CF project]

In the IoT LTC (long-term condition) project, respondents emphasized the challenges of handling and securing data in compliance with relevant regulations. Ensuring that data was collected, stored, and used according to legal standards was crucial for maintaining both trust and legality. Meeting certification standards was necessary for the project's legitimacy and acceptance, but obtaining these certifications was a complex process that required careful attention to detail. In the IoT Older Adults Care Home project, there was uncertainty about how the technology should be classified and, consequently, which laws and regulations would apply. The technology captured a clear image of the older adult, which was later blurred on a server. Since a clear image was initially captured, the municipality determined that the camera surveillance law was applicable, leading to the shutdown of the technology testing and a change in the project's direction. After extensive discussions with the municipality's legal team, it was concluded that they should have used a camera instead of a sensor. Had the camera blurred the image immediately, or if the server had been located within the same facility, a different decision might have been reached.

> *So, we use motion sensors and... yes, employ a different interpretation pattern around it or under the radar, you could say that the data you collect about the person is essentially the same. But in our case, it's the camera law that applies.* [Respondent 14, IoT Older Adults Care Home project]

In the IoT CF, IoT Bedwetting, and IoT LTC projects, the public procurement process was identified as a significant challenge. The difficulty lay in ensuring that procurement was conducted legally while still allowing for an innovative approach, which was critical to the success of these projects. The IoT CF project described the procurement process as very slow, with the risk that the product would become outdated during the procurement process. In the IoT LTC project, respondents highlighted how procurement remains a challenging issue, with the belief that the continued implementation of IoT solutions almost always stalls once it reaches the procurement stage. Several respondents suggested that this issue is linked to lawyers' reluctance to test new procurement techniques, driven by a fear of potential legal risks.

> *Then you get into the procurement issues concerning regions and municipalities, and it becomes very, very challenging in many ways. Not least, there is a fear from the actors in, for example, this project. Are they allowed to participate in a procurement, or do they know too much now about the customers in different ways, making them disqualified from the procurement? That risk always exists. So, this procurement step is very, very difficult in the next stage.* [Respondent 21, IoT LTC project]

Organizational Support

Respondents from all the cases we studied emphasized the importance of organizational support in creating an enabling environment for IoT implementation. In the 2 projects, IoT Older Adults Home Care and IoT Older Adults Care Home, both initiated by the same municipality, it was clear that the municipal management had made a firm decision to support digitalization initiatives. This decision, coupled with its effective communication, was considered highly beneficial to the success of these 2 projects. The IoT project manager received strong backing from various levels of organizational leadership. The municipality itself was "very focused on introducing various types of technology within the care services" (Respondent 13, IoT Older Adults Home Care & IoT Older Adults Care Home projects). In the Older Adults Care Home project, managers and staff at the facility where the technology was being tested were particularly enthusiastic and supportive.

> *The housing manager, department manager, and many within the organization are truly passionate about this and are very driven, visionaries, and willing to take risks to make it as good as possible for the individuals. So, we really see that as a success factor in the project.* [Respondent 12, IoT Older Adults Home Care & IoT Older Adults Care Home projects]

Both the IoT CF and IoT LTC projects also highlighted the importance of support functions, such as IT services, administrative assistance, and training, in ensuring successful implementation. The IoT CF project further illustrated the challenges of advancing IoT implementation when facing resistance from organizational stakeholders and a lack of support and resources necessary to establish a more permanent solution. In this case, the IoT project manager encountered significant pushback from colleagues, management, and support functions like the IT department.

> *There are many parts of the IT structure, and we had everyone against us, I must say. Everyone thought it was good, but many thought it was wrong [some technical aspects were deemed illegal by the IT department]. Many didn't think we should do what we did, but then [if we did it] be held accountable for it.* [Respondent 15, IoT CF project]

XSL•FO
RenderX

Several respondents across all cases noted that the level of prioritization within the organization directly impacted a project's progress and success. A common observation was that organizational priorities and decisions outside the project's scope often had a negative influence. For example, in the IoT Older Adults Home Care project, one respondent raised the influence of the political planning: "There's a political agenda that usually follows a 4-year cycle, a planning horizon for operations. So, there are many other factors that govern what actually happens [in projects like this]" (Respondent 7, IoT Older Adults Home Care project). Several respondents also highlighted the difficulty that lower-priority projects had in gaining momentum.

> *The third is getting priority in the regions to obtain staff and resources to assist in the projects is almost impossible, so even if the staff wants to, it is very difficult for them to, well, take care of these projects.* [Respondent 18, IoT LTC project]

Additionally, one project raised an aspect crucial to its success but not mentioned by any other project: organizational readiness for implementation. Respondents from the IoT Older Adults Home Care project emphasized that the level of knowledge regarding implementation varied significantly among managers at different organizational levels, influencing the feasibility of implementing the IoT solution.

> *There is a large variation in maturity levels [knowledge of implementation] among our operational managers. And when I say operational managers, I mean everyone from those at individual care facilities, group homes, to the highest leadership. They don't have a unified understanding [of what to do and how to do it]; it's quite diverse.* [Respondent 6, IoT Older Adults Home Care project]

User Focus

Another major success factor identified in all the cases we studied was having a strong user focus, which encompassed both patients and health care staff.

> *It originates from the staff and patients from the start, and they have been involved throughout and have had a say in how it should look [like], how the tool should look [like], what functionality we should have, and what the priorities should be.* [Respondent 21, IoT LTC project]

In the IoT LTC project, it became clear that involving users (ie, health care staff) not only helped the service developer improve the usability of their service but also increased the likelihood of successful implementation. Respondents noted that while the service developer initially had a vision for how the service should function, there was often a disconnect between the engineers' perspective and how patients, relatives, and health care staff perceived the service. In the IoT CF project, the team had initial ideas about the effects and benefits the technology would bring to patients, but additional benefits were identified by both patients and their relatives.

> *So, we made the technology available to everyone, instead of making decisions like "you think," "we need," or similar. And we saw quite clearly that it's very difficult to guess who will benefit from the technology. [...] I think that [involving patients] was a success factor because we said, "This doesn't mean it should replace one or the other," but we let our patients choose instead of healthcare professionals making the choice.* [Respondent 15, IoT CF project]

In addition to involving immediate users (ie, caregivers and patients), it is also essential to include a broader range of users, as demonstrated in the IoT CF project. This project engaged health care staff both directly and indirectly connected to patient care, such as medical secretaries, nurses, dietitians, psychologists, physiotherapists, and counsellors. This extensive involvement was highlighted as a key success factor. However, some respondents also pointed out the complexities that IoT technology introduces to caregiving, noting that health care staff may feel that the more technology is introduced, the further they move away from their core role as caregivers. This creates a paradox, as many choose to work in health care precisely to be close to the patient.

> *We have gained a lot of insight into the technology frustration that exists [i.e, the complex role of technology in caregiving] and that, as it is today, they [the health care staff] feel they are moving further away from their caregiver role than they would like [to].* [Respondent 12, IoT Older Adults Home Care project]

A match between user instructions and user capabilities is also crucial in IoT implementation projects. Both the IoT CF and IoT Older Adults Home Care projects emphasize the importance of clear communication with users and adjusting the complexity of instructions to suit their abilities. Additionally, one project (IoT Bedwetting) highlighted the role of national guidelines—care guidelines from authorities that outline best practices in different clinical areas—in the success of IoT implementation. Respondents from the IoT Bedwetting project suggested that if national guidelines recommend IoT solutions, this could become a critical factor in achieving successful outcomes.

Financial and Economic Circumstances

Securing sufficient resources for implementing IoT solutions was highlighted as a challenge in several interviews across the IoT LTC, IoT CF, and IoT Older Adults Home Care projects. These resources primarily include financial support. Shorter projects are often insufficient to drive meaningful change, as it takes considerable time—and therefore money—to alter established practices. There is often ambiguity around who should bear the costs for the necessary time, resources, and ongoing maintenance of the services.

> *No one wants to pay. Everyone thinks someone else should pay. But now I have no... But now I'm a little... Now we're shutting down if the hospital doesn't step in and pay. So I've set some demands. And everyone thinks this group is good and should continue, but it's... everyone thinks someone else should pay.* [Respondent 15, IoT CF project]

XSL•FO
RenderX

Infrastructure

Respondents from 3 cases emphasized that infrastructure is crucial for successful IoT implementation. They noted that the process was often hindered by either insufficient or unstable WIFI connections, as well as concerns over the use of cloud services. In the IoT Older Adults Care Home project, for example, the IoT solution had already been procured by the region for other applications. However, when installation began, it became clear that WIFI coverage was inadequate to get the technology fully operational.

> *Exactly, it's another project that is implementing the equipment we intend to use in this project. In that project, I perceive that the staff has some resistance because there have been infrastructure issues that have prevented it from working as expected. For example, the WIFI hasn't had the coverage they thought it would, and when the equipment lost network connectivity, it resulted in false alarms, in reality.* [Respondent 12, IoT Older Adults Home Care & IoT Older Adults Care Home projects]

In the IoT Older Adults Home Care project, WIFI access was also a significant issue. Here, older clients were responsible for their own WIFI, complicating the service provider's efforts to install the necessary infrastructure. In the IoT Bedwetting project, concerns about the security of cloud services also created challenges.

> *...I have been in contact with [the region], and there the IT department flatly said no, that "we cannot agree to this because we do not consider cloud services to be secure." So, that... those are the things I think are the problem, that there are many smart technical solutions, but in many cases, the healthcare system is not quite ready to embrace them.* [Respondent 4, IoT Bedwetting project]

In the IoT Older Adults Home Care project, stakeholders such as the IT department seemed unaware of local cloud solutions and declined to consider them without further investigation. In contrast, the IoT Older Adults Care Home project successfully addressed potential security concerns by contracting with a local cloud provider.

In addition, communication tools play a critical role in supporting implementation. Our findings suggest that when communication tools are not properly aligned with the IoT solution, they can obstruct the implementation process.

> *Then there are these contextual factors. For example, the staff currently use fairly old phones to receive their alarms, and this means that it takes quite a while... Every time an alarm comes in, it takes quite a long time for them to access their phones. They need to enter their code every time, open an app, and then turn off the alarm. And it takes quite a long time.* [Respondent 12, IoT Older Adults Home Care & IoT Older Adults Care Home projects]

Another important aspect of communication tools is the use of web-based meetings. Our findings indicate that well-functioning web-based meetings can greatly facilitate implementation. In the IoT CF project, secure web-based meetings were essential for delivering the service. These meetings were facilitated centrally through a national service called "1177," the primary digital platform for health care services in Sweden.

Discussion

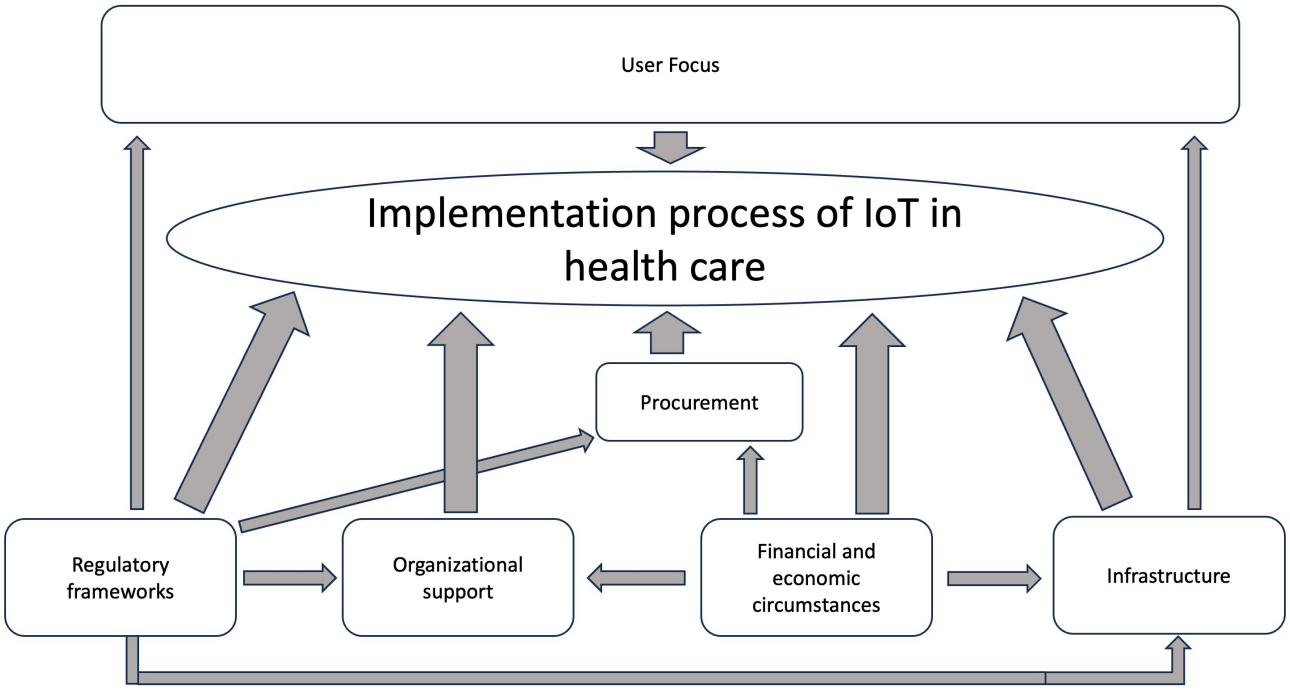
Principal Findings

This study investigates the key factors influencing IoT implementation in health care by examining 5 state-of-the-art IoT projects in Sweden. It addresses the pressing need for more empirical studies on IoT implementation [20], as well as the need for an approach that goes beyond purely technical factors, focusing instead on a broader range of influences within organizational and social contexts [8,25].

Figure 2 illustrates our theoretical contributions and proposes a sociotechnical systems model tailored to IoT implementation in health care. It is difficult to determine the exact causality between the different factors, but based on the interviews, Figure 2 illustrates our interpretation of the main relationships between the factors.

XSL•FO
RenderX

Figure 2. The most important factors that influence the implementation process of Internet of Things (IoT) in health care. The broader arrows represent a direct impact on the IoT implementation process in health care, while the narrower arrows represent an indirect impact—from one factor to another factor that in turn affects the ability to implement IoT in health care.



Our research contributes to the growing discourse on IoT in health care, which has predominantly focused on the adoption phase [5,15,25]. While prior studies emphasize technical factors as primary considerations during adoption [25,26], our findings underscore the importance of organizational and societal factors in the implementation phase. Through our study, we advance sociotechnical systems theory by demonstrating how its components interact in real-world health care settings, particularly in the context of IoT. Our findings reveal complex dynamics between technological potential and social structures, identifying key gaps, tensions, and interdependencies that impact implementation. Below, we discuss 5 primary ways in which our study contributes to applying sociotechnical systems theory to health care IoT implementation.

First, by applying the Kronlid et al [20] model, we identify specific refinements to the sociotechnical systems model necessary for the health care sector. Of the 4 core factors—people, infrastructure, technology, and processes or procedures [20]—our findings support 3. We relabel "people" as "user focus" and "processes or procedures" as "organizational support" to better reflect its relevance in health care. Infrastructure remains unchanged. The subsystem, "technology," was not observed in the 5 cases we studied.

Second, our study confirms 2 of the 3 external factors in the sociotechnical systems model, financial circumstances and regulatory frameworks [20], but did not observe evidence for the third factor, stakeholders. More importantly, we suggest that financial circumstances and regulatory frameworks should be viewed as internal organizational factors because of their interconnectedness with other internal elements, particularly organizational support. For example, financial resources significantly influence organizational support: when funding is adequate, organizations can prioritize IoT projects, allocate resources, and provide leadership backing. Conversely, limited funding can restrict support and hinder even promising projects. Regulatory frameworks similarly shape user engagement by setting boundaries around privacy and consent, which are critical in health care. Regulations that protect vulnerable users can inadvertently restrict participation in IoT projects [46]. Thus, we propose that the sociotechnical systems model should reposition these factors as internal dynamics within health care IoT implementations.

Further research could examine whether the remaining subsystems from Kronlid et al [20], technology and stakeholders, are relevant to IoT implementation in health care beyond Sweden. Regarding the technology subsystem, the affordability and technical challenges of IoT solutions may vary across countries. In low-resource settings, such as parts of Africa, the high costs of IoT devices, infrastructure, and maintenance can hinder adoption. In contrast, high-income countries like the United States and Germany may face greater challenges in integrating IoT into existing health care systems, requiring significant investment in interoperability and cybersecurity.

As for the stakeholders' subsystem, the commercialization of IoT solutions is shaped by regulations, market demand, and funding availability. In countries with strict health care regulations (eg, the European Union [EU] under General Data Protection Regulation), compliance costs can make it difficult for startups to scale their solutions. Meanwhile, in developing markets, a lack of investor confidence and infrastructure limitations may slow commercialization, even when there is strong health care demand.

Third, our findings highlight the essential role of health care workers' attitudes and patient involvement in harmonizing social and technical elements of IoT adoption. The observed frustration among health care staff with certain technologies supports the

XSL•FO
RenderX

sociotechnical systems theory assertion that social environments such as attitudes, values, and identities strongly influence technology use [47]. This insight suggests that future adaptations of sociotechnical systems theory could benefit from explicitly addressing the emotional and identity-related impacts of technology, particularly in health care, where professional identity often ties closely to user autonomy and caregiving roles.

Fourth, our study reveals that regulations often lag behind technological advances, creating obstacles to IoT implementation. Rather than facilitating adoption, regulatory complexity can become a barrier, consistent with previous findings on medical device regulations [48]. Health care's rule-driven culture fosters a cautious approach, limiting the adoption of emerging technologies [49]. Our findings further indicate that existing procurement structures can impede innovation, as traditional processes stifle creative thinking necessary for IoT implementation. Although innovation-based procurement strategies are promoted in the EU [50], none of the cases we examined identified these as practical solutions. To enable effective IoT adoption, managers must understand legal and regulatory constraints while collaborating with legal teams, a rarely studied area in management research [51]. Without this collaboration, legal departments may adopt overly cautious stances, leading to project rejections, while support departments may resist due to anticipated increases in workload. In one case, the organization strategically managed data by using local servers to avoid EU data-sharing restrictions [52,53], demonstrating that creative solutions can help navigate regulatory challenges. The respondents mentioned that they were hiding behind academic studies, ie, using ongoing research studies as a way to both test and implement IoT technology. It indicates a certain level of frustration regarding the lack of flexibility in the organizational structures.

Finally, the technical complexity of IoT systems, coupled with the need for simplicity in user interfaces, underscores the principle that technology and human systems must evolve together [54,55]. While IoT technologies provide advanced capabilities, their complexity can alienate users, particularly in health care, where ease of use is critical. This observation suggests a vital revision for sociotechnical systems theory: usability should be prioritized as a co-evolutionary requirement within sociotechnical design, not as an afterthought. This aligns with theories of sociomateriality, supporting the idea that intuitive, user-friendly design is essential in fields with high cognitive and emotional demands [54]. Thus, even if our study did not specifically address user engagement strategies tailored to vulnerable groups, such as individuals with cognitive impairments or mobility limitations, we recognize this as critical for future research. Practical methods such as participatory design and co-creation techniques should be used to ensure that IoT solutions are inclusive and accessible [56]. A stronger focus on user experience design will enhance the relevance, equity, and effectiveness of implementation efforts in health care.

Practical Implications

This study reveals several practical implications for managers involved in health care IoT implementation, highlighting the importance of collaboration, financial strategy, training, and regulatory awareness. First, close engagement with legal departments is crucial. Developing a collaborative relationship with legal teams allows managers to address complex regulatory concerns upfront, avoiding overly conservative stances that could stifle innovation. To facilitate smoother project approvals, it may also be beneficial to provide legal teams with training specific to IoT technologies, helping them understand the unique aspects of these projects. Second, demonstrating the economic benefits of IoT initiatives is equally essential. Managers can strengthen organizational support by conducting detailed cost-benefit analyses that clearly illustrate potential financial gains. This approach not only helps in securing initial funding but also promotes a culture that values and supports innovative solutions. Third, investing in comprehensive training programs is also vital for successful implementation. By providing targeted training across all staff levels, including frontline workers, support departments, and management, managers can build the technical competencies necessary for IoT integration. Moreover, fostering cross-functional collaboration by forming interdisciplinary teams encourages the sharing of expertise, which can lead to more effective problem-solving and innovation.

Fourth, ensuring that patients, health care providers, and other stakeholders are involved in the implementation process can significantly enhance the relevance and acceptance of IoT solutions. Addressing the needs and concerns of these users from the outset ensures that solutions are practical and user-centered. Developing clear protocols for managing informed consent is particularly important when technologies involve sensitive components like sensors or cameras. Special attention to accessibility, especially for patients with cognitive disabilities, further reinforces the inclusivity of the implementation. Finally, engaging with policymakers to advocate for flexible and supportive regulatory frameworks is crucial. By working together, managers and policymakers can create a balanced approach that safeguards patient safety and privacy while fostering a regulatory environment that encourages digital innovation in health care. This proactive stance helps to create a foundation for sustainable IoT implementation, aligning technological advancement with patient care priorities and organizational objectives.

Limitations and Further Research

Our study has 2 main limitations. First, it examines only 5 IoT pilot implementation cases, which, while allowing for a deep exploration of IoT implementation factors, limits the generalizability of our findings. The small sample size may mean that our conclusions are less applicable to other health care settings, particularly those outside Sweden. However, we believe that the results may very well be relevant for countries with similar demographics and systems for social and health care. Expanding future research to include a broader and more varied set of cases across different regions or countries would help to enhance generalizability. However, as shown by our sampling process, identifying numerous relevant IoT cases can be challenging. Second, each case we studied was in the pilot implementation stage, with none yet reaching full-scale public procurement or integration into Sweden's health care system. This limitation restricts our insights into the long-term

XSL•FO
RenderX

challenges and sustainability of IoT implementation in health care. Consequently, our findings do not fully capture the issues and opportunities associated with later stages of implementation, such as public procurement, scaling, and long-term use.

Building on these limitations, several avenues for further research emerge. One area needing exploration is the adaptation of health care laws and regulations to better support IoT implementation. Future studies might investigate how regulatory frameworks and procurement models, such as innovation-based procurement, could be refined to foster IoT implementation. Research could also focus on best practices for aligning health care regulations with the rapid advancements in IoT technology. Future research could also aim to identify actionable solutions to the challenges that have been identified.

Since this study is based on pilot projects, there is also a need for longitudinal research that extends beyond the pilot phase. Examining the full-scale implementation of IoT solutions would provide insights into additional challenges and opportunities, including public procurement, solution scalability, and sustained use over time. Further, expanding the scope to different regions or countries would enable comparative studies to assess whether factors influencing IoT implementation are universally relevant or specific to certain national or health care system contexts.

Another critical research direction involves deepening our understanding of effective user involvement in IoT design and implementation. Future studies could explore best practices for managing informed consent, especially for vulnerable populations, like individuals with cognitive disabilities. These studies could help shape policies that ensure the full benefits of IoT technologies reach all user groups, reinforcing the patient-centered approach essential to health care innovation.

Conclusions

In conclusion, this study explored 5 IoT pilot implementation projects in Sweden, shedding light on key factors that influence the deployment of IoT solutions in health care. By responding to the demand for empirical research on IoT implementation, our findings contribute to an understanding of the practical challenges and organizational dynamics involved, expanding beyond the adoption-focused literature. While IoT has the potential to significantly enhance health care delivery, several barriers remain, especially concerning regulatory complexity and rigid legal frameworks. These findings underscore the importance of ongoing collaboration between managers and legal teams to prevent regulatory constraints from hindering innovation. Our study also highlights the necessity of strong leadership in championing IoT projects within health care organizations. Managers play a pivotal role in clearly communicating the economic value of IoT through cost-benefit analyses and fostering a collaborative, cross-functional culture to tackle implementation hurdles. Adopting a user-centered approach—one that involves health care providers, patients, and families throughout the process—emerges as essential for achieving broad acceptance and success of IoT solutions. Additionally, managing informed consent, particularly for vulnerable populations such as individuals with cognitive disabilities, is crucial for the ethical deployment of these technologies. Overall, our study underscores the need for thoughtful strategies that prioritize regulatory alignment, economic rationale, and user involvement to realize the full potential of IoT in health care.

Conflicts of Interest

None declared.

Multimedia Appendix 1
An overview of the 5 cases.
[DOCX File, 21 KB - humanfactors_v12i1e71546_app1.docx]

Multimedia Appendix 2
An overview of interviews and an overview of subsystems and specific factors influencing implementing IoT solutions in health care. IoT: Internet of Things.
[DOCX File, 17 KB - humanfactors_v12i1e71546_app2.docx]

References

1. Aceto G, Persico V, Pescapé A. Industry 4.0 and health: Internet of Things, Big Data, and Cloud Computing for Healthcare 4.0. J Ind Inf Integr 2020 Jun;18:100129. [doi: 10.1016/j.jii.2020.100129]
2. Arfi WB, Nasr IB, Kondrateva G, Hikkerova L. The role of trust in intention to use the IoT in eHealth: application of the modified UTAUT in a consumer context. Technol Forecast Soc Change 2021 Jun;167:120688. [doi: 10.1016/j.techfore.2021.120688]
3. Belfiore A, Cuccurullo C, Aria M. IoT in healthcare: a scientometric analysis. Technol Forecast Soc Change 2022 Nov;184:122001. [doi: 10.1016/j.techfore.2022.122001]
4. Atzori L, Iera A, Morabito G. The Internet of Things: a survey. Comput Netw 2010 Oct;54(15):2787-2805. [doi: 10.1016/j.comnet.2010.05.010]
5. Rejeb A, Rejeb K, Treiblmaier H, et al. The Internet of Things (IoT) in healthcare: taking stock and moving forward. Internet Things 2023 Jul;22:100721. [doi: 10.1016/j.iot.2023.100721]

XSL•FO
RenderX

6. Ahmadi H, Arji G, Shahmoradi L, Safdari R, Nilashi M, Alizadeh M. The application of internet of things in healthcare: a systematic literature review and classification. Univ Access Inf Soc 2019 Nov;18(4):837-869. [doi: 10.1007/s10209-018-0618-4]
7. Rogers EM. Diffusion of Innovations, 5th edition: New York: Free press; 2003.
8. Martínez-Caro E, Cegarra-Navarro JG, García-Pérez A, Fait M. Healthcare service evolution towards the Internet of Things: an end-user perspective. Technol Forecast Soc Change 2018 Nov;136:268-276. [doi: 10.1016/j.techfore.2018.03.025]
9. Ben Arfi W, Ben Nasr I, Khvatova T, Ben Zaied Y. Understanding acceptance of eHealthcare by IoT natives and IoT immigrants: an integrated model of UTAUT, perceived risk, and financial cost. Technol Forecast Soc Change 2021 Feb;163:120437. [doi: 10.1016/j.techfore.2020.120437]
10. Lu Y. Examining user acceptance and adoption of the Internet of Things. Int J Bus Sci Appl Manage 2021;16(3):1-17. [doi: 10.69864/ijbsam.16-3.150]
11. Al-Rawashdeh M, Keikhosrokiani P, Belaton B, Alawida M, Zwiri A. Effective factors for the adoption of IoT applications in nursing care: a theoretical framework for smart healthcare. J Build Eng 2024 Jul;89:109012. [doi: 10.1016/j.jobe.2024.109012]
12. Yin Y, Zeng Y, Chen X, Fan Y. The Internet of Things in healthcare: an overview. J Ind Inf Integr 2016 Mar;1:3-13. [doi: 10.1016/j.jii.2016.03.004]
13. Zheng X, Martin P, Brohman K. CLOUDQUAL: a quality model for cloud services. IEEE Trans Ind Inf 2014;10(2):1527-1536. [doi: 10.1109/TII.2014.2306329]
14. Casola V, Castiglione A, Choo KKR, Esposito C. Healthcare-related data in the Cloud: challenges and opportunities. IEEE Cloud Comput 2016;3(6):10-14. [doi: 10.1109/MCC.2016.139]
15. Al-Rawashdeh M, Keikhosrokiani P, Belaton B, Alawida M, Zwiri A. IoT adoption and application for smart healthcare: a systematic review. Sensors (Basel) 2022 Jul 19;22(14):5377. [doi: 10.3390/s22145377] [Medline: 35891056]
16. Davis F. A technology acceptance model for empirically testing new end-user information systems: theory and results [Dissertation]. : Massachussetts Institute of Technology; 1985.
17. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manage Sci 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
18. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425. [doi: 10.2307/30036540]
19. Venkatesh V, Thong JYL, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Q 2012;36(1):157. [doi: 10.2307/41410412]
20. Kronlid C, Brantnell A, Elf M, Borg J, Palm K. Sociotechnical analysis of factors influencing IoT adoption in healthcare: a systematic review. Technol Soc 2024 Sep;78:102675. [doi: 10.1016/j.techsoc.2024.102675]
21. Tortorella GL, Fogliatto FS, Mac Cawley Vergara A, Vassolo R, Sawhney R. Healthcare 4.0: trends, challenges and research directions. Product Plan Control 2020 Nov 17;31(15):1245-1260. [doi: 10.1080/09537287.2019.1702226]
22. Trist EL, Bamforth KW. Some social and psychological consequences of the longwall method of coal-getting: an examination of the psychological situation and defences of a work group in relation to the social structure and technological content of the work system. Hum Relat 1951:3-38. [doi: 10.1177/001872675100400101]
23. Scott Kruse C, Karem P, Shifflett K, Vegi L, Ravi K, Brooks M. Evaluating barriers to adopting telemedicine worldwide: a systematic review. J Telemed Telecare 2018 Jan;24(1):4-12. [doi: 10.1177/1357633X16674087] [Medline: 29320966]
24. Khodadad-Saryazdi A. Exploring the telemedicine implementation challenges through the process innovation approach: a case study research in the French healthcare sector. Technovation 2021 Sep;107:102273. [doi: 10.1016/j.technovation.2021.102273]
25. Dantu R, Dissanayake I, Nerur S. Exploratory analysis of Internet of Things (IoT) in healthcare: topic modelling & co-citation approaches. Inf Syst Manage 2021 Jan 2;38(1):62-78. [doi: 10.1080/10580530.2020.1746982]
26. Dimitrov DV. Medical Internet of Things and Big Data in healthcare. Healthc Inform Res 2016 Jul;22(3):156-163. [doi: 10.4258/hir.2016.22.3.156] [Medline: 27525156]
27. Stavropoulos TG, Papastergiou A, Mpaltadoros L, Nikolopoulos S, Kompatsiaris I. IoT wearable sensors and devices in elderly care: a literature review. Sensors (Basel) 2020;20(10):2826. [doi: 10.3390/s20102826]
28. Lihong Jiang L, Da Xu L, Cai H, Jiang Z, et al. An IoT-oriented data storage framework in cloud computing platform. IEEE Trans Ind Inf 2014;10(2):1443-1451. [doi: 10.1109/TII.2014.2306384]
29. Wang L, Yang GZ, Huang J, et al. A wireless biomedical signal interface system-on-chip for body sensor networks. IEEE Trans Biomed Circuits Syst 2010;4(2):112-117. [doi: 10.1109/TBCAS.2009.2038228]
30. Kuo AH. Opportunities and challenges of cloud computing to improve health care services. J Med Internet Res 2011 Sep 21;13(3):e67. [doi: 10.2196/jmir.1867] [Medline: 21937354]
31. Sultan N. Making use of cloud computing for healthcare provision: opportunities and challenges. Int J Inf Manage 2014 Apr;34(2):177-184. [doi: 10.1016/j.ijinfomgt.2013.12.011]
32. Alraja M. Frontline healthcare providers' behavioural intention to Internet of Things (IoT)-enabled healthcare applications: a gender-based, cross-generational study. Technol Forecast Soc Change 2022 Jan;174:121256. [doi: 10.1016/j.techfore.2021.121256]

XSL•FO
RenderX

33. Huarng KH, Yu THK, Lee CF. Adoption model of healthcare wearable devices. Technol Forecast Soc Change 2022 Jan;174:121286. [doi: 10.1016/j.techfore.2021.121286]
34. May C. Towards a general theory of implementation. Implement Sci 2013 Feb 13;8(18):18. [doi: 10.1186/1748-5908-8-18] [Medline: 23406398]
35. Broens THF, Huis in't Veld RMHA, Vollenbroek-Hutten MMR, Hermens HJ, van Halteren AT, Nieuwenhuis LJM. Determinants of successful telemedicine implementations: a literature study. J Telemed Telecare 2007 Sep 1;13(6):303-309. [doi: 10.1258/135763307781644951]
36. Cannavacciuolo L, Capaldo G, Ponsiglione C. Digital innovation and organizational changes in the healthcare sector: multiple case studies of telemedicine project implementation. Technovation 2023 Feb;120:102550. [doi: 10.1016/j.technovation.2022.102550]
37. Leavitt HJ. Applying organizational change in industry:structural, techno-logical and humanistic approaches. In: March JG, editor. Handbook of Organizations: Rand McNally; 1965:1144-1170.
38. Davis MC, Challenger R, Jayewardene DNW, Clegg CW. Advancing socio-technical systems thinking: a call for bravery. Appl Ergon 2014 Mar;45(2):171-180. [doi: 10.1016/j.apergo.2013.02.009] [Medline: 23664481]
39. Attribution 4.0 International Deed. Creative Commons. URL: https://creativecommons.org/licenses/by/4.0/ [accessed 2025-05-16]
40. Yin RK. Case Study Research: Design and Methods, 5th edition: London: SAGE; 2014.
41. Eisenhardt KM, Graebner ME. Theory building from cases: opportunities and challenges. Acad Manage J 2007 Feb;50(1):25-32. [doi: 10.5465/amj.2007.24160888]
42. Eisenhardt KM. Building theories from case study research. Acad Manage Rev 1989;14(4):532-550. [doi: 10.5465/amr.1989.4308385]
43. Kumar N, Stern LW, Anderson JC. Conducting interorganizational research using key informants. Acad Manage J 1993 Dec 1;36(6):1633-1651. [doi: 10.2307/256824]
44. Dubois A, Gadde LE. Systematic combining: an abductive approach to case research. J Bus Res 2002 Jul;55(7):553-560. [doi: 10.1016/S0148-2963(00)00195-8]
45. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
46. Mubarak F, Suomi R. Elderly forgotten? Digital exclusion in the information age and the rising grey digital divide. Inquiry 2022;59:469580221096272. [doi: 10.1177/00469580221096272] [Medline: 35471138]
47. Leonardi PM. When flexible routines meet flexible technologies: affordance, constraint, and the imbrication of human and material agencies. MIS Q 2011;35(1):147. [doi: 10.2307/23043493]
48. Beitler BG, Abraham PF, Glennon AR, et al. Interpretation of regulatory factors for 3D printing at hospitals and medical centers, or at the point of care. 3D Print Med 2022 Feb 1;8(1):7. [doi: 10.1186/s41205-022-00134-y] [Medline: 35106697]
49. Oikonomou E, Carthey J, Macrae C, Vincent C. Patient safety regulation in the NHS: mapping the regulatory landscape of healthcare. BMJ Open 2019 Jul 9;9(7):e028663. [doi: 10.1136/bmjopen-2018-028663] [Medline: 31289082]
50. Uyarra E, Edler J, Garcia-Estevez J, Georghiou L, Yeow J. Barriers to innovation through public procurement: a supplier perspective. Technovation 2014 Oct;34(10):631-645. [doi: 10.1016/j.technovation.2014.04.003]
51. Lescrauwaet L, Wagner H, Yoon C, Shukla S. Adaptive legal frameworks and economic dynamics in emerging technologies: navigating the intersection for responsible innovation. LawEco 2022;16(3):202-220. [doi: 10.35335/laweco.v16i3.61]
52. Kushida KE, Murray J, Zysman J. Diffusing the cloud: cloud computing and implications for public policy. J Ind Compet Trade 2011 Sep;11(3):209-237. [doi: 10.1007/s10842-011-0106-5]
53. Aldossary S, Allen W. Data security, privacy, availability and integrity in cloud computing: issues and current solutions. Int J Adv Comput Sci App ;7(4). [doi: 10.14569/IJACSA.2016.070464]
54. Orlikowski WJ. Sociomaterial practices: exploring technology at work. Org Stud 2007 Sep;28(9):1435-1448. [doi: 10.1177/0170840607081138]
55. Baxter G, Sommerville I. Socio-technical systems: from design methods to systems engineering. Interact Comput 2011 Jan;23(1):4-17. [doi: 10.1016/j.intcom.2010.07.003]
56. Nordin S, Sturge J, Meijering L, Elf M. A 5-day codesign sprint to improve housing decisions of older adults: lessons learned from Sweden and the Netherlands. Int J Soc Res Methodol 2024 Nov;27(6):661-673. [doi: 10.1080/13645579.2023.2247290]

Abbreviations

CF: cystic fibrosis
EU: European Union
IoT: Internet of Things
LTC: long-term condition
TAM: technology acceptance model
UTAUT: Unified Theory of Acceptance and Use of Technology

XSL•FO
RenderX

Edited by A Kushniruk; submitted 21.01.25; peer-reviewed by GK Gupta, LP Hung, MV Ebrahim; revised version received 07.05.25; accepted 12.05.25; published 27.06.25.

Please cite as:
Palm K, Kronlid C, Elf M, Brantnell A
Key Factors Shaping Successful Implementation of the Internet of Things (IoT) in Health Care: Qualitative Study
JMIR Hum Factors 2025;12:e71546
URL: https://humanfactors.jmir.org/2025/1/e71546
doi:10.2196/71546

© Klas Palm, Carl Kronlid, Marie Elf, Anders Brantnell. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 27.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Optimizing Participant Engagement in Cyberhealth Co-Design: Course-of-Action Framework Analysis

Melanie Tremblay[1,2], PhD; Christine Hamel[2], PhD; Anabelle Viau-Guay[2], PhD; Dominique Giroux[3,4,5], PhD

[1]École Nationale d'Administration Publique, 555, boulevard Charest Est, Québec, QC, Canada

[2]Department of Teaching and Learning Studies, Centre de recherche et d'intervention sur la réussite scolaire, Laval University, Québec, QC, Canada, Québec, QC, Canada

[3]Department of Rehabilitation, Université Laval, Québec, QC, Canada

[4]Center of Excellence on Aging Quebec, Québec, QC, Canada

[5]Centre intégré universitaire de santé et de services sociaux de la Capitale-Nationale, Québec, QC, Canada

Corresponding Author:
Melanie Tremblay, PhD
École Nationale d'Administration Publique, 555, boulevard Charest Est, Québec, QC, Canada

Abstract

Background: Co-design is recognized for its potential to enhance the usability of products through active user participation. However, participation alone does not guarantee the effectiveness of the resulting product. Understanding participants' engagement during co-design activities can provide valuable insights into their motivations, concerns, and contributions, which are critical to achieving successful outcomes.

Objective: This study aims to analyze participant engagement in a digital health co-design parent project focusing on developing a tool to facilitate support-seeking for elderly caregivers.

Methods: The parent project included 74 participants from 3 categories: caregivers, health care and social service professionals, and community workers. Testimonies for this study were collected from 20 participants using the self-confrontation interview methodology. Engagement was analyzed qualitatively using the course-of-action framework. The engagements were organized into emergent themes. The analysis focused on variations in engagement patterns across participant categories and sessions.

Results: A total of 3 themes of engagement were identified: tool design, participant needs, and contextual situations. Engagement was distributed similarly across themes, except for community workers, who were more focused on needs (52/94, 42%) than tool design (25/62, 20%). There was significant variation in engagement over sessions, with tool design being more prominent during specific sessions (co-design sessions CoD5, CoD7, and CoD8) and less important during others (CoD4, AC2 [advisory committee session], CoD6, and AC3). Activities directly tied to design tasks significantly enhanced engagement with tool design. These results underscore the influence of activity types in shaping participants' focus and involvement.

Conclusions: This study highlights the role of affordances in co-design activities to balance engagement across design, collaboration, and participation dimensions. By strategically leveraging affordances, future co-design projects can optimize engagement and ensure more effective outcomes.

(JMIR Hum Factors 2025;12:e70772) doi:10.2196/70772

KEYWORDS

engagement; co-design; caregivers; activity analysis; course-of-action framework; affordances

Introduction

Background

Digital health can improve access to and quality of health care, lower costs, and provide a personalized experience for patients, but it also raises ethical challenges [1], usability problems [2], and, as seen, few successful implementations [3]. Involving end users in the creation of digital health solutions can minimize usability and accessibility problems [4]. Although co-design has several theoretical virtues, it does not guarantee positive outcomes. A total of 2 systematic literature reviews identified mixed and even negative outcomes, including communication problems, misunderstandings between users and the development team leading to conflicts, higher user expectations, and a negative influence on data quality [5-7]. What factors lead to better results?

For example, there is a great diversity of activities conducted during co-design sessions: interviews, advisory council meetings, focus groups, surveys, deliberative methods [8], and plenary sessions [9]. Gray literature also offers a wide range of techniques for collaborative sessions, problem-solving, and

XSL•FO
RenderX

ideation that can be used in co-design [10-12]. How does the type of activity affect outcomes?

Extra time is needed when involving patients in health research [13]. To date, research on co-design in digital health is still seldom described, evaluated, or empirically tested, raising questions about the cost-effectiveness of user contributions compared to the investments required in time, effort, and money [8]. This paper contributes to the body of work on co-design initiatives with a focus on what drives participants—namely, their engagement—in a co-design project, based on a strong theoretical lens on human activity and empirical results provided by the participants themselves.

Positioning Engagement and Co-Design Through a Theoretical Lens

Co-design in this study aligns with the perspective of Sanders and Stappers [14], which involves users (or potential users) in the development of knowledge, concepts, and idea generation. For us, co-design lies at the intersection of 3 dimensions: design, collaboration, and participation. First, participants engage in "design as a reflective practice," following Schön [15], to reflect on their situation and contemplate potential solutions to improve it. The second dimension is collaboration, where actors engage in mutual learning [16-20], leveraging each other's knowledge and experience to better understand the situation and co-create the solution. The third dimension is participation, which involves how actors act, speak, and think [21], and how they may ultimately benefit from their participation beyond the co-design project's objectives [22]. We perceive participation and collaboration as interrelated dimensions rather than levels on a ladder [23].

Engagements of participants can help understand participation in co-design by providing insights into what motivates or concerns a person to participate (act, speak, or think; see Textbox 1). For example, an actor might have concerns about the capability of users with low digital literacy. These engagements affect their participation and their willingness to engage in design as a reflective practice and in mutual learning about the situation. They can influence the potential for collaboration, creativity, and ultimately the innovative, useful, and usable potential of the proposed solution.

Textbox 1. Plain language summary.

- Affordances: perceptual cues in the environment that suggest specific interactions.
- Course-of-action framework: theoretical and analytical framework for activity analysis that focuses on understanding an individual's lived experience by examining their actions in context, including their intentions, concerns, perceptions, and adaptations over time.
- Engagement: what motivates or concerns a person to participate (act, speak, and think).
- Pragmatism paradigm: a paradigm that is not committed to any sort of philosophical stance and oriented toward what works.
- Self-confrontation interviewing: one-on-one discussion in which participants are shown excerpts of recordings from a recent activity to stimulate reflection and verbalization about their actions and experiences.

Participant engagements in co-design initiatives, particularly in the digital health domain, are crucial but come with inherent challenges. These challenges include the multifaceted nature of digital health implementations, the lack of empirical evaluations [8], and the need for a precise understanding of engagement based on participants' perspectives and theoretical frameworks.

Previous Work

The concept of engagement is often blurred with involvement [8,24-26] or remains unclear [7,27,28]. In both cases, engagement seems to be defined by how actors were involved in the projects. A total of 3 studies provide concise definitions, positioning engagement as involvement: "actively collaborate" [29]; "partnership, collaboration or consultation" [13]; and engagement type: co-production, co-design, and co-creation [30]. These are extrinsic considerations, not necessarily focusing on the motivations or concerns from the actors' perspectives.

Jolliff et al [31] used an existing framework proposed by the Patient-Centered Outcomes Research Institute [32], including extrinsic but also intrinsic considerations, such as engagement quality (perceptions of actors) and partners' outcomes (effects for actors). Peters et al [33] also associated quality of engagement with emotional and cognitive factors.

Some studies report on engagement without referencing co-design or participative design, instead focusing on patient partners [27] or patient and community involvement in research [7,13,29]. While actors were involved in these projects, they were not engaged in the co-creation of the solution, which does not align with our perspective on co-design. Empirical studies focusing on engagement [7,25,27], along with literature reviews on engagement [8,13,26,29], provide insights into factors affecting engagement and offer recommendations based on the results.

Among the problems affecting engagement, power distribution or tokenism (symbolic participation) was reported in all papers except in O'Connor et al [26], whose literature review included only 4 papers discussing co-design. Papers also mentioned:

- Levels of expertise, lack of skills, or limited knowledge [7,13]
- Unclear contribution to the project [13]
- Expectation of the degree and level of involvement not clearly articulated, leading to a lack of motivation [7]
- Communication issues [7,13]
- Personal life, values, and the engagement and recruitment approach [26]
- Limited availability of actors, lack of awareness and understanding of guiding frameworks and validated methods, and resource constraints (time and money) for adequate planning [29]

XSL•FO
RenderX

Papers mentioned positive emotional outcomes for actors (targeted end users), such as increased confidence in their daily lives [8], a sense of pride and accomplishment, feeling empowered, an increased knowledge and skills [8,13]. Most papers conclude with recommendations for future research, including:

- Clear communication [7,8,27]
- Clear end-user expectations and clearly defined roles for all parties involved in co-design [8,29]
- Engaging actors as early as possible and continuing engagement throughout the process [7,28-31,33]
- Allocating sufficient time to achieve objectives [8,27]
- Fostering a positive environment and mutual respect [29]
- Providing orientation, support, education, and training to help actors participate [8,27,29,31]
- Authentic design participation focused on "designing with" rather than "designing for" [25]
- Evaluation of engagement [29,31,33]

There was still a lack of empirical results or insufficient reporting of methodology across all the papers. Most papers did not provide detailed information about the co-design activities, except for reviews of wireframe workshops [7] and a design charrette [25]. Slattery et al [8] noted that the most frequent activities in their review were interviews or advisory councils, which do not explain how actors contributed to the co-creation process.

Objective of This Study

Despite the critical role of participant engagement in co-design for achieving fruitful outcomes, discussions on engagement remain predominantly focused on extrinsic factors, blurred with involvement, and are poorly empirically described. While the insights are useful for optimizing the co-design process, connecting them to participants' motivations or concerns from their perspective remains challenging.

In this study, engagement is considered differently, based on what participants found significant from their own perspective, aligned with the course-of-action framework [21]. This article aims to address this gap by providing an intrinsic description of participant engagements in a co-design situation.

Methods

Study Design

Overview

We used a mixed-methods approach [34], within the pragmatism paradigm [35]. In our study, the pragmatism paradigm provided various perspectives on participants' engagements, each building upon the other [36]. Most of the methodology was guided by the course-of-action framework [21].

Engagement in the Course-of-Action Framework

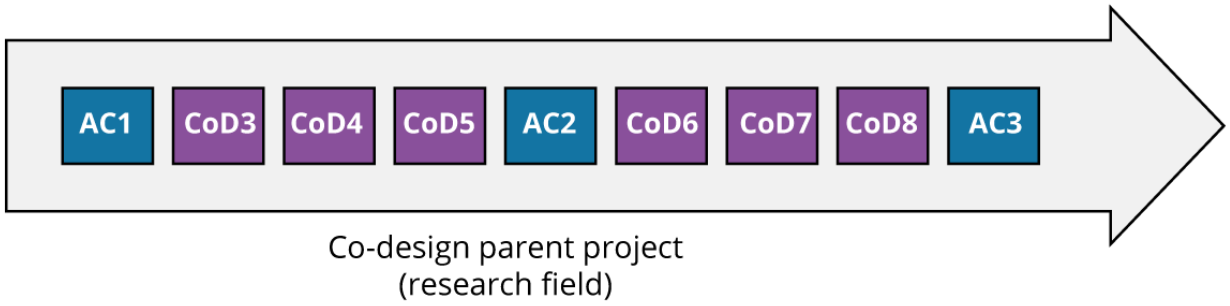
The course-of-action provides a strong theoretical and analytical framework for documenting participant experiences in digital health co-design initiatives [37]. Within the course-of-action framework, engagement represents the actor's significant concerns or motivations related to the elements considered in the situation [38]. Engagements are shaped by participants' histories and what is meaningful or significant to them at that moment [39]. Engagement is also directly linked to expectations, as expectations modulate engagement. Concerns and motivations lead actors to expect that something will happen in relation to them. Engagement is also influenced by the referential (knowledge and experience) of the actors. In a co-design situation, we examine how the participants' engagements evolve, what motivates them to think, speak, or act (participate). Expectations related to engagements can hinder the actor's participation if they are engaged by other elements in the situation, diverting them from the task [40,41].

Study Setting: The Co-Design Parent Project

This engagement study is embedded in a larger multisite co-design parent project (from May 2017 to June 2018), whose objective was to develop a tool to facilitate the search of resources for caregivers of seniors experiencing loss of autonomy. A complete description of the co-design study protocol is available in Latulippe et al [42], and the results of the approach are presented in 3 articles [43-45].

Data collection was conducted as part of a co-design project aimed at designing a tool to facilitate resource searches for caregivers of seniors experiencing loss of autonomy. A total of 8 co-design sessions (CoDs) were held from May 2017 to June 2018. These sessions were interspersed with 3 advisory committee sessions (ACs) (see Figure 1).

Figure 1. Co-design sessions (CoD) and advisory committee sessions (AC) of the parent project.



XSL•FO
RenderX

The following describes the CoD workshops and the AC meetings:

- Regional CoD workshops (n=8): Conducted in 11 administrative regions of Québec, each workshop gathered local caregivers, health and social service professionals (HSSPs), and community-organization staff to surface region-specific needs and generate design ideas. As attendance was location-based, membership naturally differed from one CoD workshop to the next.
- AC meetings (n=3): Convened in Québec City, the AC also gathered local caregivers, HSSPs, and the community-organization staff. Its mandate was to ensure alignment between the progress of the tool's design and the data collected during the co-design sessions, resolve outstanding design questions, and steer iterative prototype development. The participants in the AC sessions were generally the same individuals, as all these sessions took place in the Quebec region.

The co-design process culminated in the creation of a prototype website with 2 main sections: one for caregivers to search for resources and another for service providers to offer resources. The prototype was the subject of a study based on the triangulation of 3 frameworks in usability and user experience [46].

Ethical Considerations

The parent co-design project received institutional review board approval before the first workshop began (Université Laval #2017 - 103, issued April 20, 2017). The initial certificate covered the core workshop activities and brief contextual questionnaires completed by HSSPs and community-organization staff; these questionnaires informed the broader design effort but are not analyzed in this engagement study.

An amendment adding video-recorded participatory observation and postsession self-confrontation interviews was submitted in June 2017 and was approved on July 17, 2017, after CoD1 (May 2017) and CoD2 (early July 2017) had already taken place. Consequently, no postsession interviews were conducted for those 2 early workshops; full data collection commenced with AC1 and CoD3 and continued through AC3 and CoD8. All participants involved in the amended protocol provided written informed consent specific to these additional procedures and received a nominal compensation of CAD $20 (US $16.45). Data were anonymized for this study.

Recruitment for This Study

Participants for this study were recruited during the sessions. A caregiver serving on the AC was interviewed twice—immediately after AC1 and again after AC3—to capture how his engagement trajectory evolved over the life of the project. Retaining both interviews enabled a longitudinal within-person analysis consistent with the course-of-action framework; discarding the second interview would have forfeited this analytic leverage.

Data Collection

This co-design project served as the research setting for our study, which aimed to describe the engagements of participants in a cyberhealth co-design project. Data collection occurred in 5 steps (see Textbox 2).

Textbox 2. The 5 steps of the data collection.

Ethnographic data collection: internet-based questionnaires for HSSPs and community organizations, along with semistructured interviews with caregivers. Conducted under the parent codesign project; not analyzed in this study.

- Participatory observation of the sessions: video and audio recordings (3 h) of the sessions to capture data on participants' activities during the session.
- Partial chronicles (extrinsic description): the day following the session, a synthesized transcription of the complete session was made using the recordings to reconstruct the sequence of events with the help of time markers. Significant moments for participants (interpreted by the researcher) were selected.
- Self-confrontation interviews: interviews of approximately 1 hour conducted mostly within 2 days after the activity. Participants were presented with portions of recordings representing key moments of the session, aiming to obtain verbalizations about their activities during the session. The partial chronicle was used to structure the interview with the temporal sequence of the session, allowing the researcher to question the participant and identify significant moments expressed by the participant. Video and audio recordings were made during the interviews.
- Complete chronicles: transcriptions of the self-confrontation interviews were used to complement the partial chronicle of the session with participants' verbalizations (intrinsic description).

Data Analysis

The analysis was conducted in 2 stages by the primary author. The first stage (deductive) involved reconstructing the participants' course-of-action in the complete chronicle. Engagements and expectations were mostly identified by paraphrasing participants' statements to clarify the researchers' understanding, while in some cases, verbatim excerpts were used to stay as close as possible to participants' experiences.

These documents were then used in the second stage (inductive), conducted using the qualitative and mixed methods analysis software MAXQDA 2020 (VERBI Software). Engagements were organized into emergent themes [47], describing the nature of the engagements.

We subsequently cross-referenced the themes with variables, including participant categories (caregivers, HSSPs, and organizations) and sessions. This step allowed us to generate tables presenting the frequency with each theme. This quantitative analysis was useful for selecting data based on what appeared significant to a larger number of participants, from their perspective, and for maintaining analytic validity by

XSL•FO
RenderX

guarding against potential biases when selecting data to present [47].

Results

Participants

The co-design project included 74 participants. We collected testimonies from 20 participants, including caregivers (n=9), HSSPs (n=4), and community workers (n=7; see Table 1). We conducted 21 interviews (1 AC participant was interviewed twice).

Table . Session's location and participants.

Session	Location	Participants
AC[a]1	Centre-du-Québec	• Caregiver (n=1) • HSSP[b] (n=1)
CoD[c]3	Saguenay	• Community worker (n=1)
CoD4	Bas-Saint-Laurent	• HSSP (n=1) • Community worker (n=1)
CoD5	Outaouais	• Caregiver (n=2) • Community worker (n=2) • HSSP (n=1)
AC2	Centre-du-Québec	• Community worker (n=1) • Caregiver (n=1)
CoD6	Montréal-Laval	• HSSP (n=1) • Caregiver (n=1)
CoD7	Mauricie	• Caregiver (n=2)
CoD8	Montérégie	• Caregiver (n=1) • Community worker (n=1)
AC3	Centre-du-Québec	• Community worker (n=2)

[a]AC: advisory committee meeting.

[b]HSSP: health and social service professionals.

[c]CoD: co-design session.

The AC sessions were conducted entirely in plenary, where participants were invited to express their positions on specific issues where consensus had not been reached during the CoD sessions. All CoD sessions began with a plenary presentation outlining the project, the objectives of the session, and other relevant information. Subsequently, participants were divided into small workshops of 3 to 4 people, each facilitated by a member of the research team and typically focused on a different subcomponent or angle of the session's main objective, allowing for complementary contributions to the overall session goal.

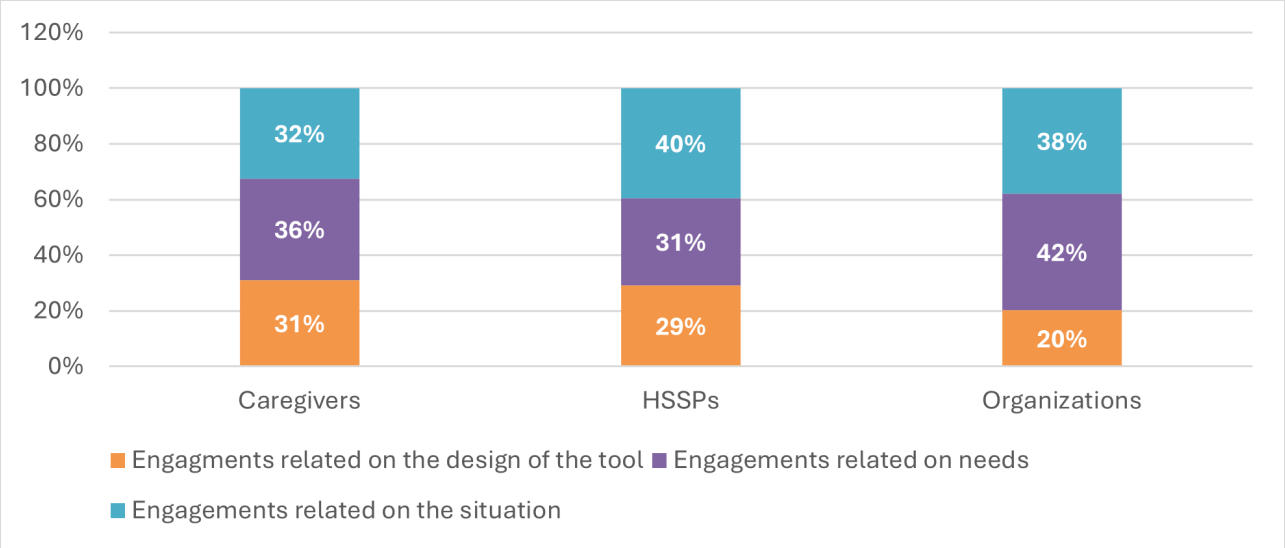
A distinctive feature of CoD5 was that all small groups worked collaboratively on the same specific task: creating a paper prototype. In contrast, other sessions (CoD1 to CoD4 and CoD6 to CoD8) explored the session's objective through more diverse or exploratory angles within subgroups. Participants had the freedom to choose their workshops, resulting in subgroups that were sometimes homogeneous and sometimes heterogeneous in terms of participant categories. Activities varied based on the session objectives, the progress of the tool design, and the specific goals of the subgroup workshops. The analysis revealed 3 main themes of participants' engagements: engagements related to their needs (personal and professional), engagements related to the situation, and engagements related to the design of the tool.

Nature of Engagements Per Participant Category

The distribution of engagement was quite similar among caregivers and HSSPs across themes (see Figure 2). However, the engagement of organizations (community workers) was significantly less focused on tool design.

XSL•FO
RenderX

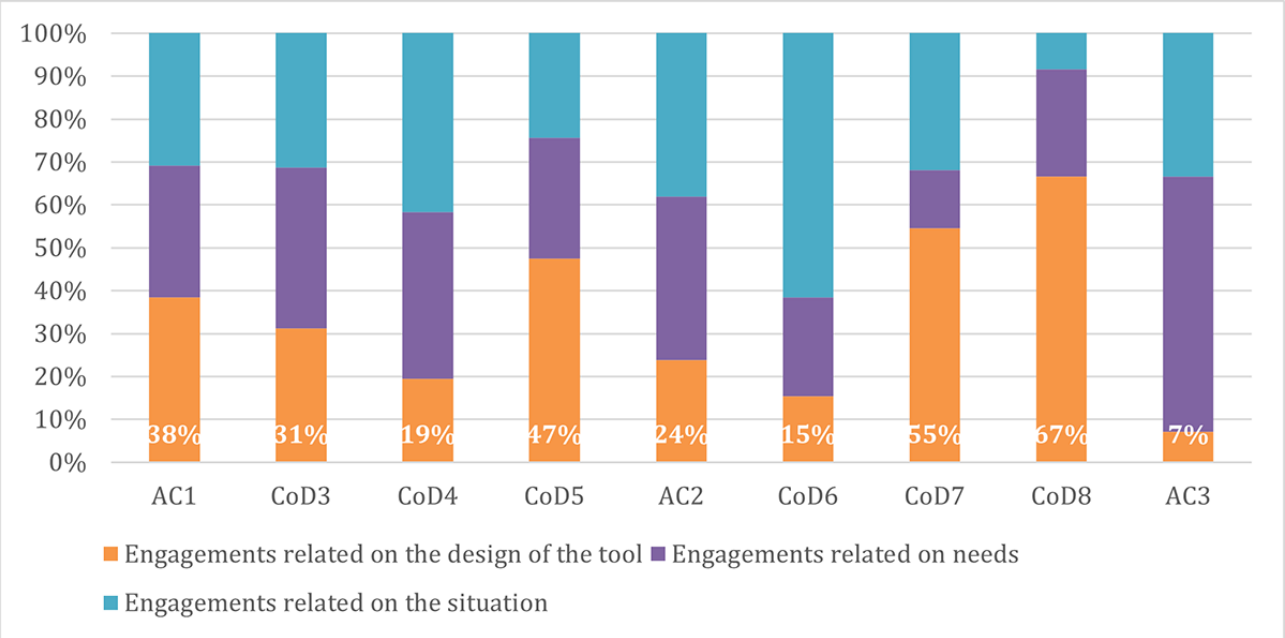
Figure 2. Distribution of engagement by participant category. HSSP: health and social services professional.



Nature of Engagements Per Session

Figure 3 shows that the nature of engagements also varied greatly throughout the project. For instance, engagements related to the design of the tool were more significant during specific sessions (CoD5, CoD7, and CoD8), while they were noticeably less important during other sessions (CoD4, AC2, CoD6, and AC3). A significant majority of the engagements related to the situation involved interaction with others, which is inherent to co-design, as interaction with others is a fundamental characteristic of the process. Figures 2 and 3 suggest that the nature of engagement varied depending on participants' categories and sessions.

Figure 3. Evolution of the nature of engagement by session. AC: advisory committee meeting; CoD: co-design session.



Example of Data Per Session

Below are examples of empirical data from each session (see Table 2), including session objectives and the methods and materials. The verbatim excerpts are enclosed in quotation marks.

XSL•FO
RenderX

Table . Session's location and participants.

Session	Objective	Methods and materials	Data examples
AC[a]1	Validate the design decisions made during the first 2 co-design sessions regarding user needs and address conflicting results.	Plenary discussion with a Microsoft PowerPoint presentation reporting the results of previous co-design sessions and addressing questions about conflicting results.	Situation: • Understand the process previously used to compile the list of needs (HSSP[b] 11 - 5). Design of the tool: • Determine what will work from the start and understand what caregivers are likely to use. (HSSP 11 - 5).
CoD3[c]	Identify user needs already addressed by other tools and determine the functionalities and content requirements related to those needs.	Comparison of existing cyberhealth tools for caregivers (2 websites, 1 app, and 1 video) conducted through small group workshops using a speed dating approach.	Professional needs: • Discover tools that can be suggested to the clientele (CW[d] 3 - 3). Design of the tool: • Avoid "losing" the caregiver through a complex and challenging resource search process (CW 3 - 3).
CoD4	Identify functional and content requirements for needs not addressed by existing tools.	Plenary brainstorming session (1.5 h) followed by small group workshops (1 h) using paper and pens.	Professional needs: • Ensure that everyone understands that for CW, the client is not the caregiver but the person receiving assistance (CW 4 - 2). • Ensure that the health services network can meet the demands of caregivers generated by the tool. "There are two problems. ONE, sometimes you're entitled to it, but you refuse the service, and TWO, there is a lack of time and resources to provide everything. If people become aware of the services, they will request them, and it will cost too much." (CW 4 - 2). Situation: • Redirect the discussion to the intended topic, as participants only had 3 hours, and the conversation had deviated from addressing the primary question (CW 4 - 2).

XSL•FO
RenderX

Session	Objective	Methods and materials	Data examples
CoD5	Design the information architecture.	Paper prototyping conducted in small group workshops using cardboard, pens, and small pieces of paper to represent options for functionalities and content requirements.	Design of the tool: • Avoid information overload by preventing it from becoming a "monster" and ensuring it remains "simple but understandable." (Caregiver 5 - 8). Situation: • Assist the caregiver in responding to the facilitating researcher who asked about a filter that could limit search results for resources in their region. He knows that older adults often do not understand computer logic; to them, a filter is akin to a coffee filter (CW 5 - 1).
AC2	Decisions on conflicting requirements.	Plenary discussion (3 h) involving both in-person and remote participants, using 2 documents: a 13-page table of results from all previous sessions and a 16-page summary table of requirements to meet the needs (results from the previous session).	Personal needs: • Understand the research project and provide relevant interventions, as she lacked sufficient information before the meeting to fully grasp the project (Caregiver 11 - 11). Situation: • Have a positive experience of participation. "I didn't fully understand my discomfort during the meeting, but I thought about it, and I realized I had become so anxious because I didn't want to mess anything up. I reflected a lot on whether I wanted to continue participating in the research because, for sure, with the way I had experienced it, I didn't want to anymore..." (Caregiver 11 - 11).
CoD6	Information design.	Plenary presentation and discussion (1 h), followed by small group brainstorming workshops (1 h and 45 min) using activity sheets for each group.	Personal needs: • Be prepared to achieve the objectives. She needed to make a choice but was unable to prepare because she had not received any information about the planned activities (HSSP 6 - 6). Situation: • Have sufficient time to achieve the objective. "At that moment, I knew we wouldn't finish our mandate. Instead of pushing forward and trying to say... at that point, I decided to disengage" (HSSP 6 - 6).
CoD7	Validation of the information architecture and interaction design.	Small group brainstorming workshops using activity sheets for each group and an interactive low-fidelity prototype (version 1).	

XSL•FO
RenderX

Session	Objective	Methods and materials	Data examples
			Situation: • Ensure that a common goal is pursued. "I felt that she [a CW] was attacking your project... [...] Meaning to say: 'You don't need to do that, we're already doing it!'" (Caregiver 7 - 7). Design of the tool: • Feel welcomed and guided by the homepage, "But caregivers need this... They're tired, they don't know where to go, they don't know what to do" (Caregiver 7 - 7).
CoD8	Comparison of homepage proposals, examination of the search results page, and validation of questions designed to help identify needs.	Small group brainstorming workshops followed by a usability discussion using an interactive low-fidelity prototype (version 2).	Design of the tool: • Ensure that visually, it immediately captures attention (Caregiver 8 - 9). • Avoid returning search results with no findings (Caregiver 8 - 9).
AC3	Decisions on conflicting results. Obtain feedback on the latest version of the prototype before website programming.	Presentation of a medium-fidelity prototype (version 3) in small groups (30 min), followed by a plenary discussion (2 h).	Professional needs: • Ensure that the language and terms used are appropriate. "I always have a concern about the language and terms that we use." (CW 11 - 6). Situation: • Take a stance without fear of judgment. "We're taking a stance on this, and that's okay. We don't have to be judged for it. [...] Can we have our approaches without getting jumped on?" (CW 11 - 6)

[a]AC: advisory committee meeting.

[b]HSSP: health and social service professionals.

[c]CoD: co-design session.

[d]CW: community worker.

Discussion

Principal Findings

Our objective was to obtain an intrinsic description of engagement, encompassing what motivates and concerns actors in a co-design situation, from their perspective. The chosen methodology, grounded in a pragmatic paradigm and informed by the course-of-action framework, allowed us to foreground participants' lived experiences and situated knowledge. This approach privileged perspectives that are often underrepresented in more top-down design processes. By adopting a co-design approach supported by self-confrontation interviews, we were able to access not only what participants said during the sessions but also how they made sense of their own engagement. These perspectives influenced both the iterative refinement of the design and our interpretation of the engagement dynamics.

The results might not appear surprising, yet it seems we are still encountering the same challenges in co-design projects. The contribution of this paper is to emphasize these results with a detailed description of the methodology and empirical data evaluating participant engagement, as recommended by Manafo et al [29] and Slattery et al [8]. This includes describing what was significant in participants' overall experience, from their perspective. Our analysis highlights that the nature of participants' engagement during the sessions went far beyond the design of the digital health tool for caregivers, even though that was the reason they were invited to participate. This aspect has not been discussed in other studies.

Comparison With Previous Work

Among the commonly discussed findings, our study revealed that actors need preparation (personal need) to understand their role, which enhances their sense of relevance and their ability to contribute meaningfully [7,8,27,29,31]. Service providers were also engaged in increasing their knowledge (professional

XSL•FO
RenderX

needs) [8,13]. A unique finding was the engagement of community workers, who sought to ensure their distinct concerns—differing from those of HSSPs—were understood. This is particularly relevant in Quebec, where community organizations operate under a model of entrepreneurial governance, with private foundations influencing their mission and work [48]. These engagements related to professional needs eventually led to engagements related to the situation (interaction with others).

Several engagements related to the situation mirrored those identified by Zowghi et al [7]. Unlike common issues of power distribution reported in other studies, our research did not encounter such conflicts between researchers and participants. However, we observed tensions between HSSPs and community workers, particularly during the final session (AC3), highlighting the need for a positive environment, mutual respect [29], and a shared vision [8]. Participants emphasized the importance of having sufficient time to achieve objectives [8,27].

The Nature of Activity as an Affordance of Action and Engagement

The concept of affordance [49] highlights perceptual cues in the environment that suggest specific interactions. From this perspective, we propose that the activities offered to participants shaped the nature of their engagement in this project. Plenary discussions generally fostered engagements related to personal and professional needs, as well as situational aspects. In contrast, design-centric activities primarily directed engagements toward the tool's design.

Participants in our study who were more engaged in the tool's design participated in typical design activities, such as paper prototyping (see Figure 4), mockups, clickable prototypes, and usability testing. These activities, which enhance creativity and direct involvement, acted as affordances by guiding appropriate actions, driving engagement, and directing participation (acting, thinking, or speaking) toward the tool's design. Effective facilitation of co-design sessions should emphasize design practices [50,51], aligning with best practices in this field, as reviewed by Mosely et al [52]. This approach supports the "designing with" philosophy rather than "designing for," as suggested by Howard and Somerville [25].

Figure 4. Paper prototyping activity.



Establishing an Engagement-Focused Approach

Grasping the essence of engagement is key to understanding participation in co-design, a core element of participatory research. However, the question arises: participation in what, exactly? From a researcher's viewpoint, participation implies committed involvement in co-designing a product or service. Yet, our findings suggest that participants perceive participation differently, emphasizing the diversity of engagement. This includes concerns, situational considerations, and the mobilization of their participation. In terms of research ethics and co-design, we should consider engagements that prioritize participants' needs and the situation. Adopting a compassionate stance to meet their expectations allows them to derive benefits from engagements that appear more significant to them. Based on our results, we recommend that researchers and designers using a co-design approach establish conditions that respect the diverse nature of participant engagements. To meet expectations tied to participants' personal needs and enable them to benefit from their involvement [22], we propose exploring the following avenues (see Textbox 3).

XSL•FO
RenderX

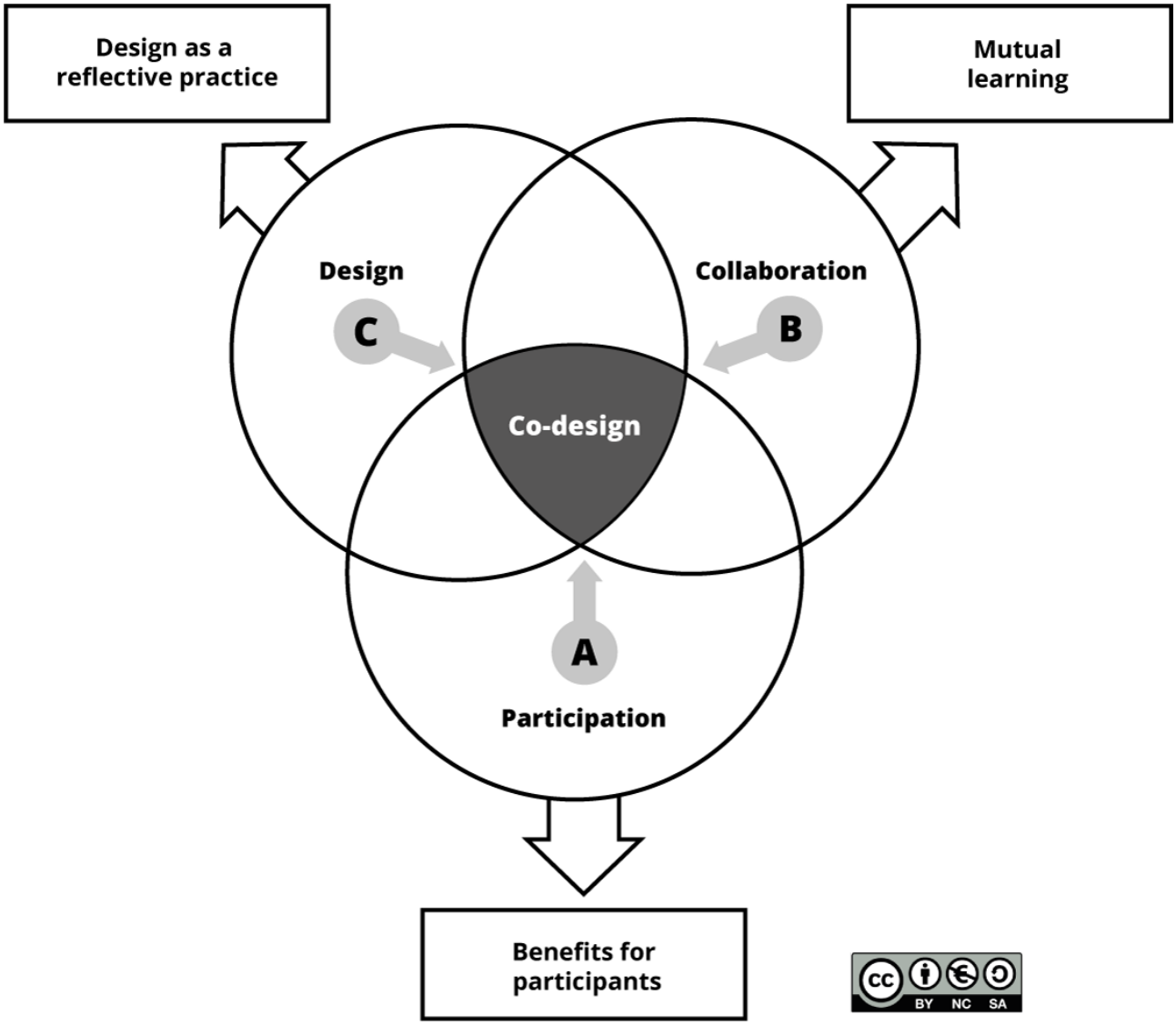
As mentioned earlier, co-design, for us, lies at the intersection of 3 dimensions: design, collaboration, and participation. These dimensions are closely related, resembling a Venn diagram. The tensions between these dimensions will impact the center (co-design). The suggested affordances (A, B, and C) could help balance these tensions (see Figure 5).

Textbox 3. Affordances that enable participants to benefit from their participation.

1. Consider participants' personal and professional engagements:
   - Recognize that participants pursue personal and professional goals that may extend beyond the intended design objectives for which they were invited to contribute.
   - Allow participants to express their personal and professional engagements (or expectations) by dedicating a specific period during sessions.
   - Endeavor to address expectations related to personal and professional engagements to honor the goals they pursue through their participation.
2. Address expectations related to the situation:
   - Ensure a shared objective from the outset by identifying and including all interest holders from the beginning.
   - Embrace a flexible approach to achieve objectives by allocating sufficient time and offering a satisfactory experience to participants.
   - Enable participants to comprehend the discussed elements by ensuring they have the necessary knowledge, such as by sending them a preparatory sheet before the session, for example, as the flipped classroom approach in education.
   - Ensure the maintenance of a positive environment characterized by mutual respect during all sessions.
3. Configure activities as affordances that suggest appropriate actions:
   - Prioritize design activities that guide participants' actions toward tool design.
   - Prompt participants to engage in action quickly during sessions (active mode) and avoid prolonged presentations (passive mode).
   - Adopt best practices in design facilitation.

XSL•FO
RenderX

Figure 5. Balancing the tensions between the dimensions of co-design.



Limitations

Participants varied in each co-design session. Analyzing the engagement of individuals who participated throughout the entire process using the course-of-action framework might yield different results. Comparing our findings with those of continuous participation could enrich the discussion on engagement by exploring the similarities and differences arising from the methodological choice of isolated versus continuous or multiple participation.

Conclusions

In this study, we explored the concept of engagement among 3 categories of potential users—caregivers, HSSPs, and organizations—who participated in a cyberhealth co-design process. We used a robust methodological approach based on the course-of-action framework to document participant engagements, addressing a gap in the scientific literature [8]. Our findings reveal that engagement extends beyond the immediate goals of tool design, encompassing broader concerns and motivations shaped by individual needs and the dynamics of interaction. In my professional co-design practice, I have integrated these recommendations by adopting a compassionate stance that prioritizes participants' perspectives and by configuring activities as affordances to support meaningful engagement. This approach has consistently led to more inclusive and impactful co-design processes, as evidenced by the quality of participation and the relevance of the outcomes generated. Our analysis highlights the importance of balancing design, collaboration, and participation to address the inherent tensions in co-design projects. Ultimately, this study underscores the value of tailoring co-design practices to respect and leverage the unique contributions of all participants, thereby advancing both the process and outcomes of collaborative innovation.

XSL•FO
RenderX

Acknowledgments

This work was supported by Le Fonds de recherche du Québec—Société et culture (2020-B2Z-273824, 2019), Centre d'excellence sur le vieillissement de Québec (2019), and Centre de recherche sur les soins et les services de première ligne de l'Université Laval (2020).

Conflicts of Interest

None declared.

References

1. Ronquillo Y, Meyers A, Korvek SJ. Digit health. In: StatPearls: StatPearls Publishing; 2023. URL: http://www.ncbi.nlm.nih.gov/books/NBK470260 [accessed 2025-08-20]
2. Broekhuis M, van Velsen L, Peute L, Halim M, Hermens H. Conceptualizing usability for the ehealth context: content analysis of usability problems of eHealth applications. JMIR Form Res 2021 Jul 27;5(7):e18198. [doi: 10.2196/18198] [Medline: 34313594]
3. Mohr DC, Lyon AR, Lattie EG, Reddy M, Schueller SM. Accelerating digital mental health research from early design and creation to successful implementation and sustainment. J Med Internet Res 2017 May 10;19(5):e153. [doi: 10.2196/jmir.7725] [Medline: 28490417]
4. Henni SH, Maurud S, Fuglerud KS, Moen A. The experiences, needs and barriers of people with impairments related to usability and accessibility of digital health solutions, levels of involvement in the design process and strategies for participatory and universal design: a scoping review. BMC Public Health 2022 Jan 6;22(1):35. [doi: 10.1186/s12889-021-12393-1] [Medline: 34991540]
5. Abelein U, Paech B. Understanding the influence of user participation and involvement on system success – a systematic mapping study. Empir Software Eng 2015 Feb;20(1):28-81. [doi: 10.1007/s10664-013-9278-4]
6. Bano M, Zowghi D. User involvement in software development and system success: a systematic literature review. Presented at: EASE '13: Proceedings of the 17th International Conference on Evaluation and Assessment in Software Engineering; Apr 14-16, 2013; Porto de Galinhas, Brazil.
7. Zowghi D, da Rimini F, Bano M. Problems and challenges of user involvement in software development. Presented at: EASE '15: Proceedings of the 19th International Conference on Evaluation and Assessment in Software Engineering; Apr 27-29, 2019; Nanjing China. [doi: 10.1145/2745802.2745810]
8. Slattery P, Saeri AK, Bragge P. Research co-design in health: a rapid overview of reviews. Health Res Policy Sys 2020 Dec;18(1):17. [doi: 10.1186/s12961-020-0528-9]
9. Lalonde S. Advocating for a more democratic process a critical analysis of codesign plenary sessions within a public project development. The Ohio State University. 2020. URL: https://etd.ohiolink.edu/acprod/odb_etd/etd/r/1501/10?clear=10&p10_accession_num=osu1598036287893528 [accessed 2023-11-25]
10. Library of facilitation techniques and workshop activities. SessionLab. URL: https://www.sessionlab.com/library [accessed 2021-12-02]
11. Untools. URL: https://untools.co/ [accessed 2023-11-25]
12. Design Kit. URL: https://www.designkit.org/methods [accessed 2021-08-16]
13. Brett J, Staniszewska S, Mockford C, et al. Mapping the impact of patient and public involvement on health and social care research: a systematic review. Health Expect 2014 Oct;17(5):637-650. [doi: 10.1111/j.1369-7625.2012.00795.x] [Medline: 22809132]
14. Sanders EBN, Stappers PJ. Co-creation and the new landscapes of design. CoDesign 2008 Mar;4(1):5-18. [doi: 10.1080/15710880701875068]
15. Visser W. Schön: Design as a reflective practice. Collection 2011;2:21-25 [FREE Full text]
16. Robertson T, Simonsen J. Participatory design: an introduction. In: Simonsen J, Robertson T, editors. Routledge International Handbook of Participatory Design: Routledge; 2013. [doi: 10.4324/9780203108543]
17. Bratteteig T, Wagner I. What is a participatory design result? Presented at: PDC '16: Proceedings of the 14th Participatory Design Conference: Full papers - Volume 1; Aug 15-19, 2019; Aarhus, Denmark. [doi: 10.1145/2940299.2940316]
18. Bratteteig T, Wagner I. Disentangling power and decision-making in participatory design. Presented at: PDC '12: Proceedings of the 12th Participatory Design Conference: Research Papers - Volume 1; Aug 12-16, 2012; Roskilde Denmark. [doi: 10.1145/2347635.2347642]
19. Bratteteig T, Wagner I. Kinds of Decisions: The Case of Urban Planning - The Case of Collaborative Syptom Assessment: Springer; 2014. [doi: 10.1007/978-3-319-06163-4]
20. McCarthy J, Wright P. Taking [a]Part: The Politics and Aesthetics of Participation in Experience-Centered Design: The MIT Press; 2015. [doi: 10.7551/mitpress/8675.001.0001]
21. Theureau J. Le Cours d'action: Méthode Développée [Book in French]: Octarès; 2006.
22. Robertson T, Wagner I. Engagement, representation and politics-in-action. In: Simonsen J, Robertson T, editors. Routledge International Handbook of Participatory Design: Routledge; 2013:64-85.

XSL•FO
RenderX

23. Gaete Cruz M, Ersoy A, Czischke D, van Bueren E. Towards a framework for urban landscape co-design: linking the participation ladder and the design cycle. CoDesign 2023 Jul 3;19(3):233-252. [doi: 10.1080/15710882.2022.2123928]
24. Bird M, McGillion M, Chambers EM, et al. A generative co-design framework for healthcare innovation: development and application of an end-user engagement framework. Res Involv Engagem 2021 Mar 1;7(1):12. [doi: 10.1186/s40900-021-00252-7] [Medline: 33648588]
25. Howard Z, Somerville MM. A comparative study of two design charrettes: implications for codesign and participatory action research. CoDesign 2014 Jan 2;10(1):46-62. [doi: 10.1080/15710882.2014.881883]
26. O'Connor S, Hanlon P, O'Donnell CA, Garcia S, Glanville J, Mair FS. Understanding factors affecting patient and public engagement and recruitment to digital health interventions: a systematic review of qualitative studies. BMC Med Inform Decis Mak 2016 Dec;16(1). [doi: 10.1186/s12911-016-0359-3]
27. Hamilton CB, Hoens AM, McKinnon AM, et al. Shortening and validation of the Patient Engagement In Research Scale (PEIRS) for measuring meaningful patient and family caregiver engagement. Health Expect 2021 Jun;24(3):863-879. [doi: 10.1111/hex.13227] [Medline: 33729634]
28. Kilfoy A, Hsu TCC, Stockton-Powdrell C, Whelan P, Chu CH, Jibb L. An umbrella review on how digital health intervention co-design is conducted and described. NPJ Digit Med 2024 Dec 23;7(1):374. [doi: 10.1038/s41746-024-01385-1] [Medline: 39715947]
29. Manafo E, Petermann L, Mason-Lai P, Vandall-Walker V. Patient engagement in Canada: a scoping review of the "how" and "what" of patient engagement in health research. Health Res Policy Syst 2018 Feb 7;16(1):5. [doi: 10.1186/s12961-018-0282-4] [Medline: 29415734]
30. Kiss N, Jongebloed H, Baguley B, et al. Meaningful consumer involvement in cancer care: a systematic review on co-design methods and processes. JNCI Cancer Spectr 2024 Jul 1;8(4):kae048. [doi: 10.1093/jncics/pkae048] [Medline: 38897655]
31. Jolliff A, Holden RJ, Valdez R, et al. Investigating the best practices for engagement in remote participatory design: mixed methods analysis of 4 remote studies with family caregivers. J Med Internet Res 2024 Dec 3;26:e60353. [doi: 10.2196/60353] [Medline: 39626228]
32. Measuring what matters for advancing the science and practice of engagement. : Patient-Centered Outcomes Research Institute; 2025 URL: https://www.pcori.org/resources/measuring-what-matters-advancing-science-and-practice-engagement [accessed 2025-08-06]
33. Peters S, Guccione L, Francis J, et al. Evaluation of research co-design in health: a systematic overview of reviews and development of a framework. Implementation Sci 2024;19(1):63. [doi: 10.1186/s13012-024-01394-4]
34. Creswell JW. Research Design Qualitative, Quantitative and Mixed Methods Approches: Sage Publications; 2014.
35. Creswell JW, Plano Clark VL. Designing and Conducting Mixed Methods Research, 2nd edition: SAGE Publications; 2011.
36. Sanchez JI, Bonache J, Paz-Aparicio C, Oberty CZ. Combining interpretivism and positivism in international business research: the example of the expatriate role. Journal of World Business 2023 Feb;58(2):101419. [doi: 10.1016/j.jwb.2022.101419]
37. Tremblay M, Hamel C, Viau-Guay A, Giroux D. User experience of the co-design research approach in eHealth: activity analysis with the course-of-action framework. JMIR Hum Factors 2022 Aug 9;9(3):e35577. [doi: 10.2196/35577] [Medline: 35943783]
38. Sève C, Saury J, Theureau J, Durand M. La construction de connaissances chez des sportifs de haut niveau lors d'une interaction compétitive [Article in French]. Trav Hum 2002;65(2):159. [doi: 10.3917/th.652.0159]
39. Ria L, Sève C, Durand M, Bertone S. Indétermination, contradiction et exploration: trois expériences typiques des enseignants débutants en éducation physique [Article in French]. rse 2004;30(3):535-554. [doi: 10.7202/012081ar]
40. Haué JB, Bellu SL, Barbier C. Le véhicule autonome: se désengager et se réengager dans la conduite [Article in French]. activites 2020;17-1:0-25. [doi: 10.4000/activites.4987]
41. Carver CS, Scheier MF. On the structure behavioral self-regulation. In: Boekaerts M, editor. Handbook of Self-Regulation: Academic Press; 2000.
42. Latulippe K, Guay M, Éthier S, et al. Supporting the process of help-seeking by caregivers of functionally dependent older persons through electronic health: protocol for a multicenter co-design. JMIR Res Protoc 2019 Apr 26;8(4):e11634. [doi: 10.2196/11634] [Medline: 31025956]
43. Tremblay M, Latulippe K, Giguere AM, et al. Requirements for an electronic health tool to support the process of help seeking by caregivers of functionally impaired older adults: co-design approach. JMIR Aging 2019 Jun 7;2(1):e12327. [doi: 10.2196/12327] [Medline: 31518279]
44. Giroux D, Tremblay M, Latulippe K, et al. Promoting identification and use of aid resources by caregivers of seniors: co-design of an electronic health tool. JMIR Aging 2019 Aug 22;2(2):e12314. [doi: 10.2196/12314] [Medline: 31518284]
45. Latulippe K, Tremblay M, Poulin V, et al. Prioritizing the needs of caregivers of older adults to support their help-seeking process as a first step to developing an ehealth tool: the technique for research of information by animation of a group of experts (TRIAGE) method. JMIR Aging 2019 May 23;2(1):e12271. [doi: 10.2196/12271] [Medline: 31518269]
46. Tremblay M, Latulippe K, Guay M, et al. Usability of a co-designed ehealth prototype for caregivers: combination study of three frameworks. JMIR Hum Factors 2021 Aug 18;8(3):e26532. [doi: 10.2196/26532] [Medline: 34406123]

XSL•FO
RenderX

47. Miles MB, Huberman AM, Saldana J. Qualitative Data Analysis: A Method Sourcebook: Sage Publications; 2020.
48. Depelteau J, Fortier F, Hébert G. Les organismes communautaires au Québec Financement et évolution des pratiques [Web page in French]. Institut de recherche et d'informations socio-économiquess. 2013. URL: https://iris-recherche.qc.ca/publications/les-organismes-communautaires-au-quebec-financement-et-evolution-des-pratiques/ [accessed 2023-12-29]
49. Gibson JJ. The theory of affordances. In: The Ecological Approach to Visual Perception: Lawrence Erlbaum Associates, Inc. (LEA); 1986:127-141.
50. Vial S, Boudhraâ S, Dumont M. Human-centered design approaches in digital mental health interventions: exploratory mapping review. JMIR Ment Health 2022 Jun 7;9(6):e35591. [doi: 10.2196/35591] [Medline: 35671081]
51. Vial S, Boudhraâ S, Dumont M, Tremblay M, Riendeau S. Developing a mobile app with a human-centered design lens to improve access to mental health care (Mentallys Project): protocol for an initial co-design process. JMIR Res Protoc 2023 Aug 22;12:e47220. [doi: 10.2196/47220] [Medline: 37606978]
52. Mosely G, Wright N, Wrigley C. Facilitating design thinking: a comparison of design expertise. Think Skills Creat 2018 Mar;27(January):177-189. [doi: 10.1016/j.tsc.2018.02.004]

Abbreviations

AC: advisory committee session
CoD: co-design session
HSSP: health and social services professional

Edited by C Jacob; submitted 01.01.25; peer-reviewed by C Johnston, S Gazaway; revised version received 24.06.25; accepted 05.07.25; published 27.08.25.

Please cite as:
Tremblay M, Hamel C, Viau-Guay A, Giroux D
Optimizing Participant Engagement in Cyberhealth Co-Design: Course-of-Action Framework Analysis
JMIR Hum Factors 2025;12:e70772
URL: https://humanfactors.jmir.org/2025/1/e70772
doi:10.2196/70772

© Melanie Tremblay, Christine Hamel, Anabelle Viau-Guay, Dominique Giroux. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 27.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Clinician Perceptions of a Novel Multicomponent Digital Care Assistant and Support Program for People After Stroke or Transient Ischemic Attack (CAPS) for the Secondary Prevention of Stroke: Qualitative Study

Liam Pearce Allan[1,2], BSc; Jane Li[2], PhD; Tara Purvis[1], PhD; David Silvera-Tawil[2], PhD; Jan Cameron[1,3], PhD; Marlien Varnfield[2], PhD; Vanessa Smallbon[2], MEng; Julia Bomke[2], DN; Natasha A Lannin[4,5], PhD; Dominique A Cadilhac[1,6], PhD

[1]Department of Medicine, School of Clinical Sciences, Sub-faculty of Clinical and Molecular Medicine, Monash University, Level 2, Victorian Heart Hospital, 631 Blackburn Rd, Clayton, Australia

[2]Australian e-Health Research Centre, Commonwealth Scientific and Industrial Research Organisation, Brisbane, Australia

[3]Australian Centre for Heart Health, The Royal Melbourne Hospital, Melbourne, Australia

[4]Department of Neuroscience, School of Translational Medicine, Monash University, Melbourne, Australia

[5]Alfred Health, Melbourne, Australia

[6]Stroke Theme and Critical Care Research, Florey Institute of Neuroscience and Mental Health, University of Melbourne, Melbourne, Australia

Corresponding Author:
Dominique A Cadilhac, PhD
Department of Medicine, School of Clinical Sciences, Sub-faculty of Clinical and Molecular Medicine, Monash University, Level 2, Victorian Heart Hospital, 631 Blackburn Rd, Clayton, Australia

Abstract

Background: We co-designed the novel multicomponent CAPS (Care Assistant and support Program for people after Stroke or transient ischemic attack) to augment the secondary prevention of stroke.

Objective: Following completion of a feasibility study, we sought feedback from Australian clinicians and service provider representatives (the potential deliverers of CAPS) regarding their perceptions of CAPS for secondary prevention and the pathways to real-world adoption.

Methods: This was a qualitative descriptive study of clinicians and service provider representatives involved in the delivery of stroke care around Australia. A pragmatic convenience sample was obtained by contacting previous CAPS co-design study participants; leveraging professional networks (eg, LinkedIn); and distributing study flyers and newsletters via primary health care networks, general practitioner (GP) networks, and social media posts (Commonwealth Scientific and Industrial Research Organization's LinkedIn pages). Semistructured interviews and focus groups were conducted virtually with clinicians and representatives of the Stroke Foundation (Australia). Qualitative content analysis was undertaken.

Results: Overall, 18 clinicians and 3 Stroke Foundation representatives participated from 5 Australian states, including medical specialists, GPs, nurses, and allied health professionals. We collected their perceptions of CAPS, categorized as potential benefits of the program for secondary prevention, and the considerations for facilitators and challenges to real-world program implementation. The perceived benefits of supporting self-management for patients and facilitating informed decision-making for clinicians were identified. Discussions regarding program implementation included program initiation and duration, patient support considerations, and workflow alignment, which involved consideration of the barriers and enablers to uptake within primary care practice and Stroke Foundation outreach support programs.

Conclusions: There was support from participants for the potential of CAPS to improve the secondary prevention of stroke. However, approaches for addressing the challenges raised by participants, including further implementation and integration considerations, such as sustainability of the model of care, are likely required for CAPS to be successfully embedded within clinical settings.

(JMIR Hum Factors 2025;12:e72873) doi:10.2196/72873

KEYWORDS

stroke; secondary prevention; mobile health; telemedicine; pilot projects

XSL•FO
RenderX

Introduction

Globally, stroke is a leading cause of death and disability [1]. In Australia, a stroke occurs every 11 minutes, with over 400,000 people with a history of stroke now living in the community [2]. Those who survive their stroke are at high risk of a recurrent vascular event (13% - 40% cumulative risk at 10 years) [3,4], as are those who experience a transient ischemic attack (TIA) [5]. This risk can be reduced through the use of secondary prevention strategies [6]. It has been estimated that up to 80% of recurrent events could be prevented through uptake of healthy lifestyle behaviors (eg, exercise) in conjunction with the use of medication (eg, antihypertensives) [7]. However, the uptake and adherence to these prevention strategies remain suboptimal. Barriers cited by survivors include issues of access to services, lack of awareness or education following stroke, and psychosocial difficulties [8,9]. Further exacerbating the issue is the limited support provided to people following their stroke and discharge into the community [10].

Using digital health as a modality to deliver secondary prevention programs may be a method to improve long-term support and patient outcomes. Certain technologies (eg, smartphone apps) have been associated with improved management of chronic health conditions, including diabetes mellitus [11] and hypertension [12]. However, evidence for the use of digital health to improve the secondary prevention of stroke remains limited [13]. To ensure digital health programs are fit-for-purpose, end-user involvement in their design and development is crucial [14]. Engaging clinicians in this process ensures the resultant program is relevant to their work, aligns with their workflows, and may lead to better uptake in practice [15].

Consequently, we developed the novel CAPS (Care Assistant and support Program for people after Stroke or transient ischemic attack). The program was co-designed with people living with stroke or TIA and clinicians from different disciplines involved in delivering stroke care [16]. This iterative process included assessment of user needs from both patient and clinician perspectives, and user experience workshops. The resultant multicomponent program combines clinician-led, person-centered goal setting with the person living with stroke and facilitates risk factor monitoring through a remotely accessible clinician portal. Behavior change is supported through SMS text messages aligned to goals. Patient end users also access a purpose-built smartphone CAPS app and a commercial wearable device for tracking health-related measures.

The feasibility of CAPS was evaluated in a 12-week, single-group, open-label, pretest-posttest pilot study among 33 people after stroke or TIA [17]. The program was highly acceptable and satisfactory to people living with stroke or TIA, with evidence of potential improvements in cardiovascular and psychosocial health [17]. Participant feedback was generally focused on program modifications to improve usability.

In addition to seeking input from people living with stroke or TIA [17], it was important to collect feedback from clinicians and health service provider representatives as the potential deliverers of CAPS, if found to be effective. While patient experiences and perspectives regarding digital health are frequently explored, further work is required to understand clinicians' perspectives and experiences [18]. Therefore, we sought to understand the perceptions of clinicians and service provider representatives regarding the overall program for the secondary prevention of stroke, as well as the potential pathways to the future adoption of CAPS into real-world use as a health promotion program. In this preimplementation study, we present the synthesis of the participants' feedback to guide future development of CAPS and explore potential future models of care.

Methods

Study Design

A qualitative descriptive design was adopted, using semistructured interviews and focus groups. Methods and results are reported against the Consolidated Criteria for Reporting Qualitative Research (COREQ) (Checklist 1) [19]. This study was conducted following the feasibility study in which the digital health program was evaluated among people living with stroke or TIA (September 2022 to May 2023) [17]. A snowball recruitment process was used, where feedback from the first group of participants informed the selection of the appropriate cohort for the second group.

Ethical Considerations

The interviews and focus groups conducted as part of this study were approved by the Monash University Human Research Ethics Committee (30521) and the Commonwealth Scientific and Industrial Research Organization (CSIRO) Health and Medical Human Research Ethics Committee (2021_118_RR). Consent was collected electronically prior to commencement. All data were stored securely according to institutional guidelines. Participants were provided a gift card to acknowledge their contribution; amounts were recommended by our institution and stakeholder advisory groups. To maintain privacy, participants were given pseudonyms for data analysis and reporting.

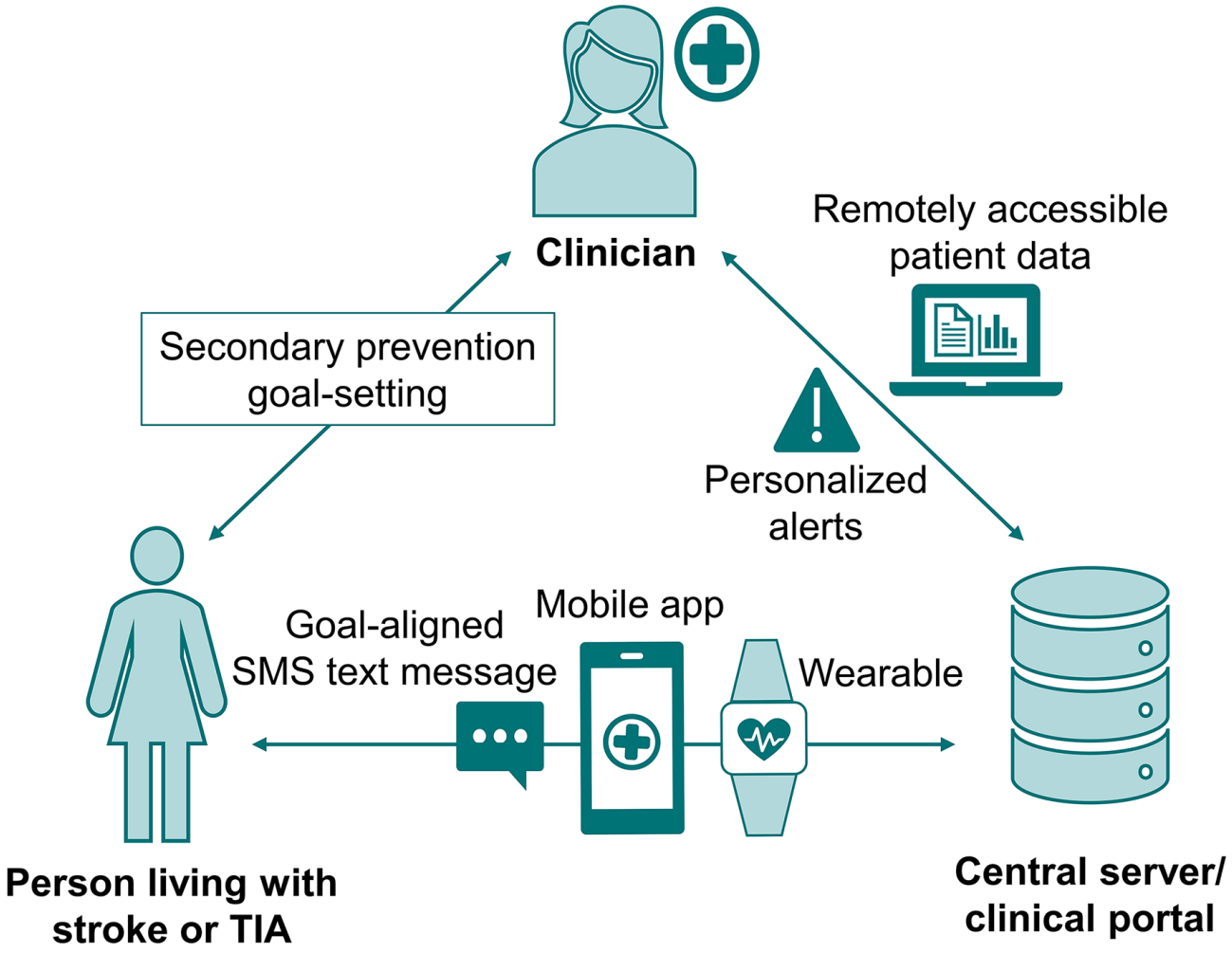
CAPS Components

The CAPS components are described in detail in previous publications [16,17]. In brief, participants used a standardized process to set 1-2 secondary prevention and person-centered goals using the SMART (specific, measurable, achievable, realistic, and timebound) format [20]. The template and procedures for this process were adapted from a phase III randomized controlled trial (RCT), which involved using SMS text messages to support goal attainment across the domains of secondary prevention, activities of daily living, self-management, psychosocial health, and accessing health care [21,22]. The clinician-facing technology consisted of a remotely accessible web-based portal (Figure 1). Patient profiles were created and tailored using personalized health measurements and alert thresholds, upload of SMART goals, activation of SMS text messages aligned to goals adapted from the above study, and education links. Alert messages within patient profiles were activated for measures outside thresholds sourced from clinical guidelines (eg, systolic blood pressure

XSL•FO
RenderX

>140 mmHg). This process also included an automated email to the monitoring coordinator and SMS text messages to the participant. Patient-facing technology consisted of 3 components. The first was motivational goal-aligned SMS text messages, sent at a rate of 2 messages per goal per week, with some containing educational hyperlinks. The second was the purpose-built smartphone CAPStroke app, where patients entered health data daily (with different measures selected through the portal) and could review their health data, enter journal notes (audio, text, and images), activate medication reminders, and access education resources. The third was a Bluetooth-synchronized smartwatch or fitness tracker that passively collected patient data with automatic upload to the portal and app.

Figure 1. Components of CAPS (Care Assistant and support Program for people after Stroke or transient ischemic attack). TIA: transient ischemic attack.



Participants

A pragmatic convenience sample of different clinicians involved in the delivery of stroke care from various parts of Australia was invited to participate in this qualitative study. Several recruitment strategies were used. For the first group, we sent invitations to participants who were involved in the co-design of CAPS and had expressed a willingness to contribute further to this research, and also reached out to our existing professional networks (eg, LinkedIn) between May and June 2023. For the second group, various methods were used to recruit stakeholders who were identified by participants in the first group as important in providing or responding to features of CAPS. These included general practitioners (GPs) and representatives of the Stroke Foundation (Australia), who provided in-reach support services to people living with stroke, including advice on secondary prevention. The recruitment strategies for the second group included distributing study flyers, internal organization newsletters, and personal invitations over a 4-month period (January to April 2024) via primary health care networks (PHNs) (Tasmania & Gippsland), Monash Health GP Liaison Unit, Monash University Practice-based Research and Education Network (MonREN), and GoldNet Research. A marketing post was also made on CSIRO's LinkedIn pages targeting GPs, and individual GP clinics were contacted via email or directly through their website.

Data Collection

Online semistructured interviews and focus groups were conducted separately with clinicians and service provider representatives based on their availability and preference to meet. During semistructured interviews and focus group sessions, participants were provided with a background of the purpose of the study and then an overview of the program, including video demonstrations of the CAPStroke smartphone app and clinical portal. Following this, participants were asked

XSL•FO
RenderX

a series of open-ended questions regarding their perceptions of using CAPS for the secondary prevention of stroke and the potential adoption of CAPS into real-world use. The questions posed within the semistructured interview guide were adapted from previous work and subsequently refined with input from the research team. This was to ensure that all important concepts, specific to the project objective, were captured (Multimedia Appendix 1) [23,24].

For the first group, sessions were conducted remotely via the Cisco Webex video conferencing system (May and June 2023). Four researchers were involved in facilitating the sessions, with 2 researchers present in each (JC: female, PhD, clinician researcher; DS-T: male, PhD, digital health researcher; JL: female, PhD, digital health researcher; LPA: male, BSc, PhD candidate). For the second group, the sessions were conducted remotely via Microsoft Teams video conferencing (March and June 2024). Two researchers were involved in facilitating the sessions, one of whom was involved with the first group (TP: female, PhD, clinician researcher; JL). Some researchers were known to some participants from the co-design process or from professional relationships.

Sessions lasted approximately 60 minutes and were audio-video recorded with consent, using Webex or Microsoft Teams. Notes were taken by the researchers during each session. Audio recordings were professionally transcribed verbatim.

Rather than basing the sample size on data saturation, in this study, we focused on achieving adequate information power by using purposive sampling, having clear objectives, and obtaining rich, detailed interview data [25]. According to this framework, the more relevant the cohort is to the study aim and the more specific the analysis, the fewer participants are needed.

Data Analysis

To minimize potential bias, a semistructured interview guide was used, and coding was completed by 2 researchers (JL and LPA) who did not have prior relationships with the participants. Qualitative content analysis was undertaken by 2 researchers (JL and LPA), using NVivo software (QSR International) [26]. As this was a preimplementation study in which potential modifications to the program and future models of care were explored, this approach was used to systematically organize the data and generate practical insights [26]. Analysis of the first group occurred prior to recruitment of the second group, with all data combined. Preliminary initial codes were established based on the interview questions and feedback received during sessions using an integrated approach, which involved applying deductively derived codes based on the interview questions with inductively derived codes from session transcripts [27]. Coding was cross-checked and discussed between the 2 researchers prior to the second round of the coding process and writing of the results. For the second group, additional steps included discussing and reviewing initial notes taken during the interviews and focus groups among 3 researchers (TP, JL, and LPA). The categories and subcategories generated in the second group were reviewed by a third researcher (TP) and visualized using Miro software (RealtimeBoard, Inc). Coding was cross-checked and discussed among the 3 researchers prior to the second round of the coding process and writing of the results. Final coding for both groups was discussed with 2 additional authors (JC and DS-T), following the writing of the results. Disagreements were resolved through consensus.

Results

Participant Characteristics

Eighteen clinicians and 3 Stroke Foundation representatives participated (Table 1). Eleven participants attended one of the focus groups, ranging from 2 to 3 participants per session. Ten participants had one-on-one interviews. The participant flowchart is provided in Figure 2. Clinicians had various clinical backgrounds, including specialists, GPs, nurses, and allied health professionals. Clinician experience spanned the stroke continuum of care. They were working in various settings, including acute care (n=4), rehabilitation (n=3), and primary practice (n=9). One clinician also worked in community-based care as an adjunct to their acute role. Stroke Foundation representatives worked across various primary and secondary prevention services offered by the provider. All 3 representatives also had clinical backgrounds (nursing and allied health).

XSL•FO
RenderX

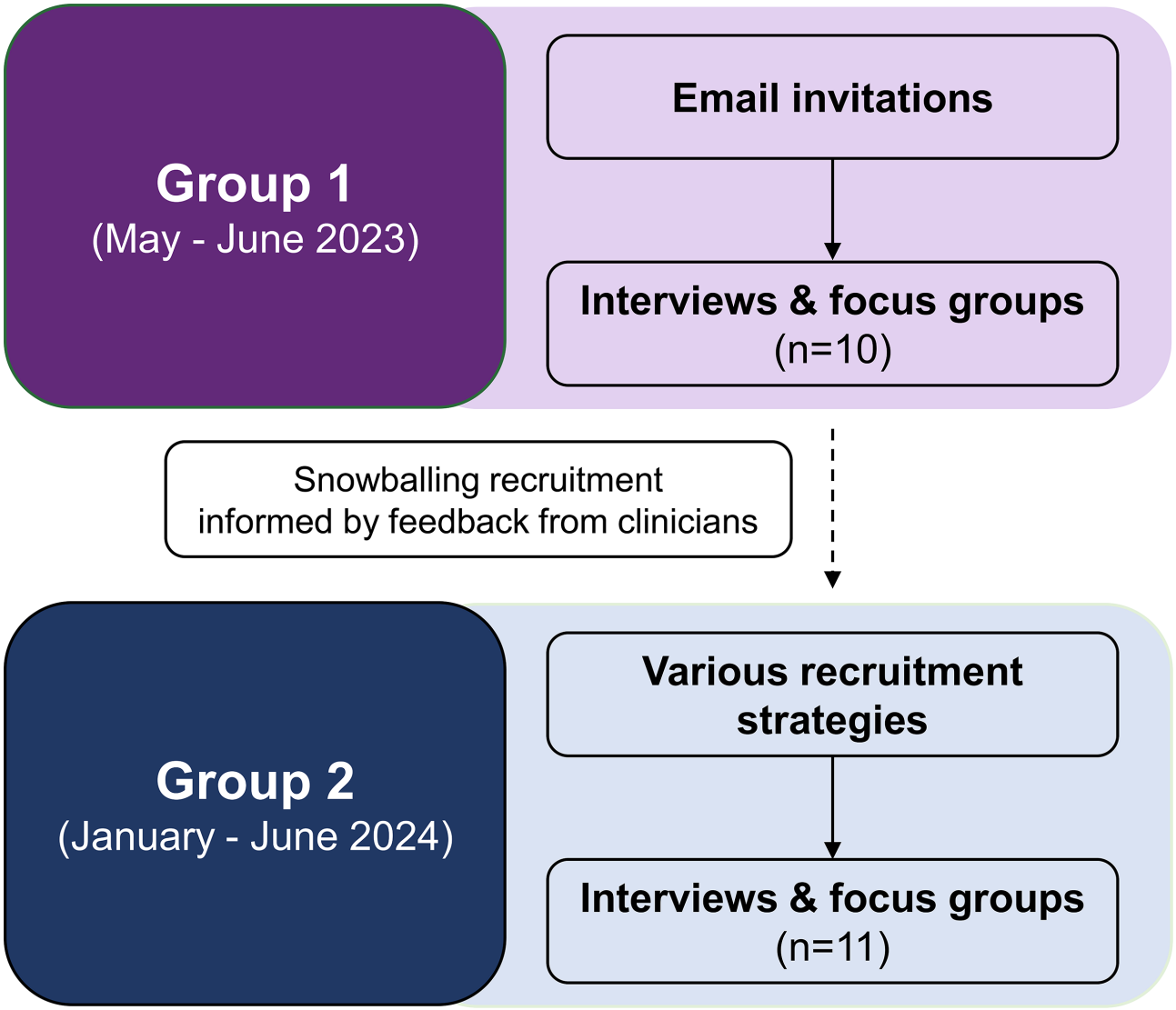
Table . Participant characteristics.

Characteristic	Value (N=21), n (%)
Female sex	11 (52)
Australian state of residence	
New South Wales	5 (24)
Queensland	4 (22)
South Australia	2 (10)
Tasmania	1 (5)
Victoria	9 (43)
Profession	
General practitioner	9 (43)
Medical specialist	
Neurologist	2 (10)
Stroke consultant	1 (5)
Nurse	
Registered nurse	2 (10)
Nurse practitioner	1 (5)
Allied health professional	
Occupational therapist	1 (5)
Physiotherapist	1 (5)
Speech pathologist	1 (5)
Stroke Foundation representative[a]	3 (14)

[a]Nurse, physiotherapist, and exercise physiologist.

XSL•FO
RenderX

Figure 2. Flow of participant recruitment for the first and second groups of the study.



Categories and Subcategories

We present the results and subcategories related to the categories of "*perceived benefits*" *of CAPS for secondary prevention* and *considerations for real-world "program implementation*," including facilitators and barriers to future program uptake (Table 2).

Table . Summary of categories and subcategories.

Major category	Subcategories
Perceived benefits	• Overall perceptions of the program and technology • Benefits to the patient - Supporting patient self-management and accountability - Appropriateness of the program technology • Benefits to the clinician - Facilitating informed decision-making
Program implementation	• Program initiation and duration • Patient support considerations - Encouraging patient ownership and accountability - Barriers to patient engagement • Workflow alignment within existing programs and models of care - Alignment with Stroke Foundation services - Alignment with primary practice • Practical factors affecting program delivery - Barriers to uptake in primary practice - Reimbursement and funding considerations • An independent model of care

XSL•FO
RenderX

Perceived Benefits

In this section, we describe the perceptions regarding the overall program and the potential benefits to the patient and the clinician.

Overall Perceptions of the Program and Technology

The majority of clinicians saw the value of CAPS as a "comprehensive" program that "integrates" multiple functions for health monitoring and tracking. They also praised the program as a novel product "a few steps above what is currently available in the market" for people living with stroke or TIA in the community. Many clinicians indicated that the technological components appeared to be "very user-friendly" and "well-designed," making them easy to navigate.

> *I'm part of a clinical entrepreneur program at the moment and […] there's so much happening in this space. I certainly haven't seen anything similar to this around secondary prevention of stroke and TIA.* [Participant #6, GP]

Some GPs also raised the wider applicability and potential use of CAPS in the primary prevention of stroke and across different chronic conditions, given that the risk factors for stroke prevention are relevant to many other disease states and health conditions.

> *It's almost like it could just be a generic patient thing, not even condition specific. And if you had enough individualized goals, I mean I think it could be nearly for everyone.* [Participant #18, GP]

Benefits to the Patient

Supporting Patient Self-Management and Accountability

The potential benefits of CAPS as an option for improving self-management and "optimizing" the secondary prevention of stroke through its emphasis on self-management were raised. Clinicians indicated that the program could be used to enhance postdischarge care, providing clear instructions, support, and information to patients in a "straightforward" way. This was seen as preferable to overwhelming them at a time when they are often feeling "lost" or vulnerable after discharge from an acute hospital. Participants saw CAPS as useful for encouraging patient self-management, with the data collection serving as a motivational tool to support accountability to themselves and to their health care provider. A GP suggested the program could provide a safety net, particularly for patients who might be deterred from presenting to their GP, and could help "pick up" patients with worsening health.

> *[CAPS could help] in one or two ways, [for the patient that has] kind of completely changed their ways and are highly motivated […] this type of program no doubt would facilitate that [self-management] for them. And then some patients obviously spiral the other way and get quite depressed and, you know, fall into a bit of a hole so, but this, you know, this might be useful to give them something tangible to hold on to.* [Participant #12, GP]

> *[…] to have this recommended, just so [patients] don't feel like they've just literally been kicked out the door and that their stroke didn't matter, I think it will help them to accept some of that and process some of those changes and actually set up good life skills going forward. I think it's - it just literally would be a safety blanket for them, but a positive safety blanket.* [Participant #1, registered nurse]

Appropriateness of the Program Technology

Feedback regarding the patient-facing components of the program indicated that the measures collected through the app from the daily check-in were "all valid" for patient management. However, a few of the specialists (a neurologist and stroke consultant) noted that including subjective measures, such as mood or distress, was less important for the secondary prevention of stroke and that objective physiological measurements, such as blood pressure, should be the main focus of the program, suggesting the daily check-in to be "pared back." Conversely, GPs stated that while those physiological parameters were very important, subjective psychosocial measures were also important in long-term patient management and well-being.

> *I think a lot of patients don't realize once you've had a stroke, there's a higher chance you'll be depressed, so I think it's good for them to be thinking about it, so maybe they can be more proactive about it as well.* [Participant #18, GP]

Clinicians also recommended tailoring the frequency of each measure for patients (rather than being exclusively daily) and taking advantage of the electrocardiogram features of some of the wearable devices for monitoring atrial fibrillation.

> *[smartwatches having] the capacity to do an ECG […] that's going to completely transform diagnosis of AF into the future.* [Participant #6, GP]

Benefits to the Clinician

The benefits of using CAPS for the clinician were discussed less often, with a few of the clinicians stating that the program would have more benefits for the patient than it would for the clinician.

Facilitating Informed Decision-Making

Most of the GPs mentioned that they could leverage the rich data provided by CAPS to improve patient care and better target the risk factors of concern. This included enhanced monitoring of psychosocial well-being, highlighted as an important factor in improving patient treatment "outcomes and compliance." Clinicians emphasized the value of accessing longitudinal data that can offer greater clinical insights than the limited data that might be collected in a single consultation (eg, blood pressure readings). One GP also indicated that the potential of improved patient outcomes would be motivating for clinicians to use the program for patient management.

> *I guess the things that I would find beneficial is just having more data […] having more blood pressure readings, having more information […] if they come in and they've got all this information just available for us, we can make much more informed decisions much quicker.* [Participant #18, GP]

XSL•FO
RenderX

The importance of presenting the data in an accessible way in the clinical portal was raised, for example, using a traffic light or matrix system to easily identify data outside of clinically recommended parameters (eg, hypertension and hypercholesterolemia). By doing so, the clinician could "very quickly" get an idea of the current health needs of the patient and use the "tangible" data to target areas of improvement (eg, blood pressure).

> *[…] I'd love to just be able to type in total LDL [low-density lipoprotein], HDL [high-density lipoprotein] ratio, triglycerides and then they can just sort of see […] how well are [they] in the green?* [Participant #10, nurse practitioner]

> *I don't want a data kind of avalanche coming our way, it would be more tangible bits of information that we can move the needle with.* [Participant #12, GP]

Program Implementation

A major focus of the focus groups and interviews was to understand the potential adoption of CAPS into real-world use. In this section, we present themes related to the optimal time for delivering such a program, patient-level considerations, and the identification of potential models of care. We also explore the facilitators and perceived barriers to program uptake at the clinician level.

Program Initiation and Duration

Clinician sentiments for a suitable program initiation time after stroke were varied, ranging from immediately upon discharge to slightly longer follow-up periods (ie, a few weeks to a few months after discharge). However, the general consensus was that patients would receive the most benefit from the program if delivered within 3 months of discharge from the hospital. This aligned with feedback provided by patient participants [17].

When queried, nearly half the clinicians noted that the program should be delivered upon or soon after discharge, particularly after having been seen by specialists such as neurologists and cardiologists. They felt this was an appropriate time given that patients and their family members are "quite motivated" to implement changes and are more "susceptible to behavior change" and to learning healthy "habits." This could be done by introducing CAPS as part of the Stroke Foundation's MyStrokeJourney, a resource provided to patients to support the "transition from hospital to home" [28].

> *CAPS really has value [delivered] very acutely after an incident, when patients are more likely to engage.* [Participant #14, GP]

However, the other half of the clinicians felt it would be more appropriate to introduce CAPS a few weeks to within 3 months after discharge, with 3 elaborating that waiting a period of time after discharge was ideal. This is because patients are often overwhelmed and need further information and support at this time.

> *[…] we do a three-month follow-up in all the strokes […] that to me that would be the sweetest spot to introduce this […] they are overwhelmed, there are multiple competing demands, multiple things to still get their head around. Three months the patient already comes with a list of questions, and I see that as a nice time.* [Participant #4, neurologist]

A more flexible model of care was also suggested by a few of the clinicians, with the possibility of introducing the program at any time to patients who were discharged directly from hospitals and who have been unable to access other support services. They also felt it would be appropriate to offer the program to patients who were not given enough education about the secondary prevention of stroke at the time of their event, as identified by their GP.

> *It really depends doesn't it, if it's a TIA, they've had a fright and they're very motivated to stop that from ever happening again or stop them from ever having a stroke. If they've had a major stroke then […] they've got a whole grief process and a rehab process, so I think it needs to be individualized […].* [Participant #6, GP]

Patient Support Considerations

Encouraging Patient Ownership and Accountability

Clinicians expressed clearly that the onus should fall on the patient to self-manage for secondary prevention, rather than assuming and relying on the oversight of clinicians, who are likely to be monitoring the data sporadically. This entails patients being "empowered" to take "ownership" of their health and being "confident" in knowing when to seek help from their primary carer. However, several GPs raised concerns about patients interpreting their clinical information without the support of a clinician, feeling it "opened up a can of worms."

> *[…] there has to be an expectation put on the patients around how much responsibility they take for their health; you know they can't completely load that back on the clinicians.* [Participant #12, GP]

> *I think […] most people are trying to take more responsibility for their health and something like this could definitely help them along the way. And I think allows them to […] take that ownership. I would think they would enjoy it.* [Participant #5, occupational therapist]

Barriers to Patient Engagement

Clinicians raised some challenges patients could experience with accessing and using the program. There were concerns that some people with stroke would be unable to access the program due to not owning the right technology and that some elderly patients may lack the technological literacy for the program. Additionally, the potential difficulties in providing equitable access to the program were emphasized, particularly in culturally and linguistically diverse communities, including Aboriginal and Torres Strait Islander people. However, it was raised by 3 clinicians that patients could be supported and trained to use the technology relatively easily.

> *Of course, there is going to be a proportion of patients who can't manage a smart phone, don't speak*

XSL•FO
RenderX

> *English, don't know their Apple ID [identification] or password. Those sorts of things are a challenge and those people are already the least well served of our population, particularly people who don't speak English.* [Participant #10, nurse practitioner]

> *I think some people might be a little IT [information technology] illiterate, but […] I wouldn't say that you couldn't teach someone how to do that pretty easily.* [Participant #18, GP]

Some of the clinicians noted that CAPS would predominantly benefit motivated patients, who would find the program "exceptionally helpful," but that less-motivated patients may struggle to engage. Similarly, there were discussions about the difficulty of patients to adhere to programs, with barriers to engagement, such as fatigue or depression, as well as previous challenges with requesting patients to complete data input for extended time periods.

> *[…] those people who are motivated will probably engage in this and those who probably need it are maybe less inclined to engage for reasons. That's always a battle.* [Participant #16, GP]

Workflow Alignment Within Existing Programs and Models of Care

Clinician opinions regarding who could best deliver the program varied. In the first group, most clinicians recommended a GP-shared model of care, feeling the program aligned with the purpose of primary practice. Most clinicians highlighted the benefits of including multidisciplinary care professionals (nurses and allied health professionals) for supporting secondary prevention. Given the current service provision by the Stroke Foundation, the alignment of CAPS with their existing programs was considered. In the second group, perceptions of the GPs regarding the potential model of care and workflow alignment varied, and several challenges to the program being GP-led were raised. This concept was explored with GPs, and further information about their perceptions is provided below.

Alignment With Stroke Foundation Services

Some of the services provided by the Stroke Foundation, relevant to CAPS, include the Living Well After Stroke program and the Stroke Navigator service. The first is a secondary prevention program delivered via telehealth to improve self-management and education for people after stroke. The Stroke Foundation representatives indicated that no objective measures (eg, blood pressure) are taken by the Living Well After Stroke program, with more focus on supporting participants to achieve their program goals and equipping them with a toolkit for self-management. In the second service, the Stroke Navigator, patients are contacted by staff following discharge from the hospital to support the access of various services for managing recovery and secondary prevention. Given this, the representatives were asked if they felt the Stroke Foundation could and would include CAPS as a service in their current model. The Stroke Foundation representatives indicated that while CAPS aligned with Living Well After Stroke and the purpose of foundation activities, there was currently no bandwidth for it to be incorporated as a Stroke Foundation–led service. However, 2 representatives felt that the Stroke Navigator service could be used as a referral or registration pathway for people living with stroke returning to life in the community. When queried regarding the placement of the program, 2 of the 3 representatives indicated the importance of partnering with GPs to deliver CAPS.

> *[…] that's their key, that's where their medical care, their recovery is mainly managed […] in general practice.* [Participant, Stroke Foundation representative]

Alignment With Primary Practice

Most GPs pointed out that CAPS aligned well with GP Management Plans currently used to facilitate chronic disease management (CDM) in primary practice [29]. These plans are described in further detail in the subsection *Reimbursement and Funding Considerations*. It was raised by 1 GP that frequent touchpoints with clinicians, as required in CAPS, which could be facilitated by aligning with these care plans, are highly beneficial for patients in programs such as these and can support continued "motivation" and "achieving goals."

> *[...] a lot of this is based on personalized goals... and this is how this could be integrated into...I could see how this could make those plans just amazing...and this program sort of goes for 12 weeks, and I was just thinking that would fit with those GP management plans so much better. And the goals that are in the app are almost exactly the goals that I write for my patients now […] adding more work is often not encouraged by GPs, but I think how this could integrate into general practice, that would be really useful.* [Participant #18, GP]

There was variability reported in the perceptions of who would potentially be involved in the various tasks of onboarding and delivering CAPS. Some GPs felt that practice nurses were best equipped to onboard participants into the program, including familiarizing them with the various technologies. It was also mentioned that nurses could support patients to set secondary prevention goals. One GP stated a preference for their practice nurse onboarding patients but they themselves [the GP] setting goals with the patient. Another GP noted that either the GP or a nurse could onboard patients, depending on the practice (eg, GPs in small practices without a nurse). One GP also raised the possibility of referring patients to an external service for registration and onboarding, and then having an "agreed shared care" model with the external entity.

> *I could see how a practice nurse involved in this could be really helpful, because they have a little bit more flexibility on time […] getting the nurses to set up this with the patient would […] take that pressure off the GP doing it.* [Participant #18, GP]

GP opinions regarding who was best equipped to monitor the data were split between GPs and practice nurses. Some indicated that both were potentially viable options in the interviews. Those of the opinion that it should rest with GPs felt that it should be their responsibility as they are the primary point of "medical care" and are managers of "recovery" for patients after stroke.

XSL•FO
RenderX

> *They'd [GPs] definitely be able to do the monitoring side of things.* [Participant #18, GP]

Others noted that the practice nurses had the "skill set" and more availability to monitor patients and would be able to "spend more time" than the doctors supporting patients. It was also mentioned that the nurses could then act as a "filter" or could triage patients and "flag" particular concerns with the GP when patients attended appointments.

> *[…] we do health assessments, data gathering, practice nurses are fully trained to do that in our clinic. We do care plans, […] falls assessments and so forth, things that take time and I don't have that. My practice nurses do it on my behalf.* [Participant #13, GP]

Practical Factors Affecting Program Delivery

Barriers to Uptake in Primary Practice

A number of challenges to embedding CAPS in primary practice were raised.

The majority of GPs highlighted the lack of time they currently have available with an "insurmountable" workload, without the addition of a "high touch" monitoring program. They expressed concerns about the increased workload, potential for missed alerts, and legal implications of managing the additional information, especially outside regular hours. Additionally, 1 clinician raised the lack of support for GPs, particularly those in rural locations.

> *[…] there's a whole other workload attached to GP workload with no more remuneration or care because if the public system is suffering, then it's, we have this finite plate of things we have to do and people just keep adding, adding, adding, there's CAPS, there's obstetrics, there's psych, there's – but the job is still, you know 40 hours a week, so unfortunately some things are going to fall off the plate if you keep adding without creating a new job or adding more capacity or hiring new GP's or nurses.* [Participant #17, GP]

Some of the GPs raised issues with timeliness in responding to alerts activated within the portal when patient measurements are outside guideline-based thresholds. This included consideration of alerts being activated during periods where GPs are away, "out of hours," or not accessing the information until an extended time afterwards. One GP shared a method for addressing this current issue in their practice, where out-of-hour patient telephone calls are forwarded to after-hours medical clinics, allowing patients to still be treated. However, this comes with the caveat that the after-hours clinics cannot access clinical information from the GP and vice versa.

> *[…] they sort of just get redirected to after-hours […] if they need to see a doctor then as well. There's pretty much no access to our notes, like those afterhours doctors can't see our file at all.* [Participant #18, GP]

Most of the GPs interviewed and one other clinician (registered nurse) had concerns regarding the duty of care and medico-legal responsibility associated with the expectation from patients that there is a clinician responding to alerts in a timely and appropriate manner. An example situation provided was an alert being activated while a GP is on leave or "out of hours," which does not get actioned and leads to a medical emergency, and the liability of the GP in that situation. Several GPs felt that adopting the program could lead to increased insurance premiums. These legal and cost implications were highlighted as deterrents to the uptake of the program.

> *I think in terms of, from the medical legal point of view, it opens up conditions to being liable for clinical information that comes across your desk, because if you see something that's abnormal and you're on leave or you're away and it doesn't get actioned and then the patient subsequently has a stroke or something like that then who is at fault?* [Participant #17, GP]

Interoperability with the current medical software was raised as critical in the adoption of CAPS by most of the GPs interviewed and another clinician (nurse practitioner). Two GPs raised the challenge of using an additional interface in the time-poor environment. To access and use the information collected in CAPS, the GPs stated that the data would need to be automatically "integrated" into the patient's medical file in the practice software. According to them, this could include "populating" the patient file in the practice management software or providing "downloadable data." One GP felt that being able to integrate with the practice software would be "game-changing." Another suggestion was to include a link to the CAPS clinician portal within the patient profile in the practice software for seamless access.

> *[…] as a GP I want it all in the one thing, I want it all in best practice, I want it all in medical director.* [Participant #18, GP]

> *I think as soon as you're asking GPs to go outside their medical software and log on to another web page, they're not going.* [Participant #12, GP]

Another concern raised by a few of the GPs was the management and storage of patient data on an external platform. Citing "privacy issues," they flagged that GPs often worry about sending data to third parties and desire clear "labeling" of how their patient data are being stored and used. One GP indicated that backing from a trusted organization, for example, data being hosted and protected by the CSIRO (a federal government research organization) or elsewhere by the government, would help allay those concerns.

Some GPs highlighted cost and reimbursement as major barriers to implementing the program in primary practice. This is explored further below.

Reimbursement and Funding Considerations

Aligning consultation and management of patients under the Medicare Benefits Scheme (Medicare), Australia's universal health care insurance program, was suggested by most GPs. This could be done by aligning the program delivery (including onboarding) and associated consultations under the CDM services provided by the GP. CDM services were introduced to provide further Medicare subsidized support to patients with chronic medical conditions [29]. By doing so, the nurses

XSL•FO
RenderX

involved in supporting the delivery of CAPS could be reimbursed via existing Medicare rebates within the management plan codes for their time and consultations, which was highlighted as vital with the "razor-thin" margins of a GP practice.

> *[…] you can probably fall within the same consult as the GP management plan, like a chronic disease management plan, because that also requires goal setting to do it correctly.* [Participant #14, GP]

Similarly, some GPs indicated they would privately bill the patients, referencing the "undercurrent" in the general practice of continued reductions in bulk billing (where Medicare benefits fully subsidize health services).

> *[…] if we created a job and we paid someone to be monitoring it possibly, but there would be, it would be a paid service and the patient would have to foot the bill for it.* [Participant #17, GP]

Another GP raised the possibility of using private health insurance to cover the sessions, given that the insurers are "incentivized" to reduce members using health services and having recurrent events. This could be used to then incorporate health coaches into the program to support and guide the patients and take the burden of monitoring from GPs. However, they acknowledged that this could exclude patients who do not have private health insurance.

> *[…] the private health insurers can pay the health coaches to actually help patients through this journey and […] the app can actually be part of the private health insurer hospital prevention program to say, a tool that health coaches can use […] set goals for [the patient] and the alerts will streamline where the alerts go, or will help [the patient] with when is it relevant to see a GP.* [Participant #14, GP]

An Independent Model of Care

An alternative proposal by several GPs was that the program could be run independently as a standalone private service, providing a solution to the current barriers GPs face. One GP mentioned that this could lead to more "feasible," "efficient," and cost-effective care by training independent clinical staff to solely onboard, monitor, and respond to data. These staff could then link in with GP services to share the information collected. Another GP highlighted how previous programs supported by a PHN have worked, for example, having independent clinical staff travel to different clinics, provide heart health checks, and then feed that information back to the GP.

Discussion

Principal Findings

We have described the perceptions of Australian clinicians and Stroke Foundation representatives regarding CAPS. This included their opinions regarding the potential perceived benefits, and the facilitators and challenges to the future adoption of this novel digital program as a health promotion program delivered in the community. We report that clinicians from different disciplines and the continuum of stroke care saw the value of a digitally delivered multicomponent secondary prevention program, which was discussed as a potential model of care. We present their insights into the challenges of integrating these types of programs into primary practice.

Both the clinicians interviewed in this study and the end-user participants living with stroke or TIA in the feasibility study felt the program could provide essential support if delivered shortly after discharge [17]. While secondary prevention begins in the hospital with medications, care plans, and risk factor education, not all patients receive these. In 2023, 30% of patients discharged from Australian hospitals did not receive a care plan, and 28% were not educated regarding risk factor management [30]. The transition from hospital to home is often poorly supported, leading to unmet needs, particularly in healthy lifestyles and uptake of secondary prevention [31]. The lack of support provided may contribute to many people living with stroke reporting long-term unmet needs, particularly in the areas of healthy lifestyles and secondary prevention [32,33]. CAPS could be used to provide essential information for self-management and ensure patients feel better supported after discharge [34].

Clinicians identified issues of accessibility, particularly "digital exclusion", as a potential barrier to patient uptake of the program. Digital access is now a key social determinant of health [35], which can be affected by factors such as digital literacy, age, education, and cultural diversity [36,37]. People with lower digital literacy, including those from culturally diverse backgrounds, are less likely to access and use digital technologies for their care [38]. To bridge the digital divide, interventions should be made accessible through various methods, such as using simple language, offering alternative formats (eg, audio or icons rather than text) [38,39], providing devices [13], co-designing interventions to ensure they are fit-for-purpose [13,38], and offering individual training in the use of technology [40]. In the CAPS feasibility study, training was provided via an image-based step-by-step manual, a tutorial video, and a telephone call. This may have contributed to the high level of participant satisfaction with the usability of the smartphone app and wearable device [17].

An additional barrier identified was the challenge of engaging patients who lack motivation to improve their health. While we observed sustained technology usage in the feasibility study [17], that finding was among people who chose to participate in digital health research. The broader stroke population may be less incentivized to engage with a digital health program and less motivated to improve their health. However, previous studies have found that particular features, such as the ability to review the data and active participation in health management, can improve motivation and engagement [41-43]. Person-centered goal-setting also increases motivation among people living with stroke [44]. While these features are part of CAPS, the program needs further evaluation in a larger, more representative population to assess motivation and adherence.

Several barriers to the adoption of a digital health program into clinical practice were identified. Stroke Foundation representatives noted limited availability for integration with existing services in their stroke-specific models. In particular,

XSL•FO
RenderX

barriers of adoption into primary practice were discussed. Scarcity of resources was a major concern of the clinicians interviewed, with limited time and personnel to regularly monitor and respond to patient data. Previous studies have also shown that mobile health technology (eg, smartphone apps), telemonitoring, and self-management tools can increase clinician workload [45], hindering the implementation of evidence-based stroke care guidelines [46]. Digital health programs should support, not disrupt, workflows, and clinicians must be involved in their development to ensure they are effective and fit-for-purpose for both patients and clinicians [47].

The first group of clinicians noted that GPs were the best placed health care providers to deliver a program such as CAPS. However, the GPs cited a number of barriers, as explored above, that raise challenges in delivering CAPS in the current primary practice setting. As such, a GP-led model of care may not be feasible at scale. Involving nurses (for example, the specialist nurses in the New South Wales Chronic Disease Management Program) [48] to support a shared model of care may be more plausible. Previous research has indicated that nurse-led remote patient monitoring is feasible and has positive clinical outcomes [49].

With training, a nurse could develop patient-centered goals, monitor and respond to data, and provide action points or feed important information back to primary care. This also leverages an existing workforce that is trained in health promotion, CDM, and patient monitoring. As suggested, delivering CAPS as an independent nurse-led service, where clinical summaries are securely fed back to primary practice, could alleviate some challenges of integration with the various practice management software. Given the fragmentation of management software used across Australia, integrating CAPS into each would be technically and financially challenging. Data sharing through secure formats (eg, HL7 or encrypted PDF reports) to primary practice could be a more practical approach.

Consideration of the sustainable implementation and delivery costs of CAPS is vital, given that they are key barriers to the uptake of digital health programs [50]. While aligning the program under the Australian Medicare–subsidized CDM plan as a potential reimbursement model was discussed, this funding is predominantly used for allied health services, such as podiatry [29]. Furthermore, there is no current funding mechanism for remote patient monitoring or the delivery of a digital health program through Medicare [51]. As such, alternative forms of funding are likely to be required. Funding through PHN initiatives or private health insurance may provide pathways for sustainable reimbursement. State-level funding may also be viable, with precedent from the South Australian government supporting the sustained delivery of a virtual clinical care model [52]. Sustainable delivery of CAPS will require investment in training, technical infrastructure (eg, servers, product updates, and cybersecurity), and secure integration with primary practices. Scalability of the program will also require addressing interoperability barriers across diverse practice management systems, ensuring workforce availability and readiness, and improving digital health literacy among both clinicians and patients [53]. Demonstration of downstream cost benefits from program delivery (eg, reduced hospitalization) will also be essential to attract continued funding.

Identifying successful digital programs, which have been uplifted into a health service, can highlight factors that facilitate implementation. However, there is limited reporting on this process, and many programs face large implementation challenges [54]. A successful Australian example is Cardihab (Cardihab Pty Ltd) [55], a cardiac rehabilitation program offered via a smartphone app and clinician-facing portal. The program was evaluated in an RCT for effectiveness before implementation [56], and it is now offered to patients in multiple Australian states through private health insurance and certain private and public health services [55]. Medly, delivered in Canada, is another example of a mobile phone–based telemonitoring program for heart failure. Developed through a user-centered design process, the program was evaluated for effectiveness in an RCT [57,58]. Medly is currently offered through 9 health services throughout Toronto, supported by the University Health Network and Boehringer Ingelheim Canada [59]. The successful implementation of these digital programs indicates the importance of conducting an RCT in order to progress through the implementation pipeline. Data from this study will inform the future design of an RCT, including modification of CAPS technology and the model of care to be evaluated.

The strengths of our study included the diverse range of clinicians and wider service providers with expertise across the continuum of care, including those involved in acute, rehabilitation, and primary care. Given their mention by GPs during the interviews, practice nurses and practice managers could have been interviewed to understand their perceptions of adopting CAPS into primary care. However, we feel that our pragmatic convenience sample of content experts provided sufficient information power related to the richness and relevance of the data [25]. We used standardized procedures to demonstrate the program components and capabilities, with clinicians asked a similar set of questions to enhance dependability and consistency while still permitting flexibility to explore their views and experiences. Analysis of the data was undertaken by researchers independently with collaborative cross-checking and oversight of a third researcher in the second group. We demonstrated the capabilities of the program for patient end users, including self-monitoring and tailored support for goal attainment. However, we did not have the opportunity to involve clinicians in using and delivering CAPS to demonstrate enhanced management of actual patients. This means their perceptions and interpretations were purely conceptual, based on the demonstration provided during the focus groups or interviews. As such, a number of clinicians made assumptions and provided comments about a program they were not fully familiar with. The current lack of interoperability with practice management software or electronic health records makes integration of CAPS challenging. The current program design, which relies on regular review and response to patient-entered data, may impact clinical workflows and workload. Furthermore, the use of CAPS requires some level of digital literacy and the ownership of compatible devices among patients, as well as digital competence among clinicians,

XSL•FO
RenderX

which may limit equitable use and adoption. Interviews for the first and second groups were conducted by different researchers; however, a consistent semistructured interview guide was used across all interviews, and the facilitators were experienced. Data analysis was also conducted by the same 2 researchers, and all the other researchers were involved in the discussion of the results.

The data collected regarding clinician perceptions will be integrated with the prior patient feedback collected [17]. This will be used to guide amendments to the program components (both patient and clinician facing; eg, patient traffic-light summaries) and updates of the technology. Their feedback will also guide the future model of care to be evaluated in a mixed-methods RCT, where patient participants will be recruited earlier in their recovery as recommended, and clinicians will be involved in the delivery of the program. We will also collect information to facilitate an economic evaluation of the program, a critical component of future implementation.

Conclusion

Clinician perspectives regarding the use of CAPS for improving the secondary prevention of stroke indicate the potential of the program to address gaps in care. Participant feedback indicated a need to consider a viable and sustainable model of care that includes primary care without detrimentally impacting associated workloads. Further evaluations of the program regarding effectiveness in phase II and III RCTs are warranted.

Acknowledgments

This research was funded by a STOPstroke Pilot Project grant and the Commonwealth Scientific and Industrial Research Organisation Precision Health Future Science Platform. LPA acknowledges scholarship support from Monash University and the Australian e-Health Research Centre, CSIRO. NAL acknowledges fellowship support from the Heart Foundation of Australia (#106762). DAC acknowledges fellowship support from the National Health Medical Research Council of Australia (#1154273).

The authors would like to acknowledge support in participant recruitment from A/Prof Seana Gall, Primary Health Tasmania, Gippsland Primary Health Network, MonREN (Monash University Practice Based Research Network), GoldNet (the Gold Standard in Primary Care Research), the Department of General Practice at Monash University, and the University of Queensland General Practitioner Network.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Outline of semistructured stimulus questions.
[PDF File, 56 KB - humanfactors_v12i1e72873_app1.pdf]

Checklist 1
COREQ checklist.
[PDF File, 433 KB - humanfactors_v12i1e72873_app2.pdf]

References

1. Feigin VL, Stark BA, Johnson CO, et al. Global, regional, and national burden of stroke and its risk factors, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Neurol 2021 Oct;20(10):795-820. [doi: 10.1016/S1474-4422(21)00252-0]
2. Kim J, Neville E, Dalli L, et al. Economic impact of stroke report 2024. : Stroke Foundation; 2024 URL: https://strokefoundation.org.au/media/g2rh1imp/2024-economic-impact-of-stroke-report-final.pdf [accessed 2025-09-15]
3. Lin B, Zhang Z, Mei Y, et al. Cumulative risk of stroke recurrence over the last 10 years: a systematic review and meta-analysis. Neurol Sci 2021 Jan;42(1):61-71. [doi: 10.1007/s10072-020-04797-5] [Medline: 33040195]
4. Mohan KM, Wolfe CDA, Rudd AG, Heuschmann PU, Kolominsky-Rabas PL, Grieve AP. Risk and cumulative risk of stroke recurrence: a systematic review and meta-analysis. Stroke 2011 May;42(5):1489-1494. [doi: 10.1161/STROKEAHA.110.602615] [Medline: 21454819]
5. Lisabeth LD, Ireland JK, Risser JMH, et al. Stroke risk after transient ischemic attack in a population-based setting. Stroke 2004 Aug;35(8):1842-1846. [doi: 10.1161/01.STR.0000134416.89389.9d] [Medline: 15192239]
6. Kleindorfer DO, Towfighi A, Chaturvedi S, et al. 2021 guideline for the prevention of stroke in patients with stroke and transient ischemic attack: a guideline from the American Heart Association/American Stroke Association. Stroke 2021 Jul;52(7):e364-e467. [doi: 10.1161/STR.0000000000000375] [Medline: 34024117]
7. Hackam DG, Spence JD. Combining multiple approaches for the secondary prevention of vascular events after stroke: a quantitative modeling study. Stroke 2007 Jun;38(6):1881-1885. [doi: 10.1161/STROKEAHA.106.475525] [Medline: 17431209]

XSL•FO
RenderX

8. Nicholson S, Sniehotta FF, van Wijck F, et al. A systematic review of perceived barriers and motivators to physical activity after stroke. Int J Stroke 2013 Jul;8(5):357-364. [doi: 10.1111/j.1747-4949.2012.00880.x] [Medline: 22974010]
9. Govori V, Budinčević H, Morović S, Đerke F, Demarin V. Updated perspectives on lifestyle interventions as secondary stroke prevention measures: a narrative review. Medicina (Kaunas) 2024 Mar 19;60(3):3. [doi: 10.3390/medicina60030504] [Medline: 38541229]
10. Zawawi NSM, Aziz NA, Fisher R, Ahmad K, Walker MF. The unmet needs of stroke survivors and stroke caregivers: a systematic narrative review. J Stroke Cerebrovasc Dis 2020 Aug;29(8):104875. [doi: 10.1016/j.jstrokecerebrovasdis.2020.104875] [Medline: 32689648]
11. Moschonis G, Siopis G, Jung J, et al. Effectiveness, reach, uptake, and feasibility of digital health interventions for adults with type 2 diabetes: a systematic review and meta-analysis of randomised controlled trials. Lancet Digit Health 2023 Mar;5(3):e125-e143. [doi: 10.1016/S2589-7500(22)00233-3] [Medline: 36828606]
12. Katz ME, Mszar R, Grimshaw AA, et al. Digital health interventions for hypertension management in US populations experiencing health disparities: a systematic review and meta-analysis. JAMA Netw Open 2024 Feb 5;7(2):e2356070. [doi: 10.1001/jamanetworkopen.2023.56070] [Medline: 38353950]
13. Allan LP, Beilei L, Cameron J, et al. A scoping review of mHealth interventions for secondary prevention of stroke: implications for policy and practice. Stroke 2023 Nov;54(11):2935-2945. [doi: 10.1161/STROKEAHA.123.043794] [Medline: 37800373]
14. Kilfoy A, Hsu TCC, Stockton-Powdrell C, Whelan P, Chu CH, Jibb L. An umbrella review on how digital health intervention co-design is conducted and described. NPJ Digit Med 2024 Dec 23;7(1):374. [doi: 10.1038/s41746-024-01385-1] [Medline: 39715947]
15. Noorbergen TJ, Adam MTP, Teubner T, Collins CE. Using co-design in mobile health system development: a qualitative study with experts in co-design and mobile health system development. JMIR Mhealth Uhealth 2021 Nov 10;9(11):e27896. [doi: 10.2196/27896] [Medline: 34757323]
16. Silvera-Tawil D, Cameron J, Li J, et al. Multicomponent support program for secondary prevention of stroke using digital health technology: co-design study with people living with stroke or transient ischemic attack. J Med Internet Res 2024 Aug 22;26:e54604. [doi: 10.2196/54604] [Medline: 39172512]
17. Allan LP, Silvera-Tawil D, Cameron J, et al. Novel multicomponent digital care assistant and support program for people after stroke or transient ischaemic attack: a pilot feasibility study. Sensors (Basel) 2024 Nov 13;24(22):7253. [doi: 10.3390/s24227253] [Medline: 39599030]
18. Wosny M, Strasser LM, Hastings J. Experience of health care professionals using digital tools in the hospital: qualitative systematic review. JMIR Hum Factors 2023 Oct 17;10:e50357. [doi: 10.2196/50357] [Medline: 37847535]
19. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
20. Bowman J, Mogensen L, Marsland E, Lannin N. The development, content validity and inter-rater reliability of the SMART-Goal Evaluation Method: a standardised method for evaluating clinical goals. Aust Occup Ther J 2015 Dec;62(6):420-427. [doi: 10.1111/1440-1630.12218] [Medline: 26286379]
21. Cadilhac DA, Cameron J, Kilkenny MF, et al. Protocol of a randomized controlled trial investigating the effectiveness of Recovery-focused Community support to Avoid readmissions and improve Participation after Stroke (ReCAPS). Int J Stroke 2022 Feb;17(2):236-241. [doi: 10.1177/17474930211022678] [Medline: 34037468]
22. Barnden R, Cadilhac DA, Lannin NA, et al. Development and field testing of a standardised goal setting package for person-centred discharge care planning in stroke. PEC Innov 2022 Dec;1:100008. [doi: 10.1016/j.pecinn.2021.100008] [Medline: 37364011]
23. Li J, Silvera-Tawil D, Varnfield M, Hussain MS, Math V. Users' perceptions toward mHealth technologies for health and well-being monitoring in pregnancy care: qualitative interview study. JMIR Form Res 2021 Dec 2;5(12):e28628. [doi: 10.2196/28628] [Medline: 34860665]
24. Li J, Varnfield M, Jayasena R, Celler B. Home telemonitoring for chronic disease management: perceptions of users and factors influencing adoption. Health Informatics J 2021;27(1):1460458221997893. [doi: 10.1177/1460458221997893] [Medline: 33685279]
25. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res 2016 Nov;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
26. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
27. Bradley EH, Curry LA, Devers KJ. Qualitative data analysis for health services research: developing taxonomy, themes, and theory. Health Serv Res 2007 Aug;42(4):1758-1772. [doi: 10.1111/j.1475-6773.2006.00684.x] [Medline: 17286625]
28. My stroke journey. : Stroke Foundation; 2024 URL: https://strokefoundation.org.au/what-we-do/for-survivors-and-carers/stroke-journey-resources [accessed 2025-09-15]

XSL•FO
RenderX

29. Use of Medicare chronic disease management items by patients with long-term health conditions. : Australian Institute of Health and Welfare; 2022 URL: https://www.aihw.gov.au/reports/chronic-disease/use-of-medicare-chronic-disease-management [accessed 2025-09-15]
30. National stroke audit: acute services report 2023. : Stroke Foundation; 2023 URL: https://informme.org.au/media/uppnqukx/2023-national-acute-services-national-report.pdf [accessed 2025-09-15]
31. Chen L, Xiao LD, Chamberlain D, Newman P. Enablers and barriers in hospital-to-home transitional care for stroke survivors and caregivers: a systematic review. J Clin Nurs 2021 Oct;30(19-20):2786-2807. [doi: 10.1111/jocn.15807] [Medline: 33872424]
32. Chen T, Zhang B, Deng Y, Fan JC, Zhang L, Song F. Long-term unmet needs after stroke: systematic review of evidence from survey studies. BMJ Open 2019 May 19;9(5):e028137. [doi: 10.1136/bmjopen-2018-028137] [Medline: 31110106]
33. Olaiya MT, Cadilhac DA, Kim J, et al. Long-term unmet needs and associated factors in stroke or TIA survivors: an observational study. Neurology (ECronicon) 2017 Jul 4;89(1):68-75. [doi: 10.1212/WNL.0000000000004063] [Medline: 28566545]
34. Pindus DM, Mullis R, Lim L, et al. Stroke survivors' and informal caregivers' experiences of primary care and community healthcare services - a systematic review and meta-ethnography. PLOS ONE 2018;13(2):e0192533. [doi: 10.1371/journal.pone.0192533] [Medline: 29466383]
35. Shaw RJ. Access to technology and digital literacy as determinants of health and health care. Creat Nurs 2023 Aug;29(3):258-263. [doi: 10.1177/10784535231211682] [Medline: 37909069]
36. Arias López MDP, Ong BA, Borrat Frigola X, et al. Digital literacy as a new determinant of health: a scoping review. PLOS Digit Health 2023 Oct;2(10):e0000279. [doi: 10.1371/journal.pdig.0000279] [Medline: 37556425]
37. Vercruyssen A, Schirmer W, Geerts N, Mortelmans D. How "basic" is basic digital literacy for older adults? Insights from digital skills instructors. Front Educ 2023;8. [doi: 10.3389/feduc.2023.1231701]
38. Whitehead L, Talevski J, Fatehi F, Beauchamp A. Barriers to and facilitators of digital health among culturally and linguistically diverse populations: qualitative systematic review. J Med Internet Res 2023 Feb 28;25:e42719. [doi: 10.2196/42719] [Medline: 36853742]
39. O'Connor S, Hanlon P, O'Donnell CA, Garcia S, Glanville J, Mair FS. Understanding factors affecting patient and public engagement and recruitment to digital health interventions: a systematic review of qualitative studies. BMC Med Inform Decis Mak 2016 Sep 15;16(1):120. [doi: 10.1186/s12911-016-0359-3] [Medline: 27630020]
40. Daniels K, Bonnechère B. Harnessing digital health interventions to bridge the gap in prevention for older adults. Front Public Health 2023;11:1281923. [doi: 10.3389/fpubh.2023.1281923] [Medline: 38259780]
41. Morrissey EC, Casey M, Glynn LG, Walsh JC, Molloy GJ. Smartphone apps for improving medication adherence in hypertension: patients' perspectives. Patient Prefer Adherence 2018;12:813-822. [doi: 10.2147/PPA.S145647] [Medline: 29785096]
42. Rief JJ, Hamm ME, Zickmund SL, et al. Using health information technology to foster engagement: patients' experiences with an active patient health record. Health Commun 2017 Mar;32(3):310-319. [doi: 10.1080/10410236.2016.1138378] [Medline: 27223684]
43. Austin L, Sharp CA, van der Veer SN, et al. Providing "the bigger picture": benefits and feasibility of integrating remote monitoring from smartphones into the electronic health record. Rheumatology (Oxford) 2020 Feb 1;59(2):367-378. [doi: 10.1093/rheumatology/kez207] [Medline: 31335942]
44. Sugavanam T, Mead G, Bulley C, Donaghy M, van Wijck F. The effects and experiences of goal setting in stroke rehabilitation - a systematic review. Disabil Rehabil 2013 Feb;35(3):177-190. [doi: 10.3109/09638288.2012.690501] [Medline: 22671934]
45. Keuper J, van Tuyl LHD, de Geit E, et al. The impact of eHealth use on general practice workload in the pre-COVID-19 era: a systematic review. BMC Health Serv Res 2024 Sep 19;24(1):1099. [doi: 10.1186/s12913-024-11524-9] [Medline: 39300456]
46. McCluskey A, Vratsistas-Curto A, Schurr K. Barriers and enablers to implementing multiple stroke guideline recommendations: a qualitative study. BMC Health Serv Res 2013 Aug 19;13(1):323. [doi: 10.1186/1472-6963-13-323] [Medline: 23958136]
47. Borges do Nascimento IJ, Abdulazeem H, Vasanthan LT, et al. Barriers and facilitators to utilizing digital health technologies by healthcare professionals. NPJ Digit Med 2023 Sep 18;6(1):161. [doi: 10.1038/s41746-023-00899-4] [Medline: 37723240]
48. Chronic disease management. Northern NSW Local Health District. URL: https://nnswlhd.health.nsw.gov.au/community-health/ballina-community-health/chronic-disease-management [accessed 2025-09-15]
49. Coffey JD, Christopherson LA, Williams RD, et al. Development and implementation of a nurse-based remote patient monitoring program for ambulatory disease management. Front Digit Health 2022;4:1052408. [doi: 10.3389/fdgth.2022.1052408] [Medline: 36588748]
50. Weik L, Fehring L, Mortsiefer A, Meister S. Understanding inherent influencing factors to digital health adoption in general practices through a mixed-methods analysis. NPJ Digit Med 2024 Feb 27;7(1):47. [doi: 10.1038/s41746-024-01049-0] [Medline: 38413767]
51. Leveraging digital technology in healthcare - research paper. : Australian Government - Productivity Commission; 2024 URL: https://www.pc.gov.au/research/completed/digital-healthcare/digital-healthcare.pdf [accessed 2025-09-15]

XSL•FO
RenderX

52. Home tele-monitoring service - virtual clinical care (VCC). Government of South Australia - South Australia Health. URL: https://www.sahealth.sa.gov.au/wps/wcm/connect/public+content/sa+health+internet/services/regional+health+services/virtual+clinical+care/home+tele+monitoring+service+-+virtual+clinical+care+vcc [accessed 2025-09-15]
53. Schlieter H, Marsch LA, Whitehouse D, et al. Scale-up of digital innovations in health care: expert commentary on enablers and barriers. J Med Internet Res 2022 Mar 11;24(3):e24582. [doi: 10.2196/24582] [Medline: 35275065]
54. Dendere R, Janda M, Sullivan C. Are we doing it right? We need to evaluate the current approaches for implementation of digital health systems. Aust Health Rev 2021 Dec;45(6):778-781. [doi: 10.1071/AH20289] [Medline: 34488938]
55. Cardihab. URL: https://cardihab.com/ [accessed 2024-03-24]
56. Varnfield M, Karunanithi M, Lee CK, et al. Smartphone-based home care model improved use of cardiac rehabilitation in postmyocardial infarction patients: results from a randomised controlled trial. Heart 2014 Nov;100(22):1770-1779. [doi: 10.1136/heartjnl-2014-305783] [Medline: 24973083]
57. Medly. URL: https://medly.ca/ [accessed 2024-11-24]
58. Seto E, Leonard KJ, Cafazzo JA, Barnsley J, Masino C, Ross HJ. Mobile phone-based telemonitoring for heart failure management: a randomized controlled trial. J Med Internet Res 2012 Feb 16;14(1):e31. [doi: 10.2196/jmir.1909] [Medline: 22356799]
59. Successful virtual care model grows thanks to collaboration. Boehringer Ingelheim. URL: https://www.boehringer-ingelheim.com/ca/partnering/human-health-partnering/partnering-interests/ontario/medly-program [accessed 2024-12-24]

Abbreviations

CAPS: Care Assistant and support Program for people after Stroke or transient ischemic attack
CDM: chronic disease management
CSIRO: Commonwealth Scientific and Industrial Research Organization
GP: general practitioner
PHN: primary health care network
RCT: randomized controlled trial
TIA: transient ischemic attack

Edited by A Malik; submitted 19.02.25; peer-reviewed by E Linedale, KP Sze; revised version received 22.06.25; accepted 26.08.25; published 09.10.25.

Please cite as:
Allan LP, Li J, Purvis T, Silvera-Tawil D, Cameron J, Varnfield M, Smallbon V, Bomke J, Lannin NA, Cadilhac DA
Clinician Perceptions of a Novel Multicomponent Digital Care Assistant and Support Program for People After Stroke or Transient Ischemic Attack (CAPS) for the Secondary Prevention of Stroke: Qualitative Study
JMIR Hum Factors 2025;12:e72873
URL: https://humanfactors.jmir.org/2025/1/e72873
doi:10.2196/72873

© Liam Pearce Allan, Jane Li, Tara Purvis, David Silvera-Tawil, Jan Cameron, Marlien Varnfield, Vanessa Smallbon, Julia Bomke, Natasha A Lannin, Dominique A Cadilhac. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 9.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Designing for Patient-Centered Care and Equity in Virtual Hospital-at-Home Models: Quality Improvement Initiative Using Experience-Based Co-Design

Mahabhir Kandola[1]; Emma Wong[1]; Robert Paquin[2]; Kamal Arora[1]; Harroop Sharda[3]; Roman Deol[1]; Mary Jung[4]; Maria Montenegro[1]; Megan MacPherson[1,5]

[1]Department of Virtual Health and Department of EDI and Language Services, Fraser Health, 13450 102 Ave, Surrey, Canada

[2]King's College London, 5-11 Lavington Street, London, United Kingdom

[3]British Columbia Institute of Technology, Burnaby, BC, Canada

[4]University of British Columbia, Kelowna, BC, Canada

[5]University of British Columbia, Vancouver, BC, Canada

Corresponding Author:
Megan MacPherson
Department of Virtual Health and Department of EDI and Language Services, Fraser Health, 13450 102 Ave, Surrey, Canada

Abstract

Background: The rapid expansion of virtual care during COVID-19 accelerated the development of virtual hospital-at-home models, which deliver hospital-level care in patients' homes through remote monitoring, virtual communication, and in-person support when required. While the virtual hospital-at-home model offers the potential to improve patient-centered care and health equity, rapid implementation often overlooks culturally diverse and underserved populations, including South Asian communities who experience disproportionate chronic disease burden and barriers to accessing culturally relevant services. Strategies are needed to ensure equitable design and adoption of virtual hospital-at-home models.

Objective: This study used an experience-based co-design (EBCD) approach to engage patients, caregivers, clinicians, and community organizations in shaping a regional virtual hospital-at-home strategy within the Fraser Health Authority, British Columbia, Canada. The aim was to identify barriers, facilitators, and equity-focused solutions to inform future implementation.

Methods: We conducted a five-stage EBCD quality improvement process in the Fraser Health Authority, British Columbia, including (1) forming a multidisciplinary steering committee, (2) reviewing health care provider experiences, (3) interviewing South Asian patients and caregivers, (4) hosting a co-design workshop to develop solutions, and (5) sharing back findings.

Results: Participants identified barriers, including digital literacy, language, and trust in virtual care. The co-designed solutions focused on culturally tailored education, hybrid digital training, caregiver inclusion, and community-driven engagement strategies.

Conclusions: EBCD enabled the development of inclusive and actionable strategies to improve virtual hospital-at-home services. The findings highlight the importance of ongoing community collaboration to ensure equity in virtual care innovation.

(JMIR Hum Factors 2025;12:e79679) doi:10.2196/79679

KEYWORDS

community-based participatory research; telemedicine; health equity; South Asian people; artificial intelligence; AI

Introduction

Background

The rapid expansion of virtual care delivery in response to COVID-19 has paved the way for the development of virtual hospital-at-home models, which are models of care that mirror the structure and functions of traditional inpatient wards, but enable eligible patients to receive hospital-level care in their own homes through a combination of remote monitoring, virtual communication, and in-person support when needed [1]. Unlike general telemedicine or outpatient virtual visits, which typically provide episodic or consultative care, the virtual hospital-at-home model aims to deliver a continuous and integrated alternative to acute hospital care.

Virtual hospital-at-home services hold promise in improving patient-centered care and enhancing health equity within communities [2]; however, the rapid implementation of virtual care during COVID-19 lacks comprehensive patient-centered development and evaluation, potentially marginalizing already underserved populations (eg, older adults, people with disabilities, and culturally diverse communities) [3]. Without effective strategies for continuous improvement, virtual care risks perpetuating inequitable and low-quality care, undermining the very goals they aim to achieve [4]. In Canada, South Asian

XSL•FO
RenderX

communities represent a significant proportion of residents who have historically been underserved and experience substantial barriers related to language, cultural relevance, and access to health services [5]. South Asians in Canada have higher rates of chronic diseases compared to other racialized groups, including cardiovascular disease, diabetes, and hypertension [6,7]—partly due to social determinants of health (eg, access to education, employment opportunities, food security, experiences of discrimination, and access to culturally appropriate care) [8-10]. Given the inequities and disparities faced by South Asian community members within Canada, innovative care strategies are needed to enhance access to culturally relevant care and improve health outcomes.

Challenges

Key virtual care challenges include the following: (1) the digital divide: disparities in technology access and digital literacy disproportionately exclude vulnerable populations, such as older adults and culturally diverse communities, from the benefits of virtual care [11,12]; (2) workforce strain: clinicians face challenges navigating new technologies and workflows and often lack the time to learn new systems and technologies contributing to burnout and hindering virtual care adoption [13]; (3) rapid technological evolution: virtual health requires continuous innovation and adaptability to ensure sustainability, interoperability, and alignment with patient needs [14,15]. In this paper, we specifically address challenges related to equitable access and patient-centered engagement in virtual hospital-at-home services.

Fraser Health Setting

These challenges are particularly salient within the Fraser Health Authority, the largest regional health authority in British Columbia, Canada, which provides care to more than 2 million residents across 20 diverse communities [16]. Fraser Health serves 75% of the South Asian residents in the province, and 1 in 5 Fraser Health residents identifies as South Asian [17]. Considering Fraser Health's diverse population and geographical spread, the virtual hospital-at-home model has the potential to improve care quality and access while alleviating pressure on hospital resources. Fraser Health is designing a regional program and strategic implementation guidelines for a virtual hospital-at-home service with five primary objectives: (1) increase overall system capacity, (2) improve patient flow and efficiency, (3) reduce hospital-acquired harm, (4) enhance patient and essential care partner experience, and (5) enhance clinician experience.

At the time of this study, the only active unit was the virtual psychiatric unit (VPU), launched in fall 2022 [18]. The VPU virtually provides hospital-level psychiatric care, but the long-term vision for Fraser Health's virtual hospital-at-home strategy is to expand to other clinical areas (not yet specified). This project did not evaluate the VPU directly, but instead sought to engage community members in shaping the design of future virtual hospital-at-home models across a range of clinical areas.

This Project

This study uses a design thinking approach, experience-based co-design (EBCD) [19], to help inform the development of the Fraser Health virtual hospital-at-home service. EBCD is a participatory, user-centered approach to quality improvement that engages those closest to the area for improvement. In this project, patients, caregivers, clinicians, and decision-makers were included to explore areas for improvement and to co-design service improvements based on lived experiences. EBCD is based on the premise that co-design can lead to experience-focused health care services, which can result in improved outcomes in clinical effectiveness and safety [20]. This multistage service improvement approach captures the experiences and emotions of service users by understanding key events or "touchpoints" during their service journey [20].

Importantly, EBCD has been recognized as a vehicle for patient empowerment and equity, shifting traditional power structures by positioning patients as co-producers in health care improvement [21]. In public health, more broadly, co-creation models such as EBCD have been intentionally used to address health inequities by involving marginalized communities across all phases of intervention development, from agenda-setting to evaluation [22]. In addition, EBCD has been shown to help flatten epistemic inequities by equally valuing all users' experiences and dismantling entrenched power imbalances in co-design processes [23]. By centering on the lived experiences of patients and caregivers, particularly from underserved communities, EBCD ensures that service improvements prioritize equity alongside quality and safety.

Methods

Overview

EBCD is a quality improvement methodology that is typically structured into 6 stages [19]. This study adapted the traditional 6-stage EBCD framework into 5 stages, as the final celebratory event was integrated into steering committee debrief sessions rather than conducted as a stand-alone event. A core component of EBCD is the creation of a short film based on interviews with participants, which is used to spark discussions between those providing and receiving care at a co-design event. This study consisted of 5 stages, and this paper describes the process of using EBCD to help inform the Fraser Health virtual hospital-at-home service (Table 1).

XSL•FO
RenderX

Table . Stages of the experience-based co-design approach.

Stage	Key activities	Participants	Key outputs
Stage 1: setting up and steering committee	Established steering committee and monthly meetings	PICS[a] society, SAHI[b], Virtual Health, EDI[c] team, and researchers	Guidance on study approach and culturally relevant methods
Stage 2: observing service delivery and staff experiences	Interviews with health care providers	7 staff (n=5, 71% physicians and n=2, 29% nurses)	Insights on barriers and opportunities for virtual hospital-at-home models
Stage 3: gathering patient experiences and touchpoint film	Interviews with South Asian community members and film creation	20 patients and caregivers (9 in film)	Short touchpoint film highlighting key experiences
Stage 4: co-design event	Workshop using film to guide problem-solving	Patients, caregivers, clinicians, and steering committee	Solutions for 5 key themes: patient-centered care, technology, training, community engagement, and process improvement
Stage 5: post–co-design analysis and uptake	Thematic analysis and follow-up meetings	Steering committee and the virtual hospital development team	Recommendations shared, implementation planning, and reflections

[a]PICS: Progressive Intercultural Community Services.

[b]SAHI: South Asian Health Institute.

[c]EDI: Equity, Diversity, and Inclusion.

Stage 1: Setting Up and Establishing a Steering Committee

Members from the Progressive Intercultural Community Services (PICS) society, Fraser Health's Virtual Health department, Fraser Health's South Asian Health Institute, Fraser Health's Equity, Diversity, and Inclusion team, and researchers were invited to join the steering committee. These groups were chosen for their understanding of the underserved communities (particularly South Asian communities), their roles in the health care system, and their experience with virtual care services and community-based research methods, ensuring a comprehensive perspective for the project. The committee met monthly to discuss project updates, provide feedback on the research approach, and assist in refining the study's methods.

Stage 2: Observing Service Delivery and Gathering Staff Experiences

To contextualize the development of virtual hospital-at-home services within Fraser Health, we incorporated findings from a previous qualitative study with 7 health care providers (n=5, 71% physicians and n=2, 29% registered nurses) regarding their experiences with existing remote patient monitoring services. We conducted approximately 30-minute, semistructured one-on-one interviews with health care providers across Fraser Health. The semistructured interview guide was developed in consultation with the Virtual Health department to ensure that the questions were relevant, actionable, and aligned with the information needs of the remote patient monitoring program. Interviews explored health care providers' experiences referring patients to remote patient monitoring programs, barriers encountered, perceived advantages and disadvantages of referral, and suggestions for program improvements. These interviews were included to inform this study, as remote patient monitoring is an integral aspect of virtual hospital-at-home models. Remote patient monitoring allows clinicians to track patient status in real time and supports proactive care delivery. For full details on the methods used, refer to the study by N Moghaddam, M Montenegro, and M MacPherson (unpublished data, 2025).

Stage 3: Gathering Patient Experiences and Creating Touchpoint Film

Overview

Qualitative interviews were conducted with 20 South Asian community members in multiple languages to better understand the attitudes, experiences, and perceptions of patients and caregivers toward virtual hospital-at-home models. South Asian communities are among the largest racialized groups within the Fraser Health region. These interviews identified perceived barriers and preferences regarding virtual hospital-at-home models. Moreover, 30- to 60-minute, one-on-one semistructured interviews were conducted virtually via Microsoft Teams. Technical support was available through the PICS society to ensure participant access. The topic guide was adapted from previous Fraser Health Virtual Health workshops, revised with feedback from South Asian community members, and reviewed by the plain language team for clarity. The guide covered perceptions of and attitudes toward the virtual hospital-at-home model, including perceived benefits, concerns, and strategies for community engagement. During the consent process, participants were asked if they agreed to have their interviews audio and video recorded and if they would like to be contacted for potential participation in the co-design event. For full details on the methods used, refer to the study by MacPherson [24].

Creating the Short Film

Following the interviews, a short film was created featuring interview clips to highlight key "touchpoints" or themes and outline experiences within the patient journey. Only participants who consented to their video being used were included in the study. Following thematic analysis, salient quotes were identified, and the corresponding video clips were spliced into a 10-minute film created by Fraser Health's Virtual Health department. The film is an essential element of the EBCD approach, which captures patient experiences to help facilitate

XSL•FO
RenderX

conversation in stage 4, and is an effective tool to communicate patient experiences to health care providers.

Stage 4: Co-Design Event

Interview participants who consented to be contacted, steering committee members, and clinicians were invited to an in-person co-design workshop where the touchpoint film from stage 3 was presented to stimulate discussion on improving virtual care. The film highlighted key experiences and themes identified in the interviews, providing a foundation for the co-design discussions. Community members were introduced to the principles of co-design and engaged in collaborative problem-solving activities to address 6 key challenges: language barriers, digital literacy, comfort and convenience, suitability, quality of care, and awareness of available services.

These 6 challenges were derived from thematic analysis of participant interviews conducted in stages 2 and 3, representing the most salient and recurrent themes reported by participants. The steering committee reviewed and confirmed these challenges before the workshop to ensure that they accurately reflected participant priorities. Discussion prompts and structured design thinking exercises were specifically aligned with these areas, ensuring that participant contributions addressed the most relevant issues for improving the virtual hospital-at-home service.

Participants contributed ideas through structured design thinking exercises, including the “5 Whys” technique and “Crazy 8’s” brainstorming method. The “5 Whys” is a root cause analysis method that involves asking “why” iteratively (typically 5 times) to drill down to the underlying cause of a problem [25]. “Crazy 8’s” is a rapid ideation exercise in which participants sketch 8 distinct ideas in 8 minutes, promoting creative thinking and diverse solution generation [26,27]. Both methods facilitated active engagement and helped structure participant feedback. The proposed solutions were analyzed and subsequently shared with the virtual hospital-at-home development team for further consideration.

Stage 5: Co-Design Analysis, Follow-Up, and Internal Uptake

Although a formal celebration event (stage 6 in the traditional EBCD model) was not conducted, several postworkshop activities were carried out to ensure continuity and uptake of insights. Ideas generated during the co-design event were thematically analyzed by the research team to identify crosscutting recommendations. These recommendations were then shared with the virtual hospital-at-home development team, who reviewed and discussed the feasibility of incorporating them into ongoing service planning.

In addition, follow-up debrief sessions were held with the steering committee to reflect on the process, discuss next steps, and disseminate knowledge products (eg, the touchpoint film and a summary of co-design insights). Informal celebrations occurred during these meetings, where the steering committee members acknowledged the value of the process and its outcomes.

Ethical Considerations

This project received ethical approval from the Fraser Health Research Ethics Board (FHREB; approval IDs H23-03144 and H23-00463). All participants provided written informed consent prior to participation, including separate consent for video recording, for inclusion of video footage in a short film shared at the co-design event, and for potential follow-up contact regarding subsequent co-design activities. Participants were informed of the purpose and procedures for each stage of the project, as well as their right to withdraw at any time without consequence.

All data were treated confidentially and stored on secure, password-protected servers accessible only to the research team. Identifying information was removed or anonymized in transcripts and reports to protect participant privacy. Video recordings were used solely for analysis and, where consent was provided, for inclusion in the co-design film. The film was shown only to event participants and was not distributed beyond that setting.

No financial or other forms of compensation were provided for participation in the co-design event.

Results

Stage 1: Establishing a Steering Committee

A total of 13 individuals participated in the steering committee, representing a range of end-user groups, including representatives from the PICS society; Fraser Health’s South Asian Health Institute; Fraser Health Virtual Health; Fraser Health’s Equity, Diversity, and Inclusion team; and academic researchers from across Canada. Over the course of 12 months, the committee held 8 meetings, with an average attendance rate of 8 members (61% of the committee).

Early discussions focused on refining the research approach, with members advocating for culturally relevant interview questions within stage 3 and the inclusion of language support services to improve accessibility. As data collection progressed, committee members played a critical role in interpreting the findings, offering insights that contextualized patient experiences within the broader health care system.

Stage 2: Observing Service Delivery and Gathering Staff Experiences

As reported in the study by N Moghaddam, M Montenegro, and M MacPherson (unpublished data, 2025), 7 staff interviews (n=5, 71% physicians and n=2, 29% registered nurses and n=5, 71% women and n=2, 29% men) highlighted both opportunities (extended assessment time and stronger therapeutic relationships) and barriers (limited involvement in early development stages) in remote monitoring. These findings reinforced the importance of engaging clinicians in program design and informed subsequent co-design discussions in this project.

XSL•FO
RenderX

Stage 3: Gathering Patient Experiences and Creating a Short Film

As reported in the study by MacPherson [24], 20 interviews with South Asian community members identified as patients or caregivers explored attitudes, barriers and facilitators, and suggestions for enhancing awareness and improving virtual care services. While participants were generally open to virtual hospital-at-home models, they identified barriers related to trust, limited awareness, and challenges in access.

From these interviews, a 10-minute touchpoint film was created to highlight the key experiences and themes identified by participants. Of the 20 participants, 9 (45%) consented to inclusion (n=6, 67% women and n=3, 33% men; aged between 19 and 69 y), with 3 (15%) opting for audio-only contributions. Participants included patients (n=3, 33%), caregivers (n=2, 22%), and those who identified as both patients and caregivers (n=4, 44%). Video and audio clips were selected to represent the most salient themes identified in the interviews, such as building trust in virtual care, addressing language barriers, and ensuring equitable access. The film, produced by Fraser Health's Virtual Health department, served as a central EBCD tool to ground the subsequent co-design event in authentic patient and caregiver voices.

Stage 4: Co-Design Event

Overview

A total of 46 individuals were invited to participate in the co-design workshop, including interview participants from stage 3 who consented to be contacted, steering committee members, and clinicians. Participants with divergent views, including those who were initially skeptical of virtual care, were purposefully included in the co-design session to ensure a broad range of perspectives and stimulate rich discussion. Ultimately, 11 participants from stage 3 attended, alongside 6 steering committee members, a physician, 2 clinical nurse specialists, a learning consultant for Virtual Health, the manager for clinical operations for the VPU, and a coordinator for the South Asian Health Institute.

This workshop aimed to generate solutions to the problem areas identified during the interviews in stages 2 and 3, addressing the five key themes described here: (1) patient-centered and equitable care; (2) technology, innovation, and digital literacy; (3) training and education; (4) community engagement and outreach; and (5) process improvement and data use.

Patient-Centered and Equitable Care

While many participants appreciated the structure and direct interactions of hospital-based care, they emphasized the importance of ensuring that virtual hospital-at-home models provide the same level of time, reassurance, and guidance. However, ensuring equitable access to care requires addressing cultural and linguistic barriers, which go beyond interpretation and using plain language to include cultural nuance and understanding. For example, some South Asian families may not trust virtual care as "real" care, especially in serious conditions. They may believe that the patient is being dismissed too early or not taken seriously. Therefore, health care providers may need to explain the concept in culturally resonant terms—perhaps likening it to home visits from trusted family physicians in culturally familiar contexts. Community consultations and continuous feedback mechanisms were highlighted as essential to ensuring that health care services align with patient needs.

Participants suggested that health care models should incorporate family caregivers, including younger family members who may play a significant role in patient support. Furthermore, addressing technical concerns and ensuring that digital health care solutions are user-friendly with 2-way communication channels can alleviate patient anxiety, fostering a seamless integration of innovative care models.

Technology, Innovation, and Digital Literacy

Digital literacy emerged as a critical factor influencing virtual health care access and engagement. Participants acknowledged that while artificial intelligence and other digital technologies could enhance health care services, discomfort or unfamiliarity with new technology remained a significant barrier. Many patients expressed the need for hybrid digital literacy training that incorporates both online and in-person instruction, ensuring that those unfamiliar with technology can develop confidence in using digital health tools.

The use of body mapping tools with glossaries of translations, accessible via tablets or other digital devices, was suggested as a way to improve communication for patients experiencing language barriers. Participants also emphasized the need for a 24/7 patient support line to assist with technological issues, ensuring continuous access to health care services, particularly in case of an emergency while at home. Transparency in wait times and care escalation pathways, particularly through online queue systems or digital "call bells" (ie, digital notification systems for patients to request timely assistance remotely, similar to a call bell used in hospital), was seen as a potential strategy for improving patient trust and engagement in digital health care solutions.

Training and Education

Participants highlighted the need for enhanced training and education for both health care providers and patients. Digital resources, such as instructional videos, were identified as effective tools for increasing accessibility to health care information. Simulation-based learning was widely recommended, with participants advocating for hands-on training before patient discharge to ensure they understand their care plans and expectations. Training for health care providers should also emphasize virtual health competencies, equipping them with the skills necessary to navigate virtual care and digital platforms effectively, ultimately improving patient care.

Cultural competency training was another key recommendation, ensuring that health care providers can communicate effectively with diverse patient populations. To address health literacy challenges, participants suggested the use of multilingual training manuals with visual aids and simplified language to enhance patient comprehension. Participants also recommended updating patient training modules continuously based on current questions and concerns. Addition, structured patient education

XSL•FO
RenderX

that clarifies health care processes was seen as essential for helping patients navigate the complexities of the health care system, with a need for clear, step-by-step education, especially regarding the "entry points" of the health care system.

Community Engagement and Outreach

Participants underscored the importance of community-driven approaches to health care engagement and education. Meeting patients where they are, through initiatives such as community pop-ups, outreach in long-term care facilities, and targeted engagement in hospice settings, was identified as a key strategy for increasing health care accessibility. To build motivation to try virtual hospital-at-home models, participants recommended engaging in transparent discussions with patients as to why the virtual hospital-at-home model would be most suitable to their individual circumstances.

Peer learning was also seen as an effective method for knowledge transfer, particularly among older adults who may feel more comfortable learning from their peers. Social media and online communities were highlighted as valuable platforms for sharing health care experiences, testimonials, and educational content. Addition, participants suggested leveraging community champions and patient testimonials to promote new health care models and build trust in virtual health solutions. Media and promotional campaigns, particularly in South Asian media networks, were also recommended to increase public awareness and encourage informed decision-making about health care options.

Process Improvement and Data Use

Participants emphasized the role of research, data collection, and continuous feedback in improving health care processes. Both qualitative and quantitative data collection were seen as essential for refining care models and ensuring that they meet patient needs. Global information sharing was recommended as a way to integrate best practices from different health care systems. Participants also stressed the importance of transparency in care escalation processes, ensuring that patients understand when and how to seek additional support. Determining the appropriate fit between hospital-based and virtual care was seen as crucial for optimizing resource allocation and patient outcomes. Additionally, surveys and patient reviews were suggested as mechanisms for evaluating health care experiences and identifying areas for improvement. By bridging the gap between traditional and digital health care models, participants believed that the health care system could become more efficient, accessible, and responsive to patient concerns.

Stage 5: Post–Co-Design Analysis, Follow-Up, and Internal Uptake

Thematic analysis of workshop data resulted in a set of actionable recommendations, which were shared with the virtual hospital-at-home development team. Internal team discussions helped prioritize feasible ideas for implementation. While a formal evaluation was not conducted, follow-up meetings with the steering committee provided an opportunity to reflect on the success of the project and explore future applications of co-design. Members expressed appreciation for the participatory nature of the process, and the knowledge products were circulated for a broader impact. One participant shared that the co-design event was the first time they felt truly heard by the health care system, and it motivated them to explore a career in health care—a powerful testament to the impact of the process.

Discussion

Principal Findings

This study demonstrates the value of using an EBCD approach to inform the development of a patient-centered and equitable virtual hospital-at-home model within a large, culturally diverse regional health authority. By meaningfully engaging patients, caregivers, health care providers, and community organizations across 5 structured stages, the project not only surfaced critical insights into patient experiences but also helped shift traditional power dynamics by positioning patients and caregivers as partners in the design process. This participatory approach generated equity-focused recommendations that highlight how virtual hospital-at-home services can be more responsive to the needs of underserved communities.

Consistent with previous evidence showing that virtual hospital-at-home models improve outcomes and reduce hospital use [2], our findings confirm the promise of this care model. However, we also extend previous research by showing that successful and equitable implementation requires more than technology alone. Participants emphasized that access is contingent on culturally safe, linguistically appropriate, and caregiver-inclusive approaches. For example, as described in the stage 3 interviews, some South Asian families perceived virtual care as less legitimate than hospital-based services, underscoring the importance of culturally resonant communication strategies.

Digital literacy, caregiver inclusion, and culturally competent communication emerged as pivotal enablers of trust and adoption. Participants in the stage 4 co-design event proposed innovative solutions, such as hybrid training that combines online and in-person instruction, the use of body mapping with multilingual glossaries, and caregiver-centered education modules.

Similarly, digital literacy and caregiver involvement emerged as critical enablers of trust and engagement, reflecting themes raised in the stage 4 co-design workshop. Participants' proposals, such as hybrid digital literacy training, multilingual educational tools, and caregiver-centered education, illustrate how community-driven insights can translate into practical solutions. These grassroots innovations, surfaced through authentic engagement, demonstrate how EBCD operates not only as a design thinking framework but also as a platform for empowerment, amplifying voices that are often marginalized in top-down service planning.

At the same time, perspectives from clinicians and health system staff revealed important implementation challenges. As seen in stage 2 interviews, while health care providers valued the potential of virtual care (eg, extended assessment time and earlier intervention), many reported feeling excluded from initial

XSL•FO
RenderX

design processes, which limited their confidence in making referrals. This disconnect reflects a broader challenge in virtual health: aligning patient-centered priorities with operational realities. Previous research highlights that clinician engagement early in design processes is critical to ensuring that virtual care integrates seamlessly into workflows [28]. Our findings reinforce this point, suggesting that virtual hospital-at-home models cannot succeed without parallel attention to patient and health care provider perspectives.

Implications for Practice and Policy

This study highlights 5 practice and policy implications to support equitable virtual hospital-at-home implementation, grounded in our findings and supported by the existing literature.

Culturally Responsive Design

Virtual models must be co-developed with, not simply for, cultural communities. Tailored communication, education, and engagement strategies are necessary to build trust and legitimacy. This aligns with previous work emphasizing that digital health interventions must be culturally adapted to reduce health inequalities [29] and that co-design uncovers culturally resonant strategies that improve uptake [22].

Hybrid Digital Literacy Support

Bridging digital divides requires blended approaches that combine in-person training, peer-led instruction, and multilingual resources. As recommended by stage 4 participants, hybrid digital literacy models ensure that those unfamiliar with technology are supported while maintaining flexibility for more experienced users. Evidence suggests that such models enhance confidence and adoption among populations with limited previous exposure to digital tools [30,31].

Ongoing Feedback Loops

Embedding continuous feedback mechanisms (eg, community consultations, patient surveys, and user analytics) ensures that services remain responsive to evolving needs. This is consistent with evidence that iterative engagement strengthens program equity and sustainability [32].

System Integration and Escalation Clarity

Clear care pathways and escalation protocols are essential for patient safety and trust. In our study, participants requested transparent communication about when and how to escalate concerns, echoing previous research demonstrating that virtual hospital models can be sustainably embedded into broader care systems through shared “digital front doors,” referral pathways, and record sharing that support continuity of care [33].

Caregiver Inclusion

Caregivers were consistently identified as central to the success of virtual hospital-at-home models. Equipping caregivers with tools, training, and access to participate directly in care planning not only supports patients but also reduces barriers related to language and technology. Previous studies confirm that caregiver involvement, including proxy access to digital platforms, enhances access and equity for underserved groups [34-36].

Broader Policy Recommendations

Beyond these practice implications, broader policy action is needed. Standardizing culturally competent practices, investing in multilingual digital tools, and establishing mechanisms for community governance in digital health initiatives are critical next steps. As frameworks for patient-centered telemedicine suggest, community involvement and cultural adaptation must be embedded as core policy principles rather than optional add-ons [37].

Limitations

This study has several limitations. First, the interviews focused exclusively on South Asian communities. Although intentional, given the demographic composition of the Fraser Health region, future work should include additional cultural and linguistic groups to ensure broader applicability. Second, the co-design process produced recommendations but did not include a formal evaluation of their implementation or impact. Assessing the outcomes of these interventions will be essential for determining their sustainability and effectiveness. Finally, at the time of data collection, Fraser Health had only launched a single virtual hospital-at-home service, the VPU, and had not yet determined which other clinical populations would be prioritized for future expansion. Consequently, our findings do not speak of the implementation of a specific model beyond psychiatry. However, this early engagement with patients and caregivers was intentional; by eliciting perspectives before clinical parameters were finalized, the co-design process generated insights that can shape a more patient-centered and equitable foundation for service development across diverse populations.

Reflections and Challenges of Implementing EBCD in a Health System Context

Implementing EBCD in a large and complex health system context such as Fraser Health presented unique challenges and valuable learning opportunities. While the process effectively centered on patient voices and stimulated rich discussions, it also disrupted the status quo. Recruiting physician participants proved difficult, despite outreach efforts; only one physician (an acting Medical Information Officer) attended the co-design session. This highlights the tension between clinical demands and the time investment required for meaningful engagement in participatory design.

In addition, this project was not embedded within the core design cycle of the virtual hospital-at-home program, which limited the extent to which recommendations could be immediately integrated. Some suggestions generated during the workshop also fell outside the direct influence of the Fraser Health team, highlighting the need for stronger system-wide support for patient-driven ideas and cross-organizational collaboration.

Nevertheless, the project demonstrated the value of partnership with community organizations and the importance of cultural and linguistic accessibility in participatory research. Language accommodations during interviews improved accessibility, and the use of a video touchpoint film was particularly effective in centering patient stories and perspectives during the workshop. Divergent views, such as a participant strongly opposing virtual care, proved essential in challenging assumptions and

XSL•FO
RenderX

underscoring the diversity of opinions within the patient population. These perspectives enriched the discussion and ultimately strengthened the relevance of the proposed solutions.

A final lesson learned was that for many health system users, engaging in a process without a guaranteed "end product" (such as a finalized service design) was a leap of faith. Although the impact of the workshop was palpable, participants expressed feeling heard, empowered, and connected. For some, it was a transformative experience, reinforcing the need for participatory approaches in health service design that go beyond consultation to foster real collaboration.

Future Directions

Building on this foundational work, future efforts should prioritize piloting and evaluating co-designed interventions across diverse populations and clinical settings. There is also a need to explore how EBCD can be embedded as a routine part of virtual service development rather than used episodically. As technology continues to evolve, virtual care systems must be supported by ongoing, rather than one-time, engagement processes to remain equitable, culturally relevant, and trusted by the communities that they aim to serve.

Conclusions

This project illustrates how EBCD can serve as an empowering approach to guide the development of equitable and patient-centered virtual hospital-at-home services. By engaging patients, caregivers, health care providers, and community leaders throughout the co-design process, the project identified both barriers and actionable strategies for improving virtual care delivery. These findings underscore the importance of integrating cultural competence, digital literacy, caregiver inclusion, and systems-level coordination into virtual care models, particularly when serving diverse populations. As virtual hospital-at-home programs continue to expand, intentional and participatory approaches such as EBCD will be essential to ensuring that these innovations are not only effective but also inclusive, sustainable, and trusted by those they are designed to serve. To ensure that virtual care achieves its full promise, future initiatives must embed equity and patient partnership as core, not optional, design principles.

Acknowledgments

This work was supported by Health Research BC and the Health Research Foundation of Innovative Medicines Canada Team Grant in Virtual Care. The funders had no role in the study design, data collection, analysis, interpretation, writing of the manuscript, or the decision to submit the paper for publication. The authors sincerely thank the South Asian community members who participated in this study and shared their experiences, as well as the steering committee for their guidance and support throughout the project.

Data Availability

The deidentified interview transcripts and other data generated or analyzed during this study are available from the corresponding author upon reasonable request, subject to ethics approval and participant confidentiality agreements.

Conflicts of Interest

None declared.

References

1. Montenegro M, MacPherson M. Virtual hospitals: a rapid review. OSF. 2023. URL: https://doi.org/10.17605/OSF.IO/M3A4B [accessed 2025-10-14]
2. Lewis G, Bardsley M, Vaithianathan R, et al. Do "virtual wards" reduce rates of unplanned hospital admissions, and at what cost? A research protocol using propensity matched controls. Int J Integr Care 2011 Apr;11:e079. [doi: 10.5334/ijic.654] [Medline: 21949489]
3. D'cruz M, Banerjee D. "An invisible human rights crisis": the marginalization of older adults during the COVID-19 pandemic - an advocacy review. Psychiatry Res 2020 Oct;292:113369. [doi: 10.1016/j.psychres.2020.113369] [Medline: 32795754]
4. Chen A, Ayub MH, Mishuris RG, et al. Telehealth policy, practice, and education: a position statement of the Society of General Internal Medicine. J Gen Intern Med 2023 Aug;38(11):2613-2620. [doi: 10.1007/s11606-023-08190-8] [Medline: 37095331]
5. van Draanen J, Shafique A, Farissi A, et al. How to offer culturally relevant type 2 diabetes screening: lessons learned from the South Asian Diabetes Prevention Program. Can J Diabetes 2014 Oct;38(5):329-333. [doi: 10.1016/j.jcjd.2013.11.008] [Medline: 24797496]
6. Bainey KR, Jugdutt BI. Increased burden of coronary artery disease in South-Asians living in North America. Need for an aggressive management algorithm. Atherosclerosis 2009 May;204(1):1-10. [doi: 10.1016/j.atherosclerosis.2008.09.023] [Medline: 18980768]

XSL•FO
RenderX

7. Rana A, de Souza RJ, Kandasamy S, Lear SA, Anand SS. Cardiovascular risk among South Asians living in Canada: a systematic review and meta-analysis. CMAJ Open 2014 Jul;2(3):E183-E191. [doi: 10.9778/cmajo.20130064] [Medline: 25295238]
8. South Asian health report. : Fraser Health; 2015 URL: https://www.fraserhealth.ca/-/media/Project/FraserHealth/FraserHealth/Health-Topics/South-Asian-Health/201606_South_Asian_Health_Report.pdf [accessed 2025-10-14]
9. Islam F, Khanlou N, Tamim H. South Asian populations in Canada: migration and mental health. BMC Psychiatry 2014 May 26;14:154. [doi: 10.1186/1471-244X-14-154] [Medline: 24884792]
10. Godley J. Everyday discrimination in Canada: prevalence and patterns. Can J Soc 2018;43(2):111-142. [doi: 10.29173/cjs29346]
11. MacPherson M. Immigrant, refugee, and indigenous Canadians' experiences with virtual health care services: rapid review. JMIR Hum Factors 2023 Oct 9;10:e47288. [doi: 10.2196/47288] [Medline: 37812489]
12. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021 Aug 17;21(1):1556. [doi: 10.1186/s12889-021-11623-w] [Medline: 34399716]
13. MacPherson MM, Kapadia S. Barriers and facilitators to patient-to-provider messaging using the COM-B model and theoretical domains framework: a rapid umbrella review. BMC Digit Health 2023;1(1):33. [doi: 10.1186/s44247-023-00033-0]
14. Kasoju N, Remya NS, Sasi R, et al. Digital health: trends, opportunities and challenges in medical devices, pharma and bio-technology. CSIT 2023 Apr;11(1):11-30. [doi: 10.1007/s40012-023-00380-3]
15. Mumtaz H, Riaz MH, Wajid H, et al. Current challenges and potential solutions to the use of digital health technologies in evidence generation: a narrative review. Front Digit Health 2023;5:1203945. [doi: 10.3389/fdgth.2023.1203945] [Medline: 37840685]
16. About Fraser Health. Fraser Health. URL: https://www.fraserhealth.ca/about-us/about-fraser-health [accessed 2025-04-07]
17. Census profile, 2021 census of population. Statistics Canada. URL: https://www12.statcan.gc.ca/census-recensement/2021/dp-pd/prof/index.cfm?Lang=E [accessed 2025-04-07]
18. Why being admitted to hospital in B.C. may soon mean being sent home instead. Vancouver Sun. 2023. URL: https://vancouversun.com/health/local-health/why-being-admitted-hospital-bc-may-soon-mean-being-sent-home-instead [accessed 2023-09-12]
19. Bate P, Robert G. Bringing User Experience to Healthcare Improvement: The Concepts, Methods and Practices of Experience-Based Design: CRC Press; 2023.
20. Fylan B, Tomlinson J, Raynor DK, Silcock J. Using experience-based co-design with patients, carers and healthcare professionals to develop theory-based interventions for safer medicines use. Res Soc Adm Pharm 2021 Dec;17(12):2127-2135. [doi: 10.1016/j.sapharm.2021.06.004]
21. Gustavsson SM, Andersson T. Patient involvement 2.0: experience-based co-design supported by action research. Action Research 2019 Dec;17(4):469-491. [doi: 10.1177/1476750317723965]
22. Morales-Garzón S, Parker LA, Hernández-Aguado I, González-Moro Tolosana M, Pastor-Valero M, Chilet-Rosell E. Addressing health disparities through community participation: a scoping review of co-creation in public health. Healthcare (Basel) 2023 Apr 4;11(7):1034. [doi: 10.3390/healthcare11071034] [Medline: 37046961]
23. Gander-Zaucker S, Unwin GL, Larkin M. The feasibility and acceptability of an experience-based co-design approach to reducing domestic abuse. Societies (Basel) 2022;12(3):93. [doi: 10.3390/soc12030093]
24. MacPherson M. Hospital-at-Home for South Asian communities in British Columbia, Canada: qualitative interview study. JMIR Hum Factors (forthcoming) 2025. [doi: 10.2196/79675]
25. Serrat O. The Five Whys technique. In: Serrat O, editor. Knowledge Solutions: Tools, Methods, and Approaches to Drive Organizational Performance: Springer; 2017. [doi: 10.1007/978-981-10-0983-9_32]
26. Knapp J, Zeratsky J, Kowitz B. Sprint: How to Solve Big Problems and Test New Ideas in Just Five Days: Simon and Schuster; 2016.
27. Jones L, Nabil S, Girouard A. Wearable crazy eights: wearable ideation methods for encouraging divergent design concepts. Presented at: Proceedings of the Fifteenth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '21); Feb 14-17, 2021; Salzburg, Austria. [doi: 10.1145/3430524.3442464]
28. The importance of co-design in telemedicine. Institute for Healthcare Improvement. 2022. URL: https://www.ihi.org/library/blog/importance-co-design-telemedicine [accessed 2025-08-27]
29. Badr J, Motulsky A, Denis JL. Digital health technologies and inequalities: a scoping review of potential impacts and policy recommendations. Health Policy 2024 Aug;146:105122. [doi: 10.1016/j.healthpol.2024.105122] [Medline: 38986333]
30. Wei Z. Navigating digital learning landscapes: unveiling the interplay between learning behaviors, digital literacy, and educational outcomes. J Knowl Econ 2023;15(3):10516-10546. [doi: 10.1007/s13132-023-01522-3]
31. El Messaoudi M. The impact of a blended learning model on undergraduate university students' digital literacy skills: empirical evidence from higher education. Educ Process Int J 2024;13(4):84-101. [doi: 10.22521/edupij.2024.134.5]
32. Bombard Y, Baker GR, Orlando E, et al. Engaging patients to improve quality of care: a systematic review. Implement Sci 2018 Jul 26;13(1):98. [doi: 10.1186/s13012-018-0784-z] [Medline: 30045735]
33. Hall JN, Chartier LB. Learning from a regional approach: integration to scale, spread, and sustain virtual urgent care. Inquiry 2023;60:469580221143273. [doi: 10.1177/00469580221143273] [Medline: 36624685]

XSL•FO
RenderX

34. Zainal H, Xiaohui X, Thumboo J, Seah SJ, Leng LL, Kok Yong F. Exploring caregiver challenges, digital health technologies, and healthcare support: a qualitative study. Front Digit Health 2025;7:1587162. [doi: 10.3389/fdgth.2025.1587162] [Medline: 40547053]
35. Ganta T, Appel JM, Genes N. Patient portal access for caregivers of adult and geriatric patients: reframing the ethics of digital patient communication. J Med Ethics 2023 Mar;49(3):156-159. [doi: 10.1136/medethics-2021-107759] [Medline: 35437282]
36. Han S. Facilitating patient portal shared access to support age-friendly health care. JAMA Netw Open 2025 Feb 3;8(2):e2461814. [doi: 10.1001/jamanetworkopen.2024.61814] [Medline: 39998834]
37. Talal AH, Sofikitou EM, Jaanimägi U, Zeremski M, Tobin JN, Markatou M. A framework for patient-centered telemedicine: application and lessons learned from vulnerable populations. J Biomed Inform 2020 Dec;112:103622. [doi: 10.1016/j.jbi.2020.103622] [Medline: 33186707]

Abbreviations

EBCD: experience-based co-design
PICS: Progressive Intercultural Community Services
VPU: virtual psychiatric unit

Edited by A Bamgboje-Ayodele; submitted 25.06.25; peer-reviewed by A Barker, E Foglia; revised version received 28.08.25; accepted 17.09.25; published 05.11.25.

Please cite as:
Kandola M, Wong E, Paquin R, Arora K, Sharda H, Deol R, Jung M, Montenegro M, MacPherson M
Designing for Patient-Centered Care and Equity in Virtual Hospital-at-Home Models: Quality Improvement Initiative Using Experience-Based Co-Design
JMIR Hum Factors 2025;12:e79679
URL: https://humanfactors.jmir.org/2025/1/e79679
doi:10.2196/79679

© Mahabhir Kandola, Emma Wong, Robert Paquin, Kamal Arora, Harroop Sharda, Roman Deol, Mary Jung, Maria Montenegro, Megan MacPherson. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 5.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring User Experience and the Therapeutic Relationship of Short-Term Avatar-Based Psychotherapy: Qualitative Pilot Study

Byeul Jang[1*], MS; Chisung Yuh[2*], DDS, MS; Hyeri Lee[3], MFA; Yu-Bin Shin[1], PhD; Heon-Jeong Lee[1], MD, PhD; Eun Kyoung Kang[4], MD, PhD; Jeongyun Heo[5], PhD; Chul-Hyun Cho[1,6], MD, PhD

[1]Department of Psychiatry, Korea University College of Medicine, 73 Goryeodae-ro, Seongbuk-gu, Seoul, Republic of Korea

[2]Department of Medical Education, Korea University College of Medicine, Seoul, Republic of Korea

[3]Annenberg School for Communication and Journalism, University of Southern California, Los Angeles, CA, United States

[4]KakaoHealthcare Corp., Seongnam, Republic of Korea

[5]Department of Smart Experience Design, Kookmin University, Seoul, Republic of Korea

[6]Department of Biomedical Informatics, Korea University College of Medicine, Seoul, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Chul-Hyun Cho, MD, PhD
Department of Psychiatry, Korea University College of Medicine, 73 Goryeodae-ro, Seongbuk-gu, Seoul, Republic of Korea

Abstract

Background: The rapid advancement of telehealth has led to the emergence of avatar-based psychotherapy (ABP), which combines the benefits of anonymity with nonverbal communication. With the adoption of remote mental health services, understanding the efficacy and user experience of ABP has become increasingly important.

Objective: This study aimed to explore the user experience and therapeutic relationship formation in short-term ABP environments, focusing on psychological effects, user satisfaction, and critical factors for implementation.

Methods: This qualitative study involved 18 adult participants (8 women and 10 men). Participants engaged in two short-term ABP sessions (approximately 50 minutes per session) over 2 weeks, using an ABP metaverse system prototype. Semistructured in-depth interviews were conducted with both the participants and therapists before and after the ABP sessions. The interviews were conducted via an online platform, with each interview lasting approximately 30 minutes. The key topics included the sense of intimacy, communication effectiveness of avatar expressions, emotions toward one's avatar, concentration during sessions, and perceived important aspects of the ABP. Data were analyzed using thematic analysis.

Results: The analysis revealed 3 main themes with 8 subthemes: (1) reduction of psychological barriers through avatar use (subthemes: anonymity, ease of access, self-objectification, and potential for self-disclosure); (2) importance of the avatar–self-connection in therapeutic relationship formation (subthemes: avatar self-relevance and avatar–self-connection fostering intimacy and trust); and (3) importance of nonverbal communication (subthemes: significance of nonverbal expressions and formation of empathy and trust through nonverbal expressions). Participants reported enhanced comfort and self-disclosure owing to the anonymity provided by avatars, while emphasizing the importance of avatar customization and the role of nonverbal cues in facilitating communication and building rapport.

Conclusions: This pilot study provides valuable insights into the short-term ABP user experience and therapeutic relationship formation. Our findings suggest that ABP has the potential to reduce barriers to therapy through anonymity, ease of access, and potential for self-disclosure, while allowing for meaningful nonverbal communication. The avatar–self-connection emerged as a crucial factor in the effectiveness of ABP, highlighting the importance of avatar customization in enhancing user engagement and therapeutic outcomes. Future research and development in ABP should focus on improving avatar customization options, enhancing the fidelity of nonverbal cues, and investigating the long-term effectiveness of ABP compared with traditional face-to-face therapy.

Trial Registration: CRIS KCT0009695; https://tinyurl.com/2a48s7dh

(JMIR Hum Factors 2025;12:e66158) doi:10.2196/66158

KEYWORDS

avatar-based psychotherapy; telehealth; therapeutic relationship; user experience; anonymity; nonverbal communication; mental health; mobile phone

XSL•FO
RenderX

Introduction

Rapid advancements in information technology have led to significant developments in telehealth, particularly in addressing mental health issues, such as anxiety, depression, and stress-related disorders [1]. The COVID-19 pandemic has accelerated the transition from traditional face-to-face clinical services to telehealth solutions [2]. This shift has created new opportunities for individuals with difficulty accessing in-person psychotherapy owing to geographical constraints or personal reservations [3]. Avatar-based psychotherapy (ABP) has gained attention because of its unique ability to combine the benefits of anonymity and nonverbal communication. Recent research has demonstrated the potential effectiveness of metaverse-based counseling approaches compared to traditional in-person settings [4].

The therapeutic relationship between the client and psychotherapist is crucial for successful therapeutic outcomes. However, the dynamics of this relationship in ABP environments require further investigation [5]. Self-disclosure, a key element in relationship building, may be facilitated differently in ABP settings than in traditional face-to-face or video-based online therapy [6]. The concept of avatar–self-connection, which refers to the psychological link between users and their digital representations, may play a significant role in the effectiveness of ABP [7].

This study aimed to explore the user experience and therapeutic relationship formation in short-term ABP environments. Specifically, we sought to understand the psychological effects and satisfaction levels of users engaging in ABP and to identify critical factors to consider when implementing ABP systems.

Methods

Ethical Considerations

This study was conducted in accordance with the principles of the Declaration of Helsinki and approved by the Institutional Review Board of Korea University Anam Hospital (IRB No: 2023AN0504). All participants provided written informed consent before participation. Participants were informed that their participation was voluntary and that they could withdraw at any time without consequences. All interviews were audio-recorded with prior consent and subsequently transcribed for analysis. The data collected in this study were anonymized to ensure the privacy and confidentiality of all participants. This manuscript and its supplementary materials do not include any images or identifiable features of research participants.

Study Design and Data Collection

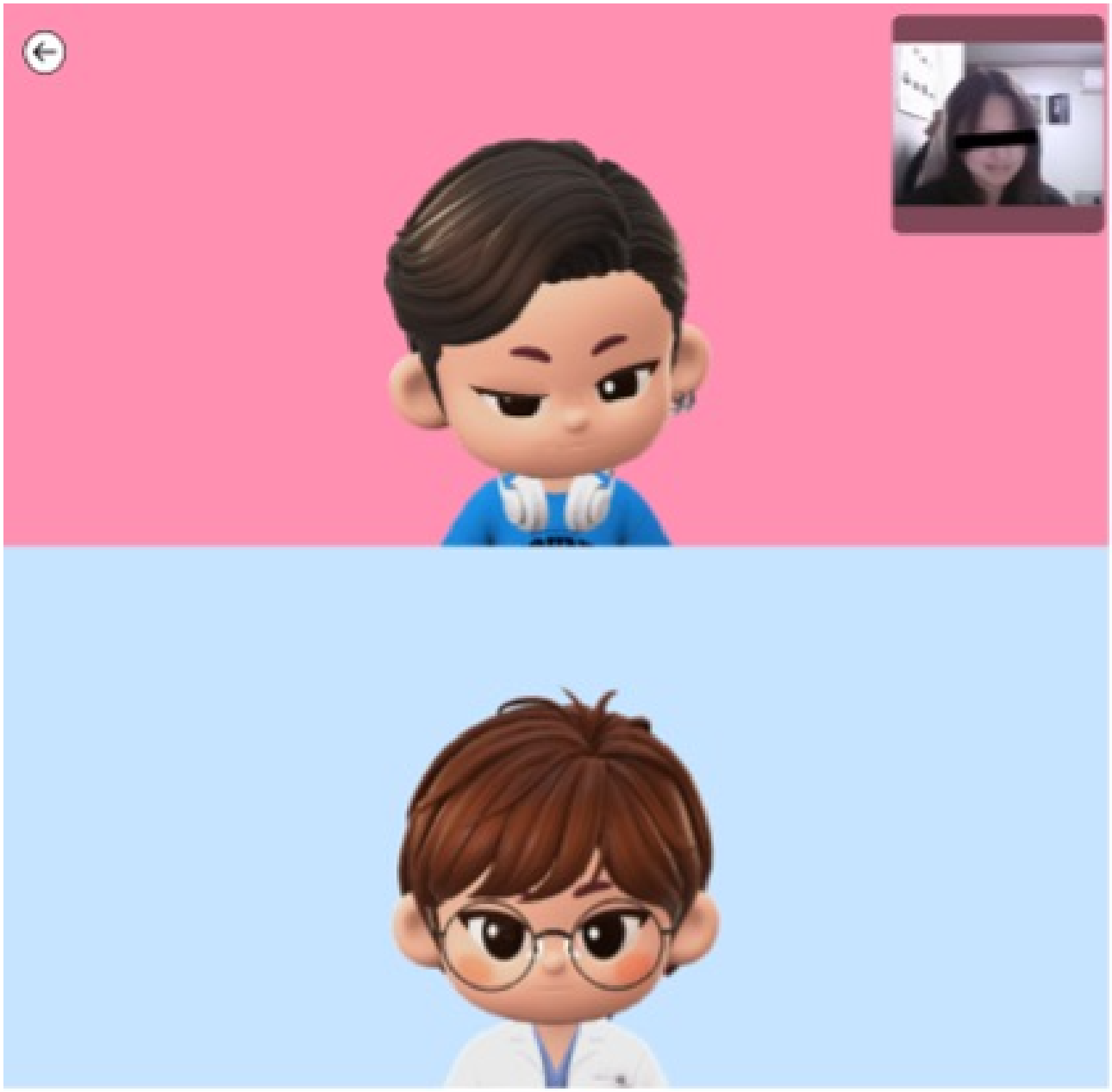
This study was designed as a short-term intervention pilot study to explore initial user experiences with ABP. The two-session format was chosen to capture immediate reactions and early therapeutic relationship formation in the ABP environment. We conducted a qualitative study to gain an in-depth understanding of user experiences in ABP sessions. Eligible participants were adults aged 19 - 70 years who self-reported their need for psychological health management and had no previous experience with avatar-based counseling. We excluded individuals with intellectual disabilities, organic brain damage, or those currently receiving psychiatric treatment for major mental disorders (including major mood disorders, anxiety disorders, and schizophrenia spectrum disorders). Additionally, participants were required to have access to and the ability to use smartphones or computers.

An avatar that reflected facial movements in real time was established using a metaverse system prototype (Figure 1). This system allowed participants to engage in psychotherapy sessions through avatars, enabling verbal and nonverbal communication. Eighteen adult participants (8 women) were initially recruited in Seoul, South Korea. The participants engaged in two ABP sessions (approximately 50 minutes per session) over 2 weeks.

Semistructured in-depth interviews were conducted with both the participants and psychotherapists before and after the ABP sessions. These interviews, each lasting 30 minutes, were conducted using the Zoom platform (Zoom Video Communications, Inc.). The key topics included the sense of intimacy and communication effectiveness of avatar expressions, emotions and feelings toward one's avatar, concentration during the session, and perceived important aspects of the ABP sessions.

XSL•FO
RenderX

Figure 1. The avatar-based psychotherapy metaverse system prototype. Top avatar: participant; bottom avatar: psychotherapist. The real face box at the top can be turned off; in a real psychotherapy session, the real face is turned off and is conducted with only the avatar.



Data Analysis

We employed thematic analysis to analyze the data, using ATLAS.ti software (Version 23; ATLAS.ti Scientific Software Development GmbH, Berlin, Germany) [8]. The analysis was conducted in 6 phases: familiarization with the data, generating initial codes, searching for themes, reviewing themes, defining and naming themes, and producing reports. Four researchers (JB, HL, JH, and CHC) participated in the initial familiarization phase and carefully reviewed the transcripts before individual coding. The 174 initial codes were systematically categorized through iterative analysis. Through team discussions, these codes were refined and consolidated into 3 main themes of psychological barriers (41 codes), avatar–self-connection (29 codes), and nonverbal communication (38 codes). The remaining codes were excluded as they did not directly address the research questions.

Results

The thematic analysis revealed 3 main themes with 8 subthemes, as shown in Table 1. The 174 initial codes were systematically categorized through iterative analysis. Through team discussions, these codes were refined and consolidated into 3 main themes of psychological barriers (41 codes), avatar–self-connection (29 codes), and nonverbal communication (38 codes). The remaining codes were excluded as they did not directly address the research questions.

XSL•FO
RenderX

Table . Thematic framework of user experiences and therapeutic relationship formation in avatar-based psychotherapy: main themes, subthemes, code counts, and representative codes.

Main themes	Subthemes	Code counts	Representative codes
Reduction of psychological barriers through avatar use	1.1 Anonymity	12	• Anonymity reduces social anxiety. • Sense of security through anonymity. • Honest communication enabled. • Anonymity prevents judgment. • Reduced fear of personal recognition.
	1.2 Ease of access	10	• Flexible scheduling. • No travel required. • Accessible from anywhere. • Easier therapy initiation. • Reduced logistical challenges.
	1.3 Self-objectification	8	• Detached view encourages self-reflection. • Avatar mirrors emotional state. • Enhances self-awareness. • Observing oneself via the avatar. • Identifies overlooked emotions.
	1.4 Potential for self-disclosure	11	• Easier to discuss sensitive topics. • Safe space for expression. • Encourages honest sharing. • Enables communication without barriers. • Reduces hesitation in disclosure.
Importance of avatar–self-connection	2.1 Avatar self-relevance	15	• Personalized avatars foster comfort. • Realism enhances connection. • Reflects personal identity. • Improves immersion. • Mirrors user's traits effectively.
	2.2 Fostering intimacy and trust	14	• Facilitates faster rapport. • Acts as a conversational bridge. • Builds therapeutic trust. • Supports meaningful dialogue. • Encourages emotional openness.

XSL•FO
RenderX

Main themes	Subthemes	Code counts	Representative codes
Importance of nonverbal communication	3.1 Significance of nonverbal expressions	20	• Gestures enhance communication. • Nonverbal cues reduce misunderstandings. • Expressive avatars improve immersion. • Adds emotional depth. • Reinforces communication effectiveness.
	3.2 Empathy and trust formation	18	• Mirrors emotions effectively. • Conveys empathy through gestures. • Builds trust via nonverbal cues. • Enhances emotional understanding. • Reduces virtual communication gaps.

Theme 1: Reduction of Psychological Barriers Through Avatar Use

Theme 1, the reduction of psychological barriers through avatar use, emerged as the most prominent theme with 41 codes. These codes were distributed across 4 subthemes: anonymity (12 codes), ease of access (10 codes), self-objectification (8 codes), and potential for self-disclosure (11 codes). Participants reported that the use of avatars in psychotherapy sessions helped lower psychological barriers primarily through 3 mechanisms.

Subtheme 1.1: Anonymity

The anonymity provided by avatars was frequently mentioned as a key factor in facilitating open and honest communication. One participant noted, “Knowing that the therapist does not know my real face gives me a sense of security. Even if we met in real life, they would not recognize me, which is reassuring in a way.”

Subtheme 1.2: Ease of Access

The convenience of accessing ABP services through a metaverse platform was highlighted as a significant advantage. One participant stated, “The fact that psychotherapy sessions can be done from anywhere makes it more accessible and easier to form a rapport. I think anonymous ABP is more effective in this regard.”

Subtheme 1.3: Self-Objectification

Avatars facilitated self-objectification, allowing participants to view themselves from a detached perspective. One participant observed, “When I occasionally looked at my avatar, I noticed that I was not smiling much. It made me think, ‘I look a bit pitiful.’ This showed me how I appeared.”

Subtheme 1.4: Potential for Self-Disclosure

Avatars facilitated self-disclosure. One participant noted, “For people who want psychotherapy without revealing themselves too much, expressing themselves through an avatar in the metaverse can make it easier to talk honestly with the therapist.”

Theme 2: Importance of Avatar–Self-Connection in Therapeutic Relationship Formation

The importance of avatar–self-connection in therapeutic relationship formation emerged as the second major theme, comprising 29 codes. These were distributed between 2 key subthemes: avatar self-relevance (15 codes) and avatar–self-connection fostering intimacy and trust (14 codes). The results showed that the degree of connection between users and their avatar played a crucial role in the therapeutic relationship formation within the ABP environment.

Subtheme 2.1: Avatar Self-Relevance

Participants expressed a desire for their avatars to reflect aspects of their real selves. One participant noted, “If I could customize the avatar more to my liking and feel that it truly represents me, I think I would feel more completely connected to it.”

Subtheme 2.2: Avatar–Self-Connection Fostering Intimacy and Trust

A strong avatar–self-connection enhanced the sense of intimacy and trust in the therapeutic relationship. One participant shared, “When I felt that this avatar was really me, I could immerse myself more deeply in the session and form a stronger therapeutic relationship with the therapist.”

Theme 3: Importance of Nonverbal Communication

The importance of nonverbal communication emerged as a significant theme with 38 codes, distributed across 2 subthemes: significance of nonverbal expressions (20 codes) and formation of empathy and trust through nonverbal expressions (18 codes). The thematic analysis highlighted the significant role of nonverbal communication in ABP, emphasizing its impact on self-expression, empathy, and trust-building.

Subtheme 3.1: Significance of Nonverbal Expressions

The participants emphasized the importance of nonverbal cues in enhancing communication. One participant stated, “If the avatar could show hand gestures or body language, I think it would be even better. These physical movements can have psychological significance.”

XSL•FO
RenderX

Subtheme 3.2: Formation of Empathy and Trust Through Nonverbal Expressions

Nonverbal communication through avatars enhanced empathy and trust in the therapeutic relationship. One participant shared, "I think the avatar helped by expressing my gestures and expressions to some extent. This somewhat overcomes communication errors that can occur when we are not face-to-face."

Discussion

Study Findings and Comparison With Previous Findings

This study provides critical insights into ABP, highlighting 3 interconnected dimensions: psychological barrier reduction, avatar–self-connection, and the role of nonverbal communication in metaverse therapeutic systems.

The anonymity, accessibility, and potential for self-disclosure provided by avatars create a unique therapeutic environment that lowers the resistance to mental health engagement, consistent with previous research on online self-disclosure [5]. A novel finding of this study was the self-objectification through avatar use, suggesting a potential mechanism for enhanced self-awareness that resonates with self-distancing theories in psychology [9]. This indicates that ABP may offer a natural pathway for adaptive self-reflection, a key component in many therapeutic approaches.

The self-objectification observed in this study demonstrates an interesting contrast with findings from video conferencing research, which found that viewing one's actual image during video conferences can lead to greater cognitive burden and negative psychological effects, particularly for individuals with high public self-consciousness [10]. However, our findings suggest that avatar-mediated self-observation may offer distinct advantages. Unlike video conferencing where users see their actual image, avatar representation creates a beneficial psychological distance that facilitates more objective self-reflection while maintaining emotional engagement. This unique characteristic of avatar-mediated interaction may explain why participants in our study reported enhanced self-awareness without the negative psychological impacts often associated with direct self-viewing in video conferences. The avatar-mediated environment provides a unique balance between self-awareness and psychological comfort, potentially making it particularly suitable for therapeutic contexts where self-reflection is crucial but emotional safety needs to be maintained.

Furthermore, the avatar–self-connection played a crucial role in the therapeutic relationship formation, extending previous work on avatar realism [7]. Our findings directly linked this connection to the quality of the therapeutic relationship, underscoring the potential significance of avatar customization in fostering engagement and improving outcomes. The avatar may serve as a bridge between the client's inner world and the therapeutic space, facilitating deeper exploration of personal narratives.

Our observations of nonverbal communication in ABP expand on previous research, highlighting how even limited nonverbal cues can significantly impact the therapeutic process [11]. This adaptability suggests that ABP taps into fundamental aspects of human communication, offering a rich alternative to traditional face-to-face therapy.

The interplay among these elements in ABP creates a unique therapeutic ecosystem that balances self-disclosure and self-protection. As we continue to explore the potential of ABP, it is crucial to consider its place within the broader context of mental health care. ABP can be viewed as a complementary tool that can lead to more comprehensive and accessible care, potentially reaching individuals who might otherwise not receive treatment [12,13].

Strengths and Limitations

The findings of this study have several implications for clinical practice. First, ABP may be particularly beneficial for clients hesitant to engage in traditional face-to-face therapy because of social anxiety, stigma, or geographical limitations [14]. The anonymity and ease of access provided by ABP could serve as a stepping stone for these individuals to transition to other forms of therapy, if needed. Second, clinicians using ABP should be aware of the importance of the avatar–self-connection and consider ways to enhance this connection during therapy sessions. This might involve discussing the client's choice of avatar and exploring how it relates to their self-perception or therapeutic goals. However, although ABP offers many advantages, clinicians must be aware of its limitations. The reduced nonverbal cues in ABP compared with face-to-face therapy require clinicians to develop new skills in reading and interpreting client expressions and emotions through avatars.

Conclusion

These findings not only contribute to the growing body of knowledge on telehealth interventions, but also prompt us to reconsider the nature of the therapeutic relationship in the digital age. As technology continues to reshape our modes of interaction, ABP stands at the forefront of a new paradigm in mental health care that harnesses the power of the virtual world to foster real-world healing and growth.

Acknowledgments

This work was supported by grants from the National Research Foundation (NRF) of Korea (NRF-2021R1A5A8032895 and NRF-2022M3C1B6080866), a grant from the Institute of Information and Communications Technology Planning and Evaluation (IITP; RS-2023-00224823), and a grant from the Information and Communications Promotion Fund (H0601-24-1017) and Broadcasting Communications Development Fund (D1013-24-1003) through the National IT Industry Promotion Agency (NIPA),

XSL•FO
RenderX

Republic of Korea. The funders had no role in the study design; collection, analysis, and interpretation of data; writing of the report; or the decision to submit the paper for publication.

Authors' Contributions

JH and CHC are cocorresponding authors and are responsible for the data and materials, manuscript submission, peer review, publication process, authorship details, and ethics committee approval. EKK, JH, and CHC conceived and designed the study. JB, CY, HL, JH, and CHC performed data analyses. JB, CY, HL, JH, and CHC wrote the first draft of the manuscript. JB, HL, EKK, JH, and CHC participated in data collection. All authors edited all versions of the manuscript. All authors were involved in interpreting the results and have read, commented on, and approved the final version of the manuscript.

Conflicts of Interest

EKK is a member of Kakao Healthcare, the company that developed the avatar-based psychotherapy metaverse system prototype used in this study. The other authors declare no conflicts of interest.

References

1. Williams MG, Stott R, Bromwich N, Oblak SK, Espie CA, Rose JB. Determinants of and barriers to adoption of digital therapeutics for mental health at scale in the NHS. BMJ Innov 2020 Jul;6(3):92-98. [doi: 10.1136/bmjinnov-2019-000384]
2. Reay R, Kisely SR, Looi JCL. Better Access: substantial shift to telehealth for allied mental health services during COVID-19 in Australia. Aust Health Rev 2021 Dec;45(6):675-682. [doi: 10.1071/AH21162] [Medline: 34743787]
3. Braude G, Mohi S, Quinlan E, Shoullis A, Collison J. Psychologists and their self-mirror: therapy experiences via videoconferencing. Clin Psychol (Aust Psychol Soc) 2023 Sep 2;27(3):384-391. [doi: 10.1080/13284207.2023.2245953]
4. Cho S, Kang J, Baek WH, Jeong YB, Lee S, Lee SM. Comparing counseling outcome for college students: Metaverse and in-person approaches. Psychother Res 2024 Nov;34(8):1117-1130. [doi: 10.1080/10503307.2023.2270139] [Medline: 37848177]
5. Bente G, Rüggenberg S, Krämer NC, Eschenburg F. Avatar-mediated networking: increasing social presence and interpersonal trust in net-based collaborations. Human Comm Res 2008 Apr;34(2):287-318. [doi: 10.1111/j.1468-2958.2008.00322.x]
6. Hooi R, Cho H. Avatar-driven self-disclosure: The virtual me is the actual me. Comput Human Behav 2014 Oct;39:20-28. [doi: 10.1016/j.chb.2014.06.019]
7. Kim DY, Lee HK, Chung K. Avatar-mediated experience in the metaverse: The impact of avatar realism on user-avatar relationship. J Retail Consum Serv 2023 Jul;73:103382. [doi: 10.1016/j.jretconser.2023.103382]
8. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
9. Kross E, Ayduk O. Making meaning out of negative experiences by self-distancing. Curr Dir Psychol Sci 2011 Jun;20(3):187-191. [doi: 10.1177/0963721411408883]
10. Kuhn KM. The constant mirror: Self-view and attitudes to virtual meetings. Comput Human Behav 2022 Mar;128:107110. [doi: 10.1016/j.chb.2021.107110]
11. Bailenson JN, Yee N, Merget D, Schroeder R. The effect of behavioral realism and form realism of real-time avatar faces on verbal disclosure, nonverbal disclosure, emotion recognition, and copresence in dyadic interaction. Pres Teleoperators Virtual Environ 2006 Aug 1;15(4):359-372. [doi: 10.1162/pres.15.4.359]
12. Kim SW, Baek SH, Cho CH. Psychosocial and digital intervention for patients with schizophrenia. J Korean Med Assoc 2024 Feb;67(2):103-114. [doi: 10.5124/jkma.2024.67.2.103]
13. Sestino A, D'Angelo A. My doctor is an avatar! The effect of anthropomorphism and emotional receptivity on individuals' intention to use digital-based healthcare services. Technol Forecast Soc Change 2023 Jun;191:122505. [doi: 10.1016/j.techfore.2023.122505]
14. Cho CH. The beginning of the era of digital therapeutics in Korea: challenges and opportunities. J Korean Med Sci 2023 May 29;38(21):e166. [doi: 10.3346/jkms.2023.38.e166] [Medline: 37270919]

Abbreviations

ABP: avatar-based psychotherapy

XSL•FO
RenderX

Edited by A Kushniruk; submitted 05.09.24; peer-reviewed by D Jung, H Kwon; revised version received 02.12.24; accepted 23.12.24; published 05.02.25.

Please cite as:
Jang B, Yuh C, Lee H, Shin YB, Lee HJ, Kang EK, Heo J, Cho CH
Exploring User Experience and the Therapeutic Relationship of Short-Term Avatar-Based Psychotherapy: Qualitative Pilot Study
JMIR Hum Factors 2025;12:e66158
URL: https://humanfactors.jmir.org/2025/1/e66158
doi:10.2196/66158

© Byeul Jang, Chisung Yuh, Hyeri Lee, Yu-Bin Shin, Heon-Jeong Lee, Eun Kyoung Kang, Jeongyun Heo, Chul-Hyun Cho. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 5.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Research Agenda for Perinatal Innovation and Digital Health Project: Human-Centered Approach to Multipartner Research Agenda Codevelopment

Haneen Amhaz[1,2,3], BSc, MPH; Sally Xuanping Chen[1,2], BSc, MFS; Amanee Elchehimi[1,2], MA, MPH; Kylin Jialin Han[1,2], BSc, BE, MFS, MBA; Jade Morales Gil[1,2], BSN; Lu Yao[1,2], BA; Marianne Vidler[1,4], MPH, PhD; Kathryn Berry-Einarson[5], BA, MPH; Kathryn Dewar[1], PhD; May Tuason[6,7], BSN, MBA; Nicole Prestley[1], BA; Quynh Doan[2,8], MHSc, MDCM, FRCPC, PhD; Tibor van Rooij[2,9], BSc, PhD; Tina Costa[6], BSN, MA; Gina Ogilvie[1,2,10], MSc, DrPH, FCFP, MD; Beth A Payne[1,2,10], BSc, PhD

[1]Women's Health Research Institute, Vancouver, BC, Canada

[2]BC Children's Hospital Research Institute, Vancouver, BC, Canada

[3]Faculty of Health Sciences, Simon Fraser University, Burnaby, BC, Canada

[4]Department of Obstetrics and Gynaecology, Faculty of Medicine, The University of British Columbia, Vancouver, BC, Canada

[5]Quality and Research, Perinatal Services BC, Provincial Health Services Authority, Vancouver, BC, Canada

[6]Provincial Digital Health and Information Services, Provincial Health Services Authority, Vancouver, BC, Canada

[7]BC Ministry of Health, Victoria, BC, Canada

[8]Department of Pediatrics, Faculty of Medicine, The University of British Columbia, Vancouver, BC, Canada

[9]Department of Computer Science, Faculty of Science, The University of British Columbia, Vancouver, BC, Canada

[10]School of Population and Public Health, Faculty of Medicine, The University of British Columbia, Vancouver, BC, Canada

Corresponding Author:
Beth A Payne, BSc, PhD
Women's Health Research Institute
H214-4500 Oak Street, Box 42
Vancouver, BC, V6H3N1
Canada
Phone: 1 604 875 3459
Email: bpayne@cw.bc.ca

Abstract

Background: Digital health innovations provide an opportunity to improve access to care, information, and quality of care during the perinatal period, a critical period of health for mothers and infants. However, research to develop perinatal digital health solutions needs to be informed by actual patient and health system needs in order to optimize implementation, adoption, and sustainability.

Objective: Our aim was to co-design a research agenda with defined research priorities that reflected health system realities and patient needs.

Methods: Co-design of the research agenda involved a series of activities: (1) review of the provincial Digital Health Strategy and Maternity Services Strategy to identify relevant health system priorities, (2) anonymous survey targeting perinatal care providers to ascertain their current use and perceived need for digital tools, (3) engagement meetings using human-centered design methods with multilingual patients who are currently or recently pregnant to understand their health experiences and needs, and (4) a workshop that brought together patients and other project partners to prioritize identified challenges and opportunities for perinatal digital health in a set of research questions. These questions were grouped into themes using a deductive analysis approach starting with current BC Digital Health Strategy guiding principles.

Results: Between September 15, 2022, and August 31, 2023, we engaged with more than 150 perinatal health care providers, researchers, and health system stakeholders and a patient advisory group of women who were recently pregnant to understand the perceived needs and priorities for digital innovation in perinatal care in British Columbia, Canada. As a combined group, partners were able to define 12 priority research questions in 3 themes. The themes prioritized are digital innovation for (1) patient

XSL•FO
RenderX

autonomy and support, (2) standardized educational resources for patients and providers, and (3) improved access to health information.

Conclusions: Our research agenda highlights the needs for perinatal digital health research to support improvements in the quality of care in British Columbia. By using a human-centered design approach, we were able to co-design research priorities that are meaningful to patients and health system stakeholders. The identified priority research questions are merely a stepping stone in the research process and now need to be actioned by research teams and health systems partners.

(JMIR Hum Factors 2025;12:e60825) doi:10.2196/60825

KEYWORDS

digital health; co-design; digital strategy; human-centered design; eHealth; cocreation; codevelopment; perinatal intervention; quality of care; digital tools; pregnancy; patient autonomy; patient support; mobile phone

Introduction

The digital health sector continues to expand after being an integral part of health care delivery during the COVID-19 pandemic, resulting in a growing interest in digital health research and innovation. Digital health provides an opportunity to reduce health disparities by addressing barriers and facilitators to health care access and improving the health and lifestyle of those who use it [1]. Digital innovation is also essential to offering value-based health care, which involves changes in health systems that are measured by improvement in patient outcomes per unit health care expenditure [2]. Core to achieving improvements in value-based care is establishing an understanding of what value is based on identifying outcomes that matter to patients themselves.

Human-centered design (HCD) is an approach to problem-solving that can be used to increase value in health care by focusing on problems and outcomes of interest to patients. HCD also has a role in improving health equity [3,4]. Core characteristics of an HCD approach include an iterative process of empathy and discovery to identify the core problem and then systematic brainstorming of solutions where the human experiencing the problem plays a central role [5]. The HCD approach goes beyond simply working with the end user to focus on their interaction with the technology and considers the broader social context in which the technology is used and other partners who may indirectly be impacted by the use of the technology [4,5]. The result of working in partnership with people under these assumptions using the HCD approach is meant to be a product or service that meets real human needs and reflects their values. This approach has been used to solve complex health care challenges, such as improving pediatric patient experience in acute care [6] and increasing adherence to self-care practices for chronic conditions [7,8]. In a perinatal health setting, these methods have been used to develop digital interventions for gestational diabetes [9].

The perinatal period encompasses a person's experiences from conception to the first year after delivery including pregnancies that end in miscarriage or stillbirths. The majority of patients receiving perinatal care in British Columbia has access to a mobile device and the internet and are increasingly looking to digital technologies to inform and support their care [10,11]. Using digital technology in health care can improve the quality of care and health outcomes for patients and infants during this early stage of life. The use of digital health tools in perinatal care has already been shown to result in lower and healthier gestational weight, lower smoking rates, increased physical activity, and lower rates of maternal and infant morbidity (eg, fewer preterm births, gestational diabetes, and pre-eclampsia), leading to fewer planned and unplanned visits to health care facilities [1,12-14]. Digital health has also been demonstrated to reduce the practitioners' workload as well as relieve some of the resource strains on an overburdened system [15]. Due to the vast potentiality of digital health and opportunities for further innovation, it is important to engage with patients who have received perinatal care and care providers as end users to identify emerging priorities for future innovations and research that are informed by actual needs [16]. By identifying and responding to real patient and care provider needs, we can optimize the implementation, adoption, and sustainability of perinatal digital health solutions, create value, and ensure existing health inequities are not exacerbated [16,17].

In this paper, we report on the Research Agenda for Perinatal Innovation and Digital Health (RAPID) project [18], which was led by the BC Children's Hospital Research Institute and Women's Health Research Institute in partnership with the Provincial Health Services Authority (PHSA) and Perinatal Services BC (PSBC) with the aim of codeveloping research priorities with patients and care providers.

Methods

Ethical Considerations

This project was run as a knowledge translation initiative, and no research ethics board approval was required. The engagement activities conducted with partners were classified as quality improvement activities rather than systematic investigations to advance knowledge. Based on these distinctions, our institutional research ethics board determined that the project falls outside of the scope of research subject to ethics review, and a waiver was granted. All workshop and survey participants consented to the anonymous collection of their opinions. Members of our patient advisory committee were compensated hourly according to local standards to reflect their equal standing to research team members within the project.

Project Design and Setting

The patient and health system partners with whom we worked were partners in the research priority setting through co-design and were compensated for their time. All final workshop

XSL•FO
RenderX

participants provided consent to the collection of their opinions and views during the workshop and through the follow-up survey.

The objectives of the RAPID project were to facilitate activities that bring key partners and existing knowledge together to (1) identify gaps in the perinatal continuum of care that can be addressed with digital innovation to improve patient and family health and (2) define current barriers to optimal adoption of digital health interventions for perinatal health care from the patient and health system perspectives. This knowledge was then used to define research priorities and codevelop research questions to motivate digital innovation informed by care needs and addressing barriers.

Recruitment of Partners and Governance

Following guidance outlined by Concannon et al [19], we developed an engagement plan for our activities that describes an analysis of which partner groups should be involved, their roles and modes of participation, and sampling strategies (ie, snowball sampling). Partners to include in engagement activities were grouped as (1) people with lived experience of perinatal care who were pregnant or recently delivered a baby; (2) perinatal health care providers including nurses, midwives, physicians, and other allied health; and (3) health care decision makers who are involved with the design or delivery of perinatal health care in British Columbia.

This project was governed by two primary groups: (1) a steering committee and (2) a patient advisory committee (PAC). Using snowball sampling, starting with a core group of experts at BC Children Hospital Research Institute and Women's Health Research Institute, we identified key partners with expertise in research, innovation, and implementation processes related to perinatal, maternal, or digital health services to form the steering committee. The core group was involved from the initiation of the project including during the development of the funding proposal. This core group included 4 health system decision makers as research users and 4 researchers. Additional partners identified for the steering committee included 2 clinical care providers and 2 additional researchers. The final steering committee included 5 researchers, 3 health care providers, and 4 health system decision makers. The purpose of the steering committee was to oversee project objectives and align efforts to each of their organizations' known digital health priorities and ongoing work.

Patient partners for the advisory committee were recruited through social media and the local hospital patient engagement office. The criteria for participation were (1) multilingual and (2) pregnant or a caregiver of a child younger than 2 years of age. There were 15 respondents to the recruitment advertisements, 9 of whom responded to our invitation to complete a 15-minute informal interview so we could learn more about their background, and they could also learn more about our project. After the interviews, the steering committee was able to select 4 participants who were willing and able to commit the time required by the project and who had diverse perinatal experiences (based on type of care provider and mode of delivery), geographic location in British Columbia, and backgrounds (language spoken and profession). The PAC met separately with only 4 members of the study team and steering committee during the initial activities of the project. This was done to create a safe space that would limit the impact of known health system hierarchies and preserve the authentic patient voice in the identification of perinatal health system pain points and the development of initial insights related to these pain points [20]. Once initial insights were defined with the patient advisors, all partners were brought together in a final workshop to collaborate and build consensus on research priorities as described in more detail below.

Engagement Activities

A series of engagement activities with each distinct group of partners was completed before a final workshop that brought all partners and results together to reach a consensus on final research priorities.

Perinatal Care providers

Perinatal care providers across British Columbia including physicians, midwives, nurses, and allied health were invited to share, rate, and explore their thoughts on digital perinatal health using a ThoughtExchange (version 6.4; Fulcrum Management Solutions Ltd). ThoughtExchange is a web-based survey tool that allows participants to share their thoughts in response to a prompt, which can be anonymously viewed and ranked by other participants. The platform then creates weighted rankings of the thoughts on the prompted topic to show participant priorities. The care providers were asked to respond to the following prompt: "What digital health tools are most valuable to your delivery of perinatal health care currently, and what do you think are the greatest areas of opportunity for digital innovation going forward?" The survey was promoted through all British Columbia health authority internal staff newsletters and through member email lists at Doctors of BC; Family Doctors of BC; and University of British Columbia Faculty of Medicine alumni, School of Midwifery, Departments of Obstetrics and Gynaecology, Pediatrics and Family Medicine. Three rounds of promotion were completed. Once the advertised deadline passed, the highest-ranked thoughts were reviewed by 2 research members of the steering committee (BAP and SXC), and key messages were developed through an inductive approach. Basic descriptive statistics were used to analyze the ranking of thoughts.

Health System Partners

Steering committee members from the British Columbia Ministry of Health (MoH), the PHSA, and the PSBC were each asked to review their organizations' recently published reports outlining their strategies for addressing current gaps in health. The reports reviewed were the BC Digital Health Strategy (BCDHS) [21] and the PSBC Maternity Services Strategy (MSS). Access to the MSS was provided through confidential personal communication with PSBC Director Robert Finch. These reports were presented at a steering committee meeting within the first 3 months of the project to identify priorities and alignments specific to perinatal digital health within these documents. The steering committee members from each organization outlined ongoing digital innovation projects in perinatal care and priorities specific to digital health solutions

XSL•FO
RenderX

for improving perinatal care. Each report, postpresentation, was summarized in lay terms to highlight priorities. These summaries were reviewed and edited by the health system partners to ensure that summaries were comprehensive, not leaving out key information.

Patients

A series of four 90-minute PAC meetings were held and recorded over Zoom (Zoom Video Communications). They served to ascertain current digital health use and perceived need for new innovation as well as to inform workshop development from a patient perspective. The meetings included a series of HCD and user-experience research activities, participatory discussions, and postmeeting homework that typically took an hour to complete. The first meeting served as an introduction to all members; the following 2 meetings were developed to allow patients to share their experiences and understand how they interact with the health care system in order to identify gaps, common touchpoints with health care systems and services, and needs (Table 1). For the final meeting, the PAC was provided the lay summaries from the engagement with other partners prior to the meeting. They were asked to review these and share their reflections during the meeting. In addition, they were asked to develop insight statements that identify a need or gap. All meetings were facilitated by the core research team consisting of BAP, SXC, HA, and MV, who were also steering committee members. Activities were carried out on Miro (version 2.0; Miro), a web-based collaborative workspace. Meeting notes were collected and circulated to all participants for review and input. After each meeting, study team members also recorded reflective notes.

Table 1. Patient engagement meeting plan.

Session number	Objective	Presession activity	Activities
1	Establish an environment where patients feel safe and comfortable sharing their experiences	• N/A[a]	• Ice breaker • Introduction to digital health • Empathy mapping activity
2	Gather patients' perspectives and experiences in the health system during pregnancy and after birth	• Complete an empathy map [22] and a 30-minute reflection on their experiences in perinatal care	• Share empathy map and experiences • Introduce user journey activity
3	Identify health system touchpoints and potential areas with opportunities for improvement	• Complete a user journey [23] • Reflect on fellow participants' empathy maps and reflections	• Open discussion regarding user journey and experiences • Introduce insightful statements
4	Prepare for workshops with patients	• Review summary of ThoughtExchange and health systems' report summaries • Insightful statements worksheet [24]	• Comment on summary results • Go over insightful statements • Plan workshop

[a]N/A: not applicable.

Final Consensus Workshop and Analysis

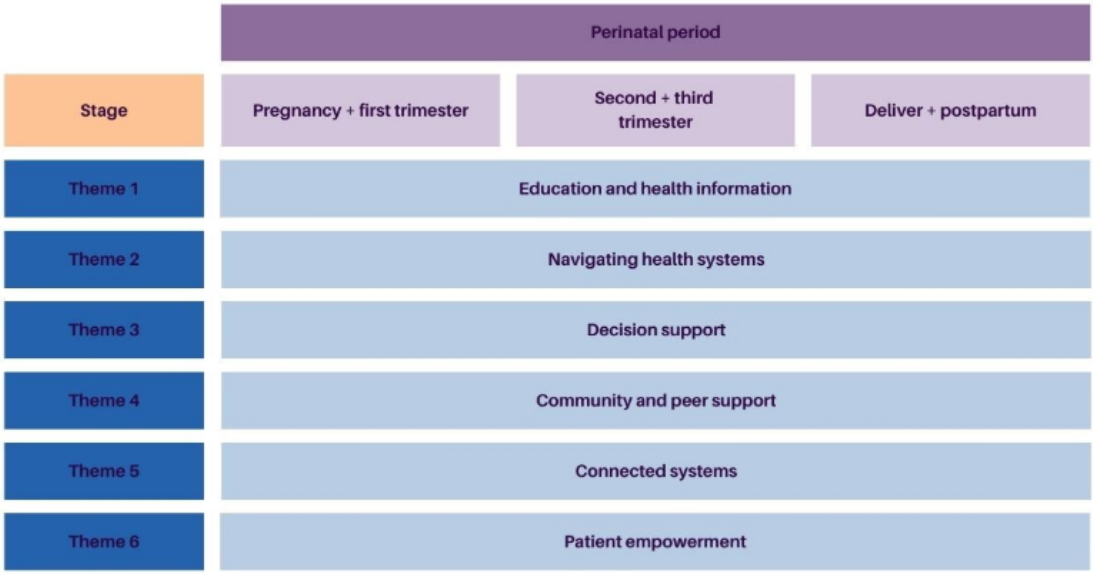
A final 90-minute workshop was held to bring together all the partners and interested parties to share and align on challenges and priorities collected from earlier engagement sessions. This workshop was composed of both existing members of each partner group: patients, clinical care providers, researchers and health system decision makers as well as additional participants recruited through snowball sampling.

In preparation for the workshop, the core study team (BAP, SXC, HA, and MV) reviewed insight statements from the patients, and using a combination of deductive (using the guiding principles of the BCDHS as a starting point) and inductive analysis, placed these into overarching and emerging themes and then further grouped statements based on relevance to a stage of the perinatal period as described in Figure 1. This grouping was done to make the review and brainstorming session more feasible to complete within our limited 90-minute time frame. Research representatives from the steering committee, BAP and MV, developed 2 example research questions for each thematic area to prime the workshop discussion. These questions were developed to prompt discussion during the workshop, considering many participants do not have a background in research.

Attendees were divided into the 3 perinatal stage–based groups based on their expertise. For example, an attendee who is a neonatologist was assigned to the delivery+postpartum group. Each group contained at least 1 member from each of the partner types in order to facilitate collaboration across groups and consensus building. Each group was asked to reflect on and revise the provided insight statements and example research questions as well as develop their own research questions in order to identify research priorities based on their knowledge and experience. At the end of the session, each group was asked to identify 3-5 priority research questions from their brainstorming exercise. Each group then presented their priority research questions to the full workshop, and collectively, all attendees further discussed and ranked the research questions in order of priority.

XSL•FO
RenderX

Figure 1. Workshop groups and themes.



After the workshop, the research team reviewed meeting recordings, notes, and research question rankings to finalize a list of priority research questions that reflected workshop consensus. These research questions were then reviewed a final time by the PAC to ensure patient needs were at the forefront. This final list of research questions was converted and deployed as a REDCap (Research Electronic Data Capture; Vanderbilt University) survey to all workshop attendees. Respondents were asked to select 3 research questions they believed to be of the highest priority for perinatal care. Survey rankings were summarized descriptively to identify priorities from within the defined research questions.

Results

Engagement activities occurred between September 15, 2022, and August 31, 2023. The steering committee met monthly during this time to review progress and plan adjustments to the strategy, as new knowledge was developed.

Thought Exchange With Clinical Care Providers

A total of 125 perinatal health care provider participants responded, 102 of whom shared what provincial health authority they were affiliated with. British Columbia is split by geographic regions into 5 regional health authorities with an additional province-wide authority focused on Indigenous health and the PHSA that provides oversight for all regions. Most participants came from PHSA (n=25, 24.5%), Fraser Health (n=19, 18.6%), Vancouver Coastal Health (n=16, 15.7%), or Island Health (n=16, 15.7%). There were no participants from the First Nations Health Authority, and 7 participants did not affiliate with any of British Columbia's 7 health authorities.

Of the 125 participants, 55 responded with 80 different thoughts; 51 participants rated the thoughts of their colleagues, resulting in a total of 1108 ratings. Analysis of the thoughts based on the built-in functions of the ThoughtExchange platform identified the highest-ranked themes, which were grouped and summarized. Five key takeaways were identified under the themes of health education and connected systems from the ThoughtExchange:

1. There is an urgent need to create more inclusive, responsive, and accessible digital health tools. The end user (patient or health care provider) must be involved in the design of these tools to make sure that they are useful and can solve the right problems.
2. Among all current digital health tools shared, "Zoom," which is a tool that allows videos and remote health visits, and the "Hyperbilirubinemia risk assessment tool," which is a way for doctors to check if a baby is at risk of developing a condition where their skin and eyes turn yellow, are the highest ranked. "Up to Date," a web-based platform that keeps clinicians up-to-date with the latest medical research and guidelines (the United States only), was mentioned but ranked lower. A tool like "Up to Date" that provides the latest guidelines specific to Canada was identified as a current need to support high-quality care.
3. Important clinical areas that were identified as being priority topics for digital innovation were postpartum lactation support and mental health services. There is a demand for innovation in these areas to be accessible to individuals residing in rural regions.
4. A 2-way communication system is needed to support sharing health information across the province and between patients and providers; this means patients should have equal access to their electronic medical records as their health care providers. Patients should also be able to freely share their information between various levels of health care units at their own discretion.
5. Provincially standardized, free, evidence-based, and user-friendly prenatal patient education was identified as a need that could be solved with digital innovation. Accessible prenatal education would support high-quality perinatal care for all patients in British Columbia.

XSL•FO
RenderX

Health System Partner Engagement

A review of the MoH and PSBC digital health reports and strategies with health system partners in the steering committee meetings identified opportunities where perinatal innovation could be used to improve the quality of the health care system in British Columbia.

The MoH report, BCDHS, identified the following gaps in care across British Columbia: lack of coordination between health care services, poor communication systems, and inadequate follow-up. Therefore, their digital health priorities include (1) empowering patients through participation in their health journey (eg, inclusive, trusted, and equitable health content provided digitally), (2) increasing the capacity for providers to deliver high-quality care (eg, improving digital literacy and promoting the use of digital tools), (3) establishing a connected health system that allows for the sharing of data (eg, interoperability across health authorities and provider level and collaborating on clinical solutions), and (4) infrastructure and business processes streamlined for efficiency (eg, modernizing British Columbia's health supply chain, data flow, asset tracking, and analytics). To accommodate these strategies and goals, PSBC is working on developing a digital platform that acts as a central place for patient information. Based on the identified gaps, strategies, and needed tools, three potential areas for research were identified: (1) What is the best way to use telecommunication and information technologies to provide health care services during pregnancy and childbirth remotely? (2) What is the best design and implementation strategy to enable a shared electronic antenatal health record that can be accessed by any health care provider and patient during pregnancy? (3) Can we validate the use of wearable devices that can track and monitor health data during pregnancy and childbirth (ie, remote blood pressure for hypertension in pregnancy and blood glucose monitoring for gestational diabetes)?

PSBC's MSS report focuses particularly on addressing modifiable factors such as maternal weight and nutrition, substance use, mental health, diabetes, infant sleep, and breastfeeding. They currently have a series of projects in progress, and those specific to digital health solutions include a personalizable information and tool resource hub for providers and patients, a single source of data that can be shared across multiple systems, leveraging access to clinical data for quality improvement, allowing for enhanced digital training opportunities, increasing access to telemedicine care, and updated screening protocols and quality control. Based on these approaches five areas for future research were identified: (1) With the implementation of new technologies, what gaps and biases arise that impact their inclusivity and equity? (2) Although we can assess the immediate and short-term benefits of these technologies, how do we assess their long-term impact on maternal and child health? (3) How do we improve interoperability between different digital platforms to enable seamless data sharing and communication? (4) What policies would need to be put in place to ensure that patient data and privacy are protected? How would this integrate patient consent? (5) How do we understand the needs of the user (patients and health care service providers) and design resources that are user-centered? How do we ensure that users will engage with these resources?

Patient Engagement Sessions

The 4 members of the PAC varied in languages spoken (1 Mandarin, 1 Cantonese, 1 Spanish, and 1 Arabic), age of their children (1 still pregnant at the start of the project, 1 younger than 6 months, and 2 between 6 months and 1 year), mode of delivery (2 vaginal birth and 2 cesarean section), and location of their birth experience (2 Vancouver Coastal Health, 1 Fraser Health, and 1 Interior Health). Over the course of our 4 group meetings, we completed empathy maps, user journeys (Multimedia Appendix 1), and insight statements and collected reflective notes and recordings of live discussions. Common challenges were experienced among all patient partners navigating the perinatal period. For example, all patient partners described difficulties in understanding the different types of perinatal care providers and how to find an available midwife or doctor in their region. They all also spoke of challenges in navigating the transition from pregnancy to early newborn care, particularly when they had nonurgent concerns about their new baby because many available health care resources focus on urgent and emergency conditions.

In the final PAC meeting, we reflected on the summaries of the ThoughtExchange, MSS report, and BCDHS to identify recurring themes, priorities, and gaps that were also expressed by patients, which were then presented in a master list of insight statements (Textbox 1). These insights were divided into 6 themes: education or health information, navigating health systems, decision support, community or peer support, connected systems, and patient empowerment. Overall, over the 4 meetings, patients clearly expressed that they wanted digital innovation that would support their confidence and ability to navigate the health system in all regions of the province of British Columbia where health care is currently distributed and often fragmented. This should come in the form of free and accessible educational resources and tools to navigate the perinatal health journey. All patient partners discussed how stressful it is to navigate the current system and the frustration of having to communicate their health information to different care providers repeatedly. Almost all of our patient partners moved during the course of their pregnancy and experienced challenges finding maternity care providers both initially and after they moved. Another universal experience was the challenge presented by our current system, which does not allow for the pairing of the mother's and infant's health records. This often resulted in logistical challenges in navigating postpartum and postnatal follow-up that were borne entirely by the new parents. As a result of the success of these sessions, the research team worked with the PAC to develop a toolkit for patient engagement that was made available for use by other research teams (Multimedia Appendix 2).

XSL•FO
RenderX

Textbox 1. Final list of insight statements.

Reliable health information is needed

- Patients who are accessing perinatal care are overwhelmed by navigating an abundance of resources, and they are not sure which ones are accurate to follow.
- When seeking health-related information on the internet, information is often conflicting or misleading.
- When seeking health information on the internet, information is only in English and can be difficult to understand or communicate to multilingual family members.
- Health information provided is in technical terms that patients are unfamiliar with leaving patients confused.
- Patients often rely on informal peer support, which adds to the conflicting information they receive, leaving them more confused and unsure of what information to trust.
- Due to the abundance of information available on the internet, patients feel like they are wasting or spending too much of their time filtering through the information.
- Patients indicated that it is challenging to understand the difference between an obstetrician and a midwife especially for newcomers or immigrants from a different health care system; they do not know how to choose the right professional for them.

Lack of supports

- Due to the duration between pregnancy and a patient's first obstetrician gynecologist appointment being too long, patients are left with the responsibility of seeking information independently.
- There are limited perinatal classes that are poorly advertised and are not free; therefore, not all patients have the means to access them.
- Patients are navigating aspects of the health care system during their perinatal journey that they have never had to use before; this often comes with little guidance and doubts about what steps need to be taken.
- Patients who are pregnant for the first time do not know where and what services to seek, leaving them feeling helpless and guilty if complications occur.

Need for continuum of care

- Patients' care is interrupted by the transfer of care to a different care provider especially after delivery. These periods of interruption leave patients uncertain and lost.
- After being discharged after delivery or from midwives, patients feel like their only support is the emergency room, especially if they do not have a family doctor.

Mental health challenges

- Complications during the perinatal period leave patients feeling incompetent and highly anxious about their own situations.
- Patients wanting to seek mental health support are discouraged because they are unaware of how to access them or told they are not in need of them.

Emergency screening

- Due to the long wait times in the emergency room and lack of knowledge of problems arising either during prepartum or postpartum, patients indicated a need to have an emergency screening process tailored to pregnant and newborn mothers by using the digital platform before they head to the hospitals to save or rule out unnecessary visits.

Uncoordinated transitions of care

- It was challenging for patients to find an appropriate obstetrician or midwife in a timely manner; they had to call a list of clinics on their own or find one through word of mouth in order to be accepted.
- When patients find out they are pregnant, they are frustrated by a lack of information on available care providers in their region.
- Patients are confused about when to choose a midwife or other type of doctor and why they may need to switch providers during pregnancy.
- Mom and baby's care is not coordinated in the health system after the birth of the baby because they are treated as 2 distinct patients.
- There is no standardized system to share health records between perinatal care providers, which leaves patients with the responsibility to communicate their health information when multiple types of caregivers are needed to support their pregnancy.
- When accessing care, services, or resources across different health authorities, health authorities do not follow standardized structures or guidelines acting as an additional barrier for patients as they learn to navigate a new health authority.
- There is a lack of communication and awareness among health authorities, therefore, interrupting patients' care.

Poor patient-centered care

XSL•FO
RenderX

- Care providers are dismissive of patient concerns and thoughts, resulting in patients advocating for themselves and leaving them stressed and isolated.
- Patients feel uninvolved in their health because they are often told what to do without direct and clear explanations or guidelines from their providers.
- Patients have to manage multiple appointments with different providers instead of an integrated care team, meaning that they often wait long periods of time before their needs are addressed.
- Patients would appreciate more visits and opportunities to meet with their care providers or a trusted source in order to discuss their concerns and ask questions.
- Patients want to hear follow-up from their care provider after every test, even if it is normal.

Final Workshop

The workshop had a total of 20 attendees: 4 core project team members, 4 patient partners, 2 researchers, 5 clinicians, and 5 health system partners. The workshop was held over Zoom. During the workshop, 3 research themes were chosen as priorities through the discussion process including patient autonomy and support, educational resources for patients and providers, and access to health information (Figure 2). Attendees chose to frame the final research questions around broad transformative goals rather than specific clinical care concepts or processes to make the final agenda broadly applicable across health settings in the British Columbia public health system. For example, although patients and clinicians identified a need for lactation support earlier in the engagement process, a decision was made not to name this clinical concern specifically, as it was felt to narrow the scope of the recommendations. Similarly, patients and care providers identified a need for more accessible support for nonurgent health concerns. Suggestions were made to create more digital resources, but in discussion, the group agreed that more foundational knowledge on the type of digital technology that would be found most usable and effective for providing health information is needed. This resulted in final research questions such as 1.3 and 2.3. A total of 12 final research questions to be included in the research agenda were developed to reflect the identified priorities and patient perspectives (Figure 3).

Figure 2. Three priority themes for perinatal digital health research.



XSL•FO
RenderX

Figure 3. Final list of priority research questions. RPM: remote patient monitoring.



A follow-up REDCap survey was sent to a total of 35 individuals, including all workshop participants, steering committee members, and patient partners, and we received 20 responses. Respondents included 5 researchers, 5 health system decision makers, and 6 perinatal health care providers with 4 patient advisors. When asked to select the top 3 priority research questions, respondents prioritized questions 1.4, 1.2, and 3.3 (Textbox 2). All of these questions focus on supporting patients' decision-making with better information and fall within the themes of improving patient autonomy and access to information. When comparing the priority of the overarching themes, overall questions from theme 1: patient autonomy and support were the most selected; however, if we adjust for the number of questions per category, questions from theme 3: access to health information were the most selected.

XSL•FO
RenderX

Textbox 2. Prioritization of research questions.

Number of times prioritized: 9

- Question 1.4: What aspects of perinatal care can be monitored, administered, or assessed using mobile devices (ie, remote patient monitoring, digital health visits, and peer support) to support empowered self-care? How do we implement these without widening the access to care gap experienced by already disadvantaged populations?

Number of times prioritized: 7

- Question 1.2: What digital tools and information can support patients' decisions in identifying and connecting with the right maternity care provider? What information do patients need to identify the right maternity care provider? How do we support this decision with digital tools?
- Question 3.3: How can the integration of patient-reported outcomes in a patient's record contribute to quality improvement and shared decision-making?

Number of times prioritized: 5

- Question 1.5: How do we monitor the mental health patients receiving perinatal care in order to provide timely support when needed and improve health outcomes and perinatal recovery?
- Question 2.1: What resources can be compiled into a starter kit to guide patients through their pregnancy? What resources can be compiled into a starter kit to guide patients who have just delivered their child? How do these resources impact patient and child health outcomes?
- Question 2.3: How do patients prefer to view and access health information and educational resources (eg, type of format such as audio vs video vs website)? How can we engage patients and providers in developing and designing these resources? Are there existing resources we can expand in function to meet patient needs (eg, 811 response, peer support, chatbots, or SmartMom)?
- Question 3.1: What is the effect of more transparent access to provincial dashboards such as HealthGateway (for both parent and child and linked together) on care quality, and access to care in both urban and rural or remote settings?

Number of times prioritized: 4

- Question 3.2: Can improved data access for patients and care providers lead to an increase in British Columbia's status against global benchmarks for health care quality?

Number of times prioritized: 3

- Question 2.2: What are the individual, organizational, and policy barriers and facilitators to providing standardized multilingual prenatal and postnatal education in British Columbia?

Number of times prioritized: 2

- Question 1.1: How do we best measure patient autonomy in the perinatal period?
- Question 1.3: What services and resources exist within British Columbia that support patient mental health, self-care, and patient autonomy in decision-making? How do we ensure providers are aware of them? How do we emphasize connecting patients to them?

Number of times prioritized: 1

- Question 2.4 What implementation strategies will work to make educational resources accessible across regions in British Columbia? How do we measure impact?

Discussion

Research Priorities and Implications for Practice

The RAPID project resulted in the co-design of a comprehensive research agenda for perinatal digital innovation that reflects provincial strategies and patient experience in British Columbia. During the project, a series of engagement activities allowed us to ascertain the current gaps and needs for digital health interventions for perinatal care and develop a research agenda that will address these needs. Our research agenda provides a foundation for identifying, developing, and implementing digital solutions that are responsive and feasible within the current British Columbia health care system landscape. The RAPID research agenda identified a core set of research priorities under 3 interdependent thematic areas that, if addressed, could lead to meaningful improvements in patient and clinician experience and health outcomes in the perinatal period. The priority themes identified are patient autonomy and support when navigating the health system, patient and provider access to education resources, and patients' ability to access health information. These themes align with the BCDHS and have broad support from health system leaders.

The research themes identified and research questions prioritized in this project suggest that further innovation is needed to properly leverage digital health tools to support perinatal care in British Columbia. Priority research questions and a major area of discussion with our patient partners focused on a need for better digital tools to support self-care and decision-making across the entire perinatal period, from pregnancy to the end of the first year of the infant's life. The ThoughtExchange also highlighted that, within the British Columbian health system, there is a limited number of tools in use that clinicians view as

XSL•FO
RenderX

valuable to care. This results in patients accessing numerous external digital resources that are not always of quality and relevance to their care. This agenda can be used to advocate for change at the provincial level and support the incorporation of digital health into the health system. Addressing research priorities identified by RAPID would improve the patient experience by providing needed solutions that allow patients to safely interact with their health care professionals through web-based platforms or through their mobile devices, access self-care tools and applications, use smartphones or other devices to manage and monitor their health, and adopt technology in their health plans.

Digital health innovation as suggested by our research agenda is appealing because it is not restricted by time or place [13] allowing for patient empowerment as well as patient participation in the decision-making of their own health. Addressing the RAPID research priorities can also improve the continuity of care and delivery of essential services, something that all partners in the RAPID project would clearly value. There is already evidence that digital innovation in perinatal care can lead to improved health outcomes for both the woman or pregnant person and their infant [1,12-14]. This agenda highlights that the perceived value of digital innovations for patients receiving perinatal care and their families cannot only be determined by traditional health outcomes. Digital innovations should also be measured against their ability to improve confidence and autonomy in decision-making, self-care, information sharing, and patient and clinical experience.

The digital health sector receives enormous investment, making available a vast number of solutions varying in quality and capabilities [17]. The existing literature has provided a broad understanding of the feasibility, acceptability, and effectiveness of digital tools outlining advantages and disadvantages [1,12]. However, this continued growth in digital solutions does not guarantee or equate to an increase in solutions that can be implemented to meet the actual existing needs of patients, providers, and health systems [17]. In this project, we used HCD methods to identify patient needs and priorities for digital innovation that are aligned with clinical care providers and health system priorities, but our identified priorities cannot exist in isolation from the current knowledge of socioecological factors. Although the RAPID priority research questions are defined based on need, these must be explored with consideration for the facilitators and barriers that contribute to engagement with digital health solutions. Previous research on digital health use has identified barriers and facilitators to implementation such as infrastructure, equipment, network access, motivation, and usability [25,26].

The effort was made to integrate the knowledge of barriers and facilitators to engagement and uptake of digital health innovations throughout the research agenda, such as geographic settings, data-security concerns, and diversity in health concerns. However, considerations for these barriers and facilitators should continue beyond what the questions have defined, as we work to action the agenda. Issues such as connectivity, user motivation, social support, digital literacy, user-friendly design, and credibility should continue to be examined throughout subsequent digital health research.

Strengths and Limitations

This project's use of a human-centered approach prioritized patient voices and provided the space for collaborative co-design of the final research questions. Limiting the number of patient partners and restricting the number of researchers in the patient advisory meetings allowed for better quality engagement through intimate and extensive discussions. The diversity of partners from across British Columbia allows for more generalizable results. However, a significant limitation was our lack of engagement with Indigenous clinical care providers or patients. There is a need to engage with Indigenous people's voices within all of the partner groups before moving research forward from this strategy. Related to this, although we were able to collect a diverse group of partners' opinions and bring these together to design a comprehensive research agenda, running the project under a knowledge translation initiative restricted the amount of data we could collect on engaged participants. As a result, it may be more difficult for others to apply this knowledge outside of the British Columbia health system. We also cannot explore if participant characteristics such as age or experience level (for health care providers) had any impact on their opinions of digital health priorities. An additional strength is our engagement with health system partners throughout the project, aligning in strategy and garnering early support from decision makers. As a result, outcomes of research informed by this agenda are more likely to be implemented and achieve long-term impact. There was strong collaboration throughout the project between researchers and knowledge users, allowing us to lead this work within the health system, leveraging developed partnerships, and articulating alignment with organizational goals and patient research priorities.

This HCD approach prioritized patient voices in developing research priorities; yet, our multipartner codevelopment approach recognizes that the ownness should not be on patients alone to identify a solution to the complex issues of perinatal health care delivery. The emphasis on people solving their own problems that is core to the HCD approach cannot be directly assumed in a health care context due to the complexity of the system and the multitude of interested parties that will be impacted by any system changes. In positioning our work to address patient needs while still aligning with health system priorities, we have aimed to support value-based care efforts in the British Columbia health system. Focusing further research and innovation efforts to address the RAPID research priorities can support improvements in value-based care, as they would improve outcomes and processes that matter to patients without an unreasonable increase in health care expenditure.

Conclusions

This project identified clear support and enthusiasm across partners for digital innovation to improve the quality of perinatal care in the province of British Columbia. Our research agenda highlights the need for better patient education and health information systems that are more accessible to patients and care providers with the overarching need for digital resources that can increase patient autonomy in decision-making throughout pregnancy and the transition to newborn care. Co-design of this research agenda is a first step in the research

XSL•FO
RenderX

process. Our priority now is to take action on identified research priorities. Action can occur through traditional research funding pathways but also by advocating to our health system partners to directly address needs identified by the RAPID research agenda within the health system.

Acknowledgments

The authors gratefully acknowledge that our project could not have been possible without the invaluable contributions of our patient partners AE, KJH, JMG, and LY. The authors thank Joanne Hepplewhite from the Provincial Health Services Authority Transformation development office for her support with the ThoughtExchange. This project was funded through a convening and collaborating grant (#C2-2022-2969) from the Michael Smith Health Research BC.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Anonymized example of a user journey.
[PNG File , 2366 KB - humanfactors_v12i1e60825_app1.png]

Multimedia Appendix 2
Toolkit for patient engagement.
[PDF File (Adobe PDF File), 1612 KB - humanfactors_v12i1e60825_app2.pdf]

References

1. Lim S, Tan A, Madden S, Hill B. Health professionals' and postpartum women's perspectives on digital health interventions for lifestyle management in the postpartum period: a systematic review of qualitative studies. Front Endocrinol (Lausanne) 2019;10:767 [FREE Full text] [doi: 10.3389/fendo.2019.00767] [Medline: 31787932]
2. Teisberg E, Wallace S, O'Hara S. Defining and implementing value-based health care: a strategic framework. Acad Med 2020;95(5):682-685 [FREE Full text] [doi: 10.1097/ACM.0000000000003122] [Medline: 31833857]
3. Holeman I, Kane D. Human-centered design for global health equity. Inf Technol Dev 2019;26(3):477-505 [FREE Full text] [doi: 10.1080/02681102.2019.1667289] [Medline: 32982007]
4. Evans L, Evans J, Pagliari C, Källander K. Scoping review: exploring the equity impact of current digital health design practices. Oxf Open Digit Health 2023;1:oqad006. [doi: 10.1093/oodh/oqad006]
5. Melles M, Albayrak A, Goossens R. Innovating health care: key characteristics of human-centered design. Int J Qual Health Care 2021;33(Suppl 1):37-44 [FREE Full text] [doi: 10.1093/intqhc/mzaa127] [Medline: 33068104]
6. Simons R. Improving the patient experience for children at the paediatric acute medical unit. Delft University of Technology. 2018. URL: https://resolver.tudelft.nl/uuid:7a975ce4-1529-4065-bcdb-b323cf0747a0 [accessed 2018-04-19]
7. Matheson G, Pacione C, Shultz RK, Klügl M. Leveraging human-centered design in chronic disease prevention. Am J Prev Med 2015;48(4):472-179. [doi: 10.1016/j.amepre.2014.10.014] [Medline: 25700655]
8. Wildevuur S, Thomese F, Ferguson J, Klink A. Information and communication technologies to support chronic disease self-management: preconditions for enhancing the partnership in person-centered care. J Particip Med 2017;9(1):e14. [doi: 10.2196/jopm.8846] [Medline: 36262005]
9. Yee LM, Leziak K, Jackson J, Strohbach A, Saber R, Niznik CM, et al. Patient and provider perspectives on a novel mobile health intervention for low-income pregnant women with gestational or type 2 diabetes mellitus. J Diabetes Sci Technol 2021;15(5):1121-1133 [FREE Full text] [doi: 10.1177/1932296820937347] [Medline: 32627582]
10. So long landline, hello smartphone. Statistics Canada. 2023. URL: https://www.statcan.gc.ca/o1/en/plus/3582-so-long-landline-hello-smartphone [accessed 2023-04-09]
11. Peragallo Urrutia R, Berger AA, Ivins AA, Beckham AJ, Thorp JM, Nicholson WK. Internet use and access among pregnant women via computer and mobile phone: implications for delivery of perinatal care. JMIR Mhealth Uhealth 2015;3(1):e25 [FREE Full text] [doi: 10.2196/mhealth.3347] [Medline: 25835744]
12. van den Heuvel JF, Groenhof TK, Veerbeek JH, van Solinge WW, Lely AT, Franx A, et al. eHealth as the next-generation perinatal care: an overview of the literature. J Med Internet Res 2018;20(6):e202 [FREE Full text] [doi: 10.2196/jmir.9262] [Medline: 29871855]
13. Hussain-Shamsy N, Shah A, Vigod SN, Zaheer J, Seto E. Mobile health for perinatal depression and anxiety: scoping review. J Med Internet Res 2020;22(4):e17011 [FREE Full text] [doi: 10.2196/17011] [Medline: 32281939]
14. Hearn L, Miller M, Fletcher A. Online healthy lifestyle support in the perinatal period: what do women want and do they use it? Aust J Prim Health 2013;19(4):313-318. [doi: 10.1071/PY13039] [Medline: 23899373]

XSL•FO
RenderX

15. Pham Q, El-Dassouki N, Lohani R, Jebanesan A, Young K. The future of virtual care for older ethnic adults beyond the COVID-19 pandemic. J Med Internet Res 2022;24(1):e29876 [FREE Full text] [doi: 10.2196/29876] [Medline: 34994707]
16. Figueroa CA, Luo T, Aguilera A, Lyles CR. The need for feminist intersectionality in digital health. Lancet Digit Health 2021;3(8):e526-e533 [FREE Full text] [doi: 10.1016/S2589-7500(21)00118-7] [Medline: 34325855]
17. Mathews SC, McShea MJ, Hanley CL, Ravitz A, Labrique AB, Cohen AB. Digital health: a path to validation. NPJ Digit Med 2019;2:38 [FREE Full text] [doi: 10.1038/s41746-019-0111-3] [Medline: 31304384]
18. RAPID: Research Agenda for Perinatal Innovation & Digital Health. Women's Health Research Institute. URL: https://whri.org/our-initiatives/rapid/ [accessed 2023-05-23]
19. Concannon TW, Grant S, Welch V, Petkovic J, Selby J, Crowe S, et al. Practical guidance for involving stakeholders in health research. J Gen Intern Med 2019;34(3):458-463 [FREE Full text] [doi: 10.1007/s11606-018-4738-6] [Medline: 30565151]
20. Trauma-informed practice guide. BC Provincial Mental Health and Substance Use Planning Council. 2013. URL: https://cewh.ca/wp-content/uploads/2012/05/2013_TIP-Guide.pdf [accessed 2024-11-12]
21. B.C.'s Digital Health Strategy. 2023. URL: https://divisionsbc.ca/sites/default/files/72475/MoH-Digital-Health-Strategy-Handout-Feb-2023.pdf [accessed 2024-12-12]
22. Build your creative confidence: empathy maps. IDEO. URL: https://www.ideo.com/journal/build-your-creative-confidence-empathy-maps [accessed 2024-04-22]
23. Build your creative confidence: customer journey map. IDEO. URL: https://www.ideo.com/journal/build-your-creative-confidence-customer-journey-map [accessed 2024-04-22]
24. The field guide to human-centered design. IDEO. 2015. URL: https://www.designkit.org/resources/1.html [accessed 2024-12-12]
25. Kaboré SS, Ngangue P, Soubeiga D, Barro A, Pilabré AH, Bationo N, et al. Barriers and facilitators for the sustainability of digital health interventions in low and middle-income countries: a systematic review. Front Digit Health 2022;4:1014375 [FREE Full text] [doi: 10.3389/fdgth.2022.1014375] [Medline: 36518563]
26. Moulaei K, Moulaei R, Bahaadinbeigy K. Barriers and facilitators of using health information technologies by women: a scoping review. BMC Medical Informatics and Decision Making 2023;23(1):176 [FREE Full text] [doi: 10.1186/s12911-023-02280-7] [Medline: 37670281]

Abbreviations

BCDHS: BC Digital Health Strategy
HCD: human-centered design
MoH: Ministry of Health
MSS: Maternal Services Strategy
PAC: patient advisory committee
PHSA: Provincial Health Services Authority
PSBC: Perinatal Services BC
RAPID: Research Agenda for Perinatal Innovation and Digital Health
REDCap: Research Electronic Data Capture

Edited by C Jacob; submitted 23.05.24; peer-reviewed by H Witteman, MG Onaiwu, E Kjörk; comments to author 23.08.24; revised version received 15.10.24; accepted 26.11.24; published 30.01.25.

Please cite as:
Amhaz H, Chen SX, Elchehimi A, Han KJ, Gil JM, Yao L, Vidler M, Berry-Einarson K, Dewar K, Tuason M, Prestley N, Doan Q, van Rooij T, Costa T, Ogilvie G, Payne BA
The Research Agenda for Perinatal Innovation and Digital Health Project: Human-Centered Approach to Multipartner Research Agenda Codevelopment
JMIR Hum Factors 2025;12:e60825
URL: https://humanfactors.jmir.org/2025/1/e60825
doi:10.2196/60825
PMID:

©Haneen Amhaz, Sally Xuanping Chen, Amanee Elchehimi, Kylin Jialin Han, Jade Morales Gil, Lu Yao, Marianne Vidler, Kathryn Berry-Einarson, Kathryn Dewar, May Tuason, Nicole Prestley, Quynh Doan, Tibor van Rooij, Tina Costa, Gina Ogilvie, Beth A Payne. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 30.01.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/),

XSL•FO
RenderX

which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Building a Decentralized Biobanking App for Research Transparency and Patient Engagement: Participatory Design Study

Ananya Dewan[1], BS, BA; M Eifler[2], MFA; Amelia Hood[3], MA; William Sanchez[2], BS; Marielle Gross[2,3], MBE, MD

[1]Johns Hopkins University School of Medicine, Baltimore, MD, United States

[2]de-bi, co., Baltimore, MD, United States

[3]Johns Hopkins Berman Institute of Bioethics, Baltimore, MD, United States

Corresponding Author:
Marielle Gross, MBE, MD
Johns Hopkins Berman Institute of Bioethics
1809 Ashland Avenue
Baltimore, MD, 21205
United States
Phone: 1 8135416103
Email: mariellesophiagross@gmail.com

Abstract

Background: Patient-derived biospecimens are invaluable tools in biomedical research. Currently, there are no mechanisms for patients to follow along and learn about the uses of their donated samples. Incorporating patients as stakeholders and meaningfully engaging them in biomedical research first requires transparency of research activities.

Objective: In this paper, we describe the use of participatory design methods to build a decentralized biobanking "de-bi" mobile app where patients could learn about biobanking, track their specimens, and engage with ongoing research via patient-friendly interfaces overlaying institutional biobank databases, initially developed for a breast cancer use case.

Methods: This research occurred in 2 phases. In phase 1, we designed app screens from which patients could learn about ongoing research involving their samples. We embedded these screens in a survey (n=94) to gauge patients' interests regarding types of feedback and engagement opportunities; survey responses were probed during 6 comprehensive follow-up interviews. We then held an immersive participatory design workshop where participants (approximately 50) provided general feedback about our approach, with an embedded codesign workshop where a subset (n=15) provided targeted feedback on screen designs. For phase 2, we refined user interfaces and developed a functional app prototype in consultation with institutional stakeholders to ensure regulatory compliance, workflow compatibility, and composability with local data architectures. We presented the app at a second workshop, where participants (n=25, across 9 groups) shared thoughts on the app's usability and design. In this phase, we conducted cognitive walkthroughs (n=13) to gain in-depth feedback on in-app task navigation.

Results: Most of the survey participants (61/81, 75%) were interested in learning the outcomes of research on their specimens, and 49% (41/83) were interested in connecting with others with the same diagnosis. Participants (47/60, 78%) expressed strong interest in receiving patient-friendly summaries of scientific information from scientists using their biospecimens. The first design workshop identified confusion in terminology and data presentation (eg, 9/15, 60% of co-designers were unclear on the biospecimens "in use"), though many appreciated the ability to view their personal biospecimens (7/15, 47%), and most were excited about connecting with others (12/15, 80%). In the second workshop, all groups found the app's information valuable. Moreover, 44% (5/9) noted they did not like the onboarding process, which was echoed in cognitive walkthroughs. Walkthroughs further confirmed interest in biospecimen tracking, and 23% (3/13) had confusion about not finding any of their biospecimens in the app. These findings guided refinements in onboarding, design, and user experience.

Conclusions: Designing a patient-facing app that displays information about biobanked specimens can facilitate greater transparency and engagement in biomedical research. Co-designing the app with patient stakeholders confirmed interest in learning about biospecimens and related research, improved presentation of data, and ensured usability of the app in preparation for a pilot study.

(JMIR Hum Factors 2025;12:e59485) doi:10.2196/59485

XSL•FO
RenderX

KEYWORDS

mobile health; mHealth; application; smartphone; digital health; digital intervention; participatory design; biobanking; research transparency; donation; patient-derived biospecimens; plain language communications; patient education

Introduction

Background

Samples of blood, tissue, urine, and other biological products, collected from patients during a medical procedure, are invaluable tools in biomedical research. Reviews estimate that >600 research biobanks exist in the United States, housing tens of millions of samples from tens of millions of patients [1,2]. Biospecimens are used in a vast array of foundational, translational, and clinical research [3-6]. Research on biospecimens shows exceptional promise in studying cancer and developing new cancer treatments [4,7,8]. Cutting-edge specimen preservation methods offer high-fidelity disease models for genomic analysis, high-throughput drug testing, advanced imaging, and more.

Typically, leftover surgical samples are collected after a 1-time informed consent process per the federal Common Rule in the United States, where patients receive information about the proposed research [9]. The terms of this consent include no expectation to follow up with pertinent research results after the form is signed and samples are collected [10]. In the transition from clinical by-product to banked research sample, patient identifiers are removed from biospecimen data [11]. Deidentification is meant to protect patient privacy by removing identifiable information; however, its conditions preclude scientists from sharing follow-up information about if, when, and how they use biospecimens for research [12]. Scientists receive deidentified biospecimens from the biobank and have no mechanism to directly communicate personally relevant insights or engage with the people whose samples they have received (Figure 1) [10]. Thus, while clinical results from pathology reports are communicated to patients, information on how their biospecimens are used and associated research findings are never shared.

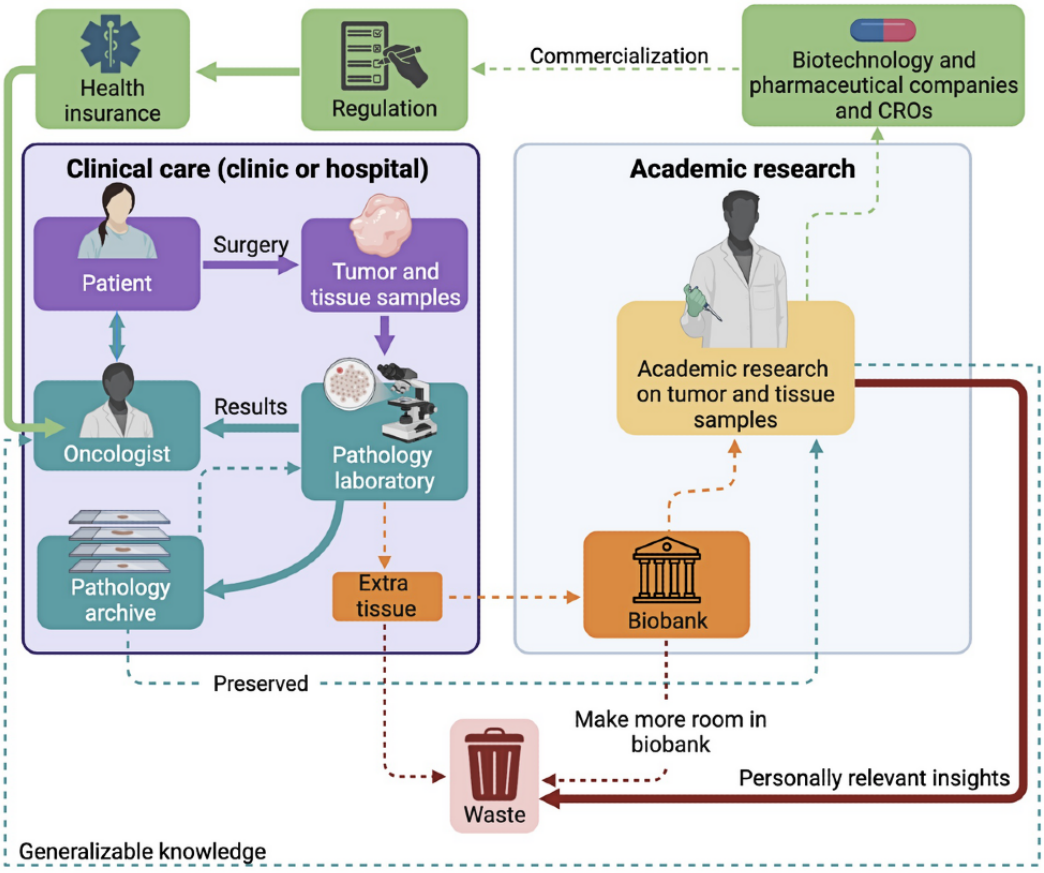
Incorporating patients as stakeholders and meaningfully engaging them in biomedical research is increasingly recognized as critical to more ethical and effective scientific advancement [13]. Patients' lived experiences can help guide research questions, priorities, and methodologies, ensuring that the resulting knowledge aligns with their and their community's priorities. Various innovative technologies and techniques have enhanced patient engagement in research [14,15]; for instance, the James Lind Alliance method facilitates collaboration among clinicians, the public, and additional stakeholders, enabling their opinions to converge and inform the establishment of research priorities [16,17]. Digital platforms that assemble citizens from different backgrounds to share experiences and actively codevelop innovative health strategies have been proposed [18]. However, improving transparency surrounding the nature and extent of research activities is a prerequisite to achieving true patient engagement.

Decentralizing biobank information and empowering biospecimen donors with knowledge about how their samples are used lays the foundation for fostering trust and transparency in biospecimen research [19]. This is particularly vital in rebuilding trust with communities considered marginalized, which have historically been subject to exploitation and unethical research practices [20]. Building trust lays the foundation for further participation by diverse communities and ensures that research endeavors are accountable to those whose biospecimens make the research possible [21].

In previous work, we have demonstrated that there is unmet patient demand for increased transparency surrounding biospecimen research use [22]. We surveyed patients with breast cancer (n=109) about their interest in learning about biobank research and their individual biospecimens [22]. This survey measured patient interest in a broad range of information that they could theoretically learn about research on their samples. Notably, survey respondents were interested in learning information that is readily available within in local biobank and research databases, such as whether their samples had been used in research (66%), details about research conducted on their samples (53%), and the ability to see images of their biospecimens if available (68%).

XSL•FO
RenderX

Figure 1. Biobanking ecosystem: biospecimen supply chain from clinical by-products to institutionalized assets. In the current biobanking paradigm, personally relevant insights and extra tissue are often wasted, representing lost opportunity to advance patient care and maximize value from donated biospecimens. CRO: contract research organization.



Objectives

To address the unmet need for greater transparency and engagement between biobanks and patients, we sought to build an initial decentralized biobanking app—"de-bi"—to serve as a patient's "window" into an established biobank database. We aimed to build a composable platform that would be accessible and valuable to patients while being compliant with existing regulations, composable with data architectures and compatible with established workflows [23]. By viewing biobank data, patients can be empowered to "track" what happens to their biospecimens [24]. A patient-facing mobile app could overlay and display existing data, serving as a technological tool for increasing the transparency of research activities and as a mechanism for directly engaging with patients who have contributed to research, all within the current frameworks of consent and deidentification. Composability with institutional databases would also serve as a source of efficiency, easing the burden of the initial manual data transfer to an app interface and setting up easier future automation and scalability. This tool could then be piloted to measure its impact on patient participation, understanding, and attitudes about biobanking and biospecimen research and begin to answer outstanding empirical questions about effectively fulfilling obligations of transparency and engagement.

Including patients in the design phase was imperative to the future pilot's success. Before the tool was deployed, we sought to ensure that it was interesting, accessible, understandable, valuable, and easy and ideally pleasant to use. Here, we describe the process and results of using participatory design methods to define, design, and deliver a mobile app that enables patients to learn about their biospecimens collected for research, the status of the biospecimens' use in research protocols, and overall biobank collections, as demonstrated in a breast cancer biobank setting.

Methods

Overview

This multiphase participatory design study was conducted among patients of a breast cancer clinic at a large US academic institution. It consisted of surveys of patients (n=94), six semistructured in-depth follow-up interviews with five survey respondents (5/94, 5%), immersive design workshops with patients (workshop 1: approximately 50 participants, focusing on a subset of 15 codesign participants), workshop 2: n=25), and cognitive walkthroughs with individual patients (n=13) [25]. The data being presented represent a subset of broader participant engagement, including additional workshops and encounters with the local community of patients with breast cancer.

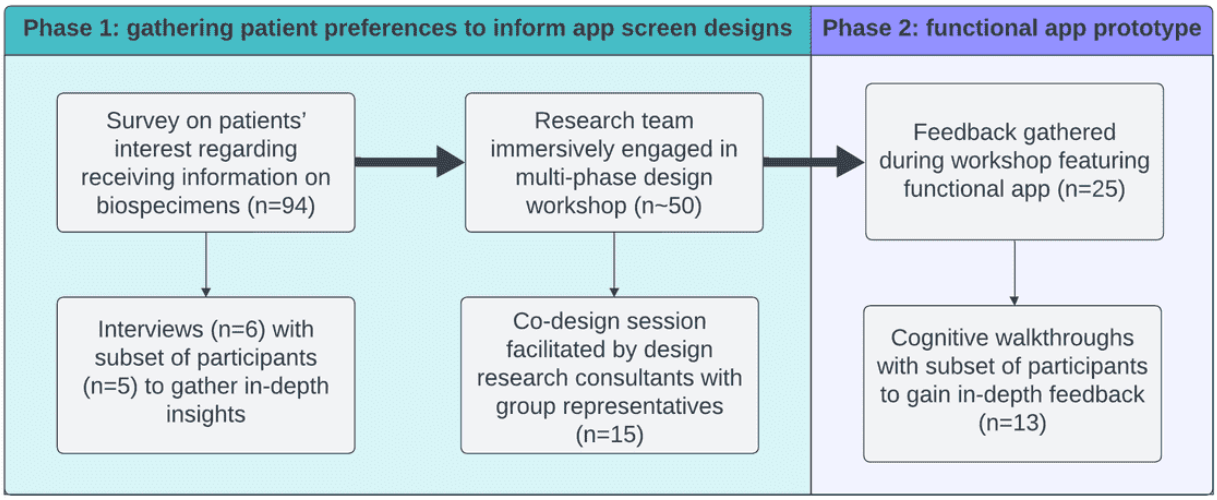
This research occurred in 2 phases (Figure 2). In phase 1, we designed smartphone screens containing different information that patients could learn about biobank and research activities

XSL•FO
RenderX

involving their samples. Embedding these screen designs in a survey, we sought to gauge patients' interest in receiving information about research or their biospecimens. We engaged a subset of survey respondents in short web-based interviews to discern their views on the importance of having this information and their opinions on its presentation and design. We held a design workshop at which participants gave feedback on the screens and suggested improvements.

For phase 2, we refined the user interfaces and developed a functional app prototype. As we developed the app, we consulted institutional stakeholders to enhance compatibility with regulations and composability of core app features, with one another and with local data architectures. We then distributed the app at a second design workshop, where participants shared thoughts on the usability and design of the app. In this phase, we conducted cognitive walkthroughs with individual participants to observe their success in using the app and to gain in-depth feedback on its functionality [26].

Figure 2. Overview of study methods.



Phase 1

App Screen Design

We used Whimsical and Figma for the initial wireframes and Flutter and Adalo for prototyping to develop the initial screen designs [27,28]. The content of these designs was informed by previous engagement with this patient population and limited by the content available within institutional biobank or research databases and publicly accessible research content. Previous survey results identified important data points to include in our patient-facing display and informed its design [22]. Working with investigators on the biobanking research protocol, their data managers, and the University of Pittsburgh Institutional Review Board, we selected data points that would be feasible and acceptable to return to patients in a future, functional solution. We then created a sample screenshot presenting a model set of samples and the corresponding hypothetical data points.

We designed 3 initial app screens presenting information about a hypothetical user's collection of biospecimens in the biobank (Figure 3). These screens were meant to be viewed sequentially, with increasing specificity of information presented about individual biospecimens in the biobank. The interface was inspired by the term "biobanking" and designed to resemble many banking apps where users first view a summary of their accounts and can click through to increase specificity about each one. Our first screen contained a summary list of a hypothetical individual's inventory of donated samples, including pictures of each. The second screen displayed high-level demographic information about each individual biospecimen: an image, its location, and the date it was collected. The third and final screen showed more details about each individual sample, such as its ID number, the medical procedure from which it was collected, the number of the freezer where it was stored, and more.

A fourth screen was created for potential users to learn about the biobank collection as a whole (Figure 4). This screen, called "connect," showed a list of subcommunities of biobank donors sorted into groups by their diagnosis code, including summary statistics indicating how many donors to the biobank shared a diagnosis (eg, invasive ductal carcinoma). The design suggests that users might be members of ≥1 of these groups (called "labs") and might use their membership to "connect" with other users, patients, or scientists.

XSL•FO
RenderX

Figure 3. The 3 "biowallet" screens, as developed in Adalo and featured in the survey, presenting different categories of specimens using accordion unfolding and individual sample–level details.

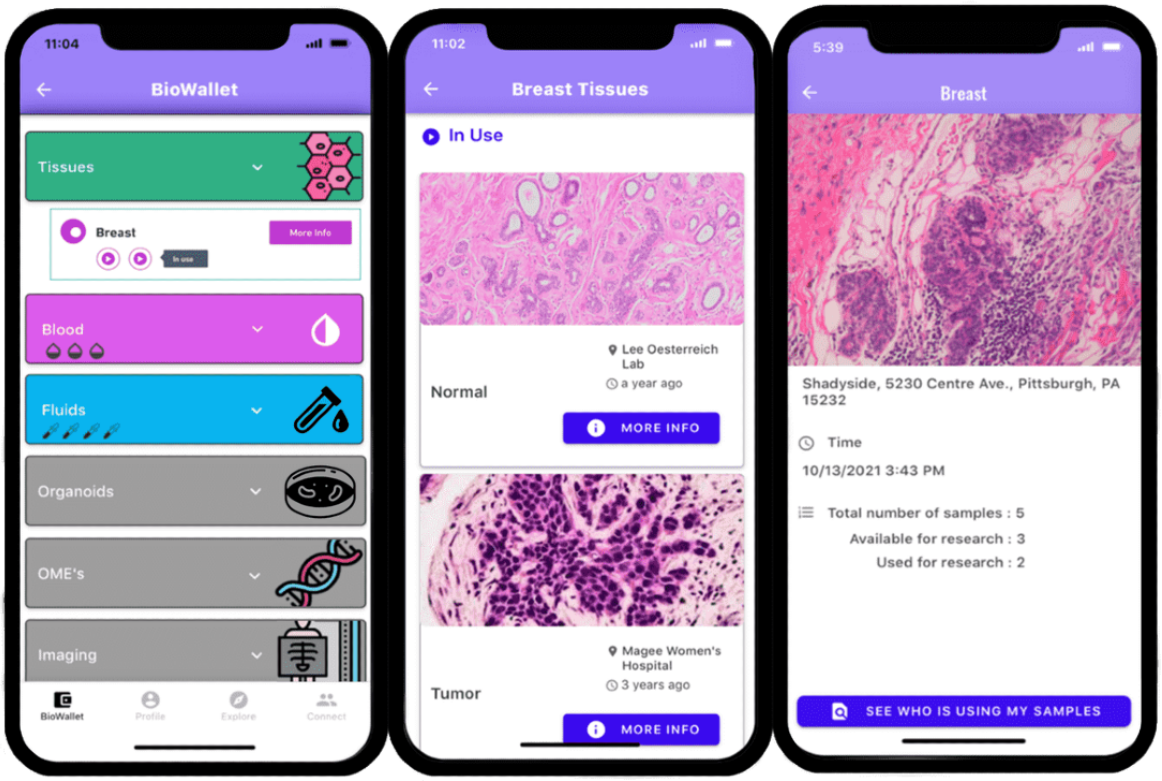
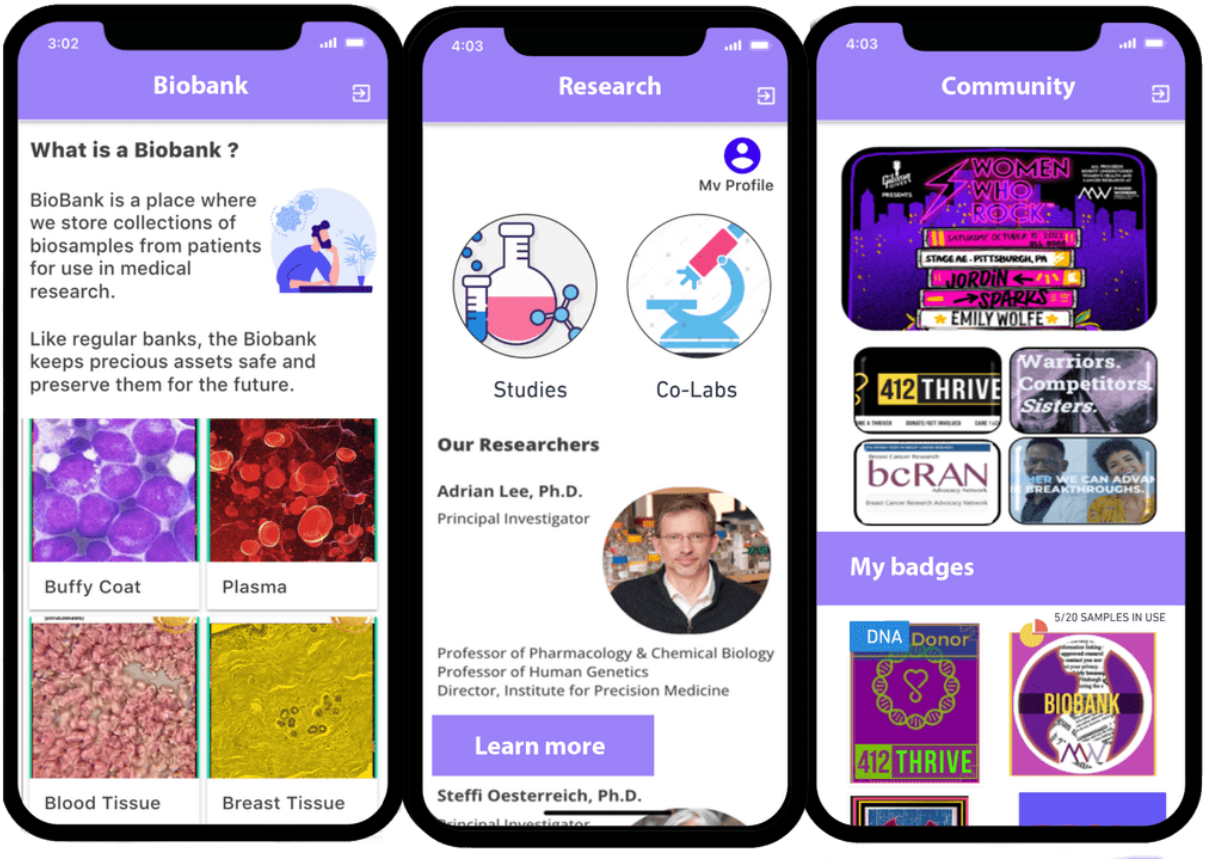


Figure 4. Biobank, research, and community screens. These screens, developed with Adalo, illustrate features explored in surveys, interviews, and workshops, including opportunities to learn about biobanking and specific specimens, follow scientists and learn about individual research programs, and connect with fellow donors and scientists.



XSL•FO
RenderX

Surveys and Semistructured Interviews

After developing the initial app screens, we embedded them into a survey to gather further evaluation and feedback from patients [29]. We conducted the survey in 2023 with a convenience sample, recruiting patients via flyers posted in breast care clinics across the health system. Participants took the survey on their own devices. IP addresses were used to identify duplicate entries from the same users. At the end of the survey, participants were invited to participate in an optional follow-up semistructured interview to discuss their survey answers. Survey content was iterated and rotated as app screens were finalized and in response to saturation of data from responses and follow-up interviews. Thus, survey data are presented with the denominator reflecting the number of participants who received each specific question and chose to reply. In total, 94 participants completed the survey, of whom 5 (5%) participated in the follow-up interviews (one participant was interviewed twice; once before and once after her cancer diagnosis). Demographics of the survey respondents, including age, race, and cancer stage, are presented in Table 1. The "Unknown" category represents participants who chose not to answer that question.

The survey asked participants their level of interest in the information presented on the app screens. After viewing the "biowallet" screens, participants were asked which of the presented data points were of interest to them. After viewing the "connect" screen, patients were asked whether identifying and belonging to a subcommunity of donors to the biobank would be of interest to them and, if so, how they might want to use the app to connect with others; for example, they were asked whether they were interested in connecting with researchers who used their samples from the biobank and, if so, what information they would like to exchange with them. They were also asked about their interest in connecting with other, similar patients who might be members of the same subcommunity (eg, have the same diagnosis code) and, if so, how.

In the follow-up semistructured interviews, participants and members of the research team viewed the participants' survey responses via screen sharing on an audio-recorded web-based call. The participant's survey responses were walked through one at a time, and the participant was asked to go into further detail about their reasoning for their answers. These interviews were meant to validate the survey questions by identifying any confusing language or images and to gain a deeper understanding of the motivations, preferences, interests, and concerns of participants. Each interview lasted 35 to 45 minutes. Interview participants were selected to represent key subsegments of the survey population, for example, age, education, and time from diagnosis.

Survey responses were analyzed, and summary statistics are presented. Transcripts of the follow-up interviews were coded for emergent themes in response to each set of screens. The results from both analyses informed changes to the presentation of biobank data and adjustments to the design of the screens to meet participant preferences and clarify content.

XSL•FO
RenderX

Table 1. Socioeconomic and clinical characteristics of survey respondents (n=94).

Demographics	Participants, n (%)[a]
Socioeconomic characteristics	
Age (y)	
<40	12 (13)
40-49	32 (34)
50-59	22 (23)
60-69	14 (15)
≥70	11 (12)
Unknown	3 (3)
Race	
White	64 (68)
Other (including multiracial)	10 (11)
Unknown	20 (21)
Household income (US $)	
<35,000	12 (13)
35,000-69,999	16 (17)
70,000-99,999	17 (18)
≥100,000	27 (29)
Prefer not to say	2 (2)
Unknown	20 (21)
Highest education	
Did not graduate high school	1 (1)
High school	6 (6)
Some college	9 (10)
Associate degree	10 (11)
Bachelor’s degree	21 (22)
Master’s degree	17 (18)
Doctoral or professional degree	9 (10)
Unknown	21 (22)
Clinical characteristics	
Breast cancer diagnosis	
Yes	56 (60)
No	35 (37)
Unknown	3 (3)
Cancer stage	
Unsure	5 (5)
0 or precancer	4 (4)
I	22 (23)
II or III	23 (24)
IV	2 (2)
Unknown	38 (40)
Time from diagnosis	
≥5 y ago	11 (12)

XSL•FO
RenderX

Demographics	Participants, n (%)[a]
>1 and <5 y ago	31 (33)
≤1 y ago	14 (15)
Unknown	38 (40)

[a]Due to rounding, not all percentages sum to 100%.

Design Workshop 1

We sought feedback on the "biowallet" screens, the "connect" screen, and 2 additional screens: "explore" and "profile." The "connect" feature was expanded to include 3 screens showing a list of all "lab" groups and subcommunities, a description and membership details of 1 specific "lab," and an action button to "add my samples" to a "lab." The screen called "explore" contained additional biobank-level data about the collection of biospecimens. The last addition, "profile," contained a hypothetical user-chosen avatar to represent themselves, as well as a collection of hypothetical achievement badges representing activities that users might be able to complete both in the app ("claimed my biowallet") and in research participation ("signed a consent form"). It also contained a progress bar indicating the users' completion of in-app tasks and achievements. We presented the second iteration of the screens to approximately 50 members of a local support group for patients with breast cancer in an embedded participatory design workshop. Members of the research team presented the screens, gave a short overview of their contents, and explained the overall intent of decentralized biobanking as well as the purpose of the dedicated app codesign workshop.

Participants provided feedback about the overall approach during a three-day participatory engagement event where the research team presented to and engaged with the local breast cancer patient and previvor population in an immersive setting. In this paper, we focus on describing results for 15 group representatives who attended a professionally facilitated in-person codesign workshop where they provided detailed feedback on individual app screen designs. Participants were given printed images of the screens, along with stickers—thumbs up or a heart to signify something they liked, a question mark to signify something they were confused by, and an exclamation point or a frowning face to signify something they did not like. Participants were asked to discuss the screens as a group and to place stickers on the printed images, either to signal their reaction to the entire screen or specific parts of the image. Some groups wrote on the images or in the margins to elaborate on their responses. As not all participants responded to all app screens, workshop data for all subsequently described workshops and cognitive walkthroughs are presented with the denominator reflecting the number of participants who provided written feedback on a given screen, page or design element.

The printouts were collected after the workshop, and the research team performed a content analysis of the sticker placements and notes left by participants. The identified themes were considered along with design research synthesis from expert consultants, as well as analysis of survey responses, semistructured interviews, and immersive engagement. Taken together, these themes informed further design changes that were then incorporated into a functioning mobile app prototype.

Phase 2

App Prototype Development

We developed a functional mobile app using Flutter [30]. The front-end patient-facing interfaces contained 4 sections: "biowallet," "explore," "connect," and "profile." We also developed a backend data architecture in which real biobank data could be imported into the app so that users could see personalized, real-time information about their biospecimens contained in the biobank collection. During the app's development, we consulted with institutional stakeholders to ensure regulatory compliance and compatibility with existing data resources to minimize burdens on researchers and data managers. Elsewhere, we describe a "digital honest broker" approach that enabled pilot participants to be connected directly to their own deidentified biospecimens [31].

Design Workshop 2

We held a second design workshop with a second local advocacy and support group for patients with any cancer type. A total of 25 participants attended. Members of the research team presented the app screens, gave an overview of their contents, and explained the intent of the workshop. Participants were then instructed to access the app on their personal devices and given a sample log-in leading to a deidentified demonstration account with biobank-level data and individual biospecimen data populated. Participants were given a sheet of paper with three columns—"Things I Liked," "Things I Didn't Like," and "Didn't Meet Expectations"—and were asked to write corresponding reactions as they explored the app freely by themselves or with one or two other participants. The papers were collected from nine small groups. Content analysis was performed to identify common areas of interest, enjoyment, displeasure, and dissatisfaction.

Cognitive Walkthroughs

During the second design workshop, members of the research team invited individuals at random to participate in a cognitive walkthrough to evaluate the usability of the application. Of the 25 workshop participants, 6 (24%) completed the walkthrough exercise. Participants were asked to perform several in-app tasks and seek specific information presented in the app while narrating out loud what they were doing. Members of the research team observed them using the app and took notes about their narration, evaluating their success in completing the assigned tasks.

We held seven additional cognitive walkthroughs with users recruited from participants in the phase 1 survey who agreed to be contacted about future research. These participants were

XSL•FO
RenderX

patients at the collaborating health system and were guided through the onboarding process that connected them to their own biospecimen information, if any existed. All users were able to see the "explore" and "profile" screens. Users who had also consented to the biobanking protocol (which was manually verified by the research team) were able to see the "connect" screen. If the user had samples in the biobank, they could see the "biowallet" screen populated with real data about their samples. Participants were led through the same cognitive walkthrough and asked to complete tasks and seek out specific information while they narrated their process as researchers observed and took notes.

Cognitive walkthroughs measure usability by collecting qualitative, experiential data from users directly. Researchers can infer usability by noting whether the users understand the task and are trying to achieve the right outcome, whether they take the appropriate action to achieve their intended outcome, and whether they correctly identify their success or the information sought from them. By having a mix of participants viewing uniform, hypothetical data and those viewing personalized data, we were able to gain generalized insights into usability, as well as insights into the personal value and added joy or frustration when engaging with one's own data. Both sets of participants were asked to create an account; perform the necessary authentications to populate their biospecimen data; identify how many of their samples had been used in research and how many were available; and to find a protocol that had used their samples, or, if their samples were unused, to find a protocol that studied their particular diagnosis.

Ethical Considerations

This study was approved by the University of Pittsburgh Institutional Review Board and quality improvement committees (IRB22010118, IRB22020035, and QRC3958). Participants provided informed consent before engaging with the study. The voluntary nature of the study was emphasized, including that their decision to participate in the study would have no impact on their care. Participants were able to skip questions or withdraw at any time. To safeguard privacy, data were stored on a secure institutional server, with survey responses deidentified and stored separately from personal information. Survey participants were paid US $10 for their participation upon completion. A subset of survey participants who agreed to participate in the follow-up interview were paid an additional US $20.

Results

Phase 1

Surveys

In reaction to the "biowallet" screen, participants were asked about their interest in using such an app to track their biospecimens donated to research. Of the 89 respondents, 20 (22%) were extremely interested, 22 (25%) were moderately interested, 24 (27%) were neutral or not sure, and 23 (26%) were not interested. Survey questions presented potential app features related to sample tracking, research feedback, and community engagement (Table 2), and follow up questions drilled down on preferences regarding specific design elements and user priorities.

We next assessed participants' interest in receiving updates regarding how, and by which researchers, their samples were used (Table 2). Participants (61/81, 75%) were most interested in learning "the outcomes of studies that use my tissues, like publications or new discoveries." Of these 61 participants, 14 (23%) provided qualitative details in optional open-ended follow-up questions, describing how or why they were interested in learning about research on their samples. For example, one respondent noted "I believe research participants should have full access to the data that is collected from them...The idea that my genetic data is setting in a lab for a variety of research studies, but I have no access to any of that data or the related information generated by that data is not fair or in my best interest."

The second highest area of interest concerned learning the status of their biospecimens regarding whether they had been used in research (45/81, 56%). Only 23% (19/81) of the participants were interested in seeing how many samples they had compared to other patients with the same diagnosis. Other data points of interest were the medical procedure from which their sample was collected (12/27, 44%), the current location of their sample (11/27, 41%), and collection details (when, volume, etc; 10/27, 37%).

Subsequently, we asked participants about their broader preferences for updates from researchers using their biospecimens with respect to scientific findings, new developments, and modes of communication (Table 2). Participants (47/60, 78%) expressed the greatest interest in receiving patient-friendly summaries of scientific information. Updates on new drugs or tests that become commercially available (39/60, 65%) and findings that may inform care (38/60, 63%) were also highly valued. Fewer patients were interested in updates on new publications (28/60, 47%), patient-friendly videos (22/60, 37%), study enrollment numbers (19/60, 32%), and new research grants (15/60, 25%).

In reaction to the "connect" screen, respondents expressed interest in connecting with other patients. Participants were most interested in connecting with patients with a similar type of breast cancer to theirs (41/83, 49%) and patients in their age group (36/83, 43%). Of the 83 respondents, 24 (29%) were not interested in connecting with any others. Table 2 details additional areas of interest.

XSL•FO
RenderX

Table 2. Interest in potential app features: biospecimen tracking, research updates, and community engagement[a].

App feature interest	Survey responses, n (%)
Sample tracking ("biowallet") features (n=81)	
Outcomes of the studies that use my tissues, like publications or new discoveries	61 (75)
Learning if/when my biospecimens are shared with researchers	45 (56)
Seeing how my own samples look under the microscope	35 (43)
Learning about the individual researchers who are studying my biospecimens	27 (33)
Seeing how many samples I have compared to patients with the same diagnosis as me	19 (23)
None	12 (15)
Other	3 (4)
Research updates ("scientist labs") (n=60)	
A patient-friendly summary of the scientific information	47 (78)
If there are new drugs or tests discovered that become commercially available products	39 (65)
If there are findings that might help my doctor and I make decisions about my care	38 (63)
When new publications come out (with access to that publication)	28 (47)
Short patient-friendly videos describing the research	24 (40)
How many patients are included	19 (32)
If there are new grants awarded to that research	15 (25)
Community engagement ("donor labs") (n=83)	
Similar type of breast cancer	41 (49)
Patients in my age group	36 (43)
Patients in the same research study	30 (36)
Similar stage of breast cancer	25 (30)
None	24 (29)
Similar genetic risk of cancer	18 (22)
Patient advocates	12 (14)

[a]These are not the only survey questions posed, but represent three key areas that were included in the survey. Where appropriate, questions were iterated to include screenshots of the app designs or concepts as they were refined.

Semistructured Interviews

While learning the outcomes of research on their biospecimens was identified as a high priority, concerns arose in the follow-up interviews. Of the 5 participants, 3 (60%) expressed concern about imposing additional burdens on busy scientists. They had expressed interest in learning about research on their specimens but clarified in interviews: "only if it doesn't take away from the work" and "I don't want to bother researchers, who are very busy" (Table 3). However, participants valued the opportunity to learn more about their samples; participant 3 expressed how specimen tracking can provide transparency about "what it's [biospecimen] being used for," implying how transparency may build trust.

In the follow-up interviews, 3 (60%) of the 5 participants expressed an interest in and need for communities that connect patients to research in general, stating that existing patient organizations were focused on emotional support, living with cancer, and survivorship rather than on biomedical research or the latest advances in treatments. Participant 2 expressed frustration with the lack of medical and scientific expertise in many patient support groups (Table 3): "...It's frustrating sometimes to get the kind of information you're looking for as a patient...with metastasis."

One participant undergoing diagnostic workup for cancer expressed interest in an additional support group, but without the granularity of connecting with other patients involved in the same studies. In contrast, one interview participant with a longer duration of cancer survivorship expressed that they did not have a need for additional patient connections. Multiple participants (3/5, 60%) expressed an interest in contributing additional information to scientists to provide additional context to their biospecimens. This included integrating data from genetic tests (eg, 23andMe), wearable devices (eg, Fitbit), and personal medical histories to support scientists. The motivation for doing so seemed to stem from the opportunity to "help...the scientific community, the research for cancer" (Participant 1) and elucidate "genetic links to breast cancer" (Participant 5). Participant 3 expressed interest in being informed about "a list of studies" they might qualify for, indicating a proactive approach to supporting biomedical research.

XSL•FO
RenderX

Table 3. Qualitative interview insights. The interview quotes are organized around feature categories of biobanking education, specimen tracking, connecting with others, and dynamic data sharing. Key standout phrases are italicized,

Proposed features ("naming convention") and interview quotes	Emerging themes
Learn biobanking ("biobank")	
• "I *totally geek out* on this stuff, so as much as you guys are willing to share and whatever level detail, I'm all about it." (Participant 1, Interview 2) • "I would want to know outcomes of the studies. I am kind of *a science geek*, like that's my interest, it's not the field that I'm in whatsoever, and I was not educated in that, but I'm very interested. I've read a lot about science and especially like bioscience, so I would say, I would also like to know outcomes, with the studies that use my tissues." (Participant 3)	Valuing scientific knowledge and approaches
• "Whatever you guys are willing to share is perfectly fine...*I just don't expect that detailed level of personal information* about the researchers, but if they're willing to share...it's fascinating...it's just not one of the absolutely necessary things, I feel." (Participant 1, Interview 2) • "*I don't expect to have the ability*, or even the need to identify specific researchers who I can go to. It's like as many people can use that tiny little tumor that you guys got out of me as possible it's fine, so what, for whatever purpose and whoever needs it, for whatever reason, is fine." (Participant 1, Interview 1) • "It was my pre-op...my surgeon, and a nurse, she just had a whole pile of papers. It [the biobanking consent form] *was just one more paper she needed to get signed*." (Participant 3)	Recognized unfamiliarity with biobanking, limited expectations and agency of donors within in current model
Specimen tracking ("biowallet")	
• "I don't need to know about all the researchers and everyone who's touching it. Just in general, like *what you're looking at and...what you're doing with it.*" (Participant 1, Interview 2) • "I don't feel the need to have control over any of that stuff—You guys can do with it what you see fit." (Participant 1, Interview 1) • "I *don't want to bother researchers*, who are very busy." (Participant 2) • "It would be interesting to see my samples, and my tissue, but not necessary, I guess. So, sort of like an added bonus." (Participant 5)	Balancing gaining personalized insights versus resource constraints for scientists
• "I mean, worst case scenario, like these are my few... last few minutes on earth... It would bring meaning...I'm *suffering but I'm suffering for a reason.*" (Participant 4) • "...People are fearful of having their information being used in nefarious ways; I think *this is a way of being super transparent and also involving the candidate*...like it's personal, it's something [very] personal, especially for women with breast tissue, very personal." (Participant 4)	Recognition for personalized contributions as a source of purpose and meaning
Community engagement ("labs")	
• "Probably in the lines of a support group kind of thing, it would be kind of neat to see...what other people were experiencing and going through, but not necessarily with regards to the study itself." (Participant 1, Interview 1) • "You have questions in there, 'Do you want to be connected with other donors?' and [in the survey] I said no, 'I don't use social media,' but I would use this, and perhaps if I do have to get surgery or I...perhaps [have cancer]... *it's like people who have dogs from the same litter, you know it's like, 'How's your pup doing?'*" (Participant 3) • "You know, now that I am talking this through, I would also...need...the type, the stage, the genetic risk of patients in my age group. I might want any one of those options. I am thinking futuristically now that it's become an actual reality or possibility...*I would be more open to connecting to others...beyond the same research.*" (Participant 3)	Some felt existing support groups were adequate, while others sought camaraderie and citizen science
• "One of the limitations with the patient advocacy groups that I've been in so far is that *there is limited expertise in the groups*...so it's frustrating sometimes to get information you're looking for as a patient...for example, drugs for preventing metastasis...what are the researchers recommending at this point [for people like me]." (Participant 2) • "I think the research is important and I'm just interested in knowing what research is going on, just because I think it's interesting, but *I wouldn't necessarily need to know what's happening with other patients* and their stuff."(Participant 5) • "I want to help, do whatever is possible to help research continue, but I don't feel the need to use something like this to connect with other people [i.e., patients]." (Participant 5)	Unmet need for direct connection with scientists, as opposed to other patients

XSL•FO
RenderX

Proposed features ("naming convention") and interview quotes	Emerging themes
Data sharing ("profile")	
• "I have the ancestry/23andMe, I do Fitbit stuff, so yeah, that'd be very cool to be able to incorporate all that...Anything that would help in any way... the scientific community, the research for cancer, anything, advanced in any way, for whatever reason, whatever purpose." (Participant 5) • "There's so much that's not known in terms of...genetic links to breast cancer...*I would share more information if it would help.*" (Participant 5) • "I would like to see a list of studies that I might qualify for, *I would provide details about my medical history or treatment.*" (Participant 3) • "[If] you guys need additional information or stats or blood or anything from me, I'm more than happy to help." (Participant 1, Interview 1)	Eager to contribute more specimens and data
• "I keep throwing out those ads for...'All of us' *because they won't give you anything...that's my reason for not participating*...[since] you have all my genetic information, you have no idea where it's going to be five years from now...But you won't let me have any access to it." (Participant 2) • "I think it's brilliant and I also think that *more people would participate if this were available*...Because I've had opportunities in the past, and...the biggest question that I have, and other people have, even when we're asked, 'Would you donate organs'...is the lack of knowledge about what it's being used for...that's what makes people suspicious and, especially, in an era where we both have like open source, you can find out information about everyone." (Participant 3)	Demand for greater reciprocity and accountability for data exchange

Design Workshop 1

Participants of the first design workshop (n=69) confirmed their interest in the concept of the app as well as the content of the app screens and provided valuable feedback on its presentation. A subset of 15 participants who participated in a dedicated codesign session documented their reactions to individual app screens. The "biobank" screens were appreciated by participants who liked the ability to view the entire collection of one's personal samples (7/15, 47%), as well as the sample-level details under the individual "biowallet" (4/15, 27%), with several emphasizing the linkage between samples and other medical or genetic data. Some were confused or upset by the technical details presented about each individual sample and about some of the terminology related to their biospecimens (eg, questioning what it means for a sample to be depicted as "in use"; 9/15, 60%). Of 15 co-designers, 4 (27%) noted that they would have preferred their samples to be organized chronologically, corresponding to their lived experience of appointments and procedures as specimen collection events, rather than initially stratified by sample type (tissue, blood, urine, organoids, etc.). For example, one co-designer emphasized: "Everything is timeline for us [patients with cancer]."

Moreover, 80% (12/15) of co-designers provided positive reactions to the "connect" screens, where they hearted the framing of virtual "lab" groups as 'patient-led' (5/12, 42%) and appreciated that there seemed to be actions that a user could take, such as joining a virtual "lab" and adding their samples to the group (6/12, 50%). Of the 12, 4 (33%) expressed confusion about the "lab" groups, including who is supposed to join them and for what purpose, while 3 (25%) expressed concerns about the legitimacy of *patient-led* "labs" because they might relate to actual future research on their samples, and they worried about the potential for unmoderated, off-topic, and illegitimate activity, with 1 participant stating "It should not feel like social media. To me it should feel more like science."

The "explore" screen, showing high-level information about the entire biobank collection of samples, prompted the most confusion. Of the 14 co-designers who were able to review this screen, 12 (86%) were confused by ≥1 terms on this page (eg, "consent" and "protocols"), and 2 (14%) suggested the addition of a glossary to help users understand. Other co-designers (3/14, 21%) incorrectly assumed that the "explore" page contained information about an individual's samples instead of the entire collection.

Participants also generally misunderstood the intended features of the "profile" page. Of the 11 co-designers who reviewed this page, 7 (64%) expressed confusion or dislike of the badges feature, with concerns that they were "useless," "not important," or excessively similar to features on social media sites, which they did not like. Finally, 18% (2/11) of the co-designers expressed concerns about the visibility of individual profile data, similarly worried about the parallels to existing social media, where all users can see individual profile pages.

Phase 2

App Prototype Development

Taking participant feedback from phase 1 activities, we developed a functional app prototype with 4 sections corresponding to the screen designs. The "biowallet" screens had been well received, and minimal changes in content were made. We removed the diagnosis code from the "biowallet" page to avoid triggering negative memories and emotions in users. We maintained sample organization by type (tissue, blood, urine, etc) to more closely resemble the structure of biobank data, where the type of sample and medium of preservation are more valuable than the collection date. However, to address participants' desire for a chronological arrangement, we listed

XSL•FO
RenderX

samples from the most recently collected to oldest *within* the sample type category. The "biowallet" page also showed a record of the users' informed consent to join the biobanking protocol. Users could click on individual samples to find information such as size, location, date collected, freezer number, and more. A generic picture of the corresponding sample type (eg, a picture of urine crystals under a microscope) accompanied each sample. Users could see whether each sample had been used in a research protocol or whether it was still available in the biobank.

The "connect" feature included patient-led "labs" organized by diagnosis, as in the original designs. We added a second class of "researcher-led labs" based on individual researchers who had used the patients' samples. Due to the structure of deidentification of the initial biobank data, the names of the researchers and details about their work were blinded, with just a numerical code representing individuals (eg, principal investigator 4); summary statistics of their work (eg, 6 protocols using 75 samples); and similar blinded, summary information about specific research protocols using patient samples (eg, "Your samples were used in protocol #4, which used 25 samples"). Importantly, displaying this information did not impose any additional workload or data input responsibilities on researchers. The patient-led and researcher-led "labs" featured some cross-referencing. A patient-led group with a shared diagnosis was linked to researcher "labs" studying this diagnosis and to protocols using samples with this diagnosis. Using existing databases, we sorted individual users into the appropriate patient-led and researcher-led "labs," adding a third section, "My Labs," where users could view just the groups of which they were members. These 2 new groups of "labs" were developed to satisfy phase 1 participants' enthusiasm for learning about research both in general and specifically about their samples, while maintaining maximum interoperability with existing biobank databases and protocols.

We updated terminology on the "explore" page in response to widespread confusion expressed by participants at the design workshop. Rather than relying on institutional terms such as "informed consent," we framed participation in the biobanking protocol as being a "member." For each sample category, we clarified how many patients were represented and how many samples there were in response to participants' confusion about whether they were looking at 1 person's samples on the "explore" page or the entire biobank's samples. The "explore" page contained expandable information about the entire biobank collection, organized by sample type (eg, breast tissue, blood, and urine).

On the "profile" page, users could view the deidentified code that links their account on the app to the biobank database, populating their sample information. Rather than showing a hypothetical set of badges that users might be incentivized to collect, which were not well received, we showed just 1 indicator of whether users had successfully been connected to their sample collection. On the "profile" page, we also developed short questionnaires about potential future research topics of interest. This feature served as proof of concept for how the app might be used to engage participants in research design or be used as a data collection tool.

Design Workshop 2

The second participatory design workshop aimed to further validate the clarity and value of the information presented in the app. In addition, we received feedback on the user experience of navigating the live app.

All 9 groups of patients participating in this workshop (consisting of 25 patients total) liked learning the information presented to them on the app. Of the 9 groups, 3 (33%) specifically mentioned the "real-time" nature of the information, with one-third (n=1, 33%) stating that they liked being able to "follow along" with research. Of the 9 groups, 5 (56%) noted that they would prefer or require more information before using the app, such as background materials or guided onboarding, to better understand the information presented; 7 (78%) wished for more information (eg, details about research protocols); and 4 (44%) wished for more capabilities (eg, a search bar or the ability to make their biospecimen data more or less visible to others).

Of the 9 groups, 7 (78%) noted that the app was easy to use. However, 7 (78%) of the 9 groups also noted usability and navigation as something they disliked. Of the 9 groups, 4 (44%) noted that the onboarding process—downloading the app, creating an account, logging in, and verifying the user's identity—did not meet their expectations or was not liked; 6 (67%) did not like a design feature of the app (eg, the choice of icon and the order in which the screens appeared); and 3 (33%) listed unmet accessibility expectations, including the desire for use on a larger screen or the preference for a web platform rather than a mobile app.

Echoing concerns from the first workshop, of the 9 groups, 3 (33%) were dissatisfied with the lack of explicit security assurances or data use terms, while 1 (11%) liked the fact that users' identities were not featured on the app and were instead represented by a unique deidentified code.

Cognitive Walkthroughs

Cognitive walkthroughs with 13 patients also gauged the usability of the app. The walkthroughs illuminated an onerous and confusing onboarding process, where participants had to create an account and submit information to link their biobank data, after which they received a notification and request for more information once their biospecimen data were populated in the app. For users who were participants in the biobank, the interim waiting period for biospecimen data was as long as 1 week, and users often mistakenly thought they had done something incorrectly. Nearly all walkthrough participants expressed confusion and uncertainty about the onboarding and biospecimen data verification process, even when completed successfully.

Once they successfully logged in and were able to view biospecimen data, most of the participants successfully described their biospecimen data, including how many of their samples were available for research. Those participants who were members of the biobank, viewing their own biospecimen data, were curious as to why their samples had not been used and expressed a willingness to take action to ensure that their samples were used in research (eg, by offering to donate more

XSL•FO
RenderX

samples or to pool samples with other participants). Overall, the success rates of completing this task indicated the usability of the app. Lower usability, namely the inability to identify successful completion of the task, arose in specific contexts applicable only to those participants who were attempting to view their own biobank data. Of the 13 participants, 3 (23%) had consented to participate in biobanking, but none of their clinical biospecimens made it to the biobank. Thus, their "biowallet" page showed no samples, although their "membership" was verified. This was a source of confusion, and participants incorrectly assumed that they had made a mistake in the onboarding process.

Participants were able to correctly identify a protocol that had used their samples or that had studied their disease. These participants successfully navigated from the "biowallet" screen to the "connect" screen and correctly understood the information presented about the protocol they identified. Some of the participants (7/13, 54%) expressed disappointment in the limited amount of information presented about the protocol or the associated researcher. Consistent with the initial survey findings, participants were most interested in learning the details and outcomes of the research conducted on their samples or their disease. In sum, the cognitive walkthroughs showed the usability of and interest in the app and revealed the need to improve the onboarding process and to better explain less intuitive scenarios, such as having no banked samples as a consented biobanking participant. The walkthroughs also confirmed participant interest in learning what happened to their biospecimens and about research on banked biospecimens.

Discussion

Key Findings

Throughout the design process, participants indicated a strong interest in and demand for learning about biobanking. They found intuitive value in the ability to track their own biospecimens. Indeed, participants positively received almost all of the information presented and commonly expressed a desire for more. In all phases, participants were most interested in learning about the research being conducted on their disease or specifically using their biospecimens, including the results of this research. They noted that they were not able to find this information on their own and were not able to satisfy general curiosity about research by participating in existing patient advocacy and support groups. These findings add to existing literature confirming public interest in greater transparency and engagement in biomedical research [32-36].

Mixed responses about the usability of the app demand further research and attention. Overall, participants in all stages were interested in using the app, although the content and presentation of information needed improvement. Some of the participants sought language and design similar to that of their medical records. Presenting research information in this way may make it more intuitively understandable but may also contribute to therapeutic misconception by misleading participants into believing that the research information and participation are of a similar clinically actionable quality. Thus, careful onboarding and framing of research information must ensure participant understanding and offer opportunities to overcome confusion.

Participants expressed a demand for a "legitimate" platform that did not suffer from what they saw as the risks and ills of social media: data misuse, a lack of privacy, and the potential for negativity or misinformation. While mobile phones and apps are widely used [37], concerns about mobile apps may make them an unacceptable platform for some. An additional facet of building trust with participants can be achieved through transparency surrounding how their biospecimens are being used. An interview participant expressed the following:

> *The biggest question that I have and other people have even when we're asked you know, would you donate organs...is...the lack of knowledge about what it's being used for.*

This approach to rebuilding trust reflects ongoing movements in biomedical research aimed at improving transparency and accountability, such as interactive and accessible "data walks" designed to engage communities and stimulate dialogue [38].

However, the de-identification of biospecimens presents challenges for direct transparency and engagement. Other challenges exist as well—meaningful transparency and engagement with patients is time and labor intensive [39,40]. Even when community engagement is required or encouraged, there are few tools or guidelines for success, and interactions may be transient and transactional [40]. Biobanking and scientific research are deeply technical disciplines that historically have minimally interfaced with the public, and require extensive time, skill, and technology to translate into accessible, meaningful plain language [39]. Finding relevant patient communities and sustaining meaningful, collaborative relationships with them might prove beneficial to biobanks, or scientists themselves, but this is not required, supported, or incentivized by traditional academic reward systems or many institutions [41-43]. Engaging patients in biobanking might fulfill obligations of respect; however, evidence of its long-term impact on research participation, understanding, and attitudes is the subject of ongoing work [40].

Our experience confirms the need for specialized skills and dedicated time and resources to deliver transparency and engagement effectively and safely. New technologies, eg blockchain and generative artificial intelligence, are essential for these endeavors to be more efficient and scalable, but their development remains time and labor intensive. Participants were largely unfamiliar with biobanking, research processes or related jargon. Thoughtful, readily digestible and personalized education about each unique research context may be a critical aspect of the app onboarding process. Participants were confused about less intuitive but common scenarios, such as having consented to biobanking but having no samples in the bank. While this confusion may stem from a lack of recall or understanding of the initial informed consent for biobanking, delivering follow-up information about actual participation (or lack thereof) presents new risks of misunderstanding that must be managed.

XSL•FO
RenderX

While our study focused on patients with breast cancer, the demand for transparency and engagement may extend to other medical fields. We found that the survey participants (61/68, 90%) were most interested in learning the outcomes of research using their tissue samples. Thus, conditions with active research and strong patient interest in new treatment paradigms, such as additional cancer types and sickle cell disease, are potential areas where this platform may be effective. Further research is needed to assess the feasibility, acceptability, and effectiveness of this platform across diverse medical contexts and patient demographics.

Finally, in light of the strong reception and demand by participants, it must be acknowledged that returning deidentified, individual-level research information to participants is uncommon and unaccommodated by existing paradigms. Existing data architectures are built for a research enterprise from which participants are intentionally excluded. Working to disrupt the paradigm by allowing participants to view research data and activities presented both technological and regulatory implications. While participants were unaccustomed to receiving this type of information and, as a result, sometimes struggled to understand it, biobanks were also unaccustomed to sharing this type of information and, as a result, were sometimes hesitant to do so. Any technological tool aiming to enable transparency and engagement in research must be acceptable to all stakeholders. Careful and inclusive co-design is imperative.

Limitations and Next Steps

The patient population we engaged in this work consisted of only women due to the nature of breast disease. Our participants were mostly White (64/94, 68%) and were more educated and had higher incomes than the national average, although they more closely resembled local demographics. Our participant population may also reflect self-selection bias—those interested in learning about research on their biospecimens or in general may have been more likely to participate in any research, including ours. Our technology was intended for piloting among this group and was therefore tailored to their preferences. For these reasons, further research will be necessary to validate interest in and usability of the technology in other patient populations (eg, men and healthy individuals).

Moreover, most of the patients we engaged with throughout all stages of design and development were unfamiliar with the term "biobank." This underscores a key challenge associated with delivering an accessible and patient-friendly biobanking app. This variation in patient knowledge regarding biobanking, even among those who already consented to biospecimen donation, limits the ability to fully capture and address diverse patient preferences.

Across the surveys and workshops, questions and app content were iterated upon as app screens were finalized, new topics were raised by participants, and data saturation was reached through survey responses and follow-up interviews. As a result, the sample size of participants exposed to specific materials varied, leading to disproportionately lower sample sizes for certain areas of inquiry. Additional work is needed to investigate a larger sampling of patients' design preferences. New technologies can facilitate meaningful transparency in biomedical research to patients in a personalized yet efficient way. This research showed that patients desire transparency and revealed the specific information they were interested in learning about their biospecimens. By deeply engaging patients and other stakeholders in participatory methods, siloed and specialized research data can be translated into meaningful patient engagement. We worked to design a mobile app that allows patients with breast cancer to track and learn about their donated biospecimens. Surveys, interviews, design workshops, and cognitive walkthroughs helped to make the app maximally valuable and usable for patients while maintaining efficient integration with existing data assets.

By creating this decentralized biobanking app, we built a platform upon which methods of transparency and engagement can be tested in subsequent pilots; for example, further research can measure the app's impact on patients' willingness to consent to biobank research, to remain enrolled and engaged longitudinally, and to actively participate in community-engaged research.

Acknowledgments

The authors thank Diana Mendoza Cervantes, Sneha Vaidhyam, and Dmitriy Babichenko, as well as Raelynn O'Leary and Ashley Deal of Dezudio, for their contributions to study design and execution. The authors also thank Sam Moatz, Jen Kehm, the members of 412Thrive, and the members of Cancer Bridges. Finally, the authors express utmost gratitude to the participants of this study: "Your openness, honesty, and trust motivate all that we do."

Conflicts of Interest

AH, WS, and MG are consultants for de-bi, co. All other authors declare no conflicts of interest.

References

1. Boyer GJ, Whipple W, Cadigan RJ, Henderson GE. Biobanks in the United States: how to identify an undefined and rapidly evolving population. Biopreserv Biobank 2012 Dec;10(6):511-517 [FREE Full text] [doi: 10.1089/bio.2012.0034] [Medline: 24845137]
2. Henderson GE, Cadigan RJ, Edwards TP, Conlon I, Nelson AG, Evans JP, et al. Characterizing biobank organizations in the U.S.: results from a national survey. Genome Med 2013 Jan 25;5(1):3 [FREE Full text] [doi: 10.1186/gm407] [Medline: 23351549]

XSL•FO
RenderX

3. Tarling TE, Byrne JA, Watson PH. The availability of human biospecimens to support biomarker research. Biomark Insights 2022 Apr 19;17:11772719221091750 [FREE Full text] [doi: 10.1177/11772719221091750] [Medline: 35464611]
4. Li Y, Tang P, Cai S, Peng J, Hua G. Organoid based personalized medicine: from bench to bedside. Cell Regen 2020 Nov 02;9(1):21 [FREE Full text] [doi: 10.1186/s13619-020-00059-z] [Medline: 33135109]
5. Barnes RO, Watson PH. Precision medicine: driving the evolution of biobanking quality. Healthc Manage Forum 2020 May;33(3):102-106. [doi: 10.1177/0840470419898874] [Medline: 31957497]
6. Olson JE, Bielinski SJ, Ryu E, Winkler EM, Takahashi PY, Pathak J, et al. Biobanks and personalized medicine. Clin Genet 2014 Jul;86(1):50-55 [FREE Full text] [doi: 10.1111/cge.12370] [Medline: 24588254]
7. Zhou Z, Cong L, Cong X. Patient-derived organoids in precision medicine: drug screening, organoid-on-a-chip and living organoid biobank. Front Oncol 2021;11:762184 [FREE Full text] [doi: 10.3389/fonc.2021.762184] [Medline: 35036354]
8. Castillo-Pelayo T, Babinszky S, LeBlanc J, Watson PH. The importance of biobanking in cancer research. Biopreserv Biobank 2015 Jun;13(3):172-177. [doi: 10.1089/bio.2014.0061] [Medline: 26035006]
9. Warner TD, Weil CJ, Andry C, Degenholtz HB, Parker L, Carithers LJ, et al. Broad consent for research on biospecimens: the views of actual donors at four U.S. medical centers. J Empir Res Hum Res Ethics 2018 Apr;13(2):115-124 [FREE Full text] [doi: 10.1177/1556264617751204] [Medline: 29390947]
10. Sobel ME, Dreyfus JC, Dillehay McKillip K, Kolarcik C, Muller WA, Scott MJ, et al. Return of individual research results: a guide for biomedical researchers utilizing human biospecimens. Am J Pathol 2020 May;190(5):918-933 [FREE Full text] [doi: 10.1016/j.ajpath.2020.01.014] [Medline: 32201265]
11. Tomlinson T, De Vries RG. Human biospecimens come from people. Ethics Hum Res 2019 Mar;41(2):22-28 [FREE Full text] [doi: 10.1002/eahr.500009] [Medline: 30895751]
12. Rothstein MA. Is deidentification sufficient to protect health privacy in research? Am J Bioeth 2010 Sep;10(9):3-11 [FREE Full text] [doi: 10.1080/15265161.2010.494215] [Medline: 20818545]
13. Gross MS, Hood AJ, Rubin JC, Miller RCJ. Respect, justice and learning are limited when patients are deidentified data subjects. Learn Health Syst 2022 Jul 04;6(3):e10303 [FREE Full text] [doi: 10.1002/lrh2.10303] [Medline: 35860318]
14. Mamo N, Martin GM, Desira M, Ellul B, Ebejer JP. Dwarna: a blockchain solution for dynamic consent in biobanking. Eur J Hum Genet 2020 May;28(5):609-626 [FREE Full text] [doi: 10.1038/s41431-019-0560-9] [Medline: 31844175]
15. Gross MS, Miller RC. Ethical implementation of the learning healthcare system with blockchain technology. Blockchain Healthc Today 2019 Jun 07;2. [doi: 10.30953/bhty.v2.113]
16. Aguayo GA, Goetzinger C, Scibilia R, Fischer A, Seuring T, Tran VT, et al. Methods to generate innovative research ideas and improve patient and public involvement in modern epidemiological research: review, patient viewpoint, and guidelines for implementation of a digital cohort study. J Med Internet Res 2021 Dec 23;23(12):e25743 [FREE Full text] [doi: 10.2196/25743] [Medline: 34941554]
17. The James Lind Alliance guidebook. James Lind Alliance. 2021 Mar. URL: https://www.jla.nihr.ac.uk/jla-guidebook [accessed 2025-02-04]
18. Barony Sanchez RH, Bergeron-Drolet LA, Sasseville M, Gagnon MP. Engaging patients and citizens in digital health technology development through the virtual space. Front Med Technol 2022 Nov 25;4:958571 [FREE Full text] [doi: 10.3389/fmedt.2022.958571] [Medline: 36506474]
19. Antoine-LaVigne D, Hayes T, Fortenberry M, Ohikhuai E, Addison C, Mozee SJ, et al. Trust and biomedical research engagement of minority and under-represented communities in Mississippi, USA. Int J Environ Res Public Health 2023 Jan 16;20(2):1592 [FREE Full text] [doi: 10.3390/ijerph20021592] [Medline: 36674346]
20. Scharff DP, Mathews KJ, Jackson P, Hoffsuemmer J, Martin E, Edwards D. More than Tuskegee: understanding mistrust about research participation. J Health Care Poor Underserved 2010 Aug;21(3):879-897 [FREE Full text] [doi: 10.1353/hpu.0.0323] [Medline: 20693733]
21. Kass NE, Faden RR. Ethics and learning health care: the essential roles of engagement, transparency, and accountability. Learn Health Syst 2018 Oct 18;2(4):e10066 [FREE Full text] [doi: 10.1002/lrh2.10066] [Medline: 31245590]
22. Hood A, Macis M, Cervantes DM, Bear T, Khan J, Atkinson JM, et al. Privacy, policy, and profits: survey of patient values and preferences for research on de-identified biosamples. Advance Preprint posted online on January 30, 2024. [doi: 10.31124/advance.24155085.v1]
23. Gross M, Miller RC. Protecting privacy and promoting learning: blockchain and privacy preserving technology should inform new ethical guidelines for health data. Health Technol 2021 Aug 20;11(5):1165-1169. [doi: 10.1007/s12553-021-00589-9]
24. Gross MS, Hood AJ, Miller RCJ. Nonfungible tokens as a blockchain solution to ethical challenges for the secondary use of biospecimens: viewpoint. JMIR Bioinform Biotechnol 2021 Oct 22;2(1):e29905 [FREE Full text] [doi: 10.2196/29905] [Medline: 38943235]
25. Spinuzzi C. The methodology of participatory design. Tech Commun 2005 May;52(2):163-174 [FREE Full text] [doi: 10.4135/9781529696202.n4]
26. Cavagnuolo M, Capozza V, Matrella A. The walkthrough method: state of the art, innovative aspects, and application fields. In: Handbook of Research on Advanced Research Methodologies for a Digital Society. Hershey, PA: IGI Global Scientific Publishing; 2022.

XSL•FO
RenderX

27. Whimsical homepage. Whimsical. URL: https://whimsical.com/ [accessed 2025-02-21]
28. Adalo homepage. Adalo. URL: https://www.adalo.com/ [accessed 2025-02-21]
29. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
30. Flutter homepage. Flutter. URL: https://flutter.io/ [accessed 2025-02-21]
31. Sanchez W, Linder L, Miller RC, Hood A, Gross MS. Non-fungible tokens for organoids: decentralized biobanking to empower patients in biospecimen research. Blockchain Healthc Today 2024;7 [FREE Full text] [doi: 10.30953/bhty.v7.303] [Medline: 38715762]
32. Jones RD, Krenz C, Griffith KA, Spence R, Bradbury AR, De Vries R, et al. Governance of a learning health care system for oncology: patient recommendations. JCO Oncol Pract 2021 Apr;17(4):e479-e489. [doi: 10.1200/op.20.00454]
33. Peppercorn J, Campbell E, Isakoff S, Horick NK, Rabin J, Quain K, et al. Patient preferences for use of archived biospecimens from oncology trials when adequacy of informed consent is unclear. Oncologist 2020 Jan;25(1):78-86. [doi: 10.1634/theoncologist.2019-0365] [Medline: 31492767]
34. Sanderson SC, Brothers KB, Mercaldo ND, Clayton EW, Antommaria AH, Aufox SA, et al. Public attitudes toward consent and data sharing in biobank research: a large multi-site experimental survey in the US. Am J Hum Genet 2017 Mar 02;100(3):414-427 [FREE Full text] [doi: 10.1016/j.ajhg.2017.01.021] [Medline: 28190457]
35. Kaufman DJ, Murphy-Bollinger J, Scott J, Hudson KL. Public opinion about the importance of privacy in biobank research. Am J Hum Genet 2009 Nov;85(5):643-654 [FREE Full text] [doi: 10.1016/j.ajhg.2009.10.002] [Medline: 19878915]
36. D'Abramo F, Schildmann J, Vollmann J. Research participants' perceptions and views on consent for biobank research: a review of empirical data and ethical analysis. BMC Med Ethics 2015 Sep 09;16:60 [FREE Full text] [doi: 10.1186/s12910-015-0053-5] [Medline: 26354520]
37. State of mobile 2024. Sensor Tower. URL: https://sensortower.com/state-of-mobile-2024 [accessed 2025-02-04]
38. Fattah L, Johnson J, Clark U, Hess L, Palermo AG, Figueroa Acosta DM, et al. Building trust and transparency in biomedical sciences through data walks. Nat Med 2024 Aug 31;30(8):2103-2104. [doi: 10.1038/s41591-024-03042-2] [Medline: 38822130]
39. Coors ME, Westfall N, Zittleman L, Taylor M, Westfall JM. Translating biobank science into patient-centered language. Biopreserv Biobank 2018 Feb;16(1):59-63 [FREE Full text] [doi: 10.1089/bio.2017.0089] [Medline: 29190122]
40. Luger TM, Hamilton AB, True G. Measuring community-engaged research contexts, processes, and outcomes: a mapping review. Milbank Q 2020 Jun;98(2):493-553 [FREE Full text] [doi: 10.1111/1468-0009.12458] [Medline: 32428339]
41. Stires H, Bado I, Brown T, Carlson M, Chan IS, Echeverria GV, et al. Improving the odds together: a framework for breast cancer research scientists to include patient advocates in their research. NPJ Breast Cancer 2022 Jun 30;8(1):75 [FREE Full text] [doi: 10.1038/s41523-022-00440-y] [Medline: 35773258]
42. Klingler C, von Jagwitz-Biegnitz M, Baber R, Becker KF, Dahl E, Eibner C, et al. Stakeholder engagement to ensure the sustainability of biobanks: a survey of potential users of biobank services. Eur J Hum Genet 2022 Dec;30(12):1344-1354 [FREE Full text] [doi: 10.1038/s41431-021-00905-x] [Medline: 34031552]
43. Spector-Bagdady K, De Vries RG, Gornick MG, Shuman AG, Kardia S, Platt J. Encouraging participation and transparency in biobank research. Health Aff (Millwood) 2018 Aug;37(8):1313-1320 [FREE Full text] [doi: 10.1377/hlthaff.2018.0159] [Medline: 30080467]

Edited by A Kushniruk, E Borycki; submitted 12.04.24; peer-reviewed by R Rahimi, M Bagewadi Ellur, L Weinert; comments to author 13.09.24; revised version received 23.12.24; accepted 14.01.25; published 05.03.25.

Please cite as:
Dewan A, Eifler M, Hood A, Sanchez W, Gross M
Building a Decentralized Biobanking App for Research Transparency and Patient Engagement: Participatory Design Study
JMIR Hum Factors 2025;12:e59485
URL: https://humanfactors.jmir.org/2025/1/e59485
doi:10.2196/59485
PMID:40053747

©Ananya Dewan, M Eifler, Amelia Hood, William Sanchez, Marielle Gross. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 05.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Impact of Bottom-Up Cocreation of Nursing Technological Innovations: Explorative Interview Study Among Hospital Nurses and Managers

Saskia van Steenis[1], MSc; Onno Helder[2,3], PhD; Helianthe S M Kort[4,5], PhD; Thijs van Houwelingen[4], PhD

[1]Nursing Sciences, program in Clinical Health Sciences, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

[2]Department of Create4Care, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

[3]Research Centre Innovations in Care, Rotterdam University of Applied Sciences, Rotterdam, The Netherlands

[4]Research Group Technology for Healthcare Innovations, Research Centre for Healthy and Sustainable Living, University of Applied Sciences Utrecht, Utrecht, The Netherlands

[5]Building Healthy Environments for Future Users Group, Department of the Built Environment, Eindhoven University of Technology, Eindhoven, The Netherlands

Corresponding Author:
Thijs van Houwelingen, PhD
Research Group Technology for Healthcare Innovations, Research Centre for Healthy and Sustainable Living
University of Applied Sciences Utrecht
Heidelberglaan 7
Utrecht, 3584 CS
The Netherlands
Phone: 31 641097762
Email: thijs.vanhouwelingen@hu.nl

Abstract

Background: In health care, the use of nursing technological innovations, particularly technological products, is rapidly increasing; however, these innovations do not always align with nursing practice. An explanation for this issue could be that nursing technological innovations are developed and implemented with a top-down approach, which could subsequently limit the positive impact on practice. Cocreation with stakeholders such as nurses can help address this issue. Nowadays, health care centers increasingly encourage stakeholder participation, which is known as a bottom-up cocreation approach. However, little is known about the experience of nurses and their managers with this approach and the innovations it results in within the field of nursing care.

Objective: This study aims to explore nurses' and their managers' experiences with a bottom-up cocreation approach in order to assess the impact of this way of working and the resulting nursing technological innovations in an academic hospital. This insight can also inform decisions on whether the bottom-up cocreation approach should be more widely disseminated.

Methods: A qualitative study using semistructured interviews was conducted with 15 participants, including cocreator nurses, end-user nurses, and their managers. First, the data were thematically analyzed. In addition, a strengths, weaknesses, opportunities, and threats analysis was conducted.

Results: The various experiences of the participants were described in 3 main themes: enhanced attractiveness of the nursing profession, feeling involved due to a cocreation environment, and experienced benefits and challenges in using cocreated products. In addition, numerous strengths and opportunities perceived by the participants were identified as associated with the bottom-up cocreation approach and resulting useful products within nursing care; for example, cocreation contributed to job satisfaction and substantially contributed to the ease of use of the innovations that were developed.

Conclusions: The findings underscore that cocreation with nurses enhances the appeal of the nursing profession and aligns nursing technological innovations with practical nursing challenges. Embracing a culture of cocreation has the potential to foster a culture of continuous improvement and innovation in nursing care.

(JMIR Hum Factors 2025;12:e60543) doi:10.2196/60543

XSL•FO
RenderX

KEYWORDS

stakeholder participation; cocreation; nursing; innovation; bottom-up approach; diffusion of innovation; qualitative research; nurses

Introduction

Background

Nowadays, technological products and other innovations are being used more frequently in the health care sector [1-3] to minimize physical strain, such as lifting patients or transporting beds, and encourage ergonomic approaches during nursing tasks. Nurses need to use these innovations during patient care to manage various health care challenges, such as the increasing demand and complexity of health care [4,5]. An example of a technological product, which will be referred to as a "nursing technological innovation" in this paper, is the TrulyEasy clamp, designed for convenient 1-handed adjustment of the arterial line to the correct height [3]. Another example of a technological product is LineConnector, a product designed to efficiently organize infusion lines and prevent tangling [3].

Unfortunately, technological solutions aimed at solving nursing issues do not always align with daily nursing practice [2,6]. A potential explanation for this lies in the top-down approach used during the development and implementation of many innovations. Policy makers may prioritize regulatory and payment interests, while technicians focus on the functionality of innovations, without fully understanding the user perspectives and the nursing context [6-8]. Consequently, some solutions may inadequately align with the needs of the nursing profession, leading to limited impact on nursing practice [2,6]. Improving the adoption of innovations in nursing practice can be addressed in the development and implementation phase of these innovations through stakeholder participation [1,5,9-12].

A growing number of health care centers are facilitating an environment that allows stakeholders to participate in the development and implementation of nursing technological innovations [1,6,9,10,12]. This involvement, also called cocreation [13], with a bottom-up approach [8], involves an interaction in which various stakeholders work together to produce an outcome that is valued by all stakeholders involved [13]. The bottom-up approach allows stakeholders to share their knowledge or opinions about issues presented in daily activities and suggest potential solutions [8,14]. This approach can support the development of nursing technological innovations that may have an impact on nursing care [1,5,9-12]. FlushEgg is an example of a cocreated nursing technological innovation. Nurses reported experiencing challenges and discomfort while flushing the narrow central lines in neonates. Therefore, a syringe attachment was designed to enhance comfort in the palm during the application of pressure [3]. Another example of a cocreated nursing technological innovation is KoosGuard, which includes a holder that can be placed in the bed through which the cables and infusion lines are guided without becoming tangled or pinched [3].

However, the bottom-up approach requires nurses to invest time, which is challenging considering many nurses face time constraints [15,16]. If these nursing technological innovations lack the required quality, this could result in decreased application in nursing practice [2,6]. Furthermore, a poorly developed nursing technological innovation could lead to resistance from nurses, resulting in a negative attitude toward these innovations [9,17]. To fully use the potential of innovations in nursing practice, the innovations need to be adequately designed, which requires the involvement of nurses [9,18]. Despite the increasing involvement of nurses in the cocreation of nursing technological innovations, little is known about how this approach and its resulting innovations are experienced by nurses. Therefore, this study seeks to explore the experience resulting from the bottom-up approach of cocreation with nurses and the resulting nursing technological innovations developed through this approach. The perspectives of nurses, both as cocreators and end users, and their managers are key to understanding the impact of this collaborative process. These results could offer a deeper insight into how nurses, both as cocreators and end users, along with their managers, perceive the bottom-up cocreation approach and its resulting nursing technological innovations. A comprehensive understanding of their experiences, both positive and negative, can significantly contribute to the overarching goal of enhancing patient care and nursing practice, making it safer and more effective.

Aim

This study explores nurses' and their managers' experiences with a bottom-up cocreation approach in order to assess the impact of this way of working and the resulting nursing technological innovations in an academic hospital. This insight can also inform decisions on whether the bottom-up cocreation approach should be more widely disseminated.

Methods

Study Design

This study used an exploratory, descriptive, qualitative research design using individual interviews conducted at an academic hospital in the Netherlands between February and April 2023. This design facilitated the exploration of participants' experiences, which were further analyzed through a thematic analysis and subsequently through a strengths, weaknesses, opportunities, and threats (SWOT) analysis [19]. The methods and subsequent results were reported according to the Consolidated Criteria for Reporting Qualitative Research (COREQ) checklist [20].

Population and Setting

The study population consisted of three subgroups: (1) nurses with experience in cocreation, (2) nurses with experience as end users, and (3) health care managers from the nursing departments at an academic hospital where the innovations were implemented. In addition, all eligible nurses and managers were required to have at least 1 year of experience within the nursing department where the innovations were implemented. The 1-year experience criterion ensured that nurses and managers were

XSL•FO
RenderX

sufficiently familiar with the innovations and context. The setting comprised an innovation department established within the academic hospital.

Sampling Strategy and Recruitment

To gain more understanding and in-depth information from a broader perspective, a heterogeneous group of nurses and their health care managers were included through purposive sampling [19,21,22]. Potential nurse cocreators and managers for this study were recruited and informed via an email by OH about participating in the interviews. Interested individuals were required to actively contact SvS by email. Upon receiving an email, SvS responded with a participant information letter and an informed consent form. Once participants agreed to participate in the study, interviews were scheduled at a mutually convenient time. For the nurse end users, SvS visited the nursing departments where nursing technological innovations were integrated into nursing practice. Interested nurse end users were able to participate in an interview during their shifts at a time convenient for them. The nurse end users were informed via a participant information letter. After informed consent was given, the interviews started. The face-to-face interviews were conducted at the academic hospital.

Sample Size

A total of 14 to 20 participants were required to obtain saturation of knowledge with a heterogeneous sample group [19]. Therefore, the aim of this study was to include at least 14 participants [19]. The study involved 15 participants, including 10 (67%) women and 5 (33%) men.

Data Collection

Data were collected through semistructured interviews using an interview guide based on literature and expert opinions [19]. The interview guide covered various topics (Table 1), including the participants' attitude [23] toward technology, the bottom-up cocreation approach, and the nursing technological innovations developed through this approach. It also explored the acceptance [9,10,23], usability [12,23], and adoption [10,23] of these innovations in nursing care as well as their impact on patient care and nursing practice [1,11,12,24]. Demographic characteristics (age, gender, and specialization), along with the adoption rate, provided an overview and background information of the participants in this study. The adoption rate proposed by Rogers [25] allowed the measurement of individuals' attitudes toward innovations, providing valuable insights into how the participants generally perceive and approach innovations. The adoption rate according to Rogers [25] was measured by asking the participants which category they identified with. They could choose one of the following five categories: (1) driven by change and introduction of innovations, (2) leading in adopting innovations, (3) deliberately adopting innovations, (4) adoption is an economic necessity and a response to social peer pressure, and (5) have a traditional view and are more skeptical about innovations [25]. A pilot test was conducted to refine the interview guide and involved expert assessment and field testing with one participant. After the interviews, the first author created memos recording thoughts that arose due to the interviews [19]. The interviews were audio recorded and conducted by SvS, a nurse with no prior experience at the hospital where the research took place. SvS was trained in conducting interviews. Throughout the data collection process, discussions were held with the research team, including investigators SvS, TvH, and OH, to refine the interview questions and ensure that they effectively addressed the research question.

XSL•FO
RenderX

Table 1. Interview guide topics and prompt questions.

Topics	Prompt questions for nurses (end users)	Prompt questions for cocreator nurses	Prompt questions for managers
Attitude	• What are your impressions regarding the fact that a coworker from your department contributed to this nursing technological innovation? • What is your opinion on dedicating time to enable nurses to create and execute nursing technological innovations for nursing care?	• Can you elaborate on how the cocreation process began and how it evolved? • What is your opinion on dedicating time to enable nurses to create and execute nursing technological innovations for nursing care?	• What are your thoughts on the fact that a nurse from the nursing department has made a substantial contribution to this nursing technological innovation? • What is your opinion on dedicating time to enable nurses to create and execute nursing technological innovations for nursing care?
Acceptance	• What were your initial impressions of the nursing technological innovation when you first began working with it?	• What were your initial impressions of the nursing technological innovation when you first began working with it?	• What responses did you observe from nurses when they were introduced to the nursing technological innovation for the first time?
Adoption	• In what ways do you integrate the nursing technological innovation into your daily nursing tasks and responsibilities?	• In what ways do you integrate the nursing technological innovation into your daily nursing tasks and responsibilities?	• What are the viewpoints within the department regarding the utilization of the nursing technological innovation?
Usability	• What is your experience with the nursing technological innovation at work?	• What is your experience with the nursing technological innovation at work?	• What are the experiences of the nurses regarding the use of the nursing technological innovation?
Nursing practice	• What transformations do you observe in nursing practice as a result of the application of the nursing technological innovation?	• What transformations do you observe in nursing practice as a result of the application of the nursing technological innovation?	• What transformations do you observe in nursing practice as a result of the application of the nursing technological innovation?
Patient care	• What changes do you perceive in patient outcomes as a result of the application of the nursing technological innovation?	• What changes do you perceive in patient outcomes as a result of the application of the nursing technological innovation?	• What changes do you perceive in patient outcomes as a result of the application of the nursing technological innovation?

Data Analysis

The semistructured interviews were transcribed verbatim [19]. A 6-step thematic analysis approach (Table 2) was applied using ATLAS.ti software (version 23; Lumivero) [26]. Open coding and constant comparison techniques were used during the analysis [19]. An iterative process was used between data collection and data analysis. Refining the interview questions made the collection of the data more focused and specific as the process developed [19]. Audio recordings were transcribed, read, and reread by SvS in order to become familiar with the data. Data were initially coded using open coding by SvS. The 1st, 3rd, 6th, and 10th transcripts were coded separately, both by SvS and TvH. The initial codes of SvS and TvH were compared and discussed until consensus was reached. Main themes were generated through an iterative process, including constant comparison, and further elaborated and finalized by SvS, TvH, and OH. A codebook was used for insights into coding and code development [19].

XSL•FO
RenderX

Table 2. Thematic analysis.

Phase	Explanation
Phase 1: familiarization with the data	SvS conducted interviews with the participants and transcribed them. Transcripts were read and reread by SvS to familiarize herself with the data and obtain an overall impression.
Phase 2: generation of initial codes	Data were extracted from the transcripts and initially coded using open coding by SvS. The 1st, 3rd, 6th, and 10th transcripts were coded separately, both by SvS and TvH. The initial codes of SvS and TvH were compared and discussed until consensus was reached. The memos were processed systematically. Codes remained linked to the transcriptions (quotes) and the data extracts, using ATLAS.ti (Lumivero).
Phase 3: broader levels of themes	SvS developed broad potential main themes by comparing the initial codes. During meetings with SvS, TvH, and OH, consensus was reached, resulting in potential main themes derived from the data. During follow-up meetings with SvS, TvH, and OH, a description of the potential main themes was created and deliberated upon.
Phase 4: reviewing and refining the candidate themes	SvS compared the potential main themes with the coded data extracts and the raw data (transcriptions), discussing inconsistencies with TvH and OH. This process led to the identification of the main themes, which were visualized with a thematic map. The themes were refined by SvS and peer reviewed by TvH and OH.
Phase 5: defining and naming themes	SvS revisited the data extracts to define the content of each main theme. In meetings, SvS, TvH, and OH discussed the essence of each theme and developed the overall story line, resulting in the final naming of the main themes.
Phase 6: producing the report	SvS drafted the report and included detailed examples and quotes supporting the main themes. TvH and OH provided feedback, and adjustments were made accordingly. Final recommendations were implemented in the report.

In addition to the thematic analysis, we performed a secondary analysis using the SWOT matrix. This secondary analysis was carried out to make the data more accessible for the readers and support strategic changes within an organization. The SWOT analysis has been shown to be a valid method in previous studies for facilitating data interpretation for readers and informing strategic planning and decision-making in practice [27,28]. In several meetings, detailed examples of the results, with related quotes of the main themes, were discussed and analyzed and subsequently linked to a component of the SWOT matrix.

Rigor and Reflexivity

Different techniques were used to enhance trustworthiness [19]. The credibility of this study was established by the investigators by creating a nonjudgmental ambiance during the semistructured interviews to obtain the participants' perspective. The chance for bias was reduced by transcribing the content of the interviews verbatim. Member checking was used to ensure the accuracy of the transcripts, with participants reviewing the transcripts to validate their thoughts and ideas [19]. The performed member check resulted in minimal textual changes to the data.

Memos were reported during data collection and analysis to support the investigation process and the thoughts that had occurred to the investigator. The conformability of the interpretation and credibility of the data were both enhanced by investigator triangulation during the data analysis and peer feedback during researcher meetings, including the investigators SvS, TvH, OH, and HSMK. Due to peer feedback, potential meanings and a wider range of perspectives were revealed. Reflexivity, ensured by the investigator's critical view of the interview process and comments from other investigators, increased depth and improved accuracy.

A thick description was pursued to ensure the transferability of this study through the diversity of the population base, the number of participants, and the length of the interviews for imitability. To enhance the reflexivity and transparency of the investigators, a logbook was kept, detailing all changes and decisions made throughout the process of this study [19].

Ethical Considerations

Throughout this study, the principles of the Declaration of Helsinki and the code of conduct for health research were followed [29,30]. The study was examined and approved by the medical ethics review committee of the Erasmus Medical Center for a nonmedical scientific research application (MEC-2023-0021). Participants were informed and invited to the study by OH. They received the investigator's email and had 1 week to express interest. Nonresponders received a reminder. The investigator sent a participant information letter and informed consent upon contact. Participants could ask further questions via email. Interviews were scheduled after participants confirmed their approval. Before the interview, they provided written informed consent. The participants did not receive any compensation for their participation in the study. During the study, data collection was conducted with confidentiality, and data were processed anonymously according to the European Union General Data Protection Regulation [29]. Only the investigators SvS, TvH, OH, and HSMK had access to the source data.

Results

Characteristics of the Participants

The duration of the interviews ranged from 26 to 62 minutes, with an average duration of 44 (SD 10.8) minutes. Participants' ages ranged from 20 to 67 years. Various departments, including the intensive care unit, children's ward, and the recovery department, were represented in the study. Table 3 presents detailed demographic information, including gender, age range, department, adoption rate, number of nurses participating as cocreators and end users, and the number of health care managers.

XSL•FO
RenderX

Table 3. Demographic data of the participants (N=15).

Characteristics	Participants, n (%)
Gender	
Women	10 (67)
Men	5 (33)
Age range (y)	
20-29	1 (7)
30-39	5 (33)
40-49	3 (20)
50-59	3 (20)
60-67	3 (20)
Department	
ICU[a] (adults)	7 (47)
ICU (thoracic surgery and cardiac monitoring)	2 (13)
Center for home ventilator and respiratory disorders in children	2 (13)
ICU (children)	1 (7)
Recovery	1 (7)
Day treatment (children)	2 (13)
Adoption rate[b]	
Innovator[c]	10 (67)
Early adopters[d]	3 (20)
Early majority[e]	2 (13)
Late majority[f]	0 (0)
Laggards[g]	0 (0)
Profession	
Nurse (cocreator)	6 (40)
Nurse (end user)	5 (33)
Health care manager	4 (27)

[a]ICU: intensive care unit.

[b]Adoption rate was measured by asking the participants which category they identified with.

[c]Driven by change and introduction of innovations.

[d]Leading in adopting innovations.

[e]Deliberately adopting innovations.

[f]Adoption is an economic necessity and a response to social peer pressure.

[g]Laggards have a traditional view and are more skeptical about innovations [26].

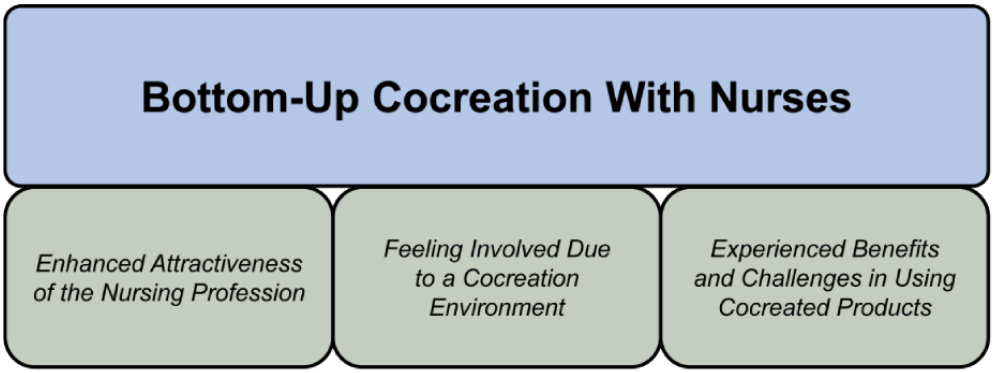
Bottom-Up Cocreation With Nurses

Overview

The thematic analysis uncovered the following three main themes concerning the participants' experiences with the bottom-up cocreation approach and resulting nursing technological innovations: (1) enhanced attractiveness of the nursing profession, (2) feeling involved due to a cocreation environment, and (3) experienced benefits and challenges in using cocreated products. Figure 1 provides a visual representation of the main themes.

XSL•FO
RenderX

Figure 1. Thematic map with the 3 main themes.



Enhanced Attractiveness of the Nursing Profession

All participants experienced that the bottom-up cocreation approach with nurses contributed to the overall attractiveness of the nursing profession. The nursing profession's attractiveness was enhanced as participants were enthusiastic about participating in cocreation, finding it enjoyable and engaging. Moreover, some participants felt that they were taken more seriously as nurses as they were the ones who actually experienced the problems in practice and were given an opportunity to use their expertise to address these challenges and drive improvements in nursing practice:

> *It makes work more enjoyable. You feel taken seriously as a nurse. You can share your ideas. You can use your expertise to improve something. So I think that's a lot of fun and very inspiring, and it also leads to something useful in the end.* [P11, cocreator]

Some participants emphasized the importance of having affinity with cocreation. Most participants felt a sense of pride in having a cocreation department within the organization and being a part of a cocreation process. A few participants expressed modesty about contributing to the cocreation product:

> *Well, I think we were genuinely proud to be able to offer the product to the patient. It was a joint effort and of course it was also very important for us that our name was set out on the product. So another team could not say, "Look at this!" It really makes you feel good if you, if you can make something that you really like to use, then that is actually really nice.* [P8, end user]

In addition, nurses were more attracted to the profession because they assumed multiple roles during cocreation. They acted as developers among designers, students, and the nursing department, actively contributing ideas and promoting products within the departments. Some participants expressed that the opportunity to become involved in cocreation made their work more interesting, and, due to this, they did not look for employment elsewhere. Cocreation with nurses was also inspiring for some participants. It motivated other nurses to address practical problems encountered during their work and explore possibilities for solving them:

> *But you're then involved in care in just a slightly different way than in standard care. And that also gives new energy and new ideas. And maybe other colleagues could also get new ideas and think, well, I have seen it a few times now, so what if I can come up with some ideas like this. Then they might think, oh well, perhaps this is something that we can also work on. This could also inspire others.* [P13, cocreator]

Feeling Involved Due to a Cocreation Environment

All participants reported enhanced collaboration through bottom-up cocreation involving various stakeholders. Collaboration included clear communication and teamwork among nurses, designers, and technically oriented students. Most participants mentioned that collaboration within the organization increased their willingness to adopt certain products that were aligned with nursing practice. One participant preferred internal product development but acknowledged that external innovations could still benefit end users:

> *If you get everyone involved, then you will have much more support and that is also the case with new things. I have experienced a few times in my life that we were getting an email on Sunday night. "From Monday we will do it like this." Well, then nurses will be digging in their heels. Whereas if you say, we're working on this, and who wants to contribute their views? Then you create support; much more people will then be willing to take it from there.* [P9, cocreator]

> *I like it when the product comes from us, I like that, but it does not necessarily have to come from us.* [P8, end user]

Some participants highlighted that collaboration among nursing departments, other hospitals, and organizations contributed to knowledge exchange, product distribution, and external production. However, one participant mentioned poor communication among nursing departments concerning new developments within the organization:

> *But we, we do note things, and we share them. It's important that you know this. We're not all working in separate groups.... On the contrary, a lot of information is shared among the four of us. Because we have to intervene and know what the other thinks.* [P10, manager]

XSL•FO
RenderX

> *No, that is also very poorly communicated. And the same applies to these products.* [P9, cocreator]

Some participants mentioned encountering practical problems but found it challenging to address these problems without the support of an innovation department with a cocreation group. The presence of such a department allowed the active participation of nurses and patients in creating specific products. Furthermore, a subgroup of participants engaged in cocreation with both patients and their parents, resulting in products tailored to their specific desires and needs. Patients were enthusiastic and felt heard, leading to the active use of end-user products:

> *Yes, parents have been involved in the co-creation process. That is, in the entire process of designing, they were not just allowed to give input, but were also able make direct choices, about what benefits them the most.* [P12, manager]

An aspect experienced by most participants was the need for support during the cocreation process. Several participants emphasized the importance of a supportive manager in facilitating and allocating dedicated time for the cocreation process. They noted that cocreation was unlikely to occur without managerial backing. Participants reported that supportive managers who recognized the value of cocreation and its potential to save time in nursing care were more proactive in addressing and resolving practical problems:

> *Managers can stimulate it. I think that people working in healthcare can be very creative, but often think, what shall I say, in boxes, because it's not possible, because we're too busy. But actually that is such a shame, because we just have a lot to offer and we have the knowledge and experience to be able to create something beautiful. And I think if a manager who (A) stimulates and (B) gives the space that is required, this will ensure that people are more motivated to contribute. Because they're aware that improvement is possible.* [P11, cocreator]

Experienced Benefits and Challenges in Using Cocreated Products

Benefits

Participants found the products resulting from the bottom-up approach to be highly practical and user-friendly when used during their nursing tasks. All participants expressed immediate positivity toward the product due to its straightforward design. They could use the product without needing special training. Nurses also immediately experienced positive outcomes during their tasks. For instance, one participant mentioned that a newly created clamp for placing pressure transducers at the correct height on infusion poles improved nursing efficiency. They also reported that this clamp alleviated wrist complaints:

> *It is faster, yes, you press it and put it on the pole, loosen it and put it back on the transport cart.* [P6 end user]

Some participants found that the products improved their ability to organize and manage nursing tasks. For example, a product named KoosGuard contributed to an improved organization of patient monitoring wires, providing enhanced structure during nursing care. Another product, the cannula emergency bag, ensured that nurses, patients, and parents had all the materials neatly arranged in case of an emergency, enhancing preparedness and efficiency:

> *Overview yes, you immediately have a clear overview. Because it's easy to open it, everything is in place and it's compact. And because it was developed together with the parents and the nurses, of course, they really like that. It fits their needs.* [P12, manager]

Several participants emphasized the importance of the products' quality and safety. Participants highlighted that the professional appearance of the product and its suitability for nursing practice served as indicators of quality. Participants reported that these well-suited products reduced the need for makeshift solutions, resulting in fewer inconvenient situations during their work:

> *Yes, the well-known example is that we often use band-aids and things, but that is to help you out for the time being. But practice has already shown that there are much more convenient solutions.* [P11, cocreator]

Certain participants emphasized the role of the products in enhancing patient safety. Participants noted that a specific product effectively prevented infusion line entanglement and tension, thus reducing the risk of accidental removal. In addition, participants mentioned a product designed to decrease the risk of postsurgery wound contamination, promoting improved healing outcomes:

> *You really want those lines to be properly arranged. To make sure that not one of them is hidden from your view, because if one of those lines is pulled, the infusion may accidentally get disconnected or the patient can accidentally end up lying on top of that line because you can't see it.* [P14, manager]

Challenges

In addition to the benefits of the products, participants also experienced some challenges in their use. Some participants mentioned that some product designs had flaws, which hindered their effectiveness. For example, a participant heard from some patients that they still thought that the cannula emergency bag was too large to be used conveniently, and, due to this, it was no longer used. In addition, it was indicated that some products looked disposable, so nurses threw them away:

> *The only negative feedback came from a patient. She just thought the bag was too big.* [P8, end user]

A few participants pointed out that some products were prone to getting lost or becoming untraceable. It was emphasized that these products should be readily available to nurses at designated and standardized locations. Otherwise, the product will be used less or not used at all. Furthermore, some participants shared their struggle with making changes, leading to a heightened risk of reverting to old routines:

> *Then it takes a lot of time anyway. And you think: well, I'll just do it in the old-fashioned way, using a gauze.* [P3, end user]

XSL•FO
RenderX

SWOT Analysis Findings

Detailed examples of the main themes were divided into the 4 components of the SWOT matrix. Multimedia Appendix 1 presents an overview of all the main themes, along with detailed examples, illustrative quotations, and the corresponding SWOT categorization. The “strengths” and “weaknesses” encompass the experienced advantages and disadvantages associated with the bottom-up cocreation approach and the products developed through this approach. The “opportunities” and “threats” are examples that could be associated with potential future opportunities and challenges of the bottom-up cocreation approach and developed products. Most of the examples were connected to the components “strengths” and “opportunities.”

Discussion

Principal Findings

In this study, we sought to explore hospital nurses’ experience with a bottom-up cocreation approach and the resulting nursing technological innovations developed through this approach. Notably, through a SWOT matrix, we mainly revealed positive experiences of participants cocreating nursing technological innovations, with several strengths and opportunities. Participants’ experiences can be described in 3 themes: enhanced attractiveness of the nursing profession, feeling involved due to a cocreation environment, and experienced benefits and challenges in using cocreated products. The findings indicate that participants found the nursing technological innovations easily applicable in nursing practice due to their user-friendly design, the ability to enhance nursing efficiency, patient safety, and time-saving potential during nursing work. A collaborative culture may lead to higher work satisfaction among nurses, inspire cocreation, and promote nursing pride.

Comparison With Prior Work

This study shows how cocreating nursing technological innovations positively impacts the attractiveness of the nursing profession. The attractiveness of the nursing profession was enhanced by nurses finding joy in their roles, feeling a sense of pride, and being taken seriously in their work. This aligns with findings from a cross-sectional study, which emphasizes the significance of decision-making involvement in health care services, such as equipment, technologies, and processes for employment preferences of nursing students [31]. In addition, it aligns with a systematic review indicating that an increase in career opportunities and challenges positively impacts nurse retention. It reduces nurses’ desire to leave a particular workplace [32].

Furthermore, this study emphasizes several collaboration advances due to a cocreation environment within the organization, such as clear communication and teamwork among nurses, designers, and technically oriented students. Moreover, a cocreation environment within the organization contributed to knowledge exchange, product distribution, and external production. An in-depth case study [14] reinforced the importance of an innovative environment within the organization. It could increase the diffusion of nurse-led innovations within and outside the organization. Particularly, managers seem to have an important role in collaborating and communicating with each other about innovations and supporting working methods in nursing departments. Existing literature [33,34] reinforces the notion that supportive leadership is crucial for fostering a culture of innovation and enhancing the sustainability and impact of health care innovation settings.

In addition, this study shows that innovations cocreated with nurses led to multiple benefits for nursing practice. Cocreated nursing technological innovations were described as user-friendly, highly practical, and easy to integrate into nursing tasks, consistent with findings from a quasi-experimental study [35]. In this quasi-experimental study, trained innovative nurses were able to identify issues in nursing practice and apply creative thinking strategies to create innovative products that matched their nursing work routines. Furthermore, the cocreated nursing technological innovations met nursing quality requirements and reduced the need for improvised solutions, thereby minimizing inconvenient situations in nursing practice. Patient safety was also enhanced by the cocreated products, which prevented issues such as infusion line entanglement and reduced the risk of wound contamination. The incubator traffic light is an example from the literature of a cocreated nursing technological innovation that meets quality standards and enhances patient safety. The incubator traffic light features a visual feedback system for neonatal incubators designed to enhance hand hygiene compliance [1,2]. Incubators with the incubator traffic light visual feedback system demonstrated significantly higher compliance with the correct drying times compared to those without this feature. This could potentially enhance patient safety by reducing infection rates in neonatal intensive care units [1]. Although the cocreated nursing technological innovations offered several advantages, the study also encountered certain challenges in their use, for example, flaws in the design of the innovations, which hindered their effectiveness or ability to make changes in the nursing routines. This study shows that changes in nursing processes may increase the risk of reverting to old routines, which aligns with findings from a cross-sectional survey that indicated how changes in nursing work processes could influence established work habits [36].

While this study did not include patients or parents directly, the participants acknowledged the significance of their perspective during the innovation process. It was mentioned that some of the cocreated innovations were developed in collaboration with patients and parents, resulting in innovations aligning with their expectations and needs. This is consistent with existing literature, where patients, family caregivers, and clinicians cocreated a mobile health app for heart failure self-management tailored to the specific health care context [37]. Furthermore, it matches with a qualitative co-design article including family, physicians, researchers, patients, and industry partners. The study revealed important differences between the participants’ preference for functional requirements of a mobile health app [38].

Limitations

Limitations of the study include its single-site focus on a specific academic hospital and limited transferability to other health

XSL•FO
RenderX

care settings [19]. Furthermore, there may be a sampling bias, as interested individuals were required to actively contact the investigator via email to participate in this study. These individuals might have been more likely than other professionals to embrace the innovations and encourage their adoption within their teams. In addition, all the participants scored themselves above average on the adoption rate scale (Table 3), and 10 (67%) of the 15 participants considered themselves to be "innovators" [25]. Literature shows that only 2.5% of the population belongs to the group of innovators [25]. This suggests a potential bias in the participant group. Including nurses as cocreators may have introduced selection bias, as their enthusiasm and assertiveness toward innovations may have influenced the results in a more positive manner [39]. Finally, the sample size consisted of 6 (40%) nurses as cocreators, 5 (33%) nurses as end users, and 4 (27%) managers. The limited insights from the managers' perspective may affect the comprehensiveness of the findings in the study [19].

Implications and Further Research

The findings of this study suggest several implications. Health care organizations could consider adopting the bottom-up cocreation approach with nurses as it has the potential to enhance nurses' work satisfaction, a sense of pride among nurses, and adoptable nursing technological innovations in nursing care. Furthermore, cocreation can contribute to accessible collaborations between stakeholders inside and outside the organization. In addition, cocreation with nurses has the potential to facilitate the development of innovations that enhance the efficiency of nursing work. Simple, professional, and time-saving innovations created through this approach could improve nursing practices.

Further research is recommended to explore the impact of the bottom-up cocreation approach and the resulting nursing technological innovations in other health care centers, aiming to improve the transferability of this way of working. Moreover, to better address patients' expectations and needs during the cocreation process and resulting innovations, further research is recommended to investigate their experiences throughout the cocreation process and its outcomes.

Conclusions

This study highlights substantial positive experiences of cocreating nursing technological innovations in nursing care. The findings underscore that cocreation with nurses enhances the appeal of the nursing profession. Participants perceived cocreation as enjoyable, leading to heightened work satisfaction and a sense of pride in their nursing role. Moreover, cocreation with nurses fosters the development of nursing technological innovations that align with nursing practice challenges, thereby facilitating their adoption within nursing care. Ultimately, embracing a culture of cocreation has the potential to foster ongoing improvements and innovations in nursing care, further contributing to the professionalization of the field.

Acknowledgments

The authors extend their appreciation to Myrthe van der Zanden, MSc, nursing scientist, for reviewing the interview guide and participating in the pilot interview.

Data Availability

The datasets generated or analyzed during this study are not publicly available due the participants involved in this research did not authorize the public dissemination of their data through written consent.

Authors' Contributions

SvS contributed to conceptualization, writing the original draft, data curation, formal analysis, investigation, methodology, project administration, and visualization. OH contributed to recruitment, conceptualization, reviewing and editing the manuscript, formal analysis, investigation, methodology, project administration, and supervision. HSMK contributed to conceptualization, methodology, and reviewing and editing the manuscript. TvH contributed to conceptualization, reviewing and editing the manuscript, data curation, formal analysis, investigation, methodology, project administration, and supervision.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Main themes; detailed examples; illustrative quotes; and strengths, weaknesses, opportunities, and threats categorization.
[DOCX File , 23 KB - humanfactors_v12i1e60543_app1.docx]

References

1. van Gils RH, Helder OK, Wauben LS. Incubator traffic light: the development of an alcohol-based hand rub dispenser system for neonatal incubators with visual feedback to improve hand hygiene compliance. BMJ Innov 2019 Sep 06;5(2-3):70-77. [doi: 10.1136/bmjinnov-2018-000301]

XSL•FO
RenderX

2. van Gils RH, Kornelisse RF, Dankelman J, Helder OK. Validation of a hand hygiene visual feedback system to improve compliance with drying time of alcohol-based hand rub in a neonatal intensive care unit: the incubator traffic light system. J Hosp Infect 2024 Mar;145:210-217 [FREE Full text] [doi: 10.1016/j.jhin.2024.01.007] [Medline: 38272126]
3. van der Zanden M, Helder O, Westland H, Ista E. Barriers and facilitators for the implementation of co-created nursing innovations in a Dutch university hospital. JBI Evid Implement 2024 Jun 28;22:11. [doi: 10.1097/XEB.0000000000000444] [Medline: 38932713]
4. Rudnicka E, Napierała P, Podfigurna A, Męczekalski B, Smolarczyk R, Grymowicz M. The World Health Organization (WHO) approach to healthy ageing. Maturitas 2020 Sep;139:6-11 [FREE Full text] [doi: 10.1016/j.maturitas.2020.05.018] [Medline: 32747042]
5. Dugstad J, Eide T, Nilsen ER, Eide H. Towards successful digital transformation through co-creation: a longitudinal study of a four-year implementation of digital monitoring technology in residential care for persons with dementia. BMC Health Serv Res 2019 Jun 10;19(1):366 [FREE Full text] [doi: 10.1186/s12913-019-4191-1] [Medline: 31182093]
6. de Jong JP, Rigtering C, Spaans L. Heroes of diffusion: making user innovations widely available. Res Policy 2023 Oct;52(8):104840. [doi: 10.1016/j.respol.2023.104840]
7. Wang S, Bolling K, Mao W, Reichstadt J, Jeste D, Kim HC, et al. Technology to support aging in place: older adults' perspectives. Healthcare (Basel) 2019 Apr 10;7(2):60 [FREE Full text] [doi: 10.3390/healthcare7020060] [Medline: 30974780]
8. Patey AM, Soong C. Top-down and bottom-up approaches to low-value care. BMJ Qual Saf 2023 Feb 14;32(2):65-68. [doi: 10.1136/bmjqs-2022-014977] [Medline: 36517225]
9. Kernebeck S, Jux C, Busse TS, Meyer D, Dreier LA, Zenz D, et al. Participatory design of a medication module in an electronic medical record for paediatric palliative care: a think-aloud approach with nurses and physicians. Children (Basel) 2022 Jan 06;9(1):82 [FREE Full text] [doi: 10.3390/children9010082] [Medline: 35053707]
10. Kernebeck S, Busse TS, Jux C, Dreier LA, Meyer D, Zenz D, et al. Evaluation of an electronic medical record module for nursing documentation in paediatric palliative care: involvement of nurses with a think-aloud approach. Int J Environ Res Public Health 2022 Mar 18;19(6):3637 [FREE Full text] [doi: 10.3390/ijerph19063637] [Medline: 35329323]
11. Grynne A, Browall M, Fristedt S, Ahlberg K, Smith F. Integrating perspectives of patients, healthcare professionals, system developers and academics in the co-design of a digital information tool. PLoS One 2021;16(7):e0253448 [FREE Full text] [doi: 10.1371/journal.pone.0253448] [Medline: 34270564]
12. Bryant L, Sedlarevic N, Stubbs P, Bailey B, Nguyen V, Bluff A, et al. Collaborative co-design and evaluation of an immersive virtual reality application prototype for communication rehabilitation (DISCOVR prototype). Disabil Rehabil Assist Technol 2024 Jan;19(1):90-99. [doi: 10.1080/17483107.2022.2063423] [Medline: 35442823]
13. Prahalad C, Ramaswamy V. Co-creation experiences: the next practice in value creation. J Interact Mark 2004 Aug 01;18(3):5-14. [doi: 10.1002/dir.20015]
14. Rigtering C, Spaans LJ, de Jong JP. How to bridge the nurse innovation-diffusion gap? An in-depth case study of Create4Care. Front Public Health 2023;11:1209965 [FREE Full text] [doi: 10.3389/fpubh.2023.1209965] [Medline: 37601181]
15. McDermid F, Mannix J, Peters K. Factors contributing to high turnover rates of emergency nurses: a review of the literature. Aust Crit Care 2020 Jul;33(4):390-396. [doi: 10.1016/j.aucc.2019.09.002] [Medline: 31836449]
16. Racy S, Davidson PM, Peeler A, Hager DN, Street L, Koirala B. A review of inpatient nursing workload measures. J Clin Nurs 2021 Jul;30(13-14):1799-1809. [doi: 10.1111/jocn.15676] [Medline: 33503306]
17. Ammenwerth E, Iller C, Mahler C. IT-adoption and the interaction of task, technology and individuals: a fit framework and a case study. BMC Med Inform Decis Mak 2006 Jan 09;6:3 [FREE Full text] [doi: 10.1186/1472-6947-6-3] [Medline: 16401336]
18. Devi DH, Duraisamy K, Armghan A, Alsharari M, Aliqab K, Sorathiya V, et al. 5G technology in healthcare and wearable devices: a review. Sensors (Basel) 2023 Feb 24;23(5):2519 [FREE Full text] [doi: 10.3390/s23052519] [Medline: 36904721]
19. Holloway I, Galvin L. Qualitative Research in Nursing and Healthcare, Fourth Edition. Hoboken, NJ: John Wiley & Sons; 2016.
20. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
21. Boeije HR. Analysis in Qualitative Research. Thousand Oaks, CA: SAGE Publications; 2010.
22. Creswell JW, Poth CN. Qualitative Inquiry and Research Design: Choosing Among Five Approaches. Thousand Oaks, CA: SAGE Publications; 2024.
23. Brandsma T, Stoffers J, Schrijver I. Advanced technology use by care professionals. Int J Environ Res Public Health 2020 Jan 23;17(3):742 [FREE Full text] [doi: 10.3390/ijerph17030742] [Medline: 31979307]
24. Farley HL, Baumlin KM, Hamedani AG, Cheung DS, Edwards MR, Fuller DC, et al. Quality and safety implications of emergency department information systems. Ann Emerg Med 2013 Oct;62(4):399-407 [FREE Full text] [doi: 10.1016/j.annemergmed.2013.05.019] [Medline: 23796627]
25. Rogers EM. Diffusion of Innovations, Fourth Edition. New York, NY: Free Press; 2010.

XSL•FO
RenderX

26. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
27. Baez-Leon C, Palacios-Ceña D, Fernandez-de-Las-Peñas C, Velarde-García JF, Rodríguez-Martínez M, Arribas-Cobo P. A qualitative study on a novel peer collaboration care programme during the first COVID-19 outbreak: a SWOT analysis. Nurs Open 2022 Jan;9(1):765-774 [FREE Full text] [doi: 10.1002/nop2.1128] [Medline: 34773372]
28. Topuz Ş, Yilmaz Sezer N, Aker MN, Gönenç IM, Öner Cengiz H, Er Korucu A. A SWOT analysis of the opinions of midwifery students about distance education during the Covid-19 pandemic a qualitative study. Midwifery 2021 Dec;103:103161 [FREE Full text] [doi: 10.1016/j.midw.2021.103161] [Medline: 34628182]
29. Regulation (EU) 2016/679 of the European parliament and of the council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). Official Journal of the European Union. 2016 May 4. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679 [accessed 2025-03-15]
30. Code of conduct for health research. Central Committee on Research Involving Human Subjects. URL: https://english.ccmo.nl/investigators/legal-framework-for-medical-scientific-research/codes-of-conduct/code-of-conduct-for-health-research [accessed 2024-03-25]
31. Terry D, Peck B, Baker E, Schmitz D. The rural nursing workforce hierarchy of needs: decision-making concerning future rural healthcare employment. Healthcare (Basel) 2021 Sep 18;9(9):1232 [FREE Full text] [doi: 10.3390/healthcare9091232] [Medline: 34575006]
32. de Vries N, Boone A, Godderis L, Bouman J, Szemik S, Matranga D, et al. The race to retain healthcare workers: a systematic review on factors that impact retention of nurses and physicians in hospitals. Inquiry 2023;60:469580231159318 [FREE Full text] [doi: 10.1177/00469580231159318] [Medline: 36912131]
33. Albert NM. Operationalizing a nursing innovation center within a health care system. Nurs Adm Q 2018;42(1):43-53. [doi: 10.1097/NAQ.0000000000000266] [Medline: 29194332]
34. Barr TL, Malloch K, Ackerman MH, Raderstorf T, Melnyk BM. A blueprint for nursing innovation centers. Nurs Outlook 2021;69(6):969-981 [FREE Full text] [doi: 10.1016/j.outlook.2021.05.006] [Medline: 34183188]
35. Gao L, Lu Q, Hou X, Ou J, Wang M. Effectiveness of a nursing innovation workshop at enhancing nurses' innovation abilities: a quasi-experimental study. Nurs Open 2022 Jan;9(1):418-427 [FREE Full text] [doi: 10.1002/nop2.1080] [Medline: 34687153]
36. Alsohime F, Temsah MH, Al-Eyadhy A, Bashiri FA, Househ M, Jamal A, et al. Satisfaction and perceived usefulness with newly-implemented electronic health records system among pediatricians at a university hospital. Comput Methods Programs Biomed 2019 Feb;169:51-57. [doi: 10.1016/j.cmpb.2018.12.026] [Medline: 30638591]
37. Woods L, Duff J, Roehrer E, Walker K, Cummings E. Design of a consumer mobile health app for heart failure: findings from the nurse-led co-design of Care4myHeart. JMIR Nurs 2019;2(1):e14633. [doi: 10.2196/14633] [Medline: 34345774]
38. Mrklas KJ, Barber T, Campbell-Scherer D, Green LA, Li LC, Marlett N, et al. Co-design in the development of a mobile health app for the management of knee osteoarthritis by patients and physicians: qualitative study. JMIR Mhealth Uhealth 2020 Jul 10;8(7):e17893 [FREE Full text] [doi: 10.2196/17893] [Medline: 32673245]
39. Bouter LM, Zielhuis GA, Zeegers MP. Textbook of Epidemiology. Houten, the Netherlands: Bohn Stafleu van Loghum; 2017.

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
SWOT: strengths, weaknesses, opportunities, and threats

Edited by A Kushniruk; submitted 06.06.24; peer-reviewed by E Huber, GC Nickel; comments to author 14.10.24; revised version received 05.12.24; accepted 10.02.25; published 31.03.25.

Please cite as:
van Steenis S, Helder O, Kort HSM, van Houwelingen T
Impact of Bottom-Up Cocreation of Nursing Technological Innovations: Explorative Interview Study Among Hospital Nurses and Managers
JMIR Hum Factors 2025;12:e60543
URL: https://humanfactors.jmir.org/2025/1/e60543
doi:10.2196/60543
PMID:

XSL•FO
RenderX

©Saskia van Steenis, Onno Helder, Helianthe S M Kort, Thijs van Houwelingen. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 31.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Turning Patients' Open-Ended Narratives of Chronic Pain Into Quantitative Measures: Natural Language Processing Study

Raquel Norel[1], PhD; Jennifer Gewandter[2], PhD; Zhengwu Zhang[3], PhD; Anika Tahsin[4], BSc; Chadi G Abdallah[5], MD; John Markman[6], MD; Zhiyao Duan[4], PhD; Guillermo Cecchi[1]; Paul Geha[4], MD

[1]IBM (United States), Yorktown Heights, NY, United States

[2]Department of Anesthesiology and Perioperative Medicine, School of Medicine and Dentistry, University of Rochester, Rochester, NY, United States

[3]Department of Statistics and Operations Research, University of North Carolina, Chapel Hill, NC, United States

[4]Department of Psychiatry, University of Rochester, Rochester, NY, United States

[5]Menninger Department of Psychiatry & Behavioral Sciences, Baylor College of Medicine, Houston, TX, United States

[6]Eli Lilly (United States), Chestnut Hill, MA, United States

Corresponding Author:
Paul Geha, MD
Department of Psychiatry
University of Rochester
430 Elmwood Ave
Rochester, NY, 14620
United States
Phone: 1 (585) 275 8675
Email: paul_geha@urmc.rochester.edu

Abstract

Background: Subjective report of pain remains the gold standard for assessing symptoms in patients with chronic pain and their response to analgesics. This subjectivity underscores the importance of understanding patients' personal narratives, as they offer an accurate representation of the illness experience.

Objective: In this pilot study involving 20 patients with chronic low back pain (CLBP), we applied emerging tools from natural language processing (NLP) to derive quantitative measures that captured patients' pain narratives.

Methods: Patients' narratives were collected during recorded semistructured interviews in which they spoke about their lives in general and their experiences with CLBP. Given that NLP is a novel approach in this field, our goal was to demonstrate its ability to extract measures that relate to commonly used tools, such as validated pain questionnaires and rating scales, including the numerical rating scale and visual analog scale.

Results: First, we showed that patients' utterances were significantly closer in semantic space to anchor sentences derived from validated pain questionnaires than to their antithetical counterparts. Furthermore, we found that the semantic distances between patients' utterances and anchor sentences related to quality of life were strongly correlated with reported CLBP intensity on the numerical rating and visual analog scales. Consistently, we observed significant differences between individuals with low and high pain levels.

Conclusions: Although our small sample size limits the generalizability of these findings, the results provide preliminary evidence that NLP can be used to quantify the subjective experience of chronic pain and may hold promise for clinical applications.

(JMIR Hum Factors 2025;12:e80269) doi:10.2196/80269

KEYWORDS

chronic pain; pain narratives; natural language processing; semantic distance; digital phenotyping

Introduction

The most recent 2020 International Association for the Study of Pain definition of pain emphasizes that pain is "always a personal experience" and is distinct from nociception, involving learning "through life experiences." In line with this definition, the biopsychosocial model of pain [1] also stresses that pain is "a personal experience emerging from the dynamic interplay between biological, psychological, and social factors." Despite this modern understanding, the everyday clinical approach to

XSL•FO
RenderX

pain still predominantly focuses on its location or injury-related causes, while monitoring pain primarily through subjective reports [2]. The "injury model" approach is incomplete, as it overlooks the affective and personal dimensions of the pain experience, which can have a significant negative impact on treatment response and long-term outcomes [3-5].

Chronic low back pain (CLBP), the most common chronic pain condition, is a typical example of the inadequacy of the "injury model" in explaining the origin of pain, the long-term outcomes, or response to treatment [5-7]. In fact, considerable variability exists among patients with low back pain, yet the most common subdivision still relies on peripheral symptom assessment, such as back pain with and back pain without radiculopathy [8,9]. The lack of understanding of this variability makes generalizing results from clinical trials to clinical practice very challenging because the observed effects in a heterogeneous population of patients do not necessarily translate to individual patients with specific phenotypes [10]. Clinical approaches that are able to capture different aspects of the "personal experience" of back pain, without relying solely on the traditional medical examination, are therefore needed to improve precision in characterizing and, consequently, treating subgroups of patients [11,12]. Experts agree that the absence of such approaches may be one of the reasons for the disappointingly poor outcomes of clinical trials for chronic pain conditions [13,14]. In addition, if these markers are affordable and easily implemented in a clinical setting, they could help generalize research results to clinical practice [13].

To date, there are no automated, quantitative tools that can both reliably assess patients with chronic pain [15] and provide clinicians with timely, clinically useful information about their pain experience. Rather, patients rely on verbal and nonverbal communication to convey their complaints. In everyday clinical practice, tools such as the numerical rating scale (NRS, 0-10), or self-administered pain questionnaires, including the McGill Pain Questionnaire (MPQ) [16] or the Patient-Reported Outcomes Measurement Information System (PROMIS) instruments, may be used. However, their utility is limited by the time constraints of busy clinical settings and by the restricted scope of information they provide. For example, an NRS rating of 4/10 does not capture the differences in the pain experience between a patient with low back pain and one with knee osteoarthritis. Therefore, patients rely on language to convey these differences to clinicians.

The language patients use can provide a valuable window into their personal experience of pain and the emotions associated with living with it. The recent advent of large language models (LLMs), which can analyze human-generated texts, offers a unique opportunity to probe the different dimensions of the chronic pain experience. Quantitative approaches of this kind are now beginning to emerge. LLMs have been applied to the analysis of clinical notes [17,18] and patients' pain narratives [19-21]. For example, one study on patients with CLBP [20] investigated the language features that distinguished between responders and nonresponders to placebo treatment using a posttreatment interview. Another study analyzed written narratives obtained from patients with fibromyalgia and found significant agreement between LLM-based assessments and those of expert clinicians, supporting the use of LLM in the clinical evaluation of patients with pain.

Here, we used patients' personal "pain stories" obtained during recorded interviews, along with natural language processing (NLP) tools, to derive quantitative measures of their pain experiences. We aimed to provide evidence that NLP can be used to analyze patients subjective accounts of their personal pain experiences and that the derived linguistic features are directly related to core phenotyping domains assessed using measures recommended by the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) and the National Institutes of Health minimal dataset for CLBP [13,22].

The interviews can be conducted in both clinical and nonclinical settings, and they are relatively inexpensive to obtain, thereby increasing accessibility for a broader patient population. The "pain story" is elicited through open-ended questions by a laboratory member without formal clinical training. It aims to capture not only a description of the physical symptoms but also how pain affects the patient's functionality, interactions with others, and overall life trajectory and meaning. This rich dataset expands upon the information typically collected during clinical encounters, which are often constrained by time.

NLP is a novel approach for identifying or distinguishing patients with various psychiatric disorders, such as major depression or psychosis, where standard diagnosis relies on subjective reports of symptoms and the impression of clinical experts [23]. It involves a set of tools built using an extensive corpus of language available on the World Wide Web and machine learning methods to determine the various aspects (eg, meaning and emotion expressed) of words and sentences used in a text by situating them in the multidimensional linguistic space. NLP has been successfully used to distinguish patients with various mental health and neurological conditions [24-27].

Methods

Participants

This study was not registered before data collection. Data were collected between December 15, 2022, and June 8, 2023. A total of 514 individuals responded to advertisements posted on social media and flyers distributed within the Rochester community. Of these individuals, 21 patients with CLBP were recruited for this study between February and June of 2023; however, 1 patient did not show up for the testing session.

Patients were first screened over the phone. To be included, they had to be aged ≥40 years; this cutoff was chosen because the prevalence of low back pain is highest in individuals aged 40 to 80 years [28]. Patients had experienced low back pain for more than 1 year below the thoracic T10 vertebra [12], with an average pain intensity greater than 4/10 on the NRS. CLBP must be the primary pain condition. At screening, patients completed the Telephone Interview for Cognitive Status (TICS) [29] and were excluded if their TICS score was less than 33. We used TICS as a screening tool to avoid bringing patients to the laboratory who might have cognitive impairments that were not detected during the screening phone call. Patients were also

XSL•FO
RenderX

excluded if they had a history of other significant pain conditions, such as fibromyalgia or migraine headaches; chronic unstable medical conditions affecting any bodily organ system; a history of traumatic brain injury; substance misuse; or mental illness. All patients, except 1, reported less than mild depression symptoms on the Montgomery-Asberg Depression Rating Scale (score <19) [30]. Patients were either opioid-naive or had not taken opioids for more than 2 weeks in the previous year. Table 1 presents the clinical and demographic characteristics of the CLBP sample; additional demographic and clinical characteristics are provided in Table S1 in Multimedia Appendix 1.

Table 1. Demographic and clinical characteristics (N=20, 10 males and 10 females).

Measure	Values, mean (SEM)
Age at screening (y)	58.3 (2.4)
BMI (kg/m^2)	29.9 (2.1)
Years of education	14.9 (0.6)
NRS[a] pain	4.0 (0.7)
VAS[b] pain	35.1 (6.1)
MPQ[c] total score	13.2 (2.8)
MPQ sensory score	11.1 (2.1)
MPQ affective score	2.2 (0.8)
PCS[d]	11.6 (4.1)
PDQ[e]	4.2 (2.1)
CSI[f]	33.9 (5.3)
VAS depression	16.1 (5.8)
MADRS-S[g]	7.9 (1.9)
HADS[h]-depression	4.4 (1.4)
HADS-anxiety	6.2 (1.2)
RMDQ[i]	8.9 (1.5)
TICS[j]	39.0 (1.0)
NEO[k] personality inventory	200.1 (7.1)

[a]NRS: numerical rating scale.
[b]VAS: visual analog scale.
[c]MPQ: McGill Pain Questionnaire.
[d]PCS: Pain Catastrophizing Scale.
[e]PDQ: PainDETECT Questionnaire.
[f]CSI: central sensitization inventory.
[g]MADRS-S: Montgomery-Asberg Depression Rating Scale–Self-Report.
[h]HADS: Hospital Anxiety and Depression Scale.
[i]RMDQ: Roland-Morris Disability Questionnaire.
[j]TICS: telephone interview cognitive status.
[k]NEO: neuroticism, extraversion, and openness.

Ethical Considerations

All participants provided written informed consent, and the study was approved by the University of Rochester Medical Center Institutional Review Board (00007633). Data protection was ensured using several layers of security in accordance with the Health Insurance Portability and Accountability Act and under the oversight of the University of Rochester Institutional Review Board . Participants were instructed not to mention their names during the interview, and their data were deidentified. All computer-based materials and data, including voice and video recordings, were stored in a password-protected server managed by the University of Rochester IT Department and protected by the university firewall. All clinical and demographic data were entered directly by the participants into the REDCap (Research Electronic Data Capture; Vanderbilt University), where the code numbers were associated with the data records. REDCap is a free, Health Insurance Portability and Accountability Act–compliant web-based application used for the electronic management of research data.

XSL•FO
RenderX

Data Collection

Interview

On the day of the interview, after obtaining consent, participants provided a urine sample to screen for drugs of abuse using a NexScreen LLC drug screen cup. Drug screening was performed because drugs of abuse are known to alter pain perception [31,32] and language output [33,34], and misuse was one of our exclusion criteria. Next, the research coordinator reviewed the inclusion and exclusion criteria with the participants and explained the study procedures. Interviews were conducted via Zoom (Zoom Video Communications, Inc) to capture voice recordings on separate channels for the patient and the interviewer. Patients were seated in front of an iPad in a quiet room, while the coordinator was in an adjacent room. All patients were interviewed in the same room by the same female coordinator (AT), with a male laboratory member present in the coordinator's room during the interviews. We included both male and female examiners during the interview to minimize bias in the patients' responses related to sex. Although this bias has not yet been examined in the context of language-based studies, in other pain-testing contexts, the sex of the examiner has been shown to influence pain-related outcomes [35-37].

The interview questions, provided in Table S2 in Multimedia Appendix 1, served as a guide for the interviewer. The questions were piloted by AT, JG, and PG during the mock interviews. PG is a trained psychiatrist with extensive experience in interviewing patients in clinical settings. However, the interviewer did not have to adhere strictly to the interview questions and asked patients open-ended questions to elicit a subjective, patient-centered narrative of pain. As shown in Table S2 in Multimedia Appendix 1, the questions were inspired by previous NLP work in chronic pain [20] and by the general biopsychosocial model of chronic pain [1].

Questionnaires Collected

The patients completed the TICS and Montgomery-Asberg Depression Rating Scale over the phone as described earlier. All remaining pain ratings and questionnaires were completed after the interview to avoid any influence from these tools on the content of the interview. When in the laboratory, the patients rated their low back pain intensity on a visual analog scale (VAS) using a cursor on a line projected on the computer screen in front of them and completed the National Institutes of Health minimal dataset for CLBP [22]. They also completed questionnaires and measures recommended by IMMPACT [13], including the short form of the MPQ, the PainDETECT Questionnaire, the Pain Catastrophizing Scale, the Hospital Anxiety and Depression Scale, and the Roland-Morris Disability Questionnaire. The core domains recommended by IMMPACT overlap extensively with those identified in a more recent consensus report on core outcome sets [14]. In addition, the patients with CLBP completed the central sensitization inventory, as well as the neuroticism, extraversion, and openness personality inventory, and reported their list of prescribed medications.

Preprocessing of Interview Text

Voice data were recorded in 2 channels, one for the interviewer and one for the participant. Automatic audio transcription of the entire participants' interview was performed using the Whisper base model [38]. We used the transcriptions verbatim for further analysis and performed the analysis sentence by sentence. Sentences were obtained using the sentence tokenizer from the Natural Language Toolkit [39], which is an open-source Python library designed for working with human language data.

Data Analysis

Construction of Anchor Sentences

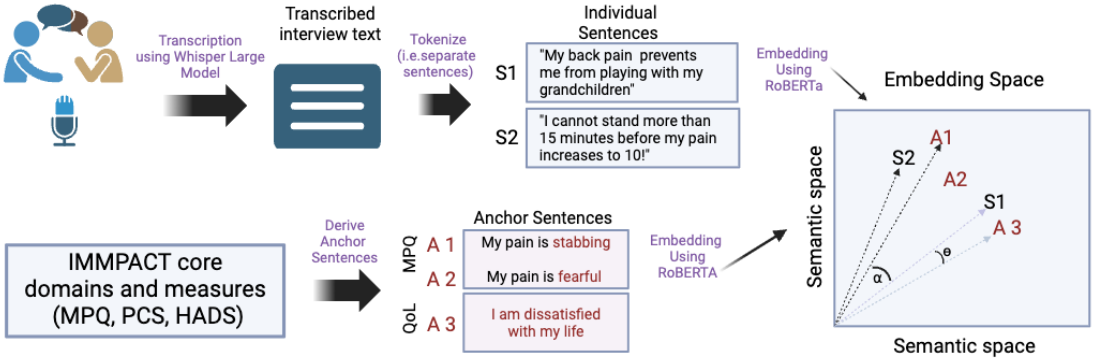
Our aim was to understand how the patients' personal accounts of their "pain stories" were situated in a multidimensional linguistic space relative to the core domains and validated measures recommended by experts' consensus [13,22]. These measures are often used in chronic pain research but are less commonly applied to monitor patients in everyday clinical settings. To achieve this, we constructed anchor sentences based on validated questionnaires that covered topics related to the sensory, affective, functional, and quality-of-life aspects of patients with CLBP. We present all the anchor sentences that we used in Table S3 in Multimedia Appendix 1, along with the source questionnaire from which they were derived. The questionnaires included the MPQ [16], the Pain Catastrophizing Scale [40,41], PROMIS [42,43], and the World Health Organization Quality of Life Questionnaire [44], which are measures tested for content validity and involved patients' population during their conception. Because LLMs can misunderstand negation [45,46], some of the anchor sentences were rephrased from the original questionnaires. Although the list of anchor sentences may not be exhaustive, we ensured that they addressed the biological, psychological, and social domains of patients' lives.

Feature Extraction

We used semantic embedding [47], a fundamental NLP and machine learning concept, for semantic feature extraction from the interview text. Figure 1 illustrates the overall approach. This method consists of projecting entire sentences into a semantic space and converting them to vectors with 1024 coordinates [48]. It enables computers to understand and process human language by mapping words, phrases, or entire documents to vectors of real numbers. These vectors can then be analyzed with a rich set of mathematical tools to extract valuable information from the text, facilitating tasks such as sentiment analysis, topic modeling, and semantic similarity measurement, which we used in this work [49-52]. We generated semantic embeddings using the RoBERTa (a robustly optimized bidirectional encoder representations from transformers pretraining approach) large model [48], which was trained on 160 GB of text.

XSL•FO
RenderX

Figure 1. Schematic representation of the data collection and analysis workflow. Transcribed interview text and anchor sentences were semantically embedded to calculate the semantic similarity between the anchor sentences and the interview text. The embedding space represented here is a projection to a 2D graph and is therefore a simplification of the actual space, which comprises 1024 dimensions. As depicted, sentence S1, uttered by the patient about daily life, is closer to anchor A3 than to anchor A1 (ie, θ<α). HADS: Hospital Anxiety and Depression Scale; IMMPACT: Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials; MPQ: McGill Pain Questionnaire; PCS: Pain Catastrophizing Scale; QoL: quality of life.



Analysis Using Anchor Sentences and Their Antithetical Pair

We wanted to evaluate whether the anchor set, derived as described in the previous section, adequately covered the interview topics by comparing the semantic distance between the interview sentences and both the anchor and its antithesis. More specifically, we calculated the semantic embeddings (Figure 1) of each sentence uttered by the patient—using the entire interview—the anchor sentence, and its antithetical pair. Next, semantic similarity was quantified by calculating the cosine similarity between the vectors of the anchor sentences and interview sentences, on the one hand, and the antithetical anchor sentences and the interview sentences on the other hand, providing a quantitative measure of their semantic resemblance. This calculation resulted in 2 distributions of semantic similarities for each anchor pair. We used the 95th percentile point (Figure S2 in Multimedia Appendix 1) of each distribution to represent the semantic distance of the interview sentences to the anchor and to its antithetical pair, respectively. This value, the 95th percentile of the similarity scores, represents a robust maximum after omitting outliers. The robust maximum is more resistant to outliers than the absolute maximum and is inherently invariant to the length of the interview. To further illustrate our approach, we plotted the distribution of semantic similarities from 2 participants representing the extreme ends of the number of sentences uttered during the interview to demonstrate the robustness of the 95th percentile, which falls within the 0.4 to 0.6 range of semantic similarity (Figure S1 in Multimedia Appendix 1).

Relating Semantic Similarities to Reported Back Pain Intensity

Ratings of chronic pain intensity are often used in clinical settings and are a mainstay of clinical pain research. Therefore, we tested the relationship of semantic similarities to the anchor sentences, showing significant differences in the previous analysis and the pain intensity reported on the NRS and the VAS. We calculated the correlation between semantic similarities to the anchor sentences, showing significant differences between the previous analysis and the reported pain back pain ratings. We also assessed whether the anchor sentences were sensitive enough to detect differences between interviews from the high- and low-pain subgroups.

Statistical Analysis

The Kolmogorov-Smirnov (KS) test was used to assess the differences between the distributions of semantic similarities. Specifically, we compared the distribution of the 95th percentile robust maximum for each anchor sentence with that of its corresponding antithetical sentence across patients. Because we analyzed 33 pairs of sentences, we corrected for multiple comparisons using the false discovery rate (FDR) method. Additionally, we compared the distributions of semantic similarities to the anchor sentences between the high-pain and low-pain groups using the KS test. Finally, we used Spearman's ρ correlation coefficient to examine the relationship between semantic similarities and reported pain intensity ratings.

Results

Demographic and Clinical Characteristics

Table 1 shows the demographic and clinical characteristics of the sample of 20 patients with CLBP. In total, 50% of the participants were female. By design, the patients did not have any evidence of cognitive impairment (*TICS*>33). Complete details of the patients' characteristics are reported in Table S1 in Multimedia Appendix 1.

Analysis of the Interviews

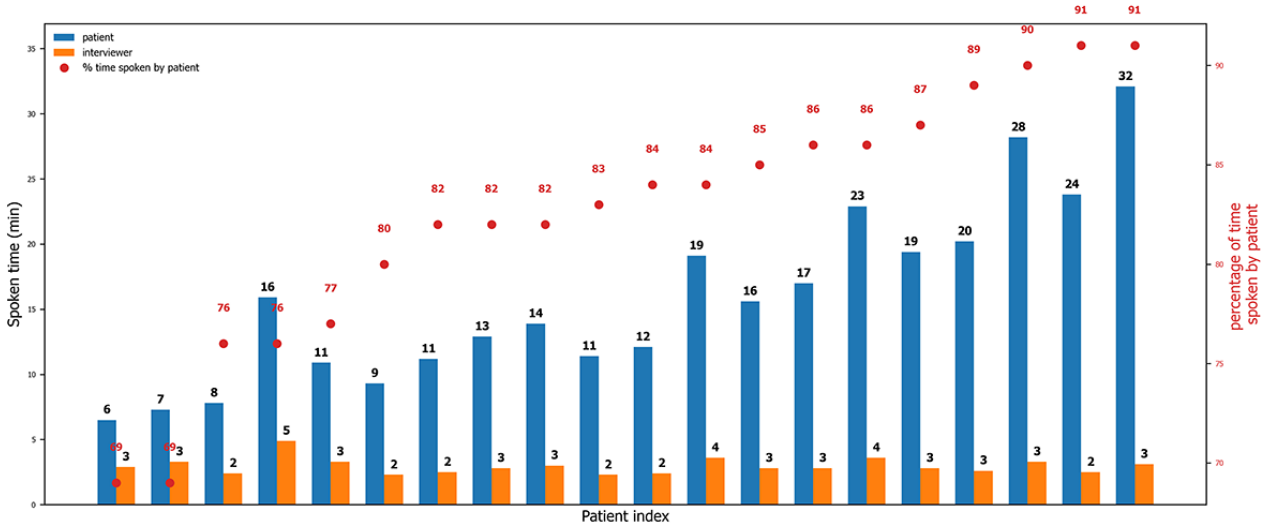
Interview Duration

The duration of the interviews ranged from less than 10 minutes to more than half an hour, with a total of 376.4 minutes of speech. The interviews across all participants included a total of 43,641 words of naturalistic language data with a unique vocabulary of 3158, making the total sample a substantial linguistic corpus. The vocabulary richness ratio was 0.072, and the natural speech rate was 115.9 words per minute, demonstrating genuine and varied linguistic expression rather than constrained responses [53,54]. The patients spoke between 69% and 91% of the interview time (Figure 2). There was no

XSL•FO
RenderX

statistically significant correlation between pain scores and interview duration (NRS and VAS: P=.94) or between pain scores and the percentage of time spoken by the participant (NRS: P=.42; VAS: P=.72). Because years of education may affect speech generation or the richness of the text generated, we examined the correlation between years of education and the interview features, such as duration, time spoken by the participants, and time spoken by the interviewers, as well as between years of education and the semantic embeddings of the anchor sentences (Table S4 in Multimedia Appendix 1). Years of education were negatively correlated with the time spoken by the interviewer, indicating that the latter prompted participants with fewer years of education more frequently than those with more years of education. Years of education were not correlated with the overall interview duration, the time spoken by the participant, the word content, or the pain scores. In addition, years of education were not significantly related to the semantic embeddings of the anchor sentences, except for "I rate my quality of life as poor" and "My physical environment is unhealthy."

Figure 2. Interview duration and percentage of time spoken by the interviewer and the patient. The left axis depicts the duration (in minutes) spoken by each patient (blue) and the interviewer (dark orange); the right axis depicts the percentage of interview time spoken by the patient (red dots). The x-axis depicts all patients (N=20), arranged by interview duration from shortest to longest.



Semantic Distance of Interview Texts to Anchor Sentences and Their Antithetical Sentences

Figure S2 in Multimedia Appendix 1 illustrates the distributions of semantic similarities for 1 anchor sentence—"I have no pain"—and its antithesis—"I am in pain." Table S5 in Multimedia Appendix 1 shows example semantic embeddings for 2 anchor sentences. The semantic embeddings of 7 pairs of anchor sentences and their antithesis showed statistically significant differences (P<.05; FDR corrected) in their similarities to interview sentences (Table 2). In total, 5 pairs of sentences were derived from the MPQ, 1 from the measure of pain variability, and 1 from the PROMIS-Pain Interference questionnaire. Table S6 in Multimedia Appendix 1 presents the results for all the anchor-antithesis sentence pairs tested.

XSL•FO
RenderX

Table 2. Anchors and their antithetical sentences showing significant differences in their semantic similarities to the interview text.

Anchor	Antithesis	Questionnaire	KS[a] test	*P* value	q value[b]
My pain is cramping	My pain is relaxing	MPQ[c]	0.95	5.80×10^{1}	1.92×10
My pain is splitting	My pain is light and dull	MPQ	0.8	1.33×10	1.46×10
My pain is throbbing	My pain is steady	MPQ	0.65	2.70×10^{4}	2.23×10^{3}
My pain is stabbing	My pain is dull and aching	MPQ	0.65	2.70×10^{4}	2.23×10^{3}
My pain is tiring-exhausting	My pain is bearable	MPQ	0.55	3.97×10^{3}	2.18×10^{2}
I am in pain	I have no pain	Pain variability	0.85	1.43×10	2.37×10
I am unable to focus	I am able to focus	PROMIS[d]-pain interference	0.55	3.97×10^{3}	2.18×10^{2}

[a]KS: Kolmogorov-Smirnov.

[b]q value refers to the false discovery rate–adjusted *P* value. Each anchor sentence has its opposite (eg, I have no pain and I am in pain). We computed the semantic similarity between each sentence from the interview and the anchor. From that distribution of values (1 value per sentence), we retained the robust maximum (the 95th percentile, which is more resistant to outliers than the maximum value). We then repeated the same calculation for the antithesis of the anchor, also retaining the robust maximum semantic similarity. We repeated this for each participant, resulting in 1 distribution of values extracted from the anchors and 1 from their antitheses. We then compared these two distributions using the 2-sample Kolmogorov-Smirnov test to determine whether significant differences existed between the anchor and its antithesis used to query the data in this cohort. This nonparametric test is designed to determine if 2 samples originate from the same distribution.

[c]MPQ: McGill Pain Questionnaire.

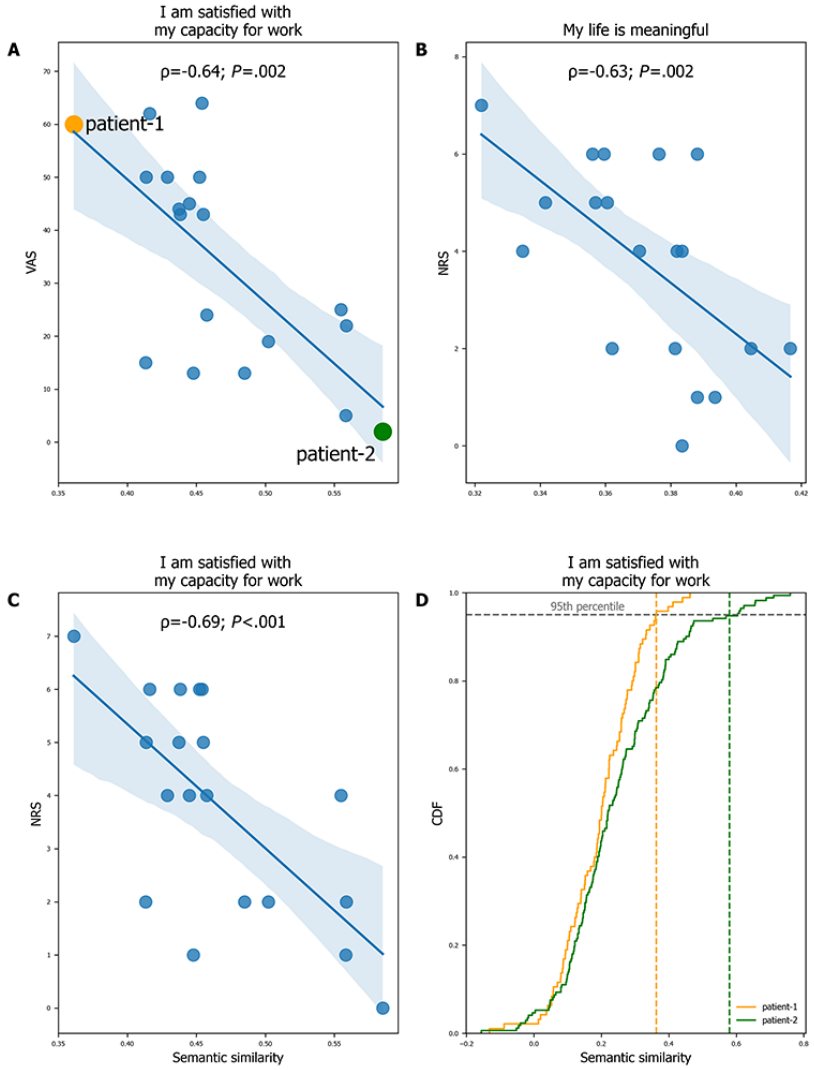
[d]PROMIS: Patient-Reported Outcomes Measurement Information System.

Relationship Between Semantic Distance of Interview Texts to Anchor Sentences and Back Pain Intensity

NRS and VAS are commonly used tools in clinical and research settings. Therefore, we tested whether the semantic distance of the patients' interviews to the anchor sentences scaled with the patients' reported pain intensity. Figure 3 shows 3 examples in which the correlations were statistically significant ($P<.05$; FDR corrected). The semantic similarity to the anchor sentence "I am satisfied with my capacity for work" was inversely related to both NRS and VAS scores. Similarly, the semantic similarity to "I am satisfied with my ability to perform my daily living activities" was inversely related to the NRS pain score. In the same figure, we illustrate the differences in the distributions of semantic similarity between the interview and the antithesis "I am satisfied with my capacity for work" for 2 participants: 1 with high pain (patient-1) and 1 with low pain (patient-2). Table S7 in Multimedia Appendix 1 presents all calculations performed between the anchor NRS and VAS, on the one hand, and the semantic embeddings of the anchor sentences, on the other hand. Notably, the NRS and VAS were most strongly correlated with the semantic embeddings of anchor sentences describing the quality of life. Another way of examining the relationship between the interview text and pain intensity was to compare the high- and low-pain subgroups. Patients who reported a VAS pain score ≤30 or an NRS score ≤3 were classified into the low-pain group; otherwise, they were classified into the high-pain group. Figure 4 shows the differences between the subgroups with high and low pain in the semantic similarity to "I am satisfied with my capacity for work" and to "I am satisfied with my ability to perform my daily activities." Not surprisingly, participants with high pain reported less satisfaction with their quality of life (Figure 4).

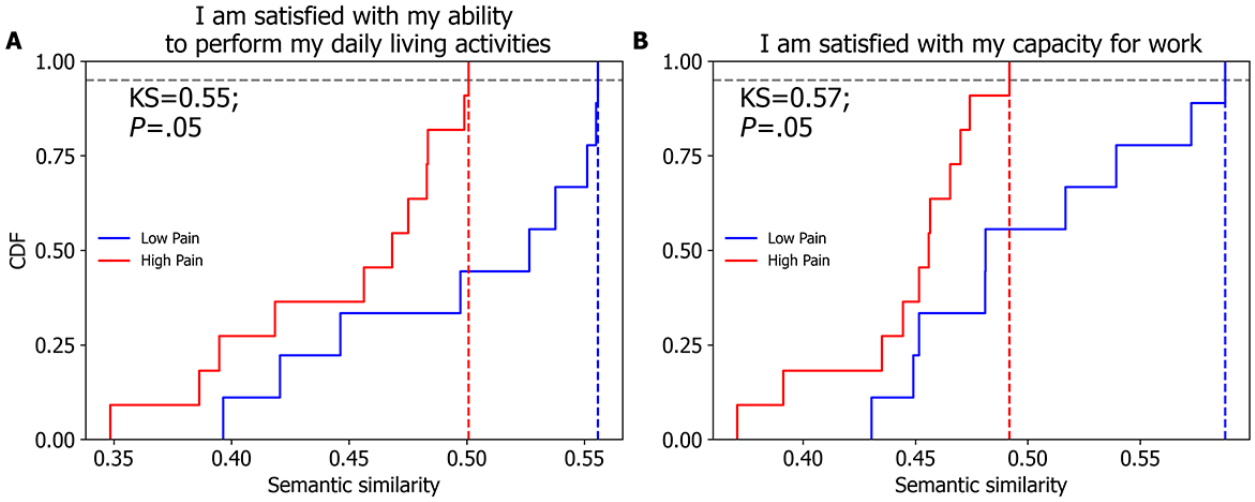
XSL•FO
RenderX

Figure 3. Relationship (Spearman ρ, *P*<.01) between the semantic similarity of patients' texts and anchor sentences and the rated back pain intensity. Significant correlations between back pain intensity and semantic similarity for "I am satisfied with my capacity for work" (A and C) and "I am satisfied with my ability to perform my daily living activities" (B). (D) Examples of 2 cumulative distributions of semantic similarity between the interview text and "I am satisfied with my capacity for work" obtained from 2 patients on both sides of the pain intensity spectrum, as shown in panel A. Patient 1 (dark orange) experienced relatively higher low back pain intensity and therefore showed smaller similarity to the anchor sentence, whereas patient 2 (green) showed the opposite pattern. Dashed lines represent the 95th percentile (the robust maximum used in our statistical analyses). The blue shading represents the CIs for the regression line. CDF: cumulative distribution function; NRS: numerical rating scale; VAS: visual analog scale.



XSL•FO
RenderX

Figure 4. Comparison of the distribution of semantic similarities to anchor sentences between high (visual analog scale [VAS]>30; red) and low pain (VAS≤30, blue) groups. The cumulative distributions of semantic similarity between the interview text and "I am satisfied with my ability to perform my daily living activities" (A), and "I am satisfied with my capacity for work" (B), differed significantly and were shifted to the left in the high-pain group compared with the low-pain group (Kolmogorov-Smirnov [KS] test, P<.01). CDF: cumulative distribution function.



Discussion

Principal Findings

In this study, we present evidence that subjective narratives of patients with CLBP, elicited through open-ended questions about their lives posed by an interviewer with no formal clinical training, can provide semantic features directly related to well-validated tools currently used to assess patients with pain [13]. Semantic embeddings, which assess the semantic similarity between anchor sentences—derived from validated pain and quality of life questionnaires—and sentences from the interviews, demonstrated statistically significant differences between negative and positive assertions about pain. In other words, the interview texts from patients showed asymmetric semantic similarities when queried with pain-related anchor sentences, providing evidence that NLP can detect clinically meaningful signals from the content of the interviews.

Additionally, the semantic embeddings differed significantly when their distributions were compared between subgroups with low and high pain levels. Notably, our analysis revealed a strong connection between the language used by patients with CLBP and their self-reported satisfaction with work capacity and other quality of life measures. Specifically, the semantic similarities of interview texts to anchors probing patient satisfaction with their life, such as "I am satisfied with my capacity for work" or "I am satisfied with my ability to perform my daily activities" were strongly inversely correlated with pain intensity, indicating that these aspects of patients' lives are deeply intertwined with their experience of pain. The statement "I am satisfied with my capacity for work," derived from the World Health Organization Quality of Life Questionnaire [44], reflects patients' perception of how pain impacts their ability to perform work-related tasks, maintain employment, and engage in productive activities. Satisfaction with work capacity may indicate a degree of pain management and coping strategies that enable individuals to continue participating in meaningful and fulfilling roles despite their condition [55]. Both anchors underscore the importance of addressing not only the physical symptoms of pain but also its broader psychological and existential impacts. This also suggests that patients' perception of their ability to work can serve as an indicator of their overall well-being and a potential target in chronic treatment. This finding aligns with the results of the INTEGRATE-pain Delphi process [14], which identified activities of daily living and quality of life as core outcome measures for chronic pain treatment. Furthermore, it highlights how applying NLP to patient interviews can extract patient-centered, quantitative measures that are not readily available from current state-of-the-art assessments in busy clinical settings, where the focus tends to be more on the sensory aspects of pain.

Not all anchor-antithesis pairs derived from validated questionnaires showed a statistically significant difference in their semantic distance to patients' narratives. In total, 21 of the 33 anchor pairs yielded a KS value ≥0.3, indicating a moderate-to-large separation between the distributions of semantic distances [56]. We predict that all these pairs will reach statistical significance with larger samples. The remaining 12 pairs had a KS value <0.30, indicating only a small separation. Several factors may account for these findings. These pairs may not adequately capture the nuances of the interview texts, or the way we constructed the pairs may have been limited. Another possibility is that the RoBERTa large model [48] used for semantic embeddings has intrinsic limitations. As these models continue to evolve, their embeddings are refined, potentially enhancing their ability to capture the relationships between questionnaire-based anchors and patients' narratives.

To date, there are no validated and widely accepted automated tools that can reliably assess patients with chronic pain; therefore, these patients must rely on verbal and nonverbal communication to convey their distress to others, including clinicians. Having such tools is critical for guiding clinical decision-making and for conducting valid, quantitative research on the mechanisms and therapeutics of chronic pain. Until recently, research on pain language has remained qualitative

XSL•FO
RenderX

[57] or was mostly limited to single-word descriptors embedded in questionnaires such as the well-known MPQ [16]. Nevertheless, studies using quantitative linguistic analysis of narratives from patients with chronic pain to measure pain or response to interventions have begun to emerge rapidly [20,21,58]. We have previously shown, for example, that by using LLMs to extract linguistic features from patients' interviews, we could distinguish between those with fibromyalgia and those with neuropathic pain, with an area under the receiver operating characteristic curve of 0.83 despite limited sample sizes [59]. Furthermore, we have demonstrated that NLP applied to patients' narratives can retrospectively [21] and prospectively predict response to treatment [20,58]. Nunes et al [60] analyzed short interviews recorded in response to 7 questions about patients' pain from 65 participants with various chronic pain conditions who reported mild, moderate, or severe pain. An NLP-based support vector machine model with early fusion of linguistic features was able to distinguish among the 3 groups of patients, thereby demonstrating, in a proof of concept, that NLP can be used to estimate chronic pain intensity. This emerging field, combined with the ease of obtaining voice recordings through the widespread use of smart devices, has the potential to provide novel, automated, and quantitative measures of patients' experiences of pain. These measures can be assessed frequently and can complement existing tools to enhance both patient care and research.

A review of the recent literature [61] studying pain language analysis using computational methods found that most studies relied on electronic health records and social media posts, with only a minority relying on patients' interviews. The most common objective was to classify patients with chronic pain. Our aim in this study was to demonstrate that a relatively brief interview (<30 min), inspired by the biopsychosocial assessment of pain and designed to elicit a subjective narrative of the patient's daily life and pain experience, can be analyzed using NLP tools to capture quantitative and clinically significant measures of the pain experience. These measures can then be used in future studies using the same approach to differentiate among patient groups, monitor progress, or evaluate treatment response. In addition to being patient-centered, substantial advantages of this approach include the ability to obtain data remotely and the absence of the need for a medical professional during data collection, both of which significantly reduce the cost of data collection.

Study Limitations

This study included a small sample of patients aged ≥40 years, which may limit the generalizability and linguistic representativeness of the findings. In addition, the small sample may have included subgroups of patients with CLBP—such as those with or without a neuropathic component—which could introduce heterogeneity into both the pain experience and the language used by the patients. Furthermore, the semantic distance calculations do not fully capture the complete meaning, with all its nuances, expressed by the patients. Thus, NLP cannot replace expert judgment but should rather be used alongside expert clinician assessments. Additionally, the process of constructing anchor sentences—particularly given the inability of LLMs to handle negations—may have affected the content validity of the anchors used, despite their derivation from content-valid questionnaires. Nevertheless, the language features related to validated pain questionnaires were strongly correlated with reported pain intensity and significantly distinguished between the high- and low-pain groups. Future research should explore larger and more diverse cohorts of patients with CLBP, including more representative age groups and groups with varying levels of common CLBP comorbidities such as depression, anxiety, or cognitive changes, to confirm these findings and investigate other aspects of the condition, such as its progression or response to treatment.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to participants' confidentiality.

Funding

This study was funded by the National Institutes of Health (grant NIAMS-R01AR080627).

Authors' Contributions

This study was designed by RN, JG, ZZ, CGA, JM, ZD, GC, and PG. AT collected the data. RN, GC, and PG analyzed the data, and all the authors critically reviewed the results. RN had the primary role in preparing the manuscript, which was edited by GC and PG. All the authors approved the final version of the manuscript and agreed to be accountable for all aspects of the work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional methods and tabulated data.
[DOCX File , 2501 KB - humanfactors_v12i1e80269_app1.docx]

References

XSL•FO
RenderX

1. Gatchel RJ, Peng YB, Peters ML, Fuchs PN, Turk DC. The biopsychosocial approach to chronic pain: scientific advances and future directions. Psychol Bull 2007 Jul;133(4):581-624. [doi: 10.1037/0033-2909.133.4.581] [Medline: 17592957]
2. Nicholas M, Vlaeyen JW, Rief W, Barke A, Aziz Q, Benoliel R, IASP Taskforce for the Classification of Chronic Pain. The IASP classification of chronic pain for ICD-11: chronic primary pain. Pain 2019 Jan;160(1):28-37. [doi: 10.1097/j.pain.0000000000001390] [Medline: 30586068]
3. Sullivan MD, Sturgeon JA, Lumley MA, Ballantyne JC. Reconsidering Fordyce's classic article, "Pain and suffering: what is the unit?" to help make our model of chronic pain truly biopsychosocial. Pain 2023 Mar 01;164(2):271-279 [FREE Full text] [doi: 10.1097/j.pain.0000000000002748] [Medline: 35972469]
4. Turk DC, Fillingim RB, Ohrbach R, Patel KV. Assessment of psychosocial and functional impact of chronic pain. J Pain 2016 Sep;17(9 Suppl):T21-T49. [doi: 10.1016/j.jpain.2016.02.006] [Medline: 27586830]
5. Webster BS, Cifuentes M. Relationship of early magnetic resonance imaging for work-related acute low back pain with disability and medical utilization outcomes. J Occup Environ Med 2010 Sep;52(9):900-907. [doi: 10.1097/JOM.0b013e3181ef7e53] [Medline: 20798647]
6. Vlaeyen JW, Maher CG, Wiech K, Van Zundert JV, Meloto CB, Diatchenko L, et al. Low back pain. Nat Rev Dis Primers 2018 Dec 13;4(1):52. [doi: 10.1038/s41572-018-0052-1] [Medline: 30546064]
7. Brinjikji W, Luetmer PH, Comstock B, Bresnahan BW, Chen LE, Deyo RA, et al. Systematic literature review of imaging features of spinal degeneration in asymptomatic populations. AJNR Am J Neuroradiol 2015 Apr;36(4):811-816 [FREE Full text] [doi: 10.3174/ajnr.A4173] [Medline: 25430861]
8. Baron R, Dickenson AH. Neuropathic pain: precise sensory profiling improves treatment and calls for back-translation. Pain 2014 Nov;155(11):2215-2217. [doi: 10.1016/j.pain.2014.08.021] [Medline: 25168667]
9. Attal N, Fermanian C, Fermanian J, Lanteri-Minet M, Alchaar H, Bouhassira D. Neuropathic pain: are there distinct subtypes depending on the aetiology or anatomical lesion? Pain 2008 Aug 31;138(2):343-353 [FREE Full text] [doi: 10.1016/j.pain.2008.01.006] [Medline: 18289791]
10. Vardeh D, Mannion RJ, Woolf CJ. Toward a mechanism-based approach to pain diagnosis. J Pain 2016 Sep;17(9 Suppl):T50-T69 [FREE Full text] [doi: 10.1016/j.jpain.2016.03.001] [Medline: 27586831]
11. Baron R, Binder A, Attal N, Casale R, Dickenson AH, Treede RD. Neuropathic low back pain in clinical practice. Eur J Pain 2016 Jul;20(6):861-873 [FREE Full text] [doi: 10.1002/ejp.838] [Medline: 26935254]
12. Markman JD, Czerniecka-Foxx K, Khalsa PS, Hayek SM, Asher AL, Loeser JD, et al. AAPT diagnostic criteria for chronic low back pain. J Pain 2020;21(11-12):1138-1148 [FREE Full text] [doi: 10.1016/j.jpain.2020.01.008] [Medline: 32036046]
13. Edwards RR, Dworkin RH, Turk DC, Angst MS, Dionne R, Freeman R, et al. Patient phenotyping in clinical trials of chronic pain treatments: IMMPACT recommendations. Pain 2016 Sep;157(9):1851-1871 [FREE Full text] [doi: 10.1097/j.pain.0000000000000602] [Medline: 27152687]
14. Bova G, Domenichiello A, Letzen JE, Rosenberger DC, Siddons A, Kaiser U, et al. Developing consensus on core outcome sets of domains for acute, the transition from acute to chronic, recurrent/episodic, and chronic pain: results of the INTEGRATE-pain Delphi process. EClinicalMedicine 2023 Dec 02;66:102340 [FREE Full text] [doi: 10.1016/j.eclinm.2023.102340] [Medline: 38089861]
15. Davis KD, Aghaeepour N, Ahn AH, Angst MS, Borsook D, Brenton A, et al. Discovery and validation of biomarkers to aid the development of safe and effective pain therapeutics: challenges and opportunities. Nat Rev Neurol 2020 Jul;16(7):381-400 [FREE Full text] [doi: 10.1038/s41582-020-0362-2] [Medline: 32541893]
16. Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. Pain 1975 Sep;1(3):277-299. [doi: 10.1016/0304-3959(75)90044-5] [Medline: 1235985]
17. Dave AD, Ruano G, Kost J, Wang X. Automated extraction of pain symptoms: a natural language approach using electronic health records. Pain Physician 2022 Mar;25(2):E245-E254 [FREE Full text] [Medline: 35322976]
18. Comito C, Forestiero A, Macrì D, Metlichin E, Giusti GD, Ramacciati N. Comparative analysis of AI algorithms on real medical data for chronic pain detection. Int J Med Inform 2025 Nov;203:106002 [FREE Full text] [doi: 10.1016/j.ijmedinf.2025.106002] [Medline: 40505249]
19. Amidei J, Nieto R, Kaltenbrunner A, Ferreira De Sá JG, Serrat M, Albajes K. Exploring the capacity of large language models to assess the chronic pain experience: algorithm development and validation. J Med Internet Res 2025 Mar 31;27:e65903 [FREE Full text] [doi: 10.2196/65903] [Medline: 40163858]
20. Berger SE, Branco P, Vachon-Presseau E, Abdullah TB, Cecchi G, Apkarian AV. Quantitative language features identify placebo responders in chronic back pain. Pain 2021 Jun 01;162(6):1692-1704. [doi: 10.1097/j.pain.0000000000002175] [Medline: 33433145]
21. Branco P, Berger S, Abdullah T, Vachon-Presseau E, Cecchi G, Apkarian AV. Predicting placebo analgesia in patients with chronic pain using natural language processing: a preliminary validation study. Pain 2023 May 01;164(5):1078-1086 [FREE Full text] [doi: 10.1097/j.pain.0000000000002808] [Medline: 36524810]
22. Deyo RA, Dworkin SF, Amtmann D, Andersson G, Borenstein D, Carragee E, et al. Report of the NIH Task Force on research standards for chronic low back pain. Spine J 2014 Aug 01;14(8):1375-1391. [doi: 10.1016/j.spinee.2014.05.002] [Medline: 24950669]

XSL•FO
RenderX

23. Corcoran CM, Cecchi GA. Using language processing and speech analysis for the identification of psychosis and other disorders. Biol Psychiatry Cogn Neurosci Neuroimaging 2020 Aug;5(8):770-779 [FREE Full text] [doi: 10.1016/j.bpsc.2020.06.004] [Medline: 32771179]
24. Corcoran CM, Carrillo F, Fernández-Slezak D, Bedi G, Klim C, Javitt DC, et al. Prediction of psychosis across protocols and risk cohorts using automated language analysis. World Psychiatry 2018 Mar;17(1):67-75 [FREE Full text] [doi: 10.1002/wps.20491] [Medline: 29352548]
25. Bedi G, Cecchi GA, Slezak DF, Carrillo F, Sigman M, de Wit H. A window into the intoxicated mind? Speech as an index of psychoactive drug effects. Neuropsychopharmacology 2014 Sep;39(10):2340-2348 [FREE Full text] [doi: 10.1038/npp.2014.80] [Medline: 24694926]
26. Birnbaum ML, Norel R, Van Meter A, Ali AF, Arenare E, Eyigoz E, et al. Identifying signals associated with psychiatric illness utilizing language and images posted to Facebook. NPJ Schizophr 2020 Dec 03;6(1):38 [FREE Full text] [doi: 10.1038/s41537-020-00125-0] [Medline: 33273468]
27. Norel R, Agurto C, Heisig S, Rice JJ, Zhang H, Ostrand R, et al. Speech-based characterization of dopamine replacement therapy in people with Parkinson's disease. NPJ Parkinsons Dis 2020 Jun 12;6:12 [FREE Full text] [doi: 10.1038/s41531-020-0113-5] [Medline: 32566741]
28. Hoy D, Bain C, Williams G, March L, Brooks P, Blyth F, et al. A systematic review of the global prevalence of low back pain. Arthritis Rheum 2012 Jun;64(6):2028-2037. [doi: 10.1002/art.34347] [Medline: 22231424]
29. Brandt J, Spencer M, Folstein M. The telephone interview for cognitive status. Neuropsychiatry Neuropsychol Behav Neurol. 1988. URL: https://journals.lww.com/cogbehavneurol/abstract/1988/00120/the_telephone_interview_for_cognitive_status.4.aspx [accessed 2025-11-10]
30. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. Br J Psychiatry 1979 Apr;134:382-389. [doi: 10.1192/bjp.134.4.382] [Medline: 444788]
31. Compton MA. Cold-pressor pain tolerance in opiate and cocaine abusers: correlates of drug type and use status. J Pain Symptom Manage 1994 Oct;9(7):462-473 [FREE Full text] [doi: 10.1016/0885-3924(94)90203-8] [Medline: 7822886]
32. Hooten WM, Mantilla CB, Sandroni P, Townsend CO. Associations between heat pain perception and opioid dose among patients with chronic pain undergoing opioid tapering. Pain Med 2010 Nov;11(11):1587-1598. [doi: 10.1111/j.1526-4637.2010.00962.x] [Medline: 21029354]
33. Agurto C, Norel R, Pietrowicz M, Parvaz M, Kinreich S, Bachi K, et al. Speech markers for clinical assessment of cocaine users. Proc IEEE Int Conf Acoust Speech Signal Process 2019 May;2019:6391-6394 [FREE Full text] [doi: 10.1109/icassp.2019.8682691] [Medline: 32095118]
34. Agurto C, Cecchi GA, Norel R, Ostrand R, Kirkpatrick M, Baggott MJ, et al. Detection of acute 3,4-methylenedioxymethamphetamine (MDMA) effects across protocols using automated natural language processing. Neuropsychopharmacology 2020 Apr;45(5):823-832 [FREE Full text] [doi: 10.1038/s41386-020-0620-4] [Medline: 31978933]
35. Levine FM, Lee De Simone L. The effects of experimenter gender on pain report in male and female subjects. Pain 1991 Jan;44(1):69-72. [doi: 10.1016/0304-3959(91)90149-R] [Medline: 2038491]
36. Gijsbers K, Nicholson F. Experimental pain thresholds influenced by sex of experimenter. Percept Mot Skills 2005 Dec;101(3):803-807. [doi: 10.2466/pms.101.3.803-807] [Medline: 16491681]
37. Chapman CD, Benedict C, Schiöth HB. Experimenter gender and replicability in science. Sci Adv 2018 Jan 10;4(1):e1701427 [FREE Full text] [doi: 10.1126/sciadv.1701427] [Medline: 29349293]
38. Radford A, Kim J, Xu T, Brockman G, McLeavey C, Sutskever I. Robust speech recognition via large-scale weak supervision. 2023 Presented at: Proceedings of the 40th International Conference on Machine Learning; July 23-29; Honolulu, HI.
39. Bird S, Klein E, Loper E. Natural Language Processing with Python. Sebastopol, CA: O'Reilly Media; Jun 2009.
40. Sullivan MJ, Bishop SR, Pivik J. The Pain Catastrophizing Scale: development and validation. Psychol Assess 1995;7(4):524-532. [doi: 10.1037/1040-3590.7.4.524]
41. Osman A, Barrios FX, Kopper BA, Hauptmann W, Jones J, O'Neill E. Factor structure, reliability, and validity of the Pain Catastrophizing Scale. J Behav Med 1997 Dec;20(6):589-605. [doi: 10.1023/a:1025570508954] [Medline: 9429990]
42. Hays RD, Spritzer KL, Schalet BD, Cella D. PROMIS®-29 v2.0 profile physical and mental health summary scores. Qual Life Res 2018 Jul;27(7):1885-1891 [FREE Full text] [doi: 10.1007/s11136-018-1842-3] [Medline: 29569016]
43. Salsman JM, Schalet BD, Park CL, George L, Steger MF, Hahn EA, et al. Qual Life Res 2020 Aug;29(8):2299-2310 [FREE Full text] [doi: 10.1007/s11136-020-02489-3] [Medline: 32306302]
44. Skevington SM, Lotfy M, O'Connell KA, WHOQOL Group. The World Health Organization's WHOQOL-BREF quality of life assessment: psychometric properties and results of the international field trial. A report from the WHOQOL group. Qual Life Res 2004 Mar;13(2):299-310. [doi: 10.1023/B:QURE.0000018486.91360.00] [Medline: 15085902]
45. Jang J, Ye S, Seo M. Can large language models truly understand prompts? A case study with negated prompts. ArXiv. Preprint posted online on September 26, 2022 . [doi: 10.48550/arXiv.2209.12711]
46. Truong TH, Baldwin T, Verspoor K, Cohn T. Language models are not naysayers: an analysis of language models on negation benchmarks. ArXiv. Preprint posted online on June 14, 2023 [FREE Full text] [doi: 10.48550/arXiv.2306.08189]

XSL•FO
RenderX

47. Reimers N, Gurevych I. Sentence-BERT: sentence embeddings using siamese BERT-networks. ArXiv. Preprint posted online on August 27, 2019 [FREE Full text] [doi: 10.48550/arXiv.1908.10084]
48. Liu Y, Ott M, Goyal N, Du J, Joshi M, Chen D, et al. RoBERTa: a robustly optimized BERT pretraining approach. ArXiv. Preprint posted online on July 26, 2019 [FREE Full text] [doi: 10.48550/arXiv.1907.11692]
49. Mikolov T, Chen K, Corrado G, Dean J. Efficient estimation of word representations in vector space. ArXiv. Preprint posted online on January 16, 2013 [FREE Full text] [doi: 10.48550/arXiv.1301.3781]
50. Le Q, Mikolov T. Distributed representations of sentences and documents. 2014 Presented at: Proceedings of the 31st International Conference on International Conference on Machine Learning - Volume 32; June 21-26; Beijing, China.
51. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. 2017 Presented at: Proceedings of the 31st International Conference on Neural Information Processing Systems; December 4-9; Long Beach, CA. [doi: 10.5555/3295222.3295349]
52. Devlin J, Chang MW, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. 2019 Presented at: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies; June 2-7; Minneapolis, MN.
53. Biber D. Variation Across Speech and Writing. Cambridge, England: Cambridge University Press; 1988.
54. Tauroza S, Allison D. Speech rates in British English. Appl Linguist 1990 Mar 01;11(1):90-105. [doi: 10.1093/applin/11.1.90]
55. Cohen SP, Vase L, Hooten WM. Chronic pain: an update on burden, best practices, and new advances. Lancet 2021 May 29;397(10289):2082-2097. [doi: 10.1016/S0140-6736(21)00393-7] [Medline: 34062143]
56. Conover WJ. Practical Nonparametric Statistics. 3rd edition. New York, NY: John Wiley & Sons; 1999.
57. Munday I, Kneebone I, Newton-John T. The language of chronic pain. Disabil Rehabil 2021 Mar;43(3):354-361. [doi: 10.1080/09638288.2019.1624842] [Medline: 31170836]
58. Vachon-Presseau E, Berger SE, Abdullah TB, Huang L, Cecchi GA, Griffith JW, et al. Brain and psychological determinants of placebo pill response in chronic pain patients. Nat Commun 2018 Sep 12;9(1):3397 [FREE Full text] [doi: 10.1038/s41467-018-05859-1] [Medline: 30209286]
59. Fadnavis S, Dhurandhar A, Norel R, Reinen JM, Agurto C, Secchettin E, et al. PainPoints: a framework for language-based detection of chronic pain and expert-collaborative text-summarization. arXiv. Preprint posted online on September 14, 2022 [FREE Full text] [doi: 10.48550/arXiv.2209.09814]
60. Nunes D, Ferreira-Gomes J, Oliveira D, Vaz C, Pimenta S, Neto F. Chronic pain patient narratives allow for the estimation of current pain intensity. 2023 Presented at: Proceedings of the 36th International Symposium on Computer-Based Medical Systems; June 22-24; L'Aquila, Italy URL: https://ieeexplore.ieee.org/document/10178746?signout=success [doi: 10.1109/CBMS58004.2023.00306]
61. Nunes DA, Ferreira-Gomes J, Neto F, de Matos DM. Computational analysis of the language of pain: a systematic review. ArXiv. Preprint posted online on April 24, 2024 [FREE Full text] [doi: 10.48550/arXiv.2404.16226]

Abbreviations

CLBP: chronic low back pain
FDR: false discovery rate
IMMPACT: Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials
KS: Kolmogorov-Smirnov
LLM: large language model
MPQ: McGill Pain Questionnaire
NLP: natural language processing
NRS: numerical rating scale
PROMIS: Patient-Reported Outcomes Measurement Information System
REDCap: Research Electronic Data Capture
VAS: visual analog scale

Edited by A Kushniruk; submitted 08.07.25; peer-reviewed by FL Neto, SP Brubacher; comments to author 12.08.25; revised version received 10.09.25; accepted 16.10.25; published 25.11.25.

Please cite as:
Norel R, Gewandter J, Zhang Z, Tahsin A, Abdallah CG, Markman J, Duan Z, Cecchi G, Geha P
Turning Patients' Open-Ended Narratives of Chronic Pain Into Quantitative Measures: Natural Language Processing Study
JMIR Hum Factors 2025;12:e80269
URL: https://humanfactors.jmir.org/2025/1/e80269
doi:10.2196/80269
PMID:

XSL•FO
RenderX

©Raquel Norel, Jennifer Gewandter, Zhengwu Zhang, Anika Tahsin, Chadi G Abdallah, John Markman, Zhiyao Duan, Guillermo Cecchi, Paul Geha. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 25.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Adolescent Voices in Action—Co-Designing Digital Sexual and Reproductive Health Knowledge Translation Interventions: Community-Based Participatory Action Project

Salima Meherali[1*], PhD; Amyna Ismail Rehmani[1*], MN; Mariam Ahmad[1*], MPH

Faculty of Nursing, College of Health Sciences, University of Alberta, Edmonton, AB, Canada

[*]all authors contributed equally

Corresponding Author:
Salima Meherali, PhD
Faculty of Nursing
College of Health Sciences
University of Alberta
Edmonton Clinic Health Academy
11405 87 Avenue
Edmonton, AB, T6G 1C9
Canada
Phone: 1 780 248 1401
Email: meherali@ualberta.ca

Abstract

Background: Adolescents need comprehensive education and resources to promote their sexual and reproductive health (SRH) and make informed decisions about their SRH. However, many adolescents fail to secure an opportunity to learn accurate and reliable SRH information, as they face sociocultural barriers, fear of prejudice, and societal stigma. Information available on digital platforms may not always be evidence-based and can further spread misinformation. Digital knowledge translation interventions that provide evidence-based SRH information in North America are limited.

Objective: This study aimed to co-design a digital knowledge translation intervention, reflective of adolescents' voices, to provide evidence-based, accessible, and accurate SRH resources.

Methods: We conducted a large multisite project across 3 cities in Canada. Using a community-based participatory approach and principles of human-centered design, we established adolescent advisory groups (AAGs) to actively engage them in the design, development, and implementation of the intervention.

Results: A total of 26 participants were recruited from Edmonton, Vancouver, and Toronto to be part of the advisory groups. AAG members participated in design-thinking sessions to brainstorm ideas for website design, identify informational needs for content development, provide iterative feedback on the design of the intervention, and suggest strategies to improve engagement and interaction. With their input, 6 priority areas were identified to develop SRH resources (ie, puberty, menstruation, sexually transmitted infections, healthy relationships, sexual assault, and contraception). Their feedback informed the design's language, visual appeal, and engagement factors. To promote meaningful engagement of AAGs at each step, we used strategies such as gamification activities, group discussions, and flexible scheduling, resulting in high retention and ownership of the process among AAG members.

Conclusions: Co-designing with adolescents strengthened the intervention's cultural relevance, youth-friendliness, and credibility. Our process emphasizes the significance of involving adolescents in co-designing SRH interventions, which results in more meaningful, long-term, and youth-friendly solutions.

(JMIR Hum Factors 2025;12:e71148) doi:10.2196/71148

KEYWORDS

co-design; digital interventions; sexual and reproductive health; youth-friendly; knowledge translation

XSL•FO
RenderX

Introduction

Adolescence is an essential stage of growth and development that sets the foundation for all the consecutive stages of life [1]. For adolescents to grow healthy and develop successfully, they require access to information about their health and well-being, including comprehensive sexual and reproductive health (SRH) education [2]. As adolescents begin to explore their identities and experience significant changes (eg, puberty), SRH education becomes fundamental to protect them from negative SRH consequences, such as unintended pregnancy and sexually transmitted infections (STIs), and equip them to improve their sexual health and well-being (eg, healthy relationships) [3,4]. However, adolescents do not always have the opportunity to access reliable and accurate SRH information and often face barriers, such as stigmatization, discrimination, provider prejudice, and embarrassment while seeking help or accessing SRH resources [5,6]. These barriers are deeply rooted in sociocultural norms that stigmatize healthy sexuality and SRH, particularly among females; lesbian, gay, bisexual, transgender, queer or questioning, and plus (others; LGBTQ+) youth; and those from immigrant communities [7,8].

A significant challenge that these adolescents encounter is the lack of reliable and trustworthy sexual health information and services, which can be attributed to restrictive policies, caregiver's or provider's unwillingness to openly discuss sexual health issues, parental control, and lack of confidentiality [5,6,9]. Such challenges are intensified for adolescents from migrant backgrounds as they struggle with balancing cultural, religious, and familial expectations pertaining to their SRH, systemic issues (eg, language barriers and lack of culturally safe resources), and navigating access to SRH services dedicated to immigrant and marginalized populations [7,10]. Previous research with immigrant adolescents in Canada has documented how they often encounter structural barriers within health and education systems, which limit their access to SRH services and lead them to rely on informal digital sources (eg, Google and YouTube [Google]), which may not be evidence-based and can promote misinformation [4,8].

Digital knowledge translation (KT) interventions have the potential to offer promising and equitable venues for influencing various health practices while addressing barriers, including stigma, confidentiality concerns, and difficulties in receiving support [5,9,11,12]. KT interventions facilitate the dissemination of research-based information to health care consumers in a user-friendly format to meet information needs, including but not limited to mobile apps, websites, multimedia content (videos, infographics, etc), and social media platforms [13]. Particularly, in the realm of SRH, digital KT interventions have the potential to reach vulnerable and marginalized populations while providing anonymity and privacy on sensitive SRH matters [5,6,14]. A systematic review on the effectiveness of digital sexual health interventions reported that a higher proportion of website-based interventions were found to improve sexual health outcomes in adolescents, which was demonstrated by an increase in knowledge and access to sexual health services, improvement in STI prevention practices, and reduction in unhealthy sexual behaviors [15]. We conducted an environmental scan to understand the range of mobile health apps dedicated to adolescent SRH in Canada and found that very few apps offered comprehensive, reliable, and evidence-based information [16]. Furthermore, the findings combined with other literature in the field also highlighted gaps in cultural adaptations within such interventions and the need for youth-centered design in SRH digital intervention development [7,16,17].

Co-design is a participatory action research approach that brings together community members, stakeholders, and researchers to collaboratively develop, design, and implement interventions to address issues within the community [17,18]. This approach aims to actively and meaningfully engage individuals in the research process to understand their needs and integrate their experiences to design innovative solutions and improve outcomes [19,20]. This is particularly significant when working with adolescents, as it allows researchers to gain insights into their unique and creative perspectives and determine what would work best for them. Engaging adolescents as equal partners in the design process empowers them to take ownership and responsibility for the process, promoting the success of an intervention. The World Health Organization's framework for planning, developing, and implementing youth-centered digital solutions recommends that for digital interventions to succeed and be sustainable, adolescents should be meaningfully engaged at each step of the planning, design, and implementation process [2]. This ensures that the intervention is adolescent-friendly and accessible, acceptable, equitable, appropriate, and effective among youth [2]. The co-design approach is uniquely positioned to bridge gaps for immigrant adolescents by incorporating cultural perspectives and SRH information needs, addressing issues of trust and accessibility, and ensuring that the style and content of information delivery resonate with their lived realities [17].

Co-designed health initiatives are gaining traction for influencing health behaviors and improving health outcomes [12,17,21]. Several studies have documented the benefits of digital KT interventions in promoting health and well-being among adolescents, particularly with SRH outcomes [5,7,22,23]. Regardless, digital interventions that provide comprehensive, evidence-based, and reliable SRH education for adolescents, especially from migrant backgrounds, are somewhat limited in North America [16]. Also, a few studies have explored co-design in the context of immigrant adolescents' SRH, especially within Canada. Therefore, our aim was to co-design adolescent-friendly digital SRH KT interventions with and for immigrant adolescents in Canada to provide easy access to resources and services at their convenience with anonymity. In this study, we outline the process of engaging immigrant adolescents in the co-design of an evidence-based website as a KT intervention aimed at promoting SRH information.

Methods

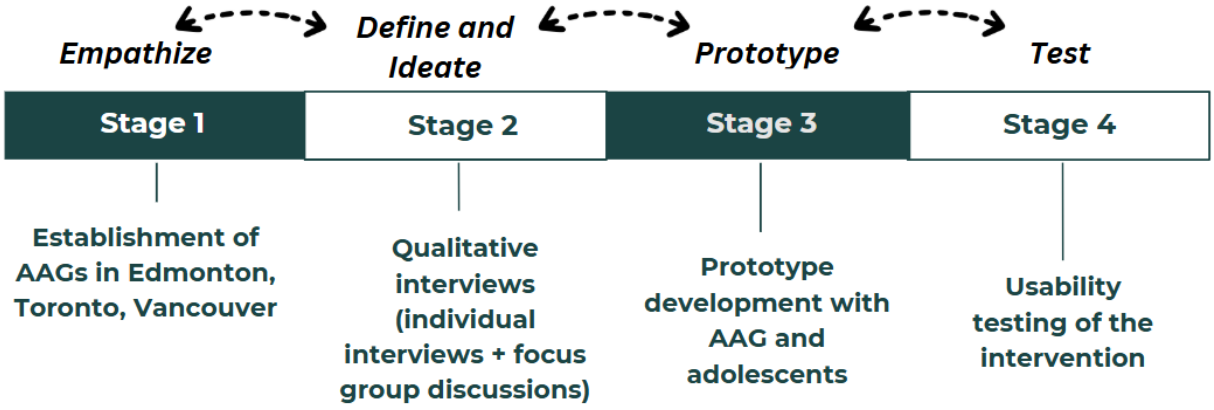
Study Design

We conducted a larger multisite project (2023-2024), intending to co-design digital SRH interventions for immigrant adolescents in Canada. Conducting a multisite project helped to ensure that the intervention would reach immigrant adolescents across

XSL•FO
RenderX

Canada. We integrated the principles of community-based participatory action research and human-centered design (HCD) to maximize end user engagement at each step of the co-design process, which guaranteed that the outcome would be effective and sustainable [17,24]. HCD, which is also referred to as design thinking, is a creative and systematic problem-solving approach that prioritizes the needs, desires, and experiences of individuals at the core [24]. It not only focuses on the functionality of a certain solution but also considers the user's motivation, behavior, and challenges, ensuring that it is culturally relevant [24]. Building on these principles, we divided our project into 4 stages (Figure 1). This process helped us to keep adolescents' perspectives and needs at the center of the co-design process and maximize collaboration, participation, and engagement [18].

Figure 1. Stages of the project. AAG: adolescent advisory group.



Stage 1: Establishment of Adolescent Advisory Groups

We first developed and convened adolescent advisory groups (AAGs) in Edmonton, Toronto, and Vancouver regions. These cities have one of the highest percentages of immigrant populations in Canada [25]. Our purpose was to develop a mutual partnership with immigrant adolescents to actively involve them in various research activities, such as data collection, data analysis, resource development, concurrent feedback, and knowledge mobilization.

Research Team Composition

The research team, associated with recruitment and training of AAGs and coordinating co-designing activities, included the principal investigator (PI; SM), the project coordinator (MA), and a research assistant (AIR). The PI who conceived this project has extensive experience with community-based research methodologies and adolescent SRH initiatives with a desire to promote equitable access to SRH resources and services, especially among marginalized populations. We partnered with the Department of Computing Science at the University of Alberta to seek professional support and expertise in co-designing digital resources. The Vice Dean, along with graduate research assistants, dedicated their skills and knowledge to develop the website prototype, incorporate user feedback and preferences, and provide technical guidance at each step of the project. The research assistant (AIR) and project coordinator (MA) were responsible for developing content, connecting with SRH experts for content validation, organizing AAG meetings and activities, coordinating with the team in the Department of Computing Science to track progress, and completing research coordination tasks. With this mutual collaboration, we were able to design our digital intervention effectively and consult on technological-related issues.

Setting and Participants

Through multifaceted strategies, we recruited adolescents aged 12-19 years from various cultural and ethnic backgrounds, ages, and genders, and invited them to participate in our AAG meetings. We included adolescents who were either born outside of Canada and immigrated to Canada (first-generation immigrants) or who were born in Canada but had at least 1 foreign-born parent (second-generation immigrants). Our inclusion criteria required that participants reside in Canada at the time of the study, be able to communicate in English, and have access to a device with internet access to connect with the research team online. Research assistants, in collaboration with various youth organizations, community networks, and adolescent-friendly clinics, recruited adolescents to be part of the AAGs. We also promoted recruitment posters through social media advertising (eg, Facebook [Meta] and Instagram [Meta]) to target our population. We made efforts to develop recruitment materials and strategies to be inclusive of LGBTQ+ youth and those living with disabilities. After screening and recruitment of potential participants, the included AAG members were given a comprehensive explanation of the study objectives, its intended outcomes, risks and benefits, and implications. They were provided with a detailed information sheet, highlighting the particulars about the study and their participation expectations.

The AAGs were established based on practical considerations (eg, the number of adolescents interested in participating) and as per the recommendations of research team members, who had extensive experience working with AAGs. We decided that each advisory group at each site should consist of a reasonable target size of 8-12 adolescents to result in meaningful engagement (expected total AAG participants: 24-30). This sample size enabled us to exercise deeper relationship-building, promoted active participation, and gave opportunity to each member to express their voice. Literature on participatory

XSL•FO
RenderX

research design encourages smaller group sizes to maximize closer collaboration, equitable participation, and maintain a balance between diverse perspectives and group dynamics [17,24]. AAG members in each city were expected to meet online on Zoom (Zoom Communications) throughout the project and share their insights through design-think sessions on the development of the website and SRH resources through a systematic process.

Procedures (Co-Design Process)

To foster meaningful engagement, AAG members were provided with initial training on SRH concepts, qualitative and quantitative research methods (types of research methods, data collection, and interview methods), social determinants of health, and the process of adolescent engagement. An important component of the training sessions was to evaluate the effectiveness of the training, which we assessed through various strategies. For example, we asked AAG members to develop some sample open-ended interview questions to demonstrate their understanding of ethical research practices and their comprehension of research concepts. Each site AAG meeting or session was held separately and was scheduled based on members' availabilities. The sessions started in February 2023. After their training and familiarization with the project, the AAG members met regularly once every month, starting in July 2023 until January 2024 for approximately 60-90 minutes to discuss project updates, recruitment activities, and planning of further activities. As we developed the prototype and started an iterative cycle of feedback and integrating changes, the meetings were organized every other month until August 2024. All the meetings were conducted online over Zoom and were audio-recorded and transcribed verbatim after identifying information was removed. Furthermore, the research assistant and project coordinator maintained detailed notes and meeting summaries to document the co-design process and AAG feedback to capture their input in real time and ensure that their feedback was accurately reflected at each step.

The meetings were moderated by the PI (SM), project coordinator (MA), and research assistant (AIR). The moderators prompted active participation from all members through various techniques (eg, blackboard activities, quizzes, and feedback forms), to ensure that each participant was able to share their thoughts, feelings, and experiences in a safe space. Given the sensitive nature of SRH topics, we ensured to apply a trauma-informed and youth-centered facilitation approach. All team members were trained to identify any signs of discomfort and distress, and a distress protocol was developed to guide facilitators in responding to signs, including pausing the discussion and offering follow-up support. Participants were encouraged to join from private spaces and were reminded of their right to skip any questions, take breaks, or keep their cameras off. Recognizing the potential for power imbalances between researchers and AAG members, we implemented measures to enhance equitable participation, such as using youth-friendly language, giving options for providing feedback, ensuring all feedback was equally valued, and regularly consulting members on meeting structure, preference, and decision-making.

All members tried to attend all meetings, considering their academic commitments and other scheduling conflicts, but all members remained part of the advisory group throughout the project. Despite the occasional absences, the members were proactive in following up with the research team and identified ways to contribute their insights either during sessions or through asynchronous feedback. Throughout the AAG meetings, we facilitated healthy discussions about the learning and informational needs of adolescents, the cultural or social influences on SRH education, and the quality of the SRH education curriculum offered in schools.

Data Analysis

We used a descriptive qualitative approach to synthesize feedback and identify recurrent ideas related to SRH information needs, engagement strategies, and co-design experiences. Meeting transcripts, facilitator logs, and session summaries were reviewed by our research team to identify recurring topics, key insights, and decisions made collaboratively with youth. Two researchers (AIR, MA) independently reviewed the session notes to ensure consistency in interpretation and enhance analytic trustworthiness through iterative discussion and consensus. We kept a detailed log of all AAG interactions and systematically organized the information into categories to support the decision-making process. After evaluating the findings from these sessions, as well as the qualitative interviews conducted in stage 2 (findings are published elsewhere), we finalized priority areas to develop SRH educational content.

Ethical Considerations

We received ethics approval for this project from the University of Alberta Research Ethics Board (August 26, 2021; Pro00113664), the Toronto Metropolitan University Research Ethics Board (April 27, 2022; REB 2022-069), and the University of British Columbia (May 31, 2022; #H22-00395). We followed the Tri-Council Policy Ethical Conduct for Research Involving Humans [26] principles for seeking consent, which does not specify an age of consent for children. Following these guidelines, we determined whether adolescents had the capacity to understand the objectives, significance, implications, risks, and benefits of the study by explaining these topics and asking them to reiterate based on their understanding. The ethics committee from all 3 boards approved and stipulated that if participants were capable enough to understand the procedures and study details, they could provide consent for themselves. After addressing any questions, participants were asked to sign a consent form to confirm their voluntary participation as AAG members, along with a confidentiality agreement to protect the identities of fellow members and the experiences shared within sessions. All participants were deemed capable of providing informed consent for themselves. Participants were given an honorarium of CAD $20 (approximately US $14) per hour for participating in various activities.

Results

Participant Characteristics

We recruited a total of 26 participants for this stage of the study, with 88.4% (n=23) of participants identifying as female and

XSL•FO
RenderX

11.5% (n=3) as male. Out of the 26 participants, 53.8% (n=14) of participants were first-generation immigrants and 46.1% (n=12) were second-generation immigrants. Almost half of the participants were 19 years of age, whereas only 1 participant was 15 years old. None of the participants formally withdrew or disengaged from the project. The demographic information of participants is presented in Table 1.

XSL•FO
RenderX

Table 1. Demographic characteristics of adolescent advisory group (n=26).

Characteristics	Participants, n (%)
Sex	
Female	23 (88.4)
Male	3 (11.5)
Age (y)	
15	1 (3.8)
16	2 (7.6)
17	3 (11.5)
18	7 (26.9)
19	13 (50)
Place of birth	
Northern America	12 (46.1)
Sub-Saharan Africa	5 (19.2)
Asia	7 (26.9)
Europe	1 (3.8)
South America and Caribbean	1 (3.8)
Languages spoken at home (multiple responses)	
English	25 (96.1)
Urdu	4 (15.3)
Hindi	3 (11.5)
Cantonese	1 (3.8)
Farsi	1 (3.8)
Vietnamese	1 (3.8)
Others (Tamil, Dari, Bengali, Gujarati, Somalian, and Tagalog)	6 (23)
Number of years in Canada	
1-3	2 (7.6)
4-9	6 (23)
10 or more	6 (23)
Lived in Canada all or most of life	12 (46.1)
Current level of education	
Grade 7-10	1 (3.8)
Grade 11	2 (7.6)
Grade 12	3 (11.5)
Postsecondary school	20 (76.9)
Place of birth of parents	
Mother	
Northern America	1 (3.8)
Sub-Saharan Africa	10 (38.4)
Asia	13 (50)
Europe	1 (3.8)
South America and Caribbean	1 (3.8)
Father	
Sub-Saharan Africa	10 (38.4)

XSL•FO
RenderX

Characteristics	Participants, n (%)
Asia	14 (53.8)
Europe	1 (3.8)
South America and Caribbean	1 (3.8)

Design Process

To initiate the development of SRH digital resources offered through our website, we consulted AAG members, stakeholders, and the interdisciplinary research team iteratively to co-develop the best course of action for planning, designing, and implementation. We started with brainstorming potential names for our resource website and its initial layout and design. After discussing many acronyms, we collectively decided to name our website PEER, which stands for Partnership to Empower adolescent Engagement in Research. This not only emphasizes the role that adolescents play in promoting research initiatives but also signifies our work to empower adolescents to drive changes by forming partnerships. We also reflected on various themes for our website and which color would be the most appealing to adolescents. Although some participants shared that bright colors (eg, dark blue and yellow) would be captivating, others commented that these colors might only complement some text and graphics. Therefore, we collectively decided on a neutral color tone (ie, light purple) with a white background. Our development team designed some sample logos and design patterns for the website, which were shared with AAG members.

Development and Delivery of Content

We finalized 6 priority areas to focus our SRH content on: puberty, menstruation, STIs, sexual assault, contraception, and healthy relationships. Our team developed evidence-based resources for every topic. We ensured all topics were covered at length and that common misconceptions regarding these topics were addressed. We identified the information to be delivered within each topic with the help of AAG members. We wanted the content to be informative and engaging at the same time. Hence, we did not rely solely on text-based information but also developed infographics, animations, and videos along with textual information to keep users engaged. We collaborated with 6 stakeholders and experts in the field (SRH researchers, clinicians, sexual health consultants, and SRH or public health nurses) to validate the content and ensure that it is factual, evidence-based, and unbiased. With their feedback, we were able to improve the language and delivery of the content as well as rectify any errors. For instance, validators suggested using language that does not blame an individual and is inclusive and gender-neutral (eg, menstruating people instead of menstruating girls).

Following content validation, we shared the SRH resources with our AAG members (sample content in Figure 2). They appreciated the simple language used to explain complex issues and the different formats that we used to deliver information. They also provided valuable feedback to improve our resources. We initially developed our content videos with audio voice-over using computer software, but members found the voice to be unnatural and robotic. They suggested that the research team record the voice-over instead to give it a more human touch. We welcomed their feedback, and a member of our team did the voiceover. In addition, members re-emphasized that the text under each category should be concise, in bullet form, and have graphics throughout. We considered their suggestions and tried to keep the information concise and bite-sized. Regarding the display of content on the website, we initially organized the subtopics (headings) through horizontal widgets. However, many AAG members did not find this format appealing and suggested that the widgets be organized vertically instead.

XSL•FO
RenderX

Figure 2. Examples of youth-centered sexual and reproductive health content developed for the Partnership to Empower Adolescent Engagement in Research platform.

(A) Content on contraception



(B) Content on puberty





(C) Content on healthy relationships



Integrating all the feedback from the AAG and stakeholders, we finalized the first prototype of our website. We agreed to display a video instead of a still image on the landing page as it seemed more attractive and engaging to the users (Figure 3). Also, we made the edits suggested by the AAG members within the SRH resources and rearranged some tabs on the website. As per their suggestions, we finalized the display and organization of the content. A sample of these resources is presented in Figure 2.

XSL•FO
RenderX

Figure 3. Landing page of the Partnership to Empower Adolescent Engagement in Research digital platform.



We implemented various strategies to collect comprehensive feedback from AAGs on the prototype of the website and SRH resources. Some participants shared detailed feedback through Zoom sessions, whereas others also agreed to provide written feedback through Google Forms. We organized the feedback received from the members and collectively decided whether to implement the change or not. Some features were specifically difficult to implement at this stage as they required advanced technological adaptations (eg, helper chat). Table 2 illustrates some examples of feedback from our members and what steps were taken to address the suggestions.

Table 2. Feedback and response.

Prototype feedback	Team decision
"I would like for the bottom tab to be cleaner with the Land Acknowledgement and the Social links. YouTube is the only icon on a new level and the Land Acknowledgement left aligned text with the title on the left is a bit of putting. Maybe adding this acknowledgement on the About Us page makes more sense."	All the icons at the bottom of the page were aligned on the same line and adjusted to provide a better interface.
"I think the infographics should follow a color theme and be cohesive in one section. The headings could be bolded to make them stand out. All the infographics can be centered in the middle of the webpage."	Headings were bolded and graphics centered on the page. Fonts kept consistent across topics.
"Maybe a helper chat or who could answer any questions."	Feature cannot be added at this point.
"I think adding some basic information about some common reproductive issues/disorders such as PCOS or endometriosis would also help adolescents because some girls get irregular periods, which becomes scary."	Topic of menstruation expanded to include content on common menstrual disorders (eg, PCOS).
"To make the website better, I would maybe move the search bar upwards to somewhere easier for the viewer to see, instead of having to scroll down. That way in every section of the site you can search what you're looking for."	Search bar added to the top right corner as well.
"Change the color of the font. The font of the subtext on the website is very light gray and doesn't stand out as much. It is a little hard on the eyes to read."	All the text font change to black color.
"Include an option to quickly exit the website if learning SRH content and need to close the tab immediately (exist fast button)."	"Discretely exit website" button added to redirect user to a different web page.
"Make it more 'adolescent friendly' by adding brighter color schemes, more stimulating buttons, and more enticing for teens to want to be interested in."	More colors added with brighter graphics and fonts; each individual tab with a different brighter color.

Discussion

Principal Findings

Our project aimed to co-design adolescent-friendly digital SRH resources by actively engaging AAG. This process provided valuable insights into best practices for adolescent participation in digital health intervention development. The formation of our AAGs follows best practices for meaningful youth engagement, as recommended in the literature [17-20,27]. We ensured that meaningful youth engagement was not only fundamental but deeply ingrained in our methodology. By

XSL•FO
RenderX

empowering adolescents to actively participate in the design and execution of our intervention, we pushed beyond tokenism and manipulation, which can hinder genuine engagement [17,27]. Instead, adolescents were involved as co-creators, with their ideas, input, and leadership influencing every stage of creation.

We recruited a diverse group of adolescents aged 12-19 years from various cultural and ethnic backgrounds through schools, community organizations, and social media platforms. This approach represented different perspectives across age groups, cultures, and gender identities [28,29]. To prepare AAG members for effective contributions, we conducted workshops introducing key research concepts and SRH topics. This training addressed a gap in previous studies where adolescent advisory members often lack formal preparation [30,31]. Armed with this foundational knowledge, AAG members were then engaged in the co-design process using various participatory techniques. Throughout the AAG meetings, we used engagement methods tailored to different age groups, including gamified activities for younger adolescents and in-depth discussions for older teens. We also used structured feedback forms and open discussions. This diverse approach is supported by research showing that creative engagement methods can enhance youth involvement and input [24]. The iterative nature of our process, with regular meetings and opportunities for feedback, allowed for continuous refinement of the intervention based on adolescent input, a fundamental principle of co-design [32].

Our approach incorporated several effective strategies from previous studies in website development. We involved adolescents in naming the intervention, choosing color schemes, and determining content formats, aligning with recommendations for youth-centered design in digital health interventions [33]. Prototypes and iterative feedback cycles mirrored successful co-design processes described in other digital health projects targeting youth [12,20,21,34]. For example, the study by Wright et al [23] highlights how youth advisors in their study broadened the scope of their proposed digital intervention through an iterative co-design approach and helped to bring forward an innovative strategy to improve information-seeking behaviors. Our experience highlighted the importance of flexibility when working with adolescents. Offering online participation options and accommodating schedules enhanced engagement, consistent with findings from other youth advisory group studies [35]. In addition, providing compensation for AAG members' time and contributions promoted long-term engagement and valued their input.

AAG members played a crucial role in evaluating content readability and suggesting more adolescent-friendly language, addressing the need for accessible digital health resources. Their involvement in discussions about privacy features and discreet access options aligned with research emphasizing the importance of confidentiality in adolescent-oriented sexual health interventions [36]. For instance, AAG members suggested implementing a "quick exit" feature and disguising the website icon, which was incorporated into the final design to enhance user privacy. During pilot testing, users highly praised these features, with many expressing that these privacy measures made them feel more comfortable using these resources. The AAG members also expressed satisfaction with how their suggestions were implemented, reinforcing the value of their input in the design process.

In addition to the significance of the co-design process in our study, the content topics that were identified and co-developed offer unique insights into the SRH informational needs of immigrant adolescents in Canada. Participants emphasized that although puberty, menstruation, and bodily changes were basic SRH topics, they would appreciate culturally grounded discussions and insights into these topics (eg, tampons and myths). They also highlighted that immigrant adolescents do not always get the opportunity to learn about topics such as contraception or intimate relationships, and navigating access to confidential care, which is also supported by our previous studies and other literature [10,37]. Although platforms in Canada are increasingly adapting their approach to include culturally sensitive information, not many are designed specifically with immigrant adolescents in mind [13,16,17]. This further indicates the significance of this digital intervention.

While our process yielded valuable insights, it also presented challenges. Balancing adolescent preferences with evidence-based practices sometimes requires careful negotiation, which aligns with challenges noted in other co-design studies [20]. For example, when adolescents suggested including popular but scientifically unsupported remedies, we had to find ways to address these interests while maintaining medical accuracy. We navigated these situations by explaining the importance of scientific evidence, acknowledging the popularity of these remedies, and collaboratively finding ways to address the underlying concerns that made these remedies appealing. This approach ensured that adolescents felt heard and valued, even when their initial suggestions were not feasible. Unexpectedly, these negotiations led to innovative ways of communicating complex medical information in adolescent-friendly language, enhancing the overall accessibility of the content. Furthermore, to ensure the sustainability of adolescent involvement, we have established a rotating membership system for the AAG and are developing a mentorship program where experienced members can guide newcomers. This mentorship program will include training sessions for mentors, regular check-ins, and a 6-month mentorship cycle to balance continuity with fresh perspectives.

Future Implications

Preliminary feedback on the intervention has been positive, with pilot users reporting high levels of engagement and improved SRH knowledge. Specifically, users praised the website's discreet design, easy-to-understand language, and engaging features. However, long-term effectiveness studies are still ongoing, and findings from the usability testing are published elsewhere. These studies will track key metrics, including SRH knowledge retention, user satisfaction, frequency of website usage, and self-reported behavior change related to SRH practices. By monitoring these outcomes, we aim to comprehensively evaluate the intervention's impact and identify areas for future improvement.

XSL•FO
RenderX

Limitations

Through collaborative efforts in our co-design process, we actively and meaningfully engaged immigrant adolescents as active partners in developing the PEER digital SRH platform. However, our process also had its limitations. First, while our recruitment efforts aimed to include individuals from diverse gender identities and backgrounds, including LGBTQ+ youth and adolescents with disabilities, our advisory group lacked sufficient representation from these groups. Although this raises the question of the generalizability of our content to diverse groups, we ensured that our content was inclusive and included gender-neutral language. This was strongly emphasized by our AAG members and SRH experts working with adolescents. Second, our advisory group was predominantly female (n=23, 88%), and more than half of the participants were aged 18-19 years old. This demographic skew may limit the generalizability of our findings and could influence the content developed for the platform. The predominance of female voices may have emphasized certain gendered experiences and needs (eg, content on menstruation and pubertal changes), while perspectives of male and nonbinary youth may be underrepresented. To address this limitation, we intentionally explored the SRH information needs of male and gender-diverse adolescents by encouraging participants to discuss challenges faced by their peers across different gender identities. However, it can also be considered as a strength as participants were able to offer greater reflection based on their capacity to reflect on their previous experiences with SRH education. Our future research will aim to engage and seek a more gender-diverse group to evaluate the effectiveness of the intervention design and incorporate recommendations to reflect the full spectrum of adolescent SRH experiences and needs. Finally, maintaining consistent engagement over time and managing diverse opinions within the group required significant resources and facilitation skills from the research team. Future projects could benefit from establishing clear guidelines for resolving conflicts between adolescent preferences and evidence-based practices and developing long-term strategies for long-term engagement.

Conclusion

Adolescents need access to accurate and reliable SRH information for positive sexual health outcomes and to prevent unintended consequences, such as STIs, unintended pregnancies, and so on. However, they do not always have the necessary means to secure such information. Digital interventions have the potential to influence health behaviors and promote positive outcomes by reaching a wide range of audiences, especially those who are at risk of being distanced from mainstream services. We used the principles of community-based participatory action research and HCD to co-design digital SRH interventions with adolescents by keeping their needs at the center of the process. By giving them a voice in the process and the ability to influence decision-making, we were able to co-develop a resource that has the potential to be user-friendly, accessible, effective, and sustainable among immigrant adolescents. This process allows the cocreation of culturally relevant interventions that have the potential to reach vulnerable populations, particularly in conservative societies where stigma and misinformation are pertinent.

Acknowledgments

We would like to acknowledge Eleni Stroulia (Vice Dean of the Department of Computing Science, University of Alberta) and Nilanjan Ghatak (Graduate Research Assistant) for their contributions toward the digital intervention development. This work was supported by the Women and Children's Health Research Institute (WCHRI) and the Canadian Institute of Health Research (CIHR) (Grant # RES0053962).

Data Availability

The datasets used and/or analyzed during this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

References

1. Adolescent health. World Health Organization. 2019. URL: https://www.who.int/health-topics/adolescent-health#tab=tab_1 [accessed 2025-10-18]
2. Sexual and Reproductive Health and Research (SRH). Youth-centred digital health interventions: a framework for planning, developing and implementing solutions with and for young people. Geneva, Switzerland: World Health Organization; 2020.
3. Salam RA, Faqqah A, Sajjad N, Lassi ZS, Das JK, Kaufman M, et al. Improving adolescent sexual and reproductive health: A systematic review of potential interventions. J Adolesc Health 2016;59(4S):S11-S28 [FREE Full text] [doi: 10.1016/j.jadohealth.2016.05.022] [Medline: 27664592]
4. Meherali S, Louie-Poon S, Idrees S, Kauser S, Scott S, Salami B, et al. Understanding the sexual and reproductive health needs of immigrant adolescents in Canada: a qualitative study. Front Reprod Health 2022;4:940979 [FREE Full text] [doi: 10.3389/frph.2022.940979] [Medline: 36303669]
5. Feroz A, Abrejo F, Ali SA, Nuruddin R, Saleem S. Using mobile phones to improve young people's sexual and reproductive health in low- and middle-income countries: a systematic review protocol to identify barriers, facilitators and reported interventions. Syst Rev 2019;8(1):117 [FREE Full text] [doi: 10.1186/s13643-019-1033-5] [Medline: 31103044]

XSL•FO
RenderX

6. Sao SS, Barre-Quick M, Yu R, Abboud S, Coleman JS. Advancing access to care through digital health: perspectives from youth on a novel platform to increase access to sexual and reproductive health care and education for adolescents and young adults. J Pediatr Adolesc Gynecol 2023;36(5):449-454 [FREE Full text] [doi: 10.1016/j.jpag.2023.04.002] [Medline: 37084876]
7. Meherali S, Hussain A, Rahim KA, Idrees S, Bhaumik S, Kennedy M, et al. Digital knowledge translation tools for sexual and reproductive health information to adolescents: an evidence gap-map. Ther Adv Reprod Health 2024;18:26334941241307881 [FREE Full text] [doi: 10.1177/26334941241307881] [Medline: 39703678]
8. Meherali S, Salami B, Okeke-Ihejirika P, Vallianatos H, Stevens G. Barriers to and facilitators of South Asian immigrant adolescents' access to sexual and reproductive health services in Canada: a qualitative study. Can. J. Hum. Sex 2021;30(3):329-338. [doi: 10.3138/cjhs.2020-0056]
9. Patterson SP, Hilton S, Flowers P, McDaid LM. What are the barriers and challenges faced by adolescents when searching for sexual health information on the internet? Implications for policy and practice from a qualitative study. Sex Transm Infect 2019;95(6):462-467 [FREE Full text] [doi: 10.1136/sextrans-2018-053710] [Medline: 31040251]
10. Meherali S, Rehmani AI, Ahmad M, Kauser S, Scott Fiddler P, Pinzon-Hernandez P, et al. Between cultures and traditions: a qualitative investigation of sexual and reproductive health experiences of immigrant adolescents in Canada. Cult Health Sex 2024:1-17 [FREE Full text] [doi: 10.1080/13691058.2024.2436677] [Medline: 39648708]
11. Malloy JA, Partridge SR, Kemper JA, Braakhuis A, Roy R. Co-design of digital health interventions for young adults: protocol for a scoping review. JMIR Res Protoc 2022;11(10):e38635 [FREE Full text] [doi: 10.2196/38635] [Medline: 36279167]
12. Stoyanov SR, Zelenko O, Staneva A, Kavanagh DJ, Smith C, Sade G, et al. Development of the niggle app for supporting young people on their dynamic journey to well-being: co-design and qualitative research study. JMIR Mhealth Uhealth 2021;9(4):e21085 [FREE Full text] [doi: 10.2196/21085] [Medline: 33877050]
13. Meherali S, Bhaumik S, Idrees S, Kennedy M, Lassi ZS. Digital knowledge translation tools for disseminating sexual and reproductive health information to adolescents: Protocol for an evidence gap map review. JMIR Res Protoc 2024;13:e55081 [FREE Full text] [doi: 10.2196/55081] [Medline: 38349722]
14. Zettergren L, Larsson EC, Hellsten L, Kosidou K, Nielsen AM. Implementing digital sexual and reproductive health care services in youth clinics: a qualitative study on perceived barriers and facilitators among midwives in Stockholm, Sweden. BMC Health Serv Res 2024;24(1):411 [FREE Full text] [doi: 10.1186/s12913-024-10932-1] [Medline: 38566080]
15. Sewak A, Yousef M, Deshpande S, Seydel T, Hashemi N. The effectiveness of digital sexual health interventions for young adults: a systematic literature review (2010-2020). Health Promot Int 2023;38(1):daac104. [doi: 10.1093/heapro/daac104] [Medline: 36757346]
16. Benoit JRA, Louie-Poon S, Kauser S, Meherali S. Promoting adolescent sexual and reproductive health in North America using free mobile apps: environmental scan. JMIR Pediatr Parent 2022;5(4):e33826 [FREE Full text] [doi: 10.2196/33826] [Medline: 36194450]
17. Lassi ZS, Neideck EG, Aylward BM, Andraweera PH, Meherali S. Participatory Action Research for Adolescent Sexual and Reproductive Health: A Scoping Review. Sexes 2022;3(1):189-208. [doi: 10.3390/sexes3010015]
18. Vial S, Boudhraâ S, Dumont M, Tremblay M, Riendeau S. Developing A mobile app with a human-centered design lens to improve access to mental health care (Mentallys Project): protocol for an initial co-design process. JMIR Res Protoc 2023;12:e47220 [FREE Full text] [doi: 10.2196/47220] [Medline: 37606978]
19. Malloy J, Partridge SR, Kemper JA, Braakhuis A, Roy R. Co-design of digital health interventions with young people: a scoping review. Digit Health 2023;9:20552076231219117 [FREE Full text] [doi: 10.1177/20552076231219117] [Medline: 38107986]
20. Raeside R, Jia SS, Todd A, Hyun K, Singleton A, Gardner LA, et al. Are digital health interventions that target lifestyle risk behaviors effective for improving mental health and wellbeing in adolescents? A systematic review with meta-analyses. Adolescent Res Rev 2023;9(2):193-226. [doi: 10.1007/s40894-023-00224-w]
21. Martin A, Caon M, Adorni F, Andreoni G, Ascolese A, Atkinson S, et al. A mobile phone intervention to improve obesity-related health behaviors of adolescents across Europe: iterative co-design and feasibility study. JMIR Mhealth Uhealth 2020;8(3):e14118 [FREE Full text] [doi: 10.2196/14118] [Medline: 32130179]
22. Isaacs N, Ntinga X, Keetsi T, Bhembe L, Mthembu B, Cloete A, et al. Are mHealth interventions effective in improving the uptake of sexual and reproductive health services among adolescents? A scoping review. Int J Environ Res Public Health 2024;21(2):165 [FREE Full text] [doi: 10.3390/ijerph21020165] [Medline: 38397656]
23. Wright M, Thomas KC, Carpenter D, Lee C, Coyne I, Garcia N, et al. Co-designing a website with and for youth, so they can better manage their health. PEC Innov 2023;2:100164 [FREE Full text] [doi: 10.1016/j.pecinn.2023.100164] [Medline: 37384159]
24. Chen E, Leos C, Kowitt SD, Moracco KE. Enhancing community-based participatory research through human-centered design strategies. Health Promot Pract 2020;21(1):37-48. [doi: 10.1177/1524839919850557] [Medline: 31131633]
25. Immigrants and emigrants in Canadian metropolitan areas 2022. Statista. URL: https://www.statista.com/statistics/443728/canada-immigrants-and-emigrants-by-census-%20metropolitan-area/ [accessed 2024-07-30]

XSL•FO
RenderX

26. Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans – TCPS 2. Government of Canada, Interagency Advisory Panel on Research Ethics. 2022. URL: https://ethics.gc.ca/eng/policy-politique_tcps2-eptc2_2022.html [accessed 2023-01-11]
27. Resource Hub. CHOICE for Youth & Sexuality. 2016. URL: https://www.youthdoit.org/assets/Uploads/20171122-Flower-of-Participation-Narrative.pdf [accessed 2025-10-18]
28. Lassi ZS, Kedzior SGE, Tariq W, Jadoon Y, Das JK, Bhutta ZA. Effects of preconception care and periconception interventions on maternal nutritional status and birth outcomes in low- and middle-income countries: a systematic review. Nutrients 2020;12(3):606 [FREE Full text] [doi: 10.3390/nu12030606] [Medline: 32110886]
29. Anyon Y, Bender K, Kennedy H, Dechants J. A systematic review of Youth Participatory Action Research (YPAR) in the United States: methodologies, youth outcomes, and future directions. Health Educ Behav 2018;45(6):865-878. [doi: 10.1177/1090198118769357] [Medline: 29749267]
30. Bhattacharya A, Liang C, Zeng EY, Shukla K, Wong MER, Munson SA, et al. Engaging teenagers in asynchronous online groups to design for stress management. Proc 18th ACM Int Conf Interact Des Child (2019) 2019;2019:26-37 [FREE Full text] [doi: 10.1145/3311927.3323140] [Medline: 32954383]
31. Kaboré SS, Ngangue P, Soubeiga D, Barro A, Pilabré AH, Bationo N, et al. Barriers and facilitators for the sustainability of digital health interventions in low and middle-income countries: a systematic review. Front Digit Health 2022;4:1014375 [FREE Full text] [doi: 10.3389/fdgth.2022.1014375] [Medline: 36518563]
32. Tanner AE, Song EY, Mann-Jackson L, Alonzo J, Schafer K, Ware S, et al. Preliminary impact of the weCare social media intervention to support health for young men who have sex with men and transgender women with HIV. AIDS Patient Care STDS 2018;32(11):450-458 [FREE Full text] [doi: 10.1089/apc.2018.0060] [Medline: 30398955]
33. Brayboy LM, Sepolen A, Mezoian T, Schultz L, Landgren-Mills BS, Spencer N, et al. Girl talk: A smartphone application to teach sexual health education to adolescent girls. J Pediatr Adolesc Gynecol 2017;30(1):23-28 [FREE Full text] [doi: 10.1016/j.jpag.2016.06.011] [Medline: 27393638]
34. Archangeli C, Marti FA, Wobga-Pasiah EA, Zima B. Mobile health interventions for psychiatric conditions in children: a scoping review. Child Adolesc Psychiatr Clin N Am 2017;26(1):13-31. [doi: 10.1016/j.chc.2016.07.009] [Medline: 27837939]
35. Wadham E, Green C, Debattista J, Somerset S, Sav A. New digital media interventions for sexual health promotion among young people: a systematic review. Sex Health 2019;16(2):101-123. [doi: 10.1071/SH18127] [Medline: 30819326]
36. Mwaisaka J, Gonsalves L, Thiongo M, Waithaka M, Sidha H, Alfred O, et al. Young people's experiences using an on-demand mobile health sexual and reproductive health text message intervention in Kenya: qualitative study. JMIR Mhealth Uhealth 2021;9(1):e19109 [FREE Full text] [doi: 10.2196/19109] [Medline: 33448930]
37. Punjani NS, Hegadoren K, Hirani S, Mumtaz Z, Jackson M, Papathanassoglou E. Perceptions and experiences of Pakistani-descent female adolescents living in Canada, on developing sexuality and self-identity. Sexes 2022;3(3):413-433. [doi: 10.3390/sexes3030031]

Abbreviations

AAG: adolescent advisory group
HCD: human-centered design
KT: knowledge translation
LGBTQ+: lesbian, gay, bisexual, transgender, queer or questioning, and plus (others)
PEER: Partnership to Empower Adolescent Engagement in Research
PI: principal investigator
SRH: sexual and reproductive health
STI: sexually transmitted infection

Edited by C Jacob; submitted 10.01.25; peer-reviewed by K Bond, E Tsiamis, C Emezue; comments to author 31.05.25; revised version received 31.07.25; accepted 14.10.25; published 01.12.25.

Please cite as:
Meherali S, Rehmani AI, Ahmad M
Adolescent Voices in Action—Co-Designing Digital Sexual and Reproductive Health Knowledge Translation Interventions: Community-Based Participatory Action Project
JMIR Hum Factors 2025;12:e71148
URL: https://humanfactors.jmir.org/2025/1/e71148
doi:10.2196/71148
PMID:

XSL•FO
RenderX

©Salima Meherali, Amyna Ismail Rehmani, Mariam Ahmad. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 01.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Experiences of Using Online Peer Forums Among People With Postpartum Psychosis: Interpretative Phenomenological Study

Katherine Williams[1], BSc, PGCert, PGDip, MA, DClinPsy; Fiona Lobban[2], BA, PhD, DClinPsy; Elizabeth Chamberlain[3], BSc, DClinPsy

[1]Department of Health Research, Doctorate in Clinical Psychology, Lancaster University, Lancaster, United Kingdom

[2]Division of Health Research, School of Health and Medicine, Spectrum Centre for Mental Health Research, Lancaster University, Lancaster, United Kingdom

[3]Perinatal Mental Health Team Liverpool and Sefton, Perinatal Mental Health Service, Mersey Care NHS Trust, Liverpool, England, United Kingdom

Corresponding Author:
Katherine Williams, BSc, PGCert, PGDip, MA, DClinPsy
Department of Health Research
Doctorate in Clinical Psychology
Lancaster University
Lancaster, LA1 4AT
United Kingdom
Phone: 44 01524 65201
Email: k.dearing25@gmail.com

Abstract

Background: Although research has found online peer support forums to be helpful for those with mental health conditions, no studies have explored the experiences of those who use forums for support with postpartum psychosis (PP) specifically.

Objective: This study aimed to understand the lived experiences of using online forums for PP, and how this form of support differs from professional and other informal support.

Methods: This study used a qualitative approach, including semistructured interviews with 8 participants. Recruitment took place via an online forum run by a charity called Action on Postpartum Psychosis. Transcripts were analyzed using interpretative phenomenological analysis.

Results: Four themes were developed in line with participants' experiences (1) from isolation to connection: validation, growth, and hope from shared experiences; (2) complementing not replacing: filling the gaps in support; (3) impacts of privacy, representation, and readiness to share on engagement; and (4) relational experiences within peer support: altruism, boundaries, and comparison. All participants believed forums were helpful to their well-being and recovery; however, some also reported difficulties with engagement, comparison, and regulating their own use. Findings suggest that forums may benefit from being designed in a way that protects users and their identities, for example, via trigger warnings and setting boundaries.

Conclusions: Peer online forums offer a unique and potentially effective addition to existing support provided by professionals and personal connections. Professionals should signpost people experiencing PP to forums, but should also understand the support that may be needed in terms of monitoring use and ensuring that appropriate boundaries are put into place.

(JMIR Hum Factors 2025;12:e80717) doi:10.2196/80717

KEYWORDS

postpartum psychosis; puerperal psychosis; postnatal psychosis; online peer support forums; online forums; peer-to-peer support; peer support

Introduction

Postpartum psychosis (PP) is a rare perinatal mental health condition, affecting between 1 and 2 in 1000 individuals [1,2]. It is a psychiatric emergency that can occur suddenly [3], usually between 0 to 185 days following childbirth [4]. Symptoms characterizing PP include mood changes, hallucinations, and delusions [5,6]. The risk of onset of psychosis is thought to be approximately 23 times higher in the 4 weeks following childbirth than at any other stage of life [7]. PP is less common and less understood than more common perinatal mental health conditions, such as postnatal depression, and may therefore be

XSL•FO
RenderX

less readily recognized [8]. Current understanding of what is helpful for PP is limited [9], and therefore it is important to consider support for people with lived experience of PP throughout and following this period.

Despite recent introductions of specialist perinatal support within the United Kingdom [10,11], and the time spans for which these services can support individuals being extended [12], gaps remain in both understanding and treating PP [4,13]. People with lived experience of PP have reported a lack of awareness from both the public and professionals [14], leading to delayed recognition and treatment [15]. PP is not formally recognized as a diagnosis by the National Institute for Health and Care Excellence [16] or the *DSM-5 (Diagnostic and Statistical Manual of Mental Disorders* [Fifth Edition]) [17]. The most common treatment options for PP include medication [18], psychological interventions [19], and, occasionally, electroconvulsive therapy [20]. Recovery from PP is thought to take between 6 and 12 months [21], with the most severe symptoms often resolving within 12 weeks when intervention is received [22]. However, there are many challenges that people with lived experience of PP encounter that these traditional interventions may not address. People with lived experience of PP report feeling isolated [14], overwhelmed [5], and stigmatized [19,23], indicating that psychosocial challenges for people with lived experience of PP are difficult to navigate throughout the recovery process.

An increasingly popular means of support for health-related challenges is through online mediums, and the use of this for mental health has become a fast-growing area [24]. It is thought that online support could be particularly helpful for mental health for reasons including the opportunity to remain anonymous when seeking support [25] to help decrease felt stigma [26], and the social isolation that is also commonly experienced [27].

Peer forums are a form of online support, characterized by the ability to anonymously provide and receive support among those in similar situations [28]. Previous research has found that peer support forums can reduce isolation [29,30], promote information sharing [30,31], and provide emotional support and hope [32]. In a recent realist synthesis to identify the impacts of online peer support forums for mental health [29], findings included improving self-efficacy and psychological safety, and connection. However, users of online support have also reported that they can sometimes feel triggered by others [30], receive misleading, defensive, or unempathetic responses [31], and struggle with their own boundaries [30], suggesting this may be a nuanced area.

When considering online support for those with PP, it is important to consider the unique experiences that could impact its use. Declines in mental well-being within the perinatal period can be distressing and overwhelming [33]. In addition, there is the challenge of navigating parenthood and the changes in identity that this and PP bring [19], alongside the impacts of symptoms on daily life, with those who experience relapse of PP being found to have more ongoing functional difficulties [34]. Online forums for support with parenting are commonly used for voicing feelings around parenting that might feel difficult to speak about openly in other contexts; however, research has found that forums tolerate these expressed emotions only to a certain level before it is implied the person may not be in the "norm" [35]. This limits the support available to people with lived experience of PP in these nonspecialist forums and suggests they might need to speak with others who understand to feel safe.

Despite there being no existing research that specifically examines the impact of peer online support for people with lived experience of PP, studies on wider topics have reported related findings. For instance, perceived social stigma is correlated with increased use of online support [36], and those experiencing perinatal mental health conditions tend to have higher levels of internal stigma [33]. Furthermore, findings have reported that increased use of forums could result in users being more likely to disclose their symptoms to professionals [33]. Research has demonstrated that those who conversed with others who have also experienced PP found this support helpful and normalizing [19]; however, with PP being rare, shared experiences may likely be more easily accessible through online mediums. Stawarz et al [37] suggested that technology and online support should not be a replacement for more formal support but instead should complement it and form part of a wider, holistic system. This may be particularly true for PP, which requires fast intervention and support [38], indicating the need for professional involvement alongside personal and social support.

NHS guidance directs people with lived experience of PP to peer support forums [2]; however, there have been no studies that have explored the impacts and safety of this type of support. Therefore, this study aims to explore the experiences of those who have used online peer forums for support with PP, to increase the evidence base and aid professionals in signposting decisions. For this research, support with PP was defined as when an individual had used the forum specifically in relation to any symptoms or difficulties related to their PP. The research questions posed were: (1) How do individuals with lived experience of PP use and make sense of online peer forums? and (2) In what way does online support differ from support from health care professionals and other informal support?

Methods

Study Design

In this study, Interpretative Phenomenological Analysis (IPA [39]) was used to explore the lived experiences of using online forums for PP, and how this form of support differs from that of professional and other informal support.

This included any support received within a forum environment that was related to symptoms or difficulties associated with PP. This broad definition of support meant that participants were able to speak about their own individual support needs related to their experiences of PP, which can vary between participants. In this study, "support" was operationalized as any form of emotional, informational, or experiential connection participants received through engagement with an online peer forum.

The method used for this study was IPA. This method most appropriately met the research aim to generate a rich and

XSL•FO
RenderX

detailed understanding of participants' experiences, rather than a broader and more varied focus, as is seen in other qualitative methodologies such as thematic analysis [40]. In addition, within IPA, there is a focus on lived experience, which helps to capture nuance between participants [41] and to explore the meaning behind experiences [42].

Patient and Public Involvement

Patient and public involvement (PPI) has been found to add insight from lived experience, achieve meaningful outcomes [43], make research more accessible, and ensure sensitivity to the topic [44]. The lead researcher held consultation meetings with a member of the public who had lived experience of using forums for support with PP. Consultations included input on the topic guide, completing a mock interview, and developing research materials. In addition, the researcher liaised with employees of Action on Postpartum Psychosis (APP) with lived experience of PP regarding research materials, the topic guide, and terminology used. In response to PPI feedback, wording in the research materials was changed, alongside edits to the research flyer to ensure language was accessible.

Inclusion and Exclusion Criteria

Inclusion and exclusion criteria were considered to ensure that the sample was relevant and focused on the research aims (Table 1).

Table 1. Inclusion and exclusion criteria.

Inclusion or exclusion	Criteria	Rationale
Inclusion	• Aged 18 or older • Having the capacity to consent • Having used peer support forums within the past 2 years • Symptoms of postpartum psychosis (PP) experienced at any time	• Able to give informed consent • Ensuring consent is ethically given • Forums are evolving quickly [45]; those who used forums longer ago may have had experiences less relevant to current experiences • Due to recognition that people with PP can need support for several years after onset, and recognizing that support has changed in recent years [11,12], and wanting to capture information from those who used the forums when support might not have been as available
Exclusion	• None beyond self-reported experiences of PP symptoms	• No formal mental health diagnosis required as PP is not recognized as a distinct diagnosis [17], and potentially issues of stigma [46] and lack of formal screening measures leading to PP going unrecognized [20]

Sampling

A purposive sampling approach was used between February and July 2024 to recruit people with lived experience of PP who had used forums for support. This method aligned with the aims of IPA, which prioritizes depth of experience over breadth of participants to gain a detailed understanding of the phenomenon of interest. A sample size of between 8 and 10 participants was sought, in line with recommendations for both qualitative research [47] and research conducted at a doctoral level [36]. It was anticipated that this would provide sufficient data to build an understanding of experiences of support, particularly due to the rarity of PP and therefore the likely small cohort of available potential participants who had both experienced PP and used forums for support with this. Furthermore, this sample size was considered sufficient to answer the research questions posed.

Recruitment

Participants were recruited through an international forum run by APP. This forum is hosted by HealthUnlocked, a worldwide community providing support for over 250 health conditions, with the APP forum accommodating 3793 users (figure correct on February 19, 2025). Participants could have accessed other forums for support with PP and still participated, as the research was not specific to the APP forum. Meetings were held with moderators of the APP forum, who posted a recruitment flyer on the forum to advertise the study. Forum users interested in taking part were invited to contact the lead researcher via email and then offered a screening call to check eligibility and to address any queries.

Study Procedure

Following the initial screening call, eligible participants were provided with an information sheet (Multimedia Appendix 1) and a consent form via email (Multimedia Appendix 2). Participants were then asked to return the consent form to the lead researcher if they wished to take part. Once the form was returned, an interview time and date were booked that were mutually convenient for both parties. Sociodemographic information (Table 2) was collected verbally at the start of each interview, and participants were given the choice to decline to answer any questions if they wished.

XSL•FO
RenderX

Table 2. Participant demographic information.

Participant	Pseudonym	Age (years)	Ethnicity	Country of residence	Time since onset
1	Jane	31	White Scottish	Scotland	2 years
2	Florence	39	White British	England	13 months
3	Emmeline	57	White American	United States	22 years
4	Frida	34	White British	England	2 years and 3 months
5	Ada	35	White, Mixed Asian	England	4 years
8	Rosalind	51	Black African	England	19 years
9	Marie	35	African American	United States	1 year and 7 months
10	Eleanor	43	White British	England	8 years

A topic guide (Multimedia Appendix 3) was used during the interviews to remain close to the research aims. The definition of support with PP was used to develop this topic guide, and the questions were reviewed and refined within research supervision and PPI meetings. A mock interview was conducted with a PPI representative, which led to some changes in the terminology and style of the questions within the topic guide, for example, reducing jargon in the questions.

Data Collection

Interviews took place at an agreed time over Microsoft Teams, Zoom (Zoom Communications), or a telephone call, depending on the participant's preference, and were recorded onto an encrypted audio device. Before commencing the interview, consent was checked. A topic guide was used as a prompt; however, in line with IPA [39] methodology, the interviews were guided by participants and their experiences [48]. This was done by allowing participants to speak about the phenomena they felt were pertinent to their experiences of using forums for support with PP. Participants were able to interpret questions in their own way, and further questions followed from responses given, meaning that a more thorough exploration of experiences could be undertaken. This helped to gain further context and understanding around the topics discussed by each participant, as well as a way to relate any responses given back to the research questions. Following the interview, participants were debriefed, and their safety was checked.

Data Analysis

To analyze the data, the first 3 interviews were transcribed by the lead researcher soon after completion to review the questioning style. The remaining 5 interviews were transcribed by a contracted transcriber, who had undergone screening and signed a confidentiality agreement. The reason for this was that the lead researcher was not feeling immersed in the data by transcribing, but instead by reading the transcript back while listening to the recording. As a transcriber was available, this was used to spend more time engaging with each interview. Reading the transcripts while listening to the interviews helped to understand context and emotive language and to become more familiar with the data, while also checking the accuracy of the transcripts. Analysis was performed in line with recommended processes from Murray and Wilde [48], which were developed for postgraduate IPA projects undertaken by those who are novices to IPA. This approach follows a stage process: first, initial coding, which involves notations for any elements directly relevant to the research aims. This process was completed by hand using Post-it notes. The second stage is grouping the codes into subthemes, with each subtheme given an interpretative summary and title. The following stage is producing interpretative summaries and titling themes for each participant. Following this, the stage of merging analyses across transcripts was completed, where the subthemes from each participant were merged into collective themes across participants, with narratives and titles created for each.

Reflexivity

IPA is a methodology underpinned by a critical realist epistemology [41]. This aligned with the lead researcher's critical realist view that, while the material world exists, our own perceptions and language shape the way in which we see and experience it [49]. This means that experiences of the world are all individual, such that although all participants had experience of using online forums, the meanings made from this would all differ.

The nature of IPA involves a process of double hermeneutics, with the participants interpreting their own experiences and the lead researcher further interpreting this themselves [49]. Therefore, the data output from this study is a secondhand account of the data collected, which may have been influenced by the demographics of the researchers (Multimedia Appendix 4). The lead researcher was a White, UK-based female trainee Clinical Psychologist with a professional background in social work. These personal and professional identities shaped the lens of focus when approaching the research, potentially influencing the analysis by orienting attention toward psychosocial, relational, and systemic elements of the data. In addition, the analysis was conducted from within a Western sociocultural context, with UK-specific mental health systems and norms likely influencing interpretation. While the researcher had no personal experience of PP, previous clinical work with individuals experiencing PP may have led to certain preconceptions or expectations.

To try to minimize the impact of these factors and enhance the trustworthiness of the analysis, several measures were taken. First, regular supervision was sought throughout the research process. Two supervisors supported the project; both were qualified clinical psychologists who were White, female, and living within the United Kingdom. One supervisor had expertise

XSL•FO
RenderX

in qualitative research methods, while the other had experience working within perinatal mental health services. Within supervision, they both reviewed developing codes and themes and offered critical feedback to help challenge assumptions and ensure that interpretations were grounded in participants' accounts rather than shaped by the researcher's preexisting frameworks.

The lead researcher also engaged in reflexivity through the use of a reflective diary, in which decision-making, emotional responses, and interpretations of the data were documented. This process helped identify where professional training or clinical experience might be influencing the interpretative process. Reflexivity was particularly prioritized after emotionally resonant interviews or when strong thematic impressions emerged.

In addition, moments of purposeful reflection were built into the analytic process—for example, pausing after the first interview and again after the first full transcript analysis, supported by supervision. This helped to preserve the integrity and transparency of the analysis and to ensure that findings remained grounded in participants' lived experiences.

Ethical Considerations

Ethical approval was granted by the Lancaster University Ethics Committee (FHM-2024-3722-RECR-3). An information sheet was provided to participants, and informed consent was obtained through a signed form. Participants were given the option to opt out at any point before data analysis was completed. Due to the potential for distress associated with the interview process, debrief materials were sent to all participants following the interview, with details of services they could contact for support if necessary. Contact details for a named professional involved in each participant's care were collected before the interview in case any risk issues were disclosed during the interview process. Data were deidentified and securely stored on the university's cloud storage. Participants were compensated with a £30 ($39.95) payment for giving their time to the research.

Results

Participant Characteristics

Initially, 15 people expressed interest in the study. Of these, 2 did not attend the initial screening call, and 3 were ineligible to participate due to not using forums for PP (n=2) or not using forums within the past 2 years (n=1). Two participants consented to be interviewed; however, they did not take part due to disengagement (n=1) and withdrawal (n=1). This resulted in a final number of 8 participants. Recruitment was a relatively easy process, with enough people showing interest through the APP without the research having to be advertised elsewhere. Participants were happy to discuss their experiences openly, with the average length of interview being 41 minutes. Participants were given pseudonyms, and demographic information was collected (refer to Table 2).

Analysis

Following the analysis, 6 interrelated themes were constructed to illustrate experiences of using a peer support forum. Multimedia Appendix 5 demonstrates which participants contributed to each theme.

Theme 1: From Isolation to Connection: Validation, Growth, and Hope From Shared Experiences

All participants are represented within this theme, which captured the evolving relationship participants had due to PP; beginning with feelings of isolation and stigma, progressing through validation and understanding when finding the forum, and resulting in a sense of growth, knowledge, and purpose. Validation of these experiences from peers who had shared similar challenges was identified as an important part of recovery.

> *"Reading the experiences of women who'd recovered gave me hope that I would too"* [Florence]

> *"I think it's given me a real validation that what I went through was real… and not just something they couldn't figure out"* [Emmeline]

This demonstrates the lasting impact of feeling misunderstood during the initial phases of PP, and how finding the forum, even many years after experiencing PP, helped to provide affirmation. For participants who experienced PP longer ago, the impact of having many years of no connection to other people with lived experience of PP was loneliness that was turned around by finding the forum.

The accessibility and anonymity of the forum made it particularly valuable. It allowed participants to engage when it felt emotionally safe and to return to the space whenever needed without pressure or stigma.

> *"I think the other thing is the accessibility day or night, in that if… maybe I couldn't sleep or I was up with the baby, you can access it at any time."* [Florence]

While many participants used the forum from the early days of diagnosis, Marie did not use the forum initially due to feeling in denial. She described how the forum helped her to accept her diagnosis, leading to her eventually being able to post herself.

> *"You're assisting people… whilst still receiving treatment for yourself, so yeah, I felt so strong again in the community"* [Marie]

The rarity of PP also impacted experiences, with participants feeling as if there was no one they could speak to before the forum who understood what they had been through.

> *"That's a long time to go without ever having talked to somebody about something that was so pivotal in your life, and so that was just amazing to me when I first got on and started talking to people. The feeling of validation, the feeling of connection, the feeling of wanting to help was overwhelming."* [Emmeline]

The use of the word "pivotal" here demonstrates the profound impact that forums can have for those who use them.

Two participants commented on how they had felt increased isolation due to others cutting them off following their experiences of PP:

XSL•FO
RenderX

> *"I've been shunned by people who I thought were friends"* [Rosalind]

> *"I felt isolated from everything, from the community, from my family"* [Marie]

Rosalind and Marie were also the only participants in the study with African heritage, which suggests it is worth considering the specific cultural impacts that PP can have and how PP might be viewed differently within different communities. For these participants, the isolation was felt on an additional level due to personal support feeling unavailable, highlighting how important the connection on the forum was.

Theme 2: Complementing, Not Replacing: Filling the Gaps in Support

All participants, except one, are represented in this theme, which recognizes the way in which forums supplement existing services and personal support. Peer support was experienced differently from any other support that had been offered, much of which had led to people with lived experience of PP feeling misunderstood and judged.

> *"[Peer support] is almost like a superpower because… you've got that empathy"* [Frida]

Even for participants who felt there had been a good support system around them throughout, it felt safer to speak to peers, with the benefit of feeling less like a burden:

> *"You don't want to burden your family either, so talking to your people on an online forum… is, I think, personally the better way."* [Eleanor]

There was a sense of safety felt within the PP community when speaking about these distressing thoughts, which may not be present elsewhere.

Personal support was felt to be lacking in the understanding that peer support can provide:

> *"I've always been so lucky to have friends and close family and support… but for this experience, it's not been good enough."* [Jane]

This demonstrates how there is a unique nature to PP that means personal support might not have the same impact when experiencing severe mental health difficulties. Jane describing how "*everyone close to me has had a different kind of experience of the same situation"* shows how PP adds a layer to the changes felt post partum.

When referring to professional support offered, most participants found this helpful and felt that such services should be complemented by the forums rather than replaced. There were concerns raised about people who might use the forums instead of speaking to professionals.

> *"Some of the women sounded like they just really needed to go and see their doctor, and that this forum wasn't necessarily quite the right place for them."* [Jane]

This highlights how PP is multifaceted and that one type of support may not be helpful in isolation. There is recognition of the limitations that the forum can have, particularly around medication and interventions. The acknowledgment of health care professionals being central to the start of recovery demonstrates how forums may become more useful at a later stage.

However, not all participants felt supported by professional services. Rosalind described how she felt let down by services, stating that she felt misunderstood. Rosalind believed that services are so overstretched that they cannot provide the support needed:

> *"It's ridiculous that we're having to use peer support in order to provide services that really the NHS should be providing. It's the walking wounded trying to help each other."* [Rosalind]

Rosalind's experiences may be different from those of other participants due to witnessing changes in services without having access to more specialist services that have recently been created (eg, being ineligible for MBU support or perinatal services support due to experiencing PP 19 years ago). As someone still experiencing the impacts of PP, Rosalind demonstrates a need for longer-term support than statutory services can provide. This was reinforced by Jane describing discharge from professional services as being a "*carpet from underneath you.*" It was felt that forums could help to fill this gap, as they can be accessed for as long as needed:

> *"Then once I get discharged from that, I'm like, 'right, I've still got my platform, I've still got the online stuff that I can still reach out to should I need it'… it's like you feel safe… you're not left to it."* [Ada]

This shows a potential disconnect between what specialist services provide (up to 24 months post partum) and what people with lived experience of PP believe would be helpful.

Theme 3: Impacts of Privacy, Representation, and Readiness to Share on Engagement

Two participants are not represented in this theme, which recognizes that although forums were valued by all participants, state of mind, age, and identity all played a role in how able people felt to share their experiences openly. Reflections were provided on the stages at which participants first started to use the forums, with Ada referring to how, initially, she felt paranoid and embarrassed and therefore did not want to use them. Eleanor echoed this point, stating that she did not feel able to speak about PP until 3 months following the onset. There were physical elements of PP that were thought to be involved in engagement, alongside being able to check in on emotional state.

> *"Then at the time when I was so acutely unwell, my brain was just firing off left and right, I could barely even use my phone at one point, and I couldn't organize my thoughts."* [Eleanor]

This adds weight to the importance of the type of support at different stages, acknowledging that symptoms of PP can make using forums more difficult and therefore other support may be imperative at this time. Privacy within forums was also a challenge, with the anonymity often provided not removing this concern:

XSL•FO
RenderX

> *"[Forum] is an international forum, and I do worry a little bit about privacy issues."* [Emmeline]

Florence referred to younger generations potentially finding forums easier to use, but in older generations, it is more common to keep one's feelings to oneself. As the interview progressed, Florence reflected on this and provided insights into how she can often still feel shame around PP, wondering if it could be helpful for her to open up more. This demonstrates the complexities of deciding what to share within forums and how this might continually shift as participants move through their recovery.

A nuance that came through within this theme was around ethnicity, with Rosalind feeling that there are impacts of being a Black woman on her use of forums. Rosalind described how Black women are more likely to experience birth trauma and be within the mental health system, and how this could translate to being more likely to use forums for support. However, this was not reflected in Rosalind's experience, with feelings of being underrepresented within the forums, which resulted in further isolation.

> *"If [forum] could have someone who is black, that would really help me as well, because I think in the black community, mental health isn't talked about or understood."* [Rosalind]

This demonstrates how identity can further lead to feelings of isolation and how this can replicate difficult experiences from other areas of life. In feeling unable to relate to others on the forum, there could be an impact on engagement and freedom to share, causing those from a global majority background to be at a disadvantage when engaging in peer support.

Theme 4: Relational Experiences Within Peer Support: Altruism, Boundaries, and Comparisons

One participant is not represented within this theme, which acknowledges the complexities that navigating a community can bring. On one hand, users could access information and support, and were given opportunities to use their experiences to help others. However, sometimes these opportunities brought pressure to respond and a sense of responsibility or comparison to others.

The forum was described by Marie as a "*happy place*" that allows participants to get their own support as well as give it:

> *"I just want to reach out… and tell them that they can get through it, there is an end, and I am proof there is an end."* [Emmeline]

This acknowledges the tie people with lived experience of PP tend to feel to the community, wanting to give back and *"do my bit for the community"* [Frida]. However, it also felt at times that this desire to give back became burdensome:

> *"The compulsion to post or to reply to people… that did cause me quite a bit of stress. Those feelings of like guilt."* [Frida]

The use of the words *"compulsion"* [Frida] and *"obsession"* [Florence] were used when speaking about using the forums when in a more vulnerable place. This shows how forums can have emotive impacts on those using them, who may already be experiencing their own distress and challenges. There was mention of moderation at these times being helpful:

> *"He'd [partner] realized what had been going on with me, going down a rabbit hole on the internet, obsessively looking for answers… so there was external moderation."* [Eleanor]

Boundaries were also described to help safeguard against feelings of responsibility. However, they were also acknowledged to be difficult to establish while mood is fluctuating, emphasizing the impact that symptoms of PP can have on engagement.

A further challenge of the forums arose from the tendency to compare. Upwards comparison was the most prominent in interviews, characterized by feelings of being patronized or not being good enough when speaking to those who had an *"easy"* [Jane] recovery.

> *"You think have I made a mistake? It's like am I doing it wrong? Like am I doing recovery wrong?* [Jane]

Ethnicity also played a role in upward comparison, demonstrating the impact that diversity within spaces such as forums can have.

> *"I just think in white women the care is generally a lot better, and I'm seeing women recover so much better, whereas with me now I'm disabled and it's ruined my full life."* [Rosalind]

This shows how forums can highlight discrepancies in care offered for PP and recognizes the impact that this could have on users.

Downwards comparison was also mentioned during interviews and seemed to be associated with feelings of being ungrateful for their experiences, feeling sad for others, and wondering "what if."

> *"I feel fortunate with the care that I got… but sometimes it makes me feel… ungrateful."* [Jane]

Questions arose about how representative forum posts are of the wider population of people with lived experience of PP. For example, Jane wondered if those who are struggling more tend to use forums more, interpreting this as feeling that those who recover well move on.

> *"It's just scary to see so much content and relapses and stuff, but that might be the people who might need support. I wonder if there's like a hoard of people who have not had like relapses or other diagnoses that don't need the support and are fine and not on the forum."* [Jane]

This poses an interesting point about the types of users that forums might attract, and whether this is reflective of different experiences of PP, or whether this tends to reflect certain narratives instead, such as people struggling to recover. However, a differing opinion was put forward by Rosalind, who felt as if there were not many people posting who were reporting difficulties, which left her feeling as if she was struggling more than others:

XSL•FO
RenderX

> *"At least online it feels like they have recovered more than I have… that's why I wish people would talk because I can't be the only one who's had an adverse reaction to their meds."* [Rosalind]

Discussion

Principal Findings

This research aimed to explore the lived experiences of using online forums for PP and how this form of support differs from that of professional and other informal support.

The findings demonstrate that this is a complex and nuanced area, and that people with lived experience of PP tend to use forums when reaching stability rather than in the initial acute phases of PP. It is also acknowledged that the use of forums can be helpful, supportive, and aid recovery. However, forums can also have challenges such as feeling responsible for other people with lived experience of PP, comparing to others, and needing to implement boundaries to maintain well-being. Individual experiences are impacted by many factors, including identity, other available support, and recovery. The research was novel in that it was the first investigation into the use of peer online forums for PP, adding to the limited research base on PP. The use of IPA allowed an in-depth, idiographic insight into the lived experiences, which did not just focus on forum use at a single point in time, but instead over the recovery trajectory of the participants. This created an understanding of how the use of forums changes, increasing knowledge of how these mediums can provide longer-term support for people with lived experience of PP.

An important element of this research was related to the rarity of PP [6], which was felt to increase the isolation that can be felt within perinatal mental health conditions [50]. Online forums give the opportunity for connection between people with lived experience of PP that otherwise may not occur, especially if situated in a more remote area. Connection found within forums was discussed by all participants, which demonstrates the extent to which online relationships can be felt as true connections despite potentially never meeting face-to-face. A recent systematic review has found PP to be a relational experience, both in recovery and management [51], and although this was regarding the couple relationship, considerations should be given to how the relational elements of PP might also be helped by connection on forums and with peers. The ability to engage on forums was found to be helped by accessibility and anonymity; this mirrors existing findings regarding forums that these factors remove barriers to support and reduce stigma [30]. This is particularly important within the PP community due to increased levels of stigma [52] and fears around social services involvement following disclosures [53].

Connection was felt to be an imperative part of forum use, linking to the CHIME recovery framework [54]. Forums played a clear role in promoting connectedness, instilling hope, shaping identity, and offering a platform for meaning-making and empowerment. However, it is important to note that this was related to personal recovery and that there are differences between this and clinical recovery [55], with both being important within psychosis presentations [56]. This contextualizes theme 2 and the importance of both professional supports to promote clinical recovery, while peer support can help achieve personal recovery.

However, connection can come with challenges, such as potentially feeling responsible for the well-being of other users. A key finding within this study was that participants felt they wanted to give back to the PP community, which was highlighted by several participants as being important for both their recovery and well-being. Although previous research has referenced helping others as potentially being related to reciprocity, and therefore expecting mutual benefit from this [57], within this study, participants suggested different reasons for this altruism. Particularly, this stemmed from an understanding of "dark" times and wanting to provide hope when approaching recovery. For participants who still felt in the midst of PP, helping others originated from a place of not wanting anyone else to feel the isolation and loneliness they experienced. Despite feeling that helping others was important to them, it is crucial to recognize the impact that this responsibility could have on forum users.

The results reflected how participants felt it was important to have boundaries when using online forums, and how this had to be learnt. This was found to be crucial during acute stages [30], which is particularly relevant as participants often engage on forums shortly after the onset of PP. Within the data collected, participants acknowledged how boundaries were more difficult to implement when experiencing more severe symptoms of PP. This poses a question of how forum users set boundaries during the time they potentially need them the most. Participants described feeling unable to step back, which resulted in obsessively scrolling for hours. This is similar to current understandings of social media addiction [58], with peer online forums being alike in characteristics; involving connection, community, and gaining information [29]. Understandably, this sensation akin to addiction could be felt more severely by people with lived experience of PP, who are potentially experiencing a lack of connection due to sleep deprivation [59] and symptoms of mania [7].

Conversations regarding professional support took place across countries and cultures, where participants explained how forums were helpful to them because they could access them for as long as was required; in turn, this helped ease the anxiety of being discharged from professional care. Research has found that levels of anxiety are often high at discharge from both hospital and community care [60,61], which could be particularly relevant within the United Kingdom, where specialist perinatal support is only provided for the first 2 years following birth [12], despite symptoms and impacts of PP sometimes continuing well past this point [62]. This links to Jane's feelings that the forum has an overrepresentation of people with lived experience of PP who have found recovery more challenging, as these users may have required more support and have therefore continued to use the forum. However, a related study recently reported that they did not find engagement with online peer support forums to be correlated with symptoms or disability [63], which makes it important to consider why people with lived experience of PP might continue to use forums.

XSL•FO
RenderX

The findings of the study showed that many of the participants felt they engaged in comparisons with other users on the forum. This can be understood within the context of social comparison theory [64], which has been found to have a strong presence in the use of online communities, such as social media [65]. While upwards comparison has been previously linked with promoting hope and downwards comparison has been thought to put things into perspective [66], this study found comparison to also have potentially detrimental effects. Upwards comparison resulted in the perception of insufficient recovery, while downwards comparison resulted in feelings of guilt about being ungrateful, as their situation could have been worse. It is important to understand the impact that both types of comparison may have on users, particularly if using the forum at a more vulnerable point.

Strengths and Limitations

The use of IPA as a methodology was found to be of value within this study, allowing the understanding of not only the commonalities between participants' lived experiences, but also divergences from these and an exploration of why these might occur. In addition, even with a small sample size, there was diversity in the ethnicity of participants, which contributed to understanding how culture may impact experiences. Using PPI within the research was also felt to be a strength, as this added the perspective and input of people with lived experience of PP, which allowed adjustments to the research when needed. The study added to the lack of existing research around PP to help increase awareness, understanding, and intervention options for people with lived experience of PP.

This research had some limitations in terms of its design. Participants in the study were recruited via one peer support forum for PP, which, although international, was based in the United Kingdom. Due to the recruitment methods, likely, the people who saw the flyer were still using forums, which indicates that they are more likely to have found the forum helpful or that they still needed support. Therefore, there may be a demographic of people who are not represented in this research. Although recruitment was only through one forum, participants were able to speak about experiences from any forums they had used for support with PP; however, this was not recorded. This meant that it is difficult to ascertain whether particular forums are more helpful than others or what characteristics of forums are more useful. In addition, all participants were aged between 31 and 53 years old, and an additional point made by Florence—that younger generations tend to find it easier to share online—poses questions around whether there are age differences related to how people use forums for support.

Despite recruitment taking place on an international forum, participants all resided in Western societies. Perinatal support within these countries can differ greatly from what is offered elsewhere [52]; therefore, it is important to acknowledge that these findings may not be generalizable beyond this population. In addition, it has been recognized and evidenced that those from the global majority can experience worse outcomes regarding mental health care [67], maternal support [68,69], and birthing experiences [70]. Within this research, there are views offered by 2 participants from a global majority background and one from a mixed ethnic background; however, this is a small proportion of the sample, and a more culturally balanced sample could have been beneficial to ensure a more comprehensive understanding of experiences.

Finally, the lead researcher of this study had not personally experienced PP. Within interviews, some participants described the value of shared experience, in that anyone who has not experienced PP cannot understand it. This highlights the impact of the researcher's understanding and interpretation of the interviews as someone who has not shared these experiences. Consequently, the findings of this study may have looked different had they been analyzed by someone with lived experience.

Clinical Implications

Following completion of the analysis, several recommendations were made related to the data.

For forums, trigger warnings and clear titles should be displayed on posts to ensure that readers are making an informed decision about what they are about to read. This links to theme 4, which acknowledged how safeguards against feelings of responsibility are important, as they help users to maintain appropriate boundaries and assess their own well-being before accessing content they might find upsetting.

A key consideration for forums is around representation and how this is managed within an online environment, as discussed within theme 4. Experiences related to diversity were central for the participants from the global majority in this research, with comments on how those within their families and communities did not understand, and how they felt shunned. There was also a reference to how, within the black community, mental health is not spoken about. This highlights a lack of spaces that people with lived experience of PP with black heritage can speak about their experiences. Given this, it might be helpful for forums to consider specific areas for the global majority to share their stories and experiences. This is supported by a recent study that found those from the global majority believed a space for peer support would help have open discussions and create a sense of belonging [71].

When considering implications for professionals, it would be helpful if professionals could signpost to forums where it would be beneficial for people with lived experience of PP. As noted in theme 2, forums have the potential to complement professional support; however, for people with lived experience of PP to use forums, they need to be aware of their existence. Signposting has the potential to reduce load on services, which was described by participants within theme 2 as an extra medium of support. However, it is also important that professionals recognize the potential for distress when using forums and help those they are working with to set boundaries and monitor usage. This is particularly important for people with lived experience of PP, who may be in the initial stages or who might be at risk of relapse.

Within future policies regarding perinatal mental health, it could be helpful to consider the expansion of peer support roles, helping people with lived experience of PP to guide while

XSL•FO
RenderX

having services involved that help to safeguard their own well-being. Theme 1 recognized how people with lived experience of PP could go many years without connection with anyone who had also experienced PP and the impact this had on their feelings of isolation. Although PP peer support roles do exist, involvement of people with lived experience of PP is limited; increasing this support could help increase accessibility and extend it to those who may be less likely to use forums. In addition, this research has highlighted existing knowledge around perinatal services often being western-centric in their approach [72]. It would be helpful for future policy to consider how to increase cultural competence, to tailor support, and improve access for communities that may otherwise be underrepresented within services.

Future Research

In the future, it would be helpful for research to more closely examine the longer-term impacts of PP. Participants within this study highlighted how specialist support was withdrawn at a time when help was still needed. Understanding the continued needs of people with lived experience of PP post recovery could help to improve service provision and increase appropriate signposting. In addition, future research related to the longer-term impacts of using peer online support forums could also be beneficial. As forums are a relatively new medium for support, the lasting impacts of their usage are unknown. As previously stated, the 2 participants within this study who had found the forum long after the onset of PP still found it beneficial. It could be helpful to understand why this might be, and how long people with lived experience of PP continues to engage with the forum, to understand what variables are involved in continued use.

Finally, with a limitation of this research being related to generalizability across cultures, future research should consider cultural differences when it comes to peer support for PP. It would be particularly useful to explore potential barriers to accessing peer support, representation issues, and how cultural perspectives of mental health play a role in disclosure and help-seeking.

Conclusion

This research study demonstrates the impact that online peer forums can have for people with lived experience of PP , recognizing that they help to increase connection, provide unique support, and play a role in recovery. There are challenges associated with the use of forums, such as issues with representation, comparison with others, and a sense of responsibility for other users. Helpful considerations include forums introducing spaces for global majority users and including trigger warnings to help users maintain boundaries. In addition, it was suggested that increasing professionals' awareness of forums as a channel of support and assisting people with lived experience of PP to access these would be beneficial to both services and service users.

Acknowledgments

Thank you to the participants of this study, who volunteered their time, shared their stories, and made this research possible, and to the patient and public involvement representative who used her lived experience to help shape the study. Thank you to Action on Postpartum Psychosis for providing consultation in designing the study and for allowing participants to be recruited from their forum.

Funding

This research was completed as part of the Doctorate in Clinical Psychology program at Lancaster University and was funded by the National Health Service. The funder had no role in the study design, data collection, analysis, interpretation, or writing of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Participant information sheet.
[DOCX File , 62 KB - humanfactors_v12i1e80717_app1.docx]

Multimedia Appendix 2
Consent form.
[DOCX File , 60 KB - humanfactors_v12i1e80717_app2.docx]

Multimedia Appendix 3
Interview schedule.
[DOCX File , 14 KB - humanfactors_v12i1e80717_app3.docx]

Multimedia Appendix 4

XSL•FO
RenderX

Example of double hermeneutics process.
[DOCX File , 15 KB - humanfactors_v12i1e80717_app4.docx]

Multimedia Appendix 5
Extract of individual theme for one participant.
[DOCX File , 16 KB - humanfactors_v12i1e80717_app5.docx]

Multimedia Appendix 6
Cross participant themes and quotes.
[DOCX File , 17 KB - humanfactors_v12i1e80717_app6.docx]

Multimedia Appendix 7
Table of summary of experiences of personal support and forum support.
[DOCX File , 15 KB - humanfactors_v12i1e80717_app7.docx]

References

1. Perinatal mental health. Public Health England. 2019. URL: https://www.gov.uk/government/publications/better-mental-health-jsna-toolkit/4-perinatal-mental-health [accessed 2025-02-19]
2. Postpartum psychosis. NHS. 2023. URL: https://www.nhs.uk/mental-health/conditions/post-partum-psychosis/ [accessed 2025-01-03]
3. Postpartum psychosis. Royal College of Psychiatrists. 2018. URL: https://www.rcpsych.ac.uk/mental-health/mental-illnesses-and-mental-health-problems/postpartum-psychosis [accessed 2025-02-19]
4. Cohen LS, Arakelian M, Church TR, Dunk MM, Gaw ML, Yoon HE, et al. The phenomenology of postpartum psychosis: preliminary findings from the Massachusetts General Hospital Postpartum Psychosis Project. Mol Psychiatry 2025;30(6):2537-2544. [doi: 10.1038/s41380-024-02856-3] [Medline: 39643690]
5. Friedman SH, Reed E, Ross NE. Postpartum psychosis. Curr Psychiatry Rep 2023;25(2):65-72 [FREE Full text] [doi: 10.1007/s11920-022-01406-4] [Medline: 36637712]
6. VanderKruik R, Barreix M, Chou D, Allen T, Say L, Cohen LS, Maternal Morbidity Working Group. The global prevalence of postpartum psychosis: a systematic review. BMC Psychiatry 2017;17(1):272 [FREE Full text] [doi: 10.1186/s12888-017-1427-7] [Medline: 28754094]
7. Bergink V, Rasgon N, Wisner KL. Postpartum psychosis: madness, mania, and melancholia in motherhood. Am J Psychiatry 2016;173(12):1179-1188. [doi: 10.1176/appi.ajp.2016.16040454] [Medline: 27609245]
8. Perry A, Gordon-Smith K, Jones L, Jones I. Phenomenology, epidemiology and aetiology of postpartum psychosis: a review. Brain Sci 2021;11(1):47 [FREE Full text] [doi: 10.3390/brainsci11010047] [Medline: 33406713]
9. Doucet S, Jones I, Letourneau N, Dennis C, Blackmore ER. Interventions for the prevention and treatment of postpartum psychosis: a systematic review. Arch Womens Ment Health 2011;14(2):89-98. [doi: 10.1007/s00737-010-0199-6] [Medline: 21128087]
10. Antenatal and postnatal mental health: clinical management and service guidance. National Institute for Health and Care Excellence. 2014. URL: https://www.nice.org.uk/guidance/cg192 [accessed 2025-12-03]
11. The five year forward view for mental health. NHS. 2016. URL: https://www.england.nhs.uk/wp-content/uploads/2016/02/Mental-Health-Taskforce-FYFV-final.pdf [accessed 2025-12-03]
12. Fit for the future: 10 year health plan for England. NHS. 2019. URL: https://www.longtermplan.nhs.uk/ [accessed 2025-12-03]
13. Cohen LS, Vanderkruik R, Arakelian M, Church TR, Dunk MM, Freeman MP. Establishment of the MGH Postpartum Psychosis Project: MGHP3. PLoS One 2023;18(2):e0281133 [FREE Full text] [doi: 10.1371/journal.pone.0281133] [Medline: 36758024]
14. Forde R, Peters S, Wittkowski A. Recovery from postpartum psychosis: a systematic review and metasynthesis of women's and families' experiences. Arch Womens Ment Health 2020;23(5):597-612 [FREE Full text] [doi: 10.1007/s00737-020-01025-z] [Medline: 32020314]
15. Sharma V, Mazmanian D, Palagini L, Bramante A. Postpartum psychosis: revisiting the phenomenology, nosology, and treatment. J Affect Disord Rep 2022;10:100378. [doi: 10.1016/j.jadr.2022.100378]
16. Differential diagnosis | Diagnosis | Depression - antenatal and postnatal. National Institute for Health and Care Excellence (NICE). 2023. URL: https://cks.nice.org.uk/topics/depression-antenatal-postnatal/diagnosis/differential-diagnosis/ [accessed 2025-12-03]
17. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders: Fifth Edition. Washington, DC: American Psychiatric Association; 2013.
18. Osborne LM. Recognizing and managing postpartum psychosis: a clinical guide for obstetric providers. Obstet Gynecol Clin North Am 2018;45(3):455-468 [FREE Full text] [doi: 10.1016/j.ogc.2018.04.005] [Medline: 30092921]

XSL•FO
RenderX

19. Forde R, Peters S, Wittkowski A. Psychological interventions for managing postpartum psychosis: a qualitative analysis of women's and family members' experiences and preferences. BMC Psychiatry 2019;19(1):411 [FREE Full text] [doi: 10.1186/s12888-019-2378-y] [Medline: 31856785]
20. Raza SK, Raza S. Postpartum psychosis. StatPearls. 2024. URL: https://www.ncbi.nlm.nih.gov/books/NBK544304/ [accessed 2025-12-03]
21. Recovery from postpartum psychosis. Action on Postpartum Psychosis (APP). URL: https://www.app-network.org/postpartum-psychosis/recovery-from-postpartum-psychosis/ [accessed 2025-06-10]
22. Bergink V, Burgerhout KM, Koorengevel KM, Kamperman AM, Hoogendijk WJ, Lambregtse-van den Berg MP, et al. Treatment of psychosis and mania in the postpartum period. Am J Psychiatry 2015;172(2):115-123. [doi: 10.1176/appi.ajp.2014.13121652] [Medline: 25640930]
23. Iyengar U, Hunt J. An open conversation about postpartum psychosis: an interview with Jessie Hunt: communications lead, advocate, and expert by experience. Yale J Biol Med 2024;97(1):107-112 [FREE Full text] [doi: 10.59249/PWXZ3821] [Medline: 38559459]
24. DeAndrea DC, Anthony JC. Online peer support for mental health problems in the United States: 2004-2010. Psychol Med 2013;43(11):2277-2288 [FREE Full text] [doi: 10.1017/S0033291713000172] [Medline: 23410539]
25. Kauer SD, Mangan C, Sanci L. Do online mental health services improve help-seeking for young people? A systematic review. J Med Internet Res 2014;16(3):e66 [FREE Full text] [doi: 10.2196/jmir.3103] [Medline: 24594922]
26. Bartlett YK, Coulson NS. An investigation into the empowerment effects of using online support groups and how this affects health professional/patient communication. Patient Educ Couns 2011;83(1):113-119. [doi: 10.1016/j.pec.2010.05.029] [Medline: 20599338]
27. Primack BA, Shensa A, Sidani JE, Whaite EO, Lin LY, Rosen D, et al. Social media use and perceived social isolation among young adults in the U.S. Am J Prev Med 2017;53(1):1-8 [FREE Full text] [doi: 10.1016/j.amepre.2017.01.010] [Medline: 28279545]
28. Kilpatrick S, Emery S, Farmer J, Kamstra P. Wellness impacts of social capital built in online peer support forums. Int J Environ Res Public Health 2022;19(23):15427 [FREE Full text] [doi: 10.3390/ijerph192315427] [Medline: 36497502]
29. Marshall P, Booth M, Coole M, Fothergill L, Glossop Z, Haines J, et al. Understanding the impacts of online mental health peer support forums: realist synthesis. JMIR Ment Health 2024;11:e55750 [FREE Full text] [doi: 10.2196/55750] [Medline: 38722680]
30. Smith-Merry J, Goggin G, Campbell A, McKenzie K, Ridout B, Baylosis C. Social connection and online engagement: insights from interviews with users of a mental health online forum. JMIR Ment Health 2019;6(3):e11084 [FREE Full text] [doi: 10.2196/11084] [Medline: 30912760]
31. Naslund JA, Bondre A, Torous J, Aschbrenner KA. Social media and mental health: benefits, risks, and opportunities for research and practice. J Technol Behav Sci 2020;5(3):245-257 [FREE Full text] [doi: 10.1007/s41347-020-00134-x] [Medline: 33415185]
32. Kavuluru R, Williams AG, Ramos-Morales M, Haye L, Holaday T, Cerel J. Classification of helpful comments on online suicide watch forums. 2016 Presented at: BCB '16: Proceedings of the 7th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics; October 2-5, 2016; Seattle, WA p. 32-40. [doi: 10.1145/2975167.2975170]
33. Moore D, Drey N, Ayers S. Use of online forums for perinatal mental illness, stigma, and disclosure: an exploratory model. JMIR Ment Health 2017;4(1):e6 [FREE Full text] [doi: 10.2196/mental.5926] [Medline: 28219879]
34. Burgerhout KM, Kamperman AM, Roza SJ, Lambregtse-Van den Berg MP, Koorengevel KM, Hoogendijk WJG, et al. Functional recovery after postpartum psychosis: a prospective longitudinal study. J Clin Psychiatry 2017;78(1):122-128. [doi: 10.4088/JCP.15m10204] [Medline: 27631144]
35. Pedersen S, Lupton D. ‘What are you feeling right now?’ communities of maternal feeling on Mumsnet. Emotion, Space and Society 2018;26:57-63. [doi: 10.1016/j.emospa.2016.05.001]
36. DeAndrea DC. Testing the proclaimed affordances of online support groups in a nationally representative sample of adults seeking mental health assistance. J Health Commun 2015;20(2):147-156. [doi: 10.1080/10810730.2014.914606] [Medline: 25116383]
37. Stawarz K, Preist C, Coyle D. Use of smartphone apps, social media, and web-based resources to support mental health and well-being: online survey. JMIR Ment Health 2019;6(7):e12546 [FREE Full text] [doi: 10.2196/12546] [Medline: 31301126]
38. Costales T. Postpartum psychosis: improving the likelihood of early intervention. Psychiatric Times. 2023. URL: https://www.psychiatrictimes.com/view/postpartum-psychosis-improving-the-likelihood-of-early-intervention [accessed 2025-02-19]
39. Smith JA, Flowers P, Larkin MH. Interpretative Phenomenological Analysis: Theory, Method and Research. Thousand Oaks, California: Sage; 2009.
40. Smith JA, Flowers P, Larkin MH. Interpretative Phenomenological Analysis. 2022. URL: https://elevate.talis.com/lancs/player/modules/633ab8aa1313ba43b3b39103/epubs/645ba949ba85f043bbca491d?chapter=6 [accessed 2025-02-19]
41. Eatough V, Smith JA. Interpretative phenomenological analysis. In: Handbook of Qualitative Psychology. Handb Qual Psychol. Thousands Oaks, California: The SAGE; 2017:193-211.

XSL•FO
RenderX

42. Smith JA. Evaluating the contribution of interpretative phenomenological analysis. Health Psychol Rev 2011;5(1):9-27. [doi: 10.1080/17437199.2010.510659]
43. Arumugam A, Phillips LR, Moore A, Kumaran SD, Sampath KK, Migliorini F, et al. Patient and public involvement in research: a review of practical resources for young investigators. BMC Rheumatol 2023;7(1):2 [FREE Full text] [doi: 10.1186/s41927-023-00327-w] [Medline: 36895053]
44. Accelerated access collaborative patient and public involvement strategy 2021-26. NHS England. 2022. URL: https://www.england.nhs.uk/aac/wp-content/uploads/sites/50/2022/06/Accelerated-Access-Collaborative-patient-and-public-involvement-strategy.pdf [accessed 2025-12-03]
45. Morzy M. Evolution of online forum communities. In: The Influence of Technology on Social Network Analysis and Mining. Cham: Springer; 2013:615-630.
46. Cooke DL, McCarty RJ, Budd S, Ordway A, Roussos-Ross D, Mathews CA, et al. Public stigma and recognition of perinatal obsessive-compulsive disorder. J Obsessive Compuls Relat Disord 2024;40:100858. [doi: 10.1016/j.jocrd.2024.100858]
47. Thorne S. Interpretive Description: Qualitative Research for Applied Practice. NY, New York: Routledge; 2025.
48. Moltu C, Stefansen J, Svisdahl M, Veseth M. Negotiating the coresearcher mandate - service users' experiences of doing collaborative research on mental health. Disabil Rehabil 2012;34(19):1608-1616. [doi: 10.3109/09638288.2012.656792] [Medline: 22489612]
49. Doyle Grace JD, Priest H. Ontology to outcome: a critical analysis of psychological research methodologies. J Crit Psychol Couns Psychother 2015;15(4):204-213.
50. Adlington K, Vasquez C, Pearce E, Wilson CA, Nowland R, Taylor BL, et al. 'Just snap out of it' - the experience of loneliness in women with perinatal depression: a meta-synthesis of qualitative studies. BMC Psychiatry 2023;23(1):110 [FREE Full text] [doi: 10.1186/s12888-023-04532-2] [Medline: 36849948]
51. Turgoose M, Sellwood W, Chamberlain E, Murray CD. Midwives' perspectives and perceptions in relation to perinatal psychotic-like experiences: a qualitative study. J Reprod Infant Psychol 2024:1-19. [doi: 10.1080/02646838.2024.2405122] [Medline: 39291693]
52. Yang JMK, Vaiphei K, Siliya M, Mkandawire T, Dolman C, Heron J, et al. Postpartum psychosis: a public involvement perspective across three continents. Arch Womens Ment Health 2023;26(6):831-837 [FREE Full text] [doi: 10.1007/s00737-023-01347-8] [Medline: 37615717]
53. Webb R, Uddin N, Constantinou G, Ford E, Easter A, Shakespeare J, MATRIx Study Team. Meta-review of the barriers and facilitators to women accessing perinatal mental healthcare. BMJ Open 2023;13(7):e066703 [FREE Full text] [doi: 10.1136/bmjopen-2022-066703] [Medline: 37474171]
54. Leamy M, Bird V, Le Boutillier C, Williams J, Slade M. Conceptual framework for personal recovery in mental health: systematic review and narrative synthesis. Br J Psychiatry 2011;199(6):445-452. [doi: 10.1192/bjp.bp.110.083733] [Medline: 22130746]
55. Anthony WA. Recovery from mental illness: the guiding vision of the mental health service system in the 1990s. Psychiatr Rehabil J 1993;16(4):11-23. [doi: 10.1037/h0095655]
56. Van Eck RM, Burger T, Vellinga A, Schirmbeck F, de Haan L. The relationship between clinical and personal recovery in patients with schizophrenia spectrum disorders: a systematic review and meta-analysis. Schizophr Bull 2018;44(3):631-642 [FREE Full text] [doi: 10.1093/schbul/sbx088] [Medline: 29036720]
57. Kim C, Wang HC. From receivers to givers: understanding practice of reciprocity in an online support community. Proc ACM Hum-Comput Interact 2022;6(CSCW1):1-17. [doi: 10.1145/3512938]
58. Rajeshwari S, Meenakshi S. The age of doom scrolling - social media's attractive addiction. J Educ Health Promot 2023;12:21 [FREE Full text] [doi: 10.4103/jehp.jehp_838_22] [Medline: 37034854]
59. Carr C, Borges D, Lewis K, Heron J, Wilson S, Broome MR, et al. Sleep and postpartum psychosis: a narrative review of the existing literature. J Clin Med 2023;12(24):7550 [FREE Full text] [doi: 10.3390/jcm12247550] [Medline: 38137618]
60. Mao W, Shalaby R, Owusu E, Elgendy HE, Agyapong B, Eboreime E, et al. Depression, anxiety, and poor well-being at discharge from psychiatric hospitals: prevalence and risk factors. Front Psychiatry 2024;15:1408095 [FREE Full text] [doi: 10.3389/fpsyt.2024.1408095] [Medline: 39056021]
61. Meher BK, Das L, Mohanty AK. Discharge and transfer of care. Asian J Clin Pediatr Neonatol 2012;2(3):1-2.
62. Gilden J, Kamperman AM, Munk-Olsen T, Hoogendijk WJG, Kushner SA, Bergink V. Long-term outcomes of postpartum psychosis: a systematic review and meta-analysis. J Clin Psychiatry 2020;81(2):19r12906. [doi: 10.4088/JCP.19r12906] [Medline: 32160423]
63. Smit D, Vrijsen JN, Broekman T, Groeneweg B, Spijker J. User engagement within an online peer support community (Depression Connect) and recovery-related changes in empowerment: longitudinal user survey. JMIR Form Res 2022;6(11):e39912 [FREE Full text] [doi: 10.2196/39912] [Medline: 36322110]
64. Festinger L. A theory of social comparison processes. Hum Relat 1954:117-140 [FREE Full text] [doi: 10.1177/001872675400700202]
65. Sisler S, Kent-Marvick J, Wawrzynski SE, Pentecost R, Coombs LA. Social isolation and social media consumption among graduate students during COVID-19: An examination using social comparison theory. Int J Nurs Health Care Res (Lisle) 2021;4(9):1257 [FREE Full text] [doi: 10.29011/2688-9501.101257] [Medline: 35079736]

XSL•FO
RenderX

66. Rayland A, Andrews J. From social network to peer support network: opportunities to explore mechanisms of online peer support for mental health. JMIR Ment Health 2023;10:e41855 [FREE Full text] [doi: 10.2196/41855] [Medline: 36853738]
67. Thomeer MB, Moody MD, Yahirun J. Racial and ethnic disparities in mental health and mental health care during The COVID-19 pandemic. J Racial Ethn Health Disparities 2023;10(2):961-976 [FREE Full text] [doi: 10.1007/s40615-022-01284-9] [Medline: 35318615]
68. Howell FM, McCarthy KJ, Boychuk N, Burdick M, Nowlin S, Maru S, et al. Racism in obstetric care: a psychometric study of the gendered racial microaggressions scale among global majority birthing people in obstetric contexts. BMC Pregnancy Childbirth 2024;24(1):448 [FREE Full text] [doi: 10.1186/s12884-024-06642-5] [Medline: 38943057]
69. Nguyen TT, Criss S, Kim M, De La Cruz MM, Thai N, Merchant JS, et al. Racism during pregnancy and birthing: experiences from Asian and Pacific Islander, Black, Latina, and Middle Eastern Women. J Racial Ethn Health Disparities 2023;10(6):3007-3017 [FREE Full text] [doi: 10.1007/s40615-022-01475-4] [Medline: 36449130]
70. Liu C, Underhill K, Aubey JJ, Samari G, Allen HL, Daw JR. Disparities in mistreatment during childbirth. JAMA Netw Open 2024;7(4):e244873 [FREE Full text] [doi: 10.1001/jamanetworkopen.2024.4873] [Medline: 38573636]
71. Sandhu A, Aquino MRJ, Tang MY. 'Whose pain is real pain and whose pain is not': an interpretative phenomenological analysis of the role of social support in help-seeking for perinatal mental illnesses in women from minoritised ethnicities. Ethn Health 2025;30(3):413-431 [FREE Full text] [doi: 10.1080/13557858.2024.2442332] [Medline: 40167269]
72. Gardner A, Oduola S, Teague B. Culturally sensitive perinatal mental health care: experiences of women from minority ethnic groups. Health Expect 2024;27(4):e14160 [FREE Full text] [doi: 10.1111/hex.14160] [Medline: 39087742]

Abbreviations

APP: Action on Postpartum Psychosis
DSM-5: Diagnostic and Statistical Manual of Mental Disorders (Fifth Edition)
IPA: Interpretative Phenomenological Analysis
PP: postpartum psychosis
PPI: patient and public involvement

Edited by A Kushniruk; submitted 15.07.25; peer-reviewed by M Evans, L Wreyford; comments to author 18.11.25; revised version received 25.11.25; accepted 01.12.25; published 24.12.25.

Please cite as:
Williams K, Lobban F, Chamberlain E
Experiences of Using Online Peer Forums Among People With Postpartum Psychosis: Interpretative Phenomenological Study
JMIR Hum Factors 2025;12:e80717
URL: https://humanfactors.jmir.org/2025/1/e80717
doi:10.2196/80717
PMID:

©Katherine Williams, Fiona Lobban, Elizabeth Chamberlain. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 24.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Correction: Patient Experiences With a Mobile Self-Care Solution for Low-Complex Orthopedic Injuries: Mixed Methods Study

Jelle Spierings[1*], MD; Gijs Willinge[2*], MD; Marike Kokke[1*], MD; Sjoerd Repping[3*], PhD; Wendela de Lange[4], BSc; Thijs Geerdink[2*], MD, PhD; Ruben van Veen[2*], MD, PhD; Detlef van der Velde[1*], MD, PhD; J Carel Goslings[2*], MD, PhD; Bas Twigt[2*], MD, PhD; Collaboration Group[5]

[1]Department of Trauma Surgery, St. Antonius Ziekenhuis Utrecht, Koekoekslaan 1, Nieuwegein, The Netherlands

[2]Department of Trauma Surgery, OLVG Hospital, Amsterdam, The Netherlands

[3]University of Amsterdam, Amsterdam University Medical Centers, Amsterdam, The Netherlands

[4]The Healthcare Innovation Center, Julius Centrum, University Medical Center Utrecht, Utrecht, The Netherlands

[5]

[*]these authors contributed equally

Corresponding Author:
Jelle Spierings, MD
Department of Trauma Surgery, St. Antonius Ziekenhuis Utrecht, Koekoekslaan 1, Nieuwegein, The Netherlands

Related Article:

Correction of: https://humanfactors.jmir.org/2025/1/e53074

(JMIR Hum Factors 2025;12:e75812) doi:10.2196/75812

In "Patient Experiences With a Mobile Self-Care Solution for Low-Complex Orthopedic Injuries: Mixed Methods Study" (JMIR Hum Factors 2025;12:e53074) the authors noted one error.

The ninth author's name previously appeared as:

Carel Goslings

This has been changed to

J Carel Goslings

The correction will appear in the online version of the paper on the JMIR Publications website, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 10.04.25; this is a non–peer-reviewed article;accepted 10.04.25; published 29.04.25.

Please cite as:
Spierings J, Willinge G, Kokke M, Repping S, de Lange W, Geerdink T, van Veen R, van der Velde D, Goslings JC, Twigt B, Collaboration Group
Correction: Patient Experiences With a Mobile Self-Care Solution for Low-Complex Orthopedic Injuries: Mixed Methods Study
JMIR Hum Factors 2025;12:e75812
URL: https://humanfactors.jmir.org/2025/1/e75812
doi:10.2196/75812

© Jelle Spierings, Gijs Willinge, Marike Kokke, Sjoerd Repping, Wendela de Lange, Thijs Geerdink, Ruben van Veen, Detlef van der Velde, J Carel Goslings, Bas Twigt, Collaboration Group. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 29.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Corrigenda and Addenda

Correction: Requirement Analysis for Data-Driven Electroencephalography Seizure Monitoring Software to Enhance Quality and Decision Making in Digital Care Pathways for Epilepsy: A Feasibility Study from the Perspectives of Health Care Professionals

Pantea Keikhosrokiani[1,2], PhD; Johanna Annunen[3,4*], MD, PhD; Jonna Komulainen-Ebrahim[3,5*], MD, PhD; Jukka Kortelainen[6,7], PhD; Mika Kallio[3,5], MD, PhD; Päivi Vieira[3,5], MD, PhD; Minna Isomursu[1,2], PhD; Johanna Uusimaa[3,8], MD, PhD

[1]Empirical Software Engineering in Software Systems and Services, Faculty of Information Technology and Electrical Engineering, University of Oulu, Oulu, Finland

[2]Research Unit of Health Sciences and Technology, Faculty of Medicine, University of Oulu, Oulu, Finland

[3]University of Oulu, Oulu, null, Finland

[4]Oulu University Hospital, Medical Research Center, Neurocenter (Member of ERN EpiCARE), Oulu, Finland

[5]Oulu University Hospital, Oulu, Finland

[6]Center for Machine Vision and Signal Analysis, Faculty of Information Technology and Electrical Engineering, University of Oulu, Oulu, Finland

[7]Cerenion Ltd, Oulu, Finland

[8]Research Unit of Clinical Medicine, Faculty of Medicine, University of Oulu, Oulu, Finland

[*]these authors contributed equally

Corresponding Author:
Pantea Keikhosrokiani, PhD
Empirical Software Engineering in Software Systems and Services
Faculty of Information Technology and Electrical Engineering
University of Oulu
Pentti Kaiteran katu 1
Linnanmaa
Oulu, 90014
Finland
Phone: 1 358514410809
Email: pantea.keikhosrokiani@oulu.fi

Related Article:

Correction of: https://humanfactors.jmir.org/2025/1/e59558

(JMIR Hum Factors 2025;12:e79484) doi:10.2196/79484

In "Requirement Analysis for Data-Driven Electroencephalography Seizure Monitoring Software to Enhance Quality and Decision Making in Digital Care Pathways for Epilepsy: A Feasibility Study from the Perspectives of Health Care Professionals" [1], the authors made four changes to the authorship list.

First, Johanna Annunen and Jonna Komulainen-Ebrahim contributed equally, and have been credited as such with asterisks in the authorship list.

In addition, the second last and last authors, Minna Isomursu and Johanna Uusimaa, contributed equally as well. This has been added to the Acknowledgements section of the paper as follows:

> *MI and JU contributed equally to this work.*

The affiliation of author JA has been changed from:

> *3. Oulu University Hospital, University of Oulu, Oulu, Finland*

Author JA is now linked to the following affiliations:

> *3. University of Oulu, Oulu, Finland*
>
> *4. Oulu University Hospital, Medical Research Center, Neurocenter (Member of ERN EpiCARE), Oulu, Finland*

XSL•FO
RenderX

This has resulted in the renumeration of subsequent affiliations.

The affiliations of authors JA, JKE, MK, and PV have been changed from:

3. Oulu University Hospital, University of Oulu, Oulu, Finland

These authors are now linked to the following affiliations:

3. University of Oulu, Oulu, Finland

5. Oulu University Hospital, Oulu, Finland

This has also resulted in the renumeration of subsequent affiliations.

The correction will appear in the online version of the paper on the JMIR Publications website, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Reference

1. Keikhosrokiani P, Annunen J, Komulainen-Ebrahim J, Kortelainen J, Kallio M, Vieira P, et al. Requirement analysis for data-driven electroencephalography seizure monitoring software to enhance quality and decision making in digital care pathways for epilepsy: a feasibility study from the perspectives of health care professionals. JMIR Hum Factors 2025 May 30;12:e59558 [FREE Full text] [doi: 10.2196/59558] [Medline: 40446306]

Submitted 22.06.25; this is a non–peer-reviewed article;accepted 27.06.25; published 11.11.25.

Please cite as:
Keikhosrokiani P, Annunen J, Komulainen-Ebrahim J, Kortelainen J, Kallio M, Vieira P, Isomursu M, Uusimaa J
Correction: Requirement Analysis for Data-Driven Electroencephalography Seizure Monitoring Software to Enhance Quality and Decision Making in Digital Care Pathways for Epilepsy: A Feasibility Study from the Perspectives of Health Care Professionals
JMIR Hum Factors 2025;12:e79484
URL: https://humanfactors.jmir.org/2025/1/e79484
doi:10.2196/79484
PMID:

©Pantea Keikhosrokiani, Johanna Annunen, Jonna Komulainen-Ebrahim, Jukka Kortelainen, Mika Kallio, Päivi Vieira, Minna Isomursu, Johanna Uusimaa. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 11.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Requirements and Value Elicitation for a High-Fidelity Pelvic Floor Simulator for Physiotherapists: Mixed Methods Study

Yael Zekaria[1,2,3], MEng; Antonia Tzemanaki[1,3], MEng, MSc, PhD; Jonathan Rossiter[1,3], Prof Dr

[1]School of Engineering Mathematics and Technology, University of Bristol, Ada Lovelace Building, Bristol, United Kingdom

[2]School of Engineering, University of the West of England, Bristol, United Kingdom

[3]Bristol Robotics Laboratory, Bristol, United Kingdom

Corresponding Author:
Yael Zekaria, MEng
School of Engineering Mathematics and Technology, University of Bristol, Ada Lovelace Building, Bristol, United Kingdom

Abstract

Background: Physiotherapists lack training opportunities for repeated practice of pelvic examinations for the identification of pelvic floor disorders (PFDs), leading to low confidence in the clinical setting. Pelvic simulators exist and are a valuable supplement to the medical curriculum, yet none demonstrate pelvic floor muscle (PFM) function or dysfunction. To design effective simulators, an assessment of end-user requirements is essential.

Objective: This study aimed to elicit physiotherapists' needs and requirements for a high-fidelity PFM simulator and the associated use cases.

Methods: This study followed a mixed methods design by collecting qualitative and quantitative data from a web-based survey. Quantitative data were analyzed using descriptive statistics and differences between demographic groups were calculated using 2-sample Kolmogorov-Smirnov 2-sided tests. Qualitative data were analyzed using thematic analysis.

Results: In total, 66 physiotherapists completed the survey. The most common suggested use cases of the simulator were for training and professional development (56/66, 84.9%), and patient education (48/66, 72.7%). Pelvic organ prolapse and muscle tone function and dysfunction were identified as the most useful PFDs for the simulator to demonstrate. Positional tracking and force sensing were considered important features and there was a preference for a generic over a pathology-specific or patient-specific simulator. A total of 3 themes emerged through the qualitative analysis: prioritizing patient care; representing the variability in anatomy and PFDs for simulator realism; and consideration of the implementation, cost, and accessibility of simulators.

Conclusions: There is value in PFM simulators for physiotherapists for multiple use cases. Design recommendations include using realistic materials, demonstrating PFM dynamics, modularity to vary the complexity for different end-users, offering a range of feedback modalities for position and pressure sensing, and ensuring affordability and curriculum integration.

(JMIR Hum Factors 2025;12:e72119) doi:10.2196/72119

KEYWORDS

pelvic examination; high-fidelity simulation; simulation-based learning; medical education; medical devices; physiotherapy

Introduction

Background

Over 60% of people with female-at-birth anatomy experience symptoms of pelvic floor disorders (PFDs) [1] such as incontinence, pelvic organ prolapse (POP) and pelvic pain. In the first course of treatment for PFD symptoms, it is recommended to contact a physiotherapist who will perform a pelvic examination to assess the pelvic floor musculature [2]. Most physiotherapists learn examination skills through peer practice which provides a way to teach ethical consent and allows students to practice their skills on each other's bodies before a patient [3]. However, peers may not present with any PFDs and the idea of being examined by a peer can be uncomfortable for some who if they choose to opt out, miss this opportunity to develop their clinical skills. As a result, it is difficult to gain experience diagnosing and assessing the numerous PFDs.

Physical simulators play a valuable role in supporting obstetrics and gynecology education and training, from offering visualization of anatomy to refining examination skills through repeated practice [4]. While a range of pelvic simulators are on the market, none represent the dynamics of the pelvic floor muscles (PFMs) to the degree of fidelity required to transfer knowledge and skills from the classroom to the clinic.

In a systematic review of simulation training for pelvic examination, simulation-enhanced education was found to improve students' technical skills, comfort and communication

XSL•FO
RenderX

with the patient. However, where offered, the most effective training method is by Gynecological Teaching Associates (GTAs) or standardized patients, volunteers with whom students can practice hands-on examination techniques [5]. As supplements to the medical curriculum, simulation-based teaching has proven effective in improving technical and communication skills for physiotherapists specifically [6]. Simulators can uniquely offer the opportunity for repeated practice and incorporate necessary proprioceptive, visual and tactile feedback to enhance the practitioner's motor control without risk of harm to patients, GTAs or standardized patients [7].

In simulation, (engineering) fidelity is defined as the extent to which the appearance and behavior of a simulator resembles reality, whereas realism, or psychological fidelity, refers to the authenticity of the scenario in which the simulator is used [7,8]. More recently, fidelity in simulation-based education has been subdivided into the following four elements [8]: (1) environmental element - the training facility and equipment used. (2) Patient element - the virtual reality, full body or part-task trainer used to represent the patient. (3) Semantical element - the scenario in which the simulator is used and the associated intention for the user. And (4) phenomenal element - the engagement of the user through their emotional and cognitive experiences while involved in the simulation activity.

It is commonly agreed that the levels of fidelity and realism of a simulator are dependent on the skills it is being used to teach, and for whom it is intended [4,7-9]. Accordingly, the first stage to developing simulators is identifying the skill or procedure that must be learned or demonstrated [7,9,10]. Engaging with end-users is the only reliable way to ascertain these skills and subsequently, the relevant design requirements needed to teach them [9,11].

Most commercial pelvic simulators focus on childbirth, pelvic examination and surgical skills [12]. To the authors' knowledge and at the time of writing, no commercial simulator demonstrates any PFD or even the difference between increased, normal and reduced muscle tone. Decreased or reduced tone and incorrect functioning of the PFMs are some of the main causes of PFDs [13].

Previous Work

Although simulators for certain PFDs have been explored in the research literature, their fidelity and validity vary. In addition, it is rare to describe the design requirements or user needs which would support the future of simulation-based education in pelvic health care. In one of the few simulators designed to mimic the PFM activity for practicing digital examination, Saleme et al [14] modelled the levator ani by 5 springs attached at the relevant points on a resin bony pelvis model. The modified Oxford scale was simulated by varying the spring tension and a physiotherapist identified most differences between the scale points by digitally palpating a specified site between the springs [14]. While this simulator was shown to be a complementary tool to the curriculum, its fidelity was limited by the use of nonrealistic materials. These examples show that both high- and low-fidelity simulation can be effective in medical education, especially for visualizing the 3-dimensionality and dynamics of PFM function and for building confidence before examining patients.

A dynamic and high-fidelity POP simulator was shown to significantly improve the understanding and confidence in performing the pelvic organ prolapse quantification (POP-Q) examination [15]. This simulator was designed such that a user can manually invert a silicone vaginal canal to different extents and angles to simulate anterior, posterior, and apical prolapse. Although medical students and residents at all levels of training were included in this study, responses from all participants' pre- and post-tests were averaged. In another study it was shown that medical students preferred a low-fidelity sock and tube model in their first and second years of residency compared to third and fourth-year residents who had neutral attitudes toward the utility of the model and preferred conventional learning modalities [16]. In an attempt to make what can often be an awkward part of medical training more lighthearted, a Santa Claus hat was used in another study to visualize and simulate POP-Q but the model itself was not evaluated with any meaningful metrics [17]. Through a robust Delphi study, Meyer et al [18] developed a list of criteria for a 3D pelvic model for anatomy and physiology education. While a range of experts from different disciplines were consulted and the methodology was validated rigorously, no physiotherapists were included in the study, the model in question was computer-based and was neither intended for clinical examination nor to have dynamic features. Physical 3D pelvic models and simulators described in the literature tend to be co-designed by medical professionals working in pelvic health [19-22] which may indicate why no user requirements have been published previously. In the evaluation of the 3D-printed model by Kiesel et al [23], medical students reported that the option to remove parts of the anatomy and parts being printed in different colors was helpful for their understanding of the anatomy. Students also found it useful to be able to angle the model in upright and tilted positions and would approve of models that simulate certain medical conditions and operations.

Although the aforementioned examples show positive efforts toward the use of simulation for teaching pelvic examination skills, they highlight the need for more advanced methods, incorporating higher levels of fidelity for more effective training as seen in numerous soft robotic organ models in the literature [24]. This need is even more pronounced considering that veterinary students practice palpation and even the assessment of pain symptoms in the pelvic region of cows and horses using simulators enhanced with haptic feedback [25,26] whereas pelvic physiotherapy training for humans with female-at-birth anatomy continues to rely on peer examination.

The research in this area shows an appetite for high-fidelity simulation training in the gynecological curriculum with a lack of investigation into the specific needs of physiotherapists who are one of the main treatment providers for people with PFDs.

Aims of This Study

In this study, we aimed to explore physiotherapists' perceptions of the value, requirements and potential use cases of a high-fidelity PFM simulator for female-at-birth anatomy. Through a mixed methods survey, we sought to elicit the design

XSL•FO
RenderX

recommendations for such a simulator and gain a deeper understanding of the motives behind these preferences.

Methods

Survey Development

We took a mixed methods approach by collecting qualitative and quantitative data through an internet-based survey. The quantitative component, which comprised the majority of the survey, informed the preferences of certain features and functionalities of the simulator. The qualitative component was intended to gather a deeper understanding of how a simulator can be integrated into pelvic health education and research, along with the associated challenges.

The survey consisted of 10 closed-ended (quantitative) and 3 open-ended (qualitative) questions. The complete questionnaire is shown in Multimedia Appendix 1. Up to 3 additional open-ended questions were asked to respondents who selected certain choices from three of the closed ended questions to better understand their choices. Respondents had the opportunity to add any further comments at the end of the questionnaire and provide a contact email if they wished to be contacted for an informal follow-up interview. Questions on professional experience and demographics were asked at the beginning and end of the survey, respectively.

Closed-ended questions formed the quantitative component of the survey and concerned the features of the simulator. These questions were formulated after exploring state-of-the-art commercial (Clinical Female Pelvic Trainer; Limbs and Things, Bristol, and United Kingdom) and research-based [15,20,27] simulators. We compiled potential use cases, target anatomy, and any additional features into multichoice, ranking and rating scale questions.

The rating scale questions varied from 3- to 5-point scales to reduce the chance of respondents becoming fatigued which can discourage survey completion and result in satisfyingbehavior [28]. The number of scale points was chosen depending on the most appropriate level of granularity for each item that would provide a meaningful understanding of the respondents' opinions [28]. Where relevant, closed-ended questions included optional free text if respondents felt that their attitude or opinion was not offered in the choice options.

Open-ended questions were designed to understand the reasons behind some selected choices of closed-ended questions and to prompt respondents to reflect on their training and development as well as the challenges they face in diagnosing and treating patients' conditions.

As this study aimed to understand the values and preferences of potential users of a pelvic simulator rather than their views of a specific simulator, no proposed solution was provided to avoid influencing respondents' answers. Respondents were only informed that the simulator could be used for medical training, patient education and medical device testing.

The survey was pretested before dissemination by 1 urogynecologist, 1 pelvic medical device engineer, 1 junior doctor, and 2 researcher engineers. The pretesters were asked to provide feedback on the usability of the survey and its presentation, the clarity of the questions, and how long it took to complete. No suggestions were made on the content of the questions. The resulting feedback regarding the usability and readability of the survey was implemented accordingly. The responses from the pretest were not included in the study results.

Ethical Considerations

Ethics approval was granted by the University of Bristol's Faculty of Engineering Research Ethics Committee (reference number 13989). Respondents were required to complete the inbuilt consent form to take part in the survey. Respondents could withdraw their data from the study within 4 weeks of taking part, but none chose to do so. There was no financial compensation offered for taking part in the survey. Any identifyable data were stored separately from the analyzed data on an encrypted hard drive which only the main researcher had access to.

Recruitment and Data Collection

The survey was originally designed for and advertised to professionals working in different areas of pelvic health. As a result, recruitment emails were sent to 2 professional networks of pelvic health practitioners for distribution among their members: the British Society of Urogynaecology [29] and Pelvic Obstetric and Gynaecological Physiotherapy [30]. Individual recruitment emails were sent to pelvic health practitioners with experience in medical education who were identified through their affiliated university profiles and organizations such as Lead Midwives for Education [31]. Pelvic floor medical device engineers and designers were recruited through their companies. Although the recruitment emails were targeted at United Kingdom—based professionals, those working abroad whether they had relocated or been forwarded the email by a colleague in the United Kingdom, were also encouraged to participate in the survey.

The recruitment email briefly explained the purpose of the study, signposted recipients to an attached 2-page document for further information and encouraged them to ask any questions relating to the study. Recipients were invited to complete the internet-based survey by following a URL link and were encouraged to forward the email to colleagues working in pelvic health.

The survey was hosted on Qualtrics (Qualtrics International Inc) and was available for completion for 3 months between May and August 2023 and for an additional month between March and April 2024. Respondents were advised that the survey should take roughly 10 minutes to complete and were notified that once the survey was begun, they would be able to return to it to complete at a later time if convenient.

Data Analysis

Identifiable information such as email addresses, if provided, was removed before analyzing the data. Data from respondents who did not fully complete the survey were not included in the analysis and responses were excluded where there were fewer than 10 respondents from a given profession to ensure statistically meaningful results. Quantitative data were analyzed

XSL•FO
RenderX

in Python (Python Software Foundation) using descriptive statistics such as the mean and SD. For rating scale data, the scales were mapped to a numerical scale: for example, not important=1, moderately important=2, and very important=3 to calculate the variables. Differences in results based on demographic groups were identified using 2 sample Kolmogorov-Smirnov 2-sided tests. Qualitative analysis was conducted in NVivo software (Lumivero) using thematic analysis [32]. Using this method, the principal researcher (YZ) identified themes after familiarization and initial coding of the responses to the open-ended questions. The themes were then refined by 2 researchers (YZ and AT) and codes were grouped into concepts relating to the themes.

Results

Demographic Data

In total, 146 people initiated the survey, and 85 surveys were fully completed, 66 of which were by physiotherapists. There were less than 10 respondents from other pelvic health professions which were excluded to focus on one core group's needs and requirements. At the time of survey dissemination, there were 1219 active members of the Pelvic Obstetric and Gynaecological Physiotherapy who would have received the recruitment email which indicates a 5% response rate. With a 90% CI and 10% margin of error, an ideal sample size for a population of 1219 is 65 people [33].

Table 1 presents the demographic composition of the respondents included in the analysis. Most respondents were female (61/66, 92%), between 45 and 54 years of age (21/66, 32%), with more than 10 years of professional experience (44/66, 67%) working in the United Kingdom (62/66, 94%). Although an additional open ended demographic question on ethnicity was asked, we have not included the responses as the types of answers were too varied (eg, some answered based on ethnicity and others on ethnic category).

Table . Respondent demographics (n=66).

Demographics	Values, n (%)
Professional experience (years)	
Less than 5	8 (12.1)
5 - 10	14 (21.2)
More than 10	44 (66.7)
Country of practice	
United Kingdom	62 (93.9)
Hong Kong	1 (1.5)
Australia	1 (1.5)
Prefer not to say	1 (1.5)
Sex	
Female	61 (92.4)
Male	1 (1.5)
Prefer not to say	4 (6.1)
Age range (years)	
25 - 34	5 (7.6)
35 - 44	20 (30.3)
45 - 54	21 (31.8)
55 - 64	17 (25.8)
65 - 74	2 (3)

Quantitative Results

Use Cases

Respondents were asked how they see themselves using a pelvic floor simulator in their practice. The most common use case is for training and professional development (n=56), followed by educating patients (n=48) and practicing examinations (n=44). Less common use cases of the simulator include as a research tool (n=16) and for medical device testing (n=11). Although 1 respondent said that they would not use a simulator, due to not seeing it as useful at the current stage of their career (specialist pelvic health physiotherapist), they added that it could be helpful for patient education. There was no significant difference between the respondents' choices of use case and their professional experience (P>.49).

Simulator Features

Respondents considered the PFMs, bony pelvis, pelvic organs (excluding the ovaries and fallopian tubes), nerves, and ligaments to be the most important anatomical structures to

XSL•FO
RenderX

include in a simulator. Figure 1 illustrates the mean Likert rating of each structure alongside the raw data.

Out of the 6 presented options for functions and dysfunctions for the simulator to demonstrate, all were considered to be very or moderately useful. Table 2 presents the overall mean and SD between all respondents. Prolapse and muscle tone had the highest mean scores of 3 and 2.95, respectively. There was no significant difference between the respondents' considered utility of the functions and their professional experience (P>.34).

Respondents were asked whether it would be useful for a simulator to be compatible with imaging modalities such as ultrasound. This feature could enable users to practice probe positioning and diagnostic skills. Table 3 shows that compatibility of the simulator with medical imaging is not of strong importance. On average, all modalities scored between 1: not at all useful and 2: moderately useful.

Figure 2 illustrates the preference for a generic over a case-specific (pathology) or patient-specific simulator: a generic simulator was the most preferred (39/66, 59%) and a patient-specific simulator was the least preferred (35/66, 53%). Where respondents ranked the options equally, their preferences were counted for each option.

The preference between auditory, tactile or visual feedback modalities varied, as seen in Figure 3. Overall, visual feedback was most preferred (mean score 4.2, SD 3.8) and acoustic feedback was the least preferred (mean score 2.95, SD 2.8). There was no significant difference between respondents' feedback preferences and their years of professional experience (P>.27) nor their age (P>.35).

Table 4 summarizes respondents' preferences toward the modularity and sensing capabilities in a pelvic floor simulator as well as the term used to describe it. 54 respondents (54/66, 82%) agreed that there was some value in the simulator being modular, that is, for parts to be removable. In total, 82% (55/66) of respondents considered both the position and the pressure of a user's interaction to be useful feedback through sensors embedded in a simulator. Overall, respondents preferred the term 'simulator' (43/66, 64.2%) over model, phantom or manikin.

Figure 1. Rated level of importance of anatomical structures to be included in the simulator. PFM: pelvic floor muscle.

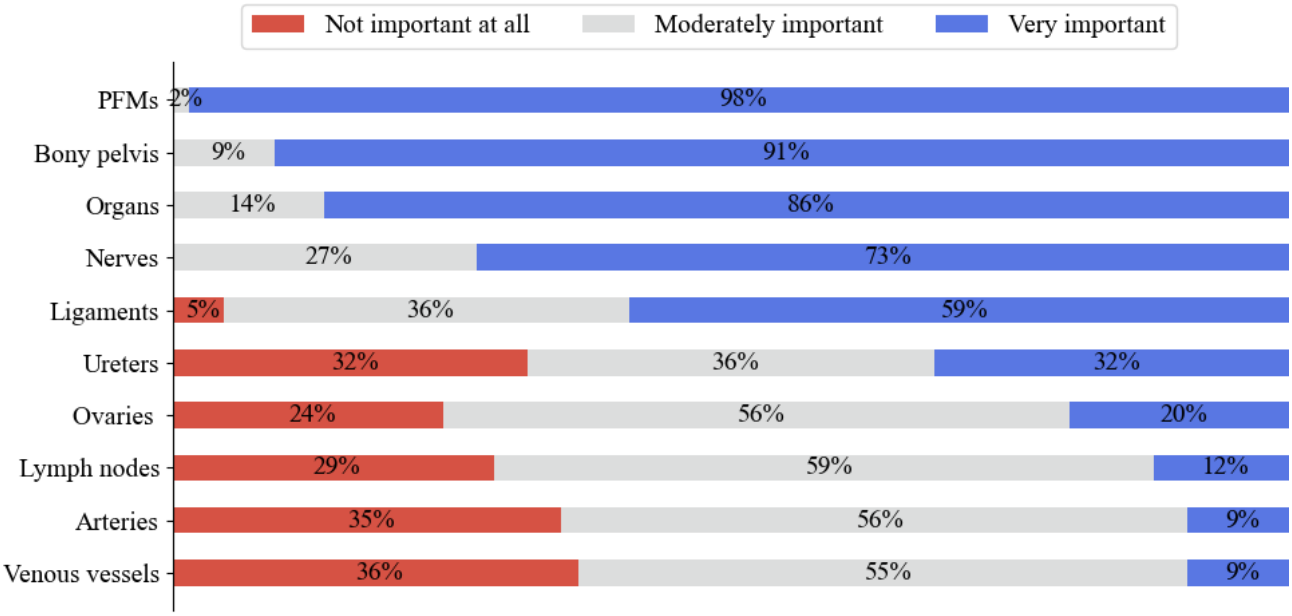


Table . Utility of functions and dysfunctions to be demonstrated.[a]

Function or Dysfunction	Values, mean (SD)
Prolapse	3 (0)
Muscle tonicity	2.95 (0.21)
Intraabdominal pressure	2.94 (0.24)
Breath	2.9 (0.48)
Stiffness	2.82 (0.39)
Posture or gait	2.75 (0.43)

[a]Responses were on a 3-point Likert scale (1=not useful at all; 2=moderately useful; and 3=very useful).

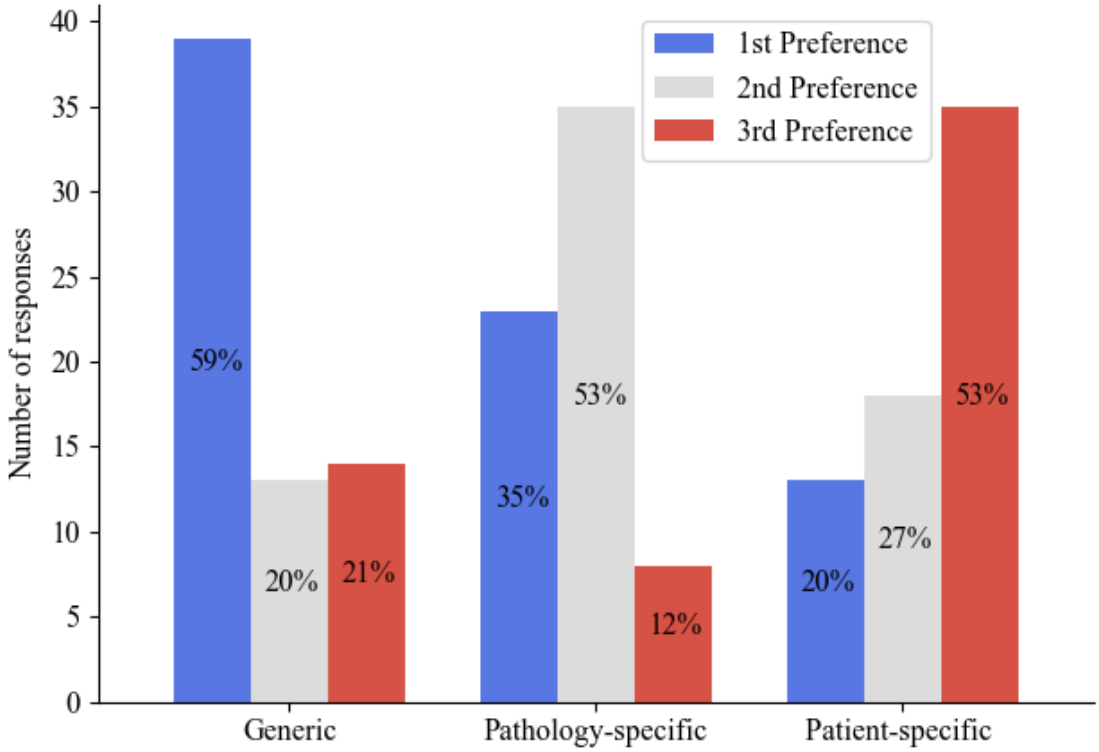
XSL•FO
RenderX

Table . Importance of imaging compatibility.[a]

Imaging modality	Mean (SD)
Magnetic resonance imaging	1.3 (0.58)
Computed tomography	1.27 (0.54)
Fluoroscopy	1.24 (0.55)
Ultrasound	1.89 (0.76)

[a]Responses were on a 3-point Likert scale (1=not useful at all; 2=moderately useful; and 3=very useful).

Figure 2. Stated preference for the specificity of the simulator.



XSL•FO
RenderX

Figure 3. Preference for feedback modalities of the simulator's sensors.

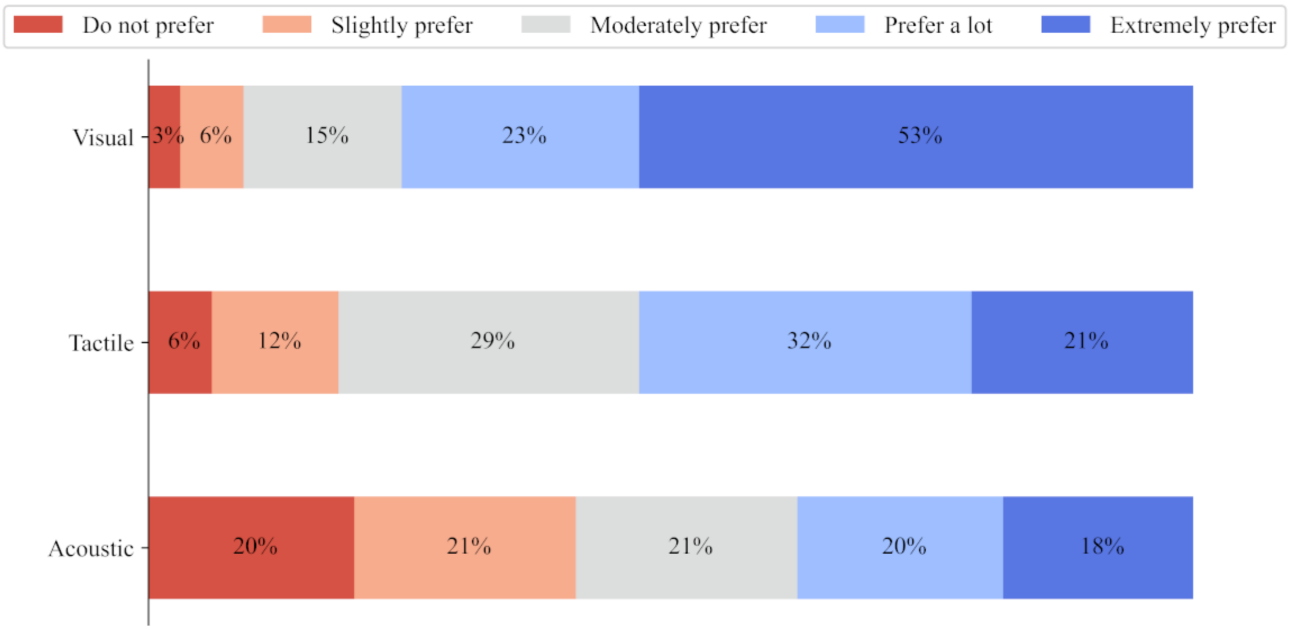


Table . Respondents' attitudes toward modularity, sensing, and terminology preferences for PFM simulators.

Questions and answers	Values, n (%)
How important is it for the pelvic floor simulator to be modular?	
Not at all important	12 (18.2)
Moderately important	29 (43.9)
Very important	25 (37.9)
What information from sensors would be useful to provide?	
Unsure	1 (1.5)
Positional tracking	2 (3)
Pressure applied	8 (11.9)
Position and Pressure	55 (82.1)
Which term do you prefer to describe a pelvic floor simulator?	
Simulator	43 (64.2)
Model	22 (32.8)
Phantom	0 (0)

Qualitative Results

Overview

The most common areas where better training needs were identified were the assessment of muscle tone (normal contractions, hypertonicity, and tone variation) (n=27) and POP (n=22). The need for better vaginal (n=11) and rectal (n=10) examination training was also identified, as was the need for anatomical mapping skills (n=4), that is, feedback on which muscle is palpated internally. Other areas include levator ani avulsion (n=6), vulvodynia (n=4), dyssynergy (n=3), pessary fitting (n=3), vaginismus (n=3), and nerve dysfunction (n=3). In total, 7 respondents felt that there were no areas where better training was needed.

Table 5 presents the most common training needs with respect to the respondents' years of professional experience. For example, 62.5% (5/8) of physiotherapists with less than 5 years of experience reported a need for better training in identifying and assessing POP. This was almost triple that of physiotherapists with over 10 years of experience (10/44, 22.7%). Although only 8 respondents had less than 5 years of experience, better training for pelvic examination (rectal and vaginal) and anatomy mapping skills were not mentioned by them.

Three main themes emerged from the thematic analysis of the open-ended questions, summarized in the following sections. The themes, subthemes and a subset of the respondents' quotes are shown in Table 6. In the optional further comments, 9 respondents expressed positive sentiments and excitement regarding a dynamic pelvic floor simulator as a potential

XSL•FO
RenderX

outcome of this research. A total of 13 respondents expressed concern over a potential simulator, however these were mainly to do with the cost and implementation and are included in the analysis in Section 3.3.4.

Table . The most common areas where needs for better training were identified per physiotherapists'' years of professional experience.

Training needs	<5 years, n (%)	5 - 10 years, n (%)	>10 years, n (%)
None	1 (13)	—[a]	6 (14)
Tone variation	3 (38)	9 (64)	15 (34)
Prolapse	5 (63)	7 (50)	10 (23)
Vaginal examination	—	2 (14)	9 (21)
Rectal examination	—	2 (14)	8 (18)
Levator Ani avulsion	1 (13)	1 (13)	4 (9)
Anatomy mapping	—	2 (14)	2 (5)

[a]Not available

XSL•FO
RenderX

Table . Emergent themes with a subset of the respondents' quotes from thematic analysis of the open-ended questions.

Themes and subthemes	Quotes
Prioritizing patient care	
Simulators can reduce the burden on patients of learning 'on-the-job'	• "There could be better ways [than peer practice], as colleagues might not have dysfunctions, so one ultimately needs to learn often about those 'on the job'" • "[We need better training for] More simulated examples of what might be a cause for concern. I found it difficult in the beginning to identify prolapse." • "[We need better training for] Correct pelvic floor contraction (have to feel a lot before you work out what is normal)."
Simulators have value in medical training and for patient education	• "I think this is a wonderful development for pelvic health training and development" • "I am excited at the prospect of having access to a pelvic floor simulator"
Patient communication is as important as practical examination skills	• "I learn as much from listening to my patients as the physical examination but I don't think this aspect [of] treating someone is very well explored in many pelvic health education courses." • "[We need better training for treating] Neurodiverse people with potential differing requirements regarding treatment. This could be related to communication style, or environmental or anatomical sensitivities, which can (should) impact how treatment is performed." • "[We need better training for] Supporting people with learning difficulties or physical disabilities in fulfilling the potential for safe physical relationships." • "[We need better training for] Transgender people and how physiotherapists and other health care professionals can support their journey better."
Simulators as supplements, not replacements to human models	• "Pelvic floor muscle function assessment techniques need human volunteers for learning and practice" • "Other online courses are available but in my experience are not as beneficial as practicing on a live model."
Representing the variability in anatomy and PFDs[a] characterizes simulator realism	
Need for simulators that are representative of anatomical variation and properties	• "[We need better training for] Variations on 'normal'" • "[We need] More examples of anatomical differences" • "[We need] More lifelike pelvic floor models. That show that tension in pelvic floor" • "[We need] A dynamic pelvic floor muscle assessment model for both Vaginal and Anorectal assessment of the pelvic floor muscle function - ideally includes superficial and deep muscle activity"
Requirements of the simulator depend on the use case and user	• "I don't really need to know about the vasculature for my examinations" • "A simple model would be better for teaching professionals new to the area of pelvic health. A more detailed model for more advanced practice or teaching." • "Patients wouldn't be likely to need to see arteries/ venous vessels or lymph nodes"
Importance of implementation, cost and accessibility in the design of simulators	
Doubt in technological capability	• "[I'm] unsure if any model could do this, for example, exam for prolapses" • "There are no simulators or models that can mimic the dynamic parameters that are tested by physiotherapist via vaginal or anorectal examination."
Practical importance of integration with health services	• "cleaning accessibility [should be considered]." • "As I work in the NHS[b] I am just mindful of getting the simulator to work alongside current IT systems"

XSL•FO
RenderX

Themes and subthemes	Quotes
Simulator features depend on affordability	• "It sounds ideal for teaching! However cost will be a big factor for NHS staff or departments" • "Important it is not too complex and expensive so that it isn't out of reach for physios to purchase. Yet sounds amazing if it can fulfill the previous suggestions" • "It's likely to be far too expensive for an NHS physiotherapy department to buy''

[a]PFDs: pelvic floor disorders.

[b]NHS: National Health Service.

Prioritizing Patient Care

Value was found in high-fidelity PFM simulators both for medical students' practice and for communicating with patients. Respondents reported that peer practice was the most common modality for practical training. However, as fellow colleagues may not present any of the variety of PFDs, respondents reported that diagnosing PFDs is learned "on-the-job" at the start of their careers. Although several respondents considered peer practice or using human volunteers as the best way to learn, respondents also indicated that an alternative option such as a simulator would be favorable.

The importance of communication skills was highlighted as much as the examination skills in treating a patient. Respondents mentioned difficulty communicating with patients or adapting to potentially differing needs of neurodivergent patients. While a simulator may not always be the answer in the context of improving patient communication, it may be useful as a tangible tool for some of the areas where respondents reported difficulties in their practice. For example, challenges in supporting differently abled patients in fulfilling safe physical relationships and considering sensitivities in examining survivors of sexual abuse or trauma. A simulator may assist a practitioner as a demonstrative tool in communicating with such patients.

Representing the Variability in Anatomy and PFDs for Simulator Realism

The use of simulators was the second most common practical training modality reported. However, respondents recognized that existing simulators lack realism. Specifically, pelvic models with no tension in the pelvic floor.

The anatomical structures to be included in a simulator were reported to vary depending on user and use case. This agrees with the aforementioned commonly shared opinion that the appropriate level fidelity is contingent on the type of task it is used for and the expertise of the user [4,7,8]. For example, if used to demonstrate the symptoms of a condition to patients, it would not need the lymph nodes, whereas these are important anatomical structures for oncologists. However, the same requirements were not shared by all respondents. Some considered vasculature to be important for patients, and others did not consider it to be relevant. Overall, respondents suggested that there should be less detail if the simulator is used to educate patients than if it is used for health care professionals. Some respondents thought no changes would be needed, while some considered this to be on a case-by-case basis.

Consideration of Implementation, Cost, and Accessibility of Simulators

The affordability of a simulator that could offer some of the features included in the survey was a shared concern amongst respondents, specifically for the United Kingdom's National Health Service (NHS) departments. The challenge of logistical integration within the NHS computer systems was also noted. For example, if the simulator needed to be plugged in to a computer to provide visual feedback on a screen. Although no specific simulator was proposed in his study, 2 respondents had low faith that any simulator or model could offer some of the features mentioned.

Design Requirements

In total, 6 recommendations for the design of future pelvic floor simulators have been formulated as a result of the survey findings and are displayed in Table 7.

XSL•FO
RenderX

Table . Design recommendations elicited from survey responses.

Design recommendation	Explanation
Realistic materials	Materials with similar properties to human muscle and skin, such as silicone should be used to provide a high level of fidelity.
Demonstrates prolapse and muscle tone	A simulator should at least demonstrate increased, normal and reduced PFM[a] tone for example, with a material that has stiffness-changing properties which users can apply resistance as if assessing actual muscle tone. This would also allow medical devices such as PFM trainers to be evaluated for their ability to differentiate between increased, normal or reduced tone.
Modular in design	A simulator should comprise adaptable anatomical complexity for patients and professionals with different experience levels
Position and pressure sensing	The palpation location should be tracked, and forces applied should be measured, both to be displayed and fed back to the user
Range of feedback modalities	Feedback modalities should be customizable according to preference. Feedback would be useful for supervisors or teachers and observers to better understand (ie, see and hear) the interactions between the simulator and the user, who would be the only one receiving tactile feedback.
Affordable and accessible	A simulator should be affordable and be designed in line with the relevant curriculum to maximize the number of people able to access it.

[a]PFM: pelvic floor muscle.

Discussion

Principal Findings

The results from our survey agreed with and built on previous research on the design and use of 3D pelvic simulators and models. All but one of the respondents reported that they would use a pelvic floor simulator for at least one use case. In addition, respondents held strong value in the simulator as a tool for demonstrating pelvic floor function and dysfunction which supports the existing literature [15,21,23].

The anatomical structures rated of high importance, illustrated in Figure 1, confirm those found by Meyer et al [18], namely, the bones, muscles, organs, and nerves. However, the experts consulted in their Delphi study considered some ligaments to be unimportant to include, and ovaries and veins as important, with the opposite being the case in our results. This discrepancy may be explained by the difference in health care professionals consulted in their study (basic scientist, obstetrics and gynecology clinician, and urologist). For example, ovaries are not always easy to palpate in a pelvic examination if a patient is overweight or unless there is a cyst or dysfunction [34]. As a result, the inclusion of ovaries may be of less importance to physiotherapists in the context of skill-based training than would be the case for gynecologists who would need to refer to these structures in surgical training [35]. Commercial pelvic simulators demonstrate this distinction, for example, the surgical [36] simulator from limbs and things includes the ovaries whereas the clinical simulator [37] does not.

As aforementioned, the level of fidelity depends on the skills being taught and the intended users [4,7,8]. High-fidelity simulation is often associated with high cognitive load and may be better suited for more experienced trainees or continuing professional development. This was echoed in our results whereby respondents suggested the simulator would need different levels of complexity depending on the user and considered modularity as a useful feature, which lead to including modularity as a design recommendation. Removing certain anatomical structures through a modular function has also been suggested in the evaluation of a virtual PFM anatomy simulator [38].

Learning from GTAs or through peer practice is considered the best form of training for learning and practicing the pelvic examination [39,40]. Training with GTAs has been shown to improve confidence and decrease anxiety for performing examinations in clinical settings [41], as has training with simulators [42]. Respondents mirrored this whereby it was reported that although peer-practice provides the most realism, an alternative would be beneficial. Pelvic simulators can be integrated into the curriculum alongside GTAs or peer practice, providing a supplementary tool that can offer repeated practice while reducing both the risk of causing harm to a peer or GTA and medical students' fear of causing it [43,44]. This evidence supports the design recommendation of using realistic materials and materials with smart properties to provide higher fidelity.

Feedback in medical simulators has been identified as one of their most valuable features [45,46]. Similar to the feedback that would be received through GTA or peer practice, feedback from sensors embedded in simulators can provide the information required to conduct the examination effectively. Figure 3 shows that the preferred modes of feedback from sensors were foremost through a visual display and second, by a tactile vibration or applied pressure which have been previously reported as the 2 commonly preferred modalities of urologists [47]. In physical examinations, physiotherapists rely on a combination of tactile, visual, and acoustic information. For internal examinations, tactile feedback would be the most realistic to integrate into a simulator as visual and acoustic cues would be more indicative of a patient's facial and uttered expressions to signal discomfort [48]. While the use of feedback

XSL•FO
RenderX

is included as a design recommendation, it would be important to consult with end-users during the design process to optimize how this feature is integrated. Without mentioning the price of a potential simulator in the survey itself, several respondents expressed concern over the cost and resulting lack of accessibility or included functionalities. Given that most respondents were working in the United Kingdom at the time of participating in the survey, there was specific mention of NHS departments being unable to afford a simulator that offered such features. It may be that some of the questions were answered with this consideration in mind. However, a study of over 400 final year medical students in the United Kingdom found that simulator-based teaching is less expensive per student for pelvic examination training than GTA-led teaching [49].

Certain features, such as sensors which were of interest to users, may increase the cost of a simulator beyond budget constraints despite being included as a design recommendation. Through careful design and consultation with end-users, a balance can be struck between cost and fidelity, ensuring the simulator provides high-quality, curriculum-relevant training while remaining affordable and accessible.

As a tool for patient education, the use of 3D models has proven efficacy in enhancing communication with patients [50], ensuring young patients' parents understand their conditions [51,52], and improving the attitudes toward pelvic examination by teaching nonmedical professionals how to perform the examination [53]. The survey results indicated that patient education was the second most popular use case of a pelvic floor simulator, with respondents highlighting several ways in which it could be used to assist in specific conditions. At later design stages, it would be beneficial to include patients in the design process as potential end-users themselves.

Limitations and Further Work

Further work should explore other pelvic health experts' opinions in equal detail and widen the participation to other countries as well as to patients with PFDs, who could also be users of such a simulator. For simplicity, anatomical structures were grouped to reduce the length of the survey. In the design stages of a given simulator, the importance of including some individual anatomical structures should be considered. For example, the pudendal and sciatic nerves are important structures that can cause pelvic pain when damaged whereas smaller, less significant nerves may be less relevant for a given use-case.

Conclusions

Our study highlighted the value of dynamic and high-fidelity pelvic floor simulators for pelvic health physiotherapists in their own education and as a tool to aid communication with their patients. The primary outcome was 6 recommendations for the design of future dynamic pelvic floor simulators, specifically for physiotherapists needs. Our findings revealed that while there is a strong need for educational tools to supplement the pelvic health curriculum, the cost of such innovations is an important factor that needs to be considered to ensure they are accessible to those who can benefit from them.

This contribution provides pelvic simulator designers with the knowledge of the needs and concerns of a crucial target user group which can better inform their product development. The importance of end-user involvement at each stage of the design process is paramount to ensuring such tools will serve the community of pelvic physiotherapists and their patients.

Acknowledgments

The authors would like to thank the respondents for their time and contributions to the study. This research was funded by the EPSRC Centre for Doctoral Training in Future Autonomous and Robotic Systems - FARSCOPE (grant number EP/L015293/1).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Questionnaire questions and answer formats.
[DOCX File, 262 KB - humanfactors_v12i1e72119_app1.docx]

References

1. RCOG calling for action to reduce number of women living with poor pelvic floor health. 2023. URL: https://www.rcog.org.uk/news/rcog-calling-for-action-to-reduce-number-of-women-living-with-poor-pelvic-floor-health [accessed 2025-05-09]
2. National Institute for Health and Care Excellence. Urinary incontinence and pelvic organ prolapse in women: management [NICE guideline no. 123]. 2019. URL: https://www.nice.org.uk/guidance/ng123 [accessed 2025-05-07]
3. Delany C, Frawley H. A process of informed consent for student learning through peer physical examination in pelvic floor physiotherapy practice. Physiotherapy 2012 Mar;98(1):33-39. [doi: 10.1016/j.physio.2011.04.347] [Medline: 22265383]
4. Gardner R. Simulation and simulator technology in obstetrics: past, present and future. Expert Rev Obstet Gynecol 2007 Nov;2(6):775-790. [doi: 10.1586/17474108.2.6.775]
5. Le Lous M, Dion L, Le Ray C. Simulation training for pelvic examination: a systematic review. J Gynecol Obstet Hum Reprod 2023 Dec;52(10):102666. [doi: 10.1016/j.jogoh.2023.102666] [Medline: 37739264]

XSL•FO
RenderX

6. Blackstock FC, Jull GA. High-fidelity patient simulation in physiotherapy education. Aust J Physiother 2007;53(1):3-5. [doi: 10.1016/s0004-9514(07)70056-9] [Medline: 17326733]
7. Maran NJ, Glavin RJ. Low- to high-fidelity simulation - a continuum of medical education? Med Educ 2003 Nov;37 Suppl 1:22-28. [doi: 10.1046/j.1365-2923.37.s1.9.x] [Medline: 14641635]
8. Alinier G, Oriot D. Simulation-based education: deceiving learners with good intent. Adv Simul (Lond) 2022 Mar 18;7(1):8. [doi: 10.1186/s41077-022-00206-3] [Medline: 35303963]
9. Scerbo MW, Murray WB, Alinier G, et al. A path to better healthcare simulation systems: leveraging the integrated systems design approach. Simul Healthc 2011 Aug;6 Suppl(7):S20-S23. [doi: 10.1097/SIH.0b013e318227cf41] [Medline: 21817859]
10. Gent D, Kainth R. Simulation-based procedure training (SBPT) in rarely performed procedures: a blueprint for theory-informed design considerations. Adv Simul (Lond) 2022 May 8;7(1):13. [doi: 10.1186/s41077-022-00205-4] [Medline: 35527267]
11. Groeneveld B, Dekkers T, Boon B, D'Olivo P. Challenges for design researchers in healthcare. Design Health 2018 Jul 3;2(2):305-326. [doi: 10.1080/24735132.2018.1541699]
12. Gardner R, Raemer DB. Simulation in obstetrics and gynecology. Obstet Gynecol Clin North Am 2008 Mar;35(1):97-127. [doi: 10.1016/j.ogc.2007.12.008] [Medline: 18319131]
13. Salvador JC, Coutinho MP, Venâncio JM, Viamonte B. Dynamic magnetic resonance imaging of the female pelvic floor-a pictorial review. Insights Imaging 2019 Jan 28;10(1):4. [doi: 10.1186/s13244-019-0687-9] [Medline: 30689115]
14. Saleme CS, Rocha DN, Vecchio SD, Barbosa MP. Simulador mecânico para o levantador do ânus [Article in Spanish]. RBEB 2009;25(2):83-87. [doi: 10.4322/rbeb.2012.068]
15. Goodwin AI, Trahan SC, Neuwirth AE, et al. Dynamic, high-fidelity, 3-dimensional pelvic model improves education on pelvic organ prolapse and the pelvic organ prolapse quantification system and interest in urogynecology. Urogynecology (Phila) 2023 Feb 1;29(2):218-224. [doi: 10.1097/SPV.0000000000001329] [Medline: 36735437]
16. Parnell BA, Dunivan GC, Geller EJ, Connolly A. A novel approach to teaching the pelvic organ prolapse quantification (POP-Q) exam. Int Urogynecol J 2011 Mar;22(3):367-370. [doi: 10.1007/s00192-010-1299-6] [Medline: 20960151]
17. Geiss IM, Riss PA, Hanzal E, Dungl A. A simple teaching tool for training the pelvic organ prolapse quantification system. Int Urogynecol J 2007 Jul 26;18(9):1003-1005. [doi: 10.1007/s00192-006-0276-6]
18. Meyer ER, James AM, Cui D. Hips don't lie: expert opinions guide the validation of a virtual 3D pelvis model for use in anatomy education and medical training. HAPS ED 2018;22(2):105-118. [doi: 10.21692/haps.2018.023]
19. Pugh CM, Youngblood P. Development and validation of assessment measures for a newly developed physical examination simulator. J Am Med Inform Assoc 2002;9(5):448-460. [doi: 10.1197/jamia.m1107] [Medline: 12223497]
20. Kiesel M, Beyers I, Kalisz A, et al. A 3D printed model of the female pelvis for practical education of gynecological pelvic examination. 3D Print Med 2022 Dec;8(1):13. [doi: 10.1186/s41205-022-00139-7]
21. Pereda-Nuñez A, Manresa M, Webb SS, et al. Pelvic + anatomy: a new interactive pelvic anatomy model. Prospective randomized control trial with first-year midwife residents. Anat Sci Educ 2023;16(5):843-857. [doi: 10.1002/ase.2304] [Medline: 37312278]
22. Advolodkina P, Chahine EB. Interactive pelvic anatomy model: a tool for teaching basic pelvic anatomy. Obstet Gynecol 2017 Oct;130(4):873-877. [doi: 10.1097/AOG.0000000000002241] [Medline: 28885415]
23. Kiesel M, Beyers I, Kalisz A, et al. Evaluating the value of a 3D printed model for hands-on training of gynecological pelvic examination. 3D Print Med 2022 Dec;8(1):20. [doi: 10.1186/s41205-022-00149-5]
24. Maglio S, Park C, Tognarelli S, Menciassi A, Roche ET. High-fidelity physical organ simulators: from artificial to bio-hybrid solutions. IEEE Trans Med Robot Bionics 2021;3(2):349-361. [doi: 10.1109/TMRB.2021.3063808]
25. Baillie S, Mellor DJ, Brewster SA, Reid SWJ. Integrating a bovine rectal palpation simulator into an undergraduate veterinary curriculum. J Vet Med Educ 2005;32(1):79-85. [doi: 10.3138/jvme.32.1.79] [Medline: 15834825]
26. Veterinary Simulators. Virtalis. URL: https://www.virtalis.com/systems/veterinary-simulators [accessed 2025-03-28]
27. Singer AP, Winklehner T, Betschart C. Interactive three-dimensional female pelvis model for teaching purposes. Am J Obstet Gynecol 2021 May;224(5):533-535. [doi: 10.1016/j.ajog.2020.10.015] [Medline: 33068523]
28. Robinson SB, Leonard KF. Chapter 5: constructing response options. In: Designing Quality Survey Questions: Sage; 2018.
29. British Society of Urogynaecology. URL: https://bsug.org.uk [accessed 2024-03-06]
30. Pelvic, Obstetric and Gynaecological Physiotherapy. URL: https://thepogp.co.uk [accessed 2024-03-06]
31. Lead midwives for education (LMEs) across the UK. URL: https://www.nmc.org.uk/education/leadmidwifery-educators [accessed 2024-03-06]
32. Kiger ME, Varpio L. Thematic analysis of qualitative data: AMEE Guide No. 131. Med Teach 2020 Aug;42(8):846-854. [doi: 10.1080/0142159X.2020.1755030] [Medline: 32356468]
33. Ahmed SK. How to choose a sampling technique and determine sample size for research: a simplified guide for researchers. Oral Oncology Reports 2024 Dec;12:100662. [doi: 10.1016/j.oor.2024.100662]
34. Padilla LA, Radosevich DM, Milad MP. Limitations of the pelvic examination for evaluation of the female pelvic organs. Intl J Gynecology & Obste 2005 Jan;88(1):84-88. [doi: 10.1016/j.ijgo.2004.09.015]
35. Miyazaki D, Matthews CA, Kia MV, El Haraki AS, Miyazaki N, Chen CCG. Validation of an educational simulation model for vaginal hysterectomy training: a pilot study. Int Urogynecol J 2019;30:1329-1336. [doi: 10.1007/s00192-0183761-9]

XSL•FO
RenderX

36. Limbs, Things. Surgical Female Pelvic Trainer. URL: https://limbsandthings.com/uk/products/60251/60251-surgical-female-pelvic-trainer-sfpt-mk-2-lightskin-tone [accessed 2024-06-06]
37. Limbs, Things. Clinical Female Pelvic Trainer. URL: https://limbsandthings.com/us/products/60935/60935-clinical-female-pelvic-trainer-mk-3-cfpt-advanceddark-skin-tone [accessed 2024-06-06]
38. Hassinger JP, Dozois EJ, Holubar SD, et al. Virtual pelvic anatomy simulator: a pilot study of usability and perceived effectiveness. J Surg Res 2010 Jun 1;161(1):23-27. [doi: 10.1016/j.jss.2009.06.016] [Medline: 19959192]
39. Janjua A, Smith P, Chu J, et al. The effectiveness of gynaecology teaching associates in teaching pelvic examination to medical students: a randomised controlled trial. Eur J Obstet Gynecol Reprod Biol 2017 Mar;210:58-63. [doi: 10.1016/j.ejogrb.2016.10.006]
40. Kirubarajan A, Li X, Got T, Yau M, Sobel M. Improving medical student comfort and competence in performing gynecological exams: a systematic review. Acad Med 2021 Sep 1;96(9):1353-1365. [doi: 10.1097/ACM.0000000000004128] [Medline: 33883396]
41. Kelly K, Wilder L, Bastin J, Lane-Cordova A, Cai B, Cook J. Utility of gynecological teaching associates. Cureus 2023 Jun;15(6):e40601. [doi: 10.7759/cureus.40601] [Medline: 37469809]
42. Orsi TD, Valadares ALR, Orsi PME, Orsi IME, Moura AS. Simulation-based training for pelvic and breast physical examination: effect on the anxiety and self-confidence of medical students. Rev Bras Ginecol Obstet 2020 Nov;42(11):739-745. [doi: 10.1055/s-0040-1718433] [Medline: 33254269]
43. Pugh CM, Obadina ET, Aidoo KA. Fear of causing harm: use of mannequin-based simulation to decrease student anxiety prior to interacting with female teaching associates. Teach Learn Med 2009;21(2):116-120. [doi: 10.1080/10401330902791099] [Medline: 19330689]
44. Jain S, Fox K, Van den Berg P, et al. Simulation Training Impacts Student Confidence and Knowledge for Breast and Pelvic Examination. MedSciEduc 2014 Mar;24(1):59-64. [doi: 10.1007/s40670-014-0023-4]
45. Issenberg SB, McGaghie WC, Petrusa ER, Lee Gordon D, Scalese RJ. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. Med Teach 2005 Jan;27(1):10-28. [doi: 10.1080/01421590500046924] [Medline: 16147767]
46. Dodge LE, Hacker MR, Averbach SH, Voit SF, Paul ME. Assessment of a high-fidelity mobile simulator for intrauterine contraception training in ambulatory reproductive health centres. J Eur CME 2016;5(1):30416. [doi: 10.3402/jecme.v5.30416] [Medline: 29644120]
47. Nakawala H, De Momi E, Tzemanaki A, et al. Requirements elicitation for robotic and computer-assisted minimally invasive surgery. Int J Adv Robot Syst 2019 Jul;16(4):172988141986580. [doi: 10.1177/1729881419865805]
48. He L, Herzig N, Lusignan SD. An abdominal phantom with tunable stiffness nodules and force sensing capability for palpation training. IEEE Trans Robot 2021;37(4):1051-1064. [doi: 10.1109/TRO.2020.3043717]
49. Janjua A, Roberts T, Okeahialam N, Clark TJ. Cost-effective analysis of teaching pelvic examination skills using Gynaecology Teaching Associates (GTAs) compared with manikin models (The CEAT Study). BMJ Open 2018 Jun 22;8(6):e015823. [doi: 10.1136/bmjopen-2017-015823] [Medline: 29934378]
50. Traynor G, Shearn AI, Milano EG, et al. The use of 3D-printed models in patient communication: a scoping review. J 3D Print Med 2022 Mar;6(1):13-23. [doi: 10.2217/3dp-2021-0021] [Medline: 35211330]
51. Entezami P, Spurgas MP, O'Brien MW, Newman LC, Adamo MA. Utility of 3-dimensionally printed models for parent education in pediatric plagiocephaly. PEC Innov 2022 Dec;1:100077. [doi: 10.1016/j.pecinn.2022.100077] [Medline: 37213734]
52. Yang T, Tan T, Yang J, et al. The impact of using three-dimensional printed liver models for patient education. J Int Med Res 2018 Apr;46(4):1570-1578. [doi: 10.1177/0300060518755267] [Medline: 29436243]
53. Siwe K, Berterö C, Pugh CM, Wijma B. Medicine meets virtual reality 17. In: Use of Clinical Simulations for Patient Education: Targeting an Untapped Audience: IOS Press; 2009:325-330. [doi: 10.3233/978-1-58603-964-6-325]

Abbreviations

GTAs: Gynecological Teaching Associates
NHS: National Health Service
PFD: pelvic floor disorder
PFM: pelvic floor muscle
POP: pelvic organ prolapse
POP-Q: pelvic organ prolapse quantification

XSL•FO
RenderX

Edited by D Gooch; submitted 07.02.25; peer-reviewed by G Alinier, MA Hefny; revised version received 28.03.25; accepted 29.03.25; published 15.05.25.

Please cite as:
Zekaria Y, Tzemanaki A, Rossiter J
Requirements and Value Elicitation for a High-Fidelity Pelvic Floor Simulator for Physiotherapists: Mixed Methods Study
JMIR Hum Factors 2025;12:e72119
URL: https://humanfactors.jmir.org/2025/1/e72119
doi:10.2196/72119

© Yael Zekaria, Antonia Tzemanaki, Jonathan Rossiter. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 15.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Identifying Design Requirements for an Interactive Physiotherapy Dashboard With Decision Support for Clinical Movement Analysis of Musicians With Musculoskeletal Problems: Qualitative User Research Study

Eduard Wolf[1,2,3], MSc; Karsten Morisse[3], Dr, Prof; Sven Meister[1,2], Dr, Prof

[1]Health Informatics, Faculty of Health, School of Medicine, Witten/Herdecke University, Witten, Germany

[2]Healthcare Department, Fraunhofer Institute for Software and Systems Engineering, Speicherstraße 6, Dortmund, Germany

[3]Department of Electrical Engineering and Computer Science, Faculty of Engineering and Computer Science, Osnabrück University of Applied Sciences, Osnabrück, Germany

Corresponding Author:
Eduard Wolf, MSc
Health Informatics, Faculty of Health, School of Medicine, Witten/Herdecke University, Witten, Germany

Abstract

Background: Performance-related musculoskeletal disorders are common among musicians, requiring precise diagnostic and therapeutic approaches. Physiotherapists face unique challenges due to the complex relationship between musculoskeletal health and the demands of musical performance. Traditional methods often lack the necessary precision for this specialized field. Integrating clinical movement analysis (CMA) with clinical decision support (CDS) could improve diagnostic accuracy and therapeutic outcomes by offering detailed biomechanical insights and facilitating data-driven decision-making.

Objective: This study aimed to identify design requirements for an interactive dashboard that aids clinical decision-making by incorporating CMA to assist physiotherapists in managing musculoskeletal disorders in musicians.

Methods: A qualitative user research study was conducted, using human factors engineering methods from problem-driven research, user-centered design, and decision-centered design. Data collection included a domain-specific literature review, workflow observations, and focus group discussions with domain experts, including 4 physiotherapist experts and an expert for clinical reasoning and applied biomechanics. This qualitative data was triangulated to characterize the domain, identify the CMA workflow, user needs, key cognitive tasks, and decision requirements. These insights were translated into concrete design requirements.

Results: A workflow for integrating musician-specific CMA into physiotherapy was established. In total, 21 user requirements, 7 key cognitive tasks, and 5 key decision requirements were defined, along with 49 design seeds. Key features identified include (1) efficient integration of musician-specific biomechanical findings into therapy, (2) combining heterogeneous data types for holistic assessment, (3) providing an adaptive overview of patient-related information, (4) using adequate visual representations and interaction techniques, (5) facilitating efficient visual-interactive analysis of findings and treatment results, and (6) enabling preparation and export of therapy findings and analysis results. Additionally, 14 CDS recommendations and 11 technical prerequisites were identified. These requirements guide the design of an interactive tool featuring advanced visualization, interactive data exploration capabilities, and contextual integration of clinical and biomechanical data.

Conclusions: An interactive physiotherapy dashboard with CDS incorporating CMA data holds significant potential to enhance decision-making in physiotherapy for musicians with performance-related musculoskeletal disorders. By addressing cognitive demands and integrating advanced visualization techniques, the tool can support physiotherapists in making more accurate assessments, potentially improving patient outcomes, reducing injury recurrence, and supporting musicians' career longevity. Ongoing research is essential to refine such a tool and validate its usability, decision support, and clinical effectiveness. Future work should explore advanced analytics, adapt to various CMA systems, and expand applications across musicians and therapeutic domains to enhance its impact.

(JMIR Hum Factors 2025;12:e65029) doi:10.2196/65029

KEYWORDS

physical therapy; clinical movement analysis; clinical decision support; performance-related musculoskeletal disorders; requirements engineering; human factors engineering; design study methodology; problem-driven research; user-centered design; decision-centered design

XSL•FO
RenderX

Introduction

Background

Musculoskeletal disorders are common among musicians due to repetitive movements, muscle strain, poor posture, and intensive practice, affecting their ability to perform and sustain careers [1-3]. At least 50% of musicians experience performance-related musculoskeletal disorders (PRMDs), which cause pain and other symptoms that hinder performance [4-6]. However, targeted therapy is underused, and specialized assessment and treatment strategies are needed [7,8].

Musician-specific physiotherapy is beneficial for identifying and treating these issues, focusing on aspects of posture, movement behavior, and musical expression that affect health [8-10]. Effective treatment requires a thorough understanding of musician-specific characteristics and identification of functional disorders in the musculoskeletal system through clinical neuro-orthopedic examinations and analyses of the musician's movement while performing [11].

In physiotherapy, clinical reasoning is essential for decision-making, often using a hypothetical-deductive method [12]. Accurate diagnostics are crucial to prevent misjudgments [13], and comprehensive information gathering is vital for effective therapy planning [14,15]. Technology-driven assessments can enhance diagnostic accuracy, moving beyond subjective methods and trial-and-error approaches [16].

Clinical movement analysis (CMA) provides objective, quantifiable data on functional disorders, supporting personalized medicine [11,17] and offering physiotherapy applications that conventional assessments cannot achieve [18,19]. As an apparatus-based biomechanical analysis for individualized functional diagnostics, CMA uses techniques such as motion capture and electromyography [20,21]. It helps evaluate complex movements during performance, allowing for the identification of aberrant postures and movements contributing to PRMDs in musicians [22-24]. Thus, CMA aids in diagnosing functional disorders, planning individualized therapy, and monitoring therapeutic measures [17,19,25], minimizing the risk of false results [26,27].

The *RefLabPerform* project [28] aims to develop a reference laboratory for PRMD assessment in musicians, integrating CMA with physiotherapeutic assessments, which will be implemented at the Institute for Applied Physiotherapy Osnabrück (INAPO). Advanced instruments will be used by specialists for musician-specific CMAs, with results provided to physiotherapists for further clinical evaluation.

However, interpreting CMA data requires expertise and time, presenting challenges for clinical practice, highlighting the need for efficient, intuitive presentation for analysis and decision-making [11,29,30]. Current tools for biomechanical analysis lack guidance for musicians' physiotherapy, leading to concerns about accuracy and usability [30,31].

A tailored health technology for musicians' physiotherapy is essential to efficiently record, process, provide, analyze, and interpret musician-specific CMA data for routine use. Hereby, an interactive dashboard with clinical decision support (CDS) can improve diagnostics by integrating CMA findings, facilitating comprehensive decision-making and effective therapies [32]. Adapting research approaches from health while considering physiotherapy-specific conditions will ensure the tool aligns with the clinical workflow, balancing perceived benefit, effort, and costs [16,33].

Human Factors and CDS Design

Human factors engineering (HFE) is crucial for designing interactive health technologies and CDS tools, focusing on usability and workflow integration for decision support [34-36]. Effective CDS requires speed, efficiency, meaningful alerts, consistency, and logical grouping with intuitive interface design [37-40]. Hence, designing such a technology must address user needs, working environments, cognitive demands, decision-making requirements, and advanced interactive visualizations for data analysis.

In total, 3 HFE approaches—design study methodology (DSM), user-centered design (UCD), and decision-centered design (DCD)—provide guidance for this. DSM is a problem-driven research paradigm focused on creating innovative visualization solutions through iterative collaboration with domain experts to address real-world problems [41]. UCD emphasizes understanding user tasks, goals, and needs by involving users from the beginning, which informs the specification of user requirements and guides the design process [42]. DCD aims to improve decision-making in critical scenarios, such as medical diagnosis, by enhancing human decision-making [43,44]. Cognitive task analysis identifies key cognitive demands and translates them into decision requirements, guiding the design process to improve clinical performance [45]. These approaches share a human-centered, iterative, and multistage methodology, leveraging in-depth analysis and expert contributions, while each addresses distinct aspects of the design process.

Objectives

The study aimed to identify design requirements for an interactive tool, specifically a physiotherapist dashboard, that aids clinical decision-making for musculoskeletal problems in musicians within the *RefLabPerform* project. This involved considering the working environment, clinical procedures, and user needs from the beginning and gathering detailed information on how decision-making in musicians' physiotherapy can be supported technically, especially with the integration of CMAs.

To achieve this, requirements elicitation methods from DSM, UCD, and DCD were followed to comprehensively describe the domain and the underlying problem, resulting in well-informed requirements and recommendations. This approach is expected to enhance usability and acceptance, improve user efficiency and accuracy, and reduce the cognitive workload of the future tool.

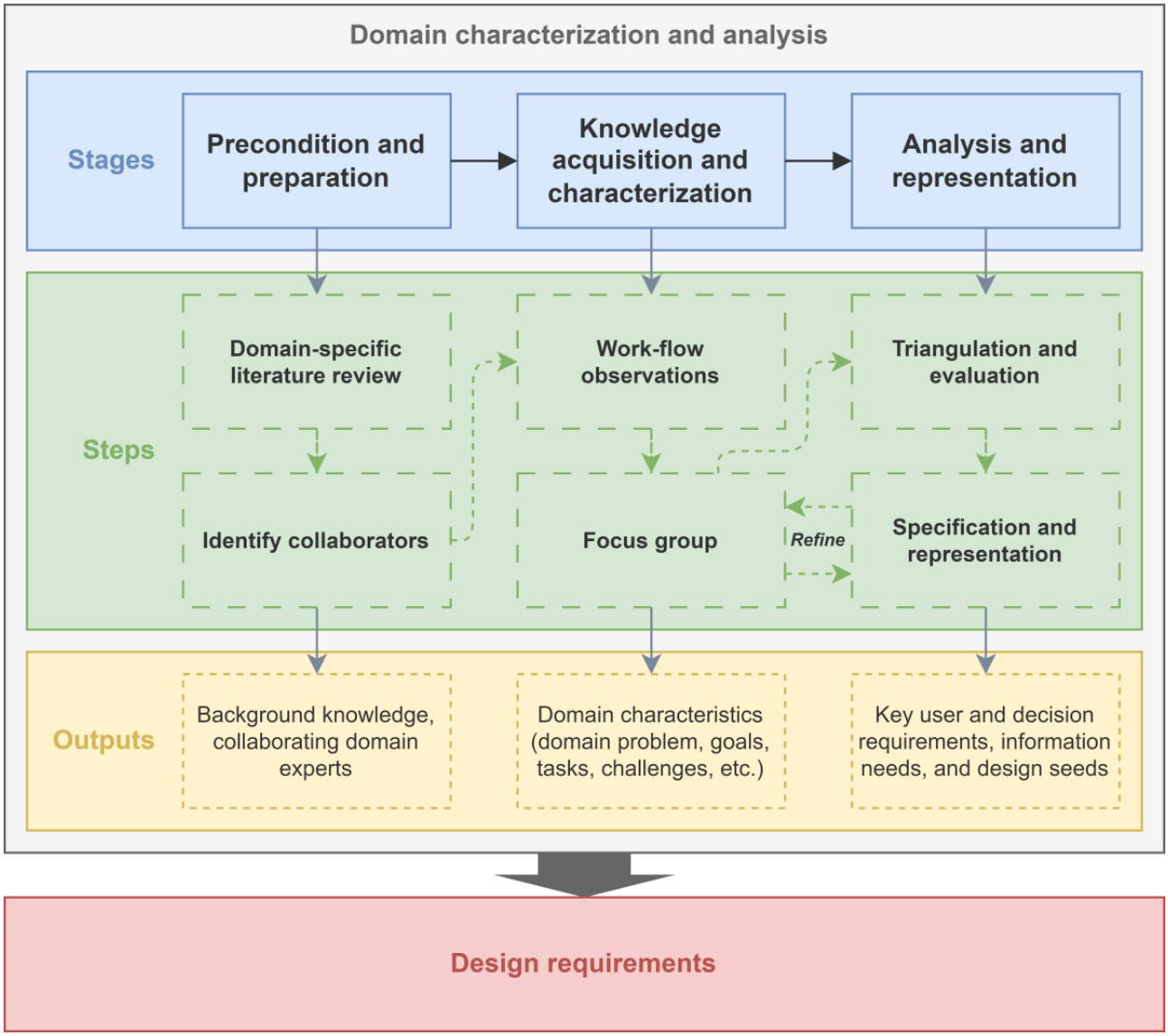
XSL•FO
RenderX

Methods

Research Approach

This qualitative user research study followed a problem-driven research paradigm and formed part of the initial *domain characterization and analysis* phase within a larger design study [46]. It systematically integrated elements from DSM, UCD, and DCD to elicit and structure design requirements for an interactive dashboard with CDS functionality in physiotherapy (see Figure 1). The requirements elicitation process drew on the first 4 steps of DSM (*learn*, *winnow*, *cast*, and *discover*) [41], the initial 2 phases of UCD (*understanding context of use* and *specifying user requirements*) [42], and the first 3 stages of DCD (*preparation*, *knowledge elicitation*, and *analysis and representation*) [44]. These approaches were merged and operationalized into three stages: (1) precondition and preparation, (2) knowledge acquisition and characterization, and (3) analysis and representation, with subsequent design and evaluation phases planned for future work. DSM ensured methodological rigor and problem characterization through iterative domain expert engagement; UCD emphasized contextual inquiry and stakeholder needs; and DCD supported the analysis of complex cognitive processes in physiotherapy decision-making. Together, they provided a robust foundation for deriving clinically relevant, user-informed, and cognitively aligned design requirements. A detailed mapping of their contributions is presented in Table 1.

Figure 1. The domain characterization and analysis process comprised 3 stages with multiple steps and outputs, ultimately defining design requirements for an interactive physiotherapy dashboard.



XSL•FO
RenderX

Table . Detailed mapping of each method's contributions to domain characterization and analysis.

Research stage	Design study methodology (DSM)	User-centered design (UCD)	Decision-centered design (DCD)
Precondition and preparation	• Domain understanding • Relevant terms and jargon • Identification of collaborators	• Understanding user group characteristics and context of use	• Identification of cognitively challenging activities from previous task descriptions
Knowledge acquisition and characterization	• Focus on real-world domain problem characterization • Domain expert engagement	• Participatory engagement through workflow observations and focus groups • Elicitation of user needs • Focus on user goals and task alignment	• Identification of cognitive demands, critical cues, and information needs
Analysis and representation	• Iterative synthesis of findings into design seeds • Structured representation of design implications	• Specifying user requirements • Integration of usability concerns • Clarity and traceability of user needs to system design	• Specifying decision requirements • Integration of cognitive support needs in clinical workflow • Translation of decision-making needs into design recommendations

The main author (EW) conducted the workflow observations and moderated the focus groups in German in March and April 2021. His researcher characteristics are the following: "gender: male," "experience: 6 years research experience in health informatics, main focus in physiotherapy and movement analysis projects," "degree: M.Sc. in Computer Science," and "occupation: Research assistant."

Ethical Considerations

The study was conducted in accordance with the Declaration of Helsinki and approved by the institutional review board (or Ethics Committee) of Osnabrueck University of Applied Sciences (protocol code HSOS/2021/1/1, date of approval March 17, 2021). Signed informed consent for publication was obtained from all participants after providing them with a written information sheet and verbal explanation of the study context, the intended procedure, and data usage. All participants were given a copy of the signed informed consent as well as confirmation of anonymization of study data. The study data were anonymized, so participants cannot be identified. Participants did not receive any financial compensation for their participation.

Precondition and Preparation

Domain-Specific Literature Review

Domain-specific literature from physiotherapy, clinical biomechanics, and movement analysis of musicians was reviewed to gather background information about the domain and user group characteristics and to become familiar with relevant terms and jargon [6,8-11,13-15,17,19,21,23,29,31,47-58]. Previously articulated task descriptions, such as clinical reasoning and functional diagnostics, were examined to gain domain understanding and to initially identify cognitively complex components.

Identify Collaborators

The INAPO, a specialized facility for treating neuromusculoskeletal disorders, served as a cooperation partner. INAPO has specialized in the treatment of musicians in close cooperation with the "Institute for Music" at Osnabrück University of Applied Sciences. Admitted patients receive tailored treatment and exercise programs. INAPO emphasizes staying updated with scientific progress in musicians' health and movement sciences, using modern therapy methods. As part of the *RefLabPerform* project, a CMA lab was established at INAPO, equipped with optoelectronic motion capture, surface electromyography sensors, and force plates. Requirements were gathered based on the CMA lab and INAPO's clinical practice needs.

In total, 4 physiotherapists (1 male, 3 female) from INAPO and an expert in clinical reasoning and applied biomechanics from the University of Applied Sciences Osnabruck were selected face-to-face as domain experts for systematic collaboration during requirements elicitation. Purposive sampling ensured the inclusion of a representative sample of practicing physiotherapists with extensive experience in managing musicians with musculoskeletal problems.

Knowledge Acquisition and Characterization

Workflow Observations

We used contextual inquiry to observe the workflow of physiotherapists managing musicians with musculoskeletal problems at INAPO prior to the focus group discussion. Contextual inquiry, a qualitative user research method, involves observing and interviewing potential users in their work environment as they perform actual tasks [59], ie, watching them perform tasks while asking questions about their actions and thought processes. This method uncovers tacit knowledge, reveals detailed work practices, and provides reliable, relevant information about user needs and behaviors. The goal was to identify and gather domain experts' needs and preferences for a physiotherapist dashboard and CDS concerning their (cognitively challenging) tasks and daily data encounters. Additionally, we asked the experts about their thoughts on the

XSL•FO
RenderX

integration of musician-specific CMAs into the diagnostic process and the application in clinical decision-making.

EW observed 3 physiotherapists experienced in managing musicians with musculoskeletal problems, previously selected for collaboration, during their daily work routines over 1 week. Each session lasted about 2 hours, allowing for detailed observation and discussion. Observations continued until thematic saturation was reached. Workflow diagrams were created from observation notes and presented at the focus group meeting to facilitate discussion on workflow processes.

Focus Group

We conducted a web-based focus group [60] lasting approximately 270 minutes (4.5 hours), comprising 7 participants: the 4 collaborating domain experts, 2 computer scientists, and 1 user experience designer. Video and audio recordings were made using Zoom (Zoom Video Communications, Inc). The focus group aimed to gain a deeper understanding of the domain, exploring the working environment, clinical procedures, user needs and decision-making processes in musicians' physiotherapy, providing essential information for the tool's design. The primary goal was to establish a common understanding of the domain and address the following key discussion topics for the specific setting: goals, skills, and barriers in musicians' physiotherapy; required collaborations, tools, and equipment; workflow structure and objectives of each step; information, tools, and activities per work step; and motion analysis lab workflow and physiotherapist interaction.

Given the lack of standardized procedures for focus group data collection and analysis, which depend on the research question and time frame, we opted to organize the results into a mind map during the discussion to facilitate consensus-building. Following the focus group session, the research team held a debriefing session to review the discussions and identify key points.

Analysis and Representation

Triangulation and Evaluation

The qualitative data was triangulated and evaluated to determine how the tool can address the domain problem and user needs, and support decision-making. First, the insights from workflow observations and the focus group were organized and contextualized with domain-specific literature to understand and describe the domain characteristics. Further evaluation of the findings revealed key user needs, as well as cognitively challenging tasks, strategies, difficulties, and critical aspects in musicians' physiotherapy.

Specification and Representation

Based on the prior findings, the solution's characteristics were specified. First, a workflow for integrating musician-specific CMA was established and visualized. Key user requirements, cognitive tasks, decision requirements, and associated information needs were then derived and formulated. Subsequently, design seeds to support decision-making were created. These findings were organized in tabular form.

In a second web-based focus group discussion, which lasted approximately 180 minutes (3 hours), the identified domain characteristics, specified user and decision requirements, and design ideas were revisited with the same group of experts and researchers. During this session, these elements were discussed in detail and refined based on their feedback.

Finally, incorporating knowledge and conditions from the theoretical background and domain, concrete design requirements were specified, including key features, technical prerequisites, and decision support recommendations.

Results

Characteristics of the Domain

To describe the domain, the characteristics were categorized under the following themes: domain problem; goals, skills, and competences, work equipment, and barriers; and tasks, routines, and challenges.

Domain Problem

The unresolved domain problem is integrating musician-specific CMA into musicians' physiotherapy and offering visual-interactive access to aggregated findings and treatment data for decision support.

Musicians' physiotherapy, an integral component of "musicians' medicine," specializes in the assessment and treatment of neuromusculoskeletal disorders prevalent among musicians. This specialty uses standard physiotherapeutic tools along with musician-specific questionnaires and CMAs to enhance diagnostic precision and facilitate data-driven therapeutic decisions. Post intervention, the collection of relevant patient questionnaires and repeated CMAs is crucial for effective therapy evaluation. However, an efficient technical solution for the comprehensive integration and analysis of clinical and biomechanical data tailored to musicians' physiotherapy remains absent. The primary beneficiaries of such a solution would be physiotherapists, with musicians gaining from improved, targeted therapies. Additionally, specialized physiotherapists or movement scientists trained in musician-specific CMAs would benefit from enhanced integration of CMA into the therapeutic process.

Documentation, typically on paper, should be electronic for comprehensive evaluation. Integrating CMAs adds complexity, requiring standardized data processing and effective communication between physiotherapists and movement labs. Current software for biomechanical analysis is not user-friendly for physiotherapists, lacking advanced visualization, interactive data exploration, and contextualization of clinical and biomechanical data. A new solution is needed to simplify data analysis and visualization for therapists, with requirements evolving during implementation.

Goals, Skills, Competences, Equipment, and Barriers

Textbox 1 summarizes the goals, skills, competences, work equipment, and (potential) barriers cited by domain experts. The primary goal of physiotherapists is to provide high-quality patient care to enhance or maintain sensorimotor self-determination. This requires addressing the patient's main

XSL•FO
RenderX

problems and goals efficiently to implement effective, targeted therapies. It is crucial to assess the effectiveness of therapeutic measures continuously, aligning interventions with the patient's condition and goals. Achieving these outcomes demands extensive skills and competencies, particularly in CR, which is central to effective therapy [58].

Textbox 1. Domain-specific goals, skills and competencies, work equipment, and barriers.

Goals:

- Enhancing or maintaining sensorimotor self-determination
- Improving the main problem
- Addressing patient goals

Skills and competencies:

Background knowledge:

- Patho-anatomy and physiology (clinical patterns, clinical picture, and musician-specific traits)
- Reference values (standard, individual, and musician-specific)
- Scientific evidence for assessments, tests, interventions

Clinical experience:

- Communication skills
- Manual skills
- Recognizing clinical patterns
- Techniques
- Test interpretation

Equipment:

- Patient questionnaires
- Diagnostic questionnaires
- Body tables
- Numerical rating scale
- Measuring instruments (eg, goniometer, algometer)

Barriers:

Physiotherapist:

- Lack of background knowledge
- Lack of clinical experience
- Shift in knowledge and beliefs, eg, through further education and training
- Premature judgments

Patient:

- Compliance
- Background (cognitive, social, and educational)
- Negative experiences
- Unreasonable expectations

General conditions:

- Legal: Limited time, lack of resources (eg, financial)
- Institutional: Limited time, lack of resources (eg, premises, measuring instruments), availability of qualified therapists

Domain experts highlight the importance of a solid background in (patho-)anatomy, physiology, normative and comparative values, and evidence-based assessments, tests, and interventions. Clinical experience is also vital for developing communication

XSL•FO
RenderX

and manual skills, refining techniques, interpreting tests, and recognizing clinical patterns. Further general requirements for physiotherapists are outlined in the literature [58]. Domain experts identified several tools used in therapy, including musician-specific patient questionnaires, body tables, numerical rating scales, goniometers, and algometers.

Physiotherapists face several barriers that can impede therapeutic objectives, including insufficient background knowledge, limited clinical experience, and shifts in knowledge due to ongoing education. Patient-related barriers include compliance issues and unrealistic expectations. Additionally, legal or institutional barriers, such as limited examination and treatment time, and insufficient financial, technical, or spatial resources, also pose significant challenges.

Tasks, Routines, and Challenges

The cyclical CR process consists of 8 main phases: looking, collecting, processing, deciding, planning, acting, evaluating, and reflecting [12]. These phases are interrelated and often overlap. The technical solution should explicitly assist the therapist in 5 of these steps (Textbox 2):

Textbox 2. Five steps that a technical solution should assist the therapist.

- **Collecting:** recording all subjective and objective findings
- **Processing:** analyzing and interpreting the findings
- **Deciding:** assessing the findings for diagnosis and planning
- **Evaluating:** assessing the interventions based on the repeat findings
- **Reflecting:** assessment of therapy effectiveness

The diagnostic process in physiotherapy involves examination, evaluation, and reassessment of findings and therapy measures under significant time pressure. Therapists must extract relevant information to make accurate treatment decisions, recognizing anomalies and changes in data [61].

Workflow observations and focus group discussions provided a deeper understanding of the therapeutic and diagnostic processes in musicians' physiotherapy. Before the first consultation, patients' complete musician- and problem-specific questionnaires, and the results guide the subjective examination, including background information and symptoms. This information helps generate diagnostic hypotheses tested during a physical examination, often using tools like goniometers. The results confirm or reject these hypotheses, leading to a physiotherapeutic diagnosis and a customized treatment plan. Treatment effectiveness is evaluated regularly, with reassessments before each follow-up to adjust or stop treatment as needed.

Specification and Representation

The findings from the prior domain characterization were further examined and discussed with the domain experts. This process aimed to address the integration of musician-specific CMA into physiotherapy, identify user requirements, and determine key cognitive tasks and decision requirements.

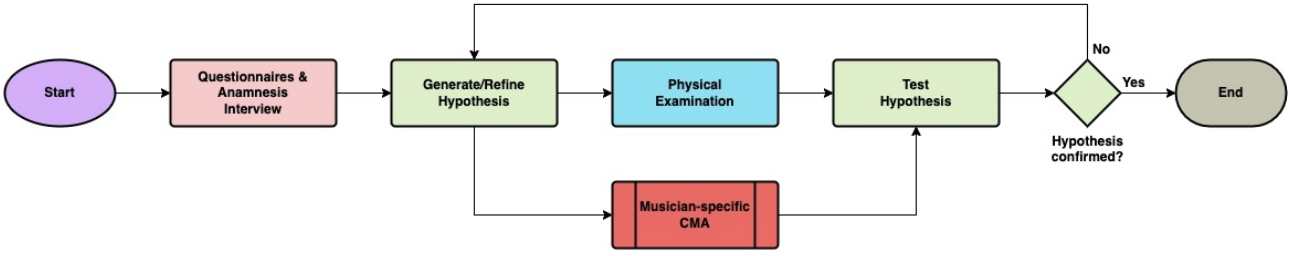
Integrating Musician-Specific CMA Into Physiotherapy

Based on the domain experts' input, the integration of musician-specific CMA into physiotherapy was designed. It is crucial to incorporate CMA efficiently, providing valuable information on a musician's dysfunction. The timing, execution, and result delivery are critical for high clinical benefit and rapid transfer from the movement analysis lab to the therapist. The CMA aims to test diagnostic hypotheses and evaluate therapeutic measures' effectiveness. Biomechanical data should be classified and interpreted within the clinical context, structured as additional findings, and made available to therapists.

Figure 2 proposes integrating CMA into the assessment process alongside the physical examination. Clinical reasoning underpins CMA use, starting with a clear question based on a working hypothesis, guiding the selection of muscles, starting positions, testing types, and biomechanical characteristics. This procedure is referred to as a movement analysis protocol based on Kontaxis et al [62].

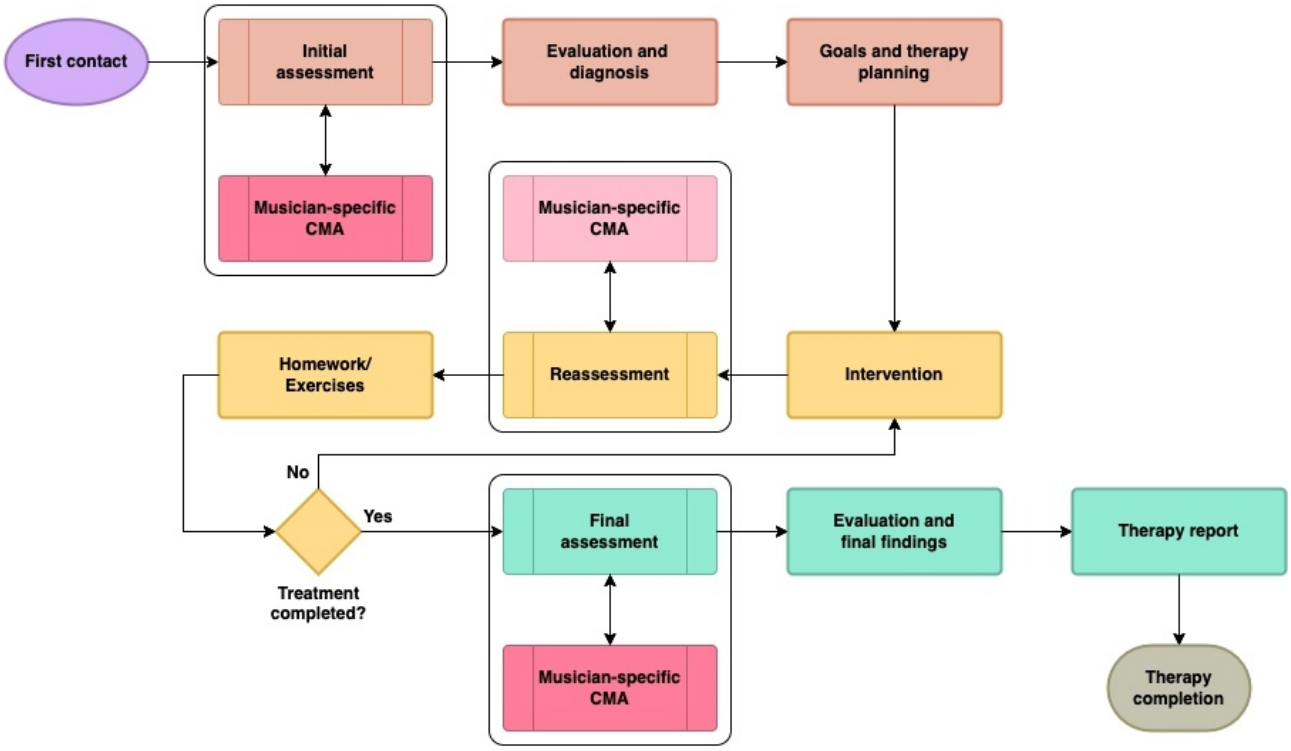
In addition to using CMA for functional diagnostics during the initial assessment, a subsequent CMA is conducted post treatment, as depicted in Figure 3. This allows for a pre-post comparison to evaluate the effectiveness of the treatment interventions.

Figure 2. The revised diagnostic process integrates clinical movement analysis (CMA) as a complementary tool alongside physical examination (own illustration). It begins with musician and problem-specific questionnaires and anamnesis, followed by the generation of diagnostic hypotheses. A physical examination is then conducted to test these hypotheses. Simultaneously, a musician-specific CMA, guided by the hypotheses and a predefined movement analysis protocol, provides objective validation. If hypotheses require refinement or new ones emerge, additional physical exams and CMA assessments are performed until the final hypotheses are confirmed.



XSL•FO
RenderX

Figure 3. The modified clinical procedure integrates clinical movement analysis (CMA) for pre-, intra-, and posttreatment assessment (own illustration). Following the initial contact, an assessment is conducted, leading to findings evaluation, final diagnosis, goal setting, and therapy planning. After intervention, a reassessment is performed, and patient homework exercises are provided. If treatment continues, further interventions and reassessments follow; if completed, a final assessment, therapy evaluation, and report are conducted before concluding therapy.



User Requirements

Contextual analysis, informed by the domain-specific literature, workflow observations, and focus groups, revealed user needs concerning the solution's characteristics. In total, 21 user stories were defined based on these needs, reflecting the users' expectations of the physiotherapist dashboard's functionality [63]. These user stories pertain to clinical findings and treatment documentation, CMA, analysis and evaluation, and reporting. The full list of user requirements is available in Multimedia Appendix 1.

Cognitive Tasks and Decision Requirements

Cognitively challenging activities were identified, such as diagnostics, treatment planning, and therapy evaluation, informed by task descriptions from literature and the findings from knowledge acquisition and characterization. For each activity, we identified critical cues used by experts, potential errors, and procedures to meet cognitive demands. Physiotherapists generate and test hypotheses by interviewing patients and conducting or commissioning examinations, continuously integrating, and interpreting the information. Each activity is a cognitively challenging demand, requiring decisions on what information to gather and when to stop, as well as determining the next steps in treatment.

Finally, seven key cognitive tasks were identified: (1) querying specific patient questionnaires, (2) subjective examination of the patient, (3) testing diagnostic hypotheses, (4) consolidating and evaluating the information relevant to therapy, (5) planning the adapted therapy, (6) treating a patient's specific problem in an iterative approach, and (7) evaluating the therapy results. The final list of key cognitive tasks and the associated strategies, (potential) errors or difficulties, and critical hints of the domain experts can be found in Multimedia Appendix 2.

In coordination with the domain experts, key decision requirements were extracted from these activities, focusing on planning and assessing examinations and treatments while considering musician-specific CMA. Information needs for each decision were defined, specifying the clinical information needed. This approach led to design recommendations for decision support, offering early human-computer interaction concepts for organizing and visualizing information.

In total, five key decision requirements were identified incorporating musician-specific CMA in the context of physiotherapeutic diagnostics: (1) plan a functional diagnostic examination; (2) assess CMA findings; (3) prepare overall findings, and, if necessary, a report of findings; (4) compare results of reassessments and evaluate treatment effectiveness; and (5) evaluate therapy results and success. Based on the key decision requirements and associated information needs, a total of 49 design seeds were identified. The final list can be found in Multimedia Appendix 3.

Physiotherapists must decide if, when, and how to perform a CMA for a patient, considering basic patient information, contributing factors, symptoms, previous CMA parameters, and time and number of previous treatments. They must quickly evaluate findings to plan next steps, gaining an overview of the

XSL•FO
RenderX

patient's problem, identifying patterns, and confirming diagnostic hypotheses to formulate a physiotherapy diagnosis. They must determine which data are therapy-relevant to create an overall finding or report.

Throughout treatment, physiotherapists continuously assess findings and treatment effectiveness, deciding if measures positively impact the patient's problem or need adjustment. They review therapy-relevant information before and after treatments to identify trends and changes. They reflect on therapy outcomes and their actions, assessing whether therapy was successful and identifying key factors. They monitor achievement of therapy and patient goals using data from baseline and follow-up findings, treatment results, and relevant questionnaires, including post-intervention CMA results.

Design Requirements

The *domain characterization and analysis* informed concrete design requirements for a physiotherapist dashboard, specifying key features and CDS recommendations based on user needs for decision support. The use case defined technical prerequisites for implementation.

Key Features

Based on the insights gained, we identified 6 key features (KF) that form the basic pillars of the tool and must be fulfilled during design and implementation:

- **KF$_1$**. *Provision of biomechanical findings:* Musician-specific CMAs should be efficiently integrated into therapy to give physiotherapists valuable insights into a patient's dysfunction. Biomechanical data must be processed, modeled, and stored to produce a report with relevant characteristics for further analysis.
- **KF$_2$**. *Integration of heterogeneous data types and sources:* For a holistic inspection and assessment of all available information from the subjective and objective examinations of a patient's problem, it is necessary to combine different types of data from different data sources. This includes clinical findings and treatment data from physiotherapy documentation, results from web-based questionnaires, and biomechanical findings data from movement analysis.
- **KF$_3$**. *Adaptive overview:* A simplified aggregated overview is essential for quick and clear inspection and monitoring of patient-related information. This should include basic therapy data, recent findings, and treatments, as well as relevant individual patient details. Information should be adaptive and patient-specific, allowing physiotherapists to manually add or remove data. Detailed views should be accessible via a drill-down function.
- **KF$_4$**. *Adequate visual representations and interaction techniques:* Efficient visualizations, such as lists, tables, diagrams, and illustrations, should be chosen to aid physiotherapists in effective decision-making. Therapists should be able to manually arrange and configure these elements and use familiar interaction techniques like drag-and-drop, sorting, searching, and filtering.
- **KF$_5$**. *Efficient visual-interactive analysis of findings and treatment results:* Physiotherapists need an efficient solution for interactive visual analysis of findings and treatment results to identify abnormalities, trends, and patterns in clinical and biomechanical data. Established visual analytics methods will be used to explore and compare multivariate and time-oriented data.
- **KF$_6$**. *Preparation and export of findings and analysis results:* Physiotherapists should be able to summarize and export therapy findings as a report for discussion with colleagues and patients, or for continued treatment elsewhere. The report should be presented in a clear, customizable format.

CDS Design Recommendations

The user needs and design seeds addressing decision requirements formed the basis for specifying concrete CDS design recommendations (DR; see Textbox 3).

For planning musician-based diagnostics, it is helpful to have a note indicating if such an examination is possible for the patient (DR$_1$). CDS should remind the physiotherapist to conduct pre- and post-interventional exams if possible (DR$_2$). Additionally, suggested configuration parameters should be provided when creating an examination order to save time (DR$_3$).

To efficiently query and inspect findings and treatment results, essential data should be preselected for the overview (DR$_4$). Information should be summarized in aggregated elements and presented in suitable display formats (DR$_5$). Relevant therapy events should be shown on a timeline (DR$_6$). Additionally, visual cues should highlight significant data changes to reduce cognitive load and speed up decision-making (DR$_7$).

For interpreting findings, integrating normative and comparative values is desirable (DR$_8$). Therapists should be able to create interactive annotations (DR$_9$), marking specific data values or time intervals. To support diagnosis, findings and annotations should be linked to diagnostic hypotheses, showing which data support or refute them (DR$_{10}$). Additionally, marking applicable or inapplicable hypotheses interactively should be possible (DR$_{11}$).

For therapy evaluation, directly comparing baseline and final findings is essential (DR$_{12}$). An interactive review of patient and therapy goals helps determine therapy success, requiring target parameters for goal achievement (DR$_{13}$). Preselecting relevant entries and suitable presentation forms aids in report preparation (D$_{14}$).

XSL•FO
RenderX

Textbox 3. Overview of the clinical decision support design recommendations (DR).

Planning musician-specific examinations for functional diagnostics

DR_1: Indicate if musician-specific examinations are possible for the patient.

DR_2: Remind physiotherapists to conduct pre- and postinterventional examinations.

DR_3: Provide suggested configuration parameters for examination orders to save time.

Query and inspection of diagnostic findings and treatment results

DR_4: Preselect essential data for efficient querying and inspection.

DR_5: Summarize information in aggregated elements and suitable display formats.

DR_6: Map relevant events on a timeline.

DR_7: Use visual cues to highlight abnormal data or significant changes.

Analysis and interpretation of the findings data

DR_8: Integrate normative and comparative values for interpreting findings.

DR_9: Allow interactive annotations for specific data values or time intervals.

DR_{10}: Link findings and annotations to working hypotheses.

DR_{11}: Mark working hypotheses as applicable or not with appropriate presentation.

Therapy evaluation and reporting

DR_{12}: Enable direct comparison of baseline and final findings for therapy evaluation.

DR_{13}: Facilitate an interactive review of patient and therapy goals, with target parameters for goal achievement.

DR_{14}: Preselect relevant entries and use appropriate presentation formats for report preparation.

Technical Prerequisites

Technical prerequisites (TP) must be established to enable the physiotherapist dashboard (see Textbox 4). The tool should allow physiotherapists to access patient data and case history (TP_1), including anamnesis, physical exams, and treatment records (TP_2). Patients must complete validated musician- and problem-related questionnaires web-based, with results integrated into the findings overview (TP_3). Biomechanical data from musician-specific CMAs should be available as findings (TP_4).

Functional diagnostics occur in a CMA lab within a physiotherapy practice, staffed by specialists. Physiotherapists must submit *electronic exam orders* with relevant CMA details (eg, patient ID, instrument group, and symptoms; TP_5). The lab requires an *order management system* to track requests and update CMA status (TP_6) and a *data management system* for recorded measurements (TP_7). Standardized, musician-specific CMA protocols should be implemented (TP_8), integrating existing measurement tools and software (TP_9). Data processing and calculations should be largely automated (TP_{10}), with results provided as biomechanical findings (TP_{11}).

XSL•FO
RenderX

Textbox 4. Overview of the technical prerequisites (TP) to enable the physiotherapist dashboard.

Technical prerequisites (TP)

Integration of therapy-related patient data

TP_1: Integration of patient and case data

TP_2: Integration of clinical findings and progress data

TP_3: Integration of the web-based questionnaire results

TP_4: Integration of the biomechanical findings data

Laboratory system

TP_5: Electronic order entry

TP_6: Laboratory order management

TP_7: Laboratory data management

TP_8: Standardized motion analysis protocols

TP_9: Integration of existing biomechanical measuring instruments or systems and analysis software

TP_{10}: (Largely) automated processing and calculation of relevant features

TP_{11}: Providing data from CMA as biomechanical findings

Discussion

Principal Findings

We proposed and applied a novel HFE process model to inform the design of a physiotherapist dashboard with CDS for CMA of musicians with musculoskeletal problems, using methods, and addressing aspects from DSM, UCD, and DCD. Following this model, we first identified and thoroughly described the characteristics of the domain and problem area. We explored user needs, cognitively challenging tasks, information needs, and decision-making challenges faced by physiotherapists managing musicians with musculoskeletal conditions. We then proposed design seeds for CDS, translating user and decision requirements into concrete design requirements, including key features, technical prerequisites, and recommendations for decision support.

Key features include efficient integration of biomechanical data, combining heterogeneous data types, and providing adaptive overviews for quick data inspection. The tool should present patient data in a suitable form and facilitate interactive visual analyses using familiar techniques like filtering, highlighting, and annotating. Summarizing and exporting findings for communication are also essential. User and decision requirements led to specific decision-support recommendations: CDS should indicate if musician-specific diagnostics are possible, remind physiotherapists of physical exams, suggest efficient configuration parameters, preselect essential data, highlight significant changes, integrate normative values, support interactive annotations, compare baseline and final findings, and provide tools for reviewing patient goals and preparing reports. To enable the tool, technical prerequisites must be met, including access to patient data (anamnesis, physical exams, treatments, and problem- or musician-related questionnaires), biomechanical data from musician-specific CMAs, support for electronic diagnostic orders, and management of these orders and CMA statuses. The tool should use standardized protocols, integrate with existing tools, and automate data processing to present biomechanical findings.

Additionally, our study contributes to the existing literature by detailing a methodology for requirements elicitation, comprehensively describing the domain of musicians' physiotherapy, and identifying requirements for the successful use of the tool. Our findings support previous research indicating that HFE is crucial for designing CDS technologies [34,35], emphasizing aspects such as speed, efficiency, user needs anticipation, workflow integration, meaningful alerts, consistency, logical grouping, interface design, and design simplicity [37-40].

Limitations

This study is in the early stages of design, focusing on establishing a structured foundation for the tool's design and domain-specific clinical workflows, relying primarily on qualitative methods like workflow observations and focus group discussions. While the findings provided in-depth qualitative insights from a small group of domain experts, generalizability is limited, and broader validation is needed. To address this, we plan to conduct a Delphi study involving a larger and more diverse panel of physiotherapists, movement scientists, and other relevant stakeholders [64]. This structured, iterative process will facilitate expert consensus on the tool's design requirements, ensuring their robustness and applicability across various physiotherapy contexts. The Delphi study will also help refine the identified requirements based on collective expert feedback, enhancing the tool's validity before further prototyping, usability testing, and implementation in clinical practice.

At this stage, no functional prototype exists, and technical features or implementation details, such as specific visualizations, interaction methods, and decision logic, were not addressed. These aspects will be defined and iteratively

XSL•FO
RenderX

refined during future stages of prototyping and evaluation. Furthermore, quantitative evaluation of metrics such as diagnostic accuracy, efficiency, or patient outcomes is not yet feasible. Future work will involve the development of a prototype, followed by iterative usability testing and structured evaluations. These evaluations will incorporate quantitative measures—such as time-to-decision, task completion rates, diagnostic decision support effectiveness, efficiency gains, and user satisfaction scores—to assess the tool's impact and clinical utility in real-world physiotherapy settings.

In addition, artificial intelligence (AI)–based CDS features were overlooked. Integrating machine learning and predictive analytics could significantly enhance precision and relevance [65]. In future development stages, machine learning algorithms can be incorporated to support early diagnosis, pattern recognition, as well as personalized treatment planning and outcome prediction considering individual patient characteristics. For example, supervised learning models could be trained on annotated CMA data to automatically identify movement anomalies or classify injury risks [65]. Predictive analytics could also be used to forecast recovery progression based on patient-specific parameters such as movement patterns, therapy progress, and contextual factors (eg, instrument type or workload) [66]. These enhancements would augment the therapist's decision-making with advanced data-driven insights while maintaining transparency and clinician control. Overall, this highlights the tool's extensibility and long-term innovation potential.

Moreover, while the proposed CDS approach focuses on integrating CMA with gold-standard methods, AI-driven approaches like computer vision-based pose estimation (eg, OpenPose [67]) should be explored. Sensorless systems could provide nonintrusive, real-time movement tracking, making assessments more accessible, especially in nonlaboratory settings. AI pattern recognition could automate impairment classification and suggest diagnostic hypotheses, aiding physiotherapists and improving decision-making. It could also enable remote physiotherapy, allowing musicians to exercise at home with real-time feedback. However, challenges in accuracy, occlusion, and clinical validation must be addressed before full integration into physiotherapy workflows.

Opportunities and Challenges

The integration of CMA into physiotherapy and incorporating clinical and biomechanical data with CDS within an interactive dashboard present both opportunities and challenges. Our study emphasizes the need to tailor CDS to real clinical environments and therapists' workflows. Ensuring these technologies are perceived as useful is crucial, as the balance between perceived benefits and implementation effort significantly affects adoption. Developing an interactive tool for supporting diagnosis and treatment of musculoskeletal disorders in musicians requires a multidisciplinary approach, combining health informatics, biomechanics, and clinical practice. Collaboration with domain experts ensures the tool meets practical needs. Requirements were elicited through problem-driven research, ensuring that real-world problems were addressed, as well as UCD and DCD approaches, considering user needs, cognitive demands, and decision requirements, aiming to enhance the usability and usefulness of the solution.

Most existing tools in physiotherapy focus on general clinical documentation [68] or basic measurement instruments like goniometers and numeric rating scales [18], with limited quantitative and objective significance. Advanced biomechanical analysis systems (eg, Qualisys [69]) remain primarily research tools—technically complex and disconnected from everyday clinical decision-making, lacking user-friendliness, tailored visualizations, and integration with clinical reasoning. CDS tools in physiotherapy are still emerging, with AI-based approaches showing promise for personalizing treatment and predicting patient-specific outcomes [70-72]. However, they lack integration of CMA into clinical workflows and provide limited support for context-specific decision-making, particularly for niche populations like musicians with PRMDs. Existing visualization tools support rehabilitation or monitoring but do not incorporate CMA data [73,74], and those that do focus on clinical gait analysis or specific disorders [75,76]. Currently, no tool integrates CMA data into the decision-making process for physiotherapists managing musicians with PRMDs.

An interactive tool based on the identified design requirements would fill this gap by directly integrating CMA into clinical workflows, aligning with physiotherapists' reasoning. It would enable interactive visualizations tailored to the needs of diagnosing and treating musicians with PRMDs, offering flexibility for different instruments and playing styles. Unlike existing tools, it bridges the gap between biomechanical analysis and clinical practice, providing musician-focused support and a platform for integrating advanced analytics to enhance diagnostic precision and longitudinal tracking. Such a system aims to integrate CMA into real-world physiotherapy workflows, providing context-aware support and decision facilitation across all therapy stages, particularly for underserved populations like performing artists.

The successful deployment of a future CDS tool will require addressing key technical and ethical challenges. Integration into existing health care IT systems involves overcoming data heterogeneity, ensuring interoperability, safeguarding data privacy, and achieving clinician adoption. Many institutions rely on proprietary electronic health records [77], complicating seamless data exchange. Standardized protocols such as HL7 (Health Level Seven) and FHIR (Fast Healthcare Interoperability Resources) can help address interoperability [78], while compliance with regulations like GDPR (General Data Protection Regulation) will require robust encryption, anonymization when applicable, and role-based access controls. Diverse IT environments further necessitate system compatibility; a modular, cloud-based architecture could offer the flexibility and scalability needed for broad adoption [79]. Clinician adoption will depend on smooth integration into existing workflows, user-friendly interfaces, and comprehensive training [80].

Future work must standardize and integrate diverse data sources—such as motion capture outputs, clinical notes, and patient records—into a unified platform compatible with physiotherapy workflows. Ethical considerations are equally

XSL•FO
RenderX

critical. The envisioned tool is designed to assist, not replace, physiotherapists by providing objective biomechanical data to enhance—not automate—clinical decision-making. The system will maintain clinician autonomy, ensure transparency and interpretability, and be developed in adherence to data protection regulations. Although risks related to automation or bias are minimal at this stage, they will remain important considerations in later development phases, particularly as data handling, user control, and AI components evolve. Finally, the evolving nature of clinical practices and data types presents ongoing challenges for maintaining a CDS tool's relevance. Continuous updates are needed to keep pace with advancements in clinical knowledge and technology, which can be resource-intensive. Addressing these challenges in future research will enhance such a tool's usability, adaptability, and impact on clinical decision-making.

In conclusion, successful implementation of the physiotherapist dashboard with CDS based on musician-specific CMA depends on alignment with clinical workflows, ease of use, successful integration in existing IT infrastructure, maintaining data protection and clinician autonomy, and demonstrable benefits in diagnostic accuracy and treatment efficiency. Insights from this study provide a solid foundation for designing an interactive tool for physiotherapists, incorporating CMA for musicians with musculoskeletal demands, aiming to improve patient outcomes and support clinical decision-making. Iterative feedback from domain experts during design and development should ensure the tool addresses real-world problems, is validated through practical application, and is continuously refined based on user insights. The involvement of physiotherapists and movement scientists from initial design to validation is pivotal.

While this study focuses on the niche context of musicians with PRMDs, the identified design requirements, domain-specific decision-making insights, and integration of CMA hold potential for broader application. The underlying framework and methodology can be adapted to other physiotherapy contexts that involve complex, movement-related diagnostic and therapeutic challenges—such as in athletes, dancers, or individuals with work-related repetitive strain or posture-related injuries. These populations similarly require tailored functional assessments and context-specific CDS [81-84]. Modular, task-oriented design supports adaptation across these groups by enabling customization of movement analysis protocols, clinical decision parameters, and visualization formats to meet domain-specific needs. This enhances both the generalizability and scalability of the proposed framework. Future work may explore how the proposed process model and design requirements can be transferred and tailored to other physiotherapy contexts, extending the broader relevance and impact of this research beyond musicians.

Conclusion and Future Work

This work contributes to HFE, physiotherapy, CMA, and CDS for musicians with PRMDs by combining rigorous research methods with expert insights. The proposed process model helps stakeholders identify user needs and decision-making requirements at the point of care, while the design requirements lay the groundwork for developing a specialized interactive physiotherapist dashboard integrating CMA with CDS. This tool can enhance diagnostic accuracy and therapy effectiveness through interactive, visual data presentation, streamline workflows, and maintain clinical relevance. Ultimately, it aims to improve patient outcomes, reduce injury recurrence, and support musicians' long-term careers.

In the next phase, we plan to translate the design requirements into a functional prototype as part of a physiotherapy-integrated diagnostic system [85]. To ensure practicality and effectiveness, we will conduct rigorous usability testing and evaluations with physiotherapists. This will involve iterative feedback loops through think-aloud protocols and task-based evaluations, ensuring that the tool aligns with clinical workflows and cognitive demands. Structured feedback will be used to refine interface features, data presentation, and interaction flows, providing empirical evidence of feasibility, usability, and alignment with the identified requirements.

Future research should validate usability, decision support, and clinical effectiveness; explore advanced analytics; adapt to various CMA systems (eg, sensorless tracking); and expand applications across diverse musicians, instruments, and therapeutic domains to enhance its impact.

Acknowledgments

The authors would like to thank Florian Avermann, Camilla Kapitza, Heike Schemmann, Sabrina Friehe, Dirk Möller, Johanne Weber, and Lars Dorenbusch for contributing their time and expertise. This work is part of the RefLabPerform project and was funded by the Federal Ministry of Education and Research (BMBF; research grant 13FH113PX8).

Authors' Contributions

EW conceptualized the study, recruited participants, collected, and analyzed data, and wrote the manuscript. KM contributed to the study's conceptualization and reviewed the final manuscript, while SM reviewed and finalized it.

Conflicts of Interest

None declared.

Multimedia Appendix 1

XSL•FO
RenderX

User requirements (UR) formulated as user stories based on user needs for decision support.
[DOCX File, 90 KB - humanfactors_v12i1e65029_app1.docx]

Multimedia Appendix 2
Cognitive tasks and associated strategies, (potential) errors or difficulties, and critical hints.
[DOCX File, 94 KB - humanfactors_v12i1e65029_app2.docx]

Multimedia Appendix 3
Key decision requirements (DR), associated information needs and design seeds (S).
[DOCX File, 93 KB - humanfactors_v12i1e65029_app3.docx]

References

1. Zaza C, Farewell VT. Musicians' playing-related musculoskeletal disorders: an examination of risk factors. Am J Ind Med 1997 Sep;32(3):292-300. [Medline: 9219660]
2. Kok LM, Huisstede BMA, Voorn VMA, Schoones JW, Nelissen R. The occurrence of musculoskeletal complaints among professional musicians: a systematic review. Int Arch Occup Environ Health 2016 Apr;89(3):373-396. [Medline: 26563718]
3. Kok LM, Groenewegen KA, Huisstede BMA, Nelissen R, Rietveld ABM, Haitjema S. The high prevalence of playing-related musculoskeletal disorders (PRMDs) and its associated factors in amateur musicians playing in student orchestras: A cross-sectional study. PLoS ONE 2018;13(2):e0191772. [doi: 10.1371/journal.pone.0191772] [Medline: 29444093]
4. Fishbein M, Middlestadt SE, Ottati V, Straus S, Ellis A. Medical problems among ICSOM musicians:overview of a national survey. Med Probl Perform Art 1988;3(1):1.
5. Abréu-Ramos AM, Micheo WF. Lifetime prevalence of upper-body musculoskeletal problems in a professional-level symphony orchestra: age, gender, and instrument-specific results. Med Probl Perform Art 2007 Sep 1;22(3):97-104.
6. Möller D, Ballenberger N, Zalpour C. The German version of the musculoskeletal pain intensity and interference questionnaire for musicians (MPIIQM-G): translation and validation in professional orchestral musicians. Musculoskelet Sci Pract 2018 Oct;37:1-7. [Medline: 29879488]
7. Gembris H, Heye A. Growing older in a symphony orchestra: the development of the age-related self-concept and the self-estimated performance of professional musicians in a lifespan perspective. Music Sci 2014 Dec;18(4):371-391. [doi: 10.1177/1029864914548912]
8. Wolff D, Gutzwiller J. Musikerphysiotherapie – ein Fach mit Zukunft – Kein Himmel voller Geigen. physiopraxis 2010 Jun;8(06):32-35. [doi: 10.1055/s-0030-1262231]
9. Zalpour C. Musikergesundheit als aufgabe einer spezialisierten musikerphysiotherapie [Article in German]. : Hochschule Osnabrück; 2013.
10. Zalpour C, Ballenberger N, Avermann F. A physiotherapeutic approach to musicians' health - data from 614 patients from a physiotherapy clinic for musicians (INAP/O). Front Psychol 2021;12:568684. [doi: 10.3389/fpsyg.2021.568684] [Medline: 34177677]
11. Niemier K, Seidel W, editors. Funktionelle Schmerztherapie Des Bewegungssystems, 2nd edition: Springer-Verlag; 2012. [doi: 10.1007/978-3-642-20576-7]
12. Levett-Jones T, Hoffman K, Dempsey J, et al. The "five rights" of clinical reasoning: an educational model to enhance nursing students' ability to identify and manage clinically "at risk" patients. Nurse Educ Today 2010 Aug;30(6):515-520. [doi: 10.1016/j.nedt.2009.10.020] [Medline: 19948370]
13. Müller-Staub M. Klinische Entscheidungsfindung und kritisches Denken im pflegediagnostischen Prozess. Pflege 2006 Oct 1;19(5):275-279. [doi: 10.1024/1012-5302.19.5.275]
14. Bartrow K. Untersuchen Und Befunden in Der Physiotherapie: Untersuchungstechniken Und Diagnoseinstrumente, 3rd edition: Springer-Verlag; 2019. [doi: 10.1007/978-3-662-58298-5]
15. Greiff C. Befund. In: Richter K, Greiff C, Weidemann-Wendt N, editors. Der Ältere Mensch in Der Physiotherapie: Springer; 2017:79-102. [doi: 10.1007/978-3-662-50466-6_2]
16. Philp F, Faux-Nightingale A, Woolley S, de Quincey E, Pandyan A. Implications for the design of a Diagnostic Decision Support System (DDSS) to reduce time and cost to diagnosis in paediatric shoulder instability. BMC Med Inform Decis Mak 2021 Feb 27;21(1):78. [doi: 10.1186/s12911-021-01446-5] [Medline: 33639920]
17. Möller D, Backes K. Von der Biomechanik zur klinischen Anwendung. manuelletherapie 2017;21(02):88-95. [doi: 10.1055/s-0043-105164]
18. Braun T, Rieckmann A, Weber F, Grüneberg C. Current use of measurement instruments by physiotherapists working in Germany: a cross-sectional online survey. BMC Health Serv Res 2018 Oct 23;18(1):810. [doi: 10.1186/s12913-018-3563-2] [Medline: 30352584]
19. Möller D, Griefahn A. Sensorgesteuerte Bewegungsanalyse bei Musikerinnen und Musikern. MSK – Muskuloskelettale Physiotherapie 2022 Feb;26(1):21-26. [doi: 10.1055/a-1685-6633]

XSL•FO
RenderX

20. Robertson DGE, Caldwell GE, Hamill J, Kamen G, Whittlesey SN. Research Methods in Biomechanics: Human Kinetics; 2014. URL: https://www.humankineticslibrary.com/encyclopedia?docid=b-9781492595809 [doi: 10.5040/9781492595809]
21. Richards J. The Comprehensive Textbook of Clinical Biomechanics, 2nd edition: Elsevier; 2018.
22. Kelleher LK, Campbell KR, Dickey JP. Biomechanical research on bowed string musicians: a scoping study. Med Probl Perform Art 2013 Dec;28(4):212-218. [Medline: 24337033]
23. Wolf E, Möller D, Ballenberger N, Morisse K, Zalpour C. Marker-Based Method for Analyzing the Three-Dimensional Upper Body Kinematics of Violinists and Violists: Development and Clinical Feasibility. Med Probl Perform Art 2019 Dec;34(4):179-190. [doi: 10.21091/mppa.2019.4029] [Medline: 31800669]
24. Wolf E, Möller D, Ballenberger N, Morisse K, Zalpour C. Marker-based method for analyzing the three-dimensional upper body kinematics of violinists: reproducibility. Med Probl Perform Art 2022 Sep;37(3):176-191. [doi: 10.21091/mppa.2022.3025] [Medline: 36053495]
25. Amelung P, Seidel W. Die instrumentierte Ganganalyse im Rahmen eines manualmedizinisch orientierten Diagnostik- und Therapiesettings im Krankenhaus bei Patienten mit Schmerzerkrankungen des Bewegungssystems. Manuelle Medizin 2005 Dec;43(6):404-413. [doi: 10.1007/s00337-005-0396-y]
26. Passier LN, Nasciemento MP, Gesch JM, Haines TP. Physiotherapist observation of head and neck alignment. Physiother Theory Pract 2010 Aug;26(6):416-423. [doi: 10.3109/09593980903317557] [Medline: 20658928]
27. Konrad P. EMG-Fibel – Eine praxisorientierte Einführung in die kinesiologische Elektromyographie: Noraxon; 2011.
28. RefLabPerform. Hochschule Osnabrück. URL: https://www.hs-osnabrueck.de/en/reflabperform/ [accessed 2024-07-24]
29. Jöllenbeck T. Bewegungsanalyse -wesentliches Element moderner sportmedizinischer Diagnostik. Dtsch Z Für Sportmed 2012(63):59-60.
30. Simon SR. Quantification of human motion: gait analysis-benefits and limitations to its application to clinical problems. J Biomech 2004 Dec;37(12):1869-1880. [doi: 10.1016/j.jbiomech.2004.02.047] [Medline: 15519595]
31. Loudon D, Carse B, Macdonald AS. Investigating the use of visualisations of biomechanics in physical rehabilitation. In: Cruz-Cunha MM, Varajão J, Powell P, Martinho R, editors. Enterprise Information Systems Communications in Computer and Information Science: Springer; 2011:30-39. [doi: 10.1007/978-3-642-24352-3_4]
32. Steinwendner J. Klinische entscheidungsunterstützungssysteme: von der datenrepräsentation zur künstlichen intelligenz. In: Pfannstiel MA, Kassel K, Rasche C, editors. Innovationen Und Innovationsmanagement Im Gesundheitswesen: Technologien, Produkte Und Dienstleistungen Voranbringen: Springer Fachmedien; 2020:683-699. [doi: 10.1007/978-3-658-28643-9_36]
33. Davis F, Bagozzi R, Warshaw P. User acceptance of computer technology: a comparison of two theoretical models. Manag Sci 1989;35:982-1003. [doi: 10.1287/mnsc.35.8.982]
34. Karsh BT, Holden RJ, Alper SJ, Or CKL. A human factors engineering paradigm for patient safety: designing to support the performance of the healthcare professional. BMJ Qual Saf 2006;15(suppl 1):i59-i65. [doi: 10.1136/qshc.2005.015974]
35. Carayon P, Wood KE. Patient safety - the role of human factors and systems engineering. In: Engineering the System of Healthcare Delivery: IOS Press; 2010:23-46. [doi: 10.3233/978-1-60750-533-4-23]
36. Miller A, Koola JD, Matheny ME, et al. Application of contextual design methods to inform targeted clinical decision support interventions in sub-specialty care environments. Int J Med Inform 2018 Sep;117:55-65. [doi: 10.1016/j.ijmedinf.2018.05.005]
37. Bates DW, Kuperman GJ, Wang S, et al. Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. J Am Med Inform Assoc 2003 Nov;10(6):523-530. [doi: 10.1197/jamia.M1370]
38. Rice H, Garabedian PM, Shear K, et al. Clinical decision support for fall prevention: defining end-user needs. Appl Clin Inform 2022 May;13(3):647-655. [doi: 10.1055/s-0042-1750360] [Medline: 35768011]
39. Kuperman GJ, Bobb A, Payne TH, et al. Medication-related clinical decision support in computerized provider order entry systems: a review. J Am Med Inform Assoc 2007 Jan 1;14(1):29-40. [doi: 10.1197/jamia.M2170]
40. Horsky J, Schiff GD, Johnston D, Mercincavage L, Bell D, Middleton B. Interface design principles for usable decision support: a targeted review of best practices for clinical prescribing interventions. J Biomed Inform 2012 Dec;45(6):1202-1216. [doi: 10.1016/j.jbi.2012.09.002] [Medline: 22995208]
41. Sedlmair M, Meyer M, Munzner T. Design study methodology: reflections from the trenches and the stacks. IEEE Trans Vis Comput Graph 2012 Dec;18(12):2431-2440. [doi: 10.1109/TVCG.2012.213] [Medline: 26357151]
42. Gulliksen J, Göransson B, Boivie I, Blomkvist S, Persson J, Cajander Å. Key principles for user-centred systems design. Behaviour & Information Technology 2003 Nov;22(6):397-409. [doi: 10.1080/01449290310001624329]
43. Crandall B, Klein G, Klein GA, Hoffman RR, Hoffman RR, Hoffman SRSRR. Working Minds: A Practitioner's Guide to Cognitive Task Analysis: MIT Press; 2006.
44. Militello LG, Klein G. Decision-centered design. In: Lee JD, Kirlik A, editors. The Oxford Handbook of Cognitive Engineering: Oxford University Press; 2013:261-271. [doi: 10.1093/oxfordhb/9780199757183.013.0016]
45. Clark R, Feldon D, Merrienboer JJG, Yates K, Early S. Cognitive task analysis. Handb Res Educ Commun Technol 2008 Jan 1:577-593.
46. Wolf E, Morisse K, Meister S. Proposing an integrated process model for designing clinical decision support tools. Stud Health Technol Inform 2025 Apr 8;323:337-338. [doi: 10.3233/SHTI250107] [Medline: 40200504]

XSL•FO
RenderX

47. Schemmann H, Rensing N, Zalpour C. Musculoskeletal assessments used in quantitatively based studies about posture and movement in high string players: a systematic review. Med Probl Perform Art 2018 Mar;33(1):56-71. [doi: 10.21091/mppa.2018.1009] [Medline: 29600309]
48. Musiker SC. Behandlung und prävention in der musikermedizin. In: Badura B, Ducki A, Schröder H, Klose J, Meyer M, editors. Fehlzeiten-Report 2015: Neue Wege Für Mehr Gesundheit - Qualitätsstandards Für Ein Zielgruppenspezifisches Gesundheitsmanagement: Fehlzeiten-Report. Springer; 2015:227-231. [doi: 10.1007/978-3-662-47264-4_20]
49. Spahn C, Richter B, Altenmüller E. MusikerMedizin: Diagnostik, Therapie Und Prävention von Musikerspezifischen Erkrankungen, 1st edition: Schattauer; 2010.
50. Shan G, Visentin P, Nusseck M, Spahn C. Investigating aspects of movement in violin performance. In: Handbook of Human Motion: Springer International Publishing; 2018. [doi: 10.1007/978-3-319-14418-4_108]
51. Lohse-Busch H. Funktionelle diagnostik des bewegungssystems. In: Ebinger F, editor. Schmerzen Bei Kindern Und Jugendlichen: Ursachen, Diagnostik Und Therapie, 2011th edition: Thieme Verlag; 2011. [doi: 10.1055/b-0034-38295]
52. Higgs J, Jones M, Loftus S, Christensen N. Clinical Reasoning in the Health Professions: Elsevier; 2008.
53. Jones MA, Jensen G, Edwards I. Clinical reasoning in physiotherapy. In: Higgs J, Jones MA, Loftus S, Christensen N, editors. Clinical Reasoning in the Health Professions: Elsevier Health Sciences; 2008:245-254.
54. Trocha M, Aigner AK, Brandt H, et al. Befundest du noch – oder diagnostizierst du schon? physioscience 2014;10(1):24-28. [doi: 10.1055/s-0034-1365977]
55. Konrad P. The ABC of EMG. A Practical Introduction to Kinesiological Electromyography: Noraxon INC; 2005.
56. Best R, Begg R. Overview of movement analysis and gait features. Comput Intell Mov Sci Neural Netw Emerg Tech 2006 Jan 1. [doi: 10.4018/978-1-59140-836-9.ch001]
57. Coutts F. Gait analysis in the therapeutic environment. Man Ther 1999 Feb;4(1):2-10. [doi: 10.1016/s1356-689x(99)80003-4] [Medline: 10463015]
58. Flörsheim-Dalsheim PKK. Risiko-Screening in Den Ausbildungsstrukturen Der Physiotherapie – Curriculare Umsetzung in Den Ausbildungseinrichtungen Und Dessen Einfluss Auf Die Fähigkeit Der Auszubildenden Und Studierenden: Universität Ulm; 2018.
59. Holtzblatt K, Wendell JB, Wood S. Rapid Contextual Design: A How-to Guide to Key Techniques for User-Centered Design: Elsevier; 2005. [doi: 10.1145/1066348.1066325]
60. Daniels N, Gillen P, Casson K, Wilson I. STEER: Factors to Consider When Designing Online Focus Groups Using Audiovisual Technology in Health Research. Int J Qual Methods 2019 Jan 1;18:1609406919885786. [doi: 10.1177/1609406919885786]
61. Griefahn A, Wolf E, Zalpour C. Wie kann die Dokumentation in der Physiotherapie durch die Digitalisierung effektiver und effizienter werden? – Eine Delphi-Studie. physioscience 2020 Sep;16(03):101-110. [doi: 10.1055/a-1113-6688]
62. Kontaxis A, Cutti AG, Johnson GR, Veeger HEJ. A framework for the definition of standardized protocols for measuring upper-extremity kinematics. Clin Biomech (Bristol, Avon) 2009 Mar;24(3):246-253. [doi: 10.1016/j.clinbiomech.2008.12.009]
63. Sharp H, Preece J, Rogers Y. Interaction Design: Beyond Human-Computer Interaction, 5th edition: Wiley; 2019.
64. Jünger S, Payne SA, Brine J, Radbruch L, Brearley SG. Guidance on Conducting and REporting DElphi Studies (CREDES) in palliative care: recommendations based on a methodological systematic review. Palliat Med 2017 Sep;31(8):684-706. [doi: 10.1177/0269216317690685] [Medline: 28190381]
65. Tack C. Artificial intelligence and machine learning | applications in musculoskeletal physiotherapy. Musculoskelet Sci Pract 2019 Feb;39:164-169. [doi: 10.1016/j.msksp.2018.11.012] [Medline: 30502096]
66. Remedios SM, Armstrong DP, Graham RB, Fischer SL. Exploring the application of pattern recognition and machine learning for identifying movement phenotypes during deep squat and hurdle step movements. Front Bioeng Biotechnol 2020;8:364. [doi: 10.3389/fbioe.2020.00364] [Medline: 32426346]
67. Cao Z, Hidalgo G, Simon T, Wei SE, Sheikh Y. OpenPose: realtime multi-person 2D pose estimation using part affinity fields. IEEE Trans Pattern Anal Mach Intell 2021 Jan;43(1):172-186. [doi: 10.1109/TPAMI.2019.2929257] [Medline: 31331883]
68. Weber F, Rieckmann A, Grueneberg C. Feasibility report of the usage of the mobile-app TheraAssist to improve documentation and usage of assessments in physiotherapy practice. Physiotherapy 2016 Nov;102:e274. [doi: 10.1016/j.physio.2016.10.347]
69. Gait analysis. Qualisys. URL: https://www.qualisys.com/analysis/gait/ [accessed 2025-04-05]
70. Granviken F, Meisingset I, Vasseljen O, Bach K, Bones AF, Klevanger NE. Acceptance and use of a clinical decision support system in musculoskeletal pain disorders - the SupportPrim project. BMC Med Inform Decis Mak 2023 Dec 19;23(1):293. [doi: 10.1186/s12911-023-02399-7] [Medline: 38114970]
71. Mikołajewska E, Masiak J, Mikołajewski D. Applications of artificial intelligence-based patient digital twins in decision support in rehabilitation and physical therapy. Electronics (Basel) 2024;13(24):4994. [doi: 10.3390/electronics13244994]
72. Hao J, Yao Z, Tang Y, Remis A, Wu K, Yu X. Artificial intelligence in physical therapy: evaluating ChatGPT's role in clinical decision support for musculoskeletal care. Ann Biomed Eng 2025 Jan;53(1):9-13. [doi: 10.1007/s10439-025-03676-4] [Medline: 39760952]

XSL•FO
RenderX

73. Ploderer B, Fong J, Klaic M, et al. How therapists use visualizations of upper limb movement information from stroke patients: a qualitative study with simulated information. JMIR Rehabil Assist Technol 2016 Oct 5;3(2):e9. [doi: 10.2196/rehab.6182] [Medline: 28582257]
74. Mansourvar M. Data vizualization dashboards in robotic rehabilitation. In: Public Health and Informatics: IOS Press; 2021:278-282. [doi: 10.3233/SHTI210164]
75. Wagner M, Slijepcevic D, Horsak B, Rind A, Zeppelzauer M, Aigner W. KAVAGait: knowledge-assisted visual analytics for clinical gait analysis. IEEE Trans Vis Comput Graph 2019 Mar;25(3):1528-1542. [doi: 10.1109/TVCG.2017.2785271] [Medline: 29993807]
76. Chan GYY, Nonato LG, Chu A, Raghavan P, Aluru V, Silva CT. Motion browser: visualizing and understanding complex upper limb movement under obstetrical brachial plexus injuries. IEEE Trans Visual Comput Graphics 2020;26(1):981-990. [doi: 10.1109/TVCG.2019.2934280]
77. Evans RS. Electronic health records: then, now, and in the future. Yearb Med Inform 2016 May 20;Suppl 1(Suppl 1):S48-S61. [doi: 10.15265/IYS-2016-s006] [Medline: 27199197]
78. Ayaz M, Pasha MF, Alzahrani MY, Budiarto R, Stiawan D. The Fast Health Interoperability Resources (FHIR) standard: systematic literature review of implementations, applications, challenges and opportunities. JMIR Med Inform 2021 Jul 30;9(7):e21929. [doi: 10.2196/21929] [Medline: 34328424]
79. Bathelt F, Lorenz S, Weidner J, Sedlmayr M, Reinecke I. Application of Modular Architectures in the Medical Domain - a Scoping Review. J Med Syst 2025 Feb 18;49(1):27. [doi: 10.1007/s10916-025-02158-3] [Medline: 39964566]
80. Weik L, Fehring L, Mortsiefer A, Meister S. Understanding inherent influencing factors to digital health adoption in general practices through a mixed-methods analysis. NPJ Digit Med 2024 Feb 27;7(1):47. [doi: 10.1038/s41746-024-01049-0] [Medline: 38413767]
81. Fauntroy V, Fyock M, Hansen-Honeycutt J, Nolton E, Ambegaonkar JP. Using the selective functional movement assessment for the evaluation of dancers' functional limitations and dysfunctions: A critically appraised topic. J Sport Rehabil 2019 Nov 1;28(8):891-896. [doi: 10.1123/jsr.2018-0054] [Medline: 30222496]
82. Bullock GS, Hughes T, Sergeant JC, Callaghan MJ, Riley RD, Collins GS. Clinical prediction models in sports medicine: a guide for clinicians and researchers. J Orthop Sports Phys Ther 2021 Oct;51(10):517-525. [doi: 10.2519/jospt.2021.10697] [Medline: 34592832]
83. Gross DP, Armijo-Olivo S, Shaw WS, et al. Clinical decision support tools for selecting interventions for patients with disabling musculoskeletal disorders: a scoping review. J Occup Rehabil 2016 Sep;26(3):286-318. [doi: 10.1007/s10926-015-9614-1] [Medline: 26667939]
84. Soares CO, Pereira BF, Pereira Gomes MV, Marcondes LP, de Campos Gomes F, de Melo-Neto JS. Preventive factors against work-related musculoskeletal disorders: narrative review. Rev Bras Med Trab 2019;17(3):415-430. [doi: 10.5327/Z1679443520190360] [Medline: 32368676]
85. Wolf E, Morisse K, Meister S. Designing a physiotherapy integrated diagnostic system incorporating clinical motion analysis of musicians. Presented at: 22nd International Conference on Informatics, Management and Technology in Healthcare (ICIMTH 2025); Jul 4-6, 2025; Athens, Greece. [doi: 10.3233/SHTI250106]

Abbreviations

AI: artificial intelligence
CDS: clinical decision support
CMA: clinical movement analysis
DCD: decision-centered design
DSM: design study methodology
FHIR: Fast Healthcare Interoperability Resources
GDPR: General Data Protection Regulation
HFE: human factors engineering
HL7: Health Level Seven
INAPO: Institute for Applied Physiotherapy Osnabrück
PRMD: performance-related musculoskeletal disorder
USD: user-centered design

XSL•FO
RenderX

Edited by A Kushniruk; submitted 02.08.24; peer-reviewed by J Steinwendner, SS Goswami; revised version received 11.04.25; accepted 17.04.25; published 16.07.25.

Please cite as:
Wolf E, Morisse K, Meister S
Identifying Design Requirements for an Interactive Physiotherapy Dashboard With Decision Support for Clinical Movement Analysis of Musicians With Musculoskeletal Problems: Qualitative User Research Study
JMIR Hum Factors 2025;12:e65029
URL: https://humanfactors.jmir.org/2025/1/e65029
doi:10.2196/65029

© Eduard Wolf, Karsten Morisse, Sven Meister. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 16.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Capturing Requirements for a Data Annotation Tool for Intensive Care: Experimental User-Centered Design Study

Marceli Wac[1,2], BSc; Raul Santos-Rodriguez[1], MSc, PhD; Chris McWilliams[1,2], BSc, PhD; Christopher Bourdeaux[2], BCHIR, MB, MA

[1]Faculty of Engineering, University of Bristol, Bristol, United Kingdom

[2]University Hospitals Bristol and Weston NHS Foundation Trust, Bristol, United Kingdom

Corresponding Author:
Marceli Wac, BSc
Faculty of Engineering
University of Bristol
Queen's Building
University Walk
Bristol, BS8 1TR
United Kingdom
Phone: 44 1173315830
Email: m.wac@bristol.ac.uk

Abstract

Background: Increasing use of computational methods in health care provides opportunities to address previously unsolvable problems. Machine learning techniques applied to routinely collected data can enhance clinical tools and improve patient outcomes, but their effective deployment comes with significant challenges. While some tasks can be addressed by training machine learning models directly on the collected data, more complex problems require additional input in the form of data annotations. Data annotation is a complex and time-consuming problem that requires domain expertise and frequently, technical proficiency. With clinicians' time being an extremely limited resource, existing tools fail to provide an effective workflow for deployment in health care.

Objective: This paper investigates the approach of intensive care unit staff to the task of data annotation. Specifically, it aims to (1) understand how clinicians approach data annotation and (2) capture the requirements for a digital annotation tool for the health care setting.

Methods: We conducted an experimental activity involving annotation of the printed excerpts of real time-series admission data with 7 intensive care unit clinicians. Each participant annotated an identical set of admissions with the periods of weaning from mechanical ventilation during a single 45-minute workshop. Participants were observed during task completion and their actions were analyzed within Norman's Interaction Cycle model to identify the software requirements.

Results: Clinicians followed a cyclic process of investigation, annotation, data reevaluation, and label refinement. Variety of techniques were used to investigate data and create annotations. We identified 11 requirements for the digital tool across 4 domains: annotation of individual admissions (n=5), semiautomated annotation (n=3), operational constraints (n=2), and use of labels in machine learning (n=1).

Conclusions: Effective data annotation in a clinical setting relies on flexibility in analysis and label creation and workflow continuity across multiple admissions. There is a need to ensure a seamless transition between data investigation, annotation, and refinement of the labels.

(JMIR Hum Factors 2025;12:e56880) doi:10.2196/56880

KEYWORDS

ICU; intensive care; machine learning; data annotation; data labeling; annotation software; capturing software requirements

XSL•FO
RenderX

Introduction

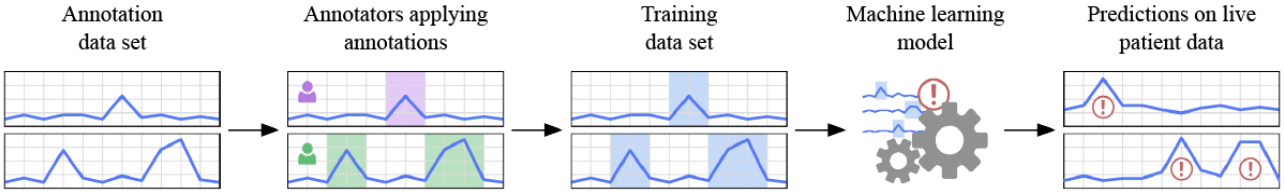
Background

Artificial intelligence (AI) is a field concerned with leveraging computers to mimic human cognitive functions, such as problem-solving and decision-making [1,2]. Depending on the task, this process can use a variety of methods and take many forms ranging from relatively simple rule-based expert systems (following if-then pathways) through regression (modeling the relation between different variables to predict their values) and clustering (grouping objects with similar properties together) to more complex systems such as artificial neural networks (computing systems imitating the anatomy of human brains capable of modeling nonlinear processes) [1,3,4]. Particularly complex problems may require solutions that cannot be achieved by traditional approaches, such as preprograming the desired behavior. Instead, techniques such as machine learning (ML) form a subset of AI in which the algorithms are used to analyze large amounts of data to derive a method for computing a solution [5]. In ML, the practice of arriving at a solution is called "learning" (or training), and the produced output is known as the ML model. The heavy reliance on data in ML highlights the importance of ensuring the appropriate quantity and quality of data and signifies its impact on the effectiveness of the created model [6]. The continuously increasing popularity of ML and its rapid adoption rate in life sciences suggest that AI is at the forefront of bringing innovation to health care [7,8].

Intensive care units (ICUs) are busy and complex health care environments where critically ill patients frequently require continuous monitoring and multiple-organ support. To provide care for those patients, clinicians working in the ICUs use a broad range of medical devices, such as ventilators, monitoring devices, and intravenous pumps and lines among many others. The information captured by those devices, as well as that entered by the ICU staff, is collected and collated in a clinical information system, which enables health practitioners to access large quantities of data routinely required as part of their job. This data-rich nature and the direct influence of data on the provision of care provide a tremendous opportunity for the deployment of ML models in health care and in particular ICUs [9].

While the data gathered in the ICUs can often be used directly, for example, to present the correlation between different vital signs and the patient prognosis, complex tasks that rely heavily on clinical experience and expertise may require human involvement to provide further guidance [10]. This guidance, most frequently referred to as labels or annotations [11], can deliver additional information or context to the existing data. An example of such a label could be a "yes" or "no" indicator of whether a patient is ready for discharge, or the type and size of tumor present on a radiograph. With this further knowledge, ML models can take advantage of the human experience to tackle complex problems and deliver solutions that could not be inferred from raw data alone [12] (Figure 1).

Figure 1. Annotating data can provide additional information and context necessary to train machine learning models.



While the large volumes of data available in intensive care offer significant opportunities, working with these data also presents a unique set of challenges, as the effort required to annotate the dataset increases proportionally to its size [13]. Furthermore, the complex nature of the health care data and the fact that applying a single label can require clinicians to look at multiple parameters, patient history, and laboratory results make the annotation task highly labor-intensive. With the clinician's time being a remarkably valuable resource, ensuring the effectiveness of the data annotation workflow is paramount. The lack of an annotation system tailored to the unique nature of the health care data (eg, accessing a subset of relevant variables from a list of potentially hundreds of parameters [14]) further complicates this problem.

In addition to the difficulties associated with working within clinical settings, it is equally important to consider the inherent challenges of data annotation in their own right. These include the need to account for multiple different types of bias that could be introduced throughout the process [15]. For example, the annotators themselves may have their own preferences and preconceptions that will guide the approach they assume when annotating data. This could stem from the amount and diversity of their clinical experience in a context specific to the annotation task, such as the familiarity with the clinical problem, the specific population within which they have treated it, or even the specific tools and methods they have used in the past. Similarly, by annotating historical data that span the entire duration of the admissions, annotators may create labels with an inherent temporal or selection bias. Alternatively, annotators tasked with creating specific annotations may focus on a subset of the admission where they would expect to create the annotation, such as reliance on medication, which may decrease toward the end of the admission but could equally change throughout its overall course.

Current literature on the applications of computational methods within health care highlights data annotation as a primary bottleneck in the ML pipeline [16]. This effect is attributed to the need for domain-specific knowledge and therefore access the expert population, as well as the considerable investment of their time [16-18]. Because of this, existing methods aim to reduce the number of required annotations required for positive results by using a variety of techniques including active learning [17], machine-assisted annotation [19], or synthetic data generation [20]. Nevertheless, the majority of studies continue

XSL•FO
RenderX

to rely on human annotations in the cases where labels cannot be easily derived, or are inadequately documented in structured data [16]. Designing a tool for the annotation of large clinical datasets is therefore a problem that needs to be approached carefully. The expert nature of the annotators means that both their numbers and time are limited and are therefore resources that need to be used efficiently [16]. Furthermore, clinical datasets frequently aggregate data spanning several years, resulting in a volume of data that is infeasible for manual annotation, suggesting a need for a semiautomated approach that could scale up to an entire dataset with limited input from the domain experts.

Involving end users in the design process of new tools prior to their development is an important aspect of designing effective software [21]. It minimizes the risk of creating a system that is inefficient and helps ensure that the developed solution meets users' expectations and requirements [22,23]. In the context of the data annotation software used by experts in intensive care, this participatory design is especially critical, as each of the variety of roles (eg, junior doctors, doctors, nurses, and consultants) can generate a unique set of requirements. Furthermore, while the staff working in the ICUs are experts in the medical domain trained to treat patients, extracting their knowledge through data annotation is an entirely different process. For this reason, it is crucial to deploy a strategy that will ensure that the design of the tool follows a structured and systematic approach that can capture a wide variety of perspectives from its end users while also accounting for the needs and priorities of the data science experts who will use the created annotations.

Objectives

The primary objective of the research is to establish a set of criteria for the design of a data annotation tool that can be used effectively by clinicians to annotate time-series datasets from intensive care. To achieve this goal under the limitations of working within the health care setting, the system needs to account for limited access to the annotators and large volumes of data. Finally, to ensure the efficiency of the process, the subtleties of how clinicians approach the problem of data annotation need to be understood and accounted for in the design. The overarching goal is, therefore, to gather requirements for a data annotation platform that is purpose-built for the intensive care data and clinicians and, as such, one that facilitates an efficient annotation workflow in that domain.

Methods

Study Design

To understand how clinicians approach and reason about the data annotation process, we conducted an experimental study that involved members of the clinical staff from the ICU manually annotating excerpts of time-series data printed on paper. Observations taken during task completion served as a basis for analysis and were used to derive the requirements for the data annotation software. This choice of methodology was influenced by constraints associated with working within the clinical setting, such as the high demand and short supply of participants' time, and the direct relationship between the task completion and the requirements for the digital tool, which allowed for a natural and unobstructed data collection.

Participant Recruitment

Our participant recruitment followed a mixture of convenience and snowball sampling comprising an invitation email sent to the staff working in the ICU who had prior experience with mechanical ventilation treatment and signposting through the internal network at our research site—University Hospitals Bristol and Weston NHS Foundation Trust. This recruitment process yielded a sample size of 7 participants across 2 distinct job roles (6 junior doctors and 1 doctor), which provided limited feasibility for suitable analysis of the interrole differences in approach to annotation but was sufficiently rich for establishing the software requirements for the tool we set out to design. Comparative sample sizes can be observed in similar research in the field [24-26]. Before data collection, the annotation task was piloted with a clinical member of our team experienced in mechanical ventilation treatment, as well as 2 other members of our team from the ML background. The insights collected from the pilot activity allowed us to refine the set of clinical parameters used throughout the activity, devise the inclusion criteria for the underlying admissions, and diversify the data extracts to form a sample representative of a typical dataset. Our research site had a preestablished research partnership with the University of Bristol and a technologically enhanced ICU with widespread adoption of digital systems.

Theoretical Framework

To analyze the participants' actions performed during the simulated activity, we used a framework that served as a basis for characterizing the data annotation process and identifying the challenges and opportunities associated with performing the task. Our framework followed the Norman's Interaction Cycle (NIC) model which assumes that the interaction is a process of evaluation and execution between the user and the technology [27]. It outlines a 7-step process that begins at the goal and, through evaluation of available means and strategies, as well as the execution of these strategies, leads to achieving that goal in practice. The framework emphasizes the breakdown of the preproduction stage into planning (formulating the problem that needs to be solved), specification (outlining the strategies that can be used to tackle the problem), and performance (specifying actions that need to be taken to deploy the strategy). Furthermore, it highlights the importance of presenting the user with feedback over 3 separate stages including perception (tracking the outcome of performed actions), interpretation (analyzing the effect of the outcome), and comparison (evaluating whether their actions resulted in achieving the goal) [27]. Finally, the model captures the disparities between user intentions and actions permitted by the technology (gulf of execution) and the effort required to correctly interpret the results of their actions with regard to the desired outcome (gulf of evaluation) [27,28].

Data Collection

To facilitate data collection, we simulated a data annotation activity involving manual annotation of real intensive care patient data using pens and highlighter pens on excerpts of data

XSL•FO
RenderX

printed on paper sheets. The data were formatted to resemble the clinical information system present at the research site and displayed patient demographics in addition to a table with time-series data. Each column of the table corresponded to an hour of the day and each row corresponded to a specific parameter; the table's cells contained the readings of a parameter for a given hour (Figure 2). The outline brief for the activity prompted participants to create annotations, allowing them to create multiple annotations for a single admission. Crucially, no additional instructions were provided regarding the annotation format. Before the data collection, the activity was piloted internally with an experienced intensive care consultant (CB) and 2 ML experts (RSR and CMW), which helped us establish inclusion criteria for the selected admissions, broaden the characteristics of our admission sample, and refine the included parameters. The activity took place as a single in-person workshop, lasting approximately 45 minutes, during which each of the 7 participants annotated the same set of 5 unique admissions (Figure 3).

Participants were observed for the duration of task completion by a single observer who took field notes throughout the entire activity. These field notes captured the interactions of the participants with the annotation tasks and described the actions taken when annotating specific admissions, how multiple admissions were annotated, and questions asked by the participants during the task. The collected notes were then aggregated across the different stages of NIC and the annotated printouts were collected and used to supplement the analysis of the methods used by the participants to label the data.

Figure 2. The provided excerpts of time series data were formatted to resemble the interfaces of the clinical information system present at the research site (the depicted table contains data for illustrative purposes only and is trimmed for conciseness). FiO2: fraction of inspired oxygen; ICU: intensive care unit; PCWP: pulmonary capillary wedge pressure; PEEP: positive end-expiratory pressure.

Stay ID: 12341234 **Age:** 54 **Race:** WHITE **Admission time:** 14/8/2174 12:00 **Length of stay in hospital**

Visit in ICU during this admission: #1 **Gender:** F **Death time:** - **Discharge time:** 21/10/2174 14:00 **Length of stay in ICU:** 16

	8/10 07:00	08:00	09:00	10:00	11:00	12:00	13:
Vitals							
Heart Rate [bpm]	63	66	57	59	54	52	57
Arterial Blood Pressure Systolic [mmHg]	132	141	152	143	119	141	155
Arterial Blood Pressure Diastolic [mmHg]	65	67	73	73	69	72	76
Arterial Blood Pressure Mean [mmHg]	86	85	94	99	92	97	104
Non-invasive Blood Pressure Systolic [mmHg]			131	139	117	127	130
Non-invasive Blood Pressure Diastolic [mmHg]			75	81	70	75	72
Non-invasive Blood Pressure Mean [mmHg]			85	95	81	86	89
Cardiac Output (thermodilution) [L/min]							
Respiratory Rate [insp/min]	14	12	13	13	14	12	14
Respiratory							
PEEP Set [cm H2O]	5			5			
Temperature [C]	36.32						
PCWP [mmHg]							
FiO2 [%]	50			50			40

Figure 3. Participants annotated data in a variety of ways using pen and paper during the workshop held in the intensive care unit.



XSL•FO
RenderX

Task Selection

Capturing data labels in a manner representative of the general annotation process relied on selecting an appropriate task. On the one hand, the desired label could not be trivial and needed to pose a significant challenge, complex enough to prompt an in-depth analysis of the data. On the other hand, this task also needed to fall within the subset of domain knowledge that participants were familiar with and could therefore solve with a relative degree of confidence. Annotating weaning from mechanical ventilation satisfies both of these criteria due to its challenging nature and the relative bias of the person who performs it. The process of weaning can be characterized by great variability in practice [29], in which both the timing of when the weaning begins and the method in which it is delivered largely depend on the clinician in charge of the treatment [29,30]. Together with the lack of personalized guidelines [31] and a wide variety of ways in which patients can be weaned, mechanical ventilation weaning constitutes a label that is difficult to derive from data and requires domain expertise, which makes it suitable for this task. While in the context of mechanical ventilation the term "weaning" comes with an inherent ambiguity due to a number of conflicting definitions [32-36], for the purpose of this activity it was explicitly defined as "the reduction of support delivered by the mechanical ventilator with an end goal of extubation." Consequently, the activity brief prompted participants to annotate periods during which mechanical ventilation weaning takes place and allowed multiple annotations for any individual admission.

Admission Data

The data used in the activity came from a deidentified intensive care dataset called "Medical Information Mart for Intensive Care IV version 2.0" [37]. The eligibility criteria required that subjects were at least 18 years of age at the time of admission, had undergone an invasive mechanical ventilation treatment that lasted for a minimum of 24 hours, their stay in the ICU lasted for a minimum of 4 days, and that their stay did not end with death, including up to 48 hours after discharge. The time-series parameters used in the data extract were limited to those relevant to mechanical ventilation weaning and were selected by 2 independent clinicians working in the ICU. The data were extracted using a structured query language script run on the PostgreSQL installation of Medical Information Mart for Intensive Care IV version 2.0 with the concept tables computed [38]. The script aggregated the selected parameters on an hourly basis for each eligible admission and limited it to the range surrounding the period of mechanical ventilation treatment.

Ethical Considerations

This work was approved by the Faculty of Engineering Research ethics committee at the University of Bristol (case 2022-150, research ethics committee reference 22/HRA/2166). Information regarding the study was circulated with the invited participants electronically alongside the recruitment email. Informed consent was captured electronically using digitally signed consent forms prior to enrollment in the study. Data obtained from this study were deidentified and stored securely in an encrypted database. The access to the database was password-protected and only the research team had access to these data. All participants were able to quit without any explanation at any point during the study. Participation in the study was voluntary, and no financial or other form of compensation was provided to the participants.

Results

Data Annotation Analysis

During the simulated data annotation activity, we observed participants following a series of discrete steps when annotating data. To understand how these actions fit into the annotation process, we analyzed the observations in the context of NIC and aggregated them into the distinct phases of the model.

Planning

Following the distribution of the printout sheets containing the admission data, participants began the process of familiarizing themselves with the data. This stage involved selecting a single excerpt from the available admissions, analyzing the patient demographics, and browsing through the time-series data. During the process, participants frequently cycled through the entire length of the time-series data from the initial admission until discharge and back.

Specifying

Upon familiarization with the specific admission, participants began searching for individual points in time suggestive of the weaning taking place. Some participants preferred to start with the end of treatment, where weaning led to extubation, while others analyzed the data to find the points in time when weaning began. This involved focusing on specific parameters (as suggested by following the specific rows with the pen tip) while browsing through the time-series data across different columns.

Performing

Upon finding the thresholds for the start or end of weaning, the participants annotated the printed sheets directly. To create the labels, a variety of techniques were used, which differed both between the individual annotators and the distinct admissions. These included circling the start and end dates in the header row of corresponding columns, drawing vertical lines on the edges of the columns to mark the start and end periods, and drawing a box around the portion of the dataset spanning the duration of the label. Similarly to the analysis process, some participants preferred to first annotate the end of the weaning process, rather than its start. In some cases, participants also provided additional information in the form of written comments that justified the created label (eg, "mode of ventilation changed from A to B" or "delivered oxygen reduced—indicative of weaning") and underlined values within cells of parameters that prompted annotation.

Perceiving

Following the act of annotating the data, participants frequently verified their labels. This was expressed by browsing through the time-series data again and ensuring that the start and end dates corresponded to the parameter readings that suggested the annotation in the first place.

XSL•FO
RenderX

Interpreting

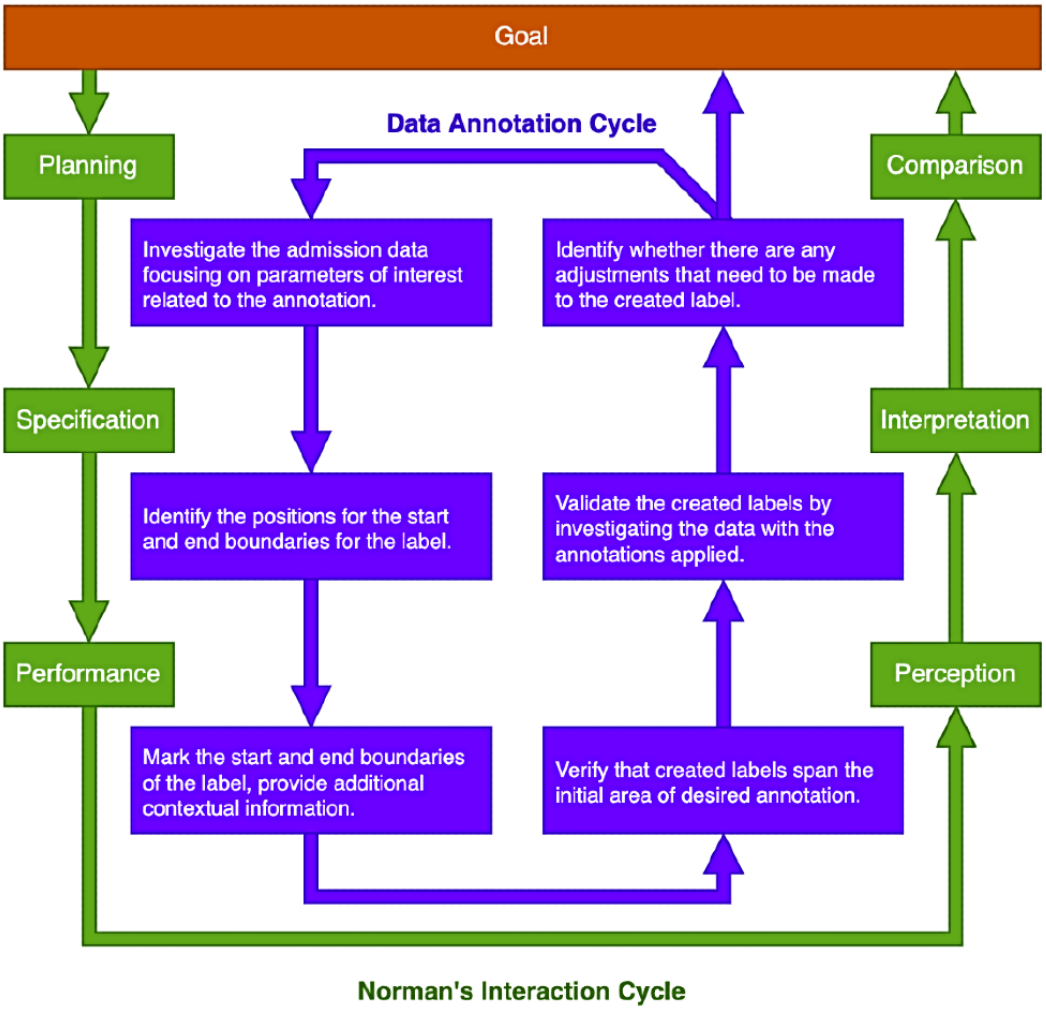
In some cases, participants derived additional insights from the data when investigating them with their labels already present. These insights frequently manifested in the form of retracing the admission data over the span of the created label, which resulted in a reevaluation of the created label and its underlying data.

Comparing

Finally, participants who identified the need for amendments returned to the planning stage to repeat the annotation process, whereas those satisfied with their annotations frequently sought feedback from the activity facilitator on the desired annotation count and the next steps involved in the process.

Crucially, by contextualizing the annotation process within the NIC model we identified that the overall approach to annotation followed a cyclic process of investigation of the data, creation of the labels, reevaluation of the data with the labels applied, and refinement of previously created labels (Figure 4).

Figure 4. Data annotation followed a cyclic process involving investigation of data, label creation, reevaluation of data with labels applied, and refinement of created annotations.



User Requirements

Observations from the workshops captured key characteristics of participant interactions with the task, forming a basis for establishing the functional requirements of the digital tool. Coupled with the operational constraints of data annotation in clinical settings and the specific needs for future label use in the ML context, these characteristics informed the user requirements for the digital annotation tool.

Annotation of Individual Admissions

Familiarization with admission data during the planning and specification stages suggested that the primary goal of participants at the early stage of the process was to obtain a high-level overview of the patient's characteristics and their stay in the ICU. This was supported by the frequent cycling through different parts of the admission, allowing participants to establish a broad timeline of events and the overall treatment trajectory without focusing on the minutia of its delivery. The interface of the tool should allow end users to conduct a similar analysis digitally, incorporating both the patient demographics and the time-series data.

Once familiarized with the admission data, participants shifted their focus to a more task-oriented analysis by narrowing the list of parameters to a subset significant for the task, suggesting

XSL•FO
RenderX

different relevance of specific parameters to the target label. Furthermore, the fact that in certain cases participants found it easier to first search for the end of weaning, rather than its start, suggests that their confidence in the accuracy of the created label, its precise start or end, might be different across annotations. Similarly, each of the participants displayed different preferences for creating annotations favoring first marking either the start of the weaning process or its end. Despite this, they all shared a common set of actions of marking the beginning and end of the label, even when the order of these operations was not the same. To make the annotation process as unobtrusive as possible, the digital data annotation tool should therefore allow flexibility in not only how the data are viewed and analyzed but also how the annotations are being created. As such, the interface needs to display a continuous dataset without constraining it to a subset of time or parameters and allow the end users to freely mark the boundaries of the label in the order of their preference.

Following the creation of a label, participants frequently continued to investigate the underlying data and further adjusted the labels, indicating the need to reflect on their immediate performance and refine the annotations. Adjusting labels upon their reevaluation indicates that when the data are viewed with the annotations applied, the annotator's perception of these data changes allowing for additional insights to be derived. As such, the digital tool should allow for the display of the created annotations on top of the data and facilitate a similar ability to adjust and modify them once they are created.

Participants also frequently provided additional information to support their labels, such as comments or specific parameters that led them to create the annotation. This contextual information could be used as a surrogate for the thought process of the annotators and further inform the importance of specific parameters on the target label. The digital tool should therefore incorporate the ability to provide additional data, such as parameters of interest, confidence in the accuracy of created labels, and free-form text that provides further context.

Finally, the feedback sought upon completing the annotation indicated the need for both a progress indicator and the importance of preservation of the continuity of the annotation process on retaining participants' focus. A digital tool could account for this by displaying the count and information about the already created labels and making it easy and efficient for participants to annotate consecutive admissions. As such, the functional requirements for the digital data annotation tool can be specified as follows:

- R1: Data analysis in which participants are free to navigate through the entire span of the admission data, as well as underlying patient demographics.
- R2: Label creation functionality which is flexible enough to allow end users to select the start and end of annotation independently and in any order.
- R3: Label adjustment capability that facilitates an easy and convenient way to amend the created label, particularly with that label presented on top of data.
- R4: Label supplementation feature that allows participants to provide additional context for the created annotation, including confidence in the label accuracy and relevance of different parameters.
- R5: Workflow continuity that ensures a cohesive process of annotation and keeps the end users engaged and focused on the task.

Semiautomated Annotation

Facilitating a semiautomated approach to annotation that allows for creating labels for an entire dataset is a separate but closely related task. As such, it can benefit from the analysis of the manual annotation process and use it to further inform the requirements for the digital tool. The annotation of a single admission could be described as a bottom-up approach, in which end users are presented with data specific to a single admission and asked to annotate them directly (optimally until the entire dataset is annotated). Conversely, adopting an approach that automates the process would likely involve the annotation of an entire dataset based on a single user input, constituting a top-down approach. A semiautomated approach would therefore focus on creating labels for the whole dataset without focusing on individual admissions during their creation but potentially using them to evaluate the annotations and refine them upon further analysis. To that extent, the digital tool needs to allow its end users to establish annotation strategies that are independent of individual admissions during their formulation.

The process in which participants interacted with the individual admissions suggested several key requirements that needed to be adapted for the semiautomated approach. To close the loop of investigating, annotating, evaluating, and refining that the participants exhibited during the simulated activity, the semiautomated annotation approach facilitated by the platform should also allow for analysis of the effectiveness of the annotation in the context of both the entire dataset and individual admissions. In particular, the interface should allow the end users to view individual admissions with the created annotations applied, providing additional explanations for why the label was created, and in turn, allowing them to refine their annotation strategies. Consequently, the requirements for the digital tool facilitating a semiautomated annotation are defined as follows:

- R6: Annotation and evaluation loop captured as one continuous process that preserves the focus of the end users and facilitates effective annotation.
- R7: Label analysis of individual admissions with the overlay of created annotations on the time-series data and justification for their presence.
- R8: Dataset-wide performance metrics that capture the effectiveness of the annotation in the form of aggregate statistics computed across the entire dataset.

Operational Constraints

In addition to the requirements identified through the analysis of participants' approach to the annotation task, there are further operational constraints that need to be reconciled to deliver an effective annotation workflow. To accommodate the busy schedules of the ICU staff, the tool should allow its end users to access it in a way that suits their needs and does not impose strict time commitments. As such, end users should be able to perform the annotation asynchronously and access the platform

XSL•FO
RenderX

remotely. Furthermore, to preserve the focus and ensure that the end users' time is used efficiently, the system should provide a responsive and time-efficient workflow that enables them to create and evaluate annotations in a short span of time. This is particularly important for the semiautomated approach in which the timely annotation of a large volume of data is critical to preserve the continuity of the annotation and evaluation loop. The operational requirements for the annotation tool are therefore defined as follows:

- R9: Asynchronous and remote annotation that allows the end users to perform the task at their convenience without further complicating their busy schedules.
- R10: Responsiveness of the system that provides feedback on the created annotations in a timely manner, particularly in the case of the semiautomated annotation.

Labels for ML

Finally, having identified the functional and operational requirements, we analyzed the platform in the wider context of using the created labels in ML workflows. To that extent, the discussion between data science experts within our research team surfaced several key factors that the tool should account for to ensure that the captured data can seamlessly integrate into ML pipelines.

The most important one to consider is the need to achieve a careful balance between the number of annotated admissions and the confidence in the created annotations. On one hand, increasing the total number of annotated admissions and their diversity would result in a larger training dataset and, consequently, improve the performance of the resulting ML model. On the other hand, to account for the biases of individual annotators, there needs to be some overlap of admissions annotated by distinct annotators that would allow for a comparison of their approaches and biases. Furthermore, for this to happen, additional metadata surrounding the confidence in the specific annotations and relating the created annotations to their authors also needs to be collected. Only then, will the data scientists be able to evaluate each of the annotators individually, compute their biases, and adjust their annotations accordingly to improve the robustness of the training dataset.

Depending on the volume of the data that needs to be annotated, the nature of the annotation task, and the number and experience of the annotators, this balance might be substantially different. Because of this, the system needs to facilitate a dynamic and adjustable configuration for how specific data are assigned to the annotators. This suggests that in addition to the previously established requirements, the tool should also aim to satisfy the following ML-specific requirement:

- R11: Flexible data-splitting solution that allows for adjustment of the data assigned to each participant and keeps track of the label authorship.

Discussion

Principal Results

This study aimed to explore how intensive care experts approach data annotation to establish requirements for a digital tool designed for annotating large datasets in intensive care. We conducted a manual data annotation activity, observing participants as they completed the task and analyzing their actions to gain insights into their annotation strategies. This analysis informed the development of software requirements for a digital data annotation platform.

Our findings revealed that the time-series annotation process follows a cyclical annotation-evaluation loop, which includes data investigation, label creation, reevaluation of data with the labels applied, and refinement of the created labels. From this analysis, we constructed 11 key requirements: 5 directly related to facilitating the annotation of individual patient admissions, 3 adapted for implementing a semiautomated annotation feature, 2 operational requirements focused on the effectiveness of the annotation workflow, and 1 requirement addressing the future use of labels in the ML context (Table 1).

XSL•FO
RenderX

Table 1. Requirements generated from analysis of observations made during simulated annotation activity.

Category	Requirement description
Annotation of individual admissions	
R1	Data analysis in which participants are free to navigate through the entire span of the admission data, as well as underlying patient demographics.
R2	Label creation functionality which is flexible enough to allow end users to select the start and end of annotation independently and in any order.
R3	Label adjustment capability that facilitates an easy and convenient way to amend the created label, particularly with that label presented on top of data.
R4	Label supplementation feature that allows participants to provide additional context for the created annotation, including confidence in the label accuracy and relevance of different parameters.
R5	Workflow continuity that ensures a cohesive process of annotation and keeps the end users engaged and focused on the task.
Semiautomated annotation	
R6	Annotation and evaluation loop captured as one continuous process that preserves the focus of the end users and facilitates effective annotation.
R7	Label analysis of individual admissions with the overlay of created annotations on the time-series data and justification for their presence.
R8	Dataset-wide performance metrics that capture the effectiveness of the annotation in the form of aggregate statistics computed across the entire dataset.
Operational constraints	
R9	Asynchronous and remote annotation that allows the end users to perform the task at their convenience without further complicating their busy schedules.
R10	Responsiveness of the system that provides feedback on the created annotations in a timely manner, particularly in the case of the semiautomated annotation.
Use in machine learning	
R11	Flexible data-splitting solution that allows for adjustment of the data assigned to each participant and keeps track of the label authorship.

Approach to Annotation

The observations made during the manual annotation activity highlighted important characteristics of how experts from a clinical background approach the annotation task. While the sample of our study produced no discernable differences in how participants from different roles approached the task of data annotation, we observed a variety of strategies used to complete the task. These included several techniques for initial familiarization with the data, prioritization of different parameters during the investigation process, and different order of operations when creating the label. Crucially, we observed that irrespective of the specific strategies, annotators followed a cyclic process of investigation, annotation, reevaluation, and refinement when annotating data. By modeling the process using NIC, we were able to analyze the observations made during the task completion stage and formulate the requirements for the software interface of a digital data annotation platform.

In our analysis, we found that the nature of how annotators choose to analyze the data is largely preferential and, therefore, requires a degree of flexibility in how the software presents the data and allows its users to navigate through it. The cyclic nature of the process suggested that to ensure effective annotation, it must facilitate a continuous and intuitive loop of annotation, allow users to perform actions in varying order, and seamlessly adapt to different stages of the process. Since these requirements were derived from the simulated annotation of individual admissions but captured the characteristics of the overall problem of data annotation, the requirements for the semiautomated approach were adapted to ensure the continuity of the annotation-evaluation loop and flexibility in analysis.

Finally, the fact that annotators frequently sought feedback after creating labels further emphasized the importance of ensuring a seamless and uninterrupted annotation workflow and suggested the need to integrate progress indicators within the digital platform. Together with individual preferences for focusing on different parameters during the analysis and supplying additional information as part of the created labels, this also highlighted the need to capture contextual information alongside labels, which could be beneficial for their later use in the ML pipelines.

Design Implications

This literature on involving clinical staff in the collaborative design of software interventions shows that it can have profound effects on user acceptance and uptake upon release [23,39]. However, it is important to highlight that effectively engaging in such collaboration can be a challenging task, particularly when research activities require substantial time commitment [40]. Collaborating with intensive care clinicians, whose time is in high demand and short supply, may therefore require additional accommodations to be effective. In the context of annotating data, these accommodations translate into additional

XSL•FO
RenderX

requirements that adapt the process to the busy schedules of the clinical staff by allowing them to annotate data both asynchronously and remotely. This further reinforces the need to ensure that the software delivers an intuitive and time-efficient annotation workflow and places additional constraints on its design. Specifically, to facilitate the responsiveness of the annotation-evaluation loop in the semiautomated approach, where users can frequently simultaneously annotate the entire dataset, the software should incorporate a solution that dynamically adapts to both the number of active users and the volume of data that needs to be processed.

Limitations and Future Work

The activity focused on capturing the approach to data annotation was conducted with a limited number of participants, which resulted in a limited diversity of clinical roles and, potentially, perspectives from the ICU. This suggests that some of the requirements established in this research could be particularly applicable to annotators in junior doctor positions and therefore not necessarily generalizable across entire staff working in ICUs. The selection of the task for the manual annotation activity was made to provide an annotation experience that could be extrapolated beyond the specifics of the task. Despite this, we acknowledge that the results it produced could be biased specifically toward the annotation of weaning from mechanical ventilation.

Furthermore, our approach to capturing the annotation process during the simulated activity with only a single observer had a significant impact on the quantity and quality of observations that were collected. Due to the imbalance in the number of annotators and observers, some of the actions undertaken by the participants during the annotation could have gone unnoticed. Furthermore, the presence of a single observer created an inherent bias in the collected observations, as the observed actions could have been perceived differently by observers with different backgrounds or characteristics. These limitations suggest that the observations collected during the activity come with a degree of incompleteness and inaccuracy, which may have impacted the elicited requirements.

We acknowledge that further work in establishing the requirements for the digital tool could strengthen the understanding of how clinicians approach the data annotation task. To that extent, we suggest that further research in this area focuses on capturing the requirements that expand beyond the confines of a single annotation task and within a broader and more diverse population from the ICUs. Capturing a wider range of perspectives could inform the applicability of the elicited requirements and, in consequence, strengthen the resulting design of the digital tool.

The nature of the activity itself was focused strictly on the direct annotation of individual admissions rather than the use of any assistive or automated technologies. Although several requirements elicited in this context applied to the overall annotation process, including a semiautomated approach, additional requirements not captured in this study may also exist. Further research should therefore focus on the evaluation of the proposed requirements and their use and limitations in real-world applications. Therefore, these requirements should be used to design and implement a digital data annotation platform that should be trialed within a clinical setting. Conducting a study that investigates the feasibility of a semiautomated approach to the data annotation, particularly in comparison with the direct annotation of individual admissions, could further inform the requirements for data annotation tools.

Conclusions

In this study, we investigated how clinical staff from ICUs approach the task of data annotation and established 11 key requirements for a digital data annotation tool that could be deployed within the health care setting. Our findings revealed that data annotation is a cyclic process that demands flexibility in how annotators investigate and annotate the data. Preservation of the workflow continuity across different admissions and fluid transition between analysis, annotation, and refinement of the label are essential to facilitating effective data annotation in the clinical domain. Adaptations for the semiautomated annotation need to consider these factors by providing a responsive interface that dynamically adapts to the volume of data and allows for analysis on both an individual and a dataset-wide basis. The significance of these findings is evident in their potential to guide the development of a data annotation tool that capitalizes on the considerable data generated in the ICUs and the expanding use of computational methods in health care.

Acknowledgments

This work was partly supported by the Engineering and Physical Sciences Research Council Digital Health and Care Centre for Doctoral Training at the University of Bristol (UK Research and Innovation grant EP/S023704/1 to MW). The work of RSR was funded by the UK Research and Innovation Turing AI Fellowship (grant EP/V024817/1). The project received an AWS Cloud Credit for Research grant.

Authors' Contributions

MW was involved in the study design, workshop facilitation, and writing. RSR, CMW, and CB were involved in the study design and critical revision.

Conflicts of Interest

None declared.

XSL•FO
RenderX

References

1. Russell SJ, Norvig P. Artificial Intelligence: A Modern Approach, Global Edition. In: Prentice Hall Series in Artificial Intelligence. Singapore Taipei Tokyo: Pearson; 2016.
2. Bellman R. An Introduction to Artificial Intelligence: Can Computers Think?. San Francisco: Boyd & Fraser Publishing Company; 1978.
3. Grosan C, Abraham A. Intelligent Systems: A Modern Approach. In: Intelligent Systems Reference Library. Berlin: Springer-Verlag; 2011.
4. Hopfield JJ. Neural networks and physical systems with emergent collective computational abilities. Proc Natl Acad Sci USA 1982;79(8):2554-2558 [FREE Full text] [doi: 10.1073/pnas.79.8.2554] [Medline: 6953413]
5. Introduction TML, Fourth E. Introduction to Machine Learning, Fourth Edition. In: Adaptive Computation and Machine Learning Series. Cambridge, MA: MIT Press; 2020.
6. Jain A. Overview and importance of data quality for machine learning tasks. 2020 Presented at: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Virtual Event CA USA: ACM; 10 Jan 2025; United States p. 3561-3562. [doi: 10.1145/3394486.3406477]
7. Noorbakhsh-Sabet N, Zand R, Zhang Y, Abedi V. Artificial intelligence transforms the future of health care. Am J Med 2019;132(7):795-801 [FREE Full text] [doi: 10.1016/j.amjmed.2019.01.017] [Medline: 30710543]
8. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. Future Healthc J 2019;6(2):94-98 [FREE Full text] [doi: 10.7861/futurehosp.6-2-94] [Medline: 31363513]
9. Chapalain X, Huet O. Is artificial intelligence (AI) at the doorstep of intensive care units (ICU) and operating room (OR)? Anaesth Crit Care Pain Med 2019;38(4):337-338. [doi: 10.1016/j.accpm.2019.05.003] [Medline: 31102792]
10. Hunter B, Hindocha S, Lee RW. The role of artificial intelligence in early cancer diagnosis. Cancers (Basel) 2022;14(6):1524 [FREE Full text] [doi: 10.3390/cancers14061524] [Medline: 35326674]
11. Heo MS, Kim J, Hwang J, Han S, Kim J, Yi W, et al. Artificial intelligence in oral and maxillofacial radiology: what is currently possible? Dentomaxillofac Radiol 2021;50(3):20200375 [FREE Full text] [doi: 10.1259/dmfr.20200375] [Medline: 33197209]
12. Azizi S. Big self-supervised models advance medical image classification. 2021 Presented at: 2021 IEEE/CVF International Conference on Computer Vision (ICCV); 10 Oct 2021; Montreal, QC, Canada p. 3458-3468. [doi: 10.1109/iccv48922.2021.00346]
13. Martinez-Martin N, Luo Z, Kaushal A, Adeli E, Haque A, Kelly SS, et al. Ethical issues in using ambient intelligence in health-care settings. Lancet Digital Health 2021;3(2):e115-e123. [doi: 10.1016/s2589-7500(20)30275-2]
14. Burki TK. Artificial intelligence hold promise in the ICU. Lancet Respir Med 2021;9(8):826-828. [doi: 10.1016/s2213-2600(21)00317-9]
15. Gautam S, Srinath M. Blind spots and biases: exploring the role of annotator cognitive biases in NLP. arXiv Preprint posted online on April 29, 2024 [FREE Full text] [doi: 10.18653/v1/2024.hcinlp-1.8]
16. Spasic I, Nenadic G. Clinical text data in machine learning: systematic review. JMIR Med Inform 2020;8(3):e17984 [FREE Full text] [doi: 10.2196/17984] [Medline: 32229465]
17. Chen Y, Lasko TA, Mei Q, Denny JC, Xu H. A study of active learning methods for named entity recognition in clinical text. J Biomed Inform 2015;58:11-18 [FREE Full text] [doi: 10.1016/j.jbi.2015.09.010] [Medline: 26385377]
18. Kholghi M, Sitbon L, Zuccon G, Nguyen A. Active learning: a step towards automating medical concept extraction. J Am Med Inform Assoc 2016;23(2):289-296 [FREE Full text] [doi: 10.1093/jamia/ocv069] [Medline: 26253132]
19. South BR, Mowery D, Suo Y, Leng J, Ferrández Ó, Meystre SM, et al. Evaluating the effects of machine pre-annotation and an interactive annotation interface on manual de-identification of clinical text. J Biomed Inform 2014;50:162-172 [FREE Full text] [doi: 10.1016/j.jbi.2014.05.002] [Medline: 24859155]
20. Pezoulas VC, Zaridis DI, Mylona E, Androutsos C, Apostolidis K, Tachos NS, et al. Synthetic data generation methods in healthcare: a review on open-source tools and methods. Comput Struct Biotechnol J 2024;23:2892-2910 [FREE Full text] [doi: 10.1016/j.csbj.2024.07.005] [Medline: 39108677]
21. Vandekerckhove P, de Mul M, Bramer WM, de Bont AA. Generative participatory design methodology to develop electronic health interventions: systematic literature review. J Med Internet Res 2020;22(4):e13780 [FREE Full text] [doi: 10.2196/13780] [Medline: 32338617]
22. Kautz K. Participatory design activities and agile software development in human benefit through the diffusion of information systems design science research. In: Pries-Heje J, Venable J, Bunker D, Russo NL, DeGross JI, editors. IFIP Advances in Information and Communication Technology. Berlin, Heidelberg: Springer; 2020:303-316.
23. Damodaran L. User involvement in the systems design process-a practical guide for users. Behav Inf Technol 1996;15(6):363-377. [doi: 10.1080/014492996120049]
24. Sanz MF, Acha BV, García MF. Co-design for people-centred care digital solutions: a literature review. Int J Integr Care 2021;21(2):16 [FREE Full text] [doi: 10.5334/ijic.5573] [Medline: 33981193]
25. Curtis K, Brooks S. Digital health technology: factors affecting implementation in nursing homes. Nurs Older People 2020;32(2):14-21. [doi: 10.7748/nop.2020.e1236] [Medline: 32159302]

XSL•FO
RenderX

26. Hobson EV, Baird WO, Partridge R, Cooper CL, Mawson S, Quinn A, et al. The TiM system: developing a novel telehealth service to improve access to specialist care in motor neurone disease using user-centered design. Amyotroph Lateral Scler Frontotemporal Degener 2018;19(5-6):351-361 [FREE Full text] [doi: 10.1080/21678421.2018.1440408] [Medline: 29451026]
27. Norman DA. The Design of Everyday Things, Revised and Expandes Editons. Cambridge, MA London: MIT Press; 2013.
28. Soegaard M, Dam RF. Gulf of evaluation and gulf of execution. Interaction Design Foundation. 2015. URL: https://www.interaction-design.org/literature/book/the-glossary-of-human-computer-interaction/gulf-of-evaluation-and-gulf-of-execution [accessed 2025-09-10]
29. Slutsky AS, Hudson LD. Mechanical ventilation. In: Goldman's Cecil Medicine. Philadelphia: Elsevier; 2012:638-645.
30. Girard TD. Efficacy and safety of a paired sedation and ventilator weaning protocol for mechanically ventilated patients in intensive care (awakening and breathing controlled trial): a randomised controlled trial. Lancet (London, England) 2008;371(9607):126-134. [doi: 10.3410/f.1097928.562141]
31. Nichol AD, Geoghegan P, Duff S, Cooper DJ, Devlin D. What is the optimal approach to weaning and liberation from mechanical ventilation? Evidence-Based Practice of Critical Care.: Elsevier; 2019. URL: https://research.monash.edu/en/publications/what-is-the-optimal-approach-to-weaning-and-liberation-from-mecha-2 [accessed 2025-09-10]
32. McConville JF, Kress JP. Weaning patients from the ventilator. N Engl J Med 2012;367(23):2233-2239. [doi: 10.1056/nejmra1203367]
33. Nitta K, Okamoto K, Imamura H, Mochizuki K, Takayama H, Kamijo H, et al. A comprehensive protocol for ventilator weaning and extubation: a prospective observational study. J Intensive Care 2019;7(1):50 [FREE Full text] [doi: 10.1186/s40560-019-0402-4] [Medline: 31719990]
34. Mador MJ. Weaning from mechanical ventilation: what have we learned and what do we still need to know? Chest 1998;114(3):672-674. [doi: 10.1378/chest.114.3.672] [Medline: 9743148]
35. Boles JM, Bion J, Connors A, Herridge M, Marsh B, Melot C, et al. Weaning from mechanical ventilation. Eur Respir J 2007;29(5):1033-1056 [FREE Full text] [doi: 10.1183/09031936.00010206] [Medline: 17470624]
36. Lermitte J, Garfield MJ. Weaning from mechanical ventilation. Contin Educ Anaesth Crit Care Pain 2005;5(4):113-117. [doi: 10.1093/bjaceaccp/mki031]
37. Johnson A, Bulgarelli L, Pollard T, Gow B, Moody B, Horng S, et al. MIMIC-IV. PhysioNet. 2022. URL: https://physionet.org/content/mimiciv/3.1/ [accessed 2025-01-10]
38. Johnson AEW, Stone DJ, Celi LA, Pollard TJ. The MIMIC code repository: enabling reproducibility in critical care research. J Am Med Inform Assoc 2018;25(1):32-39 [FREE Full text] [doi: 10.1093/jamia/ocx084] [Medline: 29036464]
39. Slattery P, Saeri AK, Bragge P. Research co-design in health: a rapid overview of reviews. Health Res Policy Syst 2020;18(1):17 [FREE Full text] [doi: 10.1186/s12961-020-0528-9] [Medline: 32046728]
40. Kushniruk A, Hr C. Participatory design, user involvement and health IT evaluation, in evidence-based health informatics. In: Studies in Health Technology and Informatics. Amsterdam: IOS Press; 2016:139-151.

Abbreviations

AI: artificial intelligence
ICU: intensive care unit
ML: machine learning
NIC: Norman's Interaction Cycle

Edited by A Choudhury; submitted 29.01.24; peer-reviewed by T Abd El-Hafeez, S Tayebi Arasteh, M Lotfinia; comments to author 11.10.24; revised version received 05.12.24; accepted 01.01.25; published 05.02.25.

Please cite as:
Wac M, Santos-Rodriguez R, McWilliams C, Bourdeaux C
Capturing Requirements for a Data Annotation Tool for Intensive Care: Experimental User-Centered Design Study
JMIR Hum Factors 2025;12:e56880
URL: https://humanfactors.jmir.org/2025/1/e56880
doi:10.2196/56880
PMID:

©Marceli Wac, Raul Santos-Rodriguez, Chris McWilliams, Christopher Bourdeaux. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 05.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic

XSL•FO
RenderX

information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Internet-Based Mental Health Intervention for Depressive Symptoms in Young Adults: Cost-Effectiveness Analysis

Joyce H S You[1], PharmD; Scotty W C Luk[2], MBChB; Dilys Y W Chow[1], MClinPharm; Arthur D P Mak[3], MBChB; Winnie W S Mak[4], PhD

[1]School of Pharmacy, Faculty of Medicine, The Chinese University of Hong Kong, 8th Floor, Lo Kwee-Seong Integrated Biomedical Sciences Building, Area 39, Shatin, N.T., Hong Kong SAR, China

[2]University Medical Service Office, The Chinese University of Hong Kong, Hong Kong SAR, China

[3]Department of Psychiatry, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong SAR, China

[4]Department of Psychology, Faculty of Social Science, The Chinese University of Hong Kong, Hong Kong SAR, China

Corresponding Author:
Joyce H S You, PharmD
School of Pharmacy, Faculty of Medicine, The Chinese University of Hong Kong, 8th Floor, Lo Kwee-Seong Integrated Biomedical Sciences Building, Area 39, Shatin, N.T., Hong Kong SAR, China

Abstract

Background: Internet-based cognitive behavioral therapy (CBT) provides psychological interventions to individuals with mild depressive symptoms.

Objective: This study aimed to examine the potential cost-effectiveness of internet-based guided-CBT in university students with mild depressive symptoms from the perspective of service providers in Hong Kong.

Methods: The outcomes of low-intensity guided internet-based CBT and in-person CBT in a hypothetical cohort of university students with mild depressive symptoms were examined using a 5-year decision-analytic model. Model inputs were obtained from published literature and local data. Model outcomes included direct medical cost, school dropouts, and quality-adjusted life years (QALYs). Sensitivity analyses were conducted on all model parameters.

Results: Compared to the in-person group, the internet group gained higher QALYs by 0.0211 QALYs, lowered school dropouts by 0.052%, and saved US $249 in the base-case analysis. In one-way sensitivity analysis, the internet group gained higher QALYs at a lower cost than the in-person group throughout the variation of all model inputs. Probabilistic sensitivity analysis showed that the internet group was cost-effective (at willingness-to-pay threshold was US $48,119/QALY) in 99.7% of the 10,000 Monte Carlo simulations.

Conclusions: Internet-based CBT appears to be the cost-effective option when compared to in-person CBT for university students with mild depressive symptoms from the perspective of service providers in Hong Kong.

(JMIR Hum Factors 2025;12:e68167) doi:10.2196/68167

KEYWORDS

internet; cognitive behavioral therapy; telehealth; depression; university students; health economic analysis

Introduction

Depression is a leading cause of disability worldwide. The World Health Organization (WHO) reported that approximately 33% first-year college students in 8 countries were screened positive for a common *DSM-IV* (*Diagnostic and Statistical Manual of Mental Disorders IV*) mental disorders [1]. The pooled depression prevalence in university students from 81 studies (near 1.5 million subjects) was estimated to be 34% [2].

Cognitive behavioral therapy (CBT) is a widely used psychological therapy for symptoms of depressive disorders and has been traditionally administered as in-person interventions. Young adults with mild symptoms of depression typically delay face-to-face intervention with a therapist, plausibly associated to several possible reasons. Perceived stigma and embarrassment at face-to-face intervention is one potential barrier [3]. Also, the accessibility of face-to-face psychological interventions may also be limited by monetary or time requirements, as well as by the service delivery barriers, such as insufficient psychological services and inadequate workforce numbers [3]. Internet-based CBT interventions provide alternative options that are self-paced and highly accessible, and computerized low-intensity CBT is recommended by the National Institute for Health and Care Excellence (NICE) as an option for mild depression [4]. Low-intensity CBT is a structured self-help program, including 6 - 8 brief sessions (20 - 30 min per session) over 9 - 12 weeks [4]. Meta-analyses reported superior efficacy of internet-based

XSL•FO
RenderX

CBT when compared to the waitlist or usual care [5,6]. Offering guided internet-based low-intensity CBT could enhance the treatment and health economic outcomes of mental health management for university students with mild symptoms of depression. In addition to clinical evidence, health economic findings are important evidence to facilitate the informed decision-making process for policy makers on the uptake of a health technology. This study aimed to examine the potential cost-effectiveness of internet-based guided CBT in university students with mild depressive symptoms from the perspective of service provider in Hong Kong.

Methods

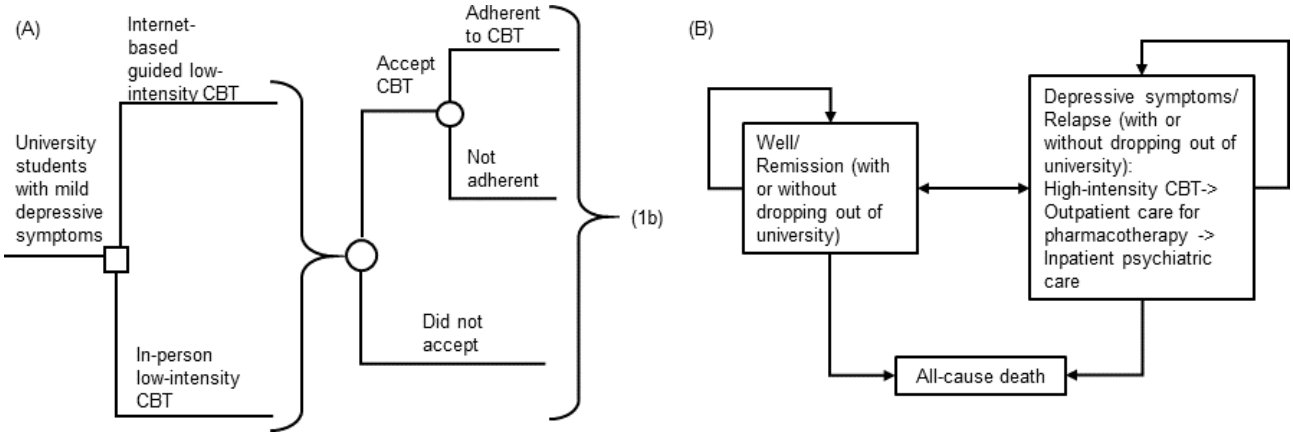
Ethical Considerations

The protocol of retrospective review had been approved by the Joint Chinese University of Hong Kong–New Territories East Cluster Clinical Research Ethics Committee. The need for informed consent was waived by the Joint Chinese University of Hong Kong–New Territories east Cluster Clinical Research Ethics Committee. Data deidentification was performed in the retrospective data review process, and no compensation was provided.

Model Overview

A decision-analytic model (Figure 1A–B), including a short-term decision tree model followed by a 5-year Markov model, was designed to evaluate the potential health economic outcomes of a hypothetical cohort of university students with mild depressive symptoms managed by two strategies: (1) therapist-guided, internet-based low-intensity CBT (internet group) or (2) conventional in-person low-intensity CBT (in-person group). Depression symptom classifications adopted the Patient Health Questionnaire-9 (PHQ-9) [7]: mild (PHQ-9 score 5 - 9), moderate (PHQ-9 score 10 - 14), or severe (PHQ-9 score ≥15).

Figure 1. (A-B) Simplified decision-analytical model. Depressive symptoms/relapse status included cases of mild, moderate, or severe depressive symptoms. CBT: cognitive behavioral therapy.



Low-intensity CBT (delivered in-person or via the internet guided by a therapist) recommended by NICE for mild depressive symptoms is a structured self-help program that includes 6 - 8 sessions over 9 - 12 weeks. In the in-person group, the participants work through the CBT-based self-help workbooks with brief support from the therapist in face-to-face sessions (20 - 30 min per session). In the internet-based CBT, the low-intensity CBT program was offered via the internet platform. The participants work through the CBT-based self-help materials (resembling the structure of face-to-face CBT) delivered by online modules, and the therapist provided written feedback to the participant's self-help work by emails or texting [4].

All hypothetical students with mild depressive symptoms entered a short-term model mirroring the duration of the low-intensity CBT program (Figure 1A). The hypothetical student (in both groups) might accept the CBT program, and those who accepted the therapy might adhere with the course of CBT. Upon the completion of low-intensity CBT, the student might recover well or deteriorate (showing moderate or severe depressive symptoms), and all were advanced to a 5-year Markov model (Figure 1B).

In the Markov model (Figure 1B), the hypothetical subjects proceeded through health states (Markov states) in each model cycle according to transition probabilities [8]. The present Markov model consisted of three key Markov states (well/remission, depressive symptoms/relapse, and all-cause death). Within the depressive symptoms/relapse status, cases are further differentiated (by proportion) into depressive symptoms of mild, moderate, or severe. The students who deteriorated or did not achieve remission with the low-intensity CBT program were stepped up to high-intensity CBT and outpatient psychiatric care for pharmacotherapy [4]. Students with severe depressive symptoms might be admitted to hospital for inpatient psychiatric care. In each model cycle, all hypothetical students (those in remission and those with depressive symptoms) might (or might not) permanently dropout from the university study. The model outcomes included direct medical cost, dropout, and quality-adjusted life years (QALYs). The model time horizon was 5 years, and the cycle length was 1 month. With the non-responding and recurring nature of depressive symptoms, a timeframe of 5 years was adequate to capture the outcomes of internet-based and conventional in-person low-intensity CBT, and the downstream costs and QALYs associated with deteriorated and non-responsive depressive symptoms.

XSL•FO
RenderX

Clinical Inputs

The clinical model inputs (Table 1) were retrieved from published clinical studies. A MEDLINE search (2000 - 2023) was performed using keywords such as “depressive disorders,” “internet,” “telehealth,” “telemedicine,” “cognitive behavioral therapy,”“cognitive behavioral therapy,” “pharmacotherapy,” “uptake,” “compliance,” “remission,” and “deterioration.” Selection criteria of published studies included: (1) Reports published in English language, (2) adults with symptoms of depressive symptoms, (3) event rates of remission and/or deterioration, and (4) individuals’ uptake of the delivery platform and (5) compliance of delivery platform (online and in-person) for CBT were reported. Meta-analyses and randomized controlled trials were the preferred types of studies. When multiple clinical randomized trials were available for the same model input, the base-case value was derived by the weighted average of the reported findings.

XSL•FO
RenderX

Table . Model input parameters.

Parameters	Base-case value	Range of sensitivity analysis	Distribution	Source of Information
Clinical inputs				
Low-intensity CBT[a] for mild depressive symptoms, event rate (%)				
Recovery to well	63.9	51.1 - 76.7	Beta[b]	[9]
Deterioration to moderate-severe symptoms	5.3	4.2 - 6.4	Beta	[9]
Relative difference in effectiveness with internet-based versus in-person CBT, n	1.00	0.95 - 1.05	Triangular	[10]
No intervention for mild symptoms, event rate (%)				
Recovery to well	25.0	20.0 - 30.0	Beta	[11]
Deterioration to moderate-severe symptoms	9.1	7.3 - 10.9	Beta	[12]
Acceptance of CBT, event rate (%)				
In-person	63.30	60.1 - 66.5	Beta	[13]
Internet-based	75.3	71.5 - 79.1	Beta	[13]
Compliance to CBT, event rate (%)				
In-person	83.9	75.7 - 92.1	Beta	[14]
Internet-based	80.8	73.0 - 88.7	Beta	[14]
Monthly probabilities for Markov modeling				
Age-specific all-cause mortality, event rate (%)	0.0025	0.0020 - 0.0030	Beta	[15,16]
Relative risk of mortality with depressive symptoms, n	1.81	1.58 - 2.07	Triangular	[17]
Remission, event rate (%)	0.54	0.43 - 0.65	Beta	[18]
Relapse, event rate (%)	2.44	1.95 - 2.93	Beta	[18]
Proportion of mild symptoms among relapse (mild-moderate-severe), event rate (%)	24.3	19.2 - 28.8	Beta	[19]
Proportion of moderate symptoms among moderate-severe cases, event rate (%)	50	40 - 60	Beta	[19]
Psychiatric hospitalization among severe cases, event rate (%)	1.78	1.42 - 2.14	Beta	[20]
Dropout from university by mental health status, event rate (%)				[21]
Well/remission	0.124	0.099 - 0.149	Beta	
Depressive symptoms	0.193	0.154 - 0.232	Beta	
Utility inputs				
Age, years	21	20 - 22	Triangular	Cohort of local university students

XSL•FO
RenderX

Parameters	Base-case value	Range of sensitivity analysis	Distribution	Source of Information
Age-specific utility, n	0.92	-	-	[22]
Remission, n	0.80	0.72 - 0.88	Triangular	[23]
Mild depressive symptoms, n	0.62	0.56 - 0.68	Triangular	[23]
Moderate depressive symptoms, n	0.48	0.43 - 0.53	Triangular	[23]
Severe depressive symptoms, n	0.33	0.30 - 0.36	Triangular	[23,24]
Cost inputs				
Low-intensity CBT for depressive symptoms				[4]
Number of sessions	7	6 - 8	Triangular	
Duration of each session (min)	30	-	-	
High-intensity CBT for depressive symptoms				
Number of sessions	18	16 - 20	Triangular	[4]
Duration of each session (minutes)	60	-	-	
Psychotherapist service per hour (US $)	54	43 - 62	Triangular	Cohort of local university students
Platform overhead for internet-based CBT (US $)	24	20 - 29	Triangular	[25]
Percentage of time spent by psychotherapist in guided interned-based CBT (compared to in-person CBT)	9.23	7.38 - 11.1	Beta	[26]
Psychiatric outpatient (US $ per patient-month)	231	160 - 335	Gamma	Cohort of local university students
Psychiatric hospitalization (US $ per episode)	1500	900 - 2100	Gamma	[27,28]

[a]CBT: cognitive behavioral therapy.

[b]Parameter-specific distribution used in probabilistic sensitivity analysis.

The effectiveness of low-intensity CBT for mild depressive symptoms was retrieved from the findings of a prospective, pre- and post-treatment study of low-intensity CBT; they reported the recovery (63.9%) and deterioration (5.3%) rates for individual with mild common mental disorders (depressive disorders and anxiety disorders) in Hong Kong [9]. A systematic review (including 17 studies) reported that electronically delivered CBT was more effective than face-to-face CBT at lowering severity of depression symptoms and concluded that electronically delivered CBT was at least as effective as in-person CBT [10]. The relative difference in the effectiveness of CBT delivered via internet versus in-person was therefore assumed to be 1.00 in the present model. A recent Canada health technology assessment study on the cost-effectiveness of internet-delivered CBT for major depression had applied a spontaneous recovery rate of 25% for individuals who did not receive any treatment [11], and the present model adopted the spontaneous recovery rate for individuals with no intervention (who did not accept or were non-adherent to the CBT) in both study arms. The occurrence rate of deterioration in the individuals with no interventions was obtained from an individual participant data meta-analysis on internet-delivered CBT (including 16 trials with total 3805 participants), reporting that clinically significant deterioration had occurred in 9.1% participants in the control groups [12]. The preference of university students for digital technology versus in-person interventions when encountered mental health needs was examined in a survey study (n=572), and the likelihood of seeking help through online and in-person methods was 75.4%% and 63.3%, respectively [13]. The compliance to CBT was retrieved from the findings of a meta-analysis on adherence to internet-based (80.8%) and face-to-face (83.9%) CBT for depression (including 24 trials) [14].

The monthly probability of mortality in the subjects with depressive symptoms was approximated by all-cause mortality from age-specific mortality [15,16] and relative risk of mortality in individuals with depression (1.18) reported by a meta-analysis (including 25 studies and n=106,628) on excess mortality associated with depression [17]. The monthly probabilities of

XSL•FO
RenderX

relapse from well (2.44%) and remission from depressive symptoms (0.54%) after the completion of low-intensity CBT in both study arms were estimated from the findings of a long-term follow-up study of depression in primary care setting over 3 - 5 years (n=248; median 46 months, IQR 43-51 months) [18]. Of the relapsed cases from those who recovered well, the proportion of relapsed cases with mild symptoms (24.3%) and the proportion of relapsed cases with moderate symptoms (50%) among the relapsed cases with moderate-to-severe symptoms were retrieved from a retrospective cohort study (n=687) of the recurrence among patients with a history of major depressive disorder [19]. The monthly probability of psychiatric hospitalization (1.78%) among those with severe depressive symptoms was estimated from the 12-month psychiatric hospitalization rate reported by a retrospective database analysis (n=248) on hospital utilization in patients with major depressive disorder [20]. The monthly probabilities of school dropout among students with remission (0.124%) and those with depressive symptoms (0.193%) were approximated from the findings of a long-term follow-up study (for 4.8 years) on mental health and school dropout in over 3000 individuals aged 16 - 29 years [21].

Utility Inputs

The QALYs gained by each hypothetical student was calculated by the time spent in various health states, adjusting by the health state-specific utility (Table 1). The QALYs gained were discounted to the current year using a 3% annual discounting rate. A study reported the age-specific scores for quality of life from a US population-based survey [22], and we adopted the US age-specific score as the utility value in the present model. The utility values of remission and mild-to-severe symptoms of depression were adopted the corresponding values applied in cost-effectiveness analyses on the management of depression [23,24].

Cost Inputs

The cost analysis was conducted from the perspective of service providers in Hong Kong. The direct cost items (Table 1) were costs for in-person low-intensity CBT, internet-supported guided low-intensity CBT, in-person high-intensity CBT, and costs of psychiatric outpatient and inpatient treatment for depressive symptoms. The cost of in-person CBTs were estimated by the number of sessions (per intensity level) [4] and the local hourly wages of the psychotherapist. The overhead for internet-supported CBT in our model adopted the overhead cost of an online-guided CBT program for depression examined in a randomized controlled study [25]. The percentage of the psychotherapist's time (9.23%) for guided CBT versus in-person CBT was retrieved from the report of randomized controlled trials of internet-delivered versus in-person CBT for depression [26]. The cost per psychiatric hospitalization was approximated by the length of hospital stay for depression and cost per day of psychiatric hospitalization [27,28]. We conducted a retrospective review of psychiatric resource utilization in a cohort of local university students who received ambulatory psychiatric care for depressive symptoms. A cohort of students of the Chinese University of Hong Kong who had been referred by the University Health Services in 2017 - 2020 to outpatient psychiatry care for depression symptoms (n=100; 32% men, median age 21 [IQR 20 - 22] years) were reviewed on the psychiatric care reimbursement claims. The investigators had no access to information that could identify individual students during or after data collection. The median cost per patient month (US $231; 25th and 75th percentile US $160 and US $335) were applied as psychiatric outpatient costs (per patient month) for depressive symptoms. The costs were adjusted to the current year by an annual discounting rate of 3%.

Data Analysis

The expected cost and QALY of both study arms were generated in the base-case analysis. A CBT strategy was the dominant strategy if it gained higher QALYs at a lower cost than the comparator. When a strategy gained higher QALYs at a higher cost than the comparator, incremental cost per additional QALY gained (ICER) was calculated as follows: Δcost/ΔQALYs. The WHO recommended that ICER less than 1×gross domestic product (GDP) per capita to be highly cost-effective [29]. The 2023 GDP per capita of Hong Kong (US $48,119) was adopted as the willingness-to-pay (WTP) threshold (cost per additional QALY gained) in the present analysis [30]. A CBT strategy was accepted as cost-effective if (1) it was the dominant option (the strategy was less costly and more effective), or (2) the strategy was more costly and more effective, and its ICER versus the comparator was less than the WTP threshold.

The robustness of the base-case results was examined by sensitivity analysis. One-way sensitivity analysis was conducted for all model inputs. The range for sensitivity analysis of each model input was either the 95% CI range, high/low values of the variable, interquartile range, standard deviation, or (if above were not reported) varied by 20% of the base-case value. Probabilistic sensitivity analysis examined the impact of simultaneous variation of all model inputs. The cost and QALYs were generated for 10,000 times using the Monte Carlo simulation by randomly drawing each model input from the parameter-specific distribution (Table 1). The incremental cost and incremental QALYs of the internet group were presented in a scatter plot. All analyses were performed using TreeAge Pro 2023 (TreeAge Software Inc) and Excel 2016 (Microsoft Corporation).

Results

The internet group gained higher QALYs (by 0.0211 QALYs) and lowered dropout rate (by 0.0528%) at a reduced cost (US $249 saved) versus the in-person group in the base-case analysis (Table 2), and therefore was the (dominant) cost-effective option at ICER −11,801 US $/QALY.

XSL•FO
RenderX

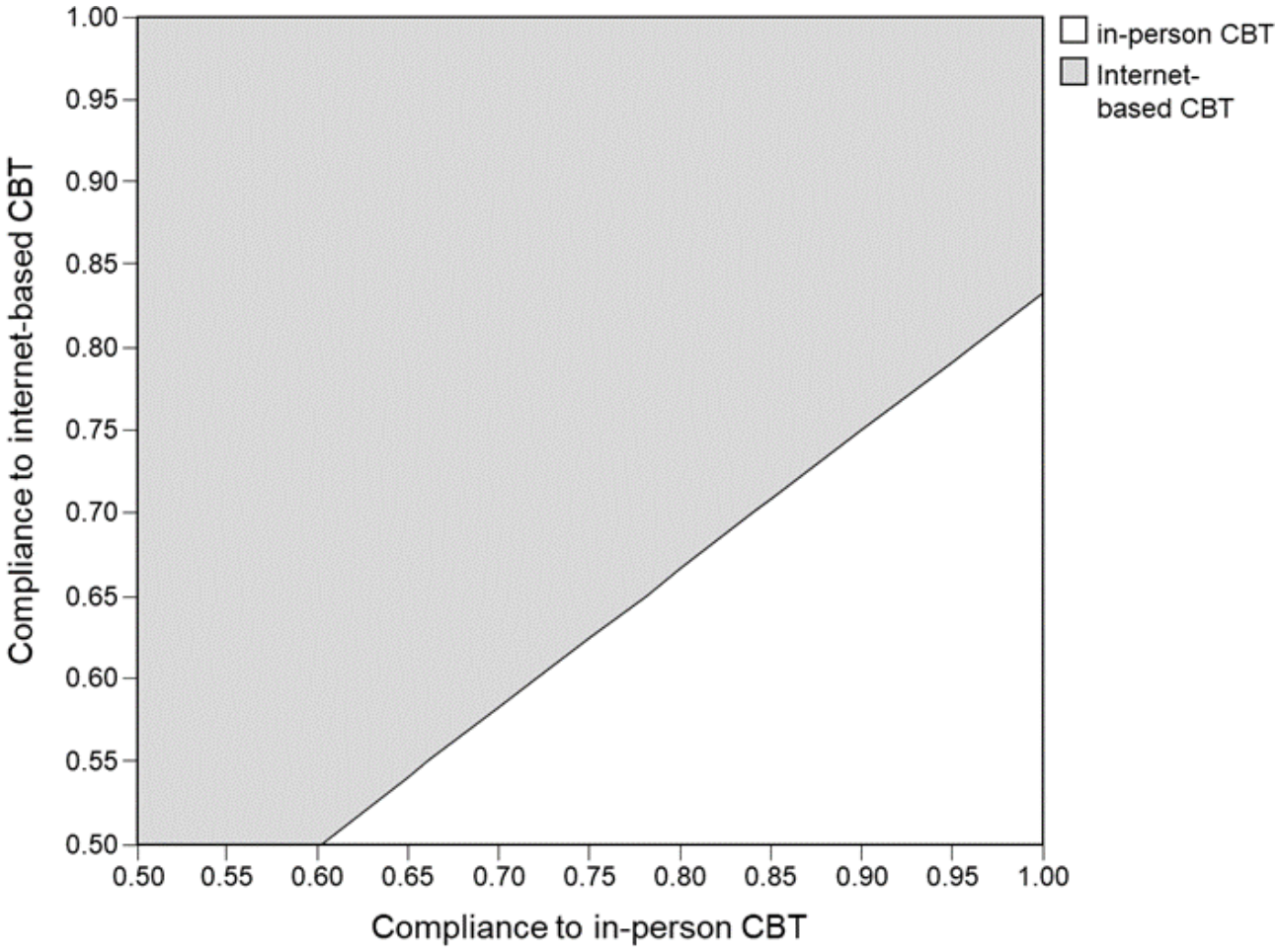
Table . Base-case analysis results.

	Costs (US $)	Incremental cost (US $)	School dropout	Incremental Dropout	QALYs[a]	Incremental QALYs
Internet group	8498	−249	9.8645%	−0.0528%	2.5627	0.0211
In-person group	8747	-	9.9174%		2.5416	-

[a]QALYs: quality-adjusted life-years.

In one-way sensitivity analysis, the internet-based group remained to gain higher QALYs at lower costs than the in-person group throughout the variation of all model inputs, and no threshold value was identified. The compliance rates to internet-based CBT and in-person CBT were the most influential factors on the ICER of internet-based CBT. When the compliance rate to internet-based CBT varied over the range of sensitivity analysis (from 73.0% to 88.7%), the ICER reduced from −10,116 USD/QALY to −23,892 USD/QALY. When the compliance rate to in-person CBT changed from 75.7% to 92.1% (range for sensitivity analysis), the ICER of internet-based CBT increased from −20,254 USD/QALY to −10,117 USD/QALY. To further explore the interacting impact of the compliance to the two delivery forms of CBT, a two-way sensitivity analysis was performed on extended range (50 - 100%) for both parameters. The findings of the two-way sensitivity analysis are shown in Figure 2. The gray area indicated the combination of compliance rates of internet-based and in-person CBT to prefer the internet-based CBT as the cost-effective option. The white area indicated the combinations of compliance rates leading to the in-person CBT as the cost-effective strategy.

Figure 2. Two-way sensitivity analysis of compliance to internet-based and in-person CBT. Gray zone: Combinations of parameters led to the internet group as the cost-effective option. White zone: Combinations of parameters led to the in-person group as the cost-effective option. CBT: cognitive behavioral therapy.

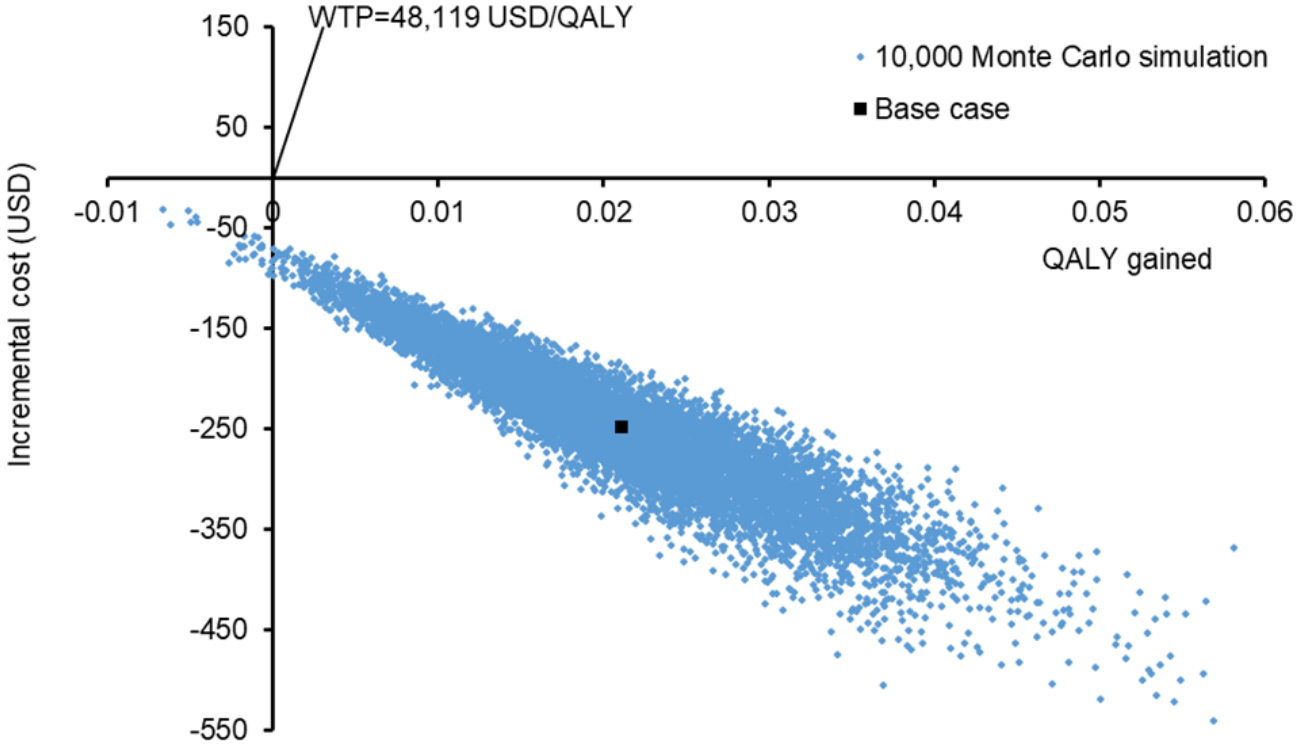


The internet platform cost and percentage of psychotherapist's time were previously reported to be major cost drivers of internet-based CBT [31,32]. The impact of these two parameters on the present base-case results were further evaluated by an extended sensitivity analysis, conducted by increasing the upper limit of internet platform cost from US $29 to US $2000, and the upper limit of psychotherapist's time from 11.1% to 50%. The extended sensitivity analysis found a threshold value for the internet platform cost at US $1745, and no threshold was identified for psychotherapist's time contribution.

XSL•FO
RenderX

The probabilistic sensitivity analysis recalculated the cost and QALYs for each study group 10,000 times using the Monte Carlo simulation. The incremental cost and additional QALYs gained by the internet group versus the in-person group in all Monte Carlo simulations are shown in Figure 3. The internet group was cost-saving in 100% of the 10,000 simulations, and gained higher QALYs 99.7% of the times (and therefore accepted as cost-effective in 99.7% of all simulations). The mean cost saved and additional QALY gained by the internet group were US $244 (95% CI US $242-246) and 0.0211 QALYs (95% CI 0.0209 - 0.0213), respectively. The internet group remained the cost-effective option when the WTP was lowered from US $48,119/QALY to 0.

Figure 3. Scatter plot of increment cost and QALY saved by internet-based CBT group versus in-person CBT group. CBT: cognitive behavioral therapy; QALY: quality-adjusted life-year; WTP: willingness-to-pay.



Discussion

Principal Results

This is the first study to report the cost-effectiveness of guided internet-based low-intensity CBT versus in-person CBT for mild depressive symptoms in university students. The present findings showed that the strategy of offering guided internet-based CBT gained higher QALYs and reduced school dropouts at a lower cost than the strategy of offering in-person CBT. The sensitivity analyses supported the robustness of these findings that the probability of the internet group to be accepted as cost-effective was 99.7% at the WTP threshold of US $48,119/QALY.

Limitations

This study was limited by the nature of model-based analysis that the model structure only included the key clinical treatment outcomes of depressive symptoms. The model inputs were retrieved from multiple clinical studies with heterogeneity in the study subjects. The findings from studies of CBT for depressive symptoms in Hong Kong or Chinese population do not fully meet the needs of the present model, and some of the clinical model inputs were therefore obtained from findings published in Western populations [10-14,17-21]. The sources of model inputs might weaken the generalizability of the study results to the university students in Hong Kong. Also, the compliance to internet-based CBT was retrieved from a meta-analysis of studies on CBT among depressed adults, including students and older participants [14]. University students are likely to have a higher affinity (vs older) for internet-based therapy due to higher digital literacy [33], and the compliance estimate (including both student and older samples) might therefore have underestimated the compliance in students, thus the cost-effectiveness benefits of internet-based CBT in students. Sensitivity analysis was therefore conducted on all model inputs to examine the impact of model parameter variation on the model findings. The cost-effectiveness findings generated by the present model were found to be highly robust to the variation of all model inputs. The psychiatric outpatient cost estimation from a local cohort of university students did not differentiate the severity of depressive symptoms and might underestimate the total direct cost related to severe depression. Indirect cost (loss of productivity) and the quality-of-life change associated with depression-related university drop-out were not considered in the present model, and might therefore undervalue health economic benefits generated by the guided internet-based low-intensity CBT.

Comparison With Prior Work

The individual compliance rates to the two CBT delivery modes were identified to be the most influential parameters in the

XSL•FO
RenderX

one-way sensitivity analysis. The base-case values of CBT compliance were adopted from a meta-analysis on 24 studies comparing the adherence to guided internet-based CBT with the adherence to face-to-face CBT. Even though the average adherence of face-to-face CBT (83.9%) was higher than that of internet-based CBT (80.8%), the difference between adherence rates did not achieve statistical significance (*P*=.59) [14]. The impact of uncertainty in the difference between the compliance rates on the cost-effectiveness of the two CBT delivery modes was further elaborated by a two-way sensitivity analysis on the compliance rates of internet-based and in-person CBT. The findings of two-way sensitivity analysis provided information on the combinations of compliance rates required for internet-based and in-person CBT to be cost-effective. In cohort with low compliance to in-person CBT (<60%), the internet-based CBT was a cost-effective option if the compliance to internet-based CBT was at least 50%. In individuals with high compliance to in-person CBT (>90%), the internet-based CBT needed to achieve a compliance rate of >75% in order to be accepted as the cost-effective alternative. The CBT delivered by the internet-based technology was previously reported by a 6-month decision model in Australia to be less costly and more effective than "treatment as usual" (using antidepressant medication) for an acute depressive episode. Similar to the present findings, the compliance rate of internet-supported CBT was also identified as the most influential factor on the cost-effectiveness of the technology-based CBT [34].

A life-long model-based health economic analysis demonstrated that an internet-based CBT was more effective in gaining higher QALYs at higher cost than "treatment as usual" for depression in Spain, and the high cost (approximately US $500) of technology platform and recurring cost (approximately US $3000 per individual) were the major cost drivers of internet-based CBT [32]. A 3-year model-based health economic analysis in Germany reported that the internet-based CBT gained higher QALYs at a lower cost than in-person CBT for unipolar depression. The Germany study included a low-cost (US $53) technology platform and low staff cost (psychotherapist-time approximately 40% of in-person CBT) [31]. Similarly, the internet group in the present model adopted a low cost for the internet platform to deliver the low-intensity CBT materials (US $24) [25] and approximately 10% psychotherapist's time of in-person CBT [26]. The cost-effectiveness of internet-based CBT was robust to high percentage (50%) of psychotherapist's time, and it remained cost-effective if the internet platform cost was less than US $1745, as indicated by the findings of extended sensitivity analysis in the present study.

University students belong to a vulnerable group of depressive symptoms. The latest report of WHO showed the statistics of World Mental Health Surveys International College Student Project (n=13,984) that the prevalence of depressive episode was the highest of all common mental disorders. The lifetime prevalence of major depressive episode was 21.2% and the 12-month prevalence was 18.5% among the surveyed students [1]. CBT is the recommended intervention for mild depressive symptoms, yet the acceptance rate of professional treatment among young adults had been low (<30%) due to embarrassment and preferring to handle the issue alone [3,35]. The internet-based CBT, as a NICE recommended intervention, has been showed to enhance acceptance by university students for emotional needs [13] and therefore could meet the needs of university students with mild depressive symptoms. Our prior study on the cost-effectiveness of internet-supported CBT for mild symptoms of anxiety for university students reported that the internet-supported CBT gained higher QALYs with cost-saving from the societal perspective of Hong Kong [36]. The findings of the present study further indicated that offering internet-based CBT (versus in-person CBT) for mild depressive symptoms was effective in gaining higher QALYs and lowering dropout rate with cost-saving. The improvement in QALYs achieved statistical significance (as indicated in 10,000 Monte Carlo simulations), yet the improvement (by 0.0211 QALYs) was modest. The health economic benefits of guided internet-based low-intensity CBT were mainly generated by the cost-saving and similar effectiveness to in-person CBT. The internet-based low-intensity CBT therefore is an alternative for university students with mild depressive symptoms for the policy makers to consider, especially when manpower resources are limited and penetration of internet-based programs is adequate. To transfer the internet-based CBT cost-effectiveness demonstrated in the present study to university students in Hong Kong, cultural adaptation of the internet-based CBT is essential to address differences (that may influence the effectiveness and acceptability) in communication styles, stigma, and social values in the context of locality [37]. The effectiveness of culturally adapted internet-based CBT for Hong Kong adults (18 - 70 years) with mild-to-moderate depressive symptoms has been recently demonstrated in a randomized clinical trial (n=402). The findings demonstrated a satisfactory recovery rate and supported that culturally and linguistically adapted internet-based CBT is an effective and feasible treatment for Hong Kong Chinese adults [38]. To further inform the decision-making process of resource allocation, future research to evaluate the culturally adapted internet-based CBT for university students with mild depressive symptoms together with feasibility study are highly warranted.

Conclusions

The guided internet-based low-intensity CBT was found to gain higher QALYs and reduce dropout rate at a lower cost, when compared to in-person CBT, for mild depressive symptoms in university students from the perspective of service providers in Hong Kong.

Acknowledgments

We thank Ms. Fiona Lam at the University Medical Service Office of the Chinese University of Hong Kong for supporting data retrieval.

XSL•FO
RenderX

Funding

This work was supported by the Health and Medical Research Fund (reference number 17180271), Food and Health Bureau, The Government of Hong Kong SAR.

Conflicts of Interest

None declared.

References

1. Auerbach RP, Mortier P, Bruffaerts R, et al. WHO World Mental Health Surveys International College Student Project: prevalence and distribution of mental disorders. J Abnorm Psychol 2018 Oct;127(7):623-638. [doi: 10.1037/abn0000362] [Medline: 30211576]
2. Deng J, Zhou F, Hou W, et al. The prevalence of depressive symptoms, anxiety symptoms and sleep disturbance in higher education students during the COVID-19 pandemic: a systematic review and meta-analysis. Psychiatry Res 2021 Jul;301:113863. [doi: 10.1016/j.psychres.2021.113863] [Medline: 33984824]
3. Ebert DD, Mortier P, Kaehlke F, et al. Barriers of mental health treatment utilization among first-year college students: First cross-national results from the WHO World Mental Health International College Student Initiative. Int J Methods Psychiatr Res 2019 Jun;28(2):e1782. [doi: 10.1002/mpr.1782] [Medline: 31069905]
4. Depression in adults: recognition and management (CG90). : National Institute for Health and Care Excellence, UK; 2009 URL: https://www.nice.org.uk/guidance/cg90 [accessed 2025-12-03]
5. Andrews G, Basu A, Cuijpers P, et al. Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: an updated meta-analysis. J Anxiety Disord 2018 Apr;55:70-78. [doi: 10.1016/j.janxdis.2018.01.001] [Medline: 29422409]
6. Karyotaki E, Riper H, Twisk J, et al. Efficacy of self-guided internet-based cognitive behavioral therapy in the treatment of depressive symptoms: a meta-analysis of individual participant data. JAMA Psychiatry 2017 Apr 1;74(4):351-359. [doi: 10.1001/jamapsychiatry.2017.0044] [Medline: 28241179]
7. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613. [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
8. Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. Med Decis Making 1993;13(4):322-338. [doi: 10.1177/0272989X9301300409] [Medline: 8246705]
9. Powell C, Lo APK, Yeung GTY, et al. A pilot study on the effectiveness of low-intensity cognitive behavioural therapy (LiCBT) for common mental disorders in Hong Kong. Behav Cogn Psychother 2021 Nov;49(6):758-763. [doi: 10.1017/S1352465820000971] [Medline: 33436141]
10. Luo C, Sanger N, Singhal N, et al. A comparison of electronically-delivered and face to face cognitive behavioural therapies in depressive disorders: a systematic review and meta-analysis. EClinicalMedicine 2020 Jul;24:100442. [doi: 10.1016/j.eclinm.2020.100442] [Medline: 32775969]
11. Health Quality Ontario. Internet-delivered cognitive behavioural therapy for major depression and anxiety disorders: a health technology assessment. Ont Health Technol Assess Ser 2019;19(6):1-199. [Medline: 30873251]
12. Karyotaki E, Kemmeren L, Riper H, et al. Is self-guided internet-based cognitive behavioural therapy (iCBT) harmful? An individual participant data meta-analysis. Psychol Med 2018 Nov;48(15):2456-2466. [doi: 10.1017/S0033291718000648] [Medline: 29540243]
13. Lungu A, Sun M. Time for a change: college students' preference for technology-mediated versus face-to-face help for emotional distress. Telemed J E Health 2016 Dec;22(12):991-1000. [doi: 10.1089/tmj.2015.0214] [Medline: 27223266]
14. van Ballegooijen W, Cuijpers P, van Straten A, et al. Adherence to Internet-based and face-to-face cognitive behavioural therapy for depression: a meta-analysis. PLoS ONE 2014;9(7):e100674. [doi: 10.1371/journal.pone.0100674] [Medline: 25029507]
15. Census and Statistics Department. The Government of Hong Kong Special Administrative Region Table E484: Hong Kong Life Tables. URL: https://www.censtatd.gov.hk/en/EIndexbySubject.html?scode=160&pcode=D5320184 [accessed 2024-01-16]
16. Census and Statistics Department. The Government of Hong Kong Special Administrative Region The Mortality Trend in Hong Kong. URL: https://www.censtatd.gov.hk/en/EIndexbySubject.html?pcode=FA100094&scode=160 [accessed 2024-01-16]
17. Cuijpers P, Smit F. Excess mortality in depression: a meta-analysis of community studies. J Affect Disord 2002 Dec;72(3):227-236. [doi: 10.1016/s0165-0327(01)00413-x] [Medline: 12450639]
18. Wiles NJ, Thomas L, Turner N, et al. Long-term effectiveness and cost-effectiveness of cognitive behavioural therapy as an adjunct to pharmacotherapy for treatment-resistant depression in primary care: follow-up of the CoBalT randomised controlled trial. Lancet Psychiatry 2016 Feb;3(2):137-144. [doi: 10.1016/S2215-0366(15)00495-2] [Medline: 26777773]

XSL•FO
RenderX

19. Hardeveld F, Spijker J, De Graaf R, Nolen WA, Beekman ATF. Recurrence of major depressive disorder and its predictors in the general population: results from the Netherlands Mental Health Survey and Incidence Study (NEMESIS). Psychol Med 2013 Jan;43(1):39-48. [doi: 10.1017/S0033291712002395] [Medline: 23111147]
20. Liu X, Tepper PG, Able SL. Adherence and persistence with duloxetine and hospital utilization in patients with major depressive disorder. Int Clin Psychopharmacol 2011 May;26(3):173-180. [doi: 10.1097/YIC.0b013e328343ba1e] [Medline: 21278576]
21. Hjorth CF, Bilgrav L, Frandsen LS, et al. Mental health and school dropout across educational levels and genders: a 4.8-year follow-up study. BMC Public Health 2016 Sep 15;16:976. [doi: 10.1186/s12889-016-3622-8] [Medline: 27627885]
22. Gold MR, Franks P, McCoy KI, Fryback DG. Toward consistency in cost-utility analyses: using national measures to create condition-specific values. Med Care 1998 Jun;36(6):778-792. [doi: 10.1097/00005650-199806000-00002] [Medline: 9630120]
23. Koeser L, Donisi V, Goldberg DP, McCrone P. Modelling the cost-effectiveness of pharmacotherapy compared with cognitive-behavioural therapy and combination therapy for the treatment of moderate to severe depression in the UK. Psychol Med 2015 Oct;45(14):3019-3031. [doi: 10.1017/S0033291715000951] [Medline: 26040631]
24. Pahlevan T, Ung C, Segal Z. Cost-utility analysis of mindfulness-based cognitive therapy versus antidepressant pharmacotherapy for prevention of depressive relapse in a canadian context: Analyse coût-utilité de la thérapie cognitive basée sur la pleine conscience contre la pharmacothérapie antidépressive pour prévenir la rechute de la dépression en contexte canadien. Can J Psychiatry 2020 Aug;65(8):568-576. [doi: 10.1177/0706743720904613] [Medline: 32031000]
25. Hollinghurst S, Peters TJ, Kaur S, Wiles N, Lewis G, Kessler D. Cost-effectiveness of therapist-delivered online cognitive-behavioural therapy for depression: randomised controlled trial. Br J Psychiatry 2010 Oct;197(4):297-304. [doi: 10.1192/bjp.bp.109.073080] [Medline: 20884953]
26. Andersson G, Hesser H, Veilord A, et al. Randomised controlled non-inferiority trial with 3-year follow-up of internet-delivered versus face-to-face group cognitive behavioural therapy for depression. J Affect Disord 2013 Dec;151(3):986-994. [doi: 10.1016/j.jad.2013.08.022] [Medline: 24035673]
27. Hospital authority. Fees and Charges. 2024. URL: https://www.ha.org.hk/visitor/ha_visitor_index.asp?Content_ID=10045&Lang=ENG [accessed 2024-03-18]
28. Citrome L, Jain R, Tung A, Landsman-Blumberg PB, Kramer K, Ali S. Prevalence, treatment patterns, and stay characteristics associated with hospitalizations for major depressive disorder. J Affect Disord 2019 Apr 15;249:378-384. [doi: 10.1016/j.jad.2019.01.044] [Medline: 30818246]
29. World Health Organization. The World Health Report 2002: Reducing Risks, Promoting Healthy Life: World Health Organization; 2002. URL: https://www.who.int/publications/i/item/9241562072 [accessed 2025-12-03]
30. Census and Statistics Department, The Government of Hong Kong Special Administrative Region. Gross Domestic Product (GDP), implicit price deflator of GDP and per capita GDP. URL: https://www.censtatd.gov.hk/en/web_table.html?id=31 [accessed 2024-03-15]
31. Baumann M, Stargardt T, Frey S. Cost-utility of internet-based cognitive behavioral therapy in unipolar depression: a Markov model simulation. Appl Health Econ Health Policy 2020 Aug;18(4):567-578. [doi: 10.1007/s40258-019-00551-x] [Medline: 32060822]
32. Piera-Jiménez J, Etzelmueller A, Kolovos S, Folkvord F, Lupiáñez-Villanueva F. Guided internet-based cognitive behavioral therapy for depression: implementation cost-effectiveness study. J Med Internet Res 2021 May 11;23(5):e27410. [doi: 10.2196/27410] [Medline: 33973857]
33. Montagni I, Tzourio C, Cousin T, Sagara JA, Bada-Alonzi J, Horgan A. Mental health-related digital use by university students: a systematic review. Telemed J E Health 2020 Feb;26(2):131-146. [doi: 10.1089/tmj.2018.0316] [Medline: 30888256]
34. Solomon D, Proudfoot J, Clarke J, Christensen H. e-CBT (myCompass), antidepressant medication, and face-to-face psychological treatment for depression in Australia: a cost-effectiveness comparison. J Med Internet Res 2015 Nov 11;17(11):e255. [doi: 10.2196/jmir.4207] [Medline: 26561555]
35. Lun KW, Chan CK, Ip PK, et al. Depression and anxiety among university students in Hong Kong. Hong Kong Med J 2018 Oct;24(5):466-472. [doi: 10.12809/hkmj176915] [Medline: 30245480]
36. You JHS, Luk SWC, Chow DYW, Jiang X, Mak ADP, Mak WWS. Cost-effectiveness of internet-supported cognitive behavioral therapy for university students with anxiety symptoms: a Markov-model analysis. PLoS ONE 2022;17(5):e0268061. [doi: 10.1371/journal.pone.0268061] [Medline: 35511888]
37. Spanhel K, Balci S, Feldhahn F, Bengel J, Baumeister H, Sander LB. Cultural adaptation of internet- and mobile-based interventions for mental disorders: a systematic review. NPJ Digit Med 2021 Aug 25;4(1):128. [doi: 10.1038/s41746-021-00498-1] [Medline: 34433875]
38. Pan JY, Rafi J. Culturally adapted guided internet-based cognitive behavioral therapy for Hong Kong people with depressive symptoms: randomized controlled trial. J Med Internet Res 2025 Feb 25;27:e64303. [doi: 10.2196/64303] [Medline: 39998865]

XSL•FO
RenderX

Abbreviations

CBT: cognitive behavioral therapy
GDP: gross domestic product
ICER: incremental cost per quality-adjusted life year
NICE: National Institute for Health and Care Excellence
PHQ-9: Patient Health Questionnaire-9
QALY: quality-adjusted life year
WHO: World Health Organization
WTP: willingness-to-pay

Edited by C Jacob; submitted 30.10.24; peer-reviewed by C Kacmarek, F Lynch, T Shimamoto; revised version received 19.11.25; accepted 24.11.25; published 15.12.25.

Please cite as:
You JHS, Luk SWC, Chow DYW, Mak ADP, Mak WWS
Internet-Based Mental Health Intervention for Depressive Symptoms in Young Adults: Cost-Effectiveness Analysis
JMIR Hum Factors 2025;12:e68167
URL: https://humanfactors.jmir.org/2025/1/e68167
doi:10.2196/68167

© Joyce H S You, Scotty W C Luk, Dilys Y W Chow, Arthur D P Mak, Winnie W S Mak. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 15.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Use of the ME-BYO Index, a Mobile Health App, During an Online Strength Training Program in Adults: Fidelity, Feasibility, and Acceptability Study

Yoshinobu Saito[1,2], PhD; Naoki Kikuchi[2,3], PhD; Kaname Watanabe[4,5,6], MD, PhD; Sho Nakamura[4,7], MD, PhD; Hiroto Narimatsu[4,5,6,7], MD, PhD

[1]Faculty of Sport Management, Nippon Sport Science University, 1221-1 Kamoshida-cho, Aoba-ku, Yokohama, Japan

[2]Graduate School of Physical Education, Health and Sport Studies, Nippon Sport Science University, Tokyo, Japan

[3]Faculty of Sport Science, Nippon Sport Science University, Tokyo, Japan

[4]Cancer Prevention and Control Division, Kanagawa Cancer Center Research Institute, Yokohama, Japan

[5]Center for Innovation Policy, Kanagawa University of Human Services, Kawasaki, Japan

[6]Department of Genetic Medicine, Kanagawa Cancer Center, Yokohama, Japan

[7]Graduate School of Health Innovation, Kanagawa University of Human Services, Kawasaki, Japan

Corresponding Author:
Yoshinobu Saito, PhD
Faculty of Sport Management, Nippon Sport Science University, 1221-1 Kamoshida-cho, Aoba-ku, Yokohama, Japan

Abstract

Background: Although various apps have been developed to support health behaviors, they are mostly commercial, possibly limiting the number of users. The ME-BYO index was developed by Kanagawa Prefecture in 2019 to comprehensively and numerically measure and visualize an individual's current health status and future disease risk by quantifying data. The ME-BYO index is free of charge, so it can be made available to as many prefectural residents as possible for health promotion. Effective online strength training programs are being developed that, when combined with ME-BYO index measurements, will help with both exercise habits and health management.

Objective: In this study, we aimed to validate the fidelity, feasibility, and acceptability of self-measurement using the ME-BYO index during the implementation of an online strength training program.

Methods: Participants were 23 adults aged 40 years or older who did not regularly perform muscle strengthening exercises. The strength training program was performed twice a week for 8 weeks (16 sessions in total), and the ME-BYO index was explained to the participants so that they could self-measure the ME-BYO index with a smartphone on the day of the strength training program, before its implementation. The ME-BYO index during the study period was continuously collected from the app, and the ME-BYO index adherence rate was calculated. Questionnaires were used to assess the feasibility (difficulty of measurement and motivation to improve lifestyle) and acceptability (intention to maintain measurement and appropriate frequency of measurement) of implementing and continuing the ME-BYO index measurements. Changes in the ME-BYO index between the first and second halves of the program period, examination of items strongly related to changes in the ME-BYO index, and a comparison of physical fitness and health outcomes before and after the program period were conducted.

Results: The mean ME-BYO index adherence rate during the strength training program was 89.4% (SD 17%). Regarding acceptability, the participants were highly motivated to continue measuring the ME-BYO index (77%), and the appropriate frequency of measurement was once a week and twice a week (31% and 31%, respectively). Panel data analysis of the self-measured ME-BYO index showed no significant change in the ME-BYO index score; however, it increased to a higher score. Examination of the items that increased the overall score indicated that systolic blood pressure, mental resilience, and Mini-Cog scores were the relevant factors. The pre- and postprogram measurements showed no significant changes in items other than physical fitness.

Conclusions: The fidelity, feasibility, and acceptability of measuring the ME-BYO index during a regular online strength training program were high, suggesting that self-measurement of the ME-BYO index could be used to implement and maintain healthy behaviors. These findings suggest that the ME-BYO index can be recommended as a basic health app.

(JMIR Hum Factors 2025;12:e63123) doi:10.2196/63123

KEYWORDS

digital health; resistance training; physical fitness; health outcomes; intrinsic capacity; self-measurement; smartphone; surveys and questionnaires; feasibility studies; implementation; Japan

XSL•FO
RenderX

Introduction

Appropriate health behaviors are important for maintaining and improving health. In recent years, there have been remarkable developments in mobile health to support healthy behaviors, with various apps being developed [1]. Reviews on the effectiveness of these apps have examined healthy behaviors such as changes in physical activity [2,3], dietary management [4], smoking cessation [5], weight management [6], diabetes management [7], hypertension management [8-10], and medication adherence [11,12], as well as their combined effects with other interventions. However, most of these apps have been commercially developed. Users of smartphone health apps have high health literacy and financial affordability [13,14], and these factors may limit the available population. A review of apps for mental health among underserved young people found that, while acceptance was high among economically disadvantaged residents (with positive evaluations and high satisfaction), they did not intend to continue using the apps. This population was also reluctant to pay for the apps [15].

Furthermore, while commercially available apps have high download numbers and receive positive evaluations regarding ease of use and engagement, significant variations in quality have been reported. Many of these apps are not evidence-based, are inconsistent with public health guidelines, lack empirical evidence of effectiveness or validation of measurement tools, lack theoretical grounding, lack expert and user involvement, and fail to provide assurances around privacy, security, or health risks [16]. Therefore, it is valuable for local governments to develop apps that address these challenges and that all residents can use with confidence.

Kanagawa Prefecture, one of Japan's largest local governments, defines "ME-BYO" as a state of continuous change between health and disease, rather than a clear distinction between health and disease, and promotes efforts to improve presymptomatic conditions based on three pillars: diet, exercise, and social participation. In March 2020, the prefecture's health management ME-BYO index, which allows individuals to measure their state of ME-BYO, was released on the smartphone app "My ME-BYO record" [17]. The ME-BYO index is a simple and comprehensive index that can be calculated by entering and measuring 15 items across the areas of metabolic function, locomotor function, cognitive function, and mental resilience. It is expected to promote behavioral changes in individual health and the construction of new systems in society through the concrete quantification of health status [18].

It is unique in that it includes risk factors for dementia (BMI, blood pressure, and mental health) and the Mini-Cog for screening cognitive function. To date, a new longitudinal study based on the Kanagawa Prospective ME-BYO Cohort Study has been initiated to verify the validity and reliability of the measurement method [19,20] and incorporate a future prediction function. However, the real-world feasibility of the ME-BYO index has not yet been validated.

The ME-BYO index was developed for health promotion among residents. This app is provided free of charge to enable widespread use. Therefore, disseminating the app and verifying its feasibility are crucial for addressing health disparities.

Epidemiological studies on physical activity have recently focused on strength training, low-intensity physical activity, and sedentary behaviors. Among these, meta-analyses have shown that strength training reduces the risk of all-cause mortality and development of cardiovascular diseases, cancer, and diabetes [21]. The COVID-19 pandemic has led to the development of various online exercise programs. In particular, online strength training programs, which are convenient and can be completed relatively quickly, are as effective as in-person training and a promising approach for busy working populations [22]. Therefore, combining the measurement of the ME-BYO index with online strength training programs may contribute to motivation to practice exercise and health management.

Therefore, in this study, we aimed to examine the fidelity, feasibility, and acceptability of the ME-BYO index during the implementation of an online strength training program, which is a promising approach for exercise continuity in the working-age population.

Methods

Study Design and Participants

This was a prospective single-arm study. The participants were 23 adults aged 40 years or older (41 - 68 y) who did not regularly perform muscle strengthening exercises at least twice per week. Information on muscle strengthening exercise habits was collected through self-report.

Ethical Considerations

This study was approved by the Ethics Committee of the Nippon Sports Science University (023-H043) and by the Ethics Committee of Kanagawa Cancer Center (2023KEN-10). The purpose and content of the study were explained to the participants, and written informed consent was obtained before the study was conducted. Data were collected in a face-to-face setting, and all user data were deidentified before analysis.

Strength Training Programs

The strength training program was performed twice weekly for 8 weeks (16 sessions in total). Strength training consisted of six exercises (squats, rear raises, dips, lunges, single-leg Romanian deadlifts, and push-ups; Multimedia Appendix 1) and lasted approximately 60 minutes per session. The load was applied with body weight or Therabands, and the number of repetitions per set was 10 to 20, with 1 to 3 sets (low intensity) titrated during the training period. The rest period between sets was 2 minutes. A video call app (ZOOM) was used, and the instructor-to-subject ratio was set at 1 to 5 - 7. The strength training program was delivered by well-trained staff [22,23]. The staff members measured the following before and after the strength training program: the ME-BYO index, cognitive function, well-being, psychological distress, physical fitness, body composition, and physical activity.

XSL•FO
RenderX

Measurement Items

The ME-BYO Index

The ME-BYO index was measured by the staff before and after the implementation of the strength training program. During the study period, self-measured ME-BYO index was continuously obtained through the My ME-BYO record app, which was created by Kanagawa Prefecture. The ME-BYO index was measured after a thorough explanation was provided during the preprogram assessments. Instructions for self-measurement were distributed and explained. For this study, data were anonymized, and only information from the study participants was collected.

The ME-BYO index is scored based on a 15-item measure across the following four domains: metabolic function, locomotor function, cognitive function, and mental resilience [18,24]. The World Health Organization categorizes intrinsic capacity into five categories: physical mobility, vitality, psychosocial, sensory, and cognitive [25]. The ME-BYO index was designed using this approach and calculated by summing the weighted partial scores that assessed the status of the four domains. The overall score ranges from 0 to 100 points, with higher scores indicating better health. Various scientific indicators were collected using a survey to construct a scoring formula for domain-specific assessments. The following measurement indices were used to calculate the ME-BYO index: sex, age, BMI, and systolic blood pressure to evaluate metabolic function; the 5-question Geriatric Locomotive Function Scale questionnaire [26] and walking speed measured by a smartphone [19] to assess locomotor function; the 3-question version of the Mini-Cog [20,27,28] to evaluate cognitive function; and individual voice information to evaluate mental resilience. A mind monitoring system (MIMOSYS; PST Inc) was recently developed [29-31] and implemented to evaluate mental resilience using smartphone apps [18].

My ME-BYO record can be installed and used free of charge in the App Store and Google Play. My ME-BYO record and ME-BYO index Terms of Use, Privacy Policy, and Security Policy are available within the app and on the Kanagawa Prefectural Government website [32-34].

Fidelity, Feasibility, and Acceptability of the ME-BYO Index Measurement

This study was conducted based on Proctor et al.'s [35] implementation outcome framework (IOF). We examined the individual-level implementation outcomes outlined in the IOF: fidelity, feasibility, and acceptability.

The ME-BYO index self-measurement frequency and adherence rate were calculated as fidelity. The adherence rate was calculated individually as the percentage of the self-measurement to the participation sessions throughout the intervention period, and the average adherence rate was obtained.

In feasibility, they were evaluated using a questionnaire regarding the difficulty of measuring the ME-BYO index twice a week (quite difficult, somewhat difficult, not very difficult, or not at all difficult) and their motivation to improve their lifestyle through regular ME-BYO index measurements (not at all motivated, not very motivated, somewhat motivated, or quite motivated).

The acceptability of the ME-BYO index was evaluated using a questionnaire with the following two items: (1) intention to maintain regular ME-BYO index measurements after the end of the study (not at all, not very much, somewhat agree, or strongly agree); and (2) frequency of ME-BYO index measurement appropriate for the participant (once a year, once every few months, once a month, twice a month, once a week, or at least twice a week). The questionnaire on feasibility and acceptability is detailed in Multimedia Appendix 2.

Fidelity and Acceptability of Strength Training Programs

Fidelity was assessed regarding the time required for participation and the adherence rate in the strength training program. Acceptability was evaluated using the following questionnaire on technical difficulty and time allocation of the strength training program: For technical difficulty, participants were asked to choose from the following options: quite difficult, somewhat difficult, not very difficult, or not at all difficult. Time allocation was selected from the following options: not at all appropriate, not very appropriate, somewhat appropriate, and quite appropriate.

Cognitive Function

We used the Japanese version of the mild cognitive impairment (MCI) screen, a simple cognitive function scale developed by the Medical Care Corporation, Inc., in the United States, to detect MCI early [36]. The assessment was based on the Memory Performance Index (MPI) score. The MPI score is a unique index that objectively expresses the health status of the brain and is evaluated and calculated using the statistical analysis algorithm of the MCI screen. The higher the MPI score, the higher the patient's cognitive health status, expressed quantitatively as a number from 0 to 100 [37].

Well-Being and Psychological Distress

The World Health Organization-5 Well-Being Index was used to measure the subjective psychological well-being [38]. The respondents answered five statements regarding their feelings over the past 2 weeks. The total score ranged from 0 to 25, with higher scores indicating higher psychological well-being. Scores <13 indicate poor well-being and suggest depression according to the International Statistical Classification criteria [39]. To assess the quality of life, we multiplied the raw score and calculated the percentage ranging from 0 to 100 (low to high quality of life).

The Psychological Distress Scale (K6) score, a robust nonspecific psychological distress measurement tool [40,41], was calculated from six items using a 5-point Likert scale, with the total score ranging from 0 to 24, with higher scores indicating more severe distress. We used the Japanese version of the scale [41], translated and validated from the original scale developed in English [40].

Physical Fitness and Body Composition

Grip strength was assessed using a hand grip dynamometer (Takei Scientific Instruments). Participants were allowed two

XSL•FO
RenderX

attempts to attain the highest possible rating. They took a standing position and squeezed the dynamometer as hard as possible with one hand, and then repeated the test with the other hand.

A push-up test was performed to evaluate upper extremity muscle function. Participants were required to touch a 10 cm foam pad placed on the floor with their knees on top. The participants were then asked to alternately perform push-ups repeatedly for 30 s, and the number of times they could come to the start position was counted and recorded.

The 5-times chair-standing test required the participants to stand up from a chair without using arms and sit down as quickly as possible. The time required for participants to perform this action was measured 5 times using a stopwatch [42].

Squat and countermovement jumps were measured using a MULTI JUMP TESTER instrument (DKH Co., Ltd). All the participants were instructed to jump as high as possible and were not allowed to use their arms. The highest score of the two trials was accepted as the jump score (cm).

The sit-and-reach flexibility test is widely used as a general test of trunk flexibility and is performed using a digital measuring device (Takei Scientific Instruments). The participants were instructed to sit on the floor and stretch their torso and arms out in front of them with their knees straight. The participants were allowed 2 to 3 attempts to attain the highest possible rating.

Body composition was measured in terms of muscle thickness and body fat content. The muscle thickness (cm) of the triceps brachii (ie, 60% distal between the lateral epicondyle of the humerus and the acromial process of the scapula) was measured using B-mode ultrasound (FAMUBO-D; SEIKOSHA). Body fat (%) was measured using a body composition measuring device (Inbody270; InBody Japan Inc). Height (seca 213; seca Nihon) and weight (HD-665; TANITA corporation) were measured, and BMI (kg/m^2) was calculated. Systolic blood pressure and diastolic blood pressure (mmHg) were measured using a sphygmomanometer (HEM-7120; OMRON Healthcare Co. Ltd). Participants rested for at least 5 minutes in a seated position before the measurement was taken. They took two readings themselves. If there were no issues, the second reading was used.

Physical Activity

Physical activity was measured using a validated accelerometer (Active Style Pro HJA-750C; OMRON Healthcare) [43-45]. Participants were instructed to wear the accelerometer on their waist for 7 consecutive days from when they walked up to when they went to bed, except when bathing, swimming, or other water exposure. The time the accelerometer was worn was calculated by subtracting the time the accelerometer was not worn from 24 h per day. Those with at least 4 days of at least 10 hours of wear per day were included. The outcome measures were average moderate-to-vigorous physical activity time (minutes per day, ≥3.0 METs) and average sedentary time (minutes per day, ≤1.5 METs). A macro program (version 190829), developed and distributed by the Japan Physical Activity Research Platform [46], was used to process the accelerometer data.

Statistical Analysis

For the panel data analysis of the ME-BYO index self-measured, the means of the first 4 weeks of the program and the second 4 weeks of the program were compared for the overall score, metabolic function score, locomotor function score, cognitive function score, and mental resilience score using an appropriate *t* test. A multilevel analysis was also performed, with the ME-BYO index overall score as the dependent variable; BMI, systolic blood pressure, mental resilience score, Mini-Cog score, 5-question geriatric locomotive function scale, and walking speed as the independent variables; and sex and age as the adjusted variables. Paired *t* tests were used to compare the means of each item in the pre- and postprogram measurements. Statistical analyses were performed using SPSS version 27 software (IBM Corp). The significance level was set at 5%.

Results

Of the 23 participants, 19 participated in the premeasurement period. Therefore, the 19 participants for whom pre- and postassessments were available were included in the analysis. The baseline characteristics of the study participants are shown in Table 1. The mean age was 57.6 (SD 7.8) years.

The results of the fidelity, feasibility, and acceptability of the ME-BYO index measurement and strength training program are shown in Table 2. Regarding the number of self-measurements and the adherence rate of the ME-BYO index, 13 participants were included in the evaluation because 8 participants had not yet taken the measurement owing to inadequate smartphones or measurement environments. The mean number of measurements was 14.1 (SD 3.1), and the ME-BYO index adherence rate was 89.4% (17%). Regarding self-measurement of the ME-BYO index twice a week during the study period, a few participants had difficulty measuring it (n=5, 39%), and many participants were motivated to improve their lifestyle (n=8, 61%). The intention to maintain the ME-BYO index twice a week was high (n=10, 77%), and the most common appropriate frequency for measuring the ME-BYO index was once or twice a week (n=8, 62%). The adherence rate in the strength training program was 94% (8.5%). The technical difficulty was somewhat difficult (n=9, 47%) but not very difficult (n=7, 37%). Time allocation was appropriate (n=16, 84%).

XSL•FO
RenderX

Table . Characteristics of the study participants (n=19).

Characteristics	Value, n (%)
Age (y)	
41 - 49	3 (15.8)
50 - 59	7 (36.8)
60 - 68	9 (47.4)
Sex	
Male	8 (42.1)
Female	11 (57.9)
Education (y)	
12 - 14	3 (15.8)
16	8 (42.1)
18+	8 (42.1)

XSL•FO
RenderX

Table . Fidelity, feasibility, and acceptability of the self-measured ME-BYO index and the strength training program.

Item	Value
Fidelity	
Self-measurement of ME-BYO index, mean (SD)	
Number of measurements (times)	14.1 (3.1)
Adherence rate (%)	89.4 (17.0)
Strength training program, mean (SD)	
Participation (times)	15.1 (1.4)
Adherence rate (%)	94.0 (8.5)
Feasibility	
Difficulty of implementing twice-weekly ME-BYO index measurements, n (%)	
Quite difficult	1 (8)
Somewhat difficult	4 (31)
Not very difficult	2 (15)
Not at all difficult	6 (46)
Motivation to improve lifestyle by ME-BYO index measurement, n (%)	
Not at all motivated	1 (8)
Not very motivated	4 (31)
Somewhat motivated	5 (38)
Quite motivated	3 (23)
Acceptability	
Intention to maintain ME-BYO index measurement, n (%)	
Not at all	2 (15)
Not very much	1 (8)
Somewhat agree	8 (62)
Strongly agree	2 (15)
Appropriate frequency of ME-BYO index measurement, n (%)	
Once a year	0
Once every few months	2 (15)
Once a month	1 (8)
Twice a month	2 (15)
Once a week	4 (31)
At least twice a week	4 (31)
Technical difficulty of the strength training program, n (%)	
Quite difficult	0
Somewhat difficult	9 (47)
Not very difficult	7 (37)
Not at all difficult	3 (16)
Time allocation fo ther strength training program, n (%)	
Not at all appropriate	0
Not very appropriate	0
Somewhat appropriate	3 (16)
Quite appropriate	16 (84)

XSL•FO
RenderX

When comparing the mean scores of the ME-BYO index self-measurement between the first and second 4 weeks, there was no significant difference, but there was a change to a higher value (Table 3).

A multilevel analysis using panel data in the self-measured ME-BYO index to examine the measures that increased the overall score revealed that systolic blood pressure, mental resilience, and the Mini-Cog score were significantly associated factors. An increase of 1 mmHg in systolic blood pressure tended to decrease the overall score by 0.15 points. A 1-point increase in mental resilience tended to increase the overall score by 0.10 points. A 1-point increase in Mini-Cog score tended to increase the overall score by 3.57 points (Table 4).

Significant improvements in physical fitness were observed in the postprogram measurements compared with the preprogram measurements, including push-ups, 5-time chair standing test, squat jump, and countermovement jump. No significant changes were observed in the other items (Table 5).

Table . Changes in the ME-BYO index self-measurement scores[a].

Item	First half of the program, mean (SD)	Second half of the program, mean (SD)	*P* value
Overall score (points)	91.6 (4.2)	92.9 (4.1)	.15
Metabolic function score (points)	90.2 (8.1)	90.2 (8.2)	.99
Locomotor function score (points)	97.3 (7.9)	97.2 (6.6)	.91
Cognitive function score (points)	96.5 (8.0)	98.7 (2.2)	.22
Mental resilience score (points)	70.3 (19.7)	78.2 (11.2)	.08

[a]Paired *t*-test: panel data on ME-BYO index self-measurement to compare mean scores for the first 4 weeks and second 4 weeks of the program.

Table . Measures associated with an increase in the overall score of the ME-BYO index[a].

	Estimated value	SE	*P* value	95% CI
BMI (kg/m^2)	−0.25	0.31	.43	−0.95 to 0.44
Systolic blood pressure (mmHg)	−0.15	0.02	<.001	−0.19 to −0.12
Mental resilience score (points)	0.10	0.01	<.001	0.08 to 0.12
Mini-Cog score (points)	3.57	0.71	<.001	2.15 to 5.00
Five-question geriatric locomotive function scale (points)	−1.54	1.18	.20	−3.91 to 0.83
Walking speed (m/s)	−0.01	0.07	.87	−0.16 to 0.14
Sex: women (reference: men)	−1.68	3.23	.62	−9.24 to 5.88
Age (years)	−0.04	0.21	.85	−0.54 to 0.46
Intercept	95.45	12.60	<.001	66.99 to 123.91

[a]Multilevel analysis: Using panel data in the self-measurement of the ME-BYO index, the overall score was the dependent variable; BMI, systolic blood pressure, mental resilience score, Mini-Cog score, 5-question geriatric locomotive function scale, and walking speed were the independent variables; and sex and age were the adjusted variables.

XSL•FO
RenderX

Table . Changes in each evaluation item before and after the online strength training program[c].

	Before, mean (SD)	After, mean (SD)	*P* value
ME-BYO index (score)	85.4 (7.7)	84.3 (6.2)	.54
BMI (kg/m^2)	23.6 (4.3)	23.5 (4.2)	.44
Systolic blood pressure (mmHg)	124.4 (21.9)	124.4 (17.9)	.99
Diastolic blood pressure (mmHg)	77.5 (10.5)	81.0 (11.1)	.20
MPI[a] (points)	69.1 (8.2)	69.6 (8.1)	.65
WHO-5[d] (points)	18.9 (3.9)	18.7 (3.9)	.81
K6 (points)	7.2 (1.9)	7.6 (2.3)	.19
Grip strength			
Right (kg)	32.3 (10.4)	31.6 (10.4)	.05
Left (kg)	30.3 (11.1)	30.6 (10.9)	.69
Push-up (times/30 s)	15.2 (6.4)	19.2 (6.1)	<.001
5 times chair standing test (sec)	6.4 (1.4)	5.7 (0.6)	.01
Squat jump (cm)	19.8 (5.3)	21.1 (5.6)	.02
Countermovement jump (cm)	22.8 (5.9)	23.9 (6.1)	.045
Sit-and-reach test (cm)	38.1 (7.9)	38.1 (8.6)	.94
Muscle thickness (cm)	3.3 (0.5)	3.3 (0.9)	.77
Body fat (%)	27.1 (6.6)	26.0 (7.7)	.07
MVPA[b] (min/d)	83.0 (41)	75.7 (39)	.29
Sedentary time (min/d)	542.3 (132.1)	526.3 (125.1)	.33

[a]MPI: Memory Performance Index.

[b]MVPA: moderate-to-vigorous-intensity physical activity.

[c]Paired *t*-test.

[d]WHO-5: The World Health Organization-5 Well-Being Index.

Discussion

Principal Findings

In this study, we evaluated the fidelity, feasibility, and acceptability of self-measuring the ME-BYO Index during an online strength training program. The mean number of measurements was 14.1 (SD 3.1), and the adherence rate was 89.4% (17%), suggesting high fidelity in an 8-week (16 sessions), twice-weekly training program. Although 5 (39%) participants had difficulty measuring the ME-BYO index, 8 (61%) participants were motivated to improve their lifestyle. The importance of daily physical condition and health management was described in the individual questionnaire (Multimedia Appendix 3), suggesting that these factors led to a high number of ME-BYO index measurements. A study on the fidelity of a smartphone nutrition app (MyNutriCart) for making smart and healthy choices when purchasing food in grocery stores found that the average use of the app was 75% for each purchase, with a 69% likelihood of use [47]. This was an 8-week intervention in which participants were instructed to use the app every time they went grocery shopping or at least once a week [48]. Reviews of self-monitoring mental health apps also report adherence rates of 85% to 100%, suggesting high engagement with these apps [49]. Despite differences in input items and duration, this study demonstrated fidelity similar to that of other studies.

Eight participants were unable to self-measure, not due to a lack of interest or motivation but primarily because their smartphones were incompatible. Furthermore, all their baseline characteristics were comparable (Multimedia Appendix 4). Therefore, it is unlikely that this resulted in a biased sample. However, it should be noted that the ME-BYO index measurement was requested on the day of program participation, which may have motivated participants to continue the measurement. Additionally, smartphone literacy may have influenced the results. Concerns remain, particularly among older adults, that limited familiarity with technology may affect outcomes. This highlights the need for further research.

High fidelity was also confirmed for the strength training program (94%), suggesting that the program may be an effective

XSL•FO
RenderX

gateway to exercise habits in the working population. Pre- and postprogram measurements revealed an increase in physical fitness. However, no changes were observed in the mental or psychological factors assessed besides the ME-BYO index. These results suggest that, although the evaluation of the ME-BYO index measures not linked to the program is problematic, combining the program with a regular and highly feasible program is effective for regular ME-BYO index measurements. The intention to continue measuring the ME-BYO index twice a week was high (n=10, 77%), and the appropriate frequency of the ME-BYO index measurement was once or twice a week (n=8, 62%), suggesting that continuous ME-BYO index measurement at short intervals could be used as an indicator of the effectiveness of regular exercise and other programs. A prior study using real-world data from the My ME-BYO Record app reported low engagement among repeat users [50]. In contrast, the present study observed high adherence rates when the app was integrated within a structured online strength training program. These findings suggest that combining the app with guided health interventions may be an effective strategy to enhance user engagement and address the low adherence typically observed in free-living settings compared to structured programs.

When comparing the mean ME-BYO index self-measured scores between the first and second 4 weeks, there was a general change to higher scores, although there were no significant differences. Notably, the mental resilience score showed an upward trend, while the scores of the other indicators remained high. The multilevel analysis showed that the measures that increased overall scores were systolic blood pressure, mental resilience, and Mini-Cog scores. The ME-BYO index app has an immediate feedback and advice function for measurement results, and the ability to self-monitor is thought to be a factor that increases its fidelity, feasibility, and acceptability. Self-monitoring is one of the most commonly used behavioral change techniques in health apps [51]. A meta-analysis by Michie et al [52] on behavior change techniques (BCTs) that contributed to improved physical activity and dietary habits found that the most effective BCT among 122 intervention studies was immediate self-monitoring of behavior, indicating the importance of implementing BCTs. This study demonstrated the importance of implementing BCTs.

In a study evaluating the quality of physical activity apps, only 1 out of 65 randomly selected apps that met the criteria had a noncommercial affiliation (government agency). Privacy and security assessments revealed that 29.2% of the apps had no privacy policy. Most apps collect personally identifiable information and share user data with third parties, and more than half do not specify how to ensure data security [53]. Furthermore, only 12 of the 65 apps had a peer-reviewed study associated with them, and only one app was evaluated for efficacy in a trial. This suggests that the quality of commercially developed apps is inadequate. Importantly, this study confirmed the feasibility of a reliable app provided free of charge by the local government, which can be recommended for use as a basic health app.

Limitations

This study had several limitations. First, it used a before-and-after comparative design, and the influence of psychological bias (the Hawthorne effect) and selection bias cannot be excluded. Second, the sample size was small, which may have resulted in insufficient statistical power. Future studies involving control groups are required. Third, although the evaluation was based on the IOF, the feasibility and acceptability indicators were not validated scales, which may have introduced interpretation bias. Fourth, the study was conducted over a short period. The ME-BYO index includes indices (cognitive and locomotor functions) that do not change easily over a short period. Therefore, it may be difficult to detect the effects of short-term measurements at short intervals, and long-term validation needs to be addressed. The impact on long-term fidelity should also be considered.

Conclusion

We investigated the fidelity, feasibility, and acceptability of self-measurement of the ME-BYO index during the implementation of an online strength training program in adults with no exercise habits. The fidelity, feasibility, and acceptability of measuring the ME-BYO index during implementation were high, suggesting that self-measurement of the ME-BYO index could be used as a tool for implementing and maintaining healthy behaviors. The fact that these findings were confirmed in an app provided free of charge by the public sector is significant and suggests that it can be recommended as a basic health app.

Acknowledgments

The authors thank all participants in this study. The authors thank students from the Graduate School of Physical Education, Health and Sport Studies, Nippon Sport Science University for their assistance with data collection and program implementation. We thank Editage [54] for English language editing.

Funding

This study was supported by the ME-BYO index project of the Kanagawa Prefecture and grants-in-aid for the Japan Society for the Promotion of Science KAKENHI grant (16H06277 [CoBiA]) from the Japanese Ministry of Education, Culture, Sports, Science, and Technology. This study was also supported by the Science and Technology Innovation co-creation center promotion project of the Kanagawa Institute of Industrial Science and Technology.

XSL•FO
RenderX

Authors' Contributions

Conceptualization: YS, NK
Methodology: YS, NK
Formal analysis: YS
Data curation: YS, NK, KW, SN, HN
Writing – original draft: YS
Writing – review & editing: YS, NK, KW, SN, HN
Project administration: YS, NK
Funding acquisition: YS, HN
Supervision: HN
All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Picture of the strength training program.
[PDF File, 393 KB - humanfactors_v12i1e63123_app1.pdf]

Multimedia Appendix 2
Questionnaire of feasibility and acceptability.
[PDF File, 95 KB - humanfactors_v12i1e63123_app2.pdf]

Multimedia Appendix 3
Reasons for intention to maintain the ME-BYO index measurement.
[PDF File, 71 KB - humanfactors_v12i1e63123_app3.pdf]

Multimedia Appendix 4
Comparison of baseline characteristics between the presence and absence of self-measurement.
[PDF File, 242 KB - humanfactors_v12i1e63123_app4.pdf]

References

1. Chong SOK, Pedron S, Abdelmalak N, Laxy M, Stephan AJ. An umbrella review of effectiveness and efficacy trials for app-based health interventions. NPJ Digit Med 2023 Dec 16;6(1):233. [doi: 10.1038/s41746-023-00981-x] [Medline: 38104213]
2. Romeo A, Edney S, Plotnikoff R, et al. Can smartphone apps increase physical activity? Systematic review and meta-analysis. J Med Internet Res 2019 Mar 19;21(3):e12053. [doi: 10.2196/12053] [Medline: 30888321]
3. Domin A, Spruijt-Metz D, Theisen D, Ouzzahra Y, Vögele C. Smartphone-based interventions for physical activity promotion: scoping review of the evidence over the last 10 years. JMIR Mhealth Uhealth 2021 Jul 21;9(7):e24308. [doi: 10.2196/24308] [Medline: 34287209]
4. Alnooh G, Alessa T, Noorwali E, et al. Identification of the most suitable mobile apps to support dietary approaches to stop hypertension (DASH) diet self-management: systematic search of app stores and content analysis. Nutrients 2023 Aug 6;15(15):3476. [doi: 10.3390/nu15153476] [Medline: 37571413]
5. Guo YQ, Chen Y, Dabbs AD, Wu Y. The effectiveness of smartphone app-based interventions for assisting smoking cessation: systematic review and meta-analysis. J Med Internet Res 2023 Apr 20;25:e43242. [doi: 10.2196/43242] [Medline: 37079352]
6. Ghelani DP, Moran LJ, Johnson C, Mousa A, Naderpoor N. Mobile apps for weight management: a review of the latest evidence to inform practice. Front Endocrinol (Lausanne) 2020;11:412. [doi: 10.3389/fendo.2020.00412] [Medline: 32670197]
7. Holmen H, Wahl AK, Cvancarova Småstuen M, Ribu L. Tailored communication within mobile apps for diabetes self-management: a systematic review. J Med Internet Res 2017 Jun 23;19(6):e227. [doi: 10.2196/jmir.7045] [Medline: 28645890]
8. Kassavou A, Wang M, Mirzaei V, Shpendi S, Hasan R. The association between smartphone app-based self-monitoring of hypertension-related behaviors and reductions in high blood pressure: systematic review and meta-analysis. JMIR Mhealth Uhealth 2022 Jul 12;10(7):e34767. [doi: 10.2196/34767] [Medline: 35819830]

XSL•FO
RenderX

9. Li R, Liang N, Bu F, Hesketh T. The effectiveness of self-management of hypertension in adults using mobile health: systematic review and meta-analysis. JMIR Mhealth Uhealth 2020 Mar 27;8(3):e17776. [doi: 10.2196/17776] [Medline: 32217503]
10. Xu H, Long H. The effect of smartphone app-based interventions for patients with hypertension: systematic review and meta-analysis. JMIR Mhealth Uhealth 2020 Oct 19;8(10):e21759. [doi: 10.2196/21759] [Medline: 33074161]
11. Al-Arkee S, Mason J, Lane DA, et al. Mobile apps to improve medication adherence in cardiovascular disease: systematic review and meta-analysis. J Med Internet Res 2021 May 25;23(5):e24190. [doi: 10.2196/24190] [Medline: 34032583]
12. Shrivastava TP, Goswami S, Gupta R, Goyal RK. Mobile app interventions to improve medication adherence among type 2 diabetes mellitus patients: a systematic review of clinical trials. J Diabetes Sci Technol 2023 Mar;17(2):458-466. [doi: 10.1177/19322968211060060] [Medline: 34861793]
13. Carroll JK, Moorhead A, Bond R, LeBlanc WG, Petrella RJ, Fiscella K. Who uses mobile phone health apps and does use matter? A secondary data analytics approach. J Med Internet Res 2017 Apr 19;19(4):e125. [doi: 10.2196/jmir.5604] [Medline: 28428170]
14. Ernsting C, Dombrowski SU, Oedekoven M, et al. Using smartphones and health apps to change and manage health behaviors: a population-based survey. J Med Internet Res 2017 Apr 5;19(4):e101. [doi: 10.2196/jmir.6838] [Medline: 28381394]
15. Bear HA, Ayala Nunes L, Ramos G, et al. The acceptability, engagement, and feasibility of mental health apps for marginalized and underserved young people: systematic review and qualitative study. J Med Internet Res 2024 Jul 30;26:e48964. [doi: 10.2196/48964] [Medline: 39078699]
16. Grundy Q. A review of the quality and impact of mobile health apps. Annu Rev Public Health 2022 Apr 5;43:117-134. [doi: 10.1146/annurev-publhealth-052020-103738] [Medline: 34910582]
17. Decade of Healthy Ageing: Baseline Report 2020: World Health Organization; 2021. URL: https://www.who.int/publications/i/item/9789240017900 [accessed 2025-12-10]
18. Nakamura S, Watanabe R, Saito Y, Watanabe K, Chung UI, Narimatsu H. The ME-BYO index: a development and validation project of a novel comprehensive health index. Front Public Health 2023;11:1142281. [doi: 10.3389/fpubh.2023.1142281] [Medline: 37213649]
19. Saito Y, Nakamura S, Tanaka A, Watanabe R, Narimatsu H, Chung UI. Evaluation of the validity and reliability of the 10-meter walk test using a smartphone application among Japanese older adults. Front Sports Act Living 2022;4:904924. [doi: 10.3389/fspor.2022.904924] [Medline: 36267485]
20. Saito Y, Nakamura S, Tanaka A, Watanabe R, Narimatsu H, Chung UI. Checking the validity and reliability of the Japanese version of the Mini-Cog using a smartphone application. BMC Res Notes 2022 Jun 25;15(1):222. [doi: 10.1186/s13104-022-06101-4] [Medline: 35752807]
21. Momma H, Kawakami R, Honda T, Sawada SS. Muscle-strengthening activities are associated with lower risk and mortality in major non-communicable diseases: a systematic review and meta-analysis of cohort studies. Br J Sports Med 2022 Jul;56(13):755-763. [doi: 10.1136/bjsports-2021-105061] [Medline: 35228201]
22. Kikuchi N, Ohta T, Hashimoto Y, et al. Effect of online home-based resistance exercise training on physical fitness, depression, stress, and well-being in middle-aged persons: a pilot study. Int J Environ Res Public Health 2023 Jan 18;20(3):1769. [doi: 10.3390/ijerph20031769] [Medline: 36767135]
23. Kikuchi N, Mochizuki Y, Kozuma A, et al. The effect of online low-intensity exercise training on fitness and cardiovascular parameters. Int J Sports Med 2022 May;43(5):418-426. [doi: 10.1055/a-1582-2874] [Medline: 34375992]
24. ME-BYO index measurement method. Kanagawa Prefecture. 2023. URL: https://www.pref.kanagawa.jp/docs/mv4/mebyo-index/mebyo-index_howtouse.html [accessed 2023-06-21]
25. Araujo de Carvalho I, Epping-Jordan J, Pot AM, et al. Organizing integrated health-care services to meet older people's needs. Bull World Health Organ 2017 Nov 1;95(11):756-763. [doi: 10.2471/BLT.16.187617] [Medline: 29147056]
26. Kobayashi T, Morimoto T, Shimanoe C, Ono R, Otani K, Mawatari M. Development of a simple screening tool based on the 5-question geriatric locomotive function scale for locomotive syndrome. J Orthop Sci 2022 Jul;27(4):913-920. [doi: 10.1016/j.jos.2021.05.001] [Medline: 34090778]
27. Borson S, Scanlan J, Brush M, Vitaliano P, Dokmak A. The Mini-Cog: a cognitive "Vital Signs" measure for dementia screening in multi-lingual elderly. Int J Geriatr Psychiatry 2000 Nov;15(11):1021-1027. [doi: 10.1002/1099-1166(200011)15:11<1021::aid-gps234>3.0.co;2-6] [Medline: 11113982]
28. Borson S, Scanlan JM, Chen P, Ganguli M. The Mini-Cog as a screen for dementia: validation in a population-based sample. J Am Geriatr Soc 2003 Oct;51(10):1451-1454. [doi: 10.1046/j.1532-5415.2003.51465.x] [Medline: 14511167]
29. Higuchi M, Nakamura M, Shinohara S, et al. Effectiveness of a voice-based mental health evaluation system for mobile devices: prospective study. JMIR Form Res 2020 Jul 20;4(7):e16455. [doi: 10.2196/16455] [Medline: 32554367]
30. Maruyama T, Ekuni D, Higuchi M, Takayama E, Tokuno S, Morita M. Relationship between psychological stress determined by voice analysis and periodontal status: a cohort study. Int J Environ Res Public Health 2022 Aug 2;19(15):9489. [doi: 10.3390/ijerph19159489] [Medline: 35954845]

XSL•FO
RenderX

31. Miyashita H, Nakamura M, Svensson AK, et al. Association between electroencephalogram-derived sleep measures and the change of emotional status analyzed using voice patterns: observational pilot study. JMIR Form Res 2020 Jun 9;4(6):e16880. [doi: 10.2196/16880] [Medline: 32515745]
32. Kanagawa MY ME-BYO record terms of use. Kanagawa Prefecture. 2023. URL: https://www.pref.kanagawa.jp/docs/fz7/mymebyo/riyoukiyaku_en.html [accessed 2023-04-01]
33. Kanagawa MY ME-BYO record privacy policy, security policy. Kanagawa Prefecture. 2023. URL: https://www.pref.kanagawa.jp/docs/fz7/cnt/f532715/p991437.html [accessed 2024-03-06]
34. ME-BYO index privacy policy. Kanagawa Prefecture. 2020. URL: https://www.pref.kanagawa.jp/docs/mv4/mebyo-index/mebyo-index_refinement.html [accessed 2023-05-31]
35. Proctor E, Silmere H, Raghavan R, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 2011 Mar;38(2):65-76. [doi: 10.1007/s10488-010-0319-7] [Medline: 20957426]
36. Cho A, Sugimura M, Nakano S, Yamada T. The Japanese MCI screen for early detection of Alzheimer's disease and related disorders. Am J Alzheimers Dis Other Demen 2008;23(2):162-166. [doi: 10.1177/1533317507312624] [Medline: 18223126]
37. Shankle WR, Mangrola T, Chan T, Hara J. Development and validation of the Memory Performance Index: reducing measurement error in recall tests. Alzheimers Dement 2009 Jul;5(4):295-306. [doi: 10.1016/j.jalz.2008.11.001] [Medline: 19560100]
38. Topp CW, Østergaard SD, Søndergaard S, Bech P. The WHO-5 Well-Being Index: a systematic review of the literature. Psychother Psychosom 2015;84(3):167-176. [doi: 10.1159/000376585] [Medline: 25831962]
39. Tiganov AS, Panteleeva GP, Vertogradova OP, Kondrat'ev FV, Tsutsul'kovskaia MI. The systematics and diagnostic criteria of endogenous psychoses in light of the International Statistical Classification of Diseases--10th revision (ICD-10). Zh Nevrol Psikhiatr Im S S Korsakova 1997;97(10):4-10. [Medline: 9424346]
40. Kessler RC, Green JG, Gruber MJ, et al. Screening for serious mental illness in the general population with the K6 screening scale: results from the WHO World Mental Health (WMH) survey initiative. Int J Methods Psychiatr Res 2010 Jun;19 Suppl 1(Suppl 1):4-22. [doi: 10.1002/mpr.310] [Medline: 20527002]
41. Furukawa TA, Kawakami N, Saitoh M, et al. The performance of the Japanese version of the K6 and K10 in the World Mental Health Survey Japan. Int J Methods Psychiatr Res 2008;17(3):152-158. [doi: 10.1002/mpr.257] [Medline: 18763695]
42. Bohannon RW, Bubela DJ, Magasi SR, Wang YC, Gershon RC. Sit-to-stand test: performance and determinants across the age-span. Isokinet Exerc Sci 2010;18(4):235-240. [doi: 10.3233/IES-2010-0389] [Medline: 25598584]
43. Ohkawara K, Oshima Y, Hikihara Y, Ishikawa-Takata K, Tabata I, Tanaka S. Real-time estimation of daily physical activity intensity by a triaxial accelerometer and a gravity-removal classification algorithm. Br J Nutr 2011 Jun;105(11):1681-1691. [doi: 10.1017/S0007114510005441] [Medline: 21262061]
44. Oshima Y, Kawaguchi K, Tanaka S, et al. Classifying household and locomotive activities using a triaxial accelerometer. Gait Posture 2010 Mar;31(3):370-374. [doi: 10.1016/j.gaitpost.2010.01.005] [Medline: 20138524]
45. Park J, Ishikawa-Takata K, Tanaka S, Bessyo K, Tanaka S, Kimura T. Accuracy of estimating step counts and intensity using accelerometers in older people with or without assistive devices. J Aging Phys Act 2017 Jan;25(1):41-50. [doi: 10.1123/japa.2015-0201] [Medline: 27180730]
46. Japan Physical Activity Research Platform. 2020. URL: http://paplatform.umin.jp [accessed 2021-08-28]
47. López D, Torres M, Vélez J, et al. Development and evaluation of a nutritional smartphone application for making smart and healthy choices in grocery shopping. Healthc Inform Res 2017 Jan;23(1):16-24. [doi: 10.4258/hir.2017.23.1.16] [Medline: 28261527]
48. Palacios C, Torres M, López D, Trak-Fellermeier MA, Coccia C, Pérez CM. Effectiveness of the nutritional app "MyNutriCart" on food choices related to purchase and dietary behavior: a pilot randomized controlled trial. Nutrients 2018 Dec 12;10(12):1967. [doi: 10.3390/nu10121967] [Medline: 30545125]
49. Almuqrin A, Hammoud R, Terbagou I, Tognin S, Mechelli A. Smartphone apps for mental health: systematic review of the literature and five recommendations for clinical translation. BMJ Open 2025 Feb 11;15(2):e093932. [doi: 10.1136/bmjopen-2024-093932] [Medline: 39933815]
50. Chei CL, Nakamura S, Watanabe K, Narimatsu H. User engagement, demographics, and health status of the my ME-BYO record, a personal health management mobile app: retrospective study. J Med Internet Res 2025 Oct 10;27:e79109. [doi: 10.2196/79109] [Medline: 41072922]
51. Bardus M, van Beurden SB, Smith JR, Abraham C. A review and content analysis of engagement, functionality, aesthetics, information quality, and change techniques in the most popular commercial apps for weight management. Int J Behav Nutr Phys Act 2016 Mar 10;13(1):35. [doi: 10.1186/s12966-016-0359-9] [Medline: 26964880]
52. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol 2009 Nov;28(6):690-701. [doi: 10.1037/a0016136] [Medline: 19916637]
53. Bondaronek P, Alkhaldi G, Slee A, Hamilton FL, Murray E. Quality of publicly available physical activity apps: review and content analysis. JMIR Mhealth Uhealth 2018 Mar 21;6(3):e53. [doi: 10.2196/mhealth.9069] [Medline: 29563080]
54. Editage. URL: https://www.editage.jp/ [accessed 2025-12-11]

XSL•FO
RenderX

Abbreviations

IOF: implementation outcome framework
MCI: mild cognitive impairment
MPI: Memory Performance Index
WHO-5: World Health Organization-5 Well-Being Index

Edited by A Kushniruk; submitted 11.06.24; peer-reviewed by DH Craighead, JLC Lee, Y Chu; revised version received 21.10.25; accepted 20.11.25; published 16.12.25.

Please cite as:
Saito Y, Kikuchi N, Watanabe K, Nakamura S, Narimatsu H
Use of the ME-BYO Index, a Mobile Health App, During an Online Strength Training Program in Adults: Fidelity, Feasibility, and Acceptability Study
JMIR Hum Factors 2025;12:e63123
URL: https://humanfactors.jmir.org/2025/1/e63123
doi:10.2196/63123

© Yoshinobu Saito, Naoki Kikuchi, Kaname Watanabe, Sho Nakamura, Hiroto Narimatsu. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 16.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Using Crowdsourcing Internet of Things Technology to Reduce Caregiver Worry in Dementia-Related Lost Episodes: Longitudinal Observational Study

Bel Wong[1*], MSocSc; Tobi Cheng[1*], MA; Nicole Fung[1*], BSc; Zhongming Lin[2*], MEng; Ki-Kit Lai[2*], PhD; Florence Ho[1*], MSc; S-H Gary Chan[2*], PhD; Timothy Kwok[1,3*], MD

[1]Jockey Club Centre for Positive Ageing, 27 A Kung Kok Street, Shatin, China (Hong Kong)

[2]Hong Kong University of Science and Technology, Kowloon, China (Hong Kong)

[3]Chinese University of Hong Kong, Shatin, China (Hong Kong)

[*]all authors contributed equally

Corresponding Author:
Bel Wong, MSocSc
Jockey Club Centre for Positive Ageing, 27 A Kung Kok Street, Shatin, China (Hong Kong)

Abstract

Background: Dementia increases the risk of individuals getting lost due to cognitive decline, impacting daily functioning and heightening caregiver worry. Traditional search methods are often time-consuming and stressful, whereas GPS-based technologies face limitations such as battery dependency. A crowdsourcing Internet of Things (IoT) technology using energy-efficient Bluetooth Low Energy (BLE) offers a potential solution to locate missing individuals with dementia more effectively by harnessing the power of the crowd and fostering a caring and inclusive community.

Objective: This study aimed to evaluate the effectiveness of a BLE-based privacy-preserving crowdsourcing IoT system consisting of a BLE tag and an Android and iOS app in improving lost-related behavior and psychological well-being by facilitating searches, after-care arrangements, and reducing caregiver worry, as well as to assess its usability among caregivers of individuals with dementia in Hong Kong.

Methods: A single-arm, prospective observational study was conducted from November 2020 to October 2023. Caregivers (N=1034) of individuals with dementia used a staff-assisted crowdsourcing IoT technology comprising a BLE tag, mobile app sensor, and location cloud server. Outcomes included search strategies, post–getting lost care arrangements, caregiver worry and distress (10-point scale), and usability (modified Quebec User Evaluation of Satisfaction with Assistive Technology 2.0 survey). Data were collected at 6- and 12-month follow-ups and analyzed using generalized estimating equations and linear mixed models.

Results: Of the 1034 participants, 143 (13.82%) reported lost episodes, with 51 (35.7%) using BLE tags for searches. Worry about future lost episodes decreased significantly over time (P=.008), especially among BLE tag users (P=.04). There was an association between BLE tag use and adoption of proactive search strategies (eg, going out to search: adjusted odds ratio 2.78, 95% CI 1.33-5.82; P=.007) and preventative measures (eg, IoT devices or CCTV: adjusted odds ratio 2.92, 95% CI 1.61-5.29; P<.001). Usability satisfaction was high for design and data security, whereas approximately half of the participants (309/707, 43.7%) were satisfied with accuracy.

Conclusions: The BLE crowdsourcing system may reduce caregiver worry and encourage proactive search behaviors, although accuracy depends on broader community adoption. Integration into dementia care plans could enhance safety and autonomy. Further research with a randomized controlled trial design is needed to confirm these findings.

(JMIR Hum Factors 2025;12:e73670) doi:10.2196/73670

KEYWORDS

dementia; getting lost; IoT; Internet of Things; crowdsourcing IoT technology; caregiving worry; caregiving distress

Introduction

Dementia is a neurodegenerative syndrome characterized by progressive impairments in cognitive domains, including memory, judgment, and orientation, significantly impacting daily functioning [1]. Individuals with dementia are at an increased risk of becoming lost due to compromised perceptual-motor and visuospatial abilities [2-5], as well as difficulties in acquiring and retaining new information [6]. Cognitive deterioration, particularly in orientation and executive function [7,8], along with psychological factors, such as depression, agitation, and stress-related coping mechanisms, is

XSL•FO
RenderX

a key contributor to this risk [9-11]. The consequences of getting lost can be severe, leading to psychological distress, heightened fear [12,13], and increased risk of serious injury or mortality, especially if the individual remains missing for an extended period [14,15]. Moreover, such incidents often result in premature admission to nursing homes [5], negatively affecting the overall quality of life for individuals with dementia.

The impact of these episodes extends beyond the individuals themselves, as caregivers frequently experience heightened stress, worry, exhaustion, and fatigue due to prolonged surveillance [13,16]. The fear of getting lost often prompts caregivers to adopt restrictive measures, such as limiting outdoor activities, which can compromise the autonomy and quality of life of individuals with dementia [17,18]. Traditional methods for locating lost individuals, such as neighborhood searches and contacting authorities, are time-consuming and often yield limited success. Recent technological advancements, including GPS tracking using mobile phones or dedicated wearable devices such as watches, offer more efficient solutions but are not without challenges, such as the need for regular battery charging for the wearable devices, which individuals with dementia who live alone may forget.

In response to these challenges, a search technology was developed that harnesses the power of the crowd and uses more energy-efficient Bluetooth devices. This technology aims to provide a more reliable and community-driven solution for locating individuals with dementia who become lost. This study sought to evaluate the effectiveness of this technology in facilitating searches and reducing caregiver worry during dementia-related lost episodes. Additionally, it aimed to explore the potential benefits and limitations of this technology in real-world scenarios, contributing to the development of more effective support systems for individuals with dementia and their caregivers.

Methods

The Crowdsourcing Internet of Things (IoT) Technology

This study used a crowdsourcing IoT technology developed by the co-authors at the Hong Kong University of Science and Technology [19]. The primary objective of this system was to locate individuals using a software-based crowdsourcing. The technology's development began in 2019, with deployment starting in 2020, predating the launch of Apple's AirTag [20]. The technology comprises 3 main components: a Bluetooth Low Energy (BLE) tag, a mobile app for both Android and iOS platforms, and a location cloud server. The BLE tag, designed and customized as a walking stick holder, keychain, or smart card, emits a signal detectable by the sensor within a range of 30 to 100 meters (Figure 1). While AirTag's "Find My" network relies on the iOS ecosystem [21], our technology engages both Android and iOS communities, broadening the search network and supporting our aim of the project to foster a dementia-friendly community through the active participation of caregivers and volunteers (referred to as "Dementia Angels") in the search.

XSL•FO
RenderX

Figure 1. The Bluetooth Low Energy tag in different designs. (A) Overview, (B) walking stick holder design, (C) keychain design, and (D) smart card design.



The Dementia Angels download the sensor app, which uses GPS for positioning. The graphical user interface of the app is shown in Figures 2–3. The app consists of 2 primary interfaces: "report lost" interface for caregivers (Figure 2) and the "missing person report" interface for Dementia Angels (Figure 3). The "report lost" interface provides the location where and when the BLE tag was last tracked. When the caregiver reports a lost episode and activates a search, they input the lost episode–related information to aid Dementia Angels' searching. The "missing person report" interface displays recent missing reports and successful recoveries. The app incorporates large, touch-friendly buttons and icons and uses different color schemes to distinguish the 2 interfaces. The app is in Chinese, and Figures 2–3 show a translated version for illustration purposes.

XSL•FO
RenderX

Figure 2. "Report lost" interface for caregivers.

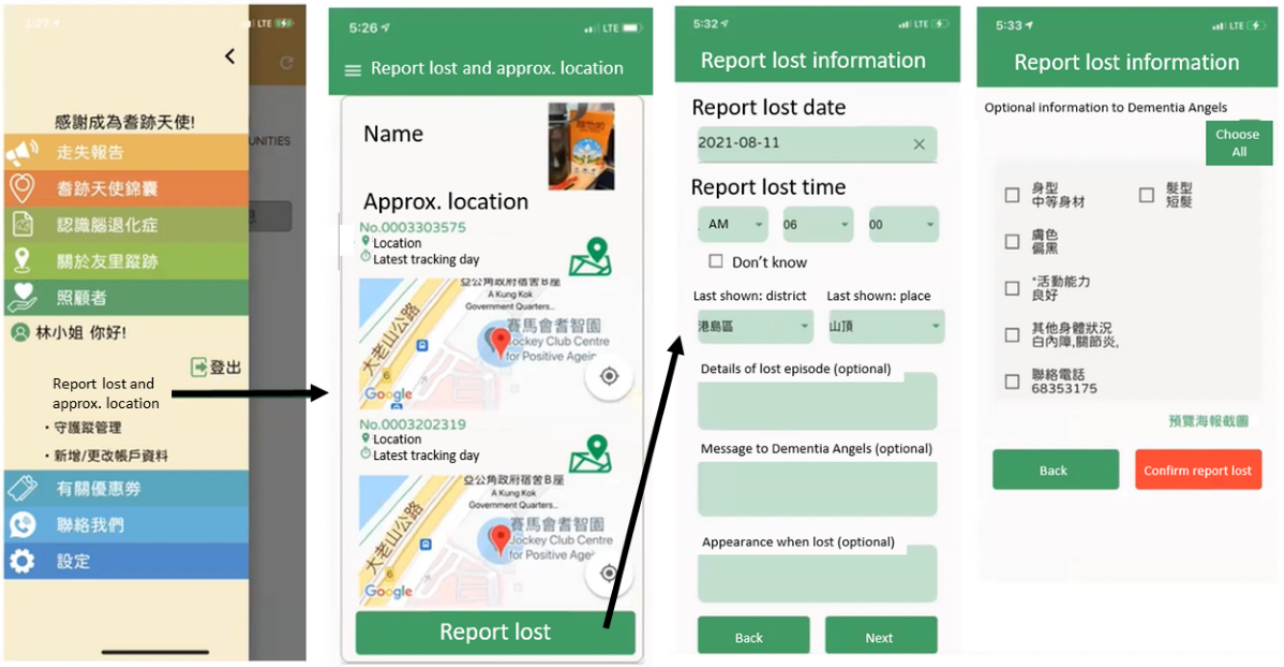


Figure 4 illustrates the mechanism of the technology. When a care recipient goes missing, the caregiver can activate a search notification. This prompts caregivers and Dementia Angels to enable Bluetooth and GPS on their devices to detect the BLE tag signal. Upon detecting the signal, the mobile devices anonymously relay their location to a location cloud server, where a localization engine determines the location of the BLE tag. This information is then communicated to caregivers to aid in the search. The accuracy of the positioning, therefore, depends on the number of scans from Dementia Angels.

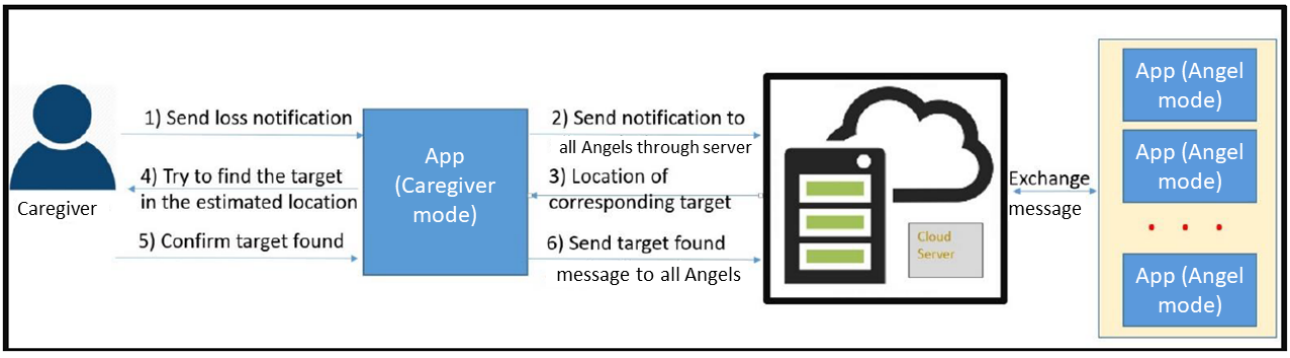
This technology adopts the privacy by design approach. The key advantages of the technology include real-time and on-demand positioning capabilities, which protect the care recipients' rights when the positioning is only performed when they are in a safety threat. The BLE tag is water-resistant, with a battery life of up to 1 year and low battery notification on the caregiver interface to ensure timely replacement by caregivers, in contrast to GPS-based trackers commonly used during the technology's development, which typically require more frequent recharging due to higher power consumption. Furthermore, the search function is only activated when necessary, minimizing inconvenience for caregivers and Dementia Angels. All location data are anonymized to protect the privacy of the searchers.

XSL•FO
RenderX

Figure 3. "Missing person report" interface for Dementia Angels.



Figure 4. The mechanism of the lost-and-found process of the crowdsourcing Internet of Things technology.



XSL•FO
RenderX

Study Design and Hypotheses

This study adopted a single-arm, prospective, observational study design. The research hypotheses were that participants would (1) adopt more proactive search strategies and post–getting lost care arrangements when using the IoT technology, (2) have less lost episode–related distress and worry when using the IoT technology, and (3) be satisfied with the usability of the IoT technology.

Participants

The participants in this study were caregivers of individuals with dementia. The inclusion criteria were (1) citizens residing in Hong Kong, (2) aged ≥18 years, (3) providing care to individuals diagnosed with dementia, (4) smartphone users, and (5) willing to participate in the study by giving informed consent. The exclusion criteria were (1) inability to participate in research due to visual, hearing, physical, or psychological issues; and (2) refusal to provide informed consent. The participants were recruited through self-enrollment in the program, networking, and referral from other organizations; as this was a usability test, we adopted a maximum recruitment rationale and included all participants who were willing to contribute to this study.

Outcomes

The primary outcome of the study was the difference in search strategies and post–getting lost care arrangements with the use of the IoT technology, measured by a survey containing a set of predesigned strategies and arrangements with binary "yes or no" responses and an open-ended "others" option.

The secondary outcomes were (1) caregiving worry toward future lost episodes after a getting lost episode, measured by the question "How worried are you about future lost episodes?," with a 10-point rating scale in which a higher score indicated higher worry; (2) caregiving distress toward the getting lost episode, measured by the question "How distressed are you about this getting lost episode?," with a 10-point rating scale in which a higher score indicated higher distress; (3) factors associated with getting lost in the dementia context by analysis based on the demographic data; and (4) usability of the IoT technology, measured by a survey modified from the Quebec User Evaluation of Satisfaction with Assistive Technology 2.0 [22]. The original scale has 12 items, including dimensions, weight, adjustments, safety, durability, simplicity of use, comfort, effectiveness of service, delivery, repairs and servicing, professional services, and follow-up. A 5-point Likert scale, ranging from "very satisfied" to "not satisfied at all," was used. The survey used in this study was modified to suit the technology trialed in this study, and the items evaluated included size, weight, willingness of the care recipient to carry the BLE tag during and after the trial, durability, design, ease of app operation, accuracy of the technology, and data security. In addition, the number of caregivers who installed the sensor app for use and whether the paid helpers of the care recipients installed the sensor app were recorded to evaluate the caregiver engagement with the technology.

To complement the structured survey-based outcomes, real-life data on the usage of the IoT technology were collected during the study period. These data, primarily consisting of scan counts by Dementia Angels, served as an indicator of the technology's deployment in a naturalistic setting.

Study Procedures

The social worker, project staff (interveners), or researchers of 2 nongovernmental organizations explained the study details to potential participants and obtained their consent for research participation. The participants were then distributed 1 to 2 BLE tags and a note about how to install the sensor mobile app and how to pair the BLE tag with their sensor accounts. They then tested the BLE tag in daily lives, whereas the interveners provided technical support, such as installation, pairing, and use, as well as support during lost episodes if needed by the caregivers. The participants were notified by the mobile app at 6-month and 12-month time points from baseline to self-administer a survey to provide information about use and getting the lost episodes, if any. The participants who did not respond were contacted by the researchers 3 times at different time slots on different days before they were regarded as "did not complete posttest." The data collected from the latest interview were included in this study.

Statistical Analysis

The baseline characteristics of participants were presented using descriptive statistics, whereas between-group differences were analyzed using the chi-square test. The outcome of factors associated with getting lost was analyzed by generalized estimating equations with a binary logistic model and time as a within-subject effect; demographic variables were the predictors. The outcomes of search strategies and post–getting lost care arrangement were analyzed by generalized estimating equation with a binary logistic model and time as within-subject effect, with BLE tag use as the predictor. Paired comparison of changes in these 2 outcomes was conducted using the 2-sided McNemar test. The outcome for caregiver worry and caregiver distress was analyzed using a linear mixed model, with time effect, group effect, and group×time interaction effect using first-order autoregressive covariance structures because it had the smallest Akaike Information Criterion. All analyses were adjusted for caregiver age, caregiver sex, care recipient age, and care recipient sex. The usability outcomes were presented descriptively, with "very satisfied" and "satisfied" grouped as the "satisfaction." Missing data were left missing. All statistical analyses were performed using the statistical package SPSS version 26 (IBM Corp). Significance was set at P of .05.

Ethical Considerations

This study was approved by the Survey and Behavioral Research Ethics Committee of the Chinese University of Hong Kong (reference 175 - 19) [23]. The study was conducted according to the Declaration of Helsinki—ethical principles for medical research involving human participants. Informed consent for participation in this study was obtained from the participants, covering the use of demographic data and their reports of lost episode–related experiences. In addition, both participants and their care recipients provided written program-level informed consent before program participation. The program consent form introduced the program objectives, explained the use of

XSL•FO
RenderX

the tracking device, and specified that data collected through the program would be used for evaluation purposes. All participants and their care recipients were informed of the program details verbally and in writing before signing the program consent. Participants did not receive any financial or nonfinancial compensation for participating in this study, and there was no cost to the participants for any study procedure. The confidentiality of personal data was maintained in accordance with the Personal Data (Privacy) Ordinance of Hong Kong. Data security was ensured through encryption protocols and secure storage practices, and care recipients' data were contributed in deidentified form through the program's tracking record.

Results

Overview

Data were collected from November 2020 to October 2023. As shown in Figure 5, 1252 caregivers joined the trial for 6 to 12 months, 1034 (82.58%) of them completed the 6- or 12-month posttest and were included in this study. Among the 1034 participants, 143 (13.82%) reported a lost episode of care recipients during the trial, and 51 of the 143 (35.7%) participants used a BLE tag to help with the search during the lost episode.

Figure 5. Participant flowchart. BLE: Bluetooth Low Energy.

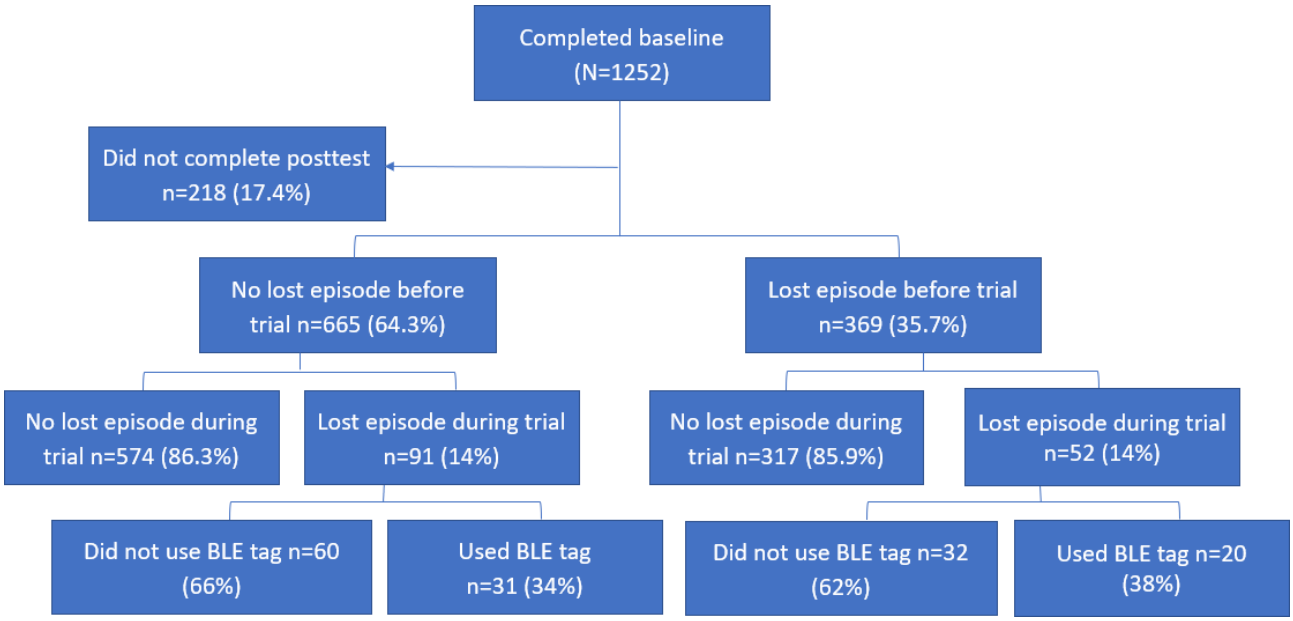


Table 1 presents the baseline participant characteristics. Approximately 20% (189/1034) of the caregiver participants were aged ≥65 years, and a quarter of the participants (256/1034, 24.75%) were male participants. They were mainly the children of the care recipients with dementia, and half of them (512/1034, 49.51%) lived with the care recipients. Of the 1034 participants, 59.86% (n=619) of care recipients were aged ≥80 years, and nearly half of them (n=480, 46.42%) had moderate or late-stage dementia, while 370 (35.78%) of them had experienced lost episodes before the trial. There was no significant difference in baseline characteristics between those who experienced lost episodes during the trial versus those who did not, except baseline worry about future lost episodes (mean 8.7, SD 1.6 vs mean 8.0, SD 2.1; *P*=.03).

XSL•FO
RenderX

Table . Baseline characteristics of participants, breakdown by getting lost or not during trial (N=1034).

Characteristics	Total	Got lost during trial (n=143)	Did not get lost during trial (n=891)	Chi-square (*df*) or *t* test (*df*)	*P* value
Caregiver demographics, n (%)					
Age (≥65 y)	189 (18.27)	32 (22.4)	157 (17.6)	1.8 (1)[a]	.18
Sex (male)	256 (24.75)	36 (25.2)	220 (24.7)	0.01 (1)[a]	.91
Spouse	172 (16.63)	22 (15.4)	150 (16.9)	0.2 (1)[a]	.66
Child	800 (77.36)	110 (76.9)	690 (77.5)	0.03 (1)[a]	.85
Lives with CR[b]	512 (49.51)	67 (46.9)	445 (50.0)	0.5 (1)[a]	.47
CR demographics, n (%)					
Age (≥80 y)	619 (59.86)	86 (60.1)	533 (59.9)	0.003 (1)[a]	.95
Sex (male)	407 (39.36)	59 (41.3)	348 (39.1)	0.2 (1)[a]	.62
Moderate or late-stage dementia	480 (46.42)	70 (49.0)	410 (46.1)	0.6 (1)[a]	.44
Lives alone	124 (11.99)	13 (9.1)	111 (12.5)	1.3 (1)[a]	.25
Lives with domestic helper only	68 (6.57)	10 (7.0)	58 (6.5)	0.04 (1)[a]	.83
Lives with spouse only	300 (29.01)	41 (28.7)	259 (29.1)	0.01 (1)[a]	.91
Lives with family without domestic helper	759 (73.40)	109 (76.2)	650 (73.0)	0.6 (1)[a]	.43
Lives with family and domestic helper	81 (7.83)	11 (7.7)	70 (7.9)	0.01 (1)[a]	.94
Baseline lost-related characteristics, n (%)					
Lost episode before trial (yes)	370 (35.78)	52 (36.4)	318 (35.7)	0.02 (1)[a]	.88
Clinical characteristics, mean (SD)					
Baseline distress about lost[c]	7.8 (2.2)	8.2 (2.0)	7.7 (2.2)	−1.4 (363)[d]	.20
Baseline worry about future lost[e]	8.1 (2.0)	8.7 (1.6)	8.0 (2.1)	−2.2 (362)[d]	.03[f]

[a] Chi-square test.

[b]CR: care recipient.

[c]Total n=365; 52 got lost during trial; and 313 did not get lost during trial.

[d]*t* test.

[e]Total n=364; 52 got lost during trial; and 312 did not get lost during trial.

[f]*P*<.05.

Factors Associated With Getting Lost

After adjusting for covariates, care recipients in moderate or late-stage dementia were more likely to get lost (adjusted odds ratio [AOR] 1.84, 95% CI 1.49-2.27; *P*<.001). Care recipients who lived alone were less likely to get lost before adjustment (Table 2).

XSL•FO
RenderX

Table . Factors associated with getting lost (N=1034).

Outcome: getting lost during trial[a]	Univariate[b]		Multivariate	
	OR[c] (95% CI)	*P* value	AOR[d] (95% CI)	*P* value
Caregiver is ≥65 years	1.21 (0.94-1.56)	.14	1.10 (0.84-1.43)	.48
CR[e] is male	1.19 (0.97-1.47)	.10	1.19 (0.95-1.48)	.13
CR has moderate or late-stage dementia	1.84 (1.49-2.27)	<.001[f]	1.84 (1.49-2.27)	<.001[f]
CR lives alone	0.73 (0.53-1.00)	.05	0.85 (0.55-1.29)	.44
CR lives with family without domestic helper	1.21 (0.95-1.55)	.12	1.11 (0.79-1.56)	.54

[a]Generalized estimating equations.

[b]Variables with *P*<.20 in univariate regression were included in the multivariate regression.

[c]OR: odds ratio.

[d]AOR: adjusted odds ratio; adjusted for caregiver age, caregiver sex, care recipient age, and care recipient sex.

[e]CR: care recipient.

[f]*P*<.05.

Intervention Effects on Search Strategies and Post–Getting Lost Care Arrangement

A total of 143 caregivers had experienced episodes of lost care recipients during the trial. Among them, 31 (21.7%) reported that the getting lost episodes happened in the morning (6 AM to 12 PM), 80 (55.9%) in the daytime (12 PM to 6 PM), and 32 (22.4%) in the nighttime (6 PM to 6 AM). Of 143 caregivers, 52 (36.4%) reported that the getting lost episode happened on the street; 51 (35.7%) at home; and 21 (14.7%) at shopping malls, markets, or retail shops.

Of 143 caregiver participants, 51 (35.7%) used the BLE tag to aid searching. There was no significant baseline difference between the BLE tag users and nonusers. Of 51 lost care recipients using the BLE tag, 17 (33%) were found within 1 hour, whereas 36 (39%) of 92 lost care recipients without using BLE tag were found within 1 hour (*P*=.49). Of 51 BLE tag user cases, 49 (96%) and 90 (98%) of 92 nonuser cases were found within 24 hours (*P*=.62).

Table 3 presents the comparison of the search strategies and post–getting lost care arrangement between BLE tag users and nonusers. The results revealed that BLE tag users had higher adjusted odds of going out to search for lost care recipients (AOR 2.78, 95% CI 1.33-5.82; *P*=.007), contacting police (AOR 2.59, 95% CI 1.26-5.35; *P*=.01), and using media or social media to aid search (AOR 22.45, 95% CI 2.62-192.38; *P*=.005); paired comparisons on search strategies pretrial and during trial (Multimedia Appendix 1) showed an insignificant, directional increase among BLE users for using these strategies, whereas nonusers had an insignificant, directional decrease in using these measures except for an insignificant increase in police calls.

After the lost episode, BLE tag users had higher odds of using an IoT device or installing CCTV to prevent future lost episodes (AOR 2.92, 95% CI 1.61-5.29; *P*<.001); paired comparison (Multimedia Appendix 1) showed an increase among BLE tag users and a decrease among nonusers in using this measure from pretrial to posttrial, but neither change reached statistical significance.

XSL•FO
RenderX

Table . Adjusted odds ratios (AOR) of search strategies and post–getting lost care arrangement of Bluetooth Low Energy tag users compared with nonusers (n=143).

Outcomes[a]	Getting lost during trial, AOR (95% CI)	*P* value
Strategies of searching after lost episodes		
Go out to search	2.78 (1.33-5.82)	.007[b]
Seek help from relatives, neighbors, or passers-by	1.05 (0.59-1.86)	.86
Call police	2.59 (1.26-5.35)	.01[b]
Seek help from media or social media	22.45 (2.62-192.38)[c]	.005[b]
Locate via IoT[d] device	1.12 (0.49-2.55)	.79
Care arrangement after recent lost episode		
Forbid CR[e] to go out alone or lock main door	1.00 (0.53-1.87)	.10
Provide CR cell phone	1.22 (0.62-2.39)	.57
Provide information tag	1.03 (0.47-2.29)	.93
Use IoT device or install CCTV	2.92 (1.61-5.29)	<.001[b]
Inform security guard	0.99 (0.48-2.08)	.99
Arrange extra manpower on care or hire paid helper	0.84 (0.42-1.67)	.61
Arrange daycare or residential service	0.72 (0.27-1.90)	.51

[a]Generalized estimating equations. Predictor: use BLE tag during trial.

[b]*P*<.05. Adjusted for caregiver age, caregiver sex, care recipient age, and care recipient sex.

[c]The event count was <10.

[d]IoT: Internet of Things.

[e]CR: care recipient.

Intervention Effects on Caregiver's Worry and Anxiety

A total of 52 caregivers had experienced episodes of lost care recipient both before (baseline) and during the trial (posttest). BLE tag users had significantly higher worry about future lost episodes (*P*=.04) and distress toward recent lost episodes (*P*<.001) than nonusers, regardless of time. Analysis of baseline and posttest scores revealed that worry toward future lost episodes decreased significantly over time, regardless of groups (*P*=.01). There was no group×time interaction effect on the 2 outcomes, although BLE tag users reported a significant reduction in worry about future lost episodes (*P*=.04); in contrast, the decrease in worry in nonusers was not statistically significant. No significant result was observed in the reduction in distress (Table 4).

XSL•FO
RenderX

Table . Comparison of caregiver worry and anxiety between Bluetooth Low Energy (BLE) tag users and nonusers (n=52).

Outcomes[a]	BLE tag users (n=20)	Nonusers (n=32)	Group effect		Time effect		Group×time effect	
	EMM[b] (SE)	EMM (SE)	*F* test (*df*)	*P* value	*F* test (*df*)	*P* value	*F* test (*df*)	*P* value
Caregiver's worry about future getting lost			4.63 (1, 46)	.04[c]	7.72 (1, 50)	.01[c]	0.27 (1, 50)	.61
Baseline	9.28 (0.38)	8.32 (0.30)						
Posttest	8.23 (0.38)	7.60 (0.30)						
Within-group change	−1.05 (0.50)	−0.72 (0.39)						
Within-group change *P* value	0.04[c]	0.07						
Caregiver's distress about getting lost episode			16.45 (1, 46)	<.001[c]	2.28 (1, 50)	.14	.06 (1, 50)	.81
Baseline	9.20 (0.41)	7.50 (0.32)						
Posttest	8.60 (0.41)	7.06 (0.32)						
Within-group change	−0.60 (0.54)	−0.44 (0.43)						
Within-group change *P* value	.27	.31						

[a]Linear mixed model.

[b]EMM: estimated marginal means.

[c]*P*<.05. Adjusted for caregiver age, caregiver sex, care recipient age, and care recipient sex.

Usability Outcomes

A total of 1034 participants who completed the posttest were included for usability analysis. More than 80% of the care recipients were observed to be willing to carry the BLE tags during (780/952, 81.9%) and after the trial (521/631, 82.6%). More than 70% (691/954) of participants trusted the data security, but less than half of the participants (309/707) were satisfied with the accuracy of the technology (Table 5).

XSL•FO
RenderX

Table . Usability outcomes (N=1034).

Usability outcomes	Valid [a], n (%)	Satisfaction[a], n (%)
Satisfaction of care recipients toward the size of BLE[b] tag	957 (92.55)	652 (68.1)
Satisfaction of care recipients toward the weight of BLE tag	956 (92.45)	668 (69.9)
Care recipients were willing to carry BLE tag during the trial	952 (92.06)	780 (81.9)
Care recipients will be willing to carry BLE tag after the trial	631 (61.02)	521 (82.6)
Satisfaction of participants toward the durability of BLE tag	956 (92.45)	763 (79.8)
Satisfaction of participants toward the design of BLE tag (convenient for care recipients to carry)	954 (92.26)	814 (85.3)
Satisfaction of participants toward the easiness of app operation	951 (91.97)	671 (70.6)
Satisfaction of participants toward the accuracy of the technology	707 (68.37)	309 (43.7)
Satisfaction of participants toward the data security	954 (92.26)	691 (72.4)
Paid helpers install the app to use the technology during the trial	268 (25.91)	39 (14.6)
Number of caregivers who installed the app to use the technology during the trial	805 (77.9)	1.71 (1.13)[c]

[a]Missing and "not applicable" are excluded from the analysis; therefore, the percentage is based on valid responses.

[b]BLE: Bluetooth Low Energy.

[c]This data point is a mean (SD) value.

Real-Life Use

The program delivered 4836 BLE tags to 3259 individuals with dementia (including those who did not consent to participate in the study), averaging 1.5 BLE tags per person. There were 47,012 Dementia Angels who downloaded our app to detect signals from the BLE tags (20,269 iOS, 43.1%; and 26,743 Android, 56.9%). Additionally, more than 900 Dementia Angel sensors (physical boxes and mobile app) were installed across shopping malls, public transport hubs, and retail shops, reflecting significant community engagement.

Among the 1252 study participants, caregivers initiated 111 lost episodes, whereas 51 (45.9%) self-reported using the BLE tag during searches for missing care recipients. This discrepancy might be attributed to factors such as multiple searches by some caregivers and variation in reporting. Despite these differences, the real-life data provide valuable insights into use patterns, which are presented in this section.

Table 6 presents the characteristics of the 111 lost episodes initiated by caregiver participants in the mobile app. The majority of caregivers who initiated searches were women aged <65 years, predominantly children of the care recipients. Approximately half of the care recipients (62/111, 56%) were aged ≥80 years. More than 60% (69/111) of them were in the moderate or late stage of dementia, and more than 70% (82/111) maintained good mobility. On average, each search involved 18.59 (SD 52.56) scans by the Dementia Angels.

XSL•FO
RenderX

Table . Characteristics of the getting lost episodes logged from real-life data (n=111).

Characteristics	Total
Caregiver demographics, n (%)	
Age (≥65 y)	20 (18)
Sex (male)	25 (22.5)
Spouse	28 (25.2)
Child	78 (70.3)
Lives with CR	64 (57.7)
CR[a] demographics, n (%)	
Age (≥80 y)	62 (55.9)
Sex (male)	47 (42.3)
CR is moderate- or late-stage dementia	69 (62.2)
Mobility—no aid needed	82 (73.9)
Mobility—stick	29 (26.1)
Search data, mean (SD)	
Count of Dementia Angels scans	18.59 (52.56)

[a]CR: care recipient.

Discussion

Principal Findings

This observational study indicates that crowdsourcing technology has mixed effects on caregivers of individuals with dementia at risk of getting lost. The technology was associated with a modest reduction in caregiver worry, a higher adjusted odds of proactive search efforts, and a higher adjusted odds of adoption of an IoT device or CCTV as a preventative strategy; however, paired comparison among BLE tag users showed only directional but nonsignificant increase in use of these measures. In addition, approximately 50% of real-world missing cases reported to have used the BLE tag to aid search, and the technology did not demonstrably reduce the time required to locate missing individuals. In terms of risk factor analysis, the progression of dementia emerged as a significant risk factor for getting lost.

Risk Factor for Getting Lost

Our findings align with existing research, revealing that approximately 35% of community-dwelling individuals with dementia have experienced getting lost [24,25]. The finding that individuals in the moderate or late stage of dementia are at higher risk of getting lost in the community, compared with individuals in the mild stage of dementia, is consistent with other studies [12,26,27]. This increased risk is likely due to the progressive impairment of insight [28] and the exacerbation of behavioral and psychological symptoms, such as agitation, apathy, and irritability, commonly observed in major types of dementia [29,30]. These symptoms are known as risk factors for wandering [31], emphasizing the need for tailored support and preventative strategies for caregivers throughout their caregiving journey.

Intervention Effects on Behavioral and Psychological Changes

The crowdsourcing technology trialed in this study was associated with a reduction in caregiver worry and higher adjusted odds of promoting proactive engagement in search and prevention. Consistent with previous systematic reviews, our study suggests that IoT technology provides a sense of reassurance, mitigating caregiver anxiety by offering a means to locate care recipients who become lost [32].

The observed group effect, wherein BLE tag users exhibited higher levels of worry and distress related to getting lost, underscores their propensity to use the technology during such episodes and highlights the need for targeted support for these caregivers. The subsequent decline in worry and distress over time, particularly the significant time effect in the worry outcome, regardless of BLE tag use, may be attributed to the availability of study interveners who provided practical, solution-focused support during missing episodes. This underscores the potential benefits of a facilitator-aided approach in managing getting lost events.

Notably, while worry significantly decreased over time for BLE tag users, the reduction in distress was not significant. This disparity may reflect the distinct nature of these constructs: distress being an immediate affective response, while worry is a more cognitive construct linked to problem-solving and control [33]. The technology may not immediately alleviate distress, but it empowers caregivers to access resources, thereby reducing worry. This is supported by the finding that BLE tag users had a higher adjusted odds of adopting proactive search measures, such as going out to search, calling police, or using media or social media, potentially enhancing their sense of control over the situation and reducing anticipatory worry. Paired comparison among BLE tag users also showed a directional increase in using these measures, but the increase did not reach statistical

XSL•FO
RenderX

significance, possibly because of the small subsample size. The higher adjusted odds of seeking help from media or social media among BLE tag users may be because they are more likely to normalize tracking technology through exchange with other dementia caregivers on these platforms [34]. Social media is also a primary and rapid channel used by police to mobilize the public during missing-person incidents [35]; therefore, it naturally complements BLE-enabled search routines. However, this association should be interpreted cautiously because the low event count (<10) may result in imprecise and unstable odds ratio, suggested by the wide CI.

Moreover, BLE tag users had a higher adjusted odds of adopting person-centered preventative strategies over time. Caregivers often prioritize safety over liberty [36,37], sometimes resorting to restrictive measures, such as forbidding the care recipients to go outdoors [13], that negatively impact the quality of life of individuals with dementia. The technology developed in this trial aims to address this by providing a positioning solution that is acceptable to individuals with dementia, thereby encouraging caregivers to shift toward less restrictive strategies, such as using CCTV or IoT devices rather than preventing outdoor activities altogether. Paired comparison showed an insignificant, directional increase in the use of IoT device or CCTV among BLE tag users. However, the proportion of BLE tag users who continued to use the prevention of outdoor activities as a preventive measure remained unchanged, which may reflect the higher worry and distress observed in this group. While the results of paired comparison should be interpreted cautiously because of small subsample size, these results suggest that technology adoption alone may be insufficient to reduce restrictive practices, and additional caregiver support may be needed to fully promote greater autonomy and quality of life for people living with dementia. In terms of privacy concerns, the crowdsourcing technology balances safety concerns with privacy by activating positioning only during a missing episode, initiated by the caregiver. This design respects the autonomy of individuals with dementia while providing a safety net [38,39].

The majority of both BLE tag users and nonusers located their missing care recipients within 24 hours, with 30% to 40% of missing care recipients found within 1 hour. However, there was an insignificantly higher portion of missing care recipients who did not use the BLE tag that were found within 1 hour. This may be due to the effectiveness of overall search strategies and the study being conducted in a densely populated area, which inherently shortens the time of searching. The significant difference in time taken to find the missing care recipients may be related to the underlying risk profiles of those using or not using the BLE tag, as suggested by the higher caregiver worry observed in the BLE tag group.

Usability and Acceptance

Findings on usability and acceptance indicated a high willingness among individuals with dementia to carry the BLE tag and a high level of caregiver trust in data security. The design of the BLE tag as a keychain, pass, or attachment to mobility aids likely contributed to its acceptance. This aligns with literature suggesting that small, unobtrusive designs enhance user acceptance [39,40]. Our BLE tag design adheres to this principle, ensuring that individuals with dementia only need to carry the device naturally, with caregivers initiating the search if a missing episode occurs. Additionally, the reassurance provided by the technology may alleviate distress associated with outdoor activities. A systematic review revealed that individuals with dementia feel distressed when confined indoors but also fear getting lost [41], leading them to restrict their outdoor activities [42,43]. Technology offering reassurance can reduce this fear and distress, encouraging more frequent and extensive outdoor activities [37,44], thus promoting autonomy, a key element of quality of life. Indeed, some individuals with dementia expressed a discrepancy between their perception of their outdoor capabilities and the restrictive measures imposed by caregivers due to safety concerns [45]. IoT technology can mitigate this potential conflict by providing reassurance to both caregivers and care recipients.

However, caregivers expressed lower satisfaction with positioning accuracy, a key determinant of use satisfaction and adherence [46]. This limitation arises from the technology's reliance on wide coverage of sensors to detect the BLE tag and relay location information. Greater sensor coverage would result in more accurate positioning. This project aimed to promote dementia-friendliness and invited citizens to install the mobile app, turning them into mobile sensors. Despite more than 47,000 downloads, coverage was insufficient, and technology usage was mainly by caregivers themselves during lost episodes. Given that individuals with dementia who get lost in the community are often found by people other than caregivers [15,24], our findings suggest the vital importance of public education on dementia-friendliness to build wider sensor coverage and increase the accuracy of this technology.

A key postprogram initiative to enhance coverage and improve positioning accuracy involved collaborating with corporate partners to deploy sensors across public spaces frequently visited by people with cognitive impairment, such as shopping malls, transport hubs, and retail shops [47]. Beyond installing sensors to support search efforts, we also implemented tailored training programs for staff in the transport, banking, and property sectors. These trainings aim to improve their ability to respond to missing-person incidents, fostering a more dementia-friendly community.

Real-Life Use

The real-life data revealed that approximately 50% of lost episodes used the technology to aid the search, and these searches were predominantly conducted by caregivers aged <65 years, mostly women and often children of the care recipients, suggesting that this crowdsourcing technology is concentrated among younger, tech-savvy family members with direct caregiving responsibilities. The predominance of female caregivers and adult children further suggests that this demographic, often balancing multiple responsibilities, may derive particular benefit from such technology.

The care recipients are typically older adults who are often in the moderate or late stage of dementia but maintain good mobility. These traits highlight a target group prone to wandering and physically capable of straying significant

XSL•FO
RenderX

distances, amplifying the need for an effective positioning solution. Beyond this cohort, the system's reliance on a battery-saving tag, a mobile sensor interface, and a community endeavor for searching suggests its potential for scalability across broader dementia care apps, such as in residential facilities or community settings where the risk of getting lost persists.

Clinical Implications

The findings of this study underscore the importance of implementing IoT technology as part of comprehensive dementia-friendliness plans for individuals with dementia who are at risk of getting lost. Health care providers should consider recommending such technology as part of a care plan for community-dwelling individuals with dementia. It is crucial to educate and encourage caregivers about the potential benefits and limitations of positioning technology, such as the BLE tag, so that they can effectively use these tools to aid in the search for lost care recipients, potentially reducing the duration of lost episodes. The real-life data show that caregivers aged >65 years, who may be less tech-savvy, could underuse the system. As the population ages and the number of older adult caregivers, who have higher worry about their literacy in using assistive technology, increases [48,49], training and support programs should be available to help caregivers navigate and effectively use assistive technology, mitigating initial worry and improving overall outcomes [39,50].

Several risk factors for getting lost in the context of dementia were identified in this study. Health care providers could identify at-risk individuals with dementia and tailor interventions or support accordingly. Additionally, as caregivers may not always be aware of the preceding signs of getting lost [11], health care providers should facilitate caregiver awareness through education and case management. Getting lost is often an adverse outcome for individuals with dementia who may seek to fulfill unmet physical or psychological needs [10]; rather than restricting mobility, a more person-centered care approach should address these needs in an unobtrusive way that facilitates autonomy while ensuring safety. This approach should include education and assessment of ethical issues, ensuring that the privacy and liberty of individuals with dementia are respected [51].

The study found that individuals with dementia were willing to use assistive technology, likely because they were aware of the risks and sought to manage these risks independently without burdening their caregivers, and this attempt should be respected and supported [52]. Additionally, research exploring the daily mobility patterns of individuals with dementia suggests that they tend to navigate predictably based on their daily routines [53], implying that caregivers who are knowledgeable about the daily whereabouts of their care recipients are more likely to find their missing care recipients. Furthermore, the benefit of assistive IoT technology during lost episodes can be multiplied when the search area is more confined.

Strengths and Limitations

This study has several strengths. First, it was conducted in a real-world setting, which enhances the generalizability of the findings. Second, the large sample size provides robust results on the usability outcomes and factors associated with getting lost. Third, it included the usability and acceptance outcomes, such as device comfort and ease of use of the app [54], which added evidence to the daily life usability testing in prior studies [55]. However, the study also has limitations. First, the single-arm study design, as opposed to a randomized controlled trial design, limits the ability to draw causal inferences and account for potential confounders into consideration. We adopted the observational study design because we would like to observe the use of this technology in a naturalistic and exploratory manner; the findings of this study support further examination using a more robust study design. Second, the participants in this study were caregivers of individuals with dementia. Their responses regarding the satisfaction of care recipients may be subject to bias, as they are informants rather than the direct recipients of the intervention. Third, this study only captured data on the most recent episode of getting lost; outcomes related to after-care arrangements, caregiver distress, and worry may be time dependent. Therefore, future studies should consider a longitudinal design to examine the trajectories of these outcomes over time.

Conclusions

This study suggests that BLE-based crowdsourcing technology may support caregivers of individuals with dementia by reducing worry about future getting lost episodes over time and encouraging proactive search strategies. Caregivers reported high satisfaction with usability and data security, although concerns about positioning accuracy highlight the need for broader community adoption to enhance effectiveness. While the technology did not significantly shorten the duration to find the missing care recipients, its integration into dementia care plans could complement existing strategies, balancing safety with autonomy. Limitations, including the observational design and reliance on caregiver reports, underscore the need for randomized controlled trials and perspectives from individuals with dementia to confirm these findings and evaluate long-term benefits.

Acknowledgments

The authors would like to thank all the research participants for their contribution to this study.

Funding

The authors would like to acknowledge the Hong Kong Jockey Club Charities Trust for its generous donation on the development of the "Caring Communities for Dementia Campaign."

XSL•FO
RenderX

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The co-authors of the Hong Kong University of Science and Technology are the developers of the technology in this study.

Multimedia Appendix 1
Paired comparison of search strategies and post–getting lost care arrangement of Bluetooth Low Energy tag users and nonusers.
[DOCX File, 36 KB - humanfactors_v12i1e73670_app1.docx]

References

1. Dementia. World Health Organization. 2025. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2025-03-31]
2. Hugo J, Ganguli M. Dementia and cognitive impairment: epidemiology, diagnosis, and treatment. Clin Geriatr Med 2014 Aug;30(3):421-442. [doi: 10.1016/j.cger.2014.04.001] [Medline: 25037289]
3. Beavers-Kirby JR, Segal-Gidan FI. Gerontology and Geriatrics for NPs and PAs: An Interprofessional Approach: Elsevier; 2023. URL: https://evolve.elsevier.com/cs/product/9780323608459?role=student [accessed 2023-05-31]
4. Monacelli AM, Cushman LA, Kavcic V, Duffy CJ. Spatial disorientation in Alzheimer's disease: the remembrance of things passed. Neurology 2003 Dec 9;61(11):1491-1497. [doi: 10.1212/wnl.61.11.1491] [Medline: 14663030]
5. Ward R, Rummery K, Odzakovic E, et al. Getting lost with dementia: encounters with the time-space of not knowing. Health Place 2022 Nov;78:102940. [doi: 10.1016/j.healthplace.2022.102940] [Medline: 36403420]
6. Fillit HM, Rockwood K, Young JB, editors. Brocklehurst's Textbook of Geriatric Medicine and Gerontology: Elsevier; 2010. [doi: 10.1093/ageing/afx140]
7. Pai MC, Lee CC. The incidence and recurrence of getting lost in community-dwelling people with Alzheimer's disease: a two and a half-year follow-up. PLoS One 2016;11(5):e0155480. [doi: 10.1371/journal.pone.0155480] [Medline: 27183297]
8. Puthusseryppady V, Morrissey S, Spiers H, Patel M, Hornberger M. Predicting real world spatial disorientation in Alzheimer's disease patients using virtual reality navigation tests. Sci Rep 2022 Aug 4;12(1):13397. [doi: 10.1038/s41598-022-17634-w] [Medline: 35927285]
9. Jeong JG, Song JA, Park KW. A relationship between depression and wandering in community-dwelling elders with dementia. Dement Neurocogn Disord 2016 Mar;15(1):1-6. [doi: 10.12779/dnd.2016.15.1.1] [Medline: 30906332]
10. Cipriani G, Lucetti C, Nuti A, Danti S. Wandering and dementia. Psychogeriatrics 2014 Jun;14(2):135-142. [doi: 10.1111/psyg.12044] [Medline: 24661471]
11. Chung JCC, Lai CKY. Elopement among community-dwelling older adults with dementia. Int Psychogeriatr 2011 Feb;23(1):65-72. [doi: 10.1017/S1041610210000657] [Medline: 20602859]
12. Kwok TCY, Yuen KSL, Ho FKY, Chan WM. Getting lost in the community: a phone survey on the community‐dwelling demented people in Hong Kong. Int J Geriat Psychiatry 2010 Apr;25(4):427-432. [doi: 10.1002/gps.2361] [Medline: 19653244]
13. Li SH, Wu SFV, Liu CY, Lin CF, Lin HR. Experiences of family caregivers taking care getting lost of persons with dementia: a qualitative study. BMC Psychiatry 2024 Jun 18;24(1):452. [doi: 10.1186/s12888-024-05891-0] [Medline: 38890607]
14. Byard RW, Langlois NEI. Wandering dementia—a syndrome with forensic implications. J Forensic Sci 2019 Mar;64(2):443-445. [doi: 10.1111/1556-4029.13885] [Medline: 30080942]
15. Larsson M, Årestedt K, Svensson A, Andersson H, Wolmesjö M. Missing incidents and the risk of harm in persons living with dementia reported to the Swedish police- A nationwide retrospective registry study. BMC Geriatr 2025 Mar 5;25(1):153. [doi: 10.1186/s12877-025-05809-9] [Medline: 40045223]
16. Peng LM, Chiu YC, Liang J, Chang TH. Risky wandering behaviors of persons with dementia predict family caregivers' health outcomes. Aging Ment Health 2018 Dec;22(12):1650-1657. [doi: 10.1080/13607863.2017.1387764] [Medline: 29169254]
17. Górska S, Forsyth K, Maciver D. Living with dementia: a meta-synthesis of qualitative research on the lived experience. Gerontologist 2018 May 8;58(3):e180-e196. [doi: 10.1093/geront/gnw195] [Medline: 28069886]
18. Margot-Cattin I, Berchtold A, Gaber S, Kuhne N, Nygård L, Malinowsky C. Associations between community participation and types of places visited among persons living with and without dementia: risks perception and socio-demographic aspects. BMC Geriatr 2022 Apr 9;22(1):309. [doi: 10.1186/s12877-022-03015-5] [Medline: 35397532]
19. Tan J, Wong W, Zhu X, Wu H, Chan SHG. Cooperative target tracking and signal propagation learning using mobile sensors. Proc ACM Interact Mob Wearable Ubiquitous Technol 2018 Sep 18;2(3):1-21. [doi: 10.1145/3264946]
20. Li Q. Analysis of innovation opportunities for airtag. Presented at: 2021 3rd International Conference on Economic Management and Cultural Industry (ICEMCI 2021); Oct 22-24, 2021. [doi: 10.2991/assehr.k.211209.243]

XSL•FO
RenderX

21. Jang HD, Ibrahim H, Asim R, Varvello M, Zaki Y. A tale of three location trackers: AirTag, SmartTag, and Tile. arXiv. Preprint posted online on Jan 29, 2025. [doi: 10.48550/arXiv.2501.17452]
22. Demers L, Monette M, Lapierre Y, Arnold DL, Wolfson C. Reliability, validity, and applicability of the Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST 2.0) for adults with multiple sclerosis. Disabil Rehabil 2002;24(1-3):21-30. [doi: 10.1080/09638280110066352] [Medline: 11827151]
23. Survey and behavioural research ethics: guidelines for survey and behavioural research ethics. The Chinese University of Hong Kong. 2025. URL: https://www.gs.cuhk.edu.hk/download/VII-C.pdf [accessed 2025-12-06]
24. Hopkinson JB, King A, Mullins J, et al. What happens before, during and after crisis for someone with dementia living at home: a systematic review. Dementia (London) 2021 Feb;20(2):570-612. [doi: 10.1177/1471301220901634] [Medline: 32013558]
25. Miguel-Cruz A, Perez H, Rutledge E, Daum C, Liu L. Factors associated with wandering among persons with dementia: a retrospective study. Innov Aging 2023 Dec 21;7(Supplement_1):579. [doi: 10.1093/geroni/igad104.1894]
26. Perez H, Cruz AM, Neubauer N, et al. Risk factors associated with missing incidents among persons living with dementia: a scoping review. Can J Aging 2024 Feb 1;43(3):1-15. [doi: 10.1017/S0714980823000776] [Medline: 38297497]
27. Murata S, Takegami M, Ogata S, et al. Joint effect of cognitive decline and walking ability on incidence of wandering behavior in older adults with dementia: a cohort study. Int J Geriat Psychiatry 2022 May;37(5). [doi: 10.1002/gps.5714] [Medline: 35451122]
28. Harwood DG, Sultzer DL, Wheatley MV. Impaired insight in Alzheimer disease: association with cognitive deficits, psychiatric symptoms, and behavioral disturbances. Neuropsychiatry Neuropsychol Behav Neurol 2000 Apr;13(2):83-88. [Medline: 10780626]
29. Kazui H, Yoshiyama K, Kanemoto H, et al. Differences of behavioral and psychological symptoms of dementia in disease severity in four major dementias. PLoS One 2016;11(8):e0161092. [doi: 10.1371/journal.pone.0161092] [Medline: 27536962]
30. Cruz AM, Perez H, Jantzi M, Liu L, Hirdes JP. Pan-Canadian estimates of the prevalence and risks associated with critical wandering among home care clients. Alzheimers Dement 2024 Oct;20(10):7079-7089. [doi: 10.1002/alz.14195] [Medline: 39175391]
31. Barnard-Brak L, Parmelee P. Measuring risk of wandering and symptoms of dementia via caregiver report. J Appl Gerontol 2021 Oct;40(10):1372-1376. [doi: 10.1177/0733464820947287] [Medline: 32757798]
32. Doyle M, Nwofe ES, Rooke C, Seelam K, Porter J, Bishop D. Implementing global positioning system trackers for people with dementia who are at risk of wandering. Dementia (London) 2024 Aug;23(6):964-980. [doi: 10.1177/14713012241248556] [Medline: 38646927]
33. Zebb BJ, Beck JG. Worry versus anxiety: Is there really a difference? Behav Modif 1998 Jan;22(1):45-61. [doi: 10.1177/01454455980221003] [Medline: 9567736]
34. Sun Y, Kim HM, Xu Y, et al. GPS tracking in dementia caregiving: social norm, perceived usefulness, and behavioral intent to use technology. Presented at: Hawaii International Conference on System Sciences; Jan 4-9, 2021. [doi: 10.24251/HICSS.2021.461]
35. Neubauer NA, Miguel-Cruz A, Liu L. Strategies to locate lost persons with dementia: a case study of Ontario first responders. J Aging Res 2021;2021(1):5572764. [doi: 10.1155/2021/5572764] [Medline: 34094601]
36. White EB, Montgomery P. Electronic tracking for people with dementia: an exploratory study of the ethical issues experienced by carers in making decisions about usage. Dementia (London) 2014 Mar 1;13(2):216-232. [doi: 10.1177/1471301212460445] [Medline: 24599815]
37. Löbe C, Petersen N. Between empowerment, patronization, and surveillance. a semi-structured interview study with persons with dementia and family caregivers on the empowering opportunities and perils of intelligent assistive technologies. BMC Med Ethics 2025 Apr 5;26(1):44. [doi: 10.1186/s12910-025-01203-7] [Medline: 40188332]
38. Shaik MA, Anik FI, Hasan MM, et al. Advancing remote monitoring for patients With Alzheimer disease and related dementias: systematic review. JMIR Aging 2025 May 14;8:e69175. [doi: 10.2196/69175] [Medline: 40367504]
39. Boyle LD, Husebo BS, Vislapuu M. Promotors and barriers to the implementation and adoption of assistive technology and telecare for people with dementia and their caregivers: a systematic review of the literature. BMC Health Serv Res 2022 Dec 23;22(1):1573. [doi: 10.1186/s12913-022-08968-2] [Medline: 36550456]
40. Dequanter S, Fobelets M, Steenhout I, et al. Determinants of technology adoption and continued use among cognitively impaired older adults: a qualitative study. BMC Geriatr 2022 Apr 28;22(1):376. [doi: 10.1186/s12877-022-03048-w] [Medline: 35484488]
41. Petty S, Harvey K, Griffiths A, Coleston DM, Dening T. Emotional distress with dementia: a systematic review using corpus‐based analysis and meta‐ethnography. Int J Geriat Psychiatry 2018 May;33(5):679-687. [doi: 10.1002/gps.4870] [Medline: 29498777]
42. Puthusseryppady V, Morrissey S, Aung MH, Coughlan G, Patel M, Hornberger M. Using GPS tracking to investigate outdoor navigation patterns in patients with Alzheimer disease: cross-sectional study. JMIR Aging 2022 Apr 21;5(2):e28222. [doi: 10.2196/28222] [Medline: 35451965]

XSL•FO
RenderX

43. Brittain K, Corner L, Robinson L, Bond J. Ageing in place and technologies of place: the lived experience of people with dementia in changing social, physical and technological environments. Sociol Health Illn 2010 Feb 1;32(2):272-287. [doi: 10.1111/j.1467-9566.2009.01203.x] [Medline: 20003041]
44. Thorpe J, Forchhammer BH, Maier AM. Adapting mobile and wearable technology to provide support and monitoring in rehabilitation for dementia: feasibility case series. JMIR Form Res 2019 Oct 17;3(4):e12346. [doi: 10.2196/12346] [Medline: 31625951]
45. Førsund LH, Grov EK, Helvik AS, Juvet LK, Skovdahl K, Eriksen S. The experience of lived space in persons with dementia: a systematic meta-synthesis. BMC Geriatr 2018 Feb 1;18(1):33. [doi: 10.1186/s12877-018-0728-0] [Medline: 29390970]
46. Dale Ø. Usability and usefulness of GPS based localization technology used in dementia care. Presented at: Computers Helping People with Special Needs; Jul 14-16, 2010. [doi: 10.1007/978-3-642-14097-6_48]
47. Sohlberg MM, Todis B, Fickas S, Hung PF, Lemoncello R. A profile of community navigation in adults with chronic cognitive impairments. Brain Inj 2005 Dec 20;19(14):1249-1259. [doi: 10.1080/02699050500309510] [Medline: 16286341]
48. Chen YC, Leung CY. A study on the lost seeking devices and systems for dementia-patients. Presented at: CENTRIC 2011: International Conference on eHealth, Telemedicine, and Social Medicine; Oct 23-29, 2011 URL: https://www.thinkmind.org/articles/centric_2011_1_30_30038.pdf [accessed 2025-12-06]
49. Sriram V, Jenkinson C, Peters M. Informal carers' experience of assistive technology use in dementia care at home: a systematic review. BMC Geriatr 2019 Jun 14;19(1):160. [doi: 10.1186/s12877-019-1169-0] [Medline: 31196003]
50. Shen H, Han Y, Shi W, et al. Perspectives and experiences of family caregivers using supportive mobile apps in dementia care: meta-synthesis of qualitative research. JMIR mHealth mHealth 2025 Jun 18;13:e65983. [doi: 10.2196/65983] [Medline: 40532203]
51. Bantry-White E. Supporting ethical use of electronic monitoring for people living with dementia: Social work's role in assessment, decision-making, and review. J Gerontol Soc Work 2018 Apr;61(3):261-279. [doi: 10.1080/01634372.2018.1433738] [Medline: 29381128]
52. Bartlett R, Brannelly T. On being outdoors: How people with dementia experience and deal with vulnerabilities. Soc Sci Med 2019 Aug;235:112336. [doi: 10.1016/j.socscimed.2019.05.041] [Medline: 31272077]
53. Bayat S, Mihailidis A. Outdoor life in dementia: How predictable are people with dementia in their mobility? Alzheimers Dement (Amst) 2021 Jan;13(1):e12187. [doi: 10.1002/dad2.12187] [Medline: 34027017]
54. Moorthy P, Weinert L, Schüttler C, et al. Attributes, methods, and frameworks used to evaluate wearables and their companion mHealth apps: scoping review. JMIRmHealth mHealth 2024 Apr 5;12:e52179. [doi: 10.2196/52179] [Medline: 38578671]
55. Koo BM, Vizer LM. Examining mobile technologies to support older adults with dementia through the lens of personhood and human Needs: scoping review. JMIR mHealth uHealth 2019 Nov 11;7(11):e15122. [doi: 10.2196/15122] [Medline: 31710305]

Abbreviations

AOR: adjusted odds ratio
BLE: Bluetooth Low Energy
IoT: Internet of Things

Edited by E Borycki; submitted 10.03.25; peer-reviewed by H Tsang, LI Lopera Gonzalez; revised version received 19.11.25; accepted 22.11.25; published 30.12.25.

Please cite as:
Wong B, Cheng T, Fung N, Lin Z, Lai KK, Ho F, Chan SHG, Kwok T
Using Crowdsourcing Internet of Things Technology to Reduce Caregiver Worry in Dementia-Related Lost Episodes: Longitudinal Observational Study
JMIR Hum Factors 2025;12:e73670
URL: https://humanfactors.jmir.org/2025/1/e73670
doi:10.2196/73670

© Bel Wong, Tobi Cheng, Nicole Fung, Zhongming Lin, Ki-Kit Lai, Florence Ho, S-H Gary Chan, Timothy Kwok. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 30.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Mobile Health App for Adults with Persisting Postconcussion Symptoms: Development and Usability Study

Gøril Storvig[1,2,3], MPhil, RN; Anker Stubberud[3,4], MD, PhD; Johanne Rauwenhoff[1,3], MSc, PhD; Liv Marie Rønhovde[2,5], cand.psychol; Martijn Smits[1], MPhil; Simen Berg Saksvik[5], MSc, PhD; Toril Skandsen[2,3,4], MD, PhD; Erling Tronvik[3,4,6], MD, PhD; Alexander Olsen[1,2,3], Cand Psychol, PhD

[1]Department of Psychology, Norwegian University of Science and Technology, Trondheim, Norway

[2]Clinic of Rehabilitation, St Olav's University Hospital, Trondheim, Norway

[3]Norwegian Center for Headache Research, Trondheim, Norway

[4]Department of Neuromedicine and Movement Science, Norwegian University of Science and Technology, Trondheim, Norway

[5]Department of Research and Development, St Olav's University Hospital, Trondheim, Norway

[6]Department of Neurology, St Olav's University Hospital, Trondheim, Norway

Corresponding Author:
Alexander Olsen, Cand Psychol, PhD
Department of Psychology, Norwegian University of Science and Technology, Trondheim, Norway

Abstract

Background: Diagnostics, treatment, and research of persisting postconcussion symptoms are challenging. Assessing symptoms is essential, but currently implemented methods only allow for retrospective reporting of symptoms. A mobile health (mHealth) symptom mapping app for adults with persisting postconcussion symptoms may be an accessible and cost-efficient alternative.

Objective: This study aimed to develop a research-based mobile app for symptom mapping for adults with persisting postconcussion symptoms and investigate its usability, feasibility, and safety.

Methods: This was a mixed method development and usability study consisting of three iterative cycles, each including (1) app design and programming, (2) app usability evaluation by the user group, and (3) app review by the clinician group. The outcomes were the mHealth App Usability Questionnaire and Mobile App Rating Scale scores, the number of days with logged symptom data during a home-testing period, and descriptions of adverse events throughout the study period. Semistructured interviews were conducted to explore the user group's experiences further.

Results: Twenty-three adults with persisting postconcussion symptoms (median age 52, IQR 34-59 years; 70% female) were included in the user group. Six clinicians (median age 53, IQR 35-60 years), including 3 (50%) females, with a mean of 13 (SD 7) years of experience working with individuals with persisting postconcussion symptoms, were included in the clinician group. The app received a mean score of 5 (SD 1.1) on the mHealth App Usability Questionnaire (7-point Likert scale) from the user group and 4.1 (SD 0.4) on the Mobile App Rating Scale (5-point Likert Scale) from the clinician group. During the 28-day home-testing period, the adherence rate among the participants in the user group was 89% (IQR 78 - 96), and two adverse events related to increased symptom awareness were registered. Three themes were created through reflexive thematic analysis of the qualitative data: (1) Visualizing the invisible—Enabling reflection and insight; (2) Personalized yet simple—Balancing relevance and usefulness; and (3) More than just a number—The complexity behind the symptom scores.

Conclusions: We developed a research-based symptom mapping app for people with persisting postconcussion symptoms. The app received high usability ratings from both the user and clinician groups. The app is a feasible alternative to traditional symptom mapping methods, and it is safe to use for its intended purpose.

Trial Registration: ClinicalTrials.gov NCT05635656; https://clinicaltrials.gov/study/NCT05635656

(JMIR Hum Factors 2025;12:e75323) doi:10.2196/75323

KEYWORDS

mTBI; mild traumatic brain injury; concussion; postconcussion syndrome; postconcussive syndrome; postconcussive symptoms; persistent postconcussion symptoms; mobile health; symptom tracking; symptom monitoring; persisting symptoms after concussion; PSaC

XSL•FO
RenderX

Introduction

Mild traumatic brain injury (mTBI) can be defined as a trauma to the head accompanied by acute signs of reduced cerebral functioning, such as brief loss of consciousness, confusion, or amnesia for the event, but often with no neuroradiological findings [1]. Common causes of civilian mTBI are falls, bicycle accidents, sports, motor vehicle accidents, being struck by objects, and violence [2]. People with mTBI can experience symptoms such as headaches, dizziness, nausea, fatigue, light sensitivity, sound sensitivity, trouble concentrating, attention problems, sadness, irritability, anxiety, depression, and sleep problems [3,4]. Most people recover within a few weeks after the injury, but a considerable proportion develop persisting postconcussion symptoms (PPCS) where symptoms persist for more than 3 months [5]. The reported prevalence of PPCS differs between contexts and study procedures, and a recent review found an estimated prevalence between 18.3% and 31.3%, depending on the applied definition of PPCS [6,7].

Despite the high prevalence, the etiology of PPCS is largely unknown and is an ongoing topic of debate among experts in the field [4,8]. Several factors, including pre, peri, and postinjury factors, are believed to influence or contribute to PPCS [9], and a biopsychosocial framework is often applied to understand these intricate processes [10,11]. PPCS is associated with reduced health-related quality of life [12], reduced return-to-work rate [13-15], and reduced participation in social life [16]. Regardless of the pathophysiological mechanisms involved, PPCS is a substantial health problem affecting the lives of many people.

Despite the high burden, we lack consensus on the diagnosis and classification of PPCS [4,17]. No single test can confirm PPCS; the diagnosis is established after careful evaluation of the medical history and self-reported symptoms [18,19]. Differential diagnosis can be particularly challenging, as similar symptoms to PPCS have also been reported in the general population [20].

Self-reporting of symptoms is essential in diagnosing and treating PPCS and is an important data source in research aiming to expand our understanding of the condition and evaluate treatment effects. Existing validated and widely used instruments for assessment of postconcussion symptoms, such as the Rivermead Postconcussion Symptom Questionnaire (RPQ) and the British Columbia Postconcussion Symptom Inventory, only allow for retrospective reporting, using a timescale of days or weeks [7,21,22]. While retrospective symptom assessment has pragmatic advantages, it may introduce recall bias, where the person reporting their symptoms fails to accurately recall the symptoms' occurrence or intensity [23]. This can potentially lead to an incomplete representation of the person's experienced symptoms, complicating the diagnosis and treatment of PPCS, as well as impeding research developments. Prospective symptom assessment, by comparison, has the benefit of the person reporting closer in time to symptom occurrence and may provide a more nuanced representation of symptom fluctuation and reduce the risk of recall bias [24].

An mHealth app for mapping PPCS is a highly warranted tool. It allows symptom registration in real time, capturing day-to-day fluctuations, and has the potential to support people with PPCS to take ownership of and manage their condition [25], support the clinician in diagnosing and evaluating treatment effects, and provide prospectively collected data for clinical trials. Furthermore, a mobile app is easily distributed, making it an accessible and cost-efficient alternative to traditional methods [26]. Therefore, we aimed to develop a mobile app for symptom mapping for people with PPCS and investigate its usability, feasibility, and safety.

Methods

Study Design and Participants

This was a mixed method iterative development and usability study carried out at St. Olavs University Hospital from December 25, 2022, to March 5, 2024. The mobile app was developed throughout the study based on feedback from people with PPCS (user group) and clinicians with expertise in PPCS (clinician group) at 3 different time points. The primary objective was to assess the app's usability, and the secondary objectives were to assess feasibility and safety.

The study participants in the user group were recruited from an outpatient rehabilitation clinic, through announcements on the hospital's official website, the Norwegian Center for Headache Research website, and user organizations and their social media. Inclusion criteria were: (1) age 18 years or older, (2) mTBI as defined by having sustained a head injury with observed or self-reported alteration of consciousness, amnesia, or other relevant acute symptoms, such as drowsiness, (3) PPCS according to the *International Classification of Diseases* research criteria operationalized as reporting 3 or more symptoms (including headache) on a moderate or greater intensity level on the RPQ, (4) posttraumatic headache according to the *International Classification of Headache Disorders 3rd edition*, (5) proficiency in the Norwegian language, and (6) signed written informed consent. Exclusion criteria were (1) severe psychiatric or somatic disease or other patient-related factors that would provide obvious challenges for adhering to the study protocol, including using the app, (2) having no access to an iOS (Apple Inc) or Android smartphone (Google LLC), or (3) having less than 3 months of smartphone experience. An experienced research nurse screened participants for eligibility by collecting relevant medical information and symptom history directly from the participants.

The study participants in the clinician group were identified based on their expertise in the field of PPCS and were recruited through convenience sampling using professional networks of the authors at Norwegian universities, hospitals, and clinics, aiming to include clinicians from diverse multidisciplinary backgrounds.

Study Measures

The primary outcome was app usability, which was measured using the mHealth App Usability Questionnaire (MAUQ) and Mobile App Rating Scale (MARS). The MAUQ consists of 21 statements divided into 3 subsections: Ease of use and

XSL•FO
RenderX

satisfaction, Interface and satisfaction, and Usefulness. The questionnaire has a 7-point Likert scale from 1 (strongly disagree) to 7 (strongly agree). The MARS is an app quality rating tool that assesses app quality on 4 dimensions: Engagement, Functionality, Aesthetics, and Information quality. The questionnaire uses a 5-point Likert scale, ranging from 1 (Inadequate) to 5 (Excellent). The total mean score describes the app's overall quality, while mean subscale scores within the dimensions can be used to describe specific strengths and weaknesses. Professional translators translated the MAUQ and the MARS into Norwegian using a forward-backward translation method. Experts in the field of mTBI performed the harmonization of the translations.

The RPQ was completed as part of the screening and inclusion of study participants in the user group, as well as before and after the home testing period. The RPQ contains 16 items representing different symptoms and operates on a 5-point Likert scale from 0 (no experience at all) to 4 (severe problem). When summing up the scores, responses of 1 (no more of a problem than before) were recoded to 0, with a total score range of 0 - 64. The Norwegian version of the RPQ has been translated from the original English version by the Collaborative European NeuroTrauma Effectiveness Research in Traumatic Brain Injury (CENTER-TBI) Consortium [27] and has been validated [28].

The secondary outcomes were feasibility and safety. To assess feasibility, we recorded the percentage of days with logged symptom data in the user group during the home-testing period of 28 days, referred to as adherence. For participants experiencing technical issues in the app hindering them from using it, the percentage was calculated as days with logged symptom data divided by the number of days it was possible to use the app. The app safety was assessed with the frequency of adverse events (AE), serious adverse events, and unexpected serious adverse events throughout the whole study period.

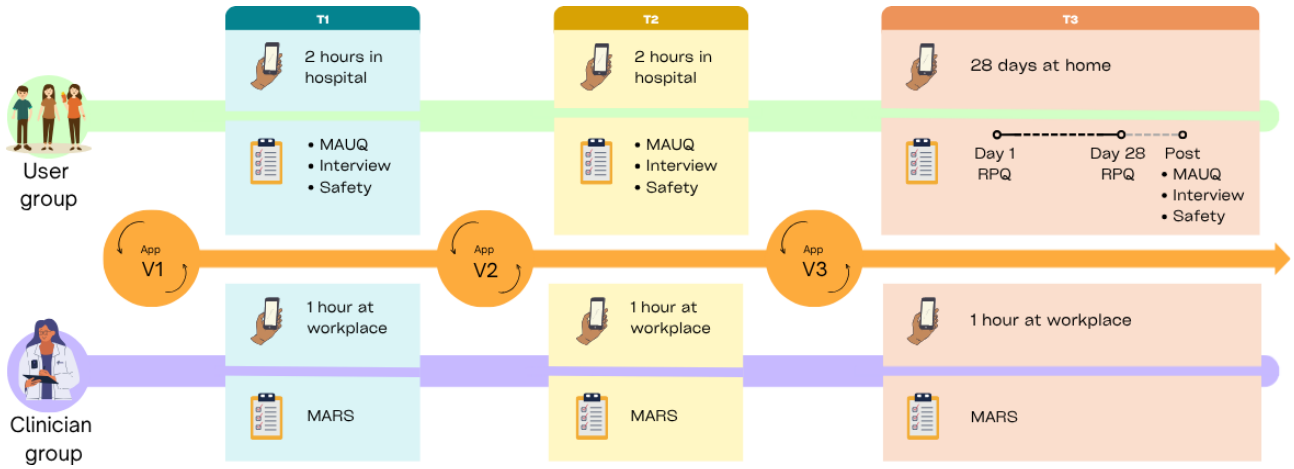
To further explore the app's usability, feasibility, and safety, semistructured interviews were conducted with all participants in the user group as part of each iteration.

Study Procedure

The study consisted of 3 iterative cycles. Each iteration included (1) app design and programming, (2) usability evaluation by the user group, and (3) review by the clinician group. During usability evaluation (2), the user group completed the MAUQ and conducted a semistructured interview, while the clinician group filled out the MARS during the app quality review (3). Before filling out the questionnaires, both the user and clinician groups had the opportunity to explore the app independently. Thereafter, they were instructed to perform specific tasks, typically describing an overarching goal, that is, "perform a registration of the symptoms you are experiencing today."

The study participants' immediate reactions and issues that emerged during the usability testing were also documented. The feedback from the user and clinician groups during usability testing was used to develop the next app version for the following iteration. In the first 2 iterations, the participants in the user group attended 2-hour individual sessions in a hospital setting to evaluate the app's usability. While in the third iteration, they used the app for 28 days at home (Figure 1).

Figure 1. Illustration of the study procedure. The study participants evaluated the different versions of the app (v1, v2, and v3) as part of every iteration. In the first two iterations, the user group evaluated v1 and v2 during 2-hour sessions at the hospital, completing the mHealth App Usability Questionnaire, an interview, and a safety assessment. In the third iteration, they used v3 at home for 28 days, completed the Rivermead postconcussion symptom questionnaire on the first and last day, and completed the mHealth App Usability Questionnaire, an interview, and a safety assessment within 2 weeks after the last day. The clinician group evaluated v1, v2, and v3 in 1-hour sessions at their workplace, completing the Mobile App Rating Scale during all iterations. MARS: Mobile App Rating Scale; MAUQ: mHealth App Usability Questionnaire; RPQ: Rivermead Postconcussion Symptom Questionnaire.



During the home testing, the user group was asked to use the app to register their symptoms daily. The app was preconfigured to send a notification every evening at 8 PM with the message "Did you experience symptoms today?" The study participants could adjust the notification time or disable it in the settings. They were able to register symptoms experienced within the last 48 hours but were not able to edit or add symptoms beyond that timeframe. Study personnel made 3 planned phone calls (approximately once a week) to assess any technical issues with the app. Additional phone calls were scheduled for further troubleshooting if technical issues were not solved immediately.

The clinician group evaluated the app's usability in 1-hour sessions in their work environments in every iteration.

Parallel to investigating the usability, feasibility, and safety of the symptom mapping app, the study participants also evaluated

XSL•FO
RenderX

a digital solution for biofeedback, as described in the preregistration of the research study Digital Solutions for Concussion [29]. The findings related to the biofeedback device will be reported elsewhere.

Data Analysis

The outcome measures for each cycle were summarized statistically. Continuous variables were summarized as mean (SD) or median (IQR) according to distribution. Categorical variables were summarized as frequencies (percentages). No statistical hypothesis testing was conducted.

The qualitative interviews were analyzed using reflexive thematic analysis, as described by Braun and Clarke [30], underpinned by a critical realism approach. The first author and fourth author conducted the qualitative interviews, which were later transcribed verbatim by the first author, fourth author, and 2 research assistants. The transcribed interviews were proofread by a research team member other than the transcriber. The first author kept a reflexive journal throughout this process, as well as in the later phases of the analytic process. The first author systematically coded the transcribed data semantically, applying code labels close to the expressed meanings. Next, initial themes were developed by exploring clusters of meaning across the dataset. Through reviewing and further developing the initial themes, the final themes were defined.

The quantitative analyses were performed in SPSS (SPSS Inc) and R (R Development Core Team), and the qualitative analyses were performed in NVivo (Lumivero).

App Development and Deployment

The app was programmed using React Native and made available for beta testing on both iOS and Android devices. St. Olavs Hospital and Norwegian University of Science and Technology (NTNU) contracted Nordic Brain Tech AS and Fominykh Consulting for consultant services on app programming. During usability testing, the study participants used the app on dedicated study phones provided by St. Olavs and NTNU. During the 28-day home-testing period, the user group used the app on their personal smartphones. They received written instructions and a QR code to download the app, and study staff assisted them by phone. Data collected with the app was transferred as deidentified data to Microsoft through the cloud service Microsoft Azure. The data were encrypted in transit and at rest. The identity and credentials were stored with the identity management service Microsoft Azure AD B2C. St. Olavs Hospital and NTNU have full ownership of the app, all data, images, and results created in the study.

Ethical Considerations

The study was approved by the Norwegian Medicinal Products Agency (NoMa 22/16250‐7) and the Norwegian Ethics Committees for Clinical Trials on Medicinal Products and Medical Devices (REC-KULMU 422538) and preregistered in ClinicalTrials (NCT05635656). The participants in the clinician group were identified based on their expertise in PPCS and recruited through convenience sampling using professional networks of the authors at Norwegian universities, hospitals, and clinics, aiming to include clinicians of diverse multidisciplinary backgrounds. Recruitment material for the user group was distributed through the hospital, patient organizations, and the Norwegian Centre for Headache Research (NorHead). Interested individuals made contact by email. All participants received oral and written information about the study, and informed written consent was obtained prior to the first study visit. Consent was collected electronically using eFORSK, a web-based application developed and maintained by Central Norway Regional Health Authority's IT department, Hemit HF. Participants were not compensated for their participation.

Results

Overview

A total of 23 adults (median age 52, IQR 34-59 years; 70% females) with PPCS were included in the user group, and 6 clinicians (median age 53, IQR 35-60 years; 50% females) were included in the clinician group. Table 1 displays demographic and injury-related characteristics of the user group. The most common injury mechanism was being struck by or on an object, whereas 5 cases were sport-related. Two participants in the user group dropped out during the study period. Both dropped out before commencing the home-testing period and stated personal reasons not related to the study as the reason for their withdrawals.

The participants in the clinician group (n=6) had a median age of 53 (IQR 35-60), and 50% (3/6) were women. They were medical doctors, specialists in physical medicine and rehabilitation, or clinical neuropsychologists with a mean of 13 (SD 7) years of experience in the field of PPCS.

XSL•FO
RenderX

Table . Demographic and injury-related variables for study participants in the user group (n=23).

Characteristics	Values
Age at inclusion (years), median (IQR)	52 (59 - 34)
Sex, n (%)	
Female	16 (70)
Highest level of completed education, n (%)	
Upper secondary, vocational	5 (22)
Upper secondary, general	4 (17)
Tertiary education	14 (61)
Injury mechanism, n (%)	
Fall	5 (22)
Struck by or on an object	12 (52)
Traffic accident	5 (17)
Violence	1 (4)
Time since injury (years), mean (SD)	5 (4)
History of previous concussions, yes, n (%)	8 (35)
Scores, mean (SD)	
Total score RPQ[a]	35.9 (7.0)
Subscore RPQ somatic	19.4 (3.3)
Subscore RPQ emotional	8 (4.0)
Subscore RPQ cognitive	8.4 (2.4)

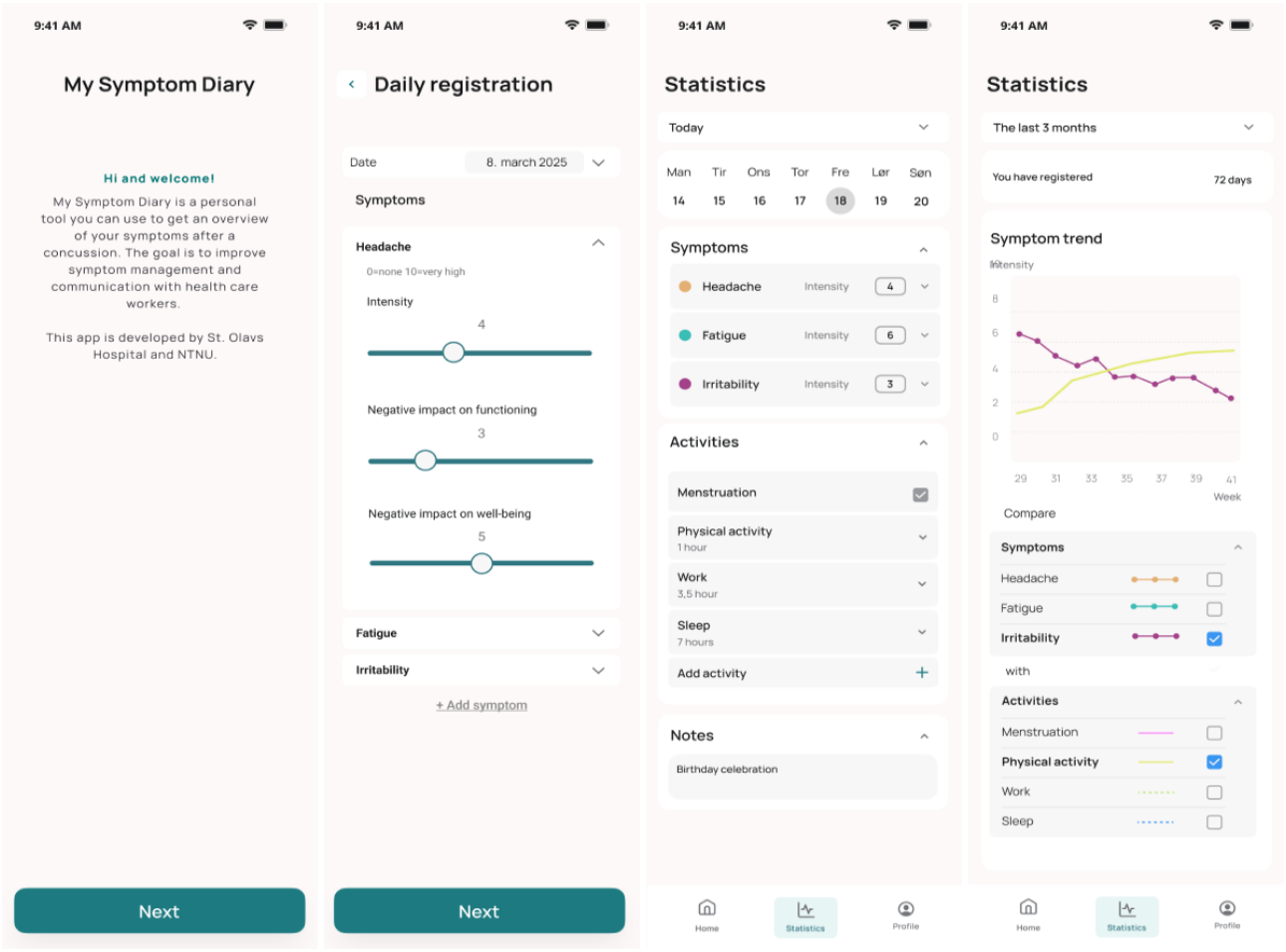
[a]RPQ: Rivermead Postconcussion Symptom Questionnaire.

App Development

Three versions of the symptom diary app were developed in the course of the study based on feedback from the study participants during usability testing. The third version of the app was used in the 28-day home-testing by the user group (Figure 2).

XSL•FO
RenderX

Figure 2. Screenshots from v3 of the app. Screenshot 1 shows the initial screen users see when opening the app for the first time. Screenshot 2 shows the first screen of a diary entry. The drop-down menu for each symptom contains 3 different parameters: intensity of symptoms, negative impact on functioning, and negative impact on well-being. Screenshot 3 shows an overview of the current day. Screenshot 4 shows statistics from the diary entries over the past 3 months in a line graph, with customization options to select which symptoms and activities to display. The study participants used a Norwegian version of the app, but for the purpose of this article, the screenshots have been translated into English.



Usability

The user group rated the third version of the app's usability 5 (SD 1.1) out of 7 on the MAUQ, with the highest scoring for Interface and satisfaction, followed by Ease of use and Usefulness (Figure 3).

The clinician group's mean rating of the app on the MARS was 4.1 (SD 0.4) out of 5, with Aesthetics scoring the highest, followed by Functionality, Engagement, and Information (Figure 4).

XSL•FO
RenderX

Figure 3. mHealth App Usability Questionnaire scores for the symptom mapping app, completed by the user group. MAUQ: mHealth App Usability Questionnaire.

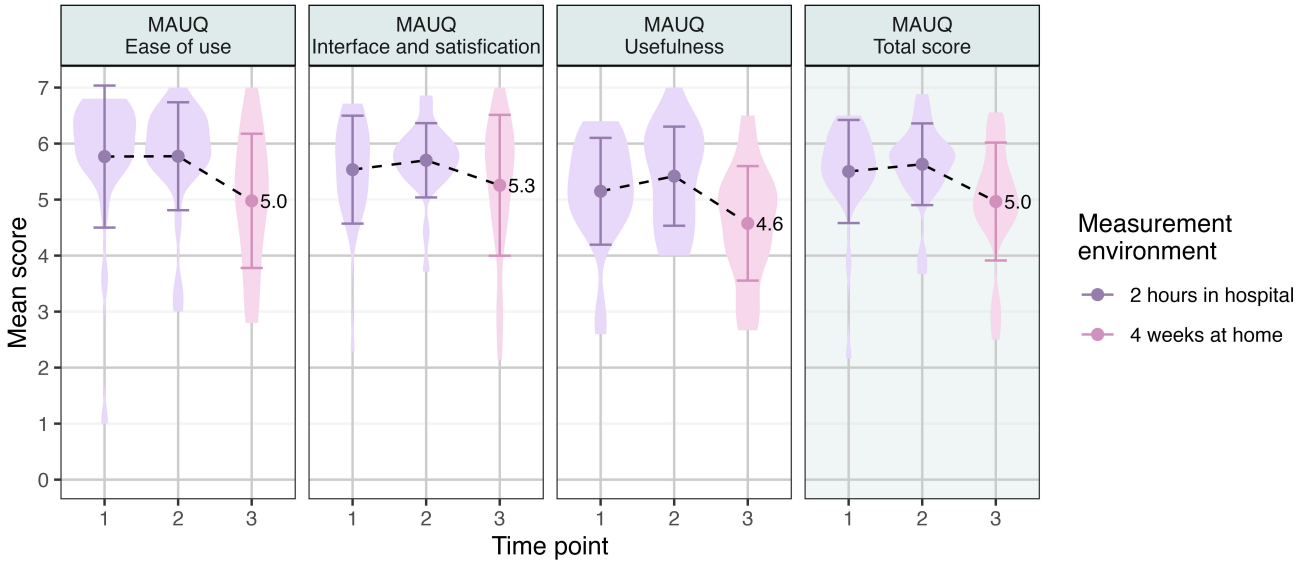
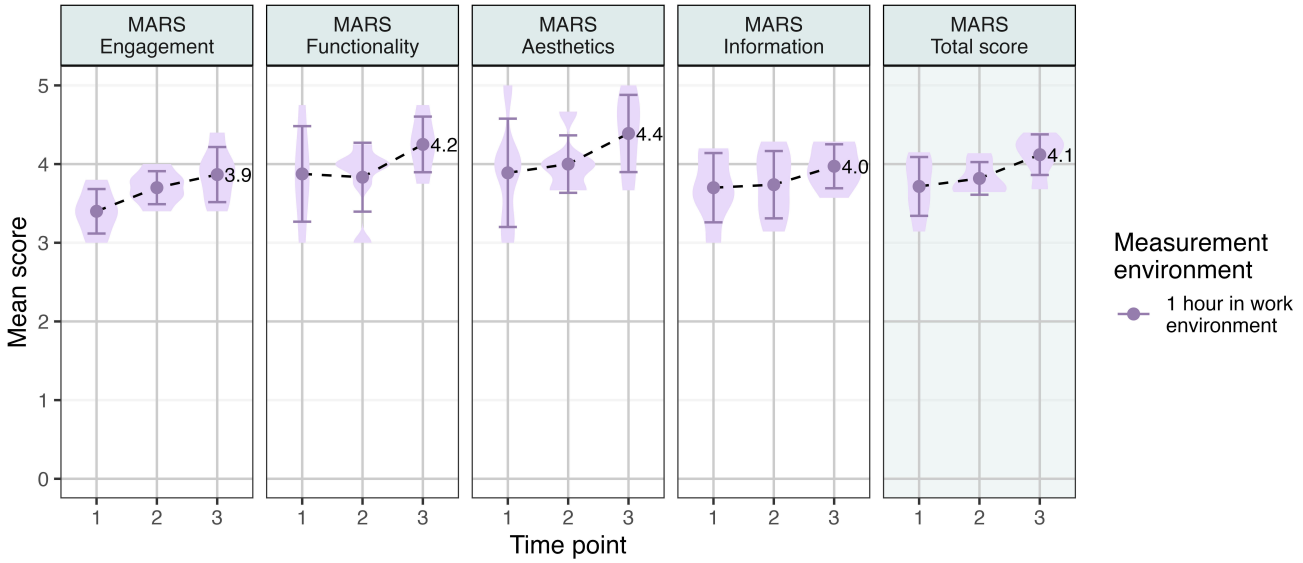


Figure 4. Mobile App Rating Scale scores for the symptom mapping app, completed by the clinician group. MARS: Mobile App Rating Scale.



Feasibility

A total of 21 participants in the user group completed the home-testing period, of whom 16 experienced technical issues of different severities during this time. There were 14 participants who experienced mild technical issues, such as being logged out of the app or the app "freezing," while 2 participants experienced severe technical problems where symptom entries were not saved due to a synchronization issue with the server. Despite the 2 participants' efforts using the app, significant amounts (possibly >85%) were not stored, and it was therefore not possible to accurately calculate their adherence rate. These were therefore excluded from the adherence analysis. The 2 excluded study participants with severe technical problems successfully completed all other study measures and were not excluded from other analyses. The adherence rate for the user group during the home-testing period was 89% (IQR 78-96), n=19. The adherence rate for all analyzed participants, independent of whether they experienced severe technical issues or not, was 81% (IQR 75-96; n=21).

Safety

Two AEs were logged during the home testing period. No severe AEs or unexpected severe events were logged. In both incidents, the study participants reported increased awareness of symptoms as an unwanted effect of using the symptom mapping app.

Qualitative Analysis and Results

Using reflexive thematic analysis through a 6-phase process as described by Braun and Clarke [30], 3 themes were constructed: (1) Visualizing the invisible—Enabling reflection and insight; (2) Personalized yet simple—Balancing relevance and usefulness; and (3) More than just a number—The complexity behind symptom scores.

Visualizing the Invisible—Enabling Reflection and Insight

The first theme surrounds the participants' need for a concrete visualization of their experiences with PPCS. They expressed a need to document their symptoms to confirm for themselves

XSL•FO
RenderX

that this is actually what it is like over time and not something they are exaggerating or understating. The app also gives the opportunity to engage in self-reflection and identify possible adjustments to improve functioning and well-being. One participant stated:

> *It's about getting an overview of your own life and knowing that, okay, I can see the patterns, on the weeks where I sleep less, it gets worse. And then it's a bit reassuring to know, okay, I know this will be a hectic period, so maybe I need to take it a bit easier on other things. It's not necessarily about building your life around this, but it gives you the opportunity to make smarter choices and set yourself up with better conditions.* [Zara, 64, v3]

The participants described how they can use the visual representations of their reality to take ownership of their condition and proactively manage their own lives. They want to make informed decisions and set priorities that align with their values.

Many also see the potential of sharing such a visual representation of their PPCS experience with health professionals to facilitate communication and create a shared understanding of their symptom burden. One participant explained:

> *When you are at the general practitioner's office, you have very limited time to communicate everything that has been going on for the past two or three months; it is hopeless, actually. The doctor could rather just go in [the app] by themselves, watch the log, and then see how you have been.* [Omar, 53, v3]

Another participant said:

> *When you are in treatment and you get the question 'how have you been lately?' it's not so easy to remember because you are standing so close to it. So, to be able to show your registrations to get it more…in writing, that is usually very helpful* [Tamara, 56, v1]

They explained that it can be helpful to zoom out and show gradual and nonlinear recovery. Tracking their journey to recovery was thus another potential of the symptom diary pointed out by the participants.

Notably, several participants had already tried to create a visualization of their PPCS experience, using, for instance, Microsoft Excel or Post-It notes. Several participants also mentioned they have searched the internet for an app to keep track of their symptoms. Many had experience with headache diary apps, and some had tried symptom-tracking apps for various chronic illnesses. They had not yet experienced finding an app that took "the whole picture" into consideration, allowing them to track the symptoms and activities perceived as relevant to them. As 1 participant pointed out when talking about the headache diaries:

> *I feel like it is not enough to just log the headache. I have several types of headaches, but that is not the biggest challenge in my life. The fatigue, the sleep problems, the periods with bad memory have been far more challenging [...] There is so much more to it (than the headache).* [Valerie, 47, v1]

Personalized Yet Simple—Balancing Relevance and Usefulness

This theme represents the fine line between a useful and a cumbersome tool. A useful tool is perceived by the study participants as an app containing the necessary elements to log the aspects relevant to their functioning and well-being, while still being as simple as possible, making it low-threshold to use daily.

Many participants found the app relevant and easy to use, though some found it cumbersome and time-consuming. They expressed a need to minimize the available options and the clicks necessary to complete a symptom registration. They also highlighted the importance of always being able to edit their preselected symptoms.

The symptoms that users wanted to log seemed to be connected not only to symptom intensity, but also to their impact on their functioning, well-being, and symptom fluctuation.

> *It can be something that doesn't occur very often, but it is very intense when it does, or the opposite, something that is present most of the time without being extremely critical, but it's always there, simmering in the background in a way. So as long as you can choose [...] because people are different* [Sophie, 25, v2]

The activities perceived as relevant seemed to be connected to their perception of what influences symptom burden, but also activities enhancing their overall well-being without necessarily actively decreasing symptom intensity. For instance, 1 participant said:

> *Working in the garden is something that gives me peace of mind, and you need that on a bad day. Yeah, working in the garden would be nice to have as a distinctive activity to register* [Omar, 53, v3]

Several participants expressed how different contexts could generate different perceptions of what is relevant to track, for instance, starting a new treatment, returning to work after sick leave, or when experiencing a particularly stressful period in their personal lives. The feeling of purpose seemed to be essential for the perception of relevance.

Thus, there was great variation in which symptoms and activities were perceived as relevant. This underlies the importance of a flexible layout allowing for personalization, where the user of the app defines what is relevant in their unique case, enhancing the user's perception of the app's usefulness.

More Than Just a Number—the Complexity Behind Symptom Scores

The third theme reflects the complexity of PPCS and how the mapped symptoms in the app will always remain a simplification of reality. One participant put it this way:

XSL•FO
RenderX

> *To me, there is more behind that number than to you, and you can't include every person's life story, you need a number, so I get that* [Pedro, 58, v1]

Many participants found it challenging to sum up their experiences in a score. They acknowledge that there is no correct answer to "what a score of 5 means," and that they operate the scale based on their perception of what the scores represent, which is not static. One participant explained how they found their scoring became more consistent over time and that they needed about 14 days of "calibration" during the home testing. Another participant highlighted how their perception is, in fact, dynamic:

> *If it's been a long time since you had a 9 - 10, and you are in more pain than usual, you might put down a score of 8. But during a bad period, that score of 8 would've actually been a 5. So, that is something to keep in mind: a 5 is not always a 5.* [Eve, 40, v2]

The participants reflected on several aspects complicating their action of scoring, including the frequent fluctuations of symptoms, the intertangled relationship between symptoms, and other factors influencing the experienced symptom burden. Some pointed out internal factors influencing their perception of today's score, (eg, how they slept), as well as external factors (eg, if they went outside and were exposed to a lot of noise or stayed home that day). One person said they generally hesitated to score a 10 as they felt others probably had it worse.

Some also reflected on what they compared themselves to when scoring, for instance, life before the injury or to their peers.

While recognizing the challenging act of scoring and acknowledging that it inevitably always will be a simplification of their experiences, several participants still eyed the potential benefits of scoring their symptoms:

> *It is difficult to register [symptoms] based on a number, but at the same time, it makes it much more measurable and easier to spot patterns* [Sophie, 25, v1]

Discussion

Overview

We developed a research-based mHealth app for symptom mapping for people with PPCS. The app was created by combining rigorous research methods with iterative usability testing, together with people with PPCS and experienced clinicians in the field. We investigated the app's usability, feasibility, and safety. The app received high usability ratings from both the user and clinician groups and achieved a high adherence rate during home testing. Two AEs related to increased symptom awareness were logged during the study period.

Usability

The app received generally favorable usability scores and app quality scores from both the user and clinician groups. After 28 days of home testing, the final version received a mean score of 5 (SD 1.1) on the 7-point Likert scale in MAUQ by the user group and 4.1 (SD 0.4) on the 5-point Likert scale in MARS by the clinician group. Previous studies have evaluated apps for tracking symptoms in various conditions, including concussion and chronic pain, using the same questionnaires, with mean MAUQ scores ranging from 4.7 - 6.1 [31,32] and mean MARS scores ranging from 1.42 - 4.1 [33,34], implying that the app is considered acceptable in regards to usability and app quality by both people with PPCS and clinicians working in the field of PPCS. Notably, the user group rated the first 2 versions of the app higher than the third and final version (Figure 3). The differences in the study setting may partly explain this, as the user group tried the first 2 versions of the app in 2-hour sessions at the hospital, while using the third version at home for 28 days. The participants may have encountered challenges with the app when using it more extensively at home than during the limited time they had at the hospital. Furthermore, many of the participants experienced technical issues in the app during the home-testing period, possibly also affecting their experience of usability. Notably, the completion of the MAUQ did transpire in identical settings at all times, always situated in the hospital.

Feasibility

The adherence during the 28-day home-testing period was 89% (IQR 78 - 96). A meta-analysis of 99 mHealth apps for preventing or managing noncommunicable diseases revealed an average adherence rate of 56% [35]. A study similar to this study developed an app for people in the acute phase of concussions and observed an adherence rate of 28% during a 5-week period [31]. They recruited participants within 14 days of their concussion and thus aimed at a slightly different patient population than in this study, which may partly explain the difference in adherence rate. The high adherence rate in this study implies that the symptom diary app is a feasible approach for logging PPCS despite the relatively high frequency of technical issues for many of the study participants. In this study, we programmed the app to send a daily notification to register symptoms, based on feedback from the user and clinician groups, in order to ensure sufficient data density while also minimizing the experienced burden on the user. Given the high observed adherence rate, future studies can consider exploring multiple registrations per day to enhance data density. However, this should be done while carefully evaluating the user's experienced burden.

Safety

Regarding the app's safety, 2 AEs were registered concerning increased awareness of symptoms during home testing. Both participants completed the home testing but stated during their last encounter with the study personnel that they became more aware of their symptoms, which had a negative impact on their well-being. They both emphasized the value of keeping track of their symptoms but acknowledged that it can also function as a negative reminder. One participant emphasized the importance of receiving guidance from a health care provider in interpreting their symptom logs. This is an important finding, as the effects of symptom tracking in PPCS are understudied. This is, however, not unique to the field of PPCS; the effects of symptom mapping are, in general, understudied. In a small study with healthy individuals, it was shown that subjects tracking their symptoms daily for 2 weeks had increased recall

XSL•FO
RenderX

of symptoms and perception of symptom severity compared to the control group [21]. Still, we should not discard the reported experiences of the app's usefulness despite these incidents. It is plausible that a subgroup of the study population is prone to experiencing unwanted side effects related to increased awareness of symptoms. Future studies should focus on identifying such subgroups and determining which clinical implications symptom tracking apps should have. Developers of mHealth apps or wearable devices measuring symptoms should be aware of their ethical responsibilities, and clinicians recommending such tools and individuals using them should be made aware of the potential side effects [36-38].

Qualitative Analysis and Results

In the qualitative analysis, 3 overarching themes were created, which together informed the quantitative findings on the apps' usability, feasibility, and safety. The first theme, Visualizing the invisible—Enabling reflection and insight, actualizes a need for visual representations of PPCS experiences. It is essential for the usefulness of a mHealth app that its purpose is rooted in users' needs, as a good fit between a mHealth app's purpose and the user's needs increases the likelihood of continuous use [39].

The research field of personal visualization offers perspectives on the expressed need for visual representations of PPCS experiences. Visual representations of personal data can empower people with improved access and comprehension of their own data [40]. A systematic review of information visualization found visual tools essential in helping people understand and make sense of their personal health information, which is necessary to improve self-management of their health condition. For chronic illnesses, the "management" is oftentimes not focused on curing the condition but rather on coping and maintaining everyday life [41]. Coping and maintaining everyday life as an overarching goal is in line with the study participants' reflections on the purpose of the symptom diary app. They describe a need for a tool to visualize their PPCS experiences, create insight, engage in self-reflection, and make informed decisions aligned with their personal values and goals, as well as to facilitate communication with health care providers.

The second theme, Personalized yet simple—Balancing relevance and usefulness, touches upon both aspects of usability and feasibility of the symptom diary app. Experience of relevance seems to be vital to the perception of the app being useful, which again is essential for the probability of a mHealth app actually being used [39,42]. What is perceived as relevant is highly individual, as reflected in this second theme. The diverse views on relevance highlight the importance of a personalized layout in the app, allowing the user to define which symptoms and other variables are relevant to log and visualize in graphics. However, there appears to be a fine line between including many options to ensure perceived relevance and including too many options, making the app overly cumbersome. Thus, this requires careful navigation when developing mHealth apps for people with PPCS.

Finally, the third theme, More than just a number—The complexity behind symptom scores, addresses a more overarching concept, informing all 3 study objectives: usability, feasibility, and safety. Even though most study participants appreciated a visual representation of their PPCS experience and leveraged it to create overview and insight, many acknowledged the challenge of quantifying their experiences. They elaborated on how their interpretation of the scale is dynamic, depending on both internal and external factors. It is an important reminder that the visual representation of a person's PPCS experience in the app, or any other symptom-tracking instrument, will always remain a simplification of reality. A person's experience will always be richer and more complex than what can be expressed in numbers and summarized in graphics [43]. However, this does not diminish the usefulness of a visual representation for coping and managing everyday life, as described in the first theme.

Still, an important aspect to consider is how simplifying and quantifying someone's experience with PPCS may contribute to how they perceive their own experiences. From a critical realist perspective, PPCS exists as a real phenomenon independent of our descriptions of it. However, the conceptual framework embedded in the app, ie, the words we use, the categorization of symptoms, and the operationalization of numerical scales, will to some extent shape how the users of the app perceive, understand, and communicate their experiences. Creating a simplified visual representation can be useful for people with PPCS in gaining an overview and insight into a complex condition, as well as in communicating their PPCS experiences to others. It can support the user of the app in understanding and making sense of their condition, which again is crucial to health management and improving quality of life [41]. Still, this highlights the ethical responsibility of mHealth app developers to create research-based layouts and content [44]. Our findings regarding the symptom diary's safety illustrate the importance of not taking this ethical responsibility lightly. An mHealth app can have unwanted side effects for some individuals, just as any other health intervention, and requires a research-based approach.

Limitations

The questionnaires we applied to investigate app usability (MAUQ) and app quality (MARS) have been used extensively and validated in many languages [45,46]. For the purpose of this study, professional translators translated both the MAUQ and the MARS questionnaires into Norwegian using forward-backward translation, followed by harmonization of the translation by experts in the field of PPCS. However, no formal validation of the psychometric properties in a Norwegian context has been done, so direct comparisons of scores in this study with studies using the original English version should be interpreted with caution.

The number of technical issues encountered in the app during the home testing may represent a limitation regarding the feasibility assessment. Yet, adherence was not drastically affected when including the 2 participants who experienced severe technical issues (adherence 89% vs 81%), indicating good adherence despite the technical obstacles. Still, another fourteen participants experienced less severe technical problems, where no data were lost, but it may have affected their daily adherence to the app. Thus, the adherence to the app may be

XSL•FO
RenderX

underestimated, but it is, despite this, considered high compared to similar studies investigating the feasibility of mHealth apps in similar study populations [31,35]. The ongoing app development and refinement will further examine technical issues and how to mitigate them in the future.

Finally, in our study sample, 70% were female. Since this study focused on development and usability within a small sample, it was not designed to generalize to the whole PPCS population. Furthermore, women are overrepresented in the PPCS population, with female gender identified as a known risk factor for PPCS [47-49]. Future research should investigate how applicable the findings of this study are to a broader PPCS population.

Conclusions

The symptom diary app for people with PPCS is considered usable, feasible, and safe for its intended use and is ready for testing in larger-scale clinical trials.

Acknowledgments

The authors thank all study participants for their valuable contribution to this research study and the development of the symptom diary app. We also thank our research assistants, Nathalie Bergh and Sunniva Traa Carlsen, and the research coordinator Janne-Birgitte Bloch Bjerkhaug, at the Rehabilitation Clinic, St. Olavs Hospital, for their efforts.

Funding

This study was funded by the Liaison Committee between the Central Norway Regional Health Authority (RHA) and the Norwegian University of Science and Technology (NTNU) (2020/28373), the Central Norway Regional Health Authority (RHA) (innovation funding), and the Norwegian Research Council (328615). The funder had no influence on the study design, data collection, data analysis, data interpretation, or writing of the article.

Disclaimer

GS used Grammarly and the large language model ChatGPT-4 for grammatical corrections in the writing of this article. After using these tools and services, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

Conflicts of Interest

NTNU and St. Olavs Hospital, Trondheim University Hospital, may benefit financially from the commercialization of the proposed treatment through future possible intellectual properties. AO, ET, and AS are co-founders and shareholders of Nordic Brain Tech AS, a spin-off company established at NTNU. AO, ET, and AS hold a pending patent for the Cerebri Biofeedback treatment. AS has received lecture honoraria from TEVA. ET has received personal fees for lectures/advisory boards: Novartis, Eli Lilly, AbbVie, TEVA, Roche, Lundbeck, Pfizer, and Biogen. ET is consultant for and owner of stocks and IP in Man & Science. The mHealth app and all data, images, and results were created and are owned 100% by NTNU and St. Olav's Hospital.

References

1. Undén J, Ingebrigtsen T, Romner B, Committee SN. Scandinavian guidelines for initial management of minimal, mild and moderate head injuries in adults: an evidence and consensus-based update. BMC Med 2013 Feb 25;11(1):50. [doi: 10.1186/1741-7015-11-50] [Medline: 23432764]
2. Skandsen T, Einarsen CE, Normann I, et al. The epidemiology of mild traumatic brain injury: the Trondheim MTBI follow-up study. Scand J Trauma Resusc Emerg Med 2018 Apr 27;26(1):34. [doi: 10.1186/s13049-018-0495-0] [Medline: 29703222]
3. Lannsjö M, af Geijerstam JL, Johansson U, Bring J, Borg J. Prevalence and structure of symptoms at 3 months after mild traumatic brain injury in a national cohort. Brain Inj 2009 Mar;23(3):213-219. [doi: 10.1080/02699050902748356] [Medline: 19205957]
4. Mayer AR, Quinn DK, Master CL. The spectrum of mild traumatic brain injury: a review. Neurology (ECronicon) 2017 Aug 8;89(6):623-632. [doi: 10.1212/WNL.0000000000004214] [Medline: 28701496]
5. Iverson GL, Lange RT, Wäljas M, et al. Outcome from complicated versus uncomplicated mild traumatic brain injury. Rehabil Res Pract 2012;2012:415740. [doi: 10.1155/2012/415740] [Medline: 22577556]
6. Cancelliere C, Verville L, Stubbs JL, et al. Post-concussion symptoms and disability in adults with mild traumatic brain injury: a systematic review and meta-analysis. J Neurotrauma 2023 Jun;40(11-12):1045-1059. [doi: 10.1089/neu.2022.0185] [Medline: 36472218]
7. Karaliute M, Saksvik SB, Smevik H, et al. Methodology matters: comparing approaches for defining persistent symptoms after mild traumatic brain injury. Neurotrauma Rep 2021;2(1):603-617. [doi: 10.1089/neur.2021.0028] [Medline: 35018362]
8. Cassidy JD, Cancelliere C, Carroll LJ, et al. Systematic review of self-reported prognosis in adults after mild traumatic brain injury: results of the international collaboration on mild traumatic brain injury prognosis. Arch Phys Med Rehabil 2014 Mar;95(3 Suppl):S132-S151. [doi: 10.1016/j.apmr.2013.08.299] [Medline: 24581902]

XSL•FO
RenderX

9. Déry J, Ouellet B, de Guise É, Bussières È, Lamontagne ME. Prognostic factors for persistent symptoms in adults with mild traumatic brain injury: an overview of systematic reviews. Syst Rev 2023 Jul 20;12(1):127. [doi: 10.1186/s13643-023-02284-4] [Medline: 37468999]
10. Register-Mihalik JK, DeFreese JD, Callahan CE, Carneiro K. Utilizing the biopsychosocial model in concussion treatment: post-traumatic headache and beyond. Curr Pain Headache Rep 2020 Jun 27;24(8):44. [doi: 10.1007/s11916-020-00870-y] [Medline: 32594258]
11. Faulkner JW, Snell DL. A framework for understanding the contribution of psychosocial factors in biopsychosocial explanatory models of persistent postconcussion symptoms. Phys Ther 2023 Feb 1;103(2):pzac156. [doi: 10.1093/ptj/pzac156] [Medline: 37104623]
12. Voormolen DC, Polinder S, von Steinbuechel N, Vos PE, Cnossen MC, Haagsma JA. The association between post-concussion symptoms and health-related quality of life in patients with mild traumatic brain injury. Injury 2019 May;50(5):1068-1074. [doi: 10.1016/j.injury.2018.12.002] [Medline: 30554897]
13. Chu SY, Tsai YH, Xiao SH, Huang SJ, Yang CC. Quality of return to work in patients with mild traumatic brain injury: a prospective investigation of associations among post-concussion symptoms, neuropsychological functions, working status and stability. Brain Inj 2017;31(12):1674-1682. [doi: 10.1080/02699052.2017.1332783] [Medline: 28872344]
14. van der Vlegel M, Polinder S, Mikolic A, et al. The association of post-concussion and post-traumatic stress disorder symptoms with health-related quality of life, health care use and return-to-work after mild traumatic brain injury. J Clin Med 2021 Jun 2;10(11):2473. [doi: 10.3390/jcm10112473] [Medline: 34199591]
15. Bloom B, Thomas S, Ahrensberg JM, et al. A systematic review and meta-analysis of return to work after mild traumatic brain injury. Brain Inj 2018;32(13-14):1623-1636. [doi: 10.1080/02699052.2018.1532111] [Medline: 30307758]
16. Polinder S, Cnossen MC, Real RGL, et al. A multidimensional approach to post-concussion symptoms in mild traumatic brain injury. Front Neurol 2018;9:1113. [doi: 10.3389/fneur.2018.01113] [Medline: 30619066]
17. Rose SC, Fischer AN, Heyer GL. How long is too long? The lack of consensus regarding the post-concussion syndrome diagnosis. Brain Inj 2015;29(7-8):798-803. [doi: 10.3109/02699052.2015.1004756] [Medline: 25870975]
18. Skjeldal OH, Skandsen T, Kinge E, Glott T, Solbakk AK. Long-term post-concussion symptoms. Tidsskr Nor Laegeforen 2022 Sep 6;142(12). [doi: 10.4045/tidsskr.21.0713] [Medline: 36066235]
19. Zasler N, Haider MN, Grzibowski NR, Leddy JJ. Physician medical assessment in a multidisciplinary concussion clinic. J Head Trauma Rehabil 2019;34(6):409-418. [doi: 10.1097/HTR.0000000000000524] [Medline: 31479079]
20. Voormolen DC, Cnossen MC, Polinder S, et al. Prevalence of post-concussion-like symptoms in the general population in Italy, the Netherlands and the United Kingdom. Brain Inj 2019;33(8):1078-1086. [doi: 10.1080/02699052.2019.1607557] [Medline: 31032649]
21. Iverson GL, Lange RT. Examination of "postconcussion-like" symptoms in a healthy sample. Appl Neuropsychol 2003;10(3):137-144. [doi: 10.1207/S15324826AN1003_02] [Medline: 12890639]
22. King NS, Crawford S, Wenden FJ, Moss NE, Wade DT. The Rivermead Post Concussion Symptoms Questionnaire: a measure of symptoms commonly experienced after head injury and its reliability. J Neurol 1995 Sep;242(9):587-592. [doi: 10.1007/BF00868811] [Medline: 8551320]
23. Van den Bergh O, Walentynowicz M. Accuracy and bias in retrospective symptom reporting. Curr Opin Psychiatry 2016 Sep;29(5):302-308. [doi: 10.1097/YCO.0000000000000267] [Medline: 27427854]
24. Lenaert B, van Kampen N, van Heugten C, Ponds R. Real-time measurement of post-stroke fatigue in daily life and its relationship with the retrospective Fatigue Severity Scale. Neuropsychol Rehabil 2022 Jul;32(6):992-1006. [doi: 10.1080/09602011.2020.1854791] [Medline: 33297839]
25. Holmes MM, Stanescu S, Bishop FL. The use of measurement systems to support patient self-management of long-term conditions: an overview of opportunities and challenges. Patient Relat Outcome Meas 2019;10:385-394. [doi: 10.2147/PROM.S178488] [Medline: 31908555]
26. Deniz-Garcia A, Fabelo H, Rodriguez-Almeida AJ, et al. Quality, usability, and effectiveness of mHealth apps and the role of artificial intelligence: current scenario and challenges. J Med Internet Res 2023 May 4;25(1):e44030. [doi: 10.2196/44030] [Medline: 37140973]
27. von Steinbuechel N, Rauen K, Krenz U, et al. Translation and linguistic validation of outcome instruments for traumatic brain injury research and clinical practice: a step-by-step approach within the observational CENTER-TBI study. J Clin Med 2021 Jun 28;10(13):2863. [doi: 10.3390/jcm10132863] [Medline: 34203325]
28. Zeldovich M, Bockhop F, Covic A, et al. Factorial validity and comparability of the six translations of the Rivermead Post-Concussion Symptoms Questionnaire translations: results from the CENTER-TBI study. J Patient Rep Outcomes 2023 Sep 8;7(1):90. [doi: 10.1186/s41687-023-00632-5] [Medline: 37682406]
29. Digital symptom mapping and home-based biofeedback treatment for adults with persistent postconcussion symptoms: a development and usability study. ClinicalTrials.gov. 2022. URL: https://clinicaltrials.gov/ct2/show/study/NCT05635656 [accessed 2023-02-27]
30. Braun V, Clarke V. Thematic analysis: a practical guide: SAGE; 2022.
31. d'Offay C, Ng XY, Alexander L, et al. A digital health intervention for concussion: development and clinical feasibility study. JMIR Form Res 2023 Feb 1;7:e43557. [doi: 10.2196/43557] [Medline: 36724010]

XSL•FO
RenderX

32. Ester M, Wagoner CW, Dreger J, et al. Effectiveness of a self-monitoring app in supporting physical activity maintenance among rural canadians with cancer after an exercise oncology program: cluster randomized controlled trial. JMIR Cancer 2023 Sep 7;9(1):e47187. [doi: 10.2196/47187] [Medline: 37676714]
33. Salazar A, de Sola H, Failde I, Moral-Munoz JA. Measuring the quality of mobile apps for the management of pain: systematic search and evaluation using the mobile app rating scale. JMIR mHealth uHealth 2018 Oct 25;6(10):e10718. [doi: 10.2196/10718] [Medline: 30361196]
34. Terhorst Y, Messner EM, Schultchen D, et al. Systematic evaluation of content and quality of English and German pain apps in European app stores. Internet Interv 2021 Apr;24:100376. [doi: 10.1016/j.invent.2021.100376] [Medline: 33718002]
35. Jakob R, Harperink S, Rudolf AM, et al. Factors influencing adherence to mHealth apps for prevention or management of noncommunicable diseases: systematic review. J Med Internet Res 2022 May 25;24(5):e35371. [doi: 10.2196/35371] [Medline: 35612886]
36. Carter A, Liddle J, Hall W, Chenery H. Mobile phones in research and treatment: ethical guidelines and future directions. JMIR mHealth uHealth 2015 Oct 16;3(4):e95. [doi: 10.2196/mhealth.4538] [Medline: 26474545]
37. Galetsi P, Katsaliaki K, Kumar S. Exploring benefits and ethical challenges in the rise of mHealth (mobile healthcare) technology for the common good: an analysis of mobile applications for health specialists. Technovation 2023 Mar;121:102598. [doi: 10.1016/j.technovation.2022.102598]
38. Silberman J, Wicks P, Patel S, et al. Rigorous and rapid evidence assessment in digital health with the evidence DEFINED framework. NPJ Digit Med 2023 May 31;6(1):101. [doi: 10.1038/s41746-023-00836-5] [Medline: 37258851]
39. Vaghefi I, Tulu B. The continued use of mobile health apps: insights from a longitudinal study. JMIR mHealth uHealth 2019 Aug 29;7(8):e12983. [doi: 10.2196/12983] [Medline: 31469081]
40. Huang D, Tory M, Aseniero BA, et al. Personal visualization and personal visual analytics. IEEE Comput Graph Appl 2015 Jul;35(4):26-27. [doi: 10.1109/TVCG.2014.2359887]
41. Faisal S, Blandford A, Potts HWW. Making sense of personal health information: challenges for information visualization. Health Informatics J 2013 Sep;19(3):198-217. [doi: 10.1177/1460458212465213] [Medline: 23981395]
42. Zapata BC, Fernández-Alemán JL, Idri A, Toval A. Empirical studies on usability of mHealth apps: a systematic literature review. J Med Syst 2015 Feb;39(2):1. [doi: 10.1007/s10916-014-0182-2] [Medline: 25600193]
43. Einola K, Alvesson M. Behind the numbers: questioning questionnaires. Journal of Management Inquiry 2021 Jan;30(1):102-114. [doi: 10.1177/1056492620938139]
44. The Bellagio eHealth evaluation group call to action on global eHealth evaluation: consensus statement of the WHO global eHealth evaluation meeting. : World Health Organization; 2011 URL: https://cdn.who.int/media/docs/default-source/a-future-for-children/who-bellagio-ehealth-evaluation-call-to-action.pdf [accessed 2025-04-10]
45. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR mHealth uHealth 2015 Mar 11;3(1):e27. [doi: 10.2196/mhealth.3422] [Medline: 25760773]
46. Zhou L, Bao J, Setiawan IMA, Saptono A, Parmanto B. The mHealth app usability questionnaire (MAUQ): development and validation study. JMIR mHealth uHealth 2019 Apr 11;7(4):e11500. [doi: 10.2196/11500] [Medline: 30973342]
47. Cnossen MC, van der Naalt J, Spikman JM, et al. Prediction of persistent post-concussion symptoms after mild traumatic brain injury. J Neurotrauma 2018 Nov 15;35(22):2691-2698. [doi: 10.1089/neu.2017.5486] [Medline: 29690799]
48. Skandsen T, Stenberg J, Follestad T, et al. Personal factors associated with postconcussion symptoms 3 months after mild traumatic brain injury. Arch Phys Med Rehabil 2021 Jun;102(6):1102-1112. [doi: 10.1016/j.apmr.2020.10.106] [Medline: 33127352]
49. Fordal L, Stenberg J, Iverson GL, et al. Trajectories of persistent postconcussion symptoms and factors associated with symptom reporting after mild traumatic brain injury. Arch Phys Med Rehabil 2022 Feb;103(2):313-322. [doi: 10.1016/j.apmr.2021.09.016] [Medline: 34695386]

Abbreviations

AE: adverse event
CENTER-TBI: Collaborative European NeuroTrauma Effectiveness Research in Traumatic Brain Injury
MARS: Mobile App Rating Scale
MAUQ: mHealth App usability questionnaire
mHealth: mobile health
mTBI: mild traumatic brain injury
NTNU: Norwegian University of Science and Technology
PPCS: persisting postconcussion symptoms
RPQ: Rivermead Postconcussion Symptom Questionnaire

XSL•FO
RenderX

Edited by D Gooch; submitted 11.04.25; peer-reviewed by J Greenberg; revised version received 22.08.25; accepted 11.09.25; published 04.11.25.

Please cite as:
Storvig G, Stubberud A, Rauwenhoff J, Rønhovde LM, Smits M, Saksvik SB, Skandsen T, Tronvik E, Olsen A
Mobile Health App for Adults with Persisting Postconcussion Symptoms: Development and Usability Study
JMIR Hum Factors 2025;12:e75323
URL: https://humanfactors.jmir.org/2025/1/e75323
doi:10.2196/75323

© Gøril Storvig, Anker Stubberud, Johanne Rauwenhoff, Liv Marie Rønhovde, Martijn Smits, Simen Berg Saksvik, Toril Skandsen, Erling Tronvik, Alexander Olsen. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 4.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Usability Evaluation of Digital Health Applications for Older People With Depressive Disorders: Prospective Observational Study in a Mixed Methods Design

Magdalini Chatsatrian[1], BA; Katharina Kunde[1], MA; Jennifer Bosompem[1], MSc; Jan Dieris-Hirche[2], Dr Med; Nina Timmesfeld[3], Prof Dr Rer Nat; Rainer Wirth[4], Prof Dr Med; Georg Juckel[5], Prof Dr Med; Magdalena Pape[2,6], Dr Phil; Anna Mai[3], Dr Sportwiss; Chantal Giehl[4], MSc; Bianca Ueberberg[5], Dr Rer Nat; Horst Christian Vollmar[1], Prof Dr Med; Ina Carola Otte[1*], Prof Dr Sc Med; Theresa Sophie Busse[1,7*], Prof Dr Rer Medic

[1]Institute of General Practice and Family Medicine, Faculty of Medicine, Ruhr University Bochum, Bochum, Germany

[2]Department of Psychosomatic Medicine and Psychotherapy, LWL-University Hospital Bochum, Ruhr University Bochum, Bochum, Germany

[3]Department of Medical Informatics, Biometry and Epidemiology, Ruhr University Bochum, Bochum, Germany

[4]Department of Geriatric Medicine, University Hospital Marien Hospital Herne, Ruhr University Bochum, Bochum, Germany

[5]Department of Psychiatry and Psychotherapy, LWL-University Hospital Bochum, Ruhr University Bochum, Bochum, Germany

[6]Department of Clinical Psychology and Psychotherapy, University of Bamberg, Bamberg, Germany

[7]Digital Health, Department of Health, Faculty of Human Medicine, Witten/Herdecke University, Alfred-Herrhausen-Str. 50, Witten, Germany

[*]these authors contributed equally

Corresponding Author:
Theresa Sophie Busse, Prof Dr Rer Medic
Institute of General Practice and Family Medicine, Faculty of Medicine, Ruhr University Bochum, Bochum, Germany

Abstract

Background: Digital health applications (DiGA) have been integrated into Germany's health care system since 2019, offering certified medical devices for various health conditions. This study focuses on deprexis and Selfapy, the first 2 permanently approved DiGA for depressive disorders in Germany, to evaluate their usability for people ≥60 years. The study's significance is underscored by the underrepresentation of older people in previous DiGA studies, accompanied by an emergent risk of inequalities in distribution for this vulnerable population.

Objective: This study assessed the usability of DiGA deprexis and Selfapy for adults aged ≥60 years with mild to moderate depression. The more user-friendly option will be chosen for the DiGA4Aged project's upcoming randomized controlled trial.

Methods: The prospective observational study uses the People at the Centre of Mobile Application Development (PACMAD) usability model in a mixed methods design. The study's multistage data collection encompasses sociodemographic data and quantitative questionnaires about health literacy (European Health Literacy Survey Questionnaire [HLS-EU-Q16]), electronic health literacy (revised German eHealth Literacy Scale [GR-eHEALS]), media affinity, depressive symptoms (9-item Patient Health Questionnaire [PHQ-9]), and perceived usability (System Usability Scale [SUS]), as well as a qualitative think-aloud and semistructured interview. Participants were equally allocated to use either deprexis or Selfapy. Recruitment of 18 participants was conducted at 3 hospital departments (ie, psychiatry, psychosomatics, and geriatrics) in spring 2024. Participants were eligible if they were aged ≥60 years, were diagnosed with mild or moderate depressive disorder, owned a digital device, and gave written consent to participate.

Results: Quantitative analysis revealed age, gender, depressive severity, and health literacy parity between both groups. Selfapy users displayed marginally lower technical proficiency and lower usability scores. Qualitative data showed lower usability among participants in the Selfapy group due to design-related errors and higher cognitive load. Despite visual, psychomotor, and cognitive challenges, participants endorsed both DiGA for older users, stressing the importance of assistance and practicing the usage.

Conclusions: Reported difficulties in usability may help to improve future DiGA development for older people, especially as the willingness to use DiGA exists.

(JMIR Hum Factors 2025;12:e66271) doi:10.2196/66271

KEYWORDS

depression; depressive; usability; digital health; eHEALS; health literacy; mHealth; mobile health; mental health; app; applications; gerontology; geriatric; old; older; elder; elderly; aging; aged

XSL•FO
RenderX

Introduction

Depressive disorders represent an increasingly urgent public health issue, having already been identified by the World Health Organization as the condition most likely to rank second in the disability-adjusted life years index by 2030 [1]. Depressive disorders are associated with a significant reduction in quality of life and performance, along with major economic consequences [2].

Between 2001 and 2020, the point prevalence of self-reported depressive symptoms stood at 34% worldwide [3]. A survey was conducted across Europe to determine the prevalence of depressive symptoms within the past 2 weeks (assessed via self-report using the 8-item Patient Health Questionnaire [PHQ-8]). The result was a prevalence of 6.6% for the European Union as a whole (men 5.2% and women 7.9%), whereas in terms of age-standardized prevalence, people aged ≥65 years and above (9%) were most affected in the European Union. In Germany, a comparatively higher prevalence of 9.2% was recorded (women 10.8% and men 7.6%). People in the 15 to 29 years age group in Germany were particularly affected (11%), whereas approximately 7% of individuals aged ≥65 years showed signs of depressive symptoms [4].

There are already considerable shortcomings in treatment in Germany [5]. In particular, the demand for psychotherapy is still insufficiently met, leading to long waiting times before treatment begins [6].

Internet-based interventions may be one way of providing low-threshold access to treatment. While more than 10,000 smartphone apps are available for mental or behavioral health, evaluating these apps to ensure privacy protection, usability, and interoperability is a major challenge [7].

In Germany, efforts are underway to precisely ensure this evaluation through a standardization process, thereby giving patients direct access to quality-assured applications. Therefore, digital health applications (DiGA or Digitale Gesundheitsanwendungen) have been integrated into Germany's health care system since the German Digital Healthcare Act came into effect in 2019. DiGA, offered as apps and web-based applications, are certified as low-risk medical devices that can be prescribed by physicians and psychotherapists or dispensed after approval by health insurance companies [8]. A prescription is possible after successfully passing an assessment procedure of the Federal Institute for Drugs and Medical Devices. This is accompanied by the entry in the DiGA directory either as temporarily approved DiGA for 1 year pending reassessment or as permanently approved [9]. The approval by the Federal Institute for Drugs and Medical Devices requires evidence of the so-called positive care effect, meaning that the DiGA has been proven to contribute either to an improved management of the disease or to enhanced patient health [10].

Currently, 26 of the overall 56 listed DiGA address indications in mental health, making this category the largest of 12. To this date (July 5, 2024), 4 DiGA (ie, deprexis, edupression.com, Novego, and Selfapy) are permanently listed for use in depressive disorders, and another 3 DiGA are temporarily approved for 1 year (elona therapy Depression, MindDoc auf Rezept, and My7steps App) [11].

The first 2 permanently approved DiGA for treating depressive disorders were deprexis and Selfapy [11]. While these DiGA offered patients an opportunity to take advantage of treatment options for depressive symptoms, not all patients could be reached through this channel. The usage behaviors of certain age groups differ according to a report on the implementation and development of DiGA between 2020 and 2022. During the period under review, deprexis was prescribed approximately 10,000 times, with the frequency of use decreasing steadily among those aged ≥65 years. During the same period, Selfapy was prescribed about 9000 times in total, with almost no users in the age group >60 years [12]. Moreover, the approval studies of both deprexis [13,14] and Selfapy [15] did not include participants aged >65 years.

These findings illustrate the necessity of incorporating accessibility and participation considerations for the older population because depressive disorders are among the most common mental illnesses in older people [16]. Approximately 4% to 5% of individuals aged 65 to 79 years in Germany are affected by depressive symptoms [17]. Owing to the global trend of an aging population [18], the prevalence in this population group will steadily increase. Although, as mentioned earlier, other age groups in Germany are more severely affected by depressive symptoms, it is crucial to ensure that DiGA, as a state-supported concept, are beneficial and accessible to all individuals and that no one is excluded from this growing path of digital health care.

While expanding digital health care services such as DiGA could address these deficits, at the same time, inequalities in distribution might be reinforced, especially for vulnerable groups such as older people [19]. To ensure they can also benefit from such offers, the usability of DiGA should meet the needs of this user group. Various studies (eg, [20,21]) have shown that older people face specific challenges when using mobile apps: visual, psychomotor, and cognitive limitations due to aging and disease can impair usability. For instance, font size, visual contrast, small screens, and complex menu navigation pose considerable obstacles to use [20,21].

This study aimed to evaluate the usability of the DiGA deprexis and Selfapy for people aged ≥60 years with mild or moderate depressive disorders. The evaluation is based on the People at the Centre of Mobile Application Development (PACMAD) usability model [22]. The more user-friendly DiGA will be selected for the following randomized controlled trial (RCT) in the overall project DiGA4Aged.

Methods

Study Design and Participants

This prospective observational study adopted a convergent mixed methods design, involving the combination of both quantitative and qualitative data to obtain a more comprehensive understanding of the research topic. The design was stated as concurrent, with greater emphasis placed on qualitative data and triangulation of data occurring on the level of interpretation.

XSL•FO
RenderX

Both quantitative and qualitative data were collected at a single time point (single phase) [23].

Owing to the recent integration of DiGA into Germany's health care system since 2019, hardly any evidence was available before this study's conception in 2022. We searched the database of the Federal Institute for Drugs and Medical Devices for approval studies of the DiGA manufacturers and reports about the prescription and usage of DiGA. Furthermore, we searched PubMed for articles published from database inception to October 1, 2022, using the keywords ("depression" OR "depressive disorder" OR "mental health") AND ("DiGA" OR "health app" OR "eHealth" OR "mobile app" OR "smartphone app") without language restrictions. We identified systematic reviews on mobile apps with regard to mental health; however, none of them specifically included DiGA. Even now, the body of evidence on DiGA is only growing slowly.

The recruitment and assessment periods took place from February to April 2024 in 3 hospital departments: psychiatry, psychosomatics (both from Landwirtschaftsverband Westfalen-Lippe–University Hospital Bochum), and geriatrics (Marien Hospital Herne). In an endeavor to recruit a heterogeneous sample, the purposive selection by physicians and researchers at these departments was based on sociodemographic characteristics, health, and perceived electronic literacy, as well as perceived media affinity, alongside the eligibility criteria (Textbox 1). No distinction was made in terms of gender. In each department, the first 3 participants included in the study used Selfapy, followed by the next 3 using deprexis.

Textbox 1. Eligibility criteria for participation.

Inclusion criteria

- Aged ≥60 years
- Diagnosis of mild or moderate depressive disorder
- Ownership of a digital device suitable for digital health applications (DiGA) use
- Written consent to participate

Exclusion criteria

- Acute psychotic symptoms
- Acute suicidal tendencies or self-endangerment that results in an immediate need for treatment
- Severe intellectual impairment
- Motor and sensory impairments that prevent the use of a digital device
- Insufficient knowledge of the German language
- Advanced and incurable disease
- Diagnosis of bipolar disorder or schizophrenia
- Current participation in another intervention study
- Current use of a DiGA for the treatment of depressive disorder

Ethical Considerations

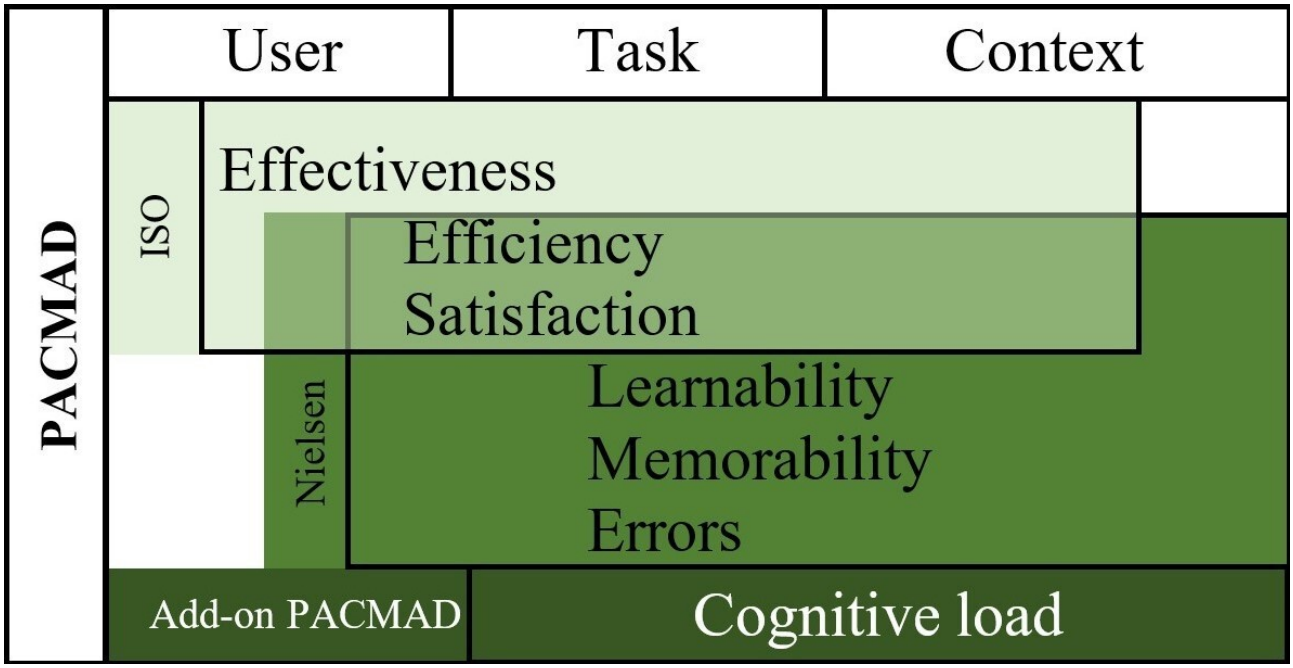
The study received ethical approval from the Ethics Committee of the Ruhr University Bochum (23 - 7901). The reporting adhered to the GRAMMS (Good Reporting of a Mixed Methods Study) guideline for mixed methods studies in health services research [24] and to the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) statement [25]. Participants were compensated with an expense allowance of €25 (US $29.24) per hour for their participation in the study. All participants were informed about the objectives of the study and the possible risks. They gave their written consent and were able to ask questions on site. Participation was voluntary, and they could withdraw at any time without suffering any disadvantages. None of the participants took advantage of this option. After the survey, the data were pseudonymized. All data was anonymized for the manuscript.

Procedures

The evaluation of the DiGA is based on the PACMAD usability model (Figure 1 [22]), which combines usability models developed by the International Organization for Standardization [26] and Nielsen [27] with an additional focus on mobile applications [22].

XSL•FO
RenderX

Figure 1. PACMAD usability model (own representation, based on a study by Harrison et al). ISO: International Organization for Standardization; PACMAD: People at the Centre of Mobile Application Development.



The PACMAD usability model addresses 7 usability attributes in total: effectiveness, efficiency, satisfaction, learnability, memorability, errors, and cognitive load [22]. The latter attribute, as an additional component to existing usability models, is considered the main contribution of the PACMAD model. Within the scope of those attributes, 3 further factors can influence the overall usability: user (eg, previous experiences interfering with overall usability), task (eg, goals of users), and context of use (eg, environment of use) [22].

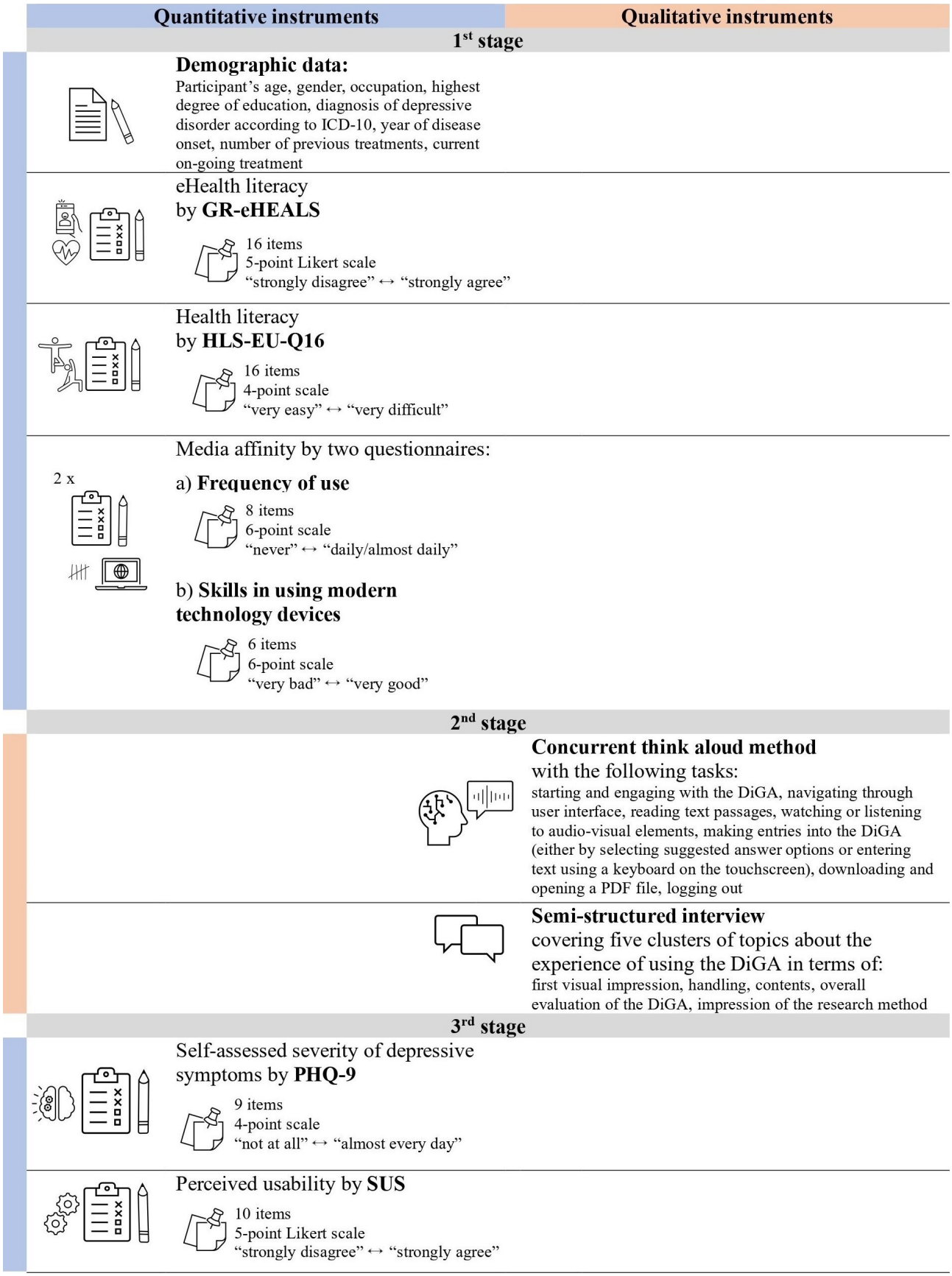
The data collection incorporated quantitative instruments in the first and third stages and qualitative instruments in the second stage (Figure 2) [28-32].

The revised German eHealth Literacy Scale (GR-eHEALS), the European Health Literacy Survey Questionnaire (HLS-EU-Q16), and the System Usability Scale (SUS) were used. During the third stage, a validated questionnaire was used for the self-assessment of the severity of depressive symptoms. The PHQ-9 is a widely used screening instrument for assessing depressive symptoms, based on *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*) criteria. Its brevity, diagnostic utility, and ease of administration make it especially suitable for use in population-based studies, including those involving older adults. The German version of the PHQ-9 was developed through a rigorous process of translation and back-translation, ensuring cultural and linguistic equivalence. Validation studies conducted in Germany have confirmed the semantic and conceptual appropriateness of the questionnaire for older German speakers [31]. In the German older adult population, the PHQ-9 has shown good criterion validity, with sensitivity and specificity levels that are consistent with its performance in younger populations. For instance, Kocalevent et al [33] reported a sensitivity of 81% and specificity of 89% at the standard cutoff score (≥10) when compared with structured clinical interviews. On the basis of these findings, the PHQ-9 can be considered a valid and reliable instrument for assessing depressive symptoms in older adults in Germany.

The concurrent think-aloud method [34] is a common procedure in the field of digital health care for testing technologies in terms of usability, content, and accuracy of fit [27]. It enables researchers to trace the user's course of action and reveals emerging usability issues [35] while offering insight into the users' cognitive and emotional reactions and processes [36]. The tasks were based on typical activities during initial use. Furthermore, semistructured interviews were conducted to gather participants' individual experiences and impressions after using the DiGA, enabling comparability while accommodating participants' personal perspectives [37]. The second stage included audio recording as well as video capturing of the participant's screen activity during DiGA use.

XSL•FO
RenderX

Figure 2. Data collection and instruments (own representation). DiGA: digital health applications; GR-eHEALS: revised German eHealth Literacy Scale; HLS-EU-Q16: European Health Literacy Survey Questionnaire; *ICD-10*: *International Statistical Classification of Diseases, Tenth Revision*; PHQ-9: 9-item Patient Health Questionnaire; SUS: System Usability Scale.



XSL•FO
RenderX

Data Analysis

Quantitative data were analyzed by descriptive statistics using the statistical software package SPSS Statistics 27.0 (IBM Corp). Measures of central tendency were calculated for the comparison of individual items or sum scores between individual participants or between the groups. Reporting includes both the mean with SD and the median with IQR. To prevent case exclusion due to missing values, median imputation was applied to 2 individual responses within the eHEALS and SUS instruments.

The extraction and synthesis of qualitative data were carried out according to Rädiker and Kuckartz [38] using the software MAXQDA 2022 (VERBI Software GmbH). The predefined attributes of the PACMAD model were used as the main category system, supplemented by an additional category on methodology.

Initially, 2 researchers (MC and TSB) independently coded 1 Selfapy assessment and discussed the resulting variations and discrepancies to set a standard. The remaining 17 assessments were coded (MC), and uncertainties across all videos were discussed (MC and TSB).

To ensure interrater reliability, 4 assessments were reevaluated by 2 researchers (ICO and TSB), with each researcher rating 1 Selfapy and 1 deprexis assessment from 2 distinct hospital departments. In total, 5 of 18 assessments were viewed twice.

The triangulation of both quantitative and qualitative datasets (Figure 3) [22,28-32] ensured a comprehensive interpretation of the PACMAD usability model. The factors user, task, and context of use are interconnected and compared across DiGA groups.

Figure 3. Triangulation matrix (own representation). GR-eHEALS: revised German eHealth Literacy Scale; HLS-EU-Q16: European Health Literacy Survey Questionnaire; PACMAD: People at the Centre of Mobile Application Development; PHQ-9: 9-item Patient Health Questionnaire; SUS: System Usability Scale.

user							**task**		**context**
demographic data			*health literacy*	*electronic health literacy and use of modern technology devices*			*usability*		*study setting*
gender	age	PHQ-9	HLS-EU-Q16	GR-eHEALS	frequency of use	skills of use	SUS	PACMAD	evaluation of the think aloud method

Results

Overview

A total of 18 evaluable assessments were conducted; 1 additional assessment could not be evaluated due to a technical malfunction of the recording equipment. Thus, each DiGA is represented by 9 assessments, equally covered by 3 per hospital department. All assessments were carried out with a study tablet, using the web-based application of both DiGA.

Quantitative Findings

Quantitative data were analyzed for 18 participants aged 61 to 88 years, with an almost equal distribution in both groups. The participants' average age was 68.3 (SD 9.6) years, and the majority (n=14, 78%) were women. Each group consisted of 7 women and 2 men. All but 3 participants were no longer employed (n=15, 83%).

On average, participants scored 12.4 (SD 5.2) on the PHQ-9, indicating moderate depression. Eight (44%) participants, 3 from the deprexis and 5 from the Selfapy group, were undergoing outpatient depression treatment at the time of data collection.

Similar results were observed in both groups for the HLS-EU-Q16, indicating moderate to good health literacy (scale 16 - 64) and for the GR-eHEALS, which represents moderate electronic health literacy (scale 16 - 80; Table 1). A total of 17 (94.4%) participants reported using a smartphone daily or almost daily. Although purposeful sampling was applied, participants in the Selfapy group showed lower media affinity overall, both in terms of usage frequency and self-assessed technology skills. The findings on media affinity are based on the assessment of individual items, as the questionnaires are not validated for a score.

XSL•FO
RenderX

Table . Patient-reported parameters.

	Deprexis (n=9)		Selfapy (n=9)		Total (N=18)	
	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)
Age (years)	68.33 (10.11)	63.00 (62.00-77.00)	68.22 (9.64)	63.00 (62.00-76.50)	68.28 (9.58)	63.00 (62.00-74.25)
Depressive severity (PHQ-9[a]) [31][b]	12.11 (4.59)	13.00 (8.50-15.50)	12.63 (6.19)	12.50 (6.50-16.50)	12.35 (5.23)	13.00 (8.00-15.50)
Health literacy (HLS-EU-Q16[c]) [29]	45.78 (5.65)	48.00 (42.00-50.00)	46.33 (8.79)	46.00 (38.00-54.50)	46.06 (7.17)	47.00 (41.25-50.00)
Electronic health literacy (GR-eHEALS[d]) [28]	51.89 (6.03)	52.00 (47.00-57.50)	48.75 (14.48)	47.00 (41.00-57.00)	50.41 (10.60)	49.00 (46.50-56.00)
System Usability Scale [32]	64.44 (20.03)	67.50 (48.75-82.50)	55.83 (18.29)	60.00 (41.25-68.75)	60.14 (19.13)	62.50 (43.125-75.00)

[a]PHQ: 9-item Patient Health Questionnaire.

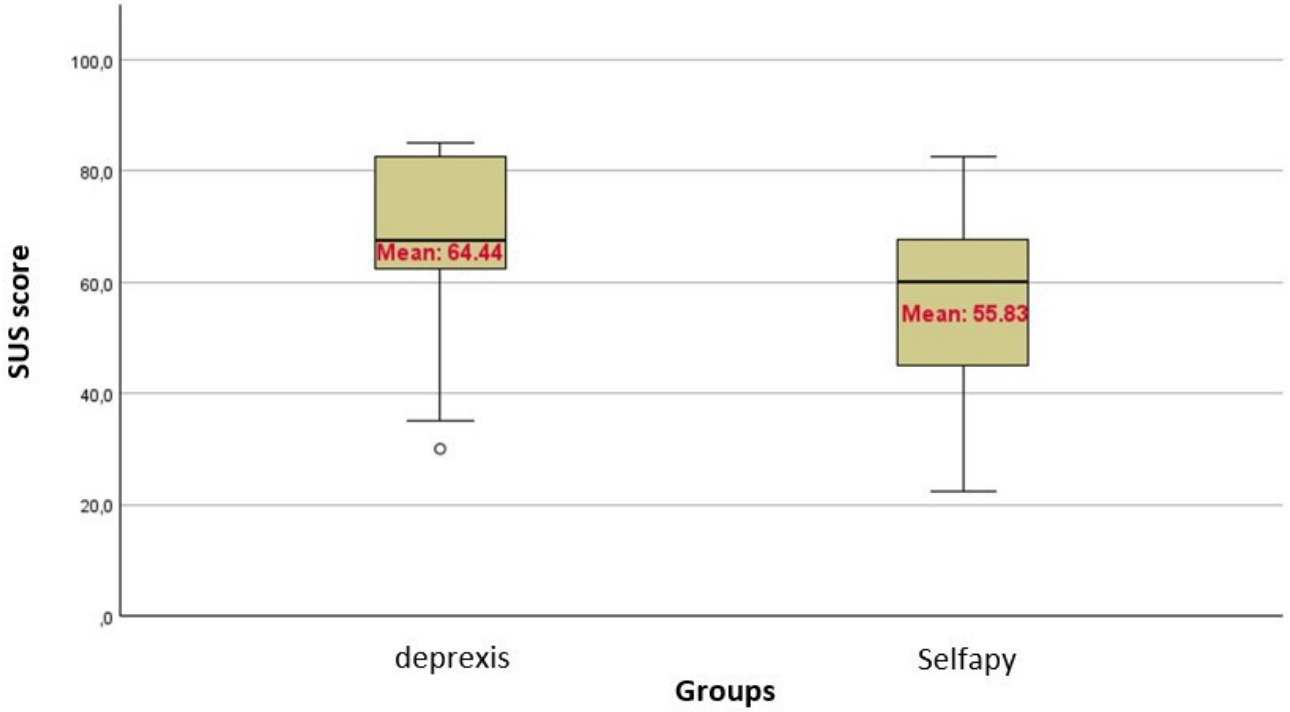
[b]For PHQ-9, the sample size was 17 (n=8 for Selfapy) due to missing data from 1 participant.

[c]HLS-EU-Q16: European Health Literacy Survey Questionnaire.

[d]GR-eHEALS: revised German eHealth Literacy Scale.

The usability assessment by SUS showed a higher level of user-friendliness for deprexis with an average score of 64.4 (SD 20.0) compared to Selfapy with 55.8 (SD 18.3). It is important to consider the SUS scores alongside participants' individual skills, as they affect the DiGA's perceived usability. Table 1 provides an overview of the self-reported patient parameters, and Figure 4 shows the SUS score distribution using a boxplot to visualize and compare the scatter, central tendency, and outliers across both groups.

Figure 4. SUS score. SUS: System Usability Scale.



Qualitative Findings

Overview

Qualitative findings are presented using the main categories, combining the results for Selfapy and deprexis. The evaluation dimension of the categories differs: while effectiveness, efficiency, and errors were solely evaluated externally by the researcher, satisfaction represents a category that only the participant could self-assess. Learnability, memorability, cognitive load, and the handling of the think-aloud method could be assessed by the participant's comments and through the researcher's observation.

XSL•FO
RenderX

Methodology

Responses to the think-aloud method varied: Selfapy participants rated it either as not disturbing at all or as stressful. Positive comments in the deprexis group outweighed, although some users struggled with verbalizing their thoughts. A few participants requested that the texts be read aloud, finding it challenging to manage both the printed instructions and the digital device.

Learnability

The category learnability describes the ease with which proficiency can be achieved [22]. Participants in both groups were eager to enhance their proficiency and expressed a desire to gain more confidence. They acknowledged that while learning requires repetition, it was generally manageable. A few participants lacked basic technical understanding and asked, for instance, what a PDF file is or how to adjust the volume. Despite initial challenges, some skills were learned even during first-time use, such as scrolling, identifying the menu icon, opening the keyboard, and zooming. Navigation was found to be learnable, and repetition of certain tasks led to faster performance. Participants recommended more guidance for using DiGA, noting the study conductor's help was valuable. They noted that, especially for older people without technical experience, the first use could be difficult.

Effectiveness, Efficiency, and Errors

The category effectiveness describes whether a task is completed by the user. Efficiency considers the speed and accuracy of completion, reflecting the user's productivity. Furthermore, the PACMAD model proposes an evaluation of the errors made while using an application [22].

Most tasks were completed successfully, although often with the assistance of the study conductor. Both groups faced touchscreen issues, causing frustration (eg, having to tap multiple times). Both DiGA presented navigation challenges, including orienting within chapters, finding PDF files, returning to the main page or the previous section, and the logout process. Reading and comprehension issues stemmed from small font size or unclear instructions and questions within both DiGA.

In Selfapy, the chapter overview layout led to confusion, with participants expecting immediate access to the chapter upon clicking. Other challenges included scrolling, selecting between answer options, and entering text via the touchscreen keyboard. Owing to Selfapy's user interface design, participants often missed the *next* button after completing 1 specific task and were confused about how to proceed. The required choice between 2 video formats (web based or app based) caused uncertainty about which video to watch. Selecting the wrong format resulted in limited visibility. Two participants in the Selfapy group terminated the assessment before completing all the tasks.

Deprexis users encountered similar but fewer issues. Specifically, some users failed to open a PDF file, as they clicked on the adjacent nonresponsive description instead of the PDF icon. They additionally faced difficulties locating the menu, as they were not familiar with the menu symbol. In some cases, the audio file was either overlooked or not played, despite having been noticed.

Satisfaction

The user's attitudes toward the DiGA provided insight into their satisfaction with using it [22]. Difficulties with navigation and the handling of the DiGA were recognized by the participants. One-third of the Selfapy users found the menu navigation and overall user interface pleasant, well structured, and intuitive, whereas the majority criticized the navigation as unintuitive and difficult to understand, requiring assistance in the beginning. Some participants suggested that repeated use could lead to improvement. One participant showed clear signs of frustration due to navigation difficulties.

In contrast, two-thirds of deprexis users perceived the DiGA as well structured and easy to navigate, whereas the remaining third found the navigation to be difficult and unintuitive. The main page and the beginning of the assessment were perceived as positive and accessible. Suggestions included clearer menu labels and a single-column layout of the menu to improve navigation.

Regarding design and layout, participants liked Selfapy's appearance, although some found the font size too small and the color contrast insufficient, restricting visibility. Deprexis was well received for its simple, clear interface, although opinions varied on visual elements and color use. The presentation of information was praised as well arranged, but participants recommended increasing the PDF font size. With regard to the general font size, while no negative feedback was given, 2 participants noted it was sufficiently large.

Content and the functions of both DiGA were generally perceived positively. The information was found to be presented in accessible language, with adequate depth and structure. The content was described as interesting, informative, and helpful; however, participants criticized the text in both DiGA as overly lengthy. They suggested an audio playback feature to resolve this concern. Some participants found the answer options limiting, whereas others praised the possibility of individual choice. Besides Selfapy's diary function and support options, the video was praised for its content, graphics, and brevity, although participants suggested moderating the voice-over pitch and reducing the pace.

Deprexis users appreciated the mood check, despite some criticism of its retrospective nature, and expressed a desire for additional support during completion. The overall guidance and motivation provided by deprexis, along with the relaxation exercises and the required responses, were well received. The audio element was praised for being engaging, clear, and memorable, with a pleasant voice quality and appropriate length.

The adoption of Selfapy for long-term use was met with mixed reactions. Some users were deterred by technical constraints and uncertainty about the effectiveness of managing illness. Others showed readiness to use DiGA, encouraged by its interactive nature and constructive suggestions. However, reluctance persisted due to general skepticism or insufficient technical equipment. Overall, the sentiment toward Selfapy was

XSL•FO
RenderX

cautious yet open to possibilities, with a clear need for support and reliable information.

In contrast, deprexis users approved of the DiGA, finding it a reliable aid. Many participants were open to using it regularly, particularly during critical health periods, and valued its potential at the onset of the disease to bridge therapy waiting times. Participants saw deprexis as a valuable addition to therapy but were unsure about its usefulness for individuals with severe depression. Some preferred personal interaction over the digital interface, whereas others liked the independence it offers. Concerns were raised regarding the data security of the diary feature, despite its perceived benefits. Doubts persisted regarding DiGA's capacity to meet individual needs and its potential to replace consultations with a therapist. A common concern was that using the DiGA alone could be overwhelming, prompting suggestions for an introductory course.

Cognitive Load

Cognitive load addresses the extent of cognitive processing while using an application [22]. Difficulties were prominent in this study: participants struggled with text comprehension and production, felt overwhelmed and insecure, and had memory and concentration lapses. These difficulties impacted their self-assessment. Both DiGA user groups faced these issues, but Selfapy was perceived as particularly challenging.

Selfapy users often struggled to grasp certain terms and instructions within the DiGA, frequently asking additional questions for clarity. They felt overwhelmed by the complexity and amount of content, struggled to articulate their thoughts in written form, were concerned about grammatical or spelling errors, and noted a decline in ability associated with age. Some Selfapy users stated that they had reached their individual limit at times and voiced uncertainty about their ability to concentrate and retain information after a single reading. Selfapy users experienced a powerful sense of self-doubt and insecurity. Despite their efforts, they harshly attributed mistakes to their own reduced intelligence and technical skills, rather than considering external factors or task complexity.

Deprexis users preferred shorter texts than those provided, despite understanding the content, and encountered minor comprehension difficulties with some tasks and instructions within the study. No serious confusion during the usage of deprexis was observed. Challenges included reduced focus due to illness, racing thoughts, and a reliance on routines to manage depression. Participants also expressed concern about self-esteem and mood assessment, which both influence cognitive processing.

Memorability

Memorability is defined as the user's ability to retain how an application is used effectively [22]. A second assessment is typically needed to evaluate memorability, but initial difficulties were evident during first-time use. From the participants' perspective, the video in Selfapy contained an excessive amount of information, with participants unable to remember important instructions, such as how to log off. Some did not recall watching a video at all when asked. Deprexis users also showed a few signs of limited memorability: 1 participant filled out the mood check twice without realizing it, and 2 participants forgot about the audio element when asked about it. A reminder feature was proposed for consistent DiGA use.

Discussion

Principal Findings

This prospective observational study gives new insights into the usability of DiGA for older adults with regard to depression, particularly focusing on deprexis and Selfapy, which were the first 2 permanently approved DiGA for depressive disorders. By focusing on participants aged ≥60 years, this research addresses the disparity in usage among older individuals, who have not been adequately represented in previous research. The study advocates for a user-centered approach in DiGA development on the basis of the PACMAD model. The study used a mixed methods approach. While the user groups were comparable in age, gender, severity of depressive symptoms, and health literacy, Selfapy users reported slightly lower technical proficiency and usability scores. Design flaws and increased cognitive load were reported by the participants of the Selfapy group. Despite visual, motor, and cognitive challenges, participants in both groups emphasized the need for support and practice in the use of DiGA for older adults.

A study by Wildenbos et al [39] on barriers to digital health among older adults identified 4 main categories of age-related barriers that can affect the usability of mobile health: physical abilities, perception, cognition, and motivation. Similar findings emerged from the present qualitative study results, corroborating the slightly lower SUS score for Selfapy compared to deprexis.

Barriers to Physical Abilities

Visual and psychomotor limitations were observed universally in both groups and were also evident in the participants' criticism regarding difficulties in using the touchscreen and navigation. Wildenbos et al [39] also reported that other studies had found slowed movements, reflexes, and tremors to be a challenge. It is assumed that these physical limitations may also affect learning duration, speed of performance, error rate, time retention, and subjective satisfaction.

Barriers to Perception

Barriers of perception were reflected in the early termination of 2 Selfapy assessments, increased error rates due to design limitations, and more moments of confusion and uncertainty with navigation, indicating usability constraints. In addition, small font size and poor color contrast were criticized. Wildenbos et al. [39] highlight the need for enhanced lighting to support visual clarity, as well as implications for video content and alerts due to the age-related decline in hearing ability. Hearing difficulties may also result in lower use of desktop computers and internet services compared to older adults without hearing impairments [40].

Barriers to Cognition

Selfapy users also experienced pronounced cognitive load. Some participants showed reduced concentration and struggled with text comprehension. This area may be less prominent in the present study compared to the study by Wildenbos et al. [39],

XSL•FO
RenderX

identifying potential challenges because this study did not include individuals with cognitive impairments. However, the participants' illnesses must be considered as a contributing factor.

Barriers to Motivation

Motivation is one of the key factors influencing the acceptance of technology [41]. Acceptance theories can provide exciting insights in this regard. Although the present study initially focused on usability, tentative conclusions can be drawn regarding this barrier: Selfapy users were more self-critical when faced with challenges. A reason for this could be the lower level of proficiency with modern technology reported by Selfapy users. Here, the mixed methods approach proved to be helpful in capturing specific impressions during use, in addition to the SUS.

Help with increasing motivation can be found in the literature: a systematic review examined the extent to which characteristics of the design and implementation of digital interventions can increase adherence among vulnerable population groups (older people, people with low socioeconomic status, single parents, social minorities, and people with a migration background). The authors state that multimodal content and direct opportunities for interaction and support between users and intervention providers can increase usage [42].

Interventions for Successful Implementation

Although participants encountered difficulties, they were mostly satisfied with both DiGA and expressed encouragement for long-term use. Successful long-term use depends particularly on actual use in everyday life. It is important to note that digital psychological treatment may have higher dropout rates compared to analog options [43]. A key factor in this area is adherence. A systematic review of adherence to digital interventions for mental health conditions found that female gender, positive treatment expectations, and the availability of personal support can increase adherence. Negative factors include limited time for engagement, personal dissatisfaction with the intervention content, or content perceived as impersonal [44]. The positive assessment of personal support was also evident in this study because both user groups rated their assigned DiGA as suitable for older people, emphasizing the necessity to practice the usage.

Although most of the tasks were successfully completed, participants received light support from the study conductors during use to reduce psychological strain. In traditional postprescription use, such support would not be available. The authors, therefore, expect participants to have more difficulty coping with the application, despite its approval through the DiGA authorization procedure. Structured support could potentially be helpful in this context. For example, it became apparent that a structured support system using weekly brief semistandardized emails achieved a high level of effectiveness for a digital mental health application [45]. Other studies on the use of mobile apps for mental health among older users also support the value of this approach [46-48].

Moreover, explicit guidelines and recommendations for developers could be beneficial in mitigating usability-related challenges. While such recommendations should not be derived solely from pilot studies focusing on usability, such as this study, there is a substantial body of design research that provides well-founded guidance in this area. For instance, Gomez-Hernandez et al [49] conducted a systematic review and thematic analysis on design guidelines for mobile applications targeted at older adults. Their study identified a total of 27 key recommendations, including 2 "golden rules" and 25 design principles, which were organized into 5 thematic categories: support and training, navigation, visual design, cognitive load, and interaction.

Strengths and Limitations

The distributions must be interpreted with caution due to the small sample size: it increases the variability of the data and limits the statistical validity of the results. In particular, outliers can distort the overall picture and thus impair the reliability of the conclusions derived.

This study gives new insights into the usability of DiGA for older adults with regard to depression, particularly focusing on deprexis and Selfapy, which were the first 2 permanently approved DiGA for depressive disorders. By focusing on participants aged ≥60 years, this research addresses the disparity in usage among older individuals, who have not been adequately represented in previous research. The study advocates for a user-centered approach in DiGA development on the basis of the People at the Centre of Mobile Application Development usability model. The distributions must be interpreted with caution due to the small sample size: it increases the variability of the data and limits the statistical validity of the results. In particular, outliers can distort the overall picture and thus impair the reliability of the conclusions derived.

On the basis of the information assessed by the GR-eHEALS and the 2 questionnaires on media affinity, the participants in the Selfapy group showed slightly lower proficiency in using modern technology than those in the deprexis group, which could have influenced the usability assessment.

Although the process of actively verbalizing one's thoughts is assumed to be unfamiliar, the participants were largely unaffected by the methodology. However, it must be taken into consideration that 3 different study conductors (1 per hospital department) were involved in data collection. Although the study settings through uniform training of the involved study members, minor procedural differences were nonetheless observed.

Another limitation arises from the target group of the study. Depressive disorder is characterized by symptoms such as attention deficits and lack of motivation. This may have impacted the perceived ease of use, although this effect cannot be conclusively demonstrated. However, we considered it crucial to focus on this target group to better understand the challenges faced by this underrepresented group. It is also important to note that participants were recruited directly from clinical practice. This approach was criticized in a meta-study on the effectiveness of deprexis, as the participants in those studies were recruited outside the clinical context, potentially compromising the transferability of the results [50].

XSL•FO
RenderX

Further Research

Larger controlled studies will be needed to replicate the findings and explore further moderators of usage outcomes. While the PACMAD usability model was designed specifically for assessing the usability of mobile applications and takes cognitive load into account, other factors influencing general technology acceptance are not considered.

The subsequent study of the DiGA4Aged project, which is the pilot study, will also examine technology acceptance and the success of support provided to participants by a digital nurse within an RCT. In the future, studies incorporating an acceptance model such as the Unified Theory of Acceptance and Use of Technology [51], which focuses more on general technology acceptance and less on mobile-specific use, will provide valuable insights. Furthermore, research on the implementation of DiGA for this specific patient group could be undertaken, possibly considering the Consolidated Framework for Implementation Research [52].

Conclusions

The findings of this study highlight the need to enhance the accessibility and usability of DiGA for older adults with depressive disorders, who face unique challenges due to age-related sensory, motor, and cognitive changes. In addition, the results show that usability was perceived as lower in the Selfapy group, primarily due to design-related factors of the application. Users in this group reported experiencing greater cognitive load. Participants in both groups emphasized the importance of receiving support when using the DiGA and expressed their desire to continue using it. However, further research is required to identify potential support measures for the use of DiGA by older adults, as well as factors influencing adherence. This way, DiGA can become more accessible and beneficial, ultimately reducing health care disparities and enhancing the quality of life for this vulnerable group. There is a necessity to take into account the needs of the target groups of DiGA in the authorization process and increase efforts to offer older adults support to bridge the digital divide. Considering the study's results and limitations, the decision of the DiGA to be used in the upcoming RCT was made in favor of deprexis.

Acknowledgments

The authors would like to thank the manufacturers for providing test accounts of Selfapy and deprexis free of charge for the duration of the study as well as the participants for their time. There was no further support, especially no financial support from the manufacturers. The authors would also like to acknowledge all members of the DiGA4Aged research group, in particular Hannes Bergmann and Tanja Roth from the hospital departments, for conducting the study.

This work was funded by the Innovations-FoRUM of the Medical Faculty at Ruhr University Bochum (funding code IF-025-22).

Data Availability

Data can be provided upon reasonable request to the first and last two authors.

Authors' Contributions

AM, GJ, HCV, JDH, NT, RW, and TSB conceived the study. ICO, JB, MC, MP, and TSB designed the study. BU, CG, JDH, and MP contributed to the collection of data. MC did the data analyses, with assistance from KK, who analyzed the quantitative data. ICO and TSB verified the data. MC prepared the first draft of the manuscript. All authors critically reviewed the manuscript before seeing and approving the final version.

Conflicts of Interest

None declared.

References

1. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med 2006 Nov;3(11):e442. [doi: 10.1371/journal.pmed.0030442] [Medline: 17132052]
2. Herrman H, Patel V, Kieling C, et al. Time for united action on depression: a Lancet-World Psychiatric Association Commission. Lancet 2022 Mar 5;399(10328):957-1022. [doi: 10.1016/S0140-6736(21)02141-3] [Medline: 35180424]
3. Shorey S, Ng ED, Wong CHJ. Global prevalence of depression and elevated depressive symptoms among adolescents: a systematic review and meta-analysis. Br J Clin Psychol 2022 Jun;61(2):287-305. [doi: 10.1111/bjc.12333] [Medline: 34569066]
4. Hapke U, Cohrdes C, Nübel J. Depressive symptoms in a European comparison - results from the European Health Interview Survey (EHIS) 2. J Health Monit 2019 Dec;4(4):57-65. [doi: 10.25646/6227] [Medline: 35146259]
5. Stahmeyer JT, Märtens C, Eidt-Koch D, Kahl KG, Zeidler J, Eberhard S. The state of care for persons with a diagnosis of depression. Dtsch Arztebl Int 2022 Jul 1;119(26):458-465. [doi: 10.3238/arztebl.m2022.0204] [Medline: 35506291]

XSL•FO
RenderX

6. Wartezeiten auf eine psychotherapie: studien und umfragen. Deutscher Bundestag. 2022. URL: https://www.bundestag.de/resource/blob/916578/53724d526490deea69f736b1fda83e76/WD-9-059-22-pdf-data.pdf [accessed 2025-10-16]
7. Torous J, Firth J, Huckvale K, et al. The emerging imperative for a consensus approach toward the rating and clinical recommendation of mental health apps. J Nerv Ment Dis 2018 Aug;206(8):662-666. [doi: 10.1097/NMD.0000000000000864] [Medline: 30020203]
8. Digitale Gesundheitsanwendungen (DiGA). Bundesministerium für Gesundheit. URL: https://www.bundesgesundheitsministerium.de/themen/krankenversicherung/online-ratgeber-krankenversicherung/arznei-heil-und-hilfsmittel/digitale-gesundheitsanwendungen [accessed 2024-02-19]
9. Für DiGA-Hersteller. Bundesministerium Für Arzneimittel und Medizinprodukte. URL: https://diga.bfarm.de/de/diga-hersteller [accessed 2025-06-09]
10. Für DiGA-Nutzende. Bundesministerium Für Arzneimittel und Medizinprodukte. URL: https://diga.bfarm.de/de/diga-nutzende [accessed 2024-02-19]
11. DiGA-Verzeichnis. Bundesministerium für Arzneimittel und Medizinprodukte. URL: https://diga.bfarm.de/de/verzeichnis [accessed 2024-02-19]
12. Bericht des GKV Spitzenverbandes über die Inanspruchnahme und Entwicklung der Versorgung mit digitalen Gesundheitsanwendungen (DiGA-Bericht): gemäß § 33a absatz 6 SGB v. GKV-Spitzenverband. 2023. URL: https://www.gkv-spitzenverband.de/media/dokumente/krankenversicherung_1/telematik/digitales/2024_DiGA-Bericht_final.pdf [accessed 2025-10-16]
13. Klein JP, Berger T, Schröder J, et al. Effects of a psychological internet intervention in the treatment of mild to moderate depressive symptoms: results of the EVIDENT study, a randomized controlled trial. Psychother Psychosom 2016;85(4):218-228. [doi: 10.1159/000445355] [Medline: 27230863]
14. Meyer B, Bierbrodt J, Schröder J, et al. Effects of an internet intervention (Deprexis) on severe depression symptoms: randomized controlled trial. Internet Interv 2015 Mar;2(1):48-59. [doi: 10.1016/j.invent.2014.12.003]
15. Krämer R, Köhne-Volland L, Schumacher A, Köhler S. Efficacy of a web-based intervention for depressive disorders: three-arm randomized controlled trial comparing guided and unguided self-help with waitlist control. JMIR Form Res 2022 Apr 4;6(4):e34330. [doi: 10.2196/34330] [Medline: 35105536]
16. Mental health of older adults. World Health Organization. 2023. URL: https://www.who.int/news-room/fact-sheets/detail/mental-health-of-older-adults [accessed 2024-03-07]
17. Heidemann C, Scheidt-Nave C, Beyer AK, et al. Gesundheitliche Lage von Erwachsenen in Deutschland – Ergebnisse zu ausgewählten Indikatoren der Studie GEDA 2019/2020-EHIS. J Health Monit 2021;6:3-27. [doi: 10.25646/8456]
18. Habekuß F. Die wachsende Welt. Bundeszentrale für Politische Bildung. 2021. URL: https://www.bpb.de/themen/soziale-lage/demografischer-wandel/196903/die-wachsende-welt/ [accessed 2025-10-08]
19. Bol N, Helberger N, Weert JCM. Differences in mobile health app use: a source of new digital inequalities? Inf Soc 2018 May 27;34(3):183-193. [doi: 10.1080/01972243.2018.1438550]
20. Awan M, Ali S, Ali M, Abrar MF, Ullah H, Khan D. Usability barriers for elderly users in smartphone app usage: an analytical hierarchical process-based prioritization. Sci Program 2021 Dec 26;2021:1-14. [doi: 10.1155/2021/2780257]
21. Elguera Paez L, Del Zapata Río C. Elderly users and their main challenges usability with mobile applications: a systematic review. In: Marcus A, Wang W, editors. Design, User Experience, and Usability: Design Philosophy and Theory: Springer; 2019:423-438. [doi: 10.1007/978-3-030-23570-3_31]
22. Harrison R, Flood D, Duce D. Usability of mobile applications: literature review and rationale for a new usability model. J Interact Sci 2013;1(1):1. [doi: 10.1186/2194-0827-1-1]
23. Creswell JW, Plano Clark VL. Designing and Conducting Mixed Methods Research, 3rd edition: SAGE; 2018.
24. O'Cathain A, Murphy E, Nicholl J. The quality of mixed methods studies in health services research. J Health Serv Res Policy 2008 Apr;13(2):92-98. [doi: 10.1258/jhsrp.2007.007074] [Medline: 18416914]
25. STROBE. 2024. URL: https://www.strobe-statement.org [accessed 2024-06-05]
26. ISO 9241-11:2018: ergonomics of human-system interaction: part 11: usability: definitions and concepts. International Organization of Standardization. 2018. URL: https://www.iso.org/standard/63500.html [accessed 2025-10-16]
27. Nielsen J. Usability Engineering: Morgan Kaufmann Publishers Inc; 1994.
28. Marsall M, Engelmann G, Skoda EM, Teufel M, Bäuerle A. Measuring electronic health literacy: development, validation, and test of measurement invariance of a revised German version of the eHealth literacy scale. J Med Internet Res 2022 Feb 2;24(2):e28252. [doi: 10.2196/28252] [Medline: 35107437]
29. Jordan S, Hoebel J. Gesundheitskompetenz von Erwachsenen in Deutschland: Ergebnisse der Studie "Gesundheit in Deutschland aktuell" (GEDA) (Health literacy of adults in Germany: Findings from the German Health Update (GEDA) study). Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2015;58:942-950. [doi: 10.1007/s00103-015-2200-z]
30. Schlomann A, Memmer N, Wahl HW. Awareness of age-related change is associated with attitudes toward technology and technology skills among older adults. Front Psychol 2022;13:905043. [doi: 10.3389/fpsyg.2022.905043] [Medline: 36160527]
31. Gräfe K, Zipfel S, Herzog W, Löwe B. Screening psychischer Störungen mit dem "Gesundheitsfragebogen für Patienten (PHQ-D)". Diagnostica 2004 Oct;50(4):171-181. [doi: 10.1026/0012-1924.50.4.171]

XSL•FO
RenderX

32. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Hum Comput Interact 2008 Jul 29;24(6):574-594. [doi: 10.1080/10447310802205776]
33. Kocalevent RD, Hinz A, Brähler E. Standardization of the depression screener patient health questionnaire (PHQ-9) in the general population. Gen Hosp Psychiatry 2013;35(5):551-555. [doi: 10.1016/j.genhosppsych.2013.04.006] [Medline: 23664569]
34. Fan M, Lin J, Chung C, Truong KN. Concurrent think-aloud verbalizations and usability problems. ACM Trans Comput-Hum Interact 2019 Oct 31;26(5):1-35. [doi: 10.1145/3325281]
35. Jaspers MWM. A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence. Int J Med Inform 2009 May;78(5):340-353. [doi: 10.1016/j.ijmedinf.2008.10.002] [Medline: 19046928]
36. Breuning M, Schäfer-Fauth L, Lucius-Hoene G, Holmberg C. Connecting one's own illness story to the illness experiences of others on a website-an evaluation study using the think aloud method. Patient Educ Couns 2020 Jan;103(1):199-207. [doi: 10.1016/j.pec.2019.08.014] [Medline: 31466883]
37. Helfferich C. Die Qualität Qualitativer Daten: Manual Für Die Durchführung Qualitativer Interviews: VS Verlag für Sozialwissenschaften; 2011. [doi: 10.1007/978-3-531-92076-4]
38. Rädiker S, Kuckartz U. Analyse Qualitativer Daten Mit MAXQDA: Text, Audio Und Video: Springer VS Wiesbaden; 2019. [doi: 10.1007/978-3-658-22095-2]
39. Wildenbos GA, Peute L, Jaspers M. Aging barriers influencing mobile health usability for older adults: a literature based framework (MOLD-US). Int J Med Inform 2018 Jun;114:66-75. [doi: 10.1016/j.ijmedinf.2018.03.012] [Medline: 29673606]
40. Henshaw H, Clark DPA, Kang S, Ferguson MA. Computer skills and internet use in adults aged 50-74 years: influence of hearing difficulties. J Med Internet Res 2012 Aug 24;14(4):e113. [doi: 10.2196/jmir.2036] [Medline: 22954484]
41. Rahimi B, Nadri H, Lotfnezhad Afshar H, Timpka T. A systematic review of the technology acceptance model in health informatics. Appl Clin Inform 2018 Jul;9(3):604-634. [doi: 10.1055/s-0038-1668091] [Medline: 30112741]
42. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen J. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152. [doi: 10.2196/jmir.2104] [Medline: 23151820]
43. Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. Clin Psychol Rev 2012 Jun;32(4):329-342. [doi: 10.1016/j.cpr.2012.02.004] [Medline: 22466510]
44. Beatty L, Binnion C. A systematic review of predictors of, and reasons for, adherence to online psychological interventions. Int J Behav Med 2016 Dec;23(6):776-794. [doi: 10.1007/s12529-016-9556-9] [Medline: 26957109]
45. Berger T, Hämmerli K, Gubser N, Andersson G, Caspar F. Internet-based treatment of depression: a randomized controlled trial comparing guided with unguided self-help. Cogn Behav Ther 2011;40(4):251-266. [doi: 10.1080/16506073.2011.616531] [Medline: 22060248]
46. Fortuna KL, DiMilia PR, Lohman MC, et al. Feasibility, acceptability, and preliminary effectiveness of a peer-delivered and technology supported self-management intervention for older adults with serious mental illness. Psychiatr Q 2018 Jun;89(2):293-305. [doi: 10.1007/s11126-017-9534-7] [Medline: 28948424]
47. Similä H, Immonen M, Toska-Tervola J, et al. Feasibility of mobile mental wellness training for older adults. Geriatr Nurs 2018;39(5):499-505. [doi: 10.1016/j.gerinurse.2018.02.001] [Medline: 29530292]
48. Gould CE, Carlson C, Ma F, Forman-Hoffman V, Ranta K, Kuhn E. Effects of mobile app-based intervention for depression in middle-aged and older adults: mixed methods feasibility study. JMIR Form Res 2021 Jun 29;5(6):e25808. [doi: 10.2196/25808] [Medline: 34185000]
49. Gomez-Hernandez M, Ferre X, Moral C, Villalba-Mora E. Design guidelines of mobile apps for older adults: systematic review and thematic analysis. JMIR Mhealth Uhealth 2023 Sep 21;11:e43186. [doi: 10.2196/43186] [Medline: 37733401]
50. Twomey C, O'Reilly G, Meyer B. Effectiveness of an individually-tailored computerised CBT programme (Deprexis) for depression: a meta-analysis. Psychiatry Res 2017 Oct;256:371-377. [doi: 10.1016/j.psychres.2017.06.081] [Medline: 28686935]
51. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425. [doi: 10.2307/30036540]
52. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated Consolidated Framework for Implementation Research based on user feedback. Implement Sci 2022 Oct 29;17(1):75. [doi: 10.1186/s13012-022-01245-0] [Medline: 36309746]

Abbreviations

DiGA: digital health applications
***DSM*:** *Diagnostic and Statistical Manual of Mental Disorders*
GR-eHEALS: revised German eHealth Literacy Scale
GRAMMS: Good Reporting of a Mixed Methods Study
HLS-EU-Q16: European Health Literacy Survey Questionnaire
PACMAD: People at the Centre of Mobile Application Development
PHQ-9: 9-item Patient Health Questionnaire

XSL•FO
RenderX

RCT: randomized controlled trial
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
SUS: System Usability Scale

Edited by K Adapa; submitted 09.09.24; peer-reviewed by C Jacob, V Rashedi; revised version received 07.08.25; accepted 15.08.25; published 28.11.25.

Please cite as:
Chatsatrian M, Kunde K, Bosompem J, Dieris-Hirche J, Timmesfeld N, Wirth R, Juckel G, Pape M, Mai A, Giehl C, Ueberberg B, Vollmar HC, Otte IC, Busse TS
Usability Evaluation of Digital Health Applications for Older People With Depressive Disorders: Prospective Observational Study in a Mixed Methods Design
JMIR Hum Factors 2025;12:e66271
URL: https://humanfactors.jmir.org/2025/1/e66271
doi:10.2196/66271

© Magdalini Chatsatrian, Katharina Kunde, Jennifer Bosompem, Jan Dieris-Hirche, Nina Timmesfeld, Rainer Wirth, Georg Juckel, Magdalena Pape, Anna Mai, Chantal Giehl, Bianca Ueberberg, Horst Christian Vollmar, Ina Carola Otte, Theresa Sophie Busse. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 28.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Mobile Health Adoption in High-Risk Pregnancies Using Cluster Analysis of Biopsychosocial Outcomes: Observational Longitudinal Cohort Study

Fernanda Schier de Fraga[1*], MD, MSci; Mayara Marenda Narita[1*], MSc; Monique Schreiner[1], MSc; Flavio Belli[1*]; Jaqueline Leonel Celestino[1*]; Karolayne Braz Pereira[1*], MD; Gabriella Soecki[1*], MD; Vitória Bevervanso[1*], MD; Rogério de Fraga[2*], MD, MSci, PhD

[1]Department of Obstetrics and Gynecology of the Federal University of Paraná, Rua General Carneiro, 181, Curitiba, Brazil

[2]Department of Surgical Clinics at the Federal University of Paraná, Curitiba, Brazil

[*]these authors contributed equally

Corresponding Author:
Fernanda Schier de Fraga, MD, MSci
Department of Obstetrics and Gynecology of the Federal University of Paraná, Rua General Carneiro, 181, Curitiba, Brazil

Abstract

Background: The use of mobile technologies during high-risk pregnancy, placing patients at the center of care, affords them self-management and easier access to health information.

Objective: This study aims to understand the health perception of pregnant women at the beginning of high-risk antenatal care, the usability of a mobile health app—the Health Assistant—and to compare maternal-fetal outcomes between users and nonusers of the app.

Methods: This is an observational longitudinal cohort study that looked into clusters of high-risk pregnant women admitted to antenatal care at the maternity unit of a public university hospital in southern Brazil between April 2022 and November 2023. Pregnant women who did not have a compatible smartphone to download the app or who did not have internet access were excluded from the study. According to systematic randomization, one patient was allocated to the app group and the other to the control group. They all answered an inclusion questionnaire (Q1), and those in the app group were instructed to use the Health Assistant app to prepare for their first antenatal appointment, which would take place in a few weeks' time, when they would answer the Brazilian version of the Mobile App Usability Questionnaire. After childbirth, maternal-fetal outcomes were assessed. Student 2-tailed *t* test, Mann-Whitney test, Fisher exact test, and the chi-square test were used for statistical analysis. A hierarchical cluster analysis was performed using the Ward method and the Euclidean squared distance measure.

Results: The sample contained 111 pregnant women, of whom 55 (49.5%) were allocated to the app group and 56 (50.5%) to the control group. Of the 55 pregnant women who used the app, 21 (38.2%) demonstrated adherence, with an average Mobile App Usability Questionnaire score of 6.2 (SD 1.0). Clustering included 110 pregnant women, and the dendrogram resulted in three clusters, which show several significant differences in terms of family income, medical history, medication adherence, and lifestyle habits. Cluster 2 had the lowest adherence to the app (P=.08) and attended significantly fewer antenatal appointments (6.9 appointments) as compared with Clusters 1 (10.3) and 3 (9.1; P=.006). Cesarean section was more frequent in Cluster 3 (n=41, 95.3%) as compared with Clusters 1 (n=12, 27.9%) and 2 (n=5, 20.8%), P<.001.

Conclusions: Cluster analysis, revealing different profiles of pregnant women, allowed us to identify groups that would benefit from personalized approaches and digital interventions to improve self-awareness and gestational outcomes. The Health Assistant app showed good usability in this context.

(JMIR Hum Factors 2025;12:e67680) doi:10.2196/67680

KEYWORDS

antenatal care; cluster analysis; mHealth; mobile health; high risk; pregnancy; biopsychosocial; mobile technology; high-risk pregnancy; self-management; pregnant; antenatal; mobile health app; health perception; app; observational longitudinal cohort study; Brazil; smartphone; mobile phone

XSL•FO
RenderX

Introduction

Mobile health (mHealth) is defined as medical and public health practice aided by mobile devices such as cellphones, patient monitoring tools, personal digital assistants, and other wireless equipment [1,2]. In Brazil, smartphone price reduction and improved connectivity facilitating access to device functionalities have had a major impact on health care with useful tools for professionals and patients alike [3,4]. As part of patient care, mHealth apps aim to promote well-being, emotional, psychological, and physical growth, to increase accessibility to health services in remote areas, and to make processes faster and more precise [2,4-6]. The use of mHealth technologies during pregnancy and the postpartum period has become more common with the advance of information and communication technologies [7,8].

Despite the great educational potential and the increasing popularity of health apps, rigorous scientific evaluation of their effectiveness is still limited, and there is an insufficient number of studies using standardized methods to guarantee their performance and safety [8,9]. DeNicola et al [8] conclude that these apps should be developed with the cooperation of health professionals. An ideal app, resting on scientific evidence, respecting the law, and protecting privacy, should also be user-friendly, attractive, simple, and functional, deploying device sensors and resources for the good of the patient [10]. The use of this technology can thus prove a valuable strategy for health self-management, transforming patient experience by affording greater autonomy and ease of access to care [11].

High-risk pregnancy—when medical conditions or specific pregnancy factors can compromise the health of the mother or the fetus—has an approximate prevalence of 22% [12,13]. This condition can intensify stress and impact the psychological health of pregnant women, worried about possible interventions, hospitalizations, and treatments [14]. The very label “high-risk pregnancy” can contribute to additional stress, as concerns may focus rather on treatments and tests than on the woman’s overall well-being, but it should also lead to greater medical vigilance, which can be perceived either positively or negatively by the patient [12].

In these cases, patient-centered communication yields positive results, since a more humanized and integrative approach can significantly improve the quality of care [15]. Several studies show that a person-centered approach improves satisfaction, reduces costs, and increases the effectiveness of treatments, especially in patients with chronic conditions [15]. In this context, the use of mobile technologies places the patient at the center of care, offering a new experience that enables health self-management and easier access to information and services [11]. A systematic review identified that in 74.3% of the studies, mHealth intervention was cost-effective, but indicated that more research is needed in low- and lower-middle-income countries for the impact of different types of mHealth to be better understood [16].

Furthermore, cluster analysis, considered a relevant method of classifying women in early pregnancy, is another way of drawing attention to the follow-up of high-risk pregnant patients [17]. By identifying homogeneous groups, it allows risk factors to be detected, as well as patterns and trends in pregnancy outcomes to be understood [18,19].

Since longitudinal follow-up with no information gaps is essential for proper antenatal care, an mHealth app produced by health professionals could help prepare high-risk pregnant women for their first antenatal appointment at a maternity unit of a public university hospital in southern Brazil. An easy-to-use app called Health Assistant was expected to have positive effects on both the pregnant woman and the child.

The aim of this study was to evaluate the health perception of pregnant women starting high-risk antenatal care in a low-income country, the usability of the Health Assistant, and maternal-fetal outcomes among app users and nonusers so that we could identify those groups of patients who benefit most from the app and thus provide focused assistance in order to optimize care and obstetric outcomes.

Methods

Recruitment

Pregnant women were included upon admission to high-risk antenatal care at the maternity unit of a public university hospital in southern Brazil between April 2022 and November 2023. High-risk pregnancies were characterized by chronic conditions or those diagnosed during pregnancy, such as hypertension, gestational diabetes, and mood disorders. Primary care referral was the gateway to this care.

Pregnant women who did not have a compatible smartphone to download the app or who did not have internet access, even when the researchers provided a router, were excluded from the study. After accepting and signing the informed consent form, participants were included in the study and allocated, according to systematic randomization with random start, to one of the groups: app or control. The first patient was drawn at random and allocated to the app group, and the others were then assigned in an interleaved manner: the second to the control group, the third to the app group, the fourth to the control group, and so on. Thus, we created two balanced groups—one invited to use the app and the other not invited—without supervision or imposition of use. Although the allocation was random, the study maintained an observational design, as the actual use of the app was at the discretion of the participants and there was no controlled intervention beyond the invitation.

Every patient answered the inclusion questionnaire (Q1), consisting of 32 questions about their social, economic, and health profile. Those allocated to the app group were instructed to use the Health Assistant app to prepare for their first antenatal appointment, which would take place in a few weeks’ time. The digital tool was used without supervision or direct control by the researchers and allowed patients to answer detailed questions about their current pregnancy and their gynecologic-obstetric history. Our objective was to evaluate, in the context of routine clinical practice, the associations between digital engagement and maternal-fetal outcomes, and not to test a controlled intervention protocol.

XSL•FO
RenderX

At the first antenatal appointment, researchers met with participants from the app group and had adherents to the app answer the Brazilian version of the Mobile App Usability Questionnaire (MAUQ).

Ethical Considerations

This is an observational longitudinal cohort study approved by the Institutional Ethics Committee of the Hospital de Clínicas Complex (47084021.5.0000.0096). All participants provided written informed consent before enrollment in the study. Participants' personal data were anonymized to preserve confidentiality and privacy. No financial or other compensation was offered for participation.

MAUQ

MAUQ was chosen because it is a widely recognized tool in the literature for evaluating the usability of mHealth apps. Developed and validated in English by Zhou et al [20], the MAUQ is highly reliable with a Cronbach α of 0.914. The questionnaire was translated into Portuguese by Machado et al in accordance with the guidelines for translation and cross-cultural adaptation of nontechnical skills assessment instruments, ensuring similarity between the original and the translation. Permission to use the translation was kindly granted to the researchers (personal communication by Machado, 2022). For this study, we used the version of the MAUQ for stand-alone digital health apps intended for patients to control, manage, and improve their health care. The questionnaire assessed the following parameters: ease of use, interface, satisfaction, and usefulness. Participants rated each item using a 7-point Likert Scale, ranging from 1=strongly disagree to 7=strongly agree. The usability of the app was determined by the total score and the averages of all the questionnaires answered, where a higher score indicates greater usability [20]. In our study, the MAUQ was administered to patients who were considered adherent, that is, who reported having used the app.

Outcome Analysis

After the expected date of delivery, researchers retrieved their medical records to analyze maternal-fetal outcomes, including the number of antenatal appointments, the number of emergency room visits during antenatal care, intercurrent events, date of delivery, route of delivery, complications, and length of stay for mother and baby, among others. The information collected was organized and filed in spreadsheets for statistical analysis.

Cluster Analysis

Clustering is a technique for organizing a large amount of information into smaller groups so that hidden patterns and relationships between the data are identified and personalized interventions are developed [21,22]. This method was chosen because of its effectiveness in identifying the hierarchical structure of binary data without the need to specify the number of desired clusters beforehand. It also promotes greater homogeneity within clusters and is easy to replicate, making it applicable to various types of data [21,23,24]. Based on the characteristics and outcomes of each cluster, it is possible to develop antenatal care strategies that optimize patient-centered care.

Sample Size Calculation

As an exploratory technique that does not require statistical hypotheses, hierarchical cluster analysis generally dispenses with sample calculation, since there are no standardized methods for it [25]. Dalmaijer et al [26] suggest that a sample of 20 - 30 individuals per expected subgroup is sufficient for cluster analysis, provided that there is a clear separation between the subgroups. Defining the sample in this study aimed to ensure that there was enough data for significant patterns to emerge. Estimating the formation of two or three clusters, depending on the dendrogram, a minimum of 100 patients was therefore proposed, 50 for the control group and 50 for the app group.

Statistical Analysis

Data were organized in an Excel (Microsoft Corp) spreadsheet and analyzed using SPSS Statistics (version 29.0.0; IBM Corp). Descriptive analysis of quantitative variables included the mean, SD, median, and IQR. Categorical variables were expressed in frequency and percentage. Student 2-tailed *t* test for independent samples or the Mann-Whitney nonparametric test was used to compare control and app groups in terms of quantitative variables. The normality of continuous variables was assessed using the Kolmogorov-Smirnov test. Categorical variables were analyzed using Fisher exact test or the chi-square test. An exploratory analysis of the variables in the inclusion questionnaire (Q1) resulted in the exclusion of questions with well-correlated answers or questions that had five or more missing data (blank answers). To determine the clustering of pregnant women, a hierarchical cluster analysis was carried out using the Ward method and the squared Euclidean distance measure. The number of clusters was determined according to the dendrogram. Values of $P<.05$ indicated statistical significance.

Results

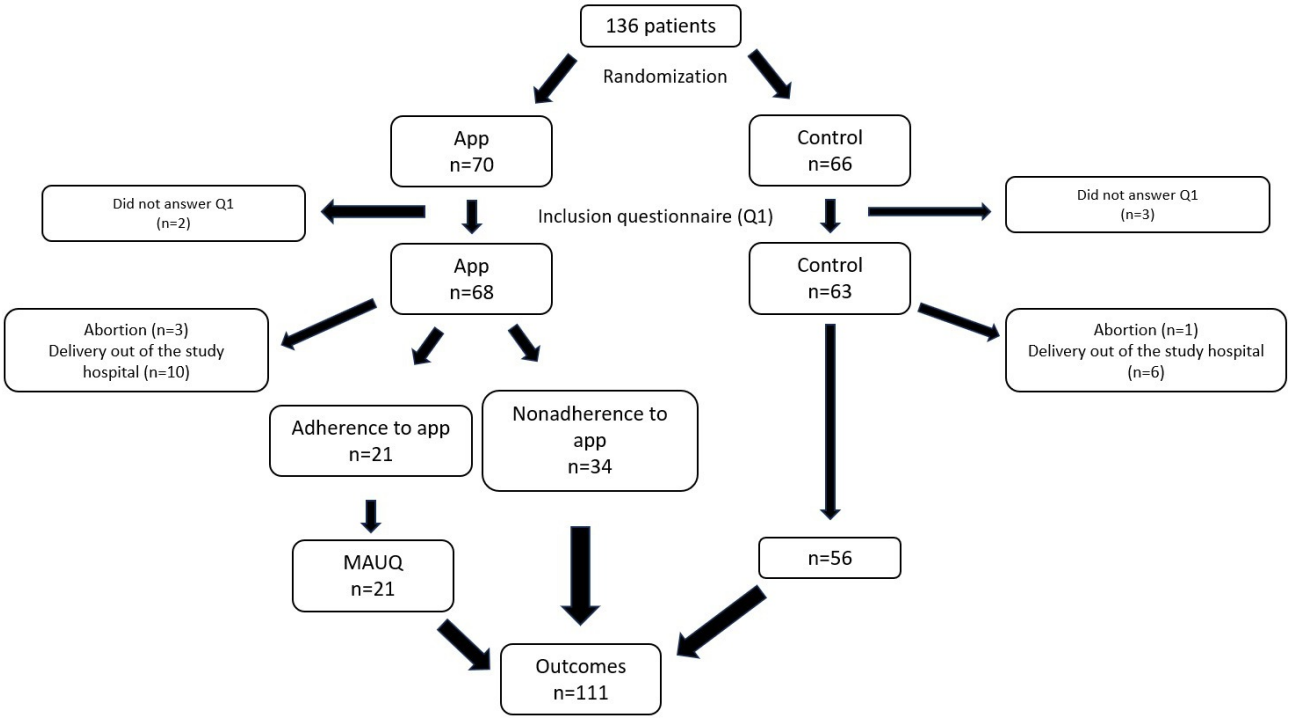
Overview

The inclusion questionnaire (Q1) was answered by 131 patients. The team had no access to the outcomes of 20 patients due to abortion or delivery out of the study hospital. Only patients who had answered Q1 and whose outcomes were available were included in the analysis, resulting in a total of 111 pregnant women, as shown in Figure 1.

The MAUQ, restricted to the group that was adherent to the app, was answered by 21 patients and had an average score of 6.2 (SD 1.0) points, with a median of 6.6 (IQR 2.3-7). The internal consistency of the instrument was high, with a Cronbach α estimated at 0.94.

XSL•FO
RenderX

Figure 1. Patient inclusion flowchart. MAUQ: Mobile App Usability Questionnaire; Q1: inclusion questionnaire.



Patient Characteristics

The average age in the control group was 30.5 (SD 5.7) years and the app group was 28.8 (SD 6.2) years, with no significant difference (P=.13). The median number of pregnancies for the control group was 3 (IQR 1-10) and 2 (IQR 0 - 6) for the app group (P=.18). Table 1 shows the categorical variables of Q1 with the characteristics of patients for each group.

XSL•FO
RenderX

Table . The inclusion questionnaire (Q1) categorical variables.

Variable and classification	Group, n (%)		*P* value[a]
	Control	App	
Marital status (n=111)			__[b]
Cohabiting	10 (17.9)	19 (34.5)	
Married	32 (57.1)	19 (34.5)	
Single	13 (23.2)	17 (30.9)	
Separated	1 (1.8)	0 (0)	
What is your average family income (in minimum wages; n=111)			.81
Less than 1 or 1	10 (17.9)	11 (20)	
2 or more	46 (82.1)	44 (80)	
Parity (n=111)			.23
Nulliparous	16 (28.6)	22 (40)	
Multiparous	40 (71.4)	33 (60)	
Have you ever had a vaginal birth? (n=111)			≥.99
No	24 (42.9)	23 (41.8)	
Yes	32 (57.1)	32 (58.2)	
How many vaginal births? (if applicable; n=64)			.75
1 or 2	25 (78.1)	27 (84.4)	
3 to 7	7 (21.9)	5 (15.6)	
Have you ever had a C-section? (n=111)			.57
No	23 (41.1)	26 (47.3)	
Yes	33 (58.9)	29 (52.7)	
How many C-sections? (if applicable; n=62)			.09
1 or 2	22 (66.7)	25 (86.2)	
3, 4, or 5	11 (33.3)	4 (13.8)	
Do you have any deceased child? (n=111)			.18
No	51 (91.1)	45 (81.8)	
Yes	5 (8.9)	10 (18.2)	
Do you understand why you have been referred to high-risk antenatal care? (n=111)			.62
No	3 (5.4)	1 (1.8)	
Yes	53 (94.6)	54 (98.2)	
Comorbidities (n=111)			.70
Only 1	34 (60.7)	31 (56.4)	
More than 1	22 (39.3)	24 (43.6)	
Are you on any sort of medication? (n=111)			≥.99
No	11 (19.6)	11 (20)	
Yes	45 (80.4)	44 (80)	
Do you know what each medication is used for? (n=95)			.49
No	6 (12.2)	3 (6.5)	
Yes	43 (87.8)	43 (93.5)	
Do you often forget to take your medication? (n=95)			.48
No	38 (77.6)	32 (69.6)	

XSL•FO
RenderX

Variable and classification	Group, n (%)		*P* value[a]
	Control	App	
Yes	11 (22.4)	14 (30.4)	
Have you brought the tests carried out so far in your pregnancy? (n=111)			.56
No	8 (14.3)	5 (9.1)	
Yes	48 (85.7)	50 (90.9)	
Do you understand why you had these tests? (n=111)			.12
No	9 (16.1)	3 (5.5)	
Yes	47 (83.9)	52 (94.5)	
Do you smoke? (n=111)			.56
No	48 (85.7)	50 (90.9)	
Yes	8 (14.3)	5 (9.1)	
Do you drink alcohol? (n=111)			.68
No	52 (92.9)	53 (96.4)	
Yes	4 (7.1)	2 (3.6)	
How do you rate the quality of your sleep? (n=111)			.15
Bad (0)	14 (25)	7 (12.7)	
Regular or good	42 (75)	48 (87.3)	
How do you rate the quality of your food? (n=110)			.49
Bad (0)	6 (10.9)	3 (5.5)	
It can be better or it is good	49 (89.1)	52 (94.5)	
Do you consider yourself anxious? (n=111)			≥.99
No	5 (8.9)	5 (9.1)	
Yes	51 (91.1)	50 (90.9)	
Do you consider yourself to be self-aware about your health (ie, do you understand your health and what it takes to stay healthy)? (n=111)			.79
No	7 (12.5)	8 (14.5)	
Yes	49 (87.5)	47 (85.5)	
Do you consider yourself happy with your health? (n=111)			.26
No or moderately	27 (48.2)	33 (60)	
Yes or very much	29 (51.8)	22 (40)	
Do you consider preventive care an important aspect of health? (n=111)			.49
No	2 (3.6)	0 (0)	
Yes	54 (96.4)	55 (100)	

[a]Fisher exact test or chi-square test; significance set at $P<.05$.

[b]Not applicable.

Cluster Analysis

Once it was determined which questions from Q1 would be included in the cluster analysis, one patient who did not adequately answer two or more of them ended up being excluded, which resulted in the analysis of data from 110 pregnant women. According to the dendrogram, the greatest leap in the Ward algorithm was from 86% to 93%, resulting in three groups of pregnant women (Figure 2).

Cluster 1 included 39.1% (n=43) of the participants, Cluster 2 had 21.8% (n=24), and Cluster 3 also included 39.1% (n=43). Pregnant women in Cluster 1 attended an average of 10.3 antenatal appointments, with a median of 11 (IQR 1-23). Participants in Cluster 2 had an average of 6.9 appointments, with a median of 6.5 (IQR 1-12). Finally, Cluster 3 showed an average of 9.1 antenatal appointments, with a median of 9 (IQR 1-24).

Figure 3 shows how significant variables are distributed among clusters, based on the positively significant answers to Q1. A

XSL•FO
RenderX

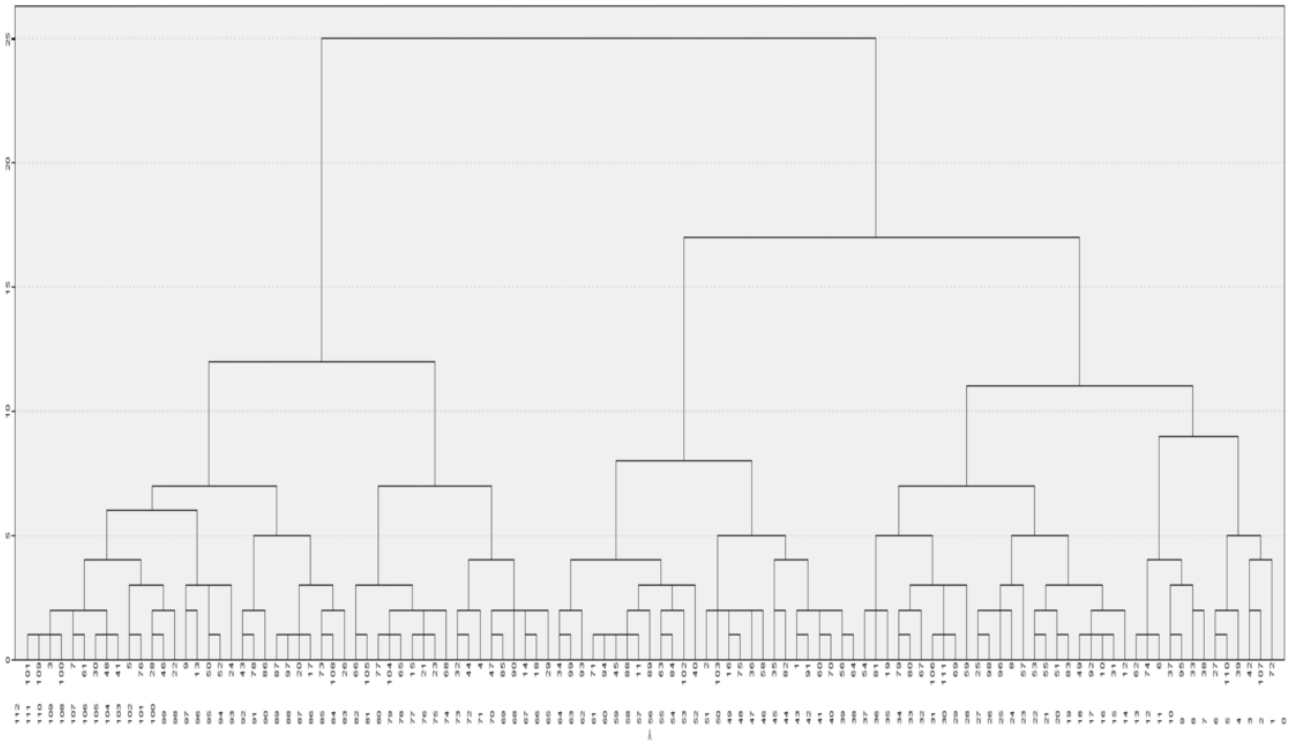
visual representation of the groups in nodes and their connections was created to show the associations between the different variables within clusters (Figure 4).

Acceptance and use of the app varied between the different clusters, with Cluster 1 showing the highest adherence and Cluster 2 the lowest (P=.08). Table 2 shows the differences in adherence to the app among clusters, including participants in the control group.

The outcomes that stood out among clusters were the number of antenatal appointments and the route of delivery. In Cluster 1, 72.1% (n=31) of births were vaginal and 27.9% (n=12) were cesarean. Similarly, Cluster 2 had 79.2% (n=19) vaginal births and 20.8% (n=5) cesarean. In contrast, Cluster 3 showed a markedly different pattern, with only 4.6% (n=2) vaginal births and 95.4% (n=41) cesarean sections.

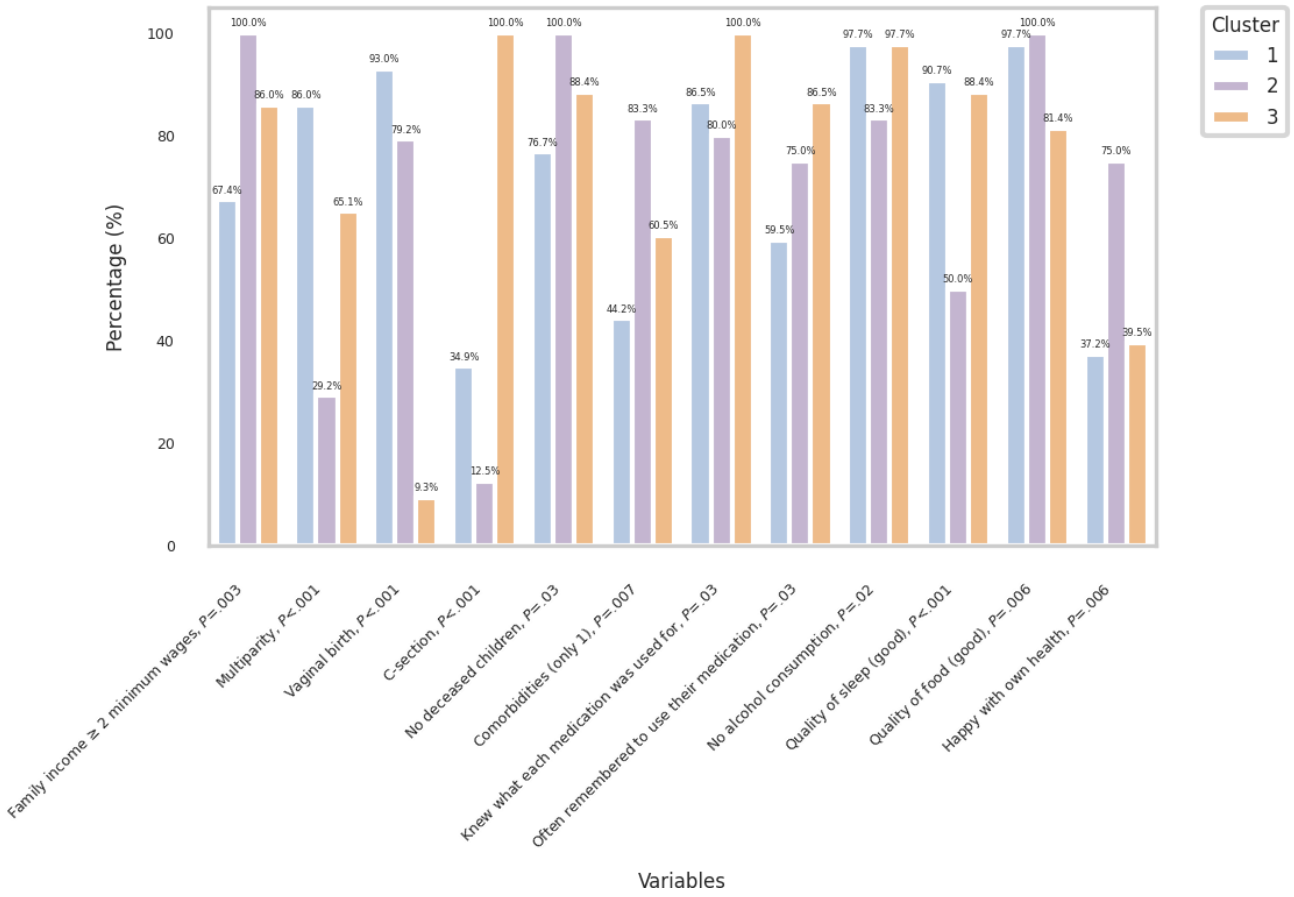
The average number of emergency consultations in Cluster 2 was the lowest (3.5) compared to Cluster 1 (5) and Cluster 3 (5; P=.62).

Figure 2. Dendrogram using Ward linkage—rescaled distance cluster combination.



XSL•FO
RenderX

Figure 3. Variables per cluster.



XSL•FO
RenderX

Figure 4. Nodes and connections per cluster.

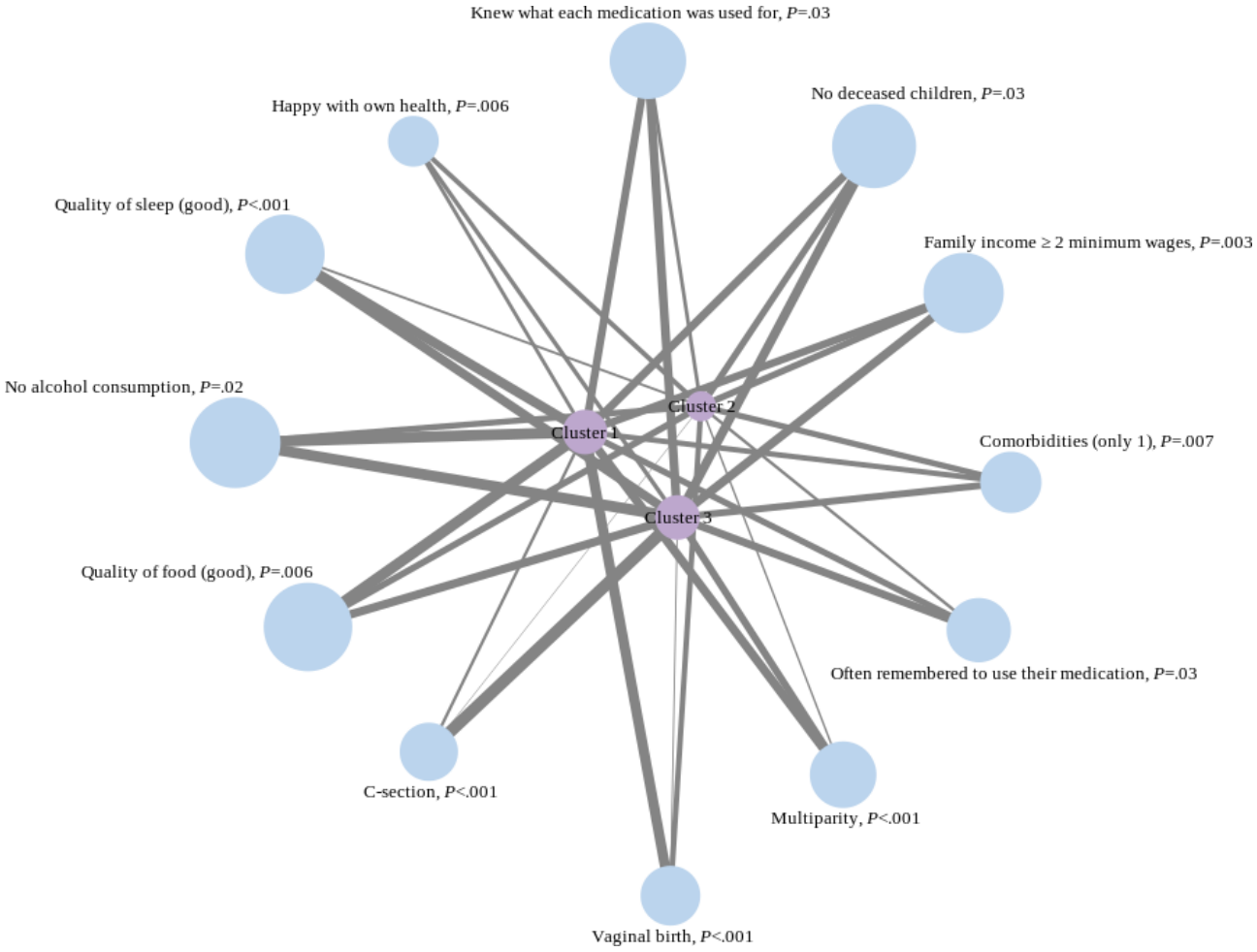


Table . Comparison of patient clusters according to adherence to the app[a,b].

Cluster	Control, n (%)	Nonadherent, n (%)	Adherent, n (%)
1	18 (32.7)	12 (35.3)	13 (61.9)
2	15 (27.3)	8 (23.5)	1 (4.8)
3	22 (40)	14 (41.2)	7 (33.3)
Total	55 (100)	34 (100)	21 (100)

[a]Chi-square test; significance set at $P<.05$.

[b]$P=.12$ for the overall comparison across the entire table.

Clustering

In this study, clustering into three groups revealed significant differences in biopsychosocial aspects, which were named by the researchers according to the most striking characteristics of each group (Graph 2). Cluster 1—“experienced, disciplined mothers”—is characterized by multiparous women with lower incomes who had vaginal births, most of whom had more than one comorbidity, a longer history of deceased children, and good lifestyle habits, but who forget to take their medication and consider themselves less happy with their health. Cluster 2—“inexperienced pregnant women with limited involvement in care”—has the highest family income of the three clusters, with a predominance of women in their first pregnancy (nulliparous), with only one comorbidity and no deceased children; despite being the happiest with their health and considering the quality of their diet to be good, most of the patients here ignored what each medication was for, forgot to use their medication, and had the worst lifestyle habits, with higher alcohol consumption during pregnancy and worse quality of sleep. Cluster 3—“conscious multiparous”—is characterized by patients with more than one child, with high rates of cesarean sections, and mostly with only one comorbidity; these are the patients who best understand about the use of their medications and the ones who most remember to take them; although they sleep better than others, their diet is the worst of all clusters.

XSL•FO
RenderX

Discussion

Health Awareness

An app specifically designed for antenatal preparation in high-risk pregnancies has shown significant potential for enhancing health awareness and improving perinatal outcomes. Sakamoto et al [7] regard mHealth as an effective perinatal tool due to its cost-effectiveness and broad availability, as well as its capacity to empower pregnant women to manage their own health, promote healthy lifestyles, and access information anytime, anywhere. In our study, most patients brought their pregnancy-related test results to their antenatal appointments, although approximately 14.3% (n=8) of the control group and 9.1% (n=5) of the app group did not, suggesting gaps in patient education or communication at primary health units. Additionally, a portion of patients was unaware of the purpose of these tests, indicating potential shortcomings in the information provided to them during routine prenatal care.

The question “Do you consider yourself self-aware about your health?” highlighted that a minority of participants (n=7, 12.5%) in the control group and 14.5% (n=8) in the app group lacked a clear understanding of their health conditions. However, most patients consider preventive care to be an important aspect of health (n=54, 96.4% of the control group and n=55, 100% of the app group). DesJarlais [27] showed that, according to the Health Belief Model, women are more likely to seek preventive care if they believe that they are at risk and that the measures will bring more benefits than disadvantages. Many women are unaware of the benefits of antenatal care or have negative attitudes toward it, which may discourage them from seeking it even when it is available. Thus, instructing women on the importance of antenatal care is an important step to reduce health disparities [27], and mHealth proves to be a useful tool to increase women’s confidence and ability to make informed decisions about their care [28].

Health Apps

Studies indicate that women’s health awareness, particularly during pregnancy, can significantly impact the overall health and well-being of both mother and baby, and that mHealth apps play a vital role in this period [29]. Women who have faced a clinical issue during pregnancy tend to increase their focus on health. Rossiter et al [30] observed that pregnant women’s motivation to adopt healthier behaviors was driven by the desire to avoid future health problems and improve overall well-being. The use of mHealth interventions, such as a health app for appointment preparation, consistently improves patients’ self-efficacy, providing tools for self-monitoring and self-care and improving health outcomes [7,29,31,32].

Adherence to an app depends on factors such as perceived usefulness, ease of use, performance, social influence, motivation, trust, privacy and security concerns, technology anxiety, personalization, and support [33-35]. Tailored content, a user-friendly interface, and technical stability improve experience and adherence [36], while demonstrated health benefits further drive adoption [37]. Conversely, technical issues, privacy worries, waning interest, and time constraints are primary barriers to continued engagement [34,36,38]. Addressing these can markedly enhance effectiveness [34].

Health professionals also play an important role by fostering awareness and motivation, providing credibility when endorsing apps [38], and reinforcing perceived usefulness and intent to use [39]. Through open dialogue, physical, and psychosocial support, especially in high-risk pregnancies, they help prevent distorted risk perceptions and ensure patients accept guidance and feedback during antenatal care [40-42]. Ultimately, adopting a person-centered approach through digital health apps such as the Health Assistant app offers promising opportunities to actively involve patients and families in antenatal care, leading to improved health care experiences and outcomes [6].

App Usability: MAUQ

In this study, the Health Assistant app demonstrated good usability among high-risk pregnant women, as indicated by an average MAUQ score of 6.2 from the 23 users who engaged with the app. This result highlights the app’s ease of use and its perceived usefulness for antenatal preparation. Few studies have used the MAUQ for mHealth in pregnant women, and none of them in Brazil. A study conducted in Indonesia with 88 pregnant women to evaluate an app during pregnancy obtained an average MAUQ score of 6.26 in areas with a mobile signal, while in areas without a signal, the average was 6.29, which is similar to our finding [43].

Given that usability significantly impacts the adoption and sustained use of mHealth apps, particularly among populations with varying levels of health literacy, the positive usability scores from our participants suggest that the Health Assistant app effectively meets users’ expectations. Ensuring user-friendly interfaces and intuitive navigation remains essential in maximizing the potential benefits of digital health interventions in antenatal care. According to the study of the development and validation study of the original MAUQ, the higher the score, the greater the usability of the app [20]. Therefore, the average score of 6.2 obtained in this study is consistent with good usability.

Clustering

The cluster analysis conducted in this study identified three distinct groups of high-risk pregnant women, each with specific biopsychosocial characteristics. This classification provided insights into behavioral and clinical patterns that may influence pregnancy outcomes and engagement with digital health interventions, detailed in the following sections.

Cluster 1, referred to as “experienced, disciplined mothers,” included mostly multiparous women with lower income and a history of multiple comorbidities. Although they reported healthier lifestyle habits and demonstrated the highest adherence to the Health Assistant app, they also showed a tendency to forget medication and reported lower satisfaction with their health. These findings align with previous research, suggesting multiparous women are more proactive in managing their complex health needs due to accumulated experiences [7,30].

Cluster 2, “inexperienced pregnant women with limited involvement in care,” was composed primarily of first-time

XSL•FO
RenderX

mothers with higher incomes and fewer comorbidities. Despite reporting high satisfaction with their health and good dietary habits, this group had the poorest adherence to the app, the most frequent forgetfulness in medication use, and the highest rates of alcohol consumption and sleep dissatisfaction. This discrepancy between perceived and actual health behaviors aligns with the Health Belief Model, suggesting lower perceived susceptibility due to lack of previous negative health experiences, leading to reduced engagement in preventive behaviors [27,42,44]. Contrary to findings by Lee et al [45], this study observed lower app adherence among first-time mothers, highlighting a need for targeted education and personalized digital health interventions to enhance awareness of risks and benefits [46].

Cluster 3, identified as "conscious multiparous women," exhibited moderate app adherence, better medication management, and high rates of cesarean deliveries. These women demonstrated a pragmatic and consistent adherence to medical recommendations, reflecting heightened awareness from previous pregnancy experiences [2,47]. Despite their good medication adherence and understanding of treatments, they reported poorer dietary habits, suggesting opportunities for targeted nutritional interventions.

Recognizing these distinct cluster profiles enables the development of tailored interventions to address the unique needs and challenges of each group, ultimately improving maternal-fetal outcomes and optimizing the use of mHealth tools in antenatal care. The main differences among the clusters are related to sociodemographic factors, obstetric history, and patient behavior. A detailed analysis of these aspects is essential for designing personalized strategies in education, adherence, and psychosocial support.

Family Income

During the study period, the minimum wage in Brazil corresponded to approximately 260 dollars [48,49]. Patients with higher incomes and a better understanding of antenatal care may require less intensive interventions. In contrast, those from groups characterized by lower incomes and higher prevalence of comorbidities often need closer monitoring and additional support.

Cluster 2 stood out as all participants reported an income equal to or greater than two minimum wages. Conversely, Cluster 1 showed a significantly higher proportion of patients with lower income (n=14, 32.6% had incomes below two minimum wages). Interestingly, despite previous evidence suggesting that pregnant women from remote areas or lower socioeconomic backgrounds might experience difficulties accessing or effectively using mHealth technologies, our findings indicated the opposite [50]. We observed greater engagement with the app, precisely among participants with lower incomes (Cluster 1). This unexpected finding could be explained by the higher prevalence of comorbidities among these patients, motivating them to seek more active participation in their antenatal care. Additionally, the high proportion of multiparous women in this group (n=37, 86%) might indicate greater health awareness resulting from prior pregnancy experiences, further explaining their increased adherence [51].

Medical History and Medication Adherence

In this study, multiparous women predominated in Clusters 1 and 3, while nulliparous women were the majority in Cluster 2. First-time pregnant women usually experience heightened awareness of bodily changes, possibly due to the novelty and uncertainty associated with pregnancy [52]. In this context, mHealth apps can serve as important tools, offering ongoing support and reassurance regardless of parity. Multiparous women, in contrast, often have better psychosocial adaptation during pregnancy and a greater appreciation for organization and planning based on previous experiences. This previous knowledge might explain their higher engagement with mHealth interventions observed in our results. Specifically, multiparous participants demonstrated greater adherence to the app, likely due to their familiarity with pregnancy-related risks and a proactive approach shaped by prior pregnancies. Thus, in our high-risk pregnancy context, previous experiences seem to strongly influence patients' adherence, highlighting the importance of addressing both the informational needs of first-time mothers and reinforcing engagement strategies tailored for women with prior pregnancies [51].

Previous Losses and Medication Adherence

The group consisting predominantly of first-time mothers with limited involvement in care (Cluster 2) stood out because none of these patients reported previous child losses. These women were characterized by fewer comorbidities (n=20, 83.3%) with only one comorbidity, possibly reflecting better overall health conditions. This may explain their lower adherence to the app, since fewer negative health experiences could lead to a reduced perceived need for additional monitoring and intervention [44]. Conversely, experienced mothers with disciplined engagement (Cluster 1) reported a higher incidence of having lost a child (n=10, 23.3%). This group's socioeconomic profile—with lower income, a higher number of children, and more frequent comorbidities—likely contributed to these experiences. Such adverse life events might have fostered greater health awareness, thus motivating higher adherence to the mHealth intervention observed in this group.

Medication adherence is notably challenging among patients with chronic conditions, particularly in high-risk pregnancies [53]. Our findings demonstrated strong adherence among patients from Cluster 3, who all clearly understood the purpose of each medication prescribed. The fact that these patients were multiparous and had previously undergone cesarean sections further indicates heightened health awareness. This awareness translated into practical adherence behaviors, with the majority of these patients (n=32, 86.5%; P=.03) consistently taking their medications without forgetting. Understanding their medical conditions and the intended benefits of treatments clearly improved medication adherence in this group [6,54].

Self-Awareness and Medication Management

Cluster 2 exhibited the lowest level of awareness regarding medication use, with 20% (n=4) of patients reporting uncertainty about the purpose of each medication. This was unexpected, given that this group was composed mostly of first-time pregnant women with higher income and fewer comorbidities—typically

XSL•FO
RenderX

factors associated with greater engagement in health management. Conversely, Cluster 1 had the highest percentage (n=15, 40.5%) of patients forgetting to take medications. This vulnerability can be explained by the combination of lower income, higher number of children, and increased comorbidities observed in this group, resulting in more complex daily health demands and possibly less family or social support [55,56].

Although the app used in this study did not provide medication reminders, patients who actively used the app generally showed greater health engagement [28]. Apps have been recognized for their potential to facilitate medication adherence by increasing patient awareness and organization [28]. In this study, Cluster 1 patients were most engaged with the app; however, longer follow-up would be needed to determine whether app use directly influenced improved medication adherence over time. Given these observations, both Clusters 1 and 2 might benefit significantly from incorporating targeted medication adherence features into mHealth interventions [57,58].

Lifestyle Habits

Adopting healthy lifestyle habits, such as regular physical activity, balanced nutrition, and adequate sleep, is highly recommended, particularly in high-risk pregnancies [59,60]. In this study, we identified differences in alcohol consumption between clusters. Cluster 2 (inexperienced pregnant women with limited involvement in care) presented the highest prevalence of alcohol use during pregnancy (n=4, 16.7%). This finding was unexpected since women with higher socioeconomic status generally demonstrate fewer risk behaviors during pregnancy [44]. Digital interventions, such as mHealth apps, have been shown to effectively reduce alcohol consumption in pregnant populations [44]. Therefore, this cluster could benefit significantly from increased engagement with such interventions.

Another notable aspect was sleep quality. Half of the patients in Cluster 2 reported inadequate sleep (P<.001). Enhanced app engagement might benefit these women, as mHealth-based interventions have previously demonstrated the ability to improve sleep quality and duration during pregnancy, positively influencing maternal health outcomes [61]. Diet quality perceptions varied among clusters. Interestingly, all patients in Cluster 2 rated their diet quality positively, whereas Cluster 3 was more critical (n=35, 81.4% considered it good; P=.006). It is possible that patients in Cluster 3, who displayed greater awareness and realistic views about medication use, similarly had more accurate self-assessments regarding their dietary habits.

We also observed significant differences between clusters in self-perceived happiness regarding health. Cluster 2 showed the highest proportion of women feeling happy with their health (n=18, 75%), whereas Cluster 1 had the lowest (n=16, 37.2%). This difference may be explained by the higher proportion of multiparous women in Cluster 1, whose additional life stressors and caregiving demands could negatively impact perceived happiness [62]. Additionally, the inherent challenges of high-risk pregnancies might further influence women's emotional state and perceptions of happiness. Despite such challenges, previous research indicates that positive emotional states, such as happiness, can persist during high-risk pregnancies, particularly when reinforced by strong family support and personal resilience.

Clusters, Outcomes, and Health Strategies

Identifying patient clusters allows targeted strategies to prevent adverse outcomes. Our findings revealed notable differences between groups regarding antenatal appointment attendance, route of delivery, and health behaviors.

The number of antenatal appointments varied significantly across clusters. Cluster 2 ("inexperienced pregnant women with limited involvement in care") attended fewer appointments (6.9) compared to Clusters 1 (10.3) and 3 (9.1). This frequency falls short of World Health Organization (WHO) recommendations (minimum 8 visits), highlighting the need for intensified early follow-up and educational interventions, particularly considering their low adherence to the app [56,63]. Conversely, Cluster 1 attended the most appointments, likely due to greater medical complexity and higher app engagement [64]. Although the app did not send reminders, evidence from previous studies supports that increased mHealth use correlates with higher antenatal attendance, suggesting these groups could benefit from enhanced digital interventions [65]. It is noteworthy that, although statistically insignificant, Cluster 2 had fewer emergency department visits (3.5 visits). This can be related to the higher satisfaction with health in Cluster 2, but it can also have to do with higher income, which suggests that these patients, while receiving some care in the private sector, are perhaps keeping less frequent appointments only to retain a link with the public health system for childbirth.

Regarding delivery mode, Cluster 3 had significantly more cesarean sections (n=41, 95.3%) compared to Clusters 1 (n=12, 27.9%) and 2 (n=5, 20.8%; P<.001). This was likely due to the medical complexity of these multiparous patients and previous cesarean histories rather than mere personal preference [66,67]. Reducing unnecessary C-sections remains essential; thus, targeted education through the app could encourage safer delivery alternatives.

Pregnancy itself is a stimulus for lifestyle changes and positive health interventions [60], but cluster-based management strategies implemented by health teams could substantially improve pregnancy outcomes. Cluster 1**,** characterized by multiparity, low income, and high comorbidity, would benefit from interventions focusing on medication adherence, dietary optimization, and family planning guidance. Financial pressures in this group highlight the potential of mHealth tools to support self-management and health education [11]. It would also be interesting to enter information about these needs in the Health Assistant, as this was the group that adhered most to the app [68]. This would be in line with the findings of Wang et al [69], according to which patients expressed the need for apps containing evidence-based, personalized information to support them during pregnancy.

Cluster 2, despite higher socioeconomic status and fewer comorbidities, exhibited poor health behaviors such as alcohol use, inadequate sleep, and limited health awareness. Strategies to enhance app adherence, including personalized information and notifications, could significantly improve self-care and

XSL•FO
RenderX

antenatal attendance. Integrating psychological support and mental health assessments into antenatal care can also address emotional well-being, increase self-awareness and self-care, and reduce risky behaviors [29].

Cluster 3, multiparous health-conscious women, demonstrated moderate app adherence (n=7, 33.3%) but lower self-rated diet quality. Targeted educational campaigns on nutrition, lifestyle, and informed choices regarding delivery methods could enhance patient engagement, improve maternal-fetal outcomes, and reduce unnecessary interventions [59]. Strategies for this group with moderate adherence to the app (n=7, 33.3%) could include educational campaigns to raise awareness about the benefits of the app, provide additional support to encourage its use, and increase attendance at antenatal appointments [70,71].

The cluster analysis presented in this study offers prenatal care teams a practical framework for stratifying pregnant patients according to their biopsychosocial risk profiles. By distinguishing groups with low engagement in antenatal care (Cluster 2) or difficulties in medication adherence (Cluster 1), health care professionals can more effectively identify individuals at heightened risk for disengagement or adverse pregnancy outcomes. Early recognition of these specific risk profiles enables teams to implement targeted, evidence-based interventions—such as personalized education, adherence support, and psychosocial assistance—before complications arise. This cluster-based approach facilitates efficient allocation of resources and proactive management, supporting improved maternal-fetal outcomes [6]. Furthermore, integrating mHealth solutions tailored to each cluster's unique needs can enhance patient engagement and self-management, ultimately strengthening the quality and continuity of antenatal care.

Implications for Future Research: "Clusters in Practice: A Proposal for a Risk Assessment Questionnaire for the First Antenatal Visit"

Based on the questions that most clearly differentiated our clusters, we propose developing a brief screening questionnaire to be administered at the first antenatal visit. This questionnaire would automatically indicate the most likely cluster for each patient, enabling rapid and specific interventions at the outset of prenatal care, thus preventing risks before undesirable behaviors or complications occur. Subsequent steps would involve drafting pilot items covering income, parity, comorbidities, and mHealth use profiles, submitting them to expert review and cognitive interviews with pregnant women for clarity and relevance, conducting a pilot study to assess construct validity (via factor analysis), internal consistency (Cronbach α), and test-retest reliability, and performing a prospective multicenter validation across diverse clinical settings to compare predicted risk scores with actual app engagement, appointment frequency, and maternal-fetal outcomes for sensitivity, specificity, and likelihood ratio estimation. Once validated, an algorithm could be embedded in eHealth records to assign each patient to a cluster and automatically trigger tailored interventions (eg, telemonitoring, reinforced health education, and SMS text message reminders), culminating in a randomized controlled trial contrasting "cluster‐guided" care with usual care using primary end points such as antenatal adherence rates, complication incidence, and health care resource utilization—an approach that, while consistent with established predictive‐health models, will require rigorous clinical validation and user‐acceptance testing to ensure accuracy, usability, and effectiveness before widespread implementation.

Limitations

Though this study has provided us with several insights, it does have some limitations. Many patients were lost during data collection, which is often expected in observational longitudinal studies, but this loss, especially among prospective app users, can make it difficult to assess accurately the effectiveness of the intervention, since part of the target group was not thoroughly followed up. Since the transition from physical to digital records is still in progress, it is not uncommon for patients to resist mHealth adherence at first, and sample reduction in the adherent group in this study seems to have been due to such barriers as low literacy, limited connectivity, low smartphone memory, cultural beliefs, and lack of interest.

The fact that this study was carried out in a public university hospital in southern Brazil may render some generalizations invalid for other regions of the country or for private institutions. The adoption of three clusters yielded less variability in answers and more generalized results. Some sociodemographic factors overlapped among clusters, making it difficult for groups to be clearly distinguished and confusing for outcomes to be analyzed. In addition, the duration of the study, restricted to the period between the first antenatal visit and delivery, may not be sufficient to assess the long-term impacts of app use on maternal-fetal health and the well-being of pregnant women, thus indicating the need for further studies with larger samples and in different socioeconomic contexts.

Conclusions

The use of mobile apps has significant potential to optimize antenatal care and self-awareness for high-risk pregnant women. The usability of the Health Assistant app is in line with the standards expected and observed in other studies. Cluster analysis revealed different profiles of pregnant women, making it possible to identify specific groups that could benefit most from digital interventions and from a personalized approach to organizing their journey and improving pregnancy outcomes. This pioneering study stresses the role of mHealth in high-risk antenatal care and highlights the importance of further exploration of its impact in different settings and populations.

Acknowledgments

The authors gratefully acknowledge the assistance of researchers Nycole de Oliveira de Andrade, Evelize Behrens, and Leticia Petronzelli with data collection, Marcia Olandoski with statistical analysis, and Afonso Deres with the graphs.

XSL•FO
RenderX

Conflicts of Interest

None declared.

References

1. mHealth: New Horizons for Health through Mobile Technologies: Second Global Survey on eHealth: World Health Organization; 2011.
2. Mehralizade A, Schor S, Coleman CM, et al. Mobile health apps in OB-GYN-embedded psychiatric care: commentary. JMIR Mhealth Uhealth 2017 Oct 6;5(10):e152. [doi: 10.2196/mhealth.7988] [Medline: 28986335]
3. Lívia Moura P, Camargo G. Impacto econômico e social do android no Brasil [Report in Portugese]. : Bain & Company; 2020.
4. Dias JD, Zem-Mascarenhas SH. Mobile applications developed for the health sector in Brazil: an integrative literature review. Nurs J Minas Gerais 2014;18(2). [doi: 10.5935/1415-2762.20140035]
5. Jones N, Moffitt M. Ethical guidelines for mobile app development within health and mental health fields. Prof Psychol Res Pract 2016;47(2):155-162. [doi: 10.1037/pro0000069]
6. Executive Board. mHealth—use of appropriate digital technologies for public health: report by the director-general. : World Health Organization; 2017 142.
7. Sakamoto JL, Carandang RR, Kharel M, et al. Effects of mHealth on the psychosocial health of pregnant women and mothers: a systematic review. BMJ Open 2022 Feb 15;12(2):e056807. [doi: 10.1136/bmjopen-2021-056807] [Medline: 35168981]
8. DeNicola N, Grossman D, Marko K, et al. Telehealth interventions to improve obstetric and gynecologic health outcomes: a systematic review. Obstet Gynecol 2020 Feb;135(2):371-382. [doi: 10.1097/AOG.0000000000003646] [Medline: 31977782]
9. Prentice C, Peven K, Zhaunova L, et al. Methods for evaluating the efficacy and effectiveness of direct-to-consumer mobile health apps: a scoping review. BMC Digital Health 2024;2(1). [doi: 10.1186/s44247-024-00092-x]
10. Salazar A, de Sola H, Failde I, Moral-Munoz JA. Measuring the quality of mobile apps for the management of pain: systematic search and evaluation using the Mobile App Rating Scale. JMIR Mhealth Uhealth 2018 Oct 25;6(10):e10718. [doi: 10.2196/10718] [Medline: 30361196]
11. Marengo LL, Kozyreff AM, Silva Moraes F, Maricato LIG, Barberato-Filho S. Mobile technologies in healthcare: reflections on development, application, legal aspects, and ethics. Pan Am J Public Health 2022. [doi: 10.26633/RPSP.2022.37]
12. Rodrigues PB, Zambaldi CF, Cantilino A, Sougey EB. Particularidades da gravidez de alto risco como fatores para o desenvolvimento de sofrimento mental [Article in Portugese]. Trends Psychiatry Psychother 2016;38(3):136-140. [doi: 10.1590/2237-6089-2015-0067]
13. Dulude D, Bélanger C, Wright J, Sabourin S. High-risk pregnancies, psychological distress, and dyadic adjustment. J Reprod Infant Psychol 2002 May;20(2):101-123. [doi: 10.1080/02646830220134612]
14. Mirzakhani K, Ebadi A, Faridhosseini F, Khadivzadeh T. Well-being in high-risk pregnancy: an integrative review. BMC Pregnancy Childbirth 2020 Sep 11;20(1):526. [doi: 10.1186/s12884-020-03190-6] [Medline: 32912254]
15. Stewart M, Burmeister A, da RS, Lopes JMC. Patient-Centered Medicine: Transforming the Clinical Method, 3rd edition 2017.
16. Iribarren SJ, Cato K, Falzon L, Stone PW. What is the economic evidence for mHealth? A systematic review of economic evaluations of mHealth solutions. PLoS One 2017;12(2):e0170581. [doi: 10.1371/journal.pone.0170581] [Medline: 28152012]
17. Banjari I, Kenjerić D, Šolić K, Mandić ML. Cluster analysis as a prediction tool for pregnancy outcomes. Coll Antropol 2015 Mar;39(1):247-252. [Medline: 26040101]
18. Nair BJB, Adarsh CK, Nishin S, Nihar KP, Shivaranjith P. A comparative approach to cluster maternal data using k-means and k-medoid. Int J Res Pharm Sci 2019;10(3):2038-2043. [doi: 10.26452/ijrps.v10i3.1416]
19. Gilligan C, Sanson-Fisher R, Eades S, D'Este C, Kay-Lambkin F, Scheman S. Identifying pregnant women at risk of poor birth outcomes. J Obstet Gynaecol 2009 Apr;29(3):181-187. [doi: 10.1080/01443610902753713] [Medline: 19358020]
20. Zhou L, Bao J, Setiawan IMA, Saptono A, Parmanto B. The mHealth App Usability Questionnaire (MAUQ): development and validation study. JMIR Mhealth Uhealth 2019 Apr 11;7(4):e11500. [doi: 10.2196/11500] [Medline: 30973342]
21. Madhulatha TS. An overview on clustering methods. IOSRJEN 2012 Apr;02(4):719-725. [doi: 10.9790/3021-0204719725]
22. Dave M, Gianey H. Different clustering algorithms for big data analytics: a review. Presented at: 2016 International Conference System Modeling & Advancement in Research Trends (SMART); Nov 25-27, 2016; Moradabad, India. [doi: 10.1109/SYSMART.2016.7894544]
23. Li T. A unified view on clustering binary data. Mach Learn 2006 Mar;62(3):199-215. [doi: 10.1007/s10994-005-5316-9]
24. Faizan M, F. M, Ismail S, Sultan S. Applications of clustering techniques in data mining: a comparative study. IJACSA 2020;11(12). [doi: 10.14569/IJACSA.2020.0111218]

XSL•FO
RenderX

25. Ferreira RRM, GSi R, Castro GSA. Ana´lise de cluster na˜o supervisionado em r: agrupamento hiera´rquico [Web page in Portugese]. EMBRAPA. 2020. URL: https://www.embrapa.br/busca-de-publicacoes/-/publicacao/1126478/analise-de-cluster-nao-supervisionado-em-r-agrupamento-hierarquico [accessed 2025-08-08]
26. Dalmaijer ES, Nord CL, Astle DE. Statistical power for cluster analysis. BMC Bioinformatics 2022 Dec;23(1). [doi: 10.1186/s12859-022-04675-1]
27. DesJarlais R. Neonatal health outcomes and access to care. Nursing Senior Theses 2020(8). [doi: 10.33015/dominican.edu/2020.NURS.ST.18]
28. Ledford CJW, Canzona MR, Cafferty LA, Hodge JA. Mobile application as a prenatal education and engagement tool: a randomized controlled pilot. Patient Educ Couns 2016 Apr;99(4):578-582. [doi: 10.1016/j.pec.2015.11.006]
29. Sandborg J, Henriksson P, Larsen E, et al. Participants' engagement and satisfaction with a smartphone app intended to support healthy weight gain, diet, and physical activity during pregnancy: qualitative study within the healthymoms trial. JMIR Mhealth Uhealth 2021 Mar 5;9(3):e26159. [doi: 10.2196/26159] [Medline: 33666554]
30. Rossiter C, Henry A, Roberts L, et al. Optimising mothers' health behaviour after hypertensive disorders of pregnancy: a qualitative study of a postnatal intervention. BMC Public Health 2022 Jun 27;22(1):1259. [doi: 10.1186/s12889-022-13590-2] [Medline: 35761317]
31. Pouriayevali B, Ehteshami A, Kohan S, Saghaeiannejad-Isfahani S. Functionality of self-care for pregnancy mobile applications: a review study. J Educ Health Promot 2022;11(1):415. [doi: 10.4103/jehp.jehp_1429_21] [Medline: 36824402]
32. Dol J, Richardson B, Murphy GT, Aston M, McMillan D, Campbell-Yeo M. Impact of mobile health interventions during the perinatal period on maternal psychosocial outcomes: a systematic review. JBI Evid Synth 2020 Jan;18(1):30-55. [doi: 10.11124/JBISRIR-D-19-00191] [Medline: 31972680]
33. Kaium M, Bao Y, Alam MZ, Hasan N, Hoque M. Understanding the insight of factors affecting mHealth adoption. IJRBS 2019;8(6):181-200. [doi: 10.20525/ijrbs.v8i6.522]
34. Szinay D, Jones A, Chadborn T, Brown J, Naughton F. Influences on the uptake of and engagement with health and well-being smartphone apps: systematic review. J Med Internet Res 2020 May 29;22(5):e17572. [doi: 10.2196/17572] [Medline: 32348255]
35. Martins NLM, Duarte P, Pinho J. An analysis of determinants of the adoption of mobile health (mHealth). Rev Adm Empres 2021;61(4). [doi: 10.1590/s0034-759020210403x]
36. Jakob R, Harperink S, Rudolf AM, et al. Factors influencing adherence to mHealth apps for prevention or management of noncommunicable diseases: systematic review. J Med Internet Res 2022 May 25;24(5):e35371. [doi: 10.2196/35371] [Medline: 35612886]
37. Palos-Sanchez PR, Saura JR, Rios Martin M, Aguayo-Camacho M. Toward a better understanding of the intention to use mHealth apps: exploratory study. JMIR Mhealth Uhealth 2021;9(9):e27021. [doi: 10.2196/27021]
38. Wang C, Qi H. Influencing factors of acceptance and use behavior of mobile health application users: systematic review. Healthcare (Basel) 2021;9(3):357. [doi: 10.3390/healthcare9030357]
39. Binyamin SS, Zafar BA. Proposing a mobile apps acceptance model for users in the health area: a systematic literature review and meta-analysis. Health Inf J 2021;27(1):1460458220976737. [doi: 10.1177/1460458220976737] [Medline: 33438494]
40. Bedaso A, Adams J, Peng W, Xu F, Sibbritt D. An examination of the association between marital status and prenatal mental disorders using linked health administrative data. BMC Pregnancy Childbirth 2022 Oct 1;22(1):735. [doi: 10.1186/s12884-022-05045-8] [Medline: 36182904]
41. Çelik FP, Güneri SE. The relationship between adaptation to pregnancy and prenatal attachment in high-risk pregnancies. Psychiatr Danub 2020 Nov;32(Suppl 4):568-575. [Medline: 33212465]
42. Shojaeian Z, Khadivzadeh T, Sahebi A, Kareshki H, Tara F. Perceived risk in women with high risk pregnancy: a qualitative study. Iran J Nurs Midwifery Res 2021;26(2):168-174. [doi: 10.4103/ijnmr.IJNMR_32_20] [Medline: 34036066]
43. Abdullah S, Rinawan FR, Judistiani RTD. Usability assessment using m-Health app usability questionnaire on the iPosyandu application in Central Sulawesi. IJPHS 2023;12(4):1447. [doi: 10.11591/ijphs.v12i4.22930]
44. Nicholls-Dempsey L, Badeghiesh A, Baghlaf H, Dahan MH. How does high socioeconomic status affect maternal and neonatal pregnancy outcomes? A population-based study among American women. Eur J Obstet Gynecol Reprod Biol X 2023 Dec;20:100248. [doi: 10.1016/j.eurox.2023.100248] [Medline: 37876770]
45. Lee Y, Moon M. Utilization and content evaluation of mobile applications for pregnancy, birth, and child care. Healthc Inform Res 2016;22(2):73. [doi: 10.4258/hir.2016.22.2.73]
46. Oh SS, Moon JY, Chon D, et al. Effectiveness of digital interventions for preventing alcohol consumption in pregnancy: systematic review and meta-analysis. J Med Internet Res 2022 Apr 11;24(4):e35554. [doi: 10.2196/35554] [Medline: 35404257]
47. Kabongo EM, Mukumbang FC, Delobelle P, Nicol E. Explaining the impact of mHealth on maternal and child health care in low- and middle-income countries: a realist synthesis. BMC Pregnancy Childbirth 2021 Mar 9;21(1):196. [doi: 10.1186/s12884-021-03684-x] [Medline: 33750340]
48. Lei n.o 14663, de 29 de setembro de [Web page in Portugese]. Presidência da República. 2023. URL: https://www.planalto.gov.br/CCIVIL_03/_Ato2023-2026/2023/Lei/L14663.htm [accessed 2025-08-08]

XSL•FO
RenderX

49. Lei n.o 14358, de 1o de junho de [Web page in Portugese]. Presidência da República. 2022. URL: https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2022/lei/l14358.htm [accessed 2025-08-08]
50. Kusyanti T, Wirakusumah FF, Rinawan FR, et al. Technology-based (Mhealth) and standard/traditional maternal care for pregnant woman: a systematic literature review. Healthcare (Basel) 2022 Jul 12;10(7):1287. [doi: 10.3390/healthcare10071287] [Medline: 35885813]
51. Fair CD, Morrison TE. The relationship between prenatal control, expectations, experienced control, and birth satisfaction among primiparous women. Midwifery 2012 Feb;28(1):39-44. [doi: 10.1016/j.midw.2010.10.013] [Medline: 21458895]
52. Perrykkad K, O'Neill R, Jamadar SD. Sense of self in first-time pregnancy. NPJ Womens Health 2024;2(1). [doi: 10.1038/s44294-024-00018-y]
53. Anglada-Martinez H, Riu-Viladoms G, Martin-Conde M, Rovira-Illamola M, Sotoca-Momblona JM, Codina-Jane C. Does mHealth increase adherence to medication? Results of a systematic review. Int J Clin Pract 2015 Jan;69(1):9-32. [doi: 10.1111/ijcp.12582] [Medline: 25472682]
54. Peng Y, Wang H, Fang Q, et al. Effectiveness of mobile applications on medication adherence in adults with chronic diseases: a systematic review and meta-analysis. J Manag Care Spec Pharm 2020 Apr;26(4):550-561. [doi: 10.18553/jmcp.2020.26.4.550] [Medline: 32223596]
55. Pati S, Puri P, Sinha R, Panda M, Pati S. Profile of comorbidity and multimorbidity among women attending antenatal clinics: an exploratory cross-sectional study from Odisha, India. J Family Med Prim Care 2022 May;11(5):1980-1988. [doi: 10.4103/jfmpc.jfmpc_1855_21] [Medline: 35800480]
56. Stegman MM, Lucarelli-Baldwin E, Ural SH. Disparities in high risk prenatal care adherence along racial and ethnic lines. Front Glob Womens Health 2023;4:1151362. [doi: 10.3389/fgwh.2023.1151362] [Medline: 37560034]
57. Tabi K, Randhawa AS, Choi F, et al. Mobile apps for medication management: review and analysis. JMIR Mhealth Uhealth 2019 Sep 11;7(9):e13608. [doi: 10.2196/13608] [Medline: 31512580]
58. Rothschild CW, Dublin S, Brown JS, et al. Use of a mobile app to capture supplemental health information during pregnancy: implications for clinical research. Pharmacoepidemiol Drug Saf 2022 Jan;31(1):37-45. [doi: 10.1002/pds.5320] [Medline: 34216500]
59. WHO Recommendations on Maternal Health: Guidelines Approved by the WHO Guidelines Review Committee, 2nd edition: World Health Organization; 2023:978-992.
60. Office of the Surgeon General (OSG). The Surgeon General's Call to Action to Improve Maternal Health: US Department of Health and Human Services; 2020. [Medline: 33661589]
61. Bartels HC, Kennelly MA, Killeen SL, Lindsay KL, Crowley RK, McAuliffe FM. An mHealth-supported antenatal lifestyle intervention may be associated with improved maternal sleep in pregnancy: secondary analysis from the PEARS trial. BJOG 2022 Dec;129(13):2195-2202. [doi: 10.1111/1471-0528.17267] [Medline: 35876246]
62. Blake SM, Kiely M, Gard CC, El-Mohandes AAE, El-Khorazaty MN, NIH-DC Initiative. Pregnancy intentions and happiness among pregnant black women at high risk for adverse infant health outcomes. Perspect Sex Reprod Health 2007 Dec;39(4):194-205. [doi: 10.1363/3919407] [Medline: 18093036]
63. WHO Recommendations on Antenatal Care for a Positive Pregnancy Experience: Geneva: World Health Organization; 2016.
64. Ditzel AN, Júnior EDS, Breitkreitz VS, et al. Manejo da gravidez de risco e assistência ao pré-natal [Article in Portugese]. Braz J Implantol Health Sci 2024;6(2):2393-2408. [doi: 10.36557/2674-8169.2024v6n2p2393-2408]
65. Mishra M, Parida D, Murmu J, et al. Effectiveness of mHealth interventions for monitoring antenatal care among pregnant women in low- and middle-income countries: a systematic review and meta-analysis. Healthcare (Basel) 2023;11(19):2635. [doi: 10.3390/healthcare11192635]
66. Soares KB, Klein VCG, Lima JD, Gadenz L, Paulo LE, Konopka CK. Gestational risk as a determining factor for cesarean section according to the robson classification groups. Rev Bras Ginecol Obstet 2021 Feb;43(2):084-090. [doi: 10.1055/s-0040-1718446]
67. Antunes MB, Rossi RM, Pelloso SM. Relationship between gestational risk and type of delivery in high risk pregnancy. Rev Esc Enferm USP 2020;54:e03526. [doi: 10.1590/s1980-220x2018042603526] [Medline: 32667388]
68. Dalton JA, Rodger D, Wilmore M, et al. The health-e Babies app for antenatal education: feasibility for socially disadvantaged women. PLOS One 2018;13(5):e0194337. [doi: 10.1371/journal.pone.0194337] [Medline: 29768407]
69. Wang N, Deng Z, Wen LM, Ding Y, He G. Understanding the use of smartphone apps for health information among pregnant chinese women: mixed methods study. JMIR Mhealth Uhealth 2019 Jun 18;7(6):e12631. [doi: 10.2196/12631] [Medline: 31215516]
70. Global Diffusion of Ehealth: Making Universal Health Coverage Achievable: Report of the Third Global Survey on Ehealth: World Health Organization; 2016.
71. Overdijkink SB, Velu AV, Rosman AN, van Beukering MD, Kok M, Steegers-Theunissen RP. The usability and effectiveness of mobile health technology-based lifestyle and medical intervention apps supporting health care during pregnancy: systematic review. JMIR Mhealth Uhealth 2018 Apr 24;6(4):e109. [doi: 10.2196/mhealth.8834] [Medline: 29691216]

XSL•FO
RenderX

Abbreviations

MAUQ: Mobile App Usability Questionnaire
mHealth: mobile health
WHO: World Health Organization

Edited by A Kushniruk; submitted 18.10.24; peer-reviewed by J Pecora, M Waring; revised version received 07.06.25; accepted 23.06.25; published 21.08.25.

Please cite as:
Schier de Fraga F, Narita MM, Schreiner M, Belli F, Leonel Celestino J, Braz Pereira K, Soecki G, Bevervanso V, de Fraga R
Mobile Health Adoption in High-Risk Pregnancies Using Cluster Analysis of Biopsychosocial Outcomes: Observational Longitudinal Cohort Study
JMIR Hum Factors 2025;12:e67680
URL: https://humanfactors.jmir.org/2025/1/e67680
doi:10.2196/67680

© Fernanda Schier de Fraga, Mayara Marenda Narita, Monique Schreiner, Flavio Belli, Jaqueline Leonel Celestino, Karolayne Braz Pereira, Gabriella Soecki, Vitória Bevervanso, Rogério de Fraga. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 21.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Designing Chatbots to Treat Depression in Youth: Qualitative Study

Florian Onur Kuhlmeier[1,2], MSc; Luise Bauch[2], BSc; Ulrich Gnewuch[3], Prof Dr; Stefan Lüttke[2,4], Dipl-Psych

[1]Human-Centered Systems Lab, Institute for Information Systems (WIN), Karlsruhe Institute of Technology, Karlsruhe, Germany

[2]Department of Clinical Psychology and Psychotherapy, Institute of Psychology, University of Greifswald, Franz-Mehring-Straße 47, Greifswald, Germany

[3]Chair of Explainable AI-Based Business Information Systems, School of Business, Economics and Information Systems, University of Passau, Passau, Germany

[4]Department of Clinical Psychology and Psychotherapy, University of Tübingen, Tübingen, Germany

Corresponding Author:
Stefan Lüttke, Dipl-Psych
Department of Clinical Psychology and Psychotherapy, Institute of Psychology, University of Greifswald, Franz-Mehring-Straße 47, Greifswald, Germany

Abstract

Background: Depression is a severe and prevalent mental disorder among youth that requires professional care; however, various barriers hinder access to effective treatments. Chatbots, one of the latest innovations in the research on digital mental health interventions, have shown potential in addressing these barriers. However, most studies on how to design chatbots to treat depression have focused on adult populations or prevention in the general population.

Objective: This study aimed to investigate the problems faced by youth with depression and their adaptive coping strategies, as well as attitudes, expectations, and design preferences for chatbots designed to treat depression.

Methods: We conducted a qualitative study, consisting of a semistructured interview and a concurrent think-aloud session, in which participants interacted with a chatbot prototype with 14 youth with a current or remitted depressive episode.

Results: The participants reported a wide range of problems beyond core depressive symptoms, such as interpersonal challenges, concerns about school and the future, and problems with human therapists. Adaptive coping strategies varied, with most seeking social support or engaging in pleasant activities. Attitudes toward chatbots for depression treatment were predominantly positive, with participants expressing less anxiety about using a chatbot than about seeing a human therapist. Participants showed diverse and partially contradictory design preferences, which included diverse dialogue topics, such as discussing daily life, acute problems, and therapeutic exercises, as well as various preferences for personality, language use, and personalization of the chatbot.

Conclusions: Our study provides a comprehensive foundation for designing chatbots that meet the unique needs and design preferences of youth with depression. These findings can inform the design of engaging and effective chatbots tailored to this vulnerable population.

(JMIR Hum Factors 2025;12:e66632) doi:10.2196/66632

KEYWORDS

human-centered design; depression; mental disorder; mental health; emotional health; youth; teenager; depressive; digital mental health intervention; questionnaire; semi-structured interview; think-aloud; chatbot prototype; remitted depressive; mixed-method; mobile phone

Introduction

Depression is a prevalent mental disorder in youth with significant personal and socioeconomic consequences that requires professional care [1,2]. Despite the availability of effective treatments, such as cognitive behavioral therapy (CBT) [3], accessing them remains challenging. Even when professional services are free and accessible [4], many youth avoid seeking professional support due to perceived stigma, a preference to solve problems by themselves, and fear of psychotherapeutic settings [5-8]. Interestingly, such attitudinal barriers may be more important reasons for not seeking help than structural barriers, such as limited treatment resources and long waiting periods [8]. Digital mental health interventions (DMHIs) are a promising and effective way to overcome these barriers because they provide anonymous and self-empowered access to effective professional care [9]. However, to leverage the full potential of DMHIs in the treatment of depressive symptoms in youth, 2 major limitations need to be overcome,that is, low adherence [10] and difficulty in establishing a therapeutic

XSL•FO
RenderX

alliance, which is viewed as a crucial factor for effective psychotherapy [11].

Chatbots, software systems that interact with users using natural language [12], have shown the potential to address these limitations. They are well accepted, feasible, and have shown promising effectiveness in strengthening mental health [13-16]. In addition, incorporating chatbots into DMHIs improves user engagement [17] and mental health outcomes [18]. Notably, users seem to develop a therapeutic alliance with chatbots [19-21] partly because of social cues such as empathetic messages and humor [15,22]. Despite these encouraging results, most studies on mental health chatbots, including those targeting depression symptoms, have focused on adult populations [13-16,23]. This is a key shortcoming because the results from adult populations cannot be generalized to youth. Youth face significant developmental changes in their biological, psychological, and social systems [24], and depression symptoms differ from those in adulthood, especially at the onset of puberty [25]. Furthermore, youth interact with smartphones and chatbots differently than adults [26,27], and have expressed that existing DMHIs often fail to address their specific concerns adequately [28].

To address this research gap, recent studies have explored the design of chatbots for youth mental health. Høiland et al [29] involved youth in designing a chatbot for high-school health services aimed at preventing mental disorders. Through focus groups, they identified four key support needs: (1) receiving information about mental health, (2) relating to someone beyond their immediate network, (3) receiving support for self-help, and (4) being referred to mental health services. Similarly, Grové [30] developed a preventive mental health chatbot with high-school students. Participants suggested topics such as school, family, friends, sexuality, and identity as well as resources for adaptive coping strategies, mindfulness, and distractions. They expressed a preference for chatbots with inspiring, charismatic, and fun personalities using emojis, humor, and GIFs. While these 2 studies provide valuable insights, they focused on the prevention rather than the treatment of mental disorders and addressed a broad spectrum of mental health issues rather than specifically targeting depression. Thus, there is a critical need for research on how to design chatbots that focus on specific problems of youth with depressive symptoms to achieve sufficient engagement and optimal treatment outcomes.

Our study aims to address this research gap by investigating the following research questions: (1) What problems do youth with depressive symptoms face? (2) What adaptive coping strategies do they apply? (3) What attitudes and expectations do they have for chatbots designed to treat depression among youth? (4) What are their design preferences?

By addressing these questions, we aim to provide a comprehensive foundation for developing chatbot-based DMHIs tailored to the unique needs and preferences of youth with depression. This foundation will facilitate the development of engaging and effective DMHIs for this vulnerable population.

Methods

Study Design

We conducted a qualitative study to examine how to design a chatbot to treat depression in youth. The study included a questionnaire, a semistructured interview, and a concurrent think-aloud session with a chatbot prototype. We chose interviews because they allowed us to gather rich, detailed data on participants' problems, coping strategies, attitudes and expectations, and chatbot design preferences.

Participants

Participants were eligible if they were between 14 and 17 years of age, owned a smartphone, and met diagnostic criteria for a current or remitted depressive episode. Participants with suicidal ideation or psychotic symptoms were excluded. Participants were recruited through a resident child and adolescent psychiatrist and the University's newsletter between June and August 2021.

The target sample size was calculated based on the goal of data saturation [31], considering age group (14 - 15 and 16 - 17 y), depression status (remitted and acutely depressed), and gender (women and men), assuming homogeneity within each subgroup. Given that data saturation can be achieved at 6 cases per homogeneous group [31], the study aimed to recruit 48 participants. The sample was recruited through convenience sampling but the sampling was constrained by the predetermined recruitment period and available recruitment channels. The final sample consisted of 14 participants (12 women and 2 nonbinary).

Procedure

After participants and caregivers provided informed consent, we assessed their eligibility to participate in the study. Next, enrolled participants answered a questionnaire on sociodemographic information and mental health on a laptop. We conducted semistructured interviews to explore participants' experiences with depression, their adaptive coping strategies, attitudes and expectations toward chatbots for depression, and their design preferences. Finally, the participants interacted with a prototype chatbot using the concurrent think-aloud method.

Material

Eligibility Interview

We conducted a semistructured interview to evaluate participants' eligibility to participate in the study. The interview guide included questions regarding age, smartphone ownership, suicidal ideation, and symptoms of depression, and psychotic disorders. The questions on suicidal ideation and symptoms of depression and psychotic disorders were based on the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (*DSM-5*) criteria [32] and Kiddie-Schedule for Affective Disorders and Schizophrenia Present and Lifetime [33]. The complete eligibility interview is presented in Multimedia Appendix 1 .

XSL•FO
RenderX

Questionnaire

Participants completed a questionnaire on demographic characteristics (age, gender, and level of education), history of mental disorders, and previous experience with psychotherapy. We assessed the prevalence of current symptoms of depression using the 8-item Patient Health Questionnaire (PHQ-8) [34] as well as current symptoms of anxiety using the Screen for Child Anxiety Related Disorders (SCARED) [35]. To complement the insights from the semistructured interview on the design preferences, participants answered a questionnaire on potential chatbot capabilities. This questionnaire is based on CBT manuals and literature on the content of DMHIs for depression [36-38] and comprises 13 capabilities that a chatbot to treat depression in youth could implement. For each capability, we asked participants to indicate the extent to which the chatbot should support them on a scale from 1 (“strongly disagree”) to 5 (“strongly agree”). The capabilities are presented in Textbox 1. The participants completed all questionnaires on a laptop using SoSci Survey (SoSci Survey GmbH).

Textbox 1. Items listed in the questionnaire on chatbot capabilities.

The chatbot should support me with…

- how I can become more physically active or do sports.
- how to sleep better.
- how I can change negative or self-critical thinking.
- how I can do more activities that are important to me or that I have enjoyed in the past.
- tracking my mood.
- learning more about my depression.
- improving my social skills.
- reminding me to take my medication.
- connecting with an expert (eg, psychotherapist) if I feel very bad.
- connecting with other people who have similar problems.
- how I can use my breath to make me feel better.
- writing a journal with things that concern me or for which I am grateful for.
- writing about events from my life.

Study Interview

We conducted semistructured interviews to investigate the problems participants faced due to depression, their adaptive coping strategies, their attitudes and expectations toward chatbots for depression, and their chatbot design preferences. Table 1 provides an excerpt of the interview guide. We asked several questions about each topic and prepared subquestions to follow up on specific details or offer suggestions based on the participants’ initial responses.

Table . Excerpt from the interview guide (translated from German).

Topic	Question
Problems	• Think of a period where you felt down. How did it look like?
Adaptive coping strategies	• What have you tried in the past to feel better?
Attitudes and expectations towards therapy chatbots	• Do you know what a chatbot is? • What do you think it would be like for you to use a chatbot that is there to help you with one of the problems you have mentioned? For example, what could be the advantages and disadvantages?
Design preferences for therapy chatbots	• Imagine you are thinking about using one of these chatbots. How should it need to be designed, so that you would download and use it? • How do you imagine an ideal conversation with the chatbot? For example, what topics would you like to discuss or what its personality should be like? • If it was accessible as a mobile app, what should the app look like? • (What) Would you like to personalize? • How often would you like to use such a chatbot (per week)? • How long should every session last?

XSL•FO
RenderX

Chatbot Prototype: Cady

Cady is the prototype of a chatbot for the treatment of depressive symptoms in youth that we developed for this and subsequent studies. Cady guides participants through a behavioral activation exercise, which is a core component of CBT to treat depression in youth [3]. Behavioral activation aims to encourage patients to engage in pleasant activities to overcome positive reinforcement deficits [39]. The conversation comprised the following sections: (1) introduction; (2) mood check on a scale from 1 (lowest) to 5 (highest) with adaptive empathetic responses; (3) psychoeducation on the relationship between behavior, thoughts, and feelings; (4) finding pleasant activities; (5) planning pleasant activities; (6) advice on how to overcome barriers when performing activities; and (7) feedback and goodbye. The conversation was designed by licensed psychologists based on established CBT manuals for youth [40,41] with a focus on positive, activating, and age-adequate language style. Cady also used emojis in its messages and sent GIFs. Figure 1 represents a screenshot of an example conversation between Cady and a participant. We named the chatbot Cady in accordance with the title of our research project and did not specify its age, gender, or other demographic characteristics to prevent specific demographic characteristics from influencing the results. We developed Cady using the prototyping software, Botsociety [42]. The chatbot was built with a rule-based architecture, where each user input triggers a predetermined response pathway following a decision tree structure. To ensure appropriate responses and maintain therapeutic quality, users primarily interacted with the chatbot by selecting from predefined options presented as buttons. We allowed free-text input in specific situations, for example, to identify and schedule pleasant activities. However, these free-text responses did not alter the chatbot's conversation pathway.

XSL•FO
RenderX

Figure 1. Screenshot of an example conversation between Cady and a participant.



Think-Aloud

We applied the concurrent think-aloud method [43], in which participants are asked to express their thoughts and feelings loudly while interacting with an app. At the beginning of the think-aloud session, the interviewer introduced the chatbot prototype Cady to the participants and explained the think-aloud method. Subsequently, participants interacted with Cady using a laptop. While the participants interacted with the prototype, we recorded the laptop screen and their comments. The interviewer stayed in the room because all participants felt it was easier to share their thoughts and feelings when the interviewer was in the room. The interviewer asked for further explanation if the statements were unclear or nonspecific.

Data Analysis

The interview and think-aloud session were recorded using an audio recorder (Zoom Q2n [Zoom Corporation]) and were

XSL•FO
RenderX

transcribed verbatim. Furthermore, 2 coders analyzed the semistructured interviews and concurrent think-aloud sessions using qualitative content analyses by Mayring and Fenzl [44] and Mayring [45] with the QCAmap software (Fenzl and Mayring) [46]. Similar to Zehetmair et al [47], we chose an inductive coding approach to achieve the most unbiased and thorough description of the data, which we deemed important for the exploratory nature of our study [44,45]. Both coders (Lilli Feline Metelmann and LB) coded the material independently, discussed their coding results, and jointly developed a category system. When disagreements between the coders could not be resolved, a third coder (FOK) was consulted. We chose this process because intercoder comparison is not feasible with inductive category development [44,45]. In an inductive approach, categories are created bottom-up from the material rather than selected before data analysis, as in deductive approaches [44,45]. We report the frequencies of the categories and codes to increase transparency and demonstrate the saliency of the codes in the data. The category system is provided in Multimedia Appendix 2. The questionnaires were analyzed using R (version 4.1.3; R Core Team). For PHQ-8, SCARED, and the potential capabilities questionnaire, we calculated descriptive statistics (mean and SD). For the potential capabilities, we also generated frequency distributions to visualize the response patterns.

Ethical Considerations

The Institutional Review Board of the Medical Faculty of the University of Tübingen approved the study (project ID 595/2021B01). The participants and caregivers provided written informed consent (Multimedia Appendix 3) before screening for eligibility and enrollment. All participant data were deidentified before analysis by removing personal identifiers such as names, addresses, and contact information, and replacing them with unique study codes to ensure participant confidentiality. After the study, participants received a reimbursement of 30€ (equivalent to US $33) for taking part in the study. The study lasted between 1 and 2 hours.

Results

Participant Characteristics

In total, 14 youth between 14 and 17 years old (mean 16.1, SD 1.14 years) participated in the study. Furthermore, 12 participants identified as women and 2 as nonbinary. All participants reported a current or remitted depressive disorder. In addition, 10 participants had a PHQ-8 score of 10 or higher, indicating acute depressive symptoms [48]. Of the total, 12 participants had a SCARED score of 25 or higher, indicating current symptoms of anxiety [49]. Furthermore, 9 participants showed symptoms of both depression and anxiety. Out of 14, 9 (64%) participants were currently receiving psychotherapy, whereas 8 (57%) had received psychotherapy in the past. A comprehensive overview of the participant characteristics is illustrated in Table 2.

XSL•FO
RenderX

Table . Participant characteristics.

Characteristic	Statistical values
Age (y), mean (SD)	16.1 (1.14)
Gender, n (%)	
Women	12 (86)
Nonbinary	2 (14)
PHQ-8[a] (sum score)	
Mean (SD)	13.57 (5.58)
≥Cut-off (10), n (%)	10 (71)
SCARED[b] (sum score)	
Mean (SD)	40.86 (17.55)
≥Cut-off (25), n (%)	12 (86)
Psychotherapy experience, n (%)	
No experience	2 (14)
Current (≤6 mo)	9 (64)
Remitted (>6 mo)	8 (57)
Both current and past	6 (43)
Diagnosis of depression, n (%)	
Current (≤6 mo)	8 (57)
Remitted (>6 mo)	10 (71)
Recurrent (minimum 2 episodes)	5 (36)

[a]PHQ-8: 8-item Patient Health Questionnaire.

[b]SCARED: Screen for Child Anxiety Related Disorders.

Research Question 1: Problems With Depression

All participants (N=14, 100%) experienced problems categorized as depressive symptoms, including lack of motivation and energy, depressed mood, and self-doubt. Furthermore, 2 participants explained their main problems:

> *I just don't have the strength to do anything" and "I'm just tired of everyday life, or of all these everyday activities like brushing my teeth.*

Some participants (9/14, 64%) reported comorbidities, such as panic attacks, anxiety, and excessive alcohol consumption, while 4 participants (28%) reported physical problems, such as tension, headaches, or stomach pain. All 14 participants (100%) also reported interpersonal problems, particularly withdrawal from social relationships, stress caused by parents, and problems with friends. Furthermore, 11 participants (78%) expressed concerns about school or their future. They shared negative experiences at school and experienced pressure to plan their future, which were reflected by statements like:

> *When it comes to school, I have a very, very big fear of the future.*
>
> *Success is a big topic right now, because of the end of school year grades.*

In total, 6 participants (42%) reported experiencing stigma associated with depression. They mentioned that adults stigmatized or trivialized their problems or feared negative reactions when sharing their problems. One participant was concerned that disclosing that they are seeing a psychotherapist would lead to others perceiving them as "sick in the head." Most participants reported encountering difficulties with mental health care. Furthermore, 12 individuals (85%) reported experiencing treatment barriers, including attitudinal and structural barriers. Attitudinal barriers included fear of being judged, self-assessment that problems are not severe enough to deserve support, and negative experiences during psychotherapy (n=9, 64%). Participants reported ineffective techniques, difficulty with disclosing personal information, pressure to perform, trust violations, and condescending treatment from therapists. Structural barriers reported are long waiting lists or being deemed not sufficiently severe by healthcare providers or parents to qualify for professional support.

Research Question 2: Adaptive Coping Strategies

All 14 participants (100%) identified social support as an adaptive coping strategy. The majority received social support from friends, family members, partners, teachers, or people they interacted with online. In total, 13 participants (93%) engaged in activities to distract themselves or to have a positive experience such as spending time outdoors or consuming media. A total of 9 participants (64%) reported receiving professional treatment and 9 (64%) used cognitive strategies, including positive self-talk, self-reflection, and reducing rumination. Furthermore, 6 participants (42%) implemented a daily structure

XSL•FO
RenderX

including forming habits, getting up early, scheduling positive activities, and setting goals. In addition, 7 (50%) prioritized their needs as a coping mechanism, such as intentionally allocating time to self-care or avoiding stressful social situations. Of the total, 3 individuals (21%) engaged in mindfulness practice, such as breathing techniques and meditation, and 2 individuals (14%) sought online information regarding their depressive symptoms, interrelated problems and further information.

Research Question 3: Attitudes and Expectations Toward Chatbots to Treat Depression

Most participants had positive attitudes toward and expectations of chatbots to treat depression. In total, 12 participants (85%) stated that they would be less anxious about using a chatbot than seeing a human therapist. They pointed out that using a chatbot would be a suitable option for discussing sensitive topics that they would not share with others, primarily because they would not fear negative reactions. In addition, texting was considered less intimidating than speaking with a therapist. Furthermore, 11 participants (79%) pointed out the unlimited capacity and flexibility of chatbots:

> *I think it's also an advantage that you can really chat with it at any time, because in therapy you just have one appointment per week. If you feel bad in the evening or at night or something, then you can still text the chatbot.*

In addition, participants indicated that a chatbot is more accessible and requires less effort than seeing a human therapist. A total of 8 participants (57%) expressed confidence in the effectiveness of chatbots for the treatment of depression. They indicated that such a chatbot would be capable of addressing a wide range of issues, and welcomed the idea of having a helpful everyday chatbot. All participants demonstrated keen interest in using a chatbot to treat depression either because they expected it to be effective or because they were curious about using it. One participant stated that a daily usable chatbot could alleviate feelings of loneliness. Several others noted that a chatbot's personal and human-like nature would increase the motivation to use it, particularly compared with other less interactive DMHIs.

However, participants also reported several concerns regarding the use of chatbots to treat depression. Some participants (n=9, 64%) were skeptical about the chatbot's intelligence and natural language capabilities. They expressed concern that the chatbot would not be able to address individual, diverse, or unusual problems effectively, and feared being disappointed if the chatbot was unable to do so. Participants were particularly worried about inappropriate answers to emotional and intimate topics or inappropriate advice for their problems. Some participants (n=9, 64%) were concerned that a conversation with the chatbot would not feel like a conversation with a human therapist, due to a potential lack of emotional intelligence or overly robotic or analytical responses. However, chatbots that appeared too human were believed to be uncanny. In total, 2 participants (14%) were skeptical about sufficient data security and privacy. Finally, 2 participants (14%) worried that their symptoms of depression, such as little motivation or forgetfulness, would result in low engagement with the chatbot, thereby preventing them from effectively using it. They also pointed out that the lack of social pressure when using a chatbot, compared to seeing a human therapist, could contribute to low engagement, which might not be solvable by the chatbot.

Research Question 4: Chatbot Design Preferences

Category 1: Dialogue Topics and Therapeutic Content

All participants (n=14, 100%) shared preferences regarding dialogue topics that the chatbot should be able to cover. These suggestions include comprehensive assessments of depressive symptoms and interrelated problems, psychoeducation, therapeutic exercises that address specific issues, reminders to address basic needs, support for regulating emotions, tackling intrusive thoughts, and being distracted when needed. Furthermore, they emphasized the importance of discussing their daily lives and, more specifically, sharing current problems and receiving suggestions on how to resolve them. One participant highlighted the importance of the chatbot explicitly asking the user what type of support they require, such as emotional support (ie, listening and validation) or solution-oriented support:

> *Getting advice on how to solve a problem isn't always helpful, even if it's well-intentioned. […] It would be important for the chatbot to ask or understand if you want advice or if you only want to share your feelings.*

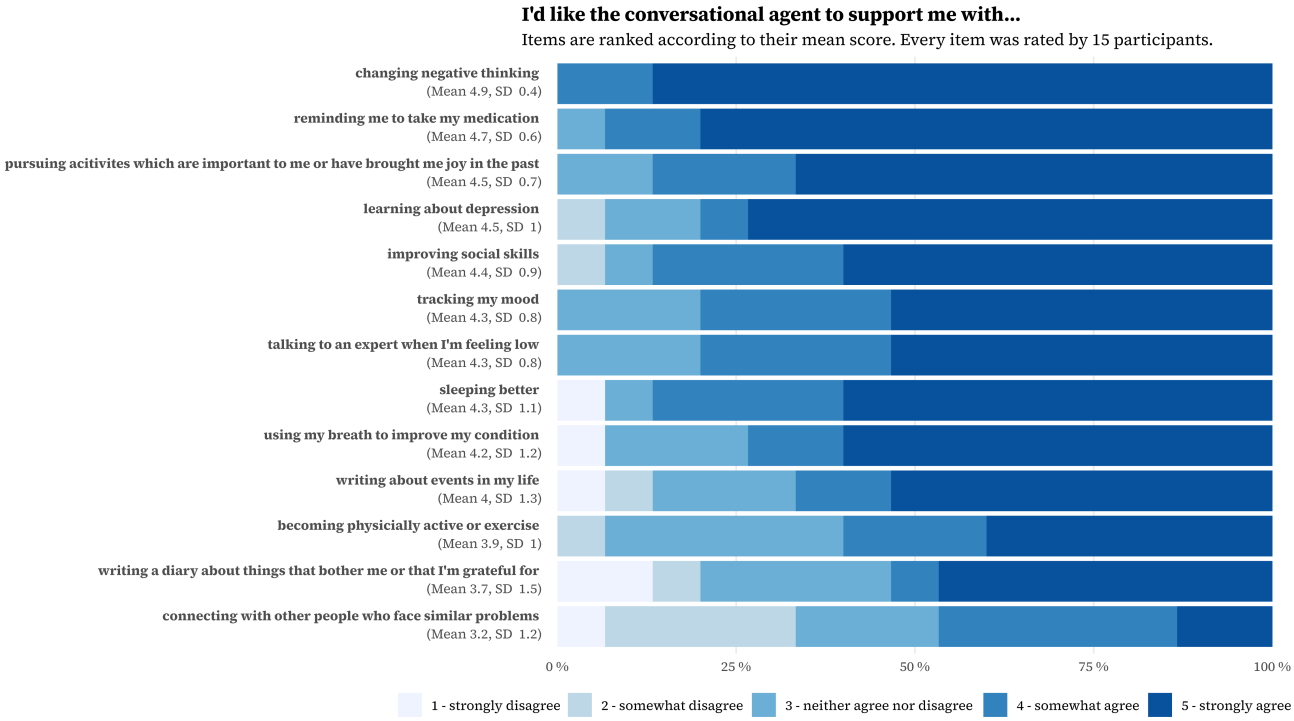
Others elaborated on how they imagined talking about daily life and receiving advice:

> *I would just talk about the things that depress me at the time, to which I don't know the answer.*
>
> *I would probably just chat about everyday situations that were unpleasant to me or something like that.*

These preferences were complemented by responses to a questionnaire regarding the chatbot's desired capabilities. Figure 2 presents the full results. The responses indicate the significance of three key components of CBT: (1) cognitive restructuring ("changing negative thinking;" mean 4.9, SD 0.4), (2) behavioral activation ("pursuing activities that are important to me or have brought me joy in the past;" mean 4.5, SD 0.7), and (3) psychoeducation ("learning about depression;" mean 4.5, SD 1). In addition, participants expressed a preference for "improving social skills" (mean 4.4, SD 0.9). On the other hand, therapeutic writing, represented by "writing about events in my life" (mean 4, SD 1.3) and "writing a diary about things that bother me or that I'm grateful for" (mean 3.7, SD 1.5), was rated lower than most items. Finally, "connecting with other people who face similar problems" was ranked the lowest (mean 3.2, SD 1.2).

XSL•FO
RenderX

Figure 2. Results for how much support participants would like for several proposed chatbot capabilities.



Category 2: Personality, Interaction Style, and Social Role

Participants made various suggestions for the chatbot's personality (n=14, 100%), such as understanding, caring, friendly, emphatic, encouraging, humorous, interested, neutral, and nonjudgmental. During the think-aloud sessions, most participants characterized Cady's personality as one of its main strengths. One participant said:

> *I like that you communicate in a friendly way, like an internet friend.*

Furthermore, 2 participants were pleased that Cady was interested and asked them personal questions. For example, 1 participant approved that Cady had asked if she had experienced becoming less active herself:

> *I just think it's good that she is asking if I know something like that or haven't experienced it yet.*

In terms of the interaction style and language used by the chatbot, participants suggested that it should be appropriate for the user's age without complicated terms. This issue was also raised during think-aloud sessions. One participant asked for more age-specific language, such as short sentences:

> *So, this was a therapeutic chat, but if it is supposed to be like a friend, I think it is better if it is writes like people my age, eg using short sentences.*

Another participant perceived the language as too therapeutic:

> *It was relatively authentic, but you also notice that it is not written for my age.*

The human-likeness of the chatbot was also a significant factor, as participants found it important to react to feelings and to understand irony and humor. However, participants also expressed concerns about the chatbot being too human-like, which could lead to irritation or fear. During the think-aloud sessions, 1 participant appreciated that Cady disclosed personal information such as enjoying chats with nice people, which resembled a human-like trait.

Regarding the social role, the participants (n=14, 100%) held divergent views: some desired the chatbot to resemble a friend, whereas others preferred it to assume the role of a therapist or a combination of both. One participant explained their preference for a therapist-like role:

> *I actually think it's better if it were like a therapist [...], because with friends you don't necessarily want talk about everything.*

One participant proposed using situation-specific roles:

> *A therapist if you're doing therapeutic exercises and a friend if you just want to talk.*

Category 3: Personalization

Personalization frequently discussed the importance of personalization (n=13, 92%). Participants reported various aspects that should be personalized, including the chatbot's personality, gender, avatar, message style, dialogue topics, therapeutic content, the mobile app's color theme, notifications, and user profile picture and username. However, the participants had divergent opinions on who should control the personalization of the chatbot. While some suggested that the user should be in control, others believed that the chatbot itself should control personalization. Others, in turn, proposed a mixed approach in which both the user and the chatbot share control. Some participants preferred personalization to occur only once at the beginning, whereas others preferred dynamic personalization, where personalization occurs continuously. In addition to these explicit preferences, personalization also emerged from diverging design preferences. All participants expressed a preference for an appealing user interface, but they

XSL•FO
RenderX

had differing opinions on the specific design elements, with some suggesting a bright and colorful design, whereas others preferred a plain and minimalistic design with black and white colors. Most participants preferred the chatbot as a standalone app, but some preferred integration into a popular instant messaging app such as WhatsApp (Meta). Regarding the interaction modalities, most participants preferred a chatbot over a voice agent. In terms of text input, the participants believed that a mix of predefined answers and free-text input was optimal, as it combines the minimal effort of clicking predefined answers with the flexibility of unrestricted text input. In total, 2 participants (14%) expressed concerns about personalization, stating that personalized content could lead to avoidance behavior, in which users avoid topics that they find difficult or uncomfortable. They also believed that personalization of the chatbot's appearance, such as its gender or avatar, could reduce its seriousness as a tool for treating depression.

Table 3 summarizes key chatbot design recommendations derived from results across all 4 research questions.

Table . Key design recommendations.

Topic	Design recommendation
Assessment	• Incorporate repeated multidimensional assessment, covering depressive symptoms and related problems, to monitor and understand the key issues a user faces. • Ensure assessments are comprehensive yet concise to minimize the effort for users.
Scope and Limitations	• Communicate the chatbot's scope and limitations clearly, for instance, when it cannot properly handle a topic. Redirect users to external resources (eg, helplines and professional interventions) in such cases to ensure users receive appropriate support.
Dialogue content	• Curate a diverse content database covering a wide range of topics that are relevant to youth with depression. • Integrate evidence-based therapeutic techniques alongside user-preferred coping strategies (eg, behavioral activation, cognitive restructuring, and social support). • Enable the chatbot to engage in conversations about users' daily life, including discussions on urgent problems. • Support follow-up discussions on user-reported problems, track progress, and encourage reflection on how successful they carried out the learned techniques.
Personality, social role and language style	• Design a chatbot with an understanding, caring, friendly, empathic, encouraging, humorous, interested, and nonjudgmental personality. • Ensure the chatbot uses age-appropriate language without complicated terminology. • Balance human-likeness (eg, recognizing feelings, irony, humor) with clear boundaries to avoid unrealistic or overly human-like interactions.
Personalization	• Personalize the content selection and presentation. • Personalize the chatbot's persona (eg, personality, social role, gender, avatar, message style). • Explore the impact of persona personalization (eg, switching between formal and casual depending on user feedback, interaction style preferences and situation). • Explore the impact of who controls personalization (the user, the chatbot or both) and when or how often personalization is performed (initial vs ongoing). • Ensure personalization enhances user experience while maintaining the chatbot's therapeutic credibility and seriousness.

Discussion

Primary Findings

This study explored the problems youth face with depression, their adaptive coping strategies, and their attitudes and design preferences for chatbots to treat depression. Our findings indicate that youth experience diverse problems beyond core depressive symptoms, employ a range of coping strategies, hold predominantly positive attitudes toward chatbots for depression, and express diverse, sometimes contradictory, design preferences. In the following section, we discuss the results according to the four research questions guiding this study: (1) the problems youth with depression face, (2) their adaptive coping strategies, (3) their attitudes and expectations toward chatbots to treat depression, and (4) their design preferences for

XSL•FO
RenderX

chatbots to treat depression. As presented in Table 3, our findings across all research questions offer specific guidance for designing chatbots tailored specifically to youth experiencing depression.

Problems of Youth With Depression

First, participants reported various problems with depression beyond the symptoms described in the DSM-5 [50], including mental health and somatic comorbidities, interpersonal issues, and concerns about school and the future. These findings align with previous research showing that youth with depression face an increased risk for anxiety disorders, substance abuse, and physical health issues [51], as well as social withdrawal, lack of friendship [52], diminished academic achievement, school dislike, and pessimism [52]. To address these problems, engaging and effective chatbots should (1) assess both core depressive symptoms and related problems repeatedly to inform therapeutic decisions and strengthen the therapeutic relationship by understanding the users' individual challenges, and (2) curate a diverse content database to address the wide range of problems, aligning with Li et al [53]. Second, participants reported significant barriers to receiving professional treatment, including poor mental health literacy (eg, awareness of symptom severity), attitudinal (eg, perceived stigma), and structural barriers (eg, waiting times). These findings align with research showing that youth struggle to recognize symptoms [54], hesitate to disclose personal information to professionals due to fear of judgement or not being taken seriously [7], and question the effectiveness of professional treatment [7], a concern supported by evidence that 60% of youth with depression do not benefit from psychological treatments [55]. To address these barriers, chatbots should (1) encourage emotional disclosure by providing a nonjudgmental conversational environment, something DMHIs cannot fully achieve due to their static, nonconversational design, and (2) clearly communicate their scope and limitations to set realistic user expectations and direct to external support when needed.

Adaptive Coping Strategies

Participants reported various adaptive coping strategies, including seeking social support, establishing positive activities, using cognitive strategies, and receiving professional treatment. Many also emphasized the importance of focusing on their needs and establishing a daily structure. Most of these strategies fall into the CBT strategies behavioral activation, cognitive restructuring, and problem solving [56], which are effective and first line treatments for depression in youth [57].

The prevalence of CBT-based coping strategies may be due to the previous psychotherapy experience of most participants. Similarly, among youth with nonsuicidal self-injury, those with experience in dialectical behavioral therapy have suggested incorporating it into a DMHI designed for them [58], showing how previous therapy experience influences preferences for DMHIs. The use of CBT-aligned adaptive coping strategies is important, as research shows that youth with depression use adaptive cognitive strategies less frequently than healthy controls [59] and that applying these strategies is associated with fewer depressive symptoms [60]. Chatbots for youth with depression should incorporate these strategies to create engaging and effective content that aligns with users' existing coping strategies and effective therapeutic techniques. For example, the chatbot could facilitate social support by helping users identify supportive individuals and suggesting personalized ways to reach out, such as a video call with a friend who is good at cheering them up or a message to a family member who provides good advice. Chatbots can even draft customizable messages to support users who struggle to reach out.

Attitudes and Expectations Toward Chatbots to Treat Depression

Participants predominantly held positive attitudes toward chatbots to treat depression. Many participants reported feeling less anxious about using a chatbot than seeing a human psychotherapist, especially when discussing sensitive topics. This finding extends research on post-traumatic stress disorder, showing that digital agents lead to greater disclosure of sensitive or stigmatized information [61,62], and supports evidence that youth value the privacy and anonymity offered by DMHIs [63]. Given these findings, chatbots could help users become comfortable with sharing sensitive information and practicing therapeutic conversations, with potential benefits for subsequent sessions with psychotherapists. However, some participants were skeptical about the chatbots' ability to address individual problems and provide appropriate advice on sensitive topics, which is supported by evidence that commercial chatbots for adults with depression often fail to match user inputs, understand messages, or respond appropriately [64]. Although privacy and data security are frequently cited as primary risks of DMHIs for young people [65], few participants raised these concerns. Instead, they viewed chatbots as privacy-enhancing compared with human therapists. Nevertheless, robust privacy and data security remain essential, and it is unclear whether participants were genuinely unconcerned or simply assumed chatbots would have strong protections in place.

Design Preferences

Our study revealed important insights into design preferences for chatbots for youth with depression. First, participants expressed diverse preferences for dialogue topics, with three areas standing out: (1) chatting about daily life, (2) discussing urgent problems and receiving advice, and (3) working through therapy exercises that address specific problems. Interestingly, while youth with nonsuicidal self-injury [58], emotional problems [66], and from the general population [63] included being connected to others facing similar challenges as a key feature of their DMHI, our participants ranked such social connection as their lowest priority, highlighting how design preferences may vary across mental health conditions. In summary, a chatbot to treat depression in youth should reflect these preferences. However, available applications have not implemented chats about daily life, and relied on rule-based approaches for problem-solving [14,15], likely due to insufficient conversational capabilities. Structured therapy techniques, such as behavioral activation or cognitive restructuring, are standard components of DMHIs [38] and chatbots [67]. However, their current implementation remains predominantly static [68], without personalized advice or feedback during therapeutic exercises. Large language models

XSL•FO
RenderX

(LLMs) such as OpenAI's GPT or Anthropic's Claude promise to overcome current limitations of DMHIs and chatbots. LLMs enable chatbots to better address our participants' design preferences, such as natural conversations about daily life and advice on urgent problems. In addition, LLMs can enhance static implementations of therapeutic exercises by offering personalized guidance and feedback. In behavioral activation, LLMs can personalize the explanation of the relationship between behavior and feelings and evaluate the proposed activity plan for feasibility and therapeutic appropriateness. In cognitive restructuring, Sharma et al [69] demonstrated that LLMs can assist users in identifying thinking traps and generating reframed thoughts. Although these enhanced capabilities have the potential to improve understanding and skill development, key challenges remain. LLMs can generate false or harmful messages [70], posing risks to vulnerable users. Future research needs to explore how capitalize on their advanced conversational abilities while ensuring therapeutic quality and safety [71].

Second, our findings show that engaging and effective chatbots for youth with depression need personalization beyond the conversational capabilities of LLMs. Drawing on the framework by Cohen et al [72] for psychotherapy personalization, effective chatbots require personalization in content selection and interaction style. Content selection ensures that chatbots address the diverse problems and dialogue topics our participants reported, aligning with findings from Li et al [53] and Ludlow et al [66]. A personalized interaction style accommodates the preferences for different social roles and language use. While the benefits of personalizing the chatbot persona have been shown [73,74]. LLMs enable the personalization of interaction styles via prompt instructions, reducing the need to craft different responses for each style manually [75]. Despite the clear need for personalization, participants disagreed on who should control it [76], preventing a clear design recommendation. While some preferred user-led personalization, aligning with Kenny et al [63], others favored chatbot-led personalization or a hybrid approach. Given this ambiguity and the lack of empirical evidence [77], future research is needed to guide chatbot personalization.

Limitations

Our study has 2 main limitations. First, our sample size was limited, and the participants were predominantly women. Although depression is more prevalent in women [78-80], the absence of men limits the generalizability of our findings across genders. The gender imbalance resulted from convenience sampling through a single youth psychiatrist. One man was informed about the study but declined to participate. While our study included 2 nonbinary participants, an underrepresented demographic in research, recruiting men would have required targeted efforts beyond our study's constraints. In addition, our sample consisted entirely of participants who had actively sought mental health support, most of whom had previous psychotherapy experience. As a result, perspectives from those who avoid professional help or resist treatment may be underrepresented. Due to these sampling limitations, we likely did not achieve full data saturation. Although qualitative research can reach saturation with as few as twelve interviews [31], our sample composition suggests that some perspectives may have been missed. However, our findings on youth depression and related problems replicated findings from studies with much larger sample sizes [81], indicating that our results are comprehensive. Future studies should recruit a more diverse sample, specifically including men, individuals reluctant to seek professional treatment, and those without previous psychotherapy experience, to validate and extend our findings.

Second, our study captured attitudes, preferences, and hypothetical usage scenarios rather than actual usage behavior. While participants interacted with a prototype during the think-aloud sessions, this brief, controlled interaction may not reflect real-world interactions. The gap between stated preferences and actual behavior is well-documented [82], and our study design does not allow us to determine whether the design preferences of the participants would translate into long-term engagement or effective use. Future research should follow the next steps of an iterative human-centered design process [83,84] with functional prototypes to assess whether the identified design preferences lead to actual user engagement and therapeutic benefits. Despite these limitations, our study provides in-depth findings on the design of chatbots to treat depression in youth, highlighting the value of qualitative research in the iterative development of DMHIs [83,84].

Conclusions and Future Work

Our study provides valuable insights into the problems and coping strategies of youth with depression, and their attitudes, expectations, and design preferences for a chatbot to treat depression in youth. We found complex user needs, predominantly positive attitudes toward chatbots, and various design preferences, including the need for diverse dialogue topics and personalization. Our findings led to concrete design recommendations that lay a crucial foundation for developing engaging and effective chatbots to treat depression in youth.

Despite these contributions, several research directions remain. First, future studies should validate and extend these findings with larger, more diverse samples to ensure broader representation of youth with depression. Second, examining actual usage patterns and long-term engagement with functional chatbot prototypes will help assess the effectiveness of our design recommendations. Third, investigating the responsible integration of LLMs is important, including the development of robust safeguards and evaluating dialogue quality and therapeutic outcomes. Finally, researchers must design and evaluate effective personalization features, particularly regarding whether users, the chatbot or both should control personalization, while balancing implementation efforts and impact. By addressing these research directions, we can further improve chatbots to treat depression in youth and ultimately contribute to more accessible, engaging, and effective mental health support for this vulnerable population.

XSL•FO
RenderX

Acknowledgments

We thank (1) all the participants for their participation in this study, (2) Ulrike Detzner, Reiter, MD for supporting the recruitment, and (3) Lilli Feline Metelmann for her efforts in planning, conducting, and analyzing the study with us. This work was supported by a donation of Cents for help e.V. to SL (project title: “An intelligent therapy chatbot for children suffering from depression and suicidal ideation”). We also acknowledge financial support from the Open Access Publication Fund of the University of Tübingen.

Conflicts of Interest

SL has received public funding, along with other scientists, for on an app-based aftercare intervention for youth and young adults with depression (iCAN study). He has received consultancy fees from companies for advice on study and intervention design in the context of e-mental health. He has also received payments for lectures on e–mental health. None of the other authors declare any conflict of interest.

Multimedia Appendix 1
Interview Guide.
[DOCX File, 62 KB - humanfactors_v12i1e66632_app1.docx]

Multimedia Appendix 2
Categories.
[DOCX File, 66 KB - humanfactors_v12i1e66632_app2.docx]

Multimedia Appendix 3
Informed written consent.
[PDF File, 351 KB - humanfactors_v12i1e66632_app3.pdf]

References

1. Clayborne ZM, Varin M, Colman I. Systematic review and meta-analysis: adolescent depression and long-term psychosocial outcomes. J Am Acad Child Adolesc Psychiatry 2019 Jan;58(1):72-79. [doi: 10.1016/j.jaac.2018.07.896] [Medline: 30577941]
2. Thapar A, Collishaw S, Pine DS, Thapar AK. Depression in adolescence. Lancet 2012 Mar 17;379(9820):1056-1067. [doi: 10.1016/S0140-6736(11)60871-4] [Medline: 22305766]
3. Oud M, de Winter L, Vermeulen-Smit E, et al. Effectiveness of CBT for children and adolescents with depression: a systematic review and meta-regression analysis. Eur Psychiatry 2019 Apr;57:33-45. [doi: 10.1016/j.eurpsy.2018.12.008] [Medline: 30658278]
4. Eisenberg D, Golberstein E, Gollust SE. Help-seeking and access to mental health care in a university student population. Med Care 2007 Jul;45(7):594-601. [doi: 10.1097/MLR.0b013e31803bb4c1] [Medline: 17571007]
5. Ebert DD, Mortier P, Kaehlke F, et al. Barriers of mental health treatment utilization among first-year college students: first cross-national results from the WHO World Mental Health International College student initiative. Int J Methods Psychiatr Res 2019 Jun;28(2):e1782. [doi: 10.1002/mpr.1782] [Medline: 31069905]
6. Gulliver A, Griffiths KM, Christensen H. Perceived barriers and facilitators to mental health help-seeking in young people: a systematic review. BMC Psychiatry 2010 Dec 30;10(1):1-9. [doi: 10.1186/1471-244X-10-113] [Medline: 21192795]
7. Radez J, Reardon T, Creswell C, Lawrence PJ, Evdoka-Burton G, Waite P. Why do children and adolescents (not) seek and access professional help for their mental health problems? A systematic review of quantitative and qualitative studies. Eur Child Adolesc Psychiatry 2021 Feb;30(2):183-211. [doi: 10.1007/s00787-019-01469-4] [Medline: 31965309]
8. Domhan D, In-Albon T, Pfeiffer S. Erfassung von Barrieren und Faszilitatoren zur Aufnahme einer Psychotherapie im Kontext ambulanter Kinder- und Jugendlichenpsychotherapie. Psychotherapie 2023 Nov;68(6):466-474. [doi: 10.1007/s00278-023-00679-1]
9. Wu Y, Fenfen E, Wang Y, et al. Efficacy of internet-based cognitive-behavioral therapy for depression in adolescents: a systematic review and meta-analysis. Internet Interv 2023 Dec;34:100673. [doi: 10.1016/j.invent.2023.100673] [Medline: 37822787]
10. Leech T, Dorstyn D, Taylor A, Li W. Mental health apps for adolescents and young adults: a systematic review of randomised controlled trials. Child Youth Serv Rev 2021 Aug;127:106073. [doi: 10.1016/j.childyouth.2021.106073]
11. Cameron SK, Rodgers J, Dagnan D. The relationship between the therapeutic alliance and clinical outcomes in cognitive behaviour therapy for adults with depression: a meta‐analytic review. Clin Psychology and Psychoth 2018 May;25(3):446-456 [FREE Full text] [doi: 10.1002/cpp.2180]

XSL•FO
RenderX

12. Vaidyam AN, Wisniewski H, Halamka JD, Kashavan MS, Torous JB. Chatbots and conversational agents in mental health: a review of the psychiatric landscape. Can J Psychiatry 2019 Jul;64(7):456-464. [doi: 10.1177/0706743719828977] [Medline: 30897957]
13. Bendig E, Erb B, Schulze-Thuesing L, Baumeister H. Die nächste generation: chatbots in der klinischen psychologie und psychotherapie zur förderung mentaler gesundheit – ein scoping-review. Verhaltenstherapie 2019 Dec 23;29(4):266-280. [doi: 10.1159/000499492]
14. Inkster B, Sarda S, Subramanian V. An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: real-world data evaluation mixed-methods study. JMIR Mhealth Uhealth 2018 Nov 23;6(11):e12106. [doi: 10.2196/12106] [Medline: 30470676]
15. Fitzpatrick KK, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. JMIR Ment Health 2017 Jun 6;4(2):e19. [doi: 10.2196/mental.7785] [Medline: 28588005]
16. Abd-Alrazaq AA, Rababeh A, Alajlani M, Bewick BM, Househ M. Effectiveness and safety of using chatbots to improve mental health: systematic review and meta-analysis. J Med Internet Res 2020 Jul 13;22(7):e16021. [doi: 10.2196/16021] [Medline: 32673216]
17. Perski O, Crane D, Beard E, Brown J. Does the addition of a supportive chatbot promote user engagement with a smoking cessation app? An experimental study. Digit Health 2019 Jan;5:1-13. [doi: 10.1177/2055207619880676]
18. Linardon J, Torous J, Firth J, Cuijpers P, Messer M, Fuller-Tyszkiewicz M. Current evidence on the efficacy of mental health smartphone apps for symptoms of depression and anxiety. A meta-analysis of 176 randomized controlled trials. World Psychiatry 2024 Feb;23(1):139-149. [doi: 10.1002/wps.21183] [Medline: 38214614]
19. Darcy A, Daniels J, Salinger D, Wicks P, Robinson A. Evidence of human-level bonds established with a digital conversational agent: cross-sectional, retrospective observational study. JMIR Form Res 2021 May 11;5(5):e27868. [doi: 10.2196/27868] [Medline: 33973854]
20. Bae Brandtzæg PB, Skjuve M, Kristoffer Dysthe KK, Følstad A. When the social becomes non-human: young people's perception of social support in chatbots. Presented at: CHI '21: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems; May 6, 2021; Yokohama Japan. [doi: 10.1145/3411764.3445318]
21. Beatty C, Malik T, Meheli S, Sinha C. Evaluating the therapeutic alliance with a free-text CBT conversational agent (Wysa): a mixed-methods study. Front Digit Health 2022;4:847991. [doi: 10.3389/fdgth.2022.847991] [Medline: 35480848]
22. Feine J, Gnewuch U, Morana S, Maedche A. A taxonomy of social cues for conversational agents. Int J Hum Comput Stud 2019 Dec;132:138-161. [doi: 10.1016/j.ijhcs.2019.07.009]
23. Vaidyam AN, Linggonegoro D, Torous J. Changes to the psychiatric chatbot landscape: a systematic review of conversational agents in serious mental illness: changements du paysage psychiatrique des chatbots: une revue systématique des agents conversationnels dans la maladie mentale sérieuse. Can J Psychiatry 2021 Apr;66(4):339-348. [doi: 10.1177/0706743720966429] [Medline: 33063526]
24. Cicchetti D, Toth SL. A Developmental Psychopathology Perspective on Adolescent Depression Handbook of Depression in Adolescents: Routledge/Taylor & Francis Group; 2009:3-32.
25. Rice F, Riglin L, Lomax T, et al. Adolescent and adult differences in major depression symptom profiles. J Affect Disord 2019 Jan 15;243:175-181. [doi: 10.1016/j.jad.2018.09.015] [Medline: 30243197]
26. Andone I, Błaszkiewicz K, Eibes M, Trendafilov B, Montag C, Markowetz A. How age and gender affect smartphone usage. Presented at: UbiComp '16: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct; Sep 12, 2016; Heidelberg Germany p. 9-12. [doi: 10.1145/2968219.2971451]
27. Huffman S. OMG! Mobile voice survey reveals teens love to talk. URL: https://blog.google/products/search/omg-mobile-voice-survey-reveals-teens/ [accessed 2014-10-14]
28. Agapie E, Chang K, Patrachari S, Neary M, Schueller SM. Understanding mental health apps for youth: focus group study with Latinx youth. JMIR Form Res 2022 Oct 18;6(10):e40726. [doi: 10.2196/40726] [Medline: 36256835]
29. Høiland CG, Følstad A, Karahasanovic A. Hi, can I help? Exploring how to design a mental health chatbot for youths. Human Technology 2020 Aug 31;16(2):139-169 [FREE Full text] [doi: 10.17011/ht/urn.202008245640]
30. Grové C. Co-developing a mental health and wellbeing chatbot with and for young people. Front Psychiatry 2020;11:606041. [doi: 10.3389/fpsyt.2020.606041] [Medline: 33597898]
31. Guest G, Bunce A, Johnson L. How many interviews are enough? An experiment with data saturation and variability. Field Methods US: Sage Publications 2006;18(1):59-82. [doi: 10.1177/1525822X05279903]
32. Association AP. Diagnostic and Statistical Manual of Mental Disorders (DSM-5®): American Psychiatric Pub; 2013.
33. Kaufman J, Birmaher B, Brent D, et al. Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. J Am Acad Child Adolesc Psychiatry 1997 Jul;36(7):980-988. [doi: 10.1097/00004583-199707000-00021] [Medline: 9204677]
34. Martin A, Rief W, Klaiberg A, Braehler E. Validity of the Brief Patient Health Questionnaire Mood Scale (PHQ-9) in the general population. Gen Hosp Psychiatry 2006;28(1):71-77. [doi: 10.1016/j.genhosppsych.2005.07.003] [Medline: 16377369]

XSL•FO
RenderX

35. Weitkamp K, Romer G, Rosenthal S, Wiegand-Grefe S, Daniels J. German Screen for Child Anxiety Related Emotional Disorders (SCARED): reliability, validity, and cross-informant agreement in a clinical sample. Child Adolesc Psychiatry Ment Health 2010 Jun 30;4(1):1-8. [doi: 10.1186/1753-2000-4-19] [Medline: 20591137]
36. Towery J. The Anti-Depressant Book: A Practical Guide for Teens and Young Adults to Overcome Depression and Stay Healthy: Jacob Towery; 2016.
37. Auerbach RP, Webb CA, Stewart JG. Cognitive Behavior Therapy for Depressed Adolescents: A Practical Guide to Management and Treatment: Routledge; 2016. [doi: 10.4324/9781315746302]
38. Huguet A, Rao S, McGrath PJ, et al. A systematic review of cognitive behavioral therapy and behavioral activation apps for depression. PLOS ONE 2016;11(5):e0154248. [doi: 10.1371/journal.pone.0154248] [Medline: 27135410]
39. Lejuez CW, Hopko DR, Hopko SD. A brief behavioral activation treatment for depression. Treatment manual. Behav Modif 2001 Apr;25(2):255-286. [doi: 10.1177/0145445501252005] [Medline: 11317637]
40. Abel U, Hautzinger M. Kognitive Verhaltenstherapie Bei Depressionen Im Kindes- Und Jugendalter: Springer-Verlag; 2013.
41. Groen G, Petermann F. Therapie Tools Depression Im Kindes-Und Jugendalter: Beltz; 2015.
42. Botsociety. Design, preview and prototype your next chatbot or voice assistant. URL: https://botsociety.io [accessed 2021-06-18]
43. Jaspers MWM, Steen T, van den Bos C, Geenen M. The think aloud method: a guide to user interface design. Int J Med Inform 2004 Nov;73(11-12):781-795. [doi: 10.1016/j.ijmedinf.2004.08.003] [Medline: 15491929]
44. Mayring P, Fenzl T. Qualitative inhaltsanalyse. In: Baur N, Blasius J, editors. Handbuch Methoden Der Empirischen Sozialforschung: Springer Fachmedien; 2019:633-648. [doi: 10.1007/978-3-658-21308-4_42]
45. Mayring P. Qualitative Inhaltsanalyse: Grundlagen Und Techniken 12, Überarbeitete Auflage: Beltz; 2015.
46. Fenzl T, Mayring P. QCAmap: eine interaktive webapplikation für qualitative inhaltsanalyse. ZPID (Leibniz Institute for Psychology Information) 2017. [doi: 10.23668/psycharchives.11259]
47. Zehetmair C, Nagy E, Leetz C, et al. Self-practice of stabilizing and guided imagery techniques for traumatized refugees via digital audio files: qualitative study. J Med Internet Res 2020 Sep 23;22(9):e17906. [doi: 10.2196/17906] [Medline: 32965229]
48. Kroenke K, Strine TW, Spitzer RL, Williams JBW, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. J Affect Disord 2009 Apr;114(1-3):163-173. [doi: 10.1016/j.jad.2008.06.026] [Medline: 18752852]
49. Caporino NE, Sakolsky D, Brodman DM, et al. Establishing clinical cutoffs for response and remission on the Screen for Child Anxiety Related Emotional Disorders (SCARED). J Am Acad Child Adolesc Psychiatry 2017 Aug;56(8):696-702. [doi: 10.1016/j.jaac.2017.05.018] [Medline: 28735699]
50. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. In: 5th: American Psychiatric Association; 2013.
51. Agnafors S, Norman Kjellström A, Torgerson J, Rusner M. Somatic comorbidity in children and adolescents with psychiatric disorders. Eur Child Adolesc Psychiatry 2019 Nov;28(11):1517-1525. [doi: 10.1007/s00787-019-01313-9] [Medline: 30895480]
52. Mullarkey MC, Marchetti I, Beevers CG. Using network analysis to identify central symptoms of adolescent depression. J Clin Child Adolesc Psychol 2019;48(4):656-668. [doi: 10.1080/15374416.2018.1437735] [Medline: 29533089]
53. Li SH, Achilles MR, Spanos S, Habak S, Werner-Seidler A, O'Dea B. A cognitive behavioural therapy smartphone app for adolescent depression and anxiety: co-design of ClearlyMe. tCBT 2022;15. [doi: 10.1017/S1754470X22000095]
54. Radez J, Reardon T, Creswell C, Orchard F, Waite P. Adolescents' perceived barriers and facilitators to seeking and accessing professional help for anxiety and depressive disorders: a qualitative interview study. Eur Child Adolesc Psychiatry 2022 Jun;31(6):891-907. [doi: 10.1007/s00787-020-01707-0] [Medline: 33502596]
55. Cuijpers P, Karyotaki E, Ciharova M, et al. The effects of psychological treatments of depression in children and adolescents on response, reliable change, and deterioration: a systematic review and meta-analysis. Eur Child Adolesc Psychiatry 2023 Jan;32(1):177-192. [doi: 10.1007/s00787-021-01884-6] [Medline: 34611729]
56. Wenzel A. Basic strategies of cognitive behavioral therapy. Psychiatr Clin North Am 2017 Dec;40(4):597-609. [doi: 10.1016/j.psc.2017.07.001] [Medline: 29080588]
57. Luxton R, Kyriakopoulos M. Depression in children and young people: identification and management NICE guidelines. Arch Dis Child Educ Pract Ed 2020;107. [doi: 10.1136/archdischild-2020-320020]
58. Čuš A, Edbrooke-Childs J, Ohmann S, Plener PL, Akkaya-Kalayci T. "Smartphone apps are cool, but do they help me?": a qualitative interview study of adolescents' perspectives on using smartphone interventions to manage nonsuicidal self-injury. Int J Environ Res Public Health 2021 Mar 22;18(6):3289. [doi: 10.3390/ijerph18063289] [Medline: 33810106]
59. Mihailescu I, Efrim-Budisteanu M, Andrei LE, et al. Cognitive coping strategies among inpatient adolescents with depression and psychiatric comorbidity. Children (Basel) 2023 Nov 29;10(12):1870. [doi: 10.3390/children10121870] [Medline: 38136072]
60. Schäfer J, Naumann E, Holmes EA, Tuschen-Caffier B, Samson AC. Emotion regulation strategies in depressive and anxiety symptoms in youth: a meta-analytic review. J Youth Adolescence 2017 Feb;46(2):261-276. [doi: 10.1007/s10964-016-0585-0]

XSL•FO
RenderX

61. Lucas GM, Rizzo A, Gratch J, et al. Reporting mental health symptoms: breaking down barriers to care with virtual human interviewers. Front Robot AI 2017;4. [doi: 10.3389/frobt.2017.00051]
62. Pickard MD, Roster CA, Chen Y. Revealing sensitive information in personal interviews: Is self-disclosure easier with humans or avatars and under what conditions? Comput Human Behav 2016 Dec;65:23-30. [doi: 10.1016/j.chb.2016.08.004]
63. Kenny R, Dooley B, Fitzgerald A. Developing mental health mobile apps: exploring adolescents' perspectives. Health Informatics J 2016 Jun;22(2):265-275. [doi: 10.1177/1460458214555041] [Medline: 25385165]
64. Martinengo L, Lum E, Car J. Evaluation of chatbot-delivered interventions for self-management of depression: content analysis. J Affect Disord 2022 Dec 15;319:598-607. [doi: 10.1016/j.jad.2022.09.028] [Medline: 36150405]
65. Wies B, Landers C, Ienca M. Digital mental health for young people: a scoping review of ethical promises and challenges. Front Digit Health 2021;3:697072. [doi: 10.3389/fdgth.2021.697072] [Medline: 34713173]
66. Ludlow K, Russell JK, Ryan B, et al. Co-designing a digital mental health platform, "Momentum", with young people aged 7-17: a qualitative study. Digit Health 2023;9:20552076231216410. [doi: 10.1177/20552076231216410] [Medline: 38033517]
67. Ahmed A, Hassan A, Aziz S, et al. Chatbot features for anxiety and depression: a scoping review. Health Informatics J 2023;29(1):14604582221146719. [doi: 10.1177/14604582221146719] [Medline: 36693014]
68. Denecke K, Schmid N, Nüssli S. Implementation of cognitive behavioral therapy in e-mental health apps: literature review. J Med Internet Res 2022 Mar 10;24(3):e27791. [doi: 10.2196/27791] [Medline: 35266875]
69. Sharma A, Rushton K, Lin IW, Nguyen T, Althoff T. Facilitating self-guided mental health interventions through human-language model interaction: a case study of cognitive restructuring. : Association for Computing Machinery; 2024 May 11 Presented at: CHI '24; May 11, 2024; Honolulu HI USA p. 1-29 URL: https://dl.acm.org/doi/proceedings/10.1145/3613904 [doi: 10.1145/3613904.3642761]
70. Birkun AA, Gautam A. Large language model (LLM)-powered chatbots fail to generate guideline-consistent content on resuscitation and may provide potentially harmful advice. Prehosp Disaster Med 2023 Dec;38(6):757-763. [doi: 10.1017/S1049023X23006568] [Medline: 37927093]
71. Stade EC, Stirman SW, Ungar LH, et al. Large language models could change the future of behavioral healthcare: a proposal for responsible development and evaluation. Npj Ment Health Res 2024 Apr 2;3(1):12. [doi: 10.1038/s44184-024-00056-z] [Medline: 38609507]
72. Cohen ZD, Delgadillo J, DeRubeis RJ. Personalized treatment approaches. In: Bergin and Garfield's Handbook of Psychotherapy and Behavior Change: John Wiley & Sons, Inc; 2021:673-703.
73. Nißen M, Rüegger D, Stieger M, et al. The effects of health care chatbot personas with different social roles on the client-chatbot bond and usage intentions: development of a design codebook and web-based study. J Med Internet Res 2022 Apr 27;24(4):e32630. [doi: 10.2196/32630] [Medline: 35475761]
74. Ahmad R, Siemon D, Gnewuch U, Robra-Bissantz S. Designing personality-adaptive conversational agents for mental health care. Inf Syst Front 2022;24(3):923-943. [doi: 10.1007/s10796-022-10254-9] [Medline: 35250365]
75. Jiang H, Zhang X, Cao X, Breazeal C, Roy D, Kabbara J. PersonaLLM: investigating the ability of large language models to express personality traits. : Association for Computational Linguistics Presented at: Findings of the Association for Computational Linguistics; Jun 16-21, 2024; Mexico City, Mexico p. 3605-3627 URL: https://aclanthology.org/2024.findings-naacl [doi: 10.18653/v1/2024.findings-naacl.229]
76. Fan H, Poole MS. What is personalization? Perspectives on the design and implementation of personalization in information systems. JOCEC 2006 Jan;16(3-4):179-202. [doi: 10.1080/10919392.2006.9681199]
77. Kocaballi AB, Berkovsky S, Quiroz JC, et al. The personalization of conversational agents in health care: systematic review. J Med Internet Res 2019 Nov 7;21(11):e15360. [doi: 10.2196/15360] [Medline: 31697237]
78. Avenevoli S, Swendsen J, He JP, Burstein M, Merikangas KR. Major depression in the national comorbidity survey-adolescent supplement: prevalence, correlates, and treatment. J Am Acad Child Adolesc Psychiatry 2015 Jan;54(1):37-44. [doi: 10.1016/j.jaac.2014.10.010] [Medline: 25524788]
79. Seedat S, Scott KM, Angermeyer MC, et al. Cross-national associations between gender and mental disorders in the World Health Organization World Mental Health Surveys. Arch Gen Psychiatry 2009 Jul;66(7):785-795. [doi: 10.1001/archgenpsychiatry.2009.36] [Medline: 19581570]
80. Hua Z, Wang S, Yuan X. Trends in age-standardized incidence rates of depression in adolescents aged 10-24 in 204 countries and regions from 1990 to 2019. J Affect Disord 2024 Apr 1;350:831-837. [doi: 10.1016/j.jad.2024.01.009] [Medline: 38242215]
81. Chevance A, Ravaud P, Tomlinson A, et al. Identifying outcomes for depression that matter to patients, informal caregivers, and health-care professionals: qualitative content analysis of a large international online survey. Lancet Psychiatry 2020 Aug;7(8):692-702. [doi: 10.1016/S2215-0366(20)30191-7] [Medline: 32711710]
82. de Corte K, Cairns J, Grieve R. Stated versus revealed preferences: an approach to reduce bias. Health Econ 2021 May;30(5):1095-1123. [doi: 10.1002/hec.4246] [Medline: 33690931]
83. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. J Med Internet Res 2015 Jan 30;17(1):e30. [doi: 10.2196/jmir.4055] [Medline: 25639757]

XSL•FO
RenderX

84. Harte R, Glynn L, Rodríguez-Molinero A, et al. A human-centered design methodology to enhance the usability, human factors, and user experience of connected health systems: a three-phase methodology. JMIR Hum Factors 2017 Mar 16;4(1):e8. [doi: 10.2196/humanfactors.5443] [Medline: 28302594]

Abbreviations

CBT: cognitive behavioral therapy
DMHI: digital mental health intervention
DSM-5: *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*
LLM: large language model
PHQ-8: 8-item Patient Health Questionnaire
SCARED: Screen for Child Anxiety Related Disorders

Edited by A Kushniruk; submitted 10.10.24; peer-reviewed by EK Onn Wong, MH Cheah; revised version received 31.03.25; accepted 31.03.25; published 19.06.25.

Please cite as:
Kuhlmeier FO, Bauch L, Gnewuch U, Lüttke S
Designing Chatbots to Treat Depression in Youth: Qualitative Study
JMIR Hum Factors 2025;12:e66632
URL: https://humanfactors.jmir.org/2025/1/e66632
doi:10.2196/66632

© Florian Onur Kuhlmeier, Luise Bauch, Ulrich Gnewuch, Stefan Lüttke. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 19.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Understanding Barriers to Home Safety Assessment Adoption in Older Adults: Qualitative Human-Centered Design Study

Jonathan J Lee[1,2], MD; Devika Patel[3,4], MS; Meghana Gadgil[3,5], MPH, MD; Simone Langness[4,6], MD; Christiana Dagmar von Hippel[3,4], MPH, SCD; Amanda Sammann[3,4], MPH, MD

[1]School of Medicine, Stanford University, Stanford, CA, United States

[2]Hasso Plattner Institute of Design, Stanford University, Stanford, CA, United States

[3]The Better Lab, University of California San Francisco, 2540 23rd St, 4th Floor, San Francisco, CA, United States

[4]Department of Surgery, University of California San Francisco, San Francisco, CA, United States

[5]Department of Medicine, University of California San Francisco, San Francisco, CA, United States

[6]Sharp HealthCare, San Diego, CA, United States

Corresponding Author:
Christiana Dagmar von Hippel, MPH, SCD
The Better Lab, University of California San Francisco, 2540 23rd St, 4th Floor, San Francisco, CA, United States

Abstract

Background: Falls are the leading cause of injury-related death among adults aged 65 and older. The fear of falling can further limit older adults' independence by contributing to activity restriction, social isolation, and physical decline—ironically increasing the risk of mechanical falls. Although home safety assessments have been shown to reduce fall risk by up to 36% and decrease serious injuries such as hip fractures, their adoption remains low. Understanding the barriers to implementing these assessments is critical to improving their uptake and effectiveness.

Objective: This study aimed to (1) identify specific barriers perceived by older adults in implementing home safety assessments and modifications to reduce the risk of mechanical falls, (2) explore the attitudes of health care professionals and other stakeholders toward these assessments, and (3) identify novel design opportunities to guide the development and implementation of more effective home safety assessment techniques and practices to reduce mechanical fall risk.

Methods: This explanatory qualitative study drew on the "inspiration" phase of the human-centered design (HCD) research process. We conducted 35 interviews (28 initial and 7 follow-up) with 28 purposefully sampled participants in the San Francisco Bay Area between February and June 2021. Participants included community-dwelling older adults (n=3), geriatricians (n=4), therapists (n=6), product developers (n=2), older health researchers (n=8), and community program leaders (n=5). Interview notes were analyzed inductively by the research team to extract themes and generate insight statements and design opportunities.

Results: Analysis yielded three key insights: (1) older adults often experience a conflict between maintaining independence and implementing safety modifications. One participant described living with a "repeating mantra in my head throughout the day saying 'above all, don't fall.'" (2) aesthetic and privacy concerns frequently override safety benefits. Participants rejected modifications that made their homes feel "institutional." (3) access to occupational therapy services—already limited in rural areas—was further constrained by the COVID-19 pandemic, with some providers reporting that travel time "took up the majority of their day just assessing one home." These barriers help explain the low adoption of home safety assessments despite strong supporting evidence. The study identified design opportunities to address these challenges, including customizable, user-friendly safety solutions, dignity-preserving approaches to assessment, and technology-enabled remote alternatives.

Conclusions: This study identified specific emotional, aesthetic, logistical, and access-related barriers to the adoption of home safety assessments among older adults. The proposed design solutions offer promising directions to increase uptake, improve user experience, and enhance safety. However, further validation through co-design with a larger and more diverse group of older adults is needed. Future research should pilot test these ideas across varied contexts and evaluate their implementation and impact.

(JMIR Hum Factors 2025;12:e66854) doi:10.2196/66854

KEYWORDS

human-centered design; fall prevention; home safety assessments; occupational therapy; aging in place; elderly; older adult; design; fall; home safety; safety; therapist; aging; perspectives; community-dwelling; development; qualitative research; qualitative data; semi-structured interviews; interview; occupational therapists; geriatricians; surgeons

XSL•FO
RenderX

Introduction

A total of 1 out of 3 people aged 65 years and older falls annually, and falls are subsequently the leading cause of injury-related death in this age group [1,2]. Falls result in 2.8 million emergency department visits every year, with 25% of admissions being associated with serious injuries, such as fractures or traumatic brain injury, which can permanently decrease quality of life [3]. Fatal and nonfatal falls cost an estimated US $50 billion in the United States in 2015, which increased up to US $80 billion in 2020, with the bulk of the cost attributed to nonfatal falls and paid for by Medicare [4,5].

Falls and fear of falling in turn threaten older adults' ability to live independently and often lead to self-imposed activity restriction, social isolation, decreased cognitive and physical function, and, paradoxically, an increased risk of serious falls [6,7]. The emotional impact of falling is grave, with studies finding that 80% of older women would prefer death to a bad hip fracture that would lead to nursing home admission [8]. Alternative solutions to receiving 24-hour monitoring and high-level medical care in a nursing home, such as moving into an assisted living community where support is available for activities of daily living, are a last resort for many. The vast majority (90%) of older adults aged 50 years or older prefer to age in place in their current home [9]. However, 85% of these same individuals who desire to remain living at home have not adequately prepared their homes for aging in place through efforts such as removing trip hazards and installing assistive devices (eg, bathroom grab bars) [9]. Finding a wide range of ways to reduce falls can potentially both increase the quality of life for older adults choosing to age independently and also lead to large savings in direct medical costs [10].

The risk of falling and resulting serious injury are associated with many modifiable and nonmodifiable risk factors. Nonmodifiable risk factors include age, cognitive impairment, female sex, and a history of falling [11]. Modifiable risk factors include balance impairment, gait impairment, muscle weakness, medication use, and home environment [11]. Unfortunately, there is no one single risk factor for falls upon which to intervene. Fall risk is multifactorial, and falls tend to occur when impairments exist in multiple domains of an individual's health [12].

There has been extensive research on interventions that can reduce the risk of fall-related injury and mortality in the older adults. For example, home safety assessments performed by an occupational therapist for community-dwelling older adults have been shown to significantly decrease the rate and risk of falling by up to 36%, along with reducing serious fall-related complications, including hip fractures and death [13,14]. Keall et al [15] randomized 842 senior households to an intervention involving a standardized set of home safety modifications including staircase handrails, bathroom grab bars, improved lighting, nonslip mats and rugs, and an educational brochure on home safety, which demonstrated a 26% decrease in fall injuries in the intervention group compared with the control group.

Despite the clear, data-driven benefits of home safety assessments and related home environmental improvements, adoption of these services remains low for a variety of reasons. Some insurance companies that previously covered home safety assessments as a benefit for their members have discontinued this benefit due to low use by older adult policyholders [14,16]. One key reason for the underuse of home safety assessments is that many older adults perceive a visit from a health care provider as a reminder of their inevitable aging and an invasion of their privacy, with some reacting to suggested changes as an encroachment on their personal independence, dignity, and sense of control [17-19] Personalized care approaches that prioritize individual preferences can significantly improve intervention acceptance and perceived autonomy [20]. Furthermore, concerns about loss of independence often outweighs the fear of injury, leading older adults to underuse fall prevention interventions [21].

Older adults' resistance to addressing the physical, social, and emotional changes associated with aging may be exacerbated by their minimal inclusion in decision-making processes, particularly those related to their living spaces [18,22]. Multiple studies suggest that this exclusion increases resistance to home modifications and other aging-related changes, as older adults often feel a loss of control when they are not actively involved in decisions affecting their environment [19,20]. By involving older adults in these decisions, their sense of autonomy and dignity can be maintained, which in turn reduces emotional distress and increases acceptance of necessary changes [21].

Understanding the barriers older adults face in using home safety assessments is critical to addressing modifiable fall risk factors and reducing falls. This study aimed to (1) identify specific barriers perceived by community-dwelling older adults in accessing and implementing home safety assessments and modifications, (2) explore attitudes of health care professionals and additional stakeholders toward home safety assessments, and (3) identify novel design opportunities to guide the development and implementation of more effective home safety assessment techniques and practices.

Methods

Study Design and Setting

This explanatory qualitative research study used human-centered design (HCD) methodology to understand the barriers to home safety assessments and develop opportunities to support fall prevention for community-dwelling older adults [23,24]. The study was performed as a collaboration between Stanford University's Design for Extreme Affordability course within the Hasso Plattner Institute of Design (d.school) and The Better Lab, an academic research laboratory focused on HCD in health care, based in the Department of Surgery at the University of California, San Francisco. Study activities took place between February and June 2021, with interviews conducted both virtually and in person at Stanford Hospital and Clinics.

Research Team Characteristics

The research team consisted of 3 designers from Stanford's d.school, 2 clinical researchers from the University of California,

XSL•FO
RenderX

San Francisco (both practicing geriatricians), and 2 HCD researchers from "The Better Lab." Team members had expertise in geriatric medicine, occupational therapy, design research, and public health. A total of 4 members had prior experience conducting qualitative research with older adults, and all researchers received training in qualitative interviewing techniques to ensure consistency. We acknowledge that our clinical and design backgrounds may have influenced our focus on intervention development. To mitigate bias, our team engaged in regular discussions to examine assumptions and incorporate peer debriefing throughout the study. None of the researchers had prior relationships with study participants.

Participant Selection

We used purposeful sampling to ensure heterogeneity and diverse representation of stakeholder perspectives [25]. Our sampling strategy aimed to include individuals with varied relationships to home safety assessment implementation, including (1) community-dwelling older adults who might benefit from assessments, (2) health care providers who conduct or refer for assessments, (3) family members who support older adults in making home modifications, (4) program administrators who oversee assessment services, (5) product developers working on related technologies such as assistive technology, smart home solutions, or fall prevention tools.

Our selection criteria included age (65+ years for older adults), direct involvement in home safety assessments, and willingness to discuss their experiences. We specifically sought older adults with varying experiences with falls and home modifications, ranging from those who had never had an assessment to those who had implemented modifications. Health care providers were selected to represent different specialties and practice settings.

HCD Process

Over the past decade, the use of HCD in health care has grown substantially. Case studies and qualitative analyses suggest that using HCD approaches may enhance patient engagement and improve service delivery, potentially leading to better health care access and outcomes [26-28]. HCD is problem-solving methodology rooted in engineering, psychology, anthropology, business, and design that brings together designers and stakeholders of an issue to identify challenges and use an iterative process to cocreate solutions [29]. HCD uses ethnographic research methods to identify unmet needs and contextualizes the study challenge within the greater ecosystem to identify impactful and sustainable solutions [24,26-28].

The HCD process consists of three phases: inspiration, ideation, and implementation [30]. The inspiration phase involves identifying unmet needs through interviews and observations with key users and related stakeholders [31]. Qualitative data from these interviews and observations are synthesized to identify themes and inform the creation of insight statements and design opportunities. The ideation phase involves convening users, stakeholders, and designers for group workshops aimed at brainstorming potential solutions and co-designing prototypes of the most desirable, feasible, and viable potential solutions [30,32]. Finally, the implementation phase consists of iterative testing, incorporating user and stakeholder feedback to refine prototypes for pilot testing [30]. This paper describes the inspiration phase of our study in which we explored barriers to home safety assessments and modifications by interviewing older adults, health care professionals, and other key stakeholders.

Inspiration Phase Qualitative Data Collection and Analysis

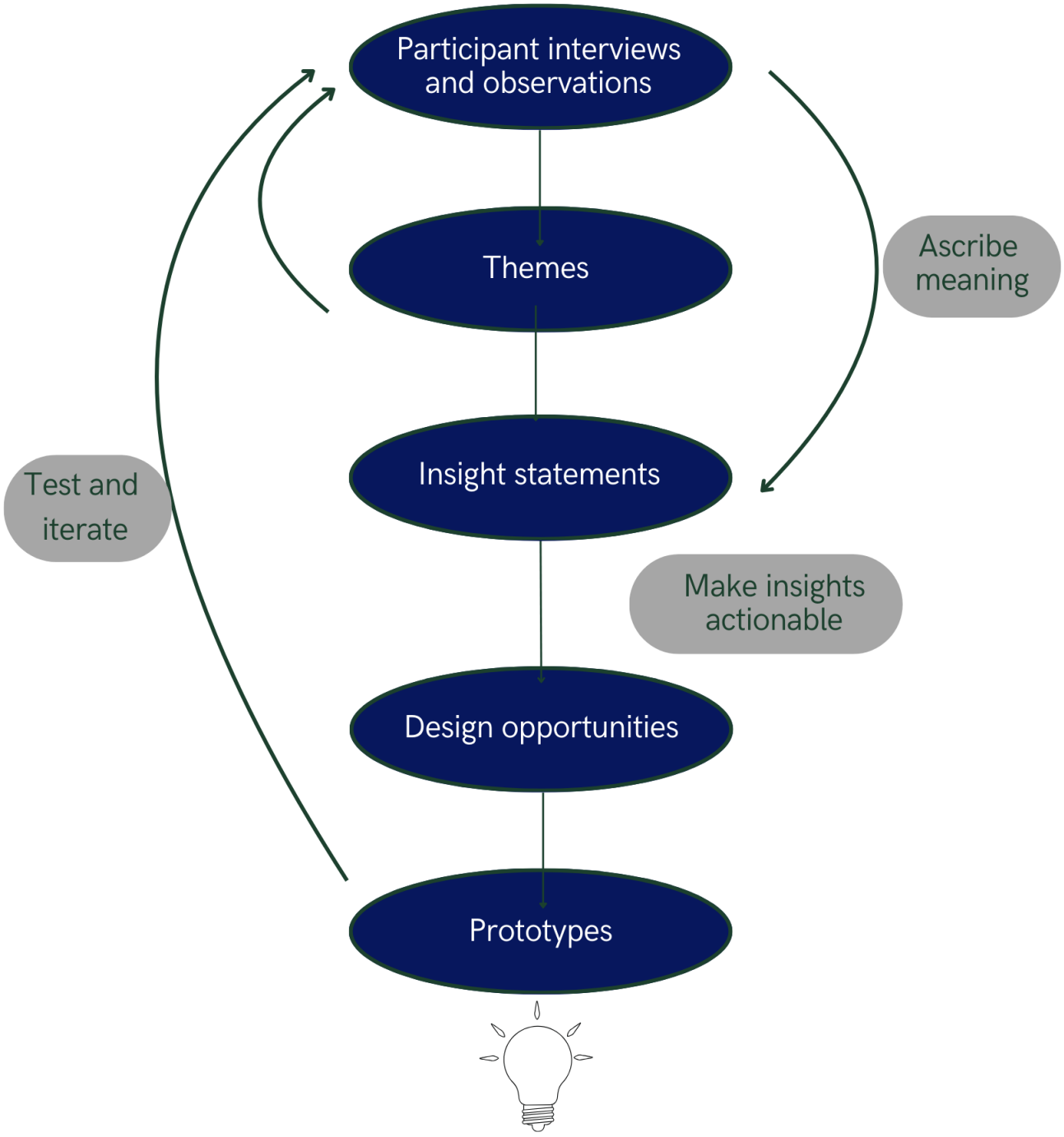
We recruited all participants using purposeful sampling to ensure heterogeneity and a diverse representation of perspectives in the study cohort, consistent with the HCD methodology [33]. Recruitment criteria involved individuals directly involved with home safety assessments or fall prevention. We obtained qualitative data through interviews with older adult users and other key stakeholders, including clinicians (physical therapists, occupational therapists, geriatricians, and surgeons), family members of older adults, engineers, senior living community administrators, interior designers, and payers. We created a semistructured interview guide to understand factors that contribute to falls, experiences within the health care system following a fall, and the psychological impact of falls.

We interviewed a total of 28 participants for 30 - 60 minute interviews conducted virtually over Zoom (Zoom Communications, Inc) or in person at Stanford Hospitals and Clinics between February and June 2021. In light of COVID-19 restrictions, we conducted interviews both virtually and in-person, and in both settings we ensured participant comfort and openness to create a flexible environment for discussion. Two to three researchers per session conducted the interviews, with 1 lead researcher leading the interview while the other researchers transcribed detailed notes.

Figure 1 illustrates this study's multistep HCD qualitative analysis process, which we followed up to the step of developing design opportunities. We followed a 3-step approach using reflexive thematic analysis [34]. First, 3 team members, who were students with a background in design and working in partnership with a HCD expert, independently conducted an inductive thematic analysis of the interview notes to identify emerging themes and refined them through discussions until consensus was reached [35,36]. Second, the entire research team collaboratively developed and reached consensus on multiple insight statements, capturing relational themes such as tensions, synergies, barriers, and facilitators to health behavior described by participants. Third, all research team members collaborated to translate each insight statement into 1 or more design opportunities through structured discussions, with each opportunity representing actionable concepts or intervention elements that directly address the challenges identified in participant responses [37],[38-39].

XSL•FO
RenderX

Figure 1. Human-centered design qualitative data analysis process diagram.



Ethical Considerations

This study was determined to be exempt from review by Stanford University's Institutional Review Board (60770). All participants provided verbal informed consent after reviewing study information. Given the sensitive nature of discussions about falls and home modifications, we took particular care to allow participants to skip any uncomfortable topics, provide support resources when appropriate, and to protect confidentiality in our data management. Participation was voluntary, and interviewees received US $20 gift cards as compensation. Interviews were anonymized to protect patient confidentiality. No interviews were audio or video recorded due to the sensitive and confidential nature of these conversations.

Results

Overview

A total of 35 interviews (28 initial and 7 follow-up) were conducted between February and June 2021. Our participant group comprised 28 diverse stakeholders: 3 community-dwelling older adults, 4 geriatricians, 6 therapists (including physical, occupational, and mental health–focused practitioners), 2 product managers, 8 researchers specializing in older health, and 5 leaders of older care programs and communities. Of these, 18 (64%) participants identified as women. Interview lasted between 30 and 60 minutes.

XSL•FO
RenderX

Through inductive analysis of interview data, we identified 3 key insights regarding the interconnected and sometimes opposing perspectives of older adult users, occupational therapists, community living organizations, and geriatricians on fall prevention and home assessments. Below, we present these insights along with proposed design opportunities that emerged from our analysis.

Insight 1

Older adults have a strong desire to maintain active, independent lifestyles that align with a sense of youthfulness, yet alterations to their environments and lifestyles that could enhance their well-being and safety are perceived as undermining independence, vitality, and creating fear around being active.

Older adults may be overwhelmed and disoriented by an increasing loss of control over their health, lifestyle, and relationships. For example, one older adult interviewee described an all-consuming fear of falling following her first fall, such that she hears a repeating mantra in her head throughout the day saying, “above all, don’t fall.” Much like an initial fall, a visit from a health care provider can remind older adults about their inevitable aging, which may contribute to resistance against home assessments. One geriatrician interviewee observed that “patients are reminded of their mortality when an occupational therapist comes to the home to visit.” In addition, multiple stakeholders shared the perspective that during home assessments there is often an extensive list of home modifications that are suggested to make older adults’ homes safer without much consideration of how older adults feel about these changes. For example, one older adult user who had grab bars placed throughout her home reflected that “the way they looked around her home” was a key reason she did not want them installed. The gray color and material of the standard grab bar reminded her of the hospital room, she stayed in for treatment of a former injury, which she wanted to avoid recreating in her home. This suggests that to increase adoption of an effective fall-prevention or safety assessment service, users need to feel motivated and not fearful to make changes to their lifestyle and that engaging users in customizing adaptations to their aesthetic preferences may be a method to improve users’ perception of control.

Much like baby-proofing a home is a sign of natural change in a family unit to protect their well-being, “age-proofing” serves a similar purpose for older adults. As one gerontology expert interviewee noted, “Homes age over time as well, and there’s this idea that a home ages and changes with you.” However, choosing not to “age-proof” their home may be an older adult’s way of rejecting the notion that their lifespan is limited. For example, older adults in our interviews have expressed sentiments such as, “We don’t want to go into a nursing home or senior housing because that’s where old people live.” This fear of aging may lead to a maladaptive way of dealing with the aging process in some individuals, similar to how a fear of falling can paradoxically increase the risk of falls.

Design Opportunity 1

Return a sense of control and autonomy to older adults in the home redesign or adaptation process. A home safety assessment tool can be designed to prioritize customizability and facilitate direct feedback from users. The essential features of this product include:

Active Participation

Engaging older adults as active participants and decision-makers in any home redesign or change so that they can shape their own personal aging experience, which is currently often overlooked.

Customizable Solutions

Providing a customizable system of home modifications that older adults can use to experiment with different materials or patterns (eg, mahogany wood for safety devices such as grab bars) to discern their aesthetic and functional preferences.

Positive Framing

Generating enthusiasm from older adults for home modifications by framing them as enhancing rather than diminishing independence, helping them feel more aligned in their goals of care with occupational therapists and geriatricians.

Growth Mindset Toward Aging

Encouraging a growth mindset to aging that emphasizes openness to adapting in a healthy, safe, and dignified manner while maximizing the opportunities available to older adults for engaging in daily life activities and aging in place.

Insight 2

Older adults are willing to risk their safety, even their lives and health, to protect the familiarity of their homes from the change and invasion of privacy associated with home safety assessments.

Older adults know the layout of their homes intimately; one older adult interviewee described this as follows: “We know our home just as well as you know the layout of your childhood home… and know how to avoid the tiny things even when waking up in the middle of the night.” An academic public health specialist in California who focuses on issues affecting the geriatric population, expressed in an interview that “[Older adults] are reluctant to let a stranger into their homes to change things.” Multiple older adult interviewees expressed that they had already heard from their children and relatives that they need to make their home safer through modifications but had rejected that advice. One older interviewee cited aesthetics as the reason for rejecting suggested changes saying, “My kids have been telling me for months now to clean up my clutter and install a grab bar in my shower, but I just don’t like the way it looks in my home.” However, according to a geriatrician interviewee, when considering accepting support to adapt their homes to support aging in place, “the elderly struggle with dignity.” Home was described by multiple older adult interviewees as representing a personal haven, or area of refuge and constancy away from the influences and preferences of others. Maintaining control over their home may be considered an aspect of living with dignity for the community-dwelling older adults in the current sample.

Both older adult users and clinicians shared in interviews that resistance to changing the home persists even after an older

XSL•FO
RenderX

adult's first fall, suggesting that users' disinterest in making home modifications is not rooted in concern for their health, but rather in the emotions associated with aging, change, and privacy. This idea is supported by the perspective of a fall prevention center leader who stated in their interview that "They [older adults] don't want to change, and this term [home modifications] isn't enough to conquer their fear of change." The design opportunities that we framed from the insights gleaned from interview data, therefore emphasized understanding and addressing older adults' powerful emotions and feelings of fear of change and loss of independence in any solution.

Design Opportunity 2

Reframe how home safety assessments are perceived by users so that older adults can become fully engaged in the broader process of safely aging in place while protecting themselves and preserving their dignity and independence. The home safety modification process can become an engaging opportunity for an older adult to design their personalized "forever" home while also enhancing their safety. The essential features include:

Inspiring a Positive Outlook

Encouraging older adults to embrace the next stage of their life using home safety assessments and modifications.

Connecting With Trusted Professionals

Linking users with local contractors who can install both the necessary fall prevention devices as well as any additional renovations that the older adult would like to enhance their safety while maintaining the comfort of home.

Providing Educational Resources

Offering supplemental educational resources about changes that occur during aging to help users conquer their fear of losing dignity and independence.

Insight 3

Access to home safety assessments would improve older adults' safety yet they are perceived as unsafe or inaccessible, especially during the COVID-19 pandemic. The COVID-19 pandemic illuminated and exacerbated the existing misalignment in supply and demand for occupational therapists and home safety assessments. It also opened opportunities to assess technological approaches to work on the problems of accessibility and acceptability of home safety assessments for older adults. Many community and aging organizations such as Ashby Village in California, have turned to trained volunteers to perform home safety assessments for individuals within their network given the high demand for occupational therapists. During the COVID-19 pandemic, when quarantines and social distancing were established, in-person home safety assessments within these communities were halted. During our interview in 2022, Ashby Village organizers mentioned to us that "we haven't done any home assessments since March 2020 and do not plan on doing any for the foreseeable future." Though they resumed this service postpandemic, in the interim many users who might benefit from a low-cost fall prevention service were left without access. Even though larger health care networks continued to perform in-person visits during the pandemic, we heard from older adult users about their resistance to having a stranger come into their home. "If I can't even feel safe inviting family into my home, why would I invite a complete stranger in?"

Finally, even large health care networks have struggled to reach older adults who live in rural areas. As one interviewee noted, a home safety assessment process for such an individual would consist of "sending an occupational therapist on a 2-hour commute each way to see this person which can take up the majority of their day just assessing one home." This suggests a current unmet need in the home safety assessment space, which likely disproportionately affects rural and underserved populations and will continue to grow as the population ages.

Design Opportunity 3

Increase access to home safety assessment tools using alternative and innovative technologies. Online home safety assessment tools would allow users to benefit from a home safety assessment and a reduced fall risk, even with fewer volunteers and limited in-person visits due to the pandemic. The essential features of such a tool could include:

Online Home Assessments

Developing an online assessment tool which allows users to submit photos of various rooms in their home for review by licensed occupational therapists, which can produce a risk score for each room and identify areas that should be addressed.

Future Adaptability

Implementing computer vision and machine learning, which can examine visual data representations and automate the home safety assessment process over time.

Personalized Recommendations

Combining visual data with user surveys that contain their demographic information and design preferences to generate personalized home modification suggestions, ensuring adaptations accommodate mobility aids (eg, walkers), while remaining functional and aesthetically pleasing to the user.

Further possibilities using a wide range of technological tools and equipments, including smartphone and tablet apps, video games, wearable devices, and future robotic efforts to improve home safety assessments. Their potential impacts on fall prevention efforts are discussed in the "Discussion" section.

Discussion

Principal Findings

This study provides insights into the complex factors affecting the adoption of home safety assessments among older adults and suggests potential directions for improving implementation. Our findings highlight 3 key areas of tension existing between stakeholders in aging in place process, including older adults, clinicians, home assessors, and others: maintaining independence while ensuring safety, preserving home sanctuary while accepting modifications, and balancing assessment access with implementation constraints. Since older adults in the current sample desired to preserve their independence and dignity, they were resistant to implementing recommended changes, especially those related to their home safety, as they wanted to

XSL•FO
RenderX

avoid the negative emotions associated with considering how aging impacts physical and other abilities. The strong emotional responses to home modifications revealed in our study align with previous research showing that older adults often perceive environmental changes as threats to independence [40]. However, our findings extend this understanding by highlighting how aesthetic and personal meaning considerations may override known safety benefits. This suggests that purely educational or risk-focused approaches may be insufficient to drive behavior change.

The COVID-19 pandemic's impact on assessment delivery highlighted both challenges and opportunities in service provision. While disrupting traditional in-person assessments, it accelerated exploration of remote alternatives. However, technological solutions must carefully consider older adults' preferences and potential barriers to adoption.

Technology has become increasingly integrated into the fall prevention research and health care field, especially following the 2008 decision by the Centers for Medicare and Medicaid Services to stop reimbursing costs related to in-hospital falls [41,42]. For example, smartphone and tablet apps have been developed to encourage greater levels of exercise in older adults, which in turn can reduce falls [43]. Similarly, video games such as the Wii Fit have been successfully used to perform mobility-based fall risk assessments in older adults [44]. Wearable sensors have also been increasingly studied, with a systematic review identifying 40 studies and 130 distinct variables examining sensor use in geriatric falls with promising results [45]. Technologically supported interventions may reduce fall risk either by directly minimizing an older adult's physical risk of falling or by improving their confidence to reduce their fear of falling, which is an important predictor of falls. Older adults' adoption of wearable sensors such as fitness trackers could lead to the accumulation of "big data"—a large-scale dataset automatically collected from the sensors themselves, potentially enabling analysis and detection of patterns in fall risk and incidence to be identified more easily—though, this would depend on greater uptake of digital technologies in this population than has yet occurred.

Only a limited number of studies have used technology specific to the home safety assessment process. For example, Du et al [46] used a robotic system operated remotely by occupational therapists to conduct home safety assessments. In addition to providing a video livestream of the home, the system was able to measure distances between objects, detect lighting situations, and communicate with the older adult. Furthermore, ambient camera-based sensors have also been tested for automatic fall detection [47]. However, both solutions face challenges in uptake, given the cost associated with the technological systems and the time needed for installation and for the user to learn the system. Another study collated a protocol to be used for remote home safety assessments through extensive secondary research and interviews with occupational therapists and caregivers [48]. Taken together, these previous studies and our current one highlight the increasing recognition of home safety assessments as a potential area of improvement in fall prevention, and evidence of the value of future investments in user-friendly technological interventions for fall prevention and awareness among older adults.

Limitations

This study has several important limitations. First, while we sought to understand older adults' perspectives, only 3 of our 28 participants were older adults themselves. This occurred primarily due to recruitment challenges related to COVID-19. This limited representation may not capture the full range of older adult views and experiences. Future work should more deeply engage older adults both as study participants and as co-designers of proposed solutions. Second, our proposed design opportunities emerged from researcher analysis rather than cocreation with stakeholders. While grounded in participant insights, these proposals require validation through future participatory design work with end users and stakeholders. Third, all participants were recruited from the San Francisco Bay Area, potentially limiting generalizability to other geographic and cultural contexts. Health care access, technology adoption, and attitudes toward aging may differ substantially in other regions. Finally, our use of interview notes rather than full transcripts may have missed subtle themes or nuances in participant responses. This approach was selected to maintain the confidentiality of participants experiencing a sensitive issue. Future studies would benefit from recorded and transcribed interviews to enable more detailed analysis.

Conclusions

This exploratory study reveals important considerations for improving home safety assessment implementation. Successful interventions likely require careful attention to older adults' emotional needs, aesthetic preferences, and desire for autonomy. While technology-enabled approaches show promise for expanding access, they must be developed in partnership with older adults to ensure acceptability and adoption. Our findings suggest that addressing these factors could potentially lead to more effective interventions and increased adherence to home modifications.

Future research should involve larger, more diverse samples of older adults to validate and refine proposed design opportunities using co-design methods. Prototypes of promising approaches should be pilot tested in varied geographic and cultural contexts. It will be important to then evaluate both implementation outcomes and impacts of pilot-tested interventions on fall prevention.

In addition to research advancements, our findings have practical implications for clinical practice. Training programs for occupational therapists should incorporate strategies to address older adults' emotional and psychological barriers to home modifications. Furthermore, the development of educational materials tailored to older adults' concerns—such as maintaining dignity and control while making safety modifications—could facilitate better engagement with home safety interventions.

Rather than prescribing specific solutions, this work provides a foundation for human-centered approaches to fall prevention that honor older adults' dignity while promoting their safety. By integrating these insights into practice, health care

XSL•FO
RenderX

professionals and designers can create more effective, personalized, and widely adopted fall prevention strategies.

Acknowledgments

This research was funded by the Social Entrepreneurship Lab at Stanford University. The authors would like to thank Stanford University students Summer Hu, Chris J Fisher, and Kesaobaka Modukanele for their support in collecting and analyzing qualitative data for this study. The authors would also like to thank Jessie von Hippel and UCSF medical student Saffanat Sumra for their support in editing this manuscript.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

Amanda Sammann is the founder and owner of the Empathy Studio LLC, where Devika Patel was a paid part-time design consultant at the time of the study. Devika Patel was also the paid instructor of a course on human-centered design at San Francisco State University at the time of this study.

References

1. Thomas KE, Stevens JA, Sarmiento K, Wald MM. Fall-related traumatic brain injury deaths and hospitalizations among older adults--United States, 2005. J Safety Res 2008;39(3):269-272. [doi: 10.1016/j.jsr.2008.05.001] [Medline: 18571567]
2. Tromp AM, Pluijm SMF, Smit JH, et al. Fall-risk screening test: a prospective study on predictors for falls in community-dwelling elderly. J Clin Epidemiol 2001 Aug;54(8):837-844. [doi: 10.1016/s0895-4356(01)00349-3] [Medline: 11470394]
3. Shankar KN, Liu SW, Ganz DA. Trends and characteristics of emergency department visits for fall-related injuries in older adults, 2003-2010. West J Emerg Med 2017 Aug;18(5):785-793. [doi: 10.5811/westjem.2017.5.33615] [Medline: 28874929]
4. Florence CS, Bergen G, Atherly A, et al. Medical costs of fatal and nonfatal falls in older adults. J Am Geriatr Soc 2018 Apr;66(4):693-698. [doi: 10.1111/jgs.15304] [Medline: 29512120]
5. Haddad YK, Miller GF, Kakara R, et al. Healthcare spending for non-fatal falls among older adults, USA. Inj Prev 2024 Aug;30(4):272-276. [doi: 10.1136/ip-2023-045023]
6. Wijlhuizen GJ, de Jong R, Hopman-Rock M. Older persons afraid of falling reduce physical activity to prevent outdoor falls. Prev Med 2007 Mar;44(3):260-264. [doi: 10.1016/j.ypmed.2006.11.003] [Medline: 17184828]
7. Schoene D, Heller C, Aung YN, et al. A systematic review on the influence of fear of falling on quality of life in older people: is there a role for falls? Clin Interv Aging 2019;14:701-719. [doi: 10.2147/CIA.S197857] [Medline: 31190764]
8. Salkeld G, Cameron ID, Cumming RG, et al. Quality of life related to fear of falling and hip fracture in older women: a time trade off study. BMJ 2000 Feb 5;320(7231):341-346. [doi: 10.1136/bmj.320.7231.341] [Medline: 10657327]
9. Home safety for seniors - statistics and solutions. Age Safe America. URL: https://agesafeamerica.com/home-safety-seniors-statistics-solutions/ [accessed 2025-05-22]
10. Stevens JA, Lee R. The potential to reduce falls and avert costs by clinically managing fall risk. Am J Prev Med 2018 Sep;55(3):290-297. [doi: 10.1016/j.amepre.2018.04.035] [Medline: 30122212]
11. Nevitt MC, Cummings SR, Hudes ES. Risk factors for injurious falls: a prospective study. J Gerontol 1991 Sep;46(5):M164-M170. [doi: 10.1093/geronj/46.5.m164] [Medline: 1890282]
12. Tinetti ME, Inouye SK, Gill TM, Doucette JT. Shared risk factors for falls, incontinence, and functional dependence. Unifying the approach to geriatric syndromes. JAMA 1995 May 3;273(17):1348-1353. [Medline: 7715059]
13. Cumming RG, Thomas M, Szonyi G, et al. Home visits by an occupational therapist for assessment and modification of environmental hazards: a randomized trial of falls prevention. J Am Geriatr Soc 1999 Dec;47(12):1397-1402. [doi: 10.1111/j.1532-5415.1999.tb01556.x] [Medline: 10591231]
14. Gillespie LD, Robertson MC, Gillespie WJ, et al. Interventions for preventing falls in older people living in the community. Cochrane Database Syst Rev 2012 Sep 12;2012(9):CD007146. [doi: 10.1002/14651858.CD007146.pub3] [Medline: 22972103]
15. Keall MD, Pierse N, Howden-Chapman P, et al. Home modifications to reduce injuries from falls in the home injury prevention intervention (HIPI) study: a cluster-randomised controlled trial. Lancet 2015 Jan 17;385(9964):231-238. [doi: 10.1016/S0140-6736(14)61006-0] [Medline: 25255696]
16. Human Plan Search. Humana. URL: https://plans.humana.com/filters/coverage-type [accessed 2022-04-14]
17. Mann WC, Ottenbacher KJ, Fraas L, Tomita M, Granger CV. Effectiveness of assistive technology and environmental interventions in maintaining independence and reducing home care costs for the frail elderly. A randomized controlled trial. Arch Fam Med 1999;8(3):210-217. [doi: 10.1001/archfami.8.3.210] [Medline: 10333815]

XSL•FO
RenderX

18. Ma C, Guerra-Santin O, Mohammadi M. Smart home modification design strategies for ageing in place: a systematic review. J Hous and the Built Environ 2022 Jun;37(2):625-651. [doi: 10.1007/s10901-021-09888-z]
19. Wiseman JM, Stamper DS, Sheridan E, et al. Barriers to the initiation of home modifications for older adults for fall prevention. Geriatr Orthop Surg Rehabil 2021 Jan;12:21514593211002161. [doi: 10.1177/21514593211002161]
20. Coulter A, Entwistle VA, Eccles A, et al. Personalised care planning for adults with chronic or long-term health conditions. Cochrane Database Syst Rev 2015 Mar 3;2015(3):CD010523. [doi: 10.1002/14651858.CD010523.pub2] [Medline: 25733495]
21. Lenouvel E, Ullrich P, Siemens W, et al. Cognitive behavioural therapy (CBT) with and without exercise to reduce fear of falling in older people living in the community. Cochrane Database Syst Rev 2023 Nov 15;11(11):CD014666. [doi: 10.1002/14651858.CD014666.pub2] [Medline: 37965937]
22. Wiles JL, Leibing A, Guberman N, et al. The meaning of “aging in place” to older people. Gerontologist 2012 Jun;52(3):357-366. [doi: 10.1093/geront/gnr098] [Medline: 21983126]
23. Ivankova NV, Creswell JW, Stick SL. Using mixed-methods sequential explanatory design: from theory to practice. Field methods 2006 Feb;18(1):3-20. [doi: 10.1177/1525822X05282260]
24. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000 Aug;23(4):334-340. [doi: 10.1002/1098-240x(200008)23:4<334::aid-nur9>3.0.co;2-g] [Medline: 10940958]
25. Palinkas LA, Horwitz SM, Green CA, et al. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. Adm Policy Ment Health 2015 Sep;42(5):533-544. [doi: 10.1007/s10488-013-0528-y] [Medline: 24193818]
26. Gitlin LN, Winter L, Dennis MP, et al. A randomized trial of a multicomponent home intervention to reduce functional difficulties in older adults. J American Geriatrics Society 2006 May;54(5):809-816. [doi: 10.1111/j.1532-5415.2006.00703.x]
27. Oswald F, Wahl HW. Housing and health in later life. Rev Environ Health 2004;19(3-4):223-252. [Medline: 15742672]
28. Gitlin LN, Hauck WW, Dennis MP, Winter L. Maintenance of effects of the home environmental skill-building program for family caregivers and individuals with Alzheimer’s disease and related disorders. J Gerontol A Biol Sci Med Sci 2005 Mar;60(3):368-374. [doi: 10.1093/gerona/60.3.368] [Medline: 15860476]
29. Beres LK, Simbeza S, Holmes CB, et al. Human-centered design lessons for implementation science: improving the implementation of a patient-centered care intervention. J Acquir Immune Defic Syndr 2019 Dec;82 Suppl 3(3):S230-S243. [doi: 10.1097/QAI.0000000000002216] [Medline: 31764259]
30. Wong LE, Hawkins JE, Langness S. Where are all the patients? Addressing Covid-19 fear to encourage sick patients to seek emergency care. NEJM Catalyst Innovations in Care Delivery 2020 May [FREE Full text]
31. Tschimmel K. Design thinking as an effective toolkit for innovation. 2012 Presented at: XXIII ISPIM Conference: Action for Innovation: Innovating from Experience; Jun 17-20, 2012; Barcelona. [doi: 10.13140/2.1.2570.3361]
32. Altman M, Huang TTK, Breland JY. Design Thinking in Health Care. Prev Chronic Dis 2018 Sep 27;15:E117. [doi: 10.5888/pcd15.180128] [Medline: 30264690]
33. Compton-Phillips A, Mohta NS. Care redesign survey: how design thinking can transform health care. NEJM Catal 2018 [FREE Full text]
34. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qualitative Research in Sport, Exercise and Health 2019 Aug 8;11(4):589-597. [doi: 10.1080/2159676X.2019.1628806]
35. Traditional socio-behavioral research and human-centered design: similarities, unique contributions and synergies. IMPT. URL: https://theimpt.org/traditional-socio-behavioral-research-and-human-centered-design-similarities-unique-contributions-and-synergies/ [accessed 2021-07-17]
36. Thomas DR. A general inductive approach for analyzing qualitative evaluation data. American Journal of Evaluation 2006 Jun;27(2):237-246. [doi: 10.1177/1098214005283748]
37. Patel D, Sarlati S, Martin-Tuite P, et al. Designing an information and communications technology tool with and for victims of violence and their case managers in San Francisco: human-centered design study. JMIR Mhealth Uhealth 2020 Aug 24;8(8):e15866. [doi: 10.2196/15866] [Medline: 32831179]
38. What is insight? The 5 principles of insight definition. Thrive. URL: https://thrivethinking.com/2016/03/28/what-is-insight-definition/ [accessed 2021-09-11]
39. Turning insights into opportunities. : Bridge Innovate; 2021. URL: https://www.bridgeinnovate.com/blog/2017/12/19/turning-insights-into-opportunities [accessed 2021-09-11]
40. Tanner B, Tilse C, de Jonge D. Restoring and sustaining home: the impact of home modifications on the meaning of home for older people. J Hous Elderly 2008 Sep 11;22(3):195-215. [doi: 10.1080/02763890802232048]
41. Oh-Park M, Doan T, Dohle C, et al. Technology utilization in fall prevention. Am J Phys Med Rehabil 2021 Jan 1;100(1):92-99. [doi: 10.1097/PHM.0000000000001554] [Medline: 32740053]
42. Lord SR, Close JCT. New horizons in falls prevention. Age Ageing 2018 Jul 1;47(4):492-498. [doi: 10.1093/ageing/afy059] [Medline: 29697780]
43. Delbaere K, Valenzuela T, Woodbury A, et al. Evaluating the effectiveness of a home-based exercise programme delivered through a tablet computer for preventing falls in older community-dwelling people over 2 years: study protocol for the

XSL•FO
RenderX

standing tall randomised controlled trial. BMJ Open 2015 Oct 22;5(10):e009173. [doi: 10.1136/bmjopen-2015-009173] [Medline: 26493461]
44. Yamada M, Aoyama T, Nakamura M, et al. The reliability and preliminary validity of game-based fall risk assessment in community-dwelling older adults. Geriatr Nurs (Lond) 2011 May;32(3):188-194. [doi: 10.1016/j.gerinurse.2011.02.002]
45. Howcroft J, Kofman J, Lemaire ED. Review of fall risk assessment in geriatric populations using inertial sensors. J Neuroeng Rehabil 2013 Aug 8;10(1):91. [doi: 10.1186/1743-0003-10-91] [Medline: 23927446]
46. Du R, Jagtap V, Long Y, et al. Robotics enabled in-home environment screening for fall risks. Association for Computing Machinery 2014 Jun 11:9-12. [doi: 10.1145/2609829.2609836]
47. Riquelme F, Espinoza C, Rodenas T, et al. eHomeSeniors dataset: an infrared thermal sensor dataset for automatic fall detection research. Sensors (Basel) 2019 Oct;19(20):4565. [doi: 10.3390/s19204565]
48. Romero S, Lee MJ, Simic I, et al. Development and validation of a remote home safety protocol. Disabil Rehabil Assist Technol 2018 Feb;13(2):166-172. [doi: 10.1080/17483107.2017.1300345] [Medline: 28326967]

Abbreviations

HCD: human-centered design

Edited by P Santana-Mancilla; submitted 25.09.24; peer-reviewed by R Buenrostro, S Rosenbaum; revised version received 21.04.25; accepted 25.04.25; published 24.06.25.

Please cite as:
Lee JJ, Patel D, Gadgil M, Langness S, von Hippel CD, Sammann A
Understanding Barriers to Home Safety Assessment Adoption in Older Adults: Qualitative Human-Centered Design Study
JMIR Hum Factors 2025;12:e66854
URL: https://humanfactors.jmir.org/2025/1/e66854
doi:10.2196/66854

© Jonathan J Lee, Devika Patel, Meghana Gadgil, Simone Langness, Christiana Dagmar von Hippel, Amanda Sammann. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 24.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Self-Determined Health App Evaluation Questionnaire Development: Mixed Methods Study

Angelika Rzepka[1], DI (FH); Kurt Edegger[1], DI; Stefan Welte[1], DI (FH); Diotima Bertel[2], Mag; Anja Mandl[3], MA; Günter Schreier[1], PhD

[1]Center for Health & Bioresources, AIT Austrian Institute of Technology GmbH, Giefunggasse 4, Wien, Austria

[2]Center for Technology Experience, AIT Austrian Institute of Technology GmbH, Wien, Austria

[3]Gesundheitsfonds Steiermark, Graz, Austria

Corresponding Author:
Angelika Rzepka, DI (FH)
Center for Health & Bioresources, AIT Austrian Institute of Technology GmbH, Giefunggasse 4, Wien, Austria

Abstract

Background: The rapid increase in the number of health apps and their volatility in Austrian pp stores for Android and Apple are signs of a flourishing business sector in the wellness industry.

Objective: In this report, a questionnaire for informed decision-making by users was developed and evaluated using health apps in the categories “Nutrition”, “Exercise”, “Mental Health” and “Symptom Checker”.

Methods: Evaluation criteria were derived from multiple reference documents and weighted in a survey, as well as a focus group meeting. Further, the selected evaluation criteria were tested against selected apps, which were most popular in the above-mentioned categories in fall 2023.

Results: The short questionnaire is to be made publicly available to citizens and covers the categories of the quality of the app provider (regulatory compliance, safety, and quality assurance), the content quality of the app (functionality and evidence-based content), and the user-friendliness of the app. It consists of 6 question items, which appeared to be problematic in the evaluation most often. For ease of use, a short questionnaire with the most critical questions and helpful tips from the evaluation was expanded by the question of importance and with meaningful advice on how to find the right information in an app.

Conclusions: The evaluation based on the comprehensive questionnaire using selected apps confirmed the importance of such a questionnaire since most of the apps lacked basic properties in terms of safety and security.

(JMIR Hum Factors 2025;12:e63739) doi:10.2196/63739

KEYWORDS

health apps; app quality; DiGA; nutrition; sports; Mental Health; Symptom Checker; health applications; diet; symptom; questionnaire; evaluation; development; physical exercise; exercise; survey; quality assurance; user-friendliness

Introduction

Background

The global as well as national supply of health apps is overwhelming, and more and more people are using health apps. This raises the basic questions of which apps can be considered health apps, and how can the credibility of these apps be better assessed.

Fundamentally, a distinction is made between [1] health and mobile apps. Health apps are applications targeting consumers. They are intended to be used for preventive purposes and help promote or change lifestyle in a health-promoting manner.

On the other hand, medical apps are applications used to support “self-empowerment” (self-determination) and, for example, are supportive in coping with chronic diseases or in rehabilitation [2], which are subject to regulations like the medical device regulations.

The use of evaluation criteria for health apps is important for several reasons, as described below.

Functionality

Evaluation criteria can assess the functionality of health apps to ensure that the intended tasks are effectively fulfilled. This could include monitoring health parameters, reminding users to take medication or providing personalized health recommendations.

Quality assurance

Evaluation criteria help assess the quality of health apps. This is particularly important because the information and functions provided by health apps can have a direct impact on the health

XSL•FO
RenderX

of users. Criteria enable ensuring that the apps provide accurate, reliable, and evidence-based information.

Evidence-based content

Health apps should be based on scientific knowledge to provide accurate and reliable information. Evaluation criteria can verify whether the content presented in the apps is based on current scientific evidence.

Usability and user experience

The usability of health apps is crucial for their effectiveness. Evaluation criteria can assess usability and accessibility to ensure that the apps are easily understandable and accessible. Furthermore, accessibility criteria may be important for allowing the apps to be used by impaired users.

Privacy

Apps may process personal health data, so it is crucial to ensure that only the minimal set of data is acquired ("data minimization") and that data are secured and protected in transit and at rest. Evaluation criteria can consider aspects such as data privacy, encryption, and compliance with regulatory requirements.

Regulatory compliance

Apps, especially health apps may need to comply with certain legal and regulatory requirements. Evaluation criteria can assess whether the apps comply with applicable regulations and standards.

Study Objectives

By applying evaluation criteria, users, health care providers, and regulatory authorities can make more- or better-informed decisions about the selection and use of health apps to ensure that they make a positive contribution to the health and well-being of users. Payors also need clear and transparent criteria to decide on reimbursement aspects.

The objective of this study was to establish evaluation criteria for users to make well-informed decisions in regard to health apps. Health app evaluation frameworks, such as the Mobile App Rating Scale (MARS) [3] and the ENLIGHT Tool [4], provide structured criteria for assessing functionality, usability, and information quality, typically for professional or research use; however, this work focuses on evaluation by laypersons rather than professional organizations, emphasizing user-friendly criteria for informed decision-making.

Methods

Ethical Considerations

Formal ethics approval was not required for this study because no patients or other vulnerable populations were involved. The user survey was completed anonymously by registered participants who volunteered to take part. No identifying information was collected at any stage, and all respondents provided informed consent before proceeding.

Literature Search and Selection of Relevant Documents

The selection of evaluation criteria was conducted as part of a literature search based on the keywords "Bewertungskriterien Apps Gesundheit" (evaluation criteria health apps) and "evaluation criteria health Apps" in English. The search was carried out on the portals Google Scholar and PubMed, and additional literature references from publications were followed up based on the results.

The World Health Organization (WHO) includes "health apps" in the "Classification of Digital Health Interventions v1.0 (DHI)" [5,6]. This primarily categorizes apps according to target groups (eg, for clients or citizens or health care providers) and health system challenges, which then result in system categories for digital health applications.

According to the health system challenges level of health apps, "Lost client data" entails the interventions "1.1.3. Provision of individualized reminder services for clients" and "2.2.1 Longitudinal recording of health status and services of clients", corresponding on a system level to the functions "Client Information System" and "Electronic Health Record". The taxonomy used was based on [7] and the "Guidelines for reporting of health interventions using mobile phones: mobile health evidence reporting and assessment (mERA) checklist" [8].

These two documents generally refer to comprehensive mobile health (mHealth) applications with a strong focus on already manifested diseases.

Partially, reliance was placed on systematic reviews already conducted, such as health technology assessment frameworks for eHealth: a systematic review [9] or framework for supporting reimbursement decisions for digital health applications (mHealth) and its (retrospective) application to selected examples [1].

However, all of these encompass digital health applications, not just health apps. The existing assessment frameworks use different formats (checklists, questionnaires, and tables) and show differences in terms of the spectrum of evaluation domains, additional technology-specific aspects, and target groups. None of the frameworks are fully established or field-tested.

At the European level, there is also the initiative Label2Enable [10]. Label2Enable is an EU-funded project. Its task is to introduce CEN-ISO/TS 82304‐2 and its quality seal for health and wellness apps. The goal is to create a digital single market where high-quality health apps can be extensively used in prevention, health care, and self-care.

The ISO/TS 82304-2:2021(en) "Health software — Part 2 [11]: Health and wellness Apps—Quality and reliability" served as the basis for developing the evaluation criteria.

The following publications also served as the basis for the evaluation criteria:

1. Health apps. Basic paper with special consideration of the aspect of health literacy [12],

XSL•FO
RenderX

2. Opportunities and risks of health apps (CHARISMHA); German: chances and risks of mobile health apps (CHARISMHA) [13],
3. Deutsches Netzwerk Versorgungsforschung e. V. Memorandum - health and medical apps [2],
4. Health apps: How can certification look concretely: Learning from experiences here and elsewhere [14],
5. Digitization and patient safety - HE 2) Checklist for the use of health apps [15],
6. Evaluating, but how? - Criteria for the evaluation and reporting of studies on health and medical apps [16],
7. How good are health apps? What determines quality & risk? What guidance is available? [17],
8. Mobile app rating scale: a new tool for assessing the quality of health mobile apps [18],
9. Methods box health literacy [19],
10. Health apps: significant lever for patient empowerment - potentials, but hardly used so far [20],
11. Professional checklist good health information [21].

Development of Evaluation Criteria

From the above publications, the evaluation criteria mentioned therein were entered into an Microsoft Excel file and assigned to overarching categories. In the next step, the discovered evaluation criteria were harmonized and evaluated. The evaluation criteria with the most matches were included in the list of evaluation criteria and categorized again. The preliminary evaluation criteria were presented and discussed in a focus group, consisting of the project team, health information experts, and public health experts. The evaluation criteria agreed upon in the focus group were prioritized within the focus group meeting and further refined in a quantitative survey of a total of 54 experts with different professions to filter out criteria considered less relevant and to make a final selection. Based on the results of the focus group and the survey, the evaluation criteria were revised and finalized and ended up in a long questionnaire, which is described in the results section.

Evaluation of theQuestionnaire With Selected Apps

The developed questionnaire was evaluated with apps from the following categories: (1) Nutrition, (2) Exercise, (3) Mental Health, and (4) Symptom Checker. These categories were chosen because they cover the main fields of healthy lifestyle. The selection of apps aimed to evaluate the most used apps in Austria in these categories. An attempt to select apps based on existing databases was unsuccessful because the data were at least one year old (eg, Statista [22]) or the databases used a distribution model that was not applicable in this context (eg, data.ai [23]).

For these reasons, an exploratory approach was chosen where 11 individuals from the project team, selected to ensure as much diversity as possible in terms of mobile phones, operating systems, and histories of installed apps, participated. This group consisted of 6 males and 5 females, with 6 participants using the Google Play Store and 5 using the Apple App Store. Each individual searched for the respective keywords in their personal app stores, ensuring a broad variety of search habits and preferences, and recorded the results—15 apps per category—in an Excel spreadsheet. The results were weighted using a pivot table after data cleaning (typos, adding app manufacturers). Then, it was ensured that the top 10 apps in each category were available in both app stores and could be used for evaluation.

The following inclusion and exclusion criteria were additionally considered in the app selection, as listed in Textbox 1:

Textbox 1. Inclusion and exclusion criteria for app selection

Inclusion criteria

- The app must target humans as its audience (excluding apps for pets or similar).
- The app must be available in English or German.
- The app must be available in Austria in the Apple App Store and Google Play Store.

Exclusion criteria

- The app must not require additional hardware (eg, fitness watches, blood glucose meters, etc).
- The app must not be explicitly used for product sales.
- The purpose of the apps must not be to assist with insurance processes (ie, claim management).
- The apps must not require a diagnosis, prescription code, etc.

Each app was assessed by 3 persons [AR, KE, SW] using the long questionnaire (see section results), whereby the persons cross-checked their results within each category of Apps.

Results

Overview

From the comprehensive evaluation criteria, we derived 2 questionnaires (short and long) to enable an informed and self-determined decision regarding the assessment of these apps. All questions in the following tables are to be answered with (yes/no) along with a detailed description. The responses serve as a guidance for the individual decision-making process regarding the app's utility and quality.

Short Questionnaire

The short questionnaire summarizes the questions, which were most often answered with "no" in the evaluation of the long questionnaire and were rated as very important in the focus group and survey mentioned above. Table 1 shows the 6-items

XSL•FO
RenderX

short questionnaire with an explanation of the reason for the importance as well as meaningful advice on how to self-assess the questions quickly.

XSL•FO
RenderX

Table . Short questionnaire with the most relevant questions for informed decision-making for health apps.

Category	Answer	Why is this important?	Meaningful advice
Imprint: Does the App have an imprint indicating the App's manufacturer?	yes or no	• In Austria, there are legal requirements regarding the "Imprint" that must be complied with.	• Often you can find the imprint in the settings or in the Terms and Conditions (Terms and conditions) of the app. • WARNING: Many apps do not have an imprint. This could be an indication that the app is not trustworthy. • An imprint must contain the following information: • Name or company according to the commercial register, • Address, • Contact information (eg, email address or phone number).
Imprint: Is the contact information in the imprint complete?	yes or no		
User account: Do I need to create an account to use the App	yes or no Attention: Both can have advantages and disadvantages!	• If you have to log in to an app, the provider can more easily associate the data directly with you. This is especially true if you need to provide additional personal information (name, address, phone number, etc.) during registration. • If you don't need to create a user account, that can also be an advantage. You can use the app anonymously. • Warning: If you delete an app without a user account or use another device, your data may be lost.	• If you need to create a user account, the manufacturer can store your data and uniquely identify you. • If you need to log in with a password, no one else can use the app under your name. • Hint: Always use a secure password for health apps.
User account: Can you easily delete this account again?	yes or no		• The function to delete the account is often difficult to find. TIP: Check the app's settings. • Additional steps may be required to delete the user account, such as entering the password.
Data Privacy: Is it specified which specific data the App collects? (for example: Name, Location, etc.)	yes or no	• There are legal requirements for data protection in Austria. These requirements must also be complied with in apps.	• You usually find information about data privacy in the privacy policy and/or the Terms and Conditions (AGB). • Sometimes you may not find this information directly in the app but during the registration process. • The privacy policy must specify the categories of data processed (eg, health data). • WARNING: Often, data not related to the app's purpose is processed (eg, GPS data in a flashlight app).

XSL•FO
RenderX

Category	Answer	Why is this important?	Meaningful advice
Data Privacy: Is it indicated whether this data is shared with third parties?	yes or no	• It is best if an app does not share your data with third parties. If data is shared, it must be indicated.	Sharing of Data with Third Parties: • App providers often take the liberty of sharing data with third parties if it is convenient for them. • Tip: Searching for the word "Third parties" will direct you to the relevant section. • If data are shared with third parties, it must be stated to whom the data is transmitted. • Data are often stored outside the EU and are therefore subject to different data protection laws. • If you search for the word "USA," you can find out if your data isarestored in the United States. Different data protection laws Apply there.
Labeling of Advertising: Are paid offers in the App labeled as "Advertisement" or "Ad"?	yes or no	• Advertising must be labeled and clearly recognizable as such.	To access all contents of the app, a paid subscription may be necessary. • Paid subscriptions in apps are often not labeled as advertisements. • Subscription offers can be annoying and are often difficult to dismiss. • Paid subscriptions are often automatically renewed. Warning: Renewal is often agreed upon when subscribing. Always read carefully. • When comparing the costs of different apps, always consider the total annual costs without initial discounts. • Tip: Pay attention to any cancellation deadlines and minimum terms. • Subscriptions can be canceled either directly in the app or on the providers' websites.
App Function: Is the App's function adequately described?	yes or no	• The function of an app is often the decisive reason for using it. Therefore, the app's function should align with your personal goal.	• You can usually read the app's description before downloading. • Warning: Goals and target groups of the app are often not specified. • Pay attention to advertising promises in the description. Promises are often not fully realized. • Tip: Check if the app's function is usable in the country where you reside. For example, apps in the "Nutrition" category may use databases from the USA, which may not be usable in Austria. • Often, a supposedly German-language app also displays content in English.
App Function: Does the description match reality?	yes or no		
	yes or no		

XSL•FO
RenderX

Category	Answer	Why is this important?	Meaningful advice
Information: Is the information contained in the App scientifically supported?		• Health information should meet certain quality criteria. In any case, scientific evidence (sources) should be provided for claims.	• Pay attention to source references for information in the app. • Warning: Often, no sources are provided for health information, recommendations, or claims. This means the information could be fabricated. • Pay attention to the timeliness of health information in apps. Information is often outdated and no longer in line with the state of science (eg, in nutrition apps). • Often sources are provided that don‘t directly match the app‘s function.

App Selection

The selected apps to be tested against the long questionnaire is shown in Table 2. Detailed evaluation is not to be published due to possible violation of commercial law.

XSL•FO
RenderX

Table . Health apps that were selected for evaluating the questionnaire.

Category and name of the app	Manufacturer
Nutrition	
Kalorien, Fett & Eiweißzähler	Virtuagym
Lifesum: Gesunde Ernährung	Lifesum
feastr Ernährungsplan Abnehmen	feastr
MyFitnessPal: Kalorien Tracker	MyFitnessPal Inc.
Yazio: Kalorien Zähler & Diät	Yazio
Kalorienzähler von FatSecret	FatSecret
Was ich esse	Bundesanstalt für Landwirtschaft und Ernährung
MySwissFoodPyramid	BA für Lebensmittelsicherheit und Veterinärwesen
Nutrilio: Gesunde Ernährung	Habitics
Foodiary: Rezepte und Diät	Foodiary
Movement	
Dehnung und Flexibilität	Leap Fitness Group
StepsApp Schrittzähler	Stepp-App
7-minute workout	Perigee
Nike Run Club: Laufcoach	Nike, Inc.
Schrittzähler - Health, iStep	Leap Fitness Group
Google Fit: Aktivitätstracker	Google LLC
Nike Training Club: Übungen	Nike, Inc.
Adidas Running: Lauf-App	Adidas Runtastic
SPAR Health Coach	Spar Österreich
Activity Tracker Schrittzähler	Bits & Coffee
Mental health	
Breeze: Mental Health	Basenji Apps
Voidpet Garden: Mental Health	Voidpet
VOS: Wellbeing Plan & Tagebuch	VOS.health
Finch: Self Care Pet	Finch Care Public Benefit Corporation
Mindshine: Mental Health Coach	Mindshine (Greator GmbH)
I am - affirmaties	Monkey Taps LLC
Amaha: anxiety self-care	Amaha Health
Wysa: Anxiety, therapy chatbot	Touchkin
BetterMe: Seelische Gesundheit	BetterMe Limited
Headspace: Medidation & Schlaf	Headspace for Meditation, Mindfulness and Sleep
Symptom tracker	
Ada - Check Deine Gesundheit	Ada Health
Flo Perioden- & Zyklus-Kalender	Flo Health Inc
MyTherapy Tabletten Erinnerung	MyTherapy
Chronic Insights	Chronic Insights Ltd.
Symptom Tracker	CareClinic Tracker & Reminder
Symptom Checker & Medication - FriendsApp Listing	FriendsApp Listing
mySymptoms Food Diary	SkyGazer Labs Ltd
Migraine Buddy	Healint

XSL•FO
RenderX

Category and name of the app	Manufacturer
Symptomate	Infermedica
Bearable Symptom & Mood Tracker	Bearable Ltd

Evaluation Results

This chapter summarizes the evaluation results of the long questionnaire against the selected apps (Table 3).

Table . Evaluation results summarized.

Question items	Nutrition	Exercise	Mental health	Symptom checker
Legally required imprint available	• 4 out of 10 apps	• 4 out of 10 apps	• 1 out of 10 apps	• 5 out of 10 apps
Privacy policies that allow for the disclosure of health data to third parties [24]	• All international operating apps	• All international operating apps • Privacy policies are sometimes unclear and contradict the information provided in the app stores by the manufacturers	• All international operating apps • Consent to the terms of use is implicitly obtained for most apps by using the app.	• User consent is mostly requested via checkboxes.
Addressing the sensitivity of health data	• is not explicitly addressed	• is not explicitly addressed	• is not explicitly addressed	• It is recommended that the app does not substitute the doctor
App tracking	• All apps	• All apps	• All apps	• not available
Target group definitions	• Usually not available	• Usually not available	• Not available	• One app only mentions in the disclaimer that it is intended for users in the United States.
Other	• German apps have good privacy policies and refer to scientific sources • For many apps, a newsletter was simultaneously subscribed to without explicit consent	• All apps reserve the right to collect at least pseudonymized user data • Making unsubstantiated claims (“Yoga for Weight Loss”, “Correction of Bow Legs”, “Relieving Menstrual Cramps through Breathing”, ...)	• Two out of 10 apps refer to scientific backgrounds for their implementation • Many of the apps aim to subscribe to a subscription after or during the registration process, which usually ends in a costly annual subscription of 20 - 130 Euros per year (~20$-130 US$ per year) after a certain free trial period. • Since most apps are financed through subscriptions, many do not use advertising placements.	• Some apps allow for the export of data, which can be presented to the doctor as a report in PDF or CSV format or, in very few cases, directly provided to the doctor via a link to the health data. • Many apps are free. Subscriptions usually cost between 7 and 40 EUR per month. • When entering symptoms, some apps provide a wide range of possible causes. This could lead to overdiagnosis and unnecessary interventions.

User Survey

A quantitative survey with 414 participants was conducted to evaluate the value of health app features for achieving different goals from users’ perspectives, aiming for a balanced gender identity, age, and prior experience with health apps. The analysis focused on the relevance of goals such as support for physical activity (fitness), weight management, nutrition, mental health, and the monitoring of physical and mental symptoms. The findings showed that health apps in the “fitness/activity” and “symptom checker” categories were considered most important, while mental health apps were rated as less important. The ratings were influenced by gender, age, and the presence of chronic conditions, and prior experience with health apps also played a role. Women rated health app functions as more important than men, with a focus on supporting physical health

XSL•FO
RenderX

and achieving weight management goals. Men prioritized physical activity support, while younger participants rated the importance of health app features higher than older ones, particularly for fitness-related goals. Individuals with chronic conditions placed more emphasis on functions promoting physical health and were more likely to pay for health apps.

Discussion

Principal Findings

Despite efforts in initiatives like Label2Enable [10], “So far, no unified concept has emerged to provide users with orientation regarding the quality of health apps. Different approaches with varying underlying criteria, often pursued by commercial entities, are being followed (eg, evaluation instruments for users, certifications, quality seals, test reports)” [12]

Our aim with this work has been to provide users with a unified concept for orientation regarding the quality of health apps, knowing well that the comprehensive instruments offered by Health technology assessments are not usable for users. The goal of this work was to create a questionnaire that could be used to evaluate 10 apps each in the areas of nutrition, exercise, mental health, and symptom checkers. The key points were summarized in the short questionnaire and supplemented with helpful tips and a “Meaningful advice” column, enabling users to assess the quality of an app in a short amount of time.

The user survey revealed that data privacy and security considerations played a minor role for most respondents, contrary to the questions outlined in the ”Short Questionnaire .” This could be due, on one hand, to the general perception of loss of control over one’s data, and on the other hand, the impacts of data privacy and security are not directly and immediately discernible to users. Data leaks with potentially negative consequences for individuals may only become known after a considerable period of time, and their effects may unfold gradually.

Additional education and information on “digital self-determination or self-defense” could help address this issue, enabling users to make clear and informed decisions about whether they prefer data-minimizing apps over those whose primary goal is to collect and market metadata. Simple questionnaires, as presented in this work, can raise awareness regarding data privacy and privacy, especially concerning sensitive health data, which may be perceived as critical in connection with initiatives like Elektronische Gesundheitsakte (ELGA) [25] but less so in the context of apps.

Digital health literacy already plays a role in selecting a suitable digital health applications, not only for users but also for therapists and medical professionals. Choosing a suitable, safe, and effective application from the multitude of digital health apps is a significant challenge. Currently, many users (35% in the user survey) select digital health applications based solely on star ratings in app stores, which have limited reliability [26]. In this context, the use of a questionnaire makes sense, especially regarding expected features and their evidence-based justification.

Apps are often downloaded based on advertising on social media, app stores, and recommendations from family and friends. This indicates that a quality-assured app recommended by a doctor could have significant potential for user adoption. Therefore, it is essential that reviews of these apps are quality-assured and reliable. This applies to both health apps and medical apps, which may target vulnerable groups such as people with chronic illnesses. Considerations regarding the German DiGAs [27] were taken into account throughout the project. Still, the outcome of this work is more geared towards private users rather than representatives of the health system. Nonetheless, we are pleased to note that Digitale Gesundheitsanwendungen (DiGAs - in English: digital health applications) have found a place in the preliminary version of the Austrian eHealth Strategy “M1.7 Creation of a unified process for the evaluation of Digital Health Applications and subsequently their introduction into standard care when demonstrably beneficial” [28]. This also applies to statements from the professional assessment of health apps. “The information offered is too extensive for laypeople”—here, too, a standardized evaluation and recommendation by authorities and health care providers would be advantageous.

In general, it can be noted that Austrian app developers adhere more frequently to regulatory requirements than international developers, since they most often met the evaluation criteria, especially in terms of data privacy and imprint. In this context, we would like to appeal to the responsibility of app providers to make a positive contribution to societal challenges A white paper in line with “Good App Development Practice” or regional requirements from pp stores could introduce even quality-assurance restrictions, particularly regarding regulatory frameworks (eg, explicitly demanding a legally compliant imprint) as collected in the results of this work.

Limitations

The selection of apps to be evaluated based on the established and prioritized criteria proved to be more challenging than initially anticipated. This was partly due to the lack of qualified sources for the previously desired criterion of “most commonly used apps in Austria,” requiring us to resort to an exploratory approach. Furthermore, the apps found as a result could not be fully tested and evaluated by all participants, as even during the implementation period, some apps were no longer available in the app stores for specific operating system versions. Nevertheless, a common denominator for joint evaluation had to be found.

In general, we want to emphasize that a framework for evaluating apps was established and applied, aiming to empower users to independently assess and evaluate both current and new apps.

We deliberately refrained from presenting individual apps or manufacturers in a positive or negative light.

Furthermore, we want to emphasize that our study has not been aimed at replacing the traditional instruments of Health Technology Assessment, as the evaluation criteria were created for the purpose mentioned above. However, this evaluation did not involve searching for clinical studies to demonstrate the

XSL•FO
RenderX

effectiveness of the apps, and the apps were not assigned to risk classes according to Medical Device Regulation [29]. As such, for a comprehensive assessment of a health app, fulfilling the requirements of evidence-based medicine, further important domains such as clinical effectiveness, clinical safety, costs, organizational and ethical aspects need to be assessed.

Conclusions

In a rapidly changing landscape such as the app market, it is therefore even more important to have a solid, easy-to-use tool for evaluating health and medical apps at hand. We are aware that this tool may not be sufficient for evaluation by authorities; however, we would like to note that it can provide a good basis for it. We would like to support further research and projects in this direction.

Acknowledgments

The work was commissioned by the Gesundheitsfonds Steiermark, who contributed massively to defining evaluation criteria with their expert network.

Conflicts of Interest

None declared.

References

1. Jeindl R, Wild C. Framework for reimbursement decisions of digital health technologies (mhealth) and its (retrospective) application on selected examples. : HTA-Projektbericht; 2020 134 URL: https://eprints.aihta.at/1279/1/HTA-Projektbericht_Nr.134.pdf [accessed 2025-06-24] [Medline: 32933277]
2. Kramer U, Borges U, Fischer F, et al. DNVF-memorandum - health and medical apps. Gesundheitswesen 2019 Oct;81(10):e154-e170. [doi: 10.1055/s-0038-1667451] [Medline: 31574556]
3. Messner EM, Terhorst Y, Barke A, et al. The German version of the mobile app rating scale (MARS-G): development and validation study. JMIR Mhealth Uhealth 2020 Mar 27;8(3):e14479. [doi: 10.2196/14479] [Medline: 32217504]
4. Baumel A, Faber K, Mathur N, et al. Enlight: a comprehensive quality and therapeutic potential evaluation tool for mobile and web-based eHealth interventions. J Med Internet Res 2017 Mar 21;19(3):e82. [doi: 10.2196/jmir.7270] [Medline: 28325712]
5. Pernencar C, Aguilar P, Saboia I, et al. Systematic mapping of digital health apps - a methodological proposal based on the World Health Organization classification of interventions. Digit Health 2022;8:20552076221129071. [doi: 10.1177/20552076221129071] [Medline: 36569821]
6. Mehl G, Tamrat T, Labrique A, et al. Classification of digital health interventions v 10. World Health Organization. 2018. URL: https://www.who.int/publications/i/item/WHO-RHR-18.06 [accessed 2025-05-24]
7. Labrique AB, Vasudevan L, Kochi E, et al. mHealth innovations as health system strengthening tools: 12 common applications and a visual framework. Glob Health Sci Pract 2013 Aug;1(2):160-171. [doi: 10.9745/GHSP-D-13-00031] [Medline: 25276529]
8. Agarwal S, LeFevre AE, Lee J, et al. Guidelines for reporting of health interventions using mobile phones: mobile health (mHealth) evidence reporting and assessment (mERA) checklist. BMJ 2016 Mar 17;352:i1174. [doi: 10.1136/bmj.i1174] [Medline: 26988021]
9. Vis C, Bührmann L, Riper H, et al. Health technology assessment frameworks for eHealth: a systematic review. Int J Technol Assess Health Care 2020 Jun;36(3):204-216. [doi: 10.1017/S026646232000015X] [Medline: 32297588]
10. Lable2Enable webpage. Label2Enable. URL: https://label2enable.eu [accessed 2024-02-15]
11. ISO/TS 82304-2:2021 Health softwarePart 2: Health and wellness AppsQuality and reliability. URL: https://www.iso.org/standard/78182.html [accessed 2024-02-15]
12. Kraus-Füreder H. Gesundheits-Apps Grundlagenpapier Unter Besonderer Berücksichtigung Des Aspekts Gesundheitskompetenz: Graz:Institut für Gesundheitsförderung und Prävention; 2018.
13. Chances and Risks of Mobile Health Apps (CHARISMHA): Hannover Medical School; 2016. URL: https://www.bundesgesundheitsministerium.de/fileadmin/Dateien/3_Downloads/A/App-Studie/charismha_abr_v.01.1e-20160606.pdf [accessed 2025-06-25]
14. Kramer U. Health apps: What does certification look like in practice? Diabetes aktuell 2017 Dec;15:344-348. [doi: 10.1055/s-0043-120802]
15. Action Alliance for Patient Safety. Digitalization and patient safety - checklist for the use of health apps. Digital Kompass. 2018. URL: https://www.digital-kompass.de/materialien/digitalisierung-und-patientensicherheit-checkliste-fur-die-nutzung-von-gesundheits-apps [accessed 2025-05-24]
16. Kernebeck S, Kramer U, Redaelli M. Evaluate, but how? – Criteria for the evaluation and reporting of studies on health and medical apps. MVF 2021 Apr 6;14(2/2021):69-75. [doi: 10.24945/MVF.02.21.1866-0533.2300]
17. Kramer U. How good are health apps? Aktuelle Ernährungsmedizin 2017 Jun;42:193-205. [doi: 10.1055/s-0043-109130]

XSL•FO
RenderX

18. Stoyanov SR, Hides L, Kavanagh DJ, et al. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27. [doi: 10.2196/mhealth.3422] [Medline: 25760773]
19. Methodenbox Gesundheitskompetenz: Dachverband der Sozialversicherungsträger; 2020. URL: https://www.sozialversicherung.at/cdscontent/?contentid=10007.843982&portal=svportal [accessed 2025-06-27]
20. Thranberend TK. Health Apps: Important Lever for Patient Empowerment - but Potential Hardly Exploited so Far, 2nd edition: Gütersloh : Bertelsmann Foundation; 2016:1-8.
21. Professional checklist. good health information. The Austrian Platform for Health Literacy (ÖPGK). 2020. URL: https://oepgk.at/schwerpunkte/english-summary/the-austrian-health-literacy-alliance/ [accessed 2025-05-24]
22. Statista [Website in German]. URL: http://de.statista.com [accessed 2024-02-15]
23. data.ai. URL: https://www.data.ai/de [accessed 2024-02-15]
24. Data protection. Unternehmensservice Portal. URL: https://www.usp.gv.at/it-geistiges-eigentum/datenschutz.html [accessed 2024-02-15]
25. ELGA. URL: https://www.elga.gv.at/ [accessed 2024-02-15]
26. Henson P, David G, Albright K, Torous J. Deriving a practical framework for the evaluation of health apps. Lancet Digit Health 2019 Jun;1(2):e52-e54. [doi: 10.1016/S2589-7500(19)30013-5] [Medline: 33323229]
27. Digital health. Bundesinstitut für Arzneimittel und Medizinprodukte. URL: https://www.bfarm.de/DE/Medizinprodukte/Aufgaben/DiGA-und-DiPA/DiGA/_node.html [accessed 2024-02-15]
28. eHealth-Strategie Österreich. V01 zur diskussion im rahmen der bundeszielsteuerung. : BMSGPK Bundesministerium für Soziales, Gesundheit, Pflege und Konsumentenschutz; 2023 URL: https://www.sozialministerium.at/dam/jcr:6f5c5706-b2c4-48a2-8b6a-c7f72f9580e3/240806-eHealth-bf.pdf [accessed 2025-06-27]
29. Regulation - 2017/745 - EN - medical device regulation - EUR-lex (europa.eu). MDR Medical Device Regulation. URL: https://eur-lex.europa.eu/eli/reg/2017/745/oj/eng [accessed 2025-06-24]

Abbreviations

DiGA: Digitale Gesundheitsanwendungen
ELGA: Elektronische Gesundheitsakte
HSC: health system challenges
MARS: Mobile App Rating Scale
mERA: mobile health (mHealth) evidence reporting and assessment
mHealth: mobile health
WHO: World Health Organization

Edited by A Kushniruk; submitted 28.06.24; peer-reviewed by MN Shalaby, R Jeindl, S Erturk; revised version received 25.02.25; accepted 23.03.25; published 01.07.25.

Please cite as:
Rzepka A, Edegger K, Welte S, Bertel D, Mandl A, Schreier G
Self-Determined Health App Evaluation Questionnaire Development: Mixed Methods Study
JMIR Hum Factors 2025;12:e63739
URL: https://humanfactors.jmir.org/2025/1/e63739
doi:10.2196/63739

© Angelika Rzepka, Kurt Edegger, Stefan Welte, Diotima Bertel, Anja Mandl, Günter Schreier. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 1.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

User-Centered Prototype Design of a Health Care Robot for Treating Type 2 Diabetes in the Community Pharmacy: Development and Usability Study

Ching-Ju Chiu[1], PhD; Lin-Chun Hua[1], MS; Jung-Hsien Chiang[2], PhD; Chieh-Ying Chou[1,3], MD, MS

[1]Institute of Gerontology, College of Medicine, National Cheng Kung University, Tainan City, Taiwan

[2]Department of Computer Science and Information Engineering, National Cheng Kung University, Tainan City, Taiwan

[3]Department of Family Medicine, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, No. 138, Sheng Li Road, Tainan City, Taiwan

Corresponding Author:
Chieh-Ying Chou, MD, MS
Institute of Gerontology, College of Medicine, National Cheng Kung University, Tainan City, Taiwan

Abstract

Background: Technology can be an effective tool for providing health services and disease self-management, especially in diabetes care. Technology tools for disease self-management include health-related applications for computers and smartphones as well as the use of robots. To provide a more effective continuity of care and to better understand and facilitate disease management in middle-aged and older adult patients with diabetes, robots can be used to improve the quality of care and supplement community health resources, such as community pharmacies.

Objective: The aim of this study was to develop a health care robot prototype that can be integrated into current community pharmacies.

Methods: Three user-centered approaches were used: (1) review of the literature on technology use among older adults, 2) reference to the seven key diabetes self-care behaviors by the American Association of Diabetes Educators (AADE), and (3) meeting with health care providers in the community. Field investigations and interviews were conducted at community pharmacies and diabetes health education centers to determine the appearance, interface, content, and function of the robot.

Results: The results show that diabetes health care prototype robots can be established through user-centered design. The following important features were revealed: (1) perceived ease of use is considered a friendly operating interface; therefore, we used less than 3 buttons in an interface; (2) minimization of the interface between blue and yellow, which is unfriendly to older adults; (3) the health education mode was the most preferred mode with sound, image, and video presentation; (4) the most predilected functions are health education resources and health records, and that patient data can be easily collected through health education games and dialogue with robots; and (5) touching the screen is the most preferred operation mode.

Conclusions: An evidence-based health care robot can be developed through user-centered design, an approach in which a model that connects medical needs to people with health conditions can be built, thereby facilitating the sustainable development of technology in the diabetes care field.

(JMIR Hum Factors 2025;12:e48226) doi:10.2196/48226

KEYWORDS

robot; diabetes; self-management; middle-aged adult; community pharmacy; older adult; prototype

Introduction

Diabetes is a significant global health issue, with the number of people affected by the disease continually rising [1]. According to the International Diabetes Federation Diabetes Atlas, the global number of people with diabetes reached 537 million in 2021 and is expected to continue rising, reaching an estimated 783 million by 2045 [2]. The most prevalent form of diabetes is type 2 diabetes, which affects more than 90% of all individuals with diabetes [2]. Uncontrolled diabetes over time can lead to severe complications, including cardiovascular diseases, lower-limb amputation, kidney damage (nephropathy), nerve damage (neuropathy), and blindness (retinopathy) [2,3]. Age is one of the risk factors for type 2 diabetes, with its prevalence and mortality rates rising as individuals get older [4,5]. According to statistics from Taiwan's Ministry of Health and Welfare, the prevalence of diabetes increases significantly from the age of 45 years for both men and women [6]. As the duration of diabetes and patient age increase, patients with diabetes face not only a higher incidence of complications but also challenges related to multiple medications and comorbidities [7]. This underscores the importance of focused

XSL•FO
RenderX

health care management for middle-aged and older adult patients with diabetes.

In recent years, technology has emerged as a crucial tool for disease management, including diabetes [8,9]. The demand for digital health care has surged, evidenced by the introduction of more than 90,000 new health-related apps on platforms like the Apple Store and Google Play in 2020 alone, contributing to a total of over 350,000 health care applications available globally [10]. Among these technological advancements in health applications, health care robots are one of the future trends in health technology [11], driven by the ongoing digitization of health care services.

Community pharmacies play a vital role in providing health care services, offering convenient access to medications and health services that support local health care needs. There are currently more than 7400 health insurance pharmacies across Taiwan [12], and a 2016 survey indicated that over 80% of these pharmacies are open for at least 12 hours, with around 40% operating almost year-round [13]. Compared to hospitals, community pharmacies face less pressure from time constraints and space limitations, positioning them as key players in bridging the gap between individuals and larger medical institutions.

Despite the availability of professional medical information in clinical settings, there remains a challenge in ensuring that patients effectively absorb and apply health education and professional advice once they return to their communities. This gap highlights the need for technological tools that support continuous care, especially for middle-aged and older adults managing diabetes. Integrating robots into community health resources, such as community pharmacies, could enhance the quality of care, improve health management, and advance health promotion within the community. The expertise of pharmacists, coupled with the high accessibility of community pharmacies, makes them an ideal setting for such interventions.

However, current technology development for disease management often excludes medical professionals from the design process and lacks real-world testing [14]. Additionally, there is no clear evidence that technology can effectively integrate with community medical resources or bridge the gap between medical institutions and personal care. Therefore, the aim of this study was to design and develop a prototype of a diabetes care robot that meets the needs of middle-aged and older adult patients with diabetes, aligns with the perspective of community pharmacist care, and can be integrated into the current community pharmacy setting.

Methods

Participants

The design and development of the prototype robot in this study were primarily conducted through literature-based theories and expert team meetings [15]. The development of the prototype robot involved the researcher synthesizing the literature and conducting field visits to pharmacies and health education units. The development goals and direction were then jointly decided during team meetings. The development of the prototype robot began in March 2018. The team included 1 expert in technology, 1 expert in gerontology, 2 physicians, 12 to 15 engineers and research assistants, and 1 researcher. To maintain progress and continuously refine the prototype robot, the team held regular meetings 1 - 2 times per month until a more complete interactive functionality was established to enhance the user experience. Furthermore, to gain a more realistic understanding of the feasibility and initial results of the robot's application in community health care, 30 middle-aged and older adult patients with diabetes, along with 10 community pharmacists, were also included to provide user feedback.

Ethical Considerations

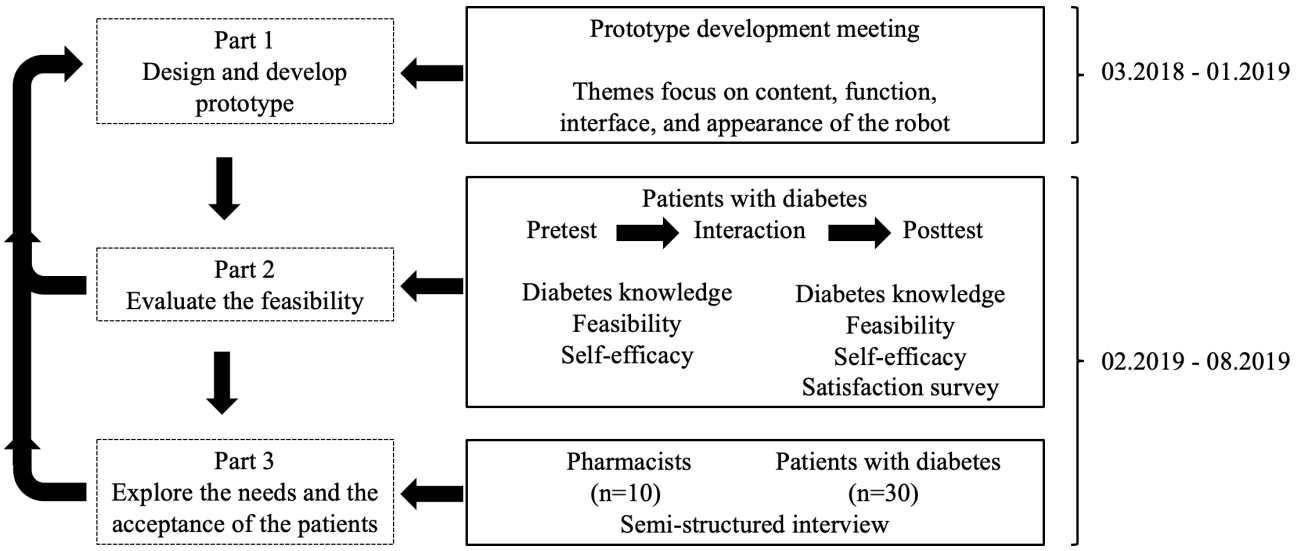
This study was performed in accordance with the relevant guidelines and regulations, including the Declaration of Helsinki and was approved by the Institution Review Board (IRB) of National Cheng Kung University Hospital in Taiwan (No. A-ER-105 - 509). All the participants provided signed informed consent, and all study data were deidentified to protect the identity of participants. Participants did not receive any financial or material compensation for their participation in this study.

Procedures

The research design framework for the prototype robot comprised three steps: using an evidence-based literature review and team meeting to design and develop the robot prototype, evaluating feasibility through collecting pre- and postassessment questionnaires, and exploring needs and acceptance through interviews with the target population, as illustrated in Figure 1.

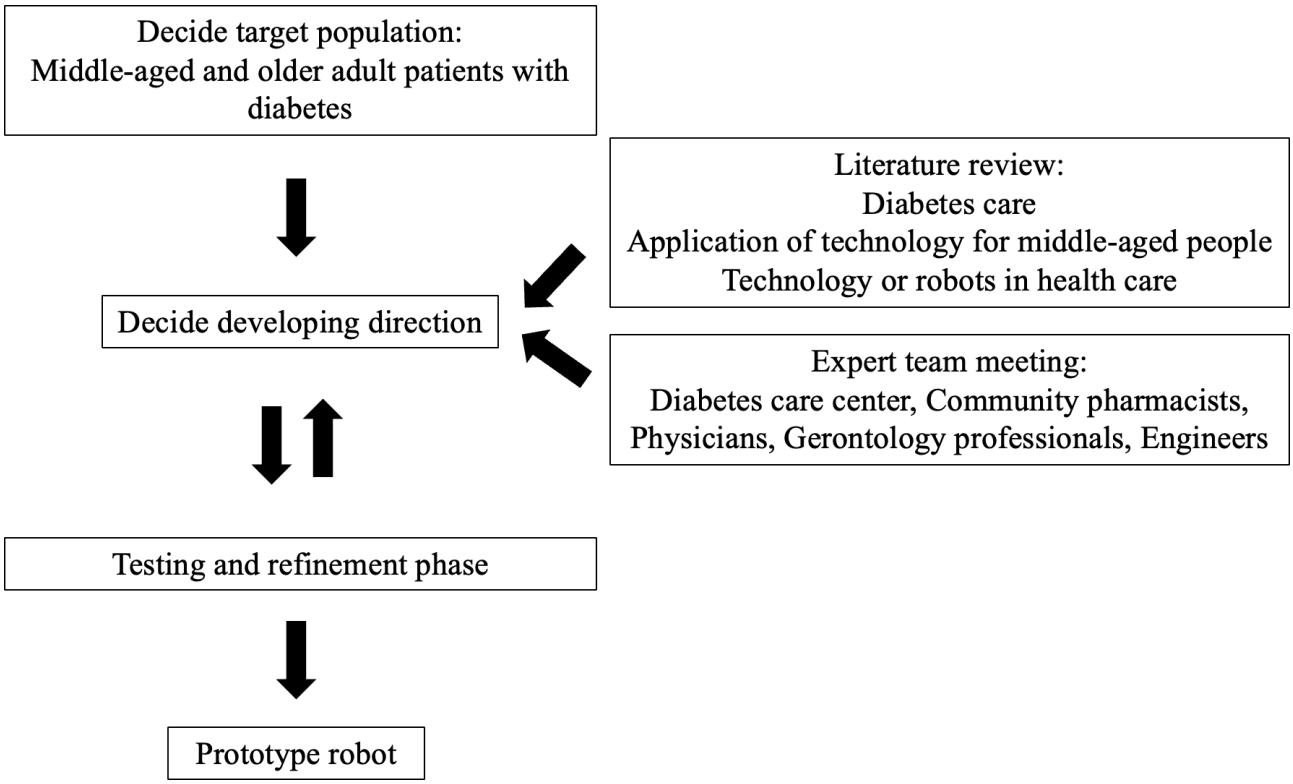
XSL•FO
RenderX

Figure 1. The research design framework for the prototype robot used in community-based diabetes care.



Although robotics is an emerging technology, there is a lack of understanding about the applications of robotics. To gain a better understanding of user needs, a high-fidelity prototype can be used to provide users with real-world experience and stimulate their needs, thus enabling them to provide more in-depth insights for future development and design [16]. Therefore, the first part of the study was to design and build a prototype community-based diabetes care robot to facilitate user experience, as illustrated in the detailed flowchart shown in Figure 2.

Figure 2. Process showing how to build a prototype community-based diabetes care robot.



The second and third parts were dedicated to testing and evaluating the prototype robot. In the second part, it took 15 - 20 min for the patients with diabetes to interact with the robot. Before and after the interaction, questionnaires including a diabetes knowledge test, self-efficacy for diabetes, and feasibility of use of the robot were administered. In the third section, semistructured qualitative interviews were conducted with middle-aged and older adult patients with diabetes and community pharmacists about their experiences and visions for using the robots. The detailed methods and results of the pre- and postassessment questionnaires, including a diabetes knowledge test, self-efficacy for diabetes, and the feasibility of using the robot in patients with diabetes, as well as the in-depth interviews with both pharmacists and patients with diabetes, were addressed in our previously published study [17]. This prior study demonstrated that robot interaction significantly improved patients' diabetes knowledge and the perceived feasibility of its use in pharmacies. Both patients and

XSL•FO
RenderX

pharmacists gave positive feedback, with patients valuing self-directed learning, comfortable interaction, and vivid engagement, while pharmacists highlighted real-world applicability and new service potential. These findings offer foundational insights that inform the design and objectives of our current research.

Design and Construction of a Prototype Community-Based Diabetes Care Robot

As the goal of this research is to develop a prototype for a community-based diabetes care robot and to gather and understand the perspectives and needs of patients with diabetes and community health care providers, the initial team meetings decided to use an existing robot with foundational functions. This robot was optimized and gradually modified to develop interactive features relevant to diabetes care, aiming for efficiency and smooth user operation. The robot developed in this study uses Zenbo (ASUS), a smart home product from ASUS. The appearance and specifications are shown in Figure 3 and Textbox 1, respectively.

XSL•FO
RenderX

Figure 3. Appearance of the Zenbo robot.



XSL•FO
RenderX

Textbox 1. Zenbo robot specifications and functions; source: ASUS.

Specifications:

- External dimensions: 37 × 37 × 62 cm (length × width × height)
- Weight: 10 kg
- Operating system: Android
- Display screen: 10.1" LCD screen
- Language skills: Chinese
- Functions: Listening, speaking, seeing, emotional expression, playing music and videos, movement, internet connection, and learning

LCD: liquid crystal display

The concept for the prototype robot is primarily derived from research and theory, which must align with the needs of the target population, namely middle-aged and older adult patients with diabetes. Therefore, this study focused on theories and care indicators related to diabetes care [18], collected and organized the design and limitations of technology applications for middle-aged individuals [19,20], and examined the development and application of current technology or robots in health care, as well as the connection between the use of technology and health in middle-aged people and the mode for delivering health services [21], to establish the development objectives [22,23]. Three user-centered approaches were used in the design and construction of a prototype community-based diabetes care robot: (1) a review of the literature on technology use among older adults, (2) reference to the seven key diabetes self-care behaviors by the American Association of Diabetes Educators (AADE), and (3) meetings with health care providers in the community. The literature review search strategy of this study involved collecting studies related to the application of robots in the elderly population [24-27]. Field investigations and interviews were conducted at community pharmacies and diabetes health education centers to determine the design of the robot. Because patients with diabetes go to community pharmacies to take their medications, and in addition to providing pharmaceutical services, community pharmacies can also offer related information on disease care. Therefore, in addition to collecting service experiences from the community pharmacists, the choice of a community pharmacy as the testing site for the prototype is driven by the fact that these pharmacies provide a secure, private, and professionally consultative environment.

The design framework of the community-based diabetes care robot was devised to collect and compile data in four main areas: appearance, interface, content, and function. For appearance, we collected data about the appearance and size of the robot. For interface, data were collected about the color, text size, arrangement, and other screen design. For content, the system was responsible for building the information that needed to be conveyed, with the user's needs as the primary consideration. For function, we collected data about the presentation of the interactive process and information transfer.

Once the direction of development has been established, a prototype robot is initially built that can be used in practice. The development of technology does not only rely on past and proven experience, but also must be tailored to the needs of real users. Therefore, the prototypes need to interact with the users in real life and give them a real experience in accordance with the proposed development directions. It is only after this that users can give practical advice.

Results

Using a literature search and practical experience, this study focused on data collection in four main areas: appearance, interface, content, and function of robot design. The development directions and options for prototype robots are listed in Table 1.

XSL•FO
RenderX

Table . Development directions and options for prototype robots.

	Literature and theory	Interviews and team discussions	Results and design directions
Appearance	• Type: The humanoid type is less well accepted. The doll type, on the other hand, is seen as a toy for children [22]. Prototype robots should avoid being too complex or trendy. However, if the prototype lacks a humanoid feel, it will not be favored by senior citizens [23]. • Application mode: The focus of the robot varies between different appearances. Appearance will affect whether it is presented in a functional or interactive way [28,29].	• Size: The prototype robot must fit into the space of the pharmacy. • Height: It should be at a height at which middle-aged to older adults can sit and see it horizontally.	• It must be confirmed that this model is not complicated and is expected to be well-accepted because it is a non-humanoid model. However, it may be seen as a toy because of the cute non-humanoid appearance. The model does not have hands and feet to operate and therefore operates mainly in a conversational manner. It must be confirmed that the angle and height of the interaction mode (eg, the height of the table or chair on which the robot can be placed) can be adjusted.
Interface	• Colors: Blue and yellow are less recognizable [19]. Yellow and green combinations, and red and purple combinations, are prone to recognition errors [20]. • Operation: Per the Technology and Acceptance Model [30], the interface and operation must be simple because it is used by older adults. The font should be large and clear. The interface should have a maximum of four buttons. All other interfaces have two buttons for simple selection [31].	• The interface should be simple. A single screen does not need much emphasis. Text should be enlarged.	• The color scheme of the interface is highly contrasting to avoid confusion over similar color schemes. The use of blue and yellow colors is reduced and the number of buttons on the interface is 2 - 3. The content is centered, and apart from the main screen, the content is mostly presented as a single theme. The font size ranges from a minimum of 30sp to a maximum of 50sp (sp being the text unit in the Android system).
Content	• Seven key diabetes self-care behaviors by the American Association of Diabetes Educators [18]	• Applications addressing various everyday challenges and health education videos on common diseases in community pharmacies.	• The content of the system is based on the seven diabetes self-care behaviors, with games and interactive dialogues to present a variety of health education methods. To make the content more relevant to the public, common disease symptoms and complications of diabetes are presented.
Function	• Operation: Technology and Acceptance Model [30]. • Health education: Multisensory perception stimulates learning [21], with a core health education focus each time. • Technology applications: The applications of robots in health management can include disease management, monitoring of physiological data, and interactive health education [26,32,33]. The use of technology in diabetes health management focuses on recording and reminding, as well as providing information and entertainment [34-36].	• There are games, friendly greetings, and flexible interaction times.	• The health education mode is presented with sound, images, and video, avoiding a focus on the use of a single sense. The functionality of the current model will be considered and the information on functionality available in the literature will be further researched.

XSL•FO
RenderX

Appearance

Although the robot's appearance was determined based on the available resources, the chosen model features a cute, non-humanoid design to better align with the preferences of older adults and enhance user acceptance. Although the robot's height is a limiting factor, placing the robot on elevated surfaces, such as tables or chairs, can correct height-related issues, making it easier for users to operate.

Interface

As the interface is critical to the overall smoothness of the robot, the presentation of the interactive interface must meet the needs of middle-aged and older adults in terms of the ease of use of technology products and text reading as much as possible. In addition, according to the Technology and Acceptance Model [30], the perceived ease of use should be considered as a whole and should take into account the user-friendliness of the interface so that the user can experience the clearest and easiest process.

In the interface development, to make it easiest for older adults to operate, each screen will display a maximum of two main options, with options clearly framed to reduce the likelihood of users being unsure of their choices. Information and selection buttons will be centrally placed and maximized on the screen to help users quickly focus on key points. The main screen will feature simple text and image displays. To avoid causing difficulties in recognition for older users, the background will be a single color, and the overall interface will use no more than two main color tones. The main screen and option areas will predominantly use a black-on-white color scheme, with high contrast to make text easier to read.

Content

The main content of the development includes the seven key diabetes self-care behaviors according to the AADE standards. The overall main content and functions presented by the prototype robot are detailed in Table 2 and aligned with the seven AADE diabetes care indicators.

The main functions of the system are health education resources and health records. Health education resources are presented through health education quizzes, pamphlets, and videos. The health records section includes health values, discomfort, and nutrition. The concept of this record is to help patients with diabetes manage their records, including basic physiological data such as blood pressure, blood glucose, and blood lipids. The relevant data are entered and the corresponding standard values are provided. The focus of the "Discomfort" program is to understand and record the various conditions in the patients with diabetes and to compare the complications associated with diabetes with the common problems of the family. By reminding and advising patients to be more alert to their health conditions, they can gain a better understanding of their health conditions. The results are recorded for follow-up. All interactions can be fed back to the user in the form of paper or communication software messages.

Table . Functional development and presentation of prototype robots.

Theme	Subtheme	Theme content	Presented at the American Association of Diabetes Educators
Health education resources	Health education test	Questions focus on the seven major areas of diabetes care and common diabetes health problems	Healthy eating, regular exercise, blood glucose monitoring, proper medication use, problem solving, health adaptation, and risk reduction
	Health education leaflet	Share and recommend links to help filter relevant content	Healthy eating, regular exercise, blood glucose monitoring, proper medication use, problem solving, health adaptation, and risk reduction
	Health education videos	Share and recommend links to help filter relevant content	Healthy eating, regular exercise, blood glucose monitoring, proper medication use, problem solving, health adaptation, and risk reduction
Health records	Health values	Alert and record blood pressure, blood glucose and lipids, and provide standard values	Blood glucose monitoring
	Discomfort	Presented in a lifestyle format to remind patients of complications and other problems they may encounter and link them to their disease conditions	Problem solving, health adaptations, and risk reduction
	Nutrition	Ask each other about their dietary status in a simple conversation to help them develop a balanced diet	Healthy eating

Function

To establish the overall function, we should first understand the current mode of application of technology in disease and health management, including the application of health technology for older adults, diabetes management technology tools, and the application of robots in health management, and consider the current operation fluency and feasibility of using robots.

XSL•FO
RenderX

According to the current functional status of the robot, the main functions are primarily simple question-and-answer interactions with pictures and images, so the main functions are presented in a health education manner. A game was used to transfer knowledge on diabetes-related care content, and a thematic health education video was shown as feedback to the patients according to their wrong answers.

The main objective of the development of this prototype robot is to enable patients with diabetes to experience actual interaction with the robot and to understand the possible interaction patterns, in order to understand the experience and perspective of the user. During the development process, the prototype robot is adapted to meet the needs of the community and medical personnel based on the health education needs of the Diabetes Health Centre and the feasible applications of the community pharmacy. In view of the time and cost of development as well as the operation of the technology, the functions were not expanded extensively and were mainly presented in the form of a basic framework. After the patient with diabetes or community health care worker has confirmed the suitability of the directions and content, the functions will be expanded and built. To allow each participant to experience the same features, the most stable and fully developed part (ie, the health education test, a game with the robot, that focuses on diabetes-related knowledge) is used as the main axis for participant interaction. The rest of the developed functions are presented as introductions and simple demonstrations.

Discussion

Principal Findings

This study demonstrated that an evidence-based health care robot can be developed through user-centered design, facilitating the sustainable advancement of technology in the field of diabetes care. The use of technology in patient management can be beneficial to patients. However, the applications of technology that are readily available on the market are not always effective. While the current technology interventions are mostly aimed at behavioral change, they are rarely supported by therapeutic guidelines and theories [35,37], making the development process too subjective or deviating from reality. The core elements of technology development must therefore be user-centered. Thus, based on the technology being developed and designed for diabetes disease self-management, Goyal et al proposed four stages of user-centered development [38]. The first is the so-called development process, which must be evidence-based and analyze possible barriers and problems and must be culturally and environmentally appropriate. In the second stage, the feasibility of the system must be understood. The third stage involves further evaluation of effectiveness. The fourth stage requires more long-term monitoring. At the same time, as the four stages are in a continuous cycle and interact with each other, it is necessary to revise and review them repeatedly to build a technology tool that meets the needs of users [38]. Furthermore, personal health information must also be protected. While this prototype requires login, ongoing efforts should be made to enhance and ensure information security.

Comparison With Prior Work

In previous studies of diabetes care robots, the main focus has been on younger age groups of patients with type 1 diabetes [32,33,39], and less on middle-aged patients with type 2 diabetes. Even in the case of older adult care, the functional applications of robots are mainly companionship-based or focus on social functions for older adults. There should be less involvement in behavioral change, but this is also an important part of health messaging. Therefore, this study uses the seven key diabetes self-care behaviors by the AADE to consider the basics of healthy eating, regular exercise, blood glucose monitoring, and proper medication use as core health-enhancing skills, and to help patients with diabetes better understand their disease and their health. These seven self-care behaviors are interlinked with diabetes disease self-management and not only directly contribute to disease management (eg, reminders and records), but also enhance the overall quality of care for the disease by building up the correct concepts of patients with diabetes. This content is presented as a basic health education leaflet and a health education video to convey information. The health education resources are sourced from relevant health education websites and reference links from major hospitals and institutes, and can be categorized to help patients quickly search and link to relevant functions. In terms of the functions, we should first understand the current mode of application of technology in disease and health management, including the application of health technology for older adults, diabetes management technology tools, and the application of robots in health management, and consider the current operation fluency and feasibility of using robots. The robot itself has basic functions such as playing videos and documents. Based on the current functionality of the robot, the main focus of further development is on the simple interaction with pictures and images. The main function is therefore presented in the form of health education. To make the health education interaction more efficient and interesting, a health education game was used to transfer knowledge about diabetes-related care content, and at the end of one stage of the game, a health education video on the relevant topic was selected and shown as feedback based on the questions that the patient answered incorrectly, to deepen the patient's understanding of the concept.

Past studies have also supported the idea that a multimedia approach to health education can help patients understand the content of health education [40,41]. The health education model is more stimulating, as it delivers information to patients in a variety of ways, such as videos and animations, which in turn produces better health education results. This is one of the main advantages of robots for the implementation of health education [40]. The robot's rich expressions and gestures, as well as its voice-to-voice interactions, are an advantage in teaching [41]. The prototype utilized in this study is a nonhumanoid robot. While humanoid robot interactions are more human-like, their realistic appearance can often induce fear, leading to lower acceptance. In contrast, nonhumanoid robots, though cute, might be perceived more like toys, resembling a game. The advantage lies in a more engaging interaction, but the downside is that they may be seen merely as playthings, potentially overshadowing the educational information they convey.

XSL•FO
RenderX

In the delivery of health services by robots, there is a need for a thorough consideration of the technological side, the professional side, and the patient side. Therefore, in addition to the professional development in the field of technology, the initial development should also identify and create a suitable development direction from the perspective of medical professionals, and build a prototype robot. Later on, the prototype robot can be used to enable patients with diabetes to experience the interaction with the robot, and through this interaction, to gather more in-depth information on patients' thoughts and experiences on the application of technology. In the context of the rapid expansion of technology, it is important to have more detailed information on the user experience so that subsequent developers can have a clearer direction, and at the same time, have the most relevant technological tools for the core users (ie, older adult patients).

Strengths

The main advantage of the robot developed in this study is that the prototype was developed using data from professional team meetings, literature, real-world data, and experience interviews, as well as the participation of people from the relevant professional fields. This comprehensive approach ensured the prototype is robust. In addition, the development of the prototype robot was not a one-off exercise. The aim was to create a prototype that would enable actual interaction with patients with diabetes, not just from the developer's perspective. Each time further progress is made, it will be compared and revised with clinical practitioners to create a more realistic diabetes care robot that can actually interact with patients with diabetes in the real world.

Limitations

The literature reviewed in this study primarily comes from Western countries, which may not fully correspond to the disease characteristics of East Asian populations. Differences in illness, age, and cultural backgrounds may affect users' perspectives. To mitigate this limitation, future studies could be conducted in Western countries to validate the findings and evaluate their applicability in different contexts. In addition, the testing site for this study was a community pharmacy in an urban area, so it does not provide insights into how elderly individuals with different lifestyles or from different regions perceive and use the robot. Consequently, the findings may not be fully applicable to all health care settings. Future research could extend to rural community pharmacies or other community health care settings, including community hospitals, clinics, and chain pharmacies, to test the application of health care robots. Moreover, the size and functionality of the robots used in this study, along with the critical role of the internet in their operation, may limit their capabilities. For example, uninterrupted and smooth conversations rely on a stable network connection. Interruptions or network issues can hinder the robot's performance. Efforts were made to ensure a stable network environment during testing, but future studies should consider testing in various network conditions to better understand the impact of connectivity on robot functionality.

Conclusion

The prototype diabetes care robot was developed using data from the literature, professional team meetings, and practical interviews, with the aim of creating a more evidence-based health care robot that can actually give patients with diabetes a realistic experience for their disease care and enable them to give clearer and more specific feedback for the subsequent development and application of technology in disease care.

Acknowledgments

This work was supported by a grant from the Taiwan Ministry of Science and Technology (MOST 108-2634-F-006 -006-). The Taiwan Ministry of Science and Technology had no role in the design of the study, data collection, data analysis, data interpretation, or writing of the report.

Authors' Contributions

Study concept and design: CJC. Acquisition of data: CJC, LCH. Analysis and interpretation of data: CJC, LCH. Drafting of the manuscript: CJC, LCH. Critical revision of the manuscript for important intellectual content: CJC, LCH, CYC. All authors have read and approved the manuscript.

References

1. GBD 2021 Diabetes Collaborators. Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: a systematic analysis for the Global Burden of Disease Study 2021. Lancet 2023 Jul 15;402(10397):203-234. [doi: 10.1016/S0140-6736(23)01301-6] [Medline: 37356446]
2. Magliano DJ, Boyko EJ, IDF Diabetes Atlas 10th Edition Scientific Committee. IDF Diabetes Atlas, 10th edition: International Diabetes Federation; 2021. URL: https://diabetesatlas.org/atlas/tenth-edition/
3. Tomic D, Shaw JE, Magliano DJ. The burden and risks of emerging complications of diabetes mellitus. Nat Rev Endocrinol 2022 Sep;18(9):525-539. [doi: 10.1038/s41574-022-00690-7] [Medline: 35668219]
4. World Health Organization. Global Report on Diabetes: World Health Organization; 2016.

XSL•FO
RenderX

5. Khan MAB, Hashim MJ, King JK, Govender RD, Mustafa H, Al Kaabi J. Epidemiology of type 2 diabetes - global burden of disease and forecasted trends. J Epidemiol Glob Health 2020 Mar;10(1):107-111. [doi: 10.2991/jegh.k.191028.001] [Medline: 32175717]
6. Department of Statistics MoHaW, Rebublic of China (Taiwan). Prevalence of high blood pressure, high blood sugar, and high cholesterol among individuals aged 20 and above [article in chinese]. 2025 URL: https://dep.mohw.gov.tw/DOS/cp-5339-59467-113.html#_2.%E7%96%BE%E7%97%85%E7%BD%B9%E6%82%A3
7. American Diabetes Association Professional Practice Committee. 13. Older adults: Standards of care in diabetes-2024. Diabetes Care 2024 Jan 1;47(Suppl 1):S244-S257. [doi: 10.2337/dc24-S013] [Medline: 38078580]
8. Goyal S, Cafazzo JA. Mobile phone health apps for diabetes management: current evidence and future developments. QJM 2013 Dec;106(12):1067-1069. [doi: 10.1093/qjmed/hct203] [Medline: 24106313]
9. Sneha S, Thalla S, Rischie I, Shahriar H. Health internet technology for chronic conditions: review of diabetes management apps. JMIR Diabetes 2021 Aug 31;6(3):e17431. [doi: 10.2196/17431] [Medline: 34463627]
10. IQVIA. Digital Health Trends 2021. 2021. URL: https://www.iqvia.com/insights/the-iqvia-institute/reports-and-publications/reports/digital-health-trends-2021
11. Huang R, Li H, Suomi R, Li C, Peltoniemi T. Intelligent physical robots in health care: systematic literature review. J Med Internet Res 2023 Jan 18;25:e39786. [doi: 10.2196/39786] [Medline: 36652280]
12. Food and Drug Administration MoHaW, Rebublic of China (Taiwan). Basic information of pharmacy [Article in Chinese]. 2024. URL: https://data.gov.tw/dataset/6134
13. Food and Drug Administration MoHaW, Rebublic of China (Taiwan). Community pharmacies: professional hubs where medicine keeps us healthy [article in chinese]. 2017 URL: https://www.mohw.gov.tw/cp-3558-37574-1.html
14. Müller P, Jahn P. Cocreative development of robotic interaction systems for health care: scoping review. JMIR Hum Factors 2024 Sep 12;11:e58046. [doi: 10.2196/58046] [Medline: 39264334]
15. Salmani Nodooshan H, Choobineh A, Razeghi M, Shahnazar Nezhad Khales T. Designing, prototype making and evaluating a mechanical aid device for patient transfer between bed and stretcher. Int J Occup Saf Ergon 2017 Dec;23(4):491-500. [doi: 10.1080/10803548.2016.1274161] [Medline: 28041534]
16. Anastassova M, Mégard C, Burkhardt JM. Prototype evaluation and user-needs analysis in the early design of emerging technologies. Presented at: Human-Computer Interaction Interaction Design and Usability, 12th International Conference, HCI International 2007; Jul 22-27, 2007; Beijing, China.
17. Chiu CJ, Hua LC, Chou CY, Chiang JH. Robot-enhanced diabetes care for middle-aged and older adults living with diabetes in the community: a small sample size mixed-method evaluation. PLoS ONE 2022;17(4):e0265384. [doi: 10.1371/journal.pone.0265384] [Medline: 35427359]
18. American Association of Diabetes Educators. An effective model of diabetes care and education: revising the AADE7 self-care behaviors ®. Diabetes Educ 2020 Apr;46(2):139-160. [doi: 10.1177/0145721719894903]
19. Klein R. Age-related eye disease, visual impairment, and driving in the elderly. Hum Factors 1991 Oct;33(5):521-525. [doi: 10.1177/001872089103300504] [Medline: 1769672]
20. Lee CF. Legibility of text-color combinations of PDA for the elderly. J Des 2005;10(2):1-12. [doi: 10.6381/JD.200506.0001]
21. Shams L, Seitz AR. Benefits of multisensory learning. Trends Cogn Sci 2008 Nov;12(11):411-417. [doi: 10.1016/j.tics.2008.07.006] [Medline: 18805039]
22. Wu YH, Fassert C, Rigaud AS. Designing robots for the elderly: appearance issue and beyond. Arch Gerontol Geriatr 2012;54(1):121-126. [doi: 10.1016/j.archger.2011.02.003] [Medline: 21349593]
23. Bedaf S, Marti P, De Witte L. What are the preferred characteristics of A service robot for the elderly? A multi-country focus group study with older adults and caregivers. Assist Technol 2019;31(3):147-157. [doi: 10.1080/10400435.2017.1402390] [Medline: 29125807]
24. Wang RH, Sudhama A, Begum M, Huq R, Mihailidis A. Robots to assist daily activities: views of older adults with Alzheimer's disease and their caregivers. Int Psychogeriatr 2017 Jan;29(1):67-79. [doi: 10.1017/S1041610216001435] [Medline: 27660047]
25. Liang A, Piroth I, Robinson H, et al. A pilot randomized trial of a companion robot for people with dementia living in the community. J Am Med Dir Assoc 2017 Oct 1;18(10):871-878. [doi: 10.1016/j.jamda.2017.05.019] [Medline: 28668664]
26. Broadbent E, Garrett J, Jepsen N, et al. Using robots at home to support patients with chronic obstructive pulmonary disease: pilot randomized controlled trial. J Med Internet Res 2018 Feb 13;20(2):e45. [doi: 10.2196/jmir.8640] [Medline: 29439942]
27. Petersen S, Houston S, Qin H, Tague C, Studley J. The utilization of robotic pets in dementia care. J Alzheimers Dis 2017;55(2):569-574. [doi: 10.3233/JAD-160703] [Medline: 27716673]
28. Beuscher LM, Fan J, Sarkar N, et al. Socially assistive robots: measuring older adults' perceptions. J Gerontol Nurs 2017 Dec 1;43(12):35-43. [doi: 10.3928/00989134-20170707-04] [Medline: 28700074]
29. Sabanovic S, Reeder S, Kechavarzi B. Designing robots in the wild: in situ prototype evaluation for a break management robot. Journal of Human-Robot Interaction 2014;3(1):70. [doi: 10.5898/JHRI.3.1.Sabanovic]
30. Szajna B. Empirical evaluation of the revised technology acceptance model. Manage Sci 1996 Jan;42(1):85-92. [doi: 10.1287/mnsc.42.1.85]

XSL•FO
RenderX

31. Jayawardena C, Kuo IH, Unger U, et al. Deployment of a service robot to help older people. Presented at: 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2010); Oct 18-22, 2010; Taipei, Taiwan. [doi: 10.1109/IROS.2010.5649910]
32. Al-Taee MA, Kapoor R, Garrett C, Choudhary P. Acceptability of robot assistant in management of type 1 diabetes in children. Diabetes Technol Ther 2016 Sep;18(9):551-554. [doi: 10.1089/dia.2015.0428] [Medline: 27479198]
33. Blanson Henkemans OA, Bierman BPB, Janssen J, et al. Using a robot to personalise health education for children with diabetes type 1: a pilot study. Patient Educ Couns 2013 Aug;92(2):174-181. [doi: 10.1016/j.pec.2013.04.012] [Medline: 23684366]
34. Greenwood DA, Gee PM, Fatkin KJ, Peeples M. A systematic review of reviews evaluating technology-enabled diabetes self-management education and support. J Diabetes Sci Technol 2017 Sep;11(5):1015-1027. [doi: 10.1177/1932296817713506] [Medline: 28560898]
35. Chomutare T, Fernandez-Luque L, Arsand E, Hartvigsen G. Features of mobile diabetes applications: review of the literature and analysis of current applications compared against evidence-based guidelines. J Med Internet Res 2011 Sep 22;13(3):e65. [doi: 10.2196/jmir.1874] [Medline: 21979293]
36. Holtz B, Lauckner C. Diabetes management via mobile phones: a systematic review. Telemed J E Health 2012 Apr;18(3):175-184. [doi: 10.1089/tmj.2011.0119] [Medline: 22356525]
37. Becker S, Miron-Shatz T, Schumacher N, Krocza J, Diamantidis C, Albrecht UV. mHealth 2.0: experiences, possibilities, and perspectives. JMIR Mhealth Uhealth 2014 May 16;2(2):e24. [doi: 10.2196/mhealth.3328] [Medline: 25099752]
38. Goyal S, Morita P, Lewis GF, Yu C, Seto E, Cafazzo JA. The systematic design of a behavioural mobile health application for the self-management of type 2 diabetes. Can J Diabetes 2016 Feb;40(1):95-104. [doi: 10.1016/j.jcjd.2015.06.007] [Medline: 26455762]
39. Sinoo C, van der Pal S, Blanson Henkemans OA, et al. Friendship with a robot: Children's perception of similarity between a robot's physical and virtual embodiment that supports diabetes self-management. Patient Educ Couns 2018 Jul;101(7):1248-1255. [doi: 10.1016/j.pec.2018.02.008] [Medline: 29548599]
40. Pei D, Liang B, Du W, et al. Multimedia patient education to assist oral impression taking during dental treatment: A pilot study. Int J Med Inform 2017 Jun;102:150-155. [doi: 10.1016/j.ijmedinf.2017.03.016] [Medline: 28495343]
41. Causo A, Vo GT, Chen IM, Yeo SH. Design of robots used as education companion and tutor. In: Robot Mechatron (2015): Springer International Publishing; 2016.

Abbreviations

AADE: American Association of Diabetes Educators

Edited by A Kushniruk; submitted 22.04.23; peer-reviewed by D Dinh, PJ Long Guo, YL Leung; revised version received 14.01.25; accepted 14.01.25; published 18.03.25.

Please cite as:
Chiu CJ, Hua LC, Chiang JH, Chou CY
User-Centered Prototype Design of a Health Care Robot for Treating Type 2 Diabetes in the Community Pharmacy: Development and Usability Study
JMIR Hum Factors 2025;12:e48226
URL: https://humanfactors.jmir.org/2025/1/e48226
doi:10.2196/48226

© Ching-Ju Chiu, Lin-Chun Hua, Jung-Hsien Chiang, Chieh-Ying Chou. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 18.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Telemedicine Booths for Screening Cardiovascular Risk Factors: Prospective Multicenter Study

Mélanie Decambron[1], MD; Christine Tchikladze Merand[2], MSc

[1]Polyclinique Vauban, Elsan, Valenciennes, France

[2]Elsan, 58 bis rue de la Boétie, Paris, France

Corresponding Author:
Christine Tchikladze Merand, MSc
Elsan, 58 bis rue de la Boétie, Paris, France

Abstract

Background: Cardiovascular risk factors such as hypertension often remain undetected and untreated. This was particularly problematic during the COVID-19 pandemic when there were fewer in-person medical consultations.

Objective: This study aimed to determine whether health screening using a telemedicine booth would have an impact on people's medical care during the COVID-19 pandemic.

Methods: Health screening was run using a telemedicine booth (the consult station) that was placed in three different vaccination centers in northern France between July 2021 and September 2021. Participants followed a series of instructions to obtain various measures, including their blood pressure, BMI, and heart rate. If any measures were found to be outside of the normal range, participants were advised to consult a doctor. After 3 months, the participants with abnormal readings were contacted by telephone and were asked a series of standardized questions. The primary outcome was the percentage of respondents who reported that they had consulted a doctor since the health check.

Results: Approximately 6000 people attended the 3 vaccination centers over the study period. Of these, around 2500 used the consult station. A total of 1333 participants (53.3%) were found to have abnormal readings, which mostly concerned their blood pressure, heart rate, or BMI. There were 638 participants who responded to the follow-up call, and 234 of these (37%) reported that they had consulted a doctor since the health check. However, 158 of the 638 respondents (24.8%) reported that they would have consulted a doctor even without the screening.

Conclusions: We succeeded in screening large numbers of people for cardiovascular risk factors during the COVID-19 pandemic by using a telemedicine booth. Although relatively few follow-up call respondents reported that they went on to consult a physician, the screening would nevertheless have raised people's awareness of their cardiovascular risk factors.

(JMIR Hum Factors 2025;12:e57032) doi:10.2196/57032

KEYWORDS

cardiovascular disease risk factors; cardiovascular disease; CVD; cardiology; cardiac; cardiologists; health check; hypertension; HTN; hypertensive; high blood pressure; blood pressure; obesity; screening; health screening; telemedicine; telehealth; virtual care; virtual health; virtual medicine; COVID-19; SARS-CoV-2; Coronavirus; respiratory; infectious; pulmonary; pandemic

Introduction

Cardiovascular disease (CVD) is the leading cause of death worldwide [1]. In France, the disease affects around 7.9% of the population and causes around 140,000 deaths each year, which is only exceeded by deaths caused by cancer [2,3]. The health care costs are substantial, representing around 10% of the country's health care reimbursements [4]. The burden of CVD is likely to increase over the coming years due to an aging population and is thus a major cause for concern [4].

There are several well-known modifiable risk factors for CVD, including obesity, high blood pressure, and physical inactivity [5]. The most significant of these is hypertension because it accounts for the largest number of CVD deaths [5,6]. In France, this affects around 22% of the population, but over a third are unaware that they have the condition because there are usually no symptoms [4,7]. As a result, many people do not benefit from interventions, such as antihypertensive medication, that can effectively lower their blood pressure, thus reducing the risk of CVD [8]. Obesity is another important modifiable risk factor for CVD, which affects 21.6% of the adult population in France [4]. Although effective treatment options are available, such as lifestyle modifications, antiobesity medication, and bariatric surgery, people with obesity typically believe that it is their responsibility to lose weight and so often wait many years before seeking medical help [9-11].

Screening for CVD risk factors is of much importance for public health, as it enables risk factors such as hypertension to be

XSL•FO
RenderX

detected and treated, which can prevent CVD. Most screening takes place during medical consultations with a general practitioner (GP), where blood pressure and weight are regularly monitored. However, some people rarely consult a medical professional. For instance, a study of over 8000 adults in England found that 22.8% (1846/8086) had not consulted a GP over the last year [12]. This means that CVD risk factors may remain undetected and untreated. In addition, the lack of GPs in certain regions is likely to further limit the number of people who are screened [13,14].

Various screening programs have been set up over the years to target people who rarely consult a GP. These programs have often focused on detecting hypertension, also known as the "silent killer," and they have been run in a variety of different settings [15-17]. Some screening programs have involved health checks carried out by nurses, pharmacists, or medical students in settings such as pharmacies, dental surgeries, and community centers [18]. However, these programs may not always be feasible, such as during pandemics when strict social distancing requirements are in place. In addition, they are likely to be limited by funding and staffing resources. As an alternative solution, self-screening facilities have been developed, which have the advantage of lower running costs as well as the potential to be used during pandemics [19]. For instance, health kiosks that assess blood pressure and weight are now available in pharmacies and shopping centers throughout the United States, including more than 6000 health stations developed by the company Higi (Chicago, United States). There is evidence that these have a positive impact. For instance, a study showed that 27.4% (17/62) of users with hypertension achieved blood pressure control within 6 to 12 months [20]. These rates could potentially be improved through patient education and by referring people on to treatment programs, such as the Kaiser Permanente treatment protocol [20,21]. However, health kiosks are not widely available in many countries, including France. Other possibilities for screening include the use of personal digital devices, such as smartphones; however, not all people have these and there are concerns about their accuracy [22].

Recently, a telemedicine booth was developed in France (the consult station, H4D, Paris) that can run a series of health checks in the privacy of a stand-alone booth [23]. The tests can be run by users without any input from a medical professional; there is also the option of using the device for teleconsultations with a physician. The booth can be used to assess blood pressure, BMI, and heart rate, which have all been linked to CVD [5,24]. In addition, other measures can also be obtained that can signal other health conditions, such as oxygen saturation, where low levels could point to anemia or pulmonary diseases, and rapid visual tests, where abnormalities could be a sign of macular degeneration or epiretinal membrane.

In this study, we assessed the feasibility and interest of using the consult station telemedicine booth for self-screening during the COVID-19 pandemic, when there was a particular need for remote medical services. The booths were placed in several vaccination centers throughout northern France at a time when large numbers of people were coming to be vaccinated. Our primary aim was to determine whether screening using the telemedicine booth would have an impact on the medical care of people with CVD risk factors. We also aimed to determine whether people who were not registered with a GP (11% of people in France) [13] had more CVD risk factors and whether they would be more likely to consult a medical professional following the screening.

Methods

Study Design

This study was a prospective, single-arm, multicenter study. All participants provided informed consent before taking part in the study; this concerned all those who had a health check using the consult station. Participants were also sent an information sheet by post or email before the follow-up telephone interview, which included a form to complete and return if they did not agree to their data being used for the study.

Setting

This study was carried out in 3 different vaccination centers between July 1, 2021, and July 1, 2022. The centers were all located in hospitals in northern France.

Participants

Participants were included in the study if they were aged 18 years or older and were affiliated to a social security scheme. Participants were excluded if they could not follow the instructions in the consult station or if they were under guardianship. There was no sample size calculation for this study.

Screening

During the waiting periods at the vaccination centers (before or after vaccination, or at any time for accompanying adults), people were offered a free health check using the consult station, a telemedicine booth developed by H4D (Paris, France). This is a class IIa medical device that can measure people's blood pressure, heart rate, oxygen saturation, BMI, and vision (acuity and central visual field using the Amsler grid) in less than 15 minutes. The health check is run by following a series of instructions on a screen, and it does not require the involvement of medical staff. However, a nurse was present to interpret the results, which were provided on a printed sheet of paper, and participants were advised to see a doctor if there were any readings outside of the normal range. For the purposes of this study, the following readings were considered to be outside the normal range: BMI >30 kg/m^2 (obese: 31 - 40 kg/m^2; severely obese >40 kg/m^2), heart rate >80 beats per minute (bpm) or <50 bpm (high heart rate: 81 - 100 bpm; tachycardia: >100bpm), oxygen saturation <94% (mild: 90 - 93%; hypoxia: <90%), systolic blood pressure <100mmHg or >140mmHg (high: 141 - 180mmHg; very high: >180mmHg), and diastolic blood pressure <70 mmHg or >90 mmHg (high: 91 - 100mmHg; very high: >100mmHg). The Amsler grid results were marked as abnormal if the participants saw the grid to be distorted (eg, wavy lines) when focusing on the central dot. We took the contact details (eg, telephone and email) of the participants who had screening results outside the normal range, with their permission. These participants were then contacted 3 months later to evaluate the impact of the screening.

XSL•FO
RenderX

Follow-Up

Three months after the health check, the participants were contacted by telephone and were asked to respond to a series of questions. These covered their demographic details (sex and history of smoking), their medical history before the health check (hypertension, cardiac arrhythmia, pulmonary disease, diabetes, or sleep apnea; whether they took medication), their reason for coming to the vaccination center (to be vaccinated, accompanying someone, or member of staff), whether they were registered with a GP, and how often they saw a medical professional. Participants were also asked whether they had consulted a doctor since the health check, and if so, whether this was with their GP. They were also asked whether the doctor had confirmed the cardiovascular risk factors and initiated regular follow-up appointments to monitor these. They were also asked whether they would have consulted a doctor without the health check (yes or no). Additional questions covered whether they had seen or planned to see a medical specialist, or whether they had been told that this was not necessary. Participants were also asked whether they had been to see a dietician. If they were smokers, they were also asked whether they had stopped smoking or taken steps to do so. Participants were also asked whether their health check had been before or after vaccination (or during vaccination for those accompanying others).

Outcome Measures

The primary outcome was the percentage of participants with abnormal screening results who then consulted a doctor (GP or specialist). The secondary outcome was the percentage of participants not registered with a GP who then consulted a doctor (GP or specialist) following abnormal screening results. An additional secondary outcome was the difference in cardiovascular risk factors between participants who were registered with a GP and those who were not registered with a GP.

Data Analysis

The program R (R Foundation for Statistical Computing) was used for the statistical analyses. The qualitative variables were summarized using frequencies and percentages; quantitative variables were summarized using the means and standard deviations, or medians, inter-quartile ranges (IQR), and ranges. Statistical analyses were run according to the type of data (eg, χ^2 or Fisher exact tests for categorical data) using a significance level of 0.05. We also ran univariate and multivariate analyses to identify factors that could predict whether the participants would consult a doctor following screening.

Ethical Considerations

The study was approved by the ELSAN Institutional Review Board (#2022-03-Dr DECAMBRON-01; approved on March 21, 2022) and is registered on the Health Data Hub (number: F20220309155706. All participants provided informed consent.

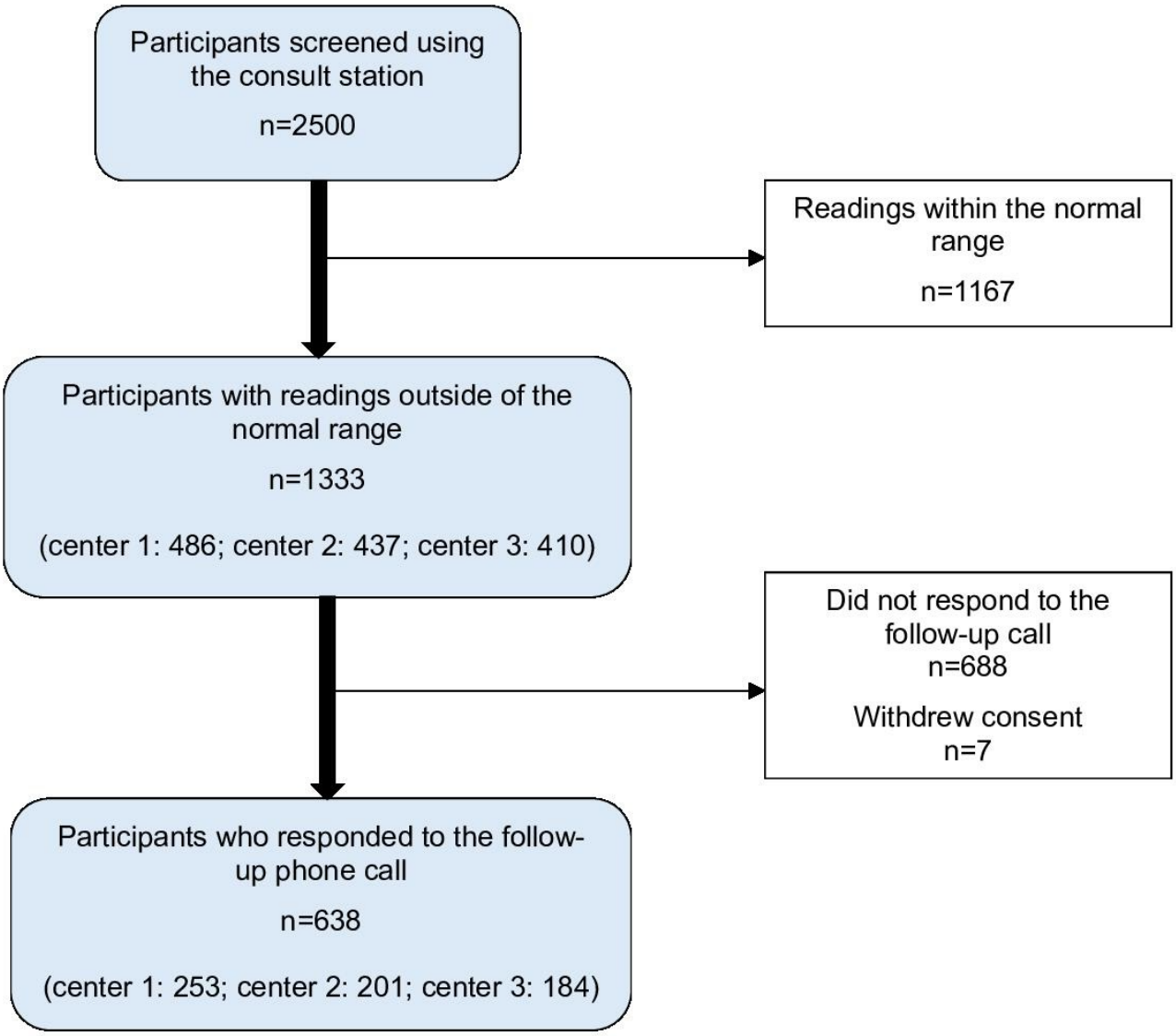
Results

Participants

Approximately 6000 people attended the three vaccination centers between July 2021 and September 2021 and were given the opportunity to take part in the study. There were around 2500 people who had a health check using the consult station (2500/6000; 41.7%). In total, 1333 participants (53.3%; 1333/2500) were found to have readings outside the normal range and so were advised to consult a doctor (center 1: N=486; center 2: N=437; center 3: N=410). These participants (N=1333) all provided their contact details, but 7 withdrew their consent and only 638 (47.9%) responded to the follow-up telephone call (center 1: N=253; center 2: N=201; center 3: N=184; see Figure 1).

XSL•FO
RenderX

Figure 1. Flow chart of the study participants.



Screening Results

Results sheets were available for just 1087 of the participants who used the consult station (1087/2500; 43.5%; all at center 1 and center 2). These data were analyzed, and it was found that over a quarter had heart rate, BMI, and blood pressure outside of the normal range (Table 1). The overlap between these abnormal readings is shown as a Venn diagram (Figure 2).

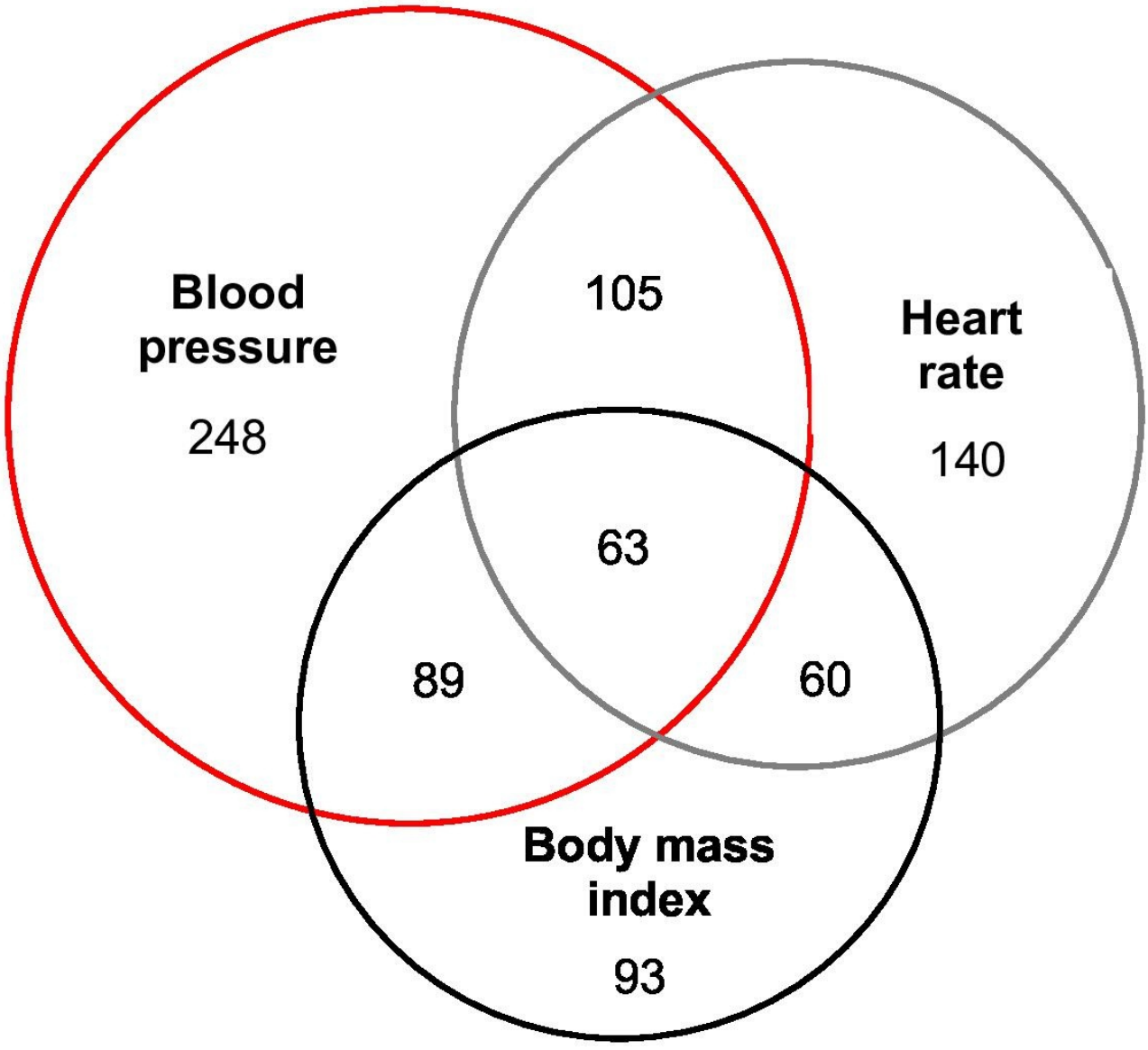
XSL•FO
RenderX

Table . Consult station screening results based on the available results sheets (total: 1087).

Variable	Statistical values	Number of readings outside normal range, n/N (%)[a]
Body mass index (kg/m^2), median (IQR) range	26 (23 - 30) 14 - 53	305/1025 (29.8)
Heart rate (beats per minute), median (IQR) range	75 (67-84) 21-132	368/985 (37.4)
Oxygen saturation (%), mean (SD) range	97.5 (1.6) 75 - 99	13/986 (1.3)
Systolic blood pressure (mmHg), median (IQR) range	126 (114 - 138) 84 - 225	283/1016 (27.9)
Diastolic blood pressure (mmHg), median (IQR) range	80 (73 - 89) 42 - 136	416/1016 (40.9)
Amsler grid (central visual field)	—[b]	105/1016 (10.3)

[a]The total number of available readings is lower than the actual number of participants who completed each measure. This is because results sheets were not available for all study participants (eg, none were available from center 3); in addition, certain measures were not obtained for all participants.
[b]—: not applicable.

Figure 2. Venn diagram showing the overlap between abnormal readings for body mass index, blood pressure, and heart rate. Note that the abnormal readings for blood pressure include both abnormal diastolic blood pressure and abnormal systolic blood pressure. Participants who had abnormal readings for both of these were counted once.



XSL•FO
RenderX

Primary Outcome Measure

Of the 638 participants who responded to the follow-up telephone call, 234 (36.7%) reported that they had consulted a doctor (GP or specialist) since the health check (center 1: 70/253, 27.7% participants; center 2: 53/201, 26.4% participants; center 3: 111/184, 60.3% participants). However, 24.8% of the participants (158/638) reported that they would have consulted a doctor anyway even without the health check (Table 2).

XSL•FO
RenderX

Table . Responses to the follow-up call questions (n=638).

Question	Responses, n/N (%)
Sex	
Female	394/638 (61.8)
Male	234/638 (36.7)
Missing	10/638 (1.6)
Reason for coming to vaccination center	
Vaccination	447/638 (70.1)
Other reason	189/638 (29.6)
Missing	2/638 (3.1)
Smoker	
Yes	183/638 (28.7)
No	455/638 (71.3)
Registered with GP[a]	
Yes	620/638 (97.2)
No	18/638 (2.8)
Frequency of GP appointments	
<1 per year	212/638 (33.2)
At least 1 per month	302/638 (47.3)
Around 4 per year	104/638 (16.3)
Missing	20/638 (3.1)
Medical history	
Hypertension	122/638 (19.1)
Cardiac arrhythmia	59/638 (9.2)
Pulmonary disease	49/638 (7.7)
Diabetes	50/638 (7.8)
Sleep apnea	42/638 (6.6)
Concomitant diseases[b]	
Yes	253/638 (39.7)
Cardiovascular system	118/638 (18.5)
Nervous system	65/638 (10.2)
Metabolic disorders	56/638 (8.8)
Respiratory system	45/638 (7.1)
Other	129[c]
GP appointment since health check	
Yes	178/638 (27.9)
No	457/638 (71.6)
Missing	3/638 (0.5)
Saw own GP	141/178 (79.2)[d]
Saw another GP	36/178 (20.2)[d]
Missing	1/178 (0.6)[d]
GP confirmed cardiovascular risk factors	59/178 (33.1)[d]

XSL•FO
RenderX

Question	Responses, n/N (%)
GP did not confirm risk factors	116/178 (65.2)[d]
Missing	3/178 (1.7)[d]
GP initiated regular follow up	56/59 (94.9)[e]
GP did not initiate regular follow up	2/59 (3.4)[e]
Missing	1/59 (1.7)[e]
Seen specialist since health check	
Yes	151/638 (23.7)
No	483/638 (75.7)
Missing	4/638 (0.6)
Seen dietitian since health check	
Yes	92/638 (14.4)
No	537/638 (84.2)
Missing	9/638 (1.4)
Would have seen doctor without health check	
Yes	158/638 (24.8)
No	474/638 (74.3)
Missing	6/638 (0.9)

[a]GP: general practitioner.

[b]Note that some patients may have had more than one concomitant disease.

[c]The percentage is not shown as the number includes participants who are counted more than once because they take more than one drug.

[d]The percentage only includes participants who saw a GP.

[e]The percentage only includes participants who saw a GP and had the risk factors confirmed.

Chi-square tests were run to identify potential factors that might predict whether participants consulted a doctor following the health check. The factors that had a *P* value above .2 were the participants' sex (*P*=.16), their reason for coming to the vaccination center *(P*<.0001), whether they smoked *(P*=.03), whether they were registered with a GP *(P*=.19), and whether they had a history of certain diseases (hypertension, cardiac arrhythmia, pulmonary disease, diabetes, and sleep apnea*; P*<.0001). A multivariate analysis was then run, which included these factors. The results showed that the reason for coming to the vaccination center and a history of certain diseases were statistically significant at the *P*<.05 level *(P*<.001 for both factors). The pattern of results indicated that participants were more likely to see a doctor following the health check if they had come to the center to be vaccinated or if they had a history of certain diseases.

Secondary Outcome Measures: Differences Between Participants Registered With a General Practitioner and Those Not Registered With a General Practitioner

There were 18 participants (18/638; 2.8%) who reported that they were not registered with a GP. Five of these (27.8%) consulted a doctor following the health check. Due to the lack of data, we did not carry out further analyses.

Additional Analyses

Additional analyses were run to determine whether there were significant differences between the 3 vaccination centers in terms of the follow-up call results (Table 3). We found that there was a significant difference between the centers in terms of the proportion of participants who were registered with a GP *(P*=.04). Specifically, fewer were registered with a GP at center 2 (94.5%) than at center 3 (98.4%) and center 1 (98.4%). This can be attributed to the relative lack of GPs in this region (101 per 100,000 individuals compared with 129 and 159 per 100,000 population for the other 2 regions) [25]. We also found a significant difference concerning the proportion of participants who consulted a GP following the health check *(P*<.001). This was higher at center 2 (45%) than at the 2 other centers (20.9% and 19.2%). Similarly, a higher proportion of participants at center 2 (34.3%) consulted a medical specialist than at the 2 other centers (17.7% and 20.7%*; P*<.0001). These results indicate that the screening led more participants at center 2 to consult a doctor. However, a higher proportion of participants at this center reported that they would have seen a doctor anyway (50.3% vs 14.5% and 12.0%*; P*<.0001). Of note, the participants at the different centers did not differ significantly in terms of how often they usually saw a doctor *(P*=.82).

XSL•FO
RenderX

Table . Differences between the vaccination centers.

Question	Center 1, n/N (%)	Center 2, n/N (%)	Center 3, n/N (%)	*P* value (test)
Sex				.43[a]
Female	152/253 (60.1)	120/201 (59.7)	122/184 (66.3)	
Male	97/253 (38.3)	76/201 (37.8)	61/184 (33.2)	
Missing	4/253 (1.6)	5/201 (2.5)	1/184 (0.5)	
Smoker				.02[a]
Yes	87/253 (34.4)	56/201 (27.9)	40/184 (21.7)	
No	166/253 (65.6)	145/201 (72.1)	144/184 (78.3)	
Missing	0/253 (0)	0/201 (0)	0/184 (0)	
Registered with GP[b]				.04[c]
Yes	249/253 (98.4)	190/201 (94.5)	181/184 (98.4)	
No	4/253 (1.6)	11/201 (5.5)	3/184 (1.6)	
Missing	0/253 (0)	0/201 (0)	0/184 (0)	
See GP				.82[e]
<1 per year	90/249[d] (36.1)	56/188[d] (29.8)	66/181[d] (36.5)	
~4 per year	49/249[d] (19.7)	25/188[d] (13.3)	30/181[d] (16.6)	
1 per month+	110/249[d] (44.2)	107/188[d] (56.9)	85/181[d] (47)	
Missing	0/249[d] (0)	2/188[d] (1.1)	0/181[d] (0)	
Seen GP since health check				<.0001[a]
Yes	53/253 (20.9)	90/201 (44.8)	35/184 (19.0)	
No	200/253 (79.1)	110/201 (54.7)	147/184 (79.9)	
(Missing)	0/253 (0)	1/201 (0.5)	2/184 (1.1)	
Saw own GP	39/53[f] (73.6)	81/90[f] (90)	21/35[f] (60)	.001[a]
Saw different GP	13/53[f] (24.5)	9/90[f] (10)	14/35[f] (40)	
(Missing)	1/53[f] (1.9)	0/90[f] (0)	0/35[f] (0)	
GP confirmed cardiovascular risk factors	28/53[f] (52.8)	15/90[f] (16.7)	16/35[f] (45.7)	<.0001[a]
GP did not confirm risk factors	23/53[f] (43.4)	75/90[f] (83.3)	18/35[f] (51.4)	
(Missing)	2/53[f] (3.8)	0/90[f] (0)	1/35[f] (2.9)	
GP initiated regular follow up	27/28[g] (96.4)	13/15[g] (86.7)	16/16[g] (100)	.06[c]
GP did not initiate follow up	0/28[g] (0)	2/15[g] (13.3)	0/16[g] (0)	
(Missing)	1/28[g] (3.6)	0/15[g] (0)	0/16[g] (0)	
Seen specialist since health check				<.0001[a]
Yes	44/253 (17.4)	69/201 (34.3)	38/184 (20.7)	
No	205/253 (81)	132/201 (65.7)	146/184 (79.3)	
Missing	4/253 (1.6)	0/201 (0)	0/184 (0)	
Would have consulted physician without health check				<.0001[a]

XSL•FO
RenderX

Question	Center 1, n/N (%)	Center 2, n/N (%)	Center 3, n/N (%)	*P* value (test)
Yes	36/253 (14.2)	100/201 (49.8)	22/184 (12.0)	
No	213/253 (84.2)	99/201 (49.3)	162/184 (88)	
Missing	4/253 (1.6)	2/201 (1.0)	0/184 (0)	

[a]Chi-square test.

[b]GP: general practitioner.

[c]Fisher exact test.

[d]The total only includes participants who were registered with a GP.

[e]Kuskal-Wallis test.

[f]The total only includes participants who saw a GP since the health check.

[g]The total only includes participants who saw a GP and had the cardiovascular risk factors confirmed.

Discussion

Principal Findings

In this study, large numbers of people were screened for CVD risk factors during the COVID-19 pandemic using the consult station telemedicine booth. According to the available results sheets, over half had readings outside the normal range, which were mostly for blood pressure (diastolic: 41%; systolic: 28%), heart rate (37%), and BMI (30%). Even though relatively few people reported that they consulted a doctor because of their results (around 12% of the follow-up call respondents), the screening would nevertheless have raised people's awareness of their risk factors and may have led to beneficial lifestyle changes.

Our study showed that the telemedicine booth was generally well-accepted by the French public, with around 42% of visitors to the vaccination centers choosing to use the device. Although the booth's enclosed environment may have led to fears of infection during the pandemic, we found that this did not prevent almost half of the visitors from using the consult station and benefiting from the free health check. We speculate that even more people would use the device at other times when there is less fear of infection. Indeed, the privacy provided by the booth is likely to encourage more people to use the device as this could improve user experience.

This high level of acceptance supports the use of telemedicine booths for CVD screening not only during pandemics but also at other times. They could potentially be used to screen large numbers of people, and they have the advantage of low running costs. They would be of particular benefit to regions that lack medical professionals, the so-called "medical deserts" of France, to improve screening services in these regions [13]; the booths could also potentially be used to provide teleconsultations with medical professionals in these regions. The telemedicine booths should ideally be placed in carefully selected locations to target groups of people who are known to have a higher risk of CVD, such as older adults. This has already been done in the United States, where blood pressure checks have been run in Black-owned barbershops to target Black American men who have a high prevalence of hypertension [26].

One of the strengths of our study is that we evaluated the impact of the screening by examining whether people with abnormal results went on to consult a physician. This is an important consideration, as it has been pointed out that people with hypertension may not always seek medical treatment [27]. In our study, only around 12% of the respondents reported that they consulted a doctor because of the screening. This raises the question as to why this was the case, especially as the participants had been explicitly advised to see a doctor. One possibility relates to the fact that the study took place during the COVID-19 pandemic, and so participants may have avoided medical consultations when there was a risk of infection as well as an overwhelmed health system. Another possibility is that some participants may already have been aware of their condition and so they may already have consulted a doctor about it. It is also possible that some participants did not fully understand the dangers of their abnormal readings or the fact that effective treatment options are available from their GP; this consideration is important, as it could potentially be remedied by providing educational material.

Although relatively few participants with abnormal readings consulted a doctor, improvements may nevertheless have occurred. Specifically, it is possible that the participants may have adopted lifestyle changes to improve their health. For instance, they may have searched the internet for advice on how to lose weight or lower their blood pressure. In support of this, a previous study found that many people with obesity prefer to find out about weight management from the internet, friends, and smartphone apps rather than from their health care provider [28].

Our statistical analyses revealed 2 factors that could predict whether participants with abnormal readings would consult a doctor. The first was having a history of certain diseases, such as diabetes. This could be attributed to the fact that patients with these diseases often have regular follow-up appointments with a GP. In addition, they may also be more aware of their need for medical care. The second factor was coming to the center to be vaccinated (as opposed to accompanying someone or for work); this may be because these participants were generally inclined to seek medical care (eg, vaccination) and so they would be more likely to consult a doctor following the screening.

One of the aims of our study was to examine differences between participants who were registered with a GP and those who were not. In France, it has been reported that a substantial proportion of the population (11%) does not have a GP [13]. However, we only identified 18 participants (2.8%) who were

XSL•FO
RenderX

not registered with a GP, and so we did not carry out statistical analyses. This low number may be because the people who came to the vaccination centers were generally inclined to seek medical care, including vaccination, and so would also be more likely to register with a GP. We noted that most of the participants who were not registered with a GP were at center 2, which can be attributed to the relative lack of GPs in this region (101 per 100,000 population for the center 2 as opposed to 129 and 159 per 100,000 population for the 2 other centers) [25]. However, this did not appear to affect the postscreening medical consultations, as the participants at this center were also more likely to see a physician following the health check. This latter finding may relate to the fact that over half of the participants at center 2 reported seeing a GP at least once a month, unlike at the other 2 centers; these participants may therefore have been more likely to consult a physician following the health check simply because they were already frequent health care users.

The limitations of this study include the lack of a control group and missing data (no screening results for over half of the participants). In addition, we were unable to link the follow-up results to the initial screening data (except for 55 participants). In addition, it is possible that certain screening measures (eg, blood pressure) may have been affected by the vaccination (eg, prevaccination anxiety or postvaccination side effects) [29]. The results might also have been affected by bias in terms of the study participants, as those who took part were at a vaccination center and so were presumably inclined to seek medical care. Finally, we did not assess whether the screening had an impact on risk factor control over the long term.

In the future, it would be helpful to run further studies to identify the most effective locations for screening, with the aim of targeting people who rarely consult a GP. In addition, it would be helpful to develop educational pamphlets or videos that explain CVD risk factors and the treatment options that are available, as this could potentially increase the number of people who seek medical care. Additional screening measures could also be introduced to detect other diseases, such as the ankle-brachial index, which can indicate whether someone has peripheral artery disease, a prevalent but currently under-diagnosed condition [30]. It would also be beneficial to introduce mHealth systems, such as mobile phone-based lifestyle interventions, for patients who are found to have CVD risk factors. These has previously been shown to effectively lower the risk of atherosclerotic CVD and increase peak oxygen consumption and so could potentially greatly increase the health benefits following screening [31,32].

As a final note, there are current plans for the consult station to be used for screening in regions of France where there is a high prevalence of hypertension and a lack of medical services. Our medical group has also started sending mobile health buses to medically underserved towns in the Hauts de France region, which aim to promote healthy lifestyles and prevent disease through patient education.

Conclusions

By placing the consult station in vaccination centers during the COVID-19 pandemic, it was possible to screen large numbers of people for CVD risk factors. Our results support the use of telemedicine booths for health screening, particularly in regions that lack medical professionals. In the future, it would be helpful to develop educational material to ensure that users understand the significance of any abnormal results and the treatment options that are available.

Acknowledgments

This research was funded by the ELSAN group, France. The authors thank Faycal Jarroudi for his work on the clinical data and for supporting the centers. We thank Jessica Foxton, PhD, from Newmed Publishing, Paris, for medical writing assistance.

Authors' Contributions

MD contributed to study conception, design, data collection, data analysis, and interpretation. CT conducted data analysis, interpretation, and draft manuscript preparation. All authors reviewed the results and approved the final version of the manuscript.

Conflicts of Interest

None declared.

References

1. Roth GA, Abate D, Abate KH, et al. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet 2018 Nov;392(10159):1736-1788. [doi: 10.1016/S0140-6736(18)32203-7]
2. Assurance Maladie. Améliorer la qualité du système de santé et maîtriser les dépenses [Health Insurance Improving the quality of the health system and controlling expenditure] [WWW Document]. URL: https://assurance-maladie.ameli.fr/sites/default/files/2022-07_rapport-propositions-pour-2023_assurance-maladie_5.pdf [accessed 2023-09-14]
3. Ministère de la Santé et de la Prévention. Maladies cardiovasculaires [Ministry of Health and Prevention, 2023 Cardiovascular diseases] [WWW Document]. URL: https://sante.gouv.fr/soins-et-maladies/maladies/maladies-cardiovasculaires/article/maladies-cardiovasculaires [accessed 2023-09-07]

XSL•FO
RenderX

4. Timmis A, Vardas P, Townsend N, et al. European Society of Cardiology: cardiovascular disease statistics 2021. Eur Heart J 2022 Feb 22;43(8):716-799. [doi: 10.1093/eurheartj/ehab892] [Medline: 35016208]
5. Danaei G, Ding EL, Mozaffarian D, et al. The preventable causes of death in the United States: comparative risk assessment of dietary, lifestyle, and metabolic risk factors. PLoS Med 2009 Apr 28;6(4):e1000058. [doi: 10.1371/journal.pmed.1000058] [Medline: 19399161]
6. Cardiovascular disease, chronic kidney disease, and diabetes mortality burden of cardiometabolic risk factors from 1980 to 2010: a comparative risk assessment. The Lancet Diabetes & Endocrinology 2014 Aug;2(8):634-647. [doi: 10.1016/S2213-8587(14)70102-0]
7. Fenech G, Vallée A, Cherfan M, et al. Poor awareness of hypertension in France: the CONSTANCES population-based study. Am J Hypertens 2020 May 21;33(6):543-551. [doi: 10.1093/ajh/hpaa018] [Medline: 32202627]
8. Ettehad D, Emdin CA, Kiran A, et al. Blood pressure lowering for prevention of cardiovascular disease and death: a systematic review and meta-analysis. Lancet 2016 Mar 5;387(10022):957-967. [doi: 10.1016/S0140-6736(15)01225-8] [Medline: 26724178]
9. Perdomo CM, Cohen RV, Sumithran P, Clément K, Frühbeck G. Contemporary medical, device, and surgical therapies for obesity in adults. Lancet 2023 Apr 1;401(10382):1116-1130. [doi: 10.1016/S0140-6736(22)02403-5] [Medline: 36774932]
10. Hughes CA, Ahern AL, Kasetty H, et al. Changing the narrative around obesity in the UK: a survey of people with obesity and healthcare professionals from the ACTION-IO study. BMJ Open 2021 Jun 30;11(6):e045616. [doi: 10.1136/bmjopen-2020-045616] [Medline: 34193488]
11. Iwabu M, Yamauchi T, Shimomura I, Eguchi K, Ogawa Y. Perceptions, attitudes and barriers to obesity management: Japanese data from the ACTION-IO study. J Diabetes Investig 2021 May;12(5):845-858. [doi: 10.1111/jdi.13427] [Medline: 33021048]
12. Song F, Bachmann M, Howe A. Factors associated with the consultation of GPs among adults aged ≥16 years: an analysis of data from the Health Survey for England 2019. BJGP Open 2023 Sep;7(3):BJGPO.2022.0177. [doi: 10.3399/BJGPO.2022.0177] [Medline: 37217212]
13. Sénat No 589 [WWW Document]. URL: https://www.senat.fr/rap/r21-589/r21-5891.pdf [accessed 2023-09-14]
14. Bes JM, Flinterman LE, González AI, Batenburg RS. Recruitment and retention of general practitioners in European medical deserts: a systematic review. Rural Remote Health 2023 Feb;23(1):7477. [doi: 10.22605/RRH7477] [Medline: 36854290]
15. Kotwal A, Joseph P. Conquering the silent killer: hypertension prevention and management in India. Indian J Community Med 2023;48(1):1-3. [doi: 10.4103/ijcm.ijcm_1000_22] [Medline: 37082390]
16. Ferdinand KC, Patterson KP, Taylor C, Fergus IV, Nasser SA, Ferdinand DP. Community-based approaches to prevention and management of hypertension and cardiovascular disease. J Clin Hypertens (Greenwich) 2012 May;14(5):336-343. [doi: 10.1111/j.1751-7176.2012.00622.x] [Medline: 22533661]
17. Maramba ID, Jones R, Austin D, Edwards K, Meinert E, Chatterjee A. The role of health kiosks: scoping review. JMIR Med Inform 2022 Mar 29;10(3):e26511. [doi: 10.2196/26511] [Medline: 35348457]
18. Fleming S, Atherton H, McCartney D, et al. Self-screening and non-physician screening for hypertension in communities: a systematic review. Am J Hypertens 2015 Nov;28(11):1316-1324. [doi: 10.1093/ajh/hpv029] [Medline: 25801901]
19. Letafat-Nejad M, Ebrahimi P, Maleki M, Aryankhesal A. Utilization of integrated health kiosks: a systematic review. Med J Islam Repub Iran 2020;34:114. [doi: 10.34171/mjiri.34.114] [Medline: 33315998]
20. Shea S, Thompson JLP, Schwartz JE, et al. The Retail Outlet Health Kiosk Hypertension Trial (ROKHYT): pilot results. Am J Hypertens 2022 Jan 5;35(1):103-110. [doi: 10.1093/ajh/hpab129] [Medline: 34382648]
21. Fontil V, Gupta R, Moise N, et al. Adapting and evaluating a health system intervention from Kaiser Permanente to improve hypertension management and control in a large network of safety-net clinics. Circ Cardiovasc Qual Outcomes 2018 Jul;11(7):e004386. [doi: 10.1161/CIRCOUTCOMES.117.004386] [Medline: 30002140]
22. Hayıroğlu M. Telemedicine: current concepts and future perceptions. Anatol J Cardiol 2019 Oct;22(Suppl 2):21-22. [doi: 10.14744/AnatolJCardiol.2019.12525] [Medline: 31670712]
23. Falgarone G, Bousquet G, Wilmet A, et al. A teleconsultation device, consult station, for remote primary care: multisite prospective cohort study. J Med Internet Res 2022 May 17;24(5):e33507. [doi: 10.2196/33507] [Medline: 35254263]
24. Seravalle G, Quarti Trevano F, Grassi G. Heart rate as a predictor of cardiovascular risk. Minerva Med 2021 Feb;112(1):130-143. [doi: 10.23736/S0026-4806.20.06695-1] [Medline: 32512980]
25. Professionnels de santé au 1er janvier 2022 Insee [Healthcare professionals on January 1, 2022]. URL: https://www.insee.fr/fr/statistiques/2012677#titre-bloc-1 [accessed 2023-08-05]
26. Ravenell J, Green T, Arabadjian M, Schoenthaler A, Ogedegbe O. Barbershop-facilitated community-to-clinic linkage implementation program: rationale and protocol for a novel program to prevent hypertension among black men. Am J Hypertens 2023 Apr 15;36(5):240-247. [doi: 10.1093/ajh/hpac133] [Medline: 37061797]
27. Chen S, Sudharsanan N, Huang F, Liu Y, Geldsetzer P, Bärnighausen T. Impact of community based screening for hypertension on blood pressure after two years: regression discontinuity analysis in a national cohort of older adults in China. BMJ 2019 Jul 11;366:l4064. [doi: 10.1136/bmj.l4064] [Medline: 31296584]

XSL•FO
RenderX

28. Lim S, Oh B, Lee SH, Kim YH, Ha Y, Kang JH. Perceptions, attitudes, behaviors, and barriers to effective obesity care in South Korea: results from the ACTION-IO study. J Obes Metab Syndr 2020 Jun 30;29(2):133-142. [doi: 10.7570/jomes20013] [Medline: 32507770]
29. Zappa M, Verdecchia P, Spanevello A, Visca D, Angeli F. Blood pressure increase after Pfizer/BioNTech SARS-CoV-2 vaccine. Eur J Intern Med 2021 Aug;90:111-113. [doi: 10.1016/j.ejim.2021.06.013] [Medline: 34158234]
30. Campia U, Gerhard-Herman M, Piazza G, Goldhaber SZ. Peripheral artery disease: past, present, and future. Am J Med 2019 Oct;132(10):1133-1141. [doi: 10.1016/j.amjmed.2019.04.043] [Medline: 31150643]
31. Tekkeşin A, Hayıroğlu M, Çinier G, et al. Lifestyle intervention using mobile technology and smart devices in patients with high cardiovascular risk: A pragmatic randomised clinical trial. Atherosclerosis 2021 Feb;319:21-27. [doi: 10.1016/j.atherosclerosis.2020.12.020] [Medline: 33465658]
32. Hayıroğlu M, Çınar T, Cilli Hayıroğlu S, Şaylık F, Uzun M, Tekkeşin A. The role of smart devices and mobile application on the change in peak VO2 in patients with high cardiovascular risk: a sub-study of the LIGHT randomised clinical trial. Acta Cardiol 2023 Nov;78(9):1000-1005. [doi: 10.1080/00015385.2023.2223005] [Medline: 37318090]

Abbreviations

bpm: beats per minute
CVD: cardiovascular disease
GP: general practitioner

Edited by A Kushniruk, Y Senathirajah; submitted 07.02.24; peer-reviewed by J Edwards, M Hayıroğlu; revised version received 13.02.25; accepted 14.02.25; published 22.04.25.

Please cite as:
Decambron M, Tchikladze Merand C
Telemedicine Booths for Screening Cardiovascular Risk Factors: Prospective Multicenter Study
JMIR Hum Factors 2025;12:e57032
URL: https://humanfactors.jmir.org/2025/1/e57032
doi:10.2196/57032

© Mélanie Decambron, Christine Tchikladze Merand. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 22.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Education and Symptom Reporting in an mHealth App for Patients With Cancer: Mixed Methods Development and Validation Study

Carolina Muñoz Olivar[1,2*], MSc; Miguel Pineiro[2,3*], PhD; Juan Sebastián Gómez Quintero[4*], PhD; Carlos Javier Avendaño-Vásquez[5,6*], PhD; Pablo Ormeño-Arriagada[7*], PhD; Silvia Palma Rivadeneira[2,8,9*], MSc; Carla Taramasco Toro[2,3*], PhD

[1]PhD in Health Sciences, Faculty of Medicine, Antonio Nariño University, Bogotá, Colombia

[2]Center for Cancer Prevention and Control (CECAN), Santiago, Chile

[3]Faculty of Engineering, Institute of Technologies for Innovation in Health and Well-being, Andrés Bello University, Viña del Mar, Chile

[4]Energy Transformation Center, Faculty of Engineering, Andrés Bello University, Santiago, Chile

[5]Faculty of Nursing, Antonio Nariño University, Bogotá, Colombia

[6]School of Nursing, Industrial University of Santander, Bucaramanga, Colombia

[7]Faculty of Engineering, Business, and Agro-Environmental Sciences, Department of Computer Civil Engineering, Universidad Viña del Mar, Viña del Mar, Chile

[8]Facultad de Medicina, Escuela de Enfermería, Pontificia Universidad Católica de Chile, Santiago, Chile

[9]UC-CHRISTUS Health Network, Santiago, Chile

[*]all authors contributed equally

Corresponding Author:
Carolina Muñoz Olivar, MSc
PhD in Health Sciences, Faculty of Medicine, Antonio Nariño University, Bogotá, Colombia

Abstract

Background: The widespread prevalence of cancer across the globe demands cutting-edge solutions for its treatment. Current cancer therapies, notably chemotherapy, pose challenges due to their side effects. The early detection and management of the side effects are vital but complex. This study introduces a mobile health app designed to bridge the communication gaps between patients with cancer and health care providers. Hence, it allows patients to report symptoms immediately and also enables proactive symptom management by health care providers.

Objective: This study has 2 objectives: first, to design a cancer-focused mobile health app that integrates educational content and real-time symptom reporting for chemotherapy patients. Second, to validate and evaluate the app quality using the Mobile App Rating Scale (MARS). The app seeks to foster health care communication, reduce hospital readmissions, and optimize symptom management, contributing to a more impactful patient experience.

Methods: This mixed-methods study details the development and validation of mobile health applications. The app was designed by a multidisciplinary team, including nurses, medical professionals, pharmaceutical chemists, computer engineers, and software developers, using agile methodologies. For validation, the app was assessed by 13 evaluators, including clinical professionals (nurses and physicians) and engineers. The evaluation included technical performance analysis using Google tools and quality assessment using the MARS, which measures engagement, functionality, aesthetics, and information quality.

Results: Performance metrics highlighted areas for improvement, with loading times showing delays in displaying content. Meanwhile, the response time of the app was moderate, and visual stability remained excellent. The app achieved an overall MARS score of 3.75 (SD 0.42), indicating consistent quality, with functionality scoring the highest (4.35; SD 0.52) and engagement the lowest (3.31; SD 0.61). The reliability of the MARS was confirmed (interclass correlation coefficient: 0.84; 95% CI: 0.72 - 0.92). Evaluators unanimously praised the app's potential benefits for patients and clinical professionals while identifying areas for improvement such as customization, onboarding guidance, and navigation.

Conclusions: The CONTIGO app showed strengths in functionality, usability, and information quality, supported by robust security measures. However, areas such as user interactivity and engagement require improvement. Future refinements will integrate insights from patients with cancer to address user-specific needs and enhance the oncology care experience.

(JMIR Hum Factors 2025;12:e60169) doi:10.2196/60169

XSL•FO
RenderX

KEYWORDS

cancer; patient-reported outcome measures; software design; unpleasant symptom; toxicity; mHealth; mobile health; surveys and questionnaires; application; design; evaluation; chemotherapy; health care communication; mixed methods; validation; efficiency; security

Introduction

According to the World Health Organization (WHO), cancer is one of the leading causes of mortality worldwide. In 2020, approximately 19 million people were diagnosed, and 9.3 million people die each year [1,2]. In the particular context of Chile, a comparable pattern was noted; in the same year, 54,000 cases of cancer were diagnosed, with 28,000 fatalities attributed to the disease annually. Additionally, cancer is the leading cause of disease burden and is responsible for 13% of disability-adjusted life years (DALYs) [3-6].

Continuing medical advances and the current therapeutic approach to cancer are multimodal and incorporate various strategies to maximize treatment efficacy. This comprehensive approach encompasses modalities such as radiotherapy, surgery, immunotherapy, and chemotherapy, the latter considered the gold standard [7-11]. Additionally, the application of an interdisciplinary approach includes palliative care to reduce suffering and improve the overall quality of life of patients and their families. This multidimensional strategy seeks to address the disease and provide comprehensive support, considering the patient's physical, spiritual, and psychosocial dimensions throughout the disease [12-14].

In this sense, it is essential for health care professionals to monitor and accompany patients and continues to be a challenge for health care professionals, requiring collaborative efforts between clinical staff, patients, and their families. The WHO promotes ensuring quality in care processes and incorporating patient-reported outcome measures (PROMs) and patient-reported experience measures (PREMs) [15] with tools that allow for the timely identification of these symptoms to contribute positively to patient well-being and service delivery by clinical staff [16].

Significant growth in health care technologies may favor this process. According to projections by the Center for Technology and Communication, approximately 7.49 billion people worldwide will be using mobile devices by 2025 [17] and Android users currently have access to 3.55 million apps [18], of which 1 million are aimed at influencing people's health, fitness, nutrition, and well-being [19].

Mobile applications in health, known as mHealth apps, have proven helpful and significant benefits for patients and health care professionals. These apps assist patients in adopting healthier lifestyles, enhancing self-care, and improving the quality of services, with a substantial focus on educational services and sensibilization programs [20-22]. This information is crucial for promptly notifying health professionals so that they can make informed decisions and effectively manage patients' adverse effects [23-27].

In line with these advancements, Wasserman et al [28] conducted a study to evaluate the quality and usefulness of 17 digital mHealth applications designed for adults with cancer. The study focused on assessing apps with tailored digital features and measured their performance using standardized tools, including the Mobile App Rating Scale (MARS) and a checklist of user-desired features derived from the literature. These features included functionalities such as symptom tracking, treatment management, and communication tools, highlighting the potential role of mHealth apps in supporting comprehensive cancer care and addressing unmet needs in symptom management and patient engagement.

We developed the CONTIGO app, designed specifically for the Chilean context and to operate within the UC-Christus network [29]. Its primary aim is to help patients adopt healthier lifestyles, improve their self-care, and positively impact their quality of life and that of their families. In addition, the app aims to facilitate communication between health care professionals and patients to reduce hospital readmissions and unnecessary emergency room visits while optimizing symptom management.

This study presents the results of designing a mHealth application for patients with cancer that integrates educational content with the functionality of immediate notification of unpleasant symptoms associated with chemotherapy administration to the treating physician and nurse. Additionally, it shows the results of validation and quality assessment of the mobile application "CONTIGO" by using the MARS, which allows mHealth developers to maximize the application's benefits to significantly impact the patient [22,30-36].

Methods

Study Type

This study employed a mixed methods approach. This methodology is presented in two sections: (1) design a mobile application, and (2) validation and quality assessment.

Ethical Considerations

The research protocol for designing a health mobile application for patients with cancer was reviewed and approved by the Scientific Ethics Committee CEC Med-UC (09/2023) ID 230808005. All participants provided informed consent before submitting their responses. Participation was voluntary, and participants were informed of their right to withdraw at any time without consequence. The study did not include any vulnerable populations and was conducted in accordance with established ethical standards for human subjects research. Data were collected anonymously through a Google Form, and no personally identifiable information was recorded. As a result, all data are fully anonymized and cannot be linked to individual participants. Researchers did not have access to responses while participants completed the evaluation, ensuring both privacy and impartiality. No financial or material compensation was offered for participation. Anonymized data may be made available upon request

XSL•FO
RenderX

mHealth App Design

To develop "CONTIGO," a multidisciplinary team consisting of nurses, pharmaceutical chemists, medics, computer engineers, and developers was formed. Previous research was considered, and principles of agile methodologies were adopted, such as software development, responsiveness to change, an iterative process with the client, and the assessment of individuals and interactions in the process. As a result, 5 stages for the development of the application were established.

Phase 1: Information Analysis

To construct the educational content, we conducted focus groups involving 1 PhD student in health science, a nurse with a master's degree, 3 nurses with oncology experience, and 2 oncologists. These professionals identified a critical gap in patient education: newly diagnosed oncology patients often forget or lack access to the educational information provided during consultations. This application was developed as a response to this need, ensuring that patients have a reliable and accessible source of information to revisit at their convenience. The research team adhered to the recommendations outlined in the Guide of Elaboration and Application of Focus Groups (G-DEP-005) [37] and followed an iterative process to obtain a theoretical framework for each section of the application. The framework was presented using text, images, and personalized videos. The application's initial focus prioritized the needs and preferences of health care professionals because the tool was designed to support their educational role with patients. The requirements were collected through a detailed request form completed by a trained clinical team member, who classified them by priority. This approach ensured the application's alignment with clinical standards while laying the groundwork for future integration of patient-specific preferences.

Phase 2: Design

We considered a previously created interface [38], which had been evaluated in earlier studies using patient focus groups. Feedback from these studies informed the decision to adopt and optimize this interface for the current application. Adjustments were made to the graphical interface, focusing on scenario optimization and comprehensive software restructuring. These modifications aimed to ensure an attractive, user-friendly, and efficient interface for the application's target users.

Architectural development prioritized asynchronous services to enhance system flexibility and scalability, ensuring the application could adapt to future updates and user demands.

Phase 3: Development

In this phase, decisions were made regarding implementation and coding, scenario adjustment, and software restructuring. Docker containers were used to simplify deployment and optimize application management and distribution. Security measures, such as management through Cloudflare, were integrated to prevent potential distributed denial-of-service (DDoS) attacks. Data encryption techniques were also applied to ensure information security. An open-source, high-performance PostgreSQL relational database management system was defined. A backup program was established every 3 hours using incremental backup copies to ensure data integrity and availability.

Phase 4: Functional Testing

This phase focused on verifying the system's robustness through emulation and simulation tests, followed by the implementation of the application on real devices. The teams were divided into 2 groups: one consisting of engineers and developers and the other consisting of a clinical team. This division aimed to enhance mobile device applications' quality, stability, and user experience.

Phase 5: Deployment

The engineering team officially delivered the "CONTIGO" mobile application to proceed with the final validation process and quality evaluation. Figure 1 presents the final version of the application and its sections, which include a variety of features designed to support patients and clinical professionals.

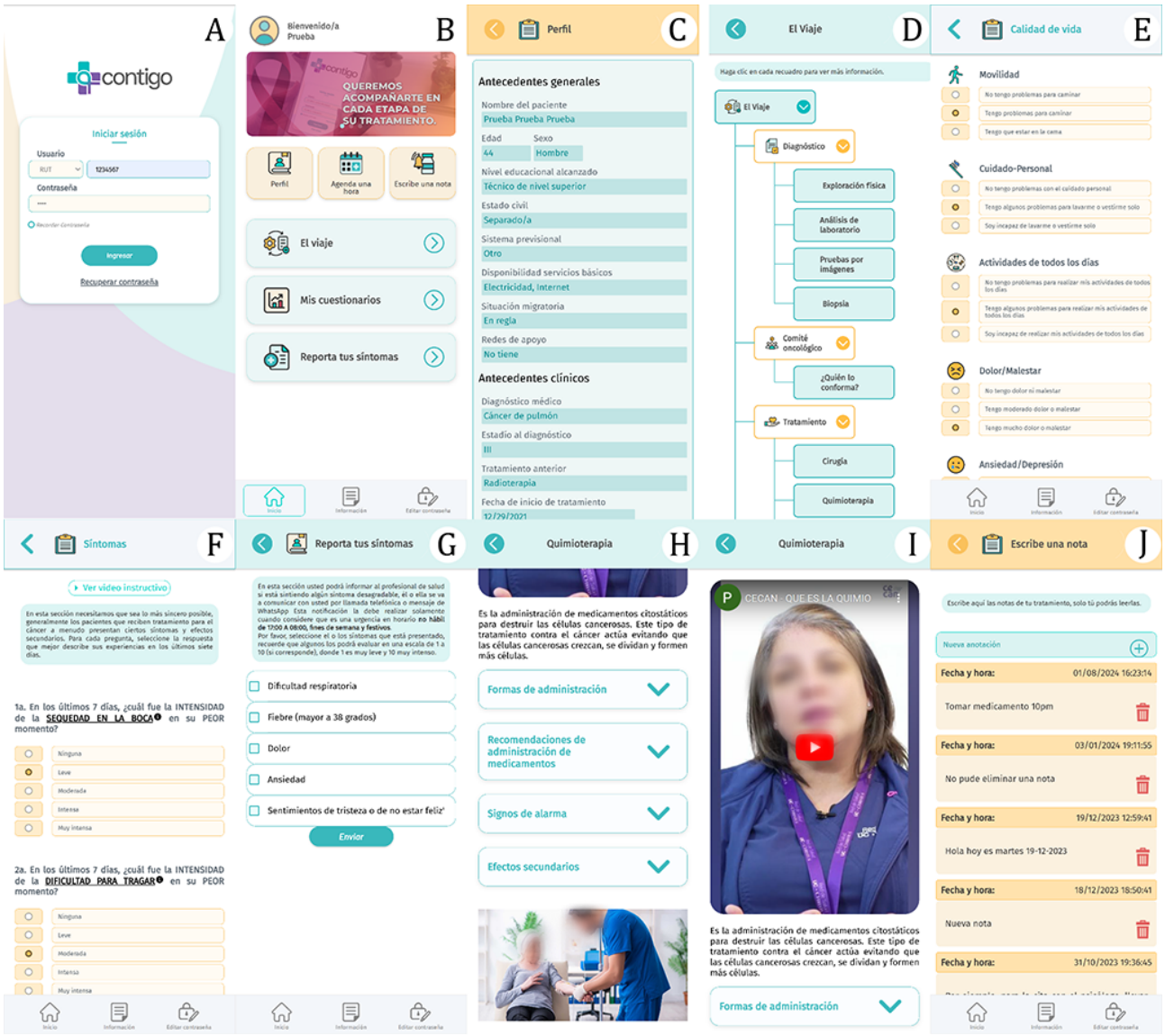
The application allows patients to record chemotherapy-related symptoms, enabling clinical teams to make data-driven decisions and manage side effects more effectively. Furthermore, the app offers features such as patient profile management, appointment scheduling, and access to educational materials presented through text, images, and videos.

The educational videos included in the app cover a variety of topics designed to improve patients' understanding of their treatment and care. These topics include laboratory tests, imaging tests, oncology committee decisions, treatment options, chemotherapy, radiotherapy, immunotherapy, palliative care, what is cancer, what is chemotherapy, what is a catheter with a reservoir, what is a continuous infusion pump, what are side effects, and what are the warning signs. The videos aim to provide patients with clear and concise explanations of medical procedures and key concepts to empower them in managing their care.

Additionally, the app includes digital questionnaires for reporting physical and psychological symptoms, as well as assessing quality of life. To enhance patient-clinician communication, the app integrates an instant messaging system, facilitating real-time interaction with the clinical team.

XSL•FO
RenderX

Figure 1. "CONTIGO" mHealth app Interface: (**A**) login screen, (**B**) home screen, (**C**) patient profile, (**D**) the journey section, (**E**) quality of life questionnaire, (**F**) PRO-CTCAE toxicity detection questionnaire, (**G**) symptom reporting module, (**H**) educational content (text and images), (**I**) educational content (videos), and (**J**) notes section.



Validation and Quality Assessment

In this study, the validation strategy reported in TGUM-GISSIC groups was applied and adapted to the current Chilean context in Chile. Technical attributes that determine the quality of an application, such as performance, graphical interface structure, usability, robustness, and information security, were identified [39]. Recommendations followed from the assessment and accreditation of health-related mobile applications [20]. The literature search began by identifying a validated questionnaire, and among the identified instruments, the MARS scale was selected.

The performance and robustness of the application were evaluated using Lighthouse [40,41], a widely used, open-source Google-developed tool for assessing web and mobile app quality. Lighthouse provides metrics across several categories, such as performance, accessibility, best practices, and search engine optimization (SEO), which are essential for understanding the application's functionality and user experience.

Lighthouse simulates a page load in a controlled environment, emulating device and network conditions to capture performance data and generate metrics for performance. These insights highlight areas needing improvement, ultimately helping developers enhance the user experience.

Its metrics include:

- First contentful paint (FCP): Measures the time it takes for the first visible content, such as text or images, to appear on the screen after the page starts loading. It provides an early indication to users that the page is loading and responsive, enhancing perceived performance. Optimal values are under 1.8 seconds, moderate values range between 1.8 and 3.0 seconds, and values above 3.0 seconds indicate a need for improvement.
- Largest contentful paint (LCP): Represents the time taken for the largest visible content, like a banner image or block of text, to fully render in the viewport. It reflects how quickly the main content becomes available, which is critical for user engagement. Optimal values are under 2.5

XSL•FO
RenderX

seconds, moderate values range between 2.5 and 4.0 seconds, and values above 4.0 seconds require improvement.
- Speed index: It evaluates how quickly content is visually displayed during the loading process by calculating the average time for visible portions of the page to appear. This metric gives a holistic view of the loading experience, beyond individual milestones. Optimal values are under 3.4 seconds, moderate values range between 3.4 and 5.8 seconds, and values above 5.8 seconds indicate poor performance.
- Total blocking time (TBT): Measures the total duration when long tasks (exceeding 50 ms) block the main thread, preventing the app or website from responding to user inputs. This metric directly impacts the perceived interactivity of the app during its loading phase. A lower TBT score indicates smoother transitions and better responsiveness. TBT scores were categorized based on established benchmarks: scores under 200 milliseconds were considered optimal, scores between 200 and 600 milliseconds as moderate, and scores above 600 milliseconds as requiring improvement.
- Cumulative layout shift (CLS): Assesses the visual stability of a page by measuring unexpected layout shifts caused by late-loading elements like images or ads. A low CLS score ensures a smooth and stable user experience, avoiding accidental clicks or frustrating movements of content. Optimal CLS values are under 0.1, while scores between 0.1 and 0.25 indicate moderate stability. Values above 0.25 reflect significant layout instability, which can negatively impact user experience.

The MARS scale was adapted and validated in Spanish to evaluate other attributes of "CONTIGO" critically and systematically [30-32]. This scale consists of 23 items, offering a reliable, flexible, multidimensional, Likert-type tool, with scores ranging from 1 (worst) to 5 (best). It comprises 3 sections. The first one evaluates unbiased quality, consisting of 4 dimensions (engagement, functionality, aesthetics, and quality of information). The second one assesses subjective quality, and the last one evaluates with 6 questions designed the perceived impact of the application on the users and their behavior change.

Application Evaluation

For the evaluation process of the "CONTIGO" mobile application, eligibility criteria were established for potential evaluators. This aims to ensure that evaluators possess the necessary expertise to provide meaningful feedback on the application's functionalities and effectiveness in supporting patients. The evaluators were selected from convenience samples that included nurses with master's degrees or specialization in oncology, computer engineers with prior experience in mobile application development, and professionals working in various oncology settings who have at least 3 years of practical experience. The sample included professionals from institutions directly involved in the app's development as well as external institutions. This approach ensured a diversity of perspectives during the evaluation process.

Evaluators who were directly involved in the development of the "CONTIGO" mobile application were excluded.

The unit of analysis for this research comprises the responses collected through the application of the MARS, provided by clinical oncology experts and computer engineers who evaluated the quality of the mobile application designed for patients with cancer.

The guidelines from the CHERRIE Checklist were followed for the survey application in Google Form [42].

Procedure for MARS Questionnaire Evaluation

Professionals meeting the criteria were invited to participate in the evaluation. During the evaluation period, which spanned approximately 1 month, each evaluator had independent access to the application to explore its features and functionalities at their convenience.

To ensure consistent understanding of the evaluation criteria, each evaluator received an initial training session with a researcher. During this session, researchers explained the elements of the MARS tool using official explanatory slides from the tool's authors. After this, evaluators independently used the app and completed the MARS rating questionnaire anonymously through a Google Form in Spanish. Informed consent was obtained before submission.

To further support the evaluators, researchers were available to answer questions or clarify doubts during the evaluation period. However, researchers could not access the responses while evaluators completed the questionnaire, ensuring the anonymity and impartiality of the process.

Data Analysis

The free-text comments provided by evaluators were analyzed using thematic analysis, following Braun and Clarke's [43] approach. This method allowed us to systematically identify recurring patterns and key themes from the qualitative feedback.

To enhance the reliability of the analysis, 2 independent researchers reviewed the comments and performed an open coding process. The initial codes were then discussed and refined through a consensus-based approach to consolidate the final themes. The analysis focused on 3 main areas: design and usability, content and features, and user onboarding.

No qualitative analysis software was used; instead, a manual coding process was conducted to ensure a direct focus on the key user experience concerns highlighted by the evaluators.

Results

App Development

During the early stages of app development, focus groups with health care professionals identified a key challenge: newly diagnosed oncology patients often forget or lack access to the educational information provided during consultations. This finding reinforced the need for a digital tool that allows patients to revisit educational materials at their convenience. As a result, the app was designed to integrate structured educational content

XSL•FO
RenderX

through text, images, and videos, ensuring accessibility and continuous support for patients.

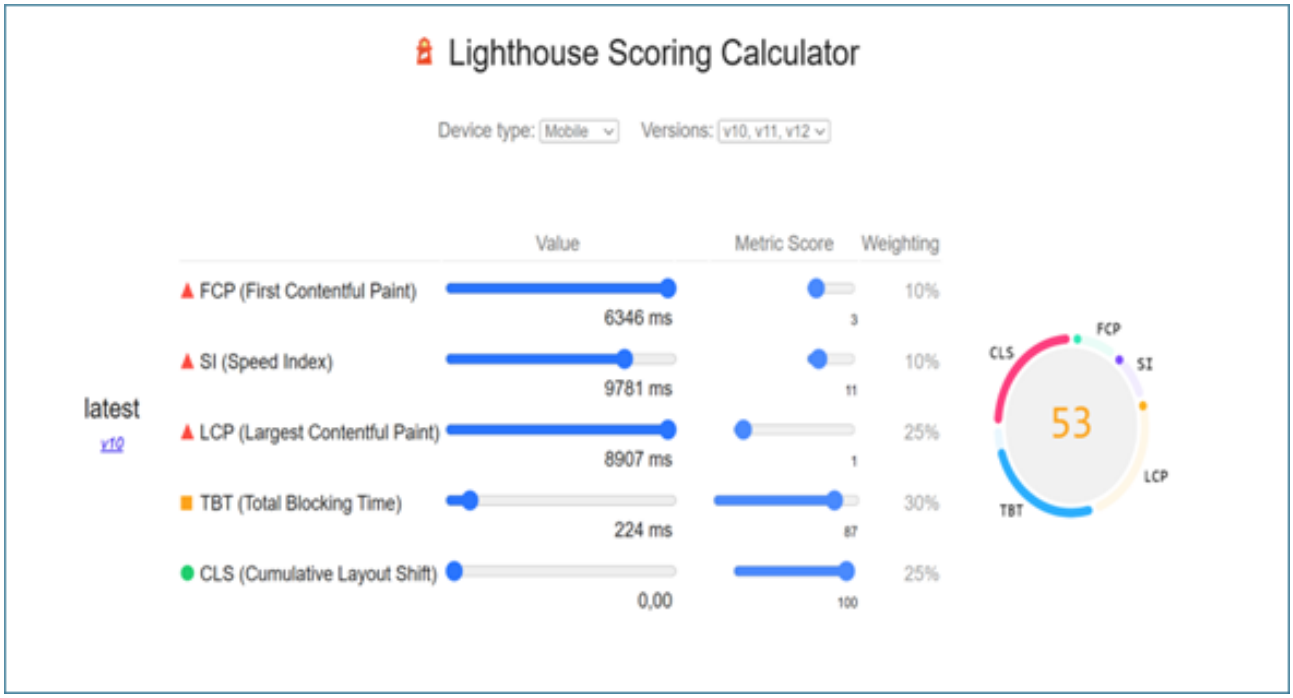
Performance and Robustness

The performance evaluation highlights areas requiring improvement, with an FCP of 6.3 seconds and an LCP of 8.9 seconds, indicating significant delays in rendering the initial and largest visible content. The Speed Index (SI) of 9.8 seconds reflects slow visual load times, impacting the perceived user experience. Despite these results, the TBT of 224 milliseconds falls within the moderate range, suggesting acceptable interactivity during loading. Visual stability, measured by CLS, scored 0, ensuring a smooth and stable interface (Figure 2).

In addition to these detailed performance metrics, the app was evaluated more broadly across categories such as Accessibility, Best Practices, and SEO (Figure 3). Accessibility scored 86, reflecting good compliance with accessibility standards. Best Practices achieved a perfect score of 100, indicating full adherence to technical and security guidelines, while SEO scored 83, highlighting solid SEO with some room for improvement. Figure 3 provides a general overview of the app's technical quality, complementing the specific performance insights presented in Figure 2.

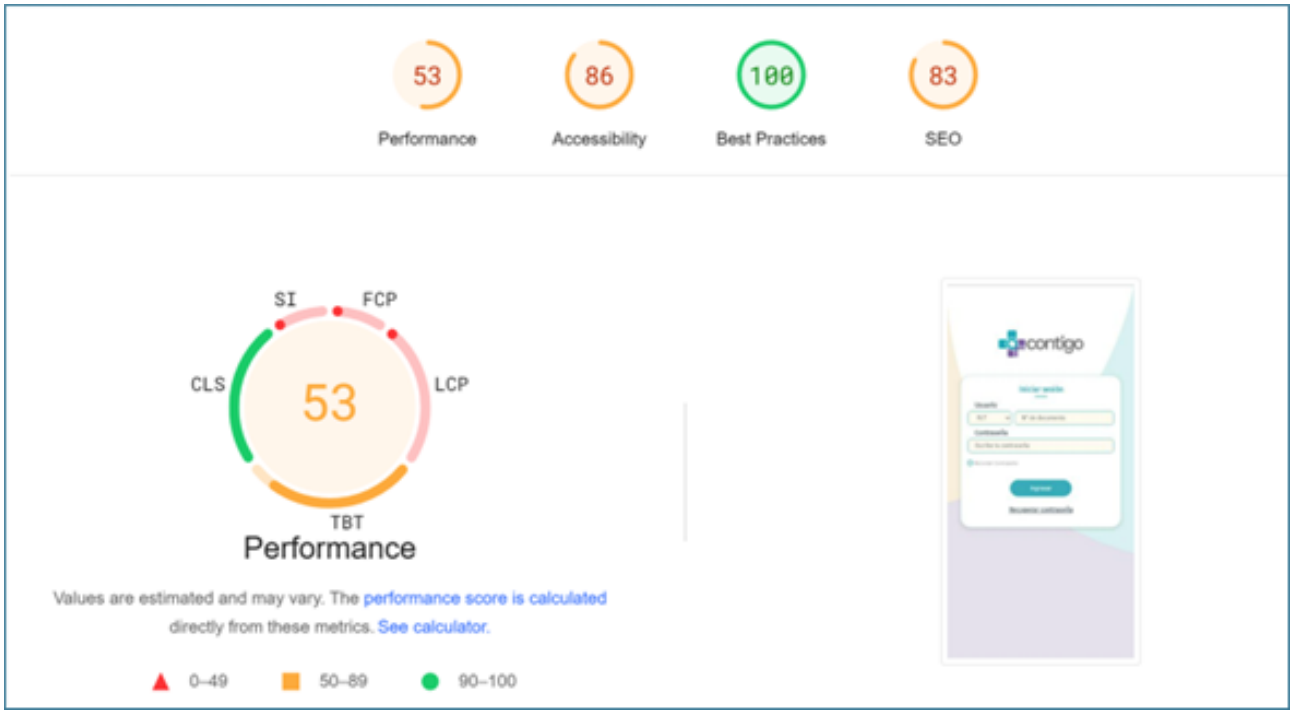
It is important to note that the data presented in Figures 2 and 3 were obtained using automated testing tools, specifically the Lighthouse tool (Chrome Web Store). These evaluations were conducted to simulate realistic usage conditions and assess the app's technical performance, including speed, visual stability, and compliance with industry standards.

Figure 2. Performance (Source: Chrome Web Store. Lighthouse).



XSL•FO
RenderX

Figure 3. Performance and robustness CONTIGO mobile application.



Security

The security certificate was issued by a trusted entity, "R3." This certificate indicates the validity of both server and client authentication in secure communication. A series of hexadecimal characters evidence the certificate's integrity, its fingerprint. With this certificate, the server identity can be identified when the user uses the mobile application, thereby thwarting phishing attempts and guaranteeing those users connect to the correct server.

Mobile App Rating Scale (MARS)

The MARS, which consists of 3 sections and 4 dimensions, was used to evaluate other technical aspects of quality, including content, usability, functionality, operability, accessibility, and reliability.

In total, 13 experts assessed the CONTIGO app: 6 were clinical staff, including nurses and doctors, and 7 were informatics engineers. In terms of educational level, 23.1% held a PhD, 15.4% were specialists, 30.85% had a master's degree, and 30.8% only had an undergraduate degree. The mean age of the experts was 41 years, providing a balanced representation of professionals across different stages of their careers.

The overall quality of the app, evaluated as the mean MARS score, was 3.75 out of 5 (SD 0.42) (Table 1). Next, we proceeded to assess each of the 4 dimensions of the MARS. The mean score for objective quality was 3.82 (SD 0.42), whereas subjective quality was lower at 3.38 (SD 0.72), with a higher variation as expected. The objective quality comprises 4 dimensions. Functionality had the highest score at 4.35 (SD 0.52), followed by information at 3.87 (SD 0.45), aesthetics at 3.85 (SD 0.59), and engagement with an average score of 3.31 (SD 0.61). Lastly, the app's impact on habits of life received a score of 3.96 (SD 0.55).

To see if engineers and clinicians had different opinions about the app, we analyzed the statistics for each group individually. We did not find significant differences in the scores given by these 2 groups of evaluators (Figure 4).

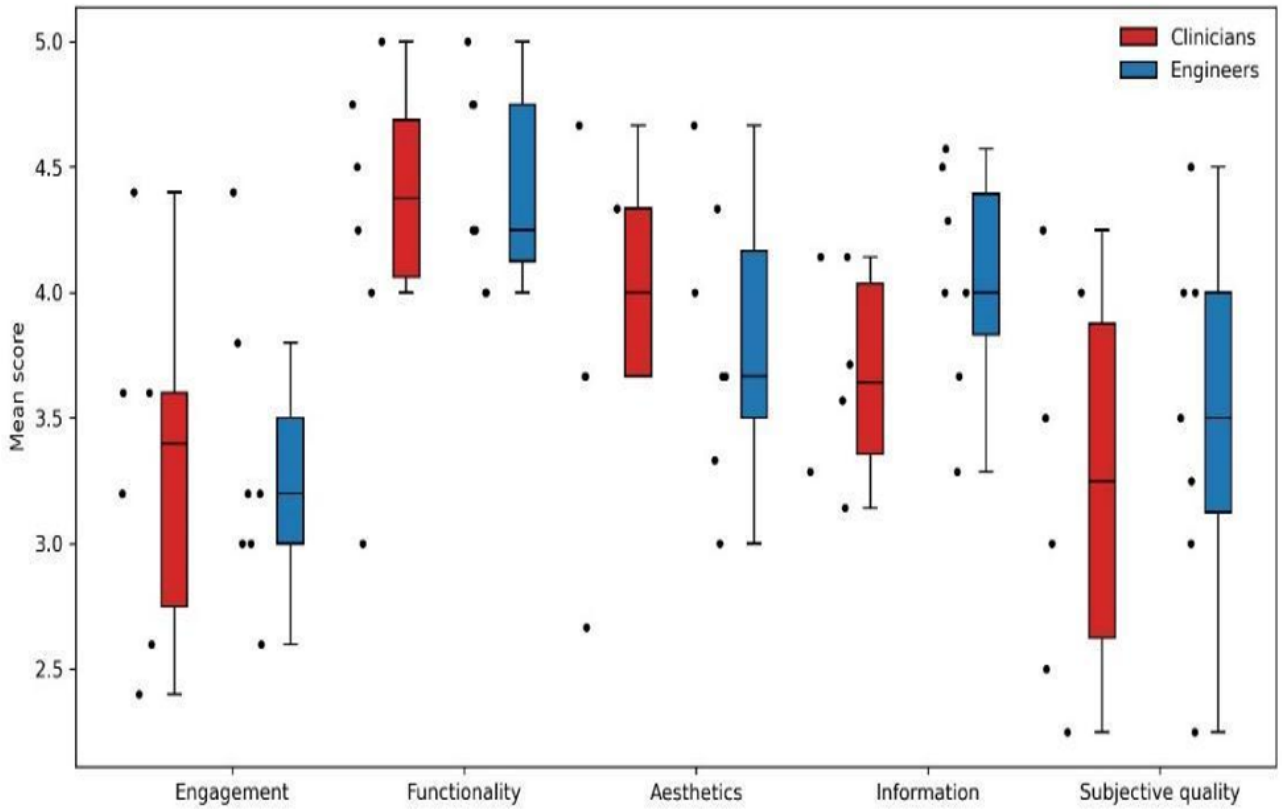
XSL•FO
RenderX

Table . Mobile application rating scale of CONTIGO.

	Mean	SD	IQR
Objective quality	3.82	0.42	0.48
Engagement	3.31	0.61	0.60
Entertainment	3.08	0.83	1.00
Interest	3.46	1.08	1.00
Customization	2.85	1.23	1.00
Interactivity	3.23	0.70	0.00
Target group	3.92	0.47	0.00
Functionality	4.35	0.52	0.75
Performance	4.69	0.61	0.00
Ease of use	4.23	0.58	1.00
Navigation	4.08	0.92	1.00
Gestural design	4.38	0.62	1.00
Aesthetics	3.85	0.59	0.67
Layout	3.77	0.89	1.00
Graphics	3.77	0.70	1.00
Visual appeal: How good does the app look?	4.00	0.68	0.00
Information	3.87	0.45	0.57
Accuracy of the app description	3.38	1.27	1.00
Goals	3.36	1.23	1.50
Quality of information	4.08	0.47	0.00
Quantity of information	4.00	0.55	0.00
Visual information	4.08	0.73	1.00
Credibility	4.08	0.83	2.00
Evidence base	4.00	0.00	0.00
Subjective quality	3.38	0.72	1.00
Would you recommend this app?	3.85	1.17	2.00
How many times do you think you would use this app?	3.54	0.75	1.00
Would you pay for this app?	2.38	1.21	2.00
What is your overall star rating of the app?	3.77	0.80	1.00
Habits of life	3.96	0.55	0.33
Awareness	4.15	0.77	1.00
Knowledge	4.38	0.74	1.00
Attitudes	3.69	0.72	1.00
Intention to change	3.92	0.83	0.00
Seek of help	3.77	0.70	0.00
Behavior change	3.85	0.53	0.00

XSL•FO
RenderX

Figure 4. Boxplot comparing clinicians versus engineers.



Consequently, with the above, the MARS demonstrated good reliability among evaluators with an interclass correlation coefficient (0.84; 95% CI 0.72 - 0.92)). The total score demonstrated high internal consistency (Cronbach α=.86). Although the MARS has been previously validated, reporting these metrics ensures that the evaluators in this study demonstrated consistent scoring. The Pearson correlation revealed a moderate but significant correlation between the MARS star rating item (#23), r(13)=0.657, *P*<.01.

Upon completing each questionnaire, experts were allowed to voluntarily provide comments about the app. Unanimously, they agreed that the application's design would be highly beneficial for both patients and clinical professionals, particularly in aiding in the early detection of symptoms.

The free-text comments were formally recorded and analyzed using thematic analysis to identify key areas for improvement. A total of 12 comments were collected, and 3 main themes emerged:

- Design and usability: Suggestions included enhancing font readability, improving navigation, and ensuring consistency in icon usage across different sections. For instance, one evaluator noted: "The font size is too small in some sections, making it difficult to read. Consistent font sizes for text would improve usability."
- Content and features: Evaluators recommended adding more detailed symptom tracking options and tailoring educational content to different user needs. A clinical evaluator suggested: "The app could include fields for tracking specific chemotherapy cycles and whether the treatment is curative or palliative."
- User onboarding: Some participants highlighted the need for better guidance for first-time users, such as an introductory tutorial. One comment stated: "Adding a quick tutorial for new users would help them navigate the app more effectively."

Discussion

Principal Findings and Comparison With Previous Works

This article describes the development and initial evaluation of a mobile application designed to support oncology patients undergoing chemotherapy and assesses its quality. The findings emphasize the importance of interdisciplinary collaboration and iterative design processes in health care app development. These processes ensure that the app's content is robust, its usability is optimized, and it aligns with clinical and educational standards [37].

Therefore, the results provide a detailed insight into the performance and overall quality of the application. While the app demonstrated strengths in areas such as accessibility (86) and adherence to best practices (100), it also highlighted areas for improvement in performance metrics, including FCP (6.3 s) and LCP (8.9 s), which exceeded optimal thresholds. The TBT scored within the moderate range (224 ms), reflecting acceptable interactivity, and the CLS scored 0, ensuring excellent visual stability. Overall, the application adheres to crucial quality standards [20,39,40] but underscores the need for further optimizations to enhance perceived user performance and loading times. These findings emphasize the importance of iterative testing and refinement during app development.

XSL•FO
RenderX

One important usage of MARS is to evaluate which areas of the app require more work to provide the final user with a better experience. In that line, we reviewed the score of each item and looked for the reasons behind that score, taking into account comments given about the app.

Of the app's 4 subjective dimensions, engagement received the lowest overall score, indicating significant room for improvement. This was largely due to limited customizability (2.85), as the app lacks options for important capabilities such as color and font size customization, which are particularly useful for users with reduced vision or color blindness, as well as direct access to functionality through favorite shortcuts.

Additionally, the entertainment subdimension scored 3.08, reflecting the app's reliance on informative videos as its main source of entertainment and the absence of gamification elements or other engaging digital features. However, the interest subdimension scored slightly higher at 3.46, suggesting that the app maintained a moderate level of engagement among evaluators despite these shortcomings. Concerning the functionality, the app showed the highest score. Evaluators appreciated its good performance (4.69), gestural design (4.38), ease of use (4.23), and ease of navigation (4.08). The aesthetics of the app (3.85) were evaluated moderately high; some users noticed that there were some inconsistencies in the layout, graphics, and fonts.

The app performed well in the Information dimension (3.87), especially in items such as the quality of information (4.08), visual information (4.08), credibility (4.08), quantity of information (4.00), and evidence base (4.00). Its performance in terms of the accuracy of the app description (3.38) and its goals (3.36), on the other hand, was much lower. The app was not available on the app store; the description provided in the Google form questionnaire given to the evaluators was taken into account for the assessment. While some users were highly positive, stating they would recommend the app to everyone and rating it as one of the best they have used, others were more reserved or critical.

The subjective quality of the application varied among users based on their recommendations. Even though the app received a good score in terms of how likely evaluators would recommend the app (3.38), at the same time it received a much lower score in terms of how likely evaluators would pay for the app (2.38).

The results also showed that the evaluators consider that app would likely have a good impact on the habits of life (3.96). The evaluators provide a critical foundation for understanding the app's strengths and areas for improvement [20,30,31]. This study highlights the value of involving health care professionals in the early stages of app evaluation, as their expertise ensures alignment with clinical standards and educational goals. Additionally, it is crucial to involve patients to capture the unique needs and preferences of the app's end-users.

To address this limitation, 2 ongoing clinical trials will directly involve patients with cancer [44,45]. These trials will evaluate the app using the uMARS tool, which focuses on patient-centered aspects such as usability, engagement, and impact on quality of life. By integrating patient feedback, these trials aim to refine the app further, ensuring it meets the specific needs of its target population.

Despite this limitation, the study contributes to the field of health apps by demonstrating the importance of phased evaluation approaches. These ensure that health apps are rigorously tested for technical and clinical standards before involving end-users, creating a strong foundation for iterative improvement.

Conclusions

This study provides an initial evaluation of the CONTIGO app, a tool designed to support oncology patients undergoing chemotherapy by addressing key aspects of symptom tracking, education, and patient-clinician communication.

The findings demonstrate that the app meets critical standards for functionality, usability, and information quality, supported by robust security measures and exceptional visual stability. However, performance evaluations revealed areas requiring improvement, particularly in loading times and user interactivity, emphasizing the need for further technical optimizations to enhance the user experience.

The evaluation also highlighted opportunities to improve user engagement and aesthetic consistency through features like gamification and personalization, which could significantly enhance the app's adoption and impact.

A significant limitation of this study was the lack of patient involvement in the evaluation phase. This decision prioritized input from health care professionals to ensure alignment with clinical and educational standards. Ongoing clinical trials involving patients with cancer are addressing this limitation by capturing patient-specific feedback on usability and impact using the uMARS tool.

This study contributes to the growing field of mHealth by underscoring the importance of phased evaluations that integrate iterative testing and stakeholder feedback. Future work will focus on refining the app based on patient insights, ensuring it effectively meets its target population's needs and improves the oncology care experience.

Limitations and Future Directions

The primary limitation of this study was the exclusion of patients with cancer during the initial evaluation phase. This decision was based on the need to ensure the app met clinical and educational standards before involving end-users. Health care professionals were selected as evaluators due to their expertise in oncology care and their ability to provide detailed feedback on the app's content and usability.

Efforts were made to recruit patients for this phase through oncology clinics and patient support groups. However, challenges such as scheduling conflicts, the burden of ongoing treatments, and limited interest in participating in a non-clinical evaluation restricted patient involvement. Additionally, it was not possible to recruit patients who were not already participating in other clinical trials, as this could have introduced confounding variables to this study.

XSL•FO
RenderX

To address this limitation, 2 clinical trials are currently underway, incorporating the uMARS tool to assess the app's usability and quality from the patient perspective. These trials will provide a comprehensive understanding of the app's impact and inform future refinements.

By prioritizing health care professionals in the early phases, we ensured a solid foundation for the application's content and usability, which will now be validated and refined based on direct input from patients.

Acknowledgments

This work was funded by the ANID FONDAP (152220002), Center for the Cancer Prevention and Control (CECAN), and Convocatoria número 22 del Sistema General de Regalías (SGR) y Ministerio de Ciencia, Tecnología e Innovación de Colombia (MINCIENCIAS).

The researchers extend their gratitude to the entire engineering team at the Technological Institute for Health and Wellness Innovation, which was consistently willing to address challenges—particularly Engineer Katherine Figueroa. They also wish to acknowledge the nurses and pharmacists who, in various capacities, contributed to the development of the application: Lillian Macaya, Patricia Pérez, and Antonio Gonzalez.

Conflicts of Interest

None declared.

References

1. World Health Organization. Noncommunicable diseases. : WHO; 2023. URL: https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases [accessed 2024-01-25]
2. Ministerio de Salud. Departamento de Epidemiología Second report on population-based cancer registries in Chile 1998-2019.: Ministerio de Salud; 2023. URL: https://epi.minsal.cl/wp-content/uploads/2023/04/Segundo-Informe-Registros-Poblacionales-de-Cancer_Febrero2023.pdf [accessed 2024-01-31]
3. Ministerio de Salud. National Cancer Plan 2018-2028.: Ministerio de Salud; 2018. URL: https://www.minsal.cl/wp-content/uploads/2019/01/2019.01.23_PLAN-NACIONAL-DE-CANCER_web.pdf [accessed 2024-01-31]
4. SEREMI. Premature years of life lost due to malignant tumors in the Metropolitan Region.: SEREMI; 2019. URL: https://seremi13.redsalud.gob.cl/wrdprss_minsal/wp-content/uploads/2014/09/AVPP-Tumores-Malignos-2019.pdf [accessed 2024-01-31]
5. Departamento de Estadisticas e Informaciòn de Salud (DEIS). Cancer mortality [Website in Spanish]. : DEIS-MINSAL; 2020. URL: https://deis.minsal.cl/ [accessed 2024-01-25]
6. Global Cancer Observatory. New global cancer data. : GLOBOCAN; 2020. URL: https://gco.iarc.fr/today/data/factsheets/populations/152-chile-fact-sheets.pdf [accessed 2024-01-25]
7. Chang KH, Brown KGM, Chen Lau Y, Solomon MJ. Excellent outcomes after extended radical pelvic resection for locally advanced and recurrent IBD-associated anorectal cancer. Dis Colon Rectum 2021 Feb 1;64(2):209-216. [doi: 10.1097/DCR.0000000000001744] [Medline: 33315717]
8. Bruns CJ. Multimodale Therapie des Ösophaguskarzinoms [Multimodal treatment of esophageal cancer]. Chirurg 2021 Dec;92(12):1075-1076. [doi: 10.1007/s00104-021-01534-2] [Medline: 34842948]
9. Chevallay M, Wassmer CH, Iranmanesh P, Jung MK, Mönig SP. Multimodal treatment in oligometastatic gastric cancer. World J Gastrointest Oncol 2022 Feb 15;14(2):434-449. [doi: 10.4251/wjgo.v14.i2.434] [Medline: 35317315]
10. Ghadimi M, Rödel C, Hofheinz R, Flebbe H, Grade M. Multimodal treatment of rectal cancer. Dtsch Arztebl Int 2022 Aug 22;119(33-34):570-580. [doi: 10.3238/arztebl.m2022.0254] [Medline: 35791271]
11. Gugenheim J, Crovetto A, Petrucciani N. Neoadjuvant therapy for pancreatic cancer. Updates Surg 2022 Feb;74(1):35-42. [doi: 10.1007/s13304-021-01186-1] [Medline: 34628591]
12. IAHPC International Association For Hospice & Palliative Care. Palliative care definition. 2022. URL: https://hospicecare.com/what-we-do/projects/consensus-based-definition-of-palliative-care/definition/ [accessed 2022-10-22]
13. Milazzo S, Hansen E, Carozza D, Case AA. How effective is palliative care in improving patient outcomes? Curr Treat Options Oncol 2020 Feb 5;21(2):12. [doi: 10.1007/s11864-020-0702-x] [Medline: 32025964]
14. El-Jawahri A, LeBlanc TW, Kavanaugh A, et al. Effectiveness of integrated palliative and oncology care for patients with acute myeloid leukemia: a randomized clinical trial. JAMA Oncol 2021 Feb 1;7(2):238-245. [doi: 10.1001/jamaoncol.2020.6343] [Medline: 33331857]
15. Organización Mundial de la Salud. Atención integrada para personas mayores (ICOPE) Marco de aplicación: orientación para los sistemas y servicios [Integrated care for older people (ICOPE) implementation framework: guidance for systems and services]. 2020. URL: https://iris.who.int/handle/10665/337374 [accessed 2024-01-24]

XSL•FO
RenderX

16. de la Rosa Oliva F, Meneses García A, Ruiz Calzada H, et al. Effects of omega-3 fatty acids supplementation on neoadjuvant chemotherapy-induced toxicity in patients with locally advanced breast cancer: a randomized, controlled, double-blinded clinical trial. Nutr Hosp 2019 Aug 26;36(4):769-776. [doi: 10.20960/nh.2338] [Medline: 31192682]
17. Statista. Forecast number of mobile users worldwide from 2020 to 2025. : Statista; 2023. URL: https://www.statista.com/statistics/218984/number-of-global-mobile-users-since-2010 [accessed 2023-11-30]
18. Statista. Number of apps available in leading app stores as of 3rd quarter 2022. 2022. URL: https://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/ [accessed 2023-11-30]
19. Baxter C, Carroll JA, Keogh B, Vandelanotte C. Assessment of mobile health apps using built-in smartphone sensors for diagnosis and treatment: systematic survey of apps listed in international curated health app libraries. JMIR Mhealth Uhealth 2020 Feb 3;8(2):e16741. [doi: 10.2196/16741] [Medline: 32012102]
20. Martín Fernández A, Marco Cuenca G, Salvador Oliván JA. Evaluación y acreditación de las aplicaciones móviles relacionadas con la salud [Evaluation and accreditation of mobile health applications]. Rev Esp Salud Pública 2020;94:e202008085 [FREE Full text]
21. Raeesi A, Khajouei R, Ahmadian L. Evaluating and rating HIV/AIDS mobile apps using the feature-based application rating method and mobile app rating scale. BMC Med Inform Decis Mak 2022 Oct 30;22(1):281. [doi: 10.1186/s12911-022-02029-8] [Medline: 36310157]
22. Lau N, O'Daffer A, Yi-Frazier JP, Rosenberg AR. Popular evidence-based commercial mental health apps: analysis of engagement, functionality, aesthetics, and information quality. JMIR Mhealth Uhealth 2021 Jul 14;9(7):e29689. [doi: 10.2196/29689] [Medline: 34259639]
23. FDA. Device software functions including mobile medical applications. : FDA; 2022. URL: https://www.fda.gov/medical-devices/digital-health-center-excellence/device-software-functions-including-mobile-medical-applications#a [accessed 2024-02-08]
24. Google. Play Store mobile application. 2024. URL: https://play.google.com/store/search?q=cancer&c=apps&hl=es_419&gl=US [accessed 2024-01-30]
25. Roche Farma SA. Mobile apps to support cancer care. 2024. URL: https://www.rocheplus.es/innovacion/tecnologia/apps-paciente-oncologicos.html [accessed 2024-01-30]
26. Ferrero NA, Morrell DS, Burkhart CN. Skin scan: a demonstration of the need for FDA regulation of medical apps on iPhone. J Am Acad Dermatol 2013 Mar;68(3):515-516. [doi: 10.1016/j.jaad.2012.10.045] [Medline: 23394920]
27. González de León B, León Salas B, del Pino-Sedeño T, Rodríguez-Álvarez C, Bejarano-Quisoboni D, Trujillo-Martín MM. Aplicaciones móviles para mejorar la adherencia a la medicación: revisión y análisis de calidad [Article in Spanish]. Atención Primaria 2021 Nov;53(9):102095. [doi: 10.1016/j.aprim.2021.102095] [Medline: 34139398]
28. Wasserman S, Ould Brahim L, Attiya A, Belzile E, Lambert SD. An evaluation of interactive mHealth applications for adults living with cancer. Curr Oncol 2023 Jul 25;30(8):7151-7166. [doi: 10.3390/curroncol30080518] [Medline: 37622999]
29. Red de Salud UC Christus. Red de salud UC christus [Website in Spanish]. 2025. URL: https://www.ucchristus.cl [accessed 2025-01-01]
30. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27. [doi: 10.2196/mhealth.3422] [Medline: 25760773]
31. Martin Payo R, Fernandez Álvarez MM, Blanco Díaz M, Cuesta Izquierdo M, Stoyanov SR, Llaneza Suárez E. Spanish adaptation and validation of the mobile application rating scale questionnaire. Int J Med Inform 2019 Sep;129(129):95-99. [doi: 10.1016/j.ijmedinf.2019.06.005] [Medline: 31445295]
32. Romero RL, Kates F, Hart M, Ojeda A, Meirom I, Hardy S. Quality of deaf and hard-of-hearing mobile apps: evaluation using the mobile app rating scale (MARS) with additional criteria from a content expert. JMIR Mhealth Uhealth 2019 Oct 30;7(10):e14198. [doi: 10.2196/14198] [Medline: 31670695]
33. Diaz-Skeete YM, McQuaid D, Akinosun AS, Ekerete I, Carragher N, Carragher L. Analysis of apps with a medication list functionality for older adults with heart failure using the mobile app rating scale and the IMS Institute for healthcare informatics functionality score: evaluation study. JMIR Mhealth Uhealth 2021 Nov 2;9(11):e30674. [doi: 10.2196/30674] [Medline: 34726613]
34. Stifani BM, Peters M, French K, Gill RK. There's an app for it: a systematic review of mobile apps providing information about abortion using a revised MARS scale. PLOS Digit Health 2023 Jul;2(7):e0000277. [doi: 10.1371/journal.pdig.0000277] [Medline: 37459287]
35. Martín-Martín J, Muro-Culebras A, Roldán-Jiménez C, et al. Evaluation of android and Apple store depression applications based on mobile application rating scale. Int J Environ Res Public Health 2021 Nov 27;18(23):12505. [doi: 10.3390/ijerph182312505] [Medline: 34886232]
36. Gasca Mantilla MC, Camargo Ariza LL, Medina Delgado B. Methodology for mobile application development. Tecnura 2014;18(40):20-35 [FREE Full text] [doi: 10.14483/udistrital.jour.tecnura.2014.2.a02]
37. Sistema Nacional de Evaluación A y C de la CE (Sineace). Guide for the elaboration and application of focus groups G-DEP-005. : Sineace; 2020. URL: https://repositorio.sineace.gob.pe/repositorio/bitstream/handle/20.500.12982/6434/Gu%C3%ADa%20para%20la%20elaboracion%20de%20grupo%20focal.pdf [accessed 2024-02-08]

XSL•FO
RenderX

38. Taramasco C, Rimassa C, Noël R, Bravo Storm ML, Sánchez C. Co-design of a mobile app for engaging breast cancer patients in reporting health experiences: qualitative case study. J Med Internet Res 2023 Nov 27;25:e45968. [doi: 10.2196/45968] [Medline: 38010791]
39. Serna Edgar. Development and Innovation in Engineering: torial UPB; 2016, Vol. 1. URL: https://www.researchgate.net/profile/Edgar-Serna-M/publication/329221691_Desarrollo_e_Innovacion_en_Ingenieria_ed_1/links/5bfd9cd0a6fdcc35428c9589/Desarrollo-e-Innovacion-en-Ingenieria-ed-1.pdf
40. Chrome Web Store. Lighthouse. : Chrome Web Store; 2023. URL: https://chromewebstore.google.com/detail/blipmdconlkpinefehnmjammfjpmpbjk [accessed 2024-02-11]
41. Google Trust Services LLC. Google certificates viewer. : Google Trust Services LLC; 2024. URL: https://pki.goog [accessed 2024-02-11]
42. Eysenbach G. Improving the quality of web surveys: the checklist for reporting results of internet e-surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34. [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
43. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
44. Martínez F, Taramasco C, Espinoza M, Acevedo J, Goic C, Nervi B. Differential effects of a telemonitoring platform in the development of chemotherapy-associated toxicity: a randomized trial protocol. Diagnostics (Basel) 2024 Mar 14;14(6):619. [doi: 10.3390/diagnostics14060619] [Medline: 38535039]
45. Martínez F, Taramasco C, Espinoza M, Acevedo J, Goic C, Nervi B. Effects on quality of life of a telemonitoring platform amongst patients with cancer (EQUALITE): a randomized trial protocol. Methods Protoc 2024 Mar 15;7(2):24. [doi: 10.3390/mps7020024] [Medline: 38525782]

Abbreviations

CLS: cumulative layout shift
DALYs: disability-adjusted life years
DDoS: distributed denial-of-service
FCP: first contentful paint
ICC: interclass correlation coefficient
LCP: large contentful paint
MARS: Mobile App Rating Scale
PREMs: Patient-Reported Experience Measures
PROMs: Patient-Reported Outcome Measures
SEO: search engine optimization
TBT: total blocking time
WHO: World Health Organization

Edited by A Kushniruk; submitted 03.05.24; peer-reviewed by L McCann, N Ribeiro; revised version received 31.01.25; accepted 02.02.25; published 28.04.25.

Please cite as:
Muñoz Olivar C, Pineiro M, Gómez Quintero JS, Avendaño-Vásquez CJ, Ormeño-Arriagada P, Palma Rivadeneira S, Taramasco Toro C
Education and Symptom Reporting in an mHealth App for Patients With Cancer: Mixed Methods Development and Validation Study
JMIR Hum Factors 2025;12:e60169
URL: https://humanfactors.jmir.org/2025/1/e60169
doi:10.2196/60169

© Carolina Muñoz Olivar, Miguel Pineiro, Juan Sebastián Gómez Quintero, Carlos Javier Avendaño-Vásquez, Pablo Ormeño-Arriagada, Silvia Palma Rivadeneira, Carla Taramasco Toro. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 28.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Mobile Therapeutic Attention for Treatment-Resistant Schizophrenia (m-RESIST) Solution for Improving Clinical and Functional Outcomes in Treatment-Resistant Schizophrenia: Prospective, Multicenter Efficacy Study

Jussi Seppälä[1,2], MD, PhD; Eva Grasa[3,4], PhD; Anna Alonso-Solis[4,5], PhD; Alexandra Roldan-Bejarano[3,4], MD, PhD; Marianne Haapea[2,6,7], PhD; Matti Isohanni[2], MD, PhD; Jouko Miettunen[2,6], PhD; Johanna Caro Mendivelso[8], MD, PhD; Cari Almazán[8], MD; Katya Rubinstein[9], PhD; Asaf Caspi[9], MD; Zolt Unoka[10], MD; Kinga Farkas[10], MD, PhD; Elisenda Reixach[11], PhD; Jesus Berdun[12], MSc; Judith Usall[4,13,14], MD, PhD; Susana Ochoa[4,13,14], PhD; Iluminada Corripio[4,15,16*], MD, PhD; Erika Jääskeläinen[17*], MD, PhD; m-Resist Group[18]

[1]Social Insurance Institution of Finland, Suokatu 40, Kuopio, Finland

[10]Department of Psychiatry and Psychotherapy, Semmelweis University, Budapest, Hungary

[11]TicSalut Health Department, Generalitat de Catalunya, Barcelona, Spain

[12]Digital Health Unit, Hospital de la Santa Creu i Sant Pau, Institut d'Investigació Biomèdica Sant Pau IIB SANT PAU, Barcelona, Spain

[13]Parc Sanitari Sant Joan de Déu, Sant Boi de Llobregat, Spain

[14]Etiopatogènia i tractament dels trastorns mentals greus MERITT, Institut de Recerca Sant Joan de Déu, Esplugues de Llobregat, Spain

[15]Mental Health and Psychiatry Department, Vic Hospital Consortium, Vic, Spain

[16]Mental Health, Institut de Recerca Sant Pau IR SANT PAU, Sant Quintí 77-79, Barcelona, Spain

[17]Unit of Population Health, University of Oulu, Oulu, Finland

[18]

[2]Research Unit of Population Health, University of Oulu, Oulu, Finland

[3]Mental Health, Institut de Recerca Sant Pau IR SANT PAU, SantQuintí 77-79, Barcelona, Spain

[4]Centro de Investigación Biomédica en Red de Salud Mental CIBERSAM, Madrid, Spain

[5]Mental Health Division, Fundació Althaia, Xarxa Assistencial Universitaria de Manresa, Manresa, Spain

[6]Medical Research Center Oulu, Oulu University Hospital and University of Oulu, Oulu, Finland

[7]Department of Psychiatry, Oulu University Hospital, Oulu, Finland

[8]Agency for Health Quality and Assessment of Catalonia AQuAS, Barcelona, Spain

[9]The Gertner Institute of Epidemiology and Health Policy Research, Sheba Medical Center, Tel Aviv University, Tel Aviv, Israel

[*]these authors contributed equally

Corresponding Author:
Jussi Seppälä, MD, PhD
Social Insurance Institution of Finland, Suokatu 40, Kuopio, Finland

Abstract

Background: Treatment-resistant schizophrenia (TRS) is a severe form of schizophrenia associated with low adherence to treatment and poor outcomes. Mobile health (mHealth) interventions may be effective in preventing relapses, increasing treatment adherence, and managing some of the symptoms of schizophrenia. Mobile therapeutic attention for treatment-resistant schizophrenia (m-RESIST) is an innovative mHealth developed specifically for TRS.

Objective: We aim to evaluate the effects of m-RESIST on the clinical and functional outcomes and on the perceived quality of life in people with TRS.

Methods: A feasibility study without a control group was performed to test the m-RESIST solution on patients with TRS. Participants were recruited from Spain, Israel, and Hungary. This study's population (N=31) followed 3 months of intervention. The m-RESIST was configured by an app, a wearable, and a web-based platform. The severity of symptoms was evaluated by using the Positive and Negative Syndrome Scale (PANSS) and the Clinical Global Impression-Schizophrenia (CGI-SCH) scale. Functionality was assessed by the Global Assessment of Functioning and perceived quality of life was evaluated by the EuroQol visual analogue scale (EQ-VAS).

XSL•FO
RenderX

Results: Significant reductions were found in symptoms from pretrial to posttrial on the PANSS total (mean difference −7.2, 95% CI −11.1 to −3.4; *P*=.001), the PANSS positive (mean difference −1.36, 95% CI −2.6 to −0.1; *P*=.04), the PANSS negative (mean difference −2.1, 95% CI −3.1 to −1.1; *P*<.001), and the PANSS general symptoms (mean difference −3.8, 95% CI −6.8 to −0.8; *P*=.02). In almost one-fifth of the participants (6/31), the overall score for the PANSS decreased by more than 20%, which may be considered a clinically significant change. On the CGI-SCH scale, the sum of total severity of illness decreased significantly (*P*=.03). A decrease in the sum of positive and negative symptoms of the CGI-SCH score was also found (*P*=.04 and *P*=.03, respectively). The sum of depressive or cognitive symptoms did not change. The functionality of participants increased significantly on the Global Assessment of Functioning (*P*≤.001). The perceived quality of life on the EQ-VAS also improved (mean difference 6.7, 95% CI 0.5 to 12.9; *P*=.04).

Conclusions: To our best knowledge, this was the first study to address the efficacy of the mHealth app m-RESIST on the symptoms and functional capacity and on the quality of life for people with TRS. Our preliminary findings showed that implementing the m-RESIST solution decreased the symptoms and severity of disease, and improved the functionality and perceived quality of life among those with TRS. The change of symptoms on the PANSS total may be clinically significant. Modern technologies such as mHealth interventions may be useful in treating symptoms and functionality even in TRS, which is a major clinical challenge, with usually poor outcomes. These results should be corroborated by performing a controlled trial.

Trial Registration: ClinicalTrials.gov NCT03064776; https://clinicaltrials.gov/study/NCT03064776

International Registered Report Identifier (IRRID): RR2-10.1136/bmjopen-2017-021346

(JMIR Hum Factors 2025;12:e67659) doi:10.2196/67659

KEYWORDS

digital health; mHealth; m-RESIST; application; treatment-resistant schizophrenia; psychosis; feasibility study; mobile phone; mobile health; mobile therapeutic attention for treatment-resistant schizophrenia

Introduction

Background

Schizophrenia is a severe mental disorder and has a prevalence of 1% of the population worldwide [1]. A meta-analysis showed that only 13.5% of the patients were in recovery, which included clinical remission and good social functioning outcome [2]. Despite the common use of antipsychotic drugs, 30% - 50% of patients with schizophrenia still obtained a low benefit from conventional treatments [3,4]. This inadequate response to pharmacological treatment is known as treatment-resistant schizophrenia (TRS). TRS implicates an important burden at three levels, as follows: (1) clinical (negative attitude to medication, drug abuse, and nutritional or physical health problems); (2) economic (hospitalizations and polytherapy); and (3) humanistic (depression and social isolation) [5]. As a result, dealing with TRS involves a high emotional burden for patients and their caregivers, affecting their quality of life. The current health care and social support systems are unable to provide adequate and effective solutions to these patients.

Results from earlier studies on psychosis [6-8] supported that information and communication technology (ICT) tools in the treatment of patients with TRS would be useful. Smartphone ownership among people with schizophrenia is also relatively high and increasing [9-11]. Patients seemed capable of using mobile phones to monitor their symptoms, participate in therapeutic interventions, and increase physical exercise [12]. Previous feasibility studies have shown that interventions based on mobile health (mHealth) could promote the empowerment of patients with schizophrenia [9,13]. Some mHealth interventions may also be effective for increasing treatment adherence and decreasing relapses [13,14].

Several digital solutions also specifically for schizophrenia exist [15,16]. Apps such as FOCUS [6,17-19], SMARTapp [20], MACS [21], weCope [22], My Journey 3 [23,24], and App4Independence [25] offer disease management strategies in general. Others are more focused on symptom monitoring, namely Crosscheck [26], Actissist [27], ClinTouch [28], ExPRESS [29], Ginger.io [30], ReMindCare [31], and MindFrame [32]. Other mobile apps are supported in cognitive behavioral therapy (SlowMo) [33]; TechCareApp [34]; movisensXS [35]; PRIME [36], or work as an agenda (MONEO) [37], or target specific domains such as social skills (MASS) [38], sleep (ExpiWell) [39], medication adherence (MedActive) [40], or delivers positive psychology content (+ Connect) [41]. Digitally delivered psychological therapies, the use of social media, virtual reality, and chatbots are also digital solutions that may have a role in the treatment of those with schizophrenia [42,43]. Digital phenotyping and machine learning, virtual reality interventions, and language models, with ChatGPT, as a well-known example, have developed in recent years [42,43].

A systematic review and meta-analysis by Vitger et al [44] examined digital shared decision-making interventions in mental health care, analyzing 16 randomized controlled trials (RCTs), of which only 3 evaluated apps. The findings suggest that digital shared decision-making tools may positively impact patient activation, decisional conflict, working alliance, and symptom severity. Specific digital interventions include a tablet app for patients with schizophrenia, designed to assess and record satisfaction with life and treatment domains, facilitating discussions between patients and nurses [45]. Another tablet app addresses concerns raised by patients with psychosis, providing methods for clinicians and patients to collaboratively explore and manage issues [46]. Additionally, the smartphone app “Power Up” aims to promote patient activation among children and adolescents facing mental health challenges [47].

XSL•FO
RenderX

Patient-reported outcomes collect data directly from patients. These data are used in clinical practice, helping decision-making. A systematic mapping review analyzed the characteristics and outcomes of mHealth interventions in psychosis based on 29 publications reporting on 23 studies [48]. Objective measures included data passively obtained from phone usage or embedded sensors. Subjective measurement instruments include either observer-rated or self-report questionnaires that can be written or administered digitally, whether regarding their health status or experience using the mHealth intervention [48]. Outcomes were grouped into 8 categories, namely severity of psychosis-related symptoms, functioning, well-being, medication adherence, adverse events, user experience, technical, and all other outcomes reported [48].

Real-time and in-context patient-generated symptom data, obtained through remote-monitoring platform technology, have the potential to timely warn clinicians about the need for intervention, improve treatment decisions by providing a clearer picture of changing patterns of symptoms, and support scheduling of health care contacts based on need [42].

Systematic comparison of apps with other digital interventions may help us to better understand their roles in mental health care. The main use of smartphone data (active and passive) is tracking mood and lifestyle in people with major depression, bipolar disorder, and psychosis. They also have future potential in machine learning toward individualized risk prediction and delivery of targeted “just in time” intervention. Key issues are a lack of validation across studies and establishing trust around data usage. Priority actions include data standards for interoperability and validation, and industry-academic partnerships around access [42].

Digitally delivered psychological therapies are used in the self-management of symptoms of depression and anxiety. They have the potential for precision interventions and preventive treatments. Lack of engagement, saturated consumer marketplace, and claims outpacing clinical evidence are major issues. Establishing an evidence base for use in people with diagnosed mental disorders is warranted [42].

Population level monitoring of mood and anxiety is possible with the help of social media. The future potential of social media lies on its potential for real time monitoring of mental health state and for access to peer support. Key issues are sampling bias, access to data from social media companies, and also privacy. Priority actions are industry-academic partnerships and privacy standards [42].

The main use of virtual reality is exposure therapies, and it may have higher engagement and potentially higher efficacy than apps. Increased accessibility, low-cost headsets, and expanded clinical targets are important [42]. Chatbots have future potential to increase access to care but they have limited range of appropriate response. Establishing evidence base for use in people with diagnosed mental disorders is needed [42].

The 5 principles of the Collaborative Care Model are evidence-based care, measurement-based treatment to target, population health management, accountable care, and team-based care [49]. Practices with integrated mental health clinicians can maximize the impact of digital mental health tools by integrating digital information and interventions into the course of clinical care. This allows for the practice of hybrid care, a combination of synchronous and asynchronous health care, which can enable increased access to care. Hybrid care can address the known challenge of low retention in digital health tools, given that a digital navigator or clinician can promote engagement, guide the patient to relevant content, and, depending on the tool, use data from the patient’s use in further guiding face-to-face care [49].

The findings from the pilot study suggest that smartphone apps may become a useful tool for psychiatric rehabilitation, addressing both psychiatric and co-occurring medical problems. Individualizing functions to each patient and facilitating connection with a certified peer specialist may be an important feature of useful apps [50].

Apps can be successfully integrated into real-world care [43]. These models of care involve both synchronous (eg, a therapy visit) and asynchronous telehealth (ie, app use between visits) and can provide therapeutic synergy between in-person and remote care [51]. The ReMindCare app for early psychosis has been integrated into clinics since 2018, and implementation models for integrating apps in youth services for schizophrenia, and coordinated specialty, have already been developed [31,52-54]. While the actual dissemination of these models is impossible to ascertain, they can already support the use of apps to improve vocational training, educational opportunities, medication adherence, cognitive training, symptom tracking, psychoeducation, and more. Apps for schizophrenia have also led to the development of human support roles as part of mental health services to facilitate their implementation in clinical practice, such as the FOCUS mHealth support specialist [55] and the mindLAMP digital navigator [53].

A renewed focus on privacy and data protection is essential for the success of mental health apps, as users must trust these apps for widespread adoption. Many health-related apps collect sensitive data, including demographic, medical, and lifestyle information, sometimes even unrelated to the app’s intended purpose [42,43,56,57]. Despite the importance of privacy, a significant portion of mental health apps (31% - 49%) lack a privacy policy, and collected data can be shared with third parties or used for advertising [42,43,56,57]. Vulnerable populations, such as individuals with schizophrenia, require additional research protections to ensure their privacy and security [43].

A cross-cultural psychological assessment focuses on understanding how a client’s cultural background influences their behavior, emotional state, and psychological needs. A recent review of 2248 studies found that only 57 specifically targeted migrants, refugees, or cultural minorities [58]. Of these, half used digital health technologies developed for the general population without cultural adaptations, while the other half tailored the interventions to meet the specific needs of these populations [58]. However, only a small number of studies involved a high level of participation from these groups in the development and testing of the digital tools [58].

XSL•FO
RenderX

Despite rapid progress, there is room for improvement. Presently, there are fewer than 10 apps whose design was specifically tailored to people with schizophrenia and that are accessible for immediate public download on the Apple or Android (Open Handset Alliance [led by Google]) marketplaces. Moreover, many of these 10 apps appear neglected by developers after their launch, with an average time since the last update of 1121 days [43,59]. The number of studies on the efficacy of mHealth apps on the symptoms and functional capacity, and on the quality of life in schizophrenia and related psychotic disorders is limited according to recent systematic reviews [15,56,60,61].

The systematic review by Miralles et al [56] found 3 studies evaluating efficacy of mHealth solutions in schizophrenia spectrum and in other psychotic disorders. Two of the three papers found decrease in symptoms measured by the Positive and Negative Syndrome Scale (PANSS) [6,27]. The third one reported no significant change in the symptoms on the PANSS, or no positive changes in functioning and in quality of life [36].
.

Additionally, a cross-sectional study on WeChat use and endorsement of WeChat-based mHealth among people living with schizophrenia (N=400) in China was carried out. Compared to nonusers, WeChat use was associated with improved health outcomes, including lower psychiatric symptoms, lower depression, higher functioning, better recovery, and higher quality of life [62].

A recent study of patients with a diagnosis of schizophrenia, randomized to receive either SlowMo, a blended digital therapy targeting reasoning, or usual care, found no significant difference between groups on the primary outcomes related to paranoia at 24 weeks (the primary end - point) [33].

The Mobile Therapeutic Attention for Treatment-Resistant Schizophrenia Project

The project of mobile therapeutic attention for treatment-resistant schizophrenia (m-RESIST), funded by the Horizon 2020 Programme of the European Union, designed and developed a mobile therapeutic attention for TRS. This digital solution consisted of a new and innovative mHealth solution based on novel technology and offering highly modular and flexible functioning. This solution was developed specifically for patients with TRS and their caregivers based on their needs [63-65].

The positive acceptance of the m-RESIST solution was related to its usefulness in meeting user needs, its capacity to empower them, and its possibility of maintaining human contact [63].

The only mHealth study performed in those with TRS indicated that the m-RESIST solution is feasible and is an acceptable, satisfactory, and potentially useful tool for a population with TRS [65].

Aims

In this study, our aim was to evaluate the effects of the m-RESIST solution on the clinical and functional outcomes and the perceived quality of life for those with TRS. To our best knowledge, there are no previous studies which have addressed these issues.

Methods

Study Design and Setting

This was a prospective, 3-month long, multicenter follow-up study without a control group that was carried out with patients with a diagnosis of TRS between March and November 2017. Blinding was not considered due to the absence of a comparator group. No sample size calculation was done.

In this study, the intervention involved patients and their caregivers, and the main actors involved in the deployment of the m-RESIST solution were psychiatrists, psychologists, and case managers. The key aim of the m-RESIST solution was to engage patients with TRS, together with their caregivers, in therapeutic processes, and empower them to enable their self-management.

Participants were recruited from 3 clinical sites: Sant Pau Hospital, Semmelweis University, Sheba Medical Center, and Gertner Institute of Epidemiology and Health Policy Research. These sites provided full-spectrum mental health services to adult people from their catchment area and have extensive experience in participating in large-scale clinical trials in patients with schizophrenia.

Patients underwent a 3-month long modular intervention designed for patients with TRS. The intervention was based on the recognition of early warning signs of psychosis to improve positive symptoms, treatment adherence, and a healthy lifestyle.

The m-RESIST Solution

The intervention was supported by 3 mHealth tools: a wearable (smartwatch), a mobile app, and a web-based platform, which are presented in the figure of the protocol article. Data for usage was captured automatically by the m-RESIST software [64].

The functionality of the smartphone was based on the m-RESIST app. Through this app, patients had access to educational content about TRS condition and related issues; could track their early warning signs symptoms and biological variables (eg, sleep and steps counter); could ask for help with questionnaires or the “alarm button”; receive and practice helpful CBT-based coping strategies and exchange messages with their caregiver or health care provider [64].

The wearable was a smartwatch that collected data from patients and sent it wirelessly to the smartphone. Sensor data was recorded through automatic passive upload. The variables collected were the level of activity, heart rate, sleeping pattern, and steps counter [64].

The web-based platform was the tool that the health care providers (case manager, psychiatrist, and psychologist) used to collect assessment data, to monitor patients’ state and review data collected by sensors, to communicate by texting with patients, caregivers, and other professionals, and to consult recommendations (based on guidelines and experts’ opinion) [64]. Mental health care professionals helped patients with TRS in using the intervention [64]. The m-RESIST intervention was

XSL•FO
RenderX

provided in 5 languages: Spanish, Catalan, Hungarian, Hebrew, and English.

The trial was registered in Clinical Trials (NCT03064776), and the study protocol was published earlier [64].

Staff or Participant Training Protocols

Three user guides were prepared to ensure that all m-RESIST users (patients, caregivers, and clinicians) will be able to understand and follow the instructions of the devices or system used in home or health care facilities. The dashboard user guide explained quickly and simply the applications and functionalities that clinicians can perform when logging in. The mobile application installation guide was addressed to clinicians in case they need to install the m-RESIST solution in patients' devices. This guide details the steps to download the m-RESIST application and install it on the smartphone and the smartwatch. The mobile application user guide was addressed to patients, caregivers, and clinicians and described the functionalities of the m-RESIST app working together with the smartwatch, once the app was installed on the Android smartphone provided. This user guide also specifies briefly the operations and the use of the smartwatch given to the patient at the beginning of this study. Finally, a short instruction sheet was designed. This sheet contained the main steps to charge smartphone and smartwatch devices and the main considerations to take into account to ensure devices were connected and worked properly.

In case of the PANSS, raters at each site had to perform a training to ensure liability between sites.

Language Adaptation Process

To ensure the language adaptation process, the clinical research team from each site was responsible for adapting the m-RESIST solution from English to Spanish, Catalan, Hungarian, and Hebrew.

Technical Support Procedures

To solve and improve the issues of the m-RESIST solution during the pilots, incidences were collected by the clinicians and communicated by email to the technical team. Thus, they could fix them.

Ethical Considerations

This study protocol received approval from the ethics committees of the 3 participating institutions: Hospital Santa Creu i Sant Committee (IIBSP-RES-2016 - 51), Egészségügyi Tudományos Tanács Tukeb, Semmelweis Ethical Committee (54920-4/2016/EKU), and Sheba Medical Center Ethical Committee (3472 - 16 SMC). Participants were fully informed of the purpose and procedures of this study, and signed informed consent forms were obtained before any study-related activities were performed. The database generated by this study did not contain any identification of the participants but kept only a numerical code in a separate list; thus, the patients' identities were protected. The information collected in this study was always treated as grouped data and never as individual or personal data, thereby maintaining anonymity and confidentiality. No compensation was provided to the participants.

Study Participants

Patients with TRS were recruited for this study. The eligibility criteria were (1) patients aged 18 - 45 years with a diagnosis of schizophrenia according to *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* criteria; (2) duration of disease less than 15 years to meet the criteria for TRS [66,67]; (3) familiarity with ICT tools; and (4) physical capability to use them; and presence of a caregiver. Exclusion criteria were (1) remission according to the Remission of Schizophrenia Working Group [68]; (2) presence of delusions mainly related to their therapists or with new technologies; (3) presence of vision, hearing, or motor impairment, interfering with operating a smartphone; (4) presence of a caregiver or informal carer who is not used to ICT tools or has physical incapability to use them; and (5) presence of intellectual developmental disability. The caregivers were also recruited as part of the m-RESIST project.

A total of 81 patients with TRS, with their caregivers, were selected for invitation to participate by researchers. Almost half of them (n=39) were willing to enroll, and nearly 80% (n=31) of those started this study and followed the whole 3 months of intervention. The dropout rate was 18% (n=8).

Measures in This Study

All the outcomes of this study's protocol are shown in Table 1.

XSL•FO
RenderX

Table . Schedule of study protocol periods and assessments.

	Recruitment	Preintervention	Intervention						Follow-up
Activity or assessment		V0	V1	V2	V3	V4	V5	V6	V7
Eligibility screen	✓								
Informed consent	✓								
Delivery and training of devices	✓								
Sociodemographic data		✓							
Clinical characteristics		✓							
Relapse signature and treatment plan		✓							
Willingness to enroll	✓								
PANSS[a]		✓							✓
CGI-SCH[b]		✓	✓	✓	✓	✓	✓	✓	✓
GAF[c]		✓							✓
EQ-VAS[d]		✓							✓

[a]PANSS: Positive and Negative Syndrome Scale.

[b]CGI-SCH: Clinical Global Impression Schizophrenia.

[c]GAF: Global Assessment of Functioning.

[d]EQ-VAS: EuroQol visual analogue scale.

In this study, the severity of symptoms was assessed using the PANSS and the Clinical Global Impression-Schizophrenia (CGI-SCH) scale [69,70].

- The PANSS was used to assess positive and negative symptoms of schizophrenia. This scale consists of 30 items rated from 1 to 7 according to increasing levels of psychopathology, where: 1=absent, 7=extreme. Four scores are obtained: positive, negative, general symptoms, and total. The PANSS was completed on 2 occasions, at the start (VO) and the end (V7) of the intervention.
- The CGI-SCH scale measures global clinical impression and consists of 5 different ratings (positive, negative, depressive, cognitive, and global) that are rated from 1 to 7, where: 1=normal, 7=among the most extremely ill patients. the CGI-SCH scale was evaluated in all protocol visits (V0-V7).

Functional and Perceived Quality of Life Outcomes

Functionality was assessed by using the Global Assessment of Functioning (GAF) scale [71] and perceived quality of life by the EuroQol visual analogue scale (EQ-VAS) [72].

- The GAF scale ranges from 1 (maximum level of disease) to 100 (the healthiest). It was performed on 2 occasions, at VO and at V7.
- The EQ-VAS is scored on a 20 cm visual analogue scale with end points labeled “the best health you can imagine” and “the worst health you can imagine.” It was administered on 2 occasions, at V0 and V7.

Adverse Events

Clinician asked for adverse events at each pilot visit, and they were reported. Safety measures were collected throughout this study. The presence of serious and nonserious adverse events, defined as any clinical change or illness reported during this study, was monitored in every clinical visit. The adverse events observed when carrying out this study, either by the clinician or by the patient themselves and regardless of the causality relationship ascribed, were recorded in the clinical records and on the patient`s dashboard.

Statistical Analysis

The sociodemographic characteristics of the participants with TRS were presented using frequency distributions for categorical variables and mean with SD for age. The outcomes in baseline

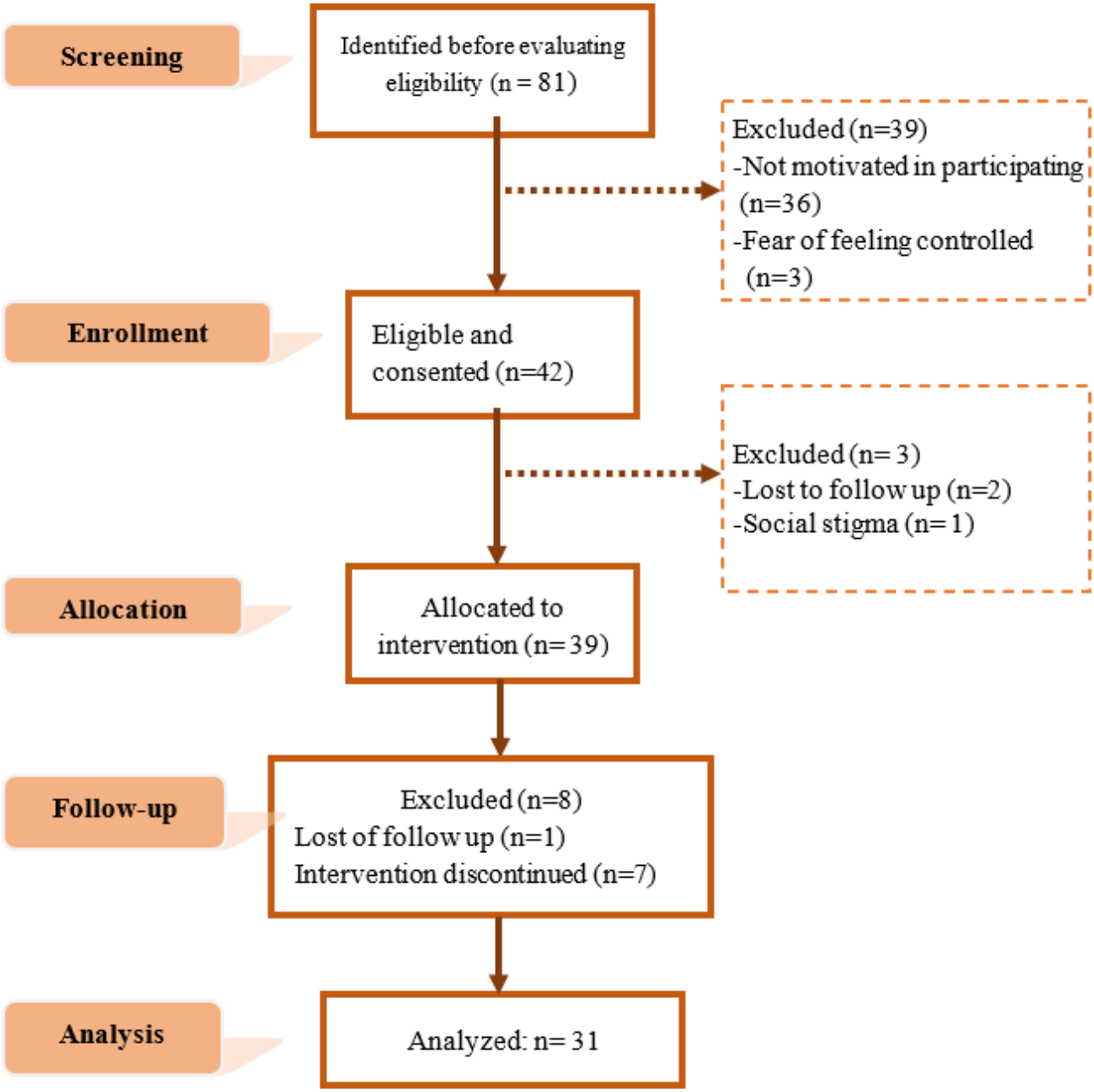
XSL•FO
RenderX

and follow-up were presented using mean with SD for PANSS scores and EQ-5D, and medians with IQR for CGI-SCH and GAF scores. The change from the baseline to the follow-up was analyzed using the paired *t* test (PANSS scores and EQ-5D) or the Wilcoxon test (CGI-SCH and GAF scores). Statistical analyses were performed using the SPSS Statistics (version 29.0; IBM Corp).

Results

The study participation flowchart based on the CONSORT (Consolidated Standards of Reporting Trials) 2010 statement recommendations [73] is presented in Figure 1.

Figure 1. m-RESIST study participation flowchart. m-RESIST: mobile therapeutic attention for treatment-resistant schizophrenia.



Background Variables

The sociodemographic characteristics of this study's population (n=31) are presented in Table 2. No statistically significant differences in sociodemographic characteristics were found at baseline between the participants in the 3 sites. The percentage of men and women was almost equal (n=16, 51.6% and n=15, 48.4%, respectively). Most participants were singles (n=25, 80.6%) and lived with parents or relatives (n=23, 74.2%). Almost half of the study population had a full-time or a half-time job (n=15, 48.4%). The most common education level was primary studies (n=15, 48.4%). Almost two of the third lives with parents or relatives (n=20, 64.5%). Nearly all study participants had a caregiver (n=28, 90.3%; Table 2). Those patients without a caregiver had support in their immediate environment.

Nearly all participants were currently on 1 or 2 antipsychotics (30/31, 96%), and clozapine was the most used drug (11/31, 35.5%). Olanzapine, clozapine, and risperidone were previously

XSL•FO
RenderX

the 3 most frequently taken antipsychotics (13/31, 41.9%; 9/31, 29%; and 9/31, 29%, respectively).

Table . The sociodemographic variables of this study's population.

	Values
Age, mean (SD)	33.9 (7.6)
Gender, n (%)	
Male	16 (51.6)
Female	15 (48.4)
Marital status, n (%)	
Single	25 (80.6)
Married or with a stable partner	4 (12.9)
Divorced or separated	2 (6.5)
Cohabitation, n (%)	
Lives alone	4 (12.9)
Lives with parents or relatives	23 (74.2)
Lives with a stable partner	3 (9.7)
Lives with roommates	1 (3.2)
Occupation, n (%)	
Full-time job	7 (22.6)
Part-time job	8 (25.8)
Unemployed	9 (29)
Pension	6 (19.4)
Protected job	1 (3.2)
Educational level, n (%)	
Primary studies	15 (48.4)
Secondary studies	10 (32.3)
University	6 (19.4)
Lives with parents or relatives[a], n (%)	
No	8 (25.8)
Yes	20 (64.5)
Caregiver availability[a], n (%)	
No	2 (6.5)
Yes	28 (90.3)

[a]Missing n=1.

Clinical Outcomes

Significant reductions were found in symptoms from pretrial to posttrial on the PANSS total (mean 73.7, SD 14.8 versus mean 66.5, SD 12.8; mean difference −7.2, 95% CI −11.1 to −3.4; *P*=.001), PANSS positive (mean 15.2, SD 2.8 versus 13.8, SD 3.2; mean difference −1.36, 95% CI −2.6 to −0.1; *P*=.04), PANSS negative (mean 21.2, SD 4.9 versus 19.1, SD 4.3; mean difference −2.1, 95% CI −3.1 to −1.1; *P*<.001), and PANSS general symptoms (mean 37.3, SD 9.3 versus 33.6, SD 8.4; mean difference 3.8, 95% CI −6.8 to −0.8; *P*=.02) (Table 3).

In almost one-fifth of the participants (6/31), the overall score for the PANSS decreased by more than 20%.

The sum of global severity of illness also decreased significantly on the CGI-SCH scale (median 4, IQR 3-4 vs median 3, IQR 3-4; median difference 0, IQR −1 to 0; *P*=.03). A decrease in the sum of positive (median 3, IQR 2.75-4 vs median 3, IQR 2-3.25; median difference 0, IQR −1 to 0; *P*=.04) and negative (median 4, IQR 3-4 vs median 3, IQR 3-4; median difference 0, IQR −1 to 0; *P*=.03) symptoms was also found. The sum of depressive or cognitive symptoms did not change (Table 3).

XSL•FO
RenderX

Table . Change of symptoms between baseline and follow-up.

		Baseline	Follow-up	Difference	*P* value
PANSS[a,b]					
	Total score	73.7 (14.8)	66.5 (12.8)	−7.2 (−11.1 to −3.4)	.001
	Positive symptoms score	15.2 (2.8)	13.8 (3.2)	−1.36 (−2.6 to −0.1)	.04
	Negative symptoms score	21.2 (4.9)	19.1 (4.3)	−2.1 (−3.1 to −1.1)	<.001
	Generalized symptoms score	37.3 (9.3)	33.6 (8.4)	−3.8 (−6.8 to −0.8)	.02
CGI-SCH[c,d]					
	Sum of positive symptoms	3 (2.75 to 4)	3 (2 to 3.25)	0 (−1 to 0)	.03
	Sum of negative symptoms	4 (3 to 4)	3 (3 to 4)	0 (−1 to 0)	.03
	Sum of depressive symptoms	3 (2 to 3.25)	3 (2 to 3)	0 (−1 to 0.5)	.22
	Sum of cognitive symptoms	3 (2.75 to 4)	3 (2 to 3)	0 (−1 to 0)	.23
	Sum of global severity	4 (3 to 4)	3 (3 to 4)	0 (−1 to 0)	.03
GAF[d,e]		55 (45 to 55)	60 (50 to 65)	5 (0 to 10)	<.001
EQ-VAS[b,f]		66.1 (18)	72.8 (16.5)	6.7 (0.5 to 12.9)	.04

[a]PANSS: Positive and Negative Syndrome Scale.

[b]Mean (SD) for the baseline and follow-up, and mean (95% CI) with paired samples *t*-test for the difference.

[c]CGI-SCH: Clinical Global Impression-Schizophrenia.

[d]Median (IQR) for the baseline and follow-up, and median (IQR) with Wilcoxon test for the difference.

[e]GAF: Global Assessment of Functioning.

[f]EQ-VAS: EuroQol visual analogue scale.

Functional and Perceived Quality of Life Outcomes

The functionality of participants increased significantly from pretrial to posttrial on the GAF scale (median 55, IQR 45-55 vs median 60, IQR 50-65; median difference 5, IQR 0-10; *P*<.001), but it was not considered clinically significant due to participants remaining in the same interval (moderate symptoms).

The perceived quality of life evaluated by the EQ-VAS also improved (mean 66.1, SD 18 vs mean 72.8, SD 16.5; difference 6.7, 95% CI 0.5-12.9; *P*=.04) (Table 3).

Adverse Events

The symptoms of TRS got worse in 1 participant before the intervention and in 2 participants during it, which was considered not to be related to the m-RESIST solution.

Discussion

Principal Results

To our best knowledge, this was the first study collecting data on the symptoms, psychosocial functioning, and quality of life for people with TRS by using the mHealth platform (m-RESIST solution). Our novel preliminary findings showed that the total, positive, negative, and general symptoms according to the PANSS were significantly decreased after intervention compared with the starting point. In almost one-fifth of the participants (6/31), the overall score for the PANSS decreased by more than 20%, which may be considered a clinically significant change [74].

According to Suzuki et al [75], the partial response of clinical trials in schizophrenia could be defined ≥10% to <20% decrease on the PANSS. In our study, the change in positive score was 8.9%. Interestingly, the change in the negative score was 10.2%, which may mean a partial response according to Suzuki`s definition. The decrease in the generalized symptoms scores and the total score was 10.2% versus 9.8%. Therefore, the mean differences of 1 - 2 points on the positive and negative subscale scores may be indicatively clinically meaningful to a small extent at the most. Nevertheless, one should be careful in interpretations regarding the significance of the results. The clinical implications of PANSS scores may not always be clear [76].

The sum of global severity, and the sum of positive and negative symptoms on the CGI-SCH scale were lower at the end of the study compared to the starting point, but the sum of depressive or cognitive symptoms did not change. The functionality measured by the GAF scale was also significantly better at the

XSL•FO
RenderX

end of this study as compared to its level at the beginning. The perceived quality of life also increased during this study.

Comparisons With Prior Work

The finding of a significant decrease in all PANSS scores and global severity, positive and negative symptoms on the CGI-SCH scale at the study end compared to the starting point is partly in line with former data, as previous studies have shown efficacy of mobile interventions for schizophrenia [6,15,27,36,43,56,61,62]. To the best of our knowledge, this is the first study to show the efficacy of an mHealth platform (m-RESIST solution) in TRS.

A study by Ben-Zeev et al [6] demonstrated the feasibility, acceptability, and efficacy of a smartphone intervention for schizophrenia, followed by those with schizophrenia or schizoaffective disorder (N=33) for 1 month without a control group. Its results showed a significant decrease in the total, positive, and general scores but not in the negative score in the PANSS. A noncontrolled study by Kidd et al [25] with patients with primary psychosis (N=38) reported a decrease in certain items of the Brief Symptom Inventory (eg, psychotism, paranoid ideation, and depression) when followed for 1 month.

Bucci et al [27] carried out an RCT among those with early psychosis (N=36). The results showed a decrease in the total, negative, and general scores of the PANSS after a follow-up of 12 weeks.

WeChat is the most widely and extensively used mobile social networking app in China. A cross-sectional study on WeChat use and endorsement of WeChat-based mHealth among people living with schizophrenia (n=400) in China was carried out. Compared to nonusers in this study's population, WeChat use was associated with improved health outcomes, including lower psychiatric symptoms (the Brief Psychiatric Rating Scale), lower depression (the Patient Health Questionnaire-9), higher functioning (GAF), better recovery (the Recovery Assessment Scale), and higher quality of life (the first 2 general questions from the World Health Organization Quality of Life Brief Scale) [62].

However, another RCT with participants with recent-onset schizophrenia spectrum disorders (N=43) showed no changes in their positive or negative score of the PANSS or their functioning measured by the Role Functioning Scale. The quality of life measured by the interview-based Quality of Life Scale Abbreviated did not improve [36]. A recent study with patients with a diagnosis of schizophrenia (N=361), randomized to receive either SlowMo, a blended digital therapy targeting reasoning, or usual care, found also no significant difference between groups on the primary outcomes related to paranoia (the Green et al [77] Paranoid Thoughts Scale) at the primary end point at 24 weeks [33].

Decrease in Symptoms

It should be taken into account that the decrease in symptoms in patients with TRS, characterized by the persistent presence of symptoms, would not be expected in such a short term. Following the considerations of this study by Ben-Zeev et al [6], the improvement in the PANSS scores obtained in this study of the m-RESIST solution should be evaluated prudently. In the case of patients with schizophrenia, in the framework of a study, a placebo effect of the offered intervention has been observed. One factor that would influence the improvement of symptoms would be all the care received for participating in a study. It could be considered that participation in a structured intervention program would promote a state of security, such that it would improve aspects such as the symptoms of the disorder itself [78,79].

However, the overall score for the PANSS decreased by more than 20% in almost one-fifth of the participants which may be considered a clinically significant change [74]. Furthermore, the negative symptoms on the PANSS decreased during a rather short period of 3 months. It may imply that somehow the m-RESIST solution may manage to activate patients.

Perceived Quality of Life and Functioning

The perceived quality of life and functioning improved in our study, which is partly in line with the earlier studies, as a study by Schlosser et al [36] did not find positive effects on functioning and quality of life. Instead, the findings of Yu et al [62] showed improvement both in functioning and quality of life. Our findings on the improvement of quality of life may need to be treated with caution as the findings of the EQ-5D-5L scale have shown weak correlations with the objective measures among patients with schizophrenia [80].

Study Timeline

The data collection period was in 2017. Although progress has been made, there is still room for improvement in mHealth apps for people with schizophrenia, with fewer than 10 apps specifically designed for this population available for public download [43,59]. The m-RESIST solution is a new, innovative mHealth tool developed for patients with TRS and their caregivers, offering modular and flexible functions. This study aimed to evaluate the effects of m-RESIST on clinical outcomes, functional outcomes, and perceived quality of life in individuals with TRS. To the best of our knowledge no previous studies have addressed these specific issues.

For the above reasons, our research results provide significant new information on the use of mobile applications in the treatment of TRS.

Strengths

This study has several strengths. To our best knowledge, this was the first study collecting data on the symptoms, functioning, and perceived quality of life among patients with TRS by an mHealth intervention. This was also a multicenter study covering different cultural environments. The m-RESIST solution is an innovative mHealth solution based on novel technology, offering highly modular and flexible functioning and planned intervention. It has been developed specifically for patients with TRS [63-65]. The outcomes of this study will also help to perform a cost-effectiveness RCT in the future.

Limitations

Some limitations of this study need to be discussed. First, there was no control group, so it is not possible to determine whether the clinical improvements were related to the use of the

XSL•FO
RenderX

m-RESIST solution. Second, this study's design is a limitation, as, due to the small sample size and the short length of the intervention and follow-up period, the differences found in outcomes should be treated with caution. Therefore, the generalization and maintenance of the results are uncertain. Third, the so-called Hawthorne effect may also have implications for the generalizability of results because the behavior of research participants is influenced by observation and measurement. In general, results in experimental settings may be more positive compared to routine clinical applications [78,79].

The moderately high mean baseline score for the PANSS (73.7) demonstrates that our patients with TRS have quite severe psychotic symptoms. Even so, they responded to intervention markedly.

Potential Selection Bias Regarding Digital Literacy

Digital literacy can be defined as the varying ability of both children and adults to use digital technologies and understand their risks. Patients who do not know how to use digital health tools, do not see the importance of those tools, or cannot access them in their preferred language, ultimately will not use them [81]. This may cause potential selection bias because, for the reasons above, those people may not participate in the studies [81].

Technical Implementation Challenges

The digital divide remains a barrier, with people with schizophrenia eager and able to obtain smartphones and learn new skills, but resources and training are often not provided. Expanding train-the-trainer and literacy programs such as the Digital Health Navigator initiative, which was specially designed to support people with psychosis, offers 1 tangible next step [42,43,53] However, a lagging infrastructure and skill base have limited the application of digital solutions in mental health care [82]. Additionally, evidence is needed on the effectiveness and appropriateness of digital health tools for patients who are marginalized and may lack access to digital health interventions [83].

The digital divide remains a barrier, with people with schizophrenia eager and able to obtain smartphones and learn new skills, but resources and training are often not provided [51]. Expanding train-the-trainer and literacy programs such as the Digital Health Navigator initiative, which was specially designed to support people with psychosis, offers 1 tangible next step [53,84].

Cost Implications for Real-World Implementation

Despite modest effect sizes, the relatively low cost and high scalability of most of the interventions tested support their public health relevance [85]. A recent review on this topic found only 37 relevant studies published in middle- and lower-income countries between 2016 and 2020, with the majority reporting feasibility and accessibility outcomes instead of efficacy, cost-effectiveness, or implementation [86]. However, the costs of implications for real-world implementation may be a concern, and that is why it has been asked if mHealth tools can decrease hospital stays and encourage patients to use mobile apps in favor of reducing costs [61]. Digital health tools still hold a unique potential to extend access to care in middle- and lower-income countries, where there is less investment or infrastructure around mental health [86]. Smartphone use is common and rapidly expanding in those countries, and thus represents a promising tool to reduce the mental health gap [42].

Conclusions

Our preliminary findings showed that implementing the m-RESIST solution may decrease the symptoms and severity of illness and improve the functionality and perceived quality of life among people with TRS. The change of symptoms on the PANSS scores may be indicatively clinically meaningful to a small extent at the most. Nevertheless, one should be careful in interpretations regarding the significance of the results.

Implementing the m-RESIST solution may be a promising tool to improve the treatment even in TRS, and to offer a new tool for its present treatment options, already in the near future. If successful, this would be a very important step forward as TRS is a major clinical challenge, with usually poor outcomes. Before being implemented in daily practice, these results should be corroborated by performing controlled clinical trials with people with TRS, comparing treatment as usual with the m-RESIST solution.

Acknowledgments

This work has been supported by the Horizon 2020 Framework Programme of the European Union (grant 643552) and partly funded by the CERCA (Centres de Recerca de Catalunya) Programme/Generalitat de Catalunya. We thank all the patients who participated in this study for their essential contribution to our field of research. We are also thankful to Ona Corporation (IIBSP-AMA2016-99). We are grateful to all the members of the mobile therapeutic attention for treatment-resistant schizophrenia (m-RESIST) project, who are also the authors of this paper under the name of the m-RESIST Group.

The m-RESIST Group is composed of the following members: Francisco Alcalde, Caritat Almazán, Anna Alonso-Solís, Jesús Berdún, István Bitter, Walter Baccinelli, Maria Bulgheroni, Caro Mendivelso, Asaf Caspi, Tanguy Coenen, Xavier Constant, Iluminada Corripio, Enrico d'Amico, Ilaria De Vita, Marisol Escobar, Kinga Farkas, Kata Fazekas, Shenja van der Graaf, Eva Grasa, Margarita Hospedales, Elena Huerta-Ramos, Matti Isohanni, Erika Jääskeläinen, Charlotte Jewel, Teija Juola, Timo Jämsä, Rachelle Kaye, Panagiotis Kokkinakis, Hannu J Koponen, Silvia Marcó, Gregoris Mentzas, Jouko Miettunen, Jani Moilanen, Susana Ochoa, Ilias Papas, Fotis Paraskevopoulos, Elisabeth Reixach, Katya Rubinstein, Elena Rubio-Abadal,Garifalia Sebú, Annika Seppälä, Jussi Seppälä, Valentina Simonetti, Matthias Stevens, Anna Triantafillou, Zsolt Szabolcs Unoka, Judith Usall, Vincenzo Vella, and David Vermeir.

XSL•FO
RenderX

Data Availability

The datasets generated and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

ARB, AAS, EG, IC, JM, CA, ER, JB, JU, SO, SA, JCM, MH, JS, and EJ planned the investigation. MH was responsible for statistical analysis. KR, AC, ZU, KF, EG, and IC were responsible for carrying out this clinical study. JS, EG, IC, EJ, MI, MH, and JM wrote this paper. All authors revised and edited this paper critically and had final approval of the version to be published.

Conflicts of Interest

None declared.

References

1. Leucht S, Busch R, Kissling W, Kane JM. Early prediction of antipsychotic nonresponse among patients with schizophrenia. J Clin Psychiatry 2007 Mar;68(3):352-360. [doi: 10.4088/jcp.v68n0301] [Medline: 17388703]
2. Jääskeläinen E, Juola P, Hirvonen N, et al. A systematic review and meta-analysis of recovery in schizophrenia. Schizophr Bull 2013 Nov;39(6):1296-1306. [doi: 10.1093/schbul/sbs130] [Medline: 23172003]
3. Kennedy JL, Altar CA, Taylor DL, Degtiar I, Hornberger JC. The social and economic burden of treatment-resistant schizophrenia: a systematic literature review. Int Clin Psychopharmacol 2014 Mar;29(2):63-76. [doi: 10.1097/YIC.0b013e32836508e6] [Medline: 23995856]
4. Mortimer AM, Singh P, Shepherd CJ, Puthiryackal J. Clozapine for treatment-resistant schizophrenia: National Institute of Clinical Excellence (NICE) guidance in the real world. Clin Schizophr Relat Psychoses 2010 Apr;4(1):49-55. [doi: 10.3371/CSRP.4.1.4] [Medline: 20643629]
5. Millier A, Schmidt U, Angermeyer MC, et al. Humanistic burden in schizophrenia: a literature review. J Psychiatr Res 2014 Jul;54:85-93. [doi: 10.1016/j.jpsychires.2014.03.021] [Medline: 24795289]
6. Ben-Zeev D, Brenner CJ, Begale M, Duffecy J, Mohr DC, Mueser KT. Feasibility, acceptability, and preliminary efficacy of a smartphone intervention for schizophrenia. Schizophr Bull 2014 Nov;40(6):1244-1253. [doi: 10.1093/schbul/sbu033] [Medline: 24609454]
7. Granholm E, Ben-Zeev D, Link PC, Bradshaw KR, Holden JL. Mobile Assessment and Treatment for Schizophrenia (MATS): a pilot trial of an interactive text-messaging intervention for medication adherence, socialization, and auditory hallucinations. Schizophr Bull 2012 May;38(3):414-425. [doi: 10.1093/schbul/sbr155] [Medline: 22080492]
8. Palmier-Claus JE, Ainsworth J, Machin M, et al. The feasibility and validity of ambulatory self-report of psychotic symptoms using a smartphone software application. BMC Psychiatry 2012 Oct 17;12:172. [doi: 10.1186/1471-244X-12-172] [Medline: 23075387]
9. Firth J, Cotter J, Torous J, Bucci S, Firth JA, Yung AR. Mobile phone ownership and endorsement of “mHealth” among people with psychosis: a meta-analysis of cross-sectional studies. Schizophr Bull 2016 Mar;42(2):448-455. [doi: 10.1093/schbul/sbv132] [Medline: 26400871]
10. Brunette MF, Achtyes E, Pratt S, et al. Use of smartphones, computers and social media among people with SMI: opportunity for intervention. Community Ment Health J 2019 Aug;55(6):973-978. [doi: 10.1007/s10597-019-00431-7] [Medline: 31175518]
11. Young AS, Cohen AN, Niv N, et al. Mobile phone and smartphone use by people with serious mental illness. Psychiatr Serv 2020 Mar 1;71(3):280-283. [doi: 10.1176/appi.ps.201900203] [Medline: 31744429]
12. Firth J, Torous J. Smartphone apps for schizophrenia: a systematic review. JMIR mHealth uHealth 2015 Nov 6;3(4):e102. [doi: 10.2196/mhealth.4930] [Medline: 26546039]
13. Bonet L, Izquierdo C, Escartí MJ, et al. Use of mobile technologies in patients with psychosis: a systematic review. Rev Psiquiatr Salud Ment (Engl Ed) 2017 Jul;10(3):168-178. [doi: 10.1016/j.rpsmen.2017.05.010]
14. Buck B, Scherer E, Brian R, et al. Relationships between smartphone social behavior and relapse in schizophrenia: a preliminary report. Schizophr Res 2019 Jun;208:167-172. [doi: 10.1016/j.schres.2019.03.014] [Medline: 30940400]
15. Clarke S, Hanna D, Mulholland C, Shannon C, Urquhart C. A systematic review and meta-analysis of digital health technologies effects on psychotic symptoms in adults with psychosis. Psychosis 2019 Oct 2;11(4):362-373. [doi: 10.1080/17522439.2019.1632376]
16. de Almeida RS, Marques A. User engagement in mobile apps for people with schizophrenia: a scoping review. Front Digit Health 2022;4:1023592. [doi: 10.3389/fdgth.2022.1023592] [Medline: 36703941]
17. Achtyes ED, Ben-Zeev D, Luo Z, et al. Off-hours use of a smartphone intervention to extend support for individuals with schizophrenia spectrum disorders recently discharged from a psychiatric hospital. Schizophr Res 2019 Apr;206:200-208. [doi: 10.1016/j.schres.2018.11.026] [Medline: 30551981]
18. Ben-Zeev D, Brian RM, Jonathan G, et al. Mobile health (mHealth) versus clinic-based group intervention for people with serious mental illness: a randomized controlled trial. Psychiatr Serv 2018 Sep 1;69(9):978-985. [doi: 10.1176/appi.ps.201800063] [Medline: 29793397]

XSL•FO
RenderX

19. Ben-Zeev D, Brian RM, Aschbrenner KA, Jonathan G, Steingard S. Video-based mobile health interventions for people with schizophrenia: bringing the “pocket therapist” to life. Psychiatr Rehabil J 2018 Mar;41(1):39-45. [doi: 10.1037/prj0000197] [Medline: 27295133]
20. Hanssen E, Balvert S, Oorschot M, et al. An ecological momentary intervention incorporating personalised feedback to improve symptoms and social functioning in schizophrenia spectrum disorders. Psychiatry Res 2020 Feb;284:112695. [doi: 10.1016/j.psychres.2019.112695] [Medline: 31831201]
21. Moitra E, Park HS, Gaudiano BA. Development and initial testing of an mHealth transitions of care intervention for adults with schizophrenia-spectrum disorders immediately following a psychiatric hospitalization. Psychiatr Q 2021 Mar;92(1):259-272. [doi: 10.1007/s11126-020-09792-9] [Medline: 32613525]
22. Almeida RFSD, Sousa TJ, Couto AS, Marques AJ, Queirós CM, Martins CL. Development of weCope, a mobile app for illness self-management in schizophrenia. Arch Clin Psychiatry (São Paulo) 2019;46(1):1-4. [doi: 10.1590/0101-60830000000182]
23. Steare T, O’Hanlon P, Eskinazi M, et al. Smartphone-delivered self-management for first-episode psychosis: the ARIES feasibility randomised controlled trial. BMJ Open 2020 Aug;10(8):e034927. [doi: 10.1136/bmjopen-2019-034927] [Medline: 32847902]
24. Steare T, Giorgalli M, Free K, et al. A qualitative study of stakeholder views on the use of a digital app for supported self-management in early intervention services for psychosis. BMC Psychiatry 2021 Jun 19;21(1):311. [doi: 10.1186/s12888-021-03317-9] [Medline: 34147075]
25. Kidd SA, Feldcamp L, Adler A, et al. Feasibility and outcomes of a multi-function mobile health approach for the schizophrenia spectrum: App4Independence (A4i). PLoS ONE 2019;14(7):e0219491. [doi: 10.1371/journal.pone.0219491] [Medline: 31306439]
26. Ben-Zeev D, Brian R, Wang R, et al. CrossCheck: integrating self-report, behavioral sensing, and smartphone use to identify digital indicators of psychotic relapse. Psychiatr Rehabil J 2017 Sep;40(3):266-275. [doi: 10.1037/prj0000243] [Medline: 28368138]
27. Bucci S, Barrowclough C, Ainsworth J, et al. Actissist: proof-of-concept trial of a theory-driven digital intervention for psychosis. Schizophr Bull 2018 Aug 20;44(5):1070-1080. [doi: 10.1093/schbul/sby032] [Medline: 29566206]
28. Lewis S, Ainsworth J, Sanders C, et al. Smartphone-enhanced symptom management in psychosis: open, randomized controlled trial. J Med Internet Res 2020 Aug 13;22(8):e17019. [doi: 10.2196/17019] [Medline: 32788150]
29. Eisner E, Bucci S, Berry N, Emsley R, Barrowclough C, Drake RJ. Feasibility of using a smartphone app to assess early signs, basic symptoms and psychotic symptoms over six months: a preliminary report. Schizophr Res 2019 Jun;208:105-113. [doi: 10.1016/j.schres.2019.04.003] [Medline: 30979665]
30. Niendam TA, Tully LM, Iosif AM, et al. Enhancing early psychosis treatment using smartphone technology: a longitudinal feasibility and validity study. J Psychiatr Res 2018 Jan;96:239-246. [doi: 10.1016/j.jpsychires.2017.10.017] [Medline: 29126059]
31. Bonet L, Torous J, Arce D, Blanquer I, Sanjuan J. ReMindCare app for early psychosis: pragmatic real world intervention and usability study. JMIR mHealth uHealth 2020 Nov 6;8(11):e22997. [doi: 10.2196/22997] [Medline: 33155986]
32. Terp M, Jørgensen R, Laursen BS, Mainz J, Bjørnes CD. A smartphone app to foster power in the everyday management of living with schizophrenia: qualitative analysis of young adults’ perspectives. JMIR Ment Health 2018 Oct 1;5(4):e10157. [doi: 10.2196/10157] [Medline: 30274966]
33. Garety P, Ward T, Emsley R, et al. Effects of SlowMo, a blended digital therapy targeting reasoning, on paranoia among people with psychosis: a randomized clinical trial. JAMA Psychiatry 2021 Jul 1;78(7):714-725. [doi: 10.1001/jamapsychiatry.2021.0326] [Medline: 33825827]
34. Gire N, Caton N, McKeown M, et al. ’Care co-ordinator in my pocket’: a feasibility study of mobile assessment and therapy for psychosis (TechCare). BMJ Open 2021 Nov 16;11(11):e046755. [doi: 10.1136/bmjopen-2020-046755] [Medline: 34785541]
35. Moore E, Williams A, Bell I, Thomas N. Client experiences of blending a coping-focused therapy for auditory verbal hallucinations with smartphone-based ecological momentary assessment and intervention. Internet Interv 2020 Mar;19:100299. [doi: 10.1016/j.invent.2019.100299] [Medline: 31890641]
36. Schlosser DA, Campellone TR, Truong B, et al. Efficacy of PRIME, a mobile app intervention designed to improve motivation in young people with schizophrenia. Schizophr Bull 2018 Aug 20;44(5):1010-1020. [doi: 10.1093/schbul/sby078] [Medline: 29939367]
37. Krzystanek M, Borkowski M, Skałacka K, Krysta K. A telemedicine platform to improve clinical parameters in paranoid schizophrenia patients: results of a one-year randomized study. Schizophr Res 2019 Feb;204:389-396. [doi: 10.1016/j.schres.2018.08.016] [Medline: 30154027]
38. Fulford D, Mote J, Gard DE, et al. Development of the Motivation and Skills Support (MASS) social goal attainment smartphone app for (and with) people with schizophrenia. J Behav Cogn Ther 2020 Apr;30(1):23-32. [doi: 10.1016/j.jbct.2020.03.016] [Medline: 33437970]

XSL•FO
RenderX

39. Taylor KM, Bradley J, Cella M. A novel smartphone-based intervention targeting sleep difficulties in individuals experiencing psychosis: a feasibility and acceptability evaluation. Psychol Psychother 2022 Sep;95(3):717-737. [doi: 10.1111/papt.12395] [Medline: 35481697]
40. Kreyenbuhl J, Record EJ, Himelhoch S, et al. Development and feasibility testing of a smartphone intervention to improve adherence to antipsychotic medications. Clin Schizophr Relat Psychoses 2019;12(4):152-167. [doi: 10.3371/CSRP.KRRE.070816] [Medline: 27454213]
41. Lim MH, Gleeson JFM, Rodebaugh TL, et al. A pilot digital intervention targeting loneliness in young people with psychosis. Soc Psychiatry Psychiatr Epidemiol 2020 Jul;55(7):877-889. [doi: 10.1007/s00127-019-01681-2] [Medline: 30874828]
42. Torous J, Bucci S, Bell IH, et al. The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality. World Psychiatry 2021 Oct;20(3):318-335. [doi: 10.1002/wps.20883] [Medline: 34505369]
43. Torous J, Smith KA, Hardy A, et al. Digital health interventions for schizophrenia: setting standards for mental health. Schizophr Res 2024 May;267:392-395. [doi: 10.1016/j.schres.2024.04.013] [Medline: 38640849]
44. Vitger T, Korsbek L, Austin SF, Petersen L, Nordentoft M, Hjorthøj C. Digital shared decision-making interventions in mental healthcare: a systematic review and meta-analysis. Front Psychiatry 2021;12(691):691251. [doi: 10.3389/fpsyt.2021.691251] [Medline: 34552514]
45. MacInnes D, Kinane C, Parrott J, et al. A pilot cluster randomised trial to assess the effect of a structured communication approach on quality of life in secure mental health settings: the Comquol study. BMC Psychiatry 2016 Sep 29;16(1):335. [doi: 10.1186/s12888-016-1046-8] [Medline: 27688024]
46. Priebe S, Kelley L, Golden E, et al. Effectiveness of structured patient-clinician communication with a solution focused approach (DIALOG+) in community treatment of patients with psychosis--a cluster randomised controlled trial. BMC Psychiatry 2013 Jun 26;13(1):173. [doi: 10.1186/1471-244X-13-173] [Medline: 23802602]
47. Edbrooke-Childs J, Edridge C, Averill P, et al. A feasibility trial of power up: smartphone app to support patient activation and shared decision making for mental health in young people. JMIR mHealth uHealth 2019 Jun 4;7(6):e11677. [doi: 10.2196/11677] [Medline: 31165709]
48. Loh PY, Martinengo L, Heaukulani C, et al. Characteristics and outcomes of mHealth interventions in psychosis: systematic mapping review. J Med Internet Res 2024 Dec 23;26:e55924. [doi: 10.2196/55924] [Medline: 39714907]
49. Lim CT, Fuchs C, Torous J. Integrated digital mental health care: a vision for addressing population mental health needs. Int J Gen Med 2024;17:359-365. [doi: 10.2147/IJGM.S449474] [Medline: 38318335]
50. Mueller NE, Panch T, Macias C, Cohen BM, Ongur D, Baker JT. Using smartphone apps to promote psychiatric rehabilitation in a peer-led community support program: pilot study. JMIR Ment Health 2018 Aug 15;5(3):e10092. [doi: 10.2196/10092] [Medline: 30111526]
51. Smith KA, Blease C, Faurholt-Jepsen M, et al. Digital mental health: challenges and next steps. BMJ Ment Health 2023 Feb;26(1):e300670. [doi: 10.1136/bmjment-2023-300670] [Medline: 37197797]
52. Camacho E, Torous J. Introducing an implementation framework for augmenting care with digital technology for early psychosis patients: theory and motivation. J Ment Health 2022 Dec;31(6):816-824. [doi: 10.1080/09638237.2021.1922634] [Medline: 34057008]
53. Green JB, Rodriguez J, Keshavan M, Lizano P, Torous J. Implementing technologies to enhance coordinated specialty care framework: implementation outcomes from a development and usability study. JMIR Form Res 2023 Oct 3;7:e46491. [doi: 10.2196/46491] [Medline: 37788066]
54. Camacho E, Torous J. Impact of digital literacy training on outcomes for people with serious mental illness in community and inpatient settings. Psychiatr Serv 2023 May 1;74(5):534-538. [doi: 10.1176/appi.ps.20220205] [Medline: 36164771]
55. Buck B, Kopelovich SL, Tauscher JS, Chwastiak L, Ben-Zeev D. Developing the workforce of the digital future: leveraging technology to train community-based mobile mental health specialists. J Technol Behav Sci 2022 Aug 3:1-7. [doi: 10.1007/s41347-022-00270-6] [Medline: 35967965]
56. Miralles I, Granell C, Díaz-Sanahuja L, et al. Smartphone apps for the treatment of mental disorders: systematic review. JMIR mHealth uHealth 2020 Apr 2;8(4):e14897. [doi: 10.2196/14897] [Medline: 32238332]
57. Agarwal S, Jalan M, Wilcox H, Sharma R, Hill R, Pantalone E, et al. Evalution of mental health mobile applications. : Agency for Healthcare Research and Quality (US); 2022 URL: https://effectivehealthcare.ahrq.gov/sites/default/files/product/pdf/mental-health-mobile-apps-tech-brief.pdf [accessed 2025-04-24]
58. Radu I, Scheermesser M, Spiess MR, Schulze C, Händler-Schuster D, Pehlke-Milde J. Digital health for migrants, ethnic and cultural minorities and the role of participatory development: a scoping review. Int J Environ Res Public Health 2023 Oct 23;20(20):6962. [doi: 10.3390/ijerph20206962] [Medline: 37887700]
59. Kwon S, Firth J, Joshi D, Torous J. Accessibility and availability of smartphone apps for schizophrenia. Schizophrenia (Heidelb) 2022 Nov 16;8(1):98. [doi: 10.1038/s41537-022-00313-0] [Medline: 36385116]
60. Seppälä J, De Vita I, Jämsä T, et al. Mobile phone and wearable sensor-based mHealth approaches for psychiatric disorders and symptoms: systematic review. JMIR Ment Health 2019 Feb 20;6(2):e9819. [doi: 10.2196/mental.9819] [Medline: 30785404]

XSL•FO
RenderX

61. Chivilgina O, Wangmo T, Elger BS, Heinrich T, Jotterand F. mHealth for schizophrenia spectrum disorders management: a systematic review. Int J Soc Psychiatry 2020 Nov;66(7):642-665. [doi: 10.1177/0020764020933287] [Medline: 32571123]
62. Yu Y, Li Y, Li T, et al. New path to recovery and well-being: cross-sectional study on WeChat use and endorsement of WeChat-based mHealth among people living with schizophrenia in China. J Med Internet Res 2020 Sep 18;22(9):e18663. [doi: 10.2196/18663] [Medline: 32945774]
63. Huerta-Ramos E, Escobar-Villegas MS, Rubinstein K, et al. Measuring users' receptivity toward an integral intervention model based on mhealth solutions for patients with treatment-resistant schizophrenia (m-RESIST): a qualitative study. JMIR mHealth uHealth 2016 Sep 28;4(3):e112. [doi: 10.2196/mhealth.5716] [Medline: 27682896]
64. Alonso-Solís A, Rubinstein K, Corripio I, et al. Mobile therapeutic attention for treatment-resistant schizophrenia (m-RESIST): a prospective multicentre feasibility study protocol in patients and their caregivers. BMJ Open 2018 Jul 16;8(7):e021346. [doi: 10.1136/bmjopen-2017-021346] [Medline: 30012788]
65. Grasa E, Seppälä J, Alonso-Solis A, et al. m-RESIST, a mobile therapeutic intervention for treatment-resistant schizophrenia: feasibility, acceptability, and usability study. JMIR Form Res 2023 Jun 30;7:e46179. [doi: 10.2196/46179] [Medline: 37389933]
66. Elkis H, Meltzer HY. Refractory schizophrenia. Braz J Psychiatry 2007 Oct;29 Suppl 2:S41-S47. [doi: 10.1590/s1516-44462007000600002] [Medline: 18157433]
67. Peuskens J. The evolving definition of treatment resistance. J Clin Psychiatry 1999;60 Suppl 12:4-8. [Medline: 10372602]
68. Andreasen NC, Carpenter WT, Kane JM, Lasser RA, Marder SR, Weinberger DR. Remission in schizophrenia: proposed criteria and rationale for consensus. AJP 2005 Mar 1;162(3):441-449. [doi: 10.1176/appi.ajp.162.3.441]
69. Kay SR, Fiszbein A, Opler LA. The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. Schizophr Bull 1987;13(2):261-276. [doi: 10.1093/schbul/13.2.261] [Medline: 3616518]
70. Haro JM, Kamath SA, Ochoa S, et al. The Clinical Global Impression–Schizophrenia scale: a simple instrument to measure the diversity of symptoms present in schizophrenia. Acta Psychiatr Scand 2003 May;107(s416):16-23. [doi: 10.1034/j.1600-0447.107.s416.5.x]
71. Endicott J, Spitzer RL, Fleiss JL, Cohen J. The global assessment scale. A procedure for measuring overall severity of psychiatric disturbance. Arch Gen Psychiatry 1976 Jun;33(6):766-771. [doi: 10.1001/archpsyc.1976.01770060086012] [Medline: 938196]
72. Brooks R. EuroQol: the current state of play. Health Policy 1996 Jul;37(1):53-72. [doi: 10.1016/0168-8510(96)00822-6] [Medline: 10158943]
73. Eldridge SM, Chan CL, Campbell MJ, et al. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. BMJ 2016 Oct 24;355:i5239. [doi: 10.1136/bmj.i5239] [Medline: 27777223]
74. Hermes EDA, Sokoloff D, Stroup TS, Rosenheck RA. Minimum clinically important difference in the Positive and Negative Syndrome Scale with data from the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE). J Clin Psychiatry 2012 Apr;73(4):526-532. [doi: 10.4088/JCP.11m07162] [Medline: 22579152]
75. Suzuki T, Remington G, Mulsant BH, et al. Defining treatment-resistant schizophrenia and response to antipsychotics: a review and recommendation. Psychiatry Res 2012 May 15;197(1-2):1-6. [doi: 10.1016/j.psychres.2012.02.013] [Medline: 22429484]
76. Leucht S, Kane JM, Kissling W, Hamann J, Etschel E, Engel RR. What does the PANSS mean? Schizophr Res 2005 Nov 15;79(2-3):231-238. [doi: 10.1016/j.schres.2005.04.008] [Medline: 15982856]
77. Green CEL, Freeman D, Kuipers E, et al. Measuring ideas of persecution and social reference: the Green et al. Paranoid Thought Scales (GPTS). Psychol Med 2008 Jan;38(1):101-111. [doi: 10.1017/S0033291707001638] [Medline: 17903336]
78. Sedgwick P, Greenwood N. Understanding the Hawthorne effect. BMJ 2015 Sep;4:h4672. [doi: 10.1136/bmj.h4672]
79. McCambridge J, Witton J, Elbourne DR. Systematic review of the Hawthorne effect: new concepts are needed to study research participation effects. J Clin Epidemiol 2014 Mar;67(3):267-277. [doi: 10.1016/j.jclinepi.2013.08.015] [Medline: 24275499]
80. König HH, Roick C, Angermeyer MC. Validity of the EQ-5D in assessing and valuing health status in patients with schizophrenic, schizotypal or delusional disorders. Eur Psychiatry 2007 Apr;22(3):177-187. [doi: 10.1016/j.eurpsy.2006.08.004] [Medline: 17142014]
81. López MDPA, Ong BA, Frigola XB, et al. Digital literacy as a new determinant of health: a scoping review. PLOS Digit Health 2023 Oct;2(10):e0000279. [doi: 10.1371/journal.pdig.0000279] [Medline: 37824584]
82. Balcombe L, De Leo D. Digital mental health challenges and the horizon ahead for solutions. JMIR Ment Health 2021 Mar 29;8(3):e26811. [doi: 10.2196/26811] [Medline: 33779570]
83. Philippe TJ, Sikder N, Jackson A, et al. Digital health interventions for delivery of mental health care: systematic and comprehensive meta-review. JMIR Ment Health 2022 May 12;9(5):e35159. [doi: 10.2196/35159] [Medline: 35551058]
84. Kozelka EE, Acquilano SC, Al-Abdulmunem M, et al. Digital mental health and its discontents: assumptions about technology that create barriers to equitable access. PS (Wash DC) 2024 Mar 1;75(3):299-302. [doi: 10.1176/appi.ps.20230238]
85. Goldberg SB, Lam SU, Simonsson O, Torous J, Sun S. Mobile phone-based interventions for mental health: a systematic meta-review of 14 meta-analyses of randomized controlled trials. PLOS Digit Health 2022;1(1):e0000002. [doi: 10.1371/journal.pdig.0000002] [Medline: 35224559]

XSL•FO
RenderX

86. Carter H, Araya R, Anjur K, Deng D, Naslund JA. The emergence of digital mental health in low-income and middle-income countries: a review of recent advances and implications for the treatment and prevention of mental disorders. J Psychiatr Res 2021 Jan;133:223-246. [doi: 10.1016/j.jpsychires.2020.12.016] [Medline: 33360867]

Abbreviations

CGI-SCH: Clinical Global Impression-Schizophrenia
CONSORT: Consolidated Standards of Reporting Trials
EQ-VAS: EuroQol visual analogue scale
GAF: Global Assessment of Functioning
ICT: information and communication technology
m-RESIST: mobile therapeutic attention for treatment-resistant schizophrenia
mHealth: mobile health
PANSS: Positive and Negative Syndrome Scale
RCT: randomized controlled trial
TRS: treatment-resistant schizophrenia

Edited by A Kushniruk; submitted 17.10.24; peer-reviewed by J Lee, P Cutrufelli, V Peitl; revised version received 26.03.25; accepted 27.03.25; published 15.05.25.

Please cite as:
Seppälä J, Grasa E, Alonso-Solis A, Roldan-Bejarano A, Haapea M, Isohanni M, Miettunen J, Caro Mendivelso J, Almazán C, Rubinstein K, Caspi A, Unoka Z, Farkas K, Reixach E, Berdun J, Usall J, Ochoa S, Corripio I, Jääskeläinen E, m-Resist Group Mobile Therapeutic Attention for Treatment-Resistant Schizophrenia (m-RESIST) Solution for Improving Clinical and Functional Outcomes in Treatment-Resistant Schizophrenia: Prospective, Multicenter Efficacy Study
JMIR Hum Factors 2025;12:e67659
URL: https://humanfactors.jmir.org/2025/1/e67659
doi:10.2196/67659

© Jussi Seppälä, Eva Grasa, Anna Alonso-Solis, Alexandra Roldan-Bejarano, Marianne Haapea, Matti Isohanni, Jouko Miettunen, Johanna Caro Mendivelso, Cari Almazán, Katya Rubinstein, Asaf Caspi, Zolt Unoka, Kinga Farkas, Elisenda Reixach, Jesus Berdun, Judith Usall, Susana Ochoa, Iluminada Corripio, Erika Jääskeläinen, m-Resist Group. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 15.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Factors Related to Pain and Disability Outcomes After an Internet-Delivered or Physiotherapist-Led Exercise Program for Individuals With Chronic Whiplash Symptoms: Secondary Analysis of a Randomized Controlled Study

Gunnel Peterson[1,2], PhD; Stefan Ljunggren[1,3], PhD; Anneli Peolsson[1,3], PhD

[1]Department of Health Medicine and Caring Sciences, Unit of Physiotherapy, Linköping University, Sjukhusvägen 34, Linköpings Universitet, Linköping, Sweden

[2]Centre for Clinical Research Sormland, Uppsala University, Uppsala, Sweden

[3]Department of Health Medicine and Caring Sciences, Occupational and Environmental Medicine Centre, Unit of Clinical Medicine, Linköping University, Linköping, Sweden

Corresponding Author:
Gunnel Peterson, PhD
Department of Health Medicine and Caring Sciences, Unit of Physiotherapy, Linköping University, Sjukhusvägen 34, Linköpings Universitet, Linköping, Sweden

Abstract

Background: A neck-specific exercise program has shown sustained clinically important changes in pain and disability for approximately 50% of individuals with chronic whiplash-associated disorders (WAD). However, there is limited information about factors related to treatment response.

Objective: The aim of this study is to identify factors related to changes in disability, neck pain, and physical function after a neck-specific exercise program delivered in 2 different ways for individuals with persistent WAD grade II or III, and to investigate whether any factors could predict those with clinically improved versus not improved disability, pain, and physical function.

Methods: This was a planned secondary analysis of a multicenter prospective randomized controlled trial. Participants (n=140) with persistent (between 6 mo and 5 y from injury) WAD grade II or III were randomized into a 12-week, internet-based neck-specific exercise program (NSEIT) with 4 physiotherapy visits or the same exercise program supervised by a physiotherapist (NSE) twice per week for 12 weeks. Multivariate data analyses and orthogonal partial least squares (OPLS) models were used to investigate change in psychological and physiological factors (independent factors) related to change in the dependent factors: neck-related disability measured with the Neck Disability Index (NDI), neck pain intensity measured with a visual analogue scale, and physical function measured with the Patient-Specific Functional Scale (PSFS). Outcomes were measured at baseline and at 3-month and 15-month follow-up. OPLS discriminant analysis was used to investigate differences between the two groups (NSEIT and NSE) by studying the change scores of the dependent and independent factors. OPLS discriminant analysis was also used to investigate whether background variables and baseline measurements of the independent factors could predict clinically significant improvement in the dependent factors NDI, neck pain, and PSFS.

Results: There were no significant differences between the groups. In both NSEIT and NSE, improvements in the following independent factors were related to improvements in NDI, pain, and PSFS at 3-month and 15-month follow-up: anxiety, depression, cognitive failures, pain catastrophizing, self-efficacy, fear avoidance beliefs, cervical range of motion, headache, and symptom satisfaction (R^2=0.31 - 0.37; Q^2=0.25 - 0.30; cross-validated ANOVA P<.001). No significant OPLS models could be built to distinguish clinically improved versus nonimproved patients as assessed by NDI, neck pain, or PSFS.

Conclusions: Improvements in both psychological and physiological factors were related to improvements in disability, neck pain, and physical function after 12 weeks of NSEIT or NSE. The results indicate that these factors are interrelated and can be improved both with NSEIT and NSE. Known risk factors for poor outcomes of neck disability in WAD, such as low self-efficacy, fear avoidance beliefs, depressive symptoms, and catastrophizing, were improved, and we need to examine other factors not included in this study that can identify those who are not improved after NSEIT or NSE.

Trial Registration: ClinicalTrials.gov NCT03022812; https://clinicaltrials.gov/ct2/show/NCT03022812

(JMIR Hum Factors 2025;12:e67991) doi:10.2196/67991

XSL•FO
RenderX

KEYWORDS

whiplash; exercise therapy; rehabilitation; spine association; outcome; pain; disability; physiotherapy; exercise; exercising; chronic whiplash; exercise program; whiplash-associated disorders; randomized controlled trial; RCT; secondary analysis; neck pain; neck; visual analogue scale; Neck Disability Index; physical activity; multivariate data analysis; internet-delivered; internet

Introduction

Whiplash-associated disorders (WAD) are indirect trauma to the neck, commonly related to a car crash [1], and approximately half of those injured develop chronic disability and pain [1,2]. Chronic WAD will have extensive consequences for the individual living with persistent pain, leading to a burden for the individual and high costs for society [3,4]. Depression, low levels of life satisfaction, and posttraumatic stress have been observed 5 years after whiplash injury [5]. Costs increase as many individuals with chronic WAD are consumers of multiple health services and experience reduced work ability [4,6]. Although the evidence is moderate at best, exercise and patient education are recommended treatments in chronic WAD [7].

Approximately 50% of the included participants in a neck-specific exercise program, conducted by 2 of the authors in this study (GP and AP), with chronic WAD grades II (neck pain and musculoskeletal findings during a physical examination) and III (as grade II but with additional neurological findings) had a sustained clinically important change in pain and disability [8,9]. Despite these promising results, about half of the participants were not clinically improved, demonstrating that chronic moderate-to-severe WAD is difficult to treat successfully. Consequently, there is a need to identify factors related to improvement in disability and pain in order to optimize treatment, discriminating them from factors that may contribute to sustained problems.

An earlier study found that depression, low self-efficacy, and higher disability were associated with a larger pain area in women with chronic WAD [10]. In another study, self-efficacy mediated the association between pain intensity and pain-related disability, and kinesiophobia mediated the association between self-efficacy and pain catastrophizing in acute and subacute WAD [11], which can lead to chronic disability. However, that study did not investigate factors related to improvements after rehabilitation. In studies where interventions included acceptance and commitment therapy or cognitive behavioral therapy, psychological factors had a mediating or predictive effect on pain and/or disability outcome for individuals with chronic WAD [12,13]. Another study found that participation in a neck-specific exercise program was the only significant factor associated with neck pain reduction at 12-month follow-up [14], and neck muscle endurance and perceived work ability were mediators for a reduction in pain [15] in chronic WAD. The treatment of chronic nonmalignant pain is influenced by biological, psychological, and sociological factors [16]. Numerous studies have found factors such as anxiety, catastrophizing, and severe pain to be related to poor treatment outcomes after acute whiplash injury [17].

However, to our knowledge, there is very limited information about factors related to treatment response in chronic WAD, or about the optimization of a neck-specific exercises program delivered via the internet. Internet-based exercise interventions with fewer visits to physiotherapists can increase availability, shorten waiting times, and reduce costs. Internet-based treatment may also strengthen the patient's own resources for greater autonomy and their ability to conduct long-term self-care [18]. Factors associated with improvements in an internet-based neck-specific exercise program (NSEIT) may differ compared to a neck-specific exercise program supervised by a physiotherapist (NSE) and need to be investigated. Moreover, the ability to predict those who will not recover after NSEIT and NSE is highly relevant to improving treatment in chronic WAD.

The aims of the study were (1) to investigate factors related to change in neck pain and disability after NSEIT or the same exercise program supervised by a physiotherapist (NSE) for individuals with chronic WAD grade II or III and (2) to investigate whether other factors could predict those with clinically improved versus not improved pain and disability after 3 months of NSEIT or NSE.

Methods

Overview

The recruitment of participants and the exercise interventions were reported earlier [9] and are briefly described below. This was a planned secondary analysis, based on a multicenter, prospective randomized controlled study [9]. Data were collected from April 6, 2017, until September 15, 2020, and the main outcome data have been published [9]. The participants were recruited from 10 county councils in Sweden.

Participants

Individuals with chronic (between 6 mo and 5 y from injury) WAD grades II–III [19] with neck symptoms within the first week after a whiplash injury in a 4-wheeled motor vehicle crash were eligible for inclusion. Symptoms emanating from the neck were verified through the participant's medical history and a physical clinical examination. For additional inclusion and exclusion criteria, see Textbox 1.

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria.

Inclusion

- Average neck pain last week ≥20 mm on a 0-100 mm visual analogue scale
- Neck disability ≥20%
- Age 18 - 63 years
- Daily access to a computer, tablet, or smartphone and the internet
- Sufficient time and motivation to follow the treatment program

Exclusion

- Signs of head injury
- Serious physical pathology
- Surgery on the cervical column
- Generalized or more dominant pain elsewhere in the body
- Previous fractures or dislocation of the cervical column
- Other illness or injury that may prevent full participation in the study and/or for which exercises are contraindicated
- Previous severe neck problems that resulted in sick leave for more than 1 month in the year before the current whiplash injury
- Lack of knowledge of the Swedish language
- Diagnosed severe mental illness
- Current alcohol and/or drug abuse

Reasons for the exclusion criteria were as follows: (1) the programs could lead to increased illness or potentially negatively influence other important treatments, (2) another illness or disease made it impossible to perform the programs and/or follow them for three months, (3) the exercises were contraindicated, and (4) participants were unable to understand the information and/or the exercises.

Ethical Considerations

The study was approved by the regional ethical review board (approval numbers 2016/135 - 31, 2016/526 - 32, and 2017/45 - 32). The study was registered in ClinicalTrials.gov (NCT03022812, initial release 12/20/2016) and the study protocol was published prior to data collection [19]. The original written and informed consent allowed for this secondary analysis without requiring additional consent, and no new data were collected for this study. Deidentification was used to prevent the identification of individuals. The key code is securely stored according to the regulations of the ethics committee and the Swedish Authority for Privacy Protection. The results are presented at the group level, ensuring no connection to individual participants. Participants were not compensated for their involvement in the study, except for traveling cost compensation in accordance with national ethical regulations.

Recruitment

Information about the study was provided via health care personnel and advertising. Interested participants completed a short survey on the Linköping University website. The patient's medical history was compiled and a physical examination was carried out before inclusion. Once the study criteria had been met, the participants were included after giving written and oral informed consent. The test leaders were blinded to group allocation. The participants were blinded to group allocation when baseline data were collected [19].

Randomization

Participants (n=140) were randomized into 2 groups (n=70 in each group) using a computerized block randomization list stratified by sex. After randomization, the treating physiotherapist was contacted by the research team and called the patient to the first appointment.

Before the start of the study, all treating physiotherapists received oral and written information and a day of practical training from the project leaders. A detailed description of the NSEIT and NSE exercise program has been published [9].

NSEIT Group

In the NSEIT group, the participants underwent a 12-week exercise program with 4 visits to the physiotherapist (at weeks 1, 2, 3, and 7) to be introduced to the exercises, to progress the exercises, and for follow-up to ensure correct performance. The patient had access to the internet-based program on a website. The program included photos and videos of the exercises, information about the musculoskeletal anatomy of the neck and about neurophysiological and neurobiological pain, and strategies for dealing with neck pain relapse. The exercises were initially targeted to activate the deep neck muscles and increase postural control, continuing with an individually tailored progression of neck muscle endurance exercises [9,19]. At the end of the treatment period, the participants were encouraged to continue doing the exercises 2-3 times per week and to include neck-specific exercises (for additional details, see [9]).

XSL•FO
RenderX

NSE Group

In the NSE group, participants received the same information and exercises as the NSEIT group, but in visits to the physiotherapy clinic twice per week (a maximum of 24 sessions) [9,19].

Patient and Public Involvement

Patients were involved in developing the questionnaires before recruitment to the study started. The questionnaires were tested by 3 patients to ensure the inclusion of relevant questions and readability. Patient satisfaction with the exercises, information, tests, and intervention (NSEIT or NSE) were measured postintervention.

Adverse events could be reported to the project leaders directly by participants, the test leaders, or the treating physiotherapists (for details, see [9]).

Predictor Variables and Outcome Measures Used in Multivariate Modeling

Measurements were collected via questionnaires at baseline and at 3-month and 15-month follow-ups. The following variables were used in the multivariate modeling.

Background Variables

The following variables were only collected at baseline: age (years), sex (male/female), smoking (yes/no), level of education, pain duration (months since injury), WAD grade II and III according to the Quebec Task Force classification [20], and expectations of the intervention on a 0 to 10 scale (with higher scores indicating stronger belief that the treatment would help) [21].

Dependent Factors

The Neck Disability Index (NDI*)* measures self-perceived disability in everyday life because of neck pain, including 10 items of neck-related disability, which are summed and transformed into a percentage (0%=no disability; 100%=highest score for disability) [22].

Current neck pain was measured on a visual analogue scale (VAS) from 0 - 100 mm, where 0=no pain and 100=worst imaginable pain intensity [23].

The Patient-Specific Functional Scale (PSFS*)* measured level of physical function, graded from 0 (unable to do the activity) to 10 (functional level equal to preinjury status), regarding self-perceived physical activity, leisure time, and work [24].

Independent Factors

Measures of factors considered as potentially related to change in pain and disability were selected based on a review of the literature and the clinical experience of 2 specialist physiotherapists with over 15 years of experience managing patients following a whiplash injury:

- Self-reported pain medication use (“Do you take pain medication?” Yes/no).
- Symptom satisfaction was measured on a 7-point scale from 1=happy about symptoms to 7=unhappy about symptoms [25].
- Exercise level rating of daily activities and sports (from 1=inactivity to 4=high activity level) [26].
- Dizziness at rest and while in motion and unsteadiness and headache intensity were measured on VAS scales from 0 mm=no dizziness/unsteadiness or headache to 100 mm=extreme dizziness/unsteadiness or headache [23,27].
- The Post-Traumatic Stress Disorder Checklist (the specific trauma version, PCL-S) measures posttraumatic stress reactions and consists of 4 subscales (re-experience, avoidance, negative alterations in cognition and mood, and hyperarousal), with a total score ranging from 0=no stress reaction to 80=extreme stress [28].
- Fear Avoidance Beliefs Questionnaire measures the individual’s fear of pain and consequent avoidance of physical activity because of their fear. It consists of 2 subscales (Physical Activity subscale [4 items] and Work subscale [7 items]), with a total score ranging from 0 to 66, where higher scores indicate higher fear avoidance beliefs [29].
- The Hospital Anxiety and Depression Scale gives an anxiety score (from 0=no anxiety to 21=maximal anxiety) and a depression score (from 0=absent depression to 21=maximal depression) [30].
- The Pain Catastrophizing Scale measures the extent of catastrophizing according to the 3 subscales (rumination, magnification, and helplessness), with a total score range of 0 - 52, with higher scores indicating higher pain catastrophizing [31].
- The Self-Efficacy Scale measures self-perceived performance of 20 everyday functional tasks in spite of pain, with total scores ranging from 0 to 200, where higher scores indicate higher perceived self-efficacy [32].
- The Cognitive Failures Questionnaire measures behavioral problems associated with attentiveness and memory in everyday life, with the total score ranging from 0 to 100, where a higher score indicates a greater frequency of cognitive failures [33].
- The EuroQol quality-of-life instrument version with 5 dimensions and 3 response levels (EQ-5D-3L) and EuroQol health-related quality of life measure using a VAS (EQ-VAS) measure health-related quality of life. EQ-5D-3L scores range from −0.624 to 1 and EQ-VAS scores range from 0 to 100 mm, with higher values indicating better health-related quality of life [34].

The following physiological measures were collected by the test leader at baseline and 3-month and 15-month follow-up:

- Dynamic hand strength was measured in kilograms using the Jamar dynamometer [35].
- Ventral and dorsal neck muscle endurance (NME*)* was registered in seconds [36].
- Active range of cervical motion in sagittal, lateral, and frontal planes was recorded in degrees using the cervical range-of-motion device [37].
- The Cranio-Cervical Flexion test was used to measure the action of the deep cervical flexor muscles, registered in mm Hg [38].
- Balance test: standing on one leg with eyes closed was measured in seconds [27].

XSL•FO
RenderX

Statistical and Multivariate Data Analysis

Multivariate data analysis was used to investigate the relationships between the variables using SIMCA software (v.18.0.0.372; Sartorius Stedim Data Analytics AB). For this, orthogonal partial least squares (OPLS) and OPLS discriminant analysis (OPLS-DA) were used. In general terms, these methods are linear regression methods that model the association between predictors (X) and response (Y) variables. They do this by summarizing the information in the original variables into a few virtual variables (components) that are selected to maximize the covariance between the X and Y variables. This allows the identification of the association between many X and Y variables simultaneously. The methods also separate variation in predictors that are either predictive for the response variables or not (orthogonal). This minimizes the issue of collinearity seen with classical regression models.

First, the differences between patients randomized to either NSEIT or the same exercises and information at a physiotherapy clinic (NSE) [9] were investigated by studying the changes (change scores) of variables between the 3 time points (baseline to 3 mo, baseline to 15 mo, and 3 mo to 15 mo) using OPLS-DA. Both the dependent factors disability (NDI), neck pain (VAS), and physical function (PSFS) and the independent factors (defined above) were included as X variables to investigate which of the variables may explain differences between NSE and NSEIT.

Second, OPLS models were used to investigate the relationship between changes for the dependent and independent factors at the 3 time points (baseline to 3 mo, baseline to 15 mo, and 3 mo to 15 mo). Y variables were set to the change scores for the dependent factors NDI, neck pain, and PSFS (using the average of the 3 PSFS components physical activity, leisure time, and work). X variables included all independent factors and all background variables detailed previously. A positive change score indicates improvement (eg, a higher change score in neck pain indicates less pain) and a negative change score indicates deterioration. To identify which of the X variables (independent factors) were most important for the changes in the Y variables (dependent factors; NDI, neck pain, and PSFS), those independent factors in a first model that had a variable influence on projection value of at least 1.0 and a variable influence on projection value greater than the standard error were included in a second OPLS model with one predictive and one orthogonal component.

Third, multivariate models using OPLS-DA were used to investigate whether background variables and baseline measurements of any of the independent factors could predict clinically significant improvement in the dependent factors NDI, neck pain, and PSFS. Individuals with clinically improved NDI, neck pain, and PSFS (value change NDI ≥7, neck pain ≥50% improvement, PSFS ≥2-point improvement) [22,23,39] were compared to individuals with no improvement or deteriorated values on these outcome measures.

All multivariate models were assessed using the fit of the data to the model (R^2) and the predictability of the model fit on new data (Q^2); the significance of group separation was assessed using cross-validated analysis of variance (CV-ANOVA) [40]. The maximum value for each (R^2 and Q^2) is 1.0 and symbolizes a perfect model. The loadings (weights representing the variables) were used for interpretation of the association between the X (independent factors) and Y (dependent factors) variables.

Sample size and power calculations were calculated based on the primary outcome measure (NDI) in the previous randomized controlled study [8]. To detect a between-group noninferiority margin of 7% with 1-sided α=.025 and β=.8, a total of 47 participants were needed in each group. To account for dropouts, 70 participants were included in each of the 2 groups (140 in total) [9]. Compliance with the exercise program (NSEIT, NSE) was defined as at least 50% self-reported completion of the exercises.

Results

Overview

Baseline characteristics of the participants are shown in Table 1. In the NSEIT group, of the initial 70 participants, 63 (90%) and 56 (80%) participants completed the 3-month and 15-month questionnaires, respectively. In the NSE group, of the initial 70 participants, 64 (91%) and 58 (83%) completed the 3-month and 15-month questionnaires, respectively (Figure 1).

XSL•FO
RenderX

Table . Baseline characteristics of participants with whiplash-associated disorders grade II and III, by treatment group.

	NSEIT[a] (n=70)	NSE[b] (n=70)
Age (years), mean (SD)	40.4 (11.6)	40.5 (11.4)
Sex, female, n (%)	55 (79)	55 (79)
Months since injury, mean (SD)	27.4 (21.0)	25.2 (15.5)
Whiplash-associated disorders, n (%)		
Grade II[c]	46 (66)	43 (61)
Grade III[d]	24 (34)	27 (39)
Educational level, n (%)		
Elementary school	1 (1)	0 (0)
Upper secondary school	30 (43)	39 (57)
University	35 (50)	27 (39)
Other	4 (6)	3 (4)
Marital status, married or cohabiting, n (%)	52 (74)	53 (76)
Previous treatment, yes, n (%)	62 (89)	66 (94)
Use of analgesic drugs, n (%)		
Yes	57 (81)	60 (86)
No	12 (17)	10 (14)
Sick leave, n (%)		
No	61 (87)	62 (89)
Sick leave, part-time	6 (9)	5 (7)
Sick leave, full-time	3 (4)	3 (4)
Disability (Neck Disability Index, 0% - 100%)	39.4 (12.2)	36.6 (10.8)
Neck pain intensity, mean (SD)		
Neck pain now, VAS[e] (0 - 100 mm)	34.6 (21.4)	39.7 (22.4)
Neck pain average previous week, VAS (0 - 100 mm)	44.9 (20.6)	47.6 (21.7)
Patient-Specific Functional Scale, mean (SD)		
Work (0 - 10)	3.9 (2.1)	3.6 (2.3)
Leisure time (0 - 10)	3.7 (2.1)	3.6 (2.3)
Activity (0 - 10)	3.2 (2.5)	3.5 (2.8)

[a]NSEIT: internet-based neck-specific exercise program.

[b]NSE: neck-specific exercise program supervised by a physiotherapist.

[c]Whiplash-associated disorders grade II: neck pain and musculoskeletal findings during a physical examination.

[d]Grade III: like grade II but with additional neurological findings.

[e]VAS: visual analogue scale.

XSL•FO
RenderX

Figure 1. Flowchart of participants in this study. WAD: whiplash-associated disorder.

XSL•FO
RenderX

During the 12-week intervention period, the participants' self-reported compliance with the exercise program was higher in the NSE group (59/64, 92%) than in the NSEIT group (46/63, 73%; *P*=.006). No adverse events were reported.

Between-Group Differences

No significant OPLS-DA models could be built to separate patients receiving NSEIT from those receiving NSE (baseline to 3 mo: R^2=0.22, Q^2=–0.12; baseline to 15 mo: R^2=0.22, Q^2=0.01; 3 mo to 15 mo: R^2=0.17, Q^2=–0.10; all models CV-ANOVA *P*>.99). There were also no differences between the NSEIT and NSE groups regarding independent factors associated with improvements in the dependent factors NDI, neck pain, and PSFS (Figures S1-S3 in Multimedia Appendix 1).

Factors Related to Improvements in Disability, Neck Pain, and Physical Function

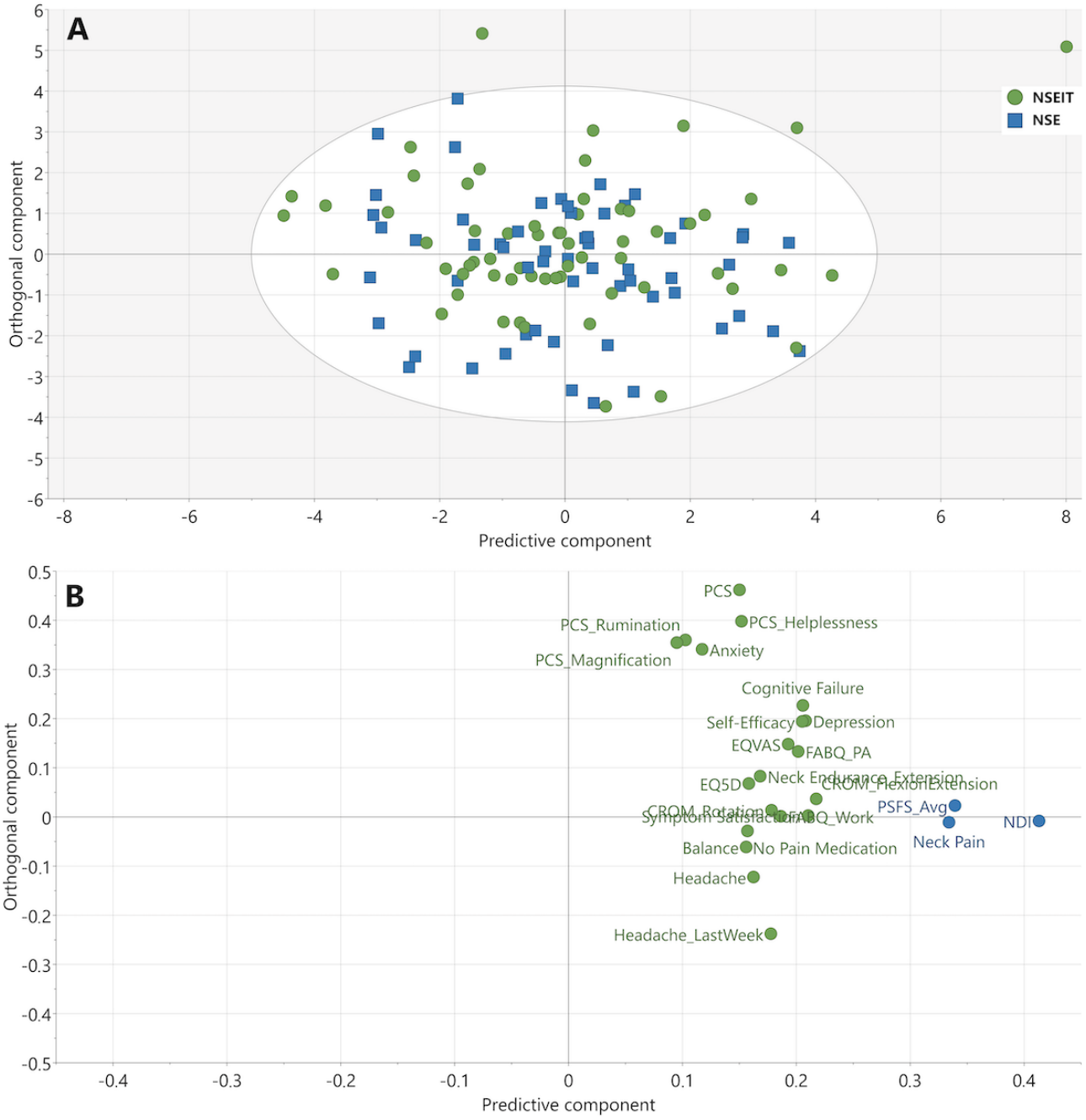
No background factors were found to be associated with improvements in NDI, neck pain, or PSFS in any of the multivariate models.

Improvement From Baseline to 3-Month Follow-Up

At 3 months compared to baseline, 20 independent factors were identified as associated with improvements in NDI, neck pain, or PSFS (R^2=0.31, Q^2=0.25, CV-ANOVA *P*<.001; Figure 2).

XSL•FO
RenderX

Figure 2. OPLS model depicting changes in independent factors from baseline to 3 months. (A) Score plot of the participating patients. The color corresponds to grouping in either NSE (blue squares) or NSEIT (green circles). There were no differences between the NSEIT and NSE groups. (B) Loading plot of analyzed variables, showing the contribution of each independent factor to the OPLS component underlying the participant separation in the score plot (A). The x-axis depicts the most important predictive component for separating participants based on improvements in NDI, neck pain, and PSFS (blue circles) in both the NSEIT and NSE groups (R^2=0.31, Q^2=0.25, cross-validated ANOVA P<.001). CROM: cervical range of motion; EQ5D: EuroQol health-related quality-of-life scale; EQ-VAS: EuroQol visual analogue scale; FABQ: Fear Avoidance Beliefs Questionnaire; NDI: Neck Disability Index; NSE: neck-specific exercise program supervised by a physiotherapist; NSEIT: internet-based neck-specific exercise program; OPLS: orthogonal partial least squares; PCS: Pain Catastrophizing Scale; PSFS_Avg: Patient-Specific Functional Scale average (the average of the 3 PSFS components physical activity, leisure time, and work).

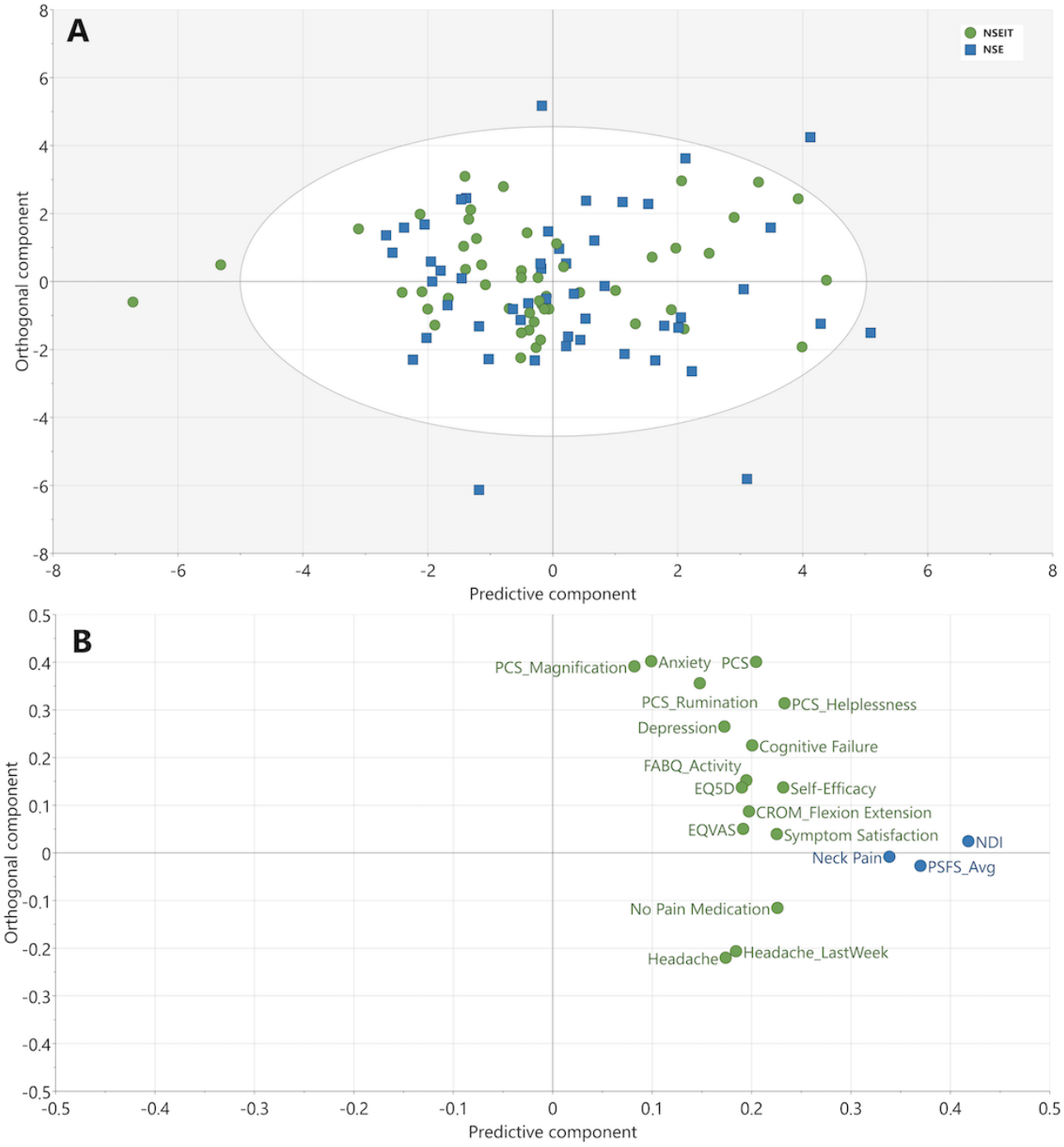


Improvement From Baseline to 15-Month Follow-Up

At 15 months compared to baseline, 16 independent factors were identified as associated with improvements in NDI, neck pain, and PSFS (R^2=0.37, Q^2=0.30, CV-ANOVA P<.001 for NDI and neck pain and P<.01 for PSFS; Figure 3).

XSL•FO
RenderX

Figure 3. OPLS model depicting changes in independent factors from baseline to 15 months. (A) Score plot of the participating patients. The color corresponds to grouping in either NSE (blue squares) or NSEIT (green circles). There were no differences between the NSEIT and NSE groups. (B) Loading plot of analyzed variables, showing the contribution of each independent factor to the OPLS component underlying the participant separation in the score plot (A). The x-axis depicts the most important predictive component for separating participants based on improvements in the primary outcomes NDI, neck pain, and PSFS (blue circles; R^2=0.37, Q^2=0.30, cross-validated ANOVA P<.001 for NDI and neck pain and P<.01 for PSFS). CROM: cervical range of motion; EQ5D: EuroQol health-related quality-of-life scale; EQ-VAS: EuroQol visual analogue scale; FABQ: Fear Avoidance Beliefs Questionnaire; NDI: Neck Disability Index; NSE: neck-specific exercise program supervised by a physiotherapist; NSEIT: internet-based neck-specific exercise program; OPLS: orthogonal partial least squares; PCS: Pain Catastrophizing Scale; PSFS_Avg: Patient-Specific Functional Scale average (the average of the 3 PSFS components physical activity, leisure time, and work).

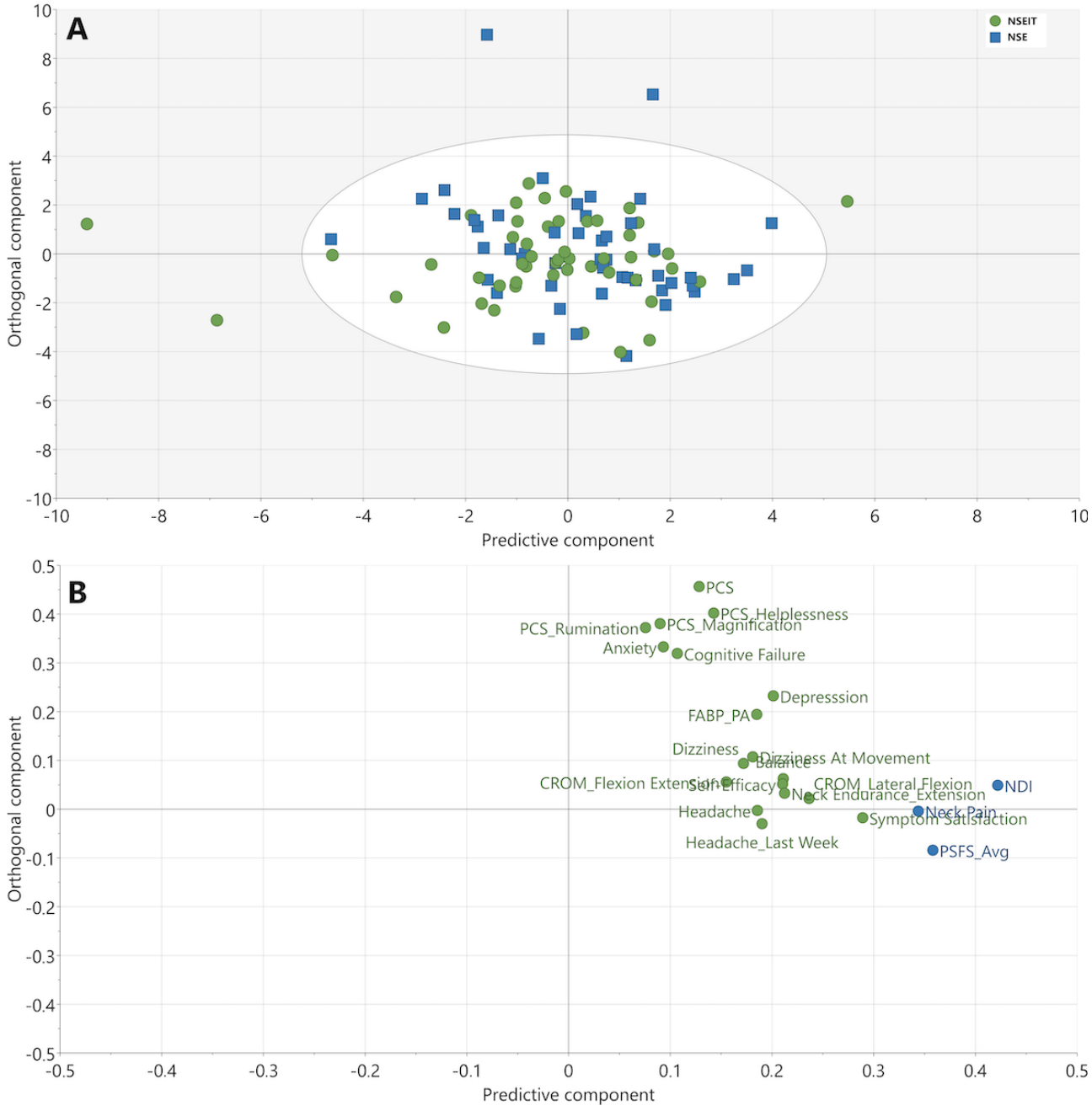


Improvement From 3-Month to 15-Month Follow-Up

At 15 months compared to 3 months, 18 independent factors were identified to be associated with improvements in NDI, neck pain, and PSFS (R^2=0.34, Q^2=0.28, CV-ANOVA P<.001; Figure 4).

XSL•FO
RenderX

Figure 4. OPLS model depicting changes in independent factors from 3 months to 15 months. (A) Score plot of participating patients. The color corresponds to grouping in either NSE (blue squares) or NSEIT (green circles). There were no differences between the NSEIT and NSE groups. (B) Loading plot of analyzed variables, showing the contribution of each independent factor to the OPLS component underlying the participant separation in the score plot (A). The x-axis depicts the most important predictive component for separating participants based on improvements in the primary outcomes NDI, neck pain, and PSFS (blue circles; R^2=0.34, Q^2=0.28, cross-validated ANOVA P<.001). CROM: cervical range of motion; EQ5D: EuroQol health-related quality-of-life scale; EQ-VAS: EuroQol visual analogue scale; FABQ: Fear Avoidance Beliefs Questionnaire; NDI: Neck Disability Index; NSE: neck-specific exercise program supervised by a physiotherapist; NSEIT: internet-based neck-specific exercise program; OPLS: orthogonal partial least squares; PCS: Pain Catastrophizing Scale; PSFS_Avg: Patient-Specific Functional Scale average (the average of the 3 PSFS components physical activity, leisure time, and work).

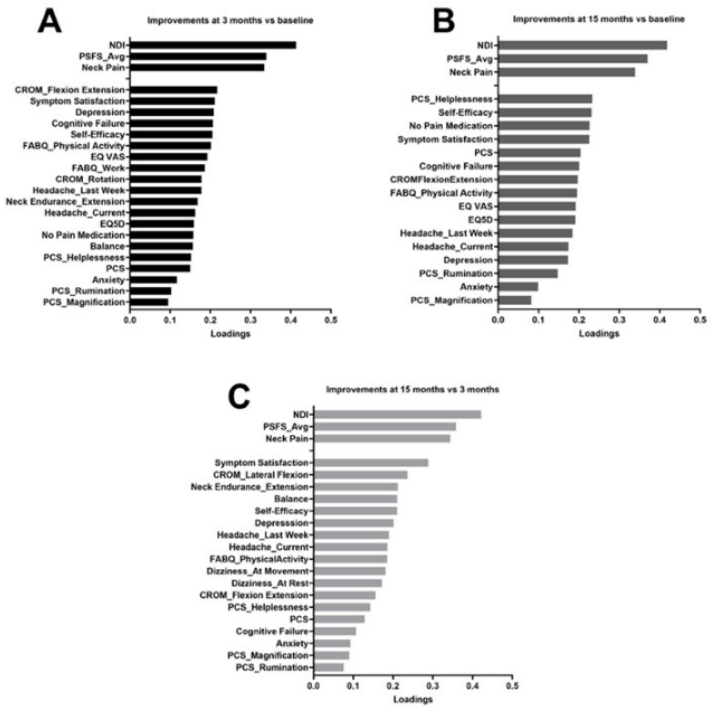


Variable Loadings for the 3 Models

The contribution of each variable (loadings) to the predictive component (x-axis) for the 3 models (Figures 1-3) is visualized in Figure 5. Thirteen of the 23 identified independent factors were common for all 3 models and included both psychological and physiological factors. The most important factors, as determined by a higher loading, that were related to improvements in NDI, neck pain, and PSFS were as follows: anxiety and depression, cognitive failure, pain catastrophizing, self-efficacy, fear avoidance beliefs, cervical range of motion, headache last week, current headache, and symptom satisfaction. Factors only found in 1 or 2 models were fear avoidance beliefs subscale work, balance (standing on one leg), cervical motion (lateral flexion and rotation), neck muscle endurance (extension), taking no pain medication, dizziness at rest and at movement, and health-related quality of life (Figure 5).

XSL•FO
RenderX

Figure 5. Loadings (contribution of each independent factor) obtained through OPLS after selecting outcome measures with variable influence on projection greater than 1.0. The three models (A-C) show the association between independent factors for improvements in NDI, neck pain, and PSFS at different time points. (A) Improvements between baseline and 3-month follow-up. (B) Improvements between baseline and 15-month follow-up. (C) Improvements between 3-month and 15-month follow-up. A higher value for the loadings indicates it is more important in the models. CROM: cervical range of motion; EQ5D: EuroQol health-related quality-of-life scale; EQ-VAS: EuroQol visual analogue scale; FABQ: Fear Avoidance Beliefs Questionnaire; NDI: Neck Disability Index; NSE: neck-specific exercise program supervised by a physiotherapist; NSEIT: internet-based neck-specific exercise program; OPLS: orthogonal partial least squares; PCS: Pain Catastrophizing Scale; PSFS_Avg: Patient-Specific Functional Scale average (the average of the 3 PSFS components physical activity, leisure time, and work).



The 5 most important independent factors related to changes in NDI, neck pain, and PSFS between baseline and 3-month follow-up were improvement in cervical range of motion (flexion/extension), symptom satisfaction, depression, cognitive failure, and self-efficacy (Figure 5A).

The 5 most important independent factors related to changes in NDI, neck pain, and PSFS between baseline and 15-month follow-up were improvement in pain catastrophizing (helplessness), self-efficacy, taking no pain medication, symptom satisfaction, and pain catastrophizing (total score; Figure 5B).

The 5 most important independent factors related to changes in NDI, neck pain, and PSFS between 3-month and 15-month follow-up were improvement in symptom satisfaction, cervical range of motion (lateral flexion), NME (extension), balance, and self-efficacy (Figure 5C).

Baseline Factors Determining Clinically Improved Versus Nonimproved Patients in Disability, Neck Pain, and Physical Function

Using OPLS-DA modeling, no significant models could be built to separate those patients whose NDI, neck pain, or PSFS was clinically improved from those with no improvement or showing deterioration based on background variables and baseline measurements of any of the independent factors (Figures S4-S9 in Multimedia Appendix 1).

Discussion

Principal Findings

This secondary analysis of a prospective randomized study investigated factors related to change in neck pain and disability (including functional ability) after a NSEIT or NSE exercise program for individuals with chronic WAD grade II or III.

The multivariate OPLS-DA models showed no significant differences between NSEIT and NSE. Thus, factors related to changes in pain and disability were the same in both groups, independent of how the exercise program was delivered. On the other hand, the multivariate OPLS models of improvements in neck pain and disability between the 3 time points (baseline to 3 mo, baseline to 15 mo, and 3 mo to 15 mo) indicated a number of related factors in both NSEIT and NSE. In total, 23 factors were identified and 13 of them featured in all 3 time points. However, neither the baseline factors nor changes in the independent factors could predict which patients were improved versus nonimproved in pain and disability.

Improvements in psychological and physiological factors were associated with an improvement in disability, pain, and physical function in both groups (NSEIT and NSE). It seems reasonable to assume that the effect may be related to the neck-specific exercise intervention itself, regardless of whether it was internet-based or physiotherapist-supervised, because the neck-specific exercise and information program were identical in NSEIT and NSE. The 4 face-to-face visits in NSEIT included

XSL•FO
RenderX

examination, an introduction to performing the exercises, progression of exercises, and follow-up. A possible explanation is that 4 visits to the physiotherapist combined with the internet-based program enabled these patients to feel confident and skilled enough to perform the exercises to the same extent as participants in the NSE group, who had many more visits to the physiotherapist.

In our study, participants in both groups described having increased knowledge regarding WAD, a stronger neck, and fewer symptoms [41]. Moreover, no significant differences were seen between NSEIT and NSE for primary or secondary outcomes [9,42].

The interpretation of the multivariate OPLS models was based on the weights (loadings) reflecting the relationship between dependent factors (NDI, neck pain, PSFS) and independent factors. Thus, independent factors with higher loadings were more closely related to improvements in NDI, neck pain, and PSFS compared to outcomes with lower loadings in the same model (Figure 5). Improvements in self-efficacy and symptom satisfaction were prominent in all 3 models (baseline to 3 mo, baseline to 15 mo, and 3 mo to 15 mo). Neck-specific exercises have been shown to improve neck-related function [42] and, in combination with the patient education in our NSEIT/NSE program, the exercises may lead to an enhanced ability to perform everyday functional activities, as the improvements in NDI and PSFS indicate [9].

Comparison to Prior Work

Psychological factors such as low self-efficacy, fear avoidance beliefs, depressive symptoms, and catastrophizing have earlier been shown to be predictors of a poor outcome of neck disability in chronic WAD [43-47], but they are also shown to be improved after 3 months of neck-specific exercise [48]. One of the most important factors to improve self-efficacy is personal mastery experiences, that is, the ability to successfully perform a task [49]. When individuals in the NSEIT and NSE groups could perform activities with less pain, their self-efficacy and symptom satisfaction increased. A painful stimulus, such as neck pain, can trigger catastrophic thinking, fear-avoidance beliefs, and avoidance behaviors, resulting in more pain and depression and leading to chronic disability as described by Vlaeyen and Linton [50] in their fear-avoidance model. Our neck-specific exercise program included information on whiplash injury, outlining why pain can persist even after potential injuries to muscles or ligaments have healed, and highlighting the importance of exercising the neck muscles. The exercises should not exacerbate neck pain but should be progressed over time, leading to the activity level recommended by the World Health Organization, as described in detail previously [9].

Physical activities and exercises are highly effective for improving symptoms of depression [51]. This may explain the association between improvements in NDI and PSFS (physical function) and depressive symptoms in this study. A speculative explanation is that the effect of our neck-specific exercise program is related to mutually positive associations. Physical function and self-efficacy increased—while pain, fear avoidance, and depression decreased—as a result of the exercise program. Anxiety and pain catastrophizing were included in the 3 models but had a lower influence on the dependent factors (neck pain and disability). Improvement in cognitive failures was also related to an improvement in pain and disability in all 3 models. Previous studies have shown that cognitive failures and attention problems are common and disabling when living with pain [52], a finding that is not limited to subjective reports, as these patients also fail in standardized cognitive tests [53]. However, another study found the relationship between chronic pain and cognitive tests in whiplash inconclusive [54]. Coexisting brain injury due to the whiplash trauma may be another reason for cognitive failures independent of pain [53], although individuals with brain injury were excluded in this study. Our results indicate that improvement in pain and disability after NSEIT or NSE was to some extent related to improvement in anxiety, pain catastrophizing, and cognitive failures. A causal relationship cannot be established in this study, but the findings may be of importance for improved rehabilitation in chronic WAD.

Background variables and baseline values of the independent factors in our study could not separate individuals with clinical improvement versus nonimprovement in disability, neck pain, and physical function. Moreover, no background variables were found to be associated with improvements in NDI, neck pain, or PSFS. A possible reason is that we failed to include the most relevant factors in the analyses despite the inclusion of many variables in the OPLS and OPLS-DA models. However, we included background variables and independent factors that were previously reported in the literature to be important in predicting and relating to disability and pain in WAD and chronic pain [10,11,14,43,44,46,47,55-57]. Nonetheless, there are conflicting results in earlier studies regarding the importance of some of the background variables and baseline values of the independent factors to pain intensity in chronic pain [55]. In chronic WAD, female gender, older age, and WAD grade III were related to nonrecovery in one study [56], while another study found that pain-related catastrophizing explained the variance in disability in patients with chronic WAD [44]. In contrast, we have previously found that no background variables were predictive of pain reduction in chronic WAD at 1-year follow-up [14]. However, higher baseline NDI score, absence of dizziness, and higher patient expectations, but not psychological factors, were associated with disability reduction [14]. We have also previously found that neck muscle endurance and perceived work ability were mediators for reduction in pain [15] in chronic WAD, and we identified a lack of improvement in ventral NME as a factor for ongoing dizziness in WAD after an exercise intervention [57]. The participants in this study had higher neck muscle endurance at baseline [42] compared to earlier results [58] and they improved more at the 3-month follow-up [42,58]. Higher neck muscle endurance may reduce disability and pain, which in turn can influence other factors and may explain the discrepancy between our results and other studies [12,44,56].

We recently reported that between 47% and 74% of participants with chronic WAD grade II and III demonstrated sustained improvement in disability, pain, and physical function after 3 months of NSEIT or NSE [9], but we failed to identify factors

XSL•FO
RenderX

associated with responders and nonresponders. Differences in analysis methods may explain the discrepancy between the present results and other studies [12,44,56]. An interesting result in this study was that neither psychological factors such as low self-efficacy, depression, and fear avoidance, nor background characteristics such as older age, low disability level, or WAD grade III, could predict nonresponders. Further studies should include factors related to possible tissue damage or altered function in chronic WAD, but one problem is that there is still a lack of reliable and effective diagnostic tools.

The results indicate that we need to go further to be able to improve our understanding of chronic WAD and to optimize treatment. Biomedical and/or psychosocial factors not investigated in this study are likely to be important for identifying nonresponders to NSEIT and NSE in chronic WAD.

Strengths and Limitations

This is the first study investigating factors related to pain and disability after an internet-based exercise program for chronic WAD grade II and III patients and, to our knowledge, factors related to the outcome of PSFS have not previously been reported for chronic WAD.

The strengths of multivariate OPLS analysis are the ability to include many variables and observations and the fact that several response variables (dependent factors) can be modeled together. OPLS can also cope with multicollinearity and missing values, which is a major advantage compared with traditional regression models. OPLS models use latent variables to model the data; these latent variables, which are linear combinations of the original variables, reduce the dimensions and minimize the issue of multicollinearity. Furthermore, OPLS separates the predictive variation (related to the response variable) from the orthogonal variation (unrelated to the response variable), thus focusing on the most relevant information, which simplifies the model and improves interpretability [59-61]. Moreover, the benefits of using multivariate statistics in this analysis are that variables can be scaled and centered on the mean, implying that each variable will have the same potential to affect the model, and that a patient or cluster of patients can be related to a variable profile rather than a single variable.

There are several limitations in this study. First, in the OPLS analyses, all 3 models (baseline to 3 mo, baseline to 15 mo, and 3 mo to 15 mo) showed relatively low explained variance, with the highest explanation level for the baseline to 15-month follow-up (R^2=0.37, Q^2=0.30). The goodness of fit of a model varies depending on the type of data and the specific application. For example, spectroscopic calibration models require values close to 1.0. However, while high R^2 and Q^2 values are desirable, they must be interpreted in the context of the specific application and data type [62]. Lower values indicate a less robust fit and predictive ability but might be acceptable, for example, in OPLS models that include self-reported psychological symptoms. Second, the results of this study are only generalizable to those individuals who fulfill the study criteria and are motivated to participate in exercise, but not to all individuals with WAD. Moreover, long-term effects can be influenced by various factors. For example, consistently measuring outcomes over time can introduce bias and affect the reliability of the results due to changes in the patient's life situation. Additionally, dropouts commonly increase over time. Third, only individuals who understood the Swedish language were included, and the age range was limited to 18 - 63 years. Thus, this study excluded minorities from a different background who did not understand Swedish and the results cannot be generalized to those aged under 18 or over 63 years.

Conclusions

After 3 months of NSEIT or NSE, improvements in neck pain and disability were related to improvements in 23 other physical and psychological factors. However, the findings should be interpreted with caution, as the results showed relatively low explained variance. No recommendations can be given for which individuals will benefit from NSEIT or NSE, due to the study's inability to distinguish between clinically improved and nonimproved participants regarding neck pain and disability. Further research is warranted to improve our understanding of those who are improved or not improved after 3 months of NSE or NSEIT.

Acknowledgments

We would like to thank all the participants with whiplash-associated disorders, the primary care physiotherapists, the test leader, and the research staff. The funder had no role in the study design, data collection and analyses, decision to publish, or preparation of the manuscript. The study was funded by Vinnova (grant number 2018- - 02244), the Swedish Research Council (grant number 2018- - 02476), the Centre for Clinical Research Sörmland at Uppsala University Sweden (grant numbers 736751, 854531, and 930399), the Regional Research Council Mid-Sweden (grant numbers 641551, 838701, and 939243), the Medical Research Council of Southeast Sweden (grant numbers 844451 and 939838) and Östergötland County Council (grant numbers 676261 and 847691).

Data Availability

The datasets generated during this study are available from the corresponding author on reasonable request and after ethical approval.

XSL•FO
RenderX

Authors' Contributions

GP and AP designed, formulated, and created the study. GP, AP, and SL analyzed and interpreted the data. GP, AP, and SL wrote the manuscript. GP, AP, and SL reviewed and approved the final manuscript. GP and AP were responsible for the management and coordination of the research study. GP and AP are responsible for the presentation of the published work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Orthogonal partial least squares models.
[PDF File, 1657 KB - humanfactors_v12i1e67991_app1.pdf]

References

1. Holm LW, Carroll LJ, Cassidy JD, et al. The burden and determinants of neck pain in whiplash-associated disorders after traffic collisions: results of the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders. J Manipulative Physiol Ther 2009 Feb;32(2 Suppl):S61-S69. [doi: 10.1016/j.jmpt.2008.11.011] [Medline: 19251076]
2. Styrke J, Sojka P, Björnstig U, Stålnacke BM. Symptoms, disabilities, and life satisfaction five years after whiplash injuries. Scand J Pain 2014 Oct 1;5(4):229-236. [doi: 10.1016/j.sjpain.2014.06.001] [Medline: 29911576]
3. Carroll LJ, Holm LW, Hogg-Johnson S, et al. Course and prognostic factors for neck pain in whiplash-associated disorders (WAD): results of the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders. J Manipulative Physiol Ther 2009 Feb;32(2 Suppl):S97-S107. [doi: 10.1016/j.jmpt.2008.11.014] [Medline: 19251080]
4. Leth-Petersen S, Rotger GP. Long-term labour-market performance of whiplash claimants. J Health Econ 2009 Sep;28(5):996-1011. [doi: 10.1016/j.jhealeco.2009.06.013] [Medline: 19683817]
5. Merrick D, Stålnacke BM. Five years post whiplash injury: symptoms and psychological factors in recovered versus non-recovered. BMC Res Notes 2010 Jul 13;3:190. [doi: 10.1186/1756-0500-3-190] [Medline: 20626861]
6. Ritchie C, Smith A, Sterling M. Medical and allied health service use during acute and chronic post-injury periods in whiplash injured individuals. BMC Health Serv Res 2020 Mar 30;20(1):260. [doi: 10.1186/s12913-020-05146-0] [Medline: 32228573]
7. Rebbeck T. The role of exercise and patient education in the noninvasive management of whiplash. J Orthop Sports Phys Ther 2017 Jul;47(7):481-491. [doi: 10.2519/jospt.2017.7138] [Medline: 28622489]
8. Ludvigsson ML, Peterson G, Dedering Å, Peolsson A. One- and two-year follow-up of a randomized trial of neck-specific exercise with or without a behavioural approach compared with prescription of physical activity in chronic whiplash disorder. J Rehabil Med 2016 Jan;48(1):56-64. [doi: 10.2340/16501977-2041] [Medline: 26660722]
9. Peterson G, Peolsson A. Efficacy of neck-specific exercise with internet support versus neck-specific exercise at a physiotherapy clinic in chronic whiplash-associated disorders: multicenter randomized controlled noninferiority trial. J Med Internet Res 2023 Jun 20;25:e43888. [doi: 10.2196/43888] [Medline: 37338972]
10. Falla D, Peolsson A, Peterson G, et al. Perceived pain extent is associated with disability, depression and self-efficacy in individuals with whiplash-associated disorders. Eur J Pain 2016 Oct;20(9):1490-1501. [doi: 10.1002/ejp.873] [Medline: 27149920]
11. Pedrero-Martin Y, Falla D, Martinez-Calderon J, Liew BXW, Scutari M, Luque-Suarez A. Self-efficacy beliefs mediate the association between pain intensity and pain interference in acute/subacute whiplash-associated disorders. Eur Spine J 2021 Jun;30(6):1689-1698. [doi: 10.1007/s00586-021-06731-5] [Medline: 33502610]
12. Wicksell RK, Olsson GL, Hayes SC. Psychological flexibility as a mediator of improvement in Acceptance and Commitment Therapy for patients with chronic pain following whiplash. Eur J Pain 2010 Nov;14(10):1059. [doi: 10.1016/j.ejpain.2010.05.001] [Medline: 20538493]
13. Angst F, Gantenbein AR, Lehmann S, et al. Multidimensional associative factors for improvement in pain, function, and working capacity after rehabilitation of whiplash associated disorder: a prognostic, prospective outcome study. BMC Musculoskelet Disord 2014 Apr 16;15:130. [doi: 10.1186/1471-2474-15-130] [Medline: 24739588]
14. Ludvigsson ML, Peterson G, Dedering Å, Falla D, Peolsson A. Factors associated with pain and disability reduction following exercise interventions in chronic whiplash. Eur J Pain 2016 Feb;20(2):307-315. [doi: 10.1002/ejp.729] [Medline: 26031995]
15. Liew BXW, Scutari M, Peolsson A, Peterson G, Ludvigsson ML, Falla D. Investigating the causal mechanisms of symptom recovery in chronic whiplash-associated disorders using Bayesian networks. Clin J Pain 2019 Aug;35(8):647-655. [doi: 10.1097/AJP.0000000000000728] [Medline: 31169550]
16. Kovačević I, Pavić J, Filipović B, Ozimec Vulinec Š, Ilić B, Petek D. Integrated approach to chronic pain-the role of psychosocial factors and multidisciplinary treatment: a narrative review. Int J Environ Res Public Health 2024 Aug 28;21(9):1135. [doi: 10.3390/ijerph21091135] [Medline: 39338018]

XSL•FO
RenderX

17. Sarrami P, Armstrong E, Naylor JM, Harris IA. Factors predicting outcome in whiplash injury: a systematic meta-review of prognostic factors. J Orthop Traumatol 2017 Mar;18(1):9-16. [doi: 10.1007/s10195-016-0431-x] [Medline: 27738773]
18. Kipping S, Stuckey MI, Hernandez A, Nguyen T, Riahi S. A web-based patient portal for mental health care: benefits evaluation. J Med Internet Res 2016 Nov 16;18(11):e294. [doi: 10.2196/jmir.6483] [Medline: 27852556]
19. Peolsson A, Landén Ludvigsson M, Peterson G. Neck-specific exercises with internet-based support compared to neck-specific exercises at a physiotherapy clinic for chronic whiplash-associated disorders: study protocol of a randomized controlled multicentre trial. BMC Musculoskelet Disord 2017 Dec 12;18(1):524. [doi: 10.1186/s12891-017-1853-1] [Medline: 29233141]
20. Spitzer WO, Skovron ML, Salmi LR, et al. Scientific monograph of the Quebec Task Force on Whiplash-Associated Disorders: redefining 'whiplash' and its management. Spine (Phila Pa 1976) 1995 Apr 15;20(8 Suppl):1S-73S. [Medline: 7604354]
21. Ardern CL, Peterson G, Ludvigsson ML, Peolsson A. Satisfaction with the outcome of physical therapist-prescribed exercise in chronic whiplash-associated disorders: secondary analysis of a randomized clinical trial. J Orthop Sports Phys Ther 2016 Aug;46(8):640-649. [doi: 10.2519/jospt.2016.6136] [Medline: 27374016]
22. Vernon H. The Neck Disability Index: state-of-the-art, 1991-2008. J Manipulative Physiol Ther 2008 Sep;31(7):491-502. [doi: 10.1016/j.jmpt.2008.08.006] [Medline: 18803999]
23. Dworkin RH, Turk DC, McDermott MP, et al. Interpreting the clinical importance of group differences in chronic pain clinical trials: IMMPACT recommendations. Pain 2009 Dec;146(3):238-244. [doi: 10.1016/j.pain.2009.08.019] [Medline: 19836888]
24. Stewart M, Maher CG, Refshauge KM, Bogduk N, Nicholas M. Responsiveness of pain and disability measures for chronic whiplash. Spine (Phila Pa 1986) 2007 Mar;32(5):580-585. [doi: 10.1097/01.brs.0000256380.71056.6d]
25. Cherkin DC, Deyo RA, Street JH, Barlow W. Predicting poor outcomes for back pain seen in primary care using patients own criteria. Spine (Phila Pa 1986) 1996 Dec;21(24):2900-2907. [doi: 10.1097/00007632-199612150-00023]
26. Kallings LV, Leijon M, Hellénius ML, Ståhle A. Physical activity on prescription in primary health care: a follow-up of physical activity level and quality of life. Scand J Med Sci Sports 2008 Apr;18(2):154-161. [doi: 10.1111/j.1600-0838.2007.00678.x] [Medline: 17555539]
27. Kammerlind AS, Bergquist Larsson P, Ledin T, Skargren E. Reliability of clinical balance tests and subjective ratings in dizziness and disequilibrium. Adv Physiother 2005 Jan;7(3):96-107. [doi: 10.1080/14038190510010403]
28. Wilkins KC, Lang AJ, Norman SB. Synthesis of the psychometric properties of the PTSD checklist (PCL) military, civilian, and specific versions. Depress Anxiety 2011 Jul;28(7):596-606. [doi: 10.1002/da.20837] [Medline: 21681864]
29. Dedering A, Börjesson T. Assessing fear-avoidance beliefs in patients with cervical radiculopathy. Physiother Res Int 2013 Dec;18(4):193-202. [doi: 10.1002/pri.1545] [Medline: 23233349]
30. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983 Jun;67(6):361-370. [doi: 10.1111/j.1600-0447.1983.tb09716.x] [Medline: 6880820]
31. Osman A, Barrios FX, Gutierrez PM, Kopper BA, Merrifield T, Grittmann L. The Pain Catastrophizing Scale: further psychometric evaluation with adult samples. J Behav Med 2000 Aug;23(4):351-365. [doi: 10.1023/a:1005548801037] [Medline: 10984864]
32. Altmaier EM, Russell DW, Kao CF, Lehmann TR, Weinstein JN. Role of self-efficacy in rehabilitation outcome among chronic low back pain patients. J Couns Psychol 1993;40(3):335-339. [doi: 10.1037/0022-0167.40.3.335]
33. Bridger RS, Johnsen S, Brasher K. Psychometric properties of the Cognitive Failures Questionnaire. Ergonomics 2013;56(10):1515-1524. [doi: 10.1080/00140139.2013.821172] [Medline: 23879800]
34. Soer R, Reneman MF, Speijer B, Coppes MH, Vroomen P. Clinimetric properties of the EuroQol-5D in patients with chronic low back pain. Spine J 2012 Nov;12(11):1035-1039. [doi: 10.1016/j.spinee.2012.10.030] [Medline: 23199409]
35. Peolsson A, Hedlund R, Oberg B. Intra- and inter-tester reliability and reference values for hand strength. J Rehabil Med 2001 Jan;33(1):36-41. [doi: 10.1080/165019701300006524] [Medline: 11480468]
36. Peolsson A, Almkvist C, Dahlberg C, Lindqvist S, Pettersson S. Age- and sex-specific reference values of a test of neck muscle endurance. J Manipulative Physiol Ther 2007;30(3):171-177. [doi: 10.1016/j.jmpt.2007.01.008] [Medline: 17416270]
37. Williams MA, Williamson E, Gates S, Cooke MW. Reproducibility of the cervical range of motion (CROM) device for individuals with sub-acute whiplash associated disorders. Eur Spine J 2012 May;21(5):872-878. [doi: 10.1007/s00586-011-2096-8] [Medline: 22139052]
38. Jull GA, O'Leary SP, Falla DL. Clinical assessment of the deep cervical flexor muscles: the craniocervical flexion test. J Manipulative Physiol Ther 2008 Sep;31(7):525-533. [doi: 10.1016/j.jmpt.2008.08.003] [Medline: 18804003]
39. Cleland JA, Fritz JM, Whitman JM, Palmer JA. The reliability and construct validity of the Neck Disability Index and patient specific functional scale in patients with cervical radiculopathy. Spine (Phila Pa 1976) 2006 Mar 1;31(5):598-602. [doi: 10.1097/01.brs.0000201241.90914.22] [Medline: 16508559]
40. Wheelock Å, Wheelock CE. Trials and tribulations of 'omics data analysis: assessing quality of SIMCA-based multivariate models using examples from pulmonary medicine. Mol Biosyst 2013 Nov;9(11):2589-2596. [doi: 10.1039/c3mb70194h] [Medline: 23999822]

XSL•FO
RenderX

41. Peolsson A, Hermansen A, Peterson G, Nilsing Strid E. Return to work a bumpy road: a qualitative study on experiences of work ability and work situation in individuals with chronic whiplash-associated disorders. BMC Public Health 2021 Apr 23;21(1):785. [doi: 10.1186/s12889-021-10821-w] [Medline: 33892672]
42. Peterson G, Nilsing Strid E, Jönsson M, Hävermark J, Peolsson A. Effect of neck-specific exercises with and without internet support on cervical range of motion and neck muscle endurance in chronic whiplash-associated disorders: analysis of functional outcomes of a randomized controlled trial. J Rehabil Med 2024 Jul 29;56:jrm34785. [doi: 10.2340/jrm.v56.34785] [Medline: 39072427]
43. Börsbo B, Peolsson M, Gerdle B. Catastrophizing, depression, and pain: correlation with and influence on quality of life and health - a study of chronic whiplash-associated disorders. J Rehabil Med 2008 Jul;40(7):562-569. [doi: 10.2340/16501977-0207] [Medline: 18758674]
44. Schmitt MA, van Meeteren NL, de Wijer A, van Genderen FR, van der Graaf Y, Helders PJ. Patients with chronic whiplash-associated disorders: relationship between clinical and psychological factors and functional health status. Am J Phys Med Rehabil 2009 Mar;88(3):231-238. [doi: 10.1097/PHM.0b013e318198b684] [Medline: 19847132]
45. Vernon H, Guerriero R, Kavanaugh S, Soave D, Moreton J. Psychological factors in the use of the Neck Disability Index in chronic whiplash patients. Spine (Phila Pa 1986) 2010 Jan;35(1):E16-E21. [doi: 10.1097/BRS.0b013e3181b135aa]
46. van der Meer S, Reneman MF, Verhoeven J, van der Palen J. Relationship between self-reported disability and functional capacity in patients with whiplash associated disorder. J Occup Rehabil 2014 Sep;24(3):419-424. [doi: 10.1007/s10926-013-9473-6] [Medline: 24288054]
47. Thompson DP, Oldham JA, Urmston M, Woby SR. Cognitive determinants of pain and disability in patients with chronic whiplash-associated disorder: a cross-sectional observational study. Physiotherapy 2010 Jun;96(2):151-159. [doi: 10.1016/j.physio.2009.11.001] [Medline: 20420962]
48. Overmeer T, Peterson G, Landén Ludvigsson M, Peolsson A. The effect of neck-specific exercise with or without a behavioral approach on psychological factors in chronic whiplash-associated disorders: a randomized controlled trial with a 2-year follow-up. Medicine (Baltimore) 2016 Aug;95(34):e4430. [doi: 10.1097/MD.0000000000004430] [Medline: 27559950]
49. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977 Mar;84(2):191-215. [doi: 10.1037//0033-295x.84.2.191] [Medline: 847061]
50. Vlaeyen JWS, Linton SJ. Fear-avoidance model of chronic musculoskeletal pain: 12 years on. Pain 2012 Jun;153(6):1144-1147. [doi: 10.1016/j.pain.2011.12.009] [Medline: 22321917]
51. Singh B, Olds T, Curtis R, et al. Effectiveness of physical activity interventions for improving depression, anxiety and distress: an overview of systematic reviews. Br J Sports Med 2023 Sep;57(18):1203-1209. [doi: 10.1136/bjsports-2022-106195] [Medline: 36796860]
52. Di Stefano G, Radanov BP. Course of attention and memory after common whiplash: a two-years prospective study with age, education and gender pair-matched patients. Acta Neurol Scand 1995 May;91(5):346-352. [doi: 10.1111/j.1600-0404.1995.tb07019.x] [Medline: 7639063]
53. Kessels RP, Aleman A, Verhagen WI, van Luijtelaar EL. Cognitive functioning after whiplash injury: a meta-analysis. J Int Neuropsychol Soc 2000 Mar;6(3):271-278. [doi: 10.1017/s1355617700633027] [Medline: 10824499]
54. Rathbone M, Parkinson W, Rehman Y, Jiang S, Bhandari M, Kumbhare D. Magnitude and variability of effect sizes for the associations between chronic pain and cognitive test performances: a meta-analysis. Br J Pain 2016 Aug;10(3):141-155. [doi: 10.1177/2049463716642600] [Medline: 27583141]
55. Hunter J. Demographic variables and chronic pain. Clin J Pain 2001 Dec;17(4 Suppl):S14-S19. [doi: 10.1097/00002508-200112001-00006] [Medline: 11783826]
56. Cobo EP, Mesquida MEP, Fanegas EP, et al. What factors have influence on persistence of neck pain after a whiplash? Spine (Phila Pa 1986) 2010 Apr;35(9):E338-E343. [doi: 10.1097/BRS.0b013e3181c9b075]
57. Treleaven J, Peterson G, Ludvigsson ML, Peolsson A. Cervical musculoskeletal, physical and psychological factors associated with ongoing dizziness in patients with whiplash associated disorder, 12 months after undertaking a neck specific or general exercise intervention. BMC Musculoskelet Disord 2022 Jul 18;23(1):683. [doi: 10.1186/s12891-022-05642-w] [Medline: 35850745]
58. Peterson GE, Landén Ludvigsson MH, O'Leary SP, et al. The effect of 3 different exercise approaches on neck muscle endurance, kinesiophobia, exercise compliance, and patient satisfaction in chronic whiplash. J Manipulative Physiol Ther 2015 Sep;38(7):465-476. [doi: 10.1016/j.jmpt.2015.06.011] [Medline: 26387858]
59. Kianoush Fathi V, Robabe M, Homayoun SB. Applications of Opls statistical method in medicine. J Math Computer SCI 2014;08(4):411-422. [doi: 10.22436/jmcs.08.04.09]
60. Homayoun SB, Shrikant IB, Kazem M, Hemmat M, Reza M. Compared application of the new OPLS-DA statistical model versus partial least squares regression to manage large numbers of variables in an injury case-control study. Sci Res Essays 2011;6(20):4369-4377. [doi: 10.5897/SRE10.1147]
61. Eriksson L. Multi- and Megavariate Data Analysis: Basic Principles and Applications: MKS Umetrics; 2013.
62. Trygg J, Wold S. Orthogonal projections to latent structures (O‐PLS). J Chemom 2002 Mar;16(3):119-128. [doi: 10.1002/cem.695]

XSL•FO
RenderX

Abbreviations

CV-ANOVA: cross-validated analysis of variance
EQ-5D-3L: EuroQol quality-of-life instrument version with 5 dimensions and 3 response levels
EQ-VAS: EuroQol health-related quality of life measure using a visual analogue scale
NDI: Neck Disability Index
NME: neck muscle endurance
NSE: neck-specific exercise program supervised by a physiotherapist
NSEIT: internet-based neck-specific exercise program
OPLS: orthogonal partial least squares
OPLS-DA: orthogonal partial least squares discriminant analysis
PSFS: Patient-Specific Functional Scale
VAS: visual analogue scale
WAD: whiplash-associated disorders

Edited by A Kushniruk; submitted 25.10.24; peer-reviewed by C Papic, H Wang; revised version received 08.04.25; accepted 17.04.25; published 30.05.25.

Please cite as:
Peterson G, Ljunggren S, Peolsson A
Factors Related to Pain and Disability Outcomes After an Internet-Delivered or Physiotherapist-Led Exercise Program for Individuals With Chronic Whiplash Symptoms: Secondary Analysis of a Randomized Controlled Study
JMIR Hum Factors 2025;12:e67991
URL: https://humanfactors.jmir.org/2025/1/e67991
doi:10.2196/67991

© Gunnel Peterson, Stefan Ljunggren, Anneli Peolsson. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 30.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

A User-Centered Design Approach for a Screening App for People With Cognitive Impairment (digiDEM-SCREEN): Development and Usability Study

Michael Zeiler[1*], MSc; Nikolas Dietzel[2*], PhD; Fabian Haug[3,4], MSc; Julian Haug[3,4], MSc; Klaus Kammerer[3], MSc; Rüdiger Pryss[3,4], PhD; Peter Heuschmann[3,4,5], MD, PhD; Elmar Graessel[6], MD, PhD; Peter L Kolominsky-Rabas[2], MD, MBA, PhD; Hans-Ulrich Prokosch[1], PhD

[1]Medical Informatics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

[2]Interdisciplinary Center for Health Technology Assessment and Public Health, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

[3]Institute of Clinical Epidemiology and Biometry, Julius-Maximilians-Universität Würzburg, Würzburg, Germany

[4]Institute for Medical Data Science, University Hospital Würzburg, Würzburg, Germany

[5]Clinical Trial Centre Würzburg, University Hospital Würzburg, Würzburg, Germany

[6]Uniklinik Erlangen Department of Psychiatry and Psychotherapy, Center for Health Services Research in Medicine, Friedrich-Alexander Universität Erlangen-Nürnberg, Erlangen, Germany

[*]these authors contributed equally

Corresponding Author:
Michael Zeiler, MSc
Medical Informatics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

Abstract

Background: Dementia is a widespread syndrome that currently affects more than 55 million people worldwide. Digital screening instruments are one way to increase diagnosis rates. Developing an app for older adults presents several challenges, both technical and social. In order to make the app user-friendly, feedback from potential future end users is crucial during this development process.

Objective: This study aimed to establish a user-centered design process for the development of digiDEM-SCREEN, a user-friendly app to support early identification of persons with slight symptoms of dementia.

Methods: This research used qualitative and quantitative methods and involved 3 key stakeholder groups: the digiDEM research team, the software development team, and the target user group (older adults ≥65 years with and without cognitive impairments). The development of the screening app was based on an already existing and scientifically analyzed screening test (Self-Administered Tasks Uncovering Risk of Neurodegeneration; SATURN). An initial prototype was developed based on the recommendations for mobile health apps and the teams' experiences. The prototype was tested in several iterations by various end users and continuously improved. The app's usability was evaluated using the System Usability Scale (SUS), and verbal feedback by the end users was obtained using the think-aloud method.

Results: The translation process during test development took linguistic and cultural aspects into account. The texts were also adapted to the German-speaking context. Additional instructions were developed and supplemented. The test was administered using different randomization options to minimize learning effects. digiDEM-SCREEN was developed as a tablet and smartphone app. In the first focus group discussion, the developers identified and corrected the most significant criticism in the next version. Based on the iterative improvement process, only minor issues needed to be addressed after the final focus group discussion. The SUS score increased with each version (score of 72.5 for V1 vs 82.4 for V2), while the verbal feedback from end users also improved.

Conclusions: The development of digiDEM-SCREEN serves as an excellent example of the importance of involving experts and potential end users in the design and development process of health apps. Close collaboration with end users leads to products that not only meet current standards but also address the actual needs and expectations of users. This is also a crucial step toward promoting broader adoption of such digital tools. This research highlights the significance of a user-centered design approach, allowing content, text, and design to be optimally tailored to the needs of the target audience. From these findings, it can be concluded that future projects in the field of health apps would also benefit from a similar approach.

(JMIR Hum Factors 2025;12:e65022) doi:10.2196/65022

XSL•FO
RenderX

KEYWORDS

dementia; usability; development; digiDEM; cognitive impairment; older adults; aging; mobile health; mHealth; design; feedback; screening; user centred; cognitive disorder; user-centered; mobile app

Introduction

Dementia is a widespread syndrome that currently affects over 55 million people worldwide, with annually almost 10 million new cases. The diagnosis of and treatment for people with dementia are going to be among the biggest challenges for health care systems worldwide [1]. A study by Eichler et al [2] found that 60% of people living with dementia in Germany had no formal diagnosis. Another problem lies in the long diagnostic periods. In the Bavarian Dementia Survey (BayDem) study, Wolff et al [3] found that the median time between the first perceived symptoms and diagnosis in Bavaria was 16 months. As Barth et al [4] were able to show, rural areas are also particularly affected here due to a high difficulty in accessing the facilities needed to diagnose and treat patients with dementia.

Screening instruments are one way of improving the diagnosis rate. A study with 146 participants has shown that diagnoses could be increased by almost 50% through upstream cognitive screening [2]. Internet-based screening tools offer the additional advantage that they can be used at a low threshold, regardless of time and place [5]. Digital technologies and the internet are already playing an increasingly central role in the everyday lives of older people. The proportion of people with internet access is growing across all age groups. In recent years, the number of German senior citizens (79 - 84 years) who regularly use the internet has more than doubled (18.8% in 2011 vs 39.4% in 2017). There is also an increased interest in health websites among older people [6]. In the 2021 report published by the German Federal Office for Information Security, it was stated that around 163,000 different health apps existed [7]. However, there is a lack of high-quality dementia apps. As analyzed in an earlier study, for only 6 of 20 identified dementia apps, scientific evaluation studies have been published. In none of those studies, the effectiveness of the respective screening app could be proven. Among the published app evaluations, screening apps received the worst overall quality rating. In summary, the analysis showed that the existing apps at this time did not provide reliable information and results [8].

Thus, in the digiDEM Bayern project (Digital Dementia Registry Bavaria), we have not only focused on the establishment of a digital registry for persons with mild cognitive impairment and mild to moderate dementia [9,10], but also on the development, scientific evaluation, and sustained provisioning of innovative eHealth tools and digital apps [11,12]. The goal of our current project, in this context, was to establish a user-centered design process for the development of digiDEM-SCREEN, a user-friendly app to support the early identification of persons with slight symptoms of dementia. The objective of this publication is to illustrate the iterative and agile user-centered development process consisting of 8 phases to move from a conceptual idea to an early prototype and a final prototypical implementation with continual involvement and feedback process from stakeholders and intended future users.

Methods

Overview

In order to achieve the goal of a user-friendly screening app for people with slight dementia symptoms, a user-centered iterative development approach comprising the following steps was chosen:

1. A systematic literature research of scientifically evaluated digital and nondigital dementia screening tests.
2. An early prototype (V1) development based on the guidelines for graphic design and textual formulation criteria for people with cognitive impairments. Graphical requirements include an easy-to-understand layout, standardized navigation elements, and a clear division of instructions into several steps [13]. In addition, textual guidelines such as short and concise sentences, logically structured sections with headings and an active approach to the user should be observed [14]. Furthermore, the expertise of 3 clinicians from different disciplines with long-term experience in dementia research and 2 professors of medical informatics with expertise in developing mobile health apps, supported by their teams was incorporated into the development.
3. Conduction of an initial evaluation of the early prototype (V1) based on a focus group discussion (FGD) with potential end users (older adults ≥65 years with and without subjective cognitive impairments) [15]. The group discussion was recorded and transcribed afterward. The results were then categorized and analyzed based on a previously published qualitative content analysis [16]. The following categories were extracted: general linguistic adaptations, task-related linguistic adaptations, menu navigation, general navigation, and specific design changes to individual components.

1. Incorporating the FGD feedback and results into the specification for the prototypical implementation of digiDEM-SCREEN (V2).
2. The second evaluation with a new group of potential users (older adults ≥65 years with and without subjective cognitive impairments) was based on the think-aloud method, where participants speak their thoughts and wishes aloud during the test and are observed by a researcher who also takes notes [17].
3. Incorporating the user feedback and think-aloud evaluation results into the enhanced specification for the subsequent digiDEM-SCREEN development step (V3).
4. Conduction of an additional focus group evaluation of the improved beta version of the app (V3) with people with migration background (nonnative German speakers).
5. Development and deployment of the first ready-to-use digiDEM-SCREEN version (V4).

Recruitment of the facilities for participation (steps 4, 5, and 7) was based upon the network of research partners in the project

XSL•FO
RenderX

digiDEM. The older adults from the facilities were informed about participation options in former group meetings (informed consent). After consenting, participants were invited to take part in the respective focus groups.

In steps 3 and 5, the System Usability Scale (SUS) has been calculated for the respective prototype versions by applying the German version of the standardized SUS questionnaire [18]. The scale can take values between 0 and 100; the higher the value, the higher the user-friendliness is categorized [19]. In addition, also in step 7, a self-assessment was used to determine technology use, interest, and expertise, each on a 5-point Likert scale (1 - 'Does not apply at all'; 5 - 'Applies completely'). The participants gave their subjective assessment and considered if they could use the app on their own (on a scale of 1 to 10) of the app and the specific components [20]. Furthermore, the participants were asked to name the most considerable problems associated with the app and if they wanted to change something.

Thus, our user-centered software design and development process included qualitative and quantitative evaluation methods at 3 different stages of the development process.

Ethical Considerations

This study was approved by the Ethics Committee of the Medical Faculty of the Friedrich-Alexander Universität Erlangen-Nürnberg (application number: 20-253_1-B; August 14, 2023). Written consent was obtained prior to the user testing and focus group discussion. All participants data were pseudonymized. The list of reidentifying data was stored separately from the analyzed data, and only authorized individuals have access to them. No one was paid to test the app.

Results

Steps 1 and 2: Development of the First Prototype (V1)

Prior to developing the screening test, we conducted systematic literature research. The search terms are shown in Multimedia Appendix 1. Criteria for the decision on a suitable screening tool were the scientifically examined psychometric properties (sensitivity and specificity), the availability of being a free to use app (not paid), and the technical feasibility of a tablet/smartphone. The main source of information was the systematic reviews of Chan et al [21] as well as the specific studies of the screening tools [22,23]. The decision on the Self-Administered Tasks Uncovering Risk of Neurodegeneration (SATURN) was based upon a group discussion about the aforementioned criteria as well as the (methodological) quality of the screening tools and the underlying scientific studies in general.

The SATURN [24] proved to be a test with particular promising diagnostic values (sensitivity: 0.92; specificity: 0.88 in dementia cases vs controls) [21]. The test is usable via a tablet. Administration time is about 10 minutes, which can be especially beneficial for older adults as shorter tests might induce less fatigue and therefore be more suitable for repeat administration compared with lengthier instruments [25]. Thus, the SATURN provides the foundation for the development and validation of a German adaption of the test usable as an app via smartphone and tablet.

To date, there is no German version of the SATURN test. The translation of the English version of the SATURN into German was carried out independently by 2 research assistants from the digiDEM Bayern project (MZ and ND) using the translate-retranslate method. Apart from some general adaptations, such as the correct assignment of the users' residence, linguistic aspects were also taken into account, and the texts were adapted to the German-speaking context. In some translations, the number of letters in the word increased noticeably (eg, farm - Bauernhof). A shorter related word (field - Feld) was then used in these places. Additional instructions were developed. The test adaptations aimed to ensure that both the implementation and the evaluation could be carried out entirely by the user or the system alone. At the start of the original SATURN test, the participant was asked to read aloud the task (close your eyes) and perform it [24]. Without a handler to check the action, there could be no subsequent evaluation (What phrase did you first read from this tablet?). Therefore, the researcher chose an alternative task (tap on the yellow circle) that also involved reading and performing an action.

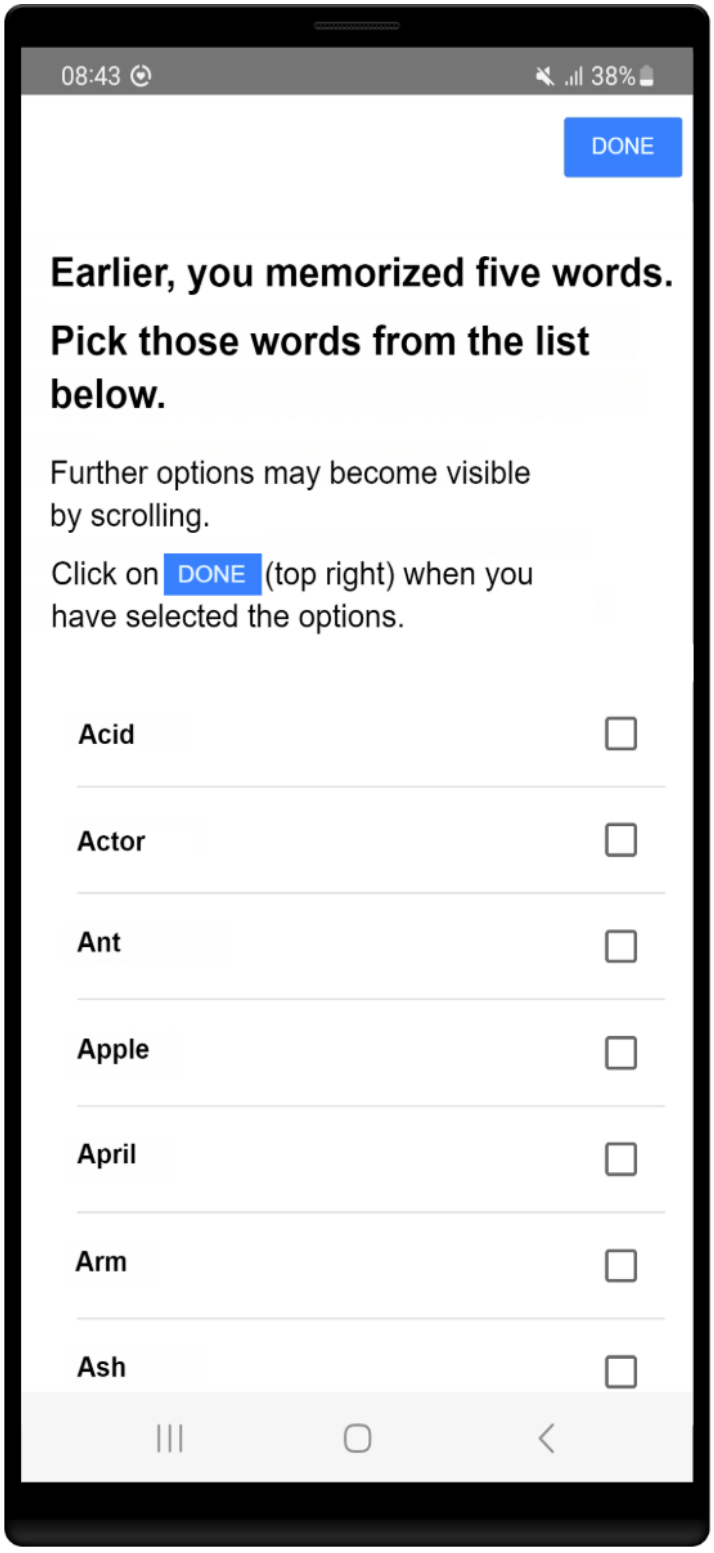
The final screening test consists of tasks from 6 different cognitive domains: Comprehension, Visuospatial, Orientation, Memory, Calculation, and Executive Function. Points are awarded for each task, which adds up to a maximum score of 30. The tasks must be completed without the help of other people. Participants may use their visual aids to complete the tasks; all other aids (eg, paper and pencil) are not permitted. A detailed description and illustrations of the individual test tasks can be found in Multimedia Appendix 1.

Another innovation is that the authors are developing the app as a tablet and a smartphone version. Due to the smaller display sizes, new components like the word selection task shown in Figure 1 had to be created. The researcher also developed some new logic to prevent larger adjustments, for example, that the user is only allowed to undo the last connection at the last task (tap on the circle with a blue background; Figure 2).

The following table (Table 1) shows the baseline characteristics of the participants in the usability analysis. This is followed by a description of the details of the individual events.

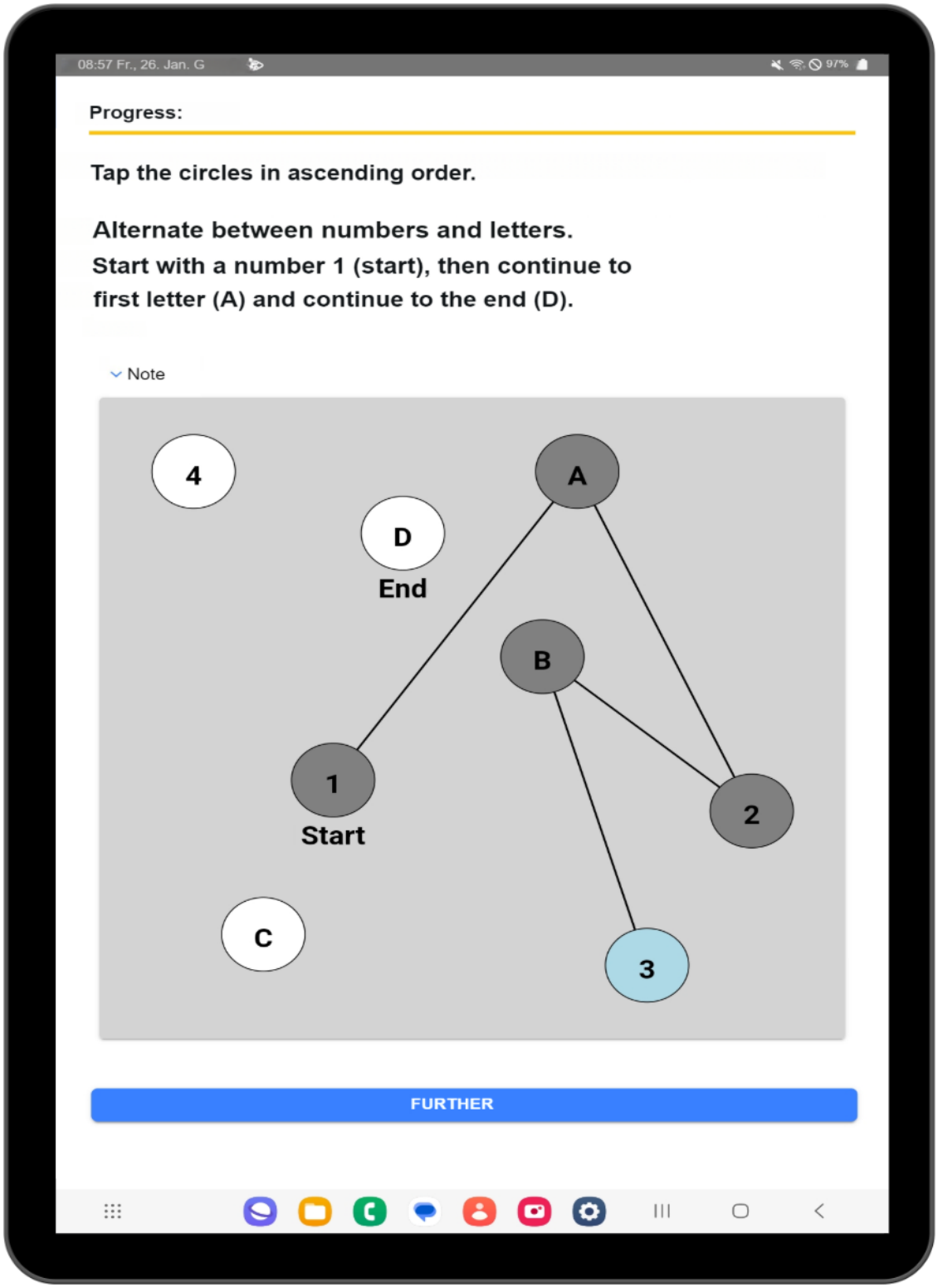
XSL•FO
RenderX

Figure 1. Word selection task as visible on a smartphone.



XSL•FO
RenderX

Figure 2. Trail Making Test as visible on a tablet.



XSL•FO
RenderX

Table . Baseline characteristics.

Study sample characteristics	FGD[a] 1 (Prototype V1)	Usability test (Prototype V2)	FGD 2 (Prototype V3)
Study population	13	21	7
Age (years), mean (range)	75.8 (66-84)	70.2 (65-80)	60.8 (53-69)
Sex, n (%)			
Male	1 (7.7)	8 (38.1)	0 (0.0)
Female	12 (92.3)	13 (61.9)	7 (100.0)
Education, n (%)			
Low	4 (30.8)	0 (0.0)	1 (14.3)
Medium	5 (38.4)	10 (47.6)	1 (14.3)
High	4 (30.8)	11 (52.4)	5 (71.4)
Self-perceived cognitive impairment, n (%)	5 (38.4)	5 (23.8)	0 (0.0)
Nonnative German speaker, n (%)	0 (0.0)	2 (9.5)	7 (100.0)
SUS[b] (0 - 100), mean (SD)	72.5 (1.6)	82.4 (16.1)	—[c]
App rating (1-10), mean (SD)	7.3 (2.1)	8.7 (1.5)	8.7 (1.2)
Independent app use (1-10), mean (SD)	7.5 (2.9)	8.95 (1.3)	8.5 (1.0)

[a]FGD: focus group discussion.

[b]SUS: System Usability Scale.

[c]Not applicable.

In both the first 2 focus groups, the SUS score was slightly lower in people with subjective cognitive impairment (FGD1: healthy older adults=74.4; people with subjective cognitive impairment=69.5; FGD2: healthy older adults=83.5; people with subjective cognitive impairment=82).

The SUS score decreased with advanced age (FDG1: ≥80 years old=63.8, 79 - 70 years old=73.9, ≤70 years old=85.0; FDG2:≥80 years old=75.0, 79 - 70 years old=77.5, ≤70 years old=85.0). People with a medium-level education had the best scores on the SUS (FDG1=85.0; FDG2=83.3), followed by people with a high-level education (FDG1=73.8; FDG2=83.0) and people with a low-level education (FDG1=55.6).

Steps 3 and 4: Focus Group Discussion V1

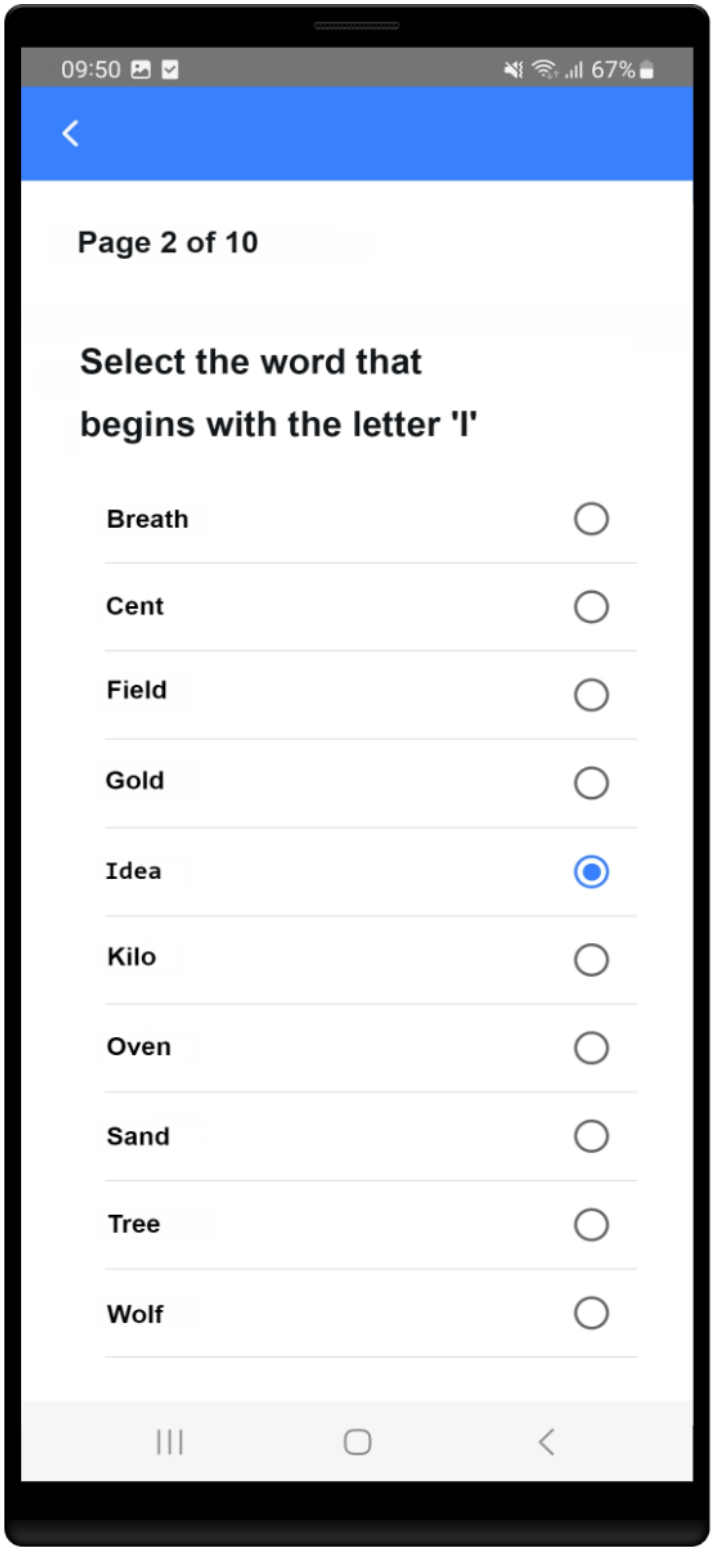
The first FGD took place as part of a memory training group. A memory training group is a frequent meeting of older adults, in which those adults perform different memory training exercises under the supervision of a group leader. Frequent excursions are also part of this service. The service is offered by a nonprofit organization (German: Wohlfahrtsverband) and is led by a research associate in the project digiDEM. The group consisted of a total of 13 participants. Their baseline characteristics are shown in Table 1. On average, they used modern technologies frequently (3.39) and showed an average interest in technological innovations (3.00). Their self-assessment of their competence in using modern technology was moderate (2.39), but the fear of failure played only a moderately important role (2.69).

Due to the large number of participants, 3 small groups were formed for the test. Participants were able to extensively test the app prototype and contact a research assistant with any questions. The individual components of the prototype achieved a subjective app rating of 7.3 (out of 10) points and an SUS score of 72.5. Participants also generally felt able (7.5 out of 10) to use the app independently without outside help. Subsequent group discussion of the results took place again in a large group.

The 2 most significant areas of improvement were observed in all the 3 small groups. Many participants recognized the letter I as T due to the inverted commas (‘I’) and had problems answering this task correctly. Participants also did not always recognize the selected word, as only the radio button on the right-hand side of the prototype showed their selection. The app prototype (V1) is shown in Figure 3. Participants therefore specifically requested that the entire line be colored when a selection was made. The 2 adjustments are shown in Figure 4.

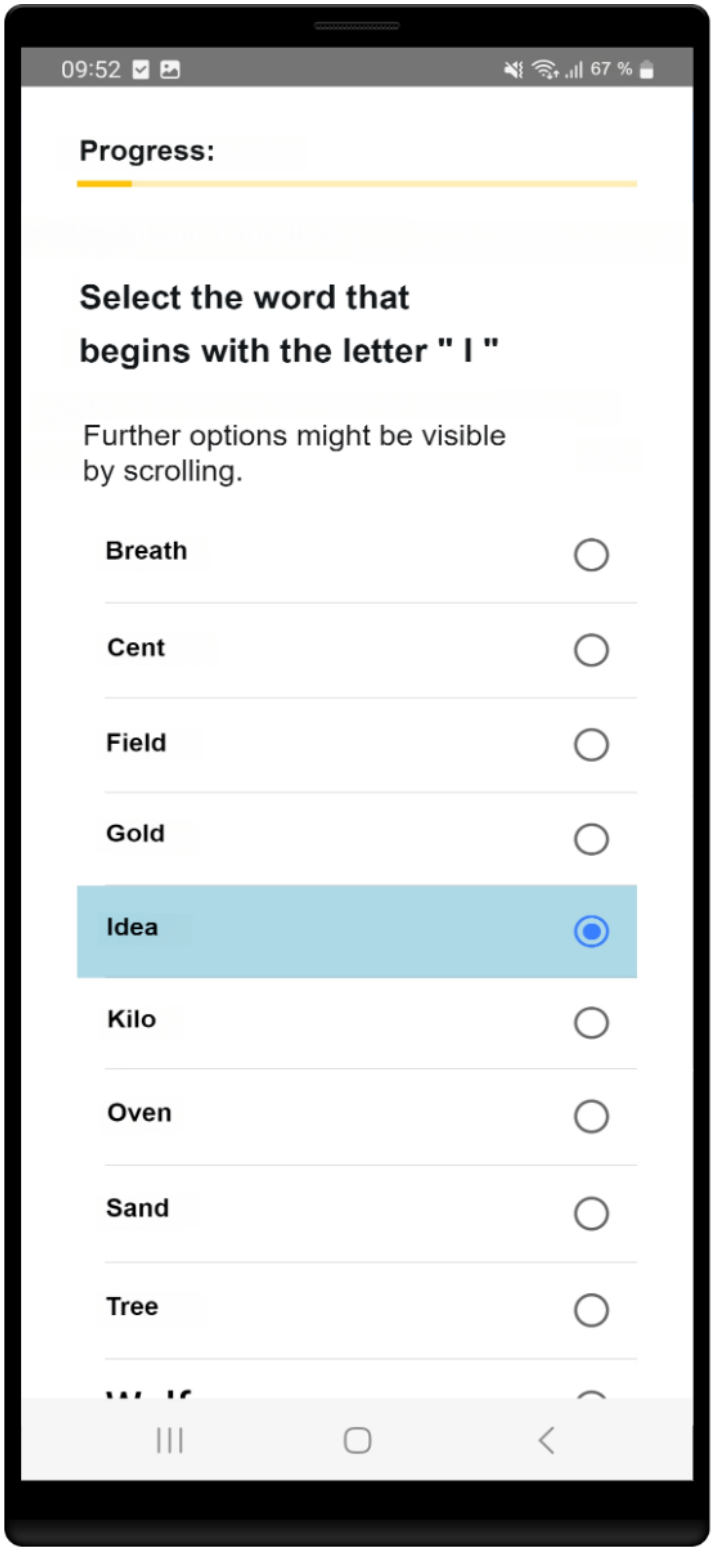
XSL•FO
RenderX

Figure 3. Illustration of the prototype (**V1**).



XSL•FO
RenderX

Figure 4. Final visualization (**V4**).



In addition, minor inconsistencies were noticed in this test, such as the fact that sometimes “Next” and sometimes “Done” were used to move on to the next task. The participants also wanted the selected images to be marked more clearly and the contrast and color intensity to be adjusted so that the colors could be recognized more clearly. One participant commented that she liked the “simple design” and that it did not distract from the actual content. Another participant mentioned that the instructions were too complex (“They were good instructions that you could actually understand. But I really had to read very carefully”). Therefore, the descriptive text has been simplified. Feedback on the user-friendliness of the FGD was predominantly positive. One participant particularly liked the fact that she could use the app without having much prior knowledge. The general consensus was that the app was easier to use on a tablet than on a smartphone due to the larger screen size. However, the smartphone version was also rated as usable by participants. These points mentioned were discussed with

XSL•FO
RenderX

the developers and incorporated into the second prototype accordingly.

Steps 5 and 6: Think-Aloud Usability Evaluation (V2)

A total of 21 older adults, who were randomly selected from participants in a dementia prevention event for older adults (≥65 years), participated in this usability test. Their baseline characteristics are shown in Table 1. They most frequently use modern technology (4) and are interested in technical innovations (3.62). They rated their competence in modern technology as average (3.33), while the fear of failure did not play a significant role (2.48).

The quantitative key figures collected increased compared with the first version. This prototype achieved a subjective app rating of 8.7 (out of 10) points and an SUS score of 82.4. The participants' assessment of using the app independently, without external help, also increased significantly (8.95 out of 10).

Based on the researchers' observations and the participants' statements minor adjustments and precisions, such as allowing €67 and €67.00 as the correct answer in the calculating task, were made. Some users also commented negatively about the last task's descriptive text. Due to the length and complexity of the content, the question was often not solved or only solved with a hint from the research assistant. Based on this feedback, the language of the text was revised again.

Step 7: Focus Group Discussion With Nonnative Speaker (V3)

The last user test took place under the aspect of accessible language and comprehensibility. To this end, an FGD was conducted with people with a migration background. Seven older adults took part in this FGD. Their baseline characteristics are shown in Table 1. They came from 4 different countries (Iraq, Kuwait, Sri Lanka, and Syria) and were all nonnative German speakers. They most frequently use modern technology (3.86) and are interested in technical innovations (3.57). They rated their competence in modern technology as average (3.14), while fear of failure played a minor role (2.58).

The quantitative indicators collected were similar to those of the German-speaking users. This prototype achieved a subjective app rating of 8.7 (out of 10) points. These participants also rated the success of using the app independently, without external help, at 8.5 (out of 10). Unfortunately, no SUS score could be obtained from this group due to the language barrier.

Two relevant changes emerged from the group discussion. First, a note on scrolling (Figure 4) was added in the appropriate places, and second, the language was adapted. A total of 5 of the 7 participants answered one of the initial questions (Select the fruit from the list.) incorrectly. The participants confused 2 words "Kirsche (cherry)" and "Kirche (church)," which are very similar in German. As this error does not indicate a possible cognitive decline, the word "Kirche (church)" was changed to "Kapelle (chapel)."

Step 8: Development and Deployment of the First Ready-to-Use digiDEM-SCREEN Version (V4)

In this step, the digiDEM-SCREEN test was finalized as a screening app for recording the current cognitive status of users. A validation study is currently underway. The test will be administered to patients in outpatient memory clinics and its sensitivity and specificity will be evaluated in the context of existing diagnoses and other nondigital cognitive tests. As part of the validation, cut-off values for categorizing current cognitive ability will also be determined as part of the validation. Depending on the test result, the user is given a short recommendation and options for action. If the result is above the threshold value calculated in the validation study, the screening does not indicate memory impairment. It is recommended that the test be repeated at regular intervals to monitor changes in memory performance. If the final result is below the threshold, further neuropsychological assessment in a memory outpatient clinic is recommended.

A study is currently underway to determine the sensitivity and specificity of the developed screening test (V4) and its correlation with the Montreal Cognitive Assessment [26]. For increased transparency, the research group has registered the project in the German Clinical Trials Register (DRKS) (registration number: DRKS00033764). After validation, the test will be available free of charge to anyone interested. Different randomization options have been used to minimize the learning effect. There are 5 versions of the test, which differ in the order of the numbers to be memorized. In addition, the position of each answer option is randomized for each test session. There are also plans to offer the screening test in different languages in the future. The possibility of multilingualism has already been taken into account in the programming of the app. This will be easy to implement once further translations have been validated.

Discussion

Principal Findings

There is a lack of evidence in the field of freely accessible apps for people with dementia, especially screening apps. A study published in 2023 showed that there are not any scientific studies to prove the effectiveness of any of the German-speaking screening apps [8].

A user-friendly app should have 3 main characteristics: typography appropriate for the target group (eg, recognizable icons), intuitive operation, such as fewer clicks to the desired action, and simplicity (eg, simple navigation) [27]. Three key stakeholder groups were involved in developing the digiDEM-SCREEN app: the digiDEM research team, the software development team, and the target user group (older adults ≥65 years with and without subjective cognitive impairments). Developing an app for older adults presents several challenges, both technical and social. Older adults may have less technology experience and difficulty understanding complex user interfaces [28]. Many older adults also have age-related limitations, such as visual or hearing impairments [29]. Cognitive abilities can decline with age, making complex

XSL•FO
RenderX

apps more difficult to use. The app should be tailored to the cognitive needs of older users, for example, by providing clear instructions and simple interactions. Considering these challenges when developing an app for older adults can help to create a user-friendly and accessible application that improves the lives of older people and promotes their independence [27]. In order to make the app as user-friendly as possible, feedback from potential future end users is essential. A critical examination of the study population shows that the participant structure is dominated by women in terms of gender. This could lead to a distortion of the results, as the findings may not be transferable to the entire target group. However, an empirical comparison of the usability of a mail app between male and female users showed that there are no statistically significant differences in the performance criteria of efficiency, effectiveness, and satisfaction between the 2 groups [30]. Another study examined whether there were systematic differences between women and men in the evaluation of the user experience of 3 websites and showed that there were no significant differences between the genders. Personal attitudes and preferences have a greater influence on the results [31].

The activities summarized under the term “patient and public involvement” enable patients to be actively involved in the planning and development of new products. International associations such as Alzheimer Europe as well as scientists are very interested in encouraging the active involvement of people with dementia in research for brainstorming and counseling [32].

Digital health apps often struggle with low adherence. One possible reason for this is users’ personal frustration with the content of the intervention, the way it is presented, and the nonintuitive handling [33]. The selected user-centered design could sustainably increase adherence. Users have unique knowledge, perspectives, and experiences that can influence a product’s quality, appropriateness, and user-friendliness. User testing is an essential part of the iterative development process and contributes to increasing the quality and success of the app [34]. Prototypes make it possible to recognize potential problems or weaknesses in user interaction or design at an early stage. By discovering these problems at an early stage, expensive changes or new developments in later phases of development can be avoided [35]. Thus, using the prototype design in the first workshop provided the team with a cost-effective way to get feedback and evaluate the idea. In the user tests, the tablet prototype of the app performed better than the smartphone version. The participants mainly criticized the smaller font and display size, which made it somewhat difficult to enter answers in some places. Despite these criticisms, the smartphone version was still rated as user-friendly. Smartphones are the most common mobile devices. A study by Weber et al [36] from 2020 found that an average of 41.4% of 71.6-year-old participants used their own smartphone. Due to the high availability of smartphones among seniors, the research team decided to stick with both versions. The mobile app also works offline, so no internet connection is necessary. Those decisions made it possible to reach a larger number of potential users. During the test phases, the participants did not use their devices, which they were familiar with in everyday life, but devices provided by the research team. For example, the display size, operating system (or at least the version), and individual settings may differ from their device. It is expected that user-friendliness will be even higher when users utilize their own smartphone or tablet.

The workshop participants were positive about the experience and gave constructive comments on the app. In addition, the SUS score also increased with each iteration of the app version. Bangor et al [37] described that products with a SUS score of 90 points and above were rated as exceptional, products with a SUS score of 80 points were rated as good, and products with a SUS score of 70 points were rated as acceptable. Anything below 70 points had usability issues that were a cause for concern. That means all SUS assessments are above the average and at least rated as acceptable. The prototype V1 gained a SUS score of 72.5 from the participants of the FGD. The rating of the second app version was notably better (SUS score: 82.4) and, therefore, rated as good. Due to some comprehension difficulties, unfortunately, no SUS score could be raised in the second FGD. Whenever questionnaires are used directly with people with dementia, the questions should be short and understandable (no technical terms), and double negatives should be avoided [27].

The results of the user evaluations showed that a user-friendly screening test for people with subjective cognitive impairments could be developed for the German-speaking population.

The main focus of the focus groups was on minimizing potential sources of error. Nevertheless, there is still a residual risk that the test cannot be carried out properly. If the first 3 simple test tasks are not answered correctly, an end screen appears with the message that the test cannot be carried out due to technical or language barriers. In this case, the user is advised to visit a specialized clinician. In the general instructions before the start of the test, participants are also informed that for example visual aids should be used.

Strengths

Different potential end users were included in the development process of the digital screening test as an app in order to improve usability and avoid technical or linguistic barriers. Moreover, additional languages and other extensions like a dementia prevention module can be added to the app.

Limitations

There was no random sampling of participants. Furthermore, although the SUS scale is the most frequently used scale to assess the user-friendliness of IT applications, it also has its weaknesses. The results of a systematic review show that some studies found that the double-negative questions from the SUS are challenging to understand for people with dementia [27]. With this knowledge in mind, we used simple language and no double negatives for the remaining questions we phrased. In the first FGD, the researchers had to explain individual questions to the group, and no SUS score could be collected in the last FGD. Most of the participants needed help understanding the questions. They were, therefore, unable to give reliable answers.

XSL•FO
RenderX

Conclusions

The development of digiDEM-SCREEN serves as an excellent example of the importance of involving experts and potential end users in the design and development process of health apps. From the initial stages of the project, experts were engaged in the content and design realization, providing a solid foundation for further development. The intensive testing phase, in which various end users tried out the app prototype in several iterations, clearly demonstrated the value of early and continuous feedback for improving the final product. This research highlights the significance of a user-centered design approach, allowing content, text, and design to be optimally tailored to the needs of the target audience. From these findings, it can be concluded that future projects in the field of health apps would also benefit from a similar approach. Research teams and app developers should integrate user-centered design practices into their development processes to ensure that the applications they create are not only functional but also user-friendly and appealing to the target audience. Such an approach could significantly enhance the acceptance and effectiveness of health apps, thereby making a valuable contribution to digital health care. A close collaboration with end users leads to products that not only meet current standards but also address the actual needs and expectations of users. This is a crucial step toward improving health technology and promoting broader adoption of such digital tools.

Acknowledgments

The present work was performed by MZ in partial fulfillment of the requirements for the degree “Dr. rer. biol. hum.” at the Faculty of Medicine of the Friedrich-Alexander-Universität Erlangen-Nürnberg. The project is funded by the Bavarian State Ministry of Health, Care and Prevention (funding code: G42d-G8300-2017/1606-83).

Authors' Contributions

MZ, ND, RP, PH, PKR, and HUP contributed to the conceptualization of the study. MZ and ND conducted the investigation, while MZ, ND, and HUP developed the methodology. Project administration was managed by RP, PH, EG, PKR, and HUP. The app programming was carried out by FH, JH, KK, and RP. Supervision was provided by MZ, ND, PKR, and HUP. MZ was responsible for writing the original draft, and MZ, NDI, EG, KK, and HUP contributed to the review and editing of the manuscript.

Conflicts of Interest

RP is a partner in Lenox UG, a company focused on applying scientific findings in digital health applications. Lenox UG holds shares in HealthStudyClub GmbH. This company is responsible for developing the app presented in the paper. RP received consultancy fees, reimbursements for conference attendance, and travel expenses in connection with the topics of mobile health and e-mental health. KK is also a shareholder in HealthStudyClub GmbH and has joined the company as chief technical developer. FH and JH work as freelancers for HealthStudyClub GmbH.

Multimedia Appendix 1
Search terms and detailed test description.
[DOCX File, 2717 KB - humanfactors_v12i1e65022_app1.docx]

References

1. Global status report on the public health response to dementia. World Health Organization. 2021. URL: https://www.who.int/publications/i/item/9789240033245 [accessed 2024-06-15]
2. Eichler T, Thyrian JR, Hertel J, et al. Rates of formal diagnosis of dementia in primary care: the effect of screening. Alzheimers Dement (Amst) 2015 Mar;1(1):87-93. [doi: 10.1016/j.dadm.2014.11.007] [Medline: 27239495]
3. Wolff F, Dietzel N, Karrer L, et al. Timely diagnosis of dementia: results of the Bavarian Dementia Survey (BayDem). Gesundheitswesen 2020 Jan;82(1):23-29. [doi: 10.1055/a-1031-9559] [Medline: 31863447]
4. Barth J, Nickel F, Kolominsky-Rabas PL. Diagnosis of cognitive decline and dementia in rural areas - a scoping review. Int J Geriatr Psychiatry 2018 Mar;33(3):459-474. [doi: 10.1002/gps.4841] [Medline: 29314221]
5. Labrique A, Vasudevan L, Mehl G, Rosskam E, Hyder AA. Digital health and health systems of the future. Glob Health Sci Pract 2018 Oct 10;6(Suppl 1):S1-S4. [doi: 10.9745/GHSP-D-18-00342] [Medline: 30305334]
6. Eighth report on older people and digitalisation - statement by the federal government (parliamentary paper). Bundesministerium für Familie Senioren, Frauen und Jugend (BMFSFJ). URL: https://www.bmfsfj.de/resource/blob/159916/9f488c2a406ccc42cb1a694944230c96/achter-altersbericht-bundestagsdrucksache-data.pdf [accessed 2024-06-15]
7. IT security in the digital consumer market: focus on health apps. Bundesamt für Sicherheit in der Informationstechnik. 2021. URL: https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/DVS-Berichte/gesundheitsapps.pdf?__blob=publicationFile&v=2 [accessed 2024-06-15]
8. Zeiler M, Chmelirsch C, Dietzel N, Kolominsky-Rabas PL. Scientific evidence and user quality in mobile health applications for people with cognitive impairments and their caregivers. Z Evid Fortbild Qual Gesundhwes 2023 Apr;177:10-17. [doi: 10.1016/j.zefq.2023.01.003] [Medline: 36890031]

XSL•FO
RenderX

9. Dietzel N, Kürten L, Karrer L, et al. Digital dementia registry Bavaria-digiDEM Bayern: study protocol for a multicentre, prospective, longitudinal register study. BMJ Open 2021 Feb 8;11(2):e043473. [doi: 10.1136/bmjopen-2020-043473] [Medline: 33558357]
10. Reichold M, Dietzel N, Chmelirsch C, Kolominsky-Rabas PL, Graessel E, Prokosch HU. Designing and implementing an IT architecture for a digital multicenter dementia registry: digiDEM Bayern. Appl Clin Inform 2021 May;12(3):551-563. [doi: 10.1055/s-0041-1731286] [Medline: 34134149]
11. Reichold M, Dietzel N, Karrer L, Graessel E, Kolominsky-Rabas PL, Prokosch HU. Stakeholder perspectives on the key components of a digital service platform supporting dementia - digiDEM Bayern. Stud Health Technol Inform 2020 Jun 23;271:224-231. [doi: 10.3233/SHTI200100] [Medline: 32578567]
12. Reichold M, Selau M, Graessel E, Kolominsky-Rabas PL, Prokosch HU. eHealth interventions for dementia - using WordPress plugins as a flexible dissemination for dementia service providers. Stud Health Technol Inform 2021 May 7;279:1-9. [doi: 10.3233/SHTI210081] [Medline: 33965911]
13. Web Content Accessibility Guidelines (WCAG) 2.1. World Wide Web Consortium. 2022. URL: https://outline-rocks.github.io/wcag/translations/WCAG21-de/#unterscheidbar [accessed 2024-10-18]
14. BMAS - easy language - a guide. Bundesministerium für Arbeit und Soziales. 2014. URL: https://www.bmas.de/DE/Service/Publikationen/Broschueren/a752-leichte-sprache-ratgeber.html [accessed 2024-10-18]
15. Data collection methods for program evaluation: focus groups. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/healthyyouth/evaluation/pdf/brief13.pdf [accessed 2024-06-15]
16. Mayring P, Fenzl T. Qualitative Inhaltsanalyse: Springer Fachmedien Wiesbaden; 2019.
17. Mey G, Mruck K. Handbuch Qualitative Forschung in Der Psychologie: Wiesbaden: VS Verlag für Sozialwissenschaften; 2010, Vol. 1.
18. Gao M, Kortum P, Oswald FL. Multi-language toolkit for the system usability scale. Int J Hum-Comput Interact 2020 Dec 13;36(20):1883-1901. [doi: 10.1080/10447318.2020.1801173]
19. Brooke J. SUS -- a quick and dirty usability scale. Usability Eval Ind 1996;189.
20. Neyer FJ, Felber J, Gebhardt C. Kurzskala technikbereitschaft (TB, technology commitment). In: Zusammenstellung Sozialwissenschaftlicher Items Und Skalen (ZIS): ZIS - GESIS Leibniz Institute for the Social Sciences; 2016. [doi: org/10.6102/zis244]
21. Chan JYC, Yau STY, Kwok TCY, Tsoi KKF. Diagnostic performance of digital cognitive tests for the identification of MCI and dementia: a systematic review. Ageing Res Rev 2021 Dec;72:101506. [doi: 10.1016/j.arr.2021.101506] [Medline: 34744026]
22. Koo BM, Vizer LM. Mobile technology for cognitive assessment of older adults: a scoping review. Innov Aging 2019 Jan;3(1):igy038. [doi: 10.1093/geroni/igy038] [Medline: 30619948]
23. García-Casal JA, Franco-Martín M, Perea-Bartolomé MV, et al. Electronic devices for cognitive impairment screening: a systematic literature review. Int J Technol Assess Health Care 2017 Jan;33(6):654-673. [doi: 10.1017/S0266462317000800] [Medline: 28920567]
24. Bissig D, Kaye J, Erten-Lyons D. Validation of SATURN, a free, electronic, self-administered cognitive screening test. Alzheimers Dement (N Y) 2020;6(1):e12116. [doi: 10.1002/trc2.12116] [Medline: 33392382]
25. Zygouris S, Tsolaki M. Computerized cognitive testing for older adults: a review. Am J Alzheimers Dis Other Demen 2015 Feb;30(1):13-28. [doi: 10.1177/1533317514522852] [Medline: 24526761]
26. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Apr;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]
27. Ye B, Chu CH, Bayat S, Babineau J, How TV, Mihailidis A. Researched apps used in dementia care for people living with dementia and their informal caregivers: systematic review on app features, security, and usability. J Med Internet Res 2023 Oct 12;25:e46188. [doi: 10.2196/46188] [Medline: 37824187]
28. Deutscher Ethikrat. Big Data and Health Report on the Public Consultation of the German Ethics Council: Deutschen Ethikrat; 2018. URL: https://www.ethikrat.org/fileadmin/Publikationen/Studien/befragung-big-data-und-gesundheit.pdf [accessed 2024-06-15]
29. Stevens G, Flaxman S, Brunskill E, Global Burden of Disease Hearing Loss Expert Group, et al. Global and regional hearing impairment prevalence: an analysis of 42 studies in 29 countries. Eur J Public Health 2013 Feb;23(1):146-152. [doi: 10.1093/eurpub/ckr176] [Medline: 22197756]
30. Weichbroth P. An empirical study on the impact of gender on mobile applications usability. IEEE Access 2022;10:119419-119436. [doi: 10.1109/ACCESS.2022.3219421]
31. Aufderhaar K, Schrepp M, Thomaschewski J. Do women and men perceive user experience differently? IJIMAI 2019;5(6):63. [doi: 10.9781/ijimai.2019.03.005]
32. Gove D, Diaz-Ponce A, Georges J, et al. Alzheimer Europe's position on involving people with dementia in research through PPI (patient and public involvement). Aging Ment Health 2018 Jun;22(6):723-729. [doi: 10.1080/13607863.2017.1317334] [Medline: 28513210]

XSL•FO
RenderX

33. Kernebeck S, Busse TS, Ehlers JP, Vollmar HC. Adherence to digital interventions in healthcare: definitions, methods and open questions. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2021 Oct;64(10):1278-1284. [doi: 10.1007/s00103-021-03415-9]
34. Beckert B, Bratan T, Friedewald M, et al. The New Role of Users in a Turbulent Innovation Environment: Karlsruhe:Fraunhofer ISI; 2021.
35. Bähr B. Prototyping Requirements. In: Prototyping of User Interfaces for Mobile Applications T-Labs Series in Telecommunication Services: Springer; 2017. [doi: 10.1007/978-3-319-53210-3_3]
36. Weber W, Reinhardt A, Rossmann C. Lifestyle segmentation to explain the online health information-seeking behavior of older adults: representative telephone survey. J Med Internet Res 2020 Jun 12;22(6):e15099. [doi: 10.2196/15099] [Medline: 32530433]
37. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009;4:114-123.

Abbreviations

BayDem: Bavarian Dementia Survey
FGD: focus group discussion
SATURN: Self-Administered Tasks Uncovering Risk of Neurodegeneration
SUS: System Usability Scale

Edited by A Kushniruk, D Gooch; submitted 02.08.24; peer-reviewed by M Bennasar, S Kröner; revised version received 06.11.24; accepted 03.12.24; published 22.01.25.

Please cite as:
Zeiler M, Dietzel N, Haug F, Haug J, Kammerer K, Pryss R, Heuschmann P, Graessel E, Kolominsky-Rabas PL, Prokosch HU
A User-Centered Design Approach for a Screening App for People With Cognitive Impairment (digiDEM-SCREEN): Development and Usability Study
JMIR Hum Factors 2025;12:e65022
URL: https://humanfactors.jmir.org/2025/1/e65022
doi:10.2196/65022

© Michael Zeiler, Nikolas Dietzel, Fabian Haug, Julian Haug, Klaus Kammerer, Rüdiger Pryss, Peter Heuschmann, Elmar Graessel, Peter L Kolominsky-Rabas, Hans-Ulrich Prokosch. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 22.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effectiveness of Using a Digital Wearable Plantar Pressure Device to Detect Muscle Fatigue: Within-Subject, Repeated Measures Experimental Design

Fu-Yu Chen, PhD; Tzu-Yao Lin; Yi-Cheng Huang, MSc; Evina Widianawati, MEd

Department of Biomedical Engineering, Chung Yuan Christian University, No. 200, Zhongbei Road, Zhongli District, Toayuan City, Taiwan

Corresponding Author:
Fu-Yu Chen, PhD
Department of Biomedical Engineering, Chung Yuan Christian University, No. 200, Zhongbei Road, Zhongli District, Toayuan City, Taiwan

Abstract

Background: Muscle fatigue, characterized by reduced force generation during repetitive contractions, impacts older adults doing daily activities and athletes during sports activities. While various sensors detect muscle fatigue via muscle activity, biochemical markers, and kinematic parameters, a real-time wearable solution with high usability remains limited. Plantar pressure monitoring detects muscle fatigue through foot loading changes, seamlessly integrating into footwear to improve the usability and compliance for home-based monitoring.

Objective: This study aimed to investigate the effects of muscle fatigue on plantar pressure measurements using a self-developed wearable plantar pressure system.

Methods: Twelve healthy participants completed a 5-minute calf muscle fatigue protocol. The plantar pressures and surface electromyography (sEMG) activity of the gastrocnemius muscles were recorded before and after exercise. The plantar pressures at 6 regions and the median frequency (MDF) of sEMG were analyzed to quantify fatigue.

Results: The self-developed foot pressure system showed a significant decrease in plantar pressure peak values at the heel of the left (P=.003) and right feet (P=.001) and at the lateral toe of the left (P=.001) and right feet (P=.026). A significant increase was observed at the metatarsal head of both the left foot (P=.001) and the right foot (P=.017). The MDF of sEMG signals significantly decreased in the left (P=.001) and right gastrocnemius (P<.001).

Conclusions: Plantar pressure changes and sEMG signals effectively detect gastrocnemius muscle fatigue using the proposed wearable system, supporting the development of a wearable solution for detecting muscle fatigue suitable for home-use.

(JMIR Hum Factors 2025;12:e65578) doi:10.2196/65578

KEYWORDS

muscle fatigue; plantar pressure sensors; wearable devices; home-based monitoring

Introduction

Muscle fatigue, characterized by the reduced ability to generate adequate force during repetitive contractions [1], affects performance across different groups and activities including daily activities of older adults as well as athletes during sports. Muscle fatigue disrupts the neuromuscular system and affects different aspects such as muscle strength, force production, and movement patterns, resulting in stiffness of the muscles and joints, impaired motor control, and poor balance [2], thereby increasing the risk of injuries. For instance, in badminton players, muscle fatigue impairs neuromuscular coordination, by reducing input from the foot sole receptors, and muscle force production, leading to weaker responses and decreased reaching distance performance [3]. Running-induced fatigue reduces lower limb muscle activity, decreases shock absorption capacity, and alters plantar pressure distribution [4], thus increasing the risk of running-related injuries such as stress fractures in the foot [5]. Mello et al highlighted how fatigue delays the gastrocnemius muscle's activation by 1 second relative to the center of pressure, thus impairing balance [6]. Furthermore, Morrison et al found that muscle fatigue in older adults (aged 60 - 79 years) significantly increased reaction times, postural sway, and fall risk compared to younger individuals (aged 30 - 59 years) [7].

Muscle fatigue can be detected through various methods, including blood lactate concentration [8], electromyography (EMG) [9], mechanomyography (MMG) [10], near-infrared spectroscopy (NIRS) [11], and kinematic parameters using inertial measurement units (IMUs) [12]. While the lactate concentration provides an estimation of global fatigue, it cannot monitor fatigue in real-time [8]. MMG, NIRS, and IMUs can be used in wearable forms; however, the use of NIRS and MMG has challenges owing to issues such as time lags, the use of

XSL•FO
RenderX

MMG has susceptibility to motion artifacts [13], and the use of IMUs needs further research regarding the relationship between kinematic parameters and muscle fatigue [14]. Surface electromyography (sEMG) is a widely used wearable method for real-time monitoring of local muscle fatigue by measuring myoelectric activity [15]. Fatigue reduces muscle fiber conduction velocity, shifting the sEMG power spectrum to lower frequencies (eg, decreased median frequency) and increasing signal amplitude. sEMG captures these changes and quantifies muscle fatigue using time-domain or spectral parameters [8]. However, the use of sEMG in a home setting is difficult, owing to the challenges of correctly attaching the electrodes and regularly wearing the device without assistance [16].

Adherence is a critical factor for effective daily monitoring, as meaningful and continuous data are required for accurate analysis [17]. Wearable devices integrated with mobile apps present promising healthcare solutions for home-based monitoring. The integration of hardware sensors and software mobile apps enables continuous and unobtrusive monitoring, thereby providing real-time data analysis that can support timely interventions [18]. A suitable alternative wearable approach for muscle fatigue monitoring is the use of plantar pressure sensors, which can be seamlessly integrated into a shoe's insole for daily comfortable wear. Wearable plantar pressure sensors in the foot insole are widely used for the detection of diabetic foot ulcers [19], gait analysis, and the measurement of the ground reaction force [20] and the center of pressure [6]. Muscle fatigue impacts body mechanics and alters foot loading patterns, measurable through plantar pressure at different foot regions. Considerable research has explored changes in plantar pressure following various physical activities linked to lower limb muscle fatigue, such as badminton [21], walking [22], and long-distance walking [23].

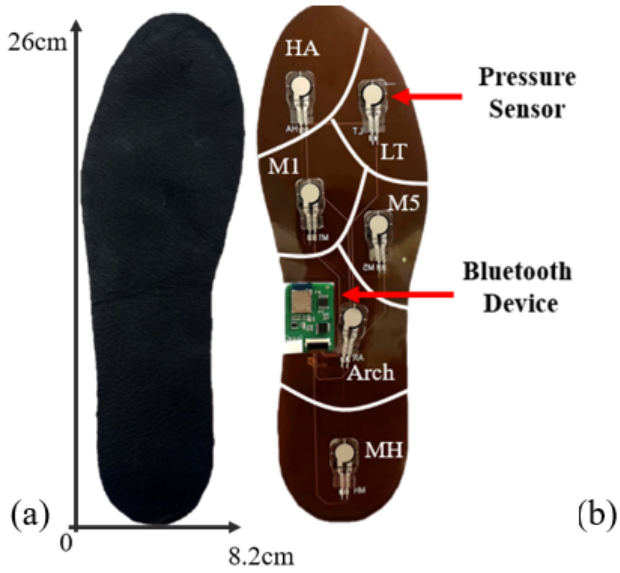
Despite the aforementioned advancements, wearable foot plantar systems for detecting muscle fatigue in everyday physical activities have yet to be designed. In this study, the effect of calf muscle fatigue in healthy participants was investigated using a self-developed wearable plantar pressure system that was suitable for daily use. The results can provide a muscle fatigue detection method for developing a wearable plantar pressure monitoring system

Methods

The Digital Foot Pressure System

A self-developed digital foot pressure system was utilized in this study. Figure 1 shows the wearable plantar devices for both the left and right feet as well as the measurements of the devices. Each device, designed to be worn comfortably with shoes, is equipped with 6 pressure sensors (A301 Flexiforce, Tekscan Inc.). The system also has a nRF52840 microprocessor (Nordic Semiconductor) used for the analog-to-digital converter measurement of foot plantar pressure, a wireless Bluetooth 5.0 module for data transmission, and a lithium-ion battery for power supply. The system operates with a sampling rate of 100 Hz, ensuring high-resolution data capture. The data from the 6 pressure sensors in both the left and right devices can be transmitted wirelessly in real time via Bluetooth to a custom-developed Android mobile app. This app facilitates real-time data display, analysis, and cloud storage, as shown in Figure 2. The pressure sensors are strategically arranged on the foot insole to cover 4 main plantar anatomical areas: the toes, metatarsals, arch, and heel. This general layout enables accurate measurement of natural gait; the layouts of the toes and metatarsals are further divided into 2 parts to capture plantar pressures along the mediolateral axis during walking [24].

Figure 1. Wearable plantar pressure device in an insole format. (**A**) Foot insole with 6 regions of pressure measurements: hallux (HA), lesser toe (LT), first metatarsal (M1), fifth metatarsal (M5), arch, and medial heel (MH); (**B**) Coordinates of the 6 pressure sensors on the insole.

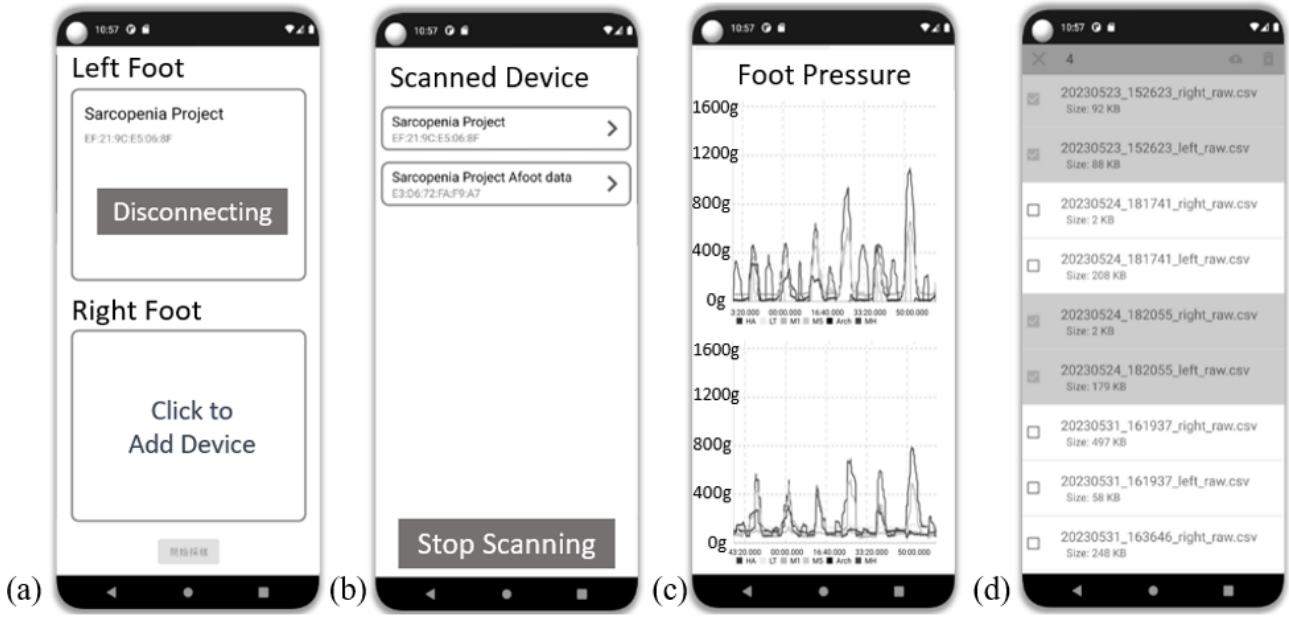


(b)

	X axis (cm)	Y axis (cm)
Hallux (HA)	5.8	21.3
Lesser Toes (LT)	2.4	21.1
First Metatarsal (M1)	6.2	16.2
Fifth Metatarsal (M5)	2.7	14.5
Arch	4.2	4.2
Medial Heel (MH)	4.7	3.9

XSL•FO
RenderX

Figure 2. Overview of the digital mobile app for the wearable plantar pressure device. (**A**) Bluetooth connection page for the devices on the right and left feet; (**B**) Bluetooth connectable device page; (**C**) real-time streaming graph showing plantar pressure data; and (**D**) Cloud storage for historical data.



Ethical Considerations

The research was approved by the Ethics Review Board of the Chung Yuan Christian University for Human Subjects Research (No. 8800-4-07-002). The participants were informed of the requirements and procedures of the entire experiment, and written consent was obtained before testing.

Subjects

A briefing session was conducted to familiarize the participants with the fatigue protocol and the walking test to be used in the research. Twelve healthy participants from the Chung Yuan Christian University, Taiwan, were recruited for the research. The mean (SD) age of the participants was 23.7 (2.6) years, and their mean (SD) weight was 62.4 (6.3) kg. The dominant limb of each participant was determined by identifying which foot would step out first once they started walking. The exclusion criteria included participants with plantar fasciitis or flatfoot and those who could not run independently for 5 minutes. Given that these foot disorders were excluded, experiments for minimizing variability could be performed. Then, suitable foot insole sizes were given to the participants.

Experimental Procedures

The experimental procedures were divided into 4 main steps. First, baseline tests were performed to measure the EMG of muscle activity and foot plantar pressure in both the legs. During the baseline test, the sEMG activities of the participants' gastrocnemius and rectus femoris muscles in the right and left legs were recorded using sEMG electrodes (T709, Comepa, France) and a data acquisition instrument (MP36, Biopac Inc., USA). The skin of each participant was cleaned with alcohol before the electrodes were attached. The sEMG muscle activities were recorded at a sampling rate of 2000 Hz. The plantar pressures at 6 locations in the left and right legs were measured simultaneously with sEMG. The plantar pressures were recorded using the wearable foot plantar device that was developed specifically for this study. Then, timed up-and-go and 10-meter walk tests were conducted; the participants were asked to rise from a standard chair, walk for 10 meters at their most comfortable speed, turn around, walk back. and sit down again [25].

After the baseline test, each participant was instructed to perform a fatigue exercise. All the participants were asked to place a weight-bearing apparatus on each leg and perform forefoot running for 5 minutes. The weight-bearing apparatus was 1/20th of a participant's weight [26]. sEMG data were collected from the gastrocnemius and rectus femoris muscles to capture the electrical activity associated with muscle contractions, and the wearable foot plantar pressure device was used to monitor changes in pressure distribution under the feet. This dual approach allowed the fatigue levels in both the muscles to be evaluated and quantified. After the exercise, the rating of perceived exertion (RPE) scale was used to assess and document each participant's perceived level of fatigue. Postfatigue measurements were recorded immediately after the fatigue protocol, while the participants performed timed up-and-go and 10-meter walk tests, and sEMG and foot plantar pressure measurements for both legs were collected.

EMG Muscle Fatigue Analysis

Three consecutive sEMG signals from the gastrocnemius and rectus femoris muscles were averaged for prefatigue and postfatigue analyses. The sEMG envelope was calculated from the raw sEMG signal through a 20 - 500–Hz finite impulse response bandpass filter. The signal was then processed via fast Fourier transformation to calculate the power spectrum median frequency (MDF) of the sEMG. A shift in the MDF to a lower frequency during exercise is considered localized muscle fatigue [27]. This method allowed researchers to determine whether muscle fatigue was induced in the gastrocnemius muscles, thereby increasing confidence in these data's correlation with foot plantar pressure data. The MDF is the frequency at which the EMG power spectrum is transmitted, and it is given by

XSL•FO
RenderX

(1)∑j=1MDF Pj=∑j=MDFM Pj=12∑j=1M pj,

where Pj is the power spectrum of the EMG at frequency bin *j*. The MDF splits the power spectrum of the EMG into two equal-amplitude regions.

The RPE Scale Analysis

The RPE scale was utilized to assess each participant's self-reported fatigue level. This subjective measurement tool is widely employed in the fields of physical activity, exercise, and sports to evaluate an individual's perceived effort and fatigue during physical tasks [28]. Participants rate their exertion based on sensations such as increased heart rate, breathing rate, muscle fatigue, and overall physical strain. The Borg RPE Scale, ranging from 6 to 20, is one of the most commonly used versions, with higher scores indicating greater levels of exertion [29].

Foot Plantar Pressure Analysis

Foot plantar pressure data, obtained using the developed digital foot plantar system and the commercial system, were simultaneously collected with the EMG data. The 6 plantar pressures in the left and right feet obtained in this research represent 6 anatomical areas of the foot: hallux (HA), lesser toe (LT), first metatarsal (M1), fifth metatarsal (M5), arch, and medial heel (MH). The peak pressures (Pa) in these 6 anatomical areas were analyzed and calculated from the raw plantar pressure data via MATLAB version R2022a (Mathworks Inc.). The measurements for the three consecutive prefatigue and postfatigue tests were averaged and compared. On the basis of the preliminary results, changes in the mean peak pressure of the 6 anatomical areas before and after the fatigue protocol were determined.

Statistical Analysis

In this study, the MDFs of the sEMG signals and the mean peak plantar pressures in each anatomical area for each participant before and after the fatigue protocol are presented as the means (standard errors). A paired *t* test with repeated measures was used to compare the MDFs of sEMG signals from the gastrocnemius and rectus femoris muscles before and after fatigue exercise. Then, a paired *t* test with repeated measures was used to detect significant differences in the 6 anatomical areas (ie, HA, LT, M1, M5, arch, and MH) before and after muscle fatigue exercise for the wearable plantar pressure system. All the statistical tests were performed with the significance level of $P<.05$.

Results

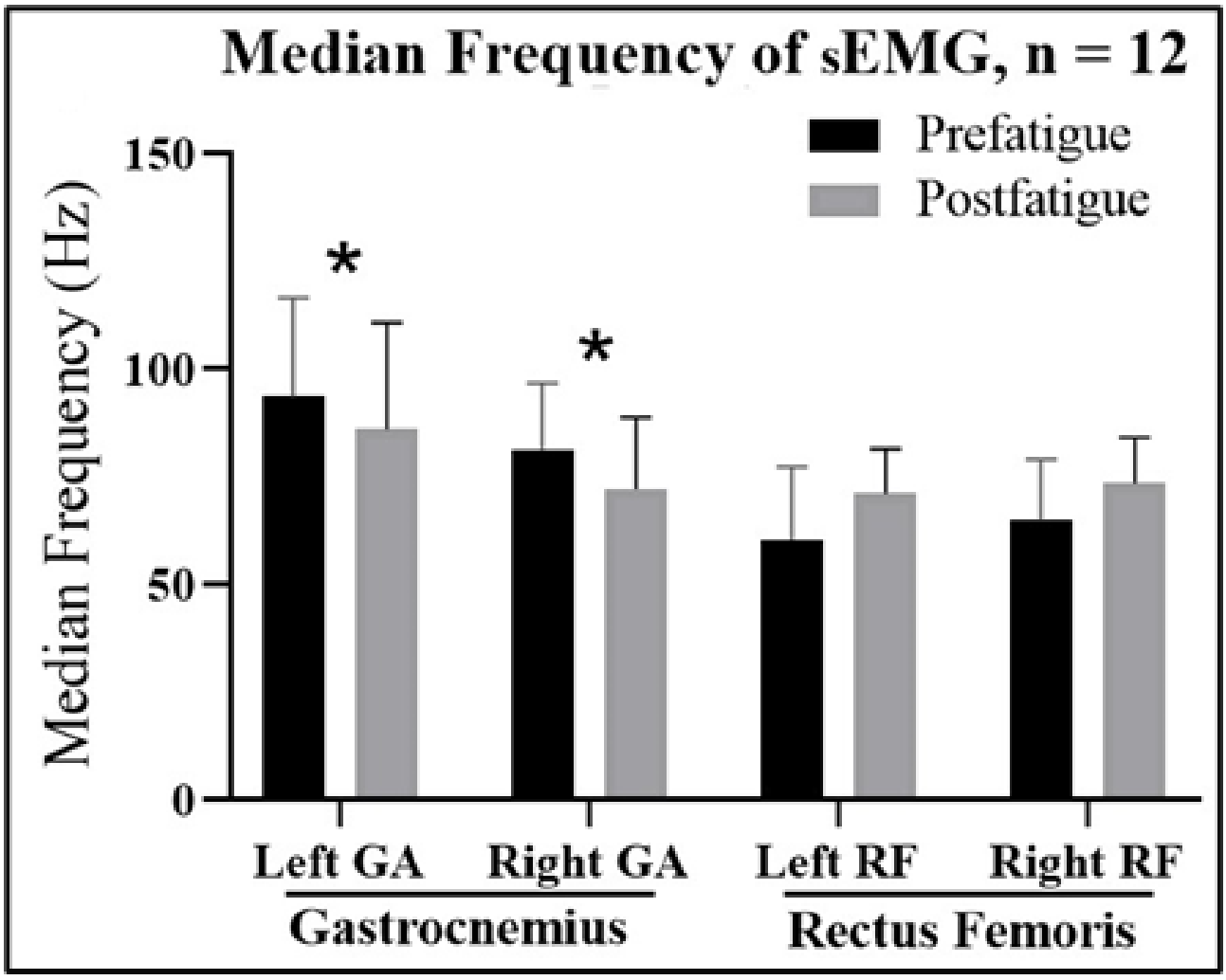
The data for muscle fatigue, RPE, and mean plantar pressure were collected and analyzed using a paired *t* test.

MDFs of sEMG in Detecting Muscle Fatigue

The MDFs of the sEMG data of the gastrocnemius and rectus femoris muscles are presented in Figure 3. The MDFs (SD) of the sEMG data of the left and right gastrocnemius muscles significantly decreased from 93.6 (22.8) Hz to 86.1 (24.6) Hz in the left foot and from 81.3 (15.5) Hz to 72.1 (16.6) Hz in the right foot after exercise for muscle fatigue ($P=.001$ and $P<.001$, respectively). However, the MDF of the sEMG before and after muscle fatigue exercise did not significantly differ between the left and right rectus femoris muscles.

XSL•FO
RenderX

Figure 3. Medium frequency of the sEMG for the gastrocnemius (GA) and rectus femoris (RF) muscles in the left and right legs before the fatigue exercise (prefatigue) and after the fatigue exercise (postfatigue). sEMG data of the left and right gastrocnemius muscles significantly decreased after the exercise for muscle fatigue (*P*=.001 and *P*<.001, respectively).



The RPE Scale in the Muscle Fatigue Protocol

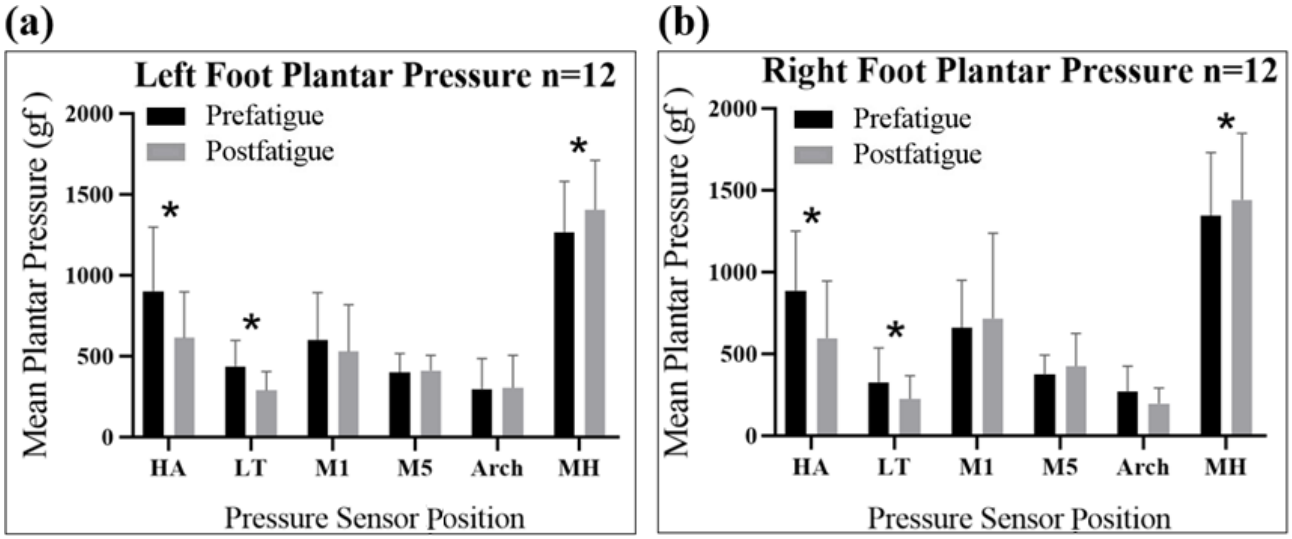
The RPE scale was used in the experiment to assess the muscle fatigue levels of the 12 participants. All the participants reported that the level of fatigue in the gastrocnemius muscle was greater than that in the rectus femoris. The analysis of the average RPE scores revealed that the mean fatigue level for the gastrocnemius muscle was 14.3 (SD 1.3), whereas the mean fatigue level for the rectus femoris was 7.0 (SD 1.3). The significant difference in fatigue levels (P <.001) between the two muscles aligns with the analysis of MDF in the sEMG data. These findings confirmed that the gastrocnemius muscle experienced greater fatigue than the rectus femoris did during the experiment.

Foot Plantar Pressure in Detecting Muscle Fatigue

The data for the left and right feet (Figure 4A and B, respectively) were obtained using a developed digital plantar pressure measurement system. Figure 4 shows the changes in plantar pressure before and after fatigue. The changes were noticeable in terms of the plantar pressure at the HA, LT, and MH on both feet after gastrocnemius muscle fatigue. The pressure values at other locations also changed, but they were not statistically significant.

XSL•FO
RenderX

Figure 4. Mean peak plantar pressure before the fatigue exercise (prefatigue) and after the fatigue exercise (postfatigue) by sensor positions: hallux (HA), lesser toe (LT), first metatarsal (M1), fifth metatarsal (M5), arch, and medial heel (MH) in the (**A**) left foot: a statistically significant decrease in the mean plantar pressure peak values was observed in the HA (*P*=.003) and LT (*P*=.001), while the MH showed a statistically significant increase (*P*=.001) after the fatigue exercise; and (**B**) right foot: a statistically significant decrease in the mean plantar pressure peak values was observed in the HA (*P*=.001) and LT (*P*=.026). A significant increase was noted in the MH position after fatigue (*P*=.017).



The mean plantar pressure values were measured using the self-developed foot pressure system. A significant decrease was observed in the mean plantar pressure peak values in the HA of the left foot (*P*=.003) and right foot (*P*=.001) and in the LT of the left foot (*P*=.001) and right foot (*P*=.026). A significant increase was observed in the mean plantar pressure peak value in the MH position after fatigue in both the left and right feet (*P*=.001 and *P*=.017, respectively). The mean plantar pressure peak value in the right HA decreased by 32.6%, from 883.2 (366.8) gf to 595.1 (349.2) gf, and that in the left HA decreased by 31.9%, from 900.7 (396.2) gf to 612.8 (284.0) gf. The mean plantar pressure peak value in the right LT decreased by 31.1%, from 325.8 (210.6) gf to 223.7 (141.4) gf, and that in the left LT decreased by 33.9%, from 435.6 (162.0) gf to 287.8 (115.8) gf. The mean plantar pressure peak value in the right MH increased by 7.2%, from 1343.4 (385.7) gf to 1440.3 (406.8) gf, and that in the left MH increased by 11.0%, from 1265.5 (313.0) gf to 1404.9 (304.4) gf. The changes in the mean peak pressure in the M1, M5, and arch positions were not significant. However, the M1 and arch positions showed a change in the mean peak pressure in both the left and right feet. In contrast, we observed an increase in the mean peak pressure in the M5 position in the left and right feet of 12.5 (24.5) gf and 52.0 (79) gf, respectively. The details of the changes in the mean peak plantar pressures before and after fatigue exercise by sensor position via the wearable plantar pressure system are shown in Multimedia Appendix 1.

Discussion

Principal Findings and Comparison With Previous Works

The results of this study demonstrated that muscle fatigue in the gastrocnemius significantly affects the plantar pressure in the HA, LT, and MH regions in both the legs. In particular, the mean peak plantar pressure in the HA and LT regions significantly decreased and that in the MH regions significantly increased after 5 minutes of muscle fatigue exercise in the gastrocnemius compared with the value before the fatigue exercise. This finding was determined using the self-developed wearable foot plantar pressure system proposed in this research. Muscle fatigue exercise in this study induced muscle fatigue solely in the calf area, which corresponded with a significant decrease in the MDF of the sEMG of the gastrocnemius muscle but not in the MDF of the sEMG of the rectus femoris. As presented, sEMG is highly muscle-specific and effective for detecting localized muscle fatigue; it is not useful to detect muscle fatigue in deeper muscles like the tibialis posterior in the calf area [5]. Pressure sensors offer an alternative by measuring the fatigue in muscles through shifts in loading patterns, providing a broader and more accessible assessment.

The findings of this study align with existing literature, showing that muscle fatigue in the lower limb induces a significant decrease in the mean pressure in the HA area and a shift in plantar pressure from the forefoot to the hindfoot [21,23,30,31]. This could be due to individuals starting to adapt their plantar pressure pattern from the forefoot to the hindfoot to avoid overuse of the forefoot [23]. The literature also revealed a significant decrease in the first metatarsal region. Similar findings were observed in this study, but the results in the first and fifth metatarsals and arch regions were not statistically significant. This discrepancy could stem from differences in fatigue protocols, as more intensive activities, such as 30-minute runs, lead to greater pronation and increased medial midfoot loading, whereas walking shows no such midfoot differences [32-35]. Additionally, subject variability, including anatomical factors such as arch height and leg-length differences, may contribute to variance in medial midfoot pressure [30,36]. For instance, flat feet increase loading on the medial longitudinal arch, while high arches shift the load to the lateral edge [37].

In this study, a shift in loading from the forefoot region including the HA and LT to the hindfoot including the MH after 5 minutes of fatigue exercise results in gastrocnemius muscle

XSL•FO
RenderX

fatigue. This finding could serve as an indicator for real-time monitoring of lower limb muscle fatigue, aiding in the prevention of injuries or falls in older adults and athletes. When a healthy person participates in sports such as badminton or running, muscle fatigue of the lower limb may be expected. These fatigued muscles can result in compromised reaction times, joint stability, and dynamic balance, thereby reducing impact absorption, which heightens the risk of injuries [21,38]. Athletes and coaches should monitor for fatigue alerts and ensure timely rest to recover muscle strength, preventing performance issues and injuries [39]. For older populations, particularly those at high risk of falls, such as patients with sarcopenia [40] or stroke [41], this system enables continuous home-based monitoring of plantar pressure patterns. A shift in plantar loading from the forefoot to the hindfoot could trigger alerts for rest to individuals or caretakers, minimizing fall risks associated with muscle fatigue.

Usability is a key factor in the adoption of digital health technologies such as the proposed wearable plantar pressure system [42]. Compared to sEMG, MMG, NIRS, or IMU wearable devices, its shoe-based form offers a user-centered design for daily wear, eliminating the need for users to remember and attach it correctly, thereby enhancing acceptance and compliance in home settings. A previous study highlighted the importance of face-to-face interactions with physicians in increasing their trust in digital health technologies [43]. Clinical or expert recommendations play a critical role in encouraging patients to adopt the device. The system's ability to support effective remote patient management allows both users and experts to review and monitor past performance or clinical conditions, thus fostering engagement [44]. These features, combined with its user-friendly design and portability, enhance the system's feasibility and scalability for home-based applications, promoting widespread adoption.

Limitations

This study has some limitations. First, direct evidence for the development of muscle fatigue in the gastrocnemius could not be provided by this study, given that muscle force was not directly measured. Instead, the MDF of the sEMG signal, which shifted to a low frequency, was used as an indicator of muscle fatigue; this approach is a common research method for measuring muscle fatigue [45-47]. The decrease in MDF was likely caused by changes in the properties of the muscle fibers, such as decreased conduction velocity and increased muscle fiber recruitment [48]. The MDF of the sEMG measurement might also be affected by factors other than fatigue, such as the muscle fiber type [49] and motor unit firing rate [50]. However, this limitation was compensated by the RPE scale questionnaire. Second, the effects of fatigue were measured only in the rectus femoris and calf gastrocnemius muscles; the effect of upper leg fatigue on plantar pressure was ruled out. sEMG measurements of other calf muscles, such as the tibialis anterior, can provide insights into a particular calf muscle's interaction with plantar pressure. The short test duration and limited wearing time (10 - 20 min) in this study may not accurately represent real-life scenarios, necessitating extended testing to better simulate prolonged activities. In addition, wearable devices must adapt to various environments and wearing conditions, as factors such as daily wear and tear, temperature, humidity, flooring, and uneven terrain can impact sensor performance.

Conclusions

The results of this study demonstrated a significant decrease in the mean plantar peak pressure in the HA and LT and an increase in the mean plantar peak pressure in the MH as an indicator of the onset of muscle fatigue in the gastrocnemius. This work is an exciting proof-of-concept outcome showing that muscle fatigue in the gastrocnemius can be detected via a wearable plantar pressure system. These findings can be used to further develop a wearable lower limb muscle fatigue monitoring system for minimizing injury risk in sports or during the daily activities of older adults.

Acknowledgments

The author would like to thank the participants for their time and willingness to take part in this study. This work was supported by the Taiwan National Science and Technology Council (Grant No. 110-2222-E-033-006-MY3).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Mean peak plantar pressures before and after fatigue exercise by sensor position via the wearable plantar pressure system system.
[DOCX File, 8 KB - humanfactors_v12i1e65578_app1.docx]

References

1. Finsterer J. Biomarkers of peripheral muscle fatigue during exercise. BMC Musculoskelet Disord 2012 Nov 8;13:218. [doi: 10.1186/1471-2474-13-218] [Medline: 23136874]
2. Pethick J, Tallent J. The neuromuscular fatigue-induced loss of muscle force control. Sports (Basel) 2022 Nov 21;10(11):184. [doi: 10.3390/sports10110184] [Medline: 36422953]

XSL•FO
RenderX

3. Sarshin A, Mohammadi S, Shahraba HBP, Sedigh M. The effects of functional fatique on dynamic postural control of badminton players. Biol of Exerc 2011;7(2):25-34. [doi: 10.4127/jbe.2011.0047]
4. Apte S, Prigent G, Stöggl T, et al. Biomechanical response of the lower extremity to running-induced acute fatigue: a systematic review. Front Physiol 2021;12:646042. [doi: 10.3389/fphys.2021.646042] [Medline: 34512370]
5. Weist R, Eils E, Rosenbaum D. The influence of muscle fatigue on electromyogram and plantar pressure patterns as an explanation for the incidence of metatarsal stress fractures. Am J Sports Med 2004 Dec;32(8):1893-1898. [doi: 10.1177/0363546504265191] [Medline: 15572318]
6. Mello RGT, Oliveira LF, Nadal J. Anticipation mechanism in body sway control and effect of muscle fatigue. J Electromyogr Kinesiol 2007 Dec;17(6):739-746. [doi: 10.1016/j.jelekin.2006.08.011] [Medline: 17142060]
7. Morrison S, Colberg SR, Parson HK, et al. Walking-induced fatigue leads to increased falls risk in older adults. J Am Med Dir Assoc 2016 May 1;17(5):402-409. [doi: 10.1016/j.jamda.2015.12.013] [Medline: 26825684]
8. Cifrek M, Medved V, Tonković S, Ostojić S. Surface EMG based muscle fatigue evaluation in biomechanics. Clin Biomech (Bristol) 2009 May;24(4):327-340. [doi: 10.1016/j.clinbiomech.2009.01.010] [Medline: 19285766]
9. Skrzat JM, Carp SJ, Dai T, et al. Use of surface electromyography to measure muscle fatigue in patients in an acute care hospital. Phys Ther 2020 Jun 23;100(6):897-906. [doi: 10.1093/ptj/pzaa035] [Medline: 32157308]
10. Naeem J, Hamzaid NA, Islam M, Azman AW, Bijak M. Mechanomyography-based muscle fatigue detection during electrically elicited cycling in patients with spinal cord injury. Med Biol Eng Comput 2019 Jun;57(6):1199-1211. [doi: 10.1007/s11517-019-01949-4] [Medline: 30687901]
11. Sheng X, Ding X, Guo W, Hua L, Wang M, Zhu X. Toward an integrated multi-modal sEMG/MMG/NIRS sensing system for human–machine interface robust to muscular fatigue. IEEE Sens J 2021;21(3):3702-3712. [doi: 10.1109/JSEN.2020.3023742]
12. Otálora S, Segatto MEV, Monteiro ME, Múnera M, Díaz CAR, Cifuentes CA. Data-driven approach for upper limb fatigue estimation based on wearable sensors. Sensors (Basel) 2023 Nov 20;23(22):9291. [doi: 10.3390/s23229291] [Medline: 38005677]
13. Li N, Zhou R, Krishna B, et al. Non-invasive techniques for muscle fatigue monitoring: a comprehensive survey. ACM Comput Surv 2024 Oct 31;56(9):1-40. [doi: 10.1145/3648679]
14. Dingwell JB, Joubert JE, Diefenthaeler F, Trinity JD. Changes in muscle activity and kinematics of highly trained cyclists during fatigue. IEEE Trans Biomed Eng 2008 Nov;55(11):2666-2674. [doi: 10.1109/TBME.2008.2001130] [Medline: 18990638]
15. Marco G, Alberto B, Taian V. Surface EMG and muscle fatigue: multi-channel approaches to the study of myoelectric manifestations of muscle fatigue. Physiol Meas 2017 May;38(5):R27-R60. [doi: 10.1088/1361-6579/aa60b9] [Medline: 28199218]
16. Arntz A, Weber F, Handgraaf M, et al. Technologies in home-based digital rehabilitation: scoping review. JMIR Rehabil Assist Technol 2023 Jul 27;10:e43615. [doi: 10.2196/43615] [Medline: 37253381]
17. Ferdinand KC, Yadav K, Nasser SA, et al. Disparities in hypertension and cardiovascular disease in blacks: The critical role of medication adherence. J Clin Hypertens (Greenwich) 2017 Oct;19(10):1015-1024. [doi: 10.1111/jch.13089] [Medline: 28856834]
18. Liao Y, Thompson C, Peterson S, Mandrola J, Beg MS. The future of wearable technologies and remote monitoring in health care. Am Soc Clin Oncol Educ Book 2019 Jan;39(39):115-121. [doi: 10.1200/EDBK_238919] [Medline: 31099626]
19. Korada H, Maiya A, Rao SK, Hande M. Effectiveness of customized insoles on maximum plantar pressure in diabetic foot syndrome: A systematic review. Diabetes Metab Syndr 2020;14(5):1093-1099. [doi: 10.1016/j.dsx.2020.06.041] [Medline: 32652497]
20. Ostaszewski M, Pauk J. Estimation of ground reaction forces and joint moments on the basis on plantar pressure insoles and wearable sensors for joint angle measurement. Technol Health Care 2018;26(S2):605-612. [doi: 10.3233/THC-182507] [Medline: 29843283]
21. Valldecabres R, Richards J, De Benito AM. The effect of match fatigue in elite badminton players using plantar pressure measurements and the implications to injury mechanisms. Sports Biomech 2022 Sep;21(8):940-957. [doi: 10.1080/14763141.2020.1712469] [Medline: 32301399]
22. Haris F, Liau BY, Jan YK, et al. A review of the plantar pressure distribution effects from insole materials and at different walking speeds. Appl Sci (Basel) 2021;11(24):11851. [doi: 10.3390/app112411851]
23. Stolwijk NM, Duysens J, Louwerens JWK, Keijsers NLW. Plantar pressure changes after long-distance walking. Med Sci Sports Exerc 2010 Dec;42(12):2264-2272. [doi: 10.1249/MSS.0b013e3181e305f4] [Medline: 20404764]
24. Wang D, Cai P, Mao Z. The configuration of plantar pressure sensing cells for wearable measurement of COP coordinates. Biomed Eng Online 2016 Oct 26;15(1):116. [doi: 10.1186/s12938-016-0237-3] [Medline: 27784299]
25. Cruz-Jentoft AJ, Baeyens JP, Bauer JM, et al. Sarcopenia: European consensus on definition and diagnosis: Report of the European Working Group on Sarcopenia in Older People. Age Ageing 2010 Jul;39(4):412-423. [doi: 10.1093/ageing/afq034] [Medline: 20392703]
26. Padua DA, Arnold BL, Perrin DH, Gansneder BM, Carcia CR, Granata KP. Fatigue, vertical leg stiffness, and stiffness control strategies in males and females. J Athl Train 2006;41(3):294-304. [Medline: 17043698]

XSL•FO
RenderX

27. Merletti R, Farina D. Surface Electromyography: Physiology, Engineering, and Applications: John Wiley & Sons, Inc; 2004.
28. Barnett F. The effect of exercise on affective and self-efficacy responses in older and younger women. J Phys Act Health 2013 Jan;10(1):97-105. [doi: 10.1123/jpah.10.1.97] [Medline: 22396245]
29. Berglund A, Orädd H. Exploring the psychological effects and physical exertion of using different movement interactions in casual exergames that promote active microbreaks: quasi-experimental study. JMIR Serious Games 2024 Aug 26;12:e55905. [doi: 10.2196/55905] [Medline: 39186776]
30. Bisiaux M, Moretto P. The effects of fatigue on plantar pressure distribution in walking. Gait Posture 2008 Nov;28(4):693-698. [doi: 10.1016/j.gaitpost.2008.05.009] [Medline: 18586495]
31. Nagel A, Fernholz F, Kibele C, Rosenbaum D. Long distance running increases plantar pressures beneath the metatarsal heads: a barefoot walking investigation of 200 marathon runners. Gait Posture 2008 Jan;27(1):152-155. [doi: 10.1016/j.gaitpost.2006.12.012] [Medline: 17276688]
32. Sun D, Fekete G, Mei Q, Gu Y. The effect of walking speed on the foot inter-segment kinematics, ground reaction forces and lower limb joint moments. PeerJ 2018;6:e5517. [doi: 10.7717/peerj.5517] [Medline: 30155372]
33. Honert EC, Bastas G, Zelik KE. Effect of toe joint stiffness and toe shape on walking biomechanics. Bioinspir Biomim 2018 Oct 10;13(6):066007. [doi: 10.1088/1748-3190/aadf46] [Medline: 30187893]
34. Rangra P, Santos D, Coda A, Jagadamma K. The influence of walking speed and heel height on peak plantar pressure in the forefoot of healthy adults: a pilot study. Clin Res Foot Ankle 2017;05(2). [doi: 10.4172/2329-910X.1000239]
35. Zhang LY, Liu QL, Yick KL, Yip J, Ng SP. Analysis of diabetic foot deformation and plantar pressure distribution of women at different walking speeds. Int J Environ Res Public Health 2023 Feb 19;20(4):3688. [doi: 10.3390/ijerph20043688] [Medline: 36834384]
36. Mootanah R, Song J, Lenhoff MW, et al. Foot type biomechanics part 2: are structure and anthropometrics related to function? Gait Posture 2013 Mar;37(3):452-456. [doi: 10.1016/j.gaitpost.2012.09.008] [Medline: 23107624]
37. Korpelainen R, Orava S, Karpakka J, Siira P, Hulkko A. Risk factors for recurrent stress fractures in athletes. Am J Sports Med 2001;29(3):304-310. [doi: 10.1177/03635465010290030901] [Medline: 11394600]
38. Bercovitz T, Herman A, Solomonow-Avnon D, Wolf A, Kodesh E. Plantar pressure modifications in experienced runners following an exhaustive run. Sports Biomech 2022 Nov;21(10):1189-1199. [doi: 10.1080/14763141.2020.1743744] [Medline: 32321365]
39. Takahashi K, Demura SI, Aoki H. Effects of lower limbs exercise with light and heavy loads on the center of gravity sway. AJSSM 2021 Jan 15;9(1):8-12. [doi: 10.12691/ajssm-9-1-3]
40. Yoo JI, Kim H, Ha YC, Kwon HB, Koo KH. Osteosarcopenia in patients with hip fracture is related with high mortality. J Korean Med Sci 2018 Jan 22;33(4):e27. [doi: 10.3346/jkms.2018.33.e27] [Medline: 29318794]
41. Rogers A, Morrison SC, Gorst T, et al. Repeatability of plantar pressure assessment during barefoot walking in people with stroke. J Foot Ankle Res 2020 Jun 29;13(1):39. [doi: 10.1186/s13047-020-00407-x] [Medline: 32600388]
42. Maqbool B, Herold S. Potential effectiveness and efficiency issues in usability evaluation within digital health: A systematic literature review. J Syst Softw 2024 Feb;208:111881. [doi: 10.1016/j.jss.2023.111881]
43. Cirkel L, Lechner F, Schlicker N, et al. Adoption and perception of prescribable digital health applications (DiGA) and the advancing digitalization among German internal medicine physicians: a cross-sectional survey study. BMC Health Serv Res 2024 Nov 6;24(1):1353. [doi: 10.1186/s12913-024-11807-1] [Medline: 39506735]
44. Rochester L, Mazzà C, Mueller A, et al. A roadmap to inform development, validation and approval of digital mobility outcomes: the Mobilise-D approach. Digit Biomark 2020;4(Suppl 1):13-27. [doi: 10.1159/000512513] [Medline: 33442578]
45. Eken MM, Braendvik SM, Bardal EM, Houdijk H, Dallmeijer AJ, Roeleveld K. Lower limb muscle fatigue during walking in children with cerebral palsy. Dev Med Child Neurol 2019 Feb;61(2):212-218. [doi: 10.1111/dmcn.14002] [Medline: 30156008]
46. Liu SH, Lin CB, Chen Y, Chen W, Huang TS, Hsu CY. An EMG patch for the real-time monitoring of muscle-fatigue conditions during exercise. Sensors (Basel) 2019 Jul 14;19(14):3108. [doi: 10.3390/s19143108] [Medline: 31337107]
47. Sarker P, Mirka GA. The effects of repetitive bouts of a fatiguing exertion (with breaks) on the slope of EMG measures of localized muscle fatigue. J Electromyogr Kinesiol 2020 Apr;51:102382. [doi: 10.1016/j.jelekin.2019.102382] [Medline: 31981942]
48. Kallenberg LAC, Schulte E, Disselhorst-Klug C, Hermens HJ. Myoelectric manifestations of fatigue at low contraction levels in subjects with and without chronic pain. J Electromyogr Kinesiol 2007 Jun;17(3):264-274. [doi: 10.1016/j.jelekin.2006.04.004] [Medline: 16790358]
49. Enoka RM, Duchateau J. Muscle fatigue: what, why and how it influences muscle function. J Physiol 2008 Jan 1;586(1):11-23. [doi: 10.1113/jphysiol.2007.139477] [Medline: 17702815]
50. Del Vecchio A, Negro F, Felici F, Farina D. Associations between motor unit action potential parameters and surface EMG features. J Appl Physiol (1985) 2017 Oct 1;123(4):835-843. [doi: 10.1152/japplphysiol.00482.2017] [Medline: 28751374]

XSL•FO
RenderX

Abbreviations

EMG: electromyography
HA: hallux
IMU: inertial measurement unit
LT: lesser toe
M1: first metatarsal
M5: fifth metatarsal
MDF: median frequency
MH: medial heel
MMG: mechanomyography
NIRS: near-infrared spectroscopy
RPE: rating of perceived exertion
sEMG: surface electromyography

Edited by A Kushniruk; submitted 20.08.24; peer-reviewed by A Panahi, H Sozen; revised version received 21.11.24; accepted 23.11.24; published 07.01.25.

Please cite as:
Chen FY, Lin TY, Huang YC, Widianawati E
Effectiveness of Using a Digital Wearable Plantar Pressure Device to Detect Muscle Fatigue: Within-Subject, Repeated Measures Experimental Design
JMIR Hum Factors 2025;12:e65578
URL: https://humanfactors.jmir.org/2025/1/e65578
doi:10.2196/65578

© Fu-Yu Chen, Tzu-Yao Lin, Yi-Cheng Huang, Evina Widianawati. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 7.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Usability Testing of a Bystander Bullying Intervention for Rural Middle Schools: Mixed Methods Study

Aida Midgett[1,2*], EdD; Diana M Doumas[1,2*], PhD; Claudia Peralta[1], PhD; Matt Peck[3], PhD; Blaine Reilly[1], PhD; Mary K Buller[4], MS

[1]Boise State University, Boise, ID, United States

[2]Institute for the Study of Behavioral Health and Addiction, Boise, ID, United States

[3]University of Arkansas, Fayetteville, AR, United States

[4]Klein and Buendel, Golden, CO, United States

[*]these authors contributed equally

Corresponding Author:
Aida Midgett, EdD
Boise State University
301 University Drive
Boise, ID, 83706
United States
Phone: 1 426 1214
Email: AidaMidgett@boisestate.edu

Abstract

Background: Targets of bullying are at high risk of negative socioemotional outcomes. Bullying programming in rural schools is important as bullying is more prevalent in those schools compared to urban schools. Comprehensive, school-wide bullying programs require resources that create significant barriers to implementation for rural schools. Because technology-based programs can reduce implementation barriers, the development of a technology-based program increases access to bullying prevention in rural settings.

Objective: We aimed to conduct usability testing of a bystander bullying intervention (STAC-T). We assessed usability and acceptability of the STAC-T application and differences in usability between school personnel and students. We were also interested in qualitative feedback about usability, program features, and feasibility.

Methods: A sample of 21 participants (n=10, 48% school personnel; n=11, 52% students) recruited from 2 rural middle schools in 2 states completed usability testing and a qualitative interview. We used descriptive statistics and 2-tailed independent-sample *t* tests to assess usability and program satisfaction. We used consensual qualitative research as a framework to extract themes about usefulness, relevance, needs, barriers, and feedback for intervention development.

Results: Usability testing indicated that the application was easy to use, acceptable, and feasible. School personnel (mean score 96.0, SD 3.9) and students (mean score 88.6, SD 9.5) rated the application well above the standard cutoff score for above-average usability (68.0). School personnel (mean score 6.10, SD 0.32) and students (mean score 6.09, SD 0.30) gave the application high user-friendliness ratings (0-7 scale; 7 indicates highest user-friendliness). All 10 school personnel stated they would recommend the program to others, and 90% (9/10) rated the program with 4 or 5 stars. Among students, 91% (10/11) stated they would recommend the program to others, and 100% (11/11) rated the program with 4 or 5 stars. There were no statistically significant differences in ratings between school personnel and students. Qualitative data revealed school personnel and students found the application useful, relevant, and appropriate while providing feedback about the importance of text narration and the need for teacher and parent training to accompany the student program. The data showed that school personnel and students found a tracker to report different types of bullying witnessed and strategies used to intervene by students a useful addition to STAC-T. School personnel reported perceiving the program to be practical and very likely to be adopted by schools, with time, cost, and accessibility being potential barriers. Overall, findings suggest that the STAC-T application has the potential to increase access to bullying prevention for students in rural communities.

Conclusions: The results demonstrate high usability and acceptability of STAC-T and provide support for implementing a full-scale randomized controlled trial to test the efficacy of the application.

XSL•FO
RenderX

(JMIR Hum Factors 2025;12:e67962) doi:10.2196/67962

KEYWORDS

technology-based bullying intervention; STAC-T; usability testing; middle school; rural

Introduction

Background

National statistics indicate that bullying is a national public health issue in the United States, with 19.2% of students aged 12 to 18 years reporting being bullied at school in the past year [1]. Bullying peaks in middle school, with 26.5% of sixth-grade students reporting being a target of school bullying, followed by 26.3% of seventh graders and 25.1% of eighth graders. Among students who report being bullied, 21.6% report being bullied online. Findings from a meta-analysis examining consequences of bullying have indicated that students who are targets of bullying report a wide range of negative mental health outcomes, including symptoms of anxiety, posttraumatic stress, depressive symptoms, nonsuicidal self-injury, suicidal ideation, and suicide attempts [2]. Similarly, being a target of cyberbullying is associated with internalizing symptoms, suicidal ideation [3-5], and alexithymia and psychotic experiences [6]. Thus, it is imperative to develop effective interventions for middle school students that are accessible and easy to implement to reduce bullying and the associated negative consequences.

Youth in Rural Schools

Students attending school in rural communities are at high risk of experiencing both school bullying and cyberbullying [7-9]. According to US national statistics, the highest rates of bullying among rural youth in the past decade were reported in 2019, with 27.7% or rural students reporting being bullied compared to 22.4% of students in urban areas [10]. Although rates of bullying peaked in 2019 for both rural and urban students, the most recent US national statistics indicate that the prevalence of school bullying victimization continues to be higher among students in rural areas (23.8%) than among students in urban areas (19%) [1]. Furthermore, among targets of bullying, students attending rural schools are also more likely to report being bullied online (23%) compared to students attending urban schools (19.5%). Rural students also report a higher rate of being bullied with repetition (18.8%) compared to urban students (14.4%). Among middle school students attending schools in rural communities, bullying victimization is associated with poor school relationships, negative school experiences [11], and depression and anxiety [11,12]. These data suggest the importance of developing school-based bullying prevention programs specifically for students in rural communities.

School-Based Bullying Interventions

Comprehensive, school-wide interventions are effective in reducing bullying and the associated negative mental health outcomes [13]. Furthermore, bystander training (eg, teaching students who witness bullying to intervene in bullying situations) is an important intervention component [13]. Although up to 80% of students report witnessing bullying [14], only 20% intervene [15]. Because students report that they do not know how to intervene when they witness bullying [16], bystander training is a promising approach to bullying prevention. However, few comprehensive school-based programs incorporate bystander training. In addition, comprehensive, school-wide bullying prevention programs are expensive, complex, and time intensive and require extensive training [17]. Because these interventions require substantial resources, many schools face implementation barriers. Schools in rural communities may also face economic disparities, creating further implementation challenges [18], including a lower tax base, increased training costs due to bringing in expert trainers, frequent staff turnover, school closures, staff overload, and lack of program advocates in bullying prevention [19]. Challenges related to logistical problems, training requirements, and limited funding can negatively impact program adoption and sustainability [19].

Shifting to a Technology-Based Bullying Intervention

Technology-based interventions have the potential to improve access to programming and decrease implementation barriers experienced in rural communities [19]. Although some rural areas have higher rates of poor internet connectivity, eligible schools in rural communities can receive discounts for internet and broadband services [20]. Federal grants to build broadband infrastructure in rural areas are also available [21]. In addition, research conducted with key middle school personnel (ie, administrators, teachers, and school counselors) in rural communities indicates both a strong interest in technology-based bullying prevention programs and positive implementation conditions (eg, administrative support and technology readiness) [22]. Thus, most students in rural communities have access to the necessary infrastructure to support technology-based programs, and key personnel in rural middle schools indicate that schools are interested and ready for technology-based bullying interventions.

Although there is strong interest and need for online bullying prevention programming, very few US bullying prevention programs include a technology component. Programs that do offer it as an adjunct to in-person delivery [13]; often rely on simple texting, not multimedia interfaces; and they do not train student bystanders to intervene. For example, the Build Respect, Stop Bullying program for middle schools uses an online platform [23] but is part of a large program with staff or family components without bystander training. Other available technology programs include (1) an SMS text messaging program pairing youth with a “text buddy” [24]; (2) apps that encourage students to report cyberbullying, block websites attracting cyberbullying, and notify parents and school personnel of cyberbullying; and (3) online social media campaigns and educational resources (eg, videos, testimonies, and quizzes) [18,25,26]. Although there are programs that incorporate technology into bullying prevention and intervention, none appear to offer a route to a technology-based, interactive bystander training for middle school students.

XSL•FO
RenderX

The STAC Intervention

STAC [27] is a brief, stand-alone bystander intervention that includes didactic and experiential training followed by 2 booster sessions. The 75-minute didactic training includes education about bullying and cyberbullying, the consequences of bullying, and bystander roles and a description of the four STAC strategies: (1) "Stealing the show"—using humor or distraction to interrupt the bullying situation removing the attention away from the target, (2) "Turning it over"—informing an adult about the bullying and asking for help, (3) "Accompanying others"—befriending or providing support to the targeted student, and (4) "Coaching compassion"—gently confronting the perpetrator to increase empathy for the target. The experiential training comprises a series of role-plays during which students practice using the STAC strategies through bullying scenarios. The STAC training is followed by two 15-minute booster sessions to reinforce learning. The STAC intervention is effective in reducing bullying [28,29] as well as mental health risks for bystanders [30-34]. STAC has also been adapted to be culturally appropriate for middle school students in rural communities [35-37]. Research on the adapted STAC intervention demonstrates both bullying reduction [35,38] and improved mental health [35,39] among students trained in the program.

The Technology-Based STAC Intervention

The technology-based STAC intervention (STAC-T) is an online application developed to shift intervention delivery from in-person implementation to a technology-based format, thereby increasing accessibility and reducing barriers to intervention implementation. STAC-T is designed to be easily disseminated to large groups of students, who can access the intervention from a computer, tablet, or smartphone. In addition, the 40-minute STAC-T application is designed to be modular, increasing implementation flexibility. The initial training is followed by one 15-minute booster session designed to reinforce skill acquisition through virtual role-plays. The program is interactive, including knowledge checks, personalized feedback, and the selection of avatars to respond to bullying scenarios. Initial development included the design and testing of a STAC-T prototype. The design of STAC-T was developed based on the content of the in-person STAC intervention for rural middle schools as well as feedback from an expert advisory board and key middle school personnel in rural communities. In addition, students attending rural middle schools participated in 3 iterative focus groups, providing feedback on program usefulness, content, and functionality [40]. Once developed, the STAC-T prototype was evaluated through usability testing, which provided feedback from end users on program functioning [41]. The results from usability testing with key personnel and students from 2 rural middle schools indicated that the STAC-T prototype was easy to use, acceptable, and feasible, supporting the full-scale development of the STAC-T application [40].

This Study

Bullying is a significant public health concern for students attending rural schools [7-9]. Comprehensive bullying intervention programs that incorporate bystander interventions are the standard for practice [13]; however, they place a high demand on schools for implementation [17] and can contribute to disparities in rural schools [19-22]. STAC-T has the potential to reduce barriers and increase access to bullying prevention for middle school students in rural settings [40]. Therefore, the purpose of this study was to evaluate the usability and acceptability of the full-scale STAC-T application to determine readiness for a large, multisite randomized controlled trial to evaluate the efficacy of the STAC-T application for middle school students in rural communities. Usability testing is an important step in the process of intervention development as it predicts the likelihood of program adoption [42]. To achieve this aim, we implemented usability testing with key stakeholders (ie, school personnel and students) at 2 middle schools in rural communities in 2 states (N=21) using a mixed methods design. This study had the following objectives: (1) to assess usability and acceptability of the STAC-T application and (2) to assess differences in usability between school personnel and students.

Methods

Participants

Participants were key school personnel (ie, administrators, teachers, and school counselors; 10/21, 48%) and students (11/21, 52%) recruited from 2 middle schools in rural, low-income communities in the Northwestern and Southern regions of the United States. Between 5 and 10 usability testers are needed to identify most usability issues [43]. The schools were selected based on previous and ongoing research partnerships. The 2 schools were Title 1 schools, with 95% and 99% of the student population at the 2 schools eligible for reduced or free lunch. Among school personnel, the ages ranged from 26 to 55 years (mean 43.4, SD 9.8 years), and most (9/10, 90%) were female. School personnel self-reported ethnicity or racial background as White (5/10, 50%), Hispanic or Latino (3/10, 30%), and Black or African American (2/10, 20%). Among students, ages ranged from 11 to 15 years (mean 12.8, SD 1.3 years), with 36% (4/11) in grade 6, a total of 18% (2/11) in grade 7, and 45% (5/11) in grade 8. Students self-reported gender as female (6/11, 55%) and male (5/11, 45%). Students self-reported ethnicity or racial background as White (4/11, 36%), Black or African American (4/11, 36%), and Hispanic or Latino (3/11, 27%).

Development of the STAC-T Application

The translation from the STAC in-person intervention to STAC-T was guided by persuasive system design, a theoretical guide for translating clinical aims to health-related technology frameworks [44-46]. The STAC-T application was developed using Agile programming, a collaborative and incremental programming methodology [44-46]. The application was functional on all web browsers that support HTML5 and was built on a full-stack web application using HTML and JavaScript as the main interface. React.js was used as the front-end framework. The look and feel of the program were designed using Adobe Illustrator and Photoshop and developed using HTML elements plus SVG, PNG, JPG, WAV, MP4, and GIF images, audio, and video graphics. The system is accessible on desktop computers and iOS and Android tablets and smartphones. All design and programming elements were

XSL•FO
RenderX

aligned, and stakeholders' inputs were incorporated throughout the multistage development. Programmers produced the STAC-T application; alpha and beta tested it in-house for stability and code errors; tested it for usability; and revised it following an iterative, Agile production process.

In previous studies, as well as the in the iterative interviews and focus group conducted in this study, participants indicated that increasing program interactivity, adding more color, and including more realistic images such as avatars were important to increase engagement [40] and promote behavior change [47]. Therefore, design elements such as space (bright colors and visual space), components (realistic characters and familiar objects), and mechanics (actions reported by students that occur in rural middle schools) were established for the program features. In addition, STAC strategy practice was designed to require students to select an avatar, view bullying scenarios and select actions to operationalize the STAC strategy, view the avatar enacting the selected action, and receive feedback on its effectiveness. An artist hand illustrated and styled 6 avatars. The avatars had light-, medium-, and dark-colored hair in different styles as well as light, medium, and dark skin tones for students to choose from to best represent themselves and stimulate engagement. To reward learning and bolster adherence, "badges" (visual reward icons; eg, "Show Stealing Badge") were included as intermittent awards to encourage user engagement (Figure 1).

The STAC-T application content comprises three modules: (1) What is Bullying?—users are presented with background information on bullying, including bullying definitions (ie, physical, verbal, and relationship bullying as well as cyberbullying), bullying facts and statistics, characteristics of students who bully, and negative consequences of bullying; (2) What are Bystanders?—users are taught what a bystander is and how bystanders affect bullying outcomes (this module explains the 4 bystander roles: *assistants*, those who intentionally help the bully; *reinforcers*, those who are not directly involved in hurting another student but encourage the bully by standing around, laughing, or watching quietly; *outsiders*, those who do not take sides while witnessing bullying; and *defenders*, those who do something to stop the bullying situation or help the target in some way); and (3) STAC Strategies—users are introduced to the 4 STAC strategies, which are "Stealing the show," "Turning it over," "Accompanying others," and "Coaching compassion" (this module also includes STAC strategy practice using avatars selected by the user). The booster session includes additional practice with bullying scenarios and STAC strategy use.

Iterative interviews (15/21, 71% of the participants; 7/15, 47% female and 8/15, 53% male; 6/15, 40% White; 6/15, 40% Black or African American; 3/15, 20% Hispanic or Latino) and 2 rounds of iterative focus groups (20/21, 95%; 11/20, 55% female and 9/20, 45% male; 6/20, 30% White; 9/20, 45% Black or African American; 3/20, 15% Hispanic or Latino; 2/20, 10% other) conducted with middle school students attending schools in rural communities in 2 states informed program development before usability testing. Students participating in the interviews provided feedback on design aspects of the program, including color scheme, narration, and cartoons. Students were given a sample slide in 5 color schemes, 3 narrator voice samples, and 3 cartoon-style character depictions. Students were asked to rank a series of questions about each program aspect and then rank their preferences. Feedback and ranked choices were used to select color schemes, the narrator, and the program artist. Iterative focus groups were then conducted to gather feedback from students related to content and stylistic aspects of the program. Overall, the program was well received; students reported that the content was helpful and they liked the look and feel of the teacher who appears throughout the training. Students in the first round of focus groups provided specific feedback to incorporate more cyberbullying scenarios (eg, having bullies use their phones to record their peers without their knowledge), make the appearance of the characters more realistic (eg, changing clothing, adding eyes to all the characters, and changing hairstyles), and improve the function of the program to make navigation more user-friendly. Students also expressed disliking a particular activity, which was removed from the program. Students' feedback was incorporated into the program before conducting the second round of focus groups. The students in the second round of focus groups provided additional feedback about how to make the appearance of the characters more realistic (eg, adding emotion to the characters), as well as adding background images to make the scenarios look more like what they are used to seeing at school and school-related activities, such as sporting events (eg, adding teachers, lockers, wall hangings, and bulletin boards). They also provided specific feedback about how to make student behaviors more realistic (eg, having the target look sad instead of crying and changing the type of bullying from physical to verbal in front of adults). Input from the focus groups informed the development of the final STAC-T application used in this study.

XSL•FO
RenderX

Figure 1. Samples from the technology-based STAC application.



Procedures

Participant recruitment and usability testing occurred in 2024. The researchers provided the liaisons (eg, school counselor and principal) from each school with an email script describing the purpose and procedures of the study. Inclusion criteria for school personnel consisted of being employed as a principal, teacher, or school counselor at the partnering school; speaking English or Spanish; consenting to participate; and having a desire to make a positive difference in the school climate. For students, inclusion criteria consisted of being enrolled in the partnering school, speaking English or Spanish, having parental consent and student assent, and having a desire to make a positive difference in the school climate. To assess desire to make a positive difference, school liaisons were provided with rubrics developed by the research team to identify key school personnel and students who exhibited the following characteristics assessed by the rubric: caring for students; having a desire to be a positive influence on the school climate; being approachable to students; caring about addressing the problem of bullying; and having leadership qualities in the case of school personnel or leadership, maturity, responsibility, caring toward others, influence, and a desire to be a positive influence on their peers in the case of students. For each item, school personnel and students were assessed on a 3-point scale, which included the ratings of *yes*, *somewhat*, or *no* for each item described previously. School personnel and students who scored *yes* or *somewhat* on all inclusion criteria were eligible to participate. Exclusion criteria for the study for both school personnel and students included having participated in a previous STAC study, speaking a language other than English or Spanish, and not providing consent or assent to participate.

The school liaison used the inclusion and exclusion criteria and the rubric to identify and contact key school personnel and students and then used the script to invite them to participate in the study.

Similarly to our previous usability research [40], we conducted the usability testing and interviews remotely. Research supports remote usability testing as a viable approach to gather high-quality user experience feedback [48,49]. To mitigate technology-related problems, before the testing session, a team member worked with the school liaison to ensure that they could open and operate the program using the school's computers and firewall. When problems occurred during the testing session, participants switched to a different device or, in one case, rescheduled the testing session.

During the usability testing, participants were asked to review the entire STAC-T application, including the booster session. Participants were asked to talk aloud while completing the tasks, identifying problems and the solutions attempted. The researchers and users were on videoconference and shared their screens. The researchers could see what the participants were doing, and they were able to communicate with each other in real time. The researchers observed the users as they worked through the tasks and asked questions to gather more data. After completing the STAC-T application, participants were asked to complete a brief usability survey followed by a semistructured interview and then a demographic questionnaire. All participants were asked to provide information about their perceptions of (1) program utility, (2) relevance and appropriateness of program content, (3) ways in which they would improve the program, and (4) using a bullying and strategy use tracker after the training. School personnel were also asked about (1) their thoughts on implementation feasibility, (2) the likelihood of school program adoption, (3) their thoughts on companion trainings for teachers and parents, and (4) barriers to program use. All individual interviews lasted 1 hour and were video recorded.

XSL•FO
RenderX

Measures

Demographics

Participants self-reported their age, ethnicity or race, and gender. Students also reported their grade level.

Program Usability

Usability was assessed using the System Usability Scale (SUS) [50]. The SUS is a widely used 10-item validated tool that measures the usability and acceptability of technology-based programs. Responses are measured on a 5-point Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). To calculate the SUS score, responses are converted as follows: (1) for odd-numbered items, 1 is subtracted from the response; (2) for even-numbered items, 5 is subtracted from the response; (3) the converted responses are added; and (4) the total is multiplied by 2.5. The final SUS score ranges from 0 to 100. An SUS score of ≥68 is considered above average [51].

Program User-Friendliness

One item selected from previous usability research [40] was used to assess the user-friendliness of the program. Participants were asked to rate the user-friendliness through the following question—"Overall, I would rate the user-friendliness of this program as:"—using a 7-point scale ranging from 0 (*worst imaginable*) to 7 (*best imaginable*).

Program Satisfaction

In total, 2 items selected from previous usability research [40] were used to assess program satisfaction. Participants were asked the following question—"Would you tell your friends/colleagues to use the program?"—with *yes*, *no*, and *don't know* as response choices. Participants were also asked how many stars they would give the program (1 star being the lowest and 5 stars being the highest).

Interview Questions

Following usability testing sessions, participants were asked a series of open-ended questions about the utility and relevance of the application prototype, as well as ways to improve the application, likelihood of program use, and potential implementation barriers. School personnel and students were asked the following: (1) "Please talk about your perception of how useful this program could be to helping to address the problem of bullying at school," (2) "Please share your thoughts on whether you think the content of this program is relevant and appropriate for students at your school and your community," and (3) "Can you talk about ways that you would improve the program?" Students were also asked the following: "If your school asked you to continue using the tracker, would it be useful?" School personnel were also asked the following: (1) "What are your thoughts on how practical or workable you think it would be to use this program at your school?" (2) "What do you believe is the likelihood that your school would use this intervention?" (3) "Do you think an online, brief teacher training and parent training module would be a helpful addition to this program?" (4) "What, if anything, would keep you from using this program?" (5) "How would you envision using the tracker, if at all, after completing the training modules and the two boosters?"

Data Analysis

Quantitative Analysis

Quantitative data from the questionnaires were analyzed using SPSS (version 29.0; IBM Corp). Before conducting statistical analyses, we examined the data for missing data points. We found no missing data. The data were also examined for outliers, defined as >3.3 SDs above the mean [52]. We found no outliers. To ensure that continuous data met statistical assumptions for parametric statistical tests, we assessed acceptable normality by using established guidelines for examining skew and kurtosis [53]. All continuous variables were within the acceptable range for skew and kurtosis. Descriptive statistics were used and are presented separately for school personnel and students. We examined differences between school personnel and students using 2-tailed independent-sample *t* tests for continuous variables and chi-square analyses for categorical variables. All statistical assumptions were met for the *t* tests and chi-square analyses. We controlled for type I error using the Bonferroni correction. On the basis of the calculated Bonferroni correction, all analyses were considered significant at $P<.004$.

Qualitative Analysis

Qualitative data from open-ended questions were analyzed separately for school personnel and students. In total, 3 team members, 2 of whom conducted the usability tests, transcribed the data verbatim. We used consensual qualitative research as a framework for data analysis. We used thematic analysis [54,55] to identify, analyze, organize, describe, and report themes found within the qualitative data. A faculty member with expertise in qualitative data analysis, along with 2 graduate students, 1 PhD student and 1 masters of art in counseling student, with previous experience in qualitative data analysis, analyzed the data. The faculty member led the data analysis team. The team met 2 times via Zoom (Zoom Video Communications). During their first meeting, the faculty member discussed the analysis protocol with the 2 students, as well as expectations and biases that they needed to be aware of as they analyzed the data. Each team member analyzed the transcripts for the school personnel and students separately to arrive at initial themes for each open-ended question from the interview protocol. Next, the team met one more time via Zoom and conducted additional email communication over a 4-week period to arrive at a consensus on themes and frequency categories supported by participant quotations. During their meetings, the team members shared their themes for each question and discussed agreement or disagreement about themes. The analysts relied on participant quotes to resolve disagreements. Once the team members reached a consensus, an external auditor reviewed the interview transcripts and themes for the school personnel and students. The auditor agreed with the team's findings. Interview data were deidentified to ensure anonymity, and quotes were identified by participant type (ie, school personnel or students).

Ethical Considerations

This study was registered with ClinicalTrials.gov (NCT05572398). All research procedures were approved (101-SB21-205) by Boise State University's institutional review

XSL•FO
RenderX

board. Researchers obtained informed consent from school personnel and parental consent and student assent in the case of students. For students, school liaisons met briefly with potential participants identified through the inclusion criteria to explain the project, and interested students were sent home with a letter describing the project as well as an informed consent form for the parent or guardian to sign. English and Spanish translations were provided in schools with a large Hispanic student population. Parents were also emailed the information and consent form. Parents could choose to sign the consent form via pen and paper or electronically. If they did not provide consent electronically, students were asked to return the signed consent form to the school liaison. Parents were provided with the study principal investigator's contact information and were encouraged to contact her if they had any questions or concerns. Students who returned the signed consent form were then provided with an opportunity to assent immediately before the interview, focus group, or usability testing session. In addition, several steps were taken to protect confidentiality: participants were informed that they were free to refrain from answering any questions, all data were identified only by a personal identifier number, and all research team members completed required training in protection of human research participants. School personnel received a US $50 Amazon gift card as an incentive for participation in the usability testing and individual interview. There were no incentives for student participants.

Results

Quantitative Analysis

Program Usability

Usability scores on the SUS are presented in Table 1. Overall, scores for both school personnel and students suggested a very high level of usability, functionality, and acceptability. As shown in Table 1, there were no differences in the scores on any of the individual items or the SUS total score between school personnel and students, with both participant groups scoring the STAC-T application at a very high level of usability.

Table 1. Means and SDs of the System Usability Scale scores from school personnel and students[a].

	School personnel (n=10), mean (SD)	Students (n=11), mean (SD)	*t* test (*df*)	*P* value
"I think that I would like to use the program frequently."	4.60 (0.52)	4.18 (0.60)	–1.70 (19)	.11
"I found the program to be more complex than it needed to be."	1.20 (0.42)	1.64 (0.92)	1.37 (19)	.19
"I thought the program was easy to use."	4.90 (0.32)	4.91 (0.30)	0.07 (19)	.95
"I think that I would need the support of a technical person to be able to use this program."	1.10 (0.32)	1.36 (0.67)	1.13 (19)	.27
"I found the various functions in the program were well put together with each other."	4.40 (1.27)	4.55 (0.69)	0.33 (19)	.74
"I thought there was too much inconsistency in this program."	1.00 (0.00)	1.55 (1.04)	1.66 (19)	.11
"I imagine that most people would learn to use this program very quickly."	4.90 (0.32)	4.55 (0.82)	–1.28 (19)	.22
"I found the program very awkward to use."	1.00 (0.00)	1.18 (0.40)	1.42 (19)	.17
"I felt very sure that I could use the program correctly."	4.90 (0.32)	4.55 (0.69)	–1.49 (19)	.15
"I needed to learn a lot of things before I could get going with this program."	1.00 (0.00)	1.55 (0.93)	1.84 (19)	.08
System Usability Scale total score	96.00 (3.94)	88.64 (9.51)	–2.27 (19)	.04

[a]Responses were scored on a 5-point Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*).

Program User Friendliness

School personnel and students rated the program highly on user-friendliness. Among school personnel, scores on user-friendliness ranged from 6.00 to 7.00 (mean 6.10, SD 0.32). Among students, scores on user-friendliness ranged from 6.00 to 7.00 (mean 6.09, SD 0.30). There were no differences in scores between school personnel and students on user-friendliness (t_{19}=–0.07; *P*=.95).

Program Satisfaction

Program satisfaction ratings are presented in Table 2. Overall ratings suggested that school personnel and students were satisfied with the program. There were no differences in scores between school personnel and students on program recommendation (χ^2_1=1.0; *P*=.33) or star ratings (χ^2_1=2.8; *P*=.25).

XSL•FO
RenderX

Table 2. Program satisfaction results for school personnel and students.

	School personnel (n=10), n (%)	Students (n=11), n (%)
Would recommend the program		
Yes	10 (100)	10 (91)
No	0 (0)	0 (0)
Unsure	0 (0)	1 (9)
Star rating		
1 star	0 (0)	0 (0)
2 stars	0 (0)	0 (0)
3 stars	1 (10)	0 (0)
4 stars	5 (50)	3 (27)
5 stars	4 (40)	8 (73)

Qualitative Analysis

Overview

Qualitative feedback for the STAC-T application supported the quantitative findings and was very positive overall, with both school personnel and students sharing the perception that the STAC-T application is useful, relevant, and appropriate, as well as providing feedback on ways to improve the program. In addition, school personnel shared positive thoughts about program practicality and adoption, including interest in a teacher and parent training, and discussed barriers to program adoption. Both school personnel and students talked about the benefits of tracking students' reports of different types of bullying and strategies they used to intervene both as part of the program and as a stand-alone feature to be used after the training. The results are presented in the following sections organized by the following themes: (1) usefulness, (2) relevance and appropriateness, (3) program improvement, (4) program tracker, (5) practicality and adoption, (6) teacher and parent training, and (7) barriers.

Usefulness

All school personnel (10/10, 100%) and students (11/11, 100%) indicated that the program was useful and increased students' knowledge to intervene in bullying situations. For example, a school personnel member shared the following:

> *...the program that you guys are creating, will definitely inform the students what they need to look for, how they can become an active positive person.*

A student also stated the following:

> *I think it'll be helpful by telling other kids that it's not right to bully other kids because you don't know how they feel, and you don't know what they go through.*

In terms of increasing knowledge, a school personnel member indicated the following:

> *I like how it has all the different ways for the students to see how you can step in you know...that there's things they can do.*

Another one added the following:

> *...it gives students the tools that they might feel like they lack in general when bullying happens.*

A student stated the following:

> *It can teach ways on how to and when bullying is happening, how do we handle it, and make it stop quicker.*

Another student added the following:

> *...if people are making fun of a person about how they look or the way they eat, and they post a video on social media, I can like easily screenshot, show it to the principal, my teachers, to get these people to stop and like get them to stop the bullying.*

Relevance and Appropriateness

All school personnel (10/10, 100%) and students (11/11, 100%) reported that the program content was relevant and relatable and taught students empathy and prosocial attitudes. For example, one school personnel member shared the following:

> *I think the content was really relevant. It's things that you actually see at school or that you hear about or that we, that get actually reported.*

Another school personnel member added the following:

> *Oh yeah, absolutely. I think that they [students] can relate to cafeteria situations, getting on the bus and you know posting things especially you know on social media or on Instagram or TikTok or whatever.*

A student shared the following:

> *I think that is really, really relevant and I think that a lot of the situations that were used in this app as examples can be used. They can be real life situations.*

In terms of empathy and prosocial attitudes, a school personnel member said the following:

> *I think this program would help...teach more empathy cause when students have more empathy, they're less likely to exhibit those bullying behaviors.*

A student indicated the following:

XSL•FO
RenderX

It basically says that bullying is not okay and if you do see it here are a few ways on how to stop it.

Program Improvement

Both school personnel (9/10, 90%) and students (6/11, 55%) offered feedback on how to improve the program and talked about the importance of having the program be fully narrated. School personnel talked about ways to improve student engagement and user experience. For example, a school personnel member stated the following:

...making it so that all of the pop-ups and scenes are narrated.

A student also said the following:

Once you add that narrative it is going to be good because most middle schoolers they aren't going to want to read it.

In terms of improving engagement and user experience, a school personnel member said the following:

Just there was one [activity] where you had to click the arrows to move on and I feel like just making it a little bit more simple.

Another school personnel member added the following:

Slowing down the captions on the cartoons because to allow kids to see the picture to get a frame of mind of what's going on and then read the words.

Program Tracker

All the school personnel (10/10, 100%) and students (11/11, 100%) indicated that they would find using a bullying and strategy use tracker useful if their school asked them to continue to use it. When asked about how they would envision using the tracker, school personnel indicated that they would use it for data collection for programmatic feedback and prevention as well as ongoing teaching and student support. For example, when asked about the usefulness of the tracker, a school personnel member stated the following:

Yeah, you know, I think that would be good.

Another one added the following:

You can bring that data up and say, you know, you know, we're, we're seeing this type of bullying going on.

A student said the following:

Oh, yeah. Because it would let more people know that bullying has been going on and stuff.

Practicality and Adoption

All the school personnel (10/10, 100%) agreed that the program would be practical and workable, and almost all school personnel (9/10, 90%) stated that they would be likely to use the program at their school. For example, one school personnel member said the following:

...our students would definitely get on and be able to, you know, go through that program without any problems at all.

Another school personnel member stated the following:

Practical. I think that it addresses the needs of our students in their day-to-day interactions.

In terms of likelihood of use, one school personnel member stated the following:

I think it would be highly likely.

Another school personnel member added the following:

This will be helpful and they [schools] will use it because it wouldn't take too much time away from the academics, academic goals that we have.

Teacher and Parent Training Modules

All school personnel (10/10, 100%) reported that a brief online teacher and parent training would be useful, and almost all (9/10, 90%) stated that it would provide a common language and a means for future collaboration. For example, one school personnel member shared the following:

It's teaching the parents and the teachers what that looks like, you know, conversation or tools to help our kids.

Another school personnel member added the following:

Yes. Yes, so that you have this so that we're all working together as a team. And using that common language.

Barriers

When asked about barriers to using the program, most school personnel members (7/10, 70%) reiterated that they would use the program, but many of them (8/10, 80%) discussed barriers. For example, one school personnel member stated the following:

I can't think of one negative reason or one reason I wouldn't want to use it.

However, other school personnel identified potential barriers, with one member stating the following:

Not being accessible on the devices that the kids have available.

Another school personnel member said the following:

Time. But I don't see that as a factor for us because we could fit it into our advisory class.

A third school personnel member added the following:

Only thing I can think of is funding.

Discussion

Principal Findings

The purpose of this study was to examine the usability of the STAC-T application, a technology-based bystander bullying intervention designed specifically for middle schools in rural communities. We were interested in feedback from middle school personnel as they are in the position of making decisions related to adopting and implementing bullying interventions and from students as end users. The primary aim of this study was to test the usability of the STAC-T application and assess

XSL•FO
RenderX

program utility, user-friendliness, and relevance as well as feasibility and ways to improve the program. Overall, both the quantitative survey results and qualitative interview findings indicate that participants perceived the STAC-T application to be useful, user-friendly, and appropriate for students at their schools and reported high levels of satisfaction with the program. The findings of this study indicate that the STAC-T application is relevant and feasible for implementation in middle schools in rural communities. The quantitative and qualitative findings are consistent with our previous usability testing research [40].

The findings of this study provide support for the usability of the STAC-T application. Both school personnel (mean 96.0, SD 3.9) and student (mean 88.6, SD 9.5) scores on the SUS demonstrated a very high level of usability, exceeding the standard cutoff score of 68 [51]. Both school personnel and students also rated the user-friendliness of the STAC-T application as very high, with all participants rating the program at ≥6 on a scale ranging from 0 to 7. We found no differences between school personnel and students on SUS scores or user-friendliness ratings, suggesting that both groups of users found the STAC-T application to be highly usable. The qualitative data supported these results, with both school personnel and students indicating that they perceived the program to be useful as well as relevant and appropriate for middle school students in rural communities. Furthermore, both school personnel and students reported high levels of satisfaction with the program, with 100% (10/10) of school personnel and 91% (10/11) of students indicating that they would recommend the program to others. Furthermore, most participants (20/21, 95%) gave the program 4 or 5 stars on a scale ranging from 1 to 5 stars. These findings are particularly important as usability and acceptability are associated with both program adoption and implementation [42].

In terms of practicality and adoption of the intervention, school personnel believed that their school would be likely to use the STAC-T application and identified cost, time, and access as potential barriers. Our results are aligned with those of previous studies that indicate that school administrators in rural communities feel favorably about adopting and implementing online programs to address the problem of bullying [20]. In addition, our findings are similar to those of previous studies that identify cost [7,20], time [56], and access to technology [57] as notable barriers to online programming implementation in schools located in rural areas.

School personnel and students reported positive perceptions of the STAC-T program, as well as feedback for program improvement. The results of the qualitative analyses showed that both school personnel and students thought that the STAC-T content was relevant and relatable and increased students' knowledge of how to intervene in bullying situations. School personnel also stated that the program taught students empathy and prosocial attitudes. Furthermore, school personnel stated that the program was practical and workable and would be used at their school and that teacher and parent trainings would be useful additions to STAC-T to provide a common language among stakeholders. These findings suggest a need for bystander bullying intervention programs in rural schools that teach students how to intervene when they witness bullying behaviors through conducting role-plays to practice strategies across different bullying scenarios. In terms of program improvement, both school personnel and students highlighted the importance of having the entire program narrated to students. School personnel also talked about ways to improve user engagement by simplifying and slowing down a few program activities. Both students and school personnel also saw value in the posttraining tracker. Including these program modifications is likely to increase user engagement [40], which, in turn, can influence positive behavior change [47], potentially increasing the efficacy of STAC-T.

Limitations

This study supports the usability, relevance, and feasibility of the STAC-T prototype. However, certain limitations must be noted. Participants were recruited from 2 schools in rural, low-income areas from 2 states, one in the Northwestern region and one in the Southern region of the United States. Although participants were recruited from 2 different states to increase generalizability, school personnel and students from different regions of the country may have a different perspective. Future research including a broader geographic sample would increase generalizability. Furthermore, most of the school personnel participants in this study were female, further limiting the generalizability of the study. In addition, because of the small sample size, we were unable to explore differences in demographic factors, limiting information that could guide improvements to the STAC-T application. Finally, it is also possible that social desirability influenced participants as they were aware that the goal of the study was to translate the in-person STAC intervention to a technology-based format.

Implications

This study has important implications for the development and implementation of STAC-T in middle schools in rural communities. First, participants provided very high usability ratings for the STAC-T application, with qualitative data supporting the usability, utility, and relevance of the program. The findings also provided valuable information about the program itself, including the need for program narration and the utility of a type of bullying and strategy tracker that could be used after the training. School personnel also provided feedback about the importance of both teacher and parent modules to foster collaboration. These modules could be developed and offered to schools as companion modules for STAC-T. School personnel could provide teachers with the STAC-T teacher training during professional development time at school and provide parents with a link to the parent training via email. Training teachers, parents, and students could provide a common language and increase buy-in from all stakeholders [58,59]. In addition, participants indicated that program implementation is feasible as long as the program is cost-effective, brief, and students can access it on their school devices. Translating STAC from an in-person modality to an online platform can help decrease barriers to implementation, increasing implementation feasibility for rural schools by decreasing program costs and reducing the demand on schools by decreasing the need for staff training, in-class time, and

XSL•FO
RenderX

expert support. Overall, the findings of this study provide valuable feedback and a strong scientific premise for moving forward with a large, multisite randomized controlled trial to examine the efficacy of the STAC-T application.

Conclusions

Bullying is a significant public health concern for students in rural middle schools. Training students to intervene and developing programming that increases access are important factors in addressing this problem. The findings of this study demonstrate the usability, relevance, and feasibility of the STAC-T application. The preliminary data from this study support conducting a large, multisite randomized controlled trial to assess the efficacy of the STAC-T application for middle school students in rural communities. Technology-based bullying interventions, and STAC-T in particular, could be an instrumental approach to decreasing educational and mental health disparities in rural schools and helping address the problem of bullying.

Acknowledgments

This research was supported by the National Institute on Minority Health and Health Disparities of the National Institutes of Health under award 2R42MD014943-02A1. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

AM, DMD, and MKB are developers of the technology-based STAC program. All other authors declare no conflicts of interest.

References

1. Student reports of bullying: results from the 2022 school crime supplement to the national crime victimization survey. National Center for Education Statistics. URL: https://nces.ed.gov/pubs2024/2024109rev.pdf [accessed 2024-10-13]
2. Moore SE, Norman RE, Suetani S, Thomas HJ, Sly PD, Scott JG. Consequences of bullying victimization in childhood and adolescence: a systematic review and meta-analysis. World J Psychiatry 2017 Mar 22;7(1):60-76 [FREE Full text] [doi: 10.5498/wjp.v7.i1.60] [Medline: 28401049]
3. Fisher BW, Gardella JH, Teurbe-Tolon AR. Peer cybervictimization among adolescents and the associated internalizing and externalizing problems: a meta-analysis. J Youth Adolesc 2016 Sep 22;45(9):1727-1743. [doi: 10.1007/s10964-016-0541-z] [Medline: 27447707]
4. Kowalski RM, Giumetti GW, Schroeder AN, Lattanner MR. Bullying in the digital age: a critical review and meta-analysis of cyberbullying research among youth. Psychol Bull 2014 Jul;140(4):1073-1137. [doi: 10.1037/a0035618] [Medline: 24512111]
5. Sticca F, Perren S. Is cyberbullying worse than traditional bullying? Examining the differential roles of medium, publicity, and anonymity for the perceived severity of bullying. J Youth Adolesc 2013 May 27;42(5):739-750. [doi: 10.1007/s10964-012-9867-3] [Medline: 23184483]
6. Movahedi N, Hosseinian S, Rezaeian H, Nooripour R. Mediating role of alexithymia in relationship between cyberbullying and psychotic experiences in adolescents. BMC Psychol 2024 Aug 31;12(1):465 [FREE Full text] [doi: 10.1186/s40359-024-01960-x] [Medline: 39217387]
7. Dulmus CN, Theriot MT, Sowers KM, Blackburn JA. Student reports of peer bullying victimization in a rural school. Stress Trauma Crisis 2004 Jan;7(1):1-16. [doi: 10.1080/15434610490281093]
8. Leadbeater BJ, Sukhawathanakul P, Smith A, Thompson RS, Gladstone EJ, Sklar N. Bullying and victimization in rural schools: risks, reasons, and responses. J Rural Community Dev 2013;8(1):31-47.
9. Nansel TR, Overpeck M, Pilla RS, Ruan WJ, Simons-Morton B, Scheidt P. Bullying behaviors among US youth: prevalence and association with psychosocial adjustment. JAMA 2001 Apr 25;285(16):2094-2100 [FREE Full text] [doi: 10.1001/jama.285.16.2094] [Medline: 11311098]
10. Burns E, Mann R, Yanez C. Student reports of bullying: results from the 2019 school crime supplement to the National Crime Victimization Survey. Institute of Education Sciences. URL: https://nces.ed.gov/pubs2022/2022031.pdf [accessed 2024-10-13]
11. Evans CB, Smokowski PR, Cotter KL. Cumulative bullying victimization: an investigation of the dose–response relationship between victimization and the associated mental health outcomes, social supports, and school experiences of rural adolescents. Child Youth Serv Rev 2014 Sep;44:256-264. [doi: 10.1016/j.childyouth.2014.06.021]
12. Smokowski PR, Cotter KL, Robertson C, Guo S. Demographic, psychological, and school environment correlates of bullying victimization and school hassles in rural youth. J Criminol 2013 Jun 11;2013:1-13. [doi: 10.1155/2013/137583]
13. Gaffney H, Ttofi MM, Farrington DP. Evaluating the effectiveness of school-bullying prevention programs: an updated meta-analytical review. Aggress Violent Behav 2019 Mar;45:111-133 [FREE Full text] [doi: 10.1016/j.avb.2018.07.001]
14. Wu WC, Luu S, Luh DL. Defending behaviors, bullying roles, and their associations with mental health in junior high school students: a population-based study. BMC Public Health 2016 Oct 10;16(1):1066 [FREE Full text] [doi: 10.1186/s12889-016-3721-6] [Medline: 27724881]

XSL•FO
RenderX

15. Waasdorp TE, Bradshaw CP. Examining variation in adolescent bystanders' responses to bullying. Sch Psychol Rev 2019 Dec 27;47(1):18-33. [doi: 10.17105/spr-2017-0081.v47-1]
16. Bauman S, Yoon J, Iurino C, Hackett L. Experiences of adolescent witnesses to peer victimization: the bystander effect. J Sch Psychol 2020 Jun;80:1-14. [doi: 10.1016/j.jsp.2020.03.002] [Medline: 32540087]
17. Reinke WM, Stormont M, Herman KC, Puri R, Goel N. Supporting children's mental health in schools: teacher perceptions of needs, roles, and barriers. Sch Psychol Q 2011 Mar;26(1):1-13. [doi: 10.1037/A0022714]
18. IT asset management for schools. NetSupport DNA. URL: http://www.netsupportdna.com/us/education/ [accessed 2021-09-21]
19. Salimi N, Karimi-Shahanjarin A, Rezapur-Shahkolai F, Hamzeh B, Roshanaei G, Babamiri M. Use of a mixed-methods approach to evaluate the implementation of violence and bullying prevention programs in schools. Educ Urban Soc 2020 Nov 11;53(6):607-628. [doi: 10.1177/0013124520972090]
20. Rachfal CL, Gilroy AA. Broadband internet access and the digital divide: federal assistance programs. Congressional Research Service. URL: https://sgp.fas.org/crs/misc/RL30719.pdf [accessed 2024-10-13]
21. USDA launches new program to create high-speed internet e-connectivity in rural America. U.S. Department of Agriculture. 2018. URL: https://www.usda.gov/media/press-releases/2018/12/13/usda-launches-new-program-create-high-speed-internet-e-connectivity [accessed 2021-09-21]
22. Midgett A, Doumas DM, Myers VH, Moody S, Doud A. Technology-based bullying intervention for rural schools: perspectives on needs, challenges, and design. Rural Ment Health 2021 Jan;45(1):14-30 [FREE Full text] [doi: 10.1037/rmh0000151] [Medline: 33777286]
23. Miller DT, Prentice DA. Changing norms to change behavior. Annu Rev Psychol 2016 Jan 04;67(1):339-361. [doi: 10.1146/annurev-psych-010814-015013] [Medline: 26253542]
24. Prescient and strategic intelligence marketing research web site. Digital Health Market. URL: https://www.psmarketresearch.com/market-analysis/digital-health-market [accessed 2021-08-20]
25. Fast facts: educational institutions. U.S. Department of Education, Institute of Education Sciences. URL: https://nces.ed.gov/fastfacts/display.asp?id=84 [accessed 2021-08-20]
26. Evers KE, Prochaska JO, Van Marter DF, Johnson JL, Prochaska JM. Transtheoretical-based bullying prevention effectiveness trials in middle schools and high schools. Educ Res 2008 May 28;49(4):397-414. [doi: 10.1080/00131880701717271]
27. Midgett A, Doumas D, Sears D, Lundquist A, Hausheer R. A bystander bullying psychoeducation program with middle school students: a preliminary report. Prof Couns 2015 Dec;5(4):486-500 [FREE Full text] [doi: 10.15241/AM.5.4.486]
28. Midgett A, Doumas DM, Johnston AD. Establishing school counselors as leaders in bullying curriculum delivery: evaluation of a brief, school-wide bystander intervention. Prof Sch Couns 2018 May 28;21(1):1-9. [doi: 10.1177/2156759X18778781]
29. Midgett A, Doumas DM, Trull R, Johnson J. Training students who occasionally bully to be peer advocates: is a bystander intervention effective in reducing bullying behavior? J Child Adolesc Couns 2017 Jan 19;3(1):1-13. [doi: 10.1080/23727810.2016.1277116]
30. Doumas DM, Midgett A, Watts AD. The impact of a brief, bullying bystander intervention on internalizing symptoms: is gender a moderator of intervention effects? Sch Psychol Int 2019 Feb 14;40(3):275-293. [doi: 10.1177/0143034319830149]
31. Midgett A, Doumas DM. The impact of a brief bullying bystander intervention on depressive symptoms. J Couns Dev 2019 Jun 13;97(3):270-280. [doi: 10.1002/jcad.12267]
32. Midgett A, Doumas DM, Trull R. Evaluation of a bystander bullying intervention program for elementary school students. Prof Sch Couns 2018 Feb 15;20(1):172-183. [doi: 10.5330/1096-2409-20.1.172]
33. Midgett A, Doumas DM, Trull R, Johnston A. A randomized controlled study evaluating a brief, bystander bullying intervention with junior high school students. J Sch Couns 2017;15(9):1-34 [FREE Full text]
34. Watts AD, Doumas DM, Midgett A. Efficacy of a brief, school‐based bystander bullying intervention on high school students' alcohol use. J Addict Offender Couns 2019 Oct 07;40(2):66-83. [doi: 10.1002/jaoc.12066]
35. Midgett A, Doumas DM. Acceptability and short-term outcomes of a brief, bystander intervention program to decrease bullying in an ethnically blended school in low-income community. Contemp Sch Psychol 2020 Sep 18;24(4):508-517. [doi: 10.1007/S40688-020-00321-W]
36. Midgett A, Doumas DM, Moran M, Gallo L. Evaluation of a brief, school-based bystander bullying intervention: a pilot study conducted at an ethnically-blended, low-income school. J Educ Psychol Res (Overl Park) 2020;2(2):52-60 [FREE Full text] [doi: 10.33140/jepr.02.02.01] [Medline: 38009113]
37. Moran M, Midgett A, Doumas DM, Moody S, Porchia S. A mixed method evaluation of a culturally adapted, brief, bullying bystander intervention for middle school students. J Child Adolesc Couns 2019;5(3):221-238 [FREE Full text] [doi: 10.1080/23727810.2019.1669372] [Medline: 33015352]
38. Moran M, Midgett A, Doumas DM. Evaluation of a brief, bystander bullying intervention (STAC) for ethnically blended middle schools in low-income communities. Prof Sch Couns 2020 Jul 20;23(1):2156759X2094064. [doi: 10.1177/2156759X20940641]
39. Midgett A, Doumas DM, Peralta C, Bond L, Flay B. Impact of a brief, bystander bullying prevention program on depressive symptoms and passive suicidal ideation: a program evaluation model for school personnel. J Prev Health Promot 2020 Jul 26;1(1):80-103 [FREE Full text] [doi: 10.1177/2632077020942959] [Medline: 33738442]

XSL•FO
RenderX

40. Doumas DM, Midgett A, Myers V, Buller MK. Usability of a technology-based bystander bullying intervention for middle school students in rural, low-income communities: mixed methods study. JMIR Form Res 2021 Oct 26;5(10):e32382 [FREE Full text] [doi: 10.2196/32382] [Medline: 34582355]
41. Wichansky AM. Usability testing in 2000 and beyond. Ergonomics 2000 Jul 10;43(7):998-1006. [doi: 10.1080/001401300409170] [Medline: 10929833]
42. Lyon AR, Bruns EJ. User-centered redesign of evidence-based psychosocial interventions to enhance implementation-hospitable soil or better seeds? JAMA Psychiatry 2019 Jan 01;76(1):3-4. [doi: 10.1001/jamapsychiatry.2018.3060] [Medline: 30427985]
43. Zichermann G, Cunningham C. Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps. New York, NY: O'Reilly Media, Inc; 2011.
44. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
45. Oinas-Kukkonen H, Harjumaa M. Persuasive systems design: key issues, process model, and system features. Commun Assoc Inf Syst 2009;24(1):28. [doi: 10.17705/1cais.02428]
46. Oinas-Kukkonen H, Harjumaa M. A systematic framework for designing and evaluating persuasive systems. In: Proceedings of the 3rd International Conference on Persuasive Technology. 2008 Presented at: PERSUASIVE '08; June 4-6, 2008; Oulu, Finland p. 164-176 URL: https://link.springer.com/book/10.1007/978-3-540-68504-3 [doi: 10.1007/978-3-540-68504-3_15]
47. Lehto T, Oinas-Kukkonen H. Persuasive features in web-based alcohol and smoking interventions: a systematic review of the literature. J Med Internet Res 2011 Jul 22;13(3):e46 [FREE Full text] [doi: 10.2196/jmir.1559] [Medline: 21795238]
48. Norman DA, Draper SW. User Centered System Design: New Perspectives on Human-Computer Interaction. Mahwah, NJ: Lawrence Erlbaum Associates Inc; 1986.
49. Peterson C. Bringing ADDIE to life: instructional design at its best. J Educ Multimed Hypermedia 2003;12(3):227-241. [doi: 10.1007/978-0-387-09506-6_3]
50. Brooke J. SUS: a 'quick and dirty' usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation In Industry. London, UK: CRC Press; 1996.
51. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. Usability Stud 2009;4(3):114-123 [FREE Full text]
52. Tabachnick BG, Fidell LS. Using Multivariate Statistics. 5th edition. Boston, MA: Allyn and Bacon; 2007.
53. Kline RB. Becoming a Behavioral Science Researcher, Second Edition: A Guide to Producing Research That Matters. New York, NY: The Guilford Press; 2009.
54. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
55. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis: striving to meet the trustworthiness criteria. Int J Qual Methods 2017 Oct 02;16(1):160940691773384. [doi: 10.1177/1609406917733847]
56. Beames JR, Johnston L, O'Dea B, Torok M, Christensen H, Boydell KM, et al. Factors that help and hinder the implementation of digital depression prevention programs: school-based cross-sectional study. J Med Internet Res 2021 Aug 27;23(8):e26223 [FREE Full text] [doi: 10.2196/26223] [Medline: 34448701]
57. Graves JM, Abshire DA, Amiri S, Mackelprang JL. Disparities in technology and broadband internet access across rurality: implications for health and education. Fam Community Health 2021;44(4):257-265. [doi: 10.1097/FCH.0000000000000306] [Medline: 34269696]
58. Midgett A, Doumas DM, Hausheer R. Development of a teacher module for a brief, bystander bullying intervention for middle schools: perspectives from school personnel. Contemp School Psychol 2022 Jun 23;27(3):491-504. [doi: 10.1007/S40688-022-00413-9]
59. Midgett A, Doumas DM, Hausheer R, Bond L, Buller MK, Peralta C, et al. Feasibility of a bullying bystander intervention (STAC) parent module for rural schools. J Prev Health Promot 2023 Apr 08;4(2):292-319 [FREE Full text] [doi: 10.1177/26320770231183961] [Medline: 38826496]

Abbreviations

SUS: System Usability Scale

XSL•FO
RenderX

Edited by A Kushniruk; submitted 24.10.24; peer-reviewed by A AL-Asadi, R Nooripour; comments to author 23.12.24; revised version received 13.01.25; accepted 18.01.25; published 21.02.25.

Please cite as:
Midgett A, Doumas DM, Peralta C, Peck M, Reilly B, Buller MK
Usability Testing of a Bystander Bullying Intervention for Rural Middle Schools: Mixed Methods Study
JMIR Hum Factors 2025;12:e67962
URL: https://humanfactors.jmir.org/2025/1/e67962
doi:10.2196/67962
PMID:

©Aida Midgett, Diana M Doumas, Claudia Peralta, Matt Peck, Blaine Reilly, Mary K Buller. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 21.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluation of Satisfaction With a Secure, Connected Mobile App for Women in Assisted Reproductive Technology Programs: Prospective Observational Study

Pauline Plouvier[1], MD; Romaric Marcilly[2,3], PhD; Geoffroy Robin[1], MD; Chaymae Benamar[2,3], MErg; Camille Robin[1], MD; Virginie Simon[1], MD; Anne Sophie Piau[1]; Isabelle Cambay[1]; Jessica Schiro[2,3], PhD; Christine Decanter[1], MD

[1]Departments of Assisted Reproductive Technologies and Fertility Preservation, Jeanne de Flandre Hospital, Lille, France

[2]Inserm, CIC-IT 1403, F-59000 Lille, France

[3]Univ. Lille, CHU Lille, ULR 2694 - METRICS : Évaluation des technologies de santé et des pratiques médicales, F-59000 Lille, France

Corresponding Author:
Pauline Plouvier, MD
Departments of Assisted Reproductive Technologies and Fertility Preservation
Jeanne de Flandre Hospital
Avenue Eugène Avinée
Lille, F-59000
France
Phone: 33 0626394791
Email: pauline.plouvier@chu-lille.fr

Abstract

Background: Telemedicine has emerged rapidly as a novel and secure tool to deliver medical information and prescriptions. A secure, connected health care app (WiStim) has been developed in order to facilitate dialogue between patients and the medical team during an ovarian stimulation cycle for medically assisted reproduction (MAR).

Objective: This study aimed to evaluate the patients' and midwives' levels of satisfaction with the connected mobile app.

Methods: We conducted a prospective, observational, single-center study at Lille University Hospital, France. From May 1 to July 31, 2021, all women undergoing ovarian stimulation started to receive their treatment advice through the mobile app. A total of 184 women were included and they filled out the 30-item Usefulness Satisfaction and Ease-of-Use (USE) questionnaire, which examines the users' opinions in 4 dimensions: usefulness, ease of use, ease of learning, and satisfaction. The women also answered a series of closed and open questions. The 5 midwives in our assisted reproductive technology center filled out the French version of the 10-item System Usability Scale (SUS) when the app was implemented and then after 3 and 6 months of use. We also performed semistructured interviews with the midwives.

Results: Overall, 183 women using the app completed the questionnaire. None refused to use the app, and 1 withdrew from the study. The mean scores for the four USE dimensions were all significantly greater than 4, that is, the middle of the response scale. The women liked the app's ease of use, the access to tutorial videos, and the reminders about appointments and treatments. In particular, the women liked to be able to (re)read the information; this reassured them, might have reduced the number of missed appointments and treatments, and made them more independent during the day, especially when they were working. Some of the women regretted the loss of direct contact with the midwife. The mean SUS score was 76 (SD 13.54) at the start of the study, 75 (SD 17.16) after 3 months, and 84 (11.21) after 6 months. According to the adjective rating scale, these scores corresponded to good usability for the app. After the requisite training and a familiarization period, the midwives reported that using the app saved them 2 hours a day. The mobile app enabled better transmission of information and thus probably helped to decrease treatment errors.

Conclusions: The WiStim connected mobile app is one of the first reliable, secure apps in the field of MAR. The app reassured the patients during the ovarian stimulation. Women and the medical team considered that the app was easy and intuitive to use. Given the growth in demand for MAR programs and the medical team's workload, the time savings provided by the app constitute a nonnegligible advantage.

(JMIR Hum Factors 2025;12:e63570) doi:10.2196/63570

XSL•FO
RenderX

KEYWORDS

mobile apps; mHealth; mobile health; assisted reproductive technologies; evaluation; satisfaction; reproduction; fertility; ovarian stimulation; ease of use; usability; midwives; obstetrics; gynecology

Introduction

In 2018, the European Society of Human Reproduction and Embryology's 22nd report on medically assisted reproduction (MAR) in Europe highlighted a continuous increase in the number of treatment cycles and the broad range of techniques used [1].

Every day, around the world, thousands of women undergo hormone assays and ultrasound scans of the pelvis as part of their MAR program. These burdensome, complex procedures can generate stress and anxiety for the women and their partners [2-4]. The women consulting in MAR departments are young, active, and, in many cases, occupied by work and family activities. Furthermore, health care professionals have to deal with a growing administrative burden and increasing demand for MAR while still providing high-quality patient care and support.

A large number of "eHealth" smartphone apps are being developed [5-8]. WiStim is a secure mobile app created in 2016 to facilitate communication between MAR patients and medical teams. The app provides access to a variety of documents and media (eg, test results and tutorials on self-injection) and gives daily advice on treatment during the ovarian stimulation phase. The objectives are (1) secure communication with patients, (2) better traceability, (3) a lower frequency of treatment errors caused by poor understanding, and (4) time savings for the care team. At present, the patient has to pay a subscription fee; however, some hospitals are considering whether to subsidize or cover this fee.

During each in vitro fertilization, frozen embryo transfer, or intrauterine insemination cycle, the midwife calls the women daily in order to adjust the treatment, check on adherence, and answer any questions. These calls are time-consuming and can be replaced with a connected app.

The primary objective of the current study was to evaluate levels of satisfaction with the connected app, according to both the patients and health care professionals. The secondary objectives were to identify potential difficulties in the use of the app and to study any app-associated changes in the health care professionals' practices and work organization.

Methods

Overview

We conducted a prospective, observational, single-center study in the MAR department at Lille University Hospital (Lille, France) from May 1 to July 31, 2021.

Ethical Considerations

In line with the French legislation on human and social science surveys of routine clinical practice using anonymized personal data [9], approval by an independent ethics committee was neither required nor sought.

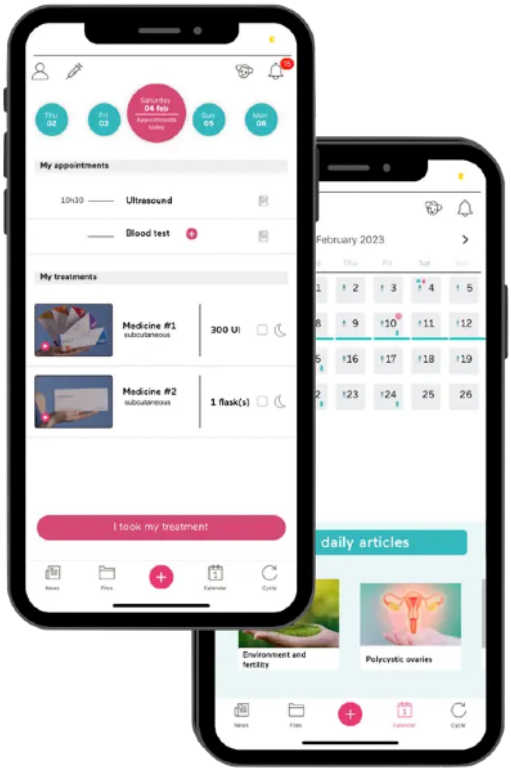
Connected Mobile App

The WiStim platform includes a mobile app for the patient (Figure 1) and a web-based management module for the medical team. The mobile app gives the patient access to advice on treatment, including the dose, the administration route, and any changes in the regimen. Tutorials on self-injections are also available. Every evening, the patient receives a reminder about her treatment and the date of her next appointment. The medical team sends information about the treatment and the next appointment through a secure web server (accredited by the French Ministry of Health) and is informed in real time about the woman's treatment.

The app requires a monthly subscription. However, Lille University Hospital has chosen to pay this cost on the patient's behalf during an initial test year.

XSL•FO
RenderX

Figure 1. The mobile app for the patient.



Patients

During a consultation, the gynecologist explained the app's principles to the patient and gave the latter an information sheet explaining how to download, install, and use the app. The patient then installed the app on her mobile phone and created an account. Once the stimulation phase had started, the patient received treatment advice through the app.

At the end of the MAR procedure, the women filled out a questionnaire (Multimedia Appendix 1). The questionnaire's closed questions concerned the type of MAR and the impact of using the app. The patient was able to write a comment for each closed question. Several open questions probed the woman's opinion of the app and the app's strengths and weaknesses. Next, the woman filled out the 30-item Usefulness Satisfaction and Ease-of-Use (USE) questionnaire, which measures the user's feelings in four dimensions: usefulness, ease of use, ease of learning, and satisfaction (Multimedia Appendix 2) [10,11]. The participants were not paid for their participation in the study.

Midwives

The five midwives in our MAR center filled out the French version of the System Usability Scale (SUS) when the app was implemented and then after 3 and 6 months of use [12]. The SUS (Multimedia Appendix 3) is a standardized, 10-item questionnaire for assessing the level of satisfaction with a technology's usability. For each of the 10 items, the midwives were invited to rate their level of agreement on a 5-point Likert scale ranging from "strongly disagree" to "strongly agree." We also performed semistructured interviews with the midwives 6 months after the implementation of the app in order to evaluate their opinion of the connected app, identify any changes in their work organization, and assess the app's acceptability (as defined in the unified theory of acceptance and use of technology [UTAUT]) [13]. The interviews were audio-recorded and transcribed.

Statistical Analyses

Statistical analyses were performed using Jamovi software (Jonathon Love, Damian Dropmann, and Ravi Selker, version 2.2.5, 2021), and thematic analyses of qualitative data were performed using QualCoder 3.0 (C Curtain). Data on the women's characteristics were used to describe the sample and perform subgroup analyses. The women's answers to the closed questions were analyzed using descriptive statistics, and the comments were analyzed qualitatively in order to highlight recurrent themes. Emerging themes in the qualitative analyses were not quantified; the goal was to understand the variety of the user's points of view. The data on the USE items and dimensions were analyzed with descriptive and inferential statistics. The threshold for statistical significance was set to $P<.05$. The results of the various questionnaires were compared in order to identify similarities and differences between the answers.

The SUS satisfaction score ranges from 1 to 100. Perceived satisfaction and usability are considered to be good when a score of 75 or more is obtained. To qualify the level of satisfaction, the mean SUS score was assessed against Aaron et al [14] grade scale and adjective rating scale. Semantic units (ie, sets of words representing the same idea) were extracted from the interviews with the midwives. An ergonomist attributed each semantic unit to one of the following themes: the opinion of the app, changes in practice, and dimensions of UTAUT acceptability (facilitating conditions, effort expectancy, social influence, and performance

XSL•FO
RenderX

expectancy). Within each theme, several subthemes were developed to represent the semantic units' diversity.

Results

Patients

Overview

During the recruitment period, 250 paper questionnaires were randomly distributed, and 184 patients responded (intrauterine insemination: 12%, in vitro fertilization: 73%, egg donation: 2%, frozen-thawed embryo transfer: 9%, fertility preservation: 4%). A total of 91 respondents had already been followed up by phone before the implementation of the app, 90 were monitored with the connected app alone, and 3 did not answer this item in the questionnaire. One woman discontinued use of the app prematurely, preferring direct contact with a midwife, and 2 women failed to complete the USE questionnaire.

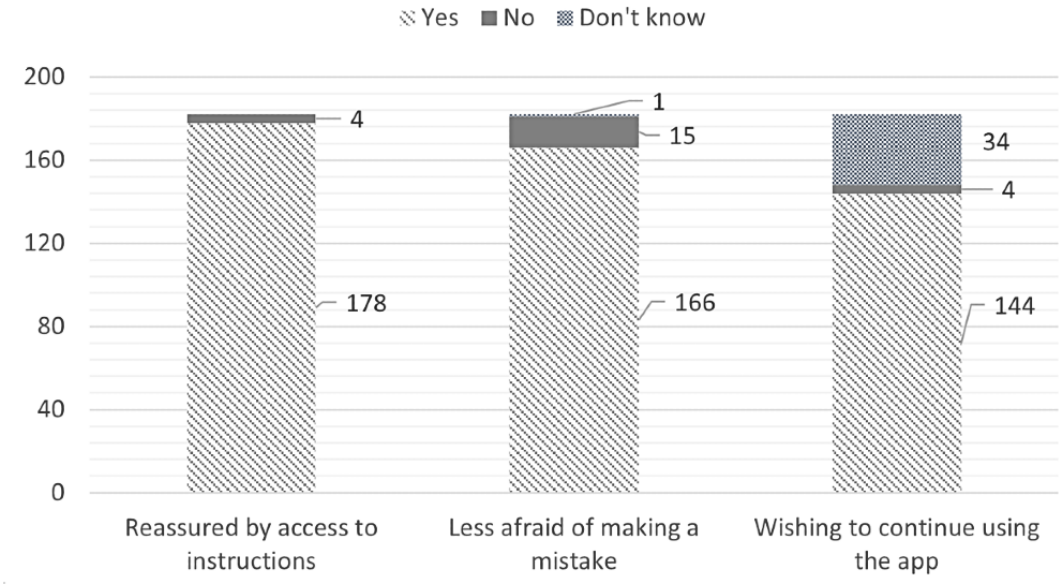
The mean scores for the four USE dimensions were all significantly greater than 4, that is, the middle of the response scale (usefulness, n=5.66, t_{180}=13.4, P<.001; ease of learning, n=6.25, t_{178}=18, P<.001; ease of use, n=5.96, t_{180}=16, P<.001; satisfaction, n=5.90, t_{177}=15.4, P<.001). The scores did not appear to be influenced by whether or not the women had been followed up by phone before the use of the app.

This good level of perceived usability was corroborated by our analysis of the women's comments and answers to the open questions. In all, 172 women described what they liked about the app. The app was seen as being easy to use in a guided, stepwise manner ("simple to use," "intuitive," and "clear"). Another reason for liking the app was its match with the women's needs and activities. The women mentioned the gain in efficiency linked to the app's centralization of information on treatment, follow-up, and appointments.

A total of 178 (97.8%) of the 182 respondents considered that the ability to access advice through the app was reassuring (Figure 2). The participants, having initially been followed up by phone, stated that they felt surer about the advice with the app because the instructions had to be rapidly written down during the phone call; with the app, the advice was available round the clock. The women also liked the additional information (eg, how to perform an injection) and the automatic reminders about appointments and treatments. A total of 166 (91.7%) of the 181 respondents considered that with the app, they were less stressed about making a treatment mistake. The participants liked the independence that the app provided; they no longer had to wait for the phone call from the midwife; previously, the call could come at any time in the afternoon.

Overall, 34 women did not reply to the question about wanting to continue to use the connected mobile app or not. 97.3% of the respondents said that they would be annoyed if they had to stop using the app (Figure 2). The 4 participants, who said they would feel relieved if they had to stop using the app, had all initially been followed up by phone and preferred direct communication with the medical team.

Figure 2. Answers to the following closed questions: "Does access to advice in the mobile app reassure you?" "Have you felt less stressed since you starting using the mobile app?" "Would you like to continue using the mobile app?".



Some weaknesses were mentioned: appointment errors, the lack of tutorials about some older models of self-injector pens, and late treatment reminders (ie, given after the scheduled time). Certain women expressed the need for direct interaction with the team of midwives via a chat function; the MAR department had decided not to offer this function because the women had already been given an email address for any questions.

A total of 53 (58.2%) of the 91 women who had initially been followed up by phone did not perceive any change in how they managed their treatment since the implementation of the app. In contrast, 38 (41.7%) women stated that the app modified the management of their treatment; they highlighted the lower mental burden, the easy, permanent access to information, the reassuring nature of the treatment reminders and app reminders,

XSL•FO
RenderX

and the fact their day no longer had to be organized around the call from the midwife.

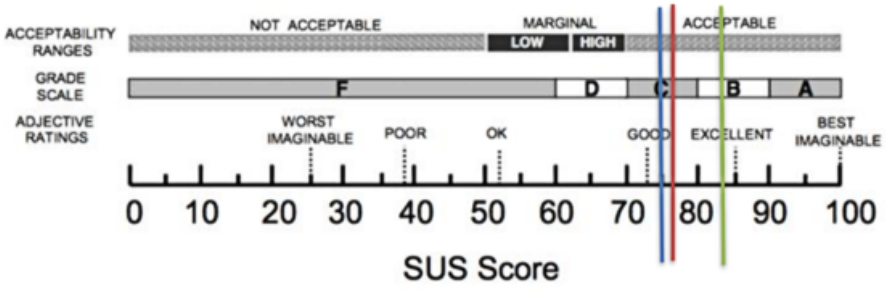
Midwives

The five midwives working in the MAR department at the time of the study filled out the SUS and were interviewed. The mean age was 43.6, and the midwives had been working in the department for an average of 6 years. The mean SUS score was 76 at the start of the study (n=4 respondents), 75 after 3 months (n=5), and 84 after 6 months (n=4; Figure 3). According to Aaron et al [14] adjective rating scale, these scores corresponded to good usability for the app.

A total of 207 semantic units were extracted from the semistructured interviews and analyzed. Even though the midwives stated that they needed some time to learn how to use the app and that the use of the app reduced the level of human contact with the women, they were generally very satisfied. The midwives found the interface pleasant to use and considered that the app was useful for the women and for the nurses who visited the women at home on a routine basis; the app modernized the midwives' practices.

During the interviews with the midwives, 3 of the 4 UTAUT dimensions emerged: facilitating conditions, effort expectancy, and performance expectancy. The social influence dimension did not emerge.

Figure 3. The mean System Usability Scale (SUS) score at the start of the study (76, SD 13.54) and after 3 months (75, SD 17.16) and 6 months (84, SD 11.21).



Facilitating Conditions

The midwives said that for the installation of the app, some IT equipment (computers) had been purchased and that they had attended training sessions on the use of the app. The midwives emphasized the good availability and responsiveness of the WiStim team when questions arose.

Effort Expectancy

The midwives viewed the app as a rather simple-to-use tool once they had been trained in its use. The app's different functions enable them to consult the patients' data, communicate with them, plan their appointments, and monitor their adherence to treatment.

Given that the app had not been integrated into the hospital information system at the time of the study, the midwives could not enter appointments into the hospital information system's diary from the app or vice versa; hence, to make appointments, they had to work on two IT tools in parallel. Even though the app had a simple-to-use interface, the fact that it was a web-based tool meant that the transitions were sometimes slow. Hence, data sometimes had to be entered twice: once on paper and then on the app.

Performance Expectancy

Use of the app was associated with a fall in the number of daily calls from midwives to patients. Nevertheless, the midwives continued to phone the women when the medical team had decided to stop the treatment or when the women had not ticked the boxes for adherence to treatment. This enabled them to optimize the dialogue by phone. Certain women continued to call or email the midwives regardless, to check that they had correctly understood the treatment procedures and advice specified by the app or to change an appointment. The midwives replied to these questions by email or by phone. The frequency of these requests tended to fall over time, as the midwives and the patients became more familiar with the app.

In order to ensure that the women were using the app correctly, the midwives helped some of them (especially those who did not understand or speak French sufficiently well) to install the app and explained how to use it. Furthermore, the advice stored in the app can be presented in several languages.

The decrease in the number of phone calls freed up an average of 2 hours per day, which enabled the midwives to perform other tasks (consultations, patient education sessions, etc).

Discussion

Principal Findings

On an annual basis, our center performs around 1400 oocyte retrieval procedures, 900 frozen embryo transfers, and 700 intrauterine inseminations. This corresponds to 30 to 50 women per day seen for ultrasound scans and hormone assays, and the volume of activity is increasing. Until now in France, MAR has only been available to heterosexual couples on medical indication. The law on bioethics, to be promulgated in 2021, extends MAR to female couples and single women. Activity in MAR centers has therefore risen sharply. Use of the connected mobile app was associated with a considerable decrease in the number of phone calls during the day. This resulted in significant time savings for the medical team, who were therefore able to perform other tasks. The women received their treatment advice in a secure manner, which had not been the case previously. The women felt reassured and less stressed about making treatment mistakes.

XSL•FO
RenderX

We did not question the partners, but in the open-ended questions, some of them said that they were reassured to be able to read the treatment instructions and felt more involved in the care compared with the phone call.

At present, it is difficult to quantify the extent of treatment errors during an ovarian stimulation program. Till date, no study has been published in the literature on this subject. Nevertheless, along with the treatment reminders, the fact that the woman's prescription is given in writing in the app and can be accessed at any time (and in several languages, if required) probably helped to reduce forgetfulness and treatment administration errors. This traceability is essential for secure communication between patients and professionals. Barrière et al [15] conducted an observational, real-life, longitudinal study involving 488 patients from 28 infertility centers in France to evaluate patient–infertility care provider relationships and communication in Assisted Reproductive Technology centers and investigate whether the quality of the care provided had an impact on patient adherence to treatment and monitoring protocols. They showed that even when patient-physician relationships appear to be satisfactory, patient miscomprehension and noncompliance during infertility treatment may be underestimated, and improvements in communication are also required.

Several studies [16,17] and reviews of the literature [18-20] have shown that stress is a major risk factor for impaired quality of life and a poor experience of infertility treatment. It also exposes patients to the risk of abandoning assisted reproduction procedures, thereby reducing their chances of pregnancy. It is therefore conceivable that mobile apps can reassure patients undergoing infertility treatments and consequently reduce the stress inherent in these procedures, as suggested in a recent review [21]. Further studies using standardized questionnaires and comparing the anxiety levels of the app users with those of nonusers could answer this hypothesis. The risk of dropping out of infertility management procedures could also be compared between these two groups.

Over the last decade, huge progress has been made in information and communication technology. "eHealth" and smart (connected) devices have brought together services that facilitate communication with the patient and thus improve the overall quality of care (teleconsultations, electronic health records, smart eHealth apps, etc) [22]. Telemedicine enables medical data to be shared with the patient or between health care professionals. Many connected health apps have been developed – notably in the fields of radiology and cardiology [23]. A recent study showed that connected health technologies can facilitate access to cancer care and improve the patient's psychological well-being and quality of life [24]. Recently, connected apps have been developed for use in gynecology and obstetrics [25]. There are probably other connected apps in the world in the field of fertility, but to date and to our knowledge, there are very few publications on this subject. Boivin et al [26] were interested in the development of a mobile app (MediEmo) to provide support during medically assisted reproduction. MediEmo is an app combining patient medication diary management and ease of integration into clinic systems with emotional support and data capture. They showed that 98% of patients expressed willingness to use the app, and almost 80% did so. Thus, the development of smartphone apps can contribute to fertility care and should be encouraged.

The profile of women in MAR programs is particularly suitable for this type of development: they are young and active and they all use a smartphone.

The protection of the patients' personal data is a crucial aspect in the development of connected health technologies [27,28]. WiStim complies with the European Union's General Data Protection Regulation, and the app's host is accredited by the French Ministry of Health.

One limitation is that the app is currently fee-based, although various funding options are being considered. In the present case, our hospital paid for the subscription to the app, and so it was cost-free for the users. However, the hospital's finance department is currently carrying out a medical-economic survey to assess the profitability of financing the app in relation to the time saved by the midwives. In fact, the midwife who used to call patients back can now do consultations instead.

Another limitation is the relatively small number of questionnaires analyzed and the fact that this was a single-center study. These are preliminary results, and they pave the way for further investigations in a larger population. Indeed, a multicenter study, with more questionnaires and over a longer period would be desirable to confirm these data.

Moreover, the questionnaires were distributed as soon as the app was installed in the department, so a learning curve was necessary for both patients and professionals. Overall satisfaction is probably higher now that the tool is better understood. Till now, all our patients have used the app.

Conclusion

Our results showed that the use of the connected health care app reassured women during the ovarian stimulation phase of a MAR program. Both the women and the medical team considered that the app was easy and intuitive to use. The information on treatment was sent in a secure manner and could be accessed around the clock by the patient. Connected health care apps have been developed widely in recent years. WiStim is one of the first reliable, secure apps in the field of MAR. Given the growth in demand for the MAR program and the medical team's workload, the time savings provided by the app constitute a nonnegligible advantage.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Patient questionnaire.
[DOCX File , 13 KB - humanfactors_v12i1e63570_app1.docx]

Multimedia Appendix 2
The Usefulness Satisfaction and Ease-of-Use questionnaire.
[DOCX File , 19 KB - humanfactors_v12i1e63570_app2.docx]

Multimedia Appendix 3
System usability scale.
[DOCX File , 15 KB - humanfactors_v12i1e63570_app3.docx]

References

1. European IVF Monitoring Consortium (EIM), for the European Society of Human ReproductionEmbryology (ESHRE), Wyns C, De Geyter C, Calhaz-Jorge C, Kupka MS, Motrenko T, et al. ART in Europe, 2018: results generated from European registries by ESHRE. Hum Reprod Open 2022;2022(3):hoac022 [FREE Full text] [doi: 10.1093/hropen/hoac022] [Medline: 35795850]
2. Stanhiser J, Steiner AZ. Psychosocial aspects of fertility and assisted reproductive technology. Obstet Gynecol Clin North Am 2018;45(3):563-574. [doi: 10.1016/j.ogc.2018.04.006] [Medline: 30092929]
3. Koumparou M, Bakas P, Pantos K, Economou M, Chrousos G. Stress management and In Vitro Fertilization (IVF): A pilot randomized controlled trial. Psychiatriki 2021;32(4):290-299 [FREE Full text] [doi: 10.22365/jpsych.2021.029] [Medline: 34390556]
4. Ying L, Wu LH, Loke AY. The effects of psychosocial interventions on the mental health, pregnancy rates, and marital function of infertile couples undergoing in vitro fertilization: a systematic review. J Assist Reprod Genet 2016;33(6):689-701 [FREE Full text] [doi: 10.1007/s10815-016-0690-8] [Medline: 26979745]
5. Meykiechel T, de Carne C, Gueguen I, Vatier C, Girard G, Buzzi JC, et al. Monitoring gestational diabetes mellitus patients with the telemedicine application MyDiabby decreases the rate of foetal macrosomia. Diabetes Metab 2022;48(2):101301. [doi: 10.1016/j.diabet.2021.101301] [Medline: 34757162]
6. Xie Q, Torous J, Goldberg SB. E-mental health for people with personality disorders: a systematic review. Curr Psychiatry Rep 2022;24(10):541-552 [FREE Full text] [doi: 10.1007/s11920-022-01360-1] [Medline: 35972718]
7. Karlsen IL, Svendsen PA, Abildgaard JS. A review of smartphone applications designed to improve occupational health, safety, and well-being at workplaces. BMC Public Health 2022;22(1):1520 [FREE Full text] [doi: 10.1186/s12889-022-13821-6] [Medline: 35948904]
8. Kumar N, Khunger M, Gupta A, Garg N. A content analysis of smartphone-based applications for hypertension management. J Am Soc Hypertens 2015;9(2):130-136. [doi: 10.1016/j.jash.2014.12.001] [Medline: 25660364]
9. Public Health Code: Section 1: Definitions. (Articles R1121-1 to R1121-2) [article in French] . Légifrance - Le service public de la diffusion du droit. URL: https://www.legifrance.gouv.fr/codes/section_lc/LEGITEXT000006072665/LEGISCTA000006190943/#LEGISCTA000006190943
10. Lund AM. Measuring usability with the USE questionnaire. Usability Interface 2001;8(2):3-6 [FREE Full text]
11. Gao M, Kortum P, Oswald F. Psychometric evaluation of the USE (Usefulness, Satisfaction, and Ease of use) questionnaire for reliability and validity. Proc Hum Factors Ergon Soc Ann Mee 2018;62(1):1414-1418. [doi: 10.1177/1541931218621322]
12. Gronier GB, Baudet A. Psychometric evaluation of the F-SUS: creation and validation of the french version of the system usability scale. International Journal of Human–Computer Interaction 2021;37(16):1571-1582. [doi: 10.1080/10447318.2021.1898828]
13. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Quarterly 2003;27(3):425-478. [doi: 10.2307/30036540]
14. Aaron B, Kortum K, Miller M. Determining what individual SUS scores mean: adding an adjective rating scale. J. Usability Studies 2009;4(3):114-123 [FREE Full text]
15. Barrière P, Avril C, Benmahmoud-Zoubir A, Bénard N, Dejager S. Patient perceptions and understanding of treatment instructions for ovarian stimulation during infertility treatment. Reprod Biomed Soc Online 2019;9:37-47 [FREE Full text] [doi: 10.1016/j.rbms.2019.08.003] [Medline: 31993512]
16. Verberg MF, Eijkemans MJ, Heijnen EM, Broekmans FJ, de Klerk C, Fauser BC, et al. Why do couples drop-out from IVF treatment? A prospective cohort study. Hum Reprod 2008;23(9):2050-2055. [doi: 10.1093/humrep/den219] [Medline: 18544578]
17. Domar AD, Smith K, Conboy L, Iannone M, Alper M. A prospective investigation into the reasons why insured United States patients drop out of in vitro fertilization treatment. Fertil Steril 2010;94(4):1457-1459 [FREE Full text] [doi: 10.1016/j.fertnstert.2009.06.020] [Medline: 19591985]

XSL•FO
RenderX

18. Paraskevi L, Antigoni S, Kleanthi G. Stress and anxiety levels in couples who undergo fertility treatment: a review of systematic reviews. Mater Sociomed 2021;33(1):60-64 [FREE Full text] [doi: 10.5455/msm.2021.33.60-64] [Medline: 34012353]
19. Rich CW, Domar AD. Addressing the emotional barriers to access to reproductive care. Fertil Steril 2016;105(5):1124-1127 [FREE Full text] [doi: 10.1016/j.fertnstert.2016.02.017] [Medline: 27054306]
20. Cousineau TM, Domar AD. Psychological impact of infertility. Best Pract Res Clin Obstet Gynaecol 2007;21(2):293-308. [doi: 10.1016/j.bpobgyn.2006.12.003] [Medline: 17241818]
21. Meyers AJ, Domar AD. Research-supported mobile applications and internet-based technologies to mediate the psychological effects of infertility: a review. Reprod Biomed Online 2021;42(3):679-685. [doi: 10.1016/j.rbmo.2020.12.004] [Medline: 33487558]
22. Loncar-Turukalo T, Zdravevski E, Machado da Silva J, Chouvarda I, Trajkovik V. Literature on wearable technology for connected health: scoping review of research trends, advances, and barriers. J Med Internet Res 2019;21(9):e14017 [FREE Full text] [doi: 10.2196/14017] [Medline: 31489843]
23. Omboni S. Connected health in hypertension management. Front Cardiovasc Med 2019;6:76 [FREE Full text] [doi: 10.3389/fcvm.2019.00076] [Medline: 31263703]
24. Gitonga I, Desmond D, Duda N, Maguire R. Impact of connected health interventions on psychological wellbeing and quality of life in patients with cancer: a systematic review and meta-analysis. Psychooncology 2022;31(10):1621-1636 [FREE Full text] [doi: 10.1002/pon.6019] [Medline: 35996330]
25. DeNicola N, Marko K. Connected health and mobile apps in obstetrics and gynecology. Obstet Gynecol Clin North Am 2020;47(2):317-331. [doi: 10.1016/j.ogc.2020.02.008] [Medline: 32451020]
26. Robertson I, Harrison C, Ng KYB, Macklon N, Cheong Y, Boivin J. Development, implementation and initial feasibility testing of the MediEmo mobile application to provide support during medically assisted reproduction. Hum Reprod 2022;37(5):1007-1017 [FREE Full text] [doi: 10.1093/humrep/deac046] [Medline: 35348672]
27. Kamel Boulos MN, Brewer AC, Karimkhani C, Buller DB, Dellavalle RP. Mobile medical and health apps: state of the art, concerns, regulatory control and certification. Online J Public Health Inform 2014;5(3):e61275. [doi: 10.5210/ojphi.v5i3.4814]
28. Schäfer F, Quinquis L, Klein M, Escutnaire J, Chavanel F, Chevallier H, et al. Attitudes and expectations of clinical research participants toward digital health and mobile dietary assessment tools: cross-sectional survey study. Front Digit Health 2022;4:794908 [FREE Full text] [doi: 10.3389/fdgth.2022.794908] [Medline: 35355684]

Abbreviations

MAR: medically assisted reproduction
SUS: System Usability Scale
USE: Usefulness Satisfaction and Ease-of-Use
UTAUT: unified theory of acceptance and use of technology

Edited by A Kushniruk; submitted 12.07.24; peer-reviewed by F Dénis, A Guerin, EM d'Argent; comments to author 19.10.24; revised version received 31.10.24; accepted 07.11.24; published 24.02.25.

Please cite as:
Plouvier P, Marcilly R, Robin G, Benamar C, Robin C, Simon V, Piau AS, Cambay I, Schiro J, Decanter C
Evaluation of Satisfaction With a Secure, Connected Mobile App for Women in Assisted Reproductive Technology Programs: Prospective Observational Study
JMIR Hum Factors 2025;12:e63570
URL: https://humanfactors.jmir.org/2025/1/e63570
doi:10.2196/63570
PMID:

©Pauline Plouvier, Romaric Marcilly, Geoffroy Robin, Chaymae Benamar, Camille Robin, Virginie Simon, Anne Sophie Piau, Isabelle Cambay, Jessica Schiro, Christine Decanter. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 24.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Digital Cognitive Behavioral Therapy–Based Treatment for Insomnia, Nightmares, and Posttraumatic Stress Disorder Symptoms in Survivors of Wildfires: Pilot Randomized Feasibility Trial

Fadia Isaac[1], BA, BBS (Hons), MPsy; Britt Klein[1,2], PsyD; Huy Nguyen[2,3,4,5], PhD; Shaun Watson[1], PhD; Gerard A Kennedy[6], BBSc, PhD

[1]Institute of Health and Wellbeing, Federation University Australia, Victoria, Australia

[2]Health Innovation & Transformation Centre, Federation University Australia, Ballarat, Australia

[3]Department of Population and Quantitative Health Sciences, UMass Chan Medical School, Worcester, MI, United States

[4]Department of Epidemiology, Institute for Preventive Medicine and Public Health, Hanoi Medical University, Hanoi, Vietnam

[5]Faculty of Health Sciences, University of Cuu Long, Vinh Long Province, Vietnam

[6]School of Health and Biomedical Sciences, Royal Melbourne Institute of Technology University, Melbourne, Australia

Corresponding Author:
Fadia Isaac, BA, BBS (Hons), MPsy
Institute of Health and Wellbeing
Federation University Australia
University Drive
Mt Helen
Victoria, 3350
Australia
Phone: 61 353276717
Email: fadia.isaac@hotmail.com

Abstract

Background: Symptoms of insomnia, nightmares, and trauma are highly prevalent. However, there are significant barriers to accessing evidence-based treatments for these conditions, leading to poor mental health outcomes.

Objective: This pilot trial evaluated the feasibility of a 4-week, digital self-paced intervention combining cognitive behavioral therapy for insomnia and exposure, relaxation, and rescripting therapy for nightmares in survivors of wildfires from Australia, Canada, and the United States.

Methods: Study participants were recruited between May 2023 and December 2023 through social media platforms, workshops, conferences, and radio interviews. Participants had to meet at least one of the following criteria: a score of ≥8 on the Insomnia Severity Index, a score of ≥3 on the Nightmare Disorder Index, or a score of ≥31 on the PTSD Checklist for DSM-5. In total, 30 survivors of wildfires were allocated to either the treatment group (n=16, 53%) or the waitlist control group (n=14, 47%) in a sequential manner. Participants' ages ranged from 18 to 79 years, with a mean age of 52.50 (SD 16.26) years. The cohort consisted of 63% (19/30) female and 37% (11/30) male participants. Participants also completed self-report secondary outcome measures, including the Generalized Anxiety Disorder–7, the Patient Health Questionnaire–9, and the Pittsburgh Sleep Quality Index, via the HealthZone digital platform. Assessments were conducted at baseline, the posttreatment time point, and the 3-month follow-up, with the waitlist group undergoing an additional assessment at the pretreatment time point, after 4 weeks of waiting and before crossing over to treatment. This study used intention-to-treat analysis as a primary analysis and per-protocol analysis as a secondary analysis.

Results: Mixed-effects linear regression models and difference-in-differences analyses were used to assess the intervention's effects. The intention-to-treat analysis revealed significant improvements over time (main effect of time), with a 1.64-point reduction (P=.001) on the Nightmare Disorder Index and 10.64-point reduction (P=.009) on the PTSD Checklist for DSM-5 at the postintervention time point. No significant changes were observed in insomnia symptoms. On the secondary measures, there was an interaction effect of condition × time, with a 2.22-point reduction (P<.001) on the Pittsburgh Sleep Quality Index, and a main effect of time, with a 6.48-point reduction (P<.001) on the Patient Health Questionnaire–9. No changes were detected on

XSL•FO
RenderX

the Generalized Anxiety Disorder–7. The per-protocol analysis yielded comparable results for both the primary and secondary measures.

Conclusions: The findings of this pilot trial demonstrated a reduction in nightmares and trauma symptoms. Future research studies should aim at evaluating the intervention in a more definitive trial with a larger sample size.

Trial Registration: Australian New Zealand Clinical Trials Registry (ANZCTR) ACTRN12623000415606; https://anzctr.org.au/Trial/Registration/TrialReview.aspx?id=385054

(JMIR Hum Factors 2025;12:e65228) doi:10.2196/65228

KEYWORDS

insomnia; nightmares; posttraumatic stress disorder; PTSD; wildfires; cognitive behavioral therapy for insomnia; CBTi; exposure, relaxation, and rescripting therapy; ERRT; Sleep Best-i; mobile health; mHealth; digital health; computer; eHealth; bushfires

Introduction

Background

The consequences of wildfires can be devastating, resulting in significant losses of life, property, and livelihood. According to the Emergency Events Database, wildfires have been responsible for 1890 fatalities and 14,360 injuries between 2000 and 2023, with the highest tolls observed in Australia, New Zealand, Southern Europe, and North America [1]. The impact of wildfires extends far beyond the immediate loss of life and property, with survivors often experiencing a range of physical and emotional challenges that can have long-lasting effects. Physical injuries, such as bodily burns, smoke inhalation, and eye irritation, as well as increased risk of myocardial infarction and certain types of cancers, are a significant concern for survivors [2-4]. In addition, the proximity of wildfires to residential areas can have a profound impact on overall well-being, with research showing that life satisfaction decreases significantly in areas within 0 to 15 km of the fires [5]. The traumatic experience of a wildfire can also leave a lasting emotional impact, with survivors often reporting persistent feelings of fear and unsafety that can linger long after the initial event [5].

Beyond the physical and emotional challenges, many survivors face significant financial pressures, including the cost of rebuilding or relocating. For those who have lost their homes, the decision of whether to rebuild can be a difficult and emotional one, with financial pressures and the loss of a sense of security and community adding to the burden [3]. Others may experience stressors related to the loss of a workplace, hardship, and financial uncertainty. Therefore, the availability and timing of government relief and community support services are critical in facilitating the recovery of communities and individuals, with initiatives such as rehousing and recovery projects, psychological support, and primary producer repair and restoration [3,5] helping mitigate the impact of wildfires. However, when such support is not provided in a timely manner, emotional and psychological distress can intensify, leading to multiple psychological problems that can become highly resistant to treatment.

This is particularly concerning as research has shown that exposure to wildfires can have a significant impact on emotional well-being. Over the past decade, there has been a growing interest in investigating the impact of wildfire trauma on mental health [6,7]. Posttraumatic stress disorder (PTSD) is among the more severe mental health conditions that may arise immediately or in the months following exposure to traumatic events [7]. Exposure to wildfires, which pose a significant threat to life or well-being, can increase an individual's risk of developing PTSD. People with PTSD may exhibit symptoms such as reliving traumatic events through nightmares or flashbacks; sleeping difficulties; irritability; intrusive memories; feelings of horror, shame, and anger; avoidance of reminders of the traumatic event; alterations in arousal state or mood and cognition; and difficulties with work and daily activities [8].

Sleep disturbances are prominent features of PTSD and are frequently reported following trauma [9,10]. The onset of sleep disturbances following exposure to traumatic events is a strong predictor of the development of PTSD [11]. According to the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (*DSM-5*), insomnia and nightmares fall under the re-experiencing and the hyperarousal clusters of PTSD symptoms [8]. Insomnia is the most prevalent sleep disturbance in survivors of wildfires and is characterized by difficulty initiating or maintaining sleep or early-morning awakening [8]. Survivors of trauma also frequently experience nightmares, which are repeated distressing and well-remembered narrative dreams that lead to the individual waking up in a fight-flight state of alertness and having difficulty returning to sleep or even trying to avoid sleeping [8]. A study examining sleep disturbances and PTSD in survivors of wildfires revealed that approximately 49.2% (n=126) experienced symptoms of insomnia, 28.7% experienced nightmares, and 77.9% reported PTSD symptoms [12].

Studies suggest that cognitive behavioral therapy for insomnia (CBTi); cognitive behavioral therapy (CBT) for PTSD; and exposure, relaxation, and rescripting therapy (ERRT) for nightmares are effective treatments for insomnia, nightmares, and other trauma symptoms [13-18]. CBTi includes stimulus control, sleep hygiene, sleep restriction, cognitive restructuring, and relaxation training [19,20]. Nightmare treatment using ERRT is effective in giving those who experience nightmares a sense of control over the nightmare through writing benign dreams that can be rehearsed verbally or mentally [21]. The treatment focuses on addressing the themes presented in the nightmare and incorporates rescripting, imagery or verbal rehearsal, relaxation, and sleep hygiene elements [22,23].

XSL•FO
RenderX

Most research on trauma and sleep disorders has been conducted with military veterans, who have been the primary focus of clinical trials and studies in this area [24-26]. There is a need to focus on developing treatments for people who have experienced the trauma of natural disasters, including wildfires. There have only been 2 clinical trials addressing the treatment of sleep disturbances and PTSD in survivors of wildfires [27,28]. Both trials have shown that the administration of sleep dynamic therapy and CBTi therapist-assisted self-help interventions led to a significant reduction in insomnia and PTSD symptoms following treatment.

Objectives

While studies support the effectiveness of CBT, CBTi, and ERRT in treating trauma symptoms, insomnia, and nightmares, accessing these treatments can be challenging, particularly when wildfires cause destruction to infrastructure. This, in turn, causes difficulty in the delivery of counseling services to remote locations where large numbers of people are affected and in great need. Digital therapies can bridge this gap by delivering evidence-based treatments in communities affected by the trauma of wildfires [29,30]. However, the feasibility of digital therapies should be first examined in clinical trials before they are operationalized in the field [27,31]. Therefore, a brief, digital, self-paced, multicomponent therapeutic approach incorporating CBTi for insomnia, CBT for PTSD, and ERRT for nightmares was developed. The main objective of this pilot trial was to assess the feasibility of an intervention called Sleep Best-i in alleviating symptoms of insomnia, nightmares, and PTSD among survivors of wildfires. A secondary research objective was to undertake a comparative analysis of the degree of symptom reduction between the intervention and the waitlist control groups, with a view to informing an assessment of the feasibility of conducting a future definitive randomized controlled trial. Specifically, this assessment aimed to evaluate the acceptability and usability of the Sleep Best-i intervention. This clinical pilot trial tested the following hypotheses: (1) Sleep Best-i will result in significant reductions in insomnia, nightmares, and PTSD symptoms from baseline to the posttreatment time point compared to the waitlist group; (2) Sleep Best-i will also lead to significant reductions in anxiety and depression scores and improvements in sleep quality from baseline to the posttreatment time point compared to the waitlist group; and (3) both groups (within-group effects) will experience significant improvements in all measures (insomnia, nightmares, PTSD symptoms, anxiety, depression, and sleep quality) from baseline to the posttreatment time point, and these improvements will be maintained at the 3-month follow-up after receiving Sleep Best-i.

Methods

Study Design

This study used a parallel-arm, sequential alternation method of randomizing participants to either the intervention or waitlist group in a 50:50 ratio, with a crossover of the latter group. The alternation method ensures the ability to answer questions about the effectiveness of treatments to inform decision-making in clinical practice [32]. The treatment group completed self-report assessments at baseline, at the posttreatment time point following the administration of Sleep Best-i, and at the 3-month follow-up. The waitlist group completed the same assessments at baseline, at the pretreatment time point (following 4 weeks of waiting and before crossing over to treatment), at the posttreatment time point following the administration of Sleep Best-i, and at the 3-month follow-up. The CONSORT (Consolidated Standards of Reporting Trials) checklist was followed [33]. Please refer to Appendix 1 for the CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) checklist (Multimedia Appendix 1).

Ethical Considerations

Following ethics approval from the Federation University Human Research Ethics Committee (approval 2022-153), an advertisement about the study was distributed. Survivors of wildfires interested in the study read a plain-language information statement (PLIS) about the study and provided baseline data. Participants who met the selection criteria provided informed consent by checking a box on a digital consent form hosted on Federation University's HealthZone platform, thereby indicating their agreement to take part in the study. Participants were informed that their participation was voluntary. To ensure the secure storage and handling of collected information, data were stored on a password-protected digital health platform (HealthZone) during the trial accessible only to the researchers involved in the study. To protect participants' confidentiality, names were replaced with unique identifiers, and only country of residence was collected, with no addresses recorded. The data will be retained for 15 years before being destroyed. To manage any potential risks, participants were provided with emergency contact numbers relevant to their country of residence. As a gesture of appreciation, participants who completed the study were offered an Aus $100 e-voucher (approximately US $66.70).

Participants

Participants were recruited between May 2023 and December 2023 via social media platforms, including international Facebook campaigns, radio broadcasts, LinkedIn, Reddit, and online community noticeboards. This study was also advertised on the Natural Hazards Research Australia and the Australian Institute for Disaster Resilience websites.

Adult survivors of wildfires from Australia, Canada, and the United States aged ≥18 years were invited to take part in the clinical pilot trial if they had experienced wildfires at any point in the past and had sleep or trauma symptoms.

An a priori power analysis indicated that a sample size of 50 was required to detect meaningful differences between the intervention and waitlist groups considering an α value of .05, a value of *d* of 1.0, an attrition rate of 20%, and 80% power [34].

Inclusion Criteria

To be eligible to take part in the trial, participants had to (1) provide digital consent; (2) have a score of ≥8 on the Insomnia Severity Index (ISI), (3) a score of ≥3 on the Nightmare Disorder

XSL•FO
RenderX

Index (NDI), or (4) a score of ≥31 on the PTSD Checklist for DSM-5 (PCL-5); (5) be fluent in the English language; and (6) have access to the internet. Exclusion criteria included not being a survivor of a wildfire, diagnosis of a psychotic disorder, suicide risk, diagnosis of sleep apnea or restless leg syndrome, diagnosis of alcohol or drug dependence, and attendance to psychotherapy for either sleep or PTSD conditions. In addition, participants were excluded if they were currently using steroids for any health condition; medications that promote sleep, such as benzodiazepines (eg, alprazolam, clonazepam, and temazepam); or medications that may affect sleep, such as opiates or other pain medications. By excluding these individuals, this study aimed to ensure that participants were suitable for the trial and that the results would not be confounded by other factors.

Procedure

The trial was registered prospectively in the Australian New Zealand Clinical Trials Registry (ACTRN12623000415606). Since its initiation, this study's protocol has undergone only minor revisions, specifically an update to the statistical analysis methodology to include linear mixed methods and intention-to-treat (ITT) analysis.

An advertisement of this pilot study with a URL was distributed on social media platforms displaying Federation University and the Natural Hazards Research Australia affiliations. This study was also promoted in national conferences, sleep workshops delivered to communities affected by wildfires, the Red Cross, and group meetings for people with sleep disorders in Australia. Radio interviews were also undertaken to promote the study. Survivors of wildfires who were interested read a PLIS (refer to Multimedia Appendix 2 for PLIS); signed a digital consent form (refer to Multimedia Appendix 3 for the consent form) to register an account on Federation University's digital HealthZone platform; and provided baseline data by completing a demographic questionnaire as well as self-report measures on insomnia, nightmares, and trauma symptoms. Participants had the option of providing a pseudonym if they wished. However, they needed to provide a valid email address to allow for communication about the study. Their email addresses were also directly connected to their personal dashboard on HealthZone. The baseline data were viewed by 2 researchers to decide eligibility. Those who were eligible (assessed by FI and GAK, both clinical psychologists) were contacted via email to notify them of their eligibility and their randomization and receive instructions on how to access their personal dashboard on HealthZone. Participants were sequentially allocated via simple randomization in the order of their enrollment (using a computer-generated simple randomization sequence), and they were informed about the purpose, allocation, and structure of the study as explained in the PLIS. However, they were not informed about the study's hypotheses. Once participants gained access to their personal dashboard within HealthZone, they were able to access information about the trial, complete the self-report secondary measures, and access the treatment modules consecutively. The treatment modules were provided at no cost, allowing participants to use their own devices and access the program and its modules on various platforms, including mobile devices, tablets, and desktop computers. HealthZone features a responsive web design compatible with both iOS and Android operating systems, making it easily accessible. Participant response data were captured on all devices. The modules were released sequentially to participants over 4 weeks. All data were collected through self-report measures at prespecified intervals. Participants were instructed to complete a module or a set of 2 modules each week along with the required interval assessments. Automated email reminders alerted participants about the release of each module and the specified assessments. If the modules were not viewed within the first 3 days of their release, an automated email reminder was sent to participants, and a second reminder was sent on the seventh day of their release. Automated email reminders at a similar frequency were also used for the primary and secondary measures (self-report scales). The personal dashboard was accessible to participants at any time. A "Contact us" tab was available on the personal dashboard that allowed participants to contact the research team and ask questions about the study or express any concerns. A total of 7% (2/30) of the participants were unsure about how to access the modules, and 3% (1/30) had an inquiry about one of the modules.

The research team sent 2 follow-up emails spaced 2 weeks apart to participants who registered, were randomized, but did not commence treatment, serving as a reminder about the study and to gauge their ongoing interest in participating.

The modules were available to participants from the start of the trial to the end of the follow-up period. Participants were able to access their dashboard for 4 weeks following the 3-month follow-up data collection. HealthZone recorded data about the number of log-ins, number of pages visited, and date of commencement and duration of participation for each participant. Each module was 17 minutes in duration. Therefore, participants were expected to spend at least 3 hours over the 4 weeks to indicate adherence to treatment. This time is optimal for viewing all modules and completing all assessments. No statistical analyses were conducted in the interim of the study. HealthZone was monitored daily during the trial to track any potential concerns or issues associated with the site or the participants.

Treatment Protocol

Sleep Best-i Intervention

Sleep Best-i was specifically designed for this clinical trial by some of the researchers who conducted this study (FI, BK, and GAK). The collated treatment manual draws from evidence-based treatment manuals authored by other sleep researchers [19,21,35-39]. The main therapeutic methods used in the manual were CBTi and ERRT. The recorded digital modules feature human-recorded voice and animated videos using VideoScribe (Sparkol) [40] that explain concepts related to sleep and trauma symptoms. The intervention also offers 2 role-plays of therapeutic sessions to demonstrate the administration of cognitive restructuring, dream rescripting, and sleep scheduling. Sleep Best-i consists of 6 modules administered over a 4-week period. In addition, the program is accessible via a platform that uses a responsive web design, with no requirement for a specific app or download. This allows participants to access the program from a variety of devices

XSL•FO
RenderX

(including mobile phones), as well as their preferred web browser. By being web based, the program can be easily accessed from anywhere at any time, making it a convenient and flexible intervention.

Participants were encouraged to complete the entire Sleep Best-i program, but completion of each module was not mandatory. This allowed individuals to engage with relevant content tailored to their specific needs, whether they experienced nightmares, PTSD, insomnia, or a combination of these conditions. Furthermore, participants had the flexibility to stop and start the program as desired, and all modules were available for the entire study period, enabling them to revisit specific modules should they wish to do so.

Table 1 shows the therapeutic modality offered in each module. Once released, the modules remained unchanged throughout the trial, and the intervention was successfully implemented without encountering technical difficulties. There were periodical checks to ensure that the modules were compliant with the treatment manual. All links have been archived in the Wayback Machine (Internet Archive).

Table 1. Sleep Best-i modules and the treatment offered in each module. All modules were approximately 17 minutes in duration.

Module	Strategies offered
Module 1 (psychoeducation)	Offered at week 1, the module provided psychoeducation about sleep and insomnia, the stages of sleep, and the neurobiology of sleep [14].
Module 2—part 1 (cognitive restructuring and sleep hygiene)	Offered at week 1, the module focused on the cognitive component of CBTi[a], types of unhelpful thoughts, methods to challenge unhelpful thoughts, and sleep hygiene [30].
Module 2—part 2 (sleep scheduling and stimulus control)	Offered at week 2, the module educated participants about specifying regular sleeping and waking up times with as little variation as possible between the 2. Stimulus control restricted the bedroom or the bed to sleep only [34].
Module 3 (trauma, PTSD[b], and flashbacks)	Offered at week 3, the module provided psychoeducation about trauma, PTSD, flashbacks, and how trauma leads to sleep difficulties. It also provided behavioral interventions for trauma symptoms.
Module 4 (nightmares)	Offered at week 4, the module explored nightmare disorder, how nightmares develop, and how to rescript the nightmare into a more neutral dream [16].
Module 5 (relapse prevention)	Offered at week 4, the module focused on identifying early warning signs and methods preventing relapse of symptoms [33].
Mindfulness module	This module offered a recorded progressive muscle relaxation mindfulness, and it was available to participants throughout the study.

[a]CBTi: cognitive behavioral therapy for insomnia.

[b]PTSD: posttraumatic stress disorder.

Waitlist Control Group

Following a waiting period of 4 weeks, the waitlist control group received Sleep Best-i in the same sequence in which it was released to the treatment group.

Measures

Overview

All scales were digital, self-report measures that were administered through HealthZone. The self-report measures as well as the modules were tested by 2 research team members (FI and BK) before being released to participants. The order of the modules and the questions were the same for all participants in the 2 groups.

Demographic Data

Participants provided the following information: name, email address, age, sex, employment status, educational level, marital status, country of residence, experience with wildfires, history of taking steroids, diagnosis of sleep apnea or restless leg syndrome, diagnosis of a psychotic disorder, use of alcohol or drug dependence, type of medications used to assist with sleeping, history of antidepressant use, and whether they were attending psychotherapy for sleep or PTSD.

Primary Measures

ISI Measure

The ISI [41] includes 7 items that assess insomnia severity through the subjective experience of sleep. The score varies from 0 to 28, with higher scores indicating more severe insomnia symptoms. A score of ≥8 indicates the presence of a subthreshold insomnia. Previous studies have demonstrated strong internal consistency for the ISI, with Cronbach α values ranging from 0.87 to 0.92, indicating high reliability [42].

NDI Measure

This scale consists of 5 items that assess symptoms of nightmares according to the DSM-5 criteria by screening for the occurrence of nightmares in a given week. The 5 items are summed to obtain a score between 0 and 20, with higher scores representing greater nightmare severity. The NDI exhibits good internal consistency, with a Cronbach α of 0.80 [43].

PCL-5 Measure

The PCL-5 [44] includes 20 self-report items assessing the presence of PTSD symptoms over a 1-month interval. A global score that ranges between 0 and 80 is obtained by summing the 20 items. A cutoff score of 31 provides optimal sensitivity and specificity in providing a provisional diagnosis of PTSD [45].

XSL•FO
RenderX

The PCL-5 has good psychometric properties, with an internal consistency of Cronbach α=0.94 [46].

Secondary Measures

Generalized Anxiety Disorder–7

This scale is a 7-item, self-report measure that assesses anxiety symptoms over a 2-week interval. A score from 0 to 21 is obtained by summing the 7 items. The Generalized Anxiety Disorder–7 (GAD-7) [47] has a good internal consistency, with a Cronbach α ranging between 0.82 and 0.93 [48]. A cutoff score of 10 discriminates between mild and severe symptoms of anxiety [48].

Patient Health Questionnaire–9

This scale consists of 9 items assessing symptoms of depression over the last 2 weeks. The overall score is obtained by summing all the 9 items and ranges from 0 to 27. A cutoff score of 10 is considered optimal in discriminating between mild and severe symptoms of depression [49]. In this study, the Cronbach α for the Patient Health Questionnaire–9 (PHQ-9) was 0.86.

Pittsburgh Sleep Quality Index

The Pittsburgh Sleep Quality Index (PSQI) [50] consists of 19 items assessing sleep quality and disturbances related to sleep over a 1-month interval. A global score ranging between 0 and 21 is generated by summing 7 components of sleep quality. A global PSQI score of >5 is considered a sensitive and specific cutoff to discriminate between good and poor sleep quality [51]. The PSQI has a good internal validity, with the Cronbach α ranging between 0.79 and 0.81 [51].

Satisfaction and Level of Engagement With Treatment

One question was designed for this study to measure satisfaction with Sleep Best-i. Participants were asked to rate how likely they were to revisit the Sleep Best-i modules on a 5-point Likert scale from *strongly disagree* to *strongly agree* (0-4). To assess level of engagement with treatment, the number of log-ins and amount of time that each participant spent on the site were monitored.

Data Analysis

Data analyses were conducted using the Stata statistical software package (version 18; StataCorp) [52]. All variables were inspected for normality and outliers. An inspection of histograms and *Q-Q* plots revealed normally distributed data. Upon inspection of box plots on all scales, 1 outlier was identified on the PCL-5 at the 3-month follow-up assessment, a second outlier was detected on the PHQ-9 at the 3-month follow-up, and a third outlier was detected on the PSQI at the waitlist assessment (3-month follow-up assessment). The results did not differ upon removing the outliers from the analyses; therefore, the outliers were retained [53,54]. To ensure a robust evaluation of the intervention's feasibility, the ITT analysis with 30 participants and the per protocol (PP) analysis with a sample of 20 were used, allowing for a comprehensive assessment of treatment outcomes, thereby enhancing the accuracy and generalizability of the findings [33,55,56].

For both analyses, missing data were addressed using multiple imputation by chained equations. This approach is more robust than other methods such as single imputation because it generates multiple predictions for each missing value, accounting for uncertainty in the imputations and yielding more accurate SEs [57]. To impute missing data, background variables such as age, sex, educational level, employment status, country, and relationship status; baseline measures; and the most recent observations of each outcome were used.

To address uncertainty, each missing value was imputed 20 times using predictive mean matching based on the 3 nearest values across multiple iterations to create 100 imputed datasets. The imputed values were then averaged to create the final estimates [58]. Predictive mean matching, a partially parametric method, is preferable to fully parametric linear regression as it remains effective even when the normality assumption of the underlying variable is violated. This method also helps preserve the distribution of observed values in the missing data [59]. Missing value analysis indicated that the data were missing completely at random (Little missing completely at random test: $\chi^2_{867}=0.0$, $P>.99$), meaning that the differences between the missing and observed data were related to observed characteristics [60].

Adjusted analyses of intervention effects were conducted using mixed-effects linear regression models. Both fixed and random effects were estimated to assess the impact of the intervention on the change in all outcomes (primary and secondary) over time and by condition. Given that observed covariates were balanced between the 2 conditions (waitlist and treatment), the difference-in-differences (DID) effect for the fixed part of the mixed-effects models was tested by including interaction terms for time point and condition. DID is considered the most appropriate measure to assess causal effects in time-series designs as it effectively controls for time-invariant confounding variables. By doing so, DID reduces bias and enhances internal validity, providing a more accurate estimate of the treatment effect [61].

The appropriateness of the mixed-effects model for each outcome was assessed using chi-square statistics ($P<.05$). Random-intercept linear models were compared with random-intercept, random-slope models using likelihood ratio tests. The random-intercept, random-slope model was reported only if the likelihood ratio test indicated $P<.05$; otherwise, the simpler random-intercept model was selected. A significant effect of the intervention or difference in outcome between groups was assessed using a 2-tailed test of significance with $P<.05$.

A sensitivity analysis was conducted to test whether the model and covariance structures were misspecified. Several major types of covariance structures, including independent, unstructured, exchangeable, identity, and autoregressive structures, were examined. In addition, the model parameters using both maximum likelihood and restricted maximum likelihood methods were estimated. Although the results were generally consistent across different covariance structures, the model with an exchangeable structure was deemed the best fit.

XSL•FO
RenderX

Clinical Significance

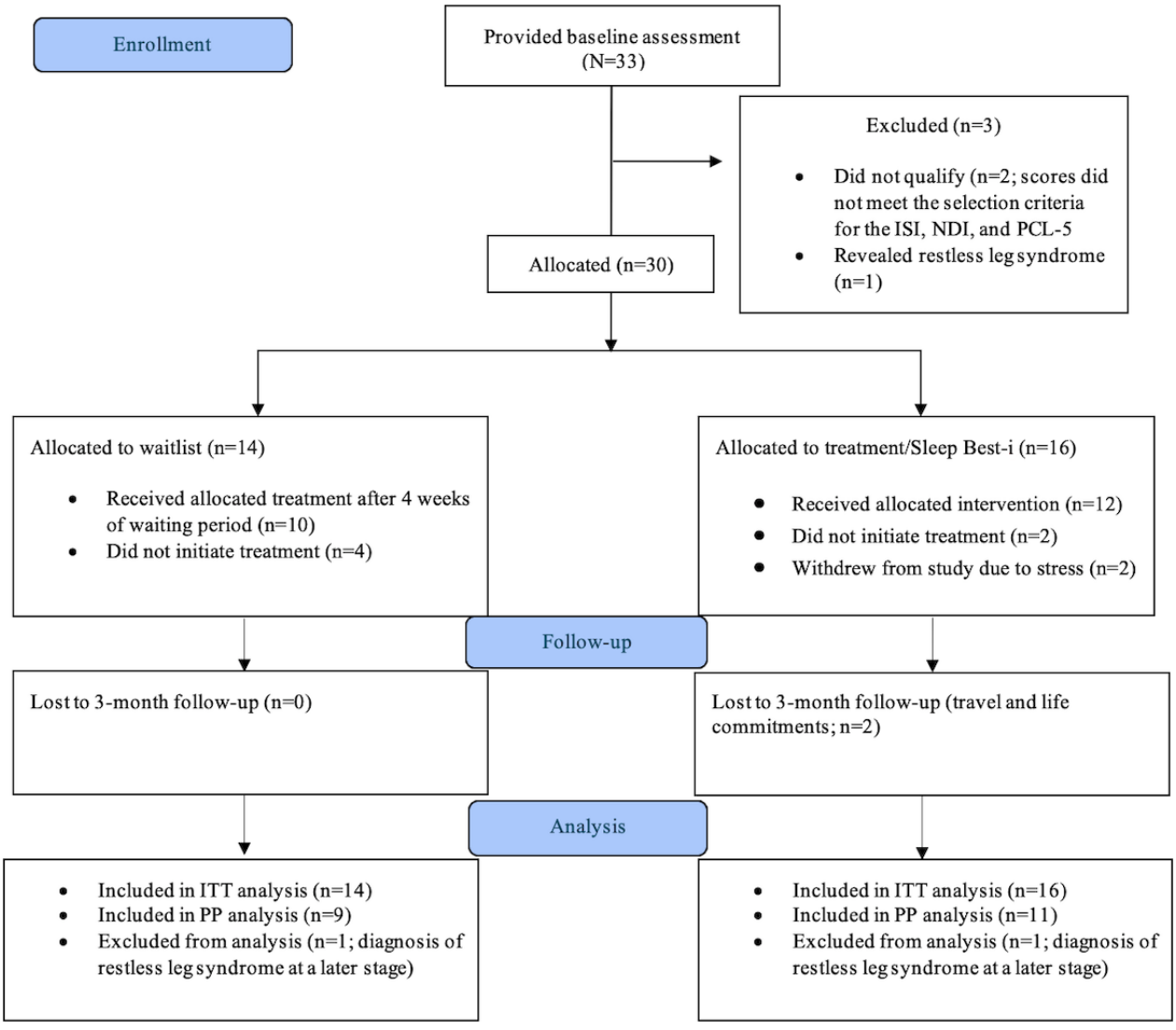
Clinical significance was assessed for the 2 groups from baseline to the 3-month follow-up. To evaluate the clinical significance on sleep and trauma measures, specific criteria were used for each scale. For example, for the ISI, a cutoff score of ≤8 with a ≥6-point reduction in scores was used to assess whether participants reached a minimal clinically significant change (MCSC) [26,62]. For the NDI, there are no established criteria for MCSC; therefore, a cutoff score of ≤7 as suggested by Dietch et al [43] was used. Marx et al [63] recommend a score of ≤28 and a drop of ≥18 points on the PCL-5 as a guide for MCSC.

Results

Overview

A total of 33 participants responded to the advertisement, registered an account with Federation University's digital HealthZone platform, and provided baseline data. In total, 9% (3/33) were excluded because they did not meet the selection criteria, and 91% (30/33) were randomized to either the intervention or the waitlist group. A total of 53% (16/30) were randomized into the intervention group, and 47% (14/30) were in the waitlist group (for the participant CONSORT chart, refer to Figure 1). Ages ranged between 18 and 79 years (mean 52.50, SD 16.26 years; N=30), with most participants (19/30, 63%) being female (vs 11/30, 37% male). A total of 67% (20/30) of the participants completed the trial, and their data were analyzed as PP. The ages of the latter group ranged from 18 to 79 years (mean 53.75, SD 16.54 years), with most participants (14/20, 70%) being female (6/20, 30% were male). Demographic variables for the treatment and waitlist groups are shown in Table 2.

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flowchart showing participant flow through allocation to the treatment and waitlist conditions. ISI: Insomnia Severity Index; ITT: intention to treat; NDI: Nightmare Disorder Index; PCL-5: PTSD Checklist for DSM-5; PP: per protocol.



XSL•FO
RenderX

Table 2. Demographic variables for the treatment and waitlist groups and differences between the 2 at the baseline assessments.

Demographics	Treatment (n=16)	Waitlist (n=14)	Statistic	*P* value[a]
Age (y), mean (SD)	55.56 (15.14)	49.00 (17.34)	t_{28}=1.11	.28
Biological sex, n (%)			χ^2_2=1.4 (N=30)	.51
Female	11 (69)	8 (57)		
Male	5 (31)	6 (43)		
Country, n (%)			—[b]	—
Australia	12 (75)	14 (100)		
Canada	3 (19)	0 (0)		
United States	1 (6)	0 (0)		
Educational level, n (%)			χ^2_3=1.2 (N=30)	.75
Bachelor's degree	8 (50)	6 (43)		
Certificate or diploma	4 (25)	6 (43)		
High school	2 (12)	1 (7)		
Postgraduate	2 (12)	1 (7)		
Employment status, n (%)			χ^2_4=2.6 (N=30)	.64
Employed	7 (44)	5 (36)		
Looking for work	1 (6)	0 (0)		
Retired	5 (31)	5 (36)		
Student	2 (12)	1 (7)		
Unemployed	1 (6)	3 (21)		
Relationship status, n (%)			χ^2_3=8.0 (N=30)	.05
Married	8 (50)	5 (36)		
Divorced or separated	0 (0)	5 (36)		
Single	6 (38)	4 (29)		
Widowed	2 (12)	0 (0)		
Wildfires experienced, n (%)			—	—
2019-2020 wildfires (Australia)	8 (50)	11 (79)		
Waroona 2016 (Australia)	1 (6)	0 (0)		
Oregon 2020 (United States)	1 (6)	0 (0)		
2023 Canadian wildfires	1 (6)	0 (0)		
2021 BC fires	2 (12)	0 (0)		
Black Saturday 2009 (Australia)	2 (12)	1 (7)		
Currowan 2021 (Australia)	1 (6)	0 (0)		
Sacramento 2014 (United States)	0 (0)	1 (7)		
Blue Mountains 2013 (Australia)	0 (0)	1 (7)		

[a]Significance level at $P<.05$.

[b]Not applicable.

A total of 4 participants were excluded from the PP analysis for the following reasons: 2 (50%) participants, who completed the pilot trial and provided a complete set of data, revealed that they had restless leg syndrome that was not initially reported during the screening process; 1 (25%) participant received 1 treatment module but withdrew from the study due to illness, providing only baseline measures; and a fourth participant (25%) who provided and completed 50% of the data and modules withdrew due to illness and family commitments. Therefore, the PP analysis was conducted on a sample of 20 participants.

XSL•FO
RenderX

In total, 20% (6/30) of the participants provided baseline data; however, they did not initiate treatment and did not respond to emails sent by the research team. Table 1 indicates that there were no significant differences in demographic variables between the intervention and waitlist groups at the baseline assessment (age, *P*=.28; biological sex, *P*=.51; education level, *P*=.75; employment status, *P*=.64; and relationship status, *P*=.05).

To address hypothesis 1, ITT analysis of the mixed-effects regression, including the effect sizes of the DID in the fixed-effects component and the variability in the random-effects component, showed that, at baseline, the waitlist group had mean scores of 3.16, 0.39, and 45.29 on the ISI, NDI, and PCL-5, respectively. At the postintervention time point, the main effect of time, these scores significantly decreased by 1.64 points on the NDI (*P*=.001) and 10.64 points on the PCL-5 (*P*=.009), but no significant (*P*=.06) change was detected on the ISI. The treatment group showed slightly lower baseline scores than the waitlist group, with mean differences of 0.45, 0.04, and 9.20 for the ISI, NDI, and PCL-5, respectively, suggesting that the 2 groups were relatively balanced at baseline. The adjusted DID effects following the treatment indicated that the treatment group experienced greater improvements in mental health measures compared to the waitlist group, with differences of 1.54 points for the ISI, 1.22 points for the NDI, and 4.98 points for the PCL-5. Notably, the improvement in the NDI was marginally significant (*P*=.06), whereas the improvements in the ISI and PCL-5 were not statistically significant (*P*=.26 and *P*=.37, respectively; Table 3).

Table 3. Mixed-effects linear regression analysis of the effect of the intervention on primary outcomes (intention-to-treat analysis; N=30)[a].

Variable	Model 1 (n=60[b]; ISI[c])		Model 2 (n=60[b]; NDI[d])		Model 3 (n=60[b]; PCL-5[e])	
	β (95% CI)	*P* value[f]	β (95% CI)	*P* value[g]	β (95% CI)	*P* value[h]
Fixed effects						
Intercept	3.16 (−0.77 to 7.08)	.12	0.39 (−1.05 to 1.83)	.60	45.29 (37.27 to 53.30)	<.001
Time						
Baseline	Reference	—[i]	Reference	—	Reference	—
Postintervention time point	−1.88 (−3.86 to 0.10)	.06	−1.64 (−2.58 to −0.71)	.001	−10.64 (−18.65 to −2.64)	.009
Condition						
Waitlist	Reference	—	Reference	—	Reference	—
Intervention	−0.45 (−2.56 to 1.66)	.68	−0.04 (−1.04 to 0.97)	.95	−9.20 (−20.17 to 1.78)	.10
Time × condition	−1.54 (−4.26 to 1.17)	.26	−1.22 (−2.50 to 0.06)	.06	−4.98 (−15.94 to 5.98)	.37
Random effects						
SD (time, intercept)	1.41 (0.82 to 2.42)	—	0.82 (0.55 to 1.20)	—	10.82 (7.26 to 16.14)	—
Correlation (time, intercept)	0.99 (−1.00 to 1.00)	—	1.00 (−1.00 to 1.00)	—	N/A[j]	—
SD of residual	2.48 (1.86 to 3.31)	—	1.23 (0.86 to 1.47)	—	10.81 (8.39 to 13.92)	—

[a]All models were adjusted for the baseline measure to address the regression-to-mean issue except for model 3, where the baseline measure was excluded for simplicity and convergence.

[b]*n* represents the number of observations in the data with long format to support mixed-effects modeling.

[c]ISI: Insomnia Severity Index.

[d]NDI: Nightmare Disorder Index.

[e]PCL-5: PTSD Checklist for DSM-5.

[f]Model fit—*P*<.001 (Wald chi-square test) and *P*=.03 (likelihood ratio test [LRT]) to compare the mixed-effects models with random-intercept linear models.

[g]Model fit—*P*<.001 (Wald chi-square test) and *P*=.001 (LRT) to compare the mixed-effects models with random-intercept linear models.

[h]Model fit—*P*<.001 (Wald chi-square test) and *P*=.19 (LRT) to compare the mixed-effects models with random-intercept linear models.

[i]Not available.

[j]N/A: not applicable as the random-intercept model better fit the data than the random-intercept, random-slope model. Given its better fit, random-intercept model parameters were reported.

In the random-effects part, there was substantial variability among individuals, with average deviations of 1.41, 0.82, and 10.82 points for the ISI, NDI, and PCL-5, respectively, around both the baseline and postintervention means. This suggests considerable individual variability in the outcomes. The positive correlation coefficients between the time point and the intercept indicate that participants with higher baseline scores (above the overall sample mean) were more likely to experience greater decreases in their scores over time compared to those with lower baseline scores. All mixed-effects models for the ISI, NDI, and PCL-5 were statistically significant (*P*<.001). Specifically, the random-intercept, random-slope model was appropriate for the

XSL•FO
RenderX

ISI (*P*=.03) and NDI (*P*=.001), whereas a random intercept–only model was sufficient for the PCL-5, which also fit the data but with a nonsignificant *P* value (*P*>.05 and *P*=.19; Table 3). Figure 2 illustrates the changes in primary outcome measures (ISI, NDI, and PCL-5) from before to after treatment based on the ITT analysis.

For completers only, PP analysis of the mixed-effects regression showed that, at baseline, the waitlist group had mean scores of 5.58, 0.46, and 13.37 on the ISI, NDI, and PCL-5, respectively. At the postintervention time point, these scores did not significantly change for the 3 measures (ISI; *P*=.68, NDI; *P*=.24, PCL-5; and *P*=.35). The treatment group baseline scores were balanced compared to those of the waitlist group, with mean differences of 0.99, 0.04, and 3.05 for the ISI, NDI, and PCL-5, respectively. However, the adjusted DID effects indicated an interaction effect of condition × time, with the treatment group experiencing greater and more significant improvements on the NDI with a decrease of 2.27 points (*P*=.049) and a decrease of 13.46 points on the PCL-5 (*P*=.03) but no significant improvements on the ISI (*P*=.25; Table 4).

There was a substantial variability among individuals, with average deviations of 3.41, 1.81, and 9.42 points for the ISI, NDI, and PCL-5, respectively, at the postintervention time point, suggesting considerable individual variability in the outcomes. All mixed-effects models for the ISI, NDI, and PCL-5 were statistically significant (*P*<.001), with a random-intercept model being appropriate for the 3 measures (Table 4).

In relation to the second hypothesis, the ITT analysis showed that, at baseline, the waitlist group had mean scores of 1.66, 12.01, and 0.69 on the GAD-7, PHQ-9, and PSQI, respectively (Table 5). Only the depression scores decreased significantly (*P*<.001) by 6.48 points at the posttreatment time point (main effect of time) in comparison to the waitlist group. The treatment group showed slightly lower baseline scores than those of the waitlist group on the secondary measures, with mean differences of 0.47, 0.90, and 0.04 on the GAD-7, PHQ-9, and PSQI, respectively, suggesting that the 2 groups were balanced at baseline. The adjusted effects at the postintervention time point indicate that the treatment group experienced greater improvements than the waitlist group on all measures but with only significant results on the PSQI (with a difference of 2.22 points; *P*<.001).

There was variability among individuals, as shown in Table 5, with average deviations of 0.91, 4.89, and 0.73 points for the GAD-7, PHQ-9, and PSQI, respectively, around both the baseline and postintervention means. The positive correlation coefficients for the GAD-7 and PSQI between the time point and the intercept indicate that participants with higher baseline scores (above the overall sample mean) were more likely to experience greater decrease in their scores over time compared to those with lower baseline scores. Mixed-effects models for the 3 measures were statistically significant (*P*<.001). Specifically, the random-intercept, random-slope model was appropriate for the PHQ-9 and PSQI (*P*=.007 and *P*=.001, respectively), whereas a random intercept–only model was sufficient for the GAD-7 (*P*=.08). Figure 3 shows the difference in decrease in scores between the intervention and waitlist groups for the GAD-7, PHQ-9, and PSQI from the ITT analysis.

For those who completed the study, PP analysis showed that the waitlist mean baseline scores were 11.78, 12.22, and 13.00 on the GAD-7, PHQ-9, and PSQI, respectively. These scores decreased significantly only for the PHQ-9 measure by 6.22 points at the posttreatment time point (*P*<.001). There was also a main effect of condition on the GAD-7 measure with a decrease of 4.14 points (*P*=.04) for the treatment group and an interaction effect of time × condition on the PSQI (*P*=.02), suggesting that the treatment group experienced significantly better sleep quality than the waitlist group following the treatment (Table 6).

Table 6 also shows a great variability among individuals on the random effects, with average deviations of 3.69, 3.16, and 2.82 points for the GAD-7, PHQ-9, and PSQI, respectively, around both the baseline and postintervention means. This suggests substantial individual variability in the outcome measures. Mixed-effects models for the 3 measures were statistically significant (*P*<.001), with the random-intercept, random-slope model being appropriate for the PHQ-9 (*P*=.004), whereas a random intercept–only model was sufficient for the GAD-7 and the PSQI (*P*=.28 and *P*>.99, respectively; Table 6).

To address the third hypothesis, separate mixed-effects linear model analyses were conducted for each group (intervention and waitlist) on all primary and secondary outcome measures. The analysis incorporated data from the following time points: baseline, posttreatment time point, and 3-month follow-up for the intervention group and pretreatment time point, posttreatment time point, and 3-month follow-up for the waitlist group. ITT analysis for the intervention group showed a significant reduction in insomnia, nightmares, and trauma symptoms as measured using the ISI, NDI, and PCL-5 both at the posttreatment time point (*P*<.001) with an effect of 3.42, 2.86, and 15.62 points, respectively, and at the 3-month follow-up (*P*<.001) with an effect of 8.64, 4.20, and 22.83 points, respectively (Table 7).

A great variability among individuals was detected on the random effects, with average deviations of 1.56, 0.32, and 11.09 points for the ISI, NDI, and PCL-5, respectively, around both the postintervention and 3-month follow-up means. Mixed-effects models for the 3 measures were statistically significant (*P*<.001), with the random-intercept, random-slope model being appropriate for the ISI and the PCL-5 (*P*=007 and *P*=.02, respectively), whereas the random-intercept model was appropriate for the NDI (*P*=.16; Table 7). Figure 4 shows the decrease in scores over time at the postintervention time point and 3-month follow-up for the intervention group.

For those who completed the study, the PP analysis showed comparable results to those reported in the ITT analysis, with a significant reduction in symptoms at both the postintervention time point (ISI; *P*=.04, NDI; *P*<.001, PCL-5; *P*<.001) and 3-month follow-up (*P*<.001 in all cases). The adjusted effects showed a continued reduction in symptoms from the postintervention time point to the 3-month follow-up, with a difference of 7.81, 4.86, and 23.71 points on the ISI, NDI, and PCL-5, respectively (Table 8).

XSL•FO
RenderX

The analysis also showed a substantial variability among individuals, with average deviations of 1.64, 1.03, and 1.81 points for the ISI, NDI, and PCL-5, respectively, around both the baseline and postintervention means. The positive correlation coefficients between the time point and the intercept on the NDI indicate that participants with higher baseline scores (above the overall sample mean) were more likely to experience greater decreases in their scores over time compared to those with lower baseline scores. All mixed-effects models for the ISI, NDI, and PCL-5 were statistically significant (P<.001), with a random-intercept model being sufficient for all 3 measures with a nonsignificant P value (ISI, P=.18; NDI, P=.07; and PCL-5, P=.42, Table 8).

On the secondary measures, the ITT analysis showed that the intervention group experienced significant improvements on the PHQ-9 and PSQI at both the posttreatment time point and the 3-month follow-up (P<.001). No significant reduction in symptoms was observed for the GAD-7 at the posttreatment time point (P=.21; Table 9). However, all participants in the intervention group experienced a significant reduction in symptoms at the 3-month follow-up (P<.001) on the 3 measures, with adjusted effects showing a decrease of 4.05, 7.50, and 5.12 points on the GAD-7, PHQ-9, and PSQI, respectively. Table 9 also shows variability among individuals on the random effects, with average deviations of 0.001, 5.46, and 0.82 points for the GAD-7, PHQ-9, and PSQI, respectively, around the postintervention and 3-month means. The 3 models were statistically significant (P<.001), with the random-intercept, random-slope model being appropriate for the PHQ-9 and PSQI (P=.03 and P=.002, respectively) and the random-intercept model being appropriate for the GAD-7 (P=.34). Figure 5 shows a change in secondary outcomes over time in the intervention group (ITT analysis).

Comparable results were also detected on the PP analysis, with individuals experiencing significant reductions in symptoms following the intervention and also at the 3-month assessments (P<.001 in all cases) except for the GAD-7, which was found to be not significant at the postintervention time point (P=.29). Participants experienced a greater reduction in symptoms at the 3-month assessment, with adjusted effects indicating a reduction of 3.88, 6.98, and 4.62 points on the GAD-7, PHQ-9, and PSQI, respectively (Table 10). The 3 models were statistically significant (P<.001), with the random-intercept model being appropriate for the 3 measures (GAD-7, P=.32; PHQ-9, P=.56, and PSQI, P=.05).

Similarly, ITT analysis results of the mixed-effects regression showed that, at baseline, participants in the waitlist group had mean scores of 10.79, 5.42, and 13.93 on the ISI, NDI, and PCL-5, respectively. The adjusted effects at the postintervention time point indicated that the symptoms significantly reduced by 6.36, 5.78, and 19.37 points, respectively (P<.001), and a significant reduction of 8.21, 7.74, and 27.96 points was observed at the 3-month follow-up for the ISI, NDI, and PCL-5, respectively (P<.001; Table 11).

In the random-effects part, there was substantial variability among individuals, with average deviations of 9.80, 6.18, and 1.73 points for the ISI, NDI, and PCL-5, respectively, around both the postintervention and 3-month means. This suggests considerable individual variability in the outcomes. All mixed-effects models for the ISI, NDI, and PCL-5 were statistically significant (P<.001), with a random-intercept model being appropriate for the 3 measures with a nonsignificant P value (P>.05; ISI, P=.45; NDI, P=.59, PCL-5, P=14 Table 11). Figure 6 shows the changes in primary outcomes over time for the waitlist group at the postintervention time point and 3-month follow-up.

The PP analysis showed similar results, with significant improvements at both the postintervention and 3-month follow-up assessments on the ISI, NDI, and PCL-5 (P<.001 in all cases; Table 12). The mixed-effects models for the ISI, NDI, and PCL-5 were all statistically significant (P<.001). The random-intercept model was used for the 3 measures (P>.05; ISI, P>.99; NDI, P>.99; and PCL-5, P=.34, Table 12).

For the secondary measures, ITT analysis showed that the waitlist group had baseline mean scores of 0.78, 0.36, and 3.31 on the GAD-7, PHQ-9, and PSQI, respectively. The adjusted effects at the postintervention time point indicated that these scores decreased significantly by 4.72 and 5.07 points for the GAD-7 and the PSQI, respectively (P<.001). However, no significant changes were observed on the PHQ-9 at the postintervention time point. Furthermore, adjusted effects showed a significant reduction of 6.51 points on the GAD-7 and 5.00 points on the PSQI (P<.001) against the baseline but no changes on the PHQ-9 at the 3-month follow-up (Table 13).

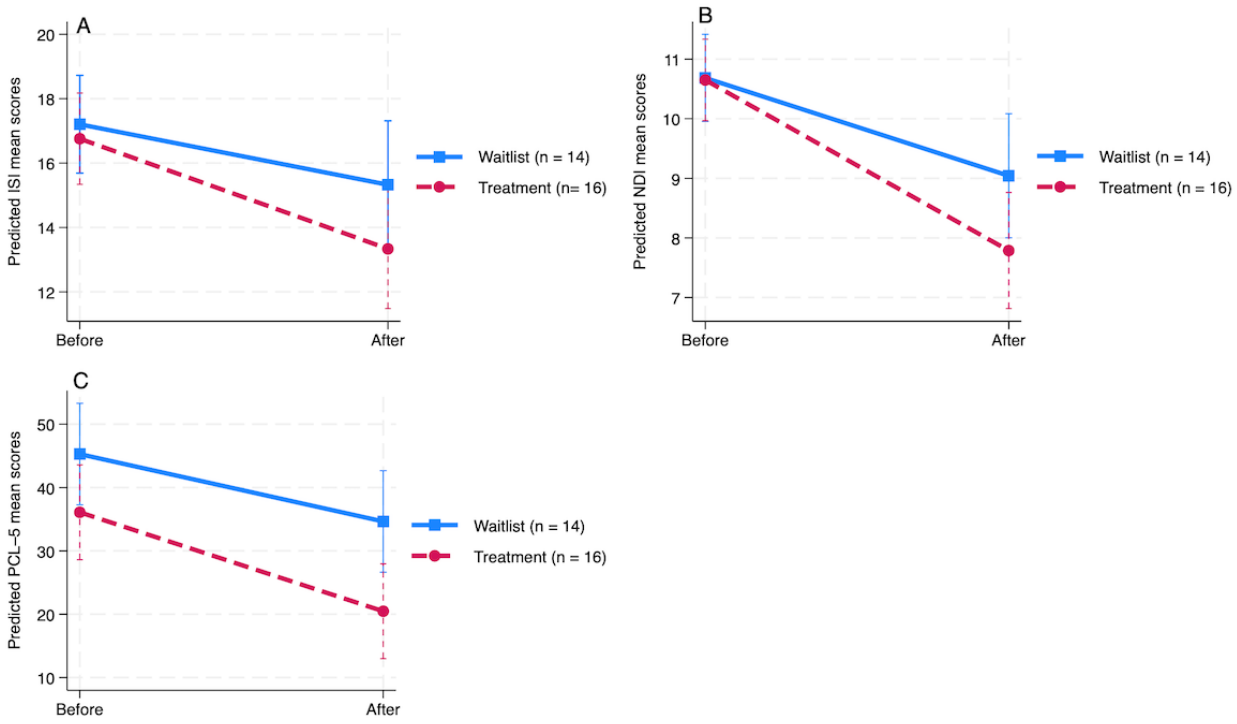
There was variability among individuals, with deviations of 1.47, 0.06, and 0.92 points for the GAD-7, PHQ-9, and PSQI, respectively, around the postintervention and 3-month follow-up means. The positive correlation coefficients between the time point and the intercept indicate that participants with higher baseline scores (above the overall sample mean) were more likely to experience greater decreases in their scores over time compared to those with lower baseline scores. All mixed-effects models for the GAD-7, PHQ-9, and PSQI were statistically significant (P<.001). Specifically, the random-intercept, random-slope model was appropriate for the GAD-7 and PSQI (P<.001 and P=.02, respectively), whereas a random intercept–only model was sufficient for the PHQ-9, which also fit the data but with a nonsignificant P value (P>.05; P=.17, Table 13). Figure 7 shows the changes in secondary outcomes at the postintervention time point and at the 3-month follow-up for the waitlist group.

For those who completed the study, PP analysis showed similar results for the waitlist group, with significant reduction at the postintervention and 3-month assessments only for the GAD-7 and the PSQI (P<.001); however, no significant changes were detected for the PHQ-9 at either the postintervention time point or the 3-month follow-up (P=.61 and P=.92, respectively; Table 14). For those who experienced significant reductions, there was great variability, with deviations of 1.98 and 1.00 points for the GAD-7 and PSQI, respectively, around the postintervention and 3-month follow-up means. All mixed-effects models for the GAD—7, PHQ—9, and PSQI were statistically significant (P<.001). Specifically, the random-intercept model was appropriate for the 3 measures

XSL•FO
RenderX

with a nonsignificant *P* values (*P*>.05; GAD-7, *P*=.05, PHQ-9, *P*=.47, PSQI, *P*=.24, Table 14).

Figure 2. Change in adjusted estimates of primary measures, including the Insomnia Severity Index (ISI); Nightmare Disorder Index (NDI); and PTSD Checklist for DSM-5 (PCL-5), from before to after the intervention based on the intention-to-treat analysis.



XSL•FO
RenderX

Table 4. Mixed-effects linear regression analysis of the effect of the intervention on primary outcomes for completers only (per-protocol analysis; N=20)[a].

Variable	Model 1 (n=40[b]; ISI[c])		Model 2 (n=40[b]; NDI[d])		Model 3 (n=40[b]; PCL-5[e])	
	β (95% CI)	*P* value[f]	β (95% CI)	*P* value[g]	β (95% CI)	*P* value[h]
Fixed effects						
Intercept	5.58 (1.12 to 10.04)	.01	0.46 (−1.70 to 2.61)	.68	13.37 (3.71 to 23.03)	.007
Time						
Baseline	Reference	—[i]	Reference	—	Reference	—
Postintervention time point	−0.67 (−3.82 to 2.48)	.68	−1.00 (−2.67 to 0.67)	.24	−4.11 (−12.81 to 4.59)	.35
Condition						
Waitlist	Reference	—	Reference	—	Reference	—
Intervention	−0.99 (−4.06 to 2.04)	.53	−0.04 (−1.65 to 1.56)	.96	−3.05 (−11.52 to 5.42)	.48
Time × condition	−2.48 (−6.72 to 1.77)	.25	−2.27 (−4.53 to −0.01)	.049	−13.46 (−25.19 to −1.73)	.03
Random effects						
SD (time, intercept)	2.81 (2.60 to 3.22)	—	0.93 (1.31 to 2.00)	—	9.43 (6.67 to 10.33)	—
Correlation (time, intercept)	0.99 (-1.00 to 1.00)	—	1.00 (-1.00 to 1.00)	—	—	—
SD of residual	3.41 (2.74 to 4.24)	—	1.81 (1.46 to 2.26)	—	9.42 (7.57 to 11.73)	—

[a]All models were adjusted for the baseline measure to address the regression-to-mean issue.

[b]*n* represents the number of observations in the data with long format to support mixed-effects modeling.

[c]ISI: Insomnia Severity Index.

[d]NDI: Nightmare Disorder Index.

[e]PCL-5: PTSD Checklist for DSM-5.

[f]Model fit—*P*<.001 (Wald chi-square test) and *P*>.99 (likelihood ratio test [LRT]) to compare the mixed-effects models with random-intercept linear models.

[g]Model fit—*P*<.001 (Wald chi-square test) and *P*>.99 (LRT) to compare the mixed-effects models with random-intercept linear models.

[h]Model fit—*P*<.001 (Wald chi-square test) and *P*>.99 (LRT) to compare the mixed-effects models with random-intercept linear models.

[i]Not available.

XSL•FO
RenderX

Table 5. Mixed-effects linear regression analysis of the effect of the intervention on secondary outcomes (intention-to-treat analysis; N=30)[a].

Variable	Model 1 (n=60[b]; GAD-7[c])		Model 2 (n=60[b]; PHQ-9[d])		Model 3 (n=60[b]; PSQI[e])	
	β (95% CI)	*P* value[f]	β (95% CI)	*P* value[g]	β (95% CI)	*P* value[h]
Fixed effects						
Intercept	1.66 (−0.14 to 3.45)	.07	12.01 (9.46 to 14.57)	<.001	0.69 (−1.33 to 2.71)	.50
Time						
Baseline	Reference	—[i]	Reference	—	Reference	—
Postintervention time point	0.16 (−1.29 to 1.62)	.83	−6.48 (−9.04 to −3.92)	<.001	−0.58 (−1.43 to 0.27)	.18
Condition						
Waitlist	Reference	—	Reference	—	Reference	—
Intervention	−0.47 (−2.01 to 1.07)	.55	−0.90 (−4.40 to 2.61)	.62	−0.04 (−0.96 to 0.87)	.92
Time × condition	−1.62 (−3.62 to 0.38)	.11	1.28 (−2.23 to 4.78)	.48	−2.22 (−3.39 to −1.06)	<.001
Random effects						
SD (time and intercept)	0.91 (0.47 to 1.74)	—	4.89 (3.89 to 6.13)	—	0.73 (0.49 to 1.09)	—
Correlation (time and intercept)	N/A[j]	—	−0.79 (−0.89 to −0.61)	—	0.99 (−1.00 to 1.00)	—
SD of residual	1.86 (1.39 to 2.50)	—	0.000065 (0.000000665 to 0.01)	—	1.03 (0.78 to 1.34)	—

[a]All models were adjusted for the baseline measure to address the regression-to-mean issue except for model 2, where the baseline measure was excluded for simplicity and convergence.

[b]*n* represents the number of observations in the data with long format to support mixed-effects modeling.

[c]GAD-7: Generalized Anxiety Disorder–7.

[d]PHQ-9: Patient Health Questionnaire–9.

[e]PSQI: Pittsburgh Sleep Quality Index.

[f]Model fit—*P*<.001 (Wald chi-square test) and *P*=.08 (likelihood ratio test [LRT]) to compare the mixed-effects models with random-intercept linear models.

[g]Model fit—*P*<.001 (Wald chi-square test) and *P*=.007 (LRT) to compare the mixed-effects models with random-intercept linear models.

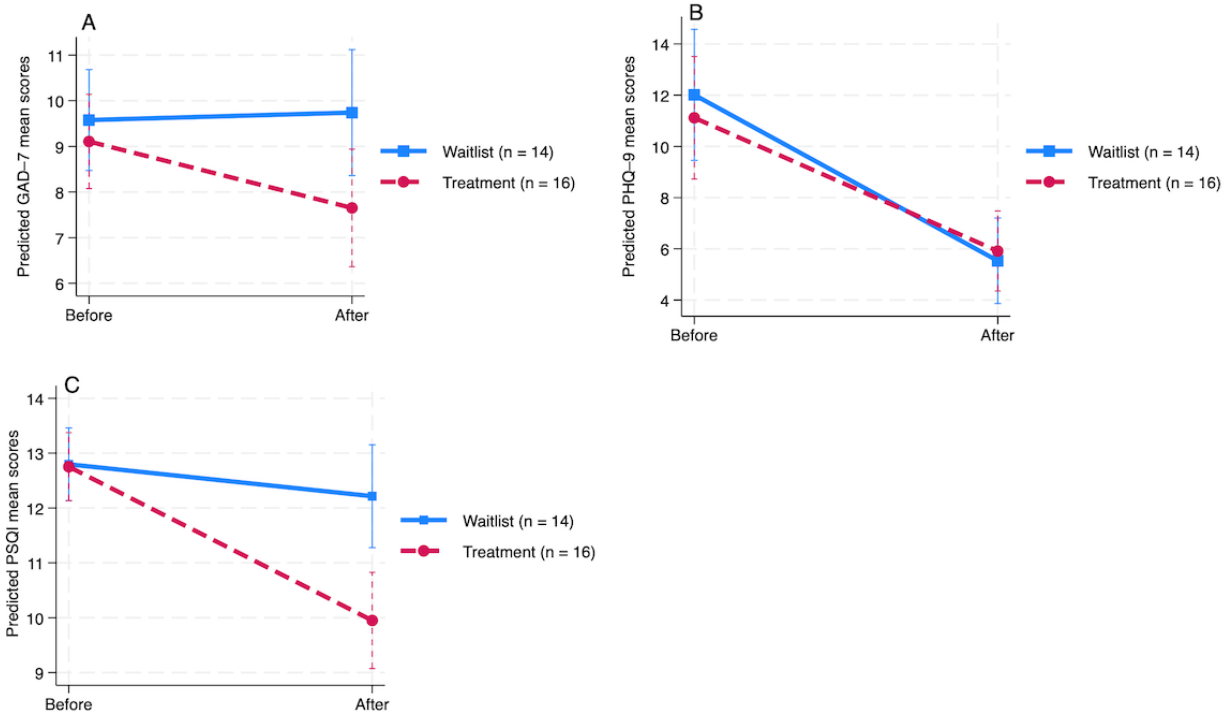
[h]Model fit—*P*<.001 (Wald chi-square test) and *P*=.001 (LRT) to compare the mixed-effects models with random-intercept linear models.

[i]Not available.

[j]N/A: not applicable as the random-intercept model better fit the data than the random-intercept, random-slope model. Given its better fit, random-intercept model parameters were reported.

XSL•FO
RenderX

Figure 3. Change in adjusted estimates of secondary outcomes, including the Generalized Anxiety Disorder–7 (GAD-7), Patient Health Questionnaire–9 (PHQ-9), and Pittsburgh Sleep Quality Index (PSQI), from before to after the intervention based on the intention-to-treat analysis.



XSL•FO
RenderX

Table 6. Mixed-effects linear regression analysis of the effect of the intervention on secondary outcomes for completers only (per-protocol analysis; N=20)[a].

Variable	Model 1 (n=40[b]; GAD-7[c])		Model 2 (n=40[b]; PHQ-9[d])		Model 3 (n=40[b]; PSQI[e])	
	β (95% CI)	*P* value[f]	β (95% CI)	*P* value[g]	β (95% CI)	*P* value[h]
Fixed effects						
Intercept	11.78 (8.89 to 14.66)	<.001	12.22 (9.44 to 15.00)	<.001	13.00 (10.87 to 15.13)	<.001
Time						
Baseline	Reference	—[i]	Reference	—	Reference	—
Postintervention time point	1.33 (−0.90 to 3.57)	.24	−6.22 (−8.86 to −3.59)	<.001	−0.67 (−2.19 to 0.86)	.39
Condition						
Waitlist	Reference	—	Reference	—	Reference	—
Intervention	−4.14 (−8.03 to −0.25)	.04	−1.60 (−5.35 to 2.15)	.40	−1.00 (−3.88 to 1.88)	.50
Time × condition	−2.70 (−5.71 to 0.32)	.08	1.14 (−2.41 to 4.70)	.53	−2.42 (−4.48 to −0.37)	.02
Random effects						
SD (intercept)	3.69 (2.52 to 5.41)	—	3.16 (2.01 to 4.98)	—	2.82 (1.95 to 4.07)	—
Correlation (time, intercept)	N/A[j]	—	−0.68 (−0.85 to −0.37)	—	N/A	—
SD of residual	2.42 (1.78 to 3.30)	—	2.85 (2.09 to 3.89)	—	1.65 (1.21 to 2.25)	—

[a]All models were adjusted for the baseline measure to address the regression-to-mean issue.

[b]*n* represents the number of observations in the data with long format to support mixed-effects modeling.

[c]GAD-7: Generalized Anxiety Disorder–7.

[d]PHQ-9: Patient Health Questionnaire–9.

[e]PSQI: Pittsburgh Sleep Quality Index.

[f]Model fit—*P*<.001 (Wald chi-square test) and *P*=.28 (likelihood ratio test [LRT]) to compare the mixed-effects models with random-intercept linear models.

[g]Model fit—*P*<.001 (Wald chi-square test) and *P*=.004 (LRT) to compare the mixed-effects models with random-intercept linear models.

[h]Model fit—*P*<.001 (Wald chi-square test) and *P*>.99 (LRT) to compare the mixed-effects models with random-intercept linear models.

[i]Not available.

[j]N/A: not applicable as the random-intercept model better fit the data than the random-intercept, random-slope model. Given its better fit, random-intercept model parameters were reported.

XSL•FO
RenderX

Table 7. Mixed-effects linear regression analysis of the change in primary outcomes over time in the intervention group (intention-to-treat analysis; N=16)[a].

Variable	Model 1 (n=48[b]; ISI[c])		Model 2 (n=48[b]; NDI[d])		Model 3 (n=48[b]; PCL-5[e])	
	β (95% CI)	*P* value[f]	β (95% CI)	*P* value[g]	β (95% CI)	*P* value[h]
Fixed effects						
Intercept	3.23 (−1.66 to 8.12)	.20	3.37 (1.12 to 5.61)	.003	−17.52 (−27.73 to −7.30)	.001
Time						
Baseline	Reference	—[i]	Reference	—	Reference	—
Postintervention time point	−3.42 (−5.47 to −1.38)	.001	−2.86 (−4.51 to −1.21)	.001	−15.62 (−21.87 to −9.37)	<.001
3 months after the intervention	−8.64 (−10.68 to −6.60)	<.001	−4.20 (−5.85 to −2.55)	<.001	−22.83 (−29.08 to −16.58)	<.001
Random effects						
SD (time, intercept)	1.56 (0.91 to 2.67)	—	0.32 (0.0000751 to 0.000133)	—	11.09 (7.06 to 17.43)	—
Correlation (time, intercept)	0.99 (−1.00 to 1.00)	—	N/A[j]	—	0.99 (−1.00 to 1.00)	—
SD of residual	2.73 (2.11 to 3.55)	—	2.38 (1.86 to 3.04)	—	4.45 (3.36 to 5.89)	—

[a]All models were adjusted for the baseline measure to address the regression-to-mean issue.

[b]*n* represents the number of observations in the data with long format to support mixed-effects modeling.

[c]ISI: Insomnia Severity Index.

[d]NDI: Nightmare Disorder Index.

[e]PCL-5: PTSD Checklist for DSM-5.

[f]Model fit—*P*<.001 (Wald chi-square test) and *P*=.007 (likelihood ratio test [LRT]) to compare the mixed-effects models with random-intercept linear models.

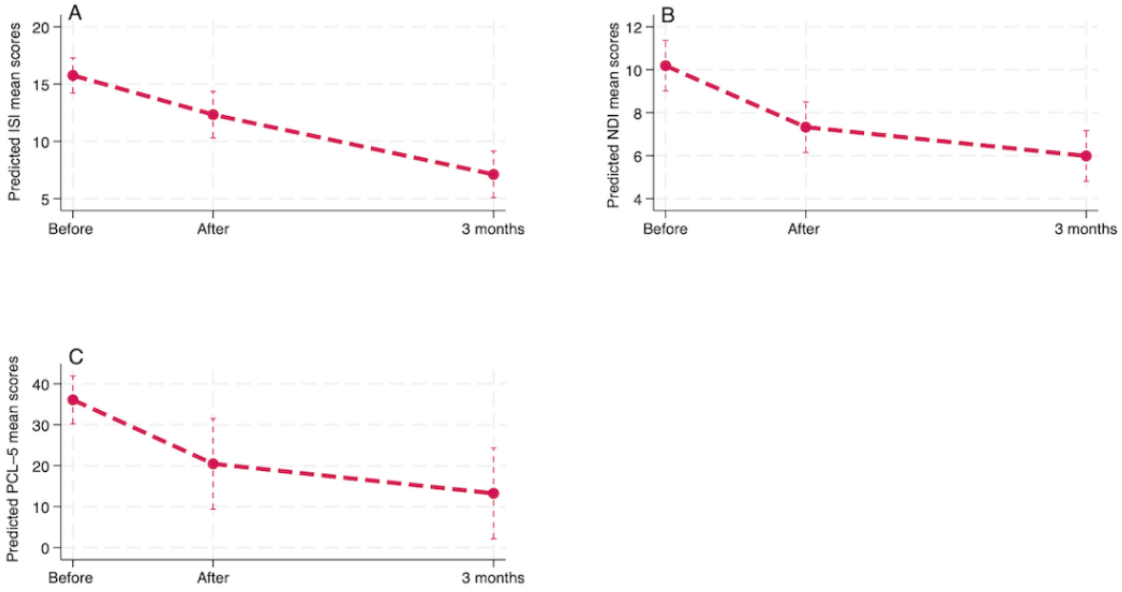
[g]Model fit—*P*<.001 (Wald chi-square test) and *P*=.16 (LRT) to compare the mixed-effects models with random-intercept linear models.

[h]Model fit—*P*<.001 (Wald chi-square test) and *P*=.02 (LRT) to compare the mixed-effects models with random-intercept linear models.

[i]Not available.

[j]N/A: not applicable as the random-intercept model better fit the data than the random-intercept, random-slope model. Given its better fit, random-intercept model parameters were reported.

Figure 4. Change in Insomnia Severity Index (ISI); Nightmare Disorder Index (NDI); and PTSD Checklist for DSM-5 (PCL-5) scores for the intervention group from baseline to the postintervention time point and the 3-month follow-up as determined by the intention-to-treat analysis.



XSL•FO
RenderX

Table 8. Mixed-effects linear regression analysis of the change in primary outcomes over time in the intervention group for completers only (per-protocol analysis; N=11)[a].

Variable	Model 1 (n=33[b]; ISI[c])		Model 2 (n=33[b]; NDI[d])		Model 3 (n=33[b]; PCL-5[e])	
	β (95% CI)	*P* value[f]	β (95% CI)	*P* value[g]	β (95% CI)	*P* value[h]
Fixed effects						
Intercept	6.50 (−1.03 to 11.97)	.002	0.24 (−4.19 to 4.66)	.92	16.05 (7.49 to 24.61)	<.001
Time						
Baseline	Reference	—[i]	Reference	—	Reference	—
Postintervention time point	−3.14 (−6.17 to −0.12)	.04	−3.27 (−4.66 to −1.89)	<.001	−17.57 (−25.45 to −9.69)	<.001
3 months after the intervention	−7.81 (−10.84 to −4.79)	<.001	−4.86 (−6.24 to −3.47)	<.001	−23.71 (−31.59 to −15.83)	<.001
Random effects						
SD (intercept)	1.64 (0.49 to 5.44)	—	1.03 (0.24 to 4.48)	—	1.81 (0.01 to 260.12)	—
SD of residual	3.62 (2.70 to 4.87)	—	1.49 (0.71 to 3.10)	—	9.43 (7.02 to 12.67)	—

[a]All models were adjusted for the baseline measure to address the regression-to-mean issue.

[b]*n* represents the number of observations in the data with long format to support mixed-effects modeling.

[c]ISI: Insomnia Severity Index.

[d]NDI: Nightmare Disorder Index.

[e]PCL-5: PTSD Checklist for DSM-5.

[f]Model fit—*P*<.001 (Wald chi-square test) and *P*=.18 (likelihood ratio test [LRT]) to compare the mixed-effects models with random-intercept linear models.

[g]Model fit—*P*<.001 (Wald chi-square test) and *P*=.07 (LRT) to compare the mixed-effects models with random-intercept linear models.

[h]Model fit—*P*<.001 (Wald chi-square test) and *P*=.42 (LRT) to compare the mixed-effects models with random-intercept linear models.

[i]Not available.

XSL•FO
RenderX

Table 9. Mixed-effects linear regression analysis of the change in secondary outcomes over time in the intervention group—intention-to-treat analysis (N=16)[a].

Variable	Model 1 (n=48[b]; GAD-7[c])		Model 2 (n=48[b]; PHQ-9[d])		Model 3 (n=48[b]; PSQI[e])	
	β (95% CI)	*P* value[f]	β (95% CI)	*P* value[g]	β (95% CI)	*P* value[h]
Fixed effects						
Intercept	3.64 (1.58 to 5.70)	.001	−8.64 (−12.69 to −4.59)	.001	0.92 (−2.45 to 4.29)	.59
Time						
Baseline	Reference	—[i]	Reference	—	Reference	—
Postintervention time point	−1.46 (−3.73 to 0.82)	.21	−5.20 (−7.88 to −2.53)	<.001	−2.80 (−3.82 to −1.78)	<.001
3 months after the intervention	−4.05 (−6.32 to −1.77)	<.001	−7.50 (−10.17 to −4.82)	<.001	−5.12 (−6.14 to −4.11)	<.001
Random effects						
SD (time, intercept)	0.001 (0.0000198 to 0.0173)	—	5.46 (4.01 to 7.42)	—	0.82 (0.53 to 1.37)	—
Correlation (time, intercept)	N/A[j]	—	0.78 (0.54 to 0.90)	—	0.99 (−1.00 to 1.00)	—
SD of residual	3.28 (2.69 to 4.01)	—	0.001 (0.00000549 to 0.20)	—	1.34 (1.04 to 1.73)	—

[a]All models were adjusted for the baseline measure to address the regression-to-mean issue.

[b]*n* represents the number of observations in the data with long format to support mixed-effects modeling.

[c]GAD-7: Generalized Anxiety Disorder–7.

[d]PHQ-9: Patient Health Questionnaire–9.

[e]PSQI: Pittsburgh Sleep Quality Index.

[f]Model fit—*P*<.001 (Wald chi-square test) and *P*=.34 (likelihood ratio test [LRT]) to compare the mixed-effects models with random-intercept linear models.

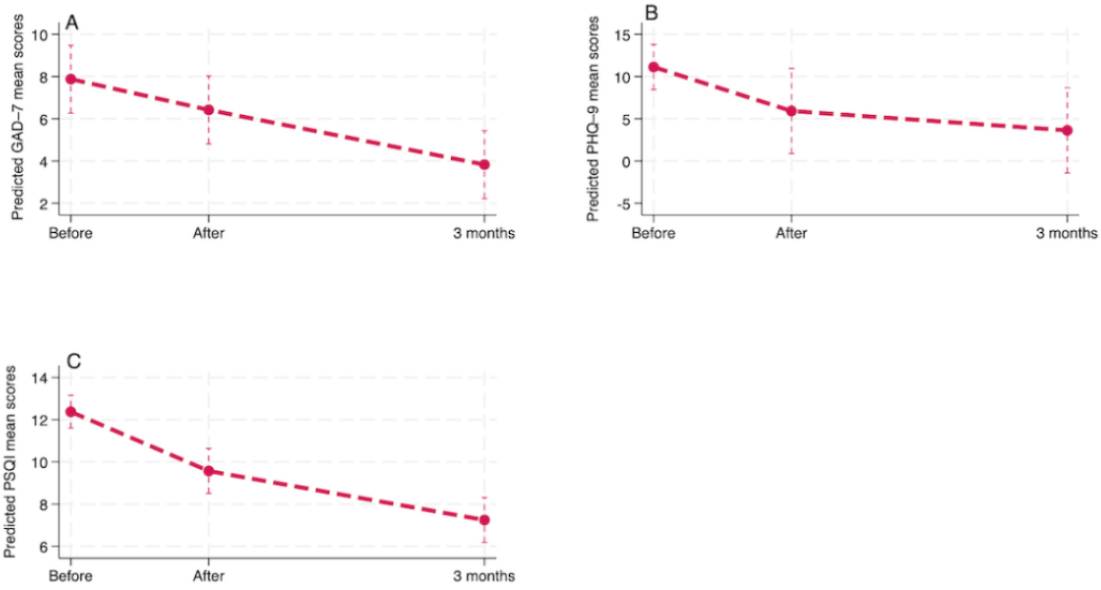
[g]Model fit—*P*<.001 (Wald chi-square test) and *P*=.03 (LRT) to compare the mixed-effects models with random-intercept linear models.

[h]Model fit—*P*<.001 (Wald chi-square test) and *P*=.002 (LRT) to compare the mixed-effects models with random-intercept linear models.

[i]Not available.

[j]N/A: not applicable as the random-intercept model better fit the data than the random-intercept, random-slope model. Given its better fit, random-intercept model parameters were reported.

Figure 5. Change in Generalized Anxiety Disorder–7 (GAD-7), Patient Health Questionnaire–9 (PHQ-9), and Pittsburgh Sleep Quality Index (PSQI) scores for the intervention group from baseline to the postintervention time point and the 3-month follow-up as determined by the intention-to-treat analysis.



XSL•FO
RenderX

Table 10. Mixed-effects linear regression analysis of the change in secondary outcomes over time in the intervention group—per-protocol analysis (N=11)[a].

Variable	Model 1 (n=33[b]; GAD-7[c])		Model 2 (n=33[b]; PHQ-9[d])		Model 3 (n=33[b]; PSQI[e])	
	β (95% CI)	*P* value[f]	β (95% CI)	*P* value[g]	β (95% CI)	*P* value[h]
Fixed effects						
Intercept	2.85 (0.31 to 5.39)	.03	10.63 (8.28 to 12.98)	<.001	0.85 (−3.05 to 4.76)	.67
Time						
Baseline	Reference	—[i]	Reference	—	Reference	—
Postintervention time point	−1.36 (−3.86 to 1.14)	.29	−5.08 (−7.80 to −2.36)	<.001	−3.09 (−4.50 to −1.68)	<.001
3 months after the intervention	−3.88 (−6.38 to −1.38)	.002	−6.98 (−9.71 to −4.26)	<.001	−4.62 (−6.03 to −3.21)	<.001
Random effects						
SD (intercept)	0.90 (0.10 to 8.32)	—	2.28 (1.100 to 4.73)	—	1.12 (0.52 to 2.41)	—
SD of residual	2.99 (2.23 to 4.02)	—	3.26 (2.42 to 4.38)	—	1.69 (1.26 to 2.27)	—

[a]All models were adjusted for the baseline measure to address the regression-to-mean issue.

[b]*n* represents the number of observations in the data with long format to support mixed-effects modeling.

[c]GAD-7: Generalized Anxiety Disorder–7.

[d]PHQ-9: Patient Health Questionnaire–9.

[e]PSQI: Pittsburgh Sleep Quality Index.

[f]Model fit—*P*<.001 (Wald chi-square test) and *P*=.32 (likelihood ratio test [LRT]) to compare the mixed-effects models with random-intercept linear models.

[g]Model fit—*P*<.001 (Wald chi-square test) and *P*=.56 (LRT) to compare the mixed-effects models with random-intercept linear models.

[h]Model fit—*P*<.001 (Wald chi-square test) and *P*=.05 (LRT) to compare the mixed-effects models with random-intercept linear models.

[i]Not available.

XSL•FO
RenderX

Table 11. Mixed-effects linear regression analysis of the change in primary outcomes over time in the waitlist group receiving the intervention—intention-to-treat analysis (N=14)[a].

Variable	Model 1 (n=42[b]; ISI[c])		Model 2 (n=42[b]; NDI[d])		Model 3 (n=42[b]; PCL-5[e])	
	β (95% CI)	*P* value[f]	β (95% CI)	*P* value[g]	β (95% CI)	*P* value[h]
Fixed effects						
Intercept	10.79 (5.89 to 15.70)	<.001	5.42 (2.74 to 8.10)	<.001	13.93 (4.95 to 22.92)	.002
Time						
Baseline	Reference	—[i]	Reference	—	Reference	—
Postintervention time point	−6.36 (−9.16 to −3.57)	<.001	−5.78 (−8.02 to −3.53)	<.001	−19.37 (−26.62 to −12.13)	<.001
3 months after the intervention	−8.21 (−11.01 to −5.41)	<.001	−7.74 (−9.98 to −5.49)	<.001	−27.96 (−35.20 to −20.71)	<.001
Random effects						
SD (intercept)	0.000098 (−0.21 to 0.21)	—	0.0000618 (−0.14 to 0.14)	—	1.73 (0.01 to 302.24)	—
SD of residual	3.78 (3.05 to 4.68)	—	3.03 (2.44 to 3.75)	—	9.78 (7.53 to 12.71)	—

[a]As the random-intercept model better fit the data than the random-intercept, random-slope model, random-intercept model parameters were reported. All models were adjusted for the pretreatment measure to address the regression-to-mean issue.

[b]*n* represents the number of observations in the data with long format to support mixed-effects modeling.

[c]ISI: Insomnia Severity Index.

[d]NDI: Nightmare Disorder Index.

[e]PCL-5: PTSD Checklist for DSM-5.

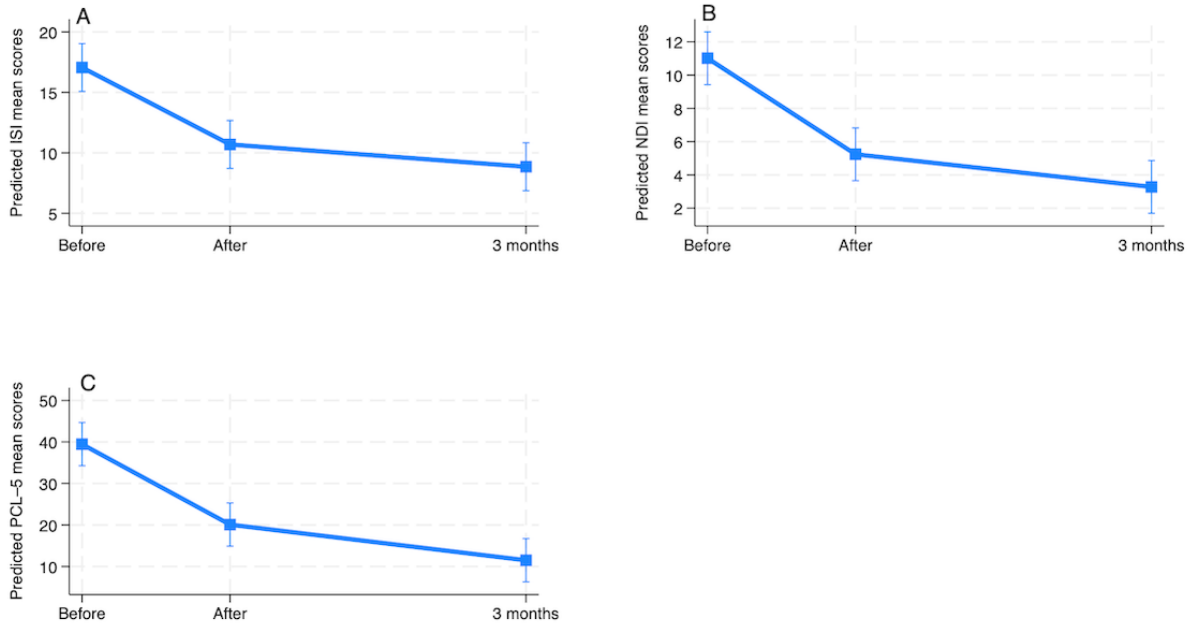
[f]Model fit—*P*<.001 (Wald chi-square test) and *P*=.45 (likelihood ratio test [LRT]) to compare the mixed-effects models with random-intercept linear models.

[g]Model fit—*P*<.001 (Wald chi-square test) and *P*=.59 (LRT) to compare the mixed-effects models with random-intercept linear models.

[h]Model fit—*P*<.001 (Wald chi-square test) and *P*=.14 (LRT) to compare the mixed-effects models with random-intercept linear models.

[i]Not available.

Figure 6. Change in Insomnia Severity Index (ISI); Nightmare Disorder Index (NDI); and PTSD Checklist for DSM-5 (PCL-5) scores for the waitlist group from before to after the intervention and the 3-month follow-up as determined by the intention-to-treat analysis.



XSL•FO
RenderX

Table 12. Mixed-effects linear regression analysis of the change in primary outcomes over time in the waitlist group receiving the intervention—per-protocol analysis (N=9)[a].

Variable	Model 1 (n=27[b]; ISI[c])		Model 2 (n=27[b]; NDI[d])		Model 3 (n=27[b]; PCL-5[e])	
	β (95% CI)	*P* value[f]	β (95% CI)	*P* value[g]	β (95% CI)	*P* value[h]
Fixed effects						
Intercept	11.65 (5.59 to 17.71)	<.001	5.06 (1.47 to 8.66)	.006	12.85 (−1.41 to 27.11)	.08
Time						
Baseline	Reference	—[i]	Reference	—	Reference	—
Postintervention time point	−6.56 (−10.49 to −2.63)	.001	−5.44 (−8.59 to −2.30)	.001	−20.89 (−30.98 to −10.79)	<.001
3 months after the intervention	−7.00 (−10.93 to −3.07)	<.001	−7.11 (−10.26 to −3.96)	<.001	−28.44 (−38.54 to −18.35)	<.001
Random effects						
SD (intercept)	0.01 (0.01 to 0.6)	—	0.01 (N/A[j])	—	3.31 (0.29 to 38.15)	—
SD of residual	4.25 (3.26 to 5.55)	—	3.41 (2.61 to 4.45)	—	10.93 (7.89 to 15.15)	—

[a]All models were adjusted for the baseline measure to address the regression-to-mean issue.

[b]*n* represents the number of observations in the data with long format to support mixed-effects modeling.

[c]ISI: Insomnia Severity Index.

[d]NDI: Nightmare Disorder Index.

[e]PCL-5: PTSD Checklist for DSM-5.

[f]Model fit—*P*<.001 (Wald chi-square test) and *P*>.99 (likelihood ratio test [LRT]) to compare the mixed-effects models with random-intercept linear models.

[g]Model fit—*P*<.001 (Wald chi-square test) and *P*>.99 (LRT) to compare the mixed-effects models with random-intercept linear models.

[h]Model fit—*P*<.001 (Wald chi-square test) and *P*=.34 (LRT) to compare the mixed-effects models with random-intercept linear models.

[i]Not available.

[j]N/A: not applicable.

XSL•FO
RenderX

Table 13. Mixed-effects linear regression analysis of the change in secondary outcomes over time in the waitlist group receiving the intervention—intention-to-treat analysis (N=14)[a].

Variable	Model 1 (n=42[b]; GAD-7[c])		Model 2 (n=42[b]; PHQ-9[d])		Model 3 (n=42[b]; PSQI[e])	
	β (95% CI)	*P* value[f]	β (95% CI)	*P* value[g]	β (95% CI)	*P* value[h]
Fixed effects						
Intercept	0.78 (−2.68 to 4.24)	.66	0.36 (−1.00 to 1.72)	.61	3.31 (–0.62 to 7.25)	.10
Time						
Baseline	Reference	—[i]	Reference	—	Reference	—
Postintervention time point	−4.72 (−6.24 to −3.19)	<.001	−0.42 (−1.84 to 1.00)	.56	−5.07 (−6.42 to −3.71)	<.001
3 months after the intervention	−6.51 (−8.03 to −4.99)	<.001	−0.40 (−1.82 to 1.02)	.58	−5.00 (−6.35 to −3.65)	<.001
Random effects						
SD (time, intercept)	1.47 (0.87 to 2.48)	—	0.06 (2.72×10^{-64} to 1.3×10^{61})	—	0.92 (0.48 to 1.77)	—
Correlation (time, intercept)	0.99 (−1.00 to 1.00)	—	N/A[j]	—	0.99 (−1.00 to 1.00)	—
SD of residual	1.78 (1.33 to 2.37)	—	1.91 (1.48 to 2.48)	—	1.71 (1.28 to 2.78)	—

[a]All models were adjusted for the baseline measure to address the regression-to-mean issue.

[b]*n* represents the number of observations in the data with long format to support mixed-effects modeling.

[c]GAD-7: Generalized Anxiety Disorder–7.

[d]PHQ-9: Patient Health Questionnaire–9.

[e]PSQI: Pittsburgh Sleep Quality Index.

[f]Model fit—*P*<.001 (Wald chi-square test) and *P*<.001 (likelihood ratio test [LRT]) to compare the mixed-effects models with random-intercept linear models.

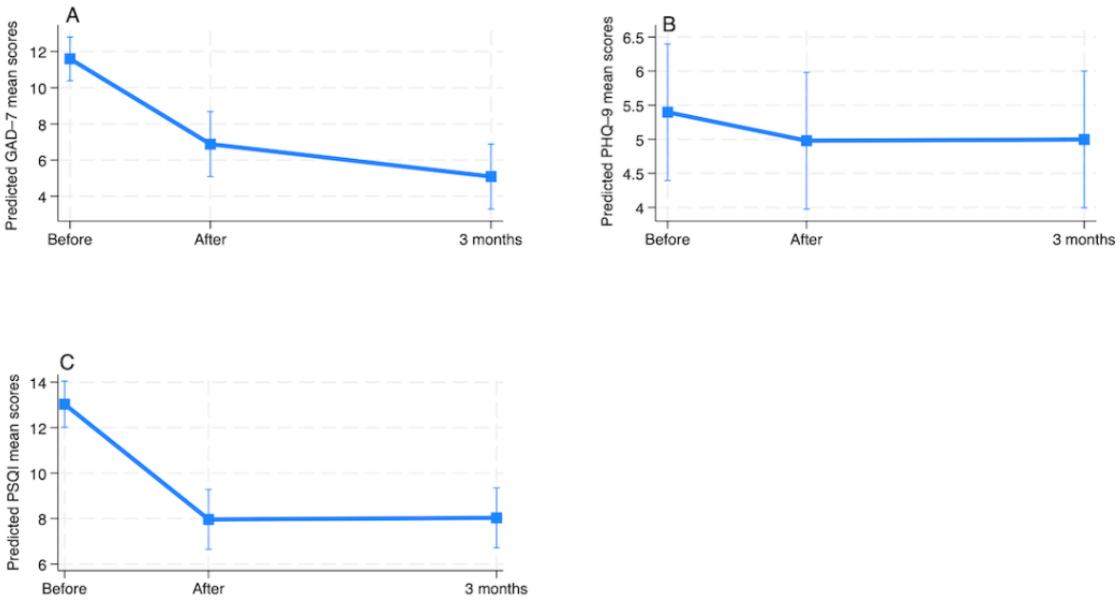
[g]Model fit—*P*<.001 (Wald chi-square test) and *P*=.17 (LRT) to compare the mixed-effects models with random-intercept linear models.

[h]Model fit—*P*<.001 (Wald chi-square test) and *P*=.02 (LRT) to compare the mixed-effects models with random-intercept linear models.

[i]Not available.

[j]N/A: not applicable as the random-intercept model better fit the data than the random-intercept, random-slope model. Given its better fit, random-intercept model parameters were reported.

Figure 7. Change in Generalized Anxiety Disorder–7 (GAD-7), Patient Health Questionnaire–9 (PHQ-9), and Pittsburgh Sleep Quality Index (PSQI) scores for the waitlist group from before to after the intervention and the 3-month follow-up as determined by the intention-to-treat analysis.



XSL•FO
RenderX

Table 14. Mixed-effects linear regression analysis of the change in secondary outcomes over time in the waitlist group receiving the intervention—per-protocol analysis (N=9)[a].

Variable	Model 1 (n=27[b]; GAD-7[c])		Model 2 (n=27[b]; PHQ-9[d])		Model 3 (n=27[b]; PSQI[e])	
	β (95% CI)	*P* value[f]	β (95% CI)	*P* value[g]	β (95% CI)	*P* value[h]
Fixed effects						
Intercept	3.85 (−8.29 to −3.04)	.22	0.43 (−1.61 to 2.48)	.68	5.89 (1.55 to 10.22)	.008
Time						
Baseline	Reference	—[i]	Reference	—	Reference	—
Postintervention time point	−5.67 (−8.29 to −3.04)	<.001	−0.56 (−2.69 to 1.58)	.61	−4.78 (−7.05 to −2.50)	<.001
3 months after the intervention	−7.78 (−10.40 to −5.15)	<.001	−0.11 (−2.25 to 2.03)	.92	−4.78 (−7.05 to −2.50)	<.001
Random effects						
SD (intercept)	1.98 (0.88 to 4.46)	—	0.29 (0.0000012 to 69,857.52)	—	1.00 (0.21 to 4.67)	—
SD of residual	2.84 (2.05 to 3.94)	—	2.31 (1.67 to 3.21)	—	2.46 (1.78 to 3.41)	—

[a]All models were adjusted for the baseline measure to address the regression-to-mean issue.

[b]*n* represents the number of observations in the data with long format to support mixed-effects modeling.

[c]GAD-7: Generalized Anxiety Disorder–7.

[d]PHQ-9: Patient Health Questionnaire–9.

[e]PSQI: Pittsburgh Sleep Quality Index.

[f]Model fit—*P*<.001 (Wald chi-square test) and *P*=.05 (likelihood ratio test [LRT]) to compare the mixed-effects models with random-intercept linear models.

[g]Model fit—*P*<.001 (Wald chi-square test) and *P*=.47 (LRT) to compare the mixed-effects models with random-intercept linear models.

[h]Model fit—*P*<.001 (Wald chi-square test) and *P*=.24 (LRT) to compare the mixed-effects models with random-intercept linear models.

[i]Not available.

Clinical Significance

Table 15 presents the number of participants achieving clinical significance from baseline to 3-month follow-up, based on intention-to-treat (ITT) analysis. The results show that a higher percentage of waitlist group participants achieved clinically significant changes in nightmare symptoms (mean change 7.74, SD 5.12) and PTSD symptoms (mean difference 27.96, SD 16.56) compared to the treatment group (mean change 4.20, SD 4.21 for nightmare symptoms and mean change: 22.83, SD 17.09 for PTSD symptoms). In contrast, a higher percentage of treatment group participants achieved clinically significant changes in insomnia symptoms (mean difference 8.65, SD 5.28) compared to the waitlist group (mean change 8.21, SD 6.96). Two participants were missing 3-month follow-up data on the ISI, NDI, and PCL-5.

The per-protocol (PP) analysis (Table 16) yielded similar results to the ITT analysis. Specifically, a higher percentage of waitlist group participants achieved clinically significant changes in nightmare symptoms (mean change 7.11, SD 5.71) and PTSD symptoms (mean change 28.44, SD 18.58) compared to the treatment group (mean change 4.67, SD 3.04 for nightmare symptoms and mean change: 24.84, SD 18.62 for PTSD symptoms). Conversely, a higher percentage of treatment group participants achieved clinically significant changes in insomnia symptoms (mean change 7.33, SD 5.87) compared to the waitlist group (mean change 7.00, SD 8.34). However, the differences between the two groups were not statistically significant.

Table 15. Distribution of clinically significant responders by condition—intention-to-treat analysis (N=30).

Outcome	Waitlist (n=14), n (%)	Treatment (n=16), n (%)	*P* value
ISI[a]	8 (57)	10 (62)	.77
NDI[b]	13 (93)	10 (62)	.05
PCL-5[c]	11 (79)	9 (56)	.20

[a]ISI: Insomnia Severity Index.

[b]NDI: Nightmare Disorder Index.

[c]PCL-5: PTSD Checklist for DSM-5.

XSL•FO
RenderX

Table 16. Distribution of clinically significant responders by condition—per-protocol analysis (N=18) (N=20).

Outcome	Waitlist (n=9), n (%)	Treatment (n=9), n (%)	*P* value
ISI[a]	4 (44)	5 (56)	.64
NDI[b]	8 (89)	5 (56)	.11
PCL-5[c]	7 (78)	4 (44)	.15

[a]ISI: Insomnia Severity Index.

[b]NDI: Nightmare Disorder Index.

[c]PCL-5: PTSD Checklist for DSM-5.

Satisfaction and Engagement With Treatment

Participants were asked to rate how likely they were to revisit the treatment modules on a 5-point, single-item Likert scale. A frequency analysis showed that most participants responded with *strongly agree* (9/20, 45%) and *moderately agree* (7/20, 35%). In total, 10% (2/20) of the participants responded with *moderately disagree*, and 10% (2/20) of the participants did not respond to this question. The number of log-ins to the site was observed as an indicator of the level of engagement by participants with the modules, and it was found that the 20 participants who were included in the PP analysis had an average of 8.5 log-ins during their engagement with Sleep Best-i. A total of 5% (2/20) logged in 12 times to the site; 5% (3/20) logged in 5 times; 5% (2/20) logged in 7 times; 10% (3/20) logged in 6 times; 15% (4/20) logged in 11 times; and 20% (6/20) logged in 8, 9, and 10 times. On average, participants spent 116 minutes on the site visiting modules and completing assessments.

Discussion

Principal Findings

The aim of this clinical pilot trial was to assess the feasibility of a brief (6 modules over 4 weeks), self-paced, digital intervention for the treatment of insomnia, nightmares, and PTSD symptoms in an international sample of survivors of wildfires. The first hypothesis was partially supported. The PP analysis revealed a significant interaction effect of condition × time on both the NDI and the PCL-5, indicating that Sleep Best-i effectively reduced symptoms of nightmares and PTSD from before to after the intervention in the treatment group compared to the waitlist group. However, no significant changes were observed in insomnia symptoms. The ITT analysis yielded similar findings, with a significant main effect of time showing a reduction in nightmare and PTSD symptoms at the posttreatment time point for the intervention group but no significant changes in insomnia symptoms. In examining the 2 groups separately, Sleep Best-i significantly reduced symptoms of insomnia, nightmares, and PTSD from baseline to the postintervention time point, and this improvement in symptoms was maintained at the 3-month assessment for the 2 groups across both the PP and ITT analyses. This study's findings diverge from those of previous research using CBTi to treat insomnia in survivors of wildfires [27,28] and veterans [25,26]. This discrepancy may be attributed to the brief duration of insomnia treatment in our study, which spanned only the first 2 weeks, unlike in other clinical trials that used CBTi in ≥6 sessions. Research suggests that an average of 6 to 8 sessions is typically required to significantly reduce insomnia symptoms [13,64]. The shorter session duration used in this study may have contributed to the differing outcomes. Nevertheless, the maintenance of improvements at the 3-month follow-up is a promising indicator of the intervention's long-term effectiveness.

These findings further substantiate the efficacy of ERRT in reducing the severity and frequency of nightmares, aligning with those of previous studies [36,65-67]. These results are also consistent with those of other clinical trials demonstrating that CBTi and ERRT can lead to significant reductions in PTSD symptoms at the posttreatment time point [36,68]. Notably, it is possible that symptoms of trauma and nightmares are more malleable and responsive to treatment, whereas insomnia symptoms may be more entrenched and less amenable to change with brief interventions, suggesting a potential explanation for the observed differences in treatment outcomes.

In relation to the second hypothesis, the PP analysis revealed that Sleep Best-i not only alleviated symptoms of nightmares and trauma but also significantly reduced anxiety symptoms (main effect of condition) in the treatment group compared to the waitlist group at the postintervention time point. In addition, the treatment group experienced a significant reduction in depressive symptoms at the posttreatment time point (main effect of time) and a significant interaction effect of time × condition on sleep quality, indicating more pronounced improvements in sleep quality at the postintervention time point compared to the waitlist group. The ITT analysis yielded similar findings, with significant reductions in depressive symptoms at the postintervention time point and a significant interaction effect of time × condition on sleep quality. However, the reduction in anxiety symptoms was no longer significant in the ITT analysis. This is notable as research suggests a strong relationship between insomnia and anxiety symptoms, with the expectation that successful insomnia treatment would lead to a reduction in anxiety symptoms [69].

When analyzing the 2 groups separately, both the PP and ITT analyses yielded similar results, with few exceptions. Notably, both groups showed significant reductions in anxiety symptoms at the 3-month follow-up, although this improvement was only sustained for the waitlist group at the postintervention assessment. In terms of sleep quality, both groups showed significant improvements at both the postintervention and 3-month follow-up assessments. However, only the intervention group exhibited significant reductions in depressive symptoms at both the postintervention and 3-month follow-up assessments.

XSL•FO
RenderX

This study's findings align with those of existing research demonstrating the effectiveness of CBTi-based treatments for comorbid conditions [13,70]. While part of this study's findings corroborates those of previous studies showing a significant reduction in symptoms of depression following CBT treatment for insomnia [27,28], the lack of improvement in symptoms of depression in the waitlist group is not well understood. One possible explanation is that the insignificant improvements in insomnia symptoms may have contributed to the absence of change in depression levels [69]. In addition, external life difficulties may have influenced symptoms of depression during the study period. Notably, the 2 groups had distinct intervention experiences, which may have impacted the outcomes. Specifically, the 3-month follow-up assessment for the waitlist group occurred 4 weeks after the treatment group's follow-up, potentially introducing passage-of-time effects that may have influenced the results. This study's findings are also in line with those of other studies that found improved sleep quality on the PSQI following the administration of CBTi [71].

In terms of clinical significance, a substantial proportion of participants achieved MCSC (according to the PP analysis) in insomnia (9/18, 50%), nightmares (13/18, 72%), and PTSD symptoms (11/18, 61%). Notably, the MCSC rates for insomnia and PTSD in this study were slightly higher than those reported in previous clinical trials using CBTi and imagery rehearsal therapy. For example, Ulmer et al [26] found that 55.4% and 50% of their sample (N=22) achieved clinical significance in insomnia and PTSD, respectively. In contrast, the MCSC rates in this study were lower than those reported by Belleville et al [27], who found that 64.7% and 70.6% of participants achieved MCSC for insomnia and PTSD at the posttreatment time point, respectively, and 64.7% and 58.8% achieved MCSC for insomnia and PTSD at the 3-month follow-up, respectively. The discrepancy in MCSC rates between our study and others may be attributed to the shorter treatment duration of Sleep Best-i (4 weeks) than those of the other trials (6-12 weeks) [26,27].

A significant proportion of participants expressed a high level of satisfaction with the intervention, with approximately 80% (18/20) of users reporting a positive experience. Furthermore, the level of engagement by participants with the modules was measured using the number of log-ins into the site. The 20 participants included in the PP analysis had an average of 8.5 log-in times during their engagement with Sleep Best-i. The log-in frequency was varied, with 5% (2/20) of the participants logging in 12 times; 5% (3/20) logging in 5 times; 5% (2/20) logging in 7 times; 10% (3/20) logging in 6 times; 15% (4/20) logging in 11 times; and 20% (6/20) logging in 8, 9, or 10 times. On average, participants spent 116 minutes on the site visiting modules and completing assessments.

Limitations

This clinical pilot trial has limitations that warrant consideration. The sample consisted of self-selected individuals, and the absence of clinical assessments to confirm diagnoses of insomnia, nightmares, and PTSD introduces a potential source of bias. The small sample size is another impediment, underscoring the need for further testing with a larger and more diverse population to establish the external validity of the intervention. One major concern in relation to Sleep Best-i is the PTSD module and its potential for triggering trauma symptoms. Although this risk was anticipated, participants were provided with emergency numbers. Moreover, throughout the initial 2-week treatment period, participants received training in cognitive and behavioral strategies to effectively manage stress. It is important to note that our study was conducted as an open-label trial. Future research should prioritize addressing the identified gaps and limitations, thereby enhancing the validity and generalizability of the findings in this field.

Future Directions

This feasibility trial provides valuable insights into the effectiveness of the Sleep Best-i intervention, a CBT-based program for survivors of wildfires. While our findings are promising, we recognize that future research could further enhance our understanding of its efficacy and effectiveness. One potential direction for future research could involve stratifying the assessment of outcomes based on the type of survivor by dividing the population into subgroups based on specific characteristics, such as demographic characteristics such as age, sex, and income; life experiences, including previous trauma and social support; experiences with wildfires, such as severity of exposure and proximity to the fire; losses experienced as a result of wildfires, including property damage and loss of loved ones; social support from family, friends, and community resources; and institutional support provided during the recovery phase, including access to mental health services and financial assistance. By doing so, researchers can identify patterns and trends within each subgroup, compare outcomes between subgroups, and develop targeted interventions that address the unique needs of each subgroup. In addition, ecological assessments over longer time frames could provide a more comprehensive understanding of the intervention's impact in real-world settings. We propose that future studies consider these approaches to build on our findings and inform the development of more targeted and effective interventions for survivors of wildfires.

Conclusions

Taken together, these findings indicate that Sleep Best-i incorporating CBTi and ERRT improved nightmares, PTSD, sleep quality, and symptoms of depression from baseline to the posttreatment time point. This positive impact was sustained at the 3-month follow-up for the 2 groups, with some variations on anxiety and depression. Participants in the intervention group, when assessed separately, experienced improvements on all measures from before to after treatment and at the 3-month follow-up, with the exception of anxiety symptoms at the posttreatment time point. The waitlist group experienced a significant reduction in symptoms on all measures from before treatment to after treatment and at the 3-month follow-up except for symptoms of depression. This clinical trial is the first in the field of sleep disturbances to use a concise, digital, self-paced intervention over 4 weeks among survivors of wildfires.

XSL•FO
RenderX

Acknowledgments

The research team would like to thank Dr Samia R Toukhsati for reviewing the ethics application and the treatment manual and for her contribution to the design of the trial; Dr Mirella DiBenedetto for her involvement in the early stages of this research project; and Dr Abby Bloom, Dr Kate Brady, Mr John Richardson, Mr Adam Dent, Ms Mona Naeimi, and Mr Nathan Maddock for their efforts in recruiting participants. The authors would also like to thank Mr Owen Cole for his assistance with the HealthZone platform. FI is supported by an Australian Government Research Training Program Fee Offset scholarship administered through Federation University. This study received funding from Natural Hazards Research Australia through an awarded scholarship to FI.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to ethical restrictions, but are available from the corresponding author on reasonable request.

Authors' Contributions

FI contributed to conceptualization; methodology; validation; formal analysis; data curation; writing—original draft preparation, review, and editing; visualization; project administration; and funding acquisition. BK contributed to conceptualization, methodology, validation, data curation, writing—review and editing, visualization, and supervision. HN contributed to methodology, validation, formal analysis, writing—review and editing, and visualization. SW provided supervision and writing—review and editing. GAK was responsible for the conceptualization, methodology, validation, formal analysis, data curation, writing—review and editing, visualization, and supervision. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

FI, BK, and GAK are the primary creators of Sleep Best-i. All other authors declare no other conflicts of interest.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1327 KB - humanfactors_v12i1e65228_app1.pdf]

Multimedia Appendix 2
Plain-language statement.
[DOCX File , 21 KB - humanfactors_v12i1e65228_app2.docx]

Multimedia Appendix 3
Consent form.
[PDF File (Adobe PDF File), 72 KB - humanfactors_v12i1e65228_app3.pdf]

References

1. Mani ZA, Khorram-Manesh A, Goniewicz K. Global health impacts of wildfire disasters from 2000 to 2023: a comprehensive analysis of mortality and injuries. Disaster Med Public Health Prep 2024 Oct 30;18:e230. [doi: 10.1017/dmp.2024.150] [Medline: 39473373]
2. Gallanter T, Bozeman WP. Firefighter illnesses and injuries at a major fire disaster. Prehosp Emerg Care 2002;6(1):22-26. [doi: 10.1080/10903120290938724] [Medline: 11789645]
3. Kulig JC, Townshend I, Edge DS, Reimer B. Impacts of wildfires: the aftermath at individual and community level. Aust J Emerg Manage 2013 Jul;28(3):23-28 [FREE Full text]
4. Shusterman D, Kaplan JZ, Canabarro C. Immediate health effects of an urban wildfire. West J Med 1993 Feb;158(2):133-138 [FREE Full text] [Medline: 8434462]
5. Johnston DW, Önder YK, Rahman MH, Ulubaşoğlu MA. Evaluating wildfire exposure: using wellbeing data to estimate and value the impacts of wildfire. J Econ Behav Organ 2021 Dec;192:782-798. [doi: 10.1016/j.jebo.2021.10.029]
6. Malhi NK, Marwaha R. Running wild: the impact of wildfires on mental health. Psychiatric Times. 2023 Oct 12. URL: https://www.psychiatrictimes.com/view/running-wild-the-impact-of-wildfires-on-mental-health [accessed 2024-01-13]
7. To P, Eboreime E, Agyapong VI. The impact of wildfires on mental health: a scoping review. Behav Sci (Basel) 2021 Sep 21;11(9):126 [FREE Full text] [doi: 10.3390/bs11090126] [Medline: 34562964]
8. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. Washington, DC: American Psychiatric Association; 2013.

XSL•FO
RenderX

9. Schoenfeld FB, Deviva JC, Manber R. Treatment of sleep disturbances in posttraumatic stress disorder: a review. J Rehabil Res Dev 2012;49(5):729-752. [doi: 10.1682/jrrd.2011.09.0164] [Medline: 23015583]
10. Swift KM, Thomas CL, Balkin TJ, Lowery-Gionta EG, Matson LM. Acute sleep interventions as an avenue for treatment of trauma-associated disorders. J Clin Sleep Med 2022 Sep 01;18(9):2291-2312 [FREE Full text] [doi: 10.5664/jcsm.10074] [Medline: 35678060]
11. Koren D, Arnon I, Lavie P, Klein E. Sleep complaints as early predictors of posttraumatic stress disorder: a 1-year prospective study of injured survivors of motor vehicle accidents. Am J Psychiatry 2002 May;159(5):855-857. [doi: 10.1176/appi.ajp.159.5.855] [Medline: 11986142]
12. Isaac F, Toukhsati SR, Klein B, DiBenedetto M, Kennedy GA. Prevalence and predictors of sleep and trauma symptoms in wildfire survivors. Sleep Epidemiol 2023 Dec;3:100052. [doi: 10.1016/j.sleepe.2022.100052]
13. Isaac F, Toukhsati SR, DiBenedetto M, Kennedy GA. Cognitive behavioral therapy-based treatments for insomnia and nightmares in adults with trauma symptoms: a systematic review. Curr Psychol 2022 Jul 21:1-11 (forthcoming) [FREE Full text] [doi: 10.1007/s12144-022-03512-1] [Medline: 35891889]
14. Germain A, Richardson R, Moul DE, Mammen O, Haas G, Forman SD, et al. Placebo-controlled comparison of prazosin and cognitive-behavioral treatments for sleep disturbances in US Military Veterans. J Psychosom Res 2012 Feb;72(2):89-96 [FREE Full text] [doi: 10.1016/j.jpsychores.2011.11.010] [Medline: 22281448]
15. Taylor DJ, Peterson AL, Pruiksma KE, Young-McCaughan S, Nicholson K, Mintz J. Internet and in-person cognitive behavioral therapy for insomnia in military personnel: a randomized clinical trial. Sleep 2017 Jun 01;40(6). [doi: 10.1093/sleep/zsx075] [Medline: 28472528]
16. Walters EM, Jenkins MM, Nappi CM, Clark J, Lies J, Norman SB, et al. The impact of prolonged exposure on sleep and enhancing treatment outcomes with evidence-based sleep interventions: a pilot study. Psychol Trauma 2020 Feb;12(2):175-185 [FREE Full text] [doi: 10.1037/tra0000478] [Medline: 31246050]
17. Harb GC, Cook JM, Phelps AJ, Gehrman PR, Forbes D, Localio R, et al. Randomized controlled trial of imagery rehearsal for posttraumatic nightmares in combat veterans. J Clin Sleep Med 2019 May 15;15(5):757-767 [FREE Full text] [doi: 10.5664/jcsm.7770] [Medline: 31053215]
18. Ho FY, Chan CS, Tang KN. Cognitive-behavioral therapy for sleep disturbances in treating posttraumatic stress disorder symptoms: a meta-analysis of randomized controlled trials. Clin Psychol Rev 2016 Feb;43:90-102. [doi: 10.1016/j.cpr.2015.09.005] [Medline: 26439674]
19. Morin CM, Espie CA. Insomnia: A Clinical Guide to Assessment and Treatment. Dordrecht, The Netherlands: Kluwer Academic Publishers; 2004.
20. Krakow B, Zadra A. Imagery rehearsal therapy: principles and practice. Sleep Med Clin 2010 Jun;5(2):289-298. [doi: 10.1016/j.jsmc.2010.01.004]
21. Davis JL. Treating Post-Trauma Nightmares: A Cognitive Behavioral Approach. Berlin, Germany: Springer; 2009.
22. Pruiksma KE, Wachen JS, Wardle S, Resick PA. Psychotherapy interventions for comorbid sleep disorders and posttraumatic stress disorder. In: Vermetten E, Germain A, Neylan T, editors. Sleep and Combat-Related Post Traumatic Stress Disorder. New York, NY: Springer; Jan 31, 2018.
23. Pruiksma KE, Taylor DJ, Mintz J, Nicholson KL, Rodgers M, Young-McCaughan S, et al. A pilot randomized controlled trial of cognitive behavioral treatment for trauma-related nightmares in active duty military personnel. J Clin Sleep Med 2020 Jan 15;16(1):29-40 [FREE Full text] [doi: 10.5664/jcsm.8116] [Medline: 31957648]
24. Taylor DJ, Peterson AL, Pruiksma KE, Hale WJ, Young-McCaughan S, Wilkerson A, et al. Impact of cognitive behavioral therapy for insomnia disorder on sleep and comorbid symptoms in military personnel: a randomized clinical trial. Sleep 2018 Jun 01;41(6). [doi: 10.1093/sleep/zsy069] [Medline: 29618098]
25. Margolies SO, Rybarczyk B, Vrana SR, Leszczyszyn DJ, Lynch J. Efficacy of a cognitive-behavioral treatment for insomnia and nightmares in Afghanistan and Iraq veterans with PTSD. J Clin Psychol 2013 Oct;69(10):1026-1042. [doi: 10.1002/jclp.21970] [Medline: 23629959]
26. Ulmer CS, Edinger JD, Calhoun PS. A multi-component cognitive-behavioral intervention for sleep disturbance in veterans with PTSD: a pilot study. J Clin Sleep Med 2011 Feb 15;07(01):57-68. [doi: 10.5664/jcsm.28042]
27. Belleville G, Ouellet MC, Békés V, Lebel J, Morin CM, Bouchard S, et al. Efficacy of a therapist-assisted self-help internet-based intervention targeting PTSD, depression, and insomnia symptoms after a disaster: a randomized controlled trial. Behav Ther 2023 Mar;54(2):230-246. [doi: 10.1016/j.beth.2022.08.004] [Medline: 36858756]
28. Krakow BJ, Melendrez DC, Johnston LG, Clark JO, Santana EM, Warner TD, et al. Sleep Dynamic Therapy for Cerro Grande Fire evacuees with posttraumatic stress symptoms: a preliminary report. J Clin Psychiatry 2002 Aug;63(8):673-684. [doi: 10.4088/jcp.v63n0804] [Medline: 12197447]
29. Rosenthal A, Stover E, Haar RJ. Health and social impacts of California wildfires and the deficiencies in current recovery resources: an exploratory qualitative study of systems-level issues. PLoS One 2021 Mar 26;16(3):e0248617 [FREE Full text] [doi: 10.1371/journal.pone.0248617] [Medline: 33770088]
30. Isaac F. Assessment of the effectiveness of online and face-to-face cognitive behavioural therapy for insomnia/nightmares in adults exposed to trauma using self-report and objective measures: preliminary findings. Trends Telemed E-Health 2022 Jan;3(2). [doi: 10.31031/TTEH.2022.03.000559]

XSL•FO
RenderX

31. Miller KE, Brownlow JA, Gehrman PR. Sleep in PTSD: treatment approaches and outcomes. Curr Opin Psychol 2020 Aug;34:12-17 [FREE Full text] [doi: 10.1016/j.copsyc.2019.08.017] [Medline: 31541965]
32. Mathe N, Johnson ST, Wozniak LA, Majumdar SR, Johnson JA. Alternation as a form of allocation for quality improvement studies in primary healthcare settings: the on-off study design. Trials 2015 Aug 25;16:375 [FREE Full text] [doi: 10.1186/s13063-015-0904-x] [Medline: 26303892]
33. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. Trials 2010 Mar 24;11:32 [FREE Full text] [doi: 10.1186/1745-6215-11-32] [Medline: 20334632]
34. Lachenbruch PA, Cohen J. Statistical power analysis for the behavioral sciences (2nd ed.). J Am Stat Assoc 1989 Dec;84(408):1096. [doi: 10.2307/2290095]
35. Edinger JD. Treatment manual: cognitive-behavioral insomnia therapy. UNC School of Medicine. 2018. URL: https://www.med.unc.edu/neurology/wp-content/uploads/sites/716/2018/05/jdedingrCBTManual.pdf [accessed 2023-02-20]
36. Davis JL, Wright DC. Randomized clinical trial for treatment of chronic nightmares in trauma-exposed adults. J Trauma Stress 2007 Apr;20(2):123-133. [doi: 10.1002/jts.20199] [Medline: 17427914]
37. Edinger JD, Carney CE. Overcoming Insomnia: A Cognitive-Behavioral Therapy Approach, Therapist Guide. Oxford, UK: Oxford University Press; 2014.
38. Lynch J, Mack L, Benesek J, Marshall C, Clevinger L, McHenry S, et al. PTSD recovery program: treatment manual. Hunter Holmes McGuire VA Medical Center. 2015 Oct. URL: https://www.mirecc.va.gov/docs/visn6/PTSD_Recovery_Group-Client_Manual_3rd_edition.pdf [accessed 2025-02-25]
39. Germain A, Buysse DJ. Chapter 15 - Brief behavioral treatment of insomnia. In: Perlis M, Kuhn B, editors. Behavioral Treatments for Sleep Disorders: Practical Resources for the Mental Health Professional. Cambridge, MA: Academic Press; 2011:143-150.
40. Almatiana D, Astuti DS, Ramadhiyanti Y. A systematic review of sparkol videoscribe media in teaching writing skill. J Engl Lang Teach Educ 2021;2(1):1-15 [FREE Full text]
41. Bastien CH, Vallières A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. Sleep Med 2001 Jul;2(4):297-307. [doi: 10.1016/s1389-9457(00)00065-4] [Medline: 11438246]
42. Morin CM, Belleville G, Bélanger L, Ivers H. The insomnia severity index: psychometric indicators to detect insomnia cases and evaluate treatment response. Sleep 2011 May 01;34(5):601-608 [FREE Full text] [doi: 10.1093/sleep/34.5.601] [Medline: 21532953]
43. Dietch JR, Taylor DJ, Pruiksma K, Wardle-Pinkston S, Slavish DC, Messman B, et al. The nightmare disorder index: development and initial validation in a sample of nurses. Sleep 2021 May 14;44(5):zsaa254 [FREE Full text] [doi: 10.1093/sleep/zsaa254] [Medline: 33245781]
44. Weathers FW, Litz BT, Keane TM, Palmieri PA, Marx BP, Schnurr PP. The PTSD Checklist for DSM-5 (PCL-5) – standard [measurement instrument]. The National Centre for PTSD. 2013. URL: https://www.ptsd.va.gov/professional/assessment/adult-sr/ptsd-checklist.asp [accessed 2025-02-25]
45. Krüger-Gottschalk A, Knaevelsrud C, Rau H, Dyer A, Schäfer I, Schellong J, et al. The German version of the Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5): psychometric properties and diagnostic utility. BMC Psychiatry 2017 Nov 28;17(1):379 [FREE Full text] [doi: 10.1186/s12888-017-1541-6] [Medline: 29183285]
46. Blevins CA, Weathers FW, Davis MT, Witte TK, Domino JL. The Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5): development and initial psychometric evaluation. J Trauma Stress 2015 Dec;28(6):489-498. [doi: 10.1002/jts.22059] [Medline: 26606250]
47. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097 [FREE Full text] [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
48. Johnson SU, Ulvenes PG, Øktedalen T, Hoffart A. Psychometric properties of the General Anxiety Disorder 7-item (GAD-7) scale in a heterogeneous psychiatric sample. Front Psychol 2019 Aug 6;10:1713 [FREE Full text] [doi: 10.3389/fpsyg.2019.01713] [Medline: 31447721]
49. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
50. Buysse DJ, Reynolds CF3, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213. [doi: 10.1016/0165-1781(89)90047-4] [Medline: 2748771]
51. Backhaus J, Junghanns K, Broocks A, Riemann D, Hohagen F. Test-retest reliability and validity of the Pittsburgh Sleep Quality Index in primary insomnia. J Psychosom Res 2002 Sep;53(3):737-740. [doi: 10.1016/s0022-3999(02)00330-6] [Medline: 12217446]
52. StataCorp. Stata statistical software: release 18. StataCorp LLC. 2023. URL: https://www.stata.com/new-in-stata/ [accessed 2025-02-25]
53. Pallant J. SPSS Survival Manual: A Step by Step Guide to Data Analysis Using IBM SPSS. London, UK: Routledge; Jul 16, 2020.

XSL•FO
RenderX

54. Hoaglin DC, Iglewicz B. Fine-tuning some resistant rules for outlier labeling. J Am Stat Assoc 1987;82(400):1147-1149. [doi: 10.1080/01621459.1987.10478551]
55. McCoy CE. Understanding the intention-to-treat principle in randomized controlled trials. West J Emerg Med 2017 Oct;18(6):1075-1078 [FREE Full text] [doi: 10.5811/westjem.2017.8.35985] [Medline: 29085540]
56. Ranganathan P, Pramesh CS, Aggarwal R. Common pitfalls in statistical analysis: intention-to-treat versus per-protocol analysis. Perspect Clin Res 2016;7(3):144-146 [FREE Full text] [doi: 10.4103/2229-3485.184823] [Medline: 27453832]
57. Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it work? Int J Methods Psychiatr Res 2011 Mar;20(1):40-49 [FREE Full text] [doi: 10.1002/mpr.329] [Medline: 21499542]
58. von Hippel PT. How many imputations do you need? A two-stage calculation using a quadratic rule. Sociol Methods Res 2020 Aug;49(3):699-718 [FREE Full text] [doi: 10.1177/0049124117747303] [Medline: 39211325]
59. Morris TP, White IR, Royston P. Tuning multiple imputation by predictive mean matching and local residual draws. BMC Med Res Methodol 2014 Jun 05;14:75 [FREE Full text] [doi: 10.1186/1471-2288-14-75] [Medline: 24903709]
60. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009 Jun 29;338:b2393 [FREE Full text] [doi: 10.1136/bmj.b2393] [Medline: 19564179]
61. Wing C, Yozwiak M, Hollingsworth A, Freedman S, Simon K. Designing difference-in-difference studies with staggered treatment adoption: key concepts and practical guidelines. Annu Rev Public Health 2024 May;45(1):485-505 [FREE Full text] [doi: 10.1146/annurev-publhealth-061022-050825] [Medline: 38277791]
62. Yang M, Morin CM, Schaefer K, Wallenstein GV. Interpreting score differences in the Insomnia Severity Index: using health-related outcomes to define the minimally important difference. Curr Med Res Opin 2009 Oct;25(10):2487-2494. [doi: 10.1185/03007990903167415] [Medline: 19689221]
63. Marx BP, Lee DJ, Norman SB, Bovin MJ, Sloan DM, Weathers FW, et al. Reliable and clinically significant change in the clinician-administered PTSD Scale for DSM-5 and PTSD Checklist for DSM-5 among male veterans. Psychol Assess 2022 Feb;34(2):197-203 [FREE Full text] [doi: 10.1037/pas0001098] [Medline: 34941354]
64. Pigeon WR. Treatment of adult insomnia with cognitive-behavioral therapy. J Clin Psychol 2010 Nov;66(11):1148-1160 [FREE Full text] [doi: 10.1002/jclp.20737] [Medline: 20853442]
65. Rhudy JL, Davis JL, Williams AE, McCabe KM, Byrd PM. Physiological-emotional reactivity to nightmare-related imagery in trauma-exposed persons with chronic nightmares. Behav Sleep Med 2008;6(3):158-177. [doi: 10.1080/15402000802162539] [Medline: 18629687]
66. Pruiksma KE, Cranston CC, Rhudy JL, Micol RL, Davis JL. Randomized controlled trial to dismantle exposure, relaxation, and rescripting therapy (ERRT) for trauma-related nightmares. Psychol Trauma 2018 Jan;10(1):67-75. [doi: 10.1037/tra0000238] [Medline: 27977223]
67. Balliett NE, Davis JL, Miller KE. Efficacy of a brief treatment for nightmares and sleep disturbances for veterans. Psychol Trauma 2015 Nov;7(6):507-515. [doi: 10.1037/tra0000055] [Medline: 26010113]
68. Talbot LS, Maguen S, Metzler TJ, Schmitz M, McCaslin SE, Richards A, et al. Cognitive behavioral therapy for insomnia in posttraumatic stress disorder: a randomized controlled trial. Sleep 2014 Feb 01;37(2):327-341 [FREE Full text] [doi: 10.5665/sleep.3408] [Medline: 24497661]
69. Mirchandaney R, Barete R, Asarnow LD. Moderators of cognitive behavioral treatment for insomnia on depression and anxiety outcomes. Curr Psychiatry Rep 2022 Feb;24(2):121-128 [FREE Full text] [doi: 10.1007/s11920-022-01326-3] [Medline: 35061137]
70. Sweetman A, Lovato N, Micic G, Scott H, Bickley K, Haycock J, et al. Do symptoms of depression, anxiety or stress impair the effectiveness of cognitive behavioural therapy for insomnia? A chart-review of 455 patients with chronic insomnia. Sleep Med 2020 Nov;75:401-410. [doi: 10.1016/j.sleep.2020.08.023] [Medline: 32957076]
71. Cook JM, Harb GC, Gehrman PR, Cary MS, Gamble GM, Forbes D, et al. Imagery rehearsal for posttraumatic nightmares: a randomized controlled trial. J Trauma Stress 2010 Oct;23(5):553-563. [doi: 10.1002/jts.20569] [Medline: 20839311]

Abbreviations

CBT: cognitive behavioral therapy
CBTi: cognitive behavioral therapy for insomnia
CONSORT: Consolidated Standards of Reporting Trials
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
DID: difference-in-differences
DSM-5: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
ERRT: exposure, relaxation, and rescripting therapy
GAD-7: Generalized Anxiety Disorder–7
ISI: Insomnia Severity Index
ITT: intention-to-treat

XSL•FO
RenderX

MCSC: minimal clinically significant change
NDI: Nightmare Disorder Index
PCL-5: PTSD Checklist for DSM-5
PHQ-9: Patient Health Questionnaire–9
PLIS: plain-language information statement
PP: per protocol
PSQI: Pittsburgh Sleep Quality Index
PTSD: posttraumatic stress disorder

Edited by A Kushniruk; submitted 04.09.24; peer-reviewed by C Baxter; comments to author 15.01.25; revised version received 22.01.25; accepted 23.01.25; published 14.03.25.

Please cite as:
Isaac F, Klein B, Nguyen H, Watson S, Kennedy GA
Digital Cognitive Behavioral Therapy–Based Treatment for Insomnia, Nightmares, and Posttraumatic Stress Disorder Symptoms in Survivors of Wildfires: Pilot Randomized Feasibility Trial
JMIR Hum Factors 2025;12:e65228
URL: https://humanfactors.jmir.org/2025/1/e65228
doi:10.2196/65228
PMID:

©Fadia Isaac, Britt Klein, Huy Nguyen, Shaun Watson, Gerard A Kennedy. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 14.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Design Preferences, Routines, and Well-Being of Older Adults Using Voice-Guided Digital Mindfulness: Qualitative Interview Study

Lucy McCarren[1], MSc; Sanna Kuoppamäki[1], PhD

Department of Biomedical Engineering and Health Systems, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, Huddinge, Sweden

Corresponding Author:
Lucy McCarren, MSc
Department of Biomedical Engineering and Health Systems
School of Engineering Sciences in Chemistry, Biotechnology and Health
KTH Royal Institute of Technology
Hälsovägen 11
Huddinge, 14152
Sweden
Phone: 46 0736146193
Email: mccarren@kth.se

Abstract

Background: Mindfulness-based interventions have been demonstrated to be effective in improving bodily and emotional well-being. However, only a few studies have explored individual differences in the application and use of digital mindfulness among adults aged ≥65 years. Voice-guided mindfulness technologies can increase the accessibility of mindfulness training, but the expected benefits may not be similar in all user groups.

Objective: This study aims to understand how older adults incorporate mindfulness into their habits and routines, explore the digital mindfulness design preferences of older adults, and understand if and how digital mindfulness can facilitate self-perceived well-being among older adults.

Methods: A qualitative interview study built on an interpretive-constructivist paradigm was conducted among older adults in Sweden who used a voice-guided mindfulness app for a 3-week period in their homes (N=15). Semistructured interviews were conducted one-on-one with participants after using the app. Qualitative thematic analysis was used to explore the lived experiences of digital mindfulness, as articulated by participants, which allowed an open exploration of the subjective experiences of digital mindfulness and their possible effects and outcomes.

Results: From the coding stage, 23 codes describing the digital mindfulness experience were identified from the data. These codes were thematized to group the codes together, resulting in 7 subthemes. From these 7 subthemes, three main themes were formed to answer the research objectives: (1) the embeddedness of digital mindfulness in older adults' daily routines and habits, (2) heterogeneity in older adults' design preferences for digital mindfulness, and (3) the consequences of digital mindfulness on the self-perceived well-being of older adults.

Conclusions: This study concludes that digital mindfulness offers a possibility to enhance the self-perceived well-being of older adults by fostering resilience and self-care. However, adverse effects of mindfulness, such as frustration and discomfort, can also be experienced by older adults. The digital mindfulness experiences and preferences of older adults are highly individual, diverse, and manifold, which indicates that personalized approaches are essential for effective engagement. By acknowledging and addressing the heterogeneous design preferences within this demographic, developers can create more personalized and adaptive voice-guided mindfulness apps.

(JMIR Hum Factors 2025;12:e67533) doi:10.2196/67533

KEYWORDS

mindfulness-based interventions; mobile health; mHealth; digital mindfulness; mindfulness app; older adults; qualitative study

XSL•FO
RenderX

Introduction

Overview

Mindfulness, which can be broadly defined as a state of focusing one's awareness on the present moment [1,2], has gained increasing attention in the field of digital health and mental well-being. Practicing mindfulness has been shown to be associated with reduced levels of stress, anxiety, and depressive symptoms across different population groups, demonstrating the benefits of mindfulness for mental well-being [3-6]. Consequently, digital mindfulness technologies, such as mindfulness apps, virtual reality environments, and wearables, have emerged, which enhance the mindfulness experience and improve the accessibility of mindfulness practice [7,8].

Mindfulness technologies can be considered as digital self-care technologies through which users can engage in the maintenance of physical and mental well-being and prevention of chronic diseases without needing human contact with a health care provider [9]. Self-care is defined as the promotion and maintenance of health and the prevention of disease, with or without the support of health care providers [10,11]. Self-care technologies such as mobile apps can provide a medium for users to collect, monitor, and track health-related behaviors independently, along with guidance and advice on areas that lack accessibility to care professionals or therapy [12,13]. The uptake of digital self-care technologies nevertheless depends on an individual's digital literacy; access to technology; and social, cognitive, and attentional resources, which are unevenly distributed across different social groups [14,15].

Older adults, defined as a group of adults aged ≥65 years, are recognized as a group who could benefit from digital mindfulness. Previous research has mostly explored mindfulness-based interventions (MBIs) from the medical or clinical perspective, demonstrating the clinical effects of mindfulness training for older adults [16-18], whereas only a few studies have explored the use of a mindfulness app with older adults using a qualitative approach [19]. Despite the increasing amount and variance of digital health apps in the market, mindfulness apps are not equally used by all groups. Factors such as age, education, gender, eHealth literacy skills, and privacy concerns may play a role in mobile health (mHealth) app use [20]. The accessibility of digital mindfulness for older adults remains a challenge, as the benefits are influenced by factors such as digital skills and interest toward technology [18]. As a result, the age-specific characteristics of using digital mindfulness have largely remained unexplored, particularly from the perspective of self-perceived benefits and challenges of digital mindfulness.

To fill the aforementioned research gap, we conducted a qualitative interview study to explore voice-guided digital mindfulness among older adults in Sweden. The main objectives are as follows: (1) understand how older adults incorporate mindfulness into their habits and routines, (2) explore the digital mindfulness design preferences of older adults, and (3) understand how digital mindfulness can facilitate self-perceived well-being among older adults. By considering the concept of aged heterogeneity [21,22], our research addresses the individual differences in the experiences and preferences of older adults when using a voice-guided mindfulness app over a 3-week period in their homes.

Background

MBIs in Medical and Clinical Research

Mindfulness originates from the Buddhist practice of meditation [23]. Kabat-Zinn [1], an early pioneer of modern mindfulness in the West, developed mindfulness-based stress reduction (MBSR) in 1979. Kabat-Zinn [1] defines mindfulness as "paying attention in a particular way, on purpose, in the present moment, and nonjudgmentally." In clinical psychology, Bishop et al [2] proposed a 2-component model of mindfulness in order to produce an operational definition. The first component involves the focus of attention so that it is maintained in the present, and the second component involves an accepting and open attitude toward the experiences in the present moment. The Swedish Healthcare Guide (1177 Vårdguiden) describes mindfulness as a method that involves focusing on the present with the help of senses and breathing, which can help an individual develop the ability of conscious presence.

In medical and clinical research, MBIs describe the use of mindfulness training in order to promote health and treat and prevent health-related conditions [6]. MBIs in contemporary psychology have been adopted as a form of mental training to reduce cognitive vulnerability to reactive modes that might otherwise heighten stress and emotional distress [2]. Consequently, MBIs have been developed in the form of mindfulness-based programs such as MBSR and mindfulness-based cognitive therapy for the prevention and treatment of mental health conditions [24].

Mindfulness practice can be considered a holistic way to support the well-being and resilience of older adults [25]. Research shows that MBIs are effective as a treatment for physical symptoms that often occur in later life. In Sweden, Norman et al [26] showed that mindfulness training decreased symptoms of fatigue, dizziness, and shortness of breath in a group of adults aged >70 years with chronic heart failure. Henriksson et al [27] explored the association between mindfulness training and the reduction of chronic pain among individuals aged ≥50 years. The group assigned to mindfulness training showed increased mindfulness skills, reduced pain intensity, reduced pain-related interference, heightened pain acceptance, reduced affective distress, and higher ratings of life satisfaction following the training.

MBIs can also improve mental and emotional well-being in later life. Studies have offered support for the use of MBIs to relieve anxiety, depression, and stress and improve pain acceptance [7,17,28,29]. Randomized controlled trials show that MBIs can improve positive affect and reduce symptoms of anxiety and depression for older adults [7,28]. MBIs can also enhance cognitive functioning [29] and working memory for older adults [8], alleviate chronic insomnia, and improve sleep quality [7,30]. Therefore, mindfulness, by definition, can be considered as an embodied experience, where bodily states shape emotion and cognitive processes [23]. For mindfulness

XSL•FO
RenderX

research, this highlights the reciprocal relationship between the mind and body.

Medical and clinical research has demonstrated the positive effects of mindfulness on well-being, but it remains unclear whether these benefits are strengthened by regular practice and if harmful effects can occur [31]. Galante et al [31] have demonstrated that a wide range of experiences can arise with mindfulness practice, ranging from profoundly positive to challenging and potentially harmful experiences. Unpleasant experiences followed by mindfulness practice have been reported to a relatively high degree in nonclinical contexts. These negative impacts can vary from anxiety issues to traumatic re-experiencing and emotional sensitivity [32]. Among mindfulness app users, reported effects included reduced self-efficacy, lower mood, increased frustration, and inability to control thoughts caused by the lack of supervision associated with self-guided digital mindfulness [33].

Consequently, the expected benefits of mindfulness training are mediated by various individual, social, and cultural factors [31]. At the individual level, the benefits and experiences of mindfulness training are influenced by the duration of practice, previous mindfulness experience, and health conditions [5,17,32]. At the sociodemographic level, social class, education, and cultural background can influence the mindfulness outcomes; yet, these factors have been less studied in previous mindfulness research. For older adults, these mediators in the mindfulness experience can also include familial and social relationships and a sense of community [34]. For these reasons, it becomes important to explore digital mindfulness with a qualitative approach and recognize older adults' experiences with digital mindfulness in relation to their background.

Digital Mindfulness for Older Adults

Digital mindfulness is a specific type of mindfulness training that includes the provision of MBIs through a digital platform, such as a mobile app. Mindfulness apps have increased the availability of MBIs to the public with and without mental health conditions. The most implemented techniques in mindfulness apps are voice-guided mindfulness exercises, self-tracking of mindfulness activities, setting personal goals, and receiving notifications [35]. These activities aim to track and influence the user's emotional state through increasing reflexivity and emotional awareness. Mindfulness apps mostly consist of a self-guided format, where users do not have a personal interaction with a mindfulness instructor [36].

Digital MBIs can increase accessibility and early prevention of mental health, as the provision of digital MBIs makes mindfulness easily accessible and could constitute a feasible alternative to promote mental health at a large scale [37]. However, the success of digital MBIs is reliant on adequate user engagement with digital mindfulness technologies. Several studies have explored mindfulness app users in their everyday contexts, demonstrating the positive effects of the use of mindfulness apps on the overall well-being [38,39]. The continuity of the mindfulness app use nevertheless remains a challenge. It is estimated that <5% of individuals continue using the mindfulness app after 30 days of download [40].

Older adults may have distinct characteristics of engagement with mindfulness apps [16,18,19]. Aged heterogeneity in information and communication technology refers to differences in older adults' backgrounds regarding the resources, needs, skills, and motivation to adopt digital technologies into their daily lives [22,41]. Older adults' use and perceived benefits of mindfulness apps are influenced by their general perceptions and attitudes toward technology, self-efficacy, and skill level [16]. Older adults may use mindfulness apps to improve interpersonal relationships [18,34]. Voice as an interaction modality has been shown to decrease many technology adoption barriers among older adults [42]. However, the conversational style can significantly impact the perceived likability of voice-guided technologies for this age group [43].

Methods

Overview

The app used for the study is Mindfulnessresan, which is a commercially available mobile app available from the Apple App Store. The app was provided free of charge to participants for the duration of the study. Mindfulnessresan was developed by a behavioral scientist and certified mindfulness instructor from the University of Gothenburg. The app has 5- to 10-minute voice-guided mindfulness exercises in Swedish (42 total exercises), with a new exercise available to the user each day. The guided mindfulness practices are based on Vipassana practices [44] and are inspired by the MBSR course by Kabat-Zinn [1]. The mindfulness exercises had varying themes, including focusing on breath, mentally scanning the body, handling negative thoughts, and cultivating a nonjudgmental orientation toward thoughts and experiences.

This study was conducted with 15 participants who reported as volunteers to take part in a 3-stage study using a voice-guided mindfulness app. The sample size was determined based on the principles of purposive sampling, which indicates the selection of "information-rich" cases for in-depth analysis [45]. We considered that the sample size was large enough to allow a new understanding of the self-perceived experiences of the mindfulness app among this age group but small enough to conduct a more in-depth analysis of interview material [46]. In qualitative research, the number of 20 research participants is often considered the saturation point after which little or no new information is generated among participants in one analytically relevant category [45]. In our study, we recruited 20 older adults to take part in all 3 stages of the study, and 15 (75%) participants completed all stages of the study. We aimed to reach a saturation point through the interview procedure, emphasizing open-ended questions, which have shown to contribute to the richness of data.

The study procedure consisted of three stages: (1) preuse study, (2) the use of a voice-guided mindfulness app over a 3-week period, and (3) postuse study. In the preuse study, participants received helped to install the mobile app on their phones, and filled out a questionnaire investigating their previous use of mHealth apps and self-perceived socioemotional well-being. Perceived socioemotional well-being consists of measuring well-being (World Health Organization-Five Well-Being Index

XSL•FO
RenderX

[WHO-5]) and mindfulness (Five Facet Mindfulness Questionnaire [FFMQ] scale). In addition, sociodemographic data variables were collected (age, gender, civil status, and household type).

In stage 2 of this study, participants used the mindfulness app daily for a 3-week period in their homes. Participants were given verbal and written instructions for using the app. Participants were instructed to engage with the mobile app daily at a time of their convenience, but their actual use was not monitored. In this procedure, we expected to better understand participants' authentic use of the app instead of giving participants predefined time frames or strict adherence rules. Participants with Android phones (8/15, 53%) were unable to download the app, as it was available only on the Apple App Store. Hence, those participants accessed screen recordings of the app through a shared drive, where they could follow the same exercises via a file-sharing app on their Android phones.

After having used the mindfulness app for a 3-week period, participants took part in the postuse study by answering a questionnaire and a semistructured interview. The interviews were conducted one-on-one with each participant and lasted approximately 20 minutes. The interviews were audio recorded, and all audio recordings were transcribed for analysis. The data collected during the postuse study consisted of the same perceived socioemotional well-being questionnaires as collected in the preuse study. In addition, participants answered the mHealth App Usability Questionnaire [47].

Analytical Approach

The study analyzed the research material through a qualitative approach, which aimed to address the complexities and subjective meanings of user engagement with a mindfulness app. Qualitative approaches have generated increasing interest in mindfulness research to explain the multifaceted characteristics and embodied nature of mindfulness through the interpretive-constructivist paradigm [48]. Qualitative approaches can provide alternative ways to explore the various subjective experiences of digital mindfulness to complement the medical model of MBIs [49]. They can be considered particularly valuable in addressing the heterogeneity in mindfulness practices among vulnerable populations, which often remain underrepresented in quantitative or medical research [50,51]. This study is reported in accordance with Standards for Reporting Qualitative Research guidelines [52].

The interpretive-constructivist paradigm in mindfulness research draws attention to subjective perceptions and experiences articulated with and through language [48]. The interpretive-constructivist paradigm indicates that the experience of digital mindfulness is socially constructed within language, and the well-being outcomes are subjectively perceived. Therefore, the narratives articulated by older adults can provide detailed accounts on how the subjective experiences of health and well-being are formed under certain contexts or environments [49] or mediated by social or cultural factors [53]. It can also mitigate the potential for a power imbalance between researchers and vulnerable groups as a research population [51] by giving research participants the opportunity to define and identify the experiences of digital mindfulness themselves.

Participant Recruitment

Participants were recruited by sending an email invitation to older adults through one Stockholm municipality that organizes weekly meetings for older adults aged ≥65 years. The Stockholm municipality, through which we recruited participants, represents an average municipality in Stockholm in terms of the sociodemographic background of the population. In the email invitation, we invited older adults interested in mindfulness to join a workshop at their weekly meeting for older adults. Eligibility criteria required participants to be (1) aged ≥65 years, (2) fluent in Swedish, and (3) smartphone owners. We used convenience sampling, selecting participants on a first-come, first-served basis. In the workshop, we presented the project and the Mindfulnessresan app and invited volunteers to take part in the study. From this group, 15 older adults were enrolled in the study and signed an informed consent form. We helped them download the app to their phones and gave verbal and written instructions for using the app. In addition, 5 participants were recruited through distributing the invitation in private yoga studios in Stockholm, following the same procedure as those who were recruited through Stockholm municipality. By recruiting participants in 2 separate locations, we were able to reach a relatively diverse sample of older adults in terms of their sociodemographic background, as shown by Table 1. Of the 15 participants recruited through Stockholm municipality, 10 (67%) completed the study. All (5/5, 100%) participants recruited from yoga studios completed the study.

XSL•FO
RenderX

Table 1. Participants' sociodemographic characteristics (age and gender), previous mindfulness experience, previous mobile health (mHealth) use, self-perceived well-being scores (World Health Organization-Five Well-Being Index [WHO-5]), and mindfulness scores (Five Facet Mindfulness Questionnaire [FFMQ]; N=15).

ID	Age (y)	Sex	Previous mindfulness experience	Previous mHealth use[a]	WHO-5 (0-100)	FFMQ score (0-100)
P1	65	Female	Yes	Intermediate	High (100)	73
P2	65	Female	Yes	High	High (90)	58
P3	66	Male	No	High	High (100)	58
P4	75	Male	Yes	Intermediate	Low (45)	45
P5	66	Female	Yes	Intermediate	High (90)	68
P6	71	Male	No	High	High (95)	60
P7	74	Female	Yes	Intermediate	Intermediate (65)	65
P8	83	Female	Yes	High	Low (40)	70
P9	69	Female	Yes	Intermediate	High (90)	73
P10	78	Female	No	Low	High (100)	58
P11	79	Female	No	Low	High (100)	70
P12	66	Female	No	High	High (95)	65
P13	89	Female	No	High	Intermediate (75)	50
P14	74	Female	Yes	Low	Intermediate (55)	63
P15	77	Female	Yes	Low	High (95)	67

[a]Previous mHealth use is calculated from participants' previous use of mobile technology for health, where daily use indicates high use, 1 to 3 times per week indicates intermediate use, and <1 to 3 times per week indicates low use.

Participant Characteristics

Participants' sociodemographic details are reported in Table 1. Most (12/15, 80%) of the participants were female, and just over half (8/15, 53%) of the participants had previous experience with mindfulness. Most (9/15, 60%) of the participants lived alone, and only 6% (1/15) of the participants lived in a senior living facility. Only 40% (6/15) of the participants were married or in a domestic relationship. Most (10/15, 66%) of the participants evaluated their self-perceived well-being as high (WHO-5 score >75). In total, 20% (3/15) of the participants reported an intermediate well-being (WHO-5 score between 50 and 75). Only 15% (2/15) of the participants reported a low well-being (WHO-5 score ≤50).

In the pre- and postuse study, participants were asked to fill out a survey questionnaire consisting of questions on their previous use of mobile apps, perceived well-being (WHO-5 score), perceived mindfulness (FFMQ scale), and sociodemographic information.

The FFMQ scale is based on a factor analytical study of 5 independently developed mindfulness scales that represent different elements of the conceptualization of mindfulness [54]. The 5 factors value the respondent's degree of observation, description, ability to act with awareness, nonreactivity, and nonjudgment toward inner thoughts and feelings. For this study, a shortened version of the FFMQ scale was used, containing 15 questions. The FFMQ-15 shows valid psychometric properties and can therefore be used in both health care settings and research settings [55]. The raw score was scaled to give a final score between 0 to 100.

The WHO-5 index uses five self-reported measures of overall mental well-being [56]: (1) "I have felt cheerful and in good spirits," (2) "I have felt calm and relaxed," (3) "I have felt active and vigorous," (4) "I woke up feeling fresh and rested," and (5) "My daily life has been filled with things that interest me." There are 5 possible answers, including "daily" (score 4), "1-3 times per week" (score 3), "1-3 times per 2 weeks" (score 2), "more seldom" (score 1), and "never" (score 0). The scores for the 5 questions are summed to obtain a raw score for each participant, and the raw score, ranging from 0 to 20, is multiplied by 5 to give the final score, where 0 represents the lowest well-being and 100 represents the highest.

Results from the prestudy questionnaire revealed participants' previous experience with mobile technology (smartphone, tablet, and portable device), mHealth apps, and conversational agents. Approximately half (6/15, 40%) of the participants used mobile technology daily for health reasons. Only a few participants used mobile technology daily for mental well-being (3/15, 20%), whereas 40% (6/15) of the participants had never used mobile technology for mental well-being. Participants rarely used mHealth apps, except for the Swedish health advice service app, which was used by six (6/15, 40%) of the participants 1 to 3 times per month. Only a few participants had used mHealth apps for tracking steps and sleep and practicing yoga (3/15, 20%). Seven of the participants (7/15, 46%) had used a voice assistant before. The voice assistants they had tried were Google Home, Siri, and ChatGPT. Of those participants who had never tried to use voice assistants, the most common reasons were that they did not need voice assistants or they did not know how to use them.

XSL•FO
RenderX

Data

Semistructured Interviews

In the postuse study, participants were interviewed about the user experience of the mindfulness app and the self-perceived benefits of digital mindfulness training. A semistructured interview topic guide (Multimedia Appendix 1) was prepared, which was guided by previous mindfulness literature. We conducted the interview with an open-minded approach, including questions that would enable the informants to respond as they found suitable. The interview guide was structured around the following themes and questions: (1) perceived health, (2) routines, (3) design preferences and usability of Mindfulnessresan, (4) digital mindfulness and well-being, and (5) ethical considerations.

Thematic Analysis of Semistructured Interviews

The interviews were conducted, transcribed, and analyzed by 2 researchers using an investigator triangulation method. Investigator triangulation refers to the participation of ≥2 researchers in the same study to provide multiple observations and conclusions that can serve as a strategy to validate the trustworthiness of research findings [57]. Researchers conducting the analysis had complementary expertise in social science methodologies and computer science. One of the researchers had a previous background in serving as a mindfulness instructor.

A thematic analysis was used to detect codes and themes from the interview data through identifying, analyzing, and interpreting patterns [58]. Qualitative thematic analysis is an established method in health research to explore individuals' beliefs, perspectives, and experiences within certain social and cultural contexts [59]. Analysis followed the principles of reflexive thematic analysis with 6 distinct stages described: familiarizing with the dataset; coding the dataset; generating initial themes; developing and reviewing themes; refining, defining, and naming themes; and writing up [59]. Data saturation was discussed throughout these stages, with the notification that no new codes or themes emerged from the data after stages 3 and 4 [45].

The interview transcripts were thematized inductively, indicating that codes were generated from the data to allow patterns and themes to emerge from the data without using any predetermined code lists or framework in the analysis [55]. After both researchers had familiarized themselves with the transcriptions, codes that answered the research objectives were identified from the data. This initial coding was conducted on the entire interview material, and certain keywords were selected to represent the codes (eg, "creating routines," "voice," and "sound"). To further recognize patterns in participants' statements, initial themes were generated that gathered codes together in a way that was representative of the data and research questions (eg, "mindfulness as a therapeutic intervention" or "mindfulness coach persona"). In developing and reviewing the themes, researchers jointly discussed the refining and naming of the themes to involve a level of interpretation and conceptualization in the analysis. Thus, the final themes described the patterns in the participants' experience of digital mindfulness that emerged from the data through interpretative analysis. The codes and themes are presented in Multimedia Appendix 2.

Ethical Considerations

The research was approved by the Swedish Ethical Review Authority (reference number 2023-07682-01). All participants gave informed verbal and written consent for participation in the study and publication of anonymized extractions from the data. Before participants signed the informed consent, they were given both written and verbal information about the effects of practicing mindfulness and potential risks in participating in the study. Participants were given the opportunity to ask questions and withdraw from the study at any time. The privacy and confidentiality of participants were protected by giving all participants an anonymous ID that was used in the analysis and reporting of the results. Participants were given a small financial compensation of a gift card worth 200 SEK (US $20.82) after they completed their participation in the study. The amount of financial compensation was not specified in the recruitment material.

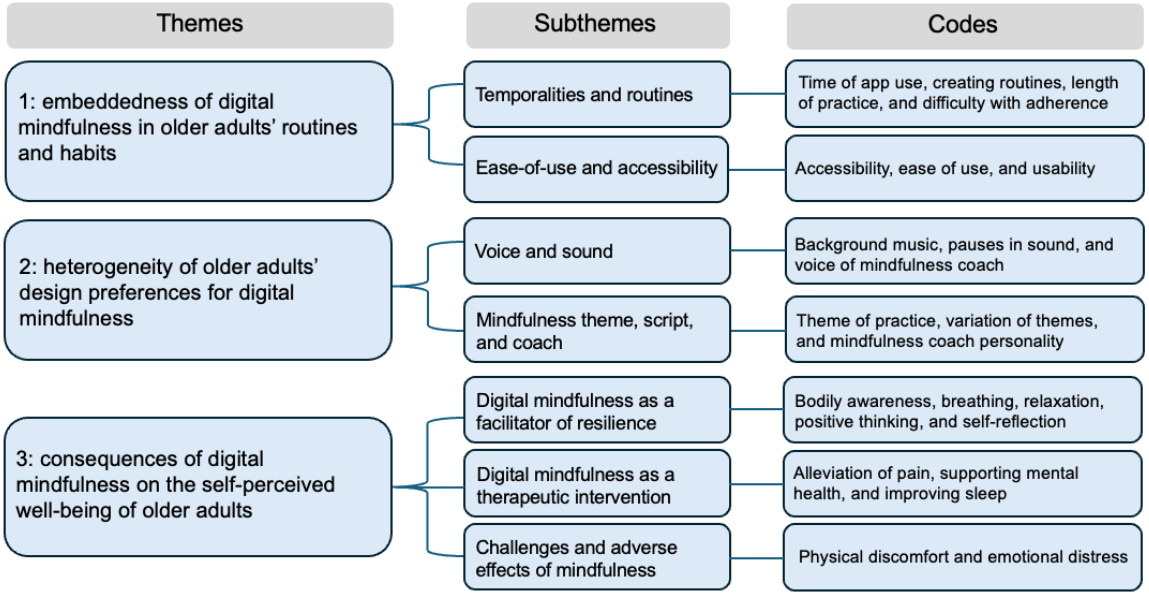
Results

Overview

From the coding stage, 23 codes describing the digital mindfulness experience were identified from the data. These codes were thematized to group the codes together, resulting in 7 subthemes. From these 7 subthemes, three main themes were formed to answer the research objective: (1) the embeddedness of digital mindfulness in older adults' daily routines and habits, (2) heterogeneity in older adults' design preferences for digital mindfulness, and (3) the consequences of digital mindfulness on the self-perceived well-being of older adults. Codes, subthemes, and main themes are presented in Figure 1 and Multimedia Appendix 2.

XSL•FO
RenderX

Figure 1. A map showing the themes, subthemes, and codes that emerged from the thematic analysis of the semistructured interview data.



Theme 1: Embeddedness of Digital Mindfulness in Older Adults' Daily Routines and Habits

Temporalities and Routines

Participants described the use of the app as relatively effortless, because "It takes only 5 minutes so I have time to do it" (P4; male participant aged 75 y). They were able to choose the time of the day for training based on their own preferences. Participants described how the temporality of the app use could significantly influence the experience of mindfulness training. Two participants mentioned that practicing mindfulness in the evening instead of morning can impact the perceived effects of the training:

> *At first, I did it in the evening, but then I started doing it in the morning instead. I almost thought it felt better. Because in the evening, when you do it as the last thing you do, then you end up not hearing what he says, you almost fall asleep.* [P3; male participant aged 66 y]

> *If you listen during the day, you feel a little more calm during the day. But if you listen in the evening, you might sleep better.* [P6; male participant aged 71 y]

Participants recognized that continuity with the app use and embedding the app into the daily routines elicited a sense of well-being:

> *Doing it regularly and creating routines [with the app use] feels good.* [P4; male participant 75 y]

In addition, a participant described that "It worked quite well actually because I did it as a sort of morning routine. [...] I went back to bed and completed the exercise there. It was a very relaxing and pleasant start to the day" (P9; female participant aged 69 y).

The app provided participants with flexibility in finding a suitable time and location for practice, which seemed to improve the user experience. One participant described that practicing mindfulness in familiar locations, such as when lying in bed, helped in creating routines with the app, which contributed to improved user engagement. Mindfulness practice requires concentration, which is easier to do in a familiar location:

> *I liked that I decided to take five minutes. And I almost always did it in the morning. And then I did it while lying in bed. So, I was in the same place. [...] I tried it [once] when I was waiting for a bus. But I found that distracting. [...] So that was the good thing about the app. It took five minutes, a routine.* [P7; female participant aged 74 y]

Some participants enjoyed the fact that the guided mindfulness sessions were short:

> *Yes, it was very good, actually. And it's short as well, just five minutes.* [P6; male participant aged 71 y]

By contrast, others described that they would have enjoyed a longer practice:

> *I thought they were a bit too short. Five minutes is too short to really get into a relaxed state, or whatever you want to call it. [...] I think ten minutes would make it much better.* [P8; female participant aged 83 y]

Some participants described challenges in adhering to the mindfulness routines, which was pronounced as a difficulty in finding a suitable time for digital mindfulness practice. One participant mentioned that her habits, such as having breakfast during a certain time of the day, made it difficult to incorporate the app into her morning routines:

> *It's difficult [...] I'm someone who has very fixed morning routines. So I wake up, and when I wake up, I have to get up and start breakfast immediately.* [P10; female participant aged 78 y]

Another participant mentioned that sometimes she forgot to use the app due to being distracted by other activities and routines of daily life:

XSL•FO
RenderX

> *I must admit I forget it on some days. I try to listen in the mornings, and something always gets in the way. Then I forget it.* [P14; female participant aged 74 y]

Sometimes, participants felt a sense of obligation to conduct the mindfulness training, and missing the daily mindfulness exercises caused a negative feeling. As a participant mentioned, the app provided flexibility in finding a suitable time for training, but this flexibility sometimes resulted in missing exercises:

> *When you miss it and think, now I haven't done it today, but I'll do it tomorrow. Then you feel like you're missing out. It has a slightly negative feeling—yes, it does.* [P5; female participant aged 66 y]

Ease of Use and Accessibility

In total, 27% (4/15) of the participants stated that the app was relatively easy to use, and none of them reported any major technical problems with the app:

> *I didn't have any problems at all. It was easy to use.* [P4; male participant aged 75 y]

Even if the app was considered accessible in technical terms, the graphical interface and symbols associated with training were sometimes considered confusing:

> *It should be clear which exercises you've done or not. The star was a bit confusing—I did not know what it meant.* [P4; male participant aged 75 y]

One participant highlighted that they would have preferred titles for the exercises so that it would have been clear what they were about:

> *I guess it's better that each exercise would have a title. It's not clear what it's about.* [P5; female participant aged 66 y]

One participant suggested that mindfulness apps can increase the accessibility to mindfulness training, in comparison to practicing yoga in a studio, which she considered to be too expensive for her budget as a pensioner:

> *I used to go to medicinal yoga. That was really, really good. So I did two semesters. That's something I miss, but I also miss the money.* [P11; female participant aged 79 y]

Theme 2: Heterogeneity in Older Adults' Design Preferences for Digital Mindfulness

Voice and Sound

Voice interaction and sound in the mindfulness exercises were a significant part of the experience. Many participants described how the voice of the mindfulness coach impacted the perceived benefits of training, such as the sense of calmness:

> *He had a good voice. He spoke clearly, and with a sense of calm.* [P10; female participant aged 78 y]

If the voice of the mindfulness coach was not pleasant or aligned with participants' preferences, this could have decreased the user experience.

Overall, participants expressed diverse preferences for voice interaction and a need to adjust the voice according to their individual desires. The voice interaction should represent qualities such as calmness, quietness, and meditative techniques:

> *I think I would like a quieter voice. A slightly more meditative voice. [...] He speaks in some dialect. He has a very special voice. Maybe that's why I [like] quite a low voice mode. Not so high, but lower. [...] When you try to calm down, you need help with that. A slightly darker voice, not too fast.* [P7; female participant aged 74 y]

Some participants reported inconsistencies in the sound interaction. In total, 27% (4/15) of the participants described how sudden pauses in the sound disturbed the mindfulness practice, which showed the significance of voice and sound interaction in the mindfulness experience:

> *I thought the app's content was quite good, but the way it was done wasn't great. The recording made it so that, when the guy is talking, it sounds good, but then there are pauses now and then, and all the sound and ambient noise disappear.* [P4; male participant aged 75 y]

> *He makes these pauses. They're so long, and then you wonder if there's something wrong with the recording.[...] You get distracted by these pauses. Is this right, or is it wrong? And then you kind of wake up.* [P14; female participant aged 74 y]

Some participants suggested that they would have preferred music in the background to avoid the sudden pauses and to add to the sense of calm:

> *I feel that music adds a lot, so if you would have calm music playing in the background, I think it would create a stronger feeling. [...] It's always so nice when there's a kind of background sound—it's very soothing.* [P3; male participant aged 66 y]

Mindfulness Theme, Script, and Coach

Mindfulness script and content provoked different metaphors among participants. A participant highlighted the effectiveness of figurative descriptions, particularly the metaphor of a wave:

> *I liked the figurative descriptions, like when he said, "feel it like a wave." I felt supported by the app. I had used breathing techniques during pregnancy, but this was completely different.* [P1; female participant aged 65 y]

The imagery of a wave aligns with common mindfulness techniques, where emotions and sensations are framed as transient, reducing resistance and promoting acceptance. Metaphorical language can enhance the practice by linking abstract concepts (eg, transience) to physical sensations.

The content provided by the mindfulness app did not always align with participants' real-life experiences. One participant

XSL•FO
RenderX

indicated that the mindfulness practices in the app may downplay the complexity of real-life struggles:

> *In some exercises, near the end, the person speaking—or this app—seemed to think that big problems could just disappear like that. But it's not that easy.* [P6; male participant aged 71 y]

This comment reflected frustration or skepticism, indicating that while mindfulness can be helpful, it may not provide sufficient depth for more profound personal issues. This aligned with critiques that mindfulness, when decontextualized, may not fully support users with complex psychological challenges.

Some participants valued the soft and reassuring tone of the mindfulness coach. This was illustrated by one participant as follows:

> *He [the mindfulness coach] was nice. His tone was soft as well. He also used those kinds of analogies sometimes. I thought that was nice.* [P15; female participant aged 77 y]

The positive response to the coach's tone suggested that a calm and warm delivery might enhance the soothing effects of mindfulness. In contrast, some participants perceived the coach's persona as too immature for their age group, leading to frustration and disengagement:

> *The guy [mindfulness coach] sounds too childish. He speaks in a way that is a bit immature. It's annoying. It's less authoritative. [...] I was a bit annoyed that his texts can be a little cheesy, a bit childish as well and a bit immature. If you're 65 plus, it's not very smart. That could bother me.* [P4; male participant aged 75 y]

The repeated use of "childish," "cheesy," and "immature" suggested that some older adults preferred a more mature, authoritative style in mindfulness guidance. The critique implied that users may expect more intellectually engaging or sophisticated content. A less authoritative tone may reduce credibility and make the guidance feel less impactful or serious.

Theme 3: Consequences of Digital Mindfulness to the Self-Perceived Well-Being of Older Adults

Digital Mindfulness as a Facilitator of Resilience

Many participants mentioned an increased bodily awareness as an outcome of mindfulness training. A participant stated that "Breathing exercises are what make you get in touch with the body" (P7, female participant aged 74 y). Another participant who had no previous experience with mindfulness and reported only a low use of other mHealth apps noticed that "I feel when I walk, that my shoulders are lower, since I started doing it, that I became aware of my posture" (P11; female participant aged 79 y). By using the app, participants could adopt and increase their bodily knowledge and make changes to improve their bodily postures independently.

A decrease in muscle tension was also described as a self-perceived benefit of digital mindfulness. When mindfulness training involves paying attention to body parts, this can result in improved awareness of the tensions or discomforts in the body. Listening to digital mindfulness can help participants to imagine alternative bodily postures, which can result in a sense of improved bodily awareness. One participant described increased relaxation in the shoulders as a self-perceived benefit in using the app:

> *Some of the exercises include feeling for the legs or the stomach and so on. It's good for bodily awareness. It is very good and useful. [...] For example, if I have pain somewhere, or you are tense somewhere in particular, I tense my shoulders and I can imagine breathing out into my shoulders to relax the shoulders more easily. It's the kind of thing you can put into practice when you use these guided meditations.* [P4; male participant aged 75 y]

Breathing exercises are a central part of mindfulness-based training [22]. To some extent, breathing was considered to be a facilitator and associated with all outcomes of mindfulness training, such as bodily awareness, sleep quality, relaxation, and positive thinking. Participants such as P6, who had no previous mindfulness experience, mentioned an increased awareness of breathing:

> *Breathing—I knew before how important it is, a few deep breaths. But after this exercise, I'm even more motivated, even more aware.* [P6; male participant aged 71 y]

Participants such as P2 and P7 noticed that mindfulness training was connected to improvements in the breathing technique. One of them mentioned feeling better after the breathing exercises, which helped her to improve her breathing technique:

> *Instead of just breathing up here all the time, like I often do. I can say, for that reason, that I feel better.* [P2; female participant aged 65 y]

A participant who had previous mindfulness experience and reported intermediate well-being scores described breathing exercises as a method to pay attention to incorrect breathing techniques, such as holding one's breath:

> *I think the breathing exercises are good. I find it interesting that thoughts affect breathing, and that we breathe too little. We hold our breath a lot. So, I'm always training to breathe with my stomach.* [P7; female participant aged 74 y]

Relaxation was one of the most frequently reported consequences of practicing digital mindfulness through breathing-related exercises, mentioned by 53% (8/15) of the participants. Participants such as P4, P5, and P6 mentioned a sense of calmness, relaxation, and improved focus on the present moment, followed by digital mindfulness training:

> *We even tested my pulse, and with this smartwatch I have, my pulse goes down. It feels really good.* [P6; male participant aged 71 y]

Other participants, such as P10 and P12, mentioned a feeling of joy and a sense of renewed energy associated with relaxation after practicing digital mindfulness:

> *It's about relaxing. When you wake up after ten minutes, you feel like you've come back. You rest,*

XSL•FO
RenderX

> *and your body feels fresher.* [P10; female participant aged 78 y]

For P5 and P6, this sense of relaxation was considered to have cognitive elements, such as maintaining better control of one's own thinking and thoughts:

> *You become calmer because thoughts are focused on breathing [...]. Suddenly, your thoughts don't stick as much, and they might rotate or move away.* [P6; male participant aged 71 y]

Digital mindfulness practice provoked positive thinking among participants. One participant who had previous mindfulness experience described how the mindfulness app increased the ability for self-reflection and harmony and also being in control of one's own emotions:

> *I don't see myself as someone who complains a lot, but I do sometimes meet many people who sigh over just about everything. So yes, reflecting can help. I really like this idea of sitting for a while—it's something that guides you, and you settle into a calmer pace, a nice harmony.* [P5; female participant aged 66 y]

A participant described the emotional awareness associated with the app use. As she described, this awareness of one's own thinking can provoke new approaches for self-reflection, which provides possibilities to change the way one thinks about oneself. From this perspective, mindfulness training has similarities with cognitive behavioral therapy [60]. The participant had previous experience with mindfulness and intermediate self-perceived well-being:

> *You get irritated and frustrated, and you get angry at yourself, you are so frustrated by making a mistake—why did I say this? And then I realise that maybe you can handle it in a different way [...] That maybe I can look at my well-being and my inner self in a different way [...]. I often get stressed by something I have done wrong.* [P14; female participant aged 74 y]

One participant recognized that increasing awareness of breath can help in letting go of negative thoughts. This reflects cognitive defusion, a core principle in mindfulness and acceptance and commitment therapy [61]. Cognitive defusion refers to the process of detaching from thoughts, seeing them as transient mental events rather than absolute truths:

> *You become calmer, because thoughts go on breathing and focus. And it was a good method, to let the mind rest on the sensation of the breath. Suddenly, thoughts don't stick very much, and they might rotate or go away. It felt good, I became a bit more conscious, you could say.* [P6; male participant aged 71 y]

Digital Mindfulness as a Therapeutic Intervention

Several participants recognized the mindfulness training as beneficial for supporting their mental health to alleviate a sense of worry or stress. P6 (male participant aged 71 y), who had no previous experience with digital mindfulness and reported high well-being scores, mentioned mindfulness training as a good method to decrease a sense of worry about "world events, worries about my family—my children, grandchildren."

During the interviews, P2 (female participant aged 65 y) and P4 (male participant aged 75 y), who both had previous experience with mindfulness, shared that they had experienced panic attacks before in life. For this reason, they had used breathing to help calm down and reported a need to find different methods for relaxation. One of them also described his personal experience with having attention-deficit/hyperactivity disorder, which is often accompanied by high energy and heightened emotions. He mentioned that with voice-guided digital mindfulness, he learned new tools to manage this condition:

> *I think mindfulness is very helpful for people with ADHD. I believe it's very, very helpful. That's one reason why I started [mindfulness training]. With ADHD, you have a lot of energy, and your head spins a lot. Learning to manage that better is a big win.* [P4; male participant aged 65 y]

Breathing techniques embedded in the app created a sense of bodily relaxation, which was considered to improve sleep quality. Two participants used the breathing exercises as a method of relaxation in the evenings, which resulted in them falling asleep faster. In this way, digital mindfulness could be used as an intervention to improve sleep in a self-guided format:

> *I have trouble sleeping. I sleep for two hours, then I wake up. But then I get up and sit and listen to it [mindfulness app]. If that doesn't work, I do crosswords for a while, then listen [to the app] again. Then I go back to bed, and I fall asleep.* [P11; female participant aged 79 y]
>
> *What I liked, and which has helped me quite a bit are these breathing exercises. Because it has sort of helped me to fall asleep faster in the evening than I have done before.* [P9; female participant 69 y]

In total, 33% (5/15) of the participants reported that the mindfulness training with the app helped them in managing pain and physical discomfort. Although the app did not alleviate the pain, as such, it could help redirect one's thoughts and "disconnect" from the feeling of pain. A participant described that the mindfulness app helped her to cope with chronic pain caused by rheumatism, which caused physical discomfort in her daily life:

> *I've had PMR [polymyalgia rheumatica] for a year now, and it has been so hard [...], there is a lot of pain sometimes. When you sit and just feel that pain, or lie down and just feel that, then it's very difficult to get your mind off it. But with the mindfulness app, I disconnect from that. So, then the pain disappears.* [P11; female participant 79 y]

A participant mentioned that increased bodily awareness through mindfulness training helped them accept the pain or physical limitations caused by an injury. This self-perceived benefit was connected to the relaxation that the app provided:

> *I am right-handed and the injury is on the right side of the neck vertebrae. [...] Of course, mindfulness helps. Thinking about relaxation, even though there's*

XSL•FO
RenderX

an injury, helps you manage it in the right way. You learn to relax, and that helps. [P12; female participant aged 66 y]

Adverse Effects of Digital Mindfulness

Sometimes participants reported adverse or unexpected negative effects from practicing digital mindfulness. One participant described how practicing digital mindfulness resulted in a sense of discomfort and a sense of physical pain because of not being able to find the right position or location for the training. Her statement also indicated that benefiting from digital mindfulness often requires familiarity with finding locations that are optimal for mindfulness training:

When I sit on a chair, I get back pain, and I don't know if it's because I'm sitting still, but I feel more pain in my back. I have to move around more. It's so rare to sit still in a relaxed way. Now I did get back pain, because of the chair. One should sit on a better chair or lie down, but then you might fall asleep. [P14; female participant aged 74 y]

Another adverse effect of digital mindfulness was mild emotional distress, as mentioned by P3 and P1. P3, who had no previous mindfulness experience, articulated that mindfulness scripts from the app highlighted unpleasant situations in everyday life that provoked distracting emotions and disturbed his sense of calmness:

It almost became annoying then when he brought up different everyday situations. "Why did I say that? Why did I do that?" I don't want to think about that; I just want to continue with the calmness. [P3; male participant aged 66 y]

Similarly, one of the participants described how breathing exercises generated unpleasant memories from a personal life event. For her, this emotional distress was associated with the breathing activity:

Breathing is emotionally charged for me because my brother passed away from ALS [Amyotrophic Lateral Sclerosis], which makes it hard to breathe. So it made me think about that. I've had quite a bit of pain at times in my life—not really now—but it also reminded me of that. [P1; female participant 65 y]

Discussion

Principal Findings

Overview

Previous studies have explored the effects of mindfulness training on the well-being of older adults through randomized controlled trials from the medical or clinical perspective [16-18]. The research evidence indicates the positive effects of mindfulness training on physical well-being [26,27]. Previous studies have also confirmed that MBIs can reduce negative emotions and thoughts in daily life among the general population [24,25,27] and demonstrated the effectiveness of MBIs to reduce stress, anxiety, and depression among older adults [4,17,28,29,62,63]. However, the research evidence on the specific characteristics of voice-guided digital mindfulness for mental and emotional well-being among older adults is rare. This study has filled this research gap by exploring voice-guided digital mindfulness and emphasizing the subjective experiences of digital mindfulness and their possible effects and outcomes among older adults.

Creating Routines With Digital Mindfulness Through Temporalities

As a response to the first research objective, the study has characterized the specific conditions for embedding the mindfulness app into the daily routines of older adults, which can be summarized as mentioned subsequently.

First, in order to actively benefit from digital mindfulness, older adults need to integrate the use of the app into their daily routines and schedules. In general, participants found the digital mindfulness app easy to incorporate into their daily schedules, with the flexibility to choose sessions at times that suited their personal routines. The brevity of the sessions, typically approximately 5 minutes, was appreciated by many participants, facilitating regular practice without significant time commitment. This highlights how the flexibility of digital platforms allows personalized engagement, enabling users to select practices that align with their individual preferences and schedules.

Second, interestingly, the challenges in the adherence to mindfulness training were not associated with technical difficulties or the accessibility of the app but rather with the extent to which older adults could find a suitable time to practice digital mindfulness. This underscores the need for apps to provide reminders or integrate seamlessly with users' existing schedules to promote adherence and aligns with existing research highlighting that continuity in app use can be a challenge [40].

Third, although the app itself enabled practicing digital mindfulness "anywhere and anytime," older adults often chose the location and time based on their existing daily habits, thus embedding the app into existing routines and practices. Participants highlighted how the effects of mindfulness varied according to what time of the day they practiced.

Heterogeneous Design Preferences for Voice-Guided Mindfulness

As a response to the second research objective, this study has recognized heterogeneity in the design preferences for voice-guided digital mindfulness. Practicing digital mindfulness is a highly personalized experience, influenced by a person's previous mindfulness experience, health conditions, and motivations for training [32,64]. This study shows that this heterogeneity is particularly pronounced in relation to design preferences of digital mindfulness apps: voice and sound landscape and mindfulness theme, script, and coach persona. The diversity in older adults' experiences and preferences with digital mindfulness underscores the necessity for customizable voice-guided apps. This variability suggests that a one-size-fits-all approach may not be effective, and developers should consider offering adjustable features, such as voice tone, session length, and content complexity, to cater to the unique needs of each user.

XSL•FO
RenderX

This study indicates that design preferences can partly be associated with various emotions and sensations that participants associate with certain types of voices, sounds, and scripts. For instance, the description of some mindfulness scripts as too "childish" or "immature" for older adults is a sign that mindfulness content should be adapted to age-specific characteristics. While other studies have revealed preferences for personalization of a mindfulness app among the younger population [65], the age-specific characteristics for certain interaction modalities or conversational styles have not been explored. This study concludes that the audio landscape of mindfulness apps provides possibilities for enhancing the mindfulness experience in a way that can respond to the age-specific characteristics of this age group.

Building Resilience and Self-Care Through Digital Mindfulness

As a response to the third research objective, this study has addressed older adults' self-perceived experiences of using digital mindfulness as a self-care technology to facilitate well-being. Previous research shows that mindfulness is useful for improving resilience and coping strategies and reducing anxiety and stress levels in older adults [62,63,66]. This study highlighted that digital mindfulness practices could contribute to resilience by fostering positive thinking, bodily awareness, relaxation, and self-reflection. Participants noted a greater ability to manage negative emotions and adopt new perspectives on personal well-being. Beyond emotional benefits, participants perceived digital mindfulness to be a therapeutic tool for managing mental health concerns and pain and improving sleep quality. Participants reported that mindfulness practices helped redirect focus away from pain, providing a sense of relief. This aligns with studies suggesting that web-based mindfulness interventions can lead to improvements in cognitive and psychological measures, with associated modulations in pain perception.

However, the adverse effects of digital mindfulness were also pronounced, in terms of discomfort and emotional distress that resemble re-experiencing unpleasant life events [32]. These findings underscore the importance of personalized mindfulness interventions and the potential need for guidance on adapting mindfulness practices to individual emotional and physical comfort levels.

By incorporating mindfulness practices into their routines, older adults can develop better coping mechanisms to navigate stressors and challenges. The convenience of digital platforms ensures that these resources are readily available, promoting sustained engagement and long-term benefits.

Limitations

This study has explored the use of a mindfulness app as a form of digital self-care among older adults through the lens of aged heterogeneity. Participants were recruited through a municipal organization, which means that research participants were relatively active older adults with a previous interest in health and well-being. Moreover, participants used the app for 3 weeks; therefore, the long-term effects of mindfulness app use are not known. This study focused on the self-perceived benefits of digital mindfulness, as articulated by participants themselves, which means that the actual use patterns remained unexplored in this study. The use, design preferences, and consequences to well-being of the mindfulness app could depend on contextual factors, such as mood, digital skills, and education level, which remain beyond the scope of this study. Despite these limitations, this study has produced new knowledge on the self-perceived benefits of digital mindfulness among older adults, particularly regarding the affective aspects of mindfulness app use and the integration of the app into daily routines. These qualitative findings extend the already existing medical and clinical research on digital mindfulness with older adults [14-16,18] and provide considerations for designing MBIs beyond medical or clinical purposes, such as in the maintenance of cognitive or physical functioning in later life.

Conclusions

The study concludes that digital mindfulness offers a possibility to enhance the self-perceived well-being of older adults by fostering resilience and self-care. However, the adverse effects of digital mindfulness, such as frustration and discomfort, can also be experienced by older adults. The digital mindfulness experiences and preferences of older adults are highly individual, diverse, and manifold, which indicates that personalized approaches are essential for effective engagement. By acknowledging and addressing the heterogeneous design preferences within this demographic, developers can create more effective and user-friendly voice-guided mindfulness apps. Additional research is needed to explore the emerging digital inequalities involving older adults from a more diverse socioeconomic background to explore the benefits of digital mindfulness in relation to an individual's digital health literacy. Furthermore, longitudinal studies could provide deeper insights into the sustained effects of digital mindfulness on well-being over time. Future research should also explore ways to personalize digital mindfulness experiences, considering factors such as previous mindfulness experience, individual emotional triggers, and physical limitations. Future studies should also explore the potential of artificial intelligence to enhance the personalization of mindfulness experience to be more adaptive to older adults' previous mindfulness experience and their personal preferences for certain types of voices, sounds, and mindfulness coaches' personas.

Acknowledgments

The authors would like to thank all participants for their invaluable participation in this study. The authors also wish to thank Sebastiaan Meijer for commenting on the manuscript.

XSL•FO
RenderX

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview topic guide.
[DOCX File , 20 KB - humanfactors_v12i1e67533_app1.docx]

Multimedia Appendix 2
Thematic analysis of interview data.
[XLSX File (Microsoft Excel File), 36 KB - humanfactors_v12i1e67533_app2.xlsx]

References

1. Kabat-Zinn J. Mindfulness-based interventions in context: past, present, and future. Clin Psychol Sci Pract 2003;10(2):144-156. [doi: 10.1093/clipsy.bpg016]
2. Bishop SR, Lau M, Shapiro S, Carlson L, Anderson ND, Carmody J, et al. Mindfulness: a proposed operational definition. Clin Psychol Sci Pract 2004;11(3):230-241. [doi: 10.1093/clipsy.bph077]
3. Zhang D, Lee EK, Mak EC, Ho CY, Wong SY. Mindfulness-based interventions: an overall review. Br Med Bull 2021 Jun 10;138(1):41-57 [FREE Full text] [doi: 10.1093/bmb/ldab005] [Medline: 33884400]
4. Hazlett-Stevens H, Singer J, Chong A. Mindfulness-based stress reduction and mindfulness-based cognitive therapy with older adults: a qualitative review of randomized controlled outcome research. Clin Gerontol 2019;42(4):347-358. [doi: 10.1080/07317115.2018.1518282] [Medline: 30204557]
5. Mirabito G, Verhaeghen P. The effects of mindfulness interventions on older adults' cognition: a meta-analysis. J Gerontol B Psychol Sci Soc Sci 2023 Mar 04;78(3):394-408. [doi: 10.1093/geronb/gbac143] [Medline: 36148552]
6. Ong NY, Teo FJ, Ee JZ, Yau CE, Thumboo J, Tan HK, et al. Effectiveness of mindfulness-based interventions on the well-being of healthcare workers: a systematic review and meta-analysis. Gen Psychiatr 2024;37(3):e101115 [FREE Full text] [doi: 10.1136/gpsych-2023-101115] [Medline: 38737894]
7. Li J, Leshed G. Beyond meditation: everyday mindfulness and technology use. In: Proceedings of the CHI Conference on Human Factors in Computing Systems Extended Abstracts. 2022 Presented at: CHI EA '22; April 29-May 5, 2022; New Orleans, LA. [doi: 10.1145/3491101.3519820]
8. Seabrook E, Kelly R, Foley F, Theiler S, Thomas N, Wadley G, et al. Understanding how virtual reality can support mindfulness practice: mixed methods study. J Med Internet Res 2020 Mar 18;22(3):e16106 [FREE Full text] [doi: 10.2196/16106] [Medline: 32186519]
9. Lupton D. Quantifying the body: monitoring and measuring health in the age of mHealth technologies. Critical Public Health 2013 Dec;23(4):393-403. [doi: 10.1080/09581596.2013.794931]
10. Riegel B, Dunbar SB, Fitzsimons D, Freedland KE, Lee CS, Middleton S, et al. Self-care research: where are we now? Where are we going? Int J Nurs Stud 2021 Apr;116:103402 [FREE Full text] [doi: 10.1016/j.ijnurstu.2019.103402] [Medline: 31630807]
11. Levin LS, Idler EL. Self-care in health. Annu Rev Public Health 1983 May;4(1):181-201. [doi: 10.1146/annurev.pu.04.050183.001145] [Medline: 6860437]
12. Price M, Yuen EK, Goetter EM, Herbert JD, Forman EM, Acierno R, et al. mHealth: a mechanism to deliver more accessible, more effective mental health care. Clin Psychol Psychother 2014;21(5):427-436 [FREE Full text] [doi: 10.1002/cpp.1855] [Medline: 23918764]
13. Ralston AL, Andrews ARIII, Hope DA. Fulfilling the promise of mental health technology to reduce public health disparities: review and research agenda. Clin Psychol Sci Pract 2019 Mar;26(1):e12277. [doi: 10.1111/cpsp.12277]
14. Winters N, Venkatapuram S, Geniets A, Wynne-Bannister E. Prioritarian principles for digital health in low resource settings. J Med Ethics 2020 Apr 16;46(4):259-264 [FREE Full text] [doi: 10.1136/medethics-2019-105468] [Medline: 31949027]
15. Yao R, Zhang W, Evans R, Cao G, Rui T, Shen L. Inequities in health care services caused by the adoption of digital health technologies: scoping review. J Med Internet Res 2022 Mar 21;24(3):e34144 [FREE Full text] [doi: 10.2196/34144] [Medline: 35311682]
16. Galluzzi S, Lanfredi M, Moretti DV, Rossi R, Meloni S, Tomasoni E, et al. Cognitive, psychological, and physiological effects of a web-based mindfulness intervention in older adults during the COVID-19 pandemic: an open study. BMC Geriatr 2024 Feb 14;24(1):151 [FREE Full text] [doi: 10.1186/s12877-024-04766-z] [Medline: 38350854]

XSL•FO
RenderX

17. Geiger PJ, Boggero IA, Brake CA, Caldera CA, Combs HL, Peters JR, et al. Mindfulness-based interventions for older adults: a review of the effects on physical and emotional well-being. Mindfulness (N Y) 2016 Apr 14;7(2):296-307 [FREE Full text] [doi: 10.1007/s12671-015-0444-1] [Medline: 27200109]
18. Mahlo L, Windsor TD. Feasibility, acceptability, and preliminary efficacy of an app-based mindfulness-meditation program among older adults. Gerontologist 2021 Jul 13;61(5):775-786. [doi: 10.1093/geront/gnaa093] [Medline: 32663286]
19. Vailati Riboni F, Sadowski I, Comazzi B, Pagnini F. Mindful age and technology: a qualitative analysis of a tablet/smartphone app intervention designed for older adults. Integr Psychol Behav Sci 2022 Sep 29;56(3):739-754 [FREE Full text] [doi: 10.1007/s12124-020-09580-x] [Medline: 33118138]
20. Bol N, Helberger N, Weert JC. Differences in mobile health app use: a source of new digital inequalities? Inf Soc 2018 Apr 26;34(3):183-193. [doi: 10.1080/01972243.2018.1438550]
21. Taipale S, Oinas T, Karhinen J. Heterogeneity of traditional and digital media use among older adults: a six-country comparison. Technol Soc 2021 Aug;66:101642. [doi: 10.1016/j.techsoc.2021.101642]
22. Hänninen R, Taipale S, Luostari R. Exploring heterogeneous ICT use among older adults: the warm experts' perspective. New Media Soc 2020 May 06;23(6):1584-1601. [doi: 10.1177/1461444820917353]
23. Khoury B, Knäuper B, Pagnini F, Trent N, Chiesa A, Carrière K. Embodied mindfulness. Mindfulness 2017 Mar 27;8(5):1160-1171. [doi: 10.1007/s12671-017-0700-7]
24. Niemi M, Crane R, Sinselmeijer J, Andermo S. The implementation of mindfulness-based programs in the Swedish healthcare system - a survey study of service providers. Glob Adv Health Med 2021 Nov 03;10:21649561211049154 [FREE Full text] [doi: 10.1177/21649561211049154] [Medline: 34760341]
25. Kourayem A, Mahmoodinia S. The quality of life in the elderly: the role of resiliency and mindfulness. Aging Psychol 2021;7(2):167-179 [FREE Full text] [doi: 10.22126/jap.2021.6217.1510]
26. Norman J, Fu M, Ekman I, Björck L, Falk K. Effects of a mindfulness-based intervention on symptoms and signs in chronic heart failure: a feasibility study. Eur J Cardiovasc Nurs 2018 Jan 22;17(1):54-65 [FREE Full text] [doi: 10.1177/1474515117715843] [Medline: 28639841]
27. Henriksson J, Wasara E, Rönnlund M. Effects of eight-week-web-based mindfulness training on pain intensity, pain acceptance, and life satisfaction in individuals with chronic pain. Psychol Rep 2016 Dec;119(3):586-607. [doi: 10.1177/0033294116675086] [Medline: 27780873]
28. Talebisiavashani F, Mohammadi-Sartang M. The effect of mindfulness-based interventions on mental health and cognitive function in older adults: a systematic review and meta-analysis. J Aging Health 2024 Jul 24:8982643241263882. [doi: 10.1177/08982643241263882] [Medline: 39046424]
29. Lee EK, Wong B, Chan PH, Zhang DD, Sun W, Chan DC, et al. Effectiveness of a mindfulness intervention for older adults to improve emotional well-being and cognitive function in a Chinese population: a randomized waitlist-controlled trial. Int J Geriatr Psychiatry 2022 Jan 27;37(1). [doi: 10.1002/gps.5616] [Medline: 34415638]
30. González-Martín AM, Aibar-Almazán A, Rivas-Campo Y, Marín-Gutiérrez A, Castellote-Caballero Y. Effects of mindfulness-based cognitive therapy on older adults with sleep disorders: a systematic review and meta-analysis. Front Public Health 2023 Dec 18;11:1242868 [FREE Full text] [doi: 10.3389/fpubh.2023.1242868] [Medline: 38179560]
31. Galante J, Grabovac A, Wright M, Ingram DM, Van Dam NT, Sanguinetti JL, et al. A framework for the empirical investigation of mindfulness meditative development. Mindfulness 2023 Apr 13;14(5):1054-1067. [doi: 10.1007/s12671-023-02113-8]
32. Goldberg SB, Lam SU, Britton WB, Davidson RJ. Prevalence of meditation-related adverse effects in a population-based sample in the United States. Psychother Res 2022 Mar 02;32(3):291-305 [FREE Full text] [doi: 10.1080/10503307.2021.1933646] [Medline: 34074221]
33. Taylor GB, Vasquez TS, Kastrinos A, Fisher CL, Puig A, Bylund CL. The adverse effects of meditation-interventions and mind–body practices: a systematic review. Mindfulness 2022 Jun 06;13(8):1839-1856. [doi: 10.1007/s12671-022-01915-6]
34. Parra DC, Wetherell JL, Van Zandt A, Brownson RC, Abhishek J, Lenze EJ. A qualitative study of older adults' perspectives on initiating exercise and mindfulness practice. BMC Geriatr 2019 Dec 23;19(1):354 [FREE Full text] [doi: 10.1186/s12877-019-1375-9] [Medline: 31865906]
35. Wasil AR, Palermo EH, Lorenzo-Luaces L, DeRubeis RJ. Is there an app for that? A review of popular apps for depression, anxiety, and well-being. Cognit Behav Pract 2022 Nov;29(4):883-901. [doi: 10.1016/j.cbpra.2021.07.001]
36. Linardon J, Messer M, Goldberg SB, Fuller-Tyszkiewicz M. The efficacy of mindfulness apps on symptoms of depression and anxiety: an updated meta-analysis of randomized controlled trials. Clin Psychol Rev 2024 Feb;107:102370 [FREE Full text] [doi: 10.1016/j.cpr.2023.102370] [Medline: 38056219]
37. Osborne EL, Ainsworth B, Hooper N, Atkinson MJ. Experiences of using digital mindfulness-based interventions: rapid scoping review and thematic synthesis. J Med Internet Res 2023 Sep 28;25:e44220 [FREE Full text] [doi: 10.2196/44220] [Medline: 37768709]
38. Chittaro L, Vianello A. Evaluation of a mobile mindfulness app distributed through on-line stores: a 4-week study. Int J Hum Comput Stud 2016 Feb;86:63-80. [doi: 10.1016/j.ijhcs.2015.09.004]

XSL•FO
RenderX

39. Huberty J, Green J, Glissmann C, Larkey L, Puzia M, Lee C. Efficacy of the mindfulness meditation mobile app "Calm" to reduce stress among college students: randomized controlled trial. JMIR Mhealth Uhealth 2019 Jun 25;7(6):e14273 [FREE Full text] [doi: 10.2196/14273] [Medline: 31237569]
40. Baumel A, Muench F, Edan S, Kane JM. Objective user engagement with mental health apps: systematic search and panel-based usage analysis. J Med Internet Res 2019 Sep 25;21(9):e14567 [FREE Full text] [doi: 10.2196/14567] [Medline: 31573916]
41. Fernández-Ardèvol M, Sawchuk K, Grenier L. Maintaining connections: octo- and nonagenarians on digital 'use and non-use'. Nordicom Rev 2017;38(S1):39-51. [doi: 10.1515/nor-2017-0396]
42. Pradhan A, Lazar A, Findlater L. Use of intelligent voice assistants by older adults with low technology use. ACM Trans Comput Hum Interact 2020 Sep 16;27(4):1-27. [doi: 10.1145/3373759]
43. Chin J, Desai S, Lin S, Mejia S. Like My Aunt Dorothy: effects of conversational styles on perceptions, acceptance and metaphorical descriptions of voice assistants during later adulthood. Proc ACM Hum Comput Interact 2024 Apr 26;8(CSCW1):1-21. [doi: 10.1145/3637365]
44. Chiesa A. Vipassana meditation: systematic review of current evidence. J Altern Complement Med 2010 Jan;16(1):37-46. [doi: 10.1089/acm.2009.0362] [Medline: 20055558]
45. Vasileiou K, Barnett J, Thorpe S, Young T. Characterising and justifying sample size sufficiency in interview-based studies: systematic analysis of qualitative health research over a 15-year period. BMC Med Res Methodol 2018 Nov 21;18(1):148 [FREE Full text] [doi: 10.1186/s12874-018-0594-7] [Medline: 30463515]
46. Morse JM. Determining sample size. Qual Health Res 2000 Jan 01;10(1):3-5. [doi: 10.1177/104973200129118183]
47. Zhou L, Bao J, Setiawan IM, Saptono A, Parmanto B. The mHealth App Usability Questionnaire (MAUQ): development and validation study. JMIR Mhealth Uhealth 2019 Apr 11;7(4):e11500 [FREE Full text] [doi: 10.2196/11500] [Medline: 30973342]
48. Frank P, Marken M. Developments in qualitative mindfulness practice research: a pilot scoping review. Mindfulness 2021 Sep 29;13(1):17-36. [doi: 10.1007/s12671-021-01748-9]
49. Joseph S, Beer C, Clarke D, Forman A, Pickersgill M, Swift J, et al. Qualitative research into mental health: reflections on epistemology. Ment Health Rev J 2009;14(1):36-42. [doi: 10.1108/13619322200900006]
50. Bashir N. The qualitative researcher: the flip side of the research encounter with vulnerable people. Qual Res 2019 Dec 09;20(5):667-683. [doi: 10.1177/1468794119884805]
51. Shaw RM, Howe J, Beazer J, Carr T. Ethics and positionality in qualitative research with vulnerable and marginal groups. Qual Res 2019 Apr 16;20(3):277-293. [doi: 10.1177/1468794119841839]
52. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251 [FREE Full text] [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
53. Krägeloh CU, Henning MA, Medvedev ON, Feng XJ, Moir F, Billington R, et al. Meta-theories and qualitative methods in mindfulness research. In: Mindfulness-Based Intervention Research: Characteristics, Approaches, and Developments. London, UK: Routledge; 2019:31-56.
54. Baer RA, Smith GT, Lykins E, Button D, Krietemeyer J, Sauer S, et al. Construct validity of the five facet mindfulness questionnaire in meditating and nonmeditating samples. Assessment 2008 Sep 29;15(3):329-342. [doi: 10.1177/1073191107313003] [Medline: 18310597]
55. Lilja JL, Frodi-Lundgren A, Hanse JJ, Josefsson T, Lundh LG, Sköld C, et al. Five Facets Mindfulness Questionnaire--reliability and factor structure: a Swedish version. Cogn Behav Ther 2011 Dec;40(4):291-303. [doi: 10.1080/16506073.2011.580367] [Medline: 21770845]
56. Topp CW, Østergaard SD, Søndergaard S, Bech P. The WHO-5 Well-Being Index: a systematic review of the literature. Psychother Psychosom 2015;84(3):167-176 [FREE Full text] [doi: 10.1159/000376585] [Medline: 25831962]
57. Carter N, Bryant-Lukosius D, DiCenso A, Blythe J, Neville AJ. The use of triangulation in qualitative research. Oncol Nurs Forum 2014 Sep;41(5):545-547. [doi: 10.1188/14.ONF.545-547] [Medline: 25158659]
58. Naeem M, Ozuem W, Howell K, Ranfagni S. A step-by-step process of thematic analysis to develop a conceptual model in qualitative research. Int J Qual Methods 2023 Nov 08;22. [doi: 10.1177/16094069231205789]
59. Saunders CH, Sierpe A, von Plessen C, Kennedy AM, Leviton LC, Bernstein SL, et al. Practical thematic analysis: a guide for multidisciplinary health services research teams engaging in qualitative analysis. BMJ 2023 Jun 08;381:e074256. [doi: 10.1136/bmj-2022-074256] [Medline: 37290778]
60. Tickell A, Ball S, Bernard P, Kuyken W, Marx R, Pack S, et al. The effectiveness of mindfulness-based cognitive therapy (MBCT) in real-world healthcare services. Mindfulness (N Y) 2020 Jan 12;11(2):279-290 [FREE Full text] [doi: 10.1007/s12671-018-1087-9] [Medline: 32064009]
61. Hayes SC, Pierson H. Acceptance and commitment therapy. In: Freeman A, Felgoise SH, Nezu CM, Nezu AM, Reinecke MA, editors. Encyclopedia of Cognitive Behavior Therapy. New York, NY: Springer; 2005:1-11.
62. Li SY, Bressington D. The effects of mindfulness-based stress reduction on depression, anxiety, and stress in older adults: a systematic review and meta-analysis. Int J Ment Health Nurs 2019 Jun;28(3):635-656. [doi: 10.1111/inm.12568] [Medline: 30656813]

XSL•FO
RenderX

63. Perez-Blasco J, Sales A, Meléndez JC, Mayordomo T. The effects of mindfulness and self-compassion on improving the capacity to adapt to stress situations in elderly people living in the community. Clin Gerontol 2015 Dec 16;39(2):90-103. [doi: 10.1080/07317115.2015.1120253]
64. Baer R, Crane C, Miller E, Kuyken W. Doing no harm in mindfulness-based programs: conceptual issues and empirical findings. Clin Psychol Rev 2019 Jul;71:101-114 [FREE Full text] [doi: 10.1016/j.cpr.2019.01.001] [Medline: 30638824]
65. Wright SL, Bach E, Bryson SP, Schaller S, Wright D. Using an app-based mindfulness intervention: a mixed methods approach. Cognit Behav Pract 2025 May;32(2):258-269. [doi: 10.1016/j.cbpra.2024.03.001]
66. Oh VK, Sarwar A, Pervez N. The study of mindfulness as an intervening factor for enhanced psychological well-being in building the level of resilience. Front Psychol 2022 Dec 21;13:1056834 [FREE Full text] [doi: 10.3389/fpsyg.2022.1056834] [Medline: 36619105]

Abbreviations

FFMQ: Five Facet Mindfulness Questionnaire
MBI: mindfulness-based intervention
MBSR: mindfulness-based stress reduction
mHealth: mobile health
WHO-5: World Health Organization-Five Well-Being Index

Edited by A Kushniruk; submitted 14.10.24; peer-reviewed by O Lahtinen, Q Ng; comments to author 24.11.24; revised version received 12.12.24; accepted 12.05.25; published 03.06.25.

Please cite as:
McCarren L, Kuoppamäki S
Design Preferences, Routines, and Well-Being of Older Adults Using Voice-Guided Digital Mindfulness: Qualitative Interview Study
JMIR Hum Factors 2025;12:e67533
URL: https://humanfactors.jmir.org/2025/1/e67533
doi:10.2196/67533
PMID:

©Lucy McCarren, Sanna Kuoppamäki. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 03.06.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Digital Mental Health Intervention for Paranoia (the STOP App): Qualitative Study on User Acceptability

Laura Eid[1], DClinPsy; Alexandra Kenny[2], BA; Rayan Taher[1], MSc; Pamela Jacobsen[3], PhD, DClinPsy; Che Wei Hsu[4,5], PhD, PGDipCIPs; Jenny Yiend[1], PhD

[1]Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience (IoPPN), King's College London, London, United Kingdom
[2]McPin Foundation, London, United Kingdom
[3]Department of Psychology, University of Bath, Bath, United Kingdom
[4]Social Services (Counselling Specialty), College of Community Development and Personal Wellbeing, Otago Polytechnic, Dunedin, New Zealand
[5]Department of Psychological Medicine, Dunedin School of Medicine, University of Otago, Dunedin, New Zealand

Corresponding Author:
Jenny Yiend, PhD
Department of Psychosis Studies
Institute of Psychiatry, Psychology & Neuroscience (IoPPN)
King's College London
Institute of Psychiatry, Psychology & Neuroscience
16 De Crespigny Park
London, SE5 8AF
United Kingdom
Phone: 44 (0)2078 480002
Email: jenny.yiend@kcl.ac.uk

Abstract

Background: Cognitive bias modification for interpretation (CBM-I) is a technique to modify interpretation and is used to reduce unhelpful negative biases. CBM-I has been extensively studied in anxiety disorders where interpretation bias has been shown to play a causal role in maintaining the condition. Successful Treatment of Paranoia (STOP) is a CBM-I smartphone app targeting interpretation bias in paranoia. It has been developed following research on the feasibility and acceptability of a computerized version. This qualitative study extended that research by investigating the acceptability of STOP in individuals with paranoia. The study design and implementation were informed by the Evidence Standards Framework for Digital Health Technologies (DHTs) published by the UK National Institute for Health and Care Excellence (NICE).

Objective: The aim of the study was to involve service users in the design, development, and testing of STOP and understand the degree of satisfaction with the product. We aimed to establish the extent to which STOP met the NICE minimum and best practice standards for DHTs, specifically its acceptability to intended end users.

Methods: In total, 12 participants experiencing mild to moderate levels of paranoia were recruited to complete 6 weekly sessions of STOP before being invited to a feedback interview to share their experiences. Interview questions revolved around the acceptability of the app, its perceived usefulness, and barriers to the intervention, as well as practicality and views on the use of a digital intervention in principle. Interviews were coded and analyzed using the framework analysis method, combining both deductive and inductive approaches.

Results: Framework analysis yielded 6 themes: independent use and personal fit; digital versus traditional approaches; user reactions and emotional impact; impact on thinking, awareness, and well-being; design, engagement, and usability; and intervention relevance and practical fit.

Conclusions: STOP was found to be a broadly acceptable intervention and was positively received by most participants. The study findings are in line with the NICE Evidence Standards Framework for DHTs, as intended end users were involved in the development, design, and testing of STOP and were mostly satisfied with it. These findings will contribute to the further iterative development of this intervention that targets interpretation bias in paranoia.

International Registered Report Identifier (IRRID): RR2-10.1186/s13063-024-08570-3

(JMIR Hum Factors 2025;12:e70181) doi:10.2196/70181

XSL•FO
RenderX

KEYWORDS

framework analysis; interpretation bias; paranoia; qualitative; digital health technologies; cognitive bias modification

Introduction

Interpretation Bias and Cognitive Bias Modification

Paranoia refers to unfounded beliefs that others intend to cause harm and exists on a continuum from mild suspiciousness to severe persecutory delusions [1]. Interpretation bias toward threat, defined as the tendency to interpret ambiguous information in a threatening or negative way, is thought to be one of the key cognitive processes involved in the development and maintenance of paranoia. Specifically, individuals with paranoia are more likely to interpret neutral or ambiguous situations as intentionally harmful, often exhibiting hostility interpretation bias (perceiving others as deliberately malicious) and external-personal attribution bias (attributing negative events to the harmful intent of others), as well as biases related to physical threats (such as those involving security devices or medical interventions) [2-6]. While similar biases are seen in anxiety disorders, paranoia is distinguished by its focus on intentional harm, rather than fear of negative evaluation or embarrassment as seen in social anxiety. Cognitive biases, including interpretation bias, have therefore been identified as critical maintenance factors in psychosis and paranoia, highlighting the need for targeted interventions that address these specific processing styles [1,7,8].

Cognitive bias modification for interpretation (CBM-I) is one such technique, designed to reduce interpretation bias toward threat. Research has demonstrated that it is possible to change bias through training [9,10], helping individuals replace maladaptive cognitive patterns with more neutral or positive interpretations. Successful Treatment of Paranoia (STOP) is a smartphone app that delivers CBM-I through a series of ambiguous scenarios, encouraging the user to generate more adaptive interpretations [11]. CBM-I has shown consistent benefits in reducing anxiety and repetitive negative thinking [12-14], but studies applying this approach to paranoia remain limited. One feasibility randomized controlled trial evaluated a desktop version of cognitive bias modification for paranoia (CBM-pa) [15], and a subsequent qualitative study highlighted the need for improvements in the task's audiovisual design to enhance attention and engagement [16]. These findings directly informed the development of STOP, the smartphone app evaluated in this study.

While STOP draws on CBM principles, it differs significantly from the earlier CBM-pa platform. CBM-pa was delivered in a controlled laboratory setting, whereas STOP is a stand-alone smartphone app designed for independent use. Unlike CBM-pa, STOP enables participants to schedule sessions autonomously, receive automated reminders, and engage with a web-based interface designed to enhance immersion. These structural and experiential differences mean that the findings from the earlier CBM-pa evaluation cannot be assumed to generalize to STOP. A distinct evaluation is therefore warranted.

Thus, a qualitative study, such as the one reported here, is essential to determine whether STOP, in its current form, is a product that would be acceptable to users and suitable for use in routine clinical practice. The importance of gathering this type of evidence for any potential new digital therapy is emphasized in the Evidence Standards Framework for Digital Health Technologies published by the UK National Institute for Health and Care Excellence (NICE) [17], which requires evidence of "acceptability with users" and their involvement in development, design, and testing. The qualitative data were contextualized using the User Experience Questionnaire (UEQ), which quantified participants' experience of the app, such as their perception of its attractiveness, ease of use, and creativity [18]. Therefore, descriptive statistics of the UEQ are reported to indicate the overall experience of participants using the digital health technology (DHT).

In addition to assessing acceptability, a growing body of literature highlights the importance of applying a human factors lens when evaluating DHTs. This includes attention to usability, accessibility, and alignment with user needs and preferences. Human factors research considers not only whether an intervention works but also how individuals interact with it—factoring in context of use, autonomy, and digital literacy; for example, the concept of the digital therapeutic alliance has gained prominence in recent years, recognizing that users may form relational connections with digital tools, which in turn can affect engagement and perceived effectiveness. Evaluating STOP through both experiential and design lenses enables a more comprehensive understanding of its acceptability and potential for real-world implementation.

The Rise and Use of DHTs

Digital interventions have become increasingly common over time, and some have proven to be effective [19]. In a scoping review, Zhang et al [19] identified 22 randomized controlled trials exploring the effect of CBM techniques delivered through a web-based intervention and found that CBM could help reduce biases in anxiety disorders and depression in adolescents.

Despite these advancements, there remains a lack of available treatments specifically targeting psychosis and paranoia. Other than pharmacological interventions, the only treatment that NICE recommends for psychosis is cognitive behavioral therapy, a psychological intervention. Smartphone app interventions offer an alternative treatment approach for paranoia. Additional benefits of using smartphone app interventions include their potential to help address long waiting times to access psychological services and to reduce the costs associated with in-person interventions. Furthermore, digital technologies could be considered low-stigma interventions, as service users can access them privately from any location. They could additionally be seen as empowering, enabling service users to take a more active and autonomous role in their care. More specifically for individuals with paranoia, digital health interventions may facilitate help seeking by reducing the need to build a therapeutic alliance and trust in a therapist, both of which can be especially challenging for this population [20,21].

XSL•FO
RenderX

Aims of This Study

This qualitative study explored users' views and opinions of STOP, a DHT that targets interpretation bias in individuals with paranoia. The study provided an opportunity for individuals reporting distress from paranoia to voice their attitudes and experiences of using STOP and to offer feedback. The purpose of the study was therefore to explore different views and opinions of the digital intervention and to identify key strengths and weaknesses of the DHT as well as potential improvements. More specifically, the aim was to involve end users in the design, development, and testing of STOP and to assess the extent to which they were satisfied with it. The research questions were therefore formulated as follows:

- To what extent does the smartphone app STOP meet the NICE minimum or best practice standards for DHTs?
- Is STOP acceptable to its intended end users?

Methods

Ethical Considerations

This qualitative study received ethics approval from the London-Stanmore Research Ethics Committee (21/LO/0896). Potential participants were emailed a digital copy of the study information sheet and had at least 24 hours to consider participating in the study. All potential participants had the opportunity to ask questions and were free to withdraw at any stage without giving reasons. Interested, eligible participants were then asked to complete an electronic consent. Participants received £40 (US $52) in vouchers as compensation for their time and effort.

Screening and Recruitment Process

Participants were recruited from various sources, including the South London and Maudsley NHS Foundation Trust's Psychological Interventions Clinic for Outpatients with Psychosis; Psychological Interventions Clinic for Outpatients with Psychosis clinic at the South West London and St George's Mental Health NHS Trust, through MQ (MQ Foundation, mental health research charity); and the McPin Foundation (a mental health research charity that specializes in integrating lived experiences in research).

Research assistants working on the main STOP randomized clinical trial (ISRCTN17754650) screened participants against the main trial eligibility criteria. Some of the main trial criteria were not relevant for this study, including the following: if on psychotropic medication, no change for at least the last 3 months; interpretation bias screening score of ≥−2; and Positive and Negative Syndrome Scale (PANSS) [22] item 6 score of >3. Therefore, suitable participants who were ineligible for the main trial were redirected to this qualitative substudy and were then assessed against this study's inclusion and exclusion criteria (Textbox 1 [23]).

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Self-reported paranoid concerns
- Aged >18 y
- Fluent in written English
- Capacity to consent

Exclusion criteria

- Severe cognitive impairment
- Major current physical illness likely to impact participation
- Substance dependence
- Currently receiving another psychological intervention
- Scoring 7 (defined as "extreme") on the Positive and Negative Syndrome Scale
- At high risk of suicide as per the Columbia Suicide Severity Rating Scale [23]

The choice of the sample size was guided by the concept of information power, which suggests that the number of participants needed for qualitative interview studies decreases as the amount of information held by the sample increases [24]. A smaller sample can be sufficient when the study aim is narrow, the sample is specific, the quality of dialogue is strong, and the analysis is theory informed [24]. Given the focused aim of exploring user experiences of STOP and the relevance of the sample to this aim, the final sample size was deemed appropriate for capturing key insights. Of note, patient and public involvement researcher and coauthor AK also took part in this study as a participant.

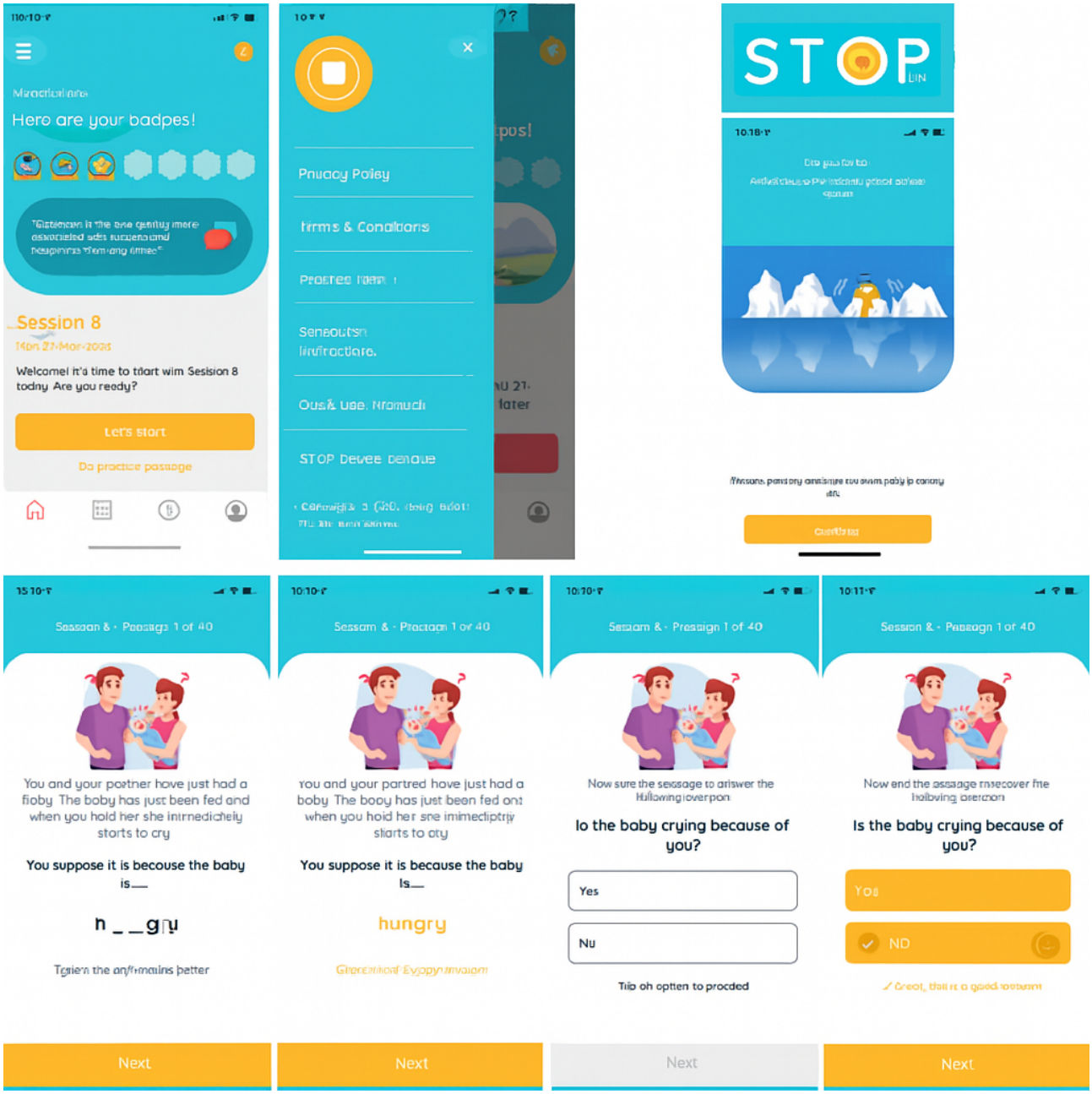
The STOP Intervention

STOP was delivered through a smartphone app. It consisted of 40- to 60-minute sessions (depending on reading speed) in which patients were presented with ambiguous scenarios that could be interpreted in either a threatening or a benign way. Each session consisted of 40 scenarios, presented as blocks of 10 items, followed by trivia items to allow users to take a short break (refer to Figure 1 for an intervention item example, a screenshot of STOP's home page, and a trivia example). After reading each scenario, participants were asked to complete a word with missing letters and a comprehension question, both of which aimed to encourage them to interpret the scenario in

XSL•FO
RenderX

a nonthreatening way. Further details regarding the intervention and its development have been published elsewhere [11].

Figure 1. Screenshot of an example of a Successful Treatment of Paranoia intervention item.



Materials

The UEQ is a validated and reliable scale designed to measure users' experiences of using digital products [25]. It consists of 26 items grouped into 6 scales: attractiveness (overall impression; do users like or dislike the product?), perspicuity (is it easy to learn how to use it?), efficiency (can users complete tasks without unnecessary effort?), dependability (does the user feel in control of the interaction with the product?), stimulation (is it exciting and motivating to use?), and novelty (is it innovative and creative? Does it catch users' interest?). Each item presents a pair of contrasting attributes (eg, annoying-enjoyable and confusing-clear), and respondents are asked to express their agreement with the attributes by selecting the point that most closely reflects their impression of the product. The UEQ demonstrates satisfactory reliability and construct validity [25] and provides normative data from 452 prior digital product evaluations, enabling benchmarking against existing products.

Procedure

Participants who provided consent were sent a study Android smartphone with the STOP app preinstalled and ready for use. In line with medical device study requirements, hard copy instructions to use the smartphone and the app were mailed to participants' home addresses. They were then given the opportunity to test the STOP app and were asked to complete 6 weekly sessions of the STOP intervention. They were also offered the option to meet with the researcher (LE) at any time for assistance with navigating the smartphone or app.

XSL•FO
RenderX

All participants had fortnightly check-in calls with the researcher (LE). They were also given the opportunity to ask questions and were offered support for any technical difficulties they encountered.

After the 6 STOP sessions ended, a time was scheduled for a feedback interview and completion of the UEQ. The feedback interviews were semistructured and lasted between 30 and 60 minutes. A topic guide (Multimedia Appendix 1) was used to structure the interview, while allowing flexibility for additional comments. The topic guide was informed by the NICE Evidence Standards Framework for DHTs [17]. Questions revolved around users' experiences of using the digital intervention, its perceived strengths and weaknesses, their opinions of the visual display of the app as well as the CBM scenarios used, session duration and ease of use, and suggestions for improvement. All interviews were conducted via Microsoft Teams, as all participants had access to a computer and a working internet connection.

After the interviews, participants were asked to complete the UEQ. They could choose to receive a digital version or a printed copy sent to their home address and to complete it either at the beginning of the interview together with the researcher or on their own.

Data Analysis

UEQ quantitative data were aggregated across participants, and mean scores were plotted against normative data [25] to enable within-sector comparisons.

The qualitative data from the interviews were coded and analyzed using the framework method [26]. The NICE Evidence Standards Framework for DHTs [15] was used to inform the topic guide, codes, and deductive themes. To enhance quality and rigor, the Reflexive Thematic Analysis Reporting Guidelines were followed [27].

Interview recordings were autotranscribed by the software and then manually corrected by LE. Data familiarization was achieved through repeated reading of the transcripts alongside any additional notes made during meetings.

Initial coding was surface level and descriptive and was conducted by LE using NVivo 14 (Lumivero).

A working analytical framework was developed by grouping and categorizing codes. New or revised codes were developed iteratively throughout coding. The data were then summarized in a matrix. Several group discussions between LE, JY, and PJ led to agreement on a final analytical framework and corresponding matrix. Themes were exclusively and intentionally related to users' experiences of testing, acceptability, and satisfaction with the STOP app. Deductive codes were informed by the topic guide and the NICE Evidence Standards Framework for DHTs [17]. Inductive codes were driven by the data and reflected participants' unique experiences and language.

Research Team and Reflexivity

A critical realist postpositivist stance was adopted to account for the fact that the coding and interpretation of data would be impacted by the researchers' roles in the study. Reflexivity was therefore viewed not as a means of eliminating subjectivity but of making it visible and critically examined throughout the research process. Indeed, based on previous research [28], most researchers involved in the study held beliefs that the app would help reduce interpretation bias in people with paranoia. In addition, the researchers' own views as well as findings from previous research on individuals' experiences with smartphone apps and CBM [28,29] would likely influence the analysis and interpretation of data. Using the topic guide informed by the NICE guidelines, interview questions were phrased to avoid guiding or stigmatizing language. The researcher attempted to embrace each participant's unique and important experiences, as opposed to trying to reach a consensus. A reflective log was maintained throughout recruitment, data collection, and analysis to record reflective thoughts, which were used to inform analysis and interpretation and understand how personal preconceptions and assumptions may influence findings.

Results

Overview

We contacted 34 participants to complete 6 sessions of the STOP intervention, followed by a feedback interview. Of these 34 individuals, 16 (47%) met the eligibility criteria and consented to take part in the qualitative study. Of these 16 participants, 4 (25%) dropped out (n=3, 75% reported that they did not feel that they had the time to commit to the study; n=1, 25% was uncontactable), resulting in a final sample of 12 (75%) participants. The participants' mean age was 36.33 (SD 10.74) years. Of the 12 participants, 7 (58%) identified as female, 8 (67%) were White, and 2 (17%) identified as being from a minority ethnic background (n=2, 17% did not disclose their ethnicity). The mean score on PANSS item 6 was 3.58 (SD 0.79) [22]. Participants' characteristics and individual scores are summarized in Table 1.

XSL•FO
RenderX

Table 1. Participants' characteristics and User Experience Questionnaire (UEQ) scores.

Participants	Sex	Age (y)	Race and ethnicity	Score on item 6 of PANSS[a]	UEQ attractiveness	UEQ perspicuity	UEQ efficiency	UEQ dependability	UEQ stimulation	UEQ novelty	Sessions completed (n=6), n (%)
1	Male	NR[b]	NR	4	−0.33	−0.25	−0.33	0.25	−0.25	0.00	4 (67)
2	Female	57	White British	3	−1.17	−1.25	−2.00	−0.75	−1.00	−1.00	5 (83)
3	Male	29	Pakistani	3	2.17	2.00	0.75	1.50	1.00	2.25	1 (17)
4	Female	25	Asian	3	1.00	2.00	0.00	0.75	0.75	1.25	3 (50)
5	Male	40	White Irish	4	1.00	1.00	1.50	1.25	0.75	1.25	2 (33)
6	Male	29	White British	5	__[c]	—	—	—	—	—	3 (50)
7	Female	38	NR	4	—	—	—	—	—	—	3 (50)
8	Female	NR	White British	4	—	—	—	—	—	—	3 (50)
9	Female	37	White British	4	1.67	2.00	2.00	1.75	1.75	2.00	5 (83)
10	Female	NR	White British	2	—	—	—	—	—	—	1 (17)
11	Male	47	White British	4	1.33	1.75	2.00	1.50	1.00	0.75	2 (33)
12	Female	25	White British	3	2.67	2.75	1.75	2.00	2.25	1.75	2 (33)

[a]PANSS: Positive and Negative Syndrome Scale (item 6 relates to suspiciousness and persecution).

[b]Not reported.

[c]Missing data.

UEQ Results

After the feedback interview, 8 (67%) of the 12 participants completed the UEQ. The results are provided in Table 1. UEQ scores range from −3 (maximum negative evaluation) through 0 (neutral) to +3 (positive evaluation). A bandwidth of −0.8 to +0.8 indicates neutral evaluation, with scores outside this range indicating overall negative and positive evaluations, respectively. Participants who completed the UEQ in this study scored an average of between 0.9 and 1.5 on the different scales, which indicates that they described their experience as positive. The UEQ indicated generally positive user perceptions across all subscales, with the highest mean scores observed for perspicuity (mean 1.464, SD 1.302), attractiveness (mean 1.238, SD 1.224), and novelty (mean 1.179, SD 1.087). Lower, yet still positive, ratings were recorded for dependability (mean 1.143, SD 0.923), stimulation (mean 0.929, SD 1.018), and efficiency (mean 0.857, SD 1.457). Participant 1 showed inconsistencies on 5 scales (attractiveness, perspicuity, efficiency, dependability, and stimulation; Table 1, row 1) and was therefore excluded from the reported averages and from the scale mean and variety plots in Figures 2 and 3, as suggested by the UEQ handbook [18]. Specifically, the handbook defines inconsistency as a lack of uniformity in item ratings within a scale. While it may be possible for inconsistencies to occur in a single scale, due to a misunderstanding of the question or an error, it is less likely to be the case when they occur across multiple scales. The handbook therefore suggests excluding such participants from analysis because it is more likely that the behavior indicates random or nonserious completion of the questionnaire.

Summary statistics are presented graphically by scale (Figure 2) and benchmarked against digital sector norms (Figure 3).

XSL•FO
RenderX

Figure 2. User Experience Questionnaire (UEQ) results. Mean scores on each scale are shown, with error bars representing variance. Green shading indicates a positive evaluation range, cream shading a neutral evaluation range, and red shading a negative evaluation range.

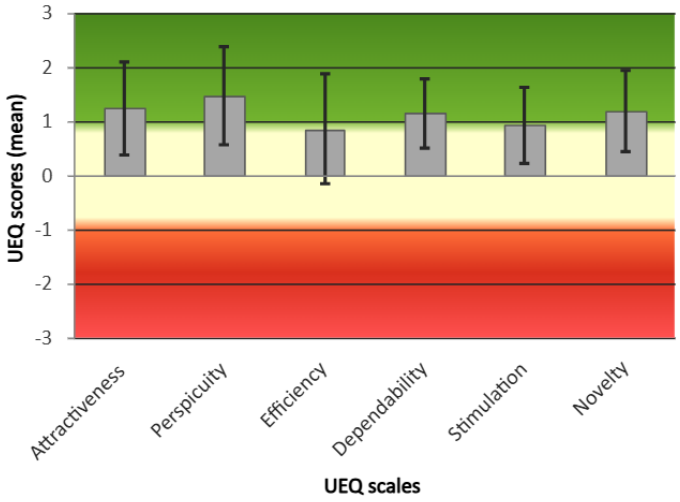
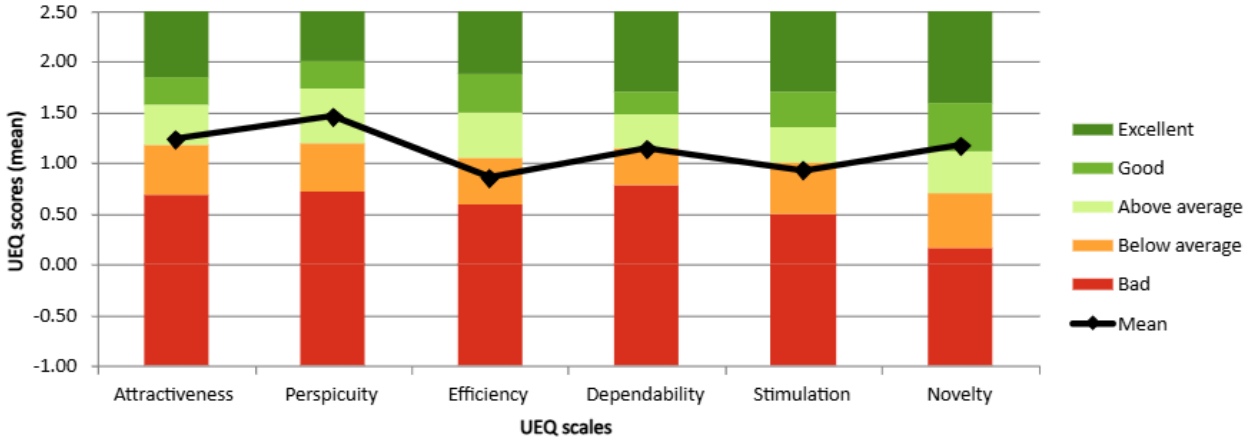


Figure 3. Benchmark graph for Successful Treatment of Paranoia (STOP) based on User Experience Questionnaire (UEQ) scores compared to sector averages. The black data points represent STOP mean scores by scale. The bar colors represent the percentage of other products in the sector falling within that range; for example, a data point falling in pale green ("above average") indicates that 25% of other products scored higher than STOP, whereas a data point falling in orange ("below average") indicates that 50% scored higher.



Qualitative Results

Overview

Through framework analysis of the interview transcripts, 6 themes were identified, along with topic summaries (summarized in the framework analysis): "independent use and personal fit; digital versus traditional approaches; user reactions and emotional impact; impact on thinking, awareness, and well-being; design, engagement, and usability; and intervention relevance and practical fit." Table 2 presents the framework analysis, and Figure S1 in Multimedia Appendix 2 illustrates the relationships between themes.

XSL•FO
RenderX

Table 2. Framework analysis.

Themes and topic summaries	Example quotations
Independent use and personal fit	
Having agency, taking responsibility, self-discipline, and self-growth	• "The advantages are the independence and the self-control and the mindfulness and those skills in themselves, the self-discipline as well, they can [be used] in their own treatments and their own self-discovery journeys." [Participant 10]
Matches lifestyle and preferences	• "So, like an advantage of this is especially for people who are suffering anxiety and PTSD and don't get out." [Participant 5]
Provides privacy, safety, and flexibility	• "When you're on your own, you sort of have time to think and like go through the answers at your own pace and when it suits you." [Participant 12]
Ability, or otherwise, to complete without assistance: neurodiversity, language barrier, and technological proficiency	• "I was able to do it myself but for other people, there's a chance they might need assistance. It could be because they have like dyslexia or there is a language barrier." [Participant 7]
Digital versus traditional approaches	
Advantages of apps over traditional therapies	• "[T]here's less need for human intervention because the app is friendly and fun." [Participant 8]
Disadvantages compared to traditional therapies: digital exclusion, dropout rate, lack of feedback, distractions, and potential dependence on the app	• "You don't get personalized feedback like you would with a person." [Participant 1]
Focused, structured, and specific intervention	• "[Usually a] therapist can guide it, whereas I like the way that this was more structured and forced you to think through different scenarios. And I like that aspect of it." [Participant 1]
User reactions and emotional impact	
Perceived as judgmental, patronizing, and invalidating	• "There is nothing to validate you or to say your initial response actually had some value in it." [Participant 1]
Validating and normalizing	• "It is validating. It's almost as if the researchers have taken into account that we struggle with these thoughts, no matter how insignificant they seem." [Participant 4]
Elicited mixed feelings: confusion, excitement, fear and embarrassment, hope, sadness, or neutral	• "I think I possibly felt more sad at the end of each session." [Participant 9] • "I would look forward to it, it was interesting." [Participant 2] • "It's not good, nor bad." [Participant 7]
Evocative and triggering of actual experiences	• "They're very good at conveying the paranoid thoughts, they were very evocative and brought up quite a lot of feelings when reading them." [Participant 9]
Perceived as restrictive	• "One negative thing was that it doesn't let you say no, I don't agree with that." [Participant 5]
Impact on thinking, awareness, and well-being	
Helpful impact on daily life and thinking patterns, sense of achievement, and general positive feelings	• "I found it really helpful...it has been giving me some clues like some other way of thinking other than like me, always overthinking what other people's intentions are." [Participant 4]
Offers new perspectives	• "I found one of the big values of the app is identifying other options and other ways of looking at things." [Participant 1]
Self-reflection, awareness, and understanding of paranoia and its presence in the user's life	• "It has made me realize how insidious paranoia and anxiety are in my everyday life via the scenarios, for which I was always instinctively drawn to the negative assumptions, and so many of which I regularly worry about." [Participant 9]

XSL•FO
RenderX

Themes and topic summaries	Example quotations
Design, engagement, and usability	
Design and look of the app	• "I think the color scheme is very calming and soothing...I think it's quite good like, it's not really intimidating or like causing people distress." [Participant 4]
Engagement with the app	• "Intuitive, really well designed and written, and very engaging." [Participant 9] • "It was a bit like a reading comprehension for a 7-year-old." [Participant 10]
Suggestions for improvement: managing expectations, personalizing the experience, suggestions for design changes, and access to the app	—[a]
Intervention relevance and practical fit	
Acceptability and testing of scenarios	• "So, about half I think were quite pretty easy to identify with and about half I had to work a bit harder." [Participant 1] • "I thought they're all very relatable...there wasn't a single one that I thought it seemed odd or shouldn't be in there." [Participant 8]
Perceived difficulty of the intervention	• "Yeah, I guess the more effort just like you'd have to think about it more to try and work out what that word is instead of like 'Ohh well, clearly it has this like one word that fits in there.'" [Participant 11] • "It was like there's always a very obvious word that fits in...if it was more challenging, I'd probably focus more on the sentence." [Participant 12] • "I had to think about it more, and that's not necessarily a bad thing...those were the scenarios which helped me the most." [Participant 1]
Session frequency and duration	• "I would say it's just about right, but it's not a little ask. It took me, I would say just over half an hour...that's still a fair commitment. Possibly for some people would need it to be a bit shorter." [Participant 1] • "I would like it to be on a continuous basis." [Participant 4] • "It was quite long as well. The sessions were longer than expected...Maybe if the session were half the length as well, then I wouldn't have got bored so quick so easily." [Participant 10]
Views and suggestions related to implementation	• "I'm severely disabled, that's what's unrelatable. Maybe the app is aimed at people who have less severe disabilities. And yet I did find it helpful." [Participant 11]

[a]Suggestions are presented in Multimedia Appendix 3.

Theme 1: Independent Use and Personal Fit

This theme depicted participants' attitudes toward completing the intervention on their own and explored the extent to which an autonomous treatment was acceptable. An autonomous treatment was defined here as an intervention to be completed independently, meaning on one's own and without additional support from professionals or other individuals, aside from any necessary practical or safety checks (eg, in STOP, fortnightly safety check-in calls with the researcher were required). This theme captured users' views on both obstacles and benefits. Overall, participants seemed satisfied with the intervention being an autonomous treatment.

Some participants indicated that completing STOP on their own gave them a sense of agency and allowed them to take responsibility for their issues and for their progress by deciding to engage with the app and complete sessions on a weekly basis. Some described that STOP helped them monitor their progress and visualize it in the form of badges that they accumulated as they completed sessions. Others thought it could positively impact their self-confidence and allow them to grow:

> *I would say responsibility because...one of the things that can really happen is that you give up responsibility or you lack confidence, whereas with this it's like I'm trying, I'm doing it.* [Participant 1]

Some participants reported satisfaction with the way the intervention was delivered as it fit their preferences and lifestyles. Often, the preferences were associated with mental health conditions getting in the way of seeking help. These participants reported that engaging in a smartphone app was helpful in that it allowed them to access help from their homes. For others, the intervention being an autonomous treatment fit with their personalities and offered them privacy and the space to be honest with themselves as well as the option to complete it at their own pace:

XSL•FO
RenderX

> *It gives you the independence, the privacy, the self-motivation and discipline.* [Participant 10]

Theme 2: Digital Versus Traditional Approaches

This theme described participants' acceptability of STOP compared to other interventions, such as cognitive behavioral therapy or counselling. It captured users' views on both obstacles and benefits related to the intervention's delivery through a smartphone app. Compared to other treatments, participants seemed to be partly satisfied with the intervention, particularly the fact that it offered flexibility and was specifically focused on paranoia:

> *This treatment is different because it's very specific to paranoia and I have not seen any treatment specific to paranoia before.* [Participant 8]

A few noted disadvantages of smartphone apps in general. These included the need for an internet connection and concerns about digital exclusion. In addition, some of the participants highlighted the high dropout rates associated with app-based interventions, while others reported that it can be distracting, potentially lead to dependence on the app, or not be in line with their preferences. Several comments revolved around the lack of human contact and feedback associated with using an app:

> *I mean there is an element, I suppose, of the isolation being extended and not having any human support around you, that you would miss out on.* [Participant 10]

Theme 3: User Reactions and Emotional Impact

This theme depicted the extent to which users were generally satisfied with the app after having completed a few sessions. It encompassed a wide range of emotions and reactions that were elicited and labeled. Two contrasting views emerged in the interviews. Most participants described their experiences as positive; however, some expressed negative views, particularly perceiving the app as invalidating or restrictive because it did not allow them the option to disagree with the answers provided within the therapeutic content:

> *You [are] only allow[ed] to answer the question one way. So not only is it restricting, what are you honestly expecting everyone to have the same opinion?...I can't give you my opinion if you don't let me answer the question in a way I feel right in answering it.* [Participant 6]

> *I found it validating.* [Participant 8]

Overall, emotions ranged from confusion and sadness to hopefulness and excitement. Such a range of emotions, even at an individual level, highlighted the unique experience of each participant and the complexities involved. Some participants explored the feelings elicited and used the opportunity to understand themselves better. Others reported that the scenarios presented in the app were evocative because of their relatability. A few described the experience as triggering their paranoia (which is an expected and possible element of STOP's therapeutic process), whether due to the content of the scenarios or the use of an external smartphone provided by the researcher. However, this was not reported as particularly harmful or distressing. Some participants described STOP as validating because it normalized paranoid thoughts as well as the negative interpretations commonly associated with paranoia.

Theme 4: Impact on Thinking, Awareness, and Well-Being

This theme described users' satisfaction with the app in terms of its impact on their daily lives, changes in thinking patterns, self-reflection, and awareness. Most participants reported that STOP had been helpful in some way.

Some described the impact as emerging from the fact that the app served as a distraction, while others described a sense of achievement from completing the weekly sessions. Some participants highlighted the positive impact on their daily lives, noting changes in their thinking patterns. Some described feeling generally more positive and less stressed after using the app. Some participants, including a few who described the app as restrictive, found it helpful in that it offered alternative ways to interpret situations that they would usually interpret in a threatening or negative manner.

Participants explicitly and spontaneously reported that they perceived the app as a tool for modifying their thinking patterns and also noted that it encouraged reflection and helped them identify themes of paranoia that they related to. They described STOP as helpful in recognizing personal thinking patterns and understanding the extent to which paranoia features in their day-to-day life. Some participants disclosed that the app helped them gain a better understanding of paranoia and how it is perceived by the general population.

Theme 5: Design, Engagement, and Usability

This theme outlined participants' views of the design of the app and explored the extent to which this would be acceptable and satisfying to end users. Participants also spontaneously shared their suggestions for improvement.

Most participants reported feeling satisfied with the design, look, and feel of STOP, including the trivia, audiovisual features, and scheduling and reminder functions. They expressed that they found it "clear" (participant 7) and "not overloading of information" (participant 5). Some reported that they enjoyed the color schemes, the images, and the fact that "it involves all senses" (participant 10), especially if one had "poor eyesight" (participant 1). One participant noted that the images used reinforced binary gender stereotypes, which made the app less relatable. Most users described the app as engaging and interesting, easy to use, and friendly, although a few found it childish. One participant found the app helpful for addressing stigma, noting that its anonymous delivery made it feel discreet. In contrast, another participant felt that the app's colors and sounds made it stand out and attract unwanted attention (the sounds can be switched off within the app).

Although STOP seemed largely well accepted, participants offered suggestions for improvement. The most common suggestions revolved around offering more response options for the scenarios, personalizing the app, and encouraging more flexible use. Suggestions also included managing expectations and providing more introductory and contextual information

XSL•FO
RenderX

about the intervention's aims, mode of action, and length. Some participants offered alternative design ideas that they believed would make the app more attractive and could help address possible issues of digital exclusion. Additional suggestions included making the app more stimulating and challenging to avoid boredom and to increase engagement:

> *If it was me, I would leave a section for another option where you could put your own feedback into, like feedback on what you think of the question or if your idea on the scenario didn't fit into the boxes, boxes of yes or no.* [Participant 6]

Theme 6: Intervention Relevance and Practical Fit

The final theme focused on the core intervention and how it was experienced by users. It reflected the extent to which STOP was perceived as therapeutic, generally acceptable, and satisfying. Some participants indicated that they could strongly relate to the scenarios, while others found some scenarios personally irrelevant. A few reported feeling marginalized or excluded when the scenarios did not reflect the experiences of individuals with disabilities, restrictive lifestyles, or different lifestyle preferences. Some felt that the intervention was too easy and needed to be more stimulating and challenging. They found the more difficult or less relatable scenarios to be the most beneficial. Participants also noted that automatically and systematically completing the missing-letter tasks discouraged them from putting in the effort and therefore did not impact their way of thinking in a positive way.

Most participants commented on the length of the intervention, the frequency of sessions, and the duration of each session. In terms of session duration, most participants felt that the sessions were long and that they would benefit from shorter sessions. Some reported that the session duration interfered with their day-to-day activities and was difficult to fit into their busy schedules, while others said that it suited their lifestyle. Some participants highlighted the benefits of having a weekly session and compared it to therapy, while others expressed a wish to complete the content at their own pace and in shorter blocks throughout the week. In terms of the length of the intervention, some participants reported that 6 sessions would be enough in terms of changing thinking patterns, while others suggested more sessions to enhance the learning. Some participants offered suggestions related to the core intervention, such as shortening the duration of individual sessions, amending the frequency of sessions and length of the intervention, encouraging participants to complete the session, offering the intervention on an indefinite basis, integrating some human contact and additional support if needed, or targeting a specific population.

Other suggestions were related to how participants would envisage the implementation of STOP outside of a research trial:

> *Well, I guess like if the app was going to sort of develop, maybe integrating some sort of like breathing techniques or something. If things get too much, almost like a little like panic button crisis button on it or something.* [Participant 12]

> *I think it should be prescribed by your GP or your psychiatrist or your psychologist or occupational therapist. You should be signposted to it and there's a mandatory one-to-one onboarding session to answer any questions, to explain and yeah.* [Participant 8]

The full list of recommendations arising from themes 5 and 6 is provided in Multimedia Appendix 3.

Discussion

Principal Findings

This study explored the acceptability of STOP for individuals struggling with paranoia, in the context of the NICE Evidence Standards Framework for DHTs. Overall, participants representing the intended end users expressed broadly positive views of the smartphone app, describing aspects of the experience and design as acceptable and well thought out. These findings are consistent with a previous qualitative study exploring the acceptability of an earlier, computer-based version of the intervention [16]. However, given the modest engagement levels and variation in participant feedback, these results should be interpreted cautiously and viewed as indicative rather than conclusive.

More specifically, after up to 6 weeks of using STOP, participants acknowledged its potential as an autonomous treatment that was not directly supported by a therapist. Most participants reported that the way the intervention was delivered fit their preferences and offered them flexibility and privacy, in line with the perceived advantages of internet-based interventions in previous studies [30,31].

The transdiagnostic approach adopted by STOP avoids the use of stigmatizing diagnostic labels and targets a single symptom by seeking to alter its associated underlying maintaining mechanism.

Our findings suggest that the autonomy of STOP (theme 1) gave users a sense of agency, in contrast to the lack of empowerment frequently experienced by individuals with severe mental health conditions. In this respect, our findings concur with those of others suggesting that DHTs can help empower individuals, offering them more control of their condition and a better understanding of it [32,33].

Of the 12 participants, 1 (8%) reported concerns about data protection, which is common in the context of using DHTs in similar populations [33]. Digital therapies designed for use with those experiencing paranoia will need to pay special attention to transparency and reassurance around data privacy and security if these technologies are to benefit this user group.

Some participants reported that factors related to technical proficiency, neurodiversity, and individual preferences may interfere with users' ability to benefit from the intervention, highlighting the need to address these considerations in future versions of STOP. While only 1 (8%) of the 12 participants explicitly reported concerns about digital exclusion, this is a commonly reported barrier in the digital mental health literature. By contrast, other circumstances, such as interpersonal challenges, can lead to digital technology improving access to

XSL•FO
RenderX

appropriate health care [33,34]. It is likely, then, that some health inequalities may be mitigated by DHTs such as STOP, while others will inevitably be introduced. For this reason, STOP, and DHTs in general, are best seen not as a panacea but as a complement to traditional therapies, with the potential to fill some important gaps in provision.

Some disadvantages of using STOP were also reported (captured in the Theme 2: Digital Versus Traditional Approaches subsection), the most common one being the lack of human contact and feedback—absent from many DHTs and often reported in the literature [32]. Participants offered some suggestions to resolve this perceived disadvantage, such as providing the option to contact a mental health professional as needed while using the app or combining sessions of STOP with traditional psychological therapy. The preference for having alternative options to seek help seems to be common in the context of DHT use [33]. Acceptability was also related to user experience (theme 3). Some participants perceived STOP as judgmental and restrictive, mainly due to the language used within the app (despite it having been codeveloped with users). Others described the intervention as validating and normalizing. Some participants held what seemed to be opposing views of the experience, such as frustration and validation or sadness and hope. This may reflect a common experience of therapy, in which patients may have ambiguous and contradicting feelings toward the approach used, the process of therapy, or their therapist. However, in traditional therapy, the therapeutic alliance is examined, feelings about therapy explored, and approaches adapted to the patients' needs and feedback. Such a process may not be achievable in digital interventions where the negative aspects of the experience are not addressed in the moment. Future digital therapies may benefit from accommodating this need, for example, by giving participants the opportunity to report their experience in real time. Indeed, the concept of the "digital therapeutic alliance" is gaining traction in psychotherapy literature and provides a suitable theoretical platform from which to explore and improve the digital therapeutic experience [35].

Moreover, some users reported noticing a positive impact of STOP on their lives (theme 4), for example, in terms of their well-being, their thinking patterns, self-reflection, understanding, and awareness of their paranoia. Users also reported general satisfaction with both the core therapeutic intervention and its digital interface (themes 5 and 6). However, some participants expressed that they found a 30- to 40-minute session too long to complete in 1 sitting. These findings echo previous literature on CBM, which has noted that the approach can feel repetitive and monotonous [28,36]. Participants offered some suggestions for improvement, such as shortening the intervention, integrating some additional support, amending the duration of each session, personalizing the scenarios, or making the intervention more challenging and stimulating to optimize changes in thinking patterns.

As shown in Figure 3, STOP received average or above average ratings for attractiveness, perspicuity, dependability, and novelty, but it scored below average on efficiency and stimulation. These findings align with qualitative themes in which participants described the task as repetitive or lacking engagement and occasionally reported that the interface did not facilitate smooth or efficient use. The absence of "good" or "excellent" ratings across scales may reflect the early-stage design of STOP and highlight areas for improvement, particularly in terms of enhancing engagement and interactivity. Given the exploratory nature of this qualitative study and the limited UEQ response rate, we present these results descriptively and cautiously. We have clarified this point in the manuscript and reflected on how this feedback can inform future iterations of STOP.

Finally, it is important to note that engagement with the intervention was low, with only 3 (25%) of the 12 participants completing ≥4 sessions and none completing the full 6-session program. Previous reviews of unguided digital mental health tools have consistently highlighted high dropout rates, with engagement often declining sharply after initial use [37]. In this context, the lack of full intervention completion among our participants aligns with broader trends but still raises important questions about the real-world feasibility of STOP as a stand-alone tool.

Comparisons with similar digital interventions show mixed engagement outcomes; for example, the EMPOWER study reported relatively high engagement when supported by peer involvement [38], while other unguided CBM interventions have faced attrition challenges [28,36]. These findings suggest that embedding some form of human support may be critical for boosting sustained use, particularly in populations with paranoia. Future versions of STOP may therefore benefit from exploring hybrid models that balance autonomy with light-touch support or therapeutic onboarding. Furthermore, the role of the digital navigator is becoming increasingly recognized as a valuable component in the delivery of digital interventions, and this will be an important element to explore in future iterations of STOP [39].

Strengths and Limitations

A strength of this study is that a critical realist perspective [40] was adopted during the coding and analysis stages, allowing the researchers to reflect on their views and experiences throughout the study and understand how these may have influenced the analytic process. In line with this perspective, a reflective log was maintained throughout the study. To account for the possibility of response and desirability bias during the interviews, the researcher (LE) attempted to normalize all points of view, including negative experiences of using the app. This was achieved by using open-ended questions and encouraging participants to share their own experiences. To avoid steering responses toward the researcher's expectations, participants were prompted and guided only when necessary. Finally, the researcher adopted a stance of curiosity and nonjudgment during the interviews to encourage participants to share their views openly and without fear of judgment. To our knowledge, this study is the first to explore the acceptability of a digital health intervention delivered via smartphone that targets interpretation bias in individuals with paranoia, in line with the NICE Evidence Standards Framework.

The findings need to be understood in light of possible limitations. Most participants reported mild levels of paranoia,

XSL•FO
RenderX

as indicated by their PANSS scores (Table 1). The sample does not therefore reflect the full spectrum of paranoia that will be experienced by the intended end users. In addition, there was an underrepresentation of ethnic diversity in the sample because most of the participants (7/12, 58%) identified as White British. Moreover, none of the participants completed the full 6-week intervention before the interview. One possible barrier may be related to difficulties engaging with apps in psychosis [41], particularly for service users with more severe paranoid symptoms. This partial exposure may have influenced their views on usability, engagement, and perceived benefit. However, even limited engagement was sufficient for participants to form meaningful opinions about key features of the intervention, and themes were largely consistent across those who completed more versus fewer sessions. Low completion rates are not uncommon in digital interventions for psychosis and may reflect broader challenges around attention, motivation, and trust in technology—particularly among individuals with more severe symptoms [41]. Participant feedback in this study pointed to task repetitiveness, lack of stimulation, and external factors (eg, life stressors and environment) as potential barriers. These insights are valuable for informing future iterations of STOP. Future studies could focus on implementing the suggestions offered by participants and reassessing acceptability in line with NICE guidelines. In addition, it could be helpful to explore the acceptability of different intervention doses (eg, 6 wk vs 12 wk), particularly in relation to the efficacy findings reported for each dose in the main trial [42]. It will also be important to understand clinicians' experiences, opinions, and recommendations regarding STOP, both to improve the intervention further and to assess its "credibility with UK health and social care professionals," as recommended in the NICE Evidence Standards Framework [17]. One other limitation of this study is that, although participants self-reported experiences of paranoia, their interpretation bias and paranoid symptoms were not formally assessed. While this means that our sample did not correspond to that of the main efficacy trial, it was arguably a more appropriate reflection of the likely population when in clinical use. Finally, the absence of contributors offering external critical perspectives represents a further limitation, which we acknowledge as an area for improvement in future work.

Conclusions

This study provides evidence of the central involvement of users experiencing paranoia in the qualitative evaluation and further development of the STOP digital mental health intervention. Participants were generally satisfied with STOP, as reflected in the quantitative survey data (UEQ scores) and views expressed during the qualitative interviews. They found the app broadly acceptable and were mostly satisfied with the way it was delivered, the core intervention itself, and the impact it had on their lives. The findings were in line with the target expectations outlined in the NICE Evidence Standards Framework for DHTs [17]. Participants reported that the intervention provided insight and awareness of their paranoid thinking patterns and offered them alternative ways of interpreting situations they related to. Participants also offered some suggestions for improvement that fit with their preferences and condition. This study therefore incorporates suggestions for future development. It will be important to take the findings into account if STOP is to succeed as a new potential intervention targeting interpretation bias toward threat in individuals with paranoia.

Acknowledgments

This work was supported by the Medical Research Council Biomedical Catalyst: Developmental Pathway Funding Scheme (MR/V027484/1). The authors would like to express their gratitude to the National Institute for Health and Care Research Biomedical Research Centre hosted at South London and Maudsley NHS Foundation Trust in partnership with King's College London. The views expressed are those of the authors and not necessarily those of the National Health Service, the National Institute for Health and Care Research, the Department of Health and Social Care, or King's College London. The authors acknowledge with thanks the contributions made to this work by the Successful Treatment of Paranoia study lived experience panel supported by the McPin Foundation. The authors would like to thank Avegen Health, the digital health care product development company that developed and hosted Successful Treatment of Paranoia throughout this work.

Disclaimer

For the purposes of open access, the author has applied a Creative Commons attribution (CC BY) license to any accepted author manuscript version arising from this submission.

Authors' Contributions

Conceptualization: LE (lead), PJ (equal), JY (equal), AK (supporting), CWH (supporting), and RT (supporting).
Data curation: LE.
Formal analysis: LE.
Funding acquisition: LE, PJ, and JY.
Investigation: LE.
Methodology: LE, PJ, JY, AK, CWH, and RT.
Project administration: LE, AK, and RT.
Resources: JY.

XSL•FO
RenderX

Supervision: JY and PJ.
Validation: JY and PJ.
Visualization: LE and RT.
Writing – original draft: LE, PJ, JY, AK, CWH, and RT.
Writing – review & editing: LE, PJ, JY, AK, CWH, and RT.

Conflicts of Interest

Patient and public involvement researcher and coauthor AK took part as a participant in this study.

Multimedia Appendix 1
Conceptual relationship of themes.
[DOCX File , 187 KB - humanfactors_v12i1e70181_app1.docx]

Multimedia Appendix 2
Semistructured interview topic guide.
[DOCX File , 145 KB - humanfactors_v12i1e70181_app2.docx]

Multimedia Appendix 3
User interview recommendations for improvement.
[DOCX File , 36 KB - humanfactors_v12i1e70181_app3.docx]

References

1. Freeman D. Suspicious minds: the psychology of persecutory delusions. Clin Psychol Rev 2007 May;27(4):425-457. [doi: 10.1016/j.cpr.2006.10.004] [Medline: 17258852]
2. Clark DM. Anxiety disorders: why they persist and how to treat them. Behav Res Ther 1999 Jul;37 Suppl 1:S5-27. [doi: 10.1016/s0005-7967(99)00048-0] [Medline: 10402694]
3. Freeman D, Garety PA. Worry, worry processes and dimensions of delusions: an exploratory investigation of a role for anxiety processes in the maintenance of delusional distress. Behav Cogn Psychother 1999 Jan 01;27(1):47-62. [doi: 10.1017/s135246589927107x]
4. Chen J, Short M, Kemps E. Interpretation bias in social anxiety: a systematic review and meta-analysis. J Affect Disord 2020 Nov 01;276:1119-1130. [doi: 10.1016/j.jad.2020.07.121] [Medline: 32777650]
5. Trotta A, Kang J, Stahl D, Yiend J. Interpretation bias in paranoia: a systematic review and meta-analysis. Clin Psychol Sci 2020 Oct 21;9(1):3-23. [doi: 10.1177/2167702620951552]
6. Yiend J, Allen P, Lopez ND, Falkenberg I, Tseng H, McGuire P. Negative interpretation biases precede the onset of psychosis. Behav Ther 2019 Jul;50(4):718-731. [doi: 10.1016/j.beth.2018.11.002] [Medline: 31208682]
7. Savulich G, Edwards A, Assadi S, Guven H, Leathers-Smith E, Shergill S, et al. Biased interpretation in paranoia and its modification. J Behav Ther Exp Psychiatry 2020 Dec;69:101575. [doi: 10.1016/j.jbtep.2020.101575] [Medline: 32505999]
8. Henry JD, Von Hippel C, Ruffman T, Perry Y, Rendell PG. Threat perception in schizophrenia-spectrum disorders. J Int Neuropsychol Soc 2010 Sep;16(5):805-812. [doi: 10.1017/S1355617710000640] [Medline: 20604986]
9. Yiend J, Mackintosh B, Mathews A. Enduring consequences of experimentally induced biases in interpretation. Behav Res Ther 2005 Jun;43(6):779-797. [doi: 10.1016/j.brat.2004.06.007] [Medline: 15890169]
10. Mathews A, Mackintosh B. Induced emotional interpretation bias and anxiety. J Abnorm Psychol 2000 Nov;109(4):602-615. [Medline: 11195984]
11. Hsu CW, Stahl D, Mouchlianitis E, Peters E, Vamvakas G, Keppens J, et al. User-centered development of STOP (successful treatment for paranoia): material development and usability testing for a digital therapeutic for paranoia. JMIR Hum Factors 2023 Dec 08;10:e45453 [FREE Full text] [doi: 10.2196/45453] [Medline: 38064256]
12. Hirsch CR, Krahé C, Whyte J, Bridge L, Loizou S, Norton S, et al. Effects of modifying interpretation bias on transdiagnostic repetitive negative thinking. J Consult Clin Psychol 2020 Mar;88(3):226-239. [doi: 10.1037/ccp0000455] [Medline: 32068424]
13. Jones EB, Sharpe L. Cognitive bias modification: a review of meta-analyses. J Affect Disord 2017 Dec 01;223:175-183. [doi: 10.1016/j.jad.2017.07.034] [Medline: 28759865]
14. Martinelli A, Grüll J, Baum C. Attention and interpretation cognitive bias change: a systematic review and meta-analysis of bias modification paradigms. Behav Res Ther 2022 Oct;157:104180. [doi: 10.1016/j.brat.2022.104180] [Medline: 36037642]
15. Yiend J, Trotta A, Meek C, Dzafic I, Baldus N, Crane B, et al. Cognitive Bias Modification for paranoia (CBM-pa): study protocol for a randomised controlled trial. Trials 2017 Jun 29;18(1):298 [FREE Full text] [doi: 10.1186/s13063-017-2037-x] [Medline: 28662715]

XSL•FO
RenderX

16. Leung CJ, Fosuaah A, Frerichs J, Heslin M, Kabir T, Lee TM, et al. A qualitative study of the acceptability of cognitive bias modification for paranoia (CBM-pa) in patients with psychosis. BMC Psychiatry 2019 Jul 23;19(1):225 [FREE Full text] [doi: 10.1186/s12888-019-2215-3] [Medline: 31337373]
17. Evidence standards framework for digital health technologies. National Institute for Health and Care Excellence. 2019. URL: https://www.nice.org.uk/Media/Default/About/what-we-do/our-programmes/evidence-standards-framework/digital-evidence-standards-framework.pdf [accessed 2024-11-19]
18. Schrepp M. User experience questionnaire handbook version 11. Research Gate. URL: https://www.researchgate.net/publication/281973617_User_Experience_Questionnaire_Handbook [accessed 2025-05-29]
19. Zhang M, Ying J, Song G, Fung DS, Smith H. Web-based cognitive bias modification interventions for psychiatric disorders: scoping review. JMIR Ment Health 2019 Oct 24;6(10):e11841 [FREE Full text] [doi: 10.2196/11841] [Medline: 31651410]
20. Yanos PT, Roe D, Lysaker PH. Narrative enhancement and cognitive therapy: a new group-based treatment for internalized stigma among persons with severe mental illness. Int J Group Psychother 2011 Oct;61(4):577-595 [FREE Full text] [doi: 10.1521/ijgp.2011.61.4.576] [Medline: 21985260]
21. Malla AK, Takhar JJ, Norman RM, Manchanda R, Cortese L, Haricharan R, et al. Negative symptoms in first episode non-affective psychosis. Acta Psychiatr Scand 2002 Jun;105(6):431-439. [doi: 10.1034/j.1600-0447.2002.02139.x] [Medline: 12059847]
22. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr Bull 1987;13(2):261-276. [doi: 10.1093/schbul/13.2.261] [Medline: 3616518]
23. Posner K, Brown GK, Stanley B, Brent DA, Yershova KV, Oquendo MA, et al. The Columbia-suicide severity rating scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. Am J Psychiatry 2011 Dec;168(12):1266-1277 [FREE Full text] [doi: 10.1176/appi.ajp.2011.10111704] [Medline: 22193671]
24. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res 2016 Nov;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
25. Laugwitz B, Held T, Schrepp M. Construction and evaluation of a user experience questionnaire. In: Proceedings of the 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer. 2008 Presented at: USAB '08; November 20-21, 2008; Graz, Austria p. 63-76 URL: https://doi.org/10.1007/978-3-540-89350-9_6 [doi: 10.1007/978-3-540-89350-9]
26. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013 Sep 18;13:117 [FREE Full text] [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]
27. Braun V, Clarke V. Supporting best practice in reflexive thematic analysis reporting in Palliative Medicine: a review of published research and introduction to the Reflexive Thematic Analysis Reporting Guidelines (RTARG). Palliat Med 2024 Jun;38(6):608-616 [FREE Full text] [doi: 10.1177/02692163241234800] [Medline: 38469804]
28. Beard C, Weisberg RB, Primack J. Socially anxious primary care patients' attitudes toward cognitive bias modification (CBM): a qualitative study. Behav Cogn Psychother 2012 Oct;40(5):618-633 [FREE Full text] [doi: 10.1017/S1352465811000671] [Medline: 22127022]
29. Steare T, Giorgalli M, Free K, Harju-Seppänen J, Akther S, Eskinazi M, et al. A qualitative study of stakeholder views on the use of a digital app for supported self-management in early intervention services for psychosis. BMC Psychiatry 2021 Jun 19;21(1):311 [FREE Full text] [doi: 10.1186/s12888-021-03317-9] [Medline: 34147075]
30. Wallin EE, Mattsson S, Olsson EM. The preference for internet-based psychological interventions by individuals without past or current use of mental health treatment delivered online: a survey study with mixed-methods analysis. JMIR Ment Health 2016 Jun 14;3(2):e25 [FREE Full text] [doi: 10.2196/mental.5324] [Medline: 27302200]
31. Beattie A, Shaw A, Kaur S, Kessler D. Primary-care patients' expectations and experiences of online cognitive behavioural therapy for depression: a qualitative study. Health Expect 2009 Mar;12(1):45-59 [FREE Full text] [doi: 10.1111/j.1369-7625.2008.00531.x] [Medline: 19250152]
32. Bucci S, Morris R, Berry K, Berry N, Haddock G, Barrowclough C, et al. Early psychosis service user views on digital technology: qualitative analysis. JMIR Ment Health 2018 Oct 31;5(4):e10091 [FREE Full text] [doi: 10.2196/10091] [Medline: 30381280]
33. Berry N, Lobban F, Bucci S. A qualitative exploration of service user views about using digital health interventions for self-management in severe mental health problems. BMC Psychiatry 2019 Jan 21;19(1):35 [FREE Full text] [doi: 10.1186/s12888-018-1979-1] [Medline: 30665384]
34. Borghouts J, Eikey E, Mark G, De Leon C, Schueller SM, Schneider M, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. J Med Internet Res 2021 Mar 24;23(3):e24387 [FREE Full text] [doi: 10.2196/24387] [Medline: 33759801]
35. Lederman R, D'Alfonso S. The digital therapeutic alliance: prospects and considerations. JMIR Ment Health 2021 Jul 20;8(7):e31385 [FREE Full text] [doi: 10.2196/31385] [Medline: 34283035]
36. Brosan L, Hoppitt L, Shelfer L, Sillence A, Mackintosh B. Cognitive bias modification for attention and interpretation reduces trait and state anxiety in anxious patients referred to an out-patient service: results from a pilot study. J Behav Ther Exp Psychiatry 2011 Sep;42(3):258-264. [doi: 10.1016/j.jbtep.2010.12.006] [Medline: 21349242]

XSL•FO
RenderX

37. Torous J, Lipschitz J, Ng M, Firth J. Dropout rates in clinical trials of smartphone apps for depressive symptoms: a systematic review and meta-analysis. J Affect Disord 2020 Feb 15;263:413-419 [FREE Full text] [doi: 10.1016/j.jad.2019.11.167] [Medline: 31969272]
38. Gumley A, Bradstreet S, Ainsworth J, Allan S, Alvarez-Jimenez M, Beattie L, et al. Early signs monitoring to prevent relapse in psychosis and promote well-being, engagement, and recovery: protocol for a feasibility cluster randomized controlled trial harnessing mobile phone technology blended with peer support. JMIR Res Protoc 2020 Jan 09;9(1):e15058 [FREE Full text] [doi: 10.2196/15058] [Medline: 31917372]
39. Wisniewski H, Gorrindo T, Rauseo-Ricupero N, Hilty D, Torous J. The role of digital navigators in promoting clinical care and technology integration into practice. Digit Biomark 2020;4(Suppl 1):119-135 [FREE Full text] [doi: 10.1159/000510144] [Medline: 33442585]
40. Sturgiss EA, Clark AM. Using critical realism in primary care research: an overview of methods. Family Pract 2020;37(1):143-145 [FREE Full text] [doi: 10.1093/fampra/cmz084]
41. Berry N, Bucci S, Lobban F. Use of the internet and mobile phones for self-management of severe mental health problems: qualitative study of staff views. JMIR Ment Health 2017 Nov 01;4(4):e52 [FREE Full text] [doi: 10.2196/mental.8311] [Medline: 29092809]
42. Yiend J, Taher R, Fialho C, Hampshire C, Hsu C, Kabir T, MPIT, Avegen. Assessing the efficacy and safety of STOP (successful treatment for paranoia)-an app-based cognitive bias modification therapy for paranoia: a randomised clinical trial protocol. Trials 2024 Dec 02;25(1):806 [FREE Full text] [doi: 10.1186/s13063-024-08570-3] [Medline: 39623444]

Abbreviations

CBM-I: cognitive bias modification for interpretation
CBM-pa: cognitive bias modification for paranoia
DHT: digital health technology
NICE: National Institute for Health and Care Excellence
PANSS: Positive and Negative Syndrome Scale
STOP: Successful Treatment of Paranoia
UEQ: User Experience Questionnaire

Edited by D Gooch; submitted 19.12.24; peer-reviewed by D Katz, A Hardy, AKC Chau; comments to author 29.03.25; revised version received 19.05.25; accepted 05.06.25; published 07.08.25.

Please cite as:
Eid L, Kenny A, Taher R, Jacobsen P, Hsu CW, Yiend J
A Digital Mental Health Intervention for Paranoia (the STOP App): Qualitative Study on User Acceptability
JMIR Hum Factors 2025;12:e70181
URL: https://humanfactors.jmir.org/2025/1/e70181
doi:10.2196/70181
PMID:

©Laura Eid, Alexandra Kenny, Rayan Taher, Pamela Jacobsen, Che Wei Hsu, Jenny Yiend. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 07.08.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Use and User Experience of a Preconception Lifestyle App for Couples Undergoing in Vitro Fertilization: Mixed Methods Study

Tessy Boedt[1,2], PhD; Sharon Lie Fong[3,4], MD, PhD; Eline Dancet[5], PhD; Merijn Mestdagh[6], PhD; Joke Verbeke[1,7], MSc; David Geerts[8], PhD; Carl Spiessens[3,4], PhD; Christophe Matthys[9,10,11], PhD

[1]Clinical and Experimental Endocrinology, Department of Chronic Diseases and Metabolism, KU Leuven, Leuven, Belgium
[2]Department of Developmental Psychology, Tilburg School of Social and Behavioral Sciences, Tilburg University, Tilburg, The Netherlands
[3]Department of Development and Regeneration, KU Leuven, Leuven, Belgium
[4]Leuven University Fertility Center, University Hospitals Leuven, Leuven, Belgium
[5]Department of Public Health and Primary Care, KU Leuven, Leuven, Belgium
[6]Center for Psychology of Learning and Experimental Pyschopathology, KU Leuven, Leuven, Belgium
[7]Health Intelligence Network Ghent (HINT.GENT), Faculty of Medicine and Health Sciences, Ghent University, Gent, Belgium
[8]Institute of Media Studies, KU Leuven, Leuven, Belgium
[9]DIEET, Department of Endocrinology, University Hospitals Leuven, Leuven, Belgium
[10]DIEET, Department of Chronic Diseases and Metabolism, KU Leuven, Leuven, Belgium
[11]Division of Human Nutrition, Faculty of Medicine & Health Sciences, Stellenbosch University, Capetown, South Africa

Corresponding Author:
Christophe Matthys, PhD
DIEET
Department of Chronic Diseases and Metabolism
KU Leuven
Herestraat 49
Leuven, 3000
Belgium
Phone: 32 016342655
Email: christophe.matthys@uzleuven.be

Abstract

Background: Mobile apps are a promising way to improve healthy lifestyle behavior among people with infertility. However, sufficient engagement with mobile health apps is crucial to influence health outcomes, and identifying features to create more effective interventions is urgently needed.

Objective: This study conducted a process evaluation focusing on the use and user experience of the PreLiFe app, a mobile lifestyle app for couples undergoing in vitro fertilization (IVF).

Methods: A mixed methods approach was used among heterosexual couples with infertility undergoing IVF. An objective quantitative study using a tracking-based system assessed the actual use of the PreLiFe app over time in relation to partner use and in relation to the specific fertility treatment. A subjective quantitative study using online questionnaires assessed the acceptability (using the Mobile App Rating Scale) and partner support (based on the Social Support for Diet and Exercise Scale) experienced while using the PreLiFe app. A subjective qualitative study using semistructured interviews evaluated in-depth user experiences with the PreLiFe app.

Results: A total of 106 couples used the PreLiFe app for 2 to 365 days. Overall use was low; 18.9% (20/106) of the men and 49.1% (52/106) of the women used all the modules of the PreLiFe app. Mixed-model analyses revealed that higher app use was observed when a partner used the app as well and during fertility treatment. The average acceptability score was 6 (SD 1) of 10, and patients felt supported by their partners while using the app. Semistructured interviews with 10 patients indicated that the PreLiFe app was easy to use.

Conclusions: Our findings showed good acceptability and user experiences but low actual objective use of a preconception lifestyle app for couples undergoing IVF. To increase use of and engagement with such apps, future studies should further focus on personalization and interaction with partners, health care providers, and other patient data systems.

XSL•FO
RenderX

(JMIR Hum Factors 2025;12:e65815) doi:10.2196/65815

KEYWORDS

mobile health; mHealth; infertility; lifestyle; usability

Introduction

Background

Infertility, defined as the failure to achieve a clinical pregnancy after at least 12 months of regular unprotected sexual intercourse, occurs in 1 in 10 couples [1-3]. Approximately half of these couples undergo fertility treatment, such as in vitro fertilization (IVF), with the ultimate goal of conceiving a healthy child [3]. Infertility and its treatments can be emotionally and financially burdensome to couples and society. An important research priority is to find modifiable factors that can improve the chance of conceiving a healthy child from IVF [4]. Increasing evidence shows that lifestyle is such a potential modifiable factor. IVF success rates are positively associated with a healthy diet, a normal BMI, and moderate physical activity levels [5-12]. Furthermore, a healthy preconception lifestyle benefits offsprings' health [13]. Several international guidelines recommend addressing unhealthy lifestyles before conception, specifically in people with infertility [14-16]. In contrast, a Cochrane review indicated a lack of clear guidance on the content and format of lifestyle programs for improving conception chances [17]. Health care providers also recognize the need for such guidance [18,19].

Evidence from other fields shows that mobile health (mHealth)—defined as the use of mobile and wireless technologies for health—is a promising format to promote healthy lifestyle behavior [20]. mHealth is widely scalable, relatively cheap, and adaptable according to the needs of patients [21]. Using mHealth among people with infertility may counter identified barriers such as lack of time and personnel and is easy to implement in daily life [20,22]. Recent studies on lifestyle apps for people with infertility have shown an improvement in secondary lifestyle outcomes (eg, food intake) but were not effective for the primary clinical outcomes, such as pregnancy rate or live birth [11,23-25]. Adherence is a challenge in traditional lifestyle studies. Therefore, in-depth analyses of use data of mHealth interventions is crucial to identifying factors that enhance use and engagement by participants and effectiveness of mobile preconception lifestyle interventions. In 2016, the World Health Organization mHealth Technical Evidence Review Group developed a guideline for the reporting of mHealth interventions. This guideline emphasized the need for detailed reporting on use [26]. Despite these existing guidelines, a review of mobile lifestyle interventions emphasized the need for improved reporting on the use of such interventions to evaluate their effectiveness more accurately [27].

Objectives

To gain better insights into the effectiveness of a preconception lifestyle app for couples undergoing IVF (PreLiFe study) [22,25], a process evaluation focusing on use and user experiences was conducted with the aim of identifying factors that may lead to higher use and engagement, ultimately enhancing the effectiveness of mobile lifestyle apps.

Methods

Overview

This process evaluation (Table 1) of the PreLiFe app, a mobile lifestyle app for couples undergoing IVF, included stakeholder interviews, quantitative observation, and data monitoring. More specifically, an objective quantitative study (data monitoring) using a tracking-based system assessed the actual use of the PreLiFe app over time in relation to partner use and in relation to the specific fertility treatment. A subjective quantitative study using online questionnaires assessed acceptability and partner support experienced while using the PreLiFe app. A qualitative study (stakeholder interviews) using semistructured interviews evaluated the user experiences with the PreLiFe app in depth in the context of the users' daily lives and in relation to their fertility treatment.

Table 1. Overview of outcomes assessed, timing, and measurement methods.

Outcomes assessed	Type of assessment	Timing of assessment	Method
Actual use of the PreLiFe app	Quantitative and objective	Continuously throughout the PreLiFe study period (2 to 365 days)	App-based tracking (data monitoring)
Acceptability—subjective quality	Quantitative and subjective	At the end of the PreLiFe study period	Subjective quality subscale of the Mobile Application Rating Scale
Acceptability—added value of different modules of the PreLiFe app	Quantitative and subjective	At the end of the PreLiFe study period	Semi–open-ended question
Acceptability—reasons for not using the PreLiFe app	Quantitative and subjective	After 3 months of using the PreLiFe app	Semi–open-ended question
Partner support experienced while using the PreLiFe app	Quantitative and subjective	After 3 months of using the PreLiFe app	Questionnaire based on the Social Support for Diet and Exercise Scale
User experience	Qualitative and subjective	After the PreLiFe study period	Semistructured interviews with stakeholders

XSL•FO
RenderX

PreLiFe App and Study (Background Information)

The effectiveness of the PreLiFe program was evaluated in the PreLiFe study [25]. The protocol for this multicenter, single-blind randomized controlled trial has been registered (ClinicalTrials.gov identifier NCT03790449) and published previously [28]. A detailed description of the PreLiFe program has been published previously [22]. Briefly, eligible couples starting IVF were randomized (1:1 allocation ratio) to an attention control arm or an intervention arm receiving the PreLiFe program. The PreLiFe program is a systematically developed complex preconception lifestyle intervention acting on a combination of lifestyle topics (diet, physical activity, and emotional distress) and offered in a blended care format. It was developed according to the Medical Research Council framework, and patients were involved in the development of this app from the earliest phases of design [22]. The program includes a mobile app (PreLiFe app) with on-demand tailored advice and skill training on diet (food literacy), physical activity, and mindfulness in combination with SMS text messages. In addition, patients receive telephone interaction every 3 months with a health care provider trained in motivational interviewing (Figure S1 in Multimedia Appendix 1). Participants were followed for 12 months or until an ongoing pregnancy was confirmed via ultrasound at 12 weeks of gestational age. Changes in lifestyle and in reproductive outcomes were observed. The study started in January 2019 and was prematurely stopped on March 13, 2020, as, due to the COVID-19 pandemic, all Belgian fertility clinics were closed for an undefined period. As the primary outcome was time to ongoing pregnancy, recruitment was stopped as of March 13, 2020. In addition, couples who had been randomized and were still participating in the trial at that time needed to be censored. The PreLiFe study period varied from 2 to 365 days. Table S1 in Multimedia Appendix 1 provides an overview of the number of couples in the study at each time point and reasons for ending the PreLiFe study. During the PreLiFe study period, all participating couples received standard medical treatment (ie, IVF with or without intracytoplasmic sperm injection) according to the local protocol of the participating fertility clinic. Both partners in couples randomized to the intervention arm additionally received the PreLiFe program.

Ethical Considerations

This study was approved by the Ethics Committee Research of the University Hospitals Leuven and the local ethics committees of the participating clinics (approval s61202 for the qualitative evaluation and s61596 for the quantitative evaluation). Written informed consent was obtained from all participants. Interviews were audio-recorded and transcribed verbatim. All identifying information was removed from the transcripts to ensure anonymity.

Quantitative Evaluation: Actual Use, Acceptability, and Partner Support

Participants and Recruitment

As we aimed to evaluate the use of the PreLiFe app, only participants allocated to the intervention arm of the PreLiFe study were eligible for this study, which included 106 couples undergoing IVF.

Outcomes

Actual use of the PreLiFe app was objectively monitored through app-based tracking (data monitoring) throughout the PreLiFe study period. Initiation of the different modules (diet, physical activity, and mindfulness) was assessed by counting the number of times the modules were opened. Actual use of the different modules was assessed by automatically counting the active clicking on the features of these modules during the PreLiFe study period. Active use of the PreLiFe app was defined as the percentage of participants having used all modules (diet, physical activity, and mindfulness) of the PreLiFe app at least once throughout the PreLiFe study period.

Acceptability of the PreLiFe app was assessed at the end of the PreLiFe study period using the subjective quality subscale of the Mobile Application Rating Scale. This scale has a score from 0 to 10, with higher scores indicating better acceptability. Furthermore, we asked which module (diet, physical activity, or mindfulness) provided the most added value to participants and if there were any specific reasons (time, interest, use of other app or program, injury, or other reasons) for potentially not using the PreLiFe app. Multiple answers could be indicated.

To evaluate whether participants felt supported by their partners while using the PreLiFe app in maintaining a healthy lifestyle, a questionnaire based on the Social Support for Diet and Exercise Scale was developed [29]. The questionnaire consisted of 3 subscales: partner support for diet, physical activity, and emotional distress.

Statistical Analyses

Using mixed-model analyses, we evaluated the actual use of the different modules of the PreLiFe app over time and in relation to partner use. Furthermore, we evaluated the actual use of the PreLiFe app in relation to the stage of fertility treatment. Start of fertility treatment was defined as the start date of the last menstrual period. Descriptive analyses were conducted on the other outcomes (baseline characteristics, acceptability, added value, reasons for not adhering to the PreLiFe app, and partner support).

Qualitative Evaluation: User Experiences

Participants and Recruitment

After the PreLiFe study was ended, participants from the intervention group were invited through email. The study process and accompanying incentives were explained in detail. Respondents who were interested in participating signed an online informed consent form and were scheduled for a semistructured interview through videoconferencing software (Microsoft Teams). Participants received a monetary incentive for taking part. The sample size was based on data saturation, with participants selected to ensure a diverse representation of age, gender, and app use, capturing a broad range of experiences.

Interview Guide and Process

Semistructured interviews were conducted by a researcher trained in qualitative methods according to an interview guide

XSL•FO
RenderX

(Table S2 in Multimedia Appendix 1) and audio recorded for later analysis. The interview guide was developed from a user experience perspective grounded in the unified theory of acceptance and use of technology and behavior change theories [30]. It covered general app use, use and experiences of the different modules of the PreLiFe app (diet, physical activity, and mindfulness), integration into daily life and treatment, social influences, and technology acceptance. Questions were refined through pilot-testing with a small user sample (n=2) to ensure clarity, flow, and sensitivity.

Statistical Analyses

The data collected in the interviews were analyzed according to the principles of thematic analysis [31]. Interviews were audio recorded, transcribed afterward, and coded, after which themes were extracted through an affinity diagram.

Results

Quantitative Evaluation: Actual Use, Acceptability, and Partner Support

Baseline Characteristics of the Participants

The baseline characteristics of the participants are presented in Table 2. More than 70% of the couples were highly educated, and almost all couples were of European origin. The mean duration trying to conceive was approximately 2 years (27, SD 16 months).

Table 2. Baseline characteristics of the study participants.

Characteristic	Qualitative study		Quantitative study	
	Women (n=5)	Men (n=5)	Women (n=106)	Men (n=106)
Age (y), mean (SD)	30.2 (3.2)	32.6 (4.1)	30.7 (3.9)	33.9 (6.1)
BMI[a], n (%)				
Underweight (<18.5 kg/m^2)	0 (0)	0 (0)	2 (1.9)	2 (1.9)
Normal weight (18.5-24.99 kg/m^2)	2 (40)	2 (40)	58 (54.7)	38 (35.8)
Overweight (25-29.99 kg/m^2)	2 (40)	3 (60)	30 (28.3)	52 (49.1)
Obesity (≥30 kg/m^2)	1 (20)	0 (0)	16 (15.1)	13 (12.3)
Unknown	0 (0)	0 (0)	0 (0)	1 (0.9)
European origin[b], n (%)	5 (100)	5 (100)	99 (93.4)	101 (95.3)
Educational level[b], n (%)				
ISCED[c] level 0-3	0 (0)	0 (0)	20 (18.9)	34 (32.1)
ISCED level 4-6	3 (60)	3 (60)	59 (55.7)	48 (45.3)
ISCED level 7-8	2 (40)	2 (40)	27 (25.5)	24 (22.6)
Duration of time trying to conceive (months)[b], mean (SD)	26 (14)	26 (14)	27 (16)	27 (16)
Type of infertility[d], n (%)				
Primary	4 (80)	4 (80)	79 (74.5)	78 (73.6)
Secondary	1 (20)	1 (20)	27 (25.5)	28 (26.4)
Female or male factor infertility diagnosis[d], n (%)				
No	2 (40)	1 (20)	52 (49.1)	33 (31.1)
Yes	3 (60)	4 (80)	52 (49.1)	71 (67)
Unexplained	0 (0)	0 (0)	2 (1.9)	2 (1.9)

[a]Measured using weight and height.

[b]Self-reported.

[c]ISCED: International Standard Classification of Education.

[d]Obtained from medical records.

Actual Use of the PreLiFe App

From January 2019 to March 2020, a total of 106 couples used the PreLiFe app for 2 to 365 days. As shown in Table 3, actual use of the app was low. Active users, namely, participants who used every module of the app at least once, included 49.1% (52/106) of the women and 18.9% (20/106) of the men. Approximately 70% of the women (79/106, 74.5%) and 50% of the men (50/106, 47.2%) used at least one of the modules (diet, physical activity, or mindfulness). Furthermore, passive users, defined as participants who initiated the app without using any of the modules, included 9.4% (10/106) of the women and

XSL•FO
RenderX

37.7% (40/106) of the men. In total, 1.9% (2/106) of the women and 1.9% (2/106) of the men did not initiate the app at all (nonusers).

Table 3. Actual use of the PreLiFe app.

	Women (n=106)	Men (n=106)
Diet		
Diet module initiated (questionnaire), n (%)	93 (87.7)	58 (54.7)
Diet module used (receiving tips and at least one goal set on food literacy), n (%)	79 (74.5)	38 (35.8)
>1 goal set, n (%)	36 (34)	11 (10.4)
Goals, range	1-7	1-7
PA[a]		
PA module initiated (questionnaire), n (%)	94 (88.7)	70 (66)
PA module used (receiving tips and at least one registration of step count or PA monitoring), n (%)	75 (70.8)	50 (47.2)
Step count function used, n (%)	71 (67)	48 (45.3)
Frequency of step count, range	8-200	8-242
Monitoring used, n (%)	31 (29.2)	23 (21.7)
Frequency of monitoring, range	1-236	1-138
Mindfulness		
Mindfulness module initiated (introductory movie), n (%)	76 (71.7)	41 (38.7)
Mindfulness module used (receiving tips and at least one exercise followed), n (%)	70 (66)	39 (36.8)
>1 exercise followed, n (%)	55 (51.9)	25 (23.6)
Frequency of exercising, range	1-19	1-22
Overall use, n (%)		
All modules initiated	68 (64.2)	32 (30.2)
All modules used (active users)	52 (49.1)	20 (18.9)
App initiated but no modules used (passive users)	10 (9.4)	40 (37.7)

[a]PA: physical activity.

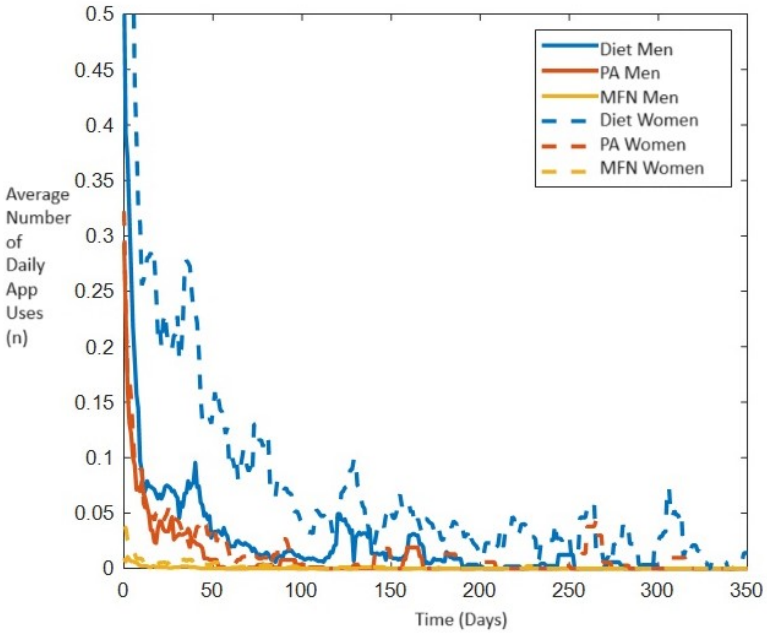
Actual Use of the Different Modules of the PreLiFe App Over Time in Relation to Partner Use

Figure 1 presents an overview of the average number of daily app uses (diet, physical activity, and mindfulness) by both partners over time. Statistical analyses revealed the following: (1) the average number of daily app uses decreased over time; (2) women used the PreLiFe app more frequently as compared to their partners; and (3) on the days when one of the partners used the app, the other partner also used the app more (Table S3 in Multimedia Appendix 1).

The diet module was the most used. The average number of daily uses per person of the diet module was 0.2149 (SD 0.4636). The mindfulness module was the least used. The average number of daily uses per person was 0.005 (SD 0.0707).

XSL•FO
RenderX

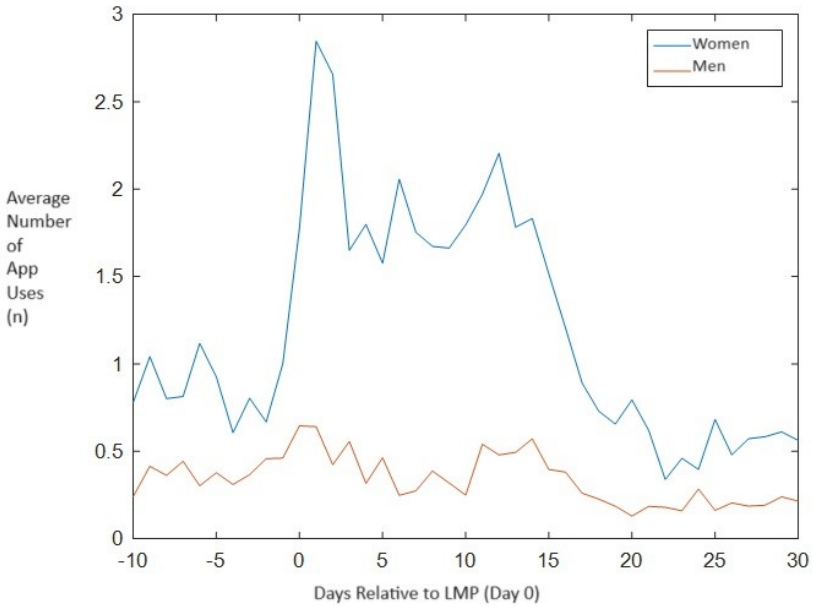
Figure 1. Use of the PreLiFe app over time. MFN: mindfulness; PA: physical activity.



Actual Use of the PreLiFe App in Relation to Fertility Treatment

Figure 2 shows that the use of the PreLiFe app was highest around the first day of the menstrual period (ie, the start of ovarian stimulation in the couples' first IVF–intracytoplasmic sperm injection treatment). We observed no differences in app use during the embryo transfer period. We did observe a decrease in app use across cycles (Figure S2 in Multimedia Appendix 1).

Figure 2. Use pattern of the PreLiFe app during in vitro fertilization (IVF) and intracytoplasmic sperm injection (ICSI). Day number relative to LMP refers to the day in the IVF and ICSI cycle that treatment was started (0=last menstrual period [LMP] day).



Acceptability of the PreLiFe App

A total of 40.6% (43/106) of the women and 33% (35/106) of the men completed the questions on acceptability of the PreLiFe app. Acceptability of the program revealed an average score of 6/10 (SD 1) from women and men, with a minimum score of 3/10 and a maximum score of 8/10. Women and men would recommend the PreLiFe app to other patients undergoing IVF (Table 4). Men rated the app slightly higher than women. In total, 26% (11/43) of the women gave a 4-star rating to the PreLiFe app versus 46% (16/35) of the men. Both women and men indicated that the diet module provided the most added value to them.

XSL•FO
RenderX

Table 4. Acceptability and added value of the PreLiFe app.

	Women (n=43), n (%)	Men (n=35), n (%)
"Would you recommend the PreLiFe app to other patients undergoing IVF who might benefit from it?"		
Not at all	0 (0)	2 (6)
To very few people	3 (7)	3 (9)
To some people	26 (60)	17 (49)
To many people	6 (14)	11 (31)
To everyone	8 (19)	2 (6)
"How would you rate the PreLiFe app?"		
1 star ("one of the worst apps I've used")	0 (0)	0 (0)
2 stars	2 (5)	4 (11)
3 stars (average)	30 (70)	15 (43)
4 stars	11 (26)	16 (46)
5 stars ("one of the best apps I've used")	0 (0)	0 (0)
Module that provided the most added value		
Diet	19 (44)	15 (43)
Physical activity	11 (26)	9 (26)
Mindfulness	7 (16)	5 (14)
No reply	6 (14)	6 (17)

A total of 26.4% (28/106) of the women and 22.6% (24/106) of the men completed the questions on reasons for not using the PreLiFe app. Table 5 provides an overview of these self-reported reasons. The most commonly indicated reason for not using the PreLiFe app was lack of time. In total, 14% (4/28) of the women also indicated that the PreLiFe app did not provide enough push messages.

Table 5. Reasons for not using the PreLiFe app.

Reasons	Women (n=28), n (%)	Men (n=24), n (%)
No time	16 (57)	18 (75)
Not interested	11 (40)	11 (46)
Use of other app or program	4 (14)	5 (21)
Injury or disease	1 (4)	0 (0)
Other	6 (21)	2 (8)
Limit on screen time	0 (0)	1 (4)
Not enough push messages	4 (14)	0 (0)
New phone	1 (4)	0 (0)
Failed fertility treatment	1 (4)	0 (0)

Partner Support Experienced While Using the PreLiFe App

A total of 44% (47/106) of the women and 33% (35/106) of the men completed the questions on partner support experienced while using the PreLiFe-app Regarding partner support experienced while using the PreLiFe app, more than half (30/47, 64%) of the women and men practiced a healthy diet often or very often together. One in 3 couples worked on their physical activity together. Furthermore, more than half (32/47, 68% and 25/35, 71%, respectively) of the women felt often supported by their partners in their diet. Only 0.9% (1/106) of the couples practiced mindfulness together. In addition, 38% (18/47) of the women and 46% (16/35) of the men never felt supported by their partners when practicing mindfulness (Table 6).

XSL•FO
RenderX

Table 6. Partner support experienced while using the PreLiFe app at 3 months.

	Women (n=47), n (%)			Men (n=35), n (%)		
	Diet	Physical activity	Mindfulness	Diet	Physical activity	Mindfulness
Practiced with their partner						
Never	3 (6)	4 (9)	33 (70)	0 (0)	3 (9)	26 (74)
Seldom	2 (4)	9 (19)	5 (11)	2 (6)	9 (26)	4 (11)
A couple of times	10 (21)	15 (32)	3 (6)	8 (23)	8 (23)	3 (9)
Often	14 (30)	15 (32)	0 (0)	15 (43)	11 (31)	1 (3)
Very often	16 (34)	2 (4)	1 (2)	10 (29)	2 (6)	0 (0)
Not applicable	2 (4)	2 (4)	5 (11)	0 (0)	2 (6)	1 (3)
Felt supported by their partner						
Never	1 (2)	3 (6)	18 (38)	0 (0)	1 (3)	16 (46)
Seldom	3 (6)	7 (15)	5 (11)	1 (3)	7 (20)	4 (11)
A couple of times	10 (21)	11 (23)	6 (13)	5 (14)	6 (17)	3 (9)
Often	19 (40)	14 (30)	2 (4)	13 (37)	13 (37)	3 (9)
Very often	13 (28)	10 (21)	2 (4)	13 (37)	3 (9)	0 (0)
Not applicable	1 (2)	2 (4)	14 (30)	3 (9)	5 (14)	9 (26)

Qualitative Evaluation: User Experiences

Baseline Characteristics of the Participants

A total of 10 participants (4 couples and n=2, 20% individual participants) took part in 6 semistructured interviews between January 2021 and March 2021. The sample consisted of 50% (5/10) male individuals and 50% (5/10) female individuals who had all participated in the PreLiFe study and received access to the PreLiFe app for 1 year or until ongoing pregnancy occurred. The baseline characteristics of the participants are presented in Table 2.

User Experiences

A total of 112 codes were categorized into eight themes: (1) general evaluation, (2) pick-and-mix functionalities, (3) descriptions of app use, (4) app impact, (5) non–app-dependent behavior, (6) trust, (7) problems with the app, and (8) suggestions.

General Evaluation

All 10 participants mentioned that the PreLiFe app was easy to use and understand and that an app such as this one should be available by default as a part of fertility treatment. However, some added that it should not be mandatory.

Pick-and-Mix Functionalities

An obvious theme within each of the interviews was the on-demand or pick-and-mix approach toward modules and even functionalities within modules that participants used. Participants tended to pick the modules or functionalities that fit best in their lives. They often mentioned finding specific parts of the app useful:

> *The possibility to choose cooking advice suitable to your level of kitchen prowess.* [Participant 1]

Conversely, participants also mentioned that specific modules, especially mindfulness, were “not for them” (participants 2 and 5).

Descriptions of App Use

When asked about the use of the PreLiFe app in different social contexts, many participants (7/10, 70%) mentioned that they would use the app or suggestions from the app together with their partner to support each other in improving their lifestyle and prepare for the fertility treatment. A lot of the app use depended on the time available. Some participants (3/10, 30%) mentioned only using the PreLiFe app when they had some free time. Others mentioned that “the short tips received through notifications also work well when you don’t have much time at your disposal” (participant 6). In addition, the timing and events during IVF treatment could have an impact on people’s app use. Several participants (6/10, 60%) mentioned mainly using the app during key moments of the treatment:

> *We used the app around important doctor visits.* [Participant 8]

One participant mentioned completely pausing their use of the app as they had received disappointing news during their treatment, wanting to step away from the treatment for a little while. Some people (2/10, 20%), mostly men, did not end up using the entire app much or at all:

> *I have more of a supportive role in the process, the application might be more useful for women.* [Participant 3]

App Impact

Several participants (7/10, 70%) mentioned that the PreLiFe app made them aware that a healthy lifestyle was relevant to their fertility treatment. In addition, several participants (6/10, 60%) mentioned that using the PreLiFe app increased awareness

XSL•FO
RenderX

of current behavior and triggered potential behavior change. However, it was mentioned that it was not always easy to adhere to the planned behavior. It was also mentioned that receiving notifications triggered further app use. Participants mentioned that, throughout their treatment, they were trying to find some sense of control over or influence on the results of the treatment. Some participants (4/10, 40%) said that the PreLiFe app helped them obtain this feeling of control.

Non–App-Dependent Behavior

Many participants (8/10, 80%) mentioned that they were already using physical activity trackers and, therefore, did not use those parts of the PreLiFe app as it was not compatible with their wearables.

Trust

The fact that the PreLiFe app was developed through the efforts of a university and hospital and operated by scientists and health care providers had a very positive effect on trust in the app. Participants were also willing to share personal information with the app.

Problems With App Use

One participant mentioned that having a pending update for the app caused her to stop receiving new information. Other experienced problems were with the pedometer, which stopped working or provided inconsistent results as compared to their own activity trackers.

Suggestions

Throughout the interviews, participants made several suggestions that could be added to the PreLiFe app. Participants stated that the functionality could become more automated by integrating current health applications or activity trackers. In addition, participants suggested that it would be interesting if health care professionals could have access to the recorded data, allowing for tailored feedback on their behavior. This aligns with the wish to have a more direct connection to health care providers by, for instance, being able to directly ask questions to the medical staff through the app.

Some participants (4/10, 40%) highlighted the importance of support throughout the emotional aspect of the fertility treatment, for instance, providing contact information for emotional support providers such as fertility psychologists. In addition, it was mentioned that it would be good to have some sort of “emotional guardian” who could step in when things became too much. Others suggested investing in features to improve the social aspect of the app, for instance, designing features that allow for sharing recorded data between partners as a way to help them stimulate and motivate each other. One participant felt that it would be interesting to have a forum attached in which different app users could share experiences and motivations with other people using the app.

Discussion

Principal Findings

This paper describes the quantitative and qualitative evaluation of the use and user experience of the PreLiFe app, as well as the app’s acceptability and the partner support experienced.

Overall, patients were positive about the PreLiFe app, often felt supported by their partners while using the app, and were willing to use it, as revealed by our quantitative and qualitative subjective evaluations. However, our objective use data revealed that actual app use was low, particularly among men. We observed that PreLiFe app use was related to partner app use and to the stage of fertility treatment (higher app use during the start of IVF treatment). Furthermore, we observed that the module that had the highest number of interactions and highest level of tailoring (the diet module) had the highest use.

Implications and Comparison With Previous Work

One of the most remarkable findings of this study was the paradox between the high scores for acceptability and subjective use and the low actual objective use, highlighting the importance of objective measurement methods for app use. In the context of behavior change, these findings could be explained by the fact that the participants were in a phase of being aware of the health benefit (contemplation phase) but not yet having enough motivation or awareness to proceed to action for behavior change (in this case, regular app use and follow-up goals). Focusing on intrinsic motivation and possibly external motivation (intermediate follow-up by a physician or health care professional) could potentially have a synergistic effect. Compliance and acceptability do not reflect the actual use of lifestyle apps. Comparing our results with those of other studies remains difficult. A recent systematic review of lifestyle apps indicated substantial variability in intervention characteristics and inconsistent reporting of use outcomes [27]. Future research should work with core outcome sets on use to address this inconsistency in selecting and reporting of use outcomes [32]. Although previous research has shown that patients undergoing IVF need support in lifestyle modification and prefer a mobile app as a time-efficient format [22], the actual use of our PreLiFe app was low, especially among men. A possible reason may be found in a recent study by Robertson et al [33], who showed that patients undergoing IVF might prefer focusing on their treatment instead of other factors such as emotional management or their lifestyle. Another potential explanation for the low use among men is that we did not consider gender-specific differences when developing our app. Previous studies have shown that men may have different perceptions of lifestyle relevance to the fertility trajectory or different coping strategies to handle infertility [34,35]. Furthermore, lifestyle behavior and motivation to improve their lifestyle might differ between partners. Further characterizing the current needs of both partners and tailoring lifestyle apps to their personalized needs and motivations can contribute to the engagement of both partners in such interventions.

A recent systematic review identified several factors that positively influence adherence to lifestyle apps: personal support, content personalization, individualized push

XSL•FO
RenderX

notifications, and user-friendly design [36]. In our study, we observed that PreLiFe app use was higher when the partner used the app as well and that patients felt supported by their partners. This is in line with previous studies showing that involving both partners in lifestyle interventions can facilitate mutual support, behavior change, and compliance at little additional cost [23,37]. A recent Dutch study providing tailored coaching on vegetable, fruit, and folic acid intake; smoking; and alcohol consumption in a blended approach to couples trying to conceive reported improved lifestyle behavior, which was more robust when both partners participated [38]. These findings suggest that partner-based interventions may offer a comparative advantage over individual-focused apps. Future studies on lifestyle apps for people with infertility should further consider this partner interaction.

PreLiFe app use also increased during fertility treatment, when participants were in contact with health care providers. This supports the idea that blended care models combining digital tools with professional support may enhance engagement. Determining the optimal balance between digital autonomy and health care provider interaction remains a key question [39].

Identifying the balance between time investment and personalization is also key for future research on this topic. Participants indicated lack of time as the most important reason for not using the PreLiFe app. In contrast, we observed the highest use for the diet module, which was the module with the highest level of time investment, interaction (2 push messages per week), and tailoring (personalized goals based on a questionnaire).

Another improvement area defined was the lack of interaction with other data systems such as wearables. Integration of wearables could enhance personalization and efficiency, as observed in other health apps [40]. However, further research on this topic in this specific population is warranted.

Strengths and Limitations

The main strength of this study is that we combined qualitative and quantitative data on the use and user experiences of both partners, allowing for an in-depth understanding of participants' use and experiences with a preconception lifestyle app. Furthermore, to our knowledge, we are the first to provide objective insights into the actual use of a lifestyle app among both partners. We did so using app-based tracking, an objective measure not prone to social desirability bias. In-depth analysis of such use data can contribute to identifying and improving factors that lead to higher use and engagement, ultimately enhancing the effectiveness of lifestyle apps [41,42].

The limitations of this study should also be considered. First, all participants were asked about their rationale for not using the PreLiFe app, but fewer than 1 in 3 women and men replied, potentially not reflecting the views of all nonusers. In general, self-selection of the participants might have played a role throughout the PreLiFe study. Our study population had a higher physical activity level and lower rates of overweight and obesity than the general Belgian population [43,44]. Selection bias could have been present in our study, resulting in participants with a high interest in lifestyle and health. Furthermore, most patients were of European origin and highly educated, limiting the generalizability of our results. Finally, to avoid burdening the participants, the final results of the qualitative study were not returned to them for validation. This decision could potentially lead to misinterpretations of their true thoughts.

Conclusions

In conclusion, our findings show good acceptability and user experiences but low actual objective use of a preconception lifestyle app for couples undergoing IVF. Lifestyle apps can empower patients to feel more in control and supported regarding their reproductive health. However, to increase use and potential effectiveness, future research on lifestyle apps for people with infertility should further focus on the personalization of such apps to individual needs and on the interaction with partners, health care providers, and other patient data systems such as wearables.

Acknowledgments

The authors would like to thank Steve De Backer for the technical development and support of the PreLiFe app, Kevin Sanders for contributing to the user experience research on the PreLiFe app, the advisory committee of our Research Foundation – Flanders–TBM project (Applied Biomedical Research with a Primary Social finality) for their guidance and advice, and all the patients participating in the PreLiFe study. This work was supported by the Research Foundation – Flanders (Belgium; reference: T005417N).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary material.
[PDF File (Adobe PDF File), 465 KB - humanfactors_v12i1e65815_app1.pdf]

References

XSL•FO
RenderX

1. Zegers-Hochschild F, Adamson GD, Dyer S, Racowsky C, de Mouzon J, Sokol R, et al. The international glossary on infertility and fertility care, 2017. Hum Reprod 2017 Sep 01;32(9):1786-1801. [doi: 10.1093/humrep/dex234] [Medline: 29117321]
2. Mascarenhas MN, Flaxman SR, Boerma T, Vanderpoel S, Stevens GA. National, regional, and global trends in infertility prevalence since 1990: a systematic analysis of 277 health surveys. PLoS Med 2012 Dec 18;9(12):e1001356 [FREE Full text] [doi: 10.1371/journal.pmed.1001356] [Medline: 23271957]
3. Boivin J, Bunting L, Collins JA, Nygren KG. International estimates of infertility prevalence and treatment-seeking: potential need and demand for infertility medical care. Hum Reprod 2007 Jun;22(6):1506-1512. [doi: 10.1093/humrep/dem046] [Medline: 17376819]
4. Moran LJ, Spencer L, Russell DL, Hull ML, Robertson SA, Varcoe TJ, Robinson Research Institute Consortium of Fertility and Conception Practitioners. Research priorities for fertility and conception research as identified by multidisciplinary health care practitioners and researchers. Nutrients 2016 Jan 13;8(1):35. [doi: 10.3390/nu8010035] [Medline: 26771633]
5. Piché ML, Babineau V, Robitaille J, Lachance É, Ruchat SM. Lifestyle-related factors associated with reproductive health in couples seeking fertility treatments: results of a pilot study. Int J Fertil Steril 2018 Apr;12(1):19-26. [doi: 10.22074/ijfs.2018.5135] [Medline: 29334202]
6. Vujkovic M, de Vries JH, Lindemans J, Macklon NS, van der Spek PJ, Steegers EA, et al. The preconception Mediterranean dietary pattern in couples undergoing in vitro fertilization/intracytoplasmic sperm injection treatment increases the chance of pregnancy. Fertil Steril 2010 Nov;94(6):2096-2101 [FREE Full text] [doi: 10.1016/j.fertnstert.2009.12.079] [Medline: 20189169]
7. Twigt JM, Bolhuis ME, Steegers EA, Hammiche F, van Inzen WG, Laven JS, et al. The preconception diet is associated with the chance of ongoing pregnancy in women undergoing IVF/ICSI treatment. Hum Reprod 2012 Aug 15;27(8):2526-2531. [doi: 10.1093/humrep/des157] [Medline: 22593431]
8. Salas-Huetos A, Bulló M, Salas-Salvadó J. Dietary patterns, foods and nutrients in male fertility parameters and fecundability: a systematic review of observational studies. Hum Reprod Update 2017 Jul 01;23(4):371-389. [doi: 10.1093/humupd/dmx006] [Medline: 28333357]
9. Sundaram R, Mumford SL, Buck Louis GM. Couples' body composition and time-to-pregnancy. Hum Reprod 2017 Mar 01;32(3):662-668 [FREE Full text] [doi: 10.1093/humrep/dex001] [Medline: 28158570]
10. Rittenberg V, Seshadri S, Sunkara SK, Sobaleva S, Oteng-Ntim E, El-Toukhy T. Effect of body mass index on IVF treatment outcome: an updated systematic review and meta-analysis. Reprod Biomed Online 2011 Oct;23(4):421-439. [doi: 10.1016/j.rbmo.2011.06.018] [Medline: 21885344]
11. van Dijk MR, Koster MP, Willemsen SP, Huijgen NA, Laven JS, Steegers-Theunissen RP. Healthy preconception nutrition and lifestyle using personalized mobile health coaching is associated with enhanced pregnancy chance. Reprod Biomed Online 2017 Oct;35(4):453-460. [doi: 10.1016/j.rbmo.2017.06.014] [Medline: 28688924]
12. Rao M, Zeng Z, Tang L. Maternal physical activity before IVF/ICSI cycles improves clinical pregnancy rate and live birth rate: a systematic review and meta-analysis. Reprod Biol Endocrinol 2018 Feb 07;16(1):11 [FREE Full text] [doi: 10.1186/s12958-018-0328-z] [Medline: 29415732]
13. Stephenson J, Heslehurst N, Hall J, Schoenaker DA, Hutchinson J, Cade JE, et al. Before the beginning: nutrition and lifestyle in the preconception period and its importance for future health. Lancet 2018 May 05;391(10132):1830-1841 [FREE Full text] [doi: 10.1016/S0140-6736(18)30311-8] [Medline: 29673873]
14. WHO recommendations on maternal health: guidelines approved by the WHO Guidelines Review Committee. World Health Organization. URL: https://www.who.int/publications/i/item/9789240080591 [accessed 2025-05-29]
15. Infertility. World Health Organization. 2024 May 22. URL: https://www.who.int/news-room/fact-sheets/detail/infertility [accessed 2025-05-29]
16. Gameiro S, Boivin J, Dancet E, de Klerk C, Emery M, Lewis-Jones C, et al. ESHRE guideline: routine psychosocial care in infertility and medically assisted reproduction-a guide for fertility staff. Hum Reprod 2015 Nov 07;30(11):2476-2485. [doi: 10.1093/humrep/dev177] [Medline: 26345684]
17. Boedt T, Vanhove AC, Vercoe MA, Matthys C, Dancet E, Lie Fong S. Preconception lifestyle advice for people with infertility. Cochrane Database Syst Rev 2021 Apr 29;4(4):CD008189 [FREE Full text] [doi: 10.1002/14651858.CD008189.pub3] [Medline: 33914901]
18. Homan GF, deLacey S, Tremellen K. Promoting healthy lifestyle in fertility clinics; an Australian perspective. Hum Reprod Open 2018 Jan 10;2018(1):hox028 [FREE Full text] [doi: 10.1093/hropen/hox028] [Medline: 30895240]
19. Boedt T, Dancet E, Speelman N, Spiessens C, Matthys C, Lie Fong S. Belgian fertility clinic staff value healthy lifestyle promotion but lack access to a structured lifestyle modification programme: an observational study. Gynecol Obstet Invest 2023 Aug 8;88(5):278-285 [FREE Full text] [doi: 10.1159/000531139] [Medline: 37552964]
20. Afshin A, Babalola D, Mclean M, Yu Z, Ma W, Chen C, et al. Information technology and lifestyle: a systematic evaluation of internet and mobile interventions for improving diet, physical activity, obesity, tobacco, and alcohol use. J Am Heart Assoc 2016 Aug 31;5(9):632 [FREE Full text] [doi: 10.1161/JAHA.115.003058] [Medline: 27581172]

XSL•FO
RenderX

21. Becker S, Miron-Shatz T, Schumacher N, Krocza J, Diamantidis C, Albrecht U. mHealth 2.0: experiences, possibilities, and perspectives. JMIR Mhealth Uhealth 2014 May 16;2(2):e24 [FREE Full text] [doi: 10.2196/mhealth.3328] [Medline: 25099752]
22. Boedt T, Matthys C, Lie Fong S, De Neubourg D, Vereeck S, Seghers J, et al. Systematic development of a mobile preconception lifestyle programme for couples undergoing IVF: the PreLiFe-programme. Hum Reprod 2021 Aug 18;36(9):2493-2505. [doi: 10.1093/humrep/deab166] [Medline: 34379119]
23. Van Dijk MR, Huijgen NA, Willemsen SP, Laven JS, Steegers EA, Steegers-Theunissen RP. Impact of an mHealth platform for pregnancy on nutrition and lifestyle of the reproductive population: a survey. JMIR Mhealth Uhealth 2016 May 27;4(2):e53 [FREE Full text] [doi: 10.2196/mhealth.5197] [Medline: 27234926]
24. Oostingh EC, Koster MP, van Dijk MR, Willemsen SP, Broekmans FJ, Hoek A, et al. First effective mHealth nutrition and lifestyle coaching program for subfertile couples undergoing in vitro fertilization treatment: a single-blinded multicenter randomized controlled trial. Fertil Steril 2020 Nov;114(5):945-954 [FREE Full text] [doi: 10.1016/j.fertnstert.2020.04.051] [Medline: 32741619]
25. Boedt T, Dancet E, De Neubourg D, Vereeck S, Jan S, Van der Gucht K, et al. A blended preconception lifestyle programme for couples undergoing IVF: lessons learned from a multicentre randomized controlled trial. Hum Reprod Open 2023;2023(4):hoad036 [FREE Full text] [doi: 10.1093/hropen/hoad036] [Medline: 38455033]
26. Agarwal S, LeFevre AE, Lee J, L'Engle K, Mehl G, Sinha C, WHO mHealth Technical Evidence Review Group. Guidelines for reporting of health interventions using mobile phones: mobile health (mHealth) evidence reporting and assessment (mERA) checklist. BMJ 2016 Mar 17;352:i1174. [doi: 10.1136/bmj.i1174] [Medline: 26988021]
27. Milne-Ives M, Lam C, De Cock C, Van Velthoven MH, Meinert E. Mobile apps for health behavior change in physical activity, diet, drug and alcohol use, and mental health: systematic review. JMIR Mhealth Uhealth 2020 Mar 18;8(3):e17046 [FREE Full text] [doi: 10.2196/17046] [Medline: 32186518]
28. Boedt T, Dancet E, Lie Fong S, Peeraer K, De Neubourg D, Pelckmans S, et al. Effectiveness of a mobile preconception lifestyle programme in couples undergoing in vitro fertilisation (IVF): the protocol for the PreLiFe randomised controlled trial (PreLiFe-RCT). BMJ Open 2019 Jul 30;9(7):e029665 [FREE Full text] [doi: 10.1136/bmjopen-2019-029665] [Medline: 31366659]
29. Carlson JA, Sallis JF, Wagner N, Calfas KJ, Patrick K, Groesz LM, et al. Brief physical activity-related psychosocial measures: reliability and construct validity. J Phys Act Health 2012 Nov;9(8):1178-1186 [FREE Full text] [doi: 10.1123/jpah.9.8.1178] [Medline: 22207589]
30. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425. [doi: 10.2307/30036540]
31. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
32. Gargon E, Williamson PR, Altman DG, Blazeby JM, Tunis S, Clarke M. The COMET Initiative database: progress and activities update (2015). Trials 2017 Feb 03;18(1):54 [FREE Full text] [doi: 10.1186/s13063-017-1788-8] [Medline: 28159003]
33. Robertson I, Harrison C, Ng KY, Macklon N, Cheong Y, Boivin J. Development, implementation and initial feasibility testing of the MediEmo mobile application to provide support during medically assisted reproduction. Hum Reprod 2022 May 03;37(5):1007-1017 [FREE Full text] [doi: 10.1093/humrep/deac046] [Medline: 35348672]
34. Peterson BD, Pirritano M, Christensen U, Schmidt L. The impact of partner coping in couples experiencing infertility. Hum Reprod 2008 May 01;23(5):1128-1137. [doi: 10.1093/humrep/den067] [Medline: 18325885]
35. Peterson B, Place JM. The experience of infertility: an unexpected barrier in the transition to parenthood. In: Ben-Ari OT, editor. Pathways and Barriers to Parenthood. Cham, CH: Springer; Sep 26, 2019:19-37.
36. Jakob R, Harperink S, Rudolf AM, Fleisch E, Haug S, Mair JL, et al. Factors influencing adherence to mHealth apps for prevention or management of noncommunicable diseases: systematic review. J Med Internet Res 2022 May 25;24(5):e35371 [FREE Full text] [doi: 10.2196/35371] [Medline: 35612886]
37. Best D, Avenell A, Bhattacharya S, Stadler G. New debate: is it time for infertility weight-loss programmes to be couple-based? Hum Reprod 2017 Dec 01;32(12):2359-2365. [doi: 10.1093/humrep/dex313] [Medline: 29040540]
38. van der Windt M, van der Kleij RM, Snoek KM, Willemsen SP, Dykgraaf RH, Laven JSE, et al. Impact of a blended periconception lifestyle care approach on lifestyle behaviors: before-and-after study. J Med Internet Res 2020 Sep 30;22(9):e19378 [FREE Full text] [doi: 10.2196/19378] [Medline: 32996885]
39. Schoeppe S, Alley S, Van Lippevelde W, Bray NA, Williams SL, Duncan MJ, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. Int J Behav Nutr Phys Act 2016 Dec 07;13(1):127 [FREE Full text] [doi: 10.1186/s12966-016-0454-y] [Medline: 27927218]
40. Kang HS, Exworthy M. Wearing the future-wearables to empower users to take greater responsibility for their health and care: scoping review. JMIR Mhealth Uhealth 2022 Jul 13;10(7):e35684 [FREE Full text] [doi: 10.2196/35684] [Medline: 35830222]

XSL•FO
RenderX

41. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun 13;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
42. Kelders SM, van Zyl LE, Ludden GD. The concept and components of engagement in different domains applied to eHealth: a systematic scoping review. Front Psychol 2020 May 27;11:926 [FREE Full text] [doi: 10.3389/fpsyg.2020.00926] [Medline: 32536888]
43. Gisle L. Gezondheidsenquête 2013. Rapport 2: gezondheidsgedrag en leefstijl. Wetenschappelijk Instituut Volksgezondheid. 2014. URL: https://www.sciensano.be/nl/biblio/gezondheidsenquete-2013-rapport-2-gezondheidsgedrag-en-leefstijl [accessed 2025-05-29]
44. WHO European regional obesity report 2022. World Health Organization. URL: https://www.who.int/europe/publications/i/item/9789289057738 [accessed 2025-05-29]

Abbreviations

IVF: in vitro fertilization
mHealth: mobile health

Edited by A Kushniruk; submitted 27.08.24; peer-reviewed by LG Tran, Y Wang; comments to author 12.05.25; revised version received 07.07.25; accepted 12.08.25; published 14.10.25.

Please cite as:
Boedt T, Lie Fong S, Dancet E, Mestdagh M, Verbeke J, Geerts D, Spiessens C, Matthys C
Use and User Experience of a Preconception Lifestyle App for Couples Undergoing in Vitro Fertilization: Mixed Methods Study
JMIR Hum Factors 2025;12:e65815
URL: https://humanfactors.jmir.org/2025/1/e65815
doi:10.2196/65815
PMID:41086000

©Tessy Boedt, Sharon Lie Fong, Eline Dancet, Merijn Mestdagh, Joke Verbeke, David Geerts, Carl Spiessens, Christophe Matthys. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 14.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Enhancing Methodological Rigor in Mobile Health Care Research

Shuhan Tang[1,2], MS

[1]Duke Global Health Institute, Duke University, Durham, NC, United States

[2]Global Health Research Center (GHRC), Duke Kunshan University, No. 8 Duke Avenue, Kunshan, China

Corresponding Author:
Shuhan Tang, MS
Duke Global Health Institute, Duke University, Durham, NC, United States

Related Article:

https://humanfactors.jmir.org/2024/1/e59153

(JMIR Hum Factors 2025;12:e68243) doi:10.2196/68243

KEYWORDS

letter to the editor; health care professionals; mobile health care; technical training; cross-sectional survey; mobile; China; web-based questionnaire; logistic regression; mhealth; mobile health

Tang et al's recent study published in *JMIR Human Factors* [1], titled "The Use of Mobile Health Care Among Medical Professionals in the Sichuan-Chongqing Region: Cross-Sectional Survey Study," captured my attention. Their analysis of mobile health (mHealth) device use and influencing factors, using chi-square and multivariable logistic regression analyses, revealed a significant association between age and mHealth use. This study provides valuable insights from China's western region.

However, I would like to offer a few comments and suggestions that I believe would further enhance the study's methodology and findings.

First, the study's reliance on a web-based questionnaire may introduce potential selection bias, as respondents, by nature of accessing web-based surveys, may be more likely to use digital devices and have a stronger interest in digital health in general. A more diverse, mixed-methods approach to questionnaire distribution could help mitigate this potential self-selection toward mHealth use and prevent the exclusion of individuals with limited experience or interest in mHealth.

Second, the study could benefit from a standardized and rigorous methodological framework for survey design and reporting. For instance, categorizing age into three broad groups may overlook essential trends. A finer measurement scale for satisfaction and usage, such as a Likert scale, could provide deeper insights into health care professionals' attitudes. Furthermore, for questions where "uncertain" responses outweigh "yes" or "no" responses, a qualitative or mixed-methods research approach could yield a more nuanced understanding of the underlying reasons. Finally, incorporating years of work experience as a variable could add valuable insights, given its potential correlation with age. Unlike professional titles, years of work experience could provide a more direct measure of professional tenure, potentially enriching the data analysis.

Third, ethical considerations of this study merit further attention. Identifying specific hospitals in the report may compromise confidentiality. Additionally, the exemption of review by an institutional review board (IRB) raises concerns, as the study involves gathering potentially sensitive information from human participants (ie, health care professionals) regarding personal perspectives, workflows, and technology use. Such studies often warrant an ethics board review to protect participant privacy, minimize psychological or social risks, and ensure adherence to ethical standards [2].

Fourth, before conducting logistic regression analysis, normality tests on both independent and dependent variables are recommended to confirm the validity of the chosen statistical methods [3]. If data are not normally distributed, adjustments should be considered.

Lastly, as noted in the "Limitations" section [1], the reliance on convenience sampling may affect the study's generalizability. Only public, district-level hospitals were included, excluding primary and tertiary institutions—this limits the representation of China's complete health care hierarchy in the analysis. Furthermore, only urban areas were sampled, overlooking rural populations. Women comprised 77.1% of the study sample, raising questions about gender representation. A multi-level sampling approach would likely yield a more comprehensive and representative dataset.

To conclude, I would like to highlight that these feedback points are not to challenge the integrity of the authors' work. Instead, I hope they can contribute to ongoing discussions on mHealth research and the development of robust methodologies in this field.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Editorial Notice

The corresponding author of the publication "The Use of Mobile Health Care Among Medical Professionals in the Sichuan-Chongqing Region: Cross-Sectional Survey Study" declined to respond to this letter.

References

1. Tang Y, Yang J, Wang N, Wang X, Hu W. The use of mobile health care among medical professionals in the Sichuan-Chongqing region: cross-sectional survey study. JMIR Hum Factors 2024 Oct 24;11:e59153. [doi: 10.2196/59153] [Medline: 39447153]
2. Beauchamp TL, Childress JF. Principles of Biomedical Ethics, 8th edition: Oxford University Press; 2019.
3. Field A. Discovering Statistics Using IBM SPSS Statistics, 4th edition: Sage Publications; 2013.

Abbreviations

IRB: institutional review board
mHealth: mobile health

Edited by A Iannaccio, T Leung; submitted 31.10.24; this is a non–peer-reviewed article;accepted 29.11.24; published 03.01.25.

Please cite as:
Tang S
Enhancing Methodological Rigor in Mobile Health Care Research
JMIR Hum Factors 2025;12:e68243
URL: https://humanfactors.jmir.org/2025/1/e68243
doi:10.2196/68243

© Shuhan Tang. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 3.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Story Retelling and Verbal Working Memory in Young Adults With a History of COVID-19: Cross-Sectional Study

Hyunsoo Yoo, PhD

Department of Communication Science and Disorders, Baylor University, 1311 S 5th St, Waco, TX, United States

Corresponding Author:
Hyunsoo Yoo, PhD
Department of Communication Science and Disorders, Baylor University, 1311 S 5th St, Waco, TX, United States

Abstract

Background: The impact of the COVID-19 pandemic has primarily been studied in the context of language delays or developmental disorders in infants and children. However, the effects on young adults have received less attention. COVID-19 not only affects physical health but also cognitive and language functions, which is an emerging area of research. While previous studies have focused on developmental stages, the effects of COVID-19 on the language abilities of healthy young adults remain underexplored. This study aimed to investigate the impact of COVID-19 on the spoken language, particularly in story retelling and working memory, in young adults.

Objective: This study aimed to investigate the effects of COVID-19 on memory-based story retelling and verbal working memory in young adults. Specifically, it examined whether there were group differences in story retelling and working memory performance between individuals with and those without a history of COVID-19, and whether verbal working memory predicted story retelling outcomes.

Methods: The study involved 79 young adult participants, of whom 39 were in the non–COVID-19 group and 40 were in the COVID-19 group. Participants completed the Story Retelling Procedure (SRP) and a verbal working memory task. Story retelling performance was quantified using information units per minute (IUs/min), a measure of informativeness in story retelling. Working memory was assessed using the Alphabet Span Test.

Results: Participants with COVID-19 produced fewer information units per minute (mean 0.53, SD 0.21) than those without COVID-19 (mean 0.63, SD 0.24; P=.049). No significant group differences were found in verbal working memory performance (P=.20). However, regression analysis showed that verbal working memory significantly predicted story retelling performance (R^2=.064, P=.02), suggesting that individual differences in working memory capacity may contribute to discourse informativeness, regardless of COVID-19 history.

Conclusions: Young adults with a history of COVID-19 exhibited reduced story retelling performance compared to those without a history of infection. In contrast, no significant differences were observed in verbal working memory performance between groups. Furthermore, verbal working memory scores significantly predicted story retelling performance, suggesting a functional link between these cognitive-linguistic domains. These findings suggest that story retelling performance may serve as a sensitive indicator of post–COVID-19 cognitive-linguistic changes in young adults.

(JMIR Hum Factors 2025;12:e71303) doi:10.2196/71303

KEYWORDS

COVID-19; story retelling; verbal working memory; young adults; cognitive; linguistic

Introduction

COVID-19 is caused by SARS-CoV-2, and this virus is known to cause respiratory illness in humans. After the first case was reported in December 2019, the total cumulative cases worldwide reached 770,875,433 by October 3, 2023, according to the World Health Organization [1]. Since the introduction of vaccines, the number of cases has decreased dramatically.

While hospitalizations and severe cases of respiratory symptoms have decreased, the impact of COVID-19 on cognition has persisted. Numerous studies have reported cognitive dysfunction in individuals with a history of COVID-19. A recent systematic review article [2] reported that the results are inconsistent. Another study [3] found that age was a risk factor, while other study [4] found that middle-aged individuals were the most vulnerable population. Gender differences were also found [5] in cognitive impairment, and women tend to have higher cognitive impairment than men.

Regarding measures tested for the cognitive abilities, the most common neuropsychological screening tests were commonly used in studies [6,7], but an inconsistency was also found here: Quan et al [2] claimed that the timing of cognitive testing might

XSL•FO
RenderX

be the reason for the inconsistency of the results. Other tasks were also used, such as the Digit Span Test-forward and backwards, Trail Making Test-parts A and B, Stroop Word reading and Colour Naming, Corsi backwards, Stroop test, Rey auditory verbal learning test, Free and Cued Selective Reminding Test, Rey–Osterrieth Complex Figure, and Boston Naming Test, Verbal [8-12]. One study [13] found that the severity of COVID-19 was related to poor performance in executive function tasks (Digit symbol, Trail Making Test B, and phonetic fluency). Akıncı et al [14] also found that COVID-19 affected global cognitive skills, which were evaluated using the Montreal Cognitive Assessment and Clock-Drawing Test; memory functions, evaluated using the Öktem Verbal Memory Processes Test; attention span, evaluated using the Digit Span Test; executive functions, evaluated using fluency tests, the Stroop test, and the Trail Making Test; visual perceptual skills, evaluated using the Rey Osterrieth Complex Figure test; and neuropsychiatric status, evaluated using the Neuropsychiatric Inventory.

Although some studies included verbal language tasks such as the Boston Naming Test, discourse-level spoken language, particularly story retelling, has not been examined in COVID-19–related cognitive research in relation to verbal working memory. Story retelling is a cognitively demanding task that engages multiple components of verbal working memory, including temporary storage, manipulation, and retrieval of linguistic information [15,16]. Successful retelling requires not only remembering the sequence of events [17] but also integrating details, maintaining coherence, and regenerating narrative structure [18], all of which rely on both working memory capacity ([19]) and attentional control [20,21].

Emerging research suggests that COVID-19 infection may affect higher-order cognitive functions, such as sustained attention, executive function, and verbal memory [9,22-24]. These functions are closely tied to narrative production abilities, making story retelling a sensitive measure for detecting subtle post–COVID-19 cognitive changes, particularly in young adults with no overt language deficits.

Some of the studies used working memory tasks, but the format of the working memory tasks was based on visual working memory. According to Baddeley's [25, 26] working memory model, visual working memory and verbal working memory are 2 distinct cognitive systems. Visual working memory is responsible for storing information and manipulating visual information such as shapes, colours, and images. Verbal working memory handles language-based information and involves also temporarily storing information and manipulating verbal or auditory information such as words, sentences, and numbers. Therefore, the nature of the information process is different, and it is processed in different systems: the visuospatial sketchpad for visual working memory and the phonological loop for verbal working memory. Considering that spoken discourse production has been shown to be a sensitive clinical indicator of language processing deficits in people with atypical neurogenic language disorders (eg, aphasia) [16,27,28], it is imperative to investigate the impact of COVID-19 on spoken discourse performance.

There are different types of spoken discourse formats depending on how the tasks are administered, and there are mainly 2 different types of tasks, one of which is story retelling. In this format, the examiner read a story, and the participants were asked to retell the story immediately after hearing it. The story retelling type taxes memory capacity [16] and is affected by recency and/or primacy effect [29]. Another format is based on picture description. In this format, participants can rely on the picture in front of them; therefore, the task places lower cognitive demands on memory compared to the story retelling format. Yoo et al [16] found the correlation between the verbal working memory tasks with the retelling type of discourse production results in people with aphasia. This aligns with the connection between verbal working memory and various levels of language processing, as extensively documented in the literature [26,30-32].

Therefore, the primary goal of this study was to examine whether COVID-19 affects memory-based story retelling performance and verbal working memory. The previously reported correlation between performance on the Story Retelling Procedure (SRP) and verbal working memory were correlated based on was assessed [16]. If working memory has declined after COVID-19, it is reasonable to hypothesize that story retelling performance would also be affected by COVID-19. The following questions were investigated in this study:

1. Are there significant differences in measures of story retelling between individuals with and those without a history of COVID-19?
2. Are there significant differences in the verbal working memory between individuals with and those without a history of COVID-19?
3. Is there a predictive relationship between the story retelling and the verbal working memory scores? In other words, can the verbal working memory scores predict the story retelling scores?

Methods

Ethical Considerations

This study was approved by the institutional review board at Baylor University (approval number 1655082 - 11). All participants provided verbal informed consent prior to participation. Some personally identifiable information was collected as part of the study. All identifiable data were removed prior to analysis, and the remaining materials were securely stored on a university-provided cloud platform with restricted access. Only approved research personnel had access to the data through institutional login credentials. No identifying details (such as names, initials, facial features, or other unique identifiers) are included in the manuscript or any submitted materials. All data are presented in a manner that ensures participant anonymity and privacy. No compensation was provided to the participants.

Participants

In total, 79 young participants in their 20s with and those without a history of COVID-19 participated in this study virtually: (1) the non–COVID-19 group comprised 39

XSL•FO
RenderX

participants (age: mean 21.58, SD 1.24 years; years of education: mean 16.32, SD 0.97), and (2) the COVID-19 group comprised 40 participants (age: mean 21.38, SD 1.63 years; years of education: mean 15.60, SD 1.25; Table 1).

Participants in the non–COVID-19 group were recruited between November 11, 2020, and December 4, 2020, and self-reported not having a history of COVID-19. Participants in the COVID-19 group were recruited between April 13, 2021, and September 20, 2022, and provided self-reported information regarding time since infection and number of infections. All COVID-19–positive participants were nonhospitalized, and symptom severity ratings were not collected using a formal scale. Vaccination status was not obtained, and information regarding post–COVID-19 condition symptoms was not collected, as the data were obtained during the early stage of the pandemic, before the concept of post–COVID-19 condition had been clearly defined. Table 2 provides a summary of COVID-19–related participant characteristics.

Table . Demographic information of participants.

Group	Participants, n	Values, mean (SD)
Age (years)		
Non–COVID-19 group	39	21.58 (1.24)
COVID-19 group	40	21.38 (1.62)
Total	79	—[a]
Years of education		
Non–COVID-19 group	39	16.32 (0.97)
COVID-19 group	40	15.61 (1.25)
Total	79	—

[a]Not applicable.

Table . Self-reported COVID-19 characteristics of the COVID-19 group (n=40).

Variable	Values
Time since infection (months)	1 infection: mean 256.6 (SD 199.28); 2 infections: mean 168.56.93 (SD 130.93)
Number of infections	1 infection: n=34 (85%); 2 infections: n=6 (15%)
Symptom severity	Not quantified; all nonhospitalized
Vaccination status	Not obtained
Post–COVID-19 symptoms	Not collected; data were obtained before the concept of post–COVID-19 was clearly defined

Procedures

All data collection procedures were conducted virtually via Zoom (Zoom Video Communications). Each session lasted less than 1 hour per participant.

The SRP Task

The stories used in the SRP task were adapted from the Discourse Comprehension Test [33], which includes 10 stories that were normed and matched on key linguistic complexity variables such as number of words, sentence length, number of subordinate clauses, listening difficulty, and number of main ideas. The validity and reliability of the SRP as a discourse elicitation procedure have been established through validation studies [34,35], which demonstrated strong correlations across multiple linguistic domains with other established elicitation tasks, including picture descriptions and procedural narratives.

SRP participants listened to a short story, and they were asked to retell it immediately afterward. Three short stories were presented, and the averaged raw scores were used for the data analysis. The story information units were checked using the scoring sheet and correct information units per minute was calculated as an outcome variable.

The Working Memory Task

Each participant also completed Alphabet Span Task as a working memory task. In the task, each participant listened to a list of words and was asked to reorder the words in alphabetical order. The list of words was increased as the span increased. The total correct numbers of items were used as the outcome variable for this working memory test.

Statistical Analysis

Statistical analyses for this project were conducted using Python (version 3.10), with the statsmodels and scipy packages for ANOVA and regression modeling. The dataset was assessed for normality, and due to skewed distributions, a natural log transformation (ln) was applied to the story retelling variable (information units per minute [IUs/min]). Two outcome variables were analyzed: (1) log-transformed IUs/min derived from the SRP task and (2) the total number of correct responses

XSL•FO
RenderX

on the Alphabet Span task as a measure of verbal working memory.

To address the first research question, a 1-way between-subjects ANOVA was conducted to compare story retelling performance between the COVID-19 and non–COVID-19 groups. The dependent variable was the number of information units produced per minute, log-transformed to meet assumptions of normality.

Results

Overview

The results revealed a significant between-group difference ($F_{1,77}$=3.99, P=.049), with the COVID-19 group producing fewer information units per minute (mean 0.53, SD 0.21) than the non–COVID-19 group (mean 0.63, SD 0.24). For the second research question, a separate 1-way ANOVA was conducted to compare verbal working memory performance between groups, measured by total number of correct items on the Alphabet Span task. The analysis indicated no significant difference between the COVID-19 (mean 16.21, SD 2.94) and non–COVID-19 (mean 15.30, SD 3.29) groups ($F_{1,77}$=1.66, P=.20). To address the third research question, a simple linear regression was conducted to evaluate the extent to which verbal working memory performance predicted story retelling performance. The regression model was significant ($F_{1,77}$=5.31, P=.02), and accounted for 6.4% of the variance in log-transformed IUs/min (R^2=.064).

The regression equation was:

Story retelling performance (log IUs/min)=−0.493+0.014×verbal working memory performance (Alphabet Span Test: the total correct numbers of items).

The unstandardized regression coefficient (B=0.014) indicates that each additional point on the Alphabet Span Test resulted in a 0.014 increase in log-transformed IUs/min. This corresponds to an approximate 1.4% increase on the raw scale:

(1)exp(0.014)= e{0.014}≈1.014

This back-transformation should be interpreted cautiously. All statistical inferences were based on log-transformed values due to nonnormality in the original distribution. The 95% CI for the slope ranged from 0.002 to 0.026, suggesting a modest but statistically reliable association between verbal working memory and story retelling performance.

Additional Exploratory Analyses

Additional exploratory analyses were conducted within the COVID-19 group to examine whether the number of days since infection was associated with performance on cognitive-linguistic tasks. Time since infection was not significantly associated with story retelling performance (r=.10; P=.57; R^2=.009) or with verbal working memory scores (r=.04; P=.83; R^2=.001). These results suggest that recovery time did not significantly account for performance variability in either task. All additional statistical analyses were conducted using Python (version 3.10) with the pandas, scipy.stats, and seaborn libraries. Exploratory correlation analyses were performed using the Python pearsonr function to assess linear associations between variables of interest, and simple linear regression was conducted using the linregress function.

Discussion

Principal Findings

This study aimed to investigate the impact of COVID-19 on spoken story retelling and verbal working memory and to examine the predictive relationship between the story retelling and the verbal working memory in young adults. The results found that the story retelling performance of the COVID-19 group was significantly lower than that of the non–COVID-19 group. However, despite this performance difference in story retelling, no significant difference was observed between the groups in verbal working memory. This highlights an interesting contrast, suggesting that while the COVID-19 group showed reduced story retelling performance, their verbal working memory capacity did not differ significantly from that of the non–COVID-19 group, indicating that other factors beyond memory might contribute to the observed differences in story retelling performance. In the results of regression analysis, the spoken story retelling performance, measured by %IUs/min, was significantly predicted by the verbal working memory span, even though the prediction was weak, indicating that the relationship between these 2 variables is complex and may be influenced by other factors not captured in this analysis. These results indicate several critical updates of the impact of the COVID-19 on spoken language processing and verbal working memory.

SRP Results

Our findings reveal a surprising impact of COVID-19 on the spoken language of young adults, challenging the prevailing assumption that only infants and children in the developmental stage are susceptible to such effects. Although prior research has predominantly linked COVID-19 with developmental language disorders or delays, particularly in infants and children [36-39], this study underscores the unique vulnerability of healthy young adults to disruptions in story retelling as a result of COVID-19 infection. While children in the developmental stage are generally considered more vulnerable to language impairments, young adults typically exhibit stable story retelling abilities. However, this study suggests that COVID-19 may be associated with disruptions in these abilities. The duration and trajectory of such effects remain unclear, particularly in individuals experiencing post–COVID-19 condition. Future longitudinal research is needed to clarify whether and how story retelling performance may change over time. Additionally, this study suggests that story retelling may be especially sensitive in detecting the cognitive impact of COVID-19. This heightened sensitivity may be attributed to the complexity of the language system, wherein discourse-level processing involves multiple cognitive and linguistic components. In this context, the SRP, which heavily relies on auditory comprehension and various memory functions [16,29,40-42], emerges as a particularly sensitive tool. Both auditory comprehension and memory have been identified as vulnerable indicators of disruptions in language processing abilities [43,44]. Given the complexity of

XSL•FO
RenderX

language processing assessed by the SRP, it holds promise as an invaluable tool for detecting the nuanced effects of COVID-19 on story retelling in young adults.

Verbal Working Memory Findings

Conversely, we observed no significant impact of COVID-19 on verbal working memory. These findings contrast with those reported in previous literature, which has documented negative effects of COVID-19 on working memory [45]. One possible explanation for this discrepancy is the age group of the participants in this study. According to recent research [45], the detrimental effects of COVID-19 on working memory function are primarily observed in adults aged 25 years and older. In contrast, the average age of the participants in the COVID-19 group in this study was 21.37 years, with the non–COVID-19 group also consisting of individuals in a similar age range. This age-related difference may account for the absence of significant working memory decline in the younger cohort in this study. Another potential factor is that none of the participants had been hospitalized for COVID-19, and most reported only mild symptoms. It is also possible that some were asymptomatic. This limited severity may help explain the absence of significant group differences in working memory performance. Moreover, this study did not collect data on whether the COVID-19 group experienced brain fog or post–COVID-19 condition, as data collection took place shortly after the acute phase of COVID-19 infection. It is possible that participants did not exhibit symptoms of post–COVID-19 condition related to cognitive decline or other persistent neuropsychological impairments. Given these factors, further research is needed to examine verbal working memory in individuals experiencing symptoms of post–COVID-19 condition, as this cohort may exhibit different cognitive outcomes.

Results of Regression Modeling

Regression modeling revealed that the story retelling performance was predicted by the verbal working memory test, specifically the Alphabet Span task. This association between SRP and verbal working memory was previously identified in a study involving stroke survivors with aphasia [16]. Revisiting these results with a narrowed focus on individuals exhibiting symptoms of post–COVID-19 condition is crucial to further clarify this predictive relationship further.

Limitations

As discussed earlier, a notable limitation of this study is the absence of information regarding whether participants with a history of COVID-19 had experienced post–COVID-19 condition. This gap in data collection stems from the timing of the study, which occurred shortly after the onset of COVID-19. In addition, this study was conducted in a virtual testing environment, which may have introduced variability in participants' attention, environmental distractions, or audio quality, potentially affecting performance. Another limitation is the possibility of selection bias, as participants volunteered to complete the study remotely and may not fully represent the broader young adult population. Furthermore, the regression analyses yielded relatively low predictive power, suggesting that while certain predictors showed statistical associations, they may not account for a substantial proportion of variance in story retelling performance.

Conclusion

This study identified differences in spoken story retelling performance between individuals with and those without a history of COVID-19. Additionally, story retelling performance was predicted by the Alphabet Span Test, a measure of verbal working memory. These findings suggest that story retelling may be sensitive to subtle post–COVID-19 cognitive-linguistic changes in young adults. This has potential clinical implications for early cognitive-linguistic screening and targeted support interventions, particularly in educational and clinical settings.

Data Availability

The datasets generated and analyzed during this study are not publicly available due to privacy concerns, but deidentified data may be made available from the author upon reasonable request.

Conflicts of Interest

None declared.

References

1. COVID-19 cases, world. World Health Organization. URL: https://data.who.int/dashboards/covid19/cases?n=c [accessed 2025-07-17]
2. Quan M, Wang X, Gong M, Wang Q, Li Y, Jia J. Post-COVID cognitive dysfunction: current status and research recommendations for high risk population. Lancet Reg Health West Pac 2023 Sep;38:100836. [doi: 10.1016/j.lanwpc.2023.100836] [Medline: 37457901]
3. Evans RA, Leavy OC, Richardson M, et al. Clinical characteristics with inflammation profiling of long COVID and association with 1-year recovery following hospitalisation in the UK: a prospective observational study. Lancet Respir Med 2022 Aug;10(8):761-775. [doi: 10.1016/S2213-2600(22)00127-8]
4. Matias-Guiu JA, Herrera E, González-Nosti M, et al. Development of criteria for cognitive dysfunction in post-COVID syndrome: the IC-CoDi-COVID approach. Psychiatry Res 2023 Jan;319:115006. [doi: 10.1016/j.psychres.2022.115006] [Medline: 36521337]

XSL•FO
RenderX

5. Ceban F, Ling S, Lui LMW, et al. Fatigue and cognitive impairment in Post-COVID-19 Syndrome: A systematic review and meta-analysis. Brain Behav Immun 2022 Mar;101:93-135. [doi: 10.1016/j.bbi.2021.12.020] [Medline: 34973396]
6. Petersen EL, Gossling A, Adam G, et al. Multi-organ assessment in mainly non-hospitalized individuals after SARS-CoV-2 infection: The Hamburg City Health Study COVID programme. Eur Heart J 2022;43(11):1124-1137. [doi: 10.1093/eurheartj/ehac086]
7. Blazhenets G, Schroeter N, Bormann T, et al. Slow but evident recovery from neocortical dysfunction and cognitive impairment in a series of chronic COVID-19 patients. J Nucl Med 2021 Jul 1;62(7):910-915. [doi: 10.2967/jnumed.121.262128] [Medline: 33789937]
8. Crivelli L, Palmer K, Calandri I, et al. Changes in cognitive functioning after COVID-19: A systematic review and meta-analysis. Alzheimers Dement 2022 May;18(5):1047-1066. [doi: 10.1002/alz.12644] [Medline: 35297561]
9. Becker JH, Lin JJ, Doernberg M, et al. Assessment of cognitive function in patients after COVID-19 infection. JAMA Netw Open 2021 Oct 1;4(10):e2130645. [doi: 10.1001/jamanetworkopen.2021.30645] [Medline: 34677597]
10. Delgado-Alonso C, Valles-Salgado M, Delgado-Álvarez A, et al. Cognitive dysfunction associated with COVID-19: A comprehensive neuropsychological study. J Psychiatr Res 2022 Jun;150:40-46. [doi: 10.1016/j.jpsychires.2022.03.033] [Medline: 35349797]
11. García-Sánchez C, Calabria M, Grunden N, et al. Neuropsychological deficits in patients with cognitive complaints after COVID-19. Brain Behav 2022 Mar;12(3):e2508. [doi: 10.1002/brb3.2508] [Medline: 35137561]
12. Whiteside DM, Basso MR, Naini SM, et al. Outcomes in post-acute sequelae of COVID-19 (PASC) at 6 months post-infection part 1: Cognitive functioning. Clin Neuropsychol 2022 May;36(4):806-828. [doi: 10.1080/13854046.2022.2030412] [Medline: 35130818]
13. Ariza M, Cano N, Segura B, et al. COVID-19 severity is related to poor executive function in people with post-COVID conditions. J Neurol 2023 May;270(5):2392-2408. [doi: 10.1007/s00415-023-11587-4] [Medline: 36939932]
14. Akıncı B, Oğul Ö, Hanoğlu L, et al. Evaluation of cognitive functions in adult individuals with COVID-19. Neurol Sci 2023 Mar;44(3):793-802. [doi: 10.1007/s10072-022-06562-2] [Medline: 36574178]
15. Gabig CS. Verbal working memory and story retelling in school-age children with autism. Lang Speech Hear Serv Sch 2008 Oct;39(4):498-511. [doi: 10.1044/0161-1461(2008/07-0023)] [Medline: 18820091]
16. Yoo H, McNeil MR. Story retelling and verbal working memory in people with aphasia. Clin Arch Commun Disord 2019 Dec;4(3):223-235. [doi: 10.21849/cacd.2019.00157]
17. Cain K, Oakhill J, Bryant P. Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. J Educ Psychol 2004;96(1):31-42. [doi: 10.1037/0022-0663.96.1.31]
18. Wright HH, Fergadiotis G. Discourse macrostructure in aphasia: An integrative framework. Top Lang Disord 2012;32(4):278-292. [doi: 10.1097/TLD.0b013e318267b460]
19. Archibald LMD, Gathercole SE. Short-term and working memory in specific language impairment. Int J Lang Commun Disord 2006;41(6):675-693. [doi: 10.1080/13682820500442602] [Medline: 17079222]
20. Engle RW. Working memory capacity as executive attention. Curr Dir Psychol Sci 2002 Feb;11(1):19-23. [doi: 10.1111/1467-8721.00160]
21. Kane MJ, Engle RW. Working-memory capacity and the control of attention: the contributions of goal neglect, response competition, and task set to Stroop interference. J Exp Psychol Gen 2003 Mar;132(1):47-70. [doi: 10.1037/0096-3445.132.1.47] [Medline: 12656297]
22. Zhou H, Lu S, Chen J, et al. The landscape of cognitive function in recovered COVID-19 patients. J Psychiatr Res 2020 Oct;129:98-102. [doi: 10.1016/j.jpsychires.2020.06.022] [Medline: 32912598]
23. Woo MS, Malsy J, Pöttgen J, et al. Frequent neurocognitive deficits after recovery from mild COVID-19. Brain Commun 2020;2(2):fcaa205. [doi: 10.1093/braincomms/fcaa205] [Medline: 33376990]
24. Almeria M, Cejudo JC, Sotoca J, Deus J, Krupinski J. Cognitive profile following COVID-19 infection: Clinical predictors leading to neuropsychological impairment. Brain Behav Immun Health 2020 Dec;9:100163. [doi: 10.1016/j.bbih.2020.100163] [Medline: 33111132]
25. Baddeley AD. Working Memory: Oxford: Clarendon Press; 1986.
26. Baddeley AD. The episodic buffer: a new component of working memory? Trends Cogn Sci 2000;4(11):417-423. [doi: 10.1016/S1364-6613(00)01483-4] [Medline: 11058819]
27. Armstrong E. Aphasic discourse analysis: The story so far. Aphasiology 2000 Sep;14(9):875-892. [doi: 10.1080/02687030050127685]
28. Ulatowska HK, Freedman-Stern R, Doyel AW, Macaluso-Haynes S, North AJ. Production of narrative discourse in aphasia. Brain Lang 1983 Jul;19(2):317-334. [doi: 10.1016/0093-934x(83)90074-3] [Medline: 6192869]
29. Brodsky MB, McNeil MR, Doyle PJ, Fossett TRD, Timm NH, Park GH. Auditory serial position effects in story retelling for non-brain-injured participants and persons with aphasia. J Speech Lang Hear Res 2003 Oct;46(5):1124-1137. [doi: 10.1044/1092-4388(2003/088)] [Medline: 14575347]
30. Daneman M, Merikle PM. Working memory and language comprehension: A meta-analysis. Psychon Bull Rev 1996 Dec;3(4):422-433. [doi: 10.3758/BF03214546] [Medline: 24213976]
31. Gathercole SE, Baddeley AD. Working Memory and Language: Lawrence Erlbaum Associates; 1993.

XSL•FO
RenderX

32. Just MA, Carpenter PA. A capacity theory of comprehension: individual differences in working memory. Psychol Rev 1992 Jan;99(1):122-149. [doi: 10.1037/0033-295x.99.1.122] [Medline: 1546114]
33. Brookshire RH, Nicholas LE. Discourse Comprehension Test: BRK Publishers; 1997.
34. McNeil MR, Doyle PJ, Fossett TRD, Park GH, Goda AJ. Reliability and concurrent validity of the information unit scoring metric for the story retelling procedure. Aphasiology 2001 Oct;15(10-11):991-1006. [doi: 10.1080/02687040143000348]
35. McNeil MR, Sung JE, Yang D, et al. Comparing connected language elicitation procedures in persons with aphasia: Concurrent validation of the Story Retell Procedure. Aphasiology 2007 Jun;21(6-8):775-790. [doi: 10.1080/02687030701189980]
36. Hsu SC, Wong AMK. Association between interruption of intervention and language performance in young children with language delay-a cohort study during COVID-19 pandemic. Front Pediatr 2023;11:1240354. [doi: 10.3389/fped.2023.1240354] [Medline: 37780049]
37. Feijoo S, Amadó A, Sidera F, Aguilar-Mediavilla E, Serrat E. Language acquisition in a post-pandemic context: the impact of measures against COVID-19 on early language development. Front Psychol 2023;14:1205294. [doi: 10.3389/fpsyg.2023.1205294] [Medline: 37575436]
38. Deoni S. Impact of the COVID-19 pandemic environment on early child brain and cognitive development. Biol Psychiatry Cogn Neurosci Neuroimaging 2022 May;91(9):S26. [doi: 10.1016/j.biopsych.2022.02.082]
39. LoBue V, Pérez-Edgar K, Kirkham N, Herbert J. The impact of COVID-19 on infant development: A special issue of infancy. Infancy 2023 Jan;28(1):4-7. [doi: 10.1111/infa.12528] [Medline: 36629027]
40. Boudreau D. Narrative abilities: Advances in research and implications for clinical practice. Top Lang Disord 2008;28(2):99-114. [doi: 10.1097/01.TLD.0000318932.08807.da]
41. Petersen DB, Spencer TD. The Narrative Language Measures: Tools for language screening, progress monitoring, and intervention planning. Perspect Lang Learn Educ 2012 Oct;19(4):119-129. [doi: 10.1044/lle19.4.119]
42. Wagner CR, Sahlén B, Nettelbladt U. What's the story? Narration and comprehension in Swedish preschool children with language impairment. Child Lang Teach Ther 1999 Jun;15(2):113-137. [doi: 10.1177/026565909901500202]
43. Wingfield A, Tun PA. Cognitive supports and cognitive constraints on comprehension of spoken language. J Am Acad Audiol 2007;18(7):548-558. [doi: 10.3766/jaaa.18.7.3] [Medline: 18236643]
44. Hickok G, Poeppel D. The cortical organization of speech processing. Nat Rev Neurosci 2007 May;8(5):393-402. [doi: 10.1038/nrn2113] [Medline: 17431404]
45. Baseler HA, Aksoy M, Salawu A, Green A, Asghar AUR. The negative impact of COVID-19 on working memory revealed using a rapid online quiz. PLOS ONE 2022;17(11):e0269353. [doi: 10.1371/journal.pone.0269353] [Medline: 36374838]

Abbreviations

IUs/min: information units per minute
SRP: Story Retelling Procedure

Edited by A Parush; submitted 14.01.25; peer-reviewed by C Udensi, F Schneider, T Adebambo; revised version received 28.05.25; accepted 03.06.25; published 05.08.25.

Please cite as:
Yoo H
Story Retelling and Verbal Working Memory in Young Adults With a History of COVID-19: Cross-Sectional Study
JMIR Hum Factors 2025;12:e71303
URL: https://humanfactors.jmir.org/2025/1/e71303
doi:10.2196/71303

© Hyunsoo Yoo. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 5.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluating the Acceptance and Usability of an Independent, Noncommercial Search Engine for Medical Information: Cross-Sectional Questionnaire Study and User Behavior Tracking Analysis

Lisa Specht[1*], MD; Raphael Scheible[2*], MSc; Martin Boeker[2], Prof Dr Med; Erik Farin-Glattacker[3], Prof Dr; Nikolas Kampel[4], MSc; Marina Schmölz[1], MSc; Andrea Schöpf-Lazzarino[1,5], PhD; Stefan Schulz[6], Prof Dr Med; Christian Schlett[3], PhD; Fabian Thomczyk[7]; Sebastian Voigt-Radloff[3], PhD; Constanze Wegner[8], MA; Katharina Wollmann[8], MPH; Andy Maun[1], Prof Dr

[1]Institute of General Practice, Faculty of Medicine and Medical Center, University of Freiburg, Freiburg, Germany

[2]Institute of Artificial Intelligence and Informatics in Medicine, Chair of Medical Informatics, University Hospital rechts der Isar, School of Medicine and Health, Technical University of Munich, Munich, Germany

[3]Institute of Medical Biometry and Statistics, Section of Health Care Research and Rehabilitation Research, Faculty of Medicine and Medical Center, University of Freiburg, Freiburg, Germany

[4]Institute of Neuroscience and Medicine, Jülich Aachen Research Alliance, Forschungszentrum Jülich GmbH, Jülich, Germany

[5]Careum School of Health, part of the Kalaidos University of Applied Sciences, Zurich, Switzerland

[6]Institute for Medical Informatics, Statistics and Documentation, Medical University of Graz, Graz, Austria

[7]Data Integration Center, University of Freiburg, Freiburg, Germany

[8]Cochrane Germany, Cochrane Germany Foundation, Freiburg, Germany

[*]these authors contributed equally

Corresponding Author:
Lisa Specht, MD
Institute of General Practice
Faculty of Medicine and Medical Center
University of Freiburg
Elsässer Straße 2m
Freiburg, 79110
Germany
Phone: 49 761 270 72490
Email: lisa.specht@uniklinik-freiburg.de

Abstract

Background: The internet is a key source of health information, but the quality of content from popular search engines varies, posing challenges for users—especially those with low health or digital health literacy. To address this, the “tala-med” search engine was developed in 2020 to provide access to high-quality, evidence-based content. It prioritizes German health websites based on trustworthiness, recency, user-friendliness, and comprehensibility, offering category-based filters while ensuring privacy by avoiding data collection and advertisements.

Objective: This study aims to evaluate the acceptance and usability of this independent, noncommercial search engine from the users’ perspectives and their actual use of the search engine.

Methods: For the questionnaire study, a cross-sectional study design was used. In total, 802 participants were recruited through a web-based panel and were asked to interact with the new search engine before completing a web-based questionnaire. Descriptive statistics and multiple regression analyses were used to assess participants’ acceptance and usability ratings, as well as predictors of acceptance. Furthermore, from October 2020 to June 2021, we used the open-source web analytics platform Matomo to collect behavior-tracking data from consenting users of the search engine.

Results: The study indicated positive findings on the acceptance and usability of the search engine, with more than half of the participants willing to reuse (465/802, 58%) and recommend it (507/802, 63.2%). Of the 802 users, 747 (93.1%) valued the

XSL•FO
RenderX

absence of advertising. Furthermore, 92.3% (518/561), 93.9% (553/589), 94.7% (567/599), and 96.5% (600/622) of those users who used the filters agreed at least partially that the filter functions were helpful in finding trustworthy, recent, user-friendly, or comprehensible results. Participants criticized some of the search results regarding the selection of domains and shared ideas for potential improvements (eg, for a clearer design). Regression analyses showed that the search engine was especially well accepted among older users, frequent internet users, and those with lower educational levels, indicating an effective targeting of segments of the population with lower health literacy and digital health literacy. Tracking data analysis revealed 1631 sessions, comprising 3090 searches across 1984 unique terms. Users performed 1.64 (SD 1.31) searches per visit on average. They prioritized the search terms “corona,” “back pain,” and “cough.” Filter changes were common, especially for *recency* and *trustworthiness*, reflecting the importance that users placed on these criteria.

Conclusions: User questionnaires and behavior tracking showed the platform was well received, particularly by older and less educated users, especially for its advertisement-free design and filtering system. While feedback highlighted areas for improvement in design and filter functionality, the search engine’s focus on transparency, evidence-based content, and user privacy shows promise in addressing health literacy and navigational needs. Future updates and research will further refine its effectiveness and impact on promoting access to quality health information.

(JMIR Hum Factors 2025;12:e56941) doi:10.2196/56941

KEYWORDS

medical information; health information; search engine; user behavior; health literacy; digital health literacy; navigational needs; information-seeking behavior; evidence-based content; Germany

Introduction

Background

The internet has become a crucial resource for accessing medical information, making it a valuable tool for individuals seeking knowledge about health. In Germany, medical professionals, notably physicians, have long been the primary source of health care information for the general population [1,2]. However, the landscape is evolving, and the internet now stands as the second most significant channel for health-related information [1], indicating a new trend in how people seek medical knowledge. A study conducted by the Bertelsmann Foundation showed a strong and increasing demand for health-related information online, including a rapidly increasing demand by older generations [3]. In 2020, approximately 70% of the German population used the internet to actively engage in searching for health-related content [4]. While these numbers have varied in recent years, there is an overall trend of an increased use of the internet [4].

A major complaint that users have about health-related information online is the lack of clarity regarding the trustworthiness and seriousness of existing websites. The information online has varying levels of quality [5,6]. Frequently used and popular websites are regarded as reliable by users, irrespective of the actual quality of their content, while independent public websites are not as well known and do not seem to be more reliable to users [3]. The reliance on digital platforms for health information also poses a challenge due to the profound influence of the order of the results generated by popular search engines. Users often initiate their search for medical information through well-known search engines, predominantly relying on the first few search results displayed on the search engine results page (SERP), which are often advertising [7-10]. Unfortunately, these top hits are not only clicked on more frequently but are also perceived as more trustworthy, despite the variation in information quality [3,9]. This highlights the ambiguity of users feeling incompetent in finding reliable information online, while at the same time trusting the top results and being satisfied with the information they find. A more recent study evaluated the quality of online health and nutrition information related to cancer supplements through a Google search, using the Health Information Quality Index [11]. The 160 relevant search results yielded median and mean Health Information Quality Index scores of 8, with one-quarter of the results scoring high (10-12). No correlation was found between high quality scores and an early appearance of these results, indicating potential limitations in using Google for obtaining accurate information on dietary supplements and cancer, particularly given the prevalence of advertisements outnumbering search results [11].

The problem with the reliance on search engines becomes even clearer when considering the low health literacy and digital health literacy of the German population [12-14]. The first Health Literacy Survey Germany study in 2014 [13] underscored a significant health literacy challenge in Germany, with 54.3% of the population reporting health literacy problems. This result pointed to challenges in navigating the health care system and understanding health-related information, with identified associations with factors such as age, migrant background, self-assessed social status, and functional literacy. The second Health Literacy Survey Germany study in 2020 [14] reinforced previous concerns with a further increased rate of individuals (58.8%) struggling considerably with health information. Even more so than in 2014, evaluating the trustworthiness of health-related information presented difficulties [15]. Up to 45% of the German population exhibit “problematic” or even “inadequate” internet skills [16]. Many individuals face so-called navigational needs. They depend on support from others to efficiently search and evaluate health information online [17]. As the amount of digital health information grows each year, the challenge of effectively navigating through this vast array of content becomes increasingly daunting, raising concerns about misinformation and health-related misconceptions.

XSL•FO
RenderX

During the work on the comprehensive GAP (Gut informierte Kommunikation zwischen Arzt und Patient, meaning "Well-informed communication between physician and patient" in English) project, which addressed evidence-based information [18-26], we studied the needs and requirements of internet users for health information online [18]. We developed criteria to assess web domains according to these needs. Assessing the quality of websites is nothing new. There are existing seals of approval, such as the afgis (Aktionsforum Gesundheitsinformationssystem eV, meaning "Action Forum Health Information System eV" in English) and Health On the Net Foundation Code of Conduct seals, which check whether the website operator meets certain transparency and quality standards. An afgis certificate is valid for 1 year. However, with Health On the Net Foundation Code of Conduct, there were several issues [27], and on December 15, 2022, the Health On the Net Foundation discontinued its services. Different seals potentially signify different standards of certification. Another issue with these seals is that the pages that have a seal cannot easily be found via a central search. To address these issues by centrally and consistently rating domains inside a search platform, we developed a search engine in 2020. This search engine was originally called "GAP search" and later renamed as "tala-med Suche" in German and "tala-med search" in English.

The topic of data retrieval in medicine is generally relevant and widely researched. Efforts ranging from decentralized search systems for patient data in registries [28] to tools such as Informatics for Integrating Biology and the Bedside (i2b2) [29] that allow the creation of cohorts, for which data security plays an important role [30], have been investigated. In addition, for public data, such as the Medical Subject Headings (MeSH), search capabilities play an important role [31], including the requirement for an intuitive user interface design [32]. In terms of scraping web content, which could be used to create a search platform, the Sampled German Health Web [33,34] created an index of German-language health-related web pages using a special focused crawler. The Sampled German Health Web index was restricted to pages from the top-level domains .de, .at, and .ch, and the domains were automatically filtered for health-related content while crawling using a support vector machine. By contrast, the new tala-med search engine, which was used in this study, relied on hand-picked, high-quality, German-language health websites. These domains—more than 50 of them—underwent a rigorous evaluation process. This evaluation was based on the categories *trustworthiness* (with the subcategories *authority*, *independence*, and *evidence based*), *recency*, *user-friendliness*, and *comprehensibility*. The ratings of these categories influence the order of the results on the SERP. The search engine also values user privacy by avoiding data collection and advertising. The detailed design of the search engine is described in the *Methods* section and in Multimedia Appendices 1-3.

By assessing the acceptance and usability of this new technology, we aimed to improve the implementation of the search tool. Numerous studies in the field emphasize the significance of outcome measures as crucial indicators of a technology's effectiveness, user-friendliness, and potential for widespread adoption. User acceptance, often determined by perceived usefulness, ease of use, and individual attitudes, is critical for a technology's long-term viability [35,36]. Moreover, the usability of a technology, including learnability, efficiency, and user satisfaction, significantly affects user experience and reduces possible adoption barriers [37]. Evaluating these dimensions can provide insight into how well technology meets users' expectations, addresses their needs, and contributes to successful implementation. In the context of health care technology, this scrutiny is particularly relevant because effective tools for accessing medical information can significantly impact health care outcomes [17,38].

In this study, the acceptance and usability of our newly developed search engine was evaluated by means of a questionnaire study and actual use by tracking user behavior during the first months of operation.

Related Works

The usability and accessibility of digital platforms are essential for ensuring that all users, including those with disabilities, can effectively engage with websites and mobile apps. Mateus et al [39] conducted a systematic mapping of accessibility issues, revealing that automated tests covered <40% of accessibility problems on websites and even fewer on mobile apps. The study stressed the importance of including users with disabilities in evaluations because user testing uncovered many issues missed by automated tools and expert inspections. This underscores the need for a comprehensive evaluation approach that combines expert and real-world user input to improve accessibility.

Petrie and Bevan [40] expanded on this by exploring the interplay between usability, accessibility, and user experience. The authors defined usability in terms of effectiveness, efficiency, and satisfaction, focusing on aspects such as learnability, flexibility, and safety. Importantly, they emphasized that accessibility is an integral component of usability and that digital systems should cater to the widest possible range of users, including those with disabilities. They also introduced the concept of user experience, which goes beyond usability to include users' emotional responses and satisfaction with a system. This holistic approach is essential for creating digital platforms that are both functional and engaging for all users.

Belinda et al [41] provided a deeper look into the internal and external usability factors that affect website performance. The authors used automated tools such as GTmetrix and Website Grader to measure internal attributes such as performance, load time, and page size, while external attributes such as ease of navigation and user satisfaction were assessed through surveys. Their findings revealed that some websites performed well from a user perspective but were found lacking in technical performance, particularly in terms of load times and page requests. This highlights the need to address both internal and external usability factors to create well-rounded digital platforms.

Kritz et al [42] explored the online resources and tools used by European physicians to gather medical information. The authors found that physicians frequently relied on general search engines and faced significant barriers in accessing high-quality,

XSL•FO
RenderX

trustworthy medical content. Medical specialists were more likely to use medical research databases, while general practitioners often faced barriers such as lack of time and language restrictions. The study highlighted the need for improved medical search tools tailored to the specific needs of different physician subgroups. Kritz et al [42] concluded that user-centered medical search tools could significantly improve accessibility and the quality of online medical information.

Strecker et al [32] also contributed to the usability of medical search tools, focusing on the MeSH Browser. The authors evaluated a newly developed multilingual MeSH Browser, which introduced improvements in user interface design to enhance the accessibility and usability of medical literature searches. The results showed that contemporary web design principles led to significant improvements in navigation and overall user satisfaction, further underscoring the importance of continual evaluation and enhancement of information systems in the medical realm.

Eysenbach and Köhler [43] studied how consumers search for and appraise health information on the internet. The qualitative study revealed that users, despite using suboptimal search techniques, were able to retrieve health information quickly. However, users rarely checked critical indicators of credibility, such as the “About us” sections or disclaimers, relying instead on superficial factors such as professional design and ease of use. This reliance on superficial credibility indicators poses risks, particularly in the health care field, where the accuracy of information is paramount. Eysenbach and Köhler [43] suggested that further research is needed to develop educational and technological tools that guide users toward high-quality health information.

Zhang [44] expanded the understanding of how consumers select sources for health information by identifying 5 categories of factors that influence source selection: source-related factors, user-related factors, user-source relationships, characteristics of the problematic situation, and social influences. The study also identified a range of criteria that mediate the influence of these factors on source-selection decisions, including accessibility, quality, usability, interactivity, relevance, usefulness, and familiarity. Zhang [44] concluded that a personalized approach to health information systems is necessary to provide effective access to health information because different consumers prioritize different factors when selecting sources. This insight strongly indicates the need for more personalized information services that cater to individual user preferences and needs.

In conclusion, the literature emphasizes the critical role of accessibility, usability, and user experience in the design of digital platforms, especially health care websites. While Saad et al [45] have highlighted general usability problems in health care websites, the studies by Kritz et al [42], Eysenbach and Köhler [43], and Zhang [44] focus on the challenges that users face in retrieving and assessing the credibility of health information online. A personalized and user-centered approach to the design of health care information systems would improve the accessibility and quality of health information, meeting the diverse needs of consumers and health care professionals alike [40,44].

Methods

Search Engine Development

Requirements

The functional requirements for the tala-med search engine included the ability to crawl and index health information websites, implement quality assessment filters, and provide search term suggestions and synonym handling. The nonfunctional requirements focused on maintaining user privacy through self-hosting, ensuring a user-friendly interface, and optimizing performance to handle large sets of synonym mappings efficiently.

Implementation

Design

To implement tala-med search, we used software with a self-hosting capability because of privacy considerations. Furthermore, modern technology with single-page application design using a web application programming interface was required to compete with modern search engines. Therefore, the selection available was limited. As no single product met all our requirements, we built a search engine stack ourselves, consisting of crawler, middleware, and front end, using existing components and established technologies.

Back End

We created the web index using the open-source crawler software Fess [46]. For boilerplate removal, we used Mozilla’s Readability tool to strip HTML tags and display core content, cleaning up crawled websites for our search index [47]. As middleware, we used Elastic App Search [48] (now part of Elastic Enterprise Search [49]), which relies on Elasticsearch. This software allows for setting up weighting mechanisms for the search equation (Multimedia Appendix 1) and configuring filters. In our case, a set of criteria for quality assessment were developed, which we integrated as filters. These quality assessment criteria were adapted from a systematic review by Eysenbach et al [6] that identified criteria used from 1969 to 2001 for evaluating the quality of health information online. We identified 74 criteria [18] to assess the content quality of different health-related websites and rigorously evaluated >50 German-language health information providers (Multimedia Appendix 2) using these criteria, assessing 1 main page and 5 randomly selected subpages per provider. To compile the list of relevant providers, we started with a list from a Bertelsmann study [3] and adapted it with the help of domain experts. The 74 quality criteria were later condensed into 4 categories: *trustworthiness*, *recency*, *user-friendliness*, and *comprehensibility*. The scores of these categories were added to the search index after boilerplate removal and divided into 4 value ranges reflecting the quality of each category, using 4 quantiles for differentiation. These scores served as quality indicators and filter categories, influencing the order of the results on the SERP. More details about the internal mechanisms of the search ranking can be found in Multimedia Appendix 1.

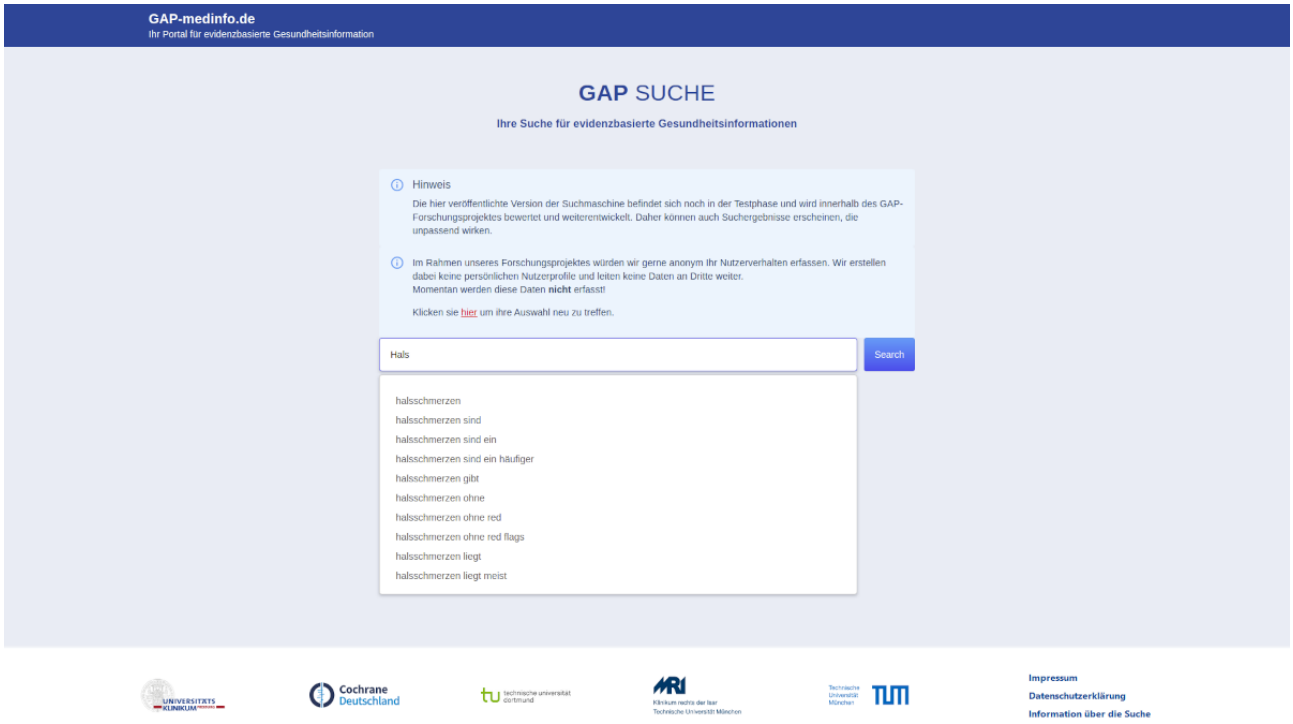
XSL•FO
RenderX

App Search supports the consideration of synonyms, which is crucial in medical language due to varied etymology and numerous abbreviations. Our goal was to enable the search engine to handle synonyms effectively. In App Search, sets of up to 32 synonyms can be created, allowing synonyms of a matched word to be considered in searches. We generated synonym sets by collecting word types and pairs from the document text, filtering against a German stop word list, resulting in approximately 6 million unique entries. These entries were matched against an experimental German interface terminology—SCT-GIT—linked to Systematized Nomenclature of Medicine–Clinical Terms (SNOMED-CT) [50], yielding approximately 40,000 mappings to SNOMED-CT codes. Synonym links in SCT-GIT were then used to add synonyms for each matching string, finding at least 1 synonym for approximately 30,000 SCT-GIT terms. We limited the number of synonym sets to 500 to maintain performance, using only sets with at least 4 synonyms, resulting in a median set size of 4.

Front End

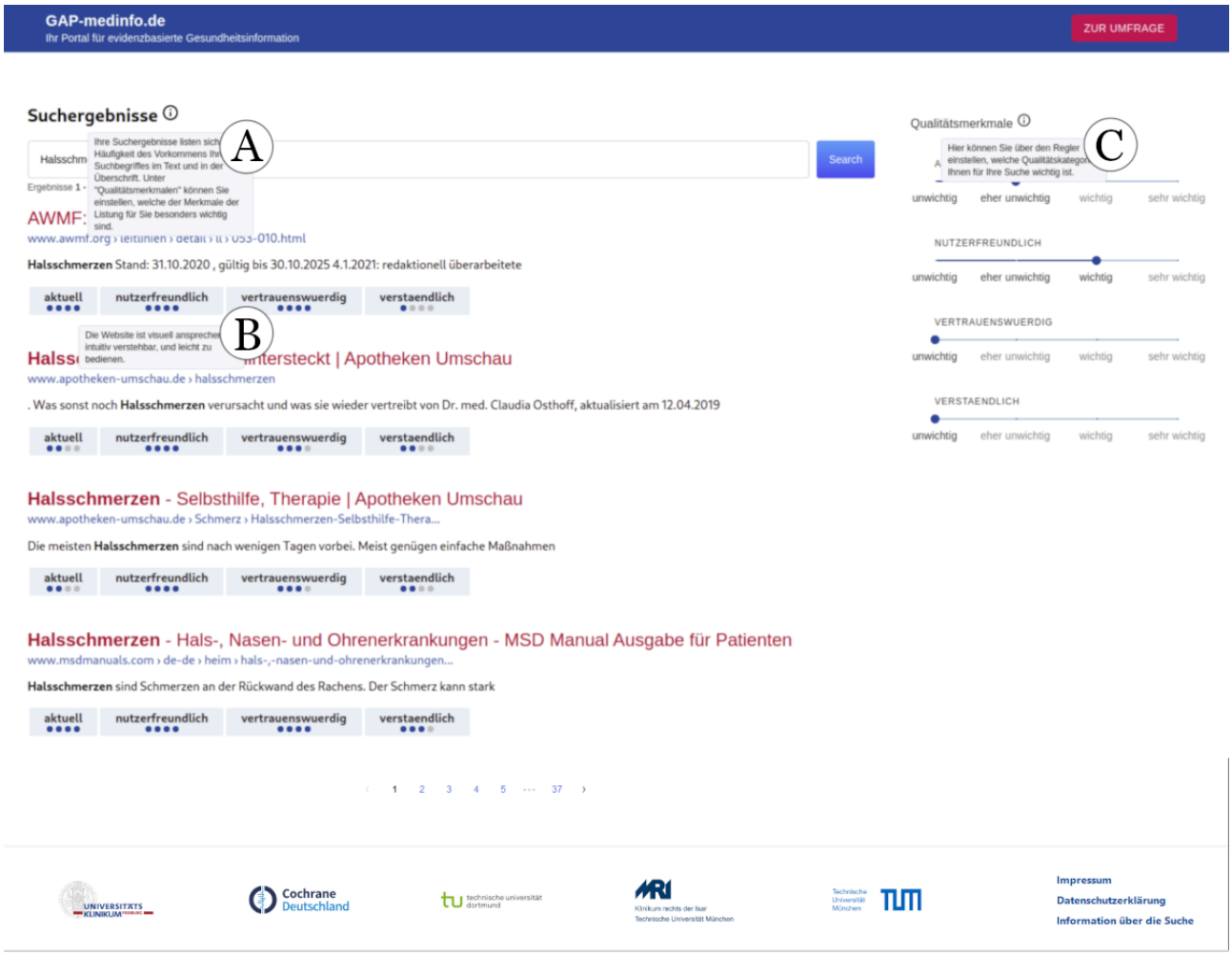
The front end was custom developed with a slim design inspired by Strecker et al [32], featuring search term suggestions and a user-friendly filter selection (Figure 1). After entering a search term, the interface displayed a SERP with evaluated German-language health information providers, showing their scores in the 4 filter categories through graphical indicators (Figure 2). Hovering over these graphics revealed explanations, and the bottom of the page included pagination and a footer with logos and links to pages displaying the imprint and data protection information.

Figure 1. Landing page of the search displaying the initial search query field.



XSL•FO
RenderX

Figure 2. The search engine results page consists of 3 areas. (A) The search query. (B) List of search results. Each search result shows the quality in the 4 filter categories. (C) Slide controls allow users to set a threshold for each filter category (aktuell: recency, nutzerfreundlich: user-friendliness, vertrauenswürdig: trustworthiness, and verständlich: comprehensibility). In each of these areas, more information can be displayed by means of tooltips, as shown in the graphic (next to each label).



Pilot Testing

The search engine was pilot-tested between October 2020 and the end of June 2021. Adapting to the challenges posed by the COVID-19 pandemic, the pilot testing format shifted from a face-to-face setting to a web-based setting with a small group of academic staff from the University of Freiburg. The evaluation, involving 12 participants, consisted of screen recordings of search tasks and a qualitative survey regarding user-friendliness. After the pilot tests, technical enhancements were implemented, guided by the qualitative insights. These findings formed the basis for the subsequent questionnaire study.

Availability

The current version of the search engine can be accessed on the web [51]. Please note that there have been some significant changes to the website structure since this study was conducted.

Evaluation

Design

We used 2 data collection methods to evaluate user behavior. First, we asked users to fill out a questionnaire (self-perception). Second, we recorded user behavior via web tracking independently and unreferenced to the questionnaire (external perception). The study protocol was published in 2019 [22].

Questionnaire Study

The study used a cross-sectional design with data collected through a web-based questionnaire. The questionnaire was developed after pilot testing of the search engine and underwent an internal pretesting round before recruitment began.

Recruitment

The study used a recruitment strategy aimed at enrolling 200 participants. The first rounds of recruitment, starting in October 2020, which involved sending email invitations to local health initiatives, professional networks for physicians, and personal contacts, were unsuccessful in obtaining the desired number of participants. Participants were then recruited from the WisoPanel, an established online panel of German-speaking individuals [52]. The panel provided a pool of 14,900 potential respondents who had previously expressed an interest in participating in research studies. The survey was open for anyone accessing the link, but only participants from the panel (>70% of the data) were included in the analysis of this study to maintain a credible sample. Recruitment via the panel took place in May 2021. The survey ended at the end of May 2021.

XSL•FO
RenderX

Panel members were invited to participate in the study via email. The email informed participants about the project and the aspects distinguishing the new search engine from other platforms.

To ensure that participants had firsthand experience with the platform before proceeding to the survey, they were asked to test the search engine first. We recommended that participants carry out 1 or 2 searches and familiarize themselves with the SERP and the features and functions of the website.

Survey Administration

The hyperlink that led participants to the search engine was sent to the panel members by email. Subsequently, the link to the survey was displayed as a pop-up element after 10 to 20 seconds of interaction with the search engine. Throughout the assessment phase, a button to access the survey was strategically positioned in the top right corner of the website to allow a direct link to the survey interface. No registration was required to complete the survey, ensuring seamless and voluntary participation. Participants had the option to cancel their responses at any time, and there were no time restrictions regarding the completion of the questionnaire. Back buttons allowed participants to review their answers before submitting the survey. Conventional technical methods, such as cookies, were used to prevent visitors from submitting their survey more than once.

The web-based survey was conducted using the Unipark tool [53], with the survey page seamlessly embedded into the search page via a modal window. After inviting panel participants, the survey remained accessible for 11 days and closed at the end of May 2021.

Before filling out the survey, participants were provided with an introductory text outlining the anticipated time commitment, project objectives, details about anonymity, the option to terminate the survey at any point, data protection measures, and contact information for queries. To proceed, participants were required to confirm their understanding by checking a box and consenting to the specified use of their data (informed consent). No personal data were collected in the survey, and participants were instructed not to provide any identifying information in their open responses. Access to the data was restricted to members of the institutions participating in the study.

Survey Characteristics

On the basis of established scales (the German version of the self-assessment eHealth Literacy Scale [G-eHEALS] [54] and the System Usability Scale [55] adapted by Quirmbach [56] and Magin et al [57]), we developed a 25-item questionnaire across five dimensions: (1) sociodemographic data; (2) internet use; (3) digital health literacy; (4) usability, acceptance, and innovative aspects of the search engine; and (5) search filters. The final questionnaire contained 24 single-choice questions and 1 open-ended question (Tables S1 and S2 in Multimedia Appendix 4). The items in the questionnaire were not randomized. No form of adaptive questioning was used. All questions, except for the open-ended question, were mandatory; nonresponse options were not provided. Thus, there was no additional consistency or completeness check before submission. The survey took approximately 5 minutes to complete.

Statistical Analysis

Descriptive statistics were used to summarize the characteristics of the study participants, including age, gender, occupation, educational level, and internet use for health information, as well as self-perceived digital health literacy. Only completed questionnaires were analyzed. No statistical correction was applied to adjust the sample. To assess the acceptance and usability of the search engine, Likert-scale items related to these dimensions were combined to create scale values. The acceptance scale measured participants' willingness to use the search engine again and to recommend it to others. The usability scale evaluated participants' agreement with statements related to comprehensibility, clarity, the effectiveness of the search, confidentiality, advertising, the ease of learning, and using the search engine.

One reviewer analyzed the content of the open-ended question and assigned the comments to different categories inductively.

Multiple regression analyses were conducted to explore predictive factors influencing the acceptance of the search engine. The factors tested in the regression analyses included age, gender, occupation, educational level, the frequency of internet use for health information, and self-perceived digital health literacy.

Reporting

This study is reported based on the guidelines of the Checklist for Reporting Results of Internet E-Surveys (CHERRIES) [58].

User Behavior Tracking

Behavior tracking data were collected from all consenting users accessing the search engine between October 2020 and June 2021 using the web analytics software Matomo (Matomo.org) [59]. Matomo is a powerful open-source web analytics platform designed to help website owners and organizations gain insights into their online presence while respecting user privacy. Its primary purpose is to track website traffic by tracking website visitor behavior. What sets Matomo apart from services such as Google Analytics is its commitment to data privacy and General Data Protection Regulation (GDPR) compliance. Matomo allows users to maintain control over their data by hosting the data on their own servers, ensuring that sensitive information is not shared with third parties. Matomo is easy to integrate into web pages and can also be customized to track specific custom data. Ripp and Falke [60] successfully used Matomo to track user behavior in their research project, where they analyzed search behavior for certain keywords on the online information system Grammis.

We installed Matomo on our own server and integrated the tracking into our search engine's front end. We enabled the tracking to record the filter settings and capture the search term beyond the standard visit data, such as how long users spent on the website, and which of the results they clicked on. Structurally, Matomo records several actions for each visit, containing the duration among other data (Figure S1 in Multimedia Appendix 5). However, for the last action of a visit, the duration is not recorded. We have filled the empty value

XSL•FO
RenderX

with the median of all actions from the particular visit to obtain a more realistic value for the entire visit duration.

Only data from consenting users was recorded and subsequently descriptively analyzed for the 9-month time period. Therefore, this sample differs from the sample of the questionnaire study but may also include users who completed the questionnaire. The data processing for the data analysis was realized with a Python script directly accessing Matomo's database using specific SQL queries.

Ethical Considerations

The study was approved by the institutional review board at the University of Freiburg (Ethikkomission-Freiburg 559/17) as an extension to the GAP study [22], and a document on data protection was agreed upon. As an incentive, the questionnaire study participants could take part in a prize draw for 1 of 25 vouchers to a bookstore worth €20 (US $24.33) after completing the survey.

Results

Questionnaire Study

Participants

During the time the survey was open (from October 2020 to the end of May 2021), there were 1577 unique site visitors to the search engine website, with 1426 (90.4%) visitors to the first survey page. Of these 1426 visitors, 1250 (87.7%) agreed to participate in the survey. Of these 1250 participants, 1123 (89.8%) completed the survey.

Of the 1123 survey respondents, 802 (71.4%) were from the panel described in the Recruitment subsection. The age distribution of these participants was approximately normal, with a little more than a quarter of the participants (215/802, 26.8%) falling into the age category of 50-59 years. The gender distribution was also fairly balanced between male and female, with a little more than half of the respondents (432/802, 53.9%) self-identifying as female. Regarding occupational background, the majority of the participants (726/802, 90.5%) did not work in a medical profession.

Educational levels were notably higher than in the general population due to recruitment through the scientific panel, with 69.3% (556/802) of the participants having completed upper secondary education (at least level 3 in the International Standard Classification of Education [ISCED]-2011) and 46.3% (372/802) holding a bachelor's or master's degree or higher (ISCED-2011 levels 5-8).

The study participants recruited from the web-based panel reported searching for health information online more frequently than the general population, with 65.4% (524/802) searching several times a month at the minimum and 27.7% (222/802) searching at least several times a week. Only 1.5% (12/802) of the participants indicated that they "never" searched the internet for health information.

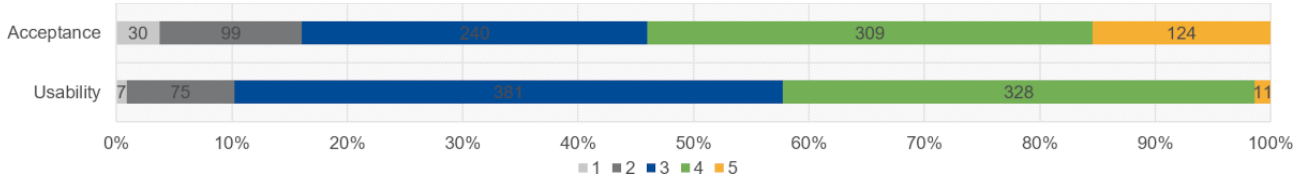
Participants' self-perceived digital health literacy was measured by the G-eHEALS items. Of the 802 participants, 291 (36.2%) felt confident in making health-related decisions based on internet information, while 461 (57.5%) felt capable of distinguishing reliable from questionable information online.

More information about the participants is available in Table S3 in Multimedia Appendix 4.

Acceptance and Usability

Scale values were created by combining items relating to the acceptability and usability of the search engine (Figure 3).

Figure 3. Scale values for acceptance and usability (n=802). The Likert-scale response options were as follows: 1=strongly disagree, 2=disagree, 3=partially agree, 4=agree, and 5=strongly agree. There was one negative question, which was adjusted for calculating the scale values.



The results show agreement regarding the acceptance and usability measures of the new search engine (Figure 3 and Table 1). In particular, more than half of the participants expressed their willingness to use the search engine again (465/802, 58%) and to recommend it to others (507/802, 63.2%). In addition, 79.3% (636/802) of the respondents agreed or strongly agreed that using the search engine was quick to learn, and 77.3% (620/802) to 98.6% (791/802) agreed at least partially with other statements related to usability aspects.

XSL•FO
RenderX

Table 1. Detailed items for acceptance, usability, and innovative aspects (n=802). The Likert-scale response options for acceptance and usability were as follows: 1=strongly disagree, 2=disagree, 3=partially agree, 4=agree, and 5=strongly agree. The Likert-scale response options for the innovative aspects item were as follows: 1=very unimportant, 2=unimportant, 3=neither unimportant nor important, 4=important, and 5=very important.

Content	Scores, mean (SD)
Acceptance	
Item 1: potential reuse	3.58 (1.02)
Item 2: recommendation to others	3.68 (1.03)
Usability	
Item 1: quick to learn	4.03 (0.86)
Item 2: efficient scannability of the SERP[a]	3.63 (1.00)
Item 3: clarity of the SERP	3.75 (1.00)
Item 4: functions *not* comprehensible[b]	2.37 (1.29)
Item 5: appreciation for the absence of advertising	4.66 (0.66)
Item 6: functionality meets expectations	3.45 (1.03)
Item 7: careful handling of personal data	3.59 (0.88)
Item 8: no commercial bias	3.55 (0.95)
Item 9: fast access to relevant information	3.77 (0.99)
Item 10: innovative approach	3.58 (1.03)
Innovative aspects of the search engine (anonymous searches)	4.50 (0.76)

[a]SERP: search engine results page.

[b]Item 4 (comprehensibility of functions) was a negated question.

Users particularly appreciated the absence of advertising on the platform, with 93.1% (747/802) agreeing or strongly agreeing that the absence of advertising was a significant benefit. The search engine's user-friendly interface and the ability to adjust search results through the use of filters were also well received by participants. Furthermore, the unique features of the search engine, notably the anonymous searches without the creation of user profiles and independence from sponsors, received high importance ratings from 91.3% (733/802) of those surveyed.

More detailed results are available in Tables S4-S6 in Multimedia Appendix 4.

Of the 802 participants, 561 (70%) to 622 (77.6%) indicated that they used the filter functions that could be enabled to change what was displayed on the SERP (Table 2; Figure 2). Furthermore, 92.3% (518/561), 93.9% (553/589), 94.7% (567/599), and 96.5% (600/622) of those users who used the filters agreed at least partially that the filter functions were helpful in finding trustworthy, recent, user-friendly, and comprehensible results, respectively. The *recency* filter was particularly well received, with 46% (286/622) of the users strongly agreeing that it was helpful. More detailed results are available in Table S7 in Multimedia Appendix 4.

Table 2. Helpfulness of filters (n=802). The Likert-scale response options were as follows: 1=strongly disagree, 2=disagree, 3=partially agree, 4=agree, and 5=strongly agree.

Content	Participants, n (%)[a]	Values, mean (SD)
Helpfulness of filters for each filter		
Item 1: trustworthiness	599 (74.7)	4.11 (0.92)
Item 2: recency	622 (77.6)	4.25 (0.84)
Item 3: comprehensibility	589 (73.4)	4.05 (0.93)
Item 4: user-friendliness	561 (70)	3.91 (0.93)

[a]All other participants answered, "I did not use this function, I cannot evaluate this question."

In the open-ended question, which aimed to supplement the data with qualitative results, participants were asked to share their comments and suggestions for improvements. Of the 802 participants, 263 (32.8%) replied to the question. Multiple aspects were mentioned. In an inductive approach, 10 main themes were identified (Figure 4). Many of the participants commented on the usability aspects of the website. Of the 236 comments, 76 (32.2%) included constructive criticism and specific suggestions for improving the website in the future. Many commenters made suggestions for improving the design because they found some aspects unclear or overwhelming. Others did not fully understand the methodology behind the

XSL•FO
RenderX

evaluation of the selected providers or the reason why only specific providers, including commercial web pages, were shown on the results page and asked for more clarification. Some commenters were happy with the site as it stood and did not ask for any changes, while some conveyed disappointment with the results they discovered on the SERP.

Figure 4. In the analysis of 263 comments in response to the open-ended question, 10 main themes were identified inductively: design; search results and filter functions; selection of providers; technical details; method of evaluating the website; target audience; text on the website; comments about the survey; general positive responses; and critique, dissatisfaction, and other comments. The illustration shows a hierarchical visualization, with the size of the elements indicating the quantity of comments.



Predictors of Acceptance

We conducted multiple regression analyses to examine the predictive factors influencing responses to the acceptance scale. Age, gender, occupation, educational level, frequency of internet use for health information, and self-perceived digital health literacy were tested as influencing factors. All factors were first tested individually and subsequently tested simultaneously due to the potential error introduced by multiple testing. Overall, only minor discrepancies were identified. Significance levels were set at $P<.05$. Table 3 presents more details.

XSL•FO
RenderX

Table 3. Predictive factors for acceptance (n=802). Categorical variables were dummy coded.

Predictive factors for acceptance	Individual testing of each predictor				Simultaneous testing of all predictors[a]			
	B (SE)	*F* test (*df*)	*P* value	R^2 (%)	B (SE)	*F* test (*df*)	*P* value	η_p^2 (%)
Age	*0.07*[b] *(0.02)*	*7.97 (1)*	*.005*	*0.99*	*0.07 (0.03)*	*7.27 (1)*	*.007*	*0.91*
Gender		0.04 (2)	.96	0.01		0.20 (2)	.82	0
Diverse	0.01 (0.49)	0.00 (1)	.99		−0.07 (0.49)	0.02 (1)	.89	
Female	0.02 (0.07)	0.09 (1)	.77		0.04 (0.07)	0.35 (1)	.55	
Male	Reference	Reference	Reference	Reference	Reference	Reference	Reference	Reference
Occupation		1.82 (2)	.16	0.45		1.14 (2)	.32	0.29
Medical doctor	0.12 (0.12)	0.92 (1)	.34		0.07 (0.12)	0.28 (1)	.59	
Other medical profession	−0.56 (0.35)	2.62 (1)	.11		−0.49 (0.35)	1.93 (1)	.17	
Not working in a medical profession	Reference	Reference	Reference	Reference	Reference	Reference	Reference	Reference
Highest level of educational attainment		*2.57 (4)*	*.04*	*1.27*		1.55 (4)	.19	0.78
No formal secondary education (ISCED[c] levels 0-1) or lower secondary education (*Hauptschulabschluss*, ISCED-2011 level 2)	*0.51 (0.19)*	*6.81 (1)*	.009		*0.41 (0.20)*	*4.22 (1)*	.04	
Lower secondary education (*Realschulabschluss*, ISCED-2011 level 2)	0.33 (0.17)	3.85 (1)	.05		0.26 (0.17)	2.33 (1)	.13	
Upper secondary education (*Fachabitur oder Abitur*, ISCED-2011 levels 3-4)	0.16 (0.17)	0.94 (1)	.33		0.13 (0.17)	0.58 (1)	.45	
Tertiary education: bachelor's or master's degree or equivalent (ISCED-2011 levels 6-7)	0.21 (0.16)	1.67 (1)	.20		0.17 (0.16)	1.14 (1)	.29	
Doctorate degree (ISCED-2011 level 8)	Reference	Reference	Reference	Reference	Reference	Reference		Reference
Frequency of internet use for health information	*0.15 (0.04)*	*17.16 (1)*	*< .001*	*2.1*	*0.13 (0.04)*	*11.98 (1)*	*<.001*	*1.5*
Self-perceived digital health literacy: information search	*0.12 (0.04)*	*9.69 (1)*	*.002*	*1.2*	0.08 (0.05)	3.03 (1)	.08	0.38
Self-perceived digital health literacy: information assessment	0.05 (0.04)	1.79 (1)	.18	0.22	−0.01 (0.05)	0.04 (1)	.85	0

[a]Variance resolution of simultaneous testing: R^2=5.02%.

[b]Factors with significant predictions (*P*<.05) are shown in italics.

[c]ISCED: International Standard Classification of Education.

There were 2 significant results in predicting the acceptance of the search engine. First, the acceptance rating increased with age, rising by 0.07 scale points (B=0.07) for each additional 10-year age group. Second, the ratings for acceptance were higher among participants with more frequent internet use for health information.

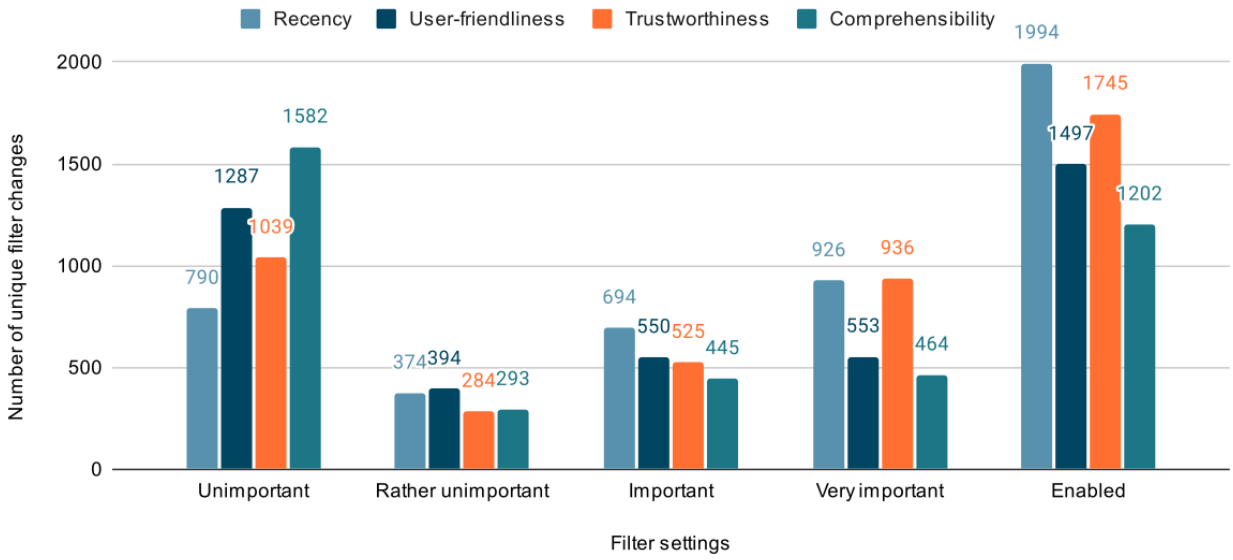
The factors "highest level of educational attainment" and "self-perceived digital health literacy" only showed significant correlations with the acceptance ratings in the individual tests and slightly lower correlations in the simultaneous tests. Participants with lower levels of education showed stronger acceptance ratings than those with higher levels of education. For this predictor, the reference category was set as "doctorate degree" (ISCED-2011 level 8). Acceptance ratings among persons without any formal secondary education (ISCED-2011 levels 0-1) or with a lower secondary education (*Hauptschulabschluss*, ISCED-2011 level 2) were 0.41 scale points higher than among persons with a doctorate degree (B=0.41). However, if we examine the effect strength using partial eta–squared (η_p^2), only small effects were shown overall for all factors (1%=small effect, 6%=medium effect). No significant results were found for the factors *gender* or *occupation*.

XSL•FO
RenderX

User Behavior Tracking

For a period of 9 months (from October 2020 to the end of June 2021), user behavior of consenting users was tracked using Matomo. The main aspects recorded were the search queries and the use of the filter functions. The data indicated that 1631 visitor sessions had taken place, with 3090 searches using 1984 different search terms across 1924 visits. The search terms had an average lexical token count of 1.74 (SD 1.14). Of the 1984 search terms, 1096 (55.24%) had a lexical token count of 1, while 888 (44.76%) had a lexical token count of >1. The maximum lexical token count was 13. We observed that 28.9% (893/3090) of the search queries were conducted through mobile devices such as smartphones and tablets. On average, visitors conducted 1.64 (SD 1.31) searches per visit, with an average duration of 137.45 (SD 278) seconds and a median duration of 49 (IQR 24-119) seconds. The top 3 search terms were "corona," "rückenschmerzen" (back pain), and "husten" (cough; Table S1 in Multimedia Appendix 5). Each search term was queried 1.57 (SD 3.46) times on average; of the 1984 search terms, 1621 (81.7%) were entered just once. Users could personalize the order of the displayed search results, depending on the prioritization of the different filter categories (*trustworthiness*, *recency*, *user-friendliness*, and *comprehensibility*), all of which were initially set to "unimportant" by default. Within the search interactions, overall, 47,532 actions were performed, among which were 1358 (2.86%) outlink actions (ie, external domains clicked by website visitors) and 34,490 (72.56%) filter changes in total. The high number of filter changes was due to the fact that the filters were often switched back and forth in a search, which led to multiple occurrences of the same filter settings. This behavior was probably observable due to the fact that, given the limited corpus, the result set could be empty for certain search terms in combination with a narrow filter setup; therefore, it was necessary to change back the filter to obtain results. Excluding these repeated changes, of the 34,490 filter changes, 5534 (16.05%) were unique. In addition, after removing the initial filter setup where all categories were set to the default value, we ended up with 2784 nondefault unique filter changes used in 15.37% (475/3090) of the searches. *Recency* was used the most, followed by *trustworthiness*, *user-friendliness*, and *comprehensibility* (Figure 5). Consistent with the results on the usefulness of the filter function from the questionnaire study, the user behavior data showed that the filters for *recency* and *trustworthiness* were adjusted most frequently.

Figure 5. Distribution of the filter settings of the 2784 nondefault unique filter changes across filter categories. We consider a filter category to be "enabled" if a filter category is not set to "unimportant.".



Discussion

Principal Findings

In this study, we evaluated the independent, noncommercial tala-med search for German-language medical information, placing a primary focus on transparency, independence, and evidence-based information. Our questionnaire study and user behavior tracking provide insights into the acceptance and usability of this innovative platform.

The survey resulted in a mean acceptance score of 3.63 (SD 0.98) and a mean usability score of 3.76 (SD 0.61) on a scale of 1 to 5 (1 signifying strong disagreement and 5 signifying strong agreement with acceptance and usability statements). Given that our user interface is similar to that of popular search engines such as Google and Bing and is advertisement-free, it is unsurprising that it was well received. However, there is still a need for optimization to achieve a higher reuse potential and a greater willingness to recommend the platform. Within the usability measures, item 1, which reflected the learnability of the search, achieved a high mean value of 4.03 (SD 0.86; Table 1). In terms of the effectiveness and clarity of our design (refer to items 2, 3, and 9 in Table 1), we still see room for improvement but are satisfied with the results for the initial release. Our findings align with those of Petrie and Bevan [40], who emphasized that usability, particularly learnability, is crucial for ensuring the platform's effectiveness and user satisfaction. Similar to the study by Petrie and Bevan [40], our study highlights the need to continually evaluate and refine

XSL•FO
RenderX

digital systems to cater to user needs and improve overall usability. The survey also revealed that the filters were used by >70% of the 802 users. Among these users, the filters were highly accepted on average, which matches with the tracking statistics revealing the high use of the filter functionality. The most used filter was *recency*, followed by *trustworthiness*, matching the survey outcome, which asked about the helpfulness of the filters. This was not the case for the 2 remaining filters. Although the filter for *user-friendliness* was used more often in our tracking data than the one for *comprehensibility*, it was perceived as less helpful than comprehensibility. In the future, we might inspect these filter categories in more detail and potentially improve their estimation. The insights support our decisions, including basing the design of our search engine on the insights of Strecker et al [32], which demonstrated how improvements in user interface design could lead to higher user satisfaction. The high acceptance of the filters for *recency* and *trustworthiness* in our study also echoes the findings of Strecker et al [32], who highlighted that content navigation and relevance contribute to overall user satisfaction. Nevertheless, the insights encourage us to improve and possibly expand the filters and the design. The fact that our search engine does not share any data with partners or advertisers and does not record the visitor's behavior to create profiles was rated as very relevant, with a mean score of 4.5 (SD 0.76). This result confirms the importance of our general endeavor driven by data privacy and noncommerciality as our fundamental motivations for creating the search engine. This finding is consistent with the work of Eysenbach and Köhler [43], who found that while internet users claimed to prioritize credibility when assessing health information online, they often relied on superficial factors such as website design and ease of use, rather than more reliable indicators such as the "About us" sections or disclaimers. Our platform addresses these shortcomings by offering more explicit filters, such as *trustworthiness*, which help users find credible information without having to rely solely on superficial design elements. In our calculations on the predictors of acceptance, it can be concluded that while the effects may be small, the search engine was particularly well accepted among older users, users with a high frequency of internet use for health information, and those with lower levels of education. This finding suggests that the new search engine may effectively cater to segments of the population with lower health literacy and digital health literacy and greater requirements for navigational assistance, bridging the gap in accessing reliable medical information online. As discussed by Kritz et al [42], general practitioners and medical specialists often face barriers in accessing high-quality medical content due to limited time and the overwhelming amount of information online. Our search engine's focus on evidence-based information and user-friendly design helps mitigate these issues, making it a suitable tool for both health care professionals and the general public.

This reinforces the necessity to consider navigational needs in health information platforms. Digital health literacy has improved in recent years, particularly among persons with low levels of education [14]. The ability to assess health information has also improved over time and over the course of the COVID-19 pandemic. Digital health literacy improved for younger but not for older people over the course of the pandemic [15]. Even with some of these improvements, health literacy is still low within the German population and particularly among people with low levels of education. This observation aligns with the findings by Zhang [44], who identified that source selection for health information is influenced by a wide range of factors, including accessibility, quality, usability, and personal relevance. Our platform's appeal to older users and those with lower health literacy supports the conclusion drawn by Zhang [44] that a personalized approach, considering different user-related factors, is crucial for effectively meeting the diverse needs of consumers when they access health information. The correlation between low educational levels, low social status, higher age, and lower health literacy rates has become even stronger [14]. This emphasizes the need to address low health literacy to improve social inequities. The new search engine attempts to address this need by assessing health information online and by making high-quality information easily searchable.

Many of the comments highlight the fact that at the time of the survey, participants used the first version of the search engine, which still had some technical and appearance-related weaknesses. Since the survey was conducted, a number of revisions have been made, some of which correspond to suggestions made in the comments, such as adding more information about the principles of the quality assessment and the functions of the websites and removing certain commercially funded information sources.

The platform's emphasis on anonymous searches and independence from advertising and sponsors resonated especially well with the participants, underscoring the importance of developing alternatives to commercial search platforms. The usability of our platform, free from external advertising influence, aligns with the findings by Belinda et al [41], who evaluated both internal factors (such as page load time and performance) and external factors (such as ease of navigation and organization of information) as critical to the overall user satisfaction with websites. Our platform similarly emphasizes ease of use and efficient performance, which were key contributors to its positive reception. The general results align with previous findings on the importance of user-friendly interfaces [61-63] and the impact of advertising on users' perceptions [64].

Limitations

Limitations of the Technology

Some participants expressed dissatisfaction with the displayed results. Several respondents found the design of the SERP and individual results confusing due to an overload of information in one place. Potential design changes have been suggested to improve clarity. Currently, some search results are duplicated on the results page. To enhance the appearance of the SERP, it is desirable to deduplicate the search index. This would mean displaying more relevant information and fewer distracting results. To address the fact that sometimes no results were displayed due to strict filter settings, we could display results with nonexact matches for the filter category score below the exact matches and visually indicate them.

XSL•FO
RenderX

Participants encountered difficulties in understanding how the evaluation process worked. Efforts should be made to make the evaluation process easier to understand when using the website. Some improvements have been made over time, such as providing brief explanations of the filter categories when hovering over the graphic showing the score for each domain.

In addition, further clarification is necessary regarding the domains chosen, which comprise both independent and commercial providers. Over time, certain websites were excluded from being displayed on the site because of the conflicts of interest of their publishers (such as the site "Zentrum der Gesundheit" and a site funded by a pharmaceutical company). Currently, the website no longer displays commercial providers.

Limitations of the Questionnaire Study

While the questionnaire study provides valuable insights into the acceptance and usability of the new search engine, several limitations should be acknowledged. First, the recruitment methodology, relying on web-based panel recruitment, unintentionally led to an overrepresentation of individuals with higher educational levels than the general population and potentially more frequent internet use due to the online recruitment process. Among our participants, 46.3% (371/802) had completed tertiary education compared to only 18.5% of the general German population with a university degree or equivalent [65]. Similarly, 65.4% (524/802) of our participants reported searching for health information only at least once a month, while a Bertelsmann Foundation 2018 study indicated that only approximately 50% of the general German population seek such information at least once a month [2]. This potentially influences acceptance and usability ratings. Participants' self-perceived digital health literacy, as measured by the G-eHEALS items, approximately aligned with that of the sample of the validation study of the German version of the questionnaire by Söllner et al [54]. However, this information should be viewed with caution because Kim et al [66] discovered a discrepancy between self-assessments and actual ability. Actual ability was not assessed in our study.

Second, as some of the participants only interacted with the search engine and its features for a brief duration, this short testing time proved to be a limiting factor, hindering the formation of comprehensive first impressions. The responses to the open-ended question suggest that some participants tested the features of the search engine for only a few seconds. This was caused by the pop-up window linking to the questionnaire opening after interacting with the website for 10 to 20 seconds. Such a brief interaction period was likely insufficient for users to thoroughly explore and understand the functionalities and benefits of the search engine. Consequently, the feedback provided by users on the acceptance and usability of the platform may not accurately reflect its true potential. The limited duration of testing may lead to superficial evaluations, whereby participants may base their judgments on initial impressions rather than informed use. This can result in an inaccurate assessment and therefore impact the results presented for the acceptance and usability measures. More extended testing periods could offer a more accurate assessment of the platform's acceptance and usability over time. Furthermore, the implementation of controls for the length of the testing period could provide more reliable data, thereby ensuring that all participants have sufficient time to engage with the platform before completing the questionnaire.

Limitations of the User Behavior Tracking

As Matomo does not record the last action, we decided to substitute this missing value with the median of all actions from the particular visit. However, a more optimal solution would be to record of the last action, which could be triggered when closing the browser, but Matomo does not support this feature.

Unfortunately, due to data protection regulations, we were unable to link the search behavior with the survey data. It would have been interesting to see to what extent the search behavior and the questionnaire would have provided more information.

Future Work

Further Evaluation of the New Search Engine

As these results are based on the first impressions of a group of users, data on long-term use and satisfaction have yet to be compiled. Future work should therefore address the need to evaluate long-term acceptance and usability measures of the site using a longitudinal study design. This could provide valuable insight into whether users actually use the search engine when they have an acute need to seek out medical information and whether they return to the site over time.

Future research should also consider comparing the search engine with other existing and established platforms to gain a deeper understanding of its unique advantages. Such investigations might involve participants executing identical search queries on both the novel search engine and other platforms and sharing their impressions of each experience, similar to the study by Strecker et al [32]. Particular attention should be given to participants' abilities to assess the quality and reliability of the information found. Furthermore, a comparison of the specific results retrieved from both search engines could be undertaken, examining both the congruence and discrepancies in the results.

Further Improvements to the Search Engine

Ongoing revisions and updates to the search engine, including integration on partner websites, hold promise in expanding its reach and use, ensuring that it remains relevant and effective in addressing users' health information needs. Since the data collection for this study was completed, there have been several revisions to the website.

As the qualitative survey showed, some users found the SERP complicated. To address this issue, a potential solution could be a clearly linked help page with a video explanation. Another solution could be to provide an animated assistant that guides users through the page.

One major disadvantage of the search engine in its current state stems from the fact that the evaluation of the domains—more than 50 of them—according to the quality criteria was performed in 2020, and there has not been any update to the evaluations thus far. The domains' quality may have improved or declined

XSL•FO
RenderX

in the time since the evaluation. A continual process with constant reevaluations of the domains is necessary to ensure an up-to-date rating of the domains accessible through the search engine. As we currently do not have the capacity to perform another evaluation or to establish a continual evaluation process for the domains, we have temporarily disabled the display of the filter category scores and the option to adjust the results based on these scores because the scores do not reflect the actual quality of the domains in their current state.

Thus, in future, inspired by Zowalla et al [34], we intend to evaluate various criteria automatically. In the study by Zowalla et al [34], the readability of the language was analyzed supported by a support vector machine. We plan to use more sophisticated approaches that also model contextual context. Therefore, we plan to create classifiers based on Bidirectional Encoder Representations from Transformers (BERT) [67]. We have in mind the domain-specific models GerMedBert [68] and BioGottBERT [69], the latter being a model of the German monolingual Robustly Optimized BERT Pretraining Approach (RoBERTa) [70] model GottBERT [71] specialized on medical language. A further improvement might be the enrichment of the texts with MeSH terms, similar to PubMed. Such approaches would enable us to process and rate documents individually, rather than evaluating domains using random samples.

From a technological standpoint, there are still notable areas for enhancement of the middleware, which could be implemented in the future. As App Search is proprietary and monolithic, to make it customizable and more extensible, we are currently exploring potential improvements. In particular, we are considering alternative technologies, as used by Scheible et al [31]. This would enable us to add features based on modern technologies. Specifically, we are contemplating a different approach to computing synonyms. Instead of using discrete structures, we plan to train and use a FastText model [72].

Conclusions

The development of the independent, noncommercial tala-med search marks a step toward improving access to reliable and evidence-based health information online. By prioritizing transparency, independence, and high-quality content, the platform has the potential to bridge the gap in health information accessibility, especially for older and less-educated individuals. This study achieved its objective of evaluating the acceptance and usability of the platform, using user questionnaires and behavior tracking, and the results have been encouraging. The search engine was well received, with users appreciating its advertising-free design and filtering system, which addresses the navigational needs of individuals with lower health and digital health literacy.

User feedback also highlighted areas for improvement, particularly in the clarity of the design and filter functionality. Nevertheless, the platform's emphasis on anonymous searches and independence from advertising resonated strongly with participants, reinforcing the importance of alternatives to commercial search engines. These findings are based on initial impressions, with long-term use and satisfaction data still to be collected. Future research should compare the search engine with existing platforms to further explore its unique advantages, while ongoing updates and potential integration on partner websites will enhance its reach and relevance.

In conclusion, the tala-med search demonstrates a promising step toward enhancing health literacy and empowering individuals to make well-informed health decisions. By addressing users' navigational needs and promoting equitable access to high-quality, evidence-based medical information, the search engine has the potential to positively impact health literacy, reduce health disparities, and promote patient empowerment in the digital age.

Acknowledgments

The study was funded by the Innovationsausschuss of the Gemeinsamer Bundesausschuss (01NVF17010) as part of the GAP (Gut informierte Kommunikation zwischen Arzt und Patient, meaning "Well-informed communication between physician and patient" in English) project. Furthermore, this work was supported by the German Ministry for Education and Research (01ZZ1801B, 01ZZ1804A, 01KX2121, and 01ZZ2304A). Finally, we acknowledge support by the Open Access Publication Fund of the University of Freiburg.

Klaus Koch of the Institute for Quality and Efficiency in Health Care supported the domain selection for the search index. Julia Serong and Jakob Henke of the TU Dortmund University contributed to the evaluation concept.

Authors' Contributions

RS, NK, and CW planned and carried out the implementation of the search engine supported by FT under the supervision of MB and SV-R. RS carried out the mathematical framework for the search engine and filter estimation with the help of LS. KW and AS-L worked out the criteria with which the domains were evaluated. MS created the procedure for the weighting of the filter categories and evaluated the pages. MS developed the questionnaire under the supervision of AM. CS prepared the data analyses for the predictors of acceptance under the supervision of EF-G. CS and RS prepared the data analyses of the web tracking. RS and LS wrote the manuscript with the help of all authors and under the supervision of AM. The manuscript has been read and approved by all named authors.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Search engine development: search equation and filter estimation.
[PDF File (Adobe PDF File), 830 KB - humanfactors_v12i1e56941_app1.pdf]

Multimedia Appendix 2
Search engine development: list of domains.
[PDF File (Adobe PDF File), 105 KB - humanfactors_v12i1e56941_app2.pdf]

Multimedia Appendix 3
Search engine development: criteria evaluated.
[XLSX File (Microsoft Excel File), 19 KB - humanfactors_v12i1e56941_app3.xlsx]

Multimedia Appendix 4
Questionnaire study: questionnaire and detailed results.
[PDF File (Adobe PDF File), 313 KB - humanfactors_v12i1e56941_app4.pdf]

Multimedia Appendix 5
User behavior tracking details.
[PDF File (Adobe PDF File), 156 KB - humanfactors_v12i1e56941_app5.pdf]

References

1. Baumann E, Czerwinski F, Rosset M, Seelig M, Suhr R. [How do people in Germany seek health information? Insights from the first wave of HINTS Germany]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2020 Sep;63(9):1151-1160. [doi: 10.1007/s00103-020-03192-x] [Medline: 32666180]
2. Haschke C, Westrick MG, Schwenk U. SPOTLIGHT gesundheit: gesundheitsinfos. Bertelsmann Stiftung. URL: https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/spotlight-gesundheit-gesundheitsinfos [accessed 2022-05-16]
3. Marstedt PG. Das internet: auch ihr ratgeber für gesundheitsfragen? Bertelsmann Stiftung. 2018 Jan 26. URL: https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/das-internet-auch-ihr-ratgeber-fuer-gesundheitsfragen [accessed 2022-05-18]
4. Individuals - internet activities. Eurostat. URL: https://ec.europa.eu/eurostat/databrowser/view/isoc_ci_ac_i/default/table?lang=en [accessed 2024-11-26]
5. Zhang Y, Sun Y, Xie B. Quality of health information for consumers on the web: a systematic review of indicators, criteria, tools, and evaluation results. J Assoc Inf Sci Technol 2015 Oct;66(10):2071-2084. [doi: 10.1002/asi.23311]
6. Eysenbach G, Powell J, Kuss O, Sa ER. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. JAMA 2002;287(20):2691-2700. [doi: 10.1001/jama.287.20.2691] [Medline: 12020305]
7. Hambrock U. Die suche nach gesundheitsinformationen. Bertelsmann Stiftung. 2018 Jan 26. URL: https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/die-suche-nach-gesundheitsinformationen [accessed 2022-05-18]
8. Baumann E, Czerwinski F. Erst mal doktor Google fragen? Nutzung neuer medien zur information und zum austausch über gesundheitsthemen. In: Böcken J, Braun B, Meierjürgen R, editors. Gesundheitsmonitor 2015: Bürgerorientierung im Gesundheitswesen. Gütersloh, Germany: Bertelsmann Stiftung; 2015:57-79.
9. Kammerer Y, Gerjets P. Effects of search interface and internet-specific epistemic beliefs on source evaluations during web search for medical information: an eye-tracking study. Behav Inf Technol 2012 Jan;31(1):83-97. [doi: 10.1080/0144929x.2011.599040]
10. Haas A, Unkel J. Ranking versus reputation: perception and effects of search result credibility. Behav Inf Technol 2017 Sep 26;36(12):1285-1298. [doi: 10.1080/0144929x.2017.1381166]
11. Cai HC, King LE, Dwyer JT. Using the Google™ search engine for health information: is there a problem? Case study: supplements for cancer. Curr Dev Nutr 2021 Feb;5(2):nzab002 [FREE Full text] [doi: 10.1093/cdn/nzab002] [Medline: 33937613]
12. Schaeffer D, Vogt D, Berens EM, Hurrelmann K. Gesundheitskompetenz der bevölkerung in Deutschland: ergebnisbericht. Universität Bielefeld, Fakultät für Gesundheitswissenschaften. 2017. URL: https://pub.uni-bielefeld.de/record/2908845#fileDetails [accessed 2024-12-20]
13. Schaeffer D, Berens EM, Vogt D. Health literacy in the German population. Dtsch Arztebl Int 2017 Jan 27;114(4):53-60 [FREE Full text] [doi: 10.3238/arztebl.2017.0053] [Medline: 28211318]
14. Schaeffer D, Berens EM, Vogt D, Gille S, Griese L, Klinger J, et al. Health literacy in Germany - findings of a representative follow-up survey. Dtsch Arztebl Int 2021 Oct 29;118(43):723-728 [FREE Full text] [doi: 10.3238/arztebl.m2021.0310] [Medline: 34551856]

XSL•FO
RenderX

15. Hurrelmann K, Klinger J, Schaeffer D. Gesundheitskompetenz der Bevölkerung in Deutschland: Vergleich der Erhebungen 2014 und 2020. Universität Bielefeld, Interdisziplinäres Zentrum für Gesundheitskompetenzforschung. 2020. URL: https://pub.uni-bielefeld.de/record/2950303 [accessed 2024-12-20]
16. Jordan S, Hoebel J. [Health literacy of adults in Germany: findings from the German Health Update (GEDA) study]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2015 Sep 31;58(9):942-950. [doi: 10.1007/s00103-015-2200-z] [Medline: 26227894]
17. Lee K, Hoti K, Hughes JD, Emmerton LM. Consumer use of "Dr Google": a survey on health information-seeking behaviors and navigational needs. J Med Internet Res 2015 Dec 29;17(12):e288 [FREE Full text] [doi: 10.2196/jmir.4345] [Medline: 26715363]
18. Wollmann K, van der Keylen P, Tomandl J, Meerpohl JJ, Sofroniou M, Maun A, et al. The information needs of internet users and their requirements for online health information-a scoping review of qualitative and quantitative studies. Patient Educ Couns 2021 Aug;104(8):1904-1932. [doi: 10.1016/j.pec.2021.01.020] [Medline: 33563502]
19. Röttele N, Schlett C, Körner M, Farin-Glattacker E, Schöpf-Lazzarino AC, Voigt-Radloff S, et al. Variance components of ratings of physician-patient communication: a generalizability theory analysis. PLoS One 2021;16(6):e0252968 [FREE Full text] [doi: 10.1371/journal.pone.0252968] [Medline: 34111197]
20. Radloff S, Wollmann K, Voigt-Radloff S, Farin-Glattacker E, Schlett C. Online-interventionen zur förderung der informiertheit. Impu!se für Gesundheitsförderung 2021;113(4):18.
21. van der Keylen P, Tomandl J, Wollmann K, Möhler R, Sofroniou M, Maun A, et al. The online health information needs of family physicians: systematic review of qualitative and quantitative studies. J Med Internet Res 2020 Dec 30;22(12):e18816 [FREE Full text] [doi: 10.2196/18816] [Medline: 33377874]
22. Voigt-Radloff S, Schöpf AC, Boeker M, Frank L, Farin E, Kaier K, et al. Well informed physician-patient communication in consultations on back pain - study protocol of the cluster randomized GAP trial. BMC Fam Pract 2019 Feb 25;20(1):33 [FREE Full text] [doi: 10.1186/s12875-019-0925-8] [Medline: 30803433]
23. van der Keylen P, Maun A, Hoyer S, Roos M, Beier M, Voigt-Radloff S, et al. "Mind the gap" — verzahnung von wissenschaft und hausärztlicher Praxis. Z Allg Med 2019;95:179-184 [FREE Full text] [doi: 10.3238/zfa.2019.0179-0184]
24. Braun C, Voigt-Radloff S. [Treatment of non-specific low back pain: evidence map of systematic reviews from 2015 to 2019]. Z Evid Fortbild Qual Gesundhwes 2019 Dec;149:12-21. [doi: 10.1016/j.zefq.2020.01.004] [Medline: 32151527]
25. Lang B, Kunzweiler K, Voigt-Radloff S. Gut informierte kommunikation zwischen arzt und patient im digitalen zeitalter – Das GAP-Projekt. In: Knieps F, Pfaff H, editors. Digitale Arbeit - Digitale Gesundheit. Berlin, Germany: Medizinisch Wissenschaftliche Verlagsgesellschaft; 2017.
26. Schlett C, Röttele N, van der Keylen P, Schöpf-Lazzarino AC, Klimmek M, Körner M, et al. The acceptance, usability, and utility of a web portal for back pain as recommended by primary care physicians: qualitative interview study with patients. JMIR Form Res 2022 Dec 29;6(12):e38748 [FREE Full text] [doi: 10.2196/38748] [Medline: 36580365]
27. Nater T, Boyer C. Debate about evaluation and monitoring of sites carrying the HON-Logo. J Med Internet Res 2000 Jun 30;2(2):E13 [FREE Full text] [doi: 10.2196/jmir.2.2.e13] [Medline: 11720932]
28. Scheible R, Kadioglu D, Ehl S, Blum M, Boeker M, Folz M, et al. Enabling external inquiries to an existing patient registry by using the open source registry system for rare diseases: demonstration of the system using the European Society for Immunodeficiencies Registry. JMIR Med Inform 2020 Oct 07;8(10):e17420 [FREE Full text] [doi: 10.2196/17420] [Medline: 33026355]
29. Murphy SN, Weber G, Mendis M, Gainer V, Chueh HC, Churchill S, et al. Serving the enterprise and beyond with informatics for integrating biology and the bedside (i2b2). J Am Med Inform Assoc 2010;17(2):124-130 [FREE Full text] [doi: 10.1136/jamia.2009.000893] [Medline: 20190053]
30. Scheible R, Thomczyk F, Blum M, Rautenberg M, Prunotto A, Yazijy S, et al. Integrating row level security in i2b2: segregation of medical records into data marts without data replication and synchronization. JAMIA Open 2023 Oct;6(3):ooad068 [FREE Full text] [doi: 10.1093/jamiaopen/ooad068] [Medline: 37583654]
31. Scheible R, Strecker P, Yazijy S, Thomczyk F, Talpa R, Puhl A, et al. A multilingual browser platform for medical subject headings. Stud Health Technol Inform 2022 Jan 14;289:384-387. [doi: 10.3233/SHTI210939] [Medline: 35062172]
32. Strecker P, Boeker M, Buechner S, Scheible R. Usability evaluation of a modern multilingual MeSH browser. Stud Health Technol Inform 2022 Jun 29;295:37-40. [doi: 10.3233/SHTI220653] [Medline: 35773799]
33. Zowalla R, Wetter T, Pfeifer D. Crawling the German Health Web: exploratory study and graph analysis. J Med Internet Res 2020 Jul 24;22(7):e17853 [FREE Full text] [doi: 10.2196/17853] [Medline: 32706701]
34. Zowalla R, Pfeifer D, Wetter T. Readability and topics of the German Health Web: exploratory study and text analysis. PLoS One 2023 Feb 10;18(2):e0281582 [FREE Full text] [doi: 10.1371/journal.pone.0281582] [Medline: 36763573]
35. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319-340. [doi: 10.2307/249008]
36. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425-478. [doi: 10.2307/30036540]
37. Nielsen J. Usability Engineering. Burlington, MA: Morgan Kaufmann Publishers; 1994.

XSL•FO
RenderX

38. Baumel A, Yom-Tov E. Predicting user adherence to behavioral eHealth interventions in the real world: examining which aspects of intervention design matter most. Transl Behav Med 2018 Sep 08;8(5):793-798. [doi: 10.1093/tbm/ibx037] [Medline: 29471424]
39. Mateus DA, Silva CA, de Oliveira AF, Costa H, Freire AP. A systematic mapping of accessibility problems encountered on websites and mobile apps: a comparison between automated tests, manual inspections and user evaluations. J Interact Syst 2021 Nov 08;12(1):145-171. [doi: 10.5753/jis.2021.1778]
40. Petrie H, Bevan N. The evaluation of accessibility, usability and user experience. University of York. 2009. URL: https://pure.york.ac.uk/portal/en/publications/the-evaluation-of-accessibility-usability-and-user-experience [accessed 2024-12-20]
41. Belinda BI, Kayode AB, Solomon N, Bethy TA. Analysis of internal and external website usability factors. Commun Appl Electron 2021 Apr;7(35):9-18. [doi: 10.5120/cae2021652880]
42. Kritz M, Gschwandtner M, Stefanov V, Hanbury A, Samwald M. Utilization and perceived problems of online medical resources and search tools among different groups of European physicians. J Med Internet Res 2013 Jun 26;15(6):e122 [FREE Full text] [doi: 10.2196/jmir.2436] [Medline: 23803299]
43. Eysenbach G, Köhler C. How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. BMJ 2002 Mar 09;324(7337):573-577 [FREE Full text] [doi: 10.1136/bmj.324.7337.573] [Medline: 11884321]
44. Zhang Y. Beyond quality and accessibility: source selection in consumer health information searching. J Assoc Inf Sci Technol 2014 Jan 07;65(5):911-927. [doi: 10.1002/asi.23023]
45. Saad M, Zia A, Raza M, Kundi M, Haleem M. A comprehensive analysis of healthcare websites usability features, testing techniques and issues. IEEE Access 2022;10:97701-97718. [doi: 10.1109/access.2022.3193378]
46. Fess - enterprise search server. Fess. URL: https://fess.codelibs.org/index.html [accessed 2021-04-07]
47. mozilla / readability. GitHub. URL: https://github.com/mozilla/readability [accessed 2021-04-07]
48. Tetrault D. Elastic app search: now available as a self-managed download. Elastic. 2019 Jun 25. URL: https://www.elastic.co/blog/elastic-app-search-7-2-0-released [accessed 2021-04-07]
49. Elasticsearch homepage. Elastic. URL: https://www.elastic.co/de/enterprise-search [accessed 2023-07-17]
50. Hashemian Nik D, Kasáč Z, Goda Z, Semlitsch A, Schulz S. Building an experimental German user interface terminology linked to SNOMED CT. Stud Health Technol Inform 2019 Aug 21;264:153-157. [doi: 10.3233/SHTI190202] [Medline: 31437904]
51. tala-med Suche. tala-med Suche Evidenzbasierte Gesundheitsinformationen. URL: https://suche.tala-med.info/ [accessed 2024-12-23]
52. WisoPanel homepage. WisoPanel. URL: https://www.wisopanel.net/ [accessed 2023-12-21]
53. Umfragesoftware. Unipark. URL: https://www.unipark.com/umfragesoftware/ [accessed 2024-12-23]
54. Soellner R, Huber S, Reder M. The concept of eHealth literacy and its measurement. J Media Psychol 2014 Jan;26(1):29-38. [doi: 10.1027/1864-1105/a000104]
55. Brooke J. SUS: a 'quick and dirty' usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation in Industry. Boca Raton, FL: CRC Press; 1996.
56. Quirmbach SM. User experience und usability in suchmaschinen. In: Suchmaschinen: User Experience, Usability und Nutzerzentrierte Website-Gestaltung. Berlin, Heidelberg: Springer; 2012:57-65.
57. Magin M, Steiner M, Heinbach D, Bosold S, Pieper A, Felka EM, et al. Suchmaschinen auf dem Prüfstand – eine Vergleichende Inhaltsanalyse der Qualität von Trefferlisten. M&K 2015;63(4):495-516. [doi: 10.5771/1615-634x-2015-4-495]
58. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
59. Matomo homepage. Matomo. URL: https://matomo.org/ [accessed 2023-10-26]
60. Ripp S, Falke S. Analyzing user behavior with Matomo in the online information system Grammis. In: Proceedings of the XVIII EURALEX International Congress Lexicography in Global Contexts. 2018 Aug 29 Presented at: XVIII EURALEX; July 17-21, 2018; Ljubljana, Slovenia URL: https://ids-pub.bsz-bw.de/frontdoor/index/index/docId/7706
61. Shamat NA, Sulaiman S, Sinpang JS. A systematic literature review on user interface design for web applications. J Telecommun Electron Comput Eng 2017;9(3-4):57-61.
62. Mazumder FK, Das UK. Usability guidelines for usable user interface. Int J Res Eng Technol 2014 Sep;03(09):79-82. [doi: 10.15623/ijret.2014.0309011]
63. Nichols DM, McKay D, Twidale MB. Participatory usability: supporting proactive users. In: Proceedings of the 4th Annual Conference of the ACM Special Interest Group on Computer-Human Interaction. 2003 Presented at: CHINZ '03; July 3-4, 2003; Dunedin, New Zealand. [doi: 10.1145/2331829.2331841]
64. Davranova MD. Internet advertising: perceptions of the users. Int J Mark Bus Commun 2019;8(2 & 3):25-36 [FREE Full text]
65. Bildung in Deutschland (Bildungsstand). Destatis Statistisches Bundesamt. URL: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bildung-Forschung-Kultur/Bildungsstand/_inhalt.html [accessed 2023-12-08]
66. Kim S, Park K, Jo HS. Gap between perceived eHealth literacy and ability to use online cancer-related information. J Korean Med Sci 2020 Jun 22;35(24):e187 [FREE Full text] [doi: 10.3346/jkms.2020.35.e187] [Medline: 32567257]

XSL•FO
RenderX

67. Devlin J, Chang MW, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2019 Presented at: NAACL-HLT 2019; June 2-7, 2019; Minneapolis, MN. [doi: 10.18653/v1/N19-1423]
68. Bressem KK, Papaioannou JM, Grundmann P, Borchert F, Adams LC, Liu L, et al. medBERT.de: a comprehensive German BERT model for the medical domain. Expert Syst Appl 2024 Mar;237:121598. [doi: 10.1016/j.eswa.2023.121598]
69. Lentzen M, Madan S, Lage-Rupprecht V, Kühnel L, Fluck J, Jacobs M, et al. Critical assessment of transformer-based AI models for German clinical notes. JAMIA Open 2022 Dec;5(4):ooac087 [FREE Full text] [doi: 10.1093/jamiaopen/ooac087] [Medline: 36380848]
70. Liu Y, Ott M, Goyal N, Du J, Joshi M, Chen D, et al. RoBERTa: a robustly optimized BERT pretraining approach. arXiv. Preprint posted online on July 26, 2019 [FREE Full text]
71. Scheible R, Frei J, Thomczyk F, He H, Tippmann P, Knaus J, et al. GottBERT: a pure German language model. In: Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing. 2024 Presented at: EMNLP; November 12, 2024; Miami, FL URL: https://aclanthology.org/2024.emnlp-main.1183 [doi: 10.18653/v1/2024.emnlp-main.1183]
72. Bojanowski P, Grave E, Joulin A, Mikolov T. Enriching word vectors with subword information. Transact Assoc Comput Linguist 2017 Jun 1;5:135-146. [doi: 10.1162/tacl_a_00051]

Abbreviations

afgis: Aktionsforum Gesundheitsinformationssystem eV (Action Forum Health Information System eV)
BERT: Bidirectional Encoder Representations from Transformers
CHERRIES: Checklist for Reporting Results of Internet E-Surveys
GAP: Gut informierte Kommunikation zwischen Arzt und Patient (Well-informed communication between physician and patient)
GDPR: General Data Protection Regulation
G-eHEALS: eHealth Literacy Scale, German version
i2b2: Integrating Biology and the Bedside
ISCED: International Standard Classification of Education
MeSH: Medical Subject Headings
RoBERTa: Robustly Optimized Bidirectional Encoder Representations from Transformers Pretraining Approach
SERP: search engine results page
SNOMED-CT: Systematized Nomenclature of Medicine–Clinical Terms

Edited by E Borycki, Y Senathirajah; submitted 31.01.24; peer-reviewed by F Jahn, S Markham, L Hasan; comments to author 02.06.24; revised version received 12.07.24; accepted 23.10.24; published 23.01.25.

Please cite as:
Specht L, Scheible R, Boeker M, Farin-Glattacker E, Kampel N, Schmölz M, Schöpf-Lazzarino A, Schulz S, Schlett C, Thomczyk F, Voigt-Radloff S, Wegner C, Wollmann K, Maun A
Evaluating the Acceptance and Usability of an Independent, Noncommercial Search Engine for Medical Information: Cross-Sectional Questionnaire Study and User Behavior Tracking Analysis
JMIR Hum Factors 2025;12:e56941
URL: https://humanfactors.jmir.org/2025/1/e56941
doi:10.2196/56941
PMID:

©Lisa Specht, Raphael Scheible, Martin Boeker, Erik Farin-Glattacker, Nikolas Kampel, Marina Schmölz, Andrea Schöpf-Lazzarino, Stefan Schulz, Christian Schlett, Fabian Thomczyk, Sebastian Voigt-Radloff, Constanze Wegner, Katharina Wollmann, Andy Maun. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 23.01.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Telehealth Acceptance and Perceived Barriers Among Health Professionals: Pre-Post Evaluation of a Web-Based Telehealth Course

Lena Rettinger[1,2], BSc, MSc; Lukas Maul[1], BSc, MSc; Peter Putz[1], Dr rer nat; Veronika Ertelt-Bach[1], Mag., MAS; Andreas Huber[1], MSc; Susanne Maria Javorszky[1], BSc, MSc; Elisabeth Kupka-Klepsch[3], MSc; Sevan Sargis[1], BSc, MSc; Franz Werner[1], Dr. techn; Klaus Widhalm[1], MSc; Rosmarie Joseph[1], BSc; Stefanie Doci[1], BSc; Sebastian Kuhn[2], Prof Dr, MME

[1]Department Health Sciences, FH Campus Wien, University of Applied Sciences, Favoritenstrasse 232, Vienna, Austria

[2]Institute of Digital Medicine, Philipps-University and University Hospital of Giessen and Marburg, Marburg, Germany

[3]Department Applied Nursing Sciences, FH Campus Wien, University of Applied Sciences, Vienna, Austria

Corresponding Author:
Lena Rettinger, BSc, MSc
Department Health Sciences, FH Campus Wien, University of Applied Sciences, Favoritenstrasse 232, Vienna, Austria

Abstract

Background: The rapid expansion of telehealth underscores the need for comprehensive telehealth education among health care professionals. Despite increasing recognition of telehealth's importance, many practitioners remain underprepared, particularly in navigating legal aspects, technology, and patient engagement.

Objective: This study aimed to evaluate the impact of a web-based telehealth training course on health care professionals' telehealth acceptance and their perceived barriers to telehealth adoption.

Methods: An interventional study with a pre-post design was used in Austria. A total of 365 health professionals enrolled in an asynchronous web-based course covering general telehealth principles (concepts, legal and technical aspects, practical implementation) and profession-specific content (eg, nursing, speech therapy, and physiotherapy). Of these, 217 completed the course, and 185 met inclusion criteria for analysis. Participants' telehealth acceptance (covering telemetry, telephasis, and telepraxis) and perceived barriers were assessed via standardized questionnaires before and after the course. Satisfaction with the training was measured post-intervention using the Training Evaluation Inventory. Qualitative insights were gathered from open-ended survey questions and 2 focus groups, transcribed, and summarized.

Results: Post-intervention, overall telehealth acceptance increased significantly (P<.001, r=0.21), particularly for telemetry (remote assessment and monitoring), telepraxis (remote interventions), video call–based, and asynchronous telehealth. Perceived barriers to telehealth use—such as uncertainty about legal frameworks, data protection, and reduced quality of care—diminished significantly (P<.001, r=0.39). Post-intervention satisfaction was high, with a total median Training Evaluation Inventory score of 76 (IQR 13). Participants rated the course highly for its clarity, breadth of content, and inclusion of profession-specific modules. Qualitative feedback highlighted a desire for more hands-on demonstrations, interactive components, and guidance on institutional support and patient accessibility.

Conclusions: A structured, on-demand telehealth course significantly improved health professionals' awareness, acceptance, and intention to use telehealth and reduced perceived barriers. While the findings highlight that targeted web-based training can increase clinicians' confidence and readiness to use telehealth, it remains uncertain whether this will lead to an increase in its utilization. Future initiatives should incorporate blended-learning formats with additional practical examples, real-time discussions, and ongoing support to enhance long-term integration of telehealth into clinical workflows. On a policy level, we suggest coordinated actions at the EU, national, and institutional levels to standardize telehealth education and facilitate its practical implementation in everyday clinical practice.

(JMIR Hum Factors 2025;12:e74107) doi:10.2196/74107

KEYWORDS

telehealth; digital health; remote care; health care professionals; education; barriers; acceptance; web-based training

XSL•FO
RenderX

Introduction

Background

The rapid adoption of telehealth technologies has underscored the necessity for adequate training among health professionals. Telehealth education is increasingly recognized as essential in preparing health care providers to deliver remote care effectively [1]. However, inconsistencies in integrating telehealth into curricula have been identified, along with a notable lack of formal study designs assessing telehealth education [2].

A significant barrier to the adoption of telehealth among providers is the deficiency of specific education and training [2,3]. For instance, graduate nursing programs often lack adequate telehealth preparation, leaving future nurses unprepared for remote patient care [4]. Similarly, allied health curricula exhibit substantial educational deficits in telehealth training, indicating an urgent need for primary research to enhance the quality and consistency of telehealth education [5].

Students acknowledge the significance of telehealth in their future professions but report low levels of knowledge and experience with telehealth applications. This gap highlights the importance of equipping students with the necessary competencies in a rapidly evolving health care landscape [6]. Training initiatives in specific disciplines, such as occupational therapy and physical therapy, have shown potential in enhancing telehealth usage [7,8]. Those have been demonstrated to improve confidence and proficiency in telehealth among practitioners [7]. However, existing telehealth training programs often are either not specifically developed for certain professions or fail to address all the unique capabilities required for remote care delivery [8]. This limitation may hinder the effective adoption of telehealth practices.

Despite these challenges, outcomes from telehealth education and training studies demonstrate high student satisfaction, with many expressing a willingness to incorporate telehealth into their future practice. Improvements in students' problem-solving skills, acceptance of technology in patient care [2,9], improved telehealth knowledge and confidence [9,10], and enhanced telehealth and communication competencies [1,4,11] have also been noted. Moreover, it fosters a belief that they will use these skills in future practice [12]. Web-based learning tailored to learners' needs has the potential to change knowledge, attitudes, and willingness to incorporate telehealth into routine care [13].

Despite the recognized importance of structured telehealth education and the positive outcomes associated with such training, there remains a lack of formal study designs evaluating its impact on health care professionals' acceptance and perceived barriers. Moreover, there is a scarcity of educational offerings for already qualified health care professionals who did not receive telehealth education during their initial training. This gap leaves many practicing clinicians without the necessary skills and knowledge to adopt telehealth technologies effectively.

Aim

Therefore, this study aimed to evaluate the impact of a web-based telehealth training course on health care professionals' acceptance of telehealth and the perceived barriers to its adoption. By using a pre-post design, the study seeks to provide insights that can contribute to the improvement of telehealth education and promote the integration of telehealth into clinical practice.

Methods

Study Design

This study used a pre-post design to evaluate the impact of an interprofessional web-based telehealth training course on health care professionals' acceptance of telehealth and perceived barriers to its adoption.

Participants

Participants in this study were health care professionals. Inclusion criteria required participants to be physiotherapists, occupational therapists, orthoptists, nurses, speech and language therapists, or dietitians. Professionals from other health care fields were also eligible to participate in the course and study, but the course content was specified for the aforementioned professions.

Procedure

Participants were recruited through professional associations of the targeted professional groups. Interested individuals were invited to register via a web-based form. On May 1, 2024, a participation link was emailed to all registered participants, and they were able to participate in the study until August 31, 2024. As a measure to increase the response rate, 2 reminder emails were sent to registered participants on August 8 and August 26, 2024.

The web-based training as well as the pre- and post-intervention questionnaires were administered using the web-based survey platform "LimeSurvey." Participants completed the initial questionnaire prior to commencing the web-based course and filled out the follow-up questionnaire upon completion of the training. Participants were able to pause their participation at any time during the study.

Ethical Considerations

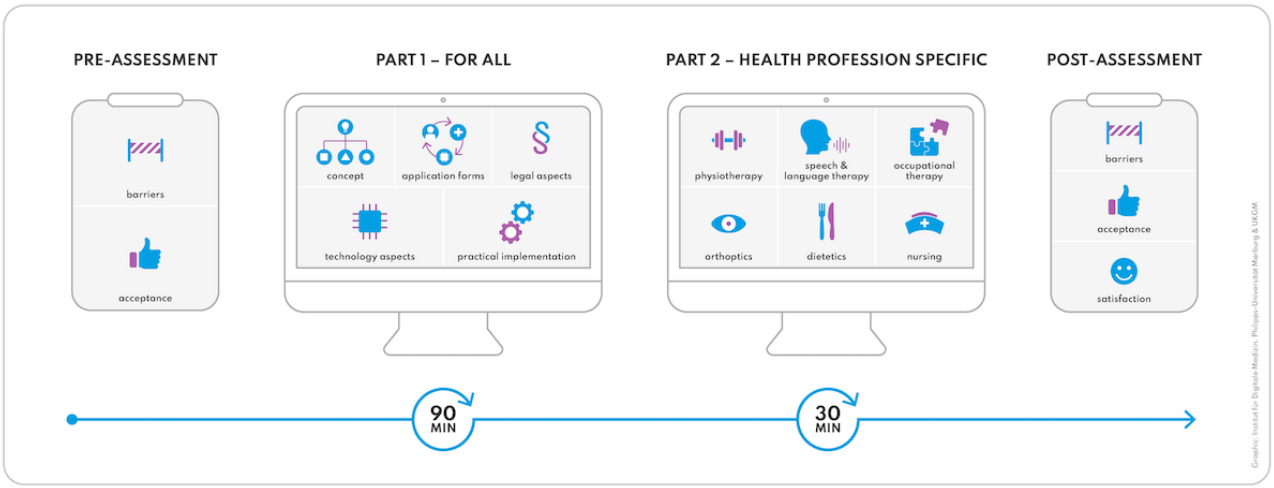
The survey followed ethical research practices as outlined in the Declaration of Helsinki, ie, voluntary participation; assurance of anonymity, data protection, and confidentiality; advance information about the purpose and content; provision of contact information; and disclosure of the organization conducting the research and the funding source. In accordance with the regulations of the Ethical Committee at FH Campus Wien, assessments involving non–health-related data from non-vulnerable persons do not require formal ethical approval. All participants received written information before participation. By completing the questionnaire, participants provided informed consent for the anonymous use of their data for research and publication purposes. Participation was entirely voluntary, and participants could pause or withdraw at any time without consequences. Participants received continuing education credits upon completion of the course.

XSL•FO
RenderX

Intervention: Telehealth Training Course

The web-based telehealth course was structured into a general and profession-specific video format (Figure 1). The general part of the course, which all participants were required to complete, consisted of the following videos: concept clarification (6:00 min), application forms (18:07 min), legal aspects (18:44 min), technology aspects (18:01 min), and practical implementation (30:20 min). Following the general part, participants were asked to watch a video tailored to their specific profession with a duration of approximately 30 minutes.

Figure 1. Structure of the assessments and the telehealth training course.



Participants from other professions were able to choose one of the specialized videos that best matched their area of practice. A detailed summary of the video content can be found in Multimedia Appendix 1.

Data Collection

Before watching the training videos, participants were asked to complete a demographic survey to collect data such as age, gender, profession, current employment status, years of experience, primary client group, federal state, population density, perceived patient access (based on health profession), and general and specific telehealth experience. Data on acceptance (willingness to adopt) and perceived influence of telehealth barriers were also collected both before and after the intervention using standardized scales. Finally, the participants rated their satisfaction with the course.

Telehealth Experience and Acceptance

We measured specific telehealth experiences and acceptance using a classification system based on the work of Colucci et al [14]. This system defines telehealth according to its communication processes, including telemetry, telephasis, and telepraxis (Table 1 and Figure 2). Participants self-reported whether they had experienced each telehealth application via phone, video calls, or asynchronously (yes or no). They also indicated their willingness to adopt each telehealth process through these modes, using a Likert scale from 0 (not at all willing) to 4 (fully willing).

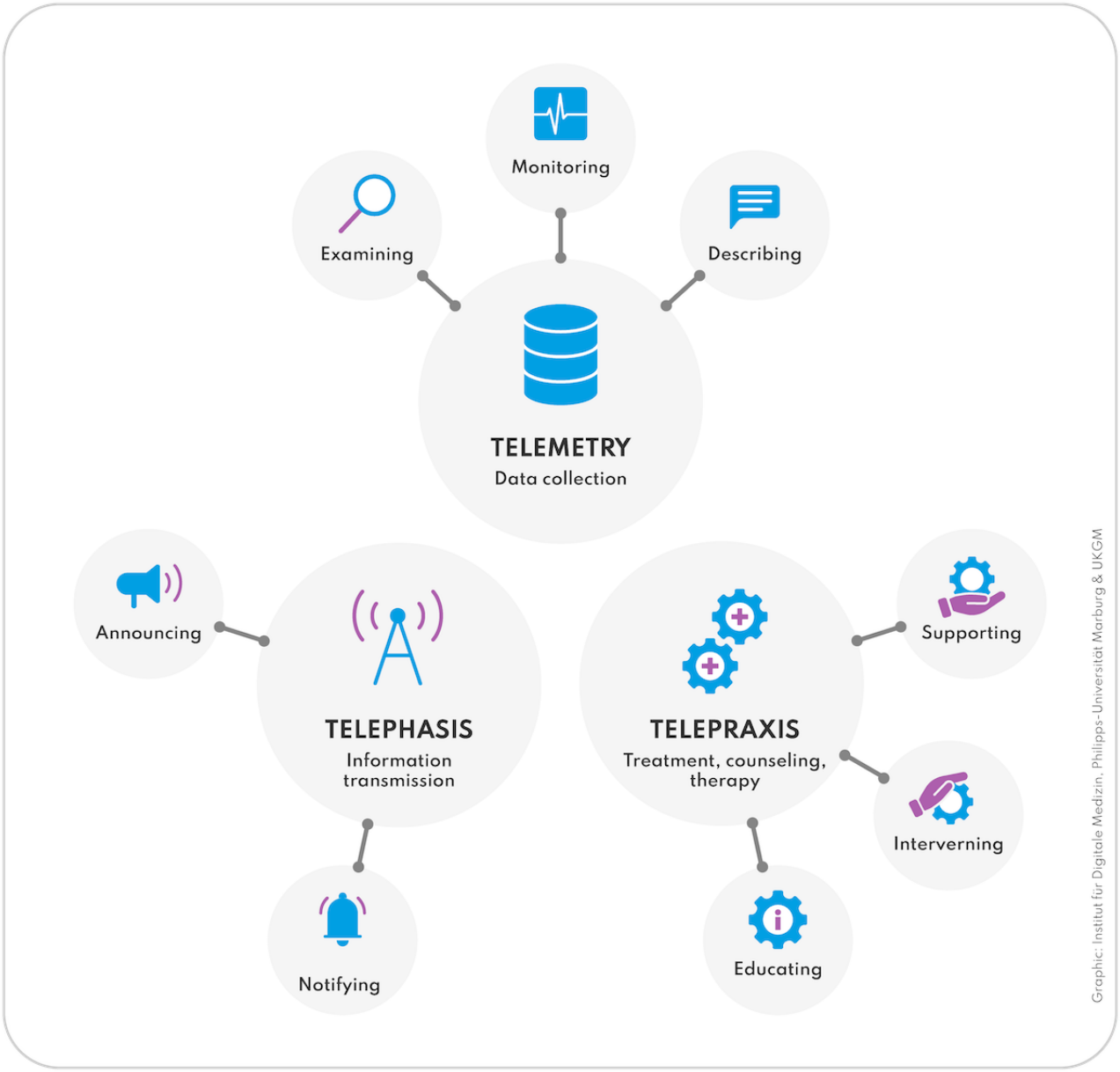
XSL•FO
RenderX

Table . Adapted classification system based on Colucci et al (2019) to assess specific telehealth experience and acceptance of the participants.

Function and application	Description or example
Telemetry	
Data collection	
Examining	Obtaining findings, performing measurements
Monitoring	Monitoring of vital data, movement, or activity behavior
Describing	Verbally or in writing, recording the client's medical history
Telephasis	
Information transmission	
Announcing	Public dissemination of health information, eg, on blogs, videos, websites
Notifying	Individual forwarding of information; eg, about findings, measurement results
Telepraxis	
Treatment, counseling, and therapy	
Supporting	eg, implementing therapy recommendations at home, providing environmental support
Intervening	eg, direct treatment or therapy
Educating	eg, health education, informing about lifestyle modifications

XSL•FO
RenderX

Figure 2. Visual representation of the classification system based on Colucci et al (2019) to assess specific telehealth experience.



Perceived Influence of Barriers

The barriers were measured based on a framework developed by Rettinger and Kuhn [3]. Participants rated the extent to which each of the 24 barriers influenced their willingness to adopt telehealth, using a Likert scale ranging from 0 (no influence) to 4 (strong influence).

Satisfaction With the Course

After the course, participants completed the Training Evaluation Inventory [15] to assess their satisfaction with the course content, subjective enjoyment, perceived usefulness, perceived difficulty, subjective knowledge gain, and attitude towards training. Additionally, they were asked for feedback in their own words. In September 2024, two focus groups were held with participants who volunteered to discuss the course in more detail.

Data Analysis

Participants who engaged with at least 40 minutes of the total 120-minute video content were included in the analysis; this threshold was established to ensure sufficient exposure to the intervention. Participants not meeting this criterion were excluded from subsequent analyses. Demographic data were summarized using means and standard deviations for continuous variables and frequencies and percentages for categorical variables.

Sum scores for telehealth acceptance were calculated: total telehealth acceptance (max 96), telemetry acceptance (max 48), telephasis acceptance (max 24), telepraxis acceptance (max 48), acceptance of telehealth via phone (max 32), acceptance of telehealth via video call (max 32), and acceptance of asynchronous telehealth (max 32). A total barrier score was calculated also with a sum-score (max 96). Acceptance of telehealth processes and perceived barriers, both before and after the intervention, were analyzed using medians and

XSL•FO
RenderX

interquartile ranges (IQRs) due to the ordinal nature of the data. The Wilcoxon signed-rank test was used to assess differences between pre- and post-intervention scores, with the significance level set at alpha=.05. This test evaluates differences in the distribution of repeated-measures data by calculating the sum of the ranks. This test can reveal significant differences between repeated measures, even when their medians are equal. Effect sizes were calculated using the r statistic, calculated as z/sqrt(n), and interpreted according to Hattie's guidelines [16,17], where $r<0$ indicates a negative effect, $r=0.00$ indicates no effect (developmental effects), $r=0.05$ indicates a small effect (teacher effects), $r=0.10$ indicates a medium effect (zone of desired effects), and $r \geq 0.15$ indicates a large effect, thereby providing context for the magnitude of the findings.

Satisfaction with the web-based course was evaluated post-intervention and presented as medians and IQRs. These satisfaction scores were further analyzed by professional groups to identify any differences across professions.

Qualitative data were obtained from 2 focus groups and written feedback collected through questionnaires. The focus group discussions were transcribed using MAXQDA AI-Assist. The transcribed and anonymized texts were then summarized using ChatGPT (version o1-preview). To ensure the validity of these summaries, the first 2 authors—who also moderated the focus groups—reviewed the summarized results for accuracy and completeness, adjusting where needed.

Sensitivity Analysis for Non-Representativeness of the Sample

When comparing the occupational and gender distribution of the sample to the population, male participants and participants from the nursing profession were found to be severely overrepresented. Therefore, an additional sensitivity analysis was conducted to take non-representativeness into account by case weighing. In this procedure, all data related to inferential statistics were post-harmonized on the basis of known population distributions, available from the annual report of the Austrian Register of Health Professions [18]. For this purpose, the following weighting factors (as shown in parentheses) were calculated for four strata: (1) female therapists (0.19), (2) male therapists (0.43), (3) female nurses (8.81), and (4) male nurses (3.48). Physiotherapists, occupational therapists, orthoptists, and speech and language therapists were thereby combined under the category "therapists." Weighting factors were calculated by dividing the target stratum proportion by the actual stratum proportion. The post-harmonized data can thus be considered representative of occupation (therapists vs nurses) and gender (female vs male). However, none of these post-harmonized inferential analyses (data not shown) differed from the unadjusted analyses in terms of statistical significance and effect interpretation, and therefore only the unadjusted results are reported in the results section.

Results

Participants

A total of 365 health professionals enrolled in the web-based course. Of these, 217 participants completed the course. Among the completers, 185 were included in the analysis: 63 physiotherapists (34%), 48 occupational therapists (26%), 38 orthoptists (21%), 20 nurses (11%), 9 speech and language therapists (5%), and 7 other health professionals (dietitian, diabetes counselor, midwife, psychologist, paramedic in training, student, and administrator). Mean engagement with the video content was 198 minutes (minimum 40; maximum 3015). In total, 32 course completers were excluded from analysis for not meeting the 40-minute minimum engagement threshold. The demographic characteristics of the participants are summarized in Table 2.

XSL•FO
RenderX

Table . Participants' characteristics.

Characteristics	PT[a] (n=63)	OT[b] (n=48)	ORT[c] (n=38)	NUR[d] (n=20)	SLT[e] (n=9)	Other (n=7)	Total (n=185)
Age, mean (SD), years	40.38 (10.57)	35.94 (9.58)	43.13 (11.48)	44.05 (8.63)	39.89 (9.21)	37.43 (10.01)	40.05 (10.52)
Professional experience in years, mean (SD)	16.02 (11.44)	11.73 (9.67)	19.84 (12.79)	20.25 (10.53)	15.33 (9.07)	15.98 (11.39)	15.98 (11.39)
Gender, n (%)							
Female	51 (80.95)	46 (95.83)	36 (94.74)	14 (70.00)	9 (100.00)	6 (85.71)	162 (87.57)
Male	12 (19.05)	2 (4.17)	2 (5.26)	6 (30.00)	0 (0.00)	1 (14.29)	23 (12.43)
Primary client group, n (%)							
Children and youth	7 (11.11)	9 (18.75)	21 (55.26)	1 (5.00)	5 (55.56)	0 (0.00)	43 (23.24)
Adults	35 (55.56)	31 (64.58)	11 (28.95)	14 (70.00)	4 (44.44)	5 (71.43)	100 (54.05)
Seniors	21 (33.33)	8 (16.67)	6 (15.79)	5 (25.00)	0 (0.00)	2 (28.57)	42 (22.70)
Federal state, n (%)							
Burgenland	1 (1.59)	1 (2.08)	4 (10.53)	1 (5.00)	1 (11.11)	0 (0.00)	8 (4.32)
Carinthia	3 (4.76)	5 (10.42)	1 (2.63)	1 (5.00)	0 (0.00)	0 (0.00)	10 (5.41)
Lower Austria	15 (23.81)	8 (16.67)	5 (13.16)	1 (5.00)	1 (11.11)	2 (28.57)	32 (17.30)
Upper Austria	5 (7.94)	3 (6.25)	1 (2.63)	4 (20.00)	1 (11.11)	1 (14.29)	15 (8.11)
Salzburg	2 (3.17)	0 (0.00)	3 (7.89)	1 (5.00)	0 (0.00)	1 (14.29)	7 (3.78)
Styria	4 (6.35)	3 (6.25)	2 (5.26)	4 (20.00)	1 (11.11)	0 (0.00)	14 (7.57)
Tyrol	4 (6.35)	3 (6.25)	0 (0.00)	1 (5.00)	2 (22.22)	0 (0.00)	10 (5.41)
Vorarlberg	1 (1.59)	1 (2.08)	2 (5.26)	0 (0.00)	0 (0.00)	0 (0.00)	4 (2.16)
Vienna	28 (44.44)	20 (41.67)	18 (47.37)	5 (25.00)	2 (22.22)	2 (28.57)	75 (40.54)
Other	0 (0.00)	4 (8.33)	2 (5.26)	2 (10.00)	1 (11.11)	1 (14.29)	10 (5.41)
Population density, n (%)							
Densely populated	38 (60.32)	26 (54.17)	27 (71.05)	9 (45.00)	5 (55.56)	2 (28.57)	107 (57.84)
Moderately populated	20 (31.75)	15 (31.25)	10 (26.32)	9 (45.00)	3 (33.33)	4 (57.14)	61 (32.97)
Sparsely populated	5 (7.94)	7 (14.58)	1 (2.63)	2 (10.00)	1 (11.11)	1 (14.29)	17 (9.19)
Perceived patient access (based on health profession), n (%)							
Poor	0 (0.00)	7 (14.58)	1 (2.63)	0 (0.00)	2 (22.22)	2 (28.57)	12 (6.49)
Moderate	16 (25.40)	21 (43.75)	13 (34.21)	6 (30.00)	4 (44.44)	2 (28.57)	62 (33.51)
Good	47 (74.60)	20 (41.67)	24 (63.16)	14 (70.00)	3 (33.33)	3 (42.86)	111 (60.00)
Telehealth experience, n (%)							
Never	46 (73.02)	32 (66.67)	31 (81.58)	16 (80.00)	2 (22.22)	4 (57.14)	131 (70.81)
In the past	11 (17.46)	3 (6.25)	2 (5.26)	2 (10.00)	1 (11.11)	1 (14.29)	20 (10.81)
Currently	6 (9.52)	13 (27.08)	5 (13.16)	2 (10.00)	6 (66.67)	2 (28.57)	34 (18.38)

XSL•FO
RenderX

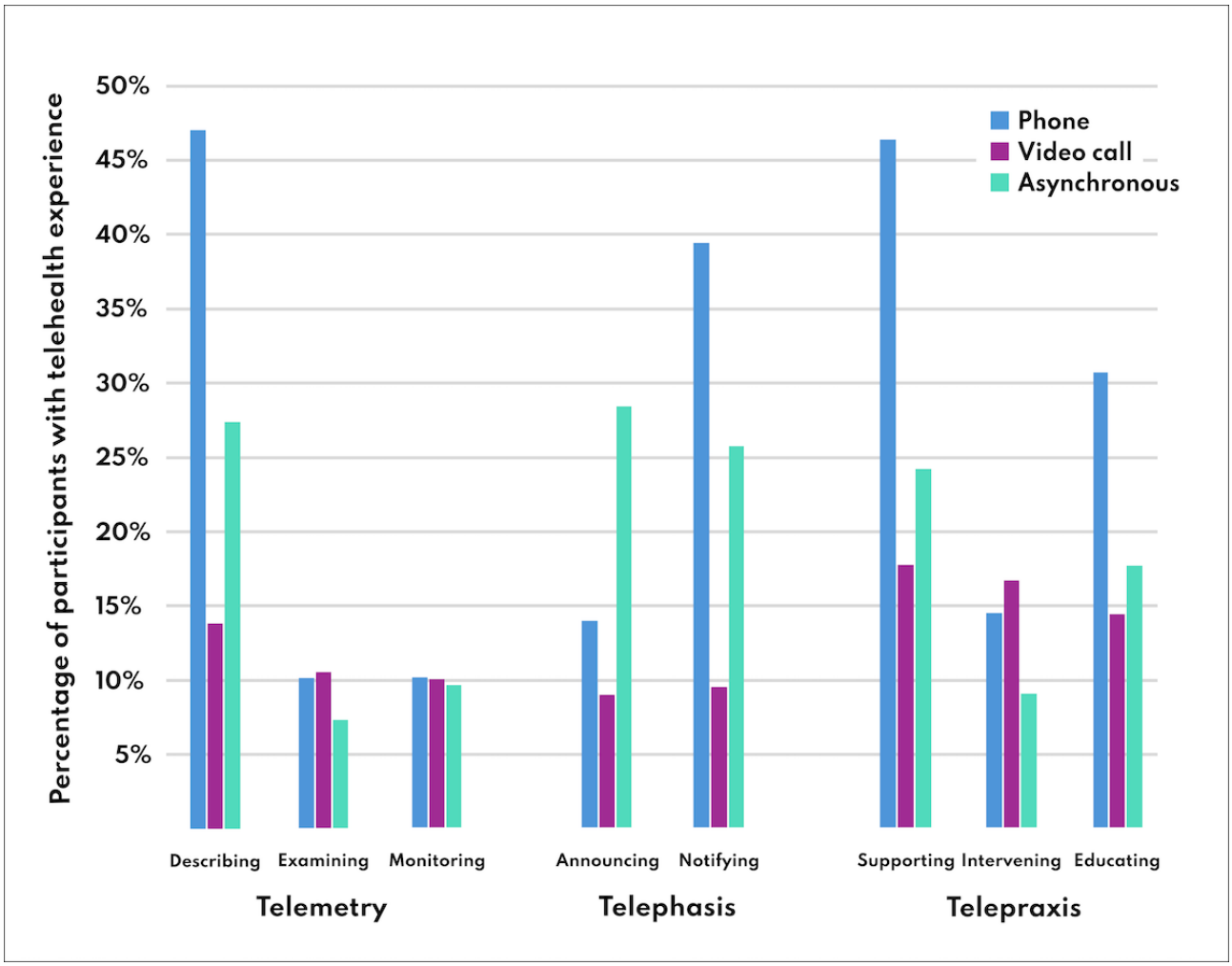
[a]PT: physiotherapists.
[b]OT: occupational therapists.
[c]ORT: orthoptists.
[d]NUR: nurses.
[e]SLT: speech and language therapists.

Telehealth Experience

In total, 71% (131/185) of the participants stated that they had never performed telehealth before, 11% (20/185) had telehealth experience but did not use it currently and 18% (34/185) currently use telehealth.

The participants had used phone or video calls or asynchronous communication for a variety of telehealth services. Telemetry was applied for describing by 95/185 (51%), for examining by 35/185 (19%), and for monitoring by 38/185 (21%); telephasis for announcing by 64/185 (35%), and for notifying by 91/185 (49%); telepraxis for supporting by 100/185 (54%), for intervening by 49/185 (26%), and for educating by 69/185 (37%). Further details on the usage of phone, video call, or asynchronous communication are listed in Figure 3.

Figure 3. Telehealth experience of the participants. Participants could select all functions they had used; thus, percentages exceed or fall short of 100%.



Telehealth Acceptance Pre and Post Web-Based Training

The total telehealth acceptance (P<.001), acceptance for telemetry (P<.001) and telepraxis (P<.001), acceptance for telehealth via video call (P<.001) and asynchronous telehealth (P=.007) changed significantly, while acceptance for telephasis (P=.203) and acceptance for phone call–based telehealth (P=.1) had no significant changes (Table 3 and Figure 4).

XSL•FO
RenderX

Table . Telehealth acceptance change pre- and post-intervention for telehealth functions, applications, and mediums (n=185).

Function and application	Pre-median acceptance (IQR)	Post-median acceptance (IQR)	*P* value	*r*	Int Hattie[b]
Acceptance total	61 (27)	63 (28)	<.001[a]	0.21	+++
Acceptance telemetry total	20 (13)	23 (12)	<.001[a]	0.20	+++
Describing					
Phone	3 (2)	3 (2)	.429	0.04	N/A
Videocall	3 (2)	3 (1)	<.001[a]	0.20	+++
Asynchronous	3 (3)	3 (2)	.584	0.03	N/A
Examining					
Phone	1 (2)	2 (2)	.018[a]	0.12	++
Videocall	2 (2)	3 (2)	<.001[a]	0.18	++
Asynchronous	2 (3)	2 (2)	.056	0.10	N/A
Monitoring					
Phone	2 (2)	2 (2)	.211	0.06	N/A
Videocall	3 (2)	3 (2)	.003[a]	0.16	++
Asynchronous	3 (3)	3 (2)	.022[a]	0.12	++
Acceptance telephasis total	16 (8)	17 (8)	.203	0.07	N/A
Announcing					
Phone	2 (3)	2 (3)	.352	0.05	N/A
Videocall	2 (3)	3 (2)	<.001[a]	0.20	+++
Asynchronous	3 (2)	3 (1)	.438	0.04	N/A
Notifying					
Phone	3 (2)	3 (2)	.785	0.01	N/A
Videocall	3 (2)	3 (2)	.042[a]	0.11	++
Asynchronous	3 (2)	3 (2)	.051	0.10	N/A
Acceptance telepraxis total	24 (11)	26 (11)	.003[a]	0.16	++
Supporting					
Phone	3 (2)	3 (2)	.885	0.01	N/A
Videocall	3 (2)	4 (1)	<.001[a]	0.18	++
Asynchronous	3 (2)	3 (2)	.487	0.04	N/A
Intervening					
Phone	2 (2)	2 (2)	.949	0.00	N/A
Videocall	3 (3)	3 (2)	.029[a]	0.11	++
Asynchronous	2 (2)	2 (2)	.001[a]	0.17	++
Educating					
Phone	3 (2)	3 (2)	.418	0.04	N/A
Videocall	3 (2)	3 (1)	.011[a]	0.13	++
Asynchronous	3 (2)	3 (1)	.123	0.08	N/A
Acceptance phone total	19 (11)	20 (11)	.053	0.10	N/A

XSL•FO
RenderX

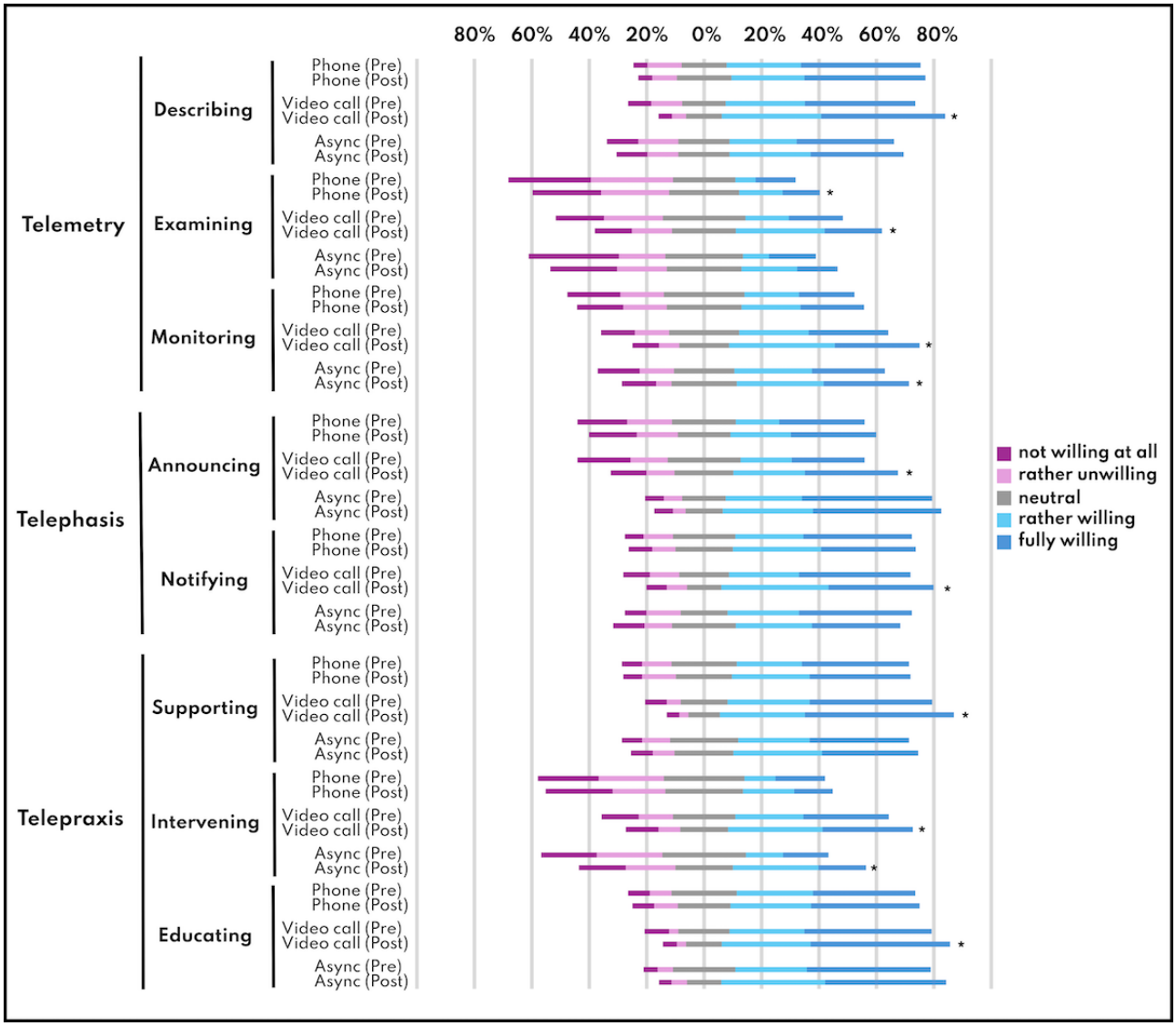
Function and application	Pre-median acceptance (IQR)	Post-median acceptance (IQR)	*P* value	*r*	Int Hattie[b]
Acceptance video call total	22 (10.5)	24 (9.5)	<.001[a]	0.26	+++
Acceptance asynchronous total	21 (11)	22 (10)	.007[a]	0.14	++

[a]Significant value, with $P < .05$.

[b]Effect sizes were calculated using the r statistic, calculated as z/sqrt(n), and interpreted according to Hattie's guidelines [16,17]. +=developmental effect (no effect). ++=teacher effect (small effect). +++=zone of desired effects (medium effect). Effect interpretation only for significant results.

[c]N/A: not applicable.

Figure 4. Telehealth acceptance change pre- and post-intervention. The asterisks indicate significant changes.



Influence of Telehealth Barriers Pre and Post Web-Based Training

The perceived influence of barriers on the willingness to adopt telehealth also was significantly reduced after the intervention ($P<.001$) (Table 4 and Figure 5).

XSL•FO
RenderX

Table . Change in the perceived influence of certain barriers on the willingness to adopt telehealth pre- and post-intervention.

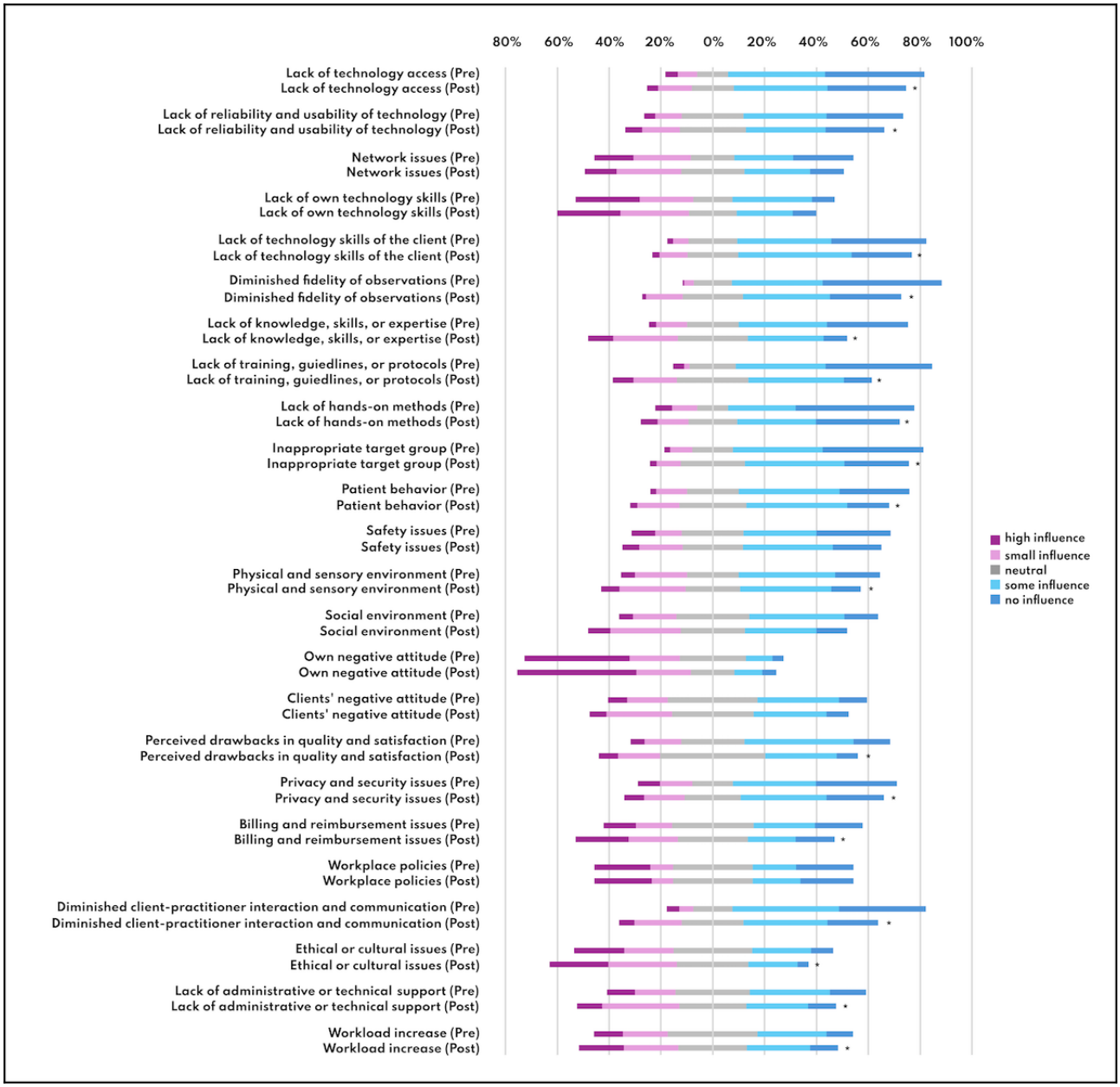
Barrier	Pre-MDN (IQR)	Post-MDN (IQR)	*P* value	*r*	Int Hattie[b]
Barriers total	61 (18)	53 (20.5)	<.001	0.39	+++
Lack of technology access	3 (1)	3 (2)	.006[a]	0.14	++
Lack of reliability and usability of technology	3 (2)	3 (1)	.003[a]	0.16	++
Network issues	2 (2)	2 (2)	.055	0.10	not applicable
Lack of own technology skills	2 (3)	1 (2)	.077	0.09	not applicable
Lack of technology skills of the client	3 (2)	3 (1)	<.001[a]	0.20	+++
Diminished fidelity of observations	3 (1)	3 (2)	<.001[a]	0.28	+++
Lack of knowledge, skills, or experience	3 (2)	2 (2)	<.001[a]	0.41	+++
Lack of training, guidelines, or protocols	3 (1)	2 (2)	<.001[a]	0.42	+++
Lack of hands-on methods	3 (2)	3 (2)	.002[a]	0.16	++
Inappropriate target group	3 (2)	3 (2)	.002[a]	0.16	++
Patient behavior	3 (2)	3 (1)	<.001[a]	0.19	++
Safety issues	3 (2)	3 (1)	.059	0.10	not applicable
Physical and sensory environment	3 (2)	2 (2)	.012[a]	0.13	++
Social environment	2 (1)	2 (2)	.294	0.16	not applicable
Own negative attitudes	1 (2)	1 (2)	.294	0.05	not applicable
Clients' negative attitude	2 (1)	2 (2)	.1	0.09	not applicable
Perceived drawbacks (quality and satisfaction)	3 (1)	2 (1)	<.001[a]	0.20	+++
Privacy and security issues	3 (2)	3 (1)	.042[a]	0.11	++
Billing and reimbursement issues	2 (2)	2 (2)	<.001[a]	0.17	++
Workplace Policies	2 (2)	2 (2)	.815	0.01	not applicable
Diminished client-practitioner interaction and communication	3 (2)	3 (1)	<.001[a]	0.30	+++
Ethical and cultural issues	2 (2)	2 (1)	.002[a]	0.16	++
Lack of administrative and technical support	2 (2)	2 (2)	.003[a]	0.16	++
Workload increase	2 (2)	2 (2)	.032[a]	0.11	++

[a]Significant value, with $P < .05$.

[b] Effect sizes were calculated using the r statistic, calculated as z/sqrt (n), and interpreted according to Hattie's guidelines [16,17]. +=developmental effect (no effect). ++=teacher effect (small effect). +++=zone of desired effects (medium effect). Effect interpretation only for significant results.

XSL•FO
RenderX

Figure 5. Perceived influence of certain barriers on the willingness to use telehealth pre and post intervention. Asterisk indicates significant changes.



Satisfaction With Web-Based Training

Overall satisfaction ratings were relatively high, with a total median of 76 (IQR 13) across all participants (Table 5).

XSL•FO
RenderX

Table . Median satisfaction scores (with IQR) of the Training Evaluation Inventory (TEI) for the web-based training.

Satisfaction	SLT[a] (n=9)	OT[b] (n=48)	PT[c] (n=63)	ORT[d] (n=38)	NUR[e] (n=20)	Other (n=7)	Total (n=185)
Total satisfaction	83 (7)	76 (13.75)	76 (11)	73 (14.25)	71.5 (18.5)	74 (28)	76 (13)
Subjective enjoyment							
Overall, I liked the course.	4 (1)	3 (1)	3 (1)	3 (1)	3 (1)	4 (1)	3.29
The learning atmosphere was agreeable.	4 (0)	4 (1)	4 (1)	4 (1)	3 (1)	4 (1)	3.43
The learning was fun.	4 (1)	3 (2)	3 (2)	3 (2)	3 (2)	3 (3)	2.88
Perceived usefulness							
I find the course useful for my job.	4 (1)	3 (1)	4 (1)	3 (2)	4 (1)	3 (1)	3.24
Investing time in this course was useful.	4 (1)	3 (1)	3 (1)	3.5 (1)	3 (1)	3 (1)	3.33
I can apply the content of this course in my job.	3 (1)	3 (1)	3 (1)	2 (1)	2.5 (1)	3 (1)	2.59
I derive personal use from this course.	3 (1)	3 (1)	3 (1)	3 (2)	3 (2)	3 (2)	2.98
Perceived difficulty							
The contents were comprehensible.	4 (0)	4 (1)	4 (0)	4 (0)	4 (1)	4 (0)	3.69
The language (foreign words and technical terms) was comprehensible.	4 (0)	4 (0)	4 (0)	4 (0)	4 (0)	4 (0)	3.77
I kept up thematically in the course.	4 (0)	4 (0)	4 (0)	4 (0)	4 (1)	4 (0)	3.76
The time was sufficient for the themes covered.	4 (0)	4 (1)	4 (0)	4 (0)	4 (1)	4 (2)	3.65
Subjective knowledge gain							
I have the impression that my knowledge has expanded on a long-term basis.	3 (1)	3 (1)	3 (1)	3 (1)	3 (1)	3 (2)	3.09
I will be able to remember the new themes well.	3 (1)	3 (1)	3 (2)	3 (2)	3 (1)	3 (1)	2.94

XSL•FO
RenderX

Satisfaction	SLT[a] (n=9)	OT[b] (n=48)	PT[c] (n=63)	ORT[d] (n=38)	NUR[e] (n=20)	Other (n=7)	Total (n=185)
I think that I will still be able to report what I learned some time after the course.	3 (1)	3 (2)	3 (1)	3 (0)	3 (1)	3 (2)	2.92
Attitude towards training							
I will apply what I learned to my day-to-day work.	3 (2)	2 (1)	2 (1)	2 (2)	2.5 (3)	2 (3)	2.24
I find it good that basics of telehealth were imparted and discussed.	4 (0)	4 (0)	4 (0)	4 (1)	4 (1)	4 (2)	3.7
I find it good that forms of application of telehealth were imparted and discussed.	4 (0)	4 (0)	4 (0)	4 (1)	4 (1)	4 (1)	3.72
I find it good that legal aspects of telehealth were imparted and discussed.	4 (0)	4 (0)	4 (0)	4 (1)	4 (1)	4 (1)	3.67
I find it good that technical aspects of telehealth were imparted and discussed.	4 (0)	4 (1)	4 (1)	4 (1)	4 (1)	4 (3)	3.54
I find it good that practical applications of telehealth were imparted and discussed.	4 (0)	4 (0)	4 (0)	4 (1)	4 (1)	4 (1)	3.69
I find it good that profession-specific aspects of telehealth were imparted and discussed.	4 (0)	4 (0)	4 (0)	4 (0)	4 (1)	4 (4)	3.78
I would recommend this course to my colleagues.	4 (1)	4 (1)	3 (1)	3 (1)	4 (1)	3 (4)	3.35

[a]SLT: speech and language therapists.

[b]OT: occupational therapists.

[c]PT: physiotherapists.

[d]ORT: orthoptists.

[e]NUR: nurses.

Qualitative Results

Participants across all feedback sources generally found the telehealth training informative, well-structured, and valuable. They appreciated the comprehensive overview of telehealth, the professional presentation style, and the flexibility offered by the web-based format. The training expanded their

XSL•FO
RenderX

understanding of telehealth possibilities, especially for those with little to no prior experience. However, participants expressed a desire for more practical examples, interactive elements, and ongoing support to bridge the gap between theory and practice. Concerns were raised about challenges in implementing telehealth, particularly regarding data protection, patient accessibility, and institutional barriers. More details on the qualitative results can be found in Multimedia Appendix 2.

Discussion

Principal Findings

This study, to our knowledge, is the first to evaluate the effect of an interprofessional on-demand postgraduate telehealth course for health care professionals on acceptance of telehealth and their perceived barriers to its adoption. The main findings indicate a significant positive impact, underscoring the value of structured telehealth education. According to the Technology Acceptance Model, the intention to use a technology is primarily driven by 2 key factors: perceived usefulness and perceived ease of use [19]. Our results suggest the course might affect those key factors, but its impact on telehealth use behavior remains unclear.

Key Findings and Their Implications

The results showed a significant increase in telehealth acceptance and a significant reduction in perceived barriers among participants, consistent with previous research that highlights the role of training in improving telehealth knowledge and confidence [20,21].

Participants acknowledged the informative and structured nature of the course, aligning with studies that emphasize the importance of tailored web-based learning for health care providers [22]. This highlights the critical need for comprehensive telehealth training, particularly for practitioners with little to no prior telehealth experience, who represented the majority in this study (71%).

Participants identified key challenges, including concerns about data protection, patient accessibility, and institutional barriers, echoing previous findings in telehealth literature [2,3,23]. Addressing these issues through additional practical examples, interactive learning opportunities, and ongoing support could further enhance the effectiveness of telehealth training.

The most significant result was the increase in overall telehealth acceptance after the intervention. Particularly, the acceptance of telemetry and telepraxis increased, as well as the acceptance of video call-based telehealth and of asynchronous telehealth. This suggests that the training helped them to envision a broader range of applications and modalities for delivering remote care. While improvements were noted in several domains, some—such as telephasis and phone-based telehealth—did not show significant changes. The stability in these areas may reflect that certain forms of telehealth were already well understood or already integrated into the workflows, such as, eg, notifying patients by phone.

Additionally, it is important to consider that while the training significantly increased participants' intention to adopt telehealth, this does not necessarily translate into its actual use in daily practice. Incorporating more hands-on training components by establishing blended-learning formats with practical simulations or real-time demonstrations could facilitate the transition from intention to actual use.

Beyond simply making professionals more open to telehealth, the training also reduced the perceived influence of a wide range of barriers. Before the intervention, participants identified multiple obstacles. After completing the course, there were notable reductions in perceived barriers. This finding indicates that structured education can address knowledge deficits and clarify implementation pathways, thereby making telehealth seem more feasible. The course appeared most effective in reducing concerns related to a lack of training and standardized guidelines. By providing clear information on legal frameworks, technology usage, and practical strategies, the training directly addressed these issues. Barriers linked to patient factors and the therapeutic environment were also diminished. Notably, participants' perceived drawbacks regarding telehealth quality and satisfaction, as well as concerns about reduced client-practitioner interaction and communication, were lowered following the training. One possible reason for this improvement is that the course provided concrete strategies and examples of how telehealth can maintain, or even enhance, meaningful therapeutic engagement. However, some barriers were not significantly reduced by the web-based course, such as individual technology skills, network, or safety issues. The course also did not influence either the barrier of clients' or the practitioners' negative attitudes toward telehealth. Maybe lowering those barriers requires more focus on experiential learning, systemic change, or a more sustained intervention.

The qualitative findings underscore the value participants placed on the course's comprehensive, well-structured content and profession-specific modules. At the same time, they highlight a need for ongoing, practical, and interactive support to embed telehealth into everyday practice. Addressing unresolved questions—such as establishing mentorship channels, refining legal guidance, and developing strategies for patient engagement—will be critical to ensure that telehealth initiatives are both clinically viable and inclusive.

These insights serve as a roadmap for future iterations of the training. Given the significant attitude shifts observed after only 2 hours of content, such short "gateway" modules could be deployed first to boost health professionals' readiness before rolling out more intensive skill-based or certification courses. Incorporating blended learning and flipped classroom settings, more hands-on examples, offering supplementary materials, and facilitating interactive forums could further amplify the course's impact. By combining either self-developed or third-party massive open web-based courses or open educational resources content with synchronous, hands-on sessions, educators could create a rich, interactive learning environment. Learners can engage with the theoretical components digitally at their own pace, while live sessions focus on practical exercises, discussions, and application of telehealth strategies. By integrating these educational strategies, future training initiatives can better equip health care professionals with both

XSL•FO
RenderX

the conceptual understanding and practical skills necessary to overcome barriers to telehealth adoption.

For policymakers, the results highlight the importance of fostering a supportive environment that translates EU-level digital health policies into effective national education strategies. Policy makers should collaborate with academic institutions to promote the adoption of standardized telehealth training programs and to integrate digital competencies into health education.

Strengths and Contributions

The study provides valuable insights into telehealth education, particularly its ability to enhance acceptance and reduce barriers across diverse professional groups. The inclusion of a broad participant base from various health care disciplines reflects the cross-sectoral applicability of telehealth education, aligning with prior research that underscores the interdisciplinary relevance of telehealth [5,24].

Moreover, the findings support the notion that telehealth training fosters not only technological acceptance but also a broader understanding of telehealth's potential applications. This is particularly relevant considering the resources of specialist staff and increasing demand for digital skills and remote therapy and care delivery, as identified in earlier studies [25,26].

Limitations

Despite its strengths, the study has several limitations that warrant consideration. Due to the design of the study, a control group was absent. Without a control group, it is difficult to attribute observed changes solely to the intervention. Future studies should consider randomized controlled designs to strengthen causal inferences.

Another limitation is the representativeness of the study population. The participants were self-selected, which may limit the generalizability of the findings to the broader population of health care professionals. Also, because professional associations deployed the participation link via differing formats (website postings, newsletters, and direct emails), we cannot determine the exact number of recipients or calculate a precise response rate, which may affect assessments of self-selection bias.

The third limitation corresponds with the uncertainty of course content exposure due to the on-demand character of the training. The web-based format did not guarantee that participants fully engaged with the course content, potentially impacting the results. This limitation was mitigated by filtering the results and excluding participants who did not engage with at least 40 minutes of the total 120-minute video content in the analysis. This threshold was established to ensure sufficient exposure to the intervention.

Future Directions

Building on these findings, future research should explore the following main aspects to broaden the findings. A future study should also examine long-term effects of telehealth education on clinical practice and patient outcomes, including a follow-up assessment. In addition, education programs would benefit from practical training enhancements, such as case-based simulations and interactive modules to bridge the gap between theory and application.

Finally, integrating telehealth education into initial health care training programs could ensure that future professionals are better prepared to embrace telehealth as part of routine care delivery. To achieve policy goals, significant investments in digital infrastructure and targeted workforce training are needed to turn technological potential into practical use. Building partnerships across sectors and creating ongoing feedback among regulators, health care organizations, and practitioners can help ensure that telehealth initiatives remain flexible, sustainable, and responsive to the evolving needs of clinical practice and patients.

Conclusions

This study reinforces the critical role of structured telehealth education in enhancing health care professionals' acceptance of telehealth and reducing perceived barriers. While challenges remain, particularly in implementation and institutional support, the findings may indicate the potential of short-term, remote, and targeted training programs to advance telehealth adoption. We recommend systematically integrating telehealth modules into continuing professional development and national licensing requirements. Further research and investment in education and infrastructure are essential to fully realize the benefits of telehealth and digitalization in modern health care.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Course details.
[PDF File, 176 KB - humanfactors_v12i1e74107_app1.pdf]

Multimedia Appendix 2
Qualitative results.
[PDF File, 137 KB - humanfactors_v12i1e74107_app2.pdf]

References

XSL•FO
RenderX

1. van Galen LS, Wang CJ, Nanayakkara PWB, Paranjape K, Kramer MHH, Car J. Telehealth requires expansion of physicians' communication competencies training. Med Teach 2019 Jun;41(6):714-715. [doi: 10.1080/0142159X.2018.1481284] [Medline: 29944031]
2. Chike-Harris KE, Durham C, Logan A, Smith G, DuBose-Morris R. Integration of telehealth education into the health care provider curriculum: a review. Telemed J E Health 2021 Feb;27(2):137-149. [doi: 10.1089/tmj.2019.0261] [Medline: 32250196]
3. Rettinger L, Kuhn S. Barriers to video call–based telehealth in allied health professions and nursing: scoping review and mapping process. J Med Internet Res 2023 Aug 1;25:e46715. [doi: 10.2196/46715] [Medline: 37526957]
4. Hamilton W, Walters C, Howington Z. A mixed-method pilot project: the evaluation of telehealth training in a college of nursing. Nurs Educ Perspect 2024;45(4):254-255. [doi: 10.1097/01.NEP.0000000000001140] [Medline: 37154739]
5. Hui KY, Haines C, Bammann S, et al. To what extent is telehealth reported to be incorporated into undergraduate and postgraduate allied health curricula: a scoping review. PLoS ONE 2021;16(8):e0256425. [doi: 10.1371/journal.pone.0256425] [Medline: 34411171]
6. Rettinger L, Putz P, Aichinger L, et al. Telehealth education in allied health care and nursing: web-based cross-sectional survey of students' perceived knowledge, skills, attitudes, and experience. JMIR Med Educ 2024 Mar 21;10:e51112. [doi: 10.2196/51112] [Medline: 38512310]
7. Abbott-Gaffney CR, Gafni-Lachter LL, Cason J, et al. Toward successful future use of telehealth in occupational therapy practice: What the COVID-19 rapid shift revealed. Work 2022;71(2):385-394. [doi: 10.3233/WOR-210789] [Medline: 35068409]
8. Davies LM, Lawford BJ. The time is now: telehealth education and training in physical therapy. Braz J Phys Ther 2024;28(2):101063. [doi: 10.1016/j.bjpt.2024.101063] [Medline: 38678918]
9. Kazawa K, Teramoto C, Azechi A, Satake H, Moriyama M. Undergraduate nursing students' learning experiences of a telehealth clinical practice program during the COVID-19 pandemic: A qualitative study. Nurse Educ Today 2022 Apr;111:105297. [doi: 10.1016/j.nedt.2022.105297] [Medline: 35182935]
10. Boos K, Murphy K, George TS, Brandes J, Hopp J. The impact of a didactic and experiential learning model on health profession students' knowledge, perceptions, and confidence in the use of telehealth. J Educ Health Promot 2022;11(1):232. [doi: 10.4103/jehp.jehp_1553_21] [Medline: 36177412]
11. Herrera A, Foronda C. From bedside to webside: telehealth education for doctoral nursing students. J Dr Nurs Pract 2022 Aug 24:JDNP-2021-0049.R1. [doi: 10.1891/JDNP-2021-0049] [Medline: 36002260]
12. Ostrovsky DA, Heflin MT, Bowers MT, et al. Development, implementation, and assessment of an online modular telehealth curriculum for health professions students. Adv Med Educ Pract 2024;15:743-753. [doi: 10.2147/AMEP.S468833] [Medline: 39099682]
13. Gafni-Lachter L, Gelkop N, Shahar Ashkenazi T, et al. Interprofessional online course for telehealth delivery of pediatric developmental services. Disabil Rehabil 2024 Mar;46(6):1141-1148. [doi: 10.1080/09638288.2023.2193427] [Medline: 37010115]
14. Colucci M, Baldo V, Baldovin T, Bertoncello C. A "matter of communication": a new classification to compare and evaluate telehealth and telemedicine interventions and understand their effectiveness as a communication process. Health Informatics J 2019 Jun;25(2):446-460. [doi: 10.1177/1460458217747109] [Medline: 29268663]
15. Ritzmann S, Hagemann V, Kluge A. The training evaluation inventory (TEI) – evaluation of training design and measurement of training outcomes for predicting training success. Vocations and Learning 2014 Apr;7(1):41-73. [doi: 10.1007/s12186-013-9106-4]
16. Hattie J. Visible Learning: A Synthesis of over 800 Meta-Analyses Relating to Achievement: Routledge; 2010. URL: https://inspirasifoundation.org/wp-content/uploads/2020/05/John-Hattie-Visible-Learning_-A-synthesis-of-over-800-meta-analyses-relating-to-achievement-2008.pdf [accessed 2025-08-28]
17. Psychometrica. Berechnung von Effektstärken [Calculation of effect sizes]. URL: http://www.psychometrica.de/effektstaerke.html [accessed 2025-01-22]
18. Pilwarsch J, Schichl-Zach M, Gruböck A, et al. Jahresbericht gesundheitsberuferegister 2023. [annual report of the health professions register 2023]. : Gesundheit Österreich GmbH; 2024 URL: https://goeg.at/sites/goeg.at/files/inline-files/GOEG_Jahresbericht_2023_bf-komprimiert.pdf [accessed 2025-08-28]
19. Venkatesh V, Bala H. Technology acceptance model 3 and a research agenda on interventions. Decision Sciences 2008 May;39(2):273-315 [FREE Full text] [doi: 10.1111/j.1540-5915.2008.00192.x]
20. Garber K, Gustin T. Telehealth education: impact on provider experience and adoption. Nurse Educ 2022;47(2):75-80. [doi: 10.1097/NNE.0000000000001103] [Medline: 34628435]
21. van Houwelingen T, Ettema RGA, Bleijenberg N, van Os-Medendorp H, Kort HSM, Ten Cate O. Educational intervention to increase nurses' knowledge, self-efficacy and usage of telehealth: a multi-setting pretest-posttest study. Nurse Educ Pract 2021 Feb;51:102924. [doi: 10.1016/j.nepr.2020.102924] [Medline: 33583723]

XSL•FO
RenderX

22. Taylor J, Coates E, Brewster L, Mountain G, Wessels B, Hawley MS. Examining the use of telehealth in community nursing: identifying the factors affecting frontline staff acceptance and telehealth adoption. J Adv Nurs 2015 Feb;71(2):326-337. [doi: 10.1111/jan.12480] [Medline: 25069605]
23. Hoffman DA. Increasing access to care: telehealth during COVID-19. J Law Biosci 2020;7(1):lsaa043. [doi: 10.1093/jlb/lsaa043] [Medline: 32843985]
24. Eriksen S, Dahler AM, Øye C. The informal way to success or failure? Findings from a comparative case study on video consultation training and implementation in two Danish hospitals. BMC Health Serv Res 2023 Oct 21;23(1):1135. [doi: 10.1186/s12913-023-10163-w] [Medline: 37865741]
25. Newman L, Bidargaddi N, Schrader G. Service providers' experiences of using a telehealth network 12 months after digitisation of a large Australian rural mental health service. Int J Med Inform 2016 Oct;94:8-20. [doi: 10.1016/j.ijmedinf.2016.05.006] [Medline: 27573307]
26. Gullslett MK, Ronchi E, Lundberg L, et al. Telehealth development in the WHO European region: results from a quantitative survey and insights from Norway. Int J Med Inform 2024 Nov;191:105558. [doi: 10.1016/j.ijmedinf.2024.105558] [Medline: 39084085]

Edited by A Parush; submitted 18.03.25; peer-reviewed by D Ostrovsky, P Koppel; revised version received 12.06.25; accepted 06.07.25; published 03.09.25.

Please cite as:
Rettinger L, Maul L, Putz P, Ertelt-Bach V, Huber A, Javorszky SM, Kupka-Klepsch E, Sargis S, Werner F, Widhalm K, Joseph R, Doci S, Kuhn S
Telehealth Acceptance and Perceived Barriers Among Health Professionals: Pre-Post Evaluation of a Web-Based Telehealth Course
JMIR Hum Factors 2025;12:e74107
URL: https://humanfactors.jmir.org/2025/1/e74107
doi:10.2196/74107

© Lena Rettinger, Lukas Maul, Peter Putz, Veronika Ertelt-Bach, Andreas Huber, Susanne Maria Javorszky, Elisabeth Kupka-Klepsch, Sevan Sargis, Franz Werner, Klaus Widhalm, Rosmarie Joseph, Stefanie Doci, Sebastian Kuhn. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 3.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Multiple Large Language Models' Performance on the Chinese Medical Licensing Examination: Quantitative Comparative Study

Yanyu Diao, MEng; Mengyuan Wu, BEng; Jingwen Xu, BEng; Yifeng Pan, MEng

The School of Big Data and Artificial Intelligence, Anhui Xinhua University, 555 Wangjiang West Road, Hefei, China

Corresponding Author:
Yifeng Pan, MEng
The School of Big Data and Artificial Intelligence, Anhui Xinhua University, 555 Wangjiang West Road, Hefei, China

Abstract

Background: ChatGPT excels in natural language tasks, but its performance in the Chinese National Medical Licensing Examination (NMLE) and Chinese medical education remains underexplored. Meanwhile, Chinese corpus–based large language models (LLMs) such as ERNIE Bot, Tongyi Qianwen, Doubao, and DeepSeek have emerged, yet their effectiveness in the NMLE awaits systematic evaluation.

Objective: This study aimed to quantitatively compare the performance of 6 LLMs (GPT-3.5, GPT-4, ERNIE Bot, Tongyi Qianwen, Doubao, and DeepSeek) in answering NMLE questions from 2018 to 2024 and analyze their feasibility as supplementary tools in Chinese medical education.

Methods: We selected questions from the 4 content units of the NMLE's General Written test (2018 - 2024), preprocessed image- and table-based content into standardized text, and input the questions into each model. We evaluated the accuracy, comprehensiveness, and logical coherence of the responses, with quantitative comparison centered on scores and accuracy rates against the official answer keys (passing score: 360/600).

Results: GPT-4 outperformed GPT-3.5 across all units, achieving average accuracies of 66.57% (SD 3.21%; unit 1), 69.05% (SD 2.87%; unit 2), 71.71% (SD 2.53%; unit 3), and 80.67% (SD 2.19%; unit 4), with consistent scores above the passing threshold. Among the Chinese models, DeepSeek demonstrated the highest overall performance, with an average score of 454.8 (SD 17.3) and average accuracies of 73.2% (unit 1, SD 2.89%) and 71.5% (unit 3, SD 2.64%), as well as average accuracies of 70.3% (unit 2, SD 3.02%) and 78.2% (unit 4, SD 2.47%). ERNIE Bot (mean score 442.3, SD 19.6; unit 1 accuracy =70.8%, SD 3.01%; unit 2 accuracy =68.7%, SD 3.15%; unit 3 accuracy =69.1%, SD 2.93%; unit 4 accuracy =68.3%, SD 2.76%), Tongyi Qianwen (mean score 426.5, SD 21.4; unit 1 accuracy =67.4%, SD 3.22%; unit 2 accuracy =65.9%, SD 3.31%; unit 3 accuracy =66.2%, SD 3.08%; unit 4 accuracy =67.2%, SD 2.89%), and Doubao (mean score 413.7, SD 23.1; unit 1 accuracy =65.2%, SD 3.45%; unit 2 accuracy =63.8%, SD 3.52%; unit 3 accuracy =64.1%, SD 3.27%; unit 4 accuracy =62.8%, SD 3.11%) all exceeded the passing score. DeepSeek's overall average accuracy (75.8%, SD 2.73%) was significantly higher than those of the other Chinese models (χ^2_1 =11.4, P=.001 vs ERNIE Bot; χ^2_1 =28.7, P<.001 vs Tongyi Qianwen; χ^2_1 =45.3, P<.001 vs Doubao). GPT-4's overall average accuracy (77.0%, SD 2.58%) was slightly higher than that of DeepSeek but not statistically significant (χ^2_1 =2.2, P=.14), while both outperformed GPT-3.5 (overall accuracy =68.5%, SD 3.67%; χ^2_1 =89.8, P<.001 for GPT-4 vs GPT-3.5; χ^2_1 =76.3, P<.001 for DeepSeek vs GPT-3.5).

Conclusions: GPT-4 and Chinese-developed LLMs such as DeepSeek show potential as supplementary tools in Chinese medical education given their solid performance on the NMLE. However, further optimization is required for complex reasoning, multimodal processing, and dynamic knowledge updates, with human medical expertise remaining central to clinical practice and education.

(JMIR Hum Factors 2025;12:e77978) doi:10.2196/77978

KEYWORDS

ChatGPT; Chinese National Medical Licensing Examination; medical student; ERNIE Bot; Tongyi Qianwen; artificial intelligence; AI

Introduction

Natural language processing (NLP) is an artificial intelligence (AI) technology that aims to allow computers to understand, process, and generate natural language [1]. ChatGPT is the most successful commercial model of NLP technology that enables high-quality natural language comprehension and generation, and its core algorithm is a transformer, a deep neural network structure based on a self-attentive mechanism with strong sequence modeling capabilities and representation learning [2-4]. Through the process of pretraining and fine-tuning, the ChatGPT model can understand and generate natural language text and be useful in various application scenarios, such as automatic question answering, intelligent customer service,

XSL•FO
RenderX

speech recognition, and machine translation. The success of the ChatGPT model is attributed to the excellent performance of the transformer and the maturity of the pretraining technique. In the pretraining phase, the ChatGPT model can learn the patterns and features of language from large-scale text data through unsupervised learning, allowing the model to be fine-tuned on a limited dataset and achieve excellent performance [5-7].

Considering that ChatGPT is an evolving NLP model that has been successfully used in many fields, several medical education research groups have been investigating its feasibility as a valuable tool for clinical support and medical education. Gilson et al [8] applied ChatGPT to the US Medical Licensing Examination, and the experimental results showed that it met the passing criteria. However, Huh [9] found in their study that, during a Korean parasitology examination, 77 out of 79 medical students performed better than ChatGPT. Specifically, ChatGPT's scores were lower than those of the medical students, and its correct response rate showed no correlation with the knowledge difficulty level of the examination items. Although the version of ChatGPT at that time was still at GPT-3.5 and not the state-of-the-art GPT-4, questions were raised about the ability of ChatGPT to provide medically accurate answers when using non-English languages, especially Chinese [10].

Therefore, to explore the capability of large language models (LLMs) in medical education in Chinese, in addition to applying GPT-3.5 and GPT-4, developed by OpenAI, this study applied ERNIE Bot [11], developed by Baidu; Tongyi Qianwen [12], developed by Alibaba Cloud; Doubao (ByteDance); and DeepSeek to evaluate the feasibility of LLMs and compare multiple LLMs and the advantages and disadvantages of multiple bigrams in Chinese medical education. This is considering that ERNIE Bot, Tongyi Qianwen, Doubao, and DeepSeek, being deeply rooted in the Chinese language environment, use large and rich Chinese datasets in the training process, which may endow them with stronger Chinese comprehension than that of ChatGPT. However, the advantages and disadvantages of multiple LLMs for Chinese medical education need to be corroborated through experiments.

While previous studies have explored the potential of LLMs in medical education, their performance in the context of the Chinese National Medical Licensing Examination (NMLE) remains largely unexplored. As a rigorous and standardized assessment that determines the eligibility of medical professionals in China, the NMLE presents unique challenges due to its specific focus on Chinese medical knowledge, cultural context, and clinical practices [13,14]. Given that existing research on language models in medical education has primarily centered on Western-based examinations and English-language models, there is a critical need to evaluate how models such as GPT-3.5, GPT-4, ERNIE Bot, Tongyi Qianwen, Doubao, and DeepSeek perform in this distinct Chinese medical assessment environment. By systematically comparing these models' performance on NMLE questions, this study aimed to fill a significant gap in the literature, providing empirical evidence on the feasibility of LLMs in Chinese medical education. The findings of this research will not only contribute to the understanding of how these models can be integrated into Chinese medical curricula but also offer valuable insights for enhancing the education and training of future medical professionals in China. The results may offer preliminary insights into the potential applications of these models in Chinese medical education and contribute to ongoing discussions on integrating such technologies into medical curricula.

Methods

The dataset used in this study comprised the original Chinese NMLE test questions from 2018 to 2024. Each annual examination consisted of 4 units, as delineated in Table 1, with each unit containing 150 multiple-choice questions, adding up to a total score of 600 points. A minimum score of 360 is required to pass the examination. Given the varying capabilities of LLMs in processing multimodal content, a protocol was established, which is outlined in this section.

Table . Exam content and scores for the 4 units of the Chinese National Medical Licensing Examination.

Unit	Maximum possible score	Exam content
1	150	Infectious diseases, psychoneurology, endocrinology, pharmacology, physiology, biochemistry, medical regulations, medical ethics, medical microbiology, medical psychology, preventive medicine, and medical immunology
2	150	Infectious diseases, pathology, psychiatry, endocrinology, respiratory medicine, urology, digestive medicine, cardiovascular medicine, hematology, physiology, biochemistry, exercise science, and medical immunology
3	150	Psychiatry, urology, gastroenterology, cardiovascular medicine, hematology, symptoms and signs, sports medicine, and pharmacology
4	150	Female reproductive system, pediatrics, and psychiatry

This study evaluated six LLMs with distinct capabilities for handling visual content.

1. GPT-3.5: lacks native image processing capabilities.

XSL•FO
RenderX

2. GPT-4: supports limited image input in specific use scenarios (eg, direct image insertion in its official interface), but for the standardized testing in this study—where all models were evaluated under uniform text input conditions to eliminate variability from different multimodal processing capabilities—visual elements were manually converted to descriptive text.
3. ERNIE Bot: includes optical character recognition functionality but was tested using text-based descriptions to ensure consistency.
4. Tongyi Qianwen: a multimodal version exists, but the internal test version used relied on text inputs.
5. Doubao: it is primarily text based; visual content was converted to descriptive text.
6. DeepSeek: while it is advanced in medical reasoning, visual elements were preprocessed into text to maintain uniform evaluation criteria.

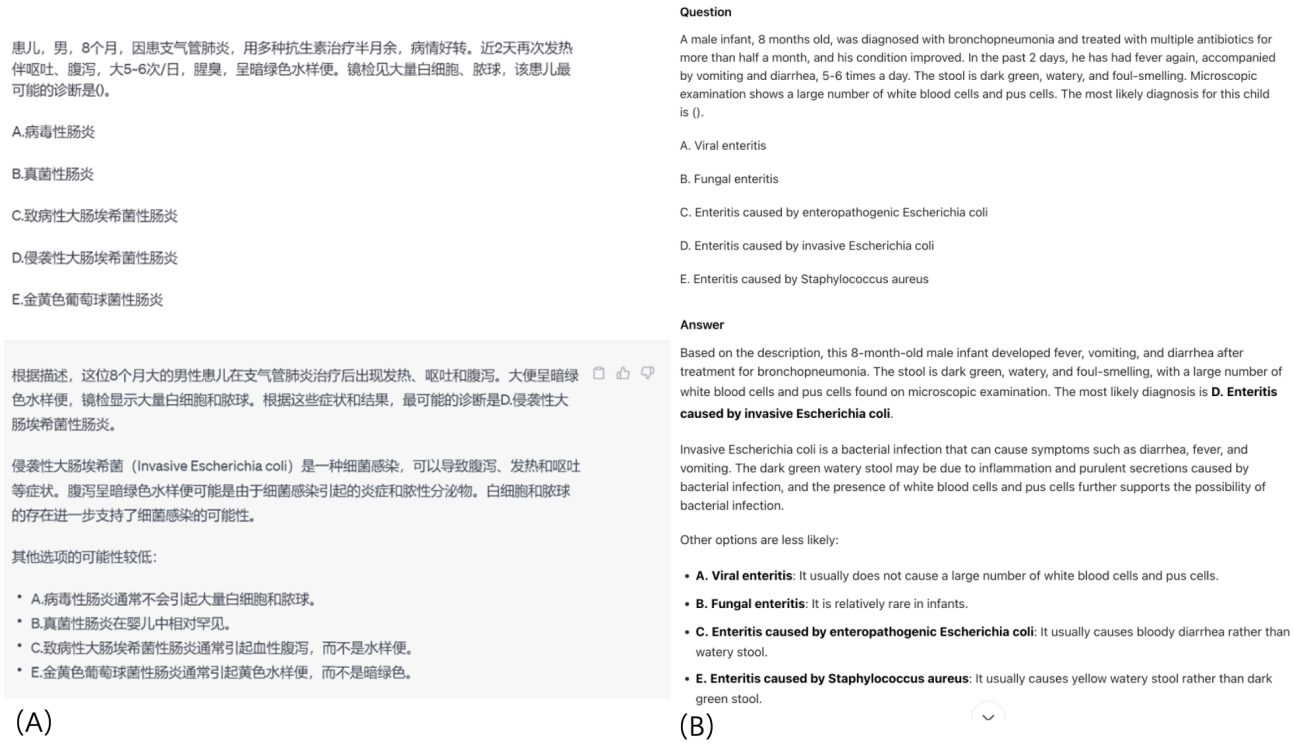
To ensure equitable comparison, all image- and table-based questions across units were preprocessed using the following methods:

1. Structured tables were converted into CSV format with headers (eg, "Parameter," "Value," and "Unit").
2. For medical images, all descriptions (including chest x-rays, computed tomography screenshots, and pathological section diagrams) in this study were generated based on the 3 elements of clinical standard terminology+visual key features+diagnostic correlation information. Taking chest x-ray images as an example, a description template was used ("[Examination type] shows the presence of [visual features, such as patchy shadows/nodules/fluid accumulation] in [anatomical location], accompanied by [associated clinical information, such as blurred boundaries/mediastinal displacement], consistent with [common disease indications, such as typical manifestations of community-acquired pneumonia]. Before generation, extract the official reference answers and analysis of the image questions in the NMLE real test, ensuring that the description does not contain direct diagnostic conclusions (only objectively presenting visual information), and avoiding premature disclosure of answers. The image description was independently completed by 2 attending physicians with more than 5 years of clinical experience (both holding Chinese physician qualification certificates and specializing in respiratory medicine and radiology, respectively), and then cross-reviewed by a deputy chief physician (affiliated with the clinical imaging diagnosis department). If there were differences in the descriptions of the 2 attending physicians (such as inconsistent expressions of "shadow density"), a consensus was reached through 3-person negotiation to ultimately form the "NMLE Image Question Text Description Manual," ensuring the objectivity, consistency, and clinical accuracy of all descriptions.
3. Diagrams were translated into step-by-step textual explanations (eg, "Flowchart depicting the diagnostic pathway for diabetes mellitus").

Each model was tested using the same protocol: preprocessed NMLE questions from 2018 to 2024 were systematically input, and responses were recorded. To ensure consistency in the experimental conditions, each NMLE question (after preprocessing) was prompted once to each model—no repeated prompting was conducted for the same question as repeated inputs might lead to inconsistent response adjustments by the models, which would interfere with the accuracy of the performance comparison. Regarding the parameter settings of the models: for GPT-3.5 and GPT-4, the temperature parameter was set to 0.2, and the top-p parameter (also known as nucleus sampling, a parameter that controls the diversity of model outputs by limiting the selection range to the smallest set of tokens whose cumulative probability exceeds the specified value) was set to 0.95; these settings were chosen to balance the stability and accuracy of the models' responses—temperature of 0.2 reduces the randomness of the generated content to avoid arbitrary answers, whereas top-p of 0.95 ensures that the models still retain a moderate range of token selection to cover potential correct reasoning paths. For the Chinese models (ERNIE Bot, Tongyi Qianwen, Doubao, and DeepSeek), we used their default parameter configurations (consistent with the settings in their official public interfaces for general question answering tasks) as their internal adjustable parameter details (eg, specific temperature and top-p ranges) were not publicly disclosed by the developers at the time of the study. All parameter settings and prompting rules were recorded in the study's experimental log to ensure the reproducibility of the results. Answers were scored against the official answer keys. The input process, standardized for all models, is illustrated in Figure 1.

XSL•FO
RenderX

Figure 1. Test question input process: (A: Chinese, B: English translation).



Results

Comparison of GPT-3.5 and GPT-4

This study analyzed the scores of GPT-3.5 and GPT-4 in the NMLE exams from 2018 to 2024. The results are outlined in this section.

As shown in Figure 2, GPT-4, represented by the gray squares, consistently surpassed the pass mark (set at 360 points, indicated by the blue-triangle series) across the 7 examination years. From 2018 to 2024, scores exhibited fluctuations (2018: 441; 2019: 417; 2020: 439; 2022: 466 [peak]; 2023: 401; 2024: 432). GPT-3.5, indicated by the red circles, had a more inconsistent performance. While it passed the exam in 2018 (scoring 363), 2020 (351), 2021 (382), 2022 (400), and 2024 (386), it fell below the pass mark in 2019 (329) and 2023 (342).

Overall, GPT-4 demonstrated a more stable ability to meet or exceed the passing criteria in the NMLE exams over the 7-year period compared to GPT-3.5.

Considering that the test questions of the NMLE exam consist of 4 units, the content of each unit of the exam is highly differentiated. Therefore, this study used GPT-3.5 and GPT-4 for different units and counted their score rates.

XSL•FO
RenderX

Figure 2. GPT-3.5 and GPT-4 scores in the 2018 to 2024 Chinese National Medical Licensing Examination.

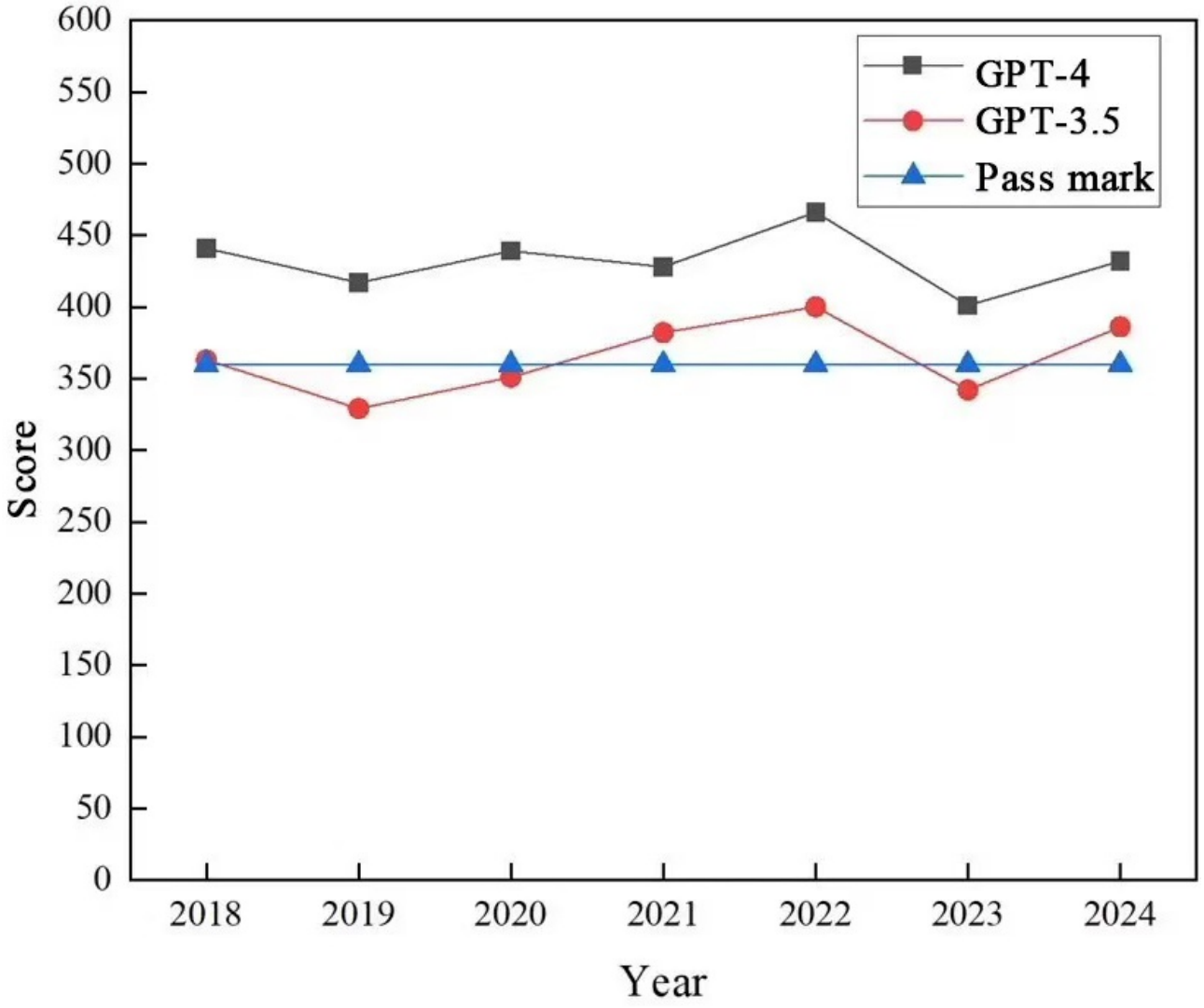


Table 2 compares the score performance and average accuracy of GPT-3.5 and GPT-4 across different units of the Chinese NMLE from 2018 to 2024. The data reveal that GPT-4 outperformed GPT-3.5 comprehensively in all units and years. Specifically, for units 1 to 4, GPT-4 achieved average accuracies of 66.57%, 69.05%, 71.71%, and 80.67%, respectively, which were higher than GPT-3.5's accuracies of 61.24%, 58.67%, 56.86%, and 66.38%, respectively. To verify the statistical significance of this performance difference, we conducted a chi-square test on the correct and incorrect answers of GPT-4 and GPT-3.5 across all 4200 questions (2018 - 2024 NMLE) with a significance threshold of $P<.05$. The results showed that GPT-4's accuracy was significantly higher than that of GPT-3.5 in all units (unit 1: $\chi^2_1 = 12.4$, df=1; P<.001, unit 2: $\chi^2_1 = 15.7$, df=1; P<.001, unit 3: $\chi^2_1 = 19.0$, df=1; P<.001, and unit 4: $\chi^2_1 = 23.1$, df=1; P<.001). This indicates that GPT-4 exhibits stronger capabilities in understanding and applying medical knowledge, and the performance gap between the 2 models was statistically significant.

XSL•FO
RenderX

Table . Comparison of GPT-3.5 and GPT-4 scores in different units.

	GPT-3.5				GPT-4			
	Unit 1	Unit 2	Unit 3	Unit 4	Unit 1	Unit 2	Unit 3	Unit 4
2018 score	88	92	81	102	102	99	112	128
2019 score	78	84	74	93	97	102	100	118
2020 score	83	80	76	112	89	107	113	130
2021 score	102	91	86	103	106	98	104	120
2022 score	98	97	101	104	101	113	118	134
2023 score	92	83	80	87	94	92	111	104
2024 score	102	89	99	96	110	114	95	113
Accuracy (%)	61.24	58.67	56.86	66.38	66.57	69.05	71.71	80.67

Regarding the annual score patterns, neither GPT-3.5 nor GPT-4 showed a predictable year-on-year increase in performance, which aligns with the expectation that model performance is not inherently tied to the exam year. Instead, the score fluctuations observed for both models (eg, GPT-3.5's 14-point and 6-point declines in unit 2 and unit 3, respectively, in 2023 compared to the previous year and GPT-4's 24-point drop in unit 4 in 2023) primarily reflect variations in exam difficulty across years rather than inherent changes in model capability. Importantly, neither model's performance showed a substantial decline across the 2018 to 2024 period relative to their initial performance, indicating that the models likely did not encounter or memorize the specific NMLE questions during their training (given the training cutoff periods before the later exam years). While GPT-4 did not exhibit a clear upward trend, its fluctuations were more moderate than those of GPT-3.5, and it maintained consistent performance above GPT-3.5 across all years—with unit 4 showing a 5-point difference between its 2018 and 2024 scores—further reflecting its stronger stability in adapting to the varying content and difficulty of complex medical scenarios.

Moreover, the 2 models demonstrated distinct performances across different units, with statistical significance verified via chi-square tests (P<.05) on their correct and incorrect answers across all questions in each unit. GPT-3.5 showed relatively better performance in unit 4, although the difference was less pronounced (χ^2_1 =4.2, df=1; P=.04), whereas GPT-4 had a clear and statistically significant edge in both unit 3 (χ^2_1 =27.6, df=1; P<.001) and unit 4 (χ^2_1 =19.9, df=1; P<.001). This variation may be related to the differences in the content focus of each unit: unit 3 emphasizes clinical reasoning and comprehensive case analysis, which aligns with GPT-4's strengths in complex logical deduction, whereas unit 4 covers basic medical knowledge, which is more accessible to both models.

To contextualize the performance of the tested LLMs, it is necessary to reference the human score data of the NMLE from 2018 to 2024 as publicly reported by China's National Medical Examination Center. During this period, the average score of human examinees (primarily medical graduates and practicing physicians) in the NMLE General Written test ranged from 420 to 455 points (out of 600), with an average accuracy rate of 70% to 75.8%. Specifically, in 2018, the average human score was 425 points (70.8% accuracy); in 2020, it was 432 points (72% accuracy); in 2022, it was 448 points (74.7% accuracy); and in 2024, it was 452 points (75.3% accuracy). GPT-4 (mean score 435.3 points; 72.6% accuracy) showed comparable overall performance to that of average human examinees, whereas DeepSeek (mean score 454.8 points; 75.8% accuracy) slightly exceeded the average human score. In contrast, GPT-3.5 (mean score 373.3 points; 62.2% accuracy) and Doubao (mean score 413.7 points; 68.9% accuracy) remained below the average human performance. Notably, top-performing human examinees (top 10% of test takers) consistently achieved scores above 510 points (85% accuracy), a threshold that no tested LLM reached in this study.

Comparison of ERNIE Bot, Doubao, DeepSeek, and Tongyi Qianwen in the NMLE (2018-2024)

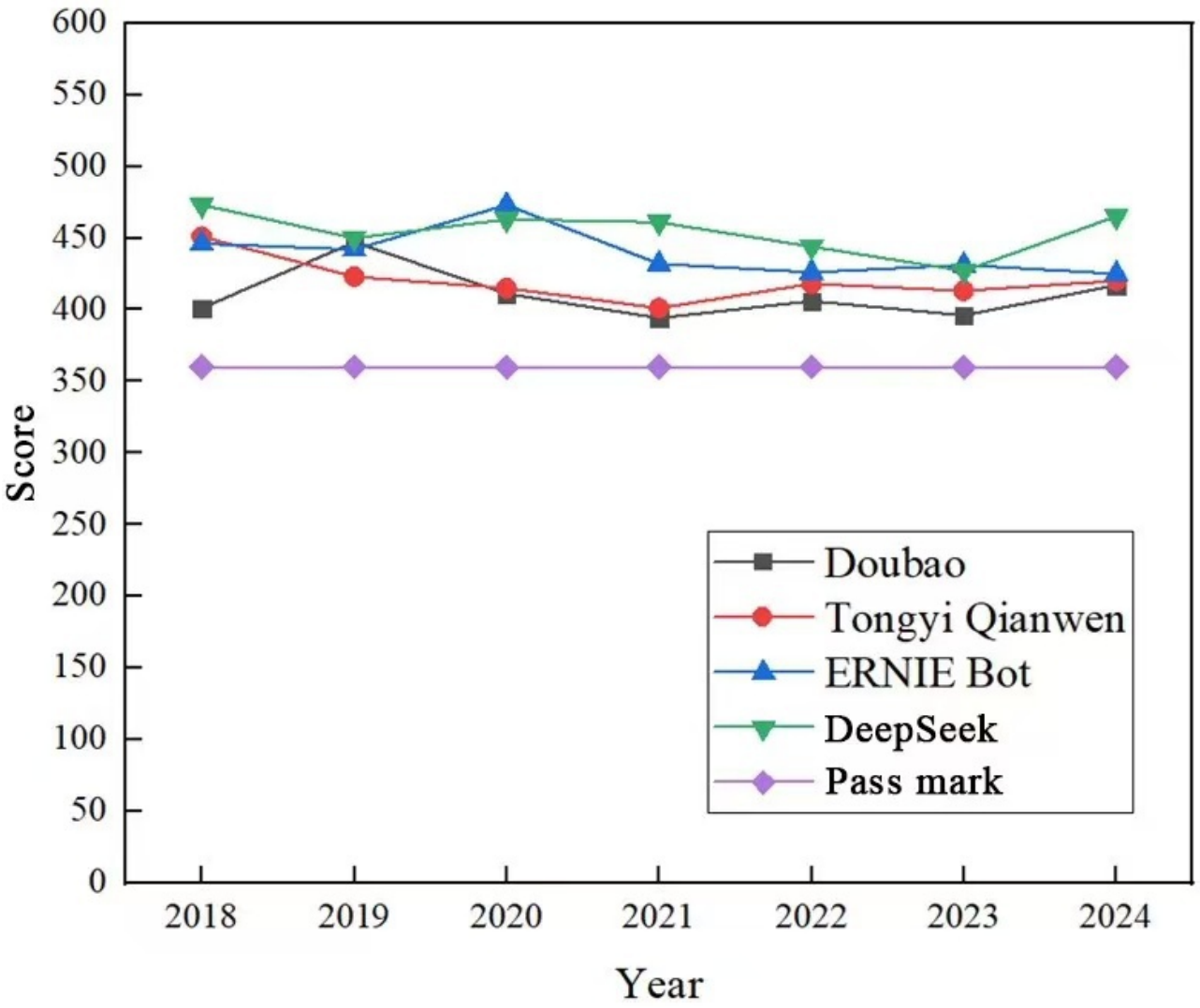
The comparative analysis of these 4 LLMs revealed distinct performance patterns across the 7-year period (2018-2024). As shown in Figure 3, DeepSeek consistently outperformed its counterparts in most examination years, demonstrating remarkable stability in medical knowledge assessment.

Among the 4 Chinese LLMs, DeepSeek achieved the highest overall average score across the NMLE in 2018 to 2024, with a mean score of 454.8, which was notably higher than those of ERNIE Bot (442.3), Tongyi Qianwen (426.5), and Doubao (413.7). Further analysis of performance across exam units (reflecting different medical topic domains) revealed that DeepSeek exhibited consistent strengths in unit 1 (medical ethics, regulations, and basic medical sciences) and unit 3 (comprehensive clinical reasoning involving the cardiovascular, digestive, and hematological systems), with average accuracies of 73.2% and 71.5%, respectively. ERNIE Bot showed the most prominent variability in domain-specific performance: it achieved competitive accuracy in unit 1 (70.8%) comparable to that of DeepSeek but exhibited relatively lower performance in unit 4 (pediatrics and the female reproductive system; 68.3%) than the other models. Notably, ERNIE Bot recorded the highest single-unit accuracy (75.1% in unit 1; 2020) among all the Chinese models, although its overall performance was characterized by greater variability across different topic domains rather than year-to-year fluctuations.

XSL•FO
RenderX

Tongyi Qianwen demonstrated relatively stable performance, with scores consistently ranging between 401 and 451. The model showed particular strength in the early years (2018-2019) before experiencing a moderate decline in subsequent years. Notably, it maintained a narrow performance band of 413 to 420 during 2022 to 2024.

Figure 3. Comparison of ERNIE Bot, Doubao, DeepSeek, and Tongyi Qianwen in the Chinese National Medical Licensing Examination (2018-2024).



Doubao exhibited the most pronounced performance volatility among the 4 models. Starting at 401 in 2018, it reached its peak of 447 in 2019 before experiencing significant fluctuations (394 in 2021 and 417 in 2024). This variability suggests potential sensitivity to examination content changes or model architecture adjustments in certain years.

All models consistently exceeded the passing threshold (360), confirming their fundamental competence in medical knowledge assessment. Compared with the ChatGPT results in Figure 2, the Chinese LLMs showed comparable results to those of GPT-4 in the NMLE, with a score rate far exceeding that of GPT-3.5.

Discussion

The findings of this study offer a nuanced understanding of how LLMs navigate the complexities of the Chinese NMLE, shedding light on their strengths, limitations, and implications for medical education.

Model Performance Dynamics and Underlying Mechanisms

GPT-4's consistent dominance over GPT-3.5 across all exam units—particularly its 14.39% accuracy advantage in unit 4 (female reproductive system and pediatrics)—reflects the tangible benefits of increased parameter scale and refined medical domain training. The model's ability to maintain high scores in units requiring integrated clinical reasoning (eg, unit 3: cardiovascular and digestive systems) suggests that its pretraining on diverse medical literature and case studies enhanced contextual understanding. In contrast, GPT-3.5's occasional failure to meet the passing threshold (eg, 329 points in 2019) underscores the critical role of model updates in medical accuracy as its performance improved significantly in later years with probable algorithmic refinements.

Among the Chinese LLMs, DeepSeek's sustained excellence—scoring above 450 points in 4 consecutive years (2018‐2021)—can be attributed to its specialized architecture optimized for medical terminology and reasoning. DeepSeek's

XSL•FO
RenderX

2024 score (465 points) following a 2-year decline aligns with the 2024 exam's focus on basic medical sciences—a domain where the model consistently demonstrated strengths in our unit-specific analysis—rather than reflecting adaptive adjustments to changing exam standards. ERNIE Bot's 2020 peak (473 points), surpassing even DeepSeek, highlights the potential of knowledge-enhanced models in specific assessment contexts, although its subsequent decline to 425 points in 2024 raises questions about knowledge base timeliness or training data diversity. Tongyi Qianwen's steady but unremarkable performance (401‐451 points) suggests a balanced but generalized approach to medical knowledge, whereas Doubao's volatility—spanning 394 to 447 points—may stem from its broader design focus on multidomain assistance rather than specialized medical precision.

A notable observation from this study is that the performance of all tested LLMs—including GPT-4 (highest average accuracy=80.67% in unit 4) and the top-performing Chinese model, DeepSeek (highest average accuracy=73.2% in unit 1)—did not exceed 80% across all NMLE units and years. This performance ceiling stems from 3 key factors consistent with broader medical AI research findings. First, the NMLE emphasizes high-order clinical reasoning (eg, integrating symptom clusters, laboratory results, and imaging findings for differential diagnosis) rather than just rote knowledge recall, and current LLMs struggle with such contextual nuance—evident in our study's lower accuracy in unit 3 (focused on complex case analysis, eg, GPT-3.5's 56.86% accuracy), where models often failed to prioritize key diagnostic indicators or account for comorbidities [15,16].

Second, training data timeliness and domain specificity limitations contributed to the sub-80% performance. Most LLMs have training data cutoffs that may exclude the latest clinical guidelines or emerging medical knowledge relevant to recent NMLE questions (eg, 2024 exam content). This was reflected in our study by GPT-4's 24-point drop in unit 4 between 2023 and 2022 and DeepSeek's fluctuating scores in unit 3 (cardiovascular and digestive systems), where questions on new pharmacotherapies or revised treatment algorithms likely fell outside the models' training scope [17,18].

Third, the text-based input processing used in this study (to standardize comparisons) inevitably lost critical visual details from medical images (eg, chest x-ray infiltrate density and blood smear leukocyte patterns). Medical licensing exams increasingly rely on multimodal content to mirror real-world practice, and LLMs' inability to directly interpret images reduces their engagement with such questions—observed in the lower accuracy across models in unit 2 (which includes pathological and radiological content, eg, GPT-4's 69.05% accuracy vs 80.67% in unit 4) as text descriptions could not capture subtle visual cues for distinguishing conditions such as pneumonia and pulmonary edema [19].

Implications for Chinese Medical Education

The consistent passing scores of all models (≥360 points) validate their utility as supplementary educational tools, particularly for foundational knowledge review. However, the performance gaps in specialized units reveal critical insights: GPT-4's 80.67% accuracy in unit 4 contrasts with Doubao's 59.87% average in the same unit, highlighting disparities in handling nuanced clinical scenarios. For Chinese medical curricula, DeepSeek and ERNIE Bot's proficiency in Chinese medical ethics and regional clinical guidelines (eg, unit 1's medical regulations) demonstrates their cultural and contextual relevance, making them more suitable for training students in China's health care framework. This aligns with prior research emphasizing the importance of language-specific models in accurately interpreting culturally embedded medical practices.

Limitations and Pathways for Advancement

This study's focus on text-based questions—with all images and tables converted to text—represents a notable constraint as real-world medical practice frequently requires multimodal interpretation. Additionally, the exclusion of the 2023 to 2024 clinical guidelines (due to dataset limitations) may have impacted the models' responses to emerging medical protocols. Future research should prioritize the following:

1. Multimodal model development—integrating computer vision algorithms to process radiology images or electrocardiograms, as seen in experimental medical AI systems
2. Dynamic knowledge integration—implementing real-time updates from sources such as the Chinese Medical Association's guidelines to address timeliness gaps (discrepancies between the latest medical knowledge and guidelines and the information covered in model training data)
3. Human-LLM collaborative models—designing hybrid systems in which LLMs assist in diagnostic reasoning while human experts oversee complex decisions, reducing reliance on autonomous AI in high-stakes scenarios.

Alignment of LLM Performance With Real-World and Literature Trends and Underlying Machine Learning Process Lacunae

The performance percentages of all tested LLMs align with trends reported in existing literature on medical AI and licensing exam assessments while also reflecting gaps consistent with documented limitations in machine learning (ML) processes for clinical tasks. Worldwide, studies on LLMs in medical licensing exams—such as evaluations on the US Medical Licensing Examination, Korean parasitology examinations, and prior NMLE analyses—show that even state-of-the-art models rarely exceed 85% accuracy, with most falling in the 60% to 80% range; this matches our findings, where GPT-4 (highest average accuracy=80.67%) and DeepSeek (73.2%) fit within this established spectrum. For instance, similar to our observation that GPT-3.5's accuracy (68.5%) occasionally fell below the NMLE passing threshold, prior research on non–English-language medical exams has noted lower performance in models with less domain-specific fine-tuning, confirming consistency between our results and real-world model capabilities.

Notable lacunae in the ML process underlying these performance constraints include 3 key areas. The first is limited domain-specific fine-tuning for localized medical contexts.

XSL•FO
RenderX

Most LLMs (including Chinese models) are pretrained on general medical corpora but lack targeted fine-tuning on region-specific clinical practices, guidelines, and NMLE-style question logic. For example, the NMLE emphasizes Chinese national medical regulations, regional disease prevalence (eg, specific infectious disease management protocols), and culturally aligned clinical decision-making—areas in which general ML training fails to deepen model understanding, leading to lower accuracy in unit 1 (which includes medical ethics and regulations) for models such as Doubao (average accuracy below 70%). The second key area is inadequate handling of low-resource clinical cases in training data. ML processes for LLMs rely heavily on large-scale, high-quality labeled data, but rare diseases, comorbid cases, and atypical symptom presentations (all critical to NMLE questions) are underrepresented in training datasets. This led to inconsistent performance in unit 3 (complex case analysis), where models such as GPT-3.5 (56.86% accuracy) frequently misdiagnosed rare comorbidities due to lack of training exposure. The third key area is static knowledge integration and absence of real-time learning. ML pipelines for current LLMs use fixed training data cutoffs (eg, pre-2023 data for many models), with no mechanism for real-time updates to incorporate new clinical guidelines, pharmacotherapies, or diagnostic criteria. This gap is reflected in our study's 2023 to 2024 score fluctuations (eg, GPT-4's 24-point drop in unit 4) as the models could not adapt to posttraining changes in medical knowledge—a limitation also highlighted in broader ML research on clinical AI.

Broader Context and Ethical Considerations

The results also raise ethical questions about LLMs in medical assessment. While models such as GPT-4 and DeepSeek demonstrate impressive recall and reasoning, their lack of clinical experience may lead to oversights in rare cases or nuanced patient interactions. For instance, GPT-4's occasional overconfidence in ambiguous questions (eg, 2023 unit 3 errors) highlights the risk of algorithmic bias. Therefore, educators must approach LLMs as diagnostic aids rather than substitutes for human judgment, emphasizing their role in enhancing efficiency rather than replacing professional expertise.

Conclusions

This study provides the first comprehensive quantitative evaluation of LLMs in the NMLE, demonstrating that GPT-4 and Chinese models such as DeepSeek and ERNIE Bot can achieve reliable performance in structured medical assessments. GPT-4's global generality and DeepSeek's localized excellence each offer unique value: the former excels in cross-cultural medical reasoning, whereas the latter thrives in Chinese-specific medical contexts. However, all models face challenges in dynamic knowledge updating, multimodal processing, and contextual nuance—areas requiring urgent research attention. As AI continues to permeate health care education, these findings advocate for a hybrid approach: leveraging LLMs for scalable knowledge dissemination and formative assessment while preserving human-led training in clinical judgment and ethical decision-making. Ultimately, the integration of these technologies should enhance, not replace, the humanistic core of medical practice.

Funding

This research was funded by the Natural Science Foundation of Anhui Xinhua University (2023zr013 and 2024zr016), the Anhui Provincial University Students Innovation and Entrepreneurship Program (S202312216119, S202312216029, S202412216215, and S202412216192), the University Natural Science Research Project of Anhui Province (2024AH050621), and the Anhui Province Quality Engineering Project (2022xsxx089).

Authors' Contributions

YP and YD studied the literature and conceived this study. MW and JX participated in the formulation of the study implementation plan (including the design of National Medical Licensing Examination question selection protocols, model testing procedures, and evaluation criteria for response accuracy or comprehensiveness) and data analysis. YP and YD wrote the initial draft of the manuscript and serve as guarantors. All authors reviewed and edited the manuscript and approved the final version.

Conflicts of Interest

None declared.

References

1. Juhn Y, Liu H. Artificial intelligence approaches using natural language processing to advance EHR-based clinical research. J Allergy Clin Immunol 2020 Feb;145(2):463-469. [doi: 10.1016/j.jaci.2019.12.897] [Medline: 31883846]
2. Rudolph J, Tan S, Tan S. ChatGPT: bullshit spewer or the end of traditional assessments in higher education? J Appl Learn Teach 2023;6(1):342-363 [FREE Full text] [doi: 10.37074/jalt.2023.6.1.9]
3. Cotton DRE, Cotton PA, Shipway JR. Chatting and cheating: ensuring academic integrity in the era of ChatGPT. Innov Educ Teach Int 2024 Mar 3;61(2):228-239. [doi: 10.1080/14703297.2023.2190148]
4. Tlili A, Shehata B, Adarkwah MA, et al. What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. Smart Learn Environ 2023;10(1):15. [doi: 10.1186/s40561-023-00237-x]

XSL•FO
RenderX

5. Zhou C, Li Q, Li C, et al. A comprehensive survey on pretrained foundation models: a history from BERT to ChatGPT. arXiv. Preprint posted online on Feb 18, 2023. [doi: 10.48550/arXiv.2302.09419]
6. Wang FY, Li J, Qin R, Zhu J, Mo H, Hu B. ChatGPT for computational social systems: from conversational applications to human-oriented operating systems. IEEE Trans Comput Soc Syst 2023;10(2):414-425. [doi: 10.1109/TCSS.2023.3252679]
7. Miao Q, Zheng W, Lv Y, Huang M, Ding W, Wang FY. DAO to HANOI via DeSci: AI paradigm shifts from AlphaGo to ChatGPT. IEEE/CAA J Autom Sinica 2023;10(4):877-897. [doi: 10.1109/JAS.2023.123561]
8. Gilson A, Safranek CW, Huang T, et al. How does CHATGPT perform on the United States Medical Licensing Examination (USMLE)? The implications of large language models for medical education and knowledge assessment. JMIR Med Educ 2023 Feb 8;9(1):e45312. [doi: 10.2196/45312] [Medline: 36753318]
9. Huh S. Are ChatGPT's knowledge and interpretation ability comparable to those of medical students in Korea for taking a parasitology examination?: a descriptive study. J Educ Eval Health Prof 2023;20(1):1. [doi: 10.3352/jeehp.2023.20.1] [Medline: 36627845]
10. Wang X, Gong Z, Wang G, et al. ChatGPT performs on the Chinese National Medical Licensing examination. J Med Syst 2023 Aug 15;47(1):86. [doi: 10.1007/s10916-023-01961-0] [Medline: 37581690]
11. Sun Y, Wang S, Li Y, et al. Ernie: enhanced representation through knowledge integration. arXiv. Preprint posted online on Apr 19, 2019. [doi: 10.48550/arXiv.1904.09223]
12. Tang F, Duan Y. The capability evaluation of GPT-4 and Tongyi Qianwen on financial domain with DIKWP analysis. ReserchGate. Preprint posted online on Apr, 2023. [doi: 10.13140/RG.2.2.12604.92804]
13. Drain PK, Primack A, Hunt DD, Fawzi WW, Holmes KK, Gardner P. Global health in medical education: a call for more training and opportunities. Acad Med 2007 Mar;82(3):226-230. [doi: 10.1097/ACM.0b013e3180305cf9] [Medline: 17327707]
14. De Champlain AF. A primer on classical test theory and item response theory for assessments in medical education. Med Educ 2010 Jan;44(1):109-117. [doi: 10.1111/j.1365-2923.2009.03425.x] [Medline: 20078762]
15. Singhal K, Azizi S, Tu T, et al. Large language models encode clinical knowledge. Nature New Biol 2023 Aug;620(7972):172-180. [doi: 10.1038/s41586-023-06291-2] [Medline: 37438534]
16. Kung TH, Cheatham M, Medenilla A, et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. PLOS Digit Health 2023 Feb;2(2):e0000198. [doi: 10.1371/journal.pdig.0000198] [Medline: 36812645]
17. Jin HK, Lee HE, Kim E. Performance of ChatGPT-3.5 and GPT-4 in national licensing examinations for medicine, pharmacy, dentistry, and nursing: a systematic review and meta-analysis. BMC Med Educ 2024 Sep 16;24(1):1013. [doi: 10.1186/s12909-024-05944-8] [Medline: 39285377]
18. Stead WW, Flatley Brennan P. Celebrating Suzanne Bakken, 2023 Morris F. Collen Award winner and pioneer in health equity. J Am Med Inform Assoc 2023 Oct 19;30(11):1760-1761. [doi: 10.1093/jamia/ocad189] [Medline: 37855452]
19. Xuyan H, Meng S, Chengxing S, Haoxuan L, Jianlin Z. Visual-language reasoning large language models for primary care: advancing clinical decision support through multimodal AI. Vis Comput 2025 Oct;41(13):11327-11348. [doi: 10.1007/s00371-025-04109-y]

Abbreviations

AI: artificial intelligence
LLM: large language model
ML: machine learning
NLP: natural language processing
NMLE: National Medical Licensing Examination

Edited by A Bamgboje-Ayodele; submitted 23.05.25; peer-reviewed by MC Schubert, MK Ghanta; revised version received 22.10.25; accepted 17.11.25; published 16.12.25.

Please cite as:
Diao Y, Wu M, Xu J, Pan Y
Multiple Large Language Models' Performance on the Chinese Medical Licensing Examination: Quantitative Comparative Study
JMIR Hum Factors 2025;12:e77978
URL: https://humanfactors.jmir.org/2025/1/e77978
doi:10.2196/77978

© Yanyu Diao, Mengyuan Wu, Jingwen Xu, Yifeng Pan. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 16.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction

XSL•FO
RenderX

in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Corrigenda and Addenda

Correction: Identifying Contextual Factors and Strategies for Practice Facilitation in Primary Care Quality Improvement Using an Informatics-Driven Model: Framework Development and Mixed Methods Case Study

Related Article:

Correction of: https://humanfactors.jmir.org/2022/2/e32174

(JMIR Hum Factors 2025;12:e84390) doi:10.2196/84390

Following the publication of "Identifying Contextual Factors and Strategies for Practice Facilitation in Primary Care Quality Improvement Using an Informatics-Driven Model: Framework Development and Mixed Methods Case Study" [1], concerns were raised regarding the corresponding author Jiancheng Ye's rate of self-citation within the article's reference list. Further investigation into this matter identified that numerous self-citations were added to the manuscript following acceptance of the article without adequate disclosure.

All authors were contacted for a response regarding this matter. Author Theresa Walunas indicated that they were unaware of the addition of references post acceptance. Authors Donna Woods, Jennifer Bannon, Lucy Bilaver, Gayle Kricke, Megan McHugh, and Abel Kho did not respond. Corresponding author Jiancheng Ye responded with a scientific justification for each added self-citation within the published article.

Following review of these responses, the JMIR Publications Editorial Office determined that several of the added references were not relevant to the article or were used to support a generic statement where another reference could have been used instead of self-citation. Therefore, the JMIR Publications Editorial Office is issuing this corrigendum to remove the following references:

- Reference 17: Ye J. The role of health technology and informatics in a global public health emergency: practices and implications from the COVID-19 pandemic. JMIR Med Inform 2020 Jul 14;8(7):e19866 [doi: 10.2196/19866] [Medline: 32568725]
- Reference 19: Ye J, Li N, Lu Y, Cheng J, Xu Y. A portable urine analyzer based on colorimetric detection. Anal Methods 2017;9(16):2464-2471 [doi: 10.1039/c7ay00780a]
- Reference 38: Ye J, Orji I, Baldridge AS, Ojo T, Shedul G, Ugwuneji E, et al. Abstract MP19: characteristics and patterns of retention in hypertension care in primary care settings: findings from the Hypertension Treatment in Nigeria Program. Circulation 2022;145 [doi: 10.1161/circ.145.suppl_1.mp19]
- Reference 40: Ye J. The impact of electronic health record-integrated patient-generated health data on clinician burnout. J Am Med Inform Assoc 2021 Apr 23;28(5):1051-1056 [doi: 10.1093/jamia/ocab017] [Medline: 33822095]
- Reference 42: Ye J, Ren Z. Using multivariate models to examine the impact of COVID-19 pandemic and gender differences on health and health care. medRxiv. Posted on September 5, 2021 [doi: 10.1101/2021.09.02.21263055]
- Reference 44: Ye J. Health information system's responses to COVID-19 pandemic in China: a national cross-sectional study. Appl Clin Inform 2021 Mar 19;12(2):399-406 [doi: 10.1055/s-0041-1728770] [Medline: 34010976]
- Reference 45: Ye J, Yao L, Shen J, Janarthanam R, Luo Y. Predicting mortality in critically ill patients with diabetes using machine learning and clinical notes. BMC Med Inform Decis Mak 2020 Dec 30;20(Suppl 11):295 [doi: 10.1186/s12911-020-01318-4] [Medline: 33380338]
- Reference 46: Ye J, Wang Z, Hai J. Social networking service, patient-generated health data, and population health informatics: national cross-sectional study of patterns and implications of leveraging digital technologies to support mental health and well-being. J Med Internet Res 2022 Apr 29;24(4):e30898 [doi: 10.2196/30898] [Medline: 35486428]
- Reference 54: Ye J, Ma Q. The effects and patterns among mobile health, social determinants, and physical activity: a nationally representative cross-sectional study. AMIA Jt Summits Transl Sci Proc 2021;2021:653-662 [Medline: 34457181]
- Reference 59: Ye J. Pediatric mental and behavioral health in the period of quarantine and social distancing with COVID-19. JMIR Pediatr Parent 2020 Jul 28;3(2):e19867 [doi: 10.2196/19867] [Medline: 32634105]

The JMIR Publications Editorial Office regrets that these issues were not identified prior to publication.

The correction will appear in the online version of the paper on the JMIR Publications website together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

XSL•FO
RenderX

Reference

1. Ye J, Woods D, Bannon J, Bilaver L, Kricke G, McHugh M, et al. Identifying contextual factors and strategies for practice facilitation in primary care quality improvement using an informatics-driven model: framework development and mixed methods case study. JMIR Hum Factors 2022 Jun 24;9(2):e32174 [FREE Full text] [doi: 10.2196/32174] [Medline: 35749211]

Submitted 18.09.25; this is a non–peer-reviewed article;accepted 18.09.25; published 08.10.25.

Please cite as:
Correction: Identifying Contextual Factors and Strategies for Practice Facilitation in Primary Care Quality Improvement Using an Informatics-Driven Model: Framework Development and Mixed Methods Case Study
JMIR Hum Factors 2025;12:e84390
URL: https://humanfactors.jmir.org/2025/1/e84390
doi:10.2196/84390
PMID:

©JMIR Publications Editorial Office. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 08.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Practitioner Perspectives on the Uses of Generative AI Chatbots in Mental Health Care: Mixed Methods Study

Jessie Goldie[1], LLB, BA, GDipPsych; Simon Dennis[1], PhD; Lyndsey Hipgrave[1], BSocSci, BSW(Hons); Amanda Coleman[1,2], BS, GradDipProPsych

[1]Melbourne School of Psychological Sciences, Medicine, Dentistry and Health Sciences, University of Melbourne, Grattan Street, Parkville, Melbourne, Australia

[2]Monash School of Psychological Sciences, Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, Australia

Corresponding Author:
Simon Dennis, PhD
Melbourne School of Psychological Sciences, Medicine, Dentistry and Health Sciences, University of Melbourne, Grattan Street, Parkville, Melbourne, Australia

Abstract

Background: Generative artificial intelligence (AI) chatbots have the potential to improve mental health care for practitioners and clients. Evidence demonstrates that AI chatbots can assist with tasks such as documentation, research, counseling, and therapeutic exercises. However, research examining practitioners' perspectives is limited.

Objective: This mixed-methods study investigates: (1) practitioners' perspectives on different uses of generative AI chatbots; (2) their likelihood of recommending chatbots to clients; and (3) whether recommendation likelihood increases after viewing a demonstration.

Methods: Participants were 23 mental health practitioners, including 17 females and 6 males, with a mean age of 39.39 (SD 16.20) years. In 45-minute interviews, participants selected their 3 most helpful uses of chatbots from 11 options and rated their likelihood of recommending chatbots to clients on a Likert scale before and after an 11-minute chatbot demonstration.

Results: Binomial tests found that Generating case notes was selected at greater-than-chance levels (15/23, 65%; *P*=.001), while Support with session planning (*P*=.86) and Identifying and suggesting literature (*P*=.10) were not. Although 55% (12/23) were likely to recommend chatbots to clients, a binomial test found no significant difference from the 50% threshold (*P*=.74). A paired samples *t* test found that recommendation likelihood increased significantly (19/23, 83%; *P*=.002) from predemonstration to postdemonstration.

Conclusions: Findings suggest practitioners favor administrative uses of generative AI and are more likely to recommend chatbots to clients after exposure. This study highlights a need for practitioner education and guidelines to support safe and effective AI integration in mental health care.

(JMIR Hum Factors 2025;12:e71065) doi:10.2196/71065

KEYWORDS

artificial intelligence; ChatGPT; mental health care; practitioner perspectives; mixed methods; digital health

Introduction

Overview

Globally, the prevalence of mental illness is rising and health care systems are under increasing pressure [1]. In Australia, individuals face barriers to mental health care, including financial costs and extended wait times for appointments, while practitioners are struggling to meet demand [2,3]. In this context, artificial intelligence (AI) is a potential game changer, offering 24/7 support for individuals, thus alleviating pressure on mental health practitioners [4,5]. Current models engage users in therapeutic conversations, with many using "rule-based" scripted responses and predefined dialog structures [6]. This technology differs from generative AI chatbots, such as ChatGPT (OpenAI), which generate original content and provide lifelike interactions that closely imitate humans.

Generative AI can offer more empathetic emotional support in comparison to rule-based chatbots [4,7-9], making it more effective in counseling conversations and more acceptable to clients. A recent qualitative study exploring the efficacy of ChatGPT in supporting 24 individuals with mental health conditions found that more than 50% of participants found ChatGPT to be empathetic and 80% said the tool helped to manage symptoms [10]. However, participants also raised concerns regarding data privacy and the risk of personal data theft, uncertainty in the reliability of the information shared and the quality of training data, and a lack of understanding of cultural determinants of mental health.

XSL•FO
RenderX

Generative AI chatbots may also provide improved access to mental health information and psychoeducation. Preprint and peer-reviewed studies assessing the quality of psychoeducational information generated by ChatGPT on depression, anxiety, psychosis, and substance abuse found that it was accurate, clear, clinically useful, relevant, and empathetic [11-13]. Furthermore, studies have found generative AI chatbots to be effective in supporting adherence to psychological treatments [14] and delivering positive psychology interventions [15,16]. These studies indicate that generative AI chatbots may improve individuals' access to information and basic therapeutic exercises, allowing clinicians to focus on more complex tasks, such as diagnosis and treatment.

Studies have identified a variety of uses of AI for practitioners, including research, documentation, administering psychological assessments, clinic management, and triaging [5,17-19]. Generative AI could enhance these functions by offering faster processing speeds, more comprehensive client histories, and higher empathy [20,21]. However, like rule-based AI models, studies have found generative AI chatbots are prone to mistakes, with one study finding that ChatGPT included incorrect information in 36% of client documents in a medical setting [21]. Another study evaluated ChatGPT's performance in conducting mental health assessments and found that, while responses for simple cases were acceptable, recommendations for more complex cases were inappropriate and potentially harmful [22]. These findings highlight the risks associated with AI, including hallucinations and misinterpretations.

Generative AI may also be capable of more sophisticated functions, such as recommending diagnoses and treatment options. Bartal et al [23] found that ChatGPT was able to accurately identify posttraumatic stress disorder (PTSD) following childbirth by analyzing individuals' birth stories, achieving high sensitivity (85%) and specificity (75%), aligning with commonly accepted benchmarks of reliability [24,25]. Similarly, another case study demonstrated that ChatGPT can identify treatment-resistant schizophrenia and suggest appropriate treatment options [26]. Despite these innovative applications, some argue that there is a need for significant improvement before AI is adopted in mental health care [27].

Clinicians appear cautious about the use of AI and feel more education and training are required [28,29]. In a systematic review of AI in mental health care, Rogan et al [28] found that most studies highlighted a need for clinician training, emphasizing that "AI will only be useful if clinicians understand and feel comfortable using it". Similarly, in a survey of 138 American psychiatrists, Blease et al [29] found that 90% of participants agreed that clinicians need more support to understand generative AI. Qualitative responses were mixed, with some participants expressing uncertainty and concerns relating to potential harm and others overestimating the readiness of generative AI to deliver clinical tasks.

While practitioners have expressed a need for more training and support, many may already be using AI tools for internal purposes. Blease et al [29] found that more than half of psychiatrists surveyed had used generative AI chatbots to answer clinical questions. A survey of 86 Australian mental health professionals found that 43% had used AI and that ChatGPT was the most commonly used tool [19]. Likewise, a systematic review preprint on the use of generative AI in health care found that the technology is "heavily applied" for documentation, research, and administration among medical practitioners [30].

Research examining practitioners' use of generative AI chatbots with clients is more limited. One study found strong support among clinicians to use chatbots with clients, with 80% of participants reporting they would be either very likely (24%) or somewhat likely (56%) to recommend AI mental health chatbots to clients within the next 5 years [31]. Participants may have been more supportive as they were asked to consider their future likelihood and, therefore, the potential for evolution of both the technology and their understanding.

Objective

This study aims to explore which functions of generative AI chatbots mental health practitioners think are most useful, and to what extent they would recommend them to clients. In addition, this study will investigate the impact of exposure to a generative AI chatbot on participants' likelihood of recommending chatbots to clients.

This study uses both quantitative and qualitative data to address 3 research questions (RQ) aligned with corresponding hypotheses. These RQs address key gaps in the literature, including limited insight into which functions practitioners prefer, a lack of empirical data on their willingness to recommend generative AI chatbots to clients, and minimal understanding of how exposure to the technology might influence those attitudes.

RQ1 explores practitioners' views on the most useful functions of generative AI chatbots in mental health care. It is hypothesized that functions such as generating case notes (hypothesis 1), supporting session planning (hypothesis 2), and identifying literature (hypothesis 3) will be prioritized. Previous research suggests clinicians favor using AI for internal tasks like documentation and research [29,30].

RQ2 examines the likelihood of mental health practitioners recommending a generative AI chatbot to clients. It is hypothesized (hypothesis 4) that fewer than 50% will do so, as practitioners typically have low technology literacy [28,29,32].

RQ3 investigates whether practitioners are more likely to recommend an AI chatbot after a demonstration, with the prediction (hypothesis 5) that exposure to a demo will significantly increase their likelihood to recommend it. Research indicates that familiarity with health care technologies improves clinician adoption [33].

Method

Participants

Participants were recruited using a purposive, convenience sampling method. Digital flyers were distributed on forums including the Australian Psychological Society (APS), Australian Association of Psychologists Inc (AAPi), through Melbourne University channels, and with friends and family.

XSL•FO
RenderX

Individuals who were not working in a client-facing, mental health role were excluded from the study.

In total, 23 individuals participated in the study, and all were working in Australia. Sample demographics are presented in Table 1.

Table . Sample demographic information (N=23).

Demographic variables	Participants (N=23)
Gender, n (%)	
Men	6 (26)
Women	17 (74)
Age (years), mean (SD)	39.39 (16.20)
Clinical role, n (%)	
Psychologist (including provisional)	10 (30)
Counselor or psychotherapist	6 (26)
Mental health social worker	4 (17)
Psychiatrist	2 (9)
Mental health support worker	1 (4)
Ethnicity, n (%)	
Caucasian or European	15 (65)
Asian	5 (22)
Mixed ethnicity	3 (13)
Area of work, n (%)	
Private sector	8 (35)
Public sector	9 (39)
Mixed	6 (26)
Years of work experience, mean (SD)	10.78 (14.31)

Procedure

The study team devised interview questions (Multimedia Appendix 1) to explore practitioners' perspectives on generative AI chatbots in mental health care. The interviews were of 45 minutes and conducted over Zoom (Zoom Video Communications, Inc.) between May and August 2024. Open-ended, Likert-scale, and ranking questions were used to collect both qualitative and quantitative data. Interviews were recorded and transcribed using AI software, Notta (Notta Inc), to support qualitative analysis.

Uses of Generative AI Chatbots in Mental Health Care

Participants were asked to select their top 3 "most useful" functions of generative AI chatbots from a list of 11 options and explain their reasoning. The uses were:

1. Generating case notes
2. Support with session planning
3. Client triaging
4. Client onboarding processes
5. Administering psychological assessments
6. Supporting therapist-directed exercises
7. Supporting self-directed exercises
8. Symptom tracking and monitoring
9. Identifying and suggesting literature relevant to a client's profile
10. Psychoeducation and socialization to therapeutic models
11. Counseling

An option to select "Other" was offered and subsequently removed from the analysis as it was not selected by any participants.

Likelihood of Recommending a Chatbot to a Client

Participants were asked, "How likely are you to recommend a chatbot to a client?" on a Likert scale of 1 (Highly unlikely) to 4 (Highly likely). This question was asked both before (T1) and after (T2) viewing a demonstration of a generative AI chatbot.

The demonstration was an 11-minute screen recording of a generative AI chatbot that had been built using GPT-4o (OpenAI) for the purposes of this research. During the demonstration, a fictional client ("Saman") interacted with the chatbot to illustrate a range of use cases, from triaging, psychological assessment, clinic onboarding, counseling, and completing a gratitude exercise. The demonstration also showed Saman's fictional therapist using the chatbot to generate case notes, identify literature relevant to Saman's profile, and plan future sessions (Multimedia Appendix 2).

Although the demonstration featured a chatbot built using GPT-4o, participants were asked to rate and reflect on generative AI chatbots for mental health support more generally, rather than the specific tool presented. Only the text-based

XSL•FO
RenderX

conversational capacity of GPT-4o was used, with safety filters (mechanisms that monitor and restrict the model's output to reduce risk of harm), and real-time context (allowing the chatbot to access external data and information) active by default. These features were not explicitly evaluated but were considered reflective of current generative AI chatbot capabilities.

Data Analysis

Quantitative

RQ1: Uses of Generative AI Chatbots

Data were analyzed to determine the proportion of participants who selected each use and plot the SE. To test statistical hypotheses (hypothesis 1, hypothesis 2, and hypothesis 3), 3 binomial tests were conducted to assess whether specific uses were selected at greater than chance levels.

RQ2: Likelihood of Recommending a Chatbot to a Client

A binomial test was conducted to test hypothesis 4 and analyze whether less than 50% of participants would be likely to recommend a chatbot. A Bayesian binomial test was conducted to evaluate the strength of the evidence in favor of the alternative hypothesis.

RQ3: Increase in Likelihood of Recommending a Chatbot After a Demonstration

To test hypothesis 5, a paired-samples *t* test was conducted to measure whether there was a significant difference in the mean likelihood of recommending a chatbot before (T1) and after (T2) the demonstration. A Bayesian paired samples *t* test assessed the strength of the evidence in favor of the alternative hypothesis (a significant increase in recommendation likelihood from T1 to T2).

Qualitative

A thematic analysis was conducted following Braun and Clarke's [34] 6-phase framework. Analysis began with familiarization, where transcripts were reviewed to identify initial ideas. Themes were then generated based on recurring concepts, iterated, and refined to ensure they captured core insights. Finally, qualitative themes were considered alongside the quantitative results to identify key research findings. Where participants are quoted, a reference is provided with their participant identification number (#), gender (men or women), and age.

Ethical Considerations

All participants provided informed consent and no compensation was offered. Participants were informed in advance that participation was voluntary, and they could withdraw at any time without consequence. Data were stored securely, and participant confidentiality was secured by anonymizing all data. Ethics approval was obtained from the University of Melbourne Human Research Ethics Committee (#28479).

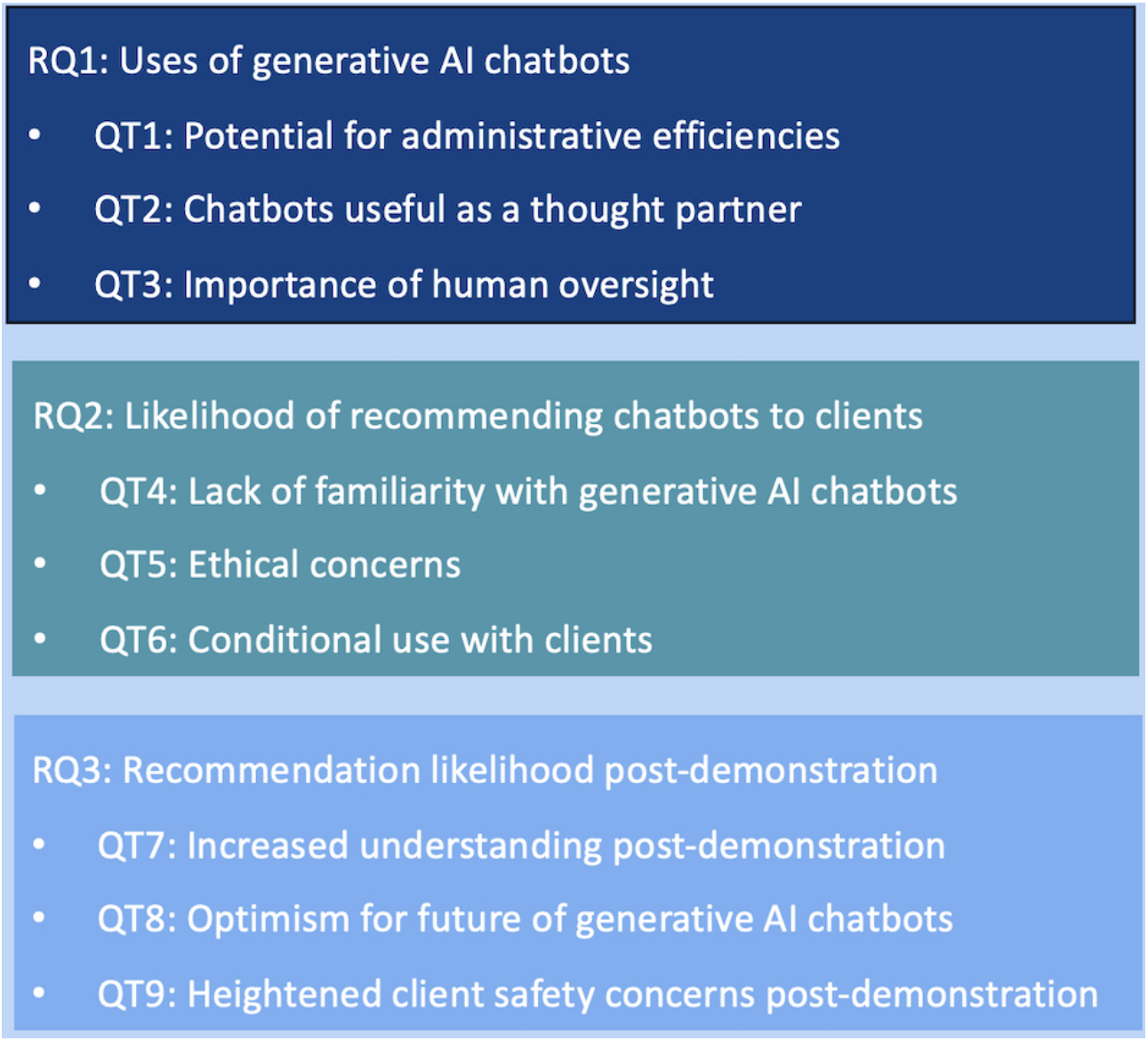
Results

Overview of Qualitative Themes

Thematic analysis produced 9 qualitative themes (QTs), each aligned to a particular RQ, as shown in Figure 1. These themes are reported in further detail throughout the following section.

XSL•FO
RenderX

Figure 1. Summary of qualitative themes.



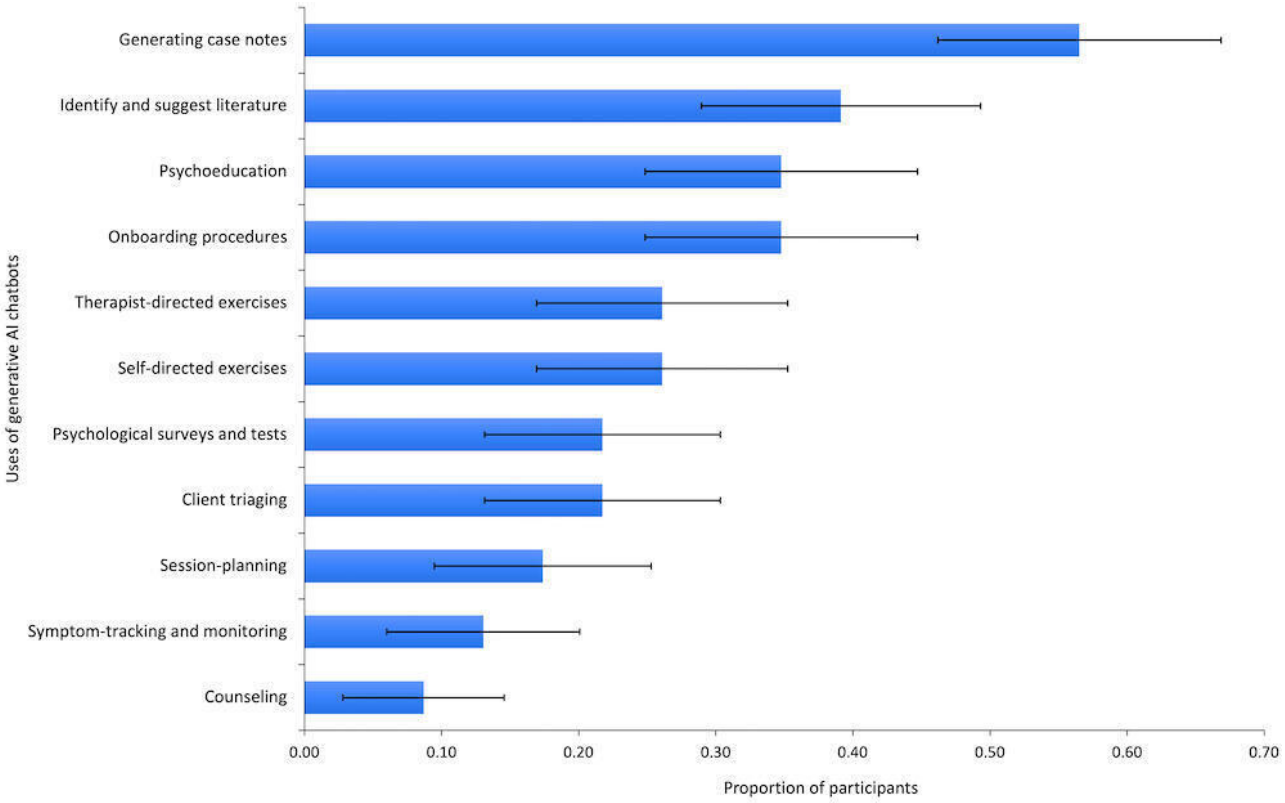
RQ1: Uses of Generative AI Chatbots

Proportions of participants who selected each use in their top 3 are presented in Figure 2.

A binomial test was conducted to assess whether Generating case notes was selected in participants' top 3 choices at greater than chance levels. The chance probability was calculated at 0.25, based on the probability of a participant selecting an option, acknowledging that they had 3 choices and could not select the same option twice (P(selected)=1 - [1- 1/11] x [1 - 1/10] x [1- 1/9]). The results were significant, P=.001, indicating that Generating case notes was selected at greater than chance levels. A binomial test indicated that Support with session planning was not selected in participants' top 3 choices at greater than chance levels (0.25; P=.86). A binomial test indicated that Identifying literature and Suggesting literature was not selected in participants' top 3 choices at greater than chance levels (0.25; P=.10).

XSL•FO
RenderX

Figure 2. Proportion of participants selecting each use of generative AI chatbots. Error bars represent the SE of the proportion of participants selecting each use.



QT1: Potential for Administrative Efficiencies

Qualitative analysis indicated that participants value the potential of generative AI chatbots to reduce time spent writing case notes. One participant explained:

> *Writing clinical notes, that takes up probably more time in therapists' days than they would like... having a process to be able to. dump your very verbatim notes from a session and have it spit out a very succinct summary of the mental state formulation. That would be awesome* [#20, F, 29.]

Another participant emphasized that using a chatbot to document case notes would enable them to spend more time with clients,

> *Generating case notes strikes me as just a time-consuming and aversive activity… you save yourself time to spend more time with clients and less time on admin* [#4, M, 42.]

QT2: Chatbots Useful as a Thought Partner

Qualitative responses indicated that participants saw the potential of generative AI chatbots to assist with session planning by providing new ideas, though it wasn't seen as a necessity. One participant stated,

> *For help with session planning I feel like well it's good with just reflecting on ideas and you know talk about things that you maybe haven't thought about* [#2, F, 27.]

Another explained,

> *I mean, you could do it without it, but it's always helpful to just to maybe get another perspective or consider something else* [#17, F, 35.]

QT3: Importance of Human Oversight

While qualitative responses indicated that participants hoped generative AI chatbots might enhance their research, many stressed the importance of maintaining human oversight. One participant explained:

> *[Identifying literature] is something that we do digitally anyway and if we're able to customize a search to a client specific profile, that might make the search better. And then obviously, the clinician can still have the final say on that, but it would be helpful* [#9, F, 24.]

Several participants raised concerns around accuracy, with one participant remarking that,

> *It would be great to have a reliable source that can give us the latest research in what is best practice CBT [Cognitive Behavioral Therapy] for OCD [Obsessive Compulsive Disorder]... [ChatGPT] just never seems quite accurate* [#11, F, 25.]

Another explained,

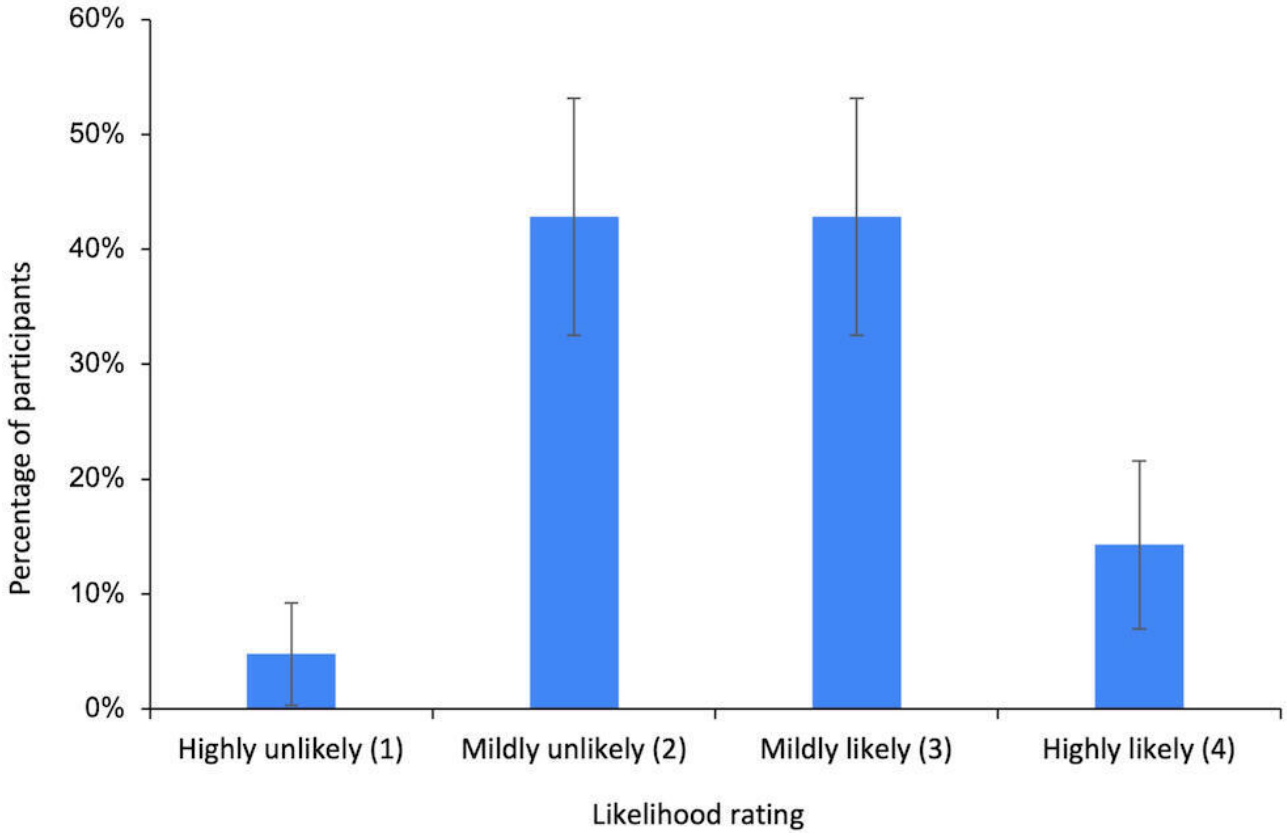
> *I know that there's been a lot of speculation around finding accurate references… I would always, you know, [use ChatGPT] in addition to my own search as well, just to be sure* [#2, F, 27.]

XSL•FO
RenderX

RQ2: Likelihood of Recommending a Chatbot to a Client

The distribution of participants' likelihood of recommending a chatbot to a client at T2 is presented in Figure 3. Data from T2 (postdemonstration) was used for the analysis, as it best reflected participants' informed perspectives of generative AI chatbots. One participant responded, "I don't know," and their data were excluded from the quantitative analysis.

Figure 3. Distribution of participants' likelihood of recommending a generative AI chatbot (n = 22). Error bars represent the SE of the percentage of participants selecting each rating.



A binomial test was conducted to determine whether participants would be likely to recommend a generative AI chatbot to a client (defined as a response of 3 or 4 on the 4-point Likert scale). The proportion of participants likely to recommend was 55% (12/22), which was not significantly different from the 50% threshold (P=.74; odds ratio 1.20 [95% CI 0.47-3.17]). A Bayesian binomial test provided moderate evidence for the null hypothesis that more than 50% of participants would be likely to recommend a chatbot (BF_{01}=5.24).

QT4: Lack of Familiarity With Generative AI Chatbots

Most participants were not familiar with generative AI chatbots in mental health care. One participant noted,

> [I don't feel like I understand the market well enough yet to recommend it to clients. I would need more experience and a chance to evaluate different models before trusting them for client use] [#4, M, 42.]

Others explained that they simply "don't know enough about it". During introductory questions, only one participant reported having experience using an AI chatbot for mental health support for themselves or with a client.

QT5: Ethical Concerns

Ethical concerns were frequently cited. One participant explained,

> *I don't think that it would be prudent or appropriate to recommend [chatbots] to clients just the same way that we wouldn't recommend unapproved treatments that we can't be fully confident in to clients because then we basically have liability* [#9, F, 24.]

Another commented,

> *Privacy would be my main concern. How do we manage to keep client information confidential if we're using it for note purposes?* [#20, F, 29.]

QT6: Conditional Use With Clients

For many participants, their likelihood of recommending a chatbot was dependent on the type of client. One participant stated,

> *I do see a use for it as an emergency point of contact, but for anything beyond that, especially involving complex risk, I wouldn't feel comfortable recommending it* [#9, F, 24.]

Another explained that their etiological formulation would inform whether a chatbot would be appropriate:

> *A lot of people's addictions and mood disorders are born out of childhood trauma, and I think there is also a certain percentage of those who are*

XSL•FO
RenderX

> *enduringattachment wounds. And that's where AI is not going to be able to help* [#12, F, 66.]

Others suggested their likelihood of recommending would be based on age, with one participant explaining,

> *It depends on the demographic. Maybe if it was a young person, I would. But in my practice, I don't actually come across many young people. So personally, probably no, because most of my demographics are older* [#6, F, 54.]

RQ3: Change in Likelihood of Recommending a Chatbot to a Client After the Demonstration

A one-tailed paired samples *t* test was conducted to assess whether there was a statistically significant increase in the likelihood of recommending a chatbot from T1 (mean 2.09 [SD 1.07]) to T2 (mean 2.64 [SD 0.79]). Results were significant, t_{21}=3.20; *P*=.002, with a mean difference of 0.55. The effect size (*d*=0.68) was medium-to-large [35], suggesting a meaningful increase in recommendation likelihood.

A Bayesian paired samples *t* test provided strong evidence in favor of the alternative hypothesis (*r*=0.707; BF_{10}=20.08), indicating that participants were more likely to recommend a chatbot after the demonstration (T2) compared to before (T1).

QT7: Increased Understanding Postdemonstration

Qualitative analysis suggested that the demonstration improved participants' understanding of how generative AI chatbots can be used in mental health care. One participant stated,

> *I have a better sense of the additional value that AI can provide particularly around screening and triage* [#4, M, 42.]

Another stated that, "It gave me new ideas on onboarding and case notes" (#1, F, 25). Some participants were satisfied with how the chatbot performed client-facing tasks such as counseling, with one participant stating,

> *I was quite happy with how [the chatbot] dealt with everything in terms of like when it realized that the issues were quite severe and that it didn't jump to try to diagnose* [#2, F, 27.]

QT8: Optimism for Future of Generative AI Chatbots

After the demonstration, several participants expressed optimism about the potential of generative AI. One participant remarked,

> *I imagine as the technology gets better and we're able to further program them to work in a really safe and diligent way around this sort of stuff, it probably will work quite smoothly and reliably* [#7, F, 27.]

One participant likened the introduction of AI chatbots to the evolution of telephone counseling, while another stated:

> *I know a lot of people firmly believe that AI will never be able to deliver that human touch. And I just think that I just think they're delusional, frankly, I think, I think they need to… realize what's coming* [#4, M, 42]

QT9: Heightened Client Safety Concerns Postdemonstration

Of the 10 participants who reported an increased likelihood of recommending a chatbot, 4 indicated that the demonstration had heightened their concerns regarding client safety. One participant stated,

> [I was spooked by how the AI responded when the person expressed suicidality… it slightly ups my sense of the risk based on current technology] [#5, F, 41.]

Another commented,

> [The user was feeling a bit flat and all he was left with was a phone number, like rather than 'I will put you through to Lifeline right now', for example, just leaving risk like that, that's probably not helpful] [#8, F, 58.]

Discussion

Principal Findings

This study found that practitioners appear to prefer administrative uses of generative AI chatbots, such as generating case notes, and are uncertain about recommending chatbots to clients. While exposure to a demonstration may increase the likelihood of recommending a chatbot, concerns regarding client safety and risk persist.

The first hypothesis (hypothesis 1) was supported. Generating case notes was selected as the most useful function of generative AI chatbots at greater than chance levels. The second (hypothesis 2) and third (hypothesis 3) hypotheses were not supported. Support with session planning and Identifying and suggesting literature were not selected at greater than chance levels. The fourth hypothesis (hypothesis 4), that less than 50% of participants would be likely to recommend a chatbot to a client, was not supported. The fifth hypothesis (hypothesis 5) was supported; participants were significantly more likely to recommend a chatbot to a client postdemonstration, and Bayesian analysis provided strong evidence for this increase.

Practitioners Prefer Administrative Uses of Generative AI Chatbots

The results indicate a preference among practitioners for using generative AI chatbots for simple, administrative tasks. Quantitative and qualitative data highlight that practitioners see value in generative AI chatbots for drafting and summarizing case notes. Participants expressed enthusiasm for the potential administrative efficiencies that using a chatbot would allow more time to focus on client-facing work. This aligns with previous studies where practitioners reported using generative AI tools for documentation and research [19,29,30]. This preference is also evident in the influx of AI-based tools for case note documentation in the market, such as Heidi (Heidi Health), Upheal (Upheal), and Autonotes (Autonotes AI, LLC) [36-38].

There is; however, some nuance in the types of administrative tasks practitioners are willing to use generative AI chatbots for. While descriptive data showed that Identifying and suggesting

XSL•FO
RenderX

literature was the second most frequently selected use, it was not selected at greater than chance levels. Similarly, there was a lack of support for Session planning. Participants emphasized the importance of human oversight, indicating a lack of trust in generative AI. This aligns with findings from Blease et al [29], where psychiatrists were cautious of using generative AI for tasks that require clinical expertise and raised concerns about accuracy. As such, while practitioners may not be willing to delegate more complex administrative tasks to generative AI chatbots, this study suggests that there is potential for the technology to be applied in a supporting capacity.

Practitioners Are Uncertain About Recommending Generative AI Chatbots to Clients

This study found no clear preference for or against recommending a chatbot to clients, with uncertainty evident in both quantitative and qualitative results. Both the frequentist and Bayesian analyses reflected statistical uncertainty regarding the true proportion of participants who were likely to recommend a chatbot to clients. Qualitative analysis highlighted several ethical concerns, ranging from professional liability to confidentiality and data security, and a lack of familiarity with the technology. This aligns with previous research, which found that practitioners often harbor ethical concerns and lack education on the appropriate use of generative AI tools in practice [29,32].

This uncertainty contrasts with Sweeney et al [31], where 80% of practitioners were likely to recommend a chatbot to a client. This variance may be explained by the difference in how questions were phrased. In Sweeney et al [31] study, participants were asked whether they would recommend a chatbot "in the next five years," allowing practitioners to consider how the technology, and their understanding of it, may evolve. In contrast, the present study focused on immediate likelihood, which may have biased participants toward considering current limitations. Participants also expressed optimism about the future of generative AI, suggesting they may have been more supportive if asked to consider their likelihood of recommending a chatbot in the future.

Practitioners' Use of Generative AI Chatbots May Be Conditional on the Client

A unique insight from this study was practitioners' support for conditional use of generative AI chatbots, depending on factors unique to the client. Qualitative responses indicated that participants were more inclined to recommend chatbots for low-risk, less complex cases (ie, stress, sleep issues, and life transitions) compared to high-risk (ie, suicidal intent) or trauma-related cases. The existing literature on the appropriate client profiles for generative AI chatbots presents mixed views. Some propose that AI is appropriate for use with clients with complex clinical histories, on the basis that the technology can analyze and track data on, for example, genetic markers and behavioral patterns [39]. Others suggest that AI chatbots may be most appropriate in prevention and psychoeducation in low-risk populations [40]. Ultimately, like any mental health treatment, the success of generative AI tools will depend not only on the technology, but also on its fit for the client.

Exposure Increases Recommendation Likelihood and May Heighten Client Safety Concerns

This study found that participants' likelihood of recommending a generative AI chatbot to a client increased significantly after a demonstration. Qualitative insights provided depth to the quantitative results, as participants explained that the demonstration had increased their understanding of how chatbots can be applied in mental health care. This result aligns with earlier findings of a positive correlation between health care practitioners' knowledge of generative AI chatbots and positive attitudes toward the technology [41]. Similarly, previous studies have found that negative perceptions of AI among mental health professionals may be due to minimal exposure [29,42].

Alternatively, participants' increased recommendation likelihood postdemonstration may be reflective of novelty bias, where individuals tend to exhibit increased enthusiasm toward a new or unfamiliar technology or idea [43]. Qualitative data suggests that for 22 out of 23 participants (96%), the demonstration was the first time they had seen a generative AI chatbot used in a mental health setting. In a similar study by Liu et al [16], it was observed that participants responded more positively to a chatbot trained to deliver positive psychology interventions during initial interactions.

Surprisingly, despite the quantitative increase in recommendation likelihood, qualitative responses indicated that the demonstration heightened client safety concerns for some participants. Of the ten participants who reported an increased recommendation likelihood, 4 described being more concerned about the risks. This suggests that a greater understanding of the associated risks of using generative AI chatbots may play a role in adoption among practitioners. This aligns with results from Zhang et al [42], where participants described the importance of understanding the risks of AI to "clinician buy-in" and adoption of the technology in practice. While an examination of potential risks was beyond the scope of this study, this finding underscores the importance of providing evidence-based, objective education to practitioners.

Limitations and Future Directions for Research

Several limitations of this study should be addressed. First, convenience sampling may have introduced selection bias, as individuals who were more supportive of AI technologies may have been more likely to participate, leading to an overrepresentation of positive attitudes. Future studies should apply random sampling methods to reduce selection bias, ensure representativeness, and improve external validity [44]. In addition, participants may have been aware that the interviewers created the chatbot demonstration, introducing demand characteristics and potential bias toward more favorable responses. The use of independent researchers to administer questionnaires may assist in mitigating potential bias in future research.

A primary focus of this research was to gather in-depth qualitative data to allow for greater exploration of practitioners' perspectives. Due to project constraints and the time required of participants, the study sample size was small (N=23). While this number is appropriate for qualitative thematic research [45],

XSL•FO
RenderX

it limited the statistical power and reliability of the quantitative analyses [46]. No a priori power calculation was conducted, as the study was primarily exploratory and qualitatively focused. In addition, the sample was Australian, predominantly female (17/23, 74%) and mostly psychologists and counselors (13/23, 56%). Caution should be exercised when generalizing these findings to other health systems or sociocultural contexts. In the future, researchers should recruit larger, more representative samples in different cultural contexts, while still maintaining the qualitative analysis that is crucial for understanding the nuanced perspectives of practitioners.

Participants were asked to rate their likelihood of recommending a chatbot to a client on a 4-point Likert scale. While the absence of a neutral midpoint encouraged participants to share their perspectives, it may have forced practitioners to choose a position that they did not fully support.

This study did not consider recent developments in generative AI chatbot technology, such as voice recognition or more sophisticated emotional detection capabilities [47]. Researchers may consider partnering with AI developers to test emerging technologies to ensure that studies remain relevant, while also allowing for practitioner and user feedback on tools before they are released to market.

Practical Implications

For practitioners, the findings suggest that generative AI chatbots could be valuable in alleviating their administrative burden. However, the study also highlights the need for careful consideration as to how AI is applied and the need for practitioners to improve their understanding of AI chatbots, particularly as their clients may already be using them [22].

Educational institutions and industry associations play an important role in ensuring practitioners have appropriate AI skills and literacy and are prepared to deliver services in a field where technology is playing an increasingly important role [48]. Practical, evidence-based guidelines for the use of AI tools should address concerns expressed in this study regarding liability, client safety, and ethics, while also enabling effective use of technology to support positive mental health outcomes. Practitioners should also be informed of emerging research so as to help validate their decision-making and points of view.

For AI developers, insights from this study may support iteration and refinement of generative AI chatbots for mental health care. For example, developers should ensure practitioners' concerns regarding client safety, data security, and accuracy are addressed in AI tools and service agreements. This study also highlights the need for developers to engage practitioners, as well as users, throughout the design process to support safe and effective technology implementation.

Conclusion

Generative AI chatbots have significant potential to play an important role in the delivery of mental health care. In the immediate term, their greatest value may be in administrative, nonclient-facing tasks, to alleviate practitioner workload and allow greater attention to be directed to therapeutic work. This study highlighted the complexity and nuance of practitioners' perspectives on generative AI chatbots. While there is interest and optimism, there is a lack of familiarity with AI tools and uncertainty regarding client-facing uses. Furthermore, exposure to AI technology may play an important role in supporting adoption among practitioners.

Future research should continue to explore practitioners' views on administrative and client-facing uses with larger, more representative samples to enhance the reliability and generalizability of findings. It would also be beneficial to explore hybrid models and the benefits of therapist-AI collaboration. Such evidence could inform objective guidelines to allow practitioners to make informed decisions on their use of generative AI chatbots. Ultimately, this study provides valuable contributions as to "how," rather than simply "whether," generative AI chatbots may be integrated into mental health care, guiding more thoughtful, stakeholder-informed implementation.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview questions and interviewer script.
[DOCX File, 19 KB - humanfactors_v12i1e71065_app1.docx]

Multimedia Appendix 2
Demonstration transcript and link.
[DOCX File, 27 KB - humanfactors_v12i1e71065_app2.docx]

References

1. Kupcova I, Danisovic L, Klein M, Harsanyi S. Effects of the COVID-19 pandemic on mental health, anxiety, and depression. BMC Psychol 2023;11(1). [doi: 10.1186/s40359-023-01130-5]
2. Mental health services. Australian Institute of Health and Welfare. 2024 Apr 30. URL: https://www.aihw.gov.au/mental-health/overview/mental-health-services [accessed 2024-10-05]

XSL•FO
RenderX

3. Mental Health Australia. Report to the nation 2023. URL: https://mhaustralia.org/sites/default/files/docs/report_to_the_nation_2023.pdf [accessed 2024-05-15]
4. Fitzpatrick KK, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. JMIR Ment Health 2017 Jun 6;4(2):e19. [doi: 10.2196/mental.7785] [Medline: 28588005]
5. Vaidyam AN, Wisniewski H, Halamka JD, Kashavan MS, Torous JB. Chatbots and conversational agents in mental health: a review of the psychiatric landscape. Can J Psychiatry 2019 Jul;64(7):456-464. [doi: 10.1177/0706743719828977]
6. Maglogiannis I, Iliadis L, Pimenidis E, editors. An overview of chatbot technology. In: IFIP Advances in Information and Communication Technology (AIAI 2020 Conference Proceedings): Springer; 2020:373-383. [doi: 10.1007/978-3-030-49186-4_31]
7. Bird T, Mansell W, Wright J, Gaffney H, Tai S. Manage your life online: a web-based randomized controlled trial evaluating the effectiveness of a problem-solving intervention in a student sample. Behav Cogn Psychother 2018 Sep;46(5):570-582. [doi: 10.1017/S1352465817000820] [Medline: 29366432]
8. Enos G. Digital mental health agent for youths shows results similar to therapy group. Mental Health Weekly 2023 Nov 6;33(43):1-5. [doi: 10.1002/mhw.33846]
9. Fulmer R, Joerin A, Gentile B, Lakerink L, Rauws M. Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety: randomized controlled trial. JMIR Ment Health 2018 Dec 13;5(4):e64. [doi: 10.2196/mental.9782] [Medline: 30545815]
10. Alanezi F. Assessing the effectiveness of ChatGPT in delivering mental health support: a qualitative study. J Multidiscip Healthc 2024;17:461-471. [doi: 10.2147/JMDH.S447368] [Medline: 38314011]
11. Coşkun AB, Elmaoğlu E, Buran C, Yüzer Alsaç S. Integration of Chatgpt and E-Health Literacy: Opportunities, Challenges, and a Look Towards the Future. Journal of Health Reports and Technology ;10(1). [doi: 10.5812/jhrt-139748]
12. Yilanli M, McKay I, Jackson DI, Sezgin E. Large language models for individualized psychoeducational tools for psychosis: a cross-sectional study. medRxiv. Preprint posted online on Jul 29, 2024. [doi: 10.1101/2024.07.26.24311075]
13. Maurya RK, Montesinos S, Bogomaz M, DeDiego AC. Assessing the use of ChatGPT as a psychoeducational tool for mental health practice. Couns and Psychother Res 2025 Mar;25(1). [doi: 10.1002/capr.12759]
14. Yasukawa S, Tanaka T, Yamane K, et al. A chatbot to improve adherence to internet-based cognitive–behavioural therapy among workers with subthreshold depression: a randomised controlled trial. BMJ Ment Health 2024 Jan;27(1):e300881. [doi: 10.1136/bmjment-2023-300881]
15. Wang Y, Li S. Tech vs. tradition: ChatGPT and mindfulness in enhancing older adults' emotional health. Behav Sci (Basel) 2024;14(10):923. [doi: 10.3390/bs14100923]
16. Liu I, Liu F, Xiao Y, Huang Y, Wu S, Ni S. Investigating the key success factors of chatbot-based positive psychology intervention with retrieval- and generative pre-trained transformer (GPT)-based chatbots. International Journal of Human–Computer Interaction 2025 Jan 2;41(1):341-352. [doi: 10.1080/10447318.2023.2300015]
17. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. Future Healthc J 2019 Jun;6(2):94-98. [doi: 10.7861/futurehosp.6-2-94] [Medline: 31363513]
18. Townsend BA, Plant KL, Hodge VJ, Ashaolu OT, Calinescu R. Medical practitioner perspectives on AI in emergency triage. Front Digit Health 2023;5:1297073. [doi: 10.3389/fdgth.2023.1297073]
19. Cross S, Bell I, Nicholas J, et al. Use of AI in mental health care: community and mental health professionals survey. JMIR Ment Health 2024 Oct 11;11:e60589. [doi: 10.2196/60589] [Medline: 39392869]
20. Sharma A, Lin IW, Miner AS, Atkins DC, Althoff T. Human–AI collaboration enables more empathic conversations in text-based peer-to-peer mental health support. Nat Mach Intell 2023;5(1):46-57. [doi: 10.1038/s42256-022-00593-2]
21. Baker HP, Dwyer E, Kalidoss S, Hynes K, Wolf J, Strelzow JA. ChatGPT's ability to assist with clinical documentation: a randomized controlled trial. J Am Acad Orthop Surg 2024 Feb 1;32(3):123-129. [doi: 10.5435/JAAOS-D-23-00474] [Medline: 37976385]
22. Dergaa I, Fekih-Romdhane F, Hallit S, et al. ChatGPT is not ready yet for use in providing mental health assessment and interventions. Front Psychiatry 2023;14:1277756. [doi: 10.3389/fpsyt.2023.1277756] [Medline: 38239905]
23. Bartal A, Jagodnik KM, Chan SJ, Dekel S. AI and narrative embeddings detect PTSD following childbirth via birth stories. Sci Rep 2024;14(1):54242. [doi: 10.1038/s41598-024-54242-2]
24. Guest R, Tran Y, Gopinath B, Cameron ID, Craig A. Prevalence and psychometric screening for the detection of major depressive disorder and post-traumatic stress disorder in adults injured in a motor vehicle crash who are engaged in compensation. BMC Psychol 2018 Feb 21;6(1):4. [doi: 10.1186/s40359-018-0216-5] [Medline: 29467035]
25. Maurer DM, Raymond TJ, Davis BN. Depression: screening and diagnosis. Am Fam Physician 2018 Oct 15;98(8):508-515. [Medline: 30277728]
26. Galido PV, Butala S, Chakerian M, Agustines D. A case study demonstrating applications of ChatGPT in the clinical management of treatment-resistant schizophrenia. Cureus 2023 Apr;15(4):e38166. [doi: 10.7759/cureus.38166] [Medline: 37252576]
27. Pandya A, Lodha P, Ganatra A. Is ChatGPT ready to change mental healthcare? Challenges and considerations: a reality-check. Front Hum Dyn 2024;5. [doi: 10.3389/fhumd.2023.1289255]

XSL•FO
RenderX

28. Rogan J, Bucci S, Firth J. Health care professionals' views on the use of passive sensing, AI, and machine learning in mental health care: systematic review with meta-synthesis. JMIR Ment Health 2024 Jan 23;11:e49577. [doi: 10.2196/49577] [Medline: 38261403]
29. Blease C, Worthen A, Torous J. Psychiatrists' experiences and opinions of generative artificial intelligence in mental healthcare: An online mixed methods survey. Psychiatry Res 2024 Mar;333:115724. [doi: 10.1016/j.psychres.2024.115724] [Medline: 38244285]
30. Wang L, Wan Z, Ni C, et al. A systematic review of chatgpt and other conversational large language models in healthcare. Health Informatics. Preprint posted online on 2024. [doi: 10.1101/2024.04.26.24306390]
31. Sweeney C, Potts C, Ennis E, et al. Can chatbots help support a person's mental health? Perceptions and views from mental healthcare professionals and experts. ACM Trans Comput Healthcare 2021 Jul 31;2(3):1-15. [doi: 10.1145/3453175]
32. Blease C, Kharko A, Annoni M, Gaab J, Locher C. Machine learning in clinical psychology and psychotherapy education: a mixed methods pilot survey of postgraduate students at a Swiss University. Front Public Health 2021;9:623088. [doi: 10.3389/fpubh.2021.623088] [Medline: 33898374]
33. Torous J, Bucci S, Bell IH, et al. The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality. World Psychiatry 2021 Oct;20(3):318-335. [doi: 10.1002/wps.20883] [Medline: 34505369]
34. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
35. Cohen J. Statistical Power Analysis for the Behavioral Sciences, 2nd edition: Lawrence Erlbaum Associates; 1988.
36. Upheal. URL: https://www.upheal.io [accessed 2024-10-20]
37. Heidi Health. URL: https://www.heidihealth.com/au [accessed 2024-10-20]
38. Autonotes. URL: https://autonotes.ai [accessed 2024-10-20]
39. Olawade DB, Wada OZ, Odetayo A, David-Olawade AC, Asaolu F, Eberhardt J. Enhancing mental health with artificial intelligence: current trends and future prospects. Journal of Medicine, Surgery, and Public Health 2024 Aug;3:100099. [doi: 10.1016/j.glmedi.2024.100099]
40. Ettman CK, Galea S. The potential influence of AI on population mental health. JMIR Ment Health 2023 Nov 16;10:e49936. [doi: 10.2196/49936] [Medline: 37971803]
41. Li Y, Li Z. Knowledge, attitude, and practices regarding ChatGPT among health care professionals. Am J Manag Care 2024 Sep 1;30(9):e258-e265. [doi: 10.37765/ajmc.2024.89604]
42. Zhang M, Scandiffio J, Younus S, et al. The adoption of AI in mental health care-perspectives from mental health professionals: qualitative descriptive study. JMIR Form Res 2023 Dec 7;7:e47847. [doi: 10.2196/47847] [Medline: 38060307]
43. Mirnig A, Gaertner M, Meschtscherjakov A, Tscheligi M. Blinded by novelty: a reflection on participant curiosity and novelty in automated vehicle studies based on experiences from the field. 2020 Presented at: Association for Computing Machinery; Sep 6-9, 2020; Magdeburg, Germany. [doi: 10.1145/3404983.3405593]
44. Andrade C. The inconvenient truth about convenience and purposive samples. Indian J Psychol Med 2021 Jan;43(1):86-88. [doi: 10.1177/0253717620977000] [Medline: 34349313]
45. Guest G, Bunce A, Johnson L. How many interviews are enough? An experiment with data saturation and variability. Field methods 2006;18(1):59-82. [doi: 10.1177/1525822X05279903]
46. Button KS, Ioannidis JPA, Mokrysz C, et al. Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci 2013 May;14(5):365-376. [doi: 10.1038/nrn3475]
47. OpenAI. ChatGPT can now see, hear, and speak OpenAI website. 2023 Sep 25. URL: https://openai.com/index/chatgpt-can-now-see-hear-and-speak [accessed 2024-10-18]
48. Rajaei A. Teaching in the age of AI/ChatGPT in mental-health-related fields. The Family Journal 2024 Jan;32(1):6-10. [doi: 10.1177/10664807231209721]

Abbreviations

AAPi: Australian Association of Psychologists Inc
AI: artificial intelligence
APS: Australian Psychological Society
PTSD: posttraumatic stress disorder
QT: qualitative theme
RQ: research question
T1: recommending a chatbot before demonstration
T2: recommending a chatbot after demonstration

XSL•FO
RenderX

Edited by A Choudhury; submitted 09.01.25; peer-reviewed by E Dal, L Zhang, M Elbattah; revised version received 07.08.25; accepted 10.08.25; published 16.09.25.

Please cite as:
Goldie J, Dennis S, Hipgrave L, Coleman A
Practitioner Perspectives on the Uses of Generative AI Chatbots in Mental Health Care: Mixed Methods Study
JMIR Hum Factors 2025;12:e71065
URL: https://humanfactors.jmir.org/2025/1/e71065
doi:10.2196/71065

©Jessie Goldie, Simon Dennis, Lyndsey Hipgrave, Amanda Coleman. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 16.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

User Intent to Use DeepSeek for Health Care Purposes and Their Trust in the Large Language Model: Multinational Survey Study

Avishek Choudhury, PhD; Yeganeh Shahsavar, MS; Hamid Shamszare, MS

Industrial and Management Systems Engineering, Benjamin M Statler College of Engineering and Mineral Resources, West Virginia University, 1306 Evansdale Drive, 321 Engineering Sciences Building, Morgantown, WV, United States

Corresponding Author:
Avishek Choudhury, PhD
Industrial and Management Systems Engineering, Benjamin M Statler College of Engineering and Mineral Resources, West Virginia University, 1306 Evansdale Drive, 321 Engineering Sciences Building, Morgantown, WV, United States

Abstract

Background: Generative artificial intelligence (AI)—particularly large language models (LLMs)—has generated unprecedented interest in applications ranging from everyday questions and answers to health-related inquiries. However, little is known about how everyday users decide whether to trust and adopt these technologies in high-stakes contexts such as personal health.

Objectives: This study examines how ease of use, perceived usefulness, and risk perception interact to shape user trust in and intentions to adopt DeepSeek, an emerging LLM-based platform, for health care purposes.

Methods: We adapted survey items from validated technology acceptance scales to assess user perception of DeepSeek. A 12-item Likert scale questionnaire was developed and pilot-tested (n=20). It was then distributed on the web to users in India, the United Kingdom, and the United States who had used DeepSeek within the past 2 weeks. Data analysis involved descriptive frequency assessments and Partial Least Squares Structural Equation Modeling. The model assessed direct and indirect effects, including potential quadratic relationships.

Results: A total of 556 complete responses were collected, with respondents almost evenly split across India (n=184), the United Kingdom (n=185), and the United States (n=187). Regarding AI in health care, when asked whether they were comfortable with their health care provider using AI tools, 59.3% (n=330) were fine with AI use provided their doctor verified its output, and 31.5% (n=175) were enthusiastic about its use without conditions. DeepSeek was used primarily for academic and educational purposes, 50.7% (n=282) used DeepSeek as a search engine, and 47.7% (n=265) used it for health-related queries. When asked about their intent to adopt DeepSeek over other LLMs such as ChatGPT, 52.1% (n=290) were likely to switch, and 28.9% (n=161) were very likely to do so. The study revealed that trust plays a pivotal mediating role; ease of use exerts a significant indirect impact on usage intentions through trust. At the same time, perceived usefulness contributes to trust development and direct adoption. By contrast, risk perception negatively affects usage intent, emphasizing the importance of robust data governance and transparency. Significant nonlinear paths were observed for ease of use and risk, indicating threshold or plateau effects.

Conclusions: Users are receptive to DeepSeek when it is easy to use, useful, and trustworthy. The model highlights trust as a mediator and shows nonlinear dynamics shaping AI-driven health care tool adoption. Expanding the model with mediators such as privacy and cultural differences could provide deeper insights. Longitudinal experimental designs could establish causality. Further investigation into threshold and plateau phenomena could refine our understanding of user perceptions as they become more familiar with AI-driven health care tools.

(JMIR Hum Factors 2025;12:e72867) doi:10.2196/72867

KEYWORDS

artificial intelligence; data privacy; health informatics; human factors engineering; ChatGPT; risk assessment; technology acceptance; trust; user adoption

Introduction

Over the past few years, generative artificial intelligence (AI)—particularly large language models (LLMs)—has transitioned from research laboratories to widespread public use, sparking unprecedented interest in applications ranging from everyday question and answer to health-related inquiries. Media narratives often amplify these systems' capabilities, leading individuals to adopt them in ways that can exceed their original design intentions. While such enthusiasm has undeniably broadened the reach of LLMs, it also carries risks—biased training data [1], fluctuations in output quality, and users' lack of domain expertise all converge in potentially problematic ways. Indeed, existing research, including our own previous work, suggests that many users consult LLMs (eg, Chat Generative Pretrained Model) for medical advice despite

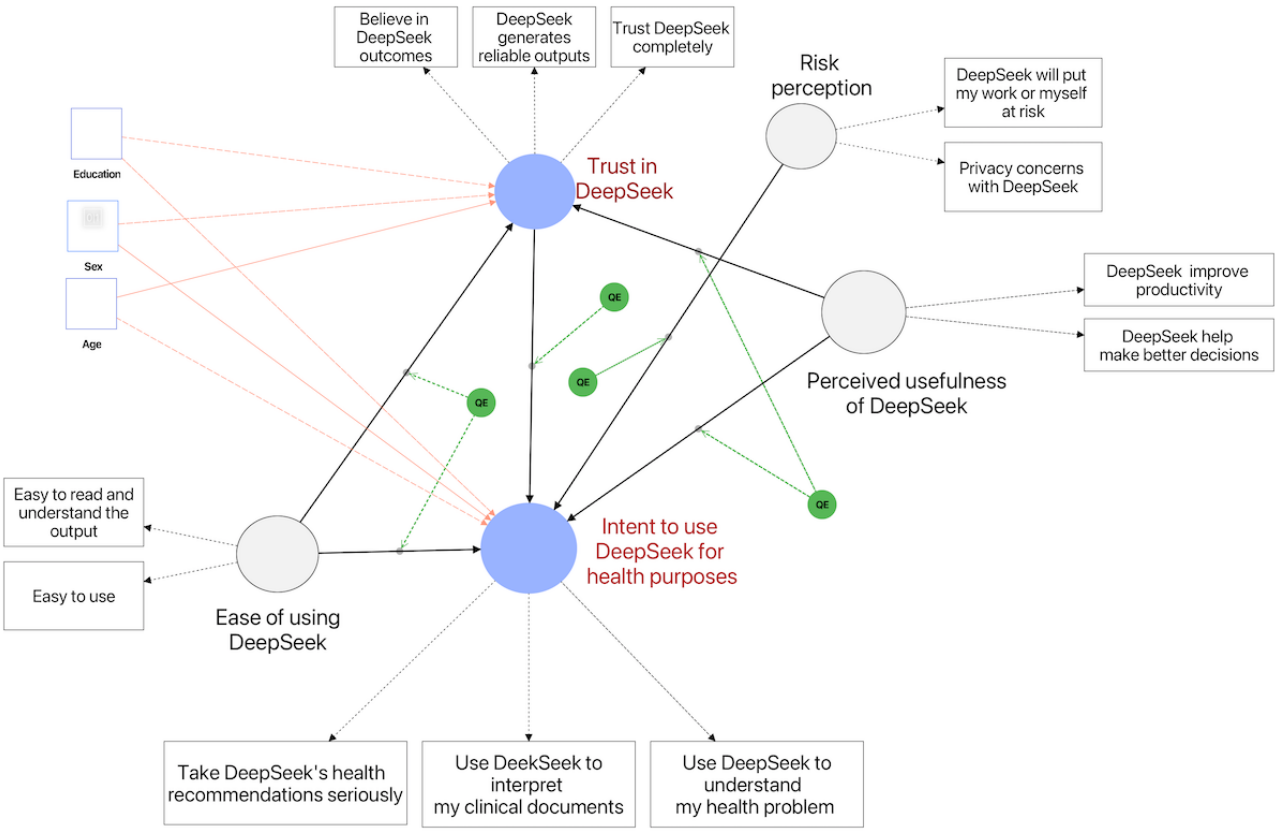
XSL•FO
RenderX

having limited or no formal clinical background. As new LLMs such as DeepSeek emerge—touting features such as cost-effectiveness, human-like conversation, and comparable performance to established systems—the likelihood that nonexpert populations will rely on AI tools to inform health care decisions appears poised to intensify.

However, little is known about how everyday users decide whether to trust and adopt these technologies—particularly in high-stakes contexts such as personal health. Studies show that while an LLM may provide *convincing* outputs, its accuracy can shift unpredictably due to biases or changes in model parameters. Thus, a recommendation that proves correct today does not guarantee reliability tomorrow. The media equation theory further underscores how human-like dialogue can encourage users to treat an AI system as though it were a trustworthy human advisor, heightening the risk of overreliance [2-5]. Against this backdrop, there remains a gap in understanding how perceptions of ease of use, perceived usefulness, risk, and trust coalesce to shape user intentions to use LLMs for health-related queries. Specifically, while prior studies on ChatGPT and similar platforms have begun to shed light on user attitudes, few have examined the human factors by which trust is established—or eroded—when individuals face important decisions about their well-being.

In this study, we add to the body of knowledge by focusing on DeepSeek, an open-source LLM that rapidly gained popularity after surpassing other competitors in App Store rankings. Public and media hype surrounding DeepSeek's purported superiority can create a halo effect, wherein users generalize competence across all domains without adequately accounting for potential inaccuracies or biases. Drawing on technology acceptance models and behavioral theories [6,7], we propose and empirically test a conceptual framework that explores the association among ease of use, risk perception, and perceived usefulness in determining user trust in-and-intentions to rely on DeepSeek for health-related advice. Figure 1 shows the conceptual model. A key focus centers on *ease of use*—a construct traditionally linked to technology acceptance. Specifically, the study posits that ease of use will significantly foster trust in DeepSeek and increase users' intent to adopt it for health purposes. This expectation aligns with the view that a user-friendly interface reduces cognitive strain, instilling an impression of competence and reliability [8]. Moreover, straightforward navigation of an AI system in a clinical or personal health setting can lower barriers to initial adoption, prompting individuals to incorporate the technology more readily into their decision-making.

Figure 1. Conceptual structural framework. Education, age, and sex are the control variables. Latent constructs are represented by circles and observed variables by rectangles. QE: quadratic effect.



Another central component is trust, which is widely recognized as a pivotal determinant of individuals' willingness to delegate sensitive tasks to AI systems. Here, trust is posited to positively influence intent to use DeepSeek, reflecting the premise that users must believe in the system's credibility and accuracy before relying on its outputs for personally significant choices, such as querying about personal health or symptoms. In parallel, the model predicts that perceived usefulness—the degree to which users believe that DeepSeek effectively aids in accomplishing tasks—will both feed into trust and directly

XSL•FO
RenderX

motivate the decision to adopt [9]. If DeepSeek demonstrably improves efficiency, offers relevant information, or yields better outcomes, individuals are more likely to view the technology as beneficial and deserving of confidence, thus reinforcing their intention to use.

In contrast, risk perception is anticipated to negatively impact the intent to use. Particularly in health-related scenarios, concerns over data privacy, the possibility of incorrect diagnoses, or broader ethical dilemmas can inhibit adoption. Users who sense a high level of risk may hesitate to rely on an AI system, even if they recognize certain advantages [10].

In addition, the framework addresses mediation pathways, where trust operates as an intermediary in 2 distinct relationships. First, the study investigates whether trust mediates the link between ease of use and intent to use, positing that user-friendly design can bolster trust, reinforcing the inclination to adopt DeepSeek. Second, perceived usefulness is hypothesized to enhance trust, which in turn increases intent to use, reflecting the theory that tangible benefits establish an underlying belief in the system's reliability.

We explore the following 8 hypotheses (Hs):

- H1: ease of use positively influences trust in DeepSeek.
- H2: ease of use positively influences intent to use DeepSeek for health-related purposes.
- H3: trust in DeepSeek positively influences intent to use DeepSeek for health-related purposes.
- H4: risk perception negatively influences intent to use DeepSeek for health-related purposes.
- H5: perceived usefulness positively influences trust in DeepSeek.
- H6: perceived usefulness positively influences intent to use DeepSeek for health-related purposes.
- H7: trust in DeepSeek mediates the relationship between ease of use and intent to use DeepSeek.
- H8: trust in DeepSeek mediates the relationship between perceived usefulness and intent to use DeepSeek.

Although canonical technology acceptance models often posit linear relationships among constructs such as ease of use, perceived usefulness, and intention to adopt, scholars have long recognized scenarios where user perceptions follow threshold or plateau patterns rather than increasing in a straightforward incremental manner [11,12]. Our study also acknowledges the possibility of nonlinear relationships [13]. In practice, once an interface or system achieves a certain baseline of usability, additional refinements may produce diminishing returns—that is, users do not perceive incremental gains in simplicity as adding meaningful value [14]. For instance, a health-focused AI system that goes from moderately user-friendly to extremely user-friendly might not see a proportional rise in users' trust or adoption if people already consider the interface easy enough. Likewise, risk perception can display an inverted-U or other nonlinear patterns, where a just-right amount of perceived risk may be beneficial in prompting vigilant, responsible usage, whereas extremely low levels of risk can lead to complacency, and extremely high levels can deter adoption outright. From a cognitive and behavioral standpoint, Kahneman and Tversky's Prospect Theory also illustrates that people exhibit loss aversion in nonlinear ways: even a small increase in perceived risk can disproportionately reduce willingness to adopt [15].

By incorporating quadratic (QE) terms for both ease of use and risk perception, this study accommodates the possibility that these factors do not scale linearly with trust or intent to adopt. Specifically, we test whether user attitudes intensify, plateau, or invert once certain inflection points in usability or perceived risk are reached. A similar rationale applies to perceived usefulness, as exceedingly high levels of perceived utility may lead to skepticism—users might become suspicious of a too good to be true system promising flawless performance, thus reducing their trust or perceived credibility.

Methods

Ethical Considerations

The West Virginia University Institutional Review Board approved the study (protocol number 2302725983; classified as a flex protocol type). All participants were required to provide informed consent statement through an individualized survey link, before proceeding to the primary survey. All data were anonymized. Participants were compensated by the paid audience panelling service.

Survey Instrument

The survey items were adapted from established measurement scales commonly used in technology acceptance and human-computer interaction research [16,17]. A preliminary literature review identified key constructs (eg, ease of use, trust, risk perception, perceived usefulness, and usage intentions) pertinent to AI-driven applications, particularly in health care [18,19]. Existing, validated items were then modified linguistically to reflect DeepSeek's functionality. For instance, several questions included references to reliability and accuracy, adapted from trust in automation scales. At the same time, risk perception measures were framed to address data privacy and potential adverse outcomes in health-related tasks [20-22].

As shown in Textbox 1, the resultant questionnaire comprised 12 primary items, each measured on a 4-point forced Likert scale. The decision to use a 4-point format was motivated by the desire to encourage decisive responses and minimize the fence-sitting effect often observed with neutral midpoint options. In the context of LLM usage, truly neutral opinions are comparatively less likely; users typically find these AI systems either useful or problematic in some capacity. By eliminating the neutral choice, the instrument sharpens distinctions in participant attitudes and prevents potential ambivalence from masking genuine inclinations. Still, we acknowledge that the absence of a neutral category can suppress expressions of ambivalence.

Questions were grouped to form latent construct and validated. The instrument also had questions about participant demographics. In addition, the survey incorporated a checking question to verify that respondents thoroughly read all questions before providing their answers, further ensuring data quality.

XSL•FO
RenderX

Textbox 1. Survey questions used in the study model.

Trust in DeepSeek

- To what extent do you agree or disagree with the following: I would trust DeepSeek completely. (T1)
- To what extent do you agree or disagree with the following: DeepSeek generates reliable outputs. (T2)
- To what extent do you agree or disagree with the following: I believe in DeepSeek outcomes. (T3)

Intent to use DeepSeek for health-related purposes

- To what extent do you agree or disagree with the following: I would take DeepSeek's health-related recommendations seriously. (IU1)
- To what extent do you agree or disagree with the following: I will use DeepSeek to interpret my clinical documents. (IU2)
- To what extent do you agree or disagree with the following: I want to use DeepSeek to understand minor health problem I occasionally get. (IU3)

Ease of using DeepSeek

- To what extent do you agree or disagree with the following: I find it easy to read and understand the output generated by DeepSeek. (E1)
- To what extent do you agree or disagree with the following: DeepSeek is easy to use. (E2)

Perceived usefulness of DeepSeek

- To what extent do you agree or disagree with the following: DeepSeek-generated outcomes will help me improve productivity. (PU1)
- To what extent do you agree or disagree with the following: DeepSeek-generated outcomes will help me make better decisions. (PU2)

Risk perception

- To what extent do you agree or disagree with the following: using DeepSeek will put my work or myself at risk. (R1)
- To what extent do you agree or disagree with the following: I have privacy concerns with DeepSeek. (R2)

Pilot Testing and Data Collection

Before launching the main data collection, the survey was piloted with a small convenience sample (n=20) who met the criterion of having used DeepSeek at least once in the previous 2 weeks. Pilot participants were asked to provide feedback on item clarity, redundancy, and overall length. Minor revisions were made, including refining the wording of risk-related items and adjusting the Likert scales for consistency.

Following pilot testing, a web-based version of the final questionnaire was administered via Centiment, a paid audience paneling service. Centiment maintains a prescreened database of respondents for market research. Panelists opt in by agreeing to Centiment's terms of use and privacy policy, and they are profiled in advance based on demographic and firmographic criteria. This approach enabled us to target specific audiences (eg, certain regions or demographic subgroups) and apply quotas to approximate a more representative sample. Centiment's quality controls—such as IP-based fingerprinting, proxy detection, and fraud scoring—were used to minimize duplicate or low-effort responses.

The survey was distributed to India, the United Kingdom, and the United States. Participant recruitment targeted adult users (aged 18 years or older) who reported using DeepSeek at least once in the preceding 2 weeks. Exclusion criteria included individuals with no prior exposure to DeepSeek.

Data Collection Procedure

Data collection took place over 2 weeks. Each participant received an individualized survey link, which led to a landing page containing an informed consent statement. After consenting, participants proceeded to the primary survey. The survey platform automatically recorded session details, including session ID and IP address (to prevent duplication only). All identifying information was removed before data analysis. Participants could terminate the survey at any point without penalty.

Statistical Analyses

First, we conducted a frequency analysis of the survey responses related to participant characteristics and DeepSeek perception. Given that all data were gathered using a single survey instrument, common method bias (CMB) represented a potential concern. To evaluate the extent of CMB, we conducted Harman's single-factor test, wherein the principal components analysis revealed that the highest variance explained by any single factor was 33%, falling below the commonly referenced threshold of 50% [23]. Consequently, these results indicate that no substantial CMB threat exists in our dataset.

Second, we used nonparametric Partial Least Squares Structural Equation Modeling (PLS-SEM) to investigate the relationships between the latent constructs and their observed indicators. PLS-SEM is suitable for analyzing complex models with multiple dependent and independent variables. The method allows for simultaneous evaluation of the measurement model, which assesses construct reliability and validity, and the structural model, which examines the hypothesized relationships between constructs [24]. The PLS-SEM method is also suitable for nonnormal data [25].

Our measurement model consisted of five reflective latent constructs, namely, (1) trust in DeepSeek, (2) intent to use DeepSeek for health-related purposes, (3) ease of using

XSL•FO
RenderX

DeepSeek, (4) perceived usefulness of DeepSeek, and (5) risk perception (Textbox 1). Each construct was measured using multiple observed indicators, with 3, 3, 2, 2, and 2 items, respectively. The constructs were assessed based on factor loading greater than 0.5. The reliability and validity were evaluated based on composite reliability (rho_C) greater than 0.70 [26] and average variance extracted (AVE) greater than 0.50 [26,27]. Variance inflation factor values were calculated to detect multicollinearity.

The validated measurement model was bootstrapped with 10,000 iterations to obtain the *P* values, bias-corrected standardized parameter estimates (β), and CIs. The model was controlled for education level, age, and sex. These were used as covariates in the model.

Sample Size Justification

Relatively large sample sizes are commonly used in structural equation modeling (SEM) to ensure robust parameter estimation and adequate statistical power. Although guidelines in the SEM literature vary, several rules of thumb have been proposed. Some recommend a minimum of 200 participants for most SEM applications. In contrast, others suggest having at least 5-10 respondents per estimated parameter in the model [28-30]. Our study's final instrument contained 12 key items spanning multiple latent constructs (eg, ease of use, trust, risk perception, perceived usefulness, and intent to use). Consequently, typical rules of thumb would indicate a desired sample size of 300 - 500 participants to meet assumptions of stable parameter estimation.

Results

Participant Characteristics

A total of 556 complete responses were collected. Geographically, 184 individuals were based in India, 185 in the United Kingdom, and 187 in the United States, reflecting an almost even split across these 3 locales. Regarding education, most respondents (n=214, 38.4%) held a bachelor's degree, followed by master's degree holders (n=153, 27.5%), high school graduates (n=142, 25.5%), those with some high school (n=24, 4%), and those with doctoral-level degrees (n=23, 4%). The sex distribution was nearly balanced, with 49.2% (n=274) identifying as male and 50.7% (n=282) identifying as female. In terms of age, 24.6% (n=137) were aged 26 - 35 year, 14% (n=78) were aged 18 - 25 years, 19.4% (n=108) were aged 36 - 45 years, another 19.4% (n=108) were aged 46 - 55 years, 15% (n=85) were aged 56 - 65 years, and 7% (n=40) were aged 66 years or older. Nearly 32.9% (n=183) used DeepSeek "once a month," 46.0% (n=156) used DeepSeek "once a week," 24.6% (n=137) used it "more than once per week," and 14% (n=80) used it "almost every day." Out of 556, about 55.4% (n=308) used DeepSeek primarily for academic and educational purposes, approximately 50.7% (n=282) used DeepSeek as a search engine, and 47.7% (n=265) used DeepSeek for health-related queries.

Participants' Perception of AI in Health Care

Responding to the question, "Will you be ok if your healthcare provider (doctor) uses similar AI tools for clinical purposes?" out of 556 respondents, only about 9% (51) were uncomfortable with their health care provider using similar AI tools for clinical purposes. About 59.3% (330) were fine with it as long as their doctor verifies the AI's input, while approximately 31.5% (175) would be enthusiastic about their provider using AI without any conditions. This suggests that while there are some reservations about the direct use of AI in clinical settings, most people are open to it if professional oversight is maintained.

As a response to this survey question, "To what extent can using such AI diagnoses increase patient safety?" nearly 5% (27) felt that using AI for diagnoses would not improve patient safety and might even put patients at risk. In contrast, 20.5% (114) believed that it would increase safety somewhat, while 29.3% (163) thought that it would do so moderately, and 27.5% (153) felt that it would enhance safety quite a lot. An additional 18% (99) believed that AI could increase patient safety to a great degree. Most respondents see a positive potential for AI in improving patient safety, with most expecting at least a moderate benefit from its use.

Participants' LLM Use Pattern

Responding to the question "In the last 6 months, how often have you used any large language model?" about 25.2% (140) indicated that they used an LLM once a month, while the largest group, 43.7% (243), reported using it weekly. Approximately 20.3% (113) of respondents mentioned using an LLM almost daily, and 11% (60) stated using it multiple times daily.

Participants' Intent to Adapt DeepSeek Over Other LLMs

As a response to the question, "How likely are you to switch over and use DeepSeek instead of other large language models like ChatGPT?" only a small fraction reported no intent to switch over to DeepSeek, with 3% (18) indicating that they would not consider it and 16% (87) saying that they are unlikely to make the change. A significant majority are open to the idea, with 52.1% (290) stating that they are likely to switch and an additional 28.9% (161) expressing that they are very likely to do so. This suggests that many users are ready to transition from LLMs such as ChatGPT to DeepSeek.

Measurement Model and Survey Validation

Table 1 shows survey responses on ease of use, trust, usefulness, intent for health use, and risk perception of DeepSeek. Respondents generally view DeepSeek favorably regarding ease of use, trust, and perceived usefulness. For instance, both ease of use items (E1 and E2) received strong positive responses, with nearly 90% of participants somewhat or strongly agreeing that DeepSeek is easy to use. Trust in the platform is similarly high, as indicated by most respondents strongly or somewhat agreeing with all trust-related items (T1, T2, and T3). perceived usefulness ratings are also robust, with most users acknowledging the benefits of DeepSeek. In addition, the intent to use DeepSeek for health-related purposes is positive, although slightly more moderate than the other categories. However, risk perception responses are more mixed, with a relatively even distribution across all levels of agreement.

Table 2 indicates a robust measurement model with strong evidence of convergent and discriminant validity across all

XSL•FO
RenderX

latent constructs. Convergent validity is supported by high AVE values: each construct exceeds the commonly accepted 0.50 threshold (ranging from 0.650 to 0.820), indicating that the items explain a substantial portion of their respective constructs. Likewise, rho_C and Cronbach alpha values are mostly well above the 0.70 benchmark, confirming internal consistency among indicators.

Table . Participants' responses to the survey questions used in the conceptual structural framework.

Observed variables	Strongly disagree	Somewhat disagree	Somewhat agree	Strongly agree
Ease of using DeepSeek, n (%)				
E1	17 (3.10)	39 (7.00)	267 (48.00)	233 (41.90)
E2	5 (0.90)	32 (5.80)	214 (38.50)	305 (54.90)
Trust in DeepSeek, n (%)				
T1	26 (4.70)	96 (17.30)	240 (43.20)	194 (34.90)
T2	8 (1.40)	46 (8.30)	268 (48.20)	234 (42.10)
T3	11 (2.00)	71 (12.80)	267 (48.00)	207 (37.20)
Perceived usefulness of DeepSeek, n (%)				
PU1	10 (1.80)	65 (11.70)	246 (44.20)	235 (42.30)
PU2	12 (2.20)	66 (11.90)	257 (46.20)	221 (39.80)
Intent to use DeepSeek for health-related purposes, n (%)				
IU1	27 (4.90)	80 (14.40)	246 (44.20)	203 (36.50)
IU2	36 (6.50)	108 (19.40)	233 (41.90)	179 (32.20)
IU3	25 (4.50)	71 (12.80)	257 (46.20)	203 (36.50)
Risk perception, n (%)				
R1	132 (23.70)	127 (22.80)	167 (30.00)	130 (23.40)
R2	66 (11.90)	139 (25.00)	185 (33.30)	166 (29.90)

XSL•FO
RenderX

Table . The measurement model and validation and reliability.

Constructs and observed variable	Factor loading	AVE[a]	Rho_C
Trust in DeepSeek		0.65	0.85
T1	0.82		
T2	0.78		
T3	0.82		
Intent to use DeepSeek for health-related purposes		0.70	0.88
IU1	0.85		
IU2	0.84		
IU3	0.82		
Ease of using DeepSeek		0.82	0.90
E1	0.90		
E2	0.91		
Perceived usefulness of DeepSeek		0.77	0.87
PU1	0.86		
PU2	0.90		
Risk perception		0.66	0.79
R1	0.97		
R2	0.64		

[a]AVE: average variance extracted.

Discriminant validity is supported by the Heterotrait-Monotrait ratios, nearly all of which fall below the critical value of 0.90 or 1.00. Variance inflation factor also ranged between 1.76 and 1.20, which is within the acceptable limits (less than 2.50) [31], indicating no multicollinearity. In addition, the R^2 values of 0.55 (adjusted 0.54) for intent to use DeepSeek for health-related purposes and 0.60 (adjusted 0.59) for trust in DeepSeek indicate that the explanatory power of the proposed paths is moderate to strong. The significant *t* statistics and *P* values across the key relationships further verify that the conceptualized constructs have been measured accurately, with only a few nonsignificant quadratic paths. These findings validate the measurement instrument and confirm that the theorized latent variables—ease of using DeepSeek, perceived usefulness, trust in DeepSeek, risk perception, and intent to use DeepSeek for health-related purposes—exhibit sufficient reliability and validity to warrant confidence in subsequent structural analyses. In addition to the overall validity of the measurement model, the findings also support the validity of the individual survey questions that served as indicators for each latent construct. First, the high factor loadings (as implied by the satisfactory AVE values) suggest that each item meaningfully contributes to measuring its intended factor, that is, users' responses to questions about ease of use, trust, perceived usefulness, risk, or intent to use strongly correlate with the respective latent constructs they were designed to represent. Second, the internal consistency indices (Cronbach alpha and composite reliability) confirm that groups of questions intended to measure the same construct hang together well, indicating that the survey items reliably capture the same underlying concept. Finally, the discriminant validity checks confirm that questions targeting one construct do not overlap excessively with those measuring other constructs; this indicates that the items' wording and content domains effectively capture distinct dimensions of user perceptions toward DeepSeek. These indicators demonstrate that the survey questions—and not just the overarching constructs—exhibit satisfactory validity.

SEM (Direct, Indirect, and Total Effects)

Table 3 shows that the path coefficients and mediation analyses provide strong evidence for most of the hypothesized relationships, although one notable exception emerged regarding risk perception. *Ease of use* (*H1*) has a significantly positive direct effect on trust in DeepSeek. This indicates that when users perceive DeepSeek as more intuitive and less effortful to operate, their trust in the system increases correspondingly. However, the direct relationship between *ease of use* and *intent to use DeepSeek for health-related purposes* (*H2*) is not statistically significant. Despite this, a significant total effect and a strong indirect path through *trust* confirm that trust fully mediates the impact of ease of use on user intentions. Hence, while there is no direct link between ease of use and intent, ease of use indirectly drives intent via trust, satisfying *H2's* broader premise when mediation is considered. *Trust in DeepSeek* (*H3*) exerts a strong, positive influence on *intent to use DeepSeek.* This suggests that once users develop confidence in DeepSeek's outputs, they are more inclined to rely on it for health-focused tasks. *Risk perception* (*H4*) shows a significant negative path coefficient, supporting the hypothesized relationship. This implies that higher perceived risk—as captured in the current measurement—correlates with a lower intent to use DeepSeek,

XSL•FO
RenderX

deterring adoption. The findings strongly support both direct and mediated effects concerning perceived usefulness.

Table . Direct, indirect, and total effects observed in the proposed structural framework.

	β[a] (SD)		
Paths	Direct effect	Indirect effect	Total effect
Ease of using DeepSeek → intent to use DeepSeek for health-related purposes	0.07 (0.058)	—[b]	0.25 (0.06)[c]
Ease of using DeepSeek → trust in DeepSeek	0.36 (0.05)[c]	—	0.36 (0.05)[c]
Perceived usefulness of DeepSeek → intent to use DeepSeek for health-related purposes	0.17 (0.05)[c]	—	0.40 (0.04)[c]
Perceived usefulness of DeepSeek → trust in DeepSeek	0.52 (0.04)[c]	—	0.52 (0.04)[c]
Risk perception → intent to use DeepSeek for health-related purposes	0.20 (0.03)[c]	—	0.21 (0.03)[c]
Trust in DeepSeek → intent to use DeepSeek for health-related purposes	0.45 (0.05)[c]	—	0.45 (0.05)[c]
Ease of using DeepSeek → trust in DeepSeek → intent to use DeepSeek for health-related purposes	—	0.16 (0.02)[c]	0.16 (0.02)[c]
Perceived usefulness of DeepSeek → trust in DeepSeek → intent to use DeepSeek for health-related purposes	—	0.24 (0.03)[c]	0.24 (0.03)[c]
Control variables			
Age → trust in DeepSeek → intent to use DeepSeek for health-related purposes	—	−0.02 (0.01)	−0.02 (0.01)
Age → intent to use DeepSeek for health-related purposes	0.09 (0.03)[c]	—	0.07 (0.03)[d]
Age → trust in DeepSeek	−0.05 (0.03)	—	−0.05 (0.03)
Education → intent to use DeepSeek for health-related purposes	−0.06 (0.03)	—	−0.04 (0.03)
Education → trust in DeepSeek	0.04 (0.03)	—	0.04 (0.03)
Sex → trust in DeepSeek → intent to use DeepSeek for health-related purposes	—	0.04 (0.02)	0.04 (0.02)
Sex → intent to use DeepSeek for health-related purposes	0.11 (0.06)	—	0.15 (0.06)[d]
Sex → trust in DeepSeek	0.09 (0.05)	—	0.09 (0.05)

[a]β: standardized coefficient.

[b]Not applicable.

[c]$P<.001$.

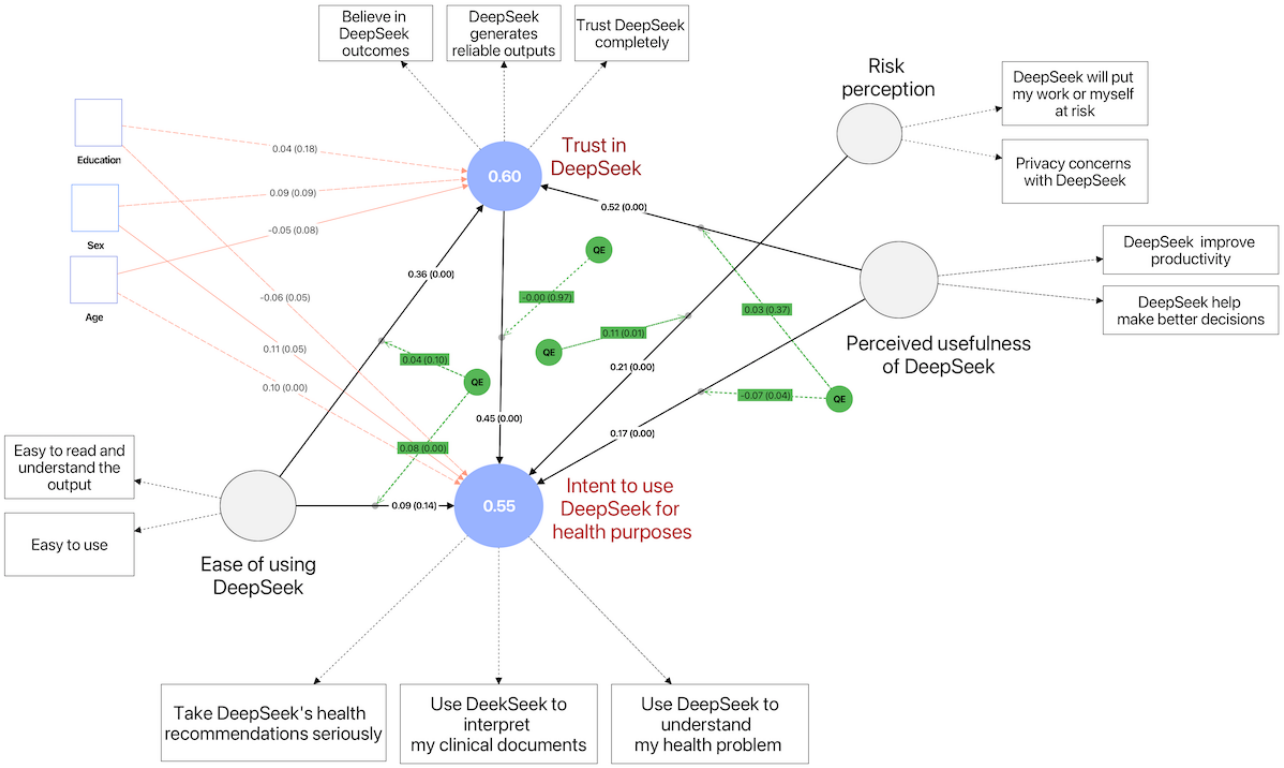
[d]$P<.05$.

Perceived usefulness (*H5*) significantly increases *trust in DeepSeek*, indicating that when participants see value in DeepSeek's capabilities, their faith in the technology grows. *Perceived usefulness* also shows a robust direct effect on *intent to use DeepSeek* (*H6*), underscoring the salience of benefit-driven evaluations in motivating adoption. In addition, mediation analyses reveal that trust partially mediates the effect of perceived usefulness on intent: the direct path remains significant, while the indirect path further boosts intentions. Finally, *H7* and *H8* both concern mediation by trust. For *H7*

XSL•FO
RenderX

(*ease of use* → *trust* → *intent*), the indirect path is significant while the direct path is not, signifying a full mediation scenario in which ease of use shapes intentions solely through trust. For *H8* (*perceived usefulness* → *trust* → *intent*), there is evidence of partial mediation, as both the direct and indirect paths are significant. Figure 2 summarizes the overall findings.

Figure 2. The structural model. It illustrates the 5 latent constructs—ease of using DeepSeek, perceived usefulness of DeepSeek, risk perception, trust in DeepSeek, and intent to use DeepSeek for health-related purposes. Each circle represents a latent construct with adjusted R^2 values. The rectangles show specific survey items measuring those constructs. The numeric values on the arrows are standardized path coefficients with *P* values in parentheses. QE captures the possible quadratic effects. QE: quadratic effect.



SEM (Quadratic Effects)

The model explored potential QE nonlinear relationships among the key constructs. The results indicate that the quadratic effect of *ease of use* on *intent to use DeepSeek for health-related purposes* is significant (β=0.08; *P*=.01), suggesting that beyond a certain point, the influence of ease of use on user intentions may intensify or plateau rather than progress linearly. In contrast, the quadratic term of ease of use on trust in DeepSeek was not significant (β=o.04; *P*=.10), indicating that trust levels respond to ease of use in a linear fashion. The quadratic effect of *risk perception* on *intent to use DeepSeek for health-related purposes* emerged as significant (β=−o.11; *P*=.01), implying that moderate risk perceptions might not affect adoption the same way extremely low or very high-risk perceptions do—potentially reflecting a threshold.

Perceived usefulness exhibited a negative but significant quadratic relationship with intent to use (β=–o.07; *P*=.04), suggesting that at very high levels of perceived usefulness, additional increases might yield diminishing or even marginally negative returns on user intentions. By contrast, no significant quadratic effect of Perceived usefulness was detected on trust in DeepSeek (β=o.03; *P*=.37), indicating that trust is likely more directly related to usefulness without an apparent nonlinear inflection point. These results demonstrate that while ease of use, risk perception, and perceived usefulness can shape user intentions curvilinearly, not all model relationships exhibit nonlinear patterns. For constructs such as trust, effects appear to follow a more straightforward linear trajectory. These findings underscore the importance of accounting for potential threshold or plateau effects in user adoption research, particularly in health-oriented applications where attitudes and behaviors can shift dramatically once certain risks or perceived benefits are reached.

Discussion

Summary

This study is among the first to explore user perception of DeepSeek for health-related purposes. Our findings suggest that many users see substantial benefits in using DeepSeek for health-related purposes. The key finding of this research is that ease of use and risk perception can exhibit nonlinear dynamics—these can be interpreted as plateau or threshold points in user attitudes toward DeepSeek. In practical terms, a plateau in ease of use implies that once the interface is sufficiently user-friendly, further refinements may have diminishing effects on adoption. Developers and user-experience designers should therefore balance interface simplicity with the system's perceived depth and reliability; an overly simplistic interface may inadvertently erode trust by making the platform seem superficial or incapable of handling complex health inquiries. Conversely, moderate risk can sometimes spur

XSL•FO
RenderX

vigilance and more thoughtful usage, while extremely high-perceived or extremely low-perceived risk may hinder responsible engagement—by either deterring users outright or breeding complacency. Other notable contribution of this study lies in its comprehensive examination of how ease of use, perceived usefulness, trust, and risk perception influence users' intentions to adopt DeepSeek for health-related purposes.

Ease of Use and Its Impact on Trust in DeepSeek

Our findings confirm that ease of use exerts a significantly positive effect on trust in DeepSeek. From a practical standpoint, developers should focus on streamlined navigation, ensuring that even users with limited digital literacy feel comfortable engaging with DeepSeek. For example, providing step-by-step guidance or context-sensitive help can reinforce the impression that the system is both user-friendly and thoroughly vetted. In contrast to some ChatGPT-based studies, where the interface is perceived as a general question and answer platform, our evidence suggests that for health-specific applications, straightforward interfaces function as a gateway to trust—a critical factor when individuals are dealing with health decisions.

Ease of Use and Intent to Use (Trust Mediation)

The findings regarding the relationship between ease of use and trust in DeepSeek illuminate the complexities of user engagement with AI systems, particularly in health care. Specifically, while ease of use appears to have a significant indirect effect on users' intent to use DeepSeek through the mediator of trust, it resonates with discussions within the larger AI acceptance literature, which posits that usability or interface simplicity alone does not drive adoption without concurrent confidence in the technology's reliability and accuracy [32]. The dynamics uncovered in this exploration emphasize that a merely simplistic interface does not cultivate trust in health-oriented LLMs; users must be assured of the technology's accuracy to engage and rely on it effectively. For instance, users seeking medical advice—such as someone with recurring migraines—may appreciate the user interface of DeepSeek, but their intention to use it consistently will hinge on the platform's perceived medical credibility derived from initial interactions [33]. This perspective is supported by findings from studies on ChatGPT that indicate user trust as a crucial factor influencing sustained interaction, particularly under high-stakes conditions where accuracy is paramount. For example, prior research has shown that while user engagement may be generated through conversational style and interactive query handling, these features fall short in fostering enduring user reliance without a foundational trust in the response quality [34,35] or robust data governance. In comparison, ChatGPT-focused studies highlight that while user engagement may increase through social and conversational interactions, the lack of trust could hinder consistent usage, especially in domains that influence personal well-being [36]. The comparison indicates that while features aimed at ease of use and interactivity can capture initial interest, the health care domain demands a systematic approach that combines usability with robust trust mechanisms to encourage adoption.

The operationalization of trust in AI health systems will likely involve continuous user feedback mechanisms and accuracy assessments to enhance reliability perceptions. Practical applications may include tools that elucidate the rationale behind AI-generated recommendations, thereby fostering a cycle of trust and utility that is essential for successful technology adoption in health care settings.

Risk Perception as a Barrier to Adoption

The finding of risk perception negatively influences intent to use AI applications in health care, such as DeepSeek, is supported by various studies highlighting consumer concerns regarding data privacy and the accuracy of outcomes. This foundational anxiety stems from the sensitive nature of health data, which is subject to strict privacy regulations and potential consequences from AI errors. Studies indicate that perceived risks can deter individuals from using AI-driven health services, emphasizing the necessity for well-defined communication regarding security measures and validations [37]. Studies also acknowledge risk perception as a key predictor of preventive behavioral intentions among health care workers, labeling risk perceptions as a barrier to adoption [38]. For instance, users are often more reassured by clear notifications of compliance with regulations such as Health Insurance Portability and Accountability Act; such risk mitigation strategies can create an environment conducive to adoption. Therefore, developing transparency and clear data-handling policies is crucial for platforms aiming to foster trust amid user apprehensions.

Perceived Usefulness (Direct and Mediated Effects on Trust and Intent)

The assertion that perceived usefulness exerts direct and indirect effects on user behavior is supported by a growing body of literature surrounding AI applications in health care, particularly findings related to ChatGPT and other LLMs. Evidence suggests that tangible benefits, such as improved accuracy and efficiency in information retrieval, significantly motivate user adoption. For instance, studies indicate that the effective performance of ChatGPT in answering medical queries enhances its perceived usefulness, thereby influencing user intention to use the technology [39,40]. The ability to provide relevant and timely medical information underpins the motivation for user engagement with systems such as DeepSeek, signifying that showcasing the background and effectiveness of AI tools is critical for fostering continued usage. Moreover, the notion that perceived usefulness enhances trust aligns with findings in the health AI literature, asserting that when users perceive real value in platform insights, their confidence in the system increases.

In the context of ChatGPT, users often appreciate the quality of responses, which can spur trust; however, this evaluation must also contend with concerns regarding the accuracy and potential misinformation [41]. Reviews of ChatGPT's effectiveness have highlighted its capacity to address diverse questions, reflecting a correlation between perceived usefulness and trust among users [42]. Practical demonstrations of efficacy, such as case studies that illustrate successful outcomes through the application of DeepSeek, resonate with sentiments found in studies involving ChatGPT. For instance, showcasing scenarios where ChatGPT effectively addresses patient inquiries can materially impact users' perceptions of its utility. As users witness consistent performance and tangible benefits, their trust

XSL•FO
RenderX

in the technology deepens. Research indicates that users are more inclined to rely on AI tools when they exhibit high reliability paired with clear demonstrations of past successes [43].

Furthermore, the mediating role of trust in enhancing the relationship between perceived usefulness and the intent to use is underscored by findings that highlight the importance of transparent communication about system performance and limitations. Studies suggest that even when users recognize the immediate utility of a system such as ChatGPT, their commitment to regular use depends on assurances of accuracy and reliability [44]. This illustrates that trust is not merely an adjunct to perceived usefulness but rather a cornerstone of it.

In contrasting DeepSeek with ChatGPT, it becomes evident that while both systems may demonstrate utility, the health context often requires additional layers of demonstrated effectiveness, such as compliance with health care standards or third-party validations. Users making health decisions are more likely to be influenced by the accessibility and clarity of information than those seeking general knowledge, leading to a tighter coupling of perceived usefulness and trust in these sensitive applications. This suggests that developers of health care–oriented AI applications need to focus on operationalizing trust through effective communication strategies, data governance, and continuous assessment of user feedback.

Ease of Use and Perceived Usefulness (Trust Mediation)

Our model indicating full mediation for ease of use → intent and partial mediation for perceived usefulness → intent reveals an interesting relationship between user experience and the adoption of AI technologies, particularly in health-related contexts. This suggests that while perceived ease of use is a critical pathway to intention, perceived usefulness plays a role that not only correlates with intent but also enhances trust. The findings resonate with studies exploring user engagement with AI technologies, highlighting the nuanced nature of technology adoption [45]. In the context of health care AI applications, users often assess functionality and convenience before establishing trust with the technology [46,47]. Studies show that potential users may adopt a system based on perceived usefulness, such as efficiency and practical advantages, even when confidence in the technology's reliability is still developing [48]. Users may appreciate how quickly and accurately an AI system can respond to queries about health-related issues, thereby fulfilling their perceived need for swift information retrieval, which correlates directly with intention to use [49].

Conversely, the aspect of trust and its relationship with perceived usefulness and ease of use is well documented in literature surrounding AI adoption. A robust trust underpinning offers the necessary assurance for sustained engagement with an AI tool. The propensity for individuals to leverage technology relies significantly on their established trust in its reliability, as users frequently remain wary of potential harms associated with incorrect medical information. This caution echoes findings that highlight how user skepticism toward AI systems can hinder adoption, even when perceived usefulness is acknowledged [50].

The dual pathways of adoption in health-related AI involve balance between functionality and trust. Users seeking immediate solutions may prioritize perceived utility—understandably, hospitals and clinics are increasingly looking for AI tools that can aid in diagnosing or managing conditions effectively. However, for deeper and continued engagement, systems must also establish credibility through transparency and accuracy, showcasing validated performance metrics and explicit contextual understanding of the system's capabilities and limitations.

It is imperative for health care–related AI platforms to combine a user-friendly interface—highlighting ease of use—with rigorous communication about their value propositions, as a multitiered approach could cater to both segments of potential users: those motivated primarily by immediate functionality and those seeking strong reliability assurances.

Implications and Recommendations

The significant quadratic terms for ease of use, risk perception, and perceived usefulness underscore that users' intentions to adopt DeepSeek do not always progress linearly. When ease of use displays a positive quadratic relationship with adoption intent, it suggests an initial stage where incremental usability enhancements produce modest gains, followed by a threshold at which further improvements can yield disproportionately greater jumps in adoption. However, past a certain point (a potential plateau), refining the interface may no longer yield significant returns. In essence, once LLMs (DeepSeek) is intuitive enough for typical health inquiries, polishing it further provides only marginal benefit. Therefore, developers should identify the *must-have* design features that bring users up to a comfortable baseline. Past that usability threshold, design resources may be more effectively spent on reliability or improving specific functionalities rather than continuously tweaking the interface. Certain groups (eg, older adults or low digital literacy populations) might require more emphasis on fundamental interface clarity to reach the plateau, whereas advanced users might appreciate specialized features or customization only after basic needs are met.

When risk perception follows a negative quadratic, it indicates that moderate levels of perceived risk may actually promote adoption—perhaps by encouraging vigilance and serious engagement—whereas extremely low or extremely high levels deter usage. Therefore, overemphasizing safety or claiming near-zero risk can paradoxically cause users to discount the tool. They may believe that it is oversimplified or unrealistic for complex scenarios. Conversely, if the system warns of too many uncertainties or potential errors, individuals may deem it too risky. AI stakeholders should aim for transparency that conveys some caution while also stressing evidence of reliability. LLM developers should consider implementing context-based disclaimers that vary by query severity. For minor health checks, a lighter cautionary note may suffice; for serious conditions, stronger reminders encourage users to seek professional verification. Similarly, for perceived usefulness, a negative quadratic implies that while moderately high

XSL•FO
RenderX

usefulness strongly encourages adoption, users can become skeptical if the AI is portrayed as flawless. Exaggerated claims raise suspicion or frustration when real-world performance inevitably falls short. While highlighting success stories or quick turnarounds boosts perceived usefulness, providers must avoid overpromising. Instead, disclosing known limitations, success rates, and typical error margins can actually enhance credibility.

Ultimately, these nonlinear relationships highlight the complexity of user psychology in AI-driven contexts. Rather than presuming a simple the more, the better paradigm, developers and stakeholders should fine-tune system design, risk messaging, and utility claims to steer clear of undesirable plateaus or inversions. By identifying and respecting these inflection points, it becomes possible to sustain healthy, informed user adoption of LLM-based platforms such as DeepSeek.

Limitations

A few constraints must be acknowledged when interpreting these findings. First, this cross-sectional study cannot establish causal direction definitively, although the robust model fit and theoretical grounding strengthen causal inferences. Second, although the sample size was large and relatively balanced across 3 regions, the results may not generalize to other cultural or regulatory contexts where attitudes toward AI-driven health tools differ. Third, the study relied on self-reported measures, which may be subject to social desirability bias or inaccuracies in recalling frequency and manner of use. Our data being collected via web-based panels are inherently composed of volunteers who may differ systematically from the general population (eg, in terms of motivation or digital literacy), thus limiting broad representativeness. True coverage bias remains possible if some segments of the population are less likely to enroll or remain active in web-based panels. While the model covered central factors in technology acceptance (ease of use, usefulness, trust, and risk), other contextual variables—such as domain expertise, quality of health care infrastructure, or cultural norms—may also shape AI adoption in health contexts. Future studies could integrate longitudinal or experimental designs to verify these relationships and examine how usage behaviors evolve with extended AI exposure in real clinical or personal health settings.

Conclusions

The findings indicate that users are generally receptive to DeepSeek when it is perceived as easy to use, useful, and trustworthy. Ease of use significantly enhances trust, which drives the intent to use DeepSeek for health-related purposes.

not only directly boosts both trust and user intent but also has an indirect effect through trust. Conversely, higher risk perception detracts from the intent to use the system. Moreover, the presence of significant quadratic impact suggests that the influence of ease of use, risk perception, and perceived usefulness on user intentions may exhibit threshold or plateau effects. Overall, the validated model underscores the pivotal role of trust as a mediator and highlights the nuanced, nonlinear dynamics shaping user adoption of AI-driven health care tools such as DeepSeek. Future research should replicate this study across different health care settings and populations to validate the generalizability of these findings. Expanding the model to include additional mediators or moderators—such as privacy concerns, cultural differences, and prior experience with AI—might provide deeper insights into user adoption dynamics. Moreover, longitudinal or experimental designs could help establish causality and track changes in user attitudes over time. Finally, given the observed quadratic effects, further investigation into threshold and plateau phenomena could refine our understanding of how perceptions of ease of use, risk, and usefulness evolve as users become more familiar with AI-driven health care tools.

Conflicts of Interest

None declared.

References

1. Yeh KC, Chi JA, Lian DC, Hsieh SK. Evaluating interfaced LLM bias. Presented at: 35th Conference on Computational Linguistics and Speech Processing (ROCLING 2023); Oct 20-21, 2023; Taipei City, Taiwan.
2. Reeves B, Nass C. The Media Equation: How People Treat Computers, Television, and New Media Like Real People: Center for the Study of Language and Information; 1996, Vol. 10:19-36.
3. Bender EM, Koller A. Climbing towards NLU: on meaning, form, and understanding in the age of data. Presented at: 58th Annual Meeting of the Association for Computational Linguistics; Jul 5-10, 2020. [doi: 10.18653/v1/2020.acl-main.463]
4. Falletta-Cowden N. Using Large Language Models in the Experimental Analysis of Persuasion and Bias: University of Nevada; 2024.
5. Glikson E, Woolley AW. Human trust in artificial intelligence: review of empirical research. Acad Manage Ann 2020 Jul;14(2):627-660. [doi: 10.5465/annals.2018.0057]
6. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. Manage Sci 1989 Aug;35(8):982-1003. [doi: 10.1287/mnsc.35.8.982]
7. Castelfranchi C, Falcone R. Trust Theory: A Socio-Cognitive and Computational Model: John Wiley & Sons, Ltd; 2010. [doi: 10.1002/9780470519851]
8. Lun L, Zetian D, Hoe TW, Juan X, Jiaxin D, Fulai W. Factors influencing user intentions on interactive websites: insights from the technology acceptance model. IEEE Access 2024;12:122735-122756. [doi: 10.1109/ACCESS.2024.3437418]

XSL•FO
RenderX

9. Lai CY, Cheung KY, Chan CS, Law KK. Integrating the adapted UTAUT model with moral obligation, trust and perceived risk to predict ChatGPT adoption for assessment support: a survey with students. Comput Educ Artif Intell 2024 Jun;6:100246. [doi: 10.1016/j.caeai.2024.100246]
10. Sisk BA, Antes AL, Burrous S, DuBois JM. Parental attitudes toward artificial intelligence-driven precision medicine technologies in pediatric healthcare. Children (Basel) 2020 Sep 20;7(9):32962204. [doi: 10.3390/children7090145] [Medline: 32962204]
11. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manage Sci 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
12. Venkatesh V, Bala H. Technology acceptance model 3 and a research agenda on interventions. Decis Sci 2008 May;39(2):273-315. [doi: 10.1111/j.1540-5915.2008.00192.x]
13. Rondan-Cataluña FJ, Arenas-Gaitán J, Ramírez-Correa PE. A comparison of the different versions of popular technology acceptance models. Kybernetes 2015 May 5;44(5):788-805. [doi: 10.1108/K-09-2014-0184]
14. Dormann C, Biddle R. Understanding game design for affective learning. Presented at: FuturePlay08; Nov 3-5, 2008; Toronto, Ontario, Canada. [doi: 10.1145/1496984.1496992]
15. Kahneman D, Tversky A. Chapter 6: prospect theory: an analysis of decision under risk. In: Handbook of the Fundamentals of Financial Decision Making: World Scientific Publishing Co Pte Ltd; 2013:99-127. [doi: 10.1142/9789814417358_0006]
16. Maassen O, Fritsch S, Palm J, et al. Future medical artificial intelligence application requirements and expectations of physicians in German university hospitals: web-based survey. J Med Internet Res 2021 Mar 5;23(3):e26646. [doi: 10.2196/26646] [Medline: 33666563]
17. Richardson JP, Curtis S, Smith C, et al. A framework for examining patient attitudes regarding applications of artificial intelligence in healthcare. Digit Health 2022;8:20552076221089084. [doi: 10.1177/20552076221089084] [Medline: 35355806]
18. Ng ZQP, Ling LYJ, Chew HSJ, Lau Y. The role of artificial intelligence in enhancing clinical nursing care: a scoping review. J Nurs Manag 2022 Nov;30(8):3654-3674. [doi: 10.1111/jonm.13425] [Medline: 34272911]
19. Choudhury A, Shamszare H. Investigating the impact of user trust on the adoption and use of ChatGPT: survey analysis. J Med Internet Res 2023 Jun 14;25:e47184. [doi: 10.2196/47184] [Medline: 37314848]
20. Gaczek P, Pozharliev R, Leszczyński G, Zieliński M. Overcoming consumer resistance to AI in general health care. J Interact Mark 2023 May;58(2-3):321-338. [doi: 10.1177/10949968221151061]
21. Reddy S, Allan S, Coghlan S, Cooper P. A governance model for the application of AI in health care. J Am Med Inform Assoc 2020 Mar 1;27(3):491-497. [doi: 10.1093/jamia/ocz192] [Medline: 31682262]
22. Tobia K, Nielsen A, Stremitzer A. When does physician use of AI increase liability? J Nucl Med 2021 Jan;62(1):17-21. [doi: 10.2967/jnumed.120.256032] [Medline: 32978285]
23. Kock N. Harman's single factor test in PLS-SEM: checking for common method bias. DAPJ 2021;2(2):1-6 [FREE Full text]
24. Hair Jr JF, Sarstedt M, Ringle CM, Gudergan SP. Advanced issues in partial least squares structural equation modeling. In: Handbook of Market Research: Springer; 2017. [doi: 10.1007/978-3-319-05542-8_15-1]
25. Streukens S, Leroi-Werelds S. Bootstrapping and PLS-SEM: a step-by-step guide to get more out of your bootstrap results. Eur Manag J 2016 Dec;34(6):618-632. [doi: 10.1016/j.emj.2016.06.003]
26. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. J Mark Res 1981 Feb;18(1):39-50. [doi: 10.1177/002224378101800104]
27. Cheung GW, Wang C. Current approaches for assessing convergent and discriminant validity with SEM: issues and solutions. Acad Manag Proc 2017 Aug(1):12706. [doi: 10.5465/AMBPP.2017.12706]
28. Wolf EJ, Harrington KM, Clark SL, Miller MW. Sample size requirements for structural equation models: an evaluation of power, bias, and solution propriety. Educ Psychol Meas 2013 Dec;76(6):913-934. [doi: 10.1177/0013164413495237] [Medline: 25705052]
29. Anderson JC, Gerbing DW. Structural equation modeling in practice: a review and recommended two-step approach. Psychol Bull 1988;103(3):411-423. [doi: 10.1037/0033-2909.103.3.411]
30. Mathur S, Tewari A, Vishnoi S, Agarwal V. An integrated model to predict students' online learning behavior in emerging economies: a hybrid SEM–ANN approach. J Int Educ Bus 2025 Jan 13;18(1):102-126. [doi: 10.1108/JIEB-01-2024-0004]
31. Midi H, Sarkar SK, Rana S. Collinearity diagnostics of binary logistic regression model. J Interdiscip Math 2010 Jun;13(3):253-267. [doi: 10.1080/09720502.2010.10700699]
32. Li JJ, Wu LR, Qi JY, Zhang YX, Wu ZY, Hu SB. Determinants affecting consumer trust in communication with AI chatbots: the moderating effect of privacy concerns. J Organ End User Comput 2023;35(1):1-24. [doi: 10.4018/JOEUC.328089]
33. Jermutus E, Kneale D, Thomas J, Michie S. Influences on user trust in healthcare artificial intelligence: a systematic review. Wellcome Open Res 2022;7:65. [doi: 10.12688/wellcomeopenres.17550.1]
34. Hans S, Kumar B, Parihar V, Singh S. Human-AI collaboration: understanding user trust in ChatGPT conversations. Int J Sci Res Eng Manage 2024;08(1):1-13. [doi: 10.55041/IJSREM27929]
35. de Brito Duarte R, Correia F, Arriaga P, Paiva A. AI trust: can explainable AI enhance warranted trust? Hum Behav Emerg Technol 2023 Oct 31;2023:1-12. [doi: 10.1155/2023/4637678]

XSL•FO
RenderX

36. Shahsavar Y, Choudhury A. User intentions to use ChatGPT for self-diagnosis and health-related purposes: cross-sectional survey study. JMIR Hum Factors 2023 May 17;10(1):e47564. [doi: 10.2196/47564] [Medline: 37195756]
37. Steerling E, Siira E, Nilsen P, Svedberg P, Nygren J. Implementing AI in healthcare-the relevance of trust: a scoping review. Front Health Serv 2023;3(1211150):1211150. [doi: 10.3389/frhs.2023.1211150] [Medline: 37693234]
38. Morowatishaifabad MA, Zare Sakhvidi MJ, Gholianavval M, Masoudi Boroujeni D, Alavijeh MM. Predictors of hepatitis B preventive behavioral intentions in healthcare workers. Saf Health Work 2015 Jun;6(2):139-142. [doi: 10.1016/j.shaw.2014.12.001] [Medline: 26106514]
39. Pan A, Musheyev D, Bockelman D, Loeb S, Kabarriti AE. Assessment of artificial intelligence chatbot responses to top searched queries about cancer. JAMA Oncol 2023 Oct 1;9(10):1437-1440. [doi: 10.1001/jamaoncol.2023.2947] [Medline: 37615960]
40. Torun C, Sarmis A, Oguz A. Is ChatGPT an accurate and reliable source of information for patients with vaccine and statin hesitancy? Medeni Med J 2024 Mar 21;39(1):1-7. [doi: 10.4274/MMJ.galenos.2024.03154] [Medline: 38511678]
41. Bhattacharyya M, Miller VM, Bhattacharyya D, Miller LE. High rates of fabricated and inaccurate references in ChatGPT-generated medical content. Cureus 2023 May;15(5):e39238. [doi: 10.7759/cureus.39238] [Medline: 37337480]
42. Thirunavukarasu AJ, Hassan R, Mahmood S, et al. Trialling a Large Language Model (ChatGPT) in general practice with the applied knowledge test: observational study demonstrating opportunities and limitations in primary care. JMIR Med Educ 2023 Apr 21;9:e46599. [doi: 10.2196/46599] [Medline: 37083633]
43. Masanneck L, Schmidt L, Seifert A, et al. Triage performance across large language models, ChatGPT, and untrained doctors in emergency medicine: comparative study. J Med Internet Res 2024 Jun 14;26:e53297. [doi: 10.2196/53297] [Medline: 38875696]
44. Hu JM, Liu FC, Chu CM, Chang YT. Health care trainees' and professionals' perceptions of ChatGPT in improving medical knowledge training: rapid survey study. J Med Internet Res 2023 Oct 18;25:e49385. [doi: 10.2196/49385] [Medline: 37851495]
45. Schnall R, Bakken S, Rojas M, Travers J, Carballo-Dieguez A. mHealth technology as a persuasive tool for treatment, care and management of persons living with HIV. AIDS Behav 2015 Jun;19 Suppl 2(0 2):81-89. [doi: 10.1007/s10461-014-0984-8] [Medline: 25572830]
46. Bankins S, Ocampo AC, Marrone M, Restubog SLD, Woo SE. A multilevel review of artificial intelligence in organizations: implications for organizational behavior research and practice. J Organ Behavior 2024 Feb;45(2):159-182. [doi: 10.1002/job.2735]
47. Choudhury A. Factors influencing clinicians' willingness to use an AI-based clinical decision support system. Front Digit Health 2022;4(920662):920662. [doi: 10.3389/fdgth.2022.920662] [Medline: 36339516]
48. Kerstan S, Bienefeld N, Grote G. Choosing human over AI doctors? How comparative trust associations and knowledge relate to risk and benefit perceptions of AI in healthcare. Risk Anal 2024 Apr;44(4):939-957. [doi: 10.1111/risa.14216] [Medline: 37722964]
49. Rony MKK, Numan SM, Johra FT, et al. Perceptions and attitudes of nurse practitioners toward artificial intelligence adoption in health care. Health Sci Rep 2024 Aug;7(8):e70006. [doi: 10.1002/hsr2.70006] [Medline: 39175600]
50. Esmaeilzadeh P. Use of AI-based tools for healthcare purposes: a survey study from consumers' perspectives. BMC Med Inform Decis Mak 2020 Jul 22;20(1):170. [doi: 10.1186/s12911-020-01191-1] [Medline: 32698869]

Abbreviations

AI: artificial intelligence
AVE: average variance extracted
CMB: common method bias
LLM: large language model
PLS-SEM: Partial Least Squares Structural Equation Modeling
SEM: structural equation modeling

Edited by A Kushniruk, Y Senathirajah; submitted 19.02.25; peer-reviewed by E Urena, MK Milic, M Saremi; revised version received 19.03.25; accepted 07.04.25; published 26.05.25.

Please cite as:
Choudhury A, Shahsavar Y, Shamszare H
User Intent to Use DeepSeek for Health Care Purposes and Their Trust in the Large Language Model: Multinational Survey Study
JMIR Hum Factors 2025;12:e72867
URL: https://humanfactors.jmir.org/2025/1/e72867
doi:10.2196/72867

XSL•FO
RenderX

© Avishek Choudhury, Yeganeh Shahsavar, Hamid Shamszare. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 26.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Prioritizing Trust in Podiatrists' Preference for AI in Supportive Roles Over Diagnostic Roles in Health Care: Qualitative Interview and Focus Group Study

Mohammed A Tahtali[1,2*], MSc; Chris C P Snijders[2*], Prof Dr; Corné W G M Dirne[1*], Prof Dr; Pascale M Le Blanc[3*], Prof Dr

[1]Department of Industrial Engineering & Management, Fontys University of Applied Sciences, Eindhoven, The Netherlands

[2]Department of Industrial Engineering & Innovation Sciences, Human Technology Interaction group, Eindhoven University of Technology, Eindhoven, The Netherlands

[3]Department of Industrial Engineering & Innovation Sciences, Human Performance Management group, Eindhoven University of Technology, Eindhoven, The Netherlands

[*]all authors contributed equally

Corresponding Author:
Mohammed A Tahtali, MSc
Department of Industrial Engineering & Management
Fontys University of Applied Sciences
De Rondom 1, 5612 AP
Eindhoven
The Netherlands
Phone: 31 8850 79388
Email: m.tahtali@fontys.nl

Abstract

Background: As artificial intelligence (AI) evolves, its roles have expanded from helping out with routine tasks to making complex decisions, once the exclusive domain of human experts. This shift is pronounced in health care, where AI aids in tasks ranging from image recognition in radiology to personalized treatment plans, demonstrating the potential to, at times, surpass human accuracy and efficiency. Despite AI's accuracy in some critical tasks, the adoption of AI in health care is a challenge, in part because of skepticism about being able to rely on AI decisions.

Objective: This study aimed to identify and delve into more effective and acceptable ways of integrating AI into a broader spectrum of health care tasks.

Methods: We included 2 qualitative phases to explore podiatrists' views on AI in health care. Initially, we interviewed 9 podiatrists (7 women and 2 men) with a mean age of 41 (SD 12) years and aimed to capture their sentiments regarding the use and role of AI in their work. Subsequently, a focus group with 5 podiatrists (4 women and 1 man) with a mean age of 54 (SD 10) years delved into AI's supportive and diagnostic roles on the basis of the interviews. All interviews were recorded, transcribed verbatim, and analyzed using Atlas.ti and QDA-Miner, using both thematic analysis for broad patterns and framework analysis for structured insights per established guidelines.

Results: Our research unveiled 9 themes and 3 subthemes, clarifying podiatrists' nuanced views on AI in health care. Key overlapping insights in the 2 phases included a preference for using AI in supportive roles, such as triage, because of its efficiency and process optimization capabilities. There is a discernible hesitancy toward leveraging AI for diagnostic purposes, driven by concerns regarding its accuracy and the essential nature of human expertise. The need for transparency and explainability in AI systems emerged as a critical factor for fostering trust in both phases.

Conclusions: The findings highlight a complex view from podiatrists on AI, showing openness to its application in supportive roles while exercising caution with diagnostic use. This result is consistent with a careful introduction of AI into health care in roles, such as triage, in which there is initial trust, as opposed to roles that ask the AI for a complete diagnosis. Such strategic adoption can mitigate initial resistance, gradually building the confidence to explore AI's capabilities in more nuanced tasks, including diagnostics, where skepticism is currently more pronounced. Adopting AI stepwise could thus enhance trust and acceptance across a broader range of health care tasks, aligning technology integration with professional comfort and patient care standards.

XSL•FO
RenderX

(JMIR Hum Factors 2025;12:e59010) doi:10.2196/59010

KEYWORDS

AI's role in health care; decision-making; diabetes and podiatrists; trust; AI; artificial intelligence; qualitative; foot; podiatry; professional; experience; attitude; opinion; perception; acceptance; adoption; thematic; focus group

Introduction

Background

As artificial intelligence (AI) continues to advance, its potential to revolutionize health care is unquestionable. AI systems are increasingly capable of performing tasks that range from routine data analysis to complex decision-making processes that were once the sole domain of human experts. These technologies offer the promise of more accurate predictions [1], more precise patient condition explanations [2], autonomous diagnoses [3], and personalized treatment plans [4]. Despite these capabilities, there remains an evident hesitancy among health care professionals and patients alike to embrace AI-driven recommendations fully. This skepticism is partly rooted in worries about AI's capacity to grasp patients' nuanced, individual scenarios [5], to integrate the crucial human elements of empathy and understanding [6,7], and to deal with complex ethical considerations in health care [8].

The literature on trust in AI highlights concerted efforts to foster trust by exploring strategies to address trust issues between AI systems and users. For example, some studies suggest that redefining AI as a supportive adjunct to professionals rather than a standalone solution could promote acceptance [9,10]. In addition, showcasing instances where AI surpasses human experts in specific tasks may enhance humans' trust in AI [5,9,11]. Emphasizing transparency and providing understandable explanations of AI processes are crucial, with evidence showing users how AI works can greatly diminish skepticism [12-15]. For example, applying explainable AI (XAI) techniques to explain the basis of AI-generated predictions can, in certain contexts, enhance human trust and understanding [12]. Although XAI can potentially be valuable in increasing trust and understanding, the specific implementation of XAI (eg, the complexity and relevance of the explanation) is crucial in achieving those goals [16]. However, despite these positive developments, fully integrating AI into practice remains challenging, and professionals and laypersons often still prefer human judgment and intuition over AI recommendations [6,7,10,17].

AI's potential applications within health care are vast and multifaceted. Its well-established ability to detect diseases such as cancer and predict various health conditions highlights its substantial role in diagnostics. Not only does AI excel in recognizing patterns within medical images for early disease detection, offering a level of precision that can sometimes surpass traditional methods [3,18,19], but also its influence extends beyond diagnostics. AI assumes a crucial supportive role in health care operations, such as triaging in emergency departments where algorithms prioritize care based on urgency and outcomes, thereby optimizing resource allocation and ensuring timely patient attention [20]. Furthermore, AI's capacity for enhancing operational efficiency is evident across a spectrum of health care activities. It predicts bed availability and automates the conversion of patient consultations into concise summaries, thereby freeing health care professionals from administrative work. This enables more patient-centered care and marks AI as a transformative tool in health care [21].

This study seeks to delve deeper into health care providers' perspectives on AI's current and potential role and value in their field. While previous studies often focus on assessing trust in already developed AI systems, this study takes a different approach by involving health care professionals from the outset in identifying potential applications of AI. Early engagement allows us to determine where AI is anticipated to be most valuable, fostering a more collaborative and user-centered integration that aligns with professionals' needs and expectations. By engaging directly with health care professionals who are the prospective users of these AI applications, the study seeks to pinpoint areas where AI integration would be most advantageous and eagerly anticipated by practitioners. Through this inquiry, we aspire to pave the way for more effective and user-endorsed integration of AI in health care, aligning technological advancements with health care professionals' real-world needs and preferences.

Objective

Our study explores health care providers' perceptions of AI's potential benefits in patient care and operational efficiency, with examples including risk assessments and diagnostics. The effectiveness of AI technologies depends on their advanced capabilities and, more crucially, on their acceptance by health care professionals. Without their acceptance, achieving a truly successful implementation becomes only partial, leading to the largely unrealized potential of AI. Thus, we aim to understand these professionals' perspectives on AI adoption, emphasizing the development of trust and integration strategies and identifying approaches that position AI as a valuable tool complementing human expertise, not replacing it [10]. Consequently, the central research question we propose is as follows: How can AI be effectively integrated into practice in a way that ensures health care professionals' acceptance and trust in its advice?

We adopt a qualitative approach to address our research question, focusing on the FootCheck (RondOm Podiatrists) app (in Dutch: Voetencheck app). This app, specifically designed for the daily monitoring of patients with diabetes in the Netherlands, sends daily push notifications to guide these patients through a foot screening process, featuring instructional videos and questions. This enables patients to submit photos to podiatrists to evaluate potential issues such as wounds or ulcers. During its successful pilot phase, which occurred amid the COVID-19 pandemic, the app was used by over 1700 individuals, leading to more than 43,000 screenings [22]. Recognizing the growing interest of stakeholders in leveraging

XSL•FO
RenderX

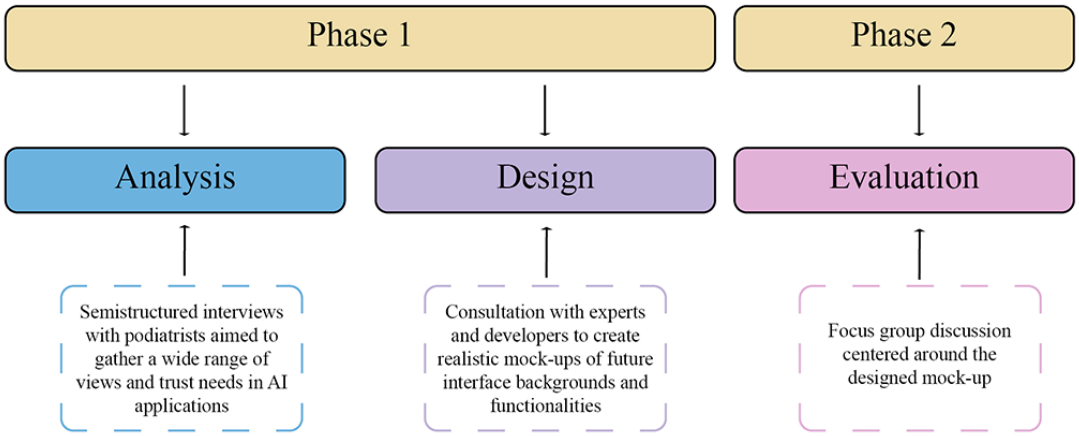
AI to enhance efficiency and health care quality, we believe the FootCheck app provides an ideal basis for exploring AI's potential. It is important to note that the app still lacks AI integration, highlighting the potential benefits of AI for future enhancements. Our study includes interviews with diverse podiatrists, encompassing current and potential future app users. These interviews aim to understand their experiences and perspectives regarding using the app as it is and the prospects for integrating AI into the screening process of future versions of the app.

Methods

Design and Setting

We conducted a qualitative study structured in 2 phases to investigate trust and acceptance in AI. In phase 1, we sought to capture diverse perspectives and needs from podiatrists regarding their trust in and acceptance of AI technologies through semistructured interviews. The insights gathered from these interviews served as input to the mock-up design process. Phase 2 aimed to further explore specific AI applications through a focus group session. Figure 1 provides a visual representation of these phases.

Figure 1. A visual presentation of the 2 phases of the study. AI: artificial intelligence.



The FootCheck App and Decision-Making Process

The practical application of AI in health care settings we studied was the FootCheck app, developed for RondOm Podiatrists. As one of the largest podiatry chains in the Netherlands, RondOm Podiatrists introduced this app during the COVID-19 pandemic to ensure continuous care for their patients. The app has since been used to facilitate daily monitoring of foot health for patients with diabetes, allowing patients to submit photographs for screening. The workflow from patient submission to professional evaluation is illustrated in Figure 2.

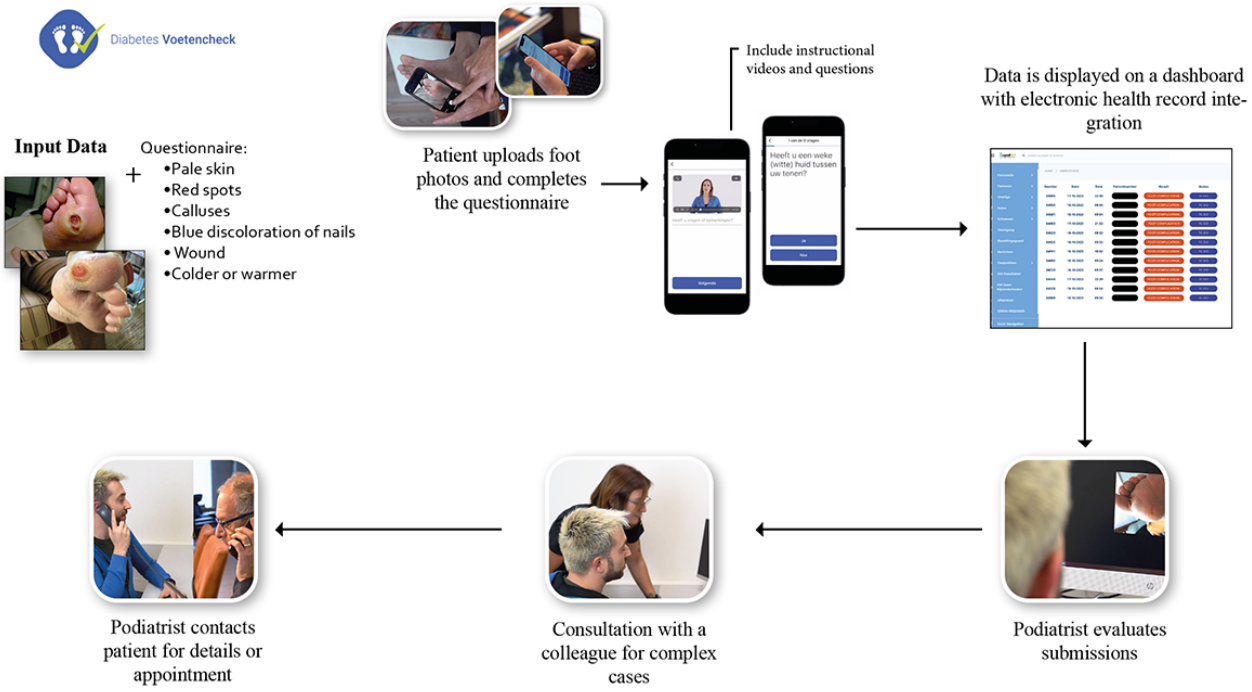
RondOm Podiatrists are keenly interested in the potential of AI to enhance their app and streamline decision-making processes. The system does not use AI currently, but there are numerous potential areas where AI could provide substantial benefits. For example, AI could be instrumental in assessing the suitability of submitted photos for evaluation. This capability would significantly reduce the workload on professionals by automatically filtering out low-quality or irrelevant images, ensuring that only those suitable for review are presented. This efficiency gain is crucial as it allows podiatrists to focus their expertise where it is most needed rather than on preliminary sorting tasks.

Moreover, AI could assist in prioritizing cases on the basis of urgency. Currently, podiatrists must manually open and evaluate all submitted photos, which is time-consuming and potentially less effective in quickly identifying urgent cases. By implementing AI to perform a preselection of instances based on predetermined criteria of urgency, the system could ensure that patients requiring immediate attention are identified and prioritized, thereby optimizing patient care and resource allocation.

Another critical area in which AI could have a significant role is aiding in the diagnosis of foot conditions. The increasing number of patients with diabetes and a relative shortage of podiatrists present a growing challenge. AI can alleviate some of this burden by providing preliminary assessments based on the analysis of photos. While the professional community may have reservations about relying entirely on AI for diagnoses, its role in offering support and preliminary assessments could be invaluable. This assistance could help bridge the gap in service provision, ensuring that more patients receive timely and effective care.

XSL•FO
RenderX

Figure 2. Workflow from patient submission to professional evaluation.



Phase 1: Participants and Recruitment Process

In the first phase of our study, we engaged in interviews with 9 podiatrists from 4 distinct practices, aiming to gather diverse insights into the use and perception of the health care app in question. The actual use of the current app varied considerably among participants, encompassing some who used the app daily and others who, while not yet current users, were familiar with its functionalities. Furthermore, participants' professional expertise spanned different aspects of podiatric care, including both foot care for patients with diabetes and general foot health services.

The recruitment of participants was facilitated through a collaborative effort with the Fontys University of Applied Sciences, one of the leading institutions for podiatry education in the Netherlands. This collaboration was essential for 2 reasons. First, Fontys University maintains strong ties with its alumni, many of whom stay actively engaged with the university through professional networks, alumni events, and workshops. Second, Fontys University's Health Innovations & Technology research group maintains regular collaborations with health care institutions, ensuring strong connections within the field. Through these partnerships, we recruited a diverse group of motivated podiatrists who voluntarily participated in the study. By using Fontys University's network and its respected position within the field, we ensured the participation of podiatrists with varying levels of experience and expertise across different areas of podiatric care, contributing to a representative sample for our study. To accommodate our participants' varying schedules and preferences, the interviews were conducted using a hybrid approach; some were held in person, taking advantage of the natural setting of the participants' practice environments, while others were carried out through the Teams (Microsoft Corp) platform.

Data Collection

We started an extensive literature review on trust in AI, both generally and within the health care context, from which we identified key topics. These topics were categorized into 4 groups as follows: knowledge of AI, trust in humans and AI advice, AI's added value to daily tasks, and factors affecting trust in AI. On the basis of these categories and codes, we formulated initial interview questions. Next, we sought feedback on these initial questions from experts in human and machine interaction. We incorporated their feedback and refined the questions to ensure their relevance and comprehensiveness. To validate the clarity and comprehensiveness of these questions, we conducted a preliminary interview with an experienced podiatrist. This pilot interview served to ensure the questions were understandable and relevant for the prospective interviewees, and it also provided an opportunity to identify any missing questions from the practical perspective of podiatry. The resulting interview guide is detailed in Multimedia Appendix 1.

Participants were briefed on the study's objectives—namely, to explore their perceptions of AI, identify factors influencing trust in AI among podiatrists, examine the impact of trust on AI adoption, and assess their needs for trusting AI in patient care. They were requested to share demographic information via a link before the interview sessions. The specific interview questions were withheld in advance to preserve the authenticity of responses. The interviews were conducted from July to September 2023.

Consent for recording was obtained from all podiatrists, with the sessions captured either on a voice recorder or through the Teams platform, depending on the interview format. These recordings were transcribed verbatim, using Amberscript [23] for efficiency and accuracy. For interviews conducted on the

XSL•FO
RenderX

Teams platform, we leveraged the platform's transcription functionality. Each interview spanned approximately 30 to 45 minutes, with podiatrists having provided explicit consent, either written or digital, for participating and recording their responses. They were also assured of their right to withdraw from the study at any time.

Data Analysis of the Semistructured Interviews

We used a dual-coding strategy, where 2 researchers independently analyzed the interview transcripts to ensure a robust examination of the data. Researcher 1 (MAT), a PhD candidate with a solid background in qualitative research methodologies, and researcher 2 (Ruben Gloudemans) used distinct software tools for coding—Atlas.ti (version 11; Atlas.ti Scientific Software Development GmbH) and QDA-Miner (version 6.0; Provalis Research). This divergence in tool use stemmed from their prior experiences and the software licenses available.

Following the thematic analysis by Braun and Clarke [24], we initiated our analysis by thoroughly reading the transcripts to gain an initial understanding and verify their accuracy. First, we labeled the transcripts without predefined categories, using an inductive approach leading to an extensive list of initial codes directly derived from the data. For example, codes such as human interaction, urgency ranking, efficiency, risk detection, and false advice were formed. We then examined these initial codes to determine if they overlapped with the topics we had previously identified from the literature and combined them if needed. For example, urgency ranking and risk detection were put together since both are related to the topic added value of AI. In this step, we started a deductive approach, where we applied existing theoretical frameworks to refine our codes into coherent themes. We repeated this analytical process several times, refining our approach each time to ensure the emergence of concrete themes. This process yielded a preliminary set of themes.

Subsequently, our individually identified themes were compared in a meeting, discovering significant overlap. For some themes where our interpretations initially diverged, we reached consensus through discussion, using inductive and deductive reasoning. These discussions revealed that, despite using different terminologies, we were referring to the same thematic content. Our thematic analysis was refined further through this iterative process of inductive and deductive coding.

Mock-Up Design

Guided by our research question on how AI can effectively integrate into the FootCheck app's existing screening process with user acceptance, we sought the podiatrists' insights on how they would like to receive advice and in which forms to ensure their confidence in it. Various suggestions emerged, such as explaining the reasoning behind the advice, photos pinpointing potential complications on the feet, and color-coding systems for triage. However, the first phase did not conclusively determine which forms of applications and the level of transparency would foster greater trust.

To further explore these possibilities, we organized a focus group featuring mock-ups. Presenting a realistic representation was crucial, particularly considering that currently, the app and its screening process do not use AI.

1. The first mock-up was designed to assess the effectiveness of AI in a supportive role, specifically in scheduling and prioritizing patient consultations based on an array of time indicators.
2. The second mock-up aimed to explore the potential of AI in diagnosing conditions by analyzing photographs submitted by users and identifying common foot complications such as ulcers or calluses.

We aimed to delve deeper into assessing these AI roles by examining their functionality and acceptability among podiatrists. We crafted 3 distinct design variations for each mock-up to achieve this, guided by needs and concerns expressed in the first interviews with the podiatrists. Our primary reason for this approach was to facilitate a broad discussion by presenting transparency levels varying from low to medium to high. This method allowed us to not only compare the discussions across different AI applications but also delve into the discussions within each level of transparency. Thereby, we aimed to explore the range in podiatrists' perceptions and acceptance extensively.

In designing these mock-ups, we used the current interface environment to ensure familiarity for users. In addition, we consulted with experts during the design process to ensure the mock-ups were practical and reflective of current best practices in the field. For an illustration of mock-up 1 with a high transparency level, see Figure 3; for the diagnostic mock-up set in a medium transparency, refer to Figure 4. For a detailed overview of all variants, refer to Multimedia Appendix 2.

XSL•FO
RenderX

Figure 3. Mock-up 1: high-transparency variant.

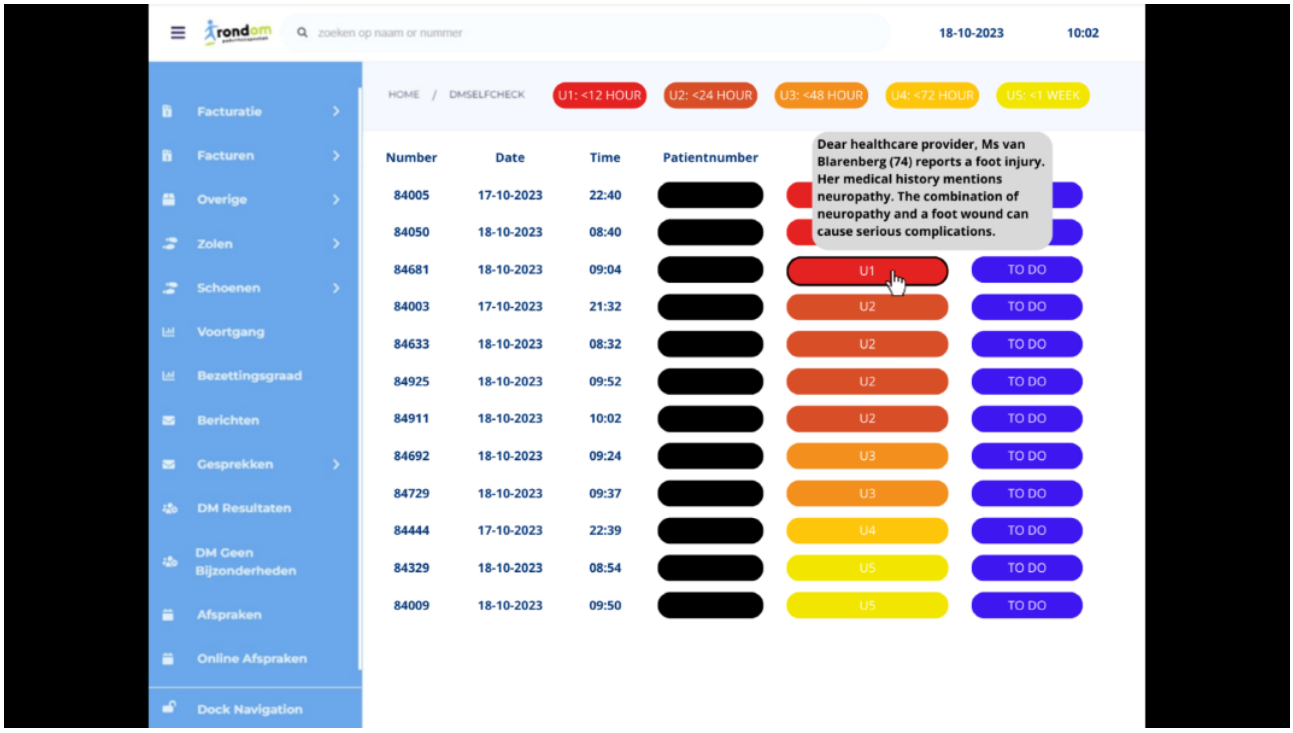


Figure 4. Mock-up 2: medium-transparency variant.



Phase 2: Participants and Recruitment Process

Phase 2 was aimed at enriching our understanding and validating the outcomes of the first phase. Initially, we used the same approaches that had effectively recruited participants for phase 1, including relying on Fontys University's network and existing partnerships within the podiatry field. We also reached out to contacts from our previous interviews. However, these efforts alone were insufficient to meet our needs for new participants for phase 2. To broaden our reach, we engaged with potential candidates on social media platforms such as LinkedIn (LinkedIn Corp) and Facebook (Meta Platforms, Inc) and attended podiatrist conferences to connect with new professionals. Through these combined efforts, we recruited 4 new podiatrists who had not been involved in the first phase and reinvited one participant from phase 1. The focus group session was organized at the Fontys University of Applied Sciences' inspiration space in December 2023.

XSL•FO
RenderX

Procedure for Focus Group and Data Collection

To evaluate our mock-ups, we created an interview guide on essential topics derived from the findings of phase 1: trust in AI's role, transparency preferences, and AI's effect on decision-making, outlined in Multimedia Appendix 3. During the focus group session, participants were thoroughly briefed on the structure of the session that was divided into two 45-minute segments with a break in between. The first session explored mock-up 1, while the second delved into mock-up 2.

The roles of MAT and Ruben Gloudemans within the session were clearly defined for participants; MAT served as the observer, summarizing discussions and posing questions, whereas Ruben Gloudemans led the session and presented the materials. Participants provided written consent by signing a consent form.

After a brief introduction, participants were provided two A3-sized sheets, each showcasing 3 mock-up variants highlighting different transparency levels. Before initiating the group discussions, participants were asked to individually rank these variants from 1 to 3, with 1 being their top preference, and to write down comments next to each variant explaining their choices. This step was designed to minimize mutual influence and ensure that each participant's initial reactions and thoughts were captured independently. Following this individual reflection, detailed group discussions were held, focusing on participants' rationale regarding their preferences and trust in the AI advice shown in each variant.

Data Analysis Focus Group

All participants attended in person for the data collection of phase 2. Initially planned as two 45-minute segments with a 15-minute break in between, the session ultimately lasted 70 minutes in a single stretch. This was shorter than anticipated, primarily because the first segment was concluded earlier than expected, and participants opted to proceed without taking the scheduled break.

We adopted a more structured coding method as we progressed to the second phase, building on the codes and themes identified in the first phase. This shift was enabled by implementing a thematic framework approach, a systematic strategy that helped us to systematically organize and interpret the data [25].

After verifying transcription accuracy and familiarizing ourselves with the transcript, we developed an analytical framework. This framework included predefined themes from phase 1: AI's roles, transparency levels, data reliability, trust development, and the significance of human interaction. This step laid the groundwork for indexing, where we systematically tagged the data with our framework's themes. Next, through charting, we distilled participants' responses into structured summaries, aligning them with the thematic framework. In the indexing and charting phase, we organized the data into themes linked to our research goals, creating visual charts. This approach helped us spot patterns and differences, uncovering insights that enhanced our understanding.

Ethical Considerations

No medical devices were used in the studies, and no physical interventions were performed on participants. The image used in mock-up 2 originated from publicly available internet sources and was appropriately licensed. Consequently, under regulations stipulated by the Dutch Medical Research Involving Human Subjects Act (WMO), the necessity for ethics approval from a national ethics committee was circumvented. However, ethical clearance was proactively obtained from the internal ethics committee of Fontys University of Applied Sciences, documented under project numbers 050324 and 080324. There were no conflicts of interest between the participants and the FootCheck app. None of the participants had any financial, developmental, or other professional ties to the app.

Results

Phase 1: Semistructured Interviews

Participants

Participant demographics for phase 1 are listed in Table 1. In total, 9 individuals, all with backgrounds in podiatry, participated in the first phase of our study. A distinguishing characteristic of this cohort is their active engagement in clinical practice, that is, their extensive practical experience.

XSL•FO
RenderX

Table 1. Participants' characteristics from semistructured interviews in phase 1 (n=9).

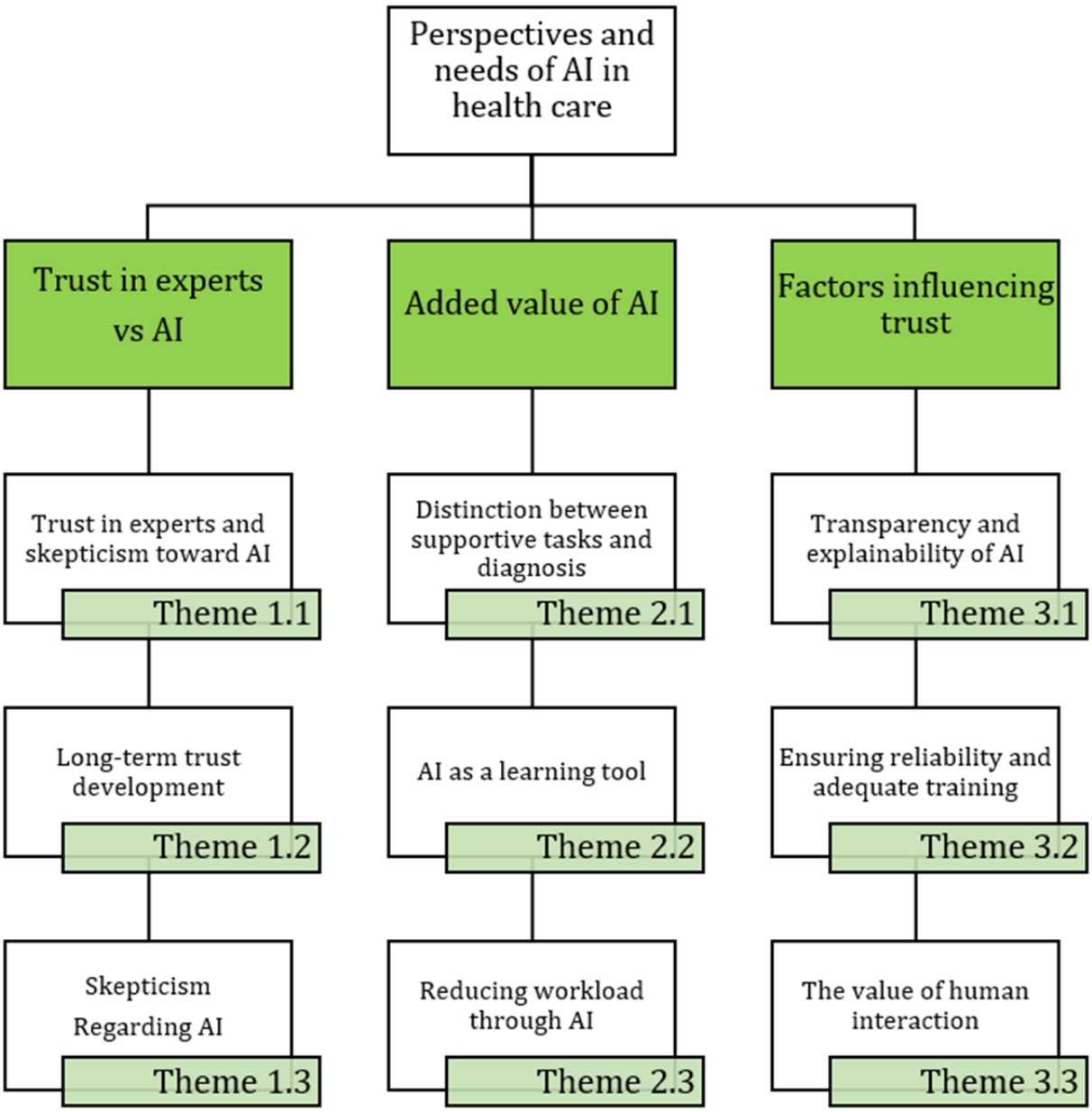
Characteristics	Participants, n (%)
Sex	
Female	7 (78)
Male	2 (22)
Age group (y)	
20-30	1 (11)
31-40	2 (22)
41-50	3 (33)
51-60	2 (22)
61-70	1 (11)
Education level	
Bachelor's	4 (44)
Master's	4 (44)
PhD	1 (11)
Specialized education: foot care for patients with diabetes	
Yes	4 (44)
No	5 (56)
Experience (mo)	
0-6	1 (11)
6-12	2 (22)
12-24	4 (44)
24-36	2 (22)
Experience with the app	
Yes	2 (22)
No	7 (78)

Themes From the Interviews

The interviews with podiatrists revealed 3 critical themes regarding podiatrists' perspectives and requirements for AI integration in the FootCheck app and related decision-making process: trust in human experts versus in AI, the added value of AI, and the factors influencing trust in AI. Each theme is further divided into 3 subthemes. A detailed overview of these key themes can be found in Figure 5.

XSL•FO
RenderX

Figure 5. Generated themes and subthemes from phase 1: perspectives of experienced podiatrists on the FootCheck app and decision-making process. AI: artificial intelligence.



Theme 1: Trust in Experts Versus AI

In the initial phase of the interview, we began by asking participants about their understanding of AI and their trust in it compared with their colleagues (ie, experts). We aimed to uncover how AI impacts their trust levels and the underlying reasons for these views.

Theme 1.1: Trust in Experts and Skepticism Toward AI

Podiatrists expressed trust in the advice from expert colleagues, on the basis of their extensive knowledge and experience. They recognized that making mistakes is part of being human and considered it normal for colleagues to err every now and then. Conversely, when receiving advice from AI, direct trust is not immediately established for most podiatrists. They maintained a critical stance, emphasizing the need to verify the accuracy of AI-generated advice:

> *I trust a colleague, sometimes blindly, because they have much experience.* [Participant 7, female, age group 61-70 years]
>
> *I see the added value of AI advice. But I still want to check it or have something alongside it.* [Participant 2, female, age group 51-60 years]
>
> *I do have trust in advice from colleagues. I've considered practical situations, and nine times out of ten, I agree with them just fine. And there are always nuanced differences because you see things slightly differently with your experience, perspective, and view. But let me put it this way: I firmly trust it.* [Participant 8, male, age group 41-50 years]

XSL•FO
RenderX

Theme 1.2: Long-Term Trust Development

Participants indicated that AI may initially struggle to provide accurate advice due to the limited data available and the learning phase required for the model. They also mentioned their initial lack of experience with using AI, which leads to uncertainty about its performance. However, participants felt their trust can increase as they gain more experience with the AI model:

> *In the beginning, I was sceptical, but as it sees and evaluates more photos, my trust grows.* [Participant 9, female, age group 20-30 years]
>
> *Yes, AI needs to prove itself first, and then it can add value.* [Participant 7, female, age group 61-70 years]
>
> *Indeed, my trust in AI would significantly increase if, over time, the technology evolves to a stage where the AI system is fully operational and applicable.* [Participant 6, female, age group 41-50 years]

Theme 1.3: Skepticism Regarding AI

Participants expressed significant concerns about the potential errors an AI system might make and the severe consequences that could arise from such mistakes. In addition, they highlighted a notable limitation of AI: its inability to conduct physical examinations on patients. Such examinations are often essential for a thorough understanding of a patient's foot health, indicating a gap in AI's capabilities in health care settings:

> *Yes, if a patient says: "I have pain in my foot, for example, in the ankles." Well, AI can't see that or physically feel or move it.* [Participant 9, female, age group 20-30 years]
>
> *We're not talking about a patient getting advice to buy Shoe A, and then they say it's too big and buy another shoe.... We're talking about serious outcomes, like the loss of a foot, a leg, or even a life, and that's a significant responsibility.* [Participant 4, female, age group 31-40 years]
>
> *The consequences are quite significant if you miss something there.* [Participant 7, female, age group 61-70 years]

Theme 2: Added Value of AI

Overview

In exploring AI's potential impact on podiatry, we engaged participants in discussions about integrating AI into their everyday practice. This dialogue revealed 3 insightful subthemes, each shedding light on AI's role in enhancing podiatric care. The following sections delve into these emergent themes.

Theme 2.1: Distinction Between Supportive Tasks and Diagnosis

Nearly all participants agreed that AI should be used for support tasks such as triage. They highlighted that significant time is spent retrieving missing information not provided by patients and addressing issues with incorrect photo submissions that impede accurate assessment. They were of the opinion that these tasks could be efficiently managed by AI, thereby streamlining the process and improving overall efficiency. However, opinions varied on whether AI can diagnose foot conditions in patients with diabetes. Most participants were of the opinion that AI is incapable to do so, emphasizing the need for physical examinations by human experts to come up with an accurate diagnosis. They pointed out that photos fail to reveal crucial details, such as hidden wounds beneath calluses, which are not directly visible in pictures. Furthermore, participants noted the importance of physically assessing the feet for swelling and temperature variations (ie, warmth or coldness) that cannot be discerned from photos alone:

> *If it can highlight specific areas of concern, suggesting "pay attention here, this is alarming," then I do place my trust in it.* [Participant 3, female, age group 41-50 years]
>
> *AI can't make complete diagnoses; you need to hold the foot in your hands.* [Participant 2, female, age group 41-50 years]
>
> *I'm considering integrating a feature for predicting situations where an immediate consultation is crucial and distinguishing those from cases where it might not be as urgent.* [Participant 1, male, age group 41-50 years]
>
> *If it integrates well with practical needs, like photos can be easily uploaded directly into the file, making follow-up simpler. Yes, so it actually assists in your work.* [Participant 6, female, age group 41-50 years]

Theme 2.2: AI as a Learning Tool

Participants acknowledged the potential of AI as a valuable learning tool, particularly for identifying unfamiliar diseases and suggesting innovative treatment strategies. They expressed optimism about AI's ability to broaden their diagnostic and therapeutic horizons:

> *Maybe the AI will uncover a disease pattern I haven't encountered before, offering me an opportunity to expand my knowledge.* [Participant 1, male, age group 41-50 years]
>
> *Hey, AI sees things differently, and that's something I can learn from. I do see the added value in that.* [Participant 5, female, age group 20-30 years]
>
> *In this light, I believe that when an AI indicates, "this is the optimal treatment," it becomes incredibly fascinating. It prompts a collective inquiry: "Why hadn't this occurred to us before? Is this indeed the correct approach? Or is there already a hospital implementing this method?"* [Participant 4, female, age group 31-40 years]

Theme 2.3: Reducing Workload Through AI

Participants highlighted the health care sector's challenge of increasing workloads, worsened by a decreasing number of podiatrists and a growing number of patients with diabetes. There was widespread agreement on AI's ability to alleviate this load. By automating processes such as triage, AI can significantly reduce the workload, freeing podiatrists to focus on the more crucial parts of patient care:

XSL•FO
RenderX

> *As the pressure continues to increase, I see AI as a great tool to indeed perform a kind of triage to ensure that the workload doesn't become overwhelming.* [Participant 3, female, age group 41-50 years]
>
> *But yes, to be completely honest, looking towards the future, we definitely need to reduce the workload; otherwise, we'll end up with even fewer people having to help more individuals.* [Participant 4, female, age group 31-40 years]

Theme 3: Factors Influencing Trust

Overview

Finally, we delved into the pivotal elements that shape trust in AI-provided advice. We also probed when participants might be inclined to defer to AI's judgment over their own. This investigation unveiled 3 subthemes, each shedding light on the interaction between humans and AI systems.

Theme 3.1: Transparency and Explainability of AI

All participants agreed that clear explanations and transparency are essential, considering them as fundamental to trusting AI. They wished to comprehend the methodology AI uses to arrive at its recommendations. They proposed several methods to enhance transparency, such as providing concise explanations of the criteria AI uses for its evaluations, statistical performance data, or visual aids such as photos with overlaying layers or color coding:

> *I believe factors like the colour of the callus, the size of the area, its location, and history also play a role. Has there been a wound in that spot before? Yes, things like that.* [Participant 8, male, age group 41-50 years]
>
> *For example, I want to see a number of photos of the foot along with some explanation about them. I want to understand how the evaluation system is designed and the data it uses.* [Participant 2, female, age group 51-60 years]
>
> *It increases my confidence in knowing the sources from which AI retrieves its information or the ones I've provided to it.* [Participant 1, male, age group 41-50 years]

Theme 3.2: Ensuring Reliability and Adequate Training

Participants were of the opinion that AI should be advanced enough from the beginning, emphasizing that a more extensive database of photos would increase their trust in the system. They would like AI to provide consistent advice, ideally on par with a fellow podiatrist. Training was highlighted as a crucial element for enhancing AI's performance. In addition, understanding who the designers are and the data used for training is considered necessary. If participants know the designers are experts in their field, they feel more confident in the AI tool's reliability:

> *Well, if you've only input three things into it, I won't trust it. But if you've input 300,000 things, then at some point, you really know better.* [Participant 2, male, age group 41-50 years]
>
> *Well, I'm not quite sure how that works, but if that's the case, then the people involved in setting it up need to have the right background knowledge, the right experience with the subject; they need to be specialists, so to speak.* [Participant 2, female, age group 51-60 years]

Theme 3.3: The Value of Human Interaction

Participants emphasized that future AI applications must not compromise their interaction with patients, highlighting each patient's uniqueness and background. They cautioned that AI may not always account for these personal nuances, underlining the importance of maintaining the human touch in health care:

> *I still believe that personal human contact will never disappear.* [Participant 3, female, age group 41-50 years]
>
> *All those kinds of things. Yes, they're slightly different for everyone.* [Participant 1, male, age group 41-50 years]
>
> *But you still maintain human contact between the client and yourself, allowing them to easily ask additional questions, for instance, to you.* [Participant 5, female, age group 61-70 years]

Moving From Phase 1 to Phase 2

The findings from phase 1 revealed that AI can significantly improve various practical aspects of podiatry, particularly by alleviating workloads and serving as a learning tool. A clear preference was observed for using AI in supportive tasks rather than diagnostic roles. This reflects a cautious attitude of podiatrists toward AI-generated advice and raised the question: Will AI support be more readily accepted in supportive than in diagnostic tasks, and what factors influence this difference in acceptance?

Therefore, phase 2 of our study sought to explore the specific applications of AI in supportive versus diagnostic roles within podiatry in detail. Given the ongoing reluctance toward its role in diagnostics, this study aimed to understand the elements that can foster trust in AI for supportive tasks. By introducing a new group of podiatrists, phase 2 of our study aimed to gather fresh insights and assess whether the findings observed in phase 1 align with new professional perspectives.

Phase 2: Focus Group

Participants

Table 2 outlines an overview of the demographics of the focus group participants. The group consisted of 5 participants, 4 of whom were new podiatrists who engaged in the discussion. This group consisted of specialists in foot care for patients with diabetes, who primarily provide daily consultations to these patients.

XSL•FO
RenderX

Table 2. Participants' characteristics from focus groups in phase 2 (n=5).

Characteristics	Participants, n (%)
Sex	
Female	4 (80)
Male	1 (20)
Age group (y)	
41-50	2 (40)
51-60	2 (40)
61-70	1 (20)
Education level	
Bachelor's	2 (40)
Master's	3 (60)
Specialized in foot care for patients with diabetes	
Yes	5 (100)
No	0 (0)
Experience (y)	
20-30	4 (80)
31-40	1 (20)
Experience with the app	
Yes	2 (40)
No	3 (60)

In phase 1 of our study, we opted for a flexible approach, allowing themes to emerge directly from the transcribed data. This exploratory phase enabled us to capture many insights directly from participants' experiences and perspectives. Moving into phase 2, we shifted toward a more structured approach, using themes, which were relevant for phase 2 and predefined in phase 1. This transition allowed us to refine our analysis, focusing on the detailed examination of specific, predetermined themes to gain deeper insights and understanding of the topics at hand.

For a clear overview, we first present the key findings from the participants about various transparency levels for the 2 mock-ups in Table 3. Subsequently, in Table 4, we display the other themes carried forward from phase 1 for further analysis in phase 2, ensuring a comprehensive and structured exploration of the data.

XSL•FO
RenderX

Table 3. Thematic framework analysis for key findings and preferences of transparency levels in AI[a] applications.

Theme: transparency and explainability of AI	Mock-up 1: AI support (triage and scheduling)	Mock-up 2: AI diagnosis (photo analysis)
Low (variant A)	Podiatrists perceive this variant as a "black box," with unclear determination of urgency and relevance of the presented information	Feedback indicates general dissatisfaction and a lack of clarity on how AI reaches its conclusions
First preference	—[b]	P4 and P5 (n=2)
Second preference	—	—
Third preference	P1-P5 (n=5)	P1-P3 (n=3)
Medium (variant B)	Podiatrists express concerns over the reliability of patient feedback as input for AI, pointing out a lack of crucial information such as wound location	While the text is deemed acceptable, this variant raises questions about the necessity of repeating information
First preference	—	—
Second preference	P1-P5 (n=5)	P1-P5 (n=5)
Third preference	—	—
High (variant C)	Preference is given to detailed information, including specific medical details and clear patient identification	Podiatrists appreciate the clarity and added layer of understanding this variant provides but question the necessity of certain statistics, such as a reliability percentage
First preference	P1-P5 (n=5)	P1-P3 (n=3)
Second preference	—	—
Third preference	—	P4 and P5 (n=2)

[a]AI: artificial intelligence.

[b]Not applicable.

Table 4. Thematic framework analysis for AI[a] applications and predefined themes.

Themes	Participant 1 (female; age group 51-60 y)	Participant 2 (female; age group 41-50 y)	Participant 3 (male; age group 41-50 y)	Participant 4 (female; age group 51-60 y)	Participant 5 (female; age group 61-70 y)
Distinction between supportive tasks and diagnosis	Prefers AI for triage due to speed and accuracy. Skeptical about diagnostic capability without physical examination	—[b]	Values AI's ability to prioritize photos based on urgency, enhancing patient care efficiency	Believes AI is incapable of handling complex diagnostic tasks, such as differentiating between types of wounds	Asserts that computers cannot replace human nuanced observation in diagnostics, emphasizing AI's use in triage
Ensuring reliability and adequate training	Values the volume of feedback, similar to how reviews inform vacation choices	Trust is reinforced by a large volume of photos, indicating data reliability	Importance of data diversity, including training AI on poor-quality photos	Questions the repetitive mention of the photo database size suggest it is only necessary in training	Trust depends on knowing the AI's advice is based on an extensive and diverse photo database
Long-term trust development	—	Trust in the algorithm grows over time with proven accuracy	Initial skepticism evolves into trust after multiple positive experiences	—	Anticipates a future shift in acceptance of AI's advice, especially without a photo, within a decade
The value of human interaction	—	—	—	Asserts that physical tools and examination are indispensable for accurate diagnostics	Stresses the irreplaceable role of human involvement in making accurate diagnoses

[a]AI: artificial intelligence.

[b]Not applicable.

For the frameworks (Tables 3 and 4), instead of using quotes from focus group participants, we chose to provide summaries of the key findings, following the common practice in thematic framework analysis. By creating summaries, we were able to handle the complex discussion, which involved a record with 7 speakers, more easily.

XSL•FO
RenderX

Interpretations of Thematic Framework Analysis

On the basis of the thematic framework analysis of podiatrists' preferences for AI integration in clinical settings, a clear pattern emerges regarding the desired application and form of transparency in AI systems for triage. Podiatrists prefer using AI in triage, emphasizing the importance of transparency that goes beyond just being open. They are looking for a kind of transparency that involves clinically relevant and detailed information—such as patient history, specific conditions such as the presence of vascular diseases, and exact wound locations—that significantly augment the decision-making process:

> *The added value for me is that the ranking is done by the AI, so photos are prioritised based on urgency.* [Participant 3, male, age group 41-50 years]

> *My preference is for triage; an AI can perform it much faster than a human. Well, I believe that AI is capable of making a risk assessment.* [Participant 1, female, age group 41-50 years]

> *Key abnormalities must be quickly recognised and addressed, achievable through triage.* [Participant 5, female, age group 61-70 years]

In the domain of triage, skepticism toward patient-reported data as the sole basis for AI-generated advice is evident (eg, medium transparency mock-up 1). The concern centers on the reliability of such data, highlighting a preference for AI systems that incorporate a comprehensive training dataset to inform their assessments. This reflects a broader apprehension regarding the accuracy of AI recommendations based on potentially incomplete or inaccurate patient self-assessments:

> *I'm not entirely convinced by the patients' responses. If they say they have a wound on their feet, then yes, I might believe it, but if they don't have a wound, I'm sceptical.* [Participant 1, female, age group 41-50 years]

> *Patients often do not have a clear understanding of complications.* [Participant 3, male, age group 41-50 years]

Podiatrists recognize the potential benefits but are cautious regarding the use of AI for diagnostic applications, particularly those involving the analysis of photographic evidence of foot conditions. This caution is based on the understanding that certain conditions cannot be fully diagnosed through images alone, emphasizing the need for AI systems to supplement rather than replace traditional diagnostic methods:

> *Computers can't replace humans in diagnostics, especially not for tasks requiring nuanced observation, like identifying underlying issues beneath a callus that AI might miss.* [Participant 5, female, age group 61-70 years]

> *AI is incapable of making diagnoses. AI cannot handle complex conditions like distinguishing between warm and cold wounds.* [Participant 4, female, age group 51-60 years]

> *Even a human can't make a diagnosis just from an image. The foot needs to be examined in person and potentially opened up for an accurate assessment.* [Participant 1, female, age group 41-50 years]

The desired level of transparency in diagnostic AI varies among podiatrists. While some value additional explanatory features, such as annotated images or textual diagnostics, others caution against the potential for such features to overly focus attention on a specific area, possibly leading to the overlooking of important symptoms in other areas. This division highlights the need for a careful balance in using AI; while it is beneficial to receive detailed, actionable insights, it is crucial to ensure that these insights do not overshadow the importance of using one's own professional knowledge and experience in decision-making:

> *Works well for me; the layer aids in understanding the basis of the text.* [Participant 3, male, age group 41-50 years]

> *Also, here, the text is okay. I also like the layer, so now I know what the text is referring to.* [Participant 1, female, age group 41-50 years]

> *The explanatory text doesn't interest me much; I'll take it on faith what it's based on.* [Participant 4, female, age group 51-60 years]

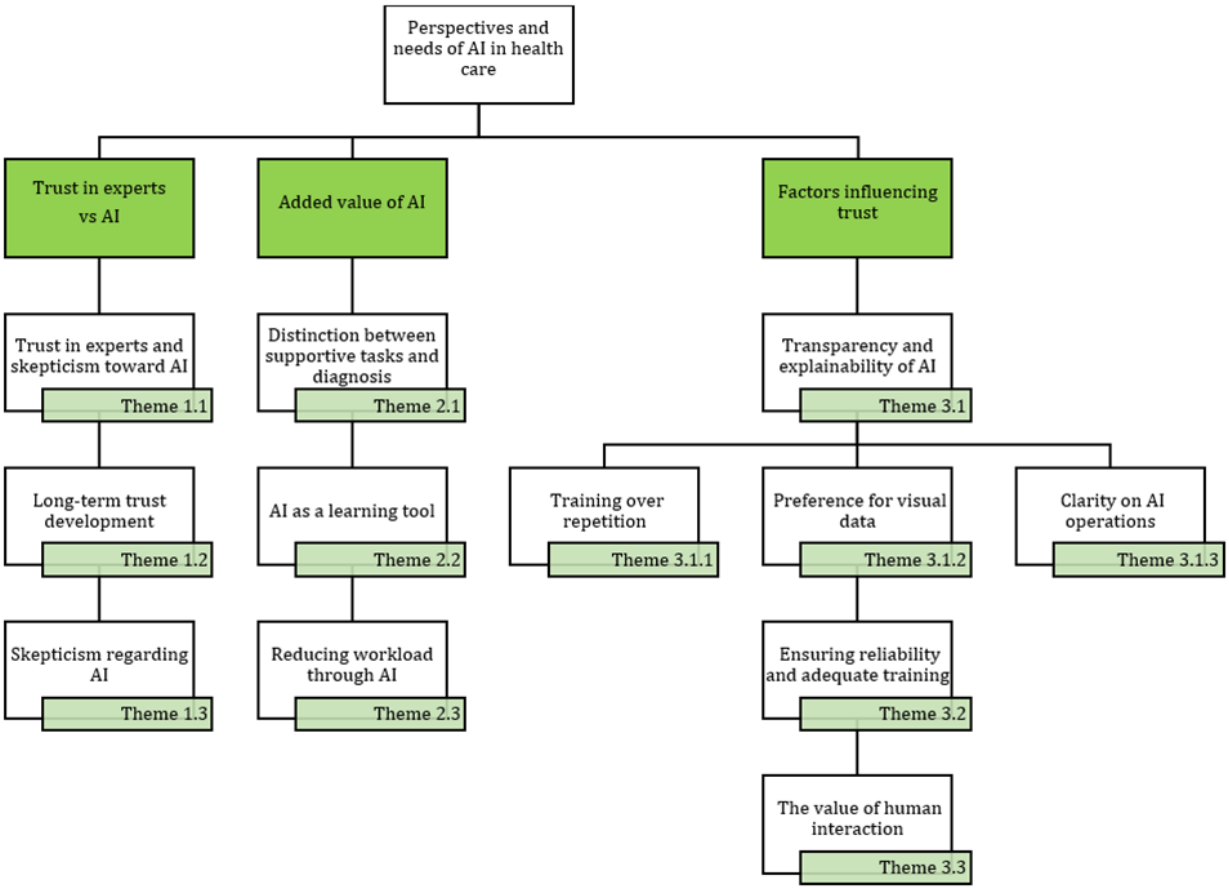
> *And that layer, I can see that for myself, too. Adding a layer can also be risky; I see even more things, and now AI dictates where I should look. A less experienced podiatrist might then focus solely on that.* [Participant 5, female, age group 61-70 years]

In addition to emphasizing transparency in integrating AI into clinical podiatry practice, our analysis has other predefined themes within the framework (Table 4). These themes enrich our understanding of how AI can be effectively applied in triage and diagnostics, emphasizing the need for clear, actionable insights and the importance of trust and data reliability as foundational elements. Reflecting on the preferences expressed by podiatrists, it becomes apparent that alongside transparency and explainability, considerations around data reliability, human interaction, trust development, and trust establishment align with sentiments from the first phase of our study.

However, we have defined 3 additional themes that resonated strongly with the initial input for transparency and explainability. Figure 6 presents a comprehensive overview of all the themes, including the 3 subthemes (ie, 3.1.1: Training Versus Repetition; 3.1.2: Preference for Visual Data; and 3.1.3: Clarity on AI Operations) and their implications for integrating AI into podiatry. The following sections detail these additional subthemes further.

XSL•FO
RenderX

Figure 6. Generated themes and subthemes from phases 1 and 2: perspectives of experienced podiatrists on the FootCheck app and decision-making process. AI: artificial intelligence.



Subtheme 3.1.1: Training Versus Repetition

While participants in phase 1 primarily highlighted concerns among podiatrists regarding data reliability, phase 2 revealed a nuanced understanding of how and when this information should be introduced. This is particularly relevant during the initial training sessions where podiatrists are first taught how to use the AI system. Feedback on mock-up 2 indicated that emphasizing data reliability in every AI recommendation is not always seen as necessary. Some participants supported this viewpoint, while others preferred having this information included. In addition, there was consensus that explanations should offer insights beyond basic knowledge, aiming to augment understanding rather than merely repeat well-known information:

> *In the introduction (training), mention somewhere that the data is based on 5000 photos, but that doesn't need to be in every piece of advice.* [Participant 4, female, age group 51-60 years]
>
> *I expect an AI to be backed by a vast photo database, so it's unnecessary to reiterate that point for me.* [Participant 5, female, age group 61-70 years]
>
> *The text is helpful; it lets me know what the AI is based on.* [Participant 3, male, age group 41-50 years]

Subtheme 3.1.2: Preference for Visual Data

There is a significant inclination toward incorporating visual data alongside algorithmic risk assessments. Participants consider combining photos with textual advice crucial, enhancing their ability to make knowledgeable decisions. The need for visual confirmation to solidify trust in AI-generated advice is highlighted, with participants expressing reservations about relying entirely on text-based recommendations. Although color-coded elements are appreciated for their clarity, the facility to view photos directly via pop-ups is deemed essential for building confidence in the advice provided:

> *The colour coding in mock-up 1 is helpful, but I'd prefer to see the photos appear as pop-ups.* [Participant 4, male, age group 41-50 years]
>
> *Just text, without a photo, doesn't inspire my trust.* [Participant 3, female, age group 51-60 years]
>
> *It seems like a small ask to quickly review the photo for a better trust in the advice.* [Participant 4, female, age group 51-60 years]

Subtheme 3.1.3: Clarity on AI Operations

Podiatrists highlight the necessity for transparency regarding the process behind AI-generated advice, which is essential for building trust and enabling users to understand when and how to rely on these recommendations. Opinions vary on the best way to present this information, with some finding objective data in photos and layered visuals helpful for understanding the advice's basis, as seen in mock-up 2. Others question the interpretability of numerical data, such as "73%," without seeing the patient in person:

XSL•FO
RenderX

> *Now, with a photo and layer, the information is more objective; you know what the text is about.* [Participant 3, male, age group 41-50 years]
>
> *I appreciate it, as then I know what it's based on, mock-up 2.* [Participant 1, female, age group 51-60 years]

Discussion

AI in Supportive Roles Versus Diagnostic Tasks

Our study aimed to discern the multifaceted roles AI could assume in health care, specifically focusing on how AI can be effectively integrated into practice in a way that ensures health care professionals' acceptance and trust in its advice. The findings from our studies suggest a preference for deploying AI in supportive roles. For instance, AI's use in triaging processes is evident; by evaluating submitted photos for urgency, AI can advise health care providers in prioritizing patient consultations. This enables a more efficient allocation of resources, ensuring that patients requiring urgent care receive it promptly, while less-critical cases are appropriately prioritized. Such applications align with and enrich suggestions from the literature, including studies by Verma et al [10] and Longoni et al [5], suggesting the supportive use of AI to foster trust within health care settings. However, a contrasting perspective is presented in another experiment by Longoni et al [5], where preference leans toward human expertise over automated systems for triage tasks. This discrepancy from our findings may stem from differing viewpoints, with the study results of Longoni et al [5] being based on patient perspectives, whereas our study was performed among health care providers.

Another critical insight from our study highlights health care professionals' hesitancy to embrace AI for diagnostic purposes, a sentiment that aligns with previous research findings [5,7,9,10,18]. While there is acknowledgment of AI's proficiency in critical functions, such as making diagnoses, with results that match or even surpass human expertise, there remains skepticism among participants about AI's current proficiency in performing these tasks accurately. The observed reluctance may stem from unfamiliarity with the full scope of AI's capabilities, as evidenced by the fact that particularly participants without prior AI experience expressed reservations. Moreover, the conversations uncovered that identifying conditions such as diabetic foot complications poses distinct obstacles for AI. The participants pointed out that superficial evaluations, such as analyzing diabetic wounds via photographs, could be inadequate without the physical removal of calluses to reveal the actual problem, an action AI is incapable of without a physical examination. This underlines the importance of direct human interaction in specific diagnostic processes, highlighting limitations in the potential application of AI.

Algorithm Aversion Among Podiatrists

The observed phenomenon of algorithm aversion among podiatrists, particularly their preference for human expertise over AI-generated recommendations, reflects a trend noted across various fields where AI competes with human judgment [26-29]. For example, Önkal et al [30] demonstrated a similar inclination among forecasters to favor expert opinions in forecasting decisions. A notable factor contributing to this aversion is the concern over mistakes made by AI systems [26], with our podiatrists voicing worries about the grave repercussions of such errors, including the risk of limb amputation. Although the significance of decisions ranges from health care to forecasting tasks in logistics, the underlying dread of AI-induced errors remains pervasive.

However, in our studies, this aversion was primarily directed at diagnostic tasks rather than supportive functions. Participants expressed confidence in the expertise and experience of their colleagues but also showed openness to considering AI-generated advice for tasks such as triage. This nuanced view suggests a complex relationship between podiatrists and AI, where trust may be contingent on the specific task at hand and the perceived reliability of AI in enhancing rather than replacing human judgment.

Transparency and Explainability

The discussions surrounding transparency and explainability in our studies mirrored the diversity of perspectives found in the broader literature on these subjects [12,14,19]. The need for explanation and transparency depends significantly on podiatrists' individual preferences. Our study provides 3 key insights regarding the implementation of explainability and transparency in AI systems.

First, precise and valuable explanations that avoid redundancy are essential for enhancing user understanding and trust in AI. Second, visual aids alongside textual advice can boost confidence in AI recommendations more than text alone, by making the motivation behind complex decisions more accessible. Finally, it is important to remind users to balance AI advice with their professional judgment, highlighting the crucial role of human oversight in AI decision-making.

However, these discussions were particularly animated regarding the scenario involving AI-generated diagnoses, where participants displayed heightened criticality toward the information received from AI. This suggests that individuals are particularly cautious about relying on AI for diagnostic tasks, scrutinizing the transparency and explainability of such systems with greater intensity. Interestingly, these issues were less extensively debated in contexts involving AI for supportive tasks. This observation highlights an area for future research.

Development of Trust Over Time

Most participants indicated a higher level of trust in their colleagues than in AI systems, suggesting that trust in AI might evolve positively with increased exposure and evidence of AI's efficacy. However, this remains speculative, as our participants have not extensively experimented with AI in their professional practice. The notion that repeated interactions with AI could enhance trust aligns with existing literature [31,32]. This area, highlighting the dynamic interplay between growing familiarity with AI and the shifting landscape of trust and reliance, remains underexplored in our studies and represents a fruitful avenue for further research.

XSL•FO
RenderX

Strengths and Limitations

Our research is strengthened by several methodological approaches that enhance the credibility of our findings. The employment of 2 independent researchers for parallel data coding in both studies significantly contributed to the reliability of our results, as evidenced by the substantial agreement in their coding outcomes [33].

The relatively small sample sizes of 9 podiatrists in the first phase and 5 in the second phase may seem to be a limitation for generating general conclusions. However, qualitative research often reaches a point of data saturation, where small samples are already sufficient to uncover dependable themes. In our study, we observed that no new codes or themes emerged after 7 interviews in the first phase, indicating that data saturation had been reached. This aligns with the findings of Guest et al [34], who noted that most codes are identified within the first 6 interviews. Guest et al [35] also supported this by providing a systematic method for assessing saturation, further validating our sample size.

The significant expertise of the podiatrists involved, that is, all specialists in the area considered, adds depth to our insights and strengthens the credibility of our findings. In addition, the second phase, with mostly new participants, confirmed the results of phase 1. By focusing on how AI might be more accepted in supportive tasks compared with diagnostic tasks, we were able to further validate the earlier findings and ensure that no new themes emerged, indicating that saturation had been reached.

Conclusions

In conclusion, the findings reveal podiatrists' complex attitudes toward AI in health care, demonstrating openness to AI assistance for supportive tasks while being cautionary regarding its application for diagnostics. This underscores the necessity for developers and policy makers to prioritize trust-building strategies, starting with opportunities that allow podiatrists to experience firsthand how AI can support their role. Such initial experiences could clarify AI's benefits and capabilities to users, reduce perceived threats, and facilitate a smoother transition toward acceptance. With a foundation of trust established through the application of AI for supportive tasks, podiatrists may become more receptive to exploring AI's potential for diagnostic tasks. Gradually expanding AI's role from foundational tasks, including decision support, to areas requiring deeper trust, such as diagnostics, could enhance acceptance and integration of AI technologies in podiatric practice.

Acknowledgments

The authors extend their heartfelt gratitude to all individuals who generously shared their insights and experiences, making this research possible. Special thanks are extended to Ruben Gloudemans for his invaluable assistance with conducting interviews, engaging in thorough analyses and participating in wide-ranging discussions that significantly enriched this work. The authors appreciate RondOm Podiatrists for providing access to the FootCheck app, an essential resource for our study. This work is part of the project Vertrouwen in Algoritmes (project number 023.017.109), under the auspices of the Promotiebeurs voor Leraren program, generously funded by the Dutch Research Council (NWO). Additionally, this study was part of the Technology Support for Diabetes project, which is subsidized by SIA RAAK EXZ.EXZ.01.004. The research efforts of the authors were greatly supported by these contributions, for which they are profoundly grateful.

Data Availability

The corresponding author will make the datasets supporting these studies available upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide used in phase 1.
[DOCX File , 20 KB - humanfactors_v12i1e59010_app1.docx]

Multimedia Appendix 2
Mock-ups used in phase 2.
[DOCX File , 1981 KB - humanfactors_v12i1e59010_app2.docx]

Multimedia Appendix 3
Interview guide used in phase 2.
[DOCX File , 16 KB - humanfactors_v12i1e59010_app3.docx]

References

XSL•FO
RenderX

1. Bazazeh D, Shubair R. Comparative study of machine learning algorithms for breast cancer detection and diagnosis. In: Proceedings of the 5th International Conference on Electronic Devices, Systems and Applications. 2017 Presented at: ICEDSA '16; December 6-8, 2016; Ras Al Khaimah, United Arab Emirates p. 1-4 URL: https://ieeexplore.ieee.org/document/7818560 [doi: 10.1109/icedsa.2016.7818560]
2. Ayers JW, Poliak A, Dredze M, Leas EC, Zhu Z, Kelley JB, et al. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. JAMA Intern Med 2023 Jun 01;183(6):589-596 [FREE Full text] [doi: 10.1001/jamainternmed.2023.1838] [Medline: 37115527]
3. Eriksen AV, Möller S, Ryg J. Use of GPT-4 to diagnose complex clinical cases. NEJM AI 2024 Jan 11;1(1):24. [doi: 10.1056/aip2300031]
4. Meng Y, Sun J, Qu N, Zhang G, Yu T, Piao H. Application of radiomics for personalized treatment of cancer patients. Cancer Manag Res 2019;11:10851-10858 [FREE Full text] [doi: 10.2147/CMAR.S232473] [Medline: 31920394]
5. Longoni C, Bonezzi A, Morewedge CK. Resistance to medical artificial intelligence. J Consum Res 2019;46(4):629-650. [doi: 10.1093/jcr/ucz013]
6. Promberger M, Baron J. Do patients trust computers? J Behav Decis Mak 2006 Nov 24;19(5):455-468. [doi: 10.1002/bdm.542]
7. Seitz L, Bekmeier-Feuerhahn S, Gohil K. Can we trust a chatbot like a physician? A qualitative study on understanding the emergence of trust toward diagnostic chatbots. Int J Hum Comput Stud 2022 Sep;165:102848. [doi: 10.1016/j.ijhcs.2022.102848]
8. Naik N, Hameed BM, Shetty DK, Swain D, Shah M, Paul R, et al. Legal and ethical consideration in artificial intelligence in healthcare: who takes responsibility? Front Surg 2022 Mar 14;9:862322 [FREE Full text] [doi: 10.3389/fsurg.2022.862322] [Medline: 35360424]
9. Juravle G, Boudouraki A, Terziyska M, Rezlescu C. Trust in artificial intelligence for medical diagnoses. Prog Brain Res 2020;253:263-282. [doi: 10.1016/bs.pbr.2020.06.006] [Medline: 32771128]
10. Verma H, Mlynar J, Schaer R, Reichenbach J, Jreige M, Prior J, et al. Rethinking the role of AI with physicians in oncology: revealing perspectives from clinical and research workflows. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. 2023 Presented at: CHI '23; April 23-28, 2023; Hamburg, Germany p. 11-52 URL: https://dl.acm.org/doi/fullHtml/10.1145/3544548.3581506 [doi: 10.1145/3544548]
11. Logg JM, Minson JA, Moore DA. Algorithm appreciation: people prefer algorithmic to human judgment. Organ Behav Hum Decis Process 2019 Mar;151:90-103. [doi: 10.1016/j.obhdp.2018.12.005]
12. Naiseh M, Al-Thani D, Jiang N, Ali R. How the different explanation classes impact trust calibration: the case of clinical decision support systems. Int J Hum Comput Stud 2023 Jan;169:102941. [doi: 10.1016/j.ijhcs.2022.102941]
13. Panigutti C, Beretta A, Giannotti F, Pedreschi D. Understanding the impact of explanations on advice-taking: a user study for AI-based clinical decision support systems. In: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. 2022 Presented at: CHI '22; April 29-May 5, 2022; New Orleans, LA p. 1-9 URL: https://dl.acm.org/doi/10.1145/3491102.3502104 [doi: 10.1145/3491102.3502104]
14. Alam L, Mueller S. Examining the effect of explanation on satisfaction and trust in AI diagnostic systems. BMC Med Inform Decis Mak 2021 Jun 03;21(1):178. [doi: 10.1186/s12911-021-01542-6] [Medline: 34082719]
15. Tucci V, Saary J, Doyle TE. Factors influencing trust in medical artificial intelligence for healthcare professionals: a narrative review. J Med Artif Intell 2022 Mar;5:4. [doi: 10.21037/jmai-21-25]
16. Nagendran M, Festor P, Komorowski M, Gordon AC, Faisal AA. Quantifying the impact of AI recommendations with explanations on prescription decision making. NPJ Digit Med 2023 Nov 07;6(1):206. [doi: 10.1038/s41746-023-00955-z] [Medline: 37935953]
17. Prahl A, Van Swol L. Understanding algorithm aversion: when is advice from automation discounted? J Forecast 2017 Mar 24;36(6):691-702. [doi: 10.1002/for.2464]
18. Hallowell N, Badger S, Sauerbrei A, Nellåker C, Kerasidou A. "I don't think people are ready to trust these algorithms at face value": trust and the use of machine learning algorithms in the diagnosis of rare disease. BMC Med Ethics 2022 Nov 16;23(1):112 [FREE Full text] [doi: 10.1186/s12910-022-00842-4] [Medline: 36384545]
19. Schwartz JM, George M, Rossetti SC, Dykes PC, Minshall SR, Lucas E, et al. Factors influencing clinician trust in predictive clinical decision support systems for in-hospital deterioration: qualitative descriptive study. JMIR Hum Factors 2022 May 12;9(2):e33960 [FREE Full text] [doi: 10.2196/33960] [Medline: 35550304]
20. Lee JT, Hsieh CC, Lin CH, Lin YJ, Kao CY. Prediction of hospitalization using artificial intelligence for urgent patients in the emergency department. Sci Rep 2021 Sep 30;11(1):19472 [FREE Full text] [doi: 10.1038/s41598-021-98961-2] [Medline: 34593930]
21. Crampton NH. Ambient virtual scribes: Mutuo Health's AutoScribe as a case study of artificial intelligence-based technology. Healthc Manage Forum 2020 Jan;33(1):34-38. [doi: 10.1177/0840470419872775] [Medline: 31522566]
22. Berings M, Arts M, Pelders D, Sirag S, Brankaert R, Nuijten E E. Technologie helpt diabetische voetenzorg vooruit. ICT Health 2022 Apr 14;46(2):82-85 [FREE Full text]
23. Transform your audio and video to text and subtitles. Amberscript Global. URL: https://www.amberscript.com/en/ [accessed 2024-04-29]

XSL•FO
RenderX

24. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
25. Collaço N, Wagland R, Alexis O, Gavin A, Glaser A, Watson EK. Using the framework method for the analysis of qualitative dyadic data in health research. Qual Health Res 2021 May 13;31(8):1555-1564. [doi: 10.1177/10497323211011599]
26. Dietvorst BJ, Simmons JP, Massey C. Overcoming algorithm aversion: people will use imperfect algorithms if they can (even slightly) modify them. Manag Sci 2018 Mar;64(3):1155-1170. [doi: 10.1287/mnsc.2016.2643]
27. Mahmud H, Islam AK, Ahmed SI, Smolander K. What influences algorithmic decision-making? A systematic literature review on algorithm aversion. Technol Forecast Soc Change 2022 Feb;175:121390. [doi: 10.1016/j.techfore.2021.121390]
28. Castelo N, Bos MW, Lehmann DR. Task-dependent algorithm aversion. J Mark Res 2019 Jul 15;56(5):809-825. [doi: 10.1177/0022243719851788]
29. Burton JW, Stein MK, Jensen TB. A systematic review of algorithm aversion in augmented decision making. Behav Dis Mark 2019 Oct 23;33(2):220-239. [doi: 10.1002/bdm.2155]
30. Önkal D, Goodwin P, Thomson M, Gönül S, Pollock A. The relative influence of advice from human experts and statistical methods on forecast adjustments. Behav Decis Mark 2009 Feb 09;22(4):390-409. [doi: 10.1002/bdm.637]
31. Lim JS, O'Connor M. Judgmental forecasting with interactive forecasting support systems. Decis Support Syst 1996 Apr;16(4):339-357. [doi: 10.1016/0167-9236(95)00009-7]
32. Whitecotton SM. The effects of experience and a decision aid on the slope, scatter, and bias of earnings forecasts. Organ Behav Hum Decis Process 1996 Apr;66(1):111-121. [doi: 10.1006/obhd.1996.0042]
33. Malterud K. Qualitative research: standards, challenges, and guidelines. Lancet 2001 Aug 11;358(9280):483-488. [doi: 10.1016/S0140-6736(01)05627-6] [Medline: 11513933]
34. Guest G, Bunce A, Johnson L. How many interviews are enough?: an experiment with data saturation and variability. Field Methods 2006 Feb 01;18(1):59-82. [doi: 10.1177/1525822X05279903]
35. Guest G, Namey E, Chen M. A simple method to assess and report thematic saturation in qualitative research. PLoS One 2020 May 5;15(5):e0232076 [FREE Full text] [doi: 10.1371/journal.pone.0232076] [Medline: 32369511]

Abbreviations

AI: artificial intelligence
XAI: explainable artificial intelligence

Edited by E Borycki, A Kushniruk; submitted 30.03.24; peer-reviewed by S Sadler; comments to author 08.07.24; revised version received 18.10.24; accepted 29.12.24; published 21.02.25.

Please cite as:
Tahtali MA, Snijders CCP, Dirne CWGM, Le Blanc PM
Prioritizing Trust in Podiatrists' Preference for AI in Supportive Roles Over Diagnostic Roles in Health Care: Qualitative Interview and Focus Group Study
JMIR Hum Factors 2025;12:e59010
URL: https://humanfactors.jmir.org/2025/1/e59010
doi:10.2196/59010
PMID:

©Mohammed A Tahtali, Chris C P Snijders, Corné W G M Dirne, Pascale M Le Blanc. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 21.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Acceptability of a Conversational Agent–Led Digital Program for Anxiety: Mixed Methods Study of User Perspectives

Pearla Papiernik[1], MSc; Sylwia Dzula[1], MSc; Marta Zimanyi[1]; Edward Millgate[1], PhD; Malika Bouazzaoui[1], MSc, MBA; Jessica Buttimer[1], MS; Graham Warren[1], PhD; Elisa Cooper[1], PhD; Ana Catarino[1], PhD; Shaun Mehew[1], PGCert, PGDip; Emily Marshall[1], PGCert, PGDip; Valentin Tablan[1], PhD; Andrew D Blackwell[1], PhD; Clare E Palmer[1], PhD

ieso Digital Health, Cambridge, United Kingdom

Corresponding Author:
Pearla Papiernik, MSc
ieso Digital Health
The Jeffreys Building, Cowley Road
Cambridge, CB4 0DS
United Kingdom
Phone: 44 0800 074 5560
Email: pearla.papiernik@iesohealth.com

Abstract

Background: The prevalence of anxiety and depression is increasing globally, outpacing the capacity of traditional mental health services. Digital mental health interventions (DMHIs) provide a cost-effective alternative, but user engagement remains limited. Integrating artificial intelligence (AI)–powered conversational agents may enhance engagement and improve the user experience; however, with AI technology rapidly evolving, the acceptability of these solutions remains uncertain.

Objective: This study aims to examine the acceptability, engagement, and usability of a conversational agent–led DMHI with human support for generalized anxiety by exploring patient expectations and experiences through a mixed methods approach.

Methods: Participants (N=299) were offered a DMHI for up to 9 weeks and completed postintervention self-report measures of engagement (User Engagement Scale [UES]; n=190), usability (System Usability Scale [SUS]; n=203), and acceptability (Service User Technology Acceptability Questionnaire [SUTAQ]; n=203). To explore expectations and experiences with the program, a subsample of participants completed qualitative semistructured interviews before the intervention (n=21) and after the intervention (n=16), which were analyzed using inductive thematic analysis.

Results: Participants rated the digital program as engaging (mean UES total score 3.7; 95% CI 3.5-3.8), rewarding (mean UES rewarding subscale 4.1; 95% CI 4.0-4.2), and easy to use (mean SUS total score 78.6; 95% CI 76.5-80.7). They were satisfied with the program and reported that it increased access to and enhanced their care (mean SUTAQ subscales 4.3-4.9; 95% CI 4.1-5.1). Insights from pre- and postintervention qualitative interviews highlighted 5 themes representing user needs important for acceptability: (1) accessible mental health support, in terms of availability and emotional approachability (Accessible Care); (2) practical and effective solutions leading to tangible improvements (Effective Solutions); (3) a personalized and tailored experience (Personal Experience); (4) guidance within a clear structure, while retaining control (Guided but in Control); and (5) a sense of support facilitated by human involvement (Feeling Supported). Overall, the DMHI met participant expectations, except for theme 3, as participants desired greater personalization and reported frustration when the conversational agent misunderstood them.

Conclusions: Incorporating factors critical to patient acceptability into DMHIs is essential to maximize their global impact on mental health care. This study provides both quantitative and qualitative evidence for the acceptability of a structured, conversational agent–driven digital program with human support for adults experiencing generalized anxiety. The findings highlight the importance of design, clinical, and implementation factors in enhancing engagement and reveal opportunities for ongoing optimization and innovation. Scalable models with stratified human support and the safe integration of generative AI have the potential to transform patient experience and increase the real-world impact of conversational agent–led DMHIs.

Trial Registration: ISRCTN Registry ISRCTN 52546704; https://www.isrctn.com/ISRCTN52546704

(JMIR Hum Factors 2025;12:e76377) doi:10.2196/76377

XSL•FO
RenderX

KEYWORDS

mental health; anxiety; digital intervention; conversational agent; artificial intelligence; user experience; acceptability; engagement; thematic analysis

Introduction

Demand for mental health care is increasing, with 1 in 8 individuals globally living with a mental health condition [1]. Despite this growing need, there are only 4 psychiatrists per 100,000 people worldwide [2], leading to long waiting lists and a lack of care [3-5]. The traditional 1:1 mental health care model is proving insufficient to meet the rising demand.

Digital mental health interventions (DMHIs) provide a cost-effective way to increase access to care worldwide [6], but face challenges with user engagement and adherence [7-9]. Engagement is often operationalized through objective behavioral metrics (frequency and duration of use), but users' cognitive and emotional investment (subjective engagement), such as focused attention and psychological investment, is equally critical to understanding how users meaningfully engage with DMHIs to drive clinical impact [10,11]. Key factors impacting engagement include available guidance, perceived fit, and personalization [12]. Artificial intelligence (AI)–powered conversational agents, typically using tree-based dialogue systems, have shown potential to enhance DMHI engagement by offering more interactive and personalized experiences compared with self-guided activities [13]. Meta-analytic evidence suggests that these agents reduce attrition rates [14] and show promising efficacy [15-18]. However, the agent's misunderstanding of user input can be frustrating [19]. Advances in generative AI, such as OpenAI's ChatGPT (released November 2022), are rapidly evolving and enable more fluid and naturalistic interactions that have the potential to transform DMHIs. However, currently, most DMHIs rely on tree-based dialogue systems with templated responses to ensure a patient's experience is safe, evidence-based, and deterministic. It is unclear how user familiarity with generative AI applications will impact perceived effectiveness and engagement with conversational-agent DMHIs reliant on tree-based dialogue systems.

Acceptability is central to developing and evaluating DMHIs as it is linked to engagement, effectiveness, and adoption [20]. Acceptability is a multidimensional construct that reflects how people think and feel about an intervention and whether they view it as appropriate, either before, during, or after using it [21]. Qualitative studies suggest that conversational agent–led DMHIs are perceived positively by users [22,23], in part by providing easily accessible and judgment-free environments [24]. However, uptake of such interventions is impacted by mixed expectations on their effectiveness [25,26]. Presence or absence of empathy, personalization, and the ability of the agent to understand user messages also impact the user experience [19,27]. Our current understanding of conversational agent–led DMHI acceptability relies on limited qualitative data (eg, app store reviews and answers to in-app free-text questions about positive and negative experiences) [15,28-31]. A deeper understanding of the many factors contributing to user perception and experience with conversational agent–led DMHIs is important to maximize their impact.

Using mixed methods is recommended when iteratively developing and evaluating complex interventions such as DMHIs [32,33]. Validated questionnaires offer standardized quantitative benchmarks for comparison with other interventions and evaluation of iterative improvements [34]. By contrast, qualitative methods, such as semistructured interviews, enable richer insights into context-specific user needs and intervention-specific experiences, and allow for the discovery of emergent topics [35]. Through open-ended questions and in-depth explorations of user perspectives and experiences, multidimensional concepts such as acceptability and subjective engagement can be assessed [32,35]. Combining these methods provides complementary insights, where quantitative data describe behavior across the sample, and qualitative findings clarify the underlying reasons and nuances of user experiences.

This study reports additional patient-reported end points from a propensity-matched study evaluating the effectiveness, acceptability, engagement, and safety of a digital program for adults with mild to severe symptoms of generalized anxiety [18]. Participants engaged with the program for up to 9 weeks via a smartphone app. The intervention consisted of a structured program using traditional and third-wave cognitive behavioral therapy (CBT) approaches delivered through multimedia content and text-based interactions with an AI-powered conversational agent. The conversational agent followed a tree-based dialogue system, using natural language processing to deliver appropriate clinician prewritten responses. The program was combined with human support, aimed to enhance user experience and adherence with the program, addressing documented barriers to engagement [12]. The digital program showed a mean reduction in anxiety symptoms comparable to traditional human-delivered therapy [18]. This substudy aimed to understand the acceptability, engagement, and usability of the digital program by combining quantitative and qualitative methods to explore patient expectations and experiences with the intervention and examine the extent to which the program met their needs. Self-report measures for subjective engagement, usability, and acceptability were collected postintervention to quantify the user experience, and qualitative semistructured interviews were conducted with a subsample of participants both before and after the intervention to provide deeper insights into user needs and experiences.

Methods

Study Design

This mixed methods substudy explored the perspectives and experiences of individuals with mild, moderate, and severe symptoms of generalized anxiety before and after engaging with a DMHI. It was conducted within a wider pragmatic, single-intervention arm study with a sample of 299 participants from the United Kingdom [18].

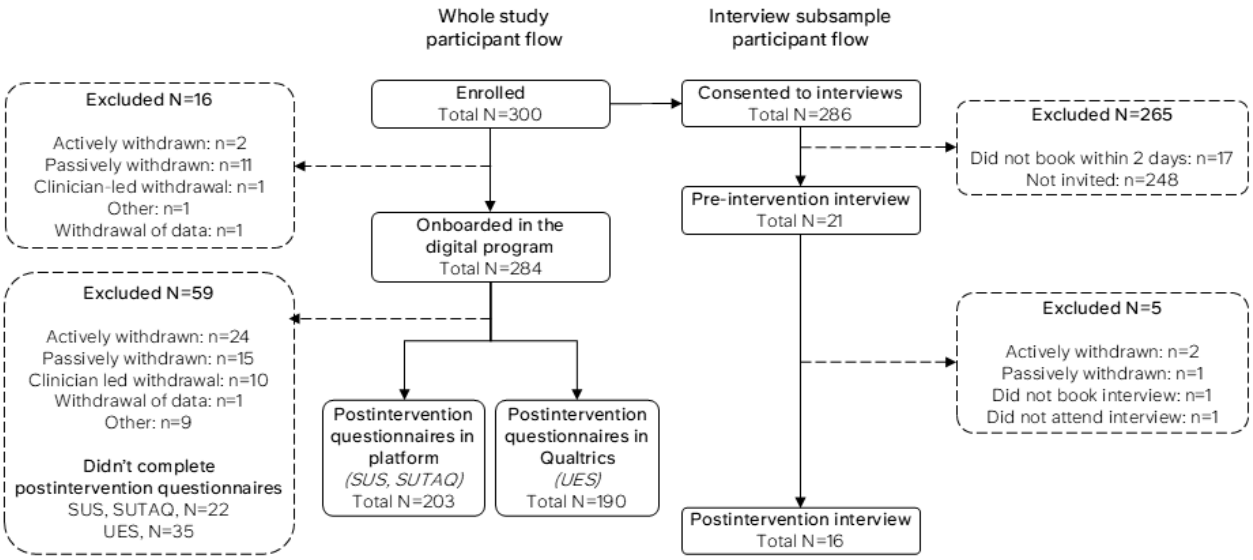
XSL•FO
RenderX

Potentially eligible individuals completed a clinical assessment with a qualified clinician to determine eligibility. At the point of consent, participants were asked if they wanted to participate in qualitative interviews. A subsample of participants (n=21) attended a semistructured interview preintervention. All participants were asked to download the software on their smartphone and engage with the 6-module program in their own time, over a 9-week period. Participants from the interviewed subsample were invited to attend a second semistructured interview after using the program. All outcomes collected at different time points are reported in Palmer et al [18]. For this study, engagement, usability, and acceptability measures, collected with validated questionnaires postintervention, were analyzed, as well as qualitative analysis of the pre- and postintervention semistructured interviews.

Ethics Considerations

The study was preregistered (ISRCTN ID 52546704) and obtained ethical approval before recruitment (IRAS ID: 327897, National Health Service [NHS] Research Ethics Committee: West of Scotland REC 4). The participant CONSORT (Consolidated Standards of Reporting Trials) flowchart for the whole sample and the interviewed subsample is summarized in Figure 1. For more details on participant flow before enrollment, please refer to Palmer et al [18]. In line with the Declaration of Helsinki, informed consent was obtained from all participants before enrollment, including optional consent for participation in semistructured interviews. Participants were debriefed at the conclusion of the study.

Figure 1. Participant flow for the study sample and interviewed subsample. Recruitment sources and preenrollment flow are described in Palmer et al [18]. Participants were withdrawn actively (requested withdrawal), passively (dropped out or disengaged), clinician-led (withdrawn on clinician recommendation), or for other reasons (eg, technical issues). Interview participation was offered on a first-come, first-served basis, with a 2-day response window, and invitations ceased once a sufficient number of preintervention interviews had been completed. Postintervention questionnaires were administered through ieso's validated clinical delivery platform (System Usability Scale [SUS] and Service User Technology Acceptability Questionnaire [SUTAQ]) and through Qualtrics (User Engagement Scale [UES]).



Description of Digital Intervention

The intervention consisted of a 6-module digital program delivered via a smartphone app (ieso Digital Program; software name IDH-DP2-001), developed by ieso Digital Health [36], a mental health technology company and provider within NHS Talking Therapies delivering 1:1 human-delivered CBT via a typed modality to treat patients with common mental health disorders. The ieso Digital Program used a conversational agent to guide users through activities, with clinical oversight and user support provided by nonclinical research coordinators. The program was designed for people primarily presenting with symptoms of generalized anxiety and was based on principles from traditional CBT and third-wave approaches such as acceptance and commitment therapy. The 6 modules consisted of an introduction module, 3 core modules, and 2 consolidation modules. All modules began with a symptom "check-in" consisting of the 7-item Generalized Anxiety Disorder (GAD-7) scale and 9-item Patient Health Questionnaire (PHQ-9) scale. The minimum meaningful program dose was defined a priori by accredited CBT therapists as completion of modules 1-3 and the module 4 check-in. There were 16 sessions total, released on a timed schedule, subject to completing the prior session, allowing space for practice and reflection between sessions. Each session included videos, educational content, conversations, and worksheets written by accredited clinicians. The clinical team had an average of 14 years' experience delivering therapy to patients with complex needs and diverse backgrounds in the United Kingdom. To ensure safety and clinical quality of the messages delivered to the user, the conversational agent used in this version of the program followed a tree-based dialogue system, using natural language processing to understand user input and deliver appropriate prewritten responses. The program complied with ISO 13485

XSL•FO
RenderX

[37] and was registered as a UKCA (UK Conformity Assessed) marked Class 1 medical device. For additional details, including screenshots of the interface and schematics of the program, see Palmer et al [18].

Participants

Adults with mild to severe anxiety symptoms were recruited between October 10, 2023, and February 2, 2024, through 3 streams: (1) referrals to ieso's typed therapy service (self-referral or via their NHS Provider) [38]; (2) responses to online advertisements; and (3) responses to email invitations through the NIHR BioResource [39].

Participants had to meet the following eligibility criteria: (1) be over the age of 18 years at point of recruitment; (2) GAD-7 total score >7; (3) PHQ-9 total score <16; (4) have a primary presenting problem of generalized anxiety disorder, as established through a clinician assessment (based on the International Classification of Diseases, 10th Revision [ICD-10] code in line with the NHS Talking Therapies manual [40]); (5) have access to a smartphone and internet connection; and (6) be registered with a general practitioner in the United Kingdom. To be suitable for the CBT protocol used in the intervention, individuals were excluded if presenting with (1) complex comorbid presentations (diagnosis of multiple disorders, psychotic or personality disorder, autism spectrum condition, or intellectual disability); (2) untreated mental health conditions (including substance misuse), except generalized anxiety disorder or major depressive disorder; or diagnoses of (3) posttraumatic stress disorder, obsessive-compulsive disorder or panic disorder. Additionally, individuals were excluded if they (1) were currently receiving psychological therapy, (2) had changes in psychiatric medication in the past 1 month; or (3) displayed a significant risk of harm to self, to others, or from others (established with the clinical assessment). Any individual who had previously participated in user research for the digital program was also excluded.

Considering the study's aims and design, participant specificity, and experienced user researchers collecting data, a moderate sample size between 15 and 30 participants was estimated to provide sufficient informational power [41]. We aimed to enroll up to 30 participants to account for an estimated 50% attrition rate with the goal of having at least 15 postintervention interviews. Those who consented to interviews self-selected into the interviewed subsample by signing up for a preintervention interview slot on a first-come, first-served basis. If they did not sign up within 2 days of receiving the invitation, they were sent the information to continue with the app download to avoid delays. Interview invitations were sent out from December 2023, 2 months into the enrollment period. Interview invitations were stopped once a sufficient sample had attended a preintervention interview.

Vouchers up to a total of £60 (US $81.1) were provided for participating in the study based on study assessments and completion of modules. An additional £15 (US $20.2) voucher per interview was provided to those who attended an interview.

Patient and Public Involvement

Experts by lived experience were involved as members of a patient and public involvement (PPI) panel. PPI members took part in study conceptualization and supported recruitment through co-designing marketing campaigns and participant-facing documents. PPI members were also involved at the end of the thematic analysis process through a workshop aimed to discuss the coherence of thematic groupings and theme names, and limitations and implications of these results. Before the workshop, PPI members were introduced to thematic analysis and qualitative methods. More details are provided in Multimedia Appendix 1.

Data Collection and Analysis

Self-Report Measures of User Experience

Overview

Validated self-report measures of user experience were collected at postintervention between January and April 2024 (after completion or after 9 weeks), via either ieso's validated clinical delivery platform (System Usability Scale [SUS] and Service User Technology Acceptability Questionnaire [SUTAQ]) or Qualtrics (User Engagement Scale [UES]). Demographic data were collected at enrollment and are summarized in Table 1.

XSL•FO
RenderX

Table 1. Sample characteristics of all study samples.

Demographic and category	Questionnaire sample (N=203)	Interview subsample (n=21)	
		Preintervention sample (n=21)	Postintervention sample (n=16)
Baseline GAD-7 score, mean (SD)[a]	12.4 (3.4)	11.8 (3.2)	11.6 (3.6)
Baseline PHQ-9 score, mean (SD)[b]	7.9 (3.8)	7.7 (2.4)	7.7 (2.5)
Age, mean (SD)	41.0 (12.0)	46.0 (13.4)	44.0 (13.3)
Gender, n (%)			
Female	163 (80.3)	17 (81.0)	13 (81.3)
Male	34 (16.7)	3 (14.3)	2 (12.5)
Other	2 (1.0)	1 (4.8)	1 (6.3)
Not known	4 (2.0)	0 (0)	0 (0)
Ethnicity, n (%)			
White	185 (91.1)	19 (90.5)	15 (93.8)
Mixed	3 (1.5)	0 (0)	0 (0)
Asian	9 (4.4)	2 (9.5)	1 (6.3)
Black/African/Caribbean/Black British	1 (0.5)	0 (0)	0 (0)
Other	1 (0.5)	0 (0)	0 (0)
Prefer not to say	4 (2.0)	0 (0)	0 (0)
Long-term condition, n (%)			
Yes	83 (40.9)	9 (42.9)	6 (37.5)
No	112 (55.2)	10 (47.6)	9 (56.3)
Not known	8 (3.9)	2 (9.5)	1 (6.3)

[a]7-item Generalized Anxiety Disorder categories: mild (scores 5-9), moderate (scores 10-14), and severe (scores ≥15).

[b]9-item Patient Health Questionnaire categories: mild (scores 5-9), moderate (scores 10-14), moderately severe (scores 15-19), and severe (scores ≥20).

Engagement

Engagement was evaluated with the UES short form [42], a validated scale to measure user engagement with digital systems. This scale consisted of 12 items across 4 subscales: (1) *Focused Attention* (ie, how absorbed the individual was in the intervention); (2) *Perceived Usability* (ie, whether the intervention was taxing or frustrating to use); (3) *Aesthetic Appeal* (ie, how visually appealing the intervention was; and (4) *Rewarding* (ie, whether using the intervention was worthwhile). The 12 items were rated on a 5-point scale (1=strongly disagree to 5=strongly agree). A high score can be interpreted as positive perceived engagement. Items were presented in a random order for each participant. Three items on the *Perceived Usability* subscale were reverse scored, such that a lower value on each item was associated with greater perceived engagement. Individual items are reported as the number and percentage of participants who selected each Likert rating. These were averaged across items within each subscale to estimate the distribution of participants who on average agreed or disagreed with the items within each subscale. Average percentage agreement or disagreement to subscales across participants was also reported, calculated by regrouping all levels of agreement (agree/strongly agree) or disagreement (disagree/strongly disagree).

Usability

Usability was evaluated with the SUS [43]. The SUS consisted of 10 items relating to the overall usability of the intervention. Items were rated on a 5-point scale (1=strongly disagree to 5=strongly agree). An overall SUS score was calculated as a sum across all items after reverse coding certain items and scaling to produce a score from 0 to 100 [43]. A high score can be interpreted as positive perceived usability. Individual items are reported as the number and percentage of participants that selected each Likert rating. These were averaged across all items to estimate the proportion of participants that on average agreed or disagreed with items across the questionnaire. Average percentage agreement or disagreement per item and for the whole questionnaire across participants was also reported, calculated by regrouping all levels of agreement (agree/strongly agree) or disagreement (disagree/strongly disagree).

Acceptability

Acceptability was quantitatively evaluated with the SUTAQ [44]. It consisted of 22 items across 6 subscales: (1) *Care Personnel Concerns* (ie, concerns about the skills and continuity of the personnel looking after them); (2) *Enhanced Care* (ie, beliefs that the intervention can improve care); (3) *Increased Accessibility* (ie, beliefs that the intervention enhanced access to care); (4) *Kit as Substitution* (ie, beliefs that the intervention may be an alternative to regular care); (5) *Privacy and*

XSL•FO
RenderX

Discomfort (ie, concerns privacy and the intervention interfering with their daily life or making them feel uncomfortable); and (6) *Satisfaction* (ie, beliefs indicating acceptance and satisfaction with the kit and service provided). Participants rated their level of agreement with each item on a 6-point Likert scale (1=strongly disagree to 6=strongly agree). High scores can be interpreted as positive user perception, except for subscales *Care Personnel Concerns* and *Privacy and Discomfort* that are inverted; therefore, a low score is interpreted as a positive user perception. The intermediate value 3.5 is interpreted as a point of neutrality. Reverse coding was carried out on 1 item in the *Kit as Substitution* subscale [44] (item 18; see Table S3 in Multimedia Appendix 2). Individual items are reported as the number and percentage of participants that selected each Likert rating. These were averaged across items within each subscale to estimate the distribution of participants agreeing or disagreeing with the items within each subscale. Average percentage agreement or disagreement to subscales across participants was also reported in text, calculated by regrouping all levels of agreement (mildly, moderately, and strongly agree) or disagreement (mildly, moderately, and strongly disagree).

Qualitative Data

Semistructured interviews pre- and postintervention provided qualitative data about user expectations and experiences with the program, to assess the acceptability of the ieso Digital Program in meeting the needs of participants. Preintervention interviews, conducted by 2 user researchers between December 2023 and February 2024 (SD and MZ), lasted 16-38 minutes (mean 23 minutes). Postintervention interviews, conducted by 1 user researcher between February and April 2024 (MZ), lasted 30-47 minutes (mean 39 minutes). Interview topic guides were developed by user researchers, based on learnings from literature and from extensive internal user research and PPI workshops during the development phase of the program. Topic guides were reviewed by internal stakeholders with clinical and product expertise. Preintervention interviews explored participants' motivations to take part, their experience and views of using technology for mental health, and their expectations for the intervention. Postintervention interviews explored participants' overall experience with the intervention, engagement barriers and facilitators, perceived usefulness and effectiveness, and perceived safety and support. Refer to Multimedia Appendix 1 for detailed interview schedules. Interviews were conducted on Microsoft Teams and were automatically transcribed on Dovetail [45] and reviewed and anonymized by researchers.

Pre- and postintervention interviews were analyzed using reflexive thematic analysis [46,47], led by a researcher trained in qualitative research methods (PP) who familiarized with the data by reading interview transcripts multiple times and completed line-by-line inductive coding at both the semantic and latent levels in Dovetail. Codes were then grouped into candidate themes, which were then reviewed by returning to the underlying data. Final themes were defined by clarifying thematic boundaries and writing synopses. To ensure reflexivity and rigor in interpretations, analysis was regularly discussed between researchers (PP, CEP, MB, and MZ). A workshop with individuals with lived experience of mental health was held to review and discuss the thematic analysis findings. Final themes grouped both pre- and postinterview data linked by the common underlying concept of the theme. Results describe expectations and experiences in each theme. A detailed description of the thematic analysis process and workshop is available in Multimedia Appendix 1.

The framework method [48,49] was used to explore patterns between preintervention expectations and postintervention experiences. Inductive themes from the thematic analysis were used to create an analytical framework, and coded extracts were categorized within it. A framework matrix summarized data for each participant and theme by interview time (pre and post), with illustrative quotes. The framework matrix was interpreted by exploring patterns within and between participants, before and after the intervention, and relevant nuances and insights were added to our description of themes developed through thematic analysis.

Results

Self-Report Measures of User Experience

Overview

The final sample for analysis included 203 participants who completed the SUS and the SUTAQ, and 190 participants who completed the UES. Table 1 provides an overview of the demographic and clinical characteristics of participants enrolled in the intervention who completed the questionnaires (N=203). Of these, 163 (80.3%) were female, with a mean age of 41 (SD 12.0; range 19-75) years, and 168 (82.8%) completed the minimum meaningful clinical dose [18]. UES and SUTAQ subscale scores are reported in Figure 2, and average agreement or disagreement across scale items are reported in Tables 2 (UES and SUS) and 3 (SUTAQ). Item-level data for each scale can be found in Tables S1-S3 in Multimedia Appendix 2.

XSL•FO
RenderX

Figure 2. Self-reported user experience mean scores with 95% CIs. (A) User Engagement Scale (UES) subscales and total score. (B) Service User Technology Acceptability Questionnaire (SUTAQ) subscales. Higher scores on all subscales indicate more positive user sentiment, except for the SUTAQ *Care Personnel Concerns* and *Privacy and Discomfort* subscales, where lower scores indicate more positive sentiment.

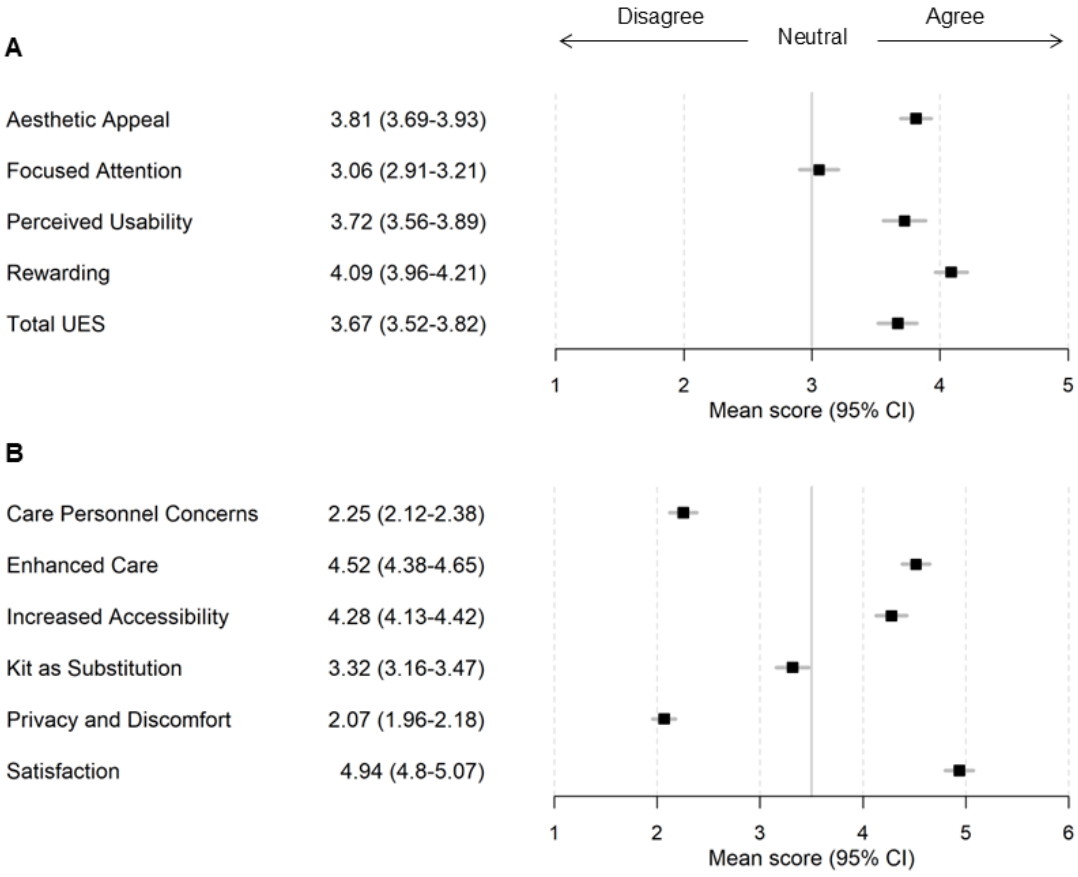


Table 2. Agreement and disagreement with the self-reported measures of engagement (UES[a]) and usability (SUS[b]).

Scale	Mean proportion agreement across all items or items within each subscale, %				
	1=Strongly disagree	2=Disagree	3=Neither agree nor disagree	4=Agree	5=Strongly agree
UES (n=190)					
Aesthetic Appeal	1.2	4.6	24.2	51.8	18.3
Focused Attention	6.5	26.3	29.3	30.9	7.0
Perceived Usability	4.4	14.2	13.5	40.5	27.4
Rewarding	1.9	3.7	10.5	51.4	32.5
Total	3.5	12.2	19.4	43.6	21.3
SUS (n=203)					
Total	1.4	5.3	9.7	44.6	39.0

[a]UES: User Engagement Scale.

[b]SUS: System Usability Scale.

XSL•FO
RenderX

Table 3. Agreement and disagreement with the self-reported measure of acceptability (SUTAQ[a]).

Scale	Mean proportion agreement across all items or items within each subscale, %					
	1=Strongly disagree	2=Moderately disagree	3=Mildly disagree	4=Mildly agree	5=Moderately agree	6=Strongly agree
SUTAQ (N=203)						
Care Personnel Concerns	36.1	27.4	19.1	11.2	4.9	1.3
Enhanced Care	4.8	5.2	9.1	25.3	25.5	30.1
Increased Accessibility	5.8	6.4	12.8	28.1	23.2	23.8
Kit as Substitution	14.8	16.9	21.0	24.1	15.6	7.6
Privacy and Discomfort	45.2	23.8	15.8	10.8	3.3	1.1
Satisfaction	2.6	2.8	6.1	16.3	31.4	40.9

[a]SUTAQ: Service User Technology Acceptability Questionnaire.

Engagement

Overall, the mean total score for the UES was 3.7 (95% CI 3.5-3.8), indicating agreement that the program was engaging. On average, across the intervention sample who completed the UES (n=190), 159 (83.7%) agreed that the program was rewarding and worthwhile (*Rewarding*: mean 4.1; 95% CI 4.0-4.2), 129 (67.9%) perceived the app not to be frustrating or taxing (*Perceived Usability*: mean 3.7; 95% CI 3.6-3.9) and 133 (70%) found the app aesthetically pleasing (*Aesthetic Appeal*: mean 3.8; 95% CI 3.7-3.9). There were mixed views on whether participants were fully focused on the experience (*Focused Attention*: mean 3.1; 95% CI 2.9-3.2). Item-level scores indicated participants generally disagreed that they were "getting lost" in the experience, and there were mixed views on whether participants were fully absorbed in the app (see Table S1 in Multimedia Appendix 2).

Usability

Self-reported usability scores were high with a mean total score for the SUS of 78.6 (n=203; 95% CI 76.5-80.7). Item-level data (Table S3 in Multimedia Appendix 2) show that across the intervention sample (n=203), most participants agreed that the system was quick to learn to use (n=185, 91.1%), without the support of a technical person (n=188, 92.6%); 180 (88.7%) felt confident using the app and 179 (88.2%) found it easy to use. The item participants agreed the least strongly with was whether they would like to use the system frequently (n=135, 66.5% agreement).

Acceptability

Across the intervention sample (n=203), 180 (88.7%) participants were satisfied with the intervention (*Satisfaction*: mean 4.9; 95% CI 4.8-5.1). On average, 164 (80.8%) participants found that the app could improve care (*Enhanced Care*: mean 4.5; 95% CI 4.4-4.7) and 152 (74.9%) found that it increased access to care (*Increased Accessibility*: mean 4.3; 95% CI 4.1-4.4), without interfering with their daily life or making them feel uncomfortable (*Privacy and Discomfort*: mean[reverse scored] 2.1; 95% CI 2.0-2.2), and without concern about the skills and continuity of the personnel looking after them (*Care Personnel Concerns*: mean[reverse scored] 2.3; 95% CI 2.1-2.4). There were mixed responses about whether the program could be used as a substitution for their regular mental health care (*Kit as Substitution*: mean 3.3; 95% CI 3.2-3.5).

Qualitative Findings

Overview

Table 1 provides an overview of the demographic and clinical characteristics of participants enrolled in the interview subsample (n=21). Participants in the interview subsample were mostly White (n=19, 90%) and women (n=17, 81%), with a mean age of 46 (SD 13.1; range 19-72) years. Half of the sample had moderate anxiety (n=10, 48%), followed by mild (n=7, 33%) and severe (n=4, 19%) anxiety. Five participants did not do postintervention interviews: 3 were withdrawn from the study, and 2 did not book or attend their postintervention interview. Participants were highly engaged, with 14 out of 21 (67%) of the pre-intervention and 12 out of 16 (75%) of the postintervention interview subsamples completing a minimum meaningful clinical dose of the program.

Figure 3 shows a map of the themes developed through thematic analysis of the pre- and postintervention interviews. Five themes represent patient needs for conversational agent–led DMHIs and describe to what extent they were met, based on their expectations and experiences with the ieso Digital Program: (1) Accessible Care, (2) Effective Solutions, (3) Personal Experience, (4) Guided but in Control, and (5) Feeling Supported. Quotes are provided with interview type (pre, post), participant gender, age, and baseline anxiety severity.

XSL•FO
RenderX

Figure 3. Thematic map. Five qualitative themes illustrate user needs that influence the acceptability of conversational agent–led digital mental health interventions (DMHIs). Each theme summarizes user expectations and experiences drawn from pre- and postintervention interviews.



Theme 1: Accessible Care

This theme describes how participants sought mental health support that was accessible in terms of both availability and emotional approachability. The digital format was perceived as offering this kind of access and reached users who would not have accessed support otherwise.

Many participants were motivated to access mental health support that they felt was difficult to obtain through the NHS. Participants appreciated that the program was a faster, easier, and more affordable way to access support.

> *It's always easy to get an appointment but then from there the support is pretty much nonexistent....Basically, the advice is you could probably do with some more talking therapy, you can go on the waiting list and that will be up to a year, otherwise you can pay. So that's why the program is such a positive thing, because otherwise there's nothing.* [post, P14, female, 53, severe anxiety]

Some participants avoided seeking support due to guilt over using NHS resources. For some, the digital program was perceived as a good "first point of call" to address mental health concerns, and other options could be available after if needed.

> *I see it like, if I lined up Domino's, this is like Domino number one.* [pre, P16, female, 44, severe anxiety]

For many participants, a digital modality was an advantage. Participants wanted to reflect on their experiences and understand their triggers, thoughts, and feelings, and some believed that a digital program was better suited for them to do this work. Indeed, several participants expected and found it

XSL•FO
RenderX

easier to express their feelings to an app compared with a human therapist, especially if they had prior negative experiences with therapy. Participants felt less judged or embarrassed to open-up, and therefore self-censored less. Some participants also felt that the program may be less biased than a human therapist.

> *I liked it because I felt there wasn't any sense of judgment there. I've had a bad experience with therapy before where I didn't really get along with the therapist and there was quite a lot of judgment passed...And just knowing that that wasn't even something that could happen was really reassuring.* [post, P4, female, 19, mild anxiety]

Several participants preferred the privacy DMHIs would provide, compared with online therapy calls that could be overheard, and valued having more time to understand information and reflect before answering.

> *Text based certainly for me, I'm able to be much more thoughtful, much more introspective, much more precise is my answer when they're written rather than spoken. I think it gives me time to think deeper about things.* [pre, P12, female, 48, moderate anxiety]

Participants also valued being able to revisit and reflect on past content and discussions saved and centralized on an app. Many wanted continued app access poststudy to further review content.

> *If I could go back into the app to see how I wrote down what I was feeling, that would be, that would really reinforce it all.* [post, P9, female, 57, mild anxiety]

Theme 2: Effective Solutions

This theme describes participants' motivation to access effective solutions for anxiety. They sought practical tools to manage anxiety and would continue engaging if they perceived the program as effective.

A key motivation expressed in pre- and postintervention interviews was to access a solution-focused intervention to effectively manage or reduce anxiety. Participants differentiated the need to "offload," from their current need to "change behavior" and receive "tools" and "solutions." Most wanted to learn techniques to manage their anxiety long-term, while some sought momentary support or both. Participants practiced and applied techniques in everyday life, particularly when anxious or when facing a stressful situation.

> *[When I'm in a triggering situation], I would go either directly back to the app or I would go back to my book and I would look at what I've learned and I would go, ok, how can I approach this situation differently?* [post, P16, female, 44, severe anxiety]

Accessing helpful content and learning new information and techniques was crucial to participants and was for many a key motivator, even for participants who experienced frustration when the conversational agent misunderstood them. Nearly all participants interviewed found the program useful and effective. Seeing improvements and perceiving the program as beneficial was an important motivator for continued engagement with the program and continued application of techniques learned.

> *I was motivated to continue because each module had something which helped me. I felt that it was helping me reduce my anxiety right from the beginning. So, you wanted to do the next module.* [post, P5, female, 68, mild anxiety]

Participants noted the program's time and effort demands but emphasized the importance of engaging to see benefits. For many, this was acceptable as they began the program highly motivated and ready to engage; however, 1 participant was surprised by the program's workload and another shared that they could not have engaged similarly at another time in their mental health journey.

> *If you can stick with it and really commit to trying the things it suggests you try, you'll definitely find it helpful....But you have got to be willing to kind of commit to it and put in the time and do what it's asking you to do. So, you've got to be in the right headspace.* [post, P14, female, 53, severe anxiety]

Theme 3: Personal Experience

The third theme describes how participants wanted the program to be tailored and feel personal, but sometimes found the conversational agent to be generic or unnatural, and often felt misunderstood by it.

Preintervention, several participants expected personalized strategies for anxiety, tailored advice, and pathways or exercises based on their input or assessments. Half of the participants viewed technology and an AI guide as an opportunity for tailored and accurate support.

> *I would assume that based on my answer to questions or how I'm responding to things, that machine learning will tell the digital guide move things to get onto the right journey.* [pre, P7, female, 53, moderate anxiety]

Postintervention, some found the conversational agent impersonal or generic, and desired more context-specific advice, historical knowledge use, and the ability to ask questions.

> *I didn't quite enjoy that I was talking to a robot that didn't tell me things tailored to my needs, only generic things. So many times it didn't really matter what I said, the robot only repeated what I said.* [post, P2, male, 33, moderate anxiety]

Feeling heard and understood was crucial to participants. In preintervention interviews, interacting with a conversational agent was viewed as acceptable if it was well-trained, lacked bias, and was able to understand them and offer thoughtful answers.

> *As long as it's been well trained, it will be interesting to see what it comes back with. That it's not just a sort of trite responses you get when you're trying to get customer service out.* [pre, P12, female, 48, moderate anxiety]

Postintervention, some participants explained that the downside of a conversational agent was that they could not receive empathy and be understood like they would with a therapist. Nearly all users reported instances where the conversational

XSL•FO
RenderX

agent wrongly paraphrased messages or misunderstood complex thoughts and feelings. This often led to frustration or confusion. Safety messages that appeared when the user did not feel they were needed could lead to frustration and distraction from the session.

> *I think there was a lot of instances where it was clear that whatever I typed it hadn't understood. In a way that's like sitting in a room with a therapist and you tell them something and they just look blank.* [post, P15, nonbinary, 42, severe anxiety]

Some participants found the conversational agent's understanding abilities were less bothersome and less present over time, and some explained that they learned to communicate more effectively with it.

> *I think as the program went on, I got to know the program and I got maybe a bit better myself at finding ways to word things.* [post, P14, female, 53, severe anxiety]

However, others explained that simplifying their answers to ensure the conversational agent would understand impacted their ability to express their emotions.

> *And so you learn to keep it really simple. But that sometimes didn't help, I couldn't really express how I was feeling because the chatbot didn't understand.* [post, P5, female, 68, mild anxiety]

Several participants were not too impacted by these limitations, finding them expected and acceptable for a conversational agent, and not interfering with the overall program's usefulness.

> *You could tell it was looking for specific responses, but that's fine. That's what they do.* [post, P2, male, 33, moderate anxiety]

Postintervention, participants had mixed views on the conversational agent's human-likeness, some finding it human-like and others robotic. Notably, 2 participants valued its neutrality and "ambiguity," neither distinctly human nor robotic, and neither male nor female. However, conversations feeling robotic and preformatted answers could feel restrictive and frustrating, and affected participant's sense of agency, an important need discussed in the next theme.

> *I do think some of the time when you only have one response to give, sometimes I wanted to have choice or say my own thing. If you read back through my data, sometimes you can probably see I'm getting a bit frustrated, if you want to say more sometimes and then you're not given that option. And that sort of forces you out of your own voice and into the voice of whoever programmed it.* [post, P12, female, 48, moderate anxiety]

Theme 4: Guided but in Control

The fourth theme highlights participants' desire to have autonomy and control over how they engaged with the program while also wanting structure and guidance throughout. Being self-directed within a flexible yet structured intervention was valued by participants and was a motivation to participate in the study.

Accessing a time-bound and structured program through an app was seen by participants as useful for personal accountability to help them regularly engage in the work they wanted to do to feel better.

> *I think this would be helpful because I can tap in and out of it as and when. And it's something that's available, the app is always there for you to use. And at the same time there's time frames on it. So, you have to do things by certain dates. So there's a bit of a routine, but also a bit of flexibility.* [pre, P1, female, 35, mild anxiety]

Reminders were also seen as important to support adherence to the program, particularly considering reported engagement challenges such as their mental health, forgetfulness, or external obligations.

> *The reminders did serve, as a prompt to say, OK, just reminding you that you got your next session and you'll be on track to complete the study.* [post, P8, male, 32, mild anxiety]

Accessing a program that is easy to navigate, with a "logical pathway" to follow, was important. Participants also wanted to be guided toward what they should practice and reflect on, through conversational agent prompts and exercises. One participant viewed conversational agents as acceptable only if used within a structured program.

> *The other app I used, I would say something and then the chatbot would reflect that back to me. I'm assuming there is a structured program in the ieso app. So, it's not just going to be stuff I say being reflected back, there would be ideas for me and something for me to follow.* [pre, P13, female, 55, moderate anxiety]

In postintervention interviews, some participants expressed that having more in-app features and sessions to support practice and review content would be helpful, as well as guided support when experiencing symptoms of anxiety, such as meditation or breathing exercises.

> *I think to actually have some, a few recorded meditations that people could turn to relax breathing techniques or relaxation, especially if you're having a bit of a crisis and you need to calm down.* [post, P13, female, 55, moderate anxiety]

In parallel to wanting guidance and structure, autonomy was important to participants.

> *Nothing feels forced upon you. Even the reminders I think are gentle, they're like a little nudge rather than a big old slap.* [post, P16, female, 44, severe anxiety]

Many participants expressed high satisfaction with the intervention's flexibility and ability to fit into their lives, allowing engagement at their own pace, when in the right headspace, and in various settings. Several participants, therefore, saw the program as particularly adapted for people with demanding jobs or caregiving responsibilities.

> *I don't have a lot of time to attend scheduled appointments. So being able to pick something up to*

XSL•FO
RenderX

> *help my own mental health in my own time is more likely to be something I would use in the future than in-person therapy.* [post, P6, female, 44, moderate anxiety]

Many participants desired more control over session timing, duration, and frequency. While some liked time locks between sessions to practice, others found them either too far apart, such that they hindered progress or timely support, or too close together, leaving insufficient time to practice.

> *Having to wait a few more days until the next session, I understand why. But sometimes I was really into a mood where I needed those sessions. And after a few days when the session was available, I wasn't available myself.* [post, P2, male, 33, moderate anxiety]

> *I thought early on it was the sessions are slightly too close together and I thought I didn't always have time to practice.* [post, P10, female, 30, mild anxiety]

Theme 5: Feeling Supported

The final theme describes how participants found the program to be supportive and safe, facilitated by human support.

Clinician involvement in the design of the program provided credibility. Participants were willing to engage as the content was clinically backed and written by humans.

> *I think if it's controlled, it's preprogramed responses written by an expert then, I have a lot of confidence in that system.* [pre, P3, female, 31, severe anxiety]

If users experienced increased anxiety during the study, it was attributed to external stressors and users were able to speak to a clinician. Some users said that negative emotions could sometimes arise, but this was viewed as expected when engaging in such exercises.

> *There were a few moments where it would ask you to think of something negative to practice. But I think that's how those sessions work. And at the end of it, I always felt better. But yeah, there was the odd moment, when it was getting you to think something, but I can't see a way for it to get you to practice and explain it without that.* [post, P10, female, 30, mild anxiety]

Participants valued human support throughout the program.

> *I think it was good that you can just pick it up and do it as and when, while also having some support from real people if and when you needed that.* [post, P1, female, 35, mild anxiety]

Regular calls with research coordinators were beneficial for adherence and engagement with the program, providing gentle reminders or deadlines in addition to app reminders.

> *It helped with accountability knowing that someone was going to call because then if I hadn't done it, that wouldn't be good.* [post, P10, female, 30, mild anxiety]

Importantly, regular check-ins with research coordinators and the option to consult a clinician provided users with a sense of safety and reassurance during the program, and knowing about this human element was a motivating factor to take part in the study. Knowing they could talk to a human was important even when participants did not experience the need for it. Several participants suggested integrating easy ways to contact a human within the app.

> *That was the good thing about those calls [with the research coordinator], you did know there was actually somebody there. I never had to contact a clinician or anything, but it was always very clear that that is an option if that is necessary.* [post, P4, female, 19, mild anxiety]

While most felt supported by the human contact provided, a few desired more frequent clinician interaction and guidance. Some accessed the human support they needed outside of the program, through friends and family, and 1 participant explained that the program itself encouraged and helped them seek social support.

Postintervention, participants expressed that the program's design and content created a supportive and encouraging environment. Several participants described the program as "calming," "relaxing," and "gentle"; the content as "lovely" and "therapeutic"; and the conversational agent as "friendly" and "reassuring."

> *And the acknowledgment at the end, "Let's acknowledge that this may be hard for you," "Are you willing to commit to it?". At that level of acknowledgement, you feel heard. It was like having a cheerleader on your side, you know? That was nice I thought.* [post, P12, female, 48, moderate anxiety]

Discussion

Principal Findings

This mixed methods study demonstrates the overall acceptability of a digital program driven by an AI conversational agent with human support for adults experiencing generalized anxiety symptoms. Quantitative self-reported measures showed that participants found the digital program rewarding and easy to use, and that it both increased access to and enhanced their mental health care. Qualitative interviews provided an in-depth understanding of factors important for DMHI acceptability, giving context to the quantitative findings—such as guiding users, fostering a sense of support, and observing tangible mental health improvements. However, participants frequently reported a lack of personalization and frustrations with the conversational agent's misunderstandings, and there were mixed views on whether it could serve as a substitute for standard care. Using mixed methods enabled a deeper understanding of patient acceptability and offered a new perspective on previously reported outcomes [18], capturing participants' affective and cognitive investment beyond usage metrics [50]. These findings highlight opportunities to optimize and innovate DMHIs to enhance their acceptability and impact on mental health care.

In this study, many participants sought a self-directed and autonomous solution while also wanting to feel supported, highlighting the need to balance structure and guidance with

XSL•FO
RenderX

personal agency in DMHIs. Users appreciated the structured program and wanted more in-app support, such as prompts or worksheets to facilitate reflection and practice. By contrast, some participants wanted increased control over their pace through the program, echoing findings from other conversational agent DMHIs [51]. For example, although time-locked sessions were designed to encourage practice between sessions, many participants preferred having greater control over this. Most participants found the program easy to use, as reflected in a high mean SUS score of 78.6, which exceeds scores reported for other conversational agent DMHIs (63.6-66.2) [19]. The program also met participants' expectations regarding access and flexibility, with 152 out of 203 (74.9%) agreeing that it increased access to care (SUTAQ *Increased Accessibility* subscale).

Despite wanting a self-led digital solution—perceived by some participants as a judgment-free and less intimidating way to access mental health support without talking to a therapist—human support remained crucial for building trust in the program, ensuring safety, and maintaining engagement. Nonclinical research coordinators, who provided motivational support through emails and phone calls, were perceived as helpful in fostering a sense of accountability. Participants also valued having the option to talk to a clinician if needed. This aligns with patient perceptions that DMHIs should involve clinicians and facilitate access to mental health professionals [23,26]. Moreover, this is consistent with findings that DMHIs with human support have lower dropout rates [14] and better outcomes [52], particularly at higher symptom severity [18]. Most participants reported minimal concerns about personnel skills and continuity of care (SUTAQ *Care Personnel Concerns* subscale; mean 2.3, range 1-6), highlighting the acceptability of this support model for real-world implementation. Nonetheless, there were mixed views on the acceptability of the digital program as a replacement for regular mental health care (SUTAQ *Kit as Substitution* subscale; mean 3.3, range 1-6). As individual needs and preferences influence the acceptability of DMHIs, it is essential to ensure that people are offered an informed choice in how they receive care. Moreover, although human support can enhance acceptability, implementations must remain scalable to meet the rising demand for mental health care. In the current program, comparable outcomes to human-delivered care were achieved with up to 8 times fewer clinician hours [18]. Innovative implementation approaches should be explored to further optimize the scalability, acceptability, and engagement of DMHIs that incorporate human support [53,54].

Acceptability can be further enhanced by improving the conversational agent. Being misunderstood by the agent was a significant source of frustration in this study and is a common limitation of tree-based dialogue systems [15,16,51,55,56]. Recent advances in generative AI have transformed human-technology interaction by enabling dialogue systems that emulate fluent human conversation, making users feel heard and understood [57]. Users have reported finding rule-based mental health apps less satisfying than those using generative conversations [58], and recent findings suggest that generative conversational agents may have a greater ameliorative impact on psychological distress than retrieval-based ones [59]. While further research is needed, these preliminary results indicate that using generative AI could enhance engagement and effectiveness in DMHIs. However, generative AI also introduces new patient risks [60,61]. Its integration into DMHIs should therefore be guided by mental health professionals [62], with rigorous frameworks in place to evaluate clinical risk and quality in mental health contexts [63].

Interestingly, despite the conversational agent's limitations, many participants remained engaged because of the content's usefulness. This is consistent with prior research linking perceived usefulness to greater engagement [12] and to the value users place on insights gained from DMHIs [15,56]. In this study, 180 out of 203 (88.7%) participants found the intervention satisfactory (180/203; SUTAQ *Satisfaction* subscale), 164 out of 203 (80.8%) reported that it enhanced care (SUTAQ *Enhanced Care* subscale), and 159 out of 190 (83.7%) found it rewarding and worthwhile (UES *Rewarding* subscale). Postintervention interviews provided further insights into how perceiving benefits—from learning new tools and techniques to observing improvements in symptoms—was key to maintaining engagement, consistent with theory identifying perceived efficacy as a core construct influencing acceptability [20,21]. While prior research on conversational agent–led DMHIs has emphasized the role of factors such as empathy and human-likeness in fostering engagement [19], our findings underscore that expected outcomes and perceived efficacy remain central for individuals using these programs.

Limitations

Limitations of this study are a lack of sample diversity and potential biases arising from the limited perspectives of the clinical and research teams. The program design and interpretation of qualitative data were shaped by the team's cultural context. Given mental health inequities [64] and varying attitudes toward AI [65], research on DMHI acceptability in diverse populations is essential but remains limited [19,66]. Dedicated research is also needed on the personalization and adaptation of conversational agent–led DMHIs to user background, identity, culture, and socioeconomic context [67], as well as on the potential and limitations of generative AI in supporting such adaptation—for example, through language translation and tailoring, metaphor usage, or individualized treatment goals [68]. Co-production with diverse clinicians, researchers, and users with lived experience must be prioritized.

Quantitative findings may be biased toward participants who adhered to the full study protocol. For example, those who completed postintervention questionnaires were more likely to reach the minimum meaningful program dose within 9 weeks (168/203, 82.8%) compared with the full enrolled sample (180/300, 60%). The interviewed subsample, selected preintervention, showed similar engagement to the full sample (14/21, 67%), while those who completed postintervention interviews demonstrated higher engagement (12/16, 75%). Exploring participants' expectations before the study helped mitigate this limitation and provided insights into factors important for acceptability, regardless of engagement. Nonetheless, findings primarily reflect the perspectives of the

XSL•FO
RenderX

most engaged participants. Actionable guidance for DMHI development should therefore focus on enhancing engagement across the broader population. Further research with nonengagers is essential to understand barriers to DMHI acceptability.

Moreover, there are limitations to the validity of the self-reported measures (UES, SUS, and SUTAQ) in this context. Engagement with DMHIs involves different motivations and behavioral processes among users experiencing mental ill health compared with the general digital systems or telehealth services on which these measures were developed and validated, making interpretability in this context challenging [7]. For example, the UES *Focused Attention* subscale (reflecting absorption in the experience) may be less applicable in a DMHI than in a noninterventional system, as suggested by the qualitative analysis, where participants described the program as demanding but rewarding. Additionally, the interpretability of the scales was constrained by the scarcity of suitable benchmarks. To address this, we reported item-level response data to aid interpretation, and qualitative findings provided deeper contextual insights.

Implications for Future Research and Implementation

This study offers insights to inform future research and implementation efforts aimed at enhancing the acceptability, engagement, and usability of conversational agent–led DMHIs. Key areas of focus are balancing user autonomy with structure and guidance, identifying when human support is most beneficial, and reflecting progress back to users to strengthen perceived efficacy and engagement. Real-world implementation pilots will be important to maximize adoption and outcomes [69,70]. The variability in user preferences across these domains underscores the need to explore personalized solutions. Advances in natural language generation using large language models offer powerful tools for creating varied, personalized interactions that surpass the limitations of rule-based dialogue systems [58]. However, their implementation must be approached with caution, ensuring outputs are guided by evidence-based protocols and safeguarded against inappropriate or potentially harmful content. Given users' parallel needs for both structure and effectiveness, future DMHIs should explore safe and engaging ways to integrate generative AI within structured, evidence-based programs that incorporate feedback from lived experience. Finally, the study highlights the need to develop validated measures that capture the unique dimensions of user experience in conversational agent–led DMHIs, enabling more accurate quantitative evaluations of acceptability and engagement and facilitating further research into their relationship with clinical outcomes.

Conclusions

This study provides evidence for the acceptability of a structured, conversational agent–driven digital program with human support for adults experiencing symptoms of generalized anxiety, along with insights to optimize such interventions. Understanding patient acceptability through multiple methods is crucial for enhancing engagement with DMHIs, a key challenge limiting their real-world impact. The findings underscore the importance of design, clinical, and implementation factors in supporting adherence. Importantly, patient expectations of conversational agents are rapidly shifting with the widespread availability of generative AI tools such as ChatGPT. Generative AI has the potential to transform the user experience with DMHIs by enabling dynamic, empathetic, and personalized interactions across modalities and languages. However, rigorous evaluation is essential to ensure patient safety and equitable health outcomes, alongside continued education and research into its acceptability. If integrated safely within structured, evidence-based, and effective DMHIs—shown here to be acceptable to users—this technology could revolutionize the delivery of mental health care and improve patient outcomes globally.

Acknowledgments

We extend our gratitude to the participants in the study and to the dedicated clinicians and support staff involved. We thank Gerald Chan, Stephen Bruso, Andy Richards, Ann Hayes, David Icke, Michael Black, Clare Hurley, Florian Erber, Richard Marsh, Sam Williams, and Jo Parfrey for their support and encouragement. We are grateful to Professor Thalia Eley for introducing us to NIHR BioResource. We thank the NIHR BioResource volunteers for their participation and gratefully acknowledge the NIHR BioResource centers, NHS Trusts, and staff for their contributions. We also thank the National Institute for Health and Care Research, NHS Blood and Transplant, and Health Data Research UK as part of the Digital Innovation Hub Program. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health and Social Care. This research was funded by ieso Digital Health Ltd.

Data Availability

Owing to the potential risk of patient identification, and following data privacy policies at ieso, individual-level data are not available. Aggregated data are available upon request, subject to a data-sharing agreement with ieso. Data requests should be sent to the corresponding author (PP) and will be responded to within 30 days.

Authors' Contributions

CEP, E Marshall, and ADB conceptualized the study. CEP and PP drafted the paper. PP, CEP, E Millgate, MB, MZ, GW, AC, and JB contributed to data analyses and interpretation. SD, MZ, CEP, and E Marshall conducted the study. All authors contributed to the interpretation of results, revised the paper, and approved the final version.

XSL•FO
RenderX

Conflicts of Interest

All investigators are employees of ieso Digital Health Limited (the company funding this research) or its subsidiaries. None of the authors had a direct financial incentive related to the results of this study or the publication of this manuscript.

Multimedia Appendix 1
Additional analysis methods.
[DOCX File , 623 KB - humanfactors_v12i1e76377_app1.docx]

Multimedia Appendix 2
Self-reported measures of user experience.
[DOCX File , 595 KB - humanfactors_v12i1e76377_app2.docx]

References

1. Mental health and COVID-19: evidence of the pandemic's impact. World Health Organization (WHO). Geneva, Switzerland: WHO; 2022 Mar 02. URL: https://www.who.int/publications/i/item/WHO-2019-nCoV-Sci_Brief-Mental_health-2022.1 [accessed 2024-07-08]
2. Psychiatrists per 100,000 people. Our World in Data. 2020. URL: https://ourworldindata.org/grapher/psychiatrists-working-in-the-mental-health-sector [accessed 2024-07-03]
3. BACP Mindometer 2023. British Association for Counselling Psychotherapy (BACP). 2023. URL: https://www.bacp.co.uk/about-us/about-bacp/bacp-mindometer-2023/ [accessed 2024-07-08]
4. Alonso J, Liu Z, Evans-Lacko S, Sadikova E, Sampson N, Chatterji S, WHO World Mental Health Survey Collaborators. Treatment gap for anxiety disorders is global: results of the World Mental Health Surveys in 21 countries. Depress Anxiety 2018 Mar;35(3):195-208 [FREE Full text] [doi: 10.1002/da.22711] [Medline: 29356216]
5. Psychologists struggle to meet demand amid mental health crisis. 2022 COVID-19 Practitioner Impact Survey. American Psychological Association (APA). 2022. URL: https://www.apa.org/pubs/reports/practitioner/2022-covid-psychologist-workload [accessed 2024-07-08]
6. Roland J, Lawrance E, Insel T, Christensen H. The digital mental health revolution: transforming care through innovation and scale-up. World Innovation Summit for Health. 2020. URL: https://wish.org.qa/research-report/the-digital-mental-health-revolution-transforming-care-through-innovation-and-scale-up/ [accessed 2024-07-08]
7. Ng MM, Firth J, Minen M, Torous J. User engagement in mental health apps: a review of measurement, reporting, and validity. Psychiatr Serv 2019 Jul 01;70(7):538-544. [doi: 10.1176/appi.ps.201800519] [Medline: 30914003]
8. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. Evid Based Ment Health 2018 Aug 05;21(3):116-119 [FREE Full text] [doi: 10.1136/eb-2018-102891] [Medline: 29871870]
9. Baumel A, Muench F, Edan S, Kane JM. Objective user engagement with mental health apps: systematic search and panel-based usage analysis. J Med Internet Res 2019 Sep 25;21(9):e14567 [FREE Full text] [doi: 10.2196/14567] [Medline: 31573916]
10. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
11. Boucher EM, Raiker JS. Engagement and retention in digital mental health interventions: a narrative review. BMC Digit Health 2024 Aug 08;2(1):e. [doi: 10.1186/s44247-024-00105-9]
12. Borghouts J, Eikey E, Mark G, De Leon C, Schueller SM, Schneider M, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. J Med Internet Res 2021 Mar 24;23(3):e24387 [FREE Full text] [doi: 10.2196/24387] [Medline: 33759801]
13. Boucher EM, Harake NR, Ward HE, Stoeckl SE, Vargas J, Minkel J, et al. Artificially intelligent chatbots in digital mental health interventions: a review. Expert Rev Med Devices 2021 Dec 31;18(sup1):37-49 [FREE Full text] [doi: 10.1080/17434440.2021.2013200] [Medline: 34872429]
14. Jabir AI, Lin X, Martinengo L, Sharp G, Theng Y, Tudor Car L. Attrition in conversational agent-delivered mental health interventions: systematic review and meta-analysis. J Med Internet Res 2024 Feb 27;26:e48168 [FREE Full text] [doi: 10.2196/48168] [Medline: 38412023]
15. Fitzpatrick KK, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. JMIR Ment Health 2017 Jun 06;4(2):e19 [FREE Full text] [doi: 10.2196/mental.7785] [Medline: 28588005]
16. Inkster B, Sarda S, Subramanian V. An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: real-world data evaluation mixed-methods study. JMIR Mhealth Uhealth 2018 Nov 23;6(11):e12106 [FREE Full text] [doi: 10.2196/12106] [Medline: 30470676]

XSL•FO
RenderX

17. Linardon J, Torous J, Firth J, Cuijpers P, Messer M, Fuller-Tyszkiewicz M. Current evidence on the efficacy of mental health smartphone apps for symptoms of depression and anxiety. A meta-analysis of 176 randomized controlled trials. World Psychiatry 2024 Feb 12;23(1):139-149 [FREE Full text] [doi: 10.1002/wps.21183] [Medline: 38214614]
18. Palmer CE, Marshall E, Millgate E, Warren G, Ewbank M, Cooper E, et al. Combining artificial intelligence and human support in mental health: digital intervention with comparable effectiveness to human-delivered care. J Med Internet Res 2025 May 13;27:e69351 [FREE Full text] [doi: 10.2196/69351] [Medline: 40152000]
19. Limpanopparat S, Gibson E, Harris DA. User engagement, attitudes, and the effectiveness of chatbots as a mental health intervention: a systematic review. Computers in Human Behavior: Artificial Humans 2024 Aug;2(2):100081. [doi: 10.1016/j.chbah.2024.100081]
20. Perski O, Short CE. Acceptability of digital health interventions: embracing the complexity. Transl Behav Med 2021 Jul 29;11(7):1473-1480 [FREE Full text] [doi: 10.1093/tbm/ibab048] [Medline: 33963864]
21. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res 2017 Jan 26;17(1):88 [FREE Full text] [doi: 10.1186/s12913-017-2031-8] [Medline: 28126032]
22. Abd-Alrazaq AA, Alajlani M, Ali N, Denecke K, Bewick BM, Househ M. Perceptions and opinions of patients about mental health chatbots: scoping review. J Med Internet Res 2021 Jan 13;23(1):e17828 [FREE Full text] [doi: 10.2196/17828] [Medline: 33439133]
23. Lee HS, Wright C, Ferranto J, Buttimer J, Palmer CE, Welchman A, et al. Artificial intelligence conversational agents in mental health: patients see potential, but prefer humans in the loop. Front Psychiatry 2024 Jan 31;15:1505024 [FREE Full text] [doi: 10.3389/fpsyt.2024.1505024] [Medline: 39957757]
24. Haque MDR, Rubya S. An overview of chatbot-based mobile mental health apps: insights from app description and user reviews. JMIR Mhealth Uhealth 2023 May 22;11:e44838 [FREE Full text] [doi: 10.2196/44838] [Medline: 37213181]
25. Benda N, Desai P, Reza Z, Zheng A, Kumar S, Harkins S, et al. Patient perspectives on AI for mental health care: cross-sectional survey study. JMIR Ment Health 2024 Sep 18;11:e58462 [FREE Full text] [doi: 10.2196/58462] [Medline: 39293056]
26. Humphreys J, Gill M, Rooney S, Ahmad Z. Digital mental health technology: user and public perspectives (research report from Woodnewton). Medicines & Healthcare products Regulatory Agency (MHRA). 2024 Apr. URL: https://assets.publishing.service.gov.uk/media/672dde575437e298ae64ce6e/dmht-report-woodnewton.pdf [accessed 2024-07-08]
27. Ahmed A, Aziz S, Khalifa M, Shah U, Hassan A, Abd-Alrazaq A, et al. Thematic analysis on user reviews for depression and anxiety chatbot apps: machine learning approach. JMIR Form Res 2022 Mar 11;6(3):e27654 [FREE Full text] [doi: 10.2196/27654] [Medline: 35275069]
28. Fulmer R, Joerin A, Gentile B, Lakerink L, Rauws M. Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety: randomized controlled trial. JMIR Ment Health 2018 Dec 13;5(4):e64 [FREE Full text] [doi: 10.2196/mental.9782] [Medline: 30545815]
29. He Y, Yang L, Zhu X, Wu B, Zhang S, Qian C, et al. Mental health chatbot for young adults with depressive symptoms during the COVID-19 pandemic: single-blind, three-arm randomized controlled trial. J Med Internet Res 2022 Nov 21;24(11):e40719 [FREE Full text] [doi: 10.2196/40719] [Medline: 36355633]
30. Liu H, Peng H, Song X, Xu C, Zhang M. Using AI chatbots to provide self-help depression interventions for university students: a randomized trial of effectiveness. Internet Interv 2022 Mar;27:100495. [doi: 10.1016/j.invent.2022.100495] [Medline: 35059305]
31. Oh J, Jang S, Kim H, Kim J. Efficacy of mobile app-based interactive cognitive behavioral therapy using a chatbot for panic disorder. Int J Med Inform 2020 Aug;140:104171. [doi: 10.1016/j.ijmedinf.2020.104171] [Medline: 32446158]
32. O'Cathain A, Croot L, Duncan E, Rousseau N, Sworn K, Turner KM, et al. Guidance on how to develop complex interventions to improve health and healthcare. BMJ Open 2019 Aug 15;9(8):e029954. [doi: 10.1136/bmjopen-2019-029954] [Medline: 31420394]
33. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. J Med Internet Res 2017 Jun 29;19(6):e232 [FREE Full text] [doi: 10.2196/jmir.7126] [Medline: 28663162]
34. Holdener M, Gut A, Angerer A. Applicability of the user engagement scale to mobile health: a survey-based quantitative study. JMIR Mhealth Uhealth 2020 Jan 03;8(1):e13244 [FREE Full text] [doi: 10.2196/13244] [Medline: 31899454]
35. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. J Med Internet Res 2015 Jan 30;17(1):e30 [FREE Full text] [doi: 10.2196/jmir.4055] [Medline: 25639757]
36. ieso Group. URL: https://www.iesogroup.com/ [accessed 2025-07-08]
37. Medical devices — quality management systems — requirements for regulatory purposes. International Organization for Standardization (ISO). Geneva, Switzerland: ISO; 2016. URL: https://www.iso.org/standard/59752.html [accessed 2025-09-14]
38. ieso online therapy. ieso Group. URL: https://www.iesohealth.com/ [accessed 2025-07-08]
39. NIHR BioResource. URL: https://www.bioresource.nihr.ac.uk/ [accessed 2025-07-08]

XSL•FO
RenderX

40. NHS talking therapies for anxiety and depression manual. NHS England. 2018. URL: https://www.england.nhs.uk/publication/the-improving-access-to-psychological-therapies-manual/ [accessed 2025-07-08]
41. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res 2016 Nov 27;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
42. O'Brien HL, Cairns P, Hall M. A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. International Journal of Human-Computer Studies 2018 Apr;112:28-39. [doi: 10.1016/j.ijhcs.2018.01.004]
43. Brooke J. SUS: a 'quick and dirty' usability scale. In: Usability Evaluation in Industry. London, UK: CRC Press; 1996:207-212.
44. Hirani SP, Rixon L, Beynon M, Cartwright M, Cleanthous S, Selva A, et al. Quantifying beliefs regarding telehealth: development of the Whole Systems Demonstrator Service User Technology Acceptability Questionnaire. J Telemed Telecare 2017 May;23(4):460-469. [doi: 10.1177/1357633X16649531] [Medline: 27224997]
45. Dovetail. URL: https://dovetail.com/ [accessed 2025-07-08]
46. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2008 Jul 21;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
47. Braun V, Clarke V. Thematic Analysis. A Practical Guide. London, UK: SAGE Publications; Nov 2021.
48. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013 Sep 18;13:117 [FREE Full text] [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]
49. Ritchie J, Spencer L. Qualitative data analysis for applied policy research. In: Analyzing Qualitative Data. London, UK: Routledge; 1994:173-194.
50. Torous J, Michalak EE, O'Brien HL. Digital health and engagement-looking behind the measures and methods. JAMA Netw Open 2020 Jul 01;3(7):e2010918. [doi: 10.1001/jamanetworkopen.2020.10918] [Medline: 32678446]
51. Gaffney H, Mansell W, Tai S. Conversational agents in the treatment of mental health problems: mixed-method systematic review. JMIR Ment Health 2019 Oct 18;6(10):e14166. [doi: 10.2196/14166] [Medline: 31628789]
52. Gilbody S, Brabyn S, Lovell K, Kessler D, Devlin T, Smith L, REEACT collaborative. Telephone-supported computerised cognitive-behavioural therapy: REEACT-2 large-scale pragmatic randomised controlled trial. Br J Psychiatry 2017 May;210(5):362-367 [FREE Full text] [doi: 10.1192/bjp.bp.116.192435] [Medline: 28254959]
53. Werntz A, Amado S, Jasman M, Ervin A, Rhodes JE. Providing human support for the use of digital mental health interventions: systematic meta-review. J Med Internet Res 2023 Feb 06;25:e42864 [FREE Full text] [doi: 10.2196/42864] [Medline: 36745497]
54. Bernstein EE, Weingarden H, Wolfe EC, Hall MD, Snorrason I, Wilhelm S. Human support in app-based cognitive behavioral therapies for emotional disorders: scoping review. J Med Internet Res 2022 Apr 08;24(4):e33307 [FREE Full text] [doi: 10.2196/33307] [Medline: 35394434]
55. Beatty C, Malik T, Meheli S, Sinha C. Evaluating the therapeutic alliance with a free-text CBT conversational agent (Wysa): a mixed-methods study. Front Digit Health 2022 Apr 11;4:847991 [FREE Full text] [doi: 10.3389/fdgth.2022.847991] [Medline: 35480848]
56. Malik T, Ambrose AJ, Sinha C. Evaluating user feedback for an artificial intelligence-enabled, cognitive behavioral therapy-based mental health app (Wysa): qualitative thematic analysis. JMIR Hum Factors 2022 Apr 12;9(2):e35668 [FREE Full text] [doi: 10.2196/35668] [Medline: 35249886]
57. Yin Y, Jia N, Wakslak CJ. AI can help people feel heard, but an AI label diminishes this impact. Proc Natl Acad Sci U S A 2024 Apr 02;121(14):e2319112121. [doi: 10.1073/pnas.2319112121] [Medline: 38551835]
58. Siddals S, Torous J, Coxon A. "It happened to be the perfect thing": experiences of generative AI chatbots for mental health. Npj Ment Health Res 2024 Oct 27;3(1):48. [doi: 10.1038/s44184-024-00097-4] [Medline: 39465310]
59. Li H, Zhang R, Lee Y, Kraut RE, Mohr DC. Systematic review and meta-analysis of AI-based conversational agents for promoting mental health and well-being. NPJ Digit Med 2023 Dec 19;6(1):236 [FREE Full text] [doi: 10.1038/s41746-023-00979-5] [Medline: 38114588]
60. Weidinger L, Mellor J, Rauh M. Ethical and social risks of harm from language models. arXiv Preprint posted online on December 8, 2021 [FREE Full text] [doi: 10.48550/arXiv.2112.04359]
61. Blease C, Torous J. ChatGPT and mental healthcare: balancing benefits with risks of harms. BMJ Ment Health 2023 Nov 10;26(1):e300884 [FREE Full text] [doi: 10.1136/bmjment-2023-300884] [Medline: 37949485]
62. Elyoseph Z, Gur T, Haber Y, Simon T, Angert T, Navon Y, et al. An ethical perspective on the democratization of mental health with generative AI. JMIR Ment Health 2024 Oct 17;11:e58011-e58011 [FREE Full text] [doi: 10.2196/58011] [Medline: 39417792]
63. Stade EC, Stirman SW, Ungar LH, Boland CL, Schwartz HA, Yaden DB, et al. Large language models could change the future of behavioral healthcare: a proposal for responsible development and evaluation. Npj Ment Health Res 2024 Apr 02;3(1):12 [FREE Full text] [doi: 10.1038/s44184-024-00056-z] [Medline: 38609507]
64. Panchal N, Hill L, Artiga S, Hamel L. Racial and ethnic disparities in mental health care: findings from the KFF Survey of racism, discrimination and health. KFF. 2024 May 23. URL: https://tinyurl.com/4rva7jt7 [accessed 2025-09-12]

XSL•FO
RenderX

65. Gillespie N, Lockey S, Curtis C, Pool J, Akbari A. Trust in artificial intelligence: a global study. The University of Queensland and KPMG Australia. 2023 Feb 22. URL: https://assets.kpmg.com/content/dam/kpmgsites/au/pdf/2023/trust-in-ai-global-insights-2023.pdf.coredownload.inline.pdf [accessed 2025-09-11]
66. Khosravi M, Azar G. Factors influencing patient engagement in mental health chatbots: a thematic analysis of findings from a systematic review of reviews. Digit Health 2024 Apr 22;10:20552076241247983 [FREE Full text] [doi: 10.1177/20552076241247983] [Medline: 38655378]
67. Naderbagi A, Loblay V, Zahed IUM, Ekambareshwar M, Poulsen A, Song YJC, et al. Cultural and contextual adaptation of digital health interventions: narrative review. J Med Internet Res 2024 Jul 09;26:e55130. [doi: 10.2196/55130] [Medline: 38980719]
68. Spanhel K, Balci S, Feldhahn F, Bengel J, Baumeister H, Sander LB. Cultural adaptation of internet- and mobile-based interventions for mental disorders: a systematic review. NPJ Digit Med 2021 Aug 25;4(1):128 [FREE Full text] [doi: 10.1038/s41746-021-00498-1] [Medline: 34433875]
69. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
70. Holtrop JS, Estabrooks PA, Gaglio B, Harden SM, Kessler RS, King DK, et al. Understanding and applying the RE-AIM framework: clarifications and resources. J Clin Transl Sci 2021 May 14;5(1):e126 [FREE Full text] [doi: 10.1017/cts.2021.789] [Medline: 34367671]

Abbreviations

AI: artificial intelligence
CBT: cognitive behavioral therapy
CONSORT: Consolidated Standards of Reporting Trials
DMHI: digital mental health intervention
GAD-7: 7-item Generalized Anxiety Disorder
ICD-10: International Classification of Diseases, 10th Revision
NHS: National Health Service
PHQ-9: 9-item Patient Health Questionnaire
PPI: patient and public involvement
SUS: System Usability Scale
SUTAQ: Service User Technology Acceptability Questionnaire
UES: User Engagement Scale
UKCA: UK Conformity Assessed

Edited by A Kushniruk; submitted 22.04.25; peer-reviewed by S Meheli, Y Senathirajah; comments to author 21.05.25; revised version received 16.07.25; accepted 12.08.25; published 04.11.25.

Please cite as:
Papiernik P, Dzula S, Zimanyi M, Millgate E, Bouazzaoui M, Buttimer J, Warren G, Cooper E, Catarino A, Mehew S, Marshall E, Tablan V, Blackwell AD, Palmer CE
Acceptability of a Conversational Agent–Led Digital Program for Anxiety: Mixed Methods Study of User Perspectives
JMIR Hum Factors 2025;12:e76377
URL: https://humanfactors.jmir.org/2025/1/e76377
doi:10.2196/76377
PMID:40824528

©Pearla Papiernik, Sylwia Dzula, Marta Zimanyi, Edward Millgate, Malika Bouazzaoui, Jessica Buttimer, Graham Warren, Elisa Cooper, Ana Catarino, Shaun Mehew, Emily Marshall, Valentin Tablan, Andrew D Blackwell, Clare E Palmer. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 04.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Research Letter

Older Adults' Experiences With an Online Survey

Kristie Rebecca Weir[1,2], PhD; Yehya Maitah[3], BS; Sarah E Vordenberg[4], PharmD, MPH

[1]Sydney School of Public Health, Faculty of Medicine and Health, The University of Sydney, Sydney, Australia

[2]Institute of Primary Health Care (BIHAM), University of Bern, Bern, Switzerland

[3]Albany Medical College, Albany, NY, United States

[4]University of Michigan College of Pharmacy, Department of Clinical Pharmacy, Ann Arbor, MI, United States

Corresponding Author:
Kristie Rebecca Weir, PhD
Sydney School of Public Health
Faculty of Medicine and Health
The University of Sydney
City Road
Sydney, NSW 2006
Australia
Phone: 61 293512222
Email: kristie.weir@sydney.edu.au

Abstract

The study explored older adults' perceptions after participating in an online survey about medication decisions, finding that approximately 80% of participants provided positive feedback about the research methodology and their experience.

(JMIR Hum Factors 2025;12:e65684) doi:10.2196/65684

KEYWORDS

older adults; gerontology; geriatric; older; aging; online; internet; survey; questionnaire; research engagement; engagement; study subject; participant; medication decisions

Introduction

Older adults are underrepresented in research, often due to age-related biases and stringent exclusion criteria; this limits generalizability and leaves knowledge gaps [1].

The increasing prevalence of online research has the potential to increase older adult participation, given the rising internet use among this demographic [2]. While research has primarily focused on improving recruitment of older adults, an understanding of their experiences of online research is needed to effectively engage older people [3]. While most surveys include some form of piloting, they rarely capture or share participants' perceptions of the research process. We sought to explore older adults' perceptions after participating in an online experimental survey, given the potential high accessibility of this type of research.

Methods

Overview

We previously reported the main findings from a vignette-based online experiment conducted among adults aged 65 years and older from Australia, the Netherlands, the United Kingdom, and the United States [4]. Participants were recruited for this 15-minute survey through a panel of internet users administered by Qualtrics Research Panels. The study focused on contextual factors influencing how older adults (mean age 71.5, SD 5.1 years) think about deciding to stop a cardiovascular disease medication. In this study, we conducted a content analysis of free-text comments (Textbox 1). The final analysis framework included 17 codes that were subsequently consolidated into 14 themes. Descriptive statistics were used to assess the frequency of each theme.

XSL•FO
RenderX

Textbox 1. Summary of methodology for the content analysis.

1. A vignette-based online experiment was conducted among adults aged 65 years and older from Australia, the Netherlands, the United Kingdom, and the United States using sampling quotas to ensure balanced representation by country and gender. This study was registered at ClinicalTrials.gov (NCT04676282). At the end of the survey, participants were provided the opportunity to leave "any comments about this study."
2. A content analysis was conducted among participants who received the vignette about a hypothetical patient, Mrs EF. Participants from the Netherlands were excluded, as the study team for the content analysis was not fluent in Dutch.
3. Study authors read through the comments to inductively generate codes.
4. Two investigators (SEV and YM) independently coded comments, resulting in >80% agreement.
5. All discrepancies in codes were discussed until consensus was reached.

Ethical Considerations

The online experiment was registered at ClinicalTrials.gov (NCT04676282) and was deemed exempt by the University of Michigan Health Sciences and Behavioral Sciences Institutional Review Board (HUM00183129); by extension, a waiver of consent was granted. All data were collected anonymously. Participants were compensated based on the terms of their panel agreement.

Results

Participants (N=1789) most frequently did not provide any feedback in the free-text comment field (n=784, 43.8%) or wrote that they had no comments (n=487, 27.2%). Three participants (0.2%) gave unclear statements.

Participants' comments (n=515, 28.8%) were primarily positive (415/515, 80.6%), such as that the study was interesting (116/515, 22.5%). Themes and representative quotes are reported in Table 1. Participants said it made them think about their health (80/515, 15.5%) and some participants shared further health information about the study topic (48/515, 9.3%). Participants provided feedback on how to improve the study, categorized as question-specific comments (21/515, 4.1%), general suggestions (12/515, 2.3%), or country-specific comments (3/515, 0.6%). Few participants (16/515, 3.1%) provided negative feedback about the survey.

XSL•FO
RenderX

Table 1. Older adults' feedback about an online survey using a hypothetical vignette by theme, with representative quotes among respondents who provided feedback (n=515).

Themes	Representative quotes (participant code)	Participants, n (%)
Positive feedback (n=415, 80.6%)		
Interesting	• "An interesting study" (919) • "This was an interesting survey to complete" (2200)	116 (22.5)
Thought provoking	• "Interesting and insightful. Made me think a little more about how best to manage my health" (1872) • "Very good survey to ponder thoughts and beliefs" (481)	80 (15.5)
Positive feedback	• "Great stuff. Keep up the good work" (2574) • "Very good" (764)	67 (13)
Thanks for opportunity to participate	• "Thanks" (42)	43 (8.3)
Positive feedback about survey questions or structure	• "Love the format, so easy to see, follow, and understand" (1904) • "Very good questions regarding whether or not to stop a medication when you've been on it for an extended time" (2026)	35 (6.8)
Enjoyed taking survey	• "Enjoyed it" (328) • "Love doing your studies" (287)	34 (6.6)
Unusual study design	• "Very different enjoyed the variety" (540) • "Unusual but interesting" (1853)	19 (3.7)
Interested in results	• "Interested in purpose of results" (1863) • "Very interesting! Now I need to see the final results" (3481)	15 (2.9)
Interested in future studies	• "Great study, I'd do more" (4815) • "Need more like this" (2516)	6 (1.2)
Neutral feedback (n=48, 9.3%)		
Shared personal experience	• "I am afraid of going to see the dr for worries about my health" (1526) • "Thanks; my prescription drugs fall under the Federal Government program where they are funded nearly 100%" (2645)	48 (9.3)
Negative feedback or suggestions for improvement (n=52, 10.1%)		
Question-specific feedback	• "The percentage charts were confusing" (2403) • "One question said check all that apply but only one was allowed" (4279)	21 (4.1)
Negative feedback	• "Boring" (1409) • "Too many generalizations" (1102)	16 (3.1)
General suggestions for improvement	• "Consider adding a progress bar to the survey" (2027) • "Good study could be shorter" (2433	12 (2.3)
Country-specific feedback	• "Some questions are designed for the USA" (1471) • "Prescription insurance? In UK if you are over 65 prescriptions are free" (4075)	3 (0.6)

Discussion

Among older adults who provided feedback about their experiences completing an online survey, approximately 80% (415/515) of the comments were positive. Our findings signal general acceptability of the methodology, and we have implemented the practical feedback to improve our online surveys. We have become more mindful of the survey length, selected straightforward question types, and have conducted pilot testing in all target countries to ensure that questions are appropriate for all participants.

Our study had several limitations. First, less than one-third of the study participants provided any comments. Second, participants were asked if they had any comments, as opposed to more specific questions about their experience taking the survey. Finally, we acknowledge that we coded a single primary theme per comment given the short statements that were provided.

With an aging population who may spend many years in retirement, participating in research can offer benefits such as reducing social isolation and loneliness, fostering a sense of

XSL•FO
RenderX

purpose, and providing mental stimulation, and it may provide monetary incentives [5]. While our online study lacked the social benefit of in-person interaction typical of traditional research, participants reported enjoying the survey; they found it prompted self-reflection on their health, and they expressed interest in the study's outcomes and future research.

Online research is becoming more prevalent; therefore, it is important to make sure this methodology is inclusive of older adults [5]. More than half of adults older than 65 years use the internet, yet they are the least likely to have a home computer [6]. While some researchers attribute this limited use to age-related functional decline, others argue that the main barriers are negative attitudes such as fear, anxiety, and low motivation—barriers that are modifiable. Anxiety about using the internet and digital technologies often leads to self-imposed limitations and low confidence, with older adults frequently underestimating their knowledge and abilities compared to younger users [6]. This underscores the importance of studies like ours that highlight the positive experiences of older adults in online research.

Acknowledgments

This study was funded by award R24AG064025 from the US Deprescribing Research Network via the Northern California Institute for Research and Education through the National Institute on Aging; a Swiss Government Excellence Scholarship (KRW), the Swiss National Science Foundation Scientific Exchanges (grant IZSEZ0_213475/1; KRW), and an Emerging Leader Research Fellowship (2017295) from the Australian National Health and Medical Research Council (KRW). The funders had no role in the design or conduct of the study; collection, management, analysis, or interpretation of the data; preparation, review, or approval of the manuscript; or the decision to submit the manuscript for publication. KRW and SEV had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Authors' Contributions

Concept and design: KRW, SEV
Acquisition, analysis, or interpretation of data: all authors
Drafting of the manuscript: all authors
Critical review of the manuscript for important intellectual content: all authors
Statistical analysis: SEV
Obtained funding: SEV
Administrative, technical, or material support: KRW
Supervision: KRW, SEV

Conflicts of Interest

KRW reports receiving grant funding from the Swiss National Science Foundation (SNSF), the National Health and Medical Research Council, and the Swiss Confederation during the conduct of the study. SEV reports receiving grant funding from the US Deprescribing Research Network during the conduct of the study.

References

1. Markham SC, McNab J, O'Loughlin K, Clemson L. International responses addressing the under-representation of older people in clinical research. Australas J Ageing 2023 Dec;42(4):762-768. [doi: 10.1111/ajag.13234] [Medline: 37724905]
2. Faverio M. Share of those 65 and older who are tech users has grown in the past decade. Pew Research Center. URL: https://www.pewresearch.org/short-reads/2022/01/13/share-of-those-65-and-older-who-are-tech-users-has-grown-in-the-past-decade/ [accessed 2024-08-12]
3. Remillard ML, Mazor KM, Cutrona SL, Gurwitz JH, Tjia J. Systematic review of the use of online questionnaires of older adults. J Am Geriatr Soc 2014 Apr;62(4):696-705 [FREE Full text] [doi: 10.1111/jgs.12747] [Medline: 24635138]
4. Weir KR, Vordenberg SE, Scherer AM, Jansen J, Schoenborn N, Todd A. Exploring different contexts of statin deprescribing: a vignette-based experiment with older adults across four countries. J Gen Intern Med 2024 Jul;39(9):1773-1776 [FREE Full text] [doi: 10.1007/s11606-024-08698-7] [Medline: 38514582]
5. Lindeman D. Improving the independence of older adults through technology: directions for public policy. Public Policy Aging Rep 2017;27(2):49-52. [doi: 10.1093/PPAR/PRX011]
6. Vroman KG, Arthanat S, Lysack C. "Who over 65 is online?" Older adults' dispositions toward information communication technology. Comput Human Behav 2015 Feb;43:156-166. [doi: 10.1016/j.chb.2014.10.018]

XSL•FO
RenderX

Edited by A Kushniruk, E Borycki; submitted 22.08.24; peer-reviewed by J Rodger, S Serhal, A Vassallo; comments to author 19.12.24; revised version received 19.12.24; accepted 20.12.24; published 08.01.25.

Please cite as:
Weir KR, Maitah Y, Vordenberg SE
Older Adults' Experiences With an Online Survey
JMIR Hum Factors 2025;12:e65684
URL: https://humanfactors.jmir.org/2025/1/e65684
doi:10.2196/65684
PMID:

©Kristie Rebecca Weir, Yehya Maitah, Sarah E Vordenberg. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 08.01.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Bridging Policy and Practice in Telemedicine Follow-Up Identification: Multicenter Mixed Methods Study in Beijing

Yanli Lyu[1,2,3*], MPH; Xi Li[4*], MS; Huixian Li[1,2], MS; Mingyu Gu[2], MS; Xuedong Xu[1,2,3], MS; Yipei Wang[4], MBA, PhD; Changxiao Jin[4,5], MS

[1]Office of Internet Hospital, Peking University Third Hospital, Beijing, China

[2]Department of Medical Affairs, Peking University Third Hospital, Beijing, China

[3]Beijing Municipal Center for Healthcare Quality Management and Improvement in Internet-Based Medical Services, Beijing, China

[4]Institute of Hospital Management, Peking University Third Hospital, North Huayuan Road 49, Beijing, China

[5]Party Committee Office and Dean's Office, Peking University Third Hospital, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Yipei Wang, MBA, PhD
Institute of Hospital Management, Peking University Third Hospital, North Huayuan Road 49, Beijing, China

Abstract

Background: Telemedicine services have been developing rapidly worldwide. Following the 2018 policy enabling telemedicine for follow-up patients, this service model has gradually gained popularity in China. However, little has been done to understand the policy's implementation across different types of medical institutions or to evaluate its effectiveness.

Objective: This study aims to (1) investigate the patient eligibility assessment process in various types of institutions for telemedicine services in Beijing, (2) elucidate institutions' rationale for adopting such approaches, (3) analyze discrepancies between policy and practice, and (4) provide references for the development of telemedicine services.

Methods: This mixed methods study involved 36 medical institutions in Beijing, determined based on placing in the top 20% for both service volume and service quality in 2023. The study was conducted in 2 phases. First, quantitative analyses were conducted based on the questionnaires collected from each institution's contact to gather information about the assessment process and patient prerequisites. Subsequently, qualitative analyses were conducted through thematic analysis of 36 semistructured interviews with each institution's contact to acquire their considerations of this practice.

Results: These 36 institutions contributed 968,786 telemedicine visits, representing 89.5% of the total service volume in Beijing. In practice, each of the visits underwent a 2-stage eligibility assessment before the physician officially accepted the patient's request. In the first stage, for assessment approaches, 86.2% (25/29) of the nonprofit, tertiary hospitals and 14.3% (1/7) of the private, for-profit institutions automatically assessed patient eligibility, while others did it manually. The assessment was based on the scope of previous visit location, diagnosis of previous visits, and visit interval. For the scope of visit location, 22 hospitals required prior visits to the same institution. For diagnosis, 7 hospitals required patients to have an identical diagnosis. For visit intervals, 11 hospitals required it to be within 6 months. The second stage assessment was conducted by physicians. Compared with policy requirements, nonprofit hospitals had stricter requirements for the scope of visit locations. The main reasons for these discrepancies included distrust in the medical outcomes from other institutions (19/26, 73.1%) and difficulties in handling interinstitutional medical disputes (18/26, 69.2%). In addition, 61.1% (22/36) of the institutions indicated that terminations of telemedicine services were primarily due to the patient's conditions.

Conclusions: This pioneering multicenter, mixed methods study delineated the patient eligibility assessment process for telemedicine services in Beijing. Discrepancies were identified between real-world practice and regulatory prerequisites. The key factors contributing to these variations included the ambiguity of policies and different priorities across institution types. Our findings suggest greater policy clarity, relaxation of regulations for new patients, and strengthened oversight of telemedicine services to improve the quality and accessibility of telemedicine.

(JMIR Hum Factors 2025;12:e75964) doi:10.2196/75964

KEYWORDS

assessment process; Beijing; China; follow-up visit; health policy; management; mixed methods; prerequisite; quality; telemedicine services

XSL•FO
RenderX

Introduction

Background

Telemedicine effectively connects individuals and their health care providers when in-person care is not necessary or not possible. Catalyzed by the pandemic, this novel service model has rapidly expanded and gained broad acceptance [1-3]. Currently, telemedicine is widely implemented in countries such as the United States, Canada, and Australia, with coverage rates reaching 35% - 60% [4-6]. In March 2020, the Centers for Medicare and Medicaid Services in the United States further facilitated adoption by removing the originating site requirement, allowing patients to receive telemedicine services from home [7]. Supported by such policy changes, telemedicine has been particularly useful in stroke treatment, psychiatry, radiology, and other chronic diseases, especially for patients in rural areas [8-11].

Similarly, in China, telemedicine services have been developing at a rapid pace [12]. In 2018, patients with "certain" chronic or common diseases were allowed to use telemedicine for follow-up visits [13], according to the Telemedicine Diagnosis and Treatment Management Measures (trial) formulated by the National Health Commission [14]. However, online initial services, including diagnoses, prescriptions, and referrals for new patients, are still prohibited [15]. By the end of 2022, the policy allowed medical institutions to prescribe medications online via telemedicine for patients presenting COVID-related symptoms. From a safety perspective, determining patient eligibility for telemedicine has remained an important consideration in policymaking. As patient demand for telemedicine continues to evolve, policies should also be updated accordingly to address these changing demands effectively.

In China, medical institutions are primarily categorized into hospitals and other medical institutions. Hospitals are further classified by function into tertiary, secondary, and first-class levels. Tertiary hospitals serve as regional medical centers with advanced capabilities in clinical care, education, and research, offering high-level specialized care across regions. Secondary hospitals provide medical and health services to multiple communities and undertake limited education and research responsibilities. First-class hospitals focus on preventive and rehabilitative care within a single community [16]. While all types of medical institutions can provide telemedicine services, tertiary hospitals remain predominant, accounting for 61.6% of the telemedicine providers by the end of 2022 [17].

Since the implementation of the telemedicine policy for follow-up patients in 2018, domestic studies have primarily focused on the service volume, operational framework, or organizational implementation [18-20], with relatively few studies evaluating the service process from the perspective of the patient's journey. Currently, only limited research has addressed telemedicine practice variances across different types of institutions [18,21], and extensive research is still needed on institutional implementation considerations and strategies. These research gaps hinder the timely updates and refinements of policies, hence impeding the effective delivery of telemedicine services.

To address this gap, our study investigated institutional approaches to telemedicine follow-up identification in Beijing. Although Beijing is a city with high-concentrated, high-quality health resources [22], its telemedicine services were comparable to those of other provinces and municipalities in terms of service accessibility and provision by tertiary hospitals. The adoption rate of telemedicine among hospitals in Beijing was within the range of 0.700 - 0.840, placing it among the top 10 provinces in China. Similar rates were observed in Jiangsu, Shanghai, Zhejiang, Ningxia, and Xinjiang [23]. In Beijing, about 30% of the tertiary hospitals provided telemedicine services, ranking 16th among 31 provinces [17]. Therefore, Beijing is representative in terms of service volume and the dual-channel service model. Its experience in formulating telemedicine policies positions it as an emerging leader in telemedicine, providing valuable reference for other regions and countries seeking to develop and refine telemedicine policies. While regions should still tailor telemedicine strategies to their own situation and local needs, the experiences from Beijing still offer important guidance for designing region-specific policies.

Aims and Objectives

Based on the analyses of 36 institutions, this study aimed to examine discrepancies between policy and practice in telemedicine follow-up identification. By investigating patient eligibility assessment in Beijing, elucidating institutions' rationale for such approaches, and analyzing the deviations from the national policies, we identified the key issues, proposed strategies to improve the efficiency of telemedicine services, and highlighted potential areas for policy refinement.

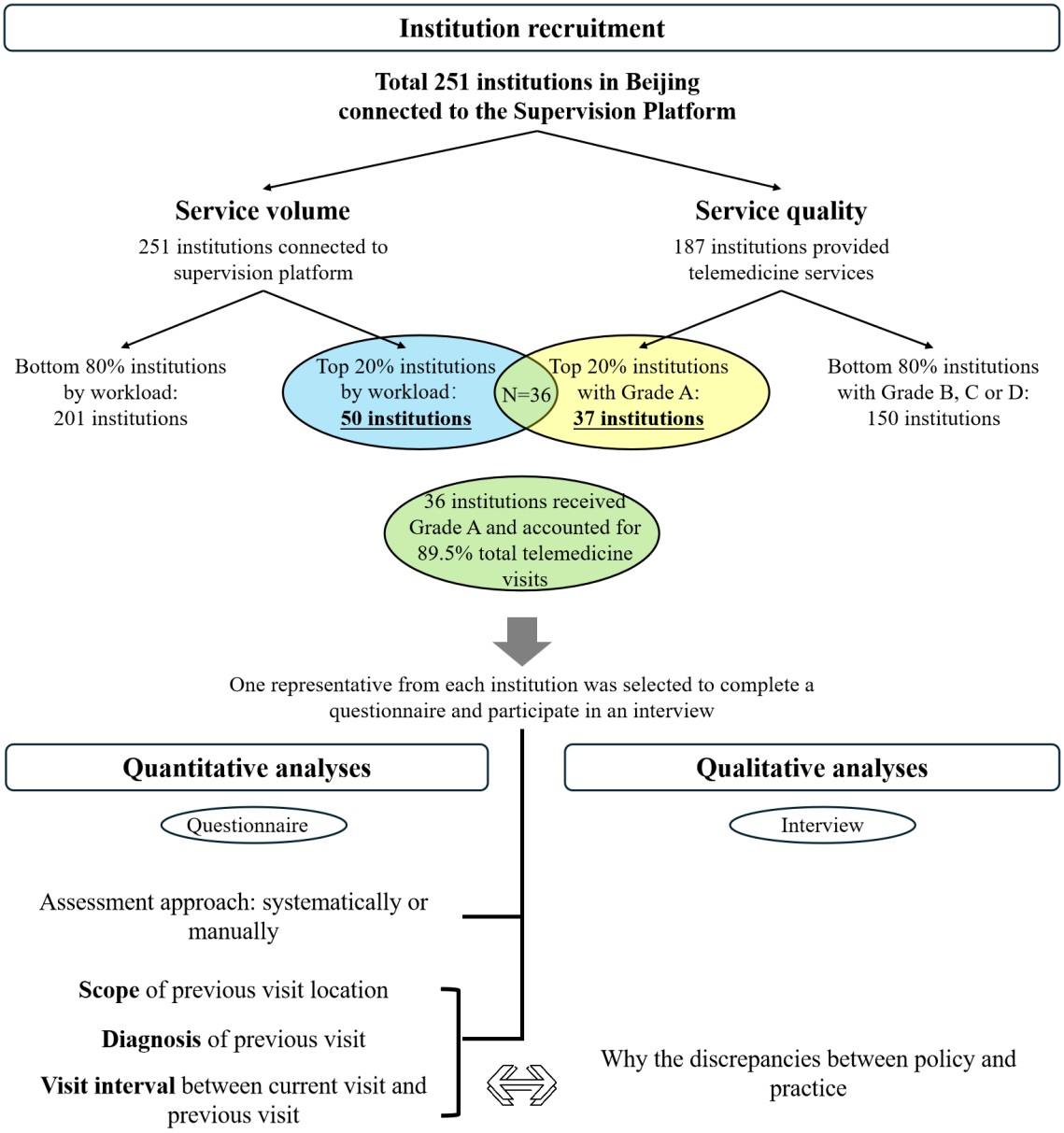
Methods

Overview

The reporting of this study followed the checklist of mixed method research studies by Lee et al [24] in *Medical Care Research and Review* (Checklist 1), adopting an explanatory sequential design consisting of 3 phases: institution recruitment, quantitative analyses, and qualitative analyses (Figure 1) [25]. During the institution recruitment phase, based on the total volume and service quality of telemedicine services in 2023, 36 institutions were selected from 251 medical institutions in Beijing. Next, in the quantitative phase, we distributed survey questionnaires to these 36 institutions asking them to specify their prerequisites for providing follow-up visits across 3 dimensions, including scope of prior visit (same physician, same department, same institution, or not specified), diagnostic consistency (identical diagnosis, similar diagnosis, or not specified), and interval between visits (within 6 mo, within 1 y, within 3 y, or not specified). Then, we calculated the frequency of each option to characterize prevailing practices. Finally, in the qualitative research, semistructured interviews were conducted to explore the potential reasons for discrepancies between policy and practice in follow-up determination.

XSL•FO
RenderX

Figure 1. Study flow diagram. The relationship and sequence of mixed methods research components, including institution recruitment, quantitative analyses, and qualitative analyses: the light blue shade indicates the number of institutions selected due to the service volume, the light yellow shade indicates the number of institutions selected due to the service quality, and the light green shade (the intersection of the blue shade and the yellow shade) indicates the institutions recruited for this study.



This design was grounded on 2 rationales. First, integrating quantitative and qualitative data facilitates a synergistic combination of their strengths. The quantitative phase utilized a structured questionnaire to capture the prevalence and distribution of patient eligibility assessment processes for telemedicine services. Concurrently, qualitative interviews offer insights into why the institutional prerequisites for patient assessment functioned. Second, combining quantitative surveys with qualitative interviews enabled data triangulation, thereby enhancing the credibility of our findings on patient assessment prerequisites across different institutions.

Since the study references multiple policy-related concepts, key terms are defined below to clarify the policy context and facilitate the interpretation of the assessment process:

1. Patient eligibility assessment process: This process evaluates whether patients meet the criteria for receiving telemedicine services. It comprises 2 stages: before treatment and during treatment.
2. Prerequisites for patient assessment: In the first stage of patient eligibility assessment, patient eligibility could be assessed either automatically by the system or manually by the physician. The assessment usually covers 3 dimensions, including the scope of previous visit location, diagnosis of previous visits, and visit interval.

XSL•FO
RenderX

3. Scope of previous visit location: Before addressing a patient-initiated request for a telemedicine visit, the scope of requested telemedicine care will be compared to the patient's prior in-person visits to determine whether the encounter involves the same hospital, the same department, or the same physician.
4. Diagnosis of previous visits: Before addressing a patient-initiated request for a telemedicine visit, the chief complaint reported for the current initiated virtual care will be compared with previous in-person visits to assess whether the present diagnosis is identical to or related to the previous diagnosis.
5. Visit interval: The time interval between the current telemedicine visit and the previous in-person visit will be reviewed to determine whether it is within 6 months, within 1 year, or within 3 years.
6. Termination: If the patient's condition does not meet the service requirements set by the proposed institution, such as violating any of the 3 dimensions described above or applying reasons, physicians could stop providing telemedicine services during either the first or second stage of the assessment process.
7. Beijing Telemedicine Service Supervision Platform (referred to as "Supervision Platform"): Developed by the Beijing Municipal Health Commission, the Supervision Platform is a centralized data aggregation platform that requires institutions to upload their telemedicine service information. The Supervision Platform can also automatically calculate the compliance monitoring scores for each institution.
8. Compliance monitoring metrics: Compliance monitoring metrics are built-in indicators within the Supervision Platform to evaluate the quality of telemedicine services. These metrics reflect situations that could lead to point deductions, including 9 qualification metrics, 17 operational metrics, and 8 supervision metrics (Multimedia Appendix 1). For all uploaded telemedicine visits, each recorded violation results in a 1-point deduction. Thus, 1 visit with several quality deficiencies can lead to multiple points deductions.
9. The Beijing Municipal Center for Healthcare Quality Management and Improvement in Internet-Based Medical Services (hereafter referred to as "the Center"): The Center is responsible for telemedicine-related oversight, such as issuing monthly statistical bulletins, conducting inspections, organizing experience-sharing sessions, and providing education and training programs. The Supervision Platform enables the Center to monitor and promptly deliver feedback to institutions.

Institution Recruitment

Based on the total volume and quality of telemedicine services in 2023, research participants were selected from 251 medical institutions in Beijing. Regular inspections by the Center revealed that institutions with higher service volume and quality tended to maintain more consistent assessment standards. Thus, institutions with high service volume and quality were selected for the study.

The volume of telemedicine services referred to the number of telemedicine visits uploaded by medical institutions to the Supervision Platform. A total of 50 institutions ranked in the top 20% by service volume were initially considered from the perspective of service volume.

The quality of telemedicine services was determined based on compliance monitoring scores. The annual average compliance monitoring deduction per visit for each institution was calculated. Institutions were ranked in ascending order based on their scores. The top 20% (n=37) of the institutions received Grade A, indicating the presence of a standardized and efficient management system. These institutions were considered from the perspective of service quality.

We then intersected the top 20% of the institutions in service volume with those by service quality, yielding a final sample of 36 institutions (Figure 1). These 36 institutions consisted of 29 nonprofit tertiary hospitals and 7 private, for-profit institutions. Among them, the 29 nonprofit hospitals included 14 general hospitals, 11 specialized hospitals, and 4 traditional Chinese medicine (TCM) hospitals, covering all categories of tertiary public hospitals in Beijing.

These 36 institutions collectively accounted for 89.5% of the overall telemedicine service volume in Beijing, representing both nonprofit and private, for-profit institutions with high service volume and quality. Therefore, their patient eligibility assessment processes could reflect the mainstream practice of telemedicine services.

Data Collection and Management Procedures

Platform-Based Data Extraction

In August 2024, the institutional information and 2023 service volume data for each of the 36 institutions were obtained from the Supervision Platform.

Questionnaire Survey

The approaches and prerequisites of patient eligibility assessment for telemedicine services in medical institutions were collected through questionnaire surveys conducted in September 2024. Each institution designated a contact for communication with the Center, and their demographic information was compiled in a spreadsheet. A structured questionnaire with 5 questions was designed and distributed sequentially to the contacts of all 36 institutions via Wenjuanxing (an online survey platform), with a 1-week deadline for completion. The survey questionnaire is available in Multimedia Appendix 2. Response progress and data completeness were monitored in real time through the Wenjuanxing administrator interface. Reminders were sent to nonrespondents, and incomplete submissions were flagged for follow-up to ensure data integrity.

Semistructured Interview

The considerations of these prerequisites for patient assessment were collected through semistructured interviews. After completing the questionnaire online, the contacts of the 36 institutions were interviewed individually via telephone calls from November 2024 to February 2025. All participants

XSL•FO
RenderX

provided verbal informed consent. The semistructured interviews were conducted in Mandarin Chinese by the authors YL, XL, HL, and MG. At least 2 of them were present to conduct the interview in each session. All interviewers received professional training in qualitative interviewing and had extensive experience conducting qualitative research. The semistructured interview outline is provided in Multimedia Appendix 3. We recorded the interviews using a digital voice recorder. The interview phase concluded once interviews with all institutions were completed and frequency statistics were compiled. The interviews had an average length of 15 minutes, with an overall length of 526 minutes. Notes were taken simultaneously in Microsoft Word 2019 (Microsoft Corp.), and recordings were transcribed into textual materials within 48 hours after the interviews by the interviewers present on-site. The interviewees were allocated a randomly selected unique identification number.

Data Analysis

Descriptive statistics were used to analyze the quantitative data using Microsoft Excel 2019 (Microsoft Corporation) and SPSS 22 (IBM Corp). The distribution of all categorical variables (eg, type of institution, gender, age, role in the department, and departmental affiliation of the interviewees) was reported as frequencies (n) with percentages (%). The service volume was reported as mean with SD and median with IQR. Frequency distribution plots were generated to visualize the distribution of prerequisites for patient assessment.

Qualitative analysis followed an inductive approach, and the data were systematized and analyzed by manual coding. All textual transcripts were imported into Excel 2019 (Microsoft Corp.). Thematic analysis was employed to generate initial codes from the qualitative data [26]. Three of the authors (YL, XL, and YW) read all of the transcripts and analyzed the data independently.

Initial manual codes were developed respectively, and data segments were assigned to these codes. The codes were clustered into broader themes, which were then merged, split, or discarded following a review of the transcripts after discussion among the above authors. The themes were renamed to accurately reflect their content, resulting in the coding framework (Multimedia Appendix 3). Two of the authors (YL and XL) conducted the coding independently. After completing the coding, both authors compared the coding selections, and when discrepancies occurred, consensus was reached through discussion with a third author (YW) until agreement was reached. The number of recurring themes was recorded and documented. A comparative analysis was performed to evaluate discrepancies between on-the-ground practices and national policy requirements, with gaps and their root causes systematically identified. Reasons for the termination of telemedicine services were further analyzed to determine contributing factors. Finally, the themes were translated into English for reporting purposes.

Ethical Considerations

This study was reviewed and approved by the Medical Science Research Ethics Committee of Peking University Third Hospital (No. IRB00006761-M2024300). The Beijing Municipal Health Commission agreed to use the data from the Supervision Platform for secondary analysis. The interviewees provided informed consent. We filed the survey data while maintaining full confidentiality. The quantitative data were anonymized upon the conclusion of the study via a 2-step process: (1) institutions were categorized and (2) all identifying names were removed. Similarly, the qualitative data were stripped of all identifying details during the transcription process to ensure anonymity. No compensation was provided for participation.

Results

Characteristics of Investigated Institutions

In 2023, 187 institutions provided telemedicine services throughout the year, totaling 1,082,591 visits in Beijing. Among them, the institution with the highest service volume had 164,868 visits, while the institution with the lowest service volume had only 1 visit. The average service volume was 5789.3 visits, with an SD of 21,999.7 visits. The median service volume was 49, the 25th percentile was 13, the 75th percentile was 444, and the IQR was 431.

This research focused on analyzing 36 institutions that distributed questionnaires. The service volume of the 36 institutions can be found in Table 1. The total telemedicine service volume of the 36 institutions included in this study was 968,786, accounting for 89.5% of the total service volume in Beijing. As a result, the services provided by these 36 institutions could reflect the status of telemedicine services in Beijing. The total amount of services in 2023 of the 36 medical institutions was 968,786 (100%), with the highest annual service volume of 164,868 visits and the lowest of 220. The average service volume was 26,910.7 visits, and the SD was 42,043.8. The median of the service volume was 9709.5, the 25th percentile was 2919, the 75th percentile was 29,649.5, and the IQR was 26,730.5.

XSL•FO
RenderX

Table . Distribution of institutions and service volume, by institution type (n=36).

Type of institution	Institutions, n (%)	Number of services		
		Total, n (%)	Mean (SD)	Median (IQR)
Nonprofit	29 (80.6)	805,917 (83.2)	27,790.2 (41,031.3)	11,163 (31,740)
Tertiary general hospital	14 (38.9)	292,801 (30.2)	20,914.4 (42,644.8)	8414 (9453)
Tertiary specialized hospital	11 (30.6)	386,094 (39.9)	35,099.5 (43,317.9)	13,866 (41,308)
Tertiary TCM[a] hospital	4 (11.1)	127,022 (13.1)	31,755.5 (33,771.8)	27,744.5 (61,829.5)
Private, for-profit	7 (19.4)	162,869 (16.8)	23,267 (49,351.7)	4125 (12,108)

[a]TCM: traditional Chinese medicine.

Among the 36 institutions investigated in this study, 29 were nonprofit hospitals, all of which were tertiary hospitals. They included 14 general hospitals, 11 specialized hospitals, and 4 TCM hospitals. The total telemedicine service volume of the 29 nonprofit hospitals in 2023 was 805,917 (83.2%), with an average service volume of 27,790.2 per hospital. The median service volume of 29 nonprofit hospitals was 11,163, and the IQR was 31,740.

The other 7 institutions were private, for-profit medical institutions, including 1 secondary general hospital, 2 primary TCM hospitals, 2 ungraded general hospitals, 1 first-class TCM hospital, and 1 clinic. The total telemedicine service volume of the 7 private, for-profit medical institutions in 2023 was 162,869 (16.8%), with an average service volume of 23,267 per institution. The median service volume of the 7 private, for-profit institutions was 4125, and the IQR was 12,108.

Characteristics of Interviewees

A total of 36 contacts from 36 institutions specializing in telemedicine services were interviewed for this study (Table 2). These interviewees had experience in outpatient management, engaged in hospital administration, or were responsible for technology-related work. After the implementation of telemedicine services in these institutions, they were assigned operational roles to coordinate and maintain the services. Among the interviewees, 20 were leaders of telemedicine services, and 16 were participants familiar with the work. The average age of interviewees was 40 years.

Table . A summary of interviewees' demographic information (n=36).

Characteristics	Interviewees, n (%)
Sex	
Male	15 (41.7)
Female	21 (58.3)
Age (y)	
30 - 39	19 (52.8)
40 - 49	10 (27.8)
50 - 59	7 (19.4)
Role in the department	
Leader	20 (55.6)
Participant	16 (44.4)
Departmental affiliation of the interviewees	
Affiliated to the outpatient department	16 (44.4)
Independent, office of telemedicine	10 (27.8)
Independent, operation planning department	7 (19.4)
Affiliated to medical administration office	2 (5.6)
Affiliated to the information technology department	1 (2.8)

Patient Eligibility Assessment Process for Telemedicine Services

After patients registered and updated their information on the telemedicine application or WeChat mini-program of each hospital, they could apply for telemedicine services. The institution providing telemedicine services conducted the assessment, which was divided into 2 stages. The first stage took place before treatment. In this stage, the assessment was made automatically by the system or manually by the physician.

XSL•FO
RenderX

The second stage occurred during treatment. The attending physician determined whether the patient's condition was suitable for telemedicine services. If institutions did not use the system for assessment in the first stage, the 2 stages were combined, and the assessment was left to the physicians. In each stage, specific reasons could prevent patients from receiving telemedicine services. If physicians determined that a patient was not suitable for telemedicine services, they could terminate the session and refer the patient to in-person care. The medical service fee for this session would be refunded.

The assessment approaches of the 36 investigated institutions are presented in Table 3. In the first stage, 72.2% (26/36) of the institutions made assessments automatically by the system, including 25 of the 29 tertiary nonprofit hospitals and 1 of the 7 private, for-profit medical institutions. Among the tertiary nonprofit hospitals, only 4 hospitals selected qualified patients for telemedicine services manually by the physicians, including 1 general hospital, 2 specialized hospitals, and 1 TCM hospital. These hospitals provided a total of 19,465, accounting for 2.42% of the service volume of nonprofit hospitals. Among the 7 private, for-profit medical institutions, only 1 first-class TCM hospital assessed patients automatically by the system. The other 6 private, for-profit institutions assessed patients manually by the physicians for follow-up appointments through telemedicine.

Table . Distribution of assessment approach by hospital type (n=36).

Type of hospital	Auto[a], n (%)	Manual[b], n (%)
Nonprofit	25 (69.4)	4 (11.1)
Tertiary general hospital	13 (36.1)	1 (2.8)
Tertiary specialized hospital	9 (25.0)	2 (5.6)
Tertiary TCM[c] hospital	3 (8.3)	1 (2.8)
Private, for-profit	1 (2.8)	6 (16.7)

[a]Automatically by the system.

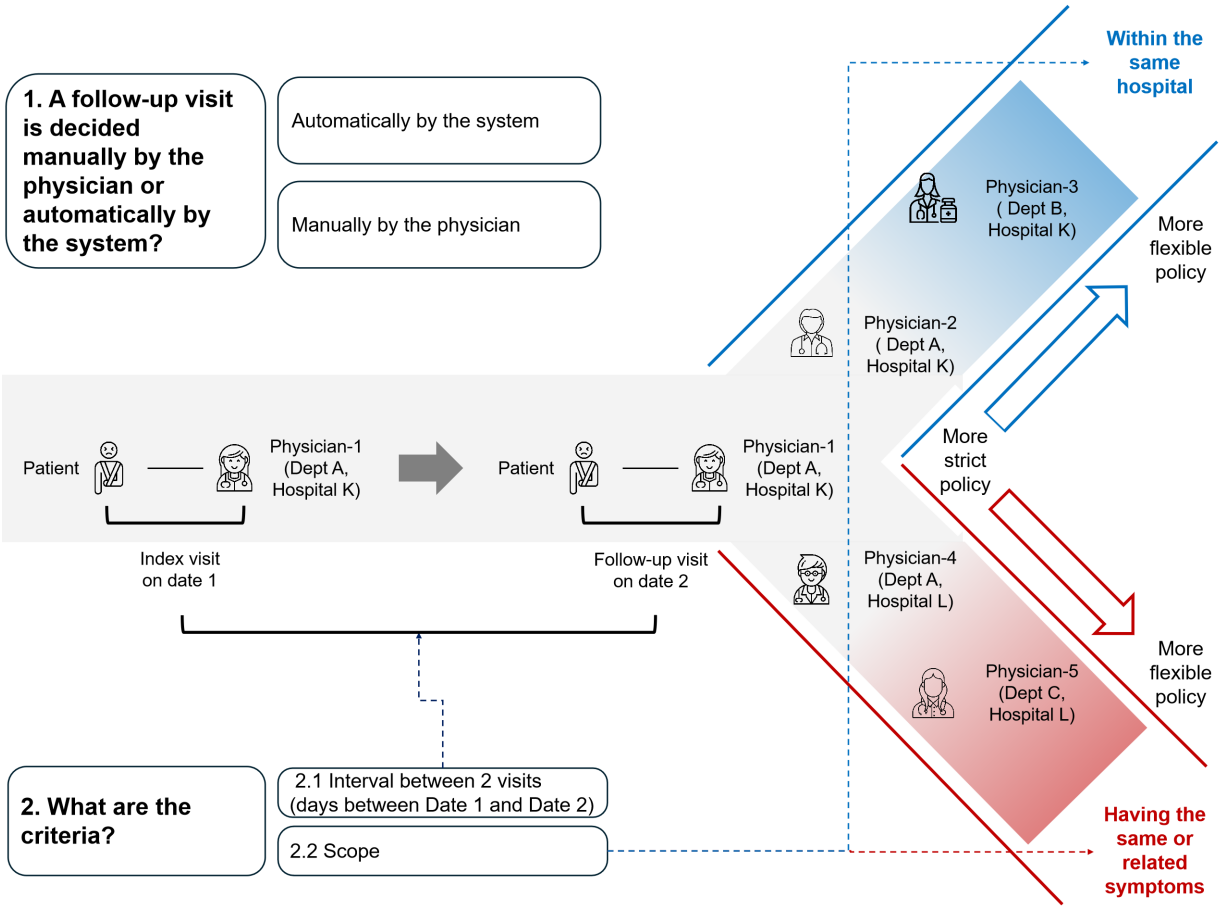
[b]Manually by the physician.

[c]TCM: traditional Chinese medicine.

Regardless of whether the assessment was conducted automatically by the system or manually by the physician, 36 medical institutions investigated in this study assessed patient eligibility from the following 3 dimensions in the first stage. Figure 2 illustrates the patient eligibility assessment process and the 3 dimensions: scope of previous visit location, diagnosis of previous visits, and visit interval. First, whether the previous visit location fell within the scope set by the current institution. Second, whether the previous diagnosis was within the scope defined by the current institution. Third, whether the visit interval between the patient's previous and current visits was within the scope set by the current institution. In this study, we found that medical institutions providing telemedicine services did not necessarily set the requirements for the above 3 dimensions. Considering the ownership type of the institutions, we would discuss the 29 nonprofit tertiary hospitals and the 7 private, for-profit institutions separately.

XSL•FO
RenderX

Figure 2. Patient eligibility assessment process for telemedicine services in the first stage.



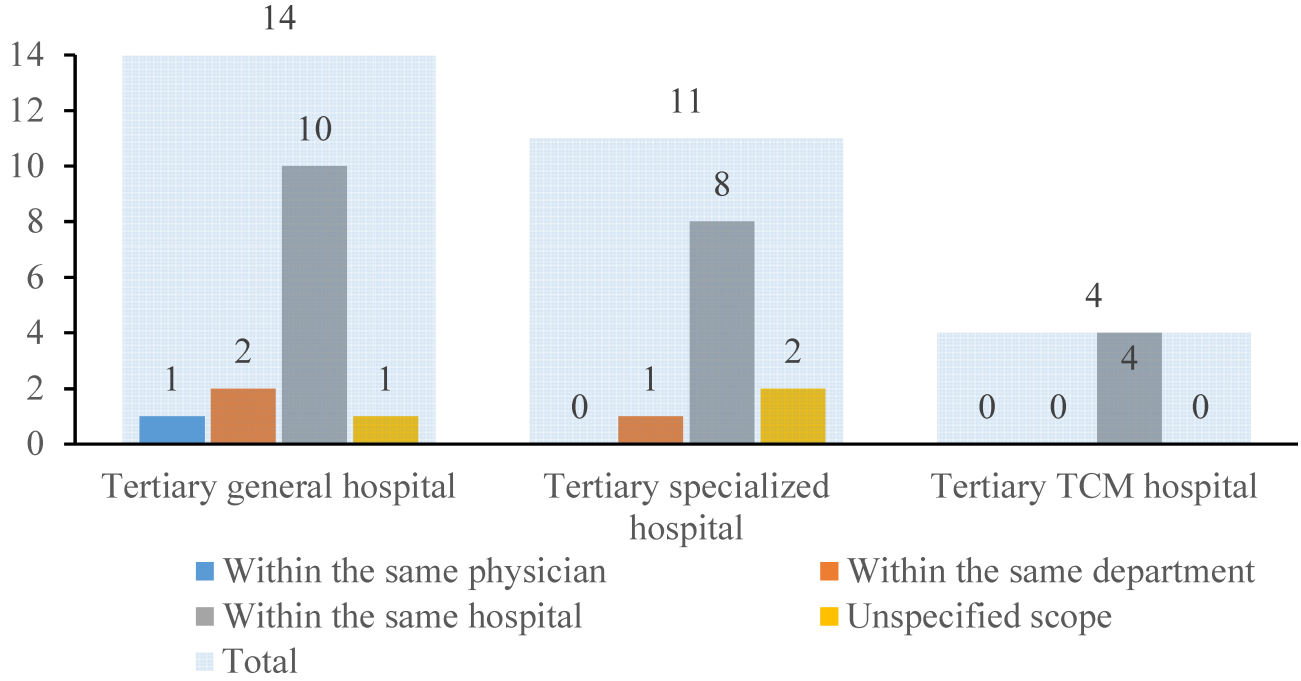
Prerequisites for Patient Assessment in Nonprofit Tertiary Hospitals in Beijing

Prerequisites for the Scope of Previous Visit Location

The prerequisites for the scope of previous visit location were categorized into 4 groups: within the same physician, within the same department, within the same hospital, and unspecified scope. Figure 3 shows that 75.9% (22/29) of the nonprofit tertiary hospitals restricted prior visits to the same institution. There was 1 general hospital that required a record of previous visits to the same physician within the hospital. There were 2 general hospitals and 1 specialized hospital that required a record of previous visits to the same department within the hospital. All the TCM hospitals only limited the scope to the same hospital. Additionally, 3 hospitals did not impose any requirements regarding the previous visit location.

XSL•FO
RenderX

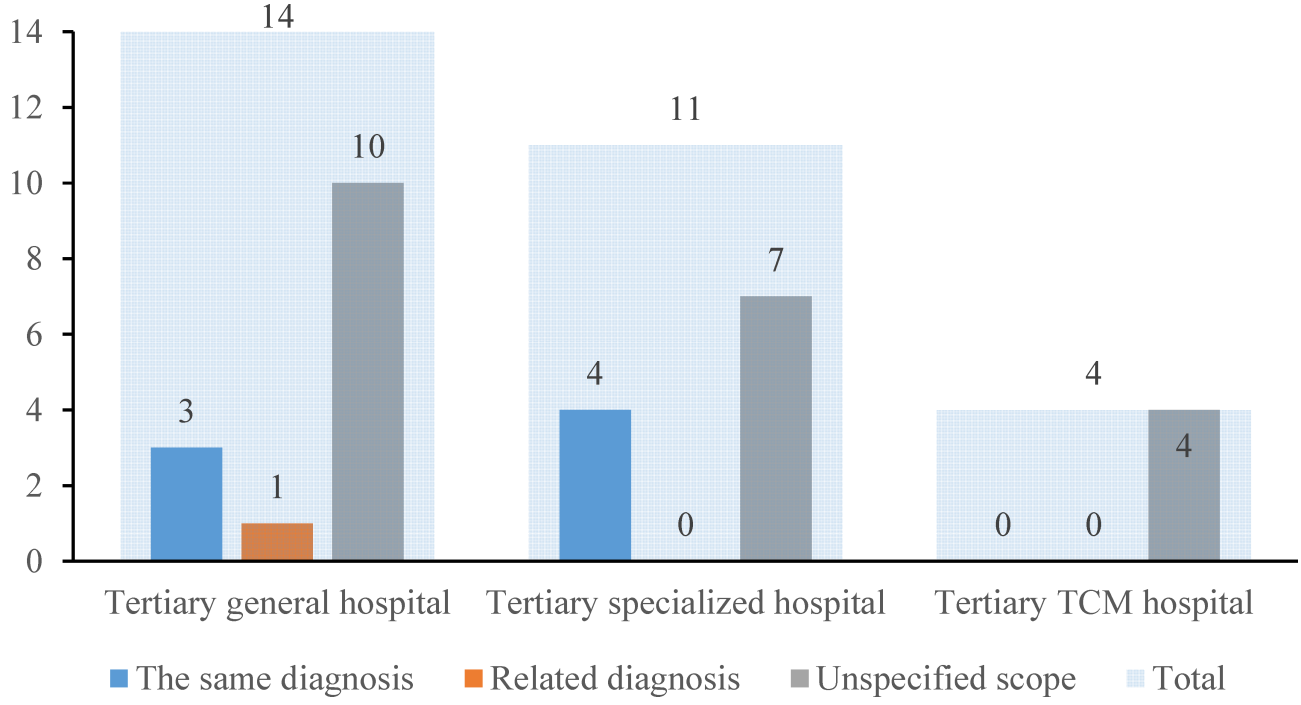
Figure 3. Institution distribution on prerequisites for the scope of previous visit location (n=29). The light blue shade indicates the sum in each type of hospital. The blue, orange, gray, and yellow columns indicate the number of hospitals that required the scope of the previous location to be within the same physician, the same department, the same hospital, and unspecified scope, respectively. TCM: traditional Chinese medicine.



Prerequisites for Diagnosis of Previous Visits

The prerequisites for diagnosis of previous visits were divided into 3 categories: the same diagnosis, related diagnosis, and unspecified scope. Figure 4 shows that 72.4% (21/29) of the nonprofit hospitals imposed no restrictions on the diagnosis of previous visits. There were 3 general hospitals and 4 specialized hospitals that required the diagnosis of the previous visit to match the current diagnosis. One general hospital required the diagnosis of the previous visit to be related to the current diagnosis. TCM hospitals had no restrictions regarding previous diagnosis.

Figure 4. Institution distribution on prerequisites for diagnosis of previous visits (n=29). The light blue shade indicates the sum in each type of hospital. The blue, orange, and gray columns indicate the number of hospitals that required the patients to have the same diagnosis, related diagnosis, and unspecified scope, respectively. TCM: traditional Chinese medicine.

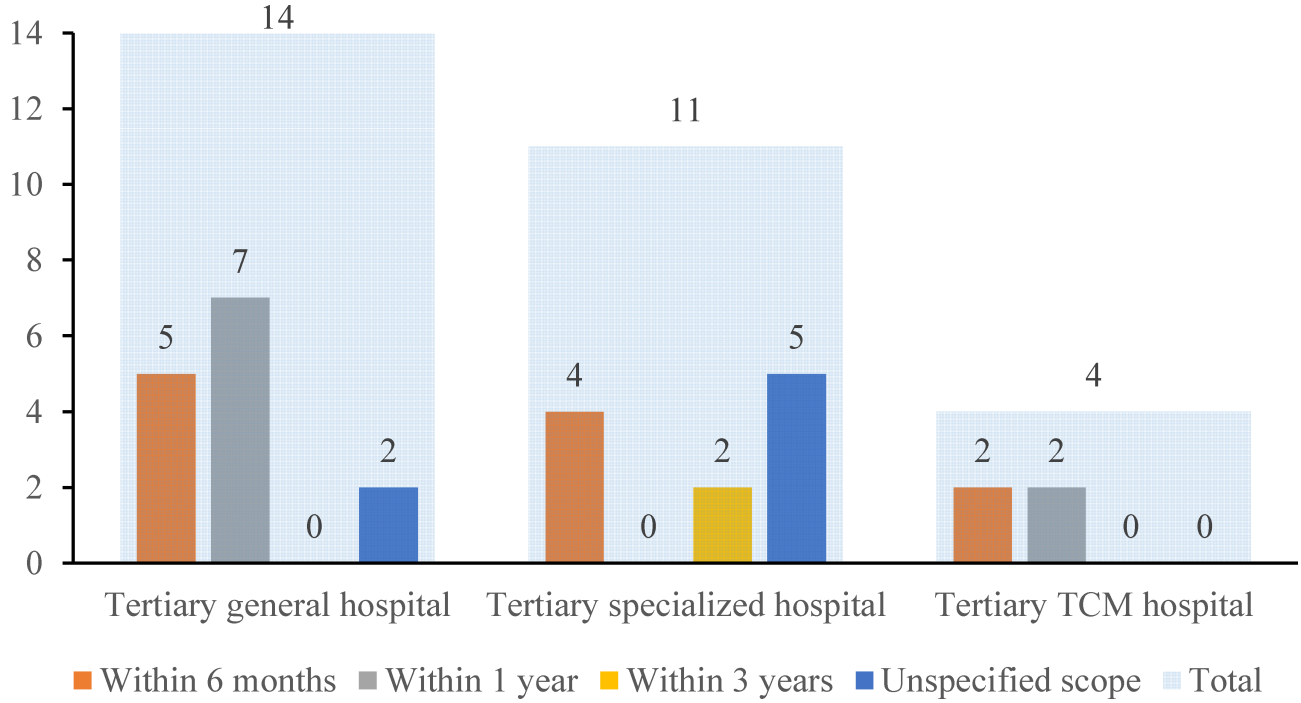


XSL•FO
RenderX

Prerequisites for Visit Interval

The prerequisites for visit interval were divided into 4 categories: within 6 months, within 1 year, within 3 years, and unspecified scope. Figure 5 shows that 37.9% (11/29) of the hospitals restricted the interval to be within 6 months, 31.0% (9/29) of the hospitals restricted it to be within 1 year, and 6.9% (2/29) of the hospitals restricted it to be within 3 years. The 2 specialized hospitals that restricted the visit interval to be within 3 years were both pediatric hospitals. Among the 29 hospitals, 2 general hospitals and 5 specialized hospitals did not have any restrictions regarding the visit interval.

Figure 5. Institution distribution on prerequisites for visit interval (n=29). The light blue shade indicates the sum in each type of hospital. The orange, gray, yellow, and blue columns indicate the number of hospitals that required the visit interval to be within 6 months, within 1 year, within 3 years, and unspecified scope, respectively. TCM: traditional Chinese medicine.



Prerequisites for Patient Assessment in Private, For-Profit Medical Institutions in Beijing

Most of the other 7 private, for-profit institutions did not have any restrictions on the above 3 dimensions. At least 1 secondary general hospital restricted the previous diagnosis to be the same as the current one. This hospital, along with a first-class TCM hospital, required the visit interval to be within 6 months. The other institutions did not have any requirements on the above 3 dimensions to assess whether patients could get telemedicine services. Overall, there were no significant disparities in prerequisites for patient assessment across different types of private, for-profit medical institutions.

Considerations for the Patient Eligibility Assessment

As mentioned by the National Health Commission, patients were required to provide relevant proof of diagnosis from past visits, with no restrictions on the visit interval or the scope of the previous visit location (ie, institution or department) [14]. Physicians providing telemedicine services were responsible for assessing patients. However, in practice, most institutions set requirements in the system for the location of previous visits and the visit interval while having fewer requirements on the diagnosis in the first assessment stage. The reason for imposing such requirements was that obtaining medical records from their own institution was more convenient. The visit intervals were set according to the typical disease progression timelines of the main diseases treated in each department. When telemedicine services restricted the scope of the previous visit location and the visit interval, it became easier to determine eligibility by referencing the patient's diagnosis from a specific institution at a particular point in time.

As shown in Figure 3, 26 institutions required a record of previous visits within their own institution. As they declined to acknowledge or accept previous visit records originating from external institutions, we obtained the following reasons (Table 4) from these institutions through interviews. The primary reason (19 interviewees) was distrust of the medical results provided by other institutions. The second most frequently mentioned reason (18 interviewees) was that it was more challenging to handle medical disputes involving other institutions once they happened. Among these 19 and 18 institutions, only 1 specified that the previous visit must be with the same physician, and 2 required that the previous visit must be within the same department. The third most frequently mentioned reason (10 interviewees) was the system's limitation to access data from other institutions. Notably, these 10 services did not impose additional restrictions regarding whether the previous visit was with the same physician or within the same department. Acceptance was granted as long as the patient had any history of visits within their own institution.

XSL•FO
RenderX

Table . Distribution of reasons for service restrictions in 26 institutions.

Reasons	Institutions, n (%)	Institutions with prerequisites for the scope of previous visit location, n		
		Within the same institution	Within the same department	Within the same physician
Distrust in the medical results provided by other institutions	19 (73.1)	16	2	1
Difficulties in handling medical disputes involving other institutions once medical disputes happen	18 (69.2)	15	2	1
Limited access to data in other institutions through information technology system	10 (38.5)	10	0	0

After the system automatically administered the first assessment stage, physicians would still terminate the telemedicine services in the second stage. Through interviews with the contacts of 36 hospitals, the main reasons for the termination included the patient's condition (n=22, 61.1%), mismatch due to the rules preestablished in the telemedicine system (n=15, 41.7%), and inability to provide the services that the patient hopes to receive (n=13, 36.1%) (Table 5). The proportion of service termination by physicians could be controlled within 2% of the total amount of telemedicine services.

Table . Institution distribution on reasons for the termination of telemedicine services (n=36).

Reasons	Institutions, n (%)
The patient's condition is not suitable for telemedicine services (eg, abdominal pain, trauma, the blood pressure of the hypertensive patient has become so high that physicians cannot refill prescriptions for the patient)	22 (61.1)
Mismatch due to the rules preestablished in the telemedicine system (eg, optometry physicians categorized as ophthalmologists in the system cannot provide services to cataract patients)	15 (41.7)
Inability to provide the services the patient hopes to receive (such as prescribing medications, conducting special examinations)	13 (36.1)
Communication problems due to network issues	7 (19.4)
Not meeting the visit interval required by medical insurance	7 (19.4)
Time conflicts with physicians	5 (13.9)
Insurance fraud committed by patients	2 (5.6)
Unknown	2 (5.6)

Discussion

Principal Findings

This study investigated the current practice of patient eligibility assessment for telemedicine services in Beijing, elucidated institutions' rationale for adopting such approaches, analyzed their differences from national policies, and proposed policy recommendations. First, most nonprofit, tertiary hospitals automatically assessed patient eligibility by the system in the first stage based on the scope of previous visit location, diagnosis of previous visits, and visit interval. Second, compared with national policies, nonprofit hospitals had stricter requirements on the scope of previous visit locations and more lenient requirements on the diagnosis of previous visits. On the other hand, private, for-profit institutions exhibited more relaxed requirements in all dimensions. Third, 2 main reasons for the discrepancies between policy and practice were distrust in the medical results provided by other institutions and difficulties in handling medical disputes involving other institutions once medical disputes happened. Based on the discrepancies, our findings recommend enhancing policy clarity, relaxing telemedicine regulation for new patients, and strengthening telemedicine supervision to improve telemedicine services.

Differences in Prerequisites for Patient Assessment

The quantitative data revealed differences across prerequisites for patient eligibility assessment among various types of medical institutions. Nonprofit hospitals had stricter requirements than private, for-profit institutions on patient eligibility assessment in the above 3 prerequisites. TCM hospitals, which had the highest average volume of telemedicine services, neither restricted prerequisites for the scope of previous visit location to specific departments or physicians nor imposed specific restrictions on patients' diagnoses of previous visits. The main reasons for the flexibility of these hospitals were the holistic view [27] and the personalized treatment approach in TCM, which relies on each physician's own professional judgment

XSL•FO
RenderX

[28]. The general hospital with the highest service volume restricted the current diagnosis to be related to the diagnosis of previous visits, considering the potential variations in the scope of diagnosis and treatment for certain departments across different hospitals. The prerequisites for the visit interval were primarily determined by the disease characteristics among the majority of the patients. We have observed that only 2 hospitals, both pediatric, adopted a 3-year visit interval. As this may be related to the rapid growth of pediatric outpatient visits and relatively long waiting times [29,30], we hypothesize these hospitals seek to transition suitable cases to telemedicine services when appropriate. Most of the other private, for-profit medical institutions did not impose restrictions on the above 3 dimensions.

The qualitative data explained the discrepancies between policy and practice. As for the results of quantitative analysis, the prerequisites for the scope of previous visit location and diagnosis of previous visits showed the gap from policy requirements, while the prerequisites for visit interval were not mentioned in the policies. Our qualitative analysis demonstrated the reasons for barriers to the acceptance of patients based only on external institution visit records. Since determining whether a diagnosis from previous visits aligned with the requirements of the current institution usually involves professional judgment, this assessment cannot be made automatically by the system. Findings from the interviews revealed that the main reasons for the termination of telemedicine service included the patient's condition not being suitable for telemedicine services and mismatches preventing physicians from providing treatment to patients with specific diagnoses.

Reasons for the Discrepancies Between Policy and Practice in Nonprofit Hospitals

The results specified that 2 main reasons had caused the discrepancy in the scope of the previous visit location between policy and practice. The first reason was distrust in the medical results provided by other institutions, and the second reason was difficulties in handling medical disputes involving other institutions once they happened. In practice, service quality and safety were the primary factors that nonprofit hospitals considered. Telemedicine services were only a supplement for traditional brick-and-mortar medical institutions [31]. Therefore, nonprofit hospitals were more cautious in assessments by the system, limiting the scope of previous visit locations to their hospital.

Most nonprofit hospitals did not have restrictions on the diagnosis of previous visits in the first assessment stage. Due to the ambiguity of national policies regarding the requirements of a patient's diagnosis for telemedicine services, it was difficult to quickly assess patient eligibility (except for a few hospitals that limit the scope of previous visit locations to the same department). The physicians always assessed patient eligibility related to the diagnosis of previous visits in the second stage and would terminate services if the patient's condition was not suitable for telemedicine.

Another frequently mentioned reason for differences was barriers to accessing patients' medical information in other institutions. The efficiency of assessment automatically performed by the system was higher than that manually performed by the physician. The system automatically limited the scope of previous visit locations to the same hospital to accommodate the large volume of telemedicine services in Beijing. However, national policies allowed patients to access telemedicine services after their first visit to a brick-and-mortar institution, regardless of the scope of the previous visit location. To be consistent with the policy's requirements, it is necessary to achieve interconnection among health systems in the region. Health information systems, including interoperability with electronic medical records, are a challenge in developing telemedicine services in other countries [32]. Systems that help with initial assessments would quickly capture patients' records at other institutions, reducing physicians' time and effort on evaluation. Beijing had already made a preliminary attempt in this regard. Physicians in 140 secondary and tertiary medical institutions had access to 181 test results and 300 imaging test results in these institutions, and the results were mutually recognized [33]. Access to test results laid the foundation for access to health information records in the future.

Reasons for the Discrepancies Between Policy and Practice in Private, For-Profit Institutions

Private, for-profit medical institutions were more lenient than nonprofit hospitals in assessing patient eligibility. Most of the private, for-profit institutions did not set requirements apart from the policy requirements. However, there was a possibility of relaxing the policy requirements in practical work. Most private, for-profit medical institutions initiated by enterprises took telemedicine services as their primary development direction. When enterprises were the main initiators of medical institutions, they focused on increasing operating income and achieving profitability [18]. Patient eligibility for telemedicine services was always assessed manually by physicians in these institutions, and the actual assessment criteria may have been more permissive than policy requirements. In addition, the private, for-profit medical institutions mainly provided psychological and traditional Chinese medical services, and the telemedicine services were also concentrated on psychological counseling and prescriptions. Physicians focused more on the patients' needs of current telemedicine visits rather than the prerequisites for the scope and diagnosis of previous visits.

International Telemedicine Policies

Telemedicine was once limited only to rural or remote communities but is now increasingly used to expand the geographic reach of health care services and improve access to care [34]. According to the 2023 report on the state of digital health in the World Health Organization (WHO), 78.4% (40/51) of the WHO Europe Member States directly addressed telehealth in their policies or strategies [35]. While the use of telemedicine was generally allowed in most countries, the eligibility requirements for telemedicine service users had changed due to the in-person visits sharply declining during the COVID-19 pandemic. Before March 2020, 36.7% (11/30) of the Organization for Economic Co-operation and Development countries allowed telemedicine only for patients who had an in-person visit in the past, but after March 2020, only 20.0% (6/30) of the countries imposed the requirement [36]. Unlike

XSL•FO
RenderX

most Organization for Economic Co-operation and Development countries, in China, an initial face-to-face meeting between the physician and the patient is still needed before telemedicine services, except for patients with COVID-related symptoms.

The establishment of patient eligibility requirements was primarily intended to ensure the safety of patients engaging in telemedicine services. The comprehensive framework designed by the WHO to enhance and streamline telemedicine services included a clinical risk assessment process. The process specified which services were safe for the patients to use independently and which needed expert medical support [37]. We observed that countries employed 2 methods to ensure safety: establishing telemedicine service lists or relying on physicians' professional judgment. The United States Drug Enforcement Administration allowed patients to obtain prescribed medications through telemedicine appointments, without requiring an initial in-person medical evaluation, except for prescribing controlled medications to new patients [38]. In Japan, it was the physicians' sole responsibility to determine whether telemedicine service was appropriate and safe [39]. Hungary not only restricted the types of telemedicine services but also mandated that physicians make the final diagnosis [39]. In China, the policy has not yet explicitly outlined a list of specific service types for telemedicine. At present, physicians play a key role in determining the services offered to patients and ensuring their safety. Experiences in other countries shed light on how China could refine its telemedicine policies.

Providing telemedicine services to new patients is expected to become a major trend in future health care, with ensuring patient safety as an essential component. A more profound understanding of what visits are appropriate for telemedicine and better triaging of visit types are crucial steps to maintain the quality of care [40]. However, the legal responsibilities of physical medical institutions and physicians regarding issues such as medical risk have not been fully clarified [41]. Supported by the telemedicine service list, the challenges and work-related stress for physicians can be alleviated.

Suggestions for Future Improvement

According to the above analyses, telemedicine policies could be optimized in the following aspects: first, the ambiguous parts of the policies could be clarified. Second, the content of telemedicine services for new patients could be added to the policies. Third, policies for telemedicine services would be made for nonprofit and private, for-profit institutions, respectively.

Explicit policies would help providers implement the requirements stipulated in the policies and make determinations about the applicability of telemedicine services more easily [42]. Although several policy interventions had been implemented at the national level, there was still a lack of specific service standards to direct frontline service delivery [31]. Specific disease names or diagnoses were not defined. A guideline could be made, including a comprehensive list of symptoms or diagnoses provided through telemedicine. Guidelines could also be considered when the most recent in-person visit occurs [42]. For example, the American Medical Association made the Telehealth Services Covered by Medicare and Included in the Current Procedural Terminology Code Set [43]. The United States Drug Enforcement Administration stated that a patient could receive a 6-month supply of buprenorphine through telehealth services; further prescriptions would require an in-person visit [38].

There were already some mature precedents for policymakers to refer to, allowing telemedicine health services for new patients. In the United States, telemedicine services were allowed reimbursement and expanded coverage for telehealth services to new and established patients from March 2020 [44,45]. An in-person visit within 6 months of an initial mental telehealth service is not required through September 30, 2025 [46]. After March 2020, 7 countries relaxed the prerequisite that patients were only allowed to have teleconsultations with physicians they had already consulted in person [36]. In this study, some institutions were found to provide telemedicine services for new patients. The service volume of 1 private, for-profit institution in 2023 was 134,715, ranked third among 36 institutions in this study. In the supervision process, it was found that this institution focused on psychological services and provided telemedicine services for new patients. However, no significant incidents had affected medical quality and patient safety. In the United States, 80% of the mental health treatment facilities that accepted new clients could offer telehealth services [47]. Telemental health use also increased substantially and globally during the pandemic and will remain in use for the foreseeable future [48,49].

In China, from the demand-side perspective, the conditions for opening up telemedicine services for new patients have been met under certain circumstances. Telemedicine services provide access to high-quality medical care, free up resources in overcrowded public hospitals [50], and also help to alleviate the uneven distribution of medical resources [51]. Telemedicine services before and during the COVID-19 pandemic had enhanced the delivery of and access to health care services [52] and increased satisfaction compared with usual in-person care [53]. They allowed unprecedented accessibility for those with mobility or geographical restrictions [54]. Considering the advantages of telemedicine services, there was also a need for new patients to access them.

However, to deal with the potential risks of providing services for new patients, the following measures could be considered. First, standardize the scope of telemedicine services (eg, ordering examinations or tests) for new patients to ensure medical safety [39]. Experts from various specialties would be engaged to develop a telemedicine service list for new patients, detailing diagnostic criteria, disease stages, and specific service items, as well as prohibiting telemedicine in unsuitable cases. Furthermore, developing an interoperable electronic medical record system that enables doctors to access patients' previous medical records and test results would enable physicians to understand patients' conditions more accurately. Second, select the diseases for which the diagnostic and treatment schemes are relatively mature with a high frequency of visits (eg, diseases of dermatology and psychology [47,55]) to meet the patients' needs for privacy protection and reduce the round-trip travel time [50]. Hospitals with high-quality telemedicine services could pilot telemedicine for new patients, starting with 2 - 3

XSL•FO
RenderX

medical specialties that carry low clinical risk. Third, improve the requirements for the professional experience and professional title of physicians who can provide telemedicine services for new patients. Finally, limit the scope of services covered by medical insurance funds. If the patient only needs a medical examination and related consultation, the patient would pay the relevant fees at his or her own expense.

The supervision policies of the services need to be formulated separately due to different initiators of institutions to ensure the quality and safety of services. The primary purpose of telemedicine services initiated by nonprofit hospitals is to expand the scope and intensity of hospital services, while the purpose of telemedicine services initiated by private, for-profit institutions is to increase the number of patients and gain benefits [18]. For nonprofit medical institutions with strict management, policies can appropriately relax the requirements of patient eligibility for telemedicine services and even prioritize providing telemedicine services for new patients. For private, for-profit institutions, policies need to restrict the types of telemedicine services provided by private, for-profit institutions and raise the requirements for the qualifications of physicians in telemedicine services. In addition, since most private, for-profit institutions manually assess patient eligibility by the physician, additional regulation is required (eg, making data of the service process traceable) to ensure the safety and quality of telemedicine services [56]. During regulation, the Supervision Platform can proactively monitor adverse events and quality deficiencies and automatically submit the above results to both medical institutions and the Center on a daily basis. The Center can also organize third-party experts to conduct annual inspections, and institutions that fail to meet the standards will be prohibited from providing telemedicine services.

Limitations

The 36 institutions selected for this study included those with high service quality and large service volumes, but those with lower service volume or quality were excluded. This is because, during routine inspections conducted by the Center, we found that these institutions often lacked stable management systems and assessment processes, making it challenging to reasonably compare their assessment processes with national policy. Therefore, these institutions were excluded from this study, making it impossible to characterize their follow-up criteria.

Additionally, there would be potential social desirability bias in interviews [57]. In an effort to minimize this potential bias, before conducting the interviews, we informed participants about the anonymization of data and ensured that their specific responses would not be disclosed; only the statistically analyzed group information would be made public. We also cross-validated interview results against previous inspection findings to ensure the authenticity of the results.

Finally, the study may have missed considerations and insights from policy formulation and regulatory constraints. Incorporating government perspectives could have enhanced the comprehensiveness of the study. We plan to submit our research results to the relevant government departments to inform policymakers about policy implementation and help elucidate the underlying reasons for the discrepancies between telemedicine policy and actual practice.

Conclusions

This pioneering multicenter, mixed methods study innovatively analyzed patient eligibility for telemedicine services in China, addressing the current gap in telemedicine policy research. The findings faithfully reflect the practices of the patient eligibility assessment process for telemedicine services in Beijing. Furthermore, by integrating quantitative and qualitative analyses, the study uncovered the underlying reasons behind discrepancies between institutional practices and national policy requirements, providing valuable insights about current implementation challenges. These results suggest potential improvements, including relaxing the regulation for new patients' access to telemedicine services, enhancing service delivery processes, and strengthening the robustness of the telemedicine supervisory framework.

Acknowledgments

The authors thank the anonymous reviewers for their constructive comments, which helped improve the manuscript.

Funding

This study was funded by the National Natural Science Foundation of China (grant ID 72204005), awarded for the period from January 2023 to December 2025, with a total of RMB 300,000 (US $42,023). The study was also funded by the Capital's Funds for Health Improvement and Research (grant ID CFH2024-2-40915), awarded for the period from January 2024 to December 2026, with a total funding of RMB 416,000 (US $58,272).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: YW (lead), YL (supporting), XL (supporting)
Data curation: YL
Formal analysis: YL (lead), XL (equal), YW (supporting)
Funding acquisition: YW

XSL•FO
RenderX

Investigation: YL (lead), XL (supporting), HL (supporting), MG (supporting)
Methodology: YW (lead), XL (supporting), YL (supporting)
Project administration: CJ (lead), XX (supporting)
Resources: XX (lead), YL (supporting)
Supervision: YW
Validation: YW
Visualization: YW (lead), XL (supporting)
Writing – original draft: XL (lead), YL (supporting)
Writing – review & editing: YW (lead), XL (supporting), YL (supporting)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Regulatory compliance assessment of telemedicine services: 34-item compliance monitoring metrics from the Beijing Telemedicine Supervision Platform.
[DOC File, 56 KB - humanfactors_v12i1e75964_app1.doc]

Multimedia Appendix 2
Survey questionnaire for patient eligibility assessment in telemedicine services.
[DOC File, 40 KB - humanfactors_v12i1e75964_app2.doc]

Multimedia Appendix 3
Interview outline and representative interview excerpts for institutional considerations of the patient eligibility assessment.
[DOC File, 45 KB - humanfactors_v12i1e75964_app3.doc]

Checklist 1
Checklist for reporting mixed methods research (MMR) studies.
[DOC File, 45 KB - humanfactors_v12i1e75964_app4.doc]

References

1. Webster P. Virtual health care in the era of COVID-19. Lancet 2020 Apr 11;395(10231):1180-1181. [doi: 10.1016/S0140-6736(20)30818-7] [Medline: 32278374]
2. Dixit RA, Ratwani RM, Bishop JA, et al. The impact of expanded telehealth availability on primary care utilization. NPJ Digit Med 2022 Sep 9;5(1):141. [doi: 10.1038/s41746-022-00685-8] [Medline: 36085158]
3. Amagai S, Vonesh E, Adams J, Luo Y. Closing the gap: addressing telehealth disparities across specialties in the sustained pandemic era. NPJ Digit Med 2024 Aug 21;7(1):217. [doi: 10.1038/s41746-024-01201-w] [Medline: 39164391]
4. Lucas JW, Villarroel MA. Telemedicine use among adults: United States, 2021. NCHS Data Brief 2022 Oct(445):1-8. [Medline: 36255940]
5. Thomas EE, Haydon HM, Mehrotra A, et al. Building on the momentum: sustaining telehealth beyond COVID-19. J Telemed Telecare 2022 May;28(4):301-308. [doi: 10.1177/1357633X20960638] [Medline: 32985380]
6. Virtual care in Canada: progress and potential. Canadian Medical Association. 2022 Feb. URL: https://www.cma.ca/sites/default/files/2022-02/Virtual-Care-in-Canada-Progress-and-Potential-EN.pdf [accessed 2025-02-16]
7. Andino JJ, Eyrich NW, Boxer RJ. Overview of telehealth in the United States since the COVID-19 public health emergency: a narrative review. mHealth 2023;9:26. [doi: 10.21037/mhealth-23-15] [Medline: 37492124]
8. Hyder MA, Razzak J. Telemedicine in the United States: an introduction for students and residents. J Med Internet Res 2020 Nov 24;22(11):e20839. [doi: 10.2196/20839] [Medline: 33215999]
9. Tu J, Wang C, Wu S. The internet hospital: an emerging innovation in China. Lancet Glob Health 2015 Aug;3(8):e445-e446. [doi: 10.1016/S2214-109X(15)00042-X] [Medline: 26187488]
10. Li P, Zhang C, Gao S, et al. Association between daily internet use and incidence of chronic diseases among older adults: prospective cohort study. J Med Internet Res 2023 Jul 17;25:e46298. [doi: 10.2196/46298] [Medline: 37459155]
11. Wang Y, Zhang P, Xing Y, et al. Telemedicine integrated care versus in-person care mode for patients with short stature: comprehensive comparison of a retrospective cohort study. J Med Internet Res 2024 Nov 19;26:e57814. [doi: 10.2196/57814] [Medline: 39560983]
12. Han T, Wei Q, Wang R, et al. Service quality and patient satisfaction of internet hospitals in China: cross-sectional evaluation with the service quality questionnaire. J Med Internet Res 2024 Nov 8;26:e55140. [doi: 10.2196/55140] [Medline: 39514849]
13. Mu X. China greenlights “Internet hospitals”. Xinhuanet. 2018 Apr 16. URL: https://www.xinhuanet.com/english/2018-04/16/c_137115614.htm [accessed 2018-04-16]

XSL•FO
RenderX

14. Su JJ, Chan MHS, Ghisi GDM, et al. Real-world mobile health implementation and patient safety: multicenter qualitative study. J Med Internet Res 2025 Apr 29;27:e71086. [doi: 10.2196/71086] [Medline: 40299494]
15. Shi J, Tan Z, Ma Q, et al. Implementation, service effectiveness and satisfaction with teleconsultation services in China during the COVID-19 pandemic: from the perspective of primary health care professionals. Int J Med Inform 2023 Oct;178:105202. [doi: 10.1016/j.ijmedinf.2023.105202] [Medline: 37651778]
16. Li L, Du T, Hu Y. The effect of different classification of hospitals on medical expenditure from perspective of classification of hospitals framework: evidence from China. Cost Eff Resour Alloc 2020;18(1):35. [doi: 10.1186/s12962-020-00229-5] [Medline: 32944007]
17. Chi H, Guo M, Li Y, Liu H. Annual Report on Development of China Internet Hospital (2024): Social Sciences Academic Press; 2024.
18. Han Y, Lie RK, Guo R. The internet hospital as a telehealth model in China: systematic search and content analysis. J Med Internet Res 2020 Jul 29;22(7):e17995. [doi: 10.2196/17995] [Medline: 32723721]
19. Huang X, Wang Y, Yang X, Jiang R, Liu Y, Wang H. Patient-centric mobile medical services accessed through smartphones in the top 100 Chinese public hospitals: cross-sectional survey study. JMIR Form Res 2024 Dec 4;8:e45763. [doi: 10.2196/45763] [Medline: 39631758]
20. Xu X, Cai Y, Wu S, et al. Assessment of internet hospitals in China during the COVID-19 pandemic: national cross-sectional data analysis study. J Med Internet Res 2021 Jan 20;23(1):e21825. [doi: 10.2196/21825] [Medline: 33417586]
21. Chen X, Wu X, Zhang Q, et al. The construction and operational models of internet hospitals in China: a hospital-based survey study. BMC Health Serv Res 2023 Jun 21;23(1):669. [doi: 10.1186/s12913-023-09675-2] [Medline: 37344831]
22. Yuan L, Cao J, Wang D, Yu D, Liu G, Qian Z. Regional disparities and influencing factors of high quality medical resources distribution in China. Int J Equity Health 2023 Jan 10;22(1):8. [doi: 10.1186/s12939-023-01825-6] [Medline: 36627636]
23. Zhao W, Xu F, Diao X, et al. The status quo of internet medical services in China: a nationwide hospital survey. Telemed J E Health 2024 Jan;30(1):187-197. [doi: 10.1089/tmj.2023.0025] [Medline: 37437119]
24. Lee SYD, Iott B, Banaszak-Holl J, et al. Application of mixed methods in health services management research: a systematic review. Med Care Res Rev 2022 Jun;79(3):331-344. [doi: 10.1177/10775587211030393] [Medline: 34253078]
25. Fetters MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs-principles and practices. Health Serv Res 2013 Dec;48(6 Pt 2):2134-2156. [doi: 10.1111/1475-6773.12117] [Medline: 24279835]
26. Ahmed SK, Mohammed RA, Nashwan AJ, et al. Using thematic analysis in qualitative research. J Med Surg Public Health 2025 Aug;6:100198. [doi: 10.1016/j.glmedi.2025.100198]
27. Traditional Chinese medicine in China. Xinhuanet. 2016 Dec 6. URL: https://www.xinhuanet.com/english/china/2016-12/06/c_135883982.htm [accessed 2025-12-16]
28. Kwan YH, Chooi S, Yoon S, et al. Professionalism in traditional Chinese medicine (TCM) practitioners: a qualitative study. BMC Complement Med Ther 2020 Nov 9;20(1):335. [doi: 10.1186/s12906-020-03127-8] [Medline: 33167947]
29. Hu K, Sun Z, Rui Y, Mi J, Ren M. Shortage of paediatricians in China. Lancet 2014 Mar 15;383(9921):954. [doi: 10.1016/S0140-6736(14)60482-7] [Medline: 24629297]
30. Zhang XY, Gao Y, Li CP, et al. Observed and projected trends in paediatric health resources and services in China between 2003 and 2030: a time-series study. BMJ Open 2017 Jun 24;7(6):e015000. [doi: 10.1136/bmjopen-2016-015000] [Medline: 28647724]
31. Lai Y, Ho Q, Wang Y, et al. Challenges and strategies of developing internet hospital: combining qualitative interview and documentary research. Digit Health 2024;10:20552076241310075. [doi: 10.1177/20552076241310075] [Medline: 39741982]
32. Curioso WH, Coronel-Chucos LG, Henríquez-Suarez M. Integrating telehealth for strengthening health systems in the context of the COVID-19 pandemic: a perspective from Peru. Int J Environ Res Public Health 2023 May 28;20(11):5980. [doi: 10.3390/ijerph20115980] [Medline: 37297584]
33. Medical institutions will have access to 181 test results and 300 imaging test results mutually by the end of the year. People's Daily Online. 2024 Nov 28. URL: http://bj.people.com.cn/n2/2024/1128/c14540-41057287.html [accessed 2025-08-22]
34. Gajarawala SN, Pelkowski JN. Telehealth benefits and barriers. J Nurse Pract 2021 Feb;17(2):218-221. [doi: 10.1016/j.nurpra.2020.09.013] [Medline: 33106751]
35. The ongoing journey to commitment and transformation: digital health in the WHO European region. World Health Organization. 2023. URL: https://iris.who.int/bitstream/handle/10665/372051/9789289060226-eng.pdf?sequence=2&isAllowed=y [accessed 2025-12-16]
36. Rivas H. The future of telemedicine after COVID-19. Organization for Economic Co-operation and Development. 2023 Jan 20. URL: https://www.oecd.org/content/dam/oecd/en/publications/reports/2023/01/the-future-of-telemedicine-after-covid-19_3b64f3d8/d46e9a02-en.pdf [accessed 2025-12-16]
37. Support tool to strengthen telemedicine: resource for assessment, strategy development, and strengthening of telemedicine services. World Health Organization. 2024. URL: https://www.who.int/europe/publications/i/item/WHO-EURO-2024-9548-49320-73720 [accessed 2025-01-09]
38. DEA announces three new telemedicine rules that continue to open access to telehealth treatment while protecting patients. United States Drug Enforcement Administration. 2025 Jan 16. URL: https://www.dea.gov/press-releases/2025/01/16/dea-announces-three-new-telemedicine-rules-continue-open-access [accessed 2025-01-16]

XSL•FO
RenderX

39. Oliveira Hashiguchi TC. Bringing health care to the patient: an overview of the use of telemedicine in OECD countries. Organization for Economic Co-operation and Development. 2020. URL: https://www.oecd.org/content/dam/oecd/en/publications/reports/2020/01/bringing-health-care-to-the-patient_dfe153ac/8e56ede7-en.pdf [accessed 2025-01-14]
40. Patel M, Berlin H, Rajkumar A, et al. Barriers to telemedicine use: qualitative analysis of provider perspectives during the COVID-19 pandemic. JMIR Hum Factors 2023 Jun 26;10:e39249. [doi: 10.2196/39249] [Medline: 37358887]
41. Zhong Y, Hahne J, Wang X, et al. Telehealth care through internet hospitals in China: qualitative interview study of physicians' views on access, expectations, and communication. J Med Internet Res 2024 Mar 29;26:e47523. [doi: 10.2196/47523] [Medline: 38551618]
42. Lin JJ, Horowitz CR, Ancker JS. An urgent need for guidelines for telemedicine use. J Gen Intern Med 2023 Mar;38(4):1061-1062. [doi: 10.1007/s11606-022-07905-7] [Medline: 36333548]
43. Telehealth services covered by Medicare and included in CPT code set. American Medical Association. 2020. URL: https://www.ama-assn.org/system/files/telehealth-services-covered-by-medicare-and-included-in-cpt-code-set.pdf [accessed 2025-12-16]
44. Golden B. Using punctuated equilibrium theory: policy momentum in traditional Medicare telehealth reimbursement during COVID-19. Policy Polit Nurs Pract 2023 Aug;24(3):198-207. [doi: 10.1177/15271544231157382] [Medline: 37439019]
45. Telehealth policy updates. TelehealthHHS.gov. URL: https://telehealth.hhs.gov/providers/telehealth-policy/telehealth-policy-updates [accessed 2025-11-21]
46. The telehealth waiver: what psychologists need to know. American Psychological Association. URL: https://www.apaservices.org/practice/reimbursement/government/medicare-inperson-telehealth-requirement [accessed 2025-12-08]
47. Cantor J, Schuler MS, Matthews S, Kofner A, Breslau J, McBain RK. Availability of mental telehealth services in the US. JAMA Health Forum 2024 Feb 2;5(2):e235142. [doi: 10.1001/jamahealthforum.2023.5142] [Medline: 38306092]
48. Sugarman DE, Busch AB. Telemental health for clinical assessment and treatment. BMJ 2023 Jan 16;380:e072398. [doi: 10.1136/bmj-2022-072398] [Medline: 36646462]
49. Abraham A, Jithesh A, Doraiswamy S, Al-Khawaga N, Mamtani R, Cheema S. Telemental health use in the COVID-19 pandemic: a scoping review and evidence gap mapping. Front Psychiatry 2021;12:748069. [doi: 10.3389/fpsyt.2021.748069] [Medline: 34819885]
50. Cheng TC, Yip W. Policies, progress, and prospects for internet telemedicine in China. Health Syst Reform 2024 Dec 17;10(2):2389570. [doi: 10.1080/23288604.2024.2389570] [Medline: 39437237]
51. Li D, Zhang R, Chen C, et al. Developing a capsule clinic-a 24-hour institution for improving primary health care accessibility: evidence from China. JMIR Med Inform 2023 Jan 9;11:e41212. [doi: 10.2196/41212] [Medline: 36622737]
52. Kuziemsky C, Hunter I, Udayasankaran JG, et al. Telehealth as a means of enabling health equity. Yearb Med Inform 2022 Aug;31(1):60-66. [doi: 10.1055/s-0042-1742500] [Medline: 35654429]
53. Sculley JA, Musick H, Krishnan JA. Telehealth in chronic obstructive pulmonary disease: before, during, and after the coronavirus disease 2019 pandemic. Curr Opin Pulm Med 2022 Mar 1;28(2):93-98. [doi: 10.1097/MCP.0000000000000851] [Medline: 34860202]
54. Jerjes W, Harding D. Telemedicine in the post-COVID era: balancing accessibility, equity, and sustainability in primary healthcare. Front Digit Health 2024;6:1432871. [doi: 10.3389/fdgth.2024.1432871] [Medline: 39233772]
55. Ranpariya V, Kats D, Lipoff JB. Direct-to-consumer teledermatology growth: a review and outlook for the future. Cutis 2022 Apr;109(4):211-217. [doi: 10.12788/cutis.0503] [Medline: 35659836]
56. Yang F, Shu H, Zhang X. Understanding "internet plus healthcare" in China: policy text analysis. J Med Internet Res 2021 Jul 26;23(7):e23779. [doi: 10.2196/23779] [Medline: 34309581]
57. Bispo Júnior JP. Social desirability bias in qualitative health research. Rev Saude Publica 2022;56:101. [doi: 10.11606/s1518-8787.2022056004164] [Medline: 36515303]

Abbreviations

TCM: traditional Chinese medicine
WHO: World Health Organization

Edited by A Kushniruk; submitted 16.04.25; peer-reviewed by MA Zoubi, Q Wang; accepted 20.11.25; published 19.12.25.

Please cite as:
Lyu Y, Li X, Li H, Gu M, Xu X, Wang Y, Jin C
Bridging Policy and Practice in Telemedicine Follow-Up Identification: Multicenter Mixed Methods Study in Beijing
JMIR Hum Factors 2025;12:e75964
URL: https://humanfactors.jmir.org/2025/1/e75964
doi:10.2196/75964

XSL•FO
RenderX

© Yanli Lyu, Xi Li, Huixian Li, Mingyu Gu, Xuedong Xu, Yipei Wang, Changxiao Jin. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 19.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Dynamic Indicators of Adherence and Retention in Adults Using a Digital Mental Health App: Longitudinal Observational Analysis From the Brighten Study

Dylan Hamitouche[1,2]; Youcef Barkat[2,3], BSc; Deven Parekh[2], MSc; Eva Hammer[2,4]; David Benrimoh[2,5], MSc, MD

[1]Department of Medicine, McGill University, Montreal, QC, Canada

[2]Douglas Mental Health University Institute, 6875 LaSalle Boulevard, Montreal, QC, Canada

[3]Department of Biochemistry, Université de Montréal, Montreal, QC, Canada

[4]Department of Psychology, McGill University, Montreal, QC, Canada

[5]Department of Psychiatry, McGill University, Montreal, QC, Canada

Corresponding Author:
Dylan Hamitouche
Department of Medicine, McGill University, Montreal, QC, Canada

Abstract

Background: Making optimal use of mobile health technologies requires the validation of digital biomarkers, which in turn demands high levels of participant adherence and retention. However, current remote digital health studies have high attrition rates and low participant adherence, which may introduce bias and limit the generalizability of their findings.

Objective: This study aimed to identify longitudinal indicators of participant retention and adherence, which may serve to develop strategies to improve data collection in digital health studies and improve understanding of how study cohorts are shaped by participant withdrawal and nonadherence.

Methods: We performed secondary analyses on the Brighten study, which consisted of 2 remote, smartphone-based randomized controlled trials evaluating mobile apps for depression treatment, enrolling 2193 participants in total. Participants were asked, after baseline assessment, to complete 7 digital questionnaires regularly. We assessed adherence to digital questionnaires, engagement (postbaseline participation), and retention rates (the proportion of participants who continued completing questionnaires over time) as outcomes. We investigated the relationship between these outcomes and both static measures (eg, demographics and average questionnaire scores) and dynamic measures (eg, changes in questionnaire scores over time).

Results: The study included 2201 participants, of whom 1093 completed at least 1 nonbaseline questionnaire, with a median completion rate of 37.6% (IQR 15.5%-67.9%). We found significantly higher adherence rates in participants who were less depressed on average over the course of the study (t_{752}=−5.63; P<.001) and in those who perceived clinical improvement (t_{744}=3.78; P=.001). There were significant demographic differences in adherence and engagement, including differences by gender, race, education, income, and income satisfaction. Participants who were more depressed at baseline were more likely to withdraw before completing any nonbaseline questionnaire (t_{1917}=−2.53; P=.01). However, participants who showed improvement in depressive symptoms during the study showed better adherence (Mann-Whitney U=127,084; P<.001) and retention (hazard ratio 0.78, 95% CI 0.67-0.91; P=.002), despite showing greater depressive symptoms at baseline.

Conclusions: We show that participants' clinical trajectory of depressive symptoms, as well as their perception of improvement, are important indicators of engagement, adherence, and retention. Expanding knowledge regarding these longitudinal indicators may improve interpretation of outcomes and help build strategies to improve retention and adherence in future clinical trials.

(JMIR Hum Factors 2025;12:e69464) doi:10.2196/69464

KEYWORDS

mHealth; digital health; mobile health; digital biomarkers; adherence; depression; smartphone; digital interventions

Introduction

Over the past decade, researchers have dedicated significant effort to identifying clinically useful biomarkers for preventing, diagnosing, and treating psychiatric disorders. Digital biomarkers have garnered considerable attention, as they enable the observation of long-term patterns and trends outside of the hospital or clinic environment, potentially enhancing our understanding of the course of mental illnesses [1,2]. Identifying reliable digital biomarkers has the potential to significantly improve the accuracy of diagnoses, predictions of clinical outcomes (eg, suicidal ideation [3]), and treatment

XSL•FO
RenderX

decision-making regarding psychiatric disorders [4-6]. Furthermore, collecting digital biomarkers is cost-effective [7,8] and allows the possibility to provide patients with feedback through electronic reports, giving them ownership of their own medical data [9-12].

One major challenge for the effectiveness of mobile health technologies is the identification and validation of digital biomarkers [13-16]. This process requires high levels of participant adherence and retention to ensure sufficient data are collected for validation purposes. Active data collection methods, such as digital questionnaires, which are often needed to help validate or contextualize digital biomarkers, intensify this challenge as they demand more time and effort than passive data collection methods such as actigraphy [17,18]. While passive measures may be useful, their validation depends on access to actively collected assessments of patient status. Therefore, understanding the drivers of participant adherence to data collection methods requiring active participation is crucial for the effective design and validation of digital biomarker collection platforms. Eysenbach described the "law of attrition," the tendency for a significant proportion of users to drop out or stop using eHealth mobile apps before completing a trial [19-23]. This dropout may shape the cohort over time in a way that does not represent the initially recruited study population in terms of demographic features, symptomatology, and clinical outcomes, therefore inducing bias in the analyses [16,24].

In studies requiring active remote digital participation, such as completing questionnaires, a few static parameters (ie, those that do not change over the course of the study) have been associated with adherence rates (the degree to which participants follow the study protocol and complete assessments), engagement rates (the proportion of participants who engage in the study after baseline, ie, who complete at least 1 nonbaseline questionnaire), and retention rates (the proportion of participants who continue to complete questionnaires over time). Previous studies suggest that participants who are more depressed at baseline, younger, and less educated tend to have lower assessment adherence rates [16,25,26]. Retention rates can also vary between treatment arms involving mobile apps versus control groups, where control groups with lower engagement requirements may often show better retention, and between ethnic groups [27]. A lower annual income also tends to be significantly associated with poorer study engagement [26].

These static parameters, often of a demographic or symptomatologic nature and considered as averages rather than longitudinal measures, provide meaningful insight into the question of participant engagement in active digital phenotyping. However, there has been less exploration of how dynamic parameters, such as longitudinal symptom change, might affect engagement. Therefore, the objective of this study was to identify novel predictive longitudinal markers of adherence, engagement, and retention to regular self-rated assessments in participants enrolling in a remote digital health study requiring active participation. To do so, we conducted secondary analyses on the Brighten study, a large longitudinal digital intervention study.

Methods

Study Design

The Brighten study consists of 2 completely remote, smartphone-based randomized controlled trials assessing the efficiency of mobile apps for treating depression. The study recruited 2193 participants from the United States through online advertisements, with recruitment for the first version (V1) beginning in August 2014 (recruitment lasted 5 months) and the second version (V2) starting in August 2016 (recruitment lasted 7 months), enrolling 1110 and 1083 participants, respectively [25,26]. Participants were included in the Brighten study if they had a score of 5 or higher on the Patient Health Questionnaire (PHQ)-9 or a score of 2 or greater on PHQ item 10 [28]. The study evaluated 3 interventions over 12 weeks: (1) iProblemSolve (iPST), an app designed by UCSF for problem-solving therapy [25,29,30]; (2) Project: EVO (Akili Interactive Labs), a therapeutic video game; and (3) Health Tips, an app offering strategies to improve mood [28]. The group using Health Tips was considered the control group. In those 2 randomized controlled trials, various questionnaires were sent to participants via smartphone notifications. Most assessments were submitted daily, weekly, or biweekly, but some questionnaires were only sent at baseline. The Brighten study also collected passive digital communication data from the V1 and V2 cohorts, along with passive mobility features from the V2 cohort. While V2 followed a protocol very similar to V1 (same interventions and active data collection methods), it specifically aimed to increase participation among individuals of Hispanic or Latino ethnic backgrounds to assess the feasibility of using digital mental health tools in this population [28].

Intervention Apps

The iPST app implements a 7-step problem-solving therapy model to manage mood: participants select a goal and are guided through a structured action plan. Participants assigned to this app were instructed to use it at least once per week, consistent with typical clinical use [31].

Participants assigned to Project: EVO were encouraged to use it 6 times per week for approximately 30 minutes per day. This video game–based app aimed at enhancing cognitive skills related to depression using adaptive algorithms to adjust difficulty to the user's proficiency [32]. Previous studies of the app's predecessor showed that this dosage improved cognitive control and depressive symptoms in older adults [32,33].

The control group used an app providing daily health tips, such as self-care (eg, showering) or physical activity (eg, walking). Participants in this group were not required to act on the tips, serving as a supportive control condition.

Importantly, all participants, regardless of their assigned intervention, completed their self-rated questionnaires in a separate study app. Our analyses specifically focused on adherence, engagement, and retention with these assessments.

Measures

Demographics such as gender, education level, working status, income satisfaction, last-year income, marital status, race, and

XSL•FO
RenderX

age, as well as parameters regarding study involvement (how they heard about the study, device used, and study version), are provided in Multimedia Appendix 1. These parameters, along with baseline PHQ-9 results, were imputed using MissForest, a nonparametric missing value imputation for mixed-type data, if less than 30% of the data were missing for that specific metric. This threshold was chosen based on validation studies demonstrating the reliability of MissForest under these conditions [34]. Participants were sent 7 questionnaires to complete at a predefined frequency specific to each questionnaire: PHQ-9 [35] (weekly for the first 4 weeks and then biweekly) and PHQ-2 [36] (daily); a 3-item sleep assessment assessing sleep onset latency, sleep duration, and time awake at night (weekly); the Sheehan Disability Scale (SDS) [37] (weekly for the first 4 weeks and then biweekly); the Patients Global Impression of Change (PGIC) Scale [38] (weekly); a survey inquiring about the use of mental health services (weekly); a survey inquiring about study app satisfaction (deployed at weeks 4, 8, and 12); and a survey inquiring about the use of other health-related apps (deployed at weeks 1, 4, 8, and 12) [28]. All questionnaires included in our analyses are provided in Multimedia Appendix 1.

Outcomes

We evaluated 3 key outcomes: adherence, engagement, and retention. To evaluate individual adherence, each participant's average completion rate was calculated as the ratio of completed questionnaires to the total number they were expected to complete. Engagement was binary: participants were considered engaged if they completed at least one nonbaseline assessment and disengaged otherwise. Therefore, disengaged participants, by definition, had a completion rate of 0%. Engaged participants were further divided into high- and low-completion groups, based on the median average completion rate. Based on previous research, engaged participants were also categorized into "improvers" and "nonimprovers" based on the difference between their baseline PHQ-9 score and their latest PHQ-9 score, with a minimum clinically important difference set at 5 points [39]. Therefore, engagement captures a binary distinction (completion beyond baseline), adherence represents a continuous measure of completion, and the subgrouping serves only as a descriptive stratification of engaged participants. Participant retention was determined as the duration for which participants continued completing a specific representative questionnaire (eg, PHQ-9) before they withdrew or stopped responding. It is important to consider that participants in the iPST and Project: EVO intervention groups were asked to use the apps as part of their treatment, but data on their engagement or adherence to the interventions were not available, so treatment adherence was not included as an outcome.

Statistical Methods

To compare engagement and average completion rates across demographic groups, Mann-Whitney U tests were used for 2-group comparisons, and Kruskal-Wallis tests followed by Dunn post hoc tests were applied for comparisons involving more than 2 groups.

A *t* test was conducted on average questionnaire scores and baseline scores between high and low completion groups, evaluating the importance of symptom severity as a static indicator of adherence. A *t* test was also conducted on baseline PHQ-9 scores between the engaged and disengaged groups to assess whether engagement was associated with baseline depressive symptoms.

To assess the potential of symptomatology as a dynamic predictor of adherence, trends in questionnaire scores over time between high and low completion groups were investigated using a repeated measures generalized linear model (GLM).

Elastic net regression with Shapley additive explanations (SHAP) values [40] was used to identify the top 5 predictive features for average completion rate, engagement status (engaged/disengaged), and completion group (high/low), incorporating demographic parameters, average questionnaire scores, and depressive symptom improvement. The models were subsequently retrained without including the study version (V1/V2) as a feature to assess its impact. Model generalizability was assessed by training on passive data from the V1 cohort and testing on the V2 cohort.

To study the impact of depression symptom improvement on adherence, differences in average completion rate and PHQ-9 completion rate between improvement groups were assessed using Mann-Whitney U tests. A robust linear model (RLM) regression was performed with improvement status, controlling for baseline PHQ-9 scores, to predict average completion rate. The RLM was chosen over other regression techniques to handle potential outliers in the data [41].

Cox proportional hazards models were constructed to predict retention, measured as time to last PHQ-9 questionnaire completion, between improvers and nonimprovers, incorporating demographic parameters and average scores from other questionnaires as covariates to evaluate their utility in predicting retention. We chose to study retention for the PHQ-9 rather than globally because completion rates and the frequency with which the questionnaires were sent to participants varied significantly between questionnaires, potentially compromising the reliability and validity of a comprehensive retention measure for survival analysis; the PHQ-9 was chosen as a representative questionnaire given its utility in usual clinical practice. Kaplan-Meier curves were plotted for these models to illustrate assessment retention over time.

Ethical Considerations

Ethical approval for both parent trials (V1: NCT00540865 and V2: NCT01808976) was granted by the Institutional Review Board of the University of California, San Francisco. Participants provided informed consent that included permission for their data to be shared with other researchers for secondary analyses, in line with the National Institute of Mental Health data-sharing policy (NOT-MH-19-033). Our study used only these deidentified datasets and involved no direct interaction with participants.

XSL•FO
RenderX

Results

Outcomes

There were 2201 participants in the dataset. Among them, 1093 were engaged, meaning they completed at least 1 nonbaseline questionnaire, and 1108 were disengaged. In the engaged group, the median average completion rate was 37.6% (IQR 15.5%-67.9%), and the mean was 40.6% (SD 24.2%). Within the engaged group, 548 participants had a mean completion rate higher than the median and were classified as the high completion group (mean 62.3%, SD 11.5%), while 550 participants were classified as the low completion group (mean 19%, SD 9.9%). The average completion rate for the entire cohort was 20.2% (SD 26.5%) for the whole cohort.

Demographics and Study Parameters

Kruskal-Wallis tests and Mann-Whitney U tests showed significant disparities in average completion rate and engagement between groups defined by gender, education, income satisfaction, last-year income, race, device used, study arm, and study version (Table 1). The findings on average completion rate align with previous literature [16,25,26], while the engagement findings are novel, with engagement defined as whether participants chose to remain in the study after completing the baseline assessments.

Table . Significant differences (P<.05) in engagement rates between groups defined by demographics or study parameters.

Category[a]	Engaged (n=1093)	Disengaged (n=1108)	P value[b]
Sex, n (%)			
Male	236 (21.59)	311 (28.07)	.025
Female	857 (78.41)	797 (71.93)	.025
Education, n (%)			
Graduate degree	216 (19.76)	132 (11.91)	2.9×10^{-5}
Income, n (%)			
<US $20,000	279 (25.53)	391 (35.29)	3.9×10^{-5}
$US 60,000 - US $80,000	277 (25.34)	172 (15.52)	7.1×10^{-7}
Income satisfaction, n (%)			
Cannot make ends meet	769 (70.36)	598 (53.97)	1.6×10^{-13}
Have enough to get along	194 (17.75)	360 (32.49)	1.2×10^{-13}
Race, n (%)			
Asian	95 (8.69)	54 (4.87)	.024
Non-Hispanic White	640 (58.55)	524 (47.29)	7.3×10^{-6}
Hispanic or Latino	191 (17.47)	366 (33.03)	3.4×10^{-15}
Study arm, n (%)			
iPST[c]	252 (23.06)	77 (6.95)	2.8×10^{-24}
Project: EVO	354 (32.39)	181 (16.34)	1.2×10^{-16}
HealthTips	369 (33.76)	151 (13.63)	8.7×10^{-27}
Study, n (%)			
Brighten-v1	742 (67.89)	376 (33.94)	3.7×10^{-55}
Brighten-v2	351 (32.11)	732 (66.06)	3.7×10^{-55}

[a]Percentages for a given category were calculated on the total number of engaged or disengaged participants. A participant is defined as being “engaged” if they completed at least 1 nonbaseline questionnaire of any kind. Participants who dropped out of the study immediately after baseline, without completing any nonbaseline questionnaire, are classified as “disengaged.”

[b]Bonferroni correction applied.

[c]iPST: iProblemSolve.

Symptomatology

To evaluate the impact of baseline symptomatology and average symptom severity over the course of the study on adherence, we compared baseline questionnaire scores and average questionnaire scores between high and low completion groups. After Bonferroni correction, the mean scores of the PHQ-2, PHQ-9, and PGIC questionnaires collected over the 12-week duration of the study were significantly different between high

XSL•FO
RenderX

and low completion groups (P<.001 for PHQ-2 and PHQ-9 and P=.002 for PGIC). Participants with lower PHfQ-2 scores and PHQ-9 scores (ie, less depressed participants), and participants with higher PGIC scores (ie, those who perceived better clinical improvement since the beginning of the study) were more likely to be in the high completion group. This suggests that participants with lower depression levels and perceived improvement in their clinical condition showed greater adherence compared to those with higher depression levels and perceived worsening of their condition (Table 2). Mean baseline PHQ-9 score was 13.9 (SD 4.8) for the engaged group and 14.4 (SD 5.0) for the disengaged group, with the difference being statistically significant (t_{1917}=−2.53; P=.01). This indicates that participants who were less depressed at baseline were more likely to complete at least 1 nonbaseline questionnaire; however, these baseline scores were not predictive of adherence in subsequent assessments.

Table . Comparisons of questionnaire scores averaged over the 12-week study period between high and low completion groups.

Questionnaire	High completion rate (SD)	Low completion rate (SD)	*t* score (*df*)	*P* value[a]
Mean score				
PHQ-2[b]	4.50 (1.52)	5.20 (1.84)	−6.793 (1019)	1.49×10^{-10}
PHQ-9[c]	9.30 (4.87)	11.27 (5.43)	−5.632 (752)	2.01×10^{-7}
Sleep assessment	7.23 (1.68)	7.43 (2.09)	−1.625 (768)	.837
SDS[d]	21.32 (8.50)	22.80 (9.21)	−2.658 (968)	.064
PGIC[e]	2.66 (0.97)	2.40 (1.07)	3.78 (744)	.001
Baseline score				
PHQ-9	13.75 (4.83)	13.96 (4.86)	−0.709 (1091)	≥.99
ALC[f]	3.50 (2.61)	3.64 (2.60)	−0.757 (773)	≥.99
GAD-7[g]	11.26 (5.78)	11.87 (5.66)	−1.529 (764)	≥.99

[a]Bonferroni correction applied.

[b]PHQ-2: Patient Health Questionnaire-2.

[c]PHQ-9: Patient Health Questionnaire-9.

[d]SDS: Sheehan Disability Scale.

[e]PGIC: Patients Global Impression of Change.

[f]ALC: Alcohol Use Questionnaire.

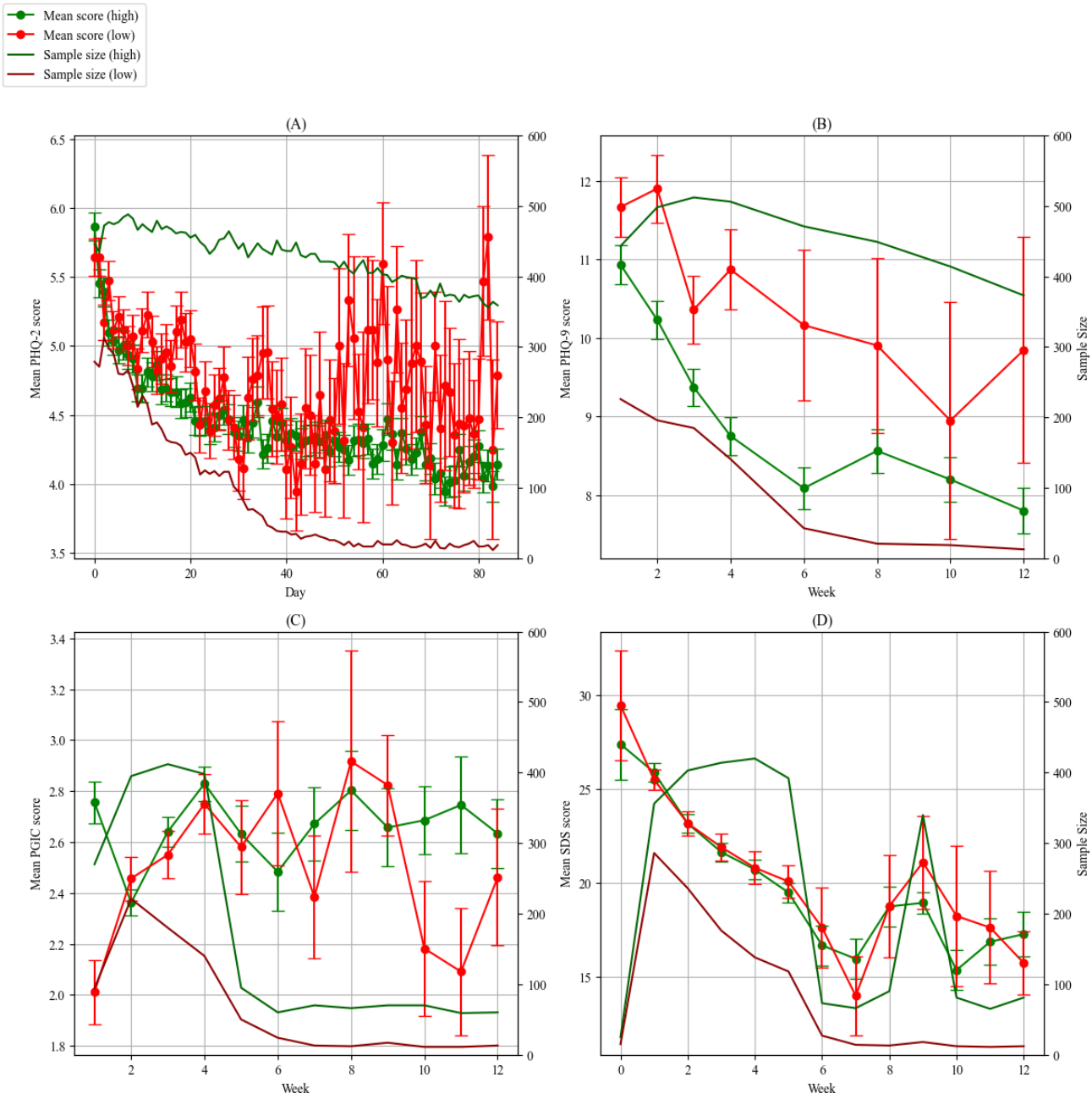
[g]GAD-7: Generalized Anxiety Disorder 7-Item Scale.

Clinical Improvement

To investigate the longitudinal relationship between questionnaire scores and adherence, we conducted a repeated measures GLM model of mean score over time for both high and low completion groups. We found that there was significant improvement in PHQ-2 scores (coeff=−0.01, 95% CI −0.01 to −0.01; P<.001) and PHQ-9 scores (week: coeff=−0.24, 95% CI −0.35 to −0.14; P<.001) over time for both high and low completion rate groups. Daily PHQ-2 scores were not significantly different between the 2 groups (coeff=0.04, 95% CI –0.14 to 0.22; P=.70), but the effect of time on the scores was statistically significant, with participants from the high completion group having a greater decrease in score over time (coeff=–0.012, 95% CI –0.014 to –0.009; P=.001) (Figure 1A). Averages of PHQ-9 score for each weekly (or biweekly) assessment over time were significantly different between the 2 groups (coeff=1.29, 95% CI 0.24 to 2.34; P=.021), but the longitudinal trend was similar (coeff=0.03, 95% CI –0.123 to 0.184; P=.70), with PHQ-9 scores decreasing significantly over time (coeff=–0.24, 95% CI –0.35 to –0.13; P<.001) (Figure 1B). Furthermore, the decrease in sample size regarding completion of PHQ-9 and PHQ-2 was significantly different between high and low completion groups (P_{PHQ-9}=.02 and P_{PHQ-2}<.001), with the sample size from the low completion group decreasing more rapidly. For other questionnaires, such as the sleep assessment, PGIC, and SDS, there was no statistically significant difference in the rate of sample size decrease between the high and low completion groups. There were no significant differences in SDS, PGIC, and sleep assessment scores between the 2 groups, although SDS scores decreased significantly over time for the 2 groups in a similar fashion (coeff=–0.84, 95% CI –1.17 to –0.50; P<.001). The effect of time on sleep assessment scores was found to be different between the groups, but it was nonsignificant (P=.08; Figure 1).

XSL•FO
RenderX

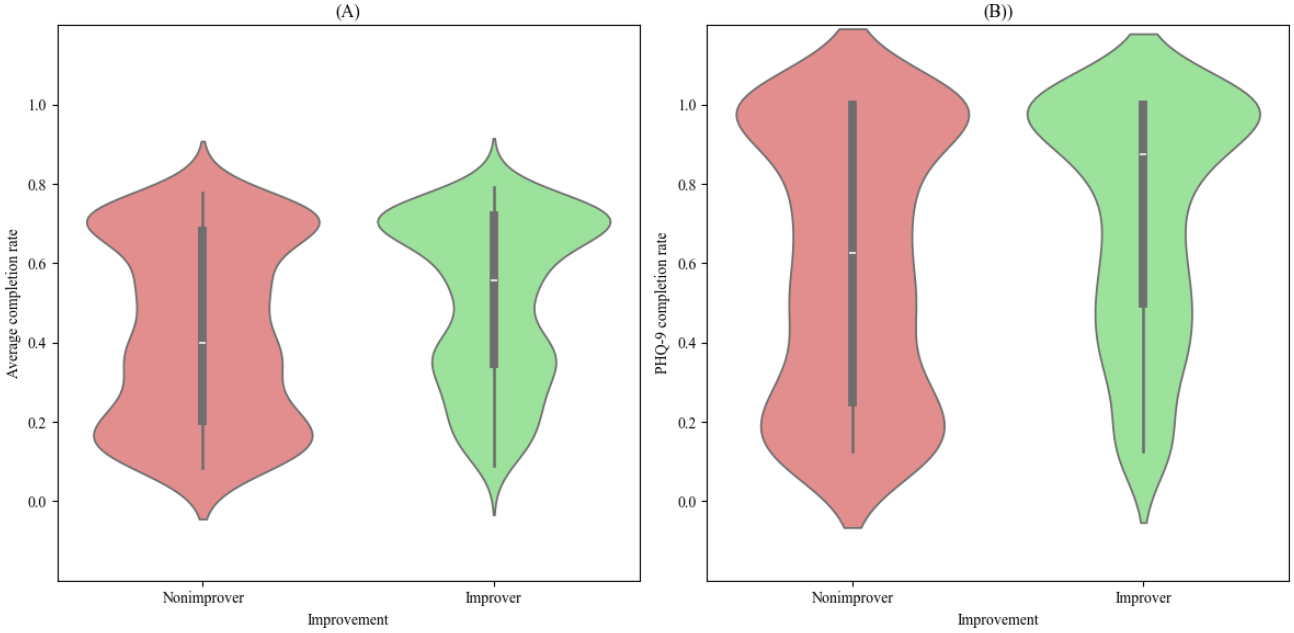
Figure 1. Mean scores over time for high- and low-completion groups by assessment, including sample sizes. (A) Mean Patient Health Questionnaire (PHQ)-2 score over time by completion group, (B) mean PHQ-9 score over time by completion group, (C) mean Patients Global Impression of Change (PGIC) score over time by completion group, and (D) mean Sheehan Disability Scale (SDS) score over time by completion group.



To evaluate the association between clinical improvement regarding depressive symptoms and adherence, we divided the engaged cohort into improvers and nonimprovers. The categorization was based on the difference between participants' baseline PHQ-9 score and their latest PHQ-9 score, using the minimal clinically important difference of 5 points as the threshold for meaningful improvement [39]. There were 410 "improvers" and 513 "nonimprovers." Between those groups, there was a significant difference in average completion rate (improvers: mean 50.7%, SD 20.6%; median 55.5%; nonimprovers: mean 42.6%, SD 22.7%; median 39.9%; and Mann-Whitney U=127,084; P<.001), PHQ-9 completion rate, and baseline PHQ-9 score (P<.001 for all comparisons), with "improvers" having higher baseline PHQ-9 scores than "nonimprovers" (Mann-Whitney U=137,574; P<.001) (Figure 2A and 2B). Mean baseline PHQ-9 score was 15.3 (SD 4.7) for improvers and 12.6 (SD 4.6) for nonimprovers. The RLM regression model showed that the average completion rate of "nonimprovers" was, on average, 8.85% (P<.001) lower than that of "improvers." However, for both groups, baseline PHQ-9 scores were not significantly associated with average completion rates (P=.06; Figure 2).

XSL•FO
RenderX

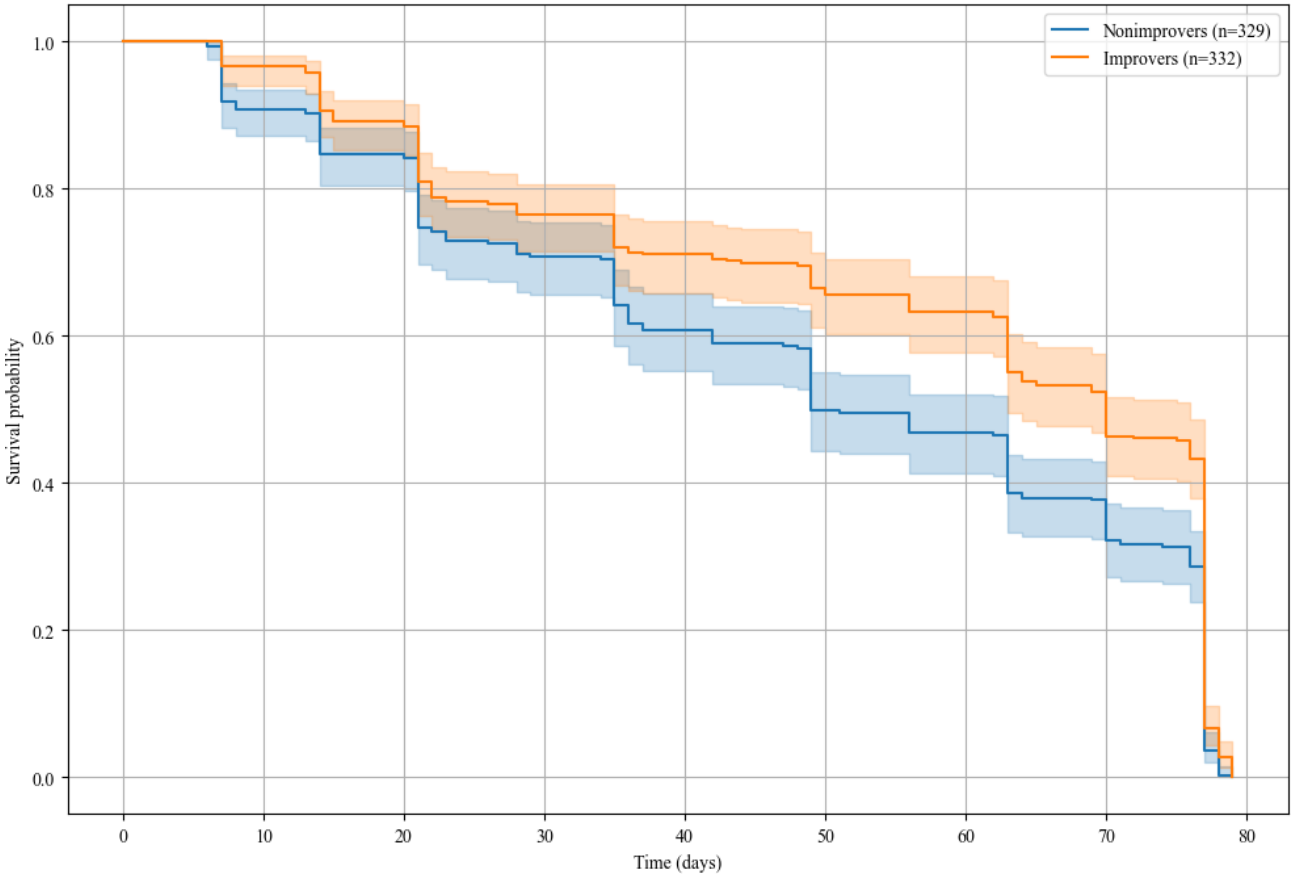
Figure 2. Comparison of adherence between improvers and nonimprovers. (A) Average completion rate (Mann-Whitney U score=127084; *P*<.0001 and (B) Patient Health Questionnaire (PHQ)-9 completion rate (Mann-Whitney U score=125418; *P*<.0001). Average completion rate is defined as the total number of questionnaires completed divided by the total number of questionnaires sent to participants. PHQ-9 completion rate is defined as the total number of PHQ-9 questionnaires completed divided by the total number of PHQ-9 questionnaires sent to participants.



To study PHQ-9 retention, defined as the duration for which participants continued completing the PHQ-9, we performed a Cox proportional hazards model to compare improvement groups, using demographic parameters [1], average scores from other questionnaires, and study version as covariates. The Cox proportional hazards model revealed significant differences in survival probabilities between the 2 groups (hazard ratio 0.78, 95% CI 0.67-0.91; *P*=.002), with the event defined as the day a participant completed their last PHQ-9 questionnaire within 12 weeks from the start of the study. The median survival time was 70 days (mean 56.8, SD 24.7) for improvers (n=332) and 49 days (mean 49.4, SD 25.6) for nonimprovers (n=329), indicating that nonimprovers stopped completing PHQ-9 questionnaires 7.4 days earlier on average compared to improvers (Mann-Whitney U=64632; *P*<.001). We found that age (*P*=.002), mean score of the sleep assessment (*P*=.01), and study version (*P*<.001) were significant predictors of PHQ-9 retention (Figure 3). Cox proportional hazards model based on the PHQ-2 questionnaire is provided in Multimedia Appendix 2.

XSL•FO
RenderX

Figure 3. Kaplan-Meier curves showing retention for improvers versus nonimprovers, with the event defined as the last Patient Health Questionnaire (PHQ)-9 questionnaire completed within the 12-week study duration (n=661). Significant survival differences were observed (improver Coef=−0.25, Exp(coef)=0.78; *P*<.005). Key predictors: age: Coef=−0.01, Exp(coef)=0.99; *P*<.005); mean score sleep: Coef=0.06, Exp(coef)=1.06; *P*=.01); and study Brighten-v2: Coef=0.79, Exp(coef)=2.19; *P*<.005). Bonferroni correction was applied.

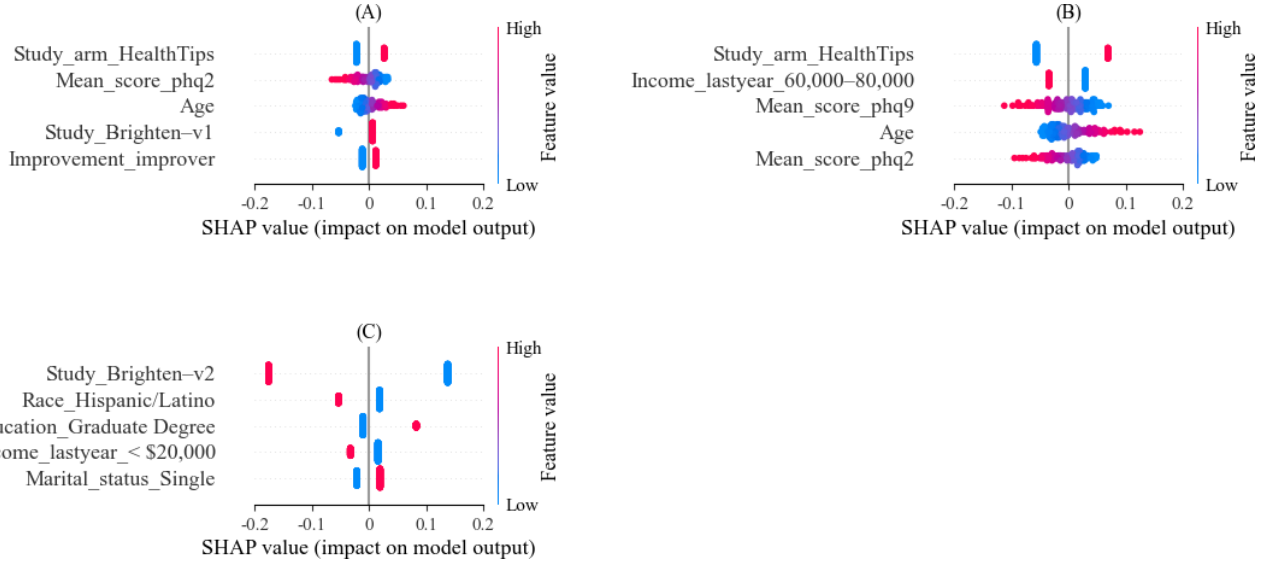


Key Indicators of Adherence and Engagement

To assess what features were most statistically useful in predicting average completion rate and engagement, we trained an elastic net regression model on 80% of the data and tested it on the remaining 20%. SHAP values were used to assess the relative importance of each feature and evaluate their contributions to the model's predictions (Figure 4). For predicting average completion rate, the 5 most useful features, in order of importance, were (n=701; mean squared error [MSE]=0.04) (1) being in the control group, (2) having a low mean PHQ-2 score, (3) being older, (4) being in the V1 cohort, and (5) improving in depressive symptoms (Figure 4A).

XSL•FO
RenderX

Figure 4. Top 5 features and their Shapley additive explanations (SHAP) value for predicting compliance metrics using elastic net regression. (A) Top 5 features for predicting average completion rate (n=701; mean squared error [MSE]=0.04), (B) top 5 features for categorizing participants into high and low completion groups (n=701; F_1-score=0.72), and (C) top 5 features for predicting whether a participant will be "engaged" or "disengaged" (n=2201; F_1-score accuracy=0.66).



The 5 most important features associated with being in the high completion group were (n=701; F_1-score=0.72) (1) being in the control group, (2) not being in the US $60,000-US $80,000 annual income group, (3) having a low mean PHQ-9 score, (4) being older, and (5) having a low mean PHQ-2 score (Figure 4B).

The 5 most important features associated with "engaging" in the study were (n=2201; F_1-score=0.65) (1) being in the V1 cohort, (2) not being a Hispanic or Latino person, (3) having a graduate degree, (4) having an annual income greater than US $20,000, and (5) being single (Figure 4C).

To assess the impact of study version (V1/V2) on the models' predictability, we conducted an exploratory analysis excluding it from the features and found that the categorization models (engagement and completion group) performed slightly worse, although the 5 most important features were generally similar (F_1-score decreased from 0.72 to 0.67 for completion group categorization and from 0.66 to 0.60 for engagement categorization). The average completion rate model MSE did not change, although the R^2 decreased from 0.16 to 0.13.

When trained on the V1 cohort and tested on the V2 cohort, the categorization models performed poorly (F_1-score=0.32 for engagement categorization and F_1-score=0.46 for completion group categorization). The model predicting average completion rate had slightly worse predictive performance (MSE=0.04; R^2=0.12).

Discussion

Principal Findings

We found that the subjective perception of a participant's own clinical trajectory (PGIC score) was associated with assessment adherence, with participants who reported a perceived improvement in their clinical condition showing greater adherence. We also found that increased depressive symptoms at baseline were not significantly associated with adherence, but rather with engagement, thus suggesting that more depressed participants were more likely to withdraw immediately after baseline assessment. Furthermore, participants who improved regarding depressive symptoms had greater assessment adherence and retention, despite having higher depressive symptoms at baseline. Previous research has established a link between baseline depressive symptoms and engagement [13]. This study makes novel contributions by exploring the associations between perceived clinical improvement and adherence, as well as the connection between depressive symptom improvement, baseline depression severity, retention, and adherence.

Despite offering monetary incentives, the researchers conducting the Brighten study observed poor engagement and adherence, highlighting the need to investigate this issue further and explore innovative strategies to address this limitation [25]. The demographic parameters we found associated with lower adherence and engagement are similar to what has been found in previous literature (ie, male sex, lower education, lower income, poorer income satisfaction, and Hispanic or Latino race were associated with poorer adherence) [16,25,26]. However, we found that a perceived worsening of clinical condition throughout the study was associated with less assessment adherence, suggesting that a participant's own perception of their change in activity, symptoms, emotion, and quality of life can influence their willingness to complete assessments.

Regarding completion of the PHQ-9 questionnaire, there was a significant difference in retention probability, with participants who improved in depressive symptoms having more chance of continuing to complete assessments than nonimprovers. Age, mean score of sleep assessment, and study version were significant features for predicting retention. Therefore, the global change in depressive symptoms over time was associated with study retention and adherence rate. These findings highlight the

XSL•FO
RenderX

possibility of attrition bias in digital health studies, leading to an overestimation of symptom severity improvement. Furthermore, this kind of bias could lead to misleading conclusions about the effectiveness of treatments and affect clinical decision-making, reinforcing the importance of controlling for attrition rates in digital health studies [42]. However, a possible limitation of the retention analyses is the lack of insight regarding improvement for nonimprovers who stopped completing answering questionnaires early in the study, as they may have eventually experienced symptom improvement that went unrecorded.

The study arm was the most important indicator of assessment completion rate, with participants from the control group having higher completion rates than participants using the Project: EVO app or the iPST app. Moreover, the finding that participants in the iPST group, who were encouraged to use the app once per week, showed higher adherence than those in the Project: EVO group, who were asked to use it 6 times per week, suggests a trade-off between effort and study engagement. Participants required to invest more time in the digital intervention were more likely to skip assessments. We found that baseline PHQ-9 score, although predictive of engagement (serving as an initial "filter" for participants continuing in the study), was not an important indicator of adherence, highlighting the need to focus on clinical trajectories in addition to static measures. In addition, participants who improved significantly in their depressive symptoms had higher completion rates than participants who did not, and their baseline PHQ-9 scores were significantly higher, although, as mentioned, baseline PHQ-9 score is not a direct indicator of completion rate. This suggests that while initial severity is not directly linked to longitudinal adherence, positive clinical progress may encourage consistent assessment completion. In fact, we found that participants who significantly improve regarding depressive symptoms are less likely to withdraw from active completion of assessments than participants who do not. Interestingly, it is also worth considering that the capacity to regularly complete questionnaires might be a potential predictor of future improvement in depressive symptoms.

Limitations

This study presents several limitations. The most important limitation is the lack of data on participants' use of the intervention apps, particularly the project: evo group, whose participants were encouraged to use the app 6 times per week. As a result, we were unable to assess whether engagement with the intervention apps influenced participants' engagement with questionnaire completion. However, our primary interest was in questionnaire completion, as this measure is more directly comparable across digital health studies and thus is the most applicable for broader research contexts. In addition, the GLM repeated measures analysis revealed intriguing patterns and trends in the effect of time and adherence on questionnaire scores, but the rapid decline in sample size over time limits the strength of any conclusions that can be drawn. The elastic net regression models showed only modest performance, with clear potential for improvement. Feature engineering, the collection of larger feature sets, hyperparameter tuning, and time-series transformations are potential strategies that could enhance the accuracy and robustness of the models by improving their ability to recognize underlying patterns, as well as identify temporal dynamics within the data [43-45]. Given the model's limited fit, SHAP value interpretations for top features should be regarded as exploratory. Furthermore, we were not able to generalize the model's predictability across cohorts, highlighting the need for further validation with diverse datasets to ensure its applicability in a realistic transdiagnostic population [46,47]. The limited data available for passively collected variables, such as communication and mobility features, resulted in inconclusive analyses including these variables (refer to Multimedia Appendix 3).

Future Directions

To address adherence and retention challenges in future studies, implementing continuous monitoring systems to identify participants at risk of dropping out could be beneficial [48]. Targeted outreach strategies, such as personalized reminders or support interventions, may help engage these individuals and improve adherence [49,50]. Monetary incentives have also been shown to be effective in increasing adherence and retention rates, though this may be impractical when attempting more naturalistic designs [51-53]. Beyond these approaches, qualitative research is crucial to better understand the underlying reasons for disengagement. Previous user-centered studies indicate that underserved users often cite stigma, cultural mismatch, and usability issues as barriers, while individuals with lower education or digital literacy report practical challenges and a higher perceived time burden when using mobile health apps [54,55]. Conducting more extensive qualitative work would provide valuable insights into how participants perceive their improvement and clinical trajectories, thereby offering a more comprehensive understanding of user challenges and informing strategies to strengthen engagement, adherence, and retention.

Conclusion

Although demographics are useful in predicting adherence, engagement, and retention in digital health studies, we show that dynamic, measurable parameters regarding symptomatology and symptom change are also important in understanding longitudinal patterns of attrition and adherence, and that they play a crucial role in shaping the remaining cohort as long-term studies progress over time [16,24]. We have shown that participants' impressions of their own improvement and the clinical trajectory of participants regarding depressive symptoms are important indicators of engagement, adherence, and retention. Further research is needed to assess the causality and implications of this effect and to determine optimal strategies for countering the bias that this phenomenon introduces into datasets.

XSL•FO
RenderX

Acknowledgments

The authors acknowledge financial support from the Mach-Gaensslen Foundation of Canada, the Dr. Clarke K. McLeod Memorial Scholarship, the Douglas Research Center, the Fonds de recherche du Québec – Nature et technologies (FRQNT; STRATÉGIA grant), and the Fonds de Recherche du Québec - Santé (FRQS; Junior 1 grant).

Funding

The authors acknowledge financial support from the Mach-Gaensslen Foundation of Canada, the Dr Clarke K McLeod Memorial Scholarship, the Douglas Research Center, the Fonds de recherche du Québec – Nature et technologies (FRQNT; STRATÉGIA grant), and the Fonds de Recherche du Québec - Santé (FRQS; Junior 1 Grant). The funders had no involvement in the study design, data collection, analysis, interpretation, or manuscript writing.

Authors' Contributions

DH helped conceive the study, performed the analyses, and wrote and revised the manuscript. YB assisted with the analyses and reviewed the manuscript. DP and EH reviewed the manuscript. DB helped conceive the study and provided supervision.

Conflicts of Interest

DB is a founder and shareholder of Aifred Health, a digital mental health company. Aifred Health was not involved in this research.

Multimedia Appendix 1
Significant differences in average completion rate between groups defined by demographics or study parameters, and questionnaires analyzed in the study.
[DOCX File, 581 KB - humanfactors_v12i1e69464_app1.docx]

Multimedia Appendix 2
Cox proportional hazards model for improvers and nonimprovers based on Patient Health Questionnaire (PHQ)-2 scores, with the event defined as the last PHQ-2 questionnaire completed within the 12-week study duration.
[DOCX File, 117 KB - humanfactors_v12i1e69464_app2.docx]

Multimedia Appendix 3
Top 5 features and their Shapley additive explanations (SHAP) values for predicting adherence metrics using elastic net regression trained on demographics, study parameters, average questionnaire scores, and passive data features.
[DOCX File, 140 KB - humanfactors_v12i1e69464_app3.docx]

References

1. Chen IM, Chen YY, Liao SC, Lin YH. Development of digital biomarkers of mental illness via mobile apps for personalized treatment and diagnosis. J Pers Med 2022 Jun 6;12(6):936. [doi: 10.3390/jpm12060936] [Medline: 35743722]
2. Fernandes BS, Williams LM, Steiner J, Leboyer M, Carvalho AF, Berk M. The new field of “precision psychiatry”. BMC Med 2017 Apr 13;15(1):80. [doi: 10.1186/s12916-017-0849-x] [Medline: 28403846]
3. Desai S, Tanguay-Sela M, Benrimoh D, et al. Identification of suicidal ideation in the Canadian Community Health Survey-Mental Health Component using deep learning. Front Artif Intell 2021;4:561528. [doi: 10.3389/frai.2021.561528] [Medline: 34250463]
4. Gashkarimov VR, Sultanova RI, Efremov IS, Asadullin AR. Machine learning techniques in diagnostics and prediction of the clinical features of schizophrenia: a narrative review. Consort Psychiatr 2023 Sep 29;4(3):43-53. [doi: 10.17816/CP11030] [Medline: 38249535]
5. Benoit J, Onyeaka H, Keshavan M, Torous J. Systematic review of digital phenotyping and machine learning in psychosis spectrum illnesses. Harv Rev Psychiatry 2020;28(5):296-304. [doi: 10.1097/HRP.0000000000000268] [Medline: 32796192]
6. Althubaiti A. Information bias in health research: definition, pitfalls, and adjustment methods. J Multidiscip Healthc 2016;9:211-217. [doi: 10.2147/JMDH.S104807] [Medline: 27217764]
7. May M. Clinical trial costs go under the microscope. Nat Med 2019 Mar 6. [doi: 10.1038/d41591-019-00008-7]
8. Benrimoh D, Fratila R, Israel S, et al. Aifred health, a deep learning powered clinical decision support system for mental health. In: Escalera S, Weimer M, editors. The NIPS '17 Competition: Building Intelligent Systems: Springer; 2018:251-287. [doi: 10.1007/978-3-319-94042-7_13]
9. Steinhubl SR, Muse ED, Topol EJ. The emerging field of mobile health. Sci Transl Med 2015 Apr 15;7(283). [doi: 10.1126/scitranslmed.aaa3487]

XSL•FO
RenderX

10. Nowels MA, McDarby M, Brody L, et al. Predictors of engagement in multiple modalities of digital mental health treatments: longitudinal study. J Med Internet Res 2024 Nov 7;26(1):e48696. [doi: 10.2196/48696] [Medline: 39509696]
11. Price M, Yuen EK, Goetter EM, et al. mHealth: a mechanism to deliver more accessible, more effective mental health care. Clin Psychol Psychother 2014;21(5):427-436. [doi: 10.1002/cpp.1855] [Medline: 23918764]
12. Qassim S, Golden G, Slowey D, et al. A mixed-methods feasibility study of a novel AI-enabled, web-based, clinical decision support system for the treatment of major depression in adults. Journal of Affective Disorders Reports 2023 Dec;14:100677. [doi: 10.1016/j.jadr.2023.100677]
13. Dixon WG, Michaud K. Using technology to support clinical care and research in rheumatoid arthritis. Curr Opin Rheumatol 2018 May;30(3):276-281. [doi: 10.1097/BOR.0000000000000485] [Medline: 29369089]
14. Sim I. Mobile devices and health. N Engl J Med 2019 Sep 5;381(10):956-968. [doi: 10.1056/NEJMra1806949] [Medline: 31483966]
15. Lattie EG, Adkins EC, Winquist N, Stiles-Shields C, Wafford QE, Graham AK. Digital mental health interventions for depression, anxiety, and enhancement of psychological well-being among college students: systematic review. J Med Internet Res 2019 Jul 22;21(7):e12869. [doi: 10.2196/12869] [Medline: 31333198]
16. Pratap A, Neto EC, Snyder P, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. npj Digit Med ;3(1):21. [doi: 10.1038/s41746-020-0224-8]
17. Torous J, Bucci S, Bell IH, et al. The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality. World Psychiatry 2021 Oct;20(3):318-335. [doi: 10.1002/wps.20883] [Medline: 34505369]
18. Schick A, Rauschenberg C, Ader L, et al. Novel digital methods for gathering intensive time series data in mental health research: scoping review of a rapidly evolving field. Psychol Med 2023 Jan;53(1):55-65. [doi: 10.1017/S0033291722003336] [Medline: 36377538]
19. Druce KL, Dixon WG, McBeth J. Maximizing engagement in mobile health studies. Rheum Dis Clin North Am 2019 May;45(2):159-172. [doi: 10.1016/j.rdc.2019.01.004]
20. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11. [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
21. Eysenbach G. The law of attrition revisited – author's reply. J Med Internet Res 2006 Sep 29;8(3):e21. [doi: 10.2196/jmir.8.3.e21]
22. Christensen H, Mackinnon A. The law of attrition revisited. J Med Internet Res 2006 Sep 29;8(3):e20. [doi: 10.2196/jmir.8.3.e20] [Medline: 17032636]
23. Dorsey ER, Yvonne Chan YF, McConnell MV, Shaw SY, Trister AD, Friend SH. The use of smartphones for health research. Acad Med 2017 Feb;92(2):157-160. [doi: 10.1097/ACM.0000000000001205] [Medline: 27119325]
24. Nunan D, Aronson J, Bankhead C. Catalogue of bias: attrition bias. BMJ Evid Based Med 2018 Feb;23(1):21-22. [doi: 10.1136/ebmed-2017-110883] [Medline: 29367321]
25. Anguera JA, Jordan JT, Castaneda D, Gazzaley A, Areán PA. Conducting a fully mobile and randomised clinical trial for depression: access, engagement and expense. BMJ Innov 2016 Jan;2(1):14-21. [doi: 10.1136/bmjinnov-2015-000098] [Medline: 27019745]
26. Pratap A, Renn BN, Volponi J, et al. Using mobile apps to assess and treat depression in Hispanic and Latino populations: fully remote randomized clinical trial. J Med Internet Res 2018 Aug 9;20(8):e10130. [doi: 10.2196/10130] [Medline: 30093372]
27. Arean PA, Hallgren KA, Jordan JT, et al. The use and effectiveness of mobile apps for depression: results from a fully remote clinical trial. J Med Internet Res 2016 Dec 20;18(12):e330. [doi: 10.2196/jmir.6482] [Medline: 27998876]
28. Pratap A, Homiar A, Waninger L, et al. Real-world behavioral dataset from two fully remote smartphone-based randomized clinical trials for depression. Sci Data 2022 Aug 27;9(1):522. [doi: 10.1038/s41597-022-01633-7] [Medline: 36030226]
29. Mynors-Wallis LM, Gath DH, Day A, Baker F. Randomised controlled trial of problem solving treatment, antidepressant medication, and combined treatment for major depression in primary care. BMJ 2000 Jan 1;320(7226):26-30. [doi: 10.1136/bmj.320.7226.26] [Medline: 10617523]
30. Arean P, Hegel M, Vannoy S, Fan MY, Unuzter J. Effectiveness of problem-solving therapy for older, primary care patients with depression: results from the IMPACT project. Gerontologist 2008 Jun;48(3):311-323. [doi: 10.1093/geront/48.3.311] [Medline: 18591356]
31. Malouff JM, Thorsteinsson EB, Schutte NS. The efficacy of problem solving therapy in reducing mental and physical health problems: a meta-analysis. Clin Psychol Rev 2007 Jan;27(1):46-57. [doi: 10.1016/j.cpr.2005.12.005] [Medline: 16480801]
32. Anguera JA, Gunning FM, Areán PA. Improving late life depression and cognitive control through the use of therapeutic video game technology: a proof-of-concept randomized trial. Depress Anxiety 2017 Jun;34(6):508-517. [doi: 10.1002/da.22588] [Medline: 28052513]
33. Anguera JA, Boccanfuso J, Rintoul JL, et al. Video game training enhances cognitive control in older adults. Nature New Biol 2013 Sep 5;501(7465):97-101. [doi: 10.1038/nature12486] [Medline: 24005416]
34. Stekhoven DJ, Bühlmann P. MissForest--non-parametric missing value imputation for mixed-type data. Bioinformatics 2012 Jan 1;28(1):112-118. [doi: 10.1093/bioinformatics/btr597] [Medline: 22039212]

XSL•FO
RenderX

35. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613. [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
36. Kroenke K, Spitzer RL, Williams JBW. The Patient Health Questionnaire-2: validity of a two-item depression screener. Med Care 2003 Nov;41(11):1284-1292. [doi: 10.1097/01.MLR.0000093487.78664.3C] [Medline: 14583691]
37. Sheehan DV, Harnett-Sheehan K, Raj BA. The measurement of disability. Int Clin Psychopharmacol 1996 Jun;11 Suppl 3:89-95. [doi: 10.1097/00004850-199606003-00015] [Medline: 8923116]
38. Remigio-Baker RA, Hungerford LD, Bailie JM, Ivins BJ, Lopez J, Ettenhofer ML. The Patient Global Impression of Change as a complementary tool to account for neurobehavioral and mental health symptom improvement for patients with concussion. Disabil Rehabil 2025 Jan 2;47(1):235-243. [doi: 10.1080/09638288.2024.2346233]
39. Kroenke K. Enhancing the clinical utility of depression screening. CMAJ 2012 Feb 21;184(3):281-282. [doi: 10.1503/cmaj.112004]
40. Lundberg S, Lee SI. A unified approach to interpreting model predictions. arXiv. Preprint posted online on May 22, 2017. [doi: 10.48550/arXiv.1705.07874]
41. Robust regression | R data analysis examples. UCLA Institute for Digital Research and Education (UCLA IDRE). URL: https://stats.oarc.ucla.edu/r/dae/robust-regression/ [accessed 2024-11-05]
42. Tripepi G, Jager KJ, Dekker FW, Zoccali C. Selection bias and information bias in clinical research. Nephron Clin Pract 2010;115(2):c94-c99. [doi: 10.1159/000312871] [Medline: 20407272]
43. Rawat T, Khemchandani V. Feature engineering (FE) tools and techniques for better classification performance. IJIET 2017;8(2). [doi: 10.21172/ijiet.82.024]
44. A Ilemobayo J, Durodola O, Alade O, et al. Hyperparameter tuning in machine learning: a comprehensive review. J Eng Res Rep 2024 Jun 7;26(6):388-395. [doi: 10.9734/jerr/2024/v26i61188]
45. Spiliotis E. Time series forecasting with statistical, machine learning, and deep learning methods: past, present, and future. In: Hamoudia M, Makridakis S, Spiliotis E, editors. Forecasting with Artificial Intelligence: Palgrave Macmillan; 2023:49-75. [doi: 10.1007/978-3-031-35879-1_3]
46. Dalgleish T, Black M, Johnston D, Bevan A. Transdiagnostic approaches to mental health problems: current status and future directions. J Consult Clin Psychol 2020 Mar;88(3):179-195. [doi: 10.1037/ccp0000482] [Medline: 32068421]
47. Wise T, Robinson OJ, Gillan CM. Identifying transdiagnostic mechanisms in mental health using computational factor modeling. Biol Psychiatry 2023 Apr 15;93(8):690-703. [doi: 10.1016/j.biopsych.2022.09.034] [Medline: 36725393]
48. Nahum-Shani I, Smith SN, Spring BJ, et al. Just-in-time adaptive interventions (JITAIs) in mobile health: key components and design principles for ongoing health behavior support. Ann Behav Med 2018 May 18;52(6):446-462. [doi: 10.1007/s12160-016-9830-8] [Medline: 27663578]
49. Brueton V, Stenning SP, Stevenson F, Tierney J, Rait G. Best practice guidance for the use of strategies to improve retention in randomized trials developed from two consensus workshops. J Clin Epidemiol 2017 Aug;88:122-132. [doi: 10.1016/j.jclinepi.2017.05.010] [Medline: 28546093]
50. O'Connor S, Hanlon P, O'Donnell CA, Garcia S, Glanville J, Mair FS. Understanding factors affecting patient and public engagement and recruitment to digital health interventions: a systematic review of qualitative studies. BMC Med Inform Decis Mak 2016 Dec;16(1):120. [doi: 10.1186/s12911-016-0359-3]
51. Dickert N, Grady C. What's the price of a research subject? Approaches to payment for research participation. N Engl J Med 1999 Jul 15;341(3):198-203. [doi: 10.1056/NEJM199907153410312] [Medline: 10403861]
52. Gelinas L, Largent EA, Cohen IG, Kornetsky S, Bierer BE, Fernandez Lynch H. A framework for ethical payment to research participants. N Engl J Med 2018 Feb 22;378(8):766-771. [doi: 10.1056/NEJMsb1710591] [Medline: 29466147]
53. Bentley JP, Thacker PG. The influence of risk and monetary payment on the research participation decision making process. J Med Ethics 2004 Jun;30(3):293-298. [doi: 10.1136/jme.2002.001594] [Medline: 15173366]
54. Saad F, Eisenstadt M, Liverpool S, Carlsson C, Vainieri I. Self-guided mental health apps targeting racial and ethnic minority groups: scoping review. JMIR Ment Health 2023 Dec 6;10(1):e48991. [doi: 10.2196/48991] [Medline: 38055315]
55. Boucher EM, Raiker JS. Engagement and retention in digital mental health interventions: a narrative review. BMC Digit Health 2024 Aug 8;2(1):52. [doi: 10.1186/s44247-024-00105-9]

Abbreviations

GLM: generalized linear model
iPST: iProblemSolve
MSE: mean squared error
PGIC: Patients Global Impression of Change
PHQ-2: Patient Health Questionnaire-2
PHQ-9: Patient Health Questionnaire-9
RLM: robust linear model
SDS: Sheehan Disability Scale
SHAP: Shapley additive explanations

XSL•FO
RenderX

Edited by A Kushniruk; submitted 30.11.24; peer-reviewed by S Pila, TH Dang; accepted 20.11.25; published 22.12.25.

Please cite as:
Hamitouche D, Barkat Y, Parekh D, Hammer E, Benrimoh D
Dynamic Indicators of Adherence and Retention in Adults Using a Digital Mental Health App: Longitudinal Observational Analysis From the Brighten Study
JMIR Hum Factors 2025;12:e69464
URL: https://humanfactors.jmir.org/2025/1/e69464
doi:10.2196/69464

© Dylan Hamitouche, Youcef Barkat, Deven Parekh, Eva Hammer, David Benrimoh. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 22.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Use and Acceptance of Innovative Digital Health Solutions Among Patients and Professionals: Survey Study

Fritz Seidl[1], MA, MBA; Florian Hinterwimmer[1,2], PhD; Ferdinand Vogt[3,4], MD, MHBA; Günther M Edenharter[5], MD; Karl F Braun[6], MD; Rüdiger von Eisenhart-Rothe[1], MD; AG Digitalisierung der DGOU; Peter Biberthaler[6], MD; Dominik Pförringer[6], MD

[1]Department of Orthopaedics and Sports Orthopaedics, Klinikum rechts der Isar, Technical University of Munich, TUM University Hospital, Ismaninger Str. 22, Munich, Germany

[2]Institute for AI and Informatics in Medicine, Technical University of Munich, Munich, Germany

[3]Department of Cardiac Surgery, Artemed Clinic Munich South, Munich, Germany

[4]Department of Cardiac Surgery, Paracelsus Medical University, Nuremberg, Germany

[5]Department of Anaesthesiology, Klinikum rechts der Isar, Technical University of Munich, TUM University Hospital, Munich, Germany

[6]Department of Trauma Surgery, Klinikum rechts der Isar, Technical University of Munich, TUM University Hospital, Munich, Germany

Corresponding Author:
Fritz Seidl, MA, MBA
Department of Orthopaedics and Sports Orthopaedics, Klinikum rechts der Isar, Technical University of Munich, TUM University Hospital, Ismaninger Str. 22, Munich, Germany

Abstract

Background: Digital solutions are gaining increasing importance and present a challenge regarding their introduction and acceptance into professional medical environments. Significant advances have been made regarding the availability, safety, and ease of use of data generated by a multitude of devices and wearables. However, data security and data protection are delaying factors. The underlying analysis focuses on the use and acceptance of digital solutions, and their respective differences between health care professionals and patients.

Objective: This study examines the current use and acceptance of digital solutions among health care professionals and patients. In addition, it derives an outlook on future developments and expectations in the setting of innovative technologies able to penetrate the health market.

Methods: An anonymous web-based survey of 23 multiple-choice and 3 open-text questions was conducted among medical professionals and patients between April and September 2023. In this study, quantitative analysis was performed using Python, with Pandas for data processing and Matplotlib for visualization. Chi-square tests were used to analyze binary categorical data, while Mann-Whitney *U* tests were used to evaluate ordinal data. Additionally, a qualitative analysis was conducted to summarize the results of the open-ended questions.

Results: During 178 days, the survey garnered 2058 clicks, resulting in 1389 participants (67.5% response rate). A total of 1002 participants completed the entire questionnaire, while 387 (27.9%) did not finish. Incomplete responses were excluded from the comprehensive analysis. The sample comprised 271 (27%) physicians and 731 (73%) patients. The study found significant agreement between both groups in adopting and foreseeing the use of digital health tools and telemedicine. Both groups recognized the future importance of digital health without substantial differences.

Conclusions: Overall, attitudes toward digital health and telemedicine were consistent, reflecting a uniform acceptance and expectation of these technologies among health care professionals and patients. The consensus on telemedicine's future role over the next 5 years indicates a unified vision for digital health paradigms. These consistencies between the 2 groups might be future drivers for improvements in accessibility, convenience, and efficiency in health care delivery.

(JMIR Hum Factors 2025;12:e60779) doi:10.2196/60779

KEYWORDS

eHealth; digital health; medical data use; information; survey; adoption

Introduction

Digital innovation in health care is accelerating in growth and presence in the media, while the regulatory framework is constantly evolving and adopting. Digitization has found its role in everyday medical life. It is of paramount importance to understand its use and acceptance to improve and continuously develop technology-based health offerings.

XSL•FO
RenderX

First of all, it makes sense to define the term *digital health*: in the search for a unified definition, diverse terminologies and conceptual frameworks are identified in the literature. Fatehi et al [1] conducted a quantitative analysis and term mapping of 95 unique definitions, revealing that digital health is predominantly concerned with the provision of health care rather than the mere use of technology. Their findings highlight a focus on the well-being of individuals at both population and individual levels, emphasizing the use of data and information to enhance patient care. The prevalence of mobile health (mHealth) as a dominant concept within digital health ecosystems was evident, closely interlinked with telehealth and artificial intelligence (AI) applications [1]. In turn, digital health can be subcategorized into a variety of dimensions. eHealth (and its subset mHealth) serves as a foundational category, in which health information technology plays a critical role, focusing on the management and exchange of health information through electronic health records and interoperability systems [2,3]. In the following, these definitions will be assumed for this study.

Developments have been catalyzed by the global COVID-19 pandemic situation [4] and by early legislation in favor of digital health applications (eg, in Germany under the acronym DiGA [*Digitale Gesundheitsanwendungen*]) [5]. Many promises—such as the availability of telemedicine—have been met, while specific needs regarding data storage are strongly underserved [6]. The use and development of digital health applications differ in addressing the needs of professionals versus those of patients. Acceptance is influenced by factors varying in complexity and depth, often driven by fears or lack of understanding. National legal and technical regulations have yielded a desired increase in the development and usage of digital health care solutions. The underlying study examines the current status and the drivers and barriers of this development, which have been found to be major influencing factors in the adoption of digital technologies in health care [7]. Specifically, it analyzes the gaps in the aforementioned categories between patients and professionals. Additionally, it helps understand the current and future needs of digital health and the requirements for offered applications.

A multinational survey among health care professionals and patients focused on this delta, its underlying reasons, and the consecutively arising demand specifically regarding the following five research questions:

1. How do estimations of the future importance of digital health differ between patients and professionals?
2. How do patients and professionals estimate their future use of digital health solutions?
3. Which areas offer specific innovation potential in digital health from the perspective of patients and professionals?
4. Which key obstacles to using digital health solutions do patients and professionals identify?
5. How does the use of web-based access differ toward health data, electronic appointment booking, and telemedicine?

Every nonmedical professional from the general population is considered a patient for research purposes.

Methods

Web-Based Survey

An anonymous web-based survey, adopted from some of the authors' [FS, VF, GME, DP] earlier research in 2018 [8], was conducted between April and September 2023 via the scientific survey platform SoSci Survey [9] hosted on German servers complying with European data protection laws. The survey was initiated with an introductory question distinguishing between patients and physicians, deciding on the further direction of the questionnaire. Depending on the chosen branch, 19 questions for patients and 25 for physicians followed, 3 of which were open-text questions. The questionnaire was provided in German and English and was available for 6 months: 6 April through 30 September 2023. Only fully completed questionnaires were considered for final analysis to guarantee consistent data quality.

The link to the web-based survey was distributed via publications in scientific journals such as *Die Unfallchirurgie*, the networks of the German Society for Orthopaedic and Trauma Surgery (*Deutsche Gesellschaft für Orthopädie und Unfallchirurgie*) and the German Society for Surgery (*Deutsche Gesellschaft für Chirurgie*), and web-based communication platforms such as LinkedIn and university web pages to reach a broad range of participants in both the professional field and among patients.

Participants were asked a set of specific questions regarding their personal and professional use, their acceptance and future expectations toward digital health care, followed by sociodemographic questions, and questions regarding their field of expertise, age, and country of origin. In addition, open-text questions were offered to understand and evaluate individual needs and gain additional insights. The English version of the survey is in Multimedia Appendix 1.

In this study, a quantitative analysis was conducted using the Python programming language, leveraging the capabilities of the Pandas library for data processing and the Matplotlib library for visualization. This approach allowed for a practical assessment of relative frequencies and distribution patterns in the responses of health care professionals and patients. For binary categorical data, such as responses regarding the "use of electronic appointment system," chi-square tests were used to assess the association between the respondents' profession and answers. In cases of ordinal data, including perceptions of the "future importance of digital health" and estimations of telemedicine use in the past and future, the Mann-Whitney *U* test was used. This nonparametric test compares the central tendency of 2 independent samples and is optimally suited for data not conforming to a normal distribution. Responses were numerically coded to facilitate this analysis. The primary objective of this investigation was to compare and contrast attitudes toward and usage of digital health tools, including aspects of telemedicine and web-based health care services. The qualitative data were assessed by a panel of experts (including a physician, a data scientist, and a management expert) and evaluated for underlying trends and overlaps between the questionnaire participants.

XSL•FO
RenderX

Ethical Considerations

This study was approved by the ethics committee of Klinikum rechts der Isar, Technical University of Munich, before distribution. The approval number is 2023 - 146-S-NP. By answering the mandatory questions of the survey and submitting them, the participants provided their consent to the anonymous study. In case they refrained from submitting the answers or refrained from answering specific mandatory questions, their consent was denied. In these cases, data were neither submitted nor recorded in any form. The participants of the survey received no compensation in any form. The privacy and confidentiality of all research subjects' data and identity were maintained at all times, because all data were collected anonymously and no identifying information was submitted at any time.

Results

Analysis of the Replies to the Web-Based Survey

During the 178 days of availability, the questionnaire yielded 2058 clicks, of which 1389 visitors answered the questionnaire and became participants in the study (response rate of 67.5% among visitors). A total of 1002 participants (72.1% of the respondents; 947 of whom were German-speaking and 55 of whom were English-speaking) completed the entire questionnaire, and the remaining 387 (27.9%) participants stopped answering the questionnaire before reaching the last question. The resultant dataset comprised 731 patient-derived responses juxtaposed with 271 physician-sourced answers, offering a representative overview of both primary health care stakeholders.

The study revealed a high level of conformity in perspectives between physicians and patients regarding the adoption and future use of digital health tools and telemedicine. Both groups showed parallel levels of adoption for electronic appointment systems and similar historical use patterns of telemedicine offerings. It yielded a comparable evaluation of the future importance of digital health, with no significant differences observed in their responses. Notably, the analysis indicated a consensus in expectations toward future use over the next 5 years, highlighting a synergistic alignment in the outlook toward digital health paradigms. The general attitude toward digital health and telemedicine yielded no significant differences either, underscoring a uniformity in the acceptance and anticipated future use of these technologies among both groups. Future expectations were aligned, with physicians and patients recognizing the importance and potential of digital health solutions in improving health care delivery. This consensus points toward a collaborative readiness for integrating digital health tools in future health care services, emphasizing the transformative potential of digital technologies in the medical field.

When asked about their “web-based access status to the hospital information system,” 66.1% (179/271) of physicians replied positively (“yes”), while 33.9% (92/271) did not have access (“no”). This result refers exclusively to the answers of the doctors and excludes patient data.

As demonstrated in Table 1, the “use of electronic appointment system” elicited a nonsignificant difference in usage patterns between physicians and patients (chi-square test, P=.16). This lack of statistical significance indicates parallel levels of adoption between the 2 groups.

Table . Proportion of health care professionals and patients using the “electronic appointment systems.”

	Yes, n/N (%)	No, n/N (%)	*P* value[a]
Group			.16
Physicians	159/271 (58.73)	114/271 (41.99)	
Patients	454/731 (62.11)	285/731 (38.97)	

[a]Chi-square test.

The assessment of the perceived “future importance of digital health” is visualized in Table 2. Responses from both cohorts indicated no statistically significant divergence (Mann-Whitney *U* test, P=.13), indicating a uniform recognition of the importance of digital health across the medical community and patient population.

Table . Comparison of the perceived “future importance of digital health” between health care professionals and patients.

	Physicians, n/N (%)	Patients, n/N (%)	*P* value[a]
Response			.13
Very important	181/271 (66.92)	562/731 (76.90)	
Important	65/271 (23.92)	129/731 (17.67)	
Not that important	26/271 (9.43)	47/731 (6.47)	
Not important	5/271 (1.74)	4/731 (0.60)	

[a]Mann-Whitney *U* test.

As delineated in Table 3, historical use patterns of telemedicine similarly showed no statistical difference between physicians

XSL•FO
RenderX

and patients (Mann-Whitney *U* test, *P*=.86), indicating a shared experience regarding the integration of telemedicine into health care delivery over the past 5 years.

Table . Comparison of responses to "use of telemedicine in the last 5 years" between health care professionals and patients.

	Physicians, n/N (%)	Patients, n/N (%)	*P* value[a]
Response			.86
Increased	214/271 (78.97)	578/731 (79.11)	
Remained the same	53/271 (19.61)	140/731 (19.22)	
Decreased	9/271 (3.14)	22/731 (2.98)	

[a]Mann-Whitney *U* test.

Prospective inclinations toward "future use of telemedicine in the next 5 years" are graphically represented in Table 4. The analysis revealed no statistically significant difference in expectations between the 2 groups (Mann-Whitney *U* test, *P*=.10), highlighting a consensus in the projected trajectory of telemedicine's role in future health care services.

Table . Expectations for the "future use of telemedicine in the next 5 years" among health care professionals and patients.

	Physicians, n/N (%)	Patients, n/N (%)	*P* value[a]
Response			.10
Will increase	229/271 (84.67)	647/731 (88.56)	
Will remain the same	36/271 (13.30)	76/731 (10.46)	
Will decrease	8/271 (2.90)	19/731 (2.59)	

[a]Mann-Whitney *U* test.

Collectively, the data shows a conformity in perspectives between physicians and patients regarding the adoption and future use of digital health tools and telemedicine. The absence of statistically significant differences underscores a potentially synergistic alignment in the outlook toward evolving digital health paradigms—in other words, a contextual significance with regard to similarities between the 2 groups. Table 5 provides a comparative statistical analysis of responses on digital health adoption between health care professionals and patients using chi-square and Mann-Whitney *U* tests. We used a significance threshold of *P*<.05 to determine whether the differences between the groups were statistically significant. As shown in the table, none of the *P* values for the comparisons (eg, *P*=.16 and *P*=.13) were below this threshold, indicating that there were no statistically significant differences in attitudes between health care professionals and patients. This suggests that both groups share similar perspectives on the various aspects of digital health adoption presented.

Table . Comparative statistical analysis between the study groups.

Question	Statistical test	*P* value	Significant difference
Use of electronic appointment system	Chi-square	.16	No
Future importance of digital health	Mann-Whitney *U*	.13	No
Use of telemedicine in the last 5 years	Mann-Whitney *U*	.86	No
Future use of telemedicine in the next 5 years	Mann-Whitney *U*	.10	No

Open-Text Questions

This section summarizes answers to the open-text questions, highlighting the most common answers and key aspects focusing on regulated digital medical products.

In Your Opinion, What Can Digitalization Achieve for Patients?

There is a strong emphasis on how digital health solutions can improve patient care and treatment efficiency, with many noting the potential for streamlined treatment processes, quicker and more accurate diagnoses, and a reduction in unnecessary medical tests. Centralized and accessible data management assumes another critical area. Respondents point out the benefits of easily retrievable patient information in an organized format, leading to more informed clinical decisions. This is closely linked with promoting preventive health care and healthy lifestyles, emphasizing the role of digitalization in disease prevention. Reducing the administrative burden on health care professionals is also a significant theme. Digital tools can allow professionals to focus more on direct patient care by minimizing time spent on nonmedical tasks and bureaucracy. Enhancing

XSL•FO
RenderX

communication and decision-making between patients and health care providers is viewed as crucial by study participants. Digital solutions are seen as facilitators of patient empowerment and shared decision-making, highlighting the importance of a patient-centered approach in digital health.

Regarding resource management, digitalization is seen as a potential solution to the shortage of health care professionals, supporting more personalized medicine and aiding in health care research through the provision of comprehensive data for analysis. Accessibility and inclusivity pose key themes, with digital health solutions considered vital for making health care more accessible, especially in remote or underserved areas. The responses collectively underscore the transformative potential of digital health solutions in enhancing health care delivery and patient experience.

Which Development Potential and Opportunities Do You See in the Digitalization of Health Care?

A recurring theme is the potential for digitalization to reduce bureaucracy and thus significantly improve the health care systems' efficiency. This aspect is particularly pertinent in the reduction of regulatory complexity associated with digital medical products. Enhanced data management and accessibility emerge as crucial factors, with many responses highlighting the integration of AI in diagnostics and decision-making, the use of big data for predictive health analytics, and the centralization of patient data. These points resonate with the emphasis on using real-world data and evidence for regulatory surveillance and the development of digital health solutions. Improved communication and coordination through digital tools were seen as crucial for better interdisciplinary and interinstitutional collaboration, reflecting the challenges of diverse regulatory landscapes in health care. The attitude toward innovation in medical devices and diagnostics, especially the role of AI and machine learning, the development of new diagnostic tools, and the application of digital twins and simulation techniques, aligns well with the need for reduced regulation of AI products in health care. Patient-centric approaches are emphasized in the responses, highlighting the need for more personalized medicine and improved patient engagement through digital solutions. This underlines the necessary development of useful, patient-centered, regulated medical products. Regulatory strategy considerations, particularly secure data handling and privacy concerns, are implicit in the responses, reflecting the importance of clear regulations in the development and implementation of digital medical products. Lastly, digitalization is viewed as a solution to health care professional shortages and a means to improve health care access, particularly in remote areas.

Where Do You See the Largest Risks and Disadvantages of the Digitalization of Health Care?

The responses reveal a multifaceted set of concerns, deeply interconnected with the challenges inherent in regulated digital medical products. A primary worry is data security and privacy, with numerous responses highlighting apprehensions about the misuse of personal health information. This emphasizes the need for stringent regulatory and ethical frameworks in the development and deployment of digital medical tools, as perceived by the study participants. Another significant concern is the depersonalization of care. Many participants fear that the rise of digital therapeutics and diagnostics might erode the personal interaction between health care providers and patients, a perceived cornerstone of adequate health care. Balancing technological efficiency with human empathy is, therefore, a key challenge in the realm of digital health. Technological dependency and system vulnerabilities emerge as major concerns. Respondents worry about an over-reliance on technology, potential system failures, and cybersecurity threats, emphasizing the importance of safety and reliability in digital medical products, especially in critical care. Resistance to change is another notable theme, particularly among health care professionals and older patients. This resistance can significantly impact the adoption and effectiveness of digital health solutions. Finally, navigating the complex regulatory landscapes and ethical dilemmas, especially with integrating AI and machine learning in health care, presents its challenges. Robust regulatory strategies and ethical guidelines are seen as paramount in ensuring the responsible and effective use of digital health innovations.

Participant Demography

Considering the age distribution of the participants, the following results were obtained from the questionnaire, as shown in Table 6.

Table . Age distribution between the 2 study groups. Values are shown as absolute numbers and relative frequencies (percentages in squares) of the whole respective population.

	18 - 25 years	26 - 35 years	36 - 45 years	46 - 55 years	56 - 65 years	>65 years
Physicians (n=271), n (%)	9 (3.32)	53 (19.56)	64 (23.62)	71 (26.20)	58 (21.40)	16 (5.90)
Patients (n=731), n (%)	35 (4.79)	166 (22.71)	209 (28.59)	155 (21.20)	111 (15.18)	55 (7.52)

Discussion

Principal Findings

This survey's results indicate alignment between professionals and patients regarding the adoption of technological innovations in health care and emphasize the necessity of improving digital health literacy. The continuous growth of digital health technologies is evident [10], with increasing use among professionals [11], the general population, and research [12], including real-world evidence generation [5]. The COVID-19 pandemic has accelerated digital adoption [13], transforming outdated processes and overcoming patient access barriers at

XSL•FO
RenderX

the “digital front door.” A balance between “techno-optimism” and “techno-skepticism” has emerged [14].

The study’s open-ended responses reflect concerns about data security, privacy, and ethical issues, with both physicians and patients identifying these as central challenges. The depersonalization of care and ethical dilemmas surrounding digital health innovations remain critical concerns. Systematic reviews confirm that digitalization introduces cybersecurity vulnerabilities, such as targeted cyberattacks and risks to patient privacy. Physicians and patients alike perceive these as barriers to effective implementation [15]. Additional studies highlight the ethical dimensions of digital health, including impacts on the patient-physician relationship, informed consent, and potential workforce shifts. Ongoing research is required to ensure that technological advancements do not compromise patient rights or exacerbate health care inequalities [16].

While many individuals rely on the internet as a primary source of health information and for anonymous consultations, skepticism persists regarding web-based data storage and digital data exchange between health care professionals and patients. Research has shown that, while internet users frequently seek health information on the web, they express doubts about the reliability of such information and the security of electronic health data exchange. Nonetheless, positive developments are also noted: physicians report improvements in access to digital health records compared to previous studies [17].

Despite advancements in digital health applications, integrating patient-generated data presents challenges, particularly due to technological limitations such as poor interoperability and data overload. Experts have emphasized the need for regulatory frameworks, ethical guidelines, and improved data governance to facilitate better integration [18]. Additionally, while protective regulations are necessary to ensure data security, the COVID-19 pandemic demonstrated how excessive regulatory barriers contributed to delays in data availability for policymakers. This underscores the need for streamlined data approval processes while maintaining public trust [19].

Germany has introduced the “Law for the Acceleration of Digitization in Healthcare” (DigiG), currently under legislative review. This initiative aims to implement electronic patient case files, digital medication overviews, electronic medical prescriptions, increased use of digital health applications (DiGAs), assisted telemedicine, and a digital medical council comprising experts in data security, information freedom, and information technology [20]. The governing legal frameworks, particularly the European Union’s Medical Device Regulation, impose comprehensive safety, performance, and postmarket surveillance requirements for digital health applications, ensuring high reliability. However, these rigorous conformity assessments may slow innovation. In contrast, the US Food and Drug Administration framework offers a more expedited approval process for digital health technologies, including AI-driven tools, through mechanisms such as the Software as a Medical Device (SaMD) categorization [21]. This regulatory flexibility has facilitated the rapid adoption of AI in US health care [22].

Research has identified various barriers to digital health adoption, including technical difficulties, resistance to change, costs, reimbursement policies, patient age, and education levels. These findings remain relevant, with technological proficiency, perceived usefulness, and accessibility influencing acceptance worldwide. Additionally, interdisciplinary collaboration, cultural factors, and leadership have been recognized as key enablers of digital health implementation [23,24]. Digital booking systems offer potential benefits, but patient and physician attitudes toward these systems remain understudied [25,26]. Our study demonstrates growing interest in digital health solutions and highlights the necessity for a transparent and innovation-friendly regulatory framework.

Age significantly impacts digital health adoption, as older adults face barriers such as limited digital literacy and reluctance toward technological change [27,28]. These challenges must be addressed through targeted interventions, including comprehensive training programs and the development of user-friendly, age-adapted digital health solutions. Integrating digital health education into medical school curricula is critical to preparing future physicians for evolving technological demands, reducing hesitation, and fostering a legal environment conducive to innovation.

Telemedicine adoption requires specific communication skills to maintain strong physician-patient relationships. A systematic review performed by Ritunga et al [29] highlights that effective digital consultations depend on training that addresses the unique challenges of virtual interactions, such as nonverbal communication and empathy. Additionally, as society shifts toward an information-driven model, medical education must balance humanistic aspects, such as the physician-patient relationship, with digital health integration. Experts warn that without proper integration, medical care quality may decline [30]. These concerns align with this study’s findings, reinforcing the importance of digital health training for both professionals and patients.

Legislative changes, particularly at the European Union level, are essential to facilitate health care’s transition from analog systems to modern digital solutions while ensuring privacy protections [31]. “Careful legal consideration has to be applied when considering privacy regulations (…), making every case an individual one,” keeping the patient at the center not only legally but also ethically, reinforcing the importance of individualized health care approaches.

The increasing alignment between physicians and patients regarding digital health adoption has significant implications for future health care delivery. With both groups recognizing the value of digital health tools and telemedicine, integration into routine practice may become more seamless. This shared understanding could reduce resistance, encourage collaboration in designing patient-centered digital solutions, and enhance accessibility, efficiency, and convenience in health care services.

Further research will enable a deeper understanding of drivers and obstacles in digital health advancements. The underlying research deserves a follow-up including a more international study population in the next few years to understand and re-evaluate the described trend patterns on a timeline,

XSL•FO
RenderX

particularly in view of the more and more rapid changes in digital health.

Limitations

Due to the nature of its origin and the involved researchers, our survey has reached mainly answers from German-speaking individuals (94.5% (947/1002) of the survey's participants). The survey distribution via digital means causes a bias toward respondents who are open to digital solutions. Persons unwilling or unable to complete a web-based questionnaire were not considered for technical reasons, causing a systematic bias that might potentially be a limiting factor regarding the generalizability of the presented results. Another limiting factor might be the above-mentioned age demographics of the study: as the majority of study participants belonged to the age groups between 26 and 65 years, the study results might lean toward a comparably high digital literacy among the respondents. Additionally, the German or German-speaking study population presents a limitation with regard to the conclusions drawn for a comparison between physicians and patients worldwide. Overall, 271 professionals versus 731 patients is a logical outcome, as all people are patients in general, whereas a smaller percentage counts as medical professionals.

Future research could benefit from exploring subgroup differences, such as attitudes by age, gender, system access, or professional experience, as well as including other health care professionals like nurses, to provide a more nuanced understanding of digital health acceptance. Practical challenges related to accessing individual nurses in this study will be addressed, and a deeper analysis of qualitative responses based on demographic factors could provide additional insights into specific barriers and facilitators.

Outlook

As also suggested by the replies of the study groups, regulatory aspects must be modified to keep pace with current innovations, allowing for meaningful evaluation, regulation, and dysregulation of new technologies. This aspect is in line with the results of Torous et al [32], who highlighted this urgent need for a steady modification of regulations pertinent to the field of digitalization in health care.

Web-based data availability can lead to misinformation, which needs to be vigorously and continuously counteracted: "To harness the full potential of digital media to support health and well-being as well as to mitigate or counteract the effects of mis- and disinformation, three fundamental skills should be continuously developed: digital literacy, health literacy, and digital health literacy" [33].

In general, perceptions, fears, and resentments must be systematically eliminated to catalyze digital health acceptance and allow for further integration and innovative solutions to benefit patients and professionals. Bureaucratic burdens and technical obstacles must be overcome and transformed from technical barriers into enablers to accelerate acceptance and usage.

Recent developments such as the Data for Health Conference 2023 (DFH23) serve as catalysts for optimistic growth: "With the momentum #DFH23 created, the European Health Data Space (EHDS) as a solid and safe foundation for consented collaborative health data use and the G7 Hiroshima AI process in place, we call on citizens and their governments to fully support the digital transformation of medicine, research and innovation including AI" [34].

Advancements in digital health literacy and enhanced technical reliability will drive openness and catalyze future developments in the patient's interest and, thus, in the interest of sustainable innovation-driven health care systems.

Authors' Contributions

FS, FH, FV, and DP designed the study. FS, FH, and DP conducted the data preprocessing. FH performed the statistical analysis. FS and DP reviewed the statistical results. FS, FH, and DP wrote the manuscript. FS, FH, FV, GME, KFB, RvER, PB, and DP contributed critical revisions to the manuscript and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Questionnaire.
[DOCX File, 24 KB - humanfactors_v12i1e60779_app1.docx]

References

1. Fatehi F, Samadbeik M, Kazemi A. What is digital health? Review of definitions. Stud Health Technol Inform 2020 Nov 23;275:67-71. [doi: 10.3233/SHTI200696] [Medline: 33227742]
2. Li E, Clarke J, Ashrafian H, Darzi A, Neves AL. The impact of electronic health record interoperability on safety and quality of care in high-income countries: systematic review. J Med Internet Res 2022 Sep 15;24(9):e38144. [doi: 10.2196/38144] [Medline: 36107486]
3. Longhini J, Rossettini G, Palese A. Digital health competencies among health care professionals: systematic review. J Med Internet Res 2022 Aug 18;24(8):e36414. [doi: 10.2196/36414] [Medline: 35980735]

XSL•FO
RenderX

4. Furtner D, Shinde SP, Singh M, Wong CH, Setia S. Digital transformation in medical affairs sparked by the pandemic: insights and learnings from COVID-19 era and beyond. Pharmaceut Med 2022 Feb;36(1):1-10. [doi: 10.1007/s40290-021-00412-w] [Medline: 34970723]
5. Stern AD, Brönneke J, Debatin JF, et al. Advancing digital health applications: priorities for innovation in real-world evidence generation. Lancet Digit Health 2022 Mar;4(3):e200-e206. [doi: 10.1016/S2589-7500(21)00292-2] [Medline: 35216754]
6. Abernethy A, Adams L, Barrett M, et al. The promise of digital health: then, now, and the future. NAM Perspect 2022 Jun 27. [doi: 10.31478/202206e] [Medline: 36177208]
7. Schlieter H, Marsch LA, Whitehouse D, et al. Scale-up of digital innovations in health care: expert commentary on enablers and barriers. J Med Internet Res 2022 Mar 11;24(3):e24582. [doi: 10.2196/24582] [Medline: 35275065]
8. Vogt F, Seidl F, Santarpino G, et al. Healthcare IT utilization and penetration among physicians: novel IT solutions in healthcare—use and acceptance in hospitals. Eur Surg Res 2018;59(1-2):100-113. [doi: 10.1159/000490241] [Medline: 30048992]
9. SoSci Survey. URL: https://www.soscisurvey.de [accessed 2025-04-07]
10. Šendelj R. Information technology and information management in healthcare. Stud Health Technol Inform 2020 Sep 25;274:139-158. [doi: 10.3233/SHTI200674] [Medline: 32990671]
11. Sheikh A, Anderson M, Albala S, et al. Health information technology and digital innovation for national learning health and care systems. Lancet Digit Health 2021 Jun;3(6):e383-e396. [doi: 10.1016/S2589-7500(21)00005-4] [Medline: 33967002]
12. Jandoo T. WHO guidance for digital health: what it means for researchers. Digit Health 2020;6. [doi: 10.1177/2055207619898984] [Medline: 31949918]
13. Shah IA, Jhanjhi NZ, Humayun M, Ghosh U. Health care digital revolution during COVID-19. In: Anandan R, Suseendran G, Chatterjee P, Jhanjhi NZ, Ghosh U, editors. How COVID-19 Is Accelerating the Digital Revolution: Challenges and Opportunities: Springer; 2022:17-30. [doi: 10.1007/978-3-030-98167-9_2]
14. Rodriguez JA, Lyles CR. Strengthening digital health equity by balancing techno-optimism and techno-skepticism through implementation science. NPJ Digit Med 2023 Nov 2;6(1):203. [doi: 10.1038/s41746-023-00954-0] [Medline: 37919418]
15. Nifakos S, Chandramouli K, Nikolaou CK, et al. Influence of human factors on cyber security within healthcare organisations: a systematic review. Sensors 2021 Jul 28;21(15):5119. [doi: 10.3390/s21155119] [Medline: 34372354]
16. Ursin F, Timmermann C, Benzinger L, Salloch S, Tietze FA. Intraoperative application of mixed and augmented reality for digital surgery: a systematic review of ethical issues. Front Surg 2024;11:1287218. [doi: 10.3389/fsurg.2024.1287218] [Medline: 38550794]
17. Haluza D, Naszay M, Stockinger A, Jungwirth D. Digital natives versus digital immigrants: influence of online health information seeking on the doctor-patient relationship. Health Commun 2017 Nov;32(11):1342-1349. [doi: 10.1080/10410236.2016.1220044] [Medline: 27710132]
18. Cresswell K, McKinstry B, Wolters M, Shah A, Sheikh A. Five key strategic priorities of integrating patient generated health data into United Kingdom electronic health records. J Innov Health Inform 2019 Jan 4;25(4):254-259. [doi: 10.14236/jhi.v25i4.1068] [Medline: 30672407]
19. Cavallaro F, Lugg-Widger F, Cannings-John R, Harron K. Reducing barriers to data access for research in the public interest—lessons from covid-19. BMJ Opinion. 2020 Jul 6. URL: https://blogs.bmj.com/bmj/2020/07/06/reducing-barriers-to-data-access-for-research-in-the-public-interest-lessons-from-covid-19/ [accessed 2025-04-07]
20. Gesetz zur Beschleunigung der Digitalisierung des Gesundheitswesens (Digital-Gesetz – DigiG) [Web page in German]. Bundesministerium für Gesundheit. 2023. URL: https://www.bundesgesundheitsministerium.de/ministerium/gesetze-und-verordnungen/guv-20-lp/digig [accessed 2025-04-07]
21. Gelis L, Stoeckert I, Podhaisky HP. Digital tools-regulatory considerations for application in clinical trials. Ther Innov Regul Sci 2023 Jul;57(4):769-782. [doi: 10.1007/s43441-023-00535-z] [Medline: 37195515]
22. Schwendicke F, Chaurasia A, Wiegand T, et al. Artificial intelligence for oral and dental healthcare: core education curriculum. J Dent 2023 Jan;128:104363. [doi: 10.1016/j.jdent.2022.104363] [Medline: 36410581]
23. Borges do Nascimento IJ, Abdulazeem H, Vasanthan LT, et al. Barriers and facilitators to utilizing digital health technologies by healthcare professionals. NPJ Digit Med 2023;6(1):161. [doi: 10.1038/s41746-023-00899-4]
24. Scott Kruse C, Karem P, Shifflett K, Vegi L, Ravi K, Brooks M. Evaluating barriers to adopting telemedicine worldwide: a systematic review. J Telemed Telecare 2018 Jan;24(1):4-12. [doi: 10.1177/1357633X16674087] [Medline: 29320966]
25. Motulsky A, Bosson-Rieutort D, Usher S, et al. Evaluation of a national e-booking system for medical consultation in primary care in a universal health system. Health Policy 2023 May;131:104759. [doi: 10.1016/j.healthpol.2023.104759] [Medline: 36907137]
26. Paré G, Raymond L, Castonguay A, Grenier Ouimet A, Trudel MC. Assimilation of medical appointment scheduling systems and their impact on the accessibility of primary care: mixed methods study. JMIR Med Inform 2021 Nov 16;9(11):e30485. [doi: 10.2196/30485] [Medline: 34783670]
27. Bakogiannis C, Tsarouchas A, Mouselimis D, et al. A patient-oriented app (ThessHF) to improve self-care quality in heart failure: from evidence-based design to pilot study. JMIR Mhealth Uhealth 2021 Apr 13;9(4):e24271. [doi: 10.2196/24271] [Medline: 33847599]

XSL•FO
RenderX

28. Herrmann M, Boehme P, Hansen A, et al. Digital competencies and attitudes toward digital adherence solutions among elderly patients treated with novel anticoagulants: qualitative study. J Med Internet Res 2020 Jan 24;22(1):e13077. [doi: 10.2196/13077] [Medline: 32012049]
29. Ritunga I, Claramita M, Widaty S, Soebono H. Challenges and recommendations in the implementation of audiovisual telemedicine communication: a systematic review. Korean J Med Educ 2024 Sep;36(3):315-326. [doi: 10.3946/kjme.2024.305] [Medline: 39246112]
30. Wartman SA. Medicine, machines, and medical education. Acad Med 2021 Jul 1;96(7):947-950. [doi: 10.1097/ACM.0000000000004113] [Medline: 33788788]
31. Putzier M, Khakzad T, Dreischarf M, Thun S, Trautwein F, Taheri N. Implementation of cloud computing in the German healthcare system. NPJ Digit Med 2024 Jan 13;7(1):12. [doi: 10.1038/s41746-024-01000-3] [Medline: 38218892]
32. Torous J, Stern AD, Bourgeois FT. Regulatory considerations to keep pace with innovation in digital health products. NPJ Digit Med 2022 Aug 19;5(1):121. [doi: 10.1038/s41746-022-00668-9] [Medline: 35986056]
33. van Kessel R, Wong BLH, Clemens T, Brand H. Digital health literacy as a super determinant of health: more than simply the sum of its parts. Internet Interv 2022 Mar;27:100500. [doi: 10.1016/j.invent.2022.100500] [Medline: 35242586]
34. Baumgart DC. An intriguing vision for transatlantic collaborative health data use and artificial intelligence development. NPJ Digit Med 2024 Jan 23;7(1):19. [doi: 10.1038/s41746-024-01005-y] [Medline: 38263436]

Abbreviations

AI: artificial intelligence
DFH23: Data for Health Conference 2023
DiGA: Digitale Gesundheitsanwendungen
mHealth: mobile health
SaMD: Software as a Medical Device

Edited by A Kushniruk; submitted 21.05.24; peer-reviewed by A Yazdanian, D Singh, E Nelson, H Oh; revised version received 17.02.25; accepted 18.02.25; published 08.05.25.

Please cite as:
Seidl F, Hinterwimmer F, Vogt F, Edenharter GM, Braun KF, von Eisenhart-Rothe R, AG Digitalisierung der DGOU, Biberthaler P, Pförringer D
Use and Acceptance of Innovative Digital Health Solutions Among Patients and Professionals: Survey Study
JMIR Hum Factors 2025;12:e60779
URL: https://humanfactors.jmir.org/2025/1/e60779
doi:10.2196/60779

© Fritz Seidl, Florian Hinterwimmer, Ferdinand Vogt, Günther M Edenharter, Karl F Braun, Rüdiger von Eisenhart-Rothe, AG Digitalisierung der DGOU, Peter Biberthaler, Dominik Pförringer. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 8.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Public Attitudes and Predictors of Public Awareness of Personal Digital Health Data Sharing for Research: Cross-Sectional Study in Japan

Yasue Fukuda[1*], PhD, Prof Dr; Koji Fukuda[2*], PhD, Prof Dr

[1]Department of Pharmacy, Suzuka University of Medical Science, 3500-3 Minami-tamagaki, Suzuka, Mie, Japan

[2]Department of Political Science and Economics, Faculty of Political Science and Economics, Waseda University, Shinjuku-ku, Tokyo, Japan

[*]all authors contributed equally

Corresponding Author:
Yasue Fukuda, PhD, Prof Dr
Department of Pharmacy, Suzuka University of Medical Science, 3500-3 Minami-tamagaki, Suzuka, Mie, Japan

Abstract

Background: As digital technology advances, health-related data can be scientifically analyzed to predict illnesses. The analysis of international data collected during health examinations and health status monitoring, along with data collected during medical care delivery, can contribute to precision medicine and the public good. Understanding citizens' attitudes and predictors of digital health data sharing is critical in promoting data-driven research.

Objective: This study aims to determine the public acceptability of data sharing and the attitudes and influencing factors toward data sharing.

Methods: A cross-sectional web-based survey was conducted in Japan from November 11 - 18, 2023. We analyzed 1000 valid responses. Five factors were investigated as predictors of participants' attitudes toward sharing digital health data for social benefit: (1) individual sociodemographic characteristics, (2) types of health data shared, (3) motivation for sharing data, (4) data sharing concerns, and (5) reasonable access and control over the data. The association of these factors with the respondents' willingness to share was analyzed. We summarized demographic characteristics based on gender, age group, affiliated educational institution, and education history and degree. Continuous variables are expressed as mean (SE). Logistic regression was used to analyze the association between attitudes and acceptability of sharing digital health data and the predicting factors, such as participants' preferences regarding data access and control, underlying concerns, motivations for data sharing, demographic characteristics, and eHealth literacy.

Results: The mean age of the participants and the SD was 52.8 (19.8) years. We identified the factors influencing respondents' willingness to share a wide range of personal digital health data in Japan, including data in medical records, biobank samples, and digitized social communication. Approximately 70% of the participants were willing to share their digital health data. The motives associated with positive willingness to share digital health data were helping future patients (odds ratio [OR] 2.5860, 95% CI 1.8849 - 3.5481; *P*<.001), receiving their own results (OR 2.2261, 95% CI 1.6243 - 3.0509; *P*<.001), and receiving financial benefits (OR 1.8059, 95% CI 1.2630 - 2.5822; *P*=.001). Concerns associated with negative willingness to share data were data being used for unethical projects (OR 0.5104, 95% CI 0.5104 - 0.722; *P*<.001) and agreeing to contract terms that they did not fully understand (OR 0.7114, 95% CI 0.5228 - 0.9681; *P*=.04). Compared with men, women were less willing to share data (OR 0.722, 95% CI 0.539 - 0.967; *P*=.03). Furthermore, the higher one's eHealth literacy, the more positive their willingness to share digital health data (OR 1.0680, 95% CI 1.0450 - 1.0920; *P*<.001).

Conclusions: This study found differences in the types of data people are willing to share. Therefore, the significance of sharing data should be fully communicated to people to motivate them to share their data and contribute to their overall health.

(JMIR Hum Factors 2025;12:e64192) doi:10.2196/64192

KEYWORDS

digital health data; cross-sectional study; predict factor; Japan; willingness to share personal data; eHealth literacy

Introduction

The progression of digital technology has enabled the comprehensive analysis of health-related data, facilitating the prediction of potential illnesses [1-3]. The exchange of personal digital health data plays a crucial role in shaping policies that enhance digital health initiatives and in the innovation of new treatment and prevention strategies [4]. Personal health data includes medical records and data from medical examinations

XSL•FO
RenderX

[5]. Data from biological samples, such as blood and DNA, are also important types of health data [6]. In addition, family health data, exercise habits, food intake, and lifestyle data are closely related to health [7-9]. The introduction of eHealth records and the development of wearable apps have led to health data digitization [10,11]. Digital health data on heart rate and blood pressure can be collected by mobile devices in several ways. Digital health data, such as social activities, behavior patterns, and location data are useful for the morbidity risk, prevention, and diagnosis of infectious diseases, including COVID-19 and mental health. Other examples of digital health data include travel records made during the COVID-19 pandemic, web-based data on travel locations, and digital health data on social activities [12-14].

Health data can be collected to contribute to the public good when shared, including data collected during health examinations and health status monitoring, as well as data collected during the delivery of medical care. Therefore, biobanks, which collect individual biological samples and use them for public benefit, and databanks, which hold electronic medical records, are being developed [15-18].

As data science develops, it will likely be used for personalized medicine and understanding the social determinants of health. In Japan, the Ministry of Health, Labor, and Welfare has formulated the “Medical Digital Transformation: Vision 2030,” a strategy for promoting medical digital transformation that actively supports the usage of health and medical care data [19]. In Europe, the General Personal Data Protection Regulation governs the use of personal data, including patients’ electronic medical records. This is because informed consent is required for personal data collection and sharing, including secondary use for research purposes.

Digitizing health data and enabling its international sharing are crucial steps in improving health care worldwide. By making medical records accessible across borders, patients can receive accurate diagnoses and appropriate treatments regardless of their location. This is particularly beneficial for travelers and expatriates, as well as people in emergencies, since doctors can quickly access their medical history and provide suitable care without unnecessary delays [20]. In addition to benefiting individual patients, the international sharing of health data plays a significant role in disease prevention and control. By monitoring global health trends, medical institutions and governments can detect outbreaks at an early stage, respond quickly, and implement effective containment strategies. The importance of such efforts was especially evident during the COVID-19 pandemic, where data sharing initiatives helped track the spread of the virus and accelerate vaccine development [21].

The digitization and sharing of health data also benefit medical research. Access to large-scale anonymized datasets allows researchers to develop innovative treatments, improve AI-driven diagnostics, and advance personalized medicine [22]. Moreover, standardized digital records contribute to the efficiency of health care systems by reducing administrative burdens and facilitating better coordination between medical professionals, ultimately leading to improved patient outcomes [23].

Furthermore, sharing health data internationally helps bridge the gap between developed and developing regions, fostering collaboration that promotes equitable health care access and reduces global health disparities [24].

Rare diseases provide a salient example of the benefits of international data sharing. Definitions of rare diseases differ by region: in the United States, a rare disease is defined as one affecting fewer than 200,000 patients nationwide [25]; in Japan, fewer than 50,000 patients nationwide [26]; and in the European Union, conditions affecting fewer than 1 in 2000 people in the general population [27]. Due to the small number of patients in each country, there is limited availability of diagnostic, genetic, and other health data, as well as insufficient specialized knowledge. Consequently, diagnoses may be delayed or missed entirely, hindering the development of effective treatment and prevention strategies. However, international sharing of health data can overcome these limitations by enabling accurate diagnosis and facilitating research on therapies and preventive measures [28].

Despite the many advantages of digitizing and sharing health data, it is essential to address privacy, security, and ethical concerns. Strong data protection measures and international regulations must be in place to ensure that personal health information is shared responsibly and securely [29]. By balancing the benefits of accessibility with the need for confidentiality, the global health care community can create a system that maximizes both innovation and patient protection.

However, digital health data sharing may not be fully understood without examining citizens’ attitudes toward data sharing related to health and medical care [30]. While health data digitization and sharing can facilitate research and create social benefits, public acceptability may be a barrier to its successful implementation [30]. Examining citizens’ attitudes toward and acceptance of sharing digital health data may be useful in formulating policies and guidelines for data management [31]. This can help in the development and implementation of policies to promote medical health research through digital health personal data. Such efforts benefit others through secondary use of large-scale data, drug development, verification of adverse effects, and reflection on health care policy, rather than for personal, individual medical treatment purposes. Existing studies have considered the acceptability of biobanks, eHealth data, and the acceptability of and attitudes toward digital health data [32-35]. Citizens’ attitudes and receptivity to digital health data sharing differ depending on the type of data being shared, and concerns regarding personal data and transparency issues in data sharing may affect data sharing. However, the types of data that can be digitized and contribute to health include not only biobanks and comprehensive medical records but also lifestyle and behavioral habits.

Few studies have been conducted to elucidate the factors that predict willingness to share digital health data. One study considered the attitudes toward the willingness of young people in Brazil and Denmark to register their digital health data [36]. Studies in Australia [37] and Europe [38] have found that citizens prefer sharing health data under government oversight. They express concerns about sharing anonymized health data

XSL•FO
RenderX

for clinical practice or research purposes unless clear provisions are made regarding its intended purpose, limitations on use, and access restrictions [36-38]. Broader citizen consultation on data use, suggesting this trend, has largely been overlooked [39]. However, there is growing evidence of public support for health data sharing [40,41]. Studies indicate that citizens in Australia [42], the United Kingdom, and the United States [43] are willing to share data aimed at improving health outcomes. The determinants of this willingness remain ambiguous, with limited research addressing issues specific to the Asian region, including Japan. This region, with its diverse cultures, social backgrounds, and varying levels of digitalization, may present different predictors of willingness to share digital health data. A comprehensive understanding of the factors influencing citizens' willingness to share health data can help identify challenges and opportunities for establishing data collection systems, thereby improving communication and fostering citizen trust in digital health data.

Methods

Study Design and Setting

This cross-sectional study was conducted via a web-based questionnaire survey administered across Japan by Cross Marketing Inc, an internet research firm. At the beginning of the questionnaire, participants were clearly informed that their personal data would be strictly protected and anonymized. Furthermore, they were explicitly instructed to respond to questions about their attitudes and intentions toward data sharing on the premise that any shared data would exclude personally identifiable data and privacy would be guaranteed. This explanation was included to ensure participants understood that the protection of personal data was a fundamental prerequisite for data sharing scenarios described in the questionnaire.

The survey was developed in accordance with the CHERRIES (Checklist for Reporting Results of Internet E-Surveys) checklist (Figure 1,Checklist 1). A pilot test was conducted between October 28 and November 6, 2023, to assess the questionnaire's clarity, readability, interface usability, and time burden. The final survey was distributed from November 11 to 18, 2023.

XSL•FO
RenderX

Figure 1. Checklist for Reporting Results of Internet E-Surveys flow chart.

XSL•FO
RenderX

Study Size

A minimum sample size of 385 was calculated using a normal approximation of the binomial distribution, considering Japan's population of 126 million, a 95% confidence level, and a 5% margin of error [44]. Based on previous studies, 1000 questionnaires were collected to improve the validity of the study [36,38,40].

Participants and Sampling

Participants were recruited from Cross Marketing Inc.'s panel and were aged 18 years or older. To ensure demographic representativeness, stratified random sampling was used based on age and gender. Participants were equally allocated into seven age categories (18 - 27, 28 - 37, 38 - 47, 48 - 57, 58 - 67, 68 - 77, and 78 years or older) and by gender, resulting in equal numbers in each age category and gender group. Within each stratum, participants were randomly selected.

Exclusion criteria included discrepancies between registered and self-reported gender or age, extremely short response times, failure to follow instructions, or provision of inconsistent or implausible responses. Out of 23,434 distributed questionnaires, 2203 participants provided informed consent, and after applying exclusion criteria, 1000 valid responses were analyzed. Participant demographics are presented in the Results section (Table 1).

XSL•FO
RenderX

Table . Participant characteristics (N=1000).

Characteristic	Participants, n (%)
Gender	
Men	500 (50)
Women	500 (50)
Age (years)	
18 - 27	142 (14.2)
28 - 37	144 (14.4)
38 - 47	144 (14.4)
48 - 57	144 (14.4)
58 - 67	142 (14.2)
68 - 77	142 (14.2)
>78	142 (14.2)
Annual household income (Japanese Yen[a])	
<3 million	196 (19.6)
3 million–5 million	238 (23.8)
5 million–8 million	203 (20.3)
8 million–10 million	92 (9.2)
>10 million	99 (9.9)
I don’t want to answer	172 (17.2)
Education	
Junior high school	32 (3.2)
High school	285 (28.5)
Vocational school or junior college	160 (16.0)
Bachelor’s degree or higher	503 (50.3)
I don’t want to answer	20 (2.0)
Marital status	
Unmarried	336 (33.6)
Married	655 (65.5)
I don’t want to answer	9 (0.9)
Owning a mobile phone (smartphone)	
Yes	940 (94.0)
Owning a personal computer	
Yes	826 (82.6)
Owning a tablet	
Yes	295 (29.5)
Owning a smartwatch	
Yes	110 (11.0)
Owning smart home assistant devices and AI[b] home appliances	
Yes	65 (6.5)
Previously participated in health or clinical research	
Yes	57 (5.7)
Health status	

XSL•FO
RenderX

Characteristic	Participants, n (%)
Bad	27 (2.7)
Somewhat bad	156 (15.6)
Neutral	263 (26.3)
Somewhat good	383 (38.3)
Very good	168 (16.8)
Concerned about own and family's health	
Not concerned at all	24 (2.4)
Not very concerned	45 (4.5)
Neutral	108 (10.8)
Somewhat concerned	400 (40)
Very concerned	415 (41.5)
I don't know	8 (0.8)

[a]US$1=147 Japanese Yen (Stand: September 2025).

[b]AI: artificial intelligence.

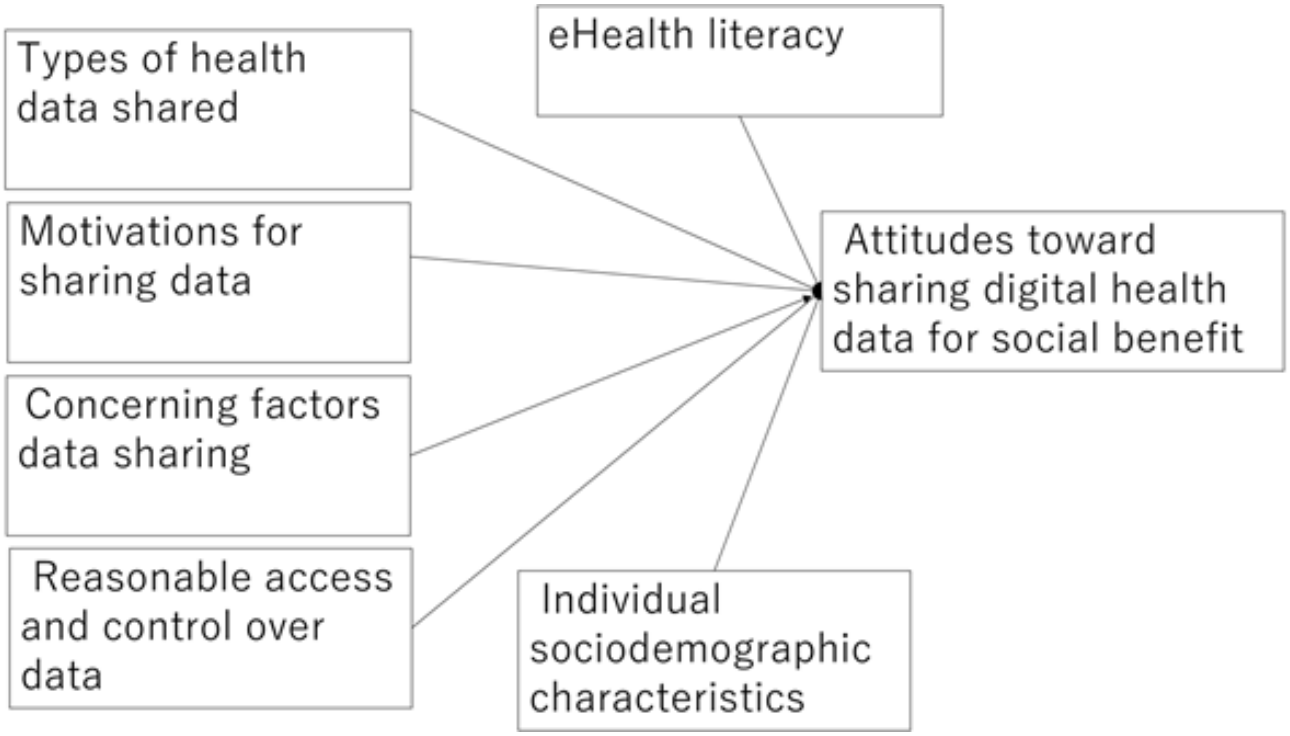
Survey Measures

The questionnaire was developed with reference to previous studies on willingness to share health-related data, including biobank and electronic medical record data, and attitudes toward data repositories. Five factors were investigated as predictors of participants' attitudes toward sharing digital health data for social benefit: (1) individual sociodemographic characteristics, (2) types of health data shared, (3) motivation for sharing data, (4) concerns factors of data sharing, and (5) reasonable access and control over the data (Figure 2) [36-38,45,45]. A pilot study involving 5 academic staff and researchers helped refine the instrument. The selection and formulation of questionnaire items were based on previous studies identifying factors influencing data sharing behavior.

The questionnaire included the following domains, all using uniform response options ("Yes," "No," "Don't know," and "Other" with free-text input). Questions were randomized, and multiple responses were allowed.

Full wording of all items and response options is available in Multimedia Appendix 1 questionnaire.

Figure 2. Predictors of public attitude, personal health digital data for social benefit research model.



XSL•FO
RenderX

Willingness to Share Digital Health Data

Participants indicated their willingness to share various types of digital health data, including physical and mental clinical diagnoses, family health history, genetic and blood test data, dietary and nutritional data, alcohol consumption, sleep status, social communication (frequency and content), physical activity, stress, and geolocation data.

Motivation for Sharing Data

Participants selected motivations such as helping future patients, supporting research, receiving study results, financial incentives, and proposing research topics.

Concerns About Sharing Digital Health Data

Participants selected concerns from a list including privacy risks, data misuse, unauthorized access, and ambiguity in consent terms. The wording related to access restrictions was revised for clarity. Participants chose from options indicating who may access their data (eg, medical professionals, researchers, and family members) and preferences on data control.

Attitudes Toward Access Restrictions

Participants reported preferences on controlling access to their data, transparency in data usage, trust in data controllers, and openness to sharing with public or academic institutions.

Demographics and Health-Related Factors

Data on age, gender, educational attainment, annual household income, marital status, and ownership of digital devices (smartphones, PCs, tablets, smartwatches, and smart home assistants or artificial intelligence [AI] appliances) were collected using predefined categories. Participants also reported previous participation in health research (epidemiological or clinical trials).

eHealth literacy was assessed using the Norman eHealth Literacy Scale (eHEALS), an 8-item, 5-point Likert scale measuring knowledge, comfort, and perceived skills in finding, evaluating, and applying eHealth data [45,46].

Self-rated health status was assessed on a 5-point Likert scale, and responses were grouped into “good” (somewhat good, very good) or “not-so-good” for analysis. Logistic regression analyzed the relationship between health status and willingness to share data.

Participants’ concern about their own and family’s health was measured on a 6-point scale, dichotomized into “interested” or “not interested,” with logistic regression used to assess associations with willingness to share digital health data.

Statistical Analysis

Descriptive statistics (frequencies, means, and SDs) summarized participant characteristics. We used the intention to share digital health data as the outcome variable. Logistic regression analysis was performed to examine its associations with a range of predictor variables, including motivations for sharing data (eg, helping patients and supporting research), concerns about data sharing (eg, privacy risks and misuse), attitudes toward data access (eg, preferences for control and trust in institutions), and sociodemographic and health-related factors (eg, age, gender, education, income, marital status, digital device ownership, and previous participation in health research). Analyses were conducted using SPSS version 28 (IBM Corp), with significance set at $P<.05$.

Ethical Considerations

This study fully complied with the Declaration of Helsinki and Japan’s Personal Data Protection Law [47]. Ethical approval was granted by the Waseda University Ethics Committee (application 2023 - 250), confirming that all protocols and consent procedures were reviewed and sanctioned before data collection.. No personally identifiable information was collected in accordance with both institutional and legal privacy requirements. All persons described are not identifiable and cannot be identified through the details of the story because this research is an anonymous Internet survey.

Informed consent was obtained electronically before participation. Participants received a small monetary incentive (points equivalent to approximately 10 Japanese yen [US$1=147.1 JPY]) via the survey provider’s point system. Low-quality or inconsistent responses were screened and excluded by the research firm.

Results

Characteristics and Attributes of Participants

Table 1 shows participant characteristics. The men-to-women participant ratio is 1:1. The mean age of the participants and the SD is 52.8 (SD 19.8) years. Approximately half of the participants had a university or postgraduate degree in education. In total, 93% of the respondents had a smartphone and 83% had a personal computer. About 80% of the respondents were interested in their health (very interested and somewhat interested).

Association Between Willingness to Share Digital Health Data and Sociodemographic Characteristics

Approximately 70% of respondents were willing to share digital health data. Respondents who were willing to share digital health data were defined as having a positive willingness to share, and subsequent analyses were conducted.

The results of the logistic analysis of the association between positive attitudes toward sharing digital health data and sociodemographic characteristics are shown in Table 2. Compared with men, women were less willing to share data (odds Ratio [OR] 0.722, 95% CI 0.539 - 0.967; $P<.05$).

Those who had more experience participating in research (OR 3.308, 95% CI 1.391 - 7.868; $P<.01$), and had awareness of research data repositories (OR 2.711, 95% CI 1.364 - 5.388; $P<.01$) were more willing to share digital health data. Moreover, owning a smartphone (OR 1.841, 95% CI 1.078 - 3.145; $P<.05$), personal computer (OR 1.658, 95% CI 1.175 - 2.341; $P<.05$), tablet (OR 1.688, 95% CI 1.212 - 2.349; $P<.01$), or smart home assistance appliances (OR 3.384, 95% CI 1.077 - 5.276; $P<.05$) was positively associated with higher willingness to share digital health data. Conversely, there was no association between annual income or education and the willingness to share digital

XSL•FO
RenderX

health data. Logistic analysis of e-health literacy and digital health data sharing suggested that the higher the e-health literacy, the more positive the willingness to share digital health data.

Willingness to share digital health data was significantly associated with a high level of health concern and willingness to share digital health data (OR 2.1871, 95% CI 1.5750 - 3.03710; $P<.001$) but not associated with health status ($P=.73$).

XSL•FO
RenderX

Table . Logistic regression results regarding a positive willingness to share personal data and sociodemographic characteristics.

Sociodemographic characteristics	B (partial regression coefficient)	SE	*P* value	OR[a] (95% CI)
Gender (reference: man)				
Woman	-0.3259	0.1490	.03[b]	0.7219 (0.5391-0.9667)
Age (reference 18 - 27 years)				
28 - 37	0.3346	0.2679	.21	1.3974 (0.8265-2.3625)
38 - 47	0.4610	0.2737	.19	1.5856 (0.9272-2.7115)
48 - 57	0.2681	0.2663	.31	1.3075 (0.7758-2.2035)
58 - 67	-0.1268	0.2612	.63	0.8809 (0.5280-1.4698)
68 - 77	0.1563	0.2697	.57	1.1692 (0.6892-1.9835)
>78	0.0073	0.2628	.98	1.0074 (0.6019-1.6860)
Annual household income (reference:<3 million Japanese Yen[c])				
3 million–5 million	0.2598	0.2162	.23	1.2967 (0.8487-1.9811)
5 million–8 million	0.3942	0.2341	.10	1.4833 (0.9375-2.3466)
8 million–10 million	0.2782	0.2945	.35	1.3207 (0.7416-2.3521)
>10 million	0.3591	0.2982	.23	1.4320 (0.7982-2.5689)
I don't want to answer	-0.4969	0.2273	.03	0.6084 (0.3897-0.9499)
Education (reference: secondary school or less)				
High school	-0.4195	0.4269	.33	0.6574 (0.2848-1.5176)
Junior college or vocational school	0.1698	0.4512	.71	1.1851 (0.4895-2.8694)
University or postgraduate	-0.0057	0.4219	.99	0.9943 (0.4350-2.2732)
I don't want to answer	-0.3880	0.6364	.55	0.6784 (0.1949-2.3618)
Owning digital devices (reference: not owned)				
Mobile phone (smartphone)	0.6104	0.2732	.03[b]	1.8412 (1.0778-3.1453)
Personal computer	0.5058	0.1759	.004[d]	1.6582 (1.1747-2.3408)
Tablet	0.5233	0.1688	.002[d]	1.6876 (1.2123-2.3492)
Smartwatch	0.1524	0.2688	.57	1.1647 (0.6877-1.9724)
Smart home assistant devices and AI[e] appliances	0.8686	0.4054	.05[b]	2.3837 (1.0770-5.2758)
Knowledge of health databases and experience participating in health research				
Knowledge of research data repositories (reference: no)	0.9972	0.3505	.005[d]	2.7106 (1.3638-5.3876)
Experience participating in health research (reference: no)	1.1964	0.4420	.007[d]	3.3082 (1.3910-7.8680)
eHealth literacy	0.0660	0.0110	<.0001[d]	1.0680 (1.0450-1.0920)
Health status (reference: poor)	0.0478	0.1380	.73	1.0490 (0.8002-1.3750)
Concern about health (reference: no)	0.2840	0.070	<.0001[d]	1.3290 (1.1600-1.5230)

[a]OR, odds ratio.

XSL•FO
RenderX

[b]*P*<.01.

[c]US$1=147 Japanese Yen (stand: September 2025).

[d]*P*<.05.

[e]AI: artificial intelligence.

Willingness to Share Digital Health Data and Types of Data That Can Be Shared

The percentage of respondents who indicated a willingness to share shareable data (multiple responses) is shown in Multimedia Appendix 2 Figure S2. The results show that 44%, 33%, 33%, 32%, and 31% of participants were highly motivated to share data regarding physical clinical diagnoses, sleep patterns, blood samples, stress, and food consumption, respectively. Moreover, the content and frequency of social communication tended to be low (11% and 14%, respectively).

Predictors of Motivation to Share Digital Health Data and Positive Willingness to Share

Multimedia Appendix 3 Figure S3 shows the percentage of respondents who indicated each item as a motivation for sharing digital health data. Helping future patients was the most common motivation for sharing digital health data (59%), followed by receiving one's results (48%), supporting researchers (39%), and financial gain (31%).

Table 3 shows the results of the logistic analysis of the association between motives and positive willingness to share data. Motives found to be associated with positive motivation to share digital health data were helping future patients (OR 2.586, 95% CI 1.885 - 3.548; *P*<.001), receiving one's results (OR 2.266, 95% CI 1.624 - 3.561; *P*<.001), and receiving financial benefits (OR 1.806, 95% CI 1.263 - 2.582; *P*=.001).

Table . Logistic regression results regarding a positive willingness to share personal health data and motives and motivation to share data.

Motivations for sharing data	B (partial regression coefficient)	SE	*P* value	OR[a] 95% CI
Helping future patients	0.9501	0.1614	<.01[b]	2.5860 (1.8849-3.5481)
Supporting researchers	0.2607	0.1784	.14	1.2978 (0.9148-1.8412)
Receiving one's own results	0.8003	0.1608	<.001[b]	2.2261 (1.6243-3.0509)
Learning the results of the research one participated in	0.5416	0.2158	.012[c]	1.7187 (1.1259-2.6236)
Receiving financial benefits	0.5911	0.1824	.001[b]	1.8059 (1.2630-2.5822
Suggesting questions for future studies	0.3823	0.3250	.24	1.4656 (0.7751-2.7712)

[a]OR: odds ratio.

[b]*P*<.01.

[c]*P*<.05.

Concerns About Sharing Digital Health Data

The percentage of respondents who selected concerns about their digital health data is shown in Multimedia Appendix 4 FIgure S4 . Of the total respondents, 52% cited unknown use of data for the benefit of others as a concern, approximately 39% agreed to terms and conditions they did not fully understand, and approximately 34% said that data sharing would make them vulnerable to cyber attacks.

Table 4 shows the results of the logistic analysis of the relationship between positive intentions toward digital health data and concerns about sharing. The concern found to be associated with a positive willingness to share data was agreeing to contract terms that were not fully understood (OR 0.711, 95% CI 0.525 - 0.968; *P*=.03). The results showed that those who felt concerned tended to have lower positive intentions than those who had no concerns about their data being used for unethical projects (OR 0.510, 95% CI 0.361 - 0.722; *P*=.001).

XSL•FO
RenderX

Table . Logistic regression results regarding a positive willingness to share personal health data and concerns about sharing.

Concerning factors data sharing	B (partial regression coefficient)	SE	*P* value	OR[a] (95% CI)
Data being used for the benefit of a company or others without one's knowledge	0.2864	0.1467	.06	1.3316 (0.9989-1.7752)
Data being used for unethical projects	-0.6726	0.1770	<.001[b]	0.5104 (0.3607-0.7221)
Agreeing to terms and conditions without complete understanding	-0.3405	0.1572	.03[c]	0.7114 (0.5228-0.9681)
Exposure to risks, such as cyber-attacks	-0.0345	0.1597	.83	0.9661 (0.7064-1.3213)
Being asked to provide more data in the future	0.2406	0.1884	.20	1.2721 (0.8794-1.8401)

[a]OR: odds ratio.

[b]*P*<.01.

[c]*P*<.05.

Desirable Access Restrictions Related to Sharing Digital Health Data

Multimedia Appendix 5 FIgure S5 shows the percentage of participants' responses regarding their preferred access to digital health data sharing. Receiving information about projects using their data was the most common response (43%), followed by receiving information about who is using their data (35%) and deciding who has access to what parts of their data (31%). However, 13% stated they would not need to be contacted after providing their data and having it anonymized.

Table 5 shows the results of the logistic analysis between positive intentions and each of the access restriction items regarding data sharing. The access restrictions found to be associated with positive intentions were receiving information about projects using their data (OR 2.313, 95% CI 1.700 - 3.147; *P*=.001), giving public or academic institutions access to their data (OR 2.800, 95% CI 1.9248 - 4.0734; *P*=.001), and not receiving information after providing their data (OR 1.929, 95% CI 1.292 - 3.029; *P*=.004).

XSL•FO
RenderX

Table . Logistic regression results regarding a positive willingness to share personal data and data access restrictions.

Reasonable access and control over data	B (partial regression coefficient)	SE	*P* value	OR[a] (95% CI)
Receiving information about projects using the shared data	0.8386	0.1571	.001[b]	2.3132 (1.7001-3.1474)
Not receiving any communication after sharing data	0.6570	0.2301	.001[b]	1.9290 (1.2289-3.0282)
Receiving information about who is using the shared data	0.1932	0.1664	.25	1.2131 (0.8755-1.6811)
Deciding who has access to different parts of the data	0.0992	0.1728	.57	1.1043 (0.7871-1.5492)
Data controllers deciding who has access to the data	0.1525	0.1940	.43	1.1648 (0.7964-1.7036)
Granting data access to public or academic institutions	1.0297	0.1912	.001[b]	2.8001 (1.9248-4.0734)

[a]OR: odds ratio.

[b]$P<.01$.

Discussion

Principal Findings

Internationally, secondary use and third-party provision of personal health-related data have become very important for public health, including surveillance of infectious diseases and the development of preventive methods and treatments [48]. Therefore, it is significant to investigate the factors related to data sharing, such as citizens' attitudes toward, barriers to, and motivations for data sharing. However, research on citizens' personal health-related data sharing is limited [49].

Our study identifies factors influencing citizens' willingness to share a wide range of personal digital health data about health in Japan, including data in medical records, biobank samples, and digitized social communication data. This study showed that approximately 70% of Japanese participants were willing to share digital health data about their health. A survey of public attitudes toward digital research registries among Danish and Brazilian citizens showed that 91% of Brazilian and 70% of Danish citizens had a positive attitude toward digital research registries; our study shows a similar willingness as the results from the Danish participants [36]. Our research focuses on whether people are willing to share data, which directly evaluates behavioral intent. This approach is essential as it provides actionable data for organizations or governments to design effective policies. For instance, it can guide the creation of incentives to promote data sharing or enhance privacy protection measures.

While the Danish and Brazilian study asks about how "comfortable" people feel with data sharing, focusing on emotional aspects, our research directly addresses the question of "willingness to share." This specificity enables deeper insights into both emotional responses and actual behaviors, bridging the gap between attitudes and actions. Our research design, which evaluates the intent to share data, is not limited by cultural or emotional factors. This makes it versatile and applicable across different global contexts. Conducting similar studies in other countries allows us to uncover universal trends and contrasts in data sharing behavior. Our study suggests that helping future patients is an important motivator for sharing health-related digital health data. Furthermore, participants were willing to share their data in the interest of research.

The participants' high level of interest in their health suggested that gaining knowledge about their health is a motivating factor. Moreover, social benefits, such as benefits to future patients and benefits to oneself, were simultaneously motivating for sharing digital health data. Researcher benefits were not found to be associated with motivation for data sharing. A study of attitudes toward sharing health data among United States and United Kingdom citizens also reported a lower willingness to share data for commercial purposes compared to that for public benefit [43].

Based on our analysis, willingness to share digital health data regarding physical illness diagnoses, sleep patterns, food intake, blood samples, and stress was relatively high, while participants had a lower willingness to share social communication data. This may occur because social communication can be considered personal data that citizens may be hesitant to share due to

XSL•FO
RenderX

privacy concerns [50]. However, social communication provides data for research into mental health and infectious disease control; thus, it holds important implications in public health research. Therefore, there is a need to increase public understanding of the importance and significance of digital health data in this area.

Our results suggest that concerns about providing data to unethical projects and agreeing to contract terms that were not fully understood were barriers to the willingness to share digital health data. A study on public attitudes toward the sharing of personal medical records and genomic data in Japan similarly reported concerns about the unauthorized use of data [35]. Although the indicators differ from this study, a link between literacy in the knowledge of genomics, AI, and digital devices, and the sharing of medical records and genomic data has also been noted [35]. Furthermore, trust is important for data sharing [51,52]. Privacy concerns deter data sharing, particularly attitudes toward sharing health data for health policy creation [53].

Aggregating and analyzing multiple types of health data can have significant benefits. We can gain comprehensive insights into health patterns and conditions by combining diverse sources, such as genetic data, clinical records, and lifestyle habits. This approach supports personalized health care, allowing tailored treatments and interventions to address individual needs effectively. Furthermore, the correlation of health data enables predictive analysis, helping forecast potential health issues and promoting preventative measures [54]. Citizens' attitudes and receptivity to digital health data sharing differ depending on the type of data; however, concerns persist regarding personal data protection and transparency in data sharing [55].

While one measure to remove barriers of concern is restricting access to digital health data, our results showed that receiving information about projects that use data, getting information about data users, and granting data access to public or academic institutions are linked to a willingness to share data. These results may suggest that people avoid providing data to unethical projects and wish to exercise their right to self-determination regarding data. Other studies have also identified the importance of controlling the sharing of digital health data [36,37].

Our analysis suggests that the sociodemographic factors associated with willingness to share digital health data are being a man; owning smartphones, PCs, and smart appliances; having experience in participating in health research; and having high digital health literacy. It is necessary to consider measures for sharing digital health data that consider individual characteristics. The association between willingness to share digital health data and educational background reported by Grande et al [56] was not found in our study.

Implications for Policy and Practice

This study identifies factors that make Japanese citizens willing and able to share their data for research. These factors may be applicable not only in Japan but also internationally across borders. An increase in data sharing can contribute to research using big data, for example, the development of diagnostic, preventive, and therapeutic methods for rare diseases for which data are scarce, as well as international surveillance for infectious diseases.

The results of our study suggest that there are differences in the types of data that people are willing to share and that they may be less willing to share data that is related to personal data or not related to their health. Therefore, it is possible to fully explain the significance of sharing through communication with people, motivating them to share, and identifying the possible contributions to their health. Concerns about unethical data use projects and consenting before understanding the full terms of the agreement, identified as barriers to data sharing, can be addressed by disclosing information about the research project when requesting digital health data.

Concerns about unethical data usage and misunderstood consent indicate a demand for stronger ethical governance. Data governance frameworks must incorporate feedback mechanisms, such as providing people with updates about the projects that use their data, thereby enhancing perceived control and accountability.

Receiving this data may contribute to increased confidence in data sharing. Furthermore, ensuring that people have the right to control their data is also key to facilitating data sharing.

The study highlights demographic variations in willingness to share data, suggesting that public engagement strategies must be tailored to specific groups. For instance, women, people with lower eHealth literacy, and those without previous research participation may require more targeted education and reassurance regarding privacy and data control mechanisms.

The willingness of the Japanese public to share health data opens opportunities for international data pooling, particularly in the context of rare diseases and emerging infectious diseases, where domestic data may be insufficient for robust research. These findings support the establishment of cross-border data sharing frameworks that are ethical, secure, and culturally sensitive.

The association between device ownership and willingness to share digital health data suggests that disparities in access to digital technologies may translate into inequities in participation in health research. Policymakers should be aware of the digital divide when designing inclusive health data collection frameworks. These findings underscore the importance of policy measures that enhance public trust in digital health data sharing. Efforts to improve data literacy, transparency in data usage, and the ethical oversight of research projects are vital to increasing citizens' willingness to share their data. Government and academic institutions should consider developing communication strategies that convey how data will be used, who will access it, and the potential social benefits of data sharing.

Limitations and Future Challenges

Two methodological limitations exist in this study. First, the web-based survey could have created the possibility of sampling bias as well as bias in education levels and ownership of digital devices.

Another limitation relates to our sampling design. Although we used random selection within strata defined by age and gender,

XSL•FO
RenderX

this constitutes stratified random sampling rather than simple random sampling. In our analysis, we treated the sample as if drawn via simple random sampling, which may lead to inaccurate estimation of standard errors and affect inferential validity, including the risk of type I error. However, this design was intentionally chosen to ensure adequate representation across demographic subgroups, particularly to explore associations between data sharing intentions and age or gender. Given the total sample size of 1000, we believe the precision of estimates remains acceptable for our exploratory purposes, although results should be interpreted with appropriate caution.

Third, this study was a survey that asked questions about willingness to share digital health data based on future assumptions. However, the study did not explore the perspectives of data providers regarding implementing actual digital health data sharing.

While the quantitative data provides a broad understanding of attitudes, future research should include qualitative approaches, such as interviews or focus groups, to explore the underlying motivations, fears, and expectations surrounding data sharing in greater depth. Therefore, it is important to survey providers during actual data sharing in a future study. Conducting ongoing research into the topic may also be necessary.

Conclusion

This study identified people's receptivity and willingness to share personal digital health data about their health and its influencing factors. Approximately 70% of the respondents viewed sharing digital health data about their health positively. The majority of those with a positive view were motivated to share their data by receiving data about themselves as well as helping future patients. Unwillingness to share digital health data about personal health is a barrier to promoting digital sharing and research for societal benefit. Concerns about data being used in unethical research projects and agreeing to contractual terms and conditions that are not understood were suggested to be associated with reduced willingness to share data. It is also important to improve the understanding of the information of users' institutions, disclosure of information on the projects used, terms of contracts, and access control of data sharing. Furthermore, associations between willingness to share digital health data for public associations with personal characteristics were found regarding gender, interest in health issues, ownership of digital devices, history of participation in research, awareness of digital health databases, and digital health literacy. Therefore, a digital sharing system that considers individual characteristics and provides information regarding the usage of the data could facilitate data sharing.

Acknowledgments

This work was supported by JSPS KAKENHI (Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research; grants 21K01308 and 20K10359).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

YF analyzed and interpreted the data in the predictor analysis of the attitude and acceptance of personal data sharing for public benefits. YF and KF contributed to data collection, and YF and KF contributed to data analysis and interpretation. YF contributed to the drafting of the manuscript. YF and KF critically reviewed the manuscript. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Questionnaire.
[DOCX File, 37 KB - humanfactors_v12i1e64192_app1.docx]

Multimedia Appendix 2
Figure S2 Types of data that can be shared.
[PNG File, 22 KB - humanfactors_v12i1e64192_app2.png]

Multimedia Appendix 3
FIgure S3 Motivation to share digital health data.
[PNG File, 14 KB - humanfactors_v12i1e64192_app3.png]

Multimedia Appendix 4

XSL•FO
RenderX

Figure S4 Concerns about data sharing.
[PNG File, 16 KB - humanfactors_v12i1e64192_app4.png]

Multimedia Appendix 5
Figure S5 Desirable data access restrictions.
[PNG File, 16 KB - humanfactors_v12i1e64192_app5.png]

Checklist 1
Checklist for Reporting Results of Internet E-Surveys (CHERRIES).
[DOCX File, 28 KB - humanfactors_v12i1e64192_app6.docx]

References

1. Ghaffar Nia N, Kaplanoglu E, Nasab A. Evaluation of artificial intelligence techniques in disease diagnosis and prediction. Discov Artif Intell 2023;3(1):5. [doi: 10.1007/s44163-023-00049-5] [Medline: 40478140]
2. Stroud C, Onnela JP, Manji H. Harnessing digital technology to predict, diagnose, monitor, and develop treatments for brain disorders. NPJ Digit Med 2019;2:44. [doi: 10.1038/s41746-019-0123-z] [Medline: 31304390]
3. Kumar Y, Koul A, Singla R, Ijaz MF. Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda. J Ambient Intell Humaniz Comput 2023;14(7):8459-8486. [doi: 10.1007/s12652-021-03612-z] [Medline: 35039756]
4. Abernethy A, Adams L, Barrett M, et al. The promise of digital health: then, now, and the future. NAM Perspect 2022;2022. [doi: 10.31478/202206e] [Medline: 36177208]
5. Tang PC, Ash JS, Bates DW, Overhage JM, Sands DZ. Personal health records: definitions, benefits, and strategies for overcoming barriers to adoption. J Am Med Inform Assoc 2006;13(2):121-126. [doi: 10.1197/jamia.M2025] [Medline: 16357345]
6. McGuire AL, Hamilton JA, Lunstroth R, McCullough LB, Goldman A. DNA data sharing: research participants' perspectives. Genet Med 2008 Jan;10(1):46-53. [doi: 10.1097/GIM.0b013e31815f1e00] [Medline: 18197056]
7. Kamegaya T, Mori Y, Takahashi E. Exercise habits and perception of physical exercise among residents of an island in the Western Pacific Region. J Phys Ther Sci 2023 Aug;35(8):568-574. [doi: 10.1589/jpts.35.568] [Medline: 37529056]
8. Hoevenaars FPM, Berendsen CMM, Pasman WJ, et al. Evaluation of food-intake behavior in a healthy population: personalized vs. one-size-fits-all. Nutrients 2020 Sep 15;12(9):2819. [doi: 10.3390/nu12092819] [Medline: 32942627]
9. Brown R, Coventry L, Sillence E, et al. Collecting and sharing self-generated health and lifestyle data: Understanding barriers for people living with long-term health conditions - a survey study. Digit Health 2022;8:20552076221084458. [doi: 10.1177/20552076221084458] [Medline: 35284085]
10. Janett RS, Yeracaris PP. Electronic medical records in the American Health System: challenges and lessons learned. Cien Saude Colet 2020 Mar;25(4):1293-1304. [doi: 10.1590/1413-81232020254.28922019] [Medline: 32267432]
11. Wu CT, Li GH, Huang CT, et al. Acute exacerbation of a chronic obstructive pulmonary disease prediction system using wearable device data, machine learning, and deep learning: development and cohort study. JMIR Mhealth Uhealth 2021 May 6;9(5):e22591. [doi: 10.2196/22591] [Medline: 33955840]
12. Gallos P, Menychtas A, Panagopoulos C, Bimpas M, Maglogiannis I. Quantifying citizens' well-being in areas with natural based solutions using mobile computing. Stud Health Technol Inform 2022 Jan 14;289:465-468. [doi: 10.3233/SHTI210958] [Medline: 35062191]
13. Abdullah M, Dias C, Muley D, Shahin MD. Exploring the impacts of COVID-19 on travel behavior and mode preferences. Transp Res Interdiscip Perspect 2020 Nov;8:100255. [doi: 10.1016/j.trip.2020.100255] [Medline: 34173481]
14. Moser A, von Wyl V, Höglinger M. Health and social behaviour through pandemic phases in Switzerland: Regional time-trends of the COVID-19 Social Monitor panel study. PLOS ONE 2021;16(8):e0256253. [doi: 10.1371/journal.pone.0256253] [Medline: 34432842]
15. Coppola L, Cianflone A, Grimaldi AM, et al. Biobanking in health care: evolution and future directions. J Transl Med 2019 May 22;17(1):172. [doi: 10.1186/s12967-019-1922-3] [Medline: 31118074]
16. Nagai A, Hirata M, Kamatani Y, et al. Overview of the BioBank Japan project: study design and profile. J Epidemiol 2017 Mar;27(3S):S2-S8. [doi: 10.1016/j.je.2016.12.005] [Medline: 28189464]
17. Bulgarelli L, Núñez-Reiz A, Deliberato RO. Building electronic health record databases for research. In: Celi L, Majumder M, Ordóñez P, editors. Leveraging Data Science for Global Health: Springer; 2020:55-64. [doi: 10.1007/978-3-030-47994-7_4]
18. Wakabayashi Y, Eitoku M, Suganuma N. Characterization and selection of Japanese electronic health record databases used as data sources for non-interventional observational studies. BMC Med Inform Decis Mak 2021 May 22;21(1):167. [doi: 10.1186/s12911-021-01526-6] [Medline: 34022876]
19. Ishii E, Ebner DK, Kimura S, Agha-Mir-Salim L, Uchimido R, Celi LA. The advent of medical artificial intelligence: lessons from the Japanese approach. J Intensive Care 2020;8:35. [doi: 10.1186/s40560-020-00452-5] [Medline: 32467762]

XSL•FO
RenderX

20. World Health Organization. Global Strategy on Digital Health 2020-2025 2021. URL: https://www.who.int/docs/default-source/documents/gs4dhdaa2a9f352b0445bafbc79ca799dce4d.pdf [accessed 2025-03-31]
21. Filip R, Gheorghita Puscaselu R, Anchidin-Norocel L, Dimian M, Savage WK. Global challenges to public health care systems during the COVID-19 pandemic: a review of pandemic measures and problems. J Pers Med 2022 Aug 7;12(8):1295. [doi: 10.3390/jpm12081295] [Medline: 36013244]
22. Johnson KB, Wei WQ, Weeraratne D, et al. Precision medicine, AI, and the future of personalized health care. Clin Transl Sci 2021 Jan;14(1):86-93. [doi: 10.1111/cts.12884] [Medline: 32961010]
23. Vos JFJ, Boonstra A, Kooistra A, Seelen M, van Offenbeek M. The influence of electronic health record use on collaboration among medical specialties. BMC Health Serv Res 2020 Jul 22;20(1):676. [doi: 10.1186/s12913-020-05542-6] [Medline: 32698807]
24. Valentin G, Nielsen CV, Nielsen ASM, et al. Bridging inequity gaps in healthcare systems while educating future healthcare professionals-the social health bridge-building programme. Int J Environ Res Public Health 2023 Sep 27;20(19):6837. [doi: 10.3390/ijerph20196837] [Medline: 37835107]
25. National Institutes of Health (NIH). Genetic and rare diseases information center. Rare Diseases. 2021. URL: https://rarediseases.info.nih.gov/ [accessed 2025-06-05]
26. Rare and Intractable Diseases. Ministry of Health, Labour and Welfare, Japan. 2019. URL: https://www.mhlw.go.jp/ [accessed 2025-06-05]
27. Rare diseases: European Reference Networks. European Commission. 2020. URL: https://ec.europa.eu/health/rare_diseases [accessed 2025-06-05]
28. Gavrilov G, Vlahu-Gjorgievska E, Trajkovik V. Healthcare data warehouse system supporting cross-border interoperability. Health Informatics J 2020 Jun;26(2):1321-1332. [doi: 10.1177/1460458219876793] [Medline: 31581924]
29. McGraw D, Mandl KD. Privacy protections to encourage use of health-relevant digital data in a learning health system. NPJ Digit Med 2021 Jan 4;4(1):2. [doi: 10.1038/s41746-020-00362-8] [Medline: 33398052]
30. van Panhuis WG, Paul P, Emerson C, et al. A systematic review of barriers to data sharing in public health. BMC Public Health 2014 Nov 5;14:1144. [doi: 10.1186/1471-2458-14-1144] [Medline: 25377061]
31. Simpson E, Brown R, Sillence E, et al. Understanding the barriers and facilitators to sharing patient-generated health data using digital technology for people living with long-term health conditions: a narrative review. Front Public Health 2021;9:641424. [doi: 10.3389/fpubh.2021.641424] [Medline: 34888271]
32. Sanderson SC, Brothers KB, Mercaldo ND, et al. Public attitudes toward consent and data sharing in biobank research: a large multi-site experimental survey in the US. Am J Hum Genet 2017 Mar 2;100(3):414-427. [doi: 10.1016/j.ajhg.2017.01.021] [Medline: 28190457]
33. Kim KK, Sankar P, Wilson MD, Haynes SC. Factors affecting willingness to share electronic health data among California consumers. BMC Med Ethics 2017 Apr 4;18(1):25. [doi: 10.1186/s12910-017-0185-x] [Medline: 28376801]
34. Kim J, Kim H, Bell E, et al. Patient perspectives about decisions to share medical data and biospecimens for research. JAMA Netw Open 2019 Aug 2;2(8):e199550. [doi: 10.1001/jamanetworkopen.2019.9550] [Medline: 31433479]
35. Kusunose M, Muto K. Public attitudes toward cloud computing and willingness to share personal health records (PHRs) and genome data for health care research in Japan. Hum Genome Var 2023 Mar 30;10(1):11. [doi: 10.1038/s41439-023-00240-1] [Medline: 36990988]
36. Nunes Vilaza G, Coyle D, Bardram JE. Public attitudes to digital health research repositories: cross-sectional International Survey. J Med Internet Res 2021 Oct 29;23(10):e31294. [doi: 10.2196/31294] [Medline: 34714253]
37. Varhol RJ, Norman R, Randall S, et al. Public preference on sharing health data to inform research, health policy and clinical practice in Australia: A stated preference experiment. PLOS ONE 2023;18(11):e0290528. [doi: 10.1371/journal.pone.0290528] [Medline: 37972118]
38. Biasiotto R, Viberg Johansson J, Alemu MB, et al. Public preferences for digital health data sharing: discrete choice experiment study in 12 European countries. J Med Internet Res 2023 Nov 23;25:e47066. [doi: 10.2196/47066] [Medline: 37995125]
39. Australia's health reimagined. The journey to a connected and confident consumer. : Curtin University; 2022 URL: https://digitalhealthcrc.com/wp-content/uploads/2022/09/Australias-health-reimagined_Report_WEB_FINAL.pdf [accessed 2025-10-02]
40. Tully MP, Bernsten C, Aitken M, Vass C. Public preferences regarding data linkage for research: a discrete choice experiment comparing Scotland and Sweden. BMC Med Inform Decis Mak 2020 Jun 16;20(1):109. [doi: 10.1186/s12911-020-01139-5] [Medline: 32546147]
41. Rising CJ, Gaysynsky A, Blake KD, Jensen RE, Oh A. Willingness to share data from wearable health and activity trackers: analysis of the 2019 health information national trends survey data. JMIR Mhealth Uhealth 2021 Dec 13;9(12):e29190. [doi: 10.2196/29190] [Medline: 34898448]
42. Aitken M, de St Jorre J, Pagliari C, Jepson R, Cunningham-Burley S. Public responses to the sharing and linkage of health data for research purposes: a systematic review and thematic synthesis of qualitative studies. BMC Med Ethics 2016 Nov 10;17(1):73. [doi: 10.1186/s12910-016-0153-x] [Medline: 27832780]

XSL•FO
RenderX

43. Ghafur S, Van Dael J, Leis M, Darzi A, Sheikh A. Public perceptions on data sharing: key insights from the UK and the USA. Lancet Digit Health 2020 Sep;2(9):e444-e446. [doi: 10.1016/S2589-7500(20)30161-8] [Medline: 32838250]
44. de Bekker-Grob EW, Donkers B, Jonker MF, Stolk EA. Sample size requirements for discrete-choice experiments in healthcare: a practical guide. Patient 2015 Oct;8(5):373-384. [doi: 10.1007/s40271-015-0118-z] [Medline: 25726010]
45. Norman CD, Skinner HA. eHEALS: The eHealth Literacy Scale. J Med Internet Res 2006 Nov 14;8(4):e27. [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
46. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. J Med Internet Res 2006 Jun 16;8(2):e9. [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]
47. Act on the protection of personal information (partly unenforced). Japanese Law Translation. URL: https://www.japaneselawtranslation.go.jp/ja/laws/view/4241 [accessed 2025-03-31]
48. Cascini F, Pantovic A, Al-Ajlouni YA, Puleo V, De Maio L, Ricciardi W. Health data sharing attitudes towards primary and secondary use of data: a systematic review. EClinicalMedicine 2024 May;71:102551. [doi: 10.1016/j.eclinm.2024.102551] [Medline: 38533128]
49. Baines R, Stevens S, Austin D, et al. Patient and public willingness to share personal health data for third-party or secondary uses: systematic review. J Med Internet Res 2024 Mar 5;26:e50421. [doi: 10.2196/50421] [Medline: 38441944]
50. Naeem I, Quan H, Singh S, et al. Factors associated with willingness to share health information: rapid review. JMIR Hum Factors 2022 Feb 9;9(1):e20702. [doi: 10.2196/20702] [Medline: 35138263]
51. Cherif E, Bezaz N, Mzoughi M. Do personal health concerns and trust in healthcare providers mitigate privacy concerns? Effects on patients' intention to share personal health data on electronic health records. Soc Sci Med 2021 Aug;283:114146. [doi: 10.1016/j.socscimed.2021.114146] [Medline: 34242891]
52. Belfrage S, Helgesson G, Lynøe N. Trust and digital privacy in healthcare: a cross-sectional descriptive study of trust and attitudes towards uses of electronic health data among the general public in Sweden. BMC Med Ethics 2022 Mar 4;23(1):19. [doi: 10.1186/s12910-022-00758-z] [Medline: 35246118]
53. McCormick JB, Hopkins MA. Exploring public concerns for sharing and governance of personal health information: a focus group study. JAMIA Open 2021 Oct;4(4):ooab098. [doi: 10.1093/jamiaopen/ooab098] [Medline: 34926998]
54. Eisinger-Mathason TSK, Leshin J, Lahoti V, Fridsma DB, Mucaj V, Kho AN. Data linkage multiplies research insights across diverse healthcare sectors. Commun Med (Lond) 2025 Mar 4;5(1):58. [doi: 10.1038/s43856-025-00769-y] [Medline: 40038513]
55. Nwebonyi N, Silva S, de Freitas C. Public views about involvement in decision-making on health data sharing, access, use and reuse: the importance of trust in science and other institutions. Front Public Health 2022;10:852971. [doi: 10.3389/fpubh.2022.852971] [Medline: 35619806]
56. Grande D, Mitra N, Iyengar R, et al. Consumer willingness to share personal digital information for health-related uses. JAMA Netw Open 2022 Jan 4;5(1):e2144787. [doi: 10.1001/jamanetworkopen.2021.44787] [Medline: 35072717]

Abbreviations

CHERRIES: Checklist for Reporting Results of Internet E-Surveys
eHEALS: eHealth Literacy Scale
OR: odds ratio

Edited by A Kushniruk; submitted 10.07.24; peer-reviewed by MM Llebot, M Mallén-Berdejo, R Moser, SH Sung; revised version received 15.06.25; accepted 10.07.25; published 09.10.25.

Please cite as:
Fukuda Y, Fukuda K
Public Attitudes and Predictors of Public Awareness of Personal Digital Health Data Sharing for Research: Cross-Sectional Study in Japan
JMIR Hum Factors 2025;12:e64192
URL: https://humanfactors.jmir.org/2025/1/e64192
doi:10.2196/64192

© Yasue Fukuda, Koji Fukuda. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 9.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Patient Satisfaction With a Comprehensive Remote Care Program for Chronic Condition Management for Adults Under Medical Care: Observational Study

Michael Merrill[1*], MD, MS, MBA; Susan M Zbikowski[2*], PhD; Garrett J Jenkins, PhD

[1]Brook, Inc, Seattle, WA, United States

[2]inZights Consulting, LLC, 834 NE 95th Street, Seattle, WA, United States

[*]these authors contributed equally

Corresponding Author:
Susan M Zbikowski, PhD
inZights Consulting, LLC, 834 NE 95th Street, Seattle, WA, United States

Abstract

Background: More than half of adults in the United States have at least one chronic medical condition. Remote care strategies, including telemedicine, remote monitoring, and connected health devices, are increasingly used to support chronic condition management and adherence to provider-recommended care plans. Evaluating participant satisfaction and perceived usefulness is essential to understanding program value and potential for sustained engagement.

Objective: This study aims to evaluate participant satisfaction, perceived usefulness, and self-reported health outcomes associated with a clinically deployed Remote Care program for chronic condition management that included nurse and coach follow-up, health monitoring devices, and digital education tools.

Methods: A survey was emailed to 1411 active participants in the Brook Remote Care program between September 23 and December 8, 2023. The survey assessed satisfaction with program features, perceived usefulness, and self-reported health changes. Descriptive statistics were used to summarize the data, and chi-square tests were used to examine differences by age, gender, program duration, and health conditions.

Results: A total of 360 participants completed the survey. Most participants rated the program as useful (320/359, 89%), and 66% (234/357) indicated they were likely to recommend it to others with similar health concerns. Between 68% (216/316) and 84% (277/332) rated specific program features as “good” to “excellent,” with nurse monitoring, health monitoring devices, and program cost receiving the most favorable ratings. Nearly half of participants reported improvements in physical health (155/314, 49%), reduced health-related stress (148/326, 45%), and increased knowledge of their condition (172/360, 48%). Additionally, approximately one-third reported improvements in lifestyle behaviors. Longer program duration was associated with higher satisfaction and greater self-reported health improvements.

Conclusions: Participants reported high satisfaction with a remote care program designed to support chronic condition self-management. These findings support the value of integrated remote care models that combine device monitoring, personalized follow-up, and patient education to enhance patient experience and promote engagement in self-care.

(JMIR Hum Factors 2025;12:e69586) doi:10.2196/69586

KEYWORDS

remote monitoring; patient satisfaction; chronic disease; self-management; telemedicine; digital health

Introduction

Over half of adults in the United States have at least one chronic medical condition [1]. Most follow a medically prescribed treatment plan and adopt self-management strategies, including lifestyle changes and self-monitoring, to improve disease control and clinical outcomes [2]. Many chronic conditions also require frequent medical visits or hospitalization. Remote care, such as telehealth and monitoring with connected devices, offers an alternative to in-person visits, supporting condition management and reducing the risk of acute events and hospitalizations [3,4].

Remote care programs have been shown to reduce mortality, disease complications, hospitalizations, and associated health care costs across a range of medical conditions [5-7]. For example, one study involving patients with chronic obstructive pulmonary disease or congestive heart failure found a 68% and 35% reduction in emergency room visits and hospitalizations, respectively, over a 6-month period (comparing 3 months before and after the intervention) [6,7].

XSL•FO
RenderX

Patient satisfaction is a widely used quality metric, used both to evaluate care and meet regulatory reporting requirements. Understanding how patients experience remote care is essential for assessing its perceived value, ongoing engagement, and potential for broader adoption. A review by Kraai et al [8] reported high satisfaction with remote care, but most studies included in that review were pilot trials with varied designs and no evaluation of remote care offered in a routine clinical setting. A more recent study by Mayo Clinic researchers [9] found that patients receiving remote care for acute conditions, such as COVID-19, were highly satisfied with their care team and the usability of the technology. However, this study focused primarily on an acutely ill population, with 88% being treated for COVID-19, limiting generalizability to chronic care settings.

There remains a need to evaluate patient satisfaction with remote care programs designed specifically for community-dwelling individuals managing chronic diseases. This study aims to contribute to the literature by examining satisfaction and self-reported health changes among insured patients enrolled in a clinically delivered remote care program for chronic condition management.

Methods

Survey Administration

In September 2023, Brook Health emailed active participants (N=1411) in its digital Remote Care program, inviting them to complete a survey (administered via Google Forms) to assess their experience and satisfaction as part of ongoing quality assurance efforts. Two reminder emails were sent approximately 2 weeks apart, followed by a final reminder shortly before the survey closed—3 months after the initial invitation. To be eligible, individuals had to be actively enrolled in the Remote Care program and have an email address on file. Participation was voluntary. To encourage responses, participants were offered the opportunity to enter a drawing for 1 of 5 US $100 Amazon gift cards.

Remote Care Program Description

The Remote Care program was developed and operated by Brook Health, a Seattle-based company that delivers digital health services to support clinic operations, provider-delivered care, and patient well-being. The program includes app-based educational resources and tools; wireless, cellular-enabled monitoring devices (scales, blood pressure cuffs, glucose monitors, and pulse oximeters) that do not require Bluetooth syncing; personalized one-on-one health coaching; and ongoing nurse monitoring to ensure care plan adherence and optimal condition management. Nurses monitored patient data and contacted clinics as needed to escalate care concerns. Patients were referred to the program by their providers and enrolled either onsite at the clinic or from home. After enrollment, patients received information about the program, downloaded the app, created an account, and were provided with digital monitoring devices. Participants received nurse monitoring and coaching calls throughout the program, with the highest frequency of contact occurring during the first month. Program costs were covered by insurance, with participants responsible for applicable copays.

Survey Instrument

The survey assessed participants' perceptions of program usefulness, satisfaction with specific features, and self-reported changes in health and well-being (refer to Textbox 1 for items). Brook developed a custom instrument informed by established digital health evaluation models, with an emphasis on face validity and alignment with program objectives. Acceptability was operationalized through participant experience, including user satisfaction ratings. Designed as a quality improvement tool for real-world implementation, the survey was intentionally brief and written in accessible language to accommodate a diverse patient population. Formal psychometric testing was not conducted. In addition to items assessing specific program components, the survey included a general self-rated health item from the 20-item Short Form Survey and a global health change item modeled after the Global Rating of Change scale, a widely used patient-reported outcome measure [10,11].

XSL•FO
RenderX

Textbox 1. Participant survey questions for the Brook Remote Care Program.

Program ratings

How would you rate the following aspects of the Brook Remote Care program? Very poor, poor, average/acceptable, good, or excellent.

- Ease of enrolling
- Program content
- Nurse interactions
- Coaching
- Devices provided
- Cost of the program
- Program duration
- Results/outcomes

How would you rate the following aspects of Nurse Monitoring in the Remote Care program? Very poor, poor, average/acceptable, good, or excellent.

- Quality of interactions with the nurse
- How often the nurse contacted you
- Amount of time spent on calls with the nurse
- Recommendations made by the nurse

On a scale from 1 to 10, how likely are you to recommend this program to others with similar health concerns? (1=Not all to 10=Extremely likely)

Rate how useful Brook Remote Care has been in helping you to manage your health/health condition? Extremely, very, somewhat, a little, or not at all.

Health ratings

In general, would you say your health is: excellent, very good, good, fair, or poor?

Would you say your health has improved, not changed at all, or worsened since joining Brook Remote Care?

Since joining Brook Remote Care, how do you feel when it comes to stressing about your health status?

While participating in Brook Remote Care, did you...? Check all that apply.

Achieve better control of your health condition

Lose weight

Increase physical activity

Eat healthier

Play a more active role in your health

Reduce medication use

Reduce stress/feel less anxious

Sleep better/more

Become happier

Become more knowledgeable about your condition and management

Spend less time on medical appointment and tests

None of the above/other

How has your use of medications to manage your condition (for which you were referred to Remote Care) changed since joining the program? Check all that apply.

Not on any medications for my condition (never took or stopped taking before starting Brook)

No change in my medication use

I started taking medication(s) for my condition

Changed the type of medication

Decreased amount, strength or frequency

Increased amount, strength, or frequency

XSL•FO
RenderX

Descriptives

How long have you been participating in Brook Remote Care? Less than 1 month, 1-3 months, 4-6 months, 7-9 months, 10-12 months, or more than 12 months.

For which medical condition are you being supported by Brook Remote Care? Congestive heart failure, diabetes, hypertension, weight loss, or other.

Age in years: 18-24, 25-34, 35-44, 45-54, 55-64, 65-74, or 75+.

Your gender: female, male, nonbinary, or transgender.

Ethical Considerations

Participation was voluntary. The following information was included with the survey. “Your responses to this survey are confidential. Your answers will not be shared with your provider or anyone else. Brook Health will only use your responses in aggregate with responses from others who complete this survey. By answering this survey, we presume your approval to use your responses for the following purposes...1) improve existing services, 2) inform the development of new services, 3) create marketing materials, or 4) generate research reports.” To encourage responses, participants were offered the opportunity to enter a drawing for 1 of 5 US $100 Amazon gift cards. This study was reviewed by the University at Buffalo institutional review board and deemed exempt from full review (IRB ID STUDY00008350).

Sample

A power calculation determined that 302 completed surveys were needed to achieve a statistically representative sample with 95% confidence and a 5% margin of error. A total of 441 surveys were received. During data cleaning, 81 duplicate responses were identified and removed; specifically, only the first complete response from each participant (based on the timestamp of submission) was retained. This resulted in 360 unique surveys included in the final analysis, yielding a response rate of 26%.

Descriptive Analysis

Descriptive statistics were calculated for all demographic variables and survey items. Chi-square analyses were conducted to examine group differences in ratings of program features, nurse monitoring, overall program usefulness, likelihood of recommending the program, and self-reported health status and health changes. Comparisons were made across gender (male vs female), age (<65 vs 65+ years), selected health conditions (hypertension, diabetes, obesity, and congestive heart failure), and program duration (<6 months, 7 - 12 months, and 12+ months). In addition, multivariate logistic regression models were conducted to assess whether observed associations remained statistically significant after adjusting for age, gender, and program duration, accounting for potential confounding factors.

Item Scoring

The Net Promoter Score (NPS), a widely used measure of overall program experience, was derived from the question: “On a scale from 1 to 10, how likely are you to recommend this program to others with similar health concerns?” (1=not at all likely; 10=extremely likely). Respondents were categorized as Promoters (scores of 9 - 10), Passives (scores of 7 - 8), or Detractors (scores of 6 or below). The NPS was calculated as the percentage of Promoters minus the percentage of Detractors. Positive NPS values indicate that more participants would recommend the program than not, while negative scores suggest the opposite [12].

Results

Sample Characteristics

Most participants identified as female (223/359, 62%), were aged 65 years or older (251/359, 70%), and had a single chronic condition (220/359, 61%), most commonly hypertension (266/359, 74%). Just more than half (204/3, 57%) had been enrolled in the program for less than 6 months (Table 1).

XSL•FO
RenderX

Table . Survey respondent demographics.

Demographics	Values (N=359), n (%)
Gender (N=359)	
Female	223 (62)
Male	136 (38)
Age (years; n=359)	
<35	4 (1)
35 - 44	18 (2)
45 - 54	25 (7)
55 - 64	61 (17)
65 - 74	178 (50)
75+	73 (20)
Health condition[a]	
Single	220 (61)
Multiple	140 (39)
Hypertension	266 (74)
Obesity	127 (35)
Diabetes	98 (27)
Congestive heart failure	27 (8)
Other	20 (6)
Duration in the program (months)	
<6	204 (57)
7 - 12	99 (28)
12+	56 (16)

[a]Multiresponse item. Participants could report multiple health conditions.

Self-Reported Health and Health-Related Changes

Most respondents rated their health as “good” or better (263/358, 74%). Nearly half reported improvements in their physical health (155/314, 49%) and a reduction in health-related stress (148/326, 45%) since enrolling in the program. Reported behavior and condition management improvements included increased knowledge of their condition (172/360, 48%), more active involvement in their care (143/360, 40%), healthier eating (120/360, 33%), improved condition control (116/360, 32%), increased physical activity (112/360, 31%), and reduced stress (96/360, 27%; Table 2).

XSL•FO
RenderX

Table . Survey respondents' self-reported health and health-related changes.

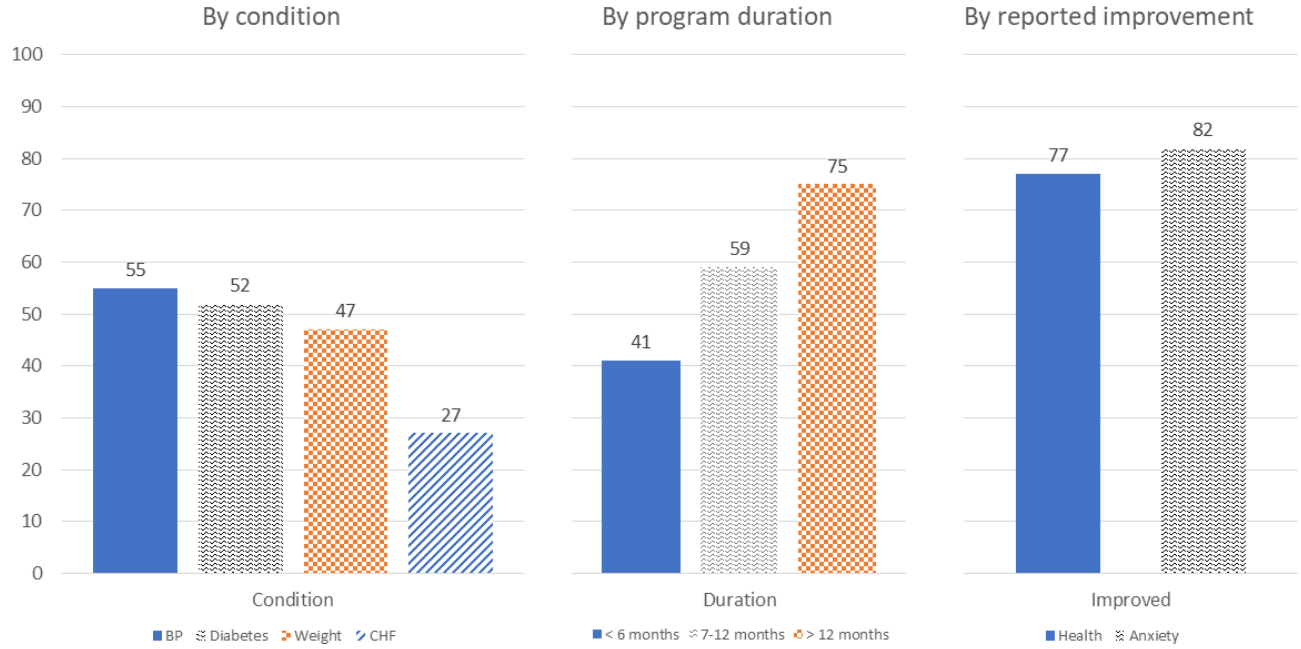
Health-related items	Values, n (%)
Health status (n=358)	
Poor	14 (4)
Fair	81 (23)
Good	186 (52)
Very good	67 (19)
Excellent	10 (3)
Health change since enrolling (n=314)	
Improved	155 (49)
Same	153 (49)
Worsened	6 (2)
Health-related stress (n=326)	
Less	148 (45)
Same	156 (48)
More	22 (7)
Reported improvements (n=360)	
Knowledge	172 (48)
More active role in health	143 (40)
Eat healthier	120 (33)
Better condition control	116 (32)
Lose weight	108 (30)
Increase physical activity	112 (31)
Reduce stress	96 (27)
Sleep better	52 (14)
Happier	39 (11)
Less time on medical appointments/tests	28 (8)
Medication change (n=356)	
Decrease	42 (12)
Increase	40 (11)
Started/changed medication	65 (18)
Not on any/no change	235 (66)

Program Ratings

The majority of participants (320/359, 89%) rated the Remote Care program as "somewhat" to "extremely" useful. Based on responses to the NPS item, 66% (234/357) were classified as Promoters, 20% (72/357) as Passives, and 14% (51/357) as Detractors, yielding an overall NPS of 52, categorized as "excellent." NPS was higher among those with longer enrollment, self-reported health improvements, and individuals managing diabetes or hypertension (Figure 1).

XSL•FO
RenderX

Figure 1. Net Promoter Scores by condition, program duration, and reported health improvements. BP: hypertension; CHF: congestive heart failure.



Program Feature Ratings

All Remote Care program features received favorable ratings, with most respondents rating components as "good" to "excellent." Nurse monitoring, remote monitoring devices, and program cost received the highest "excellent" ratings (>50%). Coaching, educational content, and outcomes were rated more variably, with a broader distribution across "average," "good," and "excellent" ratings. Fewer than 7% of respondents rated any feature as "poor" or "very poor" (Table 3).

XSL•FO
RenderX

Table . Program participant responses related to aspects of the Brook Remote Care program and nurse monitoring.

Questions (number of respondents)	Answers					
	Excellent, n (%)	Good, n (%)	Average/ acceptable, n (%)	Poor, n (%)	Very poor, n (%)	Good to excellent, %
Aspects of Brook Remote Care Program						
Ease of enrolling (n=332)	155 (47)	122 (37)	46 (14)	4 (1)	5 (2)	84
Content (n=322)	115 (36)	138 (43)	57 (18)	7 (2)	5 (2)	79
Nurse interactions (n=331)	176 (53)	99 (30)	45 (14)	7 (2)	4 (1)	83
Coaching (n=316)	118 (37)	98 (31)	78 (25)	15 (5)	7 (2)	68
Devices (n=330)	175 (53)	98 (30)	49 (15)	3 (1)	5 (1)	83
Program cost (n=317)	204 (64)	62 (20)	40 (13)	6 (2)	5 (2)	84
Program duration (n=307)	120 (39)	111 (36)	69 (22)	3 (1)	4 (1)	71
Outcomes (n=312)	113 (36)	112 (36)	72 (23)	7 (2)	8 (3)	72
Aspect of nurse monitoring						
Quality of interaction (n=333)	183 (55)	83 (25)	51 (15)	10 (3)	6 (2)	80
Frequency of contact (n=331)	152 (46)	93 (28)	70 (21)	13 (4)	3 (1)	74
Amount of time spent during calls (n=329)	149 (45)	96 (29)	67 (20)	11 (3)	6 (2)	74
Recommendations made by nurse (n=326)	148 (45)	91 (28)	65 (20)	17 (5)	5 (2)	73

Comparative Analyses by Participant Characteristics

Chi-square analyses indicated that self-reported physical health improvements were significantly associated with age (χ^2_1=10.56; *P*=.001) and program duration (χ^2_2=8.72; *P*=.01). Reductions in health-related stress differed by gender (χ^2_1=7.18; *P*=.007) and program duration (χ^2_2=5.99; *P*=.05), with trends also observed for age (χ^2_1=3.62; *P*=.06). Participants enrolled longer reported higher satisfaction across several program components, including ease of enrollment, content, duration, and perceived outcomes (all *P*<.05). Nurse monitoring ratings were also significantly higher among those with longer program participation (χ^2_4=10.49; *P*=.02). No significant differences in program or nurse ratings were found by health condition. A nonsignificant trend suggested men rated nurse recommendations more positively than women (χ^2_4=5.54; *P*=.06), and younger adults rated overall program usefulness slightly higher (χ^2_2=4.87; *P*=.09). Multivariate logistic regression models including age, gender, and program duration were conducted to adjust for potential confounding factors. Multivariate findings were consistent with univariate results, with no substantial changes in the significance or direction of observed associations.

Discussion

Principal Findings

This study found high participant satisfaction with a clinically implemented Remote Care program designed to support chronic condition management through digital monitoring, nurse outreach, and personalized coaching. Most participants reported the program was useful in helping manage their health, and satisfaction was highest among those with longer program engagement. These findings add to a growing body of literature supporting the acceptability and perceived value of remote care solutions for patients with chronic disease.

Participants reported a variety of benefits beyond satisfaction, including improvements in physical and emotional health, condition-related knowledge, and self-management behaviors. These findings are consistent with prior studies showing that remote care can support healthier lifestyles and improved condition control [6,13]. While the observational design limits causal conclusions, the high percentage of participants reporting improvements, particularly in knowledge and engagement, suggests the program may support behavior change in a real-world clinical setting.

XSL•FO
RenderX

Program duration emerged as a consistent factor associated with greater perceived benefit. Participants who were enrolled longer reported more health improvements, higher satisfaction ratings, and stronger endorsement of program features, including nurse support and educational content. While the direction of this relationship is unclear, whether satisfaction encourages continued engagement or vice versa, this aligns with prior research showing satisfaction and engagement are mutually reinforcing and essential to program effectiveness [14].

Few differences in satisfaction emerged across demographic groups. While age and gender were associated with some trends in health perception and nurse rating preferences, program usefulness and satisfaction ratings did not vary significantly by gender, age, or specific health condition. Notably, men, who often engage less with health coaching, rated nurse recommendations particularly favorably, suggesting an opportunity to better tailor outreach or messaging to male participants. Although the association did not reach significance, this pattern may indicate that nurse-led outreach is especially effective for engaging men in self-management—a group often underrepresented in such programs. This warrants further exploration.

Younger adults also appeared to perceive the program as slightly more useful, a pattern that contrasts with previous findings where older adults more often rate remote tools favorably [9]. This may reflect generational differences in expectations for digital health and could inform strategies to optimize engagement across age cohorts. While these trends fell short of statistical significance, they may still carry clinical relevance, particularly given their consistency with prior research and potential value in refining remote care models. In real-world settings with modest sample sizes, patterns like these may not reach conventional significance thresholds yet still highlight meaningful opportunities to personalize care and improve patient experience.

Nurse monitoring was highly rated and is likely a core driver of overall satisfaction. One-on-one nurse interactions may have contributed directly to the reported improvements in condition knowledge, stress reduction, and behavior change. Prior studies support the conclusion that structured remote care with personalized nurse support can enhance self-management and improve outcomes for patients with chronic illness [5,13].

The NPS, an indicator of participant willingness to recommend the program to others, was in the “excellent” range. Importantly, this finding came from a sample with a full range of user experiences, including detractors and passives, not just those who were highly satisfied. This suggests that the program delivered a consistently positive experience across a diverse patient population.

Why This Is Meaningful

Remote care has grown in popularity, especially following the COVID-19 pandemic, as health care systems seek scalable approaches to support chronic disease management [4]. Evidence suggests that remote care and telemonitoring can improve patient outcomes and reduce health care use, including emergency department visits and hospitalizations [5-7]. As these programs become more widely adopted, understanding the patient experience is critical to sustaining engagement and maximizing impact.

This study contributes to the growing body of literature showing high satisfaction with remote care programs [8,9] and extends it by focusing on a clinically integrated program for patients with chronic, not acute, conditions. The consistently high satisfaction and NPSs across diverse subgroups indicate that user-centered remote care models are feasible, scalable, and well-suited for long-term condition management.

Limitations

This study has several limitations. First, outcomes are based on subjective, self-reported data collected from a single group of participants, without a comparison group. This limits our ability to determine whether the observed outcomes were attributable to the program itself or reflect the natural course of condition management.

Second, the survey had a response rate of 26%. It is possible that the experiences and satisfaction levels of survey respondents differ from those who did not respond. However, the distribution of NPSs suggests a range of perspectives was captured. Specifically, 66% (234/357) of respondents were classified as “Promoters,” 20% (72/357) as “Passives,” and 14% (51/357) as “Detractors,” indicating the survey captured both positive and critical feedback about participants’ likelihood to recommend the program.

Third, the high satisfaction ratings related to program cost are likely influenced by the fact that the program was covered by insurance for nearly all participants. This may have bolstered overall satisfaction scores. It is unclear whether participants would have rated the program as highly if they were required to pay out of pocket. Prior research has shown that cost is a frequent barrier to program use [15].

Fourth, while some demographic and health-related differences were observed between survey responders and nonresponders, these were modest and not suggestive of systematic response bias; nonetheless, the possibility of unmeasured differences affecting the generalizability of findings cannot be ruled out.

Finally, the demographic data available for analysis were limited to age, gender, condition, and program duration. We did not have access to important clinical characteristics such as duration of diagnosis or disease severity. Future studies should consider collecting more detailed clinical data to better characterize the participant population and contextualize outcomes.

Because the survey was developed for quality improvement purposes and not psychometrically validated, the reliability and comparability of the findings may be limited. However, the use of face-valid items aligned with program goals supports their relevance for capturing participant experience in real-world care settings.

Conclusions

This study demonstrates that patients managing chronic conditions within a clinically delivered Remote Care program reported high levels of satisfaction, perceived usefulness, and

XSL•FO
RenderX

self-reported health improvements. Longer duration in the program was consistently associated with greater perceived benefit, highlighting the value of sustained engagement.

The findings suggest that remote care programs incorporating regular nurse outreach, connected health devices, and tailored educational content can support patients in managing their health outside traditional clinical settings. As health care systems continue to adopt digital and hybrid care models, understanding patient experience will remain essential to driving meaningful outcomes. Future research should incorporate longitudinal or comparative designs to evaluate the program's clinical effectiveness and long-term impact on behavior change and health outcomes.

Acknowledgments

The authors are grateful to Sarah LaPointe, PhD, for methodological advice and contributing to the survey design. We are also thankful for the contributions of Kelsea Hoover, Jay Burke, Alex Foley, and Cami Woomer for input regarding survey design, data collection, and review of survey results. ChatGPT (OpenAI) was used to assist with language editing and to improve the flow of the manuscript after it was drafted by the authors. All AI-generated suggestions were critically reviewed and approved by the authors prior to inclusion in the final version. No generative AI tools were used to generate scientific content, analyze, or interpret results.

Brook, Inc. provided funding and financial support for this study and publication.

Data Availability

The datasets generated during this study are proprietary and confidential, and therefore are not publicly available, but may be obtained from the corresponding author upon reasonable request.

Authors' Contributions

SMZ contributed to conceptualization, analysis, methodology, writing–original draft, and writing–review & editing. MM assisted with conceptualization, supervision, writing–review & editing. GJJ handled data analysis and manuscript review and editing.

Conflicts of Interest

MM is employed by Brook, Inc, serves as the Chief Medical Officer at Brook, Inc., and has stock options at Brook, Inc. Brook contracted with inZights Consulting (SZ) to serve as independent evaluator on this project and provide consultation on research projects.

References

1. Boersma P, Black LI, Ward BW. Prevalence of multiple chronic conditions among US adults, 2018. Prev Chronic Dis 2020 Sep 17;17:E106. [doi: 10.5888/pcd17.200130] [Medline: 32945769]
2. Lorig KR, Holman H. Self-management education: history, definition, outcomes, and mechanisms. Ann Behav Med 2003 Aug;26(1):1-7. [doi: 10.1207/S15324796ABM2601_01] [Medline: 12867348]
3. Peyroteo M, Ferreira IA, Elvas LB, Ferreira JC, Lapão LV. Remote monitoring systems for patients with chronic diseases in primary health care: systematic review. JMIR Mhealth Uhealth 2021 Dec 21;9(12):e28285. [doi: 10.2196/28285] [Medline: 34932000]
4. Shaver J. The state of telehealth before and after the COVID-19 pandemic. Prim Care 2022 Dec;49(4):517-530. [doi: 10.1016/j.pop.2022.04.002] [Medline: 36357058]
5. Inglis SC, Clark RA, Dierckx R, Prieto-Merino D, Cleland JGF. Structured telephone support or non-invasive telemonitoring for patients with heart failure. Cochrane Database Syst Rev 2015 Oct 31;2015(10):CD007228. [doi: 10.1002/14651858.CD007228.pub3] [Medline: 26517969]
6. Isaranuwatchai W, Redwood O, Schauer A, Van Meer T, Vallée J, Clifford P. A remote patient monitoring intervention for patients with chronic obstructive pulmonary disease and chronic heart failure: pre-post economic analysis of the smart program. JMIR Cardio 2018 Dec 20;2(2):e10319. [doi: 10.2196/10319] [Medline: 31758770]
7. Taylor ML, Thomas EE, Snoswell CL, Smith AC, Caffery LJ. Does remote patient monitoring reduce acute care use? A systematic review. BMJ Open 2021 Mar 2;11(3):e040232. [doi: 10.1136/bmjopen-2020-040232] [Medline: 33653740]
8. Kraai IH, Luttik MLA, de Jong RM, Jaarsma T, Hillege HL. Heart failure patients monitored with telemedicine: patient satisfaction, a review of the literature. J Card Fail 2011 Aug;17(8):684-690. [doi: 10.1016/j.cardfail.2011.03.009] [Medline: 21807331]
9. Haddad TC, Maita KC, Inselman JW, et al. Patient satisfaction with a multisite, multiregional remote patient monitoring program for acute and chronic condition management: survey-based analysis. J Med Internet Res 2023 Jul 27;25:e44528. [doi: 10.2196/44528] [Medline: 37343182]

XSL•FO
RenderX

10. Medical outcomes study: 20-item short form survey instrument (SF-20). Rand Health Care. URL: https://www.rand.org/health-care/surveys_tools/mos/20-item-short-form/survey-instrument.html [accessed 2025-04-09]
11. Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. J Man Manip Ther 2009;17(3):163-170. [doi: 10.1179/jmt.2009.17.3.163] [Medline: 20046623]
12. Qualtrics. What is NPS? The ultimate guide to boosting your Net Promoter Score by Adam Bunker. URL: https://www.qualtrics.com/experience-management/customer/net-promoter-score [accessed 2024-12-03]
13. Nick JM, Roberts LR, Petersen AB. Effectiveness of telemonitoring on self-care behaviors among community-dwelling adults with heart failure: a quantitative systematic review. JBI Evid Synth 2021 Oct;19(10):2659-2694. [doi: 10.11124/JBIES-20-00329] [Medline: 33896907]
14. Huang JA, Lai CS, Tsai WC, Weng RH, Hu WH, Yang DY. Determining factors of patient satisfaction for frequent users of emergency services in a medical center. J Chin Med Assoc 2004 Aug;67(8):403-410. [Medline: 15553800]
15. Okoro UE, Harland KK, Assimacopoulos E, Findlay S. Trends in health care coverage and out-of-pocket cost barriers: a gender comparison. J Womens Health (Larchmt) 2024 Apr;33(4):473-479. [doi: 10.1089/jwh.2023.0121] [Medline: 38215276]

Abbreviations

NPS: Net Promoter Score
SF-20: 20-item Short Form Survey

Edited by A Kushniruk; submitted 16.12.24; peer-reviewed by A Perrella, O Abasilim; revised version received 07.05.25; accepted 27.05.25; published 26.08.25.

Please cite as:
Merrill M, Zbikowski SM, Jenkins GJ
Patient Satisfaction With a Comprehensive Remote Care Program for Chronic Condition Management for Adults Under Medical Care: Observational Study
JMIR Hum Factors 2025;12:e69586
URL: https://humanfactors.jmir.org/2025/1/e69586
doi:10.2196/69586

© Michael Merrill, Susan M Zbikowski, Garrett J Jenkins. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 26.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Predicting Engagement With Conversational Agents in Mental Health Therapy by Examining the Role of Epistemic Trust, Personality, and Fear of Intimacy: Cross-Sectional Web-Based Survey Study

Fanny Guglielmucci[1*], BSc, MSc, PhD, PsyD; Daniela Di Basilio[2*], BSc, MSc, PhD, PsyD

[1]Department of Philosophy, Communication and Performing Arts, Roma Tre University, Rome, Italy

[2]Faculty of Health and Medicine, Division of Health Research, Lancaster University, Sir John Fisher Drive, Lancaster, United Kingdom

[*]all authors contributed equally

Corresponding Author:
Daniela Di Basilio, BSc, MSc, PhD, PsyD
Faculty of Health and Medicine, Division of Health Research, Lancaster University, Sir John Fisher Drive, Lancaster, United Kingdom

Abstract

Background: The use of conversational agents (CAs) in mental health therapy is gaining traction due to their accessibility, anonymity, and nonjudgmental nature. However, understanding the psychological factors driving preferences for CA-based therapy remains critical to ensure ethical and effective application. Variables such as epistemic trust, attachment styles, personality traits, and fear of intimacy appear central in shaping attitudes toward these artificial intelligence (AI)–driven interventions.

Objective: This study aimed to investigate the role of epistemic trust, attachment styles, personality traits, and fear of intimacy in influencing individuals' willingness to engage with CA-based therapy.

Methods: An online survey was administered to 876 psychology students, yielding 736 responses (84.01% response rate). Variables measured included epistemic trust, attachment styles, personality traits, and fear of intimacy. A 5-point ordinal scale assessed willingness to engage in CA-based therapy. The data were analyzed using ordinal logistic regression models, including proportional odds models (POMs), nonproportional odds models (NPOMs), and partial proportional odds models (PPOMs), with residual deviance used to compare model fit.

Results: The PPOM provided the best model fit (residual deviance=3530.47), outperforming both the NPOM (deviance=6244.01) and the POM based on Brant test results indicating violations of the proportional odds assumption (χ^2_{105}=187.8; *P*<.001). In the final model (n=735), epistemic trust significantly increased willingness to engage in CA-based therapy across all ordinal thresholds (odds ratio [OR] 1.75, 95% CI 1.50, 2.03; *P*<.001). Fear of sharing demonstrated a nonuniform effect, with stronger associations at higher levels of willingness (OR 1.086; *P*=.001). Among personality traits, detachment negatively predicted CA preference (OR 0.95; *P*=.001), while psychoticism showed a positive association (OR 1.12; *P*=.003). Being single emerged as a strong predictor of preference for CA-based therapy (OR 3.717; *P*<.001). Attachment styles showed more nuanced effects. While dismissing and fearful-avoidant individuals were descriptively less inclined to engage in traditional human-based therapy, this association was nonsignificant in the case of fearful-avoidant attachment (*P*=.34) and should therefore be interpreted cautiously.

Conclusions: Epistemic trust and fear of intimacy emerged as pivotal factors influencing preferences for CA-based therapy, underscoring the role of interpersonal dynamics and emotional vulnerabilities. The findings suggest that individuals with avoidant attachment styles or maladaptive personality traits are more inclined toward AI-mediated interventions, driven by reduced fear of judgment and increased perceived safety. The relative homogeneity of the sample considered—particularly in terms of age, education level, and cultural exposure—limits the generalizability of these findings to broader or more diverse populations. Nonetheless, these insights highlight the need for ethical considerations and personalized approaches in deploying CA-based mental health tools to balance user reliance with human-centric therapeutic values.

(JMIR Hum Factors 2025;12:e70698) doi:10.2196/70698

KEYWORDS

artificial intelligence; AI; conversational agents; CA; mental health; psychotherapy; epistemic trust; personality

XSL•FO
RenderX

Introduction

Background

In the past few years, the way clinical professionals, patients, and the general population relate to emotional suffering has notably changed. A crucial role in this shift can be attributed to the rise of artificial intelligence (AI) and conversational agents (CAs). These are advanced manifestations of large language models, which are capable of analyzing and extracting knowledge from human conversations, using it to adaptively prompt a real-time answer and consequently creating multiturn human-computer dialogues. The unique ability of CAs to interact with human beings, understand their communication styles, and engage in human-like conversational responses gained notable attention. This is due primarily to their potential applications in the fields of enhancing mental well-being and providing social support [1], with CA-based tools now assuming more nuanced roles, serving not only as companions [2] but also as therapists for individuals seeking emotional support and guidance (for reviews and meta-analyses, see, for example, [3-6]). The rising trend of using CAs for improving mental health can be attributed to various factors. First, the widespread integration of AI-driven technologies in daily life has normalized human-like interactions with machines. CAs, designed with advanced natural language processing capabilities, can engage users in conversations that mimic human interaction, thereby potentially helping alleviate feelings of loneliness [2,7,8]. In addition, their anonymity and nonjudgmental nature may appeal to individuals who might feel hesitant or stigmatized about sharing their emotions with their human counterparts [1,9,10].

However, beyond these technical features, it is imperative to understand the psychological determinants that may drive individuals to rely solely, or primarily, on CA-based therapy, bypassing human interaction. This understanding is crucial for the ethical provision of mental health care, as for some (eg, people with a poor mental capacity or severe mental illness), the prevalent use of CAs may be counterproductive, for example, by eliciting dependence and overreliance on the relationship built with the CA [11,12].

Several psychological factors may influence a person's willingness to communicate exclusively with autonomous agents in order to deal with life-related painful emotions and memories. Among those, 4 factors appear to be of major importance, as pointed out by several reviews on this topic: epistemic trust, fear of intimacy, attachment strategies, and personality domains (for more details on these reviews and meta-analyses, see sections below).

Epistemic Trust in AI

Perceptions of AI capabilities and trust in its effectiveness may play a significant role in shaping individuals' attitudes toward CA-based therapy. Over the past decade, there has been an incremental growth in the number of reviews [13-16] and meta-analyses [17,18] on the factors influencing trust in AI. Riedl [14] suggested that trust in technology relies on "trust propensity," that is, the tendency to believe that the use of AI can be beneficial and that its performance will appropriately and consistently allow it to address specific tasks. However, such beliefs are not necessarily accurate, concealing a logical fallacy deeply rooted in technology idealization [19]. In the case of CA-based therapy, for example, people may overlook the possibility that relying on AI-driven support alone may compromise their capability to think critically and process information autonomously. In addition, they might be exposed to the risk of receiving information that might not be fully accurate and reliable—or might, in some cases, even be dangerous [20,21].

As Alvarado [22] pointed out, AI is primarily an "epistemic technology," designed to enhance human knowledge in different epistemic contexts (eg, inquiry and accurate prediction). To be used epistemologically, though, AI needs to "earn" our epistemic trust, that is, the tendency to believe in shared communicated information and consider the source of information reliable and trustworthy [23].

As a concept, epistemic trust has a long history that precedes the rise of AI. For instance, Fonagy et al [24] have proposed conceptualizing epistemic trust within a broader sociocommunicative perspective as the output of (early) social learning processes. In these processes, several contextual factors (ie, attachment and personality styles, mentalizing capabilities, peers, and sociocultural environment) interact with each other, influencing the way we gradually discern who is knowledgeable and trustworthy [24-26]. In today's world, AI and social media seem to play a pivotal role [27] in becoming a primary source of our knowledge. However, our ability to trust what others say— including CAs—relies on the extent to which we feel safe in our relationships with them [28]. Indeed, individual differences in the propensity to place trust in and engage with AI are underpinned by various subjective factors, including, for example, personality traits and characteristics [29,30], which could help predict the likelihood of individuals opting for AI-based therapy over human-based clinical interventions.

In this study, epistemic trust was conceptualized as the individual's predisposition to perceive the information delivered by a CA as accurate, relevant, generalizable, and safe, and therefore, the CA as a trustworthy communicator. This is in line with different definitions of epistemic trust (eg, [23,31]; reflecting both the tendency to believe in shared communicated information and to consider the source of that information as reliable and trustworthy). In doing this, this study drew from the established use of this concept in clinical psychology, where epistemic trust encompasses both the informational content and the relational qualities of the communicator. In therapeutic contexts, the trust one places in the message and in the person delivering it is widely considered inseparable, as effective psychological interventions rely on both epistemic validity and emotional resonance. Accordingly, our study treats epistemic trust as a relationally embedded construct, capturing both the individual's openness to the information provided by a CA and their implicit trust in the agent as a reliable, nonthreatening source.

Attachment-Based Personality Development and Propensity to Trust AI

Research efforts aimed at understanding attitudes toward AI have primarily concentrated on personality traits underpinning

XSL•FO
RenderX

our preferences and behaviors [32,33], leading to the identification of 2 main theoretical conceptualizations of personality and trust in AI. The first considers personality traits as predictors of trust, while the second sees trust as a mediator between personality and behavioral intentions toward AI systems (for a systematic review of the research status on personality and trust in the context of AI systems, see Riedl [14]). In this regard, the same author [14] proposed a third framework suggesting that the propensity to trust AI is a specific personality trait related to broader personality constructs (ie, the Big Five). This framework also suggests a complex role for emotions and affective states, which are influenced by personality traits and, in turn, influence the development of trust in others and how we relate to them [34].

Similarly, Fonagy and colleagues [35] recently affirmed that "attachment and personality styles are seen as communicative strategies underpinning social learning to ensure adaptation to ever-changing social situations." In other words, deeply rooted aspects of individuals' selves, such as their personality patterns and attachment styles, influence their emotional regulation skills and abilities to relate to others. The quality of our interactions with others, in turn, shapes our inner generalized perception of the availability and trustworthiness of our other people and the overall environment [27,36].

Although research in this field is still in its infancy, recent findings [28] suggest that attachment insecurity can predict a lack of trust in AI. In addition, the same study [28] found that exposure to attachment security cues (but not positive affect) via priming or nudging increased trust levels. In essence, these findings are closely akin to Fonagy's conceptualizations of epistemic trust, personality, attachment, and mentalizing skills, suggesting a possible predictive role of epistemic trust, personality, and attachment in our disposition toward CA-based therapy.

Fear of Intimacy and Avoidance-Based Attachment Patterns

Intimacy involves the capacity to commit oneself to particular affiliations and relationships, and to possess the "ethical courage" necessary to uphold these commitments [37]. Intimate relationships develop in early infant-caregiver relationships, fostering trust in others and instilling the confidence and self-esteem necessary for a child to explore their social environment [38,39]. Difficulties in building intimate relationships during adolescence and early adulthood reflect a lack of emotional integration and have been linked with a wide variety of mental health problems, including general maladjustment and personality disorders [40,41].

Currently, technology often plays a paradoxical role in fostering a sense of "virtual intimacy" while leading to potential social isolation [42,43]. This dichotomy may be underpinned by the coexistence of a desire for connection with a fear of face-to-face interactions, where intimacy and empathy are truly tested [44,45]. This dynamic may also be found in therapeutic relationships, where patients may desire, albeit feel uncomfortable, to open up about their personal lives to mental health professionals [46]. From this perspective, the desire to interact solely with AI to address psychological distress may mask a deep fear of intimacy while simultaneously reflecting the need to understand and make sense of our emotionally painful experiences.

Individuals grappling with the fear of intimacy typically exhibit characteristics rooted in 2 distinct avoidance-based attachment strategies, which correspond to 2 different models of the self: fearful and dismissing [47]. In both cases, others are perceived as untrustworthy and threatening our security. However, individuals with a fearful attachment style consciously desire social connection despite avoiding them due to a depleted sense of self-worth and a deep-seated fear of rejection. They tend not to believe they deserve love and support and mistrust others' intentions. Conversely, subjects who dismiss intimacy possess a positive self-representation and defensively deny their need or desire for intimate social contact. This defensive dismissal serves to protect their self-image and independence [47-49]. This differentiation suggests that both groups may feel more secure in an AI-based relationship and be more prone to trust CAs rather than human beings (eg, therapists), albeit for different reasons: the former (fearfully attached) because of the fear of opening up to someone who might hurt, judge, or reject them, while the latter (avoidantly attached) as they value connections that allow them to preserve their positive self-image and independence [50-52].

In the past few years, an increasing reliance on CAs in mental health therapy has emerged [1], driven by their advanced ability to simulate human-like interactions and their appeal as anonymous, nonjudgmental support systems. While the technical capabilities of CAs have been extensively explored, less is understood about the psychological factors influencing individuals' preference for CA-based therapy over human interaction. Key constructs such as epistemic trust, attachment strategies, personality traits, and fear of intimacy have been identified as critical determinants of this preference.

Considering this, this study aimed to address a gap in the literature by examining how key psychological factors—specifically epistemic trust, personality traits, avoidance-based attachment styles, and fear of intimacy—influenced individuals' willingness to engage with CAs for mental health support. The goal was to contribute to a deeper understanding of the psychological underpinnings that shape preferences for CA-based therapy, with broader implications for the ethical implementation of AI in mental health care. To achieve this, we employed an ordinal logistic regression model to assess the predictive role of these factors in shaping individuals' dispositions toward AI-mediated therapeutic interventions.

Methods

Recruitment and Sample

Data were collected through an online survey distributed to students attending the MSc course "Psychology of Communication" at the Department of Philosophy, Communication and Performing Arts of Rome3 University (Italy). The study was advertised by posting an e-copy of the study flyer on the academic page and Microsoft Teams channel

XSL•FO
RenderX

of the MSc course, reaching out to a pool of 876 students. Participants accessed the survey through a Google Forms link, where they were presented with a Participation Information Sheet (PIS) outlining study details, that is, a concise overview of the study's objectives, participation requirements, data handling procedures, and contact information for the principal investigator, followed by a consent form.

Ethical Considerations

The participants were asked to confirm their agreement to participate by clicking "I agree" on the consent form. Selecting this option to express informed consent redirected participants to another form containing the assessment measures used as part of this study. Participants remained anonymous and could withdraw from the study at any time before data analysis by contacting the principal investigator through email as described in the PIS. The online survey used in this study was designed and implemented in accordance with the CHERRIES (Checklist for Reporting Results of Internet E-Surveys [53]) guidelines, ensuring methodological transparency, ethical compliance, data integrity, and robust participant engagement throughout the process. Table 1 reports the CHERRIES criteria and how they were met for this specific study. All procedures performed in studies involving human participants were in accordance with the ethical standards of Italian Psychoanalytical Association (ethical approval 0405/2023) and with the Declaration of Helsinki of 1975, revised in 2013.

Table . CHERRIES[a] compliance.

CHERRIES criteria	Study implementation details
Design	Cross-sectional web survey design using Google Forms.
IRB[b] approval and informed consent process	IRB approval obtained (AIPsi[c] ethical approval 0405/2023); informed consent collected through an online form.
Development and pretesting	The survey was pretested internally by the research team; translated scales were validated via back-translation.
Recruitment process and description of the sample having access to the questionnaire	Participants were MSc students recruited through an e-flyer advertising the study, posted on the MSc course Microsoft Teams channel and University webpage; 876 were invited.
Survey administration	Google Forms used; PIS[d] and consent form provided before questionnaire; anonymity and withdrawal rights ensured.
Response rate	736 out of 876 responded (84.01% response rate); 735 included in final analysis after removing 1 outlier.
Preventing multiple entries from the same individual	Participants could only submit the form once; the survey required institutional login preventing duplicate submissions.
Analysis	Ordinal logistic regression models used; assumptions tested; best-fit model selected (PPOM[e]); power analysis conducted.

[a]CHERRIES: Checklist for Reporting Results of Internet E-Surveys.

[b]IRB: institutional review board.

[c]AIPsi: Italian Psychoanalytical Association.

[d]PIS: Participation Information Sheet.

[e]PPOM: partial proportional odds model.

Measures

Sociodemographic Information and Willingness to Rely on CA-Based Therapy

Sociodemographic information (eg, age, gender, years of education, and marital status) was collected. To assess the preference for CA-based therapy over human being-based clinical interventions, participants were asked to answer the following question: "On a scale from 1 to 5, how willing are you to participate in psychotherapy sessions solely facilitated by artificial intelligence (AI), without any interaction with a human psychologist?" (scores range from 1=not at all to 5=extremely).

Epistemic Trust

The Italian adaptation [54] of the Epistemic Trust, Credulity and Mistrust Questionnaire (ETCMQ [31]) was used to assess epistemic trust. This is a 15-item self-report measure with a 3-factor structure assessing various forms of epistemic stance (ie, epistemic trust, credulity, and mistrust). Items were adapted and referred specifically to CA-based therapy (ie, When I have a personal problem, I usually ask AI for advice; If I don't know what to do, my first impulse is to ask AI because I trust). In this study, Cronbach α for ETCMQ was 0.95.

Attachment

The Italian translation of the Relationship Questionnaire [55] (RQ [47]) was used to assess attachment styles. It is a 4-item self-report measure that describes 4 prototypical attachment attitudes (secure, dismissing, preoccupied, and fearful), linked on a 7-point Likert scale. In this study, Cronbach α for RQ was 0.97.

XSL•FO
RenderX

Personality Domains

The Italian version [56] of the Personality Inventory for DSM-5—Brief Form—Adult (PID-5-BF [57]) was used to evaluate personality functioning. The PID-5-BF comprises 25 self-reported items designed to assess 5 maladaptive personality domains: negative affectivity, detachment, antagonism, disinhibition, and psychoticism, in accordance with the alternative *DSM-5* model for personality disorders [58]. Participants rate each item on a 4-point Likert scale, with higher scores indicating greater personality dysfunction. In this study, Cronbach α for PID-5-BF was 0.87. As for the Italian version of the RQ [54], the psychometric properties and validation processes of the translated version of the PID-5 are detailed in the original validation studies.

Fear of Intimacy With Helping Professions

Originally developed to assess the comfort of revealing intimate personal details to helping professions, the Fear of Intimacy With Helping Professionals Scale (FIS-HP [46]), is a modified version of the 35-item Fear of Intimacy Scale (FIS) [59]. The FIS-HP has 18 items grouped into 3 factors, each formed by 6 items, named fear of sharing, openness to intimate sharing, and information sharing. Although a validated Italian translation of this scale is not available, in this study, the original FIS-HP [42] was translated into Italian by 2 bilingual psychologists and researchers and then back-translated by an independent translator to verify the equivalence of the translated scale to the original one, in line with the process followed to obtain the Italian standardized version of the 35-item FIS [60]. In this study, Cronbach α for FIS-HP was 0.94.

Statistical Analysis

Power

To determine sample size, we computed an a priori power analysis for logistic regression using G*Power (version 3.1.9.2; Heinrich-Heine-Universität Düsseldorf) [61] as described by Aysel et al [62], which resulted in a recommended sample size of n=683 (α=.05; β=.95; odds ratio [OR] 1.3, 95% CI 1.15-1.47; critical *z* score=1.64).

Data Analytics Strategy

Given the polytomous and ordinal nature of the dependent variable (ie, willingness to address mental health conditions only via CA-based interventions, assessed using 5 categories that may be ordered according to the level of magnitude), we performed an ordinal logistic regressive model using the polr function in the MASS package of R (R Core Team, 2025), following the procedure described by Liang and Zhan [63]. This modeling approach was selected because the POM is particularly well-suited for handling ordinal outcome variables, where the distance between response categories cannot be assumed to be equal. Unlike traditional methods such as ANOVA, which require continuous dependent variables and assume homogeneity of variance across groups, POM retains the ordinal structure of the outcome and models the cumulative probability of being in or above a particular category. This allows for a more accurate and interpretable analysis of ranked responses. Furthermore, POM enables the inclusion of multiple predictor variables—both continuous and categorical—and yields interpretable outputs in the form of ORs, which is particularly advantageous when assessing complex psychological and demographic predictors. Thus, the POM offers a more robust and statistically appropriate framework for evaluating the ordinal response structure inherent to participants' preferences for CA-based therapy [64,65].

Probabilities of the ordinal response variable were transformed into log-odds via logit function, as shown in Equation 1.

Equation 1: Logistic transformation of cumulative probabilities.

$$(1)\text{logit}(P(Y\geq j))=\ln\left[(P(Y\geq j)/1-P(Y\geq j))\right]=\beta 0j+\beta 1jX$$

In logistic regression, log-odds indicate the likelihood of being at or below a certain category level relative to being above it. A positive logit value suggests a higher likelihood of being in or below a particular category, while a negative value indicates a higher likelihood of being above that category. To express odds in terms of predicted probabilities, which are often easier to interpret, we used the transformation formula:

Equation 2: Conversion from log-odds to predicted probabilities.

$$(2)P(Y\geq j)=\exp(\beta 0j+\beta 1jX)/[1+\exp(\beta 0j+\beta 1jX)]$$

Where P(Y≥j) represents the predicted probability of the dependent variable Y being in or below category level *j*, *e* is the base of the natural logarithm (2.718), β_{0j} and β_{1j} are, respectively, intercept and coefficient for the independent variable X. As gender is a dummy variable, we used the following formula:

Equation 3: Predicted probability for dummy variable (gender).

$$(3)P(Y\geq j)=\exp(\beta 0j+\beta 1j\times\text{Gender})/[1+\exp(\beta 0j+\beta 1j\times\text{Gender})]$$

We used the POM, considering the effects of independent variables (IVs) on the dependent variable as equal across all categories. For instance, the OR of each IV is the same for each cut-off – (1) P[Y=1] versus P[Y≥2], (2) P[Y≤2] versus P[Y≥3], (3) P[Y≤3] versus P[Y≥4], and (4) P[Y≤4] versus P[Y≥5]. POM is the most frequently used logistic regression model for analyzing ordinal response variables [64,65]. However, it leads to strong assumptions that, if violated, may lead to incorrect data interpretations [66]. Checking assumptions is crucial to ensuring the validity of statistical analyses and to justify the choice of a model above others [67]. For example, in their review of clinical psychology studies involving linear regression, Ernst and Albers [68] found that only 2% of cases were both transparent and correct in reporting assumption checking, and a further 6% were transparent but incorrect, leading to serious inferential problems [68].

To test the proportional odds assumption (or parallel line assumption), we conducted the Brant test via the Brant R package [69]. If the *P* value of the Brant test is lower than .05, the proportional odds assumption should be considered violated. When data fail to satisfy the proportional odds assumption, a solution is fitting a NPOM or the PPOM [70]. The main difference between NPOM and PPOM is that while the first allows each of the IVs to vary across each category of ordinal response variable, the second just allows to vary only IVs that have violated the assumption.

XSL•FO
RenderX

Results

Out of 876 students, 736 students (84.01% respondents) aged 18 to 85 years completed the survey. As outliers increase data variability and decrease statistical power, 1 participant (male, age 85 y) was manually removed. The final sample consisted of 735 participants (191 males, 25.98%; 543 females, 73.88%; and 1 missing) with a mean age of 27.69 (SD 12.31) years. The single case with missing gender data was handled using multiple imputation, ensuring that no data were excluded from the analyses due to this missing value. Although the age range of the sample was broad (18 - 85 y), the distribution was moderately right-skewed, with a median age of 24 (IQR 19 - 31) years, reflecting the student population from which the sample was drawn. The majority (n=561, 76.36%) were single or in a nonserious relationship, and the remaining 414 (56.3%) were married or in a stable union. On average, participants had 13.84 (SD 3.50) years of education. A total of 76 individuals (10.30%) reported a previous mental health condition, such as anxiety or depression; 127 (17.3%) reported frequent substance use (≥3 times/week) without meeting the criteria for abuse or dependence.

Table 2 shows the POM model results.

The results of the Brant test show that the overall test fails, and the parallel line assumption is not satisfied (χ^2_{105}=187.8; P<.001). Thus, the results of POM should not be interpreted nor generalized, as some predictors might have nonuniform effects across different levels of the outcome variable, necessitating a more flexible modeling approach. In light of this, possible alternative models to explain the data were considered, and Tables 3 and 4 report the results of such alternative models, respectively NPOM and PPOM.

After partially lifting the restriction, both PNOM and PPOM were models found to fit data correctly and may be useful to explain them. To identify the best solution for this purpose, we compared NPOM and PPOM by using residual deviance, which may be considered a measure of discrepancy of a generalized linear model from real observed data [71]. Thus, the lower the value, the smaller the discrepancy and the better the model can predict the value of the response variable. Residual deviance was computed by the VLMC R package [72], which allows for the optimization of the Akaike information criterion (AIC) or Bayesian information criterion (BIC) and overcomes problems related to residuals for polytomous response data by generating data through a Monte Carlo Chain. Results showed that PPOM had a better fit with respect to NPOM (r_D=3530.472 vs 6244.009) and was therefore selected as the chosen model. More specifically, the choice of PPOM over the NPOM was justified by several considerations. First, the residual deviance for PPOM (3530.472) was significantly lower than that of NPOM (6244.009), indicating that PPOM provided a better fit to the observed data. While NPOM relaxed the proportional odds assumption entirely, allowing all predictors to vary across response levels, this approach risked overfitting and reduced parsimony, especially when predictors with no significant nonproportionality were unnecessarily freed. PPOM, on the other hand, selectively relaxed the proportional odds assumption only for predictors that violated it, as identified through the Brant test. This selective relaxation struck a balance between addressing assumption violations and maintaining interpretability, which was a key advantage over the more complex NPOM. By preserving the proportional odds assumption for variables that met it, PPOM simplified interpretation while capturing nuanced effects for predictors like relationship status and fear of sharing. The findings from PPOM offered unique insights. For instance, epistemic trust demonstrated a consistent effect across all response levels, underscoring its robust role as a predictor of preference for CA-based therapy. Fear of sharing, however, exhibited a nonuniform effect that intensified as willingness thresholds increased. Similarly, PPOM revealed a gradient effect for relationship status, where single individuals showed the strongest preference for CA-based therapy at lower thresholds, a pattern that diminished at higher thresholds. Traits like detachment and psychoticism also varied across levels, with detachment exerting a weaker influence at lower willingness thresholds and becoming more pronounced at higher thresholds. These level-specific effects, which were not as clearly identifiable in NPOM, underscored the nuanced contributions of PPOM. Therefore, the model provided a more precise and actionable understanding of the factors influencing CA-based therapy preferences, balancing statistical rigor with interpretability and offering practical insights for targeted interventions.

XSL•FO
RenderX

Table . Proportional odds model results.

Variables	OR[a]	SE	*z* value	*P* value
Intercept 1	1.400	0.853	0.395	.69
Intercept 2	0.852	0.851	−0.188	.85
Intercept 3	0.367	0.851	−1.180	.24
Intercept 4	0.155	0.853	−2.188	.03
Gender (male)	0.724	0.204	−1.585	.11
Age (years)	1.000	0.009	0.005	≥.99
Education (years)	1.082	0.025	3.170	.002
Single	3.717	0.389	3.378	<.001
Married or in a stable union	2.751	0.34	2.982	.003
Not in a serious relationship	1.572	0.747	0.605	.55
ETMCQ[b]				
Epistemic trust	1.745	0.434	4.017	<.001
Epistemic mistrust	0.727	0.52	1.398	.16
Epistemic credulity	0.677	0.675	1.002	.32
PID5−BF[c]				
Negative affectivity	1.014	0.056	0.253	.80
Detachment	0.95	0.065	14.615	.001
Antagonism	0.89	0.07	12.714	.001
Disinhibition	1.05	0.06	17.5	.002
Psychoticism	1.12	0.08	14.0	.003
RQ[d]				
Secure	1.040	0.054	0.731	.47
Fearful-avoidant	0.945	0.059	−0.955	.34
Preoccupied	0.923	0.065	−1.23	.22
Dismissing-avoidant	0.952	0.055	−0.896	.37
FIS−HP[e]				
Fear of sharing	1.086	0.025	3.24	.001
Openness to intimate sharing	1.015	0.025	0.591	.56
Sharing information	0.988	0.029	−0.406	.69

[a]OR: odds ratio.

[b]ETCMQ: Epistemic Trust, Credulity and Mistrust Questionnaire.

[c]PID5−BF: Personality Inventory for DSM-5—Brief Form—Adult.

[d]RQ: Relationship Questionnaire.

[e]FIS-HP: Fear of Intimacy with Helping Professionals Scale

XSL•FO
RenderX

Table . Results of the nonproportional odds model (NPOM).

Variables	logitlink(P[Y>=2])[a]	logitlink(P[Y>=3])	logitlink(P[Y>=4])	logitlink(P[Y>=5])
Intercept[b]	0.141	0.111	0.110	0.319
Gender (male)[b]	0.613	0.855	0.823	0.766
Age (years)[b]	0.978	0.996	0.993	1.004
Education (years)[b]	1.132	1.112	1.089	1.063
Single[b]	10.309	6.469	7.862	2.627
Married/stable union[b]	6.074	3.877	5.452	2.268
Not in a serious relationship[b]	2.662	3.068	2.199	1.465
ETMCQ[c]				
Epistemic trust[b]	1.117	1.114	1.111	1.070
Epistemic mistrust	0.85	0.88	0.87	0.89
Epistemic credulity[b]	1.154	1.137	1.099	1.076
PID5-BF[d]				
Negative affectivity[b]	1.110	1.083	1.100	0.988
Detachment[b]	0.877	0.902	0.910	1.011
Antagonism	0.141	0.111	0.110	0.319
Disinhibition	1.05	1.07	1.04	1.03
Psychoticism	1.10	1.08	1.06	1.07
RQ[e]				
Secure	0.948	1.037	1.089	1.028
Fearful-avoidant[b]	1.027	0.959	0.916	0.942
Preoccupied[b]	0.881	0.837	0.963	0.964
Dismissing-avoidant[b]	1.021	0.955	1.000	0.946
FIS-HP[f]				
Fear of sharing[b]	1.114	1.099	1.023	0.919
Openness to intimate sharing	1.241	1.182	1.124	0.969
Sharing information	0.969	0.968	1.002	0.965

[a]Logit link values represent log-odds coefficients from the ordinal logistic regression model, indicating the effect of each predictor on the likelihood of expressing greater willingness to engage in conversational agent–based therapy (measured on a 5-point ordinal scale). Positive coefficients indicate a higher probability of choosing more favorable response categories (e.g., “somewhat willing” or “extremely willing”), while negative coefficients reflect a tendency toward lower willingness.

[b]$P<.001$.

[c]ETCMQ: Epistemic Trust, Credulity and Mistrust Questionnaire.

[d]PID5−BF: Personality Inventory for DSM-5—Brief Form—Adult.

[e]RQ: Relationship Questionnaire.

[f]FIS-HP: Fear of Intimacy with Helping Professionals Scale

XSL•FO
RenderX

Table . Results of the partial proportional odds model (PPOM).

Variables	logitlink(P[Y>=2])	logitlink(P[Y>=3])	logitlink(P[Y>=4])	logitlink(P[Y>=5])
Intercept[a]	0.190	0.260	0.160	0.270
Gender (male)[a]	0.770	0.770	0.770	0.770
Age (years)[a]	1.000	1.000	1.000	1.000
Single[a]	18.03	8.020	9.830	2.420
Married or in a stable union[a]	7.710	4.540	6.100	2.170
Not in a serious relationship[a]	1.490	2.310	1.810	1.590
ETMCQ[b]				
Epistemic trust[a]	1.24	1.24	1.24	1.24
Epistemic mistrust[a]	0.93	0.93	0.93	0.93
Epistemic credulity[c]	1.09	1.09	1.09	1.09
PID-5BF[d]				
Negative affectivity[a]	1.010	1.010	1.010	1.010
Detachment[a]	0.970	0.970	0.970	0.970
Antagonism	0.950	0.950	0.950	0.950
Disinhibition	0.940	0.940	0.940	0.940
Psychoticism[a]	0.950	0.950	0.950	0.950
RQ[e]				
Secure	1.040	1.040	1.040	1.040
Fearful-avoidant[a]	0.940	0.940	0.940	0.940
Preoccupied[a]	0.950	0.950	0.950	0.950
Dismissing-avoidant[a]	0.960	0.960	0.960	0.960
FIS-HP[f]				
Fear of sharing[a]	1.03	1.03	1.02	1.01
Openness to intimate sharing	0.97	0.97	0.97	0.97
Sharing information	1.00	1.00	1.00	1.00

[a]*P*<.001.

[b]ETCMQ: Epistemic Trust, Credulity and Mistrust Questionnaire.

[c]*P*<.01.

[d]PID-5BF: Personality Inventory for DSM-5—Brief Form—Adult.

[e]RQ: Relationship Questionnaire.

[f]FIS-HP: Fear of Intimacy with Helping Professionals Scale

Overall, our data showed that epistemic trust emerged as a strong predictor of willingness to rely on CA-based therapy. Higher levels of epistemic trust significantly increased the likelihood of choosing AI-mediated therapy (OR 1.745, 95% CI 1.10 - 1.41; *P*<.001). Fear of intimacy, particularly the subdomain fear of sharing, also demonstrated significant predictive power (OR 1.086, 95% CI 1.01 - 1.06; *P*=.001). These findings suggest that individuals with higher epistemic trust and fear of interpersonal vulnerability are more inclined toward AI interventions. Attachment styles showed differential associations with preferences. Secure attachment was not a significant predictor, while avoidant attachment styles—particularly dismissing and fearful—were negatively associated with willingness to engage in CA-based therapy, although these effects were not statistically significant for fearful-avoidant style (*P*=.34), and should be interpreted cautiously. Personality traits, including negative affectivity, detachment, and psychoticism, were significant predictors of

XSL•FO
RenderX

preference for CA-based therapy. For instance, detachment was associated with lower willingness (OR 0.95, 95% CI 0.94 - 1.00; *P*<.001), while psychoticism predicted higher willingness (OR 1.12, 95% CI 1.04 - 1.21; *P*=.003), suggesting that individuals with certain traits may prefer less emotionally intense interaction modalities. Demographic variables also contributed to the prediction model. Men were less likely than women to prefer CA-based therapy, although this effect was not statistically significant (*P*=.11). Education was positively associated with willingness to engage in CA therapy (OR 1.082, 95% CI 1.03 - 1.14; *P*=.002), suggesting that more educated individuals may perceive AI as a viable therapeutic option. Although men were less likely than women to prefer CA-based therapy (OR 0.724, 95% CI 0.49 - 1.08), this effect was not statistically significant (*P*=.11). Finally, relationship status showed robust associations: single participants were significantly more likely to prefer CA-based therapy (OR 3.717, 95% CI 1.73 - 7.97; *P*<.001), as were those married or in stable unions (OR 2.751, 95% CI 1.41 - 5.36; *P*=.003), when compared to those in less defined relationships. In sum, our analyses underscored the multifaceted role of epistemic trust, attachment, personality traits, and demographics in shaping individuals' preferences for CA-based mental health interventions. These findings have implications for tailoring AI-based mental health services to individual needs and addressing potential barriers to their acceptance.

Discussion

Principal Findings

This study examined the psychological and demographic factors influencing individuals' preferences for conversational agent (CA)-based therapy over traditional human-mediated interventions. Using robust statistical approaches, we identified significant predictors, including epistemic trust, attachment styles, personality traits, fear of intimacy, and demographic factors such as gender and education. More specifically, greater epistemic trust significantly increased individuals' willingness to engage with CA-based mental health therapy, highlighting its central role as a precondition for trusting AI-mediated support. Higher levels of fear of intimacy, particularly fear of sharing personal information, also predicted a preference for CA-based support, suggesting that individuals who feel emotionally vulnerable or threatened by interpersonal closeness may gravitate toward less socially demanding therapeutic tools. Personality traits such as detachment and psychoticism were also associated with greater willingness to use AI-mediated therapy, reflecting a preference for emotionally distant or unconventional interaction styles. Attachment styles showed more nuanced effects: individuals with avoidant attachment (both dismissing and fearful) were less inclined to engage in traditional human-based therapy, indirectly favoring CA-based interventions. Demographic predictors such as being single and having a higher level of education further contributed to increased willingness to use CA-based mental health support. These findings align with broader trends in the literature. Recent studies demonstrate that user trust in AI mental health tools is influenced not only by personal traits but also by perceptions of safety, anonymity, and emotional support [3] For example, Abd-Alrazaq et al [71] found that emotional comfort and perceived nonjudgmentality foster trust in CAs - mirroring our finding that epistemic trust and fear of intimacy shape AI preferences. Similarly, a validated model by Lee et al [73] emphasizes the roles of affect, social norms, and habitual use in shaping engagement with well-being chatbots, supporting the notion that behavioral and contextual factors interact with internal dispositions, such as attachment and personality. Furthermore, a recent expert on AI-driven CAs for mental health support [12] raised concerns about user overreliance and the limited emotional depth of chatbot responses, highlighting ethical risks, particularly for individuals whose psychological profiles may predispose them to overuse AI to avoid vulnerability. A scoping review by Rahsepar Meadi et al [74] further underscored the need for transparency, informed consent, and safeguarding in AI therapy design, especially for emotionally avoidant users. Finally, a study analyzing users' reviews of CA-based apps for mental health [75] found that users are more likely to disclose sensitive information to chatbots due to their perceived anonymity, reinforcing our interpretation that individuals with a fear of intimacy or high detachment may favor CAs as emotionally safer alternatives. Together, these converging lines of evidence reinforce our conclusion that psychological vulnerabilities and individual differences significantly influence engagement with AI-based mental health interventions. This underscores the importance of tailoring such tools to users' psychological needs while also addressing ethical, emotional, and relational implications of replacing human intimacy with machine-mediated interaction in mental health care [74]. The following sections delve deeper into the specific roles of epistemic trust, attachment styles, personality traits, and fear of intimacy, providing a more granular analysis of how each contributes to shaping individual engagement with AI-mediated therapy.

Epistemic Trust and the Role of AI

The results underscored the centrality of epistemic trust in shaping individuals' willingness to engage with CA-based therapy. Higher epistemic trust significantly predicted a greater preference for CAs (OR 1.745; 95% CI 1.31–2.32; *P*<.001), affirming its role as a key determinant of trust in AI-mediated interventions. This finding aligns with previous work [68], which emphasized the importance of trust in fostering acceptance of AI technologies. Our results suggest that individuals who perceive AI as a reliable source of information and support are more inclined to rely on it for mental health therapy, echoing the epistemic relationship outlined by Alvarado [22], who highlighted AI's capacity to enhance human knowledge in epistemic contexts. These findings also resonate with the conceptualization of epistemic trust as a product of early social learning [22,69], which influences how individuals assess the reliability of new sources of knowledge, including AI. This study contributes to the growing evidence that AI systems, when perceived as epistemically trustworthy, can be viable alternatives to human therapists. However, it also raises ethical considerations about overreliance on AI and the potential for reduced critical thinking, as highlighted in recent debates on AI's role in health care [70].

XSL•FO
RenderX

Attachment Styles and Fear of Intimacy

Attachment styles, particularly avoidant patterns, emerged as significant predictors of CA-based therapy preference. Individuals with dismissing and fearful attachment styles were less likely to engage with traditional human therapists, likely due to discomfort with emotional intimacy and vulnerability. These findings align with Bartholomew and Horowitz's [43] model of attachment, which posits that dismissing individuals prioritize independence and avoid reliance on others, while fearful individuals desire connection but struggle with trust and fear of rejection. The preference for CAs may reflect a less threatening avenue for emotional expression, consistent with the argument that digital platforms offer "safe spaces" for intimacy without the pressures of face-to-face interaction [75]. Fear of intimacy, particularly fear of sharing, also significantly influenced preferences for CA-based therapy (OR 1.09, 95% CI 1.03–1.14; *P*=.001). This supports previous findings that individuals who struggle with interpersonal vulnerability may gravitate toward less invasive therapeutic options. Ingersoll et al [42] similarly found that fear of intimacy negatively impacts the therapeutic alliance, potentially hindering the effectiveness of traditional therapy. These dynamics underscore the need for CA-based therapies to account for such fears while ensuring they provide adequate emotional support and empathy. Interestingly, while dismissing and fearful-avoidant attachment styles were negatively associated with willingness to engage in CA-based therapy, consistent with their underlying discomfort with closeness or trust, secure attachment did not significantly predict CA preference. This may reflect the adaptive flexibility of securely attached individuals, who are typically comfortable with both intimacy and autonomy [76,77]. As such, they may evaluate AI and human-delivered interventions based more on perceived usability or relevance than on relational needs or avoidance patterns. In this context, their openness to various forms of support may lead to a neutral or balanced preference, rather than a strong inclination toward or against CA-based therapy. Future studies might explore this by directly comparing secure individuals' preferences across therapeutic formats under varying emotional and contextual demands.

Personality Traits and Preferences for AI Therapy

Personality traits, particularly negative affectivity, detachment, and psychoticism, were significant predictors of willingness to engage with CA-based therapy. These traits reflect tendencies toward emotional dysregulation, social withdrawal, and unconventional thinking, which may drive preferences for less interpersonal forms of therapy. The findings align with Marengo et al [30] who noted that personality traits significantly influence how individuals interact with digital technologies, including AI. The association between detachment and CA preference (OR=0.95, *P*<.001) is particularly notable. Detachment often involves avoidance of emotional closeness, making traditional therapy challenging. CAs, perceived as emotionally neutral, may offer a more comfortable alternative. This supports prior research by Matthews et al [27] who remarked that personality traits affect individuals' comfort with AI, particularly in emotionally charged contexts like therapy. Similarly, the significant role of psychoticism (OR 1.12, 95% CI 1.04–1.21; *P*=.003) may reflect the appeal of AI-mediated therapy for individuals with unconventional cognitive styles who might find human interactions overwhelming or judgmental.

Demographic Influences on CA-Based Therapy Preferences

Demographic factors, including gender, education, and relationship status, also played a role in shaping preferences for CA-based therapy. While gender differences were not statistically significant, the trend suggested that men were slightly more inclined toward AI therapy than women. This aligns with previous findings suggesting that men are generally more open to adopting new technologies [28]. Education level was positively associated with willingness to engage in CA therapy, suggesting that individuals with higher education levels may perceive AI as a credible and innovative tool for mental health support. These findings corroborate Glikson and Woolley's [68] assertion that technological literacy influences the acceptance of AI in various domains. Relationship status also emerged as a significant predictor, with single individuals showing a markedly higher preference for CA-based therapy compared to those in stable unions (OR 3.72, 95% CI 2.37–5.82; *P*<.001). This may reflect the increased reliance on digital solutions for emotional support among individuals with less access to intimate interpersonal relationships, consistent with findings by Loveys et al [2] on the role of digital technologies in mitigating loneliness.

Implications for Practice and Policy

The findings have several practical and ethical implications. First, they highlight the need for CA-based therapies to be designed with sensitivity to individual differences, particularly attachment styles and personality traits. For instance, individuals with high detachment may benefit from interventions that gradually build emotional engagement, while those with high epistemic trust may prioritize transparency and reliability in AI responses. Based on these findings, we recommend that developers incorporate adaptive interaction styles into CA systems, such as pacing strategies, tailored disclosure prompts, and emotionally neutral yet supportive communication styles, to accommodate users who struggle with emotional closeness or fear of judgment. Second, our results underscore the importance of addressing the potential risks associated with CA-based therapy. While these technologies offer valuable alternatives, particularly for individuals who struggle with traditional therapy, they also carry risks of overreliance and reduced critical thinking [74]. As discussed by Abd-Alrazaq et al [3], ensuring that CA-based therapies are used as adjuncts rather than replacements for human interaction is critical to maintaining a balanced approach to mental health care. Third, the role of demographic factors suggests that targeted outreach and education efforts may be necessary to increase the accessibility and acceptance of CA-based therapies across diverse populations. For example, addressing technological skepticism among older adults and less educated individuals could help bridge gaps in mental health support. Furthermore, we suggest that mental health practitioners consider screening tools for psychological traits such as epistemic trust, attachment insecurity, and fear of sharing, to help determine whether CA-based interventions are suitable entry points or supportive

XSL•FO
RenderX

complements to traditional therapy. CAs could then be integrated into care in a way that accounts for users' emotional readiness and interpersonal vulnerabilities. Developers might also consider implementing real-time monitoring features within CA systems to detect signs of excessive dependency or emotional withdrawal, prompting referral to human practitioners when necessary. Finally, the findings raise important ethical questions about the integration of AI in mental health care. As highlighted in a recent review [78], the design and implementation of AI systems must prioritize patient safety, privacy, and informed consent. These concerns are echoed in emerging frameworks for responsible AI integration in mental health research, which emphasize transparency, stakeholder inclusion, and regulatory oversight as essential components of ethical deployment [79]. These considerations are particularly important given the reliance on self-reported data and the potential for biases in AI algorithms. Taken together, our findings provide a foundation for designing CA systems that are both ethically responsible and psychologically attuned to the individual needs of users.

Strengths, Limitations, and Future Directions

This study offered a novel contribution to the emerging field of digital mental health by focusing not only on the technological acceptability of CAs, but also on the psychological determinants that underpin individual preferences for AI-mediated therapy. While previous research has highlighted user satisfaction, perceived usefulness, or the technical capabilities of CAs, relatively little attention has been paid to why certain individuals are more psychologically predisposed to prefer AI over human interaction in therapeutic contexts. By integrating constructs such as epistemic trust, fear of intimacy, attachment style, and maladaptive personality traits, this study provides a more comprehensive and person-centered model of CA engagement. This approach fills a critical gap in the literature and offers clinically relevant insights for tailoring CA-based interventions to the psychological needs and vulnerabilities of different user profiles. Nonetheless, while providing valuable insights, this study is not without limitations. The sample was drawn from a specific population (university students), which may limit the generalizability of the findings. More specifically, the participants were all enrolled in a psychology degree program, which likely reflects a cohort that is highly educated, psychologically literate, and potentially more open to digital therapeutic tools than the general population. This homogeneity may have biased responses toward higher levels of epistemic trust and engagement with AI, which could inflate associations between psychological variables and preference for CA-based therapy. As noted in broader critiques of online survey methodology (eg, Andrade [80]), such sampling bias is common and warrants careful consideration when interpreting results. Future studies should replicate this work using more diverse and representative populations to validate and extend the findings. In addition, the reliance on self-reported measures introduces the possibility of response bias, particularly in relation to sensitive topics such as mental health history or substance use. While all participants in this study were aware of what CAs and CA-based therapy entail, due to their academic exposure to topics including the definition, clinical implications, and ethical considerations of AI-mediated therapy, this context should not be conflated with broader public understanding. Although we did not assess participants' personal use of AI tools for psychological support, this was a deliberate methodological choice, as the study aimed to investigate individuals' willingness to engage with CA-based therapy, not their actual behavioral engagement. Nonetheless, individual differences in previous exposure to AI may have shaped interpretations, and this represents an important avenue for future research. Finally, we acknowledge that although the study was grounded in established psychological theory, its cross-sectional design limits the ability to draw causal conclusions. While we refer to associations between constructs such as epistemic trust, attachment style, and CA preference, these relationships should be interpreted as correlational. Longitudinal or experimental designs are needed to determine causal pathways or mediating processes underlying these associations. Looking ahead, while this study focused on individual psychological predictors, future work could explore the interaction between user characteristics and system-specific features, such as CA interactivity, tone, or anthropomorphism. Investigating how these elements jointly influence engagement and outcomes could offer richer insights into optimizing AI-supported mental health interventions across diverse user groups. This priority also aligns with recent reviews (eg, Beg et al [81]) calling for research that connects user characteristics with CA functionalities, including interface style, responsiveness, and therapeutic framing.

Conclusion

This study contributes to the growing literature on AI-mediated mental health interventions by identifying key psychological and demographic predictors of preferences for CA-based therapy. The findings emphasize the importance of epistemic trust, attachment styles, personality traits, and demographic factors in shaping attitudes toward AI in mental health care. While these technologies offer promising alternatives to traditional therapy, their integration must be approached with caution to ensure ethical and effective mental health support. Future research should continue to explore the dynamic interactions between individual differences and technological features, paving the way for personalized and equitable mental health interventions.

Conflicts of Interest

None declared.

References

XSL•FO
RenderX

1. Ma Z, Mei Y, Su Z. Understanding the benefits and challenges of using large language model-based conversational agents for mental well-being support. AMIA Annu Symp Proc 2023;2023:1105-1114. [Medline: 38222348]
2. Loveys K, Sebaratnam G, Sagar M, Broadbent E. The effect of design features on relationship quality with embodied conversational agents: a systematic review. Int J of Soc Robotics 2020 Dec;12(6):1293-1312. [doi: 10.1007/s12369-020-00680-7]
3. Abd-Alrazaq AA, Rababeh A, Alajlani M, Bewick BM, Househ M. Effectiveness and safety of using chatbots to improve mental health: systematic review and meta-analysis. J Med Internet Res 2020 Jul 13;22(7):e16021. [doi: 10.2196/16021] [Medline: 32673216]
4. He Y, Yang L, Qian C, et al. Conversational agent interventions for mental health problems: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2023 Apr 28;25:e43862. [doi: 10.2196/43862] [Medline: 37115595]
5. Li H, Zhang R, Lee YC, Kraut RE, Mohr DC. Systematic review and meta-analysis of AI-based conversational agents for promoting mental health and well-being. NPJ Digit Med 2023 Dec 19;6(1):236. [doi: 10.1038/s41746-023-00979-5] [Medline: 38114588]
6. Beg MJ, Verma MK. exploring the potential and challenges of digital and AI-driven psychotherapy for ADHD, OCD, schizophrenia, and substance use disorders: a comprehensive narrative review. Indian J Psychol Med 2024 Dec 14:02537176241300569. [doi: 10.1177/02537176241300569] [Medline: 39677516]
7. Rodríguez-Martínez A, Amezcua-Aguilar T, Cortés-Moreno J, Jiménez-Delgado JJ. Qualitative analysis of conversational chatbots to alleviate loneliness in older adults as a strategy for emotional health. Healthcare (Basel) 2023 Dec 27;12(1):62. [doi: 10.3390/healthcare12010062] [Medline: 38200967]
8. Dosovitsky G, Bunge EL. Bonding with bot: user feedback on a chatbot for social isolation. Front Digit Health 2021;3:735053. [doi: 10.3389/fdgth.2021.735053] [Medline: 34713203]
9. Meng J, Dai Y(. Emotional support from AI chatbots: should a supportive partner self-disclose or not? J Comput-Mediat Commun 2021 Sep 9;26(4):207-222. [doi: 10.1093/jcmc/zmab005]
10. Shan Y, Ji M, Xie W, Lam KY, Chow CY. Public trust in artificial intelligence applications in mental health care: topic modeling analysis. JMIR Hum Factors 2022 Dec 2;9(4):e38799. [doi: 10.2196/38799] [Medline: 36459412]
11. Luxton DD. Ethical implications of conversational agents in global public health. Bull World Health Organ 2020 Apr 1;98(4):285-287. [doi: 10.2471/BLT.19.237636]
12. Moylan K, Doherty K. Expert and interdisciplinary analysis of AI-driven chatbots for mental health support: mixed methods study. J Med Internet Res 2025 Apr 25;27:e67114. [doi: 10.2196/67114] [Medline: 40279575]
13. Hoff KA, Bashir M. Trust in automation: integrating empirical evidence on factors that influence trust. Hum Factors 2015 May;57(3):407-434. [doi: 10.1177/0018720814547570] [Medline: 25875432]
14. Riedl R. Is trust in artificial intelligence systems related to user personality? Review of empirical evidence and future research directions. Electron Markets 2022 Dec;32(4):2021-2051. [doi: 10.1007/s12525-022-00594-4]
15. Siau K, Wang W. Building trust in artificial intelligence, machine learning, and robotics. Cut Bus Technol J 2018 Mar 26;31(2):47-53 [FREE Full text]
16. Thiebes S, Lins S, Sunyaev A. Trustworthy artificial intelligence. Electron Markets 2021 Jun;31(2):447-464. [doi: 10.1007/s12525-020-00441-4]
17. Hancock PA, Billings DR, Schaefer KE, Chen JYC, de Visser EJ, Parasuraman R. A meta-analysis of factors affecting trust in human-robot interaction. Hum Factors 2011 Oct;53(5):517-527. [doi: 10.1177/0018720811417254] [Medline: 22046724]
18. Schaefer KE, Chen JYC, Szalma JL, Hancock PA. A meta-analysis of factors influencing the development of trust in automation: implications for understanding autonomy in future systems. Hum Factors 2016 May;58(3):377-400. [doi: 10.1177/0018720816634228] [Medline: 27005902]
19. Sullivan E, Khalifa K. Idealizations and understanding: much ado about nothing? Australas J Philos 2019 Oct 2;97(4):673-689. [doi: 10.1080/00048402.2018.1564337]
20. Alanezi F. Assessing the effectiveness of ChatGPT in delivering mental health support: a qualitative study. J Multidiscip Healthc 2024;17:461-471. [doi: 10.2147/JMDH.S447368] [Medline: 38314011]
21. Dergaa I, Fekih-Romdhane F, Hallit S, et al. ChatGPT is not ready yet for use in providing mental health assessment and interventions. Front Psychiatry 2024;14. [doi: 10.3389/fpsyt.2023.1277756]
22. Alvarado R. AI as an epistemic technology. Sci Eng Ethics 2023 Aug 21;29(5):32. [doi: 10.1007/s11948-023-00451-3] [Medline: 37603120]
23. Alvarado R. What kind of trust does AI deserve, if any? AI Ethics 2023 Nov;3(4):1169-1183. [doi: 10.1007/s43681-022-00224-x]
24. Fonagy P, Luyten P, Allison E. Epistemic petrification and the restoration of epistemic trust: a new conceptualization of borderline personality disorder and its psychosocial treatment. J Pers Disord 2015 Oct;29(5):575-609. [doi: 10.1521/pedi.2015.29.5.575] [Medline: 26393477]
25. Duschinsky R, Foster S. Mentalizing and Epistemic Trust: The Work of Peter Fonagy and Colleagues at the Anna Freud Centre: Oxford University Press; 2021.

XSL•FO
RenderX

26. Fonagy P, Allison E. The role of mentalizing and epistemic trust in the therapeutic relationship. Psychotherapy (Chic) 2014 Sep;51(3):372-380. [doi: 10.1037/a0036505] [Medline: 24773092]
27. Luyten P, Campbell C, Moser M, Fonagy P. The role of mentalizing in psychological interventions in adults: systematic review and recommendations for future research. Clin Psychol Rev 2024 Mar;108:102380. [doi: 10.1016/j.cpr.2024.102380] [Medline: 38262188]
28. Gillath O, Ai T, Branicky MS, Keshmiri S, Davison RB, Spaulding R. Attachment and trust in artificial intelligence. Comput Human Behav 2021 Feb;115:106607. [doi: 10.1016/j.chb.2020.106607]
29. Jacovi A, Marasović A, Miller T, Goldberg Y. Formalizing trust in artificial intelligence. Presented at: FAccT '21: 2021 ACM Conference on Fairness, Accountability, and Transparency; Mar 3-10, 2021; Virtual Event Canada. [doi: 10.1145/3442188.3445923]
30. Matthews G, Hancock PA, Lin J, et al. Evolution and revolution: personality research for the coming world of robots, artificial intelligence, and autonomous systems. Pers Individ Dif 2021 Feb;169:109969. [doi: 10.1016/j.paid.2020.109969]
31. Campbell C, Tanzer M, Saunders R, et al. Development and validation of a self-report measure of epistemic trust. PLOS ONE 2021;16(4):e0250264. [doi: 10.1371/journal.pone.0250264] [Medline: 33861805]
32. Sindermann C, Yang H, Elhai JD, et al. Acceptance and fear of artificial intelligence: associations with personality in a German and a Chinese sample. Discov Psychol 2022 Dec;2(1). [doi: 10.1007/s44202-022-00020-y]
33. Stein JP, Messingschlager T, Gnambs T, Hutmacher F, Appel M. Attitudes towards AI: measurement and associations with personality. Sci Rep 2024 Feb 5;14(1):2909. [doi: 10.1038/s41598-024-53335-2] [Medline: 38316898]
34. Marengo D, Davis KL, Gradwohl G, Montag C. A meta-analysis on individual differences in primary emotional systems and Big Five personality traits. Sci Rep 2021;11(1). [doi: 10.1038/s41598-021-84366-8]
35. Fonagy P, Luyten P, Allison E, Campbell C. Taking stock to move forward: where the field of developmental psychopathology might be heading. Dev Psychopathol 2024 Dec;36(5):2056-2065. [doi: 10.1017/S0954579424000312] [Medline: 38389294]
36. Lorenzini N, Fonagy P. Attachment and personality disorders: a short review. FOC 2013 Apr;11(2):155-166. [doi: 10.1176/appi.focus.11.2.155]
37. Perry JB, Erikson EH. Childhood and society. J Marriage Fam 1965 Feb;27(1):115. [doi: 10.2307/349827]
38. Ainsworth MDS, Blehar MC, Waters E, Wall S. Patterns of Attachment: A Psychological Study of the Strange Situation: Psychology Press; 1979:391-xviii. [doi: 10.4324/9781315802428]
39. Bowlby J. The Bowlby-Ainsworth attachment theory. Behav Brain Sci 1979 Dec;2(4):637-638. [doi: 10.1017/S0140525X00064955]
40. Paulus FW, Ohmann S, Möhler E, Plener P, Popow C. Emotional dysregulation in children and adolescents with psychiatric disorders. a narrative review. Front Psychiatry 2021;12:628252. [doi: 10.3389/fpsyt.2021.628252] [Medline: 34759846]
41. Obeid S, Sacre H, Haddad C, et al. Factors associated with fear of intimacy among a representative sample of the Lebanese population: the role of depression, social phobia, self‐esteem, intimate partner violence, attachment, and maladaptive schemas. Perspect Psychiatric Care 2020 Jul;56(3):486-494. [doi: 10.1111/ppc.12438]
42. Toh G, Pearce E, Vines J, et al. Digital interventions for subjective and objective social isolation among individuals with mental health conditions: a scoping review. BMC Psychiatry 2022 May 12;22(1):331. [doi: 10.1186/s12888-022-03889-0] [Medline: 35549899]
43. Tateno M, Teo AR, Ukai W, et al. Internet addiction, smartphone addiction, and hikikomori trait in Japanese young adult: social isolation and social network. Front Psychiatry 2019;10:455. [doi: 10.3389/fpsyt.2019.00455] [Medline: 31354537]
44. Shabahang R, Aruguete MS, Shim H, Mokhtari Chirani B. Preference of online friends over face-to-face friends: effect of interpersonal fears on online relationship building preference. Interpers Int J Pers Relatsh 2022;16(1):120-135. [doi: 10.5964/ijpr.6711]
45. Casale S, Akbari M, Bocci Benucci S, Seydavi M, Fioravanti G. Interpersonally-based fears and problematic social networking site use: the moderating role of online social support. Int J Ment Health Addiction 2024 Jun;22(3):995-1007. [doi: 10.1007/s11469-022-00908-9]
46. Ingersoll TS, Poulin J, Deng R, Shan X, Witt H, Swain M. Fear of intimacy with helping professionals scale: reliability and validity of English and Mandarin versions. J Evid Based Soc Work 2012;9(4):317-332. [doi: 10.1080/15433714.2010.516689] [Medline: 22830935]
47. Bartholomew K, Horowitz LM. Attachment styles among young adults: a test of a four-category model. J Pers Soc Psychol 1991;61(2):226-244. [doi: 10.1037/0022-3514.61.2.226]
48. Conradi HJ, Noordhof A, Kamphuis JH. Satisfying and stable couple relationships: attachment similarity across partners can partially buffer the negative effects of attachment insecurity. J Marital Fam Ther 2021 Jul;47(3):682-697. [doi: 10.1111/jmft.12477] [Medline: 33493361]
49. Park Y, Debrot A, Spielmann SS, Joel S, Impett E, MacDonald G. Distinguishing dismissing from fearful attachment in the association between closeness and commitment. Soc Psychol Personal Sci 2019 May;10(4):563-572. [doi: 10.1177/1948550618768823]
50. Deng S, Lin Z, Wu L. Service staff makes me nervous: the effect of insecure attachment on AI service preference. SSRN . [doi: 10.2139/ssrn.4356946]

XSL•FO
RenderX

51. Ali MR, Razavi SZ, Langevin R, et al. A virtual conversational agent for teens with autism spectrum disorder. Presented at: IVA '20: ACM International Conference on Intelligent Virtual Agents; Oct 20-22, 2020; Virtual Event Scotland UK. [doi: 10.1145/3383652.3423900]
52. Huang Y, Huang H. Exploring the effect of attachment on technology addiction to generative AI chatbots: a structural equation modeling analysis. Int J Hum Comput Interact 2024:1-10. [doi: 10.1080/10447318.2024.2426029]
53. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34. [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
54. Liotti M, Milesi A, Spitoni GF, et al. Unpacking trust: the Italian validation of the Epistemic Trust, Mistrust, and Credulity Questionnaire (ETMCQ). PLOS ONE 2023;18(1):e0280328. [doi: 10.1371/journal.pone.0280328] [Medline: 36701301]
55. Carli L. Attaccamento e Rapporto Di Coppia [Attachment and Couple Relationship] [Book in Italian]: Raffaello Cortina; 1995.
56. Fossati A, Krueger RF, Markon KE, Borroni S, Maffei C. Reliability and validity of the Personality Inventory for DSM-5 (PID-5). Assessment 2013 Dec;20(6):689-708. [doi: 10.1177/1073191113504984]
57. Krueger RF, Derringer J, Markon KE, Watson D, Skodol AE. Initial construction of a maladaptive personality trait model and inventory for DSM-5. Psychol Med 2012 Sep;42(9):1879-1890. [doi: 10.1017/S0033291711002674] [Medline: 22153017]
58. Diagnostic and Statistical Manual of Mental Disorders: DSM-5TM, 5th edition: American Psychiatric Publishing; 2013:947-xliv. [doi: 10.1176/appi.books.9780890425596]
59. Descutner CJ, Thelen MH. Development and validation of a Fear-of-Intimacy Scale. Psychol Assess 1991;3(2):218-225. [doi: 10.1037//1040-3590.3.2.218]
60. Senese VP, Miranda MC, Lansford JE, Bacchini D, Nasti C, Rohner RP. Psychological maladjustment mediates the relation between recollections of parental rejection in childhood and adults' fear of intimacy in Italy. J Soc Pers Relat 2020 Jun;37(6):1968-1990. [doi: 10.1177/0265407520912339]
61. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [doi: 10.3758/bf03193146] [Medline: 17695343]
62. Aysel Y, Yenipinar A, Koç Ş, Canga D, Kaya F. Determining sample size in logistic regression with G-Power. BSJ Eng Sci 2019 Jan 1;2(1):16-22 [FREE Full text]
63. Liang J, Bi G, Zhan C. Multinomial and ordinal logistic regression analyses with multi-categorical variables using R. Ann Transl Med 2020 Aug;8(16):982. [doi: 10.21037/atm-2020-57] [Medline: 32953782]
64. Agresti A. Categorical Data Analysis, 3rd edition: John Wiley & Sons; 2013.
65. McCullagh P. Regression models for ordinal data. J R Stat Soc Ser B 1980 Jan 1;42(2):109-127. [doi: 10.1111/j.2517-6161.1980.tb01109.x]
66. Ananth C, Kleinbaum DG. Regression models for ordinal responses: a review of methods and applications. Int J Epidemiol 1997 Dec 1;26(6):1323-1333. [doi: 10.1093/ije/26.6.1323]
67. Shatz I. Assumption-checking rather than (just) testing: the importance of visualization and effect size in statistical diagnostics. Behav Res Methods 2024 Feb;56(2):826-845. [doi: 10.3758/s13428-023-02072-x] [Medline: 36869217]
68. Ernst AF, Albers CJ. Regression assumptions in clinical psychology research practice-a systematic review of common misconceptions. PeerJ 2017;5:e3323. [doi: 10.7717/peerj.3323] [Medline: 28533971]
69. Schlegel B, Steenbergen M. Brant: test for parallel regression assumption. R package version 03-0 [Web page in German]. Benjamin E Schlegel. 2020. URL: https://benjaminschlegel.ch/r/brant/ [accessed 2024-07-12]
70. Peterson B, Harrell FE. Partial proportional odds models for ordinal response variables. Appl Stat 1990;39(2):205. [doi: 10.2307/2347760]
71. McCullagh P, Nelder JA. Generalized Linear Models: Chapman and Hall; 1983. URL: https://www.utstat.toronto.edu/brunner/oldclass/2201s11/readings/glmbook.pdf [accessed 2024-12-08]
72. The VLMC package. The Comprehensive R Archive Network. 2024. URL: https://cran.r-project.org/web/packages/VLMC/index.html [accessed 2025-01-02]
73. Lee J, Lee D, Lee JG. Influence of rapport and social presence with an AI psychotherapy chatbot on users' self-disclosure. Int J Human–Comput Interact 2024 Apr 2;40(7):1620-1631. [doi: 10.1080/10447318.2022.2146227]
74. Rahsepar Meadi M, Sillekens T, Metselaar S, van Balkom A, Bernstein J, Batelaan N. Exploring the ethical challenges of conversational AI in mental health care: scoping review. JMIR Ment Health 2025 Feb 21;12:e60432. [doi: 10.2196/60432] [Medline: 39983102]
75. Haque MDR, Rubya S. An overview of chatbot-based mobile mental health apps: insights from app description and user reviews. JMIR Mhealth Uhealth 2023 May 22;11:e44838. [doi: 10.2196/44838] [Medline: 37213181]
76. Santona A, De Cesare P, Tognasso G, De Franceschi M, Sciandra A. The mediating role of romantic attachment in the relationship between attachment to parents and aggression. Front Psychol 2019;10:1824. [doi: 10.3389/fpsyg.2019.01824] [Medline: 31447749]
77. Messina I, Calvo V, Grecucci A. Attachment orientations and emotion regulation: new insights from the study of interpersonal emotion regulation strategies. Res Psychother 2023 Dec 31;26(3):703. [doi: 10.4081/ripppo.2023.703] [Medline: 38224213]

XSL•FO
RenderX

78. Borycki E, Farghali A, Joseph A, et al. A digital health evaluation framework. Appl Med Inform 2024 Nov 21;46:37-40. [doi: 10.1787/ae3016b9-en]
79. Beg MJ. Responsible AI integration in mental health research: issues, guidelines, and best practices. Indian J Psychol Med 2025 Jan;47(1):5-8. [doi: 10.1177/02537176241302898] [Medline: 39650770]
80. Andrade C. The limitations of online surveys. Indian J Psychol Med 2020 Nov;42(6):575-576. [doi: 10.1177/0253717620957496] [Medline: 33354086]
81. Beg MJ, Verma M, Vishvak Chanthar KMM, Verma MK. Artificial intelligence for psychotherapy: a review of the current state and future directions. Indian J Psychol Med 2024 Jul 28. [doi: 10.1177/02537176241260819]

Abbreviations

AI: artificial intelligence
CA: conversational agent
CHERRIES: Checklist for Reporting Results of Internet E-Surveys
ETCMQ: Epistemic Trust, Credulity and Mistrust Questionnaire
FIS: Fear of Intimacy Scale
FIS-HP: Fear of Intimacy with Helping Professionals Scale
IV: independent variable
NPOM: nonproportional odds model
OR: odds ratio
PID-5-BF: Personality Inventory for DSM-5—Brief Form—Adult
PIS: Participation Information Sheet
POM: proportional odds model
PPOM: partial proportional odds model
RQ: Relationship Questionnaire

Edited by A Kushniruk; submitted 30.12.24; peer-reviewed by A Jamal, M Shardlow, M Beg, S Mashuk; revised version received 27.05.25; accepted 28.05.25; published 30.07.25.

Please cite as:
Guglielmucci F, Di Basilio D
Predicting Engagement With Conversational Agents in Mental Health Therapy by Examining the Role of Epistemic Trust, Personality, and Fear of Intimacy: Cross-Sectional Web-Based Survey Study
JMIR Hum Factors 2025;12:e70698
URL: https://humanfactors.jmir.org/2025/1/e70698
doi:10.2196/70698

© Fanny Guglielmucci, Daniela Di Basilio. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 30.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Weighing Costs and Benefits of Delay and the Acceptance of Two Decision Support Tools in Mental Health Care: Scoping Study Using Quantitative and Qualitative Data

Sarah McKenna, PhD; Min K Chong, PhD; Adam Poulsen, PhD; Ashlee Turner, PhD; Carla Gorban; Jacob J Crouse, PhD; William Capon, MBMSc; Mathew Varidel, PhD; Melissa Aji, PhD; Elizabeth M Scott, PhD; Ian B Hickie, PhD; Frank Iorfino, PhD

Brain and Mind Centre, Faculty of Medicine and Health, University of Sydney, 94 Mallett St, Camperdown, NSW, Australia

Corresponding Author:
Sarah McKenna, PhD
Brain and Mind Centre, Faculty of Medicine and Health, University of Sydney, 94 Mallett St, Camperdown, NSW, Australia

Abstract

Background: Mental disorders are the leading cause of disability in young people (aged 12‐30 years), and their incidence constitutes a major health crisis. Primary youth mental health services are struggling to keep up due to overwhelming demand, the complexity and severity of young people presenting for care, and a shortage of qualified mental health professionals (MHPs). Artificial intelligence (AI) tools have the potential to facilitate necessary improvements to diagnosis, triage, and care planning for young people with emerging mental disorders.

Objective: The objective of the present scoping research was to examine beliefs and attitudes underlying MHP acceptance of AI tools in youth mental health services.

Methods: In total, 57 MHPs (mean age 35.35, SD = 9.50 years, 72% female (n = 39)) with experience working with youth populations (age 12-30) took part in study 1 that involved completing a web-based survey about the acceptability of using AI in early intervention services. During study 2, 15 MHPs also participated in 1-hour semistructured Zoom interviews. Attitudes toward the use of 2 novel AI prototypes (both of which provide recommendations for care coordination based on previously published data analyses) in youth mental health were explored. Quantitative data were interpreted using descriptive statistics, and qualitative analysis followed the thematic analysis approach.

Results: MHPs were more likely to agree than disagree that AI will improve youth mental health care overall (eg, n=37, 64% participants somewhat or strongly agree that the field of mental health will improve with AI). Despite voicing concerns regarding data security and privacy, MHPs also acknowledged a need for AI to improve the “signal-to-noise ratio” in services and address delays to care for those with severe and complex problems. Such problems were seen as pervasive across the youth mental health system and emphasize the serious costs of delaying the development and implementation of novel tools. All participating MHPs discussed the potential negative impacts of not adopting novel tools.

Conclusions: MHP acceptance and uptake of novel AI tools in youth mental health services will be driven by a more complex cost-benefit analysis of both adopting and not adopting, rather than solely on their design. The costs of delay are clear, and so researchers and MHPs have a shared imperative to develop useful and meaningful clinical tools and to work jointly on integrating them into practice. Limitations of our sample (including low sample size limiting generalizability) notwithstanding, these findings should inform the future design and implementation of such tools.

(JMIR Hum Factors 2025;12:e71678) doi:10.2196/71678

KEYWORDS

youth mental health; digital mental health; youth; adolescence; technologies

Introduction

Mental disorders have profound negative effects on young people’s health, well-being, and socio-occupational functioning, and the impacts typically continue throughout their life [1]. Mental disorders are the leading cause of disability among 10- to 24-year-olds and are the leading cause of death among people aged 15-44 in Australia [2-4]. In response to the extreme scale and complexity of these problems, urgent calls have been made to radically transform youth mental health services so that they are able to provide more efficient and personalized treatments [5,6]. For example, in Australia, this has led to the development of *headspace*, an early intervention youth mental health service that can provide multidisciplinary care to young people with emerging mental disorders [7,8]. Similar early intervention

XSL•FO
RenderX

services have also emerged internationally, for example, in Canada, the UK, and Ireland, reflecting that youth mental health is currently a global health crisis [9-12]. Yet, these services are frequently overburdened and under-resourced, creating an urgent need to improve assessment and care coordination, particularly for those with emerging severe and complex problems [5,13,14].

To address this need, there is a rapidly growing interest in the development of artificial intelligence (AI) tools that can provide data-driven solutions to more efficiently and accurately plan appropriate treatments, though such tools are yet to be successfully implemented at scale in youth mental health services [15-21]. Research has shown that there are unique challenges to implementing AI tools in health care settings, including regulatory restrictions and clinical decision systems, along with a lack of trust as compared to other industries, likely because making a mistake in health care settings could have serious consequences, even death [22,23]. Even so, these perspectives fail to consider the potential benefits of improving decision-making systems, particularly in settings that are under-resourced and lack qualified staff. This points to a need for more targeted research focusing on specific challenges and opportunities for implementing unique tools based on the usefulness of the individual tool, characteristics of the setting, and broader contextual factors.

Various frameworks have been proposed to identify factors that may underpin successful implementation of novel technologies. The technology acceptance model (TAM) and the unified theory of acceptance and use of technology (UTAUT) are both widely accepted theoretical frameworks [24-27]. TAM proposes that technology acceptance is linked to perceived usefulness and perceived ease of use and is one of the most widely used and earliest theories in health care industries. UTAT is a broader framework that includes performance expectancy (how much a person believes that using the system will improve job performance) and effort expectancy (the perceived simplicity of using the system) along with cultural and organizational factors such as social expectancy (the perception that others believe the technology should be used) and facilitating conditions (technical or organizational support). Both frameworks are likely to have benefits for understanding this complex problem. While it is important to understand organizational- and system-level barriers, given that this is a unique context facing significant challenges, mental health professionals (MHPs) may also consider factors independent of organizational or social setting. For example, young people often present to mental health services with high levels of risk and low functioning, which may be linked to less trust in novel tools to assist with identifying severe clients, given the potential consequences of mistakes.

So far, there is limited research on the use of tools to assist with multidisciplinary assessment and care coordination in mental health care settings, particularly early intervention services, as current research tends to focus on the use of tools within therapy or as self-help tools [5,6,28]. Broadly, barriers to implementation among MHPs include low perceived usefulness of the tools, poor access to training and supervision, and poor perceived competence and knowledge in using clinical prediction models [29,30]. Additionally, therapist characteristics such as external feedback propensity (the degree to which someone values or is open to feedback) and feedback utilization (the extent to which someone implements feedback) significantly mediate the effectiveness of digital tools on client outcomes [31]. Given these barriers, it is important for design and implementation research to engage with youth MHPs to understand how their acceptance and uptake can be improved.

The present study focused on 2 tools that can be used by MHPs that are intended to improve triage and care coordination of young people with emerging mental disorders. Clients and MHPs can input information regarding clients' multidimensional needs, which is then processed by algorithms developed from data on youth populations presenting to early intervention services. Tool 1 is focused on assessing the level of complexity so that clients can be triaged into short-term therapy or more high-intensity and multidisciplinary treatments [32]. Tool 2 is focused on predicting clients' functional trajectory.[33] Those who have a high likelihood of deteriorating likely require more focused functional supports compared to care as usual. Currently, rough prototypes of these tools have been constructed as proof of concept (and so are not publicly available), and we sought MHP feedback to inform their ongoing design and improvement.

Accordingly, this research explores MHPs' acceptance of two decision-making support tools within a unique context, youth mental health services. The present paper describes 2 studies. Study 1 aimed to assess MHP attitudes toward the use of AI tools within youth mental health services more broadly, and the aim of Study 2 was to gain feedback on MHP acceptance of 2 novel AI tools that have been designed to improve assessment and treatment planning in youth mental health services.

Methods

Two scoping studies were conducted separately, each with unique aims; quantitative and qualitative results were not integrated. Study 1 involved descriptive analyses of general attitudes toward digital tools amongst youth MHPs. Study 2 involved in-depth qualitative analysis of MHPs' attitudes and beliefs. These research studies were not preregistered. Data from the study are available on request.

Recruitment and Informed Consent

We used snowball sampling via social media advertisements (using LinkedIn and Facebook) and referrals from participants or other MHPs that had existing relationships with our team to recruit MHPs who had experience working with young people. MHPs completed a digital consent form and indicated whether they were willing to participate in a one-on-one interview, a survey, or both. Participants willing to participate in a Zoom interview were contacted via e-mail and were reimbursed for participating in both the interview and the web-based survey. Recruitment for surveys was time-limited as this was a scoping study that was intended to inform further development and testing of the digital tools; it was carried out between August 2023 and October 2024. Recruitment for qualitative interviews continued until saturation had been reached in April 2024. Participants were included in the study if they were health

XSL•FO
RenderX

professionals who had provided mental health care to adolescents and young adults (12-30).

Ethical Considerations

This study was approved by the University of Sydney's Human Research Ethics Committee (2021/HE000680). Informed consent was obtained from all participants via web-based forms before participating. All data was deidentified before analyses and no identifying information was included in the current paper. Participants were compensated for completing virtual Zoom interviews.

Study 1: Quantitative Data Collection

We first collected web-based survey data about clinicians' acceptance of decision-support tools in youth mental health care to provide descriptive information about (1) the extent to which MHPs were likely to see digital tools as acceptable and useful, and (2) the potential barriers or facilitators of engagement over and above acceptance of the tools.

Measures

Acceptance of AI-Powered Care Pathways

We adapted 8 items from an existing survey by Cornelissen et al [24] based on the UTAUT, a previous study demonstrated internal consistency and inter-rater reliability of the items, to assess what factors were likely to impact MHPs' acceptance and uptake of AI tools in youth mental health services. Participants were asked to rate statements such as "I believe AI would be useful in my overall performance" on a 5-point scale from "strongly disagree" to "strongly agree."

Potential Impact of Innovations in Artificial Intelligence

We used 10 items to assess MHP's opinions about the potential of future AI to replace key tasks carried out in youth mental health services. These items were adapted from a survey by Doraiswamy et al [34] that assessed psychiatrists' opinions on AI to be more appropriate for youth settings and a range of MHPs, including service managers, psychiatrists, psychologists, etc. The utility of these items has been established in previous studies [35,36]. Participants rated the potential importance of using AI to improve various tasks such as "treatment planning" and "identifying referrals to other services" on a 5-point scale from "not at all important" to "very important."

Facilitators and Barriers to Engagement

We used 12 items to assess what attributes of novel tools and implementation strategies would most likely impact uptake of novel tools in youth mental health services. Items were adapted from a ranking exercise designed by Leigh et al [37] and were based on a literature search along with expert consensus. Example items included usability, an established research evidence base, and endorsement by colleagues. Participants were asked to rate attributes such as "usability" using a 5-point rating scale from "not important" to "very important."

Procedure

We used web-based surveys to collect quantitative data on the acceptability of AI to health professionals in mental health settings and on the potential facilitators or barriers of implementation and engagement. We used descriptive statistics to compare responses to these items. Given the available sample size, it was not possible to analyze any differences in attitudes between various items. Before analysis, data were screened for valid responses (ie, reported job role fit inclusion criteria for study and location of workplace was in Australia). Given that we used descriptive statistics, participants with missing data were excluded from analyses.

Study 2: Qualitative Data Collection

Participants were initially shown the digital tools and asked to provide initial impressions of the tool before commencing the interview. Semistructured interviews (see Multimedia Appendix 1) allowed us to gain more in-depth information about health professionals' attitudes toward decision support tools within youth mental health settings. Qualitative data collection and analysis followed Standards for Reporting Qualitative Research [38].

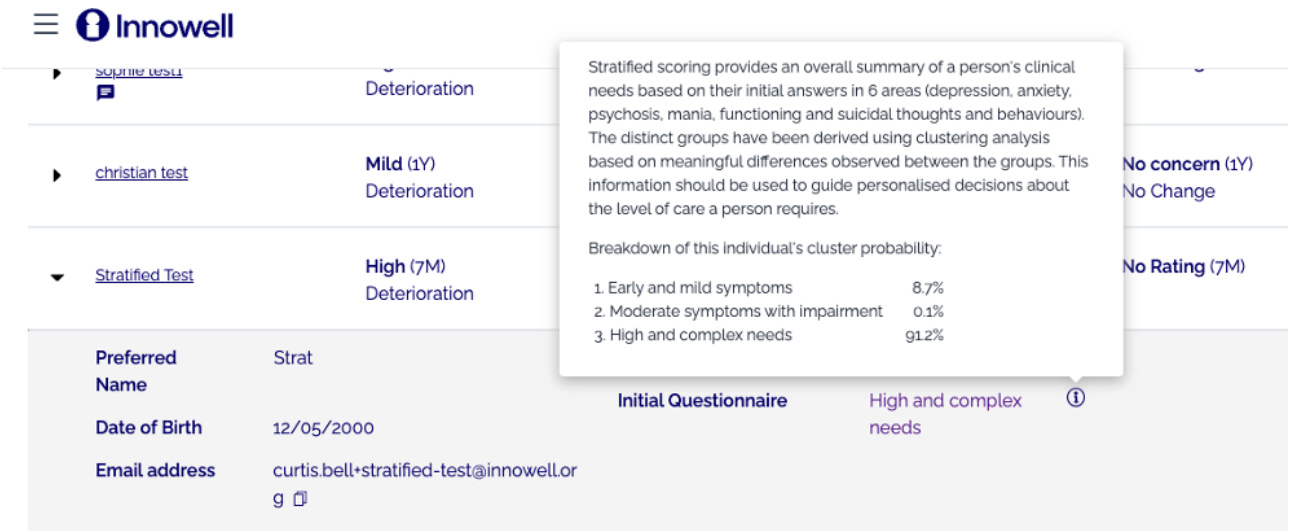
Tool 1: Stratified Scoring of Young People's Clinical Needs

Participants were shown a stratified scoring tool that had been embedded within the Innowell platform. The Innowell platform facilitates multidimensional assessments by allowing clients and their MHPs to complete in-depth initial assessments and then track progress during treatment in 5 key domains: (1) overall health, (2) mental health, (3) suicidal thoughts and behaviors, (4) everyday functioning, (5) social connectedness, and (6) drug and alcohol use. This platform has been described in detail in previous publications [39,40].

As shown in Figure 1, the stratified scoring tool is designed to improve identification of subgroups who have complex multidimensional needs to determine appropriate service pathways and care options. MHPs are given an estimate of the probability that a client has (1) early and mild symptoms (low and mixed symptomatology with limited functional impairment), (2) moderate symptoms with impairment (established depressive symptoms and functional impairment), or (3) high and complex needs (very high and complex needs with functional impairment, suicidality, and at-risk mental states (psychosis or mania). The probabilities presented to MHPs are based on a latent class analysis of a cohort of 1284 people aged 12 - 25 years presenting to youth mental health services, that has been described in more depth in a previous publication [32].

XSL•FO
RenderX

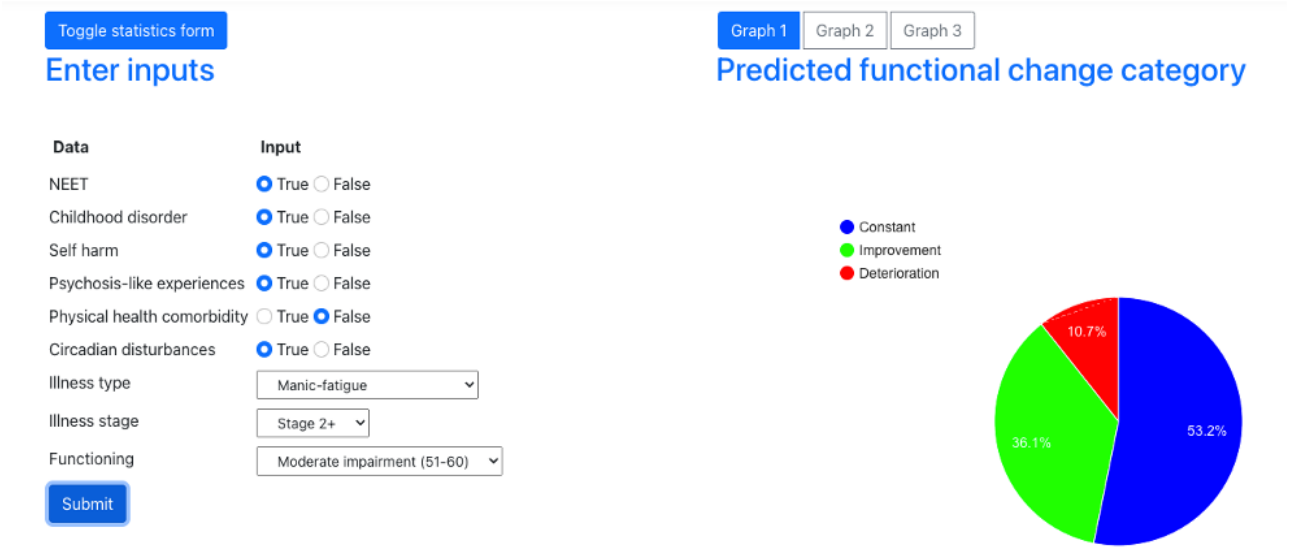
Figure 1. Depiction of tool 1 designed to provide stratified scoring of young people's clinical needs.



Tool 2: Predicted Functional Change

Participants were also shown a tool that predicts a client's functional impairment trajectory over a 3-month period. As shown in Figure 2, the tool shows MHPs the likelihood that a client's functional impairment will (1) remain constant, (2) improve, or (3) deteriorate. The prognostic model was developed from a sample of 718 young people (12 - 25 y) who were engaged in youth mental health care, and it has been described in more detail in a previous publication [8]. In total, 8 factors were used for prediction, including employment, education or training status, self-harm, psychotic-like experiences, physical health comorbidity, childhood-onset syndrome, illness type, clinical stage, and circadian disturbance.

Figure 2. Depiction of tool 2 showing predicted functional change.



Measures

We designed a semistructured interview based on components of TAM that aim to understand why users accept a given technology and how user acceptance can be improved through technology design. According to this framework, acceptance can be measured by behavioral intention to use a novel tool. Interview questions were focused on better understanding 4 key themes, including (1) first impressions of usability, (2) performance expectancy of the tool in decision-making, (3) predicted use within workflow, and (4) professional or ethical concerns.

Procedure

Semistructured interviews were conducted via Zoom, video-recorded, and later transcribed. Interviews lasted approximately 60 minutes. Transcriptions of the interviews were analyzed using thematic analysis techniques, with the aim of establishing themes regarding the acceptability of our tools within youth mental health services [41]. Data analysis was both inductive (what was in the data) and deductive (informed by TAM) [24]. Two academic researchers (SM and MC) independently established a list of codes based on transcripts of the interviews. Subsequently, these codes were shared and discussed with a master coder (FI) to establish key themes. A

XSL•FO
RenderX

second round of coding was conducted using a similar process to establish broader patterns of meaning within each theme and refined during a face-to-face meeting with the master coder. Trustworthiness of data coding was established by using two coders and a master coder for triangulation and through an audit trail that can be made available on request.

Results

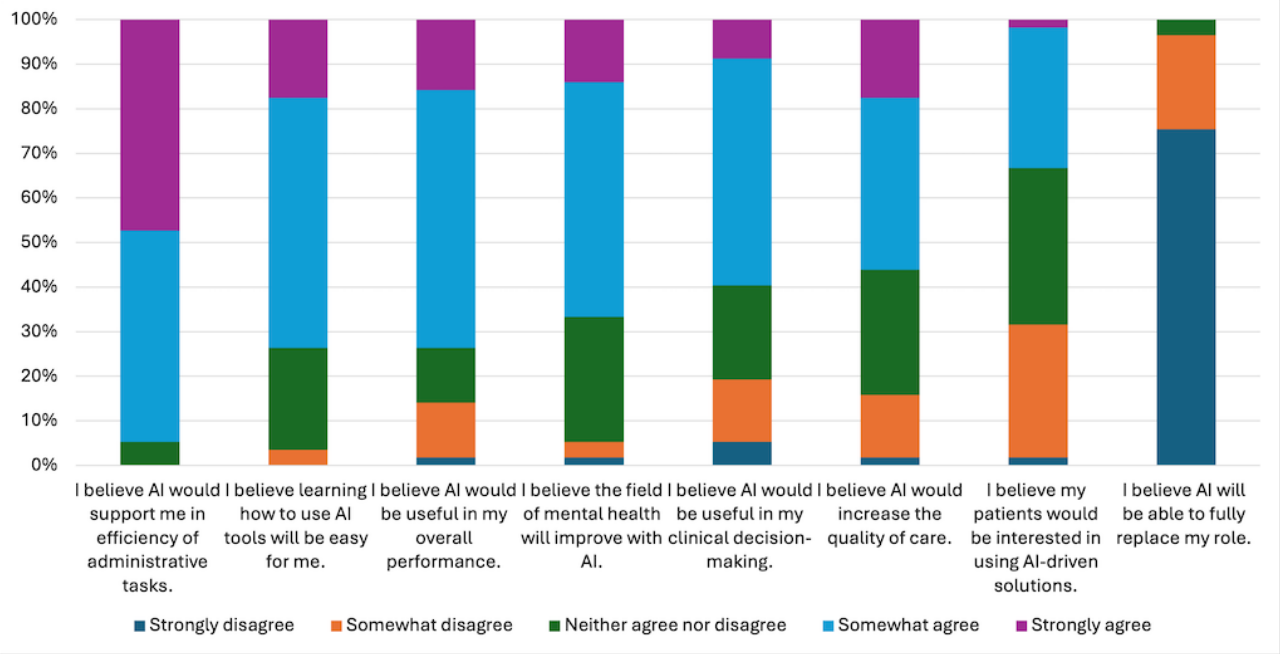
Participants

The sample consisted of 57 MHPs (mean age 35.35, SD 9.50 years, 72% female (n = 39)) with experience working with youth populations. The majority of participants (45, 81%) came from metropolitan areas compared to rural or regional areas, with a median of 6 - 10 years of clinical experience. In total, 21 participants (37%) had 10 years of experience or less, and 4 participants (7%) had 20+ years of experience. In total, 16 (28%) participants worked only in private settings, 17 participants (30%) worked only in community or hospital settings, while the remainder worked across a variety of settings. All participants completed the web-based survey, while 15 also agreed to participate in one-on-one semistructured interviews.

Quantitative Data Analysis of MHP Acceptance of Novel AI Tools in Youth Mental Health Services

As shown in Figure 3, we first asked participants to rate the perceived usefulness of AI to improve clinical care. In general, participants were more likely to agree than disagree that AI (1) would benefit their overall performance, (n=40, 71%, strongly or somewhat agreed); (2) would improve the field of mental health, (n=37, 64% strongly or somewhat agreed); (3) would be useful in clinical decision-making, (n=33, 58%) strongly or somewhat agreed); and (4) would increase the quality of care (n=31, 54%) strongly or somewhat agreed. Almost all (n=52, 92%) participants strongly or somewhat agreed that AI would improve the efficiency of administrative tasks. Most participants (n=40, 71%) also agreed that AI tools would be easy for them to learn. Even so, participants were unsure whether clients were interested in AI-driven solutions, with (n=20; 34%) reporting they neither agree nor disagree that clients would be interested. Most participants (n=53, 93%) strongly or somewhat disagreed that AI would fully replace their roles.

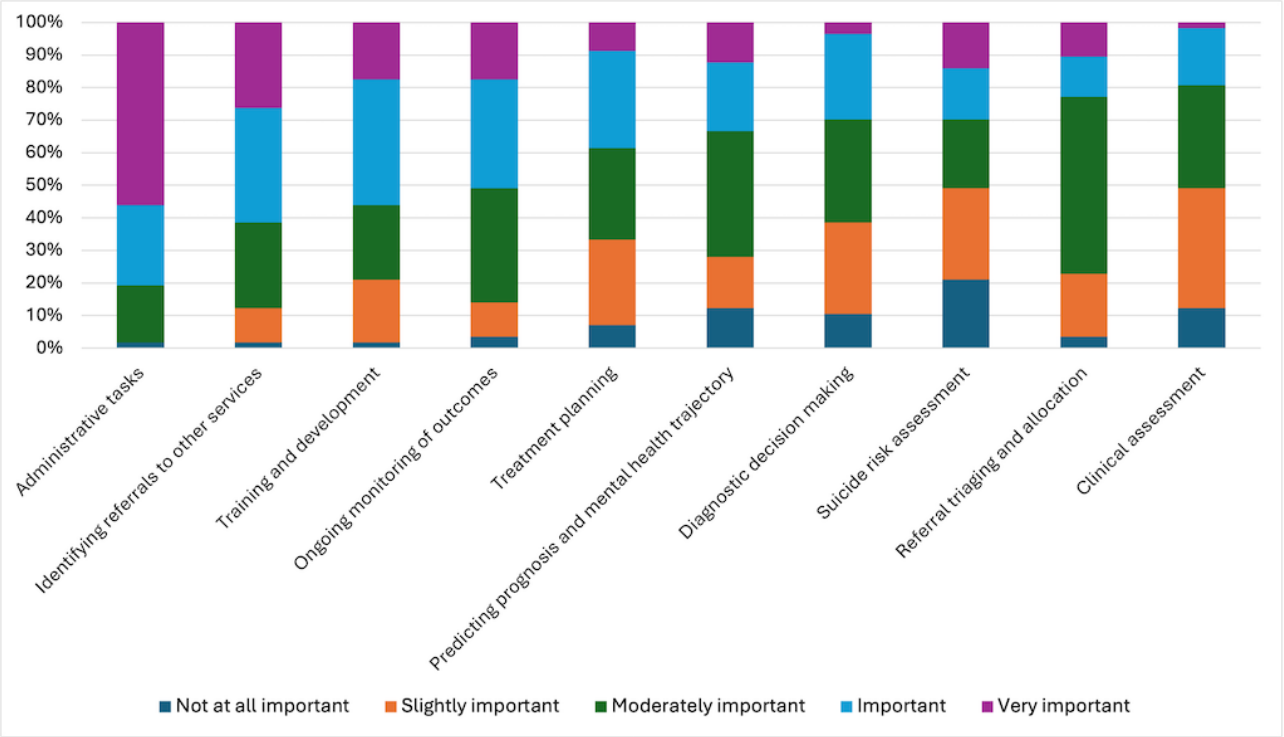
Figure 3. Levels of agreement among mental health professionals regarding the potential of AI to improve aspects of mental health care.



We also asked participants to rate the importance of using AI to improve various clinical tasks (see Figure 4). Participants were least likely to believe that AI was important for suicide risk assessment (42% (n = 24) rating it as not at all or slightly important) and clinical assessment (50% (n=29) rating it as not at all or slightly important) and were most likely to believe that it was important for improving administrative tasks (82% (n=47) rating this as very important or important). Over half (50% (n=29)) of participants believed it was important or very important for AI to improve ongoing monitoring of outcomes, identification of referrals to other services, and training and development. For other items, including referral triaging and allocation, clinical assessment, diagnostic decision-making, and predicting prognosis and mental health trajectory, participant ratings appeared to be more mixed.

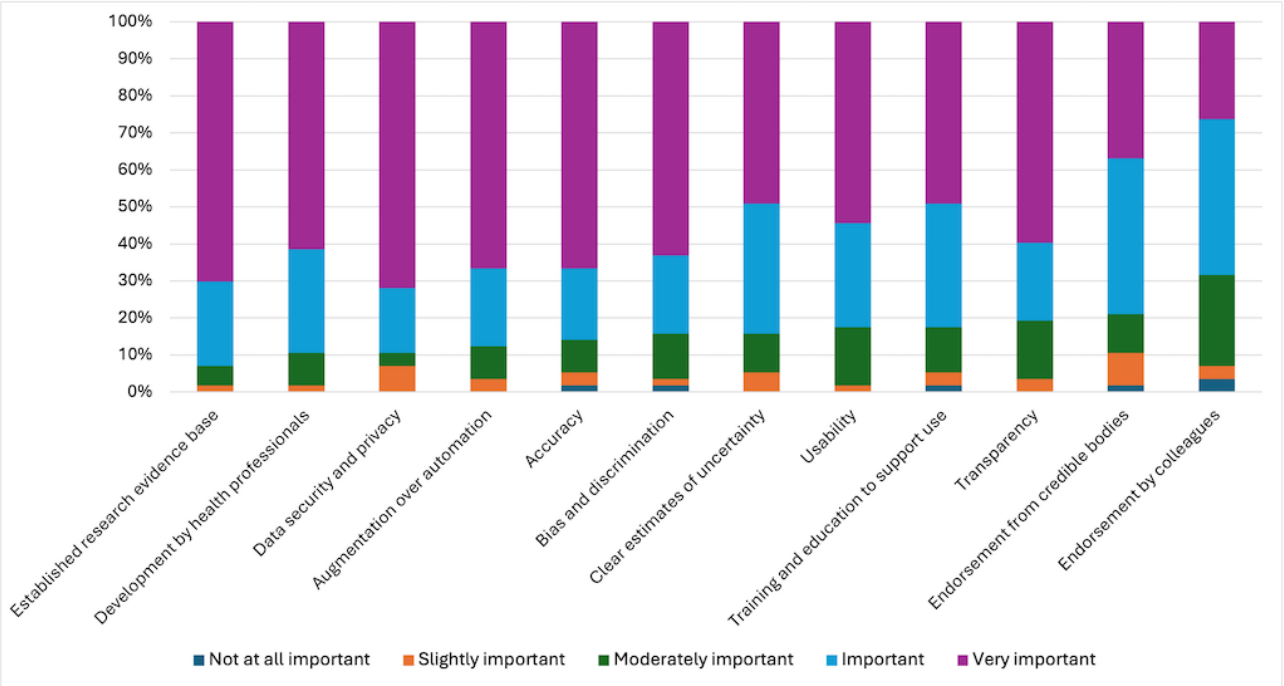
XSL•FO
RenderX

Figure 4. Perceived importance of using AI to improve various clinical tasks according to mental health professionals.



Finally, as shown in Figure 5, participants were asked to rate the level of importance of various attributes when considering implementing AI solutions into their practice. All items had high levels of agreement (ie, over 60% (n=34) agree or strongly agree). Over 90% (n=51) of participants saw an established research evidence base, development by health professionals, and data security and privacy as important or very important. Meanwhile, over 80% (n=46) of participants reported that accuracy, augmentation over automation, bias and discrimination, usability, endorsement from credible bodies, clear estimates of uncertainty, transparency, and training and education support were important or very important.

Figure 5. Perceived importance of various aspects of implementing and designing novel digital tools according to mental health professionals.



XSL•FO
RenderX

Qualitative Feedback on AI Prototypes as Proof of Concept

Overall, we found that participating MHPs were highly interested in the development of AI tools. Thus, rather than simply focusing on the prototypes they were shown, discussions focused on the costs and benefits of adopting and not adopting diagnostic and prognostic AI tools in youth mental health services more generally. In line with survey results, participating MHPs were optimistic about the potential improvement to youth mental health care that could be achieved through AI. However, they also appeared to be highly cautious about potential risks to clients and about financial, legal, or training burdens on themselves.

Costs and Benefits of Adopting Novel Tools in Youth Mental Health Services

New Tools Need to Overcome the Non-Negotiable Barriers to Adoption (Costs)

Data fluency was a key concern, as participants often struggled to interpret the results of Tool 1 (see Table 1, Theme 1.1). For example, participants were often surprised by the case study used to demonstrate Tool 1, as the individual was shown to have the greatest probability of being in the low- or high-risk group but no probability of being in the moderate group, which several participants interpreted as an error. One participant noted that for such tools to be implemented, MHPs would need to be well trained in more complex statistical models. By comparison, Tool 2, which used a pie chart to demonstrate predicted trajectories, was immediately seen as more "user friendly," visually appealing, and easy to understand.

XSL•FO
RenderX

Table . Qualitative feedback on artificial intelligence (AI) prototypes from youth mental health professionals.

Themes	Sample excerpts from qualitative data
Costs and benefits of adopting novel AI tools	
Financial costs, time burdens, and safety concerns (costs)	
1.1 Training and other resources needed to understand complex algorithms and establish trustworthiness, reliability, and validity.	"… it's based on probability, right, which is not deterministic, and it's not likely chance. And so it's something we have to get used to, in terms of things not being black and white." [Participant #3, Male] "I don't know if I'm smart enough to understand the algorithms. But I just, I'd be very interested in reading the, like the administration manual, or very dumbed down research paper that I could go, oh, this is why they gave that particular weighting… this was the clinical team and the evidence base that they drew upon." [Participant #17, Male] "I want to know that it has been tested in clinical and community samples, you know, I want to know about its reliability and validity… all that sort of stuff." [Participant #6, Female] "What one person thinks is complex at the beginning of their career, that language is imprecise between people. And so having something standardized and saying, well, 91.2% chance of this… I mean, that seems pretty helpful to me." [Participant #19, Male]
1.2 Data security and privacy	"I'd want to know like how secure it is, like who can access this information?" [Participant #24, Male] "Oh, just feels really intrusive in the person's life and in a way that I wouldn't be comfortable with. So I think that's one thing that's put me off." [Participant #15, Female]
1.3 Professional liability for decision-making	"If the clinician doesn't look at it, kind of like reading a doctor's referral letter, and they miss something, will you be liable?... you should probably be liable, because you're not looking at important information… like that is a professional issue… And on the flip side, if you make a clinical decision or a service decision that is maybe counter to the indications… say there's a coroner's inquest and they say, well, the tool said actually you should have referred and why didn't you?" [Participant #18, Male]
1.4 Ethical concerns about sharing predictive data with clients that may lead to helplessness or reduced self-efficacy.	"I think for people that maybe don't have that sort of level of insight… about their own functioning or their own prognosis… this would be really shocking for them… with someone that was like, oh, yeah, you know… there's nothing wrong with me, or this isn't my fault, or any of those sort of, like core beliefs about what's going on for them. I can imagine them seeing this would be really sort of shocking, or really disarming." [Participant #17, Male] "People feel quite heard and seen by completing these things. And I think that's something if, if that can be made front and centre of these measures… clinicians have their own biases about X, Y, Z, but actually the client experiences, it's helpful… even clients being told that I think that, yeah, helpful message." [Participant #9, Female]

XSL•FO
RenderX

Themes	Sample excerpts from qualitative data
1.5 Financial cost and time burden of implementing novel tool	"Some of the clinicians would be like, I don't have time to… look at all that data before I see someone... And so I can see that that'll be a struggle, like how do you make this really quickly accessible as well." [Participant #5, Female] "Depends about the cost… if it was free, and I knew the data wasn't going to go anywhere, and it was just for me and my practice, I'd use it every day. Okay. If it costs me money, depending on how much… I'd probably say a two out of ten [likelihood of using]. Okay. And if I'm desperate with a complex case… I would, you know, I'd give it like a seven out of ten... If it was a really small fee, I would give it to all my clients prior to them starting." [Participant #18, Male]
Expected use for tools within existing workflows (benefits)	
2.1 Faster identification and referrals of those with severe and complex needs—particularly in large multidisciplinary clinics	"I think it would be very helpful… there have been instances where we have had a few clients come in, and we just weren't the right service for them… emerging psychosis or things like that, which were just not the place for. So I think this would just be so helpful in being able to identify that early on." [Participant #24, Male] "If I ran the headspace center still… I'd be like an eight out of ten [chance of adopting], like working really hard to see how we could bring this into the center." [Participant #19, Male]
2.2 Can supplement rather than replace clinical judgment, which is naturally biased, and dependent on people's training and experience	"Clinical decision making and clinical judgment, has a pride of place… there's kind of entrenched deep problems with that, your biases… I think the clinician has to be pretty open… the statistical tool is probably more accurate than my feeling that someone is suicidal or not… If the test results are coming back and saying 86% chance of being cancer, you should probably agree with that rather than like, your various biases based on your clinical experiences that have selective memories and things like that." [Participant #5, Female]
2.3 Can reduce burden of identifying evidence-based treatments particularly for early MHPs assessing severe and complex clients	"It would be great to have something that could generate a type of like rudimentary treatment plan… your client is presenting with all their history, all this information, and it could like, all the latest research and intervention… that would save so much time, because as a clinician, I think that's where so much of my time is taken just like treatment planning and stuff." [Participant #16, Female] "Seeing what's the best intervention out there, perhaps as a space where AI can help. Okay, ADHD, the first line of treatment is CBT… gathering all the research that's effective there…" [Participant #18, Male]
2.4 "What gets measured gets done"—better information for funders and governments regarding population needs	"Whatever gets measured, gets done…. you get funded by literally the occasions of service you do… if you had this broader sense of, you know, the social occupational functioning and depression, anxiety, and maybe some other things to consider, like if that was all data that you could report back on, to your funders to say this is the targets that we're about… I really like that idea too." [Participant #4, Male] "I think it's really helpful from a service perspective… to look at how they allocate their resources, but it would need some change… you know, are they seeing people [with] early symptoms, are they mostly moderate, you know, what are the what are the breakdown, and people are presenting?" [Participant #9, Female]

XSL•FO
RenderX

Themes	Sample excerpts from qualitative data
2.5 “Clinicians will do what they always do”	“They just still continue to do what they do, they don’t necessarily take all that information into account as we know… actually feeds back into people’s skills and education.”
	[Participant #5, Female]
	“I don’t know that individual clinician workers actually, necessarily then take that into account in terms of treatment planning. They’ve ticked off the assessment that’s done and then the person is allocated according to what the service can actually provide, not necessarily what the young person needs.”
	[Participant #2, Female]
	“I think it’s up to how much the clinician uses the information again, if a clinician is sensitive to what this data is, is telling them, then I think it can have a significant impact on client outcome. But yeah, I guess, you know, therapy needs to change as a result of this to impact the outcome.”
	[Participant #19, Male]
Costs and benefits of not adopting AI tools	
Irreplaceability of the “human touch” (benefits)	
3.1 Loss of empathy, understanding, and connection	“This is a real person answering the question. There must be some other issues here… physical, family, and help seeking history to complement what’s been responded to - the why.”
	[Participant #15, Female]
	“The AI is not a human being with a culture with understanding and maybe empathy. Of course, we could build it in, we could build it into the AI, but again, you know, at what cost does that have? Like, what risk is that? You know, would I be comfortable speaking to an AI? I don’t know…”
	[Participant #16, Female]
3.2 Skeptical that treatment progress can be predicted by algorithms.	“So I guess I will check you know, if I if I agree with that statistic. And if I, if I did think that too, then I would look at you know, maybe you know, getting them more supports or connected to different resources…”
	[Participant #9, Female]
	“I think because improvement in mental health is so variable… client can have really different goals to begin with, goals can change throughout with treatment… it’s never linear improvement either… like you might not know at the start of treatment that they are on the spectrum and so you could be trying to treat them as such, but it’s like, then they’re not going to demonstrate improvement because there’s that crucial detail that you missed.”
	[Participant #10, Female]
3.3 MHPs pref their own clinical judgment	“This person’s needs haven’t been understood through this data. It’s just their symptoms have been understood… So there’s a need to be really cautious about… how we’re presenting the data, I guess, and what we’re saying the data can and can’t do.”
	[Participant #6, Female]
	“I could see its potential benefits, but… there’s a part of me that’s like… it can never override human judgment… it’s helpful, but I guess it’s not the be all and end all. That’s just my initial sort of reaction… I guess it would sound like a tool that I have to help support my decision making… it’s still up to me to decipher and make sense of the information and to come to my own conclusion about it.”
	[Participant #10, Female]

XSL•FO
RenderX

Themes	Sample excerpts from qualitative data
3.4 "Clinicians will do what they always do"	"They just still continue to do what they do, they don't necessarily take all that information into account as we know… actually feeds back into people's skills and education." [Participant #5, Female] "I don't know that individual clinician workers actually, necessarily then take that into account in terms of treatment planning. They've ticked off the assessment that's done and then the person is allocated according to what the service can actually provide, not necessarily what the young person needs." [Participant #3, Male] "I think it's up to how much the clinician uses the information again, if a clinician is sensitive to what this data is, is telling them, then I think it can have a significant impact on client outcome. But yeah, I guess, you know, therapy needs to change as a result of this to impact the outcome." [Participant #15, Female]
The hidden costs of delay (costs)	
4.1 High demand and high complexity in primary youth mental health services along with "skills shortage" leading to delayed care and referrals	"We have a huge demand for services, we have people presenting with very mixed symptoms, we have them coming into systems where there is very low levels of expertise… I see AI as being a tool which will help to help to improve the signal to noise ratio, will help to be able to discriminate quickly, those people who need to be stepped up into a kind of higher level of care, to more intensive care or monitoring, and those people who can be treated with a kind of briefer, less intense kind of treatment, and they can be triaged accordingly, and then kind of monitored and tracked accordingly." [Participant #5, Female "I think it would be helpful when I'm a bit uncertain about a client in terms of how complex I feel like it is going to be… and whether I'm feeling like, they might need a different service, potentially. Kind of having a bit of a clear line of why we might be referring them on." [Participant #4, Male]
4.2 Service managers and supervisors have poor oversight of less experienced clinicians and not involved in intake assessments	"If you have people with different levels of training and different theoretical models and so on, I think it'd be helpful… that shared data… I could see it being helpful in a supervision setting, too, because it's sort of more objective data that you can bring into the supervisee supervisory relationship, and discuss and reflect on…" [Participant #8, Female] "A really, really clear example I can think of is if there's someone who is reporting psychotic, like experiences and it's a less experienced clinician, that person is going to be more inclined to go or shit they've got psychosis, I need to send them to hospital as opposed to… what is it actually based on? … as that support tool, as an adjunct for that clinical experience." [Participant #4, Male]
4.3 Moral obligation to clients to adapt and improve.	"I think the input we can provide is making sure it's reliable, valid, and so on. But when it comes to preferences, it's not really my own preference… it's really what the folks in front of us want to utilize. And so we need to be open minded to that and not biased." [Participant #17, Male] "If these tools are improving, like health care outcomes, and are more accessible to people… it's about trying to kind of reduce human suffering, like why would you really want to get in the way of that…. if you kind of imagine it in medicine, and they came up with some revolution, yeah, to kind of detect cancer or something, and then doctors, cancer specialists were like, 'Oh, we don't want to do that' … we'd be up in arms about that sort of thing." [Participant #24, Male]

Additionally, almost all users wanted detailed information about how complexity and functioning had been assessed, including understanding psychometric properties of assessment tools, seeing individual items that had led to a particular rating, and having a good understanding of the populations from which the ratings had been derived. Notably, however, improved training

XSL•FO
RenderX

for assessing and identifying severe and complex clients was seen as an important gap, with or without novel tools. One MHP commented that having tools such as this may lead to a more uniform definition of complexity across health services and professionals, which could improve ease of referrals and care coordination.

Participating MHPs also raised concerns about security and privacy (Theme 1.2) and wanted clear information about who would have access to client data before using these tools. In addition, participants had concerns about professional liability if clients received inappropriate care or if risk was not responded to in a timely fashion because of either ignoring or following the advice of tools (Theme 1.3). Several participants reflected on the need for clear guidelines from professional bodies, as well as changes to legislation and insurance policies, before the implementation of these tools could become widespread.

Moreover, when discussing Tool 2, several participants discussed ethical concerns about sharing information about a client's predicted trajectories with clients themselves (Theme 1.4). Participants were particularly concerned about the potentially damaging impacts on a client's hope for recovery, their self-efficacy, and their motivation to continue with treatment. On the other hand, more experienced MHPs were more likely to see this tool as empowering, with one participant noting that they had used a similar tool in their practice and found that "people feel quite heard and seen by completing these things."

Finally, the financial and time burdens of adopting novel tools within health services were frequently cited as a major barrier to implementation (Theme 1.5). Participants suggested that MHPs would only pay a fee for novel technology if they could quickly and easily understand the value of the tool.

Novel Tools Can Address Knowledge Gaps and Will Have Particular Benefits for Severe and Complex Presentations (Benefits)

Participating MHPs generally expected both tools to perform well as a "red flag" system, particularly in large multidisciplinary clinics with high demand for service, to improve efficiency of care coordination, and to help identify those with complex needs more quickly and easily (see Table 1, Theme 2.1). For this reason, participants who worked in private practice were less likely to see the value in the tools, given they did not generally triage clients before seeing them for initial assessments. However, several participants reflected on the difficulty of referring clients within the Australian "Medicare system" and the value of these tools for improving communication with general practitioners and other health professionals.

Despite concerns about the tools' accuracy and reliability, several participants noted the value of having another source of information. They thought the tools could provide valuable insight about complex clients who were not responding to current treatment by assessing whether their current trajectory aligned with expectations (Theme 2.2). Participating MHPs reflected that no tools should be used in isolation; instead, they recommended that such tools should be "triangulated" with other assessment methods and their clinical expertise.

Several participating MHPs suggested that both tools could be improved by providing treatment recommendations, particularly to reduce the burden of care coordination for complex or unique presentations that may require a higher level of training and expertise (Theme 2.3). In a similar vein, participants also suggested that they would prefer these tools to help them identify a client's needs from treatment (ie, need for functional supports), rather than diagnostic information, as this was seen as more relevant for care planning.

As well as helping individual clients to get more efficient care, one participant noted that "what gets measured gets done," suggesting that if tools such as this are implemented in services, it will help funders to better understand client needs and support more diverse support options in youth mental health services other than individual therapy (Theme 2.4).

An important caveat is that participants predicted these benefits would not be realized without adequate training and support for participating MHPs. One participant commented that a lack of experience and expertise in treating severe or complex problems would limit the tools' usefulness (Theme 2.5). They also predicted that the use of the tools would be limited by participants' training in multidisciplinary care more broadly. Indeed, several participating MHPs working in private psychology practices commented that they did not directly address client functioning and did not see Tool 2 as relevant to them.

Costs and Benefits of "Business as Usual" in Youth Mental Health Services

Essential to Preserve Benefits of the "Human Touch" (Benefits)

Even among participants who were highly open to these tools and believed they could be beneficial, there were significant concerns about replacing in-person assessments with digital tools (see Table 1, Theme 3.1). For the most part, participants were worried that digital tools could not display empathy and understanding or build connections with clients. They were also concerned that tools did not consider important contextual information such as family history, social support, or interpersonal communication styles. Additionally, participants were highly skeptical about tools predicting the likelihood of future trajectories (Theme 3.2), as they viewed progress in therapy as highly variable and influenced by a complex range of factors. Most reported that they would mainly use prediction tools retrospectively to assess whether treatments were achieving the intended effects or needed adjustments.

A related problem was that participants ultimately trusted their clinical judgment over and above algorithms to make final decisions about treatment planning (Theme 3.3). When asked, most participants stated they would trust their clinical judgment if they disagreed with the tools. Having said this, 3 clinical supervisor participants wanted to use these tools with less experienced MHPs to help them recognize biases in their thinking. These participants were also more likely to see

XSL•FO
RenderX

"disagreements" with the tools as an important source of information that would enable more reflection and assessment, rather than seeing tool ratings as "right" or "wrong."

Urgent Need to Improve the Signal-to-Noise Ratio in Youth Mental Health Services, and MHPs Cannot Get in the Way of Progress (Costs)

Importantly, most participants also wanted to discuss broader issues in youth mental health services that required urgent improvement through tools such as those used in the present study, indicating that whilst they had concerns about adopting novel tools, they also had concerns about business as usual. Participants frequently commented on issues such as "skills shortage" and the current high demand for services in Australian youth mental health services (see Table 1, Theme 4.1). Participating service managers were concerned that young people were presenting to care earlier in the course of illness (when symptoms are more mixed rather than meeting criteria for specific disorders) and in higher numbers and reported a shortage of skilled health professionals in Australian early intervention services. As such, the tools were seen to improve the "signal-to-noise ratio" within health services by identifying those with more severe disorders earlier and providing more efficient referrals and care coordination to appropriate services.

Additionally, 2 participating service managers cited occasions when clients had been turned away, despite presenting with concerning symptoms, due to lack of knowledge of intake staff (Theme 4.2). For this reason, they frequently discussed how our tools could improve supervision and oversight of less experienced MHPs in their services, particularly those making decisions about referrals and allocation of care.

There were also ethical and moral concerns about "gate-keeping" a client's access to more accurate and comprehensive assessments (Theme 4.3). Two participants made comparisons to medical fields, such as cancer treatment, suggesting that if professionals in these fields were reluctant to adopt evidence-based tools, it would be seen as highly unethical. Participants reflected that the decision to use or not use novel tools should involve clients, particularly if they have been shown to improve outcomes.

Discussion

In sum, MHPs that participated in our study were optimistic that AI would lead to improvements in youth mental health care overall. Despite non-negotiable barriers to adoption, including the financial and cost burdens of learning novel tools, participants found the tools to be useful "red flag" systems that could drastically improve the efficiency of referral and treatment planning processes. These findings highlight both the expected costs and potential benefits of adopting novel tools in clinical settings. Equally, all participating MHPs discussed the impacts of not adopting novel tools (despite not being asked directly), demonstrating the complexity of this decision. Interpersonal relationships with MHPs have important benefits for clients that should be protected. Even so, there are significant risks to clients if youth mental health services do not adopt more intelligent systems that can address the "signal-to-noise" ratio and ensure complex clients receive high-quality care faster.

This work provides a better understanding of how MHPs will make decisions about accepting and adopting novel AI into practice. Digital mental health research has reported consistent problems with widespread implementation of novel tools in youth mental health, partly due to MHP resistance [15-21]. Even so, MHPs in our study clearly saw the utility and value of AI (quantitatively and qualitatively), and this value ranges from reducing administrative burden to helping with referrals, prediction, and treatment planning. These benefits notwithstanding, MHPs identified serious challenges to overcome, including training requirements, infrastructure limitations, and professional liability concerns. Meanwhile, many of the barriers identified are highly solvable—not just in the future, but they must be addressed before widespread adoption can occur. If these barriers are addressed and benefits realized, MHPs clearly acknowledge the value of improved workflows, particularly in high-volume clinics, and predict important benefits for broader mental health system demand. This has the potential to raise the quality of decisions about evidence-based treatments by supporting them with quality information and predictions. Therefore, the costs of delay are clear. Researchers and MHPs have a shared imperative to develop useful and meaningful clinical tools and to work jointly on integrating them into practice.

These findings also demonstrate feasible avenues for improving MHP acceptance and uptake of novel tools. This study identified that usability, development by health professionals, an evidence base, and augmentation over automation are crucial. Again, many of these steps are highly achievable and have been successfully trialed. For example, previous research has found that health professionals are more likely to adopt AI tools when they have seen experienced colleagues integrate them into care and demonstrate potential use [24,31]. Comprehensive literature reviews have also established key steps that may be needed to reduce algorithm avoidance in health care, such as decision autonomy and incentivization [42]. Our results suggest the design of future AI tools should ensure that prognostic information is clearly linked to treatment recommendations and is framed as a means of supporting care planning (eg, by allowing MHPs to assess whether clients are progressing as expected), rather than being seen as deterministic. Given these steps are highly achievable, delaying rapid development and implementation of such tools would be a missed opportunity. There is an urgent need for researchers and MHPs to work together on redesigning current models of care, moving beyond 'business as usual' and toward innovative approaches that better address the complexities of mental health service demand.

Despite these contributions, our research also has limitations that should be addressed. Only 49 MHPs completed our web-based survey. This may reflect that participants who were willing to complete web-based surveys were more interested in AI and therefore more likely to be optimistic about potential future contributions. We also used snowball sampling that may have introduced bias by capturing participants with similar backgrounds and attitudes. Taken together, this limits the generalizability of our findings, and significant further research

XSL•FO
RenderX

is needed amongst youth MHPs. We plan to address this in the future through more comprehensive research in real-world settings on the acceptability, feasibility, and usefulness of our tools. Moreover, we were unable to conduct subgroup analyses with our quantitative data that may help to explain whether AI acceptance is likely to vary based on professional background, experience level, or setting, or whether certain implementation strategies were seen as more valuable. Another limitation is that participants were shown prototypes and were not able to use the tools independently. Survey data suggests that MHPs are overly optimistic about the ease of understanding AI algorithms and so may have overestimated the usability and interpretability of the tools without support from the interviewer. For example, in survey items, MHPs were likely to agree that it would be easy for them to learn novel AI; however, we found that our prototypes were complicated for MHPs to understand, partly because they provided estimates of uncertainty rather than "black and white" information about client status. Addressing both limitations in future research should involve implementing prototypes in youth mental health services as early as possible, to gain more nuanced and diverse opinions from real-world settings.

In conclusion, MHPs recognize the potential of novel AI tools to solve urgent problems in youth mental health. While there are clear and non-negotiable barriers to adoption—including financial costs, potential liability, and training needs—MHPs consider these against the costs of delaying innovation. Problems accessing quality care, particularly for those with severe and complex problems, are pervasive across youth mental health services, creating a need to better handle "signal-to-noise" and to improve efficiency of treatment planning and referral processes. Future research can support MHPs by rapidly developing and implementing novel tools and helping integrate such tools into practice.

Acknowledgments

Thanks to all the mental health professionals who participated in this study and all the staff in the Youth Mental Health and Technology Team, past and present, who contributed to this work. JJC is supported by an NHMRC Emerging Leadership Fellowship (2008196); IBH is supported by an NHMRC Leadership L3 Fellowship (2016346); FI is funded by an NHMRC Emerging Leadership Fellowship (2018157). SM is supported by the Cottle Family Fellowship in Youth Mental Health. WC and MKC were supported by the Australian Government Research Training Program (RTP) Scholarship.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

SM, MC, AP, AT, CG, JJC, WC, MA, MV, and FI have nothing to declare.

IBH is a Professor of Psychiatry and the Co-Director of Health and Policy, Brain and Mind Centre, University of Sydney. He has led major public health and health service development in Australia, particularly focusing on early intervention for young people with depression, suicidal thoughts and behaviors and complex mood disorders. He is active in the development through codesign, implementation and continuous evaluation of new health information and personal monitoring technologies to drive highly personalized and measurement-based care. He holds a 3.2% equity share in Innowell Pty Ltd that is focused on digital transformation of mental health services.

EMS is a Principal Research Fellow at the Brain and Mind Centre, University of Sydney, a Consultant Psychiatrist and Adjunct Clinical Professor at the School of Medicine, University of Notre Dame. She previously served as the Discipline Leader for Adult Mental Health at Notre Dame until January 2025. In addition, she is a member of Medibank's Medical and Mental Health Reference Groups. EMS has also delivered educational seminars on the clinical management of depressive disorders, receiving honoraria from pharmaceutical companies, including Servier, Janssen, and Eli Lilly. Moreover, she has contributed to a national advisory board for Pfizer's antidepressant Pristiq and served as the National Coordinator for an antidepressant trial sponsored by Servier.

Multimedia Appendix 1
Semistructured interview guide for mental health professionals.
[DOCX File, 50 KB - humanfactors_v12i1e71678_app1.docx]

References

1. McGorry P, Mei C. Youth mental health: a rising public health challenge. Australas Psychiatry 2023 Jun;31(3):245-246. [doi: 10.1177/10398562231177350] [Medline: 37204413]
2. Plana-Ripoll O, Pedersen CB, Agerbo E, et al. A comprehensive analysis of mortality-related health metrics associated with mental disorders: a nationwide, register-based cohort study. Lancet 2019 Nov 16;394(10211):1827-1835. [doi: 10.1016/S0140-6736(19)32316-5] [Medline: 31668728]

XSL•FO
RenderX

3. Erskine HE, Moffitt TE, Copeland WE, et al. A heavy burden on young minds: the global burden of mental and substance use disorders in children and youth. Psychol Med 2015 May;45(7):1551-1563. [doi: 10.1017/S0033291714002888] [Medline: 25534496]
4. Deaths in Australia. : Australian Government - Australian Institute of Health and Welfare; 2024 URL: https://www.aihw.gov.au/getmedia/d227ae99-0fdc-4b2b-a071-283dd25c5394/deaths-in-australia.pdf?v=20250424155259&inline=true [accessed 2025-09-23]
5. McGorry PD, Mei C, Dalal N, et al. The Lancet Psychiatry Commission on youth mental health. Lancet Psychiatry 2024 Sep;11(9):731-774. [doi: 10.1016/S2215-0366(24)00163-9] [Medline: 39147461]
6. Kieling C, Salum GA, Pan PM, Bressan RA. Youth mental health services: the right time for a global reach. World Psychiatry 2022 Feb;21(1):86-87. [doi: 10.1002/wps.20933] [Medline: 35015362]
7. McGorry PD, Tanti C, Stokes R, et al. headspace: Australia's National Youth Mental Health Foundation--where young minds come first. Med J Aust 2007 Oct 1;187(S7):S68-S70. [doi: 10.5694/j.1326-5377.2007.tb01342.x] [Medline: 17908032]
8. Iorfino F, Carpenter JS, Cross SP, et al. Social and occupational outcomes for young people who attend early intervention mental health services: a longitudinal study. Med J Aust 2022 Feb 7;216(2):87-93. [doi: 10.5694/mja2.51308] [Medline: 34664282]
9. O'Reilly A, O'Brien G, Moore J, et al. Evolution of Jigsaw - a National Youth Mental Health Service. Early Interv Psychiatry 2022 May;16(5):561-567. [doi: 10.1111/eip.13218] [Medline: 34464507]
10. Wiens K, Bhattarai A, Pedram P, et al. A growing need for youth mental health services in Canada: examining trends in youth mental health from 2011 to 2018. Epidemiol Psychiatr Sci 2020 Apr 17;29:e115. [doi: 10.1017/S2045796020000281] [Medline: 32299531]
11. McGorry P, Bates T, Birchwood M. Designing youth mental health services for the 21st century: examples from Australia, Ireland and the UK. Br J Psychiatry 2013 Jan;202(s54):s30-s35. [doi: 10.1192/bjp.bp.112.119214]
12. Wilson J, Clarke T, Lower R, et al. Creating an innovative youth mental health service in the United Kingdom: The Norfolk Youth Service. Early Interv Psychiatry 2018 Aug;12(4):740-746. [doi: 10.1111/eip.12452] [Medline: 28776920]
13. KPMG Department of Health. Evaluation of the national headspace program. : Australian Government–Department of Health, Disability and Ageing; 2022 URL: https://www.health.gov.au/resources/publications/evaluation-of-the-national-headspace-program?language=en [accessed 2026-09-23]
14. Hickie IB, Scott EM, Cross SP, et al. Right care, first time: a highly personalised and measurement-based care model to manage youth mental health. Med J Aust 2019 Nov;211 Suppl 9(S9):S3-S46. [doi: 10.5694/mja2.50383] [Medline: 31679171]
15. Sharma M, Savage C, Nair M, Larsson I, Svedberg P, Nygren JM. Artificial intelligence applications in health care practice: scoping review. J Med Internet Res 2022 Oct 5;24(10):e40238. [doi: 10.2196/40238] [Medline: 36197712]
16. Lim CT, Fuchs C, Torous J. Integrated digital mental health care: a vision for addressing population mental health needs. Int J Gen Med 2024;17:359-365. [doi: 10.2147/IJGM.S449474] [Medline: 38318335]
17. Ridout B, Forsyth R, Amon KL, Navarro P, Campbell AJ. The urgent need for an evidence-based digital mental health practice model of care for youth. JMIR Ment Health 2024 Mar 22;11:e48441. [doi: 10.2196/48441] [Medline: 38534006]
18. Fusar-Poli P, Manchia M, Koutsouleris N, et al. Ethical considerations for precision psychiatry: a roadmap for research and clinical practice. Eur Neuropsychopharmacol 2022 Oct;63:17-34. [doi: 10.1016/j.euroneuro.2022.08.001] [Medline: 36041245]
19. Beg MJ. Responsible AI integration in mental health research: issues, guidelines, and best practices. Indian J Psychol Med 2025 Jan;47(1):5-8. [doi: 10.1177/02537176241302898] [Medline: 39650770]
20. Beg MJ, Verma M, M V, Verma MK. Artificial intelligence for psychotherapy: a review of the current state and future directions. Indian J Psychol Med 2025 Jul;47(4):314-325. [doi: 10.1177/02537176241260819] [Medline: 40831472]
21. Beg MJ, Verma MK. Exploring the potential and challenges of digital and AI-driven psychotherapy for ADHD, OCD, schizophrenia, and substance use disorders: a comprehensive narrative review. Indian J Psychol Med 2024 Dec 14:02537176241300569. [doi: 10.1177/02537176241300569] [Medline: 39677516]
22. Lee AT, Ramasamy RK, Subbarao A. Understanding psychosocial barriers to healthcare technology adoption: a review of TAM technology acceptance model and unified theory of acceptance and use of technology and UTAUT frameworks. Healthcare (Basel) 2025 Jan 27;13(3):250. [doi: 10.3390/healthcare13030250] [Medline: 39942440]
23. Ammenwerth E. Technology Acceptance Models in Health Informatics: TAM and UTAUT: IOS Press; 2019. [doi: 10.3233/SHTI190111]
24. Cornelissen L, Egher C, van Beek V, Williamson L, Hommes D. The drivers of acceptance of artificial intelligence-powered care pathways among medical professionals: web-based survey study. JMIR Form Res 2022 Jun 21;6(6):e33368. [doi: 10.2196/33368] [Medline: 35727614]
25. Yadegari M, Mohammadi S, Masoumi AH. Technology adoption: an analysis of the major models and theories. Technology Analysis & Strategic Management 2024 Jun 2;36(6):1096-1110. [doi: 10.1080/09537325.2022.2071255]
26. Philippi P, Baumeister H, Apolinário-Hagen J, et al. Acceptance towards digital health interventions - model validation and further development of the unified theory of acceptance and use of technology. Internet Interv 2021 Dec;26:100459. [doi: 10.1016/j.invent.2021.100459] [Medline: 34603973]

XSL•FO
RenderX

27. Dingel J, Kleine AK, Cecil J, Sigl AL, Lermer E, Gaube S. Predictors of health care practitioners' intention to use AI-enabled clinical decision support systems: meta-analysis based on the unified theory of acceptance and use of technology. J Med Internet Res 2024 Aug 5;26:e57224. [doi: 10.2196/57224] [Medline: 39102675]
28. Kleine AK, Kokje E, Lermer E, Gaube S. Attitudes toward the adoption of 2 artificial intelligence-enabled mental health tools among prospective psychotherapists: cross-sectional study. JMIR Hum Factors 2023 Jul 12;10:e46859. [doi: 10.2196/46859] [Medline: 37436801]
29. Laukka E, Huhtakangas M, Heponiemi T, Kanste O. Identifying the roles of healthcare leaders in HIT implementation: a scoping review of the quantitative and qualitative evidence. Int J Environ Res Public Health 2020 Apr 21;17(8):2865. [doi: 10.3390/ijerph17082865] [Medline: 32326300]
30. Sligo J, Gauld R, Roberts V, Villa L. A literature review for large-scale health information system project planning, implementation and evaluation. Int J Med Inform 2017 Jan;97:86-97. [doi: 10.1016/j.ijmedinf.2016.09.007] [Medline: 27919399]
31. Rogan J, Bucci S, Firth J. Health care professionals' views on the use of passive sensing, AI, and machine learning in mental health care: systematic review with meta-synthesis. JMIR Ment Health 2024 Jan 23;11:e49577. [doi: 10.2196/49577] [Medline: 38261403]
32. Capon W, Hickie IB, Fetanat M, et al. A multidimensional approach for differentiating the clinical needs of young people presenting for primary mental health care. Compr Psychiatry 2023 Oct;126:152404. [doi: 10.1016/j.comppsych.2023.152404] [Medline: 37524044]
33. Iorfino F, Oliveira R, Cripps S, et al. A prognostic model for predicting functional impairment in youth mental health services. Eur Psychiatry 2024;67(1). [doi: 10.1192/j.eurpsy.2024.1787]
34. Doraiswamy PM, Blease C, Bodner K. Artificial intelligence and the future of psychiatry: Insights from a global physician survey. Artif Intell Med 2020 Jan;102:101753. [doi: 10.1016/j.artmed.2019.101753] [Medline: 31980092]
35. Blease C, Bernstein MH, Gaab J, et al. Computerization and the future of primary care: a survey of general practitioners in the UK. PLOS ONE 2018;13(12):e0207418. [doi: 10.1371/journal.pone.0207418] [Medline: 30540791]
36. Blease C, Kaptchuk TJ, Bernstein MH, Mandl KD, Halamka JD, DesRoches CM. Artificial intelligence and the future of primary care: exploratory qualitative study of UK general practitioners' views. J Med Internet Res 2019 Mar 20;21(3):e12802. [doi: 10.2196/12802] [Medline: 30892270]
37. Leigh S, Ashall-Payne L, Andrews T. Barriers and facilitators to the adoption of mobile health among health care professionals from the United Kingdom: discrete choice experiment. JMIR Mhealth Uhealth 2020 Jul 6;8(7):e17704. [doi: 10.2196/17704] [Medline: 32628118]
38. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251. [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
39. LaMonica HM, Iorfino F, Lee GY, et al. Informing the future of integrated digital and clinical mental health care: synthesis of the outcomes from project synergy. JMIR Ment Health 2022 Mar 9;9(3):e33060. [doi: 10.2196/33060] [Medline: 34974414]
40. McKenna S, Piper S, Capon W, et al. The polarization of clinician and service staff perspectives after the use of health information technology in youth mental health services: implementation and evaluation study (preprint). JMIR Human Factors. . [doi: 10.2196/preprints.42993]
41. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
42. Burton JW, Stein MK, Jensen TB. A systematic review of algorithm aversion in augmented decision making. Behavioral Decision Making 2020 Apr;33(2):220-239. [doi: 10.1002/bdm.2155]

Abbreviations

AI: artificial intelligence
MPH: mental health professional
TAM: technology acceptance model
UTAUT: unified theory of acceptance and use of technology

XSL•FO
RenderX

Edited by A Kushniruk, Y Gong; submitted 23.01.25; peer-reviewed by M Beg, S Gaube; revised version received 06.04.25; accepted 03.06.25; published 30.09.25.

Please cite as:
McKenna S, Chong MK, Poulsen A, Turner A, Gorban C, Crouse JJ, Capon W, Varidel M, Aji M, Scott EM, Hickie IB, Iorfino F Weighing Costs and Benefits of Delay and the Acceptance of Two Decision Support Tools in Mental Health Care: Scoping Study Using Quantitative and Qualitative Data
JMIR Hum Factors 2025;12:e71678
URL: https://humanfactors.jmir.org/2025/1/e71678
doi:10.2196/71678

© Sarah McKenna, Min K Chong, Adam Poulsen, Ashlee Turner, Carla Gorban, Jacob J Crouse, William Capon, Mathew Varidel, Melissa Aji, Elizabeth M Scott, Ian B Hickie, Frank Iorfino. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 30.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Understanding the Thoughts and Preferences for Technologies Designed to Detect Feelings of Loneliness: Interview Study Among Older Adults

Jessica Rees[1], PhD; John Ratcliffe[2], PhD; Wei Liu[3], PhD; Yi Zhou[3], PhD; Sebastien Ourselin[4], PhD; Michela Antonelli[4], PhD; Ashay Patel[4], PhD; Yu Shi[5], PhD; Jingqi Liu[5], PhD; Anthea Tinker[1†], PhD; Faith Matcham[6], PhD

[1]Institute of Gerontology, King's College London, London, United Kingdom

[2]School of Healthcare, University of Leeds, Leeds, United Kingdom

[3]Department of Engineering, King's College London, Strand Building, Strand Campus, London, United Kingdom

[4]School of Biomedical Engineering & Imaging Sciences, King's College London, London, United Kingdom

[5]School of Design, University of Leeds, Leeds, United Kingdom

[6]School of Psychology, University of Sussex, Falmer, United Kingdom

[†]deceased

Corresponding Author:
Wei Liu, PhD
Department of Engineering, King's College London, Strand Building, Strand Campus, London, United Kingdom

Abstract

Background: Loneliness is a negative emotional state that is common in later life. The accumulative effects of loneliness have a significant impact on the physical and mental health of older adults. Automatic methods for detection and prediction are an emerging field to support early identification of loneliness.

Objective: This study aimed to qualitatively explore the thoughts and preferences of people aged 65 years and older regarding technologies to detect feelings of loneliness in later life.

Methods: We conducted 60 semistructured interviews with people aged 65 years and older between September 2022 and August 2023. Data were analyzed using a reflective thematic approach on NVivo software (Lumivero).

Results: In total, three themes were identified representing what older adults considered important in a system able to detect loneliness: (1) interest and control of data, which was a priority for older adults; (2) perceived usefulness to address loneliness, which related to the importance of providing recommendations to reduce feelings of loneliness after detection; and (3) personalization as a priority, which included the level of loneliness for which an alert was sent and selection of relevant individuals who would be sent a loneliness alert.

Conclusions: Findings from this in-depth qualitative study provide important perspectives from people with lived experience of loneliness on the context in which a sensor-based loneliness detection system would be most useful and acceptable to older adults. Future research will include such perspectives in the design of innovative technologies enabling the early detection of loneliness and access to timely interventions to tackle loneliness in later life.

International Registered Report Identifier (IRRID): RR2-10.1136/bmjopen-2023-072420

(JMIR Hum Factors 2025;12:e73694) doi:10.2196/73694

KEYWORDS

loneliness; older adults; preferences; qualitative; technology; mobile phone

Introduction

Loneliness is widely acknowledged to be a serious public health matter [1]. Definitions of loneliness vary, but most authors agree that it refers to a negative emotion linked to poor quality and insufficient social relationships [2-4]. This is usually placed as distinct from social isolation, which refers to an "objective" lack of social relationships. The consequences of loneliness can include poor mental health [5], risk of Alzheimer disease [6], and increased mortality [7]. Aging has long been recognized as associated with loneliness [8], with later life even being expected to be a "time of loneliness" [9]. In the United Kingdom, as many as 1 in 4 people older than 65 years are experiencing moderate to high loneliness [10]. Age-related changes such as retirement, bereavement, reduced social networks, and loss of mobility are well-evidenced as risk factors for loneliness [8,11-14]. Older adults may experience transient loneliness, a temporary feeling experienced from time to time, during such age-related changes.

XSL•FO
RenderX

If loneliness were to be felt consistently over a period of time, it would be defined as chronic loneliness, a state associated with worsening physical and mental health outcomes [15].

COVID-19 substantially impacted the feelings of loneliness and social isolation across the population, acting as a catalyst for many older adults to adopt the use of technologies in daily life [16]. For example, technology has been lauded for its potential to aid people in maintaining existing relationships and increasing communication with neighbors, family, and friends [17]. However, work has identified a multitude of barriers to uptake [18], including trust in the technology [19], and family caregiver concerns related to acting upon risks detected by technologies [20]. Issues of data protection, privacy, and control have been recorded in numerous studies looking at the use of digital technologies in health [21]. Automated technologies have the potential to challenge autonomy, particularly if they notify other people without the immediate consent of the older adult being monitored [22]. Despite such challenges, and with the acceleration of health care digitalization, wearable technologies for health monitoring present a timely opportunity to improve care for aging populations [23]. Due to the rapidly evolving digital landscape, the inclusion of older adults in the design of technology is imperative to avoid digital exclusion [24].

In recent years, using smart technology to identify loneliness has emerged as a potentially useful method of tackling loneliness in older adults [25]. Such technologies would monitor physiological or behavioral indicators of loneliness using sensors, enabling timely detection of loneliness to improve access to targeted intervention [26]. As loneliness is conceptually different from social isolation, which would detect the objective presence or absence of company, loneliness detection technologies are required to include several indicators, including stress, sleep, and activity levels, to enable the passive detection of such a subjective emotional state [27,28]. Specifically, sensor-based approaches measure the behavioral markers of social isolation, such as physical inactivity, poor sleep quality, eating behaviors, time out of home, and phone usage [29]. Numerous works have relayed instances of success using machine learning algorithms able to predict emotions using data such as galvanic skin response [30], electrocardiogram data [31], and electroencephalography data [32]. Behavioral data have also been applied to predicting loneliness [18,26,29]. For example, 1 scoping review found studies measuring daily life patterns to detect loneliness using "smart" technology (eg, ambient sensors to measure motion and touch), smartphones, and wearable sensors [19]. Another review of technology aimed at preventing loneliness in older people found that approaches measuring behavior and movement showed promise as they are usually successful at identifying loneliness [25].

The association between physiological markers of loneliness and the need to tackle loneliness in older populations suggests that it may be fruitful to construct technology capable of identifying loneliness in older adults [33]. Such technologies are required to be both useful and acceptable to older adults to be effective in supporting the reduction of feelings of loneliness. Relevant theoretical frameworks suggest that for technology to be accepted and used by older adults, it must be perceived as easy and useful, and users must have appropriate support and resources [34]. This study aims to gain insight into older adults' general acceptance of loneliness detection technologies to identify the most meaningful priorities for this population, that are applicable to future technology development. Previous work has sought to understand older adults' technology preferences and requirements [24], yet a gap remains in relation to older adults' views on automated detection of loneliness specifically.

The potential to detect loneliness in older adults has numerous potential benefits. A wealth of evidence has suggested that loneliness may be a stigmatized condition [35]. Such negative desirability bias means the ability to detect loneliness without a person needing to seek help [36]. This may be particularly useful for older men who do not identify as lonely but suffer consequences of a similar nature [37]. Timely and accurate identification of loneliness in older adults will enable people to receive the right kind of support at the right time [26]. In general, approaches to alleviate loneliness in older adults include leisure activities (ie, gardening, music, and exercise), therapies (ie, animal, reminiscence, and cognitive), social and community interventions (ie, meal sharing), educational approaches (ie, relationship or skills training), befriending, or system-wide approaches [38]. Such activities can be made available to older adults through referral by primary care professionals through social prescribing initiatives [39].

This paper is part of the "Design for Healthy Aging (DELONELINESS)" project that aims to develop a smart monitoring and communication system with multifunctional electronics built into textiles used as wearables and home furniture to detect and measure loneliness levels in later life [26]. As outlined in previous publications, this system will involve the detection of physiological and behavioral indicators of loneliness in later life using sensor-based technologies. Such technologies will be worn by people aged 65 years and older, with data being inputted into an algorithm modeled to predict feelings of loneliness. Once loneliness is identified, the aim would be to enable users to receive support from family or health and social care professionals or promote engagement in activities to alleviate feelings of loneliness [26].

Including the perspectives of older people is central to the design of such a system [40]. In this paper, we present our qualitative findings on the thoughts and preferences of older adults on technologies designed to monitor and detect feelings of loneliness. Our specific objectives are to (1) describe the context and circumstances in which a smart system might be most useful; and (2) identify the most meaningful way of providing information back to individuals, their carers, or health care or social service providers.

Methods

Study Design

One-to-one semistructured interviews were conducted with participants aged 65 years and older on their experiences of loneliness and preferences for loneliness detection technologies.

XSL•FO
RenderX

Detailed methodology used in this study is available in the published protocol [41].

Ethical Considerations

Ethical approval was obtained from Research Ethics Committees at King's College London (reference LRS/DP-21/22 - 33376) and the University of Sussex (reference ER/FM409/3). Informed consent was sought from all participants. Participants were informed that they could stop their participation at any time. Participants were given a pseudonym following consent to ensure confidentiality. All identifiable information was kept in a password-protected Excel sheet on University OneDrive systems and removed at the point of transcription of interviews to protect the privacy of participants. Participants received a £30 (approximately US $40.47) LoveToShop voucher at the end of the interview as a thank you for sharing their experiences for the study. Interviews were audio-recorded and transcribed verbatim by a professional company.

Recruitment

A stratified sampling technique based on age, gender, accommodation type, and level of digital ability was used. In terms of accommodation type, participants either owned their home, rented, or lived in sheltered accommodation, defined as housing for older adults consisting of private independent units with some shared facilities, emergency alert systems, and a warden [42]. We used the communication preferences of participants during the recruitment process (ie, if older adults had an email, mobile phone, or a landline number) as a proxy measure for the level of digital ability. Eligible participants were required to be aged 65 years and older, have experienced loneliness since reaching 65 years in age, be able to give informed consent, and speak English at a level fluent enough to take part in the interview. Study information was disseminated via research participant websites, organization newsletters related to housing, and through communication to individuals from previous research projects (related to depression and aging) who consented to being contacted about future studies. Interested individuals were invited to contact the research team to obtain study details. All participants provided written informed consent before data collection.

Data Collection

Interviews were conducted between September 2022 and August 2023 by a female qualitative researcher and chartered psychologist with expertise in sensitive interviewing. In accordance with participant preferences, interviews were conducted either in-person (at home or in University offices) or remotely (via telephone or video call using Microsoft Teams software). To describe and stratify a range of experiences, we aimed to interview up to 60 individuals with lived experience in later life [41]. This number was chosen to enable the inclusion of a diverse range of participant characteristics related to loneliness, such as age, gender, accommodation type, level of digital ability, and health status. Characteristic data collected from participants included sociodemographic information (eg, age, gender, ethnicity, and accommodation type). We also included a measure of loneliness severity using the de Jong Gierveld Loneliness scale [43], which has 11 items related to social and emotional loneliness that are totaled to obtain an individual's level of loneliness (ie, not lonely, moderately lonely, severely lonely, or very severely lonely).

The interview topic guide was developed by a multidisciplinary team of psychologists, gerontologists, product designers, and machine learning experts. The first part of the interview focused on the subjective experiences of participants to identify the signs and associated behaviors with loneliness in later life. Findings from this section of the interview have been previously published [33]. The latter half of the interview focused on participants' thoughts and preferences on the role of technology to detect loneliness. In line with theoretical models for technology acceptance [44], the interview topic guide contained specific questions related to older adults' perceptions of the usefulness of technologies. Refer to Multimedia Appendix 1 for complete details.

In relation to technology, we asked about participants' existing use of technologies in relation to their health. We then introduced proposed technologies for measuring loneliness, for example, sensors worn on the body or integrated into a fabric which collect physiological data that may be relevant to loneliness (ie, heart rate and activity levels). We asked participants where on their body or in their home they would be willing to have a device. Finally, we asked about data collected by loneliness detection technologies. Specifically, we asked how useful they would find data collected by the system (ie, sleep, movement, and stress levels) and if these data could be sent to anyone to improve participants' quality of life. We then asked about the benefits and limitations of the data to participants, family members, and health care professionals. Finally, we asked about participants' thoughts on receiving an alert to the risk of becoming lonely, and what participants would hope family members or health care professionals would do if they received a loneliness alert. Refer to Multimedia Appendix 1 for the full interview schedule.

Analysis

Data were analyzed using a reflexive thematic approach by the lead author (J Rees) using NVivo software (version 14; Lumivero) [45]. The early structured themes were constructed by identifying "semantic" themes – surface-level ideas built from the stated perspectives of the participants, but which are not thoroughly interrogated [46]. The generation of initial codes was generally grouped in line with the topic guide questions, for example, usefulness of data, benefits and limitations to data, and thoughts on loneliness alert. We used NVivo to organize our themes into the perspectives of "older adults," "family members," or "health care professionals." These were transformed into "reflexive" themes through several intertwined processes. Specifically, we wrote up "semantic" themes fully into a report, which was then shared with the multidisciplinary project team and discussed in team meetings. Such discussions lead to the refinement of themes, specifically distinctions between theme definitions. Themes were reviewed a final time by the second author (J Ratcliffe), who was involved in drafting this paper but was not present during data collection and analysis. To facilitate an iterative process of theme building and refinement, the themes were critically contrasted with extant

XSL•FO
RenderX

literature [47], specifically the literature on interventions for reducing loneliness. In addition, to extrapolate information on "what works for whom in which contexts" [48], participant characteristics, specifically level of loneliness, marital status, and accommodation type, were considered throughout the write-up stage.

Results

Participant Characteristics

In total, 60 older adults were interviewed with a mean age of 73 (SD 6.7) years, ranging from 65 to 91 years. Most participants were aged between 65-74 years, women, White British, homeowners, and with a moderate level of loneliness (Table 1).

Table . Participant demographic data (N=60).

Characteristics	Participants (N=60), n (%)
Age (y)	
65 - 74	42 (70)
75+	18 (30)
Gender	
Women	42 (70)
Men	18 (30)
Ethnicity	
White or White British	54 (90)
Black or Black British	3 (5)
Asian or Asian British	3 (5)
De-Jong-Gierveld scale loneliness level	
Not lonely	3 (5)
Moderate	36 (60)
Severe	10 (17)
Very severe	11 (18)
Accommodation type	
Own home	43 (72)
Rented	10 (17)
Sheltered accommodation	7 (11)

Marital status varied from widowed (n=20), divorced (n=16), single (n=11), married (n=8), and separated (n=5). The frequency of loneliness also varied from often or always (n=23), occasionally (n=16), some of the time (n=20), and never (n=1). Most participants were degree-educated (n=34). In total, 30% (n=18) of participants did not have children, and 13% (n=8) did not have anyone to contact if they were in trouble or needed help. The majority of participants reported having a long-standing illness or disability (n=50), such as depression, chronic fatigue, arthritis, or hearing loss. Furthermore, 5 participants communicated using their landline numbers only.

Themes

Summary of Themes

In total, 3 themes were constructed to represent older adults' thoughts and concerns regarding a loneliness detection device. First, "interest and control of data" outlined the characteristics of older adults interested in a loneliness detection system and the importance of retaining control over data collected to detect loneliness. Second, "perceived usefulness to address loneliness" highlighted the need to provide suggestions to reduce loneliness in conjunction with an alert about loneliness. Third, "personalization as a priority" relays how participants felt the system needed to be tailored to individuals, particularly regarding (1) the level of loneliness that required intervention, and (2) who the participants felt should be informed they are lonely. Refer to Table 2 for theme outline including supporting quotes.

XSL•FO
RenderX

Table . Themes and selected supporting quotes for thoughts and preferences of older adults for technologies designed to detect feelings of loneliness.

Theme	Supporting quotes
Interest and control of data	• "So it's, well it's just, you know, curiosity about how your body is working. And trying to find patterns and things which is obviously what this is trying to do anyway." [P18, Female, 70, Married, Very Severe] • "Well it's simply because I know I would only get annoyed because I know I'm not doing many steps. Although it's interesting too the way it does your heart rate. Now that's something again I would use it." [P15, Female, 66, Divorced, Moderate] • "Can't be bothered with things like that..I'm not going to make a ridiculous effort at carrying technology around to count them because it's kind of not relevant." [P52, Female, 66, Widow, Not Lonely] • "That's the barrier to me. The information, if it's kept simple, if it's something that's user-friendly to me that's not terribly keen to embrace actively too much technology." [P35, Female, Married, Moderate] • "Yes, some things I might know already but it might show some others that I was not already aware of and that is nice you know, it might not be pleasant, but it is happening and that is good." [P23, Male, 68, Separated, Not Lonely] • "The technology is there. I'd like it all to be used for good but I've no control over whether it is or it isn't and I think for my lifestyle and stuff if you've got nothing to hide it doesn't really matter." [P13, Female, 69, Separated, Moderate] • "I'd be a bit- your terminology, 'at risk of becoming lonely,' that's- a bit kind of- big brotherish, really. I suppose- perhaps it's just the wording, if you said, oh we're concerned for you, it's about time you got out a bit more, then- go and find some friends, go out and have a sing song or go out for a run." [P47, Female, 72, Widowed, Severe]
Perceived usefulness to address loneliness	• "I don't think it would be any use to me. because I know how I feel. I know how I felt yesterday, but I'm feeling better today. No, I don't think it would bother me. I'm not a great one for data. It's more likely to be useful to someone else who's trying to find patterns or (I don't know) reasons for being low." [P8, Female, 65, Married, Moderate] • "I don't know that I could do anything about it, so what's the point in telling you, I don't know. I mean it's because they're going to have someone phoning you up and having a chat with you, you know a silver line chat sort of thing. I don't know if that would make me feel less lonely I suppose you might think well someone could be bothered." [P9, Female, 72, Divorced, Very severe] • "Yes, to have the results but also to have the course of how to break that. It's no point telling me I'm lonely which I already know right. A flashing thing saying please, do this, that might be helpful because there is more directive than giving information that somebody is lonely which they most probably know already." [P49, Female, 72, Divorced, Severe] • "If the algorithms or whatever they call them, if it sees that predicted pattern and it happened once before it says, Oh, last time when you were in that- what actually brought those levels down was that you went and did this." [P35, Female, 68, Married, Moderate] • "For me would be an additional motivator because I've got some concrete evidence, so it would be a motivator to say, Look you're feeling lonely now is the time to make that effort to walk up the pub because there's some music going on tonight, and you know you're going to see so and so." [P32, Male, 70, Widowed, Moderate] • "I don't feel I would need that at this moment in time. I don't but we can't see ahead can we, into the future, and if, when I was, if I was much, much older than I am now then I would probably think it would be a good thing to have but at this moment in time I'm not there." [P55, Female, 73, Single, Moderate]

XSL•FO
RenderX

Theme	Supporting quotes
Personalization as a priority	• "And you could choose the point at which it could be shared, if you like, with those people so that, yes, don't bother to tell them if I'm feeling lonely 10% of the time, but maybe if I'm feeling lonely 50% of the time, then do. But it might be different for other people." [P52, Female, 66, Widowed, Not lonely] • "I think it would be good, because I don't get as much time with my GP as I'd like. I think if she was receiving this information and give her something to study about me, so when I did see her, she could refer back and say, Oh I've noticed your sleep isn't so good. I've noticed you've had these waking moments. But I see that your heart rate's good. So that must be good for you, you know?" [P5, Female, 65, Divorced, Severe] • "When you say you'd alert somebody, who would you alert in my case? I've got no one close that you could send an "I need help" message to." [P43, Male, 70, Widowed, Moderate] • "I really wouldn't want her to be worrying about me, she has her own problems, I think that's something I can cope with. If it's only feeling a bit lonely, I can cope with that, and I really don't think it's worth worrying somebody else. I mean, what are they going to do? Get out of work and rush over just to sit in the house and talk to you, you know I don't know." [P25, Male, 68, Divorced, Very Severe]

Theme One: Interest and Control of Data

Older adults spoke about how health data collected by the system (ie, heart rate, sleep, and activity levels) were interesting, comforting, reassuring, and empowering ("knowledge is power"). These participants all had previous positive experience with using smartwatches. However, some participants were not interested in these data either because they had no present health issues and did not wish to worry about their health ("it is not something you can do anything about anyway") or because of their general disinterest in the use of technology ("I am not a great one for data").

Older adults had varied technological ability, with some using virtual reality and online shopping, whereas others had no WiFi or a sense that there is "too much" technology in daily life. These categories were not mutually exclusive. For example, the oldest participant in the sample (aged 91 y) used technology on a daily basis, including a smartwatch or phone, but felt it isolated her further by reducing in-person interactions, which was especially impactful as she lived in a rural area. Similarly, some participants expressed worry about who would access the data and felt being monitored closely would be intrusive, yet others felt comfortable with data, describing themselves as "open minded" with "nothing to hide." Such perspectives were not associated with the age of the participants or the severity of loneliness.

> *I imagine one of your questions is going to be about the use of data and stuff like that. I'm very comfortable with that…I don't think Big Brother's going to have a huge interest in how many times my heart beats.* [P1, Female, 68, Widowed, Moderate]

The main preference described by older adults was the need to be in control of both the data collected and an alert sent by the system. It was important for older adults to understand which data were being collected and for what purpose. Older adults highlighted how it was important to be aware of the behaviors associated with loneliness and to know why they were being alerted.

> *If the algorithms or whatever they call them, if it sees that predicted pattern and it happened once before it says, Oh, last time when you were in that- what actually brought those levels down was that you went and did this.* [P35, Female, 68, Married, Moderate]

It was important for older adults to feel in control of their data. Participants wished for any data or alerts to be sent to them directly and have the option of contacting other people if they deemed it necessary ("I'd much rather it came through me").

> *The less control you have, the more control you need. So, people will control what they can control. And therefore, it's important that the end user has the control over how things are shared.* [P52, Female, 66, Widowed, Not Lonely]

Theme Two: Perceived Usefulness to Address Loneliness

Most participants described already being aware of their loneliness and did not see the benefit in being alerted about loneliness by the system ("I know I'm lonely I don't need an alert"). Linked with the severity of loneliness, some older adults felt that if you did not notice feelings of loneliness, then a person was not truly feeling lonely.

> *You're aware of the build up. If you aren't aware of it, then you're not of that degree of loneliness, you're just feeling low, can I put it that way? Whereas we can all feel low, and have a bad half hour, a bad day, but loneliness is something else that is different.* [P40, Female, 75, Single, Moderate]

Nevertheless, some thought that the system would be a useful prompt as it provided "concrete evidence" to the early warning signs of loneliness and would either "consolidate" their recognition of loneliness or motivate older adults to engage in behaviors to reduce loneliness. Older adults highlighted how it was important to be aware of the behaviors associated with

XSL•FO
RenderX

loneliness and to know why they were being alerted. One participant compared the system with a "companion" for people who felt alone, whereas others felt the system would have been useful in the past (eg, during widowhood and retirement) or in the future (eg, if cognitive ability reduced). The majority of participants (n=45) lived alone, and this group in particular expressed a concern about falling in the home and suggested that the loneliness sensors could also be used to detect this.

Older adults questioned the use of the system if there was no consequence or follow-up to loneliness detection. Participants felt that if they were to be alerted of loneliness, there would also have to be a recommendation to reduce feelings of loneliness. This requirement was consistently expressed across participants, irrespective of age, gender, marital status, or living situation.

> *If you could tell me how to reduce it, what to do, that would be more important than letting me or anybody else know that I was feeling- I know that I'm feeling it or about to feel it. But it's what to do about it that's the critical thing…There's no point in it otherwise.* [P41, Female, 84, Divorced, Severe]

Theme Three: Personalization as a Priority

Subtheme 3a: Level of Loneliness

Older adults found it important to decide on a level of loneliness at which the alert would be sent to themselves. As participants had a variety of loneliness experiences, this level would need to be tailored to individuals. For example, widowed older adults experienced transient loneliness (ie, more severely at certain times), which may be time-limited, depending on individual grief trajectories. Some older adults experienced severe loneliness, where they were more likely to describe chronic loneliness and thus questioned which level of loneliness would be most appropriate for them. One male participant, who experienced chronic loneliness despite living in shared accommodation, felt loneliness at a consistent level at all times, so he questioned what changes the system would be able to detect.

> *You've got to decide on a level of stress that, you know, like we do on a scale of 1 to 10. You know, if I feel 3 amount of lonely, then don't bother, it will probably pass over, but if I feel 7 or 8 amount of loneliness then please tell me to get off my bum and go out for a walk or something.* [P47, Female, 72, Widowed, Severe]

Participants varied in their responses describing how they would engage in activities to alleviate loneliness following an alert. Some participants described preferring to manage feelings of loneliness on their own, especially for low levels of loneliness ("if it's only a bit lonely I can cope"). When discussing potential interventions following a loneliness alert, participants highlighted the importance of personalization as "every person is going to be different."

Subtheme 3b: Involvement of Others

Participants also discussed the importance of personalization in relation to the involvement of others, specifically health care professionals and family members. For example, some participants had previously experienced suicidal thoughts during times of loneliness and felt it was important for health care professionals to be alerted if they experienced loneliness at that level in the future. However, many questioned whether busy health care professionals, such as general practitioners, had the resources to help with loneliness.

> *…the GP, he can't give me a tablet to stop me being lonely or something like that. Can't come round and spend time with me, you know. They haven't got time to see you in the surgery. So, yes I mean I wouldn't mind the GP knowing I was lonely. But I can't see what advantage there would be there.* [P46, Female, 87, Widowed, Moderate]

Nevertheless, some felt the alert could be sent to health care professionals for severe loneliness or where loneliness was related to physical or mental health issues, such as depression, and some mentioned signposting to counseling and social-prescribing services.

Similarly, older adults felt that loneliness would have to be "severe enough" for the system to contact family members and that this level of severity would differ between individuals.

> *And you could choose the point at which it could be shared, if you like, with those people so that, yes, don't bother to tell them if I'm feeling lonely 10% of the time, but maybe if I'm feeling lonely 50% of the time, then do. But it might be different for other people.* [P52, Female, 66, Widowed, Not lonely]

When participants experienced mild feelings of loneliness, many did not want friends or family members to send the loneliness alert, stating they would feel "embarrassed" or "guilty," and that they could cope with it alone to avoid being a "burden." Some participants expressed not having a problem with family receiving a loneliness alert; however, they questioned what they would be able to do about it if they lived far away. In more severe cases of loneliness, participants thought the loneliness alert could be used as a "conversation starter" during lonely periods where older adults felt less able to initiate social contact. However, others still felt the alert was "too private" to share with others without previous consent from the older adults themselves. Vitally, other participants had no family, so they questioned who they could nominate to receive a loneliness alert. This was especially pertinent for single participants over 80 who felt the alert would be an additional burden to next of kin and relied on safety alert systems in sheltered accommodation in an emergency.

> *When you say you'd alert somebody, who would you alert in my case? I've got no one close that you could send an "I need help" message to.* [P43, Male, 70, Widowed, Moderate]

Overall, involvement of family, friends, and health care professionals in a smart system to detect loneliness was thought to be part of a negotiation as to who the older adult would want involved, at what time, and for what purpose.

XSL•FO
RenderX

Discussion

Principal Results

We constructed 3 themes to represent what older adults considered important in a system able to detect loneliness. "Interest and control of data" highlighted general interest in health data being a personal preference of older adults based on current or previous positive use of technologies. Participants had varying technological abilities but were unanimous in their need to be in control of the data collected and sent by the system. "Perceived usefulness to address loneliness" refers to whether and how the system can be useful to older adults. Many older adults described being aware of feelings of loneliness, especially in severe cases, and thus questioned the usefulness of the loneliness detection system. Participants emphasized the importance of providing recommendations to reduce loneliness following detection to improve usefulness. Finally, "personalization as a priority" refers to the need to work with the individual using the system to determine the level of loneliness that results in an alert, and the consequences of that alert, including the involvement of others to support loneliness, such as friends, family members, or health and social care professionals.

The focus on control echoes concerns over privacy and data protection reported in previous studies on technology use in later life [22,49,50]. Loneliness has been associated with a lack of autonomy [51]; therefore, it is vital that older adults feel in control of the data collected and sent by a loneliness detection system. Trust in technologies can be established through providing feedback in real time and through extensive validation [17,19]. Technology-based approaches for reducing social isolation may not be appropriate for all older adults [52]. As a "digital divide" exists between those with cultural and financial access to technology and those without [24], research has found that lonelier older adults use information and communication technologies less often than their less lonely counterparts [53]. Factors that were found to influence engagement include lack of skills and support to use digital health technologies [54] and a fear or dislike of technology [21]. Ensuring that older adults across a spectrum of abilities feel knowledgeable and confident with the data and associated alerts from the loneliness detection system, is therefore paramount to its use.

The findings of this study suggest a key component of ensuring acceptability of a loneliness detection system is the effective implementation of personalization. Our findings support previous literature, which suggests that digital health systems should make it transparent as to who will gain access to the information collected, and ensure users retain control of who has access to it [55]. A novel finding from our work is the importance of determining an individually selected level of loneliness for which an alert would be sent, highlighting the subjective nature of loneliness [15]. Further consideration is required as to how degrees of loneliness would be measured and reported to the system, for example, self-reporting from older adults or the use of validated loneliness scales [56]. More broadly, our findings on personalization indicate the importance of "user-centered" design in health technology (perspectives that centralize the needs of the user in the design and implementation of the technology) [29,57,58]. This is especially important in the context of culture, which plays an important role in shaping an individual's experience of loneliness [59]. Developing culturally sensitive interventions to reduce feelings of loneliness is an important priority for future research to ensure targeted support is available for older adults.

In relation to the involvement of others in supporting loneliness once detected, older adults emphasized that whoever the system alerted, this should be decided in advance via a clear and transparent set of discussions. Participants tended to refer to general practitioners when health professionals were discussed, a finding that may mirror the trust instilled in doctors by older people [60,61]. Previous research on technology preferences of older adults has identified that if a device was recommended to older adults by a "credible" health professional, it was more likely to be taken up and used [24]. The participants in this study generally held negative attitudes toward alerting family members of loneliness, mirroring concerns from previous research [62]. However, friends or family were not consulted in this study; therefore, it is possible that their thoughts on receiving an alert for loneliness may not align with those of older adults. Perceptions of burden were described as one reason for participant disapproval of family member involvement. Indeed, loneliness has often been linked with guilt, shame, and a perception that others do not care [3,63]. Further research is required on the involvement of family or friends in providing support following an alert from a loneliness detection system, as many loneliness reduction interventions actively involve others [64].

Perceived usefulness is a core component of theoretical models of technology adoption in older adults [44,65]. Participants had varied perspectives on whether and how a loneliness detection system could be useful. Some saw it as potentially useful as a prompt to change their own actions or as an "early-warning" sign before they would have thought to seek help, perhaps even before they realized that they were lonely. Given the stigma of loneliness [3] and the reticence of older men to identify as lonely [63], such a system may be particularly useful for removing the anxiety and uncertainty of doing something about loneliness. Older adults in our study were clear that recommendations to reduce feelings of loneliness once detected were an imperative aspect of such a system. Research has highlighted how there is "no one size fits all" approach to loneliness intervention and that tailored, targeted approaches are required to reduce loneliness (Victor et al [66]). Detailed consideration is required on the selection of appropriate interventions for older adults using a loneliness detection system based on the type of loneliness (ie, emotional or social), severity of loneliness (ie, moderate, severe, or very severe), and reason for loneliness (ie, bereavement or mobility issues). Future research will be required on the acceptability and effectiveness of such recommendations, noting that individual-centered interventions are inherently limited by their lack of capacity to influence the social context of the lonely individual [3].

XSL•FO
RenderX

Limitations

This study adds to the emerging literature on sensor-based technologies to detect feelings of loneliness in older adults [18,29]. Our large sample size enabled us to present a variety of perspectives and incorporate a diverse range of participant characteristics, including those who were skeptical or ambivalent about the technology. Despite efforts to achieve an equal distribution of gendered and cultural perspectives, our sample was majority female White-British. Our themes may therefore not represent the views of older adults with different perspectives on technologies, which requires further consideration in future design testing. Our interview sample included people with lived experience of loneliness to provide relevant perspectives and expertise to inform future technology development [67]. It is important to acknowledge that such participants were more open to acknowledging and seeking help for loneliness than some of the prospective target population, who may not readily seek help for loneliness. Social desirability bias may have also influenced how participants spoke about technology. Thus, a gap exists in our knowledge of older adults with reduced help-seeking behavior that may particularly benefit the use of a loneliness detection system. Finally, the perspectives of participants were based on a hypothetical loneliness detection device described by the lead author conducting the interviews. There is no such actual device yet; therefore, potential barriers and enablers identified in related research, such as comfort, cost, and maintenance [68,69], could not be effectively discussed or reviewed. Interpretations of the loneliness detection system may vary among participants, and further user testing would be required to obtain information on acceptability once a device to detect physiological markers associated with loneliness has been developed [33].

Conclusion

Findings from this large qualitative study provide important perspectives from people with lived experience of loneliness on the context in which a sensor-based loneliness detection system would be most useful and acceptable to older adults. It is important to support older adults across a spectrum of abilities and preferences to feel knowledgeable and confident with technologies used by a future loneliness detection system. For older adults to feel in control, transparency on data collected by the system is required in addition to options for personalization related to levels of loneliness, which creates an alert, and which supporters are informed. This should be identified through a clear and full discussion, before they commence using it, of when, why, and who the device should alert in the event of it detecting loneliness. The perceived usefulness of a future loneliness detection system is reliant on activities to reduce feelings of loneliness being recommended following the identification of loneliness. Future research will include such perspectives in the design of innovative technologies enabling the early detection of loneliness and access to timely interventions to tackle loneliness in later life.

Acknowledgments

The authors would like to dedicate this paper to our extraordinary colleague, Anthea Tinker, who passed away in August 2025. Anthea was a dedicated member of the DELONELINESS team, with a wealth of knowledge from her career in gerontology. Above all she was a kind and supportive person who will be greatly missed.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Topic guide and semistructured interview.
[DOCX File, 51 KB - humanfactors_v12i1e73694_app1.docx]

References

1. Cacioppo JT, Cacioppo S. The growing problem of loneliness. The Lancet 2018 Feb;391(10119):426. [doi: 10.1016/S0140-6736(18)30142-9]
2. Perlman D, Peplau LA. Toward a social psychology of loneliness. In: Gilmour R, Duck S, editors. Personal Relationships: 3 Relationships in Disorder: Academic Press; 1981:31-56.
3. Barreto M, Doyle DM, Qualter P. Changing the narrative: loneliness as a social justice issue. Polit Psychol 2024 Apr;45(S1):157-181. [doi: 10.1111/pops.12965]
4. Cacioppo JT, Cacioppo S. Loneliness in the modern age: an evolutionary theory of loneliness (ETL). In: Adv Exp Soc Psychol 2018, Vol. 58:127-197. [doi: 10.1016/bs.aesp.2018.03.003]
5. Schinka KC, Van Dulmen MHM, Bossarte R, Swahn M. Association between loneliness and suicidality during middle childhood and adolescence: longitudinal effects and the role of demographic characteristics. J Psychol 2012;146(1-2):105-118. [doi: 10.1080/00223980.2011.584084] [Medline: 22303615]
6. Lam JA, Murray ER, Yu KE, et al. Neurobiology of loneliness: a systematic review. Neuropsychopharmacology 2021 Oct;46(11):1873-1887. [doi: 10.1038/s41386-021-01058-7] [Medline: 34230607]
7. Victor CR, Bowling A. A longitudinal analysis of loneliness among older people in Great Britain. J Psychol 2012;146(3):313-331. [doi: 10.1080/00223980.2011.609572] [Medline: 22574423]

XSL•FO
RenderX

8. Victor CR, Yang K. The prevalence of loneliness among adults: a case study of the United Kingdom. J Psychol 2012;146(1-2):85-104. [doi: 10.1080/00223980.2011.613875] [Medline: 22303614]
9. Pikhartova J, Bowling A, Victor C. Is loneliness in later life a self-fulfilling prophecy? Aging Ment Health 2016;20(5):543-549. [doi: 10.1080/13607863.2015.1023767] [Medline: 25806794]
10. Chawla K, Kunonga TP, Stow D, Barker R, Craig D, Hanratty B. Prevalence of loneliness amongst older people in high-income countries: a systematic review and meta-analysis. PLoS ONE 2021;16(7):e0255088. [doi: 10.1371/journal.pone.0255088] [Medline: 34310643]
11. Tijhuis MAR, Jong-Gierveld J, Feskens EJM. Changes in and factors related to loneliness in older men. The Zutphen elderly study. Age Ageing 1999 Sep 1;28(5):491-495. [doi: 10.1093/ageing/28.5.491]
12. Alpass FM, Neville S. Loneliness, health and depression in older males. Aging Ment Health 2003 May;7(3):212-216. [doi: 10.1080/1360786031000101193] [Medline: 12775403]
13. Barjaková M, Garnero A, d'Hombres B. Risk factors for loneliness: a literature review. Soc Sci Med 2023 Oct;334:116163. [doi: 10.1016/j.socscimed.2023.116163]
14. Cohen-Mansfield J, Hazan H, Lerman Y, Shalom V. Correlates and predictors of loneliness in older-adults: a review of quantitative results informed by qualitative insights. Int Psychogeriatr 2016 Apr;28(4):557-576. [doi: 10.1017/S1041610215001532] [Medline: 26424033]
15. Mansfield L, Victor C, Meads C, et al. A conceptual review of loneliness in adults: qualitative evidence synthesis. Int J Environ Res Public Health 2021 Nov 2;18(21):11522. [doi: 10.3390/ijerph182111522] [Medline: 34770035]
16. Elimelech OC, Ferrante S, Josman N, et al. Technology use characteristics among older adults during the COVID-19 pandemic: a cross-cultural survey. Technol Soc 2022 Nov;71:102080. [doi: 10.1016/j.techsoc.2022.102080] [Medline: 35991080]
17. Morgavi G. A virtual village network architecture for improving the elderly people quality of life. Inteligencia Artificial 2015;18(55):26-34. [doi: 10.4114/intartif.vol18iss55pp26-34]
18. Site A, Lohan ES, Jolanki O, et al. Managing perceived loneliness and social-isolation levels for older adults: a survey with focus on wearables-based solutions. Sensors (Basel) 2022 Feb 1;22(3):1108. [doi: 10.3390/s22031108] [Medline: 35161852]
19. Qirtas MM, Zafeiridi E, Pesch D, White EB. Loneliness and social isolation detection using passive sensing techniques: scoping review. JMIR Mhealth Uhealth 2022 Apr 12;10(4):e34638. [doi: 10.2196/34638] [Medline: 35412465]
20. Wrede C, Braakman-Jansen A, van Gemert-Pijnen L. Understanding acceptance of contactless monitoring technology in home-based dementia care: a cross-sectional survey study among informal caregivers. Front Digit Health 2023;5:1257009. [doi: 10.3389/fdgth.2023.1257009] [Medline: 37860038]
21. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021 Aug 17;21(1):1556. [doi: 10.1186/s12889-021-11623-w] [Medline: 34399716]
22. Villaverde Naveira A, de Masi A, Wac K, et al. Designing a social robot companion to support homecare: usability results. Stud Health Technol Inform 2022 Jun 29;295:261-264. [doi: 10.3233/SHTI220712] [Medline: 35773858]
23. Chen C, Ding S, Wang J. Digital health for aging populations. Nat Med 2023 Jul;29(7):1623-1630. [doi: 10.1038/s41591-023-02391-8]
24. LaMonica HM, Davenport TA, Roberts AE, Hickie IB. Understanding technology preferences and requirements for health information technologies designed to improve and maintain the mental health and well-being of older adults: participatory design study. JMIR Aging 2021 Jan 6;4(1):e21461. [doi: 10.2196/21461] [Medline: 33404509]
25. Latikka R, Rubio-Hernández R, Lohan ES, et al. Older adults' loneliness, social isolation, and physical information and communication technology in the era of ambient assisted living: a systematic literature review. J Med Internet Res 2021 Dec 30;23(12):e28022. [doi: 10.2196/28022] [Medline: 34967760]
26. Rees J, Matcham F, Probst F, et al. Wearables, sensors and the future of technology to detect and infer loneliness in older adults. Gerontechnology 2023 Aug 31;22(2):1-4. [doi: 10.4017/gt.2023.22.2.ree.08]
27. Doryab A, Villalba DK, Chikersal P, et al. Identifying behavioral phenotypes of loneliness and social isolation with passive sensing: statistical analysis, data mining and machine learning of smartphone and Fitbit data. JMIR Mhealth Uhealth 2019 Jul 24;7(7):e13209. [doi: 10.2196/13209] [Medline: 31342903]
28. Cacioppo JT, Ernst JM, Burleson MH, et al. Lonely traits and concomitant physiological processes: the MacArthur social neuroscience studies. Int J Psychophysiol 2000 Mar;35(2-3):143-154. [doi: 10.1016/s0167-8760(99)00049-5] [Medline: 10677643]
29. Prabhu D, Kholghi M, Sandhu M, et al. Sensor-based assessment of social isolation and loneliness in older adults: a survey. Sensors (Basel) 2022 Dec 16;22(24):9944. [doi: 10.3390/s22249944] [Medline: 36560312]
30. Domínguez-Jiménez JA, Campo-Landines KC, Martínez-Santos JC, Delahoz EJ, Contreras-Ortiz SH. A machine learning model for emotion recognition from physiological signals. Biomed Signal Process Control 2020 Jan;55:101646. [doi: 10.1016/j.bspc.2019.101646]
31. Fan T, Qiu S, Wang Z, et al. A new deep convolutional neural network incorporating attentional mechanisms for ECG emotion recognition. Comput Biol Med 2023 Jun;159:106938. [doi: 10.1016/j.compbiomed.2023.106938] [Medline: 37119553]

XSL•FO
RenderX

32. Gu Y, Zhong X, Qu C, Liu C, Chen B. A domain generative graph network for EEG-based emotion recognition. IEEE J Biomed Health Inform 2023 May;27(5):2377-2386. [doi: 10.1109/JBHI.2023.3242090] [Medline: 37022448]
33. Rees J, Liu W, Canson J, et al. Qualitative exploration of the lived experiences of loneliness in later life to inform technology development. Int J Qual Stud Health Well-being 2024 Dec;19(1):2398259. [doi: 10.1080/17482631.2024.2398259] [Medline: 39305060]
34. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003 Sep;27(3):425-478. [doi: 10.2307/30036540]
35. Barreto M, van Breen J, Victor C, et al. Exploring the nature and variation of the stigma associated with loneliness. J Soc Pers Relat 2022 Sep;39(9):2658-2679. [doi: 10.1177/02654075221087190] [Medline: 35991527]
36. Austin J, Dodge HH, Riley T, Jacobs PG, Thielke S, Kaye J. A smart-home system to unobtrusively and continuously assess loneliness in older adults. IEEE J Transl Eng Health Med 2016;4:2800311. [doi: 10.1109/JTEHM.2016.2579638] [Medline: 27574577]
37. Ratcliffe J, Galdas P, Kanaan M. Older men and loneliness: a cross-sectional study of sex differences in the English longitudinal study of ageing. BMC Public Health 2024 Feb 2;24(1):354. [doi: 10.1186/s12889-024-17892-5] [Medline: 38308255]
38. Loneliness interventions across the life-course: executive summary. GOV.UK. URL: https://www.gov.uk/government/publications/exploring-interventions-to-tackle-loneliness/loneliness-interventions-across-the-life-course-executive-summary [accessed 2025-09-03]
39. Paquet C, Whitehead J, Shah R, et al. Social prescription interventions addressing social isolation and loneliness in older adults: meta-review integrating on-the-ground resources. J Med Internet Res 2023 May 17;25:e40213. [doi: 10.2196/40213] [Medline: 37195738]
40. Probst F, Rees J, Aslam Z, et al. Evaluating a smart textile loneliness monitoring system for older people: co-design and qualitative focus group study. JMIR Aging 2024 Dec 17;7:e57622. [doi: 10.2196/57622] [Medline: 39688889]
41. Rees J, Liu W, Ourselin S, et al. Understanding the psychological experiences of loneliness in later life: qualitative protocol to inform technology development. BMJ Open 2023 Jun 19;13(6):e072420. [doi: 10.1136/bmjopen-2023-072420] [Medline: 37336536]
42. What is sheltered housing? Age UK. URL: https://www.ageuk.org.uk/information-advice/care/housing-options/sheltered-housing/ [accessed 2025-07-02]
43. De Jong Gierveld J, Van Tilburg T. The De Jong Gierveld short scales for emotional and social loneliness: tested on data from 7 countries in the UN generations and gender surveys. Eur J Ageing 2010 Jun;7(2):121-130. [doi: 10.1007/s10433-010-0144-6] [Medline: 20730083]
44. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
45. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qualitative Research in Sport, Exercise and Health 2019 Aug 8;11(4):589-597. [doi: 10.1080/2159676X.2019.1628806]
46. Javadi M, Zarea K. Understanding thematic analysis and its pitfall. J Clie Care 2016;1(1):34-40. [doi: 10.15412/J.JCC.02010107]
47. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
48. Pawson R, Tilley N. Realist Evaluation: Sage Publications; 1997.
49. Davis K, Owusu EB, Marcenaro L, et al. Pervasive sensing for social connectedness. In: Enhanced Living Environments: From Models to Technologies 2017, Vol. 3:49-79. [doi: 10.1049/PBHE010E_ch3]
50. Wrede C, Braakman-Jansen A, van Gemert-Pijnen L. Requirements for unobtrusive monitoring to support home-based dementia care: qualitative study among formal and informal caregivers. JMIR Aging 2021;4(2):e26875. [doi: 10.2196/26875]
51. Henning G, Segel-Karpas D, Bjälkebring P, Berg AI. Autonomy and loneliness - longitudinal within- and between-person associations among Swedish older adults. Aging Ment Health 2022 Dec;26(12):2416-2423. [doi: 10.1080/13607863.2021.2000937] [Medline: 34751074]
52. Chen YRR, Schulz PJ. The effect of information communication technology interventions on reducing social isolation in the elderly: a systematic review. J Med Internet Res 2016 Jan 28;18(1):e18. [doi: 10.2196/jmir.4596] [Medline: 26822073]
53. Talmage CA, Knopf RC, Wu T, Winkel D, Mirchandani P, Candan KS. Decreasing loneliness and social disconnectedness among community-dwelling older adults: the potential of information and communication technologies and ride-hailing services. Act Adapt Aging 2021 Apr 3;45(2):89-117. [doi: 10.1080/01924788.2020.1724584]
54. Chadwick H, Laverty L, Finnigan R, et al. Engagement with digital health technologies among older people living in socially deprived areas: qualitative study of influencing factors. JMIR Form Res 2024 Dec 26;8:e60483. [doi: 10.2196/60483] [Medline: 39723998]
55. Alaoui M, Lewkowicz M. Practical issues related to the implication of elderlies in the design process – The case of a Living Lab approach for designing and evaluating social TV services. IRBM 2015 Oct;36(5):259-265. [doi: 10.1016/j.irbm.2015.06.002]

XSL•FO
RenderX

56. Campaign to end loneliness. Measuring your impact on loneliness in later life. : Esmée Fairbairn Foundation; 2015 URL: https://www.campaigntoendloneliness.org/wp-content/uploads/Loneliness-Measurement-Guidance1.pdf [accessed 2023-05-23]
57. Harte RP, Glynn LG, Broderick BJ, et al. Human centred design considerations for connected health devices for the older adult. J Pers Med 2014 Jun 4;4(2):245-281. [doi: 10.3390/jpm4020245] [Medline: 25563225]
58. Probst F, Ratcliffe J, Molteni E, et al. A scoping review on human-centered design approaches and considerations in the design of technologies for loneliness and social isolation in older adults. Des Sci 2024;10:e39. [doi: 10.1017/dsj.2024.22]
59. Hartanto A, Sim L, Lee D, Majeed NM, Yong JC. Cultural contexts differentially shape parents' loneliness and wellbeing during the empty nest period. Commun Psychol 2024 Nov 3;2(1):105. [doi: 10.1038/s44271-024-00156-8] [Medline: 39489810]
60. Appleby J, Robertson R. Public satisfaction with the NHS in 2015 results and trends from the british social attitudes survey. : The Kings Fund; 2016 URL: https://www.kingsfund.org.uk/insight-and-analysis/reports/public-satisfaction-nhs-2015 [accessed 2023-05-21]
61. Butterworth JE, Campbell JL. Older patients and their GPs: shared decision making in enhancing trust. Br J Gen Pract 2014 Nov;64(628):e709-e718. [doi: 10.3399/bjgp14X682297] [Medline: 25348995]
62. Neves BB, Franz RL, Munteanu C, Baecker R. Adoption and feasibility of a communication app to enhance social connectedness amongst frail institutionalized oldest old: an embedded case study. Information, Communication & Society 2018 Nov 2;21(11):1681-1699. [doi: 10.1080/1369118X.2017.1348534]
63. Ratcliffe J, Wigfield A, Alden S. 'A lonely old man': empirical investigations of older men and loneliness, and the ramifications for policy and practice. Ageing Soc 2021 Apr;41(4):794-814. [doi: 10.1017/S0144686X19001387]
64. Gardiner C, Geldenhuys G, Gott M. Interventions to reduce social isolation and loneliness among older people: an integrative review. Health Soc Care Community 2018 Mar;26(2):147-157. [doi: 10.1111/hsc.12367] [Medline: 27413007]
65. Schroeder T, Dodds L, Georgiou A, Gewald H, Siette J. Older adults and new technology: mapping review of the factors associated with older adults' intention to adopt digital technologies. JMIR Aging 2023 May 16;6:e44564. [doi: 10.2196/44564] [Medline: 37191976]
66. Victor C, Mansfield L, Kay T, et al. An overview of reviews: the effectiveness of interventions to address loneliness at all stages of the life course. : What Works Centre for Wellbeing; 2018 URL: https://whatworkswellbeing.org/wp-content/uploads/2020/01/Full-report-Tackling-loneliness-Oct-2018_0151580300.pdf [accessed 2023-05-21]
67. Muchamore I, Karanikolas P, Gooding PM. How lived experience expertise shapes research and development in digital mental health: a review of the literature and insights. Wellcome Open Res 2024;9:373. [doi: 10.2139/ssrn.4932039]
68. Ambagtsheer RC, Borg K, Townsin L, et al. The effectiveness of technology interventions in reducing social isolation and loneliness among community-dwelling older people: a mixed methods systematic review. Archives of Gerontology and Geriatrics Plus 2024 Mar;1(1):100008. [doi: 10.1016/j.aggp.2024.100008]
69. Corbett CF, Wright PJ, Jones K, Parmer M. Voice-activated virtual home assistant use and social isolation and loneliness among older adults: mini review. Front Public Health 2021;9:742012. [doi: 10.3389/fpubh.2021.742012] [Medline: 34708017]

Edited by A Choudhury; submitted 10.03.25; peer-reviewed by A Hartanto, C Paquet; revised version received 28.07.25; accepted 29.07.25; published 18.09.25.

Please cite as:
Rees J, Ratcliffe J, Liu W, Zhou Y, Ourselin S, Antonelli M, Patel A, Shi Y, Liu J, Tinker A, Matcham F
Understanding the Thoughts and Preferences for Technologies Designed to Detect Feelings of Loneliness: Interview Study Among Older Adults
JMIR Hum Factors 2025;12:e73694
URL: https://humanfactors.jmir.org/2025/1/e73694
doi:10.2196/73694

© Jessica Rees, John Ratcliffe, Wei Liu, Yi Zhou, Sebastien Ourselin, Michela Antonelli, Ashay Patel, Yu Shi, Jingqi Liu, Anthea Tinker, Faith Matcham. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 18.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Perspectives of Older Adults on Assistive Technology: Qualitative Study

Mirou Jaana[1], PhD; Maude Lévesque Ryan[2], MSc; Haitham Tamim[3], PhD; Edward Riachy[1], MD; Guy Paré[4], PhD

[1]Telfer School of Management, University of Ottawa, 55 Laurier Ave E, Ottawa, ON, Canada

[2]Department of public health-infectious diseases, Centre intégré de santé et de services sociaux de Chaudière-Appalaches, Sainte-Marie, QC, Canada

[3]School of Business and Hospitality, Algonquin College, Ottawa, ON, Canada

[4]Department of Information Technology, HEC Montréal, Montréal, QC, Canada

Corresponding Author:
Mirou Jaana, PhD
Telfer School of Management, University of Ottawa, 55 Laurier Ave E, Ottawa, ON, Canada

Abstract

Background: The aging population presents challenges for health care systems. Assistive technologies (ATs), such as telemonitoring, fall detection, and self-monitoring devices, offer potential solutions to support older adults and their care. However, successful implementation relies on their acceptance, which remains poorly understood, particularly among nonusers.

Objective: This study aimed to explore older adults' perceptions of ATs, including perceived benefits, adoption barriers, and factors influencing willingness to use these technologies.

Methods: A qualitative study was conducted with 31 participants (aged ≥65 years) with varying levels of health and care needs. Data were collected through 6 focus groups and 6 in-depth interviews and then analyzed thematically using NVivo software.

Results: A total of 7 themes emerged: (1) limited familiarity, with greater recognition of fall detection and self-monitoring devices compared to telemonitoring; (2) perceived benefits, including safety, independence, and chronic disease management; (3) key concerns, including usability, cost, reliability, privacy, and psychological impacts; (4) suggested improvements, including user-friendly designs and training programs; (5) contextual influences identified with independent older adults perceiving greater utility; (6) strategies for ATs' promotion proposed, such as media campaigns, government subsidies, and health care endorsements; and (7) overall willingness to adopt ATs, driven by perceived need, social and health care influence, and ease of use.

Conclusions: Although ATs offer clear benefits, adoption remains limited due to usability, cost, and psychological concerns. Improving accessibility, training, and integration into traditional health care service delivery may facilitate acceptance and use. Future research should focus on inclusive designs and policy interventions to maximize ATs' potential in aging populations.

(JMIR Hum Factors 2025;12:e74214) doi:10.2196/74214

KEYWORDS

aging; assistive technology; telehealth; telemonitoring; self-monitoring; fall detection; patient acceptance; technology adoption; qualitative research; focus groups; interviews

Introduction

Background

The global population is aging rapidly, posing challenges to health care systems. In Canada, adults aged ≥65 years comprised 14% of the population in 2009 and are expected to reach 23% by 2030 [1]. Providing them with proper care and enabling them to age comfortably at home is crucial, with significant societal and economic implications [2]. Effective management of chronic conditions and health deterioration can improve well-being [3] while reducing costs related to institutional care [4-6].

Assistive technologies (ATs) offer solutions to support older adults' health and independence [3-5,7-9]. Telemonitoring enhances quality of life and security, particularly for those with chronic diseases [4,5,10-15]. Fall detection devices help prevent injuries and allow for timely interventions, reducing long hospital stays [5,8,16]. Self-monitoring devices promote healthier lifestyles by tracking movement and activity levels, providing real-time insights for better care [3,5,17,18].

Nevertheless, concerns remain about older adults' acceptance of these technologies [3,6-8,19-22], as much of the research focuses primarily on technological innovations or those using these technologies. For ATs to be effective, older adults must be inclined to use them, which depends on acceptance [7,22], motivation, and understanding of their use [8]. Placing them at the center of research ensures that these technologies are practical and accessible [23-31]. Nevertheless, studying older adults, particularly the more vulnerable subgroup, is challenging, leading to knowledge gaps, particularly regarding the

XSL•FO
RenderX

perspectives of nonusers [4,32,33]. Prior studies have mostly focused on existing AT users [26,34-37] or specific health conditions such as chronic obstructive pulmonary disease [38-40], hypertension [36,41], heart disease [39] and diabetes [36,42], limiting insights into the broader older population's views [43].

This study aimed to address this gap by investigating how older adults, with various health statuses and experiences, perceive and approach AT types—telemonitoring, fall detection technology, and self-monitoring devices.

Literature Review

Technology-enabled health care is an umbrella term referring to technologies (eg, information and communication, mobile technologies, etc.) that are designed and tailored to provide care, improve the quality of health services, and promote well-being [44]. eHealth is another term that is used in reference to information and communication technologies used for health and encompasses a variety of mobile health apps [45,46] and internet use for health services and information delivery [47,48]. Digital ATs (eg, communication systems, ambient technologies or monitoring devices, mobile apps, etc.) represent the products or systems used to "help maintain or improve an individual's functioning related to cognition, communication, hearing, mobility, self-care, and vision, thus enabling their health, well-being, inclusion, and participation" [49].

ATs present benefits to older adults' care, given their ability to support health care professionals' work (eg, by providing older adults with assistance with activities of daily living), assist older adults in completing tasks that would otherwise not be possible (eg, mobility), and enable older adults to feel safe while independently living at home [50]. In a recent systematic review of randomized controlled trials, ATs were found to be effective for personal disease management (eg, statistically significant positive effects on self-care and health-related quality of life), medication management (eg, statistically significant improvements in medication adherence and a reduction in the number of missed doses), and vision (eg, statistically significant improvements in visual functions with the use of ATs) in older adults [51]. Prior studies have also highlighted other benefits for older adults, including reduced falls [52], more autonomy and independence [53], improved social connection and engagement, and enhanced well-being [54] when using ATs.

Despite the demonstrated benefits and effectiveness of these technologies, older adults' use of ATs remains low, particularly among those with multimorbidity who may benefit most from their use [51]. Previous studies have investigated older adults' acceptance of ATs. For example, Shin et al [55] assessed South Korean older adults' acceptance of ATs using the Senior Technology Acceptance Model. They found that gerontechnology self-efficacy, health contexts and abilities, attitudinal factors, perceived usefulness, and perceived ease of use statistically affected older adults' intention to use ATs [55]. A recently published meta-analysis reviewed the literature that used the Technology Acceptance Model and Unified Theory of Acceptance and Use of Technology (UTAUT) to assess older adults' acceptance of technology [56]. They found that perceived usefulness, social influence, and ease of use all significantly impact older adults' acceptance of ATs, although this relationship is moderated by the type of AT considered [56], with high heterogeneity across the studies considered. As a result, the authors recommended that future research explore the variations in acceptance patterns across different contexts and technologies to lead to a more comprehensive understanding of older adults' AT acceptance [56].

While the existing literature using the TAM, Senior Technology Acceptance Model (STAM), UTAUT, and other behavioral theories has contributed to our understanding of older adults' technology acceptance, there have been calls in the literature for more qualitative research that explores dimensions beyond those included in existing theories [57]. For example, Felber et al [57] discussed that research often quantitatively validates these theories, without consideration of additional issues that are relevant to older adults (eg, loneliness, the fear of ATs replacing human care, etc.). As such, qualitative research would be appropriate to allow additional factors that may influence acceptance to emerge from the data [58]. They are best suited for studies aiming to understand the reasons behind why an older adult may or may not accept a technology and for providing a more detailed understanding of the nuances that are often missed when quantitative approaches are used [57,58]. This study provides new insights into this area of research by exploring how older adults in long-term care (LTC) environments perceive and approach 3 different AT types (ie, telemonitoring, fall detection technology, and self-monitoring devices) using qualitative methods.

Methods

Study Setting

The study was conducted at Perley Health, Ottawa's largest LTC home. Perley Health is a comprehensive community of care that is home to more than 600 seniors and veterans across 450 LTC beds and 139 independent living apartments. It is also a center for research, education, and clinical innovation and offers many programs and services to residents, tenants, and people from across the region. As part of a pilot project from 2018 to 2021, Perley Health had a subacute unit for frail elderly (SAFE), which was a 20-bed specialty care unit that cared for patients recovering from surgery, illness, or other acute health events, who no longer needed to be in hospital but were not ready to go home. Patients typically stayed on this unit for 2 to 3 weeks to access medical and rehabilitative supports and then returned to their place of residence in the community.

Recruitment began in late October 2018 and concluded in April 2019. The process was adapted to participants' autonomy levels. LTC and SAFE residents required multiple visits to build trust, address hesitations, and encourage engagement. For ILA residents, recruitment involved flyers and group sessions to present the study and promote participation.

Study Design and Population

We used a qualitative research design to capture diverse perspectives of older adults with varying levels of independence and care needs and residing in different environments. This approach combined focus groups, encouraging discussion and

XSL•FO
RenderX

interaction [59,60], with in-depth interviews for a deeper exploration of individual experiences [60].

Participants aged ≥65 years were selected through purposive sampling to ensure diversity in health status, living arrangements, and technology familiarity. Inclusion criteria required cognitive capability to be able to provide informed consent, which was usually assessed using the Cognitive Performance Scale [61]. Individuals with moderate to severe impairment were excluded because of their inability to participate.

Data Collection and Analysis

We conducted 6 focus groups (4 - 6 participants each), grouping individuals by autonomy level (eg, ILA vs LTC). Discussions covered experiences with ATs, perceived benefits, challenges, usefulness, willingness to adopt the technologies, and suggestions for improvement. As part of the original study design, we also conducted 6 in-depth interviews with participants who were unable to attend group discussions. These interviews provided deeper insights into specific themes and ensured a balanced representation of opinions.

The 3 ATs explored in this study were selected for their relevance to aging in place and their practical applicability for older adults in institutional or community care settings. Telemonitoring refers to remote systems that track vital signs or health conditions and transmit data to health care providers [62]. Fall detection devices are typically wearable sensors that automatically alert caregivers or emergency services when a fall is detected [63]. Self-monitoring tools, such as GPS-enabled trackers or activity monitors, allow users and family members to track location or daily activity for enhanced safety [64].

Other digital health tools, such as mobile health apps, virtual reality, or digital twin systems, were not included, as they were considered less familiar or accessible to the study population. The selected technologies reflect those most commonly promoted for older adult safety, autonomy, and chronic condition support in residential and home-based care contexts.

To facilitate discussion, participants watched short, brand-neutral, publicly available videos (1 - 2 min each) introducing each technology. These videos emphasized core features, use cases, and benefits while reducing cognitive effort and improving comprehension. Following each video, a series of open-ended and targeted questions were posed using a semistructured discussion guide. These questions were designed to explore perceived benefits (eg, timely interventions and independence), concerns (eg, cost, usability, privacy, and reliability), and potential adoption triggers (eg, health status changes and physician recommendation). Participants were specifically asked whether they believed the technology could enable timely emergency intervention and what limitations or barriers they perceived. The guide also differentiated between those who had previously used the technology and those who had not, allowing for comparison between experienced users and nonusers. Questions were adapted to each device type (telemonitoring, fall detection, and self-tracking) and to the participant's care setting.

Focus groups and interviews lasted up to 1 hour and were audio recorded, with written notes capturing nonverbal cues and context. We transcribed the audio recordings verbatim and analyzed them thematically using NVivo (QSR International, Melbourne, Australia) software. Thematic analysis followed a partially inductive approach: although the interview guide informed the initial structure, codes and subthemes emerged directly from participant narratives. After subclassifying the main themes, a sample of 100 randomly selected coded passages was independently reviewed by 2 research assistants. Discrepancies in interpretation were discussed and resolved collaboratively to enhance coding consistency and validity. To preserve confidentiality and adhere to ethical requirements, participant quotes were not linked to individual identifiers or care settings. As such, quotes are presented without attribution beyond their thematic grouping.

Ethical Considerations

Ethics approval was granted from the University of Ottawa's Research Ethics Board (S-06-18-428). Participation was voluntary, and participants provided written informed consent and were assured of confidentiality of the information collected in the study. To protect identities, pseudonyms were used in transcripts. No compensation was provided.

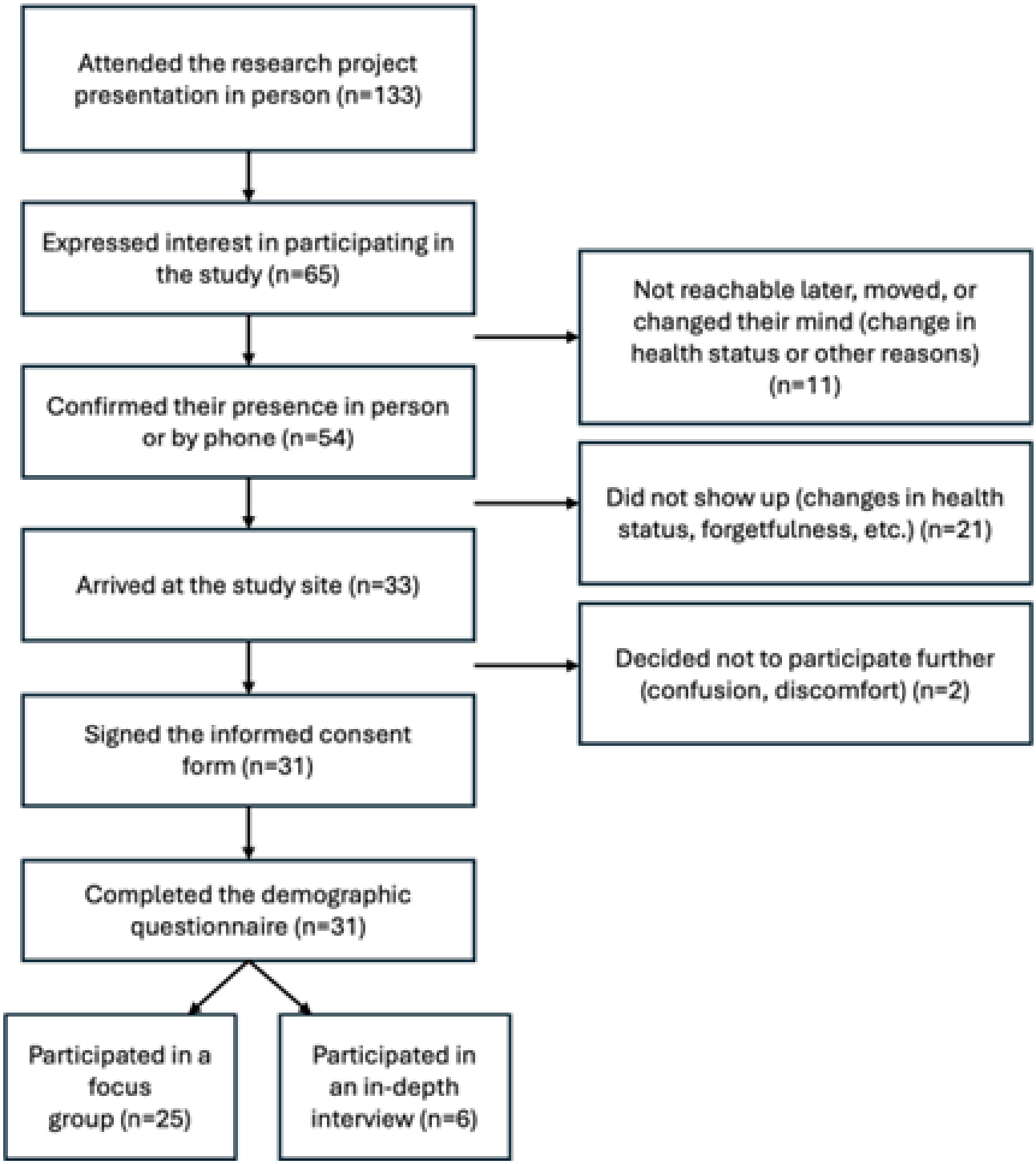
Results

Overview of Respondents

A total of 133 individuals attended the initial research presentation, with 65 expressing interest. Of these 65, 54 (85.7%) confirmed participation, but some later withdrew due to personal or health reasons or became unreachable. Ultimately, 31 participants completed the study (25 in focus groups and 6 in in-depth interviews; Figure 1).

XSL•FO
RenderX

Figure 1. Recruitment process: This flowchart illustrates the recruitment process for the study, detailing the number of participants at each stage. Of the 133 project presentation attendees, 65 expressed interest, 54 confirmed participation, and 33 arrived at the study site. Ultimately, 31 signed consent and completed the demographic questionnaire, with 25 joining a focus group and 6 participating in an in-depth interview. Drop-off reasons included inaccessibility, health changes, and discomfort.



The study included 31 participants from various care settings: SAFE, LTC, and ILA (Table 1). Of the 6 interviewees, 5 were from SAFE and 1 from LTC. Age varied by living environment, with LTC participants being the oldest. Gender representation was balanced, with 45.2% male and 54.8% female. A significant portion (42%) had university-level education. Most were widowed or single, lived alone (64.5%), and had chronic conditions, with 45.2% reporting comorbidities. Despite health issues, 67.8% rated their health as good or better.

XSL•FO
RenderX

Table . Participants' characteristics.

Characteristics	Focus groups (n=25)			Interviews (n=6)	Total (n=31)
	SAFE[a] unit (n=7)	Long-term care (n=7)	Independent living (n=11)		
Age, years, median (range)	78 (65-87)	94 (87-99)	89 (72-93)	78 (65-97)	87 (65-99)
Gender, n (%)					
Male	1 (14.3)	5 (71.4)	5 (45.5)	3 (50)	14 (45.2)
Female	6 (85.7)	2 (28.6)	6 (54.5)	3 (50)	17 (54.8)
Education, n (%)					
Nonuniversity education[b]	5 (71.4)	2 (28.6)	5 (45.5)	3 (50)	15 (48.4)
University-level education[c]	2 (28.6)	5 (71.4)	6 (54.5)	3 (50)	16 (51.6)
Marital status, n (%)					
Married/partnered	3 (42.9)	—[d]	3 (27.3)	2 (33.3)	8 (25.8)
Widowed	3 (42.9)	6 (85.7)	6 (54.5)	3 (50)	18 (58.1)
Single/divorced/separated	1 (14.3)	1 (14.3)	2 (18.2)	1 (16.7)	5 (16.1)
Living situation, n (%)					
Lives alone	4 (57.1)	7 (100)	6 (54.5)	3 (50)	20 (64.5)
Lives with partner	2 (28.6)	—	3 (27.3)	2 (33.3)	7 (22.6)
Other	1 (14.3)	—	2 (18.2)	1 (16.7)	4 (12.9)
Self-reported health, n (%)					
Good to excellent	4 (57.2)	4 (57.1)	9 (81.9)	4 (66.7)	21 (67.8)
Fair to poor	3 (42.8)	3 (42.9)	2 (18.1)	2 (33.3)	10 (32.2)
Chronic illness, n (%)					
None	—	—	3 (27.3)	—	3 (9.7)
Single condition	2 (28.6)	5 (71.4)	5 (45.5)	2 (33.3)	14 (45.2)
Comorbidities	5 (71.4)	2 (28.6)	3 (27.3)	4 (66.7)	14 (45.2)

[a]SAFE: subacute unit for frail elderly.

[b]Includes high school, trade, college, and other nonuniversity education.

[c]Includes participants with bachelor's or graduate degrees.

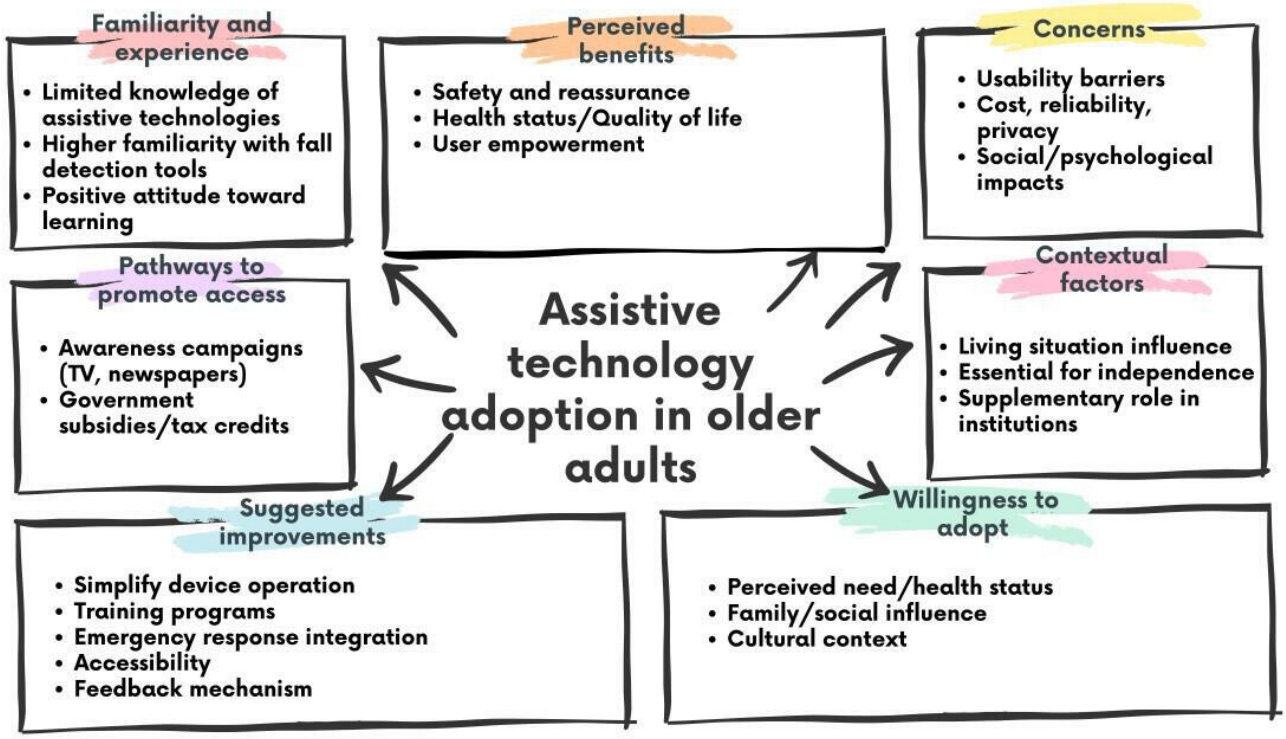
[d]—: not applicable.

Key Findings

We identified 7 overarching themes from the thematic analysis conducted on the data generated from the focus groups and interviews, which are summarized in Figure 2.

XSL•FO
RenderX

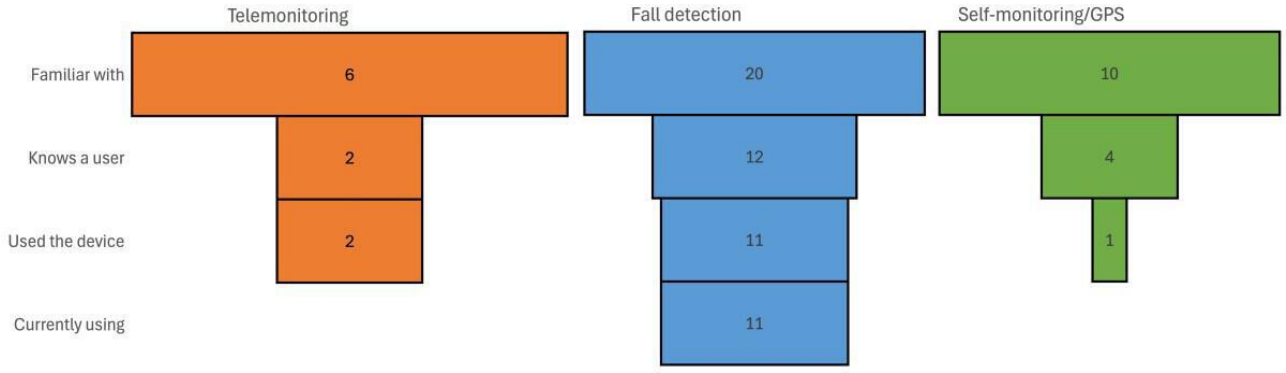
Figure 2. Key themes and subthemes: This diagram highlights factors influencing assistive technology adoption in older adults, including familiarity, perceived benefits, concerns, contextual factors, willingness to adopt, pathways to promote access, and suggested improvements.



Theme 1: Familiarity and Experience With Technology

Most participants had limited prior knowledge of ATs. While many of them followed the news through newspapers and television, they knew little about technologies that could support them. Some participants were eager to learn more, whereas others remained indifferent. As shown in Figure 3, familiarity was highest for fall detection (n=20), followed by self-monitoring/GPS (n=10), and lowest for telemonitoring (n=6). We observed a similar pattern for other measures of exposure, such as knowing a user or having used the device themselves. These differences in baseline familiarity appeared to influence openness and perceived relevance during group discussions. Despite limited experience with these investigated ATs, older adults expressed a positive attitude toward them, for example, “I’m glad to learn there are things like this happening” and “Knowing about these devices motivates us to continue living and maintaining our independence.”

Figure 3. Familiarity with and use of assistive technologies: This chart shows participants’ familiarity and engagement with telemonitoring, fall detection, and self-monitoring devices. Fall detection devices had the highest familiarity and usage, whereas self-monitoring devices had moderate awareness but low current usage. Telemonitoring had the lowest familiarity and minimal adoption.



Theme 2: Perceived Benefits

The participants identified 3 key benefits of ATs presented below.

Safety and Reassurance

Participants commonly perceived that all 3 ATs can reassure users and their families by enabling timely emergency interventions, although they had not personally used these technologies. They valued telemonitoring for its ability to detect abnormalities early and provide prompt feedback. Self-monitoring tools were praised for tracking users’ locations, helping families ensure the safety of loved ones who might become disoriented or lost. Fall detection technology was seen as lifesaving, with 1 participant stating, “Using a fall detector gives the family peace of mind, knowing that their loved one won’t be found on the floor hours later.”

XSL•FO
RenderX

Health Status and Quality of Life

ATs were seen as beneficial for older adults' health by improving chronic disease management (telemonitoring), enabling timely interventions (fall detection), and encouraging active lifestyles (self-monitoring). They also support independence, particularly for those living at home. As 1 participant noted, "This could extend the time they can live at home and do certain activities," instead of being institutionalized.

User Empowerment

Participants felt that telemonitoring's immediate feedback from health care professionals fostered a sense of control and informed decision-making. As 1 participant shared, "If she says in the morning, 'I feel this way,' and the [healthcare professional] responds within an hour, she will know what to do."

Theme 3: Concerns

Despite the perceived benefits, participants raised several key concerns regarding the use of ATs.

Usability

Participants highlighted physical, cognitive, and educational barriers that could limit ATs' adoption and effective use. As 1 respondent noted, "Telemonitoring can only be used by someone who can use it manually, but what if you can't use your hands?," raising concerns that some users may struggle with device operation.

Cost

Many participants saw high expenses as a barrier to widespread adoption and equitable access. As 1 respondent observed, "I see a huge divide between people who can afford this type of technologies and those who can't." This concern about cost reflects broader patterns of digital exclusion among older adults, as noted in recent reviews on health technology disparities [18,65]. Such perceptions may also compound age-related income constraints, underscoring the need for subsidized or publicly funded solutions.

Reliability and Response Time

Participants were concerned about the risk of incorrect or delayed information and the ability to act on it. One respondent questioned, "What happens if there's a problem? Will someone be alerted immediately? The response time matters a lot." The effectiveness of telemonitoring depends not only on the technology but also on the professionals overseeing it. Some were skeptical about whether professionals actively monitored the data, with 1 participant indicating, "It feels like they don't actively monitor the device. For me, it doesn't seem worthwhile," suggesting potential trust-related concerns.

Privacy and Data Security

Most participants were concerned about potential breaches of sensitive information or data being shared with third parties, such as insurers. One respondent also feared that ATs could increase vulnerability to theft, stating, "It could very easily enable someone with malicious intention to access the system and discover that you are not at home."

Social and Psychological Impacts

Participants expressed concerns about the psychological effects of ATs' use, including feelings of dependency and loss of dignity. Many saw these devices as symbols of weakness or reliance on others, potentially impacting self-esteem and personal pride. As one participant noted, "You feel like you don't want to be the person who has to wear that, and yet, in reality, you have to." Participants' emphasis on the psychological burden of using fall detectors echoes findings from Köttl et al [66], where older adults internalized negative stereotypes linked to assistive device use, framing them as "symbols of decline." Our findings further confirm that emotional readiness is just as critical as functional need in ATs' adoption decisions.

Acceptance and Familiarity

The novelty of ATs contributed to hesitation and skepticism among participants. Some were cautious about adopting unfamiliar devices, even when aware of their benefits. As 1 participant remarked, "People might not trust new devices. Anything new is hard to trust, isn't it?"

Theme 4: Suggested Improvements

Participants recommended designing *simpler, user-friendly* devices, as many older adults may struggle with complex technology. They stressed the importance of *training and orientation*, with 1 participant noting, "You can't just hand someone something and leave them to figure it out. There needs to be training or orientation." To enhance emergency response, respondents suggested integrating location-tracking with fall detection devices to ensure faster assistance. As 1 participant proposed, "If they press an alarm, we should be able to locate them, no matter where they are."

Accessibility was another key concern, with many advocating for features that accommodate physical and cognitive limitations, "like using large buttons or voice commands". Reminder systems (for forgetfulness), expanded coverage, and wearable options were also highlighted. One participant remarked, "If they want to change the world, these systems should be everywhere, not just in nursing homes." Finally, participants valued feedback features to confirm help was on the way. As 1 participant stated, "...if it could notify you that your message was sent and someone is coming, that would be reassuring."

Theme 5: Contextual Factors Influencing Use

Perceived usefulness varied by living situation. For older adults living independently, all 3 ATs were viewed as essential for safety, health, and independence. As 1 participant explained, "It's good if you're not in a critical state requiring hospitalization, but you have medical conditions that need monitoring, which can be done remotely."

In institutional settings, these technologies were seen as supplementary to existing care. With professional staff available around the clock, some participants saw telemonitoring and fall detection as redundant. However, others still found value in them, with 1 participant noting, "Even in an institution, I could fall in my room, and staff might not immediately know"

XSL•FO
RenderX

Theme 6: Pathways to Promote Awareness/Access

Awareness campaigns using accessible media were suggested: "Television and newspapers would be good ways to inform people about these options." Government support (ie, subsidies, need-based assistance, and tax credits) to offset costs was discussed as critical. While such support does not currently exist for ATs, participants referenced similar programs in Ontario that subsidize other physician-prescribed medical equipment, such as continuous positive airway pressure machines.

Theme 7: Willingness to Adopt/Use ATs

Many participants initially expressed enthusiasm toward ATs, "This is fantastic!," but this did not necessarily translate into a willingness to adopt them: "My diabetes isn't severe enough right now!" or "You can lead a horse to water, but you can't make it drink." Their willingness to use these technologies depended on several key factors.

Perceived Need and Health Status

Many participants indicated that their willingness to use ATs would increase if their health declined. As 1 participant noted, "I think you have to experience a fall to scare you enough to jump into it faster." Others recognized the potential benefits but felt no immediate need due to stable health conditions.

Family and Social Influence

Family encouragement played a crucial role in AT adoption. Some participants described how their families insisted on using these technologies for peace of mind, even when they were hesitant. As 1 participant shared, "My kids keep pushing me to try it [fall detection bracelet] because it makes them feel better about my safety."

Health Care Recommendations

Guidance from health care professionals significantly influenced ATs' adoption. As 1 participant stated, "If my doctor recommended it, I'd definitely consider it," highlighting the trust placed in medical advice.

Ease of Use and Design

Participants stressed that ease of use was crucial in their decision to adopt ATs. Lightweight, user-friendly, and accessible devices were more appealing, whereas complicated or impractical designs discouraged interest. In addition, limited awareness of how these technologies functioned created barriers to adoption.

Emotional and Psychological Concerns

Participants expressed mixed feelings about being monitored, with some finding it intrusive, whereas others appreciated the security it provided. As 1 participant stated, "It's a bit unsettling to know you're always being watched, but it's also comforting to know help is there if you need it."

Cultural and Situational Context

Social norms and perceived popularity influenced interest in AT adoption. As 1 participant noted, "If other people have it (telemonitoring) and I don't, I feel disadvantaged." ATs were more appealing to those living alone or in rural areas, whereas individuals in assisted living homes or with regular care access showed less interest.

Discussion

Principal Findings

This study explored older adults' perceptions of 3 ATs, namely, telemonitoring systems, fall detection devices, and self-tracking tools. By combining focus groups and in-depth interviews, we gathered nuanced perspectives from individuals with varied health statuses and living environments. This design helped uncover both attitudinal and contextual factors shaping AT acceptance, contributing to an enriched understanding of older adults' engagement with technology.

Perceived Benefits and the Role of Identity

Participants consistently acknowledged the potential of ATs to enhance safety, independence, and chronic disease management. These findings align with the literature, suggesting that older adults are not inherently resistant to technology, but that they appraise its value through the lens of personal relevance and life context [67-71]. Importantly, this study highlights a tension between recognition of benefits and personal adoption—participants often deemed ATs useful "for others," but not necessary for themselves.

This ambivalence may reflect deeper psychological dynamics. Drawing on Socioemotional Selectivity Theory [72], many participants appeared to prioritize emotional satisfaction and autonomy in the present over future-oriented planning. Similarly, internalized ageism, the absorption of negative stereotypes about aging, was evident in language that framed ATs as symbols of frailty or decline. Participants' reluctance to adopt fall detection devices, in particular, often stemmed from concerns about stigma or diminished self-image. These reactions echo prior findings that older adults resist technologies perceived to signal vulnerability [66], underscoring the importance of identity preservation in technology acceptance.

Barriers to Adoption: Beyond Attitudes

Despite generally favorable attitudes, several barriers limited participants' readiness to adopt ATs. Cost emerged as a prominent concern, even in Canada's publicly funded health care context. Consistent with UTAUT2 [73-76], where "price value" influences behavioral intention, participants were highly sensitive to the cost–benefit tradeoff. For many, high device costs were incompatible with fixed or limited incomes. This highlights the need for policy interventions, such as subsidies or expanded Assistive Devices Programs, to enhance financial accessibility.

Usability was another critical barrier. Participants emphasized the importance of accessible, intuitive designs, particularly given potential sensory, physical, or cognitive impairments. These findings reinforce prior research linking effort expectancy (ease of use) to adoption intent [77,78]. Participants suggested adaptive features (eg, voice commands and large buttons) and emphasized the value of hands-on training and orientation sessions. These recommendations support calls for participatory design approaches that incorporate older adults' perspectives into AT development [7,78,79].

XSL•FO
RenderX

Although privacy concerns were noted, they were generally outweighed by perceived safety and health benefits. Still, these concerns merit further investigation. Recent extensions of UTAUT have treated privacy either as an independent predictor [80] or a moderator [81], but its role in older adult populations remains underexplored. Future research should test whether privacy functions differently in this demographic, particularly with respect to self-tracking and location-based technologies.

Willingness Versus Readiness: A Critical Distinction

This study also underscores a conceptual distinction between willingness (attitudinal openness) and readiness (practical ability to act). While participants expressed interest in ATs, this rarely translated into immediate adoption. Readiness appeared contingent on several enabling conditions: digital literacy, health triggers (eg, experiencing a fall), social support, and recommendations from trusted others. This distinction is rarely emphasized in AT literature but is essential for tailoring interventions. Addressing readiness gaps through skills training, caregiver involvement, and experiential exposure to ATs may be just as important as shifting attitudes.

The Influence of Experience and Social Context

Experience with ATs emerged as a powerful enabler. Participants with prior exposure demonstrated higher confidence, clearer expectations, and greater openness to ongoing use. This finding supports evidence from diffusion studies that familiarity reduces perceived complexity and increases trialability [82-84]. Conversely, nonusers often failed to recognize personal relevance, even when their health status suggested potential benefit.

Here, the diffusion of innovation framework [85] offers a useful lens. Health professionals, especially in institutional settings, can serve as gatekeepers, opinion leaders, and change agents. For example, physicians and occupational therapists can assess individual needs, demystify technologies, and provide credible endorsements, thereby influencing both perceived usefulness and social norms [86-88]. Our findings reinforce the importance of leveraging these roles, especially within LTC contexts where allied health professionals routinely assess functional abilities. Promoting ATs through routine care conferences, orientation sessions, and interdisciplinary planning can normalize their use and personalize recommendations.

Implications for Design, Policy, and Practice

Bridging the gap between positive attitudes and actual adoption of ATs requires multipronged strategies that target both systemic and individual-level barriers. Participants emphasized the importance of adaptive, inclusive designs that accommodate cognitive, sensory, and physical limitations. Features such as voice activation, tactile interfaces, and customizable visual displays (eg, font size and button shape) were seen as particularly valuable for enhancing usability. In line with earlier research [7,39,89,90], these design features can help ensure technologies are accessible to a wider range of older adults.

Participants also stressed the importance of training and support. Many felt that simply distributing a device without orientation would be ineffective. Suggestions included hands-on demonstrations, step-by-step tutorials, and reinforcement opportunities through follow-up sessions. For example, in LTC settings, technology vendors could be invited to host demonstration days with staff, residents, and caregivers. For community-dwelling older adults, device trials combined with user support, potentially involving caregivers or trained volunteers, could enhance experiential learning and encourage integration into daily life.

Education and outreach are equally important. Participants advocated for awareness campaigns using accessible media, such as television and newspapers, to reach broader audiences. These campaigns should emphasize the empowering aspects of ATs, helping reframe them as tools for maintaining independence rather than markers of decline. This is particularly critical in combating internalized ageism and reshaping social narratives around aging and technology.

Health care providers, especially physicians, nurses, and allied health professionals, play a pivotal role in recommending and normalizing AT use. Their trusted relationships with older adults position them as effective intermediaries for personalized advice and confidence-building. Strengthening provider involvement through continuing education, clinical guidelines, and structured patient discussions can help increase uptake. In LTC, allied health professionals such as occupational and physiotherapists are well-positioned to incorporate AT assessments into routine care planning. Opportunities for integration include resident orientation, care conferences, and interdisciplinary team meetings. These touchpoints offer natural opportunities to introduce ATs aligned with residents' values, needs, and preferences [91].

Finally, policy interventions are also essential. Public reimbursement schemes, such as Ontario's Assistive Devices Program, could be expanded to include a broader array of ATs and cover maintenance or replacement costs. Such support would help address cost-related concerns, one of the most frequently cited barriers to adoption, and ensure equitable access. Policymakers may also consider embedding AT content into the curricula of health care training programs, ensuring that future providers across disciplines are equipped to assess, recommend, and support AT use among older adults.

Limitations and Avenues for Future Research

This study has limitations that must be acknowledged. First, given its exploratory nature, we used an inductive approach through focus groups and interviews to capture older adults' experiences without restricting responses to predefined frameworks (eg, TAM, UTAUT). This approach enabled novel insights to emerge, although future studies could build on our findings using established models to assess generalizability and validate results in other contexts.

Second, recruitment was limited to individuals with sufficient cognitive abilities to provide informed consent, excluding those with significant impairments. Physical and sensory limitations, such as hearing or vision issues, sometimes disrupted group discussions and required clarification. In some cases, small group sizes—due to cancellations or misunderstandings—reduced discussion depth. However,

XSL•FO
RenderX

in-depth interviews helped address these gaps by capturing additional perspectives. Time constraints also limited exploration of some topics during focus groups, although these were partially mitigated by the interviews.

Third, participants viewed brand-neutral video demonstrations to facilitate discussion, but these may have shaped perceptions based on how the technologies were presented. While the videos offered general overviews, they did not depict specific devices, limiting discussion of their particularities. In addition, themes such as privacy, cost, and emergency response were included in the interview guide, potentially influencing responses. For confidentiality, quotes were not linked to participants' identities or care settings, preventing comparative analysis across environments.

Despite these limitations, the study addressed several gaps in the extant literature. We explored 3 types of ATs relevant to diverse care needs and included 31 participants from LTC, SAFE, and independent living settings, offering broader insights than prior studies with smaller or more homogeneous samples. Unlike studies relying on a single method [83,92-98], our combined use of focus groups and interviews allowed for a richer, more nuanced understanding. Still, we were unable to include participants with cognitive impairment. Future research should address this by involving informal caregivers, who play a key role in decision-making and technology use. Additional research is also needed to explore AT adoption among individuals with greater physical or cognitive limitations. Quantitative studies examining perceived need, prior experience, and external influences (eg, clinician recommendations) could inform strategies to support wider adoption.

Conclusions

Promoting the use of health ATs among older adults requires a focus on improving accessibility, usability, and awareness. This research adds to the growing body of literature on older adults' perspectives on health ATs by offering insights into their attitudes, barriers, and facilitators of adoption. While older adults recognize the benefits of telemonitoring systems, fall detection technology, and self-tracking devices, adoption remains limited and is influenced by factors such as perceived personal need, prior experience, and accessibility. Future efforts should focus on creating adaptive designs, offering experiential opportunities, and implementing accessible pricing models to enhance the adoption of health ATs and maximize their potential to support aging populations.

Acknowledgments

The authors would like to thank all participants for their contributions to this project and the team at Perley Health and Danielle Sinden for facilitating access to the participants and providing the setting for the focus groups.

Funding

This research was supported by a research grant from the Social Sciences and Humanities Research Council (435-2017-1399).

Data Availability

The data gathered during this study are not publicly available and cannot be shared due to confidentiality and original ethics approval restrictions.

Conflicts of Interest

None declared.

References

1. Marchildon G. Canada: health system review. Health Syst Transit 2013;15(1):1-179. [Medline: 23628429]
2. Peek STM, Wouters EJM, Luijkx KG, Vrijhoef HJM. What it takes to successfully implement technology for aging in place: focus groups with stakeholders. J Med Internet Res 2016 May 3;18(5):e98. [doi: 10.2196/jmir.5253] [Medline: 27143097]
3. Yusif S, Soar J, Hafeez-Baig A. Older people, assistive technologies, and the barriers to adoption: a systematic review. Int J Med Inform 2016 Oct;94:112-116. [doi: 10.1016/j.ijmedinf.2016.07.004] [Medline: 27573318]
4. Botsis T, Hartvigsen G. Current status and future perspectives in telecare for elderly people suffering from chronic diseases. J Telemed Telecare 2008;14(4):195-203. [doi: 10.1258/jtt.2008.070905] [Medline: 18534954]
5. Khosravi P, Ghapanchi AH. Investigating the effectiveness of technologies applied to assist seniors: a systematic literature review. Int J Med Inform 2016 Jan;85(1):17-26. [doi: 10.1016/j.ijmedinf.2015.05.014] [Medline: 26216463]
6. Narasimha S, Madathil KC, Agnisarman S, et al. Designing telemedicine systems for geriatric patients: a review of the usability studies. Telemed J E Health 2017 Jun;23(6):459-472. [doi: 10.1089/tmj.2016.0178] [Medline: 27875667]
7. Claes V, Devriendt E, Tournoy J, Milisen K. Attitudes and perceptions of adults of 60 years and older towards in-home monitoring of the activities of daily living with contactless sensors: an explorative study. Int J Nurs Stud 2015 Jan;52(1):134-148. [doi: 10.1016/j.ijnurstu.2014.05.010] [Medline: 24951084]

XSL•FO
RenderX

8. Hawley-Hague H, Boulton E, Hall A, Pfeiffer K, Todd C. Older adults' perceptions of technologies aimed at falls prevention, detection or monitoring: a systematic review. Int J Med Inform 2014 Jun;83(6):416-426. [doi: 10.1016/j.ijmedinf.2014.03.002] [Medline: 24798946]
9. Soar J, Seo Y. Health and aged care enabled by information technology. Ann N Y Acad Sci 2007 Oct;1114(1):154-161. [doi: 10.1196/annals.1396.040] [Medline: 17986580]
10. Jaana M, Paré G, Sicotte C. Home telemonitoring for respiratory conditions: a systematic review. Am J Manag Care 2009 May;15(5):313-320. [Medline: 19435399]
11. Jaana M, Paré G. Home telemonitoring of patients with diabetes: a systematic assessment of observed effects. J Eval Clin Pract 2007 Apr;13(2):242-253. [doi: 10.1111/j.1365-2753.2006.00686.x] [Medline: 17378871]
12. Paré G, Jaana M, Sicotte C. Systematic review of home telemonitoring for chronic diseases: the evidence base. J Am Med Inform Assoc 2007;14(3):269-277. [doi: 10.1197/jamia.M2270] [Medline: 17329725]
13. Kitsiou S, Paré G, Jaana M. Effects of home telemonitoring interventions on patients with chronic heart failure: an overview of systematic reviews. J Med Internet Res 2015 Mar 12;17(3):e63. [doi: 10.2196/jmir.4174] [Medline: 25768664]
14. Jaana M, Tamim H, Sherrard H, Pare G. Telemonitoring for seniors with chronic heart failure: patient self-care, empowerment, and adoption factors. 2017 Jan 4 Presented at: Hawaii International Conference on System Sciences; Jan 4-7, 2017. [doi: 10.24251/HICSS.2017.449]
15. Jaana M, Sherrard H. Rural-urban comparison of telehome monitoring for patients with chronic heart failure. Telemed J E Health 2019 Feb;25(2):101-108. [doi: 10.1089/tmj.2017.0303] [Medline: 29847242]
16. Diduszyn J, Hofmann MT, Naglak M, Smith DG. Use of a wireless nurse alert fall monitor to prevent inpatient falls. J Clin Outcomes Manag 2008;15(6):293-296 [FREE Full text]
17. Ledger D, McCaffrey D. Inside wearables part 1: how behavior change unlocks long-term engagement. Medium. 2020 Apr 25. URL: https://medium.com/@endeavourprtnrs/inside-wearable-how-the-science-of-human-behavior-change-offers-the-secret-to-long-term-engagement-a15b3c7d4cf3 [accessed 2025-01-27]
18. Vassli LT, Farshchian BA. Acceptance of health-related ICT among elderly people living in the community: a systematic review of qualitative evidence. Int J Hum Comput Interact 2018 Feb;34(2):99-116. [doi: 10.1080/10447318.2017.1328024]
19. Liu L, Stroulia E, Nikolaidis I, Miguel-Cruz A, Rios Rincon A. Smart homes and home health monitoring technologies for older adults: a systematic review. Int J Med Inform 2016 Jul;91:44-59. [doi: 10.1016/j.ijmedinf.2016.04.007] [Medline: 27185508]
20. Peine A, Rollwagen I, Neven L. The rise of the "innosumer"—rethinking older technology users. Technol Forecast Soc Change 2014 Feb;82:199-214. [doi: 10.1016/j.techfore.2013.06.013]
21. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425. [doi: 10.2307/30036540]
22. Kavandi H, Jaana M. Factors that affect health information technology adoption by seniors: a systematic review. Health Soc Care Community 2020 Nov;28(6):1827-1842. [doi: 10.1111/hsc.13011] [Medline: 32378769]
23. Alexander GL, Wakefield BJ, Rantz M, et al. Passive sensor technology interface to assess elder activity in independent living. Nurs Res 2011;60(5):318-325. [doi: 10.1097/NNR.0b013e318225f3e1] [Medline: 21873920]
24. Alwan M, Dalal S, Mack D, et al. Impact of monitoring technology in assisted living: outcome pilot. IEEE Trans Inf Technol Biomed 2006 Jan;10(1):192-198. [doi: 10.1109/titb.2005.855552] [Medline: 16445264]
25. Chen KY, Harniss M, Patel S, Johnson K. Implementing technology-based embedded assessment in the home and community life of individuals aging with disabilities: a participatory research and development study. Disabil Rehabil Assist Technol 2014 Mar;9(2):112-120. [doi: 10.3109/17483107.2013.805824] [Medline: 23802128]
26. Demiris G, Hensel BK, Skubic M, Rantz M. Senior residents' perceived need of and preferences for "smart home" sensor technologies. Int J Technol Assess Health Care 2008;24(1):120-124. [doi: 10.1017/S0266462307080154] [Medline: 18218177]
27. Demiris G, Oliver DP, Giger J, Skubic M, Rantz M. Older adults' privacy considerations for vision based recognition methods of eldercare applications. THC 2009;17(1):41-48. [doi: 10.3233/THC-2009-0530]
28. Goodall K, Ward P, Newman L. Use of information and communication technology to provide health information: what do older migrants know, and what do they need to know? Qual Prim Care 2010;18(1):27-32. [Medline: 20359410]
29. Kutz D, Shankar K, Connelly K. Making sense of mobile- and web-based wellness information technology: cross-generational study. J Med Internet Res 2013 May 14;15(6):e83. [doi: 10.2196/jmir.2124] [Medline: 23900035]
30. Whitty JA, Ratcliffe J, Chen G, Scuffham PA. Australian public preferences for the funding of new health technologies: a comparison of discrete choice and profile case best-worst scaling methods. Med Decis Making 2014 Jul;34(5):638-654. [doi: 10.1177/0272989X14526640] [Medline: 24713695]
31. Wu YH, Damnée S, Kerhervé H, Ware C, Rigaud AS. Bridging the digital divide in older adults: a study from an initiative to inform older adults about new technologies. Clin Interv Aging 2015;10:193-200. [doi: 10.2147/CIA.S72399] [Medline: 25624752]

XSL•FO
RenderX

32. Canzanello VJ, Jensen PL, Schwartz LL, Worra JB, Klein LK. Improved blood pressure control with a physician-nurse team and home blood pressure measurement. Mayo Clin Proc 2005 Jan;80(1):31-36. [doi: 10.1016/S0025-6196(11)62954-6] [Medline: 15667026]
33. Nguyen CP, Pham TNY, Eric C. Toward a sound analysis system for telemedicine. In: Wang L, Jin Y, editors. Fuzzy Systems and Knowledge Discovery 2005, Vol. 3614:352-361. [doi: 10.1007/11540007_44]
34. Damodaran L, Olphert W, Phipps S. Keeping silver surfers on the crest of a wave – older people's ICT learning and support needs. Work Older People 2013 Mar 8;17(1):32-36. [doi: 10.1108/13663661311312566]
35. Demiris G, Rantz M, Aud M, et al. Older adults' attitudes towards and perceptions of "smart home" technologies: a pilot study. Med Inform Internet Med 2004 Jun;29(2):87-94. [doi: 10.1080/14639230410001684387] [Medline: 15370989]
36. Lu JF, Chi MJ, Chen CM. Advocacy of home telehealth care among consumers with chronic conditions. J Clin Nurs 2014 Mar;23(5-6):811-819. [doi: 10.1111/jocn.12156] [Medline: 23796027]
37. Li M, McPhillips MV, Szanton SL, Wenzel J, Li J. Electronic wearable device use for physical activity in older adults: a qualitative study. In: Xie B, editor. Work Aging Retire 2024 Jan;10(1):25-37. [doi: 10.1093/workar/waac023] [Medline: 38196825]
38. Bartlett YK, Webb TL, Hawley MS. Using persuasive technology to increase physical activity in people with chronic obstructive pulmonary disease by encouraging regular walking: a mixed-methods study exploring opinions and preferences. J Med Internet Res 2017 Apr 20;19(4):e124. [doi: 10.2196/jmir.6616] [Medline: 28428155]
39. Middlemass JB, Vos J, Siriwardena AN. Perceptions on use of home telemonitoring in patients with long term conditions - concordance with the Health Information Technology Acceptance Model: a qualitative collective case study. BMC Med Inform Decis Mak 2017 Jun 26;17(1):89. [doi: 10.1186/s12911-017-0486-5] [Medline: 28651588]
40. Simmich J, Mandrusiak A, Russell T, Smith S, Hartley N. Perspectives of older adults with chronic disease on the use of wearable technology and video games for physical activity. Digit Health 2021;7:20552076211019900. [doi: 10.1177/20552076211019900] [Medline: 34104468]
41. Harris MT, Rogers WA. Developing a healthcare technology acceptance model (H-TAM) for older adults with hypertension. Ageing Soc 2023 Apr;43(4):814-834. [doi: 10.1017/s0144686x21001069] [Medline: 37007645]
42. Trief PM, Sandberg J, Izquierdo R, et al. Diabetes management assisted by telemedicine: patient perspectives. Telemed J E Health 2008 Sep;14(7):647-655. [doi: 10.1089/tmj.2007.0107] [Medline: 18817493]
43. Schroeder T, Dodds L, Georgiou A, Gewald H, Siette J. Older adults and new technology: mapping review of the factors associated with older adults' intention to adopt digital technologies. JMIR Aging 2023 May 16;6:e44564. [doi: 10.2196/44564] [Medline: 37191976]
44. Fakhkhari H, Bounabat B, Kassou I. Digital health taxonomy. 2023 May 24 Presented at: NISS 2023; May 24-26, 2023; Larache Morocco p. 1-9. [doi: 10.1145/3607720.3607780]
45. Scheibner J, Sleigh J, Ienca M, Vayena E. Benefits, challenges, and contributors to success for national eHealth systems implementation: a scoping review. J Am Med Inform Assoc 2021 Aug 13;28(9):2039-2049. [doi: 10.1093/jamia/ocab096] [Medline: 34151990]
46. Ali MA, Alam K, Taylor B, Ashraf M. Examining the determinants of eHealth usage among elderly people with disability: the moderating role of behavioural aspects. Int J Med Inform 2021 May;149:104411. [doi: 10.1016/j.ijmedinf.2021.104411] [Medline: 33618191]
47. Villumsen S, Adler-Milstein J, Nøhr C. National monitoring and evaluation of eHealth: a scoping review. JAMIA Open 2020 Apr;3(1):132-140. [doi: 10.1093/jamiaopen/ooz071] [Medline: 32607495]
48. Eysenbach G. What is e-health? J Med Internet Res 2001;3(2):E20. [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
49. World Health Organization. Assistive technology. 2024. URL: https://www.who.int/news-room/fact-sheets/detail/assistive-technology [accessed 2025-05-21]
50. Sweeting A, Warncken KA, Patel M. The role of assistive technology in enabling older adults to achieve independent living: past and future. J Med Internet Res 2024 Jul 30;26:e58846. [doi: 10.2196/58846] [Medline: 39079115]
51. Fotteler ML, Mühlbauer V, Brefka S, et al. The effectiveness of assistive technologies for older adults and the influence of frailty: systematic literature review of randomized controlled trials. JMIR Aging 2022 Apr 4;5(2):e31916. [doi: 10.2196/31916] [Medline: 35377322]
52. Dautzenberg L, Beglinger S, Tsokani S, et al. Interventions for preventing falls and fall‐related fractures in community‐dwelling older adults: a systematic review and network meta‐analysis. J Am Geriatr Soc 2021 Oct;69(10):2973-2984. [doi: 10.1111/jgs.17375] [Medline: 34318929]
53. Garçon L, Khasnabis C, Walker L, et al. Medical and assistive health technology: meeting the needs of aging populations: Table 1. GERONT 2016 Apr;56(Suppl 2):S293-S302. [doi: 10.1093/geront/gnw005]
54. Macdonald M, Gallant A, Weeks L, et al. Long-term care home residents' experiences with socially assistive technologies and the effectiveness of these technologies: a mixed methods systematic review. JBI Evidence Synthesis 2024;22(8):1410-1459. [doi: 10.11124/JBIES-23-00021]
55. Shin HR, Um SR, Yoon HJ, et al. Comprehensive senior technology acceptance model of daily living assistive technology for older adults with frailty: cross-sectional study. J Med Internet Res 2023 Apr 10;25:e41935. [doi: 10.2196/41935] [Medline: 37036760]

XSL•FO
RenderX

56. Yang HJ, Lee JH, Lee W. Factors influencing health care technology acceptance in older adults based on the technology acceptance model and the unified theory of acceptance and use of technology: meta-analysis. J Med Internet Res 2025 Mar 28;27(1):e65269. [doi: 10.2196/65269] [Medline: 40153796]
57. Felber NA, Lipworth W, Tian YJA, Roulet Schwab D, Wangmo T. Informing existing technology acceptance models: a qualitative study with older persons and caregivers. Eur J Ageing 2024 Mar 29;21(1):12. [doi: 10.1007/s10433-024-00801-5] [Medline: 38551677]
58. Moore K, O'Shea E, Kenny L, et al. Older adults' experiences with using wearable devices: qualitative systematic review and meta-synthesis. JMIR Mhealth Uhealth 2021 Jun 3;9(6):e23832. [doi: 10.2196/23832] [Medline: 34081020]
59. Morgan DL, Krueger RA. The Focus Group Kit: Volumes 1-6: SAGE Publications; 1997. URL: https://books.google.ca/books?id=PHzQvwEACAAJ [accessed 2025-10-30]
60. Ritchie J, Lewis J, editors. Qualitative Research Practice: A Guide for Social Science Students and Researchers: SAGE; 2011. URL: https://mthoyibi.wordpress.com/wp-content/uploads/2011/10/qualitative-research-practice_a-guide-for-social-science-students-and-researchers_jane-ritchie-and-jane-lewis-eds_20031.pdf [accessed 2025-10-30]
61. Morris JN, Howard EP, Steel K, et al. Updating the cognitive performance scale. J Geriatr Psychiatry Neurol 2016 Jan;29(1):47-55. [doi: 10.1177/0891988715598231] [Medline: 26251111]
62. Serrano LP, Maita KC, Avila FR, et al. Benefits and challenges of remote patient monitoring as perceived by health care practitioners: a systematic review. Perm J 2023 Dec 15;27(4):100-111. [doi: 10.7812/TPP/23.022] [Medline: 37735970]
63. Chaudhuri S, Thompson H, Demiris G. Fall detection devices and their use with older adults: a systematic review. J Geriatr Phys Ther 2014;37(4):178-196. [doi: 10.1519/JPT.0b013e3182abe779] [Medline: 24406708]
64. Doyle M, Nwofe ES, Rooke C, Seelam K, Porter J, Bishop D. Implementing global positioning system trackers for people with dementia who are at risk of wandering. Dementia (London) 2024 Aug;23(6):964-980. [doi: 10.1177/14713012241248556] [Medline: 38646927]
65. Peek STM, Wouters EJM, van Hoof J, Luijkx KG, Boeije HR, Vrijhoef HJM. Factors influencing acceptance of technology for aging in place: a systematic review. Int J Med Inform 2014 Apr;83(4):235-248. [doi: 10.1016/j.ijmedinf.2014.01.004] [Medline: 24529817]
66. Köttl H, Gallistl V, Rohner R, Ayalon L. "But at the age of 85? Forget it!": internalized ageism, a barrier to technology use. J Aging Stud 2021 Dec;59:100971. [doi: 10.1016/j.jaging.2021.100971] [Medline: 34794716]
67. Heinz M, Martin P, Margrett JA, et al. Perceptions of technology among older adults. J Gerontol Nurs 2013 Jan;39(1):42-51. [doi: 10.3928/00989134-20121204-04] [Medline: 23244061]
68. Mann WC, Marchant T, Tomita M, Fraas L, Stanton K. Elder acceptance of health monitoring devices in the home. Care Manag J 2002 Jan;3(2):91-98. [doi: 10.1891/1521-0987.3.2.91]
69. Gellis ZD, Kenaley B, McGinty J, Bardelli E, Davitt J, Ten Have T. Outcomes of a telehealth intervention for homebound older adults with heart or chronic respiratory failure: a randomized controlled trial. Gerontologist 2012 Aug;52(4):541-552. [doi: 10.1093/geront/gnr134] [Medline: 22241810]
70. Saborowski M, Kollak I. "How do you care for technology?" – Care professionals' experiences with assistive technology in care of the elderly. Technol Forecast Soc Change 2015 Apr;93:133-140. [doi: 10.1016/j.techfore.2014.05.006]
71. Timon CM, Heffernan E, Kilcullen S, et al. Developing independent living support for older adults using internet of things and AI-based systems: co-design study. JMIR Aging 2024 Oct 24;7(1):e54210. [doi: 10.2196/54210] [Medline: 39447196]
72. Carstensen LL, Reynolds ME. Age differences in preferences through the lens of socioemotional selectivity theory. The Journal of the Economics of Ageing 2023 Feb;24:100440. [doi: 10.1016/j.jeoa.2022.100440]
73. Cimperman M, Brenčič MM, Trkman P, Stanonik MDL. Older adults' perceptions of home telehealth services. Telemed J E Health 2013 Oct;19(10):786-790. [doi: 10.1089/tmj.2012.0272] [Medline: 23931702]
74. Cocosila M, Turel O. Consumer risk perceptions in mobile health services adoption: do they matter? ETSN 2022;11(2):67-84. [doi: 10.4236/etsn.2022.112005]
75. Venkatesh V, Thong JYL, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Q 2012;36(1):157. [doi: 10.2307/41410412]
76. Wu C, Lim GG. Investigating older adults users' willingness to adopt wearable devices by integrating the technology acceptance model (UTAUT2) and the Technology Readiness Index theory. Front Public Health 2024;12:1449594. [doi: 10.3389/fpubh.2024.1449594] [Medline: 39421816]
77. Government of Ontario. Assistive Devices Program. 2016. URL: http://www.ontario.ca/page/assistive-devices-program [accessed 2025-05-23]
78. MacNeil M, Hirslund E, Baiocco-Romano L, Kuspinar A, Stolee P. A scoping review of the use of intelligent assistive technologies in rehabilitation practice with older adults. Disabil Rehabil Assist Technol 2024 Jul;19(5):1817-1848. [doi: 10.1080/17483107.2023.2239277] [Medline: 37498115]
79. Kang HG, Mahoney DF, Hoenig H, et al. In situ monitoring of health in older adults: technologies and issues. J Am Geriatr Soc 2010 Aug;58(8):1579-1586. [doi: 10.1111/j.1532-5415.2010.02959.x] [Medline: 20646105]

XSL•FO
RenderX

80. Reith R, Buck C, Eymann T, Lis B. Integrating privacy concerns into the unified theory of acceptance and use of technology to explain the adoption of fitness trackers. Int J Innovation Technol Management 2020 Nov 1;17(7). [doi: 10.1142/S0219877020500492]
81. Singh S. The moderating role of privacy concerns on intention to use smart wearable technologies: an integrated model combining UTAUT2 theoretical framework and privacy dimensions. Journal of Global Marketing 2023 Mar 15;36(2):93-111. [doi: 10.1080/08911762.2022.2141167]
82. Steele R, Lo A, Secombe C, Wong YK. Elderly persons' perception and acceptance of using wireless sensor networks to assist healthcare. Int J Med Inform 2009 Dec;78(12):788-801. [doi: 10.1016/j.ijmedinf.2009.08.001] [Medline: 19717335]
83. Dermody G, Fritz R, Glass C, Dunham M, Whitehead L. Factors influencing community-dwelling older adults' readiness to adopt smart home technology: A qualitative exploratory study. J Adv Nurs 2021 Dec;77(12):4847-4861. [doi: 10.1111/jan.14996] [Medline: 34477222]
84. Ma Q, Chan AHS, Teh PL. Insights into older adults' technology acceptance through meta-analysis. Int J Hum Comput Interact 2021 Jul 3;37(11):1049-1062. [doi: 10.1080/10447318.2020.1865005]
85. Rogers EM. Diffusion of Innovations, 5th edition: New York, NY: Free Press; 2003. URL: https://teddykw2.wordpress.com/wp-content/uploads/2012/07/everett-m-rogers-diffusion-of-innovations.pdf [accessed 2025-10-30]
86. Warty RR, Smith V, Salih M, Fox D, Mcarthur SL, Mol BW. Barriers to the diffusion of medical technologies within healthcare: a systematic review. IEEE Access 2021;9:139043-139058. [doi: 10.1109/ACCESS.2021.3118554]
87. Griesser A, Mzoughi M, Bidmon S, Cherif E. How do opt-in versus opt-out settings nudge patients toward electronic health record adoption? An exploratory study of facilitators and barriers in Austria and France. BMC Health Serv Res 2024;24(1):439. [doi: 10.1186/s12913-024-10929-w]
88. McGrath C, Ellis M, Harney-Levine S, et al. Investigating the enabling factors influencing occupational therapists' adoption of assisted living technology. Br J Occup Ther 2017 Nov;80(11):668-675. [doi: 10.1177/0308022617711669]
89. Gauvin F, DeMaio P, Alam S, Drakos A, Lavis J, Wilson M. Evidence brief: identifying and harnessing the potential of technology in long-term care settings in canada. : McMaster Health Forum; 2021 Feb 1 URL: https://www.mcmasterforum.org/docs/default-source/product-documents/evidence-briefs/tech-in-ltc-eb.pdf?sfvrsn=9bcb0fb_4 [accessed 2025-10-30]
90. Chu CH, Ronquillo C, Khan S, Hung L, Boscart V. Technology recommendations to support person-centered care in long-term care homes during the COVID-19 pandemic and beyond. J Aging Soc Policy 2021 Sep 3;33(4-5):539-554. [doi: 10.1080/08959420.2021.1927620]
91. Towsley GL, Edelman LS, Geurin R. Successful care conferences: nursing home staff, persons with dementia, and caregiver perspectives. J Am Med Dir Assoc 2025 Apr;26(4):105504. [doi: 10.1016/j.jamda.2025.105504] [Medline: 39961356]
92. Bertolazzi A, Quaglia V, Bongelli R. Barriers and facilitators to health technology adoption by older adults with chronic diseases: an integrative systematic review. BMC Public Health 2024 Feb 16;24(1):506. [doi: 10.1186/s12889-024-18036-5] [Medline: 38365698]
93. Heller KE, Crockett SJ, Merkel JM, Peterson JM. Focus group interviews with seniors. J Nutr Elder 1990;9(4):89-100. [doi: 10.1300/J052v09n04_07] [Medline: 2277334]
94. Zhang Z, Henley T, Schiaffino M, et al. Older adults' perceptions of community-based telehealth wellness programs: a qualitative study. Inform Health Soc Care 2022 Oct 2;47(4):361-372. [doi: 10.1080/17538157.2021.2006198] [Medline: 34822311]
95. Garcia Reyes EP, Kelly R, Buchanan G, Waycott J. Understanding older adults' experiences with technologies for health self-management: interview study. JMIR Aging 2023 Mar 21;6:e43197. [doi: 10.2196/43197] [Medline: 36943333]
96. Vergouw JW, Smits-Pelser H, Kars MC, et al. Needs, barriers and facilitators of older adults towards eHealth in general practice: a qualitative study. Prim Health Care Res Dev 2020 Dec 2;21:e54. [doi: 10.1017/S1463423620000547] [Medline: 33263272]
97. Vaportzis E, Clausen MG, Gow AJ. Older adults perceptions of technology and barriers to interacting with tablet computers: a focus group study. Front Psychol 2017 Oct 4;8:1687. [doi: 10.3389/fpsyg.2017.01687] [Medline: 29071004]
98. Cohen Elimelech O, Rosenblum S, Tsadok-Cohen M, Meyer S, Ferrante S, Demeter N. Three perspectives on older adults' daily performance, health, and technology use during COVID-19: focus group study. JMIR Aging 2024 May 30;7:e53141. [doi: 10.2196/53141] [Medline: 38825851]

Abbreviations

AT: assistive technology
ILA: independent living apartments
LTC: long-term care
SAFE: subacute unit for frail elderly
STAM: Senior Technology Acceptance Model
UTAUT: Unified Theory of Acceptance and Use of Technology

XSL•FO
RenderX

Edited by C Jacob; submitted 19.03.25; peer-reviewed by J McMurray, RN Al-Jabali; revised version received 17.07.25; accepted 10.10.25; published 12.11.25.

Please cite as:
Jaana M, Lévesque Ryan M, Tamim H, Riachy E, Paré G
Perspectives of Older Adults on Assistive Technology: Qualitative Study
JMIR Hum Factors 2025;12:e74214
URL: https://humanfactors.jmir.org/2025/1/e74214
doi:10.2196/74214

© Mirou Jaana, Maude Lévesque Ryan, Haitham Tamim, Edward Riachy, Guy Paré. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 12.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Factors Associated With the Intention to Use mHealth Among Thai Middle-Aged Adults and Older Adults: Cross-Sectional Study

Nida Buawangpong[1], MD; Wachiranun Sirikul[2,3,4,5], MSc, MD; Penprapa Siviroj[2,4], DrPH

[1]Department of Family Medicine, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[2]Department of Community Medicine, Faculty of Medicine, Chiang Mai University, 110, Intrawarorot road, Sriphum, Meaung, Chiang Mai, Thailand

[3]Center of Data Analytics and Knowledge Synthesis for Health Care, Chiang Mai University, Chiang Mai, Thailand

[4]Environmental and Occupational Medicine Excellence Center, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[5]Department of Biomedical Informatics and Clinical Epidemiology, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

Corresponding Author:
Wachiranun Sirikul, MSc, MD
Department of Community Medicine, Faculty of Medicine, Chiang Mai University, 110, Intrawarorot road, Sriphum, Meaung, Chiang Mai, Thailand

Abstract

Background: Mobile health care (mHealth) apps are emerging worldwide as a vital component of internet health care, but there are issues, especially among older adults.

Objective: We aim to investigate the factors influencing the intention to use (ITU) mHealth apps, focusing on those with and without prior mHealth experience.

Methods: A cross-sectional study conducted from August 2022 to July 2023 included Thai citizens aged 45 years or older. Self-reported questionnaires collected data on sociodemographic information, health conditions, smartphone or tablet ownership, and mHealth usage experience. The Thai mHealth Senior Technology Acceptance Model questionnaires with a 10-point Likert scale evaluated mHealth acceptance. A multivariable logistic regression analysis, adjusted for age, gender, education, income, and living area, was performed for 2 subgroups: those who used ITU mHealth apps and those who did not.

Results: Of 1100 participants, 537 (48.8%) intended to use mHealth apps, while 563 (51.2%) did not. The ITU group had a younger average age, higher education levels, higher income, and fewer underlying diseases compared to those who did not intend to use mHealth apps. For those who had never used mHealth apps, having a smartphone was strongly associated with higher odds of ITU (adjusted odds ratio 2.81, 95% CI 1.6 to 4.93; P<.001), while having any underlying disease was associated with lower odds of ITU (adjusted odds ratio 0.63, 95% CI 0.42 to 0.97; P=.034). Higher acceptance levels, characterized by a positive attitude toward mHealth and lower fear of making mistakes, were also associated with higher ITU. For those with prior mHealth experience, acceptance in areas such as perceived ease of use, gerontechnology anxiety, and facilitating conditions was significantly associated with ITU.

Conclusions: Among inexperienced users, a positive attitude toward mHealth significantly enhanced ITU. Conversely, having an underlying disease decreased ITU, indicating a need for tailored mHealth apps. For experienced users, acceptance levels in areas such as ease of use and gerontechnology anxiety were crucial. Future research should explore specific mHealth apps for more targeted insights.

(JMIR Hum Factors 2025;12:e63607) doi:10.2196/63607

KEYWORDS

mHealth; mobile healthcare; older adults; elderly; aging; questionnaire; smartphone; mHealth usage; intention to use

Introduction

Mobile health care (mHealth) apps, a vital component of internet health care, are emerging worldwide. These apps have attracted a wide range of health care services and are proven effective in solving health problems. The functions of mHealth are diverse, including patient monitoring, diagnosis, personal care, psychological health, educational apps, and social networking [1]. There is growing evidence of the health benefits of mHealth. Mobile devices, with their integrated sensors and features, help health care professionals treat patients with continuous connectivity. These apps are useful for collecting data related to physical activity, human body images, and other health care aspects [2,3]. For example, using mHealth in chronic disease management has shown improvements in symptoms and reduced hospitalizations for patients with asthma, chronic obstructive pulmonary disease, heart failure, and diabetes [4].

XSL•FO
RenderX

However, there are several issues with current mHealth apps, particularly among older adults. Studies show that only 60% of older people intend to use mHealth [5]. Older adults may have physiological changes according to aging such as visual and hearing decline. They may also be unfamiliar with technology and face difficulty learning new skills [6]. Some users experience technical problems with their smartphones when using these apps. Additionally, public awareness of mHealth apps is low, and their usability is not as good as expected [7].

Understanding the factors that affect the intention to use (ITU) mHealth among people with and without prior experience is crucial for enhancing its future usage [8]. Therefore, the primary objective of our study was to investigate the factors associated with ITU, focusing on 2 specific subgroups: those with prior experience using mHealth apps and those without. This study also examined information regarding device ownership, experience with mobile apps and mHealth, and technology acceptance. These findings aimed to gain a better understanding of mHealth usage and identify potential opportunities for implementing mHealth in a community-dwelling population, especially among older adults.

Methods

Study Design

We conducted the cross-sectional study from August 2022 to July 2023 using a nationwide web-based survey and a community survey. The online survey was disseminated through various social media platforms, such as department websites, Facebook, Line, Twitter, and Instagram. The investigator teams, which included medical students and health care personnels from primary care units across 10 subdistricts in Chiang Mai province, distributed the community survey. The respondents to both the online and community surveys used the REDCap (Research Electronic Data Capture; Vanderbilt University) survey platform to self-complete the questionnaires. The STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) statement guided this study's reporting [9].

Ethical Considerations

The institutional review board of the Faculty of Medicine, Chiang Mai University approved this study's ethical consideration (COM-2565 - 09079). Before participating in this survey, all respondents provided their informed consent in accordance with the screening questionnaire and study information page. Participants were compensated with an incentive of 100 Thai Baht (3 USD) for completing the questionnaires. No identification data were recorded, and respondents were permitted to remain anonymous during the online survey. The community survey used the identification data of eligible participants exclusively for recruitment purposes within each target area. These data were not documented in either the survey form or the study database.

Participants

Study participants were middle-aged Thai citizens aged 45 - 60 years or older at the time of the survey [10,11]. Our study's inclusion criteria required participants to be able to read and communicate in Thai and to have no underlying conditions or diseases that would hinder their ability to complete the survey or use mHealth apps (eg, dementia, active psychological problems, or severe visual impairments). This study excluded respondents who did not complete the survey, spent less than 2 minutes on it, or spent more than 60 minutes on it.

Data Collection

Participant Characteristics

We used self-reported questionnaires to collect data on participant characteristics, including sociodemographic data, underlying health conditions, owning a smartphone or tablet, and experience using the devices and mHealth apps. The initial section of the questionnaires included information regarding mHealth to ensure that the participants comprehended its definition, related concepts, and intended use.

The sociodemographic data included age, gender, marital status (single, married, and separate, divorced, or widowed), living status (alone, with family, and with others), living areas (urban, suburban, and rural), education levels (no education, primary, secondary, high school, vocational training, preuniversity, bachelor's degree, and master's degree), and income per month (<10,000 in Thai Baht, US $ 274; 10,001 - 30,000 in Thai Baht, US $275-$819; and >30,001 in Thai Baht, US $820). The questions inquiring about underlying medical conditions included hypertension, dyslipidemia, diabetes mellitus, chronic renal disease, visual impairments, and hearing impairments. We also gathered data on other related variables, such as wearing glasses or contact lenses, using hearing aids, and a number of current medications.

Experience Using Mobile Apps and mHealth

mHealth apps, or mobile health apps, refer to the practice of medicine and public health supported by mobile devices. The term "mHealth apps" encompasses the use of mobile communication devices such as smartphones and tablets for health care purposes [12,13]. For information on owning and experience using the devices and mHealth, questionnaires asked for smartphone or tablet usage experience (year), owning smartphones or tablets, internet usage experience (year), and previously used mobile apps and mHealth apps.

mHealth Acceptance and ITU

The acceptance of mHealth was evaluated using the Thai mHealth Senior Technology Acceptance Model (STAM) questionnaires. This instrument has been adapted from the 38-item STAM questionnaire [14] and validated for accessing mHealth acceptance in a Thai context. According to the reported psychometric analysis, the Thai mHealth STAM demonstrated satisfactory psychometric properties in terms of validity and reliability. We used instrument items in the following domains as the potential associated factors of an ITU mHealth: attitude toward using, perceived usefulness, perceived ease of use, perceived barriers, gerontechnology anxiety, and facilitating condition. The structure of the Thai mHealth STAM was a 10-point Likert scale. The highest point indicated a high level of acceptance. The participant's ITU mHealth was assessed by the structural question "If there are available mHealth apps for you, do you want to use them? (Yes/No)."

XSL•FO
RenderX

Sample Size

To derive statistics that represent the parameters for the target populations, a minimum sample size of 500 total participants is required. This minimum sample size can accurately detect low to large effect sizes, as suggested by Bujang et al [15]. Based on our primary objective, it aimed to explore the associations between ITU and predetermined variables, including characteristics, ownership, and experience using devices and mHealth, and mHealth acceptance, in 2 specified subgroups of participants: those with experience using mHealth apps and those without. To explore unbiased effect estimates, 18 independent variables were analyzed with adjustments for 5 confounding variables. The sample size calculation, based on the formula n=100+(10 - 50 event per variable)*i*, where *i* refers to the number of independent variables, was conducted to ensure our effect estimates from a multivariable logistic regression had sufficient precision to yield medium to large effect estimates in the subgroup analysis [15]. With 23 independent variables, a minimum sample size of 330 participants per subgroup was needed according to the given formula with 10 per independent variables.

Statistical Analysis

For descriptive analysis, we presented categorical data using frequency and percentage. We described continuous data using a mean with SD, a median with a range (minimum-maximum), or an IQR, as appropriate. The comparison in characteristics, owning and experience using the devices and mHealth, and mHealth acceptance between participants who had ITU and those who did not was examined using an independent 2-tailed *t* test or Wilcoxon rank sum test for continuous data and Fisher exact test for categorical data.

For the analysis of the primary objective, we hypothesized that the factors associated with ITU would be different for participants who had previously used mHealth and those who did not. We performed a multivariable logistic regression for 2 specified subgroups, adjusted for the confounders including age, gender, education levels, income, and living area. The predetermined characteristic variables included having any underlying disease, owning a smartphone, the number of years of experience using it, and an interaction term based on both ownership and experience. We entered each item's mHealth acceptance into the model as continuous data, using a 10-point scale. The magnitude of associations was presented using an adjusted odds ratio along with 95% CIs and *P* value. All statistical analyses were conducted using STATA (version 17.0; Stata Corp, LP). The visualization for this study was created using Microsoft Excel (Microsoft Corporation, 2021). The level of statistical significance for descriptive analysis was set at a *P* value below .05.

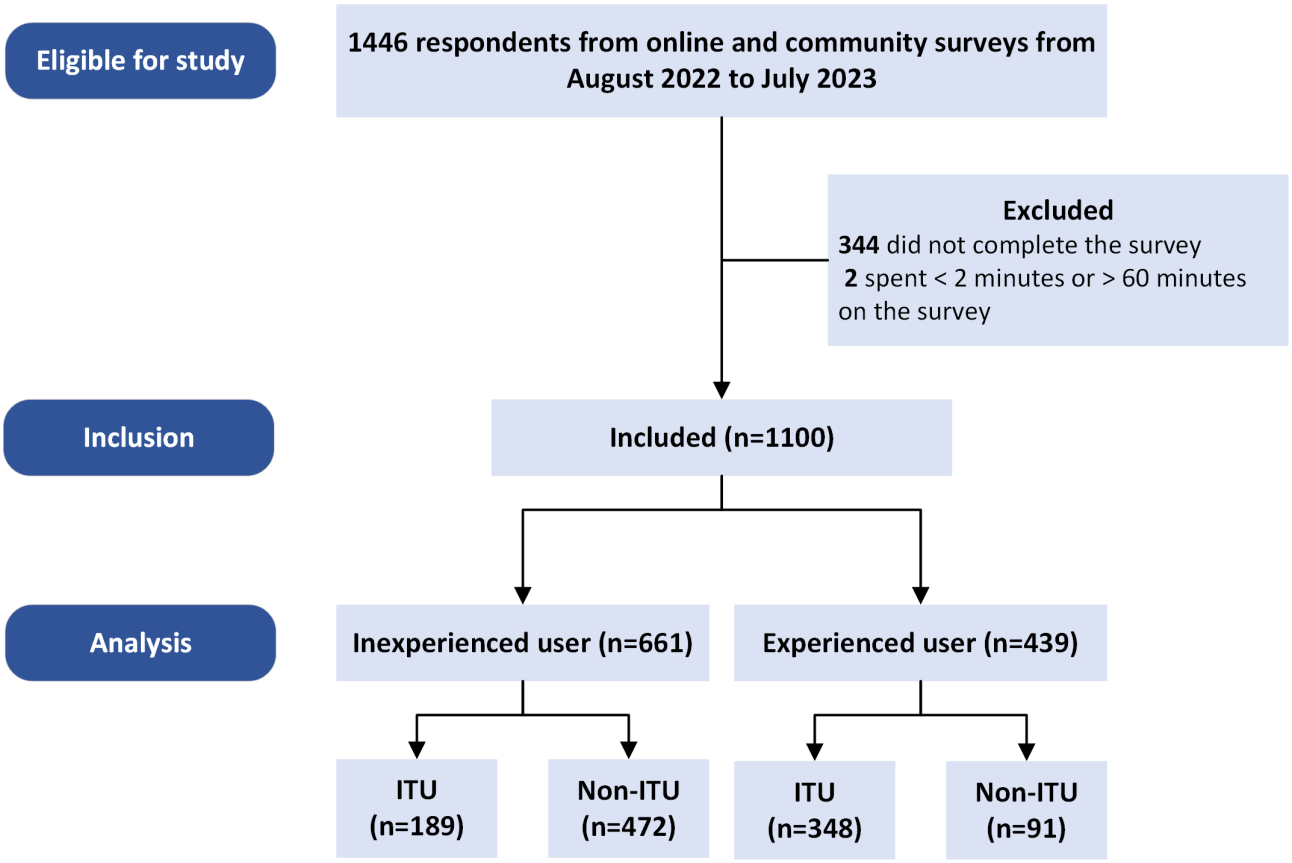
Results

Participant Characteristics

Of 1100 participants, the majority were female (776/1100, 62.3%), with a mean age of 62.3 (SD 8.8) years. This study flow diagram is presented in Figure 1. Most of the participants were married and resided with their families primarily in rural, suburban, and urban areas, respectively. For socioeconomic status, 86.2% (948/1100) had a low income (<10,000 Baht, US $274), and 65.9% (725/1100) had the highest education level at primary school. The common underlying diseases were hypertension and dyslipidemia. More than half reported having a vision problem, and 65.2% (n=399) wore glasses or contact lenses. Only 3.3% (4/1100) of respondents required a hearing aid, while 10.9% (120/1100) of respondents reported having hearing problems.

XSL•FO
RenderX

Figure 1. Study flow diagram. ITU: intention to use.



There were 537 participants who had ITU (48.8%) and 563 participants who did not (51.2%). When comparing the participants who had ITU mHealth apps with those who did not, the ITU group had a significantly younger average age, higher education levels, higher income, and a lower proportion of underlying diseases. Furthermore, 94.4% (507/537) and 92% (494/537) of the ITU group reported having experience using a smartphone or tablet and the internet; 91.2% (490/537) and 3.4% (18/537) had owned a smartphone and tablet, respectively. These percentages in the ITU group were significantly higher than those in the non-ITU group. The overall median (IQR) year of experience using the devices was 5 (0 - 10). When comparing the ITU to non-ITU groups, the median years of device usage experience in the ITU were significantly higher than non-ITU (median 6, IQR 3 - 10 versus median 3, IQR 0 - 9, P<.001). In addition, 66.9% of those in the ITU group had experience using mHealth apps, which was a significantly higher proportion than the non-ITU group of 17.4%. The details of participant characteristics are presented in Table 1.

XSL•FO
RenderX

Table . Participant characteristics.

Characteristics	Total (N=1100)	Had intention to use mHealth apps		*P* value
		Yes (n=537)	No (n=563)	
Age (year), mean (SD)	62.3 (8.8)	60 (8.4)	64.4 (8.6)	<.001
Male, n (%)	324 (29.5)	140 (26.1)	184 (32.7)	.02
Marital status, n (%)				.002
Single	96 (8.7)	57 (10.6)	39 (6.9)	
Married	747 (67.9)	377 (70.2)	370 (65.7)	
Separated, divorced, or widowed	257 (23.4)	103 (19.2)	154 (27.4)	
Education levels, n (%)				<.001
No education	18 (1.6)	2 (0.4)	16 (2.8)	
Primary school	725 (65.9)	291 (54.2)	434 (77.1)	
Secondary school	97 (8.8)	65 (12.1)	32 (5.7)	
High school and vocational training	162 (14.7)	110 (20.5)	52 (9.2)	
Preuniversity	11 (1)	8 (1.5)	3 (0.5)	
Bachelor’s degree	79 (7.2)	54 (10.1)	25 (4.4)	
Master’s degree	8 (0.7)	7 (1.3)	1 (0.2)	
Income per month, n (%)				.001
<10,000 Baht (US $274)	948 (86.2)	442 (82.3)	506 (89.9)	
10,001 - 30,000 Baht (US $275 - 819)	138 (12.5)	86 (16)	52 (9.2)	
>30,001 Baht (US $820)	14 (1.3)	9 (1.7)	5 (0.9)	
Living status, n (%)				.86
Alone	108 (9.8)	50 (9.3)	58 (10.3)	
With family	988 (89.8)	485 (90.3)	503 (89.3)	
With others	4 (0.4)	2 (0.4)	2 (0.4)	
Living area, n (%)				.38
Urban	220 (20)	99 (18.4)	121 (21.5)	
Suburban	377 (34.3)	192 (35.8)	185 (32.9)	
Rural	503 (45.7)	246 (45.8)	257 (45.6)	
Had any underlying disease, n (%)	726 (66)	330 (61.5)	396 (70.3)	.002
Hypertension, n (%)	495 (45)	212 (39.5)	283 (50.3)	<.001
Dyslipidemia, n (%)	375 (34.1)	187 (34.8)	188 (33.4)	.62
Diabetes mellitus, n (%)	184 (16.7)	76 (14.2)	108 (19.2)	.03
Chronic kidney disease, n (%)	17 (1.5)	8 (1.5)	9 (1.6)	.88
Vision problems, n (%)	612 (55.6)	304 (56.6)	308 (54.7)	.53
Wore glasses or contact lens, n (%)	399 (65.2)	210 (69.1)	189 (61.4)	.045
Hearing problems, n (%)	120 (10.9)	54 (10.1)	66 (11.7)	.38
Used hearing aids, n (%)	4 (3.3)	2 (3.7)	2 (3)	.84
Number of medications, median (IQR)	1 (0 - 2)	1 (0 - 2)	1 (0 - 2)	.55
Had experience using a smartphone or tablet, n (%)	873 (79.4)	507 (94.4)	366 (65)	<.001
Had own smartphone, n (%)	843 (76.6)	490 (91.2)	353 (62.7)	<.001
Had own tablet, n (%)	20 (1.8)	18 (3.4)	2 (0.4)	<.001
Had experience using the internet, n (%)	784 (71.3)	494 (92)	290 (51.5)	<.001

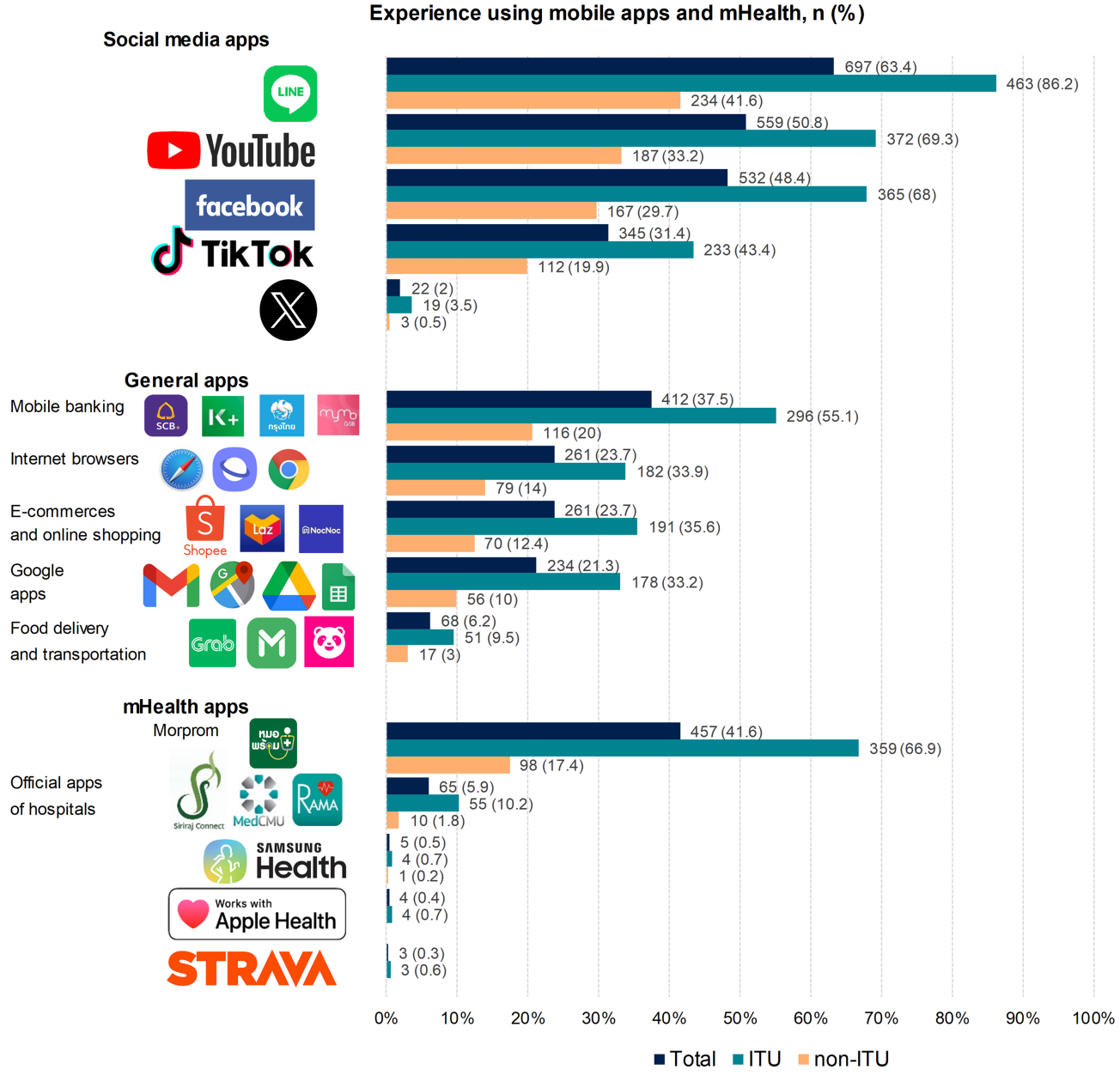
XSL•FO
RenderX

Characteristics	Total (N=1100)	Had intention to use mHealth apps		
		Yes (n=537)	No (n=563)	*P* value
Had experience using mHealth apps, n (%)	439 (39.9)	348 (64.8)	91 (16.2)	<.001

Experience Using Mobile Apps and mHealth

Figure 2 illustrates participants' experience using mobile apps and mHealth. There are many types of mobile apps and mHealth including social media apps, general apps, and mHealth apps. Among the middle-aged and older adults who had previously used mobile apps and mHealth, social media apps had the highest percentages (63.4%), followed by mHealth apps (41.6%), and apps for general purposes (37.5%). The use percentages of almost all mobile apps and mHealth were statistically and significantly higher in the ITU compared to non-ITU groups (*P*<.001), except for mHealth apps with a low use percentage (Samsung Health, Apple Health, and Strava). The use percentages of the top 3 social media apps (Line, YouTube, and Facebook) were greater than other apps. For general purposes, most participants used mobile apps for online banking services, surfing the internet, and shopping. For mHealth, the Ministry of Thai Public Health app, "MorProm," had the highest use (41.6%), followed by the official apps by hospitals (5.9%), while the other mHealth apps from public providers showed a very low use percentage.

Figure 2. Participants' experience using mobile apps and mHealth. ITU: intention to use.

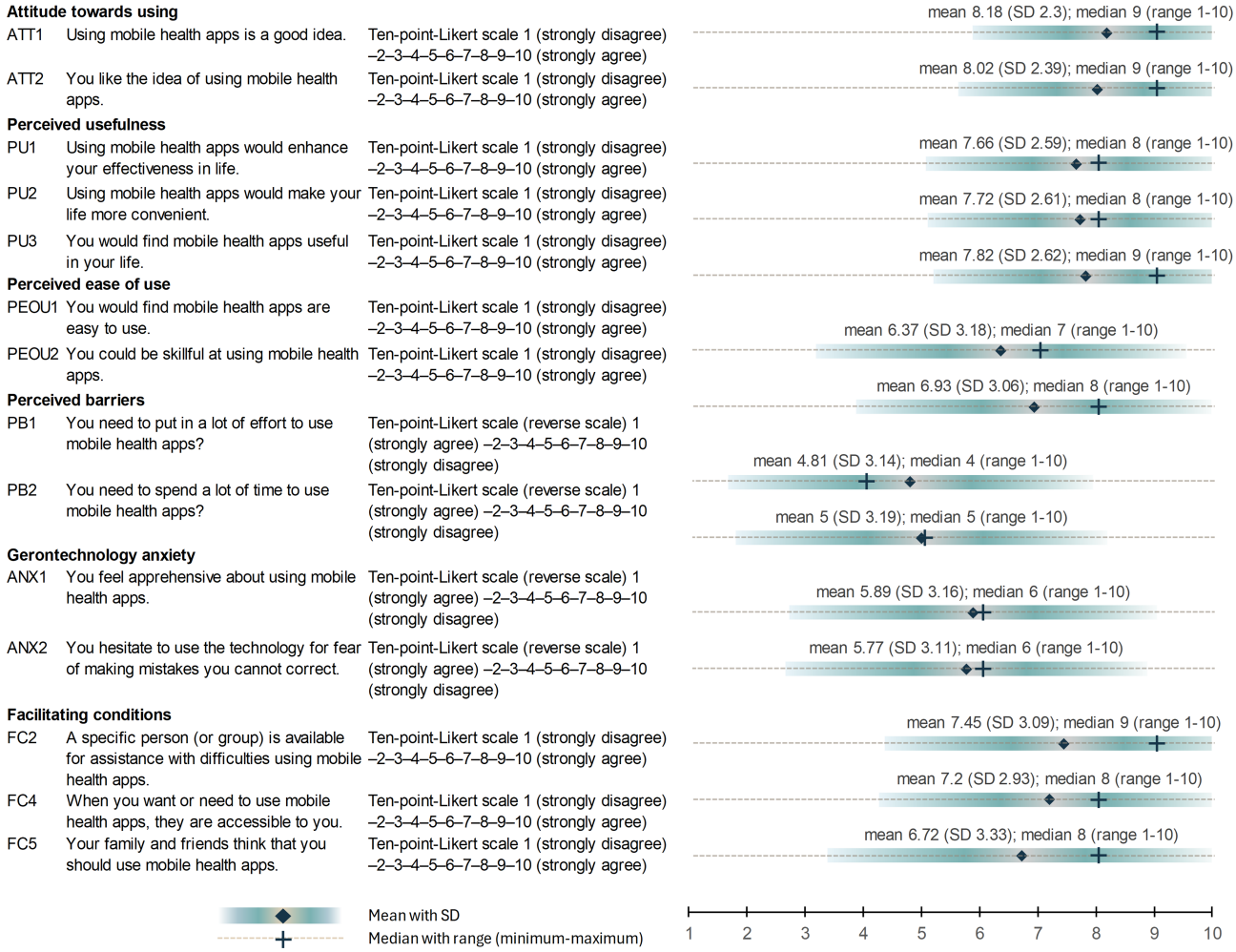


XSL•FO
RenderX

mHealth Acceptance

We assessed the participants' mHealth acceptance using the Thai mHealth STAM questionnaires in the domains, including attitude toward using, perceived usefulness, perceived ease of use, perceived barriers, gerontechnology anxiety, and facilitating condition. Figure 3 displays the structure of the questionnaires and the participants' responses. Overall, participants' acceptance of mHealth fell within the range of agreeable attitudes toward using, perceived usefulness of mHealth apps, and facilitating conditions. On the other hand, we observed responses in the range of neutral to low acceptance across the domains of perceived barriers, gerontechnology anxiety, and perceived ease of use.

Figure 3. Participants' responses to mHealth acceptance questionnaires. ANX: gerontechnology anxiety; ATT: attitude toward using; FC: facilitating condition; PB: perceived barriers; PEOU: perceived ease of use; PU: perceived usefulness.



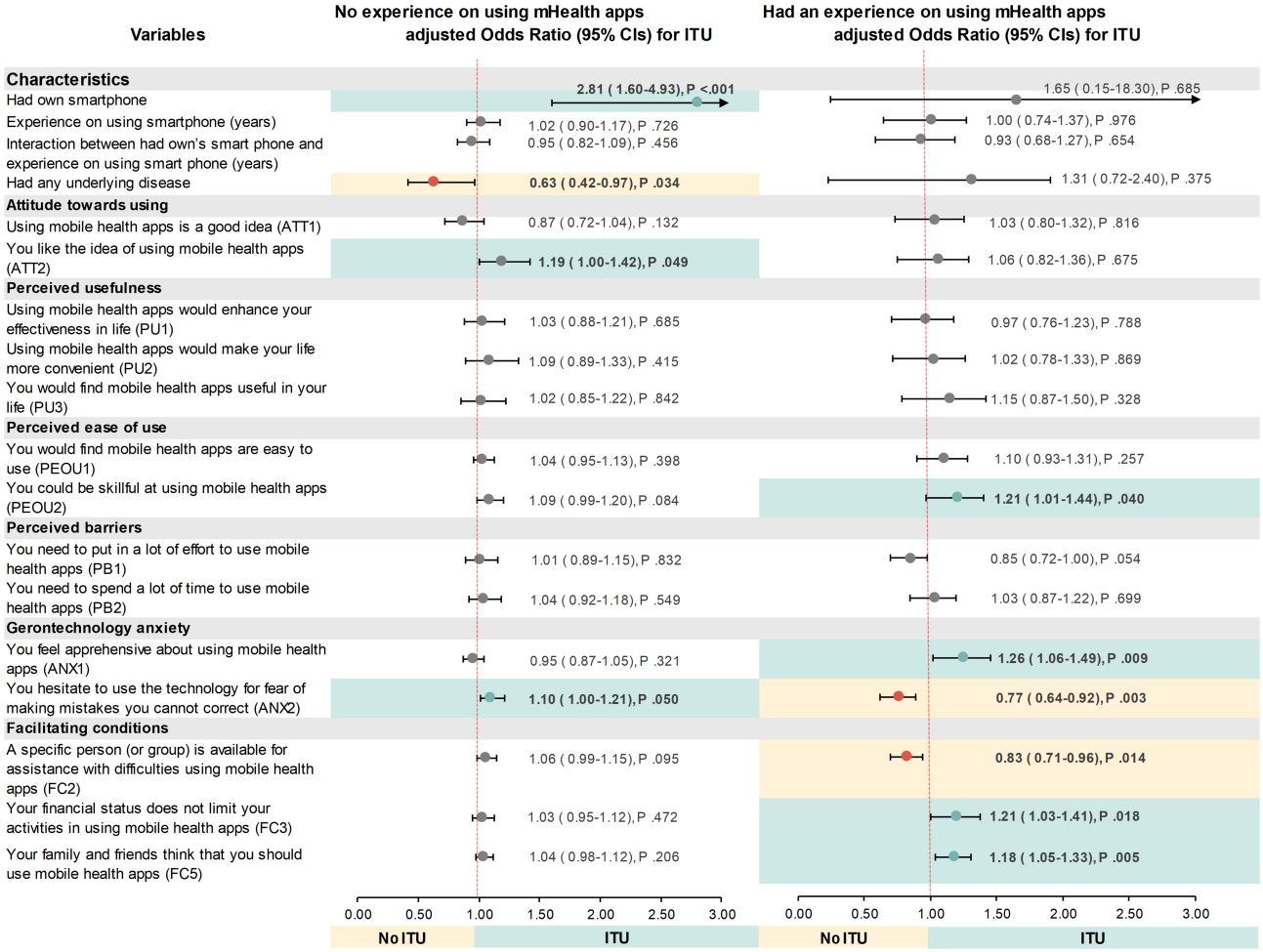
When compared between the ITU and non-ITU groups, mHealth acceptance levels assessed by the questionnaires showed a significant difference in all items, as presented in Table S1 in Multimedia Appendix 1. Participants who used an ITU mHealth app had higher mean scores in all items. The domain of perceived barriers yielded the lowest mean mHealth acceptance score in both groups.

Factors Associated With ITU mHealth Apps

In this study, we focused on exploring factors associated with ITU mHealth apps in the subgroup of participants who had experience using mHealth apps and who did not. We performed a multivariable logistic regression analysis using predetermined factors including owning a smartphone, experience using a smartphone, underlying disease, and mHeath acceptances. The analysis also adjusted for the confounders consisting of age, gender, education levels, income, and living area. Multivariable analysis results are presented in Figure 4.

XSL•FO
RenderX

Figure 4. Factors associated with intention to use mHealth apps using a multivariable logistic regression. aORs were estimated as adjusted for age, gender, education levels, income, and living area. aOR: adjusted odds ratio; ANX: gerontechnology anxiety; ATT: attitude toward using; FC: facilitating condition; ITU: intention to use; PB: perceived barriers; PEOU: perceived ease of use; PU:perceived usefulness.



For the subgroup of people who had never used mHealth apps before, having a smartphone was substantially associated with higher odds of ITU (adjusted odds ratio 2.81, 95% CI 1.60 to 4.93; *P*<.001), while having any underlying disease was significantly related to a low ITU. Higher levels of acceptance, which are characterized by a higher score of attitudes toward the idea of using mHealth and a lower fear of making mistakes that cannot be corrected when using it, were also significantly associated with a high ITU.

For the subgroup of individuals who had prior experience using mHealth apps, the pattern of associations between predetermined factors varied from the subgroup of those who did not have prior experience. There were no predetermined characteristics relating to ITU. The levels of mHealth acceptance were significantly associated with ITU in the domains, including perceived ease of use, gerontechology anxiety, and facilitating conditions. Perceived ease of use, which individuals could be skilled at using, was significantly associated with a greater odds of ITU. Regarding the level of anxiety associated with mHealth usage, less apprehensiveness about using mHealth significantly increased the odds of ITU. Conversely, a low ITU was associated with a lower fear of making mistakes when using mHealth. Having facilitated support from family and friends to use mHealth and having fewer financial constraints on using mHealth showed a significant positive association with ITU, whereas the support from others was negatively associated with ITU.

Discussion

Principal Findings

To investigate the factors associated with ITU, we found that people who had never used mHealth apps before, having a smartphone associated with a higher ITU, while having any underlying disease was significantly related to a low ITU. Higher levels of acceptance, which are characterized by a higher score of attitudes toward the idea of using mHealth and a lower fear of making mistakes, were also significantly associated with a high ITU. For the subgroup of individuals who had prior experience using mHealth apps, there were no predetermined characteristics relating to ITU. The levels of mHealth acceptance were significantly associated with ITU in the domains, including perceived ease of use, gerontechology anxiety, and facilitating conditions.

Compared to previous studies, several patient characteristics significantly influence the ITU mHealth. Older patients are less likely to use mHealth due to their unfamiliarity with technology. Despite efforts to adopt new technology, many older adult individuals still struggle with it, and the rapid pace of technological advancements leaves them unable to keep up

XSL•FO
RenderX

[16,17]. Additionally, there is a lack of promotion and support for using medical technology among older adults [18]. Socioeconomic status is another crucial factor affecting the ITU mHealth [19,20]. Patients in low-income groups not only lack support for using mHealth, but their daily routines and various burdens also reduce their intention to use these technologies [21]. This is closely related to the level of education [19]; patients with higher educational levels are more likely to understand and recognize the health benefits of mHealth, thus increasing their intention to use it [20]. Chronic diseases also play a role [19]. Patients without comorbidities and those who could effectively manage their chronic conditions tend to have better health management skills and a stronger desire to maintain good health, which positively influences their ITU mHealth [22,23].

Many factors influence the ITU mHealth among people without prior experience. One critical factor is the availability of equipment, such as mobile phones, patient testing devices, personal digital assistants, and other wireless tools. Health care professionals must ensure these facilities are accessible to patients when initiating mHealth services. Providing the necessary equipment and educating patients on its use can reduce inequity, particularly among older populations. The previous studies also showed that when health care systems provide access to such equipment, it enhances user confidence and intention to adopt mHealth, particularly among older adults who may otherwise face barriers due to digital inequality (eg, lacking smartphones or internet access) [24,25]. Empowering patients to access and use smartphones not only enhanced their ITU mHealth but could also improve self-rated, physical, and psychological health levels, as evidenced by the previous studies on the effects of smartphones and smart devices [26-28]. Underlying health conditions also impact the ITU mHealth. Patients with poor health behaviors or low motivation for self-care are less likely to adopt mHealth. The previous studies found that older adults with heart failure [29] and adults with hypertension [30] faced significant barriers to mHealth adoption, including low motivation and limited health literacy, which hindered their engagement with self-management technologies. These support our finding that underlying health conditions may reduce mHealth adoption, as individuals often lack the motivation to use health-promoting technologies. Conversely, some studies suggest that the presence of chronic health conditions can actually motivate individuals to engage with mHealth solutions. The study by Askari et al [31] found that older adults who were able to control their chronic diseases often expressed a greater ITU mHealth, as these tools can provide essential support for self-management and health monitoring. This indicated that while health conditions could pose barriers, they can also motivate the adoption of mHealth solutions, particularly when patients recognize the potential benefits for their health management. Another study also suggested that health conditions could influence older adults' readiness to engage with eHealth resources, with those experiencing chronic illnesses showing a higher willingness to use digital health tools for self-care [32]. This suggests that the context of the health condition is vital in influencing attitudes toward technology use, rather than the notion that poor health behaviors uniformly result in reduced intentions to use mHealth. Confidence in their ability to effectively use mHealth apps is crucial, as a positive attitude toward mHealth increases the likelihood of its use [33]. The positive finding on attitude toward using mHealth is also consistent with the recent systematic reviews on the impact of mHealth interventions, which highlighted that the effectiveness of these interventions is often linked to user confidence and positive attitudes toward technology [4]. To promote mHealth for people without prior experience, health care professionals should educate patients about its benefits and how it can aid in self-care. Additionally, patients must develop eHealth literacy, defined as "the ability to seek out, find, evaluate, and appraise, integrate, and apply what is gained in electronic environments toward solving a health problem." [34].

For participants with prior experience using mHealth, various factors influence their intention to continue its usage. These individuals often possess a positive attitude toward mHealth and recognize its benefits. They may also have experience in overcoming barriers, such as troubleshooting difficulties or correcting errors independently. Consequently, they perceive the benefits of using mHealth to outweigh the risks or challenges. Our finding indicated that perceived self-capability to use mHealth and lower concerns was associated with higher ITU among participants with prior mHealth experience. This finding is aligned with the study by Opoku et al [35], which emphasized that enabling resources, such as the durability and simplicity of mobile technology, significantly influence patients' perceived ease of use of mHealth interventions. This implies that when patients are informed and experienced about the user-friendly nature of mHealth apps, their confidence in using these tools increases, thereby enhancing their likelihood of adoption. In contrast to participants without prior experience using mHealth, those with experience did not avoid using mHealth out of fear of making mistakes. However, the expressed greater concerns about mHealth itself, such as its effectiveness, data security, and privacy, were associated with a low ITU. Our findings align with the studies on experienced users' attitudes toward mHealth [36,37], emphasizing the importance of addressing concerns related to security, privacy, and perceived effectiveness. Therefore, future mHealth interventions should not only focus on enhancing user confidence through health education or instruction by technical and health care providers but also on building trust in technology itself. Additionally, patients, especially older adults, value support from family or caregivers. mHealth for older adults may require assistance at various stages, including tool usage guidance, instructions, and caregiver involvement [16]. Financial considerations also play a crucial role, as increased costs associated with mHealth can affect the ITU for it. To initiate and improve mHealth usage, health care professionals need to identify patients willing to use mHealth, ensure the availability of necessary equipment, and address any financial constraints. Educating patients on the evidence and benefits of mHealth is also essential. Where possible, families and caregivers should be included in the mHealth care plan [38].

The strength of this study lies in its relatively large sample size, which included participants from rural, suburban, and urban areas. This study comprehensively explored various factors

XSL•FO
RenderX

potentially associated with mHealth use, using reliable and validated standard tools to evaluate acceptance. However, some limitations existed. The cross-sectional nature of this study limited establishing causal relationships between the ITU mHealth and associated factors. Further research designs, including longitudinal studies or randomized controlled trials of specific mHealth interventions, are necessary to overcome these limitations and provide a more comprehensive understanding of the factors influencing mHealth usage. Furthermore, the term "mHealth" was defined in a broad and general sense in the questionnaire but was not explicitly specified for participants' conditions. Defining "mHealth" in general could hinder the participants' acceptance and use of mHealth, particularly in the group of participants with underlying diseases and specific health problems. This hypothesis could explain the observed association between having any underlying disease and low ITU in the group of participants who had never used mHealth before. Accordingly, our study primarily aimed to explore mHealth usage in general for community-dwelling adults. Our findings may not generalize enough to provide specific insight into the specific use case of mHealth in a particular population. Hence, we encourage future studies to aim for an in-depth understanding of each type of mHealth and acceptance in the specific context.

Conclusion

This study identified key factors influencing the ITU mHealth and positive attitudes toward using significantly enhancing ITU in both inexperienced and experienced users. Conversely, having an underlying disease was associated with a decrease in ITU, indicating that this group of individuals may require mHealth's specific requirements for their conditions rather than the general purpose. For experienced users, acceptance levels in areas such as ease of use and gerontechnology anxiety were critical. Health care professionals should identify patients open to using mHealth, ensure access to necessary equipment, address financial barriers, and educate patients on its benefits. Future investigations should also explore specific mHealth apps to provide more targeted insights.

Acknowledgments

This work was supported by Chiang Mai University and the Faculty of Medicine Research Fund (grant 124-2565).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additonal table. The comparison of participants' responses to mHealth acceptance questionnaires.
[DOCX File, 19 KB - humanfactors_v12i1e63607_app1.docx]

References

1. Pires IM, Marques G, Garcia NM, Flórez-Revuelta F, Ponciano V, Oniani S. A research on the classification and applicability of the mobile health applications. J Pers Med 2020 Feb 27;10(1):11. [doi: 10.3390/jpm10010011] [Medline: 32120849]
2. Pires IM, Garcia NM, Pombo N, Flórez-Revuelta F, Rodríguez ND. Validation techniques for sensor data in mobile health applications. J Sens 2016;2016(1):1-9. [doi: 10.1155/2016/2839372]
3. Pires IM, Marques G, Garcia NM, et al. A review on the artificial intelligence algorithms for the recognition of activities of daily living using sensors in mobile devices. In: Singh PK, Bhargava BK, Paprzycki M, Kaushal NC, Hong WC, editors. Handbook of Wireless Sensor Networks: Issues and Challenges in Current Scenario's: Springer International Publishing; 2020:685-713.
4. Marcolino MS, Oliveira JAQ, D'Agostino M, Ribeiro AL, Alkmim MBM, Novillo-Ortiz D. The impact of mHealth interventions: systematic review of systematic reviews. JMIR mHealth uHealth 2018 Jan 17;6(1):e23. [doi: 10.2196/mhealth.8873] [Medline: 29343463]
5. van Elburg FRT, van de Klundert J, Nieboer AP, Askari M. The intention to use mHealth applications among Dutch older adults prior and during the COVID pandemic. Front Public Health 2023;11:1130570. [doi: 10.3389/fpubh.2023.1130570] [Medline: 37383259]
6. Buawangpong N, Pinyopornpanish K, Pliannuom S, et al. Designing telemedicine for older adults with multimorbidity: content analysis study. JMIR Aging 2024 Jan 10;7:e52031. [doi: 10.2196/52031] [Medline: 38198201]
7. Mäkinen H, Haavisto E, Havola S, Koivisto JM. User experiences of virtual reality technologies for healthcare in learning: an integrative review. Behav Inf Technol 2022 Jan 2;41(1):1-17. [doi: 10.1080/0144929X.2020.1788162]
8. Hoque R, Sorwar G. Understanding factors influencing the adoption of mHealth by the elderly: an extension of the UTAUT model. Int J Med Inform 2017 May;101:75-84. [doi: 10.1016/j.ijmedinf.2017.02.002] [Medline: 28347450]
9. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann Intern Med 2007 Oct 16;147(8):573-577. [doi: 10.7326/0003-4819-147-8-200710160-00010] [Medline: 17938396]

XSL•FO
RenderX

10. Anantanasuwong D, Theerawanviwat D, Siripanich P. Panel survey and study on health and aging, and retirement in Thailand. In: Encyclopedia of Gerontology and Population Aging: Springer; 2019, Vol. 10:978-973. [doi: 10.1007/978-3-319-69892-2_982-1]
11. Pengpid S, Peltzer K. Prevalence and associated factors of incident and persistent loneliness among middle-aged and older adults in Thailand. BMC Psychol 2023 Mar 14;11(1):70. [doi: 10.1186/s40359-023-01115-4] [Medline: 36918991]
12. Twilt M. mHealth applications are they the future or another burden for our patients? mHealth 2023;9:9. [doi: 10.21037/mhealth-23-5] [Medline: 37089269]
13. Adibi S. Mobile Health: A Technology Road Map: Springer; 2015, Vol. 5. [doi: 10.1007/978-3-319-12817-7]
14. Chen K, Chan AHS. Gerontechnology acceptance by elderly Hong Kong Chinese: a Senior Technology Acceptance Model (STAM). Ergonomics 2014;57(5):635-652. [doi: 10.1080/00140139.2014.895855] [Medline: 24655221]
15. Bujang MA, Sa'at N, Sidik T, Joo LC. Sample size guidelines for logistic regression from observational studies with large population: emphasis on the accuracy between statistics and parameters based on real life clinical data. Malays J Med Sci 2018 Jul;25(4):122-130. [doi: 10.21315/mjms2018.25.4.12] [Medline: 30914854]
16. Elbeddini A, Prabaharan T, Almasalkhi S, Tran C, Zhou Y. Barriers to conducting deprescribing in the elderly population amid the COVID-19 pandemic. Res Social Adm Pharm 2021 Jan;17(1):1942-1945. [doi: 10.1016/j.sapharm.2020.05.025] [Medline: 32499161]
17. Liu N, Yin J, Tan SSL, Ngiam KY, Teo HH. Mobile health applications for older adults: a systematic review of interface and persuasive feature design. J Am Med Inform Assoc 2021 Oct 12;28(11):2483-2501. [doi: 10.1093/jamia/ocab151] [Medline: 34472601]
18. Mace RA, Mattos MK, Vranceanu AM. Older adults can use technology: why healthcare professionals must overcome ageism in digital health. Transl Behav Med 2022 Dec 30;12(12):1102-1105. [doi: 10.1093/tbm/ibac070] [Medline: 36073770]
19. Jacob C, Sezgin E, Sanchez-Vazquez A, Ivory C. Sociotechnical factors affecting patients' adoption of mobile health tools: systematic literature review and narrative synthesis. JMIR mHealth uHealth 2022 May 5;10(5):e36284. [doi: 10.2196/36284] [Medline: 35318189]
20. Lee M, Kang D, Yoon J, et al. The difference in knowledge and attitudes of using mobile health applications between actual user and non-user among adults aged 50 and older. PLoS ONE 2020;15(10):e0241350. [doi: 10.1371/journal.pone.0241350]
21. McCool J, Dobson R, Whittaker R, Paton C. Mobile health (mHealth) in low- and middle-income countries. Annu Rev Public Health 2022 Apr 5;43:525-539. [doi: 10.1146/annurev-publhealth-052620-093850] [Medline: 34648368]
22. Mahmood A, Kedia S, Wyant DK, Ahn S, Bhuyan SS. Use of mobile health applications for health-promoting behavior among individuals with chronic medical conditions. Digit Health 2019;5:2055207619882181. [doi: 10.1177/2055207619882181] [Medline: 31656632]
23. Jaana M, Paré G. Comparison of mobile health technology use for self-tracking between older adults and the general adult population in Canada: cross-sectional survey. JMIR mHealth uHealth 2020 Nov 27;8(11):e24718. [doi: 10.2196/24718] [Medline: 33104517]
24. Ye Q, Deng Z, Chen Y, Liao J, Li G, Lu Y. How resource scarcity and accessibility affect patients' usage of mobile health in China: resource competition perspective. JMIR mHealth uHealth 2019 Aug 9;7(8):e13491. [doi: 10.2196/13491] [Medline: 31400104]
25. Hall CS, Fottrell E, Wilkinson S, Byass P. Assessing the impact of mHealth interventions in low- and middle-income countries--what has been shown to work? Glob Health Action 2014;7(1):25606. [doi: 10.3402/gha.v7.25606] [Medline: 25361730]
26. Li G, Jin C, Zhao B, Wu B. Smartphone use, technology affordance for healthcare and elders' life satisfaction. Front Public Health 2022;10:861897. [doi: 10.3389/fpubh.2022.861897] [Medline: 35480578]
27. Liang X, Xiong F, Xie F. The effect of smartphones on the self-rated health levels of the elderly. BMC Public Health 2022 Mar 15;22(1):508. [doi: 10.1186/s12889-022-12952-0] [Medline: 35292010]
28. Wei Y, Guo X. Impact of smart device use on objective and subjective health of older adults: findings from four provinces in China. Front Public Health 2023;11:1118207. [doi: 10.3389/fpubh.2023.1118207] [Medline: 37546303]
29. Cajita MI, Hodgson NA, Lam KW, Yoo S, Han HR. Facilitators of and barriers to mHealth adoption in older adults with heart failure. Comput Inform Nurs 2018 Aug;36(8):376-382. [doi: 10.1097/CIN.0000000000000442] [Medline: 29742549]
30. Chen Y, Kruahong S, Elias S, et al. Racial disparities in shared decision-making and the use of mHealth technology among adults with hypertension in the 2017-2020 health information national trends survey: cross-sectional study in the United States. J Med Internet Res 2023 Sep 13;25:e47566. [doi: 10.2196/47566] [Medline: 37703088]
31. Askari M, Klaver NS, van Gestel TJ, van de Klundert J. Intention to use medical apps among older adults in the Netherlands: cross-sectional study. J Med Internet Res 2020 Sep 4;22(9):e18080. [doi: 10.2196/18080] [Medline: 32624465]
32. Gordon NP, Hornbrook MC. Older adults' readiness to engage with eHealth patient education and self-care resources: a cross-sectional survey. BMC Health Serv Res 2018 Mar 27;18(1):220. [doi: 10.1186/s12913-018-2986-0] [Medline: 29587721]
33. Azhar FA, Dhillon JS. An investigation of factors influencing the intention to use mHealth apps for self-care. IJBIS 2018;29(1):59. [doi: 10.1504/IJBIS.2018.094005]

XSL•FO
RenderX

34. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. J Med Internet Res 2006 Jun 16;8(2):e9. [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]
35. Opoku D, Stephani V, Quentin W. A realist review of mobile phone-based health interventions for non-communicable disease management in sub-Saharan Africa. BMC Med 2017 Feb 6;15(1):24. [doi: 10.1186/s12916-017-0782-z] [Medline: 28162090]
36. Hossain I, Lim ZZ, Ng JJL, Koh WJ, Wong PS. Public attitudes towards mobile health in Singapore: a cross-sectional study. mHealth 2018;4:41. [doi: 10.21037/mhealth.2018.09.02] [Medline: 30363776]
37. Korpershoek YJG, Vervoort SCJM, Trappenburg JCA, Schuurmans MJ. Perceptions of patients with chronic obstructive pulmonary disease and their health care providers towards using mHealth for self-management of exacerbations: a qualitative study. BMC Health Serv Res 2018 Oct 4;18(1):757. [doi: 10.1186/s12913-018-3545-4] [Medline: 30286761]
38. Deniz-Garcia A, Fabelo H, Rodriguez-Almeida AJ, et al. Quality, usability, and effectiveness of mhealth apps and the role of artificial intelligence: current scenario and challenges. J Med Internet Res 2023 May 4;25:e44030. [doi: 10.2196/44030] [Medline: 37140973]

Abbreviations

ITU: intention to use
mHealth: mobile health care
STAM: Senior Technology Acceptance Model
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Kushniruk; submitted 25.06.24; peer-reviewed by P Nohl-Deryk, X Zhang; revised version received 04.11.24; accepted 18.02.25; published 07.03.25.

Please cite as:
Buawangpong N, Sirikul W, Siviroj P
Factors Associated With the Intention to Use mHealth Among Thai Middle-Aged Adults and Older Adults: Cross-Sectional Study
JMIR Hum Factors 2025;12:e63607
URL: https://humanfactors.jmir.org/2025/1/e63607
doi:10.2196/63607

© Nida Buawangpong, Wachiranun Sirikul, Penprapa Siviroj. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 7.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Assessing the Relationship Between the Type of Internet Use and Internet Addiction in Early and Middle Adolescents: Cross-Sectional Study From Qatar

Khansa Chemnad[1], MSc; Maryam Aziz[1], MSc; Sanaa Al- Harahsheh[2], PhD; Azza Abdelmoneium[3], PhD; Ahmed Baghdady[4], EdD; Diana Alsayed Hassan[5], PhD; Raian Ali[1], PhD

[1]College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar

[2]World Innovation Summit for Health, Qatar Foundation, Doha, Qatar

[3]Doha International Family Institute, Qatar Foundation, Doha, Qatar

[4]World Innovation Summit for Education, Qatar Foundation, Doha, Qatar

[5]Department of Public Health, College of Health Sciences, QU Health, Qatar University, Doha, Qatar

Corresponding Author:
Diana Alsayed Hassan, PhD
Department of Public Health
College of Health Sciences
QU Health, Qatar University
PO Box 2713
Doha
Qatar
Phone: 974 31576271
Email: dalsayed@qu.edu.qa

Abstract

Background: With the increasing prevalence of digital technology, adolescent internet addiction (IA) has become a global concern. Excessive internet use, especially among adolescents, has been linked to various negative outcomes such as poor academic performance, social isolation, and mental health issues. Conducted among adolescents of Arab origin, our study addressed the limitations of the literature, which predominantly focuses on Western, educated, industrialized, rich, and democratic populations.

Objective: This study aimed to differentiate between essential and nonessential internet use and how they relate to IA in early and middle adolescents, as well as the relationship between subjective happiness with the amount of time spent on nonessential internet use and IA.

Methods: A cross-sectional survey was conducted among 377 students from 16 schools in Qatar. The survey measured essential and nonessential internet use, subjective happiness with nonessential use, and IA symptoms using the Internet Addiction Diagnostic Questionnaire, as well as participant demographics. To explore age-specific associations, participants were categorized into early (age 11-13 years) and middle (age 14-17 years) adolescents. Factorial analysis, multiple regression, and logistic regression were used for statistical analysis.

Results: Nonessential internet use significantly predicted IA in both early (P<.001) and middle (P<.001) adolescents, with early adolescents showing a stronger association. Subjective happiness with nonessential internet use negatively predicted IA only in middle adolescents (P<.001) as greater dissatisfaction led to a higher IA risk. Essential internet use did not predict IA in either group.

Conclusions: Differentiating between essential and nonessential internet use is crucial in understanding IA. This study highlights the importance of developmental differences in shaping IA symptoms. The findings suggest that interventions aimed at addressing IA should be age specific and focus on addressing nonessential use specifically rather than considering internet use and screen time in general as a single entity. Cultural and regional factors also play a role in shaping internet use patterns and IA in the Middle East, necessitating context-specific, culturally sensitive approaches to IA prevention.

(JMIR Hum Factors 2025;12:e62955) doi:10.2196/62955

XSL•FO
RenderX

KEYWORDS

internet addiction; internet use; early adolescence; middle adolescence; mobile phone

Introduction

Background

Adolescent internet addiction (IA) has become a growing concern in today's society as the use of technology and the internet has become increasingly prevalent in the lives of young people. According to a study conducted by the Pew Research Center, since 2014 to 2015, there has been a modest increase in the proportion of adolescents who said that they used the internet daily or more often [1]. In a follow-up study in 2022, a total of 97% of adolescents said that they used the internet daily compared to 92% of adolescents who said the same in 2014 to 2015. It is understandable that many adolescents struggle with IA given the widespread use of the internet. Studies have shown that approximately 10% of adolescents are at risk of IA, with even higher rates reported in countries such as Hong Kong, China, and South Korea [2-5]. For example, a survey conducted in South Korea found that nearly 30% of junior high school students were either addicted to the internet or at a risk of IA [6]. While some internet use can be beneficial for educational and social purposes, excessive use can lead to negative consequences such as decreased academic performance, social isolation, and even mental health issues [7-9]. Therefore, it is important to understand the prevalence of IA among adolescents and its contributing factors.

IA is defined as the inability to control internet use, leading to disruptions in everyday functioning and symptoms of withdrawal and tolerance [10]. Although not officially recognized as a disorder in the most recent edition of the *Diagnostic and Statistical Manual of Mental Disorders*, IA has received considerable attention in scientific literature and is regarded as a growing concern [11]. Opinions in the literature differ, with some researchers suggesting distinct diagnostic criteria and subtypes of IA, such as video game and online gambling addiction [12]. However, there is still a lack of consensus regarding the internet's addictive nature and IA diagnostic criteria [13]. While there is still debate regarding the exact nature and criteria for IA, it is clear that excessive internet use can have negative consequences, particularly for adolescents. Understanding the prevalence and factors contributing to IA is important for addressing this growing concern, promoting healthy internet use, and developing appropriate interventions for those who may be affected.

Most psychological research studies predominantly rely on samples from Western, educated, industrialized, rich, and democratic (WEIRD) populations [14]. However, when examining the effects of excessive internet use on adolescent mental health, it is crucial to consider that the experiences of adolescents from non-WEIRD populations may vary owing to many factors, such as different cultural norms and access to technology [15]. This emphasizes the importance of including more diverse samples in research to gain a comprehensive understanding of the global impact of technology on mental health. Cultural context significantly shapes adolescence, with the experiences of young individuals varying widely across different cultures [16]. To address this gap, our study focused on adolescent IA in the Middle East, where limited research on IA has been conducted. In doing so, we aimed to shed light on the specific dynamics of IA in this region and contribute to a more inclusive and comprehensive understanding of this phenomenon.

Qatar has one of the highest internet penetration rates in the world at 99% [17]. The country's internet users have grown significantly from 69% in 2012 to 100% in 2022 [18]. This increase raises the possibility of an increase in IA prevalence rates, as does the spread of mobile device use and the constant release of new technology updates. Furthermore, adolescents may unintentionally be encouraged to spend more time inside streaming online content, playing online games, and spending time on-screen due to Qatar's humid and hot weather, which lasts for more than half the year.

In addition to Qatar's unique technological landscape and climate, broader Middle Eastern cultural factors play a crucial role in shaping adolescent internet use patterns. The region's complex environment, which is characterized by its unique family dynamics, societal expectations, and cultural values, may contribute to the elevated prevalence of IA among adolescents [19,20]. In various Middle Eastern societies, conservative norms that prioritize modesty and restraint frequently conflict with the liberties that online platforms provide. As adolescents navigate their identities and seek communities that are consistent with their evolving values, this cultural tension can motivate them to engage with the internet differently. In addition, it is also argued that the limited opportunities for social interaction outside the home, along with societal pressures to conform, may encourage adolescents to seek connection and validation online, potentially nurturing addictive behaviors [19]. Culturally informed research can inform tailored intervention strategies that address the specific needs of adolescents in the Middle East. For example, studies suggest that effective interventions should involve families and schools, recognizing the significant role that these environments play in shaping adolescents' internet behaviors [21,22].

Excessive screen time is often identified as a prominent indicator of IA, contributing to an unhealthy relationship with the internet and development of negative habits [23,24]. In numerous studies examining adolescent IA, the amount of time spent on the web is commonly regarded as a key factor in determining the presence of IA [25,26]. The COVID-19 pandemic has further heightened the reliance on technology for remote learning and socializing among adolescents [27]. The term *screen time* is frequently used without sufficient attention being paid to the specific activities undertaken during that period. However, this lack of specificity has several methodological limitations. To address these limitations, this study aimed to differentiate between the time spent on essential and nonessential internet use activities, providing a more nuanced understanding of IA in adolescents. However, IA is not just a matter of excessive

XSL•FO
RenderX

screen time; it has also been linked to mental health issues such as loneliness, low self-esteem, sleep problems, depression, social phobia, and anxiety [28,29]. Adolescents who struggle with these conditions may turn to the internet as a coping mechanism, further exacerbating their addiction [30]. Ultimately, IA can have a significant negative impact on adolescents' lives, from decreased academic performance to strained relationships with loved ones [31].

Adolescence is a crucial period for identity exploration and formation during which individuals may encounter situations in which their emerging beliefs and values conflict with societal expectations or their own behaviors, which may lead to cognitive dissonance [32]. This aspect of adolescent development is relevant to the exploration of subjective feelings regarding digital technology use and its association with IA. Cognitive dissonance theory offers an understanding of this relationship by examining potential conflicts arising from adolescents' subjective feelings regarding their digital technology use and addictive behavior. According to the cognitive dissonance theory, individuals experience psychological discomfort when they hold contradictory beliefs, attitudes, or behaviors [33]. In the context of digital technology use and IA, cognitive dissonance may arise when adolescents recognize the negative consequences of excessive digital technology use but continue to engage in it. The literature suggests that subjective positive feelings about one's own life conditions are negatively correlated with IA [34]. It is reasonable to hypothesize that similar correlations may exist between subjective feelings regarding digital technology use and IA. However, there is a lack of specific studies exploring this association. Given this gap in the literature, this study aimed to address this by examining the association between subjective feelings about digital technology use and IA.

Adolescents are growing up in an era characterized by rapid technological advancements, and the use of smartphones has become a common part of their daily lives. This pervasive technology has greatly influenced their interactions with the internet, particularly social networking sites [35]. Consequently, this has had a profound impact on various aspects of their lives, including social dynamics, cognitive development, and emotional well-being. The literature indicates notable distinctions between early and middle adolescents in terms of cognitive and psychological development [36]. It is critical to distinguish between early and middle adolescence when investigating adolescent IA owing to the distinct developmental traits, experiences, and coping strategies that arise throughout these separate phases of adolescence [37]. Early adolescence is a critical period defined by the start of puberty and the transition from childhood to adolescence. This stage is distinguished by rapid physical, cognitive, and social changes, as well as identity exploration and self-concept construction [38,39]. Furthermore, early adolescence is a critical time for the initiation and development of mental health issues [40]. Middle adolescence is characterized by managing the complications of moving into young adulthood [41]. Individuals become more self-sufficient, engage in complex social connections, and gain feelings of autonomy [42]. If these aspects of development are not addressed properly during adolescence, consequences may have a lasting impact in their adulthood. Researchers must examine the diverse vulnerabilities, risk factors, and patterns of internet use among early and middle adolescents while studying IA during these different phases. Early adolescents may be more vulnerable to IA because they are constantly trying new things, including online platforms, to define their identities and connect with friends [39]. Middle adolescence, on the other hand, is characterized by the emergence of more stable patterns of behavior, including internet use, as adolescents approach young adulthood and deal with increasing academic and social demands [43]. Furthermore, research has shown that the characteristics that influence IA change between early and middle adolescence [21]. For example, parental supervision and advice may have a greater effect among early adolescents, but peer influence and social support may become more significant in middle adolescence [44,45]. Furthermore, in middle adolescence, the onset of academic commitment and the need for self-regulation might alter internet use patterns and addiction risks. IA has become a growing concern among adolescents, with symptoms presenting both in relation to oneself and in interactions with others. However, the cognitive development and behavioral patterns of early and middle adolescents differ significantly, suggesting the need for a critical examination of IA symptoms according to age. Although previous research has highlighted the multifaceted nature of IA symptoms, little attention has been paid to the potential differential manifestation of symptoms based on the age of adolescents. Understanding how IA symptoms manifest differently in early and middle adolescence is crucial for tailoring prevention and intervention efforts. Understanding the differences between early and middle adolescence is critical when examining IA. Our study aimed to address the existing knowledge gap by investigating aspects that have not been thoroughly explored in the field of adolescent IA.

Adolescent technology use can be categorized into distinct types depending on type of use: essential and nonessential [46,47]. Activities that are essential for everyday life, such as educational pursuits or personal improvement–related activities, are included in essential use. On the other hand, nonessential use includes leisure activities such as gaming, social media engagement for nonspecific developmental purpose, and other forms of entertainment [46]. Although researchers have investigated the correlations between excessive internet use and IA, there is a noticeable absence of studies that explicitly distinguish between essential and nonessential use patterns. We suggest that the use of the internet for essential activities related to education and personal development can improve efficiency and time management and may not be linked to IA. In contrast, the excessive use of the internet for nonessential activities may help change mood and increase bonding but may also result in diminished productivity and potential adverse consequences [48]. Given this distinction, our study endeavored to examine the distinct effects of essential versus nonessential internet use on adolescents, thereby contributing to resolving a critical gap in the current literature and offering valuable insights for comprehending healthy internet use patterns.

XSL•FO
RenderX

Objectives

While extensive research has been conducted on IA among adolescents, several significant gaps remain. First, there is a lack of studies focusing on non-WEIRD populations, particularly in the Middle East, where unique cultural factors may influence IA. Second, the potential significance of the distinction between essential and nonessential internet use in relation to IA has not been adequately investigated. Third, there is a scarcity of research on the correlation between subjective emotions related to the use of digital technology and IA. Finally, while the developmental differences between early and middle adolescence are well documented, there are few studies that have investigated the potential differences in the manifestation of IA symptoms and their relationship to technology use across these age groups. We aimed to answer the following three research questions (RQs) through this study:

1. Do essential and nonessential types of internet use predict IA in early and middle adolescents? (RQ 1)
2. Can subjective happiness with nonessential internet use predict adolescent IA? (RQ 2)
3. Does the impact of time spent on essential and nonessential internet use, as well as subjective feelings regarding nonessential internet use, on IA differ between the early and middle adolescence stages? (RQ 3)

Methods

Participants

This study involved school students residing in Qatar, and their participation was secured through web-based surveys administered via SurveyMonkey (SurveyMonkey Inc). The survey invitation was sent out to multiple schools in Qatar through an open call using the mailing list of head teachers and educators. One of the authors works at an education policy institute with a network of schools that attend events and participate in research activities. This facilitated outreach. A total of 16 schools accepted the study invitation, comprising an equal number of public (n=8, 50%) and private (n=8, 50%) schools. Public schools in Qatar are government funded and managed by the government and primarily enroll Qatari citizens as they provide free education with Arabic as the language of instruction. These schools are also gender segregated, meaning that male and female students attend separate institutions. Private schools, on the other hand, are funded by tuition fees and managed by private organizations and predominantly enroll expatriate students. These schools follow a variety of government-reviewed international curricula and are typically not gender segregated. Data were collected between March 2022 and May 2022.

Before administering the survey, students were provided with an induction session by their head teacher, who explained the distinction between essential and nonessential internet use. Essential use was defined as internet activities connected to study, academic work, and personal development, whereas nonessential use was defined as activities that do not directly contribute to these aims. This briefing was designed to establish a common understanding among participants, guaranteeing that their answers precisely reflected these definitions. The authors acknowledge the difficulty in standardizing what adolescents consider essential versus nonessential. Therefore, our study and results should be interpreted as reflecting the subjective views and self-reports of the adolescent participants. This interpretation follows an induction on general understanding, which aimed to reduce diversity in perceptions.

We used the formula by Green [49] to calculate the optimal sample size for our research. The formula indicates that a linear regression analysis requires a minimum sample size of 50 plus 8 times the number of independent variables p ($50 + 8 \times p$). This indicates that a minimum sample size of 74 participants was suitable for our study. In addition, to ensure correlation stability [50], our aim was to have a sample size of >250 individuals. This sample size is consistent with recommendations for maintaining statistical power in studies with multiple predictors [49,50]. Furthermore, the male sample was deemed too small to segregate the analysis by gender.

After excluding participants who did not meet the inclusion criteria, such as those who provided incomplete responses, 64.3% (377/586) of the initial students were included in the analysis, which fulfilled the criteria of a 5% margin of error at the 95% CI. Of these 377 students, 86 (22.8%) were male and 291 (77.2%) were female. The gender distribution in our sample reflects the timing of data collection, which coincided with final exams, as well as the demographic characteristics of the responding schools. Specifically, 12% (2/16) of the schools that responded to the survey were predominantly female-only public schools, which significantly contributed to the observed gender imbalance. It is important to note that gender was not included as a variable in our analysis as the study's primary focus was on essential and nonessential internet use. On the basis of the estimates of the sample size and the need for adequate statistical power in our regression analysis, it was not possible to divide the sample by gender for separate analyses.

The average age of the participating adolescents was 13.19 (SD 1.24; range 11-17) years. The participants in the study were requested to explicitly provide their ages. Participants were categorized into 2 groups: early adolescents (aged 11-13 years) and middle adolescents (aged 14-17 years). The classification of adolescents into these developmental phases was determined based on the existing literature [51].

Ethical Considerations

This study was approved by the institutional review board of Hamad Bin Khalifa University (QBRI-IRB 2021-05-094). School permissions were obtained to distribute the survey. The parents and adolescents were informed beforehand of the study, and their written informed consent and assent were collected. Participants were informed about the study's purpose, and their involvement was voluntary, allowing them to skip questions or withdraw from the survey if desired. During data collection, no identifiable information was collected. The researchers ensured that all deidentified data were stored on a password-protected laptop that only the primary investigators of the study had access to. Participants were not compensated for taking part in this study.

XSL•FO
RenderX

Measures

Overview

The survey included questions regarding participants' demographics, digital technology use, subjective happiness with nonessential internet use, and IA. The questionnaire was distributed in both English and Arabic to accommodate language preferences, and the accuracy of the Arabic translation was ensured using the back translation method [52] (see Multimedia Appendix 1 for both versions of the survey). Our study included a comprehensive and multistep procedure for validating the survey questions to establish its reliability and suitability for the adolescent population. We began by conducting think-aloud protocols with a subset of 3 adolescents, asking them to verbalize their thought processes while answering the survey questions [53]. This approach helped us identify potential ambiguities or misunderstandings, leading to revisions in translation that ensured clarity and suitability for the target age group yet without departing from the original meaning of the questions in the English version. In total, 3 authors with extensive expertise dealing with adolescent groups in the Arab area offered expert advice. Furthermore, the questions in the survey were culturally appropriate, did not use any jargon, and were free of unclear language.

Internet Addiction Diagnostic Questionnaire

The Internet Addiction Diagnostic Questionnaire (IADQ) consists of 8 items, each representing a symptom used to identify IA [10]. The questionnaire uses a binary response format, with participants indicating *no* or *yes* to each symptom. The total score on the IADQ ranges from 0 to 8 and is obtained by summing the values assigned to each of the 8 binary questions. The symptoms assessed by the IADQ are *preoccupation* (question 1), *tolerance* (question 2), *unsuccessful efforts to limit or stop Internet usage* (question 3), *withdrawal* (question 4), *loss of control of time spent on the Internet* (question 5), *risk/lose relationships or opportunities* (question 6), *lies to conceal the extent of involvement* (question 7), and *dysfunctional coping* (question 8) [54]. Participants who answered *yes* to ≥5 symptoms were categorized as addicted to the internet (dependent internet users), whereas the others were categorized as nondependent internet users. Previous studies on the IADQ have reported Cronbach α values ranging from 0.60 to 0.72 [55]. The reliability of the IADQ in this study was deemed acceptable, with a Cronbach α value of 0.66 [56] and a McDonald ω value of 0.67. In addition, the research conducted by Lozano-Blasco et al [57] found that using lesser answer categories led to decreased Cronbach α values. The dichotomous character of the questions in the IADQ, which means that they only allow for replies of *yes* or *no*, leads to lower Cronbach α values that are within the acceptable range for dependability. To establish content validity [58], the initial item pool was reviewed by a panel of 3 experts who has extensive experience working with adolescents in the Arab region. This process resulted a questionnaire that comprehensively covered the construct of interest.

Essential and Nonessential Internet Use

Adolescents were presented with a set of questions regarding their use of digital technology for both essential and nonessential purposes on both weekends and weekdays. The weekend is officially observed on Fridays and Saturdays in Qatar. By focusing on weekdays and weekends separately, we aimed to capture any behavioral variations influenced by the shift from weekend to workweek and vice versa. Participants were asked the following questions to encompass four different situations: (1) How many hours do you use digital technology for study purposes daily on weekdays (Sunday-Thursday)? (2) How many hours do you use digital technology for nonessential reasons daily on weekdays (Sunday-Thursday)? (3) How many hours do you use digital technology for study purposes daily on weekends (Friday and Saturday)? (4) How many hours do you use digital technology for nonessential reasons daily on weekends (Friday and Saturday)?

Essential internet use on weekdays and weekends was combined, and their average was used to compute essential internet use time. Nonessential internet use on weekdays and weekends was combined, and their average was taken to compute nonessential internet use time.

Subjective Happiness With Nonessential Internet Use Time

Adolescents were asked to express their feelings regarding the extent of their nonessential use of digital technology. Their responses were collected using a five-point Likert scale: (1) "I am happy with it," (2) "I am somewhat happy with it," (3) "Neither happy nor unhappy with it," (4) "I am somewhat unhappy with it," and (5) "I am unhappy with it."

Statistical Analysis

Multiple linear regression analyses were conducted separately for early and middle adolescents to examine the relationship between nonessential internet use, essential internet use, and subjective happiness with nonessential internet use time and their IA status. The significance level for the statistical tests was set at $P<.05$. All statistical analyses were carried out using JASP (version 0.17.1) [59].

Exploratory factor analysis with oblique (promax) rotation was performed [60]. The determination of the number of significant components to retain for rotation was based on two criteria: (1) the scree plot indicating the number of extracted factors and (2) ensuring that the factor solution allowed for a coherent interpretation of the results.

Results

Descriptive Statistics

The descriptive statistics of the participants are presented in Table 1.

XSL•FO
RenderX

Table 1. Descriptive statistics—internet use patterns and internet addiction (IA).

Variable	Early adolescents (n=242)	Middle adolescents (n=135)
Essential internet use time (h), mean (SD)	2.22 (1.76)	2.96 (2.12)
Nonessential internet use time (h), mean (SD)	4.45 (2.74)	4.52 (2.35)
Subjective happiness with nonessential internet use time, n (%)		
"I am happy with it" (rating of 1)	90 (37.2)	44 (32.6)
"I am somewhat happy with it" (rating of 2)	77 (31.8)	40 (29.6)
"Neither happy nor unhappy with it" (rating of 3)	46 (19)	35 (25.9)
"I am somewhat unhappy with it" (rating of 4)	17 (7)	11 (8.1)
"I am unhappy with it" (rating of 5)	12 (5)	5 (3.7)
IA		
Total IA score (0-8), mean (SD)	3.26 (1.98)	3.47 (2.17)
IA prevalence, n (%)		
Addicted internet users	63 (26)	50 (37)
Nonaddicted internet users	179 (74)	85 (63)

Exploratory Factor Analysis of IA Symptoms

To identify underlying dimensions within the IA symptoms and reduce the data into meaningful factor scores, an exploratory factor analysis was conducted on the dataset. The Kaiser-Meyer-Olkin (KMO) test, a measure of sampling adequacy, was used in this study to evaluate multicollinearity in the data to determine the feasibility of conducting a factor analysis. The overall KMO measure was 0.76, denoting "middling" adequacy. Individual KMO measures were all of >0.68, ranging from the "mediocre" to the "marvelous" categories according to the classification system by Kaiser and Rice [61]. The Bartlett test of sphericity was statistically significant (P<.001), indicating that the data were likely factorizable.

The analysis extracted 2 components, which accounted for 26% of the variance. Our scree plot was compatible with the 2-factor model. We interpreted the 2 factors as representing internal IA symptoms (factor 1; 18% of the total variance)—those symptoms primarily related to the individual's self, such as preoccupation with internet use—and external IA symptoms (factor 2; 9% of the total variance)—those symptoms primarily related to interaction with others. Factor loadings for factor 1 and factor 2 ranged from 0.31 to 0.55 and from 0.33 to 0.78, respectively. Factor scores were derived by adding the scores of individual items within each empirical domain (sum score) and dividing the sum scores by the total number of items (mean item score). The component loadings and uniqueness of the rotated solutions are presented in Table 2.

Table 2. Factor analysis of internet addiction symptoms.

Item	Factor 1 loadings	Factor 2 loadings	Uniqueness
"Preoccupation" (question 1)	0.49	—[a]	0.79
"Tolerance" (question 2)	0.31	—	0.91
"Unsuccessful efforts to limit or stop Internet usage" (question 3)	0.50	—	0.76
"Withdrawal" (question 4)	0.55	—	0.62
"Loss of control of time spent on the Internet" (question 5)	0.50	—	0.77
"Risk/lose relationships or opportunities" (question 6)	—	0.78	0.48
"Lies to conceal the extent of involvement" (question 7)	—	0.33	0.75
"Dysfunctional coping" (question 8)	0.45	—	0.79

[a]The factor does not comprise this item.

Early Adolescents

Predictors of Total IA Scores

Overview

Multiple regression analysis was used to determine factors that predicted early adolescents' total IA scores and internal IA symptoms. No outliers were observed in the data. The Pearson correlation was also used to analyze the associations between total IA, which was the dependent variable, and the independent variables of subjective happiness with nonessential internet use time, essential internet use time, and nonessential internet use time.

XSL•FO
RenderX

Correlation Analysis of Variables

The correlations between the variables are presented in Table 3.

Table 3. Pearson correlation table between the variables in early adolescents.

	Total IA[a]	Internal IA symptoms	External IA symptoms	Subjective happiness with nonessential internet use time	Essential internet use time	Nonessential internet use time
Total IA	1	—[b]	—	—	—	—
Internal IA symptoms	0.95[c]	1	—	—	—	—
External IA symptoms	0.58[c]	0.30[c]	1	—	—	—
Subjective happiness with nonessential internet use time	0.04	0.03	0.03	1	—	—
Essential internet use time	0.07	0.04	0.10	0.04	1	—
Nonessential internet use time	0.41[c]	0.41[c]	0.18[d]	0.07	0.12	1

[a]IA: internet addiction.

[b]Not applicable.

[c]$P<.001$.

[d]$P<.01$.

Regression Analysis of Total IA Scores

Table 4 shows the results of the multiple regression that was run to predict the total IA score from the subjective happiness with nonessential internet use time, essential internet use time, and nonessential internet use time among early adolescents. The linearity of the sample was assessed by examining the residuals versus predicted plots, whereas residuals versus dependent plots were used to evaluate homoscedasticity in the data. All the assumptions of linearity, normality, homoscedasticity, and multicollinearity were satisfied. Residuals were independent, as assessed via a Durbin-Watson statistic of 1.93. There was no evidence of multicollinearity, as assessed via tolerance values of >0.1. The assumption of normality was met, as assessed using a *Q*-*Q* plot. The multiple regression model significantly predicted the total IA score ($F_{3, 238}$=15.72; $P<.001$; R^2=0.17; adjusted R^2=0.15). Within the model, nonessential internet use time (β=.40; $P<.001$) was the only significant predictor of early adolescents' total IA scores.

Table 4. Multiple linear regression analysis predicting internet addiction in early adolescents[a].

Predictor	Standardized	*t* test (*df*)	*P* value
Constant	—[b]	5.73 (3)	<.001
Subjective happiness with nonessential internet use time	.00923	0.16 (3)	.88
Essential internet use time	.02	0.36 (3)	.72
Nonessential internet use time	.40	6.74 (3)	<.001

[a]R^2=0.17; adjusted R^2=0.15; $F_{3, 238}$=15.72.

[b]Not applicable.

Predictors of Internal and External IA Symptoms

Multiple regression analysis was used to determine factors that predicted internal IA symptoms with the independent variables of subjective happiness with nonessential internet use time, essential internet use time, and nonessential internet use time, and an ordinal logistic regression model was used to predict external IA symptoms. Table 5 shows the results of the multiple regression that was run to predict the internal IA symptom score from the subjective happiness with nonessential internet use time, essential internet use time, and nonessential internet use time. The linearity of the sample was assessed by examining the residuals versus predicted plots, whereas residuals versus dependent plots were used to evaluate homoscedasticity in the data. All the assumptions of linearity, normality, homoscedasticity, and multicollinearity were satisfied. Residuals were independent, as assessed via a Durbin-Watson statistic of 1.98. There was no evidence of multicollinearity, as assessed via tolerance values of >0.1. The assumption of normality was met, as assessed using a *Q*-*Q* plot. The multiple regression model significantly predicted the internal IA symptom score ($F_{3, 238}$=15.95; $P<.001$; R^2=0.17; adjusted R^2=0.16). Within the model, nonessential internet use time (β=.41; $P<.001$) was the only significant predictor of early adolescents' internal IA symptoms.

XSL•FO
RenderX

Table 5. Multiple regression analysis predicting internal internet addiction symptoms in early adolescents[a].

Predictor	Standardized	*t* test (*df*)	*P* value
Constant	—[b]	6.18 (3)	<.001
Subjective happiness with nonessential internet use time	.00318	0.05 (3)	.96
Essential internet use time	−.00431	−0.07 (3)	.94
Nonessential internet use time	.41	6.86 (3)	<.001

[a]R^2=0.17; adjusted R^2=0.16; $F_{3, 238}$=15.95.

[b]Not applicable.

An ordinal logistic regression analysis was conducted to examine the relationship between the 3 predictor variables (ie, subjective happiness with nonessential internet use time, essential internet use time, and nonessential internet use time) and the outcome variable (external IA symptoms; Table 6). The variables included in the ordinal logistic regression analysis met the assumptions of the statistical model—the dependent variable (ie, external IA symptoms) was treated as an ordinal variable, and the independent variables (ie, subjective happiness with nonessential internet use time, essential internet use time, and nonessential internet use time) were treated as continuous variables. The assumption of no multicollinearity was assessed by examining the variance inflation factors (VIFs), and the result (VIFs<1.1) indicated no multicollinearity. The results revealed that nonessential internet use time was associated with increased external IA symptoms (odds ratio [OR] 1.13, 95% CI 1.02-1.24; *P*=.02). Meanwhile, essential internet use time was not associated with any change in external IA symptoms (OR 1.11, 95% CI 0.96-1.28; *P*=.17). Similarly, subjective happiness with nonessential internet use time was not associated with any change in external IA symptoms (OR 0.99, 95% CI 0.78-1.25; *P*=.94) in early adolescents.

Table 6. Ordinal logistic regression predicting external internet addiction symptoms in early adolescents.

Predictor	OR[a] (95% CI; SE)	*P* value
Subjective happiness with nonessential internet use time	0.99 (0.78-1.25; −0.01)	.94
Essential internet use time	1.11 (0.96-1.28; 0.10)	.17
Nonessential internet use time	1.13 (1.02-1.24; 0.12)	.02

[a]OR: odds ratio.

Binary logistic regression was conducted to examine the relationship between subjective happiness with nonessential internet use time and the likelihood of being a dependent internet user (addicted) compared to a nondependent internet user (unaddicted) in early adolescents. The reference category, representing nondependent internet users, was coded as class 1. The OR of 0.93 (95% CI 0.73-1.20) indicated the effect size. However, the results did not reveal a significant association between subjective happiness with nonessential internet use time and IA (*P*=.58).

Middle Adolescents

Predictors of Total IA Scores

Overview

Multiple regression analysis was used to determine the factors that predicted middle adolescents' total IA scores and internal IA symptoms. No outliers were observed in the data. The Pearson correlation was also used to analyze the associations between total IA, which was the dependent variable, and the independent variables of subjective happiness with nonessential internet use time, essential internet use time, and nonessential internet use time.

Correlation Analysis of Variables

The correlations between the variables are presented in Table 7.

XSL•FO
RenderX

Table 7. Pearson correlation table between the variables—middle adolescents.

	Total IA[a]	Internal IA symptoms	External IA symptoms	Subjective happiness with nonessential internet use time	Essential internet use time	Nonessential internet use time
Total IA	1	—[b]	—	—	—	—
Internal IA symptoms	0.95[c]	1	—	—	—	—
External IA symptoms	0.66[c]	0.39[c]	1	—	—	—
Subjective happiness with nonessential internet use time	0.34[c]	0.28[d]	0.32[c]	1	—	—
Essential internet use time	0.21[e]	0.15	0.26[d]	0.18[e]	1	—
Nonessential internet use time	0.35[c]	0.33[c]	0.24[d]	0.13	0.05	1

[a]IA: internet addiction.

[b]Not applicable.

[c]$P<.001$.

[d]$P<.01$.

[e]$P<.05$.

Regression Analysis of Total IA Scores

Table 8 shows the results of the multiple regression that was run to predict middle adolescents' total IA score from the subjective happiness with nonessential internet use time, essential internet use time, and nonessential internet use time. All the assumptions of linearity, normality, homoscedasticity, and multicollinearity were satisfied. The linearity of the sample was assessed by examining the residuals versus predicted plots, whereas residuals versus dependent plots were used to evaluate homoscedasticity in the data. Residuals were independent, as assessed via a Durbin-Watson statistic of 1.83. There was no evidence of multicollinearity, as assessed via tolerance values of >0.1. The assumption of normality was met, as assessed using a *Q-Q* plot. The multiple regression model significantly predicted the total IA score ($F_{3, 131}=12.96$; $P<.001$; $R^2=0.23$; adjusted $R^2=0.21$). Within the model, nonessential internet use time (β=.31; $P<.001$) and subjective happiness with nonessential internet use time (β=.27; $P<.001$) were significant predictors of middle adolescents' total IA score.

Table 8. Multiple linear regression analysis predicting internet addiction in middle adolescents[a].

Predictor	Standardized	*t* test (*df*)	*P* value
Constant	—[b]	1.16 (3)	.25
Subjective happiness with nonessential internet use time	.27	3.44 (3)	<.001
Essential internet use time	.15	1.90 (3)	.06
Nonessential internet use time	.31	3.94 (3)	<.001

[a]$R^2=0.23$; adjusted $R^2=0.21$; $F_{3, 131}=12.96$.

[b]Not applicable.

Predictors of Internal and External IA Symptoms

Multiple regression analysis was used to determine factors that predicted internal IA symptoms with the independent variables of subjective happiness with nonessential internet use time, essential internet use time, and nonessential internet use time, and ordinal logistic regression was used to predict external IA symptoms. Table 9 presents the results of the multiple regression. The linearity of the sample was assessed by examining the residuals versus predicted plots, whereas residuals versus dependent plots were used to evaluate homoscedasticity in the data. All the assumptions of linearity, normality, homoscedasticity, and multicollinearity were satisfied. Residuals were independent, as assessed via a Durbin-Watson statistic of 1.82. There was no evidence of multicollinearity, as assessed via tolerance values of >0.1. The assumption of normality was met, as assessed using a *Q-Q* plot. The multiple regression model significantly predicted the internal IA symptom score ($F_{3, 131}=9.15$; $P<.001$; $R^2=0.17$; adjusted $R^2=0.15$). Within the model, nonessential internet use time (β=.29; $P<.001$) and subjective happiness with nonessential internet use time (β=.22; $P=.007$) were significant predictors of middle adolescents' internal IA symptoms.

XSL•FO
RenderX

Table 9. Multiple regression analysis predicting internal internet addiction symptoms in middle adolescents[a].

Predictor	Standardized	*t* test (*df*)	*P* value
Constant	—[b]	2.10 (3)	.04
Subjective happiness with nonessential internet use time	.22	2.76 (3)	.007
Essential internet use time	.10	1.20 (3)	.23
Nonessential internet use time	.29	3.65 (3)	<.001

[a]R^2=0.17; adjusted R^2=0.15; $F_{3, 238}$=15.95.

[b]Not applicable.

An ordinal logistic regression analysis was conducted to examine the relationship between the 3 predictor variables (ie, subjective happiness with nonessential internet use time, essential internet use time, and nonessential internet use time) and the outcome variable (external IA symptoms; Table 10). The variables included in the ordinal logistic regression analysis met the assumptions of the statistical model—the dependent variable (ie, external IA symptoms) was treated as an ordinal variable, and the independent variables (ie, subjective happiness with nonessential internet use time, essential internet use time, and nonessential internet use time) were treated as continuous variables. The assumption of no multicollinearity was assessed by examining the VIFs, and the result (VIFs<1.1) indicated no multicollinearity. The results revealed that all 3 predictors were associated with changes in the external IA symptoms. Nonessential internet use time was associated with an increase in external IA symptoms (OR 1.19, 95% CI 1.03-1.39; *P*=.02). Essential internet use time was associated with an increase in external IA symptoms (OR 1.24, 95% CI 1.05-1.46; *P*=.01). Similarly, a higher unhappiness with nonessential internet use time was associated with increased external IA symptoms (OR 1.75, 95% CI 1.28-2.44; *P*<.001).

Table 10. Ordinal logistic regression predicting external internet addiction symptoms in middle adolescents.

Predictor	OR[a] (95% CI; SE)	*P* value
Subjective happiness with nonessential internet use time	1.75 (1.28-2.44; 0.56)	<.001
Essential internet use time	1.24 (1.05-1.46; 0.21)	.02
Nonessential internet use time	1.19 (1.03-1.39; 0.18)	.01

[a]OR: odds ratio.

A binary logistic regression was performed to assess the impact of subjective happiness with nonessential internet use time on the likelihood of being a dependent internet user (addicted) versus a nondependent internet user (unaddicted). The reference category was coded as class 1, representing nondependent internet users. The OR of 0.66 (95% CI 0.47-0.92) indicated the effect size. The *F*-measure was 75%. The results indicated a significant association between subjective happiness with nonessential internet use time and IA (*P*=.01), suggesting that higher levels of unhappiness were associated with a decreased likelihood of being unaddicted to the internet.

Discussion

Principal Findings

This study aimed to enhance our understanding of adolescent IA by addressing 3 key RQs. First, we examined the significance of differentiating between essential and nonessential internet use in relation to IA, offering a more nuanced view of IA beyond that of total screen time that is found in previous research. Second, we investigated whether subjective happiness linked to time spent on nonessential internet use could predict IA in adolescents. Finally, by studying early and middle adolescents separately, we identified age-specific differences in how time spent on essential and nonessential internet use and the related happiness contribute to IA.

Consistent with the extensive existing literature on IA, our study demonstrated that excessive engagement in nonessential internet activities, such as social media and online gaming, positively correlate with IA overall [62-64]. Furthermore, in our study, essential internet use did not exhibit a significant correlation with IA (*P*=.94 for early adolescents and *P*=.06 for middle adolescents), suggesting that essential internet activities such as remote learning or using the internet for school-required work may not carry the same level of addiction risk. This is similar to the results obtained in other studies [65,66]; Salubi et al [65] found that there was no correlation between essential internet use and IA among 390 surveyed university students in South Africa, whereas Pjevac et al [66] found that IA was lowest among adolescents who mostly used the internet for school purposes. This distinction of the purposes of internet use enables the development of targeted interventions and strategies to effectively address this issue. As digital technology becomes increasingly integral to tasks that once operated without it, acknowledging this technological shift is essential in our evolving digital society. It is necessary to recognize the evolving landscape of digital technology use and interpret digital technology use with caution, as emphasized by Squire and Steinkuehler [67], before delving into the complexities of IA. This caution entails considering the specific activities carried out on the internet, the motivations driving them, and the duration for which they occur.

XSL•FO
RenderX

Our study highlights similarities and key differences in IA between early and middle adolescents as our results revealed that, among early adolescents, nonessential internet use significantly predicted IA and its internal symptoms (related to the self). However, middle adolescents exhibited a more complex relationship with IA as both nonessential and essential internet use, in addition to unhappiness with nonessential internet activities, significantly predicted IA and its external symptoms (related to others).

Our results suggest that early adolescents are more likely to develop IA and experience internal IA symptoms as a result of nonessential internet use, whereas this influence seems to decrease as they get older. This finding is in line with the results of a meta-analysis (of 20 studies) on IA in adolescents reporting that IA is inversely proportional to age [57]. This difference suggests that the impact of nonessential digital technology use on IA may vary according to the developmental stage of adolescents. One possible explanation is that early adolescents are actively navigating their new identity and seeking peer acceptance, which can make them more susceptible to excessive use of digital technology [68]. The pressure to conform to peers' behaviors and expectations may override their ability to exercise self-control and resist the allure of digital technology [69]. Furthermore, in early adolescents, only nonessential internet use increased the likelihood of external IA symptoms. Our results are similar to those of a recent study that examined sociodemographic factors related to IA among a sample of 1664 adolescents in Serbia. The study reported the lowest IA rates among the younger adolescents and among those who used the internet for the purpose of schoolwork [66]. Our results could be interpreted as early adolescents not facing the same academic pressures and responsibilities that drive older adolescents to the excessive use of the internet for essential activities. Instead, early adolescent IA is more related to recreational online activities, underscoring the importance of tailoring preventive interventions to focus on nonessential use, such as social media and gaming, to equip early adolescents to develop healthier digital habits. Parental involvement and supervision play a vital role in preventing IA in early adolescents. Actively engaging with children, establishing open communication channels, and monitoring their digital technology use can create a supportive environment and set boundaries for healthy digital behaviors [27,70]. Emphasizing balance is also essential, highlighting the need for a well-rounded lifestyle that includes academic tasks, physical activities, face-to-face interactions, and nondigital recreational pursuits [71]. Self-control in early adolescents may be enhanced through self-control–training activities and participation in self-control–promoting programs [72].

In contrast to the results for early adolescents, our results demonstrated a more complex relationship between internet use and IA in middle adolescents. Our study found that, while nonessential internet use remains a significant predictor of IA, subjective happiness with nonessential internet activities also plays a role. Specifically, the unhappier middle adolescents were with the time spent on nonessential internet use, the more likely they were to experience IA. These findings are consistent with the findings of a similar study that surveyed a sample of adolescents and young adults (aged 15-24 years) and adults (aged 25-64 years), where dissatisfaction with smartphone use was linked to higher problematic internet use [73]. Our results contribute to the literature as there is a lack of research examining subjective happiness with digital technology use in relation to IA in adolescents.

This can be attributed to the increased academic pressure and the need to manage uncertainty, which drive middle adolescents to spend more time on essential internet activities, such as online educational tasks. The combination of information overload, academic demands, and the desire to avoid having to deal with uncertainty may lead to excessive internet use, further exacerbating IA.

Moreover, middle adolescents exhibited a broader range of factors influencing their external IA symptoms.

While, in early adolescents, only nonessential internet use increased the likelihood of external IA symptoms, in middle adolescents, both nonessential and essential internet use, as well as unhappiness with the time spent on nonessential internet use, increased external IA symptoms. This can be attributed to the intensified academic pressure in combination with information overload, leading them to neglect other aspects of their lives and driving them to the excessive use of the internet to meet academic demands [74,75]. The combination of these factors heightens the risk of IA [76-78]. Furthermore, these academic pressures can further impact social interactions as middle adolescents prioritize academic performance, leading them to show signs of stress in their engagement with others [74,79], manifesting in external symptoms of IA, as observed in our study. Recognizing the unique challenges faced by adolescents at different developmental stages is crucial for effectively tailoring preventive measures and interventions.

While our study focused on examining individual factors related to IA, considering factors related to the specific context in which our results were obtained should not be overlooked and could provide a more holistic understanding of IA among adolescents. Adolescents' internet use and their susceptibility to IA are influenced by cultural values, societal norms, and regional habits. For instance, in the Middle East, where our study was conducted, cultural expectations regarding academic achievement and family roles may contribute to heightened stress and pressure on adolescents [80,81], leading them to spend more time on online activities such as social media and gaming. The emphasis on academic success in our context combined with limited outdoor recreational opportunities due to the extremely hot climate might be driving adolescents to spend more time indoors and engage in online recreational activities, increasing their risk of IA.

In addition, the rapid digital transformation in the Middle East has created new challenges and opportunities for adolescents [82,83]. The increased availability of smartphones and internet access as well as a growing digital culture have made it easier for people to spend long hours on the web. While there are positive aspects of digitalization, it also raises concerns about the potential for excessive use and the development of IA. The addition of environmental factors such as the local affluent culture of convenience, the wide availability of digital devices even at very young ages, and the wide presence of high-speed

XSL•FO
RenderX

internet further influences how adolescents interact with digital technology.

Our study has important implications for the field of IA and for promoting healthier digital technology use habits in adolescents. The findings suggest that interventions aimed at addressing IA should focus on nonessential use specifically, particularly in early adolescents, rather than considering internet use as a whole. Because IA is a complex outcome influenced by multiple physical, psychological, and technological factors rather than a single issue, this targeted approach can help effectively address the factors contributing to IA while considering the developmental factors of adolescents. While there is a dearth of data from the region, there is evidence from a systematic literature review on policy and prevention approaches for disordered gaming and internet use suggesting that, due to the global prevalence of the phenomena and the geographically dispersed nature of preventative programs, it is imperative that prevention efforts be integrated across national borders while taking cultural differences into consideration [84].

The following recommendations are a recapitulation of the evidence produced by the limited international reviews that have examined the effectiveness of programs and interventions in the prevention of IA [71,84-87].

Early identification and intervention are key to addressing IA among adolescents. Schools can implement evidence-based prevention programs specifically targeting IA. These programs can incorporate educational components, skill-building exercises, and peer support networks to help adolescents develop healthy online habits and balance their time between essential and nonessential activities [87]. Comprehensive school-based programs are essential to effectively address IA. These programs should focus on increasing awareness, develop coping strategies, and promoting responsible digital technology use. These programs can be integrated into the school curricula and should include methods to improve self-control and enhance knowledge of the underlying processes that contribute to excessive online activity, such as constant notifications [88,89]. In addition, schools can implement practical digital well-being initiatives, such as mindfulness exercises and the integration of technology breaks into the daily schedule, to promote healthier digital habits. Moreover, promoting digital literacy and responsible digital technology use is crucial for educating adolescents on the potential risks and consequences associated with excessive digital technology use [90]. Given that adolescents' behaviors are largely shaped by other external factors such as their families and external environments [21], schools play a crucial role in extending these preventive programs to parents and caregivers. Understanding children's intentions when interacting with the internet is crucial and emphasizes the need for smarter digital parenting tools that provide enhanced monitoring and guidance. A healthier digital environment can be created by considering intentionality and using smart tools. Devices and programs should be designed to isolate essential and nonessential uses, advising children to avoid installing or visiting purely nonessential social media applications and sites on the same device in which they engage in essential activities.

These recommendations must be supported by designing and enacting policies. Policy makers can design programs targeting the broader environment in which adolescents live, addressing the underlying factors contributing to maladaptive coping, such as reducing academic pressure and improving family support. By reducing dependence on nonessential internet use, these policies, in turn, can empower adolescents to navigate their digital environments more responsibly.

Certain limitations of this study must be considered when interpreting the results. First, due to the cross-sectional design of our study, the generalizability of our findings is limited. Another limitation of our study was that the responses provided by the adolescents were based on self-report, potentially introducing reporting bias. It is possible that the adolescents did not accurately report the time spent on essential and nonessential activities and symptoms of IA. To mitigate the influence of reporting bias, future studies should incorporate objective measures, such as using data on actual screen time, instead of relying solely on self-reported information. Another limitation of this study is that the context of essential and nonessential internet use could have been interpreted by adolescents based on their subjective experiences and perceptions. For example, spending time with friends on social media during the examination period may be considered essential for adolescents who want to receive moral support from their colleagues. Nevertheless, our study left it at the discretion of the participants. The impact of the specific activity or content that adolescents consume or spend excessive time on and the effect it has on them should be scrutinized in more detail and is an area that requires attention. Our sample's gender imbalance, which was characterized by a higher proportion of female participants, may restrict the generalizability of the results to both genders. Furthermore, we adhered to sample size estimates using the method by Green [49] to guarantee sufficient statistical power for our study. The deliberate choice not to segregate the sample by gender was made to ensure that the sample size remained sufficient for obtaining consistent and dependable regression results. The male sample was relatively too small to allow for that option. Although the existing literature on gender differences in IA is inconclusive, with research presenting inconsistent findings [91-95], it is important to exercise caution when interpreting the results of this study due to the higher proportion of female participants. Future studies should aim for a more even gender distribution to comprehensively investigate any disparities in IA between male and female individuals.

The strength of our study lies in its inclusion of a non-WEIRD population, specifically focusing on adolescent IA in the Middle East, where limited research on this area has been conducted. Another strength of our study is that it addresses the research gap by examining both essential and nonessential internet use in the same sample. While previous research has explored the different online activities that contribute to IA [65,96], few studies have specifically investigated the impact of essential and nonessential internet use within a single study. Another strength of our study lies in the separate analysis of IA factors in early and middle adolescents, which allowed us to capture the unique dynamics and influences shaping IA in each age

XSL•FO
RenderX

group, thus enhancing the understanding of IA in relation to developmental stages.

Conclusions

In conclusion, this study sheds light on the complex relationship between digital technology use and IA among early and middle adolescents. For many years, screen time has been widely used as a prominent measure to assess the extent of IA among adolescents. As the digital society continues to evolve rapidly, relying on screen time as a stand-alone criterion fails to provide a comprehensive understanding of phenomena such as IA. Therefore, it is imperative to distinguish between essential and nonessential digital technology use, acknowledging diverse activities and their implications for a better understanding of individuals' engagement with digital technology. The findings of our study provide valuable insights for researchers, practitioners, and policy makers in addressing IA among adolescents. By recognizing the distinct nature of internet activities and considering developmental factors, effective interventions and strategies can be developed to promote healthy digital technology use and prevent IA in adolescents.

Acknowledgments

"Technology Overuse Amongst Adolescents in Qatar: current status, challenges, and recommendations for policy" is a joint project among 3 members of the Qatar Foundation for Education, Science, and Community Development; the Doha International Family Institute; the World Innovation Summit for Education; and the Word Innovation Summit for Health. The statements made herein are solely the responsibility of the authors and do not reflect the statements or opinions of these institutions.

Authors' Contributions

KC conceptualized the study, prepared and analyzed the data, and wrote the paper. MA helped with data preparation and verified the statistical analysis. AA, SA-H, and AB participated in the study conceptualization and design and conducted the data collection. DAH critically reviewed and revised the Conclusions section in the paper. RA participated in all stages and supervised the research. All authors reviewed the paper and provided feedback. All authors approved the final version of the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Survey design.
[DOCX File , 608 KB - humanfactors_v12i1e62955_app1.docx]

References

1. Vogels EA, Gelles-Watnick R, Massarat N. Teens, social media and technology 2022. Pew Research Center. 2022 Aug 10. URL: https://www.pewresearch.org/internet/2022/08/10/teens-social-media-and-technology-2022/ [accessed 2023-05-31]
2. Balhara YP, Mahapatra A, Sharma P, Bhargava R. Problematic internet use among students in South-East Asia: current state of evidence. Indian J Public Health 2018;62(3):197-210 [FREE Full text] [doi: 10.4103/ijph.IJPH_288_17] [Medline: 30232969]
3. Chung TW, Sum SM, Chan MW. Adolescent internet addiction in Hong Kong: prevalence, psychosocial correlates, and prevention. J Adolesc Health 2019 Jun;64(6S):S34-S43. [doi: 10.1016/j.jadohealth.2018.12.016] [Medline: 31122547]
4. Lee JY, Kim SY, Bae KY, Kim JM, Shin IS, Yoon JS, et al. Prevalence and risk factors for problematic internet use among rural adolescents in Korea. Asia Pac Psychiatry 2018 Jun;10(2):e12310. [doi: 10.1111/appy.12310] [Medline: 29356370]
5. Xin M, Xing J, Pengfei W, Houru L, Mengcheng W, Hong Z. Online activities, prevalence of internet addiction and risk factors related to family and school among adolescents in China. Addict Behav Rep 2018 Jun;7:14-18 [FREE Full text] [doi: 10.1016/j.abrep.2017.10.003] [Medline: 29450251]
6. Oh WO. Factors influencing internet addiction tendency among middle school students in Gyeong-buk area. Taehan Kanho Hakhoe Chi 2003 Dec;33(8):1135-1144. [doi: 10.4040/jkan.2003.33.8.1135] [Medline: 15314360]
7. Morrison CM, Gore H. The relationship between excessive internet use and depression: a questionnaire-based study of 1,319 young people and adults. Psychopathology 2010;43(2):121-126. [doi: 10.1159/000277001] [Medline: 20110764]
8. Singh N, Barmola KC. Internet addiction, mental health and academic performance of school students/adolescents. Int J Indian Psychol 2015 Jun 25;2(3). [doi: 10.25215/0203.053]
9. Tateno M, Teo AR, Ukai W, Kanazawa J, Katsuki R, Kubo H, et al. Internet addiction, smartphone addiction, and hikikomori trait in Japanese young adult: social isolation and social network. Front Psychiatry 2019;10:455 [FREE Full text] [doi: 10.3389/fpsyt.2019.00455] [Medline: 31354537]
10. Young KS. Internet addiction: the emergence of a new clinical disorder. Cyberpsychol Behav 1998 Jan;1(3):237-244. [doi: 10.1089/cpb.1998.1.237]

XSL•FO
RenderX

11. Shaw M, Black DW. Internet addiction: definition, assessment, epidemiology and clinical management. CNS Drugs 2008;22(5):353-365. [doi: 10.2165/00023210-200822050-00001] [Medline: 18399706]
12. Young KS, de Abreu CN. Internet Addiction: A Handbook and Guide to Evaluation and Treatment. Hoboken, NJ: John Wiley & Sons; 2007.
13. Mihajlov M, Vejmelka L. Internet addiction: a review of the first twenty years. Psychiatr Danub 2017 Sep 20;29(3):260-272 [FREE Full text] [doi: 10.24869/psyd.2017.260] [Medline: 28949307]
14. Rad MS, Martingano AJ, Ginges J. Toward a psychology of Homo sapiens: making psychological science more representative of the human population. Proc Natl Acad Sci U S A 2018 Nov 06;115(45):11401-11405. [doi: 10.1073/pnas.1721165115] [Medline: 30397114]
15. Schaller M, Norenzayan A, Heine SJ, Yamagishi T, Kameda T. Evolutionary Psychology Meets Cultural Psychology: Review of Evolution, Culture and the Human Mind. East Sussex, UK: Psychology Press; 2009.
16. Wigfield A, Byrnes JP, Eccles JS. Development during early and middle adolescence. In: Alexander PA, Winne PH, editors. Handbook of Educational Psychology. New York, NY: Routledge; 2006.
17. Kemp S. Digital 2023: Qatar. Data Reportal. 2023 Feb 13. URL: https://datareportal.com/reports/digital-2023-qatar [accessed 2023-11-08]
18. World Bank open data. World Bank Group. URL: https://data.worldbank.org/ [accessed 2024-08-10]
19. Saquib J. Internet addiction among Saudi Arabian youth. Int J Health Sci (Qassim) 2020;14(2):1-2 [FREE Full text] [Medline: 32206053]
20. Al-Khani AM, Saquib J, Rajab AM, Khalifa MA, Almazrou A, Saquib N. Internet addiction in Gulf countries: a systematic review and meta-analysis. J Behav Addict 2021 Sep 06;10(3):601-610 [FREE Full text] [doi: 10.1556/2006.2021.00057] [Medline: 34491902]
21. Chemnad K, Aziz M, Abdelmoneium AO, Al-Harahsheh S, Baghdady A, Al Motawaa FY, et al. Adolescents' internet addiction: does it all begin with their environment? Child Adolesc Psychiatry Ment Health 2023 Jul 04;17(1):87 [FREE Full text] [doi: 10.1186/s13034-023-00626-7] [Medline: 37403108]
22. Ko D, Yao M. 11. Internet addiction: an cross-cultural perspective. In: The Psychology of Social Networking Vol.2: Identity and Relationships in Online Communities. Warsaw, Poland: De Gruyter Open Poland; 2016:141-158.
23. Aşut Ö, Abuduxike G, Acar-Vaizoğlu S, Cali S. Relationships between screen time, internet addiction and other lifestyle behaviors with obesity among secondary school students in the Turkish Republic of Northern Cyprus. Turk J Pediatr 2019 Aug 25;61(4):568-579 [FREE Full text] [doi: 10.24953/turkjped.2019.04.014] [Medline: 31990475]
24. Twenge JM, Campbell WK. Associations between screen time and lower psychological well-being among children and adolescents: evidence from a population-based study. Prev Med Rep 2018 Dec;12:271-283 [FREE Full text] [doi: 10.1016/j.pmedr.2018.10.003] [Medline: 30406005]
25. Lissak G. Adverse physiological and psychological effects of screen time on children and adolescents: literature review and case study. Environ Res 2018 Jul;164:149-157. [doi: 10.1016/j.envres.2018.01.015] [Medline: 29499467]
26. Sousa Tavares HL, Souza Costa D, Soares AA, Kestelman I, da Silva AG, Malloy-Diniz LF, et al. Gaming addiction and screen time in a context of increase of internalizing symptoms: moderation evaluation. Clin Child Psychol Psychiatry 2023 Jan 03;28(1):212-223. [doi: 10.1177/13591045221125329] [Medline: 36062317]
27. Chemnad K, Alshakhsi S, Al-Harahsheh S, Abdelmoneium AO, Al-Khalaf MS, Baghdady A, et al. Is it contagious? Does parents' internet addiction impact their adolescents' internet addiction? Soc Sci Comput Rev 2022 Aug 05;41(5):1691-1711. [doi: 10.1177/08944393221117408]
28. Cerniglia L, Zoratto F, Cimino S, Laviola G, Ammaniti M, Adriani W. Internet addiction in adolescence: neurobiological, psychosocial and clinical issues. Neurosci Biobehav Rev 2017 May;76(Pt A):174-184. [doi: 10.1016/j.neubiorev.2016.12.024] [Medline: 28027952]
29. Yen JY, Ko CH, Yen CF, Wu HY, Yang MJ. The comorbid psychiatric symptoms of internet addiction: attention deficit and hyperactivity disorder (ADHD), depression, social phobia, and hostility. J Adolesc Health 2007 Jul;41(1):93-98. [doi: 10.1016/j.jadohealth.2007.02.002] [Medline: 17577539]
30. Costescu C, Chelba I, Ro an A, Kovari A, Katona J. Cognitive patterns and coping mechanisms in the context of internet use. Appl Sci 2021 Feb 01;11(3):1302. [doi: 10.3390/app11031302]
31. Chou WJ, Huang MF, Chang YP, Chen YM, Hu HF, Yen CF. Social skills deficits and their association with internet addiction and activities in adolescents with attention-deficit/hyperactivity disorder. J Behav Addict 2017 Mar 01;6(1):42-50 [FREE Full text] [doi: 10.1556/2006.6.2017.005] [Medline: 28245666]
32. Ciccia AH, Meulenbroek P, Turkstra LS. Adolescent brain and cognitive developments: implications for clinical assessment in traumatic brain injury. Top Lang Disord 2009;29(3):249-265 [FREE Full text] [doi: 10.1097/TLD.0b013e3181b53211] [Medline: 30220763]
33. Shah JY, Gardner WL. Handbook of Motivation Science. New York, NY: Guilford Publications; 2008.
34. Koç P. Internet addiction and subjective wellbeing in university students. Journal of Positive Psychology and Wellbeing 2017;1(1):34-41 [FREE Full text]
35. Fauzi MA, Harun SA, Martin T, Paiman N, Hussin MS. Smartphone addiction among young adolescents: integration of flow theory and fear of missing out. Int J Inf Manag Sci 2021;32:39-58 [FREE Full text]

XSL•FO
RenderX

36. Harter S. Emerging self-processes during childhood and adolescence. In: Leary MR, Tangney JP, editors. Handbook of Self and Identity. New York, NY: Guilford Press; 2012:680-715.
37. Plancherel B, Bolognini M, Halfon O. Coping strategies in early and mid-adolescence. Eur Psychol 1998 Sep;3(3):192-201. [doi: 10.1027/1016-9040.3.3.192]
38. Blyth DA, Traeger CM. The self‐concept and self‐esteem of early adolescents. Theory Pract 2009 Nov 05;22(2):91-97. [doi: 10.1080/00405848309543045]
39. Brinthaupt TM, Lipka RP. Understanding Early Adolescent Self and Identity: Applications and Interventions. New York, NY: State University of New York Press; 2002.
40. Scheiner C, Grashoff J, Kleindienst N, Buerger A. Mental disorders at the beginning of adolescence: prevalence estimates in a sample aged 11-14 years. Public Health Pract (Oxf) 2022 Dec;4:100348 [FREE Full text] [doi: 10.1016/j.puhip.2022.100348] [Medline: 36545674]
41. Joshi P, Shukla S. Middle childhood and adolescence: development and learning. In: Child Development and Education in the Twenty-First Century. Singapore, Singapore: Springer; 2019.
42. Umberson D, Crosnoe R, Reczek C. Social relationships and health behavior across life course. Annu Rev Sociol 2010 Aug 01;36(1):139-157 [FREE Full text] [doi: 10.1146/annurev-soc-070308-120011] [Medline: 21921974]
43. Borghuis J, Denissen JJ, Oberski D, Sijtsma K, Meeus WH, Branje S, et al. Big Five personality stability, change, and codevelopment across adolescence and early adulthood. J Pers Soc Psychol 2017 Oct;113(4):641-657 [FREE Full text] [doi: 10.1037/pspp0000138] [Medline: 28253001]
44. Havewala M, Bowker JC, Smith KA, Rose-Krasnor L, Booth-LaForce C, Laursen B, et al. Peer influence during adolescence: the moderating role of parental support. Children (Basel) 2021 Apr 17;8(4):306 [FREE Full text] [doi: 10.3390/children8040306] [Medline: 33920622]
45. Ren Y, Chi X, Bu H, Huang L, Wang S, Zhang Y, et al. Warm and harsh parenting, self-kindness and self-judgment, and well-being: an examination of developmental differences in a large sample of adolescents. Children (Basel) 2023 Feb 19;10(2):406 [FREE Full text] [doi: 10.3390/children10020406] [Medline: 36832535]
46. Dienlin T, Johannes N. The impact of digital technology use on adolescent well-being . Dialogues Clin Neurosci 2022 Apr 01;22(2):135-142. [doi: 10.31887/dcns.2020.22.2/tdienlin]
47. Moreno MA, Binger K, Zhao Q, Eickhoff J. Measuring interests not minutes: development and validation of the Adolescents' Digital Technology Interactions and Importance scale (ADTI). J Med Internet Res 2020 Feb 12;22(2):e16736 [FREE Full text] [doi: 10.2196/16736] [Medline: 32049068]
48. Aziz M, Chemnad K, Al-Harahsheh S, Abdelmoneium AO, Bagdady A, Hassan DA, et al. The influence of adolescents essential and non-essential use of technology and internet addiction on their physical and mental fatigues. Sci Rep 2024 Jan 19;14(1):1745 [FREE Full text] [doi: 10.1038/s41598-024-51655-x] [Medline: 38242916]
49. Green SB. How many subjects does it take to do a regression analysis. Multivariate Behav Res 1991 Jul 01;26(3):499-510. [doi: 10.1207/s15327906mbr2603_7] [Medline: 26776715]
50. Schönbrodt FD, Perugini M. At what sample size do correlations stabilize? J Res Pers 2013 Oct;47(5):609-612. [doi: 10.1016/j.jrp.2013.05.009]
51. Smetana JG, Campione-Barr N, Metzger A. Adolescent development in interpersonal and societal contexts. Annu Rev Psychol 2006 Jan;57(1):255-284. [doi: 10.1146/annurev.psych.57.102904.190124] [Medline: 16318596]
52. Brislin RW. Back-translation for cross-cultural research. J Cross Cult Psychol 1970 Sep 01;1(3):185-216. [doi: 10.1177/135910457000100301]
53. Lyons-Thomas J. Using think aloud protocols in the validity investigation of an assessment of complex thinking. University of British Columbia. 2014. URL: https://open.library.ubc.ca/soa/cIRcle/collections/ubctheses/24/items/1.0167403 [accessed 2025-01-16]
54. Strittmatter E, Brunner R, Fischer G, Parzer P, Resch F, Kaess M. [Association of peer victimization, coping, and pathological internet use among adolescents]. Z Kinder Jugendpsychiatr Psychother 2014 Mar;42(2):85-94. [doi: 10.1024/1422-4917/a000275] [Medline: 24571814]
55. Young K. Internet Addiction Test (IAT). Wood Dale, IL: Stoelting Company; 2016.
56. Ursachi G, Horodnic IA, Zait A. How reliable are measurement scales? External factors with indirect influence on reliability estimators. Procedia Econ Financ 2015;20:679-686. [doi: 10.1016/s2212-5671(15)00123-9]
57. Lozano-Blasco R, Latorre-Martínez M, Cortés-Pascual A. Screen addicts: a meta-analysis of internet addiction in adolescence. Child Youth Serv Rev 2022 Apr;135:106373. [doi: 10.1016/j.childyouth.2022.106373]
58. Rubio DM, Berg-Weger M, Tebb SS, Lee ES, Rauch S. Objectifying content validity: conducting a content validity study in social work research. Soc Work Res 2003 Jun 01;27(2):94-104. [doi: 10.1093/swr/27.2.94]
59. JASP - a fresh way to do statistics. JASP. URL: https://jasp-stats.org/ [accessed 2021-11-29]
60. Cureton EE, Mulaik SA. The weighted varimax rotation and the promax rotation. Psychometrika 1975 Jun;40(2):183-195. [doi: 10.1007/bf02291565]
61. Kaiser HF, Rice J. Little Jiffy, Mark IV. Educ Psychol Meas 1974 Apr 01;34(1):111-117. [doi: 10.1177/001316447403400115]

XSL•FO
RenderX

62. Ciacchini R, Orrù G, Cucurnia E, Sabbatini S, Scafuto F, Lazzarelli A, et al. Social media in adolescents: a retrospective correlational study on addiction. Children (Basel) 2023 Jan 31;10(2):278 [FREE Full text] [doi: 10.3390/children10020278] [Medline: 36832407]
63. Putra PY, Fithriyah I, Zahra Z. Internet addiction and online gaming disorder in children and adolescents during COVID-19 pandemic: a systematic review. Psychiatry Investig 2023 Mar;20(3):196-204 [FREE Full text] [doi: 10.30773/pi.2021.0311] [Medline: 36990662]
64. Kuss DJ, Kristensen AM, Lopez-Fernandez O. Internet addictions outside of Europe: a systematic literature review. Comput Hum Behav 2021 Feb;115:106621. [doi: 10.1016/j.chb.2020.106621]
65. Salubi O, Muchaonyerwa N, Nekhwevha F. Digital media usage and prevalence of internet addiction among undergraduate students in South Africa. Int J Web Appl 2018 Dec;10(4):127 [FREE Full text] [doi: 10.6025/ijwa/2018/10/4/127-136]
66. Ana P, Teodora S, Mirjana J. Sociodemographic factors related to internet addiction among adolescents in Serbia. Vojnosanitetski Pregled 2024;81(7):438-446. [doi: 10.2298/vsp240307037p]
67. Squire KD, Steinkuehler C. The problem with screen time. Teach Coll Rec 2017 Dec 01;119(12):1-24. [doi: 10.1177/016146811711901207]
68. Caravita SC, Sijtsema JJ, Rambaran JA, Gini G. Peer influences on moral disengagement in late childhood and early adolescence. J Youth Adolesc 2014 Feb 10;43(2):193-207. [doi: 10.1007/s10964-013-9953-1] [Medline: 23660831]
69. Steinberg L, Albert D, Cauffman E, Banich M, Graham S, Woolard J. Age differences in sensation seeking and impulsivity as indexed by behavior and self-report: evidence for a dual systems model. Dev Psychol 2008 Nov;44(6):1764-1778. [doi: 10.1037/a0012955] [Medline: 18999337]
70. Schmid E, Garrels V. Parental involvement and educational success among vulnerable students in vocational education and training. Educ Res 2021 Nov 15;63(4):456-473. [doi: 10.1080/00131881.2021.1988672]
71. Vondráčková P, Gabrhelík R. Prevention of internet addiction: a systematic review. J Behav Addict 2016 Dec;5(4):568-579 [FREE Full text] [doi: 10.1556/2006.5.2016.085] [Medline: 27998173]
72. Allemand M, Job V, Mroczek DK. Self-control development in adolescence predicts love and work in adulthood. J Pers Soc Psychol 2019 Sep;117(3):621-634 [FREE Full text] [doi: 10.1037/pspp0000229] [Medline: 30550330]
73. Alshakhsi S, Chemnad K, Almourad MB, Altuwairiqi M, McAlaney J, Ali R. Problematic internet usage: the impact of objectively recorded and categorized usage time, emotional intelligence components and subjective happiness about usage. Heliyon 2022 Oct;8(10):e11055 [FREE Full text] [doi: 10.1016/j.heliyon.2022.e11055] [Medline: 36281419]
74. Anniko MK, Boersma K, Tillfors M. Sources of stress and worry in the development of stress-related mental health problems: a longitudinal investigation from early- to mid-adolescence. Anxiety Stress Coping 2019 Mar 22;32(2):155-167. [doi: 10.1080/10615806.2018.1549657] [Medline: 30465445]
75. Dissing AS, Jørgensen TB, Gerds TA, Rod NH, Lund R. High perceived stress and social interaction behaviour among young adults. A study based on objective measures of face-to-face and smartphone interactions. PLoS One 2019 Jul 26;14(7):e0218429 [FREE Full text] [doi: 10.1371/journal.pone.0218429] [Medline: 31348787]
76. Feng Y, Ma Y, Zhong Q. The relationship between adolescents' stress and internet addiction: a mediated-moderation model. Front Psychol 2019;10:2248 [FREE Full text] [doi: 10.3389/fpsyg.2019.02248] [Medline: 31636590]
77. Hofstede G. Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations. Thousand Oaks, CA: SAGE Publications; 2001.
78. Wu CY, Lee MB, Liao SC, Chang LR. Risk factors of internet addiction among internet users: an online questionnaire survey. PLoS One 2015 Oct 10;10(10):e0137506 [FREE Full text] [doi: 10.1371/journal.pone.0137506] [Medline: 26462196]
79. Pascoe MC, Hetrick SE, Parker AG. The impact of stress on students in secondary school and higher education. Int J Adolesc Youth 2019 Apr 11;25(1):104-112. [doi: 10.1080/02673843.2019.1596823]
80. Sellami A. Parental influence on student educational expectations: results from the 2012 Qatar education study. Int J High Educ 2019 Jul;8(4):189. [doi: 10.5430/ijhe.v8n4p189]
81. Ali J, Alsakhe H, Ibrahim I, Khattab N, Madeeha M, Shouia M. Accounting for educational expectations and achievement among native and migrant students in Qatar. Int Stud Sociol Educ 2022 Feb 15;33(2):159-178. [doi: 10.1080/09620214.2022.2034518]
82. Melanyina MV, Ruzina EI, Shkvarya LV, Verenikina AY. Digital transformation in the middle east: directions, outcomes and challenges. Economic Studies Journal 2024(4):35-47 [FREE Full text]
83. Marino C, Gini G, Vieno A, Spada MM. A comprehensive meta-analysis on problematic Facebook use. Comput Hum Behav 2018 Jun;83:262-277. [doi: 10.1016/j.chb.2018.02.009]
84. King DL, Delfabbro PH, Doh YY, Wu AM, Kuss DJ, Pallesen S, et al. Policy and prevention approaches for disordered and hazardous gaming and internet use: an international perspective. Prev Sci 2018 Feb 4;19(2):233-249. [doi: 10.1007/s11121-017-0813-1] [Medline: 28677089]
85. Király O, Griffiths MD, King DL, Lee HK, Lee SY, Bányai F, et al. Policy responses to problematic video game use: a systematic review of current measures and future possibilities. J Behav Addict 2018 Sep 01;7(3):503-517 [FREE Full text] [doi: 10.1556/2006.6.2017.050] [Medline: 28859487]
86. Bağatarhan T, Siyez DM. Programs for preventing internet addiction during adolescence: a systematic review. Addicta 2015 May 31;4:243-265. [doi: 10.15805/addicta.2017.4.2.0015]

XSL•FO
RenderX

87. Throuvala MA, Griffiths MD, Rennoldson M, Kuss DJ. School-based prevention for adolescent internet addiction: prevention is the key. A systematic literature review. Curr Neuropharmacol 2019 May 09;17(6):507-525. [doi: 10.2174/1570159x16666180813153806]
88. Gazzaley A, Rosen LD. The Distracted Mind: Ancient Brains in a High-Tech World. Cambridge, MA: MIT Press; 2016.
89. Liao M, Sundar SS. Sound of silence: does muting notifications reduce phone use? Comput Hum Behav 2022 Sep;134:107338. [doi: 10.1016/j.chb.2022.107338]
90. Hyman A, Stewart K, Jamin AM, Novak Lauscher H, Stacy E, Kasten G, et al. Testing a school-based program to promote digital health literacy and healthy lifestyle behaviours in intermediate elementary students: the Learning for Life program. Prev Med Rep 2020 Sep;19:101149 [FREE Full text] [doi: 10.1016/j.pmedr.2020.101149] [Medline: 32670779]
91. Mo PK, Chan VW, Wang X, Lau JT. Gender difference in the association between internet addiction, self-esteem and academic aspirations among adolescents: a structural equation modelling. Comput Educ 2020 Oct;155:103921. [doi: 10.1016/j.compedu.2020.103921]
92. Chi X, Hong X, Chen X. Profiles and sociodemographic correlates of internet addiction in early adolescents in southern China. Addict Behav 2020 Jul;106:106385. [doi: 10.1016/j.addbeh.2020.106385] [Medline: 32187574]
93. Peris M, de la Barrera U, Schoeps K, Montoya-Castilla I. Psychological risk factors that predict social networking and internet addiction in adolescents. Int J Environ Res Public Health 2020 Jun 26;17(12):4598 [FREE Full text] [doi: 10.3390/ijerph17124598] [Medline: 32604834]
94. Li Q, Dai W, Zhong Y, Wang L, Dai B, Liu X. The mediating role of coping styles on impulsivity, behavioral inhibition/approach system, and internet addiction in adolescents from a gender perspective. Front Psychol 2019 Oct 24;10:2402 [FREE Full text] [doi: 10.3389/fpsyg.2019.02402] [Medline: 31708840]
95. Peterka-Bonetta J, Sindermann C, Elhai JD, Montag C. Personality associations with smartphone and internet use disorder: a comparison study including links to impulsivity and social anxiety. Front Public Health 2019 Jun 11;7:127 [FREE Full text] [doi: 10.3389/fpubh.2019.00127] [Medline: 31245341]
96. Seo M, Kang HS, Yom YH. Internet addiction and interpersonal problems in Korean adolescents. Comput Inform Nurs 2009;27(4):226-233. [doi: 10.1097/ncn.0b013e3181a91b3f]

Abbreviations

IA: internet addiction
IADQ: Internet Addiction Diagnostic Questionnaire
KMO: Kaiser-Meyer-Olkin
OR: odds ratio
RQ: research question
VIF: variance inflation factor
WEIRD: Western, educated, industrialized, rich, and democratic

Edited by A Kushniruk; submitted 05.06.24; peer-reviewed by A Rizvi, S Ostovar, T Safiye; comments to author 31.07.24; revised version received 17.09.24; accepted 01.12.24; published 10.02.25.

Please cite as:
Chemnad K, Aziz M, Al- Harahsheh S, Abdelmoneium A, Baghdady A, Alsayed Hassan D, Ali R
Assessing the Relationship Between the Type of Internet Use and Internet Addiction in Early and Middle Adolescents: Cross-Sectional Study From Qatar
JMIR Hum Factors 2025;12:e62955
URL: https://humanfactors.jmir.org/2025/1/e62955
doi:10.2196/62955
PMID:

©Khansa Chemnad, Maryam Aziz, Sanaa Al- Harahsheh, Azza Abdelmoneium, Ahmed Baghdady, Diana Alsayed Hassan, Raian Ali. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 10.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Employee Preference and Use of Employee Mental Health Programs: Mixed Methods Study

Benedict Sevov[1], MSc; Robin Huettemann[1], Dr; Maximillian Zinner[1], Dr; Sven Meister[2,3], Prof Dr; Leonard Fehring[1,2,4], MD

[1]Faculty of Health, School of Medicine, Witten/Herdecke University, Witten, Germany

[2]Health Care Informatics, Faculty of Health, School of Medicine, Witten/Herdecke University, Witten, Germany

[3]Department Healthcare, Fraunhofer Institute for Software and Systems Engineering ISST, Dortmund, Germany

[4]Department of Gastroenterology, Helios University Hospital Wuppertal, Witten/Herdecke University, Wuppertal, Germany

Corresponding Author:
Leonard Fehring, MD
Faculty of Health
School of Medicine
Witten/Herdecke University
Alfred-Herrhausen-Strasse 50
Witten, 58455
Germany
Phone: 49 202 8963708
Email: leonard.fehring@uni-wh.de

Abstract

Background: Mental health issues represent a prevalent challenge for employees and their employers, leading to substantial impacts on individuals, society, and the economy. Different employee mental health programs (EMHPs) can support employees in preventing and treating mental health issues. However, the impact of such EMHPs depends largely on the use behavior of employees.

Objective: This study aimed to gain deeper insights into employees' attitude and use behavior regarding EMHPs by investigating (1) employee preference and intention to use EMHPs, (2) factors that predict use, and (3) key facilitators and barriers influencing use.

Methods: An exploratory sequential mixed methods approach was applied, including a scoping review, qualitative interviews, and a quantitative web-based survey. Semistructured qualitative interviews were conducted with 15 employees to gain insights into EMHPs from the employee perspective and inform the creation of a web-based questionnaire. The quantitative web-based survey was conducted to collect representative primary data on employees' perspectives on different EMHPs using 7-point Likert scales. The collected quantitative data were analyzed through descriptive and inferential statistics, including repeated measures ANOVAs and chi-square tests.

Results: The final sample of the web-based survey consisted of 1134 participants and was representative across several sociodemographic characteristics. Analysis of the sample revealed that when given the choice, employees preferred digital (n=666, 58.73%) and self-intervention (n=590, 52.03%) EMHPs. Employees were most likely to use EMHPs focused on prevention (mean 4.89, SD 1.61). Intention to use EMHPs was predicted by age (young: mean 4.59, SD 1.2; old: mean 4.19, SD 1.4; P<.001; Cohen d=0.32), education (academic degree: mean 4.68, SD 1.24; no academic degree: mean 4.26, SD 1.32; P<.001; Cohen d=0.32), and mostly by company culture (positive company culture: mean 4.61, SD 1.27; negative company culture: mean 3.99, SD 1.27; P<.001; Cohen d=0.49). Cost coverage (n=345, 30.42%) and ease of use (n=337, 29.72%) were critical facilitators of use.

Conclusions: Employers can have a positive contribution to employee mental health by starting to offer EMHPs, preferably digital self-intervention programs for prevention; creating and maintaining the right work environment and culture; and ensuring cost coverage for the EMHP.

(JMIR Hum Factors 2025;12:e65750) doi:10.2196/65750

XSL•FO
RenderX

KEYWORDS

digital health; mental health; workplace; employee well-being; employee mental health programs; intention to use; facilitators and barriers; mixed methods

Introduction

Background

Mental health issues represent a rising global challenge. In this paper, we broadly define mental health issues as any condition affecting an individual's mental health. This definition aligns with the World Health Organization's (WHO's) term "mental health condition" and includes mental disorders as defined by the WHO [1]. According to the WHO, approximately 12% of the global population experience a mental health condition, and every fourth family has a member who is affected by a mental disorder [1,2]. Most of these individuals are employed and therefore represent the active workforce [3]. Data suggest that approximately 20% of employees in industrialized countries experience mental health issues, either through mild symptoms or more serious forms of mental disorders, with prevalence continuing to rise [3,4]. This leads to 3 main adverse effects. First, at the individual level, mental health issues deteriorate the mental health and well-being of employees and negatively impact their quality of life and job performance. Second, at the societal level, mental health issues place a significant burden on welfare systems, which must cover the high costs of treatment and pension requirements. Third, at the economic level, damage arises from decreased productivity, which causes a shrinking economic surplus for individual entities and the respective national economies [1,5,6]. Total annual costs of mental disorders, including health care expenses, costs due to lost productivity, and related economic impacts, are estimated to range between €100 billion and €150 billion (approximately US $110-$170) for countries such as Germany and France [7,8]. On a global scale, these costs amount to approximately US $5 trillion [9].

The workplace itself offers a promising environment for promoting good mental health by preventing mental health issues in early stages or mitigating the effects of existing mental disorders [10-12]. Many programs exist that are explicitly designed or used to support employees [13-16]. As no uniform umbrella term is commonly used for such programs, we refer to these programs as employee mental health programs (EMHPs). In the style of the broader definition of employee assistance programs (EAPs) of the Employee Assistance Professionals Association [17,18], we define EMHPs as any programs or solutions offered and facilitated by employers to support employees in sustaining or regaining good mental health. We consider all solutions independent of whether they stem from company internal units or external service providers. Real examples are manifold as employers offer their employees a diverse range of EMHPs, with the programs coming in different forms and formats [12,18-20]. Among others, there are 3 key dimensions to characterize EMHPs. First, the medium, that is, whether an EMHP is in a digital, analog, or hybrid form. Second, the interaction form, that is, whether an EMHP is self-guided, delivered in a one-on-one setting, or used in a group. Third, the addressed stage of mental health, that is, whether an EMHP is supposed to support employees in preventing mental challenges, treating existing mental disorders, or rehabilitating from past mental health issues [21-24]. Among EMHPs, digital programs are particularly popular and increasingly offered due to their advantages, such as cost efficiency, scalability, and flexibility in terms of time and location [25].

Prior Work

Several scientific studies and meta-analyses have researched the effectiveness of EMHPs. While some studies have not found clear evidence for the effectiveness of EMHPs in improving mental health outcomes [26,27], many other studies suggest that EMPHs can have significant positive effects on employees' mental well-being and thereby improve resilience and work effectiveness [15,18,28,29]. In recent years, studies and meta-analyses researching the effectiveness of digital EMHPs have gained particular attention and provide similar evidence for the potential positive effects of digital programs [20,25,30-33]. These positive effects have also been examined and found for different employee groups, including those in specific professions and across various countries [34-39]. Although both digital and analog EMHPs have been found effective in several studies, researchers note that the impact of these programs depends on the type of EMHP and the specific employee group they target [32,33,40].

Several studies have also identified different general facilitators of and barriers to using EMHPs [24,41,42]. However, only some studies have looked at differences in preference and use behavior based on employee-related characteristics, such as age and gender [43,44], type of profession [42], and education and income [45]. Existing research on predictive characteristics primarily focuses on specific EMHPs and select employee characteristics, leaving a gap in the literature regarding a comprehensive, holistic perspective on how different employee groups prefer and use various types of EMHPs. However, such a holistic view is essential to create transparency on which factors influence EMHP preference and use, particularly for digital ones. A deeper understanding of these predictors would benefit both researchers and practitioners. For researchers, it would help ensure that future research adopts an employee-centered approach. For employers and other stakeholders, it would help provide valuable insights into which EMHPs are most relevant and effective for their targeted workforce, ultimately increasing adoption rates. Greater adoption of EMHPs is likely to improve employees' mental well-being and reduce symptoms, while also mitigating productivity losses. The value of and need for further research to determine suitable mental health resources for employees is also explicitly mentioned by researchers [46]. To the best of our knowledge, no comprehensive study exists that combines research on multiple predictive factors for the use of different EMHP types and relevant facilitators and barriers.

XSL•FO
RenderX

Objectives

Therefore, the objective of this study was to address this research gap by (1) researching preference and intention to use of different types of EMPHs from the employee perspective, comparing digital EMHPs to nondigital ones; (2) examining several predicting factors for potential use of EMHPs; and (3) identifying key facilitators of and barriers to using EMHPs (research questions [RQs] 1-3).

On the basis of the first indications from existing research on EMHPs, as well as on initial evidence from research on the effects of several demographic, sociocultural, program, and other characteristics on intention to use or participation in general health interventions [47,48], we hypothesize that several employee-related factors, such as demographics and personal characteristics, and company-related factors, such as company culture, represent solid predictors for the use of EMHPs. As initial studies have observed different personal preferences for EMHPs among distinct demographic groups [43,44], the predictive value of the several considered characteristics may differ by type of EMHP, especially between digital and nondigital ones.

Methods

Exploratory Sequential Mixed Methods Approach

We applied an exploratory sequential mixed methods approach with 3 steps. The methodology and results of each step are reported separately, in line with the Mixed Methods Article Reporting Standards. The first step represented a scoping review to identify relevant research on EMHPs. The second step was an exploratory interview study for primary data collection. We conducted semistructured qualitative interviews to collect additional insights into the research topic and illustrate employees' views on EMHPs. The third step was a quantitative web-based survey for primary data collection. For the survey, a closed-format questionnaire was created and validated based on the results of the previous 2 steps.

Ethical Considerations

The Ethics Committee of Witten/Herdecke University raised no ethical concerns with regard to this research project (S-12/2023). All individuals who participated in either the qualitative interview study or the quantitative web-based survey provided voluntary, informed, and written consent to participate in the study and have the results published in a peer-reviewed article. The primary data collection approaches adhered to established ethical standards in scientific research as well as data protection regulations (General Data Protection Regulation). The data collected during the qualitative interview study were factually anonymized in a way that made it impossible, or only possible with disproportionately great effort, to identify individual participants. The data collected through the web-based survey were anonymized using the k-anonymity technique, which ensured that responses were removed if they did not meet the defined anonymity criteria. Additionally, demographic data obtained from the web-based survey such as age were collected in predefined intervals rather than as exact values to further ensure participant anonymity. The participants of the web-based survey received a monetary compensation from the research panel provider that executed the data collection. The research panel provider was remunerated by the authors.

Scoping Review

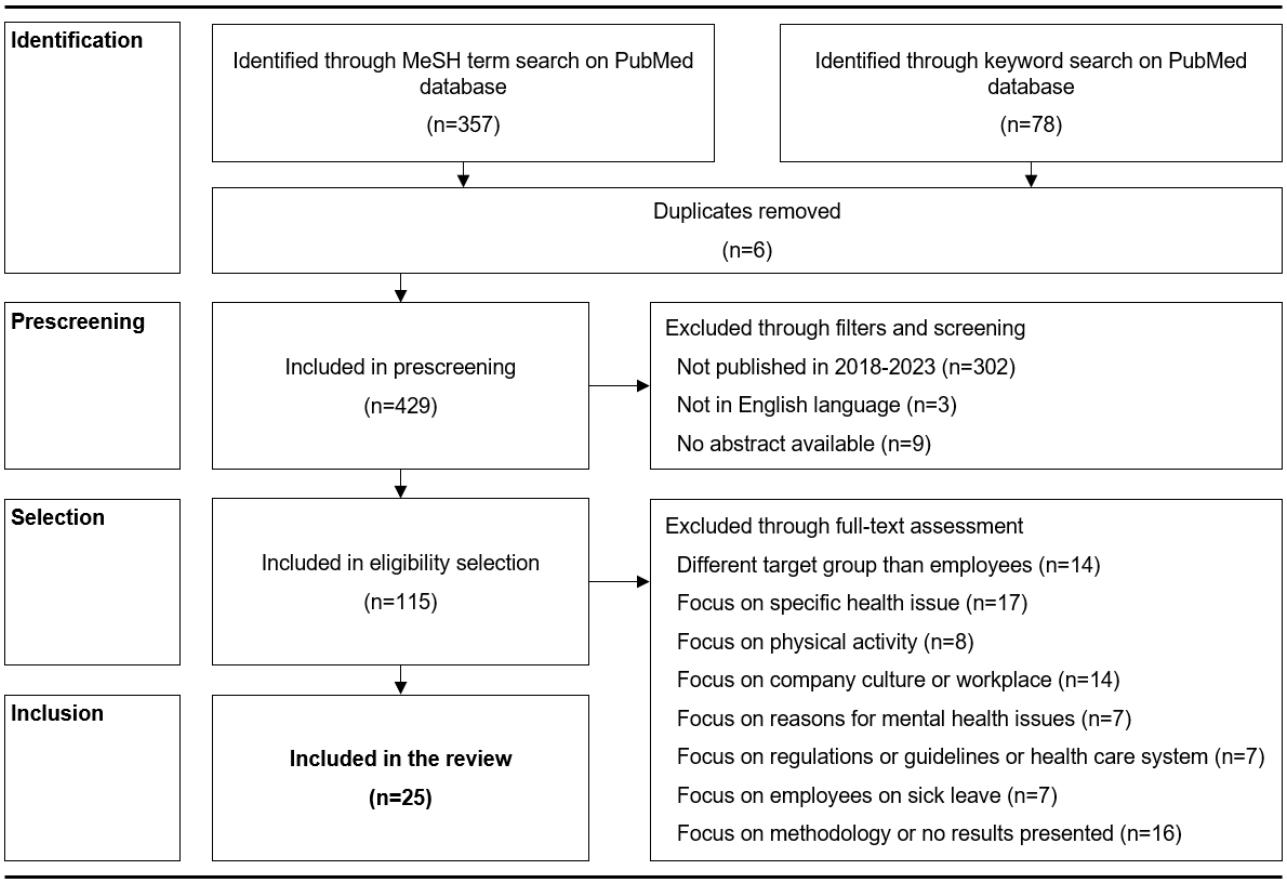
A scoping review according to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidelines was performed to scope the existing research on EMHPs; refer to the PRISMA-ScR checklist (Multimedia Appendix 1) for further details. The review was conducted in 2 streams with the following 2 search queries performed in the "PubMed" database on January 12, 2023:

1. ((mental health[MeSH Terms]) OR (psychology[MeSH Terms])) AND ((employee assistance program[MeSH Terms]) OR (employer intervention[MeSH Terms]) OR (workplace intervention[MeSH Terms])) OR ((mental health[MeSH Terms]) AND (workplace[mesh terms]) AND (health promotion[mesh terms]))
2. (mental health[Title/Abstract]) AND ((employee assistance program[Title/Abstract]) OR (employer intervention[Title/Abstract]) OR (workplace intervention[Title/Abstract]))

The final set of literature records was identified based on a structured screening process; details are found in Figure 1. Backward and forward snowballing was then applied to reveal further relevant literature records. For snowballing, we allowed the inclusion of relevant records published outside the defined publication period used for the screening process. The overall identified literature was used to create the semistructured interview guide and the web-based questionnaire with its answer options (initial codes).

XSL•FO
RenderX

Figure 1. Overview of the scoping review process. The defined scoping review process included a systematic screening rationale according to the PRISMA-ScR (Preferred Reporting Items For Systematic Reviews And Meta-Analyses Extension for Scoping Reviews) guidelines, resulting in 25 literature records in the final review. MeSH: Medical Subject Headings.



Semistructured Exploratory Interview Study

The interview guide (Multimedia Appendix 2) was designed in an open-ended format to support the exploratory nature of the interviews and was structured along the defined RQs. The interview study aimed to identify the most relevant types of EMHPs, as well as all relevant facilitators and barriers from the employee perspective, to ensure that these were included in the web-based questionnaire. Results from the interviews were also used to illustrate relevant findings from the web-based survey, which, in turn, were interpreted in the context of the interview results. A purposive sampling approach was applied to obtain a representative participant sample of the observed population (employees of employers in Germany). The participants were recruited from the broader network of all authors and selected such that they were well distributed across the 5 demographic and company characteristics—age, gender, education, size of employer, and industry of employer. This ensured representativeness. All participants were informed in writing about the study's purpose and data privacy terms. Selected parts of the resulting interview transcripts were analyzed and used for another research study.

The interviews were conducted in German, either in person or via phone or video call software, during February and March 2023. The interviews were recorded and transcribed in a factually anonymized way so that no participant or company names were stated, and personal data were categorized, for example, into age groups [49]. The transcripts were used to conduct qualitative data analysis through 2 approaches. First, we applied the directed content analysis approach outlined by Hsieh and Shannon [50] to identify the most relevant dimensions of EMHPs, for example, medium or interaction form. Second, we applied the thematic analysis approach described by Braun and Clarke [51] to identify relevant facilitators of and barriers to using EMHPs and relevant mental disorders.

The coding process was structured along 2 iterative coding cycles, and the coding was conducted and reviewed independently by 2 coauthors using the software MAXQDA (VERBI Software Consult Sozialforschung. GmbH) [52]. For both approaches, during the first coding cycle, single words or phrases were either matched with one of the initial codes derived from the scoping review or were defined as a new code. During the second coding cycle, the assigned codes were reviewed and further consolidated or separated depending on the thematic fit. The directed content analysis approach allowed for the prioritization of the most relevant dimensions of EMHPs from the employee perspective based on the frequency of the respective codes, which were limited to 1 per participant to avoid skewed data resulting from participants referring several times to 1 single code.

The COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist describes the details of the applied method (Multimedia Appendix 3).

XSL•FO
RenderX

Quantitative Web-Based Survey

Survey Design

The web-based questionnaire (Multimedia Appendix 4) was informed by the results of the qualitative interview study and designed in a closed-ended format. It was structured along the RQs with multiple- or single-choice questions and constructs based on established measures, assessed through 7-point Likert scales (except for mental health status). The questions and constructs represented predicting variables (PV), dependent variables (DV), or collected general information (GI).

Questions and Constructs

Demographic Characteristics (PV)

Data on age, gender, income, and education were collected through standard questions for such web-based surveys (categorical single-choice questions).

Digital Health Care Literacy (PV)

Participants rated their digital health care literacy using the Digital Health Care Literacy Scale by Nelson et al [53].

Mental Health Status (PV)

The current mental health status was assessed through a 3-item short form (Mental Health Inventory [MHI]; MHI-3) of the 5-item (MHI-5) measure [54]. To limit the overall number of questions, the MHI-3 was chosen for its brevity, while still being based on the well-established and widely used MHI-5 [55]. The items of the used MHI-3 were taken from a German version, and answers were collected through a 6-point Likert scale [56]. A mental health score was calculated and used to divide participants into 2 groups (positive or negative mental health status) based on a threshold defined by the scoring method of Rumpf et al [57]. Further, yes or no questions were asked to gain more detail on personal mental health status.

Relevance of Mental Health (GI)

The personal view on the relevance of mental health in society was assessed through a 1-item question. The mental disorders perceived as relevant among employees were assessed based on a multiple-choice question. The answer items were informed by literature and the interview study [58,59].

Company-Related and Profession-Related Characteristics (GI)

The number of employees, the industry of the employer, and the profession type were asked based on single-choice questions. Answer items were derived from established concepts and selectively adapted [60-62].

Company Culture (PV)

Data on company culture were collected through the German "Kurzskala zur Erfassung der Unternehmenskultur" (corporate culture scale, short form) by Jöns et al [63] using the 2 subscales leadership (5 items) and collaboration (4 items).

Intention to Use (DV)

Intention to use was measured as a DV for 9 different types of programs. The first 6 program types were derived from all combinations of the 2 dimensions of EMPHs: medium (digital and analog) and interaction form (self-intervention, bilateral intervention, and group intervention). The additional 3 program types were focused on the different addressed stages of mental health (prevention, treatment, and rehabilitation). All dimensions were clearly defined in the web-based questionnaire (Multimedia Appendix 4) to ensure a uniform understanding of all survey participants; for example, a program was defined as digital if it incorporated any form of digital technology, even if only a single component was digital. Intention to use was assessed with 1 item derived from the intention to use construct developed by Venkatesh et al [64].

Preference (DV)

To assess the unique preference for one of the first 6 programs, a single-choice question was asked.

Factors for Use (Facilitators and Barriers; DV)

Factors supporting and discouraging the use of EMHPs were assessed through 2 separate multiple-choice questions, each including an "Other" option with a free-text field. The given answer items were based on the scoping review and the qualitative interview study. The latter ensured the comprehensiveness of the selection of relevant facilitators and barriers from an employee perspective.

Offered EMHP (GI)

Participants were asked whether their employers offered any EMHP in the last 12 months and, if so, what type (medium, interaction form, and the addressed stage of mental health) and whether it was used. If participants used the program, their satisfaction was assessed through the 2-item system satisfaction construct by Wixom and Todd [65].

A control question was asked in the middle of the survey to check for participants' attention.

Data Collection

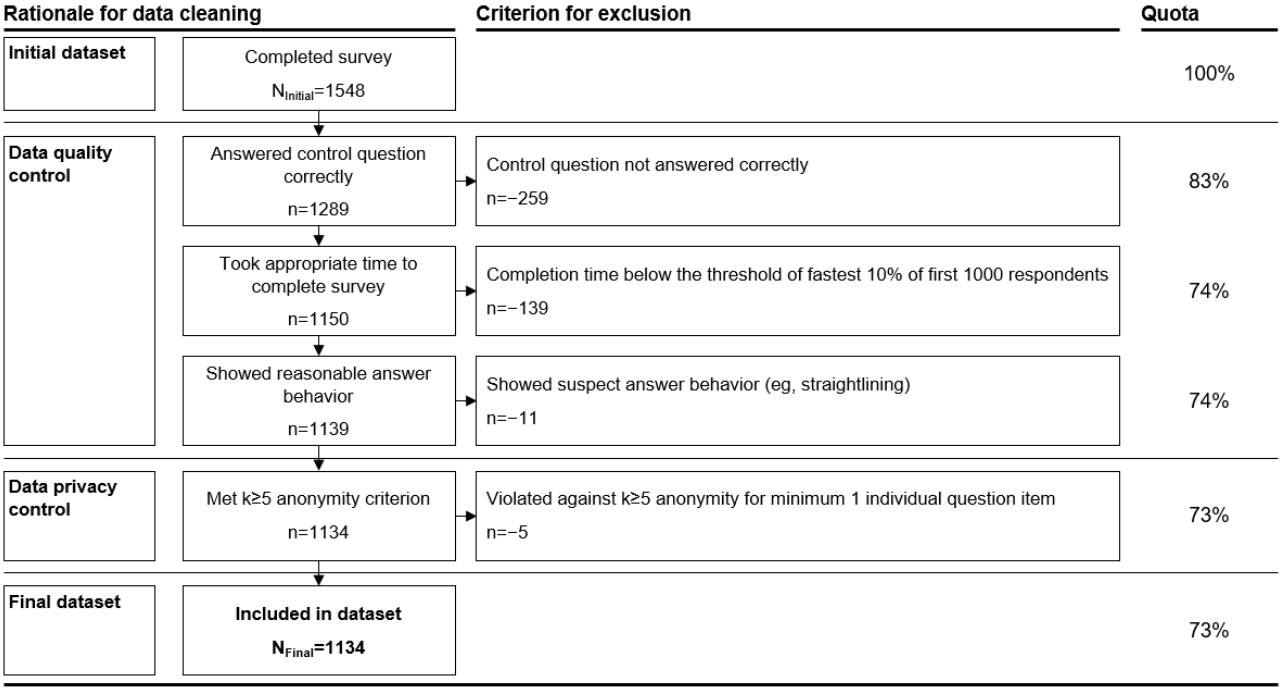
The web-based survey was conducted in Germany in the German language in April and May 2023. A purposive sampling approach was applied to obtain a representative participant sample. Participant recruitment was conducted through a research panel provider that contacted registered panel participants from across Germany via email to invite them to participate in the study. As an inclusion criterion, participants needed to be employed by an employer in Germany for a minimum of 6 months to be able to talk about the mental health support at their current employer. Moreover, during open data collection, quotas were used to ensure a diverse sample with a representative distribution of the survey participants along the sociodemographic characteristics age, gender, and education. Official figures from the German Federal Statistics Office were used as reference values for the distributions to manage the quotas [66,67]. With this and the large sample size, it was ensured that the final sample was representative of the observed population of employees working for employers in Germany. Further details are provided according to the CHERRIES (Checklist for Reporting Results of Internet E-Surveys) checklist (Multimedia Appendix 5).

XSL•FO
RenderX

Descriptive and Inferential Statistical Analyses

After cleaning the survey data (Figure 2), multi-item constructs were tested for reliability using thresholds for Cronbach α based on the study by Blanz [68]. Subsequently, data were first analyzed using descriptive statistics. Second, data were analyzed using inferential statistical analyses. Chi-square tests were conducted to examine the relation between demographic characteristics and actual use of EMHPs and the relation between preference for specific EMHP types and perception of the relevance of influencing factors. Effect sizes were measured using phi (ϕ). Repeated measures ANOVAs were conducted to analyze the difference in intention to use different EMHPs of all participants (within-participants design) [69]. The Greenhouse-Geisser correction was applied when the assumption of sphericity was violated. The Bonferroni correction was applied for multiple comparisons. Effect sizes were measured using partial eta squared for the ANOVAs and Cohen *d* for the pairwise comparisons. One-way ANOVAs were conducted to identify differences in the means of the DV (intention to use) between different groups. Effect sizes of group differences were measured using Cohen *d*. Groups were formed based on relevant independent variables, such as demographics, mental health status (current or past), and company culture. Selected demographic variables were grouped into binary variables (eg, age into young vs old and education into nonacademic vs academic). Assuming interval scale level, the means of variables reported on Likert scales were calculated as arithmetic means [70]. The significance level was set at a *P* value of <.05. SPSS Statistics software (version 29.0; IBM Corp) was used for statistical analyses.

Figure 2. Overview of the data cleaning process. The defined data cleaning process included several data quality control steps and 1 data privacy control step to achieve the final dataset of 1134 participants. Suspect answer behavior means answers with strong patterns that might indicate careless responding, for example, straightlining.



Results

Scoping Review

Through the scoping review, 25 records relevant to the characterization of EMHPs and factors impacting use of such programs were identified (Figure 1). Furthermore, 12 records relevant to the research topic were added through snowballing (an overview of all records can be found in Multimedia Appendix 6 [3-6,10-16,18-21,23-25,28-46]). A total of 45 initial codes for the web-based survey questionnaire were identified, including initial codes for relevant mental disorders.

Semistructured Exploratory Interview Study

The final sample of the exploratory interview study comprised 15 participants. The interview study was concluded after reaching content saturation, which was indicated by recurring responses, confirmation of previously identified themes, and the absence of any new emerging themes [71]. The participants were well distributed across age groups, gender, education levels, industry, and company size (number of employees; Multimedia Appendix 7).

The qualitative interviews identified the most relevant dimensions of EMHPs from the employee perspective, including medium, interaction form, and the addressed stage of mental health. In total, 17 new codes were identified and added to the 45 initial codes derived from the scoping review. The most relevant facilitators mentioned by interview participants included employees needing to be aware of offered EMHPs and employers being willing to fully or partly pay for the offered program, each mentioned by 12 (80%) of the 15 participants. This was followed by the importance of support through culture and leadership, mentioned by 11 (73%) of the 15 participants. The most relevant barrier mentioned during the interviews was that using the EMPH would be too time-consuming, which

XSL•FO
RenderX

would again increase pressure on employees willing to use it, mentioned by 10 (67%) of the 15 participants. This was followed by data protection and privacy concerns, mentioned by 9 (60%) of the 15 participants. As additional thoughts on the topic, 8 (53%) of the 15 participants, when being asked at the end of the interview, stated that they think it was a very critical topic and that more awareness across employers and employees would be needed.

Quantitative Web-Based Survey

Data Sample

The final dataset of the web-based survey included 1134 responses after data cleaning (Figure 2). To the best of our knowledge, this represents the largest survey of employees in Germany in this research context and one of the largest globally.

As targeted by design, the final sample is representative of the German workforce across the relevant demographic characteristics age, gender, and educational degree [66,67], as well as distributed across the company-related characteristics number of employees, industry, and profession (Table 1).

XSL•FO
RenderX

Table 1. Distribution of responses in the final dataset across demographic and company-related characteristics (N=1134).

Variable	Participants, n (%)
Age (years)	
≤29	214 (18.87)
30-39	236 (20.81)
40-49	222 (19.58)
50-59	291 (25.66)
≥60	170 (14.99)
Prefer not to answer	1 (0.09)
Gender	
Woman	523 (46.12)
Man	611 (53.88)
Nonbinary	0 (0)
Prefer not to answer	0 (0)
Education	
None	0 (0)
School degree except high school (“Abitur”)	79 (6.97)
High school degree (“Abitur”) or equivalent	108 (9.5)
Professional degree, vocational training, or equivalent	484 (42.68)
Bachelor’s degree or equivalent	201 (17.72)
Master’s degree or equivalent	215 (18.96)
Doctor or PhD degree or equivalent	43 (3.79)
Prefer not to answer	4 (0.35)
Number of employees	
≤9	99 (8.73)
10-49	199 (17.55)
50-249	261 (23.02)
250-499	140 (12.35)
500-999	142 (12.52)
1000-9999	161 (14.2)
≥10,000	132 (11.64)
Industry	
Energy and utilities	29 (2.56)
Raw materials and natural products	15 (1.32)
Industrial goods and services	170 (14.99)
Consumer goods and services	153 (13.49)
Commercial, technical, and scientific services as well as creative services	104 (9.17)
Health care	109 (9.61)
Financials and real estate	76 (6.7)
IT and communications	95 (8.38)
Public and governmental	125 (11.02)
Association, federation, and foundation	30 (2.65)
Other	228 (20.11)

XSL•FO
RenderX

Variable	Participants, n (%)
Profession	
Governmental professionals, including the armed forces	72 (6.35)
Scientists and Researchers	34 (3)
Commercial professions, office professionals, and equivalent	370 (32.63)
Technicians and equivalent	140 (12.35)
Service-related professions (eg, restaurant and hotel) and sales workers	132 (11.64)
Social, teaching, and caring professions	119 (10.49)
Creative and art-related professions	30 (2.65)
Craft and related trades workers as well as machinery operators and assemblers	83 (7.32)
Skilled agricultural, forestry, and fishery workers	15 (1.32)
Other	139 (12.26)

Measured Constructs

The 4 established constructs used were statistically reliable and internally consistent (Multimedia Appendix 8 [53,54,63,65]) [68]. The company culture measure showed excellent reliability, with a Cronbach α of 0.94 (9 items). The Digital Health Care Literacy Scale showed good reliability (Cronbach α=0.84; 3 items). The MHI-3 showed acceptable reliability (Cronbach α=0.76; 3 items). The EMHP satisfaction construct showed questionable reliability (Cronbach α=0.68; 2 items); however, given that Cronbach α generally is profoundly affected by the number of items and this construct only consists of 2 items, it can still be accepted as reliable [72].

Characteristics of the Sample and GI

MHI-3 suggested that 42.24% (471/1115) of employees had a negative mental health status. Independent of this, 52.84% (585/1107) of employees stated that they had issues with their mental health in the past. In total, 28.61% (321/1122) of employees stated that they pursue activities related to mental health, for example, psychotherapy, meditation, and using a mental health app. The mental health issues perceived as most relevant among employees were burnout (747/1134, 65.87% of participants), depression (717/1134, 63.23%), sleep disorders (704/1134, 62.08%), concentration issues (538/1134, 47.44%), and substance abuse or addiction (344/1134, 30.34%). Overall, employees rated the relevance of mental health in society as high (average score of 5.58, SD 1.56 on a 7-point Likert scale with 7=very relevant; median score of 6, IQR 5-7; 702/1134, 61.9% stated that mental health was “relevant” or “very relevant”).

Availability of EMHPs

Overall, 20.37% (231/1134) of participants mentioned that their employers offered at least 1 EMHP. Large companies were more likely to offer a program (117/435, 26.9% vs 30/298, 10.1% at smaller companies). Of the offered EMHPs, 65.8% (152/231) were digitally enabled, while 31.6% (73/231) were analog (6/231, 2.6% of participants did not know). Programs were equally distributed across interaction forms, with 33.3% (77/231) of them being self-intervention, 32% (74/231) bilateral intervention, and 28.6% (66/231) group intervention (14/231, 6.1% did not know). In total, 60.2% (139/231) of the offered EMHPs focused on prevention, 23.4% (54/231) on treatment, and 10% (23/231) on rehabilitation (15/231, 6.5% did not know; Multimedia Appendix 9).

Actual Use of EMHPs

In total, 54.8% (126/230) of employees with an EMHP available through their employer stated that they use it or have used it during the last 12 months. The use rate was higher when digital EMHPs were offered compared to when analog EMHPs were offered (digital: 94/152, 61.8%; analog: 32/72, 44.4%; P=.001; ϕ=0.24; Multimedia Appendix 10). Younger employees (<50 years) had a higher use rate compared to older ones (≥50 years; young: 107/174, 61.5%; old: 19/56, 34%; P<.001; ϕ=0.24; Multimedia Appendix 11). Employees with an academic degree (Bachelor’s, Master’s, or Doctor or PhD degree) had a higher use rate compared to employees without an academic degree (academic degree: 93/139, 66.9%; no academic degree: 33/91, 36%; P<.001; ϕ=0.30; Multimedia Appendix 12). Employees who used offered EMHPs stated a generally high satisfaction with the used program (average satisfaction score of 5.6, SD 1.2 on a 7-point Likert scale; median score of 6.0, IQR 5.0-6.5; 70/126, 55.6% of employees reported a score of ≥6), while there is no significant difference in satisfaction of employees using digital versus analog programs.

Preferred EMHPs

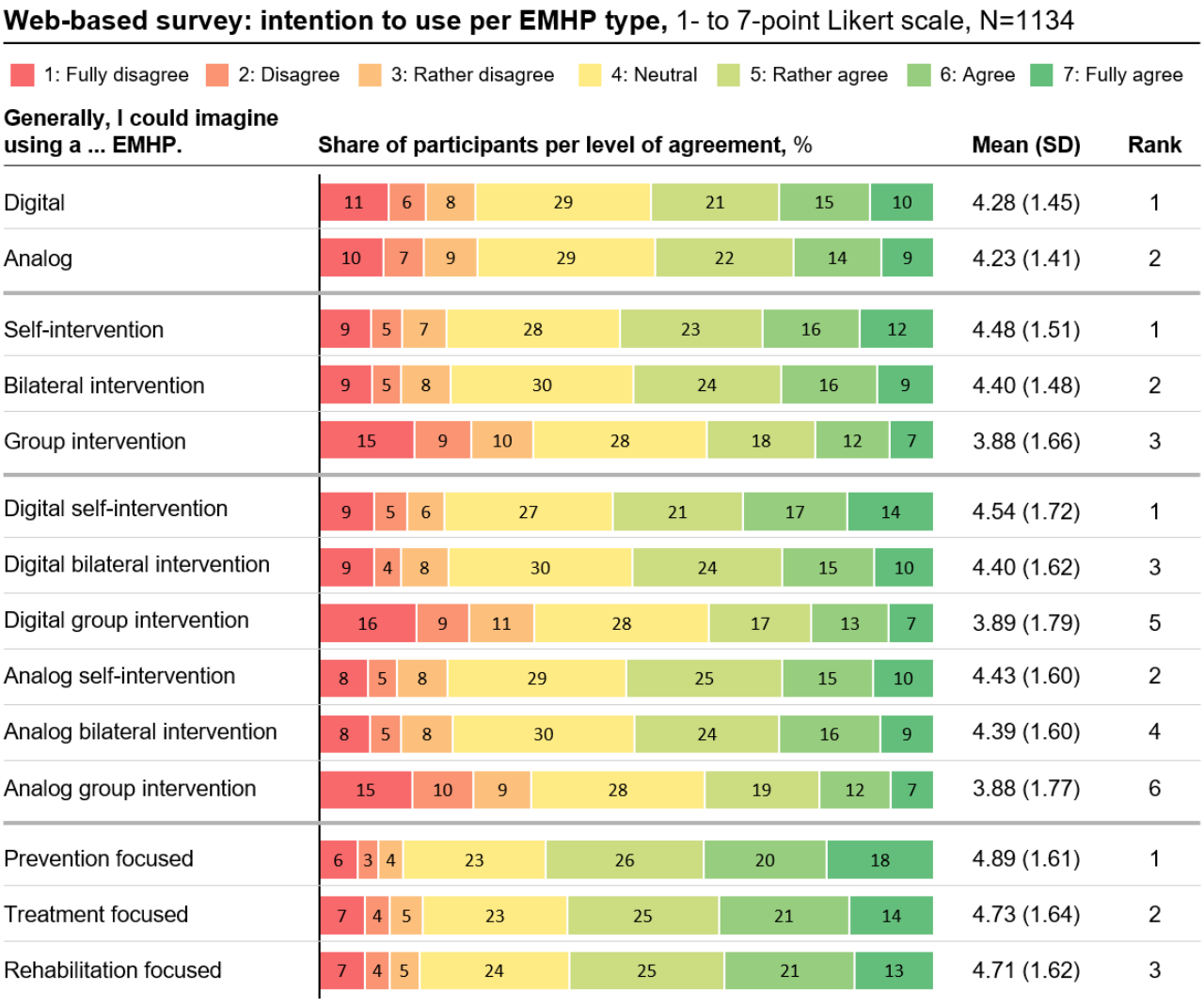
When stating their unique preference, 58.73% (666/1134) of employees stated their preference for digitally enabled EMHPs and 52.03% (590/1134) for self-intervention EMHPs. The most preferred EMHP combination of medium and interaction form was the digital self-intervention program (371/1134, 32.72% of participants), followed by the analog self-intervention program (219/1134, 19.31%) and the digital bilateral intervention program (152/1134, 13.4%), which represents part 1 of RQ1.

XSL•FO
RenderX

Intention to Use EMHPs

Overall, there was medium intention to use EMHPs, with an average of 4.43 on a 7-point Likert scale (median 4.56, IQR 3.89-5.22). However, the SD was high, with an SD of 1.30 (variance 1.68). On average, 50.88% (577/1134) of participants indicated a positive intention to use an EMHP (average score of >4.5 on a 7-point Likert scale), while 31.57% (358/1134) were neutral (average score of >3.5 and <4.5). Regarding part 2 of RQ1, there was no significant difference between the overall intention to use digital and analog programs when they were available. However, intention to use was significantly higher for self-intervention programs and bilateral intervention programs (self: mean 4.48, SD 1.51; bilateral: mean 4.40, SD 1.48) compared to group intervention programs (group: mean 3.88, SD 1.66; P<.001 each; Cohen d=0.40 and Cohen d=0.39, respectively). The difference in intention to use between self-intervention and bilateral intervention programs was also significant (P=.03; Cohen d=0.08). Considering both dimensions, medium and interaction form, intention to use was highest for a digital self-intervention program (digital self-intervention: mean 4.54, SD 1.72; 593/1134, 52.29% of participants indicated a positive intention to use, ie, a score of ≥5). Intention to use for EMHPs focusing on prevention was significantly higher than that for EMHPs focusing on treatment (prevention: mean 4.89, SD 1.61; treatment: mean 4.73, SD 1.64; P<.001; Cohen d=0.13) and that for EMHPs focusing on rehabilitation (rehabilitation: mean 4.71, SD 1.62; P<.001; Cohen d=0.14). Refer to Figure 3 for an overview of the intention-to-use results (details of all repeated measures ANOVAs can be found in Multimedia Appendix 13).

Figure 3. Results of the web-based survey on intention to use per employee mental health program (EMHP) type. Intention to use was assessed by asking the concrete intention to use on a 7-point Likert scale for the 6 EMHP types resulting from the combination of the dimensions medium (ie, digital and analog) and interaction form (ie, self-intervention, bilateral intervention, and group intervention), as well as for the 3 EMHP types based on the addressed stage of mental health (ie, prevention, treatment, and rehabilitation). The reported intention to use each EMHP type by medium and interaction form was calculated by aggregating the intention-to-use data for the 6 combined EMHP types.



Predictors for the Use of EMHPs

Addressing RQ2, it was found that intention to use was higher for younger employees (young: mean 4.59, SD 1.2; old: mean 4.19, SD 1.4; P<.001; Cohen d=0.32). No significant difference in overall intention to use nor regarding digital or analog EMHPs could be observed between women and men. However, women showed a higher intention to use self-intervention programs

XSL•FO
RenderX

(women: mean 4.61, SD 1.49; men: mean 4.38, SD 1.52; P=.008; Cohen d=0.16) as well as prevention and treatment programs (prevention: $mean_{women}$ 5.06, SD 1.51 and $mean_{men}$ 4.74, SD 1.68; treatment: $mean_{women}$ 4.90, SD 1.58 and $mean_{men}$ 4.57, SD 1.68; P<.001 each; Cohen d=0.20 each). Intention to use was higher for employees who held academic degrees (Bachelor's, Master's, or Doctor or PhD degree; academic degree: mean 4.68, SD 1.24; no academic degree: mean 4.26, SD 1.32; P<.001; Cohen d=0.32). Given that digital health care literacy was high across the whole sample (average digital health care literacy score of 6.03, SD 1.18 on a 7-point Likert scale; median score of 6.33, IQR 5.33-7.00; 774/1134, 68.25% of employees reported a score of ≥6), further analyses did not reveal additional insights. Intention to use was higher for employees who had faced mental health issues in the past (mental health issues in the past: mean 4.64, SD 1.14; no mental health issues in the past: mean 4.22, SD 1.41; P<.001; Cohen d=0.32) and for employees with a negative self-assessed mental health status according to MHI-3 (positive mental health status: mean 4.31, SD 1.33; negative mental health status: mean 4.56, SD 1.21; P=.002; Cohen d=−0.19). Employees who perceived their company culture as positive reported a higher intention to use EMHPs (positive company culture: mean 4.61, SD 1.27; negative company culture: mean 3.99, SD 1.27; P<.001; Cohen d=0.49; details of all ANOVAs can be found in Multimedia Appendix 14).

Facilitators of and Barriers to the Use of EMHPs

The most relevant factor for employees impacting the use of EMHPs is whether the employer paid for the program (345/1134, 30.42%). This top facilitator was significantly more relevant for employees who preferred analog EMHPs (160/468, 34.19%) versus those who preferred digital EMHPs (185/666, 27.78%; P=.02; ϕ=−0.07). In line with this, employers not paying for the EMHP was the number 1 barrier (433/1134, 38.18%). Ease of use was the second most relevant facilitator (337/1134, 29.72%), with no difference between employees preferring digital or analog EMHPs. Content quality was another important facilitator if given (259/1134, 22.84%) and an even more relevant barrier if not given (301/1134, 26.54%). The top 5 facilitators and barriers are shown in Table 2, representing RQ3 (details of all chi-square tests can be found in Multimedia Appendix 15).

Table 2. Top 5 facilitators of and barriers to using employee mental health programs (EMHPs) by total participants and differences by preference for medium (N=1134).

	Total participants[a] (N=1134), n (%)	Participants preferring digital EMHPs[b] (n=666), n (%)	Participants preferring analog EMHPs[b] (n=468), n (%)	P value[c]
Top 5 facilitators				
1. EMHP is paid for by the employer	345 (30.42)	185 (27.78)	160 (34.19)	.02
2. EMHP is easy to use	337 (29.72)	198 (29.73)	139 (29.7)	.99
3. EMHP with trustworthy data protection	282 (24.87)	168 (25.23)	114 (24.36)	.74
4. EMHP with easy access	265 (23.37)	153 (22.97)	112 (23.93)	.71
5. EMHP with high content quality	259 (22.84)	164 (24.62)	95 (20.3)	.09
Top 5 barriers				
1. EMHP requires full payment or copayment by the employee	433 (38.18)	250 (37.54)	183 (39.1)	.59
2. EMHP requires too much time	326 (28.75)	202 (30.33)	124 (26.5)	.16
3. EMHP with low content quality	301 (26.54)	173 (25.98)	128 (27.35)	.61
4. Employees feel forced to use EMHP	294 (25.93)	154 (23.12)	140 (29.91)	.01
5. EMHP is too complicated or requires too much effort	265 (23.37)	152 (22.82)	113 (24.15)	.60

[a]Share of total participants perceiving the facilitator or barrier as relevant.

[b]Share of participants preferring digital or analog EMHPs perceiving the facilitator or barrier as relevant.

[c]Significance level of the differences between participant groups preferring digital or analog EMHPs based on P values according to the chi-square test.

Discussion

Principal Findings

Generalizability of Findings

This study analyzed the deployment and use of EMHPs at companies across industries; the preference and intention to use of employees for different EMHPs, including relevant predictors; as well as facilitators and barriers. The representative and exceptionally large sample of the German working population allows for generalization of the findings. Generalizability is enhanced through the fact that no specific sociodemographic subgroup but a broad representative sample of the observed population across several sociodemographic characteristics was investigated. The representative distribution of education levels in the sample further ensured that no employees from a specific socioeconomic status group were underrepresented. The findings of the study are not only generalizable to Germany but may also be transferable to other countries, particularly industrialized nations where mental health

XSL•FO
RenderX

challenges are comparable and employer-employee relationships tend to be similarly structured. Given that a relevant share of participants' employers in this study were large organizations, many of these were likely to operate in several countries across the globe, where the general organizational culture and relationship between employer and employees can be expected to be similar. Moreover, the research on EMHPs generally focuses on the role of the employer as a provider of or at least an access point for mental health services. Therefore, this research is less dependent on the respective health system of a specific country but rather on universal predictors and factors of the relationship between employers and employees. To generate further insights into the transferability of the findings, this survey can be replicated in other industrialized countries by future research.

Prevalence and Actual Use of EMHPs

The fact that employees need support due to increasing mental health issues [16,28,31] is supported by our results, which illuminate the prevalence of past issues and current negative mental health status. While only a fifth of employees have access to EMHPs, more than half of these employees use them. This use rate seems to be high compared to numbers from other studies, indicating use rates for very specific psychotherapeutic consultation services and general EAPs of 1% up to 10% [73] or 39% when looking at intended use [44]. Hence, use rates seem to vary depending on the concrete EMHP. Given that the results of other studies focus on EAPs or consultation services for primarily existing issues, the overall use rates for all EMHPs, including prevention, are expected to be higher, as indicated by our results. Further studies should focus on measured use rates for different types of EMHPs, including prevention-focused programs.

Preferred EMHPs

The found preference for digital programs (666/1134, 58.73% of participants) and for self-intervention programs (590/1134, 52.03%) when choosing one preferred EMHP is in line with the desire for time- and location-wise flexibility and for anonymous use concerning colleagues [24,41]. While patients with existing mental health issues realize the potential advantages of digital programs [74], prior research has found that they prefer face-to-face mental health interventions or hybrid models that include face-to-face contact, presumably due to higher perceived effectiveness [74-76]. However, these studies focused on interventions treating existing mental health issues, for which many do not think that digital programs can replace traditional in-person psychotherapy adequately [77-79]. Furthermore, these studies did not explicitly focus on EMHPs in the work context but on the general population. Given that our study focuses on EMHPs that often also cover prevention purposes [24], the findings do not contradict each other. Instead, they indicate that preference for an intervention involving a mental health professional is higher when facing concrete, particularly severe issues, which was supported by our qualitative interviews:

> *I believe, if my mental state is seriously endangered, then I need a personal conversation, personal support. But if my mental state is somehow not good but not severely bad, then I can imagine using an app for myself, which can help me to feel better.* [Participant 9]

> *For prevention, tips for exercises, general information, or apps, etc, are very important. When there is a concrete need for treatment, then a one-on-one is the best option, with a psychotherapist.* [Participant 7]

However, digital and self-intervention programs are well suited for prevention purposes, especially in the workplace context [44], and digital EMHPs can even be as effective as nondigital ones regarding positive mental health outcomes [30]. A recent study confirms a strong preference for both types of EMHPs, digital self-intervention programs and programs delivered face-to-face [80]. While we acknowledge that preference for a certain EMHP type does not necessarily mean that these preferred EMHPs have higher effectiveness in obtaining the desired outcome, the intrinsic interest to use a specific EMHP is a relevant factor to start using such a program. As digital and nondigital EMHPs can have significant positive mental health effects, starting to use such a program is the most relevant first step, and consequently, the preference should be considered. The importance of intrinsic motivation and fun for starting to use an EMHP was also mentioned during our qualitative interviews:

> *Ideally, using it means fun. The people need a motivation to use it.* [Participant 9]

> *It is an important factor that the UX design is nice and that it is fun.* [Participant 11]

> *It needs to have an entertaining value; something needs to happen. Otherwise, you don't get the people to do it.* [Participant 13]

Intention to Use EMHPs

Overall, we found a medium intention to use EMHPs. However, the detailed results show a strong distribution, indicating opposing views on such programs. One part has a high intention to use both digital and analog EMHPs, while another part is very reluctant. Individuals with a high intention to use EMHPs may be influenced by the increasing awareness of mental health through more research and more coverage in mainstream media, which is also driven by celebrities raising awareness [81,82]. In contrast, other individuals neglect the need, as well as their eligibility, for such programs and for dealing with mental health more openly in general, partly driven by stigmatization, which could explain participants with low intention to use [28,83-85].

Predictors for Use of EMHPs

In line with previous research, intention to use differs substantially by several demographic and personal characteristics [44,45].

Age

Younger individuals tend to be more open toward EMHPs. This may be due to higher awareness and a more open mindset of younger generations toward mental illness, which is found in prior research [42,86]. Conversely, other research suggests that older age is a predictor of help-seeking behavior [87,88]. This

XSL•FO
RenderX

calls out for further research about differences in age concerning awareness and behavior regarding mental health support. Our findings support that younger generations are more open toward the topic and are more likely to use digital programs.

Gender

In contrast to other identified predictors that align with previous research, in this study, we found no evidence for differences in overall intention to use between women and men. However, when looking at specific EMHP types, women showed a higher intention to use programs focused on prevention and on treatment. This suggests that women might have a higher intention to use when presented with the concrete objective of the EMHP. The higher preference of women for self-intervention programs resonates with the findings of Smail-Crevier et al [44], who found that women have a significantly higher preference for self-help interactive programs than men. Further studies focusing on specific EMHPs with clearly defined dimensions, for example, interaction form, are needed to assess differences in preference by gender and gain deeper and solid insights.

Education

Intention to use is also found to be higher for individuals with higher formal education, which is in line with what has been suggested by prior research [45]. This might be explained by higher awareness and understanding of the challenge of mental health issues for individuals and society.

Previous Mental Health Issues

We found that individuals who indicate having had mental health issues in the past show a higher intention to use EMHPs. This suggests that awareness, primarily through lived experience with some mental health issues, increases the likelihood of seeking help and using offered support. Previous research has also found that positive experience with help seeking in the past increases the likelihood of seeking help again [87], which might also be the case for participants of our study, explaining the measured effect.

Mental Health Status

Furthermore, we found that the intention to use is higher for people who have a negative self-reported mental health status. Overall, this seems evident, as the need to use an EMHP is plausibly higher for these individuals than for people with good mental health. However, existing literature offers contrary findings regarding the effect of existing mental health issues on using support services. Some research suggests that when mental health issues reach a certain level of severity, individuals may experience reduced motivation and effort to engage with programs or seek treatment [79]. This effect is especially observable in people dealing with depression [41]. Other studies, in contrast, indicate that support is rather sought when issues already affect social functioning [87,89]. Hence, the severity of a present mental disorder plays a significant role in the resulting use of support services such as EMHPs. Therefore, further research could focus on the intention to use EMHPs based on the severity of mental health issues.

Company Culture

Another predictor of intention to use found in this study is the company culture. Individuals indicating a positive company culture at their workplace have a significantly higher intention to use EMHPs. This finding resonates with existing research on the importance of company culture and the positive effect of supportive leaders [5,13,46]. Notably, company culture emerged as the strongest predictor of intention to use EMHPs in this study, surpassing the effects of both age and education. The authenticity of the employer toward the offered program and the high relevance of the company culture were also mentioned by participants during the qualitative interviews:

> *It is important to feel that the employer means it, that it is a truly honest offer.* [Participant 1]
>
> *I want to have the feeling that this is not a one-time initiative to boost the image. I want to see the true intention of the employer, shown by continuous actions.* [Participant 11]
>
> *I need to have trust that I do not have to fear any negative consequences when I use it.* [Participant 9]
>
> *It is important that you have an environment where people do not judge, where colleagues say, "It is alright to use it, there is nothing condemnable with it."* [Participant 3]

Facilitators of and Barriers to the Use of EMHPs

Having the EMHP paid by the employer is the most relevant facilitator and represents the largest barrier if this is not the case. The reason that cost coverage is an even more important factor for employees preferring analog EMHPs might be that analog programs are usually more expensive than digital ones [25]. While the cost coverage factor is not explicitly mentioned in existing studies on facilitators and barriers, research suggests that it is in the nature of EMHPs to be offered to employees for free or at least at a low cost [42]. A potential expectation attitude of German citizens particularly might stem from the fact that they are used to reimbursed health care services, including psychotherapeutic treatment [76] and apps (specific digital health apps called DiGAs in Germany) [90]. In line with this, the qualitative interviews indicated that if the employer offers a program, employees expect financial support as a basic condition:

> *It needs to be paid by the employer, at least the largest share.* [Participant 10]
>
> *If it costs me personal money, this is a huge barrier.* [Participant 3]
>
> *If it needs to be paid fully or even partially by the employee, there is no more willingness to use it.* [Participant 5]

Further relevant factors are ease of use, easy access, use that should not require too much time, and quality of the content. Hence, convenience, that is, being able to integrate use into everyday life, and the quality of EMHPs are essential for employees, as is indicated by previous studies [14,24,41,91] and our qualitative interviews:

XSL•FO
RenderX

> *Use has to be simple and clearly explained.* [Participant 7]
>
> *Access to the program should be easy.* [Participant 3]
>
> *The time effort should be reasonable. I would like to have it easily integrated into my daily routines.* [Participant 1]
>
> *It not only has to have good content, it also has to be easily understandable and interesting.* [Participant 14]

Employers should ensure these criteria are met to increase the likelihood of use and enhance the impact of EMHPs. They should also consider digital EMHPs, which are generally more cost-efficient, easier to scale, and more flexible than nondigital programs [25].

Limitations

Despite high research standards, this study has some limitations. First, given that quantitative data collection was based on a web-based survey, it is likely that the results overrepresent digitally literate participants [92]. Therefore, participants might have been more prone to using digital EMHPs [93]. However, only a web-based survey of this nature can provide such a large sample representative across several demographic characteristics.

Second, for simplicity, the medium of EMHPs was categorized into a binary set with digital and analog EMHPs, whereas in reality, hybrid forms are available. Despite the existence of hybrid forms, the chosen definition was appropriate for the scope and objectives of our research, and the definitions were clearly presented in the web-based survey.

Third, the survey results relied on self-report measures and question items and the participants' chosen options, which may have led to some skewness in the data. For instance, participants might have indicated their intended and not their actual use behavior when asked, a phenomenon commonly referred to as the intention-behavior gap [94]. In addition, results on the available EMHPs were based on the participants' perspectives and not necessarily on the factual availability of the EMHPs offered by their employers.

Conclusions

Our study contributes to a deeper scientific understanding of employees' perspectives on different EMHPs, including preference, intention to use, predictors for use, and facilitators and barriers. This deeper understanding is relevant for academia and occupational practice. On the basis of our research, academia can drive further scientific studies and findings. Employers who are willing to actively contribute to the mental health of their employees can leverage our findings to choose and facilitate EMHPs, while providers can leverage these findings to create relevant EMHPs that match the employees' needs.

The 3 main implications of the study indicate simple but meaningful actions for employers worldwide. First, because a relevant portion of employees use EMHPs when available and most report satisfaction with them, it is a strong starting point to consider offering an EMHP. Initially, the type of EMHP is not of utmost importance. While starting with digital self-intervention programs focused on prevention may lead to greater acceptance and adoption, given overall preferences, employees generally appear to value the simple availability of any EMHP. Hence, offering any EMHP should potentially create a positive impact. Naturally, the more targeted and tailored the EMHP is, the higher the probability of positive effects. However ultimately, the most important step is to start offering an EMHP.

Second, employers can steer one of the most relevant factors predicting the use of EMHPs—the company culture. This predictor has significant impact on use across all EMHP types. Therefore, employers are advised to create an authentic strategy and communication on offering EMHPs and genuinely support the opportunity to use an EMHP. Conversely, offering an EMHP while having a negative company culture with demotivating communication, leadership, and teamwork styles will likely have no relevant positive effect.

Third, employers should ensure that employees seeking support or willing to act preventatively do not have to pay to use an offered EMHP. Furthermore, they should provide EMHPs that are easy to use, easy to access, and of good quality, as these are the key facilitators of use. Thus, employers should consider offering digital EMHPs, which typically address several of these relevant factors.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

BS and LF contributed to the conceptualization, design, and methodology of the study. RH, MZ, and SM contributed to the methodology of the study. BS conducted the scoping review; screened the identified literature records; planned, prepared, and conducted the qualitative interviews; transcribed and analyzed the interview transcripts; planned, prepared, and conducted the quantitative web-based survey; and cleaned and analyzed the data of the web-based survey. RH reviewed the interview guide and the web-based questionnaire, analyzed the interview transcripts, and reviewed the statistical analyses. BS prepared the first draft of the manuscript. All authors reviewed the manuscript. BS edited the manuscript and prepared the final version. All authors approved the final version of the manuscript. LF supervised the research project.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
The 22-item PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) checklist.
[DOCX File , 63 KB - humanfactors_v12i1e65750_app1.docx]

Multimedia Appendix 2
Interview guide for the exploratory qualitative interview study.
[DOCX File , 55 KB - humanfactors_v12i1e65750_app2.docx]

Multimedia Appendix 3
The 32-item COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist.
[DOCX File , 64 KB - humanfactors_v12i1e65750_app3.docx]

Multimedia Appendix 4
Web-based questionnaire for the quantitative web-based survey.
[DOCX File , 70 KB - humanfactors_v12i1e65750_app4.docx]

Multimedia Appendix 5
The 30-item CHERRIES (Checklist for Reporting Results of Internet E-Surveys) checklist.
[DOCX File , 61 KB - humanfactors_v12i1e65750_app5.docx]

Multimedia Appendix 6
Overview of the literature identified through scoping review and snowballing.
[DOCX File , 79 KB - humanfactors_v12i1e65750_app6.docx]

Multimedia Appendix 7
Demographic and company-related characteristics of participants in the interview study (N=15).
[DOCX File , 59 KB - humanfactors_v12i1e65750_app7.docx]

Multimedia Appendix 8
Reliability and factor analyses of the used constructs.
[DOCX File , 238 KB - humanfactors_v12i1e65750_app8.docx]

Multimedia Appendix 9
Distribution of different types of employee mental health programs offered by employers in Germany.
[DOCX File , 120 KB - humanfactors_v12i1e65750_app9.docx]

Multimedia Appendix 10
Cross-table of the chi-square test of independence for the relation between the medium of available employee mental health program and actual employee mental health program use.
[DOCX File , 57 KB - humanfactors_v12i1e65750_app10.docx]

Multimedia Appendix 11
Cross-table of the chi-square test of independence for the relation between the binary age variable and actual employee mental health program use.
[DOCX File , 57 KB - humanfactors_v12i1e65750_app11.docx]

Multimedia Appendix 12
Cross-table of the chi-square test of independence for the relation between the binary education variable and actual employee mental health program use.
[DOCX File , 57 KB - humanfactors_v12i1e65750_app12.docx]

XSL•FO
RenderX

Multimedia Appendix 13
Results of repeated measures ANOVAs analyzing differences in intention to use different employee mental health programs, differentiated by several dimensions.
[DOCX File , 66 KB - humanfactors_v12i1e65750_app13.docx]

Multimedia Appendix 14
Mean comparisons of intention to use different employee mental health programs between the 2 groups of each binary predicting variable using ANOVA.
[DOCX File , 73 KB - humanfactors_v12i1e65750_app14.docx]

Multimedia Appendix 15
Chi-square tests of independence for the relation between employee groups preferring digital or analog employee mental health programs and perceiving specific facilitators and barriers as relevant for employee mental health program use.
[DOCX File , 60 KB - humanfactors_v12i1e65750_app15.docx]

References

1. World mental health report: transforming mental health for all 2022. World Health Organization. 2022. URL: https://apps.who.int/iris/handle/10665/356119 [accessed 2024-04-29]
2. About mental disorders 2019. World Health Organization. 2019. URL: https://applications.emro.who.int/docs/EMRPUB_leaflet_2019_mnh_212_en.pdf [accessed 2024-04-29]
3. Attridge M. A global perspective on promoting workplace mental health and the role of employee assistance programs. Am J Health Promot 2019 May;33(4):622-629. [doi: 10.1177/0890117119838101c] [Medline: 31006254]
4. Jesuthasan J, Low M, Ong T. The impact of personalized human support on engagement with behavioral intervention technologies for employee mental health: an exploratory retrospective study. Front Digit Health 2022;4:846375 [FREE Full text] [doi: 10.3389/fdgth.2022.846375] [Medline: 35574254]
5. Dimoff JK, Kelloway EK. With a little help from my boss: the impact of workplace mental health training on leader behaviors and employee resource utilization. J Occup Health Psychol 2019 Feb;24(1):4-19. [doi: 10.1037/ocp0000126] [Medline: 29939045]
6. Gray P, Senabe S, Naicker N, Kgalamono S, Yassi A, Spiegel JM. Workplace-based organizational interventions promoting mental health and happiness among healthcare workers: a realist review. Int J Environ Res Public Health 2019 Nov 11;16(22):4396 [FREE Full text] [doi: 10.3390/ijerph16224396] [Medline: 31717906]
7. Chevreul K, Prigent A, Bourmaud A, Leboyer M, Durand-Zaleski I. The cost of mental disorders in France. Eur Neuropsychopharmacol 2013 Aug;23(8):879-886. [doi: 10.1016/j.euroneuro.2012.08.012] [Medline: 22959739]
8. König H, König HH, Gallinat J, Lambert M, Karow A, Peth J, et al. Excess costs of mental disorders by level of severity. Soc Psychiatry Psychiatr Epidemiol 2023 Jun;58(6):973-985 [FREE Full text] [doi: 10.1007/s00127-022-02298-8] [Medline: 35639134]
9. Arias D, Saxena S, Verguet S. Quantifying the global burden of mental disorders and their economic value. EClinicalMedicine 2022 Dec;54:101675 [FREE Full text] [doi: 10.1016/j.eclinm.2022.101675] [Medline: 36193171]
10. de Oliveira C, Cho E, Kavelaars R, Jamieson M, Bao B, Rehm J. Economic analyses of mental health and substance use interventions in the workplace: a systematic literature review and narrative synthesis. Lancet Psychiatry 2020 Oct;7(10):893-910. [doi: 10.1016/S2215-0366(20)30145-0] [Medline: 32949521]
11. Lam LT, Lam MK, Reddy P, Wong P. Efficacy of a workplace intervention program with web-based online and offline modalities for improving workers' mental health. Front Psychiatry 2022;13:888157 [FREE Full text] [doi: 10.3389/fpsyt.2022.888157] [Medline: 35711597]
12. Wan Mohd Yunus WM, Musiat P, Brown JS. Systematic review of universal and targeted workplace interventions for depression. Occup Environ Med 2018 Jan;75(1):66-75. [doi: 10.1136/oemed-2017-104532] [Medline: 29074553]
13. Bouzikos S, Afsharian A, Dollard M, Brecht O. Contextualising the effectiveness of an employee assistance program intervention on psychological health: the role of corporate climate. Int J Environ Res Public Health 2022 Apr 21;19(9):5067 [FREE Full text] [doi: 10.3390/ijerph19095067] [Medline: 35564466]
14. Dewa CS, Hoch JS. Barriers to mental health service use among workers with depression and work productivity. J Occup Environ Med 2015 Jul;57(7):726-731 [FREE Full text] [doi: 10.1097/JOM.0000000000000472] [Medline: 26147540]
15. Joseph B, Walker A, Fuller-Tyszkiewicz M. Evaluating the effectiveness of employee assistance programmes: a systematic review. Eur J Work Organ Psychol 2017 Sep 06;27(1):1-15. [doi: 10.1080/1359432X.2017.1374245]
16. Kristman VL, Lowey J, Fraser L, Armstrong S, Sawula S. A multi-faceted community intervention is associated with knowledge and standards of workplace mental health: the Superior Mental Wellness @ Work study. BMC Public Health 2019 May 24;19(1):638 [FREE Full text] [doi: 10.1186/s12889-019-6976-x] [Medline: 31126273]
17. Definitions of an employee assistance program: EAP core technology. Employee Assistance Professionals Association. URL: https://eapassn.org/page/definitionandcoretechnology# [accessed 2024-04-29]

XSL•FO
RenderX

18. Nunes AP, Richmond MK, Pampel FC, Wood RC. The effect of employee assistance services on reductions in employee absenteeism. J Bus Psychol 2017 Oct 4;33(6):699-709. [doi: 10.1007/S10869-017-9518-5]
19. Gantner M, Jarzcok MN, Schneider J, Brandner S, Gündel H, von Wietersheim J. Psychotherapeutic consultation services in the workplace: a longitudinal analysis of treatments and sick leave using health insurance data. Front Psychiatry 2022;13:838823 [FREE Full text] [doi: 10.3389/fpsyt.2022.838823] [Medline: 35401269]
20. Komase Y, Watanabe K, Hori D, Nozawa K, Hidaka Y, Iida M, et al. Effects of gratitude intervention on mental health and well-being among workers: a systematic review. J Occup Health 2021 Jan;63(1):e12290 [FREE Full text] [doi: 10.1002/1348-9585.12290] [Medline: 34762326]
21. Goetzel RZ, Roemer EC, Holingue C, Fallin MD, McCleary K, Eaton W, et al. Mental health in the workplace: a call to action proceedings from the mental health in the workplace-public health summit. J Occup Environ Med 2018 Apr;60(4):322-330 [FREE Full text] [doi: 10.1097/JOM.0000000000001271] [Medline: 29280775]
22. London E. Digital solutions for employee mental health: landscape overview, employer experiences, & best practices. Pacific Business Group on Health. 2020. URL: https://www.pbgh.org/wp-content/uploads/2021/01/Digital_MH_Report.pdf [accessed 2024-04-29]
23. Pandya A, Khanal N, Upadhyaya M. Workplace mental health interventions in India: a rapid systematic scoping review. Front Public Health 2022;10:800880 [FREE Full text] [doi: 10.3389/fpubh.2022.800880] [Medline: 35592077]
24. Scheutzow J, Attoe C, Harwood J. Acceptability of web-based mental health interventions in the workplace: systematic review. JMIR Ment Health 2022 May 11;9(5):e34655 [FREE Full text] [doi: 10.2196/34655] [Medline: 35544305]
25. Armaou M, Konstantinidis S, Blake H. The effectiveness of digital interventions for psychological well-being in the workplace: a systematic review protocol. Int J Environ Res Public Health 2019 Dec 30;17(1):255 [FREE Full text] [doi: 10.3390/ijerph17010255] [Medline: 31905882]
26. Fleming WJ. Employee well‐being outcomes from individual‐level mental health interventions: cross‐sectional evidence from the United Kingdom. Ind Relat J 2024 Jan 10;55(2):162-182. [doi: 10.1111/irj.12418]
27. Song Z, Baicker K. Effect of a workplace wellness program on employee health and economic outcomes: a randomized clinical trial. JAMA 2019 Apr 16;321(15):1491-1501. [doi: 10.1001/jama.2019.3307] [Medline: 30990549]
28. Dobson KS, Markova V, Wen A, Smith LM. Effects of the anti-stigma workplace intervention "working mind" in a Canadian health-care setting: a cluster-randomized trial of immediate versus delayed implementation. Can J Psychiatry 2021 May;66(5):495-502 [FREE Full text] [doi: 10.1177/0706743720961738] [Medline: 32960651]
29. Schneider RA, Grasso JR, Chen SY, Chen C, Reilly ED, Kocher B. Beyond the lab: empirically supported treatments in the real world. Front Psychol 2020;11:1969. [doi: 10.3389/fpsyg.2020.01969] [Medline: 32849153]
30. Carolan S, Harris PR, Cavanagh K. Improving employee well-being and effectiveness: systematic review and meta-analysis of web-based psychological interventions delivered in the workplace. J Med Internet Res 2017 Jul 26;19(7):e271 [FREE Full text] [doi: 10.2196/jmir.7583] [Medline: 28747293]
31. Huberty JL, Espel-Huynh HM, Neher TL, Puzia ME. Testing the pragmatic effectiveness of a consumer-based mindfulness mobile app in the workplace: randomized controlled trial. JMIR Mhealth Uhealth 2022 Sep 28;10(9):e38903 [FREE Full text] [doi: 10.2196/38903] [Medline: 36169991]
32. Stratton E, Lampit A, Choi I, Calvo RA, Harvey SB, Glozier N. Effectiveness of eHealth interventions for reducing mental health conditions in employees: a systematic review and meta-analysis. PLoS One 2017;12(12):e0189904 [FREE Full text] [doi: 10.1371/journal.pone.0189904] [Medline: 29267334]
33. Axén I, Björk Brämberg E, Vaez M, Lundin A, Bergström G. Interventions for common mental disorders in the occupational health service: a systematic review with a narrative synthesis. Int Arch Occup Environ Health 2020 Oct;93(7):823-838 [FREE Full text] [doi: 10.1007/s00420-020-01535-4] [Medline: 32246230]
34. Bernburg M, Groneberg D, Mache S. Professional training in mental health self-care for nurses starting work in hospital departments. Work 2020;67(3):583-590. [doi: 10.3233/WOR-203311] [Medline: 33185622]
35. Dunn A, Dixon C, Thomson A, Cartwright-Hatton S. Workplace support for mental health workers who are parents: a feasibility study. Front Psychol 2022;13:854065 [FREE Full text] [doi: 10.3389/fpsyg.2022.854065] [Medline: 35814147]
36. Mache S, Bernburg M, Baresi L, Groneberg D. Mental health promotion for junior physicians working in emergency medicine: evaluation of a pilot study. Eur J Emerg Med 2018 Jun;25(3):191-198. [doi: 10.1097/MEJ.0000000000000434] [Medline: 27879536]
37. Mohamed AF, Isahak M, Awg Isa MZ, Nordin R. The effectiveness of workplace health promotion program in reducing work-related depression, anxiety and stress among manufacturing workers in Malaysia: mixed-model intervention. Int Arch Occup Environ Health 2022 Jul;95(5):1113-1127. [doi: 10.1007/s00420-022-01836-w] [Medline: 35091853]
38. Ornek OK, Esin MN. Effects of a work-related stress model based mental health promotion program on job stress, stress reactions and coping profiles of women workers: a control groups study. BMC Public Health 2020 Nov 04;20(1):1658 [FREE Full text] [doi: 10.1186/s12889-020-09769-0] [Medline: 33148247]
39. Xu J, Liu X, Xiao Y, Fang X, Cheng Y, Zhang J. Effect of EAP psychological intervention on improving the mental health of medical workers under the novel coronavirus epidemic in China. Front Public Health 2021;9:649157 [FREE Full text] [doi: 10.3389/fpubh.2021.649157] [Medline: 34395355]

XSL•FO
RenderX

40. Phillips EA, Gordeev VS, Schreyögg J. Effectiveness of occupational e-mental health interventions: a systematic review and meta-analysis of randomized controlled trials. Scand J Work Environ Health 2019 Nov 01;45(6):560-576 [FREE Full text] [doi: 10.5271/sjweh.3839] [Medline: 31184758]
41. Carolan S, de Visser RO. Employees' perspectives on the facilitators and barriers to engaging with digital mental health interventions in the workplace: qualitative study. JMIR Ment Health 2018 Jan 19;5(1):e8 [FREE Full text] [doi: 10.2196/mental.9146] [Medline: 29351900]
42. Matthews LR, Gerald J, Jessup GM. Exploring men's use of mental health support offered by an Australian Employee Assistance Program (EAP): perspectives from a focus-group study with males working in blue- and white-collar industries. Int J Ment Health Syst 2021 Aug 04;15(1):68 [FREE Full text] [doi: 10.1186/s13033-021-00489-5] [Medline: 34348756]
43. Peters D, Deady M, Glozier N, Harvey S, Calvo RA. Worker preferences for a mental health app within male-dominated industries: participatory study. JMIR Ment Health 2018 Apr 25;5(2):e30 [FREE Full text] [doi: 10.2196/mental.8999] [Medline: 29695371]
44. Smail-Crevier R, Powers G, Noel C, Wang J. Health-related internet usage and design feature preference for e-mental health programs among men and women. J Med Internet Res 2019 Mar 18;21(3):e11224 [FREE Full text] [doi: 10.2196/11224] [Medline: 30882361]
45. Chan CM, Ng SL, In S, Wee LH, Siau CS. Predictors of psychological distress and mental health resource utilization among employees in Malaysia. Int J Environ Res Public Health 2021 Jan 04;18(1):314 [FREE Full text] [doi: 10.3390/ijerph18010314] [Medline: 33406714]
46. Van Eerd D, Cullen K, Irvin E, Le Poués ard M, Gignac M. Support for depression in the workplace: perspectives of employees, managers, and OHS personnel. Occup Health Sci 2021 Jun 07;5(3):307-343. [doi: 10.1007/S41542-021-00090-9]
47. Horan KA. You, me, and OHP: applying the sociocultural self model to OHP interventions. Occup Health Sci 2021 Jan 27;5(1-2):25-53. [doi: 10.1007/S41542-020-00079-W]
48. Robroek SJ, van Lenthe FJ, van Empelen P, Burdorf A. Determinants of participation in worksite health promotion programmes: a systematic review. Int J Behav Nutr Phys Act 2009 May 20;6:26 [FREE Full text] [doi: 10.1186/1479-5868-6-26] [Medline: 19457246]
49. Meyermann A, Porzelt M. Hinweise zur Anonymisierung von qualitativen Daten. 1st ed.: Forschungsdatenzentrum (FDZ) Bildung am DIPF. Deutsches Institut für Internationale Pädagogische Forschung. 2014. URL: https://www.pedocs.de/volltexte/2022/21968/pdf/fdb-informiert_1_Meyermann_ua_Hinweise_zur_Anonymisierung_von_qualitativen_Daten_2014_v1-1_A.pdf [accessed 2024-04-29]
50. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
51. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
52. Saldaña J. The Coding Manual for Qualitative Researchers. 4th edition. Thousand Oaks, CA: Sage Publications; 2021.
53. Nelson LA, Pennings JS, Sommer EC, Popescu F, Barkin SL. A 3-item measure of digital health care literacy: development and validation study. JMIR Form Res 2022 Apr 29;6(4):e36043 [FREE Full text] [doi: 10.2196/36043] [Medline: 35486413]
54. Yamazaki S, Fukuhara S, Green J. Usefulness of five-item and three-item mental health inventories to screen for depressive symptoms in the general population of Japan. Health Qual Life Outcomes 2005 Aug 08;3:48 [FREE Full text] [doi: 10.1186/1477-7525-3-48] [Medline: 16083512]
55. Strand BH, Dalgard OS, Tambs K, Rognerud M. Measuring the mental health status of the Norwegian population: a comparison of the instruments SCL-25, SCL-10, SCL-5 and MHI-5 (SF-36). Nord J Psychiatry 2003;57(2):113-118. [doi: 10.1080/08039480310000932] [Medline: 12745773]
56. Bullinger M. Der deutsche SF-36 Health Survey. Übersetzung und psychometrische Testung eines krankheitsübergreifenden Instruments zur Erfassung der gesundheitsbezogenen Lebensqualitat. Z. f. Gesundheitswiss. 1995. URL: https://ci.nii.ac.jp/naid/10030663582/ [accessed 2024-04-29]
57. Rumpf HJ, Meyer C, Hapke U, John U. Screening for mental health: validity of the MHI-5 using DSM-IV Axis I psychiatric disorders as gold standard. Psychiatry Res 2001 Dec 31;105(3):243-253. [doi: 10.1016/s0165-1781(01)00329-8] [Medline: 11814543]
58. De Lorenzo MS. Employee mental illness: managing the hidden epidemic. Empl Responsib Rights J 2013 Jul 5;25(4):219-238. [doi: 10.1007/s10672-013-9226-x]
59. van Hees SG, Carlier BE, Vossen E, Blonk RW, Oomens S. Towards a better understanding of work participation among employees with common mental health problems: a systematic realist review. Scand J Work Environ Health 2022 Apr 01;48(3):173-189 [FREE Full text] [doi: 10.5271/sjweh.4005] [Medline: 34878557]
60. KMU-Definition des IfM Bonn. Institut für Mittelstandsforschung Bonn. 2016. URL: https://www.ifm-bonn.org/definitionen-/kmu-definition-des-ifm-bonn [accessed 2024-04-29]
61. International standard classification of occupations: structure, group definitions and correspondence tables. International Labor Organization. 2012. URL: https://www.ilo.org/public/english/bureau/stat/isco/docs/publication08.pdf [accessed 2024-04-29]

XSL•FO
RenderX

62. Global Industry Classification Standard (GICS®) methodology. Guiding principles and methodology for GICS. MSCI. 2023. URL: https://www.msci.com/index/methodology/latest/GICS [accessed 2024-04-29]
63. Jöns I, Hodapp M, Weiss K. Kurzskala zur Erfassung der Unternehmenskultur. The Leibniz Institute for Psychology (ZPID). 2005. URL: https://doi.org/10.23668/psycharchives.11386 [accessed 2024-04-29]
64. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425-478. [doi: 10.2307/30036540]
65. Wixom BH, Todd PA. A theoretical integration of user satisfaction and technology acceptance. Inf Syst Res 2005 Mar;16(1):85-102. [doi: 10.1287/isre.1050.0042]
66. Ergebnisse des Mikrozensus 2019. Statistisches Bundesamt. 2019. URL: https://www.destatis.de/DE/Themen/Arbeit/Arbeitsmarkt/Erwerbstaetigkeit/Tabellen/EWT-Corona-Bildung.html [accessed 2024-04-29]
67. Genesis Online Datenbank: Bevölkerung, Erwerbstätige, Erwerbslose, Erwerbspersonen, Nichterwerbspersonen aus Hauptwohnsitzhaushalten: Deutschland, Jahre, Geschlecht, Altersgruppen. Statistisches Bundesamt. 2023. URL: https://www-genesis.destatis.de/genesis/online?sequenz=tabelleErgebnis&selectionname=12211-0001&startjahr=2020#abreadcrumb [accessed 2024-04-29]
68. Blanz M. Forschungsmethoden und Statistik für die Soziale Arbeit: Grundlagen und Anwendungen. 2nd edition. Stuttgart, Germany: Kohlhammer Verlag; 2021.
69. Jhangiani RS, Chiang IC, Cuttler C, Leighton DC. Research Methods in Psychology. 4th edition. Surrey, BC: Kwantlen Polytechnic University; 2019.
70. Jamieson S. Likert scales: how to (ab)use them. Med Educ 2004 Dec;38(12):1217-1218 [FREE Full text] [doi: 10.1111/j.1365-2929.2004.02012.x] [Medline: 15566531]
71. Streubert HJ, Carpenter DR. Qualitative Research in Nursing: Advancing the Humanistic Imperative. 5th edition. Philadelphia, PA: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2011.
72. Cortina JM. What is coefficient alpha? An examination of theory and applications. J Appl Psychol 1993 Feb;78(1):98-104. [doi: 10.1037/0021-9010.78.1.98]
73. Kohl F, Angerer P, Weber J. Employees' preferences on organisational aspects of psychotherapeutic consultation at work by occupational area, company size, requirement levels and supervisor function - a cross-sectional study in Germany. BMC Public Health 2023 Feb 16;23(1):347 [FREE Full text] [doi: 10.1186/s12889-023-15255-0] [Medline: 36797723]
74. Musiat P, Goldstone P, Tarrier N. Understanding the acceptability of e-mental health--attitudes and expectations towards computerised self-help treatments for mental health problems. BMC Psychiatry 2014 Apr 11;14:109 [FREE Full text] [doi: 10.1186/1471-244X-14-109] [Medline: 24725765]
75. Eichenberg C, Wolters C, Brähler E. The internet as a mental health advisor in Germany--results of a national survey. PLoS One 2013;8(11):e79206 [FREE Full text] [doi: 10.1371/journal.pone.0079206] [Medline: 24278121]
76. Phillips EA, Himmler SF, Schreyögg J. Preferences for e-mental health interventions in Germany: a discrete choice experiment. Value Health 2021 Mar;24(3):421-430 [FREE Full text] [doi: 10.1016/j.jval.2020.09.018] [Medline: 33641777]
77. Apolinário-Hagen J, Harrer M, Kählke F, Fritsche L, Salewski C, Ebert DD. Public attitudes toward guided internet-based therapies: web-based survey study. JMIR Ment Health 2018 May 15;5(2):e10735 [FREE Full text] [doi: 10.2196/10735] [Medline: 29764797]
78. Becker D. Acceptance of mobile mental health treatment applications. Procedia Comput Sci 2016;98:220-227. [doi: 10.1016/j.procs.2016.09.036]
79. Borghouts J, Eikey E, Mark G, De Leon C, Schueller SM, Schneider M, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. J Med Internet Res 2021 Mar 24;23(3):e24387 [FREE Full text] [doi: 10.2196/24387] [Medline: 33759801]
80. Murphy JK, Noble JM, Chakraborty PA, Michlig G, Michalak EE, Greenshaw AJ, et al. Values and preferences related to workplace mental health programs and interventions: an international survey. PLoS One 2023;18(9):e0283057 [FREE Full text] [doi: 10.1371/journal.pone.0283057] [Medline: 37756304]
81. Foulkes L, Andrews JL. Are mental health awareness efforts contributing to the rise in reported mental health problems? A call to test the prevalence inflation hypothesis. New Ideas Psychol 2023 Apr;69:101010. [doi: 10.1016/j.newideapsych.2023.101010]
82. Franssen G. The celebritization of self-care: the celebrity health narrative of Demi Lovato and the sickscape of mental illness. Eur J Cult Stud 2019 Jul 27;23(1):89-111. [doi: 10.1177/1367549419861636]
83. Edwards AM, Kotera Y. Mental health in the UK Police force: a qualitative investigation into the stigma with mental illness. Int J Ment Health Addiction 2021;19(4):1116-1134. [doi: 10.1007/s11469-019-00214-x]
84. Kotera Y, Green P, Sheffield D. Mental health shame of UK construction workers: relationship with masculinity, work motivation, and self-compassion. J Work Organ Psychol 2019;35(2):135-143. [doi: 10.5093/jwop2019a15]
85. Kotera Y, Adhikari P, Van Gordon W. Motivation types and mental health of UK hospitality workers. Int J Ment Health Addict 2018;16(3):751-763 [FREE Full text] [doi: 10.1007/s11469-018-9874-z] [Medline: 29904331]
86. Bethune S. Gen Z more likely to report mental health concerns. Monit Psychol 2019;50(1):1-20 [FREE Full text]

XSL•FO
RenderX

87. Doll CM, Michel C, Rosen M, Osman N, Schimmelmann BG, Schultze-Lutter F. Predictors of help-seeking behaviour in people with mental health problems: a 3-year prospective community study. BMC Psychiatry 2021 Sep 03;21(1):432 [FREE Full text] [doi: 10.1186/s12888-021-03435-4] [Medline: 34479537]
88. Mackenzie CS, Heath PJ, Vogel DL, Chekay R. Age differences in public stigma, self-stigma, and attitudes toward seeking help: a moderated mediation model. J Clin Psychol 2019 Dec;75(12):2259-2272. [doi: 10.1002/jclp.22845] [Medline: 31385298]
89. Michel C, Schnyder N, Schmidt SJ, Groth N, Schimmelmann BG, Schultze-Lutter F. Functioning mediates help-seeking for mental problems in the general population. Eur Psychiatry 2018 Oct;54:1-9. [doi: 10.1016/j.eurpsy.2018.06.009] [Medline: 30031990]
90. Uncovska M, Freitag B, Meister S, Fehring L. Rating analysis and BERTopic modeling of consumer versus regulated mHealth app reviews in Germany. NPJ Digit Med 2023 Jun 21;6(1):115 [FREE Full text] [doi: 10.1038/s41746-023-00862-3] [Medline: 37344556]
91. Perrault EK, Hildenbrand GM, Rnoh RH. Employees' refusals to participate in an employer-sponsored wellness program: barriers and benefits to engagement. Compens Benefits Rev 2020 Jan 07;52(1):8-18. [doi: 10.1177/0886368719899209]
92. Andrade C. The limitations of online surveys. Indian J Psychol Med 2020 Nov;42(6):575-576 [FREE Full text] [doi: 10.1177/0253717620957496] [Medline: 33354086]
93. Simblett S, Matcham F, Siddi S, Bulgari V, Barattieri di San Pietro C, Hortas López J, RADAR-CNS Consortium. Barriers to and facilitators of engagement with mHealth technology for remote measurement and management of depression: qualitative analysis. JMIR Mhealth Uhealth 2019 Jan 30;7(1):e11325 [FREE Full text] [doi: 10.2196/11325] [Medline: 30698535]
94. Sheeran P, Webb TL. The intention–behavior gap. Front Psychol 2016 Sep 05;10(9):503-518. [doi: 10.1111/spc3.12265]

Abbreviations

CHERRIES: Checklist for Reporting Results of Internet E-Surveys
COREQ: Consolidated Criteria for Reporting Qualitative Research
DV: dependent variable
EAP: employee assistance program
EMHP: employee mental health program
GI: general information
MHI: Mental Health Inventory
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews
PV: predicting variable
RQ: research question
WHO: World Health Organization

Edited by A Kushniruk, E Borycki; submitted 25.08.24; peer-reviewed by A Trivedi, S Kunkle; comments to author 25.02.25; revised version received 17.03.25; accepted 19.03.25; published 05.05.25.

Please cite as:
Sevov B, Huettemann R, Zinner M, Meister S, Fehring L
Employee Preference and Use of Employee Mental Health Programs: Mixed Methods Study
JMIR Hum Factors 2025;12:e65750
URL: https://humanfactors.jmir.org/2025/1/e65750
doi:10.2196/65750
PMID:

©Benedict Sevov, Robin Huettemann, Maximillian Zinner, Sven Meister, Leonard Fehring. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 05.05.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

mHealth Use, Preferences, Barriers, and eHealth Literacy Among Patients With Inflammatory Bowel Disease: Survey Study

Christopher Kretzschmar[1,2], MD; Johannes Knitza[2,3,4,5], MHBA, MD; Robert Pietschner[1,2], MD; Raja Atreya[1,2], MD, Prof Dr; Markus Friedrich Neurath[1,2], MD, Prof Dr; Till Orlemann[1,2], MD

[1]Department of Internal Medicine 1, University Hospital Erlangen, Friedrich-Alexander University Erlangen-Nürnberg, Erlangen, Germany

[2]Deutsches Zentrum Immuntherapie, Friedrich-Alexander University Erlangen-Nürnberg and University Hospital Erlangen, Erlangen, Germany

[3]Department of Internal Medicine 3, University Hospital Erlangen, Friedrich-Alexander University Erlangen-Nürnberg, Erlangen, Germany

[4]AGEIS, Université Grenoble Alpes, Grenoble, France

[5]Institute for Digital Medicine, University Hospital of Giessen and Marburg, Philipps University Marburg, Marburg, Germany

Corresponding Author:
Till Orlemann, MD
Department of Internal Medicine 1
University Hospital Erlangen
Friedrich-Alexander University Erlangen-Nürnberg
Ulmenweg 18
Erlangen, 91054
Germany
Phone: 49 09131 ext 8545204
Email: till.orlemann@uk-erlangen.de

Abstract

Background: Mobile health (mHealth), defined as health care facilitated by mobile devices, offers a promising strategy for enhancing disease management and treatment for patients with chronic conditions. However, there is limited information about how patients with inflammatory bowel disease (IBD) use mHealth and their digital preferences.

Objective: The aim of the study was to investigate the use of mHealth as well as the preferences, obstacles, and eHealth literacy reported by patients with IBD in Germany.

Methods: In April and May 2023, we sequentially enrolled patients diagnosed with IBD, including Crohn disease and ulcerative colitis, to participate in a paper-based survey. The survey included questions on sociodemographic details, health characteristics, mHealth use, internet use, eHealth literacy (measured with the eHealth Literacy Scale), and preferences regarding communication and information.

Results: Of the 200 surveyed participants, almost all (197/200, 98.5%) reported regular smartphone use, and more than two-thirds (139/200, 69.5%) indicated regular engagement with social media. Most of the respondents (168/200, 84%) expressed the belief that incorporating medical apps into their routine could positively impact their health. However, only 25 (12.5%) of the 200 patients acknowledged using medical apps, of which just 2 apps were IBD specific, used by only a few (n=3, 12%). Furthermore, awareness of useful websites or mobile apps tailored for IBD was limited (45/200, 22.5%). Nearly all participants (196/200, 98%) expressed willingness to share app data for research purposes, and most (171/200, 85.5%) consented to transmit app data to their treating physicians. A large majority (175/200, 87.5%) indicated readiness to regularly input data into an app, with a preferred duration of up to 5 minutes (109/200, 54.5%) and weekly input frequency (76/200, 38%). For an IBD-specific app, the most frequently requested functions were electronic prescriptions (110/200, 55%) and a newsletter about new scientific work and clinical studies (94/200, 47%). Usability and security were identified as key app attributes. The internet was the predominant source of health-related information (180/200, 90%). The average eHealth literacy score, measured with the eHealth Literacy Scale, was high (mean 28.9, SD 5.4; range 8-40), with a positive correlation observed between higher eHealth literacy and factors such as younger age and more frequent internet use for health information.

Conclusions: Patients with IBD are well prepared and motivated to use mHealth technologies to better understand their chronic condition and optimize treatment. However, their enthusiasm is tempered by the currently low adoption of mHealth. To fully harness the potential of mHealth in IBD treatment, effective and tailored mHealth solutions, guidance for their implementation, and patient education are needed.

XSL•FO
RenderX

(JMIR Hum Factors 2025;12:e64471) doi:10.2196/64471

KEYWORDS

eHealth Literacy Scale; mHealth; mobile apps; inflammatory bowel disease; Crohn disease; ulcerative colitis

Introduction

Background

Inflammatory bowel diseases (IBDs), including Crohn disease and ulcerative colitis, are complex, chronic conditions with multifaceted pathogenesis [1,2]. Patients typically require lifelong treatment from a multidisciplinary specialist team, including gastroenterologists, surgeons, rheumatologists, dermatologists, and radiologists [3]. This care includes, apart from regular medical appointments, immunosuppressive pharmacotherapy to maintain remission and prevent relapses [4], which can pose challenges for patients' treatment adherence [5]. The complexity of IBD not only incurs high health care costs [2] but also makes it difficult for patients to understand their illness and manage daily demands, such as monitoring symptoms, implementing appropriate nutrition, and planning activities [6]. Consequently, IBD represents both a substantial economic burden and a patient-level challenge, underscoring the need for modern disease management strategies [3,7].

In today's world, the use of digital media such as computers, smartphones, tablets, the internet, and mobile apps is ubiquitous and plays an important role in the daily lives of many people. One area increasingly affected by these technologies is health. The internet serves as a gigantic source of information, making health information widely accessible; fitness trackers enable individuals to record and document personal health parameters; and digital services such as email allow patients to communicate with health care providers. Consequently, digital technologies are increasingly incorporated into routine health care [8-10]. The concept of mobile health (mHealth), defined by the World Health Organization as "medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants (PDAs), and other wireless devices" [11], is a promising approach for managing various chronic diseases, including rheumatoid arthritis [12], diabetes mellitus [13], hypertension [14], and chronic heart failure [15]; for example, electronic medication reminders can improve treatment adherence [16,17], and supportive digital therapy can reduce pain [17] and improve key comorbidities (eg, depression) [18]. In the context of IBD, mHealth has the potential to take disease management to a new level, including improving patients' quality of life and reducing health costs [19]. However, access to mHealth solutions remains extremely limited for patients with chronic IBD [20].

Objectives

The perspectives of patients with IBD on mHealth solutions remain largely unexplored. To effectively incorporate various mHealth solutions into clinical practice, it is essential to understand patients' limitations and needs. This study aimed to examine the use of mHealth solutions as well as the preferences, obstacles, and eHealth literacy reported by patients with chronic IBD in Germany.

Methods

Paper-Based Survey

In April and May 2023, consecutive patients treated at the outpatient clinic for chronic IBD at the University Hospital Erlangen were invited to complete a paper-based survey. This format was intentionally chosen to minimize selection bias by enabling participation regardless of digital literacy or device access.

To create the survey (Multimedia Appendix 1), a broad literature review was performed. Previous mHealth patient surveys in rheumatology served as a starting point [21]. The survey comprised four main sections: (1) sociodemographic and health characteristics, (2) preferences and use of mHealth, (3) eHealth literacy, and (4) communication and information preferences. Foreign terms and explanations of technical terms were provided in footnotes. Inclusion criteria were as follows: patients had to (1) be aged ≥18 years, (2) be proficient in the German language, (3) be physically and mentally capable of completing a structured questionnaire, and (4) have a confirmed diagnosis of either Crohn disease or ulcerative colitis.

The sociodemographic and health characteristics included age, sex, place of residence, educational qualification, diagnosis, duration of illness, patients' global self-assessment of disease activity, and current use of electronic devices (smartphones, tablets, and activity trackers) and social media.

The section on preferences and use of mHealth contained questions about the use of medical apps in general, knowledge and use of IBD-specific apps, and the preferred duration and frequency of actively using an IBD-specific app. Patients were asked to evaluate their preferences for app features on a 5-point Likert scale and the importance of app characteristics on a 10-point Likert scale. In addition, patients' willingness to share recorded app data and overall perceptions of the benefits of medical apps were investigated.

Furthermore, information was gathered on internet use and the perceived utility of the internet regarding health decisions, as well as on web-based services to enhance patient health and on telemedicine, with a focus on patients' digital interaction preferences.

Measurement of the patients' eHealth literacy was carried out using the validated German version [22] of the eHealth Literacy Scale (eHEALS) [23]. This tool has been translated and validated in multiple languages [23-28] and enables systematic capture of differences in eHealth literacy levels. It consists of 8 statements regarding eHealth literacy that are rated on a 5-point Likert scale ranging from 1 ("strongly disagree") to 5 ("strongly agree"), yielding a total score ranging from 8 to 40.

Using ranking lists (unique rankings were required), patients indicated their preferences regarding medication reminders, the format of medical information, the design of digitally provided

XSL•FO
RenderX

information, the type of patient diary, and the mode of physician communication.

Statistical Analysis

The examined characteristics were summarized using counts, percentages, means, and SDs. To investigate relationships between continuous variables, Pearson correlation was used. The relationship between the eHEALS score and the frequency of internet use was investigated using the Jonckheere-Terpstra test, with eHEALS score as the continuous dependent variable and internet use frequency as the ordinal independent variable. Relationships between the eHEALS score and binary preferences were examined using logistic regression. The applied statistical tests were carefully selected according to the measurement scales and distributional characteristics of the collected data. All models included adjustments for age and sex. Two-tailed *P* values <.05 were considered statistically significant. Microsoft Excel and SPSS (version 28.0; IBM Corp) were used for data processing and analysis.

Ethical Considerations

This study received approval from the ethics committee of Friedrich-Alexander University Erlangen-Nürnberg (23-63-B) and was conducted in accordance with good clinical practice. Before inclusion in the study, all participating patients provided informed consent for participation, pseudonymized data storage, and processing of the collected data. All participant's data were collected, stored, and analyzed in accordance with applicable privacy regulations. Personal information was pseudonymized, and confidentiality was strictly maintained throughout the study. Participating patients did not receive any form of financial or other compensation for their involvement in the study.

Results

Patient Characteristics

A total of 220 patients were recruited for the study. For the final data analysis, only complete surveys from patients who met all inclusion criteria (200/220, 91%) were considered. The number of patients who declined participation was not recorded. The sociodemographic data of the study sample are presented in Table 1. The mean age was 39.1 (SD 14.0) years, with 49% (98/200) of the patients aged between 18 and 35 years. Sex distribution was nearly balanced. Regarding the use of electronic devices and social media, almost all patients regularly used a smartphone (197/200, 98.5%), more than half (108/200, 54%) regularly used a tablet, nearly one-third (64/200, 32%) used an activity tracker, and more than two-thirds (139/200, 69.5%) regularly used social media.

XSL•FO
RenderX

Table 1. Demographic and health characteristics (n=200).

Characteristics	Participants
Age (years), mean (SD)	39.1 (14.0)
Age (years), n (%)	
18-35	98 (49)
36-45	39 (19.5)
46-55	28 (14)
56-65	26 (13)
>65	9 (4.5)
Sex, n (%)	
Female	96 (48)
Male	104 (52)
Diagnosis, n (%)	
Crohn disease	127 (63.5)
Ulcerative colitis	73 (36.5)
Patient global self-assessment of disease activity (on a scale ranging from 0 to 10), mean (SD)	2.3 (2.3)
Disease duration (years), mean (SD)	13.1 (9.9)
Disease duration (years), n (%)	
≤1	8 (4)
2-5	46 (23)
>5	146 (73)
Place of residence, n (%)	
Village	95 (47.5)
Small city	56 (28)
Midsized city	18 (9)
Big city	31 (15.5)
Regular use of devices, social media, and apps, n (%)	
Smartphone	197 (98.5)
Tablet	108 (54)
Activity tracker	64 (32)
Social media	139 (69.5)
Medical apps	25 (12.5)

mHealth: Use, Knowledge, and Acceptance of Medical Apps and Willingness to Provide mHealth Data

A set of questions assessed patients' preferences and attitudes toward potential mHealth apps and the sharing of health data. An overview of the collected data is shown in Multimedia Appendix 2. Regarding app use, only a small proportion of patients (25/200, 12.5%) reported currently using medical apps. Among these, only 2 apps were IBD specific: CED-Forum (2/25, 8%) and an unnamed IBD app (1/25, 4%). Patients' knowledge of useful IBD-specific digital services (websites and apps) was generally low (45/200, 22.5%). Notwithstanding these findings, more than three-fourths of the patients (168/200, 84%) believed that using medical apps could be beneficial for their health. The potential association between eHEALS scores and the likelihood of expressing the belief that apps are beneficial was investigated using logistic regression adjusted for age and sex. It turned out that our model, using the eHEALS score as predictor, fit poorly, as indicated by the very low explained variance (Nagelkerke R^2=0.09), suggesting that eHEALS scores alone account for little of the variation in patients' perceptions of app benefits.

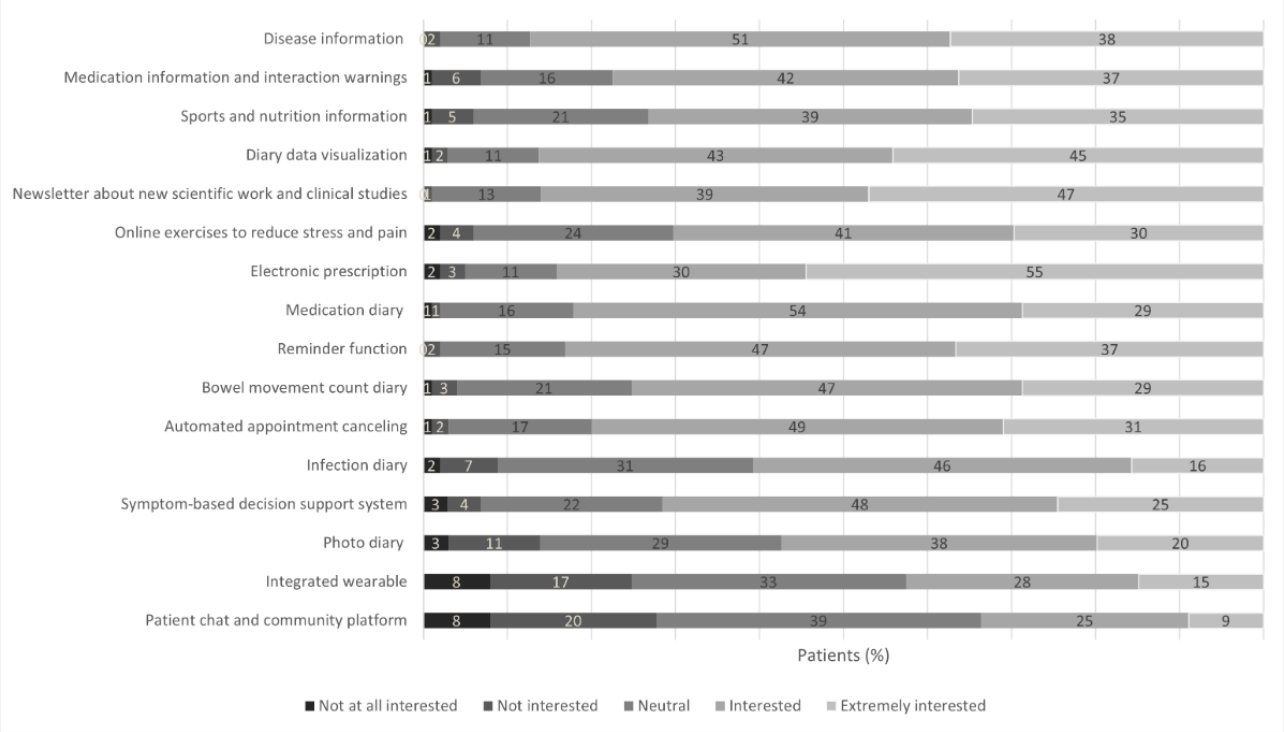
Almost all patients (196/200, 98%) indicated willingness to transmit app data for scientific research, on condition that data security was guaranteed. Similarly, most of the patients (171/200, 85.5%) agreed to share app data with their physician. Among those who did not agree (29/200, 14.5%), the main reason was the perception that personal contact with their physician was sufficient (17/29, 59%). Other reasons included concerns about not knowing where their data would be stored

XSL•FO
RenderX

(12/29, 41%), unclear use of the data (11/29, 38%), and insecure data transmission (9/29, 31%). Whereas the majority of patients (188/200, 94%) wished to be notified if an app detected irregularities regarding their health, interest in comparing medication adherence with other patients was low (57/200, 28.5%).

A large proportion of the participants (175/200, 87.5%) were willing to actively input data into an IBD-specific app. Among all patients, weekly data entry was preferred by 38% (76/200), with a duration of 1 to 5 minutes (109/200, 54.5%). When asked to evaluate key attributes of an app on a 10-point Likert scale, patients rated usability (mean 9.0, SD 1.4) and security (mean 8.4, SD 2.6) as the most important (Multimedia Appendix 2). Regarding preferred app features, patients were most interested in electronic prescriptions and a newsletter about new scientific work and clinical studies. By contrast, patients were least interested in direct peer exchanges, such as chats, or integrated wearables (Figure 1).

Figure 1. Patients' app function preferences (responses to "What app functions would you like?").



Internet Use, Perceived Benefit, and Digital Interaction

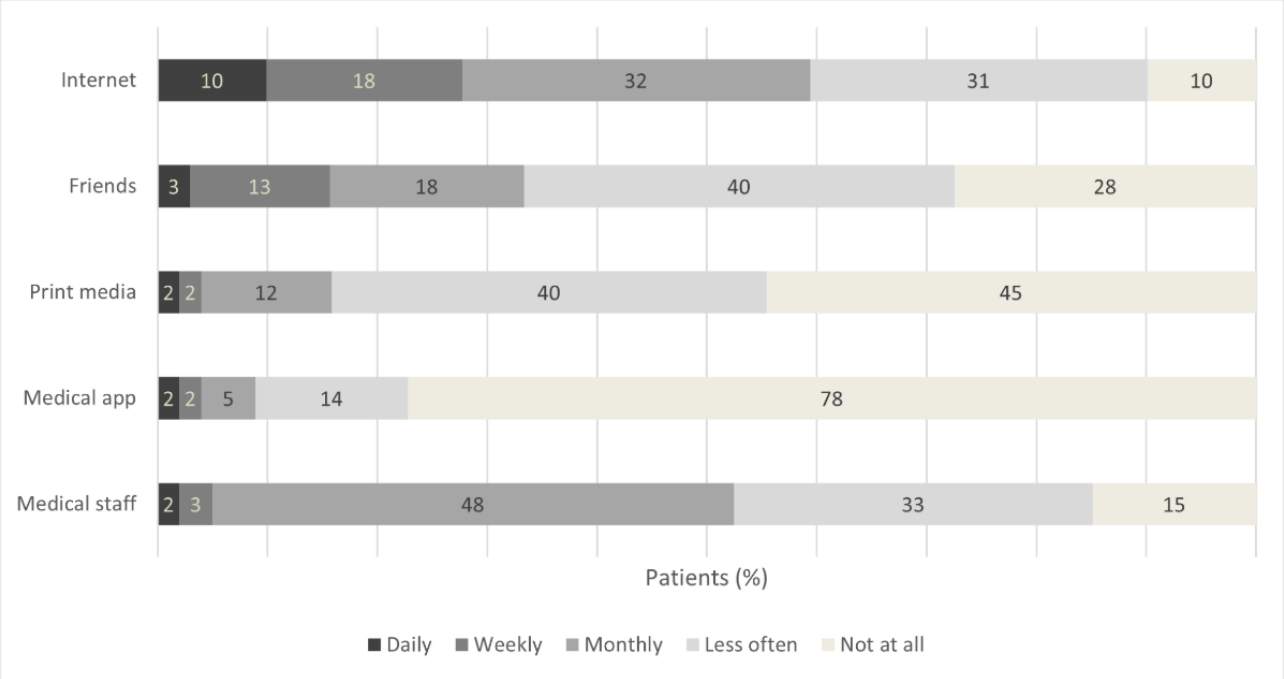
Another important aspect of our study was to obtain knowledge about patients' internet use, especially regarding health-related information. Almost all patients (193/200, 96.5%) stated that they had used the internet to seek health information, whereas only a small proportion (7/200, 3.5%) had not (Multimedia Appendix 3). The reasons for not seeking health information on the internet included perceiving their physician's information as sufficient (4/7, 57%), considering internet searches unhelpful (3/7, 43%), and other reasons (1/7, 14%). The types of health information patients sought on the internet are illustrated in Table 2. Most patients searched for information about treatment options (165/193, 85.5%) and symptoms (164/193, 85%), whereas information on support groups was least frequently sought (85/193, 44%). Regarding the frequency of internet use for obtaining health information, the most common response was "less often than monthly" (79/200, 39.5%). Comparing the use of different health information sources during the past 3 months (Figure 2), the internet was the most frequently used source (180/200, 90%), with 10% (20/200) using it daily, 17.5% (35/200) weekly, 31.5% (63/200) monthly, and 31% (62/200) less often. More than half of the patients (116/200, 58%) had previously used online support groups to read information (113/200, 56.5%), chat with other patients (17/200, 8.5%), or post information (11/200, 5.5%). Participation in an online health program was infrequent (9/200, 4.5%), and knowledge about the medication website of the German competence network for IBDs was not widespread (29/200, 14.5%).

XSL•FO
RenderX

Table 2. Health information previously sought on the internet (n=193; responses to "What health information did you look for on the internet?").

Health information	Patients, n (%)
Treatment options	165 (85.5)
Symptoms	164 (85)
Medication	158 (82)
Diagnosis	144 (75)
Physicians	142 (74)
Disease-specific websites	93 (48)
Support groups	85 (44)
Other	4 (2)

Figure 2. Use of different health information sources (responses to "How often did you use one of the following options to receive health information in the last three months?").



Regarding the perceived benefit of using the internet to make health-related decisions, results were ambiguous with approximately the same number of patients being unsure (81/200, 40.5%) and recognizing the internet as useful (79/200, 39.5%; Multimedia Appendix 3).

Another aspect investigated was patients' digital interaction with their treating physician. Slightly more than half of the patients (105/200, 52.5%) reported having previously communicated with a physician via email. In addition, 71% (142/200) of the patients indicated that they would appreciate the option to communicate with their gastroenterologist via video call. The majority of the patients (159/200, 79.5%) favored electronically filling out medical questionnaires before their appointments, and more than two-thirds (137/200, 68.5%) preferred receiving an electronic medical report rather than a paper report.

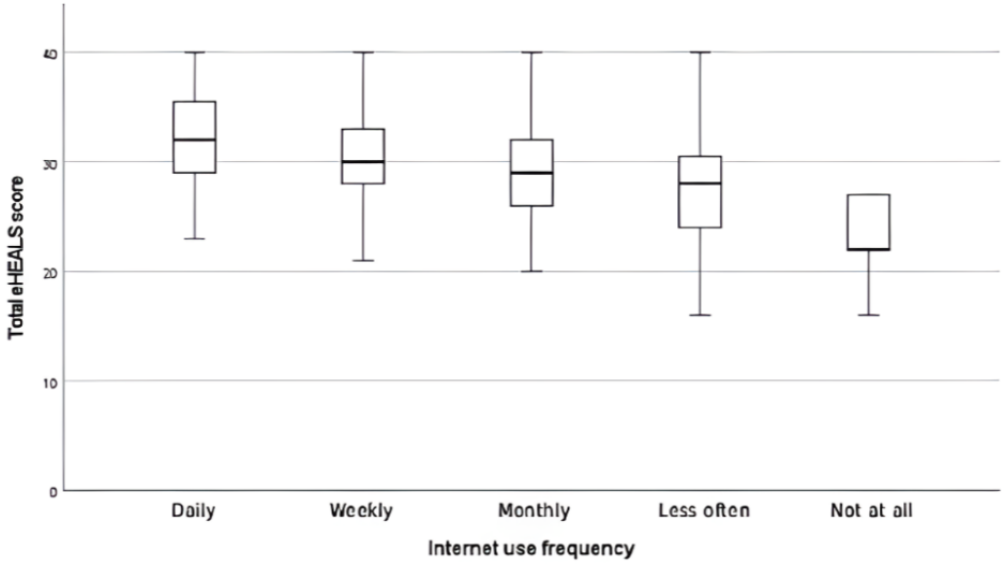
eHealth Literacy

eHealth literacy, as defined by Norman and Skinner [23] is "the ability to seek, find, understand, and appraise health information from electronic sources and apply the knowledge gained to addressing or solving a health problem." Therefore, eHealth literacy is an important dimension to record patients' handling of digitally provided health information. The mean eHealth literacy score of our patients, measured using the eHEALS tool, was 28.9 (SD 5.4; range 8-40). Male patients had a mean score of 29.0, and female patients had a mean score of 28.3, resulting in a mean difference of 0.7; this difference was not statistically significant (P=.35), as verified by an unpaired, 2-tailed t test (95% CI −0.74 to 2.07). By contrast, age was found to correlate negatively (r=−0.2, 95% CI −0.35 to −0.04) with the eHEALS score. An overview of the distribution of responses to the 8 eHEALS items is provided in Multimedia Appendix 4. The majority of patients agreed or strongly agreed that they know how to use the internet to answer health-related questions (152/200, 76%) and have the skills to evaluate the health resources found on the internet (160/200, 80%); however, concerns were raised regarding the knowledge of available health resources on the internet (41/200, 20.5), and approximately one-third of the patients (69/200, 34.5%)

XSL•FO
RenderX

expressed discomfort about using information from the internet to make health-related decisions. Lower eHEALS scores were found to be associated with a declining frequency of internet use (Figure 3). A Jonckheere-Terpstra test showed a statistically significant trend between lower eHEALS scores and decreasing internet use frequency (T_{JT}=4861.5, z=−4.958; P<.001).

Figure 3. Association of eHealth Literacy Scale (eHEALS) score and frequency of internet use.



Preferences in Communication

Almost all patients (198/200, 99%) wished for the opportunity to raise queries about complaints and treatment between routine medical visits. Preferred methods of contacting the treating gastroenterologist were telephone (81/200, 40.5%), email (61/200, 30.5%), and chat (56/200, 28%; Multimedia Appendix 5). Reminders for regular medication intake were requested frequently (172/200, 86%), with the favored mode being via an app (117/200, 58.5%). The vast majority of patients (196/200, 98%) wished to receive medical information, preferably through an app (92/200, 46%), followed by paper (62/200, 31%) and websites (42/200, 21%). If medical information was provided digitally, patients ranked simple text and images (172/200, 86%) over features based on exchanging information with other patients (20/200, 10%) and game-based learning (8/200, 4%). Acceptance of using a patient diary was high (176/200, 88%), with documentation preference being greatest for an app, followed by paper, and lowest for a website (Multimedia Appendix 5).

Discussion

Principal Findings and Comparison With Prior Work

To the best of our knowledge, this study is the first to provide an in-depth perspective on German patients' views of mHealth solutions in gastroenterology, specifically in the context of IBD. Our findings emphasize that patients with IBD are interested in mHealth technologies and are willing and motivated to use them. For this intention to be realized, two factors are crucial: the availability of the necessary technical requirements, which was found to be remarkably high in our sample, with almost all patients (197/200, 98.5%) regularly using a smartphone and more than half (108/200, 54%) regularly using a tablet; and the possession of digital skills. On the basis of participants' internet use behavior and eHealth literacy levels, it can be assumed that patients with IBD possess these skills. Nearly all patients (193/200, 96.5%) reported using the internet for health matters, underscoring its role as a critical source of health information. Patients in our study demonstrated the ability to find, evaluate, and use such information, as reflected by a mean eHEALS score of 28.9 (SD 5.4; range 8-40). This score indicates high eHealth literacy, considering that a previous study defined a cutoff for high eHealth literacy at ≥26 [29]. Numerous studies in various fields [21,29-33] have examined possible correlations between the dimensions of eHealth literacy and patient characteristics and behaviors. In our study of patients with IBD, younger age was associated with higher eHealth literacy, whereas no effect of sex was observed. Consistent with prior studies [21,29,30,33], lower eHealth literacy was associated with declining internet use frequency. Interestingly, belief in the usefulness of medical apps did not seem to depend on eHealth literacy levels but was overall high, highlighting patients' widespread confidence in the potential of mHealth.

To fully realize this potential, most patients were willing to participate actively by inputting personal health data into IBD-specific apps, sharing them with their treating physicians, and providing them for research purposes. Such active participation ensures electronic data collection of a high standard, both in terms of quality and quantity, which can be of considerable benefit to patients and physicians alike. Prior studies [34-36] have also highlighted the usefulness of patient-generated health data. In addition, patients with IBD expressed clear preferences for digital interaction with their treating physicians, including consultations via video calls, filling out questionnaires online before appointments, and receiving electronic medical reports. These preferences could contribute to improved care by optimizing treatment processes. Despite the overall high enthusiasm for mHealth among patients with IBD, some preferred personal contact with their physician, as reflected by the highest preference for telephone calls for

XSL•FO
RenderX

communication rather than eHealth services. Therefore, offering flexible communication options tailored to patient preferences is important. Furthermore, some of the patients (29/200, 14.5%) expressed reluctance to share data with their treating physicians due to concerns about security and transparency. These concerns must be taken seriously. The responsibility to guarantee data security lies with the various stakeholders involved and is critical not only to avoid undermining patients' confidence in mHealth solutions but also to strengthen it. Further investigation into which specific data security features would make patients feel safe and strengthen their confidence is crucial. In addition, clear communication about data protection measures and transparent handling of data are essential to foster patient trust and encourage wider mHealth adoption.

mHealth apps have the potential to serve as powerful tools and are therefore of particular interest for research. As noted, patients' belief in the usefulness of medical apps was remarkably high, and they were willing to actively use IBD-specific apps by inputting data for up to 5 minutes on a weekly basis. Desired features for an IBD-specific app differed, but preferences were primarily for those that deliver information, such as a newsletter about new scientific work and clinical studies, as well as electronic prescriptions. By contrast, interest in direct peer exchanges and integrated wearables was low, which is partially consistent with prior findings among patients with gastroenterological diseases [37]. Mobile apps can improve treatment adherence [38,39], and, in our study, patients with IBD expressed interest in features such as medication reminders and an app-based patient diary. Usability and security were also identified as key features of an IBD-specific app. Patient perspectives therefore provide valuable guidance for app design. Particular attention should be given to ensuring an intuitive user experience through simple, user-friendly design. App features should be customizable because preferences differ; however, features such as comprehensible information (eg, through simple text and images), electronic prescriptions, medication reminders, and patient diaries seem to be broadly desirable. In addition, the app should minimize demands on data input and guarantee data security so that patients feel safe using it. In line with previous recommendations [20,40,41], involving patients in the app development process is important, as it can enhance acceptance and use [41].

Despite the enormous interest and readiness of patients with IBD, current use of mHealth solutions remains relatively low. Only 12.5% (25/200) of the patients in our study reported using medical apps, of which just 2 were IBD specific (used by 3/25, 12%). Limited availability of IBD-specific apps, as noted in a previous study by our group [20], and the generally low knowledge of IBD-specific digital services likely represent two important factors contributing to the gap between patient interest and actual adoption. Further factors, such as usability concerns and skepticism regarding effectiveness, as reported in studies from other fields [42,43], require investigation in future research. Overall, the high discrepancy points to the need to develop effective mHealth solutions, supported by guidance and recommendations from gastroenterological specialists and societies regarding suitable and reliable sources and services and supported by qualitative patient education to build confidence in using these services, consistent with prior suggestions [42,44]. However, several challenges must be addressed; for example, strict regulations in Germany under the Digital Health Care Act ("Digitale-Versorgungs-Gesetz"), which classify prescribable health apps as medical devices and impose stringent data security demands [45], pose obstacles to the development and implementation of mHealth services. Nevertheless, translating these needs into practice to align with the interest and readiness of patients with IBD is going to be a crucial task and will require collaboration among all stakeholders to enable the implementation of reliable mHealth programs and apps in routine IBD care, thereby fully realizing the potential of mHealth.

Limitations

The main limitations of this study include its cross-sectional design, which does not allow conclusions regarding the evaluation of patients' digital engagement behaviors over time; reliance on self-reported data, which may be influenced by social desirability bias, potentially leading patients to overestimate their digital competencies or willingness to use mHealth tools; and a relatively small, circumscribed sample (outpatients from a single tertiary care center). Moreover, the relatively young mean age (39.1, SD 14.0 y) of the sample may influence findings, as older patients could exhibit different digital health behaviors and face distinct barriers to mHealth adoption. In addition, the number and characteristics of patients who declined participation were not recorded, limiting the assessment of potential selection bias. Therefore, the generalizability of the results may be compromised, as factors such as actual mHealth use could vary from what participants reported.

Conclusions

To the best of our knowledge, this study is the first to comprehensively explore perspectives on mHealth among German patients with IBD. The insights gained have the potential to significantly influence and shape the development and integration of IBD apps. A vast majority of participants not only believed in the possible health benefits of using medical apps but were also motivated and willing to use mHealth technologies and even participate actively by inputting personal health data into apps and sharing them with physicians and researchers. Furthermore, fundamental requirements for successful mHealth use were largely met, as almost all patients regularly used smartphones and demonstrated considerable digital skills, as indicated by a high level of eHealth literacy. Despite these favorable conditions, current use of mHealth among patients with IBD remains limited, likely due to limited availability and awareness of existing apps, usability concerns, lack of perceived necessity, and skepticism about app effectiveness. To close this gap, it is crucial to develop effective mHealth solutions, supported by guidance and recommendations from gastroenterological specialists and societies regarding suitable and reliable sources and services, along with qualitative education on mHealth use. The unmet needs and patient priorities highlighted in this study provide direction for expediting the incorporation of mHealth into standard IBD care.

XSL•FO
RenderX

Acknowledgments

This work was conducted to fulfill the requirements for CK's doctor of medicine (Dr med) degree. The authors thank Koray Tascilar for assistance with questions related to the statistical analysis.

Authors' Contributions

CK drafted the initial manuscript. CK and TO performed the statistical analysis. All authors reviewed the draft, suggested revisions, and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Original German questionnaire.
[PDF File (Adobe PDF File), 596 KB - humanfactors_v12i1e64471_app1.pdf]

Multimedia Appendix 2
Patients' viewpoints on medical apps.
[PNG File , 83 KB - humanfactors_v12i1e64471_app2.png]

Multimedia Appendix 3
Information on internet use, perceived benefit, knowledge of inflammatory bowel disease–specific digital offers, and eHealth literacy.
[PNG File , 78 KB - humanfactors_v12i1e64471_app3.png]

Multimedia Appendix 4
eHealth literacy measured with the eHealth Literacy Scale.
[PNG File , 50 KB - humanfactors_v12i1e64471_app4.png]

Multimedia Appendix 5
Preferences in communication.
[PNG File , 81 KB - humanfactors_v12i1e64471_app5.png]

References

1. Abraham C, Cho JH. Inflammatory bowel disease. N Engl J Med 2009 Nov 19;361(21):2066-2078. [doi: 10.1056/nejmra0804647]
2. Molodecky NA, Soon IS, Rabi DM, Ghali WA, Ferris M, Chernoff G, et al. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. Gastroenterology 2012 Jan;142(1):46-54. [doi: 10.1053/j.gastro.2011.10.001] [Medline: 22001864]
3. Lamb CA, Kennedy NA, Raine T, Hendy PA, Smith PJ, Limdi JK, IBD guidelines eDelphi consensus group, et al. British Society of Gastroenterology consensus guidelines on the management of inflammatory bowel disease in adults. Gut 2019 Dec;68(Suppl 3):s1-106 [FREE Full text] [doi: 10.1136/gutjnl-2019-318484] [Medline: 31562236]
4. Elhag DA, Kumar M, Saadaoui M, Akobeng AK, Al-Mudahka F, Elawad M, et al. Inflammatory bowel disease treatments and predictive biomarkers of therapeutic response. Int J Mol Sci 2022 Jun 23;23(13):6966 [FREE Full text] [doi: 10.3390/ijms23136966] [Medline: 35805965]
5. King K, Czuber-Dochan W, Chalder T, Norton C. Medication non-adherence in inflammatory bowel disease: a systematic review identifying risk factors and opportunities for intervention. Pharmacy (Basel) 2025 Feb 07;13(1):21 [FREE Full text] [doi: 10.3390/pharmacy13010021] [Medline: 39998019]
6. Lovén Wickman U, Yngman-Uhlin P, Hjortswang H, Riegel B, Stjernman H, Hollman Frisman G. Self-care among patients with inflammatory bowel disease: an interview study. Gastroenterol Nurs 2016;39(2):121-128 [FREE Full text] [doi: 10.1097/SGA.0000000000000120] [Medline: 26166423]
7. Dulai PS, Singh S, Ohno-Machado L, Sandborn WJ. Population health management for inflammatory bowel disease. Gastroenterology 2018 Jan;154(1):37-45. [doi: 10.1053/j.gastro.2017.09.052] [Medline: 29122544]
8. Keesara S, Jonas A, Schulman K. COVID-19 and health care's digital revolution. N Engl J Med 2020 Jun 04;382(23):e82 [FREE Full text] [doi: 10.1056/nejmp2005835]

XSL•FO
RenderX

9. Stoumpos AI, Kitsios F, Talias MA. Digital transformation in healthcare: technology acceptance and its applications. Int J Environ Res Public Health 2023 Feb 15;20(4):3407 [FREE Full text] [doi: 10.3390/ijerph20043407] [Medline: 36834105]
10. Kusnik A, Itzel T, Dooley S, Dropmann A, Stallkamp J, Ganslandt T, et al. Digital gastroenterology. J Gastrointestin Liver Dis 2020 Dec 12;29(4):493-496 [FREE Full text] [doi: 10.15403/jgld-3099] [Medline: 33331341]
11. mHealth: new horizons for health through mobile technologies: second global survey on eHealth. World Health Organization. 2011. URL: https://iris.who.int/handle/10665/44607 [accessed 2023-12-02]
12. Rodríguez Sánchez-Laulhé P, Luque-Romero LG, Barrero-García FJ, Biscarri-Carbonero Á, Blanquero J, Suero-Pineda A, et al. An exercise and educational and self-management program delivered with a smartphone app (CareHand) in adults with rheumatoid arthritis of the hands: randomized controlled trial. JMIR Mhealth Uhealth 2022 Apr 07;10(4):e35462 [FREE Full text] [doi: 10.2196/35462] [Medline: 35389367]
13. Eberle C, Löhnert M, Stichling S. Effectiveness of disease-specific mHealth apps in patients with diabetes mellitus: scoping review. JMIR Mhealth Uhealth 2021 Feb 15;9(2):e23477 [FREE Full text] [doi: 10.2196/23477] [Medline: 33587045]
14. Li R, Liang N, Bu F, Hesketh T. The effectiveness of self-management of hypertension in adults using mobile health: systematic review and meta-analysis. JMIR Mhealth Uhealth 2020 Mar 27;8(3):e17776 [FREE Full text] [doi: 10.2196/17776] [Medline: 32217503]
15. Ware P, Ross HJ, Cafazzo JA, Boodoo C, Munnery M, Seto E. Outcomes of a heart failure telemonitoring program implemented as the standard of care in an outpatient heart function clinic: pretest-posttest pragmatic study. J Med Internet Res 2020 Feb 08;22(2):e16538 [FREE Full text] [doi: 10.2196/16538] [Medline: 32027309]
16. Kang H, Park HA. A mobile app for hypertension management based on clinical practice guidelines: development and deployment. JMIR Mhealth Uhealth 2016 Feb 02;4(1):e12 [FREE Full text] [doi: 10.2196/mhealth.4966] [Medline: 26839283]
17. Yang J, Weng L, Chen Z, Cai H, Lin X, Hu Z, et al. Development and testing of a mobile app for pain management among cancer patients discharged from hospital treatment: randomized controlled trial. JMIR Mhealth Uhealth 2019 May 29;7(5):e12542 [FREE Full text] [doi: 10.2196/12542] [Medline: 31144672]
18. Economides M, Ranta K, Nazander A, Hilgert O, Goldin PR, Raevuori A, et al. Long-term outcomes of a therapist-supported, smartphone-based intervention for elevated symptoms of depression and anxiety: quasiexperimental, pre-postintervention study. JMIR Mhealth Uhealth 2019 Aug 26;7(8):e14284 [FREE Full text] [doi: 10.2196/14284] [Medline: 31452521]
19. Flamant M, Roblin X. Inflammatory bowel disease: towards a personalized medicine. Therap Adv Gastroenterol 2018;11:1756283X17745029 [FREE Full text] [doi: 10.1177/1756283X17745029] [Medline: 29383027]
20. Gerner M, Vuillerme N, Aubourg T, Messner E, Terhorst Y, Hörmann V, et al. Review and analysis of German mobile apps for inflammatory bowel disease management using the mobile application rating scale: systematic search in app stores and content analysis. JMIR Mhealth Uhealth 2022 May 03;10(5):e31102 [FREE Full text] [doi: 10.2196/31102] [Medline: 35503246]
21. Knitza J, Simon D, Lambrecht A, Raab C, Tascilar K, Hagen M, et al. Mobile health usage, preferences, barriers, and eHealth literacy in rheumatology: patient survey study. JMIR Mhealth Uhealth 2020 Aug 12;8(8):e19661 [FREE Full text] [doi: 10.2196/19661] [Medline: 32678796]
22. Soellner R, Huber S, Reder M. The concept of eHealth literacy and its measurement. J Media Psychol 2014 Jan;26(1):29-38. [doi: 10.1027/1864-1105/a000104]
23. Norman CD, Skinner HA. eHEALS: the eHealth literacy scale. J Med Internet Res 2006 Nov 14;8(4):e27 [FREE Full text] [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
24. Chung S, Park BK, Nahm ES. The Korean eHealth Literacy Scale (K-eHEALS): reliability and validity testing in younger adults recruited online. J Med Internet Res 2018 Apr 20;20(4):e138 [FREE Full text] [doi: 10.2196/jmir.8759] [Medline: 29678800]
25. van der Vaart R, van Deursen AJ, Drossaert CH, Taal E, van Dijk JA, van de Laar MA. Does the eHealth Literacy Scale (eHEALS) measure what it intends to measure? Validation of a Dutch version of the eHEALS in two adult populations. J Med Internet Res 2011 Nov 09;13(4):e86 [FREE Full text] [doi: 10.2196/jmir.1840] [Medline: 22071338]
26. Diviani N, Dima AL, Schulz PJ. A psychometric analysis of the Italian version of the eHealth literacy scale using item response and classical test theory methods. J Med Internet Res 2017 Apr 11;19(4):e114 [FREE Full text] [doi: 10.2196/jmir.6749] [Medline: 28400356]
27. Dale JG, Lüthi A, Fundingsland Skaraas B, Rundereim T, Dale B. Testing measurement properties of the Norwegian version of electronic health literacy scale (eHEALS) in a group of day surgery patients. J Multidiscip Healthc 2020;13:241-247 [FREE Full text] [doi: 10.2147/JMDH.S242985] [Medline: 32210568]
28. Ma Z, Wu M. The Psychometric Properties of the Chinese eHealth Literacy Scale (C-eHEALS) in a Chinese rural population: cross-sectional validation study. J Med Internet Res 2019 Oct 22;21(10):e15720 [FREE Full text] [doi: 10.2196/15720] [Medline: 31642811]
29. Richtering SS, Hyun K, Neubeck L, Coorey G, Chalmers J, Usherwood T, et al. eHealth literacy: predictors in a population with moderate-to-high cardiovascular risk. JMIR Hum Factors 2017 Jan 27;4(1):e4 [FREE Full text] [doi: 10.2196/humanfactors.6217] [Medline: 28130203]

XSL•FO
RenderX

30. Wong DK, Cheung MK. Online health information seeking and eHealth literacy among patients attending a primary care clinic in Hong Kong: a cross-sectional survey. J Med Internet Res 2019 Mar 27;21(3):e10831 [FREE Full text] [doi: 10.2196/10831] [Medline: 30916666]
31. Stellefson ML, Shuster JJ, Chaney BH, Paige SR, Alber JM, Chaney JD, et al. Web-based health information seeking and eHealth literacy among patients living with chronic obstructive pulmonary disease (COPD). Health Communication 2017 Sep 05;33(12):1410-1424. [doi: 10.1080/10410236.2017.1353868]
32. Tran HT, Nguyen MH, Pham TT, Kim GB, Nguyen HT, Nguyen N, et al. Predictors of eHealth literacy and its associations with preventive behaviors, fear of COVID-19, anxiety, and depression among undergraduate nursing students: a cross-sectional survey. Int J Environ Res Public Health 2022 Mar 22;19(7):3766. [doi: 10.3390/ijerph19073766] [Medline: 35409448]
33. Choi NG, Dinitto DM. The digital divide among low-income homebound older adults: internet use patterns, eHealth literacy, and attitudes toward computer/Internet use. J Med Internet Res 2013 May 02;15(5):e93 [FREE Full text] [doi: 10.2196/jmir.2645] [Medline: 23639979]
34. Omoloja A, Vundavalli S. Patient generated health data: benefits and challenges. Curr Probl Pediatr Adolesc Health Care 2021 Nov;51(11):101103. [doi: 10.1016/j.cppeds.2021.101103] [Medline: 34799255]
35. Guardado S, Mylonopoulou V, Rivera-Romero O, Patt N, Bansi J, Giunti G. An exploratory study on the utility of patient-generated health data as a tool for health care professionals in multiple sclerosis care. Methods Inf Med 2023 Dec 25;62(5-06):165-173 [FREE Full text] [doi: 10.1055/s-0043-1775718] [Medline: 37748719]
36. Oates GR, Mims C, Geurs R, Bergquist R, Hager A, Guimbellot JS, et al. Mobile health platform for self-management of pediatric cystic fibrosis: impact on patient-centered care outcomes. J Cyst Fibros 2023 Sep;22(5):823-829 [FREE Full text] [doi: 10.1016/j.jcf.2023.04.009] [Medline: 37085386]
37. Zia JK, Le T, Munson S, Heitkemper MM, Demiris G. Download alert: understanding gastroenterology patients' perspectives on health-related smartphone apps. Clin Transl Gastroenterol 2015 Jul 02;6(7):e96 [FREE Full text] [doi: 10.1038/ctg.2015.25] [Medline: 26133109]
38. Pérez-Jover V, Sala-González M, Guilabert M, Mira JJ. Mobile apps for increasing treatment adherence: systematic review. J Med Internet Res 2019 Jun 18;21(6):e12505. [doi: 10.2196/12505] [Medline: 31215517]
39. Jiménez-Chala EA, Durantez-Fernández C, Martín-Conty JL, Mohedano-Moriano A, Martín-Rodríguez F, Polonio-López B. Use of mobile applications to increase therapeutic adherence in adults: a systematic review. J Med Syst 2022 Oct 22;46(12):87. [doi: 10.1007/s10916-022-01876-2] [Medline: 36272023]
40. Con D, De Cruz P. Mobile phone apps for inflammatory bowel disease self-management: a systematic assessment of content and tools. JMIR Mhealth Uhealth 2016 Feb 01;4(1):e13 [FREE Full text] [doi: 10.2196/mhealth.4874] [Medline: 26831935]
41. Sewitch MJ, Fallone CA, Ghali P, Lee GE. What patients want in a smartphone app that supports colonoscopy preparation: qualitative study to inform a user-centered smartphone app. JMIR Mhealth Uhealth 2019 Jul 02;7(7):e12242 [FREE Full text] [doi: 10.2196/12242] [Medline: 31125310]
42. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. Evid Based Ment Health 2018 Aug 05;21(3):116-119 [FREE Full text] [doi: 10.1136/eb-2018-102891] [Medline: 29871870]
43. Giebel GD, Speckemeier C, Abels C, Plescher F, Börchers K, Wasem J, et al. Problems and barriers related to the use of digital health applications: scoping review. J Med Internet Res 2023 May 12;25:e43808. [doi: 10.2196/43808] [Medline: 37171838]
44. Santos-Vijande ML, Gómez-Rico M, Molina-Collado A, Davison RM. Building user engagement to mhealth apps from a learning perspective: relationships among functional, emotional and social drivers of user value. J Retail Consum Serv 2022 May;66:102956. [doi: 10.1016/j.jretconser.2022.102956]
45. Bundesgesetzblatt Jahrgang 2019 Teil I Nr. 49. Bundeministerium der Justiz. URL: https://tinyurl.com/3t5m72v3 [accessed 2025-03-21]

Abbreviations

eHEALS: eHealth Literacy Scale
IBD: inflammatory bowel disease
mHealth: mobile health

XSL•FO
RenderX

Edited by A Kushniruk; submitted 22.07.24; peer-reviewed by S Weidlich, Y Chu, T Martin; comments to author 04.03.25; revised version received 07.07.25; accepted 12.08.25; published 13.11.25.

Please cite as:
Kretzschmar C, Knitza J, Pietschner R, Atreya R, Neurath MF, Orlemann T
mHealth Use, Preferences, Barriers, and eHealth Literacy Among Patients With Inflammatory Bowel Disease: Survey Study
JMIR Hum Factors 2025;12:e64471
URL: https://humanfactors.jmir.org/2025/1/e64471
doi:10.2196/64471
PMID:

©Christopher Kretzschmar, Johannes Knitza, Robert Pietschner, Raja Atreya, Markus Friedrich Neurath, Till Orlemann. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 13.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Exploring Perspectives of Health Care Professionals on AI in Palliative Care: Qualitative Interview Study

Osamah Ahmad[1,2], MBChB; Stephen Mason[2], PhD; Sarah Stanley[3,4], DipHE, MRES; Amara Callistus Nwosu[2,4,5], MBChB, PhD

[1]Northern Care Alliance NHS Foundation Trust, Greater Manchester, United Kingdom

[2]Palliative Care Unit, Dept of Cardiovascular and Metabolic Medicine, University of Liverpool, Liverpool, United Kingdom

[3]Liverpool John Moores University, Liverpool, United Kingdom

[4]Marie Curie North West, Liverpool, United Kingdom

[5]Lancaster Medical School, Lancaster University, Lancaster, United Kingdom

Corresponding Author:
Amara Callistus Nwosu, MBChB, PhD
Lancaster Medical School
Lancaster University
Bailrigg
Lancaster, LA1 4YW
United Kingdom
Phone: 44 1524 594547
Email: a.nwosu@lancaster.ac.uk

Abstract

Background: The use of artificial intelligence (AI) methods in palliative care research is increasing. Most AI palliative care research involves the use of routinely collected data from electronic health records; however, there are few data on the views of palliative care health care professionals on the role of AI in practice. Determining the opinions of palliative care health care professionals on the potential uses of AI in palliative care will be useful for policymakers and practitioners to determine and inform the meaningful use of AI in palliative care practice.

Objective: This study aimed to explore the views of palliative care health care professionals on the use of AI for the analysis of patient data in palliative care.

Methods: This was a phenomenological study using qualitative semistructured interviews with palliative care health care professionals with a minimum of 1 year of clinical experience in a hospice in the North West of England. Data were analyzed using inductive thematic analysis.

Results: We interviewed 6 palliative care professionals, including physicians, nurses, and occupational therapists. AI was viewed positively, although most participants had not used it in practice. None of the participants had received training in AI and stated that education in AI would be beneficial. Participants described the potential benefits of AI in palliative care, including the identification of people requiring palliative care interventions and the evaluation of patient experiences. Participants highlighted security and ethical concerns regarding AI related to data governance, efficacy, patient confidentiality, and consent issues.

Conclusions: This study highlights the importance of staff perceptions of AI in palliative care. Our findings support the role of AI in enhancing care, addressing educational needs, and tackling trust, ethics, and governance issues. This study lays the groundwork for guidelines on AI implementation, urging further research on the methodological, ethical, and practical aspects of AI in palliative care.

(JMIR Hum Factors 2025;12:e79514) doi:10.2196/79514

KEYWORDS

artificial intelligence; AI; digital health; education; ethics; technology; end of life; innovation

XSL•FO
RenderX

Introduction

Background

Artificial intelligence (AI) is the science and engineering of creating "intelligent" machines (computers) through developed algorithms that replicate the ability of a machine to think and act like a human [1]. AI involves different methodologies (eg, machine learning, neural networks, deep learning, and natural language processing) that enable a machine to be trained to act autonomously, with or without human instruction [2]. AI has already demonstrated a significant impact in practice by facilitating the interpretation and analysis of large amounts of data contained within electronic health care datasets [3-5]. In health care, AI can facilitate the diagnosis of diseases, support clinical care delivery, and help individuals maintain their independence [6]. AI is increasingly used in palliative care (a discipline that provides holistic, person-centered support for people with life-limiting illness [7]). For example, AI-driven data analysis of electronic health records has been used to identify palliative care needs [8], support clinical documentation [9], identify quality indicators for end-of-life care [10], support symptom assessments [11], and estimate prognosis [12]. Despite the increased focus on palliative care AI, most current clinical AI tools are designed for nonpalliative care populations [8]. The lack of palliative care AI tools is a consequence of limited evidence in populations with serious illness, who often experience complex physical and psychosocial needs [13]. Palliative patients often experience significant morbidity, complex and multifaceted symptoms, and a high risk of mortality, whereas their families face grief and bereavement that extend beyond the patient's death [13]. These unique dimensions of care introduce ethical, practical, and emotional considerations for AI, which are not as prominent in other medical and surgical specialties [14]. For example, the use of AI in prognostication may be more important in palliative care, where sensitive communication about life expectancy directly shapes care planning and emotional preparedness [12,15]. Similarly, the handling of data after death (ie, digital legacy and the rights of caregivers and families) raises questions that are less relevant in acute or curative clinical contexts [16]. Despite the growing body of research on AI in medicine, there is currently limited evidence that addresses these specific complexities within palliative care, which highlights the need for targeted research in this field [17]. Consequently, understanding the views of palliative care practitioners on AI is important to better understand the opportunities, challenges, and implementation issues associated with its use [18]. Therefore, a dedicated study of palliative health care professionals' perspectives on AI is warranted to explore the specific issues affecting this cohort [17,19].

Aim

This study aimed to explore palliative care professionals' views on the use of AI for the analysis of patient data in palliative care.

Methods

Overview

This study involved inductive thematic analysis, a method of analyzing qualitative data in which themes emerge directly from the data without predefined codes or expectations. Inductive thematic analysis was chosen as it provided the researcher with the flexibility to explore participants' perspectives on AI in palliative care [20]. The lead researcher (OA) was a final-year medical student (male) who was supervised by SM and ACN. OA received training and support to conduct the interviews. The study adhered to the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist [21].

Study Setting

The study was conducted in a hospice in the North West of England, a specialized health care facility providing care for individuals in the advanced stages of a terminal illness or approaching the end of their lives. The hospice provides various services, including a 15-bed inpatient unit, day services, outpatient clinics, community outreach, patient and family support, and bereavement services.

Sampling and Recruitment

Recruitment was conducted between April and May 2022. The study was introduced at a weekly hospice education meeting, giving potential participants the opportunity to ask questions and express their interest in participating. Study advertisements were placed around the hospice, and an email was sent to all staff outlining the inclusion and exclusion criteria for the study. The study material included information about the researcher (OA). Inclusion criteria were palliative care health care professionals working in the hospice with a minimum of 1-year of clinical experience during the study data collection period of April to May 2022.

Data Collection

Over a period of 2 months, semistructured interviews were carried out with palliative care health care professionals. Participants received written information before the interview and provided written consent. Interviews were conducted face-to-face in a meeting room at the hospice. All the interviews were audio recorded using a digital voice recorder. The interviews lasted between 35 and 60 minutes, and the participants were informed that they could pause or discontinue the interview at any time. Each interview began with the researcher providing the participants with background information about the study and a definition of AI. The researcher conducted interviews using an interview guide, which was used to encourage structure across interviews but also facilitated flexibility by allowing participants to talk freely about their experiences (Multimedia Appendix 1). Open questions were used, and the interview schedule was adapted throughout the course of data collection to reflect the emergent themes and concepts (Textbox 1). Field notes and reflections were written throughout the interview process to help make sense of the data during the analysis phase.

XSL•FO
RenderX

Textbox 1. Examples of the interview questions.

- What is your experience with artificial intelligence (AI) currently used within palliative care, if any?
- What do you see as the desired goal or outcome of using AI within palliative care?
- Have you ever been educated on the use of AI within palliative care?
- What are your views on introducing education programs for palliative care health care professionals on the use of AI within palliative care?
- AI in palliative care, like any new health care technology, may raise several safety concerns. Do you have any such concerns?
- AI in palliative care, like any new health care technology, may raise several security concerns. Do you have any such concerns?
- What are some important metrics that should be used when comparing AI tools to traditional tools in palliative care?
- Can you identify a task at work which you currently perform repeatedly, with little or no variations each time?
- What level of trust would you need to allow AI tools to perform this task for you?
- What are some ways this level of trust can be established?
- Is there anything else you'd like to tell me about this topic? If not, do you have any questions?

Interviews were transcribed verbatim by OA. Data were exported to Microsoft Word, and manual thematic analysis coding was used to systematically label the qualitative data extracts to identify patterns and themes. OA used the 6-step thematic analysis proposed by Braun and Clarke [22], which involves (1) familiarization with the data, (2) generation of initial codes, (3) development of initial themes, (4) reviewing themes, (5) defining and naming themes, and (6) writing up the analysis. We used inductive analysis, with line-by-line coding of participants' interviews, to generate codes and organize the data into themes (Multimedia Appendix 2). Coding was carried out by OA. During the data collection phase, OA had regular supervisory meetings (with ACN and SM) to review data analysis, discuss initial findings, and evaluate data throughout the analysis process.

Six individuals contacted OA to consent to participate. None of the participants withdrew from the study. Interviews were analyzed iteratively. By the fourth interview, all key major themes (eg, "general openness toward AI" and "importance of human contact, empathy, and sympathy") were identified. Additional interviews reinforced, rather than expanded upon, these categories. For example, the theme of "confidentiality concerns" emerged in the first interview and was reiterated consistently in interviews 2 to 6 without new subcategories. This redundancy indicated that thematic saturation had been reached with the sample. We determined that further interviews were unlikely to provide new codes, categories, or insights relevant to the research question. Therefore, a consensus was reached to stop data collection on completion of 6 interviews (ie, to stop recruitment of further participants) as the theoretical categories became saturated [23,24].

Ethics Statement

The University of Liverpool ethics committee gave ethics approval for this work (reference number 8523). This study was approved by the hospice research governance group. All participants provided written informed consent to participate in this study. All participants were informed of their right to privacy and confidentiality. All participants were informed that their data would be anonymized and that no identifying information would be included in any works related to this study. No compensation was given to participants for participating in the study.

Results

Overview

We interviewed 6 palliative health care professionals, including 1 (17%) nurse, 2 (33%) occupational therapists, and 3 (50%) physicians. A total of 4 (67%) participants were female, and 2 (33%) were male. All interviews were conducted face-to-face and in person (Table 1).

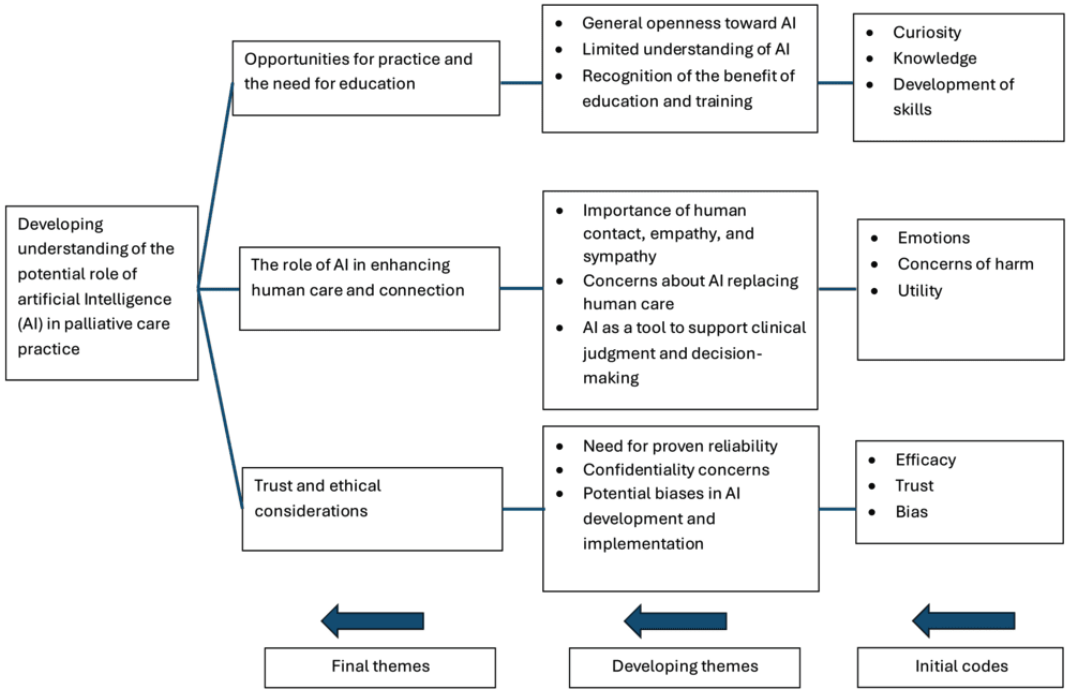
Three themes were developed from the data: (1) opportunities for practice and the need for education, (2) enhancing human care and connection, and (3) trust and ethical considerations (Figure 1).

XSL•FO
RenderX

Table 1. Characteristics of the participants (N=6).

Characteristics	Participants, n (%)
Profession	
Physician	3 (50)
Occupational therapist	2 (33)
Nurse	1 (17)
Sex	
Female	4 (67)
Male	2 (33)
Interview method	
In person	6 (100)
Online	0 (0)

Figure 1. Thematic map showing the 3 main themes. AI: artificial intelligence.



Opportunities for Practice and the Need for Education

Participants highlighted their openness to consider using AI in palliative care practice. Specifically, participants described their views on how this technology could improve care delivery and efficiency for their patients. Participants spoke of how algorithm-driven health care is increasingly being used in other clinical specialties and how, similarly, AI can be used to improve palliative care:

> *I mean, I think obviously the future is, is going that way, isn't it? You know, artificial intelligence is being developed at all areas.* [Participant 3]

Participants described the lack of AI use in palliative care compared to other medical and surgical specialties:

> *I suppose it's kind of an alien concept, isn't it? I think in terms of technology that we use day-to-day at the moment, we're kind of behind the curve, I would say in, in the NHS [National Health Service], and certainly in this environment. So, I think it's difficult to kind of visualise what that would be like in kind of day-to-day practice.* [Participant 6]

Participants spoke about how AI can potentially improve palliative care. The interviewees provided several examples of possible benefits, such as machine learning–based analyses of electronic health records to identify people needing palliative care, to improve data capture and analysis, and to inform personalized care recommendations:

> *I guess, you know, looking through the hospital records in the hospital. When I see a new patient, I'm manually looking back through the notes to try to find if they've had previous encounters with the palliative care team, how many times they've been in hospital, looking for all the medications to work out what*

XSL•FO
RenderX

> *medications they're on and if any have been stopped.* [Participant 3]
>
> *A good thing would be if something could analyse the entire database every day and say, "look, these are the patients who are flagging up, particularly symptomatic," or "are using lots of PRN [as required] medicines" or, you know, [if] certain keywords trigger [an] urgent review?* [Participant 4]

Participants highlighted that palliative care staff have a limited understanding of the different types of AI applications that may be used in clinical care. Participants expanded on this by describing how better staff education may improve their confidence in using AI in clinical practice:

> *I think for healthcare professionals, like for me, who've maybe not come across so much, or don't understand it so much, it's more about knowing about it, and understanding how we can help use it a bit better.* [Participant 1]
>
> *I think that's only a positive thing. And, I think, to introduce education to help people understand even the basics of what AI is is really important, because I think until we know more about what it is and understand more about what it is, we can't think of the ways in which we can improve care for the patient.* [Participant 1]

Enhancing Human Care and Connection

Participants emphasized the importance of human contact, empathy, and sympathy in palliative care, expressing concerns about AI potentially replacing human care. This was illustrated by the comments of 1 participant:

> *I think in palliative care, you know, our strength is based on human connections.... I don't mean to be like negative towards, kind of, AI technology and stuff, but I think, actually, you know, the majority of my day, when I'm speaking to patients, it's about human connection.* [Participant 6]

Participants identified how AI systems could be potentially designed and modeled to dynamically learn from health care professionals and service users, with the objective to develop algorithms that are meaningful in helping them provide clinical care:

> *I think it should be modelled as if, you know, the algorithm behind it is trying to learn, and it's asking you for confirmation. You know, so it's saying, "this is what I've found," you know, and you then have to decide whether this is useful, or "can you confirm this is correct?", or something like that. It should be this kind of back and forth.* [Participant 4]

Participants described the potential that palliative care health care professionals can routinely use data, informed by AI, to support care and their clinical decision-making (eg, analysis of electronic health record data to identify patients with palliative care needs, automated transcription of consultations, and personalized treatment recommendations). Consequently, the interviewees discussed scenarios in which AI was used to improve and augment the clinical care provided by palliative care professionals. Participants said that AI could be used as a tool to support clinical judgment and decision-making while maintaining a patient-centered approach:

> *I guess if you use the program alongside, like, have a step-by-step approach, so you could use the program or the intelligence alongside, you know, a human in kind of partnership. And then gradually, you just have to learn that trust and become more used to it.* [Participant 2]

Trust and Ethical Considerations

Participants highlighted the importance of health care professionals being confident that AI-driven clinical tools are trustworthy and reliable. The interviewees framed their opinions on the trustworthiness of AI in the context of their clinical responsibility of providing care for people with serious illness (who are often considered vulnerable). Several quotes from the participants illustrated the important role of research in generating evidence to inform meaningful AI use in clinical care:

> *It'd be track record and experience, wouldn't it? So, you probably would want to have enough evidence to show that it made good decisions.* [Participant 5]

Participants stated that confidentiality was important, with several statements highlighting their concerns regarding data privacy and the security implications of using AI in clinical practice. They emphasized the importance of ensuring that patients' data remain secure and protected when using AI to inform clinical decision-making in palliative care:

> *I suppose, privacy and data is the main concern as quite often it is in health care, isn't it? And you know, making sure that data is secure and, you know, things like hacking aren't an issue and people can't access patients' and relatives' private data easily.* [Participant 1]
>
> *I think with any new technology to have to make sure it's secure... And, you know, privacy concerns and breaches of confidentiality- So I think all those things are really important.* [Participant 6]

Participants discussed the risk of bias associated with AI analysis, which could create (and widen existing) inequalities in palliative care. In the interviews, the participants discussed the importance of developing strategies to reduce this risk of bias in AI algorithms and to explore public opinion about the role of AI in clinical practice. Participants described their belief that if AI should be used to improve holistic care for patients and those important to them. Further comments from the interviews highlighted the opinion that the use of AI in palliative care should be inclusive, unbiased, and ethical:

> *If our algorithms have been driven by developers in Silicon Valley in California and most of them are white, male and young, and the modelling has been tested on a particular set of individuals, which don't have certain characteristics, which mean that certain people are not represented, you then might get a*

XSL•FO
RenderX

device or an algorithm which isn't tailored for the needs of certain people. [Participant 4]

General Comments of Participants

Overall, AI was viewed positively by many participants, although many stated that they had not used it in practice. None of the participants had received training in AI, and all stated that they would have liked to have received education on this topic. Participants described their opinions on how data science can improve clinical care; potential ideas included the use of AI to identify people from electronic health records who need palliative care and to use data analytics to evaluate quality of care. Participants highlighted data privacy and ethical issues related to AI use in palliative care, including important related issues such as governance, confidentiality, and consent.

Discussion

Concerning the use of AI in palliative care, it is important to consider the opportunities it offers, the educational needs of staff, the importance of human connections, and practical issues related to trust and ethics.

Importance and Uniqueness of This Paper

This study provides an overview of the views of specialist palliative care health care professionals regarding AI in palliative care, which adds knowledge to the limited evidence base. Our qualitative approach facilitated an in-depth exploration of participants' experiences, which captured their nuances and complexities. The evidence derived from this study improves knowledge about the views palliative care staff have about AI, which will shape its clinical use. Evidence demonstrates that improved staff involvement can improve the effectiveness and success of health care system interventions [25,26].

Relation to Previous Work in This Area

In this study, the staff positively described potential opportunities in which AI could be used to support palliative care, with themes consistent with previous work conducted with generalist staff [27]. In our study, staff identified several hypothetical possibilities where AI could be positively used to improve their practice (eg, predictive modeling, text screening, symptom assessment, and communication), which are consistent with current developments of AI in palliative care [28].

Our findings align with previous research, which recommends that health care professionals receive formal education and training in AI [17,29]. Specifically, previous research advises palliative care education programs to include holistic overviews of AI technologies, ethical considerations, and case studies that highlight the real-world applications and challenges of AI in palliative care [30].

Our findings support previous work, which describes the importance of focusing on human connections in palliative care while ensuring that AI tools are meaningfully used to improve the experience of patients, caregivers, and staff [27]. Consistent with previous research, we highlight the potential problems and bias that may occur from using AI in palliative care, due to limited evidence of patient-centered outcomes measures in palliative care populations [14]. In our analysis, participants described the ethical challenges of using AI in palliative care (reporting themes of promoting transparency and accountability in AI systems, regular ethical review and continuous impact assessments, ensuring patient autonomy and informed consent, and safeguarding data privacy and security) [30]. The ethical themes reported in our study are similar to those in previous research [30,31], illustrating the need to incorporate ethical principles (such as the 4 principles) into decisions involving AI use in palliative care practice [32]. For example, autonomy (Do service users have a choice of how their data are collected, analyzed, and stored?), beneficence (Is AI used in the best interests of individuals or is the benefit mostly for groups, populations, or other stakeholders?), nonmaleficence (How can we ensure people are not harmed from AI in palliative care?), and justice (How can AI tools be used fairly and equitably, in vulnerable people, from different backgrounds, cultures, and geographies?) [33]. Furthermore, our results support the importance of integrating broader ethical theories into practice (eg, consequentialism, deontology, rights-based ethics, and virtue ethics) to provide clinicians and policymakers with a framework to make decisions on how to use emerging technologies in clinical practice [34]. These frameworks can be used to address uncertainty to help stakeholders' (eg, health care professionals, managers, and policymakers) decision-making when considering how to responsibly use AI tools in palliative care [13]. Consistent with previous work, our findings reinforce the view that there is a risk that current AI applications lack engagement with the ethical complexities of real-world use in palliative care, which highlights questions about the adequacy of clinical practice safeguards [15,30].

Limitations

This study is small, focused on one hospice in the North West of England, which means that the findings may not be generalizable to other palliative care settings (eg, home, community, hospital, and nursing homes). Although thematic saturation was achieved with 6 participants, this may have been influenced by the similarities of the participants. We acknowledge that a larger sample incorporating staff from a wider range of roles and professional backgrounds may have yielded more in-depth and diverse data. For example, this study lacks representation of some professional roles (eg, social work, spiritual care, pharmacy, and fundraising), which means there is a lack of data about how AI may impact wider specialist roles in the palliative care multidisciplinary team. Furthermore, our study did not include the perspectives of patients, caregivers, and other relevant stakeholders.

Importance to Policy, Practice, and Research

Decision-makers should consider the perspectives of palliative care staff when developing and implementing AI tools for palliative care. When considering applications of palliative care AI, decision-makers should consider how these innovations will improve care, address human needs, and fulfill ethical and governance requirements. From an educational perspective, it is important that palliative care professionals are trained to safely and effectively use new technologies (eg, AI) in clinical practice. Activities to achieve this objective may include the development of training curricula for undergraduate and

XSL•FO
RenderX

postgraduate students, including content on the opportunities and challenges of AI in health care, ethical considerations of its use, and governance issues [30,35].

Future research on palliative care AI should establish standardized reporting for studies, seek external validation, and consider ethical issues, which are needed to ensure that the clinical application of AI tools is safe, meaningful, and effective [15]. Future research should explore views (on palliative care AI) from different perspectives, including multidisciplinary teams, managers, patients, caregivers, and other relevant stakeholders. Researchers should involve interdisciplinary partnerships and collaboration to facilitate work across essential interconnected themes, such as design, computing, data analysis, ethics, and translational medicine [17,36,37].

Conclusions

This study shows the importance of considering the views of palliative care professionals regarding the potential role of AI in clinical practice. Our findings demonstrate the importance of considering opportunities to meaningfully use AI to improve human-focused care, support staff education, and address practical issues related to trust, ethics, and governance. This study provides a foundation for developing guidelines for AI implementation in palliative care practice. Future research should examine methodological, ethical, and practical issues to ensure that AI best supports palliative care for people with serious illnesses.

Acknowledgments

The authors would like to thank the participants for their valuable input and participation in this study. They would also like to thank the hospice management and research governance group for supporting study recruitment. SS and ACN were funded by Marie Curie, the UK's leading palliative care charity. This study did not receive any funding.

Data Availability

All data generated or analyzed during this study are included in this published article (and its supplementary information files).

Authors' Contributions

Conceptualization: OA (lead), SM (supporting), ACN (supporting)
Formal analysis: OA (lead), SM (supporting), ACN (supporting) SS (supporting)
Methodology: OA (lead) ACN (supporting), SS (supporting), SM (supporting)
Project administration: OA (lead), ACN (equal), SM (supporting), SS (supporting)
Supervision: ACN (lead), SM (equal), SS (equal)
Writing—original draft: OA (lead), ACN (supporting)
Writing—review and editing: OA (lead), ACN (equal), SM (supporting), SS (supporting)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.
[DOCX File , 40 KB - humanfactors_v12i1e79514_app1.docx]

Multimedia Appendix 2
Generation of study themes.
[DOCX File , 28 KB - humanfactors_v12i1e79514_app2.docx]

References

1. McCarthy J. What is artificial intelligence? Stanford University. 2007 Nov 12. URL: https://www-formal.stanford.edu/jmc/whatisai.pdf [accessed 2025-11-22]
2. Chen M, Decary M. Artificial intelligence in healthcare: an essential guide for health leaders. Healthc Manage Forum 2020 Jan 24;33(1):10-18. [doi: 10.1177/0840470419873123] [Medline: 31550922]
3. Shubhendu S S, Vijay J. Applicability of artificial intelligence in different fields of life. Int J Sci Eng Res 2013 Sep;1(1):28-35 [FREE Full text] [doi: 10.70729/1130915]
4. Quinn TP, Senadeera M, Jacobs S, Coghlan S, Le V. Trust and medical AI: the challenges we face and the expertise needed to overcome them. J Am Med Inform Assoc 2021 Mar 18;28(4):890-894 [FREE Full text] [doi: 10.1093/jamia/ocaa268] [Medline: 33340404]
5. Bajwa J, Munir U, Nori A, Williams B. Artificial intelligence in healthcare: transforming the practice of medicine. Future Healthc J 2021 Jul;8(2):e188-e194 [FREE Full text] [doi: 10.7861/fhj.2021-0095] [Medline: 34286183]

XSL•FO
RenderX

6. Sixsmith A. AgeTech: technology-based solutions for aging societies. In: Rootman I, Edwards P, Levasseur M, Grunberg F, editors. Promoting the health of older adults: the Canadian experience. Toronto, ON: Canadian Scholars; 2021.
7. Palliative care. World Health Organization. URL: https://www.who.int/health-topics/palliative-care [accessed 2025-11-22]
8. Durieux BN, Tarbi EC, Lindvall C. Opportunities for computational tools in palliative care: supporting patient needs and lowering burden. Palliat Med 2022 Sep;36(8):1168-1170. [doi: 10.1177/02692163221122261] [Medline: 36062726]
9. Lindvall C, Lilley EJ, Zupanc SN, Chien I, Udelsman BV, Walling A, et al. Natural language processing to assess end-of-life quality indicators in cancer patients receiving palliative surgery. J Palliat Med 2019 Feb;22(2):183-187. [doi: 10.1089/jpm.2018.0326] [Medline: 30328764]
10. Lee RY, Brumback LC, Lober WB, Sibley J, Nielsen EL, Treece PD, et al. Identifying goals of care conversations in the electronic health record using natural language processing and machine learning. J Pain Symptom Manage 2021 Jan;61(1):136-42.e2 [FREE Full text] [doi: 10.1016/j.jpainsymman.2020.08.024] [Medline: 32858164]
11. Heintzelman NH, Taylor RJ, Simonsen L, Lustig R, Anderko D, Haythornthwaite JA, et al. Longitudinal analysis of pain in patients with metastatic prostate cancer using natural language processing of medical record text. J Am Med Inform Assoc 2013;20(5):898-905 [FREE Full text] [doi: 10.1136/amiajnl-2012-001076] [Medline: 23144336]
12. Avati A, Jung K, Harman S, Downing L, Ng A, Shah NH. Improving palliative care with deep learning. BMC Med Inform Decis Mak 2018 Dec 12;18(Suppl 4):122. [doi: 10.1186/s12911-018-0677-8] [Medline: 30537977]
13. Finucane AM, Swenson C, MacArtney JI, Perry R, Lamberton H, Hetherington L, et al. What makes palliative care needs "complex"? A multisite sequential explanatory mixed methods study of patients referred for specialist palliative care. BMC Palliat Care 2021 Jan 15;20(1):18 [FREE Full text] [doi: 10.1186/s12904-020-00700-3] [Medline: 33451311]
14. García Abejas A, Geraldes Santos D, Leite Costa F, Cordero Botejara A, Mota-Filipe H, Salvador Vergés À. Ethical challenges and opportunities of AI in end-of-life palliative care: integrative review. Interact J Med Res 2025 May 14;14:e73517 [FREE Full text] [doi: 10.2196/73517] [Medline: 40302210]
15. Migiddorj B, Batterham M, Win KT. Systematic literature review on the application of explainable artificial intelligence in palliative care studies. Int J Med Inform 2025 Aug;200:105914 [FREE Full text] [doi: 10.1016/j.ijmedinf.2025.105914] [Medline: 40250167]
16. Stanley S, Higginbotham K, Finucane A, Nwosu AC. A grounded theory study exploring palliative care healthcare professionals' experiences of managing digital legacy as part of advance care planning for people receiving palliative care. Palliat Med 2023 Oct;37(9):1424-1433 [FREE Full text] [doi: 10.1177/02692163231194198] [Medline: 37609897]
17. Nwosu AC, McGlinchey T, Sanders J, Stanley S, Palfrey J, Lubbers P, et al. Identification of digital health priorities for palliative care research: modified Delphi study. JMIR Aging 2022 Mar 21;5(1):e32075 [FREE Full text] [doi: 10.2196/32075] [Medline: 35311674]
18. Parchmann N, Hansen D, Orzechowski M, Steger F. An ethical assessment of professional opinions on concerns, chances, and limitations of the implementation of an artificial intelligence-based technology into the geriatric patient treatment and continuity of care. Geroscience 2024 Dec 04;46(6):6269-6282. [doi: 10.1007/s11357-024-01229-6] [Medline: 38834930]
19. Mills J, Fox J, Damarell R, Tieman J, Yates P. Palliative care providers' use of digital health and perspectives on technological innovation: a national study. BMC Palliat Care 2021 Aug 07;20(1):124 [FREE Full text] [doi: 10.1186/s12904-021-00822-2] [Medline: 34364379]
20. Fereday J, Muir-Cochrane E. Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. Int J Qual Methods 2006;5(1):80-92. [doi: 10.1177/160940690600500107]
21. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec 16;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
22. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2008 Jul 21;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
23. Charmaz K. Constructing Grounded Theory. Thousand Oaks, CA: SAGE Publications; 2014.
24. Pope C, Ziebland S, Mays N. Analysing qualitative data. In: Pope C, Mays N, editors. Qualitative Research in Health Care, Third Edition. Oxford, UK: Blackwell Publishing Ltd; 2006.
25. Cadeddu SB, Dare LO, Denis JL. Employee-driven innovation in health organizations: insights from a scoping review. Int J Health Policy Manag 2023;12:6734 [FREE Full text] [doi: 10.34172/ijhpm.2023.6734] [Medline: 37579391]
26. Ghani B, Hyder SI, Yoo S, Han H. Does employee engagement promote innovation? The facilitators of innovative workplace behavior via mediation and moderation. Heliyon 2023 Nov;9(11):e21817 [FREE Full text] [doi: 10.1016/j.heliyon.2023.e21817] [Medline: 38034753]
27. Henzler D, Schmidt S, Koçar A, Herdegen S, Lindinger GL, Maris MT, et al. Healthcare professionals' perspectives on artificial intelligence in patient care: a systematic review of hindering and facilitating factors on different levels. BMC Health Serv Res 2025 May 01;25(1):633 [FREE Full text] [doi: 10.1186/s12913-025-12664-2] [Medline: 40312413]
28. Reddy V, Nafees A, Raman S. Recent advances in artificial intelligence applications for supportive and palliative care in cancer patients. Curr Opin Support Palliat Care 2023 Jun 01;17(2):125-134. [doi: 10.1097/SPC.0000000000000645] [Medline: 37039590]

XSL•FO
RenderX

29. Alowais SA, Alghamdi SS, Alsuhebany N, Alqahtani T, Alshaya AI, Almohareb SN, et al. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. BMC Med Educ 2023 Sep 22;23(1):689 [FREE Full text] [doi: 10.1186/s12909-023-04698-z] [Medline: 37740191]
30. Adegbesan A, Akingbola A, Ojo O, Jessica OU, Alao UH, Shagaya U, et al. Ethical challenges in the integration of artificial intelligence in palliative care. J Med Surg Public Health 2024 Dec;4:100158. [doi: 10.1016/j.glmedi.2024.100158]
31. Floridi L. The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities. Oxford, UK: Oxford University Press; 2023.
32. Beauchamp TL, Childress JF. Principles of Biomedical Ethics. Oxford, UK: Oxford University Press; 2001.
33. Varkey B. Principles of clinical ethics and their application to practice. Med Princ Pract 2021 Jun 4;30(1):17-28 [FREE Full text] [doi: 10.1159/000509119] [Medline: 32498071]
34. Chonko L. Ethical theories. Direct Selling Education Foundation. 2012. URL: https://dsef.org/wp-content/uploads/2012/07/EthicalTheories.pdf [accessed 2025-11-22]
35. Preparing the healthcare workforce to deliver the digital future. NHS Health Education England. 2018. URL: https://www.hee.nhs.uk/sites/default/files/documents/Topol%20Review%20interim%20report_0.pdf [accessed 2025-11-22]
36. Nwosu AC. Telehealth requires improved evidence to achieve its full potential in palliative care. Palliat Med 2023 Jul;37(7):896-897 [FREE Full text] [doi: 10.1177/02692163231182461] [Medline: 37387223]
37. Nwosu A, Stanley S, Norris J, Taubert M. Digital legacy and palliative care: using technology, design and healthcare partnerships to research how digital information is managed after death. BMJ Support Palliat Care 2023;13:A41-A42 [FREE Full text] [doi: 10.1136/spcare-2023-hunc.107]

Abbreviations

AI: artificial intelligence
COREQ: Consolidated Criteria for Reporting Qualitative Research

Edited by A Choudhury; submitted 23.06.25; peer-reviewed by M Tierney, SK Nandipati, J John Thayil; comments to author 07.09.25; revised version received 26.09.25; accepted 19.11.25; published 08.12.25.

Please cite as:
Ahmad O, Mason S, Stanley S, Nwosu AC
Exploring Perspectives of Health Care Professionals on AI in Palliative Care: Qualitative Interview Study
JMIR Hum Factors 2025;12:e79514
URL: https://humanfactors.jmir.org/2025/1/e79514
doi:10.2196/79514
PMID:

©Osamah Ahmad, Stephen Mason, Sarah Stanley, Amara Callistus Nwosu. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 08.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Organizational Readiness for Change in the Era of Smart Hospital Wards: Qualitative Study of Health Care Workers' Insights

Hui Wen Lim[1], MSc; Jennifer Sumner[1,2], MSc, PhD; Abigail Ang[2], BA; Camille Keck[2], BSc; Emily Hwee Hoon Chew[2], MPH; Alexander Wenjun Yip[2], MBBCh

[1]Medical Affairs – Research Innovation & Enterprise, Alexandra Hospital, Singapore, Singapore, Singapore

[2]Department of Healthcare Redesign, Alexandra Research Centre for Healthcare in a Virtual Environment (ARCHIVE), Alexandra Hospital, Singapore, Singapore

Corresponding Author:
Jennifer Sumner, MSc, PhD
Department of Healthcare Redesign
Alexandra Research Centre for Healthcare in a Virtual Environment (ARCHIVE)
Alexandra Hospital
378 Alexandra Road
Singapore, 159964
Singapore
Phone: 65 69082222
Fax: 65 69082222
Email: jennyssumner@gmail.com

Abstract

Background: Technology is rapidly reshaping conventional hospital environments into smart spaces, enhancing care, improving clinical workflows, and reducing workloads. However, successful implementation depends not only on the effectiveness of the technology but also on organizational readiness for change.

Objective: This study aimed to identify the key enablers and barriers to readiness for change for a smart hospital ward initiative.

Methods: We conducted a qualitative study to gauge organizational readiness for change for a smart ward initiative. Using purposive sampling, we captured diverse views from clinicians, IT staff, operational support staff, and health care redesign staff. Data were coded deductively under 3 key domains in Weiner's theory of organizational readiness: change efficacy, change commitment, and contextual factors. Subthemes were derived inductively under each domain.

Results: We interviewed 19 participants, including clinicians and support staff. Six subthemes emerged: (1) perceived valence and feasibility; (2) transparency and trust in management; (3) shared understanding and readiness to act; (4) resources, training, and staff capability; (5) innovation culture; and (6) past experiences. Participants viewed the initiative as valuable and were motivated to change, citing that the institution's innovation culture was a key enabler. However, there were key barriers, including unclear timelines, inconsistent training, limited resources, and a lack of infrastructure to support innovation. Concerns about overreliance on technology were also prominent, with staff wary of its impact on clinical judgment and system reliability.

Conclusions: Enabling readiness for the smart ward initiative requires transparent communication of timelines and project awareness, particularly for ground staff, the development of training frameworks, and adequate prioritization of innovation. Alleviating commonly reported technology concerns, such as overreliance, loss of human touch, and system reliability, will also be key to adoption and sustainability.

(JMIR Hum Factors 2025;12:e81932) doi:10.2196/81932

KEYWORDS

organizational readiness for change; health care innovation; technology adoption; change management; qualitative research

Introduction

Technology is rapidly transforming inpatient care, reshaping conventional hospital environments into smart spaces, including wards [1,2]. "Smart wards," an emerging feature of digitally enabled hospitals, are ecosystems that use interconnected digital systems to enhance patient care, improve clinical workflows, and reduce workload through automation [3,4]. These

XSL•FO
RenderX

ecosystems may incorporate a range of technologies, including real-time patient monitoring, movement detection, and artificial intelligence (AI)–based support tools [4]. Motivated by the challenges of caring for an aging population, a constrained health care workforce, and limited hospital bed capacity [5-7], there is growing investment in smart ward environments. However, the implementation of such technologies requires more than hardware or software; it demands meaningful behavioral and workflow changes among staff.

Alexandra Hospital is the latest hospital to invest in a smart ward concept as part of its commitment to adopting transformative health care technology. The smart ward initiative will include technologies such as continuous vital sign tracking with an integrated dashboard for real-time patient monitoring, a unified communication system for improved team collaboration, inpatient teleconferencing systems for remote consultations and patient education, chatbot agents for patient education, and the integration of AI analytics into clinical systems. The initiative also envisions a centralized virtual workforce that remotely supports inpatient care and ground staff.

To achieve the smart ward vision, substantial changes in staff behaviors and workflows are required**.** Thus, the successful implementation of technology depends not only on the effectiveness of the technology itself but also on how well an organization and its staff can adapt to these new systems—a concept known as "organisational readiness for change" [8,9]. In the context of digital transformation, such as smart wards, organizational readiness reflects the extent to which staff and systems are collectively prepared—psychologically and behaviorally—to adopt and sustain technological innovations. According to Weiner [9], two key constructs underlie readiness: (1) change efficacy (the shared belief in the organization's capability to execute the change successfully) and (2) change commitment (the collective resolve to implement the change). Multiple factors may influence these constructs, including leadership support, local policy, clear timeframes, resource availability, staff training, and a culture for safe experimentation and learning (ie, sandbox environments) [9-11]. Together, they determine the degree of preparedness in an organization, both psychologically and behaviorally, to implement change [9]. When readiness is high, the implementation of a proposed change is effective, while low organizational readiness leads to resistance to and avoidance of change [9,12-14].

Although several implementation and organizational change frameworks exist; such as the Consolidated Framework for Implementation Research, and Kotter's change model, Weiner's organizational readiness for change model was chosen for its specific focus on the psychological and collective capability to implement change Unlike frameworks that emphasize structural, procedural, or leadership-driven dimensions, this model centers on the shared commitment and confidence of staff to enact change. Through examining how staff perceive readiness for the smart ward and the organizational conditions that support it, the study refines the understanding of how change commitment and change efficacy manifest in high-technology, team-based clinical contexts. Although there is an increasing global interest in smart hospital initiatives, most existing research focuses on the technological infrastructure, digital capabilities, or outcomes of smart technologies [15]. Far less is known about the behavioral and contextual factors that shape successful adoption in a smart ward context. Addressing this gap, this study examines the enablers and barriers to organizational readiness for change as Alexandra Hospital embarks on its smart ward initiative, using the framework proposed by Weiner as the analytical lens.

Methods

We conducted a qualitative study to explore health care staff members' readiness for the smart ward initiative. The project was run during an initial pilot testing phase, in which some, but not all, technologies had been deployed. This study is reported according to the Consolidated Criteria for Reporting Qualitative Research (COREQ) [16].

Ethical Considerations

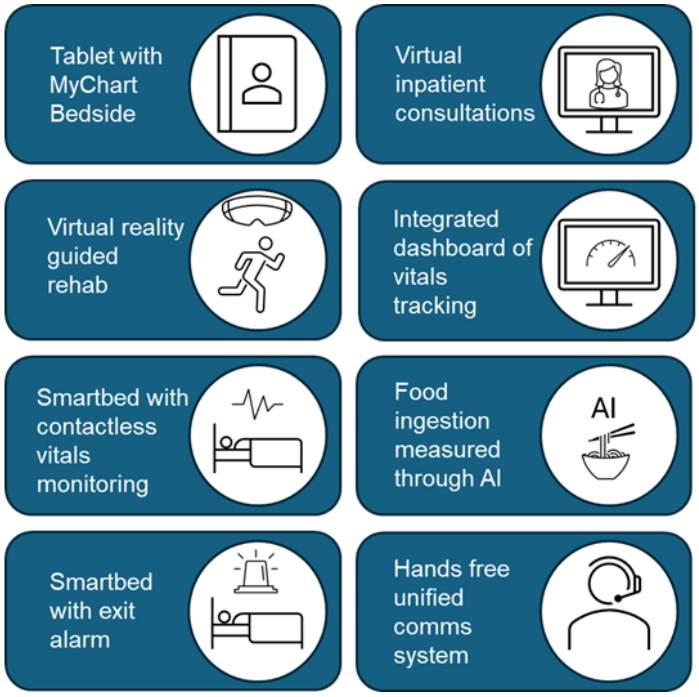
The study was reviewed and approved by the National University of Singapore Institutional Review Board for Social, Behavioral, and Educational Research studies (NUS-IRB-2023-327). All participants provided written consent to take part in the study. It was emphasized that the study was voluntary and participants could withdraw at any time. Participants received no compensation for taking part in the study. Data were deidentified and participants were assigned a unique ID number.

Setting

Alexandra Hospital, part of the National University Health System, provides care across the continuum, from acute to subacute rehabilitative services. The smart ward initiative, based at Alexandra Hospital, incorporates a suite of technologies (Figure 1). These technologies enhance access to personal health information for patients (via the MyChart Bedside) and clinical data for providers through smart bed systems for continuous vital sign monitoring, automatic bed exit alarms, computer vision, and AI-calculated food consumption, and an integrated dashboard for data analysis and visualization. The initiative also introduces new modes of care, such as virtual reality-guided rehabilitation and virtual inpatient consultations, alongside improved communication through a unified communications system. At the time of this study, some but not all of the technologies had been deployed in the ward.

XSL•FO
RenderX

Figure 1. Summary of the planned smart ward technologies. AI: artificial intelligence.



Participants

We used a purposive sampling approach to capture the views of different departments involved in smart ward development or staff working within the inpatient setting, where the initiative is being implemented. Our inclusion criteria were as follows: staff employed at Alexandra Hospital during the study period; individuals involved in smart ward development or clinicians working within the smart ward; representatives from nursing, medical, IT, operations, or health care redesign departments; the ability to provide informed consent; and the ability to participate in an English-language interview. Exclusion criteria were staff not involved in or exposed to the smart ward initiative (ie, external vendors, temporary or locum staff) and the inability to provide informed consent. The choice of representatives was made as these departments were collectively involved in the design or implementation planning for the smart ward, or represented the main clinical groups affected by the smart ward transformation.

Participants were contacted by email or phone messaging. A research team member (with no direct relationship with the participant) followed up with the potential participant to discuss the project, obtain written informed consent, and arrange a convenient time for the interview. Recruitment continued until data saturation had been reached, defined as the point at which no new codes, themes, or perspectives emerged from successive interviews, and the research team agreed that additional data would not add further conceptual insights. Saturation was assessed iteratively during data collection among the study team. No participants declined to take part, and no participants received compensation to take part in this study.

Materials

The interview guide was informed by the organizational theory of Weiner's implementation effectiveness framework by [9], which is widely used to assess organizational readiness for change in health care settings [17]. This framework was selected because it provides a structured lens to examine the psychological (eg, change commitment) and structural (eg, change efficacy) factors that influence an organization's preparedness to implement change. This makes it particularly suitable for examining health care workers' perceptions before implementing a complex intervention, such as a smart ward. The interview topics included participants' general views, awareness, and understanding of the smart ward initiative; perceived benefits and challenges; and preparedness, including available resources, managerial support, required skills, and value alignment. JS drafted an initial guide and subsequently shared it with the broader project team for feedback. The interview guide was also developed iteratively as new topics emerged during the interviews to ensure responsiveness to participants' experiences. For instance, although we originally asked about the impact of the smart ward on staff-to-patient interactions, many participants also reflected on staff-to-staff interactions. Some also raised questions regarding the suitability of certain patients for the smart ward environment. These topics were subsequently incorporated into later interviews to systematically explore this emerging dimension. Despite these additions, all interviews continued to cover the same foundational domains related to organizational readiness, ensuring comparability across participants.

Procedure

Interviews were conducted in pairs by 4 female health services researchers trained in qualitative methods (JS, HWL, CK, and EHHC) from August 2023 to February 2024. Interviews were conducted using a semistructured interview guide, were audio recorded, and transcribed verbatim. Interviews were conducted in a quiet place that was convenient and acceptable to the participant. On average, the interviews lasted 47 minutes (range 29-69 min).

XSL•FO
RenderX

Data Analysis

We used a framework analysis approach, applying a hybrid deductive-inductive approach [18]. A deductive structure was based on the theory of the organizational readiness for change framework proposed by Weiner, which guided the initial coding. Inductive coding captured emergent insights in each domain. To contextualize the framework to our study, the project team reviewed the original definitions of the framework domains: change efficacy, change commitment, and contextual factors. The team then contextualized these definitions to our study through discussion, drafting a preliminary codebook. Three researchers (JS, HWL, and AA) then independently coded a single transcript sentence by sentence using the draft codebook and met to discuss coding assignments and clarify definitions. This exercise was repeated on a second transcript to resolve any remaining questions. All subsequent transcripts were coded by a single coder (HWL or AA) and checked by a second coder (JS) to ensure consistency. Differences in opinions were mutually reconciled through discussions among the team. Codes were subsequently grouped into subthemes within the framework domains through team discussions, allowing new concepts to emerge while retaining theoretical alignment. All coding and analyses were conducted in Microsoft Word.

Results

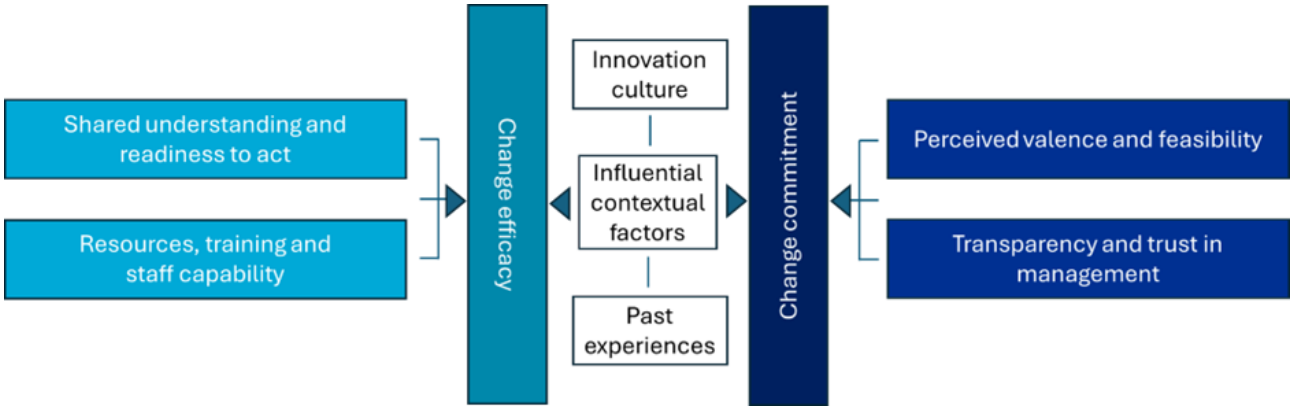
We interviewed 19 participants from different professional backgrounds, including clinicians and support staff (Table 1).

Table 1. Participant characteristics (N=19).

Job function	Participants, n (%)
Operations	1 (5)
IT	3 (16)
Nursing	8 (42)
Physician	5 (26)
Health care redesign manager	2 (11)

Results are organized according to the 3 key domains of the theory of organizational readiness for change proposed by Weiner (change efficacy, change commitment, and contextual factors), and 6 subthemes under these domains (Figure 2). We identified several enablers and barriers impacting the smart ward initiative (Table 2). A broader description of the influencing factors is provided in the following section. Additional supporting quotes are included in Multimedia Appendix 1.

Figure 2. Mapping of inductively derived subthemes to the 3 key domains in the organizational readiness for change framework proposed by Weiner [9].



XSL•FO
RenderX

Table 2. Summary of key enablers and barriers by organizational readiness domains (change efficacy, change commitment, and contextual factors) and the 6 subthemes.

	Enabler	Barrier
Change commitment		
Perceived valence and feasibility	• Value perceived • Alignment on value across departments • Perceived as appropriate for general medicine wards	• Slow roll-out • Limited integration • Failure to meet expected needs • Overreliance on technology • Reliability questioned • Alarm fatigue or desensitization
Transparency and trust in management	• Trust in management • Active stakeholder engagement • Minimal expectations for high involvement	• Minimal ground-nurse engagement • Fear of failure or perception that innovation is not a priority • Cultural norms impacting openness
Change efficacy		
Shared understanding and readiness to act	• Good awareness of the initiative • Facilitation of interactions between teams	• Inconsistent communication • Misalignments on purpose
Resources, training, and staff capability	• Ground-up training initiatives • Troubleshooting support	• Delayed training or unclear responsibility for training • Insufficient staff or time to engage with new technology • Need for external funding
Contextual factors		
Innovation culture	• Sandbox culture • Keeping up to date with technology developments • Raising staff awareness of technology	• Complexity of the “sandbox” approach in the clinical setting
Past experiences	• Prior positive experience with technology	• Prior negative experience with technology

Change Commitment: Perceived Valence and Feasibility

Participants broadly agreed that the smart ward initiative held value and had potential to improve efficiency, reduce overtime, and allow more dedicated time for direct patient care. Their comments reflected enthusiasm and a shared sense that the change was desirable:

> *It’s to help us, to make our job easy, to elevate, to help to reduce the workloads.* [Nurse ID 04]

> *Yeah, for sure. I will love to work in an environment like that. Everything is so streamlined. Yeah, that would be good.* [Physician ID 12]

However, during the pilot phase, expectations were often unmet. Many expressed frustrations with implementation delays, lack of system integration, and a failure to reduce workloads. Others shared that pilot testing had resulted in “double work,” with old and new processes carried out simultaneously because the new process was not trusted. Health care redesign staff echoed these concerns, noting early resistance and uncertainty around pilot technology:

> *I think so long already until [smart ward], I think they need to expedite.* [Nurse ID 05]

> *Because we don’t totally trust it right? Because it’s totally new and it hasn’t been totally tested out. So you want to do things the old way as well.* [Physician ID 12]

> *When we tried to implement anything new, there’s always resistance. Because I guess it’s, at first, it’s just umm, they didn’t make sense out of it. It’s like. At first, it’s not going to be everything smooth. It’s going to be a lot more work at the beginning for them. So it’s normal that they may feel hesitant about it.* [HCRD ID 18]

Concerns about overreliance on technology, alarm fatigue, and erosion of human interactions were also common:

> *Yeah, definitely. There are some things that require our own judgment. I mean when we are trained, we need to correlate that to how the patient is clinically, like what are the symptoms, how they look, how they appear, what are some of the signs that they have.* [Physician ID 14]

> *I think it’s helpful but I also don’t know like how reliable it will be, because like. for example, patient is just agitated for a while, then there’s a spike in the heart rate. Then umm, if we’re going send alerts each*

XSL•FO
RenderX

> *time, then it's not going to be helpful.* [Physician ID 16]
>
> *OK, since I can already get my vitals there, right, so it might decrease my bed time interaction.* [Nurse ID 06]

From a systems perspective, IT and operations staff highlighted the need for reliable infrastructure and clear governance to prevent failures and excessive maintenance burdens. Health care redesign staff added that bureaucratic barriers contributed to delays and hindered innovation:

> *While it is important to have all this checks in place [security checks prior to implementation], I'm not too sure that they understand that sometimes it's just a pilot.* [HCRD ID 18]
>
> *Imagine today there's a concept of virtual nursing where one nurse can take after 50 patients over screens.... But if the technology fails, what happens to that 50 patients?* [IT ID 08]

Change Commitment: Transparency and Trust in Management

Perceptions of strategic involvement in the smart ward initiative varied across professional groups. Some ground nurses felt excluded from the design and decision-making process despite being the primary users, contributing to a sense of disconnection. In contrast, physicians were less concerned about their level of involvement, noting that they lacked time to innovate but were open to providing discrete feedback. IT and operations staff reported a more structured advisory role—focused on implementation rather than shaping the initiative's direction:

> *So we need to know everything that is going on in the ward, but when they didn't inform us, we feel like we are not a part of it. So yeah, that's how we feel.* [Nurse ID 04]
>
> *To come up with those ideas and the vision and the direction, I don't think most of us honestly have the time.* [Physician ID 13]
>
> *No, I think we do have quite regular updates and meetings to keep each other informed.* [Operations ID 01]

IT and health care redesign staff generally viewed management as supportive of innovation but recognized that maintaining momentum required constant buy-in. Others felt that innovation was not always a top priority, and that fear of failure among senior management constrained their ambition. Nurses were the most skeptical, describing trust as contingent on seeing concrete results from technological investments:

> *I think sometimes it's the fear to fail that actually hinders the progress.* [HCRD ID 19]
>
> *We have a very strong management that says, no matter what, we will be behind support you, you will get this. You think that other places can have this?* [IT ID 02]

Change Efficacy: Awareness and Shared Understanding

Nurses and physicians valued the initiative's aims but felt uncertain about next steps and timelines, reflecting limited communication and shared understanding. In contrast, IT and redesign staff, who were closely involved in planning, described the project as a concrete realization of digital transformation goals:

> *What is happening now? Are we on to this phase? What phase now? What are we waiting for now? when are the trainings going to be?* [Nurse ID 05]
>
> *Smart Ward is a place that where we can put our dreams into real life.* [IT ID 02]

Health care redesign staff noted that inconsistent communication and technical jargon often hindered cross-departmental engagement, although they actively worked with IT to bridge these gaps. They described ongoing efforts to involve all stakeholders but acknowledged that achieving full transparency remained difficult:

> *You know, sometimes we try to like, you know where we have like new products that we want to bring in, new technology that we want to bring in, we will try to invite everyone.* [HCRD ID 19]
>
> *Actually I think IT and the business right can never have a common language because our language is so much more technical right? We can only start by teaching them. So when they time come to us with the business requirements, we translate that into a technical term and then we share, this is what you call it.* [IT ID 08]

Change Efficacy: Resources, Training, and Staff Capability

All participants recognized the need for adequate resources for technology implementation. Nurses highlighted that insufficient staffing would hinder technology use, while IT struggled to hire and retain skilled staff due to competition with industry roles. Health care redesign staff also called for greater IT resourcing:

> *There are certain personnel in IT that we work with very closely. But if you were to ask me if they are dedicated just solely doing our projects, I don't think so.* [HCRD ID 19]
>
> *It's always difficult because being in public healthcare, right, our pay is relatively lower than private sector. So, you always have this fight with the private sectors that you want somebody good, right? Yeah, but you cannot pay.* [IT ID 08]

IT staff further speculated on the long-term scalability of smart ward, given budget constraints and ongoing operational costs. Others noted the challenge of needing external funding to support their efforts:

> *We are running short of money for public healthcare. Because over the years we have been commissioning more and more IT systems, right and each time we commission an IT system, it actually draw an*

XSL•FO
RenderX

> *operating cost. It increases our operating cost. And if we don't start to consolidate them, some of the systems, right, this will just continue to grow.* [IT ID 08]

> *We ourselves, we are also supposed to look for the funds, for our pilot projects.* [HCRD ID 19]

All teams emphasized the importance of training to ensure competency on the ground but described inconsistent provision and unclear responsibility for ongoing training. This lack of clarity and execution led to a poor understanding of pilot technology and inconsistent user compliance across departments. Health care redesign staff were unconcerned about in-depth training, citing cross-functional support. IT viewed ongoing training as the responsibility of system owners rather than IT itself. Consequently, nursing often initiated in-house training programs, but this did not extend to other departments, creating knowledge gaps across functions:

> *So definitely the team needs to be equipped with skills, like understanding the tech landscape right, the jargon, and basically the landscape, because as an administrator when we have all this data.* [Operations ID 01]

> *Because sometimes the training will come only if the thing is there already.* [Nurse ID 05]

> *Usually we will get the vendor to provide the training to the users before they hand over the system for users to maintain it. So that will be their [system owners] responsibility to share the materials with the staff.* [IT ID 15]

Contextual Factors

Participants generally viewed Alexandra Hospital as a "sandbox" for trials and innovation, which fostered a belief in the feasibility of the smart ward. However, they also recognized that such experimentation is more complex in a hospital environment than in other industries. Several participants emphasized the role of prior technology exposure and mindset in shaping readiness. Those with positive past experiences were more open to change, whereas those with past system failures were more cautious. Strategies such as routine "tech scans" to stay informed and efforts to update staff on technology advancements to boost engagement were seen as beneficial, particularly to encourage those who are comfortable with the status quo:

> *We are designated sandbox [ie, a site for experimentation and testing of innovative ideas] from MOH, okay. We can do, try, trial, and we are not afraid of failing.* [IT ID 02]

> *So I have had to deal with system breakdowns. The backup system was extremely difficult to use for everyone. So it definitely took a toll on everybody.* [Physician ID 12]

Cross-Cutting Influences on Collective Readiness

Differences in staff perceptions reflected not only professional roles but also the hospital's culture. Frontline clinicians, who had less decision-making authority, mainly described the smart ward regarding its day-to-day effects on workload, communication, and patient care. In contrast, staff in IT, operations, and health care redesign, who had greater strategic oversight, tended to focus on systems change factors, including integration, scalability, and long-term feasibility. The hospital's culture of cautious innovation, which encouraged experimentation but was often limited by bureaucratic processes or a fear of failure, also shaped readiness. Taken together, collective readiness for change was shaped not only by perceptions of the technology's value but also by how authority and organizational culture shaped staff experiences and their sense of ownership in the change process.

Discussion

Principal Findings

As technology continues to be deployed in clinical environments, assessing the factors influencing change readiness is essential for its successful implementation and sustained use. We conducted a qualitative study to gauge the organization's readiness to change for the smart ward initiative at Alexandra Hospital. We interviewed a diverse cross-section of health care providers, operational support staff, and the in-house innovation team to identify facilitators and barriers. We found that participants generally saw value in the concept and trusted management. Many appreciated active engagement in the initiative and viewed the hospital's sandbox culture as a key enabler of innovation. However, challenges were evident, including unclear implementation timelines, inconsistent training, and limited resourcing for innovation. Concerns about the overreliance on technology were also prominent, with staff wary of its impact on clinical judgment and system reliability.

A key aspect of change commitment (a core construct of the organizational readiness for change framework by Weiner) is the perception that a proposed change holds value, which is essential for successful implementation [9]. This aligns with the 8-step change model developed by Kotter and the diffusion of innovation theory by Rogers, which emphasizes that if clear benefits are unrecognized, it can lead to a lack of urgency, reduced participation in the process, and may ultimately stall the implementation process [19,20]. In our study, we found that stakeholders from different departments collectively saw the initiative as beneficial, but the underlying perspectives on purpose differed between teams. Similar challenges have been observed in other digital transformation contexts, where divergent stakeholder views have led to misalignment and ultimately constrained implementation efforts [21].

Another driver of change commitment is engagement with stakeholders. Effective and early communication with stakeholders can help build buy-in, address concerns, align expectations, and maintain engagement. In contrast, poor engagement often leads to suboptimal adoption. For example, hospitals that introduced electronic medical records without consulting staff often faced resistance [22]. Conversely, those that involved staff early in technology design and roll-out found that implementation was facilitated [23,24]. Our findings indicated that participants had varying degrees of engagement. Participants closely involved in the initiative viewed the process as collaborative, while some frontline staff felt less informed.

XSL•FO
RenderX

To improve engagement, the nonadoption, abandonment, scale-up, spread, and sustainability framework recommends structured, transparent, and multidirectional communication to engage stakeholders and facilitate adoption [25]. Within the framework proposed by Weiner, such communication can enhance collective change commitment and efficacy by fostering a shared understanding, confidence, and ownership of the change process. Interventions could include workshops, multiple communication channels, tailored messaging, and continuous feedback loops to enhance stakeholder commitment.

Of the contextual factors influencing readiness for change, participants shared that having an innovation culture at Alexandra Hospital is a key enabler for the smart ward initiative. Although relatively new in health care, the "sandbox approach" provides a low-risk environment for testing new ideas or technologies without full implementation [26,27]. This approach may be beneficial when introducing disruptive technology, such as that in the smart ward initiative, as it provides space for learning and experimentation without full implementation [26,27]. The sandbox approach aligns with the model by Weiner, emphasizing that perceived capability (which can be acquired through experimentation) is a core component of change efficacy [9]. However, although there was a belief that a sandbox culture exists at the hospital, the infrastructure to support it was lacking. Clear communication on the purpose and objectives of the sandbox approach could help to establish the support required to achieve a true sandbox environment [28].

Training is another factor in capability development. We found that delayed or inconsistent access to training and a lack of clarity on who is responsible for training led to knowledge gaps on the ground. A lack of training may become problematic, as technology literacy is a known barrier to technology adoption and sustainability [29]. According to the organisational readiness for change model proposed by Weiner, ensuring that stakeholders feel competent and supported is essential for successful implementation, a point also mirrored in the Consolidated Framework for Implementation Research [30]. Effective training overcomes resistance to change by improving awareness, competence, and confidence [20]. These findings align with prior studies showing that insufficient or poorly timed training undermines readiness, while targeted, context-specific programs enhance staff confidence and sustain adoption [25,31]. Efforts should focus on providing timely access to training, defining responsibilities for training delivery, and tailoring content to different stakeholder groups to accommodate varying levels of digital literacy and learning needs [32].

Our findings have several practical implications for hospitals planning or implementing smart ward initiatives. We suggest several practical strategies to support the successful implementation of similar smart ward initiatives. First, efforts should focus on building awareness of the value of such initiatives to ensure a buy-in and consistent understanding. Second, implementation strategies should actively engage staff across all levels, not only clinical leads or innovation champions, to ensure that diverse perspectives are captured and foster collective ownership of change. Third, to operationalize the "sandbox" culture that participants value, institutions should establish concrete support mechanisms, such as a streamlined procurement mechanism, expedited technology safety reviews, and iterative feedback loops to rapidly trial and refine technologies in real-world settings. Fourth, training frameworks should be developed early, with clearly assigned responsibilities for delivery, content tailored to varying digital competencies, and adequate protected time for staff to learn. Finally, leaders should visibly support implementation efforts, actively monitor progress, adapt as needed, and establish mechanisms to surface and address frontline concerns as systems evolve. Together, these actions can improve organizational readiness and enable more effective, inclusive, and sustainable digital transformation in hospital settings.

Strengths and Limitations

We conducted a qualitative study that provides in-depth insights into the readiness of clinicians and support staff members for change. We used an established framework to explore organizational factors influencing technology adoption, which is a clear strength. Our qualitative approach also allowed us to identify context-specific enablers and barriers that may be overlooked in a survey approach. However, the generalizability of our results may be limited, as our data are from a single institution. Although we sought to capture a range of perspectives, the representation from some departments (notably IT and operations) was smaller. In these cases, we included senior management who were directly involved in system-level planning, reflecting the primary function of these departments at this stage of the project. For frontline staff, we recruited a broader sample to capture the experiences and views of those impacted by the smart ward. Although we acknowledge the underrepresentation of some professional groups, thematic saturation was achieved across the main domains of readiness, and the inclusion of both strategic and frontline viewpoints enabled us to interpret the findings holistically. We also cannot rule out sampling bias and the risk of social desirability bias in participant responses. We attempted to mitigate the latter by emphasizing participant anonymity. Another limitation of this study is that we used the organisational readiness for change framework by Weiner to inform the development of the interview guide and to structure the initial deductive analysis. This dual use may have reinforced existing theoretical constructs and constrained the emergence of entirely novel themes. To minimize this risk, we incorporated inductive coding within each domain to capture insights that extend beyond the original framework. Finally, as with all qualitative research, the findings reflect the interpretations of the research team, although efforts were made to enhance rigor through team coding.

Conclusions

By identifying and addressing pertinent factors that influence readiness for change, the enablement of technology implementation is possible. Efforts should focus on project awareness, particularly among ground staff, the development of training frameworks, and the adequate prioritization of an innovation culture. Furthermore, alleviating commonly reported technology concerns, such as overreliance, loss of human touch, and reliability issues, will be key to adoption and sustainability.

XSL•FO
RenderX

Acknowledgments

The authors would like to acknowledge Dr Si Ying Tan for her advice on this project.

Data Availability

All data generated or analyzed during this study are included in this published article.

Funding

This research is supported by the Singapore Ministry of Health's National Medical Research Council RIE2025 Centre Grant Programme (grant NMRC/CG3/003/2022-AH/MOH-0010130-00).

Authors' Contributions

Conceptualization: JS, EHHC, AWY
Formal analysis: JS, HWL, AA, CK, EHHC
Funding acquisition: AWY, EHHC
Investigation: JS, HWL, AA, CK, EHHC
Methodology: JS
Project administration: AWY, EHHC
Resources: AWY
Supervision: JS
Writing and editing: JS, HWL, AA, CK, EHHC, AWY

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional supporting quotes organized by subthemes.
[PDF File (Adobe PDF File), 162 KB - humanfactors_v12i1e81932_app1.pdf]

References

1. Abernethy A, Adams L, Barrett M, Bechtel C, Brennan P, Butte A, et al. The promise of digital health: then, now, and the future. NAM Perspect 2022 Jun 27;2022:10.31478/202206e [FREE Full text] [doi: 10.31478/202206e] [Medline: 36177208]
2. Loria G, Nayak Y, Nishit N. A comprehensive analysis of smart hospital technologies and applications – our experience and narrative review of literature. Apollo Med 2024 Mar 31;21(3):245-249. [doi: 10.1177/09760016241240006]
3. Nwobodo-Anyadiegwu EN, Ditend MN, Lumbwe AK. The benefits and challenges of implementing smart hospital projects: a systematic review. In: Proceedings of the IEEE 28th International Conference on Engineering, Technology and Innovation (ICE/ITMC) & 31st International Association For Management of Technology (IAMOT) Joint Conference. 2022 Presented at: ICE/ITMC & IAMOT '22; June 19-23, 2022; Nancy, France URL: https://ieeexplore.ieee.org/document/10033238 [doi: 10.1109/ice/itmc-iamot55089.2022.10033238]
4. Li J, Carayon P. Health care 4.0: a vision for smart and connected health care. IISE Trans Healthc Syst Eng 2021 Feb 15;11(3):171-180 [FREE Full text] [doi: 10.1080/24725579.2021.1884627] [Medline: 34497970]
5. Redesigning healthcare. National University Health System. URL: https://www.ah.com.sg/about-ah/redesigning-healthcare [accessed 2025-12-03]
6. Ageing and health. World Health Organization. 2025 Oct 1. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2025-12-03]
7. Azzopardi-Muscat N, Zapata T, Kluge H. Moving from health workforce crisis to health workforce success: the time to act is now. Lancet Reg Health Eur 2023 Oct 26;35:100765 [FREE Full text] [doi: 10.1016/j.lanepe.2023.100765] [Medline: 38115956]
8. Fowe IE. Evaluating organizational readiness for change in the implementation of telehealth and mobile health interventions for chronic disease management. AMIA Jt Summits Transl Sci Proc 2021 May 17;2021:210-219 [FREE Full text] [Medline: 34457135]
9. Weiner BJ. A theory of organizational readiness for change. Implement Sci 2009 Oct 19;4(1):67 [FREE Full text] [doi: 10.1186/1748-5908-4-67] [Medline: 19840381]
10. van Gemert-Pijnen JL. Implementation of health technology: directions for research and practice. Front Digit Health 2022 Nov 10;4:1030194 [FREE Full text] [doi: 10.3389/fdgth.2022.1030194] [Medline: 36440459]

XSL•FO
RenderX

11. Gabutti I, Colizzi C, Sanna T. Assessing organizational readiness to change through a framework applied to hospitals. Public Organiz Rev 2022 Apr 06;23(1):1-22. [doi: 10.1007/s11115-022-00628-7]
12. Wong JH, Näswall K, Pawsey F, Chase JG, Malinen SK. Adoption of technological innovation in healthcare delivery: a psychological perspective for healthcare decision-makers. BMJ Innov 2023 Jun 26;9(4):240-252. [doi: 10.1136/bmjinnov-2022-001003]
13. Ronaghi MH. Toward a model for assessing smart hospital readiness within the Industry 4.0 paradigm. J Sci Technol Policy Manag 2022 Oct 21;15(2):353-373. [doi: 10.1108/jstpm-09-2021-0130]
14. Novikov Z, Katz-Navon T. AI implementation in health care: the role of trust, managers' commitment, and training. Acad Manag 2023;2023(1) [FREE Full text] [doi: 10.5465/AMPROC.2023.18866abstract]
15. Grossi A, Hoxhaj I, Gabutti I, Specchia ML, Cicchetti A, Boccia S, et al. Hospital contextual factors affecting the implementation of health technologies: a systematic review. BMC Health Serv Res 2021 May 01;21(1):407 [FREE Full text] [doi: 10.1186/s12913-021-06423-2] [Medline: 33933068]
16. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec 16;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
17. Caci L, Nyantakyi E, Blum K, Sonpar A, Schultes MT, Albers B, et al. Organizational readiness for change: a systematic review of the healthcare literature. Implement Res Pract 2025 May 15;6:26334895251334536. [doi: 10.1177/26334895251334536] [Medline: 40385227]
18. Klingberg S, Stalmeijer RE, Varpio L. Using framework analysis methods for qualitative research: AMEE Guide No. 164. Med Teach 2024 May 21;46(5):603-610. [doi: 10.1080/0142159x.2023.2259073] [Medline: 37734451]
19. Rogers EM. Diffusion of Innovations. 5th edition. New York, NY: Free Press; 2003.
20. Kotter JP. Leading Change. Boston, MA: Harvard Business Press; 2012.
21. Alami H, Lehoux P, Papoutsi C, Shaw SE, Fleet R, Fortin JP. Understanding the integration of artificial intelligence in healthcare organisations and systems through the NASSS framework: a qualitative study in a leading Canadian academic centre. BMC Health Serv Res 2024 Jun 03;24(1):701 [FREE Full text] [doi: 10.1186/s12913-024-11112-x] [Medline: 38831298]
22. Tsai CH, Eghdam A, Davoody N, Wright G, Flowerday S, Koch S. Effects of electronic health record implementation and barriers to adoption and use: a scoping review and qualitative analysis of the content. Life (Basel) 2020 Dec 04;10(12):327 [FREE Full text] [doi: 10.3390/life10120327] [Medline: 33291615]
23. Janssen A, Donnelly C, Elder E, Pathmanathan N, Shaw T. Electronic medical record implementation in tertiary care: factors influencing adoption of an electronic medical record in a cancer centre. BMC Health Serv Res 2021 Jan 06;21(1):23 [FREE Full text] [doi: 10.1186/s12913-020-06015-6] [Medline: 33407449]
24. Zharima C, Mhlanga S, Abdulla S, Goudge J, Griffiths F. What engagement strategies are useful in facilitating the implementation of electronic health records in health care settings? A rapid review of qualitative evidence synthesis using the normalization process theory. Digit Health 2024 Nov 03;10:20552076241291286 [FREE Full text] [doi: 10.1177/20552076241291286] [Medline: 39497787]
25. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
26. Leckenby E, Dawoud D, Bouvy J, Jónsson P. The sandbox approach and its potential for use in health technology assessment: a literature review. Appl Health Econ Health Policy 2021 Nov 13;19(6):857-869 [FREE Full text] [doi: 10.1007/s40258-021-00665-1] [Medline: 34254275]
27. Goo JJ, Heo JY. The impact of the regulatory sandbox on the fintech industry, with a discussion on the relation between regulatory sandboxes and open innovation. J Open Innov Technol Mark Complex 2020 Jun;6(2):43 [FREE Full text] [doi: 10.3390/joitmc6020043]
28. Elvidge J, Crabb N, Delnoij D, Knies S, Lundin D, Houÿez F, et al. Implementing a sandbox approach in health technology assessment: benefits and recommendations. Int J Technol Assess Health Care 2024 Nov 04;40(1):e44 [FREE Full text] [doi: 10.1017/S0266462324000412] [Medline: 39494820]
29. Borges do Nascimento IJ, Abdulazeem H, Vasanthan LT, Martinez EZ, Zucoloto ML, Østengaard L, et al. Barriers and facilitators to utilizing digital health technologies by healthcare professionals. NPJ Digit Med 2023 Sep 18;6(1):161 [FREE Full text] [doi: 10.1038/s41746-023-00899-4] [Medline: 37723240]
30. Damschroder LJ, Reardon CM, Widerquist MA, Lowery J. The updated Consolidated Framework for Implementation Research based on user feedback. Implement Sci 2022 Oct 29;17(1):75 [FREE Full text] [doi: 10.1186/s13012-022-01245-0] [Medline: 36309746]
31. Alotaibi N, Wilson CB, Traynor M. Enhancing digital readiness and capability in healthcare: a systematic review of interventions, barriers, and facilitators. BMC Health Serv Res 2025 Apr 04;25(1):500 [FREE Full text] [doi: 10.1186/s12913-025-12663-3] [Medline: 40186200]

XSL•FO
RenderX

32. Brommeyer M, Whittaker M, Liang Z. Organizational factors driving the realization of digital health transformation benefits from health service managers: a qualitative study. J Healthc Leadersh 2024 Nov 04;16:455-472 [FREE Full text] [doi: 10.2147/JHL.S487589] [Medline: 39524481]

Abbreviations

AI: artificial intelligence

Edited by N von Kalckreuth; submitted 06.08.25; peer-reviewed by C Yi, M Al Zoubi; comments to author 14.10.25; revised version received 11.11.25; accepted 12.11.25; published 18.12.25.

Please cite as:
Lim HW, Sumner J, Ang A, Keck C, Chew EHH, Yip AW
Organizational Readiness for Change in the Era of Smart Hospital Wards: Qualitative Study of Health Care Workers' Insights
JMIR Hum Factors 2025;12:e81932
URL: https://humanfactors.jmir.org/2025/1/e81932
doi:10.2196/81932
PMID:

©Hui Wen Lim, Jennifer Sumner, Abigail Ang, Camille Keck, Emily Hwee Hoon Chew, Alexander Wenjun Yip. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 18.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Association of Virtual Nurses' Workflow and Cognitive Fatigue During Inpatient Encounters: Cross-Sectional Study

Saif Khairat[1,2,3,4*], PhD, MPH; Jennifer Morelli[2*], MPS, RN; Wan-Ting Liao[2*], MPS; Julia Aucoin[5*], DNS, RN, NPD-BC; Barbara S Edson[5*], RN, MBA, MHA; Cheryl B Jones[2,3*], PhD, RN

[1]Carolina Health Informatics Program, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[2]School of Nursing, University of North Carolina at Chapel Hill, 438 Carrington Dr, CB 7460, Chapel Hill, NC, United States

[3]Cecil G. Sheps Center for Health Services Research, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[4]Lineberger Cancer Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[5]Virtual Care Center, UNC Health, Chapel Hill, NC, United States

[*]all authors contributed equally

Corresponding Author:
Saif Khairat, PhD, MPH
Carolina Health Informatics Program, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

Abstract

Background: The virtual nursing delivery model enables the provision of expert nursing care from a remote location, using technology such as audio and video communication, remote monitoring devices, and access to electronic health records. Virtual nurses spend an extensive amount of time on computers to provide care, and little is known about how this workflow may affect and contribute to cognitive fatigue.

Objective: This study aimed to use eye tracking technology and pupil size variation to determine instances of virtual nurse cognitive fatigue during their typical workflow.

Methods: This study examined the virtual nursing workflow by recording and analyzing virtual nurse encounters using eye tracking. This cross-sectional study was conducted during regular 12-hour shifts at a major Southeastern health center in the United States.

Results: The study found that 75% (22/29) of virtual nursing encounters demonstrated a first fatigue instance at 9.8 minutes during patient discharges and at 11.9 minutes during patient admissions.

Conclusions: This study provides valuable insights into virtual nursing workflow design and how it may impact the cognitive fatigue levels of nurses providing inpatient virtual care.

(JMIR Hum Factors 2025;12:e67111) doi:10.2196/67111

KEYWORDS

virtual nursing; telemedicine; cognitive fatigue; eye-tracking technology; eye tracking; eye; nursing; virtual; cross-sectional study; workflow; inpatient; fatigue; pupil size; pupil; tracking; USA; United States; design; cognitive; virtual care; nurse; delivery model; technology; communication; EHR; electronic health record; virtual nurse

Introduction

Over 25% of US registered nurses may quit by 2027 due to burnout and challenging work environments, contributing to high nurse turnover rates [1]. Nurses experiencing burnout are five times more likely to delay care (eg, documentation and patient discharge) than nurses without burnout, leading to long patient wait times [2]. Incomplete nursing documentation threatens patient care and safety [3]. Patient home medication review, an essential element of patient admission, is commonly missed due to fatigue, a lack of training, competing tasks, and insufficient time [4,5]. A lack of medication review can lead to medication errors, readmissions, and increased healthcare costs [6-8]. In addition, nurse burnout can result in delays in care and promote inpatient complications and poor patient satisfaction [9-11].

Healthcare systems in the United States are swiftly adopting a novel care delivery model, virtual nursing [12]. Virtual nursing provides bedside nurses with remote nursing expertise to reduce their workload and alleviate work-related fatigue [13]. Virtual nursing aims to mitigate burnout by reducing the workload and improving documentation completeness at admission and patient education at discharge [14].

A key consideration in implementing this emerging care delivery model is understanding workflow design, which protects virtual nurses from the unintended consequences of continuous virtual patient encounters [15]. Extended screen time and the use of

XSL•FO
RenderX

electronic health records (EHRs) were associated with an increase in burnout and fatigue in bedside nurses [16]; however, the association between continuous screen time and virtual nurses' fatigue levels, a key factor to the sustainability of this model, is unknown. Pupillometry data, which measures the change in pupil size, measures cognitive fatigue [17]. This study examined the association between using virtual nursing and virtual nurse fatigue using eye-tracking technology.

Eye-tracking technology provides real-time insights into cognitive workload and cognitive fatigue by recording eye movements during different states and activities. Pupil size in particular can be an indicator of cognitive workload, where an increase in pupil dilation corresponds to information processing [18,19]. In instances of cognitive fatigue, pupils exhibit less dilation over time or decreased response to stimuli. As the mental effort to process information increases, pupils constrict, indicating mental fatigue or exhaustion [20].

Methods

Participants, Device, and Measurements

We conducted a cross-sectional study of virtual nurses during regular 12-hour shifts at a Southeastern tertiary hospital. We observed virtual nurses during four times of interest (TOIs): preparation (eg, chart review), admission encounter, discharge encounter, and miscellaneous (eg, documentation). All virtual nurses were previously bedside nurses at the same hospital. At the time of observation, six nurses were on staff at the Virtual Care Center during day shifts. Four nurses were observed over a period of 5 days.

We used a noninvasive screen-mounted eye-tracking device (Tobii® Pro Spark, Tobii AB®, Stockholm, Sweden) to capture pupillometry data of the virtual nurses during their shifts. Our target number of observations (ie, sessions) was based on evidence that suggests a sample size between 10 and 50 observations suffices for reliable and valid results [21-23].

The eye-tracking device was set up at the virtual nurses' workstations on the monitor that displays activity in the EHR. Prior to recording sessions, virtual nurses were oriented with the eye-tracking device and instructed to complete their tasks as they normally would. An initial calibration session was completed with each virtual nurse before recordings began to ensure data quality. Due to fluctuations in shifts and general availability, two of the four nurses had a higher number of recordings and TOIs captured. To help mitigate this as a confounder, we focused our analysis on the encounters and TOIs rather than the individual nurse level.

Ethical Considerations

The Institutional Review Board (IRB) at the University of North Carolina approved this study under IRB number 23 - 0859. Information related to the study was provided to nurses prior to participation, and they were informed they could decline to participate or withdraw at any time. All participating nurses provided verbal consent to the study team members prior to eye-tracking recording. Participating nurses were given a US $25 gift card to thank them for their participation. All data collected was anonymized following data collection to protect participant confidentiality.

Study Outcomes

The primary study outcome was cognitive fatigue measured as changes in pupil size. Secondary outcomes were subgroup analysis of fatigue by nursing tasks and patient acuity.

Statistical Analysis

Eye-tracking data were exported, and fatigue counts per minute were calculated by flagging instances where the average pupil size was below the calculated threshold. The fatigue threshold for each recording was calculated by subtracting 1.5 times the SD from the mean pupil size. The Shapiro-Wilk test was used to test for normality and the Mann-Whitney or Kruskal-Wallis tests were used to compare continuous variables. We conducted all analyses with statistical software (RStudio version 2023.12.1; Posit Software, PBC).

To generate figures for visualization, we examined the instances of fatigue in a given recording and tracked them over time. We were most interested in comparing the instances of fatigue during patient encounters (admissions and discharges) and other virtual nursing tasks (preparation and miscellaneous). In Figure 1, the x-axis is represented in continuous time groups of 5 minutes, while the y-axis represents the percentage of virtual nurses who experienced an instance of fatigue in a TOI. As encounter times differ and often decrease over time, the number of recordings that meet a given duration also decreases, leading to a smaller ratio. However, this does not fully explain the higher percentage, as it can also be attributed to a greater number of virtual nurses experiencing fatigue as time progresses. In Figure 2, we note the time to first fatigue in minutes, specifically in patient admission and discharge TOIs, and note the cumulative percentage of virtual nurses who have experienced an instance of fatigue by that period of time. In Figure 3, we compare all recordings of patient encounters and look instead at first fatigue by the level of patient acuity.

XSL•FO
RenderX

Figure 1. Proportion of virtual nurses experiencing cognitive fatigue during virtual nursing sessions over time by 5-minute intervals. TOI: time of interest.

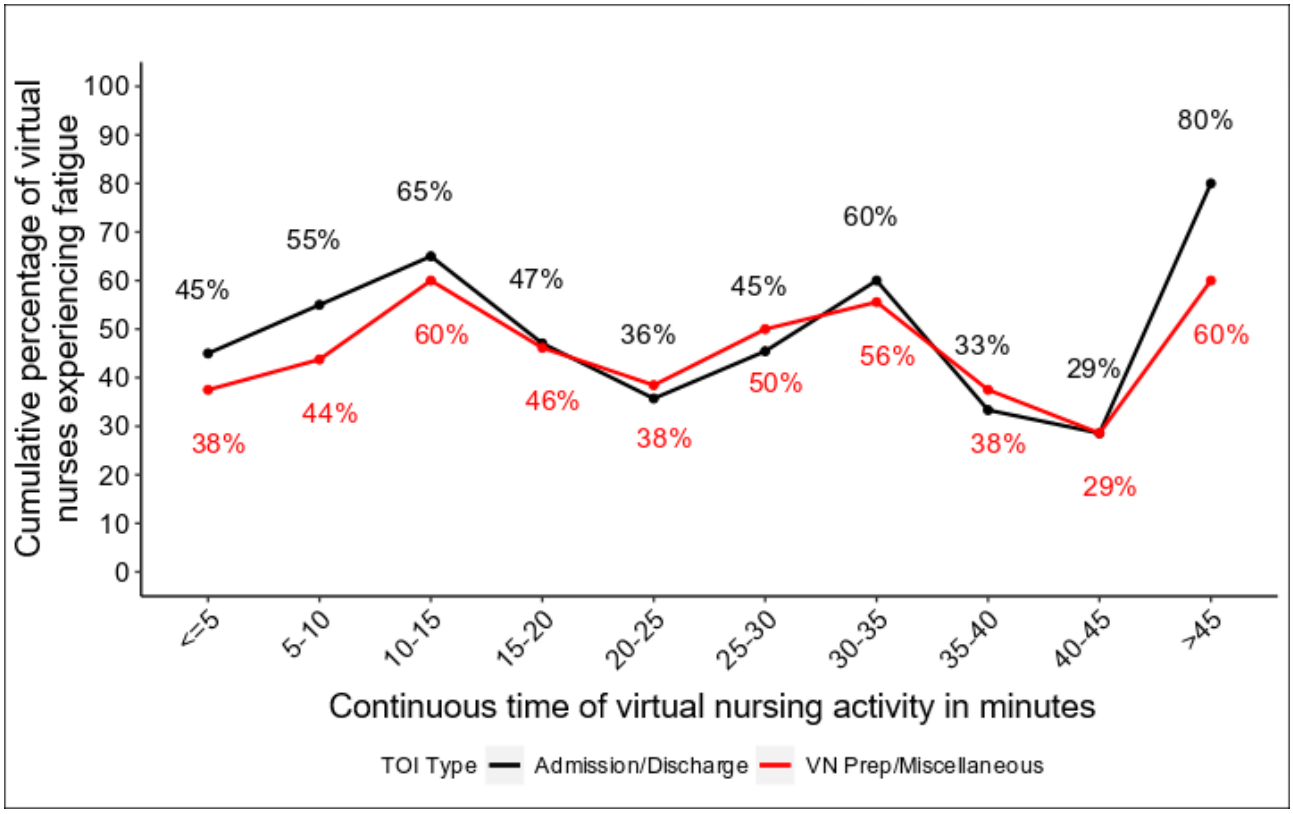
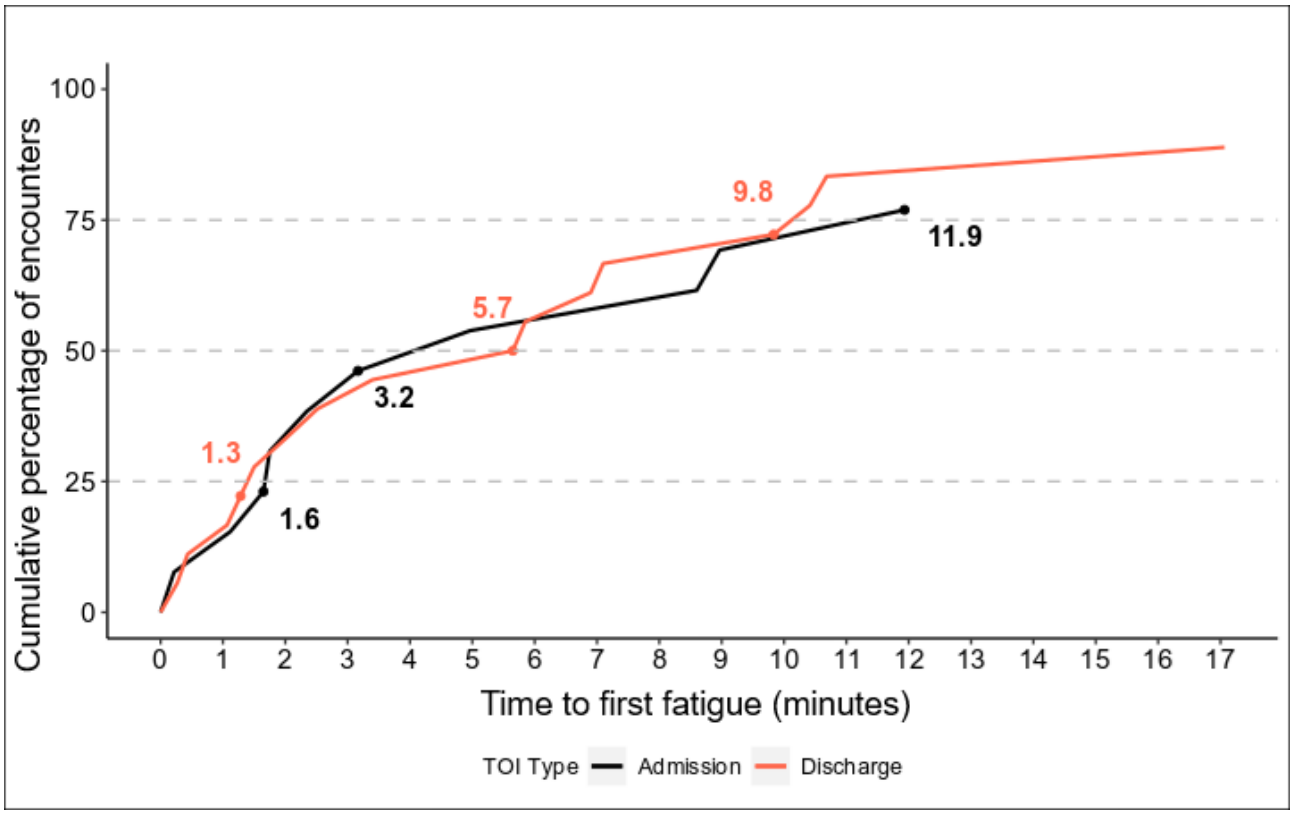
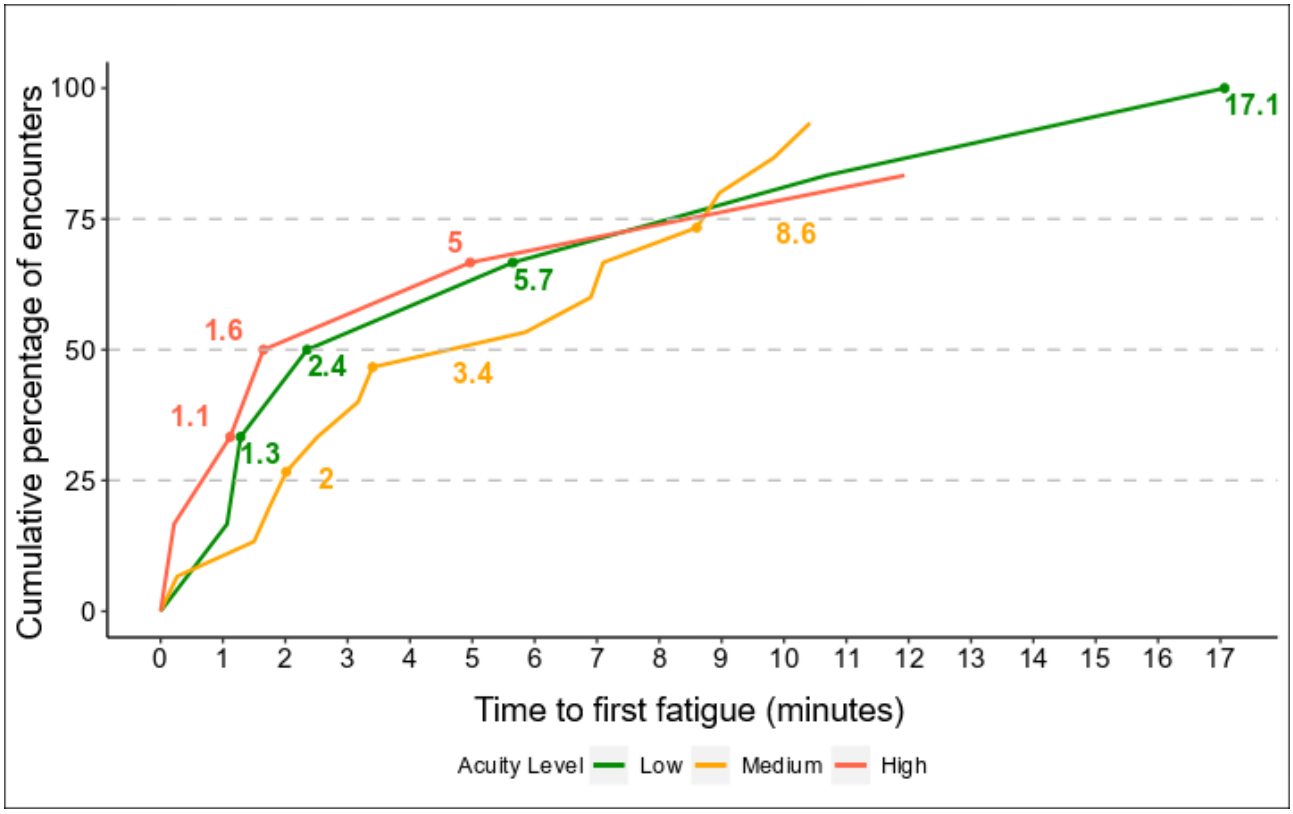


Figure 2. Cumulative percent of virtual nurses experiencing their first cognitive fatigue instance during admission and discharge times of interest (TOIs).



XSL•FO
RenderX

Figure 3. Cumulative percent of virtual nurses experiencing their first cognitive fatigue instance for all times of interest (TOIs) by patient acuity.



To determine patient acuity, we used a combination of factors, including patient age, expected or actual length of stay, principal hospital problem, and severity of the patient unit. These factors were weighted for each patient admission or discharge encounter to determine an acuity rating of low, medium, or high. We used data from the Center for Disease Control National Hospital Discharge Survey [24] to inform this process. The age groups of 18 - 44 years, 45 - 64 years, and >65 years were utilized, with the greatest weight given to patients over 65 years of age. The expected or actual length of stay was categorized into groups of 1 - 3 days, 4 - 6 days, and >7 days, with individuals staying or expected to stay greater than 7 days at the highest acuity. In cases where the principal hospital problem was given (27/29 instances), this was compared to the Center for Disease Control data; the principal hospital problem was compared to the Center for Disease Control data along with the patient's age and length of stay to determine if the patient had lower or higher than the average length of stay. Finally, the severity of the patient unit was taken into consideration, with short-stay units given the lowest acuity, general medical-surgical units given medium acuity, and intensive care units or critical care units given the highest acuity scores.

Results

A total of 22 recordings and over 14 hours of recording time were analyzed, encompassing 83 distinct virtual nursing encounters, Table 1. Of the 83 TOIs, 11 (13%) were admission encounters, 18 (22%) were discharges, 36 (43%) were miscellaneous, and 18 (22%) were preparation tasks. The average (SD) duration was 17 minutes 48 seconds (5.3 minutes) for admissions and 17 minutes 53 seconds (8.9 minutes) for discharges. All four virtual nurses were female, with an average (SD) of 26.3 (7.9) years of nursing experience.

Table . Distribution of workflow tasks for study participants.

	Total observed time	No. of recordings	No. of admissions	No. of discharges	No. of miscellaneous sessions	No. of preparation sessions
Nurse A	6 hours, 8 minutes	11	6	5	18	7
Nurse B	1 hour, 4 minutes	1	0	2	2	1
Nurse C	3 hours, 2 minutes	1	1	3	8	7
Nurse D	4 hours, 27 minutes	9	4	8	8	3
Total	14 hours, 41 minutes	22	11	18	36	18

The average (SD) number of cognitive fatigue instances varied by TOI such that admission encounters had 16.9 (16.4) fatigue instances, discharges had 15 (15.9) fatigue instances, preparation had 4.7 (8.3) fatigue instances, and miscellaneous had 0.42 (1.3) fatigue instances.

XSL•FO
RenderX

Spikes in cognitive fatigue occurred approximately every 15 minutes (Figure 1). For both admissions and discharges, 65% (19/29) of virtual nursing encounters experienced at least one instance of fatigue by 15 minutes, followed by another spike at 30 and 45 minutes. When analyzing the preparation and miscellaneous tasks, we found similar 15-minute interval patterns.

In addition, 75% (22/29) of virtual nursing encounters experienced their first fatigue instance within 9.8 minutes during patient discharges and within 11.9 minutes during patient admissions, Figure 2. Half of virtual nursing encounters (15/29) showed a first fatigue instance within 5.7 minutes during patient discharges and within 3.2 minutes for patient admissions.

We also examined the relationship between patient acuity and fatigue. Of the 29 unique patient admission and discharge encounters, we found that 8 (28%) patients were considered low acuity, 15 (51%) were considered medium acuity, and 6 (21%) were considered high acuity.

Fatigue instances varied by patient acuity (Figure 3). High-acuity patients showed a steeper increase in fatigue than medium and low-acuity patients. Regardless of patient acuity, 75% (62/83) of encounters had their first instance of fatigue within 8.6 minutes.

Discussion

Principal Findings

For the first time, this study quantified the association between the virtual nursing workflow and virtual nurse cognitive fatigue. We report that screen time–related fatigue instances reach maximum levels over 15-minute intervals. This suggests the need to consider more frequent breaks between patient encounters to reduce cognitive fatigue for virtual nurses compared to the standard routine for bedside nurses. We also found that approximately 75% of virtual nurses experience their first fatigue instance after approximately 10 minutes, regardless of the encounter type or patient acuity. By comparison, our previous study suggested that cognitive fatigue occurs after 22 minutes of EHR use by physicians [17].

Suboptimal EHR interface designs and documentation burdens contribute to nurse fatigue [16,25-27]. Virtual nursing can reduce EHR time and the documentation burden for bedside nurses; however, it is paramount to avoid shifting these job stressors to virtual nurses. Virtual nurses who spend extended durations working in the EHR and documenting patient assessments may exhibit cognitive fatigue, a precursor to nurse burnout. This study provides actionable insights into the relationship between using virtual nursing and virtual nurses' cognitive fatigue.

In general, planned and integrated work breaks have been shown to improve the well-being and performance of bedside nurses [28-30]. Our findings support a similar workflow design for virtual nurses that incorporates brief breaks every 15 minutes or between patient sessions. The type of break and its ability to achieve its goals is more important than the frequency of breaks. Prior studies show that breaks should be designed to enable virtual nurses to detach psychologically including active rest and the use of mental health applications [31-33].

For bedside nurses, having an opportunity to rest and physically remove oneself from the work environment can facilitate psychological detachment [31]. However, different work breaks may be needed for virtual nurses as their workflow is primarily computer-based and can become repetitive. One study found that an 8-minute-long virtual reality session during work breaks reduced fatigue among nurses working in high-intensity and cognitively engaging environments [34].

Future research should examine practical work break activities and their impact on virtual nurses' fatigue, including whether different types of breaks (active vs passive) are more effective in reducing instances of fatigue. In addition, future research should use eye-tracking technology to assess the relationship between virtual nursing use and cognitive fatigue among bedside nurses. Although virtual nursing is an emerging care delivery model, there is a critical gap in knowledge around the effect of virtual nursing on organizational and nurse-level outcomes, including nurse burnout.

Limitations

This study was limited by a small sample size of four virtual nurses at one hospital. The study recordings were affected by fluctuations in the virtual nurses' shifts and virtual nurse availability, leading to some nurses having more recorded duration than others. This also resulted in recordings primarily being captured in the middle of virtual nursing work shifts. Future research should examine if instances of fatigue are varied during different times of the shift, including in the beginning and the end of the work shift. The study was limited to a single hospital using the same proficient electronic health record system. Future research should expand on both the nurse sample size as well as the number of institutions to better account for generalizability in the virtual nursing workflow. Additionally, as this study recorded encounters during the live clinical environment and for extended periods of time, eye-tracking accuracy was affected. To ensure an accurate dataset, our analysis only included eye-tracking instances where the percentage of valid pupil data was ≥10%.

To address some of the aforementioned limitations, future research should incorporate larger sample sizes, comparative analyses, and fatigue mitigation strategies to optimize virtual nurse work environments. Future work should incorporate the use of additional measures of cognitive fatigue to ensure accuracy, such as an additional eye-tracker or subjective fatigue surveys for validation. While the study examined patient acuity as a potential confounder, additional contributors to fatigue and confounders may have been unintentionally excluded. Future research should ensure that all contributing factors are addressed in both the recordings and subsequent analyses. In addition, future research should use longitudinal studies to assess the relationship between virtual nursing and virtual nurse fatigue over extended periods to determine potential long-term effects, including exhaustion and burnout. Finally, while the study team used a multifactor approach to assign patient acuity ratings given available data, future research avenues should examine

XSL•FO
RenderX

how fatigue is affected by standardized patient acuity scales, such as the Emergency Severity Index.

Conclusions

This cross-sectional study used eye-tracking methods to understand the effect of continuous screen time on virtual nurses' cognitive fatigue during virtual nursing sessions. The findings emphasize that continuous screen time in virtual nursing tasks contributes to nurses' cognitive fatigue and that specific workflow designs are required to reduce screen time–related fatigue. Health systems implementing virtual nursing should consider frequent breaks away from computer workstations to alleviate the effect of continuous screen time on virtual nurses' well-being. These breaks could incorporate instances of active rest or the use of mental health applications, which have been shown to reduce burnout and fatigue in bedside nurses and other healthcare providers. Findings from this study have direct policy implications and suggestions for more effective workflow redesign.

Acknowledgments

We sincerely thank the nurses and team at the UNC Virtual Care Center for their time and participation. We also thank the following UNC SON students for their assistance: Christiana Akinyemi, Kyndall Estep, Josh Saucedo, Alex Vo, and Linnea Week.

This study was supported by the Health Resources and Services Administration (HRSA) (grant 6 U3GRH40003-01-01), the National Center For Advancing Translational Sciences of the National Institutes of Health (NIH) under award number RC2TR004380, and the UNC School of Nursing. The content of this article is solely the responsibility of the authors and does not necessarily represent the official views of HRSA or NIH, nor does mention of department or agency names imply endorsement by the US Government.

Authors' Contributions

SK and JM wrote the manuscript. JM conducted participant recordings. SK, JM, and WL conducted all statistical analyses. All authors reviewed and edited the final manuscript.

Conflicts of Interest

None declared.

References

1. Suran M. Overworked and understaffed, more than 1 in 4 US nurses say they plan to leave the profession. JAMA 2023 Oct 24;330(16):1512. [doi: 10.1001/jama.2023.10055]
2. White EM, Aiken LH, McHugh MD. Registered nurse burnout, job dissatisfaction, and missed care in nursing homes. J Am Geriatr Soc 2019 Oct;67(10):2065-2071. [doi: 10.1111/jgs.16051] [Medline: 31334567]
3. Ebbers T, Kool RB, Smeele LE, et al. The impact of structured and standardized documentation on documentation quality; a multicenter, retrospective study. J Med Syst 2022 May 27;46(7):46. [doi: 10.1007/s10916-022-01837-9] [Medline: 35618978]
4. Bell T, Sprajcer M, Flenady T, Sahay A. Fatigue in nurses and medication administration errors: a scoping review. J Clin Nurs 2023 Sep;32(17-18):5445-5460. [doi: 10.1111/jocn.16620] [Medline: 36707921]
5. Wondmieneh A, Alemu W, Tadele N, Demis A. Medication administration errors and contributing factors among nurses: a cross sectional study in tertiary hospitals, Addis Ababa, Ethiopia. BMC Nurs 2020;19(1):4. [doi: 10.1186/s12912-020-0397-0] [Medline: 31956293]
6. Goh BQ, Tay AHP, Khoo RSY, Goh BK, Lo PFL, Lim CJF. Effectiveness of medication review in improving medication knowledge and adherence in primary care patients. Proceedings of Singapore Healthcare 2014 Jun;23(2):134-141. [doi: 10.1177/201010581402300207]
7. Tsegaye D, Alem G, Tessema Z, Alebachew W. Medication administration errors and associated factors among nurses. Int J Gen Med 2020;13(1621-32):1621-1632. [doi: 10.2147/IJGM.S289452] [Medline: 33376387]
8. Veselik Z. Personalised, predictive and preventive medication process in hospitals-still rather missing: professional opinion survey on medication safety in Czech hospitals (based on professional opinions of recognised Czech health care experts). EPMA J 2014;5(1):24834138. [doi: 10.1186/1878-5085-5-7] [Medline: 24834138]
9. Recio-Saucedo A, Dall'Ora C, Maruotti A, et al. What impact does nursing care left undone have on patient outcomes? Review of the literature. J Clin Nurs 2018 Jun;27(11-12):2248-2259. [doi: 10.1111/jocn.14058] [Medline: 28859254]
10. Garcia CDL, Abreu LD, Ramos JLS, et al. Influence of burnout on patient safety: systematic review and meta-analysis. Medicina (Kaunas) 2019 Aug 30;55(9):31480365. [doi: 10.3390/medicina55090553] [Medline: 31480365]
11. Vahey DC, Aiken LH, Sloane DM, Clarke SP, Vargas D. Nurse burnout and patient satisfaction. Med Care 2004 Feb;42(2 Suppl):II57-II66. [doi: 10.1097/01.mlr.0000109126.50398.5a] [Medline: 14734943]

XSL•FO
RenderX

12. Teledoc Health. 2022 Annual Telehealth Benchmark. 2022. URL: https://assets.ctfassets.net/l3v9j0ltz3yi/7HoJ4wWtk6uurpsSe2HQOC/3eb5cbcd7291602e52fe5b4c5eb2f76e/Teladoc_SurveyResults_final.pdf [accessed 2023-04-05]
13. National Academy of Medicine. In: Wakefield MK, Williams DR, Le Menestrel S, Flaubert JL, editors. The Future of Nursing 2020-2030: Charting a Path to Achieve Health Equity: The National Academies Press; 2021:466. [doi: 10.17226/25982] [Medline: 34524769]
14. Khairat S, Morelli J, Qin Q, et al. Virtual nursing's effect on ED efficiency and quality of care. Am J Emerg Med 2025 Feb 17;91(59-66):59-66. [doi: 10.1016/j.ajem.2025.02.024] [Medline: 39987629]
15. Khairat S, Morelli J, Edson BS, Aucoin J, Jones CB. Needs assessment of virtual nursing implementation using the Donabedian framework. CIN 2024 Dec 5;43(3):39636230. [doi: 10.1097/CIN.0000000000001229]
16. Khairat S, Xi L, Liu S, Shrestha S, Austin C. Understanding the association between electronic health record satisfaction and the well-being of nurses: survey study. JMIR Nurs 2020;3(1):e13996. [doi: 10.2196/13996] [Medline: 34345776]
17. Khairat S, Coleman C, Ottmar P, Jayachander DI, Bice T, Carson SS. Association of electronic health record use with physician fatigue and efficiency. JAMA Netw Open 2020 Jun 1;3(6):e207385. [doi: 10.1001/jamanetworkopen.2020.7385] [Medline: 32515799]
18. Mathôt S, Vilotijević A. Methods in cognitive pupillometry: design, preprocessing, and statistical analysis. Behav Res Methods 2023 Sep;55(6):3055-3077. [doi: 10.3758/s13428-022-01957-7] [Medline: 36028608]
19. van der Wel P, van Steenbergen H. Pupil dilation as an index of effort in cognitive control tasks: a review. Psychon Bull Rev 2018 Dec;25(6):2005-2015. [doi: 10.3758/s13423-018-1432-y] [Medline: 29435963]
20. Hatsukawa H, Ishikawa M. Decreased initial pupil size and shortened constriction latency due to negative mood states and mental fatigue in clinical subacute pain models. Physiol Behav 2022 Sep 1;253:113850. [doi: 10.1016/j.physbeh.2022.113850] [Medline: 35598830]
21. Bol N, Boerman SC, Romano Bergstrom JC, Kruikemeier S. An overview of how eye tracking is used in communication research 2016. In: 10th International Conference on Universal Access in Human-Computer Interaction: Springer International Publishing; 2016. [doi: 10.1007/978-3-319-40250-5_40]
22. Dunn MJ, Alexander RG, Amiebenomo OM, et al. Minimal reporting guideline for research involving eye tracking (2023 edition). Behav Res Methods 2024 Aug;56(5):4351-4357. [doi: 10.3758/s13428-023-02187-1] [Medline: 37507649]
23. Kendall L, Klasnja P, Iwasaki J, et al. Use of simulated physician handoffs to study cross-cover chart biopsy in the electronic medical record. AMIA Annu Symp Proc 2013;2013(766-75):766-775. [Medline: 24551374]
24. Centers for Disease Control and Prevention (CDC). Average length of stay for discharges from short-stay hospitals, by age and first-listed diagnosis: United States. 2010. URL: https://www.cdc.gov/nchs/data/nhds/2average/2010ave2_dischargesage.pdf [accessed 2024-05-17]
25. De Groot K, De Veer AJE, Munster AM, Francke AL, Paans W. Nursing documentation and its relationship with perceived nursing workload: a mixed-methods study among community nurses. BMC Nurs 2022 Jan 28;21(1):34. [doi: 10.1186/s12912-022-00811-7] [Medline: 35090442]
26. Bjerkan J, Valderaune V, Olsen RM. Patient safety through nursing documentation: barriers identified by healthcare professionals and students. Front Comput Sci 2021;3:2021-June. [doi: 10.3389/fcomp.2021.624555]
27. Misto K, Padula C, Bryand E, Nadeau K. Nurses' perception of the impact of electronic documentation on the nurse-patient relationship. J Nurs Care Qual 2019;34(2):163-168. [doi: 10.1097/NCQ.0000000000000339] [Medline: 29975218]
28. Stutting HL. The relationship between rest breaks and professional burnout among nurses. Crit Care Nurse 2023 Dec 1;43(6):48-56. [doi: 10.4037/ccn2023177] [Medline: 38035618]
29. Cordoza M, Ulrich RS, Manulik BJ, et al. Impact of nurses taking daily work breaks in a hospital garden on burnout. Am J Crit Care 2018 Nov;27(6):508-512. [doi: 10.4037/ajcc2018131] [Medline: 30385543]
30. Albulescu P, Macsinga I, Rusu A, Sulea C, Bodnaru A, Tulbure BT. "Give me A break!" A systematic review and meta-analysis on the efficacy of micro-breaks for increasing well-being and performance. PLoS ONE 2022;17(8):e0272460. [doi: 10.1371/journal.pone.0272460] [Medline: 36044424]
31. Sagherian K, McNeely C, Cho H, Steege LM. Nurses' rest breaks and fatigue: the roles of psychological detachment and workload. West J Nurs Res 2023 Oct;45(10):885-893. [doi: 10.1177/01939459231189787] [Medline: 37621023]
32. Brown JP, Martin D, Nagaria Z, Verceles AC, Jobe SL, Wickwire EM. Mental health consequences of shift work: an updated review. Curr Psychiatry Rep 2020 Jan 18;22(2):31955278. [doi: 10.1007/s11920-020-1131-z] [Medline: 31955278]
33. Wood AE, Prins A, Bush NE, et al. Reduction of burnout in mental health care providers using the provider resilience mobile application. Community Ment Health J 2017 May;53(4):452-459. [doi: 10.1007/s10597-016-0076-5] [Medline: 28070775]
34. Bodet-Contentin L, Letourneur M, Ehrmann S. Virtual reality during work breaks to reduce fatigue of intensive unit caregivers: a crossover, pilot, randomised trial. Aust Crit Care 2023 May;36(3):345-349. [doi: 10.1016/j.aucc.2022.01.009] [Medline: 35246356]

Abbreviations

EHR: electronic health record

XSL•FO
RenderX

IRB: Institutional Review Board
TOI: time(s) of interest

Edited by E Borycki; submitted 02.10.24; peer-reviewed by L Ritter, N Perisho, R Marshall; revised version received 05.03.25; accepted 06.03.25; published 22.04.25.

Please cite as:
Khairat S, Morelli J, Liao WT, Aucoin J, Edson BS, Jones CB
Association of Virtual Nurses' Workflow and Cognitive Fatigue During Inpatient Encounters: Cross-Sectional Study
JMIR Hum Factors 2025;12:e67111
URL: https://humanfactors.jmir.org/2025/1/e67111
doi:10.2196/67111

© Saif Khairat, Jennifer Morelli, Wan-Ting Liao, Julia Aucoin, Barbara S Edson, Cheryl B Jones. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 22.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Effects of Presenting AI Uncertainty Information on Pharmacists' Trust in Automated Pill Recognition Technology: Exploratory Mixed Subjects Study

Jin Yong Kim[1], BSE; Vincent D Marshall[2], MS; Brigid Rowell[2], MA; Qiyuan Chen[1], MSE; Yifan Zheng[2], PharmD; John D Lee[3], PhD; Raed Al Kontar[1], PhD; Corey Lester[2], PharmD, PhD; Xi Jessie Yang[1], PhD

[1]Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI, United States

[2]College of Pharmacy, University of Michigan, Ann Arbor, MI, United States

[3]Industrial and Systems Engineering, University of Wisconsin-Madison, Madison, MI, United States

Corresponding Author:
Xi Jessie Yang, PhD
Industrial and Operations Engineering
University of Michigan
1640 IOE
1205 Beal Avenue
Ann Arbor, MI, 48105
United States
Phone: 1 7347630541
Email: xijyang@umich.edu

Abstract

Background: Dispensing errors significantly contribute to adverse drug events, resulting in substantial health care costs and patient harm. Automated pill verification technologies have been developed to aid pharmacists with medication dispensing. However, pharmacists' trust in such automated technologies remains unexplored.

Objective: This study aims to investigate pharmacists' trust in automated pill verification technology designed to support medication dispensing.

Methods: Thirty licensed pharmacists in the United States performed a web-based simulated pill verification task to determine whether an image of a filled medication bottle matched a known reference image. Participants completed a block of 100 verification trials without any help, and another block of 100 trials with the help of an imperfect artificial intelligence (AI) aid recommending acceptance or rejection of a filled medication bottle. The experiment used a mixed subjects design. The between-subjects factor was the AI aid type, with or without an AI uncertainty plot. The within-subjects factor was the four potential verification outcomes: (1) the AI rejects the incorrect drug, (2) the AI rejects the correct drug, (3) the AI approves the incorrect drug, and (4) the AI approves the correct drug. Participants' trust in the AI system was measured. Mixed model (generalized linear models) tests were conducted with 2-tailed *t* tests to compare the means between the 2 AI aid types for each verification outcome.

Results: Participants had an average trust propensity score of 72 (SD 18.08) out of 100, indicating a positive attitude toward trusting automated technologies. The introduction of an uncertainty plot to the AI aid significantly enhanced pharmacists' end trust (t_{28}=–1.854; P=.04). Trust dynamics were influenced by AI aid type and verification outcome. Specifically, pharmacists using the AI aid with the uncertainty plot had a significantly larger trust increment when the AI approved the correct drug ($t_{78.98}$=3.93; P<.001) and a significantly larger trust decrement when the AI approved the incorrect drug ($t_{2939.72}$=–4.78; P<.001). Intriguingly, the absence of the uncertainty plot led to an increase in trust when the AI correctly rejected an incorrect drug, whereas the presence of the plot resulted in a decrease in trust under the same circumstances ($t_{509.77}$=–3.96; P<.001). A pronounced "negativity bias" was observed, where the degree of trust reduction when the AI made an error exceeded the trust gain when the AI made a correct decision (z=–11.30; P<.001).

Conclusions: To the best of our knowledge, this study is the first attempt to examine pharmacists' trust in automated pill verification technology. Our findings reveal that pharmacists have a favorable disposition toward trusting automation. Moreover, providing uncertainty information about the AI's recommendation significantly boosts pharmacists' trust in AI aid, highlighting the importance of developing transparent AI systems within health care.

XSL•FO
RenderX

(JMIR Hum Factors 2025;12:e60273) doi:10.2196/60273

KEYWORDS

artificial intelligence; human-computer interaction; uncertainty communication; visualization; medication errors; safety; artificial intelligence aid; pharmacists; pill verification; automation

Introduction

Background

Pharmacists play a pivotal role in ensuring patients receive the correct medications as prescribed by health care providers. This involves a critical verification task, where pharmacists must match the medication dispensed in filled bottles with the prescription labels. Dispensing errors, defined as instances when patients receive the wrong drug or dosage, significantly contribute to preventable adverse drug events, leading to approximately 700,000 emergency department visits and 100,000 hospital admissions each year [1]. Several challenges contribute to these errors including, but not limited to, limitations in current technology; lack of standardized procedures; and the high cognitive workload imposed on pharmacy staff, who often manage numerous tasks simultaneously [1-4]. To enhance patient health outcomes, reduce unnecessary health care costs, and alleviate the burden on pharmacists, it is essential to develop and implement reliable tools that minimize the risk of dispensing errors [5].

Since the 1990s, the implementation of barcode scanning systems has been advocated as a means to reduce medication errors [6]. The adoption of such systems within pharmacies and broader health care environments has led to a notable reduction in medication errors [7-10]. Nevertheless, research indicates that barcode scanning systems are not immune to workarounds and human errors [11-13]. Moreover, these systems do not adequately address the challenges faced by overburdened pharmacists [14-17].

In response to these challenges, pill counting and verification or recognition systems using image classification technologies have emerged [18-21]. Innovations like Eyecon and VIVID use vision-based methods to count medications placed on the tray. More recently, advancements have been made in automated pill identification through feature engineering. For example, Yu et al [22] and Yu et al [23] proposed an automatic pill recognition method based on pill imprints, achieving an accuracy of 86.01% and 90.46%, respectively. Caban et al [18] used a modified shape distribution technique to determine the shape, color, and imprint of a pill to identify the drug. The proposed technique was evaluated with 568 of the most prescribed drugs in the United States and achieved a 91.13% accuracy.

The advent of deep learning has further enhanced the capabilities of automated pill recognition systems [5,24]. For instance, Larios Delgado et al [5] developed a pill recognition method using 2 deep learning models. They used a deep convolutional neural network model for pill blob detection to isolate the pill from the background and then passed the output to a deep learning–based classifier to identify the 5 most likely pills with 94% accuracy [5]. Similarly, Wong et al [25] proposed a deep convolutional network model and achieved a mean accuracy of 95.35%. Lester et al [24] trained a ResNet-18 deep learning model to predict the labeled features of a medication product using an image showing medication inside a filled prescription bottle. In a test set containing 65,274 images of 345 unique oral drug products, the overall macroaverage precision, that is, the mean of precision values for each class, was 98.5%.

Despite the impressive strides in model accuracy, realizing the potential of these technologies is only possible if people establish appropriate trust in them. Trust in automation, defined as "the attitude that an agent will help achieve an individual's goals in situations characterized by uncertainty and vulnerability" [26], is one of the most crucial factors determining the use of automation [27,28]. There is a growing body of research examining people's trust in autonomous and robotic technologies in various domains, including transportation [29-31], health care [32,33], education [34], and defense [35,36]. In addition, researchers have developed various methods to enhance people's (proper) trust in automation or autonomy [37,38], including the use of various graphical representations [37,39-42].

For example, a military perimeter defense experiment conducted by Mercado et al [43] aimed to investigate the role of intelligent agent transparency on operator trust. Participants were tasked with selecting optimal routes for unmanned vehicles, assisted by an artificial intelligence (AI) agent. The AI agent operated at three levels of transparency: (1) basic details only; (2) basic details supplemented with reasoning and rationale; and (3) comprehensive information, including basic details, reasoning, rationale, and uncertainty indication, in a text description. They observed a positive correlation between transparency levels and participant trust. They concluded that providing operators with the agent's reasoning process and uncertainty metrics fostered a deeper understanding of the system's capabilities, thereby enhancing trust and increasing usability [43]. This finding emphasizes the importance of transparent AI systems in supporting effective human-machine collaboration.

Another study investigated the impact of visual explanations on human trust in machine learning systems [40]. Participants performed leaf classifying tasks with or without visual explanations. The leaf examples were presented in 2 formats: images and feature charts. Results revealed that providing visual explanations enhanced trust and confidence in participants' decision-making. Interestingly, the feature charts were designed with intentional omissions of detailed explanations of features to prevent information overload. However, this simplification led participants to struggle to interpret the charts, and expert users expressed a need for more comprehensive feature descriptions to inform their decisions [40]. This insight reveals the importance of integrating visual explanations with a thoughtfully managed information load for appropriate human trust.

XSL•FO
RenderX

Signal detection theory (SDT) is commonly used to study trust in automation by modeling the reliability of automated systems used by human operators. SDT evaluates how AI systems distinguish signals from noise, categorizing the state of the world as either "signal present" or "signal absent." Based on SDT categorization, AI performance results in 4 outcomes: hit (error flagged correctly), miss (error not flagged), false alarm (FA; no error, but flagged incorrectly), and correct rejection (CR; no error and no flag). Correct identifications (hit and CR) increase trust, while incorrect ones (FA and miss) decrease trust [44,45]. Research indicates that FAs typically reduce trust less than misses, prompting designers to design more liberal systems (ie, more willing to flag an error) with higher rates of FAs to ensure potential issues are flagged [46-48]. In the context of pill dispensing, FAs may lead to minor disruptions, while misses could lead to dispensing errors, indicating that a more liberal AI system prioritizing safety by minimizing misses is beneficial.

Objectives

This study, therefore, aimed to explore pharmacists' trust in automated pill verification technology, which is designed to assist with the critical task of medication dispensing. Specifically, we aimed to study the role of presenting AI uncertainty information on pharmacists' trust in the system. The primary hypothesis was as follows:

- H1: Presenting AI uncertainty information of predicted probability in a visualization format will increase AI transparency, leading to enhanced pharmacists' trust in pill verification technology.

Beyond this primary focus, we also explored how pharmacists' trust behavior varied across different AI performance patterns categorized by SDT. Drawing from these arguments, we derived the following hypotheses:

- H2: Misses would result in a more significant decline in trust compared to FAs.
- H3: Furthermore, given the differing consequences associated with the 4 SDT patterns, we speculate that presenting AI uncertainty information might have varying effects depending on the specific type of patterns.

Methods

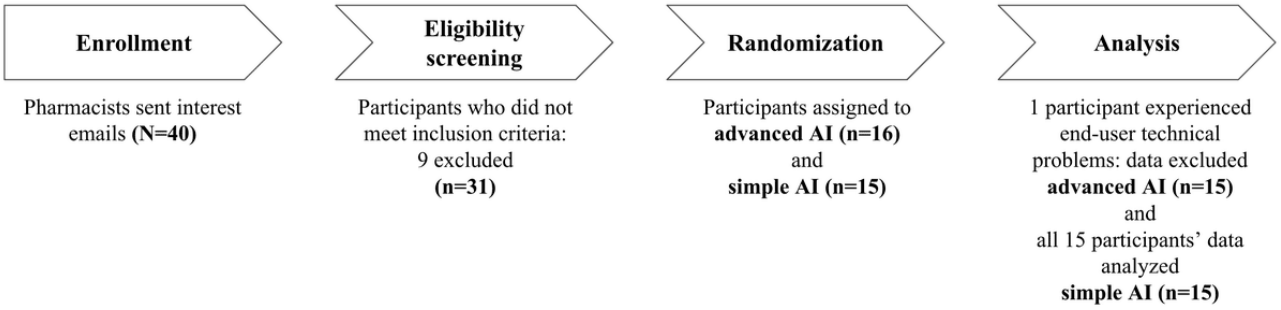
Ethical Considerations

This research was exempt from institutional review board oversight by the University of Michigan (HUM#00213493). Before participating, participants signed an electronic informed consent form, and all data were collected anonymously. Participants received US $150 upon completion of the study.

Recruitment and Participants

Recruitment emails were dispatched to pharmacists through the Minnesota Pharmacy Practice-Based Research Network and the University of Michigan College of Pharmacy Pharmacist Preceptor Network. To meet the inclusion criteria, pharmacists were required to (1) be licensed pharmacists in the United States, (2) be aged at least 18 years, and (3) have access to a laptop or desktop computer with a webcam. Pharmacists who (1) require assistive technology to use the computer; (2) wear eyeglasses with more than one power; (3) have uncorrected cataracts, intraocular implants, glaucoma, or permanently dilated pupils; and (4) have eye movement or alignment abnormalities (eg, lazy eye, strabismus, nystagmus) were excluded from participation in the study (Figure 1). A total number of 30 licensed pharmacists in the United States completed the study. Table 1 shows the demographic information.

Figure 1. Participant recruitment timeline.



XSL•FO
RenderX

Table 1. Participant demographic information (n=30).

Characteristic	Value
Age (years), mean (SD)	39.40 (11.23)
Sex, n (%)	
Female	17 (57)
Male	13 (43)
Practice setting, n (%)	
Community pharmacy	15 (50)
Hospital pharmacy	6 (20)
Specialty pharmacy	1 (3)
Grocery store or mass merchandise pharmacy	1 (3)
Other	7 (23)
Years worked, n (%)	
1-5	7 (23)
6-10	7 (23)
11-20	10 (33)
21 or more	6 (20)

AI Model

The AI model used in this study is a Bayesian neural network that predicts the National Drug Code (NDC), a unique identifier assigned by the Food and Drug Administration to catalog drug products in the United States [49,50]. The Bayesian neural network used a technique known as random dropout [51], applied to a ResNet-34 convolutional neural network architecture [52] to estimate the probability associated with each NDC (ie, each class). The model produced 50 different predictions, where each prediction is a probability vector that quantifies the probabilities that an input belongs to each of the NDCs. The predicted NDC was then attained by finding the highest average probability derived from the 50 predictions.

To train the AI model, we acquired a dataset of 432,974 images from a mail-order pharmacy in the United States. This pharmacy uses a robotic system that counts pills, fills and labels the bottle, captures the images of the contents, and seals the bottle with a cap. The image dataset consists of 1 year's worth of robot-captured images of oral medications, such as tablets and capsules, inside prescription bottles filled by the robotic system. Each image in the dataset is associated with an NDC label and various attributes of color, shape, size, manufacturer, tablet scoring, and imprint. The number of images available for each NDC varied, ranging from 3 to 12,105, with a median of 540 (IQR 257-1291). The medications featured in these datasets were classified into 12 different colors and 7 distinct shapes. The detailed classification is shown in Table 2.

XSL•FO
RenderX

Table 2. Percentage of medication characteristics featured in the training dataset (N=260,119).

Characteristics	Dataset, n (%)
Colors	
White	109,487 (42.1)
Yellow	32,041 (12.3)
Pink	23,585 (9.1)
Orange	18,541 (7.1)
Multicolor	15,289 (5.9)
Green	13,644 (5.2)
Red	13,474 (5.2)
Blue	12,452 (4.8)
Brown	9792 (3.8)
Purple	8184 (3.1)
Turquoise	1858 (0.7)
Gray	1772 (0.7)
Shapes	
Round	128,947 (49.6)
Oval	86,844 (33.4)
Capsule	42,040 (16.2)
Hexagon (6-sided)	1150 (0.4)
Triangle	738 (0.3)
Trapezoid	280 (0.1)
Pentagon (5-sided)	120 (0)

Experimental Testbed and Stimuli

In the experiment, participants performed a pill verification task with the help of an imperfect AI aid that recommends whether to accept or reject a filled medication. The participant's task was to verify whether the filled medication matched the reference image. If the reference image and filled medication did not match, the correct action was to click "reject." If the reference image and filled medication matched, the correct action was to click "accept."

The user interface was designed following pharmacists' feedback from a focus group study conducted by the research team [50]. The interface displayed an image of a filled medication, a reference image, prescription information, and AI aids. There were 2 types of AI aids powered by the same AI model: one aid augmented with an uncertainty plot indicating the degree of certainty (or uncertainty) of the AI recommendation, and the other aid without this feature. Both AI aids recommend the action the pharmacist should take, using 4 checkboxes. A recommendation to accept was indicated when all four checkboxes turned green (Figure 2); otherwise, the recommendation was to reject. For the AI aid with the uncertainty plot, an additional histogram was integrated (Figure 3). The histogram displayed the distribution of the 50 probabilities for the predicted NDC, generated by the 50 predictions. The purpose of the histogram was to provide a visual representation of the certainty or uncertainty level associated with the AI's NDC prediction.

With the help of an AI aid, participants performed a block of 100 pill verification trials. The experimental stimuli for the 100 trials, including the reference NDC and the filled medication, were carefully selected from the dataset of 432,974 images. The selection process ensured a broad representation of colors and shapes (Table 3), while blurry images were excluded to maintain clarity. To minimize learning effects, each reference NDC was intentionally shown no more than twice throughout the experiment.

XSL•FO
RenderX

Figure 2. Interface for the AI aid without the uncertainty plot. Checkboxes indicate AI's recommendation. When all 4 checkboxes are green, the AI advises to accept; otherwise, it advises to reject. AI: artificial intelligence; NDC: National Drug Code.

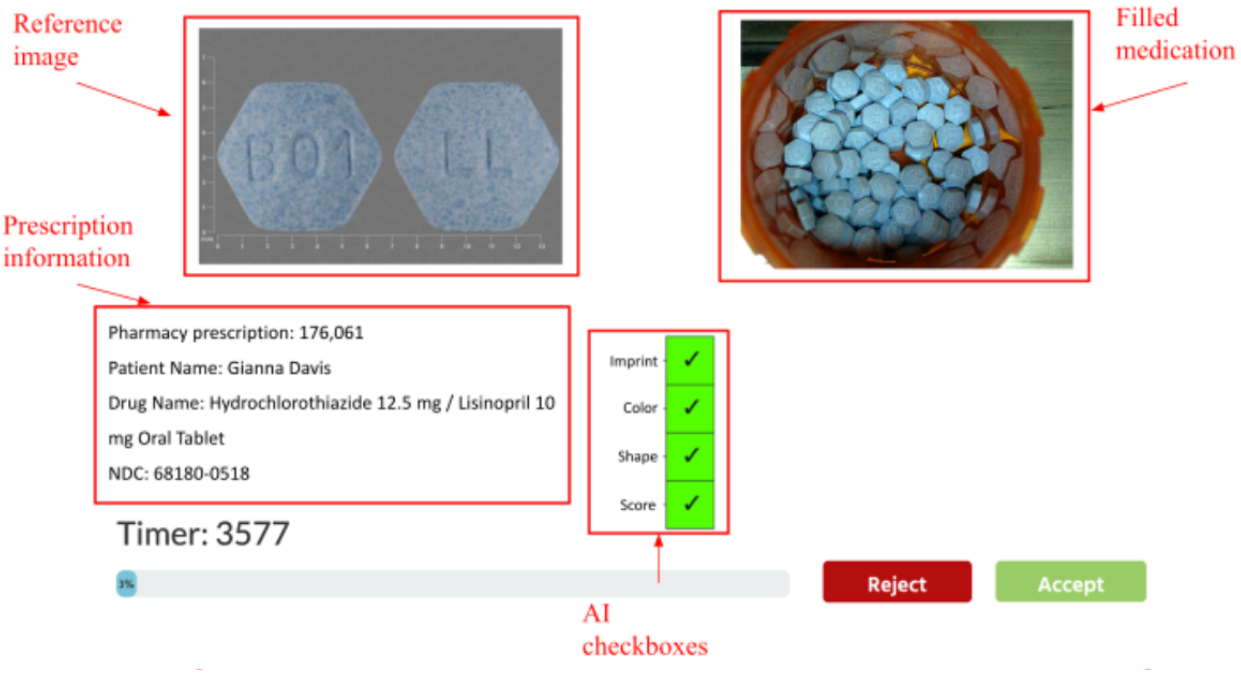
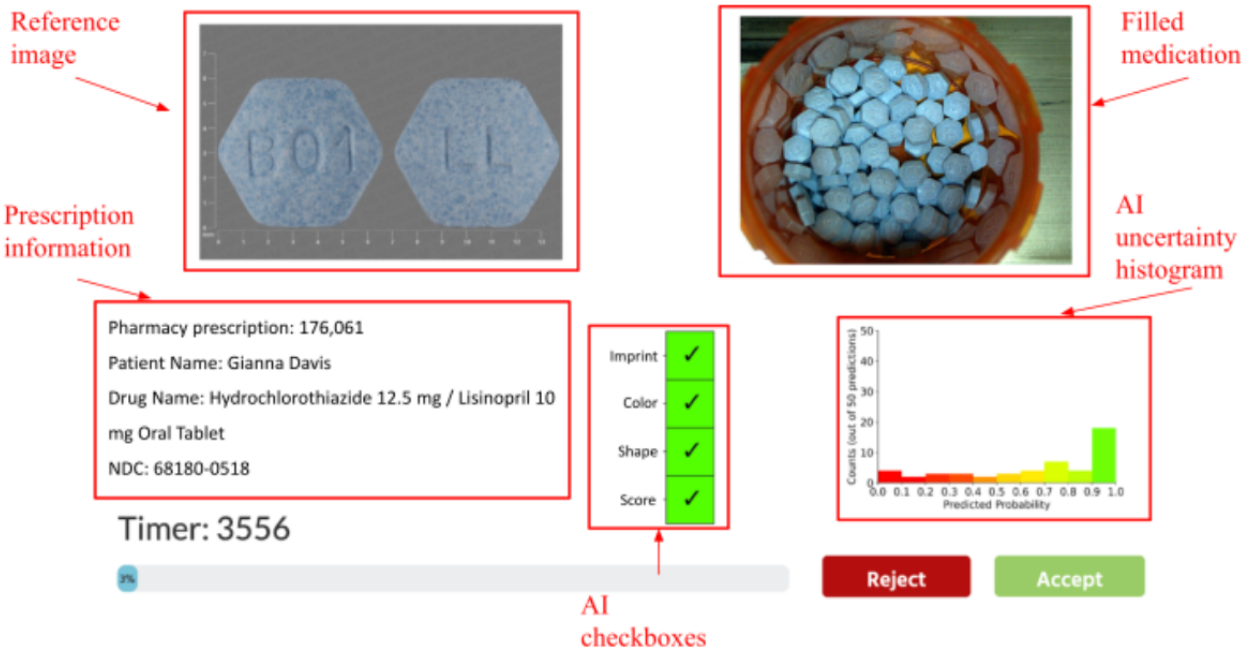


Figure 3. Interface for the AI aid with the uncertainty plot. In addition to Figure 2, the histogram shows the distribution of the 50 predicted probabilities associated with the predicted NDC. AI: artificial intelligence; NDC: National Drug Code.



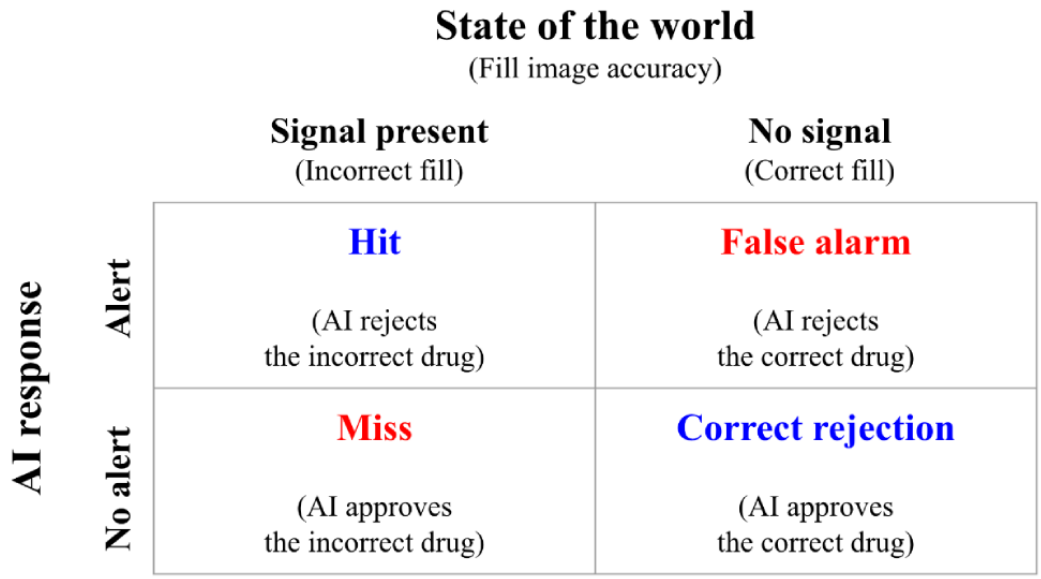
XSL•FO
RenderX

Table 3. Percentage of medication characteristics featured in artificial intelligence–aided trials (N=100) for reference images and filled images.

Characteristics	Reference, n (%)	Filled, n (%)
Color		
White	35 (35)	37 (37)
Yellow	12 (12)	12 (12)
Pink	8 (8)	8 (8)
Orange	8 (8)	8 (8)
Multicolor	3 (3)	3 (3)
Green	6 (6)	6 (6)
Red	3 (3)	2 (2)
Blue	10 (10)	10 (10)
Brown	10 (10)	10 (10)
Purple	4 (4)	3 (3)
Turquoise	1 (1)	1 (1)
Gray	1 (1)	0 (0)
Shape		
Round	54 (54)	51 (51)
Oval	24 (24)	28 (28)
Capsule	19 (19)	19 (19)
Hexagon (6-sided)	2 (2)	1 (1)
Triangle	0 (0)	0 (0)
Trapezoid	1 (1)	1 (1)
Pentagon (5-sided)	0 (0)	0 (0)

Furthermore, the AI aid was not perfect for the 100 trials, that is, it occasionally offered incorrect recommendations. Based on SDT, we mapped out the relationship between the AI's recommendation and the true state of the world [53]. In the context of this experiment, a signal in the world was an incorrectly filled medication. The extent to which the AI recommended rejecting an incorrectly filled medication reflects its ability to detect the signal. The combination of the state of the world and the AI's recommendation resulted in four potential outcomes: (1) the AI rejects the incorrect drug (hit), (2) the AI approves the incorrect drug (miss), (3) the AI rejects the correct drug (FAs), and (4) the AI approves the correct drug (CRs), as shown in Figure 4.

Figure 4. Four potential AI performance patterns according to signal detection theory. AI: artificial intelligence.

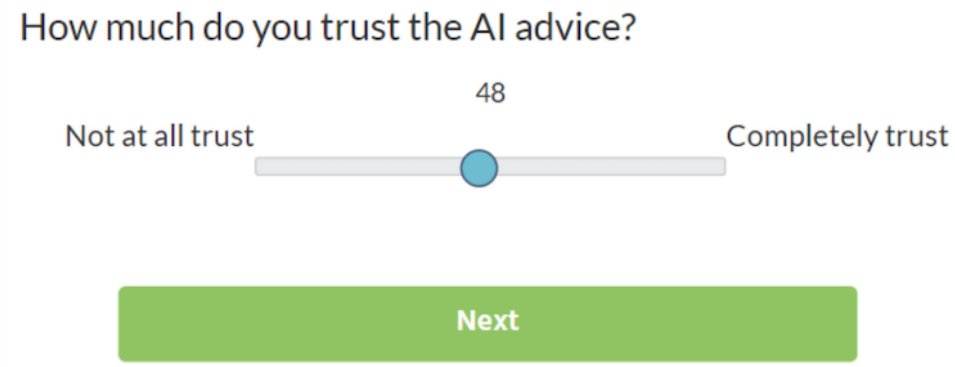


XSL•FO
RenderX

Benchmarking prior literature [54], the base rate was set to be 24%, that is, 24% of the trials contained wrongly filled medication. The AI accuracy was set as 82% to ensure that the AI would be perceived as useful while providing sufficient misses and false alarms [55]. By combining the filled image accuracy and AI accuracy, there were 60 cases of the AI approving the correct drug, 22 cases of the AI rejecting the incorrect drug, 2 cases of the AI approving the incorrect drug, and 16 cases of the AI rejecting the correct drug.

After each trial, participants received performance feedback indicating the correctness of their decision to accept or reject, as well as whether the prescription bottle was correctly or incorrectly filled (ie, "Your decision was correct. The medication was correctly filled"). Following this step, participants reported their trust in the recognition AI on a visual analog scale, with the leftmost point labeled "Not at all trust" and the rightmost point labeled "Completely trust" (Figure 5) [44,56,57].

Figure 5. Participants reported their trust in the recognition AI on a visual analog scale. AI: artificial intelligence.



Experimental Design

The experiment used a mixed subjects design. The between-subjects factor was the type of AI aid, distinguished by the presence or absence of an uncertainty plot. The within-subjects factor was the four potential outcomes: (1) the AI rejects the incorrect drug, (2) the AI rejects the correct drug, (3) the AI approves the incorrect drug, and (4) the AI approves the correct drug (Figure 4).

Half of the participants used the AI aid without the uncertainty plot, and the other half used the AI aid with the uncertainty plot. Each participant completed 2 blocks of 100 trials each. One block involved using the AI aid (either with or without the uncertainty plot), and the other block required participants to perform the task manually. The order of the 2 blocks was counterbalanced. Additionally, benchmarking prior literature [45], the trial sequence was fixed for the 100 trials in each block.

As this study is focused on the pharmacists' trust in AI, data from the manual task block were excluded from the analysis, concentrating the study's findings on interactions involving the AI aid.

Measures

Trust propensity

Before the experiment, we measured participants' trust propensity using the 6-item survey used by Merritt et al [58]. Trust propensity is "a stable, trait-like tendency to trust or not trust others" [59], and the propensity to trust machines reflects a person's tendency to trust machines in general rather than in a particular machine.

End Trust

End trust, *Trust*(*100*), is the participant's final trust rating after interacting with the AI help scenario.

Average Trust

Average trust denotes the mean of moment-to-moment trust ratings collected throughout the experiment.



Trust Change

After each trial *i*, participants reported their *Trust*(*i*) in the AI. We calculated a trust change as follows.

$$\text{Trust change } (i) = \text{Trust}(i) - \text{Trust}(i-1), \text{ where } i=2, 3, \ldots, 100.$$

Since the moment-to-moment trust was reported after each trial, 99 trust changes were obtained from each participant.

Experimental Procedure

The experiment was conducted remotely with interested participants who met the inclusion criteria. Each interested individual was phone screened to determine their eligibility. Before each experiment, participants had a brief web-based meeting with a member of the study team to ensure that the physical environment, including lighting conditions, was suitable for the experiment. Subsequently, the pharmacists were directed to Labvanced's website on their computer and presented with a 15-minute video tutorial that explained how to perform a simulated medication verification task using the testbed interface. Pharmacists were informed that the objective of the task was to determine whether an image of a filled medication

XSL•FO
RenderX

bottle matched a known reference image. The tutorial also explained the 2 AI aids.

Before engaging in the verification task, participants were directed to complete a demographics survey and a trust propensity survey [58]. Additionally, they went through a set of calibration procedures for the eye-tracking software. Participants then completed a block of 100 verification trials using an AI aid—either with or without the uncertainty plot—and another block of 100 verification trials manually, conducted in a counterbalanced order. Upon completion of the 200 trials, participants filled out a postexperimental survey and answered nonmandatory free-response feedback questions. All participants finished each block within the time limit of 60 minutes.

After completing all the tasks and surveys described earlier, participants were invited to a 30-minute debriefing session with one of the study team members. Study team members described 6 concepts of automation evaluation: observability, predictability, directing attention, exploring solution space, adaptability, and calibrated trust [60], and provided an example scenario for each concept. After each description and example, the participants were asked to provide their thoughts on how the concept relates to our system.

Statistical Analysis

Participants' trust propensity, end trust, and trust change when using both types of AI aids were analyzed. First, we conducted a descriptive analysis of the participants' trust propensity. Then, to test our directional hypothesis that AI aids with uncertainty will result in higher end and average trust, we conducted a 1-tailed *t* test. Finally, we analyzed how trust increased and decreased after participants experienced each of the 4 AI performance patterns using mixed-linear models with random intercept. Regression 2-tailed *t* tests were conducted to compare the means between the 2 AI aids for each AI performance pattern. The Kenward-Roger method was used to estimate degrees of freedom. Mixed model (generalized linear models) tests were conducted using the R (version 4.2.2; R Foundation for Statistical Computing) *lme4* package [61]. All statistical significance was determined at the α=.05 level and analyses were carried out using R statistical software [62].

Results

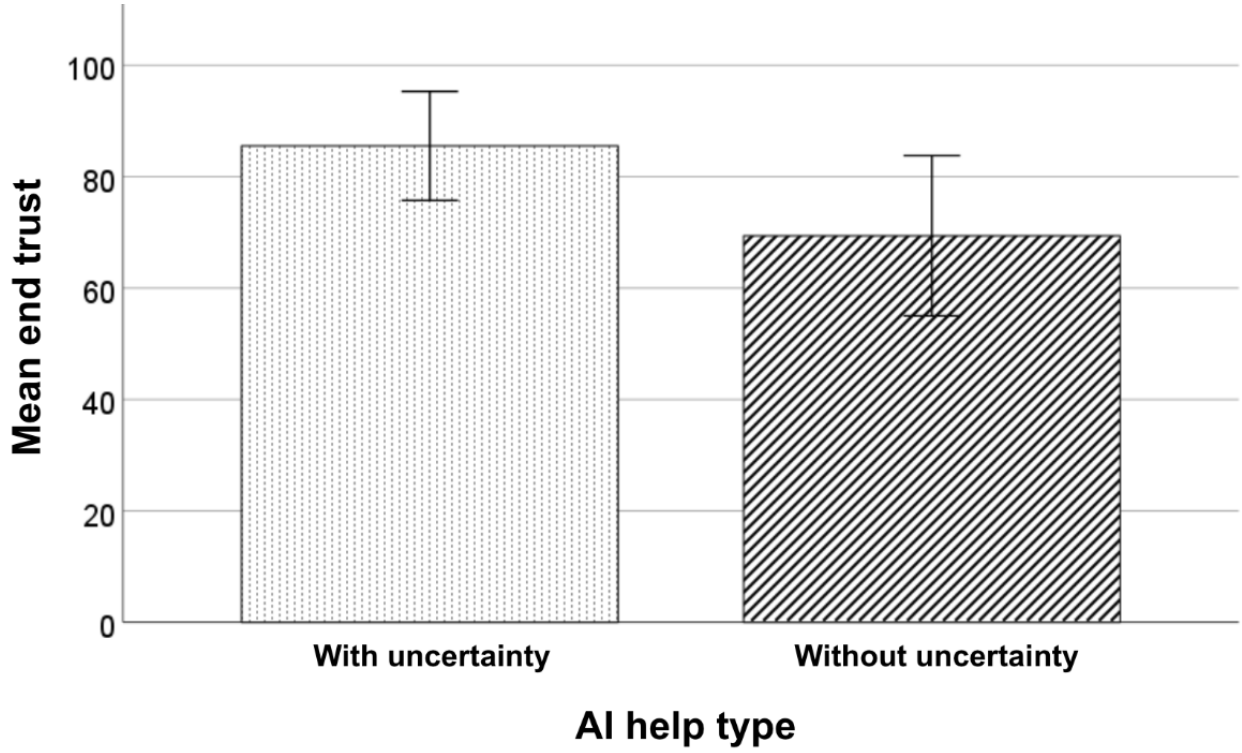
Trust Propensity

Participants had an average trust propensity score of 72 (SD 18.08) out of 100, indicating the participants generally had a positive attitude toward trusting automated technologies. There was no significant difference between the 2 AI aids (t_{28}=–0.854; *P*=.20; Cohen *d*=.312).

End Trust Toward AI Aid

The 1-tailed *t* test indicated that participants trusted the AI aid with the uncertainty plot significantly more than the AI aid without the plot at the end of the experiment (t_{28}=–1.854; *P*=.04; Cohen *d*=–.677; Figure 6).

Figure 6. Mean end trust by AI aid help type (with or without the uncertainty plot). The error bars represent a 95% CI. AI: artificial intelligence.

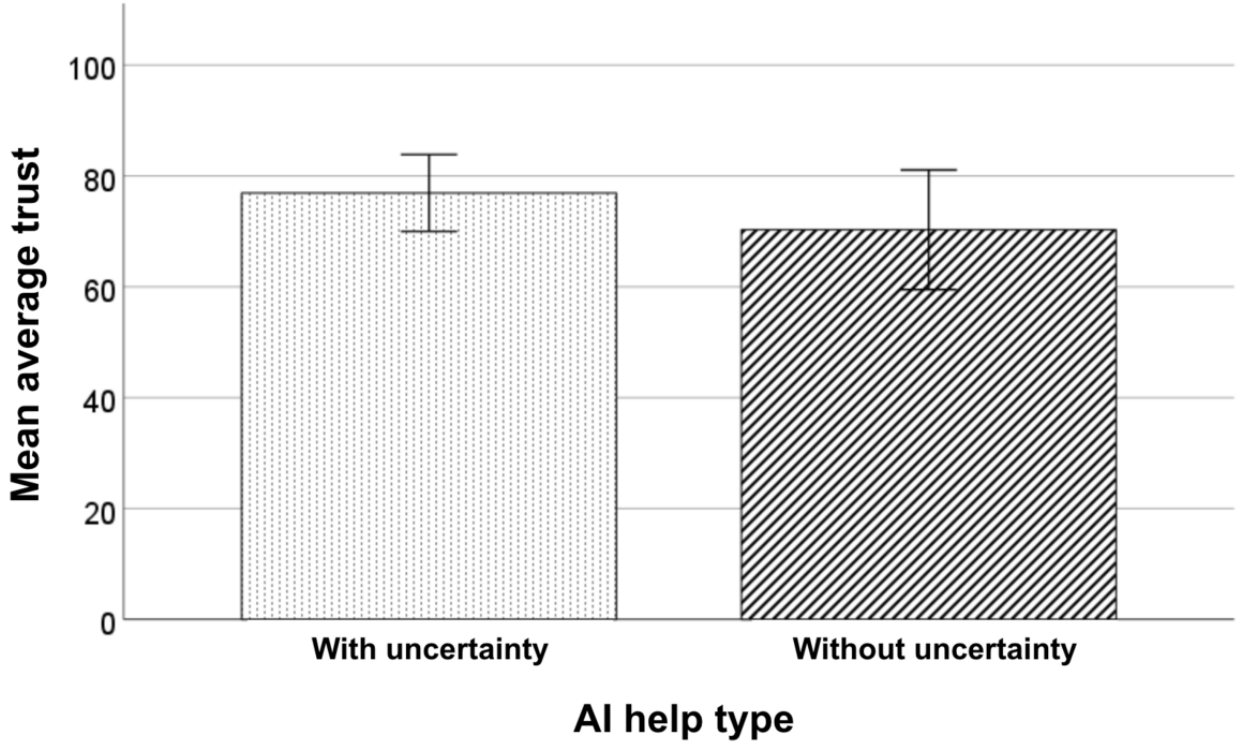


Average Trust Toward AI Aid

Participants showed a slightly higher average trust in the AI aid in the with-uncertainty condition (mean 76.92, SD 13.42) than in the without-uncertainty condition (mean 70.29, SD 20.88; Figure 7). However, the difference did not reach statistical significance (t_{28}=–1.036; *P*=.16; Cohen *d*=–.378).

XSL•FO
RenderX

Figure 7. Mean average trust by AI aid help type (with or without the uncertainty plot). The error bars represent a 95% CI. AI: artificial intelligence.



Trust Change

Figure 8 shows the trust change after participants experienced the 4 AI performance patterns. When the AI approved the correct drug, there was a significantly greater trust increment when participants used the AI aid with the uncertainty plot compared to the one without ($t_{78.98}$=3.93; *P*<.001; Cohen *d*=.214); When the AI approved the incorrect drug, there was a significantly greater trust decrement using the AI aid with the uncertainty plot ($t_{2939.72}$=–4.78; *P*<.001; Cohen *d*=.712). Interestingly, when the AI rejected the incorrect drug, we observed a decrement of trust for participants using the AI aid with the uncertainty plot ($t_{509.77}$=–3.96; *P*<.001; Cohen *d*=.312). When the AI rejected the correct drug, both AI help types showed a decrease in trust and there was no statistical difference between them ($t_{856.57}$=–0.68; *P*=.49; Cohen *d*=.045). Overall, participants using the AI aid with the uncertainty plot displayed a large magnitude of trust adjustment. In addition, we observed a significant "negativity bias" in that the magnitude of trust change when the AI made an error (ie, the AI approves the incorrect drug or the AI rejects the correct drug) was significantly larger than the magnitude of trust adjustment when the AI provided correct recommendations (generalized linear model test; *z*=–11.30; *P*<.001).

To examine variation in pharmacists' trust behavior across different AI performance patterns categorized by SDT, we initially combined the data from both the with and without uncertainty help scenarios. Trust decreased significantly more when the AI approved the incorrect drug compared to when the AI rejected the correct drug (t_{509}=–4.687; *P*<.001; Cohen *d*=.475). This trend was observed in the with-uncertainty AI help scenario (t_{254}=–4.91; *P*<.001; Cohen *d*=.758). However, in the without-uncertainty AI help scenario, although there was a greater trust decrease in trust when the AI approved the incorrect drug than when the AI rejected the correct drug, the difference was not statistically significant (t_{254}=–1.14; *P*=.255; Cohen *d*=.014).

As we measured participants' trust toward AI continuously, we calculated the autocorrelation between the trust ratings. Autocorrelation measures the relationship between a variable's current value and its past values in a time series. Figure 9 shows the mean autocorrelation as a function of time separation between the ratings. For both AI aids, the correlation decreased as the time separation increased. The AI aid with the uncertainty plot had a lower mean autocorrelation compared to the aid without the uncertainty plot (Figure 9).

XSL•FO
RenderX

Figure 8. Trust change for different AI performance patterns. Red shades represent negative trust change, and blue shades represent positive trust change.

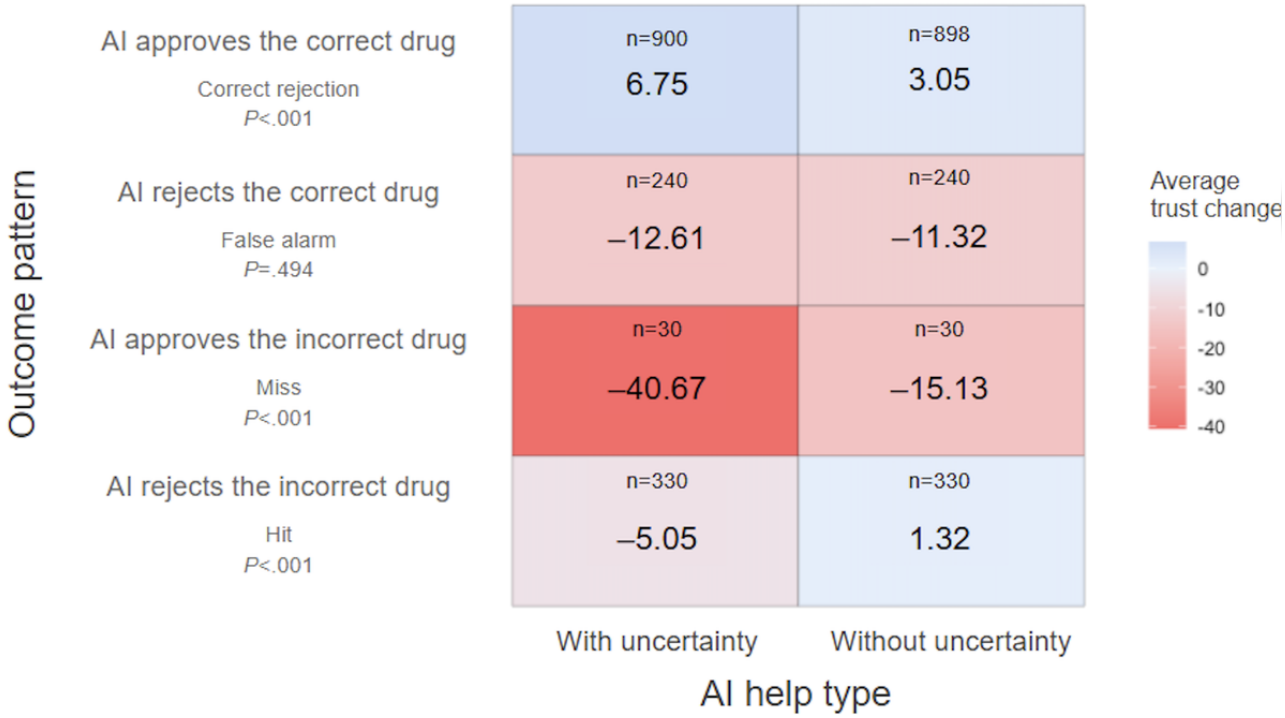
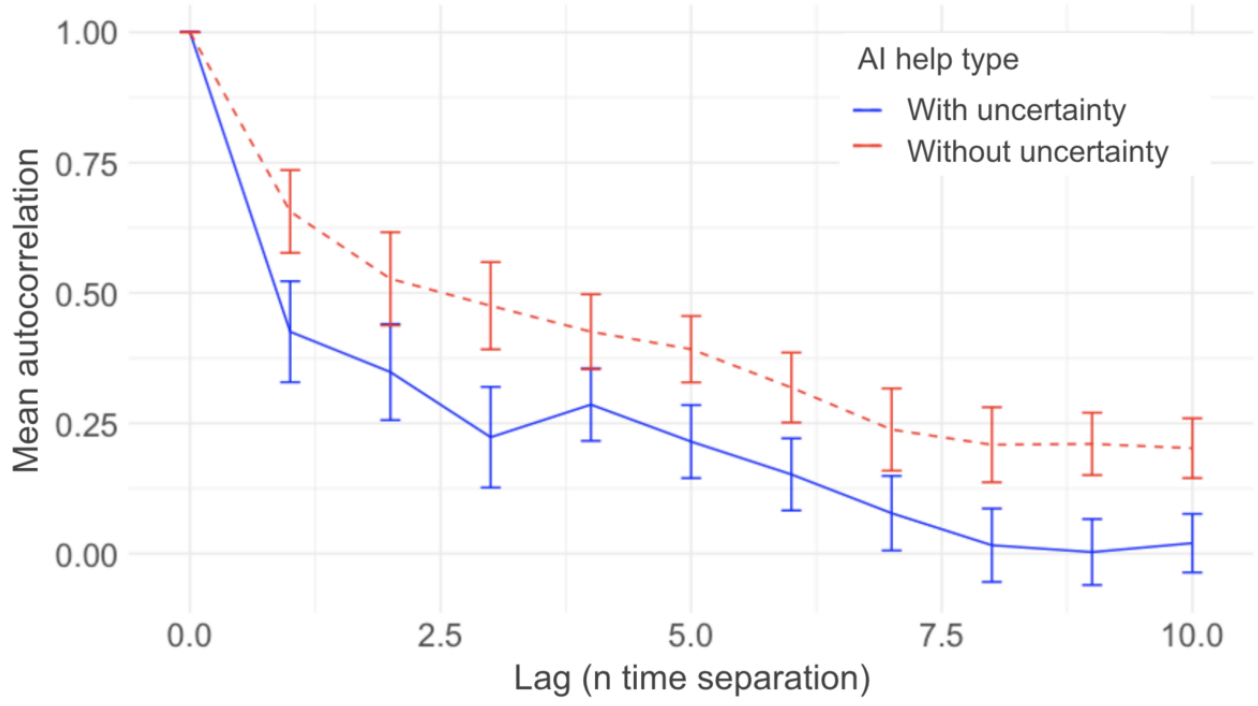


Figure 9. Autocorrelation of trust as a function of time separation. The blue solid line represents AI aid with the uncertainty plot, and the red dashed line represents the AI aid without the uncertainty plot. The error bars represent 2 SEs. AI: artificial intelligence.



Discussion

Principal Findings

This study aimed to investigate pharmacists' trust in automated pill verification technology and how the presentation of AI uncertainty information would influence them. Overall, the findings revealed that pharmacists have a favorable disposition toward trusting automation, and including the AI's uncertainty information increases pharmacists' trust in the AI recommendation *(*Figure 6*)*.

Comparison With Prior Work

The propensity to trust automation refers to an individual's general inclination to trust automated or autonomous systems, shaped by their past experiences and future expectations [58,59]. Research has shown that levels of trust propensity vary among

XSL•FO
RenderX

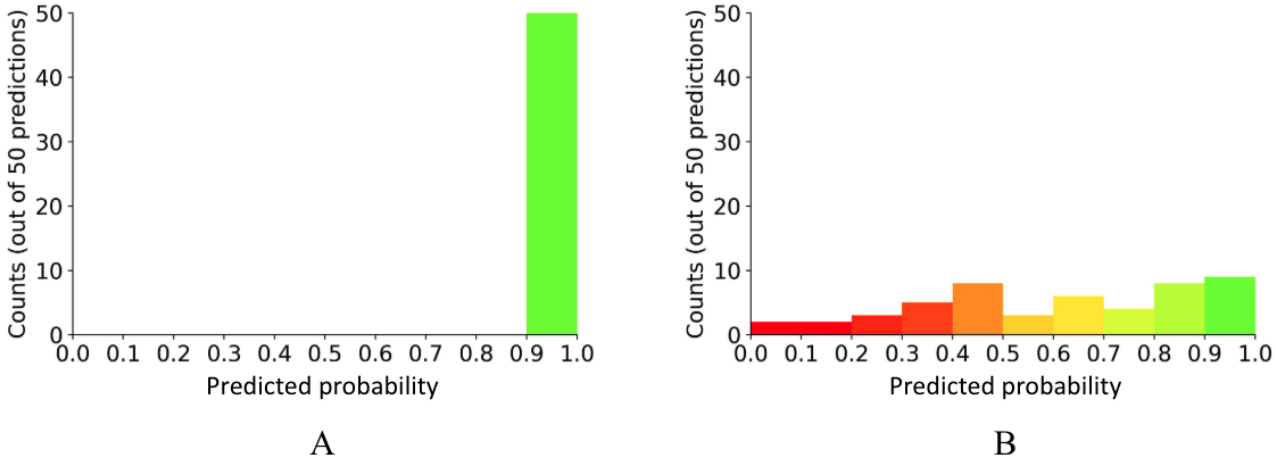
individuals. For example, an early study by Merritt et al [58], which included 69 college students (average age of 25 years), found an average trust propensity of 3.56 (SD 0.6) on a 7-point Likert scale [59]. More recently, Montag et al [63] surveyed 289 participants aged between 18 and 70 years and reported their propensity to trust automation to be 4.98 (SD 1.06) after converting to a 7-point Likert scale. Similarly, Miller et al [64] reported a trust propensity score of 4.97 (SD 1.21) from a smaller cohort of 28 participants aged between 18 to 60 years. Another investigation by Yang et al [65] with 75 adults (mean age 23.0) split into 3 groups reported trust propensity scores of 72.6 (SD 14.8), 69.4 (SD 10.4), and 69.4 (SD 14.4), equivalent to average scores of 5.08 (SD 1.89), 4.86 (SD 1.62) and 4.86 (SD 1.86) on a 7-point Likert scale.

In line with these findings [63-66], our study revealed that pharmacists generally have a favorable disposition toward trusting automation, with an average rating of 72 (SD 18.08) on a 100-point scale, or 5.03 (SD 2.09) on a 7-point Likert scale. This positive attitude may be attributed to the frequent use of automated technologies, such as barcode scanners and pill counters in their daily work [6,7]. Additionally, as expected, no significant difference was observed between the groups using different AI aids because the participants were randomly assigned to use either of the AI aids.

Examining pharmacists' end trust, our findings reveal that the AI aid with the uncertainty plot significantly enhanced the end trust scores. We attribute this enhancement to the increased transparency achieved through the presentation of a histogram showing the distribution of 50 predicted probabilities. While AI advancements promise to improve human performance, a prevailing issue is the perception of AI as a "black box." This lack of transparency contributes to a lack of trust in AI and can undermine team performance [66-68]. The higher end trust observed in participants using the AI aid with the uncertainty plot indicates that making the AI more transparent by revealing its decision-making process can foster a higher level of trust in automation. Participants 14 and 24 captured this sentiment well stating: "As soon as the uncertainty plot became red and yellow, I slowed down, which could be helpful because sometimes slowing down when there is uncertainty, just knowing there's uncertainty, is enough" (P14) and "if the uncertainty plot was all green bar and AI thought it was doing a 100% accurate job then it was easier to make my decision" (P24).

Regarding the dynamics of trust, that is, moment-to-moment trust change, when AI approved the correctly filled bottle, we noted trust increments for both AI aids. Furthermore, the inclusion of uncertainty information led to a larger increment in trust compared to when such information was absent. When the AI mistakenly approved the incorrect drug, we observed a significant trust decrement for both AI aids, potentially attributed to the adverse outcome associated with the wrong medication. Furthermore, the trust decrement was significantly larger when the uncertainty information was shown. This pronounced trust decrement could have resulted directly from the distribution of the histogram: participants were shown a histogram indicating a high level of certainty (Figure 10A). Therefore, participants may have perceived the error made by the AI aid as a "confident" error and therefore reduced their trust even more. Studies examining likelihood alarms reported that highly likely alarms (ie, "confident" alarms) engender a greater decline in momentary trust upon automation failures [57,69].

Figure 10. Uncertainty plots with (A) narrower and (B) wider IQRs. IQR is a measure of statistical dispersion, or how spread out the data points are.



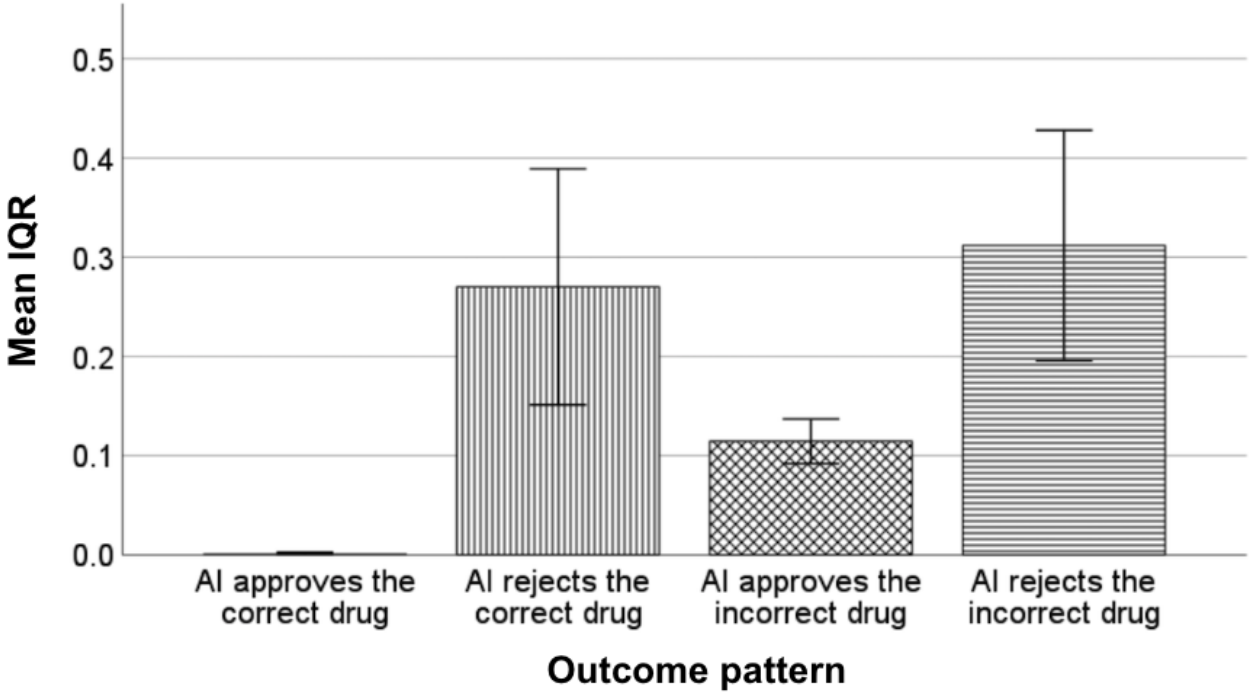
Our study also offers additional validation of prior findings regarding the SDT modeling of trust [47,48]. When the AI mistakenly approved the incorrect drug (miss), a greater trust decrease appeared compared to when the AI rejected the correctly filled bottle (FA). However, the difference was statistically significant only in the with-uncertainty AI help type. This finding may provide further evidence that "confident" AI errors lead to a greater trust decrement [57,69].

Intriguingly, when the AI rejected an incorrectly filled bottle, the absence of uncertainty information resulted in an increase in trust, whereas the presence of such information led to a decrease in trust. Such contrasting results could have stemmed from the uncertainty plots influencing the participants' decision-making process. When the AI approved the correct drug, all uncertainty plots presented to participants showed a consistent solid green bar (Figure 10A). However, when the AI rejected the incorrect drug, the IQR of the uncertainty plot was broader, indicating the lack of certainty (Figure 11). A total of 16 (73%) out of 22 uncertainty plots displayed a wider spread with mixed color bars (Figure 10B). This ambiguity might

XSL•FO
RenderX

unintentionally cause the human participants to doubt the capability of the AI aid with the uncertainty plot, resulting in a decrease in their trust, as evidenced by P18's statement: "When the checkboxes were not all green and the histogram had a bunch of colors (variability), I was even less trusting of the AI tool." This perspective is also supported by P25, who noted: "When the uncertainty plot had some red, it made me double check and decreased my confidence for sure."

Figure 11. Mean IQR by outcome pattern. The error bars represent a 95% CI. AI: artificial intelligence.



Finally, when AI rejected the correctly filled bottle, trust decrements occurred, with no significant differences between the 2 AI aids. If this circumstance happened in the real world, the pharmacists would reinspect the filled prescription. It will likely lead to an increased workload, fatigue, and stress, which could potentially lead to a lower quality of work and a higher frequency of errors [70,71]. However, P21 offered a contrasting viewpoint that more flagging would be better than AI approving the incorrect, highlighting: "I didn't lose as much trust when AI rejected the correct drug. I feel like AI should be there as a cautionary tool."

The observed trends in the moment-to-moment dynamics of trust indicate a greater degree of trust adjustment when participants were assisted by the AI with the uncertainty plot. This observation is further confirmed by the autocorrelation analysis. Specifically, the trust autocorrelation plot (Figure 9) reveals a lower autocorrelation between trust ratings when the uncertainty information was presented. This suggests that current trust levels were less influenced by past trust levels, implying more significant changes in trust from moment to moment. Pharmacists relied less on previous trials and more on the information presented in the present trial, highlighting the advantages of a more transparent display.

In addition, for both AI aids, participants displayed a larger trust decrement due to incorrect automation predictions. Even though these observations may seem alarming initially, they align with the prior literature addressing negativity bias. The study suggests that failure in automation typically has a more significant negative impact on trust than a positive impact from successful automation [56,65].

Limitations and Future Directions

We acknowledge several limitations of this study and propose directions for future research. First, as a pioneering investigation in this domain, we did not strictly control the interquartile range of the uncertainty plot (Figure 11). Future investigation should systematically examine the effects of presenting different distributions within the uncertainty plot on pharmacists' trust. Exploring similar variations in the IQR among different outcome patterns could provide a deeper understanding of how outcome patterns could impact trust while avoiding potential confounding factors.

Second, the uncertainty plot used in this study displayed the distribution of only 1 (2%) out of 50 probabilities for the predicted NDC. Future research should consider incorporating additional contextual information to enhance the interpretability of these distributions. A notable challenge identified was that users unfamiliar with statistical representation found the uncertainty plot difficult to understand. P11 suggested simplifying the uncertainty presentation because "it would be a little bit too much for some people not as comfortable with statistics or technology." Nonetheless, there is potential for pharmacists to become more comfortable with these plots with prolonged use and training, as evidenced by P30's remark: "I started skeptical because it's something I'm not familiar with, but as I got more examples of it, my trust built up quickly." Consequently, researchers should continue to develop alternative visualization techniques that provide a more comprehensive and intuitive representation of the AI's uncertainty, while maintaining a low complexity to accommodate users with varying degrees of statistical proficiency.

XSL•FO
RenderX

Third, this exploratory study was limited by a small sample size, which may have impacted the statistical power of the analyses on trust propensity, end trust, and average trust. This limitation likely contributed to the lack of significance observed in some results. However, despite being underpowered, our analysis still revealed a significant difference in end trust, indicating a large effect size. Regarding trust change, to maintain the AI's perceived usefulness while still providing enough examples of both misses and FAs, we incorporated only 2 trials of the AI approving the incorrect drug (miss) and 16 trials of the AI rejecting the correct drug (FA), setting the accuracy at 82%. The statistical power to detect significant differences may have been compromised by the small sample sizes. Future research should aim to include larger participant pools to enhance the generalizability and robustness of the findings.

Finally, this study only focused on pharmacists' trust and trust change and did not include the analysis of accuracy and reaction time. Even though focusing on trust alone is an accepted practice [53], a more comprehensive analysis linking performance with trust would likely reveal the relationship between performance and trust calibration.

Conclusions

Dispensing errors are significant contributors to adverse drug events, which lead to considerable health care expenses and harm to patients. Despite progress made in developing automated technologies to aid pill verification, pharmacists' trust in these systems has not been thoroughly investigated. Our research represents an initial exploration into pharmacists' trust in automated pill verification technology, marking a significant step in understanding the integration of such systems into health care settings.

Our findings reveal that pharmacists have a favorable disposition toward trusting automation, which can likely be attributed to their frequent use of automated technologies in their daily work. Moreover, providing uncertainty information about the AI's recommendation significantly boosts pharmacists' trust in the AI aid, highlighting the importance of transparency in AI development. The dynamics of trust vary depending on the AI's performance. Pharmacists using the AI aid with the uncertainty plot had a significantly larger trust increment when the AI approved the correct drug and a significantly larger trust decrement when the AI approved the incorrect drug. Intriguingly, the absence of the uncertainty plot led to an increase in trust when the AI correctly rejected an incorrect drug, whereas the presence of the plot resulted in a decrease in trust under the same circumstances. In addition, a pronounced "negativity bias" was observed, where the degree of trust reduction when the AI made an error exceeded the trust gain when the AI made a correct decision.

Acknowledgments

This research study was supported by the National Library of Medicine of the National Institutes of Health (R01LM013624). The contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

Data Availability

The deidentified data produced in this study are available upon reasonable request through a data use agreement with the University of Michigan.

Authors' Contributions

The authors confirm their contribution to the article as follows: study conception and design: JYK conducted subject testing, conducted statistical analysis, interpreted the results, and wrote the manuscript, VDM conducted statistical analysis and interpreted the results, BR conducted subject testing, QC interpreted the results, provided data for Tables 1, 2, and 3, YZ conducted subject testing, JDL interpreted the results and provided suggestion for Figure 9, RAK interpreted the results, CL interpreted the results, and XJY conducted statistical analysis, interpreted the results, and wrote the manuscript; All authors designed and conceptualized the study, reviewed the results, and approved the final version of the manuscript.

Conflicts of Interest

None declared.

References

1. Reiner G, Pierce SL, Flynn J. Wrong drug and wrong dose dispensing errors identified in pharmacist professional liability claims. J Am Pharm Assoc 2020;60(5):e50-e56. [doi: 10.1016/j.japh.2020.02.027] [Medline: 32217085]
2. Chui MA, Look KA, Mott DA. The association of subjective workload dimensions on quality of care and pharmacist quality of work life. Res Social Adm Pharm 2014;10(2):328-340 [FREE Full text] [doi: 10.1016/j.sapharm.2013.05.007] [Medline: 23791360]
3. Chui MA, Mott DA. Community pharmacists' subjective workload and perceived task performance: a human factors approach. J Am Pharm Assoc 2012;52(6):e153-e160 [FREE Full text] [doi: 10.1331/JAPhA.2012.11135] [Medline: 23229977]

XSL•FO
RenderX

4. Balchaites A, McCarthy S, Fleming A. Exploring the human factors of medication errors in community pharmacy: a mixed methods study. Int J Pharm Pract 2022:i39-i40. [doi: 10.1093/ijpp/riac019.055]
5. Larios Delgado N, Usuyama N, Hall AK, Hazen RJ, Ma M, Sahu S, et al. Fast and accurate medication identification. NPJ Digit Med 2019;2(1):10 [FREE Full text] [doi: 10.1038/s41746-019-0086-0] [Medline: 31304359]
6. Bates D. Using information technology to reduce rates of medication errors in hospitals. BMJ 2000;320(7237):788-791 [FREE Full text] [doi: 10.1136/bmj.320.7237.788] [Medline: 10720369]
7. Poon EG, Cina JL, Churchill W, Patel N, Featherstone E, Rothschild JM, et al. Medication dispensing errors and potential adverse drug events before and after implementing bar code technology in the pharmacy. Ann Intern Med 2006;145(6):426-434. [doi: 10.7326/0003-4819-145-6-200609190-00006] [Medline: 16983130]
8. Poon EG, Keohane CA, Yoon CS, Ditmore M, Bane A, Levtzion-Korach O, et al. Effect of bar-code technology on the safety of medication administration. N Engl J Med 2010;362(18):1698-1707. [doi: 10.1056/NEJMsa0907115] [Medline: 20445181]
9. Paoletti R, Suess T, Lesko M, Feroli AA, Kennel JA, Mahler JM, et al. Using bar-code technology and medication observation methodology for safer medication administration. Am J Health Syst Pharm 2007;64(5):536-543. [doi: 10.2146/ajhp060140] [Medline: 17322168]
10. Young D. Veterans Affairs bar-code-scanning system reduces medication errors. Am J Health Syst Pharm 2002;59(7):591-592. [doi: 10.1093/ajhp/59.7.591] [Medline: 11944599]
11. Debono DS, Greenfield D, Travaglia JF, Long JC, Black D, Johnson J, et al. Nurses' workarounds in acute healthcare settings: a scoping review. BMC Health Serv Res 2013;13:175 [FREE Full text] [doi: 10.1186/1472-6963-13-175] [Medline: 23663305]
12. Koppel R, Wetterneck T, Telles JL, Karsh B. Workarounds to barcode medication administration systems: their occurrences, causes, and threats to patient safety. J Am Med Inform Assoc 2008;15(4):408-423 [FREE Full text] [doi: 10.1197/jamia.M2616] [Medline: 18436903]
13. Oldland AR, Golightly LK, May SK, Barber GR, Stolpman NM. Electronic inventory systems and barcode technology: impact on pharmacy technical accuracy and error liability. Hosp Pharm 2015;50(1):34-41 [FREE Full text] [doi: 10.1310/hpj5001-34] [Medline: 25684799]
14. Buchanan TL, Barker KN, Gibson JT, Jiang BC, Pearson RE. Illumination and errors in dispensing. Am J Hosp Pharm 1991;48(10):2137-2145. [Medline: 1781468]
15. Gidman W. Increasing community pharmacy workloads in England: causes and consequences. Int J Clin Pharm 2011;33(3):512-520. [doi: 10.1007/s11096-011-9498-x] [Medline: 21424615]
16. Lea VM, Corlett SA, Rodgers RM. Workload and its impact on community pharmacists' job satisfaction and stress: a review of the literature. Int J Pharm Pract 2012;20(4):259-271. [doi: 10.1111/j.2042-7174.2012.00192.x] [Medline: 22775522]
17. Peterson GM, Wu MSH, Bergin JK. Pharmacist's attitudes towards dispensing errors: their causes and prevention. J Clin Pharm Ther 1999;24(1):57-71. [doi: 10.1046/j.1365-2710.1999.00199.x] [Medline: 10319909]
18. Caban JJ, Rosebrock A, Yoo TS. Automatic identification of prescription drugs using shape distribution models. 2012 Presented at: 19th IEEE International Conference on Image Processing; October 3, 2012; Orlando, FL p. 1005-1008.
19. Hamilton RN. Pharmacy pill counting vision system. Google Patents. 2004. URL: https://patents.google.com/patent/US6574580B2/en [accessed 2025-01-10]
20. Popovich J, Jordan ML. Automated drug discrimination during dispensing. US Patent. 2011. URL: https://scholar.google.com/citations?view_op=view_citation&hl=en&user=YAZqaxkAAAAJ&citation_for_view=YAZqaxkAAAAJ:WF5omc3nYNoC [accessed 2025-01-13]
21. Wootton JR, Reznack VV, Hobson G. Pharmaceutical pill recognition and verification system. Google Patents. 2003. URL: https://patents.google.com/patent/US6535637B1/en [accessed 2025-01-10]
22. Yu J, Chen Z, Kamata S. Pill recognition using imprint information by two-step sampling distance sets. 2014 Presented at: 22nd International Conference on Pattern Recognition; August 24-28, 2014; Stockholm, Sweden p. 3156-3161.
23. Yu J, Chen Z, Kamata S, Yang J. Accurate system for automatic pill recognition using imprint information. IET Image Processing 2015;9(12):1039-1047. [doi: 10.1049/iet-ipr.2014.1007]
24. Lester CA, Li J, Ding Y, Rowell B, Yang, Kontar RA. Performance evaluation of a prescription medication image classification model: an observational cohort. NPJ Digit Med 2021;4(1):118 [FREE Full text] [doi: 10.1038/s41746-021-00483-8] [Medline: 34315995]
25. Wong YF, Ng HT, Leung KY, Chan KY, Chan SY, Loy CC. Development of fine-grained pill identification algorithm using deep convolutional network. J Biomed Inform 2017;74:130-136 [FREE Full text] [doi: 10.1016/j.jbi.2017.09.005] [Medline: 28923366]
26. Lee JD, See KA. Trust in automation: designing for appropriate reliance. Hum Factors 2004;46(1):50-80. [doi: 10.1518/hfes.46.1.50_30392] [Medline: 15151155]
27. Hoff KA, Bashir M. Trust in automation: integrating empirical evidence on factors that influence trust. Hum Factors 2015;57(3):407-434. [doi: 10.1177/0018720814547570] [Medline: 25875432]

XSL•FO
RenderX

28. Schaefer KE, Chen JYC, Szalma JL, Hancock PA. A meta-analysis of factors influencing the development of trust in automation: implications for understanding autonomy in future systems. Hum Factors 2016;58(3):377-400. [doi: 10.1177/0018720816634228] [Medline: 27005902]
29. Azevedo-Sa H, Jayaraman SK, Esterwood CT, Yang XJ, Robert LP, Tilbury DM. Real-time estimation of drivers' trust in automated driving systems. Int J Soc Robot 2020;13(8):1911-1927. [doi: 10.1007/s12369-020-00694-1]
30. Lee JD, Kolodge K. Exploring trust in self-driving vehicles through text analysis. Hum Factors 2020;62(2):260-277. [doi: 10.1177/0018720819872672] [Medline: 31502885]
31. Ayoub J, Avetisian L, Yang XJ, Zhou F. Real-time trust prediction in conditionally automated driving using physiological measures. IEEE Trans Intell Transport Syst 2023;24(12):14642-14650. [doi: 10.1109/tits.2023.3295783]
32. Gombolay M, Yang XJ, Hayes B, Seo N, Liu Z, Wadhwania S, et al. Robotic assistance in the coordination of patient care. Int J Robot Res 2018;37(10):1300-1316. [doi: 10.1177/0278364918778344]
33. Wang EH, Gross CP, Tilburt JC, Yu JB, Nguyen PL, Smaldone MC, et al. Shared decision making and use of decision AIDS for localized prostate cancer : perceptions from radiation oncologists and urologists. JAMA Intern Med 2015;175(5):792-799. [doi: 10.1001/jamainternmed.2015.63] [Medline: 25751604]
34. Nazaretsky T, Ariely M, Cukurova M, Alexandron G. Teachers' trust in AI‐powered educational technology and a professional development program to improve it. Brit J Educational Tech 2022;53(4):914-931. [doi: 10.1111/bjet.13232]
35. Wang L, Jamieson GA, Hollands JG. Trust and reliance on an automated combat identification system. Hum Factors 2009;51(3):281-291. [doi: 10.1177/0018720809338842] [Medline: 19750792]
36. Du N, Huang KY, Yang XJ. Not all information is equal: effects of disclosing different types of likelihood information on trust, compliance and reliance, and task performance in human-automation teaming. Hum Factors 2020;62(6):987-1001. [doi: 10.1177/0018720819862916] [Medline: 31348863]
37. Chatzimparmpas A, Martins RM, Jusufi I, Kucher K, Rossi F, Kerren A. The state of the art in enhancing trust in machine learning models with the use of visualizations. Comput Graph Forum 2020;39(3):713-756. [doi: 10.1111/cgf.14034]
38. Rudin C, Ustun B. Optimized scoring systems: toward trust in machine learning for healthcare and criminal justice. Interfaces 2018;48(5):449-466. [doi: 10.1287/inte.2018.0957]
39. Dasgupta A, Lee J, Wilson R, Lafrance RA, Cramer N, Cook K, et al. Familiarity vs trust: a comparative study of domain scientists' trust in visual analytics and conventional analysis methods. IEEE Trans Vis Comput Graph 2017;23(1):271-280. [doi: 10.1109/TVCG.2016.2598544] [Medline: 27608465]
40. Yang F, Huang Z, Scholtz J, Arendt DL. How do visual explanations foster end users' appropriate trust in machine learning? 2020 Presented at: IUI '20: Proceedings of the 25th International Conference on Intelligent User Interfaces; March 17, 2020; Cagliari, Italy p. 189-201.
41. Sacha D, Sedlmair M, Zhang L, Lee JA, Peltonen J, Weiskopf D, et al. Human-centered machine learning through interactive visualization. 2016 Presented at: ESANN 2016 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning; April 27-29, 2016; Bruges, Belgium. [doi: 10.1016/j.neucom.2017.01.105]
42. Fisher D, Popov I, Drucker S, schraefel MC. Trust me, i'm partially right: incremental visualization lets analysts explore large datasets faster. 2012 Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; May 5, 2012; Austin, TX p. 1673-1682.
43. Mercado JE, Rupp MA, Chen JYC, Barnes MJ, Barber D, Procci K. Intelligent agent transparency in human-agent teaming for multi-UxV management. Hum Factors 2016;58(3):401-415. [doi: 10.1177/0018720815621206] [Medline: 26867556]
44. Yang XJ, Schemanske C, Searle C. Toward quantifying trust dynamics: how people adjust their trust after moment-to-moment interaction with automation. Hum Factors 2023;65(5):862-878 [FREE Full text] [doi: 10.1177/00187208211034716] [Medline: 34459266]
45. Lee J, Moray N. Trust, control strategies and allocation of function in human-machine systems. Ergonomics 1992;35(10):1243-1270. [doi: 10.1080/00140139208967392] [Medline: 1516577]
46. Swets JA. The science of choosing the right decision threshold in high-stakes diagnostics. Am Psychol 1992;47(4):522-532. [doi: 10.1037//0003-066x.47.4.522] [Medline: 1595983]
47. Davenport RB, Bustamante EA. Effects of false-alarm vs. miss-prone automation and likelihood alarm technology on trust, reliance, and compliance in a miss-prone task. Proc Hum Factors Ergon Soc Annu Meet 2010;54(19):1513-1517. [doi: 10.1177/154193121005401933]
48. Stanton NS, Ragsdale SA, Bustamante EA. The effects of system technology and probability type on trust, compliance, and reliance. HFES Ann Conf Proc 2009;53(18):1368-1372. [doi: 10.1518/107118109x12524443347238]
49. Chen Q, Al Kontar R, Nouiehed M, Yang XJ, Lester C. Rethinking cost-sensitive classification in deep learning via adversarial data augmentation. INFORMS J Data Sci 2022. [doi: 10.1287/ijds.2022.0033]
50. Zheng Y, Rowell B, Chen Q, Kim JY, Kontar RA, Yang XJ, et al. Designing human-centered AI to prevent medication dispensing errors: focus group study with pharmacists. JMIR Form Res 2023;7:e51921 [FREE Full text] [doi: 10.2196/51921] [Medline: 38145475]
51. Gal Y, Ghahramani Z. Dropout as a Bayesian approximation: representing model uncertainty in deep learning. Proc Mach Learn Res 2016:1050-1059.

XSL•FO
RenderX

52. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. 2016 Presented at: IEEE Conference on Computer Vision and Pattern Recognition (CVPR); June 27-30, 2016; Las Vegas, NV p. 770-778.
53. Tanner WP, Swets JA. A decision-making theory of visual detection. Psychol Rev 1954;61(6):401-409. [doi: 10.1037/h0058700]
54. Reilly S, Grasha AF, Matthews G, Schafer J. Automatic-controlled information processing and error detection in a simulated pharmacy-verification task. Percept Mot Skills 2003;97(1):151-174. [doi: 10.2466/pms.2003.97.1.151] [Medline: 14604036]
55. Wickens CD, Dixon SR. The benefits of imperfect diagnostic automation: a synthesis of the literature. Theor Issues Ergon Sci 2007;8(3):201-212. [doi: 10.1080/14639220500370105]
56. Manzey D, Reichenbach J, Onnasch L. Human performance consequences of automated decision aids. J Cogn Eng Decis Mak 2012;6(1):57-87. [doi: 10.1177/1555343411433844]
57. Yang XJ, Unhelkar VV, Li K, Shah JA. Evaluating effects of user experience and system transparency on trust in automation. 2017 Presented at: Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction; March 6, 2017; Vienna, Austria p. 408-416.
58. Merritt SM, Heimbaugh H, LaChapell J, Lee D. I trust it, but I don't know why: effects of implicit attitudes toward automation on trust in an automated system. Hum Factors 2013;55(3):520-534. [doi: 10.1177/0018720812465081] [Medline: 23829027]
59. Merritt SM, Ilgen DR. Not all trust is created equal: dispositional and history-based trust in human-automation interactions. Hum Factors 2008;50(2):194-210. [doi: 10.1518/001872008X288574] [Medline: 18516832]
60. McDermott P, Dominguez C, Kasdaglis N, Ryan M, Trhan I, Nelson A. Human-machine teaming systems engineering guide. Defense Technical Information Center. 2018. URL: https://apps.dtic.mil/sti/citations/AD1108020 [accessed 2025-01-10]
61. Bates DM, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Software 2015;67(1):1-48. [doi: 10.18637/jss.v067.i01]
62. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. 2019. URL: https://www.R-project.org [accessed 2025-01-20]
63. Montag C, Kraus J, Baumann M, Rozgonjuk D. The propensity to trust in (automated) technology mediates the links between technology self-efficacy and fear and acceptance of artificial intelligence. Comput Human Behav 2023;11:100315. [doi: 10.1016/j.chbr.2023.100315]
64. Miller L, Kraus J, Babel F, Baumann M. More than a feeling-interrelation of trust layers in human-robot interaction and the role of user dispositions and state anxiety. Front Psychol 2021;12:592711 [FREE Full text] [doi: 10.3389/fpsyg.2021.592711] [Medline: 33912098]
65. Yang XJ, Guo Y, Schemanske C. From trust to trust dynamics: combining empirical and computational approaches to model and predict trust dynamics in human-autonomy interaction. In: Duffy VG, Landry SJ, Lee JD, Stanton N, editors. Human-Automation Interaction: Transportation. Cham, Switzerland: Springer; 2022:253-265.
66. de Visser EJ, Pak R, Shaw TH. From 'automation' to 'autonomy': the importance of trust repair in human-machine interaction. Ergonomics 2018;61(10):1409-1427. [doi: 10.1080/00140139.2018.1457725] [Medline: 29578376]
67. Lyons JB, Sadler GG, Koltai K, Battiste H, Ho NT, Hoffmann LC, et al. Shaping Trust Through Transparent Design: Theoretical and Experimental Guidelines. Cham, Switzerland: Springer; 2017:127-136.
68. Luo R, Du N, Yang XJ. Evaluating effects of enhanced autonomy transparency on trust, dependence, and human-autonomy team performance over time. Int J Hum Comput 2022;38(18-20):1962-1971. [doi: 10.1080/10447318.2022.2097602]
69. Sorkin RD, Kantowitz BH, Kantowitz SC. Likelihood alarm displays. Hum Factors 1988;30(4):445-459. [doi: 10.1177/001872088803000]
70. Carayon P, Gurses AP. Nursing workload and patient safety—a human factors engineering perspective. In: Patient Safety and Quality: An Evidence-Based Handbook for Nurses. Rockville, MD: Agency for Healthcare Research and Quality; 2008.
71. Watterson TL, Look KA, Steege L, Chui MA. Operationalizing occupational fatigue in pharmacists: an exploratory factor analysis. Res Social Adm Pharm 2021;17(7):1282-1287. [doi: 10.1016/j.sapharm.2020.09.012] [Medline: 33004303]

Abbreviations

AI: artificial intelligence
CR: correct rejection
FA: false alarm
NDC: National Drug Code
SDT: signal detection theory

XSL•FO
RenderX

Edited by A Kushniruk, E Borycki; submitted 06.05.24; peer-reviewed by J Waterson, HF Neyedli; comments to author 13.09.24; revised version received 08.11.24; accepted 22.12.24; published 11.02.25.

Please cite as:
Kim JY, Marshall VD, Rowell B, Chen Q, Zheng Y, Lee JD, Kontar RA, Lester C, Yang XJ
The Effects of Presenting AI Uncertainty Information on Pharmacists' Trust in Automated Pill Recognition Technology: Exploratory Mixed Subjects Study
JMIR Hum Factors 2025;12:e60273
URL: https://humanfactors.jmir.org/2025/1/e60273
doi:10.2196/60273
PMID:39932773

©Jin Yong Kim, Vincent D Marshall, Brigid Rowell, Qiyuan Chen, Yifan Zheng, John D Lee, Raed Al Kontar, Corey Lester, Xi Jessie Yang. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 11.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Longitudinal and Combined Smartwatch and Ecological Momentary Assessment in Racially Diverse Older Adults: Feasibility, Adherence, and Acceptability Study

Sophia Holmqvist[1], MA; Marina Kaplan[1], BA; Riya Chaturvedi[1]; Haochang Shou[2], PhD; Tania Giovannetti[1], PhD

[1]Department of Psychology and Neuroscience, Temple University, Philadelphia, PA, United States

[2]Department of Biostatistics, Epidemiology, and Informatics, University of Pennsylvania School of Medicine, Philadelphia, PA, United States

Corresponding Author:
Sophia Holmqvist, MA
Department of Psychology and Neuroscience
Temple University
1701 N. 13th St
Philadelphia, PA, 19122
United States
Phone: 1 215 204 7000
Email: sophia.holmqvist@temple.edu

Abstract

Background: Due to the rising prevalence of Alzheimer disease and related dementias, easily deployable tools to quantify risk are needed. Smartphones and smartwatches enable unobtrusive and continuous monitoring, but there is limited information regarding the feasibility, adherence, and acceptability of digital data collection among racially diverse older adults.

Objective: This paper examined the feasibility, adherence, and acceptability of a 4-week combined smartwatch monitoring and ecological momentary assessment (EMA) study in a racially diverse sample of older adults.

Methods: A total of 44 older adults (aged ≥55 y) with either mild cognitive impairment or healthy cognition completed an informed consent comprehension quiz, baseline cognitive testing, training regarding digital data collection, and questionnaires. Participants were instructed to wear a Garmin Vivosmart 4 smartwatch for 23 h/d for 4 weeks, sync 2 smartphone apps (Garmin and Labfront) daily, and complete a daily EMA survey with automated prompts for surveys and charging. Training time, smartwatch adherence (eg, wear time), daily EMA survey response rate, and performance on the consent quiz were quantified. Associations between feasibility and adherence metrics and participant factors were evaluated. Self-reported usability of the apps and smartwatch was collected at study end.

Results: Consent comprehension quiz scores were high (mean 97.33%, SD 6.86% correct), and training sessions lasted on average 17.93 (SD 6.89) minutes. During the 4-week study, participants wore the smartwatch for an average of 21 h/d (SD 1.53) and showed an average response rate of 94% (SD 9.58%) to daily EMA surveys. In unadjusted bivariate analyses, age, race, and cognition were associated with feasibility and adherence measures, but only age and race remained significant in multivariate models. After accounting for all participant factors, older age was a significant predictor of longer training time, and Black race was a significant predictor of lower daily wear time. On the usability survey, all participants (45/45, 100%) indicated willingness to participate in future smartwatch studies, >80% (37/45) had a positive experience, and >90% (41/45) were satisfied with smartphone app syncing.

Conclusions: Smartwatch monitoring, requiring daily wear, smartphone syncing, and daily EMA survey completion, is highly feasible in older adults because adherence to daily wear and EMA surveys was high, as was general satisfaction on usability surveys. Although older participants may require more training on smartwatch and smartphone procedures and automated prompting during the study period, longitudinal monitoring with the Garmin Vivosmart 4 smartwatch and Labfront app is acceptable and feasible for collecting nearly continuous data in Black and White older adults, including those with mild cognitive impairment and those without.

(JMIR Hum Factors 2025;12:e69952) doi:10.2196/69952

XSL•FO
RenderX

KEYWORDS

cognitive impairment; smartwatch; longitudinal monitoring; ecological momentary assessment; aging

Introduction

Background

With the prevalence of Alzheimer disease (AD) and AD and related dementias (ADRD) increasing alongside the aging population and the availability of new treatments, there is a need to quantify risk and detect cognitive decline at the earliest stage [1]. It is estimated that only a small minority of older adults with mild cognitive decline are accurately identified early [2], missing an opportunity for early intervention. Missed or delayed diagnoses are even more common in Black and African American older adults, who are also at increased risk for AD/ADRD [3]. Personal digital devices such as smartwatches have been proposed as a potential tool for the early detection of AD/ADRD. Although recent data show that smartwatches have been increasingly adopted by older adults [4], the utility of personal digital tools for AD/ADRD monitoring is dependent upon consistent smartwatch use and syncing with associated software. This study examined the feasibility of and adherence to daily smartwatch wear and ecological momentary assessment (EMA) survey completion in a racially diverse sample of older adults enrolled in a 4-week study. We also investigated older adults' comprehension of study details that are relevant to informed consent. The participant characteristics associated with smartwatch adherence and the comprehension of consent were explored to optimize adherence and ethical research conduct in future studies.

Observable yet subtle cognitive, sensory, and motor changes precede and predict the clinical manifestations of AD/ADRD [1,5] and may be measured by sensors in commercially available smartwatches; for example, the Garmin smartwatch has been validated against conventional clinical measures of sleep [6], activity [7,8], and heart rate variability [7], which are associated with dementia risk [9-11]. The Garmin Vivosmart 4 smartwatch also has shown positive results in a 2-week feasibility study of stress in adults undergoing psychotherapy [12]. Thus, digital measures of stress, low heart rate variability, reduced physical activity, and sleep alterations have great potential as low-cost digital biomarkers for AD/ADRD that may be measured continuously, longitudinally, and passively, requiring relatively little effort.

EMA questionnaires delivered directly to participants via their personal digital devices enable high-frequency data collection in the participants' natural setting. EMA can be useful for contextualizing passive metrics [13]; for example, EMA data may facilitate the interpretation of atypical data as clinically significant or not (eg, elevated heart rate due to a cardiac event vs a day at an amusement park). EMA data also provide insight into participants' immediate perceptions and mood states without the confounds of recall bias [14]. There is potential for passive and EMA data collection via personal digital devices, but limited data exist on the feasibility and adherence of device use for health monitoring in older adults. In addition, the extent to which cognitive ability level influences adherence is not known. Adherence to daily EMA surveys and key tasks, such as wearing, charging, and syncing digital devices, is crucial for valid and reliable measurement and requires cognitive resources. This raises an important question: are those who could benefit the most from digital health monitoring more likely to show poor adherence?

Our review of the literature identified few studies on the feasibility and adherence of smartwatch use in digital health research. The results of the available studies suggest several important moderating factors, including the study procedures and participant factors. In-person studies and study designs that include interaction with a member of the study team show much higher retention and adherence than fully remote studies [15]. Adherence also varies depending on task demands, with different adherence rates for smartwatch wear versus syncing versus EMA completion. Adherence to wearable device use, which is relatively less demanding, versus the completion of daily EMA surveys has been estimated to be approximately >90% and 70%, respectively, in studies that are not fully remote [16-19].

Participant factors also play a role in retention and adherence. Although age is commonly considered a barrier to the adoption and acceptability of smartwatches due to low digital literacy [20], empirical data aggregated across several studies of smartwatch data show that older age was associated with greater retention [15]. A study requiring older adults with healthy cognition to wear a Fitbit device for 30 days during waking hours reported high rates of adherence to daily wear (89% of study days) and syncing (85% of study days) [21]. Other studies have demonstrated the effects of race, sex, and memory ability on adherence such that White participants [15], women [22], and people with better memory abilities [21] are more adherent when asked to engage in digital health research studies that involve interacting with a smartphone app.

In addition to the importance of learning about smartwatch adherence and feasibility, it is also crucial to understand older adults' comprehension of informed consent procedures for studies using novel technology. With the large amount of health data collected from a smartwatch and the need for long-term data storage, it is vital for older adult participants to fully understand data privacy and security limitations and the associated risks of study participation. Older adults also may be more wary of sharing personal health information in the context of digital health research. To assess older participants' understanding of informed consent procedures, interactive quizzes have been developed for use in fully remote studies [23]. For this study, which included face-to-face interactions with members of the study team, a 10-item comprehension quiz was modeled after a quiz used by Hackett et al [13] that was designed according to published guidelines [24,25]. The consent quiz focused on the novel technology used in the study (eg, where participants' data are stored and the right to request digital data to be deleted at any time). Hackett et al [13] reported high accuracy rates on a similar quiz in a sample of older adults and significant associations between quiz scores and education level.

XSL•FO
RenderX

Quiz scores also varied by race, with Black participants scoring lower than White participants.

Objectives

In this study, feasibility and adherence were evaluated over a 4-week monitoring period that required wearing the smartwatch daily (23 h/d), charging it, syncing apps, and completing EMA surveys. Feasibility and adherence were assessed by tracking participant retention and measuring the following: (1) time required to complete study training on syncing apps and completing EMA surveys, (2) the comprehension of informed consent information, (3) daily smartwatch wear time, (4) adherence to smartwatch daily wear, and (5) the completion of daily EMA surveys. We also investigated associations between feasibility and adherence measures and participant characteristics (demographics, cognition, and self-reported functional and cognitive decline). Finally, we explored participants' perceptions regarding the usability of the Garmin Vivosmart 4 smartwatch and the study-related apps (syncing process). To address these aims, we recruited a racially diverse sample of older adults, including those with mild cognitive impairment (MCI) and those without. It was hypothesized that older age, worse cognitive function, and greater self-reported functional decline would be associated with lower daily wear time, adherence to wear time, and survey completion, as well as lower comprehension of consent.

Methods

Participants

A total of 47 community-dwelling older adults aged ≥55 years classified as having healthy cognition or MCI were recruited from the Temple University Cognitive Neuropsychology Laboratory cohort in Philadelphia, Pennsylvania. This cohort includes >200 adults who have participated in ongoing federally funded studies since 2020 and represent the racial, ethnic, and economic diversity of the Philadelphia area. Participants met the following inclusion criteria: (1) oral and written fluency in English (to complete study questionnaires and measures) and (2) no history of large vessel stroke, Parkinson disease, major traumatic brain injury, seizures, schizophrenia, or significant neurological conditions other than dementia. The exclusion criteria included (1) current psychiatric disorder (eg, bipolar disorder or major depressive disorder), (2) intellectual disability, and (3) severe motor and sensory deficits precluding the use of a computer touchscreen. Participants from the laboratory cohort were contacted and screened for the following inclusion criteria for this study: (1) aged ≥55 years, (2) cognitive status classified as healthy or MCI, (3) current smartphone user (Android or IOS), and (4) not currently wearing a smartwatch or willing to wear only the study smartwatch during the study period. The exclusion criteria included (1) a diagnosis of dementia and (2) scheduled surgery or travel during the 4-week study period. Approximately 80 participants from the laboratory cohort were contacted as part of the recruitment efforts. The most common reasons for nonparticipation were not having a smartphone or a lack of interest in completing a month-long study.

Of the 47 participants recruited for this study, 1 (2%) was excluded because they required a walker for mobility, which precluded accurate step count recording, and 1 (2%) was excluded from the analyses of feasibility and adherence because their smartphone malfunctioned, and they required the monitoring period to be extended. Moreover, 1 (2%) of the 47 participants dropped out after 1 day, citing discomfort from the smartwatch band being too tight. Thus, of the 47 enrolled participants, 3 (6%) were excluded, resulting in a final sample of 44 (94%) participants for the feasibility and adherence analyses. The participant who extended their monitoring period because of the smartphone malfunction was retained for the usability analyses.

Ethical Considerations

This study was approved by the Temple University Institutional Review Board (Protocol Number: 29712). All participants provided written informed consent and received a US $50 gift card as compensation for their time. Data storage and procedures were Health Insurance Portability and Accountability Act–compliant and only approved study personnel had access to study data.

Procedures

Overview

Participation in the study involved 3 phases. First, participants completed an initial in-person study visit lasting 2 to 4 hours, which included comprehensive cognitive testing, training for the study, and questionnaire completion. Second, participants wore the smartwatch for 4 weeks, answered daily questions on their smartphone, and synced the smartwatch with a smartphone app once a day. Finally, all participants completed a debriefing session. More details of each study phase are provided in the following subsections. Textbox 1 shows the study timeline. The cognitive tests administered during study visit 1 are listed in Table 1.

XSL•FO
RenderX

Textbox 1. Study timeline.

Study visit 1

- Informed consent and comprehension quiz
- Installation of Garmin and Labfront apps
- Study training on syncing and daily survey
- Questionnaires and cognitive testing

4-wk monitoring period

- Participants engage in daily activities as usual
- Wear smartwatch for 23 h/d
- 9-question ecological momentary assessment survey once a day
- Sync Labfront and Garmin apps once a day to securely collect and transmit deidentified data

Study visit 2

- Uninstall apps and unpair smartwatch
- Debriefing questionnaire
- Payment of US $50

Table 1. Cognitive domains and corresponding neuropsychological measures.

Cognitive domains	Neuropsychological measures
Attention	• Trail Making Test part A [26] • Digit Span Forward [27]
Executive function	• Trail Making Test part B [26] • Digit Span Backward [27]
Language	• Letter fluency (S and P) and category fluency (animals) [28] • Boston Naming Test, 30-item version [29]
Processing speed	• Salthouse Letter Comparison Test [30] • Salthouse Pattern Comparison Test [30]
Episodic memory	• Hopkins Verbal Learning Test: immediate recall, delayed recall, and recognition trial [31] • Brief Visual Memory Test–Revised: immediate recall, delayed recall, and recognition trial [32]

Initial Study Visit

Participants completed the following during study visit 1, which lasted 2 to 4 hours: (1) informed consent and 10-item consent comprehension quiz on the specific features of the study (eg, data security and privacy; refer to Multimedia Appendix 1 for the quiz items); (2) gold standard cognitive tests (Table 1) and questionnaires; (3) installation and configuration of study apps and training on wearing a Garmin Vivosmart 4 smartwatch daily (23 h/d of wear; 1 h/d of charging time) for 4 weeks, syncing the smartwatch, and completion of a daily EMA survey. At the end of the first study session, a follow-up visit was scheduled for a date at least 4 weeks later. Further details of the procedures during the initial session are provided in the following subsections.

Smartwatch Data Collection

At the initial study visit, 2 apps (Garmin and Labfront) were downloaded onto the participant's smartphone to facilitate deidentified passive (raw sensor) and active data collection for the daily survey. The Garmin app facilitated data collection from the smartwatch sensors. The Labfront [33] app is a research platform that collects and organizes the Garmin smartwatch data and enabled the study team to monitor adherence and data collection in real time and remotely through a user-friendly study dashboard. The Labfront app was also used to deliver the daily EMA surveys.

A Garmin account was created for each participant using the Garmin Connect mobile app on their smartphone with deidentified information. The Labfront Companion app was then downloaded to the participant's smartphone. The Labfront app uses a randomized 6-digit ID to connect the participant's smartwatch to the app and to connect the Garmin and Labfront accounts. The Garmin smartwatch connects to the Garmin and Labfront apps via Bluetooth. Daily in-app Garmin and Labfront syncing required a Wi-Fi network or an LTE cellular data connection. Once the Garmin and Labfront apps were connected,

XSL•FO
RenderX

data were available for real-time download by the study team in CSV format. Daily syncing of the apps enabled the aggregation of minute-by-minute data.

Participants were asked to wear the Garmin Vivosmart 4 smartwatch for 23 h/d on their nondominant wrist for 4 weeks and to charge it for approximately 1 h/d. Alarms were set in participant smartphones twice daily at times of the participant's choosing to serve as reminders for the following: (1) charging the watch in the morning and (2) syncing the Garmin and Labfront apps and completing the daily survey on the Labfront app at night before going to sleep. To minimize the influence of external cues as well as behavior and health tracking information on participant behavior, all notifications were disabled on the study smartwatch, and additional features were removed from the Garmin app at the initial study visit.

Standardized training procedures with training criteria were used to teach participants how to charge the watch, sync the apps, and complete the daily survey. The training included demonstrations and a practice run during which participants were required to independently complete all daily study steps. Repetitions were recorded for practice runs, and the training session was timed. At the end of the training, participants were given a binder to take home and refer to during the study. The binder contained detailed instructions on how to charge the watch, sync the apps, and complete the daily survey during the study. The contents of this training binder are available in Multimedia Appendix 1.

4-Week Monitoring Period

Participants were told to go about their daily life during the 4-week monitoring period while wearing the smartwatch for 23 h/d and to charge it for 1 h/d. Sleep duration, beat-to-beat heart rate variability, and physical activity data such as step count were collected but are not reported in this paper. Participants were also asked to complete the daily EMA survey delivered to their smartphones through the Labfront app. The EMA survey included 9 questions (Multimedia Appendix 1). The purpose of the EMA survey was to guide the subsequent interpretation of smartwatch data (estimated time to complete the survey: 5 min).

During the course of the study, adherence was monitored through Labfront software by study personnel. If participants were nonadherent for >4 consecutive days, which was defined as wearing the watch for <16.67 h/d or not completing the daily surveys, they were contacted by the study team to determine the reason for nonadherence and to possibly reschedule the follow-up study session to extend the study period and obtain at least 28 days of data collection. The duration of the monitoring period (in days) was tallied for each participant. It is important to note that only 1 (2%) of the 47 enrolled participants required an extension of the monitoring period due to >4 consecutive days of missing data caused by a malfunction of their smartphone. Consequently, this participant was excluded from all feasibility and adherence analyses. For some of the participants (18/44, 41%), the monitoring extended beyond 4 weeks because the follow-up visit could not be scheduled on the study end date due to scheduling conflicts; in these cases, the follow-up visit was scheduled as soon as possible but no more than 1 week later.

Follow-Up Visit

After the study period, participants returned for a brief second visit (study visit 2) during which we collected data on their study experience, the usability of the apps and smartwatch, and barriers to future participation in wearable device research. At this visit, smartwatches were synced a final time to capture any aggregate-level data that had not been previously uploaded to the app. Subsequently, deidentified data obtained during the monitoring period were downloaded in CSV format for each participant from the Labfront researcher user interface for processing and analysis.

Measures of Participant Characteristics

Cognitive tests and questionnaires used in standard clinical evaluations for cognitive decline were administered during study visit 1.

Cognition and Clinical Classification

IQ was estimated with the Hopkins Reading Test [34]. The 5 cognitive domains and 10 cognitive tests that were administered are listed in Table 1. Scores from each of the 10 tests were standardized (T score) after adjusting for demographic variables (age, education, sex, and estimated IQ score [28]). Clinical classification (ie, healthy cognition vs MCI) was based on published criteria [35] that define MCI as T scores of ≤40 (<1 SD) on both measures from at least 1 cognitive domain. A total composite score and cognitive domain–specific composite scores were calculated by averaging demographically corrected T scores (refer to Table 1 for all neuropsychological measures). These neuropsychological tests are used widely in research and clinic settings, have been extensively validated, and show strong psychometric properties [28].

Self-Reported Cognitive and Functional Decline

The average score from the Everyday Cognition (ECog)–short form [36] was used to estimate self-reported functional decline across a range of abilities (eg, memory and language), with scores ranging from 1 (no decline over the past 10 y) to 4 (worse decline over the past 10 y). The optimal cut score for distinguishing between healthy cognition and MCI is 1.32. The original validation study estimated a Cronbach α value of 0.96 [36].

Smartwatch Feasibility and Adherence Measures

Training Time

The amount of time spent training (in minutes) during study visit 1 was recorded. Training procedures were standardized (Multimedia Appendix 1). Training began with a review of procedures for charging the watch and ended when the examiner observed that the participant independently demonstrated charging the watch, syncing the apps, and completing the daily survey.

Comprehension of Consent

Comprehension of consent was assessed using a 10-item quiz (Multimedia Appendix 1), which included questions on study-specific risks, data privacy, and security. The following is an example item: "The Garmin/Labfront apps collect the *content* of my texts and phone calls" (correct answer: no). To

XSL•FO
RenderX

ensure informed consent, if participants answered any item incorrectly, study personnel immediately provided feedback to clarify the correct response.

Smartwatch Adherence

The total number of days that participants did not wear the watch was tallied, and 2 measures were used to evaluate adherence to the instruction to wear the watch for 23 h/d. First, the average daily wear time (in hours) per participant across all possible days of wear was collected. Wear time was derived from minute-to-minute heart rate data; therefore, wear time reflects the accumulation of heart rate data from the watch. Weekly averages of wear time were also computed (weeks 1, 2, 3, and 4) to examine adherence over time.

On the basis of published recommendations [37], daily watch data were considered valid if the participant wore the smartwatch for at least 16.67 hours that day. Therefore, watch adherence also was assessed as the "percentage of valid smartwatch days" during the monitoring period, calculated as the number of days the participants wore the watch for ≥16.67 h/d divided by the total number of days they were in the study.

EMA Survey Adherence

Adherence to the daily survey was calculated as the number of daily surveys completed divided by the total number of days that each participant was in the study. This percentage ranged from 0% to 100%, with 100% indicating perfect adherence.

Usability Measures

A self-administered usability survey was completed by participants at the follow-up visit (study visit 2) to provide qualitative information on smartwatch use. The survey questions focused on the smartwatch and Labfront and Garmin smartphone apps, which were used for daily EMA surveys and syncing. Participants were also asked to report their likelihood of participating in another future study that included the smartwatch and daily surveys. This survey was administered to 45 (96%) of the 47 enrolled participants, including the participant whose monitoring period was extended.

Statistical Analyses

Descriptive statistics were used to characterize feasibility: average daily wear time, study training time, percentage correct on consent comprehension, the percentage of completed EMA surveys, the percentage of valid smartwatch days, and responses to the study debriefing questionnaire. To examine univariate associations between feasibility and adherence measures and participant characteristics, Spearman correlations and Mann-Whitney *U* tests were used for continuous and categorical variables, respectively, with Bonferroni correction applied to interpret statistical significance. Friedman rank tests for 1-way repeated measures of ANOVA were used to analyze whether wear time differed from week 1 to week 4, with Bonferroni correction applied to interpret statistical significance. Nonparametric statistical tests were used due to the skewed distributions of the variables of interest. Multiple linear regressions were conducted to examine multivariate associations between demographics (ie, age, sex, race, and education) and cognitive variables (ie, ECog, cognitive composite scores, and cognitive status) on all feasibility and adherence measures. Separate models were conducted for each cognitive variable to avoid multicollinearity. As the dependent variables were not normally distributed, we used bootstrapping (5000 replications) to estimate robust SEs and CIs for coefficients, leading to more reliable testing of significance.

All analyses were conducted using SPSS software (version 28.0; IBM Corp) or, for bootstrapped regressions only, R (version 2023.12.1+402; R Foundation for Statistical Computing).

Results

Participant Characteristics

A total of 44 older adults ranging in age from 55 to 83 years participated and were included in the feasibility and adherence analyses. On average, participants were college educated with estimated IQ scores in the high average range (Table 2). Self-report of cognitive and functional decline was well within the normal range (ECog; range 1-2.5). The cognitive composite T score ranged from 35.40 to 63.40. The sample comprised mostly women (26/44, 59%) and primarily individuals identifying as Black or African American (11/44, 25%) or White, non-Hispanic (31/44, 70%). Of the 44 participants, 8 (18%) met the Jak-Bondi criteria for MCI. The participants were relatively active, with an average of 5999.74 (SD 2872.38) steps per day. Of the 44 participants, 17 (39%) owned a smartwatch before the study, with most owning a Fitbit device (n=9, 50%) or an Apple Watch (n=9, 50%). Those who currently or previously owned a smartwatch reported the main motivations for use as activity monitoring, followed by improving fitness and improving health. Of those who had not owned a smartwatch, most (17/44, 39%) indicated that they would consider buying one after their study experience.

XSL•FO
RenderX

Table 2. Participant characteristics and smartwatch feasibility measures (n=44).

Variables	Values
Participant characteristics	
Age (y), mean (SD)	68.48 (7.22)
Education (y), mean (SD)	16.41 (2.08)
Sex: female, n (%)	26 (59)
Race and ethnicity, n (%)	
Asian	1 (2)
Black or African American	11 (25)
White, Hispanic	1 (2)
White, non-Hispanic	31 (70)
MCI[a], n (%)	8 (18)
Estimated IQ, mean (SD)	110.84 (7.59)
ECog[b] score, mean (SD)	1.41 (0.39)
Cognitive composite score, mean (SD)	51.69 (5.68)
Daily step count, mean (SD)	5999.74 (2872.38)
Smartwatch ownership before study, n (%)	17 (39)
"If you already currently own and use a smartwatch or have previously, what is or was your main motivation for using the smartwatch? Select all that apply," n (%)	
To monitor activities	15 (34)
Improve fitness	15 (34)
Improve health	12 (27)
Keep up with new technology	7 (16)
Improve appearance	2 (4)
"If you do not already own a smartwatch, would you consider buying one after this experience?" n (%)	
Yes	17 (39)
No	10 (24)
Already have smartwatch	17 (39)
Smartwatch feasibility and adherence measures, mean (SD)	
Percentage correct of consent comprehension	98.00 (6.86)
Average training time (min)	17.93 (6.89)
Average daily smartwatch wear time (h)	21.04 (1.53)
Percentage of valid smartwatch days	92.00 (10.39)
Percentage of completed daily surveys	94.00 (9.58)

[a]MCI: mild cognitive impairment.

[b]ECog: Everyday Cognition.

Retention Rate and Duration of Monitoring Period

A total of 47 participants were enrolled, of whom 1 (2%) dropped out after 1 day citing discomfort from the smartwatch band being too tight. All other participants (46/47, 98%) completed the study. However, 2 (4%) of these 46 participants were excluded from feasibility data analysis: 1 (50%) was excluded because their monitoring period was extended to 46 days due to a smartphone malfunction, and they did not complete the EMA survey for 1.5 weeks during the originally scheduled 28-day study; consequently, they were asked and agreed to extend their monitoring period. As a result, this participant was excluded from the analytic sample used for feasibility and adherence analyses. No other participants extended their monitoring period for reasons other than scheduling conflicts. Another participant was excluded due to mobility issues. Thus, the final analytic sample for the feasibility analyses comprised 44 participants.

The number of days monitored with the smartwatch ranged from 27 to 35. The mode for study days was 27 days,

XSL•FO
RenderX

corresponding to exactly 4 weeks because the date of the initial study visit was excluded from analyses. The average number of days in the study was 29 (SD 3.46) days. Feasibility and adherence measures were computed using all available monitoring days.

Feasibility and Adherence

Average feasibility and adherence measures are shown at the bottom of Table 2.

Comprehension Quiz

On average, participants obtained a score of 97% (range 70%-100%) correct on the consent comprehension quiz, approaching a perfect average score. The most frequently missed response was indicating "yes" to "Using the Garmin and Labfront apps will help improve my cognitive functioning."

Training Time

Participants took an average of 17.93 (range 12-53) minutes to complete the study training during study visit 1. Training time was mostly spent independently practicing study tasks, such as syncing apps.

Smartwatch Adherence

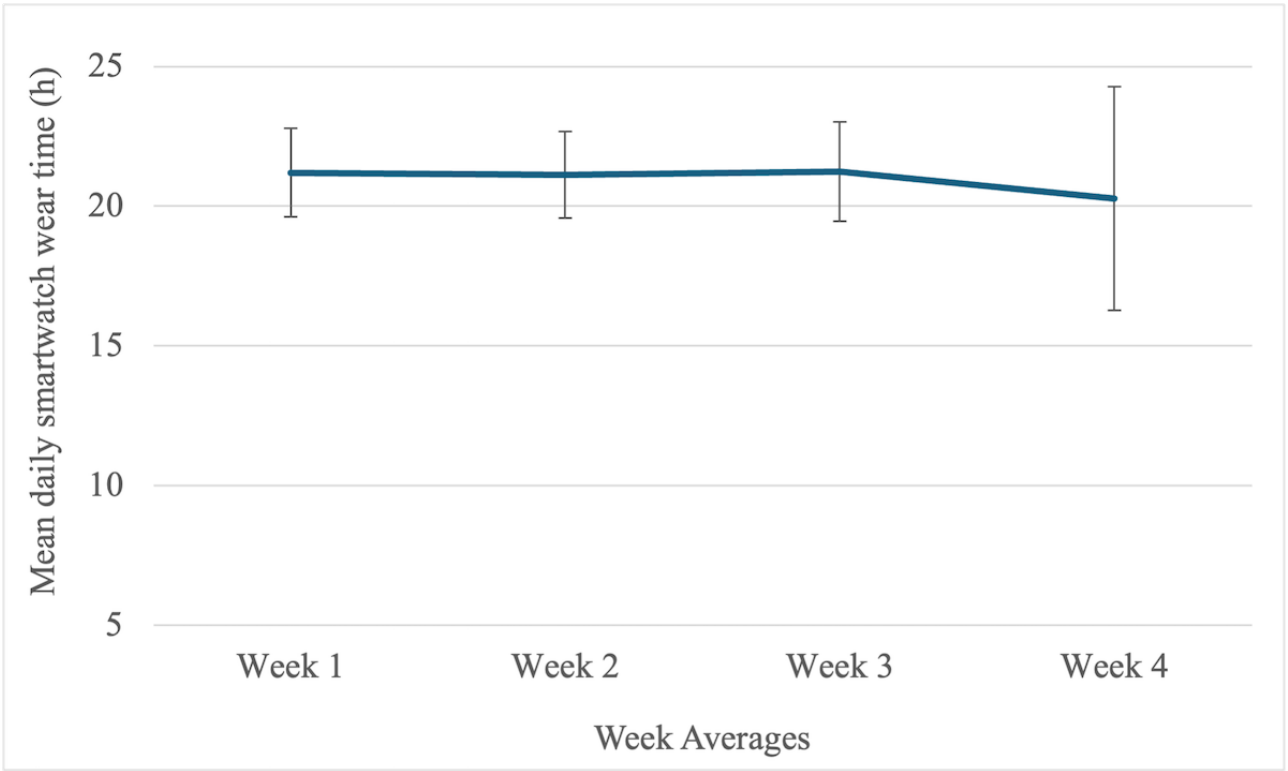
Of the 44 participants, 11 (25%) did not wear the watch for at least 1 entire day (n=9, 20% missed only 1 d; n=1, 2% missed 2 d; and n=1, 2% missed 4 d). All participants wore the watch for ≥16.67 h/d for ≥14 days during the monitoring period. Of the 44 participants, 5 (11%) had ≥27 days of wear time that exceeded 16.67 h/d. On average, participants exceeded 16.67 h/d of wear time on 92% (range 53%-100%) of total study days. As the monitoring period extended beyond 4 weeks for some participants due to scheduling conflicts, watch adherence measures also were calculated over 27 days, with the average percentage of days exceeding 16.67 h/d of wear time remaining unchanged at 92% (range 55%-100%).

On average, participants wore the watch for 21 (range 15.71-22.44) h/d across the 4-week study period. Of the 44 participants, 3 (7%) had an average wear time of <18 h/d, 2 (5%) had an average wear time of <16.67 h/d, and 39 (89%) wore the watch for ≥20 h/d on average.

The distribution of average watch wear time per week is shown in Figure 1. Pairwise 2-tailed *t* tests showed no significant difference in average wear time across the 4 weeks (refer to Multimedia Appendix 1 for full results).

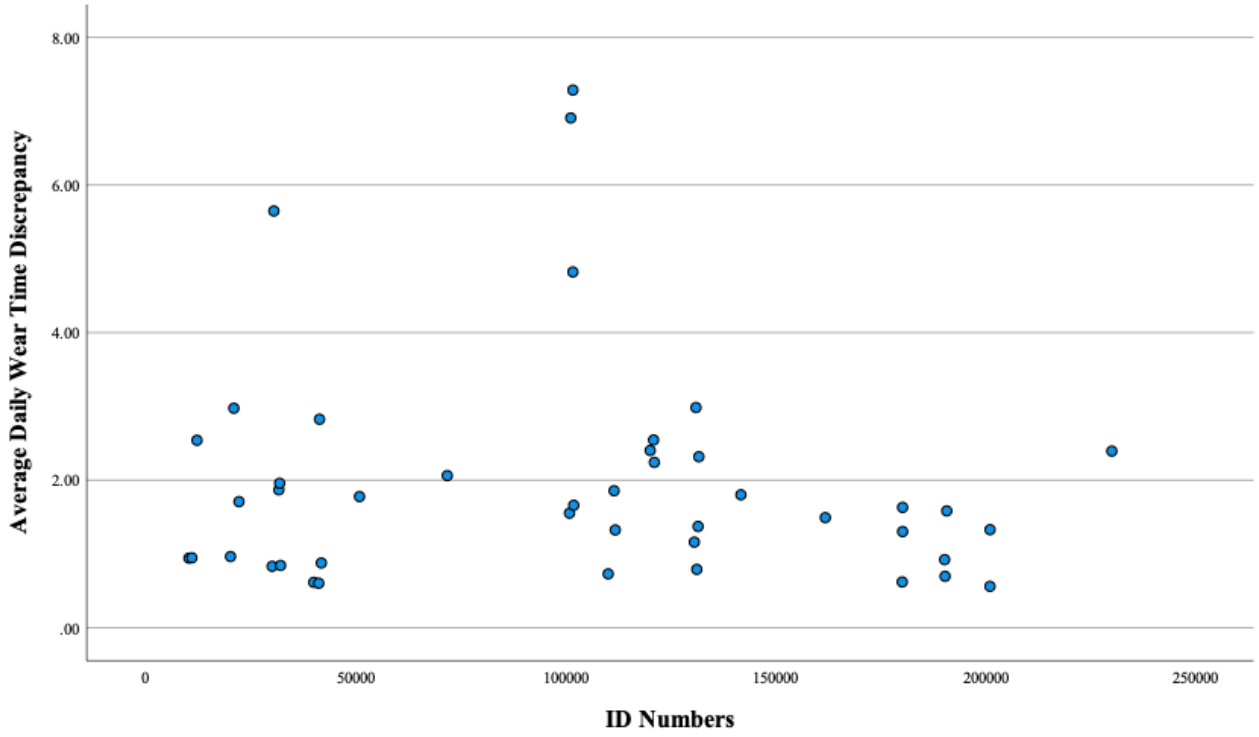
As shown in Figure 2, on average, 91% (40/44) of the participants wore the watch for at least 19 h/d, within 4 hours of the required 23 h/d. In fact, on average, participants on average wore the watch for 21 h/d, within 2 hours of the required 23 h/d.

Figure 1. Daily wear time across participants was averaged across weeks 1 to 4 in the study. The y-axis represents the mean daily smartwatch wear time in hours. The error bars represent –1 to +1 SD.



XSL•FO
RenderX

Figure 2. Average wear time discrepancy for the entire sample. The y-axis represents each participant's average daily wear time subtracted from the required 23 h/d. The x-axis shows participant ID numbers from the feasibility analyses. Each dot represents a participant.



EMA Survey Adherence

On average, participants completed surveys on most study days (94%; range 58%-100%). Of the 44 participants, 3 (7%) completed <70% of the surveys, while 41 (93%) completed >85% of the daily surveys.

Correlations Among Smartwatch Feasibility and Adherence Metrics

As shown in Table 3, a greater percentage of daily surveys completed was significantly associated with longer study training time, but the *P* value for this association did not survive Bonferroni correction (ie, *P*=.005 and below). Correlations between the 2 measures of smartwatch adherence and between smartwatch adherence measures and the percentage of daily surveys completed were significant and survived correction for multiple comparisons. The correlations indicated that participants who completed a higher percentage of daily surveys also wore the watch for longer durations and had more days of valid smartwatch wear time.

Table 3. Correlation coefficients among smartwatch feasibility and adherence metrics.

Variables	Consent comprehension percentage correct	Study training time	Percentage of daily surveys completed	Percentage of valid smartwatch wear days
Study training time				
r	0.055	—[a]	—	—
P value	.77	—	—	—
Percentage of daily surveys completed				
r	–0.016	0.298	—	—
P value	.92	.049	—	—
Percentage of valid smartwatch wear days				
r	0.230	0.018	0.544	—
P value	.13	.91	<.001	—
Daily wear time				
r	0.095	0.091	0.579	0.778
P value	.54	.56	<.001	<.001

[a]Not applicable.

XSL•FO
RenderX

Associations Between Participant Characteristics and Feasibility and Adherence Metrics

As shown in Table 4, only age and self-reported cognitive and functional decline (ECog) were significantly associated with feasibility and adherence metrics. Specifically, older participants required longer study training times, and participants who reported greater cognitive and functional decline (ECog) had lower daily average wear time. However, these significant correlations did not survive Bonferroni correction (P=.003). No other correlation coefficients were statistically significant (Table 4).

Table 4. Correlation coefficients between smartwatch feasibility metrics and demographic and clinical factors.

Variables	Consent comprehension percentage correct	Study training time	Percentage of daily surveys completed	Percentage of valid smartwatch wear days	Daily wear time
Age					
r	–0.164	0.403	0.165	–0.023	0.006
P value	.29	.007	.28	.88	.97
Education					
r	0.041	0.029	–0.197	–0.092	–0.108
P value	.79	.85	.20	.55	.48
Cognitive composite scores					
r	0.216	–0.002	0.037	0.102	0.155
P value	.16	.99	.81	.51	.31
ECog[a]					
r	–0.102	–0.206	–0.221	–0.261	–0.393
P value	.51	.19	.15	.09	.009

[a]ECog: Everyday Cognition.

There were no significant sex differences in consent comprehension, training time, watch adherence measures, or the percentage of daily surveys completed (data reported in Multimedia Appendix 1). Participants with MCI and those with healthy cognition differed on only 1 measure (data reported in Multimedia Appendix 1): participants with MCI (median 20.64, IQR 2.74 h) had significantly lower daily wear time than participants with healthy cognition (median 21.53, IQR 1.18 h; U=71.00; P=.03). This result is consistent with the significant correlation between ECog and daily wear time reported in Table 4, suggesting an association between lower cognitive abilities and less smartwatch wear time.

Between-group analyses comparing Black and White participants were not statistically significant, except for daily wear time (data reported in Multimedia Appendix 1). Black participants (median 20.76, IQR 0.91 h) had significantly lower daily smartwatch wear time than White participants (median 21.83, IQR 1.02 h; U=74.00; P=.004).

Multivariate associations between participant characteristics and feasibility and adherence metrics were investigated using multiple regression analyses with bootstrapping to estimate the significance of the coefficients. Separate regressions were run, including one for each cognitive variable (cognitive composite scores, ECog, and cognitive status; refer to Multimedia Appendix 1 for all tables), to avoid multicollinearity. The results showed that race remained a significant predictor of daily wear time, even after controlling for age, sex, education, and measures of cognition, with White participants having longer daily wear times than Black participants. Multivariate analyses also showed that age was significantly associated with training time even after controlling for sex, education, race, and measures of cognition, with older age associated with longer study training time. By contrast, associations between the measures of cognition and feasibility and adherence metrics were not statistically significant in regression analyses, including demographic variables (age, sex, education, and race).

Usability Survey

The results of the usability survey are shown in Table 5 (Garmin smartwatch) and Table 6 (Labfront smartphone app). Overall, participants expressed favorable views of both the smartwatch and the smartphone apps and especially enjoyed their involvement in research. When asked, “Would you wear this Garmin smartwatch without being asked to wear it as part of the study?” 55% (25/45) of the sample reported “yes.” However, all participants indicated they were likely (11/45, 24%) or very likely (34/45, 76%) to participate in a future study asking them to wear a Garmin smartwatch daily. Of the 45 participants, 36 (80%) either somewhat or strongly agreed that they had a positive experience using the Garmin smartwatch during the study period (Table 5). Most of the participants (31/45, 68%) reported that the smartwatch was comfortable or very comfortable. Most of the participants (32/45, 71%) were satisfied or very satisfied with the functioning of the watch. The majority of the complaints about the smartwatch were regarding low battery life (17/45, 38%), followed by discomfort (8/45, 18%) and technical issues (8/45, 18%). Of the 45 participants, 17 (38%) reported no complaints about the smartwatch. Participants suggested the following improvements for the watch: changing wristband material (18/45, 40%), increasing

XSL•FO
RenderX

the display size (15/45, 33%), and increasing the wristband size (10/45, 22%).

Table 5. Smartwatch usability and satisfaction (n=45).

Survey items	Participants, n (%)
"How comfortable was the smartwatch?"	
Very comfortable[a]	20 (44)
Comfortable	11 (24)
Somewhat comfortable	10 (22)
Not comfortable	4 (9)
"Did you wear it overnight every night?"	
Yes	44 (98)
"How satisfied were you with the functioning of the watch (i.e. you were able to tell date/time easily)?"	
Very satisfied	18 (40)
Satisfied	14 (31)
Somewhat satisfied	9 (20)
Not satisfied	4 (9)
"Do you have any complaints about using the Garmin Vivosmart 4 watch?"	
None	17 (38)
Technical issues	8 (18)
Doesn't fit	2 (4)
It's uncomfortable	8 (18)
Low battery life	17 (38)
Problems with the screen	6 (13)
"Overall, I had a positive experience using the smartwatch."	
Disagree strongly	0 (0)
Somewhat disagree	3 (7)
Neutral	5 (11)
Somewhat agree	18 (39)
Agree strongly	19 (43)
"What would you change to improve the comfort of wearing the Garmin smartwatch?"	
No changes needed	15 (33)
Improve wristband clasp function	9 (20)
Increase display size	15 (33)
Change material of wristband (cloth, material)	18 (40)
Other (increase size of wristband)	10 (22)
"How satisfied were you with charging the battery of the Garmin smartwatch?"	
Very satisfied	10 (22)
Satisfied	19 (44)
Somewhat satisfied	11 (24)
Not satisfied	4 (9)
"Did the watch ever run out of battery while you were wearing it?"	
Yes (those who answered yes, indicated between 1 and 5 times)	17 (38)

[a]Italicized text indicates the most frequent response.

XSL•FO
RenderX

App usability and overall experience are reported in Table 6. Of the 45 participants, 41 (91%) were satisfied or very satisfied with the syncing process of both apps (Garmin and Labfront). Of the 45 participants, 35 (78%) indicated that they agreed or strongly agreed that the Labfront app was easy to use and that they felt very confident using it. Importantly, 64% (29/45) of the participants had a positive or extremely positive experience with the Labfront app. Moreover, 71% (32/45) of the participants indicated that the Labfront app was easy to use or extremely easy to use (Table 6).

Table 6. App usability and overall experience (n=45).

Survey items	Participants, n (%)
"How easy was it to sync the watch every day with the Labfront and Garmin apps?"	
Very satisfied[a]	32 (71)
Satisfied	9 (20)
Somewhat satisfied	1 (2)
Not satisfied (issues included having to try a couple of times to sync the Labfront app)	3 (7)
"I thought Labfront was easy to use"	
Strongly disagree	3 (7)
Disagree	5 (11)
Neither agree nor disagree	2 (4)
Agree	18 (40)
Strongly agree	17 (37)
"I felt very confident using the Labfront app"	
Strongly disagree	2 (4)
Disagree	2 (4)
Neither agree nor disagree	4 (9)
Agree	18 (40)
Strongly agree	19 (42)
"I enjoy using the Labfront app"	
Strongly disagree	1 (2)
Disagree	2 (4)
Neither agree nor disagree	18 (40)
Agree	17 (38)
Strongly agree	7 (16)
"How would you rate your experience with the Labfront app?"	
Extremely negative	1 (2)
Negative	2 (4)
Neutral	13 (29)
Positive	23 (51)
Extremely positive	6 (13)
"How would you rate usability of the Labfront app?"	
Very difficult to use	0 (0)
Somewhat difficult to use	3 (7)
Neutral	10 (22)
Easy to use	23 (53)
Extremely easy to use	8 (18)

[a]Italicized text indicates the most frequent response.

XSL•FO
RenderX

Discussion

Principal Findings

Overall, the results demonstrated that our longitudinal study requiring daily wear of a smartwatch and daily EMA survey completion in a racially diverse sample of older adults was feasible with excellent adherence. Participants with cognitive impairment and those without wore a commercially available smartwatch during waking and sleep hours for 4 weeks for an average of 21 h/d and showed an average response rate of 94% to daily EMA surveys. Only 1 (2%) of the 47 enrolled participants refused to complete the study after informed consent due to discomfort with the watch fit. On the basis of the results of our consent comprehension quiz, participants had no difficulty understanding the study procedures and risks, and completed smartwatch and app training in 18 minutes on average. After consent and training, participants remained in the study for at least 28 days. Average wear time did not decline significantly over the course of the study and on average never fell below 20 h/d. Although there are no benchmark standards for indicating good smartwatch adherence, a recent systematic review of activity trackers to monitor physical activity suggested that 3 valid days of at least 10 h/d may be a good adherence threshold for a week-long study [38]. Another study defined a "valid day" as one that includes at least 16.67 hours of data or at least 600 one-minute epochs of nonzero heart rate values [37]. In our study, participants wore the study watch for >16.67 h/d on 92% of the days on average. Thus, our results exceed the recent suggestions for adherence and validity and show strong support for the feasibility and adherence of our protocol to monitor older adults.

Smartwatch feasibility and adherence metrics were associated with each other as expected, and EMA survey adherence also was related to smartwatch adherence measures. However, feasibility and adherence metrics were differentially related to participant factors, indicating the importance of measuring adherence to each feature of a smartwatch and EMA study separately. EMA survey adherence was not related to any participant characteristic and was uniformly high, possibly because participants received smartphone reminder alarms to complete the daily survey and sync the apps at predetermined times at night before bedtime. There was no comparable alert to prompt participants to wear the smartwatch during the day or night. Future studies should consider adding prompts to facilitate smartwatch adherence.

In contrast to past studies [21], we observed no associations between participant's sex or education level on any measure of adherence and feasibility in our diverse sample. Smartwatch adherence was associated with race such that White participants wore the smartwatch for more hours per day (daily wear time) than Black participants during the study period. The effect of race remained significant, even after controlling for other demographic variables, including age, education, and cognition. The reason for the race difference is unknown but may be explained by different beliefs, attitudes, or daily habits between the racial groups. As wear time is measured using the photoplethysmography sensor, which relies on the absorption of infrared light into the skin, it is possible that the measures might have been less accurate for people with darker skin tone. If replicated, cultural factors and potential sensor limitations related to skin tone, which might explain racial differences in adherence, should be systematically explored. Although Black and White participants differed in daily wear time, race did not influence the percentage of days that reached the 16.67-hour threshold for "valid" wear time. Thus, the race difference in daily wear time was small and potentially inconsequential. Furthermore, race was not associated with any other feasibility and adherence measure.

Although cognitive ability level (both self-reported cognitive decline and MCI status) was associated with daily wear time in bivariate analyses, the effect of cognition was not significant in multiple regression models. These results contrast with past research that has shown cognitive ability to be related to adherence; for example, memory ability was related to adherence in a study requiring the daily syncing of a Fitbit device in a sample of older adults without cognitive impairment [21]. Our study protocol, including training and supports during the monitoring period (eg, the take-home binder and twice daily alarms [in the morning to charge the watch and at night to sync the apps and complete the surveys]) may have helped participants with cognitive difficulties adhere to the smartwatch wear requirements and EMA survey.

Age was not associated with watch or survey adherence, which is consistent with the results from a recent smartphone digital phenotyping study [15]. However, similar to a recent study from our laboratory [13], we found that age was associated with training time such that older participants required longer training times in the laboratory to learn how to sync the study-related apps and complete the daily EMA survey on their smartphone. The training required participants to independently complete each activity (syncing apps, charging the smartwatch, and completing the EMA survey) while a member of the study team observed, answered questions, and provided feedback. On average, participants took <20 minutes to complete the training. Older participants may have taken longer to complete the training due to lower digital literacy; they may have required more practice and may have asked more questions during the training. Although the number of people aged >65 years owning smartphones in the United States has steadily increased over the past decade, digital literacy is generally lower in older adults. The longer training time and the take-home binder, which summarized the training for review during the study period, may have consequently minimized age effects on smartwatch and daily EMA survey adherence. Smartwatch studies that incorporate a "human in the loop" have been found to have higher adherence than fully remote studies [15]. Even if a study is conducted fully remotely, providing some contact with study personnel (via teleconference) or training resources (in the form of a training binder, video tutorial, website, or web-based helpline) could be quite beneficial to increasing adherence, especially for older adults with low digital literacy.

Participants generally demonstrated solid knowledge of the informed consent information, including issues related to data security and privacy that are relevant to research using wearable devices; for example, responses to the study consent quiz

XSL•FO
RenderX

showed that participants understood that the study apps were Health Insurance Portability and Accountability Act–compliant and that they had control over their data such that they had the right to delete their study data at any time. In contrast to a study on smartphone digital phenotyping that found that education and race were associated with quiz accuracy [13], percentage correct on the consent comprehension quiz was not significantly related to any of the participant demographic factors, possibly because the scores on the quiz had little variability (ie, ceiling effects). Similar to a past study [13], the most frequently incorrect question concerned the potential benefits of the study, with many participants incorrectly reporting that their participation would improve their cognitive abilities. In these cases, the comprehension quiz enabled the study team to explain that wearing the smartwatch had no clear benefit and possibly preclude participants' disappointment at the end of the study.

Smartwatch usability reports were largely very positive, with most of the participants (37/45, 82%) agreeing that they had a positive experience using the Garmin Vivosmart 4 smartwatch in the study as well as high satisfaction regarding its comfort and functioning. This is in line with past research showing that the Garmin Vivosmart was rated as the most usable and acceptable smartwatch compared to 5 other smartwatches by older adults in a small study [39]. Although participants indicated that they needed to charge the Garmin Vivosmart 4 smartwatch daily, they were generally satisfied with battery life. Regarding app usability, almost all participants were satisfied with the syncing process of the Labfront and Garmin Connect apps (42/45, 93%), felt that the Labfront app was easy to use (35/45, 78%), and reported a positive experience using the Labfront app (29/45, 64%). There are currently no other studies examining the usability and acceptability of the Labfront and Garmin Connect apps.

Regarding the usability of the smartwatch and apps, participants were generally satisfied with the smartwatch and reported that they would wear it again in another study. It is important to consider that most of the participants (44/45, 98%) reported having to charge the watch daily, and low battery life was the most frequently cited complaint about the smartwatch. Thus, in future studies, participants should be trained and prompted to charge watches and other digital devices, and they should be informed about battery limitations to preclude complaints and dissatisfaction. Multiple chargers, including portable batteries, may be offered to participants to improve usability. Despite minor complaints regarding charging, most of the participants (37/45, 82%) reported a positive experience during the study using the smartwatch. Participants also reported high satisfaction and ease of use of the Labfront and Garmin apps, which were used for daily syncing to enable sensor and EMA data to be transmitted to the research team. Taken together, our findings are generally in line with prior research showing good acceptability and usability of wearable devices in research studies involving older adults [21].

Limitations

Although our sample included older adults of different races, the sample was relatively small and highly educated. Furthermore, only 8 (18%) of the 44 participants had MCI; therefore, analyses comparing participants with MCI and those with healthy cognition were underpowered. However, cognitive effects on adherence were also examined with a composite test score and self-reports of cognitive and functional abilities (ECog). Moreover, our monitoring period was limited to a 4-week period; therefore, the conclusions may not generalize to longer periods. Importantly, our study incorporated several supports for participants such as in-person training, a take-home binder, contact with study personnel, and reminder alarms set for times specified by participants. These study features likely influenced the high adherence to smartwatch wear and EMA survey completion in our study compared to fully remote studies with little to no support from study personnel. In addition, our study's EMA survey imposed minimal demands because it was required only once daily and completed at a time preferred by participants.

Conclusions

This study demonstrated the feasibility of a combined smartwatch monitoring and daily EMA study requiring nearly 24-hour wear time daily for a month in racially diverse older adults, including those with cognitive impairment and those without. Smartwatch feasibility and adherence metrics were not significantly intercorrelated and were differentially related to participant characteristics. The inclusion of a guided training for older adult participants on study procedures as well as a take-home binder and smartphone reminder alarms to complete study steps (syncing apps, charging the smartwatch, and completing the daily EMA survey) may have contributed to strong adherence and feasibility in this study, regardless of cognitive impairment. Participants reported high satisfaction regarding the usability of both the Garmin Vivosmart 4 smartwatch and Labfront app. Overall, longitudinal monitoring using a commercially available smartwatch with a daily EMA survey is acceptable and feasible for collecting nearly continuous data in Black and White older adults, including those with cognitive impairment and those without.

Acknowledgments

The research in this publication was supported by the National Institute on Aging (F31AG089944: SH; R01AG062503: TG). The publication of this article was funded in part by the Temple University Libraries Open Access Publishing Fund.

Authors' Contributions

SH conceptualized the idea, ran statistical analyses, and drafted the manuscript. RC and MK contributed to data processing, analysis, and editing. TG contributed to conceptualization, writing, and editing. HS contributed to editing and statistical analyses.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Training binder, consent comprehension quiz, and supplementary analyses.
[PDF File (Adobe PDF File), 1432 KB - humanfactors_v12i1e69952_app1.pdf]

References

1. Kourtis LC, Regele OB, Wright JM, Jones GB. Digital biomarkers for Alzheimer's disease: the mobile/ wearable devices opportunity. NPJ Digit Med 2019;2:9. [doi: 10.1038/s41746-019-0084-2] [Medline: 31119198]
2. Frederiksen KS, Arus XM, Zetterberg H, Gauthier S, Boada M, Pytel V, et al. Focusing on earlier diagnosis of Alzheimer's disease. Future Neurol 2024 Mar 27;19(1):33. [doi: 10.2217/fnl-2023-0024]
3. Lin PJ, Daly AT, Olchanski N, Cohen JT, Neumann PJ, Faul JD, et al. Dementia diagnosis disparities by race and ethnicity. Med Care 2021 Aug 01;59(8):679-686 [FREE Full text] [doi: 10.1097/MLR.0000000000001577] [Medline: 34091580]
4. Faverio M. Share of those 65 and older who are tech users has grown in the past decade. Pew Rearch Centre. URL: https://www.pewresearch.org/short-reads/2022/01/13/share-of-those-65-and-older-who-are-tech-users-has-grown-in-the-past-decade/ [accessed 2024-09-11]
5. Sperling R, Mormino E, Johnson K. The evolution of preclinical Alzheimer's disease: implications for prevention trials. Neuron 2014 Dec 05;84(3):608-622 [FREE Full text] [doi: 10.1016/j.neuron.2014.10.038] [Medline: 25442939]
6. Schyvens AM, Van Oost NC, Aerts JM, Masci F, Peters B, Neven A, et al. Accuracy of Fitbit Charge 4, Garmin Vivosmart 4, and WHOOP versus polysomnography: systematic review. JMIR Mhealth Uhealth 2024 Mar 27;12:e52192 [FREE Full text] [doi: 10.2196/52192] [Medline: 38557808]
7. Evenson KR, Spade CL. Review of validity and reliability of Garmin activity trackers. J Meas Phys Behav 2020 Jul;3(2):170-185 [FREE Full text] [doi: 10.1123/jmpb.2019-0035] [Medline: 32601613]
8. Briggs BC, Hall KS, Jain C, Macrea M, Morey MC, Oursler KK. Assessing moderate to vigorous physical activity in older adults: validity of a commercial activity tracker. Front Sports Act Living 2021 Jan 3;3:766317 [FREE Full text] [doi: 10.3389/fspor.2021.766317] [Medline: 35047769]
9. Cheng YC, Huang YC, Huang WL. Heart rate variability in patients with dementia or neurocognitive disorders: a systematic review and meta-analysis. Aust N Z J Psychiatry 2022 Jan 08;56(1):16-27. [doi: 10.1177/0004867420976853] [Medline: 33287558]
10. Fan L, Xu W, Cai Y, Hu Y, Wu C. Sleep duration and the risk of dementia: a systematic review and meta-analysis of prospective cohort studies. J Am Med Dir Assoc 2019 Dec;20(12):1480-7.e5. [doi: 10.1016/j.jamda.2019.06.009] [Medline: 31604673]
11. Tan ZS, Spartano NL, Beiser AS, DeCarli C, Auerbach SH, Vasan RS, et al. Physical activity, brain volume, and dementia risk: the Framingham study. J Gerontol A Biol Sci Med Sci 2017 Jul 01;72(6):789-795 [FREE Full text] [doi: 10.1093/gerona/glw130] [Medline: 27422439]
12. Hehlmann MI, Schwartz B, Lutz T, Gómez Penedo JM, Rubel JA, Lutz W. The use of digitally assessed stress levels to model change processes in CBT - a feasibility study on seven case examples. Front Psychiatry 2021;12:613085 [FREE Full text] [doi: 10.3389/fpsyt.2021.613085] [Medline: 33767638]
13. Hackett K, Xu S, McKniff M, Paglia L, Barnett I, Giovannetti T. Mobility-based smartphone digital phenotypes for unobtrusively capturing everyday cognition, mood, and community life-space in older adults: feasibility, acceptability, and preliminary validity study. JMIR Hum Factors 2024 Dec 22;11:e59974 [FREE Full text] [doi: 10.2196/59974] [Medline: 39576984]
14. Trull TJ, Ebner-Priemer U. Ambulatory assessment. Annu Rev Clin Psychol 2013;9:151-176 [FREE Full text] [doi: 10.1146/annurev-clinpsy-050212-185510] [Medline: 23157450]
15. Pratap A, Neto EC, Snyder P, Stepnowsky C, Elhadad N, Grant D, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. NPJ Digit Med 2020;3:21 [FREE Full text] [doi: 10.1038/s41746-020-0224-8] [Medline: 32128451]
16. Goodday SM, Karlin E, Alfarano A, Brooks A, Chapman C, Desille R, et al. An Alternative to the Light Touch Digital Health Remote Study: The Stress and Recovery in Frontline COVID-19 Health Care Workers Study. JMIR Form Res 2021 Dec 10;5(12):e32165 [FREE Full text] [doi: 10.2196/32165] [Medline: 34726607]
17. Druce KL, McBeth J, van der Veer SN, Selby DA, Vidgen B, Georgatzis K, et al. Recruitment and ongoing engagement in a UK smartphone study examining the association between weather and pain: cohort study. JMIR Mhealth Uhealth 2017 Dec 01;5(11):e168 [FREE Full text] [doi: 10.2196/mhealth.8162] [Medline: 29092810]
18. Reade S, Spencer K, Sergeant JC, Sperrin M, Schultz DM, Ainsworth J, et al. Cloudy with a chance of pain: engagement and subsequent attrition of daily data entry in a smartphone pilot study tracking weather, disease severity, and physical activity in patients with rheumatoid arthritis. JMIR Mhealth Uhealth 2017 Mar 24;5(3):e37 [FREE Full text] [doi: 10.2196/mhealth.6496] [Medline: 28341616]

XSL•FO
RenderX

19. Matcham F, Leightley D, Siddi S, Lamers F, White KM, Annas P, RADAR-CNS consortium. Remote Assessment of Disease and Relapse in Major Depressive Disorder (RADAR-MDD): recruitment, retention, and data availability in a longitudinal remote measurement study. BMC Psychiatry 2022 Mar 21;22(1):136 [FREE Full text] [doi: 10.1186/s12888-022-03753-1] [Medline: 35189842]
20. Baig MM, Afifi S, GholamHosseini H, Mirza F. A systematic review of wearable sensors and IoT-based monitoring applications for older adults - a focus on ageing population and independent living. J Med Syst 2019 Jul 15;43(8):233. [doi: 10.1007/s10916-019-1365-7] [Medline: 31203472]
21. Paolillo EW, Lee SY, VandeBunte A, Djukic N, Fonseca C, Kramer JH, et al. Wearable use in an observational study among older adults: adherence, feasibility, and effects of clinicodemographic factors. Front Digit Health 2022 Jun 10;4:884208 [FREE Full text] [doi: 10.3389/fdgth.2022.884208] [Medline: 35754462]
22. Pathiravasan CH, Zhang Y, Wang X, Trinquart L, Benjamin EJ, Borrelli B, et al. Factors associated with long-term use of digital devices in the electronic Framingham Heart Study. NPJ Digit Med 2022 Dec 27;5(1):195 [FREE Full text] [doi: 10.1038/s41746-022-00735-1] [Medline: 36572707]
23. McInnis BJ, Pindus R, Kareem D, Gamboa S, Nebeker C. Exploring the future of informed consent: applying a service design approach. Proc ACM Hum Comput Interact 2024 Apr 26;8(CSCW1):1-31. [doi: 10.1145/3637330]
24. Bartlett Ellis R, Wright J, Miller LS, Jake-Schoffman D, Hekler EB, Goldstein CM, et al. Lessons learned: beta-testing the digital health checklist for researchers prompts a call to action by behavioral scientists. J Med Internet Res 2021 Dec 22;23(12):e25414 [FREE Full text] [doi: 10.2196/25414] [Medline: 34941548]
25. Nebeker C, Bartlett Ellis RJ, Torous J. Development of a decision-making checklist tool to support technology selection in digital health research. Transl Behav Med 2020 Oct 08;10(4):1004-1015 [FREE Full text] [doi: 10.1093/tbm/ibz074] [Medline: 31120511]
26. Tombaugh TN. Trail making test A and B: normative data stratified by age and education. Arch Clin Neuropsychol 2004 Mar;19(2):203-214. [doi: 10.1016/S0887-6177(03)00039-8] [Medline: 15010086]
27. Blackburn HL, Benton AL. Revised administration and scoring of the digit span test. J Consult Psychol 1957 May;21(2):139-143. [doi: 10.1037/h0047235] [Medline: 13416432]
28. Schretlen D, Testa S, Pearlson G. Calibrated Neuropsychological Normative System. Lutz, FL: Psychological Assessment Resources; 2010.
29. Kaplan E, Goodglass H, Weintraub S. Boston naming test. Clin Neuropsychol 2016. [doi: 10.1037/t27208-000]
30. Salthouse TA. When does age-related cognitive decline begin? Neurobiol Aging 2009 May;30(4):507-514 [FREE Full text] [doi: 10.1016/j.neurobiolaging.2008.09.023] [Medline: 19231028]
31. Benedict RH, Schretlen D, Groninger L, Brandt J. Hopkins verbal learning test – revised: normative data and analysis of inter-form and test-retest reliability. Clin Neuropsychol 2010 Aug 09;12(1):43-55. [doi: 10.1076/clin.12.1.43.1726]
32. Benedict RH, Schretlen D, Groninger L, Dobraski M, Shpritz B. Revision of the brief visuospatial memory test: studies of normal performance, reliability, and validity. Psychol Assess 1996 Jun;8(2):145-153. [doi: 10.1037/1040-3590.8.2.145]
33. All-in-one research solution for real-world data capture. Labfront. URL: https://www.labfront.com [accessed 2024-04-29]
34. Schretlen DJ, Winicki JM, Meyer SM, Testa SM, Pearlson GD, Gordon B. Development, psychometric properties, and validity of the hopkins adult reading test (HART). Clin Neuropsychol 2009 Aug;23(6):926-943. [doi: 10.1080/13854040802603684] [Medline: 19191072]
35. Bondi MW, Edmonds EC, Jak AJ, Clark LR, Delano-Wood L, McDonald CR, et al. Neuropsychological criteria for mild cognitive impairment improves diagnostic precision, biomarker associations, and progression rates. J Alzheimers Dis 2014;42(1):275-289 [FREE Full text] [doi: 10.3233/JAD-140276] [Medline: 24844687]
36. Farias ST, Mungas D, Reed BR, Cahn-Weiner D, Jagust W, Baynes K, et al. The measurement of everyday cognition (ECog): scale development and psychometric properties. Neuropsychology 2008 Jul;22(4):531-544 [FREE Full text] [doi: 10.1037/0894-4105.22.4.531] [Medline: 18590364]
37. Nelson BW, Low CA, Jacobson N, Areán P, Torous J, Allen NB. Guidelines for wrist-worn consumer wearable assessment of heart rate in biobehavioral research. NPJ Digit Med 2020;3:90 [FREE Full text] [doi: 10.1038/s41746-020-0297-4] [Medline: 32613085]
38. Chan A, Chan D, Lee H, Ng CC, Yeo AH. Reporting adherence, validity and physical activity measures of wearable activity trackers in medical research: a systematic review. Int J Med Inform 2022 May;160:104696 [FREE Full text] [doi: 10.1016/j.ijmedinf.2022.104696] [Medline: 35121356]
39. Tocci FL, Morey MC, Caves KM, Deberry J, Leahy GD, Hall K. Are older adults ready for wireless physical activity tracking devices? A comparison of commonly used tracking devices. J Am Geriatr Soc 2016 Jan;64(1):226-228. [doi: 10.1111/jgs.13895] [Medline: 26782889]

Abbreviations

AD: Alzheimer disease
ADRD: Alzheimer disease and related dementias
ECog: Everyday Cognition

XSL•FO
RenderX

EMA: ecological momentary assessment
MCI: mild cognitive impairment

Edited by D Gooch; submitted 11.12.24; peer-reviewed by Y Zhang, C Wu, NG Schaham; comments to author 04.02.25; revised version received 03.03.25; accepted 17.03.25; published 08.04.25.

Please cite as:
Holmqvist S, Kaplan M, Chaturvedi R, Shou H, Giovannetti T
Longitudinal and Combined Smartwatch and Ecological Momentary Assessment in Racially Diverse Older Adults: Feasibility, Adherence, and Acceptability Study
JMIR Hum Factors 2025;12:e69952
URL: https://humanfactors.jmir.org/2025/1/e69952
doi:10.2196/69952
PMID:

©Sophia Holmqvist, Marina Kaplan, Riya Chaturvedi, Haochang Shou, Tania Giovannetti. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 08.04.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Impact of a Free Influenza Vaccination Policy on Older Adults in Zhejiang, China: Cross-Sectional Survey of Vaccination Willingness and Determinants

Yusui Zhao, MS; Jinhang Xu, MS; Xuehai Zhang, MS; Yue Xu, MS; Xiaotong Yan, MS; Shaofan Ren, MS; Lixin Wang, MS

Zhejiang Provincial Center for Disease Control and Prevention, 3399 Binsheng Road, Hangzhou, China

Corresponding Author:
Lixin Wang, MS
Zhejiang Provincial Center for Disease Control and Prevention, 3399 Binsheng Road, Hangzhou, China

Abstract

Background: In 2024, Zhejiang Province introduced a new policy offering free influenza vaccinations to individuals aged 60 years and older. However, the vaccination willingness among the newly eligible 60-69 years age group remains ambiguous in comparison to those aged 70 years and older.

Objective: This study aimed to evaluate the willingness of individuals aged ≥60 years in Zhejiang Province, China, to receive free influenza vaccines under a newly implemented policy. It further explored their sources of influenza-related health information and identified key determinants of vaccination hesitancy across age subgroups.

Methods: A cross-sectional survey was conducted using multistage convenience sampling via on-site questionnaires. Structured questionnaires were administered to 7162 eligible participants aged ≥60 years from March to May 2024. Valid responses (n=7103; response rate: 99.18%) were analyzed via logistic regression and Kruskal-Wallis tests.

Results: Overall vaccination willingness was 73.15% (5196/7103), with 11.71% (832/7103) refusal and 15.14% (1075/7103) hesitancy. Key predictors of hesitancy included male gender (odds ratio [OR] 1.27, 95% CI 1.05 - 1.54), ages 60 - 69 years (OR 1.46, 95% CI 1.06 - 2.02), corporate employment (OR 0.75, 95% CI 0.58 - 0.98), and absence of chronic diseases (OR 2.06, 95% CI 1.44 - 2.96). The 60 - 69 year age group demonstrated lower awareness of the free policy (61.9% vs 73.72% in the ≥70 years group; H=61.25, *P*<.001) but higher engagement with social media (WeChat [Tencent Holdings Limited]: H=345.44; TikTok [ByteDance Ltd]: H=294.66; *P*<.001) for health information.

Conclusions: Despite high willingness, knowledge gaps persist, particularly among adults aged 60 - 69 years. Targeted dissemination of policy information via social media platforms (eg, WeChat and TikTok) and community-driven campaigns is recommended to enhance vaccination uptake. This approach may serve as a model for regions implementing similar policies. Future studies should track actual vaccination uptake postpolicy and explore artificial intelligence–driven social media interventions to boost engagement.

(JMIR Hum Factors 2025;12:e73940) doi:10.2196/73940

KEYWORDS

influenza vaccine; vaccination willingness; older adults; health policy; China

Introduction

Influenza, a prevalent seasonal acute respiratory infectious disease, exhibits an escalating incidence rate in China [1]. The virus is characterized by its high antigenic variability and rapid transmission, resulting in annual epidemics. Globally, influenza is responsible for approximately 390,000 deaths annually, with two-thirds of these fatalities occurring among older adults [2]. In China, individuals aged ≥60 years account for 80% of the nation's influenza-associated excess respiratory mortality [3]. Within Zhejiang Province, surveillance data from 2009 to 2021 indicate 990,016 reported cases and 8 deaths, with estimated actual cases 12.11-fold higher than reported [4]. Influenza vaccination—recognized as the most cost-effective intervention for preventing respiratory infections by the Chinese Center for Disease Control and Prevention [5]—reduces disease burden, hospitalizations, and mortality in older adults [6,7]. A 2013 meta-analysis of 95 studies confirmed vaccination prevents 28% of complications, 39% of influenza-like illnesses, and 49% of confirmed cases in this demographic [8].

The influenza vaccination uptake among the older population in China has traditionally been low. For instance, during the 2011 - 2012 flu season, only 4.3% of individuals aged ≥60 years in 9 Chinese cities were vaccinated against influenza [9]. This rate further declined to 3.8% in the 2014 - 2015 flu season [10].

XSL•FO
RenderX

However, in recent years, there has been a marked increase in the vaccination rate among older adults, attributable to the Chinese government's heightened emphasis on vaccinations for older adults. This surge is particularly noticeable in the aftermath of the COVID-19 outbreak and the introduction of free influenza vaccination policies by certain local governments. In the 2020‐2021 and 2021‐2022 periods, free influenza vaccinations were offered in 195 and 226 counties and districts, respectively, leading to vaccination rates of 45.7% and 32.9% among the older populations in these regions [11]. Despite this progress, the influenza vaccination rate among older adults in China remains below the World Health Organization's target of 75%.

In China, influenza vaccines were categorized as Category II, necessitating voluntary self-funded vaccination by residents [12]. This classification contributed to the low uptake of influenza vaccination. However, with China's economic progression, policies granting free influenza vaccinations to individuals aged ≥60 years were first introduced in major cities, including Beijing, Shanghai, and Shenzhen [13-15]. Extensive research confirmed that these measures substantially elevated both the willingness and rate of influenza vaccination among older adults. As a result, the initiative extended to other regions, encompassing areas within Zhejiang and Shandong provinces [16,17]. Notably, Zhejiang Province was among the pioneers in this endeavor. Specific counties and districts, such as Taizhou and Ningbo [18,19], began offering complimentary influenza vaccinations to those aged ≥60 years from 2017, aligning with local government initiatives for public welfare. By 2020, Zhejiang had fully adopted the policy for residents aged ≥70 years [16], and by 2024, it was expanded to cover all residents aged ≥60 years in the province [20].

Several studies have examined influenza vaccination willingness in Chinese regions with similar policies. In Beijing (free for ≥60 years since 2007), willingness reached 76.61% post policy implementation [21]. Shanghai reported 80.6% willingness among adults ≥50 years [22], while Hangzhou (Zhejiang's capital) observed 85.98% uptake in the ≥70 year age group under free vaccination [16]. Lower willingness was noted in non–first-tier cities (eg, 60.7% in Chongqing [23] and 53.66% in Kaifu District, Changsha [24]). These studies consistently identified policy awareness, income level, and prior vaccination history as key determinants. However, no prior study has assessed the newly eligible 60‐69 year age group in a province-wide free vaccination program, highlighting the necessity of this investigation.

Despite the existing circumstances, the influenza vaccination rate among individuals aged ≥60 years in Zhejiang Province was a mere 21.76% [25] prior to the introduction of the universal free influenza vaccination policy for this age group. The potential impact of this policy on vaccination rates, particularly among those aged between 60 and 69 years old, remains uncertain. However, an individual's willingness to receive the flu vaccine is a critical factor influencing vaccination rates; the stronger the willingness, the higher the likelihood of actual vaccination [26,27]. Given this context, our study aims to (1) analyze the willingness to receive the influenza vaccine and its determinants among individuals aged ≥60 years in Zhejiang Province under the free vaccination policy in 2024, (2) investigate the ways from which different age groups of older individuals acquire knowledge about influenza prevention and treatment, and (3) identify the primary reasons why different age groups of older individuals are willing or unwilling to get vaccinated against influenza.

Methods

Study Setting, Sampling and Recruitment Procedure

Zhejiang was a province located in the southeastern part of China with a population of 66 million at the end of 2023. Among this population, approximately 14 million individuals were aged ≥60 years, representing about 21.5% of the total population. In addition, Zhejiang had experienced the second-largest influx of migrant workers in China and remained a significant province for this demographic group. This situation further increased the risk of infectious disease transmission among them.

The policy was announced via government portals in January 2024. After 2 months for policy dissemination, a questionnaire survey was administered from March to May 2024 to the selected participants. Inclusion criteria were as follows: (1) aged ≥60 years, (2) residents of the selected community or village, and (3) individuals who were informed about the study's content and purpose and voluntarily agreed to participate. Exclusion criteria included (1) plans to be away from the community or village for more than 3 months in the upcoming year, (2) cognitive impairments that hindered normal communication, and (3) lack of cooperation with the study and poor compliance (Multimedia Appendix 1).

Measures

The instrument was developed based on the World Health Organization flu vaccine hesitancy scales. Cognitive testing was conducted with 20 older participants in January 2024, followed by a pilot survey in 2 communities (n=200). Cronbach α improved from 0.68 to 0.765 after revising items with poor discrimination (item-total correlation <0.2).

The self-developed questionnaire was composed of 3 different parts, including sociodemographic information, knowledge on influenza prevention and treatment, and reasons for willingness or unwillingness to receive influenza vaccination (Multimedia Appendix 2).

Sex, age, ethnicity, marital status, family structure, education, occupation, income, chronic illness status, and history of influenza vaccination were included in sociodemographic information.

The section of knowledge on influenza prevention and treatment consisted of 11 items, segmented into 3 categories: basic knowledge of influenza, knowledge on preventing influenza, and understanding of influenza vaccination. The fundamental knowledge contained 7 single-choice questions. The influenza prevention knowledge had 1 multiple-choice question, while the influenza vaccination knowledge comprised 3 questions, with 2 being single-choice and 1 a multiple-choice. Each correct answer in the single-choice section was allocated 1 point, whereas a completely accurate response in the multiple-choice

XSL•FO
RenderX

section was given 2 points. Incorrect answers, whether due to overselection, underselection, incorrect selection, or omission, were not awarded any points. The maximum possible score was 13, known as the influenza prevention knowledge score. Each of the 7 items had a content validity index greater than 0.8 and the overall Cronbach α coefficient was 0.765. The Kaiser-Meyer-Olkin measure was 0.855 and the Bartlett test *P* value was less than .01, indicating adequate reliability and validity.

The section of reasons for willingness or unwillingness to receive influenza vaccination consisted of 3 questions: "How do you learn about health information?," "Why are you willing to get a flu vaccine?," and "Why haven't you received the flu vaccine?" The 3 questions mentioned above were all presented in a multiple-choice format for respondents.

The following were also taken into consideration during data collection:

- Occupation: for older respondents, if they are retired, record the occupation they were engaged in before retirement. If they are still working, record their current occupation.
- Income: personal monthly income refers to the total amount of taxable income earned by an individual within a tax month. For older respondents, sources of income include both wages (or pensions) and other types of income combined.
- Chronic diseases refer to conditions that have been definitively diagnosed by a physician, including hypertension, diabetes mellitus, hyperlipidemia, chronic obstructive pulmonary disease, bronchitis, stroke, coronary artery disease, and tumors.

Data Analysis

A database was constructed using EpiData 3.1 software (EpiData Association), with data entry executed in duplicate to ensure accuracy. Statistical analysis was performed using SAS Viya Long-Term Support 2024.03 (SAS Institute). Descriptive analysis was used to examine demographic characteristics, and the Kruskal-Wallis (H) test was used for this purpose. ANOVA (*F* test) was conducted to compare scores related to influenza prevention and treatment knowledge. Multivariate logistic regression analysis was implemented to identify factors that influence vaccination intention. A *P* value less than .05 was deemed statistically significant.

The dependent variable is influenza vaccination willingness (positive=1; hesitation=2; negative=3). The term "positive" is used to denote survey participants who explicitly stated their willingness to receive the influenza vaccine. The term "hesitation" refers to those who demonstrated hesitancy, while "negative" pertains to individuals who clearly expressed their unwillingness to receive the vaccine. The independent variables consistently included in all models are as follows: sex (male=1; female=2), age (60 - 69 years old=1; 70 - 79 years old=2; aged ≥80 years old=3), marital status (unmarried=1; married=2; unmarried or divorced or widowed=3), family structure (solitary living=1; living with spouse or children=2; living with spouse and children=3; other=4), education level (primary school or lower=1; middle school=2; high school or technical school=3; college or higher=4), occupation (agency or institutional personnel=1; medical staff=2; farmers corporate staff=3; sole proprietors=4; other=5), chronic illness status (none=1; 1 to 2=2; 3 or more=3), history of influenza vaccination (yes=1; no=2), and income (less than 2000 RMB [approximately US $281]=1; 2000 - 4999 RMB [approximately US $281-$703]=2; 5000 - 10,000 RMB [approximately US $703-$1407]=3; more than 10,000 RMB [approximately >US $1407]=4; unclear=5). All conversions from RMB to US $ are based on the average exchange rate on March 1, 2024 (US $1=7.1066 RMB).

Ethical Considerations

The study was conducted in accordance with the Declaration of Helsinki. Informed consent was obtained from all participants or their guardians. Each participant received a small gift valued at 100 RMB (approximately US $14) as compensation. Surveyors recorded responses without personal identifiers; data were anonymized during electronic entry using coded IDs. The study protocol was approved by the ethics committee of the Zhejiang Provincial Center for Disease Control and Prevention (2022-016-02).

Results

Demographics and Univariate Analysis of Vaccination Willingness

In 2024, a total of 7162 individuals were surveyed, yielding 7103 valid responses, reflecting a response rate of 99.18%. The gender distribution was nearly equal, with 3566 males and 3537 females. Approximately 91.9% (6533/7103) of the participants were between the ages of 60 to 79 years, while only 8.0% (570/7103) were aged 80 years or older. Over half of the respondents (4577/7103, 64.4%) had an educational level at or below primary school, and more than 80% (5969/7103) reported an average monthly income of less than 5000 RMB (approximately US $703).

The univariate analysis results revealed that factors such as sex, age, marital status, family structure, education, occupation, income, chronic illness status, and influenza vaccination history significantly influenced respondents' willingness to receive the influenza vaccine. Of these, the most influential factor was found to be influenza vaccination history. Respondents with a history of influenza vaccination demonstrated a higher willingness to be vaccinated compared to those who had never been vaccinated (n=3572, 86.7% vs n=1624, 54.5%) against influenza. Age was identified as the second most influential factor. Respondents aged 70 - 79 years expressed the highest willingness to receive the vaccine (n=2389, 78.1%), while those aged 60 - 69 years showed the least willingness (n=2371, 68.3%). Medical staff showed the highest willingness to be vaccinated compared to other occupational groups (n=176, 81.9%). Furthermore, respondents with chronic diseases were more likely to express a willingness to receive the influenza vaccine, although they also had a higher proportion of unwillingness to be vaccinated (Table S1 in Multimedia Appendix 3).

XSL•FO
RenderX

Multivariate Analysis of Vaccination Willingness

A multivariate logistic regression analysis was performed, using the willingness to receive the influenza vaccine as the dependent variable and indicators that demonstrated statistically significant differences in univariate analysis as independent variables. The findings revealed that the 60 - 69 year age group, medical staff, individuals with chronic illnesses, and those who had a history of influenza vaccination were all contributing factors to the willingness to receive the influenza vaccine. Notably, residents with a history of influenza vaccination were 7.82 times more likely to express willingness to receive the influenza vaccine compared to those without such a history (odds ratio [OR] 7.82, 95% CI 6.52 - 9.39; $P<.001$) (Table 1).

XSL•FO
RenderX

Table . The multivariate logistic regression results of factors associated with influenza vaccination willingness of participants.

Dependent and independent variable level	B	SE	Wald χ^2 (*df*)	*P* value	OR[a] (95% CI)
Positive					
Age (years)					
60 - 69	0.38	0.17	5.36 (1)	.02	1.46 (1.06 - 2.02)
70 - 79	0.06	0.16	0.15 (1)	.70	1.84 (1.40 - 2.41)
≥80 (Reference)	__[b]	—	—	—	—
Occupation					
Agency or Institutional personnel	0.06	0.24	0.07 (1)	.79	1.07 (0.67 - 1.69)
Medical staff	0.11	0.34	0.11 (1)	.74	1.12 (0.58 - 2.18)
Corporate staff	–0.28	0.13	4.45 (1)	.04	0.75 (0.58 - 0.98)
Sole proprietors	–0.22	0.14	2.74 (1)	.10	0.80 (0.62 - 1.04)
Farmers (Reference)	—	—	—	—	—
Chronic illness status					
None	0.37	0.15	6.12 (1)	.01	1.44 (1.08 - 1.93)
1 to 2	0.31	0.13	5.42 (1)	.02	1.37 (1.05 - 1.77)
3 or more (Reference)	—	—	—	—	—
Influenza vaccination history					
No	2.06	0.09	490.76 (1)	<.001	7.82 (6.52 - 9.39)
Yes (Reference)	—	—	—	—	—
Hesitation					
Sex					
Male	0.24	0.10	6.07 (1)	.01	1.27 (1.05 - 1.54)
Female (Reference)	—	—	—	—	—
Influenza vaccination history					
No	0.57	0.11	25.76 (1)	<.001	1.76 (1.42 - 2.19)
Yes (Reference)	—	—	—	—	—
Chronic illness status					
None	0.72	0.18	15.51 (1)	<.001	2.06 (1.44 - 2.96)
1 to 2	0.39	0.17	5.15 (1)	.02	1.47 (1.05 - 2.06)
3 or more (Reference)	—	—	—	—	—

[a]OR: odds ratio.

[b]Not applicable.

Comparison of Influenza Prevention and Treatment Knowledge Scores

The results showed that the scores for influenza prevention and treatment knowledge among respondents aged 60 - 69 years was 6.5 (SD 3.19); among those aged 70 - 79 years, it was 6.1 (SD 3.24), and among those aged ≥80 years, it was 5.2 (SD 3.37). There was a statistically significant difference in the scores of influenza prevention and treatment knowledge among different age groups (F_2=45.55, *P*<.001), with older age groups having lower scores.

XSL•FO
RenderX

Access to Health Information

The findings indicated that television, doctors, and family members were the main ways of obtaining health knowledge. Compared to the 70 - 79 and ≥80 year groups, individuals in the 60 - 69 year age group exhibited a higher propensity for leveraging new media platforms such as WeChat (Tencent Holdings Limited; H=345.44, *P*<.001) and TikTok (ByteDance Ltd; H=294.66, *P*<.001) as primary ways for acquiring health information. Despite high social media use (WeChat: n=1366, 39.34%), the 60 - 69 year group's policy awareness showed no correlation with platform engagement (*r*=0.04, *P*=.11), suggesting ineffective information diffusion (Table S4 in Multimedia Appendix 4).

Reasons for Willingness to Receive Influenza Vaccination

A questionnaire survey was conducted among respondents who were willing to receive the influenza vaccine. The findings indicated that the predominant reason influencing participants' willingness to receive the influenza vaccine was "awareness of the free vaccine policy." Awareness of the free policy was 70.16% (2013/2869) overall but significantly lower in those aged 60 - 69 years (n=867, 61.9% vs n=2002, 73.72% in ≥70 year group; *P*<.001). In addition, statistically significant differences among different age groups were noted for the reasons "active promotion by the community" (H=6.740, *P*=.03) and "awareness of the advantages of getting vaccinated" (H=33.125, *P*<.001). No other reasons demonstrated a statistically significant difference (Table S5 in Multimedia Appendix 5).

Reasons for Unwillingness to Receive Influenza Vaccination

The results indicated that "concerns about the necessity of vaccination," "not previously acknowledged or recognized," and "concerns about the effectiveness of vaccines" were the 3 primary reasons for respondents not willing to receive the influenza vaccine. These reasons did not show any statistically significant differences among different age groups. In comparison, respondents in the ≥80 year group were more concerned about the safety of the vaccine (H=10.96, *P*=.004). Those in the 70 - 79 year age group were more likely to refuse the vaccine due to contraindications (H=28.40, *P*<.001), while those in the 60 - 69 year age group were more concerned about the burden of vaccination (H=6.96, *P*=.03) (Table S6 in Multimedia Appendix 6).

Discussion

Principal Findings

This study indicated that the willingness to receive the influenza vaccine among residents aged ≥60 years in Zhejiang Province was 73.15%, which was lower than the vaccination willingness rates among the older population in first-tier Chinese cities such as Beijing and Shanghai [21,22]. It was also lower than that of Hangzhou, the capital city of Zhejiang Province, but higher than that of other non–first-tier cities in China [23,24], as well as higher than similar survey results from countries such as the United States and Italy [28,29]. On one hand, after 3 years of the COVID-19 pandemic, China had unprecedentedly intensified its publicity efforts on respiratory infectious disease prevention, significantly improving residents' knowledge, attitudes, and behaviors regarding respiratory infectious disease control [30]. On the other hand, the economic burden of influenza vaccination affected residents' willingness to get vaccinated, especially for low-income middle-aged and older individuals. Whether the vaccine was free played a crucial role in their decision to get vaccinated [31].

This study found that older individuals who were female, aged ≥70 years, had a history of influenza vaccination, and had chronic diseases were more willing to receive the flu vaccine. This finding was consistent with some domestic studies [32,33]. The history of influenza vaccination was the most significant factor influencing the willingness to vaccinate among the older adults [26,34,35]. The willingness to vaccinate among the older adults who had previously received the influenza vaccine was 7.823 times higher than that of those who had not. Prior recipients may perceive reduced infection severity, but causality between vaccination history and economic burden relief remains unproven. Age was another important factor. Compared to those aged ≥70 years, the willingness to receive the vaccine was lowest among those aged 60 - 69 years, who also had the highest proportion of hesitation responses. The primary reason was likely due to the recent implementation of a free influenza vaccination policy for this age group in Zhejiang Province, which many were still unaware of. The free vaccination policy was one of the most crucial factors driving vaccination rates, especially among low-income older populations [36]. When examining reasons for willingness to vaccinate, it was found that the proportion of those aware of the free vaccine policy in the 60 - 69 year age group was significantly lower than in other age groups, further explaining their lower vaccination willingness. However, encouragingly, the 60 - 69 year age group had a higher acceptance of the influenza vaccine and better recognized its benefits.

To assess knowledge levels about influenza prevention and treatment across different age groups, an 11-question questionnaire was designed. Analysis showed that the 60 - 69 year age group scored highest on the knowledge assessment, yet they had the lowest willingness to vaccinate, contradicting previous study in Hangzhou [16]. This discrepancy might be attributed to the limited number of questions in their questionnaire (consisting of merely 5 questions), which could introduce randomness. To ensure the questionnaire's accuracy and reliability, it was repeatedly revised until both its reliability and validity were satisfactory. Another analysis revealed that despite higher educational levels, better mental health, and stronger memory among the 60 - 69 year age group, their unfamiliarity with the newly introduced free vaccination policy contributed to their lower vaccination willingness compared to those aged ≥70 years.

This study found that television, medical staff, and family members are the main ways for older adults to obtain health information. In China, traditional media such as television remains the primary source of health information for older adults [16]. Despite this, their trust in doctors is higher [37]. It is recommended that medical staff should be encouraged to

XSL•FO
RenderX

strengthen the promotion of influenza vaccine policy and related prevention knowledge among older adults, in order to better help them establish a positive willingness to get vaccinated against the flu and thus make the decision to receive the influenza vaccine. What is different is that compared to those aged ≥70 years, the 60 - 69 year age group more often obtains health information through television, while the ≥80 year age group more often relies on family members for health information. Therefore, for the newly implemented policy in Zhejiang Province of free influenza vaccination for those aged 60 - 69 years, it is advisable to promote this policy to the older adults through local mainstream television media. This would not only increase the coverage of the promotion but also enhance its authority and credibility. It is worth noting that with the widespread proliferation of the internet in China, new media represented by WeChat and TikTok have risen strongly. According to statistics, from 2010 to 2020, over a 10-year period, the number of internet users aged ≥60 years increased from 8.67 million to 60.54 million, with the proportion of older internet users in the older population increasing from 4.9% to 23.8% [38]. Compared to those aged ≥70 years, the 60 - 69 year age group is far more active on WeChat or TikTok every day. This also suggests that for the 60 - 69 year age group, the way of obtaining health information is gradually changing, and new media is playing an increasingly important role.

From the analysis of reasons for willingness to receive influenza vaccination, apart from "awareness of the free vaccine policy," "active promotion by the community" also plays a crucial role in encouraging older individuals to get vaccinated. Compared to those aged ≥70 years, individuals in the 60 - 69 year age group are more likely to believe that receiving the influenza vaccine can benefit their health. They also pay more attention to the necessity of getting vaccinated [39]. On the other hand, those aged ≥70 years are more concerned about the safety of the influenza vaccine [40]. This highlights the importance of enhancing community promotion of the free influenza vaccination policy, as well as disseminating information about the necessity and safety of the influenza vaccine, to increase the willingness to vaccinate among older adults, particularly those in the 60 - 69 year age group. Future implementation research should examine digital outreach efficacy through platforms such as TikTok or WeChat, particularly in low-resource settings.

This study has some limitations. First, the sampling method used in this study was a multistage convenience sampling approach to obtain respondents. The selection process favors individuals more likely to be accessible or compliant, skewing the results toward higher vaccination willingness. Second, the investigators were generally doctors from local hospitals or community health service centers (township health clinics), who were relatively familiar with the respondents. During the survey process, respondents might tend to provide more positive responses, which could potentially overestimate their willingness to receive influenza vaccination. Third, some questions were answered based on recall alone, such as influenza vaccination history and chronic illness status, which might lead to recall bias. Finally, the cross-sectional survey method used in this study might affect the generalizability of the results.

Conclusions

Targeted dissemination via social media (eg, TikTok or WeChat) for the 60 - 69 year age group and community-driven policy education are critical next steps. This approach may serve as a model for regions implementing similar policies. Future studies should track actual vaccination uptake and test artificial intelligence–optimized messaging through digital platforms.

Acknowledgments

We thank all residents who participated in the implementation of this project. We also thank the staff from the survey sites for their field investigation and colleagues in the Institute of Immunization Program, Zhejiang Center for Disease Control and Prevention for their technical support. This work is supported by the Chinese National Sci-Tech Plan Project and the Zhejiang Provincial Government.

Generative artificial intelligence was not used for any part of the article.

Data Availability

The dataset generated from this research has undergone anonymization and is available for scholarly purposes from the corresponding author on reasonable request. Nevertheless, it is imperative to acknowledge that data dissemination is subject to constraints imposed by the Ethics Committee of the Zhejiang Provincial Center for Disease Control and Prevention, which restricts extensive public distribution. Furthermore, the participant consent procured does not encompass unrestricted public sharing of the dataset.

Authors' Contributions

LW, JX and XZ conceived and designed the study; YZ, YX, and XY collected the data; YZ, XY, and SR analyzed and interpreted the data; and LW, JX and XZ helped draft, read, and approve the final manuscript.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Sampling framework and sample size calculation.
[DOCX File, 13 KB - humanfactors_v12i1e73940_app1.docx]

Multimedia Appendix 2
Questionnaire on influenza vaccination willingness among the older population in Zhejiang Province.
[DOC File, 48 KB - humanfactors_v12i1e73940_app2.doc]

Multimedia Appendix 3
Comparison of influenza vaccination willingness among different groups.
[DOCX File, 19 KB - humanfactors_v12i1e73940_app3.docx]

Multimedia Appendix 4
The means of obtaining health information among different age groups of participants.
[DOCX File, 14 KB - humanfactors_v12i1e73940_app4.docx]

Multimedia Appendix 5
Comparison of the reasons for willingness to receive influenza vaccination among different age groups.
[DOCX File, 14 KB - humanfactors_v12i1e73940_app5.docx]

Multimedia Appendix 6
Comparison of the reasons for unwillingness to receive influenza vaccination among different age groups.
[DOCX File, 14 KB - humanfactors_v12i1e73940_app6.docx]

References

1. Zhang Z, Wang X, Liang X, Lin Z, Yan Y. Influenza. Chin J Zoonoses 2020;36(7):593-598. [doi: 10.3969/j.issn.1002-2694.2020.00.090]
2. Paget J, Spreeuwenberg P, Charu V, et al. Global mortality associated with seasonal influenza epidemics: new burden estimates and predictors from the GLaMOR Project. J Glob Health 2019 Dec;9(2):020421. [doi: 10.7189/jogh.09.020421] [Medline: 31673337]
3. Li L, Liu Y, Wu P, et al. Influenza-associated excess respiratory mortality in China, 2010-15: a population-based study. Lancet Public Health 2019 Sep;4(9):e473-e481. [doi: 10.1016/S2468-2667(19)30163-X] [Medline: 31493844]
4. Wu H, Xue M, Wu C, et al. Estimation of influenza incidence and analysis of epidemic characteristics from 2009 to 2022 in Zhejiang Province, China. Front Public Health 2023;11:1154944. [doi: 10.3389/fpubh.2023.1154944] [Medline: 37427270]
5. National Immunization Advisory Committee (NIAC) Technical Working Group (TWG), Influenza Vaccination TWG. Technical guidelines for seasonal influenza vaccination in China (2023-2024). Zhonghua Liu Xing Bing Xue Za Zhi 2023 Oct 10;44(10):1507-1530. [doi: 10.3760/cma.j.cn112338-20230908-00139] [Medline: 37875437]
6. Reed C, Chaves SS, Daily Kirley P, et al. Estimating influenza disease burden from population-based surveillance data in the United States. PLoS One 2015;10(3):e0118369. [doi: 10.1371/journal.pone.0118369] [Medline: 25738736]
7. Zhang X, Shen P, Liu J, et al. Evaluating the effectiveness and cost-effectiveness of free influenza vaccination policy for older adults in Yinzhou, China: study protocol of a real-world analyses. Vaccine (Auckl) 2023 Jul 31;41(34):5045-5052. [doi: 10.1016/j.vaccine.2023.06.087] [Medline: 37419850]
8. Beyer WEP, McElhaney J, Smith DJ, Monto AS, Nguyen-Van-Tam JS, Osterhaus ADME. Cochrane re-arranged: support for policies to vaccinate elderly people against influenza. Vaccine (Auckl) 2013 Dec 5;31(50):6030-6033. [doi: 10.1016/j.vaccine.2013.09.063] [Medline: 24095882]
9. Zhou L, Su Q, Xu Z, et al. Seasonal influenza vaccination coverage rate of target groups in selected cities and provinces in China by season (2009/10 to 2011/12). PLoS One 2013;8(9):e73724. [doi: 10.1371/journal.pone.0073724]
10. Fan J, Cong S, Wang N, et al. Influenza vaccination rate and its association with chronic diseases in China: results of a national cross-sectional study. Vaccine (Auckl) 2020 Mar;38(11):2503-2511. [doi: 10.1016/j.vaccine.2020.01.093]
11. Zhao HT, Peng ZB, Ni ZL, et al. Investigation on influenza vaccination policy and vaccination situation during the influenza seasons of 2020-2021 and 2021-2022 in China. Zhonghua Yu Fang Yi Xue Za Zhi 2022 Nov 6;56(11):1560-1564. [doi: 10.3760/cma.j.cn112150-20220810-00802] [Medline: 36372744]
12. Gao Z. Application of influenza vaccine in China. Occup Health 2021;37(15):2150-2153. [doi: 10.13329/j.cnki.zyyjk.2021.0519]
13. Peng ZB, Wang DY, Yang J, et al. Current situation and related policies on the implementation and promotion of influenza vaccination, in China. Zhonghua Liu Xing Bing Xue Za Zhi 2018 Aug 10;39(8):1045-1050. [doi: 10.3760/cma.j.issn.0254-6450.2018.08.007] [Medline: 30180426]

XSL•FO
RenderX

14. Jiang SQ, Cai YW, Zuo R, et al. Analysis of influenza vaccination coverage, recommendation behaviors and related factors among health care workers in Nanshan district of Shenzhen city under the free policy between 2019 and 2020. Zhonghua Yu Fang Yi Xue Za Zhi 2022 Nov 6;56(11):1565-1570. [doi: 10.3760/cma.j.cn112150-20211217-01164] [Medline: 36372745]
15. Li P, Liu Z, Shan R, et al. Evolution and regional differences in the supportive environment for influenza vaccination among the elderly population in China. Zhonghua Yu Fang Yi Xue Za Zhi 2023 Dec 6;57(12):2064-2067. [doi: 10.3760/cma.j.cn112150-20230613-00463] [Medline: 38186157]
16. Che X, Liu Y, Gu W, et al. Analysis on the intention and influencing factors of free influenza vaccination among the elderly people aged 70 and above in Hangzhou in 2022. Front Public Health 2022;10:1052500. [doi: 10.3389/fpubh.2022.1052500] [Medline: 36684888]
17. Zhang WY, Li RP, Yu Y, Zhang YJ, Sun HF, Xu AQ. Analysis of adult vaccination in Shandong Province from 2018 to 2022. Zhonghua Yu Fang Yi Xue Za Zhi 2023 Dec 6;57(12):2021-2028. [doi: 10.3760/cma.j.cn112150-20230607-00444] [Medline: 38186151]
18. Zhang Z, Lao X, Yi B, et al. Community intervention research of influenza vaccination for elderly people over 60 years old in Ningbo. Chin J Public Health Manag 2018;34(1):20-23. [doi: 10.19568/j.cnki.23-1318.2018.01.006]
19. Jin Y, Zheng X, Qiu C, et al. Coverage levels of influenza vaccine and 23-valent pneumococcal polysaccharide vaccine among a ≥60-year-old population of Taizhou city after implementation of a free vaccination policy. Chin J Vaccines Immun 2022;28(2):204-207. [doi: 10.19914/j.CJVI.2022040]
20. Government work report 2024. : The People's Government of Zhejiang Province; 2024 URL: https://www.gov.cn/yaowen/liebiao/202403/content_6939153.htm [accessed 2025-09-09]
21. Wang D, Sun H, Chen Y, et al. Willingness to receive influenza vaccine, 23-valent pneumococcal polysaccharide vaccine, and herpes zoster vaccine among residents over 50 years old: a cross-sectional study in Dongcheng. Chin J Vaccines Immun 2024;30(2):141-148. [doi: 10.19914/j.CJVI.2024021]
22. Lu J, Lu XY, Shen YT, et al. Analysis on vaccination willingness and related factors of influenza, pneumonia, and herpes zoster vaccine among people over 50 years old in Minhang district of Shanghai. Zhonghua Yu Fang Yi Xue Za Zhi 2022 Feb 6;56(2):125-131. [doi: 10.3760/cma.j.cn112150-20210423-00406] [Medline: 35184439]
23. Xianxian Y, Xianbin D, Wenge T, Cheng T, Yingshu HE, Liling C. Willingness of influenza vaccination among elderly residents over 65 years old in Chongqing Municipality and its influence factors analysis. Chongqing Med 2021;50(8):1389-1393. [doi: 10.3969/j.issn.1671-8348.2021.08.028]
24. Yan J, Liu Y, Wang Z. Willingness to receive a free influenza vaccine and its influencing factors among the elderly in Kaifu District of Changsha City, 2022. Pract Prev Med 2024;31(1):49-52. [doi: 10.3969/j.issn.1006-3110.2024.01.013]
25. Xu YW, Ding LL, Ye LX, Sun YX, Wang YD, Wang SY. Analysis of the status and related factors of vaccination against respiratory diseases among elderly people aged ≥60 in Zhejiang Province. Zhonghua Yu Fang Yi Xue Za Zhi 2024 Sep 6;58(9):1360-1366. [doi: 10.3760/cma.j.cn112150-20240115-00053] [Medline: 39290017]
26. Kong G, Lim NA, Chin YH, Ng YPM, Amin Z. Effect of COVID-19 pandemic on influenza vaccination intention: a meta-analysis and systematic review. Vaccines (Basel) 2022 Apr 13;10(4):606. [doi: 10.3390/vaccines10040606] [Medline: 35455354]
27. Thomas RE, Lorenzetti DL. Interventions to increase influenza vaccination rates of those 60 years and older in the community. Cochrane Database Syst Rev 2018 May 30;5(5):CD005188. [doi: 10.1002/14651858.CD005188.pub4] [Medline: 29845606]
28. Domnich A, Cambiaggi M, Vasco A, et al. Attitudes and beliefs on influenza vaccination during the COVID-19 pandemic: results from a representative Italian survey. Vaccines (Basel) 2020 Nov 30;8(4):711. [doi: 10.3390/vaccines8040711] [Medline: 33266212]
29. Sturm L, Kasting ML, Head KJ, Hartsock JA, Zimet GD. Influenza vaccination in the time of COVID-19: a national U.S. survey of adults. Vaccine (Auckl) 2021 Apr;39(14):1921-1928. [doi: 10.1016/j.vaccine.2021.03.003]
30. Kou Y, Zhang N. Implicit theories of health predict influenza vaccination intention among elder Chinese: the mediating role of anticipated regret. J Health Psychol 2024 May 27;29(3):267-278. [doi: 10.1177/13591053241253067] [Medline: 38801106]
31. Zixin L, Shenyu W, Xiaotong Y, et al. Coverage and influencing factors of influenza and pneumococcal vaccination among urban and rural residents 50 years old and over in Zhejiang province,2022-a cross-sectional survey. Chin J Public Health 2024;40(1):44-49. [doi: 10.11847/zgggws1142709]
32. Xiaoxiao Z, Kaichao Y, Binghui DU, Yanfang JI, Lubin S, Jun LI, et al. Current status of influenza vaccination among persons aged≥60 years in Henan Province,2020-2022. Mod Dis Control Prev 2024;35(3):166-168. [doi: 10.13515/j.cnki.hnjpm.1006-8414.2024.03.004]
33. Ye N, Lingmei Z. Investigation and analysis on willingness of influenza vaccination and influencing factors of the elderly aged 65 and above in Renshou County in 2021. Parasitoses Infect Dis 2023;21(3):152-157 [FREE Full text]
34. Kim L, Kim GS, Lee KH, Kim J. Factors associated with airline cabin crews' influenza vaccination intention in the workplace-based setting. Workplace Health Saf 2023 May;71(5):245-254. [doi: 10.1177/21650799221151153] [Medline: 36726296]

XSL•FO
RenderX

35. Chen C, Liu X, Yan D, et al. Global influenza vaccination rates and factors associated with influenza vaccination. Int J Infect Dis 2022 Dec;125:153-163. [doi: 10.1016/j.ijid.2022.10.038] [Medline: 36328290]
36. Sun G, Zhang L, Qiu Y, et al. Changes of influenza vaccination rate and associated influencing factors after the COVID-19 pandemic in Shanghai, China. Hum Vaccin Immunother 2024 Dec 31;20(1):2287294. [doi: 10.1080/21645515.2023.2287294] [Medline: 38299510]
37. Ahmed WS, Abu Farha R, Halboup AM, et al. Knowledge, attitudes, perceptions, and practice toward seasonal influenza and its vaccine: a cross-sectional study from a country of conflict. Front Public Health 2023;11:1030391. [doi: 10.3389/fpubh.2023.1030391] [Medline: 36860400]
38. Li X. Short videos as new tools for social participation among the elderly. China Acad J Electron Publ House 2022;3:12-18. [doi: 10.28907/n.cnki.nxwcb.2021.003466]
39. Xu L, Zhao J, Peng Z, et al. An exploratory study of influenza vaccination coverage in healthcare workers in a Western Chinese City, 2018-2019: improving target population coverage based on policy interventions. Vaccines (Basel) 2020 Feb 19;8(1):101. [doi: 10.3390/vaccines8010092] [Medline: 32093085]
40. Fan J, Xu S, Liu Y, et al. Influenza vaccination rates among healthcare workers: a systematic review and meta-analysis investigating influencing factors. Front Public Health 2023;11:1295464. [doi: 10.3389/fpubh.2023.1295464] [Medline: 38026311]

Abbreviations

OR: odds ratio

Edited by A Bamgboje-Ayodele; submitted 14.03.25; peer-reviewed by W Du, Y Li, Y Jiang; revised version received 09.08.25; accepted 11.08.25; published 15.09.25.

Please cite as:
Zhao Y, Xu J, Zhang X, Xu Y, Yan X, Ren S, Wang L
Impact of a Free Influenza Vaccination Policy on Older Adults in Zhejiang, China: Cross-Sectional Survey of Vaccination Willingness and Determinants
JMIR Hum Factors 2025;12:e73940
URL: https://humanfactors.jmir.org/2025/1/e73940
doi:10.2196/73940

© Yusui Zhao, Jinhang Xu, Xuehai Zhang, Yue Xu, Xiaotong Yan, Shaofan Ren, Lixin Wang. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 15.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Needs and Expectations for the myNewWay Blended Digital and Face-to-Face Psychotherapy Model of Care for Depression and Anxiety (Part 2): Participatory Design Study including Mental Health Professionals

Rosemaree Kathleen Miller[1*], PhD; Kathleen O'Moore[1*], PsyD; Katarina Kikas[1,2], BPsychSc; Julie-Anne Therese Matheson[1], BSc; Alexis Estelle Whitton[1,3], PhD; Peter Baldwin[1], PhD; Sophie Li[1], PhD; Melissa Black[1], PhD; Laura Kampel[1], PhD; Nicole Cockayne[1,3], MPH; Fiona Tuttlebee[1], MPH; Caitlin Fraser[1], BBus; Victoria Carr[1], BVA; Kathleen Varghese[1], BFA; Jill Maree Newby[1,2], PhD

[1]Black Dog Institute, Randwick, NSW, Australia

[2]School of Psychology, UNSW Sydney, High Street, Kensington, NSW, Australia

[3]Faculty of Medicine & Health, UNSW Sydney, Kensington, NSW, Australia

[*]these authors contributed equally

Corresponding Author:
Jill Maree Newby, PhD
Black Dog Institute, Randwick, NSW, Australia

Related Article:

Companion article: https://humanfactors.jmir.org/2025/1/e69499

Abstract

Background: In blended care, digital mental health interventions (DMHIs) integrate with face-to-face psychotherapy provided in person or via telehealth. To incorporate DMHIs into routine care for depression and anxiety, it is important to understand the needs and expectations of mental health professionals for blended DMHIs.

Objective: The study objective was to partner with Australian mental health professionals in the design of a transdiagnostic, cognitive behavioral therapy–based blended model of care for adults experiencing depression and anxiety.

Methods: Participants were Australian health professionals who treat adults with depression and anxiety. The participatory design process included a web-based survey (N=258), one-on-one interviews (N=14), and a 2-part focus group (N=6). Quantitative and qualitative data were collected through the web-based survey. In-depth qualitative feedback from interviews and the 2-part focus group was subjected to reflexive thematic analysis.

Results: Mental health professionals found blended care with face-to-face therapy more acceptable than telehealth and blended care with telehealth, with standalone DMHIs being the least preferred option. The most common ways in which mental health professionals thought a DMHI could integrate with face-to-face psychotherapy included homework completion (129/178, 72.5%), skills practice to support in-session therapy (128/178, 71.9%), and psychoeducation (127/178, 71.3%). Mental health professionals expect the blended DMHI to be easy to use, flexible, protective of client data, and to include evidence-based content from several therapeutic modalities (eg, cognitive behavioral therapy and mindfulness). Other preferences included mental health professionals being able to prescribe specific program modules to their clients, track the treatment progress of clients, and receive alerts if their clients' symptoms worsened. In terms of implementation, mental health professionals were concerned about the time and effort needed to use blended care. They suggested that ongoing training and support would help mental health professionals implement blended care with their clients. Monitoring client risk and progress via a web-based dashboard and downloadable summaries was also important.

Conclusions: Designing DMHIs that support psychotherapy for adults with depression and anxiety has the potential to increase access to evidence-based treatment. Involving mental health professionals in DMHI design is expected to increase their acceptance of DMHIs and facilitate the integration of these digital products into routine care.

(JMIR Hum Factors 2025;12:e68789) doi:10.2196/68789

XSL•FO
RenderX

KEYWORDS

depression; anxiety; smartphone; digital intervention; blended care; participatory design

Introduction

Background

The global impact of mental disorders continues to grow annually, with significant health, social, and economic consequences [1,2]. Shortages in the mental health workforce and funding for services also impact access to evidence-based care worldwide [3]. Depression and anxiety are two of the most prevalent mental disorders globally, and psychotherapies are currently recommended as first-line treatments for both disorders, either alone or in combination with pharmacotherapies [4,5]. Cognitive behavioral therapy (CBT) has the strongest evidence base for reducing depressive and anxious symptoms, with recent research confirming moderate to large effects on depressive disorders across more than 400 randomized controlled trials (RCTs) [6] and large effects on anxiety symptoms across 66 RCTs [7]. Transdiagnostic psychotherapies are effective for depression and anxiety and are increasingly favored by the clinical community, as these psychotherapies target the shared symptoms, causes, and maintaining factors of co-occurring disorders, such as depression and anxiety [8-11].

Digital Mental Health Interventions

Unfortunately, most adults who meet diagnostic criteria for depression and anxiety do not access treatment [12,13]. Digital mental health interventions (DMHIs) help to overcome several barriers to accessing treatment (eg, waiting lists, high out-of-pocket costs, and lack of access to skilled mental health practitioners) [3,14]. DMHIs typically include psychotherapy-based content and may be especially beneficial for people who cannot access psychotherapy delivered by a mental health professional. Originally, DMHIs were delivered by computer (eg, CD-ROM) [15,16], but they have now evolved to the point where DMHIs can include artificial intelligence [17]. Today, DMHIs primarily consist of mobile apps and internet-based programs delivered on digital devices, such as smartphones, tablets, laptops, and desktop computers. Self-guided DMHIs enable individuals to self-manage depressive and anxious symptoms with evidence-based skills (eg, CBT) via automated, self-paced modules. Guided DMHIs work similarly but are used with remote support from a health care professional or a paraprofessional, such as a trained coach. Several meta-analyses [16] show that self-guided and guided DMHIs, particularly internet CBT-based programs, significantly improve symptoms of depression (83 studies; Hedge g=0.52, 95% CI 0.43 - 0.60 [15]) and anxiety (47 studies, Hedge g=0.80, 95% CI 0.68 - 0.93) [18]. Early evidence suggests that mobile DMHIs could also be used to effectively treat depression and anxiety, albeit with smaller effect sizes than digital CBT delivered via the internet or computer software [19].

Despite decades of research evidence, the uptake of DMHIs into practice is limited due to high dropout rates and poor user engagement of clients, especially in self-guided formats [20,21]. Acceptance of DMHIs by mental health professionals is another important factor influencing their inclusion in routine clinical care, as health professional endorsement can impact whether a client is willing to use a DMHI [20] and if DMHIs are considered a suitable option for referral [22]. Issues that impact DMHI acceptability by mental health professionals include concerns about program credibility [23,24], DMHI accessibility for people with digital, writing, or reading difficulties [22,25], and the appropriateness of DMHIs for people with severe mental health difficulties [26,27]. Moreover, safety concerns about DMHIs are linked to low acceptability and a reduced likelihood that health professionals will incorporate these products into routine practice [14,26,27]. For example, a qualitative study with Australian psychologists (N=10) [26] found that they preferred face-to-face psychotherapy over standalone DMHIs and were concerned that DMHIs increased the risk of severe symptoms being missed (eg, suicidality) and lacked personalization and human connection, especially self-guided DMHIs.

Combining DMHIs with face-to-face psychotherapy (blended care) can address the limitations that affect DMHI acceptability when integrated into routine clinical care [23,25,26]. Blended care extends upon guided DMHIs by using the DMHI to facilitate in-person or telehealth assessment and treatment led by a mental health professional. The DMHI may be combined with psychotherapy within a treatment plan or supplement psychotherapy as an adjunct, either in-session or between sessions. Much of the existing research on blended care has focused on models of care that support mental health professionals who deliver psychotherapy, such as therapists and psychologists. Blended care can be defined or implemented in several ways; however, one commonality between blended approaches is using DMHIs to enhance psychotherapy and ensure that DMHIs are used as effectively as possible. Examples include reinforcing psychoeducation, symptom tracking, or helping clients to practice skills learned in face-to-face therapy sessions. Findings from Germany, the Netherlands, and Denmark indicate that blended care was acceptable to CBT-based therapists and psychologists before and during the COVID-19 pandemic [28-31]. Research conducted before the COVID-19 pandemic also found that blended care is more acceptable to health professionals and care providers than standalone internet treatment for depression across all levels of severity [32].

Evidence for Blended Care

To date, most research on blended care, especially for depression, was conducted in Europe before the COVID-19 pandemic. Blended care has been tested in several RCTs and pilot trials across Europe in high-income countries, such as Germany, Denmark, Sweden, and the Netherlands, with promising results [25,33-37]. European research also shows that local mental health professionals value blended DMHIs that are flexible to use [38,39], can be customized to suit client needs [28,31,40], and help clients with symptom tracking and homework completion between therapy sessions [28,38]. Adequate training and minimizing the time burden of DMHI use with clients is also important to European mental health

XSL•FO
RenderX

professionals [28,31,38-40]. The European research suggests that blended care is scalable and could enhance mental health professionals' acceptance of DMHI use in routine care, as well as improve client engagement by increasing DMHI uptake and adherence. However, both before and during the COVID-19 pandemic, mixed results were found for blended RCTs outside of Europe in other high-income countries, such as the United States, Japan, and Australia [25,33,41-43]. There is a need to investigate whether the European findings generalize to blended care in non-European countries.

Access to mental health treatment varies worldwide [3], and, by extension, a blended care approach must be tailored to the specific mental health system of a country to function optimally. In the case of Australia, most people with depression and anxiety are provided a mental health treatment plan by their general practitioner, who then refers the person to other health professionals as needed [44]. An Australian with depression or anxiety may receive support from psychiatrists, social workers, or occupational therapists; however, psychotherapy is usually delivered by a psychologist in the private mental health system, with government payments providing partial rebates for up to 10 therapy sessions per calendar year [44,45]. Australian psychologists also currently prioritize CBT due to this therapy being recommended as first-line treatment for depression and anxiety [4,5]. In principle, blended care could work in Australia, as several research studies show the combination of DMHIs with psychotherapy sessions or health professional support may improve symptoms of depression [46-48], anxiety [48], panic disorder [49,50], comorbid depression and alcohol use [51], and hoarding disorder [52]. However, several of these studies have small sample sizes [46,48,52] or no control group for the condition that includes a DMHI combined with psychotherapy or health professional support [47,49,50].

Participatory approaches, including co-design, involve partnering with the intended target audience during the creation of interventions to identify potential issues early and ensure the intervention design reflects their preferences [53,54]. This type of research has been extensively used in DMHI research [54]; however, to date, only 1 participatory [39] design study from Germany has examined the needs and expectations of mental health professionals for blended care using participatory principles. Behr and colleagues [39] consulted with psychotherapists and people with lived experience of mental illness during the COVID-19 pandemic to design a transdiagnostic and transtheoretical internet-based DMHI for blended care. Many of the preferences expressed by psychotherapists aligned with other European blended care studies. However, Behr and colleagues [39] also found that psychotherapists wanted the blended DMHI to protect client data and privacy, contain scientifically sound content, and facilitate the therapeutic relationship with clients. More participatory research on blended care is needed to (1) optimize how the DMHI is deployed to support psychotherapy and (2) ensure the blended care model is adapted to the specific country's mental health system.

Rationale, Objectives, and Aims

Prior research on blended care may have used DMHIs that were not designed to support psychotherapy (ie, they are self-guided and do not integrate with psychotherapy). This puts increased pressure on mental health professionals, and the example of Behr and colleagues [39], along with other European research on blended care [38,40], shows that participatory research is needed to ensure that blended DMHIs align with the preferences of mental health professionals and integrate into existing treatment pathways. Moreover, differences between mental health treatment systems could have contributed to mixed results in blended care trials conducted outside Europe [25,33,41-43]. In addition, even in a high-income country like Australia, there is an estimated 32% shortfall in the mental health workforce [55], further highlighting the need to adapt blended care models to specific mental health systems. This study is part of a larger project to create a transdiagnostic DMHI for Australians with depression and anxiety that could be used self-guided or integrated into routine clinical care. Because most Australians with depression and anxiety are treated in the private mental health system [44], it is currently unknown (1) if blended care would be acceptable to Australian mental health professionals, (2) what their preferences for blended care are, and (3) if blended care would work with the current government reimbursement model for psychotherapy sessions.

To our knowledge, this will be the first Australian study to use participatory design principles to design a new CBT-based blended care model for depression and anxiety with mental health professional input. Phase 1 involves 2 multistage participatory design studies that include web-based surveys, qualitative interviews, and a 2-part focus group with mental health professionals and Australian adults with lived and living experiences of depression and anxiety. Results from the participatory design study with people with lived experience are reported separately in Part 1 of this manuscript series [56]. For mental health professionals, our main aim was to determine their needs and expectations for a DMHI to be used with their clients in blended care. The present findings will inform the creation of a DMHI that addresses clients' needs [56], is acceptable to mental health professionals, and can be integrated into routine treatment for depression and anxiety as part of a CBT-based blended model of care. A secondary aim was to identify factors affecting the training, support, and resources that mental health professionals would need to implement the DMHI into their clinical practice.

Methods

Participatory Design Approach

A mixed methods convergent parallel design [57] was used, which included a web-based survey, qualitative interviews, and a 2-part focus group with a convenience sample of mental health professionals. Throughout data collection, participants were encouraged to discuss any DMHIs they used in practice, not just those created by the Black Dog Institute. Stage 1 involved a web-based survey designed to capture the perspectives of a diverse range of Australian health professionals who provide mental health care to adults with depression or anxiety in their

XSL•FO
RenderX

clinical practice (eg, general practitioner, mental health nurse, social worker, psychologist, and psychiatrist). The web-based survey was created in line with the Checklist for Reporting Results of Internet E-Surveys [58]. In stage 2, a subset of mental health professionals participated in one-on-one web-based qualitative interviews to collect more focused data to address the study aims and conclusions [57]. Finally, in stage 3, a small sample of psychologists was recruited for a 2-part focus group to provide more targeted feedback on implementing the new blended care model. Five psychologists external to the research team served as health professional advisors throughout stages 1-3. The health professional advisors were consulted regularly throughout each stage to confirm and double-check the findings, and all 5 participated in a 2-part focus group for the study. To further mitigate potential bias, the core research team also worked with a diverse range of nonresearchers, including user experience, IT development, innovation, and service delivery professionals before, during, and after data collection. Questions for the web-based survey, interview, and focus group were developed via collaboration between the core research team, health professional advisors, and user experience experts.

Ethical Considerations

Ethical approval was received from the University of New South Wales Human Research Ethics Committee (HC200541). All mental health professionals provided informed consent for the use of their data for research purposes. If participants wished to withdraw their consent, they could contact the research team. All participant data from the web-based survey, interviews, and focus group sessions were deidentified, recoded, and stored in a password-protected file only accessible to the research team. Participants could enter a prize draw to win an Aus $100 (US $65) online gift card for completing the web-based survey. A prize draw was selected as the reimbursement method to minimize the likelihood that a participant would complete the web-based survey solely for compensation. Survey respondents were randomly invited to complete one-on-one interviews and were reimbursed with an Aus $30 (US $20) online gift card. Due to the high level of group interactivity required for a 2-part focus group compared to the interviews, participants were reimbursed with an Aus $200 (US $130) online gift card for attending both 1-hour sessions.

Stage 1: Web-Based Survey

The web-based open survey took approximately 20 minutes and was created with Qualtrics software (version 2020; SAP Inc). Eligibility criteria for health professionals included (1) that they provide mental health care to adults with depression or anxiety, (2) were 18 years of age or older, (3) were fluent in English and (4) were currently living in Australia. The web-based survey was advertised between October and December 2020 via the Black Dog Institute website, the Black Dog Institute health professional network, social media (eg, Facebook [Meta Platforms Inc] and Instagram [Meta Platforms Inc]), existing professional organizations, and peak clinical bodies for mental health professionals (eg, the Australian Psychological Society). To participate, health professionals were directed to the Black Dog Institute website to access a web link for the Qualtrics (2020) survey. The web-based survey began with a participant information sheet and consent form, which included details on the purpose of the research study, eligibility criteria, reimbursement, participant data storage, and the estimated time required to complete the survey. The web-based survey included questions on participants' demographics, clinical practice details, past and current use of DMHIs, and their needs and expectations for a blended DMHI (refer to Multimedia Appendix 1). The web-based survey also included 2 questions from Topooco et al [32] focused on the challenges health care professionals and clients face when integrating digital technology into psychotherapeutic treatment for depression and anxiety (also refer to Multimedia Appendix 1).

Stage 2: Interviews

A subset of survey respondents was randomly invited via email to participate in one-on-one interviews between November 2020 and March 2021. Interviews were conducted until data saturation was achieved. The interviews explored mental health professionals' attitudes toward DMHIs and their views on blending these programs into routine clinical care. Apart from their participation in the web-based survey, there were no pre-existing relationships between participants and the researchers. Interviews were run online using Zoom videoconferencing software (Zoom Communications Inc), and lasted, on average, 53 (SD 12) minutes. The interviews were semistructured and followed a script developed by the research team based on relevant literature and findings from the web-based survey. No personal details about the researchers were disclosed to participants apart from their involvement in the overall project. Interviews were facilitated by either KO or FT. Open-ended questions were used to prompt discussion during the interview (refer to Multimedia Appendix 1). All interviews were audio-recorded and later transcribed by KK.

Stage 3: Focus Group

A 2-part focus group was conducted online using Zoom videoconferencing software, with each session lasting approximately 1 hour. The eligibility criteria were the same as those for the web-based survey. However, only psychologists were included, as they were the most common type of mental health professional recruited for the web-based survey and interviews. A 2-part focus group was advertised via email in July 2021, and the 2 sessions were run in July and August 2021, respectively. A convenience sample of 5 health professional advisors and 2 psychologists from the Black Dog Institute psychology clinic was recruited. The researchers were familiar with 7 participants because they worked at the same workplace, but there were no pre-existing personal relationships. One psychologist's data was excluded from data analysis as they could only attend one of the two focus group sessions. To minimize bias, the sessions were designed by service delivery and user experience professionals without clinical training (FT, CF, and VC) with input from the research team. Session 1 was led by CF, while Session 2 was led by KO, FT, and VC. KK attended both sessions as the note-taker. The focus group script (refer to Multimedia Appendix 1) included questions on factors affecting the implementation of the new blended care model, such as how the psychologists would use blended care with therapy sessions, the training and support needed to implement

XSL•FO
RenderX

blended care, and the features to include in an online program dashboard designed for health care professionals as part of the blended care system. Both sessions were audio-recorded and later transcribed by author KO.

Reflexivity Statement

To reduce potential biases, the core group of researchers involved in thematic analysis included 2 clinical psychologists who were also researchers (JMN and KO) and 3 nonclinician researchers (RKM, KK, and JATM). 2 service design professionals external to the research team assisted with qualitative data collection (FT and CF). 4 members of the core research team had prior experience with reflexive thematic analysis (JMN, KO, RKM, and KK). The 2 clinicians (JMN and KO) had experience in psychotherapy and digital mental health treatment for individuals with clinical depression and anxiety. All 3 nonclinician researchers (RKM, KK, and JATM) had experience with digital mental health research. One nonclinician researcher (RKM) also had experience working with staff and health professionals from substance use and mental health services. All coauthors reviewed the final coding framework for the thematic analysis conducted for the interviews and the 2-part focus group. This included 4 clinical psychologists who were also researchers (SL, MB, PB, and LK), a psychologist who was also a researcher (AEW), 2 user experience professionals (VC and KV), 2 service design professionals (FT and CF), and a public health researcher (NC). We acknowledge that our past clinical and research experiences, as well as our familiarity with the literature, may have introduced biases and assumptions that inadvertently influenced the thematic analysis of the interview and focus group transcripts. The core research team minimized this with regular meetings to discuss the data to ensure transparency and check that our themes and subthemes were grounded in the collected data.

Data Analysis

Descriptive statistics were performed with SPSS software (version 26.0; IBM Corp). For the web-based survey results, descriptive statistics were calculated based on the actual number of participants who responded to a specific question. Reflexive thematic analysis was selected to examine qualitative data from the one-on-one interviews and a 2-part focus group, as this method was appropriate for answering the research questions [59]. For the interviews, KK cleaned and checked the transcripts against the original audio after each interview and focus group session. Authors (KK, KO, and RKM) then read the interview transcripts several times and, after familiarization, used a general inductive approach to generate the initial set of codes. These codes were then categorized using Microsoft Excel to create the initial coding framework. The coding framework was revised iteratively by authors KK, RKM, and JATM until consensus on themes and subthemes was achieved. All disagreements were resolved through discussion with the research team (including KO, JMN, and AEW) until consensus was reached. The same approach was followed to analyze the focus group transcripts. The transcripts were read several times by KO and RKM, who generated the initial set of codes and then revised the coding framework iteratively with JMN.

Results

Stage 1: Web-Based Survey

Demographics

Of the 258 participants who provided consent, 157 (60.9%) completed the web-based survey. The average age was 47 (SD 13.74; range 18 - 78) years, and most participants were women (n=200, 77.5%), spoke English as a first language (n=227, 88%), and lived in a major Australian city (n=175, 73.8%; Table 1). Statistical comparisons revealed no significant differences in the demographics of participants who completed the full survey versus those who did not (see Multimedia Appendix 2).

XSL•FO
RenderX

Table . Sample characteristics for the web-based survey.

Demographic characteristics	Participants, n (%)
Gender (N=258)	
Women	200 (77.5)
Men	58 (22.5)
English primary language at home (N=258)	227 (88)
First Nations Australian origin (N=258)	4 (1.6)
Residential location in Australia (N=258)	
Major cities	175 (73.8)
Regional or remote	62 (26.2)
Current role[a] (N=239)	
Registered psychologist	45 (18.8)
General practitioner	43 (18)
Clinical psychologist	37 (15.5)
Social worker	28 (11.7)
Mental health nurse	26 (10.9)
Counselor	20 (8.4)
Trainee psychologist	11 (4.6)
Psychiatrist	5 (2.1)
Other current role[b]	36 (15.1)
Current work setting[a] (N=239)	
Private practice	119 (49.8)
Not-for-profit organization	49 (20.5)
Outpatient	38 (15.9)
Inpatient	28 (11.7)
University	25 (10.5)
Other work setting[c]	29 (12.1)

[a]Percentages do not add up to 100% as respondents were allowed multiple responses.

[b]Other current roles included occupational therapist, peer support roles, and other types of nurses (apart from mental health).

[c]Other current work settings included government, hospitals, justice settings, and schools.

Practice Details

A total of 239 participants provided data on their current work role and work setting. The most common work roles were registered psychologists (n=45, 18.8%), general practitioners (n=43, 18%), and clinical psychologists (n=37, 15.5%; Table 1). The percentage of different work roles and work settings was largely consistent among participants who did or did not complete the full survey (refer to Multimedia Appendix 2). Participants mostly worked in private practice (n=119, 49.8%), not-for-profit organizations (n=49, 20.5%), and outpatient settings (n=38, 15.9%; Table 1). Across 239 participants, most (n=127, 53.1%) had 10 or more years of experience as mental health professionals, with an average of 15 (SD 11.89; range 0-49) years. Table 2 shows that the most common mental health conditions treated by participants were anxiety (199/221, 90%), depression (190/221, 86.1%), and trauma and stress-related disorders (152/221, 68.8%). The most common psychotherapies used to treat depression and anxiety were CBT (178/216, 82.4%), acceptance and commitment therapy (116/216, 53.7%), and mindfulness-based cognitive therapy (115/216, 53.2%; Table 2).

XSL•FO
RenderX

Table . Counts and frequency percentages of the most prevalent mental health issues treated by mental health professionals and the most used treatment modalities when working with adults experiencing depression and anxiety.

Practice details	Participants, n (%)
Most prevalent mental health issues[a] (N=221)	
Anxiety disorders	199 (90)
Depressive disorders	190 (86)
Trauma- and stressor-related disorders	152 (68.8)
Personality disorders	97 (43.9)
Substance-related and addictive disorders	80 (36.2)
Bipolar and related disorders	80 (36.2)
Schizophrenia spectrum and other psychotic disorders	55 (24.9)
Obsessive-compulsive and related disorders	51 (23.1)
Feeding and eating disorders	35 (15.8)
Most frequent treatment modalities[a] (N=216)	
Cognitive behavioral therapy	178 (82.4)
Acceptance and commitment therapy	116 (53.7)
Mindfulness-based cognitive therapy	115 (53.2)
Solution focused and brief therapy	82 (38)
Dialectical behavior therapy	64 (29.6)
Interpersonal psychotherapy	39 (18.1)
Schema therapy	38 (17.6)
Family therapy	26 (12)
Humanistic approach	23 (10.6)
Couples therapy	20 (9.3)
Psychodynamic approach	19 (8.8)
Eye movement desensitization and reprocessing	12 (5.6)
Psychoanalysis	11 (5.1)
Gestalt approach	10 (4.6)
Jungian approach	6 (2.8)
Other treatment modalities[b]	20 (9.3)

[a]Percentages do not add up to 100% as respondents were allowed multiple responses.

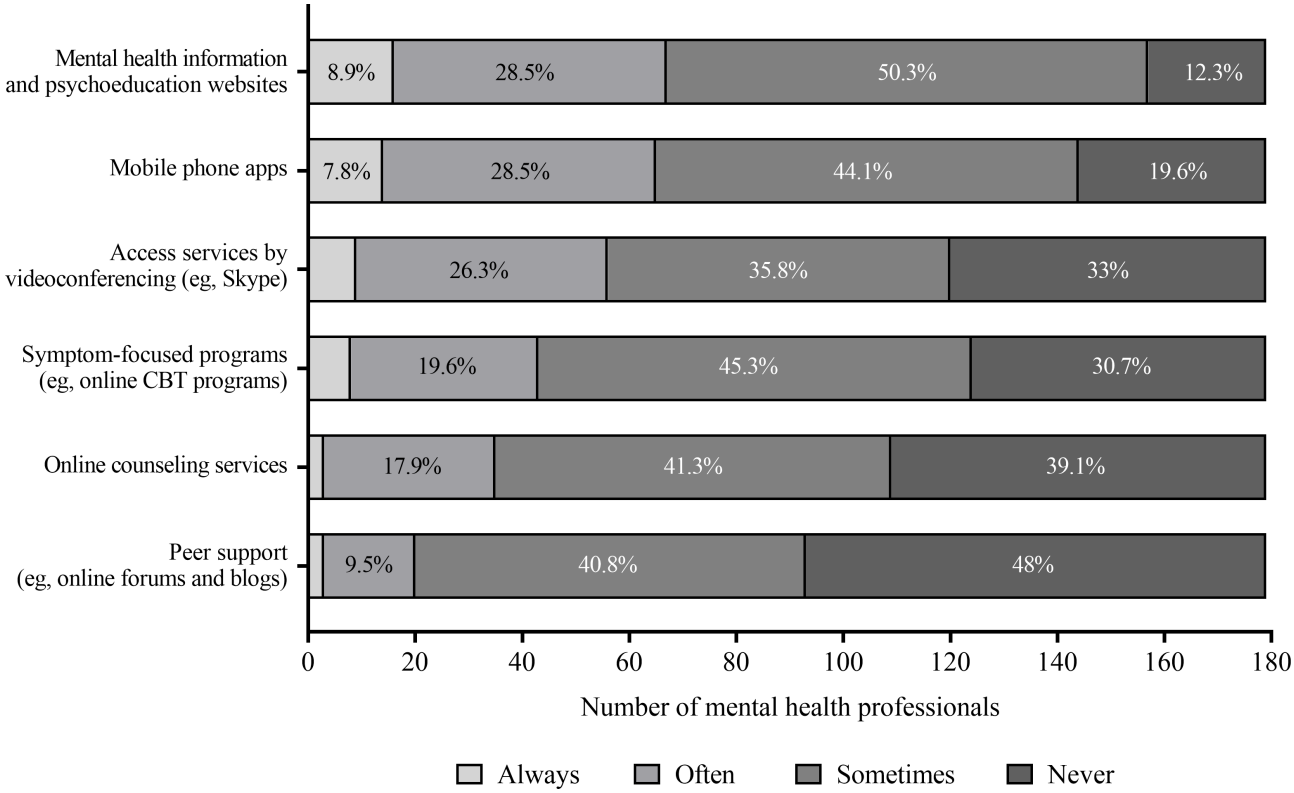
[b]Other common treatment modalities included peer support, physical activity, lifestyle advice, acupuncture, and hypnosis.

Use of Digital Mental Health in Practice

Figure 1 (data labels show the frequency percentage) shows how often particpants recommended different types of DMHIs to clients (see Multimedia Appendix 3 for counts). Participants were most likely to always or often refer clients to mental health information and psychoeducation websites (67/179, 37.4%), mobile phone apps (65/179, 36.3%), and videoconferencing services (eg, Skype [Microsoft], Zoom [Zoom Communications Inc]; 56/179, 31.3%). The 3 most common DMHIs that participants would suggest to clients with depression and anxiety were This Way Up (Clinical Research Unit for Anxiety and Depression at St Vincent's Hospital and the University of New South Wales), Smiling Mind (Smiling Mind Pty Ltd), and Headspace (Headspace Health).

XSL•FO
RenderX

Figure 1. Counts of what type of digital mental health interventions were referred to clients in participants' clinical practice (N=179). CBT: cognitive behavioral therapy.



Is There a Need for Digital Mental Health in Routine Care?

Most participants (135/178, 75.8%) reported a need for an online mental health program that could be integrated into routine care. Among the remaining 43 participants, 9 (5.1%) disagreed and 34 (19.1%) were unsure.

Digital Mental Health Preferences

Routine Care

Table 3 shows that participants preferred integrating DMHIs into routine care for depression and anxiety by referring clients to homework exercises to do in between sessions (129/178, 71.3%), specific modules to support skills learned in therapy (128/178, 71.9%), and psychoeducation (127/178, 76%). Other commonly preferred ways of integrating DMHIs into routine care included using DMHIs to assist clients with relapse prevention (117/178, 65.7%), tracking symptom changes over time (116/178, 65.2%), and supporting clients who had run out of sessions and could not afford to pay (113/178, 63.5%; Table 3).

XSL•FO
RenderX

Table . Counts and frequency percentages for integrating DMHIs into routine care, challenges faced by health care professionals and clients when digital technology is integrated with psychotherapy, and factors affecting whether blended care is suitable for a client.

Digital mental health preferences, challenges, and client factors	Participants, n (%)
Ways the program could integrate into routine care[a] (N=178)	
Refer client to homework exercises	129 (72.5)
Refer client to specific program modules	128 (71.9)
Psychoeducation	127 (71.3)
Relapse prevention once treatment is complete	117 (65.7)
Assessments to track symptoms	116 (65.2)
Support clients who run out of sessions and cannot afford to pay	113 (63.5)
Use in-session to explain or practice skills	107 (60.1)
As a prequel to therapy	88 (49.4)
Challenges for health care professionals[a] (N=158)	
Low motivation because of critical attitude toward digital technology	73 (46.5)
Not skilled in digital technology	72 (45.9)
Difficulties adapting treatment methods to digital technology	70 (44.6)
Low clinical effectiveness	57 (36.3)
Adherence of the health care professional is low	45 (28.7)
Degree of accessibility and cost of internet	41 (26.1)
Use of digital technology is too time-consuming	32 (20.4)
Other challenges	26 (16.6)
Challenges for clients[a] (N=158)	
Difficulties with applying self-help strategies	73 (46.5)
Not skilled in digital technology	72 (45.9)
Low client adherence	70 (44.6)
Low motivation because of critical attitude towards digital technology	57 (36.3)
Degree of accessibility and cost of internet	45 (28.7)
Low clinical effectiveness	41 (26.1)
Use of digital technology is too time-consuming	32 (20.4)
Other challenges	26 (16.6)
Client factors for blended care[a] (N=157)	
Access to a computer, tablet, or phone	128 (81.5)
Online modules relevant to main symptoms or diagnosis of patient	126 (80.3)
Access to the internet	123 (78.3)
Intelligence and cognitive functioning of client	118 (75.2)
Access to a private, safe place to complete program	112 (71.3)
Competence in using digital technology	111 (70.7)
Absence of severe suicidality	109 (69.4)
Absence of psychotic symptoms	100 (63.7)
Absence of acute medical need that hinders client's ability to participate in treatment	94 (59.9)
Absence of severe substance use	82 (52.2)
Other factors	22 (14)

[a]Percentages do not add up to 100% as respondents were allowed multiple responses.

XSL•FO
RenderX

Program Inclusions

Of the 178 participants, 142 (79.8%) described what they wanted to be included in a blended DMHI. Several suggested evidence-based content from common therapeutic modalities, such as CBT, acceptance and commitment therapy, dialectical behavior therapy, and mindfulness. Participants were also keen for the DMHI to support face-to-face therapy, for example, “flexibility to ’build’ your own exercises with clients or tailor them.” Supporting clients with skills practice and symptom monitoring was also important, for example, “validated assessment measures”, and “sleep tracking and thought monitoring.”

Challenges of Digital Mental Health

The most common challenges identified by participants for health care professionals were low motivation due to negative attitudes toward digital technology (73/158, 46.5%), not being skilled in digital technology use (72/158, 45.9%), and difficulties with adapting treatment methods to digital technology (70/158, 44.6%; Table 3). The 3 top challenges identified for clients were difficulties with applying self-help strategies (73/158, 64.6%), not being skilled in digital technology use (72/158, 63.9%), and low client adherence (70/158, 62%; Table 3).

Blended Care Approaches to Treatment

Client Factors

The most important factors that participants thought would affect the suitability of blended care for a client were access to a computer, tablet, or phone (128/157, 81.5%), modules relevant to the client’s main symptoms and diagnosis (126/157, 80.3%), internet access (123/157, 78.3%) and the client’s intelligence and cognitive functioning (118/157, 75.2%; Table 3). Other important client factors included access to a private, safe space to use the DMHI (112/157, 71.3%), and competence in DMHI use (111/157, 70.7%). Most participants considered the absence of severe suicidality (109/157, 69.4%), psychotic symptoms (100/157, 63.4%), acute medical needs affecting the client’s treatment (94/157, 59.9%), and severe substance use (82/157, 52.2%) to be important client factors for blended care as well.

Acceptability

Participants (N=157) rated acceptability with a 100-point scale (0=Not at all, 100=Extremely acceptable) for (1) standalone digital programs, (2) telehealth sessions, (3) blended care with digital programs and face-to-face sessions, and (4) blended care with digital programs and telehealth sessions. Participants found blended care with digital programs and face-to-face sessions the most acceptable treatment option (mean 71.61, SD 25.48), followed by telehealth sessions (mean 68.38, SD 23.72) and blended care with digital programs and telehealth sessions (mean 64.65, SD 25.86). Standalone digital programs were the least acceptable option to participants (mean 38.81, SD 25.86).

Sharing of Client Information

Most participants preferred client information to be shared with them via a secure email account (99/157, 63.1%) or sent to a secure web portal where they could log in and see the information (91/157, 58%). A minority wanted their client’s information shared with them via SMS text messaging to a nominated phone number (15/157, 9.6%) or through another information sharing method (16/157, 10.2%), such as a printable report or via an existing practice management system. Some participants were not sure (14/157, 8.9%) or did not mind (11/157, 7%) how client information was shared with them. The types of client information that participants wanted to share and receive from the blended DMHI included program adherence and completion, symptom tracking with validated and subjective (eg, sleep and mood) measures, and client feedback on treatment (eg, benefits, difficulties, or questions they have). The preferred time interval between receiving client information from the DMHI varied. Most participants preferred a weekly or fortnightly interval, and others wanted to tailor the interval to the client or treatment timeline, for example, “every week to begin with, moving to larger intervals as treatment progresses.”

Stage 2: Interviews

Demographics

A total of 14 mental health professionals participated in the one-on-one interviews. Their mean age was 46 (SD 13; range 26 - 73) years. Half of the participants were female (7/14; 50%), and the other half were male (7/14; 50%). Overall, 11/14 (79%) participants lived in major Australian cities, and 3/14 (21%) lived in regional or remote areas of Australia. The sample consisted of registered psychologists (5/14, 36%), clinical psychologists (5/14, 36%), mental health nurses (3/14, 21%), and a counselor (1/14, 7%). Participants had worked as mental health professionals for an average of 16 (SD 9.95; range 4 - 35) years. Participants worked in private practice (5/14, 36%), outpatient settings (4/14, 29%), universities (4/14, 29%), and not-for-profit organizations (2/14, 14%).

Thematic Analysis Findings

Overview of Findings

Six key themes were identified related to blended care for depression and anxiety: (1) blended care preferences, (2) promote client engagement, (3) prioritize client safety and duty of care, (4) minimize burden on time-poor clinicians, (5) program is credible and trustworthy, and (6) client barriers to using program. Illustrative quotes for each subtheme within the 6 themes are provided in Table 4.

XSL•FO
RenderX

Table . Thematic analysis themes and subthemes with example quotes.

Themes and subthemes	Example quotes
1. Blended Care Preferences	
Supports Skills Practice, Psychoeducation, and Lifestyle Changes	"If we're able to provide them with resources, [like] psychoeducation material...alongside the therapy that we're doing, it increases the support provided to the client." [Participant 71]
Helps Clinician Monitor Treatment Progress	"...I also want to know whether they're finding therapy effective and helpful and what they find helpful and what they don't find helpful." [Participant 14]
Prescribe Program Modules to Clients	"In a perfect world, a way to digitally prescribe it for people to track their progress, and then to be able to facilitate...not just like the way they engage with it, but in session tools." [Participant 15]
Walk Through Program Content in Session	"My preference...is to work through a program with a client, to have some oversight on what they're doing, to review the content with them, to answer any questions that they have." [Participant 8]
Trial the Program from the Client's Perspective	"If I were to give it to somebody or lead someone to it, I would want to do it myself, first of course." [Participant 3]
2. Promote Client Engagement	
User-Friendly and Inviting Interface, With Minimal Text	"I like that it's very simple and easy, a lot of the communication around concepts is done visually and in videos, there's no long slabs of text." [Participant 15]
Flexible with Lots of Variety	"It's important to have lots of different...ways of delivering information. So, whether it be...printed, video, audio and animated or something like that...it kind of keeps your interest and it's kind of interesting and stimulating, you're not just reading realms of stuff." [Participant 12]
Program Content Adjusts to Meet Client Needs	"The other thing that I think would be really, really helpful is if there was some way that you know this online program is personalized...because I think sometimes if you mention a whole lot of stuff, you know, clients find it a little bit trite, you know, like eat well and you know look after yourself. It's not really meaningful to them." [Participant 14]
Interactive Activities and Features	"...a bit interactive with a quiz or a little video or somebody talking about their experience, you know, that's just short and sweet, you know, a three-minute video and practical things." [Participant 24]
Relatable and Encouraging Stories Shared by Real People	"...having stories where people share their story, that can be very powerful." [Participant 48]
3. Prioritize Client Safety and Duty of Care	
Directs Clients to Crisis Support When Needed	"If there [were]...links to crisis services that are available and perhaps, you know, if people are screening as really high, knowing that they're going to get some kind of automated message to remind them to seek support immediately...that's really important." [Participant 9]
Helps Monitor Client Safety	"Hmm, if they are feeling worse...maybe that's the possibility of the clinician being alerted if their scores are really low...then the clinician could do a follow-up call, or like a check-in." [Participant 12]
Supports Maintenance of Professional Boundaries	"...we'll organize to call [the client] or send a text or we'll organize for a crisis service to ring or...and you work out with them what, you know, that's a bit of a negotiation process about what they'd be comfortable with and what you're obliged to do under your duty of care." [Participant 39]
Not Suitable for Clients with Severe Mental Distress	"With the population I work with, no, I mean if I was working with a more severe population, I might be a little bit less willing to use online programs without more support." [Participant 10]
4. Minimize Burden on Time-Poor Clinicians	
Quick Reference Guides and Resources	"So, I think first of all, helping clinicians have a quick and easy way of understanding what the content of a module or whatever it is, so they know if it's appropriate." [Participant 15]

XSL•FO
RenderX

Themes and subthemes	Example quotes
Integrates with Existing Client Management Systems	"I don't know if it's going to come into it, but also not being easily able to refer like say, for example, some of our clinicians they've got electronic medical records and databases already set up like if there was an easy trigger in those systems to kind of alert someone to refer to another service." [Participant 83]
Web-Based Program Dashboard	"If there is a clinician dashboard kind of interface, being able to have access to that, so I can see before the session, 'okay, great, they have completed this task.'" [Participant 8]
5. Program Is Credible and Trustworthy	
Evidence-Based and Reputable	"I'd want it to be, you know, peer-reviewed...and it's gone through all the you know research, efficacy, validity, all that sort of stuff, so I'd want it to have all that as the foundation." [Participant 33]
Secure and Protects Client Privacy	"...maybe a button with 'would you like your clinician to be able to see this?', 'yes or no', something that they can control, so they can say yes, just before the session and then after the session, they can hide it again if they don't want the clinician to have ongoing access." [Participant 10]
6. Client Barriers to Using Program	
Program Cost	"I think once you put a cost on it, you can lose a lot of people." [Participant 24]
Digital Interventions Not Appropriate for Some Clients	"The ability to engage with the app and take on board information that's there with the reading and the content requires a level of cognitive functioning and, you know, self-awareness and conscious thought. If all those are getting scrambled for you, then this isn't going to work." [Participant 62]

Theme 1: Blended Care Preferences

Theme Overview

This theme refers to how mental health professionals envisioned the blended care program would work for clients receiving routine clinical care for depression and anxiety.

Supports Skills Practice, Psychoeducation, and Lifestyle Changes

Participants emphasized the importance of modules in the program supporting skills taught in face-to-face therapy, such as psychoeducation and healthy lifestyle content (eg, sleep hygiene, physical exercise, and nutrition). Including this content would allow clinicians to assign psychoeducation and skills practice activities to clients between sessions, freeing up therapy sessions for more challenging and tailored skill development.

Helps Clinician Monitor Treatment Progress

Several participants wanted to track whether clients found their therapy sessions helpful and monitor the severity of client symptoms across sessions to indicate progress. Participants suggested including measures for client-reported outcomes and experiences in the program (eg, outcome rating scale and session rating scale). Some participants were also interested in viewing which modules and activities clients had completed in the program.

Prescribe Program Modules to Clients

Participants preferred to prescribe specific modules or activities in the program so that their clients were assigned content and activities relevant to their needs or which assisted clients in practicing a skill they were currently learning with their therapist.

Walk Through Program Content in Session

When using a DMHI, participants wanted to guide clients through the content during therapy sessions to help familiarize them with the program and help clients feel more comfortable using it. Participants also thought that demonstrating how the program worked to their clients would address practical barriers, such as client reporting of technical issues.

Trial the Program From the Client's Perspective

Participants preferred to trial the program from the client's perspective before recommending it. This feature would allow participants to verify the program's quality, ease of use, and content before recommending the program to clients.

Theme 2: Promote Client Engagement

Theme Overview

User engagement with the blended care program was an important consideration for interviewees. Theme 2 includes the features, functionality, and content within a DMHI that mental health professionals suggested would enhance client engagement with the program.

User-Friendly and Inviting Interface With Minimal Text

Most participants believed that the program should be easy and intuitive for clients to use and stay engaged with. Some participants suggested that the onboarding experience for clients should be simple and include minimal registration questions. Other suggestions included using clear and friendly language to explain concepts and avoid "long slabs of text." It was also important that the program was designed with an aesthetically pleasing interface (eg, colorful, comic-style graphics and animations).

XSL•FO
RenderX

Flexible With Lots of Variety

Many participants thought clients would stay engaged with the DMHI if it was flexible, modular, and included a range of content, options, and functionality. Participants also believed clients would like to choose which module or activity they wanted to complete rather than having one suggested. Some participants thought that transdiagnostic content would be more engaging for clients than disorder-specific content and recommended including transdiagnostic modules that address multiple concerns common for people with anxiety and depression (eg, perfectionism, rumination, self-esteem, and procrastination).

Program Content Adjusts to Meet Client Needs

Some participants thought content personalized to clients' needs would help them stay engaged with the program. For example, if a client indicated they wanted to improve their relationships with others or manage procrastination, the program would show information or modules relevant to these topics in response.

Interactive Activities and Features

Several participants believed interactive activities and features would enhance client engagement, such as practical activities that require users' active input, such as quizzes and chatbots. Other suggestions included videos, audio tracks, diagramming tools, and activities, such as behavioral activation or exposure stepladders where their client could input content relevant to them.

Relatable and Encouraging Stories Shared by Real People

Some participants believed that incorporating other people's experiences of mental ill health and treatment into the program would enhance client engagement. These participants thought a video would be a good way to include these stories and that the examples should be relevant and encouraging for clients.

Theme 3: Prioritize Client Safety and Duty of Care

Theme Overview

This theme focuses on participant perspectives about the importance of prioritizing and safeguarding client safety while also assisting mental health professionals in setting clear boundaries around their duty of care to clients.

Directs Clients to Crisis Support When Needed

A common concern was ensuring clients could access crisis support if they were feeling unwell or experiencing suicidal ideation while using the DMHI. Most participants wanted contact details of emergency and mental health support services included (eg, Lifeline and Suicide Call Back Line), while others endorsed a more personalized approach where a user input (eg, endorsing suicidal thinking) would trigger messages or risk alerts, such as advising the client to seek help immediately and then directing them to crisis support information.

Helps Monitor Client Safety

Participants wanted the program to help monitor client safety and alert the clinician if their client's symptoms worsened or they endorsed indicators of elevated suicide risk. A common suggestion was to use validated questionnaires to flag if a client's symptoms were worsening. Some participants thought the DMHI could prompt mental health professionals to check in with their clients or reassess their treatment plans.

Supports Maintenance of Professional Boundaries

Participants expressed concern about their increased duty of care between therapy sessions when using a DMHI with clients (eg, being required to contact clients at risk outside of sessions). Some also thought the program could help mental health professionals manage their duty of care and client safety. This included (1) guiding discussions about duty of care with clients, (2) setting boundaries around the mental health professional's availability, and (3) clarifying which client information would be shared with mental health professionals. Participants thought that this could be facilitated through the onboarding content of the DMHI and automatic notifications about elevated client risk to mental health professionals.

Not Suitable for Clients With Severe Mental Distress

Some participants felt a DMHI, even when used with blended care, would not be appropriate for high-risk clients or those experiencing severe distress. Examples included clients who were suicidal, experiencing significant challenges with perceptions of reality (eg, psychosis and personality disorder) and those with severe distress or functional impairment (eg, severe depression).

Theme 4: Minimize Burden on Time-Poor Clinicians

Theme Overview

This theme describes how DMHIs can reduce the burden on mental health professionals who may already be short of time.

Quick Reference Guides and Resources

Many participants thought simple, quick, easy-to-use guides and resources would help them decide whether the DMHI was appropriate for clients. Several participants specified that short overviews of content or program modules, the program structure, and details to pass on to the client (eg, time commitment) would be helpful. The preferred medium for quick reference guides and resources was varied and included written summaries, short videos less than 2 minutes long, booklets, and brochures.

Integrates With Existing Client Management Systems

Some participants wanted the blended DMHI to integrate with existing software and client management systems used in their practices. Examples included sending questionnaires or referring clients to the program digitally via client management systems that mental health professionals already had in place to avoid the need for additional record-keeping and the risk of duplicate records.

Web-Based Program Dashboard

Several participants wanted access to a dashboard to view their clients' progress, including completion of modules or activities and symptom scores. The dashboard could also be used to recommend modules or activities to clients using the program, and it would be helpful to quickly view their clients' data before or during a therapy session with them.

XSL•FO
RenderX

Theme 5: Program is Credible and Trustworthy

Theme Overview

This theme emphasizes the importance of a DMHI being credible, reliable, and safe for clients.

Evidence-Based and Reputable

Most participants would only recommend DMHIs to their clients if the product had been scientifically evaluated or had demonstrated effectiveness and was developed by organizations or individuals they trusted.

Secures and Protects Client Privacy

The DMHI needed to be confidential and protect client data. Participants wanted transparent information about where the client data were stored and who could access it, including in the DMHI. Participants also thought clients should be able to turn clinician access to their data on and off.

Theme 6: Client Barriers to Using the Program

Theme Overview

The final theme refers to barriers that prevent clients from using a DMHI.

Program Cost

Most participants believed cost was a barrier to their clients accessing DMHIs and that their clients would want to check the efficacy and suitability of DMHIs before paying for access. Another perspective was that some individuals with mental illness were struggling financially and that DMHIs should be included in treatment costs. Participants also felt uncomfortable with clients paying for DMHIs, considering the already high costs of therapy. However, some acknowledged that cost could indicate to clients that a program has a high level of quality and credibility. Overall, participants emphasized the importance of keeping the costs of DMHIs low.

Digital Interventions Not Appropriate for Some Clients

Participants believed that DMHIs would not suit clients who lack access to technology (ie, no mobile phone, computer, or internet access), have poor literacy skills, have cognitive impairments or intellectual disabilities, face psychosocial vulnerabilities (ie, domestic violence and child protection concerns), or require the additional empathy and validation that an in-person therapist can offer.

Stage 3: Focus Group

Demographics

In total, 6 psychologists participated in the 2-part focus group. The mean age was 35 (SD 10; range 26‐54) years. All participants spoke English as their primary language and lived in major Australian cities. Most participants were female (n=5, 83%), and 1 (17%) was male. Overall, 5 (83%) participants were clinical psychologists and 1 (17%) was a provisional psychologist. The participants had worked as mental health professionals for an average of 10 (SD 10; range 1‐28) years and worked in private practice (n=4, 67%), not-for-profit organizations (n=4, 67%), and outpatient settings (n=2, 33%). All psychologists nominated depressive and anxiety disorders as two of the most common mental health issues they treated in their clinical practice. Other common mental health issues treated by participants included trauma and stressor-related disorders (n=5, 83%), personality disorders (n=4, 67%), bipolar and related disorders (n=3, 50%), and obsessive-compulsive and related disorders (n=3, 50%). All psychologists commonly use CBT to treat depression and anxiety. Other treatment modalities included acceptance and commitment therapy (n=4, 67%), mindfulness-based cognitive therapy (n=4, 67%), and schema therapy (n=4, 67%).

Thematic Analysis Findings

Overview of Findings

Five key themes were identified that related to the resources, support, and training needed to implement blended care into routine care for depression and anxiety: (1) user-friendly resources, training, and support; (2) practical and flexible training; (3) comprehensive implementation support; (4) monitor client progress and risk; and (5) tailor client content and access.

Theme 1: User-Friendly Resources, Training, and Support

Psychologists emphasized that resources, support, or training for the blended care system should be user-friendly and facilitate the use of blended care with clients. Psychologists also emphasized that any digital components of a blended care system, such as a web-based dashboard or portal, should be simple, quick, and easy to integrate into their regular clinical practice.

Theme 2: Practical and Flexible Training

Psychologists thought training should be ongoing as they used the blended care system and focused on understanding what is included in the program (eg, clinical techniques). Like interviews, the psychologists raised concerns about client confidentiality and how the limitations of duty of care when using DMHIs should be communicated to clients. They suggested that this should be covered in any training provided to mental health professionals for the blended care DMHI. Psychologists also wanted various training resources to be provided, including workshops, written summaries, Frequently Asked Questions, and video-based guides. Receiving professional development credit for an educational module (eg, on blended care) was also considered an important incentive.

Theme 3: Comprehensive Implementation Support

The psychologists believed that one-on-one and peer support would be needed to implement blended care. For one-on-one support, they suggested the DMHI support team should regularly check in with mental health professionals using the blended care system and include “Frequently Asked Question” and “Help” sections in a web-based dashboard to provide contact details on how and who to contact for help. Additionally, they recommended facilitating peer support through access to a clinician chat group or message board.

Theme 4: Monitor Client Progress and Risk

Psychologists wanted to be able to view and track client progress via a web-based dashboard included in a blended care system. Preferred metrics included questionnaire scores, client responses, and the number of activities completed. Flagging and receiving

XSL•FO
RenderX

notifications about at-risk clients based on symptom severity was also essential. Another suggestion was to include a downloadable snapshot of the client's DMHI use and other data that could be exported to and saved in other practice management software.

Theme 5: Tailor Client Content and Access

The psychologists valued the ability to select, schedule, and personalize DMHI content for clients. This included searching for activities with clinical terms (eg, behavioral activation), assigning specific activities to clients, tailoring reminders to suit client needs, and sending content and activity links directly to clients. Having control over the digital "unlinking" of a client during therapy closure or handover to another mental health professional was also important.

Discussion

Principal Findings

This was the first participatory design study to investigate Australian mental health professionals' needs and expectations of a new transdiagnostic, CBT-based blended DMHI that could be integrated into routine care for clients experiencing depression and anxiety. Participants wanted the blended DMHI to be easy to use, flexible to use with clients, and provide a broad range of tailored content to engage and sustain client interest, including evidence-based content from several therapeutic modalities, such as CBT, acceptance and commitment therapy, and dialectical behavior therapy. Participants also provided detailed feedback on the sharing of client data, including a preference for this to occur via a secure email account or web portal, as well as the types of client data that would be most helpful to share. Factors affecting mental health professionals' acceptance of a blended DMHI, as well as the types of training, support, and resources needed to implement blended care, were also identified. Incorporating features that support the monitoring of client safety into DMHIs and ensuring that DMHIs are free or of low cost for clients could enhance the acceptance of blended care by mental health professionals. Overall, our findings indicate that blended care for depression and anxiety would be feasible and acceptable for Australian mental health professionals to implement, especially psychologists.

Blended Care Needs and Expectations

Blended care, where a DMHI is integrated with in-person or telehealth sessions, was considered more acceptable than standalone digital interventions. Many of our findings on the needs, preferences, and expectations of mental health professionals for a DMHI that could facilitate blended care align with results from our participatory design study with people with lived and living experiences of depression and anxiety [56] and research with European mental health professionals [28,30,31,38-40]. CBT was the most used therapeutic modality when treating adults with depression and anxiety; however, acceptance and commitment therapy and mindfulness-based cognitive therapy were also common. Given that most participating mental health professionals were experienced CBT-based psychologists, further research will be needed to adapt the blended care system to the needs of other types of mental health professionals. Consistent with the views of people with lived experience [56], participants also expected the blended DMHI to be evidence-based, credible, and protective of client data. These results align with 2 prior blended care studies [28,39] and existing research [22,24,27,60,61] conducted before and during the COVID-19 pandemic, which showed that mental health professionals value DMHIs with strong evidence bases and prioritize client privacy and data security.

A novel finding from this study is mental health professionals' views on sharing client data, as only 1 blended care study [39] has previously reported this. Behr and colleagues [39] found that psychotherapists believed clients should have control over what data their psychotherapist can see in a DMHI. Our results build upon these findings by providing specific details on how mental health professionals wanted this data sharing to occur. In this study, most participants preferred to receive client information through a secure email account or online portal on a weekly or fortnightly basis and wanted to view data on clients' progress through the DMHI program, mental health symptoms, and feedback on their treatment. Both our results and those of Behr and colleagues [39] show that mental health professionals value maintaining the security of client data when sharing data with treatment providers. Further research is needed to validate these findings, given the predominance of psychologists in this study and psychotherapists in the Behr and colleagues [39] study. Despite this, the data sharing findings align with qualitative research by Atik and colleagues [28] on blended care, who found that clients were not concerned about sharing personal data within a DMHI, provided their data were used only for treatment purposes. However, these 2 studies [28,39] and our study were conducted during the COVID-19 pandemic; therefore, further replication is needed to ensure that these findings were not due to the unusual circumstances of this worldwide event.

Participants in this study believed blended care could enhance psychotherapy by helping to track a client's treatment progress and providing clients with psychoeducation and skills practice that support in-session therapy. Similarly, other blended care studies found mental health professionals, who primarily consisted of psychologists and psychotherapists, prefer blended DMHIs that allow clients to track their symptoms and complete homework between therapy sessions [28,29,38,62]. Further, Cerga-Pashoja and colleagues [38] reported that mental health care workers who deliver low-intensity psychological interventions (ie, UK psychological well-being practitioners) wanted blended care to help facilitate continuity between therapy sessions. The current results indirectly support this desire for treatment continuity, as mental health professionals suggested the inclusion of a web-based dashboard that would allow them to recommend program content and review client progress between and before therapy sessions. Relatedly, mental health professionals wanted the blended DMHI to assist with the monitoring of client symptoms and flag if these worsened or indicated suicidality, further contributing to a sense of treatment continuity and mental health support between therapy sessions.

XSL•FO
RenderX

Blended Care Acceptability

Consistent with previous research [28,39,60] conducted before and during the COVID-19 pandemic, mental health professionals thought that the blended DMHI should direct clients to immediate support if they were feeling unwell or in crisis. Mental health professionals in this study were also wary of using blended care with highly distressed clients or clients with high suicide risk. which is in line with previous research [22,32]. This is despite research evidence indicating that blended care is feasible for treating severe mental illness and may also improve treatment adherence in this population [33]. One reason for the current findings could be that mental health professionals working in private practice, especially psychologists, were well-represented throughout all 3 stages of the participatory design study. These participants may have been more cautious about working with high-risk clients compared to other mental health professionals, as they could be sole traders operating their own private practices or be part of a small business. Other factors influencing participants' willingness to use DMHIs with clients were client suitability and cost. In line with previous research [22,30], mental health professionals thought blended care might not be appropriate for people with poor digital literacy, limited access to digital technology, or comorbidities that could hamper treatment engagement (eg, cognitive impairment and domestic violence). Prior research also supports our finding that mental health professionals preferred DMHIs to be free or low-cost for clients [26,60]. This highlights the importance of using participatory approaches in DMHI research, especially with practitioners who treat the mental health condition in question, and ensuring sustainable funding models are available for ongoing maintenance and support for blended care systems.

Mental health professionals were also concerned about the time and effort needed to implement blended care with clients. This concern has previously been raised in European studies conducted before the COVID-19 pandemic with CBT-trained mental health professionals working at mental health services that use blended care [30,31,38,40]. However, blended care can also save mental health professionals' time by allowing the DMHI to cover some aspects of face-to-face therapy, such as psychoeducation, or using the DMHI to continue contact with the client for aftercare and to maintain treatment gains without additional face-to-face therapy sessions [25]. Ensuring these time-saving qualities are realized is key to enhancing the acceptability of blended care models [30,39]. In this study, mental health professionals made several suggestions for minimizing the workload associated with implementing blended care and the types of training and resources that would support them using a blended DMHI with clients. Strategies included trialing the DMHI from the client's perspective, supplying quick reference guides and resources (eg, written summaries and short videos), and providing a web-based dashboard that could be used to view client progress and prescribe therapeutic content. Mental health professionals also reported a need for ongoing one-on-one and peer support when using blended care with clients and training to help them understand the therapeutic content in the blended DMHI.

Limitations

This study has several limitations that should be noted. This study occurred during the first and second years of the COVID-19 pandemic (ie, October to December 2020 and July 2021). During this time, the demand for mental health professionals increased while mental health service delivery was disrupted due to quarantining restrictions and physical distancing rules [3,23]. This forced many mental health professionals to switch from face-to-face to telehealth service delivery. It is possible that mental health professionals were more motivated than usual to participate in the research due to the sudden need for many to transition to telehealth service delivery. This may have also influenced our finding that blended care was more acceptable than standalone DMHIs. However, our results do align with a European survey conducted before the COVID-19 pandemic [32], which found that a range of stakeholders involved in depression treatment thought blended care was more acceptable than standalone internet-based treatment for mild, moderate, and severe depression. Further research will be needed to disentangle the effects specific to the COVID-19 pandemic on the blended care system in the future.

The generalizability of our results may also be limited, as the sample was predominantly female and English-speaking and was comprised of individuals who resided in major Australian cities, were psychologists, and worked in private practice within a high-income country. These demographics are consistent with research showing that psychotherapy for mental health treatment within Australia is primarily delivered by psychologists working in private practices [45]. However, further research should consider how the blended care system could be adapted for use by other types of mental health professionals, including those not working in private practice and those in rural, regional, or remote areas of Australia. This could increase the accessibility of blended care for people who may not be able to afford to cover the gap in payments for psychotherapy within the current Australian mental health system. Another limitation of this study is the focus on a transdiagnostic model of care. For highly comorbid mental health diagnoses with shared cognitive, emotional, and biological mechanisms, like depression and anxiety, transdiagnostic approaches have many benefits [7-11]. However, as also mentioned by participants in this study, this approach may not be appropriate for all clients. There is still a need for diagnosis-specific treatment, as some symptoms experienced by a client with a diagnosis, such as bipolar disorder, may not be addressed by a transdiagnostic intervention for depression and anxiety [9].

Future Directions

The findings from this study and the participatory design study with people with lived and living experiences of depression and anxiety [56] have been used to create a new Australian blended model of care for depression and anxiety. The present results have been incorporated into a co-design and prototype testing process for a web-based health professional portal that forms part of the myNewWay blended care digital system. The new system also includes a CBT-based and transdiagnostic mobile app for people receiving treatment for depression and anxiety [56]. Phase 2 of the larger Australian project involved refining

XSL•FO
RenderX

and evaluating both digital products for use as part of routine care for depression and anxiety. This included an implementation clinical trial with Australian psychologists and their clients. Findings from this clinical trial will inform the development of a workflow model for integrating the new blended care system into a mental health professional's practice without increasing their workload or compromising quality of care.

Conclusion

The present findings indicate that Australian mental health professionals who deliver psychotherapy value blended DMHIs that facilitate their clinical practice. Given the similarities between the present findings and the European blended care research [28,30,31,38-40], some elements of optimal blended care may be universal and replicable in mental health care systems in other countries. However, it is important to recognize that blended care will not solve the issue of mental health workforce or funding shortages [3]. Moreover, the increasing incorporation of artificial technology into DMHIs may further complicate the implementation of blended care in the future [17]. Blended care could be an option for leveraging the convenience and accessibility of DMHIs in a way that supports evidence-based treatment and the current workforce while longer-term solutions are sought for these larger system-level issues. Our results, as well as those of others using participatory principles [39,60], further highlight the importance of involving mental health professionals in DMHI design and development. Designing DMHIs that work with, rather than against, existing care pathways for depression and anxiety is crucial for improving access to mental health treatment.

Acknowledgments

We thank Zoe Jenkins and Cameron Banks for assisting with participant recruitment.

Funding

Funding for this study was provided by the Ramsay Foundation. The funders had no role in the design, data collection, analysis, or outcomes of this study. JN receives funding from the Australian National Health and Medical Research Council (grant 2008839), the Australian Medical Research Future Fund (grant 1145382), the United Kingdom National Institute of Health Research (NIHR), the Wellcome Trust, and Knowledge Partnership Platform Australia-Indonesia (KONESKI). AEW receives funding from the Australian National Health and Medical Research Council (grant 2017521), the Australian Medical Research Future Fund (grant 1110773), and the Wellcome Trust.

Authors' Contributions

Conceptualization: JMN (lead), KO (equal)
Data curation: RKM (lead), KK (equal), JATM (supporting)
Formal analysis: RKM (lead), KK (equal), KO (equal), JATM (supporting)
Funding acquisition: JMN (lead), KO (supporting), NC (supporting)
Investigation: KO (lead), KK (supporting), CF (supporting), FT (supporting), VC (supporting), KV (supporting)
Methodology: JMN (lead), KO (equal), PB (supporting), SL (supporting), AEW (supporting), MB (supporting), LK (supporting), CF (supporting), FT (supporting), VC (supporting), KV (supporting)
Project administration: KO (lead), RKM (supporting), KK (supporting)
Resources: JMN (lead), NC (supporting)
Supervision: JMN (lead), KO (equal), AEW (supporting)
Validation: AEW (lead), MB (supporting), LK (supporting), SL (supporting), PB (supporting), MB (supporting), LK (supporting)Visualization: RKM (lead), KK (supporting)
Writing – original draft:RKM (lead), KK (equal), JATM (supporting)
Writing – review & editing:JMN (lead), RKM (equal), KK (equal), JATM (supporting), SL (supporting), PB (supporting), MB (supporting), LK (supporting), NC (supporting), AEW (supporting), CF (supporting), FT (supporting), VC (supporting), KV (supporting)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Web-based survey, interview, and focus group questions.
[DOCX File, 46 KB - humanfactors_v12i1e68789_app1.docx]

Multimedia Appendix 2
Web-based survey sample comparisons.
[DOCX File, 44 KB - humanfactors_v12i1e68789_app3.docx]

XSL•FO
RenderX

Multimedia Appendix 3
Use of digital mental health in practice
[DOCX File, 44 KB - humanfactors_v12i1e68789_app2.docx]

References

1. Sunkel C. Global burden of mental illness. In: Haring R, Kickbusch I, Ganten D, Moeti M, editors. Handbook of Global Health: Springer; 2021:1-27. [doi: 10.1007/978-3-030-05325-3_32-1]
2. GBD 2019 Mental Disorders Collaborators. Global, regional, and natiGBD 2019 Mental Disorders Collaborators.onal burden of 12 mental disorders in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Psychiatry 2022 Feb;9(2):137-150. [doi: 10.1016/S2215-0366(21)00395-3]
3. World mental health report: Transforming mental health for all. : World Health Organization; 2022 URL: https://www.who.int/publications/i/item/9789240049338 [accessed 2025-11-17]
4. Malhi GS, Bell E, Bassett D, et al. The 2020 Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for mood disorders. Aust N Z J Psychiatry 2021 Jan;55(1):7-117. [doi: 10.1177/0004867420979353]
5. Szuhany KL, Simon NM. Anxiety disorders: a review. JAMA 2022 Dec 27;328(24):2431-2445. [doi: 10.1001/jama.2022.22744] [Medline: 36573969]
6. Cuijpers P, Miguel C, Harrer M, et al. Cognitive behavior therapy vs. control conditions, other psychotherapies, pharmacotherapies and combined treatment for depression: a comprehensive meta-analysis including 409 trials with 52,702 patients. World Psychiatry 2023 Feb;22(1):105-115. [doi: 10.1002/wps.21069] [Medline: 36640411]
7. Öst LG, Enebrink P, Finnes A, et al. Cognitive behavior therapy for adult anxiety disorders in routine clinical care: a systematic review and meta-analysis. Clin Psychol Sci Pract 2023;30(3):272-290. [doi: 10.1037/cps0000144]
8. Schaeuffele C, Meine LE, Schulz A, et al. A systematic review and meta-analysis of transdiagnostic cognitive behavioural therapies for emotional disorders. Nat Hum Behav 2024 Mar;8(3):493-509. [doi: 10.1038/s41562-023-01787-3] [Medline: 38228727]
9. Schaeuffele C, Schulz A, Knaevelsrud C, Renneberg B, Boettcher J. CBT at the crossroads: the rise of transdiagnostic treatments. J Cogn Ther 2021 Mar;14(1):86-113. [doi: 10.1007/s41811-020-00095-2]
10. Cuijpers P, Miguel C, Ciharova M, Ebert D, Harrer M, Karyotaki E. Transdiagnostic treatment of depression and anxiety: a meta-analysis. Psychol Med 2023 Oct;53(14):6535-6546. [doi: 10.1017/S0033291722003841] [Medline: 36606456]
11. Newby JM, McKinnon A, Kuyken W, Gilbody S, Dalgleish T. Systematic review and meta-analysis of transdiagnostic psychological treatments for anxiety and depressive disorders in adulthood. Clin Psychol Rev 2015 Aug;40:91-110. [doi: 10.1016/j.cpr.2015.06.002] [Medline: 26094079]
12. Mekonen T, Chan GCK, Connor JP, Hides L, Leung J. Estimating the global treatment rates for depression: a systematic review and meta-analysis. J Affect Disord 2021 Dec 1;295:1234-1242. [doi: 10.1016/j.jad.2021.09.038] [Medline: 34665135]
13. Alonso J, Liu Z, Evans-Lacko S, et al. Treatment gap for anxiety disorders is global: results of the World Mental Health Surveys in 21 countries. Depress Anxiety 2018 Mar;35(3):195-208. [doi: 10.1002/da.22711] [Medline: 29356216]
14. Newby JM, Upton E, Mason E, Black M. Technology-based cognitive behavioral therapy interventions. Psychiatr Clin North Am 2024 Jun;47(2):399-417. [doi: 10.1016/j.psc.2024.02.004] [Medline: 38724127]
15. Andrews G, Basu A, Cuijpers P, et al. Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: an updated meta-analysis. J Anxiety Disord 2018 Apr;55:70-78. [doi: 10.1016/j.janxdis.2018.01.001] [Medline: 29422409]
16. Moshe I, Terhorst Y, Philippi P, et al. Digital interventions for the treatment of depression: a meta-analytic review. Psychol Bull 2021 Aug;147(8):749-786. [doi: 10.1037/bul0000334] [Medline: 34898233]
17. Beg MJ, Verma M, M V, Verma MK. Artificial intelligence for psychotherapy: a review of the current state and future directions. Indian J Psychol Med 2025 Jul;47(4):314-325. [doi: 10.1177/02537176241260819] [Medline: 40831472]
18. Pauley D, Cuijpers P, Papola D, Miguel C, Karyotaki E. Two decades of digital interventions for anxiety disorders: a systematic review and meta-analysis of treatment effectiveness. Psychol Med 2023 Jan;53(2):567-579. [doi: 10.1017/S0033291721001999] [Medline: 34047264]
19. Linardon J, Torous J, Firth J, Cuijpers P, Messer M, Fuller-Tyszkiewicz M. Current evidence on the efficacy of mental health smartphone apps for symptoms of depression and anxiety. A meta-analysis of 176 randomized controlled trials. World Psychiatry 2024 Feb;23(1):139-149. [doi: 10.1002/wps.21183] [Medline: 38214614]
20. Borghouts J, Eikey E, Mark G, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. J Med Internet Res 2021 Mar 24;23(3):e24387. [doi: 10.2196/24387] [Medline: 33759801]
21. Linardon J, Fuller-Tyszkiewicz M. Attrition and adherence in smartphone-delivered interventions for mental health problems: a systematic and meta-analytic review. J Consult Clin Psychol 2020 Jan;88(1):1-13. [doi: 10.1037/ccp0000459] [Medline: 31697093]
22. Davies F, Shepherd HL, Beatty L, Clark B, Butow P, Shaw J. Implementing web-based therapy in routine mental health care: systematic review of health professionals' perspectives. J Med Internet Res 2020 Jul 23;22(7):e17362. [doi: 10.2196/17362] [Medline: 32706713]

XSL•FO
RenderX

23. Lattie EG, Stiles-Shields C, Graham AK. An overview of and recommendations for more accessible digital mental health services. Nat Rev Psychol 2022 Feb;1(2):87-100. [doi: 10.1038/s44159-021-00003-1] [Medline: 38515434]
24. Bruno R, Abbott JAM. Australian health professionals' attitudes toward and frequency of use of internet-supported psychological interventions. Int J Ment Health 2015 Jan 2;44(1-2):107-123. [doi: 10.1080/00207411.2015.1009784]
25. Erbe D, Eichert HC, Riper H, Ebert DD. Blending face-to-face and internet-based interventions for the treatment of mental disorders in adults: systematic review. J Med Internet Res 2017 Sep 15;19(9):e306. [doi: 10.2196/jmir.6588] [Medline: 28916506]
26. Scott S, Knott V, Finlay-Jones AL, Mancini VO. Australian psychologists experiences with digital mental health: a qualitative investigation. J Technol Behav Sci 2022 Aug 16:1-11. [doi: 10.1007/s41347-022-00271-5] [Medline: 35991293]
27. Mendes-Santos C, Nunes F, Weiderpass E, Santana R, Andersson G. Understanding mental health Professionals' perspectives and practices regarding the implementation of digital mental health: qualitative study. JMIR Form Res 2022 Apr 12;6(4):e32558. [doi: 10.2196/32558] [Medline: 35412459]
28. Atik E, Schückes M, Apolinário-Hagen J. Patient and therapist expectations for a blended cognitive behavioral therapy program for depression: qualitative exploratory study. JMIR Ment Health 2022 Dec 30;9(12):e36806. [doi: 10.2196/36806] [Medline: 36583934]
29. Dijksman I, Dinant GJ, Spigt M. The perception and needs of psychologists toward blended care. Telemed J E Health 2017 Dec;23(12):983-995. [doi: 10.1089/tmj.2017.0031] [Medline: 28556693]
30. Titzler I, Saruhanjan K, Berking M, Riper H, Ebert DD. Barriers and facilitators for the implementation of blended psychotherapy for depression: a qualitative pilot study of therapists' perspective. Internet Interv 2018 Jun;12:150-164. [doi: 10.1016/j.invent.2018.01.002] [Medline: 30135779]
31. Mol M, van Genugten C, Dozeman E, et al. Why uptake of blended internet-based interventions for depression is challenging: a qualitative study on therapists' perspectives. J Clin Med 2019 Dec 30;9(1):91. [doi: 10.3390/jcm9010091] [Medline: 31905834]
32. Topooco N, Riper H, Araya R, et al. Attitudes towards digital treatment for depression: a European stakeholder survey. Internet Interv 2017 Jun;8:1-9. [doi: 10.1016/j.invent.2017.01.001] [Medline: 30135823]
33. Ehrt-Schäfer Y, Rusmir M, Vetter J, Seifritz E, Müller M, Kleim B. Feasibility, adherence, and effectiveness of blended psychotherapy for severe mental illnesses: scoping review. JMIR Ment Health 2023 Dec 26;10:e43882. [doi: 10.2196/43882] [Medline: 38147373]
34. Berger T, Krieger T, Sude K, Meyer B, Maercker A. Evaluating an e-mental health program ("deprexis") as adjunctive treatment tool in psychotherapy for depression: results of a pragmatic randomized controlled trial. J Affect Disord 2018 Feb;227:455-462. [doi: 10.1016/j.jad.2017.11.021] [Medline: 29154168]
35. Mathiasen K, Andersen TE, Lichtenstein MB, et al. The clinical effectiveness of blended cognitive behavioral therapy compared with face-to-face cognitive behavioral therapy for adult depression: randomized controlled noninferiority trial. J Med Internet Res 2022 Sep 7;24(9):e36577. [doi: 10.2196/36577] [Medline: 36069798]
36. Kooistra LC, Wiersma JE, Ruwaard J, et al. Cost and effectiveness of blended versus standard cognitive behavioral therapy for outpatients with depression in routine specialized mental health care: pilot randomized controlled trial. J Med Internet Res 2019 Oct 29;21(10):e14261. [doi: 10.2196/14261] [Medline: 31663855]
37. Ly KH, Topooco N, Cederlund H, et al. Smartphone-supported versus full behavioural activation for depression: a randomised controlled trial. PLoS One 2015;10(5):e0126559. [doi: 10.1371/journal.pone.0126559] [Medline: 26010890]
38. Cerga-Pashoja A, Doukani A, Gega L, Walke J, Araya R. Added value or added burden? A qualitative investigation of blending internet self-help with face-to-face cognitive behaviour therapy for depression. Psychother Res 2020 Nov;30(8):998-1010. [doi: 10.1080/10503307.2020.1720932] [Medline: 32024447]
39. Behr S, Fenski F, Boettcher J, et al. TONI - one for all? Participatory development of a transtheoretic and transdiagnostic online intervention for blended care. Internet Interv 2024 Mar;35:100723. [doi: 10.1016/j.invent.2024.100723] [Medline: 38370289]
40. Doukani A, Free C, Araya R, Michelson D, Cerga-Pashoja A, Kakuma R. Practitioners' experience of the working alliance in a blended cognitive-behavioural therapy intervention for depression: qualitative study of barriers and facilitators. BJPsych Open 2022 Jul 25;8(4):e142. [doi: 10.1192/bjo.2022.546] [Medline: 35876079]
41. Nakao S, Nakagawa A, Oguchi Y, et al. Web-based cognitive behavioral therapy blended with face-to-face sessions for major depression: randomized controlled trial. J Med Internet Res 2018 Sep 21;20(9):e10743. [doi: 10.2196/10743] [Medline: 30249583]
42. Lungu A, Jun JJ, Azarmanesh O, Leykin Y, Chen CEJ. Blended care-cognitive behavioral therapy for depression and anxiety in real-world settings: pragmatic retrospective study. J Med Internet Res 2020 Jul 6;22(7):e18723. [doi: 10.2196/18723] [Medline: 32628120]
43. Owusu JT, Wang P, Wickham RE, Varra AA, Chen C, Lungu A. Real-world evaluation of a large-scale blended care-cognitive behavioral therapy program for symptoms of anxiety and depression. Telemed J E Health 2022 Oct;28(10):1412-1420. [doi: 10.1089/tmj.2021.0590] [Medline: 35263185]
44. Better access initiative. Australian Government Department of Health, Disability and Ageing. URL: https://www.health.gov.au/our-work/better-access-initiative [accessed 2025-10-24]

XSL•FO
RenderX

45. Pirkis J, Currier D, Harris M, et al. Evaluation of Better Access: Main report. : The University of Melbourne; 2022 URL: https://www.health.gov.au/sites/default/files/2022-12/main-report-evaluation-of-the-better-access-initiative.pdf [accessed 2025-11-17]
46. Jacmon J, Malouff J, Taylor N. Treatment of major depression: effectiveness of cognitive-behavioural therapy with an internet course as a central component. EJAP 2010;5(2):1-8. [doi: 10.7790/ejap.v5i2.153]
47. Robertson L, Smith M, Castle D, Tannenbaum D. Using the Internet to enhance the treatment of depression. Australas Psychiatry 2006 Dec;14(4):413-417. [doi: 10.1080/j.1440-1665.2006.02315.x] [Medline: 17116083]
48. Sethi S, Campbell AJ, Ellis LA. The use of computerized self-help packages to treat adolescent depression and anxiety. J Technol Hum Serv 2010 Aug 31;28(3):144-160. [doi: 10.1080/15228835.2010.508317]
49. Pier C, Austin DW, Klein B, et al. A controlled trial of internet-based cognitive-behavioural therapy for panic disorder with face-to-face support from a general practitioner or email support from a psychologist. Ment Health Fam Med 2008 Mar;5(1):29-39. [Medline: 22477844]
50. Shandley K, Austin DW, Klein B, et al. Therapist-assisted, Internet-based treatment for panic disorder: can general practitioners achieve comparable patient outcomes to psychologists? J Med Internet Res 2008 May 19;10(2):e14. [doi: 10.2196/jmir.1033] [Medline: 18487138]
51. Kay-Lambkin FJ, Baker AL, Lewin TJ, Carr VJ. Computer-based psychological treatment for comorbid depression and problematic alcohol and/or cannabis use: a randomized controlled trial of clinical efficacy. Addiction 2009 Mar;104(3):378-388. [doi: 10.1111/j.1360-0443.2008.02444.x] [Medline: 19207345]
52. Fitzpatrick M, Nedeljkovic M, Abbott JA, Kyrios M, Moulding R. “Blended” therapy: The development and pilot evaluation of an internet-facilitated cognitive behavioral intervention to supplement face-to-face therapy for hoarding disorder. Internet Interv 2018 Jun;12:16-25. [doi: 10.1016/j.invent.2018.02.006] [Medline: 30135765]
53. Vial S, Boudhraâ S, Dumont M. Human-centered design approaches in digital mental health interventions: exploratory mapping review. JMIR Ment Health 2022 Jun 7;9(6):e35591. [doi: 10.2196/35591] [Medline: 35671081]
54. Brotherdale R, Berry K, Branitsky A, Bucci S. Co-producing digital mental health interventions: a systematic review. Digit Health 2024;10:20552076241239172. [doi: 10.1177/20552076241239172] [Medline: 38665886]
55. National mental health workforce strategy 2022-32. : Department of Health and Aged Care; 2023 URL: https://www.health.gov.au/sites/default/files/2023-10/national-mental-health-workforce-strategy-2022-2032.pdf [accessed 2025-11-17]
56. Kikas K, O’Moore K, Miller RK, et al. Needs and expectations for the myNewWay blended digital and face-to-face psychotherapy model of care for depression and anxiety (part 1): participatory design study including people with lived and living experience. JMIR Hum Factors 2025 Jun 11;12:e69499. [doi: 10.2196/69499] [Medline: 40499157]
57. Schoonenboom J, Johnson RB. How to construct a mixed methods research design. Kolner Z Soz Sozpsychol 2017;69(Suppl 2):107-131. [doi: 10.1007/s11577-017-0454-1] [Medline: 28989188]
58. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34. [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
59. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport Exerc Health 2019 Aug 8;11(4):589-597. [doi: 10.1080/2159676X.2019.1628806]
60. Patoz MC, Hidalgo-Mazzei D, Blanc O, et al. Patient and physician perspectives of a smartphone application for depression: a qualitative study. BMC Psychiatry 2021 Jan 29;21(1):65. [doi: 10.1186/s12888-021-03064-x] [Medline: 33514333]
61. Sander J, Bolinski F, Diekmann S, et al. Online therapy: an added value for inpatient routine care? Perspectives from mental health care professionals. Eur Arch Psychiatry Clin Neurosci 2022 Feb;272(1):107-118. [doi: 10.1007/s00406-021-01251-1] [Medline: 33725165]
62. Schuster R, Pokorny R, Berger T, Topooco N, Laireiter AR. The advantages and disadvantages of online and blended therapy: survey study amongst licensed psychotherapists in Austria. J Med Internet Res 2018 Dec 18;20(12):e11007. [doi: 10.2196/11007] [Medline: 30563817]

Abbreviations

CBT: Cognitive behavioral therapy
DMHI: Digital mental health intervention
RCT: Randomized controlled trial

XSL•FO
RenderX

Edited by A Kushniruk; submitted 01.12.24; peer-reviewed by AV Schaik, M Beg; revised version received 04.08.25; accepted 14.10.25; published 04.12.25.

Please cite as:
Miller RK, O'Moore K, Kikas K, Matheson JAT, Whitton AE, Baldwin P, Li S, Black M, Kampel L, Cockayne N, Tuttlebee F, Fraser C, Carr V, Varghese K, Newby JM
Needs and Expectations for the myNewWay Blended Digital and Face-to-Face Psychotherapy Model of Care for Depression and Anxiety (Part 2): Participatory Design Study including Mental Health Professionals
JMIR Hum Factors 2025;12:e68789
URL: https://humanfactors.jmir.org/2025/1/e68789
doi:10.2196/68789

© Rosemaree Kathleen Miller, Kathleen O'Moore, Katarina Kikas, Julie-Anne Therese Matheson, Alexis Estelle Whitton, Peter Baldwin, Sophie Li, Melissa Black, Laura Kampel, Nicole Cockayne, Fiona Tuttlebee, Caitlin Fraser, Victoria Carr, Kathleen Varghese, Jill Maree Newby. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 4.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effects of Choice Set Sizes and Moderations of Anxiety and State Emotions on Mental Health Self-Care Uptake, Engagement, and User Experience: Experimental Study

Siu Kit Yeung, MPhil; Florence H T Leung, MSc; Jason C M Lee, MA; Gabriel Man Hin Cheung*, BSSc; Ching Wan Li*, MA; Winnie W S Mak, PhD

Department of Psychology, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong SAR, China (Hong Kong)

*these authors contributed equally

Corresponding Author:
Winnie W S Mak, PhD
Department of Psychology, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong SAR, China (Hong Kong)

Abstract

Background: Digital mental health platforms often consist of many different forms of self-care exercises. To our knowledge, whether the number of choices presented to the users affects their uptake and experiences and poses negative consequences (ie, not choosing any exercises, choice dissatisfaction) for users, especially those experiencing anxiety and depressive symptoms or unpleasant state emotions, has not been empirically investigated.

Objective: This study investigated the impact of choice set size on practice decisions, completion, satisfaction, and subjective experiences, as well as potential moderators including depression and anxiety symptoms, state emotions, and motivational and decisional attributes on these choice outcomes.

Methods: Participants were recruited through university mass email and social media, and 652 participants were included in our analyses. Participants completed questions regarding anxiety and depressive symptoms, state emotions, and other psychological attributes. Then, they were randomly assigned to 1-choice, 4-choice, and 16-choice conditions, in which they may choose a self-care activity to practice or decide not to practice. Finally, they completed questions regarding completion, satisfaction, engagement, attitude, and perceived improvement in psychological state.

Results: Presenting multiple choices resulted in a higher likelihood of practice (odds ratio 3.12, 95% CI 2.08 to 4.67 and 3.83, 95% CI 2.55 to 5.76; $P<.001$) and better decision satisfaction (16-choice vs 1-choice: $d=0.36$, 95% CI 0.17 to 0.56, $P<.001$; 4-choice vs 1-choice: $d=0.24$, 95% CI 0.05 to 0.43, $P=.03$) compared with presenting with a single choice. Tentative evidence indicates anxiety symptoms and state emotions were meaningful moderators. Specifically, for individuals with more anxiety symptoms and intense negative emotions, presenting a larger choice set (16 choices) resulted in more positive chosen exercise satisfaction, better attitudes toward chosen activity, and higher perceived improvement in mental health state after the activity, when compared with presenting with smaller choice sets (anxiety: $\beta=-0.38$, 95% CI −0.69 to −0.06 to −0.51, 95% CI −0.84 to −0.18; state emotions: $\beta=-0.31$, 95% CI −0.66 to 0.03 to −0.60, 95% CI −0.92 to −0.28). No evidence was found for the moderating effect of motivational and decisional attributes.

Conclusions: The moderation results were contradictory to prior research and our expectation that a larger choice set may result in worse outcomes than a smaller choice set for people who were experiencing higher levels of psychological distress. We speculated that a possible reason for these findings may be that people with more anxiety symptoms and unpleasant emotions may have a stronger need to reduce these uncomfortable symptoms and emotions, and when presented with more choices on self-care activities, there may be a higher possibility that these self-care activities can address their distress.

Trial Registration: OSF Registeries Osf.io puwxb; https://osf.io/hcbx6/

(JMIR Hum Factors 2025;12:e71165) doi:10.2196/71165

KEYWORDS

digital mental health; decision making; choice set size; individual moderators; anxiety; state emotions

XSL•FO
RenderX

Introduction

Background

To maintain and enhance positive mental health, practicing mental health self-care regularly is essential in our daily lives. With the proliferation of digital mental health apps and platforms, numerous exercises derived from approaches such as cognitive-behavioral therapy and mindfulness-based interventions [1,2] become easily accessible. These activities can be guided through audio or visual instructions or completed using worksheets. While having more exercises available may cater to different needs of service users, presenting too many choices on the platform may negatively influence the likelihood of practice, engagement, and satisfaction, especially for people experiencing anxiety and depressive symptoms or unpleasant emotions [1]. Are users more motivated to do mental health self-care exercises, and are their experiences more satisfactory and engaging if they are given fewer or a larger number of choices? Moreover, are these choice decisions and the related completion of practices, engagement, satisfaction, and experiences moderated by users' present psychological distress, emotional states, and individual differences in motivation and decision-making?

While numerous studies have explored the effects of different choice quantities in economic and consumer decision-making contexts [3-7] and some digital health studies have investigated the efficacy of choice provision versus the absence of choice or tunneling [8-10], research comparing a higher number of choices over a low number of choices in health-related settings is scarce. Only a handful of studies have addressed this comparison in health-medical contexts [7,11,12], and to our knowledge, there is no such study on digital mental health self-care. Making choices for mental health self-care is different from many other consumer behaviors. In previous research, consumer behaviors often involve commercial transactions over the buying of daily goods (ie, pens, jam), with multiple dozens of choices in the large choice set [4]. Often, when research involves making health-oriented decisions, the number of choices is generally lower (under 20 choices) compared with other consumer behavior domains [11,12]. Unlike other consumer behaviors, these choices may have more bearing on well-being and health and may be considered as more consequential than daily goods choices for most people. With the increasing popularity of digital mental health, research on choices of mental health self-care may have informative implications for personalization of self-care offerings and understanding of the public's choices over mental health-related self-care activities.

More Choices-Are-Better Perspective

Multiple theories in psychology and economics suggest that individuals tend to experience greater satisfaction when presented with multiple choices compared with having only one option available or a limited number of choices. The availability of multiple choices is believed to support feelings of autonomy, a fundamental human need that boosts motivation based on Self-Determination Theory [5,13,14]. Relatedly, having multiple choices can enhance perceptions of freedom in decision-making and elevate choice satisfaction [3,4,15]. Such beneficial effects have been observed in some health-related contexts such as illness screening and physical exercise [16]. Classical economic models also suggest that a higher number of choices increases the likelihood of individuals identifying and selecting options that match their personal preferences and goals [3,4,17,18].

However, with Chinese cultures emphasizing a culture of deference and compliance to authority [19,20] while Chinese people generally have lower mental health literacy [21,22], but higher levels of mental illness stigma [21,23], they may prefer being told which choice to opt for or may benefit more from a smaller choice set in terms of mental health self-care (see studies by Chernev et al [4], Misuraca et al [5], and Mogilner et al [24] regarding the association between lower expertise/familiarity and better outcomes with a smaller choice set). Thus, understanding how Chinese people respond to different choice sets in a digital mental health platform for self-care under a mix of this theoretical and cultural backdrop has important implications for how researchers and mental health practitioners could promote public mental health in Chinese societies.

Overchoice Effect or Too-Many-Choices Perspective

Contrary to traditional economic beliefs that more choices lead to better outcomes, some studies over the past 2 decades have indicated that an abundance of choices might diminish the likelihood of purchase and dampen choice satisfaction under certain circumstances [4,25-27]. Some individuals may find themselves overwhelmed by an excess of options, leading to decision-making challenges and dissatisfaction due to limitations or exhaustion in cognitive resources and capabilities [5,26,28]. Being presented with many choices, individuals may experience an increase in cognitive costs in processing more information, making judgments and decisions [4]. This phenomenon of negative repercussions from an excessive array of choices is often referred to as "choice overload" or "overchoice" [7,29,30]. However, while numerous studies, meta-analyses, and reviews have investigated this topic, the findings remain mixed, sometimes inconclusive, and controversial [4,30]. The overchoice effect may manifest under specific conditions, such as increased decision complexity and choice set intricacy [4], and may affect individuals with certain psychological attributes, such as prevention regulatory focus [31] or in emotional states such as anxiety [5,32]. In the mental health realm, for individuals with anxiety or depressive symptoms and state unpleasant emotions, it is worthwhile to examine how choice autonomy, which aligns with personal recovery and self-determination, and cognitive overload, which may be particularly relevant for people experiencing psychological distress, may affect decisions, engagement, and experiences with different choice set sizes.

Anxiety, Depression, and State Emotions as Potential Moderators

One of the individual difference moderators that may be relevant is the experience of anxiety. Studies by Hu et al [32] found that individuals experiencing high levels of anxiety were more prone to facing decision difficulties and delaying choices when confronted with a larger array of options. Hu et al [32] posited that anxiety demands more cognitive resources, thereby reducing the capacity for processing information and rendering judgments

XSL•FO
RenderX

based on more information within larger choice sets, so a lower number of choices may be preferred by people experiencing high levels of anxiety. This heightened challenge when encountering a larger choice set in decision-making might lead individuals to postpone choices or not choose any option and perhaps experience lower satisfaction [32-34].

Related to but distinct from anxiety, we are not aware of any study on the potential moderating effect of depression in choice set size effects. Existing research indicates that many individuals experiencing depression often contend with fatigue [35,36] and may experience amotivation [36]. Similar to cognitive resource limitations among individuals with anxiety [32], people with depression may find it challenging to process information involving more choices due to diminished energy and motivation levels. Individuals with depression may avoid activities that demand extensive cognitive resources [37-39] or encounter difficulties and unpleasant experiences when engaging in complex tasks or decisions that require heightened cognitive engagement [38,40,41]. Thus, when the number of choices is high, one may anticipate that individuals exhibiting more depressive symptoms may be more likely to experience the overchoice effect.

Although the above-discussed studies about anxiety and depression were conducted outside the digital mental health context, concerns have been raised by digital mental health researchers regarding the potential negative impact of offering users a wide array of choices, suggesting the possibility of triggering the overchoice effect [1,42,43]. Notably, Valentine et al [1] contended that individuals with depression and anxiety disorders may encounter cognitive challenges, such as difficulties in concentration, rendering them more susceptible to experiencing the adverse effects of overchoice and disengagement when confronted with excessive information and numerous options. It is imperative to investigate whether limiting the number of self-care choices presented, particularly for individuals with heightened levels of anxiety and depression, would lead to better outcomes than presenting an extensive number of choices.

In addition to anxiety and depressive symptoms, individuals' current emotional states may affect their choice preferences. Drawing from the rationale applied to anxiety and depressive symptoms, it is hypothesized that individuals reporting higher levels of anxiety or depressive symptoms, or stronger unpleasant emotions, would benefit more from a restricted number of choices (1 choice or 4 choices) compared with a larger choice set (16 choices), including greater uptake, engagement, and completion of the chosen activity as well as satisfaction with the decision.

Tailoring the Number of Choices Based on Individual Differences in Regulatory Focus and Choice Preferences

In addition to testing the potential interactions between the number of choices and mental health conditions and emotional states, it may be valuable to investigate the moderating effects of dispositional individual characteristics. One pertinent characteristic worth considering is regulatory focus, as proposed by Higgins [44], which pertains to the motivational orientations in goal pursuits and decision-making processes. Regulatory focus can be categorized into prevention focus that emphasizes safety and the prevention of losses or errors, and promotion focus that emphasizes seeking advancements, gains, and growth opportunities.

Research by Tuan Pham and Chang [45] found that individuals with a predominant promotion focus tend to view options selected from a larger choice set more favorably than those from a smaller set. This inclination was attributed to the proactive nature of promotion-oriented individuals, who are eager to capitalize on opportunities that may be more abundant in wider choice sets [45]. Furthermore, promotion-focused individuals often aim to maximize gains and pursue idealistic goals, making them more likely to find an option that aligns with their preferences in a larger selection. Another related explanation is that promotion-focused individuals tend to strive for maximizing gains and are more idealistic [46,47], and a higher number of choices is more likely to consist of an ideal option that matches well with the preference of the decision maker [3,4,17]. In contrast, Polman [31] found that prevention-focused individuals are more affected by choice overload. They worry about making mistakes and are vigilant about "bad options." With more choices, they perceive a higher risk of encountering these "bad options." Therefore, presenting fewer choices may be better for prevention-focused individuals. Given the above findings, it is hypothesized that for predominantly promotion-focused individuals, presenting a larger choice set will be more beneficial than presenting a smaller choice set, whereas for predominantly prevention-focused individuals, presenting a smaller choice set will be better than presenting a larger choice set.

Apart from regulatory focus, a relatively simple potential method of choice-set-individual-characteristics tailoring was tested by directly asking participants their preferences and perceived difficulty in making decisions when given many choices. For people who prefer many choices and perceive making decisions as less difficult [48], we expect that a higher number of choices (19) will be more beneficial in enhancing their practice likelihood, completion rates, decision satisfaction, choice satisfaction, and engagement compared with a lower number of choices (1 choice or 4 choices). In contrast, for people who do not prefer many choices and people who perceive making such decisions as difficult, a lower number of choices (1 choice or 4 choices) may be better than 16 choices. Broadly speaking, accommodating or matching individual preference is generally associated with engagement, choice satisfaction, and completion rates in mental health contexts [49,50]. This method of tailoring may seem intuitive, simple, and feasible to implement if shown to be effective.

Summary of Hypotheses

To summarize, we investigated both symptoms of psychological distress and state emotions, as well as psychological attributes as moderators. We hypothesized the following for practice likelihood, completion of exercise, satisfaction, and engagement:

- For people experiencing lower levels of anxiety, a higher number of choices (16) will result in better outcomes than

XSL•FO
RenderX

lower numbers of choices (1 or 4). For people experiencing higher levels of anxiety, a lower number of choices (1 or 4) is expected to result in better outcomes than a higher number of choices (16).
- For people with fewer or no depressive symptoms, a higher number of choices (16) will lead to better outcomes than lower numbers of choices (1 or 4). For people experiencing more depressive symptoms, a lower number of choices (1 or 4) will result in better outcomes than a higher number of choices (16).
- For people with more intense pleasant emotions, a larger choice set (16) will result in better outcomes than a smaller choice set (1 or 4). For people experiencing more intense unpleasant emotions, a smaller choice set (1 or 4) will lead to better outcomes than a larger choice set.
- For people who are predominantly promotion-focused, a larger choice set (16) is more likely to result in better outcomes than a smaller choice set (1 or 4). For people who are more prevention-focused, a smaller choice set (1 or 4) will lead to better outcomes than a larger choice set (16).
- For people who prefer many choices and perceive making decisions as less challenging, a wider choice set (16) will lead to better outcomes than a smaller choice set (1 or 4). For participants who do not prefer many choices and perceive making decisions as more difficult, a smaller choice set (1 or 4) is more likely to result in better outcomes than a larger choice set (16).

Methods

Open Science Disclosures

We preregistered our 3-condition experiment in the Open Science Framework (OSF), with the template by Van't Veer and Giner-Sorolla [51]. The preregistration document, data files, code, outputs, and materials of our study are available in OSF (Table S2 in Multimedia Appendix 1 [31,52-67] contains links).

Sample Size Determination, Recruitment, and Participants

We are not aware of any empirical study regarding choice set size in digital mental health exercises or investigating individual difference moderators on the impact of choice set size in mental health decision-making and experiences. Thus, our power analysis was not based on effect sizes of prior studies. We determined the minimum sample size through power analysis based on an effect size of f=0.18 (a small to medium effect based on Cohen [68]) that we consider as practically meaningful to detect while considering resource constraints, monetary and time costs, and number of potential participants [69]. We conducted the power analysis with G*Power (version 3.1; Heinrich-Heine-Universität Düsseldorf) [70]. With an expected alpha of 0.00167 (0.05/30, since we expected to adjust alpha with Holm correction to account for the number of hypothesis testing), 3 conditions, and a power of 80%, a minimum sample size of 575 would be required. Accounting for potential attrition and exclusion of participants not meeting validity testing, our target sample size was set as 675.

We recruited participants through the university subject pool and the social media platforms (Facebook and Instagram). We stated in the advertisement and the introductory page of the study that participants should understand Chinese and should be 18 years or older. The number of participants who were randomized to one of the conditions (1-choice, 4-choice, or 16-choice conditions) was 675.

The number of participants who completed at least one dependent variable measure and passed at least one validity check question (thereby included in our analyzed sample) was 652 (1-choice: 216 participants, 4-choice: 220 participants, 16-choice: 216 participants). A total of 23 participants were excluded due to not completing the experiment or not passing at least one validity check. The mean age of the analyzed sample was 21.77, with a SD of 12.41. A total of 40.03% of participants were women, 57.66% were men, and 2.31% of participants did not disclose gender.

Procedures and Measures

The randomized 3-condition experiment was conducted through Qualtrics. To measure pre-existing characteristics or states, participants first completed scales on anxiety and depressive symptoms, emotional states, and psychological attributes (regulatory focus, preferences, and difficulties in making choices). In mental health self-care, it is reasonable to first gauge participants' distress experience (as measured by anxiety and depressive symptoms) and their present emotional state at the time of engaging in a mental health platform, as this may affect their decision-making and we may adjust based on these potential moderators. This also simulates digital mental health contexts in which pre-existing states or psychological attributes are measured for personalization purposes. The effects of different choice set sizes on emotional states, anxiety, or depression were not tested, so such measures were presented before choice presentations only. Participants were randomized to either 1-choice, 4-choice, or 16-choice conditions, in which they chose to practice with one of the mental health exercises or chose not to practice. After that, participants responded to outcome variables including satisfaction, engagement, and subjective experiences. The entire experiment takes around 20 to 25 minutes. Details about the scales and measures are reported in the following.

Choice Set Sizes

Rationale

In the context of mental health self-care, we anticipated a possibility that for a proportion of participants experiencing moderate to severe levels of depression, they may be amotivated to make choices on their own [36]. Moreover, in the Chinese culture, many people were socialized since very young to be deferent to authorities and exercise conformity or compliance [19,20]; thus, some Chinese people may find making choices more challenging or prefer simply to be told what to do. As such, we included a single-choice option. We are aware that the choice set size literature generally focuses on comparing various levels of choice sets with a smaller number of multiple choices to a very large number of choices (meta-analysis by [4]); however, we are unaware of any study that focused on choice

XSL•FO
RenderX

set sizes for mental health-related self-care exercises. In light of the clinical and cultural considerations, we decided to include a single-choice condition to investigate whether providing a single choice may be better for a subgroup of participants. Participants can still opt out of not practicing self-care, even with this single-choice option.

As to our decision of including 4-choice and 16-choice as the other 2 experimental conditions for investigation, we acknowledged that these choice set sizes contain fewer options than some consumer behavior studies that involve products such as pens and DVDs [4] and the decision does not have much implication for their health. However, our choice set sizes were similar to studies that involve health decision making and a study that found support for the overchoice effect among people with higher levels of anxiety (eg, [12,32,71], also see Chernev [4] meta-analysis regarding the number of choices in different studies). Health-related decisions generally involve less familiar and more complex information compared with consumer decisions like buying food and other daily life products. A 4-choice condition is considered a relatively manageable size, similar to other studies using smaller choice sets [12,32]. As to a larger choice set, presenting participants with 30 or more choices on mental health self-care appears unrealistic and may not even be appropriate in light of the generally low levels of mental health literacy and unfamiliarity of the public with mental health self-care among Chinese [21,22]. Moreover, considering the user interface of presenting various choices on mobile apps, the maximum number of presenting a 16-option choice set is deemed more suitable in real-world implementation. Although the exact number of choices is not definite, we deemed comparing a single, manageable few, and relatively large choice sets in digital mental health contexts to be realistic and meaningful, as any of these choice set sizes has potential implementation implications on a mental health mobile app. Due to research constraints (eg, number of participants to be recruited, funding), we could not include more than 3 conditions with a greater number of choices.

Anxiety Symptoms

Participants completed the Generalized Anxiety Disorder-7 (GAD-7) [72], which measures anxiety symptoms over the past 2 weeks and consists of 7 items with 0 to 3 Likert scales (0=Not at all, 1=Several days, 2=More than half the days, 3=Nearly every day). Sum scores were used, with higher total scores indicating more severe levels of anxiety symptoms. GAD-7 has been validated in both general and clinical populations and has shown adequate internal consistency, construct validity, and convergent validity [72-74]. With our sample, GAD-7 has an internal consistency of α=.922.

Depressive Symptoms

Depressive symptoms over the past 2 weeks were measured using the Patient Health Questionnaire (PHQ-9) [75], with 9 items and 0 to 3 Likert scales (0=Not at all, 1=Several days, 2=More than half the days, 3=Nearly every day). The score is the sum of all items, and higher scores indicate more severe levels of depressive symptoms. PHQ-9 has been validated in both clinical populations and general populations and has demonstrated good reliability, construct validity, and convergent validity [76,77]. In our sample, PHQ-9's internal consistency was α=.876.

State Emotions

Apart from measuring anxiety and depressive symptoms, state emotions were measured by using 8 emotions with 2 emotions from each quadrant of the Circumplex Model [78]. The emotions measured include tired, sad (unpleasant and lower arousal), tense, distressed (unpleasant and higher arousal), happy, excited (pleasant and higher arousal), relaxed, and calm (pleasant and lower arousal). 5-point scales adapted based on The Positive and Negative Affect Schedule [79] were adopted (1=very slightly or not at all, 2=a little, 3=moderately, 4=quite intense, 5=very intense, slight changes in wording for 4th and 5th point). The internal consistency of the pleasant emotions subscale was α=.708, and the unpleasant emotions subscale was α=.834 in our study. For the analyses, we averaged the pleasant emotions and the unpleasant emotions to calculate composite scores for pleasant emotions and unpleasant emotions. We then calculated the difference between pleasant emotions and unpleasant emotions, in which a positive score implies more intense pleasant emotions (compared with unpleasant emotions), whereas a negative score implies more intense unpleasant emotions (compared with pleasant emotions).

Regulatory Focus

Promotion and prevention orientations in regulatory focus were measured using Gomez et al [80] Health Regulatory Focus Scale (HRFS). The scale consists of eight 7-point scale (1=strongly disagree, 7=strongly agree) items (5 promotion items and 3 prevention items). We replaced "health" with "mental health" in the items so that the items are more relevant to mental health. An example of promotion focus is "If I see a good opportunity to improve my mental health, I take advantage of it right away," and an example of prevention focus is "When I implement a mental health behavior, it's because I want to protect myself from getting sick." The scores are calculated by averaging promotion items and prevention items. HRFS demonstrated an advantage in predictive validity within health behavior contexts over general regulatory focus scales [52,53]. It also has adequate internal consistency, test-retest reliability, convergent validity, and discriminant validity [80]. In our study, HRFS showed internal consistency of α=.887 in promotion focus and α=.698 in prevention focus.

Perceived Difficulty of Making Choices and Preference for Choices

Perceived Difficulty of Making Choices (PDMC) and Preference for Choices (PFC) were measured with scales constructed and adopted by Feldman et al [48], which were positively associated with the Free Will and Determinism Scale [81]. Both scales consist of 2 items each with 7-point scales (1=strongly disagree, 7=strongly agree). An example item of PFC is "In each decision I face, I prefer to have as many options as possible to choose from," and an example item of PDMC is "It's very hard for me to choose between many alternatives." Scale scores were calculated by averaging the items. The scales have not been formally validated. PFC scale shows an internal consistency of α=.785, and the internal consistency of PDMC is α=.690.

XSL•FO
RenderX

While we focused on preregistered confirmatory moderators in this manuscript, we included additional psychological attributes, notably Need For Cognition [82] in the same data collection, and the full list of these scales is reported in Multimedia Appendix 1 [31,52-67]. (We may conduct exploratory analyses with such variables for separate manuscripts and we welcome researchers to conduct secondary exploratory analyses [eg, machine learning methods, see discussion] with our open dataset.)

After completion of the above measures, participants were randomized to either (1) 1-Choice condition, (2) 4-Choice condition, or (3) 16-Choice condition. Each condition consists of an array of mental health self-care exercises with brief descriptions about the tasks involved and the purpose of each exercise. All exercises have the same purpose in improving mental well-being and reducing psychological distress and last for a few minutes. They are commonly used exercises in mindfulness-based, self-compassion-based, and cognitive-behavioral interventions, which are evidence-based approaches demonstrated to improve mental well-being and reduce anxiety and depressive symptoms [83-85]. In a previous randomized controlled trial conducted among Chinese in Hong Kong [84], brief mobile app-based mindfulness-based, self-compassion-based, and cognitive-behavioral interventions were found to be equivalent and effective in improving mental well-being and reducing psychological distress.

Half of the exercises involve writing tasks, and half of the exercises are experiential, based on audio/video. Examples include mindful awareness of emotions, mindful breathing (mindfulness), self-compassion letter-writing, self-kindness meditation (self-compassion), problem-solving, and absorbing activities (cognitive-behavioral). The order of the exercises was randomized to prevent order effects, and the choices of the single-option and 4-choice option (2 writing, 2 experiential) were randomized from the 16 choices. More details regarding the exercises and each condition are reported in the Multimedia Appendix 1 [31,52-67] "Details regarding Different Choice Sets and Exercises" section, with screenshots on Qualtrics, titles, and descriptions of all exercises presented or randomized. The Qualtrics QSF file is shared on OSF [86], in which people can download and upload the file on Qualtrics to view the experiment with exercises in the Chinese language.

After choosing "choose later," "do not want to practice," or choosing and involving in one of the exercises, participants were presented with the following variables. Variables designated as primary are variables for which we have specific hypotheses, and secondary are variables that are exploratory, as indicated in our preregistration.

Manipulation Check

We included a manipulation check: "The previous page has many mental health exercise choices," with a 7-point scale (1=Strongly Disagree, 7=Strongly Agree), to test if the choice set size manipulation results in perceived differences in the number of choices across conditions. This question was presented neutrally to avoid negative or positive connotation regarding many choices, as there may be an impact on subsequent responses to the dependent variables if such a question was framed positively (diverse choices) or negatively (too many choices).

Decision Satisfaction (Primary)

Participants completed 2 items on decision satisfaction, which were based on Benoit and Miller's [54,83] items. They included one item on satisfaction with the number of options and one item on satisfaction with decision experiences (the process). An example of a decision satisfaction item is "Are you satisfied or unsatisfied with the number of options just a moment ago?" Both items have 7-point scales (1=very unsatisfied, 7=very satisfied), which is the same as Benoit and Miller [54]. The internal consistency of the scale is α=.721.

Chosen Exercise Satisfaction (Primary)

Two items from Benoit and Miller [54,83] were used to measure chosen option satisfaction, including one item on the degree to which participants are satisfied or unsatisfied with the chosen exercise and one item on the degree to which participant believes the chosen exercise fulfills personal preferences. An example item is "Are you satisfied or unsatisfied with your chosen exercise just a moment ago?" The 2 items have 7-point scales (1=very unsatisfied, 7=very satisfied; 1=completely unable to fulfill my preference, 7=strongly fulfill my preference). The scale had an internal consistency of α=.763 in this study.

Completion (Primary)

Participants were asked whether they completed the exercise with a single question ("Did you complete the exercise just a moment ago?") with three options: (1) did not choose any of the exercises, (2) completed, and (3) not completed. Only data from participants who chose one of the exercises were analyzed.

Engagement (Primary)

A single question on the degree of engagement in the exercise ("What was your level of engagement in the exercise just a moment ago?") was used, with a 7-point scale (1=not engaged at all, 7=very engaged). Only data from participants who selected one of the exercises were analyzed.

We also included the following secondary measures for exploratory purposes. We did not have specific hypotheses as we are unaware of any study in the number-of-choice literature with such dependent variables, and at the preregistration stage, we perceived such variables as more distal compared with the primary measures we included. Despite that, we believed including these items would be worthwhile for several reasons. These measures are simple and not lengthy, particularly relevant in mental health contexts, and may inform future studies in the number-of-choice literature.

Attitudes Toward Chosen Exercise (Secondary)

Three items measuring participants' attitudes toward chosen exercise were used, with one item each on liking, perceived helpfulness of the exercise, and desirability of the exercise. All these items are measured with 7-point scales (1=not desirable at all for me, 7=very desirable for me; 1=strongly dislike, 7=strongly like; 1=not helpful at all, 7=very helpful). An example item is "Do you think that the exercise you just did a

XSL•FO
RenderX

moment ago was helpful for you?" The internal consistency was α=.903.

Perceived Improvement in Psychological State (Secondary)

Participants who chose one of the exercises also completed an item regarding their perceived improvement in mental state ("Do you think your mental state has improved or not due to the exercise a moment ago?). This item is measured with a 7-point scale (1=no improvement at all, 7=very substantial improvement).

Apart from the above secondary variables, 2 additional secondary items on behavioral intention in practicing the exercise in the future and engaging in our digital mental health platform were included. In retrospect, we consider these items to be too distal or not well-designed. These items are included in the shared datasets (Table S1 in Multimedia Appendix 1 [31,52-67] contains links), and other researchers may conduct analyses with such variables, while understanding the limitations of such variables.

After the first survey, participants were sent a link in which they could optionally click on to practice or not, with the same number and combination of options as the first survey. Two weeks later, they fill in questions regarding their experiences with similar measures as the first survey. As very few (lower than 15%) of participants engaged in practice after the first survey, and most of the items require participants to have chosen or practiced exercises after the first survey to be valid, a very low number of responses received in the follow-up survey (<100 participants) can be analyzed. As a result, our analyses in the results section focus on participants' practice decisions, completion, and experiences of decision making and practices in the first survey, and the potential moderation effects.

Ethical Considerations

We obtained ethics approval from the Chinese University of Hong Kong Survey and Behavioral Research Ethics Committee (number: SBRE-23 - 0081). Participants provided informed consent via the web before the experiment, and the consent form indicated that they could withdraw from the experiment at any time. We emphasized in the consent form that participants would get compensated for their time regardless of whether they engaged in self-care exercise or not. Participants did not need to provide their names, and we noted in the consent form that their data would be confidential, anonymized, and deidentified. These practices may reduce social desirability bias in consumer research [87,88]. The datasets shared on OSF are anonymized and deidentified. Each participant was compensated with $50 Hong Kong dollars (equivalent to US $6.42) for completing the study, as transparently stated in the consent form.

Results

Notes Regarding Validation Check and Reported Results

We only reported results of participants who passed at least 1 validity check in the main manuscript. This procedure was not preregistered, but we believe this would be necessary to ensure the quality of the analyzed responses. For transparency purposes, we report results of all participants who were randomized to one of the conditions and completed one of the dependent variables in Multimedia Appendix 1 [31,52-67] (including participants who failed both attention checks). The results are highly similar.

Manipulation Check

A one-way Welch ANOVA was conducted to test whether the manipulation was successful, in other words, whether there were significant differences in levels of agreement on the presence of many choices between participants in different conditions. Welch's ANOVA was adopted instead of Fisher's ANOVA as it is robust against unequal variances [89]. Results indicate that the manipulation was successful with substantial differences in perceived number of choices between conditions and are reported in the Manipulation Check Results section in Multimedia Appendix 1 [31,52-67].

Alpha Adjustment for Multiple Testing

Alpha Adjustment Methods

As preregistered, to address risks of false positive results due to testing with multiple moderators and multiple dependent variables, the Holm sequential alpha adjustment method [55] was adopted for tests associated with our confirmatory hypotheses (with primary outcome variables). With the Holm method, the *P* values for each family were ranked from smallest to largest. For the lowest *P* value in the family, the alpha is adjusted by dividing 0.05 by the total number of tests in the family (0.05/m). For the second lowest *P* value in the family, the alpha is calculated by dividing 0.05 by m-1. For the third lowest *P* value, then adjusted alpha=.05/m^{-2}, etc. As preregistered, the tests were divided into 2 families, each with different research objectives [90]. In Family A, tests were divided in relation to anxiety and depressive symptoms as well as state emotions (potential moderations of GAD-9, PHQ-7, and state emotions, with primary outcome variables listed above). In Family B, tests were divided in relation to potential moderation of chronic traits in motivation and decision making (HRFS, PDMC, and PFC).

Regarding our secondary (exploratory) dependent variables, we did not preregister any alpha adjustment method. For exploratory analyses, two approaches are common: (1) not adopting alpha adjustment [90-92], (2) adjusting alpha with less conservative and more powerful approaches such as Benjamini-Hochberg False Discovery Rate method [93,94]. We contemplated both approaches and decided to adopt the Benjamini-Hochberg method due to the substantial number of tests involved. With the Benjamini-Hochberg procedure, *P* values are ranked from smallest (order of 1) to largest. The adjusted alpha is calculated by rank/number of tests * 0.05. To illustrate, when the number of exploratory tests in a family is 24, the adjusted alpha for the lowest *P* value would be 1/24 * 0.05=0.00208333333, whereas the adjusted alpha for the 5th lowest *P* value would be 5/24 * 0.05=0.01041666666. Similarly, such tests were classified into families [90]: (1) comparison between different numbers of choices (main effect), (2) potential moderation of anxiety and

XSL•FO
RenderX

depressive symptoms and state emotions, (3) potential moderation of psychological attributes.

Practice Exercise Decision

Using Jamovi [95], we conducted a logistic regression to compare participants' likelihood of practicing (in the first survey) between 3 conditions. The chi-square goodness-of-fit test indicates nonequal distribution in practice decisions between conditions, χ^2_2=50.8, *P*<.001. Compared with the 1-choice condition (54/216, 24.8%), participants in the 4-choice condition (113/220, 50.7%) were substantially more likely to practice (ie, choose an exercise option), odds ratio (OR) 3.12, 95% CI (2.08-4.67), *P*<.001 (below adjusted alpha of .00416666666). Similarly, participants in the 16-choice condition (121/216, 55.8%) were much more likely to choose an exercise option, OR 3.83, 95% CI (2.55-5.76), *P*<.001 (below adjusted alpha of .00208333333), compared with participants in the 1-choice condition (54/216, 24.8%). However, we found no significant differences in practice likelihood between participants in the 16-choice condition (121/216, 55.8%) and the 4-choice condition (113/220, 50.7%), OR 1.23, 95% CI (0.84-1.79), *P*=.29.

Completion of Exercise

Another logistic regression was conducted to compare participants' likelihood of completing chosen exercises between conditions. We found no support for differences in distributions of completion between conditions, χ^2_2=1.08, *P*=.583. The completion rates between conditions were not significantly different, 16-choice (106/139, 76.3%) versus 1-choice (90/112, 80.4%): OR 0.79, 95% CI (0.43-1.44), *P*=.44; 4-choice (111/137, 81.0%) versus 1-choice (90/112, 80.4%): OR 1.04, 95% CI (0.56-1.96), *P*=.90; 4-choice (111/137, 81.0%) versus 16-choice (106/139, 76.3%): OR 1.33, 95% CI (0.75-2.37), *P*=.34.

Decision Satisfaction

One-way Welch ANOVA was conducted, and differences in decision satisfaction between conditions were found, $F_{2, 432.00}$=7.42, *P*<.001. Games-Howell *t*-tests were subsequently conducted. Participants in the 16-choice condition (n=216, mean 4.80, SD 1.13) had higher decision satisfaction compared with participants in the 1-choice condition (n=216, mean 4.41, SD 1.05), $t_{427.85}$=3.77, *P*<.001 (below adjusted alpha of .00625), *d*=0.36, 95% CI (0.17-0.56). Tentative evidence indicates that participants in the 4-choice condition (n=220, mean 4.66, SD 1.06) have higher decision satisfaction than participants in the 1-choice condition, $t_{433.91}$=2.52, *P*=.03 (above adjusted alpha of .00833333333), *d*=0.24, 95% CI (0.05-0.43). The difference between the 16-choice condition and the 4-choice condition did not reach significance, $t_{430.87}$=1.36, *P*=.37, *d*=0.13, 95% CI (−0.06 to 0.32).

Other Immediately Measured Variables

Results from Welch's ANOVAs showed no evidence for significant differences in other dependent variables, including chosen exercise satisfaction, engagement, attitude toward chosen exercise, and perceived improvement in mental state between conditions with different numbers of choices. Results are reported in Table 1.

Table . ANOVAs for chosen exercise satisfaction, exercise engagement, attitude toward chosen exercise, and perceived improvement in mental state.

Dependent variables	*F* test (*df*), *P* value	16-Choice versus 1-Choice Games-Howell *t* tests (*df*), *P* value	4-Choice versus 1-Choice Games-Howell *t* tests (df), *P* value	16-Choice versus 4-Choice Games-Howell *t* tests (*df*), *P* value
Chosen exercise satisfaction	1.52 (2, 285.75); *P*=.22	1.74 (277.78); *P*=.19	0.85 (274.24); *P*=.67	0.93 (305.82); *P*=.62
Exercise engagement	0.98 (2, 246.61); *P*=.38	−1.00 (231.85); *P*=.58	−1.35 (237.95); *P*=.37	0.42 (268.80); *P*=.91
Attitude toward chosen exercise	0.34 (2, 243.07); *P*=.71	0.49 (228.05); *P*=.88	−0.28 (226.13); *P*=.96	0.82 (270.46); *P*=.69
Perceived improvement in mental state due to exercise	0.77 (2, 243.16); *P*=.46	1.04 (230.87); *P*=.55	0.01 (225.33); *P*>.99	1.09 (270.00); *P*=.52

Number of Practices After the First Survey

Furthermore, we found no evidence for differences in the number of practices (after the first survey) between participants in different conditions, $F_{2, 410.19}$=0.13, *P*=.88. In other words, the differences in the number of practices after the first survey were not significantly different, 16-choice versus 1-choice, $t_{330.39}$=0.50, *P*=.87; 4-choice versus 1-choice, $t_{391.77}$=0.18, *P*=.98; 16-choice versus 4-choice, $t_{410.70}$=0.32, *P*=.95.

Moderation Analyses of Anxiety and Depressive Symptoms as Well as State Emotions

The following analyses were conducted with RStudio linear regression (lm) function [96].

Linear Regressions With Anxiety Symptoms as the Potential Moderator

Findings from linear multiple regressions with anxiety symptoms (GAD-7) as the moderator found no support for significant moderation for decision satisfaction and exercise engagement. For chosen exercise satisfaction, attitude toward chosen exercise, and perceived improvement in mental state, we found tentative evidence for moderation. Results are summarized in Table 2. Simple slope analyses for chosen exercise satisfaction, attitude toward chosen exercise, and perceived improvement in mental state were conducted. A tentative moderating effect (with *P* values below .05 but above adjusted alpha) was found. People who experienced higher levels of anxiety symptoms, such that these participants in the 16-choice condition tended to experience higher chosen exercise satisfaction, more positive

XSL•FO
RenderX

attitudes toward chosen exercise, and more perceived improvement in mental state due to chosen exercise, compared with corresponding participants in the 4-choice and 1-choice conditions (Table 3).

Table . Interaction results (linear regressions) with anxiety symptoms as the potential moderator.

Dependent variables	β (95% CI)	*P* value
Decision satisfaction	• 16-Choice versus 1-Choice * GAD:[a] β=–.18 (–0.37 to 0.01) • 16-Choice versus 4-Choice * GAD: β=–.11 (–0.30 to 0.08)	• 16-Choice versus 1-Choice * GAD: *P*=.06 • 16-Choice versus 4-Choice * GAD: *P*=.25
Chosen exercise satisfaction	• 16-Choice versus 1-Choice * GAD: β=–.30 (–0.53 to –0.08) • 16-Choice versus 4-Choice * GAD: β=−.25 (–0.47 to –0.03)	• 16-Choice versus 1-Choice * GAD: *P*=.009 (above adjusted alpha of .00142857142) • 16-Choice versus 4-Choice * GAD: *P*=.02 (above adjusted alpha of .00151515151)
Exercise engagement	• 16-Choice versus 1-Choice * GAD: β=−.08 (–0.34 to 0.18) • 16-Choice versus 4-Choice * GAD: β=−.03 (–0.28 to 0.21)	• 16-Choice versus 1-Choice * GAD: *P*=.55 • 16-Choice versus 4-Choice * GAD: *P*=.78
Attitudes toward chosen exercise	• 16-Choice versus 1-Choice * GAD: β=−.33 (–0.59 to –0.08) • 16-Choice versus 4-Choice * GAD: β=−.29 (–0.53 to –0.05)	• 16-Choice versus 1-Choice * GAD: *P*=.01 (above adjusted alpha of .00714285714) • 16-Choice versus 4-Choice * GAD: *P*=.02 (above adjusted alpha of .01428571428)
Perceived improvement in mental state	• 16-Choice versus 1-Choice * GAD: β=−.36 (–0.61 to –0.10) • 16-Choice versus 4-Choice * GAD: β=−.29 (–0.52 to –0.05)	• 16-Choice versus 1-Choice * GAD: *P*=.006 (above adjusted alpha of .00178571428) • 16-Choice versus 4-Choice * GAD: *P*=.02 (above adjusted alpha of .01785714285)

[a]GAD: Generalized Anxiety Disorder.

Table . Simple slope analyses with Generalized Anxiety Disorder-7 (GAD-7) as the potential moderator, chosen exercise satisfaction, attitudes toward chosen exercise, and perceived improvement in psychological state as dependent variables.

GAD-7 level	Chosen exercise satisfaction, β (95% CI)	Attitudes toward chosen exercise, β (95% CI)	Perceived improvement in mental state, β (95% CI)
−1SD (lower anxiety)	• 1-Choice versus 16-Choice: β=.12 (–0.22 to 0.45); *P*=.49 • 4-Choice versus 16-Choice: β=.15 (–0.17 to 0.46); *P*=.36	• 1-Choice versus 16-Choice: β=.28 (–0.08 to 0.65); *P*=.13 • 4-Choice versus 16-Choice: β=.19 (–0.14 to 0.53); *P*=.26	• 1-Choice versus 16-Choice: β=.23 (–0.13 to 0.60); *P*=.21 • 4-Choice versus 16-Choice: β=.15 (–0.18 to 0.49); *P*=.38
Mean	• 1-Choice versus 16-Choice: β=−.20 (–0.43 to 0.04); *P*=.10 • 4-Choice versus 16-Choice: β=−.11 (–0.34 to 0.11); *P*=.31	• 1-Choice versus 16-Choice: β=−.06 (–0.31 to 0.19); *P*=.66 • 4-Choice versus 16-Choice: β=−.10 (–0.34 to 0.13); *P*=.39	• 1-Choice versus 16-Choice: β=−.13 (–0.38 to 0.12); *P*=.31 • 4-Choice versus 16-Choice: β=−.14 (–0.37 to 0.10); *P*=.26
+1SD (higher anxiety)	• 1-Choice versus 16-Choice: β=−.51 (–0.84 to –0.18); *P*=.002 (above adjusted alpha of .00138888888) • 4-Choice versus 16-Choice: β=−.38 (–0.69 to –0.06); *P*=.02 (above adjusted alpha of .00147058823)	• 1-Choice versus 16-Choice: β=−.39 (–0.75 to –0.04); *P*=.03 (above adjusted alpha of 0.025) • 4-Choice versus 16-Choice: β=−.40 (–0.74 to –0.06); *P*=.022 (above adjusted alpha of .02142857142)	• 1-Choice versus 16-Choice: β=−.49 (–0.85 to –0.13); *P*=.007 (above adjusted alpha of .00357142857) • 4-Choice versus 16-Choice: β=−.43 (–0.76 to –0.09); *P*=.01 (above adjusted alpha of .00892857142)

Logistic Regressions With Anxiety Symptoms as the Potential Moderator, Practice Decision, and Completion as Dependent Variables

Logistic regressions were conducted to test the moderation of anxiety symptoms (GAD-7) in the impact of different numbers of choices on practice decision and completion. With practice decision as the dependent variable, we found no evidence for moderation, OR (GAD-7 * 4-choice versus 16-choice)=0.88 (95% CI 0.59-1.30), *P*=.52; OR (GAD-7 * 1-choice versus 16-choice)=0.91 (95% CI 0.60-1.38), *P*=.65. With completion as the dependent variable, we found no support for moderation, OR (GAD-7 * 4-choice versus 16-choice)=0.97 (95% CI

XSL•FO
RenderX

0.55-1.72), *P*=.93; OR (GAD-7 * 1-choice versus 16-choice)=0.98 (95% CI 0.54-1.79), *P*=.96.

Linear Regressions With Depressive Symptoms as the Potential Moderator

No support of depressive symptoms as moderation was found for various dependent variables, except attitudes toward chosen exercise. The interaction results are summarized in Table 4.

Table . Interaction results (linear regressions) with depressive symptoms as the potential moderator.

Dependent variables	β (95% CI)	*P* value
Decision satisfaction	• 16-Choice versus 1-Choice * PHQ:[a]β=−.07 (–0.26 to 0.13) • 16-Choice versus 4-Choice * PHQ: β=.01 (–0.18 to 0.20)	• 16-Choice versus 1-Choice * PHQ: *P*=.51 • 16-Choice versus 4-Choice * PHQ: *P*=.91
Chosen exercise satisfaction	• 16-Choice versus 1-Choice * PHQ: β=−.19 (–0.42 to 0.05) • 16-Choice versus 4-Choice * PHQ: β=−.19 (–0.42 to 0.04)	• 16-Choice versus 1-Choice * PHQ: *P*=.12 • 16-Choice versus 4-Choice * PHQ: *P*=.11
Exercise engagement	• 16-Choice versus 1-Choice * PHQ: β=−.03 (–0.30 to 0.23) • 16-Choice versus 4-Choice * PHQ: β=−.05 (–0.30 to 0.20)	• 16-Choice versus 1-Choice * PHQ: *P*=.80 • 16-Choice versus 4-Choice * PHQ: *P*=.69
Attitudes toward chosen exercise	• 16-Choice versus 1-Choice * PHQ: β=−.18 (–0.44 to 0.08) • 16-Choice versus 4-Choice * PHQ: β=−.25 (–0.50 to –0.01)	• 16-Choice versus 1-Choice * PHQ: *P*=.18 • 16-Choice versus 4-Choice * PHQ: *P*=.04 (above adjusted alpha of .02857142857)
Perceived improvement in psychological state	• 16-Choice versus 1-Choice * PHQ: β=−.25 (–0.51 to 0.01) • 16-Choice versus 4-Choice * PHQ: β=−.23 (–0.48 to 0.01)	• 16-Choice versus 1-Choice * PHQ: *P*=.07 • 16-Choice versus 4-Choice * PHQ: *P*=.06

[a]PHQ: Patient Health Questionnaire.

Specifically, the moderating effect of depressive symptoms on the number of choices (16-choice versus 4-choice) on attitudes toward exercise was inconclusive, with the *P* value below .05 (*P*=.04) but above the adjusted alpha based on Benjamini–Hochberg correction. Simple slope analysis using the PROCESS Macro function in R by Hayes [97] with +1SD value in depressive symptoms (PHQ=13.49) showed that participants in the 16-choice condition (mean=4.71) had more positive attitudes toward exercise compared with participants in the 4-choice condition (mean=4.27), β=.37 (95% CI 0.01-0.72), *P*=.04 (above adjusted alpha of .02678571428). Other regression coefficients at different levels of depressive symptoms were not significant.

Logistic Regressions With Depressive Symptoms as the Potential Moderator, Practice Decision and Completion as Dependent Variables

For practice decision, we found no support for interaction between depressive symptoms and number of choices, PHQ-9 * 4-choice versus 16-choice: OR 1.10 (95% CI 0.75-1.61), *P*=.63; PHQ-9 * 1-choice versus 16-choice: OR 1.32 (95% CI 0.87-2.01), *P*=.19. Additionally, we found no evidence for interaction between PHQ-9 and number of choices on completion of exercises, PHQ-9 * 4-choice versus 16-choice: OR 1.16 (95% CI 0.65-2.07), *P*=.62; PHQ-9 * 1-choice versus 16-choice: OR 1.35 (95% CI 0.73-2.50), *P*=.35.

Linear Regressions With State Emotions as the Potential Moderator

For decision satisfaction and engagement, no evidence for the moderation of state emotions was found. For chosen exercise satisfaction (primary dependent variable), we found conclusive evidence for the moderation of state emotions (*P*<.001, below adjusted alpha) when comparing the 16-choice condition with the 1-choice condition and found tentative evidence for the moderation of state emotions (*P*=.04) when comparing the 16-choice condition with the 4-choice condition. For secondary dependent variables, including attitudes toward chosen exercise and perceived improvement in psychological state, we found tentative evidence for the moderation of state emotions, and such results are reported in Table 5. Building on the moderation findings, simple slopes analyses were conducted. For people who are experiencing more intense unpleasant emotions, participants in the 16-choice condition tend to experience higher chosen exercise satisfaction compared with participants in the 1-choice condition, whereas there is tentative evidence of higher chosen exercise satisfaction for people in the 16-choice condition compared with participants in the 4-choice condition. Furthermore, we found support that participants who experienced more intense unpleasant emotions in the 16-choice condition experienced more positive attitudes toward the chosen exercise, compared with participants in the 4-choice condition, and tentative evidence for differences in chosen exercise

XSL•FO
RenderX

satisfaction between the 16-choice condition and the 1-choice condition. Moreover, we found tentative evidence that people with more intense unpleasant emotions in the 16-choice condition experienced more subjective improvement in mental state compared with participants in the 1-choice condition. We report more details of simple slope analyses in Tables 6 and 7.

Table . Interaction results (linear regressions) with state emotions (continuous) as the potential moderator.

Dependent variables	β (95% CI)	*P* value
Decision satisfaction	• 16-Choice versus 1-Choice * State Emotions: β=.17 (–0.01 to 0.36) • 16-Choice versus 4-Choice * State Emotions: β=.01 (–0.17 to 0.20)	• 16-Choice versus 1-Choice * State Emotions: *P*=.08 • 16-Choice versus 4-Choice * State Emotions: *P*=.88
Chosen exercise satisfaction	• 16-Choice versus 1-Choice * State Emotions: β=.41 (0.18 to 0.64) • 16-Choice versus 4-Choice * State Emotions: β=.23 (0.01 to 0.45)	• 16-Choice versus 1-Choice * State Emotions: *P*<.001 (below adjusted alpha of .00135135135) • 16-Choice versus 4-Choice * State Emotions: *P*=.04 (above adjusted alpha of .00161290322)
Exercise engagement	• 16-Choice versus 1-Choice * State Emotions: β=.18 (–0.08 to 0.44) • 16-Choice versus 4-Choice * State Emotions: β=.14 (–0.11 to 0.38)	• 16-Choice versus 1-Choice * State Emotions: *P*=.19 • 16-Choice versus 4-Choice * State Emotions: *P*=.27
Attitudes toward chosen exercise	• 16-Choice versus 1-Choice * State Emotions: β=.36 (0.10 to 0.62) • 16-Choice versus 4-Choice * State Emotions: β=.29 (0.05 to 0.53)	• 16-Choice versus 1-Choice * State Emotions: *P*=.007 (above adjusted alpha of .00535714285) • 16-Choice versus 4-Choice * State Emotions: *P*=.02 (above adjusted alpha of .01607142857)
Perceived improvement in mental state	• 16-Choice versus 1-Choice * State Emotions: β=.32 (0.07 to 0.58) • 16-Choice versus 4-Choice * State Emotions: β=.17 (–0.07 to 0.41)	• 16-Choice versus 1-Choice * State Emotions: *P*=.014 (above adjusted alpha of .01071428571) • 16-Choice versus 4-Choice * State Emotions: *P*=.16

Table . Simple slope analyses with state emotions as the potential moderator, chosen exercise satisfaction, attitudes toward chosen exercise, and perceived improvement in mental state as dependent variables.

State emotions	Chosen exercise satisfaction, β (95% CI)	Attitudes toward chosen exercise, β (95% CI)	Perceived improvement in mental state, β (95% CI)
−1 SD (stronger unpleasant emotions)	• 1-Choice versus 16-Choice: β=−.60 (–0.92 to –0.28); *P*<.001 (below adjusted alpha of .00131578947) • 4-Choice versus 16-Choice: β=−.35 (–0.67 to –0.03); *P*=.03 (above adjusted alpha of .0015625)	• 1-Choice versus 16-Choice: β=−.40 (–0.75 to –0.05); *P*=.03 (above adjusted alpha of .02321428571) • 4-Choice versus 16-Choice: β=−.40 (–0.75 to –0.06); *P*=.02 (above adjusted alpha of .01964285714)	• 1-Choice versus 16-Choice: β=−.44 (–0.79 to –0.09); *P*=.02 (above adjusted alpha of .0125) • 4-Choice versus 16-Choice: β=−.31 (–0.66 to 0.03); *P*=.07
Mean	• 1-Choice versus 16-Choice: β=−.18 (–0.41 to 0.05); *P*=.12 • 4-Choice versus 16-Choice: β=−.12 (–0.34 to 0.10); *P*=.30	• 1-Choice versus 16-Choice: β=−.04 (–0.29 to 0.21); *P*=.73 • 4-Choice versus 16-Choice: β=−.11 (–0.35 to 0.12); *P*=.36	• 1-Choice versus 16-Choice: β=−.11 (–0.37 to 0.14); *P*=.38 • 4-Choice versus 16-Choice: β=−.14 (–0.38 to 0.10); *P*=.25
+1 SD (stronger pleasant emotions)	• 1-Choice versus 16-Choice: β=.24 (–0.10 to 0.57); *P*=.17 • 4-Choice versus 16-Choice: β=.12 (–0.19 to –0.43); *P*=.46	• 1-Choice versus 16-Choice: β=.31 (–0.05 to 0.68); *P*=.10 • 4-Choice versus 16-Choice: β=.18 (–0.15 to 0.51); *P*=.28	• 1-Choice versus 16-Choice: β=.21 (–0.16 to 0.58); *P*=.26 • 4-Choice versus 16-Choice: β=.03 (–0.30 to 0.37); *P*=.84

XSL•FO
RenderX

Table . Means of 1-Choice, 4-Choice, and 16-Choice conditions at different levels of state emotions.

	1-Choice	4-Choice	16-Choice
−1 SD (stronger unpleasant emotions)	• Chosen exercise satisfaction: 4.13 • Attitude toward exercise: 4.22 • Perceived Improvement in psychological state: 3.57	• Chosen exercise satisfaction: 4.41 • Attitude toward exercise: 4.21 • Perceived improvement in psychological state: 3.75	• Chosen exercise satisfaction: 4.80 • Attitude toward exercise: 4.71 • Perceived improvement in psychological state: 4.18
Mean	• Chosen exercise satisfaction: 4.48 • Attitude toward exercise: 4.50 • Perceived improvement in psychological state: 3.91	• Chosen exercise satisfaction: 4.56 • Attitude toward exercise: 4.42 • Perceived improvement in psychological state: 3.87	• Chosen exercise satisfaction: 4.69 • Attitude toward exercise: 4.55 • Perceived improvement in psychological state: 4.07
+1 SD (stronger pleasant emotions)	• Chosen exercise satisfaction: 4.84 • Attitude toward exercise: 4.78 • Perceived improvement in psychological state: 4.25	• Chosen exercise satisfaction: 4.71 • Attitude toward exercise: 4.62 • Perceived improvement in psychological state: 4.00	• Chosen exercise satisfaction: 4.57 • Attitude toward exercise: 4.40 • Perceived improvement in psychological state: 3.95

Logistic Regressions With State Emotions as the Potential Moderator

With practice decision as the dependent variable, the regression indicates no support for moderation, OR (State Emotions * 4-choice vs 16-choice)=1.03 (95% CI 0.70-1.50), *P*=.89; OR (State Emotions * 1-choice vs 16-choice)=0.92 (95% CI 0.61-1.40), *P*=.71. With completion as the variable, we also found no evidence for moderation, OR (State Emotions * 4-choice vs 16-choice)=0.85 (95% CI 0.47-1.53), *P*=.59; OR (State Emotions * 1-choice vs 16-choice)=1.18 (95% CI 0.64-2.16), *P*=.60.

Moderation Analyses of Individual Characteristics

Linear Regressions With Health Regulatory Focus [80] as the Potential Moderator

Interaction results of health regulatory focus with different dependent variables were all nonsignificant and are summarized in Table S3 in Multimedia Appendix 1 [31,52-67].

Logistic Regressions With Health Regulatory Focus [80] as the Potential Moderator With Practice Decision and Completion as Dependent Variables

With logistic regression, we found no evidence for moderation of health regulatory focus on the impact of number of choices in practice decision, OR (HRFS * 4-choice vs 16-choice)=1.01 (95% CI 0.70-1.47), *P*=.95; OR (HRFS * 1-choice vs 16-choice)=0.77 (95% CI 0.50-1.20), *P*=.26. With completion as the dependent variable, we also found no support for moderation of health regulatory focus, OR (HRFS * 4-choice vs 16-choice)=0.94 (95% CI 0.51-1.72), *P*=.84; OR (HRFS * 1-choice vs 16-choice)=0.67 (95% CI 0.35-1.28), *P*=.22.

Linear Regressions With PFC and PDMC [48] as the Potential Moderator

With PFC as the potential moderator, we found no support for any interaction effects with all continuous variables. Results are summarized in Table S4 in Multimedia Appendix 1 [31,52-67]. Additionally, we found no evidence for moderation with PDMC with all continuous variables (Table S5 in Multimedia Appendix 1 [31,52-67]).

Logistic Regressions With PFC and PDMC [48] as the Potential Moderator

With practice decision as the dependent variable, we found no evidence for moderation with PDMC as the potential moderator, OR (PDMC * 4-choice vs 16-choice)=0.86 (95% CI 0.58-1.27), *P*=.45; OR (PDMC * 1-choice vs 16-choice)=1.23 (95% CI 0.82-1.85), *P*=.31. With completion as the dependent variable and PDMC as the potential moderator, the results also indicate no significant moderation, OR (PDMC * 4-choice vs 16-choice)=1.05 (95% CI 0.57-1.92), *P*=.88; OR (PDMC * 1-choice vs 16-choice)=0.98 (95% CI 0.53-1.79), *P*=.95. Additionally, we found no support for moderation of preferences for choices with practice decision as the dependent variable, OR (PFC * 4-choice vs 16-choice)=0.98 (95% CI 0.66-1.44), *P*=.91; *OR* (PFC * 1-choice vs 16-choice)=0.79 (95% CI 0.52-1.20), *P*=.27. Finally, we found no evidence for moderation of PFC with completion of exercise as the dependent variable, OR (PFC * 4-choice vs 16-choice)=0.92 (95% CI 0.50-1.67), *P*=.78; OR (PFC * 1-choice vs 16-choice)=1.02 (95% CI 0.56-1.84), *P*=.96.

Discussion

Principal Results: Comparing Different Numbers of Choices (Overall Effects)

Our investigation delved into potential differences in various outcomes across different numbers of choices. Participants in 4-choice and 16-choice conditions appeared more likely to practice one of the mental health self-care exercises and reported higher levels of decision satisfaction, compared with participants in the single-choice condition. These findings align with existing studies that underscore the positive impact of providing multiple choices, fostering increased active involvement and satisfaction among participants in health interventions [16].

The rationale behind these results might be elucidated using the self-determination theory. Individuals might be more inclined

XSL•FO
RenderX

to engage in mental health behaviors when granted autonomy to make a choice when given multiple options [5,13-15]. The provision of multiple options caters to the fundamental human need for autonomy and the freedom to choose, possibly leading to heightened decision satisfaction in the multi-choice conditions compared with the single-choice condition. Conversely, individuals presented with only one option, including Chinese people, may experience a sense of choice deprivation and feel less satisfied with the insufficient number of choices [98].

Another plausible explanation rooted in classical economic models suggests that in conditions offering multiple choices, participants are more likely to discover an exercise aligning with their preferences, thus increasing the likelihood of selection [3,4,18]. In contrast, the singular exercise choice provided may not resonate with participants' preferences, leading to lower exercise selection rates. However, our findings found no support for significant differences between the multiple-choice conditions and the single-choice condition in other outcomes, such as satisfaction with the chosen exercise, engagement levels, and completion rates. This lack of differentiation could stem from these variables predominantly reflecting experiences with the specific chosen exercise, with the benefits of multiple-choice provision potentially not extending to these more distal variables.

While our study revealed distinctions in 2 outcome variables between the single-choice and 4-choice conditions, as well as between the single-choice and 16-choice conditions, in terms of overall effects, we did not observe any significant differences in dependent variables between the 4-choice and 16-choice conditions. These findings align with the meta-analysis on choice set size conducted by Scheibehenne et al [30], which reported a negligible mean effect size when comparing higher versus lower numbers of choices. Similarly, our results are consistent with the meta-analysis by Chernev et al [4], which found no substantial differences in outcomes, including satisfaction and option selection, when comparing larger and smaller choice sets, among studies examining potential moderators. The average effect sizes of main effects, particularly in studies testing potential moderators, often tend to be minimal or statistically nonsignificant [4]. While some studies find support for the overchoice effect (eg, [26]), others (eg, [99]) found better outcomes with a higher number of choices, underscoring the variability of effects across different contexts [4].

The reasons behind minimal differences between the 16-choice condition and the 4-choice condition remain unclear, and our findings contradict concerns by some digital mental health researchers who suggested a higher number of choices may backfire [1,42]. It is possible that discomfort or indecisiveness (not selecting any of the choices) due to many-choice comparison is less likely to occur in decision-making contexts where making a choice does not involve any financial costs, including the context of our experiment. However, it is unclear if our findings are generalizable to digital mental health platforms that require payments for audio, video, and written exercises. In contexts where people need to pay the prices, the overchoice effect may be more likely to occur, possibly due to counterfactual thinking, anticipated and experienced regret, or other uncomfortable emotions regarding the monetary costs [4,11,26,100-102]. That said, we are unaware of studies that compare the impact of choice set sizes in decision-making contexts with and without monetary considerations. These explanations are speculative, and further research, comparing digital health self-care choices that do not require payment versus self-care choices that require payment, is warranted to delve deeper into these nuances and test these conjectures.

Principal Results: Tentative Evidence for Moderations of State Emotions and Anxiety

Contrary to our hypotheses, recent findings in the broader overchoice literature beyond mental health contexts [32], and concerns raised by digital mental health researchers [1,42], our study unearthed tentative evidence suggesting that individuals with heightened anxiety levels and stronger negative emotional states may actually experience greater satisfaction, exhibit more positive attitudes toward the selected exercise, and perceive greater psychological improvement when presented with a higher number of choices (16 choices), as opposed to those offered a single choice or 4 choices. Our expectations were rooted in the notion that individuals experiencing intense anxiety and unpleasant emotions might struggle with decision-making in larger choice sets due to limited cognitive resources and heightened difficulties, leading to unpleasant experiences when choosing from a wider array of options [1,32,38,39,41]. However, our findings diverged from this anticipated direction, prompting a reflection. How can we interpret these unexpected results, and are these results reliable?

Our moderation findings are tentative, but we believe such results are likely reliable despite some uncertainties. While for several dependent variables, our results (moderation of anxiety symptoms and state emotions) reached conventional statistical significance of $P<.05$, most of the results did not retain significance following alpha corrections. There exists a possibility that these findings could be attributed to chance, particularly considering the inclusion of multiple potential moderators and a range of primary and secondary dependent variables. Despite this uncertainty, the likelihood of these results being false positives is relatively low, given the medium to large effect sizes observed for the aforementioned dependent variables, based on the benchmark outlined by Fey et al [103]. Notably, effect size serves as a robust indicator of the replicability of research findings [104]. Moreover, a common thread among the variables with $P<.05$ results is that they are all regarding participants' subjective perceptions and sentiments (satisfaction, attitude, and perceived psychological improvement) concerning the selected exercise, rather than behavioral measures like practice decision and completion. Moderations of both anxiety symptoms and emotional states concerned these same subjective dependent variables, suggesting that moderations of anxiety and emotional states for such outcomes likely exist.

For participants with high anxiety and stronger unpleasant emotions, the higher satisfaction, more positive attitudes, and more subjective improvement in mental state with such exercises under a larger choice set may stem from a higher probability of matching their preferences and needs. This resonates with

XSL•FO
RenderX

economic models and studies positing that individuals have a higher probability of finding options that cater to their specific preferences and goals when faced with more choices, as previously discussed [3,4,17]. Crucially, this could also be attributed to the heightened desire of individuals with elevated anxiety levels and unpleasant emotions to alleviate their uncomfortable feelings. A selection pool of 16 exercises is more likely to include self-care activities that effectively address their needs, in contrast to single or limited 4-choice sets. Individuals may perceive only a small fraction of exercises (perhaps one or a few) as beneficial for their personal well-being. When the choice set comprises only 1 or 4 exercises, the chances of encountering an emotion-relief method that precisely aligns with their individual needs are considerably lower. This disparity underscores the potential advantages of a larger choice set over a smaller choice set in matching personal needs, preferences, and goals for individuals experiencing high anxiety and unpleasant emotions.

Related to the heightened need to address uncomfortable emotions, when people experience unpleasant mental health and emotional states, mental health–related exercises may be perceived as more personally relevant. Based on dual process models such as the Elaboration Likelihood Model and Heuristic-Systematic Information Processing Theory, heightened perceived personal relevance can result in higher effort in information processing [105,106]. It is possible that such heightened effort in information processing and more elaborate processes in thinking and decision making may increase the likelihood of participants choosing exercises that can better satisfy their emotional needs and help address their distress, through considering more options effortfully and deciding systematically with a larger choice set. In contrast, smaller choice sets may not align well with such more elaborate and effortful processes.

Conversely, individuals with lower levels of anxiety and unpleasant emotions may not have any need to further regulate their mental health at the time of the experiment and thus may consider mental health exercises as less personally relevant. Consequently, the utility of a larger choice set may hold less significance or may be perceived as irrelevant to them. As such, they may expend less cognitive effort in processing information and weighing different options, which is consistent with dual process models discussed above. Given that practicing these self-care exercises may be considered less relevant to this group of participants, the number of choices presented may yield minimal, if any, discernible effects. This rationale may explain why we did not uncover substantial differences in both behavioral responses and subjective experiences among individuals with lower levels of anxiety and unpleasant emotions across varying choice set sizes.

It is hard to interpret the absence of evidence for the moderating effect of depressive symptoms, as null findings do not necessarily imply the absence of effect. One interpretation is that, as the direction of findings is consistent with anxiety and state emotions, moderation may exist but was not detected due to relatively small effect sizes. Therefore, the moderation effect of depressive symptoms was nonsignificant. Further research with a larger sample size and a wider range of anxiety and depressive symptoms is needed.

Limitations, Strengths, and Future Research Directions

The above explanations regarding the moderations of anxiety symptoms and state emotions are speculative. The study did not delve into testing potential mediating mechanisms, and future research can explore factors such as perceived need to alleviate unpleasant emotions, perceived personal relevance of mental health self-care exercises, and elaboration in processing and deciding on mental health exercises. Through such mediated analyses, mechanisms of the phenomenon can be better understood. Additionally, it remains possible that the moderation findings can be better elucidated by a third variable linked to anxiety and unpleasant emotions, such as fear of missing out [107,108]. Subsequent studies can incorporate these variables to further investigate and validate these potential explanations.

Given the tentative nature of our moderation findings and the lack of significance in most results following alpha corrections, it is imperative to conduct replications and extensions to ascertain the replicability, reliability, and generalizability of the findings. These efforts can involve close or very close replications (directing repeating our work with similar/same designs), or conceptual replications [109] within the realm of digital mental health platforms or diverse contexts like other digital health arenas or consumer decision-making scenarios. Future research initiatives can aim to investigate whether these effects are specific to the digital mental health domain or if they hold broader generalizability and applicability across various decision-making contexts.

An intriguing avenue for conceptual replications-extensions, or follow-up studies, involves investigating the impact of recommendation statements when presenting varying choice sets (ie, a single choice, 4 choices, or 16 choices) on decision outcomes. For instance, investigating how phrases like “We strongly recommend you to practice the following exercise(s)” influence individuals’ choices and experiences compared with conditions without such recommendation language may yield valuable insights. Notably, our study did not incorporate recommendation wordings in any of the conditions, and it is plausible that the inclusion of such language would lead to increased engagement, decision-making, satisfaction, and other subjective outcome variables [1]. Introducing recommendation wordings in select conditions can also enhance the ecological validity of the results, considering that recommendation systems are increasingly common in digital mental health platforms [1,110]. Exploring the influence of recommendation statements within the context of different choice set sizes can provide a deeper understanding of how decision-making processes and subjective experiences are influenced in digital mental health settings and beyond.

Related to recommendations, another worthwhile direction is to compare different numbers of choices with and without tailored recommendations, which have been commonly adopted and researched in digital mental health contexts [1,110,111]. Studies can test if tailored recommendations that are based on prior behavioral data, demographics, mental health states, and other psychological attributes and offered a lower number of

XSL•FO
RenderX

choices may outperform a higher number of choices that are not tailored, as such fewer number of choices may be able to align with service users' needs and preferences better without providing a large amount of information. Studies that involve the intersection of choice set size and tailored recommendations [111] will require a much larger sample size due to a much higher number of conditions but can be feasible in digital mental health platforms where there are many service users. Another method that has been adopted in digital health studies and can be considered in testing the effects of various conditions over time is micro-randomized controlled trials, in which service users or participants are randomized sequentially and multiple times [112,113]. Further research can help us understand if more choices are really needed to achieve matching of needs, or if a lower number of choices of tailored recommendations can already satisfy users' needs. It is possible that when comparing subjective experiences and behavioral outcomes across conditions, mental health states and psychological attributes (eg, trait reactance, [114]) may moderate such effects.

Another worthwhile direction in replications-extensions or follow-up studies is to investigate other choice set sizes. As mentioned in the introduction, the 3 choice set sizes were chosen as such sizes were potentially suitable for different subgroups of participants and feasible to implement in real-world mobile apps. Over 30 choices are perhaps overwhelming due to lower mental health literacy of most Chinese people [21,22], as there may be too much unfamiliar and technical information to process and understand (see studies by Chernev et al [4], Misuraca et al [5], and Mogilner et al [24] regarding higher likelihood of overchoice effect among people with lower familiarity or expertise with the options). However, it is possible that such a number of choices may be suitable for a proportion of participants with higher digital mental health self-care related knowledge, since studies in the broader choice set size literature have shown that people with higher expertise regarding the options benefit more from a larger choice set [4,5,24]. While such a choice set size of over 30 choices does not seem suitable in a mobile app, it can be feasible in a web page in a computer. Another possibility is that a number between 4 and 16 may outperform other numbers of choices for a proportion of participants, perhaps because a smaller choice set does not optimally satisfy needs and preferences, but a larger choice set is also suboptimal due to increased cognitive costs and diminishing marginal benefits, whereas a medium choice set may satisfy needs well with limited cognitive costs (see studies by Grant and Schwartz [115], Liu et al [116], and Reutskaja and Hogarth [117] regarding inverted-U relationship between number of choices and outcomes). Further digital health studies can examine this speculation of an inverted-U relationship. That said, including more conditions with different choice set sizes will require more participants, and these decisions involve dilemmas given considerations for power analysis, number of participants, and real-world constraints in data collection.

Another worthwhile area for improvement in future replications, extensions, and follow-up studies is to include a question on perceived quality to test if the perceived quality of exercises is similar or varies between conditions. Given that exercises in the 1-choice and 4-choice conditions were randomly chosen from the exercises that were also presented in the 16-choice condition, it is likely that perceived quality would not vary much. That said, perceived quality between conditions may influence the results, and future studies can explicitly account for this.

Furthermore, researchers may consider investigating decision-making dynamics, task completion rates, user experiences, and mental health outcomes over extended periods within a digital platform, where individuals repeatedly engage with self-care or health-related options. Conducting such longitudinal studies would necessitate larger sample sizes to ensure adequate statistical power, as sustained engagement in self-care exercises or digital interventions is not common among most individuals [118,119]. These studies can be effectively carried out in real-world digital platforms with a substantial user base, ensuring sufficient statistical power and enhancing ecological validity, while examining generalizability and applicability.

In addition to pursuing replications-extensions or follow-up studies involving primary data collection and analyses, another valuable direction for research involves conducting secondary data analyses with our datasets shared on the OSF (Table S1 in Multimedia Appendix 1 [31,52-67] contains links). Given the tentative nature of our findings, exploring the data with different statistical methods may lead to different findings, offer fresh perspectives, and potentially uncover new insights. Researchers can consider using methodologies like the Personalized Advantage Index [120] and machine learning techniques such as Random Forests to delve into potential moderators and enhance personalization strategies [56]. Moreover, while our manuscript primarily focused on preregistered confirmatory moderators and hypotheses, as well as a few exploratory outcome variables, we included additional measures of psychological attributes such as Need for Cognition [82] in our data collection (Multimedia Appendix 1 [31,52-67]). We welcome and encourage researchers to conduct secondary analyses based on our dataset and are open to potential collaborations in this regard.

Implications

Presenting a higher number of choices does not appear to be counterproductive for most individuals with higher levels of anxiety symptoms, depressive symptoms, or intense state unpleasant emotions. For individuals experiencing elevated anxiety and depression symptoms or unpleasant emotions, there is currently no empirical evidence supporting the reduction of presented choices if there are initially 10 - 20 options shown. Instead, our findings tentatively indicate that offering a greater number of choices (16 options) may actually enhance the subjective experiences of individuals with high anxiety and intense unpleasant emotions. This could stem from an increased likelihood of aligning with and addressing their mental health needs with a larger choice set. However, we cannot definitively ascertain the outcomes of presenting individuals with an even larger number of choices (eg, over 30 choices, see Chernev et al [4] meta-analysis), and caution is advised for digital mental health developers and designers when presenting dozens of options.

XSL•FO
RenderX

In the realm of digital mental health platforms, many already incorporate measures like GAD-7 for assessing anxiety, where our findings find support for moderations based on anxiety symptoms. Tailoring the number of choices to individual anxiety levels may be beneficial. Furthermore, our moderation results concerning state unpleasant emotions suggest that including measures of current emotional states may be valuable. Developers of digital mental health platforms also need to consider the impact of adding additional items, such as 8 new state emotion measures, to existing assessments, as this may deter some participants from completing them. One potential strategy could involve including 1 to 3 state emotion items that exhibit stronger or more consistent moderation effects compared with others. Digital mental health platforms may include multiple state emotion items for a period of time. After that, based on behavioral, satisfaction, and subjective experiences data and corresponding moderation analyses (eg, through machine learning methods such as Random Forest [56]) with the initially included emotions, they may choose 1 to 3 state emotions that show stronger moderation effects to increase the likelihood that people fill in such measures and to facilitate personalization for a higher proportion of participants.

Conclusions

Our research found support for the advantage of having multiple exercise choices (4 choices or 16 choices) in enhancing the likelihood of practicing self-care and decision satisfaction, surpassing such outcomes of the single choice condition. More importantly, contrary to our expectations, prior findings [32], and concerns regarding the impact of a larger choice set in the digital mental health field [1,42], we found tentative evidence that for people with more anxiety symptoms and more intense unpleasant emotions, presenting them with a higher number of choices (16 choices) would be more beneficial for their subjective experiences of the chosen exercise (ie, satisfaction, attitude, perceived improvement in mental state) than presenting them with a lower number of choices (4 choices or 1 choice). This unexpected finding could be attributed to the higher probability of encountering exercises that align with participants' mental health needs, preferences, and effectively address their emotional distress within a more extensive choice array.

In light of these intriguing findings, we advocate for further investigations into the underlying mechanisms at play, replication and extension studies to examine the reliability and generalizability of our findings, perhaps with additional choice set size conditions, secondary analyses using our open dataset, and additional research on the intersection between tailored recommendations and choice set size, as well as impact of choice set size in decision-making contexts without monetary considerations. Moreover, we highlight the potential of tailoring the number of choices based on individuals' anxiety symptoms and current emotional states within digital mental health platforms. We call for further investigation aimed at deepening our understanding of how choice set sizes can be potentially tailored to enhance engagement, user experiences, and outcomes in digital mental health environments.

Acknowledgments

We used Poe AcademicWriterBot [121] in copy-editing the English language to improve the manuscript's readability. We adapted some of these suggestions. We also thank Hoi Ching Yu and Yuxuan Zheng for their help with pretesting the Qualtrics surveys. The digital mental health self-care exercises included as choices in the study are from Jockey Club Tour Heart+Project, a digital mental health platform, developed and led by the corresponding author WWSM.

Funding

This study is funded by the internal grant from Mental Health Research Center (MHRC) of the Hong Kong Polytechnic University (Project ID: P0040456) as well as the General Research Fund from Research Grants Council (RGC) of Hong Kong SAR, China (Reference number: 14606623). The Jockey Club Tour Heart+Project is funded by the Hong Kong Jockey Club Charities Trust (SN: 2024 - 0077) for the Jockey Club TourHeart 3.0: Advancing Precision Mental Health Care Through Generative Artificial Intelligence.

Data Availability

The datasets, analysis files, and code with outputs of this study are available on the Open Science Framework, with links available in Table S1 in Multimedia Appendix 1.

Authors' Contributions

SKY, WWSM, and FHTL contributed to conceptualizations of the study. SKY, FHTL, and JCML contributed to the design of the experiment, whereas all authors helped with pretesting the Qualtrics survey and providing suggestions. JCML was responsible for data collection of the study. WWSM was responsible for the supervision of the project. SKY conducted analyses and wrote the initial draft of the manuscript and revised the manuscript based on suggestions. WWSM made many changes to the manuscript, and GMHC and CWL provided some suggestions for changes to the manuscript.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Supplementary file with OSF links, supplementary results, and details regarding method or materials.
[DOCX File, 589 KB - humanfactors_v12i1e71165_app1.docx]

References

1. Valentine L, D'Alfonso S, Lederman R. Recommender systems for mental health apps: advantages and ethical challenges. AI Soc 2022 Jan 17;38(4):1-12. [doi: 10.1007/s00146-021-01322-w] [Medline: 35068708]
2. Baumel A. Therapeutic activities as a link between program usage and clinical outcomes in digital mental health interventions: a proposed research framework. J Technol Behav Sci 2022 Jun;7(2):234-239. [doi: 10.1007/s41347-022-00245-7]
3. Botti S, Iyengar SS, McGill AL. Choice freedom. J Consum Psychol 2023 Jan;33(1):143-166. [doi: 10.1002/jcpy.1325]
4. Chernev A, Böckenholt U, Goodman J. Choice overload: a conceptual review and meta‐analysis. J Consum Psychol 2015 Apr;25(2):333-358. [doi: 10.1016/j.jcps.2014.08.002]
5. Misuraca R, Nixon AE, Miceli S, Di Stefano G, Scaffidi Abbate C. On the advantages and disadvantages of choice: future research directions in choice overload and its moderators. Front Psychol 2024;15(1290359):1290359. [doi: 10.3389/fpsyg.2024.1290359] [Medline: 38784631]
6. Sharma A, Sharma M. Charting the evolution of research on the paradoxes in consumer choice—systematic review and future research agenda. Int J Consumer Studies 2024 Jan;48(1):e12996. [doi: 10.1111/ijcs.12996]
7. Jacob BM, Thomas S, Joseph J. Over two decades of research on choice overload: an overview and research agenda. Int J Consumer Studies 2024 Mar;48(2):e13029. [doi: 10.1111/ijcs.13029]
8. Altendorf MB, van Weert JCM, Hoving C, Smit ES. Should or could? testing the use of autonomy-supportive language and the provision of choice in online computer-tailored alcohol reduction communication. Digit Health 2019;5:2055207619832767. [doi: 10.1177/2055207619832767] [Medline: 30834136]
9. Norman P, Webb TL, Millings A, Pechey L. Does the structure (tunneled vs. free-roam) and content (if-then plans vs. choosing strategies) of a brief online alcohol intervention effect engagement and effectiveness? a randomized controlled trial. Transl Behav Med 2019 Nov 25;9(6):1122-1130. [doi: 10.1093/tbm/ibz110] [Medline: 31287897]
10. Smit ES, Zeidler C, Resnicow K, de Vries H. Identifying the most autonomy-supportive message frame in digital health communication: a 2x2 between-subjects experiment. J Med Internet Res 2019 Oct 30;21(10):e14074. [doi: 10.2196/14074] [Medline: 31670693]
11. Abaluck J, Gruber J. When less is more: improving choices in health insurance markets. Rev Econ Stud 2023 May 5;90(3):1011-1040. [doi: 10.1093/restud/rdac050]
12. Brown JA, Oikawa M, Rose JP, Haught HM, Oikawa H, Geers AL. Choosing across cultures: the effect of choice complexity on treatment outcomes. Behavioral Decision Making 2015 Dec;28(5):515-528. [doi: 10.1002/bdm.1868] [Medline: 25620847]
13. Beymer PN, Thomson MM. The effects of choice in the classroom: is there too little or too much choice? Support for Learning 2015 May;30(2):105-120. [doi: 10.1111/1467-9604.12086]
14. Deci EL, Ryan RM. Self-determination theory. In: Handbook of Theories of Social Psychology 2012, Vol. 1:416-436.
15. Kahn B, Moore WL, Glazer R. Experiments in constrained choice. J Consum Res 1987 Jun;14(1):96. [doi: 10.1086/209096]
16. Carlisle S, Ayling K, Jia R, Buchanan H, Vedhara K. The effect of choice interventions on retention-related, behavioural and mood outcomes: a systematic review with meta-analysis. Health Psychol Rev 2022 Jun;16(2):220-256. [doi: 10.1080/17437199.2021.1962386] [Medline: 34423744]
17. Baumol WJ, Ide EA. Variety in retailing. Manage Sci 1956 Oct;3(1):93-101. [doi: 10.1287/mnsc.3.1.93]
18. Mussa M, Rosen S. Monopoly and product quality. J Econ Theory 1978 Aug;18(2):301-317. [doi: 10.1016/0022-0531(78)90085-6]
19. Chen B, Vansteenkiste M, Beyers W, Soenens B, Petegem S. Autonomy in family decision making for Chinese adolescents: disentangling the dual meaning of autonomy. J Cross-Cult Psychol 2013;44(7):1184-1209. [doi: 10.1177/0022022113480038]
20. Xiao H. Independence and obedience: an analysis of child socialization values in the United States and China. J Comp Fam Stud 1999 Dec 1;30(4):641-657. [doi: 10.3138/jcfs.30.4.641]
21. Huang D, Yang LH, Pescosolido BA. Understanding the public's profile of mental health literacy in China: a nationwide study. BMC Psychiatry 2019 Jan 14;19(1):20. [doi: 10.1186/s12888-018-1980-8] [Medline: 30642305]
22. Wong DFK, Cheng CW, Zhuang XY, et al. Comparing the mental health literacy of Chinese people in Australia, China, Hong Kong and Taiwan: implications for mental health promotion. Psychiatry Res 2017 Oct:258-266. [doi: 10.1016/j.psychres.2017.06.032] [Medline: 28648988]
23. Yu BC, Mak WW. Unpacking cultural influences on stigma of people with mental illness between group-oriented and individual-oriented cultures. In: The Cambridge Handbook of Stigma and Mental Health: Cambridge University Press; 2022:263-281. [doi: 10.1017/9781108920995.016]
24. Mogilner C, Rudnick T, Iyengar SS. The mere categorization effect: how the presence of categories increases choosers' perceptions of assortment variety and outcome satisfaction. J Consum Res 2008 Aug;35(2):202-215. [doi: 10.1086/588698]
25. Reutskaja E, Iyengar S, Fasolo B, Misuraca R. Cognitive and affective consequences of information and choice overload. In: Viale R, editor. In Routledge Handbook of Bounded Rationality: Routledge; 2020:625-636. [doi: 10.4324/9781315658353]

XSL•FO
RenderX

26. Iyengar SS, Lepper MR. When choice is demotivating: can one desire too much of a good thing? J Pers Soc Psychol 2000 Dec;79(6):995-1006. [doi: 10.1037//0022-3514.79.6.995] [Medline: 11138768]
27. Schwartz B. The paradox of choice. In: Positive Psychology in Practice: Promoting Human Flourishing in Work, Health, Education, and Everyday Life 2015, Vol. 3:121-138. [doi: 10.1002/9781118996874]
28. Simon HA. Bounded rationality in social science: today and tomorrow. Mind & Society 2000 Mar;1(1):25-39. [doi: 10.1007/BF02512227]
29. Gourville JT, Soman D. Overchoice and assortment type: when and why variety backfires. Marketing Science 2005 Aug;24(3):382-395. [doi: 10.1287/mksc.1040.0109]
30. Scheibehenne B, Greifeneder R, Todd PM. Can there ever be too many options? A meta-analytic review of choice overload. J Consum Res 2010 Oct;37(3):409-425. [doi: 10.1086/651235]
31. Polman E. Effects of self-other decision making on regulatory focus and choice overload. J Pers Soc Psychol 2012 May;102(5):980-993. [doi: 10.1037/a0026966] [Medline: 22429272]
32. Hu X, Turel O, Chen W, Shi J, He Q. The effect of trait-state anxiety on choice overload: the mediating role of choice difficulty. Decision (Wash D C) 2023 Jun;50(2):143-152. [doi: 10.1007/s40622-023-00345-0]
33. Garg N, Inman JJ, Mittal V. Emotion effects on choice deferral: the moderating role of outcome and process accountability. European Journal of Marketing 2017 Sep 12;51(9/10):1631-1649. [doi: 10.1108/EJM-12-2015-0861]
34. Tsao SD, McKay D. Behavioral avoidance tests and disgust in contamination fears: distinctions from trait anxiety. Behav Res Ther 2004 Feb;42(2):207-216. [doi: 10.1016/S0005-7967(03)00119-0] [Medline: 14975781]
35. Corfield EC, Martin NG, Nyholt DR. Co-occurrence and symptomatology of fatigue and depression. Compr Psychiatry 2016 Nov;71:1-10. [doi: 10.1016/j.comppsych.2016.08.004] [Medline: 27567301]
36. Calabrese JR, Fava M, Garibaldi G, et al. Methodological approaches and magnitude of the clinical unmet need associated with amotivation in mood disorders. J Affect Disord 2014 Oct;168:439-451. [doi: 10.1016/j.jad.2014.06.056]
37. Ellis HC, Ashbrook PW. Resource allocation model of the effects of depressed mood states on memory. In: Affect, Cognition, and Social Behavior 1988.
38. Kircanski K, Joormann J, Gotlib IH. Cognitive aspects of depression. Wiley Interdiscip Rev Cogn Sci 2012;3(3):301-313. [doi: 10.1002/wcs.1177] [Medline: 23240069]
39. Ottenbreit ND, Dobson KS. Avoidance and depression: the construction of the cognitive-behavioral avoidance scale. Behav Res Ther 2004 Mar;42(3):293-313. [doi: 10.1016/S0005-7967(03)00140-2] [Medline: 14975771]
40. Jones NP, Siegle GJ, Muelly ER, Haggerty A, Ghinassi F. Poor performance on cognitive tasks in depression: doing too much or not enough? Cognitive, Affective, & Behavioral Neuroscience 2010 Mar 1;10(1):129-140. [doi: 10.3758/CABN.10.1.129]
41. Major B, Cozzarelli C, Cooper ML, et al. Psychological responses of women after first-trimester abortion. Arch Gen Psychiatry 2000 Aug;57(8):777-784. [doi: 10.1001/archpsyc.57.8.777] [Medline: 10920466]
42. Jelinek L, Schröder J, Bücker L, et al. Improving adherence to unguided Internet-based therapy for obsessive-compulsive disorder: a randomized controlled trial. J Affect Disord 2023 Oct 1;338:472-481. [doi: 10.1016/j.jad.2023.06.038] [Medline: 37348657]
43. Valentine L, McEnery C, O'Sullivan S, et al. Young people's experience of online therapy for first-episode psychosis: a qualitative study. Psychol Psychother 2022 Mar;95(1):155-172. [doi: 10.1111/papt.12356] [Medline: 34252267]
44. Higgins ET. Beyond pleasure and pain. Am Psychol 1997 Dec;52(12):1280-1300. [doi: 10.1037//0003-066x.52.12.1280] [Medline: 9414606]
45. Tuan Pham M, Chang HH. Regulatory focus, regulatory fit, and the search and consideration of choice alternatives. J Consum Res 2010 Dec;37(4):626-640. [doi: 10.1086/655668]
46. Shah J, Higgins ET. Expectancy × value effects: regulatory focus as determinant of magnitude and direction. J Pers Soc Psychol 1997;73(3):447-458. [doi: 10.1037//0022-3514.73.3.447]
47. Idson LC, Liberman N, Higgins ET. Distinguishing gains from nonlosses and losses from nongains: a regulatory focus perspective on hedonic intensity. J Exp Soc Psychol 2000 May;36(3):252-274. [doi: 10.1006/jesp.1999.1402]
48. Feldman G, Baumeister RF, Wong KFE. Free will is about choosing: the link between choice and the belief in free will. J Exp Soc Psychol 2014 Nov;55:239-245. [doi: 10.1016/j.jesp.2014.07.012]
49. Lindhiem O, Bennett CB, Trentacosta CJ, McLear C. Client preferences affect treatment satisfaction, completion, and clinical outcome: a meta-analysis. Clin Psychol Rev 2014 Aug;34(6):506-517. [doi: 10.1016/j.cpr.2014.06.002] [Medline: 25189522]
50. Swift JK, Mullins RH, Penix EA, Roth KL, Trusty WT. The importance of listening to patient preferences when making mental health care decisions. World Psychiatry 2021 Oct;20(3):316-317. [doi: 10.1002/wps.20912] [Medline: 34505382]
51. van 't Veer AE, Giner-Sorolla R. Pre-registration in social psychology—a discussion and suggested template. J Exp Soc Psychol 2016 Nov;67:2-12. [doi: 10.1016/j.jesp.2016.03.004]
52. Higgins ET, Friedman RS, Harlow RE, Idson LC, Ayduk ON, Taylor A. Achievement orientations from subjective histories of success: promotion pride versus prevention pride. Euro J Social Psych 2001 Jan;31(1):3-23. [doi: 10.1002/ejsp.27]
53. Lockwood P, Jordan CH, Kunda Z. Motivation by positive or negative role models: regulatory focus determines who will best inspire us. J Pers Soc Psychol 2002 Oct;83(4):854-864. [doi: 10.1037/0022-3514.83.4.854]

XSL•FO
RenderX

54. Benoit ID, Miller EG. The mitigating role of holistic thinking on choice overload. J Consum Mark 2017 May 8;34(3):181-190. [doi: 10.1108/JCM-07-2016-1889]
55. Holm S. A simple sequentially rejective multiple test procedure. Scand J Stat 1979 Jan 1;6(2):65-70. [doi: 10.2307/4615733]
56. Rubel JA, Zilcha-Mano S, Giesemann J, Prinz J, Lutz W. Predicting personalized process-outcome associations in psychotherapy using machine learning approaches—a demonstration. Psychother Res 2020 Apr 2;30(3):300-309. [doi: 10.1080/10503307.2019.1597994]
57. Yamada Y, Ćepulić DB, Coll-Martín T, et al. COVIDiSTRESS Global Survey dataset on psychological and behavioural consequences of the COVID-19 outbreak. Sci Data 2021 Jan 4;8(1):3. [doi: 10.1038/s41597-020-00784-9] [Medline: 33398078]
58. Azevedo F, Pavlović T, Rêgo GG, et al. Social and moral psychology of COVID-19 across 69 countries. Sci Data 2023 May 11;10(1):272. [doi: 10.1038/s41597-023-02080-8] [Medline: 37169799]
59. Blackburn AM, Vestergren S, Blackburn AM. COVIDiSTRESS diverse dataset on psychological and behavioural outcomes one year into the COVID-19 pandemic. Sci Data 2022 Jun 21;9(1):331. [doi: 10.1038/s41597-022-01383-6] [Medline: 35729305]
60. Niles AN, Loerinc AG, Krull JL, et al. Advancing personalized medicine: application of a novel statistical method to identify treatment moderators in the coordinated anxiety learning and management study. Behav Ther 2017 Jul;48(4):490-500. [doi: 10.1016/j.beth.2017.02.001] [Medline: 28577585]
61. Cacioppo JT, Petty RE, Kao CF. The efficient assessment of need for cognition. J Pers Assess 1984 Jun;48(3):306-307. [doi: 10.1207/s15327752jpa4803_13] [Medline: 16367530]
62. Hadar L, Sood S. When knowledge Is demotivating: subjective knowledge and choice overload. Psychol Sci 2014 Sep;25(9):1739-1747. [doi: 10.1177/0956797614539165] [Medline: 25037963]
63. Choi I, Koo M. Individual differences in analytic versus holistic thinking. Pers Soc Psychol Bull 2007 May;33(5):691-705. [doi: 10.1177/0146167206298568] [Medline: 17440200]
64. Vallacher RR, Wegner DM. Levels of personal agency: Individual variation in action identification. J Pers Soc Psychol ;57(4):660-671. [doi: 10.1037/0022-3514.57.4.660] [Medline: 2926623]
65. Nenkov GY, Morrin M, Ward A, Schwartz B, Hulland J. A short form of the Maximization Scale: factor structure, reliability and validity studies. Judgm decis mak 2008 Jun;3(5):371-388. [doi: 10.1017/S1930297500000395]
66. Misuraca R, Faraci P, Gangemi A, Carmeci FA, Miceli S. The decision making tendency inventory: a new measure to assess maximizing, satisficing, and minimizing. Pers Individ Dif 2015 Oct;85:111-116. [doi: 10.1016/j.paid.2015.04.043]
67. Diehl K, Poynor C. Great expectations?! Assortment size, expectations, and satisfaction. Journal of Marketing Research 2010 Apr;47(2):312-322. [doi: 10.1509/jmkr.47.2.312]
68. Cohen J. Statistical Power Analysis for the Behavioral Sciences: Routledge; 1988.
69. Lakens D. Sample size justification. Collabra Psychol 2022 Mar 22;8(1):33267. [doi: 10.1525/collabra.33267]
70. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [doi: 10.3758/bf03193146] [Medline: 17695343]
71. Sela A, Berger J, Liu W. Variety, vice, and virtue: how assortment size influences option choice. J Consum Res 2009 Apr;35(6):941-951. [doi: 10.1086/593692]
72. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
73. Löwe B, Decker O, Müller S, et al. Validation and standardization of the generalized anxiety disorder screener (GAD-7) in the general population. Med Care 2008 Mar;46(3):266-274. [doi: 10.1097/MLR.0b013e318160d093] [Medline: 18388841]
74. Plummer F, Manea L, Trepel D, McMillan D. Screening for anxiety disorders with the GAD-7 and GAD-2: a systematic review and diagnostic metaanalysis. Gen Hosp Psychiatry 2016;39:24-31. [doi: 10.1016/j.genhosppsych.2015.11.005] [Medline: 26719105]
75. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613. [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
76. Martin A, Rief W, Klaiberg A, Braehler E. Validity of the brief patient health questionnaire mood scale (PHQ-9) in the general population. Gen Hosp Psychiatry 2006;28(1):71-77. [doi: 10.1016/j.genhosppsych.2005.07.003] [Medline: 16377369]
77. Sun Y, Fu Z, Bo Q, Mao Z, Ma X, Wang C. The reliability and validity of PHQ-9 in patients with major depressive disorder in psychiatric hospital. BMC Psychiatry 2020 Sep 29;20(1):474. [doi: 10.1186/s12888-020-02885-6] [Medline: 32993604]
78. Russell JA. A circumplex model of affect. J Pers Soc Psychol 1980 Dec;39(6):1161-1178. [doi: 10.1037/h0077714] [Medline: 7431205]
79. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. J Pers Soc Psychol 1988 Jun;54(6):1063-1070. [doi: 10.1037//0022-3514.54.6.1063] [Medline: 3397865]
80. Gomez P, Borges A, Pechmann C(. Avoiding poor health or approaching good health: does it matter? the conceptualization, measurement, and consequences of health regulatory focus. J Consum Psychol 2013 Oct;23(4):451-463. [doi: 10.1016/j.jcps.2013.02.001]
81. Rakos RF, Laurene KR, Skala S, Slane S. Belief in free will: measurement and conceptualization innovations. Behav Soc Iss 2008 May;17(1):20-40. [doi: 10.5210/bsi.v17i1.1929]

XSL•FO
RenderX

82. Cacioppo JT, Petty RE. The need for cognition. J Pers Soc Psychol 1982 Jan;42(1):116-131. [doi: 10.1037/0022-3514.42.1.116]
83. Linardon J. Can acceptance, mindfulness, and self-compassion be learned by smartphone apps? a systematic and meta-analytic review of randomized controlled trials. Behav Ther 2020 Jul;51(4):646-658. [doi: 10.1016/j.beth.2019.10.002] [Medline: 32586436]
84. Mak WW, Tong AC, Yip SY, et al. Efficacy and moderation of mobile app-based programs for mindfulness-based training, self-compassion training, and cognitive behavioral psychoeducation on mental health: randomized controlled noninferiority trial. JMIR Ment Health 2018 Oct 11;5(4):e60. [doi: 10.2196/mental.8597] [Medline: 30309837]
85. Mak WWS, Tong ACY, Fu ACM, et al. Efficacy of internet-based rumination-focused cognitive behavioral therapy and mindfulness-based intervention with guided support in reducing risks of depression and anxiety: a randomized controlled trial. Appl Psychol Health Well Being 2024 May;16(2):696-722. [doi: 10.1111/aphw.12512] [Medline: 38073271]
86. Choice_Empowerment_Choice_Overload_Study_First_Survey_-_revised.qsf. Open Science Framework. URL: https://osf.io/rgm25 [accessed 2025-11-28]
87. Joinson A. Social desirability, anonymity, and internet-based questionnaires. Behav Res Methods Instrum Comput 1999 Aug;31(3):433-438. [doi: 10.3758/bf03200723] [Medline: 10502866]
88. Larson RB. Controlling social desirability bias. International Journal of Market Research 2019 Sep;61(5):534-547. [doi: 10.1177/1470785318805305]
89. Delacre M, Leys C, Mora YL, Lakens D. Taking parametric assumptions seriously: arguments for the use of Welch's F-test instead of the classical F-test in one-way ANOVA. International Review of Social Psychology 2019 Aug 1;32(1):13. [doi: 10.5334/irsp.198]
90. García-Pérez MA. Use and misuse of corrections for multiple testing. Methods in Psychology 2023 Nov;8:100120. [doi: 10.1016/j.metip.2023.100120]
91. Bender R, Lange S. Adjusting for multiple testing--when and how? J Clin Epidemiol 2001 Apr;54(4):343-349. [doi: 10.1016/s0895-4356(00)00314-0] [Medline: 11297884]
92. Ranstam J. Multiple P-values and Bonferroni correction. Osteoarthritis Cartilage 2016 May;24(5):763-764. [doi: 10.1016/j.joca.2016.01.008] [Medline: 26802548]
93. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B 1995 Jan 1;57(1):289-300. [doi: 10.1111/j.2517-6161.1995.tb02031.x]
94. Menyhart O, Weltz B, Győrffy B. MultipleTesting.com: a tool for life science researchers for multiple hypothesis testing correction. PLoS One 2021;16(6):e0245824. [doi: 10.1371/journal.pone.0245824] [Medline: 34106935]
95. Şahi˙n M, Aybek E. Jamovi: an easy to use statistical software for the social scientists. International Journal of Assessment Tools in Education ;6(4):670-692. [doi: 10.21449/ijate.661803]
96. RStudio team. RStudio: Integrated Development Environment for R. 2024. URL: http://www.rstudio.com/ [accessed 2025-11-28]
97. Hayes AF. Introduction to mediation, moderation, and conditional process analysis: a regression-based approach. In: Guilford Press 2022.
98. Reutskaja E, Cheek NN, Iyengar S, Schwartz B. Choice deprivation, choice overload, and satisfaction with choices across six nations. J Int Mark 2022 Sep;30(3):18-34. [doi: 10.1177/1069031X211073821]
99. Chernev A. Decision focus and consumer choice among assortments. J Consum Res 2006 Jun;33(1):50-59. [doi: 10.1086/504135]
100. Park JY, Jang S(. Confused by too many choices? Choice overload in tourism. Tour Manag 2013 Apr;35:1-12. [doi: 10.1016/j.tourman.2012.05.004]
101. Park S, Eves A. Choice overload in tourism: moderating roles of hypothetical and social distance. Journal of Travel Research 2024 Sep;63(7):1626-1641. [doi: 10.1177/00472875231197379]
102. Thai NT, Yuksel U. Choice overload in holiday destination choices. IJCTHR 2017 Mar 6;11(1):53-66. [doi: 10.1108/IJCTHR-09-2015-0117]
103. Fey CF, Hu T, Delios A. The measurement and communication of effect sizes in management research. Manag Organ Rev 2023 Feb;19(1):176-197. [doi: 10.1017/mor.2022.2]
104. Li X, Liu J, Gao W, Cohen GL. Challenging the N-heuristic: effect size, not sample size, predicts the replicability of psychological science. PLoS ONE 2024 Aug 23;19(8):e0306911. [doi: 10.1371/journal.pone.0306911]
105. Chaiken S, Ledgerwood A. A theory of heuristic and systematic information processing. In: AW K, ET H, editors. Handbook of Theories of Social Psychology: SAGE Publications; 2012, Vol. 1:246-266. [doi: 10.4135/9781446249215.n13]
106. Petty RE, Cacioppo JT. The elaboration likelihood model of persuasion. In: Communication and Persuasion: Springer, New York, NY; 1986:1-24. [doi: 10.1007/978-1-4612-4964-1_1]
107. Hodkinson C. 'Fear of Missing Out' (FOMO) marketing appeals: a conceptual model. Journal of Marketing Communications 2019 Jan 2;25(1):65-88. [doi: 10.1080/13527266.2016.1234504]
108. Milyavskaya M, Saffran M, Hope N, Koestner R. Fear of missing out: prevalence, dynamics, and consequences of experiencing FOMO. Motiv Emot 2018 Oct;42(5):725-737. [doi: 10.1007/s11031-018-9683-5]

XSL•FO
RenderX

109. LeBel EP, McCarthy RJ, Earp BD, Elson M, Vanpaemel W. A unified framework to quantify the credibility of scientific findings. Advances in Methods and Practices in Psychological Science 2018 Sep;1(3):389-402. [doi: 10.1177/2515245918787489]
110. Hornstein S, Zantvoort K, Lueken U, Funk B, Hilbert K. Personalization strategies in digital mental health interventions: a systematic review and conceptual framework for depressive symptoms. Front Digit Health 2023;5:1170002. [doi: 10.3389/fdgth.2023.1170002] [Medline: 37283721]
111. Bollen D, Knijnenburg BP, Willemsen MC, Graus M. Understanding choice overload in recommender systems. 2010 Sep 26 Presented at: Proceedings of the Fourth ACM Conference on Recommender Systems p. 63-70. [doi: 10.1145/1864708.1864724]
112. Liao P, Klasnja P, Tewari A, Murphy SA. Sample size calculations for micro-randomized trials in mHealth. Stat Med 2016 May 30;35(12):1944-1971. [doi: 10.1002/sim.6847] [Medline: 26707831]
113. Walton A, Nahum-Shani I, Crosby L, Klasnja P, Murphy S. Optimizing digital integrated care via micro-randomized trials. Clin Pharmacol Ther 2018 Jul;104(1):53-58. [doi: 10.1002/cpt.1079] [Medline: 29604043]
114. Hong SM, Faedda S. Refinement of the Hong Psychological Reactance Scale. Educ Psychol Meas 1996 Feb;56(1):173-182. [doi: 10.1177/0013164496056001014]
115. Grant AM, Schwartz B. Too much of a good thing: the challenge and opportunity of the inverted U. Perspect Psychol Sci 2011 Jan;6(1):61-76. [doi: 10.1177/1745691610393523] [Medline: 26162116]
116. Liu L, Zheng Y, Chen R. Better with more choices? impact of choice set size on variety seeking. Acta Psychologica Sinica 2015;47(1):66. [doi: 10.3724/SP.J.1041.2015.00066]
117. Reutskaja E, Hogarth RM. Satisfaction in choice as a function of the number of alternatives: when “goods satiate”. Psychology and Marketing 2009 Mar;26(3):197-203. [doi: 10.1002/mar.20268]
118. Fleming T, Bavin L, Lucassen M, Stasiak K, Hopkins S, Merry S. Beyond the trial: systematic review of real-world uptake and engagement with digital self-help interventions for depression, low mood, or anxiety. J Med Internet Res 2018 Jun 6;20(6):e199. [doi: 10.2196/jmir.9275] [Medline: 29875089]
119. Lipschitz JM, Pike CK, Hogan TP, Murphy SA, Burdick KE. The engagement problem: a review of engagement with digital mental health interventions and recommendations for a path forward. Curr Treat Options Psychiatry 2023 Sep;10(3):119-135. [doi: 10.1007/s40501-023-00297-3] [Medline: 38390026]
120. Huibers MJH, Cohen ZD, Lemmens LHJM, et al. Predicting optimal outcomes in cognitive therapy or interpersonal psychotherapy for depressed individuals using the personalized advantage index approach. PLoS One 2015;10(11):e0140771. [doi: 10.1371/journal.pone.0140771] [Medline: 26554707]
121. OpenAI. Poe AcademicWriterBot [Large language model]. 2024 Sep 3. URL: https://poe.com/ [accessed 2025-11-28]

Abbreviations

GAD-7: Generalized Anxiety Disorder-7
HRFS: Health Regulatory Focus Scale
OR: odds ratio
OSF: Open Science Framework
PDMC: Perceived Difficulty of Making Choices
PFC: Preference for Choices
PHQ-9: Patient Health Questionnaire-9

Edited by A Kushniruk, J Sarvestan; submitted 16.01.25; peer-reviewed by R Kornfield, R Chen; accepted 30.10.25; published 15.12.25.

Please cite as:
Yeung SK, Leung FHT, Lee JCM, Cheung GMH, Li CW, Mak WWS
Effects of Choice Set Sizes and Moderations of Anxiety and State Emotions on Mental Health Self-Care Uptake, Engagement, and User Experience: Experimental Study
JMIR Hum Factors 2025;12:e71165
URL: https://humanfactors.jmir.org/2025/1/e71165
doi:10.2196/71165

© Siu Kit Yeung, Florence H T Leung, Jason C M Lee, Gabriel Man Hin Cheung, Ching Wan Li, Winnie W S Mak. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 15.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly

XSL•FO
RenderX

cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Designing Digital Mental Health Support for Paramedics Exposed to Trauma: Qualitative Study of Lived Experiences and Design Preferences

Nicola Cogan, PhD; Spence Whittaker, MSc; Ashleigh Craig, MSc; Lucy Milligan, MSc; Robyn McCluskey, MSc; Tara Burns, MSc; Alison Kirk, PhD; Susan Rasmussen, PhD; William Hodgson, MRes

Department of Psychological Sciences and Health, University of Strathclyde, Graham Hills Building, Glasgow, United Kingdom

Corresponding Author:
Nicola Cogan, PhD
Department of Psychological Sciences and Health, University of Strathclyde, Graham Hills Building, Glasgow, United Kingdom

Abstract

Background: Paramedics face frequent exposure to trauma and intense occupational stress, often under conditions of limited psychological support and ongoing stigma. Digital mental health interventions have the potential to offer accessible, confidential, and tailored support. However, their acceptability and design must be informed by the lived experiences of paramedics to ensure effectiveness.

Objective: This study aimed to explore the experiences of trauma exposure among UK paramedics in the workplace and their views on the design and delivery of digital mental health interventions.

Methods: Semi-structured interviews were conducted with 22 UK paramedics. Participants were recruited through purposive and snowball sampling. Interviews were transcribed verbatim and analyzed using reflexive thematic analysis. Ethical approval was obtained, and trauma-informed principles were applied throughout data collection and analysis.

Results: Five key themes were identified: (1) It Has to Feel Easy to Use*:* highlighting the need for digital tools that reduce cognitive burden and are accessible during unpredictable shifts; (2) Make It Fit My Needs: calling for interventions specifically designed for paramedics, with lived-experience-informed language and delivery; (3) We Need to Talk to Each Other: describing a strong desire for peer connection while recognizing barriers such as stigma and shift pressures; (4) I Need to Know It's Safe: emphasizes the importance of anonymity, data privacy, and psychological safety; and (5) Support Needs to Feel Human: reinforcing the value of integrating digital tools with human connection and professional services. Participants expressed strong support for an app-based solution that offers anonymity, rapid accessibility, and flexibility, while preserving opportunities for human interaction.

Conclusions: Paramedics face unique mental health challenges that are not adequately addressed by existing services. Digital mental health tools offer promise if they are carefully co-designed to reflect the realities of frontline work. Anonymity, usability, peer connection, and integration with existing support systems are critical to engagement. These findings offer actionable insights for the development of trauma-informed, context-sensitive digital mental health interventions for emergency service workers.

(JMIR Hum Factors 2025;12:e76158) doi:10.2196/76158

KEYWORDS

paramedics; trauma; digital mental health; first responders; qualitative research; usability; peer support; occupational stress

Introduction

Frontline workers, including paramedics, firefighters, police officers, and emergency health care staff, operate in some of the most high-pressure environments, frequently encountering life-threatening emergencies, traumatic incidents, and high-stress situations [1,2]. The cumulative exposure to critical incidents places these workers at significant risk of developing mental health conditions such as posttraumatic stress disorder, anxiety, depression, and burnout [3,4]. While all frontline workers experience occupational stress, research consistently shows that paramedics report some of the poorest mental health outcomes among emergency responders, including elevated levels of psychological distress, emotional exhaustion, and suicidal ideation compared to their counterparts in firefighting and law enforcement [5,6].

Paramedics face unique challenges that exacerbate their vulnerability to mental health difficulties [7]. They frequently operate under relentless workloads with little opportunity for recovery between challenging and often traumatic calls. Paramedic work requires responding to unpredictable situations with limited resources at any hour of the day. In these high-pressure, uncontrolled settings, paramedics are often

XSL•FO
RenderX

required to make complex decisions while also managing the effects of sleep deprivation and disrupted circadian rhythms caused by shift work [8,9]. The unpredictability of their role, such as never knowing the severity of an incident before arriving on the scene, further compounds stress levels. Additionally, paramedics often work in isolated environments, providing care in ambulances, homes, or public settings without immediate access to team support [10]. This lack of structured debriefing opportunities can lead to cumulative emotional distress, as paramedics must frequently suppress their emotions to remain professional in high-stakes situations [11].

Another significant challenge is moral injury, a psychological distress experienced when actions taken in the line of duty conflict with personal values or ethical standards [12]. Paramedics frequently witness unavoidable suffering, resource limitations, and cases where they are unable to save a life, leading to profound feelings of guilt, self-doubt, and helplessness [13]. Furthermore, paramedics often provide end-of-life care in uncontrolled environments, making their experiences emotionally intense and personally distressing [14]. This repeated exposure to human suffering, combined with high-performance expectations and limited emotional support, places them at heightened risk of long-term psychological harm [15].

In addition to physical trauma cases, paramedics are frequently dispatched to mental health-related emergencies, including suicide attempts and death by suicide, self-harm incidents, and acute psychiatric crises [16]. These callouts present unique stressors, as paramedics must navigate complex interactions with individuals in extreme distress, often with limited mental health training or specialist support [17,18]. Attending suicide-related incidents has been shown to significantly impact paramedics' mental well-being, particularly when required to interact with grieving families or manage high-risk patients without adequate resources [19]. The emotional toll of these callouts, combined with the stigma surrounding mental health within emergency services, contributes to burnout, compassion fatigue, and secondary traumatic stress among paramedics [20].

Another critical issue paramedics face is exposure to workplace violence and assaults. Paramedics frequently encounter aggressive patients, intoxicated individuals, or bystanders who may become physically or verbally abusive [21]. This occupational hazard is further exacerbated by working alone or in unpredictable environments where immediate backup may not be available. Exposure to workplace violence not only increases the risk of physical injury but also significantly contributes to long-term psychological distress, fear, and hypervigilance [22]. Studies have shown that repeated exposure to aggression and violence leads to elevated rates of posttraumatic stress disorder, anxiety, and depressive symptoms among paramedics, further compounding the already high levels of occupational stress they experience [10,23,24]. Despite policies aimed at protecting emergency responders, many paramedics feel unsupported when reporting incidents of violence, contributing to a culture of normalization and underreporting [25].

Despite these challenges, many paramedics do not seek professional mental health support due to stigma, fear of judgment, and concerns about confidentiality [26]. Prior research has identified that paramedics are less likely than other emergency responders to disclose psychological distress, fearing that such admissions could lead to professional consequences, social isolation, or doubts about their fitness to work [11]. Traditional face-to-face mental health support services, while available, remain underused due to perceived stigma, logistical barriers, and a culture that often prioritizes resilience over emotional well-being [27].

Such issues highlight the urgent need for tailored mental health support for paramedics [28]. Digital mental health interventions present a promising avenue for overcoming barriers to traditional face-to-face support by offering accessible, confidential, and flexible options for managing psychological distress [29]. However, for such interventions to be effective, they must be designed with direct input from end users to ensure relevance and usability [30]. Cocreation is increasingly recognized as a critical component in the development of digital mental health interventions, ensuring they are relevant, engaging, and responsive to the needs of the end user [31]. By involving paramedics in the design process, digital interventions can be tailored to align with their unique work environments, stressors, and coping mechanisms. Research has shown that interventions developed in collaboration with users have higher adoption rates and greater long-term efficacy compared to those designed without user input [32]. For paramedics, cocreation is particularly vital as they face distinct challenges such as exposure to trauma, high-pressure decision-making, and limited opportunities for structured mental health support. Generic digital interventions may fail to address the realities of their work, leading to disengagement and limited effectiveness. Additionally, cocreation fosters a sense of ownership and trust in the intervention, reducing potential stigma associated with help-seeking. When end users see their insights reflected in the final digital intervention, they are more likely to perceive the intervention as relevant and beneficial to their well-being [33]. Given the demanding and often unpredictable nature of paramedic work, it is crucial that digital mental health interventions are not only evidence-based but also practical and person-centered.

This study aims to explore paramedics' perspectives on digital mental health interventions and identify key design considerations to inform the development of tailored support systems [34]. It seeks to understand paramedics' experiences with workplace trauma and their perspectives on the acceptability of a digital mental health intervention. The research also aims to contribute to the growing body of literature on digital mental health interventions for frontline workers. This study addresses the following research questions:

1. What are paramedics' experiences with workplace trauma, and how do they currently manage their mental health?
2. What features and functionalities do paramedics consider essential for a digital mental health intervention?

XSL•FO
RenderX

Methods

Study Design

This study used a qualitative design using in-depth, semi-structured interviews to explore paramedics' experiences with workplace trauma and their perspectives on digital mental health interventions. A qualitative approach was chosen to generate detailed insights into participants' mental health needs and intervention preferences, allowing for a deeper understanding of lived experiences that quantitative methods may not capture. Grounded in a constructivist epistemology, this approach emphasizes the coconstruction of meaning between researchers and participants and acknowledges the active role of the researcher in interpreting the data. It aligns with the participatory and trauma-informed principles underpinning this study, ensuring sensitivity to the emotional content of interviews and valuing the voices of participants throughout the analytic process [35].

Recruitment

Recruitment was conducted through professional networks, paramedic organizations, and web-based forums to ensure a diverse sample that captured a broad range of experiences within emergency medical services. A combination of purposive and snowball sampling was used. Purposive sampling enabled the targeted recruitment of participants with relevant professional backgrounds and trauma exposure, while snowball sampling allowed existing participants to refer colleagues with similar experiences, broadening the reach and diversity of the sample. Recruitment strategies included direct invitations disseminated via paramedic organizations, participant referrals through word-of-mouth, and digital flyers shared on professional social media platforms. Individuals were eligible to participate if they resided in the United Kingdom, were aged 18 years or older, had been employed in a paramedic role for at least 6 months, and self-identified as having experienced trauma in the workplace at the time of this study. Interested participants were provided with an electronic participant information sheet detailing this study's purpose, procedures, and ethical safeguards. Following written informed consent, interviews were scheduled at times convenient to each participant.

All interviews were conducted via Microsoft Teams to ensure accessibility, data security, and consistency in data collection [36]. A total of 22 paramedics working in emergency medical services were recruited through purposive and snowball sampling [37]. This sample size was sufficient as qualitative research prioritizes depth, richness, and variation in experiences over numerical representativeness [38]. Studies on first responders and trauma have drawn meaningful conclusions from similar sample sizes [39-41], demonstrating that a focused cohort can yield transferable, contextually grounded insights [42]. Purposive sampling ensured that participants had direct lived experience relevant to the research aims, while snowball sampling enhanced inclusivity by recruiting those hesitant to engage in trauma-related research [43,44].

Data Collection

A semi-structured interview guide was developed to facilitate in-depth exploration of participants' experiences, drawing on previous research related to occupational trauma and mental health in emergency services [45,46]. The guide was designed to prompt discussion on key areas, including encounters with workplace trauma, access to and experiences with existing mental health support, perceived organizational and cultural barriers to seeking help, and preferences for digital mental health interventions tailored to paramedic needs. Questions were open-ended and designed to elicit rich, narrative data while allowing for flexibility based on individual experiences. The guide was refined iteratively during the early phases of data collection, with minor modifications made to reflect emerging themes and to enhance sensitivity to the language and challenges specific to the paramedic context. This adaptive approach ensured the interviews remained responsive to participants' perspectives while maintaining consistency across core thematic areas. All interviews were conducted remotely using Microsoft Teams, allowing for flexible scheduling and increased accessibility for participants working variable shift patterns. This web-based format also aligned with trauma-informed practices by enabling participants to choose a safe and private environment in which to share their experiences.

Interviews were conducted over 18 months and ranged in length from approximately 40 to 75 minutes (mean 58.2, SD 10.2). All interviews were conducted remotely, allowing participants to join from a private setting of their choice. The final sample size (n=22) was determined based on the richness, relevance, and diversity of the data, consistent with the principles of reflexive thematic analysis [42]. Rather than aiming for data saturation, we drew on the concept of information power to guide sample sufficiency, prioritizing depth, variation, and the capacity of the data to meaningfully address the research questions. The dataset was judged to provide rich, nuanced insights into paramedics' lived experiences of workplace trauma and preferences for digital support, with recurring patterns and interpretative depth evident across the interviews.

Ethical Considerations

This study forms part of a broader program of research exploring trauma exposure and psychological support needs among frontline workers in the United Kingdom. It specifically focuses on the views and experiences of paramedics regarding the acceptability and design of digital mental health interventions, with distinct aims and analyses from prior related research [28]. Given the sensitive nature of workplace trauma, ensuring the well-being and safeguarding of participants was a key ethical consideration. This study received ethical approval from the University of Strathclyde Ethics Committee (UEC22/92). Informed consent was obtained before participation, and ethical approval was granted by the institutional review board. Participants were informed of their right to withdraw at any stage without repercussions. Data were anonymized and stored securely, per GDPR and institutional data protection policies, to protect participant identity and confidentiality. Participants were offered a £20 (US $23.29) gift voucher as a thank-you for their time and contribution to the interviews. Additionally,

XSL•FO
RenderX

signposting to mental health support services was provided to all participants, ensuring access to professional help if discussing traumatic experiences triggered distress [47]. Researching trauma-related topics requires careful ethical handling, and efforts were made to create a psychologically safe space where participants could share their experiences without fear of judgment or harm [48]. This study also emphasized the importance of participatory methods and lived experience perspectives in shaping digital mental health interventions. By centering paramedics' voices, this study aligns with the principles of participatory research, ensuring that digital interventions are not only theoretically sound but also practically applicable to frontline workers' needs. Additionally, researcher well-being was considered, given the emotionally demanding nature of trauma-related research. The research team received specific training, led by a Health and Care Professions Council–registered clinical psychologist, on conducting sensitive research and applying trauma-informed principles. This included prioritizing self-care, reflexive journaling throughout the research process, and recognizing signs of secondary trauma exposure. To further safeguard well-being, interviewers were limited to 1 interview per day, followed by a structured debrief with the research lead. Additionally, a structured debriefing process was implemented to provide access to professional supervision and peer support [49,50]. Reflexive meetings were held regularly throughout the research process, fostering a team approach to addressing challenges, ensuring ethical consistency, and supporting both participant and researcher well-being [51,52].

Analysis

All interviews were one-to-one and were audio-recorded and transcribed in full verbatim. The data were then managed in NVivo (version 12, QSR International) and analyzed using a reflexive thematic approach, following the framework by Braun and Clarke [53], which is particularly suited for studies exploring complex psychological and social phenomena [54]. This approach builds upon the earlier thematic analysis framework of Braun and Clarke [55], adopting a more explicitly reflexive and interpretative approach. It emphasizes the researcher's active role in meaning-making [56]. Themes are not viewed as pre-existing within the data but are constructed through deep engagement and iterative interpretation [57]. This began with the process of familiarization and immersion within the data, generating initial codes organically and fluidly. Initial codes were logged, with records of how each theme evolved through discussion and refinement. Regular team meetings were held to review and challenge coding decisions, ensuring that the analysis was grounded in the data while allowing for flexibility in interpretation. Any discrepancies in theme development were resolved through collaborative discussion in the research team, ensuring that final themes were coconstructed rigorously and reflexively [58].

A value-based approach to qualitative research was applied, ensuring a commitment to ethical transparency, reflexivity, and participant-centered analysis [59]. This framework emphasizes the importance of researcher subjectivity and the co-construction of knowledge, aligning with the study's focus on participatory research and lived experience perspectives. Recognizing the limitations of rigid, universal evaluative tools for qualitative research, the Big Q Qualitative Reporting Guidelines were considered to ensure that methodological choices aligned with qualitative research values rather than externally imposed standards rooted in postpositivist traditions [59]. This framework prioritizes methodological congruence, reflexivity, and the interactive nature of knowledge production. To ensure transparency and rigor in the research process, an audit trail was maintained throughout data collection, analysis, and interpretation. This involved systematically documenting all key decisions made at each stage of this study, providing a structured record of the rationale behind methodological choices and analytic interpretations.

Results

Participant Characteristics

The final sample consisted of 22 paramedics. Participants ranged in age from 23 to 59 years (mean 38.22, SD 9.28), with professional experience ranging from 1 to 38 years (mean 13.11, SD 6.93). Thirteen (59%) participants identified as male and 9 (41%) as female. The majority (82%, n=18) identified as White British, while 4 (18%) participants identified as belonging to ethnic minority backgrounds. Most participants were based in England (64%, n=14), with the remainder in Scotland (36%, n=8). All participants self-identified as having experienced occupational trauma that had negatively impacted their mental health and well-being. See Table 1 for a full breakdown of demographic characteristics.

XSL•FO
RenderX

Table . Participant characteristics. Names are pseudonyms. Ethnicity categories adapted from UK Office for National Statistics classifications.

Name	Gender	Age (years)	Service (years)	Ethnicity
Billy	Male	30	5	White British
Harry	Male	34	10	White Scottish
Gillian	Female	47	21	White British
Garry	Male	51	10	White Scottish
Mike	Male	40	2	White Scottish
Jan	Male	28	3	White Other (Polish)
Nathan	Male	50	21	White Irish
Terry	Female	44	17	White Scottish
Sam	Male	51	15	White British
William	Male	39	5	White Scottish
James	Male	37	12	White British
Rachel	Female	36	11	White Scottish
Paul	Male	45	20	White Irish
Laura	Female	33	8	White Scottish
Jane	Female	48	25	White British
Emma	Female	29	6	White British
David	Male	41	14	White Scottish
Sarah	Female	42	19	White British
Leigh	Female	46	23	White Scottish
Eric	Male	50	18	White Other (Polish)
Rebecca	Female	35	9	White British
Luke	Male	38	16	White British

Themes

The themes presented below reflect the diverse and often deeply personal ways in which paramedics make sense of their experiences of occupational trauma. Given the subjective and multifaceted nature of psychological trauma, participants were invited to self-define their experiences, enabling the analysis to capture a broad spectrum of emotional and psychological responses. This approach respected individual interpretations of what constitutes trauma and ensured that the analysis remained grounded in participants' own meanings and language. Five key themes emerged from the analysis, capturing the experiences and preferences of paramedics regarding digital mental health interventions. These themes illustrate their needs for usability, personalization, social support, privacy, and human connection (Table 2).

XSL•FO
RenderX

Table . Summary of themes identified in the thematic analysis.

Theme	Description
It Has to Feel Easy to Use—Navigating Acceptability and Usability	• Intervention must be simple and intuitive. • Should reduce cognitive load postshift. • Quick access is crucial for usability.
Make It Fit My Needs—The Importance of Customization and Personalization	• Tools should be tailored to paramedic experiences. • Generic interventions were seen as ineffective. • Relevance drives engagement.
We Need to Talk to Each Other—Social Connection and Peer Support	• Peer support is highly valued. • Digital tools can enable safe, structured connection. • Reduces isolation and builds validation.
I Need to Know It's Safe—Confidentiality, Privacy, and Safety	• Concerns about stigma and data breaches. • Security and trust are prerequisites for use. • Anonymity must be guaranteed.
Support Needs to Feel Human—Integration With Existing Services	• Digital tools should complement, not replace, human support. • Facilitate access to professional help. • Include guided check-ins and support pathways.

Theme 1: It Has to Feel Easy to Use—Navigating Acceptability and Usability

Acceptability was closely linked to usability for many participants, who emphasized that any digital intervention must feel simple, nonintrusive, and easy to engage with during or after demanding shifts. Paramedics described the cognitive burden of their work, noting that complex or time-consuming tools would be impractical in high-pressure environments. Many highlighted that their workload and unpredictable schedules often leave little energy for engaging with structured mental health support.

> *Getting five minutes to breathe is rare now. We used to have time to gather ourselves, but not anymore.* [Rachel]

> *You deal with something traumatic, but before you've processed it, you're already at the next job.* [Luke]

These accounts reflect the high workload pressures paramedics face, with limited opportunities for decompression or reflection. Digital interventions, therefore, were seen as acceptable only if they were intuitive, low-effort, and seamlessly integrated into the rhythm of shift-based work.

> *You can't go through life just dealing with trauma on adrenaline… I have to have a way of bringing that down and managing that.* [Paul]

Some paramedics reported becoming emotionally desensitized over time, reflecting a gradual numbing that made it harder to recognize when support was needed.

> *It's trauma after trauma. Eventually it just becomes normal, and that's when you realise you've stopped feeling it.* [Leigh]

This emotional distancing may contribute to reduced help-seeking unless support is easily accessible and presented in a way that feels informal and manageable.

Participants expressed that traditional face-to-face mental health support often felt inaccessible, both logistically and psychologically, due to stigma and confidentiality concerns. In contrast, digital interventions were viewed as potentially acceptable because of their on-demand nature and discretion.

> *I wouldn't go to a therapist, even if I needed to. It's just not what we do. But if I could check in with something that doesn't feel like a big deal, I'd probably use it.* [Nathan]

> *Going through work to get help feels risky. But if I can do it quietly, that changes everything.* [Rebecca]

A key factor influencing acceptability was that the intervention should not require additional cognitive or emotional effort, particularly at the end of long or emotionally draining shifts.

> *If I've had a long shift, I don't want something that's going to take brain power. It has to be quick and easy.* [Jane]

Some participants also saw digital tools as a proactive way to manage stress before it became overwhelming, valuing short, in-the-moment strategies.

> *Even a short breathing exercise or prompt could stop it spiralling after a tough job.* [Laura]

In sum, usability was seen as foundational to acceptability. Tools that were easy to navigate, unobtrusive, and offered brief, practical support were more likely to be welcomed and used.

Theme 2: Make It Fit My Needs—the Importance of Customization and Personalization

Beyond usability, participants consistently emphasized the importance of personal relevance and individual fit in shaping how acceptable an intervention would feel. In this context, acceptability was not just about simplicity or ease of use, but about feeling seen, understood, and supported in a way that resonated with paramedics' specific experiences. When discussing the acceptance and needs for an appropriate

XSL•FO
RenderX

intervention, participants noted the importance of "specific" [Rachel] help, comprising needs unique to paramedics. It was noted that others "just don't get it" [Leigh].

Participants felt that credible support should be tailored to paramedics specifically, as they felt they are "different from other first responders" [Eric] and "other members of society" [Rebecca]. Participants stated that creators of an intervention should have a strong knowledge and preferably, experience of life as a paramedic, which can be evidenced by comments such as:

> *You need people designing this who understand our reality. Otherwise, it just won't connect.* [Nathan]
>
> *There's no substitute for lived experience. We spot it straight away when something feels off.* [Emma]

Participants viewed themselves as different from other occupations, given the nature of their working environment, and therefore stated they required an intervention "that's tailored for an individual" [Leigh]. All participants stated that they would be more likely to speak of mental health concerns and seek help if they could speak to someone with "real insight into the job" [Rachel].

Additionally, participants suggested they were more likely to access help if it could be "completely anonymous" [Eric], given the stigmatization prevalent within the occupation:

> *If people don't know it's you, you're far more likely to be honest.* [David]

Regarding anonymity, all participants suggested that an identifiable method of intervention would prevent them from accessing help. Overall, acceptability was high across participants for an intervention, with all participants stating they would be accepting of an anonymous and specific intervention:

> *Even something small, like a quick check-in after a shift, if it's anonymous and for us, I'd use it.* [Laura]
>
> *If nothing changes, we're just going to lose more people. This kind of support has to be built in.* [Rachel]

Participants reflected the timely need for an intervention comprising the detailed characteristics to prevent "experienced people from burning out and leaving for good" [Emma]. Hence, all participants suggested that a digital intervention in app form would be beneficial, given that it could be specific to them, constantly accessible, and confidential, as voiced by Eric:

> *Having it on our phones would make it feel like a normal part of the job (pauses) something we all use, not something secret or awkward.* [Eric]

Furthermore, in discussing potential technological barriers, participants described digital tools as familiar and well-integrated into their working lives, with 1 noting it was "second nature now" [Sam]. As such, acceptability was not about the presence of technology per se, but about how well it aligned with paramedics' specific needs and values:

> *We use tech every day. It's not strange anymore. If anything, it makes sense to use it for our mental health too.* [Rebecca]

Participants emphasized that interventions were more likely to be accepted if they could be tailored based on personal needs, emotional state, or shift patterns. This capacity for customization fostered a sense of personal relevance and ownership, making the intervention feel trustworthy and worth engaging with:

> *You're more likely to keep going back if it feels like it's built for you.* [Sarah]

Theme 3: We Need to Talk to Each Other—Social Connection and Peer Support

A strong theme that emerged was the need for social connection, particularly among colleagues who understand the unique pressures of the role. Participants spoke about how workplace trauma is often managed informally through discussions with peers, yet increasing work demands have reduced these opportunities. For instance, 1 paramedic shared how important it was to debrief but noted that opportunities to do so were limited:

> *There are fewer and fewer chances to talk things through. That's when you start carrying things on your own.* [Luke]

However, some expressed concerns about the diminishing support culture in their work environment:

> *The support bit, it's just not there.* [Leigh]

The need for peer support was reinforced by the recognition that sharing experiences with colleagues who "really get it" [Mike] provides comfort and validation:

> *If you're involved, first responders and stuff, you have an understanding of what you've been through* [David]

Despite this, some participants highlighted a reluctance to open up about their struggles, especially within a "macho culture" that discouraged vulnerability:

> *You certainly wouldn't admit it to your colleagues. It was a stiff upper lip culture, you know, just get on with it* [Terry]

Given these concerns, digital interventions were seen as a potential way to facilitate peer connection in a safe and anonymous environment. Some participants suggested that a digital platform could provide structured opportunities to connect with others who have shared similar experiences, while still allowing individuals to maintain privacy:

> *I wouldn't want to sit in a room and talk about it, but if there was something online where you could connect with people who've been through the same, that would be different.* [Rebecca]

Others highlighted the potential for digital interventions to offer moderated peer forums or guided group discussions, ensuring that support remains professional and constructive:

> *It would have to be somewhere safe, not just a big free-for-all chat where people bring each other down. Maybe something that's monitored, or with prompts to help guide discussions.* [James]

XSL•FO
RenderX

The ability to share experiences without fear of judgment was seen as a crucial element in increasing engagement with peer support mechanisms. Digital tools that allow for anonymous peer check-ins or shared reflections were regarded as particularly valuable:

> *Sometimes you don't need someone to fix it, you just need to know someone else has been through it too. If an app could help with that, I'd use it.* [Luke]

Theme 4: I Need to Know It's Safe—Confidentiality, Privacy, and Safety

A major concern for participants was data security and confidentiality. Many reported that stigma and fears of professional consequences prevented them from seeking help. The pressure to appear resilient and the fear of being judged or deemed unfit for their role created barriers to accessing mental health support. Participants highlighted that any digital intervention must prioritize absolute anonymity to encourage engagement.

> *There's a fear. There's a big fear… About our professional registration, because you've got a mental health condition… because if you speak, you're fired.* [Gillian]

> *There are people with the attitude that if you suffer from mental health problems, then this is obviously not a job for you… you're clearly not cut out for it.* [Garry]

Ensuring anonymity was seen as crucial, as paramedics were concerned that seeking support through formal channels might lead to professional repercussions or workplace stigma:

> *You want to be sure it's completely anonymous. That's the only way I'd open up.* [Jan]

Digital interventions were perceived as a potential solution, offering a confidential and secure platform where paramedics could access support without fear of being identified. Participants suggested that digital tools should include features such as encrypted communication, anonymous logins, and discreet access options to protect their privacy. Additionally, some paramedics expressed the need for clear information on how their data would be stored and who would have access to it:

> *You'd have to be upfront about what happens with any data. If I'm putting anything in, I want to know it's not going anywhere else.* [Laura]

Several participants described how the long-term effects of trauma can be overwhelming, making it even harder to seek help. They acknowledged that while digital interventions could not replace professional therapy, they could provide an important first step for those who are hesitant to engage with traditional mental health services. By incorporating stringent privacy protections and offering an anonymous space for processing emotions, digital interventions were viewed as potentially bridging the gap for paramedics who might otherwise avoid seeking mental health support due to concerns about stigma and confidentiality.

Theme 5: Support Needs to Feel Human—Integration With Existing Services

Although participants expressed a willingness to engage with digital interventions, they stressed that technology should not replace human-led mental health support. Many emphasized that digital tools should be used as a supplement rather than a substitute for direct peer or professional support. While digital interventions offer accessibility and confidentiality, participants acknowledged the value of human connection in processing trauma and stress. One paramedic highlighted how the lack of structured support left them feeling exposed:

> *Sometimes it's just a few words from a person that makes all the difference. The app can't do it all.* [Sarah]

Participants expressed a strong need for interventions that offer structured and ongoing support, ensuring that paramedics feel genuinely seen and supported in their mental health journey. They recognized that digital tools could help fill gaps in existing support services, but were cautious that such tools must integrate seamlessly with professional and peer-based systems.

> *I think having a digital app for trauma is the right thing. I think building such an app, it's something that needs to be built for the people that you're wanting to use it.* [Harry]

Many paramedics emphasized that for a digital intervention to be truly beneficial, it should facilitate bridging the gap between self-help and professional support. They suggested that digital tools could provide structured pathways to connect users with trained professionals when needed, ensuring that those who require more intensive intervention are not left without options:

> *I'd use something like a digital app if I knew there was a way to reach out to a real person if I needed it. Just knowing that's there makes a difference.* [James]

Another key aspect raised by participants was the need for human elements within the digital space to create a sense of connection and validation. Some suggested that interventions could incorporate recorded messages from experienced paramedics or psychologists to guide them through difficult moments:

> *If I could hear from someone who's been in my shoes, even if it's just pre-recorded guidance, that would feel more real than just reading generic advice.* [David]

Participants also highlighted the need for structured check-ins and follow-ups to ensure ongoing support, rather than a 1-time or reactive approach:

> *It shouldn't just be something you use once when you're in crisis. It should check in with you, like, "How are you doing?" and offer reminders to engage with it.* [Emma]

These insights reinforce the idea that while digital interventions can provide essential support and accessibility, they should not function in isolation. Instead, they should act as an entry point into broader mental health care, offering both immediate coping

XSL•FO
RenderX

strategies and seamless integration with existing services. By incorporating human-centered elements such as guided support, personalized check-ins, and links to professional care, digital tools can become a valuable and sustainable resource for paramedics navigating workplace trauma.

Discussion

Principal Findings

This study explored paramedics' experiences of workplace trauma, their approaches to managing mental health, and their perspectives on the acceptability and design preferences for digital mental health interventions. Thematic analysis identified 5 key themes that offer novel and context-specific insights into how digital tools can be designed to address the unique challenges faced by paramedics. These findings contribute to a growing body of research advocating for trauma-informed, user-centered approaches in digital mental health interventions for emergency service workers.

First, participants emphasized that digital interventions need to be easy to use, with a simple, intuitive design that does not add to their cognitive load after stressful shifts. They highlighted the importance of quick access and ease of navigation, aligning with previous research indicating that digital mental health interventions must be designed for seamless usability to encourage engagement [34].

Second, participants expressed a strong preference for customization and personalization, with a tool that adapts to their individual coping styles, schedules, and emotional states. They noted that existing mental health resources often fail to address the distinct challenges of their role, reinforcing findings that user-centered digital interventions lead to higher adherence and satisfaction [60]. Participants also emphasized that their work schedules are unpredictable, meaning the intervention should be flexible, allowing them to engage with support at a time that suits them rather than being restricted to set hours or appointments.

Third, participants highlighted the importance of social connection and peer support, valuing the opportunity to connect with colleagues who understand the unique pressures of their role. They suggested that digital interventions should facilitate safe and structured peer interactions, aligning with research demonstrating the critical role of peer support in promoting resilience among first responders [61-63]. The ability to share experiences with others who have been through similar situations was seen as particularly beneficial, as it reduced the sense of isolation that many paramedics reported experiencing.

Fourth, concerns around confidentiality, privacy, and safety were significant, with participants voicing worries about data security and the need for absolute privacy to overcome fears of stigma or professional consequences. Many preferred an anonymous platform to reduce concerns about disclosure, echoing previous studies that highlight the importance of privacy in emergency workers' engagement with mental health interventions [11]. They expressed particular concerns about their professional registration and career security, with some fearing that seeking support through formal occupational health routes could have negative consequences. A digital intervention designed with strong privacy protections was seen as a way to reduce these concerns and encourage engagement. However, with the need for anonymity comes an ethical requirement to ensure adequate support and safeguarding. Recent research on safeguarding of users of mental health apps has emphasized the importance of ensuring that platforms clearly highlight their limitations and capacities concerning identifying risk and referring individuals to outside care opportunities, in addition to any built-in interventions [64].

Finally, participants stressed the need for digital interventions to complement, rather than replace, human-led mental health support. While digital tools offer accessibility and discretion, participants recognized the value of professional guidance and structured follow-ups to ensure sustained engagement. Research supports a blended approach that integrates digital mental health tools with professional and peer-based support systems to maximize effectiveness [65]. Participants noted that digital interventions could serve as an important first step in seeking help, but that access to human support, whether through structured debriefs, therapy, or informal check-ins, remained a crucial component of effective mental health care. This aligns with earlier work whereby emergency workers have emphasized the importance of personal engagement in their training and education [66].

Strengths, Limitations, and Future Directions

This study makes a valuable contribution to the existing body of knowledge by providing in-depth insights into paramedics' experiences of workplace trauma and their perspectives on digital mental health interventions. A key strength of this study lies in the use of robust qualitative methodology, which enabled the generation of rich, contextually grounded findings. However, this study has some limitations. The sample consisted of 22 paramedics, which, while appropriate for qualitative research, may not fully capture the diverse experiences across different ambulance services and geographic locations. Additionally, participants self-selected into this study, which may have resulted in the inclusion of individuals who are more willing to engage in discussions about mental health, potentially limiting the generalizability of findings. Future research should aim to recruit a larger and more demographically diverse sample to explore a broader range of perspectives and experiences. Particular attention should be given to increasing representation of ethnically and culturally diverse paramedics, as this is essential not only for producing inclusive and representative findings but also for informing the effective transcreation and localization of digital interventions. This will contribute to more accessible, relevant, and equitable interventions, ultimately reducing disparities in mental health outcomes and supporting health equity across emergency service settings.

While purposive and snowball sampling allowed the research team to access a diverse range of paramedics with direct experience of workplace trauma, it is acknowledged that there is potential for self-selection bias. Participants who chose to take part may have been more open to discussing mental health or more comfortable with technology, compared to those who did not volunteer. As such, the sample may not fully reflect the

XSL•FO
RenderX

perspectives of paramedics who are more reluctant to engage in conversations about mental health or who may be more skeptical about digital interventions. This should be considered when interpreting the findings, particularly concerning the anticipated uptake of digital tools across the broader paramedic workforce.

Another limitation is that this study focused solely on paramedics, whereas other emergency responders, such as firefighters and police officers, also experience similar mental health challenges. Comparative studies across different frontline professions would help identify both shared and profession-specific needs for digital interventions. Additionally, this study explored user preferences but did not evaluate an actual digital intervention. The research team is currently undertaking a program of work focused on the co-design, development, and real-world testing of a digital tool, Sentinel, to evaluate its feasibility, usability, and impact on mental health outcomes among first responders [67]. By incorporating paramedic feedback throughout the development process, Sentinel aims to address usability, personalization, peer connectivity, and confidentiality concerns. The intervention will be piloted with paramedics to evaluate its effectiveness and integration into existing mental health support systems. Furthermore, customized versions of Sentinel will be expanded to support other frontline professionals, including firefighters, police officers, and emergency health care staff, recognizing both the common and unique stressors they encounter. Additionally, the platform will offer resources for the families of frontline workers, acknowledging their vital role in providing support and the importance of addressing their specific needs to help foster stronger community and familial resilience [68]. Future studies will examine how these adaptations enhance mental health support across various high-pressure occupational settings, leading to a scalable, evidence-based intervention for frontline workers [69].

The findings of the current study have clear implications for ambulance services and policymakers seeking to improve the mental health of frontline staff. Services should consider investing in digital tools that are co-designed with paramedics, offer confidentiality, and are accessible outside of traditional working hours. Importantly, such tools should not operate in isolation but be embedded within wider organizational strategies that promote psychological safety and reduce stigma around help-seeking. Embedding trauma-informed support into routine practice, through education, peer-support systems, and leadership engagement, can help foster a culture where mental health conversations are normalized and proactively supported. Policy makers can play a key role in endorsing the development and adoption of such interventions across emergency services, ensuring that funding mechanisms and regulatory guidance align with the psychological needs of high-risk occupational groups.

While future technological innovations hold promise, the immediate priority lies in developing usable, human-centered digital support grounded in the lived experience of paramedics. Artificial intelligence, wearable data, and real-time analytics may offer longer-term opportunities to enhance personalization, tailor content delivery, and monitor distress signals [70]. However, any integration of such technologies must be approached cautiously, with strong attention to ethical considerations, including data privacy, informed consent, and algorithmic bias [71]. Crucially, future research must ensure that artificial intelligence–driven tools remain grounded in cocreation and usability principles, as interventions designed with direct user input are more likely to be trusted and effective in real-world settings [72,73].

To support scalability and implementation, structured frameworks such as RE-AIM (Reach, Effectiveness, Adoption, Implementation, and Maintenance), the UK Medical Research Council guidance on complex interventions, or the technology acceptance model could be used to guide development and evaluation [74-78]. These frameworks can help developers attend to issues of accessibility, human interaction, and long-term sustainability, ensuring that digital interventions are both usable and adaptable to diverse frontline professions. Ultimately, while innovation in digital health is advancing rapidly, this study underscores the critical value of listening to paramedics themselves. Future work should continue to center on the voices of frontline workers to ensure that digital mental health tools are responsive, acceptable, and fit for purpose in the demanding realities of emergency service work.

Conclusions

This study highlights the importance of designing digital mental health interventions that are user-friendly, personalized, confidential, and integrated with existing support services. The findings provide critical insights into the specific needs of paramedics, particularly concerning their frequent exposure to workplace trauma, and outline key design considerations for digital mental health tools. Given the high levels of trauma-related psychological distress among paramedics, tailored digital solutions have the potential to offer accessible, flexible, and trauma-informed support. Future research should prioritize cocreating and rigorously evaluating digital interventions with paramedics and other emergency responders to ensure real-world usability and clinical relevance. By integrating end user feedback and addressing the psychological impact of trauma directly, digital mental health tools can enhance engagement, promote well-being, and provide frontline workers with timely support to navigate the emotional demands of their work. Through ongoing collaboration and refinement, trauma-informed digital interventions have the potential to transform the accessibility and effectiveness of mental health care for paramedics and other frontline professionals.

Acknowledgments

We would like to thank all the paramedics who participated in this study for sharing their experiences and insights.

XSL•FO
RenderX

Data Availability

The datasets generated or analyzed during this study are not publicly available due to the sensitive nature of the qualitative interview content and confidentiality agreements with participants. Anonymized data may be made available to qualified researchers upon reasonable request to the corresponding author, subject to approval by the relevant ethics committee and completion of a data sharing agreement.

Conflicts of Interest

NC is the Chief Scientific Officer of Sentinel, serves on its executive board, and holds a minority equity stake in the company. She leads the clinical strategy and research direction of the platform. AK and WH are also members of the Sentinel research and development team. Although Sentinel is referenced in the future directions of this paper, it was not evaluated in this study. All aspects of study design, data collection, and qualitative analysis were conducted independently of any commercial interests. The authors provide this disclosure in the interest of transparency. All other authors declare no competing interests.

References

1. Arjmand HA, O'Donnell ML, Putica A, et al. Mental health treatment for first responders: an assessment of mental health provider needs. Psychol Serv 2024 Aug;21(3):489-499. [doi: 10.1037/ser0000832] [Medline: 38166296]
2. Billings J, Ching BCF, Gkofa V, Greene T, Bloomfield M. Experiences of frontline healthcare workers and their views about support during COVID-19 and previous pandemics: a systematic review and qualitative meta-synthesis. BMC Health Serv Res 2021 Sep 6;21(1):923. [doi: 10.1186/s12913-021-06917-z] [Medline: 34488733]
3. Smith E, Dean G, Holmes L. Supporting the mental health and well-being of first responders from career to retirement: a scoping review. Prehosp Disaster Med 2021 Aug;36(4):475-480. [doi: 10.1017/S1049023X21000431] [Medline: 33928892]
4. Vizheh M, Qorbani M, Arzaghi SM, Muhidin S, Javanmard Z, Esmaeili M. The mental health of healthcare workers in the COVID-19 pandemic: a systematic review. J Diabetes Metab Disord 2020 Dec;19(2):1967-1978. [doi: 10.1007/s40200-020-00643-9]
5. Mars B, Hird K, Bell F, James C, Gunnell D. Suicide among ambulance service staff: a review of coroner and employment records. Br Paramed J 2020 Mar 1;4(4):10-15. [doi: 10.29045/14784726.2020.12.4.4.10] [Medline: 33456374]
6. Liberati E, Richards N, Parker J, et al. Remote care for mental health: qualitative study with service users, carers and staff during the COVID-19 pandemic. BMJ Open 2021 Apr 22;11(4):e049210. [doi: 10.1136/bmjopen-2021-049210] [Medline: 33888531]
7. Duffee B, Willis DB. Paramedic perspectives of job stress: qualitative analysis of high-stress, high-stakes emergency medical situations. Soc Sci Med 2023 Sep;333:116177. [doi: 10.1016/j.socscimed.2023.116177]
8. Hirello LM, Drummond SPA, Bowles KA, Wolkow AP. The impact of shift work on paramedics and their practice: protocol for a simulated paramedic shift work study. PLoS ONE 2025;20(3):e0319569. [doi: 10.1371/journal.pone.0319569] [Medline: 40138286]
9. Marvin G, Schram B, Orr R, Canetti EFD. Occupation-induced fatigue and impacts on emergency first responders: a systematic review. Int J Environ Res Public Health 2023 Nov 12;20(22):7055. [doi: 10.3390/ijerph20227055] [Medline: 37998287]
10. Lawn S, Roberts L, Willis E, Couzner L, Mohammadi L, Goble E. The effects of emergency medical service work on the psychological, physical, and social well-being of ambulance personnel: a systematic review of qualitative research. BMC Psychiatry 2020 Jul 3;20(1):348. [doi: 10.1186/s12888-020-02752-4] [Medline: 32620092]
11. Auth NM, Booker MJ, Wild J, Riley R. Mental health and help seeking among trauma-exposed emergency service staff: a qualitative evidence synthesis. BMJ Open 2022 Feb 2;12(2):e047814. [doi: 10.1136/bmjopen-2020-047814] [Medline: 35110304]
12. Lentz LM, Smith-MacDonald L, Malloy D, Carleton RN, Brémault-Phillips S. Compromised conscience: a scoping review of moral injury among firefighters, paramedics, and police officers. Front Psychol 2021;12:639781. [doi: 10.3389/fpsyg.2021.639781] [Medline: 33868111]
13. Murray E. Moral injury and paramedic practice. J Paramedic Pract 2019 Oct 2;11(10):424-425. [doi: 10.12968/jpar.2019.11.10.424]
14. Juhrmann ML, Butow PN, Platts CM, Simpson P, Boughey M, Clayton JM. "It breaks a narrative of paramedics, that we're lifesavers": a qualitative study of health professionals', bereaved family members' and carers' perceptions and experiences of palliative paramedicine. Palliat Med 2023 Sep;37(8):1266-1279. [doi: 10.1177/02692163231186451] [Medline: 37452564]
15. Farr-Wharton B, Brunetto Y, Hernandez-Grande A, Brown K, Teo S. Emergency service workers: the role of policy and management in (re)shaping wellbeing for emergency service workers. Rev Public Pers Adm 2023 Dec;43(4):774-793. [doi: 10.1177/0734371X221130977]
16. Lyra RD, McKenzie SK, Jenkin G, Every-Palmer S. 'How can you not be traumatised': experience of paramedics occupationally exposed to suicide. Paramedicine 2024;22(2):80-90. [doi: 10.1177/27536386241283986]
17. Bennett R, Mehmed N, Williams B. Non‐technical skills in paramedicine: a scoping review. Nurs Health Sci 2021 Mar;23(1):40-52. [doi: 10.1111/nhs.12765]

XSL•FO
RenderX

18. Emond K, O'Meara P, Bish M. Paramedic management of mental health related presentations: a scoping review. J Ment Health 2019 Feb;28(1):89-96. [doi: 10.1080/09638237.2018.1487534] [Medline: 30269628]
19. Lyra RD, McKenzie SK, Every-Palmer S, Jenkin G. Occupational exposure to suicide: a review of research on the experiences of mental health professionals and first responders. PLoS ONE 2021;16(4):e0251038. [doi: 10.1371/journal.pone.0251038]
20. Renkiewicz GK, Hubble MW. Secondary Traumatic STRESS in Emergency Services Systems (STRESS) project: quantifying and predicting compassion fatigue in emergency medical services personnel. Prehosp Emerg Care 2022 Sep 3;26(5):652-663. [doi: 10.1080/10903127.2021.1943578]
21. Wongtongkam N. An exploration of violence against paramedics, burnout and post-traumatic symptoms in two Australian ambulance services. IJES 2017 Aug 7;6(2):134-146. [doi: 10.1108/IJES-03-2017-0014]
22. Friis K, Pihl-Thingvad J, Larsen FB, Christiansen J, Lasgaard M. Long-term adverse health outcomes of physical workplace violence: a 7-year population-based follow-up study. Eur J Work Organ Psychol 2019 Jan 2;28(1):101-109. [doi: 10.1080/1359432X.2018.1548437]
23. Setlack J, Brais N, Keough M, Johnson EA. Workplace violence and psychopathology in paramedics and firefighters: mediated by posttraumatic cognitions. Can J Behav Sci / Rev Can Sci Comportement 2021;53(3):211-220. [doi: 10.1037/cbs0000240]
24. Straud C, Henderson SN, Vega L, Black R, Van Hasselt V. Resiliency and posttraumatic stress symptoms in firefighter paramedics: the mediating role of depression, anxiety, and sleep. Traumatology (Tallahass Fla) 2018;24(2):140-147. [doi: 10.1037/trm0000142]
25. Mausz J, Donnelly EA. Violence against paramedics: protocol for evaluating 2 years of reports through a novel, point-of-event reporting process. JMIR Res Protoc 2023 Mar 16;12:e37636. [doi: 10.2196/37636] [Medline: 36928257]
26. Mackinnon K, Everett T, Holmes L, Smith E, Mills B. Risk of psychological distress, pervasiveness of stigma and utilisation of support services: exploring paramedic perceptions. Australas J Paramedicine 2020 Jan;17:1-7. [doi: 10.33151/ajp.17.764]
27. Kunyahamu MS, Daud A, Tengku Ismail TA, Md Tahir MF. PsyHELP pocket guide: development of an educational tool to promote professional mental health help-seeking among the health workforce. PLoS ONE 2024;19(10):e0309224. [doi: 10.1371/journal.pone.0309224]
28. Cogan N, Craig A, Milligan L, et al. 'I've got no PPE to protect my mind': understanding the needs and experiences of first responders exposed to trauma in the workplace. Eur J Psychotraumatol 2024 Dec 31;15(1). [doi: 10.1080/20008066.2024.2395113]
29. Bucci S, Schwannauer M, Berry N. The digital revolution and its impact on mental health care. Psychol Psychother 2019 Jun;92(2):277-297. [doi: 10.1111/papt.12222]
30. Cameron G, Mulvenna M, Ennis E, et al. Effectiveness of digital mental health interventions in the workplace: umbrella review of systematic reviews. JMIR Ment Health 2025 Jan 24;12:e67785. [doi: 10.2196/67785] [Medline: 39854722]
31. Brotherdale R, Berry K, Branitsky A, Bucci S. Co-producing digital mental health interventions: a systematic review. Digit Health 2024;10:20552076241239172. [doi: 10.1177/20552076241239172] [Medline: 38665886]
32. Schneider ML, Cha BS, Borghouts J, et al. Individual and organizational outcomes of engaging peers in the cocreation of digital mental health interventions. Psychol Serv. 2024 Aug 22. [doi: 10.1037/ser0000889] [Medline: 39172398]
33. Mimoso I, Figueiredo T, Midão L, et al. Co-creation in the development of digital therapeutics: a narrative review. Int J Environ Res Public Health 2024 Nov 28;21(12):1589. [doi: 10.3390/ijerph21121589] [Medline: 39767430]
34. Borghouts J, Eikey E, Mark G, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. J Med Internet Res 2021 Mar 24;23(3):e24387. [doi: 10.2196/24387] [Medline: 33759801]
35. Muchamore I, Karanikolas P, Gooding P. How lived experience expertise shapes research and development in digital mental health. SSRN J 2024:1-88. [doi: 10.2139/SSRN.4932039]
36. Wakelin KJ, McAra-Couper J, Fleming T. Using an online platform for conducting face-to-face interviews. Int J Qual Methods 2024 Jan;23:1-9. [doi: 10.1177/16094069241234183]
37. Valerio MA, Rodriguez N, Winkler P, et al. Comparing two sampling methods to engage hard-to-reach communities in research priority setting. BMC Med Res Methodol 2016 Oct 28;16(1):146. [doi: 10.1186/s12874-016-0242-z] [Medline: 27793191]
38. Braun V, Clarke V. To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. Qual Res Sport, Exercise Health 2021 Mar 4;13(2):201-216. [doi: 10.1080/2159676X.2019.1704846]
39. Casas JB, Benuto LT. Breaking the silence: a qualitative analysis of trauma narratives submitted online by first responders. Psychol Trauma 2022 Feb;14(2):190-198. [doi: 10.1037/tra0001072] [Medline: 34410810]
40. Nicholson TP, Blazer EC, Hymes AS, Ginley MK. A qualitative investigation into the trauma exhibited by first responders impacted by the opioid epidemic. Int J Ment Health Addiction 2024 Aug;22(4):2391-2412. [doi: 10.1007/s11469-022-00993-w]
41. Phung VH, Trueman I, Togher F, Ørner R, Siriwardena AN. Perceptions and experiences of community first responders on their role and relationships: qualitative interview study. Scand J Trauma Resusc Emerg Med 2018 Feb 5;26(1):13. [doi: 10.1186/s13049-018-0482-5] [Medline: 29402312]
42. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res 2016 Nov;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]

XSL•FO
RenderX

43. Robinson OC. Sampling in interview-based qualitative research: a theoretical and practical guide. Qual Res Psychol 2014 Jan 2;11(1):25-41. [doi: 10.1080/14780887.2013.801543]
44. Naderifar M, Goli H, Ghaljaie F. Snowball sampling: a purposeful method of sampling in qualitative research. Strides Dev Med Educ 2017;14(3):e67670. [doi: 10.5812/sdme.67670]
45. Ford-Jones PC, Daly T. Paramedicine and mental health: a qualitative analysis of limitations to education and practice in Ontario. JMHTEP 2020 Oct 14;15(6):331-345. [doi: 10.1108/JMHTEP-05-2020-0031]
46. Jones S, Agud K, McSweeney J. Barriers and facilitators to seeking mental health care among first responders: “removing the darkness”. J Am Psychiatr Nurses Assoc 2020;26(1):43-54. [doi: 10.1177/1078390319871997] [Medline: 31509058]
47. Williamson E, Gregory A, Abrahams H, Aghtaie N, Walker SJ, Hester M. Secondary trauma: emotional safety in sensitive research. J Acad Ethics 2020;18(1):55-70. [doi: 10.1007/s10805-019-09348-y] [Medline: 32435166]
48. Isobel S. Trauma-informed qualitative research: some methodological and practical considerations. Int J Ment Health Nurs 2021 Oct;30 Suppl 1:1456-1469. [doi: 10.1111/inm.12914] [Medline: 34310829]
49. Cook DA, Hatala R, Brydges R, et al. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. JAMA 2011 Sep 7;306(9):978-988. [doi: 10.1001/jama.2011.1234] [Medline: 21900138]
50. Silverio SA, Sheen KS, Bramante A, et al. Sensitive, challenging, and difficult topics: experiences and practical considerations for qualitative researchers. Int J Qual Methods 2022 Apr;21. [doi: 10.1177/16094069221124739]
51. Alessi EJ, Kahn S. Toward a trauma-informed qualitative research approach: guidelines for ensuring the safety and promoting the resilience of research participants. Qual Res Psychol 2023 Jan 2;20(1):121-154. [doi: 10.1080/14780887.2022.2107967]
52. Shaw RM, Howe J, Beazer J, Carr T. Ethics and positionality in qualitative research with vulnerable and marginal groups. Qual Res 2020 Jun;20(3):277-293. [doi: 10.1177/1468794119841839]
53. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport, Exercise Health 2019 Aug 8;11(4):589-597. [doi: 10.1080/2159676X.2019.1628806]
54. Braun V, Clarke V, Hayfield N, Davey L, Jenkinson E. Doing reflexive thematic analysis. In: Bager-Charleson S, McBeath A, editors. Supporting Research in Counselling and Psychotherapy: Qualitative, Quantitative, and Mixed Methods Research: Springer International Publishing; 2022:19-38. [doi: 10.1007/978-3-031-13942-0_2]
55. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
56. Sekar D, Bhuvaneswari M. Exploring the depths: qualitative research in psychology. In: Methodologies and Ethics for Social Sciences Research: IGI Global Scientific Publishing; 2024:267-292. [doi: 10.4018/979-8-3693-1726-6.ch014]
57. Lim WM. What is qualitative research? An overview and guidelines. Australas Mark J 2025 May;33(2):199-229. [doi: 10.1177/14413582241264619]
58. Campbell KA, Orr E, Durepos P, et al. Reflexive thematic analysis for applied qualitative health research. TQR 2021;26(6):2011-2028. [doi: 10.46743/2160-3715/2021.5010]
59. Braun V, Clarke V. Reporting guidelines for qualitative research: a values-based approach. Qual Res Psychol 2025 Apr 3;22(2):399-438. [doi: 10.1080/14780887.2024.2382244]
60. Matthews P, Rhodes-Maquaire C. Personalisation and recommendation for mental health apps: a scoping review. Behav Inf Technol 2025 Jun 15;44(10):2389-2404. [doi: 10.1080/0144929X.2024.2356630]
61. Fallon P, Jaegers LA, Zhang Y, Dugan AG, Cherniack M, El Ghaziri M. Peer support programs to reduce organizational stress and trauma for public safety workers: a scoping review. Workplace Health Saf 2023 Nov;71(11):523-535. [doi: 10.1177/21650799231194623]
62. Tjin A, Traynor A, Doyle B, Mulhall C, Eppich W, O’Toole M. Turning to “trusted others”: a narrative review of providing social support to first responders. Int J Environ Res Public Health 2022 Dec 8;19(24):16492. [doi: 10.3390/ijerph192416492] [Medline: 36554368]
63. Traynor A, Doyle B, Eppich W, Tjin A, Mulhall C, O’Toole M. “This is it…this is our normal” - the voices of family members and first responders experiencing duty-related trauma in Ireland. Compr Psychiatry 2024 Aug;133:152499. [doi: 10.1016/j.comppsych.2024.152499] [Medline: 38776684]
64. Verbeke K, Jain C, Shpendi A, Borry P. Safeguarding users of consumer mental health apps in research and product improvement studies: an interview study. Neuroethics 2024 Apr;17(1):10. [doi: 10.1007/s12152-024-09543-8]
65. Alden LE, Matthews LR, Wagner S, et al. Systematic literature review of psychological interventions for first responders. Work Stress 2021 Apr 3;35(2):193-215. [doi: 10.1080/02678373.2020.1758833]
66. Riley J, McGowan M, Rozmovits L. Exploring the value of technology to stimulate interprofessional discussion and education: a needs assessment of emergency medicine professionals. J Med Internet Res 2014 Jun 30;16(6):e162. [doi: 10.2196/jmir.3482] [Medline: 24981597]
67. Cogan N, Kirk A, Rodriguez CG. The development and protocol for testing a co-created digital intervention (sentinel) to improve mental wellbeing and help manage and prevent trauma in first responders. PsyArXiv. Preprint posted online on Mar 1, 2024. [doi: 10.31234/osf.io/8j32t]
68. Johnstone H, Cogan N. ‘He messaged me the other night and said you are my saviour’: an interpretative phenomenological analysis of intimate partners’ roles in supporting veterans with mental health difficulties. J Mil, Veteran Fam Health 2021 May 1;7(2):61-70. [doi: 10.3138/jmvfh-2019-0055]

XSL•FO
RenderX

69. Nomeikaite A, Gelezelyte O, Böttche M, Andersson G, Kazlauskas E. Role of tailored timing and frequency prompts on the efficacy of an internet-delivered stress recovery intervention for health care workers: randomized controlled trial. JMIR Ment Health 2025 Jan 28;12:e62782. [doi: 10.2196/62782] [Medline: 39874572]
70. Nicola C. Psychotraumatology and artificial intelligence: a public health approach. PHOA 2024;8(1):1-8. [doi: 10.23880/phoa-16000284]
71. Morgan P, Cogan NA. Using artificial intelligence to address mental health inequalities: co-creating machine learning algorithms with key stakeholders and citizen engagement. JPMH 2025 Jun 19;24(2):150-165. [doi: 10.1108/JPMH-07-2024-0095]
72. Boucher EM, Ward H, Miles CJ, Henry RD, Stoeckl SE. Effects of a digital mental health intervention on perceived stress and rumination in adolescents aged 13 to 17 years: randomized controlled trial. J Med Internet Res 2024 Mar 29;26:e54282. [doi: 10.2196/54282] [Medline: 38551617]
73. Asman O, Torous J, Tal A. Responsible design, integration, and use of generative AI in mental health. JMIR Ment Health 2025;12:e70439-e70439. [doi: 10.2196/70439]
74. Hopkin G, Coole H, Edelmann F, et al. Toward a new conceptual framework for digital mental health technologies: scoping review. JMIR Ment Health 2025;12:e63484-e63484. [doi: 10.2196/63484]
75. Forman J, Heisler M, Damschroder LJ, Kaselitz E, Kerr EA. Development and application of the RE-AIM QuEST mixed methods framework for program evaluation. Prev Med Rep 2017 Jun;6:322-328. [doi: 10.1016/j.pmedr.2017.04.002]
76. Skivington K, Matthews L, Simpson SA, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 2021 Sep 30;374:n2061. [doi: 10.1136/bmj.n2061] [Medline: 34593508]
77. Marikyan D, Papagiannidis S. Technology acceptance model: a review. In: Papagiannidis S, editor. TheoryHub Book 2024.
78. Peddie N, Hoegh J, Rice G, Shetty S, Ure A, Cogan N. Health and social care professionals' experience of psychological safety within their occupational setting: a thematic synthesis review. Nurs Rep 2025 Apr 14;15(4):131. [doi: 10.3390/nursrep15040131] [Medline: 40333056]

Abbreviations

RE-AIM: Reach, Effectiveness, Adoption, Implementation, and Maintenance

Edited by D Gooch; submitted 20.04.25; peer-reviewed by L Paletta, S Burchim; revised version received 14.07.25; accepted 20.07.25; published 12.09.25.

Please cite as:
Cogan N, Whittaker S, Craig A, Milligan L, McCluskey R, Burns T, Kirk A, Rasmussen S, Hodgson W
Designing Digital Mental Health Support for Paramedics Exposed to Trauma: Qualitative Study of Lived Experiences and Design Preferences
JMIR Hum Factors 2025;12:e76158
URL: https://humanfactors.jmir.org/2025/1/e76158
doi:10.2196/76158

© Nicola Cogan, Spence Whittaker, Ashleigh Craig, Lucy Milligan, Robyn McCluskey, Tara Burns, Alison Kirk, Susan Rasmussen, William Hodgson. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 12.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Theory-Based Mobile App Intervention to Promote Healthy Salt Intake Among Adults: Randomized Controlled Trial

Milena Sia Perin[1], MSc, PhD; Maria Cecília Bueno Jayme Gallani[2], MSc, PhD; Eduardo Augusto Fernandes Nilson[3,4], DSc; Titilayo Tatiana Agbadje[5], MSc; Marilia Estevam Cornélio[1], MSc, PhD

[1]Faculdade de Enfermagem, Universidade Estadual de Campinas - UNICAMP, 126 Tessália Vieira de Camargo Street, Cidade Universitária, Campinas, Brazil

[2]Faculté des Sciences Infirmières, Université Laval, Quebec, QC, Canada

[3]Fundação Oswaldo Cruz, Brasilia, Brazil

[4]Universidad Autónoma de Chile, Santiago, Chile

[5]Centre Intégré Universitaire de Santé et Services Sociaux de la Capitale-Nationale - Centre de recherche en santé durable (VITAM), Quebec, QC, Canada

Corresponding Author:
Milena Sia Perin, MSc, PhD
Faculdade de Enfermagem, Universidade Estadual de Campinas - UNICAMP, 126 Tessália Vieira de Camargo Street, Cidade Universitária, Campinas, Brazil

Abstract

Background: Nowadays, mobile health technology has been increasingly used for treatment and prevention at all levels of health care. Associating this technology with the promotion of healthy salt consumption—both cost-effective and cost-saving public health strategies—can reduce this risk factor that contributes significantly to the increase of noncommunicable diseases worldwide.

Objective: We aim to assess the usability and the efficacy of a mobile app intervention—"Sal na Medida" app—on the promotion of a healthy salt intake among adults, based on the Behavior Change Wheel framework. Additionally, to investigate if intention, self-efficacy, and habit variables mediated the effect of intervention on salt intake behavior.

Methods: An experimental, randomized, and longitudinal study that evaluated the intervention effect with assessments at baseline, postintervention (1 mo), and at follow-up (2 mo). Sociodemographic and clinical data were collected from participants recruited at primary health care centers. The behavior of salt intake and per capita salt consumption were the primary outcomes. Psychosocial variables of intention, self-efficacy, and habit were assessed as possible mediators. Usability was evaluated after 1 month and 2 months of using the app.

Results: Eighty-six participants were randomized in the intervention group (IG; n=43) or control group (CG; n=43). Most of the participants were female (IG: n=36, 84% and CG: n=37, 86%). Usability of app intervention scored 77.8 points (on a scale of 0 to 100) among IG participants. There was a significant reduction in salt intake in IG according to the variables of per capita salt and the behavior of salt intake. Furthermore, at the end of the follow-up, individuals in IG were 63% more likely to have a lower salt intake than those in the CG. The regression analysis showed an increase in intention and the perception of self-efficacy, and a more pronounced reduction in the habit of using more than 3 g of salt/day in preparing meals in IG when compared to CG. Habit and self-efficacy were identified as mediators of the intervention's effect.

Conclusions: The theory-based mobile app intervention for reducing salt intake has shown promise both in terms of usability and efficacy among adults. Conducting further studies to assess its potential for implementation on a larger scale would be valuable for determining its real-world impact and feasibility.

Trial Registration: Brazilian Clinical Trials Registry (ReBEC) RBR-4s8qyyq; https://ensaiosclinicos.gov.br/rg/RBR-4s8qyyq

International Registered Report Identifier (IRRID): RR2-10.2196/37853

(JMIR Hum Factors 2025;12:e54174) doi:10.2196/54174

KEYWORDS

sodium, dietary; mHealth; mobile phone; clinical trial; self-efficacy; intention; habits; nursing

XSL•FO
RenderX

Introduction

Background

Recent literature indicates that excessive salt consumption worldwide is responsible for approximately 3 million deaths—more than half of all diet-related deaths—and 70 million disability-adjusted life years (DALYs), which accounts for two-thirds of DALYs [1]. In Brazil, the economic burden is particularly significant, with an estimated annual cost of US$ 102 million due to public hospitalizations [2]. These detrimental outcomes could be prevented or at least mitigated by adhering to the World Health Organization's recommendation of a daily salt intake of 5 grams (roughly 2 grams of sodium) [3]. Additionally, the World Health Organization's global action plan for the prevention and control of noncommunicable diseases 2013‐2020 has prioritized salt reduction as 1 of 9 voluntary global targets [4].

To reduce the problem of salt overconsumption, it is crucial to understand the different sources of salt intake, which can differ across countries and regions. In South Asia, most salt intake comes from added salt during cooking and discretionary use at the table [5]. In the Eastern Mediterranean Region, bread and dairy products are the primary sources of salt consumption [6]. In Europe and North America, processed and ultraprocessed foods contribute the most to salt intake [7]. In Brazil, table salt and salt-based spices account for approximately 74.4% of the total household sodium availability [8], with about 70% of the salt consumed coming from the addition of salt to home-cooked meals [9].

Reducing salt intake in household diets offers significant health benefits [10-12]. A 10% reduction in salt consumption could prevent 5.8 million DALYs annually due to cardiovascular disease, at a population-weighted average cost of US $1.13 per capita [13]. To achieve this reduction in salt intake, various behavioral initiatives are used, including health education interventions, public awareness campaigns, and multicomponent strategies [14].

However, interventions grounded in established theories and frameworks tend to produce more effective health outcomes, since theories allow us to identify effective intervention content and processes through which it is expected to influence behavior, and also to understand the mechanisms by which the interventions work [15]. In this way, the Behavior Change Wheel (BCW) guide [16] has been broadly used to design health interventions due to its comprehensive methodology aimed at fostering behavior change. In the field of eating behavior, the BCW framework has been successfully applied in intervention studies [17,18].

Considering the findings from previous studies on salt intake behaviors in Brazil [8,9,19], which highlighted the excessive use of salt in food preparation, the need to develop an intervention to reduce salt consumption became apparent, which led to the development of a theory-based intervention using a rigorous methodological approach, the BCW, as a guiding model, to ensure the effectiveness of the intervention.

The process of intervention design based on BCW is detailed in a previous study [20]. During the development of the intervention, it was possible to identify, by means of a literature review, that the behavior of salt addition during food preparation was the main source of the overall daily salt intake of adults and elderly Brazilians [8,9,19]. In turn, this behavior was determined by intention, self-efficacy, and habit [21]. To address these determinants and by applying the BCW steps, we were able to identify the behavior change techniques: goal setting (behavior), problem-solving, goal setting (outcome), action planning, commitment, feedback on behavior, self-monitoring of behavior, social support (unspecified), social support (practical), instruction on how to perform the behavior, information about health consequences, demonstration of the behavior, prompt or cues, behavioral practice or rehearsal, credible source, and adding objects to the environment [20].

Another critical step in the intervention development is to identify the most effective method for delivering the intervention. A review of the literature revealed the potential benefits of mobile apps and the efficacy, reliability, and association with improved health outcomes of mobile app-based interventions on eating behaviors [22-24], and specifically in interventions related to salt consumption [25]. This evidence informed our decision to use a mobile app as the delivery method for the intervention [20].

Subsequent to the intervention development process, the next step is to evaluate the effect of the intervention. Therefore, this study shows the results of the efficacy and usability assessment of a mobile phone app intervention to promote healthy salt intake behavior among adults, focusing on behavior change in adding salt during food preparation.

Objective

This study aimed to assess the usability and the efficacy of a mobile app intervention, named "Sal na Medida" app, on the promotion of a healthy salt intake behavior among adults, based on the BCW framework. Additionally, it aimed to investigate if intention, self-efficacy, and habit serve as mediator variables of intervention effects.

Methods

Trial Design

This is a randomized controlled trial, with 2 parallel groups (intervention [IG] and control group [CG]). This study was conducted between October 2021 and August 2022, part of a large study developed in three modules: (1) intervention design, (2) development of "Sal na Medida" app, and (3) implementation and evaluation of the intervention efficacy and app usability. Details of the protocol of this study, with data from modules 1 and 2, have been previously described [20]. This paper adheres to the CONSORT (Consolidated Standards of Reporting Trials) guidelines for reporting randomized trials (Checklist 1) [26]. The trial was registered at The Brazilian Clinical Trials Registry (RBR-4s8qyyq) [27], date of registration: January 28, 2022.

XSL•FO
RenderX

Participants

Participants were recruited from primary health care centers (PHCC) in a Brazilian city with a population of 51,000 people, located in the interior of São Paulo state. Inclusion criteria comprised individuals aged 20 to 59 years with adequate oral and reading skills. Participants were also required to possess sufficient digital literacy to operate the app independently or with assistance from an adult or family member. Exclusion criteria included individuals without access to a mobile phone equipped with the Android operating system. Individuals who met all the inclusion criteria and did not present any exclusion criteria were invited to participate by either the researcher or the Community Health Agent, who serves as the link between the PHCC and the families within the covered area.

Intervention

The intervention, delivered via a mobile app, was developed using the BCW [16], as detailed in a previous study protocol [20]. The intervention development was based on the results of previous studies that identified, in the Brazilian population, a high salt intake and its sources of consumption [8,9,19] and that identified that intention, self-efficacy, and habit were the key psychosocial determinants of salt intake behavior [21].

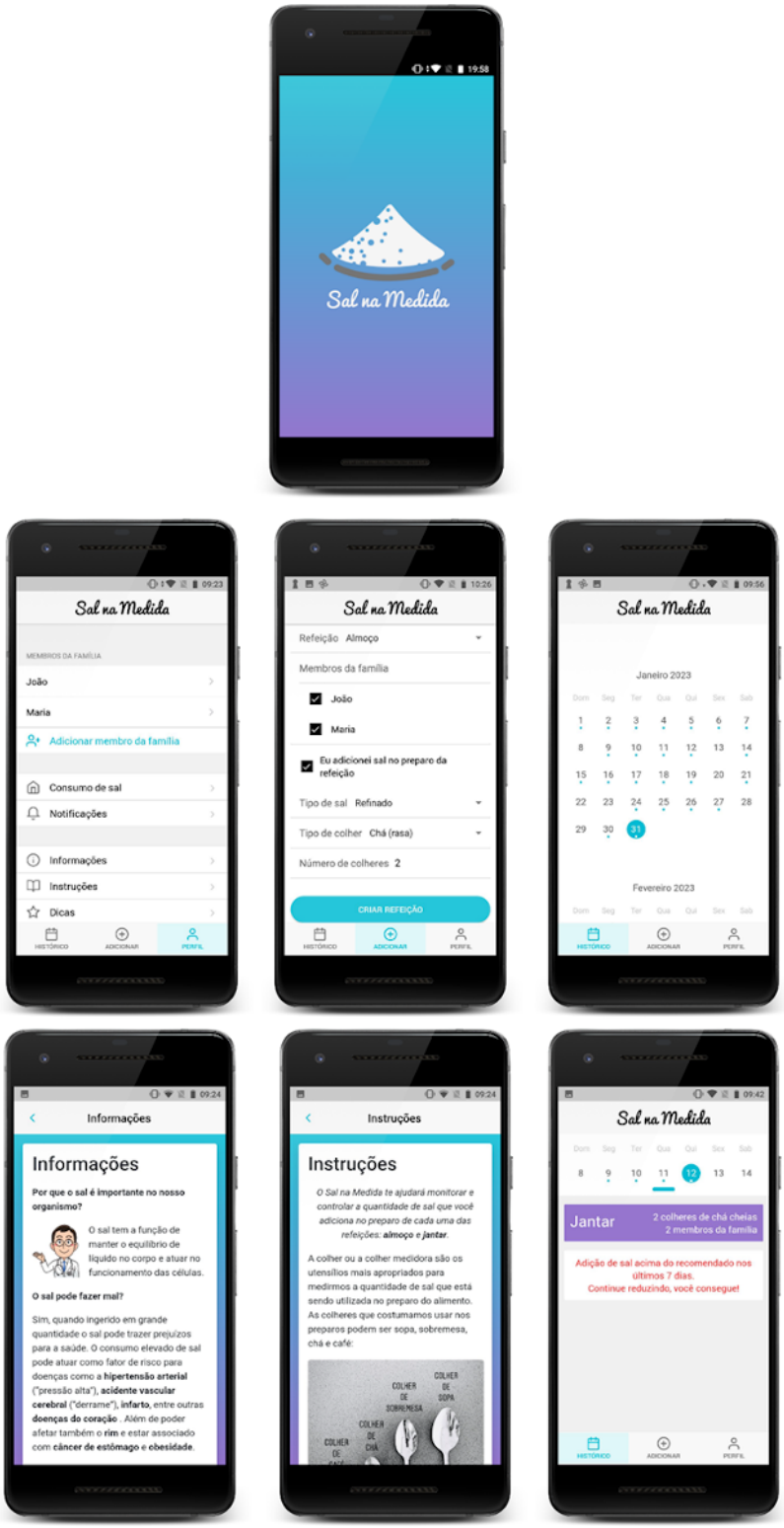
To address these determinants and by applying the BCW framework, we selected corresponding intervention functions, that is, broad categories of things one can do to change behavior, which included education, persuasion, incentivization, modeling, enablement, and training. Building on these intervention functions, specific “active ingredients” were identified to facilitate behavioral change, referred to as behavioral change techniques. These techniques included—problem solving, goal setting (outcome), action planning, commitment, feedback on behavior, self-monitoring of behavior, social support (unspecified), social support (practical), instruction on how to perform the behavior, information about health consequences, demonstration of the behavior, prompts or cues, behavioral practice or rehearsal, credible sources, and adding objects to the environment [20].

The prototype of the mobile app, named “Sal na Medida” app, was pretested among potential users, and the final version was produced.

In the app, the behavior change techniques described above were incorporated. Each day, participants record the number of people who had meals at home. For each meal, they enter the number of individuals who consumed it, as well as the number and type of salt spoons used. Additionally, the app provides information on the recommended daily salt intake, the health risks of excessive consumption, tips for reducing salt intake, and demonstrations—through text, images, and videos—to support better understanding. Throughout this process, the app offers feedback on whether the participant is meeting the recommended daily salt intake goal (Figure 1).

XSL•FO
RenderX

Figure 1. The “Sal na Medida” app screenshots.



After the app was installed on the Android (Google LLC) phones of the participants in the IG, they received instructions on how to use it, and then they were followed up for 2 months. Participants in the CG did not receive any intervention. However, as they were under the usual care of the primary care center, they may have been exposed to general advice on salt intake that is usually given during consultations with health professionals.

Outcomes

Primary Outcomes (Salt Intake Variables)

- Per capita salt: this is a self-reported questionnaire that estimates the discretionary sodium intake from salt used at the table and during home food preparation, developed and used in previous Brazilian context studies [9,19,21,28]. It measures the amount of salt consumed per person daily based on reports of the amount of salt used per month in

XSL•FO
RenderX

the household (1 kg packages), taking into account the number of people living in the house and the number of meals each person eats at home.
- Behavior of salt intake: the instrument used in this study was developed in a previous study [29] as a subjective tool for collecting data on salt consumption, specifically regarding the behavior of adding salt to food. It includes a question with 5 response options, assessing the behavior of adding a maximum of 3 g of salt per day during cooking.

Secondary Outcomes (Psychosocial Variables)

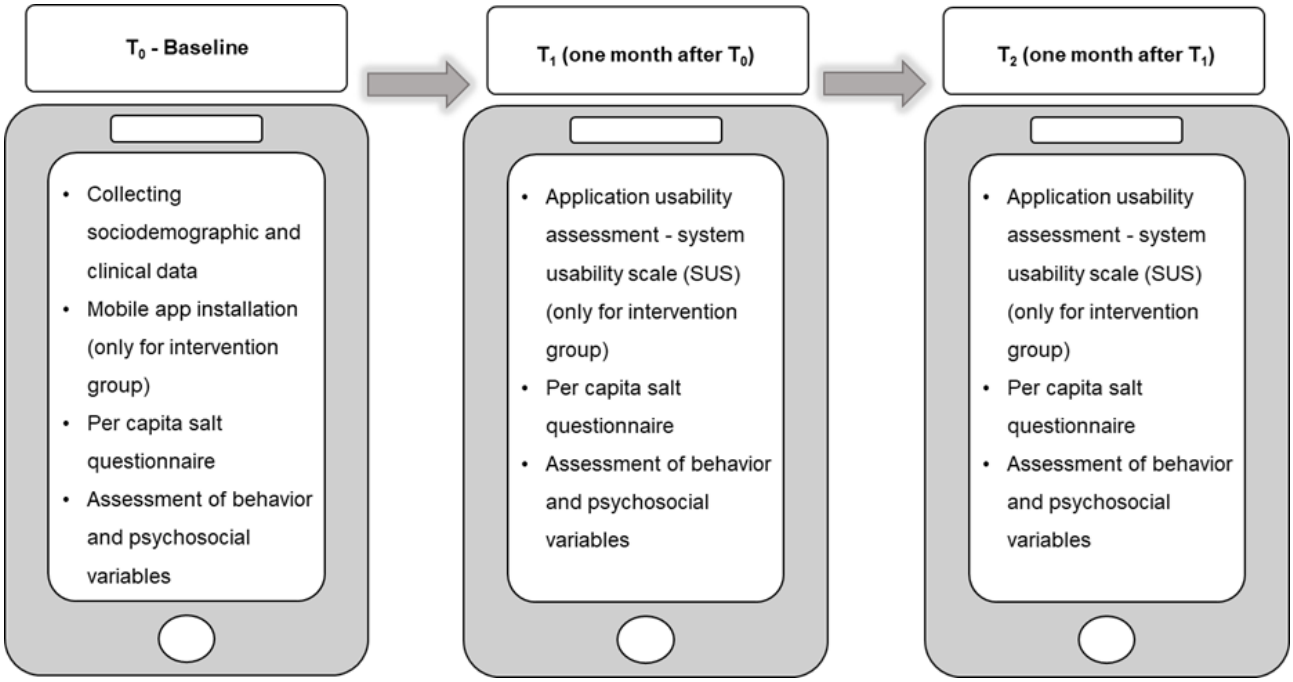
Psychosocial variables were assessed using a validated instrument developed in a previous study [29].

- Intention: participants' motivation for adding a maximum of 3 g of salt per day during cooking was estimated using a scale comprising 6 items, with each item rated on a 5-point scale (1-completely disagree to 5-completely agree). A high score indicated a high intention to perform the behavior.
- Self-efficacy: participants' confidence in their ability to add a maximum of 3 g of salt per day during cooking was assessed by means of the mean of 3 items rated on a 5-point scale (1-completely disagree to 5-completely agree). The higher the score, the greater the perception of self-efficacy to perform the behavior.
- Habit: evaluation of participants' habits regarding salt intake during cooking was conducted using a scale consisting of 10 items, each assessing the extent to which they used more than 1 level teaspoon of salt per day (equivalent to more than 3 g of salt), assessed on a 5-point scale (1-definitely not to 5-definitely yes]. Higher scores on the scale indicate a greater habit of using more than 3 g of salt per day during cooking.
- System Usability Scale (SUS): a reliable and widely used tool for assessing the usability of various products and services, including software, websites, and mobile applications [30]. It consists of a 10-item questionnaire with 5 response options for respondents (1-completely disagree to 5-completely agree). A higher score indicates better usability. The scale was applied at T_1 and T_2 only for IG participants.

Data Collection

This study was conducted over 3 waves of data collection during a 2-month follow-up period (Figure 2). At baseline (T_0), sociodemographic and clinical variables were obtained from participants in both IG and CG. Variables including per capita salt, behavior of salt intake, and the psychosocial factors (intention, self-efficacy, and habit) were evaluated at T_0, T_1, and T_2 for both groups. The SUS was applied to the IG at T_2.

Figure 2. Data collection plan: control and intervention group.



Sample Size

The sample size was estimated using the G* Power software (version 3.1.9.4; Heinrich-Heine-Universität Düsseldorf), based on the objective of comparing quantitative outcomes between IG and CG, in the 3 periods of data collection. Repeated measures ANOVA was used with a significance level of 5%, 80% power, and an effect size of 0.25, which is considered a moderate degree of effect size [28]. A minimum sample of 86 participants was estimated (each group consisted of 43 participants).

Randomization

Participants were randomly allocated to either the IG or the CG using a randomization procedure with variable block sizes. The randomization sequence was generated by a statistician, independent of the research team, using a web-based tool. The principal investigator was responsible for enrolling participants

XSL•FO
RenderX

and assigning them to their respective groups according to the predefined randomization sequence and order of enrollment.

Statistical Methods

Qualitative data were described using frequencies and percentages, and quantitative data as measures of central tendency (mean and median) and dispersion (SD, maximum, and minimum).

To evaluate the efficacy of the intervention on primary outcomes (salt intake variables), a linear regression model via generalized estimating equations modeling was used to make comparisons between groups and times.

Regression analysis with generalized linear models was used to assess the prediction of per capita salt at T_1 and T_2 by the psychosocial variables of habit, self-efficacy, and intention measured at T_0.

The app's usability was evaluated by the mean (SD) obtained by the SUS at T_2. The SUS scores were classified as: above 90 points and the app is considered to have the best possible usability, 81 to 90 points and it is considered excellent, 71 to 80 points as good, 60 to 70 points as 0kay, and below 60 points and the degree of usability is not acceptable [31].

To assess the mediators of the effect of the intervention, mediation analysis using linear regression models was used via generalized linear models and modified Poisson regression models with robust variance. The results of the linear regression and Poisson models are presented using the estimates of the regression coefficients and relative risk, respectively, with CI values and *P* values. In the mediation analyses, the results of behavior at T_0 were considered as a control variable.

A significance level of 5% will be adopted. Statistical analyses were performed using the SAS (version 9.4; IBM Corp) program.

Ethical Considerations

This study was approved by the Research Ethics Committee of the university to which the researchers are connected (10937419.0.0000.5404), accredited by the National Research Ethics Commission, in compliance with Resolution 466/2012/CNS/MS. The procedures followed were in accordance with the ethical standards outlined in the Declaration of Helsinki. Written informed consent was obtained from all participants. All necessary measures were undertaken to ensure participant privacy, including the use of anonymized data, secure data storage, and strict confidentiality protocols throughout the research process.

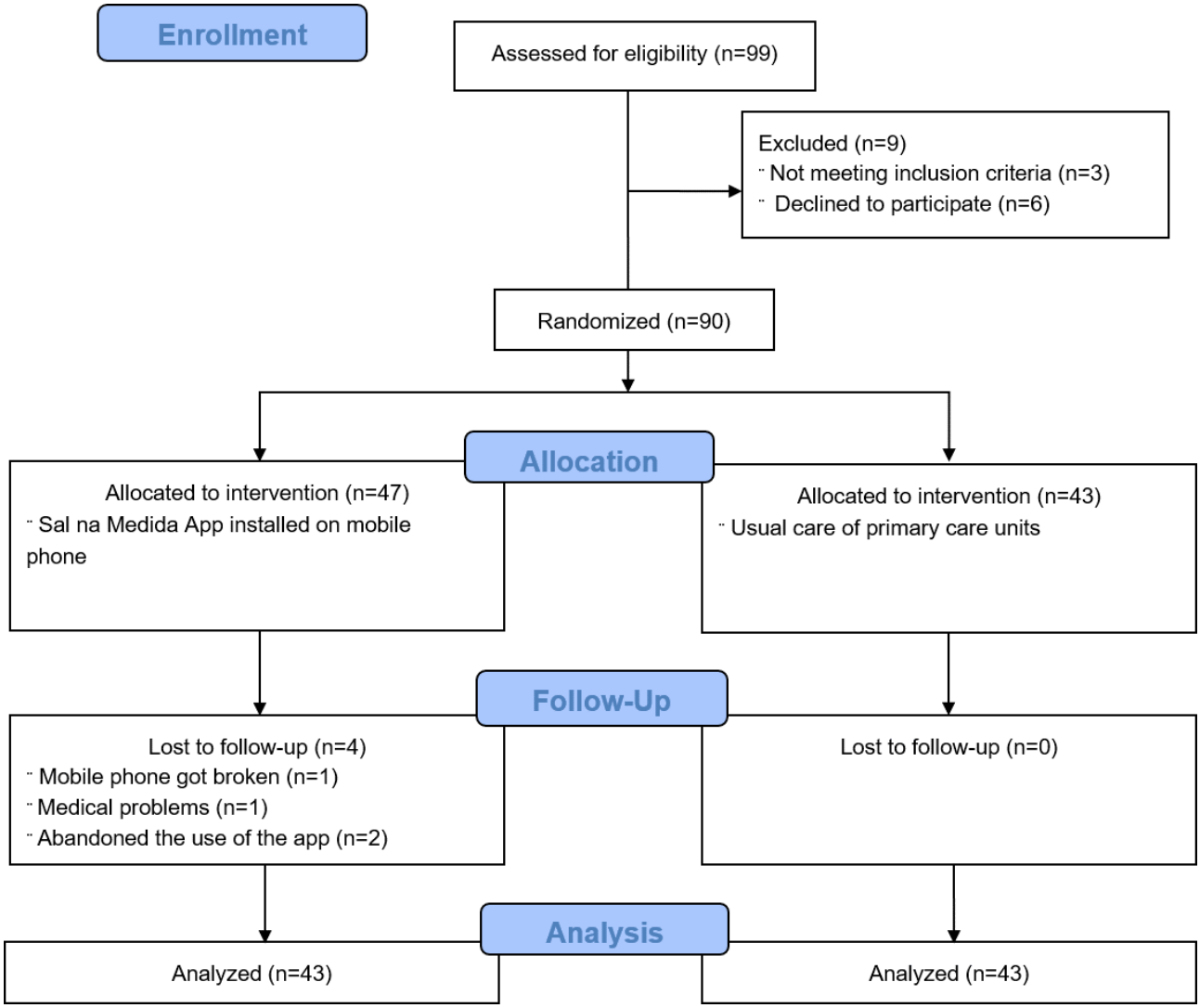
Data entered by IG participants into the "Sal na Medida" app was securely stored in the Firebase Realtime Database per Google's privacy policy [32].

Results

A total of 90 participants were recruited, and 86 (95%) of them completed this study after 2 months of its beginning (IG n=43; CG n=43). Four participants who did not complete the study, who were predominantly female 75% (3/4), with an average age of 40.2 (SD 13.6) years and a mean of 11 (SD 2.4) years of schooling (Figure 3).

XSL•FO
RenderX

Figure 3. Flow diagram of the study.



Sociodemographic and clinical characterization data, as well as comparative and association analyses applied to the 2 groups, are presented in Table 1.

XSL•FO
RenderX

Table . Sociodemographic and clinical characteristics of the intervention group and control group, and comparative and association analyses at baseline.

Variables	Intervention group (n=43)	Control group (n=43)	*P* value
Age (years), mean (SD)	34.2 (9.4)	43.4 (9.9)	<.001[a]
Schooling (years), mean (SD)	11.6 (2.9)	10.2 (1.6)	.07[b]
Individual monthly income (US $), mean (SD)	574.5 (625.8)	406.0 (185.5)	.11[b]
Family monthly income (US $), mean (SD)	1058.7 (826.6)	962.4 (757.0)	.94[b]
Skin color: White, n (%)	28 (65.1)	20 (46.5)	.08[c]
Sex: women, n (%)	36 (83.7)	37 (86)	.76[c]
Labor status: active, n (%)	40 (93)	42 (97.7)	—[d]
Hypertension	6 (13.9)	12 (27.9)	.11[c]
Diabetes	—	4 (9.3)	—
Dyslipidemia	1 (2.3)	1 (2.3)	—

[a]*P* value obtained with a 2-tailed unpaired Student *t* test.

[b]*P* value obtained with the Mann-Whitney test.

[c]*P* value obtained with the chi-square test.

[d]Not available.

Only for the age variable, there was a statistically significant difference (*P*<.001) between the means of participants in the IG (34.2, SD 9.4 y) and the CG (43.4, 9.9 y). For this reason, the analyses were adjusted according to the age variable. In both groups, there was a predominance of female (IG: n=36, 83.7%; CG: n=37, 86%), White (IG: n=28, 65.1%; CG: n=20, 46.5%), and active (IG: n=40, 93%; CG: n=42, 97.7%) participants. Considering the clinical data, there were no significant differences between groups.

The results regarding the intervention's effect on per capita salt intake indicated that, at baseline (T_0), there was no significant difference between the IG and CG (IG=4.6, SD 2.7 g; CG=5.3, SD 4.0 g). At T_1, 30 days after the baseline, both groups showed a reduction in salt intake (IG=3.9, SD 2.6 g; CG=4.7, SD 3.4 g). Finally, at the end of the intervention, the IG showed a further reduction in salt consumption, while the CG remained stable (IG=3.5, SD 2.3 g; CG=4.7, SD 3.6 g; Table 2).

Table . Description of variables of salt intake (per capita salt and behavior) and psychosocial determinants (intention, self-efficacy, and habit) at T_0, T_1, and T_2.

Variables	Intervention group (n=43)			Control group (n=43)		
	T_0	T_1	T_2	T_0	T_1	T_2
Per capita salt, mean (SD)	4.6 (2.7)	3.9 (2.6)	3.5 (2.3)	5.3 (4.0)	4.7 (3.4)	4.7 (3.6)
Intention, mean (SD)	4.0 (0.2)	4.1 (0.3)	4.2 (0.4)	3.9 (0.1)	3.9 (0.2)	3.7 (0.7)
Self-efficacy, mean (SD)	4.0 (0.2)	4.1 (0.3)	4.2 (0.5)	3.8 (0.4)	3.8 (0.5)	3.6 (0.7)
Habit, mean (SD)	3.4 (1.2)	2.3 (0.9)	1.8 (0.8)	3.2 (1.3)	2.1 (1.2)	2.6 (1.2)
Behavior (1 to 3)[a], n (%)	24 (55.8)	7 (16.2)	3 (7)	21 (48.9)	16 (37.2)	15 (34.9)
Behavior (4 to 5)[a], n (%)	19 (44.2)	36 (83.7)	40 (93)	22 (51.1)	27 (62.7)	28 (65.1)

[a]Behavior 1, 2, 3, 4, and 5 = use of a maximum of 3 g of salt/day in meal preparation—never, rarely, sometimes, on most days of the week, or every day, respectively.

In the regression analyses, the intragroup results confirmed that, at the end of the intervention, IG participants had a significant reduction in per capita salt intake (mean difference =−1.08 g; *P*=.01). In contrast, although the CG also showed a decrease in salt intake at T_2, this difference was not statistically significant (mean difference =−0.55 g; *P*=.22). There was no statistically significant intergroup difference at T_2 (Table 3).

XSL•FO
RenderX

Table . Effect of the intervention on the salt intake variables (per capita salt and behavior) and psychosocial variables (intention, self-efficacy, and habit) at T_0, T_1, and T_2.Analyses adjusted according to the age variable.

Variables	Mean difference	95% CI	*P* value[a]
Per capita salt			
Intervention-control (T_0)	−0.70	−2.38 to 0.98	.41
Intervention-control (T_1)	−0.81	−2.43 to 0.80	.32
Intervention-control (T_2)	−1.23	−2.89 to 0.43	.14
T_1 - T_0 (control group)	−0.58	−1.34 to 0.19	.14
T_2 - T_0 (control group)	−0.55	−1.43 to 0.33	.22
T_2 - T_1 (control group)	0.03	−0.48 to 0.53	.92
T_1 - T_0 (intervention group)	−1.69	−1.45 to 0.07	.07
T_2 - T_0 (intervention group)	−1.08	−1.91 to −0.25	.01
T_2 - T_1 (intervention group)	−0.39	−0.87 to 0.09	.11
Intention			
Intervention-control (T_0)	0.17	0.07 to 0.26	<.001
Intervention-control (T_1)	0.23	0.08 to 0.38	.002
Intervention-control (T_2)	0.50	0.24 to 0.76	<.001
T_1 - T_0 (control group)	0.02	−0.07 to 0.10	.71
T_2 - T_0 (control group)	−0.17	−0.38 to 0.03	.09
T_2 - T_1 (control group)	−0.19	−0.36 to −0.02	.02
T_1 - T_0 (intervention group)	0.08	−0.02 to 0.18	.09
T_2 - T_0 (intervention group)	0.16	0.02 to 0.29	.02
T_2 - T_1 (intervention group)	0.08	−0.02 to 0.17	.12
Self-efficacy			
Intervention-control (T_0)	0.22	0.01 to 0.43	.04
Intervention-control (T_1)	0.41	0.16 to 0.66	.001
Intervention-control (T_2)	0.66	0.35 to 0.97	<.001
T_1 - T_0 (control group)	−0.05	−0.22 to 0.11	.51
T_2 - T_0 (control group)	−0.23	−0.46 to −0.01	.049
T_2 - T_1 (control group)	−0.18	−0.38 to 0.03	.09
T_1 - T_0 (intervention group)	0.14	0.02 to 0.26	.02
T_2 - T_0 (intervention group)	0.21	0.03 to 0.39	.02
T_2 - T_1 (intervention group)	0.07	−0.08 to 0.22	.36
Habit			
Intervention-control (T_0)	0.03	−0.56 to 0.61	.93
Intervention-control (T_1)	−0.72	−1.22 to −0.21	.005
Intervention-control (T_2)	−1.03	−1.52 to −0.54	<.001
T_1 - T_0 (control group)	−0.41	−0.68 to −0.14	.003
T_2 - T_0 (control group)	−0.55	−0.83 to −0.27	<.001
T_2 - T_1 (control group)	−0.14	−0.32 to 0.04	.12
T_1 - T_0 (intervention group)	−1.15	−1.48 to −0.82	<.001

XSL•FO
RenderX

Variables	Mean difference	95% CI	*P* value[a]
T_2 - T_0 (intervention group)	−1.60	−1.95 to −1.25	<.001
T_2 - T_1 (intervention group)	−0.45	−0.62 to −0.29	<.001

[a]Linear regression via generalized estimating equations.

To assess the effect of the intervention on the behavior variable of salt intake, participants were considered to have achieved the desired behavior if they reported scores of "4" or "5," indicating that they used a maximum of 3 g of salt per day in meal preparation on most days of the week or every day, respectively.

We found a significant change in the scores among IG participants. At T_0, the majority (n=24, 55.8%) reported never, rarely, or occasionally using a maximum of 3 g of salt per day during meal preparation. By the end of the intervention (T_2), this percentage increased significantly, with 93% reporting consistent adherence to this behavior on most or all days of the week (Table 2).

Regression intergroup analyses showed that, at T_2, participants in the IG had a 63% higher probability of adhering to this behavior compared to those in the CG (*P*<.001; Table 3). Additionally, the intragroup analysis revealed that, after the intervention, IG participants were 111% more likely (*P*<.001) to consistently use a maximum of 3 g of salt per day when preparing meals on most days or every day (Table 4).

Table . Effect of the intervention on the salt intake variables (behavior) at T_0, T_1, and T_2. Analyses adjusted according to the age variable.

Behavior	Relative risk[a]	95% CI	*P* value[b]
Intervention-control (T_0)	0.99	0.63 to 1.56	.95
Intervention-control (T_1)	1.52	1.15 to 2.03	.004
Intervention-control (T_2)	1.63	1.26 to 2.11	<.001
T_1/T_0 (control group)	1.23	1.03 to 1.47	.02
T_2/T_0 (control group)	1.27	1.05 to 1.54	.01
T_2/T_1 (control group)	1.04	0.92 to 1.17	.56
T_1/T_0 (intervention group)	1.89	1.39 to 2.58	<.001
T_2/T_0 (intervention group)	2.11	1.52 to 2.92	<.001
T_2/T_1 (intervention group)	1.11	1.01 to 1.23	.04

[a]It was estimated the risk of presenting the result "4 and 5."

[b]Poisson model modified via generalized estimating equations.

Concerning secondary outcomes, that is, the psychosocial determinants of behavior, the IG participants showed, after the intervention, a slight increase in the mean of intention (mean T_0=4.0; mean T_2=4.2) and in the self-efficacy (mean T_0=4.0; mean T_2=4.2). As for the habit of using more than 3 g of salt per day in the preparation of meals, there was a decrease in the average score of this variable in both groups, but in IG, the decrease was more pronounced (mean T_0: 3.4; mean T_2: 1.8) when compared to the CG (mean T_0=3.2; mean T_2=2.6; Table 2).

The intragroup regression analysis indicated that, at the end of the intervention, IG participants reported an increase in the intention to use a maximum of 3 g of salt per day in preparing meals (mean difference=0.16; *P*=.02). When comparing the groups, at T_2, the IG showed higher intention when compared to CG (mean difference=0.51; *P* value<.001; Table 3).

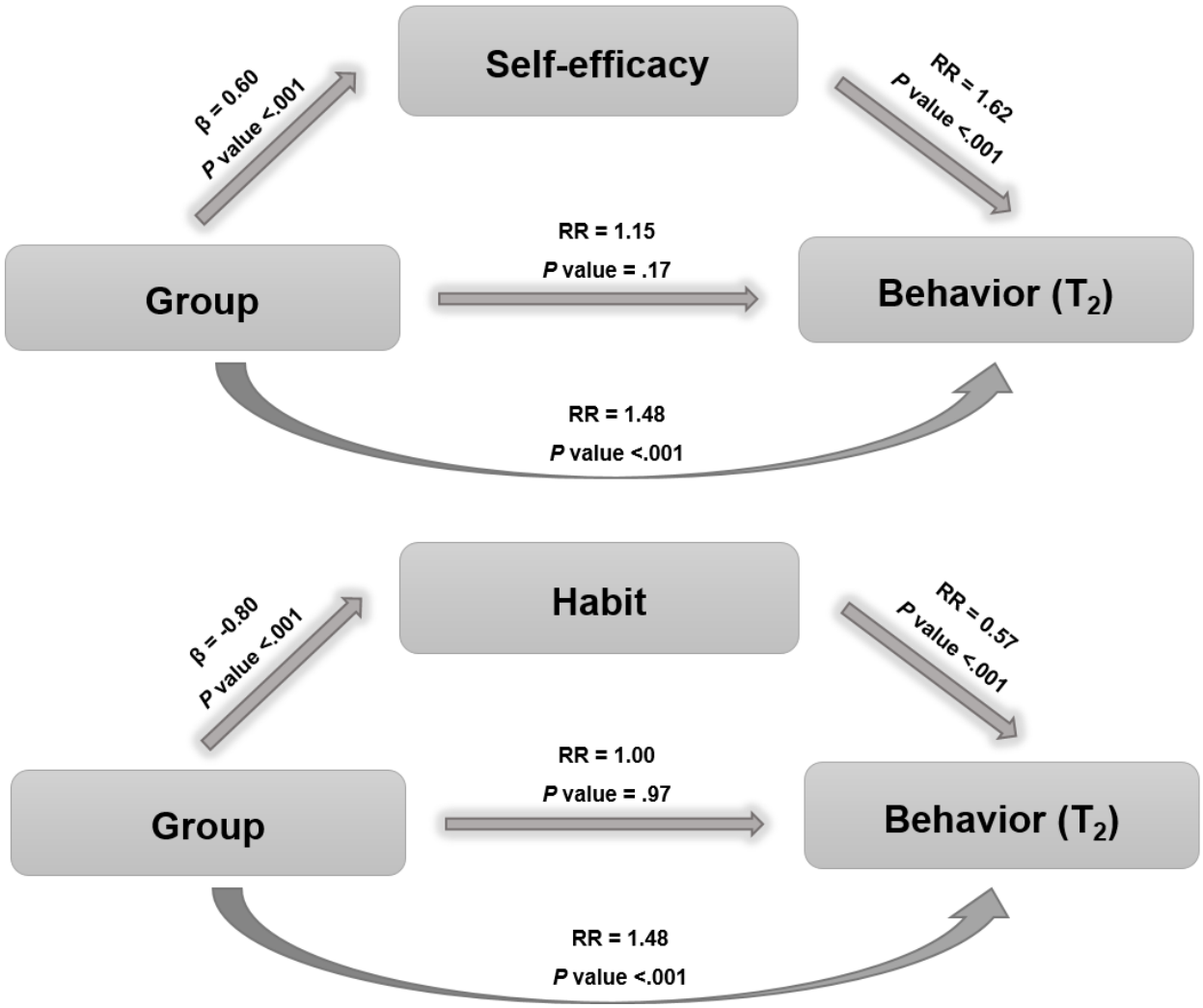
For self-efficacy, after the intervention, there was an increase in the scores of this variable in the IG (mean difference=0.21; *P*=.02) and a decrease in the CG (−0.23; *P*=.0494). In the intergroup analysis, after the intervention, the IG presented higher self-efficacy when compared to the CG (mean difference=0.66, *P*<.001; Table 3).

Regarding the habit of using more than 3 g of salt per day in meal preparation, a significant reduction was observed in both the CG (mean difference=−0.55; *P* value <.001) and IG (mean difference=−1.60; *P*<.001). However, at T_2, the reduction was significantly greater in the IG compared to the CG (mean difference=−1.03; *P*<.001; Table 3).

Mediation analyses found that self-efficacy and habit mediated the effect of the intervention. The analyses show that the intervention led to an increase in participants' perception of self-efficacy, which in turn contributed to improving the behavior of salt intake, that is, the increase in self-efficacy led to a higher frequency of the behavior of using no more than 3 g of salt per day. Additionally, the intervention resulted in a decrease in the habit of adding large amounts of salt when cooking, which also served as a pathway through which the intervention influenced the behavior of salt addition (Figure 4).

XSL•FO
RenderX

Figure 4. Mediation analysis. Analyzes adjusted according to the behavior T_0. It was estimated the risk of presenting the result "4 and 5"(use of a maximum of 3 g of salt/day in the preparation of meals on most days of the week or every day, respectively). β: coefficient; RR: relative risk.



The usability of the application evaluated in the IG indicated a mean score of 77.8 (SD 12.7). The average score on the SUS scale of the individuals after the intervention was 77.8, which indicates good usability of the application.

Discussion

Principal Findings

The “Sal na Medida” app was developed to contribute as a public policy strategy to change the behavior of adding large amounts of salt during home food preparation by adults in the general Brazilian population. Our results showed a substantial positive impact on participants’ salt intake behavior throughout the intervention and indicate that the “Sal na Medida” app is an important method for changing the behavior of salt addition in home-cooked meals and proved to be a good alternative in promoting healthy salt intake.

The improvement in the behavior among participants in the IG from baseline to the end of the intervention was quite significant. While the majority initially reported limited salt addition only occasionally or rarely at baseline, postintervention 93% reported consistently adhering to the maximum 3 g of salt per day during cooking. This highlights the efficacy of the intervention in promoting healthier salt intake behavior. Furthermore, the chances of an individual performing the healthy salt intake behavior using the mobile app were 63% higher than for people who did not use this tool.

When evaluating other review studies, we found many have been using mobile apps to change eating habits [22,23], but still, few have been in the area of salt intake specifically. As we found several sources of consumption of this nutrient, strategies were applied to reduce the consumption of high-sodium foods by offering an alternative to low-salt foods in these places where processed or ultraprocessed foods are sold or consumed [33,34]. In another study, despite adding information about using less salt when cooking or at the table, this guidance was provided by a health professional in a face-to-face visit, and the app was used only, as well as in previous studies, to encourage individuals to make choices with low sodium content in industrialized foods [35].

Another finding of our study is with respect to the mediation analyses, which revealed important insights into the mechanisms through which the intervention was successful in changing the salt intake behavior. Specifically, we found that the positive effect of the intervention on salt intake behavior was mediated by 2 psychosocial variables: self-efficacy and habit. This means

XSL•FO
RenderX

that as individuals gained confidence in their ability to control their salt intake, they were more likely to adopt this behavior; additionally, breaking the automaticity in the use of salt, that is, making a conscious behavior, people are able to achieve the desired behavior. So, understanding these mediating factors clarifies the mechanism through which the intervention influences the desired results [15].

Despite providing advice and knowledge about healthy salt recommendations, interventions are generally aimed at a specific public, with the disease already triggered and with the aim of not aggravating this condition, as in cases of hypertension [34,35] or cardiovascular disease [33] or even controlling the associated symptoms, such as high blood pressure, and not the general population to prevent these diseases.

In terms of health prevention, such a tool is essential to promote public policies at the first level of health care, as actions aimed at reducing the addition of salt aim to prevent the onset and worsening of chronic noncommunicable diseases, impacting health as well as the economy [36]. Intervening in this behavior is part of several government initiatives, as this behavior of salt consumption above the recommended level is not unique to Brazil, but also to other countries in the Americas [37]. The Pan American Health Organization, together with the International Development Research Center, developed a report for Latin American countries with social marketing to help reduce salt, but it has been insufficient to implement or adopt salt reduction strategies in local communities [37].

In Brazil, there are ongoing policies, such as voluntary targets for reducing the sodium content in processed and ultraprocessed foods, the front labeling of warnings that was approved by Brazilian Health Regulatory Agency (Anvisa) and includes the alert for excess sodium [38], in addition to the recommendations of the Food Guide Brazilian, when it comes to the use of small amounts of culinary ingredients [39].

Finally, in evaluating the app usability, it is possible to identify how easily and efficiently users can interact with an app to achieve their goal. The usability score of the “Sal na Medida” app showed good usability, which is a good indication. Higher usability leads to greater user acceptance and improved health outcomes, making it a crucial factor in evaluating digital health applications [40]. Although the study involved a younger adult sample—who might generally be more tech-savvy and thus could have influenced the app’s usability score—a systematic review on educational technology usability using the SUS found no significant correlation between SUS scores and participants’ age [41]. The “Sal na Medida” app was tested with the population that uses primary health care services in Brazil, which has the potential to be another alternative tool in the future.

Strengths and Limitations

A recent review on mobile health technologies to reduce salt makes 3 recommendations for future interventions in this area [25], in which the “Sal na Medida” app is included.

First, to create innovative and interactive mobile health technologies. Given that this study addresses a previously unexplored source of salt consumption in mobile apps—one that contributes to exceeding recommended daily intake limits—along with its demonstrated good usability and the ability to be used independently without professional assistance, the “Sal na Medida” app represents the first tool to integrate these concepts.

Second, to design interventions in primary care and associate them with salt reduction targets. Our study is aligned with specific public health recommendations for reducing salt intake, particularly by promoting behavioral changes in excessive salt use during meal preparation, and, additionally, it was conducted with adults recruited from PHCCs in a Brazilian city. Thus, the “Sal na Medida” app is promising with regard to its impact on reducing salt intake at the population level, especially in a specific geographical context. Although it was developed specifically for reducing salt consumption in food preparation, the app contains strategies that involve replacing foods with high sodium content with natural foods. This behavioral change can trigger changes in other behaviors related to salt intake, such as the consumption of processed foods.

The third and final item is to develop large-scale clinical trials with a robust design, incorporating an outcome variable (24 h urine sodium). We conducted a randomized controlled study, targeted at a specific behavior of salt intake, and the intervention was guided by the BCW [16], a robust framework. As the focus of the study was on behavior change—specifically reducing the amount of salt added during meal preparation—and on encouraging participants to adopt this behavior as a lasting habit in their daily lives, the 24-hour urine sodium was not assessed as a quantitative variable in this study.

The study also had some limitations. The system required the participants’ smartphones to have only Android because the app is not available in iOS (Apple Inc). Another limitation was the impossibility of blinding researchers and participants for data collection after the intervention, which could cause potential bias in the responses of individuals.

Another study limitation is the use of a self-reported questionnaire to estimate per capita salt intake. The literature points to the lithium-tagged salt method as the gold standard for measuring discretionary salt intake. However, studies using this method are often limited in size due to the significant burden it places on participants [42]. The self-reported per capita salt questionnaire used in this study was used in previous studies in the Brazilian context [9,19-21,43,44] because it allows us to capture the main source of salt consumption for this population. The 24-hour urinary sodium excretion is the gold standard for objectively quantifying salt intake; however, this method does not allow identification of the source of salt intake and has several limitations, such as intraindividual variability in urinary excretion due to factors like diet, hydration status, physical activity, stress, and renal function; inconvenience and burden for participants to collect the 24-hour urine sample, which may lead to incomplete collection; and high costs of analyses [45].

Finally, future studies investigating the long-term effects of the “Sal na Medida” app on maintaining behavior change are necessary, as the follow-up period in this study was limited to 2 months. Additionally, future research should include older

XSL•FO
RenderX

adults to evaluate the intervention's efficacy and the app's usability within this population.

Conclusions

The "Sal na Medida" app intervention demonstrated a positive impact on salt intake behavior, enhancing participants' intention and self-efficacy perception toward reducing salt use during home-cooked meal preparation, while also decreasing the habit of adding excessive salt. Achieving a change in behavior and integrating healthy habits of salt intake into the routine of families during meals is the goal of the "Sal na Medida" app. The next phase involves scaling up the app for broader implementation.

Acknowledgments

This work was supported by the São Paulo Research Foundation (FAPESP; 2020/13588-0), and by the Fundo de Apoio ao Ensino, à Pesquisa e à Extensão (FAEPEX) of the Universidade Estadual de Campinas (UNICAMP; 2461/23).

Conflicts of Interest

None declared.

Checklist 1
CONSORT-eHEALTH checklist (V1.6.1).
[PDF File, 1162 KB - humanfactors_v12i1e54174_app1.pdf]

References

1. GBD 2017 Diet Collaborators. Health effects of dietary risks in 195 countries, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2019;393(10184):1958-1972. [doi: 10.1016/S0140-6736(19)30041-8]
2. Nilson EAF, da Silva EN, Jaime PC. Developing and applying a costing tool for hypertension and related cardiovascular disease: attributable costs to salt/sodium consumption. J Clin Hypertens (Greenwich) 2020 Apr;22(4):642-648. [doi: 10.1111/jch.13836] [Medline: 32108425]
3. SHAKE the salt habit - the SHAKE technical package for salt reduction. : World Health Organization; 2016 URL: http://apps.who.int/iris/bitstream/10665/250135/1/9789241511346-eng.pdf [accessed 2025-07-29]
4. Global action plan for the prevention and control of noncommunicable diseases 2013-2020. : World Health Organization; 2013 URL: https://www.who.int/publications/i/item/9789241506236 [accessed 2025-07-29]
5. Ghimire K, Mishra SR, Satheesh G, et al. Salt intake and salt‐reduction strategies in South Asia: from evidence to action. J Clin Hypertens 2021 Oct;23(10):1815-1829. [doi: 10.1111/jch.14365]
6. Al Jawaldeh A, Rafii B, Nasreddine L. Salt intake reduction strategies in the Eastern Mediterranean Region. East Mediterr Health J 2019 Feb 18;24(12):1172-1180. [doi: 10.26719/emhj.18.006] [Medline: 30799557]
7. Campbell NRC, Whelton PK, Orias M, et al. 2022 World Hypertension League, Resolve To Save Lives and International Society of Hypertension dietary sodium (salt) global call to action. J Hum Hypertens 2023 Jun;37(6):428-437. [doi: 10.1038/s41371-022-00690-0] [Medline: 35581323]
8. Sarno F, Claro RM, Levy RB, Bandoni DH, Monteiro CA. Estimated sodium intake for the Brazilian population, 2008-2009. Rev Saude Publica 2013 Jun;47(3):571-578. [doi: 10.1590/s0034-8910.2013047004418] [Medline: 24346570]
9. Perin MS, Cornélio ME, Oliveira HC, São-João TM, Rhéaume C, Gallani MCBJ. Dietary sources of salt intake in adults and older people: a population-based study in a Brazilian town. Public Health Nutr 2019 Jun;22(8):1388-1397. [doi: 10.1017/S1368980018003233] [Medline: 30472972]
10. Schorling E, Niebuhr D, Kroke A. Cost-effectiveness of salt reduction to prevent hypertension and CVD: a systematic review. Public Health Nutr 2017 Aug;20(11):1993-2003. [doi: 10.1017/S1368980017000593] [Medline: 28487006]
11. He FJ, Campbell NRC, Woodward M, MacGregor GA. Salt reduction to prevent hypertension: the reasons of the controversy. Eur Heart J 2021 Jul 1;42(25):2501-2505. [doi: 10.1093/eurheartj/ehab274] [Medline: 34117487]
12. Khalesi S, Williams E, Irwin C, et al. Reducing salt intake: a systematic review and meta-analysis of behavior change interventions in adults. Nutr Rev 2022 Mar 10;80(4):723-740. [doi: 10.1093/nutrit/nuab110] [Medline: 34921314]
13. Webb M, Fahimi S, Singh GM, et al. Cost effectiveness of a government supported policy strategy to decrease sodium intake: global analysis across 183 nations. BMJ 2017 Jan 10;356:i6699. [doi: 10.1136/bmj.i6699] [Medline: 28073749]
14. Trieu K, McMahon E, Santos JA, et al. Review of behaviour change interventions to reduce population salt intake. Int J Behav Nutr Phys Act 2017 Feb 8;14(1):17. [doi: 10.1186/s12966-017-0467-1] [Medline: 28178990]
15. Hagger MS, Moyers S, McAnally K, McKinley LE. Known knowns and known unknowns on behavior change interventions and mechanisms of action. Health Psychol Rev 2020 Jan 2;14(1):199-212. [doi: 10.1080/17437199.2020.1719184]
16. Michie S, Atkins L, West R. The Behaviour Change Wheel: A Guide to Designing Interventions: Silverback; 2014.
17. Rohde A, Duensing A, Dawczynski C, Godemann J, Lorkowski S, Brombach C. An app to improve eating habits of adolescents and young adults (Challenge to Go): systematic development of a theory-based and target group-adapted mobile app intervention. JMIR mHealth uHealth 2019 Aug 12;7(8):e11575. [doi: 10.2196/11575] [Medline: 30903746]

XSL•FO
RenderX

18. MacPherson MM, Cranston KD, Locke SR, Bourne JE, Jung ME. Using the Behavior Change Wheel to develop text messages to promote diet and physical activity adherence following a diabetes prevention program. Transl Behav Med 2021 Aug 13;11(8):1585-1595. [doi: 10.1093/tbm/ibab058]
19. Perin MS, Cornelio ME, Rodrigues RCM, Gallani MCBJ. Characterization of salt consumption among hypertensives according to socio-demographic and clinical factors. Rev Latino-Am Enfermagem 2013 Sep;21(5):1013-1021. [doi: 10.1590/S0104-11692013000500002]
20. Perin MS, São-João T, Gallani M, Agbadje TT, Rodrigues RCM, Cornélio ME. A mobile phone app intervention to promote healthy salt intake among adults: protocol for a randomized controlled study. JMIR Res Protoc 2022 Jun 29;11(6):e37853. [doi: 10.2196/37853] [Medline: 35767347]
21. Cornélio ME, Gallani MCBJ, Godin G, Rodrigues RCM, Nadruz W Jr, Mendez RDR. Behavioural determinants of salt consumption among hypertensive individuals. J Hum Nutr Diet 2012 Aug;25(4):334-344. [doi: 10.1111/j.1365-277X.2012.01238.x] [Medline: 22380701]
22. Schoeppe S, Alley S, Van Lippevelde W, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. Int J Behav Nutr Phys Act 2016 Dec;13(1):127. [doi: 10.1186/s12966-016-0454-y]
23. Villinger K, Wahl DR, Boeing H, Schupp HT, Renner B. The effectiveness of app-based mobile interventions on nutrition behaviours and nutrition-related health outcomes: a systematic review and meta-analysis. Obes Rev 2019 Oct;20(10):1465-1484. [doi: 10.1111/obr.12903] [Medline: 31353783]
24. Franco RZ, Fallaize R, Weech M, Hwang F, Lovegrove JA. Effectiveness of web-based personalized nutrition advice for adults using the eNutri web app: evidence from the EatWellUK randomized controlled trial. J Med Internet Res 2022 Apr 25;24(4):e29088. [doi: 10.2196/29088] [Medline: 35468093]
25. Ali SH, Luo R, Li Y, Liu X, Tang C, Zhang P. Application of mobile health technologies aimed at salt reduction: systematic review. JMIR mHealth uHealth 2019;7(4):e13250. [doi: 10.2196/13250]
26. CONSORT. Checklist of information to include when reporting a randomised trial. 2010. URL: https://www.equator-network.org/wp-content/uploads/2017/05/CONSORT-2010-Checklist.doc [accessed 2025-8-6]
27. RBR-4s8qyyq use of mobile phone application (app) to promote healthy salt consumption: intervention in adults. ReBEC. URL: https://ensaiosclinicos.gov.br/rg/RBR-4s8qyyq [accessed 2025-07-29]
28. Cohen J. A power primer. Psychol Bull 1992 Jul;112(1):155-159. [doi: 10.1037//0033-2909.112.1.155] [Medline: 19565683]
29. Cornélio ME, Gallani M, Godin G, Rodrigues RCM, Mendes RDR, Junior WN. Development and reliability of an instrument to measure psychosocial determinants of salt consumption among hypertensive patients. Rev Lat Am Enfermagem 2009;17(5):701-707. [doi: 10.1590/s0104-11692009000500017] [Medline: 19967221]
30. Brooke J. SUS-a quick and dirty usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation in Industry: Taylor and Francis; 1996.
31. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009;4(3):114-123 [FREE Full text]
32. Privacy and security in firebase. Firebase. 2022. URL: https://firebase.google.com/support/privacy?hl=en [accessed 2025-07-29]
33. Eyles H, McLean R, Neal B, et al. A salt-reduction smartphone app supports lower-salt food purchases for people with cardiovascular disease: findings from the SaltSwitch randomised controlled trial. Eur J Prev Cardiol 2017 Sep;24(13):1435-1444. [doi: 10.1177/2047487317715713] [Medline: 28631933]
34. Dorsch MP, Cornellier ML, Poggi AD, et al. Effects of a novel contextual just-in-time mobile app intervention (LowSalt4Life) on sodium intake in adults with hypertension: pilot randomized controlled trial. JMIR mHealth uHealth 2020;8(8):e16696. [doi: 10.2196/16696]
35. Riches SP, Piernas C, Aveyard P, et al. A mobile health salt reduction intervention for people with hypertension: results of a feasibility randomized controlled trial. JMIR mHealth uHealth 2021;9(10):e26233. [doi: 10.2196/26233]
36. Saving lives, spending less: a strategic response to noncommunicable diseases. : World Health Organization; 2018 URL: https://www.who.int/publications/i/item/WHO-NMH-NVI-18.8 [accessed 2025-07-29]
37. Khalig M, Sommariva S, Bardfield L, et al. Salt reduction in latin america: a regional social marketing & communication plan. In: Project - IDRC 108167 Scaling Up and Evaluating Salt Reduction Policies and Programs in Latin American Countries 2016-2020 Tres Ríos, Costa Rica: Costa Rican Institute of Research and Teaching in Health and Nutrition (INCIENSA) 2019.
38. Policy Brief: Redução Do Sódio Em Alimentos Processados e Ultraprocessados No Brasil [Recurso Eletrônico] / Ministério Da Saúde, Secretaria de Atenção Primária à Saúde, Departamento de Promoção Da Saúde: Brasília: Ministério da Saúde; 2022.
39. Secretaria de Atenção à Saúde Departamento de Atenção Básica Guia Alimentar Para a População Brasileira / Ministério Da Saúde, Secretaria de Atenção à Saúde, Departamento de Atenção Básica – 2: Brasília: Ministério da Saúde; 2014.
40. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: a scoping review. Int J Med Inform 2019 Jun;126:95-104. [doi: 10.1016/j.ijmedinf.2019.03.018] [Medline: 31029270]

XSL•FO
RenderX

41. Vlachogianni P, Tselios N. Perceived usability evaluation of educational technology using the system usability scale (SUS): a systematic review. JRTE 2021;54(3):392-409. [doi: 10.1080/15391523.2020.1867938]
42. McLean RM, Wang NX, Cameron C, Skeaff S. Measuring sodium from discretionary salt: comparison of methods. Nutrients 2023 Dec 12;15(24):5076. [doi: 10.3390/nu15245076] [Medline: 38140335]
43. Cornélio ME, Godin G, Rodrigues RC, de Freitas Agondi R, Alexandre NM, Gallani MCB. Effect of a behavioral intervention of the SALdável program to reduce salt intake among hypertensive women: a randomized controlled pilot study. Eur J Cardiovasc Nurs 2016 Apr;15(3):e85-e94. [doi: 10.1177/1474515115589275] [Medline: 26025215]
44. Ferreira-Sae MCS, Gallani MCB, Nadruz W, et al. Reliability and validity of a semi-quantitative FFQ for sodium intake in low-income and low-literacy Brazilian hypertensive subjects. Public Health Nutr 2009 Nov;12(11):2168-2173. [doi: 10.1017/S1368980009005825] [Medline: 19476677]
45. Ginos BNR, Engberink R. Estimation of sodium and potassium intake: current limitations and future perspectives. Nutrients 2020 Oct 26;12(11):3275. [doi: 10.3390/nu12113275] [Medline: 33114577]

Abbreviations

BCW: Behavior Change Wheel
CG: control group
CONSORT: Consolidated Standards of Reporting Trials
DALY: disability-adjusted life year
IDRC: International Development Research Center
IG: intervention group
PHCC: primary health care center
SUS: system usability scale

Edited by A Kushniruk; submitted 17.03.24; peer-reviewed by A Eisingerich, S Payne; revised version received 19.05.25; accepted 09.06.25; published 20.08.25.

Please cite as:
Perin MS, Gallani MCBJ, Nilson EAF, Agbadje TT, Cornélio ME
Theory-Based Mobile App Intervention to Promote Healthy Salt Intake Among Adults: Randomized Controlled Trial
JMIR Hum Factors 2025;12:e54174
URL: https://humanfactors.jmir.org/2025/1/e54174
doi:10.2196/54174

© Milena Sia Perin, Maria Cecília Bueno Jayme Gallani, Eduardo Augusto Fernandes Nilson, Titilayo Tatiana Agbadje, Marilia Estevam Cornélio. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 20.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Setting Goals and Accepting Challenges for Behavior Change—Analysis of Participants' Interactions With a Digital Multiple Health Behavior Intervention: Mixed Methods Study

Katarina Åsberg[1], PhD; Marie Löf[2], PhD; Marcus Bendtsen[1], PhD

[1]Division of Society and Health, Department of Health, Medicine and Caring Sciences, Linköping University, Linköping, Sweden

[2]Department of Medicine Huddinge, Karolinska Institutet, Stockholm, Sweden

Corresponding Author:
Katarina Åsberg, PhD
Division of Society and Health, Department of Health, Medicine and Caring Sciences, Linköping University, Linköping, Sweden

Abstract

Background: Digital interventions are effective in promoting healthy behaviors and are recognized as one of many strategies for achieving healthier populations. These interventions often include goal-setting, but the practical application and fidelity of goal setting, especially when targeting multiple health behaviors, remain underexplored. In a factorial randomized trial, we included goal-setting as one of six behavior change components in the digital intervention "Buddy," targeting university and college students' alcohol, diet, physical activity, and smoking behaviors. However, we found no strong and consistent evidence of an effect of goal-setting alone on any of the outcomes, highlighting the need to investigate how participants used this component.

Objective: This case study of Buddy aimed to gain insight into participants' interactions with the goal-setting component. Specific objectives were to identify the characteristics of participants who used this component and to analyze participants' self-authored content.

Methods: This study combined fidelity and effectiveness findings and involved 1704 participants from 18 universities and colleges in Sweden. Self-authored goals and challenges were analyzed using summative content analysis. Logistic and negative binomial regression analyses were conducted to estimate the odds of setting a goal, selecting or self-authoring a challenge, to estimate the odds of setting a goal with respect to a specific behavior, and to estimate the frequency of selecting or self-authoring different behavioral challenges.

Results: Of the 850 participants given access to the goal setting and challenges component, 427 (50%) set at least one goal and 403 (47%) selected or self-authored at least one challenge. A total of 607 goals were set, with most participants setting one goal (336/427, 79%). Goals primarily targeted physical activity (n=302), dietary behavior (n=140), and multiple health behaviors (n=53), typically combining physical activity with diet, alcohol, smoking, or sleep. Other goals included study performance, mental health, sleep, and mobile phone use (n=73). Fewer goals concerned alcohol (n=19) or tobacco (n=17). Participants selected 1506 challenges from 41 premade challenges, with dietary behavior challenges being most popular (667/1506, 44%). An additional 170 challenges were self-authored. Participants' baseline characteristics were associated with the odds of setting goals targeting specific behaviors and the frequency of selecting or self-authoring challenges targeting specific behaviors.

Conclusions: Our analyses suggest that, while goal-setting is theoretically grounded, and participants used Buddy in ways that suited their personal needs, this did not translate to measurable behavior change in the study population. The self-authored content showed how participants used the component and provided insights into how they articulate behavior change in terms of personal goals, challenges, strategies for action, motivation plans, and rewards. Future research should explore the conditions under which goal-setting may be more or less effective, to better understand its nuances and potential benefits.

Trial Registration: The UK's Clinical Study Registry ISRCTN23310640; https://www.isrctn.com/ISRCTN23310640

(JMIR Hum Factors 2025;12:e66208) doi:10.2196/66208

KEYWORDS

digital intervention; mHealth; mobile health; public health; fidelity; multiple health behavior change; behavioral challenges; goal-setting; action-planning; coping-planning; factorial randomized controlled trial; college and university students

XSL•FO
RenderX

Introduction

Improving unhealthy behaviors using digital interventions is one of many potential public health strategies for achieving healthier populations. Behavioral changes can reduce the incidence and burden of noncommunicable diseases, including cancers, cardiovascular diseases, and diabetes [1-5]. Individual-level public health interventions can support the adoption of healthy behaviors, including reducing alcohol consumption, eating a healthy diet, engaging in regular physical activity, and smoking cessation. In this context, digital interventions have proven effective in facilitating behavior change across various health behaviors [6-13]. Digital interventions offer scalable and easily accessible support designed to strengthen individual skills and capabilities, for example, by guiding individuals in setting goals for behavior change. However, few digital interventions have addressed multiple health behaviors simultaneously, despite the frequent co-occurrence of unhealthy behaviors [14,15].

In the context of behavioral interventions, goal-setting is recognized as a key behavior change technique [16]. It is commonly operationalized in interventions to focus on both behavioral goals, such as "go for a walk," and outcome-oriented goals, such as "lose weight." The rationale for this approach is grounded in the goal-setting theory developed by Locke and Latham [17,18], which focuses on the motivational aspects of setting goals. The theory emphasizes that specific and challenging goals, coupled with feedback, may lead to increased effort and improved task performance compared to easy, vague, or nonspecific goals, such as "doing one's best." The utility of setting goals is dependent on the individual's commitment to the goal, possession of the necessary skills to attain it, and absence of conflicting goals. Building on this theory, social cognitive theory [19,20] expands the scope of goal-setting by asserting that health behaviors are influenced not only by an individual's goals but also by their understanding of potential health risks and benefits associated with the behavior, their belief in their capability to perform these behaviors (perceived self-efficacy), their expectations of the outcome of their actions, and the perceived facilitators and barriers.

Complementing these theoretical frameworks, interventional research provides further insight into the efficacy of goal-setting within digital health interventions. Studies suggest that interventions are particularly effective when key intervention components include goal-setting, along with action planning, practicing behavior, behavior substitution, habit formation, and problem-solving [21-27]. For instance, a meta-analysis by McEwan et al [28], which included randomized controlled trials where participants set goals for physical activity, reported a medium-sized positive effect of goal-setting (Cohen *d*=0.55; 95% CI 0.43-0.67). Furthermore, a review by Epton et al [29] evaluated the effects of goal-setting across a range of behavior outcomes, including health-related goals (eg, weight loss), sporting goals (eg, performance in sports), cognitive goals (eg, problem-solving), and educational goals (eg, increasing study time). The review found small but robust effect sizes (Cohen *d*=0.34; 95% CI 0.28-0.41), suggesting that goal-setting can be beneficial in various behavioral contexts beyond health behavior change.

While goal-setting has been extensively researched, most studies focus on single-behavior interventions with clearly defined behavioral targets such as smoking cessation or physical activity. In contrast, this study investigated its application within the context of a digital multiple health behavior change intervention, which has not been thoroughly examined in previous research. This approach allowed us to explore how participants interpreted and used goal-setting when behavioral goals were not predefined by the intervention, offering insights into its real-world implementation. Based on its theoretical grounding and wide use in other health behavior interventions, we included goal-setting as one of six potentially effective components in a multiple health behavior change intervention [30]. This digital intervention, called Buddy, targeted university and college students' alcohol, diet, physical activity, and smoking behaviors. The intervention was designed to be open and flexible, allowing participants to use its components based on their individual needs. Our findings revealed no strong and consistent evidence of goal-setting alone impacting any of the measured health behaviors (Asberg et al, unpublished data, 2025). When goal-setting was combined with self-monitoring, there was limited evidence of positive effects on alcohol consumption and smoking 2 months postrandomization, but these effects did not persist at later follow-ups.

If goal-setting is to be an effective component of digital interventions, it is essential to ensure that the component is delivered and received by participants as intended. Evaluating this implementation fidelity can help identify any deviations from the protocol, such as when an intervention component is not delivered or used as planned [31]. Deviations may compromise the impact that an intervention may have; thus, identifying them and revising protocols is important. Despite this, the practical application and fidelity of goal-setting within both digital and face-to-face interventions are rarely evaluated [32,33].

Considering that the evidence we collected did not indicate an effect of goal-setting on behavior, and that the use of goal-setting is understudied in digital multiple behavior interventions, we aimed to explore how participants used this component. To investigate this interaction, we conducted a case study of Buddy with the aim of gaining a deeper understanding of participants' interactions with the goal-setting component of the Buddy intervention. The specific objectives were to identify the characteristics of participants who used this component and to analyze participants' self-authored content.

Methods

Study Design

This study combined quantitative and qualitative data collected from a factorial randomized trial of Buddy, which invited participants from 18 geographically dispersed universities and colleges in Sweden (representing approximately 50% of all) to take part in a behavior change intervention. Participants were recruited through posters, leaflets, email, websites, social media,

XSL•FO
RenderX

and via Student Health Care staff. Interested students registered by texting a dedicated number. Students were eligible for inclusion if they were 18 years or older and met at least one of six criteria for unhealthy behaviors, in accordance with national public health guidelines. These included: (1) having consumed 10/15 (woman/man) or more standard drinks of alcohol in the past week (with one standard drink defined as 12 grams of pure alcohol); (2) having consumed 4/5 (woman/man) or more standard drinks of alcohol on a single occasion at least once in the past month; (3) having consumed less than 500 grams of fruit and vegetables on average per day in the past week; (4) having consumed two or more units of sugary drinks in the past week (one unit defined as 33 cl); (5) having spent less than 150 minutes on moderate and vigorous physical activity the past week; or (6) having smoked at least one cigarette in the past week. Students were excluded if they lacked sufficient proficiency in Swedish to complete the registration or did not have access to a mobile phone, as all study information and intervention materials were delivered digitally in Swedish.

The trial was part of the MoBILE (Mobile Health Multiple Lifestyle Behavior Interventions Across the Lifespan) research program [34], which aims to study digital health interventions across the lifespan, and was designed to evaluate the components of Buddy, of which, goal-setting was one. Buddy is based on previous research of digital health interventions among Swedish university students [35-42], on social cognitive models [43,44], and behavior change techniques [16]. For details of the factorial trial and the intervention, please see the protocol [30].

A screenshot of the goal-setting and challenges component of Buddy is presented in Figure 1. The Buddy intervention aimed to support health behavior change by enhancing self-regulatory skills and capacity through goal-setting and planning. By providing a structured approach, participants were encouraged to set goals and challenges for behavior change in any area they chose and were reminded of these throughout the week via text message. The intervention was accessible to participants over a period of 4 months, with weekly reminders sent on Sundays to access the intervention.

Participants were prompted to set one or more goals for their future behavior, which included: (1) Setting a goal for the upcoming week, participants self-authored what they wanted to achieve; (2) Action-planning, participants formulated two strategies or activities that they planned to use to progress toward their goal; (3) Coping-planning, participants identified and prepared for potential motivational struggles by formulating their approach to self-motivation when faced with barriers; and a (4) Reward-planning, participants selected rewards for themselves upon achieving their goal.

Participants could also create their own challenges or choose from a library of ready-made challenges for the coming week. The challenges were shorter descriptions of tasks, for example, “I will start every day with a 15-minute walk.” Participants were informed that they could return to the system to set new goals and challenges.

XSL•FO
RenderX

Figure 1. Interface of the goal-setting and challenges module in Buddy, showing where participants could self-author goals for the upcoming week (upper section), and select premade challenges addressing alcohol, diet, physical activity, and smoking. Translation: Goal-setting and challenges. Clearly defining your goals is an important step when aiming to make lasting changes in your life. Here, you can create a weekly plan by setting goals and accepting challenges. Throughout the week, you will receive text message reminders to help you stay on track with your plan. Set a goal. My goal for the upcoming week is to: ...For example, engage in physical activity for at least 60 minutes a day, or limit alcohol consumption to no more than 5 units per week. Accept challenges! Tap on the different health behaviors below and choose up to 5 challenges. If you prefer, you can also create your own challenges by tapping the pencil icon. We will send you text message reminders about your challenges throughout the week. Come back next week to select new challenges!



XSL•FO
RenderX

Ethical Considerations

The study received approval from the Swedish Ethical Review Authority on December 15, 2020 (Dnr 2020‐05496). All participants provided informed consent prior to participation. Data were anonymized before analysis to ensure confidentiality. No financial compensation was provided to participants. The trial was preregistered in the ISRCTN registry on January 28, 2021 (ISRCTN23310640).

Data Analysis

Qualitative Analysis

We conducted a qualitative content analysis using a summative approach, as outlined by Hsieh and Shannon [45], to analyze how participants interacted with the goal-setting and challenges component of the Buddy intervention. This approach involved both the quantification of specific content and the interpretation of its contextual meaning, allowing us to explore not only what participants wrote about, but also how they articulated themselves when setting goals and challenges across different health behaviors. The qualitative data consisted of user-generated, self-authored text entries submitted by participants within the intervention. These entries included both self-authored goals and challenges. The self-authored goals were structured entries composed of 5 sentences: one goal formulation, two strategies for achieving the goal, one motivational statement, and one reward. Each goal entry was approximately 60‐80 words long, whereas the self-authored challenges were formulated more succinctly, typically ranging from 2 to 15 words.

All qualitative data were exported from the platform and imported into NVivo12 Plus (Lumivero) for structured coding. We began the qualitative analysis of goals by identifying keywords related to the health behaviors that participants intended to change. Based on these keywords, we grouped the goals according to the targeted behavior, for example, goals related to physical activity were placed in one group, while those related to alcohol use were placed in another. Goals addressing multiple behaviors were categorized independently. Within each behavior group, we then coded the goals by focusing sequentially on content related to strategies, motivations, and finally rewards. This phase involved multiple rounds of coding and iterative refinement through comparing and contrasting codes to ensure attention to detail, accuracy, and consistency. Finally, we developed categories for strategies, motivations, and rewards within each behavior group.

Quantitative Analysis

We used logistic regression to estimate the odds of setting a goal and selecting or self-authoring a challenge. We also used logistic regression to estimate the odds of setting a goal with respect to a specific behavior (among those who had set a goal). We used negative binomial regression to estimate the frequency of selecting or self-authoring different behavioral challenges (among those who had selected or self-authored a challenge). All models included baseline characteristics of participants as covariates, including age, sex, BMI, total weekly alcohol consumption, frequency of heavy episodic drinking per month, number of cigarettes smoked per week, minutes of moderate and vigorous physical activity per week, average daily fruit and vegetable intake, weekly sugary drinks consumption, self-perceived stress, and psychosocial measures (importance, confidence, and know-how).

We used Bayesian inference to estimate the parameters of the regression models, with Student *t* priors centered at 0 with 3 degrees of freedom and a scale of 2.5 (a half-Student *t* prior was used for the dispersion parameter of the negative binomial models). We report the median of the posterior distribution as a point estimate of the association, along with 95% compatibility intervals (CI) defined by the 2.5 and 97.5 quartiles of the posterior distribution. We also report on the probability of an association (POA) as the proportion of the posterior distribution that is above or below the null in the direction of the median.

Results

Overview

The results are presented in four parts, illustrated in Textbox 1. The first part presents an overview of interactions with the component and the characteristics of the study population. The second part provides a detailed summary of the self-authored goals and challenges. The third part explores the associations between participants' characteristics and setting goals and challenges, as well as the associations between participants' characteristics and setting goals and challenges targeting specific behaviors. Finally, the fourth part presents the analysis of participants' self-authored content regarding goals and challenges.

XSL•FO
RenderX

Textbox 1. Overview of the four parts in the results

Interactions with the component

- Access to component
- Number of self-authored goals
- Number of self-authored challenges
- Number of selected challenges

Setting goals and selecting or self-authoring challenges

- Focus of self-authored goals
- Focus of self-authored challenges

Associations with participants' characteristics

- Participants' characteristics and setting goals and challenges
- Participants' characteristics and behavior targeted in goals and challenges

Participants' self-authored content

- Goals
- Strategies
- Motivation
- Reward

Interactions With the Component and Participant Characteristics

From April 13, 2021, to October 18, 2023, we randomized 1704 participants to the factorial trial. In Table 1, baseline characteristics are presented for the entire study population. Access to the goal-setting and challenges component was given to 850 participants, of whom 427 (50%) set at least one goal and 403 (47%) selected or self-authored at least one challenge. A total of 607 goals were set by 427 participants, the majority of whom set one goal (336/427, 79%), 13% (55/427) set two goals, and 8.4% (36/427) set three or more goals. Additionally, 28 participants started setting a goal but did not fill in the entire form. A total of 1506 challenges were selected from 41 premade challenges, with dietary behavior challenges being the most popular (667/1506, 44%), followed by physical activity (635/1506, 42%), alcohol (154/1506, 10%), and smoking (50/1506, 3%). An additional 170 challenges were self-authored by participants. A complete list of the premade challenges available for selection is provided in Multimedia Appendix 1.

XSL•FO
RenderX

Table . Baseline characteristics of randomized participants.

Variables	Participants (N=1704)
Sex, n (%)	
Female	1400 (82.16)
Male	304 (17.84)
Age (years), mean (SD)	28 (6.9)
BMI (kg/m^2), n (%)	
Under (<18.5)	60 (3.52)
Normal (18.5 - 24.9)	844 (49.53)
Over (25.0 - 29.9)	458 (26.88)
Obese (≥30)	342 (20.07)
Alcohol, mean (SD)	
Total weekly alcohol consumption	3.7 (5.2)
Frequency of heavy episodic drinking	1.8 (2.7)
Smoking	
Number of smokers (smoked ≥1 cigarette per past week), n (%)	218 (12.79)
Number of cigarettes last week, mean (SD)	3.5 (16.6)
Physical activity	
Weekly moderate and vigorous physical activity (minutes), mean (SD)	207 (214)
Dietary behavior, mean (SD)	
Average daily fruit and vegetable consumption, portions per week	1.3 (1.1)
Sugary drinks, cans per week	2.7 (4.1)
Candy and snacks, portions per week	6.5 (6.4)
Stress, mean (SD)	
Self-perceived stress[a]	7.8 (2.9)
Psychosocial measures, median (IQR)	
Importance[b]	8 (7-10)
Confidence[b]	6 (5-8)
Know-how[b]	6 (5-8)

[a]4-item perceived stress scale with total score ranging from 0 to 16.

[b]Single item with 1-10 response options.

Setting Goals and Selecting or Self-Authoring Challenges

Table 2 shows that most participants set goals for physical activity (n=302) and dietary behavior (n=140). The goals highlighted the interconnectedness of health behaviors as participants set goals targeting multiple behaviors (n=53), typically combining physical activity with other behaviors like diet, alcohol, smoking, and sleep. Participants took the opportunity to set other kinds of behavior goals that focused on a variety of miscellaneous aspects (n=72), such as study-related performance, mental health, sleep, physical body, and mobile phone use. Meanwhile, fewer participants set goals for alcohol (n=19) and tobacco (n=17), including both smoking and snus.

When contrasting the self-authored goals (Table 2) with the self-authored challenges (n=170) in Table 3, similarities emerged regarding the choice of behaviors and activities to be conducted. Challenges, like goals, were primarily about physical activity (n=60), dietary behavior (n=30), mental health (n=19), tobacco use (n=14), study performance (n=12), sleep routines (n=8), leisure activities (n=6), alcohol consumption (n=5), and physical health (n=4). These challenges mirrored the goals' structure, being articulated as clear directives, intentions, or actions formulated in the first person singular, such as "I shall" or "I shall not." For example, "I will run every weekday morning this week," "Eat salad with every meal," "No smoking before 18.00," and "I will only drink on Friday and Saturday between 19 - 22."

XSL•FO
RenderX

Table . Focus of self-authored goals.

Focus of self-authored goals	Number of self-authored goals (n=607)
Physical activity	302
Dietary behavior	140
Miscellaneous[a]	72
Multiple health behaviors[b]	53
Alcohol consumption	19
Tobacco including snus	17
Unspecified	4

[a]Include performance goals, positive mental health goals, sleep behavior goals, physical body goals, and mobile phone use goals.

[b]Include diet and physical activity goals, combination of behaviors (three or more), physical activity and alcohol goals, physical activity and smoking goals, physical activity and other routine goals, and alcohol or smoking and diet goals.

Table . Focus of self-authored challenges.

Focus of self-authored challenges	Number of self-authored challenges (n=170)
Physical activity[a]	60
Dietary behavior[b]	30
Mental health: stress reduction, meditation, reflection, gratitude, and self-compassion	19
Tobacco use: reduce and refrain from cigarettes and snus	14
Study performance behavior: time and task commitment for completion	12
Sleeping routines: establish sleep routines	8
Leisure activities: prioritize leisure activities	6
Alcohol consumption: restrictions on consumption	5
Physical health and well-being: prioritize activities for physical health	4

[a]Number of days exercising per week (n=30), amount of exercise per day (n=16), and specified type of activity (n=14).

[b]Eat regularly and eat more vegetables (n=23) and reduce snacking (n=19).

Participant Characteristics and Setting Goals and Challenges

Table S1 in Multimedia Appendix 2 shows associations between baseline characteristics and setting a goal. These analyses reveal that the odds were smaller for men to write a goal compared to women (odds ratio [OR] 0.44, 95% CI 0.30-0.66; POA>99.9%). In addition, those with higher consumption of candy and snacks had higher odds of writing a goal; however, the association did not appear as strong (OR 1.03, 95% CI 1.00-1.05; POA=98%). Those who rated the importance of changing their health behaviors as higher at baseline also had higher odds of setting a goal (OR 1.09, 95% CI 1.00-1.18; POA=97.4%). Finally, although the evidence was weaker, smoking more cigarettes per week at baseline was associated with higher odds of setting a goal (OR 1.12, 95% CI 0.96-1.32; POA=92%).

Table S2 in Multimedia Appendix 2 shows associations between baseline characteristics and selecting or self-authoring a challenge. As was the case for goal-setting, men had a lower odds of selecting challenges compared to women (OR 0.46, 95% CI 0.30-0.67; POA>99.9%). There was also evidence, albeit weaker, that those who rated higher importance of change had higher odds of interacting with challenges (OR 1.07, 95% CI 0.98-1.16; POA=93.6%). Additionally, there was some evidence suggesting that individuals with high BMI (obese) were more likely than individuals with normal BMI to interact with the challenges component (OR 1.32, 95% CI 0.89-1.97; POA=91.4%).

Participant Characteristics and Behavior Targeted in Goals and Challenges

Overview

Tables S3-S16 in Multimedia Appendix 2 show associations between participants' characteristics and setting goals targeting specific behaviors, and the frequency of selecting or self-authoring challenges targeting specific behaviors. These analyses were done among participants who had set a goal or selected a challenge, respectively. Here, we give a summary of the findings presented in the tables.

Man Versus Woman and Age

When comparing men versus women, the strongest associations were found concerning setting goals targeting alcohol and diet, with the odds being lower for men. This pattern was also evident for selecting challenges related to alcohol and diet, where men selected fewer of these compared to women. In addition, men

XSL•FO
RenderX

selected fewer challenges related to physical activity than women. Older participants had higher odds of setting goals targeting alcohol consumption and selected fewer challenges relating to diet.

Alcohol Consumption

Participants who reported a higher weekly consumption of alcohol at baseline had lower odds of setting goals targeting diet and higher odds of setting goals targeting multiple health behaviors and miscellaneous goals. The multiple health behavior goals concerned combinations of behaviors such as physical activity, diet, alcohol, and smoking, while miscellaneous goals concerned behaviors not specifically targeted by the intervention, such as sleep. Those with higher weekly alcohol consumption also selected more challenges relating to alcohol consumption, but fewer challenges relating to physical activity. More frequent heavy episodic drinking at baseline was associated with selecting more alcohol-related challenges and writing fewer self-authored challenges.

Smoking

Participants who reported being more frequent smokers at baseline had higher odds of setting tobacco-related goals and had lower odds of setting goals for all other behaviors (except for multiple health behavior goals). These participants also selected more smoking challenges and fewer diet challenges.

Dietary Behavior

Participants with a higher fruit and vegetable intake at baseline had higher odds of setting goals related to physical activity and were less likely to set goals related to alcohol, diet, and tobacco. Similarly, these participants selected more physical activity and smoking challenges, but fewer diet challenges. Participants who reported a higher weekly consumption of sugary drinks had higher odds of setting multiple health behavior goals. Finally, participants who had higher candy and snacks intake at baseline selected more diet and self-authored challenges.

BMI

Participants with higher BMI (overweight and obese) had lower odds of setting goals targeting alcohol, as well as selecting fewer alcohol-related challenges, compared to those with normal BMI. Participants in the highest BMI category (obese) had higher odds of setting goals relating to diet and lower odds of setting goals targeting multiple health behaviors. Participants with lower BMI (<18.5) were more likely to set goals targeting diet but selected fewer challenges for diet compared to those with normal BMI. The participants in the underweight category were also less likely to set goals targeting tobacco. Participants in the edge BMI categories (underweight and obese) selected fewer challenges relating to physical activity and smoking, compared to those in the normal BMI category.

Physical Activity

Participants with higher levels of physical activity at baseline had higher odds of setting goals targeting diet and miscellaneous goals, and lower odds of setting physical activity goals. They also authored more of their own challenges and selected fewer physical activity challenges.

Stress Levels

Participants who experienced higher levels of stress relative to the group mean had higher odds of setting miscellaneous goals and lower odds of setting goals targeting alcohol. These participants also selected more smoking-related challenges.

Psychosocial Measures

Participants who rated the importance of changing their behavior as higher relative to the group mean at baseline had higher odds of setting goals targeting alcohol consumption and self-authored fewer challenges. There were no marked associations between confidence in being able to change with goal-setting and challenges. Finally, those who rated their knowledge of how to change their behavior higher relative to the group mean had lower odds of setting goals targeting alcohol and smoking and more frequently self-authored challenges.

Participants' Self-Authored Content From Buddy

The content of the self-authored goals and challenges was qualitatively analyzed. The coding process resulted in a total of 2412 codes. They focused primarily on health improvement, behavior modification, and self-improvement. A summary of all self-authored goals, strategies, motivations, rewards, and challenges can be found in Multimedia Appendix 3.

Participants Articulated Goals Differently Depending on Health Behavior

Goals related to alcohol, tobacco, and physical activity were often quantifiable, focusing on aspects like quantity, frequency, type, or duration. Physical activity goals were typically more specific, detailing activity duration or training frequency. Dietary behavior goals, while also specific, focused more on restrictions, as well as establishing desirable healthy patterns. As expected, considering the complexity of diet, these goals, unlike other health behaviors, included both preventive measures, such as avoiding unhealthy food products, and promotive aspects like incorporating healthy food choices. This contrasted with physical activity goals, which solely promoted increased activity, and alcohol and tobacco goals, which were purely preventive, such as limiting, reducing, or abstaining from the behavior.

> *In the upcoming week, my goal is to...*
>
> *...not drink alcohol on weekdays (Alcohol)*
>
> *...not smoke more than 4 cigarettes a day (Tobacco)*
>
> *...go for a walk every morning before studying (Physical activity)*
>
> *...not eat sugar and swap to fruit and berries instead... (Diet)*
>
> *...study at least 4 hours a day (Miscellaneous)*
>
> *...consume 500 grams of vegetables daily and work out 5 times (Multiple)*

Participants Intend to Use Proactive Strategies Such as Planning and Preparation

When formulating goals, participants were asked to formulate strategies to achieve their goals. Planning, preparation, and prioritization were central to all behaviors and were pronounced as specific, detailed, and proactive strategies on how to achieve

XSL•FO
RenderX

goals. This involved setting new routines, being mindful, and focusing on goal-oriented activities.

Participants articulated diverse activities and combinations of strategies. Dietary strategies included meal planning and availability of healthy alternatives, preparation of healthy food, and bringing fruit or lunch boxes to university. Physical activity strategies involved scheduling activities, making room for exercise in one's study schedule, adjusting routines, setting reminders, and finding motivational tools such as podcasts or social support. Alcohol and tobacco strategies involved heightened awareness of what is important to them and planned modifications to habits, such as slowing drinking pace, not carrying cigarettes, leaving parties earlier, or making environmental changes like avoiding alcohol-related activities.

> *In the upcoming week, I will undertake these two actions to progress towards my goals...*
>
> *...buy fruits to eat instead and eat proper breakfast, lunch, and dinner (Diet)*
>
> *...by going to bed on time and making a lunchbox the day before (Physical activity)*
>
> *...make some plans for Saturday night instead of going to the pub (Alcohol)*
>
> *...chew gum, drink tea, and do breathing exercises (Tobacco)*
>
> *...to go to bed on time and not sleep too late during the day (Miscellaneous)*
>
> *...schedule training and say no to cigarettes (Multiple)*

Physical activity strategies were typically action-oriented, specific, and time-framed, closely aligning with the goals. For instance, a goal to walk 20 minutes per day could involve the strategy to go for a morning and a poststudy walk. Conversely, dietary, tobacco, and alcohol strategies involved a broader range of activities like distractions, restrictions, substitutes, and limitations of unhealthy alternatives, as well as planning and preparing for healthy choices. For example, a typical goal to limit drinking involved the strategy to abstain from drinking before going out or swapping to alcohol-free alternatives. A typical goal to avoid unhealthy food involved the strategy to abstain from buying and replacing dessert with fruit and berries instead. Physical activity was also articulated as a strategy to achieve dietary, alcohol, and tobacco goals, such as being physically active as a means to divert attention from snacking, smoking, or drinking.

Participants Motivate Themselves With Health Benefits, Self-Encouragement, and Goal Focus

When formulating strategies, participants were asked to formulate how to motivate themselves when faced with barriers to achieve their goals. Participants' motivational self-talk, across various health behaviors, was expressed in a similar manner, primarily focusing on well-being, health benefits, self-encouragement, and goal orientation. The self-talk emphasized both short- and long-term well-being and health benefits. For tobacco, the focus was mainly on the long-term benefits, while physical activity and alcohol were more focused on short-term well-being, such as "feeling better afterwards, or the next day." Participants recalled and expressed previous experiences of feeling good when adopting a healthy behavior. For example, "I always feel better after exercising" (Physical activity), "I will be feeling so much better if I don't snack and avoid sweets" (Diet), "This will bring me energy" (Multiple), "I will avoid anxiety the day after if I drink moderately" (Alcohol), "Smoking is harmful to me and my asthma will improve without it" (Tobacco), and "I will feel good from exercise and less screen time. It will help" (Miscellaneous).

Participants used self-encouraging positive affirmations like "you can do it" to motivate themselves. In the context of physical activity, this included pushing oneself to work hard with a just-do-it attitude, overcoming barriers, and reminding themselves of the necessity of effort for success. For diet, tobacco, and alcohol goals, the focus shifted to reminding oneself that there are better alternatives, encouraging perseverance, and strengthening one's capability to combat craving, all with the aim of making oneself proud. Overall, participants reminded themselves of their capability and determination with affirmations like "I can and I will. They also reassured themselves of their past accomplishments, reinforcing their belief in their capabilities and the manageability of the task.

> *If it feels tough, I will tell myself that...*
>
> *...get a grip, it goes quickly and it's fun! (Physical activity)*
>
> *...I actually prefer and feel better with good, cooked food (Diet)*
>
> *...I feel much better during the week if I don't drink (Alcohol)*
>
> *...that you cannot afford it, and that I have succeeded in several demanding challenges before (Tobacco)*
>
> *...you can do it, you are goal-oriented and wonderful (Miscellaneous)*
>
> *...you must keep your promises to yourself so that you can trust yourself (Multiple)*

Regarding goal orientation, participants reminded themselves to stay focused on the goal, its significance, and why it is important, for example, the long-term benefits, self-improvement, or success with a task or studies. For miscellaneous behaviors, participants also highlighted the importance of self-value, personal sustainability, and life balance by encouraging themselves to focus on the task and dedicate time for health.

Participants Plan to Reward Themselves With Shopping, Indulgence, and Self-Appreciation

To meet their goal, participants were asked to articulate how they would reward themselves when achieving their goal. These rewards varied, but overall included shopping, indulging in something to eat, and engaging in enjoyable activities. Regardless of health behavior, participants also planned to express self-appreciation by acknowledging and recognizing their progress and effort. This often took the form of symbolic gestures like "a pat on the back," applause, a hug, or a sense of pride.

XSL•FO
RenderX

Shopping was a popular reward, with participants planning to buy clothing, beauty products, craft materials, electronics, or home decor. Some participants planned to indulge in culinary treats, such as cooking their favorite meal, dining out, or purchasing "special items" like sushi, exotic fruits, or barista coffee. Among those who set goals related to tobacco and alcohol, some planned to reward themselves with the ability to drink or smoke. Engaging in enjoyable activities was also a common reward. These activities ranged from socializing, movie nights, and gaming to outdoor activities like walking in nature or self-care activities like taking a day off from studying. Some planned to treat themselves with a visit to the hairdresser, a tattoo artist, or a massage, while others planned a trip. Some participants planned to save or invest the money they would have otherwise spent on unhealthy behaviors.

Discussion

Principal Findings

We conducted this study to gain insight into the fidelity of a digital multiple lifestyle behavior intervention by investigating participants' interactions with a goal-setting and challenges component. The self-authored content shed light on how participants used the goal-setting component and provided insights into how participants articulate behavior change in terms of personal goals, challenges, strategies for action, motivation plans, and rewards. Overall, our findings indicate that participants who engaged with the intervention component did so as intended, demonstrating the feasibility of incorporating goal-setting into a multiple health behavior intervention. Despite this, as mentioned earlier, in our analyses of the effects of goal-setting on behavior, we found no strong evidence that the component was effective (Asberg et al, unpublished data, 2025). This may partially be explained by the use being heterogeneous, with many not using the component at all, and individual characteristics determining engagement within the component. However, the combined findings of fidelity and effectiveness suggest that while components of behavioral interventions may be grounded in theory and used as intended, this does not always translate to behavior change. This calls for further research to explore the conditions under which goal-setting may be more or less effective, to better understand its nuances and potential benefits.

Our findings show that participants' baseline characteristics were associated with setting goals targeting specific behaviors and the frequency of selecting or self-authoring challenges targeting specific health behaviors. These associations were in line with what was anticipated and intended; for instance, participants with higher alcohol consumption focused on alcohol, while smokers prioritized smoking cessation. Meanwhile, despite Buddy's strong emphasis on alcohol, diet, physical activity, and smoking, those experiencing stress concentrated on goals and challenges concerning managing studies, mental well-being, sleep, and mobile phone use. Our findings also suggest that participants who already practice some healthy behaviors, such as consuming fruits and vegetables or engaging in physical activity, were more inclined to focus on other health areas where they may not be following guidelines. Overall, our findings suggest that participants used the intervention component to address their personal needs for support in changing unhealthy health behaviors.

When analyzing the content of the goals and challenges set by participants, we found that strategies, motivations, and rewards focused on proactive activities like planning, preparation, and prioritization. Participants were motivated by reminding themselves of the health benefits, self-encouragement, and maintaining a focus on their goals. These motivations had both a short- and long-term perspective. Participants supported themselves with positive affirmations, such as "you can do it," and richer messages of self-improvement, self-worth, and sustainable life balance.

Comparison With Prior Work

Previous studies have shown that goal-setting can lead to improved outcomes in various health behaviors. For instance, a web-based intervention [46] showed that weekly goal-setting can increase fruit and vegetable consumption among university students over time, but no corresponding impact on physical activity levels was observed. Researchers suggest that while goal-setting can be beneficial in initiating new behaviors, it might not be as effective for enhancing behaviors that individuals are already practicing. In a factorial trial [47], researchers compared adaptive goals versus static goals, and the effect of immediate versus delayed financial rewards on increasing physical activity. The result indicated that static goals initially increased daily step count but subsequently declined physical activity levels over time. Conversely, those with adaptive goals showed a smaller initial increase but maintained a more consistent level of activity throughout the 4-month intervention period. Notably, the activity levels of the static group eventually fell below those of the adaptive goal group, even with the incentive of immediate rewards.

In a study by Padovano et al [48], the relationship between daily goal-setting and its impact on alcohol consumption was investigated among individuals with alcohol use disorder. The study revealed that the most commonly selected goal was complete abstinence ("I have a plan to not drink"; n=858, 31%), followed by moderation goals ("2 or fewer drinks"; n=604, 22%), or ("5 or fewer drinks"; n=618, 22%). The odds of drinking were reduced by 89% when an abstinence goal was set (OR 0.11, 95% CI 0.08-0.15). Additionally, the study observed that those who more frequently set abstinence goals were the ones who drank more per drinking occasion. These formulations align with the self-authored alcohol and tobacco goals in this study, which were framed in a preventive manner—limiting, reducing, or abstaining from the behavior—and were predominantly formulated by participants with higher alcohol and smoking behaviors.

Similarly, the self-authored goals in this study, particularly those related to physical activity, aligned with the self-authored goals set by participants in a digital weight-loss intervention [49]. In that study, participants were encouraged to set weekly behavioral goals and share their goals and accomplishments on a shared social media platform. Findings indicate that goals expressed in repeating terms, or as both measurable and repeating terms (eg, "My exercise goal is to walk for at least 15 minutes a day

XSL•FO
RenderX

this week") were associated with greater weight loss. These repeating term goals, for example, articulating the number of training days per week, to eat fruit every day, to meditate, or to study for a specified number of hours per day, were also articulated among the students in this study.

Limitations

This study has a number of limitations that should be considered. First, associations between baseline characteristics and use of the component are not causal and should not be interpreted as such. These analyses describe who among the participants did what, but not why they did so. Second, the study did not assess whether interaction with the goal-setting component translated into behavior change. To establish causal relationships and better understand the impact of goal-setting on health behaviors, future studies could randomize participants to different goal-setting strategies—such as predefined versus self-authored goals to evaluate their differential effects on behavior change outcomes. Third, half of the participants with access to the goal-setting and challenges component did not choose to engage with it. While we do not expect all participants to interact with digital interventions that have low barriers to entry, it remains an important question to investigate reasons why the remaining participants, despite signing up for Buddy, did not use this module. Future studies could include follow-up with nonusers to better understand barriers to interaction, such as lack of time, perceived relevance, or competing academic or personal demands. Gaining insights into these factors could inform the design of digital interventions that better align with users' contexts and needs.

Moreover, among those who did engage with the goal-setting component, the majority set one or two goals (n=391, 92%). This could be due to various reasons, including satisfaction with the initial goal or behavior change progress, which might have eliminated the need to set other goals. The process of self-authoring goals may be cognitively demanding, and not all participants are likely to benefit equally from a goal-setting component. Some may find the task of setting and achieving goals difficult, overwhelming, or simply outside the scope of their current priorities. Additionally, the absence of feedback on progress or positive reinforcement from Buddy may have contributed to a lack of motivation to set new goals. Without encouragement to revisit or revise their goals, participants may have disengaged over time.

Conclusions

Our analyses of self-authored content from the goal-setting and challenges component in a digital intervention for multiple health behavior change provided insights into how participants articulate behavior change in terms of personal goals, challenges, strategies for action, motivation plans, and rewards. Participants set goals related to both single and multiple health behaviors, which is in alignment with the intended purpose of Buddy. Furthermore, participants used Buddy in ways that suited and supported their personal behavior change needs, sometimes extending beyond the scope of the lifestyle behaviors that Buddy targeted. Future research should explore the conditions under which goal-setting may be more or less effective, to better understand its nuances and potential benefits.

Acknowledgments

This study was funded by the Swedish Research Council for Health, Working Life, and Welfare (grant 2018 - 01410; Principal Investigator: ML).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

This study is part of the MoBILE (Mobile health multiple Behavior Interventions across the LifEspan) research program, which was designed by ML and MB. KÅ, MB, and ML contributed to conceptualization. Formal analysis was done by KÅ and MB. ML and MB contributed to funding acquisition. KÅ contributed to the investigation, and KÅ and MB contributed to the methodology. MB and ML supervised the study. Validation was done by MB and ML. Writing—original draft was done by KÅ, and all the authors contributed to writing—review and editing.

Conflicts of Interest

MB owns a private company (Alexit AB) that develops and distributes eHealth solutions to the public and private sectors. Alexit AB played no role in developing the intervention, study design, data analysis, data interpretation, or writing of this report.

Multimedia Appendix 1
A complete list of the premade challenges available for selection is provided in Appendix 1.
[DOCX File, 19 KB - humanfactors_v12i1e66208_app1.docx]

Multimedia Appendix 2
Associations between baseline characteristics and setting a goal.
[DOCX File, 60 KB - humanfactors_v12i1e66208_app2.docx]

XSL•FO
RenderX

Multimedia Appendix 3
A summary of all self-authored goals, strategies, motivations, rewards, and challenges can be found in Appendix 3.
[DOCX File, 69 KB - humanfactors_v12i1e66208_app3.docx]

References

1. GBD 2020 Alcohol Collaborators. Population-level risks of alcohol consumption by amount, geography, age, sex, and year: a systematic analysis for the Global Burden of Disease Study 2020. Lancet 2022;400(10347):185-235. [doi: 10.1016/S0140-6736(22)02121-3]
2. Reitsma MB, Kendrick PJ, Ababneh E, et al. Spatial, temporal, and demographic patterns in prevalence of smoking tobacco use and attributable disease burden in 204 countries and territories, 1990–2019: a systematic analysis from the Global Burden of Disease Study 2019. Lancet 2021 Jun;397(10292):2337-2360. [doi: 10.1016/S0140-6736(21)01169-7]
3. Ong KL, Stafford LK, McLaughlin SA, et al. Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: a systematic analysis for the Global Burden of Disease Study 2021. Lancet 2023 Jul;402(10397):203-234. [doi: 10.1016/S0140-6736(23)01301-6]
4. Brauer M, Roth GA, Aravkin AY, et al. Global burden and strength of evidence for 88 risk factors in 204 countries and 811 subnational locations, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021. Lancet 2024 May;403(10440):2162-2203. [doi: 10.1016/S0140-6736(24)00933-4]
5. Noncommunicable diseases: fact sheet. World Health Organization. 2023. URL: https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases [accessed 2025-03-28]
6. Oosterveen E, Tzelepis F, Ashton L, Hutchesson MJ. A systematic review of eHealth behavioral interventions targeting smoking, nutrition, alcohol, physical activity and/or obesity for young adults. Prev Med 2017 Jun;99:197-206. [doi: 10.1016/j.ypmed.2017.01.009] [Medline: 28130046]
7. Schoeppe S, Alley S, Van Lippevelde W, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. Int J Behav Nutr Phys Act 2016 Dec;13(1). [doi: 10.1186/s12966-016-0454-y]
8. Mönninghoff A, Kramer JN, Hess AJ, et al. Long-term effectiveness of mHealth physical activity interventions: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2021 Apr 30;23(4):e26699. [doi: 10.2196/26699] [Medline: 33811021]
9. Sevic A, Hashemi NS, Thørrisen MM, et al. Effectiveness of eHealth Interventions targeting employee health behaviors: systematic review. J Med Internet Res 2023 Apr 20;25:e38307. [doi: 10.2196/38307] [Medline: 37079369]
10. Kaner EF, Beyer FR, Muirhead C, et al. Effectiveness of brief alcohol interventions in primary care populations. Cochrane Database Syst Rev 2018 Feb 24;2(2):CD004148. [doi: 10.1002/14651858.CD004148.pub4] [Medline: 29476653]
11. O Rourke L, Humphris G, Baldacchino A. Electronic communication based interventions for hazardous young drinkers: a systematic review. Neurosci Biobehav Rev 2016 Sep;68:880-890. [doi: 10.1016/j.neubiorev.2016.07.021] [Medline: 27453570]
12. Taylor GMJ, Dalili MN, Semwal M, Civljak M, Sheikh A, Car J. Internet-based interventions for smoking cessation. Cochrane Database Syst Rev 2017 Sep 4;2017(9):CD007078. [doi: 10.1002/14651858.CD007078.pub5] [Medline: 28869775]
13. McIntosh JRD, Jay S, Hadden N, Whittaker PJ. Do E-health interventions improve physical activity in young people: a systematic review. Public Health (Fairfax) 2017 Jul;148:140-148. [doi: 10.1016/j.puhe.2017.04.001] [Medline: 28482322]
14. Milne-Ives M, Lam C, De Cock C, Van Velthoven MH, Meinert E. Mobile apps for health behavior change in physical activity, diet, drug and alcohol use, and mental health: systematic review. JMIR Mhealth Uhealth 2020;8(3):e17046. [doi: 10.2196/17046]
15. Noble N, Paul C, Turon H, Oldmeadow C. Which modifiable health risk behaviours are related? A systematic review of the clustering of Smoking, Nutrition, Alcohol and Physical activity ('SNAP') health risk factors. Prev Med 2015 Dec;81:16-41. [doi: 10.1016/j.ypmed.2015.07.003]
16. Michie S, Richardson M, Johnston M, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
17. Locke EA, Latham GP. Building a practically useful theory of goal setting and task motivation. A 35-year odyssey. Am Psychol 2002 Sep;57(9):705-717. [doi: 10.1037//0003-066x.57.9.705] [Medline: 12237980]
18. Locke EA, Latham GP. New directions in goal-setting theory. Curr Dir Psychol Sci 2006 Oct;15(5):265-268. [doi: 10.1111/j.1467-8721.2006.00449.x]
19. McAlister AL, Perry CL, Parcel GS. How individuals, environments, and health behaviors interact: social cognitive theory. In: Glanz K, Rimer BK, Viswanath K, editors. Health Behavior and Health Education: Theory, Research, and Practice, 4th edition: Jossey-Bass/Wiley; 2008:169-188.
20. Bandura A. Health promotion by social cognitive means. Health Educ Behav 2004 Apr;31(2):143-164. [doi: 10.1177/1090198104263660]
21. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol 2009;28(6):690-701. [doi: 10.1037/a0016136]

XSL•FO
RenderX

22. Knittle K, Nurmi J, Crutzen R, Hankonen N, Beattie M, Dombrowski SU. How can interventions increase motivation for physical activity? A systematic review and meta-analysis. Health Psychol Rev 2018 Jul 3;12(3):211-230. [doi: 10.1080/17437199.2018.1435299] [Medline: 29385950]
23. Howlett N, Trivedi D, Troop NA, Chater AM. Are physical activity interventions for healthy inactive adults effective in promoting behavior change and maintenance, and which behavior change techniques are effective? A systematic review and meta-analysis. Transl Behav Med 2019 Jan 1;9(1):147-157. [doi: 10.1093/tbm/iby010] [Medline: 29506209]
24. Ashton LM, Sharkey T, Whatnall MC, et al. Effectiveness of interventions and behaviour change techniques for improving dietary intake in young adults: a systematic review and meta-analysis of RCTs. Nutrients 2019 Apr 11;11(4):825. [doi: 10.3390/nu11040825] [Medline: 30979065]
25. Garnett CV, Crane D, Brown J, et al. Behavior change techniques used in digital behavior change interventions to reduce excessive alcohol consumption: a meta-regression. Ann Behav Med 2018 May 18;52(6):530-543. [doi: 10.1093/abm/kax029] [Medline: 29788261]
26. McCrabb S, Baker AL, Attia J, et al. Internet-based programs incorporating behavior change techniques are associated with increased smoking cessation in the general population: a systematic review and meta-analysis. Ann Behav Med 2019 Feb 1;53(2):180-195. [doi: 10.1093/abm/kay026]
27. Michie S, Whittington C, Hamoudi Z, Zarnani F, Tober G, West R. Identification of behaviour change techniques to reduce excessive alcohol consumption. Addiction 2012 Aug;107(8):1431-1440. [doi: 10.1111/j.1360-0443.2012.03845.x]
28. McEwan D, Harden SM, Zumbo BD, et al. The effectiveness of multi-component goal setting interventions for changing physical activity behaviour: a systematic review and meta-analysis. Health Psychol Rev 2016 Jan 2;10(1):67-88. [doi: 10.1080/17437199.2015.1104258]
29. Epton T, Currie S, Armitage CJ. Unique effects of setting goals on behavior change: systematic review and meta-analysis. J Consult Clin Psychol 2017;85(12):1182-1198. [doi: 10.1037/ccp0000260]
30. Åsberg K, Lundgren O, Henriksson H, et al. Multiple lifestyle behaviour mHealth intervention targeting Swedish college and university students: protocol for the Buddy randomised factorial trial. BMJ Open 2021 Dec;11(12):e051044. [doi: 10.1136/bmjopen-2021-051044]
31. Borrelli B. The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. J Public Health Dent 2011 Jan;71(s1). [doi: 10.1111/j.1752-7325.2011.00233.x]
32. Hawkes RE, Warren L, Cameron E, French DP. An evaluation of goal setting in the NHS England diabetes prevention programme. Psychol Health 2022 Feb 1;37(2):131-150. [doi: 10.1080/08870446.2021.1872790]
33. Lorencatto F, West R, Bruguera C, Brose LS, Michie S. Assessing the quality of goal setting in behavioural support for smoking cessation and its association with outcomes. Ann Behav Med 2016 Apr;50(2):310-318. [doi: 10.1007/s12160-015-9755-7]
34. Bendtsen M, Bendtsen P, Henriksson H, et al. The Mobile Health Multiple Lifestyle Behavior Interventions Across the Lifespan (MoBILE) research program: protocol for development, evaluation, and implementation. JMIR Res Protoc 2020;9(4):e14894. [doi: 10.2196/14894]
35. Bendtsen P, McCambridge J, Bendtsen M, Karlsson N, Nilsen P. Effectiveness of a proactive mail-based alcohol internet intervention for university students: dismantling the assessment and feedback components in a randomized controlled trial. J Med Internet Res 2012;14(5):e142. [doi: 10.2196/jmir.2062]
36. McCambridge J, Bendtsen M, Karlsson N, White IR, Nilsen P, Bendtsen P. Alcohol assessment and feedback by email for university students: main findings from a randomised controlled trial. Br J Psychiatry 2013 Nov;203(5):334-340. [doi: 10.1192/bjp.bp.113.128660]
37. Bendtsen P, Bendtsen M, Karlsson N, White IR, McCambridge J. Online alcohol assessment and feedback for hazardous and harmful drinkers: findings from the AMADEUS-2 randomized controlled trial of routine practice in Swedish universities. J Med Internet Res 2015;17(7):e170. [doi: 10.2196/jmir.4020]
38. Thomas K, Linderoth C, Bendtsen M, Bendtsen P, Müssener U. Text message-based intervention targeting alcohol consumption among university students: findings from a formative development study. JMIR Mhealth Uhealth 2016 Oct 20;4(4):e119. [doi: 10.2196/mhealth.5863] [Medline: 27765732]
39. Müssener U, Thomas K, Linderoth C, Leijon M, Bendtsen M. A text message–based intervention targeting alcohol consumption among university students: user satisfaction and acceptability study. JMIR Hum Factors 2018;5(3):e23. [doi: 10.2196/humanfactors.9641]
40. Thomas K, Müssener U, Linderoth C, Karlsson N, Bendtsen P, Bendtsen M. Effectiveness of a text messaging-based intervention targeting alcohol consumption among university students: randomized controlled trial. JMIR Mhealth Uhealth 2018 Jun 25;6(6):e146. [doi: 10.2196/mhealth.9642] [Medline: 29941417]
41. Müssener U, Bendtsen M, Karlsson N, White IR, McCambridge J, Bendtsen P. Effectiveness of short message service text-based smoking cessation intervention among university students. JAMA Intern Med 2016 Mar 1;176(3):321. [doi: 10.1001/jamainternmed.2015.8260] [Medline: 26903176]
42. Müssener U, Bendtsen M, McCambridge J, Bendtsen P. User satisfaction with the structure and content of the NEXit intervention, a text messaging-based smoking cessation programme. BMC Public Health 2016 Nov 22;16(1):1179. [doi: 10.1186/s12889-016-3848-5] [Medline: 27876031]

XSL•FO
RenderX

43. Fishbein M, Triandis HC, Kanfer FH, Becker M, Middlestadt SE, Eichler A. Factors influencing behaviour and behaviour change. In: Baum A, Revenson TA, Singer JE, editors. Handbook of Health Psychology: Psychology Press; 2001:3-17.
44. Conner M, Norman P. Predicting Health Behaviour: Research and Practice with Social Cognition Models: Open University Press; 2005.
45. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687]
46. O'Donnell S, Greene GW, Blissmer B. The effect of goal setting on fruit and vegetable consumption and physical activity level in a web-based intervention. J Nutr Educ Behav 2014;46(6):570-575. [doi: 10.1016/j.jneb.2014.03.005] [Medline: 24857600]
47. Adams MA, Hurley JC, Todd M, et al. Adaptive goal setting and financial incentives: a 2 × 2 factorial randomized controlled trial to increase adults' physical activity. BMC Public Health 2017 Dec;17(1). [doi: 10.1186/s12889-017-4197-8]
48. Padovano HT, Levak S, Vadhan NP, Kuerbis A, Morgenstern J. The role of daily goal setting among individuals with alcohol use disorder. Drug Alcohol Depend Rep 2022 Mar;2:100036. [doi: 10.1016/j.dadr.2022.100036] [Medline: 36303709]
49. Jake-Schoffman DE, Waring ME, DiVito J, Goetz JM, Pan C, Pagoto SL. The relationship between how participants articulate their goals and accomplishments and weight loss outcomes: secondary analysis of a pilot of a web-based weight loss intervention. JMIR Mhealth Uhealth 2023;11:e41275. [doi: 10.2196/41275]

Abbreviations

CI: compatibility interval
MoBILE: Mobile Health Multiple Lifestyle Behavior Interventions Across the Lifespan
OR: Odds ratio
POA: Probability of an association

Edited by A Kushniruk; submitted 06.09.24; peer-reviewed by A Eisingerich, P Cornish, YC Liu; revised version received 17.06.25; accepted 10.07.25; published 29.08.25.

Please cite as:
Åsberg K, Löf M, Bendtsen M
Setting Goals and Accepting Challenges for Behavior Change—Analysis of Participants' Interactions With a Digital Multiple Health Behavior Intervention: Mixed Methods Study
JMIR Hum Factors 2025;12:e66208
URL: https://humanfactors.jmir.org/2025/1/e66208
doi:10.2196/66208

© Katarina Åsberg, Marie Löf, Marcus Bendtsen. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 29.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Mobile App Designed to Promote Shared Decision-Making in the Treatment of Psychotic Disorders: Feasibility and Acceptability Study

Mari Skoge[1,2], ClinPsy; Sofie Ragnhild Aminoff[1], ClinPsy, PhD; Elizabeth Ann Barrett[1], ClinPsy, PhD; Gina Engen Bryhni[3], MSN; Kristine Kling[1], MA; Kari Jorunn Kværner[1,4], MD, PhD; Ingrid Melle[2], MD, PhD; Erlend Mork[1], ClinPsy, PhD; Carmen Simonsen[1], ClinPsy, PhD; Linn Nathalie Støme[1], MS, PhD; Josina Vink[5], MA, PhD; Tor Gunnar Værnes[1], ClinPsy, PhD; Kristin Lie Romm[1,2], MD, PhD

[1]Early Intervention in Psychosis Advisory Unit for South East Norway, Division of Mental Health and Addiction, Oslo University Hospital, Sognsvannsveien 21, Bygg 12, 2. etg, Oslo, Norway

[2]Faculty of Medicine, Institute of Clinical Medicine, University of Oslo, Oslo, Norway

[3]Early Intervention in Psychosis Unit, Sanderud, Division of Psychiatric Health Services, Innlandet Hospital Trust, Ottestad, Norway

[4]Department of Strategy and Entrepreneurship, BI Norwegian Business School, Oslo, Norway

[5]Institute of Design, Oslo School of Architecture and Design, Oslo, Norway

Corresponding Author:
Mari Skoge, ClinPsy
Early Intervention in Psychosis Advisory Unit for South East Norway, Division of Mental Health and Addiction, Oslo University Hospital, Sognsvannsveien 21, Bygg 12, 2. etg, Oslo, Norway

Abstract

Background: Strengthening shared decision-making in mental health care may improve the quality of services and treatment outcomes, but its implementation in services for severe mental disorders is currently lacking.

Objective: This study aims to explore the feasibility and acceptability of iTandem (University of Oslo), a mobile app designed to promote shared decision-making in the treatment of psychotic disorders. In addition, the study aims to investigate mechanisms that potentially contribute to the intended effect of the app. iTandem is a therapy supplement that facilitates patient involvement in decisions regarding treatment goals and focus areas. It is designed for personalized use and contains 8 optional modules: sleep, medication, recovery, mood, psychosis, activity, substance use, and feedback concerning therapy.

Methods: Patients undergoing assessment or treatment for psychotic disorders and their clinicians were recruited for the study. Patients and clinicians jointly used iTandem as part of standard treatment in a 6-week trial. We used a mixed-methods study design with a clear emphasis on qualitative methods. Feasibility and acceptability were assessed through descriptive statistics based on preintervention and postintervention questionnaires and app usage data, in addition to text responses to open-ended items. We conducted a reflexive thematic analysis of postintervention interviews to elaborate these measures and to explore mechanisms potentially contributing to achieving shared decision-making when using iTandem.

Results: A total of 9 patients and 8 clinicians completed the trial. The participants evaluated iTandem as a user-friendly and acceptable tool, but there were considerable variations in how the app was integrated into treatment and in perceptions of its clinical value. The thematic analysis suggests that iTandem has the potential to facilitate shared decision-making through supporting cognition and shifting the patient's role. We also identified scaffolding structures, an analogy of personalized support, as a precondition for these mechanisms and for the overall feasibility and acceptability of iTandem.

Conclusions: iTandem was generally perceived as a feasible and acceptable tool in the treatment of patients with psychotic disorders. Our findings suggest that nonclinical aspects, such as support structures, are important to the feasibility and acceptability of such digital interventions and patients' aptness for digitalized treatment in general. Future research should explore related nonclinical aspects further instead of defining potential target groups based on diagnoses and symptom severity alone.

(JMIR Hum Factors 2025;12:e68813) doi:10.2196/68813

KEYWORDS

shared decision-making; patient involvement; psychotic disorders; mobile apps; mHealth

XSL•FO
RenderX

Introduction

Background

The clinical integration of technology into health care services has been suggested as a strategy to enhance both treatment efficiency and service user satisfaction [1,2]. Moreover, shared decision-making (SDM) has become a core principle in the design of new person-centered services within mental health care [3,4].

SDM involves mutual involvement and collaboration between the patient and service provider, enabling a shared understanding to explore and decide on therapeutic alternatives throughout the course of treatment, in accordance with patient needs and preferences [5,6]. To achieve a shared understanding of the clinical condition, as well as relevant factors that should be included in the decision-making process, delivering information in a tailored and accessible format is crucial [7,8]. Although the body of available literature is inconclusive, studies indicate that higher levels of patient involvement in mental health care services result in improved treatment outcomes related to both clinical recovery and patient satisfaction [9,10]. Thus, mental health care services are making efforts to develop and distribute publicly available communication tools that promote SDM implementation in routine health care, both in Norway [11,12] and internationally [4].

However, the process is slow, and mental health care services have so far implemented SDM only to a limited extent [13,14]. The research literature indicates that SDM is especially lacking in the treatment of severe conditions, such as psychotic disorders [15,16]. According to a Norwegian study, patients with psychotic disorders report that they are not involved in their treatment as much as they prefer [17]. These findings are mirrored in the international literature [16,18,19]. Nevertheless, a recent study found that receiving information and feeling understood are fundamental care needs of individuals who have experienced a psychotic episode for the first time [20]. Both of these aspects are core elements of SDM and underline the potential clinical value of implementing SDM in this patient population.

There are several obstacles to the implementation of SDM in the treatment of patients with psychotic disorders. Many patients endure cognitive and negative symptoms, which comprise problems such as reduced memory, levels of engagement and interest, and difficulties expressing emotions and needs [21-23]. The presence of such symptoms may impede patients from pursuing a role as active agents who take part in informed dialogues with clinicians. In addition, research has revealed stigmatizing views toward persons with psychotic disorders among mental health service providers [24] and that clinicians expect that patients with psychotic disorders are less able and less willing to engage in a collaborative approach to treatment [25,26]. A possible consequence is reduced effort from the service provider to achieve SDM. Another consequence is that patients perceive the stigma from service providers and feel disempowered [27], which further complicates their opportunities to experience SDM. Furthermore, the services delivered to people with severe mental disorders are complex and resource-intensive activities [28,29]. Patients are typically followed for longer periods of time, and services should be able to adjust the goals and content of the treatment according to changing needs. Without strategies to maintain a focus on common goals, this might be challenged in a protracted course with changing clinicians, departments, and patient contexts [30-33].

The discrepancy between the ideal of SDM-based clinical practices and the current situation in severe mental health care exposes a need to develop novel strategies to enhance patient involvement. Integrating digital tools may offer a flexible and personalized approach to targeting SDM [34]. Highly accessible digital tools, such as mobile apps, offer functionality that can overcome some of the obstacles to SDM. Apps can facilitate the collection of real-time data points, providing rich information about specific aspects of the user's life, which can be of importance for the process of recovery. In addition, this functionality can reduce the demands on memory and help users convey relevant information during therapy sessions [35,36]. Several app interventions developed for patients with psychotic disorders have been evaluated as feasible and acceptable by users, such as Actissist (University of Manchester), FOCUS (University of Washington Psychiatry & Behavioural Sciences), and PRIME (Personalized Real-time Intervention for Motivation Enhancement; Regents of the University of Minnesota) [37-40]. However, few mobile apps directly target SDM in mental health [34]. The existing SDM apps offer functionality such as mapping clinical decisions and weighing the benefits and disadvantages connected to them [41] or prompting evaluations of therapy sessions and monitoring of health parameters, which can be used to support decision-making [42,43]. Despite the relevant functionality and promising user experiences, adoption and diffusion of such apps remain low in routine clinical settings [44,45]. A user-centered approach that involves service providers and patients from early exploratory stages is necessary to ensure feasibility and clinical benefit, which facilitates diffusion in everyday clinical settings [46].

The idea of the mobile app examined in this study developed from a previous innovation project in which persons with lived experiences stated that they missed alternative ways of collaborating with services and that there is a lack of tools to strengthen the focus on treatment goals during and between treatment sessions [47]. This insight sparked an iterative and person-centered process of concept and software development, involving a thorough mapping of the needs of patients with severe mental illness and their therapists. The outcome was iTandem, a mobile app designed to enhance the patient's voice in decisions concerning focus areas and treatment goals. Before planning the present exploration of the feasibility and acceptability of iTandem in clinical settings, the app was beta-tested, and its functionality was refined according to feedback from patients and clinicians.

Study Aim

We aim to explore the feasibility and acceptability of the mobile app iTandem [47] in the treatment of patients with psychotic disorders from the perspectives of both clinicians and patients.

XSL•FO
RenderX

In addition, we aim to explore mechanisms that potentially contribute to achieving SDM when using iTandem.

Methods

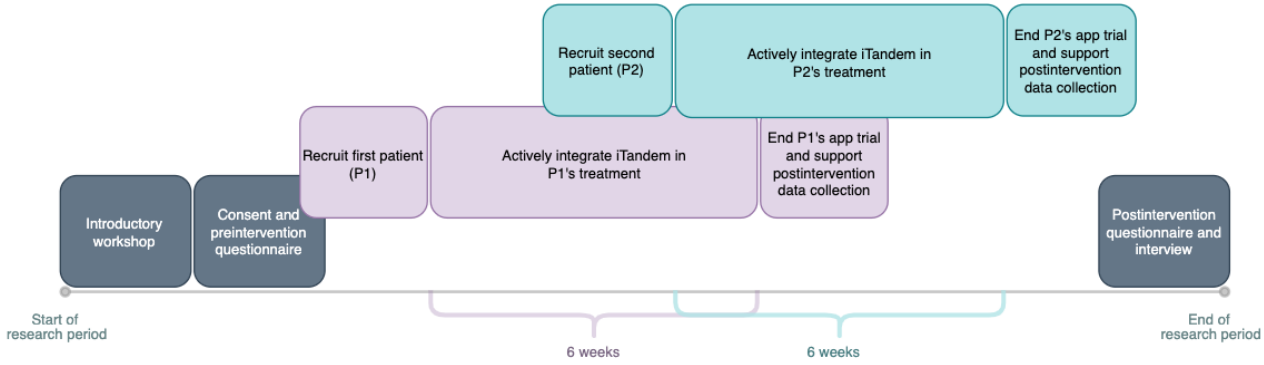
Setting and Recruitment

The study took place at 2 clinical sites at Innlandet Hospital Trust in Norway from December 2022 to May 2023. Innlandet Hospital Trust is part of the public specialist health care services and offers both inpatient and outpatient treatment. The catchment area is geographically large and mainly rural. The hospital trust serves people from a broad range of social strata, and many patients live far away from the specialized services.

One of the sites is a specialized unit for the assessment and treatment of early psychosis. The other site is a psychosis treatment clinic at a district psychiatric center. All the clinicians who participated in the study were invited by the heads of their respective units.

The core project group (consisting of authors MS, SRA, KLR, and GEB) arranged an introductory workshop for clinicians, and the overall aims of the study were explained. The research activities and a timeline of the project are described in Figure 1. Demonstration videos covering the different functionalities of iTandem and providing examples of how to inform and recruit patients were presented. Based on previous user testing, clinicians were advised to guide patients to reduce the selection of modules to 2 - 3 to avoid overload. Written instructions for recruitment and how to access a visual report of app data were handed out, including references to the project website [48]. Guided role-play exercises were offered to the clinicians. However, the role-play training was not carried out because the clinicians did not perceive it to be necessary. Clinicians were encouraged to contact the project coordinators (MS and GEB) when they needed support. GEB was a local coordinator.

Figure 1. Research activities and project timeline for clinicians participating in the study.



Initially, 10 clinicians agreed to take part in the study. They participated in an introductory workshop and signed an informed written consent form. Eight completed the trial, while the remaining 2 clinicians left the study due to lack of engagement in patient recruitment and personal reasons, respectively.

Patients were invited by their clinicians to participate in the study during routine consultations. The inclusion criteria for patients were as follows: currently undergoing assessment or treatment for a psychotic disorder, age 18 years and older, having access to a smartphone, being able to download the iTandem app, and being willing and able to consent. The first inclusion criterion, concerning the presence of a psychotic disorder, was based on clinician ratings. Experienced clinicians working at the clinical sites of this study assessed and diagnosed patients using the *International Classification of Diseases, 10th Revision* diagnostic criteria for schizophrenia, schizotypal, and delusional disorders (F20-F29). The 2 clinical sites offer services exclusively to patients diagnosed with a psychotic disorder and patients undergoing assessment based on indications of a psychotic disorder. Therefore, all patients included in the study met this criterion.

The convenience sample of patients first comprised 14 patients from both outpatient and inpatient units. During the trial, 5 left the study. Two patients did not complete the trial because their therapists were absent from the clinic for much of the trial period due to vacation or sick leave. One patient agreed to participate but never began using the app actively. One patient was discharged and left the study, while another patient did not attend the sessions within the trial period.

Six patients and all 8 clinicians agreed to participate in interviews. A total of 2 patients and 2 clinicians were interviewed in person, 2 patients were interviewed over the phone, and the remaining participants were interviewed through a secure video platform. The interviews were performed by MS and KLR.

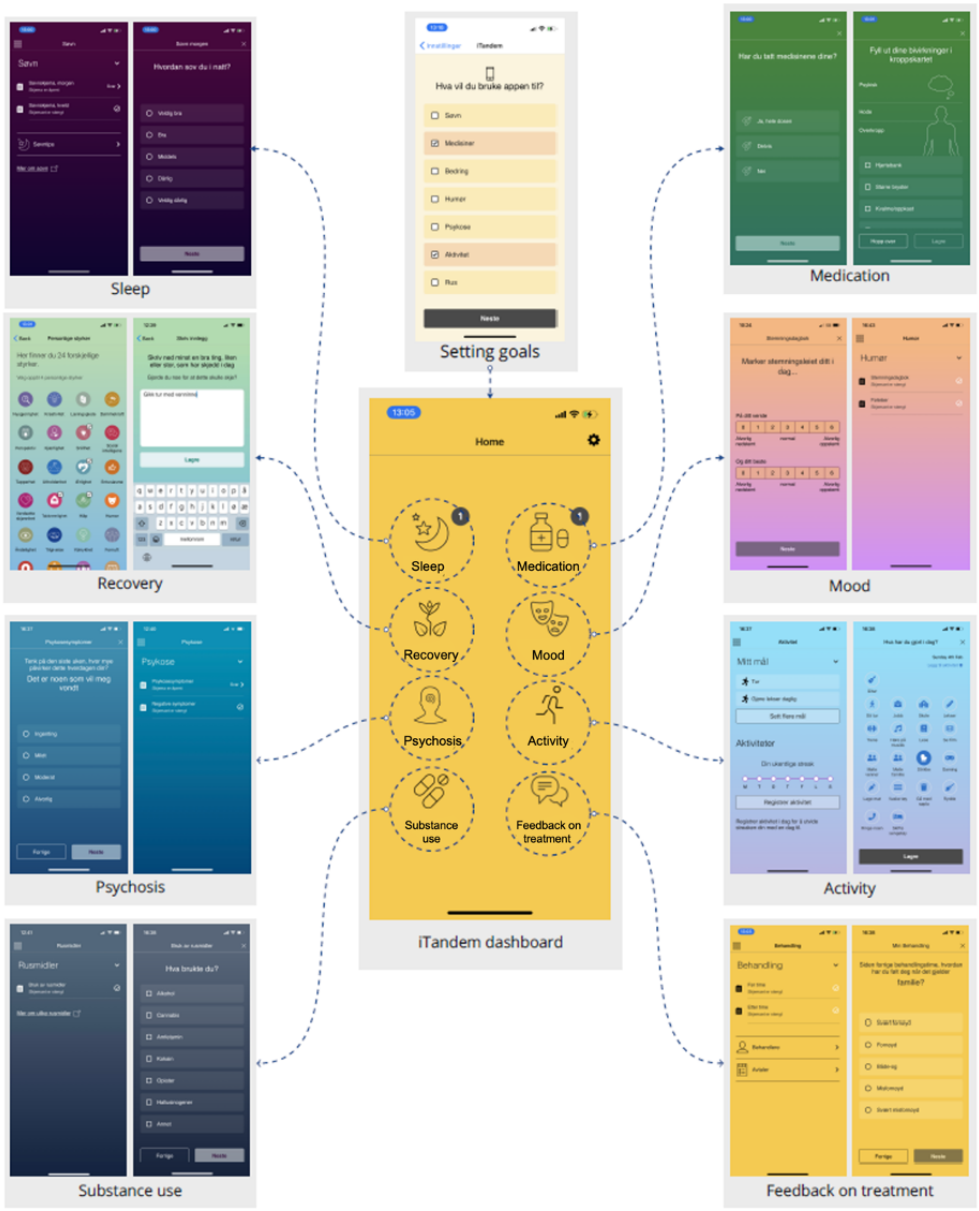
Intervention

The iTandem mobile app [47] is developed by the Early Intervention in Psychosis Advisory Unit for Southeast Norway and the app development team at the Information Technology Department at the University of Oslo. iTandem is designed as a supplement to typical treatment. The purpose of integrating the app in treatment is to enhance therapeutic communication and accentuate the needs and preferences of patients in treatment for severe mental disorders or addiction. Since iTandem is designed for personalized and need-based use, there are no generic recommendations concerning the ideal combination of modules or how long the app should be an integrated part of treatment. However, the trial period of this study was set to 6 weeks.

XSL•FO
RenderX

iTandem is administrated by patients, while their clinicians have access to a digital report that summarizes the app data. iTandem contains 8 optional modules assessing sleep, medication, recovery, mood, psychosis, activity, use of substances, and feedback concerning treatment (Figure 2). The contents of the different modules are described in detail elsewhere [47]. The app prompts patients to register data regularly. The time intervals between notifications depend on which modules the user has activated. For instance, the sleep module entails a sleep diary that requires data registrations in the morning and in the evening. When this module is activated, the user receives notifications twice a day. In contrast, the psychosis module sends out weekly prompts, as the symptoms tracked with this module change more slowly and over a longer period of time.

Figure 2. The dashboard and modules of the iTandem app.



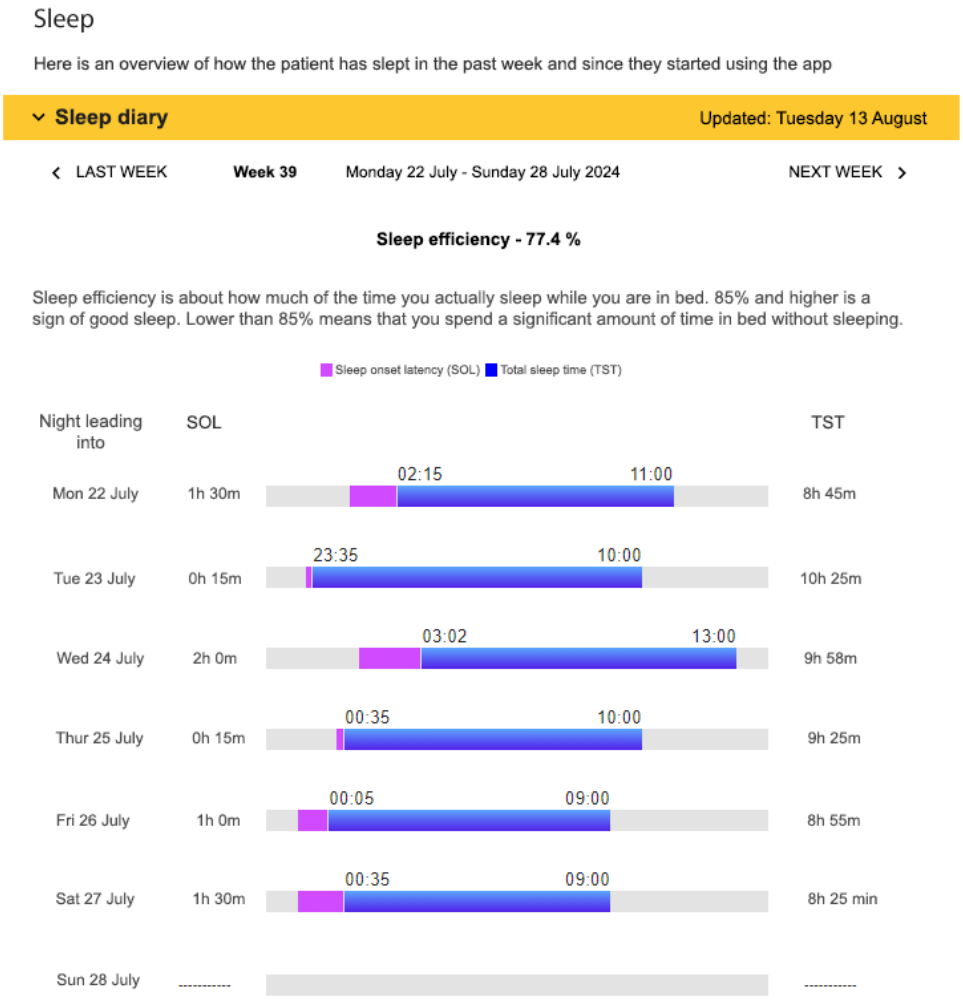
In this study, patients consented to participate in the 6-week trial via a web-based form with guidance from their therapists. Thereafter, patients downloaded the iTandem app on their personal smartphones and discussed with their therapists which of the modules would be most meaningful to their condition and current life situation. This conversation represents the beginning of an SDM-based approach, directing the focus of ongoing treatment in accordance with patient needs and preferences. Patients and clinicians were encouraged to change the selection of modules during the six-week trial period if they experienced that a treatment goal was achieved or perceived other modules as more useful.

The data registered by patients during the trial period were transferred directly to Tjenester for Sensitive Data, a service for storing and processing sensitive data at the University of Oslo. These data comprised the basis for an automatically generated digital report with visual representations of the active modules. Clinicians accessed the report through a secure web

XSL•FO
RenderX

portal. Figure 3 demonstrates how the sleep module is represented in the report, summarizing a week of data registration.

Figure 3. The sleep module is visualized in the report. The content has been translated into English.

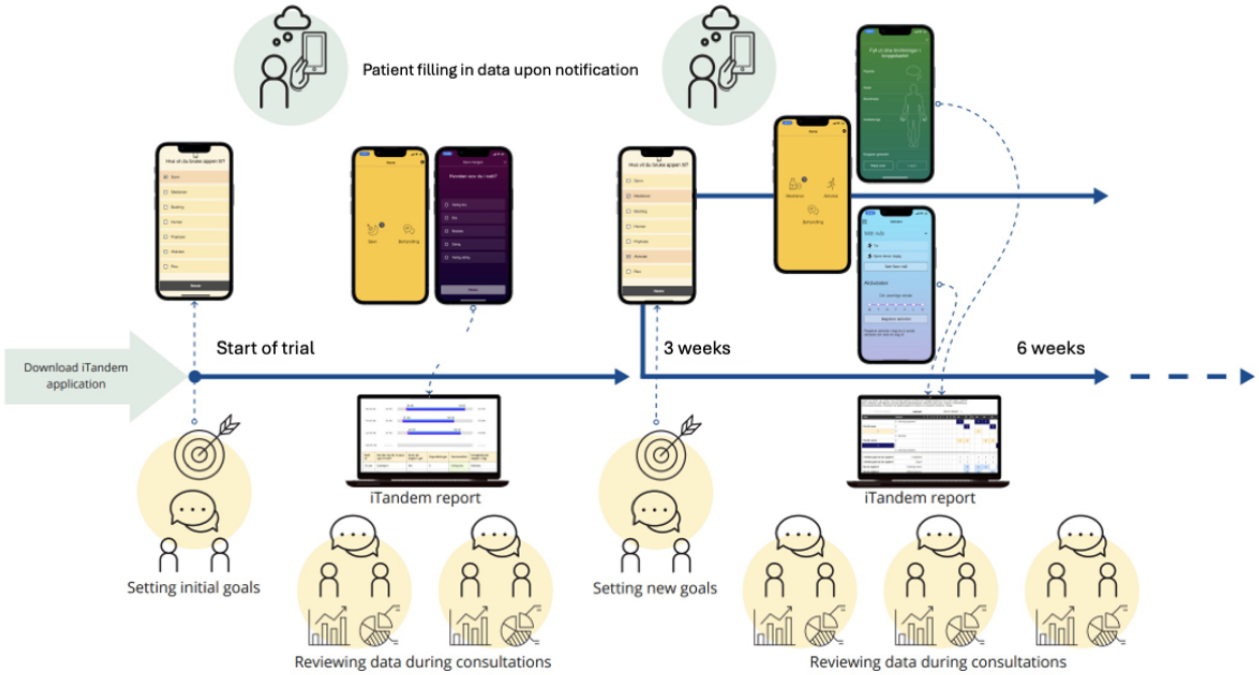


Night leading into	Did you take anything to help you fall asleep?	Daytime naps	Awakenings during sleep	Sleep quality	Daytime functioning
22.July	No	No	1	Medium	Medium
23.July	Sleep medication	No	0	Good	Good
24.July	No	No	2	Bad	Bad
25.July	Sleep medication	Yes, 20-60 min.	0	Medium	Good
26.July	No	Yes, < 20 min.	0	Very good	Very good
27.July	Sleep medication	No	0	Very good	Very good
28.July					

In the introductory package, the clinicians were instructed to share and discuss the report with their patients during the treatment sessions of the trial period. The project group presented various potential areas of application, such as monitoring fluctuations of symptoms and side effects from medication, identifying connections between different aspects of the patient's condition and daily life, and guiding ongoing clinical interventions in accordance with reliable patient data over time. Figure 4 demonstrates the process of integrating and actively using the app and report during the trial.

XSL•FO
RenderX

Figure 4. Flexible integration of iTandem throughout the trial period.



Study Design

A mixed-methods research design placed within a pragmatic paradigm [49] was applied to explore iTandem's feasibility and acceptability and the mechanisms potentially involved in achieving SDM. The study places a heavier emphasis on qualitative methods, prioritizing active researcher interpretation and broad exploration of individual experiences from a small sample of participants. The quantitative methods of the study comprise some instruments providing reliable data that can easily be compared to benchmark values. Other measures that are typically operationalized in a largely quantitative manner within similar research contexts involve more qualitative assessments in this study.

The research design applied here largely resembles a convergent design with a side-by-side comparison approach [49], since the data collection and analyses are mainly conducted separately before integrating the findings in the discussion, where the qualitative elements elaborate and nuance the quantitative results. However, the quantitative data are also used to guide the qualitative methods earlier in the analytic process.

Quantitative Data and Methods

The quantitative data include responses to the web-based preintervention and postintervention questionnaires and app usage data. All 9 participants and 8 clinicians completed the preintervention and postintervention assessments, except for one clinician who did not fill out the postintervention questionnaire. Descriptive analyses were conducted using SPSS Statistics 29 to depict the groups of participants, identify tendencies in the data material, and create a foundation for the qualitative exploration of feasibility and acceptability. The following quantitative measures were included in the study:

Demographics and Participant Background Information

In preintervention questionnaires, patients reported gender, age, treatment level (inpatient or outpatient), and duration of treatment. Clinicians reported gender, age, which service they worked in (inpatient, outpatient, or both), and patient diagnoses.

Usability

User-friendliness (ie, technical feasibility) was measured using the System Usability Scale (SUS) [50]. Separate items constructed by the project group mapped occurrences of technical errors.

Acceptability

Acceptability was assessed with postintervention closed-ended questionnaire items. Both groups were asked whether they experienced the intervention as clinically valuable and to what extent the digital tool had been used during the trial period. In addition, patients reported whether the intervention was experienced as comfortable and safe and which modules they had chosen. App usage data supplemented the patient acceptability measure with indications of whether the participants were engaging actively in iTandem or not. These data were obtained manually by observing and documenting the number of data points registered by each of the participants in the different modules, providing insight into broad tendencies of app activity across the sample.

Qualitative Data and Methods

The main data source in the current study was the qualitative data material. It comprised both text responses to open-ended items in the preintervention and postintervention questionnaires and individual semistructured postintervention interviews. All 8 clinicians and 6 of the patients participated in interviews. The number of words in the 14 interview transcripts ranged from 2304 to 6574, with a median of 3603 words.

XSL•FO
RenderX

A reflexive thematic analysis was conducted using NVivo 14 (Lumivero). The purpose was to gain a deeper understanding of the feasibility and acceptability measures and to explore these data further, identifying mechanisms that potentially contribute to the intended effect of the iTandem intervention:

Feasibility

Feasibility was explored in postintervention interviews. Both groups were asked similar questions about how the iTandem intervention fit into the current clinical setting, what factors may impact a successful implementation, and their thoughts on the future use of apps in long-term follow-up. All participants were asked about potential areas of improvement related to the tool and its implementation. The patients' postintervention questionnaires included text response items mapping experiences with the specific modules they had used, which were used as an indication of whether the different modules had been applied as intended.

Acceptability

The postintervention interviews provided oral accounts of the participants' enthusiasm, satisfaction, potential negative events, and experience value from taking part in the trial. The patients' open-ended questionnaire items prompted the patients to share perceived health effects, as well as positive and negative aspects of the intervention.

The thematic analysis of this study is largely based on Braun and Clarke's reflexive thematic analysis, applying essentialist and experiential theoretical assumptions [51]. The data collection and analysis of the study were guided by the overall aims of the research project, broadly exploring the feasibility and acceptability of iTandem, as well as mechanisms related to SDM. However, the thematic analysis applied here is essentially inductive rather than theory-driven.

MS conducted the analysis. KLR and SRA contributed in stages 1, 4, and 5. The authors involved in the analysis had several explicit discussions concerning aspects that may have had an impact on the methodological choices and results of the study, such as roles, backgrounds, and motivations in the project. KLR and SRA were central actors in the development of the app. KLR, SRA, and MS hold relatively optimistic views on the potential value of digitalization and SDM in health care. These contextual and personal factors impact all stages and decisions of the research project. Furthermore, the participants' accounts represent subjective interpretations of experiences within specific environments, which are accepted and included in the analysis. Thus, principles of reflexiveness, attention to context, and an experiential approach to data are integral in this application of thematic analysis [52]. The 6 steps of thematic analysis were conducted in a dynamic manner, moving back and forth between the stages described below, especially from the third to the sixth phases:

Data Familiarization

MS conducted verbatim transcriptions of the 14 interviews. MS, KLR, and SRA read and reread the interviews and noted their first thoughts about each interview and the dataset as a whole. The group shared ideas and reflected on the content of the data material. This workshop, as well as the other collaborative activities described in later phases, is not a standard element in Braun and Clarke's reflexive approach. The purpose of the discussions was not to get closer to objectiveness, as this is not considered an ideal in the reflexive form of thematic analysis [51]. Instead, this was done to spark creativity and reflexiveness in the analytic process. The transcripts were imported to NVivo. Electronic memos based on the initial thoughts about the data material were created and connected to each interview.

Coding

MS coded the dataset systematically in NVivo. The coding was conducted in accordance with the guidelines for reflexive thematic analysis, where one subjective coder is sufficient to achieve high-quality output, while aiming for interrater reliability using a second rater would represent a deviation from this branch of thematic analysis [51]. No coding frameworks were applied, and no parts of the data material were excluded from the coding process. Although ideas concerning the research questions emerged in these early phases, MS made efforts to direct the same level of attention to all parts of the texts in all interview transcripts to maintain an inductive approach. However, the process can be said to have been influenced by theory to some extent, as the authors had hypotheses related to the research questions and had experience with the same kind of data material from a previous beta test. Furthermore, the interviews were semistructured, which led the accounts in a specific direction, although question formulations were open, and participants were encouraged to communicate their thoughts without restrictions. This phase resulted in a total of 340 semantic codes.

Generating Initial Themes

MS combined different codes into groups that represented units of similar meaning in NVivo. Examples of headings of such code groupings are "Target group" and "Prerequisites of experiencing clinical value." In this phase, the titles of the candidate theme structures were still descriptive and topic-like rather than conveying shared meaning about essential aspects of the data set.

Developing and Reviewing Themes

Maintaining an experiential approach to the data, MS went through the tentative themes and the associated data extracts in NVivo to ensure that the themes reflected what the participants had said. As a result, the themes were merged or reconstructed into new tentative themes. In a new workshop, the groups of codes were presented on paper to KLR and SRA. The authors used this as a starting point to adjust themes and structure them in relation to each other. Some were left out as they were deemed not contributing to the research aims. MS also studied the questionnaire material to evaluate whether the tentative themes reflected the participant perceptions provided outside an interview context. This phase resulted in 3 preliminary themes on core mechanisms potentially involved in facilitating SDM through iTandem.

Refining, Defining, and Naming Themes

The themes were reviewed again by MS, KLR, and SRA, given titles, and connected to the most representative quotes of the data material. The group also worked on descriptions of each

XSL•FO
RenderX

theme that communicated the intended meaning clearly. As the analytic process was carried out in an iterative manner, MS repeated phases 3, 4, and 5. After another round of consulting with the other authors and reviewing and refining the themes, the final set of themes was completed.

Producing the Report

MS initiated the write-up in parallel with previous phases of the analytic procedure, long before themes were finalized. Thus, the data analysis, literature search, and writing processes interacted throughout this research project.

Service User Involvement

Service user representatives were involved in all stages of this study, from early app development to designing the assessment tools for data collection and writing this paper. Early Intervention in Psychosis Advisory Unit for Southeast Norway, at which several of the article authors are employed, has its own board of service user representatives. In addition, the researchers in this project discussed the research design with a user representative employed at Innlandet Hospital Trust. During the writing process, the authors discussed the interpretation of findings with 2 service user representatives from the Bipolar Association Norway and refined the manuscript according to feedback on themes.

Ethical Considerations

The patients participating in this study were recruited based on their status as patients in service for severe mental disorders and can be defined as members of a vulnerable population. Involving vulnerable persons in research requires careful ethical considerations, including the protection of the participants' interests and obtaining freely given informed consent. These principles were integral in the design of the study and in the introductory workshop for the clinicians. During the introductory workshop and the further communication with clinicians, they were instructed to clearly explain both benefits and risks related to participation to their patients, as well as stressing that all elements of the study participation were voluntary and that withdrawing would not result in any negative consequences. All participants gave consent to participate in the study through a digital form after receiving information about the project presented verbally and in a text document. The project obtained approval from the data protection officer at Innlandet Hospital Trust (23839981) and Oslo University Hospital, the institution responsible for data processing (22/20426). Importantly, this study primarily investigates digitalized service delivery and user experiences, rather than mental health conditions and treatment outcomes. The Regional Ethics Committee evaluated the study to be outside of their mandate, as the research did not primarily aim to collect and analyze health data or develop new knowledge about health or disease (523626).

Results

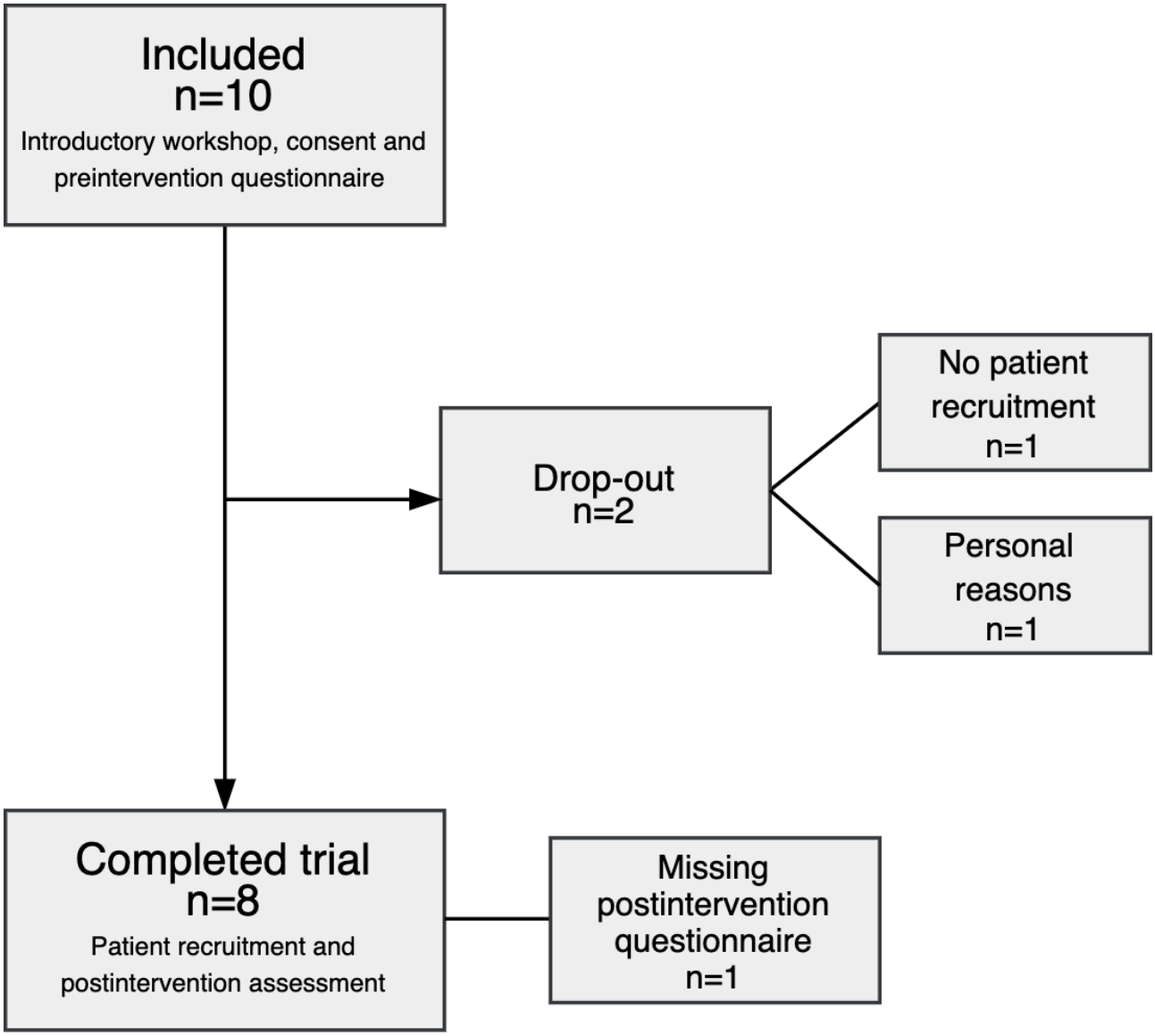
Quantitative Data

Participant Background Information and Demographics

Eight clinicians completed the study (Figure 5). The clinicians reported the following gender identities: 4 males, 4 females, and 0 other identities. The age range was 27 - 63 years old (mean and median 45 y old). A total of 4 clinicians provided outpatient services, 2 inpatient services, while the remaining 2 clinicians worked in both inpatient and outpatient services.

XSL•FO
RenderX

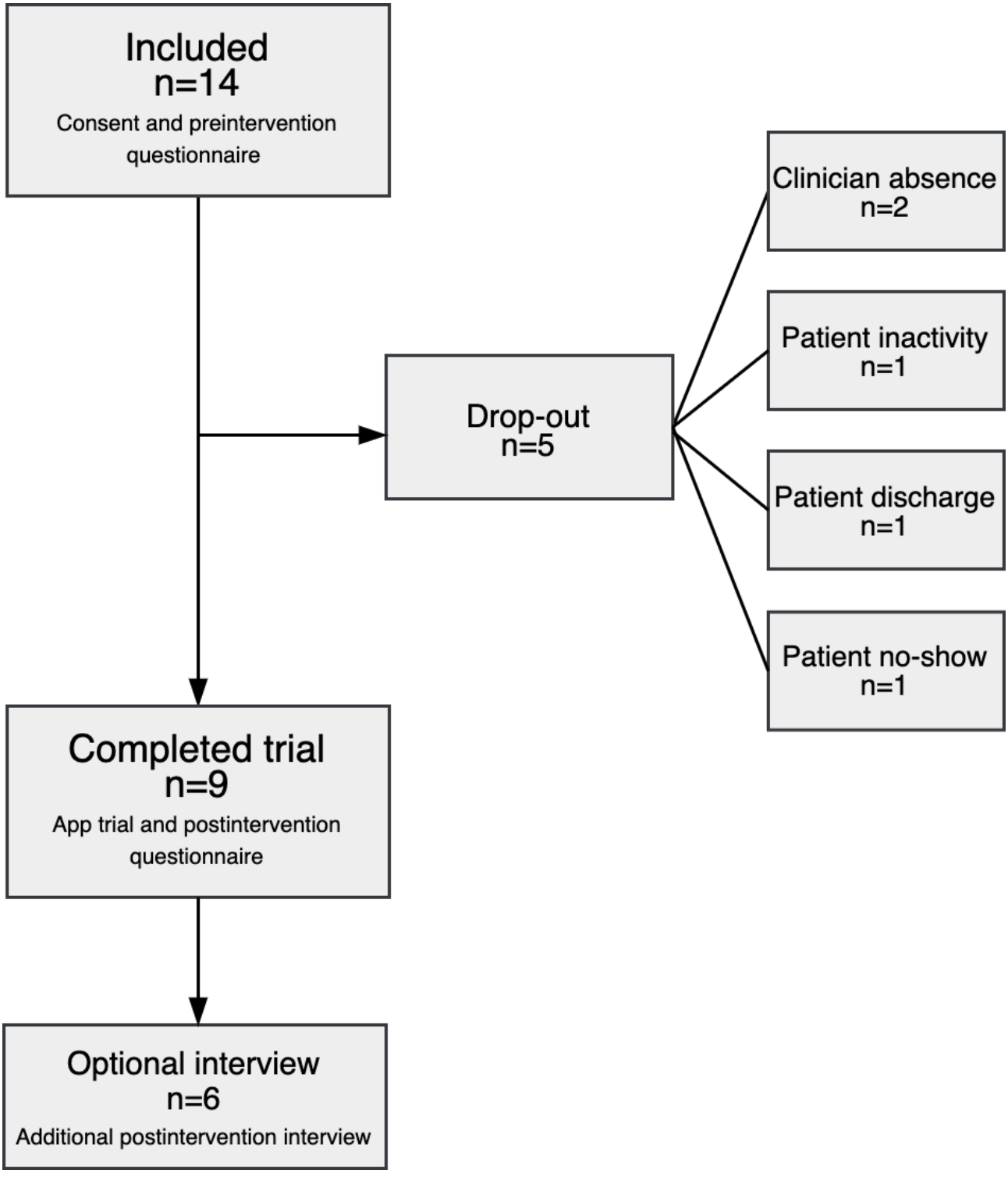
Figure 5. Flowchart of the clinicians recruited to the study.



Nine patients completed the study (Figure 6). The patients reported the following gender identities: 7 females, 2 males, and 0 other identities. The age range was 18 - 43 years old (median 27 y old). Eight out of 9 received outpatient services. The duration of treatment in mental health care services was 0-200 months (median 16 mo). All patients were undergoing assessment or receiving treatment for schizophrenia, schizotypal, or delusional disorders (F20-F29). Two were diagnosed with additional stress reaction disorders and/or a developmental disorder.

XSL•FO
RenderX

Figure 6. Flowchart of the patients recruited to the study.



Usability

iTandem was rated as above average in terms of user-friendliness, with SUS mean scores of 75.0 by patients and 82.1 by clinicians, out of 100.

Additional items covering experiences of technical errors revealed occasions where clinicians were denied access to the report or where the report's representation of one of the app modules contained errors.

Acceptability

The levels of engagement varied across the sample. Most patients engaged with 3 or more modules (Table 1). The introductory package to clinicians recommended limiting the number of modules to one or two, but some patients chose more. Patients displayed a preference for modules related to general health and well-being over symptom-oriented modules. Activity and Sleep were the most popular modules, selected by 8 of the 9 patients.

XSL•FO
RenderX

Table . Joint display of patients' questionnaire responses.

Participant number	Selected modules	Report brought up in sessions	Experience of using report in sessions	Experience of sharing personal data	Perceived health effect
103	Sleep, mood, activity	Once	Useful	Ok	Positive: "I became more aware of my feelings and mood … I realized how much I accomplished to do each day. I used the app less after a while because I managed to maintain a good structure in my everyday life. But I liked it a lot in the beginning."
105	Sleep, mood	Several times	Useful	Ok	Positive: "Provides me with an overview."
107	Sleep, recovery, activity	Once	Ok	Un-comfortable	Negative: "I became a bit more paranoid and felt watched."
108	Sleep, recovery, mood, activity, medication, psychosis	Several times	Useful	Safe	Positive: "Maybe a bit, nice to have a routine with the app."
109	Recovery, activity	Every time	Useful	Ok	Positive: "More aware of activities, that it doesn't have to be that many every day, but that I still manage some things."
112	Sleep, mood, activity	No	No experience	Ok	Positive: "It was nice to get an overview of sleep and feelings, and for instance I became more aware that I often go to bed too late in the night."
113	Sleep, mood, activity, psychosis	Several times	Useful	Safe	Neutral: "No."
114	Sleep, recovery, mood, activity, medication, psychosis	Once	Ok	Safe	Neutral: "I don't know."
115	Sleep, mood, activity	Several times	Ok	Ok	Neutral: "It has not made any big differences to my mental health. It has been the same regardless (…)"

Responses to questionnaire items targeting experienced value from the iTandem intervention revealed distinct differences between patients (Table 1). Five out of 9 patients reported that reviewing the report in sessions felt useful. The same number of patients reported that they experienced positive health effects from using iTandem. Three patients reported neutral experiences with both items. One incident was identified as an adverse event related to the iTandem intervention. One patient participating in the study (#107) experienced increased levels of paranoid thoughts and a feeling of being watched through the digital tool. This patient also reported that sharing data through the report system was experienced as uncomfortable. The patient discussed the problems with their clinician and stopped using iTandem.

Questionnaire data showed that some clinicians were not able to maintain an active focus reviewing the report regularly in treatment sessions (Table 1). Their evaluations of clinical value from the intervention varied. The variation in responses seemed to correspond with the levels of app activity and experienced value from the perspectives of their respective patients.

Qualitative Data

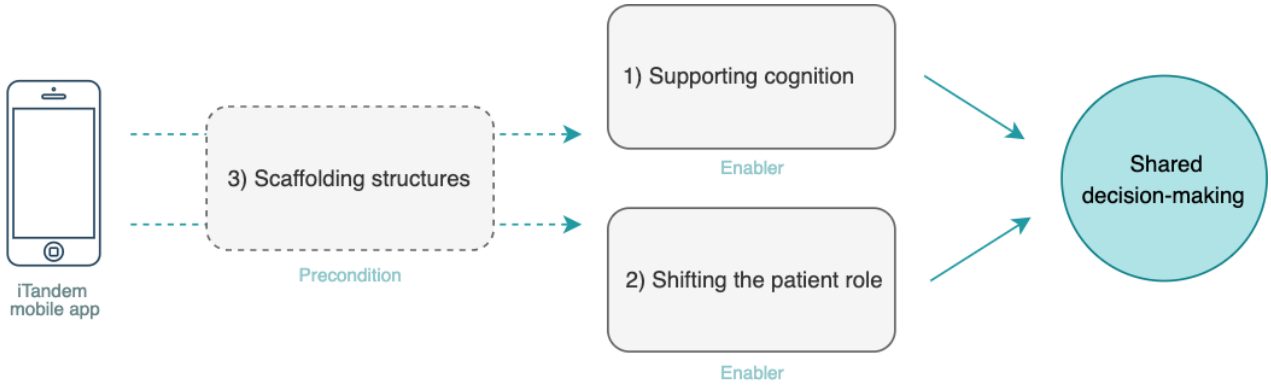
Feasibility and Acceptability

Through the reflexive thematic analysis, the following themes, which represent potential enablers for SDM in the iTandem intervention, were generated: Supporting cognition and Shifting the patient role. The third theme, Scaffolding structures, was

XSL•FO
RenderX

constructed as a precondition for these processes (Figure 7) and suggested as a precondition for the overall feasibility and acceptability of iTandem.

Figure 7. Themes generated through qualitative analysis and their suggested role in promoting SDM using the iTandem app.



Theme 1: Supporting Cognition

The first theme elaborates on how the iTandem intervention served as an aid for cognitive tasks within the treatment context for both patients and clinicians. Participants recognized bias and limitations regarding their own memory, their capacity to communicate inner states, and their ability to follow a structure in both treatment sessions and over a course of treatment. The data registration made by patients in iTandem, together with the clinicians' use of the report, were viewed as compensatory activities. In addition, several participants expressed that the intervention supplied the treatment with new clinically relevant information and that they had the experience of getting an overview of aspects related to their condition and general well-being.

Graphic representations of the patient data were automatically generated in the report, offering a more accessible format for summarizing and communicating complex information. Thus, iTandem helped translate data about the patient into something easily comprehensible and provided the therapy with more concrete reference points. Participants valued that the visual report provided an overview of parameters related to health and general well-being:

> *When patients register their sleep, you access a graph where the information from sleep over time is a lot more comprehensible. It makes it more tangible.* [Clinician #209]

> *It is nice to get to see the patterns, get it documented instead of writing it down or trying to remember it. That does not work.* [Patient #108]

Some participants talked about the value of iTandem as a tool for clinicians to maintain structure and focus on the patient's chosen therapeutic project over time. iTandem was also perceived as an aid to promote a more nuanced presentation of the lives of patients. Moreover, iTandem facilitated the search for positive elements in patients' everyday lives:

> *I used iTandem as a reminder for myself to maintain structure and keep track from one session to the next. I think it worked well for remembering to bring up topics. (…)*

> *Taking in the nuances in everyday life can be a problem for many patients. One event might be experienced as bad, and the patient focuses on this event only. And then, [with iTandem], it might become evident that there had been better moments during the past week.* [Clinician #207]

Theme 2: Shifting the Patient Role

The second theme covers how the iTandem intervention had the potential to cause a push toward a more active and empowered patient role in therapy. The intervention entailed a direct opportunity for patients to take part in decisions regarding their treatment focus. A patient account reveals that this was experienced as unfamiliar, but the alteration of the patient role resulted in a feeling of involvement and empowerment. Furthermore, when maintaining activity in the selected focus areas over time, the patient experienced increased commitment and agency:

> *I felt that I was very involved in deciding what we would focus on. Although it was difficult to express myself firmly, it helped me to think "Okay, this is what I want to focus on, this is what I want." (…) I liked using the app. In a way, I felt more—I don't know the exact word, but—I felt that I had to actively use it to get better. Well, there was no pressure, but I felt that it was my responsibility too.* [Patient #109]

With iTandem, the course of treatment was supplemented with an alternative path for patients to engage in their treatment. Several patients exerted strong digital skills and explored the tool independently. They quickly became comfortable in navigating iTandem, sometimes to a greater degree than the clinicians. Thus, iTandem also represented an area of mastery and a means to show competence and reduce the experience of hierarchies that characterize traditional health care contexts. Furthermore, although most participants across the sample underlined the need for patients and clinicians to know each other well before introducing digital tools in treatment, some talked about technology as a means to relate to and disclose important issues in therapy:

> *They grasp it immediately. They are used to apps. I had briefly mentioned the name of the app before I was going to introduce it in the session, and the*

XSL•FO
RenderX

> *patient had already registered and started exploring the app.* [Clinician #211]
>
> *I think [disclosing feelings through an app] is easier. But it depends on who the person is. If it is a friend, I can speak freely. But I don't feel like I always can do that with a therapist.* [Patient #112]

Perceiving digital communication as an equivalent or superior alternative to the traditional therapeutic conversation might break with assumptions regarding what high-quality and safe therapeutic communication is for patients in mental health care. Thus, it demonstrates a change in the views of different patient roles and underlines the importance of mapping needs, preferences, and personal traits to create beneficial conditions for secure therapeutic relationships and hearing patient voices.

Theme 3. Scaffolding Structures

The third theme refers to support structures that acted as prerequisites for the 2 processes described above and for the participants' evaluation of iTandem as feasible and acceptable. All the participants talked about social and/or technological structures that had significantly impacted their experiences with iTandem throughout the trial. Examples of scaffolding included the digital infrastructure enabling the operation of iTandem, encouragement to maintain app activity during the trial, and pragmatic strategies to compensate for reduced levels of functioning caused by invasive symptoms.

For some, scaffolding was necessary from the initial use of iTandem. One patient endured paranoid thoughts directed toward digital devices, which made it impossible to use a personal smartphone. To solve this, a relative lent their phone to the patient. The patient expressed positive experiences with the intervention, even though the intensity of paranoid thoughts persisted throughout the trial. Other patients struggled with understanding text and symbols in iTandem and needed guidance from close ones to interpret the content:

> *I am positive [toward using the app]. I have not installed it on my phone, I have used [a relative's] phone. I might have to disclose that first. I have a problem with apps on my own phone.* [Patient #108]
>
> *I needed to ask my partner at times because I have difficulties understanding the meaning of the tasks.* [Patient #114]

Moreover, the act of clearly conveying the purpose of the intervention was a central scaffolding prerequisite for both groups. When there was a lack of communication about purpose and use, patients did not necessarily perceive the intervention as helpful, and the clinician was not using the tool to focus the treatment.

Scaffolding structures, such as a safe therapeutic alliance and an attentive clinician, seemed central to the process of enabling a shift in the patient role. Work within the recovery app module provided an example of how scaffolding acted as a precondition for this relational dynamic. The recovery module prompted patients to register positive everyday happenings and personal strengths. These activities appeared to represent a highly relevant focus area in the treatment of several patients and incited encouraging experiences for some. However, without scaffolding, such as the presence of reassurance and guidance from a clinician, the app intervention was not feasible and did not result in experiences of SDM-related processes. The differences across the sample are exemplified by the following quotes:

> *I would say that [the recovery module] was the most positive aspect. I got a new perspective on which strengths I actually possess. Especially for someone who doesn't think they have any…I actually have some.* [Patient #109]
>
> *With "good experiences" and "personal strengths," it's all about positive things concerning yourself, you know. You are meant to come up with positive things about yourself, and that is not something I do.s So that part is difficult for me.* [Patient #115]

Discussion

Principal Findings

This study aimed to explore iTandem's feasibility and acceptability and the mechanisms potentially contributing to SDM in treatment for psychotic disorders based on a 6-week app trial and data from questionnaires and interviews. The findings reveal largely positive evaluations of user-friendliness and acceptability. SUS scores from both patients and clinicians place iTandem's usability higher than the benchmark mean score of digital health apps [53]. However, there were substantial differences in how much the app was used and in perceived clinical value, including its contribution to SDM. These variations were explored further through the qualitative analysis. The qualitative findings suggest that the app has the potential to initiate cognitive and relational mechanisms that are central to enabling SDM in therapy. These processes, as well as evaluations of iTandem as feasible and acceptable, often relied on personalized support over time or "scaffolding."

Our results suggest that a digital tool like iTandem can offer an alternative channel to communicate inner states, as well as a platform where the information is made accessible for both the patient and clinician. Thus, iTandem facilitates a common recognition of clinically relevant issues incited by the patient. The findings fit well with existing literature underlining how cognitively demanding SDM processes can be for patients [54] and highlighting the importance of building a shared understanding to succeed in implementing SDM [7]. Importantly, the provision of information alone is insufficient as a means of driving SDM. In the absence of cognitive aids that help patients apply complex information to their own situation, information does not necessarily contribute to moving the patient toward experiencing more empowerment [55]. Existing literature has pointed out the value of multidimensional information tools in facilitating both the involvement and apprehension of health information in severe mental health contexts [56]. In accordance with this literature, the participants in our trial expressed appreciation for the tables, graphs, and other visual cues of the iTandem intervention. Nevertheless, research on digital SDM tools for mental health care [34,57] has little focus on cognitive mechanisms and the formats of information from the perspectives of both clinicians and patients.

XSL•FO
RenderX

An enhanced focus on tools that aim to compensate for limited cognitive capacity may be useful in the further development of digital interventions aimed at strengthening SDM. This may be of even greater relevance to patient populations that struggle with cognitive and negative symptoms [21-23] and their clinicians.

The integration of iTandem caused a movement in the patient role toward increased agency and involvement. The change was facilitated by the introduction of a digital patient-administered approach to relating to and actively impacting the direction of therapy. As the iTandem app entails personalization at the operational level, it is simple for users to tailor the use of the app based on individual preferences [26]. Interestingly, the results revealed a tendency toward preferring modules related to general health and well-being over modules dealing with illness-related symptoms or drug use. This tendency was found in both the quantitative and qualitative data material, and although it is not statistically valid, it may represent the relevance of a holistic and resource-focused approach to treatment. This is in line with a desire for more recovery-oriented mental health care services [58], where we move our clinical focus from only monitoring a reduction in severity of symptoms toward more person-centered treatment goals reflecting what the patients find important for their well-being [59,60]. Thus, efforts to strengthen SDM with flexible and person-centered methods harmonize with the recovery movement of mental health care [9,57].

The patient role shift also accentuates the complex relational processes underlying SDM. Previous research has conceptualized SDM as a method of care [5,6]. This perspective underlines that SDM should be approached as an intrinsic value in therapeutic settings, not a practical procedure. Operationalizing SDM by using digital tools such as iTandem can help us overcome some of the barriers we have been facing until now, especially in vulnerable populations. However, patients hold different preferences regarding roles in the therapeutic relationship, the desire for control in health care settings, and the degree to which they would like to involve themselves in decisions about their own condition [13,61].

The participants' accounts indicated that various social and technological support structures acted as prerequisites for experiencing the 2 processes discussed above and for the overall feasibility and acceptability of iTandem. There exists evidence that speaks for the value of several types of support in severe mental health care, such as the involvement of relatives in treatment [62,63] and social support in more general terms [64-66]. This literature underlines the general relevance of scaffolding for patients with severe mental health conditions, which proved to be a central concept in our findings. In addition, the importance of supportive social dynamics has been recognized in models from implementation science [67-69]. However, the emphasis in the implementation literature is often on more formal and impersonal support structures, such as the availability of resources and organizational backing. The concept of scaffolding structures in this study is more sensitive to subtle individual differences in social and practical needs, rooted in a person-centered approach.

Importantly, there are prerequisites that are more fundamental than the scaffolding structures discussed above. The target population of iTandem is characterized by a lower level of digital competency and financial status than the general population [70,71]. Although previous research has shown high rates of smartphone ownership within the population [72], there still exists a minority that lacks the required financial means and digital skills, for whom interventions like iTandem remain inaccessible. The presence of paranoid thoughts directed toward technology has been identified as another barrier [72]. Although this barrier does not apply to most patients, the one adverse event in our trial demonstrates the relevancy of recognizing this subgroup. In our trial, the therapist reported the adverse event after it was successfully handled in the treatment sessions and the symptom level had been normalized. The clinician also informed the project group that the therapeutic alliance had maintained cooperation despite the emergence of paranoid thoughts directed toward iTandem. Therefore, we perceived the event's level of seriousness as moderate and the negative reaction as reversible, and no interventions were initiated. Nevertheless, such adverse incidents expose a need for strategies to uncover and handle negative participant reactions. Guidelines addressing this are currently emerging [73], underscoring the importance of improving future practice.

In the digital mental health field, conditions such as depression and anxiety have been the main areas of focus [74,75]. Our study suggests that digital tools also have the potential to compensate for unmet needs associated with patients with psychotic disorders, such as cognitive and negative symptoms and a pronounced lack of participation in clinical decision-making. Thus, mental health care services should also offer such tools to this patient population. Existing research on interventions that share some of iTandem's functionality supports this notion [34,76]. For instance, the computer-based decision-making tool TReatment E-AssisT (TREAT; RoQua Inc.) improved clinical discussions by uncovering more clinically relevant topics that aligned with the needs of patients with psychotic disorders [77]. However, the tool did not significantly affect the levels of SDM [78]. This optimistic but inconclusive evidence of digital tools and SDM [34,76] encourages further research on the mechanisms involved. Our study also invites further research on individual differences that are not related to psychopathology, such as the preferred degree of control and decisional impact, tech-savviness, and the digital habits of patients. In addition, we recommend exploring clinicians' perspectives on SDM and digitalization. Finally, research investigating the feasibility of these concepts at service, economic, and societal levels is needed. We will address some of these knowledge gaps in an ongoing study that includes iTandem and uses a modified study design and new outcome measures.

Limitations

This is a small study, and the limited number of participants makes it impossible to draw firm conclusions. In addition, technical problems with the report that occurred during the study represent a limitation, as there were cases where interruptions obstructed participants from carrying through the intervention as planned. Another limitation is the possible weaknesses in the

XSL•FO
RenderX

introductory workshop. We did not insist on doing role-play to make clinicians more familiar with how it should be introduced when they were reluctant. We believe this could have helped them convey the essential recommendations regarding the need to limit the number of app modules employed by patients. There are also ethical concerns related to the recruitment procedure. Since clinicians recruited their own patients, there is a risk that patients felt obligated to participate because of the pre-established relationship. Importantly, during the introductory workshop for clinicians, the project group communicated the importance of fully informed consent and the option to permanently leave the trial for all participants. The clinicians were also instructed to underscore the voluntary nature of participation when recruiting patients to the study. Nevertheless, research staff will be responsible for patient recruitment in the new trial. Moreover, the study is based on the conception that SDM and clinical app implementation are generally wanted in health care services, and a final limitation is that this stance is not challenged. The mixed methods design is, however, a strength of this study. A pragmatic combination of quantitative and qualitative methods provided the study with a comprehensive and nuanced understanding of the participants' experiences with iTandem and SDM. The qualitative emphasis allowed us to collect rich data, although the sample was small. Participants perceived the threshold to respond to questionnaires as lower than that of being interviewed; thus, we were able to collect information, even from those who were reluctant to speak with us. Another strength is the naturalistic setting and the inclusion of user perspectives of both patients and clinicians during development and testing.

Conclusions

Patients with psychotic disorders and their clinicians generally perceived iTandem as a feasible and acceptable supplementary tool for treatment. The app has the potential to serve as an aid for cognition and to facilitate a shift of the patient role toward a more active position in clinical settings. The study suggests that the feasibility and acceptability of iTandem depend on some fundamental prerequisites, including appropriate support structures. We recommend not excluding this patient population from accessing digital interventions based on the severity of the diagnoses and invite further research on how nonclinical individual differences impact experiences with digital tools and SDM. We also encourage research on the digitalization of mental health care and SDM that explores service, economic, and societal domains.

Acknowledgments

We want to thank all the service users and clinicians who participated in the trial. We also want to thank Jonathan Romm for creating illustrations for this paper.

Conflicts of Interest

None declared.

References

1. McGorry PD, Mei C, Chanen A, et al. Designing and scaling up integrated youth mental health care. World Psychiatry 2022 Feb;21(1):61-76. [doi: 10.1002/wps.20938] [Medline: 35015367]
2. Patel V, Saxena S, Lund C, et al. The Lancet Commission on global mental health and sustainable development. Lancet 2018 Oct;392(10157):1553-1598. [doi: 10.1016/S0140-6736(18)31612-X]
3. Barry MJ, Edgman-Levitan S. Shared decision making--pinnacle of patient-centered care. N Engl J Med 2012 Mar 1;366(9):780-781. [doi: 10.1056/NEJMp1109283] [Medline: 22375967]
4. National Institute for Health and Care Excellence (NICE). Shared Decision Making: NICE Guideline, No. 197; 2021. URL: https://www.ncbi.nlm.nih.gov/books/NBK572428/ [accessed 2024-09-09]
5. Zisman-Ilani Y, Chmielowska M, Dixon LB, et al. NICE shared decision making guidelines and mental health: challenges for research, practice and implementation. BJPsych Open 2021 Sep 2;7(5):e154. [doi: 10.1192/bjo.2021.987] [Medline: 34470688]
6. Montori VM, Ruissen MM, Hargraves IG, et al. Shared decision-making as a method of care. BMJ Evid Based Med 2023 Aug;28(4):213-217. [doi: 10.1136/bmjebm-2022-112068] [Medline: 36460328]
7. Stiggelbout AM, Van der Weijden T, De Wit MPT, et al. Shared decision making: really putting patients at the centre of healthcare. BMJ 2012 Jan 27;344:e256. [doi: 10.1136/bmj.e256] [Medline: 22286508]
8. Dy SM, Purnell TS. Key concepts relevant to quality of complex and shared decision-making in health care: a literature review. Soc Sci Med 2012 Feb;74(4):582-587. [doi: 10.1016/j.socscimed.2011.11.015] [Medline: 22236643]
9. Slade M. Implementing shared decision making in routine mental health care. World Psychiatry 2017 Jun;16(2):146-153. [doi: 10.1002/wps.20412] [Medline: 28498575]
10. Luciano M, Fiorillo A, Brandi C, et al. Impact of clinical decision-making participation and satisfaction on outcomes in mental health practice: results from the CEDAR European longitudinal study. Int Rev Psychiatry 2022;34(7-8):848-860. [doi: 10.1080/09540261.2022.2085507] [Medline: 36786107]
11. Hva er samvalg? Helsenorge. 2022. URL: https://www.helsenorge.no/samvalg/hva-er-samvalg/ [accessed 2025-06-06]

XSL•FO
RenderX

12. Meld. St. 9 (2023–2024). Nasjonal helse- og samhandlingsplan 2024–2027— Vår felles helsetjeneste. 2023. URL: https://www.regjeringen.no/no/aktuelt/nasjonal-helse-og-samhandlingsplan-2024-2027-kortere-ventetider-og-en-felles-helsetjeneste/id3028011/ [accessed 2024-02-14]
13. Huang C, Plummer V, Lam L, Cross W. Perceptions of shared decision-making in severe mental illness: an integrative review. J Psychiatr Ment Health Nurs 2020 Apr;27(2):103-127. [doi: 10.1111/jpm.12558] [Medline: 31444919]
14. Tambuyzer E, Pieters G, Van Audenhove C. Patient involvement in mental health care: one size does not fit all. Health Expect 2014 Feb;17(1):138-150. [doi: 10.1111/j.1369-7625.2011.00743.x] [Medline: 22070468]
15. Hamann J, Heres S. Adapting shared decision making for individuals with severe mental illness. PS (Wash DC) 2014 Dec 1;65(12):1483-1486. [doi: 10.1176/appi.ps.201400307]
16. Beitinger R, Kissling W, Hamann J. Trends and perspectives of shared decision-making in schizophrenia and related disorders. Curr Opin Psychiatry 2014 May;27(3):222-229. [doi: 10.1097/YCO.0000000000000057] [Medline: 24613981]
17. Haugom EW, Stensrud B, Beston G, et al. Experiences of shared decision making among patients with psychotic disorders in Norway: a qualitative study. BMC Psychiatry 2022 Mar 17;22(1):192. [doi: 10.1186/s12888-022-03849-8] [Medline: 35300633]
18. De las Cuevas C, Peñate W. To what extent psychiatric patients feel involved in decision making about their mental health care? Relationships with socio-demographic, clinical, and psychological variables. Acta Neuropsychiatr 2014 Dec;26(6):372-381. [doi: 10.1017/neu.2014.21] [Medline: 25288200]
19. Simmons MB, Brushe M, Elmes A, et al. Shared decision making with young people at ultra high risk of psychotic disorder. Front Psychiatry 2021;12:683775. [doi: 10.3389/fpsyt.2021.683775] [Medline: 34603094]
20. van der Heijden-Hobus IMW, Rosema BS, Vorstman JAS, et al. Personal preferences for treatment and care during and after a First Episode Psychosis: A qualitative study. Early Interv Psychiatry 2024 Jun;18(6):415-424. [doi: 10.1111/eip.13477] [Medline: 38030570]
21. ICD-11 for Mortality and Morbidity Statistics. World Health Organization. 2023. URL: https://icd.who.int/browse11/l-m/en [accessed 2023-11-16]
22. Mueser KT, McGurk SR. Schizophrenia. Lancet 2004 Jun 19;363(9426):2063-2072. [doi: 10.1016/S0140-6736(04)16458-1] [Medline: 15207959]
23. Carbon M, Correll CU. Thinking and acting beyond the positive: the role of the cognitive and negative symptoms in schizophrenia. CNS Spectr 2014 Dec;19 Suppl 1(S1):38-52. [doi: 10.1017/S1092852914000601] [Medline: 25403863]
24. Sullivan G, Mittal D, Reaves CM, et al. Influence of schizophrenia diagnosis on providers' practice decisions. J Clin Psychiatry 2015 Aug;76(8):1068-1074. [doi: 10.4088/JCP.14m09465] [Medline: 26335084]
25. Haugom EW, Stensrud B, Beston G, et al. Mental health professionals' experiences with shared decision-making for patients with psychotic disorders: a qualitative study. BMC Health Serv Res 2020 Nov 27;20(1):1093. [doi: 10.1186/s12913-020-05949-1] [Medline: 33246451]
26. Smith KA, Hardy A, Vinnikova A, et al. Digital mental health for Schizophrenia and other severe mental illnesses: an international consensus on current challenges and potential solutions. JMIR Ment Health 2024 May 8;11:e57155. [doi: 10.2196/57155] [Medline: 38717799]
27. Wang K, Link BG, Corrigan PW, et al. Perceived provider stigma as a predictor of mental health service users' internalized stigma and disempowerment. Psychiatry Res 2018 Jan;259:526-531. [doi: 10.1016/j.psychres.2017.11.036] [Medline: 29156425]
28. Kovács G, Almási T, Millier A, et al. Direct healthcare cost of schizophrenia - European overview. Eur Psychiatry 2018 Feb;48(1):79-92. [doi: 10.1016/j.eurpsy.2017.10.008] [Medline: 29428166]
29. Keepers GA, Fochtmann LJ, Anzia JM, et al. The American Psychiatric Association practice guideline for the treatment of patients with schizophrenia. Am J Psychiatry 2020 Sep 1;177(9):868-872. [doi: 10.1176/appi.ajp.2020.177901] [Medline: 32867516]
30. Macdonald A, Adamis D, Craig T. Continuity of care and clinical outcomes in the community for people with severe mental illness. Br J Psychiatry 2019 May;214(5):273-278. [doi: 10.1192/bjp.2018.261] [Medline: 31012407]
31. Rubio JM, Taipale H, Tanskanen A, et al. Long-term continuity of antipsychotic treatment for schizophrenia: a nationwide study. Schizophr Bull 2021 Oct 21;47(6):1611-1620. [doi: 10.1093/schbul/sbab063] [Medline: 34129663]
32. Sanatinia R, Cowan V, Barnicot K, et al. Loss of relational continuity of care in schizophrenia: associations with patient satisfaction and quality of care. BJPsych Open 2016 Sep;2(5):318-322. [doi: 10.1192/bjpo.bp.116.003186] [Medline: 27713834]
33. Fusar-Poli P, Estradé A, Stanghellini G, et al. The lived experience of psychosis: a bottom-up review co-written by experts by experience and academics. World Psychiatry 2022 Jun;21(2):168-188. [doi: 10.1002/wps.20959] [Medline: 35524616]
34. Vitger T, Korsbek L, Austin SF, et al. Digital shared decision-making interventions in mental healthcare: a systematic review and meta-analysis. Front Psychiatry 2021;12:691251. [doi: 10.3389/fpsyt.2021.691251] [Medline: 34552514]
35. Currey D, Torous J. Digital phenotyping data to predict symptom improvement and mental health app personalization in college students: prospective validation of a predictive model. J Med Internet Res 2023 Feb 9;25:e39258. [doi: 10.2196/39258] [Medline: 36757759]

XSL•FO
RenderX

36. Torous J, Bucci S, Bell IH, et al. The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality. World Psychiatry 2021 Oct;20(3):318-335. [doi: 10.1002/wps.20883] [Medline: 34505369]
37. D'Arcey J, Torous J, Asuncion TR, et al. Leveraging personal technologies in the treatment of schizophrenia spectrum disorders: scoping review. JMIR Ment Health 2024 Sep 30;11:e57150. [doi: 10.2196/57150] [Medline: 39348196]
38. Ben-Zeev D, Kaiser SM, Brenner CJ, et al. Development and usability testing of FOCUS: a smartphone system for self-management of schizophrenia. Psychiatr Rehabil J 2013 Dec;36(4):289-296. [doi: 10.1037/prj0000019] [Medline: 24015913]
39. Schlosser D, Campellone T, Kim D, et al. Feasibility of PRIME: a cognitive neuroscience-informed mobile app intervention to enhance motivated behavior and improve quality of life in recent onset schizophrenia. JMIR Res Protoc 2016 Apr 28;5(2):e77. [doi: 10.2196/resprot.5450] [Medline: 27125771]
40. Berry N, Machin M, Ainsworth J, et al. Developing a theory-informed smartphone app for early psychosis: learning points from a multidisciplinary collaboration. Front Psychiatry 2020;11:602861. [doi: 10.3389/fpsyt.2020.602861] [Medline: 33362612]
41. Edbrooke-Childs J, Edridge C, Averill P, et al. A feasibility trial of power up: smartphone app to support patient activation and shared decision making for mental health in young people. JMIR Mhealth Uhealth 2019 Jun 4;7(6):e11677. [doi: 10.2196/11677] [Medline: 31165709]
42. Vitger T, Hjorthøj C, Austin SF, et al. A smartphone app to promote patient activation and support shared decision-making in people with a diagnosis of schizophrenia in outpatient treatment settings (momentum trial): randomized controlled assessor-blinded trial. J Med Internet Res 2022 Oct 26;24(10):e40292. [doi: 10.2196/40292] [Medline: 36287604]
43. Korsbek L, Tønder ES. Momentum: A smartphone application to support shared decision making for people using mental health services. Psychiatr Rehabil J 2016 Jun;39(2):167-172. [doi: 10.1037/prj0000173] [Medline: 27030907]
44. Bucci S, Schwannauer M, Berry N. The digital revolution and its impact on mental health care. Psychol Psychother 2019 Jun;92(2):277-297. [doi: 10.1111/papt.12222] [Medline: 30924316]
45. Mohr DC, Lyon AR, Lattie EG, et al. Accelerating digital mental health research from early design and creation to successful implementation and sustainment. J Med Internet Res 2017 May 10;19(5):e153. [doi: 10.2196/jmir.7725] [Medline: 28490417]
46. Mohr DC, Weingardt KR, Reddy M. Three problems with current digital mental health research... and three things we can do about them. Psychiatr Serv 2017 May 1;68(5):427-429. [doi: 10.1176/appi.ps.201600541] [Medline: 28412890]
47. Romm KL, Skoge M, Barrett EA, et al. A mobile health intervention to support collaborative decision-making in mental health care: development and usability. JMIR Form Res 2025 Jan 17;9(9):e57614. [doi: 10.2196/57614] [Medline: 39823632]
48. Professional and service development itandem. Oslo-Universitetssykehus. 2024. URL: https://www.oslo-universitetssykehus.no/itandem [accessed 2024-10-09]
49. Creswell JW, Clark VLP. Designing and Conducting Mixed Methods Research, 3rd edition: Sage; 2018.
50. Brooke J. SUS: a "quick and dirty" usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation in Industry: Taylor & Francis; 1996:189-194. [doi: 10.1201/9781498710411]
51. Braun V, Clarke V. One size fits all? What counts as quality practice in (reflexive) thematic analysis? Qual Res Psychol 2021 Jul 3;18(3):328-352. [doi: 10.1080/14780887.2020.1769238]
52. Braun V, Clarke V. Thematic Analysis: A Practical Guide: SAGE; 2022.
53. Hyzy M, Bond R, Mulvenna M, et al. System usability scale benchmarking for digital health apps: meta-analysis. JMIR Mhealth Uhealth 2022 Aug 18;10(8):e37290. [doi: 10.2196/37290] [Medline: 35980732]
54. Stalnikowicz R, Brezis M. Meaningful shared decision ‐ making: complex process demanding cognitive and emotional skills. J Eval Clin Pract 2020 Apr;26(2):431-438. [doi: 10.1111/jep.13349] [Medline: 31989727]
55. Joseph-Williams N, Elwyn G, Edwards A. Knowledge is not power for patients: a systematic review and thematic synthesis of patient-reported barriers and facilitators to shared decision making. Patient Educ Couns 2014 Mar;94(3):291-309. [doi: 10.1016/j.pec.2013.10.031] [Medline: 24305642]
56. Aoki Y. Shared decision making for adults with severe mental illness: a concept analysis. Jpn J Nurs Sci 2020 Oct;17(4):e12365. [doi: 10.1111/jjns.12365] [Medline: 32761783]
57. Chmielowska M, Zisman-Ilani Y, Saunders R, et al. Trends, challenges, and priorities for shared decision making in mental health: the first umbrella review. Int J Soc Psychiatry 2023 Jun;69(4):823-840. [doi: 10.1177/00207640221140291] [Medline: 36680367]
58. Treichler EBH, Spaulding WD. Beyond shared decision-making: collaboration in the age of recovery from serious mental illness. Am J Orthopsychiatry 2017;87(5):567-574. [doi: 10.1037/ort0000256] [Medline: 28945446]
59. Schrank B, Slade M. Recovery in psychiatry. Psychiatr bull 2007 Sep;31(9):321-325. [doi: 10.1192/pb.bp.106.013425]
60. Clarke S, Oades LG, Crowe TP. Recovery in mental health: a movement towards well-being and meaning in contrast to an avoidance of symptoms. Psychiatr Rehabil J 2012;35(4):297-304. [doi: 10.2975/35.4.2012.297.304] [Medline: 22491369]
61. Mundal I, Lara-Cabrera ML, Betancort M, et al. Exploring patterns in psychiatric outpatients' preferences for involvement in decision-making: a latent class analysis approach. BMC Psychiatry 2021 Mar 6;21(1):133. [doi: 10.1186/s12888-021-03137-x] [Medline: 33676452]

XSL•FO
RenderX

62. Glick ID, Stekoll AH, Hays S. The role of the family and improvement in treatment maintenance, adherence, and outcome for schizophrenia. J Clin Psychopharmacol 2011 Feb;31(1):82-85. [doi: 10.1097/JCP.0b013e31820597fa] [Medline: 21192148]
63. Pharoah F, Mari JJ, Rathbone J, Wong W. Family intervention for schizophrenia. Cochrane Database Syst Rev 2010 Dec 8;8(12):CD000088. [doi: 10.1002/14651858.CD000088.pub2] [Medline: 21154340]
64. Gayer-Anderson C, Morgan C. Social networks, support and early psychosis: a systematic review. Epidemiol Psychiatr Sci 2013 Jun;22(2):131-146. [doi: 10.1017/S2045796012000406] [Medline: 22831843]
65. Wang J, Mann F, Lloyd-Evans B, et al. Associations between loneliness and perceived social support and outcomes of mental health problems: a systematic review. BMC Psychiatry 2018 May 29;18(1):156. [doi: 10.1186/s12888-018-1736-5] [Medline: 29843662]
66. Bjørlykhaug KI, Karlsson B, Hesook SK, et al. Social support and recovery from mental health problems: a scoping review. Nord Soc Work Res 2022 Dec 5;12(5):666-697. [doi: 10.1080/2156857X.2020.1868553]
67. May C. Agency and implementation: understanding the embedding of healthcare innovations in practice. Soc Sci Med 2013 Feb;78:26-33. [doi: 10.1016/j.socscimed.2012.11.021] [Medline: 23246396]
68. Rogers L, De Brún A, McAuliffe E. Defining and assessing context in healthcare implementation studies: a systematic review. BMC Health Serv Res 2020 Jun 29;20(1):591. [doi: 10.1186/s12913-020-05212-7] [Medline: 32600396]
69. Holden RJ, Karsh BT. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010 Feb;43(1):159-172. [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]
70. Eliason M. The financial situation before and after first-time psychiatric in-patient diagnosis of schizophrenia spectrum, bipolar, and major depressive disorder. Econ Hum Biol 2023 Apr;49:101231. [doi: 10.1016/j.ehb.2023.101231] [Medline: 36827797]
71. Spanakis P, Wadman R, Walker L, et al. Measuring the digital divide among people with severe mental ill health using the essential digital skills framework. Perspect Public Health 2024 Jan;144(1):21-30. [doi: 10.1177/17579139221106399] [Medline: 35929589]
72. Eisner E, Berry N, Bucci S. Digital tools to support mental health: a survey study in psychosis. BMC Psychiatry 2023 Oct 7;23(1):726. [doi: 10.1186/s12888-023-05114-y] [Medline: 37803367]
73. Eisner E, Richardson C, Thomas N, et al. Measurement of adverse events in studies of digital health interventions for psychosis: guidance and recommendations based on a literature search and framework analysis of standard operating procedures. Schizophr Bull 2024 Nov 8;50(6):1456-1470. [doi: 10.1093/schbul/sbae048] [Medline: 38683836]
74. Lecomte T, Potvin S, Corbière M, et al. Mobile apps for mental health issues: meta-review of meta-analyses. JMIR Mhealth Uhealth 2020 May 29;8(5):e17458. [doi: 10.2196/17458] [Medline: 32348289]
75. Huckvale K, Nicholas J, Torous J, et al. Smartphone apps for the treatment of mental health conditions: status and considerations. Curr Opin Psychol 2020 Dec;36:65-70. [doi: 10.1016/j.copsyc.2020.04.008] [Medline: 32553848]
76. Müller K, Schuster F, Rodolico A, et al. How should patient decision aids for schizophrenia treatment be designed? - A scoping review. Schizophr Res 2023 May;255:261-273. [doi: 10.1016/j.schres.2023.03.025] [Medline: 37062107]
77. Roebroek LO, Bruins J, Boonstra A, et al. The effects of a computerized clinical decision aid on clinical decision-making in psychosis care. J Psychiatr Res 2022 Dec;156:532-537. [doi: 10.1016/j.jpsychires.2022.10.053] [Medline: 36356555]
78. Roebroek LO, Bruins J, Boonstra A, et al. The effects of treatment as usual versus a computerized clinical decision aid on shared decision-making in the treatment of psychotic disorders. Eur J Psychiatry 2024 Jan;38(1):100216. [doi: 10.1016/j.ejpsy.2023.06.002]

Abbreviations

ICD-10: International Classification of Diseases, 10th Revision
PRIME: Personalized Real-time Intervention for Motivation Enhancement
SDM: shared decision-making
SUS: System Usability Scale
TREAT: TReatment E-AssisT

XSL•FO
RenderX

Edited by C Jacob; submitted 14.11.24; peer-reviewed by A Rabasco, J Bruins; revised version received 06.05.25; accepted 09.05.25; published 11.07.25.

Please cite as:
Skoge M, Aminoff SR, Barrett EA, Bryhni GE, Kling K, Kværner KJ, Melle I, Mork E, Simonsen C, Støme LN, Vink J, Værnes TG, Romm KL
A Mobile App Designed to Promote Shared Decision-Making in the Treatment of Psychotic Disorders: Feasibility and Acceptability Study
JMIR Hum Factors 2025;12:e68813
URL: https://humanfactors.jmir.org/2025/1/e68813
doi:10.2196/68813

© Mari Skoge, Sofie Ragnhild Aminoff, Elizabeth Ann Barrett, Gina Engen Bryhni, Kristine Kling, Kari Jorunn Kværner, Ingrid Melle, Erlend Mork, Carmen Simonsen, Linn Nathalie Støme, Josina Vink, Tor Gunnar Værnes, Kristin Lie Romm. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 11.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Using Personalized Intervention Criteria in a Mobile Just-in-Time Adaptive Intervention for Increasing Physical Activity in University Students: Pilot Study

Mai Ikegaya[1], MSc; Jerome Clifford Foo[2,3,4,5], PhD; Taiga Murata[1], MSc; Kenta Oshima[1], BSc; Jinhyuk Kim[1], PhD

[1]Department of Informatics, Graduate School of Integrated Science and Technology, Shizuoka University, 3-5-1 Johoku, Chuo-ku, Hamamatsu, Shizuoka, Japan

[2]Department of Psychiatry, College of Health Sciences, University of Alberta, Edmonton, AB, Canada

[3]Neuroscience and Mental Health Institute, University of Alberta, Edmonton, AB, Canada

[4]Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany

[5]Institute for Psychopharmacology, Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany

Corresponding Author:
Jinhyuk Kim, PhD
Department of Informatics, Graduate School of Integrated Science and Technology, Shizuoka University, 3-5-1 Johoku, Chuo-ku, Hamamatsu, Shizuoka, Japan

Abstract

Background: While the health benefits of physical activity are well-known, adherence to regular physical activity remains a major challenge. Just-in-time adaptive intervention (JITAI) has been proposed as one method to increase physical activity by delivering an intervention at a time when individuals are more likely to make behavioral changes. However, most studies that have implemented JITAI have used uniform intervention criteria (UIC) across participants rather than personalized intervention criteria (PIC) for the individual.

Objective: The objective of this paper was to examine the effectiveness of using JITAI implemented with PIC to increase physical activity.

Methods: Healthy university students wore a wrist activity monitor for 2 weeks. Participants were divided into 2 groups, which received JITAI to promote physical activity according to either PIC or UIC. In the first week, the mean distance moved and sedentary time per hour for each participant were calculated to derive PIC. UIC was obtained from a 2-week study with a different sample (n=47) conducted under the same conditions. In the second week, JITAI prompts were sent every hour if both of the following criteria were met: the distance moved was shorter, and sedentary time was longer than PIC or UIC. Differences in changes in physical activity as a result of implementing interventions according to PIC and UIC were analyzed using multilevel models.

Results: We analyzed data from 28 healthy university students (18 - 23 y old, female n=12). Both PIC (P<.001) and UIC (P<.001) significantly increased physical activity in the first hour after JITAI was received. In that first hour, PIC increased physical activity more than UIC; more calories were burned (P=.02), more steps were taken (P=.007), and distance moved was increased (P=.003). However, over the course of the week, the use of JITAI did not significantly increase physical activity levels.

Conclusions: Our results appear to suggest that PIC-based JITAI is more effective than UIC-based JITAI, consistent with the idea of a need for precision health approaches. Further research is needed to develop effective long-term intervention designs with sustainable effects.

(JMIR Hum Factors 2025;12:e66750) doi:10.2196/66750

KEYWORDS

physical activity; just-in-time adaptive intervention; mHealth; consumer-wearable activity trackers; multilevel model; activity tracker; wearable; behavioral change; effectiveness; activity monitor; activity monitoring; long-term intervention; chronic disease; prevention; treatment; mobile health; mobile phone; digital health; smartphone

Introduction

The benefits of physical activity for prevention and management of chronic diseases and mortality have been well established [1-3]. To promote physical activity for better health, the World Health Organization (WHO) has recommended 150 - 300 minutes of moderate-intensity aerobic physical activity per week, or 75 - 150 minutes of vigorous-intensity aerobic physical

XSL•FO
RenderX

activity throughout the week, or an equivalent combination of moderate- to vigorous-intensity activity [4]. However, a pooled analysis of 507 population-based surveys (n=5.7 million) found that 31.3% of adults worldwide do not meet this recommended amount of physical activity [5]. Low levels of daily physical activity have become a global health problem, and methods to increase physical activity using mobile health (mHealth) have started to be implemented [6-8].

mHealth is defined as health care services supported through the use of mobile devices such as mobile phones [9]. The aim of mHealth is to enhance users' health and well-being through mobile devices and other wireless technologies by monitoring and analyzing their health status, providing support for disease prevention and treatment, and management and introduction of interventions for disease in everyday environments [10]. Previous studies incorporating mHealth interventions have shown beneficial effects in promoting health behaviors and support for behavioral change (eg, increasing physical activity [9,11,12]).

Just-in-time adaptive intervention (JITAI) is an intervention design that adapts the timing, type, and intensity of intervention support to changes in an individual's status and behavior [13]. The goal is to provide support at the timing and in the context that the individual most needs and is most likely to accept it. JITAI has shown promising effects for improving physical activity in research settings. For example, van Dantzig et al [14] found that when they encouraged a 5-minute break after 60 minutes of continuous sitting, the amount of sedentary time decreased. Finkelstein et al [15] found that reminders promoting exercise sent when participants took 15 or fewer steps per hour significantly decreased sedentary behavior compared to when they did not receive reminders. Bond et al [16] found that 3-minute break reminders after 30 minutes of sitting decreased sedentary time and significantly increased the amount of time spent doing light physical activity compared to no reminder. Fiedler et al [17] found that delivering the prompt when the participant was inactive for more than 60 minutes significantly increased moderate to vigorous physical activity (MVPA). Mair et al [18] showed the feasibility of using personalized smartphone-delivered JITAI in elderly people. JITAI was tailored to the individual based on real time physical activity, daily physical activity goal, time of day, and weather conditions [18].

In most JITAI studies to date, however, the criteria for the intervention have been uniform (eg, 15 steps fewer per hour and sitting continuously for 30 or 60 minutes), and have not reflected the different activity patterns of the individual [12,18]. Uniform intervention criteria (UIC) do not accommodate varying levels of preexisting physical activity, and as such may not be the most effective way to promote uptake. In this study, we compared the effects of increasing activity in groups that received physical activity interventions using UIC versus groups that received an intervention using personalized intervention criteria (PIC) tailored to the individual's activity patterns. We hypothesize that the intervention using PIC is more effective in increasing physical activity than the intervention using UIC.

Methods

Participants

A total of 31 healthy university students in Japan aged 18‐23 years old participated in the study. Healthy university students under 30 years of age were eligible to be included in the study. Participants who had current medical problems or were using medication that could influence physical activity were not eligible to participate. This prospective 2-week randomized comparative study took place between December 2022 and January 2023. The participants were randomly assigned into 2 groups in a 1:1 ratio by using a random number table (integers 0 and 1): PIC group, which received the intervention with personalized intervention criteria, and UIC group, which received the intervention with uniform intervention criteria. While a within-persons design was considered (both PIC and UIC in a single individual), due to the time limitations and potential mixing of effects of both types of intervention criteria, a between-persons design was used. Details are reported in the CONSORT-EHEALTH V 1.6.1 checklist (see Checklist 1).

Ethical Considerations

The study was approved by the institutional review board at Shizuoka University, Japan (19‐8). Participants provided written informed consent after receiving a full explanation about the purpose of the study, methods, and precautions for participation in the study. Participants who completed the study received an Amazon gift card worth JP ¥5000 (approximately US $35).

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and national research committee. Participants provided written informed consent. All data used in this study were anonymized to protect participants' privacy and ensure confidentiality.

Assessments

Questionnaires

Participants completed a questionnaire before the 2-week study to collect their age, sex, height, weight, medical history, and drug use. Participants who had current medical problems or used medication were excluded from the study.

Activity Monitor

Physical activity was measured using a wrist-worn wearable activity monitor (Fitbit Inspire 2). The active monitor estimates sedentary time (minutes), lightly active time (minutes), fairly active time (minutes), very active time (minutes), calories (kcal), steps (n), and distance moved (m). Steps are calculated by the 3-axis accelerometer, while distance moved is automatically calculated based on the number of steps and the participant's stride length. Time spent at different activity levels are also calculated by the activity monitor (sedentary time, lightly active time, fairly active time, and very active time). Activity monitor and participant's smartphone were synchronized via Bluetooth by installing the Fitbit app on their smartphones, and activity data were recorded on the Fitbit server via the smartphone. The activity monitor was worn on the nondominant wrist at all times

XSL•FO
RenderX

during the study period. Participants were instructed to remove the monitors when showering or bathing.

Intervention

Intervention Criteria (PIC vs UIC)

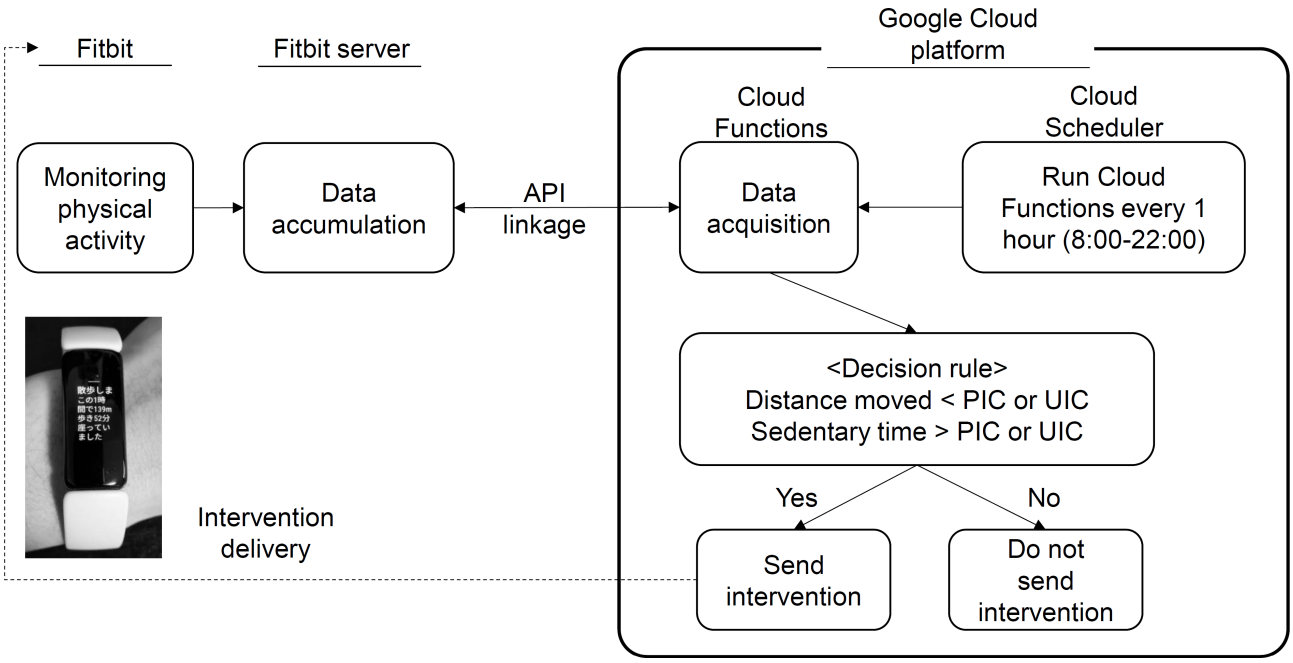
The intervention criteria used were distance moved (m) and sedentary time (minutes). For distance moved, 1 SD lower than the mean (mean distance, SD) was used, and for sedentary time, 1 SD longer than the mean (mean sedentary time + SD) were used. The reason for extending the intervention criteria to include SD rather than using only the mean of physical activity was to ensure that interventions are not triggered when activity levels differed only slightly from the mean. That is, interventions might be too frequent and difficult to adhere to if the threshold was strictly set at the mean values.

Set points for interventions were determined differently for the 2 groups. PIC group received interventions based on individual activity data measured during week 1. The UIC group received interventions based on group data from a 2-week preliminary study conducted a year earlier using the same devices looking at physical activity in a sample of university students (n=47). The UIC were determined using a different sample because the timing of the study period was different for each participant; participants were recruited and participated during different dates so the UIC could not be calculated at the start of the intervention period for all participants.

Intervention Structure and Flow

Figure 1 shows the intervention structure and flow. The intervention was performed using the tools "Cloud Functions" and "Cloud Scheduler" of the Google Cloud platform via the Fitbit application programming interface (API). Cloud Functions is a serverless runtime environment used to build and connect cloud services and to run the process of retrieving physical activity data from activity monitor, calculating intervention criteria with the physical activity data and delivering feedback to participants as an intervention for increasing physical activity. Cloud Scheduler is a scheduling function for processing functions at a specific time, and was used to repeatedly run the process as needed. Cloud Scheduler was set up to trigger Cloud Functions at the 34-minute mark every hour (ie, 10:34 AM, 11:34 AM, 12:34 PM…) to evaluate the physical activity in the past hour (ie, 9:34 AM-10:33 AM, 10:34 AM-11:33 AM, 11:34 AM-12:33 PM…). Cloud Functions were used to retrieve 1-minute interval data from the Fitbit server through the Fitbit API. When both (1) "Distance moved" was less than the intervention criteria and (2) the "Sedentary time" was longer than the intervention criteria, participants were notified. Intervention decisions were based on data on distance moved and sedentary time in the past hour.

Figure 1. Intervention structure and flow. API: application programming interface; PIC: personalized intervention criteria; UIC: uniform intervention criteria.



Intervention Timing

An intervention decision was made once per hour by the programmed Cloud Scheduler of the Google Cloud platform, and the intervention message was delivered when the intervention criteria met the predefined conditions (ie, UIC or PIC). The intervention schedule was run from 8:00 AM to 10:00 PM.

Intervention Contents

The intervention consisted of sending an email to the participant's email address. The email included a message encouraging physical activity and stating the participant's distance moved and sedentary time in the last hour. One message (in Japanese) was chosen at random each time from the following 5 messages: "Would you like to take a walk?" "Would you like to exercise?" "Move your body!" "Let's exercise!" and "Let's move a little!" The Fitbit app on the smartphone allowed

XSL•FO
RenderX

email notifications, and when the email was received, the activity monitor vibrated, and the contents of the email were displayed on the Fitbit screen.

Statistical Analysis

Analysis was performed using SAS 9.4 (SAS Institute Inc). The physical activity data acquired are longitudinal and repeatedly measured from individuals, and thus have a hierarchical structure. We used multilevel modeling, a widely used analytical method for such data structures [19]. We specified 2 different sets of models to examine the effectiveness of using JITAI implemented with PIC to increase physical activity. Model set A was specified to examine the effects of PIC or UIC on physical activity after the intervention. To examine the change in physical activity after 1 hour of intervention by different intervention criteria, the mean level of physical activity was estimated for the group using UIC and the group using PIC at the hour before and the hour after the intervention. The physical activity types estimated were sedentary time (minutes), lightly active time (minutes), fairly active time (minutes), very active time (minutes), calories (kcal), steps (n), and distance moved (m). We ran 7 different models, using each of these physical activity parameters as dependent variables. The intervention group (ie, PIC group and UIC group) and before or after intervention (ie, the hour before and the hour after the intervention) were used as independent variables.

Model Set A

Level 1 model (within-participant):

$$(1)\text{Physical Activity}ti=\beta 0i+\beta 1i\times(\text{Before or After Intervention}ti)+\varepsilon ti$$

Level 2 models (between-participants):

$$(2)\beta 0i=\gamma 00+\gamma 01\times(\text{Intervention Group}i)+\gamma 02\times(\text{BMI}i)+\gamma 03\times(\text{Sex}i)+\upsilon 0i$$

$$(3)\beta 1i=\gamma 10+\gamma 11\times(\text{Intervention Group}i)+\upsilon 1i$$

where "*Physical Activity*$_{ti}$" indicates the dependent variable (sedentary time, lightly active time, fairly active time, very active time, calories, steps, or distance moved) at the *t*th recording for the *i*th participant; β_{0i} is participant i's intercept; β_{1i} is participant i's slope; γ_{00} is the average intercept across all participants; γ_{01}, γ_{02}, γ_{03}, and γ_{11} are all participants' slopes; γ_{10} is the slope of β_{1i}; ε_{ti} is the within-individual residuals; υ_{0i} and υ_{1i} are the random effects of intercept and slope, respectively. The random or fixed effect of slope was determined by comparing Akaike information criterion (AIC) to find the best-fitting parameters for each. As an example, we have shown Equation 3 which has a random effect of slope (υ_{1i}). To test the differences before or after intervention in mean level of physical activity, we added a categorical variable representing "*Before or After Intervention*$_{ti}$" into the right-hand side of Equation 1. In addition, to test the difference in mean level of physical activity between PIC and UIC groups, we added a categorical variable representing "*Intervention Group*$_i$" into the right-hand side of Equations 2 and 3. The slope (γ_{11}) of an interaction term between "*Before or After Intervention*$_{ti}$" and "*Intervention Group*$_i$" represents the difference in physical activity for the group using UIC and the group using PIC at the hour before and the hour after the intervention in the combined multilevel model. In Equation 2, BMI and sex were adjusted for in the Level 2 (between-participants) models.

Model set B was specified to examine changes in physical activity from Week 1 (without intervention) to Week 2 (with intervention) associated with the different intervention criteria (UIC or PIC). A total of 7 separate models were specified using the hourly mean of each physical activity parameter (as above) as a dependent variable. Intervention group and week were specified as independent variables.

Model Set B

Level 1 model (within-participant):

$$(4)\text{Physical Activity}ti=\beta 0i+\beta 1i\times(\text{Week 1 or Week 2}ti)+\varepsilon ti$$

Level 2 models (between-participants):

$$(5)\beta 0i=\gamma 00+\gamma 01\times(\text{Intervention Group}i)+\gamma 02\times(\text{BMI}i)+\gamma 03\times(\text{Sex}i)+\upsilon 0i$$

$$(6)\beta 1i=\gamma 10+\gamma 11\times(\text{Intervention Group}i)+\upsilon 1i$$

where "*Physical Activity*$_{ti}$" indicates the dependent variable (sedentary time, lightly active time, fairly active time, very active time, calories, steps, or distance moved) at the *t*th recording for the *i*th participant; β_{0i} is participant i's intercept; β_{1i} is participant i's slope; γ_{00} is the average intercept across all participants; γ_{01}, γ_{02}, γ_{03}, and γ_{11} are all participants' slopes; γ_{10} is the slope of β_{1i}; ε_{ti} is the within-individual residuals; υ_{0i} and υ_{1i} are the random effects of intercept and slope, respectively. The random or fixed effect of slope was determined by comparing AIC to find the best-fitting parameters for each. As an example, we have shown Equation 6 which has a random effect of slope (υ_{1i}). To test the change in physical activity from Week 1 without intervention to Week 2 with intervention, we added a categorical variable representing "*Week 1 or Week 2*$_{ti}$" into the right-hand side of Equation 4. In addition, to test the difference in mean level of physical activity between PIC and UIC groups, we added a categorical variable representing "*Intervention Group*$_i$" into the right-hand side of Equations 5 and 6. In Equation 5, BMI and sex were adjusted for in the Level 2 (between-participants) models. In an effort to compare intervention effects more precisely, we used the hourly averaged physical activity to minimize variation due to factors such as sleep length and the time participants did not wear the activity monitor.

For all analyses, data were excluded when participants were sleeping and not wearing the activity monitor. A $P<.05$ was considered significant. The Tukey-Kramer method was used to adjust for multiple comparisons.

Results

Participant Characteristics

While 31 participants were recruited for the study, 3 participants did not participate for more than 4 days during their 2-week periods and were excluded from the analysis. There were no significant differences in characteristics between the intervention groups, except for in BMI (Table 1).

XSL•FO
RenderX

Table . Participant characteristics. Two-sample *t* tests were performed to compare the age, height, weight, and BMI between the PIC and UIC groups. χ^2 tests were performed to compare the sex distribution across the two groups.

Characteristics	All participants (n=28)	PIC[a] group (n=14)	UIC[b] group (n=14)	*t* (*df*)	χ^2 (*df*)	*P* value
Age (years), mean (SD)	20.6 (1.7)	20.6 (1.4)	20.6 (1.2)	0.00 (26)	—[c]	≥.99
Height (cm), mean (SD)	166.2 (7.3)	168.2 (6.9)	164.3 (7.2)	1.41 (26)	—	.17
Weight (kg), mean (SD)	54.8 (6.6)	53.3 (5.7)	56.3 (7.2)	−1.18 (26)	—	.25
BMI (kg/m^2), mean (SD)	19.8 (1.9)	18.8 (1.1)	20.8 (2.0)	–3.24 (26)	—	.003
Sex				—	210.58 (1)	.45
Male, n (%)	16 (57)	9 (64)	7 (50)			
Female, n (%)	12 (43)	5 (36)	7 (50)			

[a]PIC: personalized intervention criteria.

[b]UIC: uniform intervention criteria.

[c]Not applicable.

The PIC group received interventions based on individual activity data measured during week 1. The UIC group received interventions based on group data from a 2-week preliminary study conducted a year earlier using the same devices looking at physical activity in a sample of university students (n=47; see Table 2).

Table . Participant intervention criteria.

ID	Distance moved (m)	Sedentary time (min)
PIC[a]1	151	50
PIC2	127	52
PIC3	125	52
PIC4	164	46
PIC5	263	41
PIC6	105	51
PIC7	93	55
PIC8	152	50
PIC9	148	50
PIC10	140	51
PIC11	39	56
PIC12	18	57
PIC13	57	55
PIC14	294	42
UIC[b] (n=14)	139	52

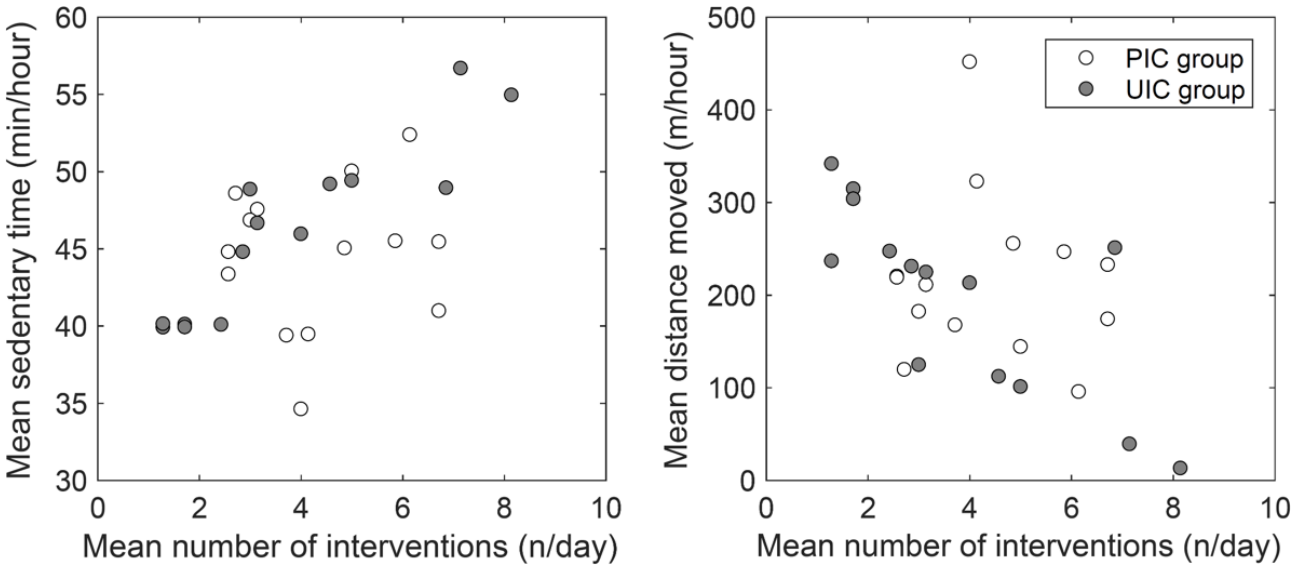
[a]PIC: personalized intervention criteria.

[b]UIC: uniform intervention criteria.

The mean number of interventions for the PIC group was 4.4 per day (SD 1.5; range 1.3 - 8.1), and the mean number of interventions for the UIC group was 3.8 per day (SD 2.3; range 2.6 - 6.7). No significant difference in the mean number of interventions was found between the PIC and UIC groups. However, some participants in the UIC group who had higher average levels of physical activity received a lower number of interventions (ie, a mean of 1 - 2 interventions per day) to increase activity, while some participants who had lower average physical activity levels received a higher number of interventions (ie, a mean of 7 - 8 interventions per day; see Figure 2).

XSL•FO
RenderX

Figure 2. Mean number of interventions per day in accordance with physical activity level in personalized intervention criteria (PIC) versus uniform intervention criteria (UIC) group.



Model Set A: Effects of PIC and UIC on Physical Activity for the Hour Before and the Hour After the Intervention

In model set A, the best models for sedentary time and lightly active time used random intercepts and slopes, while the best models for fairly active time, very active time, calories, steps, and distance moved used random intercepts and fixed slopes. For both the PIC and UIC groups, sedentary time significantly decreased after the intervention (PIC group: before *B*=55.83, SE=0.54; after *B*=48.21, SE=1.07; *P*<.001 and UIC group: before *B*=57.11, SE=0.58; after *B*=49.88, SE=1.10; *P*<.001; see Table 3). The other 6 physical activity types (ie, lightly active time, fairly active time, very active time, calories, steps, and distance moved) significantly increased after the intervention in intervention groups (*P*<.001; see Table 3). See Table S1 in Multimedia Appendix 1 for detailed information.

Table . The change in physical activity 1 hour after the intervention as affected by differences in intervention criteria.

Physical activity	PIC[a] group			UIC[b] group		
	Before intervention	After intervention	*P* value	Before intervention	After intervention	*P* value
Sedentary time (min), mean (SE)	55.83 (0.54)	48.21 (1.07)	<.001	57.11 (0.58)	49.88 (1.10)	<.001
Lightly active time (min), mean (SE)	4.02 (0.53)	9.80 (0.94)	<.001	2.95 (0.56)	8.76 (0.97)	<.001
Fairly active time (min), mean (SE)	−0.04 (0.10)	0.86 (0.10)	<.001	0.08 (0.11)	0.65 (0.11)	.003
Very active time (min), mean (SE)	0.01 (0.12)	0.95 (0.12)	<.001	0.04 (0.13)	0.59 (0.13)	.01
Calories (kcal), mean (SE)	74.54 (3.08)	99.83 (3.08)	<.001	68.97 (3.24)	87.10 (3.24)	<.001
Steps (n), mean (SE)	66.35 (28.70)	403.35 (28.70)	<.001	63.80 (30.48)	285.67 (30.48)	<.001
Distance moved (m), mean (SE)	26.82 (12.52)	175.70 (12.52)	<.001	27.06 (13.31)	119.85 (13.31)	<.001

[a]PIC: personalized intervention criteria.

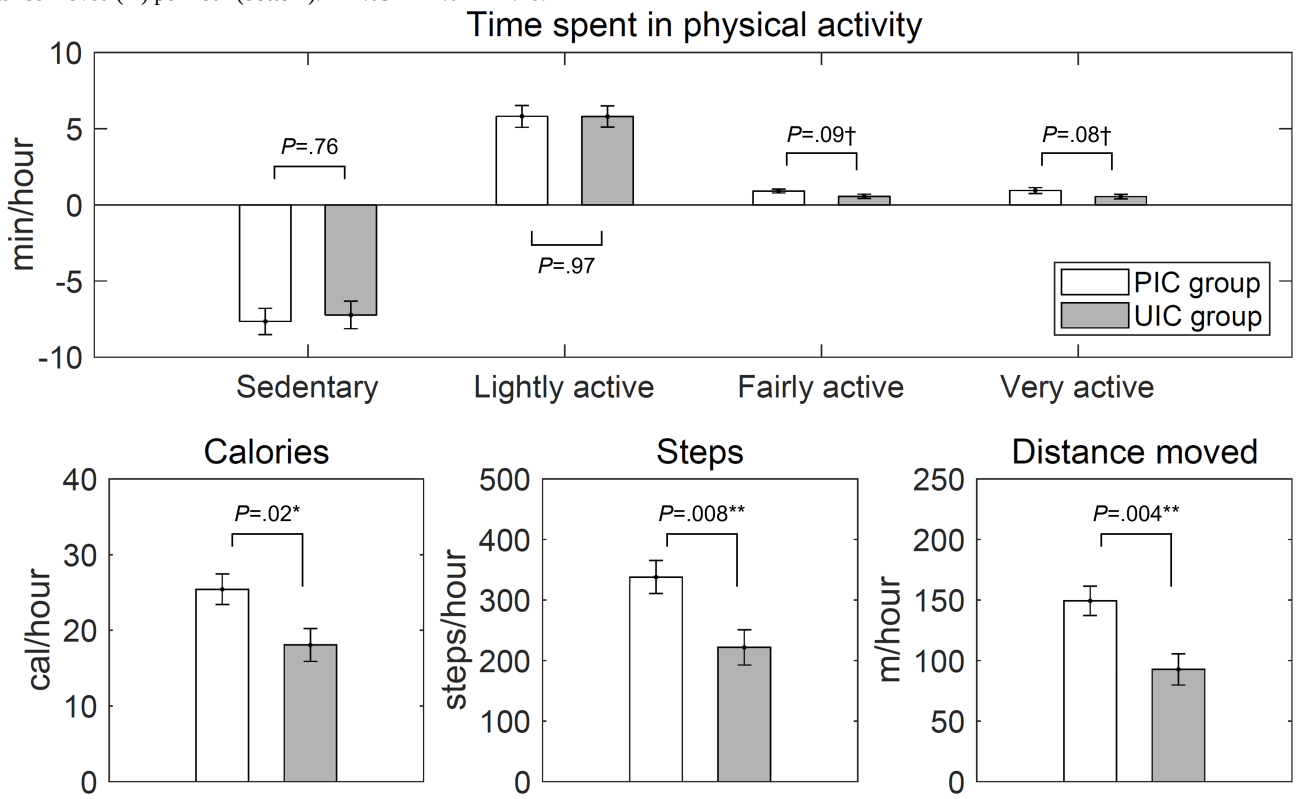
[b]UIC: uniform intervention criteria.

Calories burned (PIC group: *B*=25.29, SE=2.03 and UIC group: *B*=18.13, SE=2.18; *P*=.02), Steps taken (PIC group: *B*=336.99, SE=27.45 and UIC group: *B*=221.88, SE=29.45; *P*=.008), and distance moved (PIC group: *B*=148.89, SE=12.19 and UIC group: *B*=92.79, SE=13.07; *P*=.004) showed significant differences in the increase in physical activity the hour before and after the intervention between the PIC group and the UIC group (Figure 3). Specifically, the number of calories burned per hour was estimated to increase by 7.16 kcal more in the PIC group than in the UIC group after the intervention. Steps taken increased by 115.11 steps and distance moved increased by 56.10 m in the PIC group compared with the UIC group. There were no significant differences in change observed between

XSL•FO
RenderX

groups in the other activity types (see Table S1 in Multimedia Appendix 1 for details).

Figure 3. Changes in physical activity from the hour before to the hour after the intervention in personalized intervention criteria (PIC) versus uniform intervention criteria (UIC) group. The values represent time spent (min) doing physical activity (top), calories burned (kcal), steps taken (steps), or distance moved (m) per hour (bottom). $^{*}P<.05$ $^{**}P<.01$ $^{\dagger}P<.10$.



Model Set B: Effects of PIC and UIC on Physical Activity From Week 1 Without Intervention to Week 2 With Intervention

In model set B, the best models for lightly active time, fairly active time, calories, steps, and distance moved used random intercepts and slopes, while the best models for sedentary time and very active time used random intercepts and fixed slopes. In both PIC and UIC groups, there were no significant changes in physical activity (all parameters) between week 1 and week 2 (see Table 4). See Table S2 in Multimedia Appendix 1 for detailed information.

Table . The change in physical activity from week 1 without intervention to week 2 with intervention as affected by differences in intervention criteria.

Physical activity	PIC group			UIC group		
	Week 1	Week 2	*P* value	Week 1	Week 2	*P* value
Sedentary time (min), mean (SE)	43.01 (1.36)	42.26 (1.36)	.55	44.82 (1.36)	45.38 (1.36)	.76
Lightly active time (min), mean (SE)	14.01 (1.19)	14.52 (1.12)	.92	12.81 (1.19)	11.94 (1.12)	.69
Fairly active time (min), mean (SE)	1.48 (0.21)	1.75 (0.29)	.67	1.35 (0.21)	1.70 (0.30)	.50
Very active time (min), mean (SE)	1.44 (0.19)	1.41 (0.19)	≥.99	1.06 (0.19)	1.02 (0.19)	≥.99
Calories (kcal), mean (SE)	116.91 (4.36)	119.06 (5.81)	.94	106.99 (4.34)	105.32 (5.80)	.97
Steps (steps), mean (SE)	590.46 (58.36)	564.56 (62.62)	.94	525.15 (58.00)	497.16 (62.50)	.93
Distance moved (m), mean (SE)	257.31 (26.14)	245.49 (27.81)	.93	230.32 (25.98)	214.77 (27.75)	.85

XSL•FO
RenderX

Discussion

Principal Findings

In this study, we examined the differences between in the impact of interventions based on personalized intervention criteria and uniform intervention criteria on physical activity in daily life. We found that in the first hour after JITAI was delivered, physical activity in the PIC group increased more than in the UIC group. However, the use of JITAI (both criteria) did not significantly increase physical activity over the whole study period.

There was a significant difference in increase in calories burned, steps taken, and distance moved between the PIC and UIC groups. This indicates that interventions based on PIC were more effective, suggesting that interventions to promote physical activity need to use individually developed intervention criteria based on physical activity in daily life. It has been reported that low feasibility of intervention suggestions is likely to lead to dissatisfaction [20,21]. It is possible that the setting of individually optimized intervention criteria that presented highly feasible goals (eg, proper number of interventions; see Figure 2) based on their physical activity level may have made it easier for them to change their behavior. Further studies need to identify the optimal number of interventions per day to promote physical activity.

Interestingly, in contrast to previous studies, we did not observe differences in overall physical activity between the first week without intervention and the week with intervention in either group. This may be due to the length of the intervention period used in this study. A previous study using UIC showed found an increase in physical activity after 21 days of intervention [16], while another study that sent contextually tailored intervention messages for 6 weeks showed an increase in the number of steps, although the increase diminished over time [22]. Taken together, it is possible that a longer intervention period is needed to effectively promote increased physical activity. A possible additional factor explaining this was that all but one participant had never worn activity monitors before, and simply wearing the monitors during the first week might have affected their baseline activity levels [23]. Future studies may need to set a burn-in period to precisely identify the baseline of physical activity before starting intervention.

This study had several limitations. Participants were university students, and the number of participants was small (n=28). Considering the sample size and the limited number of personal characteristics available to include as covariates (ie, age, BMI, and sex), it is possible that group differences (PIC vs UIC) may be related to other hidden third variables. Further studies need to examine the effectiveness of JITAI using PIC in larger samples with more personal characteristics, especially for participants who have low average levels of physical activity (eg, obese and elderly people who need more personalized intervention criteria). In addition, the variables of physical activity (ie, sedentary time and distance moved) we used for intervention were limited and in the future, more criteria could be used to determine thresholds to trigger interventions, and more sophisticated approaches to determine intervention decisions (eg, machine learning classification) could be used with additional physical activity features. A technical limitation was that the time of the intervention decision could not be randomized and was fixed. In future JITAI research, a finer grained approach ensuring that it is possible to intervene at the timing when the intervention criteria are met might lead to more optimal timing of the intervention. Finally, this study was limited to 2 weeks in duration and did not include any follow-up. Programs for increasing physical activity require the promotion of long-term behavioral changes. Further studies need to consider maintenance of benefits and sustainability of programs over longer periods of time.

Conclusion

We found that personalized intervention criteria were more effective than uniform intervention criteria in increasing physical activity levels, supporting the idea that precision health approaches tailored to the individual may lead to beneficial effects on health. The findings of this pilot study warrant further investigation; JITAI PIC programs need to be assessed over longer durations and in different populations.

Acknowledgments

We would like to thank Mr Kazuki Sugiyama for his assistance with data collection. This work was supported by Grant-in-Aid for Early-Career Scientists (21K17681) and Grant-in-Aid for Scientific Research (C) (24K14745) from the Ministry of Education, Culture, Sports, Science and Technology, and the German Federal Ministry of Education and Research (BMBF) through the e:Med Programme Target-OXY grant (031L0190A).

Data Availability

The datasets generated or analyzed during this study may be made available from the corresponding author on reasonable request.

Authors' Contributions

JK designed the study. TM, KO, and JK collected the data. MI, JCF, TM, and JK cleaned and analyzed the data. MI wrote the first draft of the manuscript. JCF and JK reviewed and edited the draft of the manuscript. All authors approved the final version of the manuscript.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables providing detailed information on the effects of intervention criteria on physical activity.
[DOCX File, 80 KB - humanfactors_v12i1e66750_app1.docx]

Checklist 1
CONSORT-EHEALTH checklist. CONSORT: Consolidated Standards of Reporting Trails.
[PDF File, 1333 KB - humanfactors_v12i1e66750_app2.pdf]

References

1. Warburton DER, Nicol CW, Bredin SSD. Health benefits of physical activity: the evidence. Can Med Assoc J 2006 Mar 14;174(6):801-809. [doi: 10.1503/cmaj.051351]
2. Mandsager K, Harb S, Cremer P, Phelan D, Nissen SE, Jaber W. Association of cardiorespiratory fitness with long-term mortality among adults undergoing exercise treadmill testing. JAMA Netw Open 2018 Oct 5;1(6):e183605. [doi: 10.1001/jamanetworkopen.2018.3605] [Medline: 30646252]
3. Stamatakis E, Ahmadi MN, Gill JMR, et al. Association of wearable device-measured vigorous intermittent lifestyle physical activity with mortality. Nat Med 2022 Dec;28(12):2521-2529. [doi: 10.1038/s41591-022-02100-x] [Medline: 36482104]
4. Bull FC, Al-Ansari SS, Biddle S, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. Br J Sports Med 2020 Dec;54(24):1451-1462. [doi: 10.1136/bjsports-2020-102955] [Medline: 33239350]
5. Strain T, Flaxman S, Guthold R, et al. National, regional, and global trends in insufficient physical activity among adults from 2000 to 2022: a pooled analysis of 507 population-based surveys with 5·7 million participants. Lancet Glob Health 2024 Aug;12(8):e1232-e1243. [doi: 10.1016/S2214-109X(24)00150-5] [Medline: 38942042]
6. Zhang M, Wang W, Li M, Sheng H, Zhai Y. Efficacy of mobile health applications to improve physical activity and sedentary behavior: a systematic review and meta-analysis for physically inactive individuals. Int J Environ Res Public Health 2022 Apr 18;19(8):4905. [doi: 10.3390/ijerph19084905] [Medline: 35457775]
7. Sohaib Aslam A, van Luenen S, Aslam S, van Bodegom D, Chavannes NH. A systematic review on the use of mHealth to increase physical activity in older people. Clinical eHealth 2020;3:31-39. [doi: 10.1016/j.ceh.2020.04.002]
8. Wang JW, Zhu Z, Shuling Z, et al. Effectiveness of mHealth app-based interventions for increasing physical activity and improving physical fitness in children and adolescents: systematic review and meta-analysis. JMIR Mhealth Uhealth 2024 Apr 30;12:e51478. [doi: 10.2196/51478] [Medline: 38687568]
9. mHealth: new horizons for health through mobile technologies: second global survey on eHealth. : WHO Global Observatory for eHealth; 2011 URL: https://iris.who.int/handle/10665/44607 [accessed 2025-05-01]
10. Kim J, Marcusson-Clavertz D, Yoshiuchi K, Smyth JM. Potential benefits of integrating ecological momentary assessment data into mHealth care systems. Biopsychosoc Med 2019;13:19. [doi: 10.1186/s13030-019-0160-5] [Medline: 31413726]
11. Direito A, Carraça E, Rawstorn J, Whittaker R, Maddison R. mHealth technologies to influence physical activity and sedentary behaviors: behavior change techniques, systematic review and meta-analysis of randomized controlled trials. Ann Behav Med 2017 Apr;51(2):226-239. [doi: 10.1007/s12160-016-9846-0] [Medline: 27757789]
12. Hardeman W, Houghton J, Lane K, Jones A, Naughton F. A systematic review of just-in-time adaptive interventions (JITAIs) to promote physical activity. Int J Behav Nutr Phys Act 2019 Apr 3;16(1):31. [doi: 10.1186/s12966-019-0792-7] [Medline: 30943983]
13. Nahum-Shani I, Smith SN, Spring BJ, et al. Just-in-time adaptive interventions (JITAIS) in mobile health: key components and design principles for ongoing health behavior support. Ann Behav Med 2018 May 18;52(6):446-462. [doi: 10.1007/s12160-016-9830-8] [Medline: 27663578]
14. van Dantzig S, Geleijnse G, van Halteren AT. Toward a persuasive mobile application to reduce sedentary behavior. Pers Ubiquit Comput 2013 Aug;17(6):1237-1246. [doi: 10.1007/s00779-012-0588-0]
15. Finkelstein J, Bedra M, Li X, Wood J, Ouyang P. Mobile app to reduce inactivity in sedentary overweight women. Stud Health Technol Inform 2015;216:89-92. [Medline: 26262016]
16. Bond DS, Thomas JG, Raynor HA, et al. B-MOBILE--a smartphone-based intervention to reduce sedentary time in overweight/obese individuals: a within-subjects experimental trial. PLoS One 2014;9(6):e100821. [doi: 10.1371/journal.pone.0100821] [Medline: 24964010]
17. Fiedler J, Seiferth C, Eckert T, Woll A, Wunsch K. A just-in-time adaptive intervention to enhance physical activity in the SMARTFAMILY2.0 trial. Sport Exerc Perform Psychol 2023;12(1):43-57. [doi: 10.1037/spy0000311]
18. Mair JL, Hayes LD, Campbell AK, Buchan DS, Easton C, Sculthorpe N. A personalized smartphone-delivered just-in-time adaptive intervention (JitaBug) to increase physical activity in older adults: mixed methods feasibility study. JMIR Form Res 2022 Apr 7;6(4):e34662. [doi: 10.2196/34662] [Medline: 35389348]

XSL•FO
RenderX

19. Singer JD, Willett JB. Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence: Oxford University Press; 2003.
20. Rabbi M, Pfammatter A, Zhang M, Spring B, Choudhury T. Automated personalized feedback for physical activity and dietary behavior change with mobile phones: a randomized controlled trial on adults. JMIR Mhealth Uhealth 2015 May 14;3(2):e42. [doi: 10.2196/mhealth.4160] [Medline: 25977197]
21. Kabir KS, Kenfield SA, Blarigan ELV, Chan JM, Wiese J. Ask the users: a case study of leveraging user-centered design for designing just-in-time adaptive interventions (JITAIs). Proc ACM Interact Mob, Wearable Ubiquitous Technol 2022. [doi: 10.1145/3534612]
22. Klasnja P, Smith S, Seewald NJ, et al. Efficacy of contextually tailored suggestions for physical activity: a micro-randomized optimization trial of HeartSteps. Ann Behav Med 2019 May 3;53(6):573-582. [doi: 10.1093/abm/kay067] [Medline: 30192907]
23. Balbim GM, Marques IG, Marquez DX, et al. Using Fitbit as an mHealth intervention tool to promote physical activity: potential challenges and solutions. JMIR Mhealth Uhealth 2021 Mar 1;9(3):e25289. [doi: 10.2196/25289] [Medline: 33646135]

Abbreviations

AIC: Akaike information criterion
API: application programming interface
JITAI: just-in-time adaptive intervention
mHealth: mobile health
MVPA: moderate to vigorous intensity physical activity
PIC: personalized intervention criteria
UIC: uniform intervention criteria
WHO: World Health Organization

Edited by C Jacob; submitted 22.09.24; peer-reviewed by J Rehder, L Allan; revised version received 06.04.25; accepted 07.04.25; published 26.05.25.

Please cite as:
Ikegaya M, Foo JC, Murata T, Oshima K, Kim J
Using Personalized Intervention Criteria in a Mobile Just-in-Time Adaptive Intervention for Increasing Physical Activity in University Students: Pilot Study
JMIR Hum Factors 2025;12:e66750
URL: https://humanfactors.jmir.org/2025/1/e66750
doi:10.2196/66750

© Mai Ikegaya, Jerome Clifford Foo, Taiga Murata, Kenta Oshima, Jinhyuk Kim. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 26.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Novel Just-in-Time Intervention for Promoting Safer Drinking Among College Students: App Testing Across 2 Independent Pre-Post Trials

Philip I Chow[1], PhD; Jessica Smith[1], MS; Ravjot Saini[2], MS; Christina Frederick[1], BSc; Connie Clark[2], BS; Maxwell Ritterband[2], BA; Jennifer P Halbert[1], MSPH; Kathryn Cheney[2], BA; Katharine E Daniel[1], PhD; Karen S Ingersoll[1], PhD

[1]School of Medicine, University of Virginia, 560 Ray C Hunt Dr, Charlottesville, VA, United States

[2]University of Virginia, Charlottesville, VA, United States

Corresponding Author:
Philip I Chow, PhD
School of Medicine, University of Virginia, 560 Ray C Hunt Dr, Charlottesville, VA, United States

Abstract

Background: Binge drinking, which is linked to various immediate and long-term negative outcomes, is highly prevalent among US college students. Behavioral interventions delivered via mobile phones have a strong potential to help decrease the hazardous effects of binge drinking by promoting safer drinking behaviors.

Objective: This study aims to evaluate the preliminary efficacy of bhoos, a novel smartphone app designed to promote safer drinking behaviors among US college students. The app offers on-demand educational content about safer alcohol use, provides dynamic feedback as users log their alcohol consumption, and includes an interactive drink tracker that estimates blood alcohol content in real time.

Methods: The bhoos app was tested in 2 independent pre-post studies each lasting 4 weeks, among US college students aged 18 - 35 years. The primary outcome in both trials was students' self-reported confidence in using protective behavioral strategies related to drinking, with self-reported frequency of alcohol consumption over the past month examined as a secondary outcome.

Results: In study 1, bhoos was associated with increased confidence in using protective behavioral strategies. Students also endorsed the high usability of the app and reported acceptable levels of engagement. Study 2 replicated findings of increased confidence in using protective behavioral strategies, and demonstrated a reduction in the self-reported frequency of alcohol consumption.

Conclusions: Bhoos is a personalized, accessible, and highly scalable digital intervention with a strong potential to effectively address alcohol-related behaviors on college campuses.

(JMIR Hum Factors 2025;12:e69873) doi:10.2196/69873

KEYWORDS

alcohol; college students; smartphone intervention; binge drinking; safe drinking

Introduction

Overview

Binge drinking, defined as consuming more than 5 standard drinks for men or 4 for women within a 2-hour period [1,2], is highly prevalent among US college students. This is concerning because binge drinking is linked to various immediate and long-term negative outcomes, including lower academic performance, a higher incidence of sexual assault, drunk driving, motor vehicle accidents, organ damage, and premature death [2-5]. Additionally, research shows that college students consume more alcohol than their noncollege-attending peers [6,7], which is troubling given that excessive drinking during late adolescence and early adulthood is a strong predictor of alcohol use disorders later in life [8]. Therefore, taking proactive steps to address binge drinking among college students is critical to reducing both the short- and long-term consequences in this population.

Alcohol Interventions in US Colleges

A common approach adopted by US colleges and universities to address hazardous drinking among students involves programs that focus on changing attitudes, increasing knowledge, and modifying behaviors related to alcohol use. Programs like the Brief Alcohol Screening and Intervention for College Students [9,10] typically include brief motivational counseling sessions aimed at reducing students' positive alcohol expectancies, increasing their awareness of the consequences of drinking, and enhancing their readiness to change [11-15]. Despite the

XSL•FO
RenderX

widespread implementation of such alcohol education and intervention programs, national data show that binge drinking rates among college students have remained relatively stable for over a decade [5], highlighting the ongoing difficulty in effectively addressing this public health issue.

Delivering Alcohol Interventions Through Smartphone Apps

Safer drinking interventions delivered through smartphones may offer distinct advantages in addressing binge drinking among young adults due to their ubiquity, high use, constant presence, and engaging features. Among all US demographic groups, college students have some of the highest rates of smartphone ownership and use [16]. Capitalizing on this widespread adoption provides an opportunity to develop personalized interventions that can effectively engage students and help regulate their alcohol consumption. Compared with in-person interventions, smartphone app-based alcohol interventions are highly accessible, cost-effective, scalable for large student populations, and can offer personalized feedback based on users' behaviors [17-19].

Researchers have begun exploring the use of smartphone apps to deliver alcohol interventions both in the general population [18-20] and among college students specifically [21-25]. Reviews generally support the feasibility and potential of app-based interventions for addressing outcomes of problematic drinking in college students, though results have been inconsistent across studies [21,23]. One potential reason for this inconsistency may be the lack of sufficient tailoring or personalization for US college students, given that incorporating strategies like gamification to personalize digital interventions may increase user engagement and improve outcomes [26,27]. For example, providing users with dynamic feedback about their drinking, such as their estimated blood alcohol level, may help them to better regulate their alcohol consumption during a night out with their friends to avoid negative consequences. Furthermore, providing users with graphical displays of their current and past alcohol consumption patterns may help them understand how their drinking has impacted other aspects of their lives, such as their academic performance and personal health.

Our team conducted a formative study to gain a detailed understanding of US students' preferences for interventions and their patterns of smartphone use [28]. Based on our findings and through conducting iterative user-centered design testing, our team developed bhoos (see below for a detailed description of the app and its functions). Similar to other app interventions, bhoos provides on-demand psychoeducational content about safer alcohol use. A novel feature of bhoos, however, is its just-in-time approach. Through an interactive drink tracking feature, students can log their consumption of alcohol in real time, enabling the app to provide dynamic feedback on recommended safe drinking behaviors based on estimated real-time blood alcohol content [28]. The primary goal of bhoos is therefore to promote safer drinking behaviors, in large part by increasing students' confidence in engaging in protective strategies, with the potential secondary outcome of reducing overall alcohol consumption.

The Current Studies

We investigated the feasibility and preliminary impact of bhoos in 2 pre-post studies conducted among students at a mid-Atlantic university in the US over a 4-week period. The hypothesis for study 1 was that bhoos would lead to an increase in students' confidence in using protective behavioral strategies (ie, behaviors that are used while drinking to reduce alcohol use or limit alcohol-related problems) from baseline to postintervention. A secondary hypothesis was that bhoos would lead to a reduction in self-reported alcohol consumption from baseline to postintervention. Study 1 also evaluated the impact of small monetary incentives on drinking outcomes and how students use the app to log their drinks. As study 1 was the first to test bhoos in college students, we examined students' engagement with the app and their ratings of its usability. Focus groups were conducted at the end of study 1 to gather feedback on the app from participants. Study 2 was conducted as a replication of self-reported drinking outcomes in study 1 to determine the potential of testing bhoos in a future randomized trial.

Ethical Considerations

Both studies reported in the manuscript were approved by the University of Virginia Institutional Review Board for Social and Behavioral Sciences (IRB-SBS #4020, 5334). All participants provided informed consent. Data are deidentified. Participants in study 1 received gift cards via the web as compensation for completing the baseline (US $20) and postassessments (US $45). Participants randomized to receive added incentives in study 1 received an additional sum, up to US $30, based on their completion of 3 engagement milestones: US $10 for downloading the app, US $10 for logging 1 drink or dry day, and US $10 for logging a "streak" of 3 consecutive days of drinks or dry days. Participants in study 2 received gift cards via the web as compensation for completing the baseline (US $20) and postassessments (US $25).

Study 1

Methods

Study Overview and Participants

Study 1 consisted of 3 phases. The first phase evaluated the bhoos app in a simple pre-post trial design. The second phase randomized (1:1) participants to receive small additional monetary incentives or not for using the bhoos app in a second pre-post trial. The third phase consisted of focus groups to capture impressions of the app by a subset of participants from the second phase.

For all phases of study 1, college students were eligible to participate if they met the following criteria: (1) aged 18 to 35 years, (2) current enrollment verified via a university email address, and (3) currently owned and used a smartphone. Students were enrolled using traditional, online, and social media methods. Recruitment spanned June through September 2021. Recruitment materials instructed applicants to complete a web-based interest form including contact information, demographic information, and questions to determine initial

XSL•FO
RenderX

study eligibility. Research coordinators verified student status and identity with the university's internal people search. Eligible verified applicants were invited to participate, signed a web-based consent form, and were enrolled in the study.

The phase 1 group included 83 participants (mean age 20.8 years, SD 1.6 years; 68% self-identified as female; 58% self-identified as White, 16% self-identified as Asian or Hawaiian/Pacific Islander, 4% self-identified as Black, 12% self-identified as multiracial, 4% prefer not to answer, and 6% chose to leave the item blank), and phase 2 group included 172 participants (mean age 20.1 years, SD 1.8 years; 60% self-identified as female; 48% self-identified as White, 19% Asian or Hawaiian/Pacific Islander, 8% Black, 11% multiracial, 1% prefer not to answer, and 13% chose to leave the item blank). Assessments occurred at baseline and postintervention, 28 days later. All participants received gift cards via the web as compensation for completing the baseline (US $20) and postassessments (US $45). Participants in the phase 2 group who were randomized to not receive added incentives (n=86) were only provided compensation for completing the baseline and postassessments. However, those randomized to receive added incentives (n=86) received an additional sum, up to US $30, based on their completion of 3 engagement milestones: US $10 for downloading the app, US $10 for logging 1 drink or dry day, and US $10 for logging a "streak" of 3 consecutive days of drinks or dry days (see bhoos description, below).

To improve the app for future trials, the third phase of the trial included a subset of participants (n=18) from the phase 2 group who were invited to be part of debriefing focus groups that occurred 9 - 12 months after the initial pre-post trial. A total of 5 focus groups were conducted and qualitative information was collected regarding participants' general impressions of the app, what features of the app they liked most/least, and the degree to which the app helped them manage their drinking.

Intervention: Bhoos

Bhoos (pronounced [booz]) is an app whose name is a play on the word "booze," slang for alcohol, and the nickname for University of Virginia students, "hoos." The app provides on-demand educational content about safer alcohol use and dynamic feedback to users as they log their alcohol consumption in real time. Users can log the type of alcoholic drink they are consuming and each drink entry is time-stamped. The app provides real-time estimates of the user's blood alcohol concentration (estimated blood alcohol concentration) based on their self-identified sex, weight, and number and type of alcoholic drinks logged. In-app notifications are pushed to users based on their estimated blood alcohol concentrations, including information about recommended actions to stay safe and avoid overdrinking. Users can view their current and past drink history through a built-in dashboard. To encourage engagement with the app, users can establish a streak of logins whenever they log consecutive drinking days or dry days (ie, days in which users logged no alcoholic drinks). To encourage engagement, the app also allows users to rate activity level or sleep quality as a secondary health behavior. Materials about safe drinking tips and staying healthy can be accessed directly from the main page. Users can also learn about health-related resources in and around the university through the app on demand. The app design and content were informed by formative research and a think-aloud process with the target population prior to this trial [28].

Measures

Baseline Alcohol Use Severity

The 10-item Alcohol Use Disorders Identification Test (AUDIT) is the most widely used self-report measure of unhealthy alcohol use [29]. Scores range from 0 to 40, with higher scores indicating more unhealthy alcohol use. Prior studies on college students have suggested cutoff scores for low-risk drinkers (<7), hazardous drinkers (8-15), and alcohol-dependent drinkers (>15). Participants completed the AUDIT at baseline (Cronbach α=0.78) to characterize their alcohol use over the past year.

Alcohol Consumption in the Past Month

We used a modified 3-item measure based on the Daily Drinking Questionnaire [30] to assess students' self-reported drinking behavior. Students were initially asked "How often did you drink during the last month?" with response options ranging from 1 (did not drink at all) to 7 (once a day or more). Responses were analyzed as a single-item measure of the average number of days per week in the last month involving alcohol consumption, with higher scores indicating more days per week that involved alcohol consumption. Students who responded that they had drank at least once a month (ie, a score of ≥2) were then asked 2 follow-up questions: "Think of a typical weekend evening (Friday or Saturday) during the last month. How many standard drinks did you drink on that evening?" and "Think of the occasion (any day of the week) you drank the most during the last month. How many standard drinks did you drink?" Students responded to each of these questions by typing a number into an open field. Because these questions were intended to assess for average and maximum alcohol consumption, respectively, we analyzed responses to them as separate items.

Protective Behavioral Strategies

A modified version of the 20-item Protective Behavioral Strategies Scale [31-33] was administered at baseline (α=0.92) and postintervention (α=0.93) to assess students' confidence in using protective behavioral strategies, or behaviors adopted while drinking to limit their alcohol consumption or limit alcohol-related problems. Specifically, the instructions were modified from the original version to ask students to select the response option that "best fits your confidence level" for using various protective strategies when drinking. Participants responded to each item on a modified scale from 1 (extremely not confident) to 7 (extremely confident), with scores ranging from 20 to 140 [34,35]. Higher scores indicate more confidence in using protective behavioral strategies such as stopping or limiting alcohol consumption, changing the speed or frequency of drinking, and reducing the risk of serious harm by using a designated driver, going home with a friend, or protecting one's drink from adulterants.

XSL•FO
RenderX

App Usability

The 10-item System Usability Scale (SUS) [36] was administered postintervention to assess the usability of the bhoos app. The SUS provides a global view of subjective assessments of usability. We customized the items to specifically mention bhoos. Sample items include "I think that I would like to use bhoos frequently" and "I thought bhoos was easy to use." Responses to each item range from 1 (strongly disagree) to 5 (strongly agree). Possible scores on the SUS range from 0 to 100, with a higher score indicating higher overall usability of a system or program. The SUS has been used in roughly 3500 surveys within 273 studies evaluating a range of systems, interfaces, and programs [37]. Internal consistency of the SUS was good (α=0.84).

Plan for Analyses

To determine whether students' self-reported drinking outcomes differed across the 2 phases, including the incentive conditions in phase 2, 1-way ANOVAs were conducted using the phase 1, phase 2 incentivized, and phase 2 nonincentivized groups as within participants variables and difference scores for the drinking outcomes as the dependent variables. Difference scores in drinking outcomes between the baseline and postintervention time points served as the dependent variables. Next, for each of the phase 1 and phase 2 groups, paired *t* tests were conducted to obtain estimates of short-term changes in drinking and attitudinal measures from baseline to follow-up. Box-Cox transformations were performed for the drinking outcomes where there was evidence of nonnormality of data based on skewness (outside of −1 and +1) or kurtosis (outside of −2 and +2) being outside of the conventional acceptable ranges. The pattern of findings did not change when using transformed or raw variables. Analyses of self-reported drinking outcomes were performed using IBM SPSS Statistics (version 26.0.0). App use is described using descriptive statistics. Qualitative feedback from focus groups was reviewed by members of the study team (KI, CF, NA, and CC), and recurring themes were identified.

Results

Baseline Alcohol Use Severity

To characterize the sample, we examined AUDIT scores for all participants in study 1. Of the 234 students who completed at least the baseline measures, 133 (59%) were classified as low-risk drinkers (AUDIT Score ≤7), 93 (40%) were classified as hazardous drinkers (AUDIT Score between 8 and 15), and 8 (3%) were classified as alcohol dependent drinkers (AUDIT score >15).

Evaluating the Impact of Study Phases on Self-Reported Outcomes

There was no significant effect of the study phase or incentives on any of the self-reported drinking outcomes, for the average number of days per week in the last month involving alcohol consumption ($F_{2, 232}$=0.294, *P*=.75, η^2=.003), typical weekend evening drink consumption in the last month ($F_{2, 165}$=0.662, *P*=.52, η^2=.008), the maximum number of drinks consumed in the last month ($F_{2, 175}$=0.005, *P*=.99, η^2=.00), or protective behavioral strategies ($F_{2, 232}$=1.469, *P*=.23, η^2=.013). Table 1 contains all results of self-reported drinking outcomes, separately for each phase and incentive condition. Below, we report self-reported drinking-related outcomes overall for those in phases 1 and 2 combined, followed separately for each phase and incentive group.

Table . Self-reported drinking outcomes for study 1.

	Average number of days per week drinking in the last month, mean (SD)				Average number drinks per weekend in the last month, mean (SD)				Max number of drinks on one occasion in the last month, mean (SD)				Protective behaviors, mean (SD)			
	Pre	Post	*t* test (*df*)	*P* value	Pre	Post	*t* test (*df*)	*P* value	Pre	Post	*t* test (*df*)	*P* value	Pre	Post	*t* test (*df*)	*P* value
Phase 1	3.66 (1.31)	3.71 (1.34)	0.54 (75)	.59	3.64 (1.68)	3.61 (1.75)	0.11 (60)	.91	5.68 (3.16)	5.94 (3.12)	0.79 (65)	.43	98.20 (21.01)	100.30 (20.18)	0.98 (75)	.33
Phase 2 (no incentive)	2.99 (1.35)	3.01 (1.36)	0.22 (74)	.83	3.21 (2.06)	2.89 (1.95)	1.35 (50)	.18	4.53 (3.36)	4.59 (3.65)	0.27 (53)	.81	106.19 (24.74)	111.91 (23.28)	2.29 (76)	.03
Phase 2 (incentive)	3.05 (1.52)	2.99 (1.27)	0.55 (81)	.59	3.76 (1.92)	3.93 (2.75)	0.12 (52)	.91	5.93 (3.05)	5.98 (4.48)	0.64 (54)	.52	106.14 (26.05)	108.52 (24.56)	0.89 (80)	.38

Protective Behavioral Strategies

Overall, there was a significant increase in students' confidence in using protective behavioral strategies in drinking from baseline ($Mean_{pre}$ 103.58, SD 24.26) to postintervention ($Mean_{post}$ 106.97, SD 23.20; t_{233}=2.393, *P*=.02), such that students reported more confidence engaging in more protective behaviors while drinking to limit alcohol-related problems from drinking from baseline to postintervention.

As seen in Table 1, despite an increase in students' self-reported confidence in using protective behavioral strategies from baseline to postintervention for all 3 groups, the effect was not significant for students in the phase 1 group (t_{75}=0.98, *P*=.33), and the phase 2 incentive group (t_{80}=0.89, *P*=.38). There was,

XSL•FO
RenderX

however, a significant increase in confidence in using protective behavioral strategies for the phase 2 no incentive group (t_{76}=2.29, P=.03).

Alcohol Consumption in the Past Month

Overall, there was no significant change to the average number of days per week in the last month involving alcohol consumption, from baseline ($Mean_{pre}$ 3.23, SD 1.43) to postintervention ($Mean_{post}$ 3.25, SD 1.39; t_{232}=0.07, P=.95). There were also no significant changes to students' average number of drinks per weekend in the last month from baseline ($Mean_{pre}$ 3.52, SD 1.90) to postintervention ($Mean_{post}$ 3.44, SD 2.20; t_{164}=0.85, P=.40), nor were there significant changes to students' maximum number of drinks consumed in the last month from baseline ($Mean_{pre}$ 5.43, SD 3.46) to postintervention ($Mean_{post}$ 5.51, SD 3.83; t_{173}=0.20, P=.84).

As seen in Table 1, there were no significant changes to students' self-reported drinking frequency in the past month from baseline to postintervention for students in the phase 1 group (drinking frequency in the past month, t_{75}=0.54, P=.59; the average number of drinks per weekend in the last month, t_{60}=0.11, P=.91; the maximum number of drinks consumed in the last month, t_{65}=0.79, P=.43). This was also the case for the phase 2 no incentive group (drinking frequency in the past month, t_{74}=0.22, P=.83; the average number of drinks per weekend in the last month, t_{50}=1.35, P=.18; the maximum amount of drinks consumed in the last month, t_{53}=.27, P=.81), and the phase 2 incentive group (drinking frequency in the past month, t_{81}=0.55, P=.59; average number of drinks per weekend in the last month, t_{52}=0.12, P=.91; the maximum amount of drinks consumed in the last month, t_{54}=0.64, P=.52).

Usability of Bhoos

There was no significant effect of the study phase or incentives on self-reported usability of bhoos ($F_{2,\ 234}$=0.097, P=.908, η^2=.001). Usability scores for students in phase 1 ranged from 22.5 to 100 (mean 71.22, SD 17.21). Scores for students in the phase 2 nonincentive group ranged from 35 to 100 (mean 71.75, SD 16.96), and 37.50 to 100 (mean 72.38, SD 15.69) for those in the phase 2 incentive group. The average usability score (mean 71.80, SD 16.54) placed bhoos in the third quartile of all programs evaluated by the SUS [37].

Engagement With Bhoos

Nearly all participants (93.33%, 238 of 255) in study 1 downloaded the app. The majority of participants (85.49%, 218 of 255) logged at least 1 drink or dry day. Roughly two-thirds (160 of 255) of the participants logged at least 1 alcoholic drink, whereas nearly 3-quarters (189 of 255) logged at least 1 dry day. Slightly more than half (145 of 255) logged at least 1 streak. On average, participants used the app for 12.67 days (SD 9.96 d) of the 28 days of the intervention period.

Overall, providing added incentives did not seem to impact how students used the app to log their drinks. Among those in the phase 2 group, students randomized to receive incentives logged, on average, 11.16 drinks (SD 8.92) while those randomized to receive no added incentives logged, on average, 13.04 drinks (SD 16.62), t_{130}=0.924, and P=.357. Moreover, students randomized to receive incentives logged, on average, 15.50 dry days (SD 9.39) while those randomized to receive no added incentives logged, on average, 15.90 dry days (SD 8.81), t_{169}=0.282, and P=.78.

Qualitative Feedback From Focus Groups

The feedback obtained from focus groups revealed several common themes. Students found the bhoos app to be visually appealing and user-friendly. They expressed that the drinking chart within the app raised their awareness of their drinking habits over time, and they appreciated the ability to review their drinking history within the app. However, some students felt that the psychoeducational content within the app was too static, and they suggested that incorporating videos would be beneficial. Additionally, many students expressed a desire to track a broader range of secondary behaviors, including mood and stress.

In response to this feedback from the focus groups, we made several iterative improvements to the bhoos app in preparation for the next pilot study (study 2). Specifically, we enhanced the psychoeducational content by adding 4 short videos on relevant topics, including information on helping an intoxicated friend, knowing the signs of alcohol overdose, learning about alcohol tolerance, and understanding standard drink amounts. We also included more information on safety tips while drinking. We refined the app's functionality for users to track their mood or stress to identify patterns related to drinking events. Last, known bugs and glitches that students reported while using the app were addressed, enhancing bhoos' overall performance and reliability.

Discussion

Collectively, the results from study 1 provide preliminary support for the feasibility and usability of bhoos among college students in the United States. While the students in this sample did not report a reduction in drinking frequency, findings indicate that they were more confident in their ability to engage in protective behaviors while drinking, potentially enhancing their safety. Both app engagement data and mixed methods results confirm the usability of bhoos, while providing added monetary incentives did not seem to impact any of the observed outcomes or how students used bhoos to track their drinking. The qualitative feedback from students was instrumental in identifying areas for app improvement to enhance usability and engagement. Building on these findings, study 2 was conducted to examine whether the impact of bhoos on drinking-related outcomes would be replicated in an independent sample of college students.

Study 2

Methods

Study Overview and Participants

Eligibility criteria were identical to study 1. College students were recruited over the 2022 Fall semester using advertisements in dining halls. Although the initial goal was to enroll up to 200

XSL•FO
RenderX

students, the recruitment timeline was hampered by the continuation of the COVID-19 pandemic and a tragic event that occurred on university grounds on November 13th, 2022 (a shooting on University grounds that resulted in the murders of 3 students). The latest baseline questionnaires were completed before the event and enrollment of new participants was halted after the event. Participants who completed baseline questionnaires were permitted to complete the study and complete the postintervention questionnaires. In total, 43 students were recruited (60% students self-identified as women; 24 students self-identified as White, 12 students self-identified as Asian or Native Hawaiian, 2 students self-identified as Black, and 5 students self-identified as multiracial). Once again, assessments were administered at baseline and postintervention, 28 days later. All participants received gift cards via the web as compensation for completing the baseline (US $20) and postassessments (US $25). An added incentivization group was not included in study 2.

Measures and Plan for Analyses

The same self-report measures were administered in study 2 as in study 1. Specifically, baseline alcohol use severity was assessed by the 10-item AUDIT [29] (α=0.82). Alcohol consumption in the past month was assessed at baseline and postintervention with the 3-item measure used in study 1. Each item was treated as a separate single-item scale. The modified 20-item Protective Behavioral Strategies Scale [31-33] was administered at baseline (α=0.90) and postintervention (α=0.91) to assess students' confidence in using protective behavioral strategies. Finally, the usability of the bhoos app was assessed postintervention using the SUS [36] (α=0.75).

Paired *t* tests were used to investigate short-term changes in drinking and protective behaviors from baseline to follow-up. Due to the limited sample size, if the normality assumption was not met some analyses were computed again with the paired Wilcoxon Signed Rank Test. This did not change the interpretation of any of the results]. Analyses were performed using R (version 4.3.3; R Core Team). App use is described using descriptive statistics.

Results

Baseline Alcohol Use Severity

Of the 43 students, 25 (58%) were classified as low-risk drinkers (AUDIT Score ≤7), 11 (26%) were classified as hazardous drinkers (AUDIT Score between 8 and 15), and 6 (14%) were classified as alcohol dependent drinkers (AUDIT score >15).

Protective Behavioral Strategies

As expected, there was a significant increase in students' confidence in using protective behavioral strategies after using bhoos from baseline ($Mean_{pre}$ 107, SD 17.19) to postintervention ($Mean_{post}$ 113.37, SD 15.87; t_{31}=-2.79, *P*=.01), such that students reported confidence in engaging in more protective behaviors while drinking to limit alcohol-related problems from drinking after the intervention compared with before.

Alcohol Consumption in the Past Month

There was a significant decrease in the average number of days per week drinking alcohol in the last month from baseline ($Mean_{pre}$ 2.32, SD 1.27) to postintervention ($Mean_{post}$ 2.06, SD 1.15; t_{33}=2.32, *P*=.031). Among students who reported drinking, there was no significant change to their typical weekend evening drink consumption from baseline ($Mean_{pre}$ 3.83, SD 2.80) to postintervention ($Mean_{post}$ 3.87, SD 2.30; t_{29}=0, *P*=.92), nor was there a significant change to the maximum number of drinks they consumed in a single occasion the last month, from baseline ($Mean_{pre}$ 5.17, SD 4.07) to postintervention ($Mean_{post}$ 5.34, SD 4.50; t_{28}=0.24, *P*=.82).

Usability of Bhoos

Usability scores ranged from 55 to 98 (out of a possible 100). Overall, students reported higher usability of the revised version of bhoos (mean 77.93, SD 11.92) than in the original version tested in study 1. The average score placed bhoos in the upper 75% (fourth quartile) of all programs evaluated by the SUS [37].

Discussion

Collectively, the findings from study 2 reproduce the drinking-related results from study 1. Specifically, the findings of study 2 replicate the finding that bhoos is associated with an increase in confidence in using protective behaviors while drinking. It is worth noting that students in study 2 also reported reduced frequency of drinking, which may have resulted from improvements to the app from study 1, which is supported by our finding that students rated the app as more usable in study 2 than in study 1.

Overall Discussion

Principal Findings

Developing effective, accessible, and scalable interventions to address excessive drinking among college students is critical due to the significant public health concerns posed by alcohol misuse on college campuses. In 2 studies, we observed promising and consistent results regarding the impact and usability of the bhoos app, which effectively promoted confidence in using protective behaviors related to drinking among US college students.

The results suggest that bhoos' features, such as drink tracking and dynamic feedback, help to encourage responsible drinking behaviors, including monitoring alcohol intake, planning for safe transportation after drinking, or seeking help when needed. The ability to provide personalized, real-time feedback increases the relevance of the app to the individual needs and experiences of its users. By offering recommendations based on the user's drinking patterns and risk factors, bhoos may engage users more effectively than a generic one-size-fits-all program which enhances the likelihood of behavior change. Moreover, this more personalized approach to providing risk feedback aligns with the principles of Motivational Interviewing [38,39] and follows proven strategies for promoting behavior change [40-42]. Smartphone app-delivered interventions like bhoos offer several advantages over traditional approaches, including

XSL•FO
RenderX

the ability to deliver actionable feedback at critical moments. Additionally, they provide the potential for widespread accessibility and scalability. College students, in particular, maybe the most receptive to mobile apps for obtaining information and support for health-related decisions, given their high smartphone use [16]. The ease of dissemination and the low cost of smartphone apps make them a practical option for broad implementation on college campuses.

Limitations and Future Directions

The findings from our studies should be considered in light of several limitations. First, the reliance on infrequent and static measures of drinking behaviors introduces the possibility of recall bias. Moreover, the primary outcomes of both studies reflected students' confidence in using various protective behaviors while drinking and not the actual behaviors they engaged in. Future research could benefit from incorporating more ecologically valid approaches, such as ecological momentary assessment, perhaps administered within the app, to assess for actual protective behaviors related to drinking while limiting recall bias. Additionally, the pre-post design of our study limits our ability to evaluate the efficacy of bhoos, which could be better assessed through a randomized controlled trial. The short time frame between baseline and postassessment (4 weeks) restricts our understanding of the long-term sustainability of the app's effects. Future studies should explore the app's impact over longer periods, such as a full college semester. Moreover, a large percentage of students in both studies were White and female which may limit the generalizability of our findings to minorities and men. Finally, the tragic shooting that occurred on university grounds may have influenced students' drinking behaviors in study 2, although the extent of this impact is unclear. Future studies should examine how such events, and events more broadly (eg, political elections, performance of University athletic teams), affect student drinking behaviors.

General Conclusion

In summary, preliminary findings across 2 studies indicate that bhoos shows strong potential as a tool for promoting protective behaviors related to alcohol use among college students. While these studies help advance highly accessible and scalable approaches to address the consequences of binge drinking by US college students, further research is needed to fully understand the app's role in comprehensive alcohol harm reduction efforts within this population. Its personalized approach, combined with its accessibility and scalability, suggests that bhoos could be a valuable addition to the range of interventions available to address alcohol-related issues on campuses.

Acknowledgments

The authors thank the Board of Visitors and the President of the University of Virginia for funding this project through the Strategic Investment Fund. The authors also appreciate the support and guidance of the Executive Committee for the project "Reducing the Burden of Addictions in Virginia": Dr Robert Klesges, Dr Lee Ritterband, Dr Nassima Ait-Daoud Tiouririne, Dr Karen Ingersoll, Dr Christopher Holstege, and Dr Margaret Shupnik. Finally, the authors thank the undergraduate research assistants for this study: Grace Biegger, Aasritha Natarajan, David Vann, and Shruti Vachhani. We would also like to thank our development team, Gabe Heath, Nicole Le, and Brendan Murphey, for the design and development of the bhoos app intervention.

Conflicts of Interest

None delcared.

References

1. Goldman MS. Commentary on White, Kraus, and Swartzwelder (2006): "Many college freshmen drink at levels far beyond the binge threshold". Alcohol Clin Exp Res 2006 Jun;30(6):919-921. [doi: 10.1111/j.1530-0277.2006.00123.x] [Medline: 16737448]
2. White AM, Kraus CL, Swartzwelder HS. Many college freshmen drink at levels far beyond the binge threshold. Alcohol Clin Exp Res 2006 Jun;30(6):1006-1010. [doi: 10.1111/j.1530-0277.2006.00122.x] [Medline: 16737459]
3. Chassin L, Pitts SC, Prost J. Binge drinking trajectories from adolescence to emerging adulthood in a high-risk sample: predictors and substance abuse outcomes. J Consult Clin Psychol 2002 Feb;70(1):67-78. [Medline: 11860058]
4. Howland J, Rohsenow DJ, Greece JA, et al. The effects of binge drinking on college students' next-day academic test-taking performance and mood state. Addiction 2010 Apr;105(4):655-665. [doi: 10.1111/j.1360-0443.2009.02880.x] [Medline: 20403018]
5. Hingson R, Zha W, Smyth D. Magnitude and trends in heavy episodic drinking, alcohol-impaired driving, and alcohol-related mortality and overdose hospitalizations among emerging adults of college ages 18-24 in the United States, 1998-2014. J Stud Alcohol Drugs 2017 Jul;78(4):540-548. [doi: 10.15288/jsad.2017.78.540] [Medline: 28728636]
6. Blanco C, Okuda M, Wright C, et al. Mental health of college students and their non-college-attending peers: results from the national epidemiologic study on alcohol and related conditions. Arch Gen Psychiatry 2008 Dec;65(12):1429-1437. [doi: 10.1001/archpsyc.65.12.1429] [Medline: 19047530]
7. Slutske WS, Hunt-Carter EE, Nabors-Oberg RE, et al. Do college students drink more than their non-college-attending peers? evidence from a population-based longitudinal female twin study. J Abnorm Psychol 2004 Nov;113(4):530-540. [doi: 10.1037/0021-843X.113.4.530] [Medline: 15535786]

XSL•FO
RenderX

8. Huurre T, Lintonen T, Kaprio J, Pelkonen M, Marttunen M, Aro H. Adolescent risk factors for excessive alcohol use at age 32 years. a 16-year prospective follow-up study. Soc Psychiatry Psychiatr Epidemiol 2010 Jan;45(1):125-134. [doi: 10.1007/s00127-009-0048-y] [Medline: 19363578]
9. Dimeff LA, Baer JS, Kivlahan DR, Marlatt G. Brief Alcohol Screening and Intervention for College Students (BASICS): Guilford Press; 1999.
10. Fachini A, Aliane PP, Martinez EZ, Furtado EF. Efficacy of brief alcohol screening intervention for college students (BASICS): a meta-analysis of randomized controlled trials. Subst Abuse Treat Prev Policy 2012 Sep 12;7:40. [doi: 10.1186/1747-597X-7-40] [Medline: 22967716]
11. Dunn ME, Fried-Somerstein A, Flori JN, Hall TV, Dvorak RD. Reducing alcohol use in mandated college students: a comparison of a brief motivational intervention (BMI) and the expectancy challenge alcohol literacy curriculum (ECALC). Exp Clin Psychopharmacol 2020 Feb;28(1):87-98. [doi: 10.1037/pha0000290] [Medline: 31008641]
12. Gesualdo C, Pinquart M. Expectancy challenge interventions to reduce alcohol consumption among high school and college students: a meta-analysis. Psychol Addict Behav 2021 Nov;35(7):817-828. [doi: 10.1037/adb0000732] [Medline: 33856837]
13. Lewis BA, O'Neill HK. Alcohol expectancies and social deficits relating to problem drinking among college students. Addict Behav 2000;25(2):295-299. [doi: 10.1016/s0306-4603(99)00063-5] [Medline: 10795955]
14. Patrick ME, Maggs JL. Does drinking lead to sex? daily alcohol-sex behaviors and expectancies among college students. Psychol Addict Behav 2009 Sep;23(3):472-481. [doi: 10.1037/a0016097] [Medline: 19769431]
15. Rhew IC, Duckworth JC, Lee CM. The association between intended drinking contexts and alcohol expectancies in college students: a daily diary study. Addict Behav 2021 Sep;120:106967. [doi: 10.1016/j.addbeh.2021.106967] [Medline: 33971498]
16. Internet, broadband fact sheet. Pew Research Center. URL: https://www.pewresearch.org/internet/fact-sheet/internet-broadband/?tabItem=d40cde3f-c455-4f0e-9be0-0aefcdaeee00 [accessed 2024-09-17]
17. Nahum-Shani I, Smith SN, Spring BJ, et al. Just-in-time adaptive interventions (JITAIs) in mobile health: key components and design principles for ongoing health behavior support. Ann Behav Med 2018 May 18;52(6):446-462. [doi: 10.1007/s12160-016-9830-8] [Medline: 27663578]
18. Tofighi B, Chemi C, Ruiz-Valcarcel J, Hein P, Hu L. Smartphone apps targeting alcohol and illicit substance use: systematic search in in commercial app stores and critical content analysis. JMIR Mhealth Uhealth 2019 Apr 22;7(4):e11831. [doi: 10.2196/11831] [Medline: 31008713]
19. Colbert S, Thornton L, Richmond R. Smartphone apps for managing alcohol consumption: a literature review. Addict Sci Clin Pract 2020 May 7;15(1):17. [doi: 10.1186/s13722-020-00190-x] [Medline: 32381062]
20. Hoeppner BB, Schick MR, Kelly LM, Hoeppner SS, Bergman B, Kelly JF. There is an app for that - or is there? a content analysis of publicly available smartphone apps for managing alcohol use. J Subst Abuse Treat 2017 Nov;82:67-73. [doi: 10.1016/j.jsat.2017.09.006] [Medline: 29021117]
21. Hutton A, Prichard I, Whitehead D, et al. mHealth interventions to reduce alcohol use in young people: a systematic review of the literature. Compr Child Adolesc Nurs 2020 Sep;43(3):171-202. [doi: 10.1080/24694193.2019.1616008] [Medline: 31192698]
22. Kazemi DM, Borsari B, Levine MJ, et al. Effectiveness of a theory-based mHealth intervention for high-risk drinking in college students. Subst Use Misuse 2020;55(10):1667-1676. [doi: 10.1080/10826084.2020.1756851] [Medline: 32394772]
23. Berman AH, Gajecki M, Sinadinovic K, Andersson C. Mobile interventions targeting risky drinking among university students: a review. Curr Addict Rep 2016;3(2):166-174. [doi: 10.1007/s40429-016-0099-6] [Medline: 27226948]
24. Riordan BC, Winter T, Carey KB, et al. A combined web based intervention and ecological momentary intervention for reducing alcohol use among incoming first-year university students: results from a three-arm randomised controlled trial. Addict Behav 2023 Jan;136:107471. [doi: 10.1016/j.addbeh.2022.107471] [Medline: 36081248]
25. Bertholet N, Schmutz E, Studer J, et al. Effect of a smartphone intervention as a secondary prevention for use among university students with unhealthy alcohol use: randomised controlled trial. BMJ 2023 Aug 16;382:e073713. [doi: 10.1136/bmj-2022-073713] [Medline: 37586742]
26. Floryan MR, Ritterband LM, Chow PI. Principles of gamification for Internet interventions. Transl Behav Med 2019 Nov 25;9(6):1131-1138. [doi: 10.1093/tbm/ibz041] [Medline: 30944933]
27. Floryan M, Chow PI, Schueller SM, Ritterband LM. The model of gamification principles for digital health interventions: evaluation of validity and potential utility. J Med Internet Res 2020 Jun 10;22(6):e16506. [doi: 10.2196/16506] [Medline: 32519965]
28. Smith JG, Alamiri NS, Biegger G, Frederick C, Halbert JP, Ingersoll KS. Think-aloud testing of a novel safer drinking app for college students during COVID-19: usability study. JMIR Form Res 2022 Feb 17;6(2):e32716. [doi: 10.2196/32716] [Medline: 35175210]
29. Babor TF, de la Fuente JR, Saunders J, Grant M. AUDIT, the alcohol use disorders identification test: guidelines for use in primary health care: World Health Organization; 1992:1-29.
30. Collins RL, Parks GA, Marlatt GA. Social determinants of alcohol consumption: the effects of social interaction and model status on the self-administration of alcohol. J Consult Clin Psychol 1985 Apr;53(2):189-200. [doi: 10.1037//0022-006x.53.2.189] [Medline: 3998247]

XSL•FO
RenderX

31. DeMartini KS, Palmer RS, Leeman RF, et al. Drinking less and drinking smarter: direct and indirect protective strategies in young adults. Psychol Addict Behav 2013 Sep;27(3):615-626. [doi: 10.1037/a0030475] [Medline: 23088406]
32. Martens MP, Ferrier AG, Cimini MD. Do protective behavioral strategies mediate the relationship between drinking motives and alcohol use in college students? J Stud Alcohol Drugs 2007 Jan;68(1):106-114. [doi: 10.15288/jsad.2007.68.106] [Medline: 17149524]
33. Sugarman DE, Carey KB. The relationship between drinking control strategies and college student alcohol use. Psychol Addict Behav 2007 Sep;21(3):338-345. [doi: 10.1037/0893-164X.21.3.338] [Medline: 17874884]
34. Madson MB, Arnau RC, Lambert SJ. Development and psychometric evaluation of the revised protective behavioral strategies scale. Psychol Assess 2013 Jun;25(2):556-567. [doi: 10.1037/a0031788] [Medline: 23397929]
35. Treloar H, Martens MP, McCarthy DM. The protective behavioral strategies scale-20: improved content validity of the serious harm reduction subscale. Psychol Assess 2015 Mar;27(1):340-346. [doi: 10.1037/pas0000071] [Medline: 25558969]
36. Sauro J. A Practical Guide to the System Usability Scale: Background, Benchmarks & Best Practices. Denver, Colarado: Measuring Usability LLC; 2011.
37. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009 May 1;4(3):114-123 [FREE Full text]
38. Hettema J, Steele J, Miller WR. Motivational interviewing. Annu Rev Clin Psychol 2005;1(1):91-111. [doi: 10.1146/annurev.clinpsy.1.102803.143833] [Medline: 17716083]
39. Apodaca TR, Longabaugh R. Mechanisms of change in motivational interviewing: a review and preliminary evaluation of the evidence. Addiction 2009 May;104(5):705-715. [doi: 10.1111/j.1360-0443.2009.02527.x] [Medline: 19413785]
40. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol 2008 May;27(3):379-387. [doi: 10.1037/0278-6133.27.3.379] [Medline: 18624603]
41. Bensley RJ, Mercer N, Brusk JJ, et al. The eHealth behavior management model: a stage-based approach to behavior change and management. Prev Chronic Dis 2004 Oct;1(4):A14. [Medline: 15670446]
42. Cugelman B. Gamification: what it is and why it matters to digital health behavior change developers. JMIR Serious Games 2013 Dec 12;1(1):e3. [doi: 10.2196/games.3139] [Medline: 25658754]

Abbreviations

AUDIT: Alcohol Use Disorders Identification Test
SUS: System Usability Scale

Edited by A Kushniruk; submitted 10.12.24; peer-reviewed by RS Mohn, R Mertz; revised version received 18.02.25; accepted 04.03.25; published 10.04.25.

Please cite as:
Chow PI, Smith J, Saini R, Frederick C, Clark C, Ritterband M, Halbert JP, Cheney K, Daniel KE, Ingersoll KS
A Novel Just-in-Time Intervention for Promoting Safer Drinking Among College Students: App Testing Across 2 Independent Pre-Post Trials
JMIR Hum Factors 2025;12:e69873
URL: https://humanfactors.jmir.org/2025/1/e69873
doi:10.2196/69873

© Philip I Chow, Jessica Smith, Ravjot Saini, Christina Frederick, Connie Clark, Maxwell Ritterband, Jennifer P Halbert, Kathryn Cheney, Katharine E Daniel, Karen S Ingersoll. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 10.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Comparison of an AI Chatbot With a Nurse Hotline in Reducing Anxiety and Depression Levels in the General Population: Pilot Randomized Controlled Trial

Chen Chen[1], MSc; Kok Tai Lam[1], MPhil; Ka Man Yip[1], MStat; Hung Kwan So[1], PhD; Terry Yat Sang Lum[2], PhD; Ian Chi Kei Wong[3,4,5], PhD; Jason C Yam[6], MD; Celine Sze Ling Chui[7], PhD; Patrick Ip[1], MD

[1]Department of Paediatrics and Adolescent Medicine, School of Clinical Medicine, Li Ka Shing Faculty of Medicine, University of Hong Kong, Room 115, 1/F, New Clinical Building, Queen Mary Hospital, 102 Pokfulam Road, Hong Kong SAR, China (Hong Kong)

[2]Department of Social Work Administration, University of Hong Kong, Hong Kong SAR, China (Hong Kong)

[3]Department of Pharmacology and Pharmacy, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong SAR, China (Hong Kong)

[4]School of Pharmacy, Medical Sciences Division, Macau University of Science and Technology, Macau, China

[5]School of Pharmacy, Aston University, Birmingham, United Kingdom

[6]Department of Ophthalmology and Visual Sciences, Chinese University of Hong Kong, Hong Kong SAR, China (Hong Kong)

[7]School of Nursing, University of Hong Kong, Hong Kong SAR, China (Hong Kong)

Corresponding Author:
Patrick Ip, MD
Department of Paediatrics and Adolescent Medicine, School of Clinical Medicine, Li Ka Shing Faculty of Medicine, University of Hong Kong, Room 115, 1/F, New Clinical Building, Queen Mary Hospital, 102 Pokfulam Road, Hong Kong SAR, China (Hong Kong)

Abstract

Background: Artificial intelligence (AI) chatbots have been customized to deliver on-demand support for people with mental health problems. However, the effectiveness of AI chatbots in tackling mental health problems among the general public in Hong Kong remains unclear.

Objective: This study aimed to develop a local AI chatbot and compare the effectiveness of the AI chatbot with a conventional nurse hotline in reducing the level of anxiety and depression among individuals in Hong Kong.

Methods: This study was a pilot randomized controlled trial conducted from October 2022 to March 2023, involving 124 participants allocated randomly (1:1 ratio) into the AI chatbot and nurse hotline groups. Among these, 62 participants in the AI chatbot group and 41 in the nurse hotline group completed both the pre- and postquestionnaires, including the GAD-7 (Generalized Anxiety Disorder Scale-7), PHQ-9 (Patient Health Questionnaire-9), and satisfaction questionnaire. Comparisons were conducted using independent and paired sample *t* tests (2-tailed) and the χ^2 test to analyze changes in anxiety and depression levels.

Results: Compared to the mean baseline score of 5.13 (SD 4.623), the mean postdepression score in the chatbot group was 3.68 (SD 4.397), which was significantly lower (*P*=.008). Similarly, a reduced anxiety score was also observed after the chatbot test (pre vs post: mean 4.74, SD 4.742 vs mean 3.4, SD 3.748; *P*=.005), respectively. No significant differences were found in the pre-post scores for either depression (*P*=.38) or anxiety (*P*=.19). No statistically significant difference was observed in service satisfaction between the two platforms (*P*=.32).

Conclusions: The AI chatbot was comparable to the traditional nurse hotline in alleviating participants' anxiety and depression after responding to inquiries. Moreover, the AI chatbot has shown potential in alleviating short-term anxiety and depression compared to the nurse hotline. While the AI chatbot presents a promising solution for offering accessible strategies to the public, more extensive randomized controlled studies are necessary to further validate its effectiveness.

Trial Registration: ClinicalTrials.gov NCT06621134; https://clinicaltrials.gov/study/NCT06621134

(JMIR Hum Factors 2025;12:e65785) doi:10.2196/65785

KEYWORDS

AI chatbot; anxiety; depression; effectiveness; artificial intelligence

XSL•FO
RenderX

Introduction

An estimated 1 in 7 people in Hong Kong experience a common mental disorder in their lifetime, with anxiety and depression being the two most common mental health issues in Hong Kong, accounting for 19% and 14% among the generation population, respectively [1,2]. Nearly three-quarters of individuals experiencing a mental health disorder do not seek professional help for various reasons such as high cost of services, fear of judgment, or stigma [3]. To address these potential problems, artificial intelligence (AI) chatbots have been incorporated into digital interventions, particularly web- and smartphone-based apps, to enhance user experience and optimize individualized mental health [4,5]. Chatbots can provide a safe and confidential space for people to seek help for their mental health concerns. Additionally, they are available at any time throughout a study, making them a convenient option for those who may lack access to traditional mental health services or require support beyond regular business hours. Further efforts toward strengthening implementation of chatbots are needed, and their application to the general public should be explored to improve perceptions of general mental health and increase awareness of the importance of premedical interventions in the future [6].

Chatbot models have been developed in various countries and are widely used in psychology and clinical practice. For example, Tess, a psychological AI service, has been tailored to provide immediate assistance to caregiving specialists, individuals receiving care, and family caregivers within a nonprofit organization in the United States and Canada [7]. Among university students, chatbot applications have showed a significant decrease in anxiety symptoms [8]. Anna, an AI-driven chatbot, simulates the function of a therapist and offers gamified versions of evidence-based activities drawn from various therapeutic approaches. Participants who engaged in activities facilitated by Anna provided more detailed responses containing a higher number of positive relational terms [9]. A review of users' opinions, satisfaction, and attitudes regarding depression and anxiety chatbot apps found that users felt supported and confident while they used apps that were easy to navigate, affordable, and free of cost [10]. In Hong Kong, the first AI-driven Cantonese psychological support tool, the Pai.ACT mobile app, was developed for parents of children with special education needs [11]. However, there is currently a lack of evidence on the effectiveness of AI chatbots in tackling mental health problems in the general public.

Given the risk of anxiety and depression in Hong Kong and the limited availability of mental health services during the epidemic, this study aimed to compare the effectiveness of the developed AI chatbot with that of a conventional nurse hotline in alleviating psychological anxiety and depression in the general public. Additionally, this study also aimed to understand user satisfaction with chatbot usage and gather preliminary data for larger randomized controlled trials to advocate for the AI chatbot as a psychological support tool for individuals seeking information on the chatbot platform.

Methods

Recruitment

This study was designed as a randomized controlled trial (RCT) with two parallel groups. Participants were parents recruited through two school principal's networks. The inclusion criteria included (1) being able to use smartphones proficiently to interact with the AI chatbot, (2) being fluent in Chinese language used in the study to effectively communicate with the AI chatbot, and (3) providing an electronic consent form. Parents who were unwilling or unable to commit to the entire research process and had difficulties in reading and understanding Chinese were excluded. Invitation letters were sent to kindergarten and primary school principal groups from chief schools, and 124 parents responded to this study and provided an electronic consent form. Participants were then randomly allocated in a 1:1 ratio to either the AI chatbot or the nurse hotline group using block randomization with blocks of four. Participants immediately filled out the posttest questionnaires after communicating with the AI chatbot or nurse. Those who did not respond immediately received reminders until they completed the posttest questionnaires. Participants with risky mental health problems were followed up by phone calls and were encouraged to ask to connect with necessary medical services and mental health clinical departments.

Ethical Considerations

This study was approved by the Institutional Review Board of the University of Hong Kong/Hospital Authority Hong Kong West (approval number: UW21-344). Informed consent was obtained from all participants in the study, ensuring that they were fully aware of the nature and possible consequences of their participation. Participants were informed that they had the right to opt out at any time without penalty. The original consent and approval covers secondary analysis without additional consent. Data collected from participants were deindentified to ensure privacy and confidentiality. Identifying information, including names, initials, or hospital numbers, was removed from all datasets to protect participant identities. Participation was voluntary, and individuals were informed that no financial or material incentives would be provided. Consolidated Standards of Reporting Trials (CONSORT) reporting guideline was adhered to in this study [12].

AI Chatbot Development Procedures

The development of the AI chatbot involved 7 key steps. First, the developer analyzed requirements by understanding common queries and focusing on providing accurate information to the public and assisting nurses. Relevant data was then collected from credible sources, and natural language processing (NLP) algorithms were used to train a robust model on diverse questions and answers. By integrating large language models, the chatbot could provide detailed responses. Next, the developer designed and developed the backend infrastructure, including databases, application programming interfaces, and integration with University of Hong Kong's website and WhatsApp. Extensive testing and validation ensured the chatbot's accuracy and user-friendliness, with a beta version tested by selected users. Finally, the chatbot was deployed on University of Hong

XSL•FO
RenderX

Kong's website and made accessible through WhatsApp, with a nurse hotline version developed to assist health care professionals in handling public inquiries.

The AI chatbot system comprises several essential components to provide accurate and relevant information. It begins with context identification, where the chatbot analyzes keywords and phrases to understand the user's query, ensuring accurate interpretation and relevant response. Next, the system conducts query analysis using NLP techniques to determine the user's intent and extract specific details such as dates, locations, and vaccination information. The query response system then leverages the trained NLP model and large language model–generated responses to provide accurate and contextually appropriate answers tailored to users' needs. Finally, the chatbot relies on a regularly updated and comprehensive dataset containing the latest information on COVID-19 vaccinations, government regulations, and health care procedures, ensuring the delivery of reliable and accurate responses.

Questionnaires

The Patient Health Questionnaire-9 (PHQ-9) is a self-assessment tool consisting of 9 questions that measure the frequency and severity of depressive symptoms over the last 2 weeks. Each question is derived from the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* (*DSM-IV*) criteria and is rated on a scale of 0 (not at all) to 3 (almost daily) [13]. The Generalized Anxiety Disorder Scale-7 (GAD-7) is a 7-item self-report scale that evaluates the frequency and severity of anxious thoughts and behaviors during the last 2 weeks. Items are based on the diagnostic criteria of the *DSM-IV* and scored from 0 (not at all) to 3 (nearly every day) [14]. The service satisfaction survey is a 3-item questionnaire for participants to report their degree of satisfaction after using the two platforms. The rating score ranges from 0 to 10, representing an increasing satisfaction from not at all likely to extremely likely. Based on the standardized cutoff that determines the severity of anxiety and depression, participants were divided into two groups: the no-risk group (total score<4) and the risk group (total score>4) [15].

Statistical Analysis

Quantitative variables were tested for normal distribution using the Kolmogorov-Smirnov test. Normally distributed continuous variables were expressed as mean (SD) and categorical variables were expressed as numbers and percentages. A per-protocol analysis was performed for all outcomes when comparing the 2 groups. The independent *t* test (2-tailed) was employed to compare the preanxiety or predepression, postanxiety or postdepression, pre-post difference scores, and service satisfaction between the AI chatbot and nurse hotline groups, respectively. The paired *t* test (2-tailed) was used to compare scores before and after group chat. Linear regression was conducted to analyze the difference in posttest scores between the AI chatbot and nurse hotline groups adjusted by pretest scores. Categorical variables were compared with the χ^2 test. All analyses were conducted using SPSS software (version 29.0; IBM Corp).

Results

From October 2022 to March 2023, all the participants (N=124) answered prequestionnaires, and 62 (62/62, 100%) participants in the AI chatbot group and 41 (41/62, 66.1%) in the nurse hotline group completed both the pre- and postquestionnaires (Figure 1). Comparisons within groups before and after the test are displayed in Table 1 and Figure 2. Compared to the mean baseline score of 5.13 (SD 4.623), the mean postdepression score 3.68 (SD 4.397) was significantly lower in the chatbot group (*P*=.008). Similarly, a reduction in anxiety score was also observed after using the chatbot test (pre vs post: mean 4.74, SD 4.742 vs mean 3.4, SD 3.748; *P*=.005).

XSL•FO
RenderX

Figure 1. Diagram of the study design. AI: artificial intelligence. PHQ-9: Patient Health Questionnaire-9; GAD-7: Generalized Anxiety Disorder Scale-7.

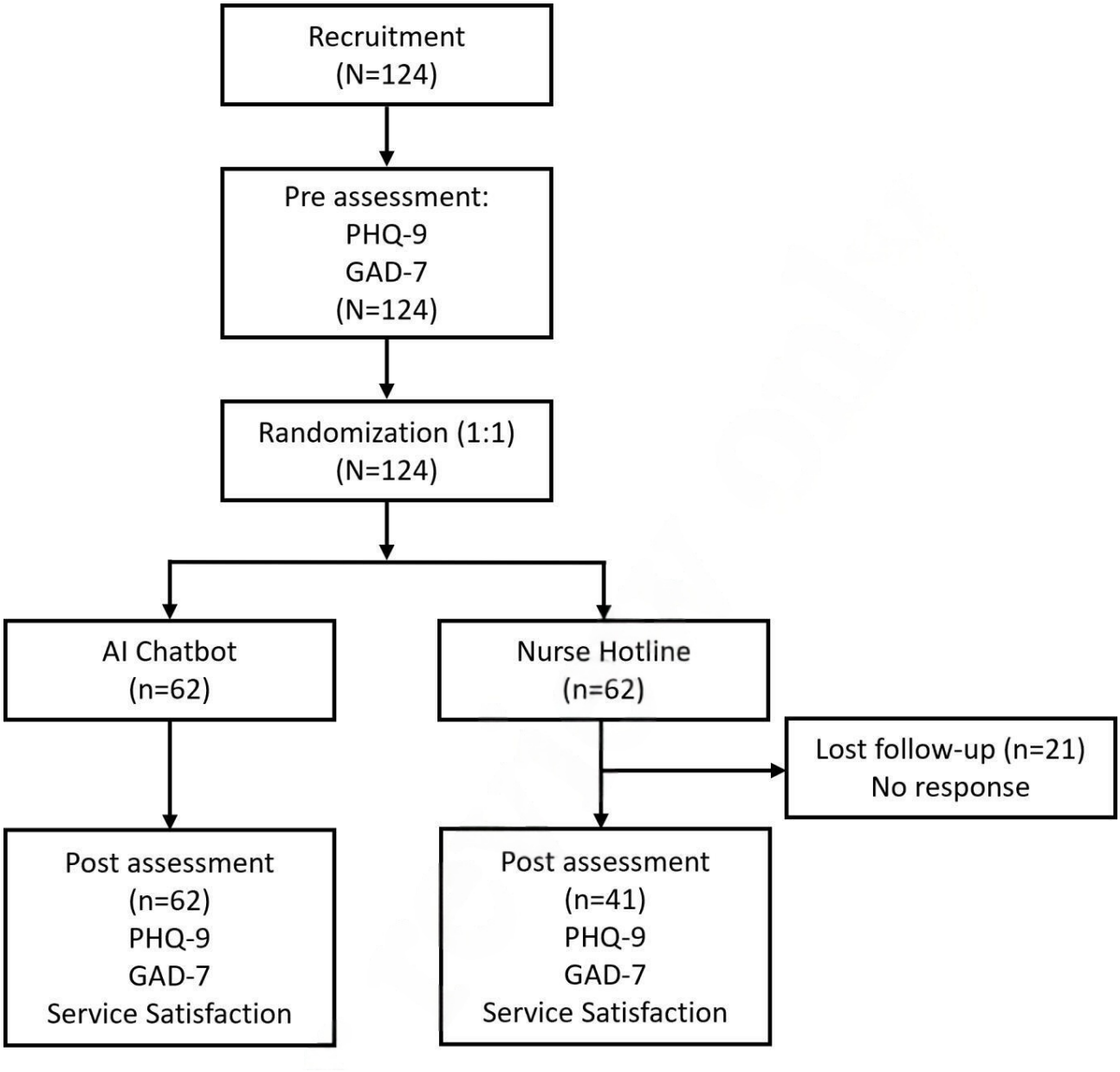


Table . Comparison of the anxiety and depression scores within groups before and after the test.

Questionnaires and groups (N=124)[a]	Pretest score, mean (SD)	Posttest score, mean (SD)	*P* value
PHQ-9[b]			
AI[c] chatbot (n=62)	5.13 (4.632)	3.68 (4.397)	.008
Nurse hotline (n=41)	5.46 (6.4)	4.76 (4.346)	.28
GAD-7[d]			
AI chatbot (n=62)	4.74 (4.742)	3.4 (3.748)	.005
Nurse hotline (n=41)	4.37 (5.439)	4.05 (4.455)	.63

[a]n=21 participants were lost to follow up.

[b]PHQ-9: Patient Health Questionnaire-9.

[c]AI: artificial intelligence.

[d]GAD-7: Generalized Anxiety Disorder Scale-7.

XSL•FO
RenderX

Figure 2. Comparison of anxiety and depression scores within groups pre- and postintervention. GAD-7: Generalized Anxiety Disorder Scale-7; PHQ-9: Patient Health Questionnaire-9; ns: nonsignificant. **P<.05.

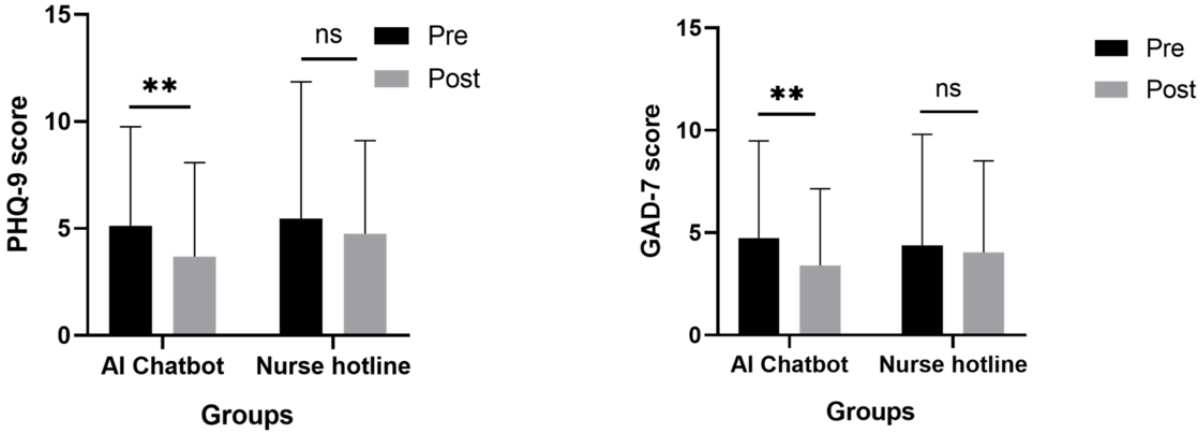


Table 2 shows there were no statistically significant differences at baseline between groups in depression (P=.76) and anxiety scores (P=.71). Although the postdepression and postanxiety scores in the chatbot group were slightly lower than those in the nurse hotline group, there was no statistical difference (all Ps>.05). No statistically significant difference was observed in service satisfaction between the 2 service platforms (P=.32). In the nurse hotline group, there were no notable differences in depression and anxiety scores before and after the test.

Table . Comparison of the before and after PHQ-9[a], GAD-7[b], and service satisfaction scores between the AI chatbot and nurse hotline groups.

Variables	AI chatbot scores (n=62), mean (SD)	Nurse hotline scores (n=41), mean (SD)	P value[c]
Pre–PHQ-9	5.13 (4.63)	5.46 (6.4)	.76
Post–PHQ-9	3.68 (4.40)	4.76 (4.35)	.43
Pre–GAD-7	4.74 (4.74)	4.37 (5.44)	.71
Post–GAD-7	3.40 (3.75)	4.05 (4.46)	.22
Service satisfaction	16.11 (6.88)	17.39 (5.59)	.32

[a] PHQ-9: Patient Health Questionnaire-9.

[b]GAD-7: Generalized Anxiety Disorder Scale-7.

[c]Independent t test.

No significant variations were observed in the pre-post score differences for either depression (P=.38) or anxiety scores (P=.19) between the 2 groups (Table 3).

Table . Comparison of the pre-post anxiety and depression score differences between the artificial intelligence (AI) chatbot and nurse hotline groups.

Variables	Groups		
	AI chatbot (n=62), mean (SD)	Nurse hotline (n=41), mean (SD)	P value
Pre-post differences			
PHQ-9[a]	−1.452 (4.203)	−0.707 (4.161)	.38
GAD-7[b]	−1.339 (3.635)	−0.317 (4.15)	.19

[a]PHQ-9: Patient Health Questionnaire-9.

[b]GAD-7: Generalized Anxiety Disorder Scale-7.

[c]Independent t-test was used to calculate P value.

XSL•FO
RenderX

The various test groups were not associated with the postdepression score (β=−.902, *P*=.18) or the postanxiety score (β=−.845, *P*=.17), after adjusting the prescores (Table 4).

Table . Regression analysis on the posttest score between the artificial intelligence chatbot and nurse hotline adjusted by pretest score (n=103).

Questionnaires and groups	β estimate	Standard error	*t* test (*df=100)*	*P* value
Post–PHQ-9[a] (ref: nurse group)				
Pre–PHQ-9	0.528	0.0615	8.591	<.001
Chatbot	−0.902	0.6720	−1.342	.18
Post–GAD-7[b] (ref: nurse group)				
Pre–GAD-7	0.531	0.0604	8.787	.001
Chatbot	−0.845	0.6148	−1.375	.17

[a]PHQ-9: Patient Health Questionnaire-9.

[b]GAD-7: Generalized Anxiety Disorder Scale-7.

Based on the PHQ-9 and GAD-7 cutoff, the pre- and postscores were categorized into no-risk group and risk groups. No significant differences were observed in the prevalence of anxiety and depression before and after using the nurse hotline or AI chatbot (Table 5).

Table . Comparison of the anxiety and depression prevalence before and after using the nurse hotline and artificial intelligence (AI) chatbot among participants.

Groups	No-risk group (score≤4), n (%)	Risk group (score>4), n (%)	*P* value
Pre–PHQ-9[a]			
Nurse hotline (n=41)	27 (65.9)	14 (34.1)	.20
AI chatbot (n=62)	33 (53.2)	29 (46.8)	
Post–PHQ-9[a]			
Nurse hotline (n=41)	26 (63.4)	15 (36.6)	.33
AI chatbot (n=62)	45 (72.6)	17 (27.4)	
Pre–GAD-7[b]			
Nurse hotline	27 (65.9)	14 (34.1)	.34
AI chatbot	35 (56.5)	27 (43.5)	
Post–GAD-7[b]			
Nurse hotline	27 (65.9)	14 (34.1)	.47
AI chatbot	45 (72.6)	17 (27.4)	

[a]PHQ-9: Patient Health Questionnaire-9.

[b]GAD-7: Generalized Anxiety Disorder Scale-7.

Discussion

Principal Results

This study is among the first in Hong Kong to compare the effectiveness of AI chatbots and nursing hotlines in addressing mental health–related inquiries. The AI chatbot proved effective in reducing anxiety and depression during the COVID-19 outbreak and was comparable to the conventional nursing hotline. It suggests the potential use of AI chatbots as a complementary approach in mental health interventions, particularly when AI chatbots can provide timelier and uninterrupted 24-hour support to users in need during the challenging periods such as a pandemic.

Previous research comparing the effectiveness of the AI chatbot and humans in improving mental health is controversial. Our study found that both the AI chatbot and nurse hotline had similar effects in alleviating depression and anxiety. The results are consistent with another study comparing the text-only chatbot and voice-based digital human interactions as responders to mental health questions, which reported that the text-only chatbot was more user-friendly, although there were no significant differences in electroencephalography measurements [16]. No differences were observed in another RCT comparing the effects of AI chatbot (ie, ChatGPT-3.5) and nurse education on preoperative anxiety reduction among groups [17]. However, one study comparing physician and AI chatbot responses to patient questions posted on a public social media forum found that the chatbot generated higher quality and empathetic

XSL•FO
RenderX

responses to patient questions than physicians [18]. Additionally, the follow up process took less time for the nurse hotline group compared to the chatbot group. This was primarily because participants in the nurse hotline group were quicker in responding to follow up questionnaires while the chatbot group typically required more than one attempt to gather responses. It is feasible to use the AI chatbots to provide instant information to users about mental health knowledge and basic strategies to dissipate their temporary anxiety and depressed mood. However, the AI chatbot can only be a facilitator, and humans will not be replaced in the long-term professional treatment. The sample size in the above-mentioned studies is very small, which may have limited the power of statistical analysis. Therefore, more RCTs with larger sample sizes are necessary to confirm differences between the AI chatbot and human-based services.

Our study showed a significant reduction in the levels of anxiety and depression after using the AI chatbot. This finding is supported by a study from Argentina, which showed a substantial decrease in anxiety symptoms in university students using *Tess* [8]. Another study reported that the use of the Elomia chatbot for 4 weeks resulted in a significant decrease in the symptoms of anxiety and depression along with reduced negative effects [19]. AI chatbots can offer information to users in a quick and easy format. They can also be programmed to answer specific questions about a certain condition, such as what to do during a medical crisis or what to expect during a medical procedure. The AI chatbot could lower the barriers to care by helping patients access help more quickly and efficiently [20]. AI chatbots appear to have a positive impact on depression and anxiety in a wide range of populations, including university students [8], short-course treatment patients [21], and preoperative patients [17]. The potential benefit of chatbot–assisted mental health support is that it can provide helpful information about depressive moods, especially for users who have difficulty in communicating emotions to other humans [22]. Besides, the chatbot offers real-time relief, emotional support, and instant messages on basic knowledge about health-related questions. AI chatbots can help reduce anxiety and depression by providing accessible, anonymous, consistent, and immediate emotional support. They serve as a valuable supplement to conventional therapy and mental health care, offering users a safe space to express their feelings and find effective coping strategies.

Our study differs from previous ones by addressing a research gap in the broader population. Earlier studies primarily concentrated on specific groups, such as college students aged 18-33 [8], younger individuals aged 19-23 [19], and older adults aged 60 years [17]. The sample size in our study (N=124) aligns with past studies, which typically included approximately 100 participants. A study with a larger sample size of 412 younger participants [19] also found AI chatbots to be effective in alleviating anxiety and depression symptoms. Notably, our study was conducted during the COVID-19 pandemic, unlike previous studies, which were conducted under normal circumstances. Hence, additional research is necessary to understand the efficacy and impact of AI chatbots across different age groups, demographics, and environmental contexts.

Our study was carried out during the COVID-19 outbreak, revealing that AI chatbot interventions are versatile and can be applied in various situations, such as during future pandemics or for disseminating information to prevent other diseases. There is great potential in AI chatbot designing and application during an epidemic like COVID-19 [23]; AI chatbots not only help in reducing depression and anxiety at any given time but also can be embedded into the health care systems, to help patients describe their symptoms and provide preliminary guidance on whether they need to seek medical attention [24]. This can assist in triaging patients based on the severity of their symptoms. They can help patients book appointments with health care providers, reducing the burden on administrative staff. AI Chatbots can send reminders to patients to take their medications on time, improving medication adherence [24]. They can provide ongoing emotional support and coping strategies for individuals dealing with mental health issues like anxiety, depression, and stress. This can complement traditional therapy and make mental health care more accessible. Chatbots can also deliver personalized health education to patients, helping them understand their conditions, treatment options, and preventive measures.

The findings of this study comparing AI chatbots and nurse hotlines in alleviating the levels of anxiety and depression during the COVID-19 pandemic can be generalized to a certain extent, given that the general population was the target group. This broad target population ensures that the results are applicable to a wide range of individuals who may have experienced similar mental health challenges during the pandemic. However, the generalizability may be limited by factors such as the specific characteristics of the sample used in the study, including their demographics and the severity of their anxiety and depression symptoms. Additionally, the dynamic nature of AI chatbots, which can be influenced by various factors such as the phrasing of questions and ongoing optimization processes, may also affect the generalizability of the results. The study's design and methodology provide a solid foundation for replicability. The use of an RCT framework, which is a robust design for evaluating interventions, increases the likelihood that similar studies can be conducted with comparable results. The clear documentation of the study's procedures, including the recruitment process, data collection methods, and statistical analysis techniques facilitates replication by other researchers. Furthermore, the study's focus on a common and significant issue—mental health during a pandemic—ensures that there is a continued interest and relevance in conducting similar studies. However, replicability may be challenged by the need for access to similar populations and resources, as well as the evolving nature of AI chatbot technology, which may require adaptation of the intervention for future studies.

Although AI chatbots have a potential complementary function in mental health and health-related fields, certain drawbacks must be considered. For example, the safety and privacy concerns remain unclear [25]. Chatbots are unable to deliver the emotional support and personal bond that a certified mental health expert may provide. They cannot deliver a diagnosis or management strategy for mental health disorders. Therefore, it is crucial to use AI chatbots as a complement to, rather than a

XSL•FO
RenderX

substitute for, specialized mental health services. Moreover, further refinement to the AI chatbot is necessary for establishing an integrated system for recognition-alert-reporting system for mental health problems. After early detection and identification of depression and anxiety, the AI chatbot should be designed to automatically send self-regulation advice to the users to handle their negative emotions as well as send reminder messages to encourage them to seek medical services. Additionally, the question coverage of the AI chatbot is limited; it covers common mental health–related questions that may overlook the scope of other health problems. The scope of the population using chatbots should be expanded among the public. Further studies are required to draw solid conclusions about the effectiveness and safety of chatbots.

Limitations

One of the limitations of our study is the small sample size. Toward the end of the COVID-19 outbreak, there was a notable lack of enthusiasm among parents to participate in this study. Only two school principal groups were involved, which may have introduced a selection bias. Additionally, the study period was too short to recruit sufficient participants; however, the preliminary results showed a significant reduction in anxiety and depression after using the AI chatbot, supporting the applicability of using AI chatbots to relieve negative emotions during the next epidemic crisis. Demographic data and other confounding factors were not collected due to privacy concerns; these factors will be included in the following larger randomized control study.

Conclusions

The AI chatbot is as effective as the nurse hotline in answering mental health–related questions. Besides, the AI chatbot can reduce short-term anxiety and depression levels compared to the nurse hotline. The developed AI chatbot shows promise as a complementary tool to provide available intervention strategies to the public, particularly for families with limited access to mental health services. This study provides insights into the potential application of using AI chatbots to provide immediate information to the public and mitigate public emotional distress during future epidemic emergencies.

Acknowledgments

This study was funded by the Collaborative Research Fund, University Grants Committee (reference no. C7149-20GF) and General Research Fund, University Grants Committee (reference no. 17606523). All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

Authors' Contributions

Conceptualization: PI, TYSL, ICKW, CSLC
Data curation: KLT
Formal analysis: KMY, CC
Funding acquisition: PI
Investigation: KLT
Methodology: TYSL, ICKW, CSLC
Project administration: PI
Supervision: PI
Writing—original draft: CC
Writing—review & editing: CC, HKS, KMY

Conflicts of Interest

None declared.

Checklist 1
CONSORT-eHEALTH checklist (V1.6.1).
[PDF File, 89 KB - humanfactors_v12i1e65785_app1.pdf]

References

1. Mental health review report. Food and Health Bureau, Hong Kong Special Administrative Region Government. 2017. URL: https://www.healthbureau.gov.hk/download/press_and_publications/otherinfo/180500_mhr/e_mhr_full_report.pdf [accessed 2024-07-15]
2. Choi EPH, Hui BPH, Wan EYF. Depression and Anxiety in Hong Kong during COVID-19. Int J Environ Res Public Health 2020 May 25;17(10):3740. [doi: 10.3390/ijerph17103740] [Medline: 32466251]
3. Lam LCW, Wong CSM, Wang MJ, et al. Prevalence, psychosocial correlates and service utilization of depressive and anxiety disorders in Hong Kong: the Hong Kong Mental Morbidity Survey (HKMMS). Soc Psychiatry Psychiatr Epidemiol 2015 Sep;50(9):1379-1388. [doi: 10.1007/s00127-015-1014-5] [Medline: 25660760]

XSL•FO
RenderX

4. D'Alfonso S. AI in mental health. Curr Opin Psychol 2020 Dec;36:112-117. [doi: 10.1016/j.copsyc.2020.04.005] [Medline: 32604065]
5. Tso WWY, Leung LK, Chow MSC, et al. Mental health of children with special educational needs and the return to in-person learning after the COVID-19 pandemic. JAMA Netw Open 2023 Dec 1;6(12):e2346106. [doi: 10.1001/jamanetworkopen.2023.46106] [Medline: 38060228]
6. Torous J, Bucci S, Bell IH, et al. The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality. World Psychiatry 2021 Oct;20(3):318-335. [doi: 10.1002/wps.20883] [Medline: 34505369]
7. Joerin A, Rauws M, Ackerman ML. Psychological artificial intelligence service, Tess: delivering on-demand support to patients and their caregivers: technical report. Cureus 2019 Jan 28;11(1):e3972. [doi: 10.7759/cureus.3972] [Medline: 30956924]
8. Klos MC, Escoredo M, Joerin A, Lemos VN, Rauws M, Bunge EL. Artificial intelligence-based chatbot for anxiety and depression in university students: pilot randomized controlled trial. JMIR Form Res 2021 Aug 12;5(8):e20678. [doi: 10.2196/20678] [Medline: 34092548]
9. Boucher EM, Harake NR, Ward HE, et al. Artificially intelligent chatbots in digital mental health interventions: a review. Expert Rev Med Devices 2021 Dec;18(sup1):37-49. [doi: 10.1080/17434440.2021.2013200] [Medline: 34872429]
10. Ahmed A, Aziz S, Khalifa M, et al. Thematic analysis on user reviews for depression and anxiety chatbot apps: machine learning approach. JMIR Form Res 2022 Mar 11;6(3):e27654. [doi: 10.2196/27654] [Medline: 35275069]
11. CUHK nursing develops mobile app "pai.ACT" the first AI-driven Cantonese psychological support tool for parents of SEN children in Hong Kong. Faculty of Medicine, The Chinese University of Hong Kong. 2023. URL: https://www.med.cuhk.edu.hk/press-releases/cuhk-nursing-develops-mobile-app-pai-act-the-first-ai-driven-cantonese-psychological-support-tool-for-parents-of-sen-children-in-hong-kong [accessed 2024-07-15]
12. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. EQUATOR Network. URL: https://www.equator-network.org/reporting-guidelines/consort [accessed 2024-07-15]
13. Yu X, Tam WWS, Wong PTK, Lam TH, Stewart SM. The Patient Health Questionnaire-9 for measuring depressive symptoms among the general population in Hong Kong. Compr Psychiatry 2012 Jan;53(1):95-102. [doi: 10.1016/j.comppsych.2010.11.002] [Medline: 21193179]
14. Robinson CM, Klenck SC, Norton PJ. Psychometric properties of the Generalized Anxiety Disorder Questionnaire for DSM-IV among four racial groups. Cogn Behav Ther 2010;39(4):251-261. [doi: 10.1080/16506073.2010.486841] [Medline: 20830629]
15. Mental health screening: Patient Health Questionnaire-9 (PHQ-9). National HIV Curriculum. URL: https://www.hiv.uw.edu/page/mental-health-screening/phq-9 [accessed 2024-07-15]
16. Thunström AO, Carlsen HK, Ali L, Larson T, Hellström A, Steingrimsson S. Usability comparison among healthy participants of an anthropomorphic digital human and a text-based chatbot as a responder to questions on mental health: randomized controlled trial. JMIR Hum Factors 2024 Apr 29;11:e54581. [doi: 10.2196/54581] [Medline: 38683664]
17. Yahagi M, Hiruta R, Miyauchi C, Tanaka S, Taguchi A, Yaguchi Y. Comparison of conventional anesthesia nurse education and an artificial intelligence chatbot (ChatGPT) intervention on preoperative anxiety: a randomized controlled trial. J Perianesth Nurs 2024 Oct;39(5):767-771. [doi: 10.1016/j.jopan.2023.12.005] [Medline: 38520470]
18. Ayers JW, Poliak A, Dredze M, et al. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. JAMA Intern Med 2023 Jun 1;183(6):589-596. [doi: 10.1001/jamainternmed.2023.1838] [Medline: 37115527]
19. Romanovskyi O, Pidbutska N, Knysh A. Elomia chatbot: the effectiveness of artificial intelligence in the fight for mental health. In: Proceedings of the 5th International Conference on Computational Linguistics and Intelligent Systems (COLINS 2021): Computational Linguistics and Intelligent Systems; 2021:1215-1224 URL: https://api.semanticscholar.org/CorpusID:235271819 [accessed 2024-07-17]
20. Habicht J, Viswanathan S, Carrington B, Hauser TU, Harper R, Rollwage M. Closing the accessibility gap to mental health treatment with a personalized self-referral chatbot. Nat Med 2024 Feb;30(2):595-602. [doi: 10.1038/s41591-023-02766-x] [Medline: 38317020]
21. Zhong W, Luo J, Zhang H. The therapeutic effectiveness of artificial intelligence-based chatbots in alleviation of depressive and anxiety symptoms in short-course treatments: a systematic review and meta-analysis. J Affect Disord 2024 Jul 1;356:459-469. [doi: 10.1016/j.jad.2024.04.057] [Medline: 38631422]
22. Chin H, Song H, Baek G, et al. The potential of chatbots for emotional support and promoting mental well-being in different cultures: mixed methods study. J Med Internet Res 2023 Oct 20;25:e51712. [doi: 10.2196/51712] [Medline: 37862063]
23. Battineni G, Chintalapudi NO, Amenta F. AI chatbot design during an epidemic like the novel coronavirus. Healthcare (Basel) 2020 Jun 3;8(2):154. [doi: 10.3390/healthcare8020154] [Medline: 32503298]
24. Lee HO, Kang JO, Yeo JO. Medical specialty recommendations by an artificial intelligence chatbot on a smartphone: development and deployment. J Med Internet Res 2021 May 6;23(5):e27460. [doi: 10.2196/27460] [Medline: 33882012]

XSL•FO
RenderX

25. Abd-Alrazaq AA, Rababeh A, Alajlani M, Bewick BM, Household M. Effectiveness and safety of using chatbots to improve mental health: systematic review and meta-analysis. J Med Internet Res 2020 Jul 13;22(7):e16021. [doi: 10.2196/16021] [Medline: 32673216]

Abbreviations

AI: artificial intelligence
CONSORT: Consolidated Standards of Reporting Trials
***DSM-IV*:** *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*
GAD-7: Generalized Anxiety Disorder Scale-7
NLP: natural language processing
PHQ-9: Patient Health Questionnaire-9
RCT: randomized controlled trial

Edited by A Kushniruk; submitted 26.08.24; peer-reviewed by S Steingrimsson, Y Yang; revised version received 14.01.25; accepted 15.01.25; published 06.03.25.

Please cite as:
Chen C, Lam KT, Yip KM, So HK, Lum TYS, Wong ICK, Yam JC, Chui CSL, Ip P
Comparison of an AI Chatbot With a Nurse Hotline in Reducing Anxiety and Depression Levels in the General Population: Pilot Randomized Controlled Trial
JMIR Hum Factors 2025;12:e65785
URL: https://humanfactors.jmir.org/2025/1/e65785
doi:10.2196/65785

© Chen Chen, Kok Tai Lam, Ka Man Yip, Hung Kwan So, Terry Yat Sang Lum, Ian Chi Kei Wong, Jason C Yam, Celine Sze Ling Chui, Patrick Ip. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 6.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Development of Digital Strategies for Reducing Sedentary Behavior in a Hybrid Office Environment: Modified Delphi Study

Iris Parés-Salomón[1,2*], MSc; Cristina Vaqué-Crusellas[1,3*], PhD; Alan Coffey[4*], MSc; Bette Loef[5*], PhD; Karin I Proper[5,6*], Prof Dr; Anna M Señé-Mir[7*], PhD; Anna Puig-Ribera[2*], PhD; Kieran P Dowd[4*], PhD; Judit Bort-Roig[1,2*], PhD

[1]Faculty of Health Sciences and Welfare, University of Vic – Central University of Catalonia (UVic-UCC), Vic, Spain

[2]Sport and Physical Activity Research Group, Institute for Research and Innovation in Life and Health Sciences in Central Catalonia (IRIS-CC), University of Vic – Central University of Catalonia (UVic-UCC), Vic, Spain

[3]Research group on Methodology, Methods, Models and Outcomes of Health and Social Sciences (M3O), Institute for Research and Innovation in Life Sciences and Health in Central Catalonia (IRIS-CC), Vic, Spain

[4]SHE Research Centre, Department of Sport and Health Sciences, Technological University of the Shannon, Athlone, County Westmeath, Ireland

[5]Centre for Prevention, Lifestyle and Health, National Institute for Public Health and the Environment, Bilthoven, The Netherlands

[6]Department of Public and Occupational Health, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam Public Health Research Institute, Amsterdam, The Netherlands

[7]Sport and Physical Activity Research Group, Sport and Physical Activity Studies Centre, University of Vic-Central University of Catalonia, Vic, Spain

[*]all authors contributed equally

Corresponding Author:
Iris Parés-Salomón, MSc
Faculty of Health Sciences and Welfare
University of Vic – Central University of Catalonia (UVic-UCC)
C Sagrada Família, 7
Vic, 08500
Spain
Phone: 34 938 815 50
Email: iris.pares@uvic.cat

Abstract

Background: Hybrid work is the new modus operandi for many office workers, leading to more sedentary behavior than office-only working. Given the potential of digital interventions to reduce sedentary behavior and the current lack of studies evaluating these interventions for home office settings, it is crucial to develop digital interventions for such contexts involving all stakeholders.

Objective: This study aimed to reach expert consensus on the most feasible work strategies and the most usable digital elements as a delivery method to reduce sedentary behavior in the home office context.

Methods: A modified Delphi study including 3 survey rounds and focus groups was conducted to achieve consensus. The first Delphi round consisted of two 9-point Likert scales for assessing the feasibility of work strategies and the potential usefulness of digital elements to deliver the strategies. The work strategies were identified and selected from a scoping review, a systematic review, and 2 qualitative studies involving managers and employees. The median and mean absolute deviation from the median for each item are reported. The second round involved 2 ranking lists with the highly feasible strategies and highly useful digital elements based on round 1 responses to order the list according to experts' preferences. The weighted average ranking for each item was calculated to determine the most highly ranked work strategies and digital elements. The third round encompassed work strategies with a weight above the median from round 2 to be matched with the most useful digital elements to implement each strategy. In total, 4 focus groups were additionally conducted to gain a greater understanding of the findings from the Delphi phase. Focus groups were analyzed using the principles of reflexive thematic analysis.

Results: A total of 27 international experts in the field of occupational health participated in the first round, with response rates of 86% (25/29) and 66% (19/29) in rounds 2 and 3, respectively, and 52% (15/29) in the focus groups. Consensus was achieved on 18 work strategies and 16 digital elements. Feedback on activity progress and goal achievement; creating an action plan; and standing while reading, answering phone calls, or conducting videoconferences were the most feasible work strategies, whereas wrist-based activity trackers, a combination of media, and app interfaces in smartphones were the most useful digital elements.

XSL•FO
RenderX

Moreover, experts highlighted the requirement of combining multiple levels of strategies, such as social support, physical environment, and individual strategies, to enhance their implementation and effectiveness in reducing sedentary behavior when working from home.

Conclusions: This expert consensus provided a foundation for developing digital interventions for sedentary behavior in home office workers. Ongoing interventions should enable the evaluation of feasible strategies delivered via useful digital elements in home office or hybrid contexts.

(JMIR Hum Factors 2025;12:e59405) doi:10.2196/59405

KEYWORDS

sedentary behavior; office work; home office; hybrid work; technology; Delphi

Introduction

Background

Workplace sedentary behavior (SB) is associated with poorer health and work-related outcomes, such as musculoskeletal disorders, reduced mental and occupational well-being, job dissatisfaction, and fatigue [1-4]. According to the World Health Organization, the workplace is one of the main settings to reduce SB in the adult population [5], with office workers, who spend an average of 77% of their working day sitting [6], being a key target group [7]. After the COVID-19 pandemic, hybrid work, alternating work in the office and at home, has become the new work paradigm for many desk-based jobs [8-10]. Hybrid work results in even higher levels of SB than working solely in the office [11], increasing the risk of musculoskeletal health problems, mental health issues, and reduced work productivity [11,12].

Developing strategies to reduce and break up workplace SB, replacing it with activities other than sitting, such as walking or standing, has been the focus of recent studies [13,14]. In particular, the use of digital strategies in desk-based jobs (eg, delivered via mobile phone, activity tracker, or desktop computer) has the potential to provide scheduled prompts, deliver information, give automated tailored feedback, and mediate organizational support and social influences [15]. Recent evidence has highlighted the capability of digital interventions to reduce workplace SB and its associated harmful effects on health and work-related outcomes among office workers [16-18]. However, this research has focused on the traditional office environment, with a dearth of research available on the use of digital interventions to reduce of employees' SB while working remotely or undertaking hybrid work [19,20].

Multicomponent digital interventions that combine environmental and organizational changes and provision of information and counseling have been reported as more effective than single-component interventions in reducing occupational SB among office workers [16,21]. In addition, the grounding of interventions in a theory of behavior change is key in the success and behavior maintenance of multicomponent interventions [22-24]. Given the limited evidence base on the effectiveness of interventions targeting hybrid work, it is crucial to develop multicomponent digital interventions grounded in sound theory of behavior change to reduce occupational SB among home office workers.

Objectives

To contribute to a successful implementation of an intervention and, thereby, to its effectiveness, involving stakeholders such as researchers, experts in the addressed topic, and the target population in the development of interventions is necessary to identify needs, priorities, and potential solutions [25]. A previous qualitative study identified which factors influenced employees' ability to reduce SB when working from home from the employers' perspective, whereas additional qualitative evidence identified the factors influencing the reduction in SB from the employees' perspective [26]. This study aimed to go beyond the existing research gaps to incorporate expert opinions to gain a complete picture of all stakeholders and reach consensus on the feasibility of digital work strategies to reduce SB in a home office context. Therefore, the aim of this study was to reach expert consensus on the most feasible and usable digital work strategies as a delivery method to reduce SB in a home office context.

Methods

Design

A modified Delphi study was used to elicit a consensus from a spectrum of experts in the field of occupational health. It is a flexible approach combining quantitative and qualitative data. Qualitative data can be collected before, after, or between Delphi rounds [27,28]. This study encompassed 3 survey rounds and subsequent web-based focus group sessions. This study is part of the initial modeling of a multicomponent intervention and of a larger European project (ie, the Click2Move program) aimed at tackling the challenge of occupational sedentarism in hybrid office workers.

This study was based on the recommendations of the Conducting and Reporting Delphi Studies guidelines [29] and the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist [30].

Ethical Considerations

This project was approved by the Research Ethics Committee of the University of Vic–Central University of Catalonia (250/2023). This study was conducted in accordance with relevant guidelines and regulations. The invitation email contained general information about the study and a link with further details of the participant information statement, including assurances of anonymity and confidentiality and emphasizing that participation was completely voluntary and participants

XSL•FO
RenderX

could withdraw at any time without any consequences. In compliance with ethical research practices, digital informed consent was obtained from all participants before initiating the survey. No incentives were provided to the participants for their involvement.

Participants and Recruitment

International experts (N=57) identified as leading or established researchers in the field of occupational health (with a specific focus on modification of physical activity [PA] behavior in the workplace), well-being, or computer- or desk-based jobs specialists identified by the research team were invited via email to participate. The experts were selected through a search of published articles that identified highly active authors in the field, whereas the job specialists were identified from professional networks of the involved institutions. The exclusion criterion was not being able to understand and communicate in English.

To maximize participation in the Delphi survey, a reminder email was sent within 1 week. After completing the 3 Delphi rounds, participants who signed the informed consent form for the Delphi study (29/57, 51%) were invited via email to participate in a focus group.

Procedure

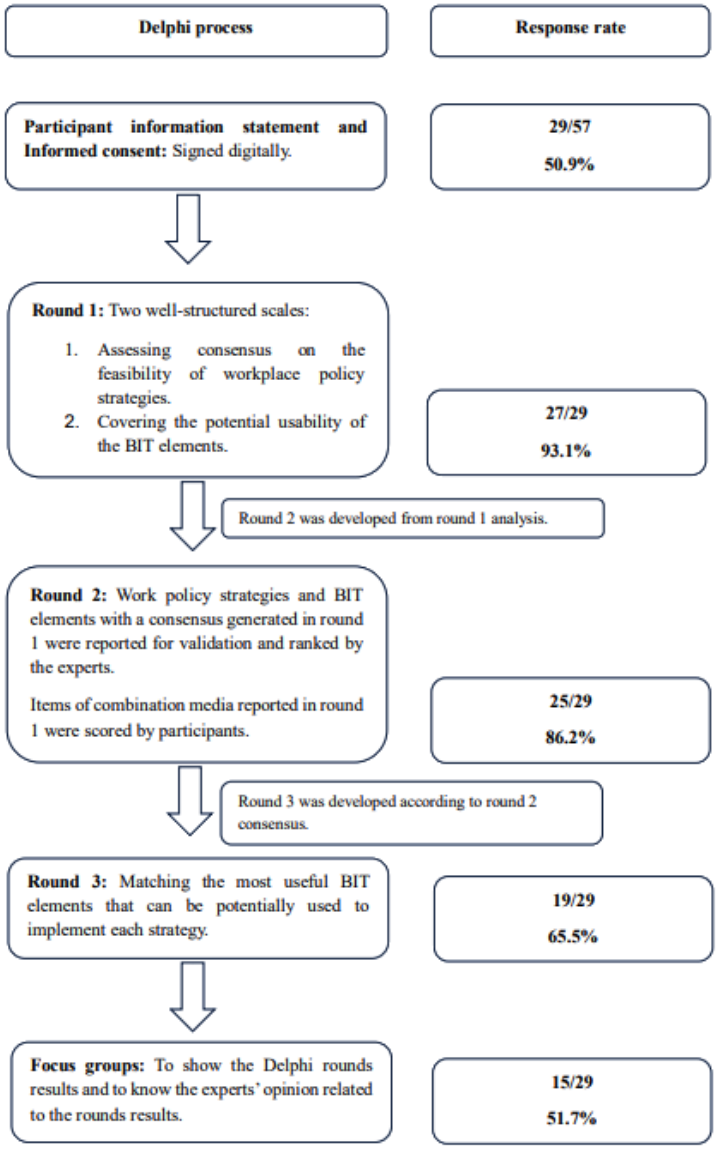
Delphi Survey

Overview

In total, 3 survey rounds were conducted using the LimeSurvey software (LimeSurvey GmbH; Multimedia Appendix 1), which were piloted by 4 members of the research team to ensure the appropriateness of the questionnaire items, wording, and functionality and identify any issues or gaps. No changes in wording were made to the Delphi surveys. If participants agreed to take part, they received an automatic email with the link to the first survey. Participants had 10 days to complete each round of the survey. Individual responses were anonymized, and all participants were invited to complete the second and third rounds of the survey regardless of their participation status in the previous rounds. Attrition was prevented through a short period between rounds and by providing feedback on study findings. A summary of the entire Delphi process is provided in Figure 1.

XSL•FO
RenderX

Figure 1. Delphi process. BIT: behavioral intervention technology.



Round 1

Round 1 was carried out between January 13 and 23, 2023. Participants were provided with an explanation of the survey and were asked to provide demographic information (ie, age, gender, and employer organization). Two 9-point Likert scales were used.

The first scale assessed participants' opinions on the feasibility of work strategies (n=36) for reducing SB when working from home. The work strategies had been selected from a scoping review of the vast majority of real-life PA strategies implemented [31], a systematic review of effective digital interventions to reduce SB [21], and 2 qualitative studies that further explored the perceptions of managers and employees on the factors influencing the ability of employees to reduce their SB in a home office context [26]. The strategies were classified based on the Behavior Change Wheel (ie, environmental planning and service provision, 9/36, 25% of the strategies; regulation, guidelines, and restriction, 15/36, 42% of the strategies; and communication and social support, 12/36, 33% of the strategies) [32].

Participants were invited to score each work strategy. Scores ranging between 1 and 3 suggested a lack of feasibility, scores between 4 and 6 suggested neutral feasibility, and scores between 7 and 9 suggested high feasibility. This scoring was based on the 9-point Grading of Recommendations Assessment, Development, and Evaluation methodology [33]. The feasibility scores considered the 4 dimensions of feasibility: social validity or acceptability, integration into the existing system, practicality, and adaptability (Table 1) [34].

The second scale covered the potential usefulness of the technology-based behavior change elements according to the behavioral intervention technology (BIT) model (ie, messaging and social support [n=6 elements], push notifications [n=5 elements], information delivery [n=9 elements], report and visualization [n=4 elements], digital log [n=3 elements], and passive data collection [n=7 elements]) [35]. Scoring was based on the 9-point Grading of Recommendations Assessment,

XSL•FO
RenderX

Development, and Evaluation scale divided into 3 categories: not useful (1-3), neutral (4-6), and highly useful (7-9) [33]. Participants were also invited to provide additional items or observations through an open-ended question for each strategy or BIT element category.

Table 1. Feasibility dimensions for the strategies and the related research question for each dimension.

Dimension	Research question
Social validity or acceptability	Is the strategy appropriate, reasonable, fair, and potentially effective for home-based work?
Integration into the existing system	To what extent are the strategies aligned with the infrastructure of home office settings?
Practicality	Can the initiative be implemented using the available resources, time, training, and materials in home-work offices?
Adaptability	Is the initiative flexible enough to fit across the diverse needs of home office work?

Round 2

Round 2 started on February 6, 2023, and finished on February 16, 2023. Work strategies from round 1 were ranked based on the *highly feasible* scores for work strategies and *highly useful* scores for BIT elements. A total of 35 work strategies (rather than 36) were ranked in round 2 based on recommendations from round 1 (a similar approach was implemented for work strategies and BIT elements). In this second round, participants were asked to place each work strategy in order of preference considering the most feasible strategies to reduce SB when working from home. The same procedure was followed for the BIT elements. Additional items that appeared in the first survey were scored based on the round 1 approach (ie, the 9-point Likert scale).

Round 3

Round 3 was launched on February 17, 2023, and could be answered until February 27, 2023. According to the preferred order of the items ranked in round 2, participants were asked to match the most useful BIT elements that could potentially be used to implement each of the work strategies to reduce SB when working from home. The work strategies included in this round were those with a weight above the median.

Focus Groups

To explore in more depth the experts' opinions and understand their perspectives, focus groups took place.

Focus groups were conducted in English or Spanish by a researcher (AC or IP-S) and a moderator (IP-S, JB-R, or AMS-M) with previous experience facilitating focus groups. The focus groups were held on the web using Microsoft Teams in April 2023. The focus groups lasted a total of 60 minutes. The size of the groups was small (4 participants on average) due to scheduling difficulties and time zone differences among countries.

A guide was developed based on the 3 Delphi rounds (Multimedia Appendix 2). However, the researchers did not follow it in full detail, using the freedom to explore thoughts and ideas based on the responses from each group. First, there was a brief welcome and introduction to explain the aim of the focus group, obtain consent for the recording of the session, and introduce each participant. The results of every Delphi round were then shown. Subsequently, experts were asked about what factors they thought made some strategies more feasible than others considering the 4 dimensions (outlined in Table 1) and what factors may enhance participation among the BIT elements. The last element of the focus group covered the need for any adaptation before the implementation of the strategies in the hybrid work model. The focus groups were recorded, transcribed verbatim, and translated into English if necessary to be analyzed later.

Data Management and Analysis

Study uptake was calculated as the proportion of participants who accessed the round 1 Delphi survey to that of all potential participants identified. Retention rate referred to the proportion of completed responses for each round and participation in the focus groups. In round 1, responses to the Likert scale in the *highly feasible* or *highly useful* category (score from 7 to 9) were calculated. The median and mean absolute deviation from the median (MAD-M) for each item were also reported to assess the strength and extent of agreement. Medians of 7 to 9 were defined as strong support, medians of 4 to 6 were defined as moderate support, and medians of 1 to 3 were defined as no support. The MAD-M describes the variation in the median values [36]. In round 2, we calculated the weighted average ranking for each item to determine the most highly ranked work strategies and BIT elements for reducing SB among home office workers. The respondents' most preferred choice had the largest weight, and the least preferred choice had a weight of 1. For achieving consensus, the order-weighted average process was followed through the aggregation of the strategies evaluated in the previous round to be validated in round 2 [37]. Items that achieved consensus were those that were ranked with a weighted average above the median. The additional items that emerged from round 1 were assessed through a 9-point Likert scale in round 2, and the median and MAD-M were calculated. In round 3, the matched items for each work strategy and BIT element were reported, as well as the percentage of respondents.

The focus group analysis was conducted using the principles of reflexive thematic analysis [38,39]. First, a researcher (IPS) familiarized themselves with the data, reading and rereading carefully each transcript to obtain a general impression of the main opinion of the experts and identify repeated concepts. Subsequently, according to the subjectivity of the researcher, an initial open coding process was generated from interesting features and relevant data searching for evidence for themes, namely, patterns of shared meaning underpinned by a central organizing concept. Reflexive thematic analysis was approached

XSL•FO
RenderX

in an inductive, semantic, and realistic way. In other words, the construction of codes and themes was directed by the content of the data but not by a topic summary. Codes, concepts, and the core category were recoded and relabeled if necessary after each step until the codes provided the intended depth of the insights. The analytic outputs obtained from the codes and through the active engagement of the researcher with their data were then conceptualized in potential themes and subthemes. Finally, the themes were reviewed, defined, and named. At the end of each step, data were shared with an experienced supervisor (CV-C) for reflecting on how the data were initially coded and consider things that might have been overlooked. MAXQDA 2022 (VERBI GmbH) was used for the analysis.

To ensure a robust analytical process, we considered key aspects of quality and rigor. The first aspect was reflexivity, which involves critical reflection of researcher perspectives and its integration within the analysis and interpretation of the data to develop themes [40]. The study team comprised researchers with expertise in PA and SB in the workplace who have experienced working from home since before or after COVID-19 pandemic. IP-S, who carried out the analysis, is a Spanish physiotherapist and osteopath PhD student who has been working from home for 3 years; CV-C, the supervisor during the analysis, is an experienced nutritionist with a background in health promotion and qualitative research who has been working from home since the COVID-19 pandemic. Second, confirmability, the objectivity of the findings [41], was attained through a critical friend approach whereby an experienced supervisor reviewed interpretations and introduced alternative perspectives, increasing objectivity and confirming the accuracy of the findings. Third, dependability, the reliability and consistency of the study [41], was attained through detailed research procedures (ie, rigorous data collection techniques and well-documented procedures and analyses). Credibility was then built on the reflexivity and triangulation of the data, incorporating more than one researcher in the analysis process [42]. Furthermore, a critical friend approach was used whereby an experienced qualitative researcher, CV-C, critically analyzed the preliminary themes and ensured that the selected quotes provided support for the themes developed. Finally, transferability was enhanced through the transparent and detailed descriptions of the participants and research process [42].

Results

Participant Characteristics

A total of 51% (29/57) of the invited experts expressed their interest in participating in the study through the digital informed consent form. Of these 29 participants, 27 (93%) completed the first round, 25 (86%) completed the second round, and 19 (66%) completed the third round. Of those 29 participants, 15 (52%) joined the focus groups.

Most of the first-round Delphi respondents and focus group participants identified as women (16/27, 59% and 8/15, 53%, respectively), with 48% (13/27) of first-round respondents aged between 36 and 45 years.

A wide variety of countries across Europe were represented. However, non-European countries such as Australia (5/29, 17% of the participants) were also represented.

Delphi Survey

Feasibility of Work Strategies

Round 1

Multimedia Appendix 3 shows the results of round 1, which considered the 36 work strategies divided into 3 categories. A total of 69% (25/36) of the work strategies were ranked as highly feasible.

In total, 7 environmental planning and service provision work strategies were rated as highly feasible with strong support. Activity trackers (20/27, 74%; median 8, MAD-M 1.41); workstation accessories (19/27, 70%; median 8, MAD-M 1.59); and relocation of home office supplies such as the bin, printer, or water source (19/27, 70%; median 8, MAD-M 1.30) were the 3 work strategies considered as highly feasible for reducing SB when working from home. Active workstations equipped with treadmills were reported to be unfeasible (19/27, 70%; median 2, MAD-M 1.48).

In total, 53% (8/15) of the work strategies regarding guidelines, regulations, and restrictions were rated as highly feasible with strong support. Self-monitoring of sedentary and activity behaviors (20/27, 74%; median 8, MAD-M 1.30), substitution of sedentary tasks (eg, reading, answering a phone call, or participating in a videoconference) with standing or incidental movements (20/27, 74%; median 7, MAD-M 1.37), and creating an action plan (20/27, 74%; median 7, MAD-M 1.11) were the most highly useful. The least feasible strategy was the locking or blocking of the screen or keyboard for specific durations (11/26, 42%; median 5, MAD-M 2.19).

Regarding the communication and social support work strategies, feedback on activity progress and achievement of goals (22/27, 81%; median 7, MAD-M 1.74); providing information to increase awareness and knowledge (20/27, 74%; median 8, MAD-M 1.37); and the provision of awards, rewards, or incentives by the company to achieve goals (19/27, 70%; median 7, MAD-M 1) were ranked as the most feasible for reducing SB when working from home. Most of the work strategies included in this category (10/12, 83%) were rated by the respondents as highly feasible, and none of them were considered unfeasible. No new items emerged from open-ended questions.

Round 2

The ranking list of the work strategies with the weighted average (range 14.6-0.48) is reported in Table 2. Most items reported as highly feasible when working from home in round 1 (18/25, 72%) were scored above the median (3.64) in round 2, with the exception of self-selected reminders (3.24); active breaks such as stretching, walking, or strengthening exercises (2.36); real-time records (2.32); high or height-adjustable chairs (1.96); exercise accessories (1.52); and social networking for sharing experiences (0.48).

XSL•FO
RenderX

Table 2. Ranking of highly feasible work strategies (results of round 2).

Strategy	Weighted average score
Feedback on activity progress and goal achievement	14.6
Create an action plan—increase standing breaks or replace sitting time, how long, how often, when, and how (eg, when the phone rings)	13.04
Standing while reading, answering phone calls, or conducting videoconferences	13.04
Providing information to increase awareness and knowledge of the dangers associated with prolonged sedentary behavior and the potential benefits of reducing it or breaking it up	12.88
Self-monitoring sedentary and activity behaviors (ie, activity tracker or a diary log)	10.76
Short breaks (5-10 min) approximately every 60 min of sitting time	10.36
Scheduling (blocking) 5-10–min breaks between meetings on the calendar	8.6
Information and support about the strategies and goals and reminders	8.04
Setting tailored goals for reducing sitting time	7.24
Activity demonstrations to perform during the breaks	7.12
Height-adjustable desks or desk platforms	6.92
Relocation of home office supplies (eg, bins and printers)	6.8
Incidental moving while reading, answering phone calls, or conducting videoconferences	5.76
Awards, rewards, or incentives to achieve goals or recommendations	5.48
Standing desks	4.6
Point-of-choice or point-of-decision prompts	4.4
Workstation accessories (seated footrests, standing footrests, or sit-stand antifatigue mats for standing desks)	4.24
Motivational messages from managers	3.64
Wellness coaches supporting the employees during breaks	3.48
Self-selected reminders to achieve goals	3.24
Blocking the screen or keyboard for breaking up sitting time unless conducting a meeting	3.24
Active workstation with underdesk stepper or pedaling device	2.52
Short breaks (5-10 min) approximately every 30 min of sitting time	2.48
Team or individual activity challenges	2.4
Active breaks, such as stretching, walking, or performing strengthening exercises for ≥10 min	2.36
Scheduling (blocking) snack breaks on the calendar	2.32
Real-time records or feedback	2.32
Active lunchtime, such as Pilates, yoga, walking, or cycling	2.04
High chairs or height-adjustable chairs	1.96
Social comparison	1.68
Exercise accessories, such as rubber bands, wooden sticks, or mats	1.52
Active workstation equipped with treadmill	1.44
Short breaks (5-10 min) approximately every 40 min of sitting time	1.36
Competition among peers	0.92
Social networking for sharing experiences	0.48

Usefulness of BIT Elements

Round 1

Multimedia Appendix 4 shows the results of round 1 considering the usefulness of the BIT elements divided into 5 categories.

In total, 33% (2/6) of the BIT elements from the messaging and social support category were ranked as highly useful with strong support: social challenge features (20/27, 74%; median 7, MAD-M 1.22) and gamification (14/26, 54%; median 7, MAD-M 1.78). None of the messaging and social support items were considered not useful. Smartphone app and desktop application interfaces were identified to be the 2 most useful

XSL•FO
RenderX

BIT elements within the push notification category (12/27, 44%; median 7; and MAD-M 1.41 for smartphone app interface and 15/27, 56%; median 7; and MAD-M 1.18 for desktop application interface). The other 3 push notification elements were ranked as highly useful or neutral with medians between 5 and 6, indicating moderate support. A combination of different media was the preferred item in the delivery of information category (14/24, 58%; median 7, MAD-M 1.71), with the other elements within the category receiving moderate support and being considered neutrally or highly useful. A total of 15 new combined digital items, such as videos and images, moving images on websites, or reminders combined with social challenges, emerged from open-ended responses to the combination of media element in the information delivery category. Blending was considered more useful than isolated media. Real-time data (19/27, 70%; median 7, MAD-M 1.48) and data summaries via the app interface (14/27, 52%; median 7, MAD-M 1.67) were considered the most useful in the report and visualization category. The 2 remaining items (data summaries via email or SMS text message and websites) were ranked as neutral. The 3 most useful BIT elements to track activity were wrist-based activity trackers (17/26, 65%; median 7, MAD-M 1.69) and using a mobile phone diary (17/27, 63%; median 7, MAD-M 1.41) or smartphone sensors (13/26, 50%; median 7, MAD-M 1.61). The only element considered as not useful to track activity was waist-based activity trackers (10/26, 38%; median 5, MAD-M 1.85).

Round 2

Overall, BIT elements reported as highly useful to support work strategies were ranked at the top in each category (Table 3). In the messaging and social support category, items ranked above the median score (2.52) coincided with highly useful items in round 1 (ie, social challenge features and gamification), but chats (eg, WhatsApp groups) were added to the top of the ranking. SMS text messages were added to the 2 BIT elements in the push notification category considered as highly useful in round 1, which were scored above the median (2.00) in round 2. In addition to combination of media, videos, interfaces in smartphone apps and desktop applications (same as in the push notification category), images, and websites were among the items scored above the median (3.24) in the information delivery category. The new items that emerged from the open-ended responses in round 1, specifying potential combinations of media for information delivery described as highly useful, were a combination of mobile phone apps and desktop applications; a combination of personalized messages, reminders, and social challenges; a combination of videos, text, and images within an interactive website; and a combination of a mobile app with social media (Table 4). In the report and visualization category, the 2 items ranked as highly useful in round 1 remained above the median (2.32) in round 2. The median of the BIT elements to track activity was 3.56. In round 2, a total of 2 more items were added as highly useful to the 3 identified in round 1: leg-based activity tracker and computer software. In total, 17 BIT elements were considered highly useful in round 2.

XSL•FO
RenderX

Table 3. Ranking of highly usable behavioral intervention technology (BIT) elements (results of round 2).

BIT element to communicate information	Weighted average scores (SD)
Messaging and social support	
Social challenge features (cooperative activities, eg, reach 10,000 steps among all members of the department)	5.16
Gamification features (competition among colleagues)	4.48
Chats (eg, WhatsApp groups)	3.28
Emails	1.76
Calls	1.4
Forums	1.64
Push notifications	
App interface—text and sound and vibration in smartphones	4
Application interface—computer screen notification in desktop computers	3.08
SMS text messages	2
Chats	1.64
Emails	1.52
Information delivery	
Combination of media	7.6
Videos	6
App interface—smartphones	4.44
Images	4.36
Application interface—desktops	3.24
Websites	3.24
Emails	3.16
SMS text messages	2.36
Audios	1.44
Report and visualization	
Real-time data via app interface	2.96
Data summary via app interface	2.84
Data summary via email or SMS text message	1.8
Data summary via website	1.24
BIT elements to track activity	
Mobile phone diary (manual entry)	7.72
Web-based questionnaire (manual entry)	3.52
Computer software diary (manual entry)	1.6
Passive data collection	
Wrist-based activity tracker	8.6
Leg-based activity tracker	3.64
Smartphone sensors	3.6
Computer software	3.6
Waist-based activity tracker	3.28
Electronic workstation	3.08
Cushions on chairs	2.04

XSL•FO
RenderX

Table 4. Results for the new items from round 1 scored in round 2—behavioral intervention technology (BIT) elements.

New BIT element	Level of usefulness, n (%)									Median (MAD-M[a])
	Not useful			Neutrally useful			Highly useful			
	Score of 1	Score of 2	Score of 3	Score of 4	Score of 5	Score of 6	Score of 7	Score of 8	Score of 9	
Mixture of videos, information, and images within an interactive website (n=23)	0 (0)	1 (4)	2 (9)	2 (9)	0 (0)	6 (26)	5 (22)	6 (26)	1 (4)	6 (1.48)
App and social media (n=23)	0 (0)	1 (4)	3 (13)	0 (0)	2 (9)	1 (4)	10 (43)	4 (17)	2 (9)	7 (1.30)
Videos and images (n=23)	0 (0)	1 (4)	1 (4)	1 (4)	4 (17)	3 (13)	12 (52)	1 (4)	0 (0)	6 (1.17)
Audio and video combination with an explanation (educational purposes; n=23)	0 (0)	0 (0)	2 (9)	1 (4)	4 (17)	6 (26)	2 (9)	3 (13)	3 (13)	6 (1.48)
Moving images on a website (n=22)	1 (5)	1 (5)	3 (14)	2 (9)	4 (18)	7 (32)	3 (14)	1 (5)	0 (0)	5 (1.41)
Smartphone and desktop with website or app for background information (n=22)	1 (5)	0 (0)	1 (5)	0 (0)	3 (14)	3 (14)	6 (27)	7 (32)	1 (5)	6.5 (1.39)
Personalized messages and app (n=22)	0 (0)	0 (0)	0 (0)	0 (0)	1 (5)	3 (14)	8 (36)	5 (23)	5 (23)	7 (0.91)
Reminders combined with an app (n=22)	0 (0)	0 (0)	2 (9)	0 (0)	1 (5)	4 (18)	4 (18)	8 (36)	3 (14)	7 (1.27)
Reminders combined with social challenges (n=21)	0 (0)	0 (0)	1 (5)	0 (0)	2 (10)	3 (14)	4 (19)	6 (29)	5 (24)	7 (1.29)
Forum for general information and smartphone messages (eg, SMS text messages) for specific tailored information (n=21)	1 (5)	1 (5)	0 (0)	5 (24)	2 (10)	6 (29)	2 (10)	3 (14)	1 (5)	5 (1.67)
Forum for general information and app for specific tailored information (n=21)	1 (5)	1 (5)	1 (5)	1 (5)	4 (19)	6 (29)	1 (5)	4 (19)	2 (10)	5.5 (1.67)
Forum for general information and email for specific tailored information (n=21)	2 (10)	1 (5)	2 (10)	2 (10)	3 (14)	5 (24)	4 (19)	2 (10)	0 (0)	5 (1.71)
Website for general information and smartphone messages (eg, SMS text messages) for specific tailored information (n=23)	0 (0)	1 (4)	3 (13)	4 (17)	3 (13)	5 (22)	4 (17)	3 (13)	0 (0)	4.5 (1.65)
Website for general information and email for specific tailored information (n=23)	0 (0)	2 (9)	2 (9)	4 (17)	5 (22)	5 (22)	3 (13)	1 (4)	1 (4)	4.5 (1.52)
Website for general information and app for specific tailored information (n=23)	0 (0)	1 (4)	2 (9)	1 (4)	6 (26)	6 (26)	3 (13)	3 (13)	1 (4)	5 (1.43)

[a]MAD-M: mean absolute deviation from the median.

Round 3

After the description of the priorities of the expert panel, the BIT elements that were considered most suitable for implementing each of the work strategies are shown in Multimedia Appendix 5, including the 18 highly feasible work strategies and 16 rather than 17 highly useful BIT elements because social challenge features were considered within the gamification features in round 3.

The use of wrist-based activity trackers was identified as useful by 63% (12/19) of respondents for self-monitoring sedentary and activity behaviors. Using a computer software was considered useful to schedule 5- to 10-minute breaks between meetings (9/19, 47%), whereas a desktop application interface was considered useful to support strategies such as short breaks every 60 minutes of sitting time (6/19, 32%). Videos were described to be useful when delivering activity demonstrations to perform during the sitting breaks (9/19, 47%). Other BIT elements identified were gamification features, such as social challenges, which are used for providing awards and rewards and offering incentives (8/19, 42%). However, setting tailored goals (8/19, 42%); creating an action plan (6/19, 32%); standing while reading, answering calls, or conducting videoconferences (6/19, 32%); providing feedback on activity progress (4/19, 21%); and motivational messages (4/19, 21%) were mainly recognized to be best delivered through a smartphone app. Providing information and support was considered to be best delivered through a combination of media or a smartphone app (5/19, 26% each). Combination of media was also chosen by most of the respondents as the most useful mode for providing information to increase awareness and knowledge (5/19, 26%) and for promoting incidental movement (3/19, 16%).

XSL•FO
RenderX

Environmental strategies such as height-adjustable desks (11/19, 58%), relocation of home office supplies (5/19, 26%), standing desks (9/19, 47%), and workstation accessories (9/19, 47%) did not correspond with any BIT elements according to most of the respondents. For point-of-choice or point-of-decision prompts, there were no corresponding BIT elements to support their implementation according to most of the respondents for this strategy (4/19, 21% of the sample), although 16% (3/19) considered that this strategy could be delivered via computer software.

Focus Groups

Overview

A total of 2 indicative themes were generated from the analysis of the focus groups regarding the experts' perceptions in relation to the Delphi results on the potential feasibility and usefulness of strategies and digital elements to reduce SB in a home office context. The first theme referred to the integration of the digital strategies for reducing SB when working from home, which included 3 subthemes: social influences, physical environment, and individual factors. The second theme was related to the adaptability of the strategies to reduce SB in the office and home context.

Theme 1: Integration of the Digital Strategies for Reducing SB When Working From Home

Overview

This theme involves the importance of combining multiple levels of strategies to facilitate and maximize the strategies' implementation and their effectiveness in reducing SB when working from home. According to the experts, while organizational-level factors such as organizational support or policy changes may not be immediately feasible on their own, they are essential for achieving sustainable change. The same was considered true for information provision—it alone is not enough to motivate behavior change, although it is an important factor at the individual level to increase individuals' knowledge. In addition, environmental changes are required to promote less sitting and more movement during working hours. Therefore, experts highlighted that the following implementation levels need to work in tandem to achieve sustainable behavior change.

Social Support

Experts explained the need for the company to support the culture of active and healthy workplaces to motivate the employees' behavior change. Social support may be achieved by aligning the company policies and culture with the support of the manager (eg, motivational messages) to be able to increase the integration of digital strategies for reducing SB when working from home:

> *[I]t is based on what the management or the philosophy of the company or the company policy wants to refer to their workers. If those at the top are not proactive in...the chain, it will decrease.* [E13; focus group 4; paragraph 43]

While the target behavior is contextualized within the domain of working from home, participants identified interaction with colleagues as a relevant factor to motivate employees toward behavior change. For example, social modeling may motivate home office workers to reduce their SB by sharing activities around the sitting time breaks. However, social interaction can be perceived negatively if a competition component is present in the strategies. This may be due to competing against people whom they do not know or due to the bottom position of some workers during the challenges. Therefore, cooperative challenges may be the solution, in which every worker contributes to the challenge without a competition element:

> *[W]hen people find themselves in an environment where they've got to adapt behaviourally, seeing what can be done and observing how others might do it, can be quite powerful...we are social animals, monkey see, monkey do and that social modelling element to me jumps out as really quite important and not to be lost.* [E01; focus group 1; paragraph 44]

Physical Environment

The home office context was seen by some participants as a limiting factor in engaging in sufficient PA during work hours due to the reduced space compared to the office context. For example, having (normally) no more than 1 floor at home, implying the use of stairs, or the lack of large spaces for walking are barriers to be active while working at home. Although applying physical changes such as standing desks may be a solution to break up sitting time at work, it appears to be a less feasible strategy in the home office context due to the expensive cost for the companies:

> *I'm a big advocate of standing desks, standing workstations in terms of being effective but I appreciate that in this situation it's probably not feasible for them to be incorporated in the home-office.* [E10; focus group 3; paragraph 7]

Nevertheless, experts highlighted the importance of combining these environmental changes with prompts or reminders for the success of the behavior change integration when working from home. Regarding point-of-decision prompts, experts highlighted that mobile phone notifications were more feasible for reducing SB than desktop notifications because they cause less disturbance:

> *[W]hen we moved into our new offices a few years ago, we all got standing desks and just observing...I don't see anybody using. So, if there's something for me in there about prompting/reminders, do you know?... So, I think the prompting is really important.* [E011; focus group 3; paragraph 11]

Individual Factors

Experts mentioned the difference in literacy between *PA* and *sitting time*, believing that workers may understand the term *PA* more easily than the term *sitting time*. This may be due to the lack of knowledge on the importance of reducing sitting time behaviors, which could be perceived as a barrier when only reporting sitting time feedback. Therefore, educational material or information, one of the highlighted strategies to be used together with social support and physical environment strategies to achieve SB reduction when working from home, delivered through a combination of media (ie, smartphone, desktop, and

XSL•FO
RenderX

websites), was considered the most adaptative medium to increase the employees' knowledge. Experts emphasized that the information delivered must be simple and short, especially with the use of videos:

> *I think the combination of media is probably the way forward because I guess different organisations will also have different ways of rolling out information. I know a lot of companies nowadays that they all have their own company app, and they do a lot of communication via their app that all their employees have access to.* [E10; focus group 3; paragraph 42]

Furthermore, feedback on the respondents' own behavior was identified as a key strategy to motivate individuals and achieve behavior change. The best frequency at which to deliver feedback could be in real time as this adapts to each individual's schedule. However, a daily summary feedback approach was also considered helpful in combination with real-time feedback:

> *For the summary—daily maybe? But I probably go for the combination of the two, so you probably have some sort of summary data to share with them but still, it's nice to see at the moment what you did.* [E05; focus group 2; paragraph 67]

Theme 2: Adaptation to the Home Office Environment

This theme refers to the flexibility of the strategies to fit the diverse needs of hybrid work. All the experts mentioned the capacity of the strategies to be adapted to all environments without the need to be changed to fit in the home or office context:

> *I think you want to have the same or approach or whatever you use for say, remote working, I would also use for hybrid working because I guess a lot of strategies are going to be the same...So, I think you could maybe just replicate or duplicate what you're doing here for the hybrid workers.* [E10; focus group 3; paragraph 64]

However, some strategies need to be adapted to each individual depending on the type of work they perform. For example, walking while on a call may be a good strategy for a call center, but it might not work in another type of job. Therefore, it is valuable to provide different strategies to allow everyone to choose their preferred one to change their behavior:

> *[D]ifferent populations need a different approach, right? So very much dependent on the type of work that people are doing, you're going to have to approach them differently. So, the task set is going to help determine what options you actually have to still do your work but would be doing it sitting down all the time.* [E05; focus group 2; paragraph 29]

Therefore, experts highlighted the importance of providing a list of strategies that can be adapted to employees' needs for reducing sitting and moving more during their working hours:

> *So instead, you look towards providing a menu of strategies that you know are effective or have at least been shown to be feasible and acceptable and not shown to have any harm and you know, have benefits maybe in terms of cultural elements of it and that a lot of those would be applicable for the home environment but would also be applicable obviously for the workplace environment.* [E02; focus group 1; paragraph 65]

Summary

In general, no dissonance was found among the responses of the different experts. Strategies such as feedback, social modeling, breaks every hour, prompts via mobile phone every hour, educational material, and organizational support (eg, motivational messages) were the most discussed and considered the most feasible. Other strategies such as physical changes (ie, height-adjustable desks) were considered unfeasible despite being ranked higher in the Delphi rounds. Finally, although certain strategies such as the creation of an action plan, standing while reading and answering calls, blocking 5 to 10 minutes between meetings, or setting tailored goals for reducing sitting time obtained high scores in the Delphi rounds, these topics were discussed less among the experts during the focus groups.

Discussion

Principal Findings

This study achieved international expert consensus on 18 potentially feasible work strategies to reduce home office SB delivered through 16 potentially useful digital elements. These results seem to be transferable to both office and home settings based on experts' responses. Moreover, the experts highlighted the importance of multiple factors (ie, social, individual, and environmental) working in tandem to facilitate and maximize the strategies' implementation, as well as to achieve behavior change.

Comparison With Prior Work

Wrist-based activity trackers were identified by most of the experts (12/19, 63%) as useful for self-monitoring activity and SB. Previous research has highlighted the feasibility and usability of these devices for self-monitoring SB as well as workers' acceptability of these approaches in the office context [43,44]. In addition, these strategies have demonstrated effectiveness in reducing SB in such settings [44]. Some studies have highlighted limitations in recognition of postures and step detection when using wrist-based devices compared to leg-based devices [45-48]. However, few experts considered leg-based activity trackers for self-monitoring activity and SB due to the discomfort they may cause. This suggests that future interventions incorporating activity trackers with the aim of reducing SB should first validate the devices for their subsequent implementation and effectiveness evaluation in home office settings.

In the literature, computer visual feedback on activity patterns is considered to be an enabling factor for increasing motivation and awareness among office workers [49]. However, experts preferred smartphone apps as they are able to provide real-time feedback on activity progress. In addition, feedback may be an indication of progression toward a defined goal [43]. In this study, experts stated that smartphone apps could be a useful tool for setting tailored goals. Evidence has demonstrated the

XSL•FO
RenderX

effectiveness of smartphone apps in reducing SB and promoting PA in office settings [16]. Hence, smartphone apps may be a potential mode of delivery for setting tailored goals and providing feedback on the move, which should be evaluated in future home office environment interventions.

Previous research has demonstrated that short activity breaks (ie, walking) every 30 minutes result in reductions in office SB time [50] but also in reductions in fatigue and improvements in energy and mood [51]. However, in this study, experts considered breaks every 60 minutes more feasible. According to recent literature, a possible explanation may be related to the fact that workload and time pressure were identified by employees as the most common barriers for reducing and breaking up SB [52], making it difficult to stop more frequently. Moreover, replacing SB with active working tasks (eg, incidental moving while reading and answering calls) may overcome these barriers [49]. Working remotely increases opportunities to perform active tasks or even other light-intensity PA that may increase the potential benefits of reducing SB and its harmful health effects among office workers [9,10,53]. Experts reflected on the importance of providing clear instructions and modeling through video demonstrations as a key digital element to enhance active breaks among employees when working from home. This may have implications for future studies evaluating the most effective active break duration through video demonstrations and the implications of these breaks regarding SB time and its health effects in home office contexts.

On the basis of experts' responses, there is a literacy gap between the terms *PA* and *SB*, with activity trackers largely focused on PA instead of SB [52]. Therefore, educational material may be a key aspect to increase knowledge for behavior change and reduce the lack of understanding of SB and its subsequent health implications, which was highlighted by experts. Educational material is an essential and effective strategy for workers to increase awareness of SB's health consequences [52,54]. In our findings, a combination of media was preferred to deliver information to increase awareness and knowledge. Therefore, it is recommended that future studies include educational material to increase knowledge and awareness among home office workers.

Despite the relevance of knowledge and awareness to achieve behavior change, social and organizational support is a factor noted by workers in the current literature for facilitating and encouraging movement in the workplace [55-57], as well as by managers in the home office setting [26]. Home office work is more related to negative emotions than office work due to the isolation and loneliness; therefore, social support is even more important in this context [58]. Gamification features, including social challenges, were identified as a useful digital element to deliver strategies such as awards, rewards, or incentives to achieve goals or recommendations for reducing SB among home office workers. This strategy could provide social support and motivation. Motivational messages from managers are also an organizational support strategy, which experts suggest could be delivered through smartphone apps. Thus, future studies should evaluate the acceptability of these social and organizational strategies among both managers and employees, as well as their effectiveness in the home office context.

Environmental work strategies such as standing desks, height-adjustable desks, or workstation accessories were mostly not matched with any digital element. Moreover, experts recognized work strategies such as height-adjustable desks as unfeasible in the home office environment, primarily due to the associated costs. Recent evidence has demonstrated the high cost of such desks (eg, standing desks or height-adjustable desks) compared with nondesk interventions [59]. Workstation accessories (eg, desk platforms) or relocation of office supplies are less expensive strategies [60] that were deemed more feasible and usable by experts and may be implemented in home office settings. Therefore, home office interventions may include environmental work strategies in a similar way to office interventions but with a lower associated cost.

On the basis of the findings presented in this paper, physical environmental strategies need to work in tandem with social support and individual strategies to achieve behavior change. A recent systematic review showed that theory-based interventions, a relevant step in multicomponent intervention development [25], that include social, physical, and individual elements are more effective than non–theory-based interventions [21]. Furthermore, stakeholder engagement (eg, consultation or coproduction among researchers, experts on the topic, and the target population) should be encouraged throughout the development process [25]. Recent studies have highlighted the need for a top-down approach whereby managers should provide support to employers [26]; hence, the implementation of strategies should be an organizational initiative. Although evaluating effectiveness on outcomes has traditionally been the primary focus of interventions in this discipline, in particular to encourage their adoption in the real world, the feasibility, acceptability, cost-effectiveness, and scalability of interventions is increasingly important to ensure the translation of findings from research to practice [25,61]. This study provides new insights for future research on the design of home office interventions aimed at reducing SB based on the learned experiences of occupational health experts. However, we were not able to cover the entire audience's perspective, so new qualitative studies exploring the perspectives of different stakeholders (eg, end users) may further inform about the feasibility and usability of digital workplace strategies.

In addition, experts considered that no adaptations would be needed to transfer these home office strategies to hybrid work, the current modus operandi for many desk-based workers. However, the experts who participated in this study were mainly from high-income countries; therefore, such perceptions may vary among people from other regions. Hybrid work presents opportunities for promoting PA by providing flexibility to integrate occupational, leisure, lifestyle, and incidental activities [9] but also for improving health and well-being [62]. Hence, the results of this study may be transferable to hybrid work in high-income countries, whereas hybrid work should receive greater consideration in future interventions given cultural variability.

Strengths and Limitations

The main strength of this study is the novelty of the home office context to deliver strategies for reducing SB through digital

XSL•FO
RenderX

elements. The fact that focus groups were conducted after a Delphi study enabled us to elicit additional views and opinions on the Delphi results from the experts' perspectives, consolidating the consensus. Nevertheless, we did not share the analytical categories, interpretations, or conclusions from the thematic analysis with the experts, which may affect the analysis's credibility. The entire study was conducted in a rigorous manner, with high participant response rates and representation from international experts. However, some strategies may have been overlooked due to the nonsystematic search approaches used, and most of the international experts were from high-income countries, limiting the applicability of these findings to those communities. Some other limitations were present in the study. First, despite the rigor in the methodology used to engage experts, not all responses were complete, meaning that data were sometimes missing from some questions. Second, as is the case in any Delphi survey, the data gathered were based on subjective opinion and the expertise of the participants. Third, experts acted as representatives of the stakeholders; thus, the inclusion of other stakeholders such as end users would also have been important to address other, more diverse perspectives.

Conclusions

This study presents the consensus of international experts on feasible work strategies and useful digital elements as a mode of delivery of these work strategies to support the reduction in SB in home office or hybrid contexts. Consensus was achieved on 18 work strategies and 16 BIT elements. Future interventions should implement and evaluate the effectiveness of these strategies but also their feasibility and acceptability in home office settings.

Acknowledgments

The authors thank all the participants for contributing their time and expertise, including Dianne Commissaris, Emma George, Erwin Speklé, Frank Vandaele, Genevieve Healy, Ida Damen, Ilkka Väänänen, Ingrid Hendriksen, James McKenna, Judith Jelsma, Katja Wouters, Leon Straker, Maja Pajeck, Matthew P Buman, Mette Jensen Stochkendahl, Neville Owen, Nipun Shrestha, Pania Karnaki, Pieter Coenen, Satu Mänttäri, Sebastià Mas Alòs, Stacy Clemes, Wendy J Brown, and others with their names withheld. This research was funded by the Cooperation Partnerships program of the European Commission (101050490-ERASMUS-SPORT-2021-SCP).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Delphi surveys. The document shows the completed survey of round 1 of the Delphi study on the first page. On the second page, the file presents the completed survey of round 2, and the third page shows the completed survey of round 3.
[XLSX File (Microsoft Excel File), 37 KB - humanfactors_v12i1e59405_app1.xlsx]

Multimedia Appendix 2
Focus group guide. The file shows the complete guide with the times of the focus groups.
[DOCX File , 22 KB - humanfactors_v12i1e59405_app2.docx]

Multimedia Appendix 3
Feasibility of the work strategies (results of round 1).
[DOCX File , 66 KB - humanfactors_v12i1e59405_app3.docx]

Multimedia Appendix 4
Usefulness of behavioral intervention technology elements (results of round 1).
[DOCX File , 22 KB - humanfactors_v12i1e59405_app4.docx]

Multimedia Appendix 5
Matching behavioral intervention technology elements with work strategies on which consensus was achieved (round 3 results).
[DOCX File , 26 KB - humanfactors_v12i1e59405_app5.docx]

References

1. Patel AK, Banga C, Chandrasekaran B. Effect of an education-based workplace intervention (move in office with education) on sedentary behaviour and well-being in desk-based workers: a cluster randomized controlled trial. Int J Occup Saf Ergon 2022 Sep;28(3):1655-1663. [doi: 10.1080/10803548.2021.1916221] [Medline: 33843490]

XSL•FO
RenderX

2. Puig-Ribera A, Martínez-Lemos I, Giné-Garriga M, González-Suárez Á, Bort-Roig J, Fortuño J, et al. Self-reported sitting time and physical activity: interactive associations with mental well-being and productivity in office employees. BMC Public Health 2015 Jan 31;15:72 [FREE Full text] [doi: 10.1186/s12889-015-1447-5] [Medline: 25886270]
3. Rosenkranz SK, Mailey EL, Umansky E, Rosenkranz RR, Ablah E. Workplace sedentary behavior and productivity: a cross-sectional study. Int J Environ Res Public Health 2020 Sep 08;17(18):6535 [FREE Full text] [doi: 10.3390/ijerph17186535] [Medline: 32911740]
4. Putsa B, Jalayondeja W, Mekhora K, Bhuanantanondh P, Jalayondeja C. Factors associated with reduced risk of musculoskeletal disorders among office workers: a cross-sectional study 2017 to 2020. BMC Public Health 2022 Aug 06;22(1):1503 [FREE Full text] [doi: 10.1186/s12889-022-13940-0] [Medline: 35932005]
5. Global action plan on physical activity 2018-2030: more active people for a healthier world. World Health Organization. 2019. URL: https://www.who.int/publications-detail-redirect/9789241514187 [accessed 2023-10-02]
6. Waters CN, Ling EP, Chu AH, Ng SH, Chia A, Lim YW, et al. Assessing and understanding sedentary behaviour in office-based working adults: a mixed-method approach. BMC Public Health 2016 Apr 27;16:360 [FREE Full text] [doi: 10.1186/s12889-016-3023-z] [Medline: 27117178]
7. Keown MK, Skeaff CM, Perry TL, Haszard JJ, Peddie MC. Device-measured sedentary behavior patterns in office-based university employees. J Occup Environ Med 2018 Dec;60(12):1150-1157. [doi: 10.1097/JOM.0000000000001467] [Medline: 30308622]
8. Wiseman B. The next great disruption is hybrid work—are we ready? Microsoft WorkLab. 2021. URL: https://www.brianheger.com/2021-work-trend-index-annual-report-the-next-great-disruption-is-hybrid-work-are-we-ready-microsoft/ [accessed 2024-04-29]
9. Gilson N, Coenen P, Hallman D, Holtermann A, Mathiassen SE, Straker L. Postpandemic hybrid work: opportunities and challenges for physical activity and public health. Br J Sports Med 2022 Nov;56(21):1203-1204. [doi: 10.1136/bjsports-2022-105664] [Medline: 35710111]
10. Smite D, Moe NB, Hildrum J, Huerta JG, Mendez D. Work-from-home is here to stay: call for flexibility in post-pandemic work policies. J Syst Softw 2023 Jan;195:111552 [FREE Full text] [doi: 10.1016/j.jss.2022.111552] [Medline: 36320721]
11. Loef B, van Oostrom SH, Bosma E, Lifelines Corona Research Initiative, Proper KI. The mediating role of physical activity and sedentary behavior in the association between working from home and musculoskeletal pain during the COVID-19 pandemic. Front Public Health 2022;10:1072030 [FREE Full text] [doi: 10.3389/fpubh.2022.1072030] [Medline: 36530694]
12. Tronco Hernández YA, Parente F, Faghy MA, Roscoe CM, Maratos FA. Influence of the COVID-19 lockdown on the physical and psychosocial well-being and work productivity of remote workers: cross-sectional correlational study. JMIRx Med 2021;2(4):e30708 [FREE Full text] [doi: 10.2196/30708] [Medline: 34898665]
13. Pereira MA, Mullane SL, Toledo MJL, Larouche ML, Rydell SA, Vuong B, et al. Efficacy of the 'Stand and Move at Work' multicomponent workplace intervention to reduce sedentary time and improve cardiometabolic risk: a group randomized clinical trial. Int J Behav Nutr Phys Act 2020 Oct 27;17(1):133 [FREE Full text] [doi: 10.1186/s12966-020-01033-3] [Medline: 33109190]
14. Larisch L, Blom V, Hagströmer M, Ekblom M, Ekblom Ö, Nilsson J, et al. Improving movement behavior in office workers: effects of two multi-level cluster-RCT interventions on mental health. BMC Public Health 2024 Jan 09;24(1):127 [FREE Full text] [doi: 10.1186/s12889-024-17647-2] [Medline: 38195449]
15. Huang Y, Benford S, Blake H. Digital interventions to reduce sedentary behaviors of office workers: scoping review. J Med Internet Res 2019 Feb 07;21(2):e11079 [FREE Full text] [doi: 10.2196/11079] [Medline: 30730294]
16. Buckingham SA, Williams AJ, Morrissey K, Price L, Harrison J. Mobile health interventions to promote physical activity and reduce sedentary behaviour in the workplace: a systematic review. Digit Health 2019;5:2055207619839883 [FREE Full text] [doi: 10.1177/2055207619839883] [Medline: 30944728]
17. Edwardson CL, Yates T, Biddle SJ, Davies MJ, Dunstan DW, Esliger DW, et al. Effectiveness of the Stand More AT (SMArT) work intervention: cluster randomised controlled trial. BMJ 2018 Oct 10;363:k3870 [FREE Full text] [doi: 10.1136/bmj.k3870] [Medline: 30305278]
18. Kitano N, Kai Y, Jindo T, Fujii Y, Tsunoda K, Arao T. Association of domain-specific physical activity and sedentary behavior with cardiometabolic health among office workers. Scand J Med Sci Sports 2022 Aug;32(8):1224-1235. [doi: 10.1111/sms.14165] [Medline: 35426181]
19. Olsen HM, Brown WJ, Kolbe-Alexander T, Burton NW. A brief self-directed intervention to reduce office employees' sedentary behavior in a flexible workplace. J Occup Environ Med 2018 Oct;60(10):954-959. [doi: 10.1097/JOM.0000000000001389] [Medline: 30001255]
20. Wilkerson AH, Elliott CR, McFadden NT, Abutalib N. Feasibility of using mobile standing desks to address sedentary behavior in flexible work environments: a mixed methods study. J Occup Environ Med 2023 May 01;65(5):e273-e278 [FREE Full text] [doi: 10.1097/JOM.0000000000002804] [Medline: 36701795]
21. Parés-Salomón I, Señé-Mir AM, Martín-Bozas F, Loef B, Coffey A, Dowd KP, et al. Effectiveness of workplace interventions with digital elements to reduce sedentary behaviours in office employees: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2024 Apr 19;21(1):41 [FREE Full text] [doi: 10.1186/s12966-024-01595-6] [Medline: 38641816]

XSL•FO
RenderX

22. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, Medical Research Council Guidance. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008 Sep 29;337:a1655 [FREE Full text] [doi: 10.1136/bmj.a1655] [Medline: 18824488]
23. Davis R, Campbell R, Hildon Z, Hobbs L, Michie S. Theories of behaviour and behaviour change across the social and behavioural sciences: a scoping review. Health Psychol Rev 2015;9(3):323-344 [FREE Full text] [doi: 10.1080/17437199.2014.941722] [Medline: 25104107]
24. Kwasnicka D, Dombrowski SU, White M, Sniehotta F. Theoretical explanations for maintenance of behaviour change: a systematic review of behaviour theories. Health Psychol Rev 2016 Sep;10(3):277-296 [FREE Full text] [doi: 10.1080/17437199.2016.1151372] [Medline: 26854092]
25. O'Cathain A, Croot L, Duncan E, Rousseau N, Sworn K, Turner KM, et al. Guidance on how to develop complex interventions to improve health and healthcare. BMJ Open 2019 Aug 15;9(8):e029954 [FREE Full text] [doi: 10.1136/bmjopen-2019-029954] [Medline: 31420394]
26. Coffey A, Parés-Salomón I, Reckman P. Cross-cultural needs analysis: factors for reducing sedentary behaviours among home-office workers. Click2Move Project. URL: https://click2moveproject.com/phases/phase-1/ [accessed 2024-04-08]
27. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. J Adv Nurs 2000 Oct;32(4):1008-1015. [Medline: 11095242]
28. Boulkedid R, Abdoul H, Loustau M, Sibony O, Alberti C. Using and reporting the Delphi method for selecting healthcare quality indicators: a systematic review. PLoS One 2011;6(6):e20476 [FREE Full text] [doi: 10.1371/journal.pone.0020476] [Medline: 21694759]
29. Jünger S, Payne SA, Brine J, Radbruch L, Brearley SG. Guidance on conducting and REporting DElphi studies (CREDES) in palliative care: recommendations based on a methodological systematic review. Palliat Med 2017 Sep 13;31(8):684-706. [doi: 10.1177/0269216317690685] [Medline: 28190381]
30. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
31. Väänänen I, Mas-Alòs S, Vandaele F, Codina-Nadal A, Matas S, Aumatell E, et al. Workplace physical activity practices in real life: a scoping review of grey literature for small- and medium-sized enterprises. Eur J Public Health 2022 Aug 26;32(Suppl 1):i22-i27 [FREE Full text] [doi: 10.1093/eurpub/ckac083] [Medline: 36031820]
32. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
33. Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, et al. GRADE guidelines: 2. Framing the question and deciding on important outcomes. J Clin Epidemiol 2011 Apr;64(4):395-400. [doi: 10.1016/j.jclinepi.2010.09.012] [Medline: 21194891]
34. Gadke DL, Kratochwill TR, Gettinger M. Incorporating feasibility protocols in intervention research. J Sch Psychol 2021 Feb;84:1-18. [doi: 10.1016/j.jsp.2020.11.004] [Medline: 33581765]
35. Mohr DC, Schueller SM, Montague E, Burns MN, Rashidi P. The behavioral intervention technology model: an integrated conceptual and technological framework for eHealth and mHealth interventions. J Med Internet Res 2014 Jun 05;16(6):e146 [FREE Full text] [doi: 10.2196/jmir.3077] [Medline: 24905070]
36. Elamir E. Mean absolute deviation about median as a tool of explanatory data analysis. Int J Res Rev Appl Sci 2012;11:e630 [FREE Full text]
37. Tran TN, Felfernig A, Le VM. An overview of consensus models for group decision-making and group recommender systems. User Model User-adapt Interact 2024;34(3):489-547 [FREE Full text]
38. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
39. Braun V, Clarke V. Toward good practice in thematic analysis: avoiding common problems and be(com)ing a researcher. Int J Transgend Health 2023;24(1):1-6 [FREE Full text] [doi: 10.1080/26895269.2022.2129597] [Medline: 36713144]
40. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport Exerc Health 2019;11(4):589-597 [FREE Full text]
41. Kakar ZU, Rasheed R, Rashid A, Akhter S. Criteria for assessing and ensuring the trustworthiness in qualitative research. Int J Bus Reflect 2023 Dec 31;4(2):150-173. [doi: 10.56249/ijbr.03.01.44]
42. Ahmed SK. The pillars of trustworthiness in qualitative research. J Med Surg Public Health 2024:2 [FREE Full text]
43. Sanders JP, Loveday A, Pearson N, Edwardson C, Yates T, Biddle SJ, et al. Devices for self-monitoring sedentary time or physical activity: a scoping review. J Med Internet Res 2016 May 04;18(5):e90 [FREE Full text] [doi: 10.2196/jmir.5373] [Medline: 27145905]
44. Guitar NA, MacDougall A, Connelly DM, Knight E. Fitbit activity trackers interrupt workplace sedentary behavior: a new application. Workplace Health Saf 2018 May 09;66(5):218-222. [doi: 10.1177/2165079917738264] [Medline: 29121833]
45. Chu AH, Ng SH, Paknezhad M, Gauterin A, Koh D, Brown MS, et al. Comparison of wrist-worn Fitbit Flex and waist-worn ActiGraph for measuring steps in free-living adults. PLoS One 2017;12(2):e0172535 [FREE Full text] [doi: 10.1371/journal.pone.0172535] [Medline: 28234953]

XSL•FO
RenderX

46. Tudor-Locke C, Leonardi C, Johnson WD, Katzmarzyk PT. Time spent in physical activity and sedentary behaviors on the working day: the American time use survey. J Occup Environ Med 2011 Dec;53(12):1382-1387. [doi: 10.1097/JOM.0b013e31823c1402] [Medline: 22104979]
47. Rosenberger ME, Haskell WL, Albinali F, Mota S, Nawyn J, Intille S. Estimating activity and sedentary behavior from an accelerometer on the hip or wrist. Med Sci Sports Exerc 2013 May;45(5):964-975 [FREE Full text] [doi: 10.1249/MSS.0b013e31827f0d9c] [Medline: 23247702]
48. Fanning J, Miller M, Chen S, Davids C, Kershner K, Rejeski W. Is wrist accelerometry suitable for threshold scoring? A comparison of hip-worn and wrist-worn actigraph data in low-active older adults with obesity. J Gerontol A Biol Sci Med Sci 2022 Dec 29;77(12):2429-2434 [FREE Full text] [doi: 10.1093/gerona/glab347] [Medline: 34791237]
49. Bort-Roig J, Martin M, Puig-Ribera A, González-Suárez Á, Martínez-Lemos I, Martori JC, et al. Uptake and factors that influence the use of 'sit less, move more' occupational intervention strategies in Spanish office employees. Int J Behav Nutr Phys Act 2014 Dec 10;11:152 [FREE Full text] [doi: 10.1186/s12966-014-0152-6] [Medline: 25490857]
50. Mailey EL, Rosenkranz SK, Casey K, Swank A. Comparing the effects of two different break strategies on occupational sedentary behavior in a real world setting: a randomized trial. Prev Med Rep 2016 Dec;4:423-428 [FREE Full text] [doi: 10.1016/j.pmedr.2016.08.010] [Medline: 27583200]
51. Mailey EL, Rosenkranz SK, Ablah E, Swank A, Casey K. Effects of an intervention to reduce sitting at work on arousal, fatigue, and mood among sedentary female employees: a parallel-group randomized trial. J Occup Environ Med 2017 Dec;59(12):1166-1171. [doi: 10.1097/JOM.0000000000001131] [Medline: 28816735]
52. Edwardson CL, Maylor BD, Biddle SJ, Clarke-Cornwell AM, Clemes SA, Davies MJ, et al. Participant and workplace champion experiences of an intervention designed to reduce sitting time in desk-based workers: SMART work and life. Int J Behav Nutr Phys Act 2023 Nov 30;20(1):142 [FREE Full text] [doi: 10.1186/s12966-023-01539-6] [Medline: 38037043]
53. Del Pozo-Cruz J, García-Hermoso A, Alfonso-Rosa RM, Alvarez-Barbosa F, Owen N, Chastin S, et al. Replacing sedentary time: meta-analysis of objective-assessment studies. Am J Prev Med 2018 Sep;55(3):395-402. [doi: 10.1016/j.amepre.2018.04.042] [Medline: 30122216]
54. Edwardson CL, Abell L, Clarke-Cornwell A, Dunstan DW, Gray LJ, Healy GN, et al. Implementation and engagement of the SMART Work and Life sitting reduction intervention: an exploratory analysis on intervention effectiveness. Int J Behav Nutr Phys Act 2023 Dec 19;20(1):148. [doi: 10.1186/s12966-023-01548-5] [Medline: 38115044]
55. Kim J, Mizushima R, Nishida K, Morimoto M, Nakata Y. Proposal of a comprehensive and multi-component approach to promote physical activity among Japanese office workers: a qualitative focus group interview study. Int J Environ Res Public Health 2022 Feb 15;19(4):2172 [FREE Full text] [doi: 10.3390/ijerph19042172] [Medline: 35206358]
56. Taylor WC, Suminski RR, Das BM, Paxton RJ, Craig DW. Organizational culture and implications for workplace interventions to reduce sitting time among office-based workers: a systematic review. Front Public Health 2018;6:263 [FREE Full text] [doi: 10.3389/fpubh.2018.00263] [Medline: 30320051]
57. Mackenzie K, Such E, Norman P, Goyder E. Sitting less at work: a qualitative study of barriers and enablers in organisations of different size and sector. BMC Public Health 2019 Jul 04;19(1):884 [FREE Full text] [doi: 10.1186/s12889-019-7148-8] [Medline: 31272484]
58. Beckel JL, Fisher GG. Telework and worker health and well-being: a review and recommendations for research and practice. Int J Environ Res Public Health 2022 Mar 24;19(7):3879 [FREE Full text] [doi: 10.3390/ijerph19073879] [Medline: 35409563]
59. Cox E, Walker S, Edwardson CL, Biddle SJ, Clarke-Cornwell AM, Clemes SA, et al. The cost-effectiveness of the SMART work and life intervention for reducing sitting time. Int J Environ Res Public Health 2022 Nov 11;19(22):14861 [FREE Full text] [doi: 10.3390/ijerph192214861] [Medline: 36429578]
60. Munir F, Miller P, Biddle SJ, Davies MJ, Dunstan DW, Esliger DW, et al. A cost and cost-benefit analysis of the stand more at work (SMArT work) intervention. Int J Environ Res Public Health 2020 Feb 13;17(4):1214 [FREE Full text] [doi: 10.3390/ijerph17041214] [Medline: 32070034]
61. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 2021 Sep 30;374:n2061 [FREE Full text] [doi: 10.1136/bmj.n2061] [Medline: 34593508]
62. Olsen HM, Brown WJ, Kolbe-Alexander T, Burton NW. Flexible work: the impact of a new policy on employees' sedentary behavior and physical activity. J Occup Environ Med 2018 Jan;60(1):23-28. [doi: 10.1097/JOM.0000000000001190] [Medline: 29023342]

Abbreviations

BIT: behavioral intervention technology
COREQ: Consolidated Criteria for Reporting Qualitative Research
MAD-M: mean absolute deviation from the median
PA: physical activity
SB: sedentary behavior

XSL•FO
RenderX

Edited by A Kushniruk; submitted 24.04.24; peer-reviewed by A Wagg, SA Alothman; comments to author 09.01.25; revised version received 12.02.25; accepted 04.03.25; published 08.04.25.

Please cite as:
Parés-Salomón I, Vaqué-Crusellas C, Coffey A, Loef B, Proper KI, Señé-Mir AM, Puig-Ribera A, Dowd KP, Bort-Roig J
Development of Digital Strategies for Reducing Sedentary Behavior in a Hybrid Office Environment: Modified Delphi Study
JMIR Hum Factors 2025;12:e59405
URL: https://humanfactors.jmir.org/2025/1/e59405
doi:10.2196/59405
PMID:40198908

©Iris Parés-Salomón, Cristina Vaqué-Crusellas, Alan Coffey, Bette Loef, Karin I Proper, Anna M Señé-Mir, Anna Puig-Ribera, Kieran P Dowd, Judit Bort-Roig. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 08.04.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Innovative Digital Cognitive Behavioral Treatment for Insomnia Disorder in Adults (dCBT-i): Framework Development

Emilie Stern[1,2], MSc; Severine Brune[3], MSc; Stéphane Mouchabac[4,5], MD; Anthony Dubroc[6], PharmD; Charles de la Personne[6], PharmD; Pierre Alexis Geoffroy[1,7,8], MD, PhD

[1]Centre ChronoS, GHU Paris psychiatrie & neurosciences, Paris, France

[2]Université Paris Cité, Laboratoire de Psychopathologie et Processus de Santé, Boulogne-Billancourt, France

[3]Centre de l'insomnie PsySomnie, Toulouse, France

[4]Centre National de la Recherche Scientifique (CNRS), Institut du Cerveau-Paris Brain Institute, Institut National de la Santé et de la Recherche Médicale U112, iCrin, Inserm, Paris, France

[5]Department of Psychiatry, Saint Antoine Hospital, DMU Neurosciences, Assistance Publique-Hôpitaux de Paris (AP-HP), Sorbonne Université, Paris, France

[6]Di&Care, Loos, France

[7]Département de psychiatrie et d'addictologie, Assistance Publique-Hôpitaux de Paris (AP-HP), DMU Neurosciences, Hôpital Bichat-Claude-Bernard, Paris, France

[8]NeuroDiderot, F-75019, Université Paris Cité, Inserm, Paris, France

Corresponding Author:
Emilie Stern, MSc
Centre ChronoS
GHU Paris psychiatrie & neurosciences
4 avenue de la Porte de Saint Ouen
Paris, 75014
France
Phone: 33 0181699179
Email: e.stern@ghu-paris.fr

Abstract

Background: Chronic insomnia, or insomnia disorder, is a major health issue with a prevalence of up to 15%. The recommended first-line treatment is cognitive behavioral therapy for insomnia (CBT-i), which, unfortunately, remains insufficiently accessible. Digitalization has the potential to reduce health care access inequalities by offering more flexible and accessible care options. Digital CBT-i (dCBT-i) has been shown to be as effective as in-person CBT-i, highlighting its potential for broader implementation.

Objective: This study aimed to develop an evidence-based dCBT-i program grounded in theoretical and clinical knowledge, designed for efficient integration into health care systems, and to establish it as the first prescribed digital treatment in France.

Methods: The program was constructed based on validated CBT-i theory and practice, incorporating the latest scientific data on CBT for insomnia. It was designed as a robust multicomponent therapy, integrating an initial standardized assessment and daily intelligent adaptation to enable digital phenotyping and provide personalized treatment.

Results: We developed an innovative digital solution that combines scientific rigor with practical application. The program includes a standardized initial evaluation and dynamic personalization through intelligent algorithms. These features allow for the adaptation of therapy based on patient progress and needs, ensuring individualized care.

Conclusions: The development of this dCBT-i program represents a significant milestone in digital health care, offering a scalable solution to the accessibility challenges of traditional CBT-i. Future steps involve conducting clinical studies to further evaluate its effectiveness and optimize its implementation within health care systems.

(JMIR Hum Factors 2025;12:e70193) doi:10.2196/70193

KEYWORDS

digital program; insomnia; integrative therapies; CBT-i; nonpharmacological intervention

XSL•FO
RenderX

Introduction

Insomnia is a major health complaint, with a prevalence up to 50% in primary care patients undergoing current treatment for either independent insomnia or insomnia comorbid with another disease, posing a risk of worsening comorbid medical conditions without addressing insomnia [1]. Chronic insomnia, or insomnia disorder, has an estimated prevalence of approximately 9% to 15% based on criteria from the Diagnostic and Statistical Manual of Mental Disorders (Fourth Edition; DSM-IV) and the International Classification of Sleep Disorders [2]. Moreover, insomnia affects 85% of patients with a mood disorder during depressive phases, and more than one-half of the patients will retain insomnia symptoms after remission [3,4]. Insomnia disorder is defined by sleep dissatisfaction for at least 3 months and 3 nights per week despite adequate opportunity for sleep, with difficulty initiating sleep, maintaining sleep continuity, early morning awakening, or poor sleep quality and with daytime dysfunction and significant impairment in areas of functioning [5,6].

The latest European Insomnia Guideline suggests that cognitive behavioral therapy for insomnia (CBT-i) is the reference and first-line treatment for insomnia disorder, with sustained improvements over time [7], and multicomponent CBT-i is strongly recommended [8]. Another important aspect of recent research involves distinguishing between response and remission rates. Response is typically defined as a significant reduction in Insomnia Severity Index (ISI) scores, while remission is generally defined as achieving a final ISI score ≤7. Studies have shown that 70% to 80% of patients respond positively to CBT-I, with about 50% of patients showing a clinically meaningful reduction of symptoms, while remission was attained by approximately 40% of patients receiving CBT-i and outcomes were often improved at the 6-month follow-up for both response and remission [7,9-12]. In addition, patients who do not respond well to CBT-i often have insomnia associated with a mood disorder and with insomnia complaints often masking disruptions in the biological clock or a state of “hyperarousal” [1,7,13]. However, CBT-i has been shown to be effective in treating insomnia comorbid with a mood disorder [14] and to also have an effect on psychiatric symptoms comorbid with insomnia disorder [15,16], and mindfulness techniques (often integrated into the emotional regulation component of CBT-i) were found to reduce hyperarousal [17-20].

CBT-i has largely been proven effective but is not sufficiently accessible as they are too few CBT-i therapists available and the treatment can be expensive [7,21-24]. Digital CBT-i (dCBT-i) could make treatment more accessible. Digitalization of CBT-i was first proposed in the 1990s, and the first proposal to introduce dCBT-i into the health care system as a preliminary step before in-person orientation was made in 2009 [25]. There are now numerous programs with proven efficacy and technological advances that allow for adaptation and personalization of content, while some obstacles still reside, including patient adherence and the perception of individualization [21-24,26-28]. Some programs are fully automated, while others are guided (mixing human and digital interaction) or only used as a tool to support in-person care [29].

To make CBT-i more accessible, fully automated treatments seem to be more relevant, but they are considered less engaging and associated with poorer adherence. The latest technological advances, encouragement for therapeutic adherence, and digital phenotyping to create a sense of treatment personalization should allow innovative, fully automated programs to be as engaging as in-person care [21-24,26-28,30]. dCBT-i has already been proven effective at reducing insomnia severity, with many studies reporting significant reductions in ISI [31] scores after treatment compared with control groups. For instance, a study reported an average difference in ISI score of 7.6 points, while a meta-analysis of 33 randomized controlled trials (RCTs) found a mean difference of 5.0 points between treatment and control groups [32,33]. dCBT-i was also found to have a significant effect on quality of life and cognitive functions, and efficacy was also demonstrated for mild insomnia or insomnia comorbid to a mood disorder [34-38]. dCBT-i was also found to have a reducing and protective effect on comorbid psychiatric symptoms [39-44] and a preventive effect on sleep disturbances after a stressor such as during the COVID-19 pandemic [45].

Digital interventions could significantly reduce access inequalities by introducing less stigma, being more flexible, and being less expensive for the patient than in-person therapy [23,34,46]. dCBT-i could then be a perfect solution to the challenge of access, as it was proven to have the same efficacy as in-person CBT-i. As highlighted in a recent meta-analysis, dCBT-i may be a viable alternative when in-person care is not available [47]. However, some issues still remain. First, many interventions exist, but less than 15% of the available “care apps” are based on strong theoretical models [23,48]. Second, adherence is an important challenge with this type of care, as less than 30% of users participate in digital programs in their entirety [30,49,50]. Finally, evidence-based, valid programs are not always made available to the public, as they must be implemented in the health care system [51].

In this context, our goal was to describe the 2-year development of an evidence-based digital program that could retain CBT-i efficacy while being accessible, engaging, and ready to be implemented in the French health care system, as already accomplished in other countries [52,53]. We further sought to develop an innovative digital solution grounded in the latest scientific data on CBT-i, with a robust multicomponent therapy approach. This solution integrates an initial standardized assessment and daily intelligent adaptation, where digital phenotyping and algorithms adjust the treatment based on daily monitoring of the patients’ behaviors, optimizing the treatment in real time and providing personalized solutions. This cutting-edge treatment is grounded in the latest scientific data and uses technological tools to offer individualized and evolving treatment based on a robust standardized assessment of the patients’ sleep complaints. Furthermore, algorithmic possibilities and artificial intelligence (AI) should allow the treatment to evolve continuously and adapt to the changes in the patients’ daily lives, making the treatment even more patient-centered and tailored to its specific needs than most traditional CBT-i protocols.

XSL•FO
RenderX

Methods

Program Overview

We developed a digital program tailored for French patients diagnosed with insomnia disorder by their medical practitioner, integrating data from the evidence-based CBT-i literature, user research, and clinical expertise. The development process followed the latest guidelines provided by various authors specializing in digital intervention issues and the treatment of insomnia disorder [7,8]. To make the app useful and safe, we followed the ASPECTS (Actionable, Secure, Professional, Evidence-based, Customizable, Transparent) guidelines and the various steps used in evaluating an app [54].

Our method should allow prescribers to integrate into their care useful data and information about their patient from our intervention such as a sleep diary and daily patient habits.

The French guidelines for safety and privacy were followed, and the data will be secured and protected according to those guidelines.

Various experts (eg, legal advisors, scientific and medical professionals, sleep practitioners, psychologist specializing in CBT-i, digital and user experience design specialists) whose credibility and capability were assessed were included in the process.

We conducted a deep literature review, and the entire program was based on CBT-i as described in the literature. In addition, dCBT-i is already evidence-based [7].

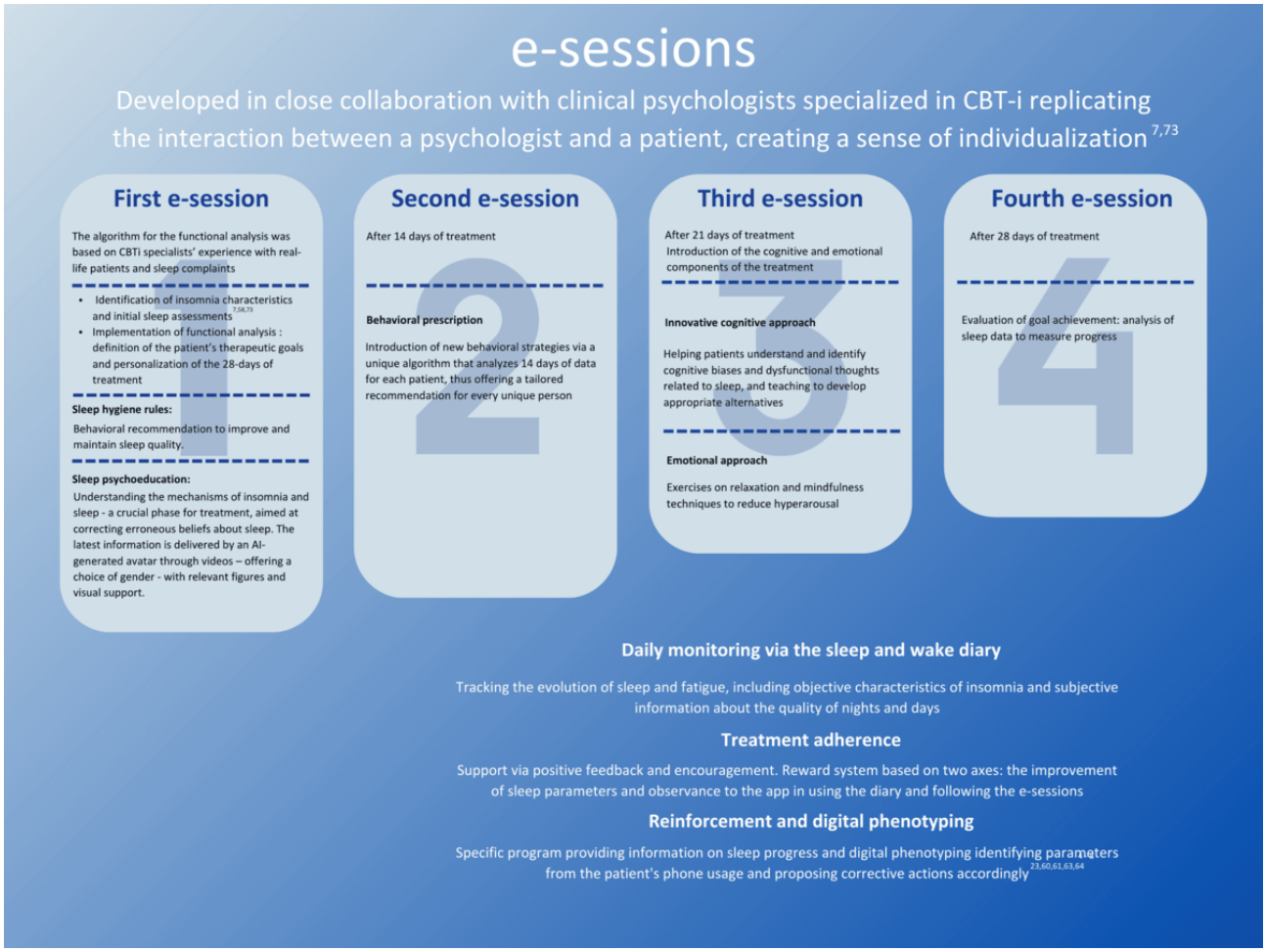
The program is flexible and adaptable to its users, and the app should be easy enough to be engaging, which will be determined though user research and planned studies. The program provides transparency to users about how it works and collects data. Electronic sessions (e-sessions) mimic real-life CBT-i consultations where the patient is introduced to fundamental tools and information, and every recommendation is explained well to the patient in order to develop good treatment adherence.

The program was built according to known psychological models of change in order to “reflect the psychological chronology of the change process” [55].

Finally, the program is personalized for each patient using numerous tests and algorithms, which define a tailored treatment pathway for the patient's insomnia.

The digital medical device Noxxea will undergo clinical validation to demonstrate its effectiveness and safety of use with patients with insomnia. The method incorporates traditional pillars of insomnia management through several e-sessions designed to replicate the interaction between a psychologist and a patient (Figure 1).

Figure 1. Key components of the digital insomnia treatment program. AI: artificial intelligence; CBT-i: cognitive behavioral therapy for insomnia; e-session: electronic session.



XSL•FO
RenderX

The first e-session involves identification of the patient's insomnia characteristics and initial sleep assessments, based on the requirements of an in-person clinical evaluation and traditional theoretical models of insomnia [7,56]. Implementation of functional analysis involves personalization of a 28-day treatment based on this analysis and defining the patient's therapeutic goals. The algorithm was based on CBT-i specialists' experience with real-life patients and sleep complaints.

Daily monitoring via the sleep diary includes use of the sleep and wake diary. The diary tracks the evolution of sleep and fatigue, including objective characteristics of insomnia and subjective information about the quality of nights and days.

Sleep hygiene rules include behavioral recommendations to improve and maintain sleep quality.

Sleep psychoeducation helps to understand the mechanisms of insomnia and sleep. This is a crucial phase for treatment, aimed at correcting erroneous beliefs about sleep. The latest information is delivered by an AI-generated avatar in videos with relevant figures and visual support.

The second e-session (after 14 days) is the mid-treatment assessment, which evaluates the first 14 days of treatment. The behavioral prescription introduces new behavioral strategies via a unique algorithm that analyzes 14 days of data for each patient, thus offering a tailored recommendation for every unique person.

The third e-session (after 21 days) involves cognitive and emotional restructuring. Although this approach is quite rare in digital formats today, it remains essential for CBT-i. Thought management was finally made possible in dCBT-i through a specific module based on Beck's cognitive restructuring techniques [57]. It enables patients to understand cognitive biases and dysfunctional thoughts, identify automatic thoughts related to sleep, and learn to respond to them with appropriate alternatives. Another module offers exercises on relaxation and mindfulness techniques to reduce hyperarousal.

The final assessment (after 28 days) evaluates achieved goals, including an analysis of sleep data to measure progress.

To ensure adherence to treatment, a reward system was developed based on 2 axes: improvement of sleep parameters and adherence to entering the agenda and e-sessions in the app.

Regarding reinforcement and digital phenotyping, throughout the treatment, the patient receives reinforcement through a specific program providing information on sleep progress, thus facilitating engagement. An essential aspect of the app is the use of digital phenotyping, which identifies parameters from the patient's phone use and proposes corrective actions accordingly.

The inclusion of digital phenotyping is a key element of the method that allows for further adaptation and personalization of treatment and makes it as close to in-person care as possible. In 1982, evolutionary biologist Richard Dawkins highlighted the arbitrary nature of many phenotypic delineations and introduced the concept of the "extended phenotype" [58]. This concept suggests that phenotypes should include all effects a gene has on its environment, not just biological processes like protein biosynthesis or tissue growth. Both animals and humans modify their surroundings, and these modifications, along with associated behaviors, are expressions of their genomes, forming part of their extended phenotype. Dawkins used the dam-building behavior of beavers as an example of an extended phenotype. With the increasing integration of personal technology into human lives, a novel concept called digital phenotyping was introduced in 2015 by Jain et al [59] and further developed by Torous et al [23,60,61] in psychiatry. Digital phenotyping involves real-time collection of data and markers of human behavior to identify the "digital signature of a pathology." It assumes behaviors can be quantified through data from connected tools like smartphones, digital sensors, and wearable devices to derive an "e-semiology" [62]. Two types of data are considered. "Passive" data are automatically collected in real time without user input, using tools like accelerometers and GPS to detect psychomotor retardation or motor agitation. Deviations in GPS coordinates from typical travel patterns, such as reduced activity, increased time spent at home, or chaotic journeys to unusual locations, may indicate pathological behaviors. Mobile phone–based sensing software measures heart rate, heart rate variability, galvanic skin conductance, and blood pressure, which are regarded as potential biomarkers of certain psychiatric conditions. Supplementing this with "active" data is essential, as metrics like screen time, number of calls, and usage hours provide valuable indicators of psychosocial functioning [23,60,61]. Additionally, analyzing language (verbal and paraverbal) and emotional arousal offers insights into the valence of messages during social interactions, indicating the user's mood and mental state. Digital phenotyping can also help minimize cognitive biases, such as confirmation or self-serving biases, by collecting objective information through passive data, for instance. This approach is highly promising for mental disorders and allows for optimal therapeutic personalization, almost in real time, and enables the enhancement of patient empowerment [63,64].

Noxxea Development

Several steps described here were carried out to create this innovative solution based on the latest scientific data.

Treatment development was first discussed with experts in the field in order to take patient needs into account from the beginning. As a result, there was a dynamic interplay between scientific research and product development.

An acceptability study validated certain ergonomic aspects of the app and the usefulness of generative AI for sleep solutions (M Villechevrolle, unpublished data, May 2025). A study involving 582 users was conducted to gather their concerns and feedback. It used the script method, allowing for both qualitative and quantitative analysis of user reactions. The findings were then integrated into the ongoing development of the app to enhance user trust and ensure data privacy.

The most important aspect of the development was to remain as close to in-person treatment as possible.

Translating CBT-i into a digital format requires maintaining the relational aspect of therapy. Indeed, in psychotherapy, a

XSL•FO
RenderX

form of therapeutic alliance is essential, and it is not yet fully known how that translates to a digital program [30]. Flexibility and personalization of treatment seem fundamental—for instance, dCBT-i with a therapist or direct context with a practitioner or therapist-guided dCBT-i had greater effects on sleep parameters and treatment adherence than dCBT-i without a therapist [38,65-67]. Programs are effective even without support, but better results with digital programs are found when the users have a sense of support and individualized care [22,68-71]. This may also help explain why a recent network meta-analysis of various CBT-i delivery methods found that both guided and unguided dCBT-i achieved medium effect sizes, while smartphone-based CBT-i did not show significant effects compared with the control group [7,47].

Therefore, a significant effort was made to develop a form of communication within the app that was as close to an in-person consultation as possible. The participant can choose the avatar's gender, and support is offered through positive feedback and encouragement via the app. A reward system was developed to engage the patient and help him or her identify changes and improvements. Every item of text and dialogue was written by clinical psychologists specializing in CBT-i in the same way a therapist would talk to his or her patient, and the e-sessions were developed to recreate a sense of individualization and the 4 to 8 validated sessions of CBT-i [7,72].

Nonetheless, we prioritized chronobiological regulation with emotional and cognitive aspects to stay true to the description and validation of CBT-i in the literature, which does not require the same relational demands in terms of therapeutic alliance as in-person psychotherapy (which typically requires a higher number of sessions and is often more suitable for resistant insomnia). The program was thus developed according to the components described in sleep recommendations and many studies using this form of treatment for sleep disorders [7,72,73].

Psychoeducation and Sleep Hygiene

The program must address the sleep hygiene "rules." Therefore, explanations about recommendations, habits, sleep functioning (roles and functions as well as circadian patterns), and changes in sleep patterns across the lifespan are given and integrated into the treatment.

Behavioral Treatment

A sleep calendar is introduced at the beginning of the treatment to collect at least 10 nights of sleep data in the first 2 weeks of treatment, allowing the calculation of individualized sleep parameters and habits. Sleep restriction and stimulus control are the main behavioral strategies used in the program. These components were developed in close collaboration with the scientific committee and translated into a digital format within the app in accordance with established CBT-i guidelines. Sleep restriction is thus personalized for each patient according to an individualized sleep calendar, and stimulus control is personalized using patient-reported data, the chronotype, and routines assessed at the start of the program. Personalized video content targeting chronobiological rhythms and behavioral habits further supports the tailoring of treatment to each patient's specific needs.

Relaxation Component

Many different relaxation strategies as well as mindfulness meditations are introduced within treatment as ways to reduce somatic tension, arousal, or intrusive thoughts. Patients are encouraged to explore them during their treatment.

Cognitive Component

Cognitive strategies are introduced as ways to identify and challenge misconceptions and faulty beliefs about sleep and state during the day. Patients are introduced to the concepts of cognitive control at bedtime and cognitive restructuring to learn to challenge thoughts. They are encouraged to explore these strategies within the framework of behavioral change.

Developing this program was challenging due to the need to incorporate sleep medicine recommendations using a digital intervention while designing it to closely resemble an individualized care format. The presented treatment not only offers sleep education, which is predominantly present in therapeutic education programs in institutions or digital platforms, but also integrates an individualized functional analysis adapted for digital use. It guides the patient toward a form of cognitive therapy similar to that provided in CBT-i treatment.

Results

The Noxxea program is now fully developed, and several steps are planned for its clinical evaluation, including an RCT scheduled for 2025 that will evaluate the program against a waiting list control group.

Digital Phenotyping Research

Collaboration with the mathematics department at the Pasteur Institute and the Institut National de la Santé et de la Recherche Médicale (INSERM) research laboratory (Decision and Bayesian Computation, Computational Biology & Neuroscience Department, Centre National de la Recherche Scientifique [CNRS] Unité Mixte de Recherche [UMR] 3571 – Jean Baptiste MASSON) will allow analysis of data to enhance the treatment's algorithms and improve efficacy at predicting and preventing insomnia relapse. This work will mainly focus on digital phenotyping: Approximately 50 data points will be collected and analyzed to aid with the detection, prediction, and prevention of chronic insomnia relapses. The digital phenotyping initiative is central to the innovative dCBT-i and has the following 3 main objectives: validate the initial selection of the 4 key tracking data points chosen for the first version of the treatment; analyze and evaluate the 70 passive and active data points collected to determine interaction models that enable detection of symptoms, prediction, and prevention of insomnia relapse, as well as the isolation of additional data that could enhance the efficacy of a new version of the treatment and improve adherence; and test a second version of the treatment by validating the selection of new tracking data points.

Effectiveness of Real-World Implementation and RCT

We will evaluate the efficacy and compare it with that of in-person CBT-i in the same population. The WHAT (work,

XSL•FO
RenderX

help, accessible, trusted) criteria can be used to evaluate a digital intervention, and we will need to evaluate the following [51,74]:

- Work: To evaluate efficacy, we plan to conduct an RCT.
- Help: A clinically significant effect in the target population would help conclude that it can be a useful addition to in-person CBT-i.
- Accessible: Technical quality will be continuously assessed (by professionals and users).
- Trusted: Risk management and adherence to national and international standards and guidelines will be assessed and continuously updated to the latest requirements.

Discussion

This dCBT-i treatment is built according to European recommendations and guidelines and follows closely what is done in person. The goal is not to replace in-person care but to offer an alternative when in-person care is not available given that not all individuals with insomnia disorder can receive the recommended treatment in France. Ideally, it will be covered by the national health insurance system as it is a first-line medical treatment for a major health issue.

The digital form of insomnia disorder treatment incorporates validated and recommended behavioral, cognitive, and emotional techniques for insomnia, closely mirroring the CBT-i treatment offered traditionally while being even more tailored to patients' individual needs. It is meant to be a medical treatment prescribed by a health care professional following an evaluation and a precise diagnosis, addressing the dual challenge of considering therapeutic adherence and support in its development while individualizing the behavioral and cognitive components of the treatment for greater effectiveness. The innovative program presented here followed those demands, as well as safety and security guidelines for digital health interventions while being grounded in the latest technological advances and scientific data on CBT for insomnia. It would offer access to an evidence-based multicomponent treatment and is intended to be integrated into the health care system as such. In line with expert recommendations, the Noxxea solution could ideally be part of a stepped care model that adapts to the severity of the condition and the patient's needs, with dCBT-i proposed as a first step delivered through multiple pathways, such as being prescribed by a doctor or accessed directly by the patient [51].

Authors' Contributions

All authors contributed to the development of the method. ES drafted the article, and all authors reviewed, revised, and approved the final manuscript.

Conflicts of Interest

ES, SB, SM, and PAG received consulting fees from Di&Care, which is developing the digital cognitive behavioral therapy for insomnia (dCBT-i). CP and AD are employed by Di&Care. All authors agreed to disclose these relationships.

References

1. Perlis ML, Posner D, Riemann D, Bastien CH, Teel J, Thase M. Insomnia. Lancet 2022 Sep 24;400(10357):1047-1060. [doi: 10.1016/S0140-6736(22)00879-0] [Medline: 36115372]
2. Ohayon MM, Reynolds CF. Epidemiological and clinical relevance of insomnia diagnosis algorithms according to the DSM-IV and the International Classification of Sleep Disorders (ICSD). Sleep Med 2009 Oct;10(9):952-960 [FREE Full text] [doi: 10.1016/j.sleep.2009.07.008] [Medline: 19748312]
3. Geoffroy PA, Micoulaud Franchi J, Maruani J, Philip P, Boudebesse C, Benizri C, et al. Clinical characteristics of obstructive sleep apnea in bipolar disorders. J Affect Disord 2019 Feb 15;245:1-7. [doi: 10.1016/j.jad.2018.10.096] [Medline: 30359809]
4. Henry C, Etain B, Godin O, Dargel AA, Azorin J, Gard S, FACE-BD group. Bipolar patients referred to specialized services of care: Not resistant but impaired by sub-syndromal symptoms. Results from the FACE-BD cohort. Aust N Z J Psychiatry 2015 Oct;49(10):898-905. [doi: 10.1177/0004867415585582] [Medline: 25991763]
5. American Psychiatric Association, DSM-5 Task Force. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. Washington, DC: American Psychiatric Publishing; 2013.
6. Sateia M. International classification of sleep disorders, third edition (ICSD-3). Darien, IL: American Academy of Sleep Medicine; 2014.
7. Riemann D, Espie CA, Altena E, Arnardottir ES, Baglioni C, Bassetti CLA, et al. The European Insomnia Guideline: an update on the diagnosis and treatment of insomnia 2023. J Sleep Res 2023 Dec;32(6):e14035 [FREE Full text] [doi: 10.1111/jsr.14035] [Medline: 38016484]
8. Edinger JD, Arnedt JT, Bertisch SM, Carney CE, Harrington JJ, Lichstein KL, et al. Behavioral and psychological treatments for chronic insomnia disorder in adults: an American Academy of Sleep Medicine clinical practice guideline. J Clin Sleep Med 2021 Feb 01;17(2):255-262 [FREE Full text] [doi: 10.5664/jcsm.8986] [Medline: 33164742]
9. Morin C, Hauri P, Espie C, Spielman A, Buysse D, Bootzin R. Nonpharmacologic treatment of chronic insomnia. An American Academy of Sleep Medicine review. Sleep 1999 Dec 15;22(8):1134-1156. [doi: 10.1093/sleep/22.8.1134] [Medline: 10617176]

XSL•FO
RenderX

10. Morin CM, Vallières A, Guay B, Ivers H, Savard J, Mérette C, et al. Cognitive behavioral therapy, singly and combined with medication, for persistent insomnia: a randomized controlled trial. JAMA 2009 May 20;301(19):2005-2015 [FREE Full text] [doi: 10.1001/jama.2009.682] [Medline: 19454639]
11. Morin CM, Edinger JD, Beaulieu-Bonneau S, Ivers H, Krystal AD, Guay B, et al. Effectiveness of sequential psychological and medication therapies for insomnia disorder: a randomized clinical trial. JAMA Psychiatry 2020 Nov 01;77(11):1107-1115 [FREE Full text] [doi: 10.1001/jamapsychiatry.2020.1767] [Medline: 32639561]
12. Morin CM, Benca R. Chronic insomnia. Lancet 2012 Mar 24;379(9821):1129-1141. [doi: 10.1016/S0140-6736(11)60750-2] [Medline: 22265700]
13. Geoffroy PA, Palagini L. Biological rhythms and chronotherapeutics in depression. Prog Neuropsychopharmacol Biol Psychiatry 2021 Mar 02;106:110158. [doi: 10.1016/j.pnpbp.2020.110158] [Medline: 33152388]
14. Maruani J, Stern E, Boiret C, Leseur J, Romier A, Lejoyeux M, et al. Predictors of cognitive behavioral therapy for insomnia (CBT-I) effects in insomnia with major depressive episode. Psychiatry Res 2023 Nov;329:115527. [doi: 10.1016/j.psychres.2023.115527] [Medline: 37839317]
15. Carney C, Edinger J, Kuchibhatla M, Lachowski A, Bogouslavsky O, Krystal A, et al. Cognitive behavioral insomnia therapy for those with insomnia and depression: a randomized controlled clinical trial. Sleep 2017 Apr 01;40(4):A [FREE Full text] [doi: 10.1093/sleep/zsx019] [Medline: 28199710]
16. Taylor DJ, Pruiksma KE. Cognitive and behavioural therapy for insomnia (CBT-I) in psychiatric populations: a systematic review. Int Rev Psychiatry 2014 Apr 03;26(2):205-213. [doi: 10.3109/09540261.2014.902808] [Medline: 24892895]
17. Baglioni C, Bostanova Z, Bacaro V, Benz F, Hertenstein E, Spiegelhalder K, et al. A systematic review and network meta-analysis of randomized controlled trials evaluating the evidence base of melatonin, light exposure, exercise, and complementary and alternative medicine for patients with insomnia disorder. J Clin Med 2020 Jun 22;9(6):1949 [FREE Full text] [doi: 10.3390/jcm9061949] [Medline: 32580450]
18. Baron KG, Hooker S. Next steps for patients who fail to respond to cognitive behavioral therapy for insomnia (CBT-I): the perspective from behavioral sleep medicine psychologists. Curr Sleep Medicine Rep 2017 Oct 26;3(4):327-332. [doi: 10.1007/s40675-017-0096-x]
19. Crane RS, Brewer J, Feldman C, Kabat-Zinn J, Santorelli S, Williams JMG, et al. What defines mindfulness-based programs? The warp and the weft. Psychol Med 2017 Apr;47(6):990-999. [doi: 10.1017/S0033291716003317] [Medline: 28031068]
20. Vanhuffel H, Rey M, Lambert I, Da Fonseca D, Bat-Pitault F. [Contribution of mindfulness meditation in cognitive behavioral therapy for insomnia]. Encephale 2018 Apr;44(2):134-140. [doi: 10.1016/j.encep.2016.12.001] [Medline: 28213988]
21. Erten Uyumaz B, Feijs L, Hu J. A review of digital cognitive behavioral therapy for insomnia (CBT-I apps): are they designed for engagement? Int J Environ Res Public Health 2021 Mar 12;18(6):1 [FREE Full text] [doi: 10.3390/ijerph18062929] [Medline: 33809308]
22. Luik AI, van der Zweerde T, van Straten A, Lancee J. Digital delivery of cognitive behavioral therapy for insomnia. Curr Psychiatry Rep 2019 Jun 04;21(7):50 [FREE Full text] [doi: 10.1007/s11920-019-1041-0] [Medline: 31161406]
23. Stern E, Micoulaud Franchi J, Dumas G, Moreira J, Mouchabac S, Maruani J, et al. How can digital mental health enhance psychiatry? Neuroscientist 2023 Dec;29(6):681-693. [doi: 10.1177/10738584221098603] [Medline: 35658666]
24. Wickwire E. The value of digital insomnia therapeutics: what we know and what we need to know. J Clin Sleep Med 2019 Jan 15;15(1):11-13 [FREE Full text] [doi: 10.5664/jcsm.7558] [Medline: 30621849]
25. Espie CA. "Stepped care": a health technology solution for delivering cognitive behavioral therapy as a first line insomnia treatment. Sleep 2009 Dec;32(12):1549-1558 [FREE Full text] [doi: 10.1093/sleep/32.12.1549] [Medline: 20041590]
26. Kang S, Kim Y. Cognitive behavioral therapy for insomnia in the digital age. Adv Exp Med Biol 2019;1192:629-641. [doi: 10.1007/978-981-32-9721-0_30] [Medline: 31705516]
27. Manber R, Simpson N, Gumport N. Perspectives on increasing the impact and reach of CBT-I. Sleep 2023 Dec 11;46(12):1. [doi: 10.1093/sleep/zsad168] [Medline: 37903637]
28. Xu Z, Anderson KN. Real-world evaluation of digital CBT for insomnia in the primary care setting – many should not log on to doze off. tCBT 2019 Aug 05;12:1. [doi: 10.1017/s1754470x19000242]
29. Luik AI, Kyle SD, Espie CA. Digital cognitive behavioral therapy (dCBT) for insomnia: a state-of-the-science review. Curr Sleep Med Rep 2017;3(2):48-56 [FREE Full text] [doi: 10.1007/s40675-017-0065-4] [Medline: 28553574]
30. Stern E, Breton Z, Alexaline M, Geoffroy PA, Bungener C. Redefining the relationship in digital care: a qualitative study of the Digital Therapeutic Alliance. Encephale 2024 May 08:1 [FREE Full text] [doi: 10.1016/j.encep.2024.02.011] [Medline: 38724431]
31. Morin CM. Insomnia Severity Index (ISI). Washington, DC: APA PsycTests; 1993.
32. Schuffelen J, Maurer LF, Lorenz N, Rötger A, Pietrowsky R, Gieselmann A. The clinical effects of digital cognitive behavioral therapy for insomnia in a heterogenous study sample: results from a randomized controlled trial. Sleep 2023 Nov 08;46(11):1 [FREE Full text] [doi: 10.1093/sleep/zsad184] [Medline: 37428712]
33. Soh HL, Ho RC, Ho CS, Tam WW. Efficacy of digital cognitive behavioural therapy for insomnia: a meta-analysis of randomised controlled trials. Sleep Med 2020 Nov;75:315-325. [doi: 10.1016/j.sleep.2020.08.020] [Medline: 32950013]

XSL•FO
RenderX

34. Darden M, Espie CA, Carl JR, Henry AL, Kanady JC, Krystal AD, et al. Cost-effectiveness of digital cognitive behavioral therapy (Sleepio) for insomnia: a Markov simulation model in the United States. Sleep 2021 Apr 09;44(4):1. [doi: 10.1093/sleep/zsaa223] [Medline: 33151330]
35. Denis D, Eley TC, Rijsdijk F, Zavos HM, Keers R, Espie CA, et al. Is digital cognitive behavioural therapy for insomnia effective in treating sub-threshold insomnia: a pilot RCT. Sleep Med 2020 Feb;66:174-183. [doi: 10.1016/j.sleep.2019.10.007] [Medline: 31901759]
36. Espie CA, Emsley R, Kyle SD, Gordon C, Drake CL, Siriwardena AN, et al. Effect of digital cognitive behavioral therapy for insomnia on health, psychological well-being, and sleep-related quality of life: a randomized clinical trial. JAMA Psychiatry 2019 Jan 01;76(1):21-30 [FREE Full text] [doi: 10.1001/jamapsychiatry.2018.2745] [Medline: 30264137]
37. Kyle S, Hurry M, Emsley R, Marsden A, Omlin X, Juss A, et al. The effects of digital cognitive behavioral therapy for insomnia on cognitive function: a randomized controlled trial. Sleep 2020 Sep 14;43(9):a. [doi: 10.1093/sleep/zsaa034] [Medline: 32128593]
38. Leerssen J, Lakbila-Kamal O, Dekkers LMS, Ikelaar SLC, Albers ACW, Blanken TF, et al. Treating insomnia with high risk of depression using therapist-guided digital cognitive, behavioral, and circadian rhythm support interventions to prevent worsening of depressive symptoms: a randomized controlled trial. Psychother Psychosom 2022;91(3):168-179 [FREE Full text] [doi: 10.1159/000520282] [Medline: 34872087]
39. Blom K, Jernelöv S, Kraepelien M, Bergdahl MO, Jungmarker K, Ankartjärn L, et al. Internet treatment addressing either insomnia or depression, for patients with both diagnoses: a randomized trial. Sleep 2015 Feb 01;38(2):267-277 [FREE Full text] [doi: 10.5665/sleep.4412] [Medline: 25337948]
40. Cheng P, Kalmbach DA, Tallent G, Joseph CL, Espie CA, Drake CL. Depression prevention via digital cognitive behavioral therapy for insomnia: a randomized controlled trial. Sleep 2019 Oct 09;42(10):5-20 [FREE Full text] [doi: 10.1093/sleep/zsz150] [Medline: 31535688]
41. Cheng P, Casement MD, Kalmbach DA, Cuamatzi Castelan A, Drake CL. Self-efficacy in insomnia symptom management after digital CBT-I mediates insomnia severity during the COVID-19 pandemic. Behav Sleep Med 2022;20(5):638-648 [FREE Full text] [doi: 10.1080/15402002.2021.1976780] [Medline: 34511016]
42. Henry AL, Miller CB, Emsley R, Sheaves B, Freeman D, Luik AI, et al. Insomnia as a mediating therapeutic target for depressive symptoms: a sub-analysis of participant data from two large randomized controlled trials of a digital sleep intervention. J Sleep Res 2021 Feb 18;30(1):e13140 [FREE Full text] [doi: 10.1111/jsr.13140] [Medline: 32810921]
43. Henry AL, Miller CB, Emsley R, Sheaves B, Freeman D, Luik AI, et al. Does treating insomnia with digital cognitive behavioural therapy (Sleepio) mediate improvements in anxiety for those with insomnia and comorbid anxiety? An analysis using individual participant data from two large randomised controlled trials. J Affect Disord 2023 Oct 15;339:58-63 [FREE Full text] [doi: 10.1016/j.jad.2023.06.053] [Medline: 37390923]
44. Lee S, Oh JW, Park KM, Lee S, Lee E. Digital cognitive behavioral therapy for insomnia on depression and anxiety: a systematic review and meta-analysis. NPJ Digit Med 2023 Mar 25;6(1):52 [FREE Full text] [doi: 10.1038/s41746-023-00800-3] [Medline: 36966184]
45. Reffi AN, Drake CL, Kalmbach DA, Jovanovic T, Norrholm SD, Roth T, et al. Pre-pandemic sleep reactivity prospectively predicts distress during the COVID-19 pandemic: the protective effect of insomnia treatment. J Sleep Res 2023 Feb;32(1):e13709 [FREE Full text] [doi: 10.1111/jsr.13709] [Medline: 36053867]
46. Hollis C, Sampson S, Simons L, Davies EB, Churchill R, Betton V, et al. Identifying research priorities for digital technology in mental health care: results of the James Lind Alliance Priority Setting Partnership. Lancet Psychiatry 2018 Oct;5(10):845-854 [FREE Full text] [doi: 10.1016/S2215-0366(18)30296-7] [Medline: 30170964]
47. Simon L, Steinmetz L, Feige B, Benz F, Spiegelhalder K, Baumeister H. Comparative efficacy of onsite, digital, and other settings for cognitive behavioral therapy for insomnia: a systematic review and network meta-analysis. Sci Rep 2023 Feb 02;13(1):1929 [FREE Full text] [doi: 10.1038/s41598-023-28853-0] [Medline: 36732610]
48. Wang K, Varma DS, Prosperi M. A systematic review of the effectiveness of mobile apps for monitoring and management of mental health symptoms or disorders. J Psychiatr Res 2018 Dec;107:73-78. [doi: 10.1016/j.jpsychires.2018.10.006] [Medline: 30347316]
49. Fleming T, Bavin L, Lucassen M, Stasiak K, Hopkins S, Merry S. Beyond the trial: systematic review of real-world uptake and engagement with digital self-help interventions for depression, low mood, or anxiety. J Med Internet Res 2018 Jun 06;20(6):e199 [FREE Full text] [doi: 10.2196/jmir.9275] [Medline: 29875089]
50. Torous J, Luo J, Chan S. Mental health apps: what to tell patients. Psychiatry 2018;17(3):21-25.
51. Espie CA, Henry AL. Disseminating cognitive behavioural therapy (CBT) for insomnia at scale: capitalising on the potential of digital CBT to deliver clinical guideline care. J Sleep Res 2023 Dec;32(6):e14025. [doi: 10.1111/jsr.14025] [Medline: 37642008]
52. Espie CA, Kyle SD, Williams C, Ong JC, Douglas NJ, Hames P, et al. A randomized, placebo-controlled trial of online cognitive behavioral therapy for chronic insomnia disorder delivered via an automated media-rich web application. Sleep 2012 Jun 01;35(6):769-781 [FREE Full text] [doi: 10.5665/sleep.1872] [Medline: 22654196]

XSL•FO
RenderX

53. Morin CM. Profile of somryst prescription digital therapeutic for chronic insomnia: overview of safety and efficacy. Expert Rev Med Devices 2020 Dec;17(12):1239-1248 [FREE Full text] [doi: 10.1080/17434440.2020.1852929] [Medline: 33226269]
54. Torous JB, Chan SR, Yellowlees PM, Boland R. To use or not? Evaluating ASPECTS of smartphone apps and mobile technology for clinical care in psychiatry. J. Clin. Psychiatry 2016 Apr 26;77(06):e734-e738. [doi: 10.4088/jcp.15com10619]
55. Kraft P, Drozd F, Olsen E. Digital Therapy: Addressing Willpower as Part of the Cognitive-Affective Processing System in the Service of Habit Change. In: Oinas-Kukkonen H, Hasle P, Harjumaa M, Segerståhl K, Øhrstrøm P, editors. Persuasive Technology. PERSUASIVE 2008. Lecture Notes in Computer Science, vol 5033. Berlin, Germany: Springer; 2008:177-178.
56. Spielman AJ. Assessment of insomnia. Clinical Psychology Review 1986;6(1):11-25 [FREE Full text] [doi: 10.1016/0272-7358(86)90015-2]
57. Beck JS. Cognitive Behavior Therapy: Basics and Beyond, 2nd edition. New York City, NY: The Guilford Press; 2011.
58. Dawkins R. The Extended Phenotype: The Gene as the Unit of Selection. Oxford, England: Oxford University Press; 1982.
59. Jain SH, Powers BW, Hawkins JB, Brownstein JS. The digital phenotype. Nat Biotechnol 2015 May;33(5):462-463. [doi: 10.1038/nbt.3223] [Medline: 25965751]
60. Insel T. Digital phenotyping: technology for a new science of behavior. JAMA 2017 Oct 03;318(13):1215-1216. [doi: 10.1001/jama.2017.11295] [Medline: 28973224]
61. Torous J, Onnela J, Keshavan M. New dimensions and new tools to realize the potential of RDoC: digital phenotyping via smartphones and connected devices. Transl Psychiatry 2017 Mar 07;7(3):e1053-e1053 [FREE Full text] [doi: 10.1038/tp.2017.25] [Medline: 28267146]
62. Oudin A, Maatoug R, Bourla A, Ferreri F, Bonnot O, Millet B, et al. Digital phenotyping: data-driven psychiatry to redefine mental health. J Med Internet Res 2023 Oct 04;25:e44502 [FREE Full text] [doi: 10.2196/44502] [Medline: 37792430]
63. Mouchabac S, Conejero I, Lakhlifi C, Msellek I, Malandain L, Adrien V, et al. Improving clinical decision-making in psychiatry: implementation of digital phenotyping could mitigate the influence of patient's and practitioner's individual cognitive biases. Dialogues Clin Neurosci 2021 Jun 01;23(1):52-61 [FREE Full text] [doi: 10.1080/19585969.2022.2042165] [Medline: 35860175]
64. Smith KA, Blease C, Faurholt-Jepsen M, Firth J, Van Daele T, Moreno C, et al. Digital mental health: challenges and next steps. BMJ Ment Health 2023 Feb;26(1):1 [FREE Full text] [doi: 10.1136/bmjment-2023-300670] [Medline: 37197797]
65. Boinon D, Charles C, Fasse L, Journiac J, Pallubicki G, Guerdoux-Ninot E, et al. Quelle est la place des professionnels de santé dans l'adhésion aux programmes thérapeutiques en ligne de l'insomnie ? Éléments de réflexion issus de l'étude Sleep-4-All-1 et protocole de l'étude Sleep-4-All-2.0. Psycho-Oncol 2022 Mar 25;16(1):173-181. [doi: 10.3166/pson-2022-0179]
66. Hasan F, Tu Y, Yang C, James Gordon C, Wu D, Lee H, et al. Comparative efficacy of digital cognitive behavioral therapy for insomnia: a systematic review and network meta-analysis. Sleep Med Rev 2022 Feb;61:101567. [doi: 10.1016/j.smrv.2021.101567] [Medline: 34902820]
67. Ren R, Zhang Y, Shi Y, Zhang H, Vitiello MV, Tang X. The beneficial effects of integrating a personalized telephone-delivered component into digital cognitive behavioral therapy for insomnia in a large, hospital-based population. Sleep Med 2023 Jun;106:25-32. [doi: 10.1016/j.sleep.2023.03.012] [Medline: 37037053]
68. Carlbring P, Andersson G, Cuijpers P, Riper H, Hedman-Lagerlöf E. Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: an updated systematic review and meta-analysis. Cogn Behav Ther 2018 Jan;47(1):1-18 [FREE Full text] [doi: 10.1080/16506073.2017.1401115] [Medline: 29215315]
69. Lancee J, van Straten A, Morina N, Kaldo V, Kamphuis JH. Guided online or face-to-face cognitive behavioral treatment for insomnia: a randomized wait-list controlled trial. Sleep 2016 Jan 01;39(1):183-191 [FREE Full text] [doi: 10.5665/sleep.5344] [Medline: 26414893]
70. Seyffert M, Lagisetty P, Landgraf J, Chopra V, Pfeiffer PN, Conte ML, et al. Internet-delivered cognitive behavioral therapy to treat insomnia: a systematic review and meta-analysis. PLoS One 2016;11(2):e0149139 [FREE Full text] [doi: 10.1371/journal.pone.0149139] [Medline: 26867139]
71. Zachariae R, Lyby MS, Ritterband LM, O'Toole MS. Efficacy of internet-delivered cognitive-behavioral therapy for insomnia - A systematic review and meta-analysis of randomized controlled trials. Sleep Med Rev 2016 Dec;30:1-10. [doi: 10.1016/j.smrv.2015.10.004] [Medline: 26615572]
72. Baglioni C, Espie C, Riemann D. Cognitive-behavioural therapy for insomnia (CBT-I) across the life span: guidelines and clinical protocols for health professionals. 2022 Jul 8;1? 1:274. [doi: 10.1002/9781119891192]
73. Kyle SD, Miller CB, Rogers Z, Siriwardena AN, Macmahon KM, Espie CA. Sleep restriction therapy for insomnia is associated with reduced objective total sleep time, increased daytime somnolence, and objectively impaired vigilance: implications for the clinical management of insomnia disorder. Sleep 2014 Feb 01;37(2):229-237 [FREE Full text] [doi: 10.5665/sleep.3386] [Medline: 24497651]
74. Espie CA, Torous J, Brennan TA. Digital therapeutics should be regulated with gold-standard evidence. Health Affairs Forefront. 2022. URL: https://www.healthaffairs.org/content/forefront/digital-therapeutics-should-regulated-gold-standard-evidence [accessed 2025-05-20]

XSL•FO
RenderX

Abbreviations

ASPECTS: Actionable, Secure, Professional, Evidence-based, Customizable, Transparent
CBT-i: cognitive behavioral therapy for insomnia
CNRS: Centre National de la Recherche Scientifique
dCBT-i: digital cognitive behavioral therapy for insomnia
DSM-IV: Diagnostic and Statistical Manual of Mental Disorders (Fourth Edition)
INSERM: Institut National de la Santé et de la Recherche Médicale
ISI: Insomnia Severity Index
RCT: randomized controlled trial
UMR: Unité Mixte de Recherche
WHAT: work, help, accessible, trusted

Edited by A Kushniruk; submitted 17.12.24; peer-reviewed by E McCarthy, K Bavarsad; comments to author 17.04.25; revised version received 24.04.25; accepted 12.05.25; published 27.05.25.

Please cite as:
Stern E, Brune S, Mouchabac S, Dubroc A, de la Personne C, Geoffroy PA
Innovative Digital Cognitive Behavioral Treatment for Insomnia Disorder in Adults (dCBT-i): Framework Development
JMIR Hum Factors 2025;12:e70193
URL: https://humanfactors.jmir.org/2025/1/e70193
doi:10.2196/70193
PMID:40424031

©Emilie Stern, Severine Brune, Stéphane Mouchabac, Anthony Dubroc, Charles de la Personne, Pierre Alexis Geoffroy. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 27.05.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Needs and Expectations for the myNewWay Blended Digital and Face-to-Face Psychotherapy Model of Care for Depression and Anxiety (Part 1): Participatory Design Study including People with Lived and Living Experience

Katarina Kikas[1,2*], BPsychSc; Kathleen O'Moore[1*], PsyD; Rosemaree Kathleen Miller[1], PhD; Julie-Anne Therese Matheson[1], BSc; Sophie Li[1], PhD; Kathleen Varghese[1], BFA; Peter Baldwin[1], PhD; Nicole Cockayne[1], MPH; Alexis Estelle Whitton[1,3], PhD; Jill Maree Newby[1,2], PhD

[1]Black Dog Institute, Randwick, Australia

[2]School of Psychology, UNSW Sydney, Kensington, Australia

[3]Faculty of Medicine & Health, UNSW Sydney, Kensington, Australia

[*]these authors contributed equally

Corresponding Author:
Jill Maree Newby, PhD
School of Psychology
UNSW Sydney
High Street
Kensington, 2252
Australia
Phone: 61 (02) 9065 9108
Email: j.newby@unsw.edu.au

Abstract

Background: Digital mental health interventions (DMHIs) are effective in reducing symptoms of depression and anxiety. Low user engagement and uptake of DMHIs observed in previous research may be addressed by involving the intended target audience in the design of the DMHI from the outset.

Objective: This study is phase 1 of a multiphase project aimed at designing, developing, and evaluating a blended DMHI for depression and anxiety in Australia. Our objective was to partner with adults with lived and living experiences of depression and anxiety on their needs and expectations of a new transdiagnostic DMHI for depression and anxiety. This included identifying strategies that would help increase their engagement with the DMHI and their preferences for integrating the DMHI with psychotherapy.

Methods: A mixed methods participatory design approach was used to collect quantitative and qualitative data via a web-based survey (n=324) and semistructured interviews (n=21). Feedback was collected on participants' needs and expectations for the DMHI, including accessibility, content, features, functionality, format, data sharing, preferred clinical support pathways, and barriers to and facilitators of user engagement. Qualitative interview data were analyzed using reflexive thematic analysis.

Results: Most participants (190/257, 73.9%) preferred a DMHI delivered as a smartphone app that could be used at any time of the day. Ease of use and a well-designed interface were important, as was a positive, encouraging, and uplifting DMHI look and feel. Other preferences included symptom tracking, diverse therapeutic content, and features that facilitated social connection and peer support (eg, online community and stories of lived and living experience). Participants also suggested several strategies to enhance engagement with the DMHI, including personalization, reminders, short and achievable activities, and goal setting. Participants reported a strong interest in sharing information from their DMHI with mental health professionals (to facilitate therapy), especially regarding changes to their emotions.

Conclusions: Transdiagnostic DMHIs for depression and anxiety have great potential to improve access to affordable, evidence-based mental health support. Involving people with lived and living experiences of depression and anxiety in the design, development, and conceptualization of DMHIs may improve uptake, acceptance, engagement, usability, and ultimately, treatment outcomes.

XSL•FO
RenderX

(JMIR Hum Factors 2025;12:e69499) doi:10.2196/69499

KEYWORDS

participatory design; depression; anxiety; smartphone; digital intervention; blended care; transdiagnostic

Introduction

Background

Globally, depression and anxiety are two of the most prevalent mental health disorders. In 2019, an estimated 280 million people were living with a depressive disorder, while >300 million people were living with an anxiety disorder [1]. Depression and anxiety commonly co-occur, with comorbidity rates as high as 50% [2]. These disorders are associated with a decreased quality of life due to poor physical health, disability, and reduced life expectancy [3], as well as lower workforce productivity and high economic costs [4]. Fortunately, there are many effective treatments for depression and anxiety that either target specific disorders or use a transdiagnostic approach [5]. However, several barriers affect access to treatment. These include high out-of-pocket costs and inconsistencies in the quality and type of care received [1]. Concerns about seeking help in person due to negative beliefs and stigma surrounding mental illness also play a role [6]. In addition, long waiting times persist due to a shortage of trained mental health professionals and inadequate funding of mental health services [1,6].

Digital Mental Health Treatment

Digital mental health interventions (DMHIs), such as smartphone apps or online programs, address many of the barriers to face-to-face mental health care. DMHIs allow people to access psychotherapy skills and content in their own time and empower them to self-manage their symptoms [7-11]. Due to these advantages and the need to minimize face-to-face contact during the COVID-19 pandemic, many DMHIs have been developed and implemented quickly in recent years [12]. DMHIs can be used in a self-guided format (ie, without human support) [8,9], guided format (ie, with remote support from a health professional or coach), or together with face-to-face therapy sessions in person or with telehealth in a "blended" format (blended care) [12-14]. Evidence supports the efficacy of DMHIs, particularly those based on cognitive behavioral therapy (CBT), in reducing symptoms of depression and anxiety [15,16]. A 2021 meta-analysis [15] found that compared to control conditions, both guided and self-guided DMHIs significantly reduced depressive symptoms across 83 studies, with a medium pooled effect size (Hedges g=0.52; 95% CI 0.43-0.60). For anxiety, a recent meta-analysis [16] reported that compared to inactive controls, guided and self-guided DMHIs significantly reduced anxiety symptoms across 47 studies, with a large pooled effect size (Hedges g=0.80; 95% CI 0.68-0.93). While most studies have evaluated internet-delivered DMHIs (ie, delivered on websites), smartphone apps targeting depressive and anxiety symptoms have also shown positive results, with small to medium pooled effect sizes for anxiety (Hedges g=0.26; 95% CI 0.21-0.31) and depression (Hedges g=0.50; 95% CI 0.40-0.61) [17,18].

Despite this evidence showing that smartphone apps and internet-based interventions are effective in improving symptoms of depression and anxiety, DMHIs are not typically integrated into routine mental health care [19]. There are several reasons for this, including a lack of practitioner confidence and training in integrating these tools into practice, difficulties in personalizing these interventions, and a lack of integration with practice management software and electronic health records [20]. There is promising preliminary research into the combination of DMHIs with face-to-face therapy in blended care for depression and anxiety [14,21,22]. However, research on the efficacy and effectiveness of blended care is limited compared to the evidence base for self-guided and guided DMHIs [8,9,23] and is confined mainly to Europe and North America [14,24]. In practice, many existing DMHIs have high attrition rates and low user engagement, especially when accessed in stand-alone self-guided formats [25,26] and in "real-world" naturalistic settings outside of controlled research studies [27]. For example, a review by Linardon and Fuller-Tyszkiewicz [26] found that the dropout rate of users participating in studies that included smartphone apps for depressive or anxious symptoms was 36% and 21%, respectively. Dropout rates are even higher in naturalistic community settings. For instance, studies have shown that completion rates of DMHIs in naturalistic settings are at best half of what is found in clinical trials of the same intervention [28-30]. Several factors contribute to low DMHI engagement, including technical issues, motivation issues, and a lack of user-driven design [11,25,31,32]. Reviews of the existing literature also show that various features, functionality, and design considerations can promote user engagement with DMHIs. These include the perceived usefulness; positive impact and relevance of DMHIs; content that promotes good mental, physical, and cognitive health [33,34] and social connection [25,31]; design aspects (eg, ease of use and interface); personalized content and choice [35-37]; data security and privacy; and therapist guidance [25].

Limitations of DMHIs for Depression and Anxiety

These findings provide guidance for improving participant engagement and reducing attrition for DMHIs overall, which is especially relevant given the increased availability of DMHIs following the COVID-19 pandemic. However, an often-overlooked issue is how DMHIs should be integrated into routine clinical care for depression, anxiety, and other mental disorders [9]. Many existing self-guided and guided DMHIs include linear, one-size-fits-all programs for a specific disorder (eg, major depressive disorder) or multiple disorders (eg, transdiagnostic programs for mixed anxiety or anxiety and depression). In this model, programs follow a mostly set structure, delivering a series of modules that are completed in the same order for each person. While this approach has benefits [38], the downside is that the DMHI cannot be individually tailored to suit a person's goals, needs, symptoms, and problems.

XSL•FO
RenderX

This lack of personalization negatively affects user engagement [11,25] and is a barrier for mental health professionals using these programs in practice [39]. Another important clinical consideration is that people experiencing severe depression and anxiety symptoms, and frequent suicidal ideation, use DMHIs once they are released into the community, although the programs are often not designed for these clients (eg, myCompass and THIS WAY UP) [40-42]. The needs, preferences, and expectations of people with varying symptom severity and suicidal risk levels need to be considered at the outset when developing interventions for depression and anxiety. Another factor affecting the inclusion of DMHIs in routine clinical care is that many of these interventions have been designed as "self-help" or "stand-alone" products [8,43,44]. This has contributed to the low uptake of DMHIs in practice and limits their potential use, reach, and benefits for people with depression and anxiety at various stages of care (eg, before, during, and after seeking in-person therapy).

Incorporating the needs of the intended target audience directly into DMHI design and development can potentially help address poor user engagement, high attrition, and the lack of integration of DMHIs into clinical practice. Participatory design approaches, such as co-design and co-creation, actively incorporate multiple stakeholders' perspectives into defining a problem and collaborating with them on potential solutions [45,46]. In digital mental health, participatory design involves partnering with the intended target audience during DMHI design, development, refinement, and evaluation to ensure the product or service reflects their needs. Unfortunately, in DMHI research that uses participatory approaches, these details are often inconsistent, use varied terminology, or are not reported clearly [45,46]. Moreover, the few existing DMHI studies that use co-design principles have been limited to young adult samples (aged 18-34 years) or mental health professionals, with very few in adults aged ≥35 years [45]. Using participatory design and co-design principles to create DMHIs that fit into routine care may help adults with anxiety and depression, along with their mental health professionals, adopt blended care more easily [21]. However, to date, only 1 European study [47] reports the participatory design process during the development of a blended care intervention with people with lived experience of mental illness and psychotherapists. In that study, people with lived experience preferred a DMHI that was gentle and compassionate rather than performance based and they valued having the autonomy to choose modules themselves or have them selected by their therapist. It was also important that the DMHI included crisis support, transparent data security, and confidentiality information. Finally, people with lived experience preferred inclusive, varied, concise content delivered in various formats (eg, text, images, and videos) to ensure the DMHI was accessible to many people.

Rationale, Objectives, and Aims

Blending digital and face-to-face therapy is a promising approach for integrating DMHIs into routine mental health care [14,21,22,24]. However, to date, only 1 study has used participatory design to incorporate the needs of people with lived experience of mental illness into the design of a blended care intervention [47]. Blended care interventions can address the limitations of DMHIs, especially stand-alone options; hence, there is a need to further investigate the utility of the blended care approach. For depression and anxiety, the option of mental health professional involvement would add an extra layer of clinical safety and permit better tailoring of DMHI use for people experiencing depression and anxiety. This study is part of a larger Australian research project that aims to design, develop, and evaluate a new transdiagnostic DMHI for adults with depression and/or anxiety (ie, myNewWay) [48]. The intervention is intended to address limitations of existing DMHIs and be suitable for use as part of blended care. Australian research on blended care is limited; however, preliminary results align with studies in other countries [14,21,22,24] and indicate that blending digital and face-to-face therapy may improve symptoms of depression [49-51] and anxiety [51]. To the best of our knowledge, there are also currently no transdiagnostic blended care models in Australia designed with target audience input and tailored to the specific needs and preferences of adults with depression and anxiety.

This study is 1 of 2 related participatory design studies in which we partnered with people with lived and living experiences of depression and anxiety and mental health professionals in the design of the new blended care DMHI. This paper reports on the findings of people with lived and living experience. Findings from mental health professionals are reported separately in part 2 of this manuscript series [52]. The DMHI will be transdiagnostic rather than diagnosis specific, as transdiagnostic approaches are acceptable, feasible, efficient, and effective for treating multiple symptoms and diagnoses within 1 treatment program [5,40]. We will use findings of this study to co-design and then develop a program that will be engaging, safe, and effective for people with varying severity levels, including those experiencing suicidal ideation. The program could be personalized to individual goals and needs and integrated into face-to-face or telehealth mental health care. We also wanted to ensure that the DMHI met the needs of adults with lived and living experience by including relevant therapeutic content and sustaining their interest with appealing design features and functionality, especially given that user engagement is associated with better treatment outcomes [15,53] and more cost-effective treatment [54,55].

Conducting participatory design within an Australian context is important to ensure that the new blended care model aligns with the needs and expectations of adults living with anxiety and depression. In Australia, adults with depression and anxiety are treated in the private mental health system and can access up to 10 psychotherapy sessions per calendar year with partial government rebates [56]. Blending DMHIs with this existing treatment pathway could increase the quality of care, and, thereby, enhance recovery by providing access to evidence-based resources for depression and anxiety before, during, and after psychotherapy. Given that most blended care research has been conducted in Europe and North America [14,24], conducting research outside these cultural contexts also helps to show which elements of blended care may be universal and which elements are not. The first aim of this study was to examine the needs and expectations of participants regarding the content, features, functionality, and look and feel of the

XSL•FO
RenderX

DMHI. This also includes how people with depression and anxiety want to access the DMHI, their preferences for content delivery, and the overall blended model of care. With respect to the blended care model, we asked about the types of therapeutic guidance and human-based support they would like to receive while using a DMHI that blends with psychotherapy, as well as their preferences for sharing data from the DMHI with health professionals. Understanding participants' preferences for including guided components within a DMHI will ensure that it meets their needs and leverages the potential benefits of guided DMHIs for depression and anxiety [15]. Our second aim was to identify factors that would affect engagement with the new blended DMHI, which could be addressed in the design of the interface, functionality, or content of the intervention to maximize sustained engagement with the DMHI.

Methods

Participatory Design Approach

A mixed methods convergent parallel design [57] was used for this study, which included a convenience sample of people with lived and living experience. This approach ensured that both a broad range of perspectives (ie, web-based survey) and in-depth insights (ie, qualitative interviews) were represented to achieve the study aims and conclusions [57]. Stage 1 was the web-based survey, which assessed participants' experiences and attitudes toward DMHIs and the barriers and facilitators to their use. The survey was developed in accordance with the Checklist for Reporting Results of Internet E-Surveys [58]. For stage 2, a subset of survey respondents completed one-on-one online qualitative interviews on their preferences for the DMHI. For people with lived and living experience, stage 3 included co-design sessions and prototype testing of designs based on findings from stages 1 and 2; however, these results are not reported in this paper. The survey and interview questions were developed by the research team, which included health professionals, people with lived experience, user experience and user interface experts, and those with knowledge of implementation research (Multimedia Appendix 1). The data collection methods used in stages 1 and 2 are consistent with prior participatory design research in digital health [45]. In addition, 5 people with lived experience of depression and anxiety served as lived experience advisors during stages 1 to 3. The lived experience advisors took part in each stage of the broader project, including both the survey and qualitative interviews conducted in this study.

Ethical Considerations

All research procedures received ethics approval from the University of New South Wales Human Research Ethics Committee (HC200541), and all participants provided informed consent before participating. If participants wished to withdraw their consent, they could contact the research team. Participant data from the online survey and interview was deidentified and replaced with a number and saved in password protected file accessible only by the research team. Participants were given the option to enter a draw to win an Aus $100 (US $65) online gift card for taking part in the online survey. For the interview, participants were reimbursed with an Aus $30 (US $20) online gift card.

Stage 1: Web-Based Survey

The web-based 20-minute open survey was distributed between October and December 2020. The survey was advertised via the Black Dog Institute website and social media (eg, Facebook and Instagram) and the Black Dog Institute Lived Experience Resource Network. Interested participants were directed to the Black Dog Institute website, which outlined the eligibility criteria for the study and included a link to the web-based survey hosted on Qualtrics (version 2020; SAP Inc) software. At the start of the survey, respondents read the participant information sheet and consent form outlining the purpose of the research study, eligibility criteria, where their data were stored and who had access to it, the length of time to complete the study, and their reimbursement incentive. Eligibility criteria were that participants were aged >18 years, could speak English, currently lived in Australia, and had current or previous experience with anxiety or depression. The web-based survey included questions (Multimedia Appendix 1) on participant demographics, their mental health treatment history, current technology use, their experiences with DMHIs, and preferences for a new DMHI targeted at reducing symptoms of depression and anxiety. Participants were asked what features in the DMHI they would find helpful, how they would like any content displayed, and their preferred medium for accessing the DMHI (eg, smartphone and laptop). Participants' current levels of depression, anxiety, and stress were also assessed using the validated 21-item Depression Anxiety Stress Scale (DASS-21) [59]. The DASS-21 is reported to have sound psychometric properties [60]. Respondents rate each item on a 4-point Likert scale (never=0 to almost always=3), with higher scores indicating more severe levels of depression, anxiety, or stress.

Stage 2: Interviews

The interviews were conducted with Zoom videoconferencing software (Zoom Communications, Inc) and lasted 60 minutes on average (SD 10.52 min; range 35-74 min). The semistructured interviews included open-ended questions (Multimedia Appendix 1) that assessed attitudes toward and experiences with DMHIs, preferences for features and functionality, and suggestions for improving engagement with DMHIs. A clinical psychologist (KO) and the user experience designer (KV) facilitated interviews. All interviews were audio-recorded and later transcribed by the project officer (KK).

Data Analysis

Quantitative data from the web-based survey were analyzed using SPSS (version 26.0; IBM Corp). Descriptive statistics were calculated for closed-ended survey questions. Sample counts and percentages were provided for each survey question. Reflexive thematic analysis was adopted for qualitative data because this methodology was appropriate for answering the research questions [61]. Following each interview, transcripts were checked and cleaned against the original audio by KK. KK (project officer), KO (clinical psychologist), and RKM (project officer) read the transcripts several times, and after familiarization, they used a general inductive approach to

XSL•FO
RenderX

generate the initial set of codes. The 3 authors have experience with digital mental health and thematic analysis. Codes were categorized in Microsoft Excel (version 2024) to create the initial coding framework, which was revised iteratively by KK, RKM, and JATM (research assistant with experience in digital mental health and thematic analysis) for consensus on final themes and subthemes. Any disagreements on codes were resolved through discussion with the research team (including KO, JMN, and AEW, who are psychologists) until a consensus was reached.

Results

Stage 1: Web-Based Survey

Overview

Of the 331 adult participants who provided consent, 324 (97.9%) supplied their demographic information, and 231 (68.3%) participants completed the full survey. The demographic, mental health, and treatment characteristics of 324 participants are presented in Tables 1 and 2. The sample had an average age of 46 (SD 16.4; range 18-80) years, was predominantly female, spoke English as their first language, was born in Australia, and lived in a major Australian city. A total of 209 (68.3%) of 306 participants rated their digital literacy as good to very good, with the rest rating it as acceptable (n=65, 21.2%) or poor (n=12, 3.9%). Approximately half (156/306, 51%) of the participants spent 1 to 3 hours online daily, and >40% (131/306, 42.8%) spent >4 hours online daily. A small number of participants (19/306, 6.2%) spent <1 hour online daily.

XSL•FO
RenderX

Table 1. Demographic characteristics for the web-based survey (n=324).

Demographic characteristics	Participants, n (%)
Gender	
Women	204 (63)
Men	113 (34.9)
Nonbinary	7 (2.2)
First language English	305 (94.1)
Born in Australia	244 (75.3)
Aboriginal or Torres Strait Islander origin	5 (1.5)
Residential location in Australia[a]	
Major cities	221 (69.3)
Regional or remote	98 (30.7)
Relationship status[a]	
Single or never married	91 (28.5)
In a relationship	58 (18.2)
Married or de facto	102 (32)
Separated, divorced, or widowed	66 (20.7)
Other relationship status	2 (0.6)
Level of education[a]	
Less than high school level	25 (7.8)
High school level	53 (16.6)
Certificate or diploma	133 (41.7)
University undergraduate degree	58 (18.2)
University postgraduate degree	41 (12.9)
Other levels of education	9 (2.8)
Employment status[a,b]	
Full-time paid work	106 (33.2)
Part-time paid work	74 (23.2)
Not in the labor force	59 (18.5)
Unemployed	56 (17.6)
At-home parent	9 (2.8)
Carer for a family member	7 (2.2)
Student	7 (2.2)
Other employment status	14 (4.4)

[a]Percentages for residential location, relationship status, level of education, and employment status were calculated based on the number of participants who responded to these questions (n=319).

[b]Percentages do not add up to 100% as respondents were allowed multiple responses.

XSL•FO
RenderX

Table 2. Treatment and mental health characteristics for the web-based survey (n=318).

Characteristics	Participants
Current chronic condition, n (%)	126 (39.5)
Carer of a person with depression and anxiety, n (%)	58 (18.2)
Ever experienced a mental health problem, n (%)	
No	6 (1.9)
Anxiety	30 (9.4)
Depression	34 (10.7)
Anxiety and depression	198 (62.3)
Not sure	7 (1.6)
Other mental health problem	43 (13.5)
Past mental health treatment[a], n (%)	
None	50 (14.9)
Medication	225 (67.2)
Therapy with psychologist	233 (69.6)
Therapy with psychiatrist	129 (38.5)
Counseling from GP[b]	129 (38.5)
Counseling from other MHP[c] (ie, nurse and social worker)	99 (29.6)
Online mental health program	63 (18.8)
Other treatment types	30 (9)
Current mental health treatment[a], n (%)	
None	110 (32.8)
Medication	153 (45.7)
Therapy with psychologist	113 (33.7)
Therapy with psychiatrist	46 (13.7)
Counseling from GP	60 (17.9)
Counseling from other MHP (ie, nurse and social worker)	24 (7.2)
Online mental health program	14 (4.2)
Other treatment type	13 (3.9)
DASS-21 scores[d], mean (SD)	
Depression	18.94 (11.50)
Anxiety	13.60 (9.31)
Stress	19.16 (8.63)
Total scores	51.70 (25.77)
DASS-21 depression ranges[d], n (%)	
Normal	61 (20)
Mild	43 (14)
Moderate	82 (26.6)
Severe	43 (14)
Extremely severe	79 (25.6)
DASS-21 anxiety ranges[d], n (%)	
Normal	92 (29.9)
Mild	18 (5.8)

XSL•FO
RenderX

Characteristics	Participants
Moderate	70 (22.7)
Severe	43 (14)
Extremely severe	85 (27.6)
DASS-21 stress ranges[d], n (%)	
Normal	111 (36)
Mild	47 (15.3)
Moderate	72 (23.4)
Severe	60 (19.5)
Extremely severe	18 (5.8)

[a]Percentages do not add up to 100% as respondents were allowed multiple responses.

[b]GP: general practitioner.

[c]MHP: mental health professional.

[d]Results for the 21-item Depression Anxiety Stress Scale (DASS-21) were calculated based on the number of participants who responded to these questions (n=308).

Mental Health and Treatment

Most participants (198/318, 62.3%) had experienced both depression and anxiety (Table 2). A total of 191 (60.1%) of 318 participants were currently receiving mental health treatment, and 285 (85.3%) participants had received mental health treatment in the past. Medication was the most common form of current and past treatment, followed by therapy with a psychologist or psychiatrist (Table 2). The distribution of participants (308/324, 95.1%) who scored in the normal, mild, moderate, severe, and extremely severe ranges for the 3 DASS-21 subscales can be seen in Table 2. Approximately 80% (247/308) of participants had elevated depression scores, 70% (216/308) had elevated anxiety scores, and 64% (197/308) had elevated stress scores on the DASS-21 subscales (scored in the mild to extremely severe ranges).

Current and Past Use of DMHIs

Only 52 (17%) of 306 participants were currently using DMHIs, while 131 (42.8%) had used them in the past. The most commonly used DMHIs among 139 participants were mindfulness meditation apps (n=79, 56.8%), followed by CBT-based apps (n=30, 21.6%) and forums (n=13, 9.4%). The most frequently used apps were Headspace, Smiling Mind, and Calm. Of the participants who had never used DMHIs (155/306, 50.7%), the 3 most common reasons were a lack of awareness, a belief that these products are not useful or necessary, and not having found a DMHI that was helpful to them.

Preferences for Digital Mental Health

Product Promotion or Awareness

Almost 70% (177/256) of participants preferred finding out about a DMHI via a health professional recommendation, followed by social media or Google search (Table 3).

XSL•FO
RenderX

Table 3. Participants' preferences for promotion and access.

	Participants, n (%)
Finding out about the DMHI[a] (n=256)	
Recommendation from a health professional	177 (69.1)
Social media	159 (62.1)
Google search	156 (60.9)
Recommendation from friend or family member	119 (46.5)
App store search	70 (27.3)
I am not sure	13 (5.1)
Other promotion pathway	10 (3.9)
Mode of access (n=257)	
Smartphone via an app	190 (73.9)
Laptop	85 (33.1)
Tablet or iPad	72 (28.1)
Smartphone via a web browser	55 (21.4)
Desktop computer	50 (19.5)
Other mode of access (eg, smartwatch)	11 (4.3)
Time of the day (n=257)	
Any time they felt the need to use it	164 (63.8)
In the evening (PM)	96 (37.4)
Just before bed	73 (28.4)
In the middle of the night (if needed)	47 (18.3)
In the morning (AM)	43 (16.7)
Upon waking up	22 (8.6)
Lunchtime	15 (5.8)
Other time of the day	6 (2.3)

[a]DMHI: digital mental health intervention.

Access to Product

Most participants (190/257, 73.9%) reported wanting to access a DMHI via a smartphone app, while a third preferred (85/257, 33.1%) using a laptop (Table 3). Most participants wanted to use a DMHI when they felt the need to use it (164/257, 63.8%), in the evening (96/257, 37.4%), or just before bed (73/257, 28.4%; Table 3).

Willingness to Pay

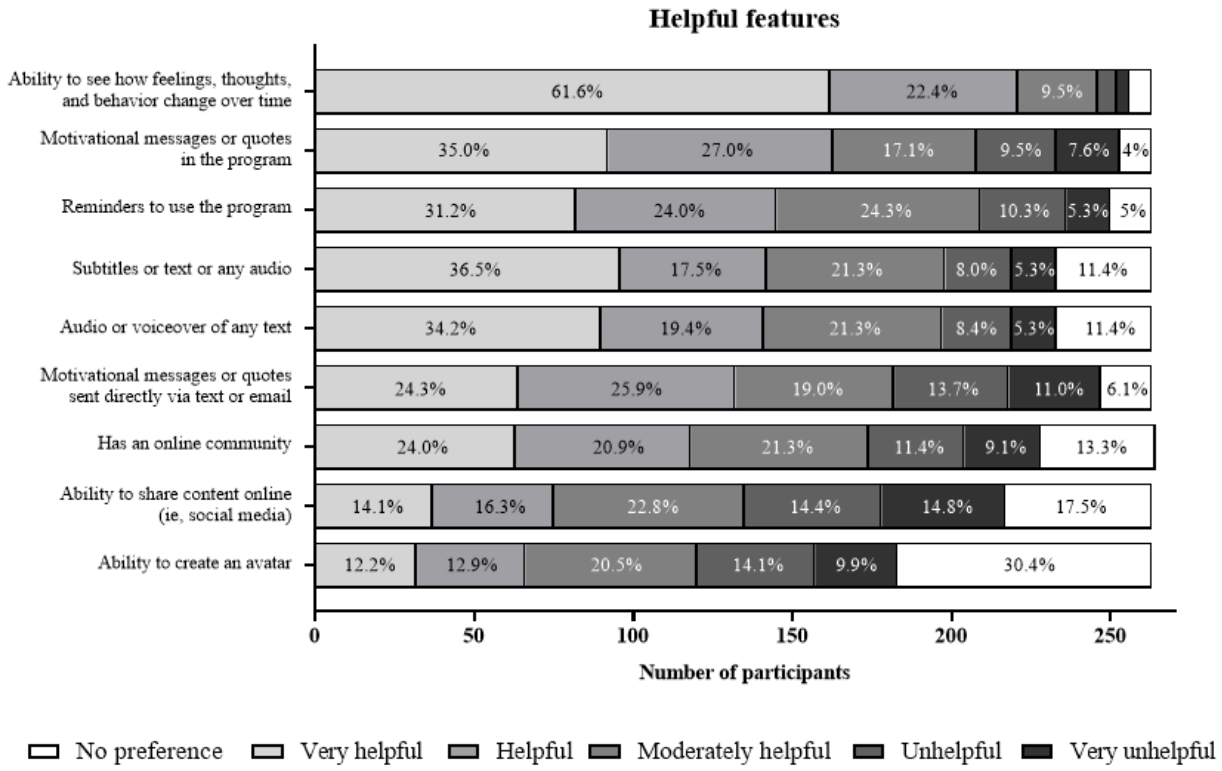
Only 40 (16.5%) of the 242 participants were willing to pay for a DMHI, whereas most said maybe (n=142, 58.7%), and the rest said no (n=60, 24.8%). Of the 182 participants who responded yes or maybe, the median amount they would pay monthly for a DMHI was Aus $10.00 (US $6.5; IQR Aus $5.00-$15.00). The majority (125/176, 71%) were willing to pay between Aus $1 to $10 (US $0.65-$6.5) per month, while some participants (12/176, 6.8%) were willing to pay >Aus $50 (US $32) per month.

Preferred Features

Figure 1 shows which DMHI features participants considered the most helpful (ie, rated as helpful or very helpful). The most popular features were the ability to track thoughts, feelings, and behaviors over time (221/263, 84%), motivational messages (163/263, 62%), reminders (145/263, 55.1%), subtitles (142/263, 54%), and audio voiceover (141/263, 53.6%). Features viewed as moderately helpful included motivational messages sent directly to the user via email or text and the inclusion of an online community, while the least helpful features were having an avatar and the ability to share content online in general.

XSL•FO
RenderX

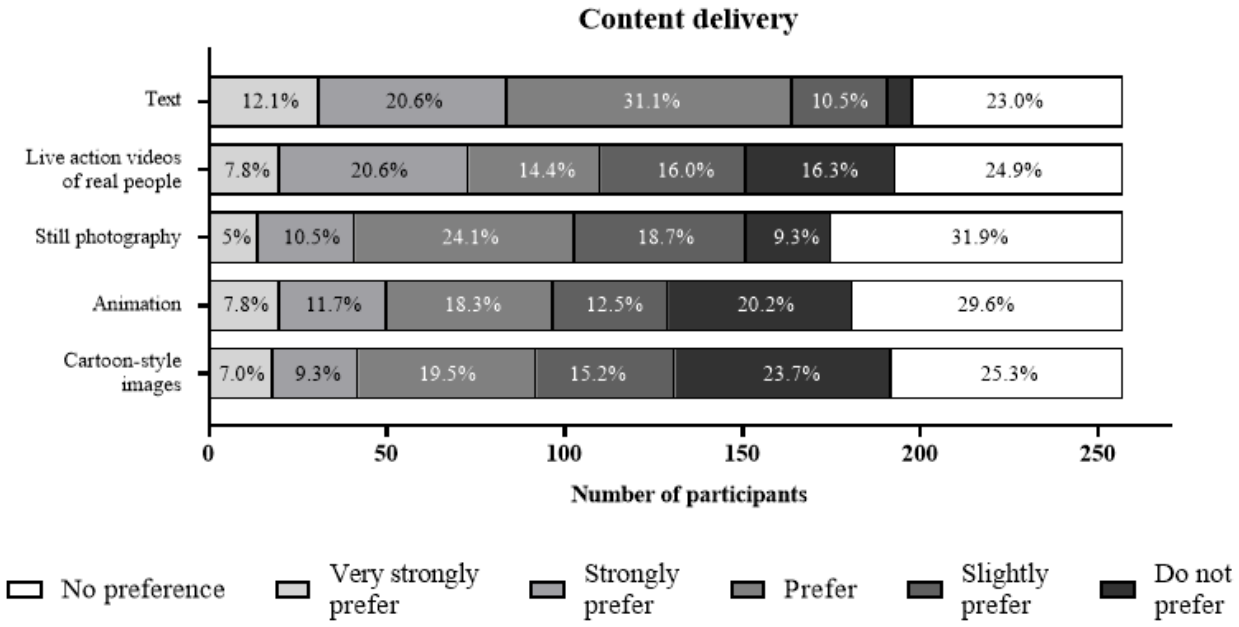
Figure 1. Participant preferences for the features of digital mental health interventions they find most helpful (n=263). Data labels show the frequency percentages.



Preferred Format of Content Delivery

Figure 2 indicates that text, still photography, and live-action videos of real people were the preferred forms of content display (ie, rated as very strongly prefer, strongly prefer, or prefer). Although there was a wide range of preferences for content display types, most participants preferred not to see information displayed in cartoon-style images (61/257, 23.7%), animation (52/257, 20.2%), or live-action videos of real people (42/257, 16.3%; Figure 2).

Figure 2. Participant preferences for preferred content delivery methods (n=257). Data labels show the frequency percentages.



XSL•FO
RenderX

Engagement With DMHIs

Participants commonly reported that if the DMHI were effective and helped them feel better (eg, “effectiveness in making me less anxious,” “I feel it is helping me,” and “success in feeling better”), it would help them stay engaged. Some participants also suggested that being able to visualize their progress and the use of rewards, challenges, and achievement milestones would encourage users to keep coming back to the DMHI (eg, “check ins to see progress” and “rewards for logging in every day”). Participants also thought personalized and tailored content, such as “choose your own path based on your mood” and “options to turn off what you use and don’t use,” would encourage them to keep using the DMHI. Other engagement strategies included reminders and notifications (eg, “to be able to set a reminder and task to help with motivation and skills”) and having a well-designed layout (eg, “colorful,” “interactive,” “simple,” and “visual and no text blocks”).

Sharing Information

Participants rated psychologists, general practitioners, and psychiatrists as individuals they were most likely to share information with via a DMHI (Table 4). Participants were relatively confident in sharing information with a counselor, a friend, or a family member and less confident when sharing with a social worker, carer, or teacher. When participants were asked about what information they would like to share via the DMHI with others (ie, psychologist, friend, or teacher), most preferred to share how they were feeling over time, followed by specific program activities, whether they felt worse or better, and their goals (Table 4). Participants also rated what types of information they were most likely to share specifically with a mental health professional. The top-rated responses were how they felt over time, when they felt better or worse, their goals, and certain activities they completed within the DMHI (Table 4). Most participants (155/241, 64.3%) preferred communicating with mental health professionals via a chat box inside the DMHI, while some (86/241, 35.7%) preferred face-to-face appointments, phone calls, or SMS text messaging (Table 4).

XSL•FO
RenderX

Table 4. Information sharing preferences by participants for digital mental health interventions (DMHIs).

	Participants
People I would share information with[a], mean (SD)	
Psychologist	67.60 (35.81)
General practitioner	60.33 (34.54)
Psychiatrist	56.64 (40.20)
Counselor	53.42 (38.39)
Friend	52.32 (36.75)
Family member	48.69 (36.48)
Social worker	34.01 (36.74)
Carer	27.45 (35.64)
Teacher	16.48 (27.95)
Types of content I would share[b,c], n (%)	
How I am feeling over time	169 (67.9)
Certain activities I complete in the program that I choose to share	123 (49.4)
When I am feeling worse	119 (47.8)
When I am feeling better	111 (44.6)
My goals	100 (40.2)
What I use within the program	61 (24.5)
All the exercises and activities I complete in the program	53 (21.3)
When I use the program	40 (16.1)
Other types of content	19 (7.6)
Types of content I would share with a mental health professional[a], mean (SD)	
How I am feeling over time	73.71 (30.97)
When I am feeling worse	70.42 (34.08)
When I am feeling better	68.67 (35.05)
My goals	64.49 (34.59)
Certain activities I complete in the program that I choose to share	62.05 (34.75)
What I use within the program	55.37 (36.47)
When I use the program	52.87 (37.82)
All the exercises and activities I complete in the program	47.92 (36.58)
How I would communicate with a mental health professional[b,c], n (%)	
Chat box inside the program	155 (64.3)
Face-to-face appointment	102 (42.3)
Phone call	98 (40.7)
SMS	84 (34.9)
Video call	43 (17.8)
Other communication method (eg, email)	8 (3.3)

[a]Ratings were made using a 100-point scale (0=would never share and 100=extremely likely to share).

[b]Percentages for types of content (n=249) and how the participant preferred to communicate with a mental health professional (n=241) were calculated based on the number of participants who responded to these questions.

[c]Percentages do not add up to 100% as respondents were allowed multiple responses.

XSL•FO
RenderX

Stage 2: Interviews

Demographic Information

A total of 21 adults participated in one-on-one qualitative interviews. Most of them (13/21, 62%) were female participants and aged 41 years on average (SD 16.4; range 21-68 years). Most spoke English as their first language (13/21, 62%), were born in Australia (11/21, 52%), and were from major Australian cities (18/21, 86%). Of the 21 participants, 14 (67%) had experienced both depression and anxiety, 3 (14%) had experienced depression only, 1 (5%) experienced anxiety only, and 3 (14%) experienced another mental health problem.

Thematic Analysis Findings

Eight key themes related to DMHIs were identified: (1) promotes user engagement; (2) friendly, pleasant look and feel; (3) content is multifaceted and interesting; (4) sense of connection with real people; (5) features that facilitate progress; (6) safe, trustworthy, and effective; (7) data security is important; and (8) must be low cost or free. All themes, subthemes, and illustrative quotes for each subtheme are listed in Table 5.

XSL•FO
RenderX

Table 5. Themes and subthemes from the thematic analysis, along with example quotes.

Themes and subthemes	Example quotes
1. Promotes user engagement	
Easy to use and simple to navigate	"It just all needs to be easy. I hate learning new technology platforms, I hate it, so easy as possible." [Participant 4]
Personalize content and activities to user needs	"There needs to be tailored kind of options for people, different ways that they can engage with it just to keep it interesting." [Participant 5]
Reminders and notifications	"I really liked the fact that it gave you reminders to use it, that was great, and you could set it to what time of the day you'd like it to remind you." [Participant 36]
Short and achievable activities	"Maybe kind of breaking up into small tasks. So that there's a clear plan how to start to work, whatever the user wants to achieve." [Participant 9]
Goal setting to support a sense of progress	"So you can set up goals...so like basically this is what I want to become or like this is what I want and it's like a journey." [Participant 12]
2. Friendly, uplifting look and feel	
Streamlined and attractive graphic design	"It felt like it had a kind of color scheme, it looked professional." [Participant 1]
Warm, gentle, and inviting tone	"The app comes up and says gently to you, 'welcome back' or 'welcome, glad to see you.'" [Participant 9]
Short, positive, and encouraging messages	"And I also get like motivating notifications every day like small quotes or something, it makes me feel good." [Participant 12]
3. Content is multifaceted and interesting	
Provides options for wellness practice that support recovery	"I think things that are like mindfulness and exercise and things that can be part of anyone's regular routine when they're recovering would be valuable." [Participant 29]
Broad variety of content to prevent boredom	"So variety would be really good for the app as well because yeah people if they know exactly what's going to happen every time they do it, they'll just get bored of it, I think." [Participant 56]
Program incorporates different treatment approaches	"Like ACT...and DBT as well, like, I think a lot of people connect with those better...because it seems like every single app on the market is literally all about CBT and mindfulness." [Participant 36]
Searchable information about mental health	"I think it would be good if there was kind of an inventory of mental illnesses that a person can look up and read the description on the symptoms and things certain things like that and maybe suggestions on what to do if you suspect that you might have it." [Participant 1]
4. Sense of connection with real people	
Personal stories shared by real people	"I think having if possible, some form of personalized stories or...there is a section there where you can read about other people who are like you and know that you're not alone." [Participant 63]
Online community support	"So I think overall, the online forums I found the most, the most useful, probably because of that connection with other people with peers, basically." [Participant 4]
5. Features that track and aid progress	
Mood-tracking feature	"I liked the feature where you could input other things that are going on like how much sleep you had...you can put all that stuff and then you could sort of look and see correlations between your mood and these external factors." [Participant 36]
Visualize changes in mental health over time	"You can step back from it and sort of view the data on you and then, you know, work on it from there, because it just feels like it's this evolving thing that you know comes and goes." [Participant 51]
Reflective journaling	"Maybe somewhere to sort of record or reflect on how far you've come or what you want to achieve, or like sort of like a reflection journal could be quite valuable as well." [Participant 29]
Share treatment progress with their psychologist	"I don't think I would use it necessarily like in (my psychologist's) office, but maybe even just to sort of like talk about with her of like how I'm progressing, what I'm doing for my own like mental health." [Participant 52]
6. Safe, trustworthy, and effective	
Tips and activities to help feel better immediately	"If there is something that I could, if I am feeling, not necessarily in crisis, but if I am feeling not okay for any reason, that I can dip into like say a certain section and know ok, I can do this." [Participant 5]
Confidence in the credibility of the program	"Knowing where it's coming from and who is being provided by." [Participant 24]

XSL•FO
RenderX

Themes and subthemes	Example quotes
Access to crisis support information if needed	"There's a kind of list of hotlines that come up that I can access at any time in case of a crisis, so I think that's really nice." [Participant 1]
User-created plan when feeling unwell	"Inviting the user to list what they like to do to keep the head above water. I like to watch comedy, pursue my interest in things like that. I look to my brother, play with my pets. Yeah, having that as a list and just go alright, here's a couple of things that you can do, but the benefit of that is that these are things you have indicated that you want to do in the past, this is your stuff." [Participant 9]
7. Data security and control over information sharing	
Control over sharing sensitive information	"Because I want to have control over that...I've got the utmost trust in my GP and we work out what's going to go on. I don't want somebody else to have knowledge or understanding of my history and come at me cold when I don't have a relationship with that person." [Participant 19]
Transparent information about data privacy	"Transparency is very important, and that element of trust thing comes in as well that you know you feel comfortable about dealing with an organization or service provider that you know is there for your own understanding and benefit." [Participant 27]
Positive about sharing data with a trusted health professional	"I know that if my psychologist found it useful that I have personally I think she's you know, she's a very trusted professional for me. If there is that comfort level there, I think it's definitely like a great feature like for her to see where I'm at." [Participant 63]
8. Must be low cost or free	"Yeah, I think I didn't want to pay for anything because I wasn't sure if it was going to work and to be honest, I didn't use it frequently." [Participant 27]

Theme 1: Promotes User Engagement

Easy to Use and Simple to Navigate

Participants indicated that they wanted a DMHI that was easy to follow, intuitive, and straightforward to navigate. Suggestions included having a fast and easy sign-up process and clear navigational icons and features. Participants also mentioned negative past experiences with products that were too clunky, slow, and frustrating to navigate. This was thought to be particularly problematic for people with mental health issues who are feeling unwell or stressed.

Personalize Content and Activities to User Needs

Many participants emphasized personalization as a key feature for maintaining user engagement with DMHIs. Some expressed a preference for having a program that was responsive to their needs. This included the user choosing the activities or content to complete based on how they were feeling on the day and the DMHI's ability to suggest content and activities to the user based on their interactions with the DMHI (eg, mood-tracking scores).

Reminders and Notifications

Participants mentioned liking reminders and notifications in apps or programs they had previously used, as this feature helped them stay on track with managing their mental health. Some preferred daily check-in reminders or notifications that could be updated at times that suited the individual's schedule, while others valued the option to customize the content shown in the reminders and notifications.

Short and Achievable Activities

There was a clear preference for activities in the DMHI to be short and quick to complete. Some participants said that being able to complete a few smaller tasks would help foster a sense of accomplishment. Participants also suggested that if an activity needed to be longer, it could be divided into smaller tasks to prevent the activity from becoming too time consuming.

Goal Setting to Support a Sense of Progress

Participants emphasized the importance of being able to track their progress within the DMHI. Suggestions included unlocking a new stage or section in the intervention and being rewarded, offering daily challenges the user could engage in, or setting up various goals the user could work toward.

Theme 2: Friendly, Uplifting Look and Feel

Streamlined and Attractive Graphic Design

Overall, participants wanted a DMHI that was visually appealing with a professional-looking color scheme and images. However, opinions varied about the preferred colors and designs. Some participants preferred bright and vibrant colors, while others wanted soft and calming colors. Presenting information graphically (eg, pie charts and icons) was also mentioned as a desirable design feature.

Warm, Gentle, and Inviting Tone

The language and tone of the content within the digital product were important. They preferred the language and tone to be empathetic, calm, and encouraging. Participants also specifically mentioned that they did not want any language or tone that could be interpreted as forceful, overpowering, or aggressive. For example, some participants preferred the DMHI to have a friendly tone, akin to how a therapist would speak to them.

Short, Positive, and Encouraging Messages

Many participants wanted to see brief, uplifting, and inspiring messaging displayed throughout the DMHI. Participants suggested these messages could be included in notifications or shown when a user first opens the DMHI to help someone feel welcome, create a sense of connection, and validate that they made the right decision to engage with the program.

XSL•FO
RenderX

Theme 3: Content Is Multifaceted and Interesting

Provides Options for Wellness Practices That Support Recovery

Several participants wanted a DMHI that incorporated a range of wellness techniques, skills, and practices. Suggestions were diverse and included sleep hygiene, exercise guidance, spirituality resources, breathing techniques, information on healthy living, meditation exercises, and gratitude journaling. Participants wanted these components to maintain long-term recovery and sustained engagement with the DMHI.

Broad Variety of Content to Prevent Boredom

Participants stressed the importance of offering a wide range of up-to-date and relevant content in the DMHI that addresses the various experiences and situations that users might encounter. This breadth and variety of content would also ensure the user could always find something new and interesting, preventing boredom and sustaining their interest in the DMHI in the long term.

Program Incorporates Different Treatment Approaches

Many participants had used CBT-based apps and programs in the past. These participants expressed that CBT was not suitable for all individuals and wished to see other therapeutic modalities, such as dialectical behavior therapy, acceptance and commitment therapy, and mindfulness, included.

Searchable Information About Mental Health

Some participants reported that a DMHI should include an information section where users could search for mental health information. For example, 1 participant suggested an inventory of mental health conditions accompanied by descriptions of their symptoms. Another participant mentioned the value of a search section to access articles or advice, allowing users to search for specific topics related to their struggles.

Theme 4: Sense of Connection With Real People

Personal Stories Shared by Real People

Several participants mentioned that they valued connecting with others who were also experiencing mental health challenges. Participants suggested that these personal stories could be shared via the DMHI through video or written accounts, either by reading stories or participating in an online forum. Some participants expressed a desire to provide support to others dealing with anxiety and depression, drawing on their past positive interactions and helping others in other apps and forums.

Online Community Support

Some participants reported that being able to interact with an online community of other individuals facing mental health challenges would enhance their engagement with the DMHI and help reduce feelings of loneliness. However, some participants highlighted the importance of ongoing safety monitoring and moderation if an online community was included in the DMHI.

Theme 5: Features That Track and Aid Progress

Mood-Tracking Feature

Most participants preferred to monitor their mood over time using the DMHI. Some participants suggested that it would be helpful to have the ability to track other factors, such as sleep, coffee intake, and menstrual cycles, within the mood-tracking feature to observe how these variables relate to changes in mood.

Visualize Changes in Mental Health Over Time

Several participants preferred to view their tracking data collected by the app visually to see how their mental health changed over time. They also wanted to view their data plotted at different calendar frequencies, such as weekly, monthly, or yearly intervals.

Reflective Journaling

Some participants recommended including a section for recording and reflecting on achievements and progress. They believed this would be valuable for their recovery and allow them to reflect on their progress and growth.

Share Treatment Progress With Their Psychologist

Some participants were keen to share their data from the DMHI with their psychologist to inform them of their progress between therapy sessions and initiate discussion in their sessions.

Theme 6: Safe, Trustworthy, and Effective

Tips and Activities to Help Feel Better Immediately

Participants mentioned that it would be useful to have a section in the DMHI where users could access tips and activities to help calm them down when they were feeling overwhelmed (eg, panicking or anxious). Suggestions included easy meditations, mood-lifting activities, and breathing exercises.

Confidence in the Credibility of the Program

Several participants emphasized that the content of the DMHI should be backed by scientific research and evidence. Some participants also highlighted that the intervention should be developed by a reputable organization and in collaboration with mental health professionals and researchers.

Access to Crisis Support Information if Needed

Some participants thought it was important to include crisis support information, such as mental health hotline numbers and websites.

User-Created Plan When Feeling Unwell

A few participants suggested that the DMHI should offer users the option to create a plan for challenging times or when life feels overwhelming. These participants recommended that the plan include options for pursuing personal interests and listing strategies or activities that they found helpful when they felt unwell.

XSL•FO
RenderX

Theme 7: Data Security and Control Over Information Sharing

Positive About Sharing Data With a Trusted Health Professional

Many participants were comfortable sharing data from the DMHI with a trusted health professional. Participants also highlighted several benefits of sharing the data with their health professional, such as streamlining information sharing, allowing the health professional to monitor their client's mood between sessions, and storing and sharing homework digitally with the health professional. All these benefits were thought to enhance face-to-face therapy.

Control Over Sharing Sensitive Information

Participants emphasized the private and personal nature of their mental health journey. Consequently, many participants thought it was important to have control over what information they could share from the DMHI and the ability to choose whom they could share it with. Participants also felt that sharing information should be an optional feature included in the DMHI and that this feature could be managed by the client.

Transparent Information About Data Privacy

A few participants were concerned about where the DMHI's data would be stored and who would oversee the data. These participants emphasized that it was important for the DMHI to be transparent with this information.

Theme 8: Must Be Low Cost or Free

Nearly every participant interviewed expressed reluctance to pay for a mental health DMHI. Some reported that the high cost of in-person therapy would mean that the cost of an app would be an additional financial burden for people seeking mental health treatment. Other participants hesitated to pay for a DMHI without knowing whether the intervention was effective.

Discussion

Principal Findings

The objective of this study was to use participatory design to incorporate the perspectives of people with lived and living experiences of depression and anxiety into a new DMHI that could be integrated into clinical practice as part of a blended care model. To the best of our knowledge, this research is the first Australian study to partner with adults with living and lived experiences of depression and anxiety and use participatory design to conceive a blended care DMHI. Our mixed methods approach also provided valuable insight into our participants' needs, preferences, and expectations of DMHIs and clarified how the new DMHI could be integrated with psychotherapy in a blended care model. Survey data highlighted how respondents wanted to access the DMHI and their preferences for features, content delivery, and sharing of information via the DMHI. Interview findings converged with survey data but also provided in-depth detail on preferred content and features and how to sustain engagement with the new DMHI. To date, the study findings have informed the designs for a new Australian blended care digital system called myNewWay [48], which includes a mental health smartphone app that can be used alone or as part of blended care.

DMHI Needs, Preferences, and Expectations

Survey findings indicated that participants preferred accessing a DMHI via a smartphone app, rather than on a computer or tablet. This preference and other findings may reflect the opinions of people who are familiar with digital technology, as the level of digital literacy was relatively high in the study sample. Two important factors for addressing the needs, preferences, and expectations of people with lived and living experiences of depression and anxiety were ensuring that the DMHI was well designed and easy to use. Qualitative interviews also highlighted that including a variety of therapeutic content was important to people with lived and living experiences of depression and anxiety. These findings are consistent with DMHI research conducted before, during, and after the COVID-19 pandemic [11,31,47,62], indicating these 3 factors should be considered in the design of DMHIs for depression and anxiety. Though no clear trends in the graphic design of the DMHI were identified in surveys and interviews, participants wanted the overall esthetic, tone, and feel of the DMHI to be positive, encouraging, and uplifting. This is an important consideration given that past research has found that the esthetic qualities of DMHI design are linked to better usability and user-friendliness [11,62].

Participants were keen for a DMHI to be user-friendly, simple to navigate, and lacking in technical issues, given that (a lack of) ease of use was often associated with past negative DMHI experiences and disengagement. However, most participants were currently using or had used DMHIs, which aligns with the relatively high levels of digital literacy in the study sample. Future research should give special consideration to defining whether these qualities can be addressed in a similar way for people with higher or lower levels of digital literacy. In addition, a novel finding from qualitative interviews that builds upon previous work is that DMHI users preferred to receive short, positive motivational messages within the DMHI or via push notifications (if app based). This finding is aligned with research conducted before the COVID-19 pandemic [63], which found short motivational messages to be the most commonly used component of the myCompass self-help program. While this preference for short, positive motivational messages may be specific to conditions before the COVID-19 pandemic and DMHI users with high digital literacy, existing literature does suggest that the language used in DMHIs should be hopeful and confident [64]. Moreover, people with depression and anxiety may be experiencing concentration and working memory problems alongside any mental health symptoms or co-occurring issues, further supporting the notion that the length of messages should not be overly long.

In this study, the appeal of the DMHI was also connected to whether the product helped to improve their mental health symptoms and well-being. These results support the findings of the systematic review by Borghouts et al [25], who found that the perceived effectiveness of a DMHI will affect whether a user engages with it. Participants also suggested that the DMHI could help them monitor their mental health symptoms via a

XSL•FO
RenderX

self-tracking feature (ie, mood and sleep), which is consistent with previous research [31,62]. One point of differentiation between this study and past research is that interview findings showed which aspects of self-tracking were most preferred, including the visualization of changes over time, reflective journaling, and the ability to share progress with a mental health professional. In qualitative interviews, participants also expressed a strong preference for diverse DMHI content that referred to wellness practices in addition to a wide range of therapeutic techniques, such as healthy eating, physical exercise, and gratitude practices. This desire for diverse and comprehensive content may have stemmed from many participants' previous experiences with DMHIs that focus on a single therapeutic modality, such as CBT or mindfulness. Participants also wanted the DMHI to include features that help facilitate social connections, which again was most evident in the qualitative interview findings. This is unsurprising, given the importance of social connectedness in positive mental health outcomes [33], especially for depression [34]. Features that support social connectedness were also desired as an inclusion in past DMHIs [11,25], and the need for isolation during the COVID-19 pandemic may have exacerbated this need. Suggestions included incorporating an online community where DMHI users could connect with peers with similar experiences or being able to read or listen to personal stories of people who had overcome mental health challenges.

In line with our secondary aim, participants identified 4 strategies in surveys and interviews that could enhance user engagement with a DMHI: personalization, reminders, goal setting, and including short and achievable activities within the DMHI. Consistent with previous research [11,31,62], participants wanted to personalize a DMHI to suit their needs and set goals. Unlike other studies that saw personalization as customizing the interface (ie, colors, theme, and music) and reminders, qualitative interviews showed that participants viewed personalization as the ability to pick and choose the content and activities or have the DMHI suggest them based on their responses. These findings are especially important, given that a lack of DMHI personalization is a common barrier to user engagement [25] and a common source of user dissatisfaction [11]. Participants preferred a modular, flexible design that could be individually tailored over a less flexible linear program where all users received the same content. Consistent with previous research [35,36,62], participants suggested using custom reminders to engage and maintain user interest in the DMHI in both surveys and qualitative interviews. For example, by allowing users to customize when reminders are delivered or by prompting the user to check in with the DMHI throughout the day. Furthermore, the finding from qualitative interviews that short and achievable activities would facilitate engagement with a DMHI was unique, yet important to consider, as users of DMHIs may be experiencing anxiety and depression and associated symptoms, such as rumination, fatigue, and low energy, which may hinder their ability to engage with longer activities. On the basis of their systematic review, Borghouts et al [25] proposed that shorter module durations were associated with better user engagement in self-guided DMHIs. However, this finding was only based on 2 studies, suggesting that more research is needed to verify this finding.

Integrating DMHIs Into Routine Care

The study findings provide insight into participants' preferences regarding whether and how they would use a DMHI with health professionals in a blended model of care. However, it is important to note that these findings represent the views of people who tended to have high digital literacy and were either currently using or had used DMHIs. Participants were open to sharing their information with health professionals via the DMHI, and there was no clear preference for a DMHI to be self-guided or guided in surveys and qualitative interviews. These findings are in line with the potential need for a program to be used flexibly, either in a self-guided or guided format, throughout their mental health recovery and navigation through mental health treatment. Survey data indicated that participants were also comfortable with and positive about sharing data from a DMHI with a trusted health professional, such as a psychologist, general practitioner, or psychiatrist. These findings extend prior research by highlighting that participants wanted autonomy and control over what information they could share, which is especially relevant given the rapid acceleration of remote care via DMHIs and other digital technology during the COVID-19 pandemic [12]. Several participants expressed concerns about how the data collected by the DMHI would be used, stored, and accessed. These results align with a blended care qualitative study [65], which found that clients were not concerned about sharing personal data via an app if their data were kept secure, anonymous, and private. Previous research on user perceptions of stand-alone DMHIs [25] and blended DMHIs [47] also confirmed that data security, confidentiality, and transparency affect user uptake of DMHIs both before and during the COVID-19 pandemic.

Other factors that impacted whether someone was willing to use a DMHI were the product's credibility, cost, the inclusion of crisis support, and immediate distress management. In qualitative interviews, participants emphasized that a DMHI should be backed by scientific research and developed by a reputable organization, consistent with previous research that showed that inadequate information about credibility can affect user trust in a DMHI [11]. The inclusion of a crisis support section was highlighted as important in this study, as well as in the European user-centered study by Behr et al [47]. The inclusion of this feature can act as an additional safety feature when the program is released into the community, especially if it is used by individuals with higher symptom severity and frequent suicidal ideation in a self-guided format. The addition of distress management tools that can help an individual feel calmer in the moment may also be helpful for people with varying severity levels and assist with developing emotion regulation and distress tolerance. Finally, in both surveys and qualitative interviews, we found that participants were reluctant to pay for a DMHI and strongly believed it should be free or offered at a very low cost (ie, Aus $1-$10 [US $0.65-$6.5]). This corresponds with existing research indicating that many studies have found that participants prefer free DMHIs [11], highlighting the need for future DMHIs to be accessed free of charge or at a low cost. Future research should explore the drivers of this low willingness to pay for DMHIs and understand whether it is due to the low perceived value of these

XSL•FO
RenderX

interventions, general expectations that mental health care should be free, or other factors. Future research also needs to explore the optimal funding and business models to support the ongoing maintenance required to successfully integrate DMHIs into routine clinical care, especially given the higher availability of DMHIs overall following the COVID-19 pandemic.

Limitations

Two main limitations may affect the generalizability of the findings of this study. The first limitation was the nature of the sample recruited for the study. Although the sample included people with a wide range of symptom severity, it predominantly consisted of female, English-speaking, and city-dwelling individuals, with most having previously received mental health treatment. Most participants had good self-rated digital literacy (229/306, 74.8%) and spent >1 hour per day online (287/306, 93.8%), and >40% (131/306) had also used DMHIs in the past. Their opinions may not be representative of the general community, who may be eventual users of DMHIs, especially those who do not usually use digital products or those who have never accessed mental health support. The second limitation was that recruitment for this study took place during the COVID-19 pandemic. Due to the need for many people to isolate before COVID-19 vaccines were widely available, the digital provision of mental health services was expedited during this time [12] (eg, telehealth). Whether the findings generalize to the mental health care system after the COVID-19 pandemic is unknown, and further research will be needed to verify this for this study and other DMHI research conducted before the COVID-19 pandemic.

Conclusions and Future Directions

This study provides clear insights into the needs, preferences, and expectations of adults with lived and living experiences of anxiety and depression for a transdiagnostic DMHI. Although our results are most relevant for people with high levels of digital literacy and experience with DMHIs, our findings highlight the importance of using mixed methods approaches to incorporate lived experience feedback into DMHI design and development. For people comfortable with digital technology, therapeutic interventions that improve mental health symptoms and are perceived as helpful, engaging, and effective are most desirable. Specific features that could enhance user engagement for a DMHI targeted at depression and anxiety include reminders, a self-tracking feature for symptoms, and content that facilitates social connection and peer support (eg, online forum and personal stories). The DMHI for depression and anxiety should also include access to support and resources that users can access when feeling unwell or in crisis. The potential for detection of severe symptomology, such as suicidality in depression, also supports the utility of a DMHI that a person can use as part of a blended care model with a mental health professional.

To date, findings from the participatory design study have been applied in subsequent parts of the larger Australian research project [48]. Designs for the new DMHI were subjected to a co-design and prototype testing process, which was then used to develop a smartphone app that is part of the myNewWay blended care digital system. The smartphone app provides therapeutic content and symptom management strategies primarily based on CBT, including psychoeducation, cognitive restructuring, emotion awareness and acceptance, goal setting, problem-solving, behavioral activation, exposure, relaxation, and mindfulness. Results from the participatory design study with mental health professionals were used similarly and informed the development of a health professional portal [52]. Phase 2 of the larger Australian research project involved an implementation clinical trial to evaluate the myNewWay app and Health Professional portal when used alongside psychological therapy for depression and anxiety. Currently, the myNewWay blended care system is being updated based on feedback from clients and psychologists involved in the clinical trial. The blended care system will also undergo further evaluation in a randomized controlled trial to compare the effectiveness of blended care with myNewWay to the self-guided use of the smartphone app for improving symptoms of depression and anxiety.

Acknowledgments

The authors thank Nyree Gale for her support with the Black Dog Institute Lived Experience Resource Network. JMN received funding from the Australian National Health and Medical Research Council (grant 2008839), the Australian Medical Research Future Fund (grant 1145382), the United Kingdom National Institute of Health Research, the Wellcome Trust, and Knowledge Partnership Platform Australia-Indonesia. AEW received funding from the Australian National Health and Medical Research Council (grant 2017521), the Australian Medical Research Future Fund (grant 1110773), and the Wellcome Trust. All other authors declare no other conflicts of interest. Funding for this study was also provided by the Ramsay Foundation.

Authors' Contributions

Conceptualization: JMN (lead), KO (equal)
Data curation: KK (lead), RKM (equal), JATM (supporting)
Formal analysis: KK (lead), RKM (equal), KO (equal), JATM (supporting)
Funding acquisition: JMN (lead), KO (supporting), NC (supporting)
Investigation: KO (lead), KK (supporting), KV (supporting)
Methodology: JMN (lead), KO (equal), KV (supporting), PB (supporting), SL (supporting), AEW (supporting)
Project administration: KO (lead), KK (supporting), RKM (supporting)
Resources: JMN (lead), NC (supporting)

XSL•FO
RenderX

Supervision: JMN (lead), KO (equal), AEW (supporting)
Validation: AEW (lead), SL (supporting), PB (supporting)
Visualization: RKM (lead), KK (supporting)
Writing – original draft: KK (lead), RKM (equal), JATM (supporting)
Writing – review and editing: JMN (lead), KK (equal), RKM (equal), JATM (supporting), SL (supporting), KV (supporting), PB (supporting), NC (supporting), AEW (supporting)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Web-based survey and interview questions.
[DOCX File , 41 KB - humanfactors_v12i1e69499_app1.docx]

References

1. World mental health report: transforming mental health for all. World Health Organization. 2022 Jun 16. URL: https://www.who.int/publications/i/item/9789240049338 [accessed 2024-08-13]
2. Saha S, Lim CC, Cannon DL, Burton L, Bremner M, Cosgrove P, et al. Co-morbidity between mood and anxiety disorders: a systematic review and meta-analysis. Depress Anxiety 2021 Mar 22;38(3):286-306 [FREE Full text] [doi: 10.1002/da.23113] [Medline: 33225514]
3. GBD 2019 Mental Disorders Collaborators. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Psychiatry 2022 Feb;9(2):137-150 [FREE Full text] [doi: 10.1016/S2215-0366(21)00395-3] [Medline: 35026139]
4. Arias D, Saxena S, Verguet S. Quantifying the global burden of mental disorders and their economic value. EClinicalMedicine 2022 Dec;54:101675 [FREE Full text] [doi: 10.1016/j.eclinm.2022.101675] [Medline: 36193171]
5. Cuijpers P, Miguel C, Ciharova M, Ebert D, Harrer M, Karyotaki E. Transdiagnostic treatment of depression and anxiety: a meta-analysis. Psychol Med 2023 Oct;53(14):6535-6546 [FREE Full text] [doi: 10.1017/S0033291722003841] [Medline: 36606456]
6. Moitra M, Owens S, Hailemariam M, Wilson KS, Mensa-Kwao A, Gonese G, et al. Global mental health: where we are and where we are going. Curr Psychiatry Rep 2023 Jul;25(7):301-311 [FREE Full text] [doi: 10.1007/s11920-023-01426-8] [Medline: 37256471]
7. Newby JM, Mason E, Kladnistki N, Murphy M, Millard M, Haskelberg H, et al. Integrating internet CBT into clinical practice: a practical guide for clinicians. Clin Psychol 2021 Mar 15;25(2):164-178. [doi: 10.1080/13284207.2020.1843968]
8. Weisel KK, Fuhrmann LM, Berking M, Baumeister H, Cuijpers P, Ebert DD. Standalone smartphone apps for mental health-a systematic review and meta-analysis. NPJ Digit Med 2019 Dec 2;2(1):118 [FREE Full text] [doi: 10.1038/s41746-019-0188-8] [Medline: 31815193]
9. Seegan PL, Miller MJ, Heliste JL, Fathi L, McGuire JF. Efficacy of stand-alone digital mental health applications for anxiety and depression: a meta-analysis of randomized controlled trials. J Psychiatr Res 2023 Aug;164:171-183. [doi: 10.1016/j.jpsychires.2023.06.019] [Medline: 37352813]
10. Paganini S, Teigelkötter W, Buntrock C, Baumeister H. Economic evaluations of internet- and mobile-based interventions for the treatment and prevention of depression: a systematic review. J Affect Disord 2018 Jan 01;225:733-755. [doi: 10.1016/j.jad.2017.07.018] [Medline: 28922737]
11. Alqahtani F, Orji R. Insights from user reviews to improve mental health apps. Health Informatics J 2020 Sep 10;26(3):2042-2066 [FREE Full text] [doi: 10.1177/1460458219896492] [Medline: 31920160]
12. Lattie EG, Stiles-Shields C, Graham AK. An overview of and recommendations for more accessible digital mental health services. Nat Rev Psychol 2022 Feb 26;1(2):87-100 [FREE Full text] [doi: 10.1038/s44159-021-00003-1] [Medline: 38515434]
13. Newby JM, Upton E, Mason E, Black M. Technology-based cognitive behavioral therapy interventions. Psychiatr Clin North Am 2024 Jun;47(2):399-417. [doi: 10.1016/j.psc.2024.02.004] [Medline: 38724127]
14. Erbe D, Eichert HC, Riper H, Ebert DD. Blending face-to-face and internet-based interventions for the treatment of mental disorders in adults: systematic review. J Med Internet Res 2017 Sep 15;19(9):e306 [FREE Full text] [doi: 10.2196/jmir.6588] [Medline: 28916506]
15. Moshe I, Terhorst Y, Philippi P, Domhardt M, Cuijpers P, Cristea I, et al. Digital interventions for the treatment of depression: a meta-analytic review. Psychol Bull 2021 Aug;147(8):749-786. [doi: 10.1037/bul0000334] [Medline: 34898233]
16. Pauley D, Cuijpers P, Papola D, Miguel C, Karyotaki E. Two decades of digital interventions for anxiety disorders: a systematic review and meta-analysis of treatment effectiveness. Psychol Med 2023 Jan 28;53(2):567-579 [FREE Full text] [doi: 10.1017/S0033291721001999] [Medline: 34047264]

XSL•FO
RenderX

17. Bae H, Shin H, Ji HG, Kwon JS, Kim H, Hur JW. App-based interventions for moderate to severe depression: a systematic review and meta-analysis. JAMA Netw Open 2023 Nov 01;6(11):e2344120 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.44120] [Medline: 37983028]
18. Linardon J, Torous J, Firth J, Cuijpers P, Messer M, Fuller-Tyszkiewicz M. Current evidence on the efficacy of mental health smartphone apps for symptoms of depression and anxiety. A meta-analysis of 176 randomized controlled trials. World Psychiatry 2024 Feb 12;23(1):139-149 [FREE Full text] [doi: 10.1002/wps.21183] [Medline: 38214614]
19. Zielasek J, Reinhardt I, Schmidt L, Gouzoulis-Mayfrank E. Adapting and implementing apps for mental healthcare. Curr Psychiatry Rep 2022 Sep;24(9):407-417 [FREE Full text] [doi: 10.1007/s11920-022-01350-3] [Medline: 35835898]
20. National digital mental health framework. Australian Government Department of Health and Aged Care. 2021 Jun. URL: https://www.health.gov.au/resources/publications/national-digital-mental-health-framework?language=en# [accessed 2024-08-13]
21. Buelens F, Luyten P, Claeys H, Van Assche E, Van Daele T. Usage of unguided, guided, and blended care for depression offered in routine clinical care: lessons learned. Internet Interv 2023 Dec;34:100670. [doi: 10.1016/j.invent.2023.100670] [Medline: 37767005]
22. Owusu JT, Wang P, Wickham RE, Varra AA, Chen C, Lungu A. Real-world evaluation of a large-scale blended care-cognitive behavioral therapy program for symptoms of anxiety and depression. Telemed J E Health 2022 Oct 01;28(10):1412-1420 [FREE Full text] [doi: 10.1089/tmj.2021.0590] [Medline: 35263185]
23. Lecomte T, Potvin S, Corbière M, Guay S, Samson C, Cloutier B, et al. Mobile apps for mental health issues: meta-review of meta-analyses. JMIR Mhealth Uhealth 2020 May 29;8(5):e17458 [FREE Full text] [doi: 10.2196/17458] [Medline: 32348289]
24. Ehrt-Schäfer Y, Rusmir M, Vetter J, Seifritz E, Müller M, Kleim B. Feasibility, adherence, and effectiveness of blended psychotherapy for severe mental illnesses: scoping review. JMIR Ment Health 2023 Dec 26;10:e43882 [FREE Full text] [doi: 10.2196/43882] [Medline: 38147373]
25. Borghouts J, Eikey E, Mark G, De Leon C, Schueller SM, Schneider M, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. J Med Internet Res 2021 Mar 24;23(3):e24387 [FREE Full text] [doi: 10.2196/24387] [Medline: 33759801]
26. Linardon J, Fuller-Tyszkiewicz M. Attrition and adherence in smartphone-delivered interventions for mental health problems: a systematic and meta-analytic review. J Consult Clin Psychol 2020 Jan;88(1):1-13. [doi: 10.1037/ccp0000459] [Medline: 31697093]
27. Baumel A, Edan S, Kane JM. Is there a trial bias impacting user engagement with unguided e-mental health interventions? A systematic comparison of published reports and real-world usage of the same programs. Transl Behav Med 2019 Nov 25;9(6):1020-1033. [doi: 10.1093/tbm/ibz147] [Medline: 31689344]
28. Christensen H, Griffiths KM, Korten AE, Brittliffe K, Groves C. A comparison of changes in anxiety and depression symptoms of spontaneous users and trial participants of a cognitive behavior therapy website. J Med Internet Res 2004 Dec 22;6(4):e46 [FREE Full text] [doi: 10.2196/jmir.6.4.e46] [Medline: 15631970]
29. Fleming T, Bavin L, Lucassen M, Stasiak K, Hopkins S, Merry S. Beyond the trial: systematic review of real-world uptake and engagement with digital self-help interventions for depression, low mood, or anxiety. J Med Internet Res 2018 Jun 06;20(6):e199 [FREE Full text] [doi: 10.2196/jmir.9275] [Medline: 29875089]
30. Sanatkar S, Baldwin PA, Huckvale K, Clarke J, Christensen H, Harvey S, et al. Using cluster analysis to explore engagement and e-attainment as emergent behavior in electronic mental health. J Med Internet Res 2019 Nov 28;21(11):e14728 [FREE Full text] [doi: 10.2196/14728] [Medline: 31778115]
31. Chan AH, Honey ML. User perceptions of mobile digital apps for mental health: acceptability and usability - an integrative review. J Psychiatr Ment Health Nurs 2022 Feb 18;29(1):147-168. [doi: 10.1111/jpm.12744] [Medline: 33604946]
32. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. Evid Based Ment Health 2018 Aug 05;21(3):116-119 [FREE Full text] [doi: 10.1136/eb-2018-102891] [Medline: 29871870]
33. Holt-Lunstad J. Social connection as a public health issue: the evidence and a systemic framework for prioritizing the "social" in social determinants of health. Annu Rev Public Health 2022 Apr 05;43:193-213 [FREE Full text] [doi: 10.1146/annurev-publhealth-052020-110732] [Medline: 35021021]
34. De Risio L, Pettorruso M, Collevecchio R, Collacchi B, Boffa M, Santorelli M, et al. Staying connected: an umbrella review of meta-analyses on the push-and-pull of social connection in depression. J Affect Disord 2024 Jan 15;345:358-368 [FREE Full text] [doi: 10.1016/j.jad.2023.10.112] [Medline: 37852587]
35. Gan DZ, McGillivray L, Larsen ME, Torok M. Promoting engagement with self-guided digital therapeutics for mental health: insights from a cross-sectional survey of end-users. J Clin Psychol 2023 May 24;79(5):1386-1397. [doi: 10.1002/jclp.23486] [Medline: 36693234]
36. Stawarz K, Preist C, Tallon D, Wiles N, Coyle D. User experience of cognitive behavioral therapy apps for depression: an analysis of app functionality and user reviews. J Med Internet Res 2018 Jun 06;20(6):e10120 [FREE Full text] [doi: 10.2196/10120] [Medline: 29875087]

XSL•FO
RenderX

37. Saleem M, Kühne L, De Santis KK, Christianson L, Brand T, Busse H. Understanding engagement strategies in digital interventions for mental health promotion: scoping review. JMIR Ment Health 2021 Dec 20;8(12):e30000 [FREE Full text] [doi: 10.2196/30000] [Medline: 34931995]
38. Andrews G, Basu A, Cuijpers P, Craske MG, McEvoy P, English CL, et al. Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: an updated meta-analysis. J Anxiety Disord 2018 Apr;55:70-78 [FREE Full text] [doi: 10.1016/j.janxdis.2018.01.001] [Medline: 29422409]
39. Scott S, Knott V, Finlay-Jones AL, Mancini VO. Australian psychologists experiences with digital mental health: a qualitative investigation. J Technol Behav Sci 2022 Aug 16;8(4):1-11 [FREE Full text] [doi: 10.1007/s41347-022-00271-5] [Medline: 35991293]
40. Newby JM, Twomey C, Yuan Li SS, Andrews G. Transdiagnostic computerised cognitive behavioural therapy for depression and anxiety: a systematic review and meta-analysis. J Affect Disord 2016 Jul 15;199:30-41. [doi: 10.1016/j.jad.2016.03.018] [Medline: 27060430]
41. Proudfoot J, Clarke J, Birch MR, Whitton AE, Parker G, Manicavasagar V, et al. Impact of a mobile phone and web program on symptom and functional outcomes for people with mild-to-moderate depression, anxiety and stress: a randomised controlled trial. BMC Psychiatry 2013 Nov 18;13:312 [FREE Full text] [doi: 10.1186/1471-244X-13-312] [Medline: 24237617]
42. Newby JM, Mewton L, Williams AD, Andrews G. Effectiveness of transdiagnostic internet cognitive behavioural treatment for mixed anxiety and depression in primary care. J Affect Disord 2014 Aug;165:45-52. [doi: 10.1016/j.jad.2014.04.037] [Medline: 24882176]
43. De Witte NA, Joris S, Van Assche E, Van Daele T. Technological and digital interventions for mental health and wellbeing: an overview of systematic reviews. Front Digit Health 2021 Dec 23;3:754337 [FREE Full text] [doi: 10.3389/fdgth.2021.754337] [Medline: 35005695]
44. Goldberg SB, Lam SU, Simonsson O, Torous J, Sun S. Mobile phone-based interventions for mental health: a systematic meta-review of 14 meta-analyses of randomized controlled trials. PLOS Digit Health 2022;1(1):e0000002 [FREE Full text] [doi: 10.1371/journal.pdig.0000002] [Medline: 35224559]
45. Brotherdale R, Berry K, Branitsky A, Bucci S. Co-producing digital mental health interventions: a systematic review. Digit Health 2024 Apr 25;10:20552076241239172 [FREE Full text] [doi: 10.1177/20552076241239172] [Medline: 38665886]
46. Vial S, Boudhraâ S, Dumont M. Human-centered design approaches in digital mental health interventions: exploratory mapping review. JMIR Ment Health 2022 Jun 07;9(6):e35591 [FREE Full text] [doi: 10.2196/35591] [Medline: 35671081]
47. Behr S, Fenski F, Boettcher J, Knaevelsrud C, Hammelrath L, Kovacs G, et al. TONI - One for all? Participatory development of a transtheoretic and transdiagnostic online intervention for blended care. Internet Interv 2024 Mar;35:100723 [FREE Full text] [doi: 10.1016/j.invent.2024.100723] [Medline: 38370289]
48. myNewWay®. Black Dog Institute. URL: https://www.blackdoginstitute.org.au/research-projects/my-new-way/ [accessed 2025-04-25]
49. Jacmon J, Malouff JM, Taylor N. Treatment of major depression: effectiveness of cognitive-behavioural therapy with an internet course as a central component. E J Appl Psychol 2010 Jan 03;5(2):1-8. [doi: 10.7790/ejap.v5i2.153]
50. Robertson L, Smith M, Castle D, Tannenbaum D. Using the internet to enhance the treatment of depression. Australas Psychiatry 2006 Dec;14(4):413-417. [doi: 10.1111/j.1440-1665.2006.02315.x]
51. Sethi S, Campbell AJ, Ellis LA. The use of computerized self-help packages to treat adolescent depression and anxiety. J Technol Hum Serv 2010 Aug 31;28(3):144-160. [doi: 10.1080/15228835.2010.508317]
52. Miller RM, O'Moore K, Kikas K, Matheson JA, Whitton AE, Baldwin P, et al. Needs and expectations for the myNewWayTM blended digital and face-to-face psychotherapy model of care for depression and anxiety (Part 2): mental health professional co-design study. JMIR Preprints Preprint posted online on January 15, 2025. [doi: 10.2196/preprints.68789]
53. Gan DZ, McGillivray L, Han J, Christensen H, Torok M. Effect of engagement with digital interventions on mental health outcomes: a systematic review and meta-analysis. Front Digit Health 2021;3:764079 [FREE Full text] [doi: 10.3389/fdgth.2021.764079] [Medline: 34806079]
54. Koh J, Tng GY, Hartanto A. Potential and pitfalls of mobile mental health apps in traditional treatment: an umbrella review. J Pers Med 2022 Aug 25;12(9):1376 [FREE Full text] [doi: 10.3390/jpm12091376] [Medline: 36143161]
55. Kählke F, Buntrock C, Smit F, Ebert DD. Systematic review of economic evaluations for internet- and mobile-based interventions for mental health problems. NPJ Digit Med 2022 Nov 23;5(1):175 [FREE Full text] [doi: 10.1038/s41746-022-00702-w] [Medline: 36424463]
56. Mental health care and Medicare. Services Australia. URL: https://www.servicesaustralia.gov.au/mental-health-care-and-medicare?context=60092 [accessed 2025-03-20]
57. Schoonenboom J, Johnson RB. How to construct a mixed methods research design. Kolner Z Soz Sozpsychol 2017;69(Suppl 2):107-131 [FREE Full text] [doi: 10.1007/s11577-017-0454-1] [Medline: 28989188]
58. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]

XSL•FO
RenderX

59. Lovibond PF, Lovibond SH. The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. Behav Res Therapy 1995 Mar;33(3):335-343. [doi: 10.1016/0005-7967(94)00075-U]
60. Osman A, Wong JL, Bagge CL, Freedenthal S, Gutierrez PM, Lozano G. The Depression Anxiety Stress Scales-21 (DASS-21): further examination of dimensions, scale reliability, and correlates. J Clin Psychol 2012 Dec 28;68(12):1322-1338. [doi: 10.1002/jclp.21908] [Medline: 22930477]
61. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport Exerc Health 2019 Jun 13;11(4):589-597. [doi: 10.1080/2159676x.2019.1628806]
62. Alqahtani F, Winn A, Orji R. Co-designing a mobile app to improve mental health and well-being: focus group study. JMIR Form Res 2021 Feb 26;5(2):e18172 [FREE Full text] [doi: 10.2196/18172] [Medline: 33635281]
63. Whitton AE, Proudfoot J, Clarke J, Birch MR, Parker G, Manicavasagar V, et al. Breaking open the black box: isolating the most potent features of a web and mobile phone-based intervention for depression, anxiety, and stress. JMIR Ment Health 2015 Mar 04;2(1):e3 [FREE Full text] [doi: 10.2196/mental.3573] [Medline: 26543909]
64. Bakker D, Kazantzis N, Rickwood D, Rickard N. Mental health smartphone apps: review and evidence-based recommendations for future developments. JMIR Ment Health 2016 Mar 01;3(1):e7 [FREE Full text] [doi: 10.2196/mental.4984] [Medline: 26932350]
65. Atik E, Schückes M, Apolinário-Hagen J. Patient and therapist expectations for a blended cognitive behavioral therapy program for depression: qualitative exploratory study. JMIR Ment Health 2022 Dec 30;9(12):e36806 [FREE Full text] [doi: 10.2196/36806] [Medline: 36583934]

Abbreviations

CBT: cognitive behavioral therapy
DASS-21: 21-item Depression Anxiety Stress Scale
DMHI: digital mental health intervention

Edited by A Kushniruk; submitted 01.12.24; peer-reviewed by M Dumont, I Reinhardt; comments to author 04.03.25; revised version received 23.03.25; accepted 17.04.25; published 11.06.25.

Please cite as:
Kikas K, O'Moore K, Miller RK, Matheson JAT, Li S, Varghese K, Baldwin P, Cockayne N, Whitton AE, Newby JM
Needs and Expectations for the myNewWay Blended Digital and Face-to-Face Psychotherapy Model of Care for Depression and Anxiety (Part 1): Participatory Design Study including People with Lived and Living Experience
JMIR Hum Factors 2025;12:e69499
URL: https://humanfactors.jmir.org/2025/1/e69499
doi:10.2196/69499
PMID:

©Katarina Kikas, Kathleen O'Moore, Rosemaree Kathleen Miller, Julie-Anne Therese Matheson, Sophie Li, Kathleen Varghese, Peter Baldwin, Nicole Cockayne, Alexis Estelle Whitton, Jill Maree Newby. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 11.06.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Associations Between Sleep Duration and Cognitive Function Among Older Adults: Cross-Sectional Study

Guolin Guo[1*], BSc; Yanling Jiang[1*], MSc; Jianteng Dong[1], MSc; Xu Zhao[1], BSc; Xiaofan Lai[1], BSc; Xiumei Wang[2], BSc; Hongguo Rong[1,3*], PhD; Jian Li[1*], PhD

[1]School of Traditional Chinese Medicine, Beijing University of Chinese Medicine, Beijing, China

[2]School of humanities, Beijing University of Chinese Medicine, Beijing, China

[3]Center for Evidence-Based Chinese Medicine, Beijing University of Chinese Medicine, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Hongguo Rong, PhD
Center for Evidence-Based Chinese Medicine
Beijing University of Chinese Medicine
No. 11 Beisanhuan East Road, Chaoyang District
Beijing, 102488
China
Phone: 86 (10) 6428675
Email: hgrong@hsc.pku.edu.cn

Abstract

Background: Sleep duration plays a crucial role in cognitive health and is closely linked to cognitive decline. However, the relationship between sleep duration and cognitive function in the Chinese population remains poorly understood.

Objective: This study aims to evaluate the association between sleep duration and cognitive function among middle-aged and older adults in China.

Methods: Using data from 15,526 participants in the 2020 China Health and Retirement Longitudinal Study, we used 3 composite indicators, encompassing episodic memory, mental acuity, and overall cognitive function to evaluate cognitive performance. Sleep duration per night, self-reported through face-to-face interviews, was also obtained. Adjustments were made using multiple generalized linear regression models, accounting for demographic, lifestyle, and health-related covariates.

Results: Among the 15,526 respondents analyzed, 53.02% (8232/15,526) were female and 46.98% (7294/15,526) were male, with an average age of 61.5 (SD 9.27) years. Those reporting sleep durations of 4 hours or less (β=−1.85, 95% CI −2.07 to −1.62), 5 hours (β=−0.55, 95% CI −0.78 to −0.33; *P*<.001), 9 hours (β=−1.78, 95% CI −2.17 to −1.39), and 10 hours or more (β=−3.01, 95% CI −3.39 to −2.63) per night had a significant negative relationship with cognitive function. In the adjusted model, the negative impact of long sleep (≥10 hours) on overall cognitive function became more pronounced (β=−3.01, 95% CI −3.39 to −2.63; *P*<.001), followed closely by extremely short sleep (≤4 hours; β=−1.85, 95% CI −2.07 to −1.62; *P*<.001).

Conclusions: This study reveals an inverted U-shaped relationship between sleep duration and global cognitive decline, indicating that cognitive function should be closely monitored in individuals with both short and long sleep durations. Consequently, public health strategies should prioritize the promotion of moderate sleep to mitigate the cognitive risks associated with aging, particularly in culturally specific contexts.

(JMIR Hum Factors 2025;12:e72886) doi:10.2196/72886

KEYWORDS

sleep; cognitive function; aging; China Health and Retirement Longitudinal Study; CHARLS

XSL•FO
RenderX

Introduction

Background

Cognitive impairment poses a significant challenge in the context of global aging, profoundly affecting societal development and individual well-being [1]. In China, approximately 30% of the middle-aged and older population, representing over 260 million people, experience cognitive decline [2-4], which undermines autonomy and imposes substantial economic burdens on families and society [5,6]. Among various influencing factors, sleep duration has emerged as a potential mitigator of cognitive decline, with both insufficient and excessive sleep being associated with impaired cognitive performance [7,8].

Studies conducted in the United States, such as the English Longitudinal Study of Aging (ELSA) and the Health and Retirement Study, have revealed that both short and long sleep durations adversely affect cognitive function [9,10]. Specifically, the ELSA study found that older adults who sleep less than 5 hours per night have significantly lower cognitive scores compared with those who sleep for 7 hours [11,12]. Various mechanisms have been proposed to explain these findings, including impairment of the brain's ability to clear metabolic waste due to insufficient sleep and the reflection of underlying health problems through excessive sleep [13-15]. However, these studies are predominantly focused on Western populations, and their findings may not be entirely applicable to other cultural contexts. In China, research on the relationship between sleep duration and cognitive function is relatively scarce and often relies on single cross-sectional data [16,17]. Furthermore, systematic evaluations of this relationship, particularly with respect to sex and age differences, are lacking. In addition, most existing studies are limited to specific age groups or fail to account for the moderating effects of demographic factors such as sex and age [18]. Thus, it is crucial to conduct a more comprehensive assessment of the complex relationship between sleep duration and cognitive function.

Objectives

Given the distinctions in historical ethnic, economic, and sociocultural background, Chinese adults had different sleep patterns contrasted with other countries [19,20]. This study uses nationally representative data from the China Health and Retirement Longitudinal Study (CHARLS) to systematically examine the relationship between sleep duration and cognitive function. By subdividing cognitive function into mental intactness and episodic memory, this study provides evidence to support improving the cognitive health among middle-aged and older people. We hypothesized that both short and longer sleep durations are strongly associated with cognitive decline.

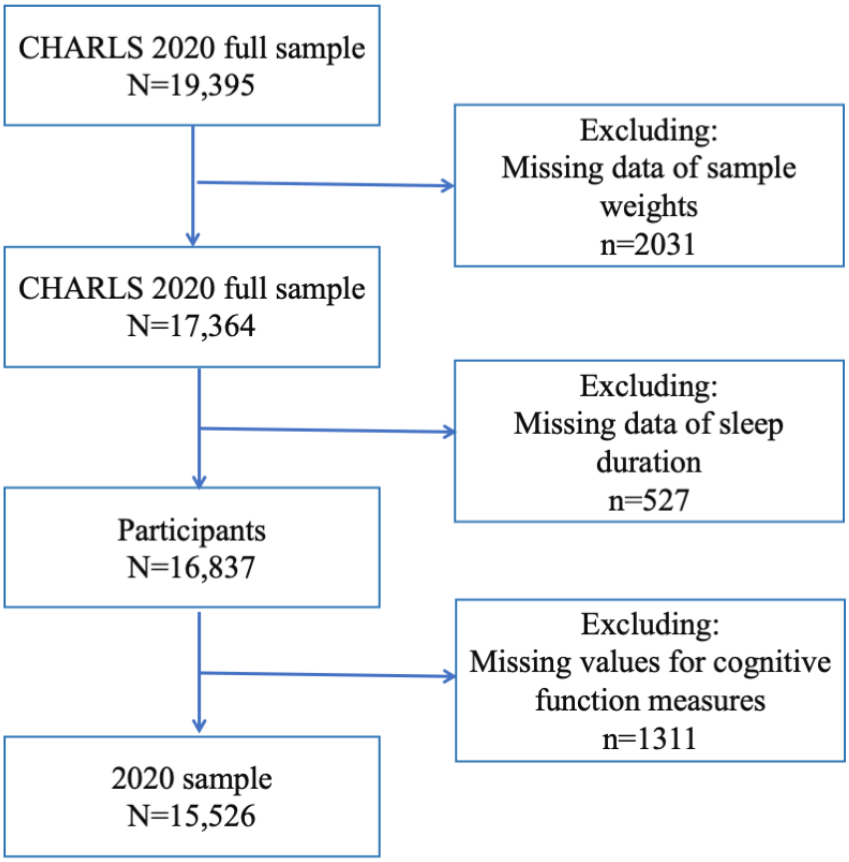
Methods

Study Design

This cross-sectional study used data from the 2018 survey of CHARLS, conducted between 2018 and 2020, which surveyed residents aged 45 years or above in 28 provinces across China [21,22]. The CHARLS focused on aging-related issues in China, the largest upper-middle-income country, and aimed to align with leading international research studies to ensure cross-study comparability of results [23]. Initiated in 2011, the CHARLS conducted follow-ups every 2 or 3 years, including in 2013, 2015, 2018, and 2020, to collect a nationally representative sample of the middle-aged and older adults in China, using a multistage stratified sampling technique with 450 villages distributed in 150 counties in 28 provinces. In this study, we used the latest CHARLS wave in 2020, with 19,816 individuals, to investigate the associations between sleep duration and cognitive function. Figure 1 shows a schematic flow diagram of the study sample [24].

XSL•FO
RenderX

Figure 1. The flowchart of the study sample from the 2020 China Health and Retirement Longitudinal Study (CHARLS).



Ethical Considerations

The CHARLS study protocol adhered to the principles of the Declaration of Helsinki and received ethical approval from the institutional review board of Peking University (ethical approval number IRB 00001052-11015) [25]. As this study involved a secondary analysis of existing data sets, ethical approval was not required under the research ethics policies and procedures of the London School of Economics and Political Science. This cross-sectional study follows the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline. All CHARLS participants provided written informed consent before data collection [26].

Cognitive Function

To evaluate cognitive performance, we used 3 composite measures based on the framework of the Health and Retirement Study and other cognitive function research, which include the word recall test, the Telephone Interview of Cognitive Status, and a pentagon-drawing test. The word recall test assessed episodic memory by requiring respondents to immediately recall 10 Chinese nouns (immediate recall; just read to them in any order) and then recall the same list with 4-minute delay (delayed recall). The episodic memory score was calculated as the average of immediate and delayed recall scores, ranging from 0 to 10. The Telephone Interview of Cognitive Status and a pentagon drawing test were used to assess mental intactness, focusing on time orientation, numerical ability, and visuospatial skills. Participants performed tasks such as serial subtractions, identifying the current data (month, day, year, and day of the week and season) and replicating a drawing of overlapping pentagons [27,28]. The scores from these tasks were combined into a single mental intactness score, ranging from 0 to 11. The global cognitive function was calculated as the sum of the episodic memory and mental intactness scores, with higher scores indicating better cognitive performance [29,30].

Sleep Duration

Drawing upon data from the ELSA conducted in England, CHARLS gathered self-reported sleep duration figures during face-to-face interviews using an open-ended format. Specifically, participants were queried about their average sleep hours per night over the past month: “In the past month, how many hours of sleep have you had on average per night?” In alignment with the previous CHARLS sleep study [31-34], responses were subsequently categorized into 7 sleep duration groups (≤4, 5, 6, 7, 8, 9, and ≥10 hours per evening). Daytime napping patterns were assessed through the question: “In the past month, how long have you napped on average after lunch?” leading to 4 napping duration categories: nonnappers (0 min), short nappers (<30 min), moderate nappers (30-90 min), and extended nappers (>90 min). In line with the previous CHARLS study on sleep, we divided respondents into 3 sleep duration categories in the analyses. The specific categorization is as follows: a sleep duration of 7 to 8 hours is considered the standard sleep group, less than 7 hours is considered the short sleep group, and more than 8 hours is considered the long sleep group [35].

Consistent with the approach adopted in a previous CHARLS on sleep patterns, we classified participants into 3 sleep duration

XSL•FO
RenderX

categories for analysis. The detailed categorization criteria are outlined as follows: individuals with a sleep duration of 7 to 8 hours are the standard sleep group, those sleeping less than 7 hours are the short sleep group, and those with a sleep duration exceeding 8 hours are the long sleep group.

Covariates

The CHARLS used a structured questionnaire to collect comprehensive sociodemographic and health-related data. The covariates considered potentially influential in shaping the associations included age; sex; educational attainment (elementary school and below, secondary school, and college and above); marital status (married vs others); residence (rural vs urban); smoking and drinking status (never vs current); visual and hearing impairment; self-reported general health status (ranging from very good to very poor); self-reported life satisfaction (spanning from completely satisfied to not at all satisfied); daytime napping habits; depression; and chronic disease condition (none, mild, and severe). In addition, sample weights were considered during the analysis to adjust for the representativeness of the sample.

Statistical Analysis

In this study, ANOVA was used for numerical variables, while ordered *t* test and chi-square tests were conducted for discrete variables to compare respondent characteristics among different sleep groups, specifically short sleepers, standard sleepers, and those with long sleep disorders. Continuous variables were summarized with means and SDs, while categorical variables were reported in terms of counts and percentages. To maximize statistical power, the sample size for analysis was allowed to vary based on the quantity of valid responses for each cognitive metric. For the continuous cognitive measures, a generalized linear model with a Gaussian distribution and an identity link function was used. A comprehensive suite of covariates, encompassing demographic factors such as age, sex, educational attainment, marital status, and lifestyle attributes, including smoking and alcohol consumption, were incorporated into our analytic framework. In addition, measures of functional independence based on activities of daily living, urban versus rural residency, self-assessed health status, life satisfaction, depressive symptomatology, and the presence of chronic medical conditions were included as covariates to mitigate potential confounding effects. Statistical significance was set at 2-sided $P<.05$. All data analyses were conducted using Stata (version 18.0; Stata Corp).

Results

Sample Characteristics

The characteristics of the surveyed individuals are shown in Table 1. The study sample included 15,526 respondents, including 8232 female participants (53.02%) and 7294 male participants (46.98%), with an average age of 61.5 (SD 9.27) years. Among them, 7808 individuals (N=15,526, 50.29%) belonged to the standard sleep group, 6034 (N=15,526, 41.92%) to the short sleep group, and 1084 (N=15,526, 6.98%) to the long sleep group. In terms of the cognitive function composite score, participants in the standard sleep group demonstrated the highest mean score of 11.03 (SD 4.34), followed by the short sleep group with a mean score of 10.32 (SD 4.52). Conversely, the long sleep group recorded the lowest mean score of 8.40 (SD 4.67). Regarding the mental intactness composite score, a similar pattern emerged, with the standard sleep group scoring highest (mean 6.81, SD 2.96), followed by the short sleep group (mean 6.35, SD 3.06) and the long sleep group recording the lowest score (mean 5.27, SD 3.14). In the domain of episodic memory, the standard sleep group achieved a mean score of 4.22 (SD 2.06), outperforming the short sleep group (mean 3.97, SD 2.08) and the long sleep group (mean 3.13, SD 2.19). Collectively, these findings underscore the superior performance of the standard sleep group across cognitive function, mental intactness, and episodic memory domains, compared with both the short and long sleep groups. Notably, the long sleep group consistently shows the lowest scores across all 3 cognitive domains, hinting at a potential association between sleep duration and cognitive health.

XSL•FO
RenderX

Table 1. Characteristics of the sample by night sleep duration from the 2020 China Health and Retirement Longitudinal Study (N=15,526).

Characteristics	Total sample	Short sleep group (n=6034)	Standard sleep group (n=7808)	Long sleep group (n=1084)
Cognitive function score (grade), mean (SD)	10.42 (4.52)	10.32 (4.52)	11.03 (4.34)	8.40 (4.67)
Mental intactness score (grade), mean (SD)	6.43 (3.06)	6.35 (3.06)	6.81 (2.96)	5.27 (3.14)
Episodic memory score (grade), mean (SD)	3.99 (2.07)	3.97 (2.08)	4.22 (2.06)	3.13 (2.19)
Age (y), mean (SD)	61.50 (9.27)	61.79 (9.16)	60.34 (9.20)	64.54 (9.75)
Sex, n (%)				
Male	7294 (46.98)	4124 (56.54)	2622 (35.95)	548 (7.5)
Female	8232 (53.02)	2510 (30.49)	5186 (63.00)	536 (6.51)
ADL[a] impairment, n (%)	3389 (21.83)	2393 (70.61)	768 (22.66)	228 (6.73)
Normal ADL, n (%)	12,137 (78.17)	6917 (56.99)	4364 (35.96)	856 (7.05)
IADL[b] impairment, n (%)	2957 (19.05)	2057 (69.56)	663 (22.42)	237 (8.01)
Normal IADL, n (%)	12,569 (80.95)	7253 (57.71)	4469 (35.56)	847 (6.74)
Education, n (%)				
Illiterate	3341 (21.52)	2037 (60.97)	934 (27.96)	370 (11.07)
Middle school	10.28 (66.21)	6143 (59.76)	3.49 (33.91)	651 (6.33)
High school	1905 (12.27)	1130 (59.32)	712 (37.38)	63 (3.31)
Married, n (%)	11,977 (77.14)	6981 (58.29)	4161 (34.74)	835 (6.97)
Smoke, n (%)	4018 (25.88)	7035 (61.13)	3713 (32.26)	760 (6.60)
Drink use, n (%)	5638 (36.31)	3342 (59.28)	1960 (34.76)	336 (5.96)
Rural areas, n (%)	11,794 (75.96)	6964 (59.05)	3732 (32.9)	942 (7.99)
Depressed, n (%)	6123 (39.44)	4324 (70.62)	1432 (23.39)	367 (5.99)
Feeling of own good health, n (%)	3734 (24.05)	1805 (48.34)	1602 (42.90)	327 (8.76)
Subjective well-being, n (%)	13,810 (88.95)	8069 (58.43)	4762 (34.48)	979 (7.09)
Chronic diseases, n (%)				
None	9936 (64.00)	5793 (58.30)	3458 (34.80)	685 (6.89)
Single chronic disease	5026 (32.37)	3126 (62.20)	1530 (30.44)	370 (7.36)
Multiple chronic diseases	564 (3.63)	391 (69.33)	144 (25.53)	29 (5.14)

[a]ADL: activities of daily living.

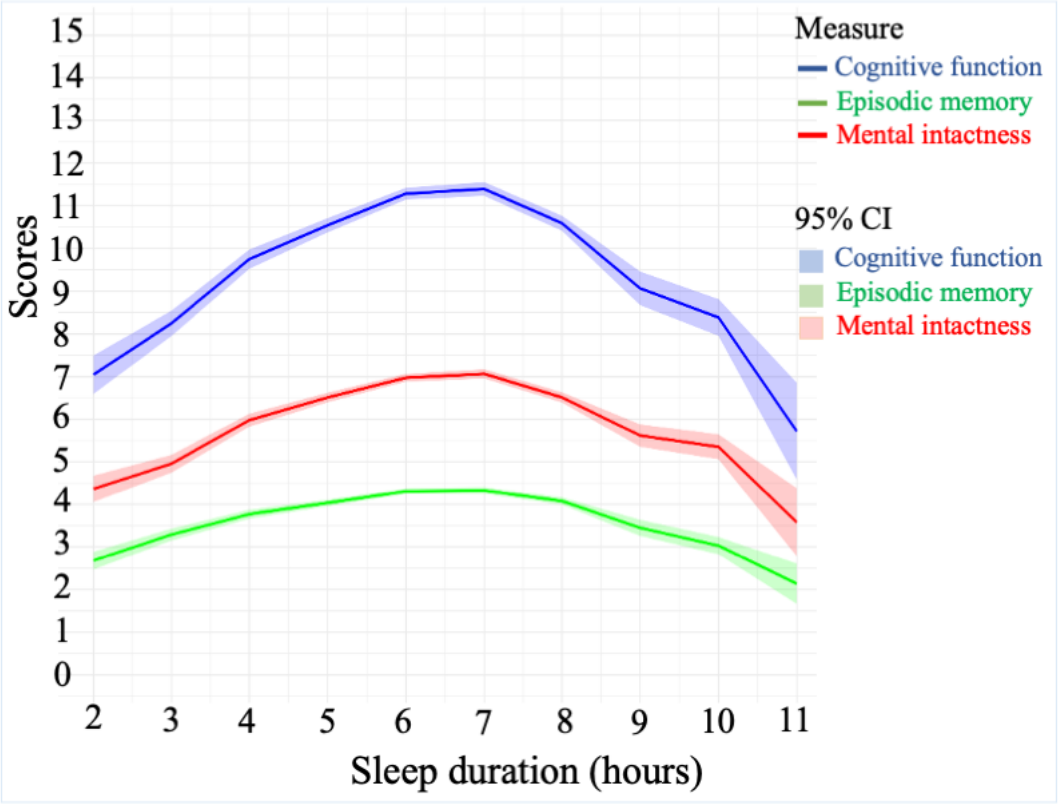
[b]IADL: instrumental activities of daily living; defined as having a score of 10 or above on the 10-item Center for Epidemiologic Studies Depression scale.

Trajectories for Sleep Duration Across Cognition Status

Figure 2 displays the trajectories of cognitive scores across different sleep durations, including cognitive function, mental intactness, and episodic memory. Overall, in 2020, the relationship between sleep duration and the scores of the selected population in cognitive function, mental intactness, and episodic memory exhibited an inverse U-shaped pattern.

XSL•FO
RenderX

Figure 2. Trajectories of cognitive scores across groups with different sleep durations.



In Table 2, in the unadjusted analysis, compared with individuals sleeping for 7 hours, those in the long sleep group (≥10 hours) exhibited the most significant negative impact on cognitive function (β=−1.85, 95% CI −2.17 to −1.52; *P*<.001). Subsequently, the 9-hour sleep group demonstrated the second-highest negative effect with β=−0.85 (95% CI −1.02 to −0.64; *P*<.001). Unhealthy sleep patterns had a detrimental influence on mental intactness than on episodic memory. Specifically, in the very short sleep group (≤4 hours), the negative effect on mental intactness was β=−0.61, whereas for episodic memory, it was β=−0.22. Intriguingly, no significant association was observed between cognitive function, episodic memory, and mental intactness with sleep duration in the 6-hour sleep group (*P*>.05). Upon adjustment, the long sleep group (≥10 hours) showed an exacerbated negative impact on overall cognitive function, with β increasing from −1.85 to −3.01, 95% CI −3.39 to −2.63; *P*<.001). The second most significant negative effect shifted from the 9-hour sleep group to the very short sleep group (≤4 hours) in the adjusted model with β=−1.85 (95% CI −2.07 to −1.62; *P*<.001). Consistent with the unadjusted findings, unhealthy sleep continued to have a greater adverse effect on mental intactness compared with episodic memory. The 6-hour sleep group maintained a nonsignificant association with cognitive function (*P*>.05).

XSL•FO
RenderX

Table 2. Associations between sleep duration and cognitive function in participants from the 2020 China Health and Retirement Longitudinal Study.

Model and sleep duration (hr)	Cognitive function		Episodic memory		Mental intactness	
	β	95% CI	β	95% CI	β	95% CI
Model 1[a]						
≤4	−0.83	−1.02 to −0.64	−0.22	−0.32 to −0.12	−0.61	−0.75 to −0.48
5	−0.24	−0.43 to −0.05	−0.03	−0.13 to −0.06	−0.21	−0.34 to −0.07
6	−0.08	−0.25 to −0.10	0.01	−0.08 to −0.10	−0.09	−0.21 to −0.03
7	1 (reference)	1 (reference)	1 (reference)	1 (reference)	1 (reference)	1 (reference)
8	−0.45	−0.64 to −0.25	−0.09	−0.19 to −0.01	−0.36	−0.49 to −0.22
9	−0.85	−1.18 to −0.52	−0.29	−0.46 to −0.13	−0.56	−0.79 to −0.33
≥10	−1.85	−2.17 to −1.52	−0.74	−0.91 to −0.57	−1.11	−1.34 to −0.88
Model 2[b]						
≤4	−1.85	−2.07 to −1.62	−0.58	−0.69 to −0.48	−1.26	−1.42 to −1.11
5	−0.55	−0.78 to −0.33	−0.16	−0.26 to −0.05	−0.40	−0.55 to −0.25
6	−0.12	−0.33 to −0.08	−0.01	−0.10 to −0.09	−0.11	−0.26 to −0.03
7	1 (reference)	1 (reference)	1 (reference)	1 (reference)	1 (reference)	1 (reference)
8	−0.79	−1.01 to −0.56	−0.21	−0.32 to −0.10	−0.57	−0.73 to −0.42
9	−1.78	−2.17 to −1.39	−0.61	−0.79 to −0.43	−1.17	−1.43 to −0.90
≥10	−3.01	−3.39 to −2.63	−1.15	−1.33 to −0.97	−1.86	−2.13 to −1.60

[a]Model 1 was unadjusted.

[b]Model 2 was adjusted for activities of daily living, instrumental activities of daily living, education, marital status, smoking, alcohol consumption, residence, self-rated health, self-rated life satisfaction, depression, chronic diseases condition, and sampling weights.

Discussion

Principal Findings

This nationally representative study, using the CHARLS database, reveals a statistically significant inverse U-shaped association between sleep duration and cognitive function. The findings reveal that the standard sleep group had the optimal cognitive scores, episodic memory, and mental intactness, whereas groups with either short or long sleep durations were significantly associated with cognitive decline, with long sleep having a more pronounced negative effect.

These findings align with those from major international studies, such as the ELSA and the Health and Retirement Study, which also identified the detrimental effects of both short and long sleep on cognitive function [12,36]. The consistent inverted U-shaped relationship between sleep duration and cognition suggests a universal phenomenon, albeit influenced by cultural and social contexts. However, it is important to recognize that cultural and social contexts may partly explain the observed variations in findings across different populations. For instance, in Chinese culture, familial support and older adult lifestyles may exacerbate the negative effects of prolonged sleep duration. The emphasis on close family bonds in Chinese society could inadvertently lead to excessive care or protection of older adults [37], potentially reducing their physical activity levels and contributing to cognitive decline. In addition, traditional health beliefs that promote prolonged rest as beneficial may further amplify the prevalence of long sleep among older adults in China [38]. These factors, coupled with a slower pace of life and lower levels of social engagement in some Chinese populations, present a unique cultural framework that differs from Western populations, where long sleep durations are less prevalent [39]. Thus, the findings of this study impart profound insights into how cultural and social factors influence the interplay between sleep duration and cognitive function, offering perspectives that are infrequently explored in the realm of international scholarly research.

From a mechanistic perspective, the adverse consequences of short sleep on cognitive function are well-documented. Short sleep may hinder the elimination of cerebral metabolic waste, including β-amyloid, a neurotoxic protein linked to Alzheimer disease [40,41]. Inefficient removal of such waste materials can disrupt neuronal function and impair cognitive processes. Moreover, sleep deprivation interferes with the equilibrium of neurotransmitters, including dopamine and serotonin, which are critical for mood regulation, attention, and decision-making [13]. These disruptions can culminate in impairments in executive function and emotional regulation, both integral aspects of mental well-being. Furthermore, hormonal dysregulation associated with insufficient sleep, such as elevated cortisol levels, has been shown to exert neurotoxic effects on the hippocampus, an essential brain region for memory consolidation. Moreover, decreased levels of growth hormone during sleep deprivation may also hinder tissue repair and neuronal regeneration, further compromising brain health [42].

XSL•FO
RenderX

The severity of short sleep's impact varies across different cognitive domains. Mental integrity, along with episodic memory, suffers from sleep deprivation; however, the disruptions to mental integrity appear more profound and systemic. Insufficient sleep not only impairs attention, executive function, and emotional stability but also exerts broad effects on mental health and cognitive capabilities, mediated by chronic stress and heightened inflammatory processes, ultimately leading to structural damage in the prefrontal cortex and neurotransmitter imbalances [43]. Conversely, while the impact on episodic memory is notable, it is relatively confined to specific processes such as memory formation and consolidation, often stemming from hippocampal dysfunction and increased susceptibility to memory interference. Therefore, the systemic nature of sleep deprivation's effects on mental integrity renders then more comprehensive and extensive when compared with its effects on episodic memory [44]. In contrast, long sleep duration is frequently associated with underlying health conditions that may indirectly contribute to cognitive decline. For instance, prolonged sleep has been linked to elevated inflammatory markers like C-reactive protein and metabolic dysregulation [45,46]. These biological responses could reflect the presence of chronic illnesses or suboptimal health states adversely affecting cognitive function. In addition, sex differences in physiological, hormonal, and social factors may further exacerbate cognitive disparities associated with long sleep [47,48]. Specifically, variations in estrogen levels and their neuroprotective effects could partly account for differing cognitive outcomes between men and women. Socially, women may encounter distinct roles and stressors that interact with sleep patterns, potentially influencing cognitive health in unique ways [49]. These findings highlight the need for a sophisticated appreciation of the interplay between biological, social, and sex-specific factors with sleep duration in influencing cognitive outcomes.

Since the declines are primarily centered on mental intactness, the resulting downstream consequences on instrumental activities of daily living warrant explicit consideration, particularly for older adults living alone. Low scores in these domains map onto tangible challenges in everyday life. Impairments in mental intactness, encompassing attention, and sequencing and visuospatial copying abilities, are closely tied to instrumental activities of daily living, which include managing medications and finances, keeping appointments, using the phone, navigating public transport, or following multistep cooking and home-safety routines. Deficits in episodic memory further undermine an individual's capacity for new learning and recall. This, in return, affects their adherence to treatment plans, remember emergency contacts, or recall whether gas or electric appliances have been turned off. Together, these impairments heighten risks that go far beyond a mere reduction in independence, which includes medication errors, missed or duplicate doses, falls resulting from disorientation, wandering or getting lost outdoors, malnutrition due to irregular meals, fire hazards, and vulnerability to financial scams. These risks are particularly pronounced among older adults living alone, underscoring the need for proactively identifying at-risk sleepers at both the clinical and community levels.

The study's strengths lie in its use of nationally representative data from the CHARLS, bolstering the external validity and generalizability of the findings. The substantial sample size reinforces the robustness of the results, enabling an in-depth analysis of sleep duration and its correlation with cognitive function among middle-aged and older adults. Importantly, this research contributes significant scientific evidence to inform sleep-based interventions aimed at promoting cognitive health. Translating these strengths into action, we outline implementable strategies for prediction, public communication, and prevention that public-sector and primary-care services can adopt now.

These findings provide a solid foundation for developing predictive models. They motivate nonlinear specifications (eg, restricted cubic splines or categorical cut-points at ≤4, 5, 6, 7 [reference], 8, 9, and ≥10 hours) combined with demographic, health, and psychosocial covariates already collected in population surveys and primary care. Models can be trained to estimate individual risk of low global cognition and domain-specific deficits (especially attentional or executive components captured by mental intactness), enabling risk stratification, providing early warning signals for clinicians, and targeting of behavioral counseling for those who stand to gain the most.

The study also offers data to support public awareness initiatives. Given the clear, nonlinear risk signal at both extremes of sleep duration and the relative sparing at approximately 7 hours, public health messaging can be framed in plain language (eg, "both insufficient and excessive sleep are linked with worse thinking and attention in later life"), while emphasizing that attention or executive skills appear especially vulnerable to chronic unhealthy sleep habits.

While acknowledging its strengths, several limitations merit recognition. First, the cross-sectional design of this study precludes the establishment of causal relationships between sleep duration and cognitive function, necessitating longitudinal studies to ascertain whether sleep duration directly influences cognitive trajectories over time or if other factors mediate this relationship. Second, the reliance on self-reported sleep data may introduce information bias, potentially affecting the accuracy and reliability of the results. Incorporation of objective sleep measures, such as actigraphy or polysomnography, could yield more precise assessments in future studies. Third, the relatively smaller sample size in older age groups may have contributed to increased variability in the findings, diminishing the statistical power to detect significant associations in this subgroup. Finally, the study did not account for other potential confounding factors, such as dietary habits, genetic predispositions, or environmental influences, which may play crucial roles in shaping the observed relationships and warrant further exploration.

In prioritizing future research endeavors, longitudinal study designs should be used to confirm causal relationships and examine potential mediators and moderators of the relationship between sleep duration and cognitive function. For instance, biomarkers indicative of inflammation and metabolic health could provide valuable insights into the underlying biological mechanisms. Similarly, lifestyle variables, including physical

XSL•FO
RenderX

activity, social engagement, and dietary habits, could elucidate the broader context of sleep-cognition interactions. In addition, comparative analyses across diverse cultural milieux could augment our comprehension of how sociocultural factors shape these relationships, ultimately guiding the development of tailored interventions to foster cognitive health in various populations.

Conclusions

In summary, this study reveals that moderate sleep duration is significantly protective of cognitive function in middle-aged and older adults. The identified risks associated with both short and long sleep underscore the necessity for individualized sleep health interventions tailored to personal characteristics. Public health policies should account for sex and age disparities to implement effective interventions aimed at enhancing cognitive health in aging populations, thereby alleviating the societal and economic burdens associated with global aging.

Acknowledgments

This work was supported by grant zyyzdxk-2023249 from the high-level traditional Chinese medicine key subject construction project of the National Administration of Traditional Chinese Medicine-Evidence-based Traditional Chinese Medicine, grant B2025019 from the Hebei Administration of Traditional Chinese Medicine, and grant YLZLXZ24G033 from the National Health Commission. The authors thank the China Health and Retirement Longitudinal Study team for providing data and simple training in using the database.

Authors' Contributions

HR and JL conceived the research design. GG and YJ conducted the data analysis and drafted the initial version of the manuscript. JD, XZ, and XL offered insights pertaining to the presentation of the findings and critically reviewed the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Wang AY, Hu HY, Ou YN, Wang ZT, Ma YH, Tan L, et al. Socioeconomic status and risks of cognitive impairment and dementia: a systematic review and meta-analysis of 39 prospective studies. J Prev Alzheimers Dis 2023 Jan;10(1):83-94. [doi: 10.14283/jpad.2022.81] [Medline: 36641612]
2. Zhang Y, Xu Y, Nie H, Lei T, Wu Y, Zhang L, et al. Prevalence of dementia and major dementia subtypes in the Chinese populations: a meta-analysis of dementia prevalence surveys, 1980-2010. J Clin Neurosci 2012 Oct;19(10):1333-1337. [doi: 10.1016/j.jocn.2012.01.029] [Medline: 22682650]
3. Wu YT, Beiser AS, Breteler MM, Fratiglioni L, Helmer C, Hendrie HC, et al. The changing prevalence and incidence of dementia over time - current evidence. Nat Rev Neurol 2017 Jun 12;13(6):327-339 [FREE Full text] [doi: 10.1038/nrneurol.2017.63] [Medline: 28497805]
4. Rosenberg A, Ngandu T, Rusanen M, Antikainen R, Bäckman L, Havulinna S, et al. Multidomain lifestyle intervention benefits a large elderly population at risk for cognitive decline and dementia regardless of baseline characteristics: the FINGER trial. Alzheimers Dement 2018 Mar 19;14(3):263-270. [doi: 10.1016/j.jalz.2017.09.006] [Medline: 29055814]
5. GBD 2015 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet 2016 Oct 08;388(10053):1545-1602 [FREE Full text] [doi: 10.1016/S0140-6736(16)31678-6] [Medline: 27733282]
6. Ding D, Zhao Q, Guo Q, Meng H, Wang B, Luo J, et al. Prevalence of mild cognitive impairment in an urban community in China: a cross-sectional analysis of the Shanghai Aging Study. Alzheimers Dement 2015 Mar 10;11(3):300-9.e2. [doi: 10.1016/j.jalz.2013.11.002] [Medline: 24613707]
7. Lo JC, Ong JL, Leong RL, Gooley JJ, Chee MW. Cognitive performance, sleepiness, and mood in partially sleep deprived adolescents: the need for sleep study. Sleep 2016 Mar 01;39(3):687-698 [FREE Full text] [doi: 10.5665/sleep.5552] [Medline: 26612392]
8. Farrell MT, Zahodne LB, Stern Y, Dorrejo J, Yeung P, Cosentino S. Subjective word-finding difficulty reduces engagement in social leisure activities in Alzheimer's disease. J Am Geriatr Soc 2014 Jun 02;62(6):1056-1063 [FREE Full text] [doi: 10.1111/jgs.12850] [Medline: 24890186]
9. Frauscher B, Mitterling T, Bode A, Ehrmann L, Gabelia D, Biermayr M, et al. A prospective questionnaire study in 100 healthy sleepers: non-bothersome forms of recognizable sleep disorders are still present. J Clin Sleep Med 2014 Jun 15;10(6):623-629 [FREE Full text] [doi: 10.5664/jcsm.3784] [Medline: 24932141]

XSL•FO
RenderX

10. Neufeld E, Hirdes JP, Perlman CM, Rabinowitz T. Risk and protective factors associated with intentional self-harm among older community-residing home care clients in Ontario, Canada. Int J Geriatr Psychiatry 2015 Oct 06;30(10):1032-1040. [doi: 10.1002/gps.4259] [Medline: 25663535]
11. Ferrie JE, Shipley MJ, Akbaraly TN, Marmot MG, Kivimäki M, Singh-Manoux A. Change in sleep duration and cognitive function: findings from the Whitehall II Study. Sleep 2011 May 01;34(5):565-573 [FREE Full text] [doi: 10.1093/sleep/34.5.565] [Medline: 21532949]
12. Miller MA, Wright H, Ji C, Cappuccio FP. Cross-sectional study of sleep quantity and quality and amnestic and non-amnestic cognitive function in an ageing population: the English Longitudinal Study of Ageing (ELSA). PLoS One 2014 Jun 26;9(6):e100991 [FREE Full text] [doi: 10.1371/journal.pone.0100991] [Medline: 24968354]
13. Xie L, Kang H, Xu Q, Chen MJ, Liao Y, Thiyagarajan M, et al. Sleep drives metabolite clearance from the adult brain. Science 2013 Oct 18;342(6156):373-377 [FREE Full text] [doi: 10.1126/science.1241224] [Medline: 24136970]
14. Shelton RC, Haman KL, Rapaport MH, Kiev A, Smith WT, Hirschfeld RM, et al. A randomized, double-blind, active-control study of sertraline versus venlafaxine XR in major depressive disorder. J Clin Psychiatry 2006 Nov 15;67(11):1674-1681. [doi: 10.4088/jcp.v67n1102]
15. Talbot LS, Maguen S, Metzler TJ, Schmitz M, McCaslin SE, Richards A, et al. Cognitive behavioral therapy for insomnia in posttraumatic stress disorder: a randomized controlled trial. Sleep 2014 Feb 01;37(2):327-341 [FREE Full text] [doi: 10.5665/sleep.3408] [Medline: 24497661]
16. Olaithe M, Pushpanathan M, Hillman D, Eastwood PR, Hunter M, Skinner T, et al. Cognitive profiles in obstructive sleep apnea: a cluster analysis in sleep clinic and community samples. J Clin Sleep Med 2020 Sep 15;16(9):1493-1505 [FREE Full text] [doi: 10.5664/jcsm.8564] [Medline: 32400387]
17. Liu H, Shi Y, Yu M, Guo X, Ruan Y, Qin F, et al. Individual and joint associations between sleep duration and physical activity with cognitive function: a longitudinal analysis among middle-aged and older adults in China. Alzheimers Dement 2025 Jan 18;21(1):e14212. [doi: 10.1002/alz.14212] [Medline: 39692596]
18. Ran W, Yu Q. Data-driven clustering approach to identify novel clusters of high cognitive impairment risk among Chinese community-dwelling elderly people with normal cognition: a national cohort study. J Glob Health 2024 Apr 19;14:04088 [FREE Full text] [doi: 10.7189/jogh.14.04088] [Medline: 38638099]
19. Gu D, Sautter J, Pipkin R, Zeng Y. Sociodemographic and health correlates of sleep quality and duration among very old Chinese. Sleep 2010 May;33(5):601-610 [FREE Full text] [doi: 10.1093/sleep/33.5.601] [Medline: 20469802]
20. Chen X, Wang R, Zee P, Lutsey PL, Javaheri S, Alcántara C, et al. Racial/ethnic differences in sleep disturbances: the multi-ethnic study of atherosclerosis (MESA). Sleep 2015 Jun 01;38(6):877-888 [FREE Full text] [doi: 10.5665/sleep.4732] [Medline: 25409106]
21. Yao Y, Wang K, Xiang H. Association between cognitive function and ambient particulate matters in middle-aged and elderly Chinese adults: evidence from the China Health and Retirement Longitudinal Study (CHARLS). Sci Total Environ 2022 Jul 01;828:154297 [FREE Full text] [doi: 10.1016/j.scitotenv.2022.154297] [Medline: 35288137]
22. Lin L, Wang HH, Lu C, Chen W, Guo VY. Adverse childhood experiences and subsequent chronic diseases among middle-aged or older adults in China and associations with demographic and socioeconomic characteristics. JAMA Netw Open 2021 Oct 01;4(10):e2130143 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.30143] [Medline: 34694390]
23. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the China Health and Retirement Longitudinal Study (CHARLS). Int J Epidemiol 2014 Feb;43(1):61-68 [FREE Full text] [doi: 10.1093/ije/dys203] [Medline: 23243115]
24. Kuo LY, Hsu PT, Wu WT, Lee RP, Wang JH, Chen HW, et al. The incidence of mental disorder increases after hip fracture in older people: a nationwide cohort study. BMC Geriatr 2021 Apr 15;21(1):249 [FREE Full text] [doi: 10.1186/s12877-021-02195-w] [Medline: 33858356]
25. Wu Z, Zhang H, Miao X, Li H, Pan H, Zhou D, et al. High-intensity physical activity is not associated with better cognition in the elder: evidence from the China health and retirement longitudinal study. Research Square Preprint posted online on August 18, 2021 [FREE Full text] [doi: 10.21203/rs.3.rs-806033/v1]
26. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Int J Surg 2014 Dec;12(12):1495-1499 [FREE Full text] [doi: 10.1016/j.ijsu.2014.07.013] [Medline: 25046131]
27. Pérez Palmer N, Trejo Ortega B, Joshi P. Cognitive impairment in older adults: epidemiology, diagnosis, and treatment. Psychiatr Clin North Am 2022 Dec;45(4):639-661. [doi: 10.1016/j.psc.2022.07.010] [Medline: 36396270]
28. Wittenberg R, Hu B, Jagger C, Kingston A, Knapp M, Comas-Herrera A, et al. Projections of care for older people with dementia in England: 2015 to 2040. Age Ageing 2020 Feb 27;49(2):264-269 [FREE Full text] [doi: 10.1093/ageing/afz154] [Medline: 31808792]
29. Martinelli JE, Cecato JF, Martinelli MO, de Melo BA, Aprahamian I. Performance of the Pentagon Drawing test for the screening of older adults with Alzheimer's dementia. Dement Neuropsychol 2018 Mar;12(1):54-60 [FREE Full text] [doi: 10.1590/1980-57642018dn12-010008] [Medline: 29682234]
30. Helmes E. Cognitive screening of older adults: the utility of pentagon drawing. Int Psychogeriatr 2013 Mar;25(3):413-419. [doi: 10.1017/s1041610212001998]

XSL•FO
RenderX

31. Jing R, Xu T, Rong H, Lai X, Fang H. Longitudinal association between sleep duration and depressive symptoms in Chinese elderly. Nature Sci Sleep 2020 Oct;Volume 12:737-747. [doi: 10.2147/nss.s269992]
32. Ma Y, Liang L, Zheng F, Shi L, Zhong B, Xie W. Association between sleep duration and cognitive decline. JAMA Netw Open 2020 Sep 01;3(9):e2013573 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.13573] [Medline: 32955572]
33. Rong H, Wang X, Lai X, Yu W, Fei Y. Associations between sleep duration and sensory impairments among older adults in China. Front Aging Neurosci 2022 Jun 7;14:910231 [FREE Full text] [doi: 10.3389/fnagi.2022.910231] [Medline: 35754970]
34. Guo J, Fei Y, Li J, Zhang L, Luo Q, Chen G. Gender- and age-specific associations between sleep duration and prevalent hypertension in middle-aged and elderly Chinese: a cross-sectional study from CHARLS 2011-2012. BMJ Open 2016 Sep 06;6(9):e011770 [FREE Full text] [doi: 10.1136/bmjopen-2016-011770] [Medline: 27601494]
35. Tian L, Ding P, Kuang X, Ai W, Shi H. The association between sleep duration trajectories and successful aging: a population-based cohort study. BMC Public Health 2024 Nov 01;24(1):3029 [FREE Full text] [doi: 10.1186/s12889-024-20524-7] [Medline: 39482676]
36. Garfield V, Llewellyn CH, Steptoe A, Kumari M. Investigating the bidirectional associations of adiposity with sleep duration in older adults: the English Longitudinal Study of Ageing (ELSA). Sci Rep 2017 Jan 09;7(1):40250 [FREE Full text] [doi: 10.1038/srep40250] [Medline: 28067295]
37. Yu Z, Fang B. Filial-piety-based family care in Chinese societies. In: He L, Gangopadhyay J, editors. Eldercare Issues in China and India. London, UK: Routledge; 2022.
38. Lin CY, Lee MY, Kao SF, Lu L, Chung A. The positive effects of activity involvement on the subjective well-being of Chinese older adults: a moderated effects model. Educ Gerontol 2023 Jan 02;49(9):790-802. [doi: 10.1080/03601277.2022.2163356]
39. Liu X, Liu L, Owens JA, Kaplan DL. Sleep patterns and sleep problems among schoolchildren in the United States and China. Pediatrics 2005 Jan;115(1 Suppl):241-249. [doi: 10.1542/peds.2004-0815F] [Medline: 15866858]
40. Ju YE, Ooms SJ, Sutphen C, Macauley SL, Zangrilli MA, Jerome G, et al. Slow wave sleep disruption increases cerebrospinal fluid amyloid-β levels. Brain 2017 Aug 01;140(8):2104-2111 [FREE Full text] [doi: 10.1093/brain/awx148] [Medline: 28899014]
41. Boespflug EL, Iliff JJ. The emerging relationship between interstitial fluid-cerebrospinal fluid exchange, amyloid-β, and sleep. Biol Psychiatry 2018 Feb 15;83(4):328-336 [FREE Full text] [doi: 10.1016/j.biopsych.2017.11.031] [Medline: 29279202]
42. McEwen BS, Karatsoreos IN. Sleep deprivation and circadian disruption stress, allostasis, and allostatic load. Sleep Med Clin 2022 Jun;17(2):253-262. [doi: 10.1016/j.jsmc.2022.03.005] [Medline: 35659078]
43. Lim J, Dinges DF. A meta-analysis of the impact of short-term sleep deprivation on cognitive variables. Psychol Bull 2010 May;136(3):375-389 [FREE Full text] [doi: 10.1037/a0018883] [Medline: 20438143]
44. Havekes R, Abel T. The tired hippocampus: the molecular impact of sleep deprivation on hippocampal function. Curr Opin Neurobiol 2017 Jun;44:13-19 [FREE Full text] [doi: 10.1016/j.conb.2017.02.005] [Medline: 28242433]
45. Lo JC, Groeger JA, Cheng GH, Dijk DJ, Chee MW. Self-reported sleep duration and cognitive performance in older adults: a systematic review and meta-analysis. Sleep Med 2016 Jan;17:87-98. [doi: 10.1016/j.sleep.2015.08.021]
46. Wang Z, Heizhati M, Wang L, Li M, Yang Z, Lin M, et al. Poor sleep quality is negatively associated with low cognitive performance in general population independent of self-reported sleep disordered breathing. BMC Public Health 2022 Jan 03;22(1):3 [FREE Full text] [doi: 10.1186/s12889-021-12417-w] [Medline: 34980052]
47. Hislop J, Arber SL. Sleep, gender, and aging: temporal perspectives in the mid-to-later life transition. In: Calasanti TM, Slevin KF, editors. Age Matters: Re-Aligning Feminist Thinking. Milton Park, UK: Routledge; 2006:225-246.
48. Badihian N, Badihian S, Poursafa P, Kelishadi R. Life-course effects of sleep on health promotion and disease prevention. In: Kelishadi R, editor. Healthy Lifestyle. Cham, Switzerland: Springer; 2022.
49. Henderson VW. Estrogen, cognition, and a woman's risk of Alzheimer's disease. Am J Med 1997 Sep 22;103(3A):11S-18S. [doi: 10.1016/s0002-9343(97)00261-1] [Medline: 9344402]

Abbreviations

CHARLS: China Health and Retirement Longitudinal Study
ELSA: English Longitudinal Study of Aging
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

XSL•FO
RenderX

Edited by J Waterson; submitted 20.02.25; peer-reviewed by J Griffiths, LF Belo; comments to author 30.07.25; revised version received 12.08.25; accepted 12.08.25; published 08.09.25.

Please cite as:
Guo G, Jiang Y, Dong J, Zhao X, Lai X, Wang X, Rong H, Li J
Associations Between Sleep Duration and Cognitive Function Among Older Adults: Cross-Sectional Study
JMIR Hum Factors 2025;12:e72886
URL: https://humanfactors.jmir.org/2025/1/e72886
doi:10.2196/72886
PMID:

©Guolin Guo, Yanling Jiang, Jianteng Dong, Xu Zhao, Xiaofan Lai, Xiumei Wang, Hongguo Rong, Jian Li. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 08.09.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Efficacy and Compliance of a Working Memory Multitasking Task Mobile Intervention for Children With Attention-Deficit/Hyperactivity Disorder: Single-Arm, Pre-Post Pilot Study

Minyoung Jung[1*], MS; Jimin Woo[1*], MA; Sung Jin Kim[1], MA; You Bin Lim[2], MD; Younglae Kim[1], MA; Dongwon Kang[1], PhD; Jung-sang Min[1], MS; Bung-Nyun Kim[2], MD, PhD

[1]eMotiv Co. Ltd, Seoul, Republic of Korea

[2]College of Medicine, Division of Child and Adolescent Psychiatry, Department of Psychiatry, Seoul National University Hospital, Seoul, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Bung-Nyun Kim, MD, PhD
College of Medicine
Division of Child and Adolescent Psychiatry, Department of Psychiatry
Seoul National University Hospital
101 Daehak-ro, Jongno-gu
Seoul, 03080
Republic of Korea
Phone: 82 1588 5700
Email: kbn1@snu.ac.kr

Abstract

Background: Attention-deficit/hyperactivity disorder (ADHD) is a prevalent neurodevelopmental disorder affecting 3%-7% of children globally. Alternative treatments are needed to address the limitations of traditional pharmacotherapy and nonpharmacotherapy, such as drug side effects and substantial time and financial costs. In this light, digital therapeutics for childhood ADHD are emerging as an effective alternative, with the benefits of potentially being free from serious side effects associated with software-based treatments and facilitating easy home use without constraints on time or space.

Objective: This study aims to evaluate whether a 4-week digital treatment program can improve symptoms, problem behaviors, and neurocognitive functions in children with ADHD, independent of medication status, while also gauging participant satisfaction with the program.

Methods: We recruited 22 Korean children aged 6-12 years with a diagnosis of ADHD. During the preintervention visit, we collected data on ADHD symptoms, relevant behavior scales, and neurocognitive assessments. Participants then used the program 5 times per day, 5 days a week for 4 weeks at home. At the postintervention visit, we collected the same data as during the preintervention visit and gathered additional feedback on their experience over the 4 weeks.

Results: A total of 19 participants were included in the statistical analysis, showing significant decreases in scores across various categories. These included the Korean ADHD Rating Scale (Total: P=.004; Inattentive: P=.004; and Hyperactive-impulsive: P=.01) and Korean Conners' Parent Rating Scale (Total: P<.001; Impulsive-hyperactive: P=.001; and Conduct Problem I: P=.04). Significant improvements were also noted in the Stroop word (P=.004), color (P<.001), and color-word (P<.001) scores. No significant differences in treatment effects were found between medicated and unmedicated participants. Caregiver and child satisfaction surveys yielded mean ratings of 4.3 and 4.1 out of 5, respectively.

Conclusions: A 4-week gamified intervention improves attention and hyperactivity-impulsivity in children with ADHD, irrespective of medication status, demonstrating its potential effectiveness and acceptability as a treatment option. As this is a small pilot study and underpowered, larger studies with appropriate control groups are needed in future research.

(JMIR Hum Factors 2025;12:e70479) doi:10.2196/70479

XSL•FO
RenderX

KEYWORDS

attention-deficit/hyperactivity disorder; digital therapeutics; cognitive training; gamification; efficacy; feasibility; multitasking; n-back

Introduction

Attention-deficit/hyperactivity disorder (ADHD) is a neurodevelopmental disorder characterized by persistent hyperactivity, impulsivity, or inattention that disrupts daily functioning [1]. The condition is heterogeneous, typically characterized by persistent patterns of inattention, hyperactivity/impulsivity, or a combination of both [2]. Most diagnoses occur early in life, with estimates indicating that approximately 3%-7% of children globally are diagnosed with ADHD [3,4]. Both in Korea and internationally, there has been a steady increase in ADHD diagnoses, primarily among children and adolescents [5,6].

In addition to the hallmark symptoms of persistent inattention and hyperactivity, ADHD can lead to cognitive and functional impairments in affected children, including challenges with response inhibition and working memory [2]. It also results in diminished academic performance and strained peer or family relationships, adding further distress to those affected [7]. Moreover, if not addressed, childhood ADHD can persist into adulthood, continuing to disrupt everyday life [8-10].

Child ADHD management strategies are categorized into pharmacological and nonpharmacological approaches. Although prescribed medication is frequently used to alleviate ADHD symptoms [10], parents often hesitate to medicate their children due to potential or actual side effects. These side effects commonly include decreased appetite, sleep disturbances, and stomach problems [4]. Furthermore, selecting the optimal medication for a particular patient can be time-consuming [10,11]. Nonpharmacological approaches include parental training, cognitive behavioral therapy, and cognitive training. These treatments can burden parents because they are both time-consuming and costly [12]. In addition, the scarcity of professional trainers makes access to nonpharmacological ADHD care more challenging. These limitations of traditional treatments have spurred the development of new ADHD treatments using advanced digital technologies.

Digital therapeutics (DTx) are defined as health software intended to treat or alleviate a disease, disorder, condition, or injury by generating and delivering a medical intervention that has a demonstrable positive therapeutic impact on a patient's health [13-15]. They offer high accessibility in terms of both time and location [16] and considerably reduce concerns about side effects due to their noninvasive nature. Moreover, it demands less supervisory effort than other behavioral therapies, since embedded algorithms can automatically adjust the usage instructions and program difficulty according to the user's level.

Gamification is an additional strategy used by DTx to enhance user engagement and experience [17,18]. Given that children with ADHD often display heightened novelty-seeking behaviors [19], they generally favor game-based activities. Presenting therapeutic interventions in a gamified format can greatly improve adherence to treatment regimens.

Previous research has consistently shown that various multitasking tasks can mitigate ADHD symptoms [20,21]. However, the widely used dual n-back task is recognized for its high level of difficulty and complex rules. These characteristics can make onboarding challenging for children—particularly those with ADHD—leading to early dropout, and sustaining motivation for continued training can be difficult. Given these challenges of conventional interventions, some researchers developed a gamified intervention as an alternative treatment. One of the most notable approaches was EndeavorRx, a Food and Drug Administration (FDA)-cleared DTx, developed by Akili Interactive for children with ADHD, which is effective in mitigating inattentive symptoms [21-23]. It was developed by incorporating gamified cognitive training that involved multitasking Go/No-Go tasks and sensory-motor tasks [24]. The results showed that digital-based multitasking tasks, focusing on perceptual discrimination and sensory-motor coordination, were effective in enhancing attention in both older adults and children with ADHD [21,24].

Notably, the driving task effectively increases cognitive load while maintaining an engaging experience for children. Nevertheless, this type of multitasking approach also presents limitations, particularly in providing finely tuned difficulty adjustments. By implementing a difficulty adjustment system, it is possible to engage a broader range of age groups at different stages of cognitive development. This approach also helps address cultural differences related to disparities in access to digital technologies.

Therefore, we present a gamified DTx named "StarRuckus" (eMotiv). It aims to reduce core ADHD symptoms through the simultaneous execution of n-back and driving tasks, including a personalized difficulty adjustment algorithm to ensure training benefits [21,25,26]. It functions as an enhanced version of the Go/No-Go task (0-back task), allowing for the customization of difficulty levels through varying n levels [27,28]. While difficulty adjustment in Go/No-Go tasks typically involves manipulating factors such as stimulus presentation time, the number of stimuli, and the number of discriminative targets, the n-back task allows for additional parameters, including the n level and the inclusion of lure stimuli.

To the best of our knowledge, this study is the first to implement multitasking that combines an n-back task with a driving task as a cognitive training intervention aimed at improving ADHD symptoms. Computerized multitasking working memory training is recognized for its effectiveness in alleviating core symptoms of ADHD, such as inattention [29] and hyperactivity-impulsivity [20]. Multitasking n-back training strengthens the right inferior frontal gyrus, an area exhibiting structural [30,31] and functional deficits in children with ADHD [32,33]. Furthermore, multitasking driving training can enhance neural activity efficiency in the prefrontal cortex [24,34], thereby improving attention in children with ADHD [21,26,35].

XSL•FO
RenderX

Another distinctive feature of StarRuckus is its focus on the reward system within its gamification elements, setting it apart from other DTx interventions. After providing a personalized difficulty level tailored to each user, the system delivers differentiated in-game rewards based on performance outcomes. These rewards can be collected and used to unlock additional content, such as character purchasing and town customization.

This pilot study explored the effectiveness of StarRuckus as a potential treatment for children with ADHD, both medicated and unmedicated, while assessing its appeal to both children and their caregivers. We evaluated improvements in symptomatology and behavior scales as well as cognitive-neurological indicators before and after a 4-week intervention with StarRuckus. In addition, feedback on usage experience and safety was gathered from children and caregivers. Through this study, we aimed to provide preliminary evidence supporting the efficacy and acceptability of StarRuckus as a DTx.

Methods

Participants

Two cohorts, 1 taking medication and 1 without medication, were included in the study. The participants in both cohorts agreed not to alter their medication status during the study, understanding that any change could result in their withdrawal. Although this agreement was not binding, participants were permitted to change their medication status after consulting with their doctors. They were also required to report any medication changes during the study. Inclusion criteria for the participants consisted of (1) a confirmed *DSM-5* diagnosis of ADHD; (2) age between 6 and 12 years; and (3) either consistently taking medication for ADHD for at least 30 days or not taking psychotropic medication for at least 3 days to allow for a washout before baseline measures. Exclusion criteria for the participants included (1) current comorbid psychiatric diagnoses listed in *DSM-5* with marked symptoms; (2) initiation of behavior therapy within the last 4 weeks or plans to change or start behavior therapy during the study; (3) motor conditions reported by parents that prevent the use of digital interventions; (4) color blindness as reported by caregivers; (5) uncorrected visual or auditory impairments; and (6) any other medical condition that might confound the study data.

Procedures

Participants were recruited from a community child center and a web-based community for caregivers of children diagnosed with neurodevelopmental disorders. Children and caregivers willing to participate completed a web-based survey to check eligibility. In addition, caregivers had to provide clinic-issued documentation as proof of their children's medical histories. If children were eligible for the study, caregivers were contacted to provide instructions about the course of the study.

The study consisted of 3 phases: a preintervention visit, the intervention period, and a postintervention visit. At the preintervention visit, participants and their caregivers received detailed information about the research purpose, eligibility criteria, procedures, potential risks and benefits, and confidentiality aspects. After comprehending the study protocol, they provided informed consent. Subsequently, the caregiver and child were placed in separate rooms for data collection. While parents completed questionnaires regarding their children's behavior, the children engaged in several neurocognitive tasks primarily focused on attention and working memory. Upon completion of the data collection tasks, children tested StarRuckus for approximately 10 minutes, with supervision to ensure compliance with the rules. At the conclusion of the preintervention period, tablets preinstalled with StarRuckus were lent to the participants for use during the intervention period. Throughout the intervention phase, participants had to complete 5 daily sessions, each lasting about 20-25 minutes, 5 days a week for 4 weeks. Compliance was electronically monitored by researchers, and reminder text messages were sent to parents when they were likely to fall short of the required session counts. In addition, caregivers and participants were instructed to report any adverse events to investigators in real time. The postintervention visit mirrored the preintervention visit, including neurocognitive tasks for the participants and questionnaires for the parents. Following data collection, satisfaction and inconvenience questionnaires were completed, incorporating reports of any adverse events. The children's questionnaires were designed for readability, and researchers assisted the children with any difficulties.

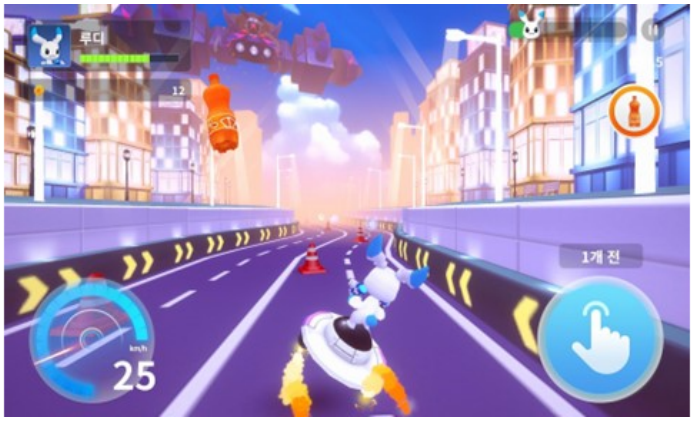
Digital Intervention

StarRuckus (Figure 1) is a gamified training program designed for children aged 6 to 12. It is a digital application available for download on both Android (Google LLC) and iOS (Apple Inc) platforms, aimed at mitigating ADHD symptoms in children. The program includes two main tasks: (1) a continuous visuomotor driving task and (2) an n-back task [36], which engages working memory. In the driving task, users maneuver the character by tilting the device left or right. The primary objective of this task is to dodge as many obstacles as possible while collecting target points. The n-back task, a working memory exercise, challenges participants to identify whether each presented stimulus matches one shown several steps earlier in the sequence [37,38]. The maximum difficulty level in the app goes up to 4-back. Once all 5 sessions are completed, the training program ends automatically.

In a cognitive training program, it is crucial to maintain a level of difficulty that aligns with the users' skill levels. The adaptive algorithm of StarRuckus adjusts task difficulty in real time based on individualized performance metrics tailored to each task type. For instance, in n-back tasks, increased difficulty may involve raising the n level or modifying stimulus parameters, such as increasing lure frequency or shortening stimulus duration. In driving tasks, difficulty can be modulated through speed changes. When users demonstrate improved performance, the system gradually introduces more challenging levels. Conversely, if performance declines, the difficulty is reduced to keep the tasks manageable and maintain engagement.

XSL•FO
RenderX

Figure 1. StarRuckus in-game screenshot.



In addition to providing internal rewards through a gamified treatment designed to entertain children, the program included various in-app rewards to maintain participant engagement. For example, children were able to build their own villages with the resources gained during the training process. They could also use these resources to outfit new space vehicles and characters.

Measures

Korean ADHD Rating Scale

A Korean ADHD Rating Scale (K-ADHD-RS) was used, originally developed by DuPaul in 1991 [39]. The K-ADHD-RS, comprising 18 questions, is a prevalent measure for assessing symptoms of ADHD. There are 2 subscales, 1 for inattention and 1 for hyperactivity. The Korean version, validated and proven reliable in 2002 [40,41], is extensively used in Korea to support ADHD diagnosis in children. In the scale, 0 corresponds to “rarely or never,” 1 to “somewhat,” 2 to “often,” and 3 to “always or very often.”

Korean Conners’ Parent Rating Scale

We used the Korean Conners’ Parent Rating Scale (K-CPRS), originally developed in 1970 [42,43]. It is a widely recognized tool for assessing problematic behaviors in children. Although analogous to the original, the Korean edition features a different distribution of questions across factors. According to the factor analysis results of the Korean edition [44], our assessment focused on 5 identified factors: Impulsive-hyperactive, Conduct Problem I, Conduct Problem II, Psychosomatic, and Anxiety [45].

ADHD Diagnostic System

The ADHD diagnostic system (ADS) is a computerized attention test designed to facilitate the diagnosis and assessment of treatment effects for ADHD. It is a type of continuous performance test [46,47] and features both ADS visual and auditory (ADS-A) modules, each containing 3 distinct stimuli (1 target and 2 nontargets). The proportion of targets was controlled to appear at 22% in the first part, 50% in the second part, and 78% in the final part of the task, respectively. Participants were instructed to press the spacebar when the target stimulus appeared and to ignore nontarget stimuli. Key outcome measures in our analyses included omission error, commission error, response time (RT), and RT SD [45].

Stroop Color and Word Test

In this study, we used the Korean version of the Stroop Color and Word Test [48], validated for children aged 5-14 years and originally developed in 1978 [49]. The Stroop test, primarily assessing selective attention and processing speed, consists of 3 card sessions: Word, Color, and Color-Word. Each session presents 100 stimuli, and children are timed to read them as quickly as possible within a 45-second limit. Specifically, children are instructed to read black words denoting color names (red, green, and blue) in the first session, name the printed color marked with “XXXX” in red, green, and blue in the second session, and name the printed color instead of the incongruently printed word in the final session. Raw scores from the word, color, and color-word sessions, along with the interference score, were recorded for data analysis.

Digit Span

The digit span (DS) test, a subcomponent of the Wechsler Intelligence Scale for Children-Fifth Edition [50], was administered using the Korean standardized version in this study [51]. The DS test, which involves memorizing a sequence of digits beginning with 2 digits and increasing by 1 digit at each step through 2 trials at each level until both are incorrect, was generally used to evaluate working memory and consists of 3 subtasks: DS Forward, DS Backward, and DS Sequencing. In the forward task, the examiner reads numbers aloud, and the participant recalls them in the order presented. The backward and sequencing tasks follow a similar procedure to the forward task but require participants to recall the sequence in reverse order or reorder the digits numerically, respectively. These tasks are generally more challenging than the forward task as they necessitate manipulating and reorganizing the received information. Scoring was based on the total number of successful trials across all subtasks.

Statistical Analysis

Behavioral and cognitive outcome measures were analyzed using open-source statistical libraries in Python (version 3.9.12; Python Software Foundation). Depending on the Shapiro-Wilk normality test results, either a 1-sample *t* test (parametric) or a Wilcoxon rank-sum test (nonparametric) was performed. A 2-tailed significance level of $P<.05$ was considered statistically significant, and Cohen *d* effect sizes were calculated.

XSL•FO
RenderX

Ethical Considerations

This study received approval from the Public Institutional Review Board designated by the Ministry of Health and Welfare, Korea (approval no P01-202305-01-007). Written informed consent was obtained from caregivers of all participating children, and assent was obtained from children when appropriate. Participants were informed about the study purpose, procedures, potential risks, and their right to withdraw at any time without penalty.

Results

Overview

The study is reported in accordance with the CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile HEalth Applications and onLine TeleHealth) checklist (Multimedia Appendix 1).

Participants Flow

A total of 22 Korean participants were enrolled in the study following screening. During the intervention, 1 participant withdrew with parental consent. Two participants experienced changes in their medication status: 1 resumed medication during the study after not taking any at the preassessment, and another changed their medication type. A total of 19 participants completed the study, as shown in the participant flowchart (Figure 2). Table 1 presents the demographic and clinical characteristics of the participants at baseline.

Figure 2. Flowchart of participants.

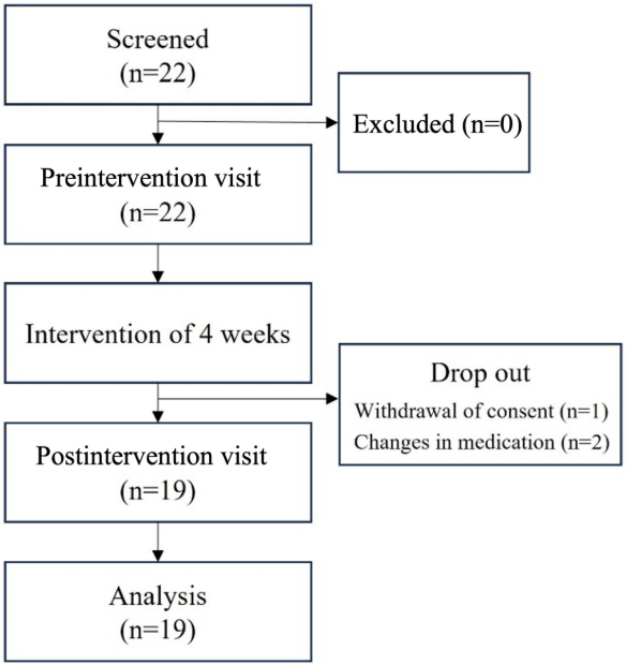


Table 1. Demographic and clinical characteristics of the participants.

Characteristic	Count
Sex , n (%)	
Male	14 (73.7)
Female	5 (26.3)
Age (years), mean (SD)	8.5 (1.5)
IQ[a], mean (SD)	99.4 (16.7)
Medication, n (%)	
On medication	10 (52.6)
No medication	9 (47.4)

[a]IQ: intelligence quotient.

XSL•FO
RenderX

Parent Rating Scales

K-ADHD-RS

For the K-ADHD-RS, an ADHD symptom measure rated by parents, a paired *t* test assessed changes in the Total, Inattentive, and Hyperactive-impulsive scores before and after 4 weeks of intervention. There were statistically significant reductions in Total (*P*=.004; effect size *d*=0.69, 95% CI –9.27 to –1.99), Inattentive (*P*=.004; effect size *d*=0.77, 95% CI –5.27 to –1.15), and Hyperactive-impulsive scores (*P*=.01; effect size *d*=0.47, 95% CI –4.22 to –0.62; Table 2).

Table 2. Changes in attention-deficit/hyperactivity disorder (ADHD)-related symptoms and challenging behaviors of participants following the digital intervention.

Measure and factor	Pre mean (SD)	Post mean (SD)	*P* value	Effect size *d*
K-ADHD-RS[a]				
Total	28.00 (8.78)	22.37 (7.55)	.004[c]	0.69
Inattentive	15.74 (4.45)	12.53 (3.84)	.004[c]	0.77
Hyperactive-impulsive	12.26 (5.42)	9.84 (4.79)	.01[c]	0.47
K-CPRS[b]				
Total	43.84 (14.57)	34.79 (10.34)	<.001[c]	0.72
Impulsive-hyperactive	14.00 (5.94)	10.26 (4.61)	.001[d]	0.70
Conduct problem I	9.00 (4.05)	7.53 (3.27)	.04[c]	0.40
Conduct problem II	0.21 (0.52)	0.26 (0.55)	.71[d]	-0.10
Psychosomatic	2.95 (2.50)	2.00 (2.34)	.15[c]	0.39
Anxiety	8.42 (3.76)	7.26 (3.11)	.06[d]	0.34

[a]K-ADHS-RS: Korean ADHD Rating Scale.

[b]K-CPRS: Korean Conners' Parent Rating Scale.

[c]*P* value derived from the *t* test.

[d]*P* value derived from the Wilcoxon signed-rank test.

K-CPRS

Postintervention changes in the Total score and the Impulsive-hyperactive, Conduct Problem I, Conduct Problem II, Psychosomatic, and Anxiety subscales from the K-CPRS, which assesses childhood behavior problems, were analyzed. Significant improvements were observed in Total (*P*<.001; effect size *d*=0.72, 95% CI –13.84 to –4.26), Impulsive-hyperactive (*P*=.001; effect size *d*=0.70, 95% CI –5.68 to –1.80), and Conduct Problem I (*P*=.04; effect size *d*=0.40, 95% CI –2.90 to –0.05) scales (Table 2).

Neurocognitive Tests

ADS

For the ADS in the visual and auditory domains, postintervention changes—including omission error, commission error, mean of correct RT, and SD of correct RT for both visual and auditory tasks—were analyzed. Apart from a decrease in auditory commission error, which approached significance (*P*=.05; effect size *d*= 0.59), no other statistically significant or notable changes were observed (Table 3).

XSL•FO
RenderX

Table 3. Changes in neurocognitive test performances of participants with ADHD after the digital intervention.

Measure and factor	Pre mean (SD)	Post mean (SD)	*P* value	Effect size *d*
ADS[a]-visual				
Omission error	0.05 (0.08)	0.09 (0.13)	.35[b]	–0.34
Commission error	0.08 (0.12)	0.06 (0.07)	.42[b]	0.26
RT[c] mean	0.61 (0.28)	0.64 (0.29)	.71[b]	–0.10
RT SD	0.22 (0.11)	0.25 (0.15)	.34[d]	–0.20
ADS[a]-auditory				
Omission error	0.25 (0.17)	0.30 (0.22)	.60[b]	–0.26
Commission error	0.16 (0.10)	0.10 (0.09)	.05[d]	0.59
RT mean	1.11 (0.17)	1.18 (0.09)	.06[d]	–0.53
RT SD	0.31 (0.10)	0.32 (0.14)	.28[d]	–0.15
Stroop Color-Word Test[e]				
Word	50.72 (14.07)	56.94 (13.87)	.004[d]	–0.45
Color	38.78 (10.00)	43.28 (10.14)	<.001[d]	–0.45
Color-Word	21.33 (8.13)	25.11 (8.77)	<.001[d]	–0.45
Interference	17.44 (6.40)	18.17 (3.99)	.60[d]	–0.14
DS[f]				
Total	25.42 (6.53)	24.11 (6.70)	.32[b]	0.20

[a]ADS: ADHD diagnostic system.

[b]*P* value derived from the Wilcoxon signed-rank test.

[c]RT: response time.

[d]*P* value derived from a *t* test.

[e]In the case of the Stroop Color Word Test, statistical analysis excluded a participant who refused to undergo the test.

[f]DS: digit span.

Stroop Color and Word Test

Scores for the Word, Color, and Color-Word tasks, as well as the Interference score, were evaluated. One participant who did not complete the Stroop task on the postintervention visit was excluded from analysis. Scores for Word (*P*=.004; effect size *d*=-0.45, 95% CI 2.24-10.21), Color (*P*<.001; effect size *d*=-0.45, 95% CI 2.52-6.48), and Color-Word (*P*<.001; effect size *d*=-0.45, 95% CI 1.87-5.68) exhibited significant improvements. The Interference score did not show a statistically significant change (Table 3).

DS

In the WISC-DS, total scores for DS Forward, Backward, and Sequencing were analyzed. No statistically significant changes were noted in the DS tasks (Table 3).

Medication Status Subgroup Analysis

As an exploratory analysis, we compared changes in outcome measures between participants taking medication (n=10) and those not taking medication (n=9). Independent *t* tests (or Mann-Whitney U tests, as appropriate) revealed no statistically significant differences in treatment effects between the medicated and unmedicated groups across all outcome variables. These findings suggest that the observed benefits of the digital intervention were independent of concurrent pharmacological treatment.

Satisfaction and Acceptance

Satisfaction and Acceptance Reported by Children

The mean compliance score of the study participants during the intervention, calculated as the proportion of completed sessions to the total number of instructed sessions, was 81.8% (SD 21.8). For this calculation, only fully completed sessions were counted as compliant. For example, if a participant was instructed to complete 5 sessions per day, 5 days a week, for 4 weeks (totaling 100 sessions) and completed 80 full sessions, their compliance rate was calculated as 80%. Partial completions were not counted in the compliance score.

Participants completed an 11-item questionnaire assessing the digital intervention program's difficulty, learnability, enjoyment, likability, and desire for more play at the postintervention visit. The overall average rating was 4.3 on a 5-point Likert scale (1=strongly disagree, 2=disagree, 3=neutral, 4=agree, and

XSL•FO
RenderX

5=strongly agree). The mean ratings for task difficulty and learnability were 3.7 (SD 0.9) and 4.7 (SD 0.6), respectively. Enjoyment had a mean rating of 3.8 (SD 1.4), while the average likability for narrative structures, rewards, and sound effects was 4.1 (SD 0.9). The desire for more play received a mean rating of 3.9 (SD 1.2). Refer to Table 4 for additional results.

Table 4. Children's ratings of satisfaction and acceptance with the digital intervention.

Dimension and question	Mean (SD)
Difficulty	
Found the game easy to play.	3.7 (0.9)
Learnability	
Understood the gameplay mechanics.	4.7 (0.6)
Enjoyment	
Enjoyed playing the game.	3.8 (1.4)
Likability	
Liked the appearance of the characters.	4.1 (1.2)
Desired to collect the characters in the game.	4.5 (0.6)
Liked the appearance of the space vehicles.	4.0 (1.3)
Desired to collect the space vehicles in the game.	3.8 (1.3)
Appreciated the game's storyline	3.9 (1.1)
Appreciated the rewards obtained from playing the game.	4.7 (0.6)
Enjoyed the sound effects in the game.	3.7 (1.2)
Continuity	
Look forward to continuing the game.	3.9 (1.2)

Satisfaction and Acceptance Reported by Caregivers

At the postintervention visit, caregivers completed a survey about their intervention experience. The survey included four items: (1) the extent to which their children enjoyed playing StarRuckus, (2) the presence of perceived risks to their children, (3) their willingness to continue using the program, and (4) the program's ease of use. On a 5-point Likert scale (1=strongly disagree, 2=disagree, 3=neutral, 4=agree, and 5=strongly agree), the mean rating for enjoyment was 4.0 (SD 1.2), for the absence of perceived risks was 4.4 (SD 0.8), for willingness to continue was 3.6 (SD 0.8), and for ease of use was 4.5 (SD 0.6). The overall mean rating across the 4 items was 4.1 (SD 0.9). In addition, caregivers reported whether children experienced discomfort while using StarRuckus. Reported issues included arm pain (n=2), emotional reactions (n=2), and eyestrain (n=1) as intervention-related discomforts. Children who experienced arm pain reported that the tablet felt heavy. Regarding emotional reactions, 1 child felt frustrated with increasing task difficulty, and another became upset due to unsatisfactory outcomes despite considerable effort in the game. The child with eyestrain reported occasional glare or continuous blinking during app usage.

Discussion

Principal Findings

This study aimed to examine the feasibility and acceptability of the digital intervention StarRuckus, which targets the reduction of ADHD symptoms. Data on symptom and behavior scales, neurocognitive indicators, and user experience feedback were collected. The results of K-ADHD-RS and K-CPRS parental reports demonstrated significant changes in total scores and some subscales compared with baseline, with a moderate effect size indicating symptom improvement and associated behavioral or functional difficulties in children with ADHD.

Although the study lacked a control group, the observed improvements in behavioral symptoms are consistent with findings from previous controlled trials of digital cognitive interventions, which reported similar benefits beyond placebo or practice effects [21,23,52]. These precedents support the interpretation that the improvements observed in this study may reflect genuine intervention-related changes rather than nonspecific factors alone.

Although most results from neurocognitive tasks (ADS visual, ADS-A, and Stroop test) showed no significant changes, 3 subtasks of the Stroop test exhibited a significant increase in performance. Also, the commission error of ADS-A showed a marginally significant tendency to decrease. Furthermore, safety, user convenience, and satisfaction were confirmed by questionnaires administered to parents and children during the postintervention period, with no major adverse events leading to dropout. These findings provide preliminary evidence that gamified digital interventions could be an effective and safe treatment option for children with ADHD, regardless of medication status.

In the context of cognitive training, the transfer effect, a phenomenon where the outcomes of one learning event influence another, is a principal issue [53]. These effects appear within

XSL•FO
RenderX

the realm of cognitive functions, or they may extend to behavioral aspects or symptoms. The current results provide preliminary evidence of transfer effects on behavior or symptoms in ADHD, in contrast to neurocognitive outcomes. Concerning the subjective and behavioral outcomes based on parental reports, our findings support a reduction in both inattention and hyperactivity-impulsivity symptoms of ADHD. These results are consistent with previous studies. Such studies have shown that digitalized cognitive training, including game-based therapeutics, can reduce the behavioral symptoms of ADHD [52,54,55]. Moreover, improvements were noted in the Conduct Problem I subscale of K-CPRS following the use of StarRuckus. Since we used K-CPRS, which differs in item composition and subscale names compared with the original version, interpreting these results requires caution. The Conduct Problem I subscale of K-CPRS, concerning oppositional behaviors and bad habits [40], involves argumentative and defiant behaviors, fidgeting or squirming, emotional dysregulation, and sleep problems. These symptoms are prevalent in ADHD and frequently co-occur with oppositional defiant disorder [3,56]. Indeed, exploratory findings from previous studies have suggested that computer-assisted cognitive training may yield symptomatic improvements in oppositional defiant disorder or conduct disorder itself [57,58]. Other evidence indicates that the efficacy of such training in alleviating ADHD symptoms is not substantially influenced by the presence of comorbid conditions, such as disruptive behavior disorders [59]. However, since the information was gathered by nonblinded raters and there was no control group, we cannot eliminate the possibility that parental overestimation may have influenced the results. A randomized controlled trial conducted by a blinded rater is needed for further investigation.

The current results provide insufficient evidence of transfer effects within neurocognitive functions. Based on the literature on digitalized cognitive training that targets executive functioning in children with ADHD, these findings can be attributed to several factors. First, the results of this study may have been influenced by the low ADS scores at baseline, as suggested by the exploratory analysis, which showed that many participants' ADS performances were below the clinical threshold. Among the 8 indicators yielded by ADS, 13 (68%) participants out of 19 had none or only 1 deficit. These findings may partly relate to the heterogeneity in the cognitive profiles of ADHD [60,61]. Therefore, considering the observed improvements in subjective and behavioral outcomes, it is essential to use multiple cognitive outcomes to identify the underlying cognitive functioning in heterogeneous ADHD participants. Similarly, a previous study demonstrating the efficacy of ADHD digital interventions included children whose Test of Variables of Attention and Attention Performance Index scores, targeted by the intervention, met the inclusion criteria at baseline and found significant improvement in Attention Performance Index scores after training [21]. Second, it is possible that our cognitive training did not result in transfer effects within cognitive abilities such as auditory working memory and interference control. There is currently no consensus on the transfer effects of multitasking training. Finally, due to the small sample size of this study, there may not have been sufficient power to yield significant results. Therefore, future research involving measures targeting various domains of neurocognitive functions, brain signals, brain imaging, or mediating research could further elucidate the mechanisms of transfer in multitasking training. In addition, future studies with larger sample sizes may be necessary.

Although the current results provide preliminary evidence of transfer effects on behavior or symptoms in ADHD, they offer insufficient evidence for transfer effects within neurocognitive functions. Nevertheless, previous research suggests that the multitasking working memory training mechanism in this study likely involves updating and dual processing, which are central executive components of working memory considered vital for behavioral transfer in ADHD. Furthermore, a marginally significant reduction in ADS commission errors and an increase in 3 Stroop Test subtasks may suggest enhanced inhibitory control and processing speed, indicating potential cognitive transfer effects. These gains in inhibitory control, a primary deficit in children with ADHD [62], align with findings from previous studies [20,37]. These findings are further supported by overlapping neural systems, which may explain how working memory training leads to improvements in inhibition [63,64]. In addition, although the corpus of research is limited, several studies demonstrate that working memory training enhances processing speed in ADHD [65,66]. In this study, relatively simple cognitive tasks were used, and test-retest intervals of several weeks were applied; thus, the influence of practice effects may have been reduced [67]. However, the 2 factors previously mentioned do not eliminate practice effects. Therefore, further research with control groups is essential to verify hypotheses related to inhibitory control and processing speed.

While the small sample size limited statistical power, exploratory subgroup analyses indicated no significant differences in treatment outcomes between medicated and unmedicated participants. This finding provides preliminary support for the feasibility of using DTx both as a standalone and as an adjunctive intervention. Nonetheless, adequately powered future studies with stratified randomization are needed to more conclusively evaluate differential effects by medication status.

To promote high treatment adherence among children with ADHD, StarRuckus was developed as a gamified digital therapy, incorporating elements designed to foster internal motivation and maintain engagement in therapeutic activities. Notably, children rated the affinity of game components (eg, characters, space vehicles, and storyline) highly in terms of satisfaction and acceptance. In addition, there was a strong correlation between the children's enjoyment and their desire for continued use (Pearson r=0.74; P<.001). This finding suggests that children who enjoyed the program for 4 weeks were more inclined to continue playing. Consequently, the game format is likely to encourage sustained engagement with the treatment.

StarRuckus was designed to maximize the advantages of gaming while minimizing its disadvantages, such as addictiveness and violence. It incorporates a playtime restriction that limits daily usage to 5 sessions per day. In addition, the story and audiovisual components are free from violent content. Notably,

XSL•FO
RenderX

the item indicating the absence of risk factors in the program received a high average score (mean 4.4, SD 1.2 on a 5-point scale). Furthermore, digital programs, through their convenient usability, minimize the effort required for parental guidance. In the case of StarRuckus, parents reported that the program was easy to operate (mean 4.5, SD 0.6 on a 5-point scale), while children indicated a strong understanding of the task rules (mean 4.7, SD 0.6 on a 5-point scale).

The caregivers' willingness to continue using the program received relatively low scores (mean 3.6, SD 0.8 on a 5-point scale). A similar pattern emerged in a preceding study of a digital intervention for children with ADHD [35], where 80% of parents found the program valuable, yet 40% were hesitant to continue the intervention (approximately 37% of parents responded "neutral" or "disagree" when asked if they wanted their child to continue using the program in this study). Researchers hypothesized that this reluctance might stem from the program's perceived difficulty. Our study conducted a correlation analysis between the caregivers' desire to continue and the children's reported enjoyment, revealing a negative correlation of –0.29, which was not statistically significant (Pearson r=–0.29; P=.22). To better understand why caregivers of children with ADHD report high satisfaction yet exhibit reluctance to continue use, future research should include structured interviews to explore their motivations and concerns more deeply.

Among the 7 of 19 (37%) caregivers who were "neutral" or "disagree" about continuing the use of the program, only 1 reported discomfort (eyestrain) during participation. This suggests that discomfort did not notably influence their decision to continue using the program. In addition, no participant dropped out of the study due to discomfort.

Limitations

This study has several important limitations that should be considered when interpreting our findings. First, the study had a small sample size. Second, and perhaps most critically, we had no control group, which makes it difficult to definitively attribute the observed improvements to our intervention rather than to natural progression, placebo effects, practice effects, or other external factors. The unblinded nature of our assessments further introduces the possibility of subjective bias, particularly in parent-reported outcomes. The improvements observed in behavioral measures and cognitive tasks could be influenced by practice effects or heightened parental attention during the intervention period. To solidify our understanding of the intervention's efficacy, additional randomized controlled trials with appropriate control conditions (eg, waitlist, placebo, or active comparator) and larger participant cohorts are essential, with outcomes assessed by blinded raters. Finally, the heterogeneity of ADHD severity was likely due to participants' differing times of diagnosis. Therefore, future studies need to use cohorts with controlled heterogeneity, enabling researchers to determine which ADHD subpopulations are most likely to benefit from our intervention.

Conclusions

This is the first study to demonstrate the potential of a gamified multitasking training program that combines the n-back task and the driving task to improve core symptoms in Korean children with ADHD. In this preliminary investigation, we assessed the feasibility and appeal of a gamified DTx for children with ADHD, considering both medicated and nonmedicated groups. In both cohorts, the 4-week intervention was associated with reductions in core ADHD symptoms, including inattention and hyperactivity-impulsivity, and received favorable user feedback. StarRuckus, the digital program that is not bound by time or location and is readily accessible at home, holds promising potential to serve as both a substitute for and a supplement to conventional ADHD treatment strategies. While the preliminary results we observed provide an initial foundation for further research into this approach, our study design, which includes a small sample size without a control group, limits conclusions about causality and generalizability. Future studies should implement robust randomized controlled trials with appropriate control conditions, larger and more diverse samples, blinded assessments, and longer follow-up periods to further validate these preliminary findings and reveal the underlying mechanisms of the intervention's effectiveness.

Acknowledgments

This work was supported by the Korea Medical Device Development Fund, funded by the Korean government (Ministry of Science and ICT, Ministry of Trade, Industry and Energy, Ministry of Health and Welfare, and Ministry of Food and Drug Safety; project no RS-2023-00255328).

Authors' Contributions

MJ contributed to the study concept, StarRuckus program design, study design, acquisition and interpretation of data, statistical analyses, drafting the manuscript, and revising the manuscript. JW contributed to the study design, acquisition and interpretation of data, drafting the manuscript, and revising the manuscript. SK contributed to the StarRuckus program design, acquisition of data, drafting the manuscript, and revising the manuscript. YBL contributed to reviewing and editing the manuscript. YK contributed to reviewing and revising the manuscript. DK contributed to reviewing and revising the manuscript. JSM contributed to reviewing the manuscript, and BNK contributed to the StarRuckus program design, study design, and reviewing and editing the manuscript.

XSL•FO
RenderX

Conflicts of Interest

MJ, JW, SK, YK, and DK are employees of eMotiv Co. Ltd. JSM is the CEO of eMotiv Co. Ltd. All other authors declare no conflicts of interest related to the content of this article.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1109 KB - humanfactors_v12i1e70479_app1.pdf]

References

1. Jeste D, Lieberman E, Fassler D. Diagnostic and Statistical Manual of *DSM-5* TM. Washington, DC: American Psychiatric Association; 2013.
2. Rohde LA, Buitelaar JK, Gerlach M, Faraone SV. The world federation of ADHD guide. ADHD World Federation. URL: https://www.russellbarkley.org/factsheets/TheWorldFederationOfAdhdGuide.pdf [accessed 2025-09-17]
3. Seo JC, Jon DI, Shim SH, Sung HM, Woo YS, Hong J, et al. Prevalence and comorbidities of attention deficit hyperactivity disorder among adults and children/adolescents in Korea. Clin Psychopharmacol Neurosci 2022;20(1):126-134 [FREE Full text] [doi: 10.9758/cpn.2022.20.1.126] [Medline: 35078955]
4. Drechsler R, Brem S, Brandeis D, Grünblatt E, Berger G, Walitza S. ADHD: current concepts and treatments in children and adolescents. Neuropediatrics 2020;51(5):315-335 [FREE Full text] [doi: 10.1055/s-0040-1701658] [Medline: 32559806]
5. Jaeeun L. ADHD cases doubled in last four years: data. The Korea Herald. 2023. URL: https://www.koreaherald.com/article/3072933 [accessed 2025-09-17]
6. Polanczyk GV, Willcutt EG, Salum GA, Kieling C, Rohde LA. ADHD prevalence estimates across three decades: an updated systematic review and meta-regression analysis. Int J Epidemiol 2014;43(2):434-442 [FREE Full text] [doi: 10.1093/ije/dyt261] [Medline: 24464188]
7. Sasser T, Schoenfelder EN, Stein MA. Targeting functional impairments in the treatment of children and adolescents with ADHD. CNS Drugs 2017;31(2):97-107. [doi: 10.1007/s40263-016-0400-1] [Medline: 27943133]
8. Caye A, Spadini AV, Karam RG, Grevet EH, Rovaris DL, Bau CHD, et al. Predictors of persistence of ADHD into adulthood: a systematic review of the literature and meta-analysis. Eur Child Adolesc Psychiatry 2016;25(11):1151-1159. [doi: 10.1007/s00787-016-0831-8] [Medline: 27021056]
9. Klein RG, Mannuzza S, Olazagasti MAR, Roizen E, Hutchison JA, Lashua EC, et al. Clinical and functional outcome of childhood attention-deficit/hyperactivity disorder 33 years later. Arch Gen Psychiatry 2012;69(12):1295-1303 [FREE Full text] [doi: 10.1001/archgenpsychiatry.2012.271] [Medline: 23070149]
10. Cortese S. Pharmacologic treatment of attention deficit-hyperactivity disorder. N Engl J Med 2020;383(11):1050-1056. [doi: 10.1056/NEJMra1917069] [Medline: 32905677]
11. Caye A, Swanson JM, Coghill D, Rohde LA. Treatment strategies for ADHD: an evidence-based guide to select optimal treatment. Mol Psychiatry 2019;24(3):390-408. [doi: 10.1038/s41380-018-0116-3] [Medline: 29955166]
12. Pelham WE, Gnagy EM, Greiner AR, Hoza B, Hinshaw SP, Swanson JM, et al. Behavioral versus behavioral and pharmacological treatment in ADHD children attending a summer treatment program. J Abnorm Child Psychol 2000;28(6):507-525. [doi: 10.1023/a:1005127030251] [Medline: 11104314]
13. Digital Therapeutics Alliance. 2017. URL: https://dtxalliance.org/wp-content/uploads/2023/06/DTA_FS_New-DTx-Definition.pdf [accessed 2025-09-17]
14. Wong KP, Qin J, Xie YJ, Zhang B. Effectiveness of technology-based interventions for school-age children with attention-deficit/hyperactivity disorder: systematic review and meta-analysis of randomized controlled trials. JMIR Ment Health 2023;10:e51459 [FREE Full text] [doi: 10.2196/51459] [Medline: 37988139]
15. Medina R, Bouhaben J, de Ramón I, Cuesta P, Antón-Toro L, Pacios J, et al. Electrophysiological brain changes associated with cognitive improvement in a pediatric attention deficit hyperactivity disorder digital artificial intelligence-driven intervention: randomized controlled trial. J Med Internet Res 2021;23(11):e25466 [FREE Full text] [doi: 10.2196/25466] [Medline: 34842533]
16. Sapanel Y, Tadeo X, Brenna CTA, Remus A, Koerber F, Cloutier LM, et al. Economic evaluation associated with clinical-grade mobile app-based digital therapeutic interventions: systematic review. J Med Internet Res 2023;25:e47094 [FREE Full text] [doi: 10.2196/47094] [Medline: 37526973]
17. Lumsden J, Edwards EA, Lawrence NS, Coyle D, Munafò MR. Gamification of cognitive assessment and cognitive training: a systematic review of applications and efficacy. JMIR Serious Games 2016;4(2):e11 [FREE Full text] [doi: 10.2196/games.5888] [Medline: 27421244]
18. Boendermaker WJ, Gladwin TE, Peeters M, Prins PJM, Wiers RW. Training working memory in adolescents using serious game elements: pilot randomized controlled trial. JMIR Serious Games 2018;6(2):e10 [FREE Full text] [doi: 10.2196/games.8364] [Medline: 29792294]
19. Antrop I, Roeyers H, Van Oost P, Buysse A. Stimulation seeking and hyperactivity in children with ADHD. J. Child Psychol. Psychiat 2000;41(2):225-231. [doi: 10.1017/s0021963099005302]

XSL•FO
RenderX

20. Dotare M, Bader M, Mesrobian SK, Asai Y, Villa AEP, Lintas A. Attention networks in ADHD adults after working memory training with a dual -back task. Brain Sci 2020;10(10):715 [FREE Full text] [doi: 10.3390/brainsci10100715] [Medline: 33050115]
21. Kollins SH, DeLoss DJ, Cañadas E, Lutz J, Findling RL, Keefe RSE, et al. A novel digital intervention for actively reducing severity of paediatric ADHD (STARS-ADHD): a randomised controlled trial. Lancet Digit Health 2020;2(4):e168-e178 [FREE Full text] [doi: 10.1016/S2589-7500(20)30017-0] [Medline: 33334505]
22. Sun TH, Yeom JW, Choi K, Kim J, Lee H, Kim H, et al. Potential effectiveness of digital therapeutics specialized in executive functions as adjunctive treatment for clinical symptoms of attention-deficit/hyperactivity disorder: a feasibility study. Front Psychiatry 2023;14:1169030 [FREE Full text] [doi: 10.3389/fpsyt.2023.1169030] [Medline: 37547212]
23. Kim SC, Lee H, Lee HS, Kim G, Song JH. Adjuvant therapy for attention in children with ADHD using game-type digital therapy. Int J Environ Res Public Health 2022;19(22):14982 [FREE Full text] [doi: 10.3390/ijerph192214982] [Medline: 36429699]
24. Anguera JA, Boccanfuso J, Rintoul JL, Al-Hashimi O, Faraji F, Janowich J, et al. Video game training enhances cognitive control in older adults. Nature 2013;501(7465):97-101 [FREE Full text] [doi: 10.1038/nature12486] [Medline: 24005416]
25. Mishra J, de Villers-Sidani E, Merzenich M, Gazzaley A. Adaptive training diminishes distractibility in aging across species. Neuron 2014;84(5):1091-1103. [doi: 10.1016/j.neuron.2014.10.034] [Medline: 25467987]
26. Kollins SH, Childress A, Heusser AC, Lutz J. Effectiveness of a digital therapeutic as adjunct to treatment with medication in pediatric ADHD. NPJ Digit Med 2021;4(1):58 [FREE Full text] [doi: 10.1038/s41746-021-00429-0] [Medline: 33772095]
27. Bopp KL, Verhaeghen P. Aging and n-back performance: a meta-analysis. J Gerontol B Psychol Sci Soc Sci 2020;75(2):229-240. [doi: 10.1093/geronb/gby024] [Medline: 31943115]
28. Krivitzky LS, Roebuck-Spencer TM, Roth RM, Blackstone K, Johnson CP, Gioia G. Functional magnetic resonance imaging of working memory and response inhibition in children with mild traumatic brain injury. J Int Neuropsychol Soc 2011;17(6):1143-1152. [doi: 10.1017/S1355617711001226] [Medline: 22014100]
29. Lilienthal L, Tamez E, Shelton JT, Myerson J, Hale S. Dual n-back training increases the capacity of the focus of attention. Psychon Bull Rev 2013;20(1):135-141. [doi: 10.3758/s13423-012-0335-6] [Medline: 23184506]
30. Depue BE, Burgess GC, Bidwell LC, Willcutt EG, Banich MT. Behavioral performance predicts grey matter reductions in the right inferior frontal gyrus in young adults with combined type ADHD. Psychiatry Res 2010;182(3):231-237 [FREE Full text] [doi: 10.1016/j.pscychresns.2010.01.012] [Medline: 20493669]
31. Gehricke JG, Kruggel F, Thampipop T, Alejo SD, Tatos E, Fallon J, et al. The brain anatomy of attention-deficit/hyperactivity disorder in young adults - a magnetic resonance imaging study. PLoS One 2017;12(4):e0175433 [FREE Full text] [doi: 10.1371/journal.pone.0175433] [Medline: 28406942]
32. Schweitzer JB, Faber TL, Grafton ST, Tune LE, Hoffman JM, Kilts CD. Alterations in the functional anatomy of working memory in adult attention deficit hyperactivity disorder. Am J Psychiatry 2000;157(2):278-280. [doi: 10.1176/appi.ajp.157.2.278] [Medline: 10671402]
33. Valera EM, Brown A, Biederman J, Faraone SV, Makris N, Monuteaux MC, et al. Sex differences in the functional neuroanatomy of working memory in adults with ADHD. Am J Psychiatry 2010;167(1):86-94 [FREE Full text] [doi: 10.1176/appi.ajp.2009.09020249] [Medline: 19884224]
34. Gallen CL, Anguera JA, Gerdes MR, Simon AJ, Cañadas E, Marco EJ. Enhancing neural markers of attention in children with ADHD using a digital therapeutic. PLoS One 2021;16(12):e0261981. [doi: 10.1371/journal.pone.0261981] [Medline: 34972140]
35. Davis NO, Bower J, Kollins SH. Proof-of-concept study of an at-home, engaging, digital intervention for pediatric ADHD. PLoS One 2018;13(1):e0189749. [doi: 10.1371/journal.pone.0189749] [Medline: 29324745]
36. Kirchner WK. Age differences in short-term retention of rapidly changing information. J Exp Psychol 1958;55(4):352-358 [FREE Full text]
37. Jones MR, Katz B, Buschkuehl M, Jaeggi SM, Shah P. Exploring n-back cognitive training for children with ADHD. J Atten Disord 2020;24(5):704-719 [FREE Full text] [doi: 10.1177/1087054718779230] [Medline: 29877128]
38. Pahor A, Seitz AR, Jaeggi SM. Near transfer to an unrelated n-back task mediates the effect of n-back working memory training on matrix reasoning. Nat Hum Behav 2022;6(9):1243-1256. [doi: 10.1038/s41562-022-01384-w] [Medline: 35726054]
39. DuPaul GJ. Parent and teacher ratings of ADHD symptoms: psychometric properties in a community-based sample. J Clin Child Psychol 1991;20(3):245-253. [doi: 10.1207/s15374424jccp2003_3]
40. So YK, Noh JS, Kim Y, Koh SG, Koh YJ. The reliability and validity of Korean parent and teacher ADHD rating scale. J Korean Neuropsychiatr Assoc 2002;41(2):283-289 [FREE Full text]
41. Kim YS, So YK, Noh JS, Choi NK, Kim SJ, Koh YJ. Normative data on the Korean ADHD rating scales (K-ARS) for parents and teacher. J Korean Neuropsychiatr Assoc 2003;42(3):352-359 [FREE Full text]
42. Conners CK. Symptom patterns in hyperkinetic, neurotic, and normal children. Child Development 1970;41(3):667-682 [FREE Full text] [doi: 10.2307/1127215]

XSL•FO
RenderX

43. Conners CK, Sitarenios G, Parker JDA, Epstein JN. The revised conners' parent rating scale (CPRS-R): factor structure, reliability, and criterion validity. J Abnorm Child Psychol 1998;26(4):257-268. [doi: 10.1023/a:1022602400621] [Medline: 9700518]
44. Park EH, So YK, Kim YS, Choi NK, Kim SJ, Noh JS, et al. The reliability and validity of Korean conners parent and teacher rating scale. J Korean Acad Child Adolesc Psychiatry 2003;14(2):183-196 [FREE Full text]
45. Shin MS, Cho S, Chun SY, Hong K. A study of the development and standardization of ADHD diagnostic system. J Korean Acad Child Adolesc Psychiatry 2000;11(1):91-99 [FREE Full text]
46. Mackworth JF, Taylor MM. The d'measure of signal detectability in vigilancelike situations. Can J Psychol 1963;17(3):302-325 [FREE Full text]
47. Epstein JN, Erkanli A, Conners CK, Klaric J, Costello JE, Angold A. Relations between continuous performance test performance measures and ADHD behaviors. J Abnorm Child Psychol 2003;31(5):543-554. [doi: 10.1023/a:1025405216339] [Medline: 14561061]
48. Shin M, Park M. STROOP: Color and Word Test Children's Version for Ages 5-14. South Korea: Hakjisa; 2007.
49. Golden C. The Stroop Color and Word Test: A Manual for Clinical Experimental Uses. Wood Dale, Illinois: Stoelting company; 1978.
50. Wechsler D. Wechsler Intelligence Scale for Children. London: Pearson Publishing; 2014.
51. Kwak K, Jang S. The Korean-Wechsler Intelligence Scale for Children. South Korea: Hakjisa; 2019.
52. Oh S, Choi J, Han DH, Kim E. Effects of game-based digital therapeutics on attention deficit hyperactivity disorder in children and adolescents as assessed by parents or teachers: a systematic review and meta-analysis. Eur Child Adolesc Psychiatry 2024;33(2):481-493. [doi: 10.1007/s00787-023-02174-z] [Medline: 36862162]
53. Morrison AB, Chein JM. Does working memory training work? The promise and challenges of enhancing cognition by training working memory. Psychon Bull Rev 2011;18(1):46-60. [doi: 10.3758/s13423-010-0034-0] [Medline: 21327348]
54. Azami S, Moghadas A, Sohrabi-Esmrood F, Nazifi M, Mirmohamad M, Hemmati F, et al. A pilot randomized controlled trial comparing computer-assisted cognitive rehabilitation, stimulant medication, and an active control in the treatment of ADHD. Child Adolesc Ment Health 2016;21(4):217-224. [doi: 10.1111/camh.12157] [Medline: 32680341]
55. Beck SJ, Hanson CA, Puffenberger SS, Benninger KL, Benninger WB. A controlled trial of working memory training for children and adolescents with ADHD. J Clin Child Adolesc Psychol 2010;39(6):825-836. [doi: 10.1080/15374416.2010.517162] [Medline: 21058129]
56. Larson K, Russ SA, Kahn RS, Halfon N. Patterns of comorbidity, functioning, and service use for US children with ADHD, 2007. Pediatrics 2011;127(3):462-470 [FREE Full text] [doi: 10.1542/peds.2010-0165] [Medline: 21300675]
57. Dovis S, Van der Oord S, Wiers RW, Prins PJM. Improving executive functioning in children with ADHD: training multiple executive functions within the context of a computer game. a randomized double-blind placebo controlled trial. PLoS One 2015;10(4):e0121651 [FREE Full text] [doi: 10.1371/journal.pone.0121651] [Medline: 25844638]
58. van der Oord S, Ponsioen AJGB, Geurts HM, Ten Brink EL, Prins PJM. A pilot study of the efficacy of a computerized executive functioning remediation training with game elements for children with ADHD in an outpatient setting: outcome on parent- and teacher-rated executive functioning and ADHD behavior. J Atten Disord 2014;18(8):699-712. [doi: 10.1177/1087054712453167] [Medline: 22879577]
59. Bigorra A, Garolera M, Guijarro S, Hervas A. Impact of working memory training on hot executive functions (decision-making and theory of mind) in children with ADHD: a randomized controlled trial. Neuropsychiatry 2016;6(5):251-263. [doi: 10.4172/neuropsychiatry.1000147]
60. Kofler MJ, Irwin LN, Soto EF, Groves NB, Harmon SL, Sarver DE. Executive functioning heterogeneity in pediatric ADHD. J Abnorm Child Psychol 2019;47(2):273-286 [FREE Full text] [doi: 10.1007/s10802-018-0438-2] [Medline: 29705926]
61. Nigg JT, Willcutt EG, Doyle AE, Sonuga-Barke EJ. Causal heterogeneity in attention-deficit/hyperactivity disorder: do we need neuropsychologically impaired subtypes? Biol Psychiatry 2005;57(11):1224-1230. [doi: 10.1016/j.biopsych.2004.08.025] [Medline: 15949992]
62. Pievsky M, McGrath R. The neurocognitive profile of attention-deficit/hyperactivity disorder: a review of meta-analyses. Arch Clin Neuropsychol 2018;33(2):143-157. [doi: 10.1093/arclin/acx055] [Medline: 29106438]
63. Salmi J, Nyberg L, Laine M. Working memory training mostly engages general-purpose large-scale networks for learning. Neurosci Biobehav Rev 2018;93:108-122. [doi: 10.1016/j.neubiorev.2018.03.019] [Medline: 29574197]
64. Klingberg T, Fernell E, Olesen PJ, Johnson M, Gustafsson P, Dahlström K, et al. Computerized training of working memory in children with ADHD--a randomized, controlled trial. J Am Acad Child Adolesc Psychiatry 2005;44(2):177-186 [FREE Full text] [doi: 10.1097/00004583-200502000-00010] [Medline: 15689731]
65. Egeland J, Aarlien AK, Saunes B. Few effects of far transfer of working memory training in ADHD: a randomized controlled trial. PLoS One 2013;8(10):e75660 [FREE Full text] [doi: 10.1371/journal.pone.0075660] [Medline: 24124503]
66. Cunha F, Campos S, Simões-Silva V, Brugada-Ramentol V, Sá-Moura B, Jalali H, et al. The effect of a virtual reality based intervention on processing speed and working memory in individuals with ADHD—A pilot-study. Front. Virtual Real 2023;4. [doi: 10.3389/frvir.2023.1108060]

XSL•FO
RenderX

67. Scharfen J, Blum D, Holling H. Response time reduction due to retesting in mental speed tests: a meta-analysis. J Intell 2018;6(1):6 [FREE Full text] [doi: 10.3390/jintelligence6010006] [Medline: 31162433]

Abbreviations

ADHD: attention-deficit/hyperactivity disorder
ADS: ADHD diagnostic system
ADS-A: ADS-auditory
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile HEalth Applications and onLine TeleHealth
DS: digit span
DTx: digital therapeutics
FDA: Food and Drug Administration
K-ADHD-RS: Korean ADHD Rating Scale
K-CPRS: Korean Conners' Parent Rating Scale
RT: response time

Edited by A Kushniruk; submitted 27.12.24; peer-reviewed by JH Song; comments to author 28.01.25; revised version received 25.03.25; accepted 12.08.25; published 28.10.25.

Please cite as:
Jung M, Woo J, Kim SJ, Lim YB, Kim Y, Kang D, Min JS, Kim BN
Efficacy and Compliance of a Working Memory Multitasking Task Mobile Intervention for Children With Attention-Deficit/Hyperactivity Disorder: Single-Arm, Pre-Post Pilot Study
JMIR Hum Factors 2025;12:e70479
URL: https://humanfactors.jmir.org/2025/1/e70479
doi:10.2196/70479
PMID:

©Minyoung Jung, Jimin Woo, Sung Jin Kim, You Bin Lim, Younglae Kim, Dongwon Kang, Jung-sang Min, Bung-Nyun Kim. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 28.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Understanding the Differences Between Online and Offline Mental Health Help Seekers: Cross-Sectional Comparative Study

Mohamed Adwi[1], MB BCh; Bahaa Mahmoud[2], MSc; Noha Amer[3], MD; Roa Gamal Alamrawy[4], MSc; Ismail Sadek[2], MD; Mohamed Elsheikh[2], MD

[1]Shezlong Inc, Giza, Egypt

[2]Department of Psychiatry, Al-Azhar University, Cairo, Egypt

[3]Department of Psychiatry, Mansoura University, Mansoura, Egypt

[4]Ministry of Health and Population, Mamoura Psychiatric Hospital, General Secretariat of Mental Health and Addiction Treatment, Alexandria, Egypt

Corresponding Author:
Mohamed Adwi, MB BCh
Shezlong Inc
Sheikh Zayed, Giza, Egypt
Giza, 41516
Egypt
Phone: 20 1112910 ext 671
Email: mohamed.adwi.206@azhar.edu.eg

Abstract

Background: Telepsychiatry has gained considerable attention, particularly during the COVID-19 pandemic. Although various factors influence the choice between online and offline modalities, differences among populations remain underexplored.

Objective: This study aims to compare adults seeking mental health support online and offline in private clinics.

Methods: In this cross-sectional study, we assessed differences in sociodemographic factors, internet accessibility and usability, previous help-seeking history, personality traits assessed using the Arabic Big Five Personality Inventory, and levels of self-stigma measured using the Self-Stigma of Seeking Help Scale.

Results: In total, 259 participants were included (136 online and 123 offline). The online group had a higher proportion of university graduates (P=.02), employed individuals (P<.001), and those with better internet access (P=.03) and higher internet usability (P=.001). The offline group showed higher levels of conscientiousness (P=.003). The primary reasons for choosing online therapy were ease of access and time-saving. Logistic regression identified previous use of online psychiatry as the strongest factor associated with choosing online services (odds ratio [OR] 28.90, 95% CI 11.739-71.165; P<.001). Employment (OR 5.01, 95% CI 1.781-14.080; P=.002), better internet usability (OR 1.69, 95% CI 1.069-2.664; P=.03), and agreeableness (OR 1.16, 95% CI 1.001-1.351; P=.05) were also significant factors. In contrast, previous in-person visits (OR 0.11, 95% CI 0.048-0.269; P<.001), openness (OR 0.85, 95% CI 0.748-0.975; P=.02), and conscientiousness (OR 0.86, 95% CI 0.758-0.971; P=.02) were negatively associated with online preference.

Conclusions: This study highlights key differences between online and offline mental health help seekers, enhances our understanding of treatment modality preferences, and paves the way for future research.

(JMIR Hum Factors 2025;12:e69305) doi:10.2196/69305

KEYWORDS

help-seeking behavior; online therapy; personality traits; self-stigma; telepsychiatry

Introduction

Telepsychiatry refers to the delivery of psychiatric assessment, consultation, and treatment services through telecommunication technologies that support real-time interactive communication between clinicians and patients. It is mostly implemented through videoconferencing, enabling direct face-to-face interactions at a distance. Online therapy, more broadly, encompasses any form of professional mental health support conducted over the internet, including text-, audio-, or video-based communication [1,2]. It has gained significant attention and adoption during the COVID-19 pandemic, which disrupted traditional face-to-face mental health service delivery [3]. Telepsychiatry has demonstrated the ability to perform

XSL•FO
RenderX

various functions, such as monitoring, surveillance, mental health promotion, and treatment programs, as efficiently and effectively as in-person care [4].

Research on help-seeking preferences has revealed that several factors can influence individuals' choices between online-based therapy and offline mental health help-seeking. Mental health literacy, gender, age, and location of residence have been identified as influential factors [5,6]. Online and offline users may differ in educational level, social support, social withdrawal, family attachment, and technological awareness [7]. Some individuals may lean toward face-to-face interactions, driven by personality traits and concerns about efficacy and privacy [8-10]. In contrast, online interventions are favored by many because they offer increased accessibility, overcoming barriers such as stigma, distance, and transportation issues [11-13].

Cross-cultural factors can significantly influence the preference for offline consultation or telepsychiatry; cultural background and identity can affect comfort with technology and acceptance of telepsychiatry [14]. Some patients prefer telepsychiatry, especially ethnic minorities and immigrants, to connect with mental health professionals who have the appropriate language and cultural background, regardless of location [15].

According to Andersen's Behavioral Model of Health Services Use [16], the use of health services is shaped by 3 categories of factors, which can be applied to understand decisions about online mental health services. Predisposing factors, such as age, sex, and education, may influence individuals' attitudes toward telepsychiatry. Enabling factors, including employment status, income, internet accessibility, and the ease of accessing therapists online, can facilitate or constrain the use of digital platforms. Need factors, such as privacy concerns and the preference for saving time, may further motivate individuals to choose online therapy over traditional in-person care. However, most studies comparing online and offline mental health help-seekers to better understand these factors have been conducted in Western populations, with few studies comparing populations from developed and developing countries [16-18]. Studies conducted in the Arab world, including Egypt, have primarily focused on readiness to use digital mental health services, as well as knowledge, attitudes, perceptions, and satisfaction with these services [19-21]. Nevertheless, there remains a gap in the literature addressing the factors that influence individuals' choices between telepsychiatry and traditional in-person therapy in the Arab world [22].

Therefore, this study aims to address this gap by examining differences between online and offline mental health help seekers in Egypt, focusing on socioeconomic factors such as educational attainment, income level, and employment status, along with personality traits. Additionally, this study assessed differences with regard to the self-stigma of seeking help. In this study, we define "online help seekers" as individuals receiving therapy through videoconferencing platform, whereas "offline help seekers" are those attending in-person sessions at outpatient clinics. By undertaking this investigation, we aim to provide valuable insights into the field of telepsychiatry and mental health services.

Methods

Study Design and Participants

This cross-sectional study was conducted using a self-administered survey for both online and offline participants. Inclusion criteria were adult Egyptians aged 18 years or older, residing in Egypt, who sought mental health help either through an online telepsychiatry platform (Shezlong) or offline at private outpatient clinics, without psychotic features. To control for migration as a potential confounding factor, the study was limited to participants from a single country, ensuring that each participant could choose between online and offline therapy. To maintain comparability, only private clinics were selected for the offline group, aligning with the private nature of the online service. The study was guided by the CHERRIES (Checklist for Reporting Results of Internet E-Surveys) guidelines [23].

Sample Size

The required sample size was determined using MedCalc version 18.2.1 (MedCalc Software Ltd). We based our calculation on a previous study that reported a mean agreeableness score of 4.67 (SD 1.17) among individuals preferring online mental health services and 5.13 (SD 1.10) among those preferring offline services [24]. Assuming a significance level α=.05, a power of 80%, and equal group allocation (1:1 ratio), the minimum sample size required was 190 participants. To ensure sufficient power and account for potential data loss, the sample size was rounded up to 200 participants, with 100 in each group.

Recruitment

For online help seekers, a total of 7 therapists expressed interest in inviting their clients to participate. Each therapist sent a link to the online survey at the end of the session. For offline participants, a total of 6 therapists working in private clinics provided a hard copy of the survey following clinic consultations. A consent form was placed on the first page of the online survey, and only participants who provided consent proceeded to the subsequent pages, where they answered the age question to confirm eligibility based on predetermined inclusion criteria. Due to cultural sensitivity toward written consent, initial oral consent was obtained from offline participants, after which those who agreed to participate completed a paper-based survey that began with a written consent statement. Data were collected randomly from January to June 2024. To minimize selection bias, therapists were instructed to approach all eligible clients during the recruitment period.

There are notable sociodemographic differences between public and private sector help-seekers in Egypt. Additionally, there is no equivalent public sector online mental health service with comparable reach or accessibility, as the only existing platform remains in a limited beta phase. Moreover, the private sector accounts for a substantial portion of outpatient mental health care use, even among lower-income individuals, due to perceptions of better quality and accessibility compared with public services. Therefore, we focused exclusively on the private

XSL•FO
RenderX

sector for the comparison between online and offline help-seekers [25].

Data Collection Tools

The participants completed a self-administered questionnaire. The overall instrument demonstrated acceptable internal consistency in our sample (Cronbach α=0.674). The questionnaire consisted of the following sections: demographic and social factors, internet accessibility and usability, previous history of online or offline help seeking, reasons for preferring online therapy, the Arabic Big Five Personality Inventory, and the Self-Stigma of Seeking Help (SSOSH) scale.

Demographic and Socioeconomic Factors

Demographics included age, sex, educational status, last educational attainment, and marital status. Other variables included residence, monthly salary, breadwinner status, and receiving charity. Participants who were employed were asked about employment type (part time or full time), work shift (rotational or fixed), working location, salary range, and work sector.

Internet Accessibility and Usability

Internet accessibility was assessed as yes or no, and internet usability was measured on a Likert scale from 1 to 5, where 1 represents the lowest and 5 the highest.

Previous History of Online or Offline Help Seeking

History of online or offline help seeking was assessed by asking about previous experience with offline help seeking through clinics or online through videoconference, with a yes or no response.

Reasons for Preferring Online Therapy

Online participants were asked to rate the strength of each reason for preferring online therapy using a Likert scale, where 0=no, 1=moderate, and 2=strong. The options were: ease and comfort, saving time, cost savings, feeling of confidentiality, lack of nearby therapists, availability of many therapists to choose from, preferred therapist is available online only, and others (please specify).

The Arabic Big Five Personality Inventory

A 25-item questionnaire was used to assess the 5 major personality traits: neuroticism (N), extraversion (E), agreeableness (A), openness to experience (O), and conscientiousness (C). Sample items include: “I am a sociable person,” “I am an anxious person,” “I like new ideas,” and “I am organized and disciplined.” The reliability of this inventory ranged between 0.70 and 0.83, while validity ranged from 0.49 to 0.86 [26,27]. Respondents rated each item on a 4-point Likert scale: 1=no, 2=some, 3=much, and 4=always. Each trait’s total score ranges from 5 to 20, with higher scores reflecting stronger expression of that trait. In the current sample, internal consistency was assessed using Cronbach α. The Arabic Big Five Personality Inventory demonstrated acceptable overall reliability (α=0.658). Subscale reliability coefficients were as follows: extraversion (α=.790), neuroticism (α=0.757), agreeableness (α=0.909), openness (α=0.839), and conscientiousness (α=0.820).

SSOSH Scale

The SSOSH consists of 10 items on a Likert scale from 1 to 5, where 1=strongly disagree and 5=strongly agree. Items 2, 4, 5, 7, and 9 are reverse scored, and higher scores indicate greater of self-stigma [28]. Sample items include: “I feel inferior if I go to a therapist for help” and “Seeking help from a professional does not threaten my self-confidence.” The translated Arabic version is available on the official website of the Self Stigma Research Collaborative at Iowa State University [29]. In the current sample, Cronbach α for the SSOSH was 0.398, indicating low internal consistency.

Statistical Analysis

The collected data were wrangled, coded, and analyzed using SPSS software (version 25; IBM Corp) and R (R Foundation for Statistical Computing) to generate a correlation heatmap of the Big Five traits and SSOSH scale. Quantitative variables are expressed as mean (SD), whereas categorical variables are expressed as n (%). An independent 2-tailed *t* test was used to estimate differences in continuous variables between the online and offline groups, whereas the chi-square test and Monte Carlo exact probability were used to assess the difference in categorical variables, with column proportions analyzed for significant variables. Pearson correlation was used to assess the relationship between the Big Five traits and the SSOSH scale. A multivariable logistic regression model was conducted using important significant variables, the Big Five traits, and the SSOSH scale to identify the significant factors associated with preference for online consultations. Odds ratios (ORs) and 95% CIs were reported. Statistical significance was set at $P<.05$. Cronbach α values were calculated to evaluate the internal consistency of the psychometric scales used in the current sample.

Ethical Considerations

Ethics approval was obtained from the Ethics Committee of Al-Azhar Faculty of Medicine (registry number Psych._84 Med.Research._Prevalence Depression/Pts.-MS._00000084). All procedures involving human participants were conducted in accordance with the ethical standards of the institutional research committee of Al-Azhar Faculty of Medicine and the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. The process of obtaining informed consent, described above, was implemented for both online and offline participants using electronic and written formats, respectively. Participants’ privacy and confidentiality were safeguarded throughout all study phases, with anonymized data used for analysis and reporting. Online participants were offered a 10%-20% discount for a future online therapy session as a token of appreciation, whereas offline participants received no compensation.

Results

Demographic Characteristics

A total of 304 eligible participants were approached, of whom 259 agreed to participate, yielding a response rate of 85.2%. The study included 259 participants, with 136 in the online group and 123 in the offline group. The majority of participants

XSL•FO
RenderX

were female, representing 80.9% (110/136) in the online group and 73.2 % (90/123) in the offline group. The mean age was similar between groups, with the online group having a mean age of 28.92 (6.35) years and the offline group having a mean age of 29.21 (9.55) years.

Significant differences were observed in the current educational level between the 2 groups (*P*=.003). The online group had a lower percentage of students (22/136, 16.2%) compared to the offline group (35/123, 28.5%) but a higher proportion of graduates (114/136, 83.8%) compared to 68.3% (84/123). When considering the highest level of education achieved, the offline group had a higher proportion of participants with secondary education or less (16/119, 13.4%) compared to the online group (6/137, 4.4%), whereas the online group had more participants with a university degree (109/136, 80.1%) compared to 68.1% (81/119), a statistically significant difference (*P*=.02). There were no significant differences in marital status between the groups, with the majority of participants being single, particularly in the online group (96/136, 70.6%) compared to 61.8% (76/123; Table 1).

Table 1. Demographic characteristics.

Variable	Online group (n=136)	Offline group (n=123)	Test statistic	Effect size	*P* value
Sex			2.2 (1)[a]	0.092[b]	.14
Male, n (%)	26 (19.1)	33 (26.8)			
Female, n (%)	110 (80.9)	90 (73.2)			
Age in years, mean (SD)	28.92 (6.35)	29.21 (9.55)	–0.284 (251)[c]	0.0358[d]	.78
Current education level, n (%)			—[e]	0.205[b]	.003[f]
Not educated[g]	0 (0)	4 (3.3)			
Student[g]	22 (16.2)	35 (28.5)			
Graduate[g]	114 (83.8)	84 (68.3)			
Highest education level achieved (n=256), n (%)			7.6 (2)[a]	0.173[b]	.02
Secondary or less[g]	6 (4.4)	16 (13.4)			
University degree[g]	109 (80.1)	81 (68.1)			
Higher education[h]	21 (15.4)	22 (18.5)			
Marital status, n (%)			3.5 (3)[a]	0.116[b]	.32
Single	96 (70.6)	76 (61.8)			
Married	35 (25.7)	38 (30.9)			
Divorced	5 (3.7)	8 (6.5)			
Widow	0 (0.0)	1 (0.8)			

[a]Chi-square test (degrees of freedom).

[b]Cramer V.

[c]2-tailed *t* test (degrees of freedom).

[d]Cohen *d*.

[e]Not applicable.

[f]Monte Carlo exact probability.

[g]Significant difference between both groups.

[h]No significant difference between both groups.

Social and Economic Characteristics

The social and economic characteristics of the participants revealed notable differences between the online and offline groups. A significant difference was observed in working status, with a higher proportion of participants in the online group being employed (105/136, 77.2%) compared to the offline group (69/123, 56%), indicating a difference of 21.2% (*P*<.001). This association reflected a moderate effect size (Cramer V=0.224). Both groups predominantly resided in urban areas, with the online group showing 96.3% (131/136) compared to 91.1% (112/123) in the offline group, although this difference was not statistically significant. The proportion of participants who were breadwinners was also comparable between the groups, with 15.4% (21/136) in the online group and 22.8% (28/123) in the offline group.

Working conditions indicated that both groups had similar distributions, with the majority working full-time, and there were no significant differences in workplace (office-based vs remote) or working time (fixed vs rotational shifts). Participants were engaged in various sectors, including health, technology,

XSL•FO
RenderX

communication, governmental services, and entrepreneurship. Income levels also showed no significant differences, with 22.3% (23/103) of the online group earning more than 20,000 Egyptian pounds (US $419.03) compared with 25.8% (17/66) in the offline group (*P*=.25) (Table 2).

Table 2. Social and economic characteristics. A currency exchange rate of EGP £1=US $0.02078 is applicable.

Variable	Online group (n=136), n (%)	Offline group (n=123), n (%)	Chi-square (*df*)	Effect size, Cramer V	*P* value
Residence			3.1 (1)	0.109	.08
Urban	131 (96.3)	112 (91.1)			
Rural	5 (3.7)	11 (8.9)			
Breadwinner			2.3 (1)	0.093	.13
Yes	21 (15.4)	28 (22.8)			
No	115 (84.6)	95 (77.2)			
Receiving pocket money			2.1 (1)	0.090	.15
Yes	52 (38.2)	58 (47.2)			
No	84 (61.8)	65 (52.8)			
Working status			13.1 (1)	0.224	<.001
Yes[a]	105 (77.2)	69 (56.1)			
No[a]	31 (22.8)	54 (43.9)			
Working condition (n=174)			0.039 (1)	0.015	.84
Part-time	26 (24.8)	18 (26.1)			
Full-time	79 (75.2)	51 (73.9)			
Working place (n=174)			1.8 (2)	0.102	.41
Remote or home	27 (25.7)	17 (24.6)			
Office based	45 (42.9)	36 (52.2)			
Both	33 (31.4)	16 (23.2)			
Working time (n=172)			0.019 (1)	0.011	.89
Fixed shifts (9 AM to 5 PM)	60 (57.1)	39 (58.2)			
Rotational shifts	45 (42.9)	28 (41.8)			
Working sector (n=169)			6.9 (5)	0.202	.23
Health	18 (18.0)	14 (20.3)			
Technology	29 (29.0)	15 (21.7)			
Communication	7 (7.0)	10 (14.5)			
Governmental or public services	8 (8.0)	3 (4.3)			
Entrepreneurship	7 (7.0)	10 (14.5)			
Other	31 (31.0)	17 (24.6)			
Income in Egyptian pounds (n=169)			5.4 (4)	0.179	.25
Less than 5000	17 (16.5)	14 (21.2)			
5000-10,000	34 (33.0)	21 (31.8)			
10,000-15,000	20 (19.4)	5 (7.6)			
15,000-20,000	9 (8.7)	9 (13.6)			
More than 20,000	23 (22.3)	17 (25.8)			

[a]Significant difference between both groups.

XSL•FO
RenderX

Internet Access

Regarding internet access, the online group demonstrated a higher connectivity rate, with 97.8% (133/136) having internet access compared to 91.9% (113/123) in the offline group (*P*=.03). Additionally, the online group reported better internet usability, with a mean score of 4.57 (0.64) compared to 4.14 (1.25) in the offline group, indicating a significant difference (*P*=.001) (Table 3).

Table 3. Internet access and history of psychiatric visit.

Variable	Online group (n=136)	Offline group (n=123)	Test statistic	Effect size, Cramer V	*P* value
Having internet access, n (%)			4.8 (1)[a]	0.135	.03
Yes[b]	133 (97.8)	113 (91.9)			
No[b]	3 (2.2)	10 (8.1)			
Internet usability, mean (SD)	4.57 (0.641)	4.14 (1.25)	3.515 (257)[c]	0.295	.001
History of psychiatric visit, n (%)			35.8 (1)[a]	0.372	<.001
Yes[b]	49 (36.0)	90 (73.2)			
No[b]	87 (64.0)	33 (26.8)			
Online through video, n (%)			100.7 (1)[a]	0.623	<.001
Yes[b]	107 (78.7)	20 (16.3)			
No[b]	29 (21.3)	103 (83.7)			

[a]Chi-square test (degrees of freedom).

[b]Significant difference between both groups.

[c]2-tailed *t* test (degrees of freedom).

History of Psychiatric Visit

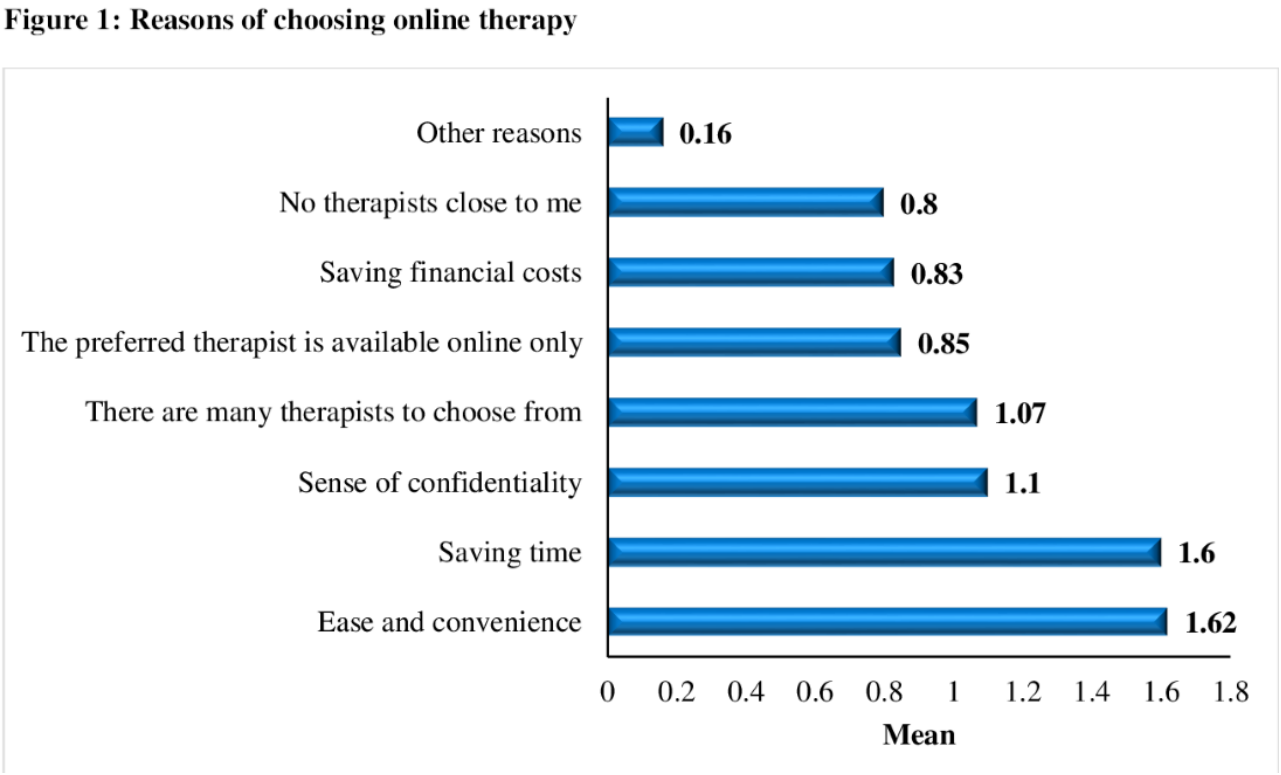
This study examined the history of psychiatric visits for each group. The offline group had a 37.2% higher rate of clinic visits, with almost three-quarters reporting such visits (90/123, 73%) compared to only 36% (49/136) in the online group (*P*<.001). Similarly, the online group demonstrated a significantly higher use of online video consultations, with 78.7% (107/136) receiving this service compared to only 16.3% (20/123) in the offline group (*P*<.001). This finding was associated with a large effect size (Cramer V=0.623) (Table 3).

Reasons for Choosing Online Therapy

Figure S1 in Multimedia Appendix 1 shows the primary reasons participants chose online therapy, with the top 4 factors being ease and convenience (mean 1.62, SD 0.55), saving time (mean 1.6, SD 0.58), sense of confidentiality (mean 1.1, SD 0.82), and availability of many therapists (mean 1.07, SD 0.76; Figure 1).

XSL•FO
RenderX

Figure 1. Reasons of choosing online therapy.



Big Five Personality Trait

With regard to personality traits, the only statistically significant difference between the groups was found in conscientiousness, with the offline group scoring higher than the online group (mean 13.44, SD 3.47 vs mean 12.15, SD 3.47; *P*=.003), reflecting a moderate effect size (Cohen *d*=0.37). No significant differences were observed for the other personality traits, including extraversion, neuroticism, agreeableness, and openness, which showed similar levels between the 2 groups. Among the personality traits, agreeableness had the highest mean in both groups. In the online group, openness was the second highest, whereas conscientiousness was the second highest in the offline group (Table 4).

Table 4. Big Five personality traits and the Self-Stigma of Seeking Help (SSOSH).

Variable	Online group (n=136), mean (SD)	Offline group (n=123), mean (SD)	2-tailed *t* test (*df*)	Effect size, Cohen *d*	*P* value
Extraversion	11.37 (2.79)	11.98 (3.27)	–1.635 (257)	0.201	.10
Neuroticism	12.57 (3.04)	12.74 (3.16)	–0.431 (257)	0.055	.67
Agreeableness	15.18 (3.51)	15.45 (3.12)	–0.635 (257)	0.081	.53
Openness	12.84 (3.71)	13.04 (3.68)	–0.440 (257)	0.054	.66
Conscientiousness	12.15 (3.47)	13.44 (3.5)	–2.982 (257)	0.37	.003
SSOSH	19.96 (5.68)	22.96 (6.48)	–3.973 (257)	0.492	<.001

SSOSH

The offline group showed a higher level of SSOSH, with a mean score of 22.96 (6.48) compared to the online group (mean 19.96, SD 5.68), reflecting a significant difference (2-tailed t_{257}=–3.973; *P*<.001) and a moderate-to-large effect size (Cohen *d*=0.492) (Table 4). We examined the correlation between personality traits and SSOSH in both the online and offline groups. Several significant correlations were identified in the online group. Extraversion showed a weak negative correlation with SSOSH (*r*=–0.199; *P*=.02), while openness exhibited a stronger negative correlation (*r*=–0.287; *P*=.001). Additionally, conscientiousness was also negatively correlated with SSOSH (*r*=-0.192; *P*=.03) (Table 4).

In the offline group, openness was the only trait that significantly correlated with SSOSH (*r*=–0.226; *P*=.12). These findings suggest that openness is consistently associated with lower self-stigma across both groups, whereas extraversion and conscientiousness appear to play more prominent roles in the online group (Table 5).

XSL•FO
RenderX

Table 5. Correlation between the Big Five traits and the Self-Stigma of Seeking Help (SSOSH) scale among online and offline groups (Pearson r and *P* value).

Group	Extraversion	Neuroticism	Agreeableness	Openness	Conscientiousness
Online (n=136)					
r	–0.199	0.149	–0.143	–0.287	–0.192
P value	.02	.08	.10	.001	.03
Offline (n=123)					
r	–0.083	0.140	–0.052	–0.226	–0.113
P value	.36	.12	.57	.01	.21

To account for the increased risk of type I error due to multiple comparisons, the Benjamini–Hochberg procedure was applied to control the false discovery rate at 0.05 across bivariate tests. Seven variables remained statistically significant after correction: SSOSH scores, history of psychiatric visits (both clinic-based and online), working status, ease of internet access, current education level, and conscientiousness. Other variables, including highest level of education and internet access, did not remain significant after correction (Table S1 in Multimedia Appendix 1).

Logistic Regression Analysis of Factors Affecting Choosing Online Sessions

Logistic regression analysis revealed several significant factors associated with choosing online mental health services. Multicollinearity among predictors was assessed using variance inflation factors; all values were below 1.52, indicating an acceptable level of collinearity (Table S2 in Multimedia Appendix 1). The model demonstrated a good fit, as indicated by a nonsignificant Hosmer–Lemeshow test (χ^2_8=3.9; *P*=.86), and explained a high proportion of variance in the outcome (Nagelkerke R^2=0.662). After adjustment for confounders, a history of online psychiatric consultation was the strongest factor (OR 28.90, 95% CI 11.739-71.165; *P*<.001), followed by a history of in-person psychiatric consultation, which was negatively associated with choosing online services (OR 0.11, 95% CI 0.048-0.269; *P*<.001). Being employed (OR 5.01, 95% CI 0.327-2.739; *P*=.002) and higher internet usability (OR 1.69, 95% 0.445-56.969; *P*=.03) were also significant factors. Among the personality traits, higher agreeableness (OR 1.16, 95% CI 1.001-1.351; *P*=.05) was positively associated with choosing online therapy, while higher openness (OR 0.85, 95% CI 0.748-0.975; *P*=.02) and conscientiousness (OR 0.86, 95% CI 0.758-0.971; *P*=.02) were negatively associated. Other factors such as education level, internet access, and SSOSH scores were not significant (Table 6). A marginal OR plot was constructed to show the adjusted effects of the statistically significant factors identified in the logistic regression model (Figure S2 in Multimedia Appendix 1).

XSL•FO
RenderX

Table 6. Logistic regression analysis of factors affecting choice of online sessions (overall percentage=85.7%; $\chi^2$13=177.5 and *P*<.001; Hosmer–Lemeshow test $\chi^2$8=3.9 and *P*=.86; Nagelkerke R^2=66.2%).

Independent variable	Unstandardized β coefficient	Unstandardized SE	OR[a]	(95% CI)	*P* value
Constant	–0.231	2.165	0.794		.92
Education level[b]					>.99
Illiterate	–17.562	14784.757	0.000	0.000	>.99
Student	–0.056	0.543	0.946	0.327-2.739	.92
Working status[c]	1.611	0.527	5.007	1.781-14.080	.002
Having internet access[c]	1.617	1.238	5.037	0.445-56.969	.19
Internet usability[c]	0.523	0.233	1.688	1.069-2.664	.03
History of psychiatric consultation (clinic)[c]	–2.178	0.442	0.113	0.048-0.269	<.001
History of psychiatric consultation (online)[c]	3.364	0.460	28.904	11.739-71.165	<.001
Extraversion	–0.091	0.070	0.913	0.796-1.048	.20
Neuroticism	–0.076	0.070	0.927	0.808-1.063	.28
Agreeableness	0.151	0.076	1.163	1.001-1.351	.048
Openness	–0.158	0.068	0.854	0.748-0.975	.02
Conscientiousness	–0.153	0.063	0.858	0.758-0.971	.02
SSOSH[d]	–0.056	0.034	0.946	0.885-1.011	.10

[a]OR: odds ratio.

[b]Reference: graduate.

[c]Reference: no.

[d]SSOSH: Self-Stigma of Seeking Help.

Discussion

Overview

The internet has become increasingly important as primary channel for delivering remote therapy. Over the past 20 years, there has been significant growth in online platforms supporting individuals with various mental health conditions [30,31]. However, research has primarily focused on comparing the efficacy of telepsychiatry with offline methods [32,33], with less emphasis on the individual factors that influence the choice of online options. The personal and psychological factors associated with choosing online therapy over offline therapy remain insufficiently explored.

This study aims to provide preliminary data on the differences between individuals who prefer online versus offline mental health services, with a focus on the relationships between sociodemographic characteristics, previous mental health experiences, personality traits, self-stigma, and the preferred modality of help-seeking. Multivariable logistic regression analysis identified a history of online psychiatric consultation as the strongest factor associated with online preference, followed by employment, ease of internet use, and agreeableness. Conversely, previous offline psychiatric visits, along with higher levels of openness and conscientiousness, were significantly associated with a preference for offline care.

We found that the online group was more likely to be graduates, employed, hold a university degree, and have greater internet access and usability. This suggests that telepsychiatry may be more widely adopted when internet access is better and education levels are higher. A national survey in Egypt found that individuals with higher education and current employment status demonstrated greater awareness of and a more positive attitude toward telemedicine [34], aligning with broader research that links telemedicine use to higher educational levels, better internet access, and greater digital literacy [35,36]. Another study on Egyptian students reported that unfamiliarity with electronic mental health is among the most commonly perceived barriers [37]. Additionally, graduates and employed individuals often face time constraints and must balance multiple responsibilities, making the flexible nature of online therapy an appealing option.

This study revealed no significant difference in income levels between the online and offline groups. Globally, individuals with lower income levels face greater challenges in accessing mental health services [38]. However, participants in both groups in this study had already accessed private mental health care, suggesting that they likely have the financial resources to afford such services, which may minimize the impact of income disparities between the groups. In addition, several other factors influence the decision to seek mental health care in the Arab world, including mental health stigma, negative perceptions and attitudes toward mental illness, low mental health literacy,

XSL•FO
RenderX

family dynamics, structural barriers related to the health care system, and service accessibility [39,40]. Such factors have also been identified in studies conducted in Egypt, where help-seeking behaviors are often affected by strong social stigma and reliance on informal support [41,42].

We found that more than 90% of both the offline and online groups reside in urban areas, indicating that rural residents are less likely to seek mental health help, whether online or offline. Egyptians living in rural areas tend to have lower levels of mental health literacy, higher levels of stigma, and limited internet access, all of which may contribute to reduced help-seeking rates compared with urban residents [43-45]. Additionally, there is an unequal distribution of mental health services across Egypt, where the majority of psychiatric facilities and specialists are concentrated in urban centers. As a result, individuals living in rural areas face limited access to mental health care, both online and offline [46]. This pattern is commonly seen in other countries as well; rural populations often face logistical challenges, including a shortage of mental health professionals, significant distances to health care facilities, and lower levels of mental health literacy, which hinder their ability to recognize mental health issues and the importance of seeking help [47-49].

This study found that the offline group had a statistically significant association with previous clinic visits, whereas the online group had a similar association with use of video-based online therapy. Notably, this was the strongest factor associated with choosing online therapy in the logistic regression, with OR 28.90 (95% CI 11.739-71.165). This pattern may be explained by the mere exposure effect, whereby individuals develop a preference or bias toward familiar options due to increased comfort, perceived safety, and reduced cognitive effort [50]. Previous literature suggests that when individuals have experience using online services, they are more likely to prefer these platforms when seeking help for mental health problems [10,51]. However, choosing face-to-face counseling over online therapy may be attributed to unfamiliarity with online counseling and concerns about privacy and security [52,53]. Additionally, the association may reflect the persistence of the initial factors that led individuals to choose online therapy, such as the need for privacy, scheduling flexibility, time efficiency, or work-related constraints.

The strongest reasons for choosing online therapy over offline were ease and convenience, time saving, and a greater sense of confidentiality. This trend is common, as online therapy is perceived as beneficial due to its flexibility, allowing clients to schedule sessions without needing to travel, saving both time and money [54,55]. This convenience is especially helpful for individuals balancing work, family, or other commitments [56]. Additionally, confidentiality is a primary motive for choosing online therapy, as it offers greater privacy, especially for individuals dealing with stigmatized issues [57].

Agreeableness was the most prevalent trait in both groups in this study. Agreeable individuals are more likely to have a positive attitude toward seeking professional help and recognize the need for information and treatment [55,56], with a lower likelihood of holding stigmatizing beliefs about mental illness [57]. In many Arab countries, patients tend to favor a paternalistic model of the doctor-patient relationship, where the physician is expected to take the lead in decision-making, with little involvement from the patient, a tendency that aligns with a high level of agreeableness [58,59]. Moreover, high agreeableness has been commonly observed in studies conducted on Arab populations, including Egyptians, reflecting cultural values of collectivism and social conformity [58-60].

The logistic regression analysis in this study identified associations between specific personality traits and the preference for telepsychiatry. Higher levels of agreeableness showed a marginal association with choosing online therapy, while higher levels of openness and conscientiousness were negatively associated. Previous studies have shown that individuals high in agreeableness may demonstrate greater acceptance of online therapy and a higher capacity to adapt to new situations that require online therapy [61,62]. Moreover, their heightened sensitivity to social evaluation from others may make them prefer the privacy offered by telepsychiatry, as accessing care from home helps minimize exposure to potential stigma associated with attending in-person clinics [63].

Unexpectedly, higher openness was negatively associated with the preference for telepsychiatry. While openness is generally linked to creativity and a willingness to explore new experiences [64,65], individuals high in this trait may favor offline therapy, potentially due to its greater opportunity for in-depth emotional connection and self-disclosure [66]. This aligns with clinical observations from Shezlong, where many clients tend to keep their cameras off during initial video sessions—an act that may reflect lower levels of openness.

Conscientiousness was statistically higher in the offline group and negatively associated with telepsychiatry. This trait is associated to a strong sense of responsibility, self-discipline, and a goal-oriented mindset [67]. Seeking help in the clinic may require higher levels of commitment, punctuality, and adherence to structured appointments compared with the flexibility of online therapy. Currently, the literature lacks direct studies on personality traits and telepsychiatry use. Most research focuses on broader help-seeking behaviors or general telehealth adoption [68,69].

Self-stigma was higher among the offline group. Previous research has demonstrated that higher levels of self-stigma are associated with a reduced likelihood of seeking offline psychological help [70,71]. A systematic review conducted in the Middle East found that both public and self-stigma serve as major barriers to accessing mental health care, with many individuals turning to family members or traditional healers before seeking professional help [67]. Addressing these cultural nuances through targeted awareness campaigns and community-based initiatives may help reduce stigma and promote help-seeking behavior across age groups [72,73].

Although self-stigma was lower among the online group, it was not independently associated with the choice of online help when controlling confounding variables in the logistic regression analysis. The relation between self-stigma and help seeking has not been sufficiently investigated in the context of online help-seeking or telepsychiatry. A previous study from China

XSL•FO
RenderX

reported a negative association between self-stigma and the willingness to seek online help, though it did not include a comparison group of offline help-seekers [70]. Another Swedish study found that self-stigma significantly predicted the intention to seek online help among college students and primary care patients [39]. The inconsistency in previous literature suggests that the dynamics of self-stigma in digital contexts may differ from traditional settings and warrant further investigation.

The absence of a significant association between self-stigma and choosing online therapy in the logistic regression in this study might be attributed to several factors. First, other variables such as perceived effectiveness of online therapy, accessibility, and technological proficiency may play a more substantial role in the decision-making process. Second, the relationship between self-stigma and help-seeking behavior is complex and may be influenced by additional factors, such as cultural norms and individual differences [74,75].

The study showed that openness was the only trait negatively associated with self-stigma in both the online and offline groups. Self-stigma occurs when individuals internalize prejudiced beliefs prevalent in society, leading to feelings of shame and reduced help-seeking behavior [71]. Research demonstrates that openness to experience is associated with cognitive flexibility and a lower tendency to accept rigid stereotypes or prejudices, which may in turn reduce the internalization of stigmatizing beliefs about mental health and lead to lower levels of self-stigma [76,77].

Additionally, individuals with high openness are more willing to disclose psychological distress, as they tend to view self-exploration and emotional vulnerability as opportunities for personal growth rather than sources of self-stigma [78]. In Egyptian and broader Arab communities, individuals often express psychological distress through physical symptoms rather than verbalizing emotional difficulties, in an effort to avoid the perceived weakness and social stigma associated with mental illness and help seeking, which is often viewed as socially unacceptable. Mental health issues are frequently associated with shame, leading families to conceal a relative's psychiatric diagnosis for fear of damaging their social standing or marriage prospects. This stigma is further reinforced by cultural norms that place a high value on endurance and emotional restraint, discouraging open discussion of mental health struggles [40,79].

There are inconsistent results regarding openness and self-stigma, with previous research showing no direct association between them [74,75]. The complex relationship between self-stigma of seeking help and personality traits is rarely studied in the context of telepsychiatry. Therefore, further investigation is needed to understand how factors associated with self-stigma operate across online and offline settings. Understanding these factors associated with self-stigma is essential due to the high levels of mental health stigma, which pose a significant barrier to help seeking, especially in Egypt and the Arab world [80].

Implications and Future Research

This study has significant implications for clinical practice and underscores the need for future research. Understanding the personal and psychological factors that influence the preference for online therapy can help mental health professionals tailor their services to better meet the needs of patients. The significant differences in self-stigma and personality traits between online and offline help seekers highlight the importance of exploring these factors further, particularly as online therapy emerges as an increasingly vital modality for mental health support, likely to expand in popularity and accessibility. Consequently, future research should focus on replicating these findings among diverse populations and with larger sample sizes to enhance their generalizability. Importantly, this study opens the door for investigating differences between actual offline and online help seekers, necessitating further studies to address the existing knowledge gap surrounding help-seeking behaviors. By expanding this research, we can gain deeper insights into the complexities of mental health service use, ultimately contributing to a more comprehensive understanding of the motivations behind choosing different modalities for support.

Strengths and Limitations

This study aimed to compare actual online and offline mental health help seekers in terms of sociodemographic factors, personality traits, and self-stigma, rather than merely assessing preferences and intentions, by focusing on individuals who sought mental health care. The findings serve as preliminary data for further research and contribute to the representation of Egyptian and Arabic populations in the literature. Despite its strengths, the study has several limitations. The relatively small sample size and recruitment from one online platform and a few private clinics may restrict the generalizability of the findings to the broader population of mental health service users. Additionally, the use of convenience sampling may introduce selection bias, as participants who engage with the study may differ systematically from those who do not. Furthermore, conducting the study in only one country limits the ability to draw broader conclusions applicable to diverse cultural contexts, highlighting the need for future research to replicate these findings in different settings and populations.

Conclusions

In conclusion, this study provides key insights into the differences between individuals who prefer online versus offline mental health services. Logistic regression analysis identified several significant factors associated with online service preference. The strongest factor was a history of online psychiatric consultation, highlighting the role of previous exposure in shaping treatment choices. Employment status and ease of internet access were also positively associated with choosing online therapy. In contrast, a history of in-person consultations and higher levels of openness and conscientiousness were associated with a preference for offline services. Higher agreeableness showed a marginal association with choosing online therapy. These findings suggest that treatment modality preferences are influenced by a complex interplay of previous experiences, sociodemographic factors, and personality traits. Tailoring mental health interventions to these factors may enhance service accessibility and engagement. Future research should aim to validate these findings in more diverse populations and explore additional cultural and structural influences.

XSL•FO
RenderX

Acknowledgments

The authors would like to thank Asmaa Hassan and her team for their dedicated efforts in collaborating with the researchers and her invaluable contributions to data collection. Her commitment and support were instrumental in the success of this study. All authors declare that they had insufficient funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee (APF) support for the publication of this article.

Data Availability

The datasets used and/or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary figures and tables.
[DOCX File , 984 KB - humanfactors_v12i1e69305_app1.docx]

References

1. Rochlen AB, Zack JS, Speyer C. Online therapy: review of relevant definitions, debates, and current empirical support. J Clin Psychol 2004;60(3):269-283. [doi: 10.1002/jclp.10263] [Medline: 14981791]
2. Malhotra S, Chakrabarti S, Shah R. Telepsychiatry: promise, potential, and challenges. Indian J Psychiatry 2013;55(1):3-11. [doi: 10.4103/0019-5545.105499] [Medline: 23441027]
3. Li H, Glecia A, Kent-Wilkinson A, Leidl D, Kleib M, Risling T. Transition of mental health service delivery to telepsychiatry in response to COVID-19: a literature review. Psychiatr Q 2022;93(1):181-197. [doi: 10.1007/s11126-021-09926-7] [Medline: 34101075]
4. Bashshur RL, Shannon GW, Bashshur N, Yellowlees PM. The empirical evidence for telemedicine interventions in mental disorders. Telemed J E Health 2016;22(2):87-113 [FREE Full text] [doi: 10.1089/tmj.2015.0206] [Medline: 26624248]
5. Batterham PJ, Calear AL. Preferences for internet-based mental health interventions in an adult online sample: findings from an online community survey. JMIR Ment Health 2017;4(2):e26. [doi: 10.2196/mental.7722] [Medline: 28666976]
6. Mseke E, Jessup B, Barnett T. A systematic review of the preferences of rural and remote youth for mental health service access: telehealth versus face-to-face consultation. Aust J Rural Health 2023;31(3):346-360. [doi: 10.1111/ajr.12961] [Medline: 36606417]
7. Rickwood DJ, Mazzer KR, Telford NR. Social influences on seeking help from mental health services, in-person and online, during adolescence and young adulthood. BMC Psychiatry 2015;15:40. [doi: 10.1186/s12888-015-0429-6] [Medline: 25886609]
8. Renn BN, Hoeft TJ, Lee HS, Bauer AM, Areán PA. Preference for in-person psychotherapy versus digital psychotherapy options for depression: survey of adults in the US. NPJ Digit Med 2019;2:6. [doi: 10.1038/s41746-019-0077-1] [Medline: 31304356]
9. Berle D, Starcevic V, Milicevic D, Hannan A, Dale E, Brakoulias V, et al. Do patients prefer face-to-face or internet-based therapy? Psychother Psychosom 2015;84(1):61-62. [doi: 10.1159/000367944] [Medline: 25547039]
10. March S, Day J, Ritchie G, Rowe A, Gough J, Hall T, et al. Attitudes toward e-mental health services in a community sample of adults: online survey. J Med Internet Res 2018;20(2):e59. [doi: 10.2196/jmir.9109] [Medline: 29459357]
11. Wallin E, Maathz P, Parling T, Hursti T. Self-stigma and the intention to seek psychological help online compared to face-to-face. J Clin Psychol 2018;74(7):1207-1218. [doi: 10.1002/jclp.22583] [Medline: 29315545]
12. Levin ME, Krafft J, Levin C. Does self-help increase rates of help seeking for student mental health problems by minimizing stigma as a barrier? J Am Coll Health 2018;66(4):302-309. [doi: 10.1080/07448481.2018.1440580] [Medline: 29447600]
13. Kauer SD, Mangan C, Sanci L. Do online mental health services improve help-seeking for young people? A systematic review. J Med Internet Res 2014;16(3):e66. [doi: 10.2196/jmir.3103] [Medline: 24594922]
14. Hilty DM, Ferrer DC, Parish MB, Johnston B, Callahan EJ, Yellowlees PM. The effectiveness of telemental health: a 2013 review. Telemed J E Health 2013;19(6):444-454. [doi: 10.1089/tmj.2013.0075] [Medline: 23697504]
15. Ye J, Shim R, Lukaszewski T, Yun K, Kim SH, Ruth G. Telepsychiatry services for Korean immigrants. Telemed J E Health 2012;18(10):797-802 [FREE Full text] [doi: 10.1089/tmj.2012.0041] [Medline: 23234426]
16. Clough BA, Zarean M, Ruane I, Mateo NJ, Aliyeva TA, Casey LM. Going global: do consumer preferences, attitudes, and barriers to using e-mental health services differ across countries? J Ment Health 2019;28(1):17-25. [doi: 10.1080/09638237.2017.1370639] [Medline: 28857650]

XSL•FO
RenderX

17. Austin D, Bhola P, Tebble C, Shandley K. Preferences for online mental health services among Australian and Indian samples: a cross-cultural comparison. Psychol Stud 2018;63(4):376-383. [doi: 10.1007/s12646-018-0453-y]
18. Feng H, Kurata K, Cao J, Itsuki K, Niwa M, Aoyama A, et al. Telemedicine research trends in 2001-2022 and research cooperation between China and other countries before and after the COVID-19 pandemic: bibliometric analysis. Interact J Med Res 2024;13:e40801. [doi: 10.2196/40801] [Medline: 39213010]
19. Murshidi R, Hammouri M, Taha H, Kitaneh R, Alshneikat M, Al-Qawasmeh A, et al. Knowledge, attitudes, and perceptions of Jordanians toward adopting and using telemedicine: national cross-sectional study. JMIR Hum Factors 2022;9(4):e41499. [doi: 10.2196/41499] [Medline: 36257609]
20. Sayed M, Gafary ME, Mohsen N, Khalil S, Farid H, Sabry N, et al. Assessment of telepsychiatry services provided by Okasha institute of psychiatry during COVID-19 pandemic. Middle East Curr Psychiatry 2023;30(1):1-8. [doi: 10.1186/s43045-023-00356-8]
21. Mamdouh M, Tai AMY, Westenberg JN, Shams F, Jang K, Badawy A, et al. Egyptian students open to digital mental health care: cross-sectional survey. JMIR Form Res 2022;6(3):e31727. [doi: 10.2196/31727] [Medline: 35311692]
22. Waqas A, Mehmood S, Jawwad A, Pittam B, Kundu S, Correia J, et al. Telemedicine in Arab countries: innovation, research trends, and way forward. Front Digit Health 2020;2:610837. [doi: 10.3389/fdgth.2020.610837] [Medline: 34713071]
23. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004;6(3):e34. [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
24. Klein B, Cook S. Preferences for e-mental health services amongst an online Australian sample? E-J Appl Psychol 2010;6(1):27-38. [doi: 10.7790/ejap.v6i1.184]
25. Rashad A, Sharaf M. Who benefits from public healthcare subsidies in Egypt? Social Sciences 2015;4(4):1162-1176. [doi: 10.3390/socsci4041162]
26. Abdel-Khalek AM. Psychometric Properties of the Arabic Big Five Personality Inventory (ABFPI). Psychol Behav Sci Int J 2018;11(4). [doi: 10.19080/PBSIJ.2019.11.555820]
27. Hussein MH, Abdel-Khalek AM. Developing a revised version of the Arabic Big Five Personality Inventory using Item Response Theory. Mankind Q 2021;62(2):327-343. [doi: 10.46469/mq.2021.62.2.7]
28. Vogel DL, Wade NG, Haake S. Measuring the self-stigma associated with seeking psychological help. J Couns Psychol 2006;53(3):325-337. [doi: 10.1037/0022-0167.53.3.325]
29. Self-Stigma Research Collaborative.: Iowa State University; 2023. URL: https://selfstigma.psych.iastate.edu/ssosh-scale/ [accessed 2023-09-24]
30. Hoermann S, McCabe KL, Milne DN, Calvo RA. Application of synchronous text-based dialogue systems in mental health interventions: systematic review. J Med Internet Res 2017;19(8):e267 [FREE Full text] [doi: 10.2196/jmir.7023] [Medline: 28784594]
31. Lipson SK, Lattie EG, Eisenberg D. Increased rates of mental health service utilization by U.S. college students: 10-year population-level trends (2007-2017). Psychiatr Serv 2019;70(1):60-63 [FREE Full text] [doi: 10.1176/appi.ps.201800332] [Medline: 30394183]
32. Waldmann T, Staiger T, Oexle N, Rüsch N. Mental health literacy and help-seeking among unemployed people with mental health problems. J Ment Health 2020;29(3):270-276. [doi: 10.1080/09638237.2019.1581342] [Medline: 30862221]
33. Ye X, Bapuji SB, Winters SE, Struthers A, Raynard M, Metge C, et al. Effectiveness of internet-based interventions for children, youth, and young adults with anxiety and/or depression: a systematic review and meta-analysis. BMC Health Serv Res 2014;14:313. [doi: 10.1186/1472-6963-14-313] [Medline: 25037951]
34. Alboraie M, Allam MA, Youssef N, Abdalgaber M, El-Raey F, Abdeen N, et al. Knowledge, applicability, and barriers of telemedicine in Egypt: a national survey. Int J Telemed Appl 2021;2021:5565652 [FREE Full text] [doi: 10.1155/2021/5565652] [Medline: 34211550]
35. Adams RB, Nelson VR, Holtz BE. Barriers for telemedicine use among nonusers at the beginning of the pandemic. Telemed Rep 2021;2(1):211-216 [FREE Full text] [doi: 10.1089/tmr.2021.0022] [Medline: 35720746]
36. Hung M, Ocampo M, Raymond B, Mohajeri A, Lipsky MS. Telemedicine among adults living in America during the COVID-19 pandemic. Int J Environ Res Public Health 2023;20(9):5680. [doi: 10.3390/ijerph20095680] [Medline: 37174198]
37. Mamdouh M, Tai AMY, Westenberg JN, Shams F, Jang K, Badawy A, et al. Egyptian students open to digital mental health care: cross-sectional survey. JMIR Form Res 2022;6(3):e31727. [doi: 10.2196/31727] [Medline: 35311692]
38. Henking C, Reeves A, Chrisinger B. Global inequalities in mental health problems: understanding the predictors of lifetime prevalence, treatment utilisation and perceived helpfulness across 111 countries. Prev Med 2023;177:107769. [doi: 10.1016/j.ypmed.2023.107769] [Medline: 37952711]
39. Khatib HE, Alyafei A, Shaikh M. Understanding experiences of mental health help-seeking in Arab populations around the world: a systematic review and narrative synthesis. BMC Psychiatry 2023;23(1):324 [FREE Full text] [doi: 10.1186/s12888-023-04827-4] [Medline: 37161342]
40. Okasha T, Shaker NM, El-Gabry DA. Mental health services in Egypt, the middle east, and North Africa. Int Rev Psychiatry 2025;37(3-4):306-314. [doi: 10.1080/09540261.2024.2400143] [Medline: 40708426]

XSL•FO
RenderX

41. Zalat MM, Mortada EM, El Seifi OS. Stigma and attitude of mental health help-seeking among a sample of working versus non-working Egyptian women. Community Ment Health J 2019;55(3):519-526. [doi: 10.1007/s10597-018-0298-9] [Medline: 29931402]
42. Baklola M, Terra M, Elzayat MA, Abdelhady D, El-Gilany A, Collaborators ATO. Pattern, barriers, and predictors of mental health care utilization among Egyptian undergraduates: a cross-sectional multi-centre study. BMC Psychiatry 2023;23(1):139 [FREE Full text] [doi: 10.1186/s12888-023-04624-z] [Medline: 36879216]
43. Baklola M, Terra M, Taha A, Elnemr M, Yaseen M, Maher A, et al. Mental health literacy and help-seeking behaviour among Egyptian undergraduates: a cross-sectional national study. BMC Psychiatry 2024;24(1):202. [doi: 10.1186/s12888-024-05620-7] [Medline: 38475754]
44. Sayed TA, Ali MM, Hadad S. Risk factors and impact of stigma on psychiatric patients in Sohag. Egypt J Neurol Psychiatry Neurosurg 2021;57(1):148. [doi: 10.1186/s41983-021-00403-3]
45. Mohamed SAE MF. Analyzing the digital divide in Egypt. Arab J Adm 2025;45(1):113-134. [doi: 10.21608/AJA.2022.112214.1190]
46. Okasha TA, Shaker NM, Elgabry DMA. Community mental health services in Egypt. Consort Psychiatr 2022;3(2):123-128. [doi: 10.17816/CP165] [Medline: 39045112]
47. Gupta R. Rural–urban divide in mental health care in India: bridging the gaps. Indian J Soc Psychiatry 2024;40(1):7-10. [doi: 10.4103/ijsp.ijsp_41_24]
48. Barry R, Rehm J, de Oliveira C, Gozdyra P, Chen S, Kurdyak P. Help-seeking behavior among adults who attempted or died by suicide in Ontario, Canada. Suicide Life Threat Behav 2023;53(1):54-63. [doi: 10.1111/sltb.12921] [Medline: 36098239]
49. Almalki Z, Imam M, Abou Chahin N, ALSammak N, Entabli S, Alhammad S, et al. Access and disparities in the use of telemedicine among patients with chronic conditions in Saudi Arabia: a cross-sectional study. J Multidiscip Healthc 2023;16:3789-3798 [FREE Full text] [doi: 10.2147/JMDH.S433653] [Medline: 38076594]
50. Liao H, Yeh S, Shimojo S. Novelty vs. familiarity principles in preference decisions: task-context of past experience matters. Front Psychol 2011;2:43 [FREE Full text] [doi: 10.3389/fpsyg.2011.00043] [Medline: 21713246]
51. Haun MW, Oeljeklaus L, Hoffmann M, Tönnies J, Wensing M, Szecsenyi J, et al. Primary care patients' experiences of video consultations for depression and anxiety: a qualitative interview study embedded in a randomized feasibility trial. BMC Health Serv Res 2023;23(1):9 [FREE Full text] [doi: 10.1186/s12913-022-09012-z] [Medline: 36600264]
52. Rochlen AB, Beretvas SN, Zack JS. The online and face-to-face counseling attitudes scales: a validation study. Meas Eval Couns Dev 2004;37(2):95-111. [doi: 10.1080/07481756.2004.11909753]
53. Bathje GJ, Kim E, Rau E, Bassiouny MA, Kim T. Attitudes toward face-to-face and online counseling: roles of self-concealment, openness to experience, loss of face, stigma, and disclosure expectations among Korean college students. Int J Adv Couns 2014;36(4):408-422. [doi: 10.1007/s10447-014-9215-2]
54. Lippke S, Gao L, Keller FM, Becker P, Dahmen A. Adherence with online therapy vs face-to-face therapy and with online therapy vs care as usual: secondary analysis of two randomized controlled trials. J Med Internet Res 2021;23(11):e31274. [doi: 10.2196/31274] [Medline: 34730541]
55. Tajan N, Devès M, Potier R. Tele-psychotherapy during the COVID-19 pandemic: a mini-review. Front Psychiatry 2023;14:1060961 [FREE Full text] [doi: 10.3389/fpsyt.2023.1060961] [Medline: 37476543]
56. Pullyblank K, Scribani M, Krupa N, Chapman A, Kern M, Brunner W. Exploring multiple dimensions of access to and preferences for telehealth use. Telemed Rep 2023;4(1):348-358 [FREE Full text] [doi: 10.1089/tmr.2023.0049] [Medline: 38098780]
57. Wallin E, Maathz P, Parling T, Hursti T. Self-stigma and the intention to seek psychological help online compared to face-to-face. J Clin Psychol 2018;74(7):1207-1218. [doi: 10.1002/jclp.22583] [Medline: 29315545]
58. Abdel-Khalek AM, Carson J, Patel A, Shahama A. The Big Five personality traits as predictors of life satisfaction in Egyptian college students. Nordic Psychology 2022;75(2):113-130. [doi: 10.1080/19012276.2022.2065341]
59. Mohamed Taha Eid S, Del Mundo E, Hashem El-Monshed A. Personality traits, perceived stress and perception of COVID-19 among Arab population: a cross-sectional survey. Int J Afr Nurs Sci 2022;17:100502. [doi: 10.1016/j.ijans.2022.100502] [Medline: 36311299]
60. Ali M, Jahan A, Abdelrahman R. Measuring personality in Libyan Arabs: validating the Big Five aspect scale with 10 factors domain. BMC Psychol 2024;12(1):748. [doi: 10.1186/s40359-024-02270-y] [Medline: 39696550]
61. Gulliver A, Calear AL, Sunderland M, Kay-Lambkin F, Farrer LM, Batterham PJ. Predictors of acceptability and engagement in a self-guided online program for depression and anxiety. Internet Interv 2021;25:100400 [FREE Full text] [doi: 10.1016/j.invent.2021.100400] [Medline: 34026569]
62. Cieślik B, Kuligowski T, Cacciante L, Kiper P. The impact of personality traits on patient satisfaction after telerehabilitation: a comparative study of remote and face-to-face musculoskeletal rehabilitation during COVID-19 lockdown. Int J Environ Res Public Health 2023;20(6):2019. [doi: 10.3390/ijerph20065019] [Medline: 36981927]
63. Stoll J, Müller JA, Trachsel M. Ethical issues in online psychotherapy: a narrative review. Front Psychiatry 2020;10:993 [FREE Full text] [doi: 10.3389/fpsyt.2019.00993] [Medline: 32116819]

XSL•FO
RenderX

64. Gocłowska MA, Ritter SM, Elliot AJ, Baas M. Novelty seeking is linked to openness and extraversion, and can lead to greater creative performance. J Pers 2019;87(2):252-266. [doi: 10.1111/jopy.12387] [Medline: 29604214]
65. Vartanian O, Wertz CJ, Flores RA, Beatty EL, Smith I, Blackler K, et al. Structural correlates of openness and intellect: implications for the contribution of personality to creativity. Hum Brain Mapp 2018;39(7):2987-2996. [doi: 10.1002/hbm.24054] [Medline: 29656437]
66. Shahaf-Oren B, Madan I, Henderson C. "A lot of medical students, their biggest fear is failing at being seen to be a functional human": disclosure and help-seeking decisions by medical students with health problems. BMC Med Educ 2021;21(1):599 [FREE Full text] [doi: 10.1186/s12909-021-03032-9] [Medline: 34865636]
67. Elshamy F, Hamadeh A, Billings J, Alyafei A. Mental illness and help-seeking behaviours among middle eastern cultures: a systematic review and meta-synthesis of qualitative data. PLoS One 2023;18(10):e0293525 [FREE Full text] [doi: 10.1371/journal.pone.0293525] [Medline: 37883515]
68. Kauer SD, Mangan C, Sanci L. Do online mental health services improve help-seeking for young people? A systematic review. J Med Internet Res 2014;16(3):e66. [doi: 10.2196/jmir.3103] [Medline: 24594922]
69. Aguirre Velasco A, Cruz ISS, Billings J, Jimenez M, Rowe S. What are the barriers, facilitators and interventions targeting help-seeking behaviours for common mental health problems in adolescents? A systematic review. BMC Psychiatry 2020;20(1):293. [doi: 10.1186/s12888-020-02659-0] [Medline: 32527236]
70. Chen X, Du A, Qi R. Factors affecting willingness to receive online counseling: the mediating role of ethical concerns. Int J Environ Res Public Health 2022;19(24):16462 [FREE Full text] [doi: 10.3390/ijerph192416462] [Medline: 36554344]
71. Vogel DL, Bitman RL, Hammer JH, Wade NG. Is stigma internalized? The longitudinal impact of public stigma on self-stigma. J Couns Psychol 2013;60(2):311-316. [doi: 10.1037/a0031889] [Medline: 23421775]
72. Shahwan S, Lau JH, Goh CMJ, Ong WJ, Tan GTH, Kwok KW, et al. The potential impact of an anti-stigma intervention on mental health help-seeking attitudes among university students. BMC Psychiatry 2020;20(1):562 [FREE Full text] [doi: 10.1186/s12888-020-02960-y] [Medline: 33238951]
73. Abd El Salam AE, AbdAllah AM, El Maghawry HA. Effect of health education program on improving knowledge and atitude towards mental health stigma and professional help-seeking among adolescents. Middle East Curr Psychiatry 2023;30(1):1-9. [doi: 10.1186/s43045-023-00298-1]
74. Komiya N, Good GE, Sherrod NB. Emotional openness as a predictor of college students' attitudes toward seeking psychological help. J Couns Psychol 2000;47(1):138-143. [doi: 10.1037/0022-0167.47.1.138]
75. Ingram PB, Lichtenberg JW, Clarke E. Self-stigma, personality traits, and willingness to seek treatment in a community sample. Psychol Serv 2016;13(3):300-307. [doi: 10.1037/ser0000086] [Medline: 27253319]
76. Sibley CG, Duckitt J. Personality and prejudice: a meta-analysis and theoretical review. Pers Soc Psychol Rev 2008;12(3):248-279. [doi: 10.1177/1088868308319226] [Medline: 18641385]
77. Yuan Q, Seow E, Abdin E, Chua BY, Ong HL, Samari E, et al. Direct and moderating effects of personality on stigma towards mental illness. BMC Psychiatry 2018;18(1):358 [FREE Full text] [doi: 10.1186/s12888-018-1932-3] [Medline: 30400846]
78. Corrigan PW, Rao D. On the self-stigma of mental illness: stages, disclosure, and strategies for change. Can J Psychiatry 2012;57(8):464-469. [doi: 10.1177/070674371205700804] [Medline: 22854028]
79. Okasha T, Elkholy H, El-Ghamry R. Overview of the family structure in Egypt and its relation to psychiatry. Int Rev Psychiatry 2012;24(2):162-165. [doi: 10.3109/09540261.2012.658030] [Medline: 22515467]
80. Fekih-Romdhane F, Daher-Nashif S, Stambouli M, Alhuwailah A, Helmy M, Shuwiekh HAM, et al. Mental illness stigma as a moderator in the relationship between religiosity and help-seeking attitudes among Muslims from 16 Arab countries. BMC Public Health 2023;23(1):1671. [doi: 10.1186/s12889-023-16622-7] [Medline: 37649023]

Abbreviations

CHERRIES: Checklist for Reporting Results of Internet E-Surveys
OR: odds ratio
SSOSH: Self-Stigma of Seeking Help

Edited by A Kushniruk; submitted 26.11.24; peer-reviewed by A Alqasir, S Sheth - MD; comments to author 04.03.25; revised version received 16.05.25; accepted 12.08.25; published 07.11.25.

Please cite as:
Adwi M, Mahmoud B, Amer N, Alamrawy RG, Sadek I, Elsheikh M
Understanding the Differences Between Online and Offline Mental Health Help Seekers: Cross-Sectional Comparative Study
JMIR Hum Factors 2025;12:e69305
URL: https://humanfactors.jmir.org/2025/1/e69305
doi:10.2196/69305
PMID:

XSL•FO
RenderX

©Mohamed Adwi, Bahaa Mahmoud, Noha Amer, Roa Gamal Alamrawy, Ismail Sadek, Mohamed Elsheikh. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 07.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

XSL•FO
RenderX